WATER ENCYCLOPEDIA

DOMESTIC, MUNICIPAL, AND INDUSTRIAL WATER SUPPLY AND WASTE DISPOSAL

Jay Lehr, Ph.D.
Editor-in-Chief

Jack Keeley
Senior Editor

Janet Lehr
Associate Editor

Thomas B. Kingery III
Information Technology Director

The *Water Encyclopedia* is available online at
http://www.mrw.interscience.wiley.com/eow/

WILEY-
INTERSCIENCE

A John Wiley & Sons, Inc., Publication

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data is available.

Lehr, Jay
Water Encyclopedia: Domestic, Municipal, and Industrial Water Supply and Waste Disposal

ISBN 0-471-73687-2
ISBN 0-471-44164-3 (Set)

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

WATER ENCYCLOPEDIA

WATER QUALITY
AND RESOURCE DEVELOPMENT

Jay Lehr, Ph.D.
Editor-in-Chief

Jack Keeley
Senior Editor

Janet Lehr
Associate Editor

Thomas B. Kingery III
Information Technology Director

The *Water Encyclopedia* is available online at
http://www.mrw.interscience.wiley.com/eow/

WILEY-
INTERSCIENCE

A John Wiley & Sons, Inc., Publication

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data is available.

Lehr, Jay
Water Encyclopedia: Water Quality and Resource Development

ISBN 0-471-73686-4
ISBN 0-471-44164-3 (Set)

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

WATER ENCYCLOPEDIA

SURFACE AND AGRICULTURAL WATER

Jay Lehr, Ph.D.
Editor-in-Chief

Jack Keeley
Senior Editor

Janet Lehr
Associate Editor

Thomas B. Kingery III
Information Technology Director

The *Water Encyclopedia* is available online at
http://www.mrw.interscience.wiley.com/eow/

WILEY-
INTERSCIENCE

A John Wiley & Sons, Inc., Publication

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data is available.

Lehr, Jay
Water Encyclopedia: Surface and Agricultural Water

ISBN 0-471-73685-6
ISBN 0-471-44164-3 (Set)

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

WATER ENCYCLOPEDIA

OCEANOGRAPHY; METEOROLOGY; PHYSICS AND CHEMISTRY; WATER LAW; AND WATER HISTORY, ART, AND CULTURE

Jay Lehr, Ph.D.
Editor-in-Chief

Jack Keeley
Senior Editor

Janet Lehr
Associate Editor

Thomas B. Kingery III
Information Technology Director

The *Water Encyclopedia* is available online at
http://www.mrw.interscience.wiley.com/eow/

WILEY-
INTERSCIENCE

A John Wiley & Sons, Inc., Publication

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data is available.

Lehr, Jay
Water Encyclopedia: Oceanography; Meteorology; Physics and Chemistry; Water Law; and Water History, Art, and Culture

ISBN 0-471-73684-8
ISBN 0-471-44164-3 (Set)

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

WATER ENCYCLOPEDIA

GROUND WATER

Jay Lehr, Ph.D.
Editor-in-Chief

Jack Keeley
Senior Editor

Janet Lehr
Associate Editor

Thomas B. Kingery III
Information Technology Director

The *Water Encyclopedia* is available online at
http://www.mrw.interscience.wiley.com/eow/

WILEY-
INTERSCIENCE

A John Wiley & Sons, Inc., Publication

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data is available.

Lehr, Jay
  Water Encyclopedia: Ground Water

ISBN 0-471-73683-X
ISBN 0-471-44164-3 (Set)

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

PREFACE

No one really questions that water is the life blood of humankind. We all remain amazed that the existence of water separates our planet from every other we have thus far viewed in our universe. We can arguably do without every naturally occurring molecule on the earth except water. Life was clearly formed within water and exists in one way or another on water.

Few people in the developed world give this simple fact of life much thought. We have an abundance of water for most of our needs, although some agricultural areas, a few municipalities, and some rural families, at times are strapped for the full amount of water they desire. In the developing world, however, in some locations, the collection and distribution of water is a critical part of every day life, with many women devoting the major portion of their day to the provision of water for their family. In other villages, the construction and protection of a single well can be the primary focus of community needs.

In general, the actual delivery of water for a myriad of uses followed by its disposal is taken for granted by all but the individuals charged with carrying out these often amazing tasks. In this volume of the *Water Encyclopedia*, a collaborative effort of hundreds of people from dozens of countries, we have tried to cover every conceivable topic of interest to people in every walk of life, be they students, researchers, professionals, or just plain folks with an intellectual curiosity about our elixir of life.

We are concerned in this volume with the actual delivery of water to the home by the home owner (subjects include disinfection, corrosion control, nitrates, gray water, septic tanks, and windmills), from the municipal supplier (and their challenges, including distribution, filtration, zebra mussels, reverse osmosis, cryptosporidium, arsenic, and public confidence), to industry and its special needs (such as microfiltration, effluent discharge, reuse, energy, nuclear reactor coolants, and even golf course irrigation), and of course the disposal of our used water in a safe and efficient manner (subjects such as air stripping, bioassays, flotation, sludge, bioavailability, and wetlands). We hope that no reader can stump the experts, which means that we have covered every area of interest. However, we know that this goal is not currently possible, but in coming years and in coming editions on paper and on the World Wide Web, we will more closely approach it. Let us know on our website where our information may be incomplete, and we will be sure to follow-up and fill in the gap in the future.

The contributors to this volume have freely offered their expertise to this project. Some have focused their information on those in need of complete and often complex detail of their subject matter. Others have followed a middle road for a wider audience, and still others believed that a very simple approach to conveying information on their subject was best.

The reader may find all approaches on the same subject matter because the editors frequently sought overlapping information presented from different points of view. We are confident that most people will find their needs met.

Through this encyclopedia, which is the most comprehensive effort ever undertaken on behalf of this most important subject, we hope that we will collectively make a contribution that will enhance the distribution and use of our water supplies in ever safer and more efficient ways.

Jay Lehr
Jack Keeley

PREFACE

Cities, towns, states, and nations must manage their water resources wisely from both a quality and a quantitative perspective.

If we do otherwise and manage them with a narrow perspective, the public's needs will not be adequately met. In this volume of the *Water Encyclopedia*, authors from around the world have described a myriad of problems relating to individual water bodies as well as to geographic water resources and their management dilemmas.

Humans and other living creatures contribute to our water quality problems. Neither can be fully controlled. Even the nature of contaminant sources and programs for their elimination can be difficult to design. This volume contains the best and brightest ideas and case studies relating to the areas of water quality and resource management problems.

Quality problems deal with a diverse suite of subjects ranging widely from acid mine drainage to biosorption, colloids, eutrophication, protozoa, and recalcitrant compounds. Resource management features drought studies, flood control, river basin management, perennial overdraft, water banking, and a host of other subjects.

The perspective of scientists from nearly every continent of the world offers a truly catholic view of attitudes and biases harbored in different regions and how they affect scientific and regulatory outcomes.

The editors cannot imagine what has been left out, but we know of course that readers will at times come up short of finding an exact match to a problem they face. We hope they will contact us at our website and allow us the opportunity of adding additional subjects to our encyclopedia. At the same time, the reader will understand that many subjects in the area of water quality may have been addressed in our Surface Water category. It was often difficult to determine where an investigator would be more likely to look for a piece of information. (The complete index of all five volumes appears in the Ground Water volume as well as on our website.)

We trust all users of this encyclopedia will find it detailed, informative, and interesting. Not only are a wide range of subjects treated, but authors choose varying approaches to presenting their data to readers who may be professionals, students, researchers, as well as individuals simply satisfying their intellectual curiosity. We hope we are successfully serving all of these populations in some useful way.

Jay Lehr
Jack Keeley

PREFACE

Surface water and agricultural water are uniquely associated as they provide many of our basic needs, including food and fiber, power, transportation, and recreation. Like other volumes in the *Water Encyclopedia*, we have selected articles as varied in content as they are in technical sophistication. To this end, the reader will also recognize that single topics are occasionally duplicated at varying levels of scientific acumen.

Articles are also provided that demonstrate that surface and agricultural water are associated in yet another way: They must be used efficiently and protected to assure their productiveness far into the future. For example, agricultural water use efficiency is discussed from several viewpoints with respect to irrigation technology. River basin planning is approached in diverse ways, including stream classifications, watershed hydrology, modeling, erosion control, and water conservation.

We have necessarily included articles addressing issues of quality with respect to both surface and agricultural water. In addition to an assessment of pollution outflow from agricultural areas, the quality of reclaimed irrigation is addressed from both chemical and microbial standpoints. Watershed areas are examined according to their contribution and vulnerability to contamination, flooding, sediment transport, and trace elements.

Discourses on surface water would not be complete without articles related to fish. Accordingly, we have included articles on fish growth, fisheries, fishponds, and the use of fish scales in toxicological studies as examples.

Another vital area of study in this volume is perhaps best described as the practical side. These areas are of less esoteric origins, including salt tolerance of plants, irrigation wells, weed control, tile drains, and moisture content in to agriculture. Similar topics in surface water include riparian systems, reservoir design, wetlands, lakes, levees, and the unit hydrograph.

Finally, and *appropriately, this volume of the Water Encyclopedia* contains articles on specific water bodies and the consequences of their being. Included are the Aral Sea, the Ganga River of India, the Great Lakes, and the Yellow River in China, only to name a few. Here, too, the association of surface water and agricultural water are reinforced. This volume presents an important segment of the topic of water. We believe that the reader's educational pursuits will be well met by its contents.

Jay Lehr
Jack Keeley

PREFACE

No natural molecule on the planet is more fascinating than water. It has unique properties ranging from the unusual angles formed between its two hydrogen ions and its single oxygen molecule, to the fact that unlike most substances, it expands when it freezes rather than shrinks and reaches a maximum density as a liquid 4°F above its freezing point. These and many other aspects of the special physics and chemistry are described in this volume, including the impact of a wide variety of chemicals occurring in water, osmosis, diffusion, hydration, isotope exchange, along with the fun physics of the mariotte bottle.

Equally fascinating are the many unusual physical and chemical encounters in both the ocean and the atmosphere. Although oceanography and meteorology are frequently considered separate sciences from hydrology, their limited inclusion in the *Water Encyclopedia* was deemed necessary to tell the complete story. Tidal changes, benthic nutrients, the sea floor, el nino, sea level, and ocean/climate relationships make up but a few of the oceans fascinating stories, whereas water spouts, hurricanes, monsoons, droughts, sublimation, and barometric efficiency just touch the tip of what this volume has in store in the area of meteorology.

But water has a wonderful human face as well. We have reached around the world to describe the history of water and its role in the development of civilizations and the many beliefs held about it. As society developed, the distribution of water needed supervision, which lead to a wide variety of water laws we have attempted to categorize and describe in an interesting and meaningful way.

We are equally proud of our open-minded effort to describe the role that water has played in art and culture. We have attempted not be judgmental, with stories of water forms and water intelligence along with some medical theories and, of course, the wonderful descriptions of early water clocks. This volume is a true intellectual cornucopia of water in the life of humankind on a personal level.

We are confident that in the coming years and editions of the *Water Encyclopedia*, this volume will expand with more participation from individuals working in unusual fields relating to water.

Jay Lehr
Jack Keeley

PREFACE

Throughout history, groundwater has played a major role in providing the resource needs of the world. It accounts for 97% of the world's freshwater and serves as the base flow for all streams, springs, and rivers. In the United States, one half of the population relies on groundwater for its drinking water and is the sole source of supply for 20 of the 100 largest cities. Well over 90% of rural America is totally dependent on groundwater. An inventory of the total groundwater resources in the United States can be visualized as being equal to the flow of the Mississippi River at Vicksburg for a period of 250 years.

One of the first groundwater scientists was a French engineer who was in charge of public drinking water in Dijon. In 1856, Henri Darcy conducted experiments and published mathematical expressions describing the flow of water through sand filters. His work remains one of the cornerstones of today's groundwater hydrologists. At about the same time, a Connecticut court ruled that the influences of groundwater movement are so secret, changeable, and uncontrollable that they could not be subject to regulations of law, nor to a system of rules, as had been done with surface streams.

In this volume of the *Water Encyclopedia*, we have attempted to erase the ignorance that existed in the early years of groundwater science by presenting the most current knowledge on the subject as provided by authors from around the globe. In addition to excellent articles from many American scholars, equally superb writings from such diverse countries as England, Nigeria, India, Iran, Thailand, and Greece are provided.

As the origins of the selected articles are diverse, so are the subjects of discussion. Along with straightforward descriptions of basic groundwater concepts (drawdown around pumping wells, hydraulic head, field capacity, and flow), the reader is introduced to more complex subjects (isotope technologies, aquifer tests, *in situ* remediation, tritium dating, modeling, and geophysical properties). There are also articles for more practical applications (well maintenance, subsurface drainage, nitrate contamination, tracer tests, well yields, and drilling technologies). Of course, for the more fanciful reader, we have selected articles that remind us of the way windmills sounded at night, the ancient use of qanats in Persia to provide sustainable groundwater resources, and the development of Darcy's Law.

In the end, we feel that the information provided will afford an educational home for readers approaching the *Water Encyclopedia* from a variety of needs as well as different levels of scientific acumen. We are also confident that many readers will simply be expanding their knowledge base by these sets of enjoyable reading.

Jay Lehr
Jack Keeley

ACKNOWLEGMENTS

The editors of this encyclopedia wish to acknowledge two special people without whose assistance this enormous undertaking could not have been initiated, administered, or completed.

Bob Esposito, Executive Editor, placed the project before us and convinced us it would be as exciting and rewarding as time has proved.

Jonathan T. Rose, Editorial Program Coordinator, was the true backbone of our team. His warm, accommodating, and skillful management helped us overcome each and every problem. His interaction with contributors and his knowledge of the publishing process offered us the security and comfort required to persevere through the four years required to complete this work. We will forever be in his debt.

Jay Lehr
Jack Keeley
Janet Lehr
Thomas B. Kingery III

CONTRIBUTORS

Segun Michael Ade Adelana, *University of Ilorin, Ilorin, Kwara State, Nigeria,* Water and Human Health, Nitrate Health Effects

Franklin Agardy, *Forensic Management Association, San Mateo, California,* Chemical Drinking Water Standards, Past, Present, and Future

Rasheed Ahmad, *Khafra Engineering Consultants, Atlanta, Georgia,* Filtration With Granular Media, Hydraulic Design of Water Distribution Storage Tanks, Ozone With Activated Carbon for Drinking Water Treatment, Particulate Removal, Filtration Water Treatment, Particulate Matter Removal by Filtration and Sedimentation, Synthetic and Natural Organic Removal by Biological Filtration, Water Filtration, Flocculation, Gravity Separation/Sedimentation, Particulate Matter Removal by Coagulation, Dechlorination, Wastewater Reclamation and Reuse, Wastewater Reclamation and Reuse Treatment Technology

Julia E. Aizpún de Moreno, *Universidad Nacional de Mar del Plata, Mar del Plata, Argentina,* Macrophytes as Biomonitors of Polychlorinated Biphenyls

Imram Ali, *National Institute of Hydrology, Roorkee, India,* Wastewater Treatment and Recycling Technologies

George R. Alther, *Biomin, Inc., Ferndale, Michigan,* The Role of Organoclay in Water Cleanup

Pietro Argurio, *Universita della Calabria, Rende, Italy,* Ultrafiltration—Complexation in Wastewater Treatment

Scott Arthur, *Heriot-Watt University, Edinburgh, Scotland, United Kingdom,* Roof Drainage Hydraulics

Samuel C. Ashworth, *Idaho National Engineering and Environmental Laboratory, Idaho Falls, Idaho,* Polycyclic Aromatic Hydrocarbons, Water Treatment in Spent Nuclear Fuel Storage, Hydrocarbon Treatment Techniques, Metal Speciation and Mobility as Influenced by Landfill Disposal Practices

Kwok-Keung Au, *Greeley and Hansen, Chicago, Illinois,* Removal of Pathogenic Bacteria, Viruses, and Protozoa, Granular Bed and Precoat Filtration

Ann Azadpour-Keeley, *United States Environmental Protection Agency, Ada, Oklahoma,* Virus Transport in the Subsurface

Christine L. Bean, *University of New Hampshire, Durham, New Hampshire,* Giardiasis

Asbjorn Bergheim, *RF-Rogaland Research, Stavanger, Norway,* Waste Treatment in Fish Farms

Vipin Bhardwaj, *NDWC Engineering Scientist,* Diatomaceous Earth Filtration for Drinking Water, Reservoirs, Towers, and Tanks Drinking Water Storage Facilities, Pumps, Water Meters, Preventing Well Contamination, Cross Connection and Backflow Prevention, Repairing Distribution Line Breaks

William J. Blanford, *Louisiana State University, Baton Rouge, Louisiana,* Review of Parasite Fate and Transport in Karstic Aquifers

C. Blöcher, *Saarland University, Saarbrücken, Germany,* Application of Microfiltration to Industrial Wastewaters

Patrick Bond, *Kensington, South Africa,* The Economics of Water Resources Allocation

T.R. Bott, *University of Birmingham, Birmingham, United Kingdom,* Industrial Cooling Water—Scale Formation, Industrial Cooling Water—Corrosion, Industrial Cooling Water—Biofouling, Energy Dissipation

Jeanine L. Boulter-Bitzer, *University of Guelph, Guelph, Ontario, Canada,* Molecular-Based Detection of *Cryptosporidium Parvum* in Water

Brenda Boutin, *National Center for Environmental Assessment, Cincinnati, Ohio,* Health Effects of Microbial Contaminants and Biotoxins in Drinking Water, Health Effects of Commonly Occurring Disinfection Byproducts in Municipal Water Supplies

Jacqueline Brabants, *University of New Hampshire, Durham, New Hampshire,* Cryptosporidium

Alexander Brinker, *Fischereiforschungsstelle des Landes Baden-Württemberg, Langenargen, Germany,* Waste Treatment in Fish Farms

B.M. Brouckaert, *University of Natal, Durban, South Africa,* Key Causes of Drinking Water Quality Failure in a Rural Small Water Supply of South Africa

Robert Bruce, *National Center for Environmental Assessment, Cincinnati, Ohio,* Health Effects of Microbial Contaminants and Biotoxins in Drinking Water, Health Effects of Commonly Occurring Disinfection Byproducts in Municipal Water Supplies

Charlie Bryce, *Napier University, Edinburgh, Scotland, United Kingdom,* The Arsenic Drinking Water Crisis in Bangladesh

Mario O. Buenfil-Rodriguez, *National University of Mexico & Mexican Institute of Water Technology, Cuernavaca, Morelos, Mexico,* Public Water Supply World, Water Distribution System Operation, Water Meter

Zia Bukhari, *American Water, Quality Control and Research Laboratory, Belleville, Illinois,* Potential Risks of Waterborne Transmission of *Escherichia coli* O157:H7, Measuring *Cryptosporidium Parvum* Oocyst Inactivation Following Disinfection With Ultraviolet Light

Michael A. Butkus, *United States Military Academy, West Point, New York,* Biochemical Oxygen Demand

D. Butler, *Imperial College, London, United Kingdom,* Gray Water Reuse in Households

Rebecca L. Calderon, Ph.D., *National Health and Environmental Effects Laboratory Research Triangle Park, North Carolina,* Improving Waterborne Disease Surveillance

Christine M. Carey, *University of Guelph, Guelph, Ontario, Canada,* Molecular-Based Detection of *Cryptosporidium Parvum* in Water

Peter S. Cartwright, P.E., *Minneapolis, Minnesota,* Water Reuse

Frank J. Castaldi, *Brown and Caldwell, Austin, Texas,* Degradation of Chloro-Organics and Hydrocarbons, Aqueous Behavior of Elements in a Flue Gas Desulfurization Sludge Disposal Site

Frédéric J.F. Chagnon, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Chemically Enhanced Primary Treatment of Wastewater

Chavalit Chaliraktrakul, *Thammasat University, Pathumthani, Thailand,* A Real-Time Hydrological Information System for Cities

Lena Ciric, *Centre for Ecology and Hydrology, Oxford, United Kingdom,* Molecular Biology Tools for Monitoring Biodiversity in Wastewater Treatment Plants

Gunther F. Craun, P.E., M.P.H., D.E.E., *Gunther F. Craun and Associates, Staunton, Virginia,* Improving Waterborne Disease Surveillance

James Crocker, *Richland, Washington,* Radioactive Waste

Sukumar Devotta, *Environmental Genomics Unit, National Environmental Engineering, Research Institute, CSIR, Nagpur, India,* Water Security: An Emerging Issue

John E. Dodes, *Forest Hills, New York,* Fluoridation

L. Donald Duke, *University of South Florida, Tampa, Florida,* Effluent Limitations and the NPDES Permit

Timothy J. Downs, *Clark University, Worcester, Massachusetts,* Municipal Watersheds, Integrated Capacity Building Needs for Water Supply and Wastewater Sanitation

Jorg E. Drewes, *Colorado School of Mines, Golden, Colorado,* Wastewater Reclamation and Reuse Research

James B. Duncan, *Kennewick, Washington,* Aqueous Reactions of Specific Organic Compounds with Ozone

María-del-Carmen Durán-de-Bazúa, *UNAM, National Autonomous University of Mexico, México D.F., México,* Sugarcane Industry Wastewaters Treatment, Use of Anaerobic-Aerobic Treatment Systems for Maize Processing Plants

C. Erickson, *New Mexico State University, Las Cruces, New Mexico,* Land Applications of Wastewater in Arid Lands: Theory and Case Studies

T. Erwe, *Saarland University, Saarbrücken, Germany,* Application of Microfiltration to Industrial Wastewaters, Bonding of Toxic Metal Ions

Susan L. Franklin, *Tetra Tech MPS, Ann Arbor, Michigan,* Consumer Confidence Reports

From *Drinking Water Regulation and Health*, Wiley 2003, 1962 U.S. Public Health Service Standards

Floyd J. Frost, Ph.D., *The Lovelace Institutes, Albuquerque, New Mexico,* Improving Waterborne Disease Surveillance

Michael H. Gerardi, *Linden, Pennsylvania,* Microbial Foaming in the Activated Sludge Process, Biological Phosphorus Removal in the Activated Sludge Process, Nitrification in the Activated Sludge Process, Denitrification in The Activated Sludge Process

Sanjiv Gokhale, *Vanderbilt University, Nashville, Tennessee,* Basics of Underground Water and Sewer Pipeline Assessment, Repair, and Rehabilitation, Trenchless Repair and Rehabilitation Techniques, Problems Encountered During Pipe Repair and Renewal

Janusz Guziur, *University of Warmia and Mazury in Olsztyn, Olsztyn, Poland,* Role of Small Water Reservoirs in Environment

Donald R.F. Harleman, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Chemically Enhanced Primary Treatment of Wastewater

Arnim R.H. Hertle, *GHD Pty Ltd, Wembley, Washington, Australia,* Fine Bubble Diffused Air Aeration Systems, Aeration

James Higgins, *USDA-ARS, Beltsville, Maryland,* Threat Agents and Water Biosecurity

Arthur M. Holst, *Philadelphia Water Department, Philadelphia, Pennsylvania,* Windmills

Robert M. Hordon, *Rutgers University, Piscataway, New Jersey,* Bottled Water

Marsha A. Hosner, *DHI, Newtown, Pennsylvania,* Approaches for Securing a Water Distribution System

Look Hulshoff Pol, *Lettinga Associates Foundation, Wageningen, The Netherlands,* Anaerobic Wastewater Treatment

M. Eng Nguyen Quang Hung, *Asian Institute of Technology, Pathumthani, Thailand,* A Real-Time Hydrological Information System for Cities

Herbert Inhaber, *Risk Concepts, Las Vegas, Nevada,* Water Use in Energy Production

Enos C. Inniss, *University of Texas, San Antonio, Texas,* Use of Redox Potentials in Wastewater Treatment

Th. A. Ioannidis, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* Solidification/Stabilization of Hazardous Solid Wastes

Irena B. Ivshina, *Russian Academy of Sciences, Perm, Russia,* Microbial Foaming and Bulking in Activated Sludge Plants

James A. Jacobs, *Environmental Bio-Systems, Inc., Mill Valley, California,* Water Impacts from Construction Sites

Kauser Jahan, *Rowan University, Glassboro, New Jersey,* Detergents

C.K. Jain, *National Institute of Hydrology, Roorkee, India,* Wastewater Treatment and Recycling Technologies

H.A. Jenner, *KEMA Power Generation and Sustainables, Arnhem, The Netherlands,* Chlorine and Chlorine Residuals

Dick de Jong, *IRC International Water and Sanitation Centre, Delft, The Netherlands,* Multistage Filtration: An Innovative Water Treatment Technology

Mohammad R. Karim, *American Water, Quality Control and Research Laboratory, Belleville, Illinois,* Microbiological Concerns of Drinking Water Distribution Systems

A. Katsoyiannis, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* The Fate of Persistent Organic Pollutants (POPs) in The Wastewater Treatment Process

I. Katsoyiannis, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* Technologies for Arsenic Removal from Contaminated Water Sources

Absar A. Kazmi, *Department of Civil Engineering, Roorkee, Uttaranchal, India,* Water and Wastewater Properties and Characteristics, Reclaimed Water, Sewage, Wastewater Characterization

Jack W. Keeley, *Environmental Engineer, Ada, Oklahoma,* Virus Transport in the Subsurface

Jaehong Kim, *Georgia Institute of Technology, Atlanta, Georgia,* Municipal Water Supply: Ozonation

K. Thomas Klasson, *Oak Ridge National Laboratory, Oak Ridge, Tennessee,* Mercury Removal From Complex Waste Waters

Jeff Kuo, *Cerritos, California,* Air Stripping

Maria S. Kuyukina, *Russian Academy of Sciences, Perm, Russia,* Microbial Foaming and Bulking in Activated Sludge Plants

Zacharia Michael Lahlou, *Technical Assistance Consultant,* Valves, Water Hammer, Point-of-Use/Point-of-Entry Systems (POU/POE), Leak Detection and Water Loss Control, System Control and Data Acquisition (SCADA)

Guenter Langergraber, *BOKU—University of Natural Resources and Applied Life Sciences, Vienna, Austria,* Ecological Wastewater Management, Wastewater Treatment—Small Scale, Constructed Wetlands

Lars Chr Larsen, *DHI Water and Environment, Hørsholm, Denmark,* A Real-Time Hydrological Information System for Cities

Mark W. LeChevallier, *American Water, Voorhees, New Jersey,* Microbiological Concerns of Drinking Water Distribution Systems, Potential Risks of Waterborne Transmission of *Escherichia coli* O157:H7

Hung Lee, *University of Guelph, Guelph, Ontario, Canada,* Molecular-Based Detection of *Cryptosporidium Parvum* in Water

Gatze Lettinga, *Wageningen University and Research Center, Wageningen, The Netherlands,* Anaerobic Sewage Treatment

Srinivasa Lingireddy, *University of Kentucky, Lexington, Kentucky,* Design of Water Distribution Systems

Gerasimos Lyberatos, *University of Patras, Patras, Greece,* Nitrification of Potable Water Using Trickling Filters

Bruce A. Macler, *Toxicologist, U.S. Environmental Protection Agency, San Francisco, California,* Application of Risk Assessments in Crafting Drinking Water Regulations

N. Makala, *University of Fort Hare, Alice, South Africa,* Assessing the Bactericidal Efficiency of Polydex for the Disinfection of Drinking Water in Rural Areas of South Africa, Key Causes of Drinking Water Quality Failure in a Rural Small Water Supply of South Africa

Babalola Makinde-Odusola, *Public Utilities, Riverside, California,* Well Head Protection, Drinking Water Quality Standards (DWQS)-United States

Joe D. Manous, Jr., *United States Military Academy, West Point, New York,* Biochemical Oxygen Demand

Ole Mark, *DHI Water and Environment, Hørsholm, Denmark,* A Real-Time Hydrological Information System for Cities

Kostas A. Matis, *Aristotle University, Thessaloniki, Greece,* Bonding of Toxic Metal Ions, Application of Microfiltration to Industrial Wastewaters, Flotation as A Separation Process

Paul Mavros, *Aristotle University, Thessaloniki, Greece,* Mixing and Agitation in Water Treatment Systems

V. Mavrov, *Saarland University, Saarbrücken, Germany,* Application of Microfiltration to Industrial Wastewaters, Bonding of Toxic Metal Ions

Steve Maxwell, *TechKNOWLEDGEy Strategic Group,Boulder, Colorado,* Ten Key Trends That Will Shape the Future of the World Water Industry, The State of the Water Industry—2004

Kevin S. McLeary, *Pennsylvania Dept. of Environmental Protection, Harrisburg, Pennsylvania,* Wastewater Treatment Processes and Water Reuse, Domestic Sewage

Sue McLeod, *William Forrest and Sons, Omoa Works, Newarthill, Motherwell, United Kingdom,* Odor Abatement in Wastewater Treatment Plants

Fayyaz A. Memon, *Imperial College, London, United Kingdom,* Gray Water Reuse in Households

Mirta L. Menone, *Universidad Nacional de Mar del Plata, Mar del Plata, Argentina and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina,* Macrophytes as Biomonitors of Polychlorinated Biphenyls

Chris Metzgar, *Graphic Designer,* Reservoirs, Towers, and Tanks Drinking Water Storage Facilities

J.G. Mexal, *New Mexico State University, Las Cruces, New Mexico,* Land Applications of Wastewater in Arid Lands: Theory and Case Studies

C. Mfenyana, *University of Fort Hare, Alice, South Africa,* Inadequate Treatment of Wastewater: A Source of Coliform Bacteria in Receiving Surface Water Bodies in Developing Countries—Case Study: Eastern Cape Province of South Africa

Z. Michael Lahlou, *Technical Assistance Consultant,* Water Quality in Distribution Systems

Karina S.B. Miglioranza, *Universidad Nacional de Mar del Plata, Mar del Plata, Argentina and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina,* Macrophytes as Biomonitors of Polychlorinated Biphenyls

Mel Mirliss J., *International Diatomite Producers Association,* Diatomaceous Earth Filtration for Drinking Water

M. S. Mohan Kumar, *Indian Institute of Science, Bangalore, India,* Modeling Chlorine Residuals in Urban Water Distribution Systems

Dinesh Mohan, *Gomti Nagar, Lucknow, Uttar Pradesh India,* Granular Activated Carbon, Competitive Adsorption of Several Organics and Heavy Metals on Activated Carbon in Water

T. C. Molden, *Magnatech Corporation, Fort Wayne, Indiana,* Zebra Mussel Control Without Chemicals

Raffaele Molinari, *Università della Calabria, Rende, Italy,* Photocatalytic Membrane Reactors in Water Purification, Ultrafiltration—Complexation in Wastewater Treatment

M.N.B. Momba, *University of Fort Hare, Alice, South Africa,* Key Causes of Drinking Water Quality Failure in a Rural Small Water Supply of South Africa, Assessing the Bactericidal Efficiency of Polydex for the Disinfection of Drinking Water in Rural Areas of South Africa, Inadequate Treatment of Wastewater: A Source of Coliform Bacteria in Receiving Surface Water Bodies in Developing Countries—Case Study: Eastern Cape Province of South Africa

John E. Moore, *Hydrologic Consultant, Denver, Colorado,* Septic Tank Systems

Víctor J. Moreno, *Universidad Nacional de Mar del Plata, Mar del Plata, Argentina,* Macrophytes as Biomonitors of Polychlorinated Biphenyls

Chandrika Moudgal, *National Center for Environmental Assessment, Cincinnati, Ohio,* Health Effects of Microbial Contaminants and Biotoxins in Drinking Water, Health Effects of Commonly Occurring Disinfection Byproducts in Municipal Water Supplies

G. R. Munavalli, *Walchand College of Engineering, Sangli, India,* Modeling Chlorine Residuals in Urban Water Distribution Systems

Michael Muntisov, *GHD Pty Ltd., Melbourne, Victoria, Australia,* Guide to Selection of Water Treatment Processes

Susan Murcott, *Massachusetts Institute of Technology, Cambridge, Massachusetts,* Household Drinking Water Treatment and Safe Storage

National Drinking Water Clearinghouse, Treatment for Technologies for Small Drinking Water Systems, Disinfection, Filtration, Corrosion Control, Ion Exchange and Demineralization, Organic Removal, Package Plants, Water Treatment Plant Residuals Management, Lime Softening, Iron and Manganese Removal, Water Conservation Measures, Membrane Filtration, Ozone, Radionuclides, Slow Sand Filtration, Ultraviolet Disinfection

Abid M. Nasser, *Water Quality Research Laboratory, Ministry of Health, Tel-Aviv, Israel,* Persistence of Pathogens in Water

Louis H. Nel, *University of Pretoria, Pretoria, Gauteng, South Africa,* Emerging Waterborne Infectious Diseases

Robert Y. Ning, *King Lee Technologies, San Diego, California,* Reverse Osmosis, Membrane Foulants, Arsenic in Natural Waters, Reverse Osmosis, Process Chemistry, Reverse Osmosis, Membrane Cleaning

Office of Water—United States Environmental Protection Agency, EPA's National Pretreatment Program, 1973–2003: Thirty Years of Protecting The Environment

Oladele Ogunseitan, *University of California, Irvine, California,* Pharmaceuticals in Water Systems

A. Okeyo, *Programme Unit of Biochemistry and Microbiology, University of Fort Hare, Alice, South Africa,* Assessing the Bactericidal Efficiency of Polydex for the Disinfection of Drinking Water in Rural Areas of South Africa

Daniel A. Okun, *(from Drinking Water Regulation and Health, Wiley 2003) University of North Carolina, Chapel Hill, North Carolina,* Drinking Water and Public Health Protection

Lindell E. Ormsbee, *University of Kentucky, Lexington, Kentucky,* Design of Water Distribution Systems

Aisling D. O'Sullivan, *University College Dublin, Belfield, Ireland,* Using Ecosystem Processes in a Constructed Wetland to Treat Mine Wastewater in Ireland

Marinus L. Otte, *University College Dublin, Belfield, Ireland,* Using Ecosystem Processes in a Constructed Wetland to Treat Mine Wastewater in Ireland

David Lloyd Owen, *Envisager, Llangoedmor, Ceredigion, United Kingdom,* Private Sector Participation, Marketing and Corporate Strategies in Municipal Water Supply and Sewerage

L. Palmisano, *Università di Palermo, Palermo, Italy,* Photocatalytic Membrane Reactors in Water Purification

E.N. Peleka, *Aristotle University, Thessaloniki, Hellas,* Bonding of Toxic Metal Ions

Jim Philp, *Napier University, Edinburgh, Scotland, United Kingdom,* The Arsenic Drinking Water Crisis in Bangladesh, Odor Abatement in Wastewater Treatment Plants, Molecular Biology Tools for Monitoring Biodiversity in Wastewater Treatment Plants, Landfill

Laurel Phoenix, *Green Bay, Wisconsin,* Extraterritorial Land Use Control to Protect Water Supplies

G. Picchioni, *New Mexico State University, Las Cruces, New Mexico,* Land Applications of Wastewater in Arid Lands: Theory and Case Studies

Nicholas J. Pokorny, *University of Guelph, Guelph, Ontario, Canada,* Molecular-Based Detection of *Cryptosporidium Parvum* in Water

Kelly Pollack, *University of California, Irvine, California,* Pharmaceuticals in Water Systems

Christopher Polley, *William Forrest and Sons, Omoa Works, Newarthill, Motherwell, United Kingdom,* Odor Abatement in Wastewater Treatment Plants

Rathnavel Ponnuswami, *CARE2, Redwood City, California,* Water Hammer: Quantitative Causes and Effects

Prakhar Prakash, *Pennsylvania State University, University Park, Pennsylvania,* Selective Coagulant Recovery from Water Treatment Plant Residuals Using the Domain Membrane Process

C.A. Prochaska, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* Municipal Storm Water Management, Combined Sewer Overflow Treatment

Hemant J. Purohit, *Environmental Genomics Unit, National Environmental Engineering, Research Institute, CSIR, Nagpur, India,* Water Security: An Emerging Issue

S. Rajagopal, *Radboud University Nijmegen, Nijmegen, The Netherlands,* Chlorine and Chlorine Residuals

D. Ramalingam, *University of Kentucky, Lexington, Kentucky,* Design of Water Distribution Systems

Niranjanie Ratnayake, *University of Moratuwa, Moratuwa, Sri Lanka,* Water Disinfection Using UV Radiation—A Sri Lankan Experience

Eugene R. Reahl, *Ionics, Inc., Watertown, Massachusetts,* Answering the Challenge

Robin J. Reash, *American Electric Power, Water & Ecological Resource Services, Columbus, Ohio,* Electric Generating Plants—Effects of Contaminants

Bethany Reed, *NESC Graphic Designer,* Pumps, Cross Connection and Backflow Prevention

Steven J. Renzetti, *Brock University, St. Catharines, Ontario, Canada,* Economics of Residential Water Demands, Economics of Industrial Water Demands

Susan Richardson, *U.S. Environmental Protection Agency,,* What is in Our Drinking Water?

Ingrid Ritchie, *Indiana University-Purdue University, Indianapolis, Indiana,* Magnetic Water Conditioning

Paul D. Robillard, *World Water Watch, Cambridge, Massachusetts,* Methods of Reducing Radon in Drinking Water

D.S. Rodriguez, *New Mexico State University, Las Cruces, New Mexico,* Land Applications of Wastewater in Arid Lands: Theory and Case Studies

Stephen J. Rooklidge, *Aurora, Oregon,* Multistage Drinking Water Filtration, Slow Sand Filtration and the Impact of *Schmutzdecke*

David L. Russell, *Global Environmental Operations, Inc., Lilburn, Georgia,* Introduction to Wastewater Modeling and Treatment Plant Design, Practical Applications of Wastewater Modeling and Treatment Plant Design

Z. Samani, *New Mexico State University, Las Cruces, New Mexico,* Land Applications of Wastewater in Arid Lands: Theory and Case Studies

Petros Samaras, *Chemical Process Engineering Research Institute, Thermi-Thessaloniki, Greece,* Landfill Leachates: Part 2: Treatment, Landfill Leachates, Part I: Origin and Characterization

Technological Educational Institute of West Macedonia, Kozani, Greece Evaluation of Toxic Properties of Industrial Effluents by on-Line Respirometry

C. Samara, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* The Fate of Persistent Organic Pollutants (POPs) in The Wastewater Treatment Process

T. Sammis, *New Mexico State University, Las Cruces, New Mexico,* Land Applications of Wastewater in Arid Lands: Theory and Case Studies

Charles H. Sanderson, *Magnatech Corporation, Fort Wayne, Indiana,* Zebra Mussel Control Without Chemicals, Physical Water Conditioning

Zane Satterfield, *NDWC Engineering Scientists,* Water Meters

Lucas Seghezzo, *Wageningen University and Research Center, Wageningen, The Netherlands,* Anaerobic Sewage Treatment

Arup K. SenGupta, *Lehigh University, Bethlehem, Pennsylvania,* Selective Coagulant Recovery from Water Treatment Plant Residuals Using the Domain Membrane Process

William E. Sharpe, *Pennsylvania State University, University Park, Pennsylvania,* Methods of Reducing Radon in Drinking Water

M. Siddiqui, *University of Utah, Salt Lake City, Utah,* Ultraviolet Irradiation, Ozone–Bromide Interactions

Kunwar P. Singh, *Dinesh Mohan and Kunwar P. Singh,* Gomti Nagar, Lucknow, Uttar Pradesh IndiaGranular Activated Carbon, Competitive Adsorption of Several Organics and Heavy Metals on Activated Carbon in Water

Marin Slunjski, *Orica Watercare, Regency Park, SA, Australia,* Ion Exchange—Use of Magnetic Ion Exchange Resin For DOC Removal

Jo Smet, *IRC International Water and Sanitation Centre, Delft, The Netherlands,* Multistage Filtration: An Innovative Water Treatment Technology

Stuart A. Smith, *Smith-Comeskey Ground Water Science LLC, Upper Sandusky, Ohio,* Evaluation of Microbial Components of Biofouling

Mervyn Smyth, *Centre for Sustainable Technologies, Newtownabbey, United Kingdom,* Domestic Solar Water Heaters

Muhammad Sohail, *Loughborough University, Leicestershire, United Kingdom,* Domestic Water Supply—Public–Private Partnership

Ludovico Spinosa, *National Research Council, BARI, Italy,* Sludge Treatment and Disposal

Fiona M. Stainsby James C. Philp Sandra Dunbar, *Napier University, Edinburgh, Scotland, United Kingdom,* Microbial Foaming and Bulking in Activated Sludge Plants

Bradley A. Striebig, *Gonzaga University, Spokane, Washington,* Sewerage Odors—How to Control

Patrick Sullivan, *Forensic Management Association, San Mateo, California,* Chemical Drinking Water Standards, Past, Present, and Future

Bryan R. Swistock, *Pennsylvania State University, University Park, Pennsylvania,* Methods of Reducing Radon in Drinking Water

David J. Tonjes, *Cashin Associates, PC, Hauppauge, New York,* New York City Harbor Survey

Jack T. Trevors, *University of Guelph, Guelph, Ontario, Canada,* Molecular-Based Detection of *Cryptosporidium Parvum* in Water

Konstantinos P. Tsagarakis, *University of Crete, Rethymno, Greece,* Wastewater Management for Developing Countries

Izrail S. Turoviskiy, *Izrail S. Turoviskiy,* Jacksonville, FloridaProcessing of Sludge, Biosolids, Wastewater Sludge

U.S. Environmental Protection Agency—Office of Wastewater Management, *What Wastewater Utilities Can Do Now to Guard Against Terrorist and Security Threats,*

U.S. Geological Survey, *Estimated Use of Water in The United States in 1990 Industrial Water Use, Source Water Assessment,*

Miguel A. Valenzuela, *Instituto Politecnico Nacional—ESIQIE. MEXICO,,* Wastewater Treatment Techniques—Advanced

G. van der Velde, *Radboud University Nijmegen, Nijmegen, The Netherlands,* Chlorine and Chlorine Residuals

D.V. Vayenas, *University of Ioannina, Agrinio, Greece,* Nitrification of Potable Water Using Trickling Filters

Raghuraman Venkatapathy, *Oak Ridge Institute for Science and Education, Cincinnati, Ohio,* Health Effects of Microbial Contaminants and Biotoxins in Drinking Water, Health Effects of Commonly Occurring Disinfection Byproducts in Municipal Water Supplies, Disinfectants

V.P. Venugopalan, *BARC Facilities, Kalpakkam, India,* Chlorine and Chlorine Residuals

Roger C. Viadero Jr., *West Virginia University, Morgantown, West Virginia,* Lime–Soda Ash Processes

Christian J. Volk, *Indiana-American Water, Richmond, Indiana,* Corrosion Control in Drinking Water Systems

Constantin Von Der Heyden, *School of Geography and the Environment, Oxford, United Kingdom,* Industrial Mine Use: Mine Waste

Nikolay Voutchkov, *Poseidon Resources Corporation, Stamford, Connecticut,* Desalination, Settling Tanks

Sutat Weesakul, *Asian Institute of Technology, Pathumthani, Thailand,* A Real-Time Hydrological Information System for Cities

Uruya Weesakul, *Thammasat University, Pathumthani, Thailand,* A Real-Time Hydrological Information System for Cities

Janice Weihe, *American Water, Quality Control and Research Laboratory, Belleville, Illinois,* Potential Risks of Waterborne Transmission of *Escherichia coli* O157:H7

June Weintraub, *City and County of San Francisco, Department of Public Health, San Francisco, California,* Disinfectants

Andy Whiteley, *Centre for Ecology and Hydrology, Oxford, United Kingdom,* Molecular Biology Tools for Monitoring Biodiversity in Wastewater Treatment Plants

Dan Wolz, *City of Wyoming, Michigan, Wyoming, Michigan,* Getting Our Clean Water Act Together

Don J. Wood, *University of Kentucky, Lexington, Kentucky,* Design of Water Distribution Systems

Grant Wright, *Heriot-Watt University, Edinburgh, Scotland, United Kingdom,* Roof Drainage Hydraulics

J. Michael Wright, *Harvard School of Public Health, Boston, Massachusetts,* Disinfectants

W. Zachritz II, *New Mexico State University, Las Cruces, New Mexico,* Land Applications of Wastewater in Arid Lands: Theory and Case Studies

B. Zani, *University of Fort Hare, Alice, South Africa,* Key Causes of Drinking Water Quality Failure in a Rural Small Water Supply of South Africa

Grietje Zeeman, *Wageningen University and Research Center, Wageningen, The Netherlands,* Anaerobic Sewage Treatment

Anastasios Zouboulis, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* Landfill Leachates: Part 2: Treatment, Landfill Leachates, Part I: Origin and Characterization, Solidification/Stabilization of Hazardous Solid Wastes, Municipal Storm Water Management, Combined Sewer Overflow Treatment, Technologies for Arsenic Removal from Contaminated Water Sources

CONTRIBUTORS

Absar Alum, *Arizona State University, Tempe, Arizona,* Water Quality Management in the U.S.: History of Water Regulation

Mohammad N. Almasri, *An-Najah National University, Nablus, Palestine,* Best Management Practices for Water Resources

Linda S. Andrews, *Mississippi State University, Biloxi, Mississippi,* Shellfish Growing Water Classification, Chlorine Residual

Hannah Aoyagi, *University of California, Irvine, California,* Cytochrome P450 Monooxygenase as an Indicator of PCB/Dioxin-Like Compounds in Fish

Robert Artinger, *TÜV Industrie Service GmbH, München, Germany,* Column Experiments in Saturated Porous Media Studying Contaminant Transport

Mukand Singh Babel, *Asian Institute of Technology, Pathumthani, Thailand,* Conservation of Water, Integrated Water Resources Management (IWRM)

Mark Bailey, *Centre for Ecology and Hydrology–Oxford, Oxford, United Kingdom,* Bioluminescent Biosensors for Toxicity Testing

Shimshon Balanson, *Cleveland State University, Cleveland, Ohio,* Macrophytes as Biomonitors of Trace Metals

Christine L. Bean, *University of New Hampshire, Durham, New Hampshire,* Protozoa in Water

Jennifer Bell, *Napier University, Edinburgh, United Kingdom,* Bioluminescent Biosensors for Toxicity Testing

Lieven Bervoets, *University of Antwerp, Antwerp, Belgium,* Active Biomonitoring (ABM) by Translocation of Bivalve Molluscs

J.M. Blasco, *Instituto de Ciencias Marinas de Andalucía, Cádiz, Spain,* A Weight of Evidence Approach to Characterize Sediment Quality Using Laboratory and Field Assays: An Example For Spanish Coasts

Ronny Blust, *University of Antwerp, Antwerp, Belgium,* Active Biomonitoring (ABM) by Translocation of Bivalve Molluscs

Marta Bryce, *CEPIS/PAHO, Delft, The Netherlands,* Flood of Portals on Water

Mario O. Buenfil-Rodriguez, *National University of Mexico, Cuernavaca, Morelos, Mexico,* Water Use Conservation and Efficiency

Jacques Buffle, *University of Geneva, Geneva, Switzerland,* Colloids and Dissolved Organics: Role in Membrane and Depth Filtration

Zia Bukhari, *American Water, Belleville, Illinois,* Understanding *Escherichia Coli* O157:H7 and the Need for Rapid Detection in Water

John Cairns, Jr., *Virginia Polytechnic Institute and State University, Blacksburg, Virginia,* Microscale Test Relationships to Responses to Toxicants in Natural Systems

Michael J. Carvan III, *University of Wisconsin–Milwaukee, Milwaukee, Wisconsin,* Genomic Technologies in Biomonitoring

M.C. Casado-Martínez, *Facultad de Ciencias del Mar y Ambientales, Cádiz, Spain,* A Weight of Evidence Approach to Characterize Sediment Quality Using Laboratory and Field Assays: An Example For Spanish Coasts

Tomás Ángel Del Valls Casillas, *Universidad de Cadiz, Cadiz, Spain,* Amphipod Sediment Toxicity Tests, Development and Application of Sediment Toxicity Test for Regulatory Purposes

Teresa A. Cassel, *University of California, Davis, California,* Remediation and Bioremediation of Selenium-Contaminated Waters

Augusto Cesar, *Universidad de Cadiz, Cadiz, Spain,* Amphipod Sediment Toxicity Tests

K.W. Chau, *The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong,* Water Quality Models: Mathematical Framework

Paulo Chaves, *Water Resources Research Center, Kyoto University, Japan,* Quality of Water in Storage

Shankar Chellam, *University of Houston, Houston, Texas,* Bromide Influence on Trihalomethane and Haloacetic Acid Formation

X. Chris Le, *University of Alberta, Edmonton, Alberta, Canada,* Arsenic Compounds in Water

Russell N. Clayshulte, *Aurora, Colorado,* Water Quality Management in an Urban Landscape

Gail E. Cordy, *U.S. Geological Survey,* A Primer on Water Quality

Rupali Datta, *University of Texas, San Antonio, Texas,* Lead and its Health Effects

Joanna Davies, *Syngenta, Bracknell, Berkshire, United Kingdom,* The Control of Algal Populations in Eutrophic Water Bodies

Maria B. Davoren, *Dublin Institute of Technology, Dublin, Ireland,* Luminescent Bacterial Biosensors for the Rapid Detection of Toxicants

T.A. Delvalls, *Facultad de Ciencias del Mar y Ambientales, Cádiz, Spain,* Biomarkers and Bioaccumulation: Two Lines of Evidence to Assess Sediment Quality, A Weight of Evidence Approach to Characterize Sediment Quality Using Laboratory and Field Assays: An Example For Spanish Coasts

Nicolina Dias, *Centro de Engenharia Biológica, Braga, Portugal,* Ciliated Protists as Test Organisms in Toxicity Assessment

Galina Dimitrieva-Moats, *University of Idaho, Moscow, Idaho,* Microbial Detection of Various Pollutants as an Early Warning System for Monitoring of Water Quality and Ecological Integrity of Natural Resources, in Russia

Halanaik Diwakara, *University of South Australia, Adelaide, Australia,* Water Markets in India: Economic and Institutional Aspects

Francis G. Doherty, *AquaTox Research, Inc., Syracuse, New York,* The Submitochondrial Particle Assay as a Biological Monitoring Tool

Antonia A. Donta, *University of Münster, Centre for Environmental Research, Münster, Germany,* Sustainable Water Management On Mediterranean Islands: Research and Education

Timothy J. Downs, *Clark University, Worcester, Massachusetts,* Field Sampling and Monitoring of Contaminants, State and Regional Water Supply, Water Resource Sustainability: Concepts and Practices

Hiep N. Duc, *Environment Protection Authority, NSW, Bankstown, New South Wales Australia,* Urban Water Resource and Management in Asia: Ho Chi Minh City

Suzanne Du Vall Knorr, *Ventura County Environmental Health Division, Ventura, California,* Regulatory and Security Requirements for Potable Water

Sandra Dunbar, *Napier University, Edinburgh, United Kingdom,* Bioluminescent Biosensors for Toxicity Testing

Diane Dupont, *Brock University, St. Catharines, Ontario, Canada,* Valuing Water Resources

Michael P. Dziewatkoski, *Mettler-Toledo Process Analytical, Woburn, Massachusetts,* pH

Energy Information Administration—Department of Energy, Hydropower—Energy from Moving Water

Environment Canada, Water—Here, There, and Everywhere in Canada, Water Conservation—Every Drop Counts in Canada

Environmental Protection Agency, Water Recycling and Reuse: The Environmental Benefits

M. Eric Benbow, *Michigan State University, East Lansing, Michigan,* Road Salt

Teresa W.-M. Fan, *University of Louisville, Louisville, Kentucky,* Remediation and Bioremediation of Selenium-Contaminated Waters

Federal Emergency Management Agency, Food and Water in an Emergency

Huan Feng, *Montclair State University, Montclair, New Jersey,* Classification and Environmental Quality Assessment in Aquatic Environments

N. Buceta Fernández, *Centro de Estudios de Puertos y Costas, Madrid, Spain,* A Weight of Evidence Approach to Characterize Sediment Quality Using Laboratory and Field Assays: An Example For Spanish Coasts

Peter D. Franzmann, *CSIRO Land and Water, Floreat, Australia,* Microbial Activities Management

Christian D. Frazar, *Silver Spring, Maryland,* Biodegradation

Rajiv Gandhi Chair, *Jawaharlal Nehru University, New Delhi, India,* Oil Pollution

Suduan Gao, *USDA–ARS, Parlier, California,* E_h

Horst Geckeis, *Institut für Nukleare Entsorgung, Karlsruhe, Germany,* Metal Ion Humic Colloid Interaction

Robert Gensemer, *Parametrix, Corvallis, Oregon,* Effluent Water Regulations in Arid Lands

Mário Abel Gonçalves, *Faculdade de Ciências da Universidade de Lisoba, Lisoba, Portugal,* Background Concentration of Pollutants

Neil S. Grigg, *Colorado State University, Fort Collins, Colorado,* Planning and Managing Water Infrastructure, Drought and Water Supply

Management, Drought Management Planning, Water Infrastructure and Systems, Water Resources Management

Håkan Håkanson, *University of Lund, Lund, Sweden,* Dishwashing Water Quality Properties

Carol J. Haley, *Virginia Water Resources Research Center,* Management of Water Resources for Drought Conditions

M.G.J Hartl, *Environmental Research Institute, University College Cork, Ireland,* Development and Application of Sediment Toxicity Tests for Regulatory Purposes

Roy C. Haught, *U.S. Environmental Protection Agency,* Water Supply and Water Resources: Distribution System Research

Joanne M. Hay, *Lincoln Ventures, Ltd., Lincoln, New Zealand,* Biochemical Oxygen Demand and Other Organic Pollution Measures

Richard M. Higashi, *University of California, Davis, California,* Remediation and Bioremediation of Selenium-Contaminated Waters

A.Y. Hoekstra, *UNESCO–IHE Institute for Water Education, Delft, The Netherlands,* Globalization of Water

Charles D.D. Howard, *Water Resources, Victoria, British Columbia, Canada,* River Basin Decisions Support Systems

Margaret S. Hrezo, *Radford University, Virginia,* Management of Water Resources for Drought Conditions

Enos C. Inniss, *University of Texas, San Antonio, Texas,* Perchloroethylene (PCE) Removal

James A. Jacobs, *Environmental Bio-Systems, Inc., Mill Valley, California,* Emerging and Recalcitrant Compounds in Groundwater

Chakresh K. Jain, *National Institute of Hydrology, Roorkee, India,* Water Quality Management, Trace Element Contamination in Groundwater of District Hardwar, Uttaranchal, India, Ground Water Quality in Areas Adjoining River Yamuna at Delhi, India, Irrigation Water Quality in Areas Adjoining River Yamuna At Delhi, India

Sanjay Kumar Jain, *National Institute of Hydrology, Roorkee, India,* Remote Sensing and GIS Application in Water Resources

Sharad K. Jain, *National Institute of Hydrology, Roorkee, Uttranchal, India,* Water Resources of India

H.A. Jenner, *KEMA Power Generation and Sustainables, Arnhem, The Netherlands,* Dose-Response of Mussels to Chlorine

Y. Jiang, *Hong Kong Baptist University, Kowloon, Hong Kong,* Algal Toxins in Water

B. Ji, *Hong Kong Baptist University, Kowloon, Hong Kong,* Algal Toxins in Water

N. Jiménez-Tenorio, *Facultad de Ciencias del Mar y Ambientales, Cádiz, Spain,* Biomarkers and Bioaccumulation: Two Lines of Evidence to Assess Sediment Quality

Zhen-Gang Ji, *Minerals Management Service, Herndon, Virginia,* Water Quality Modeling—Case Studies, Water Quality Models: Chemical Principles

Erik Johansson, *GS Development AB, Malmö, Sweden,* Dishwashing Water Quality Properties

B. Thomas Johnson, *USGS—Columbia Environmental Research Center, Columbia, Missouri,* Monitoring Lipophilic Contaminants in the Aquatic Environment using the SPMD-TOX Paradigm

Anne Jones-Lee, *G. Fred Lee & Associates, El Macero, California,* Water Quality Aspects of Dredged Sediment Management, Municipal Solid Waste Landfills—Water Quality Issues

Dick de Jong, *IRC International Water and Sanitation Centre, Delft, The Netherlands,* Flood of Portals on Water

Jagath J. Kaluarachchi, *Utah State University, Logan, Utah,* Best Management Practices for Water Resources

Atya Kapley, *National Environmental Engineering Research Institute, CSIR, Nehru Marg, Nagpur, India, Salmonella*: Monitoring and Detection in Drinking Water

I. Katsoyiannis, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* Arsenic Health Effects

Absar A. Kazmi, *Nishihara Environment Technology, Tokyo, Japan,* Activated Carbon—Powdered, Chlorination

Keith O. Keplinger, *Texas Institute for Applied Environmental Research, Stephenville, Texas,* The Economics of Water Quality

Kusum W. Ketkar, *Jawaharlal Nehru University, New Delhi, India,* Oil Pollution

Ganesh B. Keremane, *University of South Australia, Adelaide, Australia,* Harvesting Rainwater

Rebecca D. Klaper, *University of Wisconsin–Milwaukee, Milwaukee, Wisconsin,* Genomic Technologies in Biomonitoring

Toshiharu Kojiri, *Water Resources Research Center, Kyoto University, Japan,* Quality of Water in Storage

Ken'ichirou Kosugi, *Kyoto University, Kyoto, Japan,* Lysimeter Soil Water Sampling

Manfred A. Lange, *University of Münster, Centre for Environmental Research, Münster, Germany,* Sustainable Water Management On Mediterranean Islands: Research and Education

Frédéric Lasserre, *Université Laval, Ste-Foy, Québec, Canada,* Water Use in the United States

N.K. Lazaridis, *Aristotle University, Thessaloniki, Greece,* Sorptive Filtration

Jamie R. Lead, *University of Birmingham, Birmingham, United Kingdom,* Trace Metal Speciation

G. Fred Lee, *G. Fred Lee & Associates, El Macero, California,* Water Quality Aspects of Dredged Sediment Management, Municipal Solid Waste Landfills—Water Quality Issues

Terence R. Lee, *Santiago, Chile,* Water Markets: Transaction Costs and Institutional Options, The Provision of Drinking Water and Sanitation in Developing Countries, Spot Prices, Option Prices, and Water Markets, Meeting Water Needs in Developing Countries with Tradable Rights

Markku J. Lehtola, *National Public Health Institute, Kuopio, Finland,* Microbiological Quality Control in Distribution Systems

Gary G. Leppard, *National Water Research Institute, Burlington, Ontario, Canada,* Colloids and Dissolved Organics: Role in Membrane and Depth Filtration

Mark LeChevallier, *American Water, Voorhees, New Jersey,* Understanding *Escherichia Coli* O157:H7 and the Need for Rapid Detection in Water

Nelson Lima, *Centro de Engenharia Biológica, Braga, Portugal,* Ciliated Protists as Test Organisms in Toxicity Assessment

Maria Giulia Lionetto, *Università di Lecce, Lecce, Italy,* Metallothioneins as Indicators of Trace Metal Pollution

Jody W. Lipford, *PERC, Bozeman, Montana, and Presbyterian College, Clinton, South Carolina,* Averting Water Disputes

Baikun Li, *Pennsylvania State University, Harrisburg, Pennsylvania,* Iron Bacteria, Microbial Dynamics of Biofilms, Microbial Forms in Biofouling Events

Rongchao Li, *Delft University of Technology, Delft, The Netherlands,* Transboundary Water Conflicts in the Nile Basin, Institutional Aspects of Water Management in China, Flood Control History in the Netherlands

Bryan Lohmar, *Economic Research Service, U.S. Department of Agriculture,* Will Water Scarcity Limit China's Agricultural Potential?

Inmaculada Riba López, *Universidad de Cadiz, Cadiz, Spain,* Amphipod Sediment Toxicity Tests

M.X. Loukidou, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* Biosorption of Toxic Metals

Scott A. Lowe, *Manhattan College, Riverdale, New York,* Eutrophication and Organic Loading

G. Lyberatos, *University of Ioannina, Agrinio, Greece,* Cartridge Filters for Iron Removal

Kenneth M. Mackenthun, *Arlington, Virginia,* Water Quality

Tarun K. Mal, *Cleveland State University, Cleveland, Ohio,* Macrophytes as Biomonitors of Trace Metals

Philip J. Markle, *Whittier, California,* Toxicity Identification Evaluation, Whole Effluent Toxicity Controls

James T. Markweise, *Neptune and Company, Inc., Los Alamos, New Mexico,* Assessment of Ecological Effects in Water-Limited Environments, Effluent Water Regulations in Arid Lands

Pertti J. Martikainen, *University of Kuopio, Kuopio, Finland,* Microbiological Quality Control in Distribution Systems

M.L. Martín-Díaz, *Instituto de Ciencias Marinas de Andalucía, Cádiz, Spain,* A Weight of Evidence Approach to Characterize Sediment Quality Using Laboratory and Field Assays: An Example For Spanish Coasts, Biomarkers and Bioaccumulation: Two Lines of Evidence to Assess Sediment Quality

Maria del Carmen Casado Martínez, *Universidad de Cadiz, Cadiz, Spain,* Amphipod Sediment Toxicity Tests

K.A. Matis, *Aristotle University, Thessaloniki, Greece,* Sorptive Filtration

Lindsay Renick Mayer, *Goddard Space Flight Center, Greenbelt, Maryland,* NASA Helping to Understand Water Flow in the West

Mark C. Meckes, *U.S. Environmental Protection Agency,* Water Supply and Water Resources: Distribution System Research

Richard W. Merritt, *Michigan State University, East Lansing, Michigan,* Road Salt

Richard Meyerhoff, *CDM, Denver, Colorado,* Effluent Water Regulations in Arid Lands

J. Michael Wright, *Harvard School of Public Health, Boston, Massachusetts,* Chlorination By-Products

Cornelis J.H. Miermans, *Institute for Inland Water Management and Waste Water Treatment–RIZA, Lelystad, The Netherlands,* SOFIE: An Optimized Approach for Exposure Tests and Sediment Assays

Ilkka T. Miettinen, *National Public Health Institute, Kuopio, Finland,* Microbiological Quality Control in Distribution Systems

Dusan P. Miskovic, *Northwood University, West Palm Beach, Florida,* Oil-Field Brine

Diana Mitsova-Boneva, *University of Cincinnati, Cincinnati, Ohio,* Quality of Water Supplies

Tom Mohr, *Santa Clara Valley Water District, San Jose, California,* Emerging and Recalcitrant Compounds in Groundwater

M.C. Morales-Caselles, *Facultad de Ciencias del Mar y Ambientales, Cádiz, Spain,* Biomarkers and Bioaccumulation: Two Lines of Evidence to Assess Sediment Quality

National Drought Mitigation Center, Drought in the Dust Bowl Years

National Water-Quality Assessment (NAWQA) Program—U.S. Geological Survey, Source-Water Protection

Jennifer Nelson, *The Groundwater Foundation, Lincoln, Nebraska,* Reaching Out: Public Education and Community Involvement in Groundwater Protection

Anne Ng, *Swinburne University of Technology, Hawthorne, Victoria, Australia,* River Water Quality Calibration, Review of River Water Quality Modeling Software Tools

Jacques Nicolas, *University of Liege, Arlon, Belgium,* Interest in the Use of an Electronic Nose for Field Monitoring of Odors in the Environment

Ana Nicolau, *Centro de Engenharia Biológica, Braga, Portugal,* Ciliated Protists as Test Organisms in Toxicity Assessment

Diana J. Oakes, *University of Sydney, Lidcombe, Australia,* Environmental Applications with Submitochondrial Particles

Oladele A. Ogunseitan, *University of California, Irvine, California,* Microbial Enzyme Assays for Detecting Heavy Metal Toxicity, Cytochrome P450 Monooxygenase as an Indicator of PCB/Dioxin-Like Compounds in Fish

J. O'Halloran, *Environmental Research Institute, University College Cork, Ireland,* Development and Application of Sediment Toxicity Tests for Regulatory Purposes

Victor Onwueme, *Montclair State University, Montclair, New Jersey,* Classification and Environmental Quality Assessment in Aquatic Environments

Alper Ozkan, *Selcuk University, Konya, Turkey,* Coagulation and Flocculation in Practice

Neil F. Pasco, *Lincoln Ventures, Ltd., Lincoln, New Zealand,* Biochemical Oxygen Demand and Other Organic Pollution Measures

B.J.C. Perera, *Swinburne University of Technology, Hawthorne, Victoria, Australia,* River Water Quality Calibration, Review of River Water Quality Modeling Software Tools

Jim Philip, *Napier University, Edinburgh, United Kingdom,* Bioluminescent Biosensors for Toxicity Testing

Laurel Phoenix, *Green Bay, Wisconsin,* Source Water Quality Management, Water Managed in the Public Trust

Randy T. Piper, *Dillon, Montana,* Overview and Trends in the International Water Market

John K. Pollak, *University of Sydney, Lidcombe, Australia,* Environmental Applications with Submitochondrial Particles

Dörte Poszig, *University of Münster, Centre for Environmental Research, Münster, Germany,* Sustainable Water Management On Mediterranean Islands: Research and Education

Hemant J. Purohit, *National Environmental Engineering Research Institute, CSIR, Nehru Marg, Nagpur, India, Salmonella*: Monitoring and Detection in Drinking Water

Shahida Quazi, *University of Texas, San Antonio, Texas,* Lead and its Health Effects

S. Rajagopal, *Radboud University Nijmegen, Toernooiveld, Nijmegen, The Netherlands,* Dose-Response of Mussels to Chlorine

Krishna Ramanujan, *Goddard Space Flight Center, Greenbelt, Maryland,* NASA Helping to Understand Water Flow in the West

Lucas Reijnders, *University of Amsterdam, Amsterdam, The Netherlands,* Sustainable Management of Natural Resources

Steven J. Renzetti, *Brock University, St. Catharines, Ontario, Canada,* Water Demand Forecasting, Water Pricing, Valuing Water Resources

Martin Reuss, *Office of History Headquarters U.S. Army Corps of Engineers,* The Development of American Water Resources: Planners, Politicians, and Constitutional Interpretation, The Expansion of Federal Water Projects

I. Riba, *Facultad de Ciencias del Mar y Ambientales, Cádiz, Spain,* Biomarkers and Bioaccumulation: Two Lines of Evidence to Assess Sediment Quality, A Weight of Evidence Approach to Characterize Sediment Quality Using Laboratory and Field Assays: An Example For Spanish Coasts

Matthew L. Rise, *University of Wisconsin–Milwaukee, Milwaukee, Wisconsin,* Genomic Technologies in Biomonitoring

Arthur W. Rose, *Pennsylvania State University, University Park, Pennsylvania,* Acid Mine Drainage—Extent and Character, Passive Treatment of Acid Mine Drainage (Wetlands)

Barry H. Rosen, *US Fish & Wildlife Service, Vero Beach, Florida,* Waterborne Bacteria

Serge Rotteveel, *Institute for Inland Water Management and Waste Water Treatment–RIZA, Lelystad, The Netherlands,* SOFIE: An Optimized Approach for Exposure Tests and Sediment Assays

Timothy J. Ryan, *Ohio University, Athens, Ohio,* Water Sampling and Laboratory Safety

Randall T. Ryti, *Neptune and Company, Inc., Los Alamos, New Mexico,* Assessment of Ecological Effects in Water-Limited Environments

Masaki Sagehashi, *University of Tokyo, Tokyo, Japan,* Biomanipulation

Basu Saha, *Loughborough University, Loughborough, United Kingdom,* Activated Carbon: Ion Exchange and Adsorption Properties

Md. Salequzzaman, *Khulna University, Khulna, Bangladesh,* Ecoregions: A Spatial Framework for Environmental Management

Dibyendu Sarkar, *University of Texas, San Antonio, Texas,* Lead and its Health Effects

Peter M. Scarlett, *Winfrith Technology Centre, Dorchester, Dorset, United Kingdom,* The Control of Algal Populations in Eutrophic Water Bodies

Trifone Schettino, *Università di Lecce, Lecce, Italy,* Metallothioneins as Indicators of Trace Metal Pollution

Lewis Schneider, *North Jersey District Water Supply Commission, Wanaque, New Jersey,* Classification and Environmental Quality Assessment in Aquatic Environments

Wolfram Schuessler, *Institut für Nukleare Entsorgung, Karlsruhe, Germany,* Column Experiments in Saturated Porous Media Studying Contaminant Transport

K.D. Sharma, *National Institute of Hydrology, Roorkee, India,* Water Quality Management

Mukesh K. Sharma, *National Institute of Hydrology, Roorkee, India,* Ground Water Quality in Areas Adjoining River Yamuna at Delhi, India, Irrigation Water Quality in Areas Adjoining River Yamuna At Delhi, India

Daniel Shindler, *UMDNJ, New Brunswick, New Jersey,* Methemoglobinemia

Slobodan P. Simonovic, *The University of Western Ontario, London, Ontario, Canada,* Water Resources Systems Analysis, Fuzzy Criteria for Water Resources Systems Performance Evaluation, Participatory Multicriteria Flood Management

Shahnawaz Sinha, *Malcolm Pirnie Inc., Phoenix, Arizona,* Disinfection By-Product Precursor Removal from Natural Waters

Joseph P. Skorupa, *U.S. Fish and Wildlife Service,* Remediation and Bioremediation of Selenium-Contaminated Waters

Roel Smolders, *University of Antwerp, Antwerp, Belgium,* Active Biomonitoring (ABM) by Translocation of Bivalve Molluscs

Jinsik Sohn, *Kookmin University, Seoul, Korea,* Disinfection By-Product Precursor Removal from Natural Waters

Fiona Stainsby, *Napier University, Edinburgh, United Kingdom,* Bioluminescent Biosensors for Toxicity Testing

Ross A. Steenson, *Geomatrix, Oakland, California,* Land Use Effects on Water Quality

Leonard I. Sweet, *Engineering Labs Inc., Canton, Michigan,* Application of the Precautionary Principle to Water Science

Kenneth K. Tanji, *University of California, Davis, California,* E_h

Ralph J. Tella, *Lord Associates, Inc., Norwood, Massachusetts,* Overview of Analytical Methods of Water Analyses With Specific Reference to EPA Methods for Priority Pollutant Analysis

William E. Templin, *U.S. Geological Survey, Sacramento, California,* California—Continually the Nation's Leader in Water Use

Rita Triebskorn, *Steinbeis-Transfer Center for Ecotoxicology and Ecophysiology, Rottenburg, Germany,* Biomarkers, Bioindicators, and the Trondheim Biomonitoring System

Nirit Ulitzur, *Checklight Ltd., Tivon, Israel,* Use of Luminescent Bacteria and the *Lux* Genes For Determination of Water Quality

Shimon Ulitzur, *Technion Institute of Technology, Haifa, Israel,* Use of Luminescent Bacteria and the *Lux* Genes For Determination of Water Quality

U.S. Environmental Protection Agency, How We Use Water in These United States

U.S. Agency for International Development (USAID), Water and Coastal Resources

U.S. Geological Survey, Water Quality, Water Science Glossary of Terms

G. van der velde, *Radboud University Nijmegen, Toernooiveld, Nijmegen, The Netherlands,* Dose-Response of Mussels to Chlorine

F.N.A.M. van pelt, *Environmental Research Institute, University College Cork, Ireland,* Development and Application of Sediment Toxicity Tests for Regulatory Purposes

D.V. Vayenas, *University of Ioannina, Agrinio, Greece,* Cartridge Filters for Iron Removal

Raghuraman Venkatapathy, *Oak Ridge Institute for Science and Education, Cincinnati, Ohio,* Alternative Disinfection Practices and Future Directions for Disinfection By-Product Minimization, Chlorination By-Products

V.P. Venugopalan, *BARC Facilities, Kalpakkam, India,* Dose-Response of Mussels to Chlorine

Jos P.M. Vink, *Institute for Inland Water Management and Waste Water Treatment–RIZA, Lelystad, The Netherlands,* Heavy Metal Uptake Rates Among Sediment Dwelling Organisms, SOFIE: An Optimized Approach for Exposure Tests and Sediment Assays

Judith Voets, *University of Antwerp, Antwerp, Belgium,* Active Biomonitoring (ABM) by Translocation of Bivalve Molluscs

Mark J. Walker, *University of Nevada, Reno, Nevada,* Water Related Diseases

William R. Walker, *Virginia Water Resources Research Center,* Management of Water Resources for Drought Conditions

Xinhao Wang, *University of Cincinnati, Cincinnati, Ohio,* Quality of Water Supplies

Corinna Watt, *University of Alberta, Edmonton, Alberta, Canada,* Arsenic Compounds in Water

Janice Weihe, *American Water, Belleville, Illinois,* Understanding *Escherichia Coli* O157:H7 and the Need for Rapid Detection in Water

June M. Weintraub, *City and County of San Francisco Department of Public Health, San Francisco, California,* Chlorination By-Products, Alternative Disinfection Practices and Future Directions for Disinfection By-Product Minimization

Victor Wepener, *Rand Afrikaans University, Auckland Park, South Africa,* Active Biomonitoring (ABM) by Translocation of Bivalve Molluscs

Eva Ståhl Wernersson, *GS Development AB, Malmö, Sweden,* Dishwashing Water Quality Properties

Andrew Whiteley, *Centre for Ecology and Hydrology–Oxford, Oxford, United Kingdom,* Bioluminescent Biosensors for Toxicity Testing

Siouxsie Wiles, *Imperial College London, London, United Kingdom,* Bioluminescent Biosensors for Toxicity Testing

Thomas M. Williams, *Baruch Institute of Coastal Ecology and Forest Science, Georgetown, South Carolina,* Water Quality Management in a Forested Landscape

Parley V. Winger, *University of Georgia, Atlanta, Georgia,* Water Assessment and Criteria

M.H. Wong, *Hong Kong Baptist University, Kowloon, Hong Kong,* Algal Toxins in Water

R.N.S. Wong, *Hong Kong Baptist University, Kowloon, Hong Kong,* Algal Toxins in Water

J. Michael Wright, *Harvard School of Public Health, Boston, Massachusetts,* Alternative Disinfection Practices and Future Directions for Disinfection By-Product Minimization

Gary P. Yakub, *Kathleen Stadterman-Knauer Allegheny County Sanitary Authority, Pittsburgh, Pennsylvania,* Indicator Organisms

Yeomin Yoon, *Northwestern University, Evanston, Illinois,* Disinfection By-Product Precursor Removal from Natural Waters

M.E. Young, *Conwy, United Kingdom,* Water Resources Challenges in the Arab World

Mehmet Ali Yurdusev, *Celal Bayar University, Manisa, Turkey,* Integration of Environmental Impacts into Water Resources Planning

Karl Erik Zachariassen, *Norwegian University of Science and Technology, Trondheim, Norway,* Physiological Biomarkers and the Trondheim Biomonitoring System

Luke R. Zappia, *CSIRO Land and Water, Floreat, Australia,* Microbial Activities Management

Harry X. Zhang, *Parsons Corporation, Fairfax, Virginia,* Water Quality Management and Nonpoint Source Control, Water Quality Models for Developing Soil Management Practices

Igor S. Zonn, Lessons from the Rising Caspian

A.I. Zouboulis, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* Biosorption of Toxic Metals, Arsenic Health Effects

CONTRIBUTORS

Julia E. Aizpún de Moreno, *Universidad Nacional de Mar del Plata, Mar del Plata, Argentina,* Occurrence of Organochlorine Pesticides in Vegetables Grown on Untreated Soils from an Agricultural Watershed, Pesticide Chemistry in the Environment

Mahbub Alam, *Kansas State University, Garden City, Kansas,* Vacuum Gauge Tensiometer

Absar Alum, *Arizona State University, Tempe, Arizona,* Microbial Quality of Reclaimed Irrigation: International Perspective

Peyman Daneshkar Arasteh, *Soil Conservation and Watershed Management Research Institute (SCWMRI), Tehran, Iran,* Minimum Environmental Flow Regimes, Large Area Surface Energy Balance Estimation Using Satellite Imagery, Water Spreading

Muhammad Nadeem Asghar, *International Water Management Institute (IWMI), Lahore, Pakistan,* Irrigation Wells, Skimmed Groundwater, Sprinkler Irrigation

Mukand Singh Babel, *Asian Institute of Technology, Pathumthani, Thailand,* Soil Moisture Measurement—Neutron

Frank Balon, *Buffalo District, U.S. Army Corps of Engineers,* Corps Turned Niagara Falls Off, On Again

Joseph D. Bankston, *Louisiana State University, Baton Rouge, Louisiana,* Pumping Stations

Nathalie Barrette, *Laval University, Québec, Canada,* Greenhouse Gas Emissions From Hydroelectric Reservoirs

Luis Berga, *ETSIn Caminos, Barcelona, Spain,* Floods as a Natural Hazard

Asbjørn Bergheim, *RF-Rogaland Research, Stavanger, Norway,* Water Pollution From Fish Farms

K.K.S. Bhatia, *National Institute of Hydrology, Roorkee, Uttaranchal, India,* Dilution-Mixing Zones and Design Flows, Surface Water Pollution, Assessment of Pollution Outflow From Large Agricultural Areas

Sandra Bird, *U.S. Environmental Protection Agency,* Impervious Cover—Paving Paradise

Peter E. Black, *State University of New York College of Environmental Science and Forestry, Syracuse, New York,* Watershed Hydrology

Robert W. Black, *National Water Quality Assessment Program, U.S. Geological Survey,* Organic Compounds and Trace Elements in Freshwater Streambed Sediment and Fish from the Puget Sound Basin

Alexander Brenning, *Humboldt–Universität zu Berlin, Berlin, Germany,* Rock Glacier

Karen D. Brettschneider, *Houston-Galveston Area Council, Houston, Texas,* Urban Water Studies, Watershed

Emera Bridger, *SUNY-ESF, Syracuse, New York,* Forests and Wetlands

Alexander Brinker, *Fischereiforschungsstelle des Landes Baden-Württemberg, Langenargen, Germany,* Water Pollution From Fish Farms

Gary A. Buchanan, *Division of Science, Research, and Technology, Trenton, New Jersey,* Fish Consumption Advisories

Yong Cai, *Florida International University, Miami, Florida,* Metal Tolerance in Plants: The Roles of Thiol-Containing Peptides

René Canuel, *University of Québec in Montréal, Montréal, Canada,* Hydroelectric Reservoirs as Anthropogenic Sources of Greenhouse Gases

Virginia Carter, *U.S. Department of the Interior,* Classification of Wetlands and Deepwater Habitats of the United States

Harenda Singh Chauhan, *G.B. Pant University of Agriculture and Technology, Pantnagar, India,* Microirrigation: An Approach to Efficient Irrigation

Shulin Chen, *Washington State University, Pullman, Washington,* Aquaculture Technology for Producers

Pietro Chiavaccini, *Universita di Pisa, Pisa, Italy,* Flood Control Structures, Hydraulics

Brent C. Christopher, *Montana State University, Bozeman, Montana,* Subglacial Lake Vostok

Xuefeng Chu, *Grand Valley State University, Muskegon, Michigan,* Pesticide Occurrence and Distribution in Relation to Use

Michelle Clarke, *Cranfield University, North Wyke, Devon, United Kingdom,* Water Quality Management in an Agricultural Landscape, Soil Erosion and Control Practices

Thomas R. Clarke, *USDA ARS U.S. Water Conservation Laboratory, Phoenix, Arizona,* Crop Water Stress Detection Using Remote Sensing

B.D. Clinton, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Measuring and Modeling Tree and Stand Level Transpiration

Steve Colman, *U.S. Geological Survey,* Lake Baikal—A Touchstone for Global Change and Rift Studies

Charles M. Cooper, *USDA Agricultural Research Service National Sedimentation Laboratory, Oxford, Mississippi,* Drainage Ditches

Dennis L. Corwin, *Salinity Laboratory, Riverside, California,* Soil Salinity

Lewis M. Cowardin, *U.S. Department of the Interior,* Classification of Wetlands and Deepwater Habitats of the United States

Christophe Cudennec, *Ecole Nationale Supérieure Agronomique, Rennes, France,* Unit Hydrograph, Rivers and Streams: One-Way Flow System

Rupali Datta, *University of Texas, San Antonio, Texas,* Phytoremediation By Constructed Wetlands, Soil N Management Impact on The Quality of Surface and Subsurface Water

Francesca Dellacasa, *Università di Pisa, Pisa, Italy,* River Basins

Richard Dowling, *Pittsburgh District, U.S. Army Corps of Engineers,* Innovative Pens Hatch Thousands of Trout

Éric Duchemin, *DREX Environment, Montréal, Canada,* Hydroelectric Reservoirs as Anthropogenic Sources of Greenhouse Gases

Julia Duzant, *Cranfield University, North Wyke, Devon, United Kingdom,* Water Quality Management in an Agricultural Landscape, Soil Erosion and Control Practices

J. Gordon Edwards, *San Jose, California,* Effects of DDT in Surface Water on Bird Abundance and Reproduction—A History

Environment Canada, Water—The Canadian Transporter

K.J. Elliott, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Measuring and Modeling Tree and Stand Level Transpiration

Theodore A. Endreny, *SUNY-ESF, Syracuse, New York,* Forests and Wetlands, Riparian Systems, Great Lakes

Wayne D. Erskine, *State Forests of New South Wales, Beecroft, New South Wales, Australia,* Stream Classification, Gully Erosion, Sediment Load Measurements

Xing Fang, *Lamar University, Beaumont, Texas,* Culvert Design, Water Turbine, Hydraulics of Pressurized Flow, Open Channel Design, Storage and Detention Facilities, Streamflow

Jerry L. Farris, *Arkansas State University, State University, Arkansas,* Drainage Ditches

Douglas H. Fender, *International Turf Producers Foundation, Meadows, Illinois,* Landscape Water-Conservation Techniques

R. Andrés Ferreyra, *Ag Connections, Inc., Murray, Kentucky,* Stomates

Markus Flury, *Washington State University, Pullman, Washington,* Dyes As Hydrological Tracers

Julia Freedgood, *American Farmland Trust, Washington, District of Columbia,* Agriculture and Land Use Planning

Annette Geller, *UFZ Center for Environmental Research Leipzig-Halle Ltd., Magdeburg, Germany,* Limnology

Walter Geller, *UFZ Center for Environmental Research, Magdeburg, Germany,* Limnology, Lakes

Walter H. Geller, *UFZ, Dept. of Inland Water Research, Magdeburg, Germany,* Acidification of Freshwater Resources

Robert J. Gilliom, *U.S. Geological Survey,* Classification and Mapping of Agricultural Land for National Water-Quality Assessment

Francis C. Golet, *U.S. Department of the Interior,* Classification of Wetlands and Deepwater Habitats of the United States

Mariana Gonzalez, *Universidad Nacional de Mar del Plata, Mar del Plata, Argentina and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina,* Pesticide Chemistry in the Environment, Occurrence of Organochlorine Pesticides in Vegetables Grown on Untreated Soils from an Agricultural Watershed

Eve Gruntfest, *(from The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts, Wiley 2003),* Floods

Suresh K. Gupta, *CSSRI, Karnal, Haryana, India,* Salt Tolerance, Waterlogging

John Hall, *New Orleans District, U.S. Army Corps of Engineers,* Lasers Scan Levees from the Air

Dorota Z. Haman, *University of Florida, Gainesville, Florida,* Screen Filters for Microirrigation, Media Filters for Microirrigation, Microirrigation

Blaine Hanson, *LAWR, Davis, California,* Deep-Well Turbine Pumps

David R. Hargis, *Hargis + Associates, Inc., Tucson, Arizona,* Forensic Hydrogeology

Thomas Harter, *University of California, Davis, California,* Animal Farming Operations: Groundwater Quality Issues

G.J. Harvey, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Measuring and Modeling Tree and Stand Level Transpiration

Karl E. Havens, *South Florida Water Mgmt. District, West Palm Beach, Florida,* Submerged Aquatic Plants Affect Water Quality in Lakes

Peder Hjorth, *Lund University, Lund, Sweden,* River Basin Planning and Coordination

Joseph Holden, *University of Leeds, Leeds, United Kingdom,* Surface Runoff and Subsurface Drainage

Robert M. Hordon, *Rutgers University, Piscataway, New Jersey,* Flow-duration Curves

Stéphane Houel, *University of Québec in Montréal, Montréal, Canada,* Hydroelectric Reservoirs as Anthropogenic Sources of Greenhouse Gases

Xiangjiang Huang, *Colorado State University, Fort Collins, Colorado,* Drainage Networks

Fen C. Hunt, *U.S. Department of Agriculture, Washington, District of Columbia,* Agriculture and Land Use Planning

Stephen W. Hussey, *Dabane Trust, Burnside, Bulawayo, Zimbabwe,* Water from Saturated River Sediment—Sand Abstraction

Deborah Hutchinson, *U.S. Geological Survey,* Lake Baikal—A Touchstone for Global Change and Rift Studies

James A. Jacobs, *Environmental Bio-Systems, Inc., Mill Valley, California,* Groundwater Assessment Using Soil Sampling Techniques

Chakresh K. Jain, *National Institute of Hydrology, Roorkee, India,* Adsorption of Metal Ions On Bed Sediments

Sharad K. Jain, *National Institute of Hydrology, Roorkee, Uttranchal, India,* Hydroelectric Power, Reservoir Sedimentation, Reservoirs-Multipurpose, Base Flow, Ganga River, India

Steven Jennings, *(from The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts, Wiley 2003),* Floods

Ramakar Jha, *National Institute of Hydrology, Roorkee, Uttranchal, India,* Hydrological Processes and Measured Pollutant Loads, Dilution-Mixing Zones and Design Flows, Surface Water Pollution, Assessment of Pollution Outflow From Large Agricultural Areas

Klaus Jöhnk, *University of Amsterdam, Amsterdam, The Netherlands,* Heat Balance of Open Waterbodies

Bill Johnson, *Washington State University, Pullman, Washington,* Aquaculture Technology for Producers

Brit Johnson, *Washington State University, Pullman, Washington,* Aquaculture Technology for Producers

Anne Jones-Lee, *G. Fred Lee & Associates, El Macero, California,* Eutrophication (Excessive Fertilization), Unrecognized Pollutants, Urban Stormwater Runoff Water Quality Issues

Pierre Y. Julien, *Colorado State University, Fort Collins, Colorado,* Sedimentation, Rivers

Alpana Khairom, *University of Texas, San Antonio, Texas,* Phytoremediation By Constructed Wetlands

Jacob W. Kijne, *Herts, United Kingdom,* World's Major Irrigation Areas, Maintaining Salt Balance on Irrigated Land

Charles W. Kirby, *Paciulli, Simmons & Assoc., Fairfax, Virginia,* Water Quality in Suburban Watersheds

Matthias Koschorreck, *Center of Environmental Research (UFZ), Magdeburg, Germany,* Microorganisms in Their Natural Environment

Demetris Koutsoyiannis, *National Technical University, Athens, Greece,* Stochastic Simulation of Hydrosystems, Reliability Concepts in Reservoir Design, Hydrologic Persistence and The Hurst Phenomenon

V. Kulik, *CYPUM PTY LTD, Canberra City, Australia,* Combustible Watersheds

Richard Lanyon, *Metropolitan Water Reclamation District of Greater Chicago, Chicago, Illinois,* Reversal of the Chicago River

Edward T. Laroe, *U.S. Department of the Interior,* Classification of Wetlands and Deepwater Habitats of the United States

Catherine Larose, *University of Québec in Montréal, Montréal, Canada,* Hydroelectric Reservoirs as Anthropogenic Sources of Greenhouse Gases

Frederic Lasserre, *Laval University, Quebec City, Canada,* The Aral Sea Disaster: Environment Issues and Nationalist Tensions, Irrigation in The United States

Jamie R. Lead, *University of Birmingham, Birmingham, United Kingdom,* Freshwater Colloids

G. Fred Lee, *G. Fred Lee & Associates, El Macero, California,* Eutrophication (Excessive Fertilization), Unrecognized Pollutants, Urban Stormwater Runoff Water Quality Issues

Leo S. Leonhart, *Hargis + Associates, Inc., Tucson, Arizona,* Forensic Hydrogeology, Cienega

Xu Liang, *University of California, Berkeley, California,* Land Surface Modeling

Srinivasa Lingireddy, *University of Kentucky, Lexington, Kentucky,* Calibration of Hydraulic Network Models

Zhi-Qing Lin, *Southern Illinois University at Edwardsville, Edwardsville, Illinois,* Bioaccumulation

Rongchao Li, *Delft University of Technology, Delft, The Netherlands,* Flood Control in the Yellow River Basin in China

G.V. Loganathan, *Virginia Polytechnic Institute and State University, Blacksburg, Virginia,* Instream Flow Methods

Marc Lucotte, *University of Québec in Montréal, Montréal, Canada,* Hydroelectric Reservoirs as Anthropogenic Sources of Greenhouse Gases

Dorene E. Maccoy, *National Water Quality Assessment Program, U.S. Geological Survey,* Organic Compounds and Trace Elements in Freshwater Streambed Sediment and Fish from the Puget Sound Basin

Chandra Madramootoo, *Macdonald Campus of McGill University, Ste-Anne de Bellevue, Quebec, Canada,* Tile Drainage: Impacts, Plant Growth, and Water Table Levels, Water Logging: Topographic and Agricultural Impacts, Water Table Contribution to Crop Evapotranspiration

Ole Mark, *Asian Institute of Technology, Hørsholm, Denmark,* Modeling of Urban Drainage and Stormwater, Modeling of Water Quality in Sewers, Urban Flooding

Brane Maticic, *Ljubljana, Slovenia,* Nitrate Pollution Prevention

Matthew P. McCartney, *International Water Management Institute, Pretoria, South Africa,* Wetlands: Uses, Functions, and Values

Marianne McHugh, *Cranfield University, Silsoe, United Kingdom,* Water Quality Management in an Agricultural Landscape, Soil Erosion and Control Practices, Soil Conservation

T.C. Mcintyre, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Plant and Microorganism Selection for Phytoremediation of Hydrocarbons and Metals

Karina S.B. Miglioranza, *Universidad Nacional de Mar del Plata, Mar del Plata, Argentina and Consejo Nacional de Investigaciones Cientťficas y Técnicas (CONICET), Buenos Aires, Argentina,* Pesticide Chemistry in the Environment, Occurrence of Organochlorine Pesticides in Vegetables Grown on Untreated Soils from an Agricultural Watershed

Matjaz Mikos, *University of Ljubljana, Ljubljana, Slovenia,* Sediment Transport

Myron J. Mitchell, *State University of New York, Syracuse, New York,* Episodic Acidification

Jarai Mon, *Washington State University, Pullman, Washington,* Dyes As Hydrological Tracers

Matthew T. Moore, *USDA Agricultural Research Service National Sedimentation Laboratory, Oxford, Mississippi,* Drainage Ditches

M. Susan Moran, *USDA ARS Southwest Watershed Research Center, Tucson, Arizona,* Crop Water Stress Detection Using Remote Sensing

Víctor J. Moreno, *Universidad Nacional de Mar del Plata, Mar del Plata, Argentina,* Pesticide Chemistry in the Environment, Occurrence of Organochlorine Pesticides in Vegetables Grown on Untreated Soils from an Agricultural Watershed

R.P.C. Morgan, *Cranfield University, Silsoe, United Kingdom,* Soil Conservation

Roy Morgan, *Cranfield University, North Wyke, Devon, United Kingdom,* Soil Erosion and Control Practices

National Wild and Scenic Rivers System—National Park Service, River and Water Facts

J.R. Newman, *IACR-Centre for Aquatic Plant Management, Reading Berkshire, United Kingdom,* Weed Control Strategies

Jeffrey D. Niemann, *Colorado State University, Fort Collins, Colorado,* Drainage Networks

NOAA Great Lakes Environmental Research Lab, NOAA Lake Level Forecast for Lake Michigan Right on Target

C.S.P. Ojha, *Indian Institute of Technology, Roorkee, Uttaranchal, India,* Surface Water Pollution, Dilution-Mixing Zones and Design Flows, Assessment of Pollution Outflow From Large Agricultural Areas

Lindell E. Ormsbee, *University of Kentucky, Lexington, Kentucky,* Calibration of Hydraulic Network Models

Stefano Pagliara, *Universita di Pisa, Pisa, Italy,* Flood Control Structures, Hydraulics, River Basins, Levees for Flood Protection

Qiang Pan, *Washington State University, Pullman, Washington,* Aquaculture Technology for Producers

Geoffrey Petts, *University of Birmingham, Birmingham, United Kingdom,* Regulated Rivers, Environmental Flows

Laurel Phoenix, *Green Bay, Wisconsin,* Lakes—Discharges To

Gloria Post, *Division of Science, Research, and Technology, Trenton, New Jersey,* Fish Consumption Advisories

Simone Pozzolini, *Università di Pisa, Pisa, Italy,* Levees for Flood Protection

John C. Priscu, *Montana State University, Bozeman, Montana,* Subglacial Lake Vostok

Nitish Priyadarshi, *Ranchi University, Ranchi, Jharkhand, India,* Pollution of Surface Waters, Trace Elements in Water, Sediment, and Aquatic Biota—Effects of Biology and Human Activity, Cultural Eutrophication, Water Quality in Ponds

Rudi Rajar, *University of Ljubljana, Ljubljana, Slovenia,* Numerical Modeling of Currents

Tara Reed, *University of Wisconsin-Green Bay, Green Bay, Wisconsin,* Tile Drainage

Lucas Reijnders, *Universiteit van Amsterdam, Amsterdam, The Netherlands,* Biofuel Alternatives to Fossil Fuels

Jane Rickson, *Cranfield University, North Wyke, Devon, United Kingdom,* Soil Erosion and Control Practices

R.J. Rickson, *Cranfield University, Silsoe, United Kingdom,* Soil Conservation

Eliot C. Roberts, *Rosehall Associates, Sparta, Tennessee,* Soil Water Issues

T. Lackey Robert, *United States Environmental Protection Agency, Corvallis, Oregon,* Fisheries: History, Science, and Management

Barry H. Rosen, *U.S. Fish and Wildlife Service, Vero Beach, Florida,* Potential Health Issues Associated With Blue-Green Algae Blooms in Impoundments, Ponds and Lakes

William R. Roy, *Illinois State Geological Survey, Champaign, Illinois,* Remediation of Pesticide-Contaminated Soil at Agrichemical Facilities

Bahram Saghafian, *Soil Conservation and Watershed Management Research Institute, Tehran, Iran,* Time-Area and The Clark Rainfall-Runoff Transformation, Time of Concentration and Travel Time in Watersheds, Flood Source Mapping in Watersheds, Floodwater Spreading

Dibyendu Sarkar, *University of Texas at San Antonio, San Antonio, Texas,* Phytoremediation By Constructed Wetlands, Soil N Management Impact on The Quality of Surface and Subsurface Water, Soil Phosphorus Availability and Its Impact on Surface Water Quality

Hubert H.G. Savenije, *Delft University of Technology, The Netherlands,* Interception

Miklas Scholz, *The University of Edinburgh, Edinburgh, United Kingdom,* Urban Runoff

Harold L. Schramm, Jr, *U.S. Geological Survey, Mississippi State, Mississippi,* Water Needs for Freshwater Fisheries Management

Andy Seidl, *Colorado State University, Fort Collins, Colorado,* Agriculture and Land Use Planning

K.D. Sharma, *National Institute of Hydrology, Roorkee, Uttranchal, India,* Hydrological Processes and Measured Pollutant Loads

G. S. Shrivastava, *University of the West Indies, St. Augustine, Trinidad, West Indies,* Watershed Management for Environmental Quality and Food Security

Paul K. Sibley, *University of Guelph, Guelph, Ontario, Canada,* Chironomids in Sediment Toxicity Testing

Slobodan P. Simonovic, *The University of Western Ontario, London, Ontario, Canada,* Flood Prevention

Vijay P. Singh, *Louisiana State University, Baton Rouge, Louisiana,* Base Flow, Kinematic Wave Flow Routing, Kinematic Shock, Kinematic Wave and Diffusion Wave Theories, Surface Water Pollution, Hydrological Processes and Measured Pollutant Loads

Bellie Sivakumar, *University of California, Davis, California,* Hydrologic Thresholds

Nicolas Soumis, *University of Québec in Montréal, Montréal, Canada,* Hydroelectric Reservoirs as Anthropogenic Sources of Greenhouse Gases

Susan-Marie Stedman, *NMFS F/HC, Silver Spring, Maryland,* Coastal Wetlands

Pasquale Steduto, *Mediterranean Agronomic Institute, Valenzano, Italy,* Agricultural Water Use Efficiency (WUE) and Productivity (WP)

Alan Stern, *Division of Science, Research, and Technology, Trenton, New Jersey,* Fish Consumption Advisories

Eric A. Strauss, *Upper Midwest Environmental Sciences Center, La Crosse, Wisconsin,* Microbiology of Lotic Aggregates and Biofilms

Tim Sullivan, *E & S Environmental, Corvallis, Oregon,* Acidification—Chronic

Leonard I. Sweet, *Environmental Energy Group, Engineering Labs Inc., Canton, Michigan,* Fish Cells in the Toxicological Evaluation of Environmental Contaminants

Maria A. Szumiec, *Polish Academy of Sciences, Chybie, Poland,* Pond Aquaculture—Modeling and Decision Support Systems, An Outline of the History of Fishpond Culture in Silesia, the Western Part of Poland

David D. Tarkalson, *University of Nebraska-Lincoln, North Platte, Nebraska,* Nitrification

Gail P. Thelin, *U.S. Geological Survey,* Classification and Mapping of Agricultural Land for National Water-Quality Assessment

Chacharee Therapong, *University of Texas, San Antonio, Texas,* Phytoremediation By Constructed Wetlands

Mladen Todorovic, *Mediterranean Agronomic Institute of Bari, Valenzano, Bari, Italy,* Crop Water Requirements, Crop Evapotranspiration

Mark R. Tompkins, *University of California, Berkeley, California,* Fish Passage Facilities, Floodplain, Fishing Waters

U.S. Environmental Protection Agency—Office of Water, Office of Wetlands, Oceans and Watersheds, Wetlands Overview

Vandana Vandanapu, *University of Texas, San Antonio, Texas,* Soil N Management Impact on The Quality of Surface and Subsurface Water

Roger C. Viadero, *West Virginia University, Morgantown, West Virginia,* Sedimentation and Flotation, Factors Affecting Fish Growth and Production, Geochemistry of Acid Mine Drainage

J.M. Vose, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Measuring and Modeling Tree and Stand Level Transpiration

Thorsten Wagener, *The Pennsylvania State University, University Park, Pennsylvania,* Uncertainty Analysis in Watershed Modeling, Watershed Modeling, Modeling Ungauged Watersheds

Lizhu Wang, *Wisconsin Department of Natural Resources, Monona, Wisconsin,* Biotic Integrity Index to Evaluate Water Resource Integrity in Freshwater Systems

Sutat Weesakul, *DHI Water & Environment, Bangkok, Thailand,* Modeling of Urban Drainage and Stormwater

Brian B. Weigel, *Wisconsin Department of Natural Resources, Monona, Wisconsin,* Biotic Integrity Index to Evaluate Water Resource Integrity in Freshwater Systems

Sebastian Weissenberger, *University of Québec in Montréal, Montréal, Canada,* Hydroelectric Reservoirs as Anthropogenic Sources of Greenhouse Gases

Douglas F. Welsh, *Texas A&M University, Collage Station, Texas,* Xeriscape

Adrian E. Williams, *APEM Ltd., Manchester, United Kingdom,* Water Hyacinth—The World's Most Problematic Weed

Tommy S.W. Wong, *Nanyang Technological University, Singapore,* Kinematic Wave Method For Storm Drainage Design

Don J. Wood, *University of Kentucky, Lexington, Kentucky,* Calibration of Hydraulic Network Models

Chih Ted Yang, *Colorado State University, Fort Collins, Colorado,* Sedimentation, Dam Removal as River Restoration, Rivers

Pablo J. Zarco-Tejada, *Instituto de Agricultura Sostenible (IAS-CSIC), Córdoba, Spain,* Crop Water Stress Detection Using Remote Sensing

Weihua Zhang, *Florida International University, Miami, Florida,* Metal Tolerance in Plants: The Roles of Thiol-Containing Peptides

CONTRIBUTORS

M. Abad, *Universidad de Huelva. Avda. de las Fuerzas Armadas, Huelva,* A Statistical Approach to Critical Storm Period Analysis

Edinara Adelaide Boss, *Artur-Nogueira-SP, Brazil,* Sublimation

Joseph H. Aldstadt, III, *Genetic Technologies, Inc. Testing Institute, Waukesha, Wisconsin, In Situ* Chemical Monitoring

Alaska Biological Science Center—U.S. Geological Survey, Oceanographic Environment of Glacier Bay

Sergio Alonso, *University of the Balearic, Palma de Mallorca, Spain,* Relative Humidity

Marie de Angelis, *Humboldt State University, Arcata, California,* Major Ions in Seawater

Carolyn Ann Koh, *Colorado School of Mines, Golden, Colorado,* Clathrate Hydrates, Hydration

Yoseph Negusse Araya, *Open University, Milton Keynes, United Kingdom,* Hydrosphere

Ann Azadpour-Keeley, *National Risk Management Research Laboratory, ORD, U.S.EPA, Ada, Oklahoma,* Water and the History of Man

W.D. Bach, Jr., *(from The Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements, Wiley 2003),* Basic Research for Military Applications

Patrick L. Barry, *NASA,* Water on the Space Station

F. Batmanghelidj, *Global Health Solutions, Falls Church, Virginia,* The Myth of Bad Cholesterol: Why Water Is A Better Cholesterol-Lowering Medication, Water: The Key to Natural Health And Healing, The Mistake of Waiting to Get Thirsty

M. Eric Benbow, *Michigan State University, East Lansing, Michigan,* A History of Hawaiian Freshwater Resources

Mike Bettwy, *Goddard Space Flight Center, NASA,* African Monsoons

David Biggs, *University of Washington, Seattle, Washington,* Canals in the Mekong Delta: A Historical Overview from 200 C.E. to the Present

Jessica Black, *NOAA/CREST CCNY, New York, New York,* Remote Sensing of Applications in Hydrology

Peter E. Black, *SUNY ESF, Syracuse, New York,* Water Resource Organizations

Julián Blasco, *Institute of Marine Sciences of Analucia, Cadiz, Spain,* Marine and Estuarine Microalgal Sediment Toxicity Tests

M. Boström, *Linköping University, Linköping, Sweden,* Hofmeister Effects, Dissolved Gases

Kristofor R. Brye, *University of Arkansas, Fayetteville, Arkansas,* Nitrogen

Bureau of Labor Statistics, U.S. Department of Labor, Water Transportation Occupations

Bureau of Reclamation—U.S. Department of the Interior, Hoover Dam History

Robert M. Burgess, *U.S. Environmental Protection Agency, Narragansett, Rhode Island,* Ammonia

Albert J. Burky, *University of Dayton, Dayton, Ohio,* A History of Hawaiian Freshwater Resources

James Butler, *Climate Monitoring and Diagnostics Laboratory,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

Bradford Butman, *U.S. Geological Survey,* Mapping the Sea Floor of the Historic Area Remediation Site (HARS) Offshore of New York City

Pietro Chiavaccini, *Universitá di Pisa, Pisa, Italy,* Shallow Water Waves, Breakwaters

G.G. Clarke, *University of Hawaii, Hilo, Hawaii,* Climate and Water Balance on the Island of Hawaii

Jay Clausen, *AMEC Earth and Environmental, Inc., Westford, Massachusetts,* Technetium in Water, Beryllium in Water

Aldo Conti, *Frascati (RM), Italy,* Floating Ice, Condensation, Desertification, Dew, Distilled Water, Isotopes, Rain Forests, Dew Point

Giuseppe Cortese, *Alfred Wegener Institute for Polar and Marine Research (AWI), Bremerhaven, Germany,* Seawater Temperature Estimates in Paleoceanography

Sara Cotton, *University of Miami,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

L.M. Cáceres, *Universidad de Huelva. Avda. de las Fuerzas Armadas, Huelva,* A Statistical Approach to Critical Storm Period Analysis

Eli Dahi, *Environmental Development Corporation, Søborg, Denmark,* Defluroidation

Carl W. David, *University of Connecticut, Storrs, Connecticut,* The Hydronium Ion

Jana Davis, *Smithsonian Environmental Research Center, Edgewater, Maryland,* Marine Stock Enhancement Techniques

Christine Dickenson, *Florida Institute of Technology, Melbourne, Florida,* An Analysis of the International Maritime Organization–London Convention Annual Ocean Dumping Reports

Robert E. Dickinson, *(from The Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements, Wiley 2003),* Overview: The Climate System

Priyanka K. Dissanayake, *Moratuwa, Sri Lanka,* Estuarian Waters

Iver W. Duedall, *Florida Institute of Technology, Melbourne, Florida,* An Analysis of the International Maritime Organization–London Convention Annual Ocean Dumping Reports

Sandra Dunbar, *Napier University, Edinburgh, Scotland, United Kingdom,* Partitioning and Bioavailability

Antonín Dvořák, *University of Economics, Prague, Czech Republic,* Negotiating between Authority and Polluters: An Approach to Managing Water Quality

David R. Easterling, *(from Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements, Wiley 2003),* Observations of Climate and Global Change from Real-Time Measurements

Theodore A. Endreny, *SUNY-ESF, Syracuse, New York,* Remote Sensing of Applications in Hydrology, Evaporation, Radar Use in Rainfall Measurements, Rainfall and Runoff

Environment Canada, Water—Nature's Magician

Timothy Erickson, *National Weather Service New Orleans/Baton Rouge Forecast Office, Slidell, Louisiana,* Coastal Fog Along the Northern Gulf of Mexico

Kirsten Exall, *National Water Research Institute, Burlington, Ontario, Canada,* Coagulation and Flocculation

David L. Feldman, *University of Tennessee, Knoxville, Tennessee,* Water Supply Planning—Federal, Water Transfers

Montserrat Filella, *University of Geneva, Geneva, Switzerland,* Antimony in Aquatic Systems

Sylwester Furmaniak, *N. Copernicus University, Toruń, Poland,* Mechanisms of Water Adsorption on Carbons

Piotr A. Gauden, *N. Copernicus University, Toruń, Poland,* The Effect of Carbon Surface Chemical Composition on the Mechanism of Phenol Adsorption from Aqueous Solutions, Mechanisms of Water Adsorption on Carbons

Geophysical Fluid Dynamics Laboratory—NOAA, El Niño: The Interannual Prediction Problem, Hurricanes: Modeling Nature's Fury, Global Climate Change

Andrea K. Gerlak, *Columbia University, New York, New York,* Wetlands Policy in the United States: From Drainage to Restoration

Michael H. Glantz, *(from The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts, Wiley 2003),* Climate and Society

Global Change Research Program—U.S. Geological Survey, Flattop Mountain Snotel Snowpack: Water Year 2004

Jyotsna Goel, *Centre for Fire, Explosives, and Environmental Safety, Timarpur, India,* Ion Exchange and Inorganic Adsorption

Kelly Goodwin, *University of Miami,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

David M. Gray, *Mettler-Toledo Thornton Inc., Bedford, Massachusetts,* Conductivity-Electric

Great Lakes Environmental Research Laboratory (NOAA), Technology Development: Hardware Development—Marine Instrumentation Laboratory (MIL)

Seth I. Gutman, *NOAA Forecast Systems Laboratory,* Ground-Based GPS Meteorology at FSL

Janusz Guziur, *University of Warmia and Mazury in Olsztyn, Olsztyn, Poland,* History of Pond Fisheries in Poland

David Haley, *Picacadilly, Ulverston, Cumbria, United Kingdom,* Water and Well-Being, Evolution

Jon Hare, *NOAA Center for Coastal Fisheries and Habitat Research,* NOS/NMFS Cooperative Research on Coastal Fisheries and Habitats at the Beaufort Laboratory

Michael J. Hayes, *Climate Impacts Specialist—National Drought Mitigation Center,* Drought Indices

William Henderson, *University of Waikato, Hamilton, New Zealand,* Analysis of Aqueous Solutions Using Electrospray Ionization Mass Spectrometry (esi ms)

Warwick Hillier, *The Australian National University, Canberra, Australia,* Isotope Exchange in Gas-Water Reactions

Joseph Holden, *University of Leeds, Leeds, United Kingdom,* Mariotte Bottle—Use in Hydrology, Rain Simulators

S. Holgate, *Proudman Oceanographic Laboratory, Birkenhead, United Kingdom,* The Permanent Service for Mean Sea Level

Arthur M. Holst, *Philadelphia Water Department, Philadelphia, Pennsylvania,* Frost Damage, Droughts, Water Cycle, Monsoon, Permanent Frost, United States Weather Bureau, Waterspout, Cloud Seeding, Cyclones, Fog, Frost, Chinook, Flood Control Act of 1944, Jacob's Well, Gordon and Franklin Rivers and the Tasmanian Wilderness World Heritage Area

Kirk L. Holub, *NOAA Forecast Systems Laboratory,* Ground-Based GPS Meteorology at FSL

Paul R. Houser, *(from The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts, Wiley 2003),* Infiltration and Soil Moisture Processes

S.A. Hsu, *Louisiana State University, Baton Rouge, Louisiana,* Air–Sea Interaction

Jason A. Hubbart, *University of Idaho, Moscow, Idaho,* Hydrologic Cycle, Water Resources, and Society, Hydrologic History, Problems, and Perspectives

Basia Irland, *University of New Mexico, Albuquerque, New Mexico,* A Concise Glimpse of Water in the History of Photography

Mohammed Riajul Islam, *University of Idaho, Moscow, Idaho,* Rain and Rocks: The Recipe for River Water Chemistry

James A. Jacobs, *Environmental Bio-Systems, Inc., Mill Valley, California,* Fenton's Reaction and Groundwater Remediation, Regulatory Issues and Remediation: Risk, Costs, and Benefits, Water, Bacteria, Life on Mars, and Microbial Diversity

Sharad K. Jain, *National Institute of Hydrology, Roorkee, India,* Isohyetal Method, Hydrologic Cycle

Purnima Jalihal, *National Institute of Ocean Technology, Chennai, India,* Renewable Energies from the Ocean

S. Jevrejeva, *Proudman Oceanographic Laboratory, Birkenhead, United Kingdom,* The Permanent Service for Mean Sea Level

Alicia Jimenez, *Michigan State University, East Lansing, Michigan,* Transboundary Waters in Latin America: Conflicts and Collaboration

Anne Jones-Lee, *G. Fred Lee & Associates, El Macero, California,* Clean Water Act, Water Quality Criteria/Standards, TMDLs, and Weight-of-Evidence Approach for Regulating Water Quality

Andrew Juhl, *Lamont–Doherty Earth Observatory of Columbia University, Palisades, New York,* Food Chain/Foodweb/Food Cycle

James O. Juvik, *University of Hawaii, Hilo, Hawaii,* Climate and Water Balance on the Island of Hawaii

K. Kadirvelu, *Centre for Fire, Explosives, and Environmental Safety, Timarpur, India,* Ion Exchange and Inorganic Adsorption

Th.D. Karapantsios, *Aristotle University, Thessaloniki, Greece,* Sorption Kinetics

Thomas R. Karl, *(from Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements, Wiley 2003),* Observations of Climate and Global Change from Real-Time Measurements

Melinda R. Kassen, *Trout Unlimited, Colorado Water Project, Boulder, Colorado,* Legal Protection for In-stream Flow, Interface between Federal Water Quality Regulation and State Allocation of Water Quantity

Gholam A. Kazemi, *Shahrood University of Technology, Shahrood, Iran,* Age Dating Old Groundwater, Chlorine-36 and Very Old Groundwaters, Deuterium, Freshwater, Hard Water, Heavy Water, Hydrogen Ion, Isotope Fractionation, Soft Water, Strontium Isotopes in Water and Rock, Chlorofluorocarbons (CFCs)

Jack W. Keeley, *Environmental Engineer, Ada, Oklahoma,* Water and the History of Man

Giora J. Kidron, *The Hebrew University of Jerusalem, Jerusalem, Israel,* Dew Deserts

Daniel King, *University of Colorado,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

Piotr Kowalczyk, *Faculty of Science, Chiba University, Chiba, Japan,* The Effect of Carbon Surface Chemical Composition on the Mechanism of Phenol Adsorption from Aqueous Solutions, Mechanisms of Water Adsorption on Carbons

Upadhyayula V. K. Kumar, *Choa Chu kang Ave-4, Singapore,* Coastal Water Pollutants

W. Kunz, *University of Regensburg, Regensburg, Germany,* Hofmeister Effects

Frederic Lasserre, *Laval University, Quebec City, Quebec, Canada,* Great Lakes Governors' Agreement

Marshall Lawson, *Land Conservation Legal Services, LLC, Columbia, South Carolina,* Quantitative GroundWater Law

Jamie R. Lead, *University of Birmingham, Birmingham, United Kingdom,* Dissolved Organic Carbon

G. Fred Lee, *G. Fred Lee & Associates, El Macero, California,* Clean Water Act, Water Quality Criteria/Standards, TMDLs, and Weight-of-Evidence Approach for Regulating Water Quality

Maggie Lee, *Santa Fe, New Mexico,* Free Flowing Water: A Source of Wisdom

Nai Kuang Liang, *National Taiwan University, Taipei, Taiwan,* Water Waves

Sharon L. Lien, *The Groundwater Foundation, Lincoln, Nebraska,* Effective Water Education Strategies in a Nontraditional Setting

Clive D. Lipchin, *Ann Arbor, Michigan,* Water Between Arabs and Israelis: Researching Twice-Promised Resources, Conflict and Water Use in the Middle East

Eileen Loiseau, *Bigelow Laboratory for Ocean Sciences,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

M.X. Loukidou, *Aristotle University, Thessaloniki, Greece,* Sorption Kinetics

Helen H. Lou, *Lamar University, Beaumont, Texas,* Solubility of Hydrocarbons and Sulfur Compounds in Water, Solubility of Chemicals in Water

Nancy A. Lowery, *San Diego, California,* United States-Mexico Border Waters: Conventions, Treaties, and Institutions

Luís M. Lubián, *Institute of Marine Sciences of Analucia, Cadiz, Spain,* Marine and Estuarine Microalgal Sediment Toxicity Tests

Paolo Magni, *IMC—International Marine Centre, Torregrande-Oristano, Italy,* Laboratory Experiments On Bivalve Excretion Rates of Nutrients, Temporal Scaling of Benthic Nutrient Regeneration in Bivalve-Dominated Tidal Flat, Physical and Chemical Variability of Tidal Streams, Tidally Mediated Changes in Nutrient Concentrations

Babs A. Makinde-Odusola, *Riverside, California,* Radon in Water

William E. Marks, *Water Consciousness, Inc., Martha's Vineyard, Massachusetts,* Benjamin Franklin: From Kite to Lightning Rod, Benjamin Franklin's Armonica: A Water Music Instrument, Our Evolving Water Consciousness, Water Clocks, Ben Franklin's Gulf Stream Weather and Swim Fins

D.L. Marrin, *Hanalei, Hawaii,* Cosmic Water, Molecular Network Dynamics, Sound in Water, Water Symbolism

K.A. Matis, *Aristotle University, Thessaloniki, Greece,* Sorption Kinetics

Patricia Matrai, *Bigelow Laboratory for Ocean Sciences,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

Mollie D. McIntosh, *Michigan State University, East Lansing, Michigan,* A History of Hawaiian Freshwater Resources

Kevin S. McLeary, Physical Properties

Lucius O. Mendis, *Colombo, Sri Lanka,* Ancient Water and Soil Conservation Ecosystems of Sri Lanka

Carlos D. Messina, *University of Florida, Gainesville, Florida,* Degree Day Method

Paulette Middleton, *(from The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts, Wiley 2003),* Acid Rain and Society

Elsie F. Millano, *ERM, Inc., Vernon Hills, Illinois,* Great Lakes Water Quality Initiative

A. Trent Millet, *Newark, Vermont,* The Medicinal Properties of the Waters of Saratoga Springs

Koichi Miyashita, *Okayama University, Kurashiki, Japan,* Transpiration

Eric P. Mollard, *Institut de Recherche pour le Développement, France,* Curious Uses of Agricultural Water in the World, Jaubert de Passa: The First World History of Irrigation in 1846

Shigeru Montani, *Hokkaido University, Hakodate, Japan,* Physical and Chemical Variability of Tidal Streams, Laboratory Experiments On Bivalve Excretion Rates of Nutrients, Tidally Mediated Changes in Nutrient Concentrations, Temporal Scaling of Benthic Nutrient Regeneration in Bivalve-Dominated Tidal Flat, Seasonal Coupling Between Intertidal Macrofauna and Sediment Column Porewater Nutrient Concentrations

Robert M. Moore, *Dalhousie University, Halifax, Nova Scotia, Canada,* Marine Sources of Halocarbons

Ignatio Moreno-Garrido, *Institute of Marine Sciences of Analucia, Cadiz, Spain,* Marine and Estuarine Microalgal Sediment Toxicity Tests

Suparna Mukherji, *IIT Bombay, Powai, Mumbai, India,* Adsorption of Organic Compounds

J.M. Muñoz, *Universidad de Sevilla. Avda. Reina Mercedes, Sevilla,* A Statistical Approach to Critical Storm Period Analysis

Prasad K. Narasimhan, *Lamar University, Beaumont, Texas,* Adsorption Capacity of Activated Carbon for Water Purification, Solubility of Hydrocarbons and Sulfur Compounds in Water, Solubility of Chemicals in Water, Solubility of Hydrocarbons in Salt Water

NASA—Goddard Space Flight Center, Black Water Turns the Tide on Florida Coral, The Water Cycle, Weather Forecasting Through The Ages, The Earth Observing System: Aqua

NASA—Langley Research Center, CERES: Understanding The Earth's Clouds and Climate

NASA Marshall Space Flight Center, Where Water Floats

National Drought Mitigation Center, What is Climatology?

National Oceanographic and Atmospheric Administration (NOAA), NOAA's Atlantic Oceanographic and Meteorological Laboratory

Natural Resources Conservation Service, Conservation and the Water Cycle

John W. Nielsen-Gammon, *(from Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements, Wiley 2003),* Overview of Weather Systems

Robert Y. Ning, *King Lee Technologies, San Diego, California,* Silica in Natural Waters, Carbonate in Natural Waters

B.W. Ninham, *University of Regensburg, Regensburg, Germany and Australian National University, Canberra, Australia,* Hofmeister Effects, Dissolved Gases

NOAA Coral Reef Information System (CORIS), Deep Water Corals

NOAA National Ocean Service, NOS Sanctuaries Protect Nation's Maritime History

NOAA—Pacific Marine Environmental Laboratory, Pacific Marine Environmental Laboratory—30 Years of Observing The Ocean

Northeast Fisheries Science Center—NOAA, Woods Hole: The Early Years

Gertrud K. Nürnberg, *Freshwater Research, Baysville, Ontario, Canada,* Quantification of Anoxia and Hypoxia in Water Bodies

Pacific Fisheries Environmental Laboratory, NOAA, Physical Oceanography

Pacific Northwest National Laboratory—Shrub-Steppe Ecology Series, What About Meteorology?

Stefano Pagliara, *Universitá di Pisa, Pisa, Italy,* Shallow Water Waves, Breakwaters

Magni Paolo, *IMC-International Marine Centre, Torregrande-Oristano, Italy,* Seasonal Coupling Between Intertidal Macrofauna and Sediment Column Porewater Nutrient Concentrations

Jose O. Payero, *University of Nebraska-Lincoln, North Platte, Nebraska,* Evapotranspiration

Mauricio Peredo, *NASA Goddard Space Flight Center,* Electricity as a Fluid

Naraine Persaud, *Virginia Polytechnic Institute and State University, Blacksburg, Virginia,* Adiabatic Cooling, Humidity—Absolute, Heat of Vaporization, Vapor Pressure

Geoff Petts, *University of Birmingham, Birmingham, United Kingdom,* Water in History

Tony Phillips, *NASA,* Water on the Space Station

Jim Philp, *Napier University, Edinburgh, Scotland, United Kingdom,* Partitioning and Bioavailability

Laurel Phoenix, *Green Bay, Wisconsin,* Water as a Human Right

Physics Laboratory, National Institute of Standards and Technology, Early Clocks

Ralph W. Pike, *Louisiana State University, Baton Rouge, Louisiana,* Solubility of Hydrocarbons and Sulfur Compounds in Water, Solubility of Chemicals in Water

R. Pino, *Universidad de Sevilla. Avda. Reina Mercedes, Sevilla,* A Statistical Approach to Critical Storm Period Analysis

Richard Z. Poore, *U.S. Geological Survey, Reston, Virginia,* Sea Level and Climate

Bobby J. Presley, *Texas A&M University, College Station, Texas,* Trace Element Pollution

Nitish Priyadarshi, *Ranchi University, Ranchi, Jharkhand, India,* Sodium in Natural Waters

Jian-Wen Qiu, *Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong,* Larvae and Small Species of Polychaetes in Marine Toxicological Testing

Climent Ramis, *University of the Balearic, Palma de Mallorca, Spain,* Relative Humidity

Jorge Ramírez-Vallejo, *Harvard University, Cambridge, Massachusetts,* Economic Value of Water: Estimation

Todd Rasmussen, *The University of Georgia, Athens, Georgia,* Barometric Efficiency

Howard Reminick, *Ohno Institute on Water and Health, Willoughby, Ohio,* An Analysis of The Impact of Water on Health and Aging: Is All Water The Same?

Martin Reuss, Ph.D, *Office of History Headquarters, U.S. Army Corps of Engineers,* The Constitution and Early Attempts at Rational Water Planning

J. Rodríguez Vidal, *Universidad de Huelva. Avda. de las Fuerzas Armadas, Huelva,* A Statistical Approach to Critical Storm Period Analysis

A. Rodríguez-Ramírez, *Universidad de Huelva. Avda. de las Fuerzas Armadas, Huelva,* A Statistical Approach to Critical Storm Period Analysis

Romualdo Romero, *University of the Balearic, Palma de Mallorca, Spain,* Relative Humidity

William R. Roy, *Illinois State Geological Survey, Champaign, Illinois,* Iron

F. Ruiz, *Universidad de Huelva. Avda. de las Fuerzas Armadas, Huelva,* A Statistical Approach to Critical Storm Period Analysis

David L. Russell, *Global Environmental Operations, Inc., Lilburn, Georgia,* A Brief History of the Water Pollution Control Act in the U.S.

Timothy J. Ryan, *Ohio University, Athens, Ohio,* The Clean Water Act

Gerhard Rychlicki, *N. Copernicus University, Toruń, Poland,* The Effect of Carbon Surface Chemical Composition on the Mechanism of Phenol Adsorption from Aqueous Solutions, Mechanisms of Water Adsorption on Carbons

S. Sabri, *University of Malaya, Petaling Jaya Selangor, Malaysia,* Islamic Water Law

Basu Saha, *Loughborough University, Loughborough, United Kingdom,* Removal of Organic Micropollutants and Metal Ions from Aqueous Solutions by Activated Carbons

Eric Saltzman, *University of Miami,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

Peter H. Santschi, *Texas A&M University, Galveston, Texas,* Marine Colloids

Edward S. Sarachik, *(from The Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements, Wiley 2003),* The Ocean in Climate

Eva Saroch, *Panjab University, Chandigarh, India,* Representing Geopolitics of (Hydro) Borders in South Asia

Petr Šauer, *University of Economics, Prague, Czech Republic,* Negotiating between Authority and Polluters: An Approach to Managing Water Quality

Donald Savage, *NASA Headquarters, Washington,* Mars Exploration Rover Mission

Reuel Shinnar, *The City College of the CUNY, New York,* The Mirage of The H_2 Economy

Melissa A. Singer Pressman, *Genetic Technologies, Inc. Testing Institute, Waukesha, Wisconsin, In Situ* Chemical Monitoring

Pratap Singh, *National Institute of Hydrology, Roorkee, India,* Snow Density, Snow and Snowmelt, Snow Surveys

Vijay P. Singh, *Louisiana State University, Baton Rouge, Louisiana,* Entropy Theory For Hydrologic Modeling, Unit Hydrograph Theory, Hydrologic Cycle, Isohyetal Method

D.C. Singleton, *University of Hawaii, Hilo, Hawaii,* Climate and Water Balance on the Island of Hawaii

Bellie Sivakumar, *University of California, Davis, California,* Hydropsychology

E. Dendy Sloan, *Colorado School of Mines, Golden, Colorado,* Clathrate Hydrates

James A. Smith, *(from The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts, Wiley 2003),* Rainfall

Fiona Stainsby, *Napier University, Edinburgh, Scotland, United Kingdom,* Partitioning and Bioavailability

Susan-Marie Stedman, *Silver Spring, Maryland,* Coastal Waters

Kenneth F. Steele, *University of Arkansas, Fayetteville, Arkansas,* Nitrogen, Carbonate Geochemistry

David P. Stern, *Goddard Space Flight Center, NASA,* Electricity as a Fluid, Weather and the Atmosphere

Bradley A. Striebig, *Gonzaga University, Spokane, Washington,* Chemical Precipitation

Georgina Sturrock, *Commonwealth Scientific and Industrial Research Organization—Australia,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

Artur P. Terzyk, *N. Copernicus University, Toruń, Poland,* The Effect of Carbon Surface Chemical Composition on the Mechanism of Phenol Adsorption from Aqueous Solutions, Mechanisms of Water Adsorption on Carbons

Ryszard Tokarczyk, *University of Miami,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

Marta E. Torres, *Oregon State University, Corvallis, Oregon,* Distribution and Dynamics of Gas Hydrates in the Marine Environment

Christopher Tracey, *U.S. Geological Survey, Reston, Virginia,* Sea Level and Climate

Anne M. Trehu, *Oregon State University, Corvallis, Oregon,* Distribution and Dynamics of Gas Hydrates in the Marine Environment

Robert C. Upstill-Goddard, *University of Newcastle upon Tyne, Newcastle upon Tyne, United Kingdom,* The Role of Oceans in the Global Cycles of Climatically-Active Trace-Gases

David B. Vance, *ARCADIS, Midland, Texas,* Regulatory Issues and Remediation: Risk, Costs, and Benefits, Water, Bacteria, Life on Mars, and Microbial Diversity

Javier Velez-Arocho, *U.S. Environmental Protection Agency, Washington, District of Columbia,* Marine Debris Abatement

Roger C. Viadero Jr., *West Virginia University, Morgantown, West Virginia,* Henry's Law

Rolf R. von Oppenfeld, *The TESTLaw Practice Group, Phoenix, Arizona,* The National Pollution Discharge Elimination System, The Safe Drinking Water Act

Linda Voss, *U.S. Centennial of Flight Commission,* Ballooning and Meteorology in the Twentieth Century

U.S. Department of Labor, Bureau of Labor Statistics, Atmospheric Scientists

U.S. Environmental Protection Agency—Oceans and Coastal Protection Division, Coral Reefs and Your Coastal Watershed

U.S. Geological Survey, Water Quality

U.S. Global Change Research Program, The Global Water Cycle

Guy Webster, *Jet Propulsion Laboratory, Pasadena, California,* Mars Exploration Rover Mission

Radosław P. Wesołowski, *N. Copernicus University, Toruń, Poland,* Mechanisms of Water Adsorption on Carbons

Scott Whiteford, *Michigan State University, East Lansing, Michigan,* Transboundary Waters in Latin America: Conflicts and Collaboration

Patrick Willems, *Hydraulics Laboratory, Leuven, Belgium,* Uncertainties in Rainfall–Runoff Modeling

Richard S. Williams, Jr., *U.S. Geological Survey, Woods Hole, Massachusetts,* Sea Level and Climate

Marek Wiśniewski, *N. Copernicus University, Toruń, Poland,* Mechanisms of Water Adsorption on Carbons

Ming Hung Wong, *Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong,* Larvae and Small Species of Polychaetes in Marine Toxicological Testing

Eve Woods, *Denver, Colorado,* Reserved Water Rights for Indian and Federal Lands

P.L. Woodworth, *Proudman Oceanographic Laboratory, Birkenhead, United Kingdom,* The Permanent Service for Mean Sea Level

Tom Wydrzynski, *The Australian National University, Canberra, Australia,* Isotope Exchange in Gas-Water Reactions

Carl L. Yaws, *Lamar University, Beaumont, Texas,* Adsorption Capacity of Activated Carbon for Water Purification, Solubility of Hydrocarbons and Sulfur Compounds in Water, Solubility of Chemicals in Water, Solubility of Hydrocarbons in Salt Water

Brian Yocis, *Bigelow Laboratory for Ocean Sciences,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

David W. Yoskowitz, *Texas A&M University—Corpus Christi, Laredo, Texas,* U.S./Canadian Boundary Waters Treaty and the Great Lakes Water Quality Agreement

Mark A. Young, *University of Iowa, Iowa City, Iowa,* Environmental Photochemistry in Surface Waters

Shari Yvon-Lewis, *Atlantic Oceanographic and Meteorological Laboratory,* NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific

Karl Erik Zachariassen, *Norwegian University of Science and Technology, Trondheim, Norway,* Freezing and Supercooling of Water, Osmosis-Diffusion of Solvent or Caused by Diffusion of Solutes?

A. Zaharuddin, *University of Malaya, Petaling Jaya Selangor, Malaysia,* Islamic Water Law

CONTRIBUTORS

Segun Adelana, *University of Ilorin, Ilorin, Nigeria,* Summary of Isotopes in Contaminant Hydrogeology, Environmental Isotopes in Hydrogeology

Mohammad N. Almasri, *An-Najah National University, Nablus, Palestine,* Groundwater Flow and Transport Process

Tom A. Al, *University of New Brunswick, Fredericton, New Brunswick, Canada,* River-Connected Aquifers: Geophysics, Stratigraphy, Hydrogeology, and Geochemistry

Larry Amskold, *University of New Brunswick, Fredericton, New Brunswick, Canada,* River-Connected Aquifers: Geophysics, Stratigraphy, Hydrogeology, and Geochemistry

Ann Azadpour-Keeley, *National Risk Management Research Laboratory, ORD, U.S. EPA, Ada, Oklahoma,* Microbial Processes Affecting Monitored Natural Attenuation of Contaminants in the Subsurface, Nitrate Contamination of Groundwater

Mukand Singh Babel, *Asian Institute of Technology, Pathumthani, Thailand,* Groundwater Velocities, Groundwater Flow Properties, Water in The Unsaturated Zone

Philip B. Bedient, *Rice University, Houston, Texas,* Transport of Reactive Solute in Soil and Groundwater

David M. Bednar, Jr., *Michael Baker, Jr. Inc., Shreveport, Louisiana,* Karst Hydrology, Karst Topography, Groundwater Dye Tracing in Karst

Milovan Beljin, *Cincinnati, Ohio,* Horizontal Wells

Craig H. Benson, *University of Wisconsin-Madison, Madison, Wisconsin,* Reactive Transport in The Saturated Zone: Case Histories for Permeable Reactive Barriers

Robert A. Bisson, *Alexandria, Virginia,* Megawatersheds

William J. Blanford, *Louisiana State University, Baton Rouge, Louisiana,* Vadose Zone Monitoring Techniques

Thomas B. Boving, *University of Rhode Island, Kingston, Rhode Island,* Organic Compounds in Ground Water, Innovative Contaminated Groundwater Remediation Technologies

Richard C. Brody, *UC Berkeley, Berkeley, California,* Connate Water

Kristofor R. Brye, *University of Arkansas, Fayetteville, Arkansas,* Lysimeters, Soil and Water Contamination by Heavy Metals

Gunnar Buckau, *Institut für Nukleare Entsorgung, Karlsruhe, Germany,* Mobility of Humic Substances in Groundwater

Bureau of Indian Affairs and Arizona Department of Water Resources—U.S. Geological Survey, Black Mesa Monitoring Program

Karl E. Butler, *University of New Brunswick, Fredericton, New Brunswick, Canada,* River-Connected Aquifers: Geophysics, Stratigraphy, Hydrogeology, and Geochemistry

Herbert T. Buxton, *United States Geological Survey,* Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams

Natalie L. Capiro, *Rice University, Houston, Texas,* Transport of Reactive Solute in Soil and Groundwater

Harendra S. Chauhan, *G.B. Pant University of Agriculture and Technology, Uttar Pradesh, India,* Steady-State Flow Aquifer Tests, Subsurface Drainage

Bernard L. Cohen, *University of Pittsburgh, Pittsburgh, Pennsylvania,* Risk Analysis of Buried Wastes From Electricity Generation

David P. Commander, *Water and Rivers Commission, East Perth, Australia,* Water Dowsing (Witching), Artesian Water

Dennis L. Corwin, *USDA-ARS George E. Brown, Jr., Salinity Laboratory, Riverside, California,* Characterizing Soil Spatial Variability, Modeling Non-Point Source Pollutants in the Vadose Zone Using GIS, Groundwater Vulnerability to Pesticides: An Overview of Approaches and Methods of Evaluation

Colin C. Cunningham, *The University of Edinburgh, Edinburgh, Scotland, United Kingdom, In Situ* Bioremediation of Contaminated Groundwater

William L. Cunningham, *U.S. Geological Survey, Denver, Colorado,* Earthquakes—Rattling the Earth's Plumbing System

Uwe Dannwolf, *URS Australia Pty Ltd., Turner, Australia,* Groundwater and Vadose Zone Hydrology

Diganta Bhusan Das, *Oxford University, Oxford, United Kingdom,* Viscous Flow, Finite Element Modeling of Coupled Free and Porous Flow, Combined Free and Porous Flow in the Subsurface, Modeling Techniques for Solute Transport in Groundwater

Rupali Datta, *University of Texas at San Antonio, San Antonio, Texas,* Remediation of Contaminated Soils, Genetics of Metal Tolerance and Accumulation in Higher Plants, Phytoextraction of Zinc and Cadmium from Soils Using Hyperaccumulator Plants, Phytoremediation of Selenium-Laden Soils, Phytoextraction and Phytostabilization: Technical, Economic and Regulatory Considerations of the Soil-Lead Issue

Ali H. Davani, *University of Texas at San Antonio, San Antonio, Texas,* Remediation of Contaminated Soils

L.C. Davis, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Phytoremediation of Methyl *Tertiary*-Butyl Ether

Melissa R. Dawe, *University of New Brunswick, Fredericton, New Brunswick, Canada,* River-Connected Aquifers: Geophysics, Stratigraphy, Hydrogeology, and Geochemistry

Steven A. Dielman, *ENVIRON International Corporation, Arlington, Virginia,* Hydraulic Conductivity/Transmissibility

Craig E. Divine, *Colorado School of Mines, Golden, Colorado,* Groundwater Sampling with Passive Diffusion Samplers, Detecting Modern Groundwaters with ^{85}Kr, Groundwater Dating with H–He

Shonel Dwyer, *Environmental Bio-Systems, Inc., Mill Valley, California,* Groundwater and Perchlorate: Chemical Behavior and Treatment

Aly I. El-Kadi, *University of Hawaii at Manoa, Honolulu, Hawaii,* Unconfined Groundwater

Environment Canada, Groundwater—Nature's Hidden Treasure

L.E. Erickson, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Phytoremediation of Methyl *Tertiary*-Butyl Ether

Thomas R. Fisher, *Horn Point Laboratory—UMCES, Solomons, Maryland,* What is a Hydrochemical Model?

Craig Foreman, *Environmental Bio-Systems, Inc., Mill Valley, California,* Groundwater and Cadmium: Chemical Behavior and Treatment

Devin L. Galloway, *U.S. Geological Survey, Denver, Colorado,* Earthquakes—Rattling the Earth's Plumbing System

Lorraine Geddes-McDonald, *Environmental Bio-Systems, Inc., Mill Valley, California,* Groundwater and Nitrate: Chemical Behavior and Treatment

Mário Abel Gonçalves, *Faculdade de Ciências da Universidade de Lisoba, Lisoba, Portugal,* Metal Organic Interactions in Subtitle D Landfill Leachates and Associated Ground Waters, Geochemical Modeling-Computer Codes, Geochemical Models

Jason J. Gurdak, *U.S. Geological Survey, Lakewood, Colorado and Colorado School of Mines, Golden, Colorado,* Groundwater Vulnerability to Pesticides: Statistical Approaches

Navraj S. Hanspal, *Loughborough University, Loughborough, United Kingdom,* Modeling Techniques for Solute Transport in Groundwater, Viscous Flow, Laminar Flow, Finite Element Modeling of Coupled Free and Porous Flow

Thomas Harter, *University of California, Davis, California,* Specific Yield Storage Equation, Vulnerability Mapping of Groundwater Resources, Aquifers

Blayne Hartman, *H&P Mobile Geochemistry, Solana Beach, California,* Applications of Soil Vapor Data to Groundwater Investigations

Joseph Holden, *University of Leeds, Leeds, United Kingdom,* Infiltrometers, Soil Pipes and Pipe Flow, Infiltration and Soil Water Processes, Darcy's Law, Infiltration/Capacity/Rates

Ekkehard Holzbecher, *Humboldt Universität Berlin, Berlin, Germany,* Groundwater Modeling, Ghijben–Herzberg Equilibrium

Paul F. Hudak, *University of North Texas, Denton, Texas,* Mass Transport in Saturated Media

John D. Humphrey, *Colorado School of Mines, Golden, Colorado,* Groundwater Dating with H–He

S.L. Hutchinson, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Hydrologic Feasibility Assessment and Design in Phytoremediation

Th.A. Ioannidis, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* Phytoremediation of Lead-Contaminated Soils

Jahangir Islam, *City Design Limited, Auckland, New Zealand,* Sub-Surface Redox Chemistry: A Comparison of Equilibrium and Reaction-Based Approaches

Irena B. Ivshina, *Institute of Ecology and Genetics of Microorganisms of the RAS, Perm, Russia, In Situ* Bioremediation of Contaminated Groundwater

James A. Jacobs, *Environmental Bio-Systems, Inc., Mill Valley, California,* Groundwater and Cobalt: Chemical Behavior and Treatment, Limiting Geochemical Factors in Remediation Using Monitored Natural Attenuation and Enhanced Bioremediation, The Role of Heat in Groundwater Systems, Horizontal Wells in Groundwater Remediation, Groundwater Flow in Heterogenetic Sediments and Fractured Rock Systems, Groundwater and Cadmium: Chemical Behavior and Treatment, Groundwater and Benzene: Chemical Behavior and Treatment, Groundwater and Lead: Chemical Behavior and Treatment, Groundwater and Nitrate: Chemical Behavior and Treatment, Groundwater and Uranium: Chemical Behavior and Treatment, Groundwater and Mercury: Chemical Behavior and Treatment, The Environmental Impact of Iron in Groundwater, Water Well Drilling Techniques, Water-Jetting Drilling Technologies for Well Installation And *In Situ* Remediation of Hydrocarbons, Solvents, and Metals, Source, Mobility, and Remediation of Metals, Particulate Transport in Groundwater—Bacteria and Colloids, Groundwater and Arsenic: Chemical Behavior and Treatment, *In Situ* Groundwater Remediation for Heavy Metal Contamination, Groundwater Remediation by *In Situ* Aeration and Volatilization, MTBE, Phytoremediation Enhancement of Natural Attenuation Processes, Groundwater Remediation by Injection and Problem Prevention, Chemical Oxidation Technologies for Groundwater Remediation, Physical Properties of DNAPLs and Groundwater Contamination, Process Limitations of *In Situ* Bioremediation of Groundwater, Water Contamination by Low Level Organic Waste Compounds in the Hydrologic System, Applications of Soil Vapor Data to Groundwater Investigations, Groundwater Remediation Project Life Cycle, Groundwater Remediation: *In Situ* Passive Methods, Groundwater and Vinyl Chloride: Chemical Behavior and Treatment, Groundwater and Perchlorate: Chemical Behavior and Treatment, Groundwater Sampling Techniques for Environmental Projects, Low Flow Groundwater Purging and Surging

Hamid R. Jahani, *Water Research Institute, Hakimieh, Tehran, Iran,* Groundwater Tracing, Resistivity Methods

Chakresh K. Jain, *National Institute of Hydrology, Roorkee, India,* Assessment of Groundwater Quality in District Hardwar, Uttaranchal, India, Nonpoint Sources, Fluoride Contamination in Ground Water, Irrigation Water Quality in District Hardwar, Uttaranchal, India

John R. Jansen, *Aquifer Science & Technology, Waukesha, Wisconsin,* Geophysics and Remote Sensing

Anthea Johnson, *University of Auckland, Auckland, New Zealand,* Bacteria Role in the Phytoremediation of Heavy Metals

Silvia Johnson, *Environmental Bio-Systems, Inc., Mill Valley, California,* Groundwater and Mercury: Chemical Behavior and Treatment

Tracey Johnston, *University of Texas at San Antonio, San Antonio, Texas,* Phytoextraction and Phytostabilization: Technical, Economic and Regulatory Considerations of the Soil-Lead Issue

Jagath J. Kaluarachchi, *Utah State University, Logan, Utah,* Groundwater Flow and Transport Process

A. Katsoyiannis, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* The Use of Semipermeable Membrane Devices (SPMDs) for Monitoring, Exposure, and Toxicity Assessment

Jack Keeley, *Environmental Engineer, Ada, Oklahoma,* Nitrate Contamination of Groundwater

David W. Kelley, *University of St. Thomas, St. Paul, Minnesota,* Leaching

Lisa Kirkland, *Environmental Bio-Systems, Inc., Mill Valley, California,* Groundwater and Lead: Chemical Behavior and Treatment

Dana W. Kolpin, *United States Geological Survey,* Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams

C.P. Kumar, *National Institute of Hydrology, Roorkee, India,* Groundwater Balance

Maria S. Kuyukina, *Institute of Ecology and Genetics of Microorganisms of the RAS, Perm, Russia, In Situ* Bioremediation of Contaminated Groundwater

Kung-Yao Lee, *Horn Point Laboratory—UMCES, Solomons, Maryland,* What is a Hydrochemical Model?

Leo S. Leonhart, *Hargis Associates, Inc., Tucson, Arizona,* Recharge in Arid Regions, Perched Groundwater

Scott M. Lesch, *USDA-ARS George E. Brown, Jr., Salinity Laboratory, Riverside, California,* Characterizing Soil Spatial Variability

Len Li, *University of Wisconsin-Madison, Madison, Wisconsin,* Reactive Transport in The Saturated Zone: Case Histories for Permeable Reactive Barriers

Keith Loague, *Stanford University, Stanford, California,* Groundwater Vulnerability to Pesticides: An Overview of Approaches and Methods of Evaluation, Modeling Non-Point Source Pollutants in the Vadose Zone Using GIS

Walter W. Loo, *Environmental & Technology Services, Oakland, California,* Treatment for Nitrates in Groundwater, Treatment of Arsenic, Chromium, and Biofouling in Water Supply Wells, *In Situ* Electrokinetic Treatment of MtBE, Benzene, and Chlorinated Solvents, Hydraulic Properties Characterization

Kerry T. Macquarrie, *University of New Brunswick, Fredericton, New Brunswick, Canada,* River-Connected Aquifers: Geophysics, Stratigraphy, Hydrogeology, and Geochemistry

Mini Mathew, *Colorado School of Mines, Golden, Colorado,* Modeling of DNAPL Migration in Saturated Porous Media

S.C. Mccutcheon, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Hydrologic Feasibility Assessment and Design in Phytoremediation

John E. McCray, *Colorado School of Mines, Golden, Colorado,* Groundwater Vulnerability to Pesticides: Statistical Approaches

M.S. Mohan Kumar, *Indian Institute of Science, Bangalore, India,* Modeling of DNAPL Migration in Saturated Porous Media

John E. Moore, *USGS (Retired), Denver, Colorado,* Well Hydraulics and Aquifer Tests, Drawdown, Groundwater Quality, Hot Springs, Overdraft, Saline Seep, Geological Occurrence of Groundwater

Angela Munroe, *Environmental Bio-Systems, Inc., Mill Valley, California,* Groundwater and Vinyl Chloride: Chemical Behavior and Treatment

Jean-Christophe Nadeau, *University of New Brunswick, Fredericton, New Brunswick, Canada,* River-Connected Aquifers: Geophysics, Stratigraphy, Hydrogeology, and Geochemistry

NASA Earth Science Enterprise Data and Services, Squeezing Water from Rock

Vahid Nassehi, *Loughborough University, Loughborough, United Kingdom,* Combined Free and Porous Flow in the Subsurface, Viscous Flow

Sascha E. Oswald, *UFZ Centre for Environmental Research, Leipzig-Halle, Germany,* Modeling Contaminant Transport and Biodegradation in Groundwater

Timothy K. Parker, *Groundwater Resources of California, Sacramento, California,* Water Contamination by Low Level Organic Waste Compounds in the Hydrologic System

Jim C. Philp, *Napier University, Edinburgh, Scotland, United Kingdom, In Situ* Bioremediation of Contaminated Groundwater

Laurel Phoenix, *Green Bay, Wisconsin,* Fossil Aquifers

Nitish Priyadarshi, *Ranchi University, Ranchi, Jharkhand, India,* Geothermal Water, Rock Fracture, Consolidated Water Bearing Rocks, Groundwater Contamination from Runoff, Groundwater Dating with Radiocarbon, Methane in Groundwater, Permeability

S.N. Rai, *National Geophysical Research Institute, Hyderabad, India,* Artificial Recharge of Unconfined Aquifer

Todd Rasmussen, *The University of Georgia, Athens, Georgia,* Head, Deep Soil-Water Movement, Soil Water, Specific Gravity, Tidal Efficiency

Hugh H. Russell, *CHR_2 Environmental Services, Inc.,, Oilton, Oklahoma,* Microbial Processes Affecting Monitored Natural Attenuation of Contaminants in the Subsurface

Philip R. Rykwalder, *University of Texas at San Antonio, San Antonio, Texas,* Vadose Zone Monitoring Techniques

Bahram Saghafian, *Soil Conservation and Watershed Management Research Institute, Tehran, Iran,* Qanats: An Ingenious Sustainable Groundwater Resource System

C. Samara, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* The Use of Semipermeable Membrane Devices (SPMDs) for Monitoring, Exposure, and Toxicity Assessment

Dibyendu Sarkar, *University of Texas at San Antonio, San Antonio, Texas,* Remediation of Contaminated Soils, Genetics of Metal Tolerance and Accumulation in Higher Plants, Phytoextraction of Zinc and Cadmium from Soils Using Hyperaccumulator Plants, Phytoremediation of Selenium-Laden Soils, Phytoextraction and Phytostabilization: Technical, Economic and Regulatory Considerations of the Soil-Lead Issue

J.L. Schnoor, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Phytoremediation of Methyl *Tertiary*-Butyl Ether

Guy W. Sewell, *National Risk Management Research Laboratory, ORD, U.S. EPA, Ada, Oklahoma (formerly with Dynamac Corporation),* Microbial Processes Affecting Monitored Natural Attenuation of Contaminants in the Subsurface

Raj Sharma, *University of KwaZulu-Natal, Durban, South Africa,* Laminar Flow

Caijun Shi, *CJS Technology, Inc., Burlington, Ontario, Canada,* High pH Groundwater—The Effect of The Dissolution of Hardened Cement Pastes

Naresh Singhal, *University of Auckland, Auckland, New Zealand,* Bacteria Role in the Phytoremediation of Heavy Metals, Sub-Surface Redox Chemistry: A Comparison of Equilibrium and Reaction-Based Approaches

V.P. Singh, *Louisiana State University, Baton Rouge, Louisiana,* Artificial Recharge of Unconfined Aquifer

Joseph Skopp, *University of Nebraska, Lincoln, Nebraska,* Field Capacity

Jeffrey G. Skousen, *West Virginia University, Morgantown, West Virginia,* Acid Mine Drainage: Sources and Treatment in the United States

Ricardo Smalling, *Environmental Bio-Systems, Inc., Mill Valley, California,* Groundwater and Uranium: Chemical Behavior and Treatment

James A. Smith, *University of Virginia, Charlottesville, Virginia,* Vapor Transport in the Unsaturated Zone

Stuart A. Smith, *Smith-Comeskey GroundWater Science LLC, Upper Sandusky, Ohio,* Well Maintenance, Biofouling in Water Wells, Soil and Groundwater Geochemistry and Microbiology

Michelle Sneed, *U.S. Geological Survey, Denver, Colorado,* Earthquakes—Rattling the Earth's Plumbing System

Roger Spence, *Oak Ridge National Laboratory, Oak Ridge, Tennessee,* High pH Groundwater—The Effect of The Dissolution of Hardened Cement Pastes

Kenneth F. Steele, *University of Arkansas, Fayetteville, Arkansas,* Soil and Water Contamination by Heavy Metals

Mark D. Steele, *MDC Systems, Inc., Berwyn, Pennsylvania,* Water Level Drawdown

P. Takis Elefsiniotis, *University of Auckland, Auckland, New Zealand,* Bacteria Role in the Phytoremediation of Heavy Metals

Henry Teng, *The George Washington University, Washington, DC,* Water/Rocks Interaction

Stephen M. Testa, *Mokelumne Hill, California,* Dating Groundwaters with Tritium, Brine Deposits

Geoffrey Thyne, *Colorado School of Mines, Golden, Colorado,* Detecting Modern Groundwaters with ^{85}Kr, Geochemical Modeling—Computer Code Concepts

Fred D. Tillman, *U.S. Environmental Protection Agency, Athens, Georgia,* Vapor Transport in the Unsaturated Zone

David J. Tonjes, *Cashin Associates PC, Hauppauge, New York,* Groundwater Contamination From Municipal Landfills in the USA

Douglas C. Towne, *Phoenix, Arizona,* Ambient Groundwater Monitoring Network Strategies and Design

Michael D. Trojan, *Minnesota Pollution Control Agency, St. Paul, Minnesota,* Land Use Impacts on Groundwater Quality, Sensitivity of Groundwater to Contamination

Kristine Uhlman, *University of Arizona, Tucson, Arizona,* Recharge in Desert Regions Around The World

Matthew M. Uliana, *Texas State University—San Marcos, San Marcos, Texas,* Regional Flow Systems, Hydraulic Head, Storage Coefficient

David B. Vance, *ARCADIS G&M, Inc., Midland, Texas,* Groundwater Remediation by *In Situ* Aeration and Volatilization, Source, Mobility, and Remediation of Metals, Particulate Transport in Groundwater—Bacteria and Colloids, The Environmental Impact of Iron in Groundwater, Groundwater Remediation by Injection and Problem Prevention, Chemical Oxidation Technologies for Groundwater Remediation, Physical Properties of DNAPLs and Groundwater Contamination, Process Limitations of *In Situ* Bioremediation of Groundwater, Phytoremediation Enhancement of Natural Attenuation Processes, Groundwater and Arsenic: Chemical Behavior and Treatment, Low Flow Groundwater Purging and Surging, The Role of Heat in Groundwater Systems, Horizontal Wells in Groundwater Remediation, Groundwater Flow in Heterogenetic Sediments and Fractured Rock Systems, Limiting Geochemical Factors in Remediation Using Monitored Natural Attenuation and Enhanced Bioremediation

Keith Villiers, *Environmental Bio-Systems, Inc., Mill Valley, California,* Groundwater and Benzene: Chemical Behavior and Treatment

Nikolay Voutchkov, *Poseidon Resources Corporation, Stamford, Connecticut,* Well Design and Construction

Atul N. Waghode, *Loughborough University, Leicestershire, United Kingdom,* Finite Element Modeling of Coupled Free and Porous Flow

Roger M. Waller, *U.S. Geological Survey,,* Ground Water: Wells

Lise Walter, *Environmental Bio-Systems, Mill Valley, California,* Groundwater and Cobalt: Chemical Behavior and Treatment

J.W. Weaver, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Hydrologic Feasibility Assessment and Design in Phytoremediation

Jason J. Wen, *City of Downey, Downey, California,* Treatment for Nitrates in Groundwater, Treatment of Arsenic, Chromium, and Biofouling in Water Supply Wells

Dennis E. Williams, *Geoscience Support Services, Claremont, California,* Well TEST, Radial Wells, Well Screens

Eric S. Wilson, *E. L. Montgomery & Associates, Inc., Tucson, Arizona,* Safe Yield of an Aquifer, Specific Capacity

S.K. Winnike-McMillan, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Phytoremediation of Methyl *Tertiary*-Butyl Ether

Q. Zhang, *(from Phytoremediation: Transformation and Control of Contaminants, Wiley 2003),* Phytoremediation of Methyl *Tertiary*-Butyl Ether

A.I. Zouboulis, *Aristotle University of Thessaloniki, Thessaloniki, Greece,* Phytoremediation of Lead-Contaminated Soils

Contents

Prefaces

Preface (Volume 1) xi
Preface (Volume 2) xii
Preface (Volume 3) xiii
Preface (Volume 4) xiv
Preface (Volume 5) xv

Acknowledgments xvi

Contributors

Contributors (Volume 1) xvii
Contributors (Volume 2) xxi
Contributors (Volume 3) xxv
Contributors (Volume 4) xxix
Contributors (Volume 5) xxxiii

Volume 1. Domestic, Municipal, and Industrial Water Supply and Waste Disposal

1.1 Domestic Water Supply 1:1
1.1.1 The Arsenic Drinking Water Crisis in Bangladesh 1:1
1.1.2 Bottled Water 1:3
1.1.3 Corrosion Control in Drinking Water Systems 1:5
1.1.4 Economics of Residential Water Demands 1:12
1.1.5 Gray Water Reuse in Households 1:16
1.1.6 Water and Human Health 1:19
1.1.7 Nitrate Health Effects 1:30
1.1.8 Domestic Water Supply – Public-private Partnership 1:42
1.1.9 Methods of Reducing Radon in Drinking Water 1:51
1.1.10 Water Reuse 1:53
1.1.11 Roof Drainage Hydraulics 1:54
1.1.12 Septic Tank Systems 1:61
1.1.13 Domestic Solar Water Heaters 1:63
1.1.14 Household Drinking Water Treatment and Safe Storage 1:67
1.1.15 Virus Transport in the Subsurface 1:70
1.1.16 Windmills 1:73

1.2 Municipal Water Supply 1:76
1.2.1 Mixing and Agitation in Water Treatment Systems 1:76
1.2.2 Arsenic in Natural Waters 1:81
1.2.3 Evaluation of Microbial Components of Biofouling 1:83
1.2.4 Threat Agents and Water Biosecurity 1:87
1.2.5 Granular Activated Carbon 1:92
1.2.6 Competitive Adsorption of Several Organics and Heavy Metals on Activated Carbon in Water 1:107
1.2.7 A Real-time Hydrological Information System for Cities 1:121
1.2.8 Chlorine and Chlorine Residuals 1:127
1.2.9 Modeling Chlorine Residuals in Urban Water Distribution Systems 1:131
1.2.10 Particulate Matter Removal by Coagulation 1:137
1.2.11 Selective Coagulant Recovery from Water Treatment Plant Residuals Using the Domain Membrane Process 1:139
1.2.12 Physical Water Conditioning 1:141
1.2.13 Consumer Confidence Reports 1:145
1.2.14 Water Conservation Measures 1:146
1.2.15 Preventing Well Contamination 1:149
1.2.16 Corrosion Control 1:152
1.2.17 Cross Connection and Backflow Prevention 1:155
1.2.18 Molecular-based Detection of *Cryptosporidium Parvum* in Water 1:158
1.2.19 Cryptosporidium 1:162
1.2.20 Measuring *Cryptosporidium Parvum* Oocyst Inactivation Following Disinfection with Ultraviolet Light 1:165
1.2.21 Dechlorination 1:169
1.2.22 Desalination 1:170
1.2.23 Diatomaceous Earth Filtration for Drinking Water 1:174
1.2.24 Emerging Waterborne Infectious Diseases 1:177
1.2.25 Improving Waterborne Disease Surveillance 1:183
1.2.26 Disinfectants 1:192
1.2.27 Disinfection 1:196
1.2.28 Water Distribution System Operation 1:200
1.2.29 Water Quality in Distribution Systems 1:204
1.2.30 Design of Water Distribution Systems 1:207
1.2.31 What Is in Our Drinking Water? 1:213
1.2.32 The Economics of Water Resources Allocation 1:215
1.2.33 Answering the Challenge 1:218
1.2.34 Key Causes of Drinking Water Quality Failure in a Rural Small Water Supply of South Africa 1:221
1.2.35 Filtration 1:227

1.2.36 Water Filtration 1:230
1.2.37 Filtration with Granular Media 1:233
1.2.38 Slow Sand Filtration and the Impact of *Schmutzdecke* *1:2*35
1.2.39 Multistage Drinking Water Filtration 1:237
1.2.40 Multistage Filtration: an Innovative Water Treatment Technology 1:238
1.2.41 Particulate Matter Removal by Filtration and Sedimentation 1:243
1.2.42 Filtration Water Treatment 1:245
1.2.43 Synthetic and Natural Organic Removal by Biological Filtration 1:248
1.2.44 Granular Bed and Precoat Filtration 1:249
1.2.45 Flocculation 1:252
1.2.46 Fluoridation 1:254
1.2.47 Giardiasis 1:257
1.2.48 Gravity Separation/Sedimentation 1:259
1.2.49 Water Hammer 1:261
1.2.50 Health Effects of Commonly Occurring Disinfection Byproducts in Municipal Water Supplies 1:264
1.2.51 Health Effects of Microbial Contaminants and Biotoxins in Drinking Water 1:277
1.2.52 Drinking Water and Public Health Protection 1:281
1.2.53 1962 U.S. Public Health Service Standards 1:292
1.2.54 Ion Exchange and Demineralization 1:297
1.2.55 The State of the Water Industry – 2004 .... 1:301
1.2.56 Iron and Manganese Removal 1:312
1.2.57 Extraterritorial Land Use Control to Protect Water Supplies 1:315
1.2.58 Leak Detection and Water Loss Control 1:317
1.2.59 Lime-soda Ash Processes 1:320
1.2.60 Lime Softening 1:322
1.2.61 Ion Exchange – Use of Magnetic Ion Exchange Resin for DOC Removal 1:325
1.2.62 Membrane Filtration 1:331
1.2.63 Water Meters 1:337
1.2.64 Microbiological Concerns of Drinking Water Distribution Systems 1:341
1.2.65 Nitrification of Potable Water Using Trickling Filters 1:346
1.2.66 Organic Removal 1:350
1.2.67 Ozone 1:354
1.2.68 Ozone with Activated Carbon for Drinking Water Treatment 1:357
1.2.69 Ozone-bromide Interactions 1:357
1.2.70 Municipal Water Supply: Ozonation 1:362
1.2.71 Review of Parasite Fate and Transport in Karstic Aquifers 1:365
1.2.72 Particulate Removal 1:370
1.2.73 Pharmaceuticals in Water Systems 1:372
1.2.74 Point-of-use/Point-of-entry Systems (POU/POE) 1:378

1.2.75 Assessing the Bactericidal Efficiency of Polydex for the Disinfection of Drinking Water in Rural Areas of South Africa 1:382
1.2.76 Private Sector Participation, Marketing and Corporate Strategies in Municipal Water Supply and Sewerage 1:387
1.2.77 Pumps 1:391
1.2.78 Radionuclides 1:395
1.2.79 Use of Redox Potentials in Wastewater Treatment 1:399
1.2.80 Repairing Distribution Line Breaks 1:400
1.2.81 Role of Small Water Reservoirs in Environment 1:403
1.2.82 Reservoirs, Towers, and Tanks Drinking Water Storage Facilities 1:408
1.2.83 Water Treatment Plant Residuals Management 1:411
1.2.84 Reverse Osmosis, Process Chemistry 1:414
1.2.85 Reverse Osmosis, Membrane Foulants 1:416
1.2.86 Reverse Osmosis, Membrane Cleaning 1:419
1.2.87 Application of Risk Assessments in Crafting Drinking Water Regulations 1:422
1.2.88 Potential Risks of Waterborne Transmission of *Escherichia Coli* O157:H7 .... 1:429
1.2.89 Slow Sand Filtration 1:431
1.2.90 Approaches for Securing a Water Distribution System 1:434
1.2.91 Water Security: an Emerging Issue 1:437
1.2.92 Guide to Selection of Water Treatment Processes 1:439
1.2.93 Source Water Assessment 1:444
1.2.94 Hydraulic Design of Water Distribution Storage Tanks 1:448
1.2.95 System Control and Data Acquisition (SCADA) 1:449
1.2.96 Settling Tanks 1:452
1.2.97 Treatment for Technologies for Small Drinking Water Systems 1:457
1.2.98 Ultraviolet Disinfection 1:466
1.2.99 Ultraviolet Irradiation 1:469
1.2.100 Water Disinfection Using UV Radiation – a Sri Lankan Experience 1:471
1.2.101 Drinking Water Quality Standards (DWQS) – United States 1:476
1.2.102 Valves 1:482
1.2.103 Removal of Pathogenic Bacteria, Viruses, and Protozoa 1:485
1.2.104 Water Meter 1:489
1.2.105 Municipal Watersheds 1:495
1.2.106 Public Water Supply World 1:500
1.2.107 Ten Key Trends That Will Shape the Future of the World Water Industry 1:508
1.2.108 Zebra Mussel Control without Chemicals 1:510
1.2.109 Package Plants 1:514
1.2.110 Anaerobic Sewage Treatment 1:517
1.2.111 Persistence of Pathogens in Water 1:521
1.2.112 Well Head Protection 1:524

1.2.113 Chemical Drinking Water Standards, Past, Present, and Future 1:529
1.3 Industrial Water 1:534
1.3.1 Magnetic Water Conditioning 1:534
1.3.2 Water Impacts from Construction Sites 1:537
1.3.3 Industrial Cooling Water – Biofouling 1:538
1.3.4 Industrial Cooling Water – Corrosion 1:542
1.3.5 Industrial Cooling Water – Scale Formation 1:545
1.3.6 Economics of Industrial Water Demands 1:549
1.3.7 Electric Generating Plants – Effects of Contaminants 1:553
1.3.8 Energy Dissipation 1:558
1.3.9 Water Use in Energy Production 1:560
1.3.10 Evaluation of Toxic Properties of Industrial Effluents by on-line Respirometry 1:565
1.3.11 Polycyclic Aromatic Hydrocarbons 1:571
1.3.12 Hydrocarbon Treatment Techniques 1:575
1.3.13 Use of Anaerobic-aerobic Treatment Systems for Maize Processing Plants 1:581
1.3.14 Bonding of Toxic Metal Ions 1:586
1.3.15 Application of Microfiltration to Industrial Wastewaters 1:591
1.3.16 Water Treatment in Spent Nuclear Fuel Storage 1:595
1.3.17 Industrial Mine Use: Mine Waste 1:609
1.3.18 Sugarcane Industry Wastewaters Treatment 1:614
1.3.19 Estimated Use of Water in the United States in 1990 Industrial Water Use 1:620
1.4 Waste Water Treatment 1:623
1.4.1 Aeration 1:623
1.4.2 Fine Bubble Diffused Air Aeration Systems 1:626
1.4.3 Air Stripping 1:631
1.4.4 Land Applications of Wastewater in Arid Lands: Theory and Case Studies 1:632
1.4.5 Technologies for Arsenic Removal from Contaminated Water Sources 1:636
1.4.6 Biochemical Oxygen Demand 1:639
1.4.7 Molecular Biology Tools for Monitoring Biodiversity in Wastewater Treatment Plants 1:642
1.4.8 Biosolids 1:646
1.4.9 Integrated Capacity Building Needs for Water Supply and Wastewater Sanitation 1:651
1.4.10 Wastewater Characterization 1:656
1.4.11 Chemically Enhanced Primary Treatment of Wastewater 1:659
1.4.12 Getting Our Clean Water Act Together 1:660
1.4.13 Inadequate Treatment of Wastewater: a Source of Coliform Bacteria in Receiving Surface Water Bodies in Developing Countries – Case Study: Eastern Cape Province of South Africa 1:661

1.4.14 Denitrification in the Activated Sludge Process 1:667
1.4.15 Detergents 1:669
1.4.16 Ecological Wastewater Management 1:675
1.4.17 Waste Treatment in Fish Farms 1:681
1.4.18 Flotation as a Separation Process 1:684
1.4.19 Degradation of Chloro-organics and Hydrocarbons 1:688
1.4.20 Landfill 1:695
1.4.21 Landfill Leachates, Part 1: Origin and Characterization 1:699
1.4.22 Landfill Leachates: Part 2: Treatment 1:702
1.4.23 Macrophytes as Biomonitors of Polychlorinated Biphenyls 1:714
1.4.24 Wastewater Management for Developing Countries 1:718
1.4.25 Mercury Removal from Complex Waste Waters 1:722
1.4.26 Metal Speciation and Mobility as Influenced by Landfill Disposal Practices 1:723
1.4.27 Microbial Foaming in the Activated Sludge Process 1:728
1.4.28 Introduction to Wastewater Modeling and Treatment Plant Design 1:730
1.4.29 Practical Applications of Wastewater Modeling and Treatment Plant Design 1:738
1.4.30 New York City Harbor Survey 1:745
1.4.31 Nitrification in the Activated Sludge Process 1:751
1.4.32 Effluent Limitations and the NPDES Permit 1:755
1.4.33 Odor Abatement in Wastewater Treatment Plants 1:760
1.4.34 Aqueous Reactions of Specific Organic Compounds with Ozone 1:765
1.4.35 The Fate of Persistent Organic Pollutants (POPs) in the Wastewater Treatment Process 1:766
1.4.36 The Role of Organoclay in Water Cleanup 1:771
1.4.37 Combined Sewer Overflow Treatment 1:782
1.4.38 Biological Phosphorus Removal in the Activated Sludge Process 1:788
1.4.39 Photocatalytic Membrane Reactors in Water Purification 1:791
1.4.40 EPA's National Pretreatment Program, 1973–2003: Thirty Years of Protecting the Environment 1:798
1.4.41 Problems Encountered during Pipe Repair and Renewal 1:801
1.4.42 Radioactive Waste 1:802
1.4.43 Reclaimed Water 1:805
1.4.44 Wastewater Treatment and Recycling Technologies 1:808
1.4.45 Wastewater Treatment Processes and Water Reuse 1:814
1.4.46 Wastewater Reclamation and Reuse Research 1:819
1.4.47 Wastewater Reclamation and Reuse 1:825
1.4.48 Wastewater Reclamation and Reuse Treatment Technology 1:826
1.4.49 Sewage 1:828
1.4.50 Domestic Sewage 1:830
1.4.51 Solidification/Stabilization of Hazardous Solid Wastes 1:835

1.4.52 Wastewater Treatment – Small Scale 1:840
1.4.53 Microbial Foaming and Bulking in Activated Sludge Plants 1:844
1.4.54 Aqueous Behavior of Elements in a Flue Gas Desulfurization Sludge Disposal Site 1:848
1.4.55 Sludge Treatment and Disposal 1:853
1.4.56 Wastewater Sludge 1:861
1.4.57 Processing of Sludge 1:864
1.4.58 Municipal Storm Water Management 1:866
1.4.59 What Wastewater Utilities Can Do Now to Guard Against Terrorist and Security Threats 1:870
1.4.60 Wastewater Treatment Techniques – Advanced 1:871
1.4.61 Trenchless Repair and Rehabilitation Techniques 1:876
1.4.62 Basics of Underground Water and Sewer Pipeline Assessment, Repair, and Rehabilitation 1:883
1.4.63 Water Hammer: Quantitative Causes and Effects 1:891
1.4.64 Constructed Wetlands 1:892
1.4.65 Using Ecosystem Processes in a Constructed Wetland to Treat Mine Wastewater in Ireland 1:897
1.4.66 Water and Wastewater Properties and Characteristics 1:900
1.4.67 Anaerobic Wastewater Treatment 1:904
1.4.68 Sewerage Odors – How to Control 1:910
1.4.69 Ultrafiltration – Complexation in Wastewater Treatment 1:916

Volume 2. Water Quality and Resource Development

2.5 Water Quality Control 2:1
2.5.1 Acid Mine Drainage – Extent and Character 2:1
2.5.2 The Control of Algal Populations in Eutrophic Water Bodies 2:2
2.5.3 Arsenic Compounds in Water 2:7
2.5.4 Arsenic Health Effects 2:15
2.5.5 Background Concentration of Pollutants 2:18
2.5.6 Waterborne Bacteria 2:20
2.5.7 Water Assessment and Criteria 2:24
2.5.8 Physiological Biomarkers and the Trondheim Biomonitoring System 2:28
2.5.9 Biomarkers, Bioindicators, and the Trondheim Biomonitoring System 2:29
2.5.10 Active Biomonitoring (ABM) by Translocation of Bivalve Molluscs 2:33
2.5.11 Biochemical Oxygen Demand and Other Organic Pollution Measures 2:37
2.5.12 Biodegradation 2:41
2.5.13 Bioluminescent Biosensors for Toxicity Testing 2:45
2.5.14 Biomanipulation 2:50
2.5.15 Genomic Technologies in Biomonitoring 2:58

2.5.16 Macrophytes as Biomonitors of Trace Metals 2:64
2.5.17 Biosorption of Toxic Metals 2:68
2.5.18 Bromide Influence on Trihalomethane and Haloacetic Acid Formation 2:74
2.5.19 Activated Carbon: Ion Exchange and Adsorption Properties 2:79
2.5.20 Activated Carbon – Powdered 2:86
2.5.21 Chlorination 2:88
2.5.22 Chlorination by-products 2:91
2.5.23 Classification and Environmental Quality Assessment in Aquatic Environments 2:94
2.5.24 Coagulation and Flocculation in Practice 2:98
2.5.25 Colloids and Dissolved Organics: Role in Membrane and Depth Filtration 2:99
2.5.26 Column Experiments in Saturated Porous Media Studying Contaminant Transport 2:103
2.5.27 Cytochrome P450 Monooxygenase as an Indicator of PCB/Dioxin-like Compounds in Fish 2:106
2.5.28 Water Related Diseases 2:111
2.5.29 Dishwashing Water Quality Properties 2:112
2.5.30 Disinfection by-product Precursor Removal from Natural Waters 2:115
2.5.31 Alternative Disinfection Practices and Future Directions for Disinfection by-product Minimization 2:118
2.5.32 Water Quality Aspects of Dredged Sediment Management 2:122
2.5.33 The Economics of Water Quality 2:127
2.5.34 Understanding *Escherichia Coli* O157:H7 and the Need for Rapid Detection in Water 2:136
2.5.35 Eutrophication and Organic Loading 2:142
2.5.36 Trace Element Contamination in Groundwater of District Hardwar, Uttaranchal, India 2:143
2.5.37 Iron Bacteria 2:149
2.5.38 Cartridge Filters for Iron Removal 2:152
2.5.39 Irrigation Water Quality in Areas Adjoining River Yamuna at Delhi, India 2:155
2.5.40 Water Sampling and Laboratory Safety 2:161
2.5.41 Municipal Solid Waste Landfills – Water Quality Issues 2:163
2.5.42 Land Use Effects on Water Quality 2:169
2.5.43 Monitoring Lipophilic Contaminants in the Aquatic Environment Using the SPMD-TOX Paradigm 2:170
2.5.44 Use of Luminescent Bacteria and the *Lux* Genes for Determination of Water Quality 2:172
2.5.45 Water Quality Management 2:176
2.5.46 Water Quality Management and Nonpoint Source Control 2:184
2.5.47 Water Quality Management in an Urban Landscape 2:189

2.5.48 Water Quality Management in the U.S.: History of Water Regulation 2:193
2.5.49 Water Quality Management in a Forested Landscape 2:199
2.5.50 Trace Metal Speciation 2:202
2.5.51 Metal Ion Humic Colloid Interaction 2:205
2.5.52 Heavy Metal Uptake Rates Among Sediment Dwelling Organisms 2:211
2.5.53 Methemoglobinemia 2:219
2.5.54 Microbial Activities Management 2:223
2.5.55 Microbial Dynamics of Biofilms 2:228
2.5.56 Microbial Enzyme Assays for Detecting Heavy Metal Toxicity 2:233
2.5.57 Microbial Forms in Biofouling Events 2:239
2.5.58 Microbiological Quality Control in Distribution Systems 2:243
2.5.59 Water Quality Models for Developing Soil Management Practices 2:248
2.5.60 Water Quality Modeling – Case Studies 2:255
2.5.61 Field Sampling and Monitoring of Contaminants 2:263
2.5.62 Water Quality Models: Chemical Principles 2:269
2.5.63 Water Quality Models: Mathematical Framework 2:273
2.5.64 Environmental Applications with Submitochondrial Particles 2:278
2.5.65 Interest in the Use of an Electronic Nose for Field Monitoring of Odors in the Environment 2:281
2.5.66 Oil-field Brine 2:284
2.5.67 Oil Pollution 2:290
2.5.68 Indicator Organisms 2:292
2.5.69 pH 2:294
2.5.70 Perchloroethylene (PCE) Removal 2:299
2.5.71 A Primer on Water Quality 2:301
2.5.72 Overview of Analytical Methods of Water Analyses with Specific Reference to EPA Methods for Priority Pollutant Analysis 2:304
2.5.73 Source-water Protection 2:311
2.5.74 Protozoa in Water 2:313
2.5.75 Water Quality 2:314
2.5.76 Water Quality 2:316
2.5.77 Emerging and Recalcitrant Compounds in Groundwater 2:316
2.5.78 Road Salt 2:319
2.5.79 Review of River Water Quality Modeling Software Tools 2:325
2.5.80 River Water Quality Calibration 2:331
2.5.81 *Salmonella*: Monitoring and Detection in Drinking Water 2:337
2.5.82 Lysimeter Soil Water Sampling 2:340
2.5.83 Regulatory and Security Requirements for Potable Water 2:343
2.5.84 A Weight of Evidence Approach to Characterize Sediment Quality Using Laboratory and Field Assays: an Example for Spanish Coasts 2:350

2.5.85 Remediation and Bioremediation of Selenium-contaminated Waters 2:355
2.5.86 Shellfish Growing Water Classification 2:360
2.5.87 Sorptive Filtration 2:362
2.5.88 Quality of Water in Storage 2:367
2.5.89 Quality of Water Supplies 2:370
2.5.90 The Submitochondrial Particle Assay as a Biological Monitoring Tool 2:376
2.5.91 Microscale Test Relationships to Responses to Toxicants in Natural Systems 2:379
2.5.92 Toxicity Identification Evaluation 2:380
2.5.93 Whole Effluent Toxicity Controls 2:382
2.5.94 Development and Application of Sediment Toxicity Tests for Regulatory Purposes 2:383
2.5.95 Algal Toxins in Water 2:387
2.5.96 Ground Water Quality in Areas Adjoining River Yamuna at Delhi, India 2:392
2.5.97 Chlorine Residual 2:398
2.5.98 Source Water Quality Management 2:399
2.5.99 Dose-response of Mussels to Chlorine 2:401
2.5.100 Metallothioneins as Indicators of Trace Metal Pollution 2:406
2.5.101 Amphipod Sediment Toxicity Tests 2:408
2.5.102 Ciliated Protists as Test Organisms in Toxicity Assessment 2:413
2.5.103 SOFIE: an Optimized Approach for Exposure Tests and Sediment Assays 2:418
2.5.104 Passive Treatment of Acid Mine Drainage (Wetlands) 2:423
2.5.105 Biomarkers and Bioaccumulation: Two Lines of Evidence to Assess Sediment Quality 2:426
2.5.106 Lead and Its Health Effects 2:432
2.5.107 Microbial Detection of Various Pollutants as an Early Warning System for Monitoring of Water Quality and Ecological Integrity of Natural Resources, in Russia 2:440
2.5.108 Luminescent Bacterial Biosensors for the Rapid Detection of Toxicants 2:453
2.5.109 Development and Application of Sediment Toxicity Test for Regulatory Purposes 2:458
2.5.110 E_h 2:464

2.6 Water Resource Development and Management 2:470
2.6.1 Water Resources Challenges in the Arab World 2:470
2.6.2 Effluent Water Regulations in Arid Lands 2:475
2.6.3 California – Continually the Nation's Leader in Water Use 2:478
2.6.4 Lessons from the Rising Caspian 2:480
2.6.5 Institutional Aspects of Water Management in China 2:484
2.6.6 Will Water Scarcity Limit China's Agricultural Potential? 2:488
2.6.7 Water and Coastal Resources 2:489

2.6.8 Water Use Conservation and Efficiency 2:489
2.6.9 Conservation of Water 2:495
2.6.10 The Development of American Water Resources: Planners, Politicians, and Constitutional Interpretation 2:498
2.6.11 Water Markets: Transaction Costs and Institutional Options 2:499
2.6.12 Averting Water Disputes 2:501
2.6.13 Water Supply and Water Resources: Distribution System Research 2:509
2.6.14 Drought in the Dust Bowl Years 2:511
2.6.15 Drought Management Planning 2:514
2.6.16 Drought and Water Supply Management 2:515
2.6.17 Assessment of Ecological Effects in Water-limited Environments 2:516
2.6.18 Reaching Out: Public Education and Community Involvement in Groundwater Protection 2:518
2.6.19 Integration of Environmental Impacts into Water Resources Planning 2:520
2.6.20 The Expansion of Federal Water Projects 2:522
2.6.21 Flood Control History in the Netherlands 2:524
2.6.22 Food and Water in an Emergency 2:526
2.6.23 Water Demand Forecasting 2:529
2.6.24 Remote Sensing and GIS Application in Water Resources 2:531
2.6.25 Globalization of Water 2:536
2.6.26 Water Science Glossary of Terms 2:541
2.6.27 Harvesting Rainwater 2:548
2.6.28 Urban Water Resource and Management in Asia: Ho Chi Minh City 2:552
2.6.29 Hydropower – Energy from Moving Water 2:554
2.6.30 Water Markets in India: Economic and Institutional Aspects 2:555
2.6.31 Water Resources of India 2:559
2.6.32 Water Infrastructure and Systems 2:567
2.6.33 Overview and Trends in the International Water Market 2:568
2.6.34 Best Management Practices for Water Resources 2:570
2.6.35 Integrated Water Resources Management (IWRM) 2:574
2.6.36 Management of Water Resources for Drought Conditions 2:576
2.6.37 Water Resources Management 2:586
2.6.38 NASA Helping to Understand Water Flow in the West 2:587
2.6.39 Transboundary Water Conflicts in the Nile Basin 2:590
2.6.40 Planning and Managing Water Infrastructure 2:594
2.6.41 Application of the Precautionary Principle to Water Science 2:595
2.6.42 Water Pricing 2:603
2.6.43 Spot Prices, Option Prices, and Water Markets 2:606
2.6.44 Water Managed in the Public Trust 2:608
2.6.45 Water Recycling and Reuse: the Environmental Benefits 2:610

2.6.46 State and Regional Water Supply 2:613
2.6.47 River Basin Decisions Support Systems 2:619
2.6.48 Water Resource Sustainability: Concepts and Practices 2:624
2.6.49 The Provision of Drinking Water and Sanitation in Developing Countries 2:630
2.6.50 Sustainable Management of Natural Resources 2:633
2.6.51 Sustainable Water Management on Mediterranean Islands: Research and Education 2:638
2.6.52 Meeting Water Needs in Developing Countries with Tradable Rights 2:643
2.6.53 Water Use in the United States 2:645
2.6.54 How We Use Water in These United States 2:650
2.6.55 Valuing Water Resources 2:653
2.6.56 Water – Here, There, and Everywhere in Canada 2:656
2.6.57 Water Conservation – Every Drop Counts in Canada 2:660
2.6.58 Ecoregions: a Spatial Framework for Environmental Management 2:667
2.6.59 Flood of Portals on Water 2:668
2.6.60 Fuzzy Criteria for Water Resources Systems Performance Evaluation 2:674
2.6.61 Participatory Multicriteria Flood Management 2:678
2.6.62 Water Resources Systems Analysis 2:683

Volume 3. Surface and Agricultural Water

3.7 Surface Water Hydrology 3:1
3.7.1 Acidification – Chronic 3:1
3.7.2 Episodic Acidification 3:5
3.7.3 Acidification of Freshwater Resources 3:7
3.7.4 Geochemistry of Acid Mine Drainage 3:13
3.7.5 The Aral Sea Disaster: Environment Issues and Nationalist Tensions 3:15
3.7.6 Lake Baikal – a Touchstone for Global Change and Rift Studies 3:20
3.7.7 Base Flow 3:22
3.7.8 River Basins 3:28
3.7.9 River Basin Planning and Coordination 3:33
3.7.10 Bioaccumulation 3:34
3.7.11 Biotic Integrity Index to Evaluate Water Resource Integrity in Freshwater Systems 3:36
3.7.12 Reversal of the Chicago River 3:41
3.7.13 Flood Control in the Yellow River Basin in China 3:45
3.7.14 Chironomids in Sediment Toxicity Testing 3:50
3.7.15 Cienega 3:57
3.7.16 Time-area and the Clark Rainfall-runoff Transformation 3:60
3.7.17 Stream Classification 3:65
3.7.18 Coastal Wetlands 3:71

3.7.19 Freshwater Colloids 3:74
3.7.20 Culvert Design 3:75
3.7.21 Dilution-mixing Zones and Design Flows 3:78
3.7.22 Drainage Ditches 3:87
3.7.23 Drainage Networks 3:93
3.7.24 Dyes as Hydrological Tracers 3:95
3.7.25 Flow-duration Curves 3:102
3.7.26 Environmental Flows 3:106
3.7.27 Eutrophication (Excessive Fertilization) 3:107
3.7.28 Cultural Eutrophication 3:114
3.7.29 Fish Cells in the Toxicological Evaluation of Environmental Contaminants 3:115
3.7.30 Fish Consumption Advisories 3:118
3.7.31 Fisheries: History, Science, and Management 3:121
3.7.32 Factors Affecting Fish Growth and Production 3:129
3.7.33 Water Needs for Freshwater Fisheries Management 3:133
3.7.34 An Outline of the History of Fishpond Culture in Silesia, the Western Part of Poland 3:135
3.7.35 Floods 3:142
3.7.36 Flood Control Structures 3:150
3.7.37 Floods as a Natural Hazard 3:153
3.7.38 Flood Source Mapping in Watersheds 3:155
3.7.39 Urban Flooding 3:159
3.7.40 Floodwater Spreading 3:163
3.7.41 Minimum Environmental Flow Regimes 3:166
3.7.42 Forensic Hydrogeology 3:168
3.7.43 Forests and Wetlands 3:170
3.7.44 Rock Glacier 3:174
3.7.45 Great Lakes 3:175
3.7.46 Greenhouse Gas Emissions from Hydroelectric Reservoirs 3:180
3.7.47 Gully Erosion 3:183
3.7.48 Potential Health Issues Associated with Blue-green Algae Blooms in Impoundments, Ponds and Lakes 3:188
3.7.49 Heat Balance of Open Waterbodies 3:190
3.7.50 Hydraulics 3:194
3.7.51 Hydraulics of Pressurized Flow 3:196
3.7.52 Hydroelectric Power 3:199
3.7.53 Hydroelectric Reservoirs as Anthropogenic Sources of Greenhouse Gases 3:203
3.7.54 Hydrologic Persistence and the Hurst Phenomenon 3:210
3.7.55 Unit Hydrograph 3:221
3.7.56 Hydrological Processes and Measured Pollutant Loads 3:222

3.7.57 Hydrologic Thresholds 3:229
3.7.58 Ganga River, India 3:232
3.7.59 Interception 3:235
3.7.60 Kinematic Shock 3:239
3.7.61 Kinematic Wave Method for Storm Drainage Design 3:242
3.7.62 Kinematic Wave and Diffusion Wave Theories 3:246
3.7.63 Kinematic Wave Flow Routing 3:253
3.7.64 Reliability Concepts in Reservoir Design 3:259
3.7.65 Lakes 3:265
3.7.66 The Theory of Alternative Stable States in Shallow Lake Ecosystems 3:272
3.7.67 NOAA Lake Level Forecast for Lake Michigan Right on Target 3:274
3.7.68 Submerged Aquatic Plants Affect Water Quality in Lakes 3:275
3.7.69 Lakes – Discharges to 3:281
3.7.70 Lasers Scan Levees from the Air 3:284
3.7.71 Levees for Flood Protection 3:286
3.7.72 Limnology 3:291
3.7.73 Adsorption of Metal Ions on Bed Sediments 3:295
3.7.74 Microbiology of Lotic Aggregates and Biofilms 3:305
3.7.75 Microorganisms in Their Natural Environment 3:309
3.7.76 Calibration of Hydraulic Network Models 3:313
3.7.77 Numerical Modeling of Currents 3:320
3.7.78 Uncertainty Analysis in Watershed Modeling 3:325
3.7.79 Watershed Modeling 3:327
3.7.80 Modeling of Water Quality in Sewers 3:331
3.7.81 Modeling of Urban Drainage and Stormwater 3:337
3.7.82 Modeling Ungauged Watersheds 3:342
3.7.83 Corps Turned Niagara Falls Off, on Again 3:345
3.7.84 Open Channel Design 3:346
3.7.85 Organic Compounds and Trace Elements in Freshwater Streambed Sediment and Fish from the Puget Sound Basin 3:349
3.7.86 Impervious Cover – Paving Paradise 3:363
3.7.87 Phytoremediation by Constructed Wetlands 3:364
3.7.88 Unrecognized Pollutants 3:371
3.7.89 Pollution of Surface Waters 3:373
3.7.90 Pond Aquaculture – Modeling and Decision Support Systems 3:375
3.7.91 Pumping Stations 3:379
3.7.92 Regulated Rivers 3:381
3.7.93 Reservoirs-multipurpose 3:382
3.7.94 Dam Removal as River Restoration 3:387
3.7.95 Riparian Systems 3:390

3.7.96 Rivers 3:392
3.7.97 River and Water Facts 3:394
3.7.98 Sediment Load Measurements 3:397
3.7.99 Sedimentation 3:401
3.7.100 Sedimentation and Flotation 3:404
3.7.101 Reservoir Sedimentation 3:408
3.7.102 Water from Saturated River Sediment – Sand Abstraction 3:412
3.7.103 Sediment Transport 3:417
3.7.104 Stochastic Simulation of Hydrosystems 3:421
3.7.105 Storage and Detention Facilities 3:430
3.7.106 Urban Stormwater Runoff Water Quality Issues 3:432
3.7.107 Rivers and Streams: One-way Flow System 3:437
3.7.108 Streamflow 3:439
3.7.109 Water Quality in Suburban Watersheds 3:441
3.7.110 Surface Water Pollution 3:444
3.7.111 Surface Runoff and Subsurface Drainage 3:451
3.7.112 Trace Elements in Water, Sediment, and Aquatic Biota – Effects of Biology and Human Activity 3:454
3.7.113 Innovative Pens Hatch Thousands of Trout 3:458
3.7.114 Watershed 3:460
3.7.115 Combustible Watersheds 3:461
3.7.116 Time of Concentration and Travel Time in Watersheds 3:469
3.7.117 Watershed Hydrology 3:472
3.7.118 Watershed Management for Environmental Quality and Food Security 3:479
3.7.119 Water Hyacinth – The World's Most Problematic Weed 3:479
3.7.120 Water Quality in Ponds 3:484
3.7.121 Water Turbine 3:487
3.7.122 Wetlands: Uses, Functions, and Values 3:489
3.7.123 Wetlands Overview 3:493
3.7.124 Classification of Wetlands and Deepwater Habitats of the United States 3:496
3.7.125 Urban Runoff 3:498
3.7.126 Urban Water Studies 3:501
3.7.127 Subglacial Lake Vostok 3:503
3.7.128 Water – the Canadian Transporter 3:507
3.7.129 Flood Prevention 3:510
3.7.130 Effects of DDT in Surface Water on Bird Abundance and Reproduction – a History 3:513
3.7.131 Instream Flow Methods 3:526
3.7.132 Floodplain 3:527
3.7.133 Fish Passage Facilities 3:529

3.7.134 Fishing Waters 3:532
3.7.135 Land Surface Modeling 3:533
3.8 Agricultural Water 3:538
3.8.1 Animal Farming Operations: Groundwater Quality Issues 3:538
3.8.2 Aquaculture Technology for Producers 3:540
3.8.3 Biofuel Alternatives to Fossil Fuels 3:545
3.8.4 Soil Conservation 3:549
3.8.5 Landscape Water-conservation Techniques 3:553
3.8.6 Crop Water Requirements 3:557
3.8.7 Agricultural Water Use Efficiency (WUE) and Productivity (WP) 3:558
3.8.8 Large Area Surface Energy Balance Estimation Using Satellite Imagery 3:560
3.8.9 Soil Erosion and Control Practices 3:565
3.8.10 Water Table Contribution to Crop Evapotranspiration 3:570
3.8.11 Crop Evapotranspiration 3:571
3.8.12 Water Pollution from Fish Farms 3:579
3.8.13 World's Major Irrigation Areas 3:581
3.8.14 Irrigation in the United States 3:586
3.8.15 Irrigation Wells 3:594
3.8.16 Agriculture and Land Use Planning 3:595
3.8.17 Waterlogging 3:599
3.8.18 Water Quality Management in an Agricultural Landscape 3:604
3.8.19 Classification and Mapping of Agricultural Land for National Water-quality Assessment 3:608
3.8.20 Metal Tolerance in Plants: the Roles of Thiol-containing Peptides 3:609
3.8.21 Microirrigation 3:615
3.8.22 Microirrigation: an Approach to Efficient Irrigation 3:620
3.8.23 Plant and Microorganism Selection for Phytoremediation of Hydrocarbons and Metals 3:628
3.8.24 Nitrate Pollution Prevention 3:637
3.8.25 Nitrification 3:640
3.8.26 Occurrence of Organochlorine Pesticides in Vegetables Grown on Untreated Soils from an Agricultural Watershed 3:643
3.8.27 Pesticide Chemistry in the Environment 3:647
3.8.28 Remediation of Pesticide-contaminated Soil at Agrichemical Facilities 3:651
3.8.29 Pesticide Occurrence and Distribution in Relation to Use 3:655
3.8.30 Assessment of Pollution Outflow from Large Agricultural Areas 3:657
3.8.31 Deep-well Turbine Pumps 3:664
3.8.32 Microbial Quality of Reclaimed Irrigation: International Perspective 3:667
3.8.33 Soil Salinity 3:673
3.8.34 Maintaining Salt Balance on Irrigated Land 3:677

3.8.35 Salt Tolerance 3:681
3.8.36 Groundwater Assessment Using Soil Sampling Techniques 3:688
3.8.37 Skimmed Groundwater 3:691
3.8.38 Soil Moisture Measurement – Neutron 3:692
3.8.39 Soil N Management Impact on the Quality of Surface and Subsurface Water 3:694
3.8.40 Soil Phosphorus Availability and Its Impact on Surface Water Quality 3:701
3.8.41 Soil Water Issues 3:706
3.8.42 Water Spreading 3:708
3.8.43 Sprinkler Irrigation 3:712
3.8.44 Stomates 3:714
3.8.45 Crop Water Stress Detection Using Remote Sensing 3:719
3.8.46 Vacuum Gauge Tensiometer 3:724
3.8.47 Tile Drainage 3:729
3.8.48 Tile Drainage: Impacts, Plant Growth, and Water Table Levels 3:731
3.8.49 Measuring and Modeling Tree and Stand Level Transpiration 3:732
3.8.50 Water Logging: Topographic and Agricultural Impacts 3:741
3.8.51 Weed Control Strategies 3:742
3.8.52 Screen Filters for Microirrigation 3:748
3.8.53 Xeriscape 3:750
3.8.54 Media Filters for Microirrigation 3:752

Volume 4. Oceanography; Meteorology; Physics and Chemistry; Water Law; and Water

4.9 Oceanography 4:1
4.9.1 Air-sea Interaction 4:1
4.9.2 NOAA’s Atlantic Oceanographic and Meteorological Laboratory 4:4
4.9.3 Laboratory Experiments on Bivalve Excretion Rates of Nutrients 4:6
4.9.4 Temporal Scaling of Benthic Nutrient Regeneration in Bivalve-dominated Tidal Flat 4:11
4.9.5 Breakwaters 4:14
4.9.6 The Ocean in Climate 4:21
4.9.7 Coastal Waters 4:23
4.9.8 Marine Colloids 4:27
4.9.9 Deep Water Corals 4:32
4.9.10 Marine Debris Abatement 4:38
4.9.11 Larvae and Small Species of Polychaetes in Marine Toxicological Testing 4:42
4.9.12 El Niño: the Interannual Prediction Problem 4:43
4.9.13 Renewable Energies from the Ocean 4:44
4.9.14 Estuarian Waters 4:49

4.9.15 NOS/NMFS Cooperative Research on Coastal Fisheries and Habitats at the Beaufort Laboratory 4:55
4.9.16 Distribution and Dynamics of Gas Hydrates in the Marine Environment 4:57
4.9.17 Oceanographic Environment of Glacier Bay 4:61
4.9.18 NOS Sanctuaries Protect Nation's Maritime History 4:62
4.9.19 Quantification of Anoxia and Hypoxia in Water Bodies 4:64
4.9.20 Floating Ice 4:69
4.9.21 Technology Development: Hardware Development – Marine Instrumentation Laboratory (MIL) 4:70
4.9.22 Seasonal Coupling between Intertidal Macrofauna and Sediment Column Porewater Nutrient Concentrations 4:73
4.9.23 Mapping the Sea Floor of the Historic Area Remediation Site (HARS) Offshore of New York City 4:77
4.9.24 NOAA and University Scientists Study Methyl Bromide Cycling in the North Pacific 4:80
4.9.25 Tidally Mediated Changes in Nutrient Concentrations 4:81
4.9.26 The Role of Oceans in the Global Cycles of Climatically-active Trace-gases 4:85
4.9.27 Pacific Marine Environmental Laboratory– 30 Years of Observing the Ocean 4:89
4.9.28 Seawater Temperature Estimates in Paleoceanography 4:92
4.9.29 Physical Oceanography 4:95
4.9.30 Coastal Water Pollutants 4:96
4.9.31 Trace Element Pollution 4:109
4.9.32 Coral Reefs and Your Coastal Watershed 4:113
4.9.33 Sea Level and Climate 4:117
4.9.34 The Permanent Service for Mean Sea Level 4:118
4.9.35 Marine and Estuarine Microalgal Sediment Toxicity Tests 4:120
4.9.36 Marine Stock Enhancement Techniques 4:124
4.9.37 Physical and Chemical Variability of Tidal Streams 4:128
4.9.38 Black Water Turns the Tide on Florida Coral 4:133
4.9.39 Shallow Water Waves 4:135
4.9.40 Water Waves 4:138
4.9.41 Woods Hole: the Early Years 4:139
4.9.42 An Analysis of the International Maritime Organization – London Convention Annual Ocean Dumping Reports 4:144
4.9.43 Marine Sources of Halocarbons 4:149
4.9.44 Food Chain/Foodweb/Food Cycle 4:151
4.9.45 Plankton 4:154
4.9.46 Major Ions in Seawater 4:159
4.9.47 Tsunami 4:160

4.10 Meteorology 4:164
4.10.1 Ballooning and Meteorology in the Twentieth Century 4:164
4.10.2 Barometric Efficiency 4:166
4.10.3 CERES: Understanding the Earth's Clouds and Climate 4:169
4.10.4 Chinook 4:170
4.10.5 Global Climate Change 4:171
4.10.6 Observations of Climate and Global Change from Real-time Measurements 4:172
4.10.7 Overview: the Climate System 4:179
4.10.8 Climate and Society 4:183
4.10.9 What Is Climatology? 4:186
4.10.10 Cloud Seeding 4:187
4.10.11 Condensation 4:188
4.10.12 Cosmic Water 4:189
4.10.13 The Water Cycle 4:191
4.10.14 Cyclones 4:194
4.10.15 Water Cycle 4:196
4.10.16 Degree Day Method 4:197
4.10.17 Desertification 4:199
4.10.18 Dew 4:200
4.10.19 Dew Deserts 4:201
4.10.20 Dew Point 4:207
4.10.21 Droughts 4:208
4.10.22 Drought Indices 4:209
4.10.23 The Earth Observing System: Aqua 4:214
4.10.24 Entropy Theory for Hydrologic Modeling 4:217
4.10.25 Evaporation 4:223
4.10.26 Evapotranspiration 4:226
4.10.27 Fog 4:229
4.10.28 Coastal Fog along the Northern Gulf of Mexico 4:230
4.10.29 Rain Forests 4:239
4.10.30 Frost 4:240
4.10.31 Frost Damage 4:241
4.10.32 The Global Water Cycle 4:242
4.10.33 Ground-based GPS Meteorology at FSL 4:244
4.10.34 Climate and Water Balance on the Island of Hawaii 4:255
4.10.35 Heat of Vaporization 4:263
4.10.36 Hydrologic History, Problems, and Perspectives 4:265
4.10.37 Humidity – Absolute 4:269
4.10.38 Relative Humidity 4:270
4.10.39 Hurricanes: Modeling Nature's Fury 4:274

4.10.40 Hydrologic Cycle 4:275
4.10.41 Hydrosphere 4:283
4.10.42 Hydrologic Cycle, Water Resources, and Society 4:287
4.10.43 Isohyetal Method 4:290
4.10.44 What about Meteorology? 4:292
4.10.45 Basic Research for Military Applications 4:295
4.10.46 Uncertainties in Rainfall-runoff Modeling 4:297
4.10.47 Monsoon 4:304
4.10.48 African Monsoons 4:304
4.10.49 Permanent Frost 4:305
4.10.50 Radar Use in Rainfall Measurements 4:306
4.10.51 Rainfall 4:309
4.10.52 Rainfall and Runoff 4:315
4.10.53 Remote Sensing of Applications in Hydrology 4:319
4.10.54 Atmospheric Scientists 4:328
4.10.55 Rain Simulators 4:330
4.10.56 Snow Density 4:333
4.10.57 Flattop Mountain Snotel Snowpack: Water Year 2004 .... 4:334
4.10.58 Snow and Snowmelt 4:336
4.10.59 Snow Surveys 4:337
4.10.60 A Statistical Approach to Critical Storm Period Analysis 4:338
4.10.61 Sublimation 4:343
4.10.62 Transpiration 4:345
4.10.63 Waterspout 4:347
4.10.64 United States Weather Bureau 4:347
4.10.65 Weather Forecasting through the Ages 4:348
4.10.66 Overview of Weather Systems 4:352
4.10.67 Unit Hydrograph Theory 4:355
4.10.68 Weather and the Atmosphere 4:360
4.10.69 Vapor Pressure 4:362
4.10.70 Adiabatic Cooling 4:366
4.10.71 Rain and Rocks: the Recipe for River Water Chemistry 4:371
4.11 Physics and Chemistry of Water 4:377
4.11.1 Acid Rain and Society 4:377
4.11.2 Adsorption Capacity of Activated Carbon for Water Purification 4:381
4.11.3 Adsorption of Organic Compounds 4:384
4.11.4 Age Dating Old Groundwater 4:388
4.11.5 Ammonia 4:390
4.11.6 Beryllium in Water 4:394
4.11.7 Dissolved Organic Carbon 4:399

4.11.8 Mechanisms of Water Adsorption on Carbons 4:400

4.11.9 The Effect of Carbon Surface Chemical Composition on the Mechanism of Phenol Adsorption from Aqueous Solutions 4:404

4.11.10 Carbonate Geochemistry 4:408

4.11.11 Carbonate in Natural Waters 4:413

4.11.12 Chlorine-36 and Very Old Groundwaters 4:416

4.11.13 Chlorofluorocarbons (CFCs) 4:420

4.11.14 Coagulation and Flocculation 4:424

4.11.15 Conductivity-electric 4:429

4.11.16 Conservation and the Water Cycle 4:433

4.11.17 Defluroidation 4:434

4.11.18 Deuterium 4:438

4.11.19 Distilled Water 4:441

4.11.20 Electricity as a Fluid 4:442

4.11.21 Analysis of Aqueous Solutions Using Electrospray Ionization Mass Spectrometry (ESI MS) 4:443

4.11.22 Fenton's Reaction and Groundwater Remediation 4:445

4.11.23 Where Water Floats 4:448

4.11.24 Freshwater 4:449

4.11.25 Dissolved Gases 4:450

4.11.26 Hard Water 4:452

4.11.27 An Analysis of the Impact of Water on Health and Aging: Is All Water the Same? 4:455

4.11.28 Heavy Water 4:462

4.11.29 Henry's Law 4:466

4.11.30 Hofmeister Effects 4:468

4.11.31 Clathrate Hydrates 4:471

4.11.32 Hydration 4:475

4.11.33 The Mirage of the H_2 Economy 4:477

4.11.34 Hydrogen Ion 4:480

4.11.35 The Hydronium Ion 4:482

4.11.36 Infiltration and Soil Moisture Processes 4:484

4.11.37 Ion Exchange and Inorganic Adsorption 4:490

4.11.38 Iron 4:496

4.11.39 Isotopes 4:499

4.11.40 Isotope Fractionation 4:500

4.11.41 Mariotte Bottle – Use in Hydrology 4:503

4.11.42 Mars Exploration Rover Mission 4:504

4.11.43 Removal of Organic Micropollutants and Metal Ions from Aqueous Solutions by Activated Carbons 4:506

4.11.44 Molecular Network Dynamics 4:511
4.11.45 *In situ* Chemical Monitoring 4:514
4.11.46 Nitrogen 4:517
4.11.47 Osmosis-diffusion of Solvent or Caused by Diffusion of Solutes? 4:520
4.11.48 Partitioning and Bioavailability 4:521
4.11.49 Physical Properties 4:527
4.11.50 Environmental Photochemistry in Surface Waters 4:529
4.11.51 Isotope Exchange in Gas-water Reactions 4:535
4.11.52 Radon in Water 4:541
4.11.53 Silica in Natural Waters 4:548
4.11.54 Sodium in Natural Waters 4:551
4.11.55 Soft Water 4:553
4.11.56 Solubility of Chemicals in Water 4:555
4.11.57 Solubility of Hydrocarbons in Salt Water 4:559
4.11.58 Solubility of Hydrocarbons and Sulfur Compounds in Water 4:561
4.11.59 Sorption Kinetics 4:564
4.11.60 Sound in Water 4:569
4.11.61 Water on the Space Station 4:572
4.11.62 Strontium Isotopes in Water and Rock 4:574
4.11.63 Technetium in Water 4:578
4.11.64 Water – Nature's Magician 4:583
4.11.65 Freezing and Supercooling of Water 4:585
4.11.66 Chemical Precipitation 4:586
4.11.67 Antimony in Aquatic Systems 4:589

4.12 Water Law and Economics 4:595
4.12.1 The Clean Water Act 4:595
4.12.2 Clean Water Act, Water Quality Criteria/Standards, TMDLs, and Weight-of-evidence Approach for Regulating Water Quality 4:598
4.12.3 The Constitution and Early Attempts at Rational Water Planning 4:604
4.12.4 Economic Value of Water: Estimation 4:605
4.12.5 Water Supply Planning – Federal 4:612
4.12.6 Flood Control Act of 1944 ..... 4:616
4.12.7 Great Lakes Governors' Agreement 4:617
4.12.8 U.S./Canadian Boundary Waters Treaty and the Great Lakes Water Quality Agreement 4:620
4.12.9 Great Lakes Water Quality Initiative 4:621
4.12.10 Quantitative Groundwater Law 4:627
4.12.11 Islamic Water Law 4:634
4.12.12 Transboundary Waters in Latin America: Conflicts and Collaboration 4:636

4.12.13 United States-Mexico Border Waters: Conventions, Treaties, and Institutions 4:643
4.12.14 Negotiating between Authority and Polluters: an Approach to Managing Water Quality 4:647
4.12.15 Water Resource Organizations 4:648
4.12.16 A Brief History of the Water Pollution Control Act in the U.S. 4:651
4.12.17 The National Pollution Discharge Elimination System 4:655
4.12.18 Legal Protection for in-stream Flow 4:659
4.12.19 Water Quality 4:664
4.12.20 Interface between Federal Water Quality Regulation and State Allocation of Water Quantity 4:671
4.12.21 Regulatory Issues and Remediation: Risk, Costs, and Benefits 4:674
4.12.22 The Safe Drinking Water Act 4:676
4.12.23 Representing Geopolitics of (Hydro) Borders in South Asia 4:680
4.12.24 Water Transfers 4:685
4.12.25 Reserved Water Rights for Indian and Federal Lands 4:689
4.12.26 Wetlands Policy in the United States: from Drainage to Restoration 4:690
4.13 Water History, Art, and Culture 4:695
4.13.1 Curious Uses of Agricultural Water in the World 4:695
4.13.2 Water between Arabs and Israelis: Researching Twice-promised Resources 4:699
4.13.3 The Myth of Bad Cholesterol: Why Water Is a Better Cholesterol-lowering Medication 4:701
4.13.4 Early Clocks 4:702
4.13.5 Water Clocks 4:704
4.13.6 Our Evolving Water Consciousness 4:707
4.13.7 Effective Water Education Strategies in a Nontraditional Setting 4:713
4.13.8 Evolution 4:715
4.13.9 History of Pond Fisheries in Poland 4:718
4.13.10 Water: the Key to Natural Health and Healing 4:722
4.13.11 Water in History 4:726
4.13.12 Hoover Dam History 4:732
4.13.13 Hydropsychology 4:733
4.13.14 Jacob's Well 4:735
4.13.15 Jaubert de Passa: the First World History of Irrigation in 1846 ..... 4:736
4.13.16 Benjamin Franklin: from Kite to Lightning Rod 4:741
4.13.17 Water and the History of Man 4:745
4.13.18 Water, Bacteria, Life on Mars, and Microbial Diversity 4:746
4.13.19 Canals in the Mekong Delta: a Historical Overview from 200 C.E. to the Present 4:748

4.13.20 Conflict and Water Use in the Middle East 4:753
4.13.21 Benjamin Franklin's Armonica: a Water Music Instrument 4:758
4.13.22 Water Transportation Occupations 4:762
4.13.23 A Concise Glimpse of Water in the History of Photography 4:766
4.13.24 Water as a Human Right 4:769
4.13.25 Ancient Water and Soil Conservation Ecosystems of Sri Lanka 4:772
4.13.26 Ben Franklin's Gulf Stream Weather and Swim Fins 4:780
4.13.27 Water Symbolism 4:785
4.13.28 Gordon and Franklin Rivers and the Tasmanian Wilderness World Heritage Area 4:788
4.13.29 The Mistake of Waiting to Get Thirsty 4:789
4.13.30 Water and Well-being 4:791
4.13.31 Free Flowing Water: a Source of Wisdom 4:793
4.13.32 The Medicinal Properties of the Waters of Saratoga Springs 4:797
4.13.33 A History of Hawaiian Freshwater Resources 4:801

Volume 5. Ground Water

5.14 Ground Water 5:1
- 5.14.1 Acid Mine Drainage: Sources and Treatment in the United States 5:1
- 5.14.2 Aquifers 5:9
- 5.14.3 Artificial Recharge of Unconfined Aquifer 5:11
- 5.14.4 Groundwater and Arsenic: Chemical Behavior and Treatment 5:17
- 5.14.5 Treatment of Arsenic, Chromium, and Biofouling in Water Supply Wells 5:22
- 5.14.6 Artesian Water 5:29
- 5.14.7 Modeling Contaminant Transport and Biodegradation in Groundwater 5:30
- 5.14.8 Biofouling in Water Wells 5:35
- 5.14.9 *In situ* Bioremediation of Contaminated Groundwater 5:38
- 5.14.10 Process Limitations of *in situ* Bioremediation of Groundwater 5:42
- 5.14.11 Black Mesa Monitoring Program 5:48
- 5.14.12 Brine Deposits 5:51
- 5.14.13 Connate Water 5:54
- 5.14.14 Consolidated Water Bearing Rocks 5:55
- 5.14.15 Sensitivity of Groundwater to Contamination 5:56
- 5.14.16 Water Contamination by Low Level Organic Waste Compounds in the Hydrologic System 5:60
- 5.14.17 Darcy's Law 5:63
- 5.14.18 Groundwater Dating with Radiocarbon 5:64
- 5.14.19 Groundwater Dating with H–He 5:65
- 5.14.20 Dating Groundwaters with Tritium 5:69
- 5.14.21 Recharge in Desert Regions around the World 5:72

5.14.22 Hydrologic Feasibility Assessment and Design in Phytoremediation 5:76
5.14.23 Well Design and Construction 5:87
5.14.24 Physical Properties of DNAPLs and Groundwater Contamination 5:91
5.14.25 Water Dowsing (Witching) 5:92
5.14.26 Subsurface Drainage 5:94
5.14.27 Drawdown 5:101
5.14.28 Water Level Drawdown 5:102
5.14.29 Water Well Drilling Techniques 5:105
5.14.30 Groundwater Dye Tracing in Karst 5:107
5.14.31 Earthquakes – Rattling the Earth's Plumbing System 5:111
5.14.32 *In situ* Electrokinetic Treatment of MtBE, Benzene, and Chlorinated Solvents 5:116
5.14.33 Field Capacity 5:124
5.14.34 Groundwater Flow Properties 5:128
5.14.35 Fluoride Contamination in Ground Water 5:130
5.14.36 Rock Fracture 5:136
5.14.37 Geochemical Models 5:138
5.14.38 Geochemical Modeling – Computer Codes 5:140
5.14.39 Geochemical Modeling – Computer Code Concepts 5:142
5.14.40 Geological Occurrence of Groundwater 5:145
5.14.41 Geophysics and Remote Sensing 5:145
5.14.42 Geothermal Water 5:156
5.14.43 Ghijben-Herzberg Equilibrium 5:158
5.14.44 Groundwater Balance 5:162
5.14.45 Hydraulic Head 5:169
5.14.46 The Role of Heat in Groundwater Systems 5:172
5.14.47 Groundwater Flow in Heterogenetic Sediments and Fractured Rock Systems 5:175
5.14.48 Horizontal Wells 5:177
5.14.49 Horizontal Wells in Groundwater Remediation 5:178
5.14.50 Head 5:180
5.14.51 Well Hydraulics and Aquifer Tests 5:182
5.14.52 Hydraulic Properties Characterization 5:184
5.14.53 Mobility of Humic Substances in Groundwater 5:188
5.14.54 Assessment of Groundwater Quality in District Hardwar, Uttaranchal, India 5:192
5.14.55 Irrigation Water Quality in District Hardwar, Uttaranchal, India 5:204
5.14.56 Infiltration and Soil Water Processes 5:210
5.14.57 Infiltration/Capacity/Rates 5:212
5.14.58 Infiltrometers 5:214
5.14.59 Summary of Isotopes in Contaminant Hydrogeology 5:216

5.14.60 Environmental Isotopes in Hydrogeology 5:227
5.14.61 Water-jetting Drilling Technologies for Well Installation and *in situ* Remediation of Hydrocarbons, Solvents, and Metals 5:234
5.14.62 Karst Hydrology 5:235
5.14.63 Karst Topography 5:243
5.14.64 Detecting Modern Groundwaters with ^{85}Kr 5:248
5.14.65 Land Use Impacts on Groundwater Quality 5:250
5.14.66 Groundwater Contamination from Municipal Landfills in the USA 5:253
5.14.67 Metal Organic Interactions in Subtitle D Landfill Leachates and Associated Ground Waters 5:258
5.14.68 Leaching 5:260
5.14.69 Well Maintenance 5:263
5.14.70 Megawatersheds 5:266
5.14.71 Mass Transport in Saturated Media 5:273
5.14.72 Soil and Water Contamination by Heavy Metals 5:275
5.14.73 Source, Mobility, and Remediation of Metals 5:280
5.14.74 Genetics of Metal Tolerance and Accumulation in Higher Plants 5:284
5.14.75 *In situ* Groundwater Remediation for Heavy Metal Contamination 5:290
5.14.76 Methane in Groundwater 5:293
5.14.77 Fossil Aquifers 5:294
5.14.78 What is a Hydrochemical Model? 5:295
5.14.79 Modeling Non-point Source Pollutants in the Vadose Zone Using GIS 5:299
5.14.80 Modeling Techniques for Solute Transport in Groundwater 5:305
5.14.81 Ambient Groundwater Monitoring Network Strategies and Design 5:313
5.14.82 MTBE 5:318
5.14.83 Limiting Geochemical Factors in Remediation Using Monitored Natural Attenuation and Enhanced Bioremediation 5:319
5.14.84 Nitrate Contamination of Groundwater 5:322
5.14.85 Treatment for Nitrates in Groundwater 5:323
5.14.86 Nonpoint Sources 5:331
5.14.87 Organic Compounds in Ground Water 5:337
5.14.88 Overdraft 5:340
5.14.89 Chemical Oxidation Technologies for Groundwater Remediation 5:344
5.14.90 Particulate Transport in Groundwater – Bacteria and Colloids 5:349
5.14.91 Perched Groundwater 5:352
5.14.92 Permeability 5:355
5.14.93 Groundwater Vulnerability to Pesticides: an Overview of Approaches and Methods of Evaluation 5:357
5.14.94 High pH Groundwater – the Effect of the Dissolution of Hardened Cement Pastes 5:362

5.14.95 Phytoextraction and Phytostabilization: Technical, Economic and Regulatory Considerations of the Soil-lead Issue 5:365
5.14.96 Phytoextraction of Zinc and Cadmium from Soils Using Hyperaccumulator Plants 5:369
5.14.97 Phytoremediation Enhancement of Natural Attenuation Processes 5:374
5.14.98 Bacteria Role in the Phytoremediation of Heavy Metals 5:376
5.14.99 Phytoremediation of Lead-contaminated Soils 5:381
5.14.100 Phytoremediation of Methyl *Tertiary*-butyl Ether 5:385
5.14.101 Phytoremediation of Selenium-laden Soils 5:397
5.14.102 Soil Pipes and Pipe Flow 5:401
5.14.103 Low Flow Groundwater Purging and Surging 5:404
5.14.104 Groundwater Quality 5:406
5.14.105 Radial Wells 5:407
5.14.106 Recharge in Arid Regions 5:408
5.14.107 Sub-surface Redox Chemistry: a Comparison of Equilibrium and Reaction-based Approaches 5:413
5.14.108 Regional Flow Systems 5:417
5.14.109 Groundwater Remediation by Injection and Problem Prevention 5:421
5.14.110 Groundwater Remediation: *in situ* Passive Methods 5:423
5.14.111 Groundwater Remediation by *in situ* Aeration and Volatilization 5:426
5.14.112 Remediation of Contaminated Soils 5:432
5.14.113 Groundwater Remediation Project Life Cycle 5:436
5.14.114 Innovative Contaminated Groundwater Remediation Technologies 5:438
5.14.115 Resistivity Methods 5:443
5.14.116 Risk Analysis of Buried Wastes from Electricity Generation 5:448
5.14.117 Groundwater Contamination from Runoff 5:451
5.14.118 Saline Seep 5:453
5.14.119 Groundwater Sampling Techniques for Environmental Projects 5:454
5.14.120 Groundwater Sampling with Passive Diffusion Samplers 5:456
5.14.121 Specific Capacity 5:460
5.14.122 Soil Water 5:461
5.14.123 Soil and Groundwater Geochemistry and Microbiology 5:463
5.14.124 Characterizing Soil Spatial Variability 5:465
5.14.125 Deep Soil-water Movement 5:471
5.14.126 Specific Gravity 5:473
5.14.127 Hot Springs 5:475
5.14.128 Squeezing Water from Rock 5:477
5.14.129 Storage Coefficient 5:480
5.14.130 Qanats: an Ingenious Sustainable Groundwater Resource System 5:483
5.14.131 Lysimeters 5:487

5.14.132 Steady-state Flow Aquifer Tests 5:491
5.14.133 Tidal Efficiency 5:497
5.14.134 Combined Free and Porous Flow in the Subsurface 5:498
5.14.135 Groundwater Tracing 5:501
5.14.136 Hydraulic Conductivity/Transmissibility 5:507
5.14.137 Groundwater Flow and Transport Process 5:514
5.14.138 Reactive Transport in the Saturated Zone: Case Histories for Permeable Reactive Barriers 5:518
5.14.139 Transport of Reactive Solute in Soil and Groundwater 5:524
5.14.140 Water in the Unsaturated Zone 5:531
5.14.141 Groundwater and Vadose Zone Hydrology 5:533
5.14.142 Vadose Zone Monitoring Techniques 5:538
5.14.143 Vapor Transport in the Unsaturated Zone 5:543
5.14.144 Applications of Soil Vapor Data to Groundwater Investigations 5:548
5.14.145 Groundwater Velocities 5:554
5.14.146 Viscous Flow 5:555
5.14.147 Vulnerability Mapping of Groundwater Resources 5:561
5.14.148 Water/Rocks Interaction 5:566
5.14.149 Ground Water: Wells 5:571
5.14.150 Well Screens 5:572
5.14.151 Well TEST 5:574
5.14.152 Safe Yield of an Aquifer 5:575
5.14.153 Specific Yield Storage Equation 5:576
5.14.154 Microbial Processes Affecting Monitored Natural Attenuation of Contaminants in the Subsurface 5:578
5.14.155 Groundwater Vulnerability to Pesticides: Statistical Approaches 5:594
5.14.156 Groundwater – Nature’s Hidden Treasure 5:599
5.14.157 Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams 5:605
5.14.158 The Environmental Impact of Iron in Groundwater 5:608
5.14.159 Groundwater and Cobalt: Chemical Behavior and Treatment 5:610
5.14.160 Groundwater and Cadmium: Chemical Behavior and Treatment 5:613
5.14.161 Groundwater Modeling 5:619
5.14.162 Groundwater and Benzene: Chemical Behavior and Treatment 5:626
5.14.163 Groundwater and Nitrate: Chemical Behavior and Treatment 5:628
5.14.164 Groundwater and Perchlorate: Chemical Behavior and Treatment 5:631
5.14.165 Groundwater and Vinyl Chloride: Chemical Behavior and Treatment 5:634
5.14.166 Groundwater and Uranium: Chemical Behavior and Treatment 5:640
5.14.167 Groundwater and Mercury: Chemical Behavior and Treatment 5:642
5.14.168 Groundwater and Lead: Chemical Behavior and Treatment 5:645

5.14.169 Laminar Flow 5:649
5.14.170 Finite Element Modeling of Coupled Free and Porous Flow 5:655
5.14.171 Unconfined Groundwater 5:662
5.14.172 Modeling of DNAPL Migration in Saturated Porous Media 5:668
5.14.173 The Use of Semipermeable Membrane Devices (SPMDs) for Monitoring, Exposure, and Toxicity Assessment 5:672
5.14.174 River-connected Aquifers: Geophysics, Stratigraphy, Hydrogeology, and Geochemistry 5:677

Index **I:1**

DOMESTIC WATER SUPPLY

THE ARSENIC DRINKING WATER CRISIS IN BANGLADESH

Charlie Bryce
Jim Philp
Napier University
Edinburgh, Scotland,
United Kingdom

INTRODUCTION

Although the incidence of arsenic poisoning in groundwater is worldwide and includes Bangladesh and India, Taiwan, Vietnam, Chile, China, North America, and Finland, the area of the highest demand for a resolution of the problem is Bangladesh. The source seems to be geological, for arsenic has been found in tube well water used for drinking and irrigation, although the geochemistry is not completely understood. As many thousands of boreholes have been produced to support modern irrigation systems, the underground aquifers are aerated, which causes transformation of anaerobic conditions to aerobic conditions. The presence of oxygen in this way decomposes arsenopyrite-releasing arsenic acid. At low pH, this arsenic dissolves in water and hence leads to water contamination. The arsenic content of sediments is high relative to crustal concentrations. The biogeochemical cycling of arsenic and iron are coupled in deltaic systems; iron oxyhydroxides act as a carrier for the deposition of arsenic in sediments. From there, it can be mobilized by bicarbonate, which can extract arsenic from sediments under both aerobic and anaerobic conditions.

Arsenic also becomes a pollutant as a result of various industrial uses and activities. Arsenic is a metalloid, and its primary usage has been in agriculture, in formulating herbicides, especially for controlling weeds in cotton fields. Sodium arsenite has been used as an insecticidal ingredient in sheep-dips. In industry, arsenic has found use in glass manufacture and a new role in the semiconductor industry. Copper smelting releases significant amounts into soils.

ARSENIC AND THE GEOGRAPHY OF BANGLADESH

Bangladesh, 85% of which is deltaic and alluvial plain, is situated in the lower end of three large river systems, the Ganges, the Brahmaputra, and the Meghna, whose catchment area is about 600,000 square miles (Fig. 1). The sediments produced in the catchment areas are very high and expose the underlying rocks, including arsenic-bearing rocks.

Arsenic pollution became a live issue in Bangladesh as recently as 1993, following a warning by the World Health Organisation (WHO) that levels of arsenic in groundwater above the permissible limit of 0.05 mg/L had been reported in seven districts of adjoining West Bengal in India. The Department of Public Health Engineering of Bangladesh was invited to test water samples from the adjoining eight Bangladeshi districts that have the same geographical continuity and aquifers as the West Bengal districts; this yielded the result that slightly more than 20% of the samples contained arsenic at levels ranging from 0.01 to 0.4 mg/L. Ten million people populate these areas and hence are at risk of arsenic toxicity.

Figure 1. The geography of Bangladesh.

Since that time, it has been shown that there is groundwater contamination in more than 40 districts that endanger in excess of 50 million people. The problem has been described in *The Lancet* as the world's worst episode of arsenic poisoning; more than 220,000 people reportedly suffer from arsenic-related diseases. In a recent study of 27 districts in Bangladesh, 58% of the water samples were unsuitable for drinking. The worst case was in Nawabganj district, where one well contained 60 times the WHO maximum permissible level.

TOXICITY AND DISEASE

Arsenic occurs principally in the forms of organic arsenic (methyl arsonic acid, dimethyl arsonic acid, arsenobetaine, and arsenocholine) and inorganic arsenic (trivalent and pentavalent arsenic). Of these, the trivalent form is the most toxic to humans (20 times more so than the pentavalent form) and is the most difficult to remove chemically from water.

Arsenic is a suspected carcinogen and has many acute effects on human health. But at the concentrations present in drinking water, it has no immediate side affects. The latency (i.e., the time from first exposure to manifestation of disease) for arsenic-caused skin lesions, in particular keratoses, is typically of the order of 10 years, and so a major increase in the number of cases of arsenic-caused diseases can be projected into the future (Fig. 2).

Figure 2. Common manifestations of long-term, chronic arsenic poisoning.

Exposure to arsenic in this way can lead to latent or manifest clinical symptoms through even low-level exposure over a period of time. This can result in an accumulation of this toxicant in various organs and systems, affecting their normal functioning, including the kidney and nervous system. Arsenic causes skin cancers and internal cancers such as lung and bladder cancer. The most common manifestations in afflicted people in Bangladesh are melanosis (93.5%), keratoses (68.3%), hyperkeratosis (37.6%), and dipigmentation (leucomelanosis) (39.1%). Cancers are found in 0.8% of the afflicted population. Preliminary work indicates that there may be several factors triggering arsenic-related diseases, but experts generally feel that poor nutrition may be a primary cause. Studies in Taiwan have shown that there is an increased occurrence of diabetes in the population exposed to arsenic via drinking water.

Recent studies have shown that arsenic is also a teratogen. Further, at the *5th International Conference on Arsenic* held in Dhaka, 2004, one of the key messages and cause for increased concern is that there is very good evidence that the environmental contaminant is getting into the food chain, thus putting even more lives at risk.

DRINKING WATER STANDARDS

The World Health Organisation has set 10 μg/L as the allowable level for arsenic in drinking water. On January 22, 2001, the U.S. EPA adopted this standard, and public water systems must comply by January 23, 2006.

DETECTING ARSENIC IN DRINKING WATER

In the modern analytical laboratory, arsenic is quantified by soluble arsenic assaying, preferable with GF-AAS (graphite furnace-atomic absorption spectrometry) for detection levels of less than 50 μg/L. However, given the highly dispersed nature of tube wells in Bangladesh, the transport of the many samples to central laboratories is logistically impossible. Field techniques are more important, so that samples can be processed as they are taken. In Bangladesh, this requires inexpensive and completely portable techniques. At present, a chemical test kit is being used (Fig. 3). The disadvantage of this approach is that the sensitivity of the chemistry (poor below 100 μg/L) is not compatible with the levels of contamination that need to be detected (50 μg/L and less). Often, at best, the presence or absence of arsenic can be inferred, but not the level of contamination. In addition, such testing is slow and can take about 6 months to cover some 2000 villages in a district.

Bioavailability Biosensors for Detecting and Quantifying Arsenic

A few strains of bacteria are resistant to arsenate, arsenite, and antimonite through the action of the gene products of the *ars* operon. The *ars* operon consists of five genes that code for three structural and two regulatory proteins. Two structural genes in the *ars* operon, *ars*A and *ars*B, code for proteins that form an efflux pump that transports arsenite and antimonite out of cells.

A means of measuring available arsenicals would be to construct a gene fusion plasmid in which part of the *ars* operon is fused upstream of a reporter gene system, such as the bacterial *lux* operon, which results in the production of light. A transcriptional gene fusion has been done (1) that consists of *E. coli ars*B :: *lux*AB. The detection limit of arsenic is of the order of 10 μg/L. Moreover, bioluminescence may be inducible in a concentration-dependent manner (Fig. 4) by arsenic salts; high concentrations result in higher bioluminescence, so that such biosensors may be able to

- Quantify arsenic within the required range of drinking water in Bangladesh
- Provide a measure of the bioavailability of the arsenic for risk assessment

REMOVAL OF ARSENIC FROM DRINKING WATER

Coprecipitation of arsenate with ferric (Fe^{3+}) ion is currently the most effective and practical method of arsenic removal. Optimum stability of the $FeAsO_4$ precipitate occurs at Fe/As molar ratios of >4; this ratio increases significantly, in practice, depending on water turbidity, slime levels, dissolved solids, and the presence of

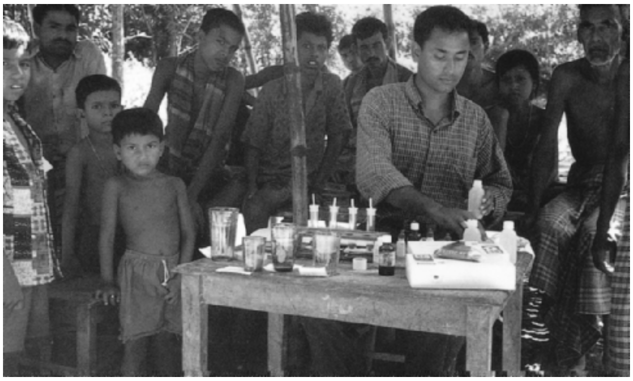

Figure 3. Testing for arsenic in drinking water using test kits at a village.

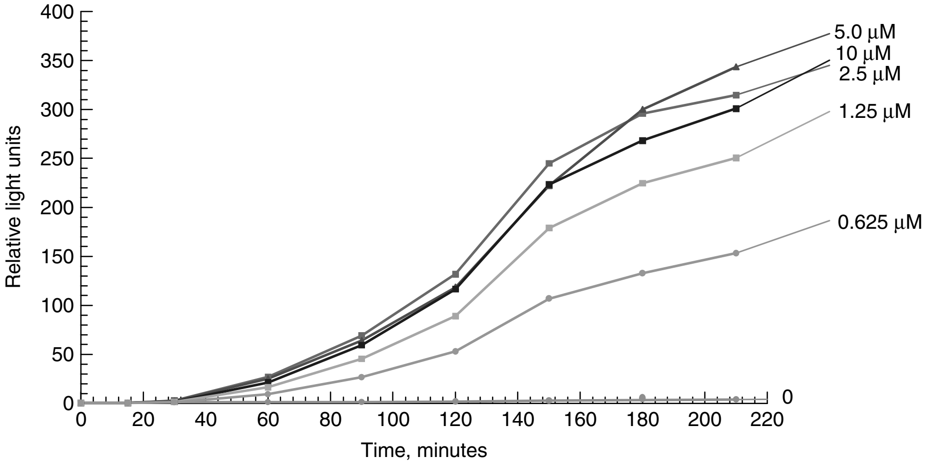

Figure 4. The response of a bacterial whole-cell bioavailability biosensor to increasing concentrations of arsenic.

iron-consuming species. However, Fe^{3+} ion coprecipitation of arsenite ($AsO_3{}^{3-}$) is moderately effective at pH ~7.0. The trivalent As(III) species must be oxidized to As(V) for complete precipitation with Fe^{3+} ion. Oxidation may be achieved through aeration or by adding oxidizers such as hypochlorite, permanganate, peroxide, and ozone. The application of other technologies, including alum and lime precipitation together with activated alumina adsorption, are not fully effective.

In Bangladesh, for geographical and financial reasons, there is likely to be a preference for local treatment rather than large-scale treatment plants. The ideal solution would be to modify each tube well at low cost for arsenic removal by, for example, ion exchange.

BIBLIOGRAPHY

1. Cai, J. and DuBow, M.S. (1997). Use of a bioluminescent bacterial biosensor for biomonitoring and characterization of arsenic toxicity of chromated copper arsenate (CCA). *Biodegradation* **8**(2): 105–111.

READING LIST

Akhtar, S.A. et al. (1997). Arsenic contamination in ground water and arsenicosis in Bangladesh. *Int. J. Environ. Health Res.* **7**: 71–276.

Anawar, H.M., Akai, J., and Sakugawa, H. (2004). Mobilization of arsenic from subsurface sediments by effect of biocarbonate ions in groundwater. *Chemosphere* **54**: 753–762.

Dave, J.M. (1997). Arsenic contamination of drinking water in Bangladesh. *WHO SEA/EH/Meet.3/6.17*.

Dipankar, D. et al. (1996). Arsenic in groundwater in six districts of West Bengal, India. *Environ. Geochem. Health* **18**: 5–15.

Karim, M.D.M. (2000). Arsenic in groundwater and health problems in Bangladesh. *Water Res.* **34**: 304–310.

Tareq, S.M., Safiullah, S., Anawar, H.M., Majibur Rahman, M., and Ishizuka, T. (2003). Arsenic pollution in groundwater: a self-organizing complex geochemical process in the deltaic sedimentary environment, Bangladesh. *Sci. Total Environ.* **313**: 213–226.

Wadud Khan, A. et al. (1997). Arsenic contamination in ground water and its effect on human health with particular reference to Bangladesh. *J. Preventative Soc. Med.* **16**: 65–73.

SOME OTHER ARSENIC LINKS

Harvard University Arsenic Project Website

Natural Resources Defense Council—FAQs: Arsenic in drinking water

U.S. Agency for Toxic Substances and Disease Registry—ToxFAQs: Arsenic

U.S. Environmental Protection Agency—Arsenic Standard pages and Q & A's: Occurrence of Arsenic in Ground Water

West Bengal and Bangladesh Arsenic Crisis Information Centre

World Health Organisation http://www.who.int/mediacentre/factsheets/fs210/en/

BOTTLED WATER

Robert M. Hordon
Rutgers University
Piscataway, New Jersey

INTRODUCTION

Bottled water sales in the United States have increased dramatically during the past decade. Total domestic and imported sales rose 142% from almost 2.5 billion gallons (9.4 billion liters) in 1992 to more than 6 billion gallons (22.8 billion liters) in 2002. Bottled water revenues rose nearly 190% from \$2.66 billion in 1992 to \$7.7 billion in 2002. Using the same 11-year period from 1992 to 2002 (1), per capita consumption in the U.S. increased more than 119% from 9.8 to 21.5 gallons (37.1 to 81.4 liters).

The global water market also shows comparable increases. For example, the world total consumption of bottled water rose nearly 63% from 21.3 billion gallons (80.6 billion liters) in 1997 to 34.7 billion gallons (131.3 billion liters) in 2002. Using the same 6-year period from 1997 to 2002, global per capita consumption increased 107% from 5.7 to 11.8 gallons (21.6 to 44.7 liters). Italy, Mexico, and France were the top three countries in per capita consumption in 2002 at 44.2, 37.7, and 37.1 gallons (167.3, 142.7, and 140.4 liters), respectively. In per capita consumption, the United States rose from a rank of 15 in 1997 to a rank of 11 in 2002 (1).

The rising consumption in the United States is attributed to an increasingly effective advertising campaign that touts bottled water as a safer and better tasting alternative to tap water. Packaging labels that show massive glaciers and springs in alpine settings have also helped.

HISTORICAL USES OF BOTTLED OR MINERAL WATER

The presumed reason for drinking mineral water and more recently bottled water is for the purported therapeutic effects and associated health benefits. Using mineral waters as a form of therapy was discussed by the ancient Egyptians and Greeks. Based on legend, the thermal springs of Bath, England, date to 800 B.C. Hannibal's army (and elephants) was reputed to have refreshed itself in the pools of Vergeze in southern France in 218 B.C. while enroute to attack Rome. Later on, the spring waters at Vergeze became known as the source of Perrier Water. The mineral waters from 12 springs at Vichy, France, date back to Roman times. Spring water from Fiuggi near Rome, Italy, was used by Michaelangelo. Another Italian luminary from the fifteenth century, Leonardo da Vinci, used the mineral waters from the three deep springs (396 m; 1300 ft) at San Pellegrino. Famous spas and watering holes, such as Hot Springs in Arkansas and Saratoga Springs in New York and many others in Europe, developed near mineral springs during the late 1800s and early 1900s (2).

Poland Spring water from Maine began to be distributed during the mid-1800s. Bottled water from Mountain Valley near Hot Springs, Arkansas, goes back to 1871. Commercial bottling of San Pellegrino water from Italy began in 1899, and export of Evian water from France to the United States began in 1905 (2).

TYPES OF BOTTLED WATER

Bottled water can be grouped into the following several categories depending upon the nature of the water and its source.

Nonsparkling Water

This includes spring water, artesian water, mineral water, and purified water. Domestic production of this type of water in the United States made up more than 95% of the bottled water market in 2002 (1).

1. Spring water is ground water. It comes from a water-bearing subsurface geologic formation known as an aquifer from which water flows naturally to the earth's surface. Water of this type can be collected only at the spring or from a well that taps the aquifer that feeds the spring.
2. Artesian water is derived from a well in an aquifer that is under pressure due to overlying confining layers. Artesian or confined well water can be collected with external pumps that supplement the natural underground pressure. The word "artesian" was derived from the first deep wells that were drilled into confined aquifers in the province of Artois in northern France from about 1750 (3).
3. Mineral water naturally contains at least 250 ppm of mineral salts such as calcium, chloride, sulfate, carbonate, and bicarbonate. No minerals can be added artificially, and it cannot come from a municipal source.
4. Purified (or demineralized) water may come from a municipal source and is treated by one or more of the following water treatment processes:
   a) Distillation: heating of water to produce water vapor which is then condensed and collected;
   b) reverse osmosis, where water is filtered by passing it though a membrane;
   c) deionization: a process where minerals are drawn to particles of the opposite electrical charge and then removed.

Sparkling Water

Sparkling water may include any type of naturally carbonated water. In addition, if the water is treated, CO_2 can be added to the product as long as the water has the same amount of CO_2 as it had when it emerged from its source.

Domestic and imported sparkling water made up about 2.6% and 2.1% of the U.S. bottled water market in 2002, respectively (1).

Beverages that contain certain ingredients or additives, such as sugar, fall into a separate category called soft drinks. Thus, tonic water, soda water, and seltzer are not considered bottled waters and are regulated differently.

REGULATORY AGENCIES

Public potable water supplies in the United States are regulated by the Environmental Protection Agency (EPA) under the federal Safe Drinking Water Act (SDWA). This means that all public water systems that operate under either public or private investor-owned companies that serve 25 or more people must be tested regularly for up to 118 chemicals and bacteria that are specified by the SDWA.

In contrast, bottled water in the United States is regulated as a packaged food product by the federal Food and Drug Administration (FDA). By law, the FDA must follow the same water quality standards as outlined in the Safe Drinking Water Act. In addition, bottled water companies are required to comply with FDA's quality standards, labeling rules, and good manufacturing practices. Finally, bottlers that are members of the International Bottled Water Association (IBWA) may opt to receive random, unannounced site inspections annually by a third-party organization. However, not all bottled water companies comply with the standards of the IBWA. To compound the regulatory issues associated with bottled water, the standards of the IBWA are not legally enforceable (4). Also, the results of any water quality tests that are made by the bottled water companies need not be released to the public (5). This stands in sharp contrast with the water quality reports that all United States public water supply purveyors must furnish each year to their customers.

The EPA states that bottled water is not inherently safer than tap water distributed by public water systems (6). Although tap water and bottled water must meet the standards set by their respective regulatory agencies, the FDA requires testing only once a year for bottled water, whereas the EPA requires much more frequent, often daily, testing. Another factor worth noting is that the FDA rules apply only to water sold in interstate commerce. The Natural Resources Defense Council (7) estimates that about 60% of all bottled water is sold in the same state where it is bottled; thus, FDA rules do not apply.

State regulations and industry standards affect bottled water at the state level. Some states abide by the FDA standards; others are even more stringent. For example, the State of New Jersey requires that the labels of all bottled water products must contain an expiration date of 2 years from the date of bottling (8). However, about 20% of the states have either very limited enforcement powers or no regulations at all.

Fluorides

It is well known that fluorides are a key factor in reducing tooth decay, particularly in children. Drinking water that has been adjusted to optimal fluoride levels prevents cavities and thereby improves dental health. The consumption of fluoridated water is more effective than the use of fluoridated toothpastes or mouth rinses as the latter is only on the teeth for a short time, whereas fluoridated water can be delivered to the teeth continuously through the bloodstream and saliva (5). Most bottled water does not have an optimal fluoride level. Although some bottlers provide fluoride information on their labels, they are not required to do so.

Safety of Plastic Bottles

For reasons of convenience and nonbreakability, most bottled waters are sold in plastic containers that contain phthalates. Water is a universal solvent, so phthalates can be leached from a plastic bottle. It is still not known if there are any negative health effects from human exposure to phthalates. However, rodents have experienced adverse effects from the chemical in some studies. It is apparent that additional research is needed on this issue (5).

COST TO THE CONSUMER

One thing is very clear about the differences between bottled water and tap water: the former is orders of magnitude more expensive (240–10,000 times greater). Bottled water packaged in convenient sizes of plastic containers costs about $6.60/gallon ($1.74/liter) compared to average costs of 1/100 of a cent/gallon for municipal tap water.

Is the difference warranted by taste, quality, and convenience in carrying, or does the explosive growth in bottled water sales reflect the success of mass marketing appeals to a more affluent generation that follows current fashion trends in beverage types? Time will tell, but for the forseeable future, bottled water sales continue to increase, even though the EPA states on its website that bottled water is not necessarily safer than regular tap water.

BIBLIOGRAPHY

1. Beverage Marketing Corp. (2003). *Bottled Water*. New York.
2. LaMoreaux, P.E. and Tanner, J.T. (Eds.). (2001). *Springs and Bottled Waters of the World: Ancient History, Source, Occurrence, Quality and Use*. Springer-Verlag, New York.
3. Todd, D.K. (1980). *Groundwater Hydrology*, 2nd Edn. John Wiley & Sons, New York.
4. International Bottled Water Association. www.bottledwater.org.
5. Minnesota Department of Health. (2003). *Bottled Water: Questions and Answers*. St. Paul, MN.
6. U.S. EPA (Environmental Protection Agency). *Ground Water and Drinking Water*. www.epa.gov/safewater.
7. Natural Resources Defense Council. (1999). *Bottled Water: Pure Drink or Pure Hype?* New York, NY.
8. New Jersey Department of Health and Senior Services. (2002). *Report to the New Jersey Legislature Summarizing Laboratory Test Results on the Quality of Bottled Drinking Water for the period January 1, 2001 Through December 31, 2001*. Trenton, NJ.

CORROSION CONTROL IN DRINKING WATER SYSTEMS

Christian J. Volk
Indiana-American Water
Richmond, Indiana

CORROSION IN DISTRIBUTION SYSTEM

Corrosion Process and Corrosion Cell

Corrosion can be defined as the wearing away or deterioration of a material because of chemical reaction with its environment. When iron or steel is exposed to water, rust (oxidized iron) forms (1). Water that promotes corrosion is defined as aggressive or corrosive. The corrosion processes consist of a series of electrochemical reactions occurring at the metal surface in contact with water and its constituents. Corrosion is an extremely complex chemical and electrochemical phenomenon. During oxidative reactions, local galvanic couples form on the surface of the metal, in which the metal is oxidized, while the oxidant is reduced. Each couple is a microbattery where the corrosion reaction proceeds with a flow of electric current between anodic and cathodic sites on the metal (Fig. 1). The electrochemical corrosion corresponds to the destruction of a metal by electron transfer reaction. All the components of an electrochemical cell must be present for this type of corrosion to occur. The components include an anode and a cathode (which are sites that have a different electrical potential on the metal), an electrical path between the anode and cathode for electron transport (internal circuit), an electrolyte solution that will conduct ions between the anode and cathode (external circuit), and an oxidizing agent to

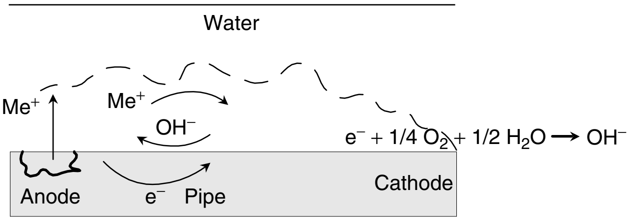

Figure 1. Pipe corrosion in water.

be reduced at the cathode (2). As metallic plumbing materials are not completely homogeneous, anodic and cathodic sites occur on the pipe surface. Oxidation and dissolution of the metal occur at the anode (Fig. 1). Electrons generated at the anode migrate to the cathode, where they are accepted by an electron acceptor, such as oxygen, after chemical reduction. Consequently, the positive ions generated at the anode migrate through the solution to the cathode and negative ions generated at the cathode tend to migrate to the anode. The mechanisms of corrosion are extremely complex and depend on the interactions of physical and chemical factors and the material itself. Table 1 shows different materials, their use in distribution systems, and corrosion-associated problems. Generally, inert and nonmetallic materials like concrete or plastic are more corrosion resistant than metallic pipes.

Scale Formation

The formation of a scale on the pipe surface protects the pipe from corrosion by separating the corrodible metal from the water. Scale is formed when the divalent metallic cations associated with hardness (calcium and magnesium) combine with other minerals contained in water and precipitate to coat the pipe wall. Scale generally includes calcium carbonate ($CaCO_3$), but also magnesium carbonate ($MgCO_3$), calcium sulfate ($CaSO_4$), or magnesium chloride ($MgCl_2$). Water can hold a certain amount of a given chemical in solution. If more is added, it will precipitate instead of dissolve. The point at which no more chemical can be dissolved is the point of saturation. The saturation point depends on water quality, including pH, temperature, and total dissolved solids (TDS) (1).

Types of Corrosion

Different types of corrosion exist, which can be divided into two broad classes: uniform or localized, depending on the material to be corroded, system construction, and water characteristics. Localized corrosion resulting in pitting is produced after galvanic corrosion and concentration cell corrosion. Uniform corrosion takes place in an equal rate over the entire surface (3). The different types of corrosion are summarized in Table 2.

Microbial Corrosion

Bacteria adhere to the pipe surface and form a biofilm (4). Microbes can promote corrosion by creating areas with different concentrations in oxygen, hydrogen ions, minerals, and metals. These concentration differentials promote corrosion. Some microorganisms also catalyze reactions associated with the corrosion process. Iron precipitating bacteria (such as *Gallionella*) can convert Fe (II) to Fe (III) and influence the structure of Fe (III) precipitates (2,5). Organisms involved in the sulfur cycle in water also affect the corrosion process. Sulfate reducers have been found in tubercles under anoxic conditions. Bacteria involved in the nitrogen reaction affect the concentration of oxygen, leading to oxygen concentration cells that produce localized corrosion and pitting. Corrosion also protects bacteria from disinfection. Corrosion products offer a large surface area for microbial attachment. One corrosion product

Table 1. Different Material Types and Corrosion-Associated Problems in the Distribution Network

Material	Sources	Corrosion Problems
Iron and steel	Most common material in water systems	–Buildup of corrosion products on pipe walls and release or iron oxide products
Galvanized pipe	Oldest and most common plumbing material, quality varies	–Better service in hard water –Subject to galvanic corrosion –Leaching of zinc, iron, cadmium, and lead (impurities)
Lead	Lead service lines, lead pipe joints, gaskets	–Corrodes in soft water with pH<8 and hard water with high alkalinity levels –Leaching of lead
Copper	Most commonly used since World War II, excellent characteristics including good corrosion resistance, low cost, and ease of installation	–Subject to uniform corrosion, erosion, and pitting –Attacked by high velocities, soft water, chlorine, oxygen, low pH, and high alkalinity water –Copper leaching
Brass	Alloy of copper and zinc, in meters and plumbing fixtures	–Dezincification (selective removal of zinc) –Good corrosion resistance
Cement/asbestos	Asbestos/cement mains	–Release of asbestos fibers –Calcium can be leached from cement in contact with aggressive water –Increase of pH –Good corrosion resistance
Plastic	PVC pipes	–Resistant to corrosion

Table 2. Different Types of Corrosion

Type of Corrosion	Characteristics
Uniform corrosion	–Corrosion of a single metal –A site is anodic one instant and cathodic the next –Metal loss is uniform over the metal surface
Galvanic corrosion	–Occurs between 2 different types of metals in contact –One metal serves as anode and deteriorates, whereas the other acts as cathode and is protected (Anode: Zinc>aluminum>iron>cast iron>lead>brass>copper: cathode)
Pitting corrosion	–Localized attack, pitting may occur if imperfections in the metal or regions of high stress exist –Imperfection or high-stress area is anodic and surrounded by cathodic area –Corrosion occurs rapidly at point of failure –Chloride is associated with pitting
Tuberculation	–Occurs when pitting corrosion leads to a product buildup at the anode next to a pit
Crevice corrosion	–Form of localized corrosion caused by changes in acidity, decrease in oxygen, or dissolved iron
Erosion corrosion	–Removal of protective films and the pipe metal as well –Results from high flow velocities and turbulence
Cavitation corrosion	–Type of erosion corrosion –Water pressure drops, causing water to form water vapor bubbles that collapse with an explosive effect –Removal of protective coating on metal –Occur in pump impellers, partially closed valves, and reducers
Selective leaching	–Preferential removal of one metal
Stay-current corrosion	–Localized corrosion caused by grounding of home appliances to water pipes
Concentration cell corrosion	Occurs when concentrations of aqueous species (like oxygen) differ between two parts of the metal

(goethite, α–Fe–O–OH) quickly reacts and consumes chlorine. Chlorine is consumed before it can diffuse to the core of corrosion tubercles. The microorganisms within the tubercles are not exposed to lethal concentrations of disinfectant and thus are allowed to grow in the distribution systems. Corrosion products including iron oxides are also capable of adsorbing natural organic matter (NOM) from the bulk fluid (6). NOM accumulates on the surface of corrosion products. It was found that heterotrophic microorganisms found in drinking water supply could readily extract the NOM from the corrosion products for cell growth (7,8). In highly corroded environments, biofilm microorganisms can cause many problems such as bad taste and odor, slime formation, or coliform occurrences.

Consequences of Corrosion

Corrosion is one of the main problems in the drinking water industry. It can affect public health, water aesthetics, and operations. Corrosive water can leach toxic metals from distribution and household plumbing systems. Lead and cadmium may occur in tap water. US EPA promulgated the lead and copper rule in 1991 in order to reduce lead and copper levels in drinking water (9). The methods to reduce lead and copper are the following: removal of these metals from the source water (if present), implementation of a corrosion control program, replacement of lead service lines, and public education. The lead and copper rule defines an action level for the tap concentration of lead and copper higher than 15 ppb and 1.3 ppm, respectively (for the 90th percentile of the samples). Periodic sampling is required to monitor lead and copper concentrations at the customer's tap after leaving the water stagnant in the service lines for 6–8 hours. Moreover, optimal corrosion control water parameters (pH, calcium, alkalinity, temperature, inhibitor level) are defined for the plant effluent and distributed waters. When the lead or copper concentration is above the action level, the water utility has to implement a corrosion control program (9).

Copper in water can cause blue stains and a metallic taste, whereas zinc leads to a metallic taste. Corrosion of cast iron pipes can cause the formation of iron deposits called tubercles in the mains. Red water problems occur when iron is dissolved from cast iron by corrosive water. Iron stains plumbing fixtures, laundry, and makes water appear unappealing for drinking. Responding to customer complaints of colored water or bad taste is expensive in terms of money and public relations. Corrosion-caused problems that add to the cost of water include increased pumping costs after a buildup of corrosion products (tuberculation), uncontrolled scale deposit that can seriously reduce pipeline capacity, and increased resistance to flow. Aggressive water reduces the life of valves and can shorten the service life of plumbing fixtures and hot water heaters. Water leaks lead to loss of water and water pressure (1,3,10).

FACTORS IMPACTING CORROSION

The corrosion rates depend on many site-specific conditions and their interactions, including water characteristics and pipe conditions (2,3,10). The following section describes major factors impacting corrosion. Table 3 shows the effects of some chemicals present in water.

Table 3. Effects of Chemicals on Corrosion

Parameter	Effect on Corrosion
Hardness (measures calcium and magnesium)	Hard water is less corrosive than soft water because of the formation of a film of calcium carbonate ($CaCO_3$) on the pipe wall
Chloride (Cl^-)	Leads to pitting by causing the metal to stay in solution
Sulfate (SO_4^{2-})	Leads to pitting by causing the metal to stay in solution
Hydrogen sulfide (H_2S)	Accelerates corrosion
Ammonia (NH_3)	Increases corrosion rates
Organic Matter	Can increase or decrease corrosion rates

Temperature

Corrosion generally increases with temperature as temperature accelerates chemical reactions. Temperature changes the solubility constants and can favor the precipitation of different substances or transform the identities of corrosion products. These changes result in either more or less protection of the pipe surface, depending on the conditions. Temperature also affects the dissolving of $CaCO_3$. $CaCO_3$ tends to precipitate and form a protective coating more readily at higher temperatures. Temperature can affect the nature of corrosion. Corrosion that can be of pitting type at cold temperatures can become uniform at hot temperatures (2,3,10).

Flow/Velocity

The velocity of water increases or decreases corrosion rates depending on the conditions. When water is corrosive, higher flow velocities bring dissolved oxygen to the corroding surface more rapidly. For water with protective properties or containing corrosion inhibitors, high flow velocities aid in the formation of a protective film. At low velocities, the slow movement does not aid the effective diffusion of protective ingredients to the metal surface. High velocities can lead to the erosion of pipes, especially in copper lines. Stagnant waters in water main and house plumbing have been shown to promote tuberculation accompanied with biological growth (4).

Metal and Manufacturing Process

A metal that easily gives up electrons will corrode easily. When dissimilar metals are connected together, the metal corroding easier becomes an anode, whereas the metal resistant to corrosion becomes a cathode (defined as galvanic corrosion). The anode metal will corrode rapidly, whereas the cathode is protected. Manufacturing process can also impact the durability of the pipe, especially for galvanized piping.

Electrical Current

Corrosion is accelerated when an electrical current is passed through the metal. Improper grounding of household electrical systems or electric railway systems leads to electrical current in water pipes.

Dissolved Oxygen (DO)

Corrosion rate increases with increasing DO concentrations. Oxygen is the molecule accepting the electrons given up by the corroding metal. Oxygen also reacts with soluble ferrous iron ions to form ferric iron, which precipitates and forms a tubercle.

Total Dissolved Solids (TDS)

TDS levels are critical because electrical flow is necessary for corrosion to occur. Corrosion rates increase with increasing concentrations of TDS because water becomes a better conductor.

pH

pH is a measure of hydrogen ions (H^+). H^+ is a substance accepting the electrons given up by the metal. Generally, corrosion rate decreases as pH increases.

Alkalinity and Dissolved Inorganic Carbon (DIC)

Alkalinity measures the ability of water to neutralize acids or bases. The corrosion rates decreases as alkalinity increases.

Chlorine Residual

Gaseous chlorine lowers the pH of water by reacting with water to form hypochlorous acid (HOCl), hydrogen ion, and chloride ion. This reaction makes the water more corrosive. For low alkalinity water, the problem is greater because water has less ability to resist pH changes.

MEASUREMENT OF CORROSION

A comprehensive corrosion control program should include several techniques to monitor corrosion, because no single technique provides all the information on corrosion. The corrosion rates are expressed in mils (1/1000 inch) per year (MPY). Corrosion rates can be determined by weight loss method or electrochemical methods. Physical observations of a pipe exposed to water can be conducted, and corrosion indexes can be determined for given water (2,3,10). Table 4 presents different methods to monitor corrosion.

Coupon Weight Loss Method

A metallic coupon is inserted inside a main. The coupon method determines the average corrosion rate over a period of exposure, which is accomplished by weighing the coupon before and after exposure. The corrosion rate is calculated from the weight loss, initial surface area of the coupon, and time of exposition.

Loop System Weight Loss

This method uses a pipe loop or sections of a pipe for determining the effect of water quality on pipe material. Water flows through the loop under a continuous or intermittent flow to simulate the flow patterns of a

Table 4. Different Methods to Assess Corrosion Levels

Method	Advantages	Disadvantages
Coupon weight loss	–Provides corrosion over a specific period –Economical –Coupons placed in system	–Takes a long time –Does not show variations occurring during test –Analytical error because of weighing –Coupon may not be representative of system material
Loop system weight loss	–Actual pipe sections used –Loops placed in the system –Loop can be used to test different chemicals	–Takes a long time –Does not show variations occurring during test
Electrochemical rate measurements	–Simple and reliable –Instantaneous readings –Continuous monitoring possible –Gives uniform and pitting corrosion	–Relatively expensive –Gives corrosion rates for a particular material
Microscopic techniques	–Examination of particles in water and films on pipes –Determination of the morphology of corrosion products	–Require equipment
X-ray analysis and diffraction	–Identification of the elements or class of compounds present in corrosion products and films with possible quantification depending of the technique	–Require equipment
Corrosion indices	–Various indices available	–Have some limitations, can be misapplied –Useful after the fact but not to predict corrosion problems

household. Pipe sections are removed for weight loss measurement and inspection.

Electrochemical Methods

Several instruments are based on the electrochemical nature of metal corrosion in water. They are based on electrical resistance, linear polarization, or galvanic current. Electrical resistance probes measure the resistance of a thin metal probe; as corrosion causes metal to be removed from the probe, its resistance increases. Linear polarization resistance (LPR) is an electrochemical technique that measures the DC current through the metal/fluid interface, which results from polarization of one or two electrodes of a material after application of a small electrical potential. The corrosion current density, which corresponds to the current flowing in a corrosion cell per unit area, is related to the DC current. The galvanic current method measures corrosion of dissimilar alloys of metals.

Radiography Methods

X-ray emission spectroscopy and X-ray diffraction help characterize the corrosion scales.

Microscopic Methods

The deterioration of pipe surface can be evaluated using optical or scanning electron microscopes.

Chemical Analysis of Water and Corrosion Indexes

Water quality data can be used to calculate stability indexes or indicators of water corosivity (Table 5) (10–14). The Langelier Saturation Index (LSI) indicates whether a given water is likely to form or dissolve a protective film of calcium carbonate (11). The calcium carbonate precipitation potential (CCPP) estimates the quantity of $CaCO_3$ that can be precipitated from oversaturated water and the amount that can be dissolved by unsaturated water (12). Larson and Skold (13) studied the effects of chloride and sulfate ions on iron and proposed two indices. In general, it is very difficult to find a relationship between these indices and potential corrosion problems in the system.

METHODS OF CORROSION CONTROL

The complete elimination of corrosion is almost impossible. However, it is possible to reduce corrosion. As corrosion depends on both water quality and pipe characteristics, optimal corrosion control methods are site specific. Three basic approaches to control corrosion exist: (a) modify water quality so that it is less corrosive to the metal, (b) place a protective barrier between the water and pipe, and (c) use pipe material and design the system so that it is not corroded by water. Methods used to achieve corrosion control involve modifying water quality (changing pH and alkalinity), forming a calcium carbonate coating, using corrosion inhibitors, providing cathodic protection, and using a corrosion-resistant coating (3,10).

pH Adjustment

Adjusting the pH is one of the most common methods of corrosion control. As most metals used in the distribution system tend to dissolve more readily at lower pH (presence of H^+), an increase in pH and alkalinity levels can reduce corrosion by reducing the solubility of metals. The optimum pH is related to water and system characteristics; it is generally above 7.0. Various chemicals can be used in corrosion control treatment (Table 6). For example, lime is commonly used to increase both pH and alkalinity. It is less expensive than the other chemical products. However, lime softening can cause severe scale problems

Table 5. Indexes of Water Corrosion

Index	Formula
Langelier Saturation Index (LSI)	$LSI = pH_{actual} - pH_s$ pH_{actual} : pH of water sample $pH_{s:}$ Theorical pH at which water is saturated with calcium carbonate LSI = 0: stable water, LSI < 0: corrosive water, calcium carbonate dissolves, LSI > 0: calcium carbonate precipitates
Calcium carbonate precipitation potential (CCPP)	$CCPP = 50{,}000\ (Talk_i - Talk_{eq})$ (in mg of $CaCO_3$/L) $Talk_i$: initial total alkalinity $Talk_{eq}$: Total alkalinity at equilibrium
Larson Index	$LI_1 = \dfrac{2[SO_4{}^{2-}] + [Cl^-]}{[HCO_3{}^-]}$ The brackets indicate the concentration in mole/L $LI_2 = \dfrac{[Cl^-]}{[HCO_3{}^-]}$
Ryznar Stability Index (RSI)	RSI = pHs-pH RSI = 6.5–7: water at saturation RSI < 6.5: Scale forming water RSI > 7: Undersaturated water

Table 6. Various Chemicals Used to Adjust pH

Chemical	Effect	Addition Equipment
Lime, $Ca(OH_2)$	Increases pH, alkalinity, and calcium levels	Dry storage, gravimetric or volumetric dry feeders, slurry feed
Caustic soda, NaOH	Raises pH and converts CO_2 to alkalinity species	50% solution, metering pumps
Soda Ash, Na_2CO_3	Increases alkalinity with moderate pH increase	Dry storage with solution feed
Sodium bicarbonate, $NaHCO_3$	Increases alkalinity	Dry storage with solution feed
Sulfuric acid, H_2SO_4	Lowers pH	Metering pumps
Carbon dioxide, CO_2	Lowers pH, converts hydroxyls to bicarbonate and carbonate	Pressurized gas feeder

when it is not stabilized. Stabilization after softening can be accomplished by feeding carbon dioxide or sulfuric acid to decrease pH so that calcium carbonate does not precipitate in the distribution system (1).

Formation of a Calcium Carbonate Coating

Protective coating can be applied by controlling the chemistry of water. A common protective coating technique is to adjust the pH of water to a level just above the saturation of calcium carbonate. This process has to be closely controlled because a pH that is too low may result in corrosion and a pH too high may lead to precipitation and cause plugging of the lines.

Corrosion Inhibitors

Some waters do not contain enough calcium or alkalinity to lead to the formation of a coating. Corrosion inhibitors used in potable water act by forming a protective scale over anodic or cathodic sites. These films are commonly inorganic precipitates containing the ions added as inhibitors. They provide a barrier between the water and the pipe. Chemical inhibitors include phosphates and silicates. The success of an inhibitor in controlling corrosion depends on three requirements. First, when treatment is initiated by adding two to three times the normal concentration of inhibitor to build up a protective film rapidly. Several weeks are required for the film to develop. Second, the inhibitor must be fed continuously at a proper concentration. An interruption in chemical addition can lead to the dissolution of the film. Third, water flow rates must be adequate to transport the inhibitor to all parts of the system, otherwise the protection film will not form (3).

Several types of phosphates are used for corrosion control, including linear and cyclic polyphosphates, orthophosphates, glassy polyphosphates, and bimetallic phosphates. It is also possible to use zinc along with polyphosphates and orthophosphates, or blends of ortho and polyphosphates (15). The choice of a particular phosphate product depends on the water and the distribution system characteristics as well as the utility treatment goals. Some phosphates work better than others for a given system (3). It is recommended to conduct laboratory and pilot testing to evaluate the effectiveness of different products. Phosphate inhibitors need particular pH, alkalinity, and concentrations to be effective; the balance required is poorly understood. Orthophosphates seem to be very effective for a wide range of plumbing materials (15). It can be used alone (phosphoric acid, H_3PO_4) or in combination with zinc. Commercial formulations of orthophosphates exist that contain various levels of zinc. Zinc orthophosphate products are used primary in water where the potential for formation of $CaCO_3$ scale is low (soft and slightly acidic water). When using zinc phosphates, limitations of metal levels in wastewaters can limit the use of products with a high zinc content. Zinc orthophosphate leads to the formation of a zinc phosphate scale (16). When zinc is not available with phosphates, a protective scale still forms, such as ferrous or ferric phosphate scale on iron pipes.

Phosphate is most effective in the absence of prior scale formation because phosphates must be part of the precipitate to be effective. When fed into the system, phosphates soften the previously formed scales, causing red water and bacteriological problems as the scale washes out of the system. Orthophosphates can reduce corrosion rate of iron, lead, zinc, and galvanized pipes; although it is not considered to be very effective in preventing copper corrosion (15,17).

A distinction should be made between orthophosphates and polyphosphates (10). Polyphosphates have been used to control corrosion in cast iron pipes; however, little evidence exists of any beneficial effects of polyphosphates on lead corrosion. Some studies (18,19) showed that lead levels could increase in water after solubilization of protective films on the pipes. Some polyphosphates can be used as sequestering agents to keep in solution scale-forming ions (calcium and magnesium) and iron. These polyphosphate products tie up with iron and prevent red water (3).

Inorganic silicates have also been used to reduce corrositivity; they lead to the formation of a protective film onto various metal surfaces. Silicates inhibit corrosion in asbestos cement pipes, where its effectiveness may be attributed to a surface catalyzed conversion to quartz (20). Soluble silicates are adsorbed onto the metal pipe surface at the anodic area and form a thin film. High dosages (20 mg/L) are required during the first 30–60 days of treatment. Then, doses of 4–8 mg/L are added in the system. Silicates have not been widely used because their effects on corrosion are debatable (21,22).

Cathodic Protection

This process limits corrosion of metallic structures. It is used to prevent internal corrosion in water storage tanks. It consists of using an inert electrode (such as high silicon cast iron or graphite) powered by an external source of current. The current forces the inert electrode to act as an anode, preventing the metal to be protected to become an anode and corrode. Another method involves a magnesium or zinc electrode acting as a galvanic anode. The electrode produces a galvanic action with iron; it is sacrificed and corrodes, whereas the iron is protected from corroding.

Lining, Coating, and Paints

A protective coating can keep corrosive water away from the pipe surface and storage tanks. Some linings include coal tar enamels, epoxy paint, cement mortar, polypropylene, or vinyl. These linings are applied when pipes are manufactured or in the field.

UTILITY EXPERIENCE WITH CORROSION CONTROL

Figure 2 shows corrosion rates as a function of temperature and corrosion inhibitor levels in two pilot systems (23). The first system received a constant dose of phosphate over the entire study (system fed after conventional treatment of the Mississippi water, IL-American Water, E. St. Louis), whereas the second system received changing inhibitor levels (test system to optimize treatment). Corrosion rates were strongly related to water temperature (and/or other seasonal factors). For the plant condition system, corrosion rates could vary up to 7 mpy, even when the plant was feeding a corrosion inhibitor (constant doses of 0.86 mg PO_4/l over the year). Higher phosphate levels in the test reactor resulted in low corrosion rates, especially in the summer. Corrosion rates were maintained below 3 mpy when phosphate dosages were slightly increased (between 1.5 and 2 PO_4/l), especially during warm periods. For this site, a seasonal corrosion control strategy was developed that would require slightly higher corrosion inhibitor concentrations during the summer and possibly lower dosages during winter months, rather than using a constant concentration over the entire year (23).

CONCLUSION

As distribution system pipes are in place for long periods of time ($\gg$50 years), corrosion control is critical to maintain microbial and aesthetic water quality and pipe integrity. Water utilities should set corrosion control goals and monitor corrosion rates on a regular basis to determine seasonal changes in corrosivity and adjust corrosion control programs to prevent excessive corrosion.

BIBLIOGRAPHY

1. American Water Works Association (AWWA). (1995). *Water Treatment*, 2nd Edn. AWWA, Denver, CO.
2. AWWA Research Foundation. (1996). Internal Corrosion of Water Distribution Systems. *AWWARF-DVGW-TZW Cooperative Research Report*. Denver, CO, p. 586.
3. Schock, M. (1990). Internal corrosion and deposition control. In: *Water Quality and Treatment*, 4th Edn. McGraw-Hill, New York.
4. Geildreich, E.E. (1996). *Microbial Quality of Water Supply in Distribution Systems*. CRC Press, Boca Raton, FL.
5. Kolle, W. and Roshe, H. (1980). *Von Wasser.* **55**: 159–167.
6. Tipping, E. and Cooke, D. (1982). *Geochimica et Cosmochimica Acta.* **46**: 75–80.
7. Chang, Y., Li, C., and Benjamin, M.M. (1997). *J. Amer. Water Works Assoc.* **89**: 100–113.
8. Abernathy, C.G. and Camper, A.K. (1997). Interactions between pipe materials, disinfectants, corrosion inhibitors, organics and distribution biofilms. *Proc. AWWA WQTC*. Denver, CO.
9. U.S. EPA. (1991). Maximum Contaminant Level Goals and NPDWR for Lead and Copper. *Federal Register*.
10. Faust, S.D. and Aly, O.M. (1999). *Chemistry of Water Treatment*, 2nd Edn. Lewis Pubilsher, Boca Raton, FL.
11. Langelier, W.F. (1936). *J. Amer. Water Works Assoc.* **28**(10): 1500–1505.
12. Rossum, J.R. and Merrill, D.T. (1984). *J. Amer. Water Works Assoc.* **76**(8): 72–82.
13. Larson, T.E. and Skold, R.V. (1957). *J. Amer. Water Works Assoc.* **49**(10): 94–97.
14. Ryznar, J.W. (1944). *J. Amer. Water Works Assoc.* **36**: 472–479.
15. Benjamin, M.M., Reiber, S.H., Fergusson, J.F., Van der Werff, E.A., and Miller, M.W. (1990). *Chemistry of Corrosion Inhibitors in Potable Water*. AWWA Research Foundation, Denver, CO.
16. Swayze, J. (1983). *J. Amer. Water Works Assoc.* **75**: 101–102.

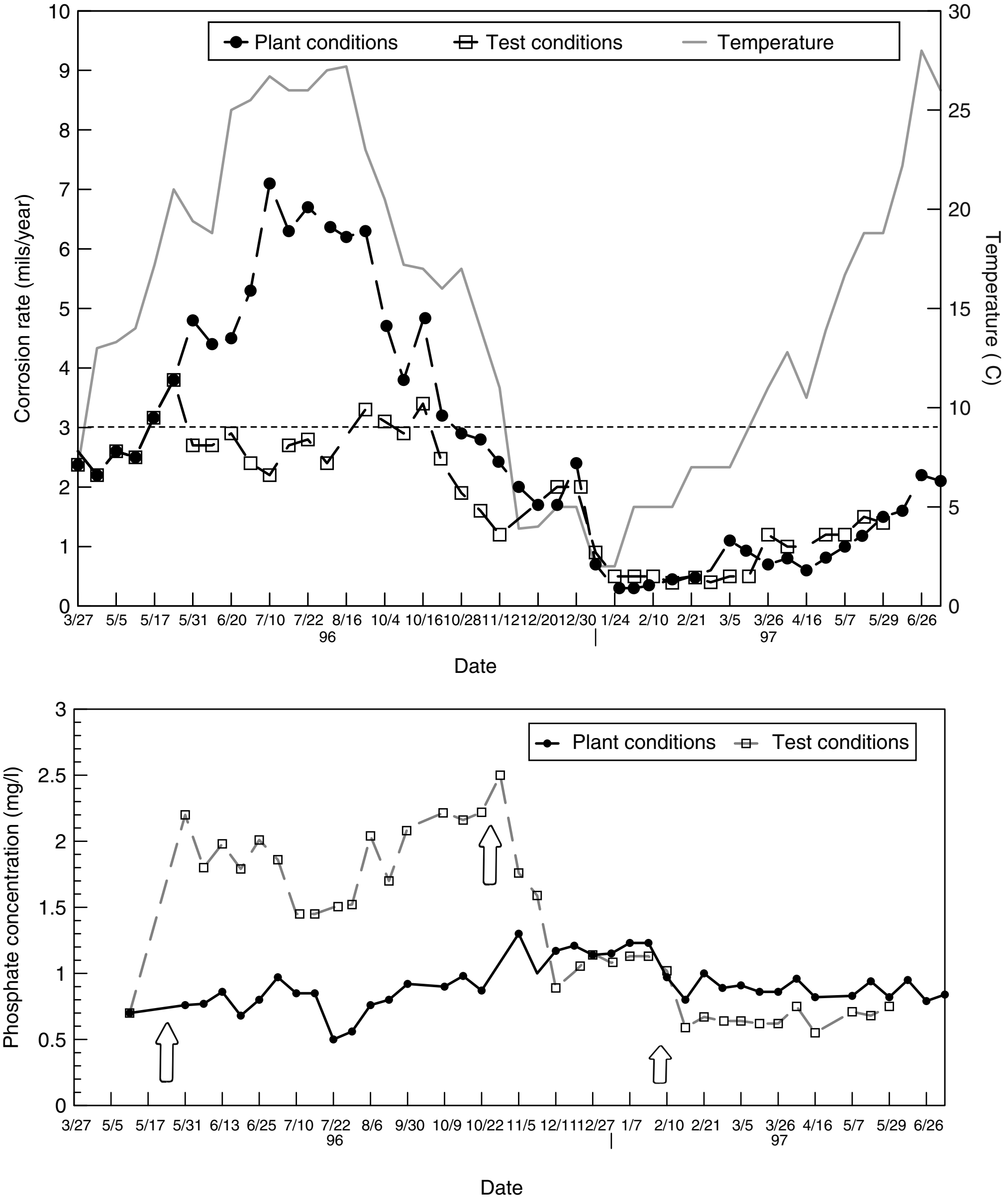

Figure 2. Corrosion rates as a function of corrosion inhibitor levels and temperature (the arrows show the changes in inhibitor levels to optimize corrosion control).

17. Ryder, R. and Wagner, I. (1985). Corrosion inhibitors. In: *Internal Corrosion of Distribution Systems*. AWWA Research Foundation, Denver, CO.
18. Schock, M.R. (1989). *J. Amer. Water Works Assoc.* **81**(7): 88–95.
19. Holin, T.R. and Shock, M.M. (1991). *J. Amer. Water Works Assoc.* **83**(7): 76–83.
20. Schock, M.R. and Buelow, R.W. (1981). *J. Amer. Water Works Assoc.* **73**: 636–751.
21. Sheiham, I. and Jackson, P.J. (1981). *J. Inst. Water Engrs. Scientists.* **35**(6): 491–496.
22. Rompre, A., Prevost, M., Brisebois, P., Lavois, J., and Lafrance, P. (1997). Comparison of corrosion control strategy efficiency and their impacts on biofilm growth. *Proc. AWWA Water Quality Tech. Conf.*, Denver, CO.
23. Volk, C., Dundore, E., Schiermann, J., and LeChevallier, M. (2000). *Water Res.* **34**: 1967–1974.

ECONOMICS OF RESIDENTIAL WATER DEMANDS

Steven J. Renzetti
Brock University
St. Catharines, Ontario, Canada

This entry examines what is known regarding the economic characteristics of residential water use and how this type of information can be applied in managing water resources. Understanding the economic dimension of household water use is important for several reasons. First and most importantly, all households on the planet share the need for access to potable water supplies. On the other hand, households have very different degrees of access to safe water. For most of North America and Europe, access to clean water is almost taken for granted by many households. In contrast, a great number of households in low-income countries do not have access to reliable supplies of potable water. As the World Bank contends, "The challenge is enormous: one billion people still lack access to safe water, two billion lack safe sanitation. Slow progress is not acceptable, as more than three million children still die every year from avoidable water-related disease" (World Bank Water Supply and Sanitation web-site, September 1, 2001 http://www.worldbank.org/html/fpd/water/). Second, the agricultural sector dominates global water use, but residential and industrial demands have grown much more rapidly during the twentieth century. Third, empirical evidence indicates that the amount of water used by a household is a complex function of a number of influences, including the price of water (both in monetary terms and in time costs), household income, and household characteristics (for example, the number of residents and water-using appliances).

THE ECONOMIC PERSPECTIVE ON RESIDENTIAL WATER DEMANDS

Households use water for personal hygiene, waste removal, cooking, cleaning and a number of outdoor applications (lawn and garden watering, etc.) Table 1, Fig. 1. Except in some arid portions of the United States, the bulk of residential water use occurs indoors and is related to personal hygiene and waste removal (4). In North America, for example, these two applications account for 60% to 70% of the average household's daily indoor water use. Another important feature of residential water use is its cyclical character. On a daily basis, household water use typically displays two peak periods of use: early morning and early evening. On an annual basis, residential water use in summer months is usually substantially higher, than during the rest of the year due to the increase in outdoor water use. Hanemann (5), for example, indicates that for households in the Western states of the United States, total summer water use exceeds total winter water use by 50–60%. The multiplicity of water's uses and the cyclical nature of water use suggest that potentially a number of factors influence a household's water use decisions.

Economic theory provides a useful framework within which the nature of residential water use may be examined. Economists usually assume that a household's demand for any good such as potable water is, in the most general case, a function of all of the prices facing a household as well as its income and demographic characteristics. The household's estimated demand for water may be used to predict household consumption levels and to predict how households will respond to changes in the price of water. This degree of responsiveness is captured by a variable known as "elasticity." This variable is measured as the ratio of the percent change in the demand for a good to the percent change in a price or income. A household's demand is said to be 'inelastic' if the estimated elasticity is less than one. For example, if the price elasticity of demand equals −0.5, then a 10% increase in price, it is predicted, with to induce a 5% decrease in water use. The empirical water demand literature seeks to estimate households' price and income elasticities of water demand as well as those households' valuation of access to potable water supplies.

Table 1. Residential Water Use in 1995[a]

Country	Annual Residential Water Use (m^3/person)
Chad	4
Nigeria	9
China	21
India	25
Greece	45
Israel	45
South Africa	49
Argentina	67
Uruguay	73
France	94
Japan	137
Canada	157
U.S.A.	203

[a] Ref. 1.

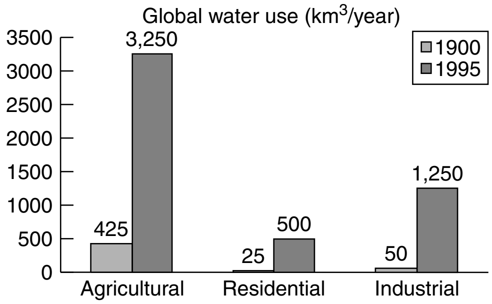

Figure 1. Global water use (1,2).

There are several ways in which the general model of household demand has been altered to reflect conditions that are important in the consumption of water. These include situations when households' choice of water use is constrained by their stock of water-using appliances and the size of the house and property (6), when households must choose not only the quantity of water to consume but also the source of its potable water (7), and when households face water prices that are complex functions of the quantity of water used. The last of these extensions is quite important because of the growing frequency of complex price schedules facing households and because of the implications of this type of price structure for the statistical estimation of residential water demands (5,8).

EMPIRICAL MODELS AND RESULTS

Residential water use has received a substantial amount of attention from economic researchers. There are a number of surveys that summarize the theoretical and empirical research (4,9–11). The main goal of statistical models of water demand is to estimate the relationship between the observed quantity of residential water use and the explanatory variables (such as the price of water) suggested by economic theory. During the evolution of the residential water demand literature, a number of issues have been at the forefront, including the appropriate definition of the price of water, the choice of statistical technique, and the role of other explanatory variables (such as climate and household characteristics).

The modern residential water demand literature begins with the work of Howe and Linaweaver (12). The authors estimate demand models for domestic (indoor) and sprinkling (outdoor) water demands. The demand models are further divided according to climatic zones. The authors assume that the quantity of water demanded by the average household is a function of water and sewage prices, age and value of property, the number of people in the household, and climate conditions. The estimation models indicate that indoor water use is responsive to the price of water to a limited degree (price elasticities ranging from -0.214 to -0.231) and outdoor water use is more responsive to prices (price elasticities ranging from -0.438 to -1.57). Income elasticities follow the same pattern: 0.314 to 0.378 for indoor water use and 0.447 to 1.45 for outdoor.

Since the work of Howe and Linaweaver, statistical models of residential water demand have become more sophisticated, and data on household water use and characteristics have become more detailed and comprehensive. A particularly important issue has been the manner in which the price of water is represented. When prices are set out as complex functions in which prices can increase or decrease with the amount purchased, researchers must take more care in measuring the influence of prices on water use (the variety of statistical methods used to do this is reviewed in Ref. 11).

A number of other factors, it has been found, are influential in determining household water demand. Researchers have found positive relationships between water use and family size, property value, and household income (4). In Hanemann's (5) comprehensive listing of estimated water demand elasticities, the average of the income elasticities reported is 0.52. Another important set of factors that, it has been found, influences residential water demand is related to climate. There is a consensus in the literature that increases in temperature or evaporation rates lead to higher residential water demand, whereas increases in precipitation have the opposite effect. However, indoor residential water use appears to be relatively insensitive to weather conditions, and most of the influence of climate on residential water use works through outdoor water uses.

LOW-INCOME COUNTRIES

Households in many low-income countries face a variety of challenges in their efforts to secure potable water. In large urban centers, water supply systems are often unreliable and fail to provide service to many of the city's poorest residents (13). Residents in rural areas frequently confront even greater risks from contaminated and distant water supplies (14). Households in low-income countries also differ from their high-income counterparts in that they spend a larger share of their income on water. They also may face a different set of circumstances regarding their supply of potable water. For example, it is common for households in low-income countries to have several possible sources of potable water. Mu, Whittington and Briscoe (14) report that the members of a small town in Kenya choose among private pipe, communal wells and pumps, kiosks, and water vendors for their potable water. These sources differ in their relative cost, convenience, reliability, and quality. It is important to note that the cost of each alternative is a combination of the time spent obtaining water, installation charges, continuing connection charges, and prices. Households face the challenge of deciding which supply source(s) to use as well as how much water to obtain from each source, and researchers face the challenge of understanding and modeling these decisions (15).

There have been a number of attempts to model, first, household decision-making regarding its choice of the source of its drinking water and, second, households' valuation of improved access to reliable water supplies in low-income countries. Researchers have conducted surveys that collect information on households' supply–source choices and characteristics. For example, Madanat and Humplick (16) examine the behavior of 900 households in Faisalabad, Pakistan. The authors conclude "the more expensive the in-house pipe connection relative to the other sources, the less likely the household is to connect" (p. 1337). In addition, the connection decision is strongly influenced by household expectations regarding the relative reliability and quality of alternative sources as well as their past experiences with alternative supply sources.

With respect to the value they assign to access to safe water, many households in low-income countries find themselves in what some researchers have termed a "low-level equilibrium trap" (17, p. 1931). By this, the authors mean that the existing water supply system has few connections, low prices, low revenues, low maintenance, poor reliability, and low usage by households. One of the ways to improve this situation is to demonstrate that the value of improved service to households (as expressed in their willingness to pay through higher prices) exceeds the costs of improving service.

The World Bank Water Research Team (15) is the most extensive effort to assess the preferences of households in low-income countries for connection to water supply systems. The authors find that household income is positively linked to the demand for improved services, but the link between the two is not strong. Households with higher levels of education are also willing to pay more for improved access. Gender is often an important factor in explaining willingness to pay, but "the direction of the influence depends on the specific cultural context" (p. 53). In Tanzania and Haiti, female respondents' willingness to pay exceeds that of males, but in Nigeria and India, the reverse is true. As expected, economic variables play an important role, and higher connection charges and monthly prices reduce demand. Finally, households are willing to pay more for private connections and for sources higher expected quality levels and reliability.

RESIDENTIAL WATER DEMAND MANAGEMENT

Historically, the challenge of managing water resources and providing potable water has concentrated on the engineering task of conveying and treating water so that it would be available for household use. However, as the construction and operation of water distribution and treatment systems becomes increasingly expensive (in monetary and environmental terms), attention has turned to the idea of balancing the costs of developing water supplies with the benefits derived from consumption. From this adoption of a more balanced perspective has come a renewed interest in using the information contained in estimated residential water demand to assist in planning and operating water delivery systems.

Estimated residential water demand equations provide information that can be used in a variety of ways to promote water conservation and in the development and operation of water supply systems. First, the structure of residential water demand can provide water utility operators with information regarding the relative efficacy of alternative policy instruments aimed at promoting water conservation. For example, Renwick and Archibald (18) examine the factors influencing Californian households' adoption of water conserving technologies such as low-flow toilets and showerheads. The results of the authors' empirical model suggest that both price and nonprice measures reduce household water use, although their impact varies across households.

Forecasting future water use is a second way in which demand information is used. Dziegielewski (19) provides a brief review of the history of urban water demand forecasting. In the "traditional" method of forecasting that has dominated historically, total future demand is predicted as the product of expected population growth and a fixed per capita water use coefficient. This method was not very accurate as it neglected other influences (such as prices) on water use. The development of the municipal and industrial needs (MAIN) model represents a major change in forecasting methods. The MAIN model disaggregates total water use into a large number (approximately 400) of categories and locations. The factors that influence water demand for each category are determined through statistical analysis. Changes in these explanatory variables (such as income, climate, and energy prices) translate into anticipated changes in water use and, in turn, form the basis for water demand forecasts. A third way in which water demand information is used in assessing the construction and operation of water delivery systems. For example, the costs of improving the reliability of a water supply system can be compared with households' valuation of that increase in reliability (20).

CONCLUSIONS

All households share the need for access to potable water, but actual consumption levels vary significantly due to differences in income, water prices, proximity to reliable water supplies, climate, and a variety of other factors. Economic models of household decision-making regarding water use indicate that households are influenced by these factors and that water prices, income levels, and climate play particularly influential roles. A different line of research highlights the challenges faced by households in low-income countries in their efforts to acquire potable water. One of the facets of this situation that has received attention recently concerns households' decision-making when confronted with more than one source of potable water. As predicted by economic theory, most households consider the relative quality, reliability, and cost of alternative sources when making their choices. The last topic considered here is the use of information regarding the economic features of residential water demand to encourage water conservation. Research indicates that both price and nonprice based conservation programs are effective in curtailing demands.

BIBLIOGRAPHY

1. Gleick, P. (2000). *The World's Water 2000–2001*, Washington, DC, Island Press.
2. Abramovitz, J. (1996). *Imperiled Waters, Impoverished Future: The Decline of Freshwater Ecosystems*, WorldWatch paper 128, Washington, DC, Worldwatch Institute.
3. Biswas, A.K. (1997). Water development and environment. In: *Water Resources: Environmental Planning, Management and Development*. A.K. Biswas (Ed.). McGraw-Hill, New York, pp. 1–37.
4. Baumann, D.J., Boland, and Hanemann, W.M. (1998). *Urban Water Demand Management and Planning*. McGraw-Hill, New York.
5. Hanemann, W.M. (1998). Determinants of urban water use. In: *Urban Water Demand Management and Planning*. D.J. Baumann, Boland, and W.M. Hanemann. McGraw-Hill, New York, pp. 31–75.
6. Lyman, R.A. (1992). Peak and off-peak residential water demand. *Water Resour. Res.* **28**(9): 2159–2167.
7. Asthana, A. (1997). Where the water is free but the buckets are empty: Demand analysis of drinking water in rural India. *Open Economies Review* **8**(2): 137–149.
8. Shin, J. (1985). Perception of price when information is costly: Evidence from residential electricity demand. *Rev. Econ. Stat.* **67**(4): 591–598.
9. Hanke, S. and de Maré, L. (1984). Municipal water demands. In: *Modeling Water Demands*. J. Kindler and C. Russell, (Eds.). Academic Press, London, pp. 149–170.
10. Gibbons, D.C. (1986). *The Economic Value of Water*, Resources for the Future, Washington, DC.
11. Renzetti, S. (2002). *The Economics of Water Demands*, Kluwer Academic, Norwell, MA.
12. Howe, C. and Linaweaver, F. (1967). The impact of price on residential water demand and its relation to system design and price structure. *Water Resour. Res.* **3**(1): 13–31.
13. Rivera, D. (1996). *Private Sector Participation in the Water Supply and Wastewater Sector: Lessons from Six Developing Countries*. The World Bank, Washington, DC.
14. Mu, X., Whittington, D., and Briscoe, J. (1990). Modeling village water demand behavior: a discrete choice approach. *Water Resour. Res.* **26**(4): 521–529.

15. World Bank Water Research Team. (1993). The demand for water in rural areas: Determinants and policy implications. *World Bank Res. Obs.* **8**(1): 47–70.
16. Madanat, S. and Humplick, F. (1993). A model of household choice of water supply systems in developing countries, *Water Resour. Res.* **29**(5): 1353–1358.
17. Singh, B., Ramasubban, R., Bhatia, R., Briscoe, J., Griffin, C., and Kim, C. (1993). Rural water supply in Kerala, India: How to emerge from a low-level equilibrium trap. *Water Resour. Res.* **29**(7): 1931–1942.
18. Renwick, M. and Archibald, S. (1998). Demand side management policies for residential water use: Who bears the conservation burden? *Land Econ.* **74**(3): 343–59.
19. Dziegielewski, B. (1996). Long-term forecasting of urban water demands. In: *Marginal Cost Rate Design and Wholesale Water Markets, Advances in the Economics of Environmental Resources*, Vol. 1, D. Hall (Ed.). JAI Press, Greenwich, CT, pp. 33–47.
20. Howe, C. and Griffin Smith, M. (1993). Incorporating public preferences in planning urban water supply reliability. *Water Resour. Res.* **29**(10): 3363–3369.

GRAY WATER REUSE IN HOUSEHOLDS

Fayyaz A. Memon
D. Butler
Imperial College
London, United Kingdom

Gray water is a loosely defined term representing discharges from wash basins, baths, showers, dishwashers and washing machines. This generally excludes wastewater from kitchen sinks and toilets, commonly known as black water. Gray water accounts for about 50% of the total household water consumption activities (Fig. 1).

The quality of gray water depends on factors, including the habits and affluence of the water users, the types of products used for clothes and personal washing, and the nature of the substances disposed of through sinks and other appliances. Substances found in gray water include detergents, shaving foam, toothpaste, soap, hair, body oils, and dried skin residues. Small amounts of fecal material arising from washing of baby diapers and traces of urine are also present in gray water. These pollutants exert oxygen demand and contain some disease causing microorganisms. Typical gray water pollutant concentrations from different sources are shown in Table 1. The average pollutant concentration measured in the effluent from different appliances housed in a residence hall (for example) is shown in Table 2.

Freshly produced gray water usually does not have any objectionable odor. Compared to black water, gray water has a relatively higher temperature and readily degradable pollutants. Therefore, it requires immediate treatment after collection. If stored untreated for long periods, oxygen deficient conditions will develop, and scum will be formed that can float or sink in the collection tank. Experiments have shown that bacterial population also increases with increased storage time (5).

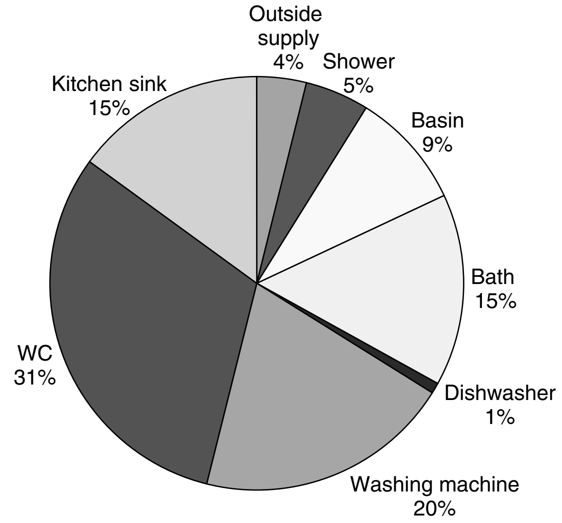

Figure 1. Components of domestic water use (2).

Table 1. Gray Water Quality from Various Sources[a]

	BOD (mg/l)	COD (mg/l)	Turbidity (NTU)	NH_3 (mg/l)	P (mg/l)	Total Coliforms
Single person	110	256	14	–	–	–
Single family	–	–	76.5	0.74	9.3	–
Block of flats	33	40	20	10	0.4	1×10^6
College	80	146	59	10	–	–
Large college	96	168	57	0.8	2.4	5.2×10^6

[a] Ref. 3.

Before gray water is reused, a certain level of treatment is required to minimize aesthetic concerns and the potential for health risk. Table 3 shows a summary of gray water quality criteria for toilet flushing followed in different countries. The level of treatment required depends on the scale and purpose of use. On a small scale (domestic level), a two-stage treatment consisting of filtration of coarse pollutants (hair and suspended impurities) followed by disinfection with chlorine, bromine, or UV is probably sufficient. On larger scales (hotels, commercial buildings), more complex and expensive methods of treatment could be employed.

Domestic gray water recycling systems, normally employed, produce water for toilet flushing. A recycling system (Fig. 2), typically, consists of an underground collection tank and an overhead distribution tank to supply toilet cisterns. The collection tank is designed to prevent groundwater contamination and ingress and is sized to accommodate water volumes intended for reuse. The optimal size of the collection tank has been modeled by Dixon (6), and systems storing 100–200 liters are considered sufficient for a family of five persons (7). Any excess gray water is diverted to the sanitary drain (i.e., the drain going out of the household). Devices are installed to prevent back-flow from the foul drain to the tank. Filtration is typically carried out at the tank inlet. The clogged filters are either replaced or cleaned using water jets. A submersible pump fitted with a float switch

Table 2. Average Pollutant Concentration in Gray Water Measured in a Residence Hall[a]

Parameter	Bath/ Shower	Washbasin	Washing Machine	Laundry and Dishwashing
BOD (mg/L)	216	252	472	110
COD (mg/L)	424	433	725	–
Ammonia as N (mg/L)	1.56	0.53	10.7	–
Phosphate as P (mg/L)	1.63	45.5	101	–
Total coliforms (cfu/100 mL)	6×10^6	5×10^4	7×10^5	5×10^6
Faecal coliforms (cfu/100 mL)	600	32	728	462
Turbidity (NTU)	92	102	108	148
Inorganic carbon (mg/L)	26	20	25	20
TOC (mg/L)	104	40	110	84
Total solids (mg/L)	631	558	658	538
Suspended solids (mg/L)	76	40	68	90
Dissolved solids (mg/L)	559	520	590	449
Volatile solids (mg/L)	318	240	330	277
pH	7.6	8.1	8.1	7.8
Copper (mg/L)	111	–	322	–
Lead (mg/L)	3	–	33	–
Zinc (mg/L)	59	–	308	–
Cadmium (mg/L)	0.54	–	0.63	–

[a] Ref. 4.

Table 3. International Water Quality Criteria for Toilet Flushing[a]

	Fecal coliforms (cfu/100 mL)	Total coliforms (cfu/100 mL)	BOD (mg/L)	Turbidity (NTU)	TSS	DO% (% saturation)	pH	Cl_2 residual (mg/L)
US EPA (g)	14 for any sample	–	10	2	–	–	6–9	1
	0 for 90% samples	–	–	–	–	–	–	–
Florida (m)	25 for any sample	–	20	–	5	–	–	1
	0 for 75% samples	–	–	–	–	–	–	–
Texas (m)	75 (m)	–	5	3	–	–	–	–
Germany (g)	100 (g)	500 (g)	20 (g)	1–2 (m)	30	80–120	6–9	–
Japan (m)	10 for any sample	10	10	5	–	–	6–9	–
South Africa (g)	0 (g)	–	–	–	–	–	–	–
WHO lawn irrigation	200 (g)	–	–	–	–	–	–	–
	1000 (m)	–	–	–	–	–	–	–
EC bathing water	100 (g)	500 (g)	2 (m)	–	–	80–120	6–9	–
	2000 (m)	10000 (m)	(g)	–	–	–	–	–
	–	–	1 (m)	–	–	–	–	–
	–	–	(m)	–	–	–	–	–
UK (BSRIA)	14 for any sample	–	–	–	–	–	–	–
Proposed (g)	0 for 90%	–	–	–	–	–	–	–

[a] Ref. 4.
g = guideline.
m = mandatory.

is normally used in the tank to transfer filtered water to the overhead tank. This then contains the disinfectant feeding arrangement and switches to control the water level. When the water volume in the overhead tank drops below a certain level, the pump turns on and stops when the water level in the tank reaches the design level. The overhead tank normally has an inlet to provide a top up supply of mains water when the treated gray water is not sufficient to meet the demand or the recycling system is inoperative. An air gap is typically provided between the inlet pipes of gray water and mains water. The pipework carrying gray water needs to be clearly marked or colored differently to avoid cross connections and contamination of potable water. The underground and overhead tanks will be designed to drain down fully to avoid problems of prolonged gray water storage. There are several packaged recycling systems available on the market. Each offers a different degree of treatment and safety controls. Experience has shown that they are not 'fit and forget' systems but require monitoring to ensure their smooth operation. Therefore, a clear warning mechanism that can show the failure of system components should be installed within the household, and clear maintenance instructions provided.

Gray water recycling is not a problem-free option, and particularly, issues related to health risk must

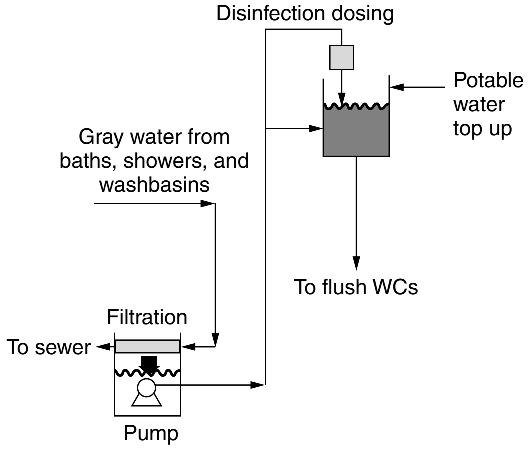

Figure 2. Water flows and components of a typical gray water recycling system.

be addressed. Although most waterborne pathogens are killed by conventional disinfectants, there are certain species (e.g., *Legionella* spp.) that are resistant to normal modes of treatment. *Legionella pneumophila*, a naturally occurring bacterium in domestic hot water supplies, showerheads, cooling waters, and other water services in buildings, has been linked to outbreaks of Legionnaires' disease. Surface fouling, biofilm formation, slow moving or stagnant waters, and increased temperatures are favorable conditions for *Legionella* growth. Therefore, recycling systems could provide an ideal environment for their growth. *Legionella* has an infection route through inhalation, and it is suspected that some bacteria may be inhaled through water vapors during toilet flushing (8). Fortunately, however, research shows that the *Legionella* count in gray water is typically low (3). To avoid in-pipe water stagnation for a prolonged period, the recycling system should be kept free from 'dead legs.' Past case studies on gray water recycling systems suggest that the relative health risk from gray water reuse is not high if it is properly treated and does not come in direct contact with users (8). For additional safety, the use of treated gray water should be discontinued when any of the users living in the household is ill. Spray irrigation with gray water is not advised because it will increase the bacterial exposure potential. There is also concern about the use of chlorine as a disinfectant. It has been found that chlorine can corrode metal switches and fittings in the overhead tank and toilet cistern. Excessive buildup of chlorine gas in the overhead tank (if located in a loft) may produce an unpleasant smell in the household and may be linked to asthma (9). Other disinfectants, such as ultraviolet radiation, are also available. UV lamps are expensive, and their germicidal efficiency reduces with time. They are most effective in waters of low turbidity. Therefore, fine filtration will be needed to achieve improved pathogen removal. The residual effect of UV as a disinfectant is not stable and microbial regrowth is possible, so UV treated waters should not be stored for long periods. Bromine is also used for pathogen removal from gray water. Some forms of bromine disinfectants such as hypobromous acid are considered harmful to plants, and the treated water would not be suitable for irrigation. The health risks from bromine use are yet to be quantified (7).

At present, gray water recycling systems on a single household scale are hardly financially viable. Although there are some savings from reduced consumption of mains water, the capital and operating expenditures incurred for these systems are relatively high, and the pay-back period is 20–25 years (10). The payback period reduces with increase in occupancy.

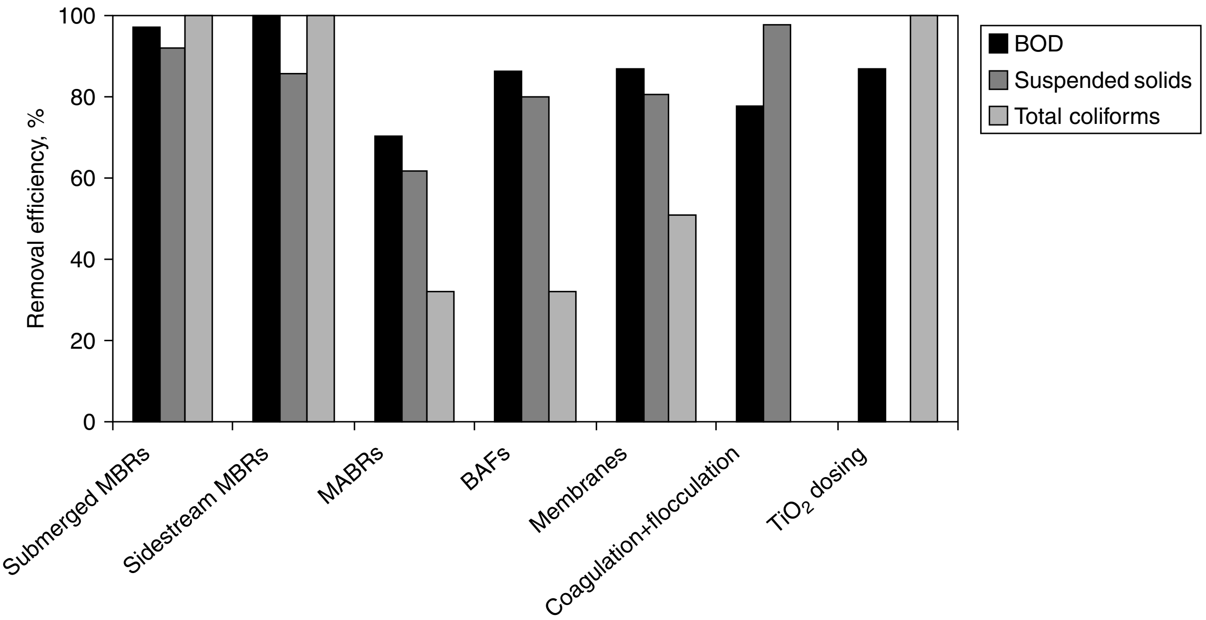

Figure 3. Technological performance in removing pollutants from gray water (3).

Gray water recycling on a medium to large scale (e.g., hotels, blocks of flats, commercial buildings,) may be more viable. A stepwise complex treatment sequence, instead of relying just on simple filtration and disinfection, produces water that has relatively low potential for health risk. A wide selection of gray water treatment technologies is currently available. These include biological aerated filters (BAFs), membranes, sidestream and submerged membrane bioreactors (MBRs), UV treatment, titanium dioxide (TiO_2) dosing, membrane aeration bioreactors (MABRs), and coagulation/flocculation with alum and ferric. Trials with these technologies have shown efficient and reliable removal of pollutants from gray water (3). The comparative efficiency of these technologies in removing BOD, suspended solids, and total coliforms is shown in Fig. 3. Large recycling units perform well, and their use in large buildings in the developed world and particularly in Japan is well established.

The main barrier to wider uptake of gray water recycling systems is lack of adequate consolidated legislation, high capital and maintenance costs, and potential health risks due to technology failure. Studies carried out to gauge public perception have shown that individuals have a positive attitude toward using treated gray water produced within their own households for toilet flushing, as long as safety is guaranteed and it is cost-effective (3).

BIBLIOGRAPHY

1. Butler, D. and Dixon, A. (1998). Assessing the viability of the household grey water and rainwater re-use. *21AD: Water, Architectural Digest for the 21st Century*. Oxford Brookes University, Oxford, pp. 20–23.
2. POST. (2000) *Water Efficiency in the Home*. Parliamentary Office of Science and Technology Note 135, London.
3. Jeffrey, P. and Jefferson, B. (2001). Water recycling: How feasible is it? *Filtration + Separation* **38**(4): 26–29.
4. Surendran, S. and Wheatley, A.D. (1998). Gray water reclamation for non-potable reuse. *J. Chartered Inst. Water Environ. Manage.* **12**(6): 406–413.
5. Dixon, A., Butler, D., Fewkes, A., and Robinson, M. (2000). Measurement and modelling of quality changes in stored untreated gray water. *Urban Water* **1**(4): 293–306.
6. Dixon, A. (2000). *Simulation of Domestic Water Reuse Systems: Gray Water and Rainwater in Combination*. PhD Thesis, Imperial College, London.
7. CIRIA. (2000). *Rainwater and Gray Water Use in Buildings: Best Practice Guidance*. Construction Industry Research Information Association Report, London.
8. Dixon, A., Butler, D., and Fewkes, A. (1999). Guidelines for gray water reuse—health issues. *J. Chartered Inst. Water Environ. Manage.* **13**(5): 322–326.
9. HAZOP. (2000). *HAZOP-Single house gray water recycling system*. Report No. WROCS HAZOP1, The School of Water Science, Cranfield University. http://www.cranfield.ac.uk/sims/water/gray waterhazopstudy.pdf.
10. WROCS. (2000). *Water Recycling Opportunities For City Sustainability*. Final Report prepared by Cranfield University, Imperial College and Nottingham Trent University in collaboration with industrial partners. http://www.cranfield.ac.uk/sims/water/finalreport-wrocs.pdf.

WATER AND HUMAN HEALTH

S.M.A Adelana
University of Ilorin
Ilorin, Kwara State, Nigeria

INTRODUCTION

Water in the literal sense is the source of life on the earth. Research has shown that the human body is 70% water. Generally, human beings begin to feel thirst after a loss of only 1% of body fluids and risk death if fluid loss exceeds 10%. It has been proved that human beings can survive for only a few days without fresh water. Although it is true that life depends on water, society does not usually act as though water has value equal to life itself. The reason is that the supply of water in many parts of the world far exceeds what is required to sustain life. Estimates revealed that about 9000 cubic meters (9.0×10^6 liters) of water is available for use per person per year. Based on projected population growth, this amount will drop to 5100 cubic meters per person by the year 2025 because another 2 billion people are expected to join the world's population by the year 2025. Despite this sharp drop (by nearly 50% in 35 years), the amount of water available would be sufficient to meet human needs if it were distributed equally among the world's population and less polluted by human activities.

Present estimates give a false picture of freshwater available for human use because the distribution of the world's available freshwater is uneven throughout the seasons and from year to year. According to Falkenmark (1), water is not always where we want it. Sometimes, it is not available in sufficient quantities where we want it or at another time too much water is in the wrong place. Yet, in many parts of the world, people are withdrawing water from surface and ground sources at a rate faster than they can be recharged.

In the last century, world population has tripled, but water withdrawals have increased by more than six times (2,3). For example, since 1940, annual global water withdrawals have increased by an average of 2.5% to 3% a year compared with annual population growth of 1.5% to 2% (4,5). In the past decade, however, water withdrawal has increased from 4% to 8% a year, especially in developing countries (6). If the present consumption patterns continue, by year 2025, about two billion people will be living in areas where it will be difficult or impossible to meet all their needs for fresh water. Half of them will face severe shortages (3,7,8).

Apart from the pressure of population growth on water resources, the supply of freshwater available to humanity is shrinking, in effect, because of increasing pollution. Population growth, urbanization, and industrialization with little regard for the environment are polluting and decreasing the quantity of freshwater available for human consumption (or use). Farming is said to be responsible for a great deal of water pollution in the United States (9). Similarly, in India (where there is heavy dependence on irrigation farming for food supplies), more than 4 million hectares of high-quality land have been abandoned as a result of salinization and waterlogging caused by excessive irrigation (10,11). More than 90% of the rivers in Europe have high nitrate concentrations, mostly from agrochemicals, and 5% of these rivers have concentrations at least 200 times higher than nitrate levels naturally occurring in unpolluted rivers (12–14). Moreover, in the Czech Republic, 70% of all surface waters are heavily polluted mostly with municipal and industrial wastes (15). Havas-Szilagyl (16) reported that 600 out of the 1600 well fields tapping groundwater in Hungary are already contaminated, mostly by agrochemicals.

In developing countries, an average of 90% of all domestic sewage and 75% of industrial wastes are discharged into surface waters without any kind of treatment (17,18). Generally, oil and salts are washed off city streets, and heavy metals are leached from municipal and industrial dump sites. There is also the possibility that pollutants, such as sulfur dioxide and oxygen or nitrogen, combine in the atmosphere to form acid rain which has devastating effects on both surface water and land ecosystems (19) and accompanying health implications (3).

On the whole, both water scarcity and water pollution pose serious health problems. Unclean water is by far the largest environmental killer around the world; it claims millions of lives every year. According to the World Health Organization (WHO), a large percentage of urban population in developing countries do not have access to proper sanitation facilities, and about half lacks a regular supply of potable water (12). In the year 2000, an estimated 1.1 billion people remained without access to improved drinking water (7), and the number of persons drinking water contaminated by human sewage was much higher (21). Obviously, scarce and unclean water supplies are critical public health problems in many parts of the world and are likely to be one of the major factors that will limit economic development in the near future (10,11,13,14).

WATER SCARCITY AND STRESS: IMPLICATIONS FOR HUMAN HEALTH

The world's population of nearly 6 billion is growing by about 80 million people each year (3). This rapid population growth coupled with increasing demands for water for irrigation agriculture, domestic (municipal), and industrial uses puts tremendous pressure on the world's freshwater resources. As population grows, water use per person rises and freshwater withdrawal becomes faster than it can be recharged, resulting in water mining (7,22). If this continues and water needs consistently outpace available supplies, a level will be reached when depletion of surface and groundwater resources results in chronic water shortages (23), as illustrated in Fig. 1.

Investigations have revealed that up to 31 countries, which represent nearly 8% of the world population, face chronic water shortages (3). It has also been estimated that by the year 2025, the number of countries facing water shortages is expected to be near 48, affecting more than 2.8 billion people—35% of the world's projected population (24–27). Figure 2 shows the population in water-scarce and water-stressed countries. It is obvious from this figure that, as population grows, many more countries will face water shortages. Accordingly, a more optimistic outlook predicts that 2.8 billion people in 48 countries will be struggling with water scarcity by the year 2025, whereas worst case scenarios for water shortages

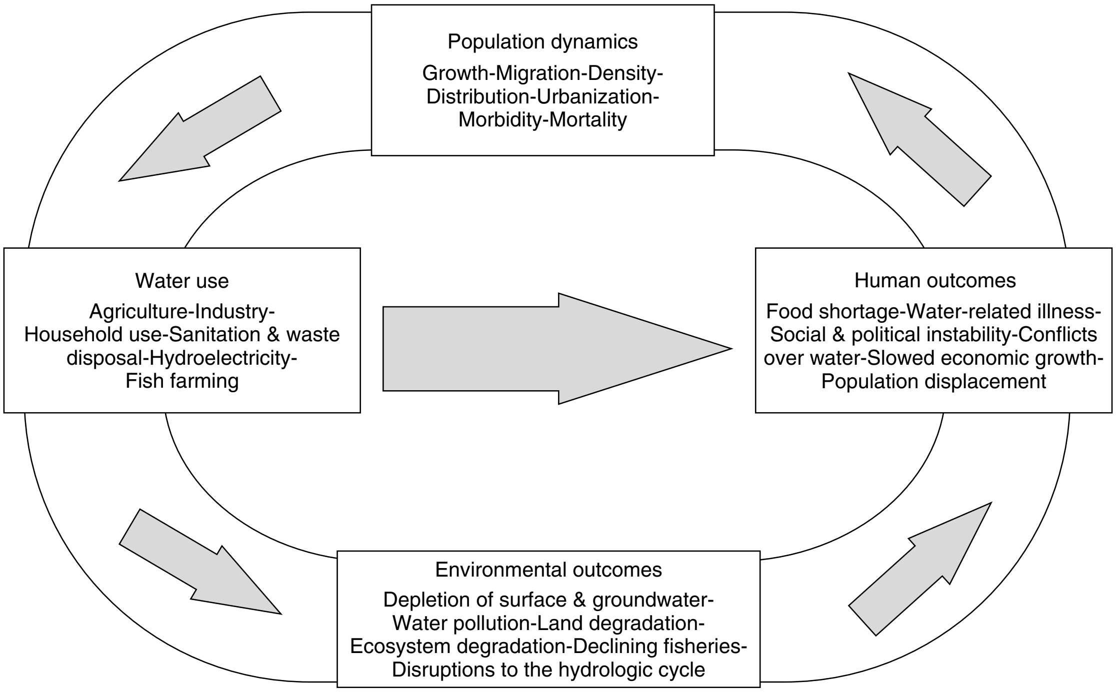

Figure 1. Links between world population and freshwater (3).

foresee 4 billion people in 54 countries facing water shortages in 2050. This critical trend is supported by the water scarcity and population projections of the United Nations (28–30).

From the growing consensus among hydrologists, a country is said to experience water stress when its annual water supply is between 1000 and 1700 cubic meters per person. Such a country can expect to experience temporary or limited water shortages. But when the annual supply of renewable freshwater drops below 1000 cubic meters per person, the country faces water scarcity (1,20,25).

In such cases, chronic and widespread shortages of water that hinder development result, and this could lead to severe health problems. Among countries likely to run short of water in the next 25 years are Ethiopia, Kenya, Nigeria, India, and Peru, whereas most parts of China and Pakistan are already approaching water stress. The list of water-stressed countries and those already suffering from water scarcity has been tabulated in Gardner-Outlaw and Engleman (25) and PR (3) based on the 1995 world population and water per capita and a projection for the year 2025. Gardner-Outlaw and Engelman (25) based their calculations on United Nations Population Division population estimates and growth rate and total fertility rate (TFR) data from the Population Reference Bureau (24), World Population Data Sheet. According to the estimates, it is obvious that in this century, water crises in more and more countries will present obstacles to better living standards and better health and may even bring risks of outright conflict over access to scarce freshwater supplies.

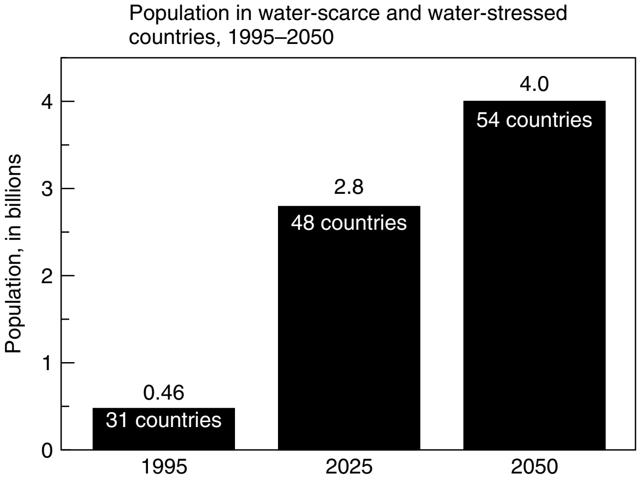

Figure 2. The rising trend in water scarcity and stress, 1995—2025 (3).

Available statistics show that more than half of the world's population suffers from water services that are inferior to those of the ancient Greeks and Romans (8). According to Gleick, this has long been recognized as a serious global water problem that even generated attention at the World Water Conference organized by the United Nations at Mar del Plata in 1977, where strong commitments and resolution were made to finding lasting solutions. Since this initial attempt, considerable efforts have been geared toward providing access to safe drinking water and adequate sanitation services. The United Nations in its Millennium Developmental Goal (during the World Summit on Sustainable Development held in Johannesburg in September 2002) planned to reduce by half, by the year 2015, the proportion of people without sustainable access to safe drinking water and basic sanitation services. Yet, a high percentage of the world's population is still without access to adequate water supply and sanitation (see Fig. 3a, b).

The 1990 world population without access to clean drinking water was estimated at 1300 million, and close to 2600 million people have no access to basic sanitation

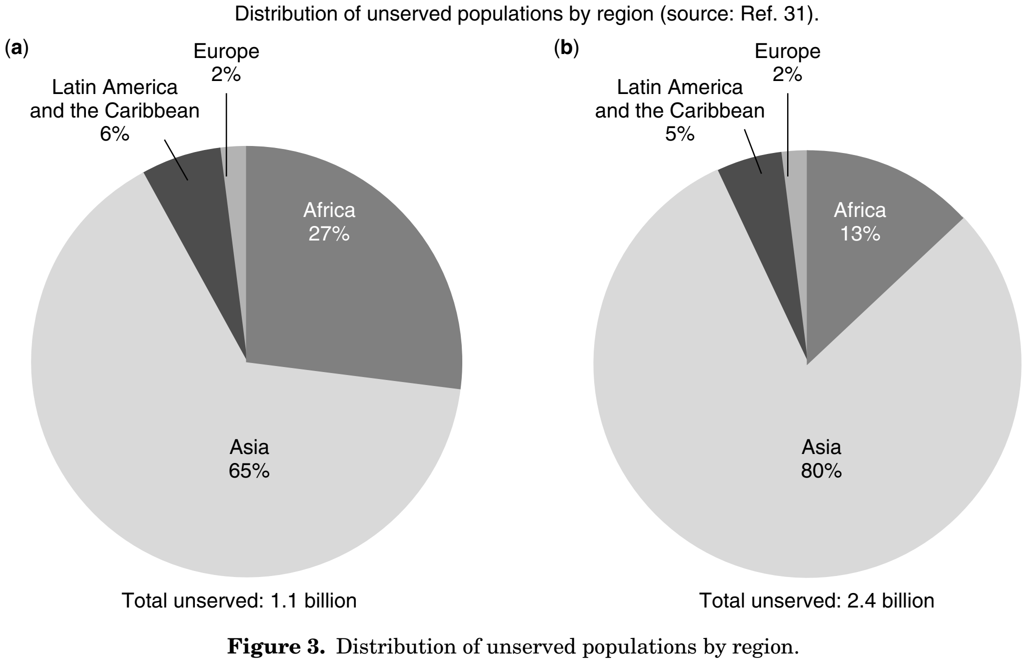

Figure 3. Distribution of unserved populations by region.

services (32). In Africa alone, 54% of the 1994 total estimated population of 707 million have no access to clean drinking water (33). Table 1 shows access to safe drinking water in developing countries by region; the percentage of population that has access to sanitation services is tabulated in Table 2. Illustrative world maps showing estimates of population without access to clean drinking water and adequate sanitation services are included in Gleick (8).

It is important to point out that the failure to provide basic clean water and sanitation services takes a serious toll on human health and results in economic loss in many countries of the world. Reports have shown that water, shortages, polluted water, and unsanitary living conditions claim millions of lives annually (7,20,34). The World Health Organisation reports that 80% of diseases are overtly or covertly waterborne (35) and/or consequent to freshwater shortages. Moreover, in much of the world, polluted water, improper waste disposal, and poor water management cause serious public health problems. For example, diarrheal diseases leave millions of children underweight, mentally as well as physically handicapped, and vulnerable to other diseases.

BASIC WATER, REQUIREMENTS FOR HUMAN HEALTH

Generally, in developing and using water resources, priority has to be given to the satisfaction of basic human needs (35). Therefore, providing water sufficient to meet basic human needs should be an obligation of governments and nongovernmental organizations. The postconference reports on the *International Symposium and Technology Expo* on small drinking water and wastewater systems held at the Hyatt Regency in Phoenix, Arizona (from January 12–15, 2000), demonstrated that the provision of safe drinking water and effective wastewater system managements are key elements that ensure safe and healthful living linked to social and economic development.

The basic water requirements for humans depend on the purpose for which water is used in the different sectors of our society. Among these are drinking (and other domestic use), removing and diluting waste (including disposing of human waste), growing food, producing manufactured goods, producing and using energy, and so on. The water requirement for each of these activities varies with domestic conditions, lifestyle, culture, tradition, diet, technology, and wealth (36).

It has been argued that water requirements for humans should also include any water necessary for disposing of human wastes (33). For example, in regions where absolute water quantity is a major problem, waste disposal options that require no water are available, in most developing nations, preference is given to alternatives that use at least some water (8). However, there are societies that use enormous amounts of fresh water to dispose of wastes. Based on the research carried out so far, a recommendation for a basic water requirement has been made. In 1996, Gleick proposed the overall minimum water required per person per day as 50 liters. This covers

Table 1. Population that has Access to Safe Drinking Water in Developing Countries, by Region, 1980 and 1994[a]

Region and Country	1980 Population, Millions	Percent With Access	Number Unserved, Millions	1994 Population, Millions	Percent with Access	Number Unserved, Millions
			Africa			
Urban	120	83	20	239	64	86
Rural	333	33	223	468	37	295
Total	453	46	243	707	46	381
			Latin America & the Caribbean			
Urban	237	82	43	348	88	42
Rural	125	47	66	125	56	55
Total	362	70	109	473	80	97
			Asia & the Pacific			
Urban	549	73	148	955	84	150
Rural	1823	28	1,313	2167	78	477
Total	2373	38	1,461	3122	80	627
			Western Asia			
Urban	28	95	1	52	98	1
Rural	22	51	11	29	69	9
Total	49	75	12	81	88	10
			Total			
Urban	**933**	**77**	**215**	**1594**	**82**	**279**
Rural	**2303**	**30**	**1612**	**2789**	**70**	**836**
Grand total	**3236**	**44**	**1827**	**4383**	**74**	**1115**

[a] Reference 8.

Table 2. Population that has Access to Basic Sanitation Services in Developing Countries, by Region, 1980 and 1994[a]

Region and Country	1980 Population, Millions	Percent with Access	Number Unserved, Millions	1994 Population, Millions	Percent with Access	Number Unserved, Millions
			Africa			
Urban	120	65	42	239	55	108
Rural	333	18	273	468	24	356
Total	453	30	315	707	34	464
			Latin American & the Caribbean			
Urban	237	78	52	148	73	94
Rural	125	22	97	125	34	83
Total	362	59	150	473	63	176
			Asia & the Pacific			
Urban	549	65	192	955	61	371
Rural	1823	42	1058	2167	15	1835
Total	2373	47	1250	3,122	29	2206
			Western Asia			
Urban	28	79	6	52	69	16
Rural	22	34	14	29	66	10
Total	49	59	20	81	68	26
			Total			
Urban	**933**	**69**	**289**	**1594**	**63**	**589**
Rural	**2303**	**37**	**1451**	**2789**	**18**	**2284**
Grand total	**3236**	**46**	**1740**	**4383**	**34**	**2873**

[a] Reference 8.

the minimum standards for drinking, sanitation, domestic (bathing and washing), and cooking. Out of this overall water requirement, 25 liters/person/day is required for basic hygiene (washing, showering, and bathing) and for cooking (33). A minimum of 20 liters/person/day offers the maximum benefits of combining waste disposal and related hygiene, thereby meeting cultural and societal preferences for water-based disposal (8). In other words, the minimum amount of water needed for drinking, cooking, bathing, and sanitation is 50 liters. The average person needs a minimum of 5 liters of water per day to survive in a moderate climate at an average activity level. However, average people in the United States uses between 250 to 300 liters of water per day for drinking, cooking, bathing, and watering their yards, whereas the average person in the Netherlands uses only 104 liters per day for the same tasks (33,37). This amount is slightly above the minimum target of 20–40 liters/person/day set by the United States Agency for International Development, the World Bank, and the World Health Organisation. Many people in the poorest nations survive on far less than the recommended amount. For example, the average person in Somalia uses only 8.9 liters of water per person per day (7,37).

Although different sources use different figures for total water consumption and for water use by sector of the economy (1,25,33,34,37–39), yet from drinking water and sanitation needs, it became obvious that a basic requirement of 25 liters/person/day of clean water must be provided for drinking and sanitation by water agencies, governments, or community organisations. An estimate made in 1990 revealed that about 55 countries whose population was nearly 1 billion people did not meet this standard (33). Yet, it is a desirable goal from a health perspective and from a broader objective of meeting a minimum quality of life. Further information on basic water requirements can be obtained in *The World's Water 1998–1999, The World's Water 2000–2001*, and The World's Water 2002–2003, which are available from Island Press, Washington (http://www.islandpress.com/).

WATER-RELATED, DISEASES

Water-related diseases that affect human health are relatively widespread and abundant, especially in rural communities of developing nations, although there is evidence that they have been reduced to a greater extent as a serious health problem in industrialized countries. The incidence of these diseases depends on local climate, geography, culture, sanitary habits and facilities, and on the quantity and quality of the local water supply as well as the methods of waste disposal (3). Changes in water supply do affect different groups of diseases in different ways. Some may depend on changes in water quality, others on water availability, and yet others on the indirect effects of standing water.

A World Health Organisation (40) estimate of the number of people suffering from water-related diseases is staggering (see Table 3). Generally, in many developing countries, waterborne diseases such as cholera, dysentery,

Table 3. Estimates of Global Morbidity and Mortality from Water-Related Diseases (Early 1990s) Culled from Reference 8[a]

Disease	Morbidity, Episodes/Year or People Infected	Mortality, Deaths/Year
Diarrheal diseases	1,000,000,000	3,300,000
Intestinal helminths	1,500,000,000 (people infected)	100,000
Schistosomiasis	200,000,000 (people infected)	200,000
Dracunculiasis	150,000 (in 1996)	—
Trachoma	150,000,000 (active cases)	—
Malaria	400,000,000	1,500,000
Dengue fever	1,750,000	20,000
Poliomyelitis	114,000	—
Trypanosomiasis	275,000	130,000
Bancroftian filariasis	72,800,000 (people infected)	—
Onchocerciasis	17,700,000 (people infected; 270,000 blind)	40,000 (mortality caused by blindness)

[a] Original *Source*: Reference 33.

typhoid, malaria, and schistosomiasis are increasing and harm or kill millions of people every year. The Pacific Institute's recent research indicates that lack of clean drinking water leads to nearly 250 million cases of water-related disease each year and roughly five to ten million result in deaths (7). Earlier estimates have shown much higher numbers of people in the world suffering from diseases that are linked with water (5,41) and resultant death (8,15). The true extent of these water-related diseases is unknown, and even the WHO data (40) suggest there may be many more cases of the diseases and resultant death.

However, about 60% of all infant mortality is linked to infectious and parasitic diseases; most are water-related (42), and a large percentage of these diseases is attributable to inadequate water supply and sanitation.

Research has shown that, at any one time, there are probably millions of people who have trachoma, elephantiasis, bilharzias (snail fever), malaria, diarrhea, dracunculiasis (guinea worm disease), and onchocerciasis (river blindness). For example, according to Edungbola (43), at least 15 million Africans suffered from guinea-worm infection; of these, nearly 75,000 people are permanently disabled every year, and about 3 million individuals were irreversibly crippled in Africa. His estimates have further shown that subsistence farmers in Africa lost at least 80 million man-days each year to guinea worm disease.

Water-related diseases are generally classed into four categories: waterborne, water-washed, water-based and water-related insect vectors (8,44,45). Waterborne diseases include those caused by both fecal—oral organisms and those caused by toxic substances; water-washed (also referred to as water-scarce) consists of diseases that develop where clean fresh water is scarce (44). Aquatic organisms that spend part of their life cycles in water and other part as parasites of animals cause water-based diseases. Insects that transmit infections, such as mosquitoes and tsetse flies, cause water-related vector diseases. A full description of each class of water-related disease together with their causative agents and routes of transmission as well as the geographical extent and number of reported cases has been compiled in Population Reports (3).

According to Population Reports, diarrheal disease (which belongs to the class of waterborne disease) is prevalent in many countries where sewage treatment is inadequate or where human wastes are disposed of in open latrines, ditches, canals, and watercourses or are spread indiscriminately on farmland. In the mid-1990s, a large number of people drank water contaminated with human sewage (28), and the World Health Organisation reported that drinking contaminated water contributes directly to diarrhea-related deaths (46). An estimated 4 billion cases of diarrheal diseases are reported annually that cause 3–4 million deaths, mainly among children (34,47–49). In Nigeria alone, more than 300,000 children less than 5 years of age die annually from diarrheal diseases (50). For example, in 1996, a large outbreak of severe diarrhea (which was later confirmed as cholera from tests conducted at the Institut Pasteur, Paris, France) struck the commercial city of Kano in northern Nigeria. According to Hutin et al. (51), a total of 5600 cases and 340 deaths (attack rate = 86.3 per 100, 000 inhabitants) were reported to the Kano State Ministry of Health within 5 months of the incident. This incidence was highest among children less than 5-years-old and was linked to drinking street-vended water and failure to wash hands with soap before meals (51).

Earlier, the consumption of street-vended water was reportedly associated with a cholera outbreak in Latin America (52–54). A similar cause of a cholera outbreak was also reported in India (55,56) and in Peru (57). Recent research has also shown that childhood mortality from diarrhea in Latin America remains high (58).

Gleick (8), using available data on the prevalence of different water-related diseases, presented and discussed two of these diseases extensively—dracunculiasis (guinea worm) and cholera—as case studies. He traced the history and reported the total global cases of these diseases by region and the recurrent deaths as a result of the epidemic from 1990 to 1997 with an update on the complete eradication programs. Reported guinea worm cases, globally, have fallen from an estimated 3.5 million in 1986 to 150, 000 in 1996 (59,60). This is approximately a 97% reduction, and there are hopes that it may have been completely eradicated in accordance with the “New Millennium Plan.”

On the other hand, little has been achieved in the effort to control the transmission of other parasitic infections such as schistosomiasis, intestinal helminthiasis, and malaria which are related to water supply and sanitation, especially in the developing world and particularly in Africa. For further information and statistics on other water-related diseases, readers are referred to (5,12,14,24,34,47–49,59–68).

The issue of water quality or maximum permissible limits of certain elements that can constitute health risks in drinking water should be included here as another source of waterborne disease. For example, increased nitrate concentrations in drinking water add to the variety of water-related health risks. Health problems from nitrate in water sources are generating serious concern in almost all countries of the world, particularly in urban and rural communities where agricultural practises are intensive (69–73). There is increasing evidence that nitrate levels in many aquifers are rising and that the problem of increased exposure of the world population to high nitrate inputs will become more pressing, as speculated earlier by the WHO (74).

Agricultural activities such as fertilizer and pesticides applications are frequent sources of contamination in surface and groundwaters. An estimate from Population Reports has shown that in more than 150 countries, nitrate from the application of fertilizers has seeped into water wells and polluted drinking water (75). Increased concentrations of nitrate often cause blood disorders (76). High levels of nitrate and phosphates in drinking water also encourage the growth of blue-green algae, resulting in deoxygenation (eutrophication) and subsequent reduction in metabolic activities of the organisms that purify fecal-polluted water in the human system. Details of nitrate health hazards are discussed in Adelna (in this volume).

Other sources of water pollution include animal wastes, excess nutrients, salinity, pathogens, and sediments that often render water unusable for drinking, unless it is purified (77–81). Even when any of these substances or chemicals occurs in low concentrations, they can accumulate in humans over time to cause serious health problems such as cancer if the water is used for drinking. For maximum permissible and acceptable levels of ions/elements in water, refer to the standards of the World Health Organisation (62) and of most national authorities, which are consistent with standards for the composition of drinking water (82).

The average contribution of drinking water to the daily intake of mineral nutrients is important in health considerations. Of note here are those for fluoride (F^-) and arsenic (As). Generally, excessive concentrations of these elements often limit the use of groundwater for drinking. Too high an intake of fluoride is often the general cause of painful skeleton deformations called fluorosis, which is a common disease in East Africa, especially in Kenya and Ethiopia. The occurrence of fluoride in groundwater for human consumption has also been reported in Argentina (83). High concentrations of arsenic in groundwater used for drinking are reported in many countries such as India, Bangladesh, China, Thailand, Vietnam, Taiwan, Hungary, Mexico, and Finland (84–88). Nearly 50 million people are at risk of cancer and other arsenic-related diseases due to consumption of high arsenic groundwater in India and Bangladesh (64,89). About 44% of the population of West Bengal (India) is suffering from arsenic-related diseases such as conjunctivitis, melanosis, hyperkeratosis, and hyperpigmentation (90). In certain areas, gangrene in the limb, malignant neoplasm, and even skin cancer have also been observed. High arsenic concentrations lead to black-foot disease. This is sometimes visible in a blackening of the fingers and toe tops and induces general lethargy in the patient. Arsenic toxicity affects almost all organs of the human body. Ingestion of large doses of arsenic usually results in symptoms within 30 to 60 minutes but may be delayed when taken with food (90). High arsenic concentrations have also been reported in Southeast Asia (91), the United States (92), Argentina, and Chile (93–97); all have consequent health implications.

THE IMPACT OF IMPROVED WATER SUPPLY AND SANITATION

The direct consequence of water scarcity and failure to meet basic water requirements is the prevalence of most water-related diseases. In the past, this has caused serious economic and social loss to both governments and communities. Estimates in the late 1970s have shown that water-related diseases cost more than $125 billion per year, excluding social costs, the loss of education and other opportunities, lost economic productivity of sick workers, and other hidden costs (8,98). For example, in sub-Saharan Africa, malaria costs an estimated $1.7 billion US annually in treatment and lost productivity (48). A study in Pakistan (within its capital city, Karachi) has shown that people living in areas without proper sanitation or hygiene education spent six times more on medical care than residents in areas with access to sanitation and basic hygiene (63). In Peru, an economic loss of more than $1 billion dollars in seafood exports and tourist revenues has been reported due to a cholera epidemic (99).

However, the huge investments by governments in Asia, Africa, and Latin America in basic water and sanitation services (5,100) have reduced the prevalence of these diseases in the last decade. The World Bank estimates spending for water and sanitation in developing countries at nearly $26 billion per year (101). Not until clean drinking water and improved sanitation services are universally available will millions of people stop dying from preventable water-related illnesses (8). Rogers (101) estimated that completely meeting basic water supply needs up to the year 2020 would require total capital costs of about $24 billion per year. If the additional costs of meeting a higher level of services, such as advanced wastewater treatments, were included, the cost would be up to $50 billion a year.

Several studies have reported the high reduction in water-related morbidity and mortality as a result of improvements in water services and sanitation consequent to these huge financial investments. According to Population Reports, a review in 1991 of more than 100 studies of the effects of clean water and sanitation on human health revealed a medium reduction in deaths from water-related diseases (up to 69%) among residents, who have access to clean drinking water and improved sanitation services, because effective disposal of human wastes controls the spread of infectious agents and interrupts the transmission of water-related diseases. Table 4 shows the impact of improved water infrastructure on reducing water-related diseases.

Another approach that has shown improved water quality and reduced the incidence of waterborne disease (for example, diarrhea) is the Center for Disease Control (CDC) safe water system. This system combines locally produced sodium hypochlorite solution (chlorine bleach), a CDC water storage vessel, and a public health campaign to change the behavior of rural dwellers to basic hygiene (102,103). This system has improved water quality and reduced the incidence of diarrhea by 68% in Uzbekistan (104), by 44% in Bolivia (105), and by 48% in Zambia (106). It further serves as an alternative method of disinfecting drinking water in rural Guatemala and prevents excessive morbidity and mortality from waterborne diseases (103).

Moreover, according to Population Reports (3), some water development schemes have started disease control programs along with construction of water and sanitation facilities. As a result of such a program in the Philippines, for example, the prevalence of water-related diseases fell from 24% in 1979 to 9% in 1985 (67). There are indications of good progress made so far, and, at this point, some water-related diseases are on the verge of complete eradication. A good example is guinea worm (dracunculiasis) eradication. Pakistan, reportedly the first country to have completely eradicated guinea worm during the new global eradication program, recorded zero cases every month since October 1993 (8). In India, guinea worm was completely eliminated from the Tamil Nadu area in 1984, the Gujarat area in 1989, and Maharashtra in 1991; only nine cases were reported for the entire country in 1996 (60). In the Kwara State of Nigeria (where guinea worm once had devastating effects on the rural populace), any reported case of guinea worm now attracts a monetary prize (107).

Globally, only five of the countries that had guinea worm recorded slightly above 100 cases in 1996. The number of cases has generally dropped by nearly 97% during the past decade (8). Although the eradication program has shown impressive progress, guinea worm is still prevalent in nearly 17 developing nations, mostly in Africa as of the end of 1996 (8). In the final analysis, guinea worm has been eradicated most effectively by providing protected clean drinking water in all countries where the disease was prevalent. Successful eradication programs for guinea worm and other water-related diseases are documented in the series of articles in Population Reports, Series M (3), as well as in (5,59,60,100,108–110). Therefore, an improved water supply and sanitation system will consequently generate tremendous improvements in the health, social welfare, and economic development of any nation, especially a developing one. For details of this and several other studies related to reduction in waterborne morbidity and mortality as a result of improvements in sanitation and water supplies, refer to Esrey et al. (111), Alam (112), Aung and Thein (113), Baltazar (114), Cairncross and Cliff (115), Young and Briscoe (110), Esrey and Habicht (116), Henry (117), Rahman (118), Haines and Avery (119), Khan (120), Torun (121), Ankar and Knowles (122), Koopman (123), Misra (124), and White et al. (36).

SUMMARY

Water is essential for life and health and has cultural and religious significance. Water plays a vital role in transmitting infectious diseases, and 80% of diseases reported are directly or indirectly water-related. Scarce and unclean water supplies are critical public health problems in many parts of the world and are likely to be one of the major factors that will limit economic development in the near future. It has been reported that

Table 4. Impact of Improved Water Infrastructure on Reducing Water-Related Diseases[a]

Place	Type of Facilities or Improvement	Type of Study	Diseases	Difference in Incidence After Improvement
Teknaf, Bangladesh	Hand pumps and health education	Case-control	Diarrheal diseases	17% difference between groups
Northeast Brazil	Latrines, communal taps, laundry facilities, showers, and hand pumps	Case-control	Schistosomiasis	77% difference between groups
Khuzestan, Iran	Courtyard latrine and public standpipes	Case-control	Ascariasis	16% difference between groups
Uttar Pradesh, India	Piped water	Before and after	Dysentery	76% reduction
Peninsular Malaysia	Toilets and running water	Case-control	Diarrheal diseases	82% difference in infant mortality between groups
Kwara State, Nigeria	Boreholes, hand pumps, and health education	Before and after	Dracunculosis	81% reduction
Cebu, Philippines	Private, sanitary latrines	Before and after	Diarrheal diseases	42% reduction
St. Lucia	Household water and latrines	Case-control	Ascariasis	31% difference between groups
Lusaka, Zambia	Extension of piped water supply	Before and after	Typhoid	37% reduction

[a] *Source*: Selected studies compiled in Reference 3.

water shortages, polluted water, and unsanitary living conditions claim millions of lives annually via various water-related diseases.

Research has shown that, at any one time, there are probably millions of people who have trachoma, elephantiasis, bilharzias (snail fever), malaria, diarrhea, dracunculiasis (guinea worm disease), and onchocerciasis (river blindness). The incidence of these diseases, it has been shown, depends on local climate, geography, culture, sanitary habits and facilities, and on the quantity and quality of the local water supply as well as methods of waste disposal. Effective disposal of human wastes controls the spread of infectious agents and interrupts the transmission of water-related diseases. The role of good quality drinking water and access to adequate sanitation facilities in reducing water-related diseases has been reviewed in this article. The universal provision of treated pipe-borne water is not currently feasible due to economic and political constraints, and this consequently leaves millions of people without access to safe drinking water. Generally, the failure to provide clean drinking water and adequate sanitation services has serious implications for human health and is consequent to severe economic loss in many countries.

Furthermore, the water requirements for each of the basic human activities vary with domestic conditions, lifestyle, culture, tradition, diet, technology, and wealth. In any case, the minimum amount of water needed for drinking, cooking, bathing, and sanitation is set at 50 liters. The average person needs a basic minimum of 5 liters of water per day to survive in a moderate climate at an average activity level. This is the absolute minimum amount of water required to maintain adequate human health, independent of lifestyle and culture.

About 60% of all infant mortality is linked to infectious and parasitic diseases; most are water-related, and a large percentage of these diseases is attributable to inadequate water supply and sanitation. Diarrheal diseases, for example, leave millions of children underweight, mentally as well as physically handicapped, and vulnerable to other diseases. However, the huge financial investments by governments in Asia, Africa, and Latin America (as well as by nongovernmental organizations) in basic water and sanitation services have reduced the prevalence of these diseases in the last decade.

BIBLIOGRAPHY

1. Falkenmark, M. (1994). Landscape as life support provider: Water-related limitations. In: F. Graham-Smith (Ed.), *Population—The Complex Reality*. North American Press, Golden, CO, pp. 103–116.
2. Juma, C. (1998). *The CBD and the Biological Diversity of Inland Waters*. Presented at the *Int. Conf. Water Sustainable Dev.*, Paris, March 19–21, pp. 1–4.
3. Population Reports. (1998). *Solutions for a Water-Short World*. Series M, No. 14, September 1998.
4. Dowdeswell, E. (1996). Editorial. *Our Planet*, Oct. pp. 1–2.
5. United Nations Children's Fund. (1997). Facts and figures. Online: http://www.unicef.org/factx/facright.htm.
6. Marcoux, A. (1994). *Population and Water Resources*. United Nations Food and Agriculture Organization, Rome, pp. 4–33.
7. The Pacific Institute for Studies in Development, Environment and Security. (2003). Threats to the World's Freshwater Resources. Online: http://pacinst.org/reports/freshwater_threats.htm.
8. Gleick, P.H. (1998). *World Water 1998–1999: The Biennial Guide to Freshwater*. Island Press, Washington, DC.
9. New York Times. (1998). Farms are polluters of nations waterways. *New York Times*, May 14, A19.
10. Serageldin, I. (1995). *Toward Sustainable Management of Water Resources*. World Bank, Washington, DC, pp. 1–33.
11. Serageldin, I. (1995). Water resources management: New policies for a sustainable future. *Water Front*, Jun. 1, Stockholm 2–3.
12. World Health Organization. (1997). *Health and Environment in Sustainable Development Five Years after the Earth Summit*. WHO, Geneva, pp. 19–133.
13. World Health Organization. (1992). *Our Planet, our Health—Report of the WHO Commission on Health and Environment*. WHO, Geneva, pp. 106–144.
14. World Health Organisation. (1992). *The International Drinking Water Supply and Sanitation Decade: End of Decade Review*. Geneva.
15. Nash, L. (1993). Water quality and health. In: *Water in Crisis*. P. Gleick (Ed.). Oxford University Press, New York, pp. 25–39.
16. Havas-Szilagyi, E. (1998). *National Groundwater Protection Programs in Hungary*. Presented *Int. Conf. Water Sustainable Dev.*, Paris, March 19–21, pp. 1–5.
17. Allaoui, K. (1998). *Long-Term Finance for Water Projects: The IDB's Approach*. Presented at the *Int. Conf. Water Sustainable Dev.*, Paris, March 19–21, pp. 1–7.
18. Carty, W. (1991). Towards an urban world. *Earthwatch* **43**: 2–4.
19. Hinrichsen, D. (1988). Acid rain and forest decline. In: *The Earth Report: Monitoring the Battle for Our Environment*. E. Goldsmith and N. Hildyard (Eds.). Mitchell Beazley, London, pp. 66–78.
20. Davidson, J., Myers, D., and Chakraborty, M. (1992). *No Time to Waste—Poverty and the Global Environment*. Oxfam, Oxford.
21. United Nations. (1996). *Human Development Report 1996. United Nations Development Program*. Oxford University Press, New York.
22. Abramovitz, J. (1996). *Imperilled Waters, Impoverished Future: The Decline of Freshwater Ecosystems*. Worldwatch Institute, Washington, DC, pp. 5–66.
23. Merla, A.A. (1998). *Commitment to the Global Environment: The role of GEF and International Waters*. Presented at the *Int. Conf. Water Sustainable Dev.*, Paris, March 19–21, pp. 1–3.
24. Population Reference Bureau. (1998). *World Population Data Sheet [wallchart]*. Washington, DC.
25. Gardner-Outlaw, T. and Engleman, R. (1997). *Sustaining Water, Easing Scarcity: A Second Update*. Population Action International, Washington, DC, pp. 2–19.
26. United Nations Population Fund. (1997). *Population and Sustainable Development—Five Years after Rio*. UNPF, New York, pp. 1–36.
27. Falkenmark, M. (1993). Water scarcity: Time for realism. *Populi* **20**(6): 11–12.

28. United Nations. (1997). *Comprehensive Assessment of the Freshwater Resources of the World*. Report of the Secretary-General, UN, Commission on Sustainable Development, New York.
29. United Nations. (1997). *Critical Trends—Global Change and Sustainable Development*. Department for Policy Coordination and Sustainable Development, UN, New York, pp. 43–56.
30. United Nations Educational Scientific and Cultural Organization. (2003). *Water for People, Water for Life*. (UNESCO Release, March 5, 2003).
31. World Health Organisation. (WHO/UNICEF). (2002). *Executive Summary of the World Water Day Report*. WHO/UNICEF Joint Monitoring Programme, updated in September 2002.
32. World Health Organization. (1996). *The World Health Reports: Fighting Disease, Fostering development*. WHO, Geneva, Switzerland.
33. Gleick, P.H. (1996). Basic water requirements for human activities: Meeting basic needs. *Water Int.* **21**: 83–92.
34. United States Agency for International Development. (1990). *Strategies for Linking Water and Sanitation Programs to Child Survival*. USAID, Washington, DC, pp. 1–62.
35. United Nations. (1992). *Protection of the Quality and Supply of Freshwater Resources: Application of Integrated Approaches to the Development, Management and Use of Water Resources of Agenda 21*. United Nations Publications, New York, chap. 18.
36. White, G.F., Bradley, D.J., and White, A.U. (1972). *Drawers of Water: Domestic Water Use in East Africa*. University of Chicago Press, Chicago.
37. Gleick, P. (2000). Amount of water per capita in year 2000 on global basis. Global Environment Program of the Pacific Institute. http://pacinst.org/reports.
38. Food and Agriculture Organization of the United Nations. (1995). *Dimensions of Need: An Atlas of Food and Agriculture*. FAO, Rome.
39. Falkenmark, M. and Widstrand, C. (1992). Population and water resources: A delicate balance. *Population Bull.* **47**(3): 1–36.
40. World Health Organization. (1995). *Community Water Supply and Sanitation: Needs, Challenges and Health Objectives*. *48th World Health Assembly*, A48/INF.DOC./2, April 28, 1995, Geneva, Switzerland.
41. Kristof, N. (1997). For third world, water is still a deadly drink. *New York Times*, Jan. 9, A1, A8.
42. Rowley, J. (1990). Linking population to conservation—special report on Pakistan. *Earthwatch* **40**: 3–5.
43. Edungbola, L.D. (1988). The impact of a UNICEF-assisted rural water project on the prevalence of guinea worm disease in Asa, Kwara State, Nigeria. *Am. J. Trop. Med. Hyg.* **39**: 79–85.
44. Kjellén, M. and Mcgranahan, G. (1997). *Urban Water—Towards Health and Sustainability*. Stockholm Environmental Institute, Stockholm, pp. 1–54.
45. Bradley, D. (1994). Health, environment, and tropical development. In: *Health and the Environment: The Linacre Lectures 1992–3*. B. Cartledge (Ed.). Oxford University Press, Oxford, pp. 126–149.
46. World Health Organisation. (1999). *The World Health Report 1999*. Geneva, Switzerland.
47. Warner, D. (1998). *Drinking Water Supply and Environmental Sanitation for Health*. Presented at the *Int. Conf. Water Sustainable Dev.*, Paris, March 19–21, pp. 1–10.
48. Olshansky, S.J., Carnes, B., Rogers, R., and Smith, L. (1997). Infectious diseases—New and ancient threats to world health. *Population Bull.* **52**(2): 2–43.
49. World Bank. (1993). *Water Resources Management: A World Bank policy Paper*. World Bank, Washington, DC.
50. Idika, N. (2001). Water Purification: Strategy for reducing waterborne diseases in Nigeria. *Waterfront—Nig. Water Sanit. Assoc.* **1**: 4–6.
51. Hutin, Y., Luby, S., and Paquet, C. (2003). A large cholera outbreak in Kano City, Nigeria: The importance of hand washing with soap and the danger of street-vended water. *J. Water Health* **1**(1): 45–52.
52. Tauxe, R.V., Mintz, E.D., and Quick, R.E. (1995). Epidemic cholera in a new world: Translating field epidemiology into new prevention strategies. *Emerging Infect. Dis.* **1**: 141–146.
53. Weber, J.T. et al. (1994). Epidemic cholera in equador: Multidrug-resistance and transmission by water and sea food. *Epidemiol. Infect.* **112**: 1–11.
54. Ries, A.A. et al. (1992). Cholera in Piura, Peru: A modern epidemic. *J. Infect. Dis.* **166**: 1429–1433.
55. Deb, B.C. et al. (1986). Studies on interventions to prevent El Tor cholera transmission in urban slums. *Bull. WHO* **64**: 127–131.
56. Deb, B.C. et al. (1982). Intra-familial transmission of *Vibro cholerae* biotype El Tor in Calcutta slums. *Indian J. Med. Res.* **76**: 814–819.
57. Swerdlow, D.L. et al. (1992). Waterborne transmission of epidemic cholera in Trujilo, Peru: Lessons for a continent at risk. *Lancet* **340**: 28–33.
58. Pan American Health Organisation. (2002). *Health in the Americas*. Pan American Health Organisation, Washington, DC.
59. World Health Organisation. (1997). *Dracunculiasis: Disease Sheet*. Division of Control of Tropical Diseases, Geneva, Switzerland.
60. World Health Organisation. (1997). Dracunculiasis: Global surveillance summaries. *Wkly. Epidemiol. Rec.* **57, 60, 61, 66, 68–70, 72**: Geneva, Switzerland
61. World Health Organization. (1998). Division of Control of Tropical Disease homepage. Online: http://www.who.ch/ctd/. Feb. 25.
62. World Health Organisation. (1998). *Guidelines for Drinking-Water Quality, Health Criteria and other information*, 2nd Edn. Geneva, Switzerland, pp. 281–283.
63. Khan, A.H. (1997). The sanitation gap: Development's deadly menace. In: *The Progress of Nations*. UNICEF, New York, pp. 5–13.
64. Chatterjee, A. et al. (1995). Arsenic in groundwater in six districts of West Bengal, India: The biggest arsenic calamity in the world. Part 1: Arsenic species in drinking water and urine of the affected people. *Analyst* **120**: 643–650.
65. Garrett, L. (1994). *The Coming Plague: Newly Emerging Diseases in a World Out of Balance*. Farrar, Straus and Giroux, New York.
66. Muller, R. and Morera, P. (1994). Helminthoses. In: *Health and Disease in Developing Countries*. K.S. Lankinen, S. Berström, P.H. Mäkelä, and M. Peltomaa (Eds.). Macmillan, London, pp. 195–209.
67. Hunter, J.M., Rey, L., Chu, K.Y., Adekolu-John, E.O., and Mott, K.E. (1993). *Parasitic Diseases in Water Resources*

Development: The Need for Intersectoral Negotiation. World Health Organization, Geneva.

68. Basch, P.F. (1990). *Textbook of International Health*. Oxford University Press, New York.
69. Adelana, S.M.A. and Olasehinde, P.I. (2003). High nitrates in rural water supply in Nigeria: Implications for human health. *Water Res.* **14**(1): 1–11.
70. Tredoux, G., Tama, A.S., and Engelbrecht, J.F.P. (2000). *The Increasing Nitrate Hazard in Groundwater in the Rural Areas*. Paper presented at *Water Inst. S. Afr. Conf.*, May, Sun City.
71. Colvin, C. (1999). *Increased Risk of Methaemoglobinaemia as a Result of Bottle-Feeding by HIV Positive Mothers in South Africa*. Paper presented at *IAH Cong.*, 1999, Bratislava, Slovakia.
72. Marais, S. (1999). *Dependency of Communities on Groundwater for Water Supply and Associated Fluoride and Nitrate Problems*. Paper presented at Workshop on Fluorides and Nitrates in Rural Water Supplies, Mafikeng, March 9–10, 1999.
73. Fan, A.M. and Steinberg, V.E. (1996). Health implications of nitrate and nitrite in drinking water: An update on methaemoglobinaemia occurrence and reproductive and developmental toxicology. *Regul. Toxicol. Pharmacol.* **23**, Academic Press Inc.
74. World Health Organisation. (1985). *Health Hazards from Nitrates in Drinking Water*. Copenhagen.
75. Maywald, A., Zeschmar-Lahl, B., and Lahl, U. (1988). Water fit to drink? In: *The Earth Report: Monitoring the Battle for Our Environment*. E. Goldsmith and N. Hildyard (Eds.). Mitchell Beazley, London, pp. 79–88.
76. Bowman, J. (1994). 'Water is best': Would pindar still think so? In: *Health and the Environment: The Linacre Lectures 1992–3*. B. Cartledge (Ed.). Oxford University Press, Oxford, pp. 85–125.
77. Maurice-Bourgoin, L., QuirogA, I., Guyot, J.L., and Maim, O. (1999). Mercury pollution in the Upper Beni River, Amazonian Basin: Bolivia. *Ambio* **28**(4): 302–306.
78. Silva-Forsberg, M.C., Forsberg, B.R., and Zeidemann, V.K. (1999). Mercury contamination in humans linked to river chemistry in the Amazon Basin. *Ambio* **28**(6): 519–521.
79. Klohn, W. and Wolter, W. (1998). *Perspectives on Food and Water*. Presented at the *Int. Conf. Water Sustainable Dev.*, Paris, March 19–21, pp. 1–6.
80. Satterthwaite, D. (1993). The impact on health of urban environments. *Environ. Urbanization* **5**(2): 87–111.
81. Witt, V.M. and Reiff, F. (1991). Environmental health conditions and cholera vulnerability in Latin America and the Caribbean. *J. Public Health Policy* **12**(4): 450–463.
82. Safe Drinking Water Communication (SDWC) (1980). *Drinking Water and Health*, Vol. 3, National Academy Press, Washington, DC, p. 415.
83. Schultz, C.J., Castro, E., Marino, E., and Dalmaso, G. (2001). Occurrence of fluoride in groundwater for human consumption in the province of La Pampa, Republica Argentina. In: *New Approaches Characterizing Groundwater Flow*. K.P. Seiler and S. Wohnlich (Eds.). A.A. Balkema, Lisse, pp. 637–639.
84. Arsenic Crisis Information Center. (2001). West Bengal and Bangladesh Arsenic Crisis Information Center. http://bicn.com/acic/.
85. Kurttio, P. et al. (1998). Urinary excretion of arsenic species after exposure to arsenic present in drinking water. *Arch. Environ. Contamination Toxicol.* **34**: 297–305.
86. Chen, S.L., Dzeng, S.R., Yang, M.H. (1994). Arsenic species in groundwaters of the black-foot disease area, Taiwan. *Environ. Sci. Technol.* **28**: 877–881.
87. Varsa'nyi, I., Fodre, Z., and Bartha, A. (1991). Arsenic in linking water and mortality in the southern Great Plain, Hungary. *Environ. Geochem. Health* **13**: 14–22.
88. Del Razo, L.M., Arellano, M.A., and Cebrián, M.E. (1990). The oxidation states of arsenic in well water from a chronic arsenicism area of northern Mexico. *Environ. Pollut.* **64**(2): 143–153.
89. Bhattacharya, P., Frisbie, S.H., Smith, E., Naidu, R., Jacks, G., and Sarkar, B. (2001). Arsenic in the environment: A global perspective. In: *Handbook of Heavy Metals in the Environment*. B. Sarkar (Ed.). Marcel Dekker, New York.
90. Chandrasekharam, D. (2002). Geogenic arsenic pollution in groundwater. In: *Groundwater and Human Development*. E. Bocanegra, D. Martínez, and H. Massone (Eds.). Mar del Plata, pp. 1990–1994.
91. Bhattacharya, P. (2002). Arsenic contaminated groundwater from the sedimentary aquifers of Southeast Asia. In: *Groundwater and Human Development*. E. Bocanegra, D. Martínez, and H. Massone (Eds.). Mar del Plata, pp. 1973–1979.
92. Welch, A.H., Lico, M.S., and Hughes, J.L. (1988). Arsenic in groundwater of western United States. *Groundwater* **26**(3): 333–347.
93. Herrera, H., Farias, B., Martin, R., Cortés, J., Ferreyra, G., Thir, J., and Storniolo, A. (2000). El arsenico en la zona rural del dpto. Robles, Santiago del Estero, Argentina. *Proc. Cong.* Agua, Junio 2000, Santiago del Estero, Argentina.
94. Sancha, A.M. and Frenz, P. (2000). *Estimate of Current Exposure of the Urban Population of Northern Chile to Arsenic*. Interdisciplinary Perspectives on Drinking Water Risk Assessment and Management. *Proc. Santiago* (Chile) *Symp.*, IAHS Publ. 260, pp. 3–8.
95. Fiorentino, C., Sequeira, M., Paoloni, D., and Echeverría, N. (1998). *Detección y distribución de arsénico, fluor y boro en aguas subterráneas regionales*. Mapas de Riesgo. *Cong. Nacional Agua, Actas*: 71–79. Santa Fe, Argentina.
96. Smedley, P.L., Nicolli, H., Barros, J., and Tullio, O. (1998). *Origin and Mobility of Arsenic in Groundwater from the Pampean Plain, Argentina*. *9th Int. Symp. Water-Rock Interactions*, Taupo, New Zealand.
97. Nicolli, H. et al. (1989). Groundwater contamination with arsenic and other elements in an área of the Pampa, province of Córdoba, Argentina. *Environ. Geol. Water Sci.* **14**: 1–15.
98. Pearce, D.W. and Warford, J.J. (1993). *World Without End: Economics, Environment, and Sustainable Development*. Oxford University Press, New York.
99. Epstein, P.R. and Nutter, F. (1997). Climate change: Assessing the economic damages. *Newsl. Center Health Global Environ.*, Harvard Medical School. Online: http://www.med.harvard.edu/chge/.
100. United Nations Children's Fund. (1997). Water and sanitation progress and disparity. Online: http://www.unicef.org/pon97/p12b.htm.
101. Rogers, P. (1993). Integrated urban water resources management. *Nat. Res. Forum*, pp. 33–42.
102. Mintz, E., Reiff, F., and Tauxe, R. (1995). Safe water treatment and storage in the home: A practical new strategy to prevent waterborne disease. *J. Am. Med. Assoc.* **273**(12): 948–953.

103. Rangel, J.M. et al. (2003). A novel technology to improve drinking water quality: A microbiological evaluation of in-home flocculation and chlorination in rural Guatemala. *J. Water Health* **1**(1): 15–22.
104. Semenza, J. et al. (1998). Water distribution system and diarrheal disease transmission: a case study in Uzbekistan. *Am. J. Trop. Med. Hyg.* **59**(6): 941–946.
105. Quick, R. et al. (1999). Diarrhea prevention in Bolivia through point-of-use water treatment and safe storage: A promising new strategy. *Epidemiol. Infect.* **122**(1): 83–90.
106. Quick, R. et al. (2002). Diarrhea prevention through household-level water disinfection and safe storage in Zambia. *Am. J. Trop. Med. Hyg.* **66**(5): 584–589.
107. Adebisi, S.A. (2003). Medical Consultant & Director, Community-Based Health Program, Chemical Pathology & Immunology Department. University of Ilorin, Ilorin, personal communication, June 14, 2003.
108. Carter Center. Former President Jimmy Carter travels to Africa for guinea worm eradication. News release, Mar. 17, 1995. Online: http://www.cc.emory.edu/Carter_Center/RLS95/GWtrip.htm.
109. Carter Center. USAID joins Carter Center for final assault on guinea worm. Carter Center News, Fall 1994. Online: http://www.cc.emory.edu/Carter_Center/CCN-F94/gw.htm.
110. Young, B. and Briscoe, J. (1987). A case-control study of the effect of environmental sanitation on diarrhea morbidity in Malawi. *J. Epidemiol. Community Health* **42**: 83–88.
111. Esrey, S.A., Potash, J.B., Roberts, L., and Shiff, C. (1991). Effects of improved water supply and sanitation on ascariasis, diarrhea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma. *Bull. World Health Organ.* **69**(5): 609–621.
112. Alam, N. (1989). Mother's personal and domestic hygiene and diarrhea incidences in young children in rural Bangladesh. *Int. J. Epidemiol.* **18**: 242–247.
113. Aung, M.H. and Thein, H. (1989). Prevention of diarrhea and dysentery by hand washing. *Trans. R. Soc. Trop. Med. Hyg.* **83**: 128–131.
114. Baltazar, J. (1988). Can the case-control method be used to assess the impact of water supply and sanitation on diarrhea? A study in the Philippines. *Bull. World Health Organ.* **66**: 627–635.
115. Cairncross, S. and Cliff, J.L. (1987). Water use and health in Mueda, Mozambique. *Trans. R. Soc. Trop. Med. Hyg.* **81**: 51–54.
116. Esrey, S.A. and Habicht, J.P. (1986). Epidemiologic evidence for health benefits from improved water and sanitation in developing countries. *Epidemiol. Rev.* **8**: 117–128.
117. Henry, F.J. (1981). Environmental sanitation infection and nutritional status of infants in rural St. Lucia, West Indies. *Trans. R. Soc. Trop. Med. Hyg.* **75**: 507–513.
118. Rahman, M. (1985). Impact of environmental sanitation and crowding on infant mortality in rural Bangladesh. *Lancet* **8445**(2): 28–31.
119. Haines, M.R. and Avery, R.C. (1982). Differential infant and child mortality in Costa Rica: 1968–1973. *Population Stud.* **36**: 31–43.
120. Khan, M.U. (1982). Interruption of shigellosis by hand-washing. *Trans. R. Soc. Trop. Med. Hyg.* **76**: 164–168.
121. Torun, B. (1982). Environmental and educational interventions against diarrhea in Guatemala. In: *Diarrhea and Malnutrition: Interactions, Mechanisms, and Interventions*. L.C. Chen and N.S. Scrimshaw (Eds.). Plenum Press, New York, pp. 235–266.
122. Anker, R. and Knowles, J.C. (1980). An empirical analysis of morbidity differentials in Kenya at the macro- and micro-levels. *Econ. Dev. Cul. Change* **29**: 165–185.
123. Koopman, J.S. (1978). Diarrhea and school toilet hygiene in Cali, Colombia. *Am. J. Epidemiol.* **107**: 412–420.
124. Misra, K.K. (1975). Safe water in rural areas. *Int. J. Health Educ.* **18**: 53–59.

READING LIST

Agardy, T. (1997). *Marine Protected Areas and Ocean Conservation*. Academic Press, Austin.

Falkenmark, M. and Lindh, G. (1993). Water and economic development. In: *Water in Crisis*. P. Gleick (Ed.). Oxford University Press, New York, pp. 80–91.

Safe Drinking Water Communication. (1980). *Drinking Water and Health*, Vol. 3. National Academy Press, Washington, DC, p. 415.

NITRATE HEALTH EFFECTS

SEGUN MICHAEL ADE ADELANA
University of Ilorin
Ilorin, Kwara State, Nigeria

INTRODUCTION

Nitrate is a stable nitrogen (N) species under certain natural conditions and forms highly soluble compounds. These are peculiar features that allow nitrate ion to be transported in some groundwater systems to environments where it can be converted into other nitrogen species that either promote surface water eutrophication or are hazardous to humans, livestock, and the environment. Nitrate test results are usually expressed in milligrams per liter as either nitrogen (NO_3-N, sometimes written as plain N) or as nitrate (NO_3).

The following conversion factors can be useful for nitrate reporting:

1 milligram (mg) of compound expressed as nitrogen (N) is equivalent to 4.43 mg when expressed as nitrate (NO_3^-).

1 milliequivalent (meq) of compound expressed as N is equivalent to 62 mg when expressed as NO_3^-.

Therefore the nitrate reporting expressions (mg NO_3-N/L, mg NO_3/L) are used interchangeably throughout this article without any special preference for one or the other.

Problems of nitrate pollution, particularly in groundwater, are widespread in many countries of the world. In the United States, several illustrations of nitrate pollution in groundwater are available in the literature. For example, more than 30 references dealing with nitrate groundwater pollution studies in 15 of the 50 states in America have been reviewed. Information from all these references have been tabulated and presented in detail in Canter (1). Survey results from a study of 25 pesticides and nitrate in 201 rural wells in eight agricultural areas in Missouri revealed that 22% of the wells exceeded the drinking water standard for nitrate (2). A summary of the nature and geographical extent of nitrate pollution of groundwater in Nebraska is given in Exner and Spalding (3).

A total of 27 references dealing with case studies of nitrate groundwater pollution (outside the United States) have been reviewed. Of these references, 22 addressed groundwater nitrate pollution studies in Europe. One of these references (4) is a conference proceeding with 52 papers addressing nitrogen as a surface and groundwater pollutant; Tessendorff (5) and Kraus (6) provided a general discussion of groundwater nitrate problems in the European community countries (EEC). The countries noted with adequate references on groundwater nitrate problems include Czechoslovakia, Germany, The Netherlands, United Kingdom, Denmark, Israel, and Chile.

Overgaard (7), in a nationwide investigation of nitrate concentrations in groundwater in Denmark, based on analyses of samples from about 11,000 wells and drinking water from 2800 groundwater works, revealed that the overall mean level of nitrate in groundwater had trebled within the last 20–30 years and is increasing at a rate of about 3.3 mg NO_3/L per year. Consequently, the result showed that 8% of the water produced in Danish Waterworks now has a nitrate concentration above the EEC guide limit of 50 mg NO_3/L (6). It has been estimated that 800,000 people in France, 850,000 in the United Kingdom, and 2.5 million in Germany are drinking water whose nitrate concentrations are above the permissible limit of the European Community (1).

A survey conducted in France revealed that 81% of the population had nitrate levels less than 25 mg NO_3/L and 96 to 98% had levels less than 50 mg NO_3/L in their drinking water supplies (8). Out of the 53 million people accounted for in the survey, 280,000 at most had a water supply exceeding 100 mg NO_3/L at least once during a 3-year period. Of 20,000 distribution units surveyed, nearly 1000 had nitrate above 50 mg NO_3/L; however, only 61 units were above 100 mg NO_3/L. Most of the high nitrate levels were found in groundwater supplies.

Custodio (9) reported that agricultural nitrate pollution is a widespread problem in irrigated areas in Spain, where nitrate in groundwater often exceeds 50 mg NO_3/L and sometimes reaches 500 mg NO_3/L. A regional survey of nitrate in the Anglian area of the United Kingdom in 1975 indicated that 50 public supply boreholes, wells, and springs had recorded nitrate levels in excess of 11.3 mg/L (as nitrogen), 50 mg NO_3/L (10). Due to increasing agricultural activity after the 1960s, both shallow and deep-water resources in the Czech Republic, including karstic systems, have been contaminated by infiltrating nitrate. The nitrate content of one of the largest springs (yielding up to 19 L/s during minimum discharge) in the Republic now varies from 50 to 60 mg NO_3/L (11). Others reported nitrate pollution studies in Canada (12,13), India (14,15), Israel (16), Chile (17), Portugal (18), Southern Africa (19,20), Nigeria (21–25), Ghana (26), Burkina Faso (27), and Senegal (28).

Note the pollution of groundwater from industry and waste dumps is a serious problem, particularly in the more developed countries in the European community. This is also the case in the United States, Canada, and Australia. High concentrations of nitrates are the other main cause of groundwater pollution. Concentrations of nitrate that approach or exceed 10 mg/L as N, equivalent to 44.3 mg/L as NO_3, present health hazards. Thus, the international drinking water quality standard is set at 10 mg/L for NO_3-N (or plain nitrogen), and it is approximately 45 mg/L as NO_3 (29). Therefore, it is important to consider the health implications of this common pollutant of groundwater with respect to humans, livestock, and the environment, which is the main focus of this paper.

OCCURRENCE AND EXTENT OF NITRATE IN WATER AND FOODS

The occurrence of nitrate concentrations has been reported in many parts of the world such as Europe, the United States, Australia, Chile, Ghana, South Africa, Nigeria, and Cote d'Ivoire. The World Health Organization (WHO) studied the occurrence of nitrate in water and came to the conclusion that nitrate concentrations in surface waters have increased substantially over the last 30–40 years (29). Many countries, mostly in Western Europe and the United Kingdom, showed a more marked increase in the levels of nitrates in groundwater, especially between 1970 and 1980 (30). The reason for these increases in groundwater nitrate is not unconnected with the vast increase in fertilizer application and other forms of animal manure.

In the United States, the occurrence of high nitrate concentrations in groundwater is widespread, particularly as a result of agricultural usage of fertilizers or land disposal of domestic wastewaters. Much research has been conducted to determine the amounts of nitrates in drinking water wells. "USGS data show that the 20 states with the largest agricultural marketing in 1989 had a notably higher percentage of wells with nitrate concentrations above 10 mg NO_3-N/L than the remaining 30 states, 7.1% compared to 3.0%, respectively" (31). Research conducted by private firms also links high nitrogen content in wells with agricultural activity. A survey of 1430 randomly selected and sampled drinking water wells in agricultural areas of 26 states, conducted in 1988 and 1989 by the Monsanto Agricultural Products Company, found nitrate above 10 mg NO_3-N/L in 4.9% of the wells. For wells on farmsteads only, however, the proportion was 10% (31). Monsanto, however, concluded that the frequency of wells with nitrate exceeding 10 mg NO_3-N/L doubled for wells located on farm property.

Several case histories have indicated the geographical extent and seriousness of high nitrate in groundwater in the United States. For example, the Metropolitan Water District of Southern California has indicated that annually it loses 4% of its drinking water supply primarily to nitrate pollution, compared to less than 1/2% from toxic organic chemicals (1). About 12% of the wells sampled in the service area exceeded the state maximum contaminant level for nitrate.

It was estimated that by the year 2000, the groundwater in most of the water table aquifers in Salinas Valley

Table 1. Summary of Nitrate Data in the United States[a]

State	Counties Tested	Number of Samples	Average Nitrate Concentration, mg/L	Percent Over 10 mg/L
Illinois	8	286	5.76	19.9
Indiana	33	5,685	0.92	3.5
Kentucky	90	4,559	2.50	4.6
Louisiana	23	997	1.19	0.8
New Jersey	5	1,108	2.60	6.8
Ohio	80	18,202	1.32	3.0
Virginia	24	1,054	2.92	7.1
West Virginia	13	1,288	0.83	0.8

[a] Reference 33.

would have exceeded the state's drinking water standard of 45 mg/L for nitrate (32). It is expected that the rising trend in groundwater nitrate will continue for many more years, even if nitrate leaching from soils is reduced by changes in agricultural practices. A summary of nitrate data by state is presented in Table 1. This was the result of the multistate groundwater quality testing undertaken for several years at Heidelberg College in Tiffin, Ohio and reported by Swanson (33). The data indicate higher nitrate occurrence in the sampled groundwater from Illinois, Virginia, New Jersey, and Kentucky.

The extent of nitrate and nitrite as (both anthropogenic and natural) water pollutants has been widely reported and published in the hydrogeologic and biochemical/pharmacological literature, especially, in developed nations (1,13,34–43). Such information is scanty in developing countries until recently (19–27,39,44–48). It has been reported that nitrate health hazards are also posed greatly by solid foods from agricultural and diary products and preservatives (24,44,49), yet the greatest threats come from polluted water; fluids and foods cooked and washed are contaminated directly or indirectly with such waters. The occurrence of nitrate is known in foods, especially in African foods. Okonkwo et al. (49) researched and reported high concentrations (in parts per million) for various African foods (Table 2).

Most of these foods are heavily consumed in West Africa and are now largely exported to Europe, America, Canada, and Australia. If the rate and volume at which these foods are consumed continues, nitrate could become potentially hazardous. A high percentage of the world population, as projected by the WHO (29), would ingest increased nitrates in the near future. Nitrates are easily converted to nitrite (the more poisonous form of nitrogen) by various mechanisms. It has been observed that drying tends to change the nitrite content of food items (Table 3). Drying is a common method of food preservation in the tropics, particularly in developing countries. Drying reduces the ascorbate level in food thereby inhibiting its antagonism to the carcinogenic action of nitrites (50). Thus, dried foods especially vegetables, have high nitrite concentrations and consequently increase the hazard. Ezeonu (51) reported high nitrate concentration in some Nigerian beers with frightening statistics. All ten brands of beer selected for study whose production locations are in different parts of the country showed high nitrate content. Yet large volumes of beer are consumed daily in Nigeria. Most breweries in Nigeria do not treat their water sources (i.e., boreholes) for nitrate, yet anthropogenic nitrate is widespread in the country (21).

Table 2. Nitrate Levels in Typical African Foods[a]

Beverages	Range, ppm	Mean, ppm
Palm Wine	0.1–1.4	1.2
Tap water (pooled urban supply)	0.1–0.3	0.2
Cereals		
Guinea Corn & Rice	2.0–4.9	3.0
Proteins (Animal Source)		
Crayfish	18.2–30.8	28.7
Fish	7.9–10.4	8.6
Meat	0.4–1.7	0.7
Proteins (Plant Source)		
Beans (black)	3.2–6.6	4.9
Bean (white)	3.1–6.5	4.9
Ground nuts	4.4–9.7	6.1
Melon	9.0–11.6	10.3
Pigeon pea	1.9–2.7	2.4
Vegetables and Fruits		
Bitter leaf	8.0–9.8	8.9
Flutted pumpkin leaf	0.8–1.2	1.1
Garden egg leaf	0.7–1.2	1.0
Green amaranth leaf	0.5–0.8	0.7
Okra	2.0–2.2	2.1
Pumpkin lead	5.6–15.0	11.3
Nitrate Not Detected in these Foods		
Beer (premier brand)	Corn	Sida
Native Gin	Cassava (Gari)	Yams

[a] After Reference 49.

WATER QUALITY CONCERNS RELATED TO NITRATE LEVELS

The issue of water quality related to nitrate concentrations, especially in drinking water, is of concern worldwide in view of the various health implications of ingesting high doses of nitrates. Usually, drinking water contributes <30% of the total dietary nitrate intake (52), but when the drinking water nitrate level exceeds 10 mg/L (as N), the

Table 3. Changes in Nitrite Content of Food Items After Drying[a]

Food Item	Number of Samples Analyzed	Change on Drying, ppm Mean Values
Meat	17	0.7–2.0
Bitter leaf	17	9.4–14.2
Flutted pumpkin leaf	17	1.1–2.8
Garden egg leaf	17	1.0–10.5
Green amaranth leaf	17	2.1–14.7
Pumpkin leaf	17	11.5–21.7

[a] After reference 49.

contribution may become considerably higher and result in health problems. This is why the World Health Organization (WHO) originally set the standard at 10 mg NO_3-N/L for nitrate in drinking water (29), although the first limit was set by the United States Public Health Service in 1962 (53). Therefore, most governmental authorities consider approximately 10 mg/L (as NO_3-N) or 45 mg/L (as NO_3) an acceptable limit for nitrate in drinking water for their countries (Table 4). The United States Environmental Protection Agency (USEPA) standard for nitrate in drinking water is set at 10 mg/L NO_3-N to protect babies under about 3 months of age, the most nitrate sensitive segment of the U.S. population. This level of 10 mg/L NO_3-N has been determined as the No Observable Adverse Effect Level (NOAEL) and is therefore considered the safe drinking water limit for nitrate. This was based on the epidemiological studies carried out mainly in the United States (e.g., References 54,55).

However, on the basis of their epidemiological studies, both Super et al. (56) and Hesseling et al. (57), argued that up to 20 mg NO_3-N/L of nitrate would still be an acceptable limit as it presented a low risk to infants. Nevertheless, the maximum permissible nitrate level (in drinking water) still remains 45 mg/L (as NO_3) and 10 mg/L (as NO_3-N), as set by the World Health Organization (29).

The concentration of nitrates in groundwater is of primary concern due to potential human impacts from its usage. Depending on the use of the groundwater, animals, crops, and industrial processes could also be affected (1). The United States Environmental Protection Agency (58,59) conducted a 5-year National Survey of Nitrates in drinking water wells, the main aim was developing national estimates of the frequency and concentration of nitrates in drinking water wells in the United States in two phases. The drinking water wells include both community water system (CWS) wells and rural domestic wells (RDW). Based on the results from phase I, 52.1% (49,300) of all the CWS wells (94,600), it was estimated, contain detectable nitrate, of which 45% of the pollution was due only to nitrate (1). Out of the 827 RDW wells sampled, it was estimated that 57% had detectable nitrate (60). A total of 1.2% of the CWS wells and 2.4% of the RDW wells, it was projected, have nitrate-nitrogen concentrations above the health-based level of 10 mg/L (60).

The phase I survey report further identified some of the chemical characteristic factors related to detection and concentration of nitrates, and that from the phase II survey included estimates of national population exposure and the resultant health risks due to nitrate for RDW and CWS wells (59). Estimates were provided of the populations corresponding to quartiles of general interest (e.g., 95th and 99th percentiles) and of the number of individuals exposed above health-based levels (1). The information related to CWS wells is summarized in Tables 5 and 6.

SOURCES OF NITRATES IN GROUNDWATER AND FOODS

Agriculture is the primary source of elevated nitrate levels, although in some rare cases certain geologic units can

Table 4. Limits and Standard Guideline Value for Nitrate in Drinking Water[a]

			Concentration, mg/L[b]	
Organization	Year	Limit of Specification	As NO_3	As N
WHO (European standard)	1970	Recommended	50	(11.3)
		Acceptable	11.3–22.6	
WHO (International)	1971		45	10.2
WHO	1984	Guide value	(44.3)	10
US EPA[c]	1977		(44.3)	10
European Communities	1980	Maximum admissible	50	(11.3)
	1980	Guide level	25	(5.6)
Health & Welfare, Canada	1978		(44.3)	10
Mexican Standard[e]	1988	Maximum permissible	22	5
SABS[f]	1984	Recommended[d]	(26.6)	6
		Maximum allowable[d]	44.3	10
NFEPA[g]	1991	Recommended	45	10

[a] Modified from References 1 and 52.
[b] Brackets indicate derived units.
[c] U.S. Environmental Protection Agency.
[d] Nitrate plus nitrite.
[e] From Reference 98.
[f] South African Bureau of Standards.
[g] Nigerian Federal Environmental Protection Agency.

Table 5. Estimates of Population Exposed to Nitrate in Community Water System Wells by Distribution Percentile[a]

			95% Confidence Interval	
Percentile	People Exposed	Concentration, mg/L	Lower Bound, mg/L	Upper Bound, mg/L
Median	68,000,000	0.63	0.45	0.95
95	6,800,000	6.52	5.34	7.60
99	1,360,000	14.2	10.6	17.7

[a] Reference 59.

Table 6. Estimates of Population Exposed to Nitrate by Concentration in Community Water System Wells[a]

		95% Confidence Interval	
Concentration, mg/L	Population Exposed	Lower Bound, mg/L	Upper Bound, mg/L
All concentrations >0	85,300,000	78,100,000	98,900,000
≥10	2,980,000	1,600,000	4,260,000

[a] Reference 59.

be the origin of the nitrate. Nitrate sources other than applied fertilizers such as wastes from livestock, dairy, or poultry and accidents or careless precautionary handling of fertilizers near well sites may be involved (31). In the United States, for example, there are locations with high agricultural production in at least 14 states where nitrate contamination has been associated with the application of nitrogen fertilizers (1). In certain categories, organic and inorganic compounds of nitrogen, phosphorus, and potassium that originate from many commercial fertilizers may be released into groundwater. In some cases, it has been demonstrated that the leaching of nitrate is accelerated by irrigation (1,31,61).

Septic tank systems also represent a significant fraction of the nitrogen load to groundwater in the United States (62). About 25% of the population in America is served by individual home sewage disposal systems. Research revealed that effluent from a typical septic tank system has a total nitrogen content of 25 to 60 mg/L (62). In the Netherlands, the reason for the rising nitrate level can be due to the application of nitrogen fertilizer (63,64). Jacks and Sharma (15) reported nitrate levels in excess of 300 mg/L (as N) in wells in Southern India owing to anthropogenic and agricultural influences. In Australia, biological fixation in the soil is considered the principal origin, although point sources such as sewage effluent, animal and industrial waste could be significant locally (65). However, in Nigeria and most parts of West Africa, high nitrate levels in groundwater result mostly from indiscriminate waste disposal (21–25,27) and agricultural activities (26,46,47). Table 7 is a summary of nitrate sources in groundwater.

HEALTH EFFECTS OF NITRATE

Nitrate and the nitrite form of nitrogen constitute a general public health concern, related especially to infant methemoglobinemia (infantile cyanosis) and carcinogenesis (66). The concentrations of nitrate and nitrite in foods that include vegetables, crayfish, meat, etc. and drinking water may indicate serious potentials for pollution and also could result in severe health problems. For example, the nitrate levels in Nigerian foods that include drinking water (from surface and subsurface sources) and beverages are reportedly high and generally perceived to be associated with adverse health effects in humans (24,49,51,67). These have resulted in reported cases of water-related diseases such as diarrhea in children or cancerous diseases that claimed lives yearly (24,68,69). According to Population Reports (70), diarrheal disease is a class of waterborne disease, which is prevalent in many countries where sewage treatment is inadequate or where human wastes are disposed of in open latrines, ditches, canals, and watercourses or is spread indiscriminately on farmland. These practices are frequent in developing nations and favor the accumulation of anthropogenic nitrate.

Certain vegetables (e.g., lettuce, spinach, beetroot, and celery) contain relatively high levels of nitrate [>3000 mg/kg for lettuce (Ref. 71)] but the nitrite levels are usually very low. Nitrates and nitrites are also added as preservatives in some foods, such as cured meats, consequently exposing consumers to higher health risks. The World Health Organization (29) estimated daily dietary intake of nitrate and nitrite in different countries.

In most European countries, the mean nitrate intake is about 10–30 mg/day. Vegetarians usually have a two to fourfold higher intake of nitrates than nonvegetarians. In India, it has been estimated that 20–50% of the wells in areas of high population density produce water whose nitrate level is above 50 mg/L, thus causing severe health hazards (72,73). Terblanche (74) reviewed the health hazards of nitrate in drinking water in many developed countries, including South Africa. Of an estimated 219 million people using public drinking water

Table 7. Examples of the Various Sources of Nitrate in Groundwater[a]

Natural Sources
Geologic nitrogen which can be mobilized and leached to groundwater via irrigation practices
Unmanaged (natural) climax forests that are normally nitrogen conserving; however, nitrogen losses to groundwater can occur from human-initiated clear cutting and other forest disturbances
Waste Materials
Animal manures, which may be concentrated in large commercial poultry, dairy, hog, and beef operations
Land application of municipal or industrial sludge or liquid effluent on croplands, forests, parks, golf courses, etc.
Disposal of household wastes or small business wastes into septic tank systems (septic tank plus soil absorption field)
Leachates from sanitary or industrial landfills or upland dredged material disposal sites
Row Crop Agriculture[b]
Nitrogen losses to the subsurface environment can occur as a result of excessive fertilizer application, inefficient uptake of nitrogen by crops, and mineralization of soil nitrogen
Nitrogen losses to the subsurface environment can occur as a function of fertilizer application rates, seasonal rainfall and temperature patterns, and tillage practices
Irrigated Agriculture
Enhanced leaching of nitrogen from excessive fertilizer application rates and inefficient irrigation rates
Associated leaching of nitrogen from soils periodically subjected to leaching to remove salts so that the soils do not become saline and unproductive

[a] Culled from Reference 1.

[b] Refers to annual crops.

supplies in the United States, approximately 1.7 million are exposed to nitrate levels above 10 mg/L. About two-thirds of those exposed, 1.1 million, are served by public water systems using groundwater supply sources. Almost 27,000 infants a year are exposed to tap water with nitrate levels exceeding 10 mg/L (31). The resulting health hazards and associated statistics in the United States are documented in the Federal Register (75). The following section describes the details of the various health effects of nitrate.

Methemoglobinemia

High nitrate levels in water can cause infant methemoglobinemia. Methemoglobinemia is a disease primarily affecting babies and is often described by the lay term "blue baby syndrome." Infants are the primary concern because they are the most vulnerable. The USEPA standard for nitrate in drinking water is set at 10 mg/L to protect babies under about 3 months of age. Such infants are much more sensitive to nitrate toxicity than the rest of the population for many reasons. For example, bacteria that live in the digestive tracts of such infants convert nitrate into toxic nitrite.

Nitrite transforms hemoglobin to methemoglobin, preventing transport of oxygen and producing symptoms of asphyxiation (another term for blue baby syndrome). This methemoglobin is considerably more stable than the oxygen hemoglobin complex that fulfils the oxygen transport function of the blood. Once the concentration of methemoglobin in the blood exceeds 5% the first symptoms of 'cyanosis' are generally noticeable; anoxia (death) results at levels of 50% and higher (74) or if the condition is left untreated (66). After babies reach the age of 3 to 6 months, acid in their stomachs increases, thereby creating an unfavorable environment for the bacteria causing the problem (31). It must be borne in mind that nitrate itself has low primary toxicity, but acute toxicity occurs as nitrate is reduced to nitrite (NO_2), a process that can occur under specific conditions in the stomach and saliva (66). Consequently, the nitrite ion formed becomes an oxidizing agent, transforming hemoglobin in the blood to methemoglobin (29), thereby preventing transport of oxygen and resulting in methemoglobinemia. Most reported cases of infantile methemoglobinemia have been associated with the use of water containing more than 10 mg/L NO_3-N.

The occurrence of infant methemoglobinemia from consumption of water with high nitrate concentrations was first recognized clinically by Comly (72). The infants were both less than 1 month old and had received rural well water containing 90 and 140 mg/L, respectively. Earlier, Comly (72) suggested a recommended limit of 10 mg/L NO_3-N in drinking water and a maximum of 20 mg/L.

Later on, Shuval and Gruener (76) studied 1702 infants living in the Israel coastal plain in areas with medium to high nitrate (11.3 to 20.3 mg NO_3-N/L) and compared them with a control group of 759 infants in Jerusalem where only 1.1 mg/L of nitrate is in the water supply. There were no significant differences found between the methemoglobin levels in the 1702 infants in the study areas compared to the 758 infants in the control area. In most countries, methemoglobinemia is not a notifiable disease, making its true incidence unknown. From 1945 until 1970, some 2000 cases of methemoglobinemia have been reported in the world literature (76) with a case fatality of about 8%. The WHO (29) cites literature indicating that 10 cases of methemoglobinemia have been reported in the United Kingdom since 1950 when the first cases of methemoglobinemia were reported in East Anglia. Only one death was reported during this period. In 1986, a 2-month-old infant in South Dakota

(USA) died of methemoglobinemia (77). The exact nitrate concentration is unknown. In another nonfatal case in Iowa, the water apparently contained 285 mg/L nitrate (as N) but the 5-week-old infant survived (42). Hungary is one of the countries, with exceptions, that possibly has the best statistics on the occurrence of infantile methemoglobinemia. Table 8 shows the occurrence of methemoglobinemia in Hungary between 1976 and 1990 (78).

Methemoglobinemia became a notifiable disease in Hungary in 1968 (29), and in the first 5 years after 1968, 883 cases were reported. Of the recorded cases, 92% had a nitrate level in the drinking water exceeding 22.6 mg/L as N; in the remaining 8%, it was between 9 and 22.6 mg/L. The highest number of cases was reported in 1977, and the measures taken to supply the population with drinking water low in nitrate have resulted in a definite decrease in the number of cases each year. For detailed statistics on infant methemoglobinemia, readers are referred to WHO (29) and Csanady (78).

In a later publication, Shuval and Gruener (79) confirm a direct relationship between the occurrence of methemoglobinemia in infants and high concentrations (>10 mgNO_3-N/L) of nitrate in water. According to Ross and Desforges (80), other factors important in the pathogenesis of the disease are age, the presence of bacteria in sufficient numbers in the gastrointestinal tract, gastric acidity (a pH >4), gastrointestinal disturbances, the types of powdered milk product used as baby food, high fluid intake, and the effect of nutrition because foods rich in nitrate can increase the severity of illness.

The first epidemiological survey in South Africa to assess the effect of well water nitrates on infant health was published (57) after the review of health hazards by Terblanche (74). The survey was undertaken due to the risk of methemoglobinemia in infants in the Rietfontein area as a result of the large number of boreholes where nitrate-nitrogen exceeded 10 mg/L. Unfortunately, no correlation was found between the nitrate content of the groundwater used and the methemoglobin levels in the blood. A clinical health risk assessment was attempted in South Africa due to high ingestion of nitrate water (81). There was an increased risk of methemoglobinemia as a result of increased in bottle-feeding by HIV positive mothers. In South Africa, it has been shown that breastfeeding increases the risk by 12–43% that HIV-positive mothers transmit the virus to their children (82). For this reason, the South African Department of Health advises HIV-positive mothers to bottle-feed infants to reduce the risk of mother to child transmission of the HIV virus via breast milk (81). This assessment attempts to quantify the potential additional total exposed population (PATEP) facing increased risk from nitrates. Table 9 shows the potential additional total exposed population at significant risk of methemoglobinemia as a result of bottle-feeding infants by HIV-positive mothers. The number of infants has been calculated as proportional to the area of the province in which groundwater contains >10 mg/L NO_3-N (81).

Table 8. The Occurrence of Methemoglobinemia in Hungary[a]

Year	Number of Cases	Number of Fatalities
1976	207	4
1977	293	7
1978	239	3
1979	180	2
1980	172	3
1981	166	1
1982	91	1
1983	67	—
1984	33	—
1985	46	1
1986	41	—
1987	30	—
1988	31	2
1989	35	2
1990	22	—
TOTAL	**1653**	**26**

[a] Reference 78.

According to Colvin (81), Gauteng has 10% of the PATEP for >10 mg/L NO_3-N groundwater, due mainly to the high infant population density (7.5 per km^2). This province includes over half of the total area underlain by groundwater with >50 mg/L NO_3-N and has 88% of the potential additional population exposed to very high risk (in South Africa), shown in Table 10.

The qualitative banding of degrees of risk following the matrix model developed by Carpenter and Maragos (83) and simplified by Genthe (84) is shown in Fig. 1.

Severity is defined according to the potential for exposure to nitrate in groundwater sources. For example, a risk that combines high severity with frequent probability is rated "high," whereas a risk that combines low severity with occasional or the remote probability of occurrence is rated "acceptable." The highest risk is for infants <3 months old who are bottle-fed using groundwater with >50 mg/L NO_3-N, and confounding factors (such as lack of vitamin C and gastrointestinal infections) are evident (81).

In another epidemiological study on a comparable population group in Namibia (56), a correlation was found between the nitrate level in the groundwater and blood methemoglobin levels. The main difference between the two studies was that the level of nitrate in groundwater in the Namibian study was much higher, up to 56 mg/L of nitrate-nitrogen (52). However, the ingestion of nitrate has no apparent short-term effects on adults such as methemoglobinemia. Research has shown that adults on a farm near Otjiwarongo, Namibia, continued drinking water with 268 mg/L of nitrate with no apparent ill effects, even after stock losses occurred on that farm (85).

Gastric Cancer

Infants are not the only ones at risk; it is possible that high nitrate concentrations can cause cancer in adults. "Nitrate itself is not directly carcinogenic. However, there is recognition of the fact that nitrate could be converted to nitrite in the human body that can react with secondary and tertiary amines to form nitrosamines—which have been identified as potent carcinogens" (31). Several studies have shown that simultaneous ingestion of nitrite (or nitrate with amines) results in cancers of many organ

Table 9. Potential Additional Total Population Exposed to Significant Risk of Methemoglobinemia as a Result of Bottle-Feeding by HIV-Positive Mothers[a]

Province	>10 mg/L NO_3-N, km^2	%Area of Province	Total Infants <1 yr/Province, 1996	%HIV+, Oct/Nov 1998	Estimated% Groundwater Dependency	PATEP[b]
Free State	45,739	35	48,418	23	30	1167
Gauteng	1,760	10	136,373	23	30	896
Mpumalanga	6,153	8	61,678	30	6	88
North-West	80,958	70	75,097	21	80	8831
Northern Cape	210,603	58	17,711	10	35	357
Northern Province	74,434	60	119,644	12	75	6460
Eastern Cape	0	0	137,552	16	35	0
Kwazulu-Natal	0	0	182,858	33	30	0
Western Cape	765	1	76,908	5	20	5
Total South Africa	**420,411**	**34**	**856,239**	**23**	**40**	**17,804**

[a] Reference 81.

[b] PATEP: Potential additional total exposed population.

Table 10. Potential Additional Total Population Exposed to Very High Risk of Methemoglobinemia as a Result of Bottle-Feeding by HIV-Positive Mothers[a]

Province	>50 mg/L NO_3-N Area (km^2)	%Area of Province	Total Infants <1 yr per Province, 1996	%HIV+ (Oct–Nov 1998)	Estimated% Groundwater Dependency	PATEP[b]
Gauteng	985	5.45	136,373	23	30	513
Mpumalanga	160	0.21	61,678	30	6	2
North-West	416	0.36	75,097	21	80	45
Northern Province	262	0.21	119,644	12	75	22
Total South Africa	**1823**	**0.001**	**856,239**	**23**	**30**	**582**

[a] Reference 81.

[b] PATEP: Potential additional total exposed population.

systems (1). The N-nitroso compounds are presumed to be the ultimate carcinogenic substances (86). The role of nitrite as a precursor to carcinogenic nitrosamines and other N-nitroso compounds has been established (1). Nitrite reacts with amines or amides under several conditions resulting in N-nitroso derivatives, the majority of which are carcinogenic to animals. The expectation that these N-nitroso derivatives are also human carcinogens suggests a mechanism whereby exposure to nitrite might result in carcinogenesis (86).

In earlier publications, Wolff and Wasserman (87) reviewed 60 articles on the potential hazard of nitrate, nitrite, and nitrosamines in the environment. Furthermore, Shuval and Gruener (79) refer to the possible carcinogenic, teratogenic, and mutagenic properties of nitrosamines, which can theoretically develop in food or in the human digestive tract if nitrates and nitrites are exposed to secondary and tertiary amines under certain conditions. In both cases, no conclusions were drawn as to the real nature of the potential risks involved. From all available data, the WHO (29) came to the conclusion that no convincing evidence of a relationship between gastric cancer and consumption of drinking-water containing nitrate levels up to 10 mg/L has emerged, yet a link cannot be ruled out; perhaps the inadequacy of the data available would explain this.

However, Hartman (88) in his publication showed a graphical relationship between gastric cancer mortality rates and nitrate ingestion in twelve countries (Fig. 2). Hartman's (88) data on gastric cancer mortality rates are for the years 1974–1975. This graph refers to total nitrate ingestion and does not specifically refer to ingestion via drinking water. Tredoux (52), however, questioned whether the seemingly good correlation is actually related only to nitrate ingestion or possibly also to other dietary factors relating to the various population groups.

Several epidemiological studies have indicated further significant positive correlations between exposure to nitrate and cancer risk. For instance, nitrate in drinking water has been correlated with gastric cancer risk in Colombia and England, and exposure to nitrate-containing fertilizers appeared to be linked to gastric cancer mortality in Chile (1). It should be noted that high risk for gastric cancer correlates with nitrate, and also with several other dietary or environmental factors. These associations, however, cannot be fully traced to the causation of gastric cancer (86,89).

Other Health Effects

The ingestion of nitrate in excess of 10 mg NO_3-N/L is also the possible cause of a number of other adverse effects such as birth defects (congenital malformations), which were studied in South Australia and also in eastern England. Detailed investigation has provided no evidence that these were associated with nitrate levels in water supplies (29). It has also been suggested that

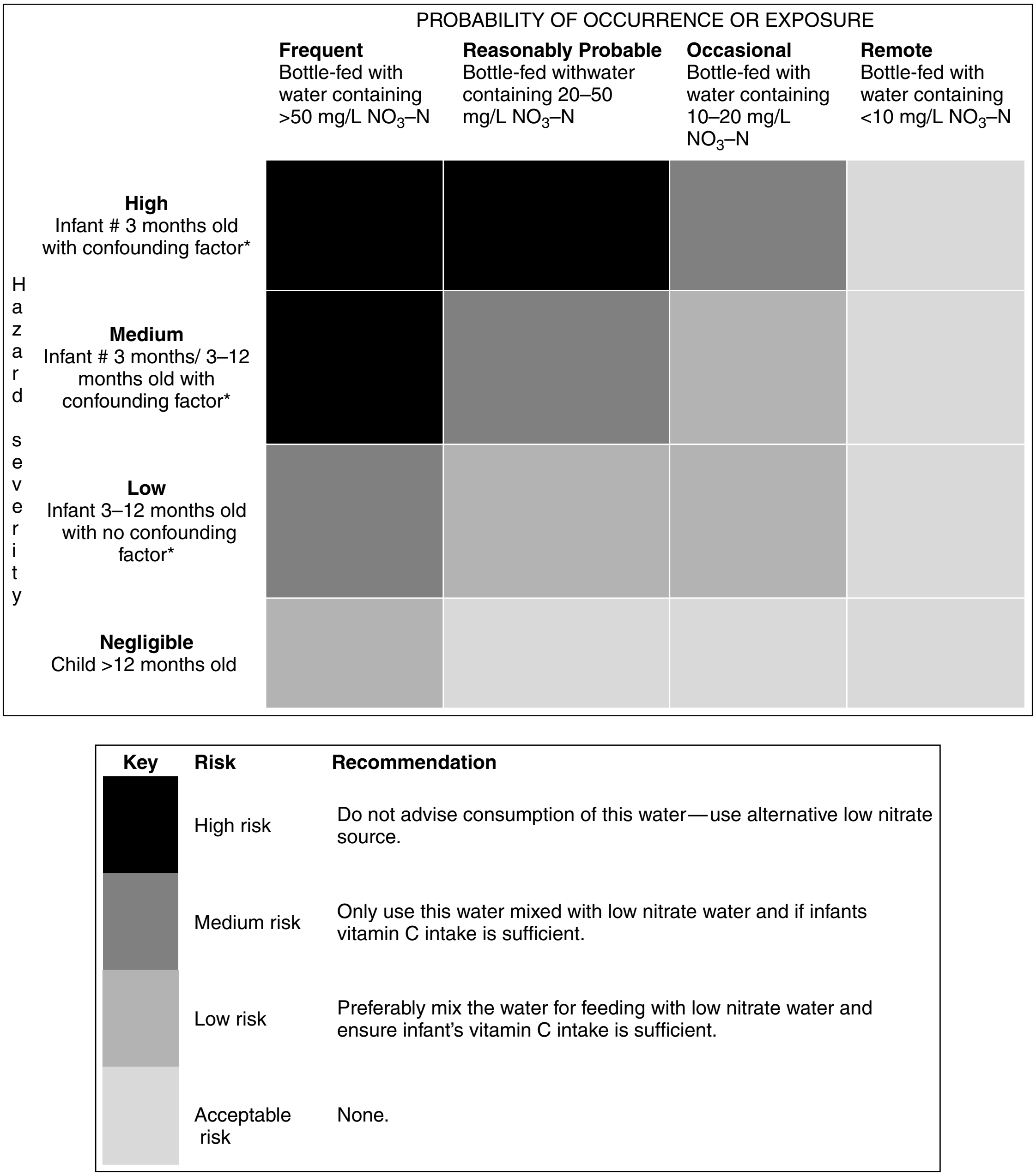

Hazard severity	PROBABILITY OF OCCURRENCE OR EXPOSURE			
	Frequent Bottle-fed with water containing >50 mg/L NO_3–N	**Reasonably Probable** Bottle-fed withwater containing 20–50 mg/L NO_3–N	**Occasional** Bottle-fed with water containing 10–20 mg/L NO_3–N	**Remote** Bottle-fed with water containing <10 mg/L NO_3–N
High Infant # 3 months old with confounding factor*	■	■	▓	░
Medium Infant # 3 months/ 3–12 months old with confounding factor*	■	▓	▒	░
Low Infant 3–12 months old with no confounding factor*	▓	▒	▒	░
Negligible Child >12 months old	▒	░	░	░

Key	Risk	Recommendation
■	High risk	Do not advise consumption of this water—use alternative low nitrate source.
▓	Medium risk	Only use this water mixed with low nitrate water and if infants vitamin C intake is sufficient.
▒	Low risk	Preferably mix the water for feeding with low nitrate water and ensure infant's vitamin C intake is sufficient.
░	Acceptable risk	None.

Figure 1. Risk characterization of methemoglobinemia (after References 83 and 84).

chronic exposure to high levels of nitrate in drinking water may have adverse effects on the cardiovascular system. The WHO (29), however, reported that an inverse relationship between cardiovascular mortality and nitrate concentration in water supplies had been demonstrated.

Moreover, excessive nitrates in drinking water have also resulted in problems with ruminants (cud-chewing animals with divided stomachs). Sheep and cattle, in particular, can be seriously affected by nitrates from birth through adulthood (1). Infants of monogastric (single-stomach) animals like horses, pigs, and chickens are also susceptible to problems from nitrate ingestion. However, as chickens and pigs mature, they are much less susceptible to the health effects of nitrate, but horses can be affected through adulthood (90).

Other possible effects of nitrates relate to the thyroid function in animals. Some animal studies indicate that chronic exposure to high levels of nitrates can reduce the intrathyroid iodine pool and thus render the gland more sensitive to goitrogens (29). However, whether or not exposure to nitrate is an etiological factor in human goiter remains unclear (52).

Symptoms of nitrate-nitrite poisoning in livestock include cyanosis in or about the non pigmented areas (mouth and eyes), shortness of breath, rapid heartbeat, staggered gait, frequent urination, and collapse. In severe cases, convulsions, coma, and death may result within a

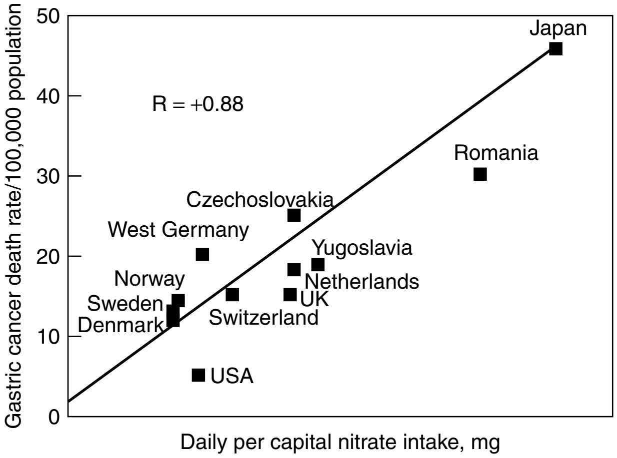

Figure 2. Relationship between gastric cancer mortality rates and nitrate ingestion in 12 countries (after Reference 88).

few hours (1). Loss of milk production in cows and aborted calves are also indicative of nitrate poisoning (90). Stock losses due to nitrate poisoning have been reported in Namibia (91), in the dolomitic area of South Africa (92), and in Bophuthatswana (52).

Exposure to high doses of nitrate is also associated with adverse effects such as the 'hot-dog headache' (52). The 'hot-dog headache' has been described in the literature as related to nitrites used in curing meat to give it a uniform color (93). Nitrites are also vasodilators, so that some people find that soon after eating these meat products, they develop flushing of the face and headache. The 'hot-dog' is the classical example, but other meat products, including bacon, ham, and salami, can also cause these symptoms (52). A farmer in the Springbok Flats regularly complained about hot-dog headaches. The problem was solved when he started using nitrate-free water (94).

Finally, detailed information on the health effects of nitrate and nitrite in humans and animals, the mechanism and quantification of toxicological effects of nitrate and nitrite, as well as other health-related information on nitrates are available in literature (29,43,56,74,76,79, 95–97). Obviously, pollution and contamination problems from NO_3-NO_2 compounds must be causing environmental public health havoc yet to be fully determined and documented. Therefore, nitrate-nitrite pollution control programs must be established to reduce the health effects of these common and widespread contaminants. Excessive chemical fertilizer and animal manure applications must be controlled and curtailed. Exceptionally high NO_3 borehole waters must be abandoned, and unpolluted groundwater exploited. These and possible biodenitrification water treatment (for affected waters before use) could contribute solutions.

SUMMARY

The occurrence of high nitrate concentrations in groundwater is widespread, particularly from agricultural usage of fertilizers and animal manure or land disposal of domestic waste and wastewaters. Much research has been conducted to determine the amounts of nitrates in drinking water wells as well as in foods. Exposure to high doses of nitrate is generally perceived to be associated with adverse health effects in humans and other species. These range from infant methemoglobinemia, cancers, the 'hot dog headache,' and hypertension, to other adverse effects such as birth defects (congenital malformations) and spontaneous abortions. Most reported cases of infantile methemoglobinemia have been associated with the use of water containing more than 10 mg/L NO_3-N.

Pollution and contamination problems from NO_3-NO_2 compounds may be causing environmental public health havoc yet to be fully determined and documented. The evidence as outlined in this article is overwhelming. Therefore, nitrate-nitrite pollution control programs must be established to reduce the health effects of these common and widespread contaminants. Excessive chemical fertilizer and animal manure applications must be controlled and curtailed. Exceptionally high NO_3 borehole waters must be abandoned, and unpolluted groundwater exploited. These and possible biodenitrification water treatment could contribute solutions.

BIBLIOGRAPHY

1. Canter, L.W. (1996). *Nitrates in Groundwater*. Lewis, New York.
2. Sievers, D.M. and Fulhage, C.D. (1992). Survey of rural wells in Missouri for pesticides and nitrates. *Groundwater Monitoring Rev.* **12**(4): 142–150.
3. Exner, M.E. and Spalding, R.F. (1990). *Occurrence of Pesticides and Nitrate in Nebraska's Ground Water*. Report WC1, Water Center, Institute of Agriculture and Natural Resources, University of Nebraska, Lincoln, Nebraska, pp. 3–30.
4. Anonymous. (1975). *Proc. Conf. Nitrogen Water Pollutant, Int. Assoc. Water Pollut. Res.*, August, Copenhagen, Denmark.
5. Tessendorff, H. (1985). Nitrates in groundwater: A European problem of growing concern. *Aqua* **4**: 192–193.
6. Kraus, H.H. (1993). *The European Parliament and the EC Environment Policy*. Working Paper W-2, April, European Parliament, Luxembourg, p. 12.
7. Overgaard, K. (1984). Trends in nitrate pollution of groundwater in Denmark. *Nordic Hydrol.* **15**(4–5): 174–184.
8. Anonymous. (1983). Nitrate in water for human consumption: The situation in France (1979–1981). *Aqua* **2**: 74–78.
9. Custodio, E. (1982). Nitrate build-up in Catalonia coastal aquifers. *Memoires, Int. Assoc. Hydrogeol.* **16**(1): 171–181.
10. Greene, L.A. (1978). Nitrates in water supply abstractions in the Anglian region: Current trends and remedies under investigation. *Water Pollut. Control* **77**(4): 478–491.
11. Buzek, F., Kadlecová, R., and Zák, K. (1998). *Nitrate pollution of a karstic groundwater system*. In: *Isotope Techniques in the Study of Environmental Change*. IAEA, Vienna, pp. 453–464.
12. Hill, A.R. (1982). Nitrate distribution in the ground water of the Alliston region of Ontario, Canada. *Ground Water* **20**(6): 696–702.
13. Egboka, B.C.E. (1984). Nitrate contamination of shallow groundwater in Ontario, Canada. *Sci. Total Environ.* **35**: 53–70.
14. Handa, B.K. (1983). Effect of Fertilizer use on Groundwater Quality in India. *Proc. Symp. Ground Water Water Resour. Plann.*, Vol. II, IAHS Publication, No. 142, pp. 1105–1119.

15. Jacks, G. and Sharma, V.P. (1983). Nitrogen circulation and nitrate in ground water in an agricultural catchment in southern India. *Environ. Geol.* **5**(2): 61–64.
16. Ronen, D. and Margaritz, M. (1985). High concentration of solutes at the upper part of the unsaturated zone (water table) of a deep aquifer under sewage-irrigated land. *J. Hydrol.* **80**(3/4): 311–323.
17. Schalscha, E.B. (1979). Nitrate movement in a Chilean agricultural area irrigated with untreated sewage water. *J. Environ. Qual.* **8**(1): 27–30.
18. Duque, J. and Almeida, C. (2001). Nitrate contamination in a shallow gabbro hard rock aquifer, Beja—South Portugal. In: *New Approaches Characterizing Groundwater Flow*, K.P. Seiler and S. Wohnlich (Eds.). Swets & Zeitlinger Lisse, Lisse, pp. 1169–1173.
19. Tredoux, G., Engelbrecht, J.F.P., and Tama, A.S. (2001). Nitrate in groundwater in Southern Africa. In: *New Approaches Characterizing Groundwater Flow*. K.P. Seiler and S. Wohnlich (Eds.). Swets & Zeitlinger Lisse, Lisse, pp. 663–666.
20. Tredoux, G. and Kirchner, J. (1985). *The occurrence of nitrate in groundwater in South West Africa/Namibia*. Paper presented at *Conf. Nitrates in Water*, Paris.
21. Adelana S.M.A. and Olasehinde P.I. (2003). High nitrate in water supply in Nigeria: Implications for human health. *Water Resour.* **14**(1): 1–11.
22. Edet, A.E. (2000). Water Pollution by Nitrate Near Some Waste Disposal Sites in Calabar (Nigeria). *Groundwater 2000, Proc. Int. Conf. Groundwater Res.*, Copenhagen, June 6–8, pp. 239–240.
23. Uma, K.O. (1993). Nitrates in shallow (regolith) aquifers around Sokoto Town, Nigeria. *Environ. Geol.* **21**: 70–76.
24. Egboka, B.C.E. and Ezeonu, F.C. (1990). Nitrate and nitrite pollution and contamination in parts of SE Nigeria. *Water Resour.* **2**(2): 101–110.
25. Langenegger, D. (1981). High nitrate concentrations in shallow aquifers in a rural area of central Nigeria caused by random deposits of domestic refuse and excrement. In: *Studies in Environmental Science*. W. van Duijvenbooden, P. Glasbergen, and H. van Lelyveld (Eds.). Elsevier, Amsterdam, Vol. 17, pp. 147–153.
26. Akiti, T.T. (1982). *Nitrate levels in some granitic aquifers from Ghana*. *Proc. Int. Symp. Impact Agric. Activities Groundwater*, IAH Mem. XVI, Part 1, pp. 87–98.
27. Groen, J., Schuchnann, J.B., and Geinaert, W. (1988). The occurrence of high nitrate concentration in groundwater in villages in northwestern Burkina Faso. *J. African Earth Sci.* **7**(7/8): 999–1009.
28. Tandia, A.A., Gaye, C.B., and Faye, A. (1998). Origin, process and migration of nitrate compounds in the aquifers of Dakar region, Senegal. In: *Application of Isotope Techniques to Investigate Groundwater Pollution*. IAEA, Vienna, pp. 67–80.
29. World Health Organization. (1985). *Health Hazards from Nitrates in Drinking Water*. Copenhagen.
30. Schroder, H., Harremoes, P., and Simonsen, J.F. (1985). *Water Pollution Caused by Nitrogen from Urban Wastewater and from Agriculture*. Paper presented at *Conf. Nitrates Water*, Paris.
31. Fedkiw, J. (1991). *Nitrate Occurrence in Waters*. USDA, Washington, DC.
32. Anton, E.C., Barnickol, J.L., and Schnaible, D.R. (1988). *Nitrate in Drinking—Report to the Legislature*, Report No.88-11 WQ, October, Division of Water Quality, State Water Resources Control Board, State of California, Sacramento, California, pp. 35–38.
33. Swanson, G.J. (1992). Well testing program is cooperative effort. *Water Well J.* **46**(8): 39–41.
34. George, W.W. and Hustings, W.W. (1951). Nitrate in groundwater of Texas. *Trans. Am. Geophys. Union* **32**(3): 316–342.
35. Ferlmutter, N.M. and Koch, E. (1972). *Preliminary Hydrogeologic Appraisal of Nitrate in Groundwater and Streams, Southern Nassau County, Long Island, New York*. U.S. Geol. Surv. Prol. Paper, No. 800-B, pp. B225–B235.
36. Schmidt, K.D. (1972). Nitrate in groundwater of Fresno–Clovis Metropolitan area, California. *Ground Water* **10**(1): 60–64.
37. Behnke, J.J. and Haskell, E.E. (1968). Groundwater nitrate distribution beneath Fresno, California. *J. Am. Water Works Assoc.* **4**: 477–480.
38. Devitt, D., Leley, J.J., Lund, L.J., and Blair, J.W. (1976). Nitrate nitrogen movement through soil as affected by soil profile characteristic. *J. Environ. Qual.* **5**(3): 283–288.
39. Marais, S. (1999). *Dependency of Communities on Groundwater for Water Supply and Associated Fluoride and Nitrate Problems*. Paper presented at Workshop on Fluorides and Nitrates in Rural Water Supplies, Mafikeng, March 9–10.
40. Trudell, M.R., Gillham, R.W., and Cherry, J.A. (1986). An in-situ study of the occurrence and rate of denitrification in a shallow unconfined sand aquifer. *J. Hydrol.* **83**: 251–268.
41. Miller, L.W. (1971). Methemogolobinaemia associated with well water. *J. Am. Med. Assoc.* **216**: 1642–1643.
42. Rajagopal, R. and Tobin G. (1989). Expert opinion and groundwater quality protection: The case of nitrate in drinking water. *Ground Water* **27**(6): 835–847.
43. Fan, A.M. and Steinberg, V.E. (1996). Health implications of nitrate and nitrite in drinking water: An update on methemoglobinemia occurrence and reproductive and developmental toxicology. In: *Regulatory Toxicology and Pharmacology*. 23, Academic Press, London.
44. Bassir, O. and Maduagwu, E.N. (1978). Occurrence of nitrate, nitrite, dimethylamine and dimethylnitrosamine in some fermented Nigerian beverages. *J. Agric. Food Chem.* **26**(1): 200–203.
45. Faillat, J.P. (1986). *Aquiferes Fissures en Zone Tropicale Humide : Structure, Hydrodynamique et Hydrochimie. (Afrique de I'Ouest)*. These Doc. D'Etat. U.S.T.L. Montpellier.
46. Faillat, J.P. (1990). Origin desnitrates dans les nappes de fissures de la zone tropicale humide. Exemple de la Cote d'Ivoire. *J. Hydrol.* **113**: 231–264.
47. Faillat, J.P. and Rambaud, A. (1988). La teneur en nitrates des nappes de fissures de la zone tropicale humide en relation avec les problems de deforestation. *C.R. Acad. Sci. Paris* **306**(II): 1115–1120.
48. Edmunds, W.M. and Gaye, C.B. (1997). High nitrate baseline concentrations in groundwaters from the Sahel. *J. Environ. Qual.* **26**: 1231–1239.
49. Okonkwo, P.O., Nwokolo, C., and Onwuamaeze, I.C. (1981). Dietary nitrite and carcinogenesis in the Nigeria environment. *J. Nutr. Sci.* **2**: 9–13.
50. Mirvish, S.S., Wallcare, I., Eagen, M., and Shubik, P. (1972). Ascorbate-nitrite. Reactions; possible means of blocking the transformation of carcinogenic N-nitroso compounds. *Science* **177**: 65–68.
51. Ezeonu, F.C. (1989). The Role of Rat Liver Microsomal Enzyme in the *in vitro* Metabolism of Nitrate and Formation of Nitrosamine. Paper presented at the *14th Annual Biochem. Soc. Nigeria Conf.*, Imo State University, Okigwe, Nigeria.

52. Tredoux, G. (1993). *Preliminary Investigation of Nitrate Content of Groundwater and Limitation of the Nitrate Input*. WRC Report No. 368/1/93, Pretoria.
53. Packham, R.F. (1991). Public health and regulatory aspects of inorganic nitrogen compounds in drinking water. *Int. Workshop, Inorg. Nitrogen Comp. Water Supply*, Hamburg, November 1991.
54. Bosch, H.M. et al. (1950). Methemoglobinemia and Minnesota well supplies. *J. Am. Water Works Assoc.* **42**: 91–95.
55. Walton, G. (1951). Suvey of literature relating to infant methemoglobinemia due to nitrate contaminated water. *Am. J. Public Health* **41**: 1141–1152.
56. Super, M. et al. (1981). An epidemiological study of well water nitrates in a group of Namibian infants. *Water Res.* **15**: 1265–1270.
57. Hesseling, P.B., Toens, P.D., and Visser, H. (1991). An epidemiological survey to assess the effects of well water nitrates on infants' health at Rietfontein in the northern Cape Province, South Africa. *South Afr. J. Sci.* **87**: 300–304.
58. United States Environmental Protection Agency. (1990). *National Survey of Pesticides in Drinking Water Wells—Phase I Report*, EPA 570/09-90-015, Office of Water, Washington, DC.
59. United States Environmental Protection Agency. (1992). *Another Look: National Survey of Pesticides in Drinking Water Wells—Phase II Report*, EPA 579/09-91-020, January, Office of Water, Washington, DC.
60. Cohen, S., (1992). Results of the National Drinking Water Survey: Pesticides and well characterized. *Water Well J.* **46**(8): 35–38.
61. Keeney, D. (1989). Sources of nitrate to ground water. In: *Nitrogen Management and Groundwater Protection*. R.F. Follett (Eds.). Elsevier Science, Amsterdam, The Netherlands, Chap. 2, pp. 23–34.
62. United States Environmental Protection Agency. (1994). *Nitrogen Control*. Technomic, Lancaster, PA, pp. 1–22.
63. Van Duyvenboden, W.V. and Loch, J.P.G. (1983). Nitrate in the Netherlands—A serious threat to groundwater. *Aqua* **2**: 59–60.
64. Van Beek, C.G.E.M. (1985) Nitrate situation. Paper presented at *Conf. Nitrates Water*, Paris.
65. Lawrence, C.R. (1983). Occurrence and genesis of nitrate-rich groundwaters of Australia. *Int. Conf. Groundwater Man, Sydney* **2**: 237–247.
66. United States Environmental Protection Agency. (1993). *Nitrogen Control*. EPA 625/R-93-010, September 1993, Office of Water, Washington, DC.
67. Bassir, O. and Maduagwu, E.N. (1978). Occurrence of nitrate, nitrite, dimethylamine and dimethylnitrosamine in some fermented Nigerian beverages. *J. Agric. Food Chem.* **26**(1): 200–203.
68. United States Agency for International Development. (1990). *Strategies for Linking Water and Sanitation Programs to Child Survival*. USAID, Washington, DC, pp. 1–62.
69. Warner, D. (1998). Drinking water supply and environmental sanitation for health. Presented at the *Int. Conf. Water Sustainable Dev.*, Paris, March 19–21, pp. 1–10.
70. Population Reports. (1998). *Solutions for a Water-short World*, Series M, **14**, September.
71. Parsons, M.L (1977). Current research suggests the nitrate standard in drinking water is too low. *J. Environ. Health* **40**: 140–142.
72. Comly, H.H. (1945). Cyanosis in infants caused by nitrates in well water. *J. Am. Med. Assoc.* **129**: 112–116.
73. Pande, S.P., Hassan, M.Z., and Saxena, K.L. (1986). Nitrates and nitrites in the environment. *J.I.W.W.A.* **XVIII**(3): 145.
74. Terblanche, A.P.S. (1991). Health hazards of nitrate in drinking water. *Water SA* **17**(1): 77–82.
75. Federal Register. (1985). *National Primary Drinking Water Standards*. **219**: November 13, 46,880–47,022.
76. Shuval, H.I. and Gruener, N. (1972). Epidomiological and toxicological aspects of nitrates and nitrites in the environment. *Am. J. Public Health* **62**: 1045–1052.
77. Groundwater Newsletter. (1986). Blue-baby death from contaminated well in South Dakota stirs investigation. *Groundwater Newsl.* **15**(18): 1.
78. Csanady, M. (1991). Nitrite formation and bacteriological deterioration of water quality in distribution networks. Paper presented at *Int. Workshop: Inorg. Nitrogen Comp. Water Supply*, Hamburg, IWSA, Pre-Prints, 45–49.
79. Shuval, H.I. and Gruener, N. (1977). Infant methaemoglobinemia and other health effects of nitrates in drinking water. *Prog. Water Technol.* **8**: 183–193.
80. Ross, J.D. and Desforges, J.F. (1959). Reduction of methemoglobin by erythrocytes from cord blood. *Pediatrics* **23**: 718–721.
81. Colvin, C. (1999). Increased risk of methemoglobinemia as a result of bottle-feeding by HIV positive mothers in South Africa. Paper presented at *IAH Congr. 1999*, Bratislava, Slovakia.
82. McIntyre, J. (1997). *Affordable option for the prevention of mother to child transmission of HIV 1 from research to clinical care. National AIDs Programme Perinatal HIV Res. Unit Conf.*, November 1997, Johannesburg, South Africa.
83. Carpenter, R.A. and Maragos, J.E. (1989). How to assess environmental impact on tropical islands and coastal areas. In: *Training Manual for South Pacific Regional Environment Programme*, South Pacific Regional Environment Programme (SPREP), Apia, pp. 119–122.
84. Genthe, B. (1998). *Specialist Study on the Potential Health Impacts of the Proposed Wastewater Treatment Facility on the West Bank of East London*. CSIR report ENV/S—198003.
85. Tredoux, G. (1974). *'n Intensiewe opname van waterbronne in die omgewing van Otjiwarongo*, Project Report No. 12 (unpublished), NIWR-CSIR, Windhoek, Namibia.
86. Tannenbaum, S.R. and Green L.C. (1985). *Selected Abstracts on the Role of Dietary Nitrate and Nitrite in Human Carcinogenesis*. International Cancer Research Data Bank Program, National Cancer Institute, Washington, DC.
87. Wolff, I.A. and Wasserman, A.E. (1972). Nitrates, nitrites and nitrosamines. *Science* **177**: 15–19.
88. Hartman, P.E. (1983). Review: Putative mutagens and carcinogens in foods. *Environ. Mutagenesis* **5**: 111–121.
89. Kleinjans, J.C.S. et al. (1991). Nitrate contamination of drinking water evaluation of genotoxic risk in human populations. *Environ. Health Perspect.* **94**: 189–193.
90. Chandler, J. (1989). Nitrates in water. *Water Well J.* **3**(5): 45–47.
91. Anonymous. (1974). Nitrate vergif water in noorde: Beeste en skape vrek. *Die Suidwester*, Windhoek, Namibia, June 26.
92. Marais, S. (1991). *Written communication*. Department of Water Affairs, Mmabatho, Bophuthatswana.
93. Johnson, G.T. and Goldfinger, S.E. (1982). *The Harvard Medical School Health Letter Book*. Warner Books, New York.
94. Talma, A.S. (1991). *Written Communication*, CSIR, Pretoria.
95. Life Systems, Inc. (1987). *Drinking Water Criteria Document for Nitrate/Nitrite*, TR-832-77. Cleveland, Ohio.

96. Dahab, M.F. and Bogardi, I. (1990). *Risk Management for Nitrate-Contaminated Groundwater Supplies*. University of Nebraska, Lincoln, NE, pp. 9–16.
97. Kross, B.C., Ayebo, A.D., and Fourtes, L.J. (1992). Methemoglobinemia: nitrate toxicity in rural America. *Am. Family Physician*, July.
98. *Secretaria de Salud. (1988). Reglamento de la Ley General de Salud en Materia de Control Sanitario de Actividades, Establesimientos, Productos y Servicios* (General law for sanitary control of activities, products and services): Official News of the Mexican Government, January 8, Mexico, p. 27.

DOMESTIC WATER SUPPLY—PUBLIC–PRIVATE PARTNERSHIP

Muhammad Sohail
Loughborough University
Leicestershire, United Kingdom

INTRODUCTION

> "Food and water are basic rights. But we pay for food. Why should we not pay for water?"
>
> Ismail Serageldin

> "Water should not be privatised, commodified, traded or exported in bulk for commercial purposes."
>
> Maude Barlow

The above two quotes typify the two extremes of the arguments surrounding public–private partnerships (PPPs). Although this ideological debate is intellectually exciting, the more challenging problem can be summarized as follows (1):

- How do you provide access to safe water to around I billion people?
- How do you provide sanitation services to around 2.6 billion people?

Time is ticking by, and the global community must meet the above targets soon. Policymakers and water and sanitation practitioners alike should be prepared to include PPPs as a tool in achieving the above. Before going into detail, a brief summary of the context follows.

Water is essential for life. This is not only in terms of its biological utility, for water also has social, economic, health, technical, financial, and political dimensions. Historically, the availability of a domestic water supply has been a significant factor in the development and sustenance of civilizations.

Water is one of Earth's most important natural resources. If naturally occurring freshwater is polluted as a result animal (including human), plant, or other activities, processes are required to convert the then raw water to a quality fit for a particular use, such as drinking. In most cases, because of high levels of pollution by humans, water must be treated before and after its use.

Various water sources exist, for example, rainwater, groundwater, spring water, surface water, rivers, lakes, ponds, fog, and even glaciers. Water can be supplied from such sources in many different ways, but these can then be divided into either piped or nonpiped options.

The behavior of humans in terms of their consumption of water also has historical, geographical, and cultural dimensions. For example, in some parts of Bolivia people only consume 5 or 6 L per capita each day as compared with the 30 to 250 L consumed each day by a person in developed economies.

Water has various uses, including agricultural, recreational, industrial, and domestic. In terms of its domestic use, water is used for sanitation facilities as well as for drinking. Indeed, a high level of consumption—up to 80%—is caused by piped sanitation wherever such facilities are available. Think about how much water a person flushes as compared with how much he or she drinks! With a limited amount of useable water, there is competition, sometimes tension, among various water users. Both market-oriented and hierarchy-based rules are used to distribute water among its various consumers.

In the supply chain of water—its production, distribution, management, and consumption—private sector organisations are key players. Before moving on to the main discussion of this paper, it is useful to summarize some of the more important concepts.

CONCEPTS

Public and Private Goods

Private goods are those for which consumption (or use) by one person prevents consumption (or use) by another. Public goods are those that can be used by one person without diminishing the opportunity for use by others. There is a seemingly unending debate over whether water should be treated as a public or private good, or both.

Commodification

Commodification is the process of converting a good or service formerly subject to non-market social rules into one that is subject to market rules. Treating water as an economic good implies that the resource will be allocated across competing uses in a way that maximizes its economic value across society. However, it also implies that safety nets will be needed for people who cannot afford to pay. Meanwhile, for some people, considering water to be a commodity at all is sacrilegious!

Governance

Governance refers to the relationship that can be manifested in various types of partnerships and networks. Water governance is the wider context within which water services procurement plays a key role and under which PPP is a niche tool.

What is Partnering?

At the very least, partnering should be viewed as the absence of adversarial behavior.

WHAT ARE PPPS?

The involvement of the private sector in partnership with government has long been advocated as a means of improving the development of sustainable water and sanitation systems. The author uses PPPs in this document as a general term to cover a wide range of agreements or partnerships between private sector (nongovernment) concerns or organizations, public sector utilities, government departments, and consumer groups in relation to the delivery of water and sanitation services. The community has a direct role to play in such arrangements as a beneficiary and in expressing the price people would be willing to pay for an acceptable level of service. It also has an indirect role to play in shaping policy for the urban environment. In a small PPP, the community could take the role of the private partner.

One of the difficulties in determining the scope of discussion concerning private sector involvement in water and sanitation is the sheer diversity of possible partnership arrangements and potential actors. The three main roles are those of the private sector, users (consumers), and the government (often referred to a the client or sponsor; it also may act as regulator). The possible arrangements include complex concession arrangements operated by multinational corporations lasting perhaps 30 years; shorter duration, simpler forms of management or service contract undertaken by medium-sized private enterprises; and service delivery by small-scale independent providers (local entrepreneurs). An outline of formal contracting arrangements is produced in Appendix A.

WHY PPPS FOR DOMESTIC WATER SUPPLY?

The role of the private sector in domestic water supply is not new. Many water supply-related activities in many municipalities in the world have started out as small private water firms or informal organizations of people.

Considering the potential demand for water sanitation, it is obvious that already budget-constrained governments cannot improve services alone. Some estimates indicate huge capital outlays are needed to meet water and sanitation targets; Camdessus and Winpenny (2), for example, estimate US$90 billion per year is needed globally to meet such targets. On the other hand, there is an argument that unless national governments make environmental sustainability a priority, the shortage of water and sanitation services will remain as they are. Other people believe that improved water management and better use of current assets is what is crucial and will reduce the need for more capital infrastructure significantly.

In any case, private finance is only one of the benefits of PPPs. Others include the managerial capacity of the private sector in managing assets for water and sanitation services. The private sector is also more likely to be innovative than the public sector, which leads to more effective and efficient services provision.

Governments are turning to PPP arrangements for the provision of services for a variety of reasons; therefore, these may include the following:

- Lack of capacity of government institutions to deliver a reasonable level of service or to improve service quality.
- Financial weakness of public utilities.
- The inability of public institutions to respond to an increasing growth in demand for services because of, for example, rapid urbanization.
- Problems related to the large numbers of employees in public sector providers.
- The requirement that international financing institutions (IFIs) place on certain indebted countries to reduce domestic spending as part of structural adjustment programmes. Such institutions promote the idea that the private sector is more efficient and effective and delivers a better quality service.

HOW PUBLIC-PRIVATE PARTNERSHIPS GET STARTED

There are various ways in which a partnership can be initiated. In some cases, the initiative comes from the likely partners and in some cases the demand comes from an third party. Public-private partnerships may not always seem to be a desirable solution at first. Most organizations prefer to stay on paths they know well, sharing goals and work practices with other groups that think and act like them—governments working with governments, businesses with businesses, and nonprofit groups with nonprofit groups.

Governments and private firms have long worked together under simple arrangements, such as government purchase of products produced by the private sector. However, both parties often hesitate to enter into more complex relationships. Governments are frequently concerned that private businesses will take advantage of them, whereas businesses often consider government approaches to be burdensome and a waste of time. Therefore, it is useful to allow some time for trust to be established among the key partners.

Three main conditions favor the formation of a partnership: urgency, the involvement of a champion, and some kind of catalyst.

Urgency

Generally, it takes a widely acknowledged urgency—for instance, the lack of particular services or the waste of resources—before key stakeholders start looking for partnerships and partners are open to cooperate to resolve the problem. Although it is hoped that progress can be made in the absence of a crisis, in practice, the inertia that keeps many people on familiar paths is usually broken only by a pressing need for change.

Entrepreneurs/Champion

Sometimes, even in the absence of a urgency, an individual, group, or organization may realize that separate, uncoordinated actions are creating redundancies, missed opportunities, and less-than-optimal use of scarce resources. In reality, even if the crisis is clear and the interest is there, partnership arrangements will not succeed without the drive and commitment of a few individuals. Such “champions” (leaders or pioneers) can be government officials, NGOs, business people, or citizens who—through their personal motivation—make partnerships happen. In other cases, champions are service providers who stand to profit from the partnership.

Catalyst

Frequently, there is a need for some kind of catalyst to bring the partners together. Such a catalyst could be the actions of an external actor, such as one or more international finance institutions or other bodies that are respected and trusted by all partners.

WHAT ARE THE KEY FEATURES OF SUCCESSFUL PARTNERSHIPS?[1]

Characteristics of Successful Partnerships

Compatible Goals. Government , businesses, and community leaders must understand and respect one another’s goals. For instance:

- Government may initially have difficulty accepting the profit motive of private businesses.
- Private companies may be tempted to walk away from the more bureaucratic decision-making processes used in the public sector.
- Local communities may not have the patience needed to address issues affecting other areas of the city.

To resolve these differences, all parties must focus on the broader, complementary goals that are to be achieved. It is important for them to realize that public and private goals do not necessarily need to be the same for partnerships to work—they must be merely compatible.

Enabling Environments. An enabling regulatory, legal, and political environment is the cornerstone of sustainable private sector participation.

Legal Framework. Early on, the public sector must establish an appropriate legal framework for contract procurement and private sector investment. It is important that mechanisms be put into place to minimize the likelihood or appearance of corruption in any procurement processes. Unpredictable and unfair procurement processes reduce both political acceptability and the interest of many private investors.

Regulatory Framework. The government must also establish a clear regulatory framework, and it must implement appropriate tariff regimes and subsidy mechanisms. The creation of a regulatory framework alone, however, does not necessarily guarantee effective regulation. As all local governments are different, the public and private sectors will face a steep learning curve as they try to define and regulate their relationship with one another and their roles in providing services. In particular, the public sector needs to define a clear allocation of responsibilities between the national and municipal governments and a clear statement of its role as a provider and a regulator.

In general, private sector companies prefer that the contract serve as the major regulatory mechanism, and that governments have limited regulatory discretion once the contract is in place. Highly specific contract terms that establish duties, performance targets, rules for changing prices, and dispute resolution procedures allow the private sector to better predict the profitability of the venture and decide whether and what to bid for the contract. Given these preferences, governments will have to make important decisions about the degree of regulatory discretion they are willing to give up, particularly for long-term contracts.

Political Environment. In addition to the regulatory climate, a bad political climate caused by the pressure of election cycles, the potential instability of new democracies, the personal agendas of government officials, and the special status of some services (particularly in terms of access to water, for example) can create barriers to starting or maintaining public–private collaborations. Governments must provide assurances whenever possible to private sector partners that such political factors will not disrupt the contractual partnership.

Acceptance. The government and business leaders cannot build partnerships alone; political and social acceptance of private sector involvement is essential. The population must see private sector participation as beneficial if the partnership is to last over time. Public support of private involvement over the long term will depend on primarily the delivery of promised services and benefits at reasonable costs. Therefore, it is of the utmost importance that mechanisms be developed to ensure that the organization providing the service, whether it is a public or a private sector organization, be accountable to its customers.

Public support will also depend on the ability of the partnership to meet the needs of all stakeholders. For example, public sector workers can be a source of tremendous opposition to increased private involvement in the provision of services. Contracts should ensure the employment or placement of public employees and local residents to the greatest degree possible.

Credibility and Transparency. Effective cooperation between local government, businesses, and the community is always difficult to achieve because of the wide range of participants involved, the low level of trust that often exists between potential partners, and the lack of predictability in the process. The credibility of champions and other leaders involved, as well as transparency in the process, are critical determinants of long-term success. Experience suggests that genuine partnerships must include the principles of equity, transparency of operations, and mutual benefit. Trust and confidence in any project is necessary for successful partnerships.

---

[1] The following section is based on Sohail and Olena (3).

Factors Contributing to the Durable Partnerships

Governments clearly want to establish PPPs that are sustainable over time. Essential ingredients of durable PPPs include those listed below.

Commitment of Resources. All partners to the arrangement should be obliged to commit resources (financial, human, capital) to increase their interest in seeing the partnership succeed, which implies shared risks and rewards.

Capacity Development. Projects requiring substantial institutional change or large capital investments will require capacity development within all groups of stakeholders. For example, development of:

- Consumers, in terms of their knowledge of the service they are to receive and the costs associated with its provision
- Service providers, particularly local organizations, in terms of entrepreneurial skills
- Governments, in terms of their capacity to adopt the frameworks necessary for, and oversee the provision of, the service

Roles and Responsibilities. The delineation of appropriate roles and responsibilities is another element necessary in the development of effective, durable partnerships. It is essential that partnerships be organized in a concerted fashion to make the most effective use of the resources committed by both parties.

Individual responsibilities should be clearly outlined from the beginning so that there is no ambiguity in the tasks that each party is expected to perform. Furthermore, these responsibilities need to be defined realistically with a clear understanding of the strengths and weaknesses of each partner.

Flexibility. All partnerships are context-based and different locally. Partnerships should draw on other experiences but at the same time should be opportunistic about exploiting the comparative advantage of local resources. Over the long term, changes in investment plans, technology choices, and priority actions will be necessary in response to unforeseen circumstances. Including clear procedures for making such changes over the life of the project will reduce the chance that they will have a negative impact on the partnership.

Time. Partnerships take time. The process of understanding the problems to be addressed and the impacts on potential partners, as well as those partners' needs and aspirations, all takes time. Progress can certainly be made along the way, but the process of achieving and maintaining acceptance among users, providers, and regulators is a continuing one; a cooperative dialogue to address shared needs must be maintained throughout the project.

Patience. Projects requiring substantial institutional change or large capital investments require a lot of patience. Careful attention must be paid to the balance between responding rapidly to the most pressing crises and developing integrated solutions that will last. Political cycles and the desire for immediate improvement in a crisis situation often lead to the development of time frames that are too short. Such short-term agendas and limited horizons lead to unrealistic expectations and unsustainable solutions. It is not realistic to expect that private sector involvement will overcome public institutional and operational inefficiencies quickly, nor that it will compensate immediately for a history of insufficient public sector resources and funding.

Social Responsibility. Public services provide public goods—in other words, goods that should be available to everyone. Improving provision of such services is about making people's lives better, especially those of the urban poor. Governments should always make sure that the changes they make promote increased access to, and better quality of, services. An emphasis on social responsibility will also increase political gain, as better services will lead to greater political acceptance by the general population.

What are the Major Obstacles to Forming a Successful Partnership?

A range of possible obstacles or deficiencies in the capacity of both public and private actors could hinder the formation of a successful partnership. Major obstacles in this respect include:

- Reciprocal mistrust and lack of understanding of one another's interests and needs across the public and private sectors
- Absence of locally available information on, and experience with, arranging sustainable partnerships
- Underlying legal, political, and institutional obstacles to forming effective public–private relationships

These obstacles often lead to lengthy negotiations, increased transaction costs, and make smaller projects much less attractive to potential investors. To minimize the harm from such obstacles, PPP arrangements should provide certain safeguards for the public and private sectors and for the community.

The public sector usually expects the private sector to contribute in one or all of the following ways:

- To provide agreed services
- To make agreed investments
- To meet agreed standards/targets
- To not exploit any monopoly situation that might exist

The private sector expects the public sector to contribute in one or all of the following ways:

- To create an enabling environment suitable for the PPP
- To pay agreed fees promptly and in full
- To implement tariff increases as agreed

- To prevent unexpected competition from others during operation (exclusivity)

The community expects the PPP to:

- Provide appropriate levels of services
- Be affordable to the community, either through direct charges or indirectly through general taxation

THE CURRENT SITUATION

Quantitative information in this section is based mainly on the World Bank Private Participation in Infrastructure (PPI) database (4).

- At least 203 water and sanitation projects are taking place in 43 developing countries, with a corresponding commitment of some US $40 billion.
- A few large projects in, for example, Argentina, Chile, and Manila, Philippines can explain peaks in water and sanitation investment during 1991–2000.
- Latin America and the Caribbean are the most active regions in terms of global investment (52%), followed by East Asia and the Pacific (38%) and only then Europe and Central Asia (8%).
- Concession contracts provide the largest proportion of such investment (69%) through the largest number of projects (90).
- Private investment reached a peak in 1999.
- The top five countries by cumulative investment in water and sewerage projects with private participation were Argentina, the Philippines, Malaysia, Chile, and Brazil during 1991–2000.

The above figures relate only to formal PPPs and do not take into account many informal private sector and community-based operations; such operations could be serving many users well.

CHALLENGES[2]

Political Acceptance

Infrastructure services such as water and sanitation are intrinsically political and that fact, along with the politicians' individual agendas, should be acknowledged in policy discussions. For example, PPPs are still seen to be a political risk by many politicians, whose principle objectives include acquiring or retaining power. Quick fixes may be favored to attain popularity rather than making hard decisions for the long-term sustainability of water and sanitation services. Nonetheless, socially sensitive PPPs are more likely to be politically acceptable.

Social Acceptance in Developing Countries

Policy Issue: The Poor are Seldom Mentioned. In many developing countries, private sector involvement in services provision is expected to also fulfill the social role of the public sector. However, where this is the case, it must be negotiated as part of the PPP right from the beginning of the procurement process.

Knowledge and Understanding of Advisers. Terms of Reference mostly appear to be drafted by advisers, who tend to be distant from the sponsor/client's organization. Their expertise will be determined by the requirements specified by the sponsor/client and normally focus on technology and finance. A lack of understanding on the part of those responsible for developing and negotiating the contracts can lead to the omission of important existing informal arrangements among public utility staff, water vendors, and low-income customers. Hence the inclusion of expert staff members that have PPP experience and a workable knowledge of how to improve water and sanitation facilities for low-income (poor) areas and communities is important if PPPs are to become socially sensitive.

The time scale for the sponsor/client's advisers to prepare the necessary bidding information is relatively short, and this inevitably results in workloads being prioritized. Consequently, the complex and little understood issues involved in service provision for the poor are unlikely to receive serious consideration at this stage. The "time factor" may therefore have contributed to the historic underrepresentation of services to low-income communities in many PPP contract documents.

Governance and Relationships with Consumers. Public sector utilities and private companies are experienced in service provision to regular settlements and housing developments occupied by middle and high-income consumers with individual service connections. However, public sector utilities rarely have much experience in dealing with service delivery to the poor, whereas operators from the formal private sector appear to have even less. A key feature of extending coverage to the urban poor therefore involves working out new relationships between a diverse range of actors who have little if any experience of one another.

Direct Links: Partnership does not have to Involve Formal Private Sector. In cases where small local entrepreneurs or civil society groups take on the role of private sector operator, they generally act as an intermediary between the public sector sponsor/client and the consumer. This may, for instance, involve acting as a retailer of water services. In such circumstances, there tends to be more negotiation to address problems and issues at the local level, with local entrepreneurs communicating directly with consumers or their representative groups. Although there is less recourse to legally binding agreements, the key issue remains the same, namely, whether the mechanism works for poor consumers.

User Perceptions

In recent years, consumers have been asked to voice their opinions on the overall process of PPP development involving the formal private sector. Concerns that have

---

[2] This section is based on Sohail (5).

emerged include lack of consultation; concern over lack of public control or safeguards; fear of corruption; high tariffs; unemployment; and the assumption that there will be increased burdens on the "common people," who in turn have no clear idea of the benefits.

Involvement in the processes leading up to change is critical. Although this does not guarantee success, case studies have found that projects in which stakeholder participation was absent or minimal were the least successful in terms of feedback from people on low incomes. In such cases, consumers tended to raise objections to the involvement of private companies in their water supply, especially foreign companies. Commonly, such involvement is precipitated by increases in tariffs, which are perceived to be making profits for private companies at the consumer's expense.

Financial Issues and Tariffs

Payment Problems. Some of the problems faced by households include high arrears, high repayment levels, disconnection of the water supply, and inability to pay reconnection charges. The introduction of higher charges through both metering and increased standing charges creates additional pressure on the household budget. Low-income people are clearly under most pressure to economize, and this results in reduced water consumption and less cash being available for other needs.

Technology

Scope for Innovation. Simple yet innovative technology changes may form part of the implementation of PPP arrangements. Such changes are the result of various factors, including:

- Analysis of settlements and their needs
- Taking community needs and preferences into account
- Knowledge of alternative options, relevant technologies, and proven models—for example, simplified, lower cost designs
- Willingness to experiment to find innovative alternatives to standard technological approaches

Levels of Service and Service Differentiation. There are several cases where PPP has led to the introduction of more appropriate and flexible levels of service. This has involved lowering the conventionally accepted, high levels of service that predominate in wealthier areas.

Information

The Need to Communicate Effectively. The introduction of PPPs normally changes existing roles and responsibilities, and these have to be communicated to consumers. If one of the partners (usually the operator) has no working knowledge of the area, then it is crucial for that partner to find out about consumer needs.

Regulation, Monitoring and Complaints

Formal Regulatory Systems. These should be developed during the preoperational phases of PPP contracts, and they are notably absent during the preparation for many PPPs. No provisions are made to check the operation of a monopoly supplier or to impose penalties for substandard performance. In addition, there may be no defined procedures for routine external monitoring of performance. In particular, lack of capacity for monitoring and reporting is a problem for the relatively small municipalities involved in many PPPs.

Effectiveness of Regulation. International PPP experience has shown there to be situations where regulation does not necessarily safeguard the different interests of the various parties to the arrangement. For instance, poor consumers are often dissatisfied with services and tariffs in formally regulated environments; billing and bill collection by the public partner may be inefficient; and the private operator may perceive lack of control over customer management to be a significant risk.

CREATING TRUST

Communication and information alone are not enough. In many PPPs, consumers are unsure of what to expect from privatization because they have not been involved in the development process. If user groups and other stakeholders had been integrated into the process from the start, then there might not be such widespread opposition.

In conclusion, then:

- There is a pressing need to develop a base of information about low-income groups; this can then be used directly in the development of PPP arrangements. Such action requires a clearer understanding of information needs on the part of both the designers of PPPs and the local institutions and organizations.
- There is also a clear need to provide information for consumers, particularly about proposed roles and responsibilities. Lack of understanding and consensus leads to operational problems and is ultimately disempowering.
- Communication is a vital component of PPPs, and investment in this will pay dividends in operational terms. Lack of information could result in low-income communities refusing to accept or comply with the partnership, and this in turn could lead to the risk of nonpayment.

BIBLIOGRAPHY

1. World Health Organization and UNICEF (2004). *Meeting the MDG Drinking Water and Sanitation Target: A Mid-Term Assessment of Progress*. WHO, Geneva.
2. Camdessus, M. and Winpenny, J. (2003). Report of the World Panel on Financing Water Infrastructure: Financing Water For All. World Water Council, 3rd World Water Forum, Global Water Partnership. Ahmed, N. and Sohail, M. (2003). Alternate water supply arrangements in peri-urban localities: Awami (People's) tanks in Organgi Township, Karachi. *Environ. Urbanisation* **15**(2): 33–42.
3. Sohail and Olena (2003).

4. World Bank (2003). *Private Participation in Infrastructure: Trends in Developing Countries in 1990–2001*. World Bank, Washington, DC.
5. Sohail, M. (2003). *Making Partnering Work for All—A Guideline*. Margraf Publishers, Frankfurt, Germany.
6. Blokland, M., Braadbaart, O., and Schwartz, K. (Eds.). (1999). *Private Business, Public Owners: Government Shareholdings in Water Enterprises*. Netherlands Ministry of Housing, Spatial Planning and the Environment and WSCC, the Netherlands.
7. Calaguas (1999).
8. Kempe, A.H. and Schreiber, P. (1997). Rural water services benefit from PPP. *Water Supply* **15**(3): 99–111.
9. Johnstone, N. and Wood, L. (1999). *The Regulation of Private Sector Participation in Urban Water and Sanitation: Realising Social and Environmental Objectives in Developing Countries*. IIED, London.
10. Lewis, M.A. and Miller, T.R. (1987). Public-private partnerships in water supply and sanitation in sub-Saharan Africa. *Health Policy and Planning* **2**(1): 70–79.
11. Nickson, A. (1998). The public-private mix in urban water supply. *Int. Rev. Admin. Sci.* **63**(2): 165–186.
12. Rameama (1997).

READING LIST

Ahmed, N. and Sohail, M. (2000). *Privatisation of KWSB—An Analysis*. City Press, Karachi, Pakistan.

Barlow, M. and Clarke, T. (2002). *Blue Gold: The Fight to Stop the Corporate Theft of the World's Water*. The New Press, Washington, DC.

Becker, M.L. (1993). The international joint commission and public participation—past experiences, present challenges, future tasks. *Natural Resour. J.* **33**(2): 235–274.

Bennett, A. (1998). Sustainable public/private partnerships for public service delivery. *Natural Resour. Forum* **22**(3): 193–199.

Bhattacharyya, A., Parker, E., and Raffiee, K. (1994). An examination of the effect of ownership on the relative efficiency of public and private water utilities. *Land Economics* **70**(2): 197–209.

Bhattacharyya, A., Harris, T.R., Narayanan, R., and Raffiee, K. (1995). Specification and estimation of the effect of ownership on the economic efficiency of the water utilities. *Regional Sci. Urban Economics* **25**(6): 759–784.

Briscoe, J. and Gann, H.A. (1995). Financing water supply and sanitation under Agenda 21. *Natural Resour. Forum* **19**(1): 59–70.

Brook Cowen, P.J. (1999). Lessons from the Guinea Water Lease. World Bank Public Policy for the Private Sector Note 78.

Brook Cowen, P.J. (1997). Getting the private sector involved in water: What to do in the poorest of countries? World Bank Public Policy for the Private Sector Note 102.

Brook Cowen, P.J. (1997). The private sector in water and sanitation—how to get started. World Bank Public Policy for the Private Sector Note 126.

Cain, L.P. (1997). Historical perspective on infrastructure and U.S. economic development. *Regional Sci. Urban Economics* **27**(2): 117–138.

Cashore, B. (2002). Legitimacy and the privatization of environmental governance: How non–state market–driven (NSMD) governance systems gain rule–making authority. *Governance* **15**(4): 503–529.

Cavill, S. and Sohail, M. (2003). Accountability in the provision of urban services. *Proceedings of the Institution of Civil Engineers* **156**(ME4): 235–244.

Cook, P. and Kirkpatrick, C. (1988). Privatisation in less developed countries: An overview. In: *Privatisation in Less Developed Countries*. P. Cook, and C. Kirkpatrick (Eds.). Harvester Wheatsheaf, London.

Correll, D. (1996). Public-private water partnerships benefit all of New Jersey. *Water, Eng. Manage.* **143**(9): 31–32.

Coyaud, D. (1988). Private and public alternatives for providing water supply and sewerage services. INU Discussion Note 31.

Crampes, C. and Estache, A. (1996). Regulating water concessions: Lessons from the Buenos Aires Concession. World Bank Public Policy for the Private Sector Note 91.

Crane, R. (1994). Water markets, market reform and the urban poor: Results from Jakarta. *World Development* **22**(1): 71–83.

Crosslin, R.L. (1991). Decision-support methodology for planning and evaluating PPPs. *J. Urban Planning Develop.* **117**(1): 15–31.

Dellapenna, J. (1994). Why are true water markets rare or why should water be treated as public property? In: *Joint Management of Shared Aquifers: Second Workshop*, M. Haddad and E. Feitelson (Eds.). 27 November–1 December.

Días A. and Ioannou, P.G. (1995). Debt capacity and optimal capital structure for privately financed infrastructure projects. *J. Construction Eng. Manage.* **121**(4): 404–414.

Drakeford, M. (1998). Water regulation and pre-payment meters. *J. Law Soc.* **25**(4): 558–602.

Easter, K.W. and Hearne, R. (1995). Water markets and decentralised water resources management: International problems and opportunities. *Water Resour. Bull.* **31**(1): 9–20.

Farade, T. and Sohail, M. (2003). Making regulatory mechanisms work: Lessons from cases of private sector participation. *Water Policy* **5**(3): 269–288.

Fass, S.M. (1993). Water and poverty: Implications for water planning. *Water Resour. Res.* **29**(7): 1975–1981.

Feigenbaum, S. and Teeples, R. (1983). Public vs. private water delivery: A hedonic cost approach. *Rev. Econ. Stat.* **65**(4): 672–678.

Feldman, R.D. (1993). Power and infrastructure privatization in Mexico. *Strategic Planning for Energy Environ.* **13**(2): 14–16.

Gentry, B. and Abuyuan, A. (2000). Global Trends in Urban Water Supply and Waste Water Financing and Management: Changing Roles for the Public and Private Sectors. Working Paper CCNM/ENV(2000)36/FINAL, OECD, Paris.

Gleick, P., Wolff, G., Chalecki, E.L., and Reyes, R. (2002). *The New Economy of Water: The Risks and Benefits of Globalization and Privatization of Fresh Water*. Pacific Institute, Oakland.

Gray, P. (1997). Colombia's gradualist approach to private participation in infrastructure. World Bank Public Policy for the Private Sector Note 113.

Haarmeyer, D. and Mody, A. (1998). Pooling water projects to move beyond project finance. World Bank Public Policy for the Private Sector Note 152.

Hall, D. (2002). Water in Public Hands. Public Services International Research Unit (Working Paper), Greenwich, UK.

Hall, D., Bayliss, K., and Lobina, E. (2002). Water privatization in Africa. Paper by the Public Services International Research Unit, University of Greenwich for the Municipal Services Project Conference, Johannesburg, South Africa.

Hazin, L.S. (1998). New strategy in urban water management in Mexico: The case of Mexico's Federal District. *Natural Resour. Forum* **22**(3): 185–192.

Hulme, D. and Montgomery, R. (1994). Cooperatives, credit and the poor: Private interest, public choice and collective action in Sri Lanka. *Marga: Quart. J.* **13**(3): 35–55.

Humphrey, C.E. (1990). *Privatization in Bangladesh: Economic Transition in a Poor Country*. Westview Press, Boulder, CO.

Idelovitch, E. and Ringskog, K. (1995). *Private Sector Participation in Water Supply and Sanitation in Latin America*. Directions in Development, World Bank, Washington, DC.

Irwin, T. (1997). Price structures, cross-subsidies, and competition in infrastructure. World Bank Public Policy for the Private Sector Note 107.

Kay, J.A. and Thompson, D.J. (1986). Privatisation: A policy in search of a rationale. *Econ. J.* **96**(3).

Kerf, M. and Smith, W. (1996). Privatizing Africa's infrastructure: Promise and challenge. World Bank Technical Paper 337, Africa Region Series, World Bank, Washington, DC.

Kikeri, S., Nellis, J., and Shirley, M. (1992). *Privatization: The Lessons of Experience*. Country Economics Dept, World Bank, Washington, DC.

Klein, M. (1998). *Bidding for Concessions—The Impact of Contract Design*. World Bank Public Policy for the Private Sector Note no. 158. Available at http://www.ppiaf.org

Krumholz, N. (1999). Equitable approaches to local economic development. *Policy Studies J.* **27**(1): 83–95.

Labonne, B. (1998). Public-private partnerships in natural resources management. *Natural Resources Forum* **22**(2): 75–76.

Lamberton, D.M. (1996). Infrastructure: A nebulous and overworked concept? *Int. J. Technol. Manag.* **12**(5–6): 696–703.

Lane, M. (1999). Trends in international water project procurement. *Proc. British Water Overseas Contracts and Conditions Conference*.

Lee, T. and Jouralev, A. (1992). Self-financing water supply and sanitation services. *CEPAL Rev.* (Santiago de Chile). **48**: 117–128.

Lee, T.R. (1996). Alternatives for private participation in the provision of water services. *Natural Resources Forum* **20**(4): 333–341.

Liddle, B.T. (1997). Privatization decision and civil engineering projects. *J. Manag. Engineer.* **13**(3): 73–78.

Lorrain, D. (1993). Urban services, the market and politics. In: *Private Financing of Public Infrastrucure: The French Experience*. C. Martinand (Ed.). Economica, Paris.

Lyonnaise des Eaux (1990). *Draft Concession Contract for Management, Operation and Maintenance of a Water Supply System*. Southern Africa.

Lyonnaise des Eaux (Coordinator) (1999). *Alternative Solutions for Water Supply and Sanitation in Areas with Limited Financial Resources*, France.

Maldonado, M. (1997). Public water in private hands. *Civil Engineer [New York]* **67**(1): 49–51.

Marvin, S. and Slater, S. (1997). Urban infrastructure: The contemporary conflict between roads and utilities. *Progress in Planning* **48**(4): 247–307.

McCarthy, S. and Tiong, L.K.R. (1991). Financial and contractual aspects of BOT projects. *Int. J. Proj. Manag.* **9**(4): 222–227.

McCoy Thompson, S. (1998). Public-private partnerships: Spreading the three R's. *Natural Resources Forum* **22**(2): 87–94.

Metzger, J.T. (1998). Remaking the growth coalition: The Pittsburgh partnership for neighbourhood development. *Economic Development Quarterly* **12**(1): 12–29.

Mody, A. (1996). Infrastructure delivery: New ideas, big gains, no panaceas. In: *Infrastructure Delivery: Private Initiative and the Public Good*, A. Mody (Ed.). World Bank Economic Development Institute, Washington, DC.

Montgomery, J.D. (1988). The informal service sector as an administrative resource. In: *Urban Services in Developing Countries: Public and Private Roles in Urban Development*, D.A. Rondinelli and G.S. Cheema (Eds.). MacMillan, London.

Montiel, R.P. and Barten, F. (1999). Urban governance and health development in Leon, Nicaragua. *Environ. Urbanizat.* **11**(1): 11–26.

Morris, N.G. and Parry-Jones, S.A. (1999). The affordability of water in an African town. *J. Charter. Inst. Water Environ. Manag.* **13**(1): 1–6.

Nath, S. (1994). Privatisation of urban public services: Cost recovery as a major constraint in India. *Public Enterprise* **14**(3–4): 432–445.

Neto, F. (1998). Water privatization and regulation in England and France: A tale of two models. *Natural Resources Forum* **22**(2): 107–117.

Nicholson, N. (1998). How Hardwired is Human Behavior? *Harvard Bus. Rev.* **76**(4): 135–147.

Padco, Inc. (2002). A Review of Reports by Private-Sector-Participation Skeptics. Discussion Draft prepared for the Municipal Infrastructure Investment Unit, South Africa and USAID.

Palmer, A. (1996). Linking external and internal relationship building in networks of public and private sector organizations: A case study. *Int. J. Public Sector Manag.* **9**(3): 51–60.

Ranasinghe, M. (1998). Thoughts on a method to analyse viability of private sector participation in new infrastructure projects in developing countries. *Impact Assessment and Project Appraisal* **16**(3): 203–213.

Rees, J.A. (1998). Regulation and private participation in the water and sanitation sector. *Natural Resources Forum* **22**(2): 95–105.

Reijniers, J.J.A.M. (1994). Organization of PPP projects: The timely prevention of pitfalls. *Int. J. Proj. Manag.* **12**(3): 137–142.

Rivera, D. (1996). *Private Sector Participation in the Water Supply and Wastewater Sector: Lessons from Six Developing Countries*. World Bank Directions in Development, Washington, DC.

Roger, N. (1999). Recent trends in private participation in infrastructure. World Bank Public Policy for the Private Sector Note no. 196. Available at http://www.ppiaf.org

Rogerson, C.M. (1996). Willingness to pay for water: The international debates. *Water SA* **22**(4): 73–380.

Rogozinski, J. (1999). *High Price for Change: Privatization in Mexico*. Inter American Development Bank.

Rondinelli, D.A. (1988). Increasing the access of the poor to urban services: Problems, policy alternatives and organisational choices. In: *Urban Services in Developing Countries: Public and Private Roles in Urban Development*. D.A. Rondinelli and G.S. Cheema (Eds.). MacMillan, London.

Rosell, E. (1994). The chickens can come home to roost: The anatomy of a local infrastructure crisis. *Urban Affairs Rev.* **30**(2): 298–306.

Roth, G. (1987). *The Private Provision of Public Services in Developing Countries*. Oxford University Press for the World Bank, New York.

Saghir, J., Sherwood, E., and Macoun, A. (1998). Management contracts in water and sanitation—Gaza's experience. World Bank Public Policy for the Private Sector Note no. 177. Available at http://www.ppiaf.org.

Schusterman, R. and Hardoy, A. (1997). Reconstructing social capital in a poor urban settlement: The integral improvement programme in Barrio San Jorge. *Environ. Urbaniz.* **9**(1).

Serageldin, I., Barrett, R., and Martin, J. (Eds.). (1994). The business of sustainable cities: Public-private partnerships for creative technical and institutional solutions: An associated event of the second annual World Bank Conference on Environmentally Sustainable Development. *World Bank Conference on Environmentally Sustainable Development*. World Bank, Washington, DC.

Sohail, M. and Baldwin, A.N. (2001). Partnering with the community—an option for infrastructure procurement. *Proc. Institution of Civil Engineers Municipal Engineer.* **1454**: 293–297.

Sohail, M. and Baldwin, A.N. (2003). Urban infrastructure procurement in low-income countries. *ICE Proc. Institute of Civil Engineers, Municipal Engineer, Engineering Sustainability* **156**(ES2): 0–1.

Sohail, M. and Baldwin, A.N. (2004). Performance indicators for 'micro-projects' in developing countries. *Constr. Manag. Econom.* **22**: 11–23.

Sohail, M. and Cavill, S. (2000). *Public Private Partnerships and Poor, Part B—Interim Findings and Case Studies*. WEDC, Loughborough.

Sohail, M. and Cotton, A.P. (2000). *Public Private Partnerships and Poor Interim Findings, Part A—Summary and Lessons Learned*. WEDC, Loughborough.

Sohail, M. and Aslyukivska, M. (Eds.). (2004). Tools for Propoor municipal PPP, United Nations Development programme Margraf Publishers, Frankfurt, Germany.

Solo, T.M. (1999). Small-scale entrepreneurs in the urban water and sanitation market. *Environ. Urbaniz.* **11**(1): 117–131.

Solo, T.M. (2003). *Independent Water Entrepreneurs in Latin America: The Other Private Sector in Water Services*. World Bank, Washington, DC.

United Nations (1992). Earth Summit Agenda 21, Chapter 33 (Finance). United Nations, Washington, DC.

Waddell, S. (2002). Core competences: A key force in business-government-civil society collaborations. *J. Corporate Citizenship* **7**: 43–56.

Webb, M. and Ehrhardt, D. (1998). Improving water services through competition. World Bank Public Policy for the Private Sector Note 164.

World Bank (1997). *Toolkits for Private Participation in Water and Sanitation*. The World Bank, Washington, DC.

World Bank (1997). *Toolkits for Private Sector Privatization in Water and Sanitation*. World Bank, Washington, DC.

World Bank (1998). *Global Development Finance 1998*. World Bank, Washington, DC.

World Bank (2003). Private Participation in Infrastructure data base. Available at http://ppi.worldbank.org/

World Bank (2004). *Global Development Finance 2004*. World Bank, Washington, DC.

Yepes, G. (1999). Do Cross-subsidies help the poor to benefit from water and wastewater subsidies? Lessons from Guayaquil, Ecuador. UNDP-World Bank Water and Sanitation Program.

APPENDIX A: OVERVIEW OF MODELS OF PRIVATE SECTOR PARTICIPATION IN WATER AND SANITATION PROVISION

Full Privatization (Divestiture)

Private company not only takes full responsibility for operation, maintenance, and investment, but ownership of infrastructure is transferred from the public to the private sector at an agreed fee. The government is responsible for regulation.

Partial Private-Sector Responsibilitys

Responsibility for service provision is shared between the private and public sectors, with differing levels of responsibility being delegated to the private partner depending on the contract type.

In all of the following models, ultimate ownership of assets remains with the public sector.

Service Contract

Service contracts are usually short-term agreements whereby specific operations and maintenance activities are contracted to the private sector. The public sector retains overall responsibility for the administration of the service.

Management Contract

A management contract entails private sector responsibility for utility operation and maintenance but without the obligation of investment or commitment of private investment capital.

Lease Contract (Affermage)

Under lease contracts, the private firm operates and maintains the utility at its commercial risk, deriving revenue directly from tariffs, but it does not invest in new infrastructure.

Concession Contract

Under concession contracts, the private company manages the infrastructure facility and operates it at its commercial risk and accepts investment obligations. The role of the government in concession contracts is predominantly regulatory.

Build-Own-[Operate]-[Train]-[Transfer]-Type Contracts (BOO/BOT/BOOT/BOTT)

These are similar to concession contracts, but they are usually used for greenfield projects as the private contractor is also responsible for constructing the infrastructure. At the end of the contract, the assets may either remain with the private company or be transferred back to the government.

Cooperative Model

The cooperative model is a type of government-owned public-limited company (plc) subject to the rules and regulations of other plcs and of which most shares are publicly owned (either by government or citizens/users).

Informal Sector Provision

Provision of water and sanitation services to the poor by "informal" and/or small-scale operators is common in most low- and middle-income countries, especially where the poor lack access to formal service provision. In an increasing number of cases, governments are supporting small-scale private initiatives to increase services provision to the poor.

Sources: Blokland et al. (6), Calaguas (7), Kempe and Schreiber (8), Johnstone and Wood (9), Lewis and Miller (10), Nickson (11), and Ramaema (12).

METHODS OF REDUCING RADON IN DRINKING WATER

Paul D. Robillard
World Water Watch
Cambridge, Massachusetts

William E. Sharpe
Bryan R. Swistock
Pennsylvania State University
University Park, Pennsylvania

Since the late 1980s, radon has become a highly publicized health threat. This naturally occurring radioactive gas is seeping out of the earth's crust and into the basements of thousands of homes across the nation. Until recently, radon concerns have focused primarily on airborne radon; radon in drinking water was not considered a problem. Now, water tests reveal its presence, and many homeowners are asking water treatment dealers, "How can I reduce radon in my water supply?"

WATERBORNE RADON

Waterborne radon usually originates in deep wells that tap radon-contaminated groundwater. Radon increases household air levels during showering, laundering, and dishwashing. The EPA estimates that 2–5% of airborne radon comes from household water. They further estimate that even these small percentages increase the incidence of cancer. If radon is discovered in water, it is likely that radon is entering the house through the basement as well.

Currently, the EPA has not set official standards for either airborne or waterborne radon. EPA suggests that an airborne level of 4 pCi/L is a point at which remedial action should be taken. Recognize that for every 10,000 pCi/L in water, about 1 pCi/L will be released in the air. The EPA's proposed limit for radon in water is 300 pCi/L.

Detection and Testing

Radon and its daughters are radioactive—continually decaying and emitting radioactive particles called alpha and beta rays. Therefore, testing for radon in water requires special sampling and laboratory analytical techniques that measure its presence before it escapes from the sample. Direct water sampling is by far the most accurate testing method.

Treatment

Radon water treatment should remove radon before it becomes airborne.

Methods of Home Aeration

Home Aeration Units. Home aeration exposes the water to enough air so that radon can escape to the air before the water reaches the taps. Using new technological advancements in home aeration, these units can have radon removal efficiencies up to 99.9%. They are also ideal for high waterborne radon levels.

Spray Aeration Unit. A spray aeration unit, as shown in Fig. 1, sprays radon contaminated water into the tank using a spray nozzle. The increased surface area of the sprayed water droplets causes the radon to come out of solution, and the air blower carries the radon contaminated air to a vent outside the home. About 50% of the radon will be removed in the initial spraying. The water must be sprayed several times to increase removal efficiencies. To keep a supply of treated water, at least a 100-gallon holding tank must be used.

Packed Column Aeration Unit. In a packed column system, water moves through a thin film of inert packing material in a column. The air blower forces radon contaminated air back through the column to an outdoor vent. If the column is high enough, removal efficiencies can be between 90 to 95%. For a 6-foot column (shown in Fig. 2), the removal efficiency is around 95%. Packed columns become impractical if the radon level exceeds 20,000 pCi/L.

Shallow Aeration Unit. A final aeration system uses a shallow tray to contact air and water. Water is sprayed into the tray and then flows over the tray as air is sprayed up through tiny holes in the tray bottom (see Fig. 3). The system removes more than 99.9% of the radon and vents

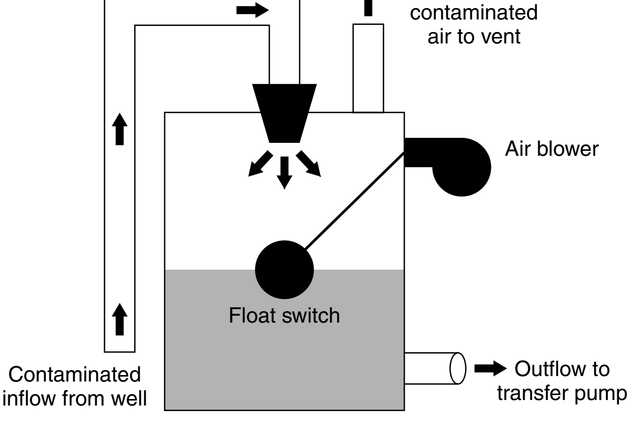

Figure 1. Radon removal using a home spray aeration system.

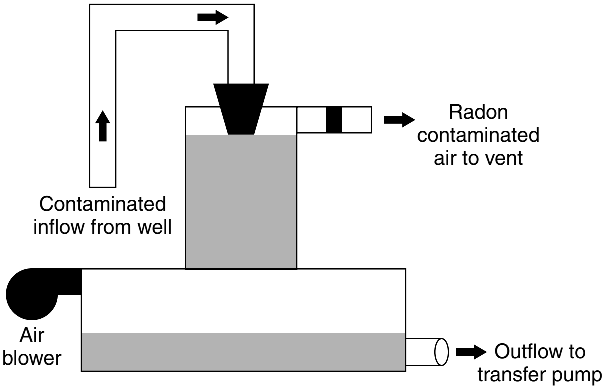

Figure 2. Radon removal using a packed column.

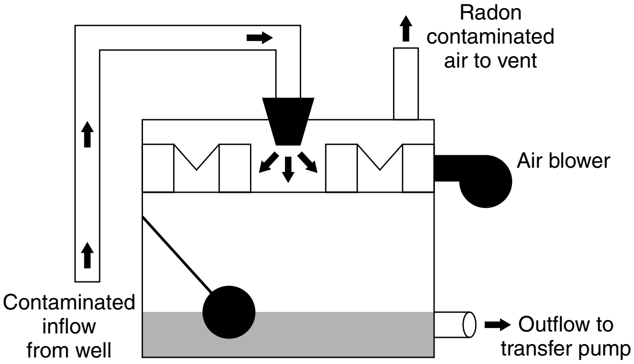

Figure 3. Radon removal by horizontally extended shallow aeration.

it outside the home. The treated water collects in the tank bottom and is pumped to the water pressure tank.

Advantages of this type of aeration include:

- Low pressure air blower
- No fouling problems in tray holes
- Small unit size

However, this unit uses 100 cubic feet per minute of air compared to the others, which can depressurize the basement.

Point-of-Entry Treatment

Another method for removing radon from water is a granular activated carbon (GAC) unit. Figure 4 shows a typical GAC unit.

For radon removal, GACs are constructed of a fiberglass tank containing granular activated carbon—a fine material that traps and holds the radon. Because of the carbon's fine particle size, it easily clogs with sediments or other contaminants in the water.

Important points to consider with GAC units are:

- Some GAC units come with a special backwashing feature for removing sediment, which eventually reduces the effectiveness of the carbon to remove radon.

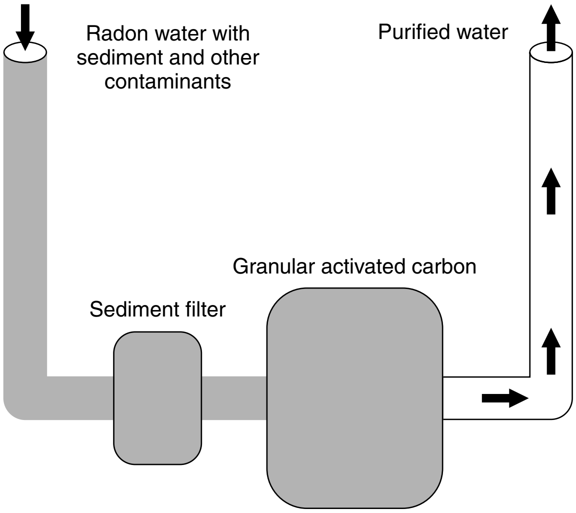

Figure 4. Treatment by activated carbon.

- Elimination of the sediment source or a sediment filter placed ahead of the GAC tank is the best protection against clogging.
- The maximum radon level at which a GAC unit operates effectively is uncertain. Some estimates show that it should not be used if waterborne radon levels exceed 30,000 pCi/L. Other experts say 5,000 pCi/L.
- It is important that the filter size matches the water use and conditions. According to the EPA, a 3-cubic-foot unit can handle as much as 250 gallons of water per day and effectively reduce radon levels. Typical water use in a home ranges from 50 to 100 gallons per person per day.
- GAC filter will remove radon indefinitely providing that sediments or organic pollutants have not clogged the filter.
- A major drawback to using a GAC filter is that if radon is present, the filter becomes radioactive as it picks up the gas. Lead-210 (a radon daughter) builds up on the carbon filter and then gives off its harmful radioactive rays, as it continues to decay.
- A GAC filter may produce a radiation problem when the device is used to remove other contaminants. For example, a GAC unit is installed to remove a pesticide without testing the water for radon. The GAC unit sits under the sink harmlessly removing the problem contaminant. Right? Wrong. Unfortunately, what the homeowner doesn't know is that the water supply has very high radon levels. So, while the GAC traps the pesticide, it also traps radon, thus producing a radioactive filter and a radiation hazard.
- Proper maintenance and handling of the GAC unit can minimize exposure risks. Redevelopment of the well intake or a sediment filter is vital to protecting the fine carbon from fouling and clogging; protected filters won't need to be changed as often. The water should also be periodically retested to insure that radon is still being removed.

WATER REUSE

Peter S. Cartwright, P.E.
Minneapolis, Minnesota

Water reuse is the general term applied to the act of recovering water from a process and reusing it in the same process, or another one, before discharging it. Some experts attach labels to this activity based on the specific use of the water: *Recycle* involves redirecting the recovered water back to the same process. *Reuse* means reusing the recovered water in a different application. *Recovery* generally refers to the technologies used to accomplish this.

Black water is defined as the effluent from toilets and garbage disposers in residences. *Gray water* in the residential environment typically refers to the effluent from bathing, showering, laundry, and dishwashing and other effluents from normal household activities. *Gray water* is distinguished from black water in that it contains much less organic loading and is expected to contain much less fecal coliform bacteria and other pathogenic organisms.

We all are familiar with the admonitions addressing the finite quality of water on this planet: the fact that only 1% is considered “fresh” (nonseawater), but is used for virtually all human activities and, as a result, is rapidly deteriorating in quality. The mindset is there, as are the technologies. The barrier is the commitment to make the economic investment.

From a technical standpoint, there are virtually no barriers to the quality improvement of either black water or gray water, even to the point of drinking water quality! Today it is possible to install a “black box” on the sewer line from residences and drink the treated water coming from it. Presently, the deterrents to this are economics and the “yuck” factor.

Drinking water regulations continue to place new standards for water quality, and industrial and commercial activities are requiring higher quality water in many applications. Due to population growth, water quantity requirements are also steadily increasing. For example, today it is estimated that one-fifth of the world’s population does not have access to safe drinking water; by 2025, the global population is expected to be 8.3 billion people (up from about 6.4 billion in 2004), and the usage of industrial water is expected to double by then. Interestingly, even in cities experiencing severe drinking water shortages, there always seems to be enough water to flush toilets.

Industries, in general, are very poor stewards of water conservation. Most water brought into the plant is used only once, even though that use may have had very little impact on water quality. In most of the industrialized world, we have labored under the misconception that our water supply was inexhaustible as well as inexpensive. This cost aspect has been exacerbated by the fact that, in many areas, the price of water has been partially subsidized by local government.

Undoubtedly, a paradigm shift is taking place with regard to water conservation and reuse. The barriers to reuse activity are

1. the “yuck” factor
2. lack of practical reuse technologies
3. economic factors
4. commitment to reuse

1. The average consumer recoils at the thought of drinking (or even reusing) “sewer water,” not considering the fact that the drinking water for one community is very likely the wastewater discharged from another community on the same river or lake. It is estimated that in major U.S. rivers, water is reused as many as 20 times by the time that river empties into the sea. Most “groundwater recharge” is treated municipal wastewater injected into groundwater supplies to provide sufficient storage capacity or to serve as a barrier to salt water intrusion from the oceans.
2. The key to water reuse is to have an arsenal of technologies available to remove hazardous or undesirable contaminants efficiently from the water supply. There is no single technology that efficiently removes all classes of contaminants; however, for the past century or so, there have been significant developments in treatment technologies that effectively reduce the concentration of virtually any contaminant to acceptable levels for any water use.

   There is little argument that reducing the huge variety of contaminants that may be encountered in typical wastewater, resulting from the combination of sewage, industrial wastewaters, and perhaps even effluent from surface water runoff, requires a stunning array of advanced technologies. The linchpins are the membrane separation technologies of reverse osmosis, nanofiltration, ultrafiltration, and microfiltration. Whereas reverse osmosis has been around for more than 50 years, the others, most notably nanofiltration, are relatively new developments.

   By separating the treatment process into its key components, pretreatment, primary treatment, and posttreatment, it is now possible to create optimum technology trains that can purify the stream from virtually any source and condition it for virtually any reuse.
3. In addition to treatment technology costs, it is also necessary to factor in the fact that raw water costs are steadily increasing and will continue to do so, reflecting the diminishing supply and costs of meeting new regulations.
4. It is estimated that as of the year 2000, there were more than 10,000 water reuse systems installed in the United States, almost all of this recovered water was used to irrigate agricultural fields or residential and commercial landscaping. In California alone, it is estimated that 120 billion gallons per year are reused.

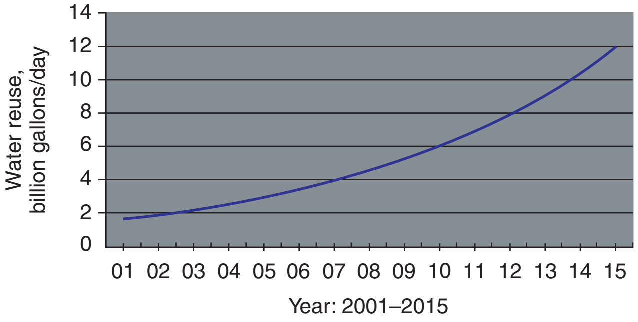

Figure 1. Projected water reuse in 21st century.

The rate of water reuse in the United States is expected to increase markedly over the next decade, as indicated by Fig. 1.

SPECIFIC TECHNOLOGIES

To understand the technologies most applicable to water reuse, it is necessary to understand the contaminants prevalent in water supplies. Contaminants can be categorized by their physical and chemical properties. Table 1 provides this classification.

Suspended Solids

Removing suspended solids from water supplies is probably the oldest water treatment procedure. Throughout history, humankind has used everything from containers of sand to cloth to charcoal to "clarify" water to make it look and taste better. The chronology of water treatment technology development also underscores the improvements to suspended solids removal processes; the latest is microfiltration.

Dissolved Organics

Removing dissolved organic contaminants requires the greatest variety of technologies, reflecting the diversity of dissolved organic chemicals. For most of these contaminants, there are several choices of technologies to effect removal. Some may involve adding chemicals, such as alum, powdered activated carbon, or an acid; other technologies may include physical separation as with a coalescer or ultrafiltration technology.

Dissolved Ionics (salts)

These are contaminants that have ionic charges and are almost all inorganic chemicals. The somewhat limited choice of removal technologies consists of membrane technologies of reverse osmosis, nanofiltration, and electrodialysis; adsorptive resin technology known as ion exchange and the hybrid of electrodialysis and ion exchange known as electrodeionization. Distillation is the oldest technology, mirroring the natural water cycle, but, because of its high energy cost, is now used only in specialized applications.

Table 1. Chemical Properties of Contaminants

Class	Typical Example
Suspended solids	Dirt, clay, colloidal materials
Dissolved organics	Trihalomethanes, synthetic organic chemicals, humic acids, fulvic acids
Dissolved ionics (salts)	Heavy metals, silica, arsenic, nitrate
Microorganisms	Bacteria, viruses, protozoan cysts, fungi, algae
Gases	Hydrogen sulfide, methane, radon

Microorganisms

Microorganism contaminants in most water supplies are from one or more of the following categories:

bacteria
viruses
protozoan cysts
fungi
algae

There are a number of disinfection technologies that inactivate or remove microorganisms. These include chemical (chlorine compounds and ozone). Certain heavy metals in solution inactivate microorganisms. The most common of these are silver for bacteria reduction and copper for algae inactivation. The most prevalent nonchemical technology is ultraviolet irradiation, although heat is still used occasionally in specialized applications, such as preparation of "water for injection" in the pharmaceutical industry. Many disinfectants are effective on only certain classes or types of microorganisms or under very specific conditions.

The most troublesome class of microorganisms is bacteria. Because bacteria are viable and grow under virtually any condition, they are impossible to eliminate completely. In most applications, the goal is to minimize bacterial growth so as not to interfere with the water use.

CONCLUSION

Water reuse is not an abstract concept; it is both a reality and a necessity. For the reasons cited, the requirements and opportunities for water reuse will continue to grow at an increasingly rapid rate.

ROOF DRAINAGE HYDRAULICS

Scott Arthur
Grant Wright
Heriot-Watt University
Edinburgh, Scotland, United Kingdom

INTRODUCTION

Over the past decade, urban drainage systems have moved toward what are now commonly known as "sustainable urban drainage systems" (SUDS) or "best management practice" (BMPs). Fundamental to the implementation of these systems is addressing both runoff quantity and quality at a local level in a manner which may also have the potential to offer benefits to stakeholders. This has led to a change in the way new developments now look and interact within catchments. However, despite the availability of such tools to reduce, attenuate, and treat urban runoff, substantial areas of the urban environment are still 100% impermeable and drain rapidly, namely, roof surfaces. Normally, roof drainage systems do not always receive the attention they deserve in design, construction, and maintenance. Although the cost of a system is usually only a small proportion of a building's total cost, it can be far outweighed by the costs of the damage and disruption resulting from a failure of the system to provide the degree of protection required.

There are basically two different types of roof drainage system; conventional and siphonic (refer to Fig. 1). Conventional systems operate at atmospheric pressure, and the driving head is thus limited to the gutter flow depths. Consequently, conventional roof drainage systems normally require a considerable number of relatively large diameter vertical downpipes, all of which have to connect to some form of underground collection network before discharging to a surface drain. In contrast, siphonic roof drainage systems are designed to run full bore, resulting in subatmospheric pressures, higher driving heads, and higher flow velocities. Turbulent gutter conditions mean that there will always be a small percentage of entrained air within the system (typically 5%). Hence, siphonic systems normally require far fewer downpipes, and the depressurized conditions also mean that much of the collection pipework can be routed at high level, thus reducing the extent of any underground pipework.

Both types of drainage system comprise three basic interacting components:

- the roof surface
- the rainwater collection gutters (including outlets)
- the system pipework

Each of these components can alter the runoff hydrograph substantially as it is routed through the system. This text focuses on the role and performance of each of these components. As the principles of siphonic drainage are generally less well understood and certainly less well documented, particular emphasis placed on the performance of siphonic roof drainage systems.

ROOF SURFACE

The design of the roof surface is usually within the remit of the architect rather than the drainage designer. Notionally, there are three types of roof surfaces:

Flat Roofs

Flat roofs are normally associated with domestic properties in climates with low rainfall and with industrial buildings in developed countries. Such roofs are seldom truly "flat" but simply fall below the minimum gradient associated with sloped roofs in the jurisdiction under consideration; for example, in the United Kingdom, a flat roof is one whose gradient is less than 10° (1). Minimum gradients are usually specified to avoid any unwanted ponding (BS EN 6229:2003 specifies a 1 in 80 minimum gradient) and to help prevent the development of any adverse gradient due to differential settlement (2).

Although flat roofs can be problematic if not maintained properly, they are often preferred; they reduce the amount of dead space within the building and they attenuate flows more than sloped surfaces.

Sloped Roofs

Most residential and many commercial properties have sloped roofs. Such roofs are generally favored because their ability to drain naturally means that there is less risk of leakage. In temperate climates, their specification also means that snow loading is less of an issue. Once a rainfall is underway, the rate at which the runoff flows across a roof is a function of roof slope and roughness. Where rainfall data are available, runoff rates from roof surfaces may be readily assessed using kinematic wave theory (3).

Green/Brown Roofs (Sloped or Flat)

The oldest type of permanent roof is a green roof. These involve planting roof areas to attenuate and/or dissipate rainfall and can take the form of a rooftop garden with trees and shrubs (termed intensive) or a lightweight carpet of growth media and flora (termed extensive). The latter technology is already employed widely (e.g., the Rolls Royce plant at Goodwood; purportedly Europe's largest green roof). Many of these applications tend to focus on the aesthetic benefits such systems offer to high profile developments and are often installed to "green"

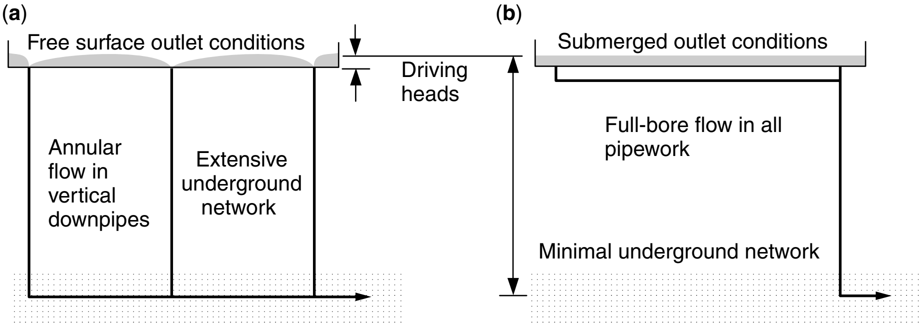

Figure 1. Schematics of a typical conventional (**a**) and (**b**) siphonic roof drainage system (at normal design condition).

a development and thus help secure planning consent in sensitive areas (4). However, as well as being aesthetically pleasing and hydraulically beneficial, green roofs may also offer thermal insulation (5), reduce the heat island effect, the phenomenon whereby absorption of solar radiation by urban surfaces causes a marked increase in ambient air temperature (6), provide acoustic damping, and extend the service life of the roof membrane (7–10).

Green roof systems are used extensively in Germany and to a lesser extent in North America, but again their specification is primarily due to a desire for a reduced aesthetic impact associated with a particular development. Germany probably has the most experience to date, a direct result of their use in the 1800s as a low fire risk alternative to tarred roofs in deprived urban areas (11). Currently, German research is focused predominantly on planting issues, and there is only a limited understanding of how the systems may be used to mitigate the impact of urban runoff. One research project, which ran from 1987–1989 in Neubrandenburg (8), found that an installed green roof with 70 mm of substrate could reduce annual runoff from a roof by 60–80%. Work in Vancouver (Canada), based on an uncalibrated computer model, suggests that for catchments where the roof area comprises 70% of the total surface, installing an extensive system could reduce total runoff to approximately 60% over 12 months (12). The same model was also used to assess specific synthetic rainfall; these results indicated that the catchment experienced increased runoff during longer rainfalls.

Neither of these studies detail how green roofs could be expected to perform during a particular rainfall or where efficiencies may be gained in the design of collection pipework. Limited testing in the United States (13), where green roofs are often irrigated, has indicated that runoff can be reduced by 65% during a single rain. The most authoritative design guidance for green roofs in the United States is produced by The New Jersey Department of Environmental Protection (14). This is focused on lightweight structures and gives guidance on how to ensure "rapid draining" where the rainfall return period exceeds 2 years.

Rainfall return periods are normally set within the context of failure probability and consequence. Conventional systems are usually designed assuming 100% runoff for a 2-minute storm; the 2-minute duration is selected because it is the typical time of concentration for conventional systems. Although advice is given in codes for setting higher runoff rates, there is little guidance on setting runoff rates below 100%. These observations mean that inadequacies are encountered if conventional codes are used to design green roofs:

- Runoff *coefficients* should be expected to be below that used for conventional roofs; 100% is used by BS EN 12056-3:2000 and 98.7% was recorded by Pratt and Parkar (15).
- Peak runoff *rates* are reduced; even where there is no infiltration, the surface roughness has a significant impact.
- Time of concentration is expected to be greater than 2 minutes; particularly relevant when designing collection pipework for large roof areas for public sector, commercial, and industrial properties.
- As with other elements of urban drainage design, it is not efficient for a complex system such as a green roof to be matched to a single rainfall. It is probable that the duration of runoff hydrographs will be orders of magnitude longer compared with conventional systems, and runoff interactions between independent rainfalls are probable; this may make a time-series approach more appropriate.

RAINWATER COLLECTION GUTTERS

The basic requirement for rainwater collection gutters is that they have sufficient flow capacity to accommodate flows from the design storm (16). Although it is common practice to install gutters at a slight gradient to prevent ponding, the nature of the construction industry and the process of settlement means that it is normal to assume that gutters laid at slack gradients are actually flat; for example, BS EN 12056-3:2000 stipulates that gutters at gradients less than 0.3% shall be treated as flat (17). In a level gutter, the water surface profile will slope toward the outlet, and it is the difference in hydrostatic pressure along the gutter that gives the incoming water the required momentum to flow toward the outlet (18).

Gutter Outlet Depths

Key to ensuring whether or not collection gutters have sufficient capacity are the conditions that occur at the gutter outlets. As well as affecting the flow rates entering the drainage system pipework, the outlet depths also affect upstream gutter depths (via the backwater surface profile). Hence, although the depth at a gutter outlet may not cause any particular problems, the greater depths at the upstream end of the gutter may result in overtopping.

Extensive experimental studies in the 1980s determined that the flow conditions in the vicinity of a gutter outlet in a conventional roof drainage system could be categorized as either "weir" type or "orifice" type, depending on the depth of water relative to the size of the outlet (19). At depths below that equivalent to half of the outlet diameter, the flow conditions are "weir" type, and outlet conditions are calculated using an appropriate sharp-edged weir equation (18). At higher flow depths, the flow effectively "chokes," and the flow regime changes to "orifice" type; the outlet conditions are calculated by an appropriate sharp-edged orifice equation (18). Although conventional roof drainage systems are usually designed to ensure free discharge at gutter outlets, design restrictions may mean that the outlets cannot discharge freely; in such circumstances, additional gutter capacity (storage) is normally required to accommodate the resulting higher flow depths.

In siphonic roof drainage systems, the outlets are designed to become submerged to allow full-bore flow to develop and be sustained; if this is the case, the determination of outlet depth is complicated as the gutter conditions depend on downstream conditions (within the connected pipework) as well as gutter inflows. Recent experimental work has also indicated that conventional roof drainage systems incorporating “nonstandard” gutter sections, whose base width and height are significantly greater than the diameter of the outlet, can result in the development of full-bore flow in the vertical downpipe and siphonic action (20); for a given gutter section, the onset and extent of such conditions depend on the diameter of the downpipe. Similar phenomena have also been observed in “standard” gutter sections (semicircular and elliptical); in these cases, limited siphonic action occurs for only a short distance below the outlet (18).

Flow Division within Gutters

In terms of flow division between multiple outlets in a gutter under free discharge, it can be seen from Fig. 2a that the flow splits evenly in any given gutter section (between two outlets or between an end wall and an outlet), whether or not the gutter inflow is uniform or nonuniform. Figure 2b, c indicates the effect of outlet placement within a gutter; evenly spaced outlets require far less gutter capacity than those placed at gutter extremities.

Where outlets are not freely discharging, the flow division between multiple outlets in a gutter may not be as described, as the individual gutter sections may “hydraulically merge” to form one continuous channel and/or downstream system conditions may become significant. For example, the pipework in a siphonic system runs full bore when operating at or near its design point, and the flow division between outlets depends on the relative losses for each branch of the system.

Backwater Profiles

The water surface profile in gutters can only be assessed realistically by applying the momentum equation for channels with lateral input. In many cases, the low velocities associated with gutter flows mean that gutter friction losses are minor and may be ignored (18). If a gutter outlet allows free discharge and frictional effects are neglected, the backwater profile may be determined by applying Equation 1 to determine the horizontal distance (ΔL) between any given upstream depth (h_1) and downstream depth (h_2).

$$\Delta L = \int_{h_1}^{h_2} \left[\frac{\left(1 - \dfrac{Q^2 T}{g A^3}\right)}{\left(S_o - \dfrac{Q^2}{A^2 m C^2}\right)} \right] dh \qquad (1)$$

where Q = flow rate (m^3/s)
T = surface width (m)
g = gravitational constant ($\mathrm{m/s}^2$)
A = flow area (m^2)
S_o = bed slope (-)
m = hydraulic mean depth (m)
C = Chézy coefficient (-)

This equation can be modified if frictional effects are significant (very long gutter lengths or very high flow velocities) or if the gutter outlet is not freely discharging.

Current Design Methods

The foregoing discussion has highlighted the key elements that should be considered when designing a rainwater gutter. However, without recourse to some form of numerical modeling, it is not feasible to calculate backwater surface profiles, and hence gutter capacities, for roof drainage systems; this is particularly the case for large commercial or manufacturing developments which may incorporate many kilometers of different types of guttering. Consequently, current gutter design methods for gutters installed in conventional drainage systems are based primarily on empirical relationships (19) and the assumption of free discharge at the outlet. For example, BS EN 12056-3:2000 specifies that the design capacity of a “short,” level, semicircular gutter located on the eaves of

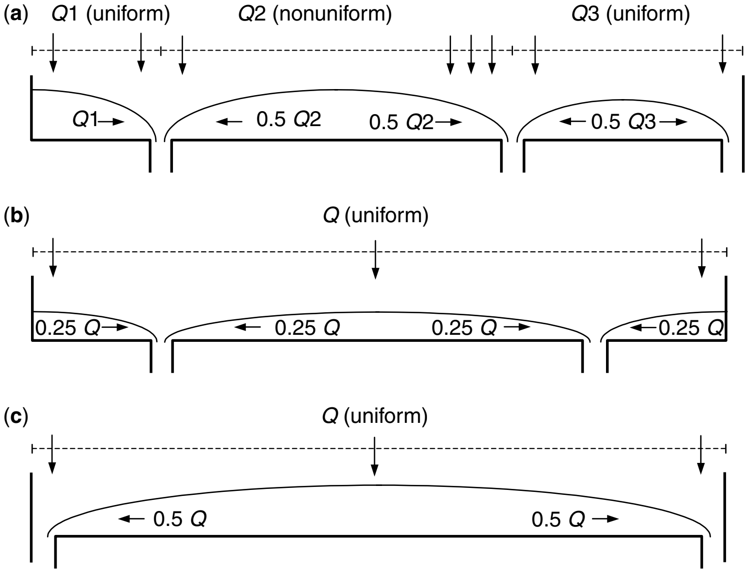

Figure 2. Effect of outlet positioning on flow division in gutters. (**a**) Flow division between multiple outlets in a gutter. (**b**) Flow division between evenly spaced outlets in a gutter. (**c**) Flow division between outlets positioned at gutter extremities.

a building (with outlets capable of allowing free discharge) is given by (17)

$$Q_{\mathrm{L}} = 0.9 \times Q_{\mathrm{N}} = 0.9 \times 2.78 \times 10^{5} \times A_{\mathrm{E}}^{1.25} \qquad (2)$$

where Q_{N} = notional gutter design capacity (L/s)
A_{E} = gutter cross-sectional area (mm^2)

Clearly not all gutters can be designed by application of Equation 2. For example, BS EN 12056-3:2000 (17) contains clauses to account for many eventualities, including

- location of gutter on building that may result in varying consequences of failure; eaves gutter, valley gutter, parapet gutter
- differently shaped gutter sections
- "hydraulically long" gutters (where frictional effects may be significant)
- gutters laid at a significant gradient
- changes in gutter alignment (bends, offsets, etc.)
- additional system elements, such as strainers or rainwater hoppers
- restricted flow at outlets
- gutters installed in siphonic roof drainage systems

In addition to the type of clauses listed above, BS EN 12056-3:2000 also allows designers to use data obtained from experimental testing of a particular arrangement (17).

Numerical Models

Numerical models have been developed that can accurately simulate the flow conditions in any type of gutter as a result of either steady or unsteady roof runoff. An example of this is incorporated into the "ROOFNET" model recently developed as part of an academic research project dealing with the effect of climate change on urban drainage (20). This model enables the user to specify data describing the relevant aspects of a particular installation, including details of the prevailing rainfall conditions, details of the roof surfaces to be drained, and details of the actual gutters. A kinematic wave model is then used to route the rainfall over the roof surfaces and into the gutters. A method of characteristics solution of the fundamental equations of one-dimensional flow in open channels is then used to route the runoff along the gutters to the outlets (21), at which point the flow enters the drainage pipework. The model automatically determines the flow conditions at the gutter outlets and, in addition to dealing with free discharge, can also simulate the effect of restricted flow and submerged outlet scenarios. Output includes depths, velocities, and flow rates along the gutter, as well as the location and severity of any gutter overtopping.

At present, models such as those described before are research tools; they are normally developed and used by universities for specific research projects. However, it is envisaged that such models may soon be used as diagnostic design aids, particularly for national code development.

SYSTEM PIPEWORK

The type and extent of pipework incorporated into a roof drainage system depends primarily on whether the system is conventional or siphonic.

Conventional Rainwater Systems

In conventional roof drainage systems, the aboveground pipework generally consists of vertical downpipes, connecting the gutter outlets to some form of underground drainage network, and offset pipes, used where the gutter overhang is significant. Note that an offset pipe is defined as a pipe with an angle less than 10° to the horizontal. The capacity of the system as a whole usually depends on the capacity of the gutter outlets rather than the capacity of the vertical downpipes.

The flow within vertical downpipes is normally free surface; BS EN 12056-3:2000 (17) specifies that downpipes run no more than 33% full; this effectively installs redundant capacity within the system. If the downpipes are sufficiently long (normally greater than 5 meters), annular flow may occur. Similarly, the flow within offset pipes will also normally be free surface; BS EN 12056-3:2000 (17) specifies that offsets run no more than 70% full, indicating the need to install all offsets at a gradient. The design of the pipework can either be undertaken using the design tables in BS EN 12056-3:2000 or by applying the Wyly–Eaton equation for vertical downpipes (22) and the Colebrook–White equation for offset pipes (23).

Siphonic Roof Drainage Systems

In contrast to conventional systems, siphonic installations depend on purging air from the system (priming) and subsequently establishing full-bore flow within the pipework connecting the outlets in the roof gutters to the downstream surface water sewer network (at ground level).

Current design practice assumes that, for a specified design storm, a siphonic system fills and primes rapidly with 100% water (24). This assumption allows siphonic systems to be designed using steady-state hydraulic theory. The steady flow energy equation is normally employed (25), and the elevation difference between the gutter outlets and the point of discharge is equated to the head losses in the system. Although this approach neglects the small quantities of entrained air that always enter a siphonic roof drainage system, it reportedly yields operational characteristics similar to those observed in laboratory test rigs in the fully primed state (25,26).

However, steady-state design methods are not applicable when a siphonic system is exposed to rainfall below the design criteria or with time-varying rainfall intensity. In the former case, the flow may contain substantial quantities of entrained air and exhibit pulsing or cyclical phases, a result of greatly varying gutter water levels and an indication of truly unsteady, transient flow. Such problems are exacerbated when the system incorporates more than one outlet connected to a single downpipe (multi-outlet system), as the breaking of full-bore conditions at one of the outlets (due to low gutter depths and air entry) is transmitted throughout the system and, irrespective of the gutter depths above the remaining outlet(s), results in cessation of fully siphonic conditions. As subdesign events are the norm, it is clear that current design methods may not be suitable for assessing the day-to-day performance characteristics of siphonic roof drainage systems. This is a major disadvantage, as it is during these events that the majority of operational problems tend to occur, for example, noise and vibration.

Despite any defects that current design methods may have, thousands of systems have been installed worldwide with very few reported failures. Where failures have occurred, they have invariably been the result of one or more of the following:

1. a lack of understanding of operational characteristics
2. poor material specification
3. installation defects
4. a poor maintenance program

In response to these perceived shortcomings, a series of research projects has recently been undertaken to augment the understanding of siphonic roof drainage systems and to develop numerical models for use as diagnostic design aids (27). The remainder of this section will present a selection of the salient points arising from this work.

In contrast to the assumption made in current design methods, the priming of a typical siphonic system actually found was as follows (refer to Fig. 3):

1. Flow conditions throughout the system are initially free surface (Phase 1).
2. Full-bore flow forms at some point within the horizontal pipework (Phase 1).
3. Full-bore flow conditions propagate downstream toward the vertical downpipe and upstream toward the gutter outlets (Phase 1).
4. Full-bore flow conditions reach the vertical downpipe, the downpipe starts to fill, and the system starts to depressurize (Phase 2).
5. Once the conditions throughout the downpipe are full bore, any remaining air pockets are purged from the system (Phase 2).
6. Full siphonic action occurs (Phase 3) and continues until the gutter depth(s) falls below the level at which air can enter the system.

The data shown in Fig. 4a illustrate the type of unsteady flow conditions that occur when a siphonic system is exposed to rainfall below the design point and the gutter flow depths are insufficient to sustain full siphonic action. The data shown in Fig. 4b illustrate the type of unsteady flow conditions that occur when an installed siphonic system is exposed to a “real” rainfall and the rainfall intensity varies with time.

Figure 5 shows an example of the output from one of the numerical models that has recently been developed (SIPHONET). As can be seen, the model can accurately simulate the priming of a siphonic system (0–32 s) as well as steady siphonic conditions (32–62 s). These data also illustrate that the model can simulate complex operating conditions, such as the rise in system pressure when the depth in gutter 1 drops below that necessary for full-bore flow, hence allowing air to enter the system and break the siphon (at approximately 62 s).

CONCLUSION

The text has illustrated how roof drainage systems are a key, but often overlooked, element of urban drainage infrastructure. It has also been shown that their design is a complex process, which relies heavily on gutter outlet performance. The following conclusions may be drawn with respect to the operation of roof drainage systems:

1. Their operation depends on three interacting components: the roof surface, the collection gutter, and the collection pipework.
2. Green or brown roofs provide an opportunity to reduce the flow from roof surfaces, improve urban aesthetics, and increase biodiversity.
3. Outlet conditions are key to understanding how a system performs.

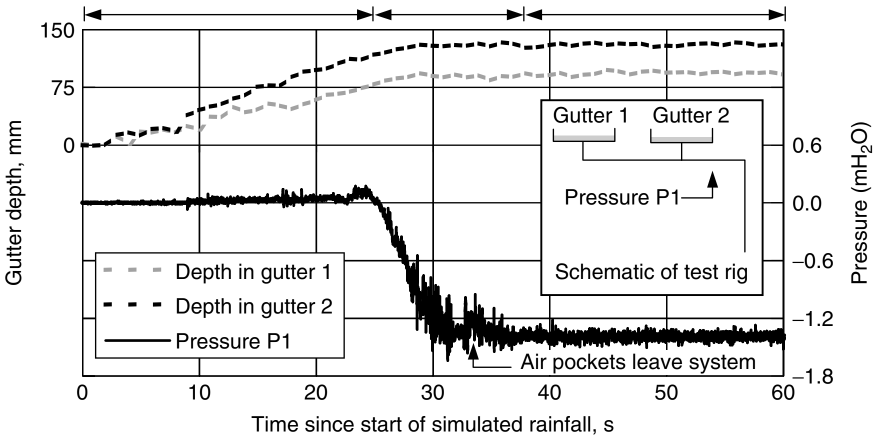

Figure 3. Priming of a laboratory siphonic drainage test rig (28).

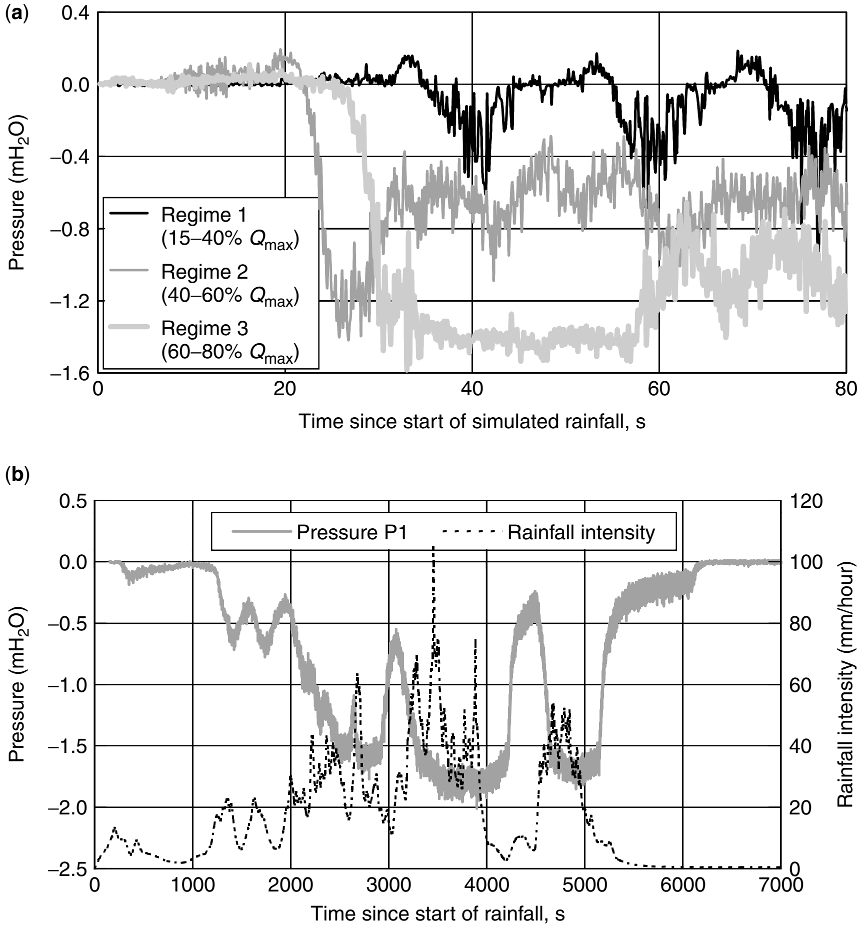

Figure 4. (**a**) Measured system pressures for subdesign rainfall within a laboratory siphonic drainage test rig (28). Note that data refer to the pressure P1, as indicated in Fig. 3. (**b**) Subdesign rainfall within an installed siphonic drainage system (28).

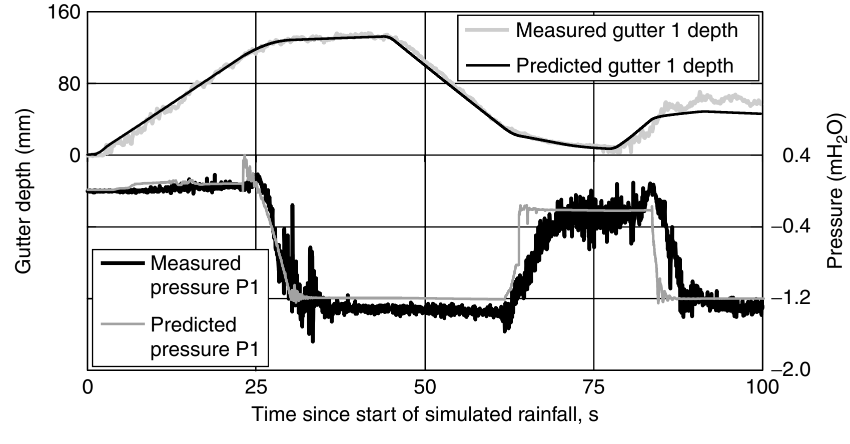

Figure 5. Measured and predicted system conditions within a laboratory siphonic drainage test rig: no inflow into gutter 1 between 62 and 82 s (28). Note that data refer to the system shown in Fig. 3.

4. Siphonic roof drainage systems present a more efficient way to drain large roof surfaces.
5. The design of siphonic roof drainage systems should consider subdesign rainfall and operational problems, such as blocked outlets.

THE FUTURE

Although green roofs are an attractive alternative, it is probable that conventional roof surfaces will continue to dominate domestic installations. However, it is likely that green roofs will experience a step-change in acceptance by the commercial sector once more becomes known about their performance and sustainability. Similarly, the efficiencies offered by siphonic systems means that they will continue to play a significant role in draining large commercial buildings, particularly if numerical models are applied diagnostically to improve performance and reduce costly system failures.

The biggest threat to roof drainage comes from climate change. Existing systems may not simply become more prone to flooding; changes in rainfall patterns may result in long periods of low precipitation, and self-cleansing velocities may be attained less frequently as a result. Furthermore, changes in wind patterns may also increase levels of rooftop debris and hence necessitate enhanced maintenance programs. As concern regarding climate change grows and the sustainability agenda widens, it is possible that harvesting roof runoff may become more widespread. At present, water consumption varies globally between 7 and 300 liters/household/day (L/h/d). In the United Kingdom, average consumption is 145 L/h/d, but only 1–2 liters may actually be consumed by humans, 30% may be used for WC flushing (29). Studies have shown that, when coupled with storage, roof rainwater harvesting has the potential to contribute substantially to domestic water usage in both developing and developed countries (30,31).

BIBLIOGRAPHY

1. BS EN 6229:2003. (2003). *Flat Roofs with Continuously Supported Coverings*, Code of practice, British Standards Publishing Limited (BSPL).
2. Simmons, T. (1994). Methods for designing proper roof drainage. *Prof. Roofing* 22–25.
3. Singh, V.P. (1996). *Kinematic Wave Modeling in Water Resources: Surface Water Hydrology*. John Wiley & Sons, New York.
4. Tarr, A.R. (2003). Green roof implementation: construction and contractor issues in the UK. *Green Roofs for Healthy Cities*. University of Sheffield.
5. Köhler, M. (2004). The multi-beneficiary system of green roofs: The green roof challenge to biophilic architecture and ecology. *Proc. Sheffield Conf. "Nature Enhanced"*.
6. Rosenfeld, A.H. et al. (1995). Mitigation of urban heat islands: materials, utility programs, updates. *Energy Build.* **22**(3): 255–265.
7. Hendricks, N.A. (1994). Designing green roof systems: A growing interest. *Prof. Roofing* 20–24.
8. Köhler, M., Schmidt, M., Paiva, V.L.A., and Taveres, S. (2002). Green roofs in temperate climates and in hot-humid tropics—far beyond aesthetics. *Environ. Manage. Health* **13**(4): 382–391.
9. Niachou A. et al. (2001) Analysis of the green roof thermal properties and investigation of its energy performance. *Energy Build.* **33**(7): 719–729.
10. Onmura, S., Matsumoto, M., and Hokoi, S., (2001), Study on evaporative cooling effect of roof lawn gardens. *Energy Build.* **33**: 653–666.
11. Köhler, M. (2004). Green roof technology—From a fire-protection system to a central instrument in sustainable urban design. *Second Green Roof Conf.*, Portland, Oregon.
12. GVRD. (2002). *Effectiveness of Stormwater Source Control*. Greater Vancouver Regional Sewerage and Drainage District.
13. PSU (2001). Online Research http://www.rps.psu.edu/0105/-roofs.html.
14. NJDEP (2000). *Standard for Rooftop Runoff Management*.
15. Pratt, C.J. and Parkar, M.A. (1987). Rainfall loss estimation on experimental surfaces. *Proc. Fourth Int. Conf. Urban Storm Drainage*, International Association of Hydraulic Runoff Research, Lausanne, Switzerland.
16. Beij, H. (1934). *Flow in Roof Gutters*. US Department of Commerce, Bureau of Standards: Research Paper RP644, Bureau of Standards Journal of Research, Vol 12.
17. BS EN 12056-3:2000. (2000). *GravitY Drainage Systems Inside Buildings*. Roof drainage, layout and calculation, British Standards Publishing Limited (BSPL).
18. May, R.W.P. (1995). *Design of conventional and siphonic roof drainage systems*, Public Health Services in Buildings—Water Supply, Quality and Drainage, IWEM Conference, London.
19. May, R.W.P. (1984). *Hydraulic Design of Roof Gutters*. Proceedings of the Institution of Civil Engineers, Part 2, Vol. 77.
20. Blanksby, J., Ashley, R., Saul, A.J., Cashman, A., Packman, J., Maksimovic, C., Jack, L., Wright, G., and Kay, D. (2004). Adaptable urban drainage (AUDACIOUS), *NOVATECH 2004: 5th Int. Conf. sustainable Tech. Strategies Urban Water Manage*.
21. Escarameia, M. and Swaffield, J.A. (1999). Prototype monitoring and numerical simulation of roof drainage gutter systems. *Proc. 8th Int. Conf. Urban Storm Drainage*, Sydney, Australia.
22. Wyly, R.S. and Eaton, H.N. (1952). *Capacities of Plumbing Stacks in Buildings,* National Bureau of Standards Building Materials and Structures, Report BMS 132.
23. Chadwick, A. and Morfett, J. (2004). *Hydraulics in Civil and Environmental Engineering*, 4th Edn. Spon Press.
24. Arthur, S. and Swaffield, J.A. (2001). Siphonic roof drainage: The state of the art. *Urban Water* **3**(1): 43–52.
25. May, R.W.P. and Escarameia, M. (1996). *Performance of siphonic drainage systems for roof gutters*. Report No SR 463. HR Wallingford.
26. Arthur, S. and Swaffield, J.A. (2001). Siphonic roof drainage system analysis utilising unsteady flow theory. *Build. Environ.* **36**(8): 939–948.
27. Wright, G.B., Swaffield, J.A., and Arthur, S. (2002). Investigation into the performance characteristics of multi-outlet siphonic rainwater systems. *Build. Serv. Eng. Res. Technol.* **23**(3): 127–141.
28. Swaffield, J.A., Wright, G.B., Jack, L.B., and Arthur, S. (2004). Pressure transient analysis to inform system design for building and roof drainage systems. *Proc. The Practical Application of Surge Analysis For Design And Operation, 9th Int. Conf.*
29. Butler, D. and Davies, J. (2004). *Urban Drainage*, 2nd Edn. Spon Press.
30. Thomas, T. (1998). Domestic water supply using rainwater harvesting. *Build. Res. Inf.* **26**(2): 94–101.
31. Fewkes, A. (2000). Modelling the performance of rainwater collection systems: towards a generalised approach. *Urban Water*, **1**(4): 323–333.

SEPTIC TANK SYSTEMS

John E. Moore
Hydrologic Consultant
Denver, Colorado

According to Wilson and Moore (1), a septic tank is an "underground vessel for treating wastewater from a single dwelling or building by a combination of settling and anaerobic digestion. Effluent is usually disposed of by leaching. Settled solids are pumped out periodically and hauled to a treatment facility for disposal." When properly

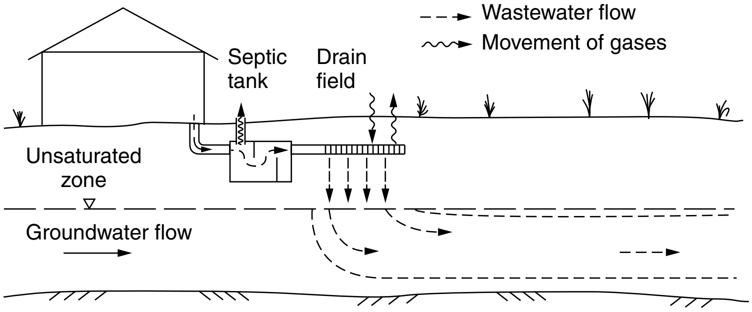

Figure 1. Components of septic tank systems (4).

sited, constructed, and maintained, septic systems can provide a low-cost environmentally responsible method of waste disposal. Improperly sited, constructed, operated, or maintained septic systems can, however, lead to water quality degradation and threats to public health. The basic components of a septic tank system are shown in Figs. 1 and 2. The septic tank is an enclosed receptacle designed to collect wastewater, segregate floatable solids, accumulate, consolidate, and store solids; wastewater treatment is provided by septic tank systems. The tank is the most important component used in these systems (2). The waste enters the tank near the top. There is a pair of baffles in the tank to keep the solids in the tank, preventing them from flowing out of the tank with liquids. Bacteria in the tank break down the solids as much as they can into a liquid form and this with the water leaves the tank on the other side of the baffles. The liquid then flows to a leaching field where the liquid enters the soil and is absorbed. If the bacteria cannot break the solids down, they will build up over time. If these solids are not removed by periodic pumping, the tank will allow solids to be washed out to the leaching field and begin to clog the soil. When the soil is clogged, the system stops working.

Septic systems fail for the following reasons:

1. Faulty design (leaching field that is too small).
2. Faulty installation (plugged lines or uneven grades).
3. Soil conditions (highly permeable soil or relatively impervious soil, less than 6 feet of unsaturated soil cover).
4. High water table less than 6 feet from the land surface.
5. Water overload.
6. Inadequate cleaning of the tank (should be pumped every 2–3 years).
7. Highly permeable soil.

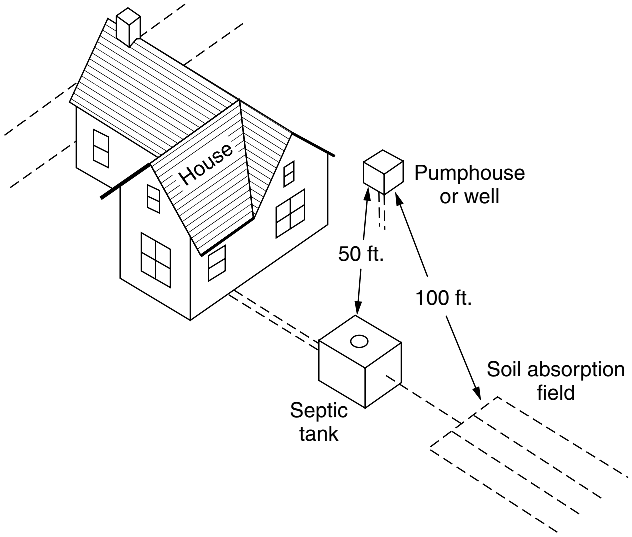

Figure 2. Setback distances (4).

It has been estimated that 25% of the U.S. population uses septic systems for treatment and disposal of their household sewage. Septic system technology is undergoing dramatic changes in efficiency and reduced contamination (2). The American Society of Testing and Material has prepared three standards for the treatment and disposal of on-site waste (D 5879-95, D 5921-96, D5925-96).

Bacterial and viral contamination from septic systems is the most common cause of drinking water contamination in the United States. The liquid effluent from septic systems follows the same path as precipitation moving into an unsaturated zone and aquifer. When the effluent reaches the water table, it moves downgradient to the point of discharge (lake, stream, wetland, and well). The location of the septic system in relation to the slope of the land surface is important because septic tank discharge follows the slope of the land surface. Wells downslope from septic tanks are subject to contamination (3).

The septic tank effluent can contain bacteria and also toxic materials and other contaminants. Some of the contaminants adhere to the soil and aquifer material or travel with the water.

Many septic systems are found in small rural homes sites and are commonly located on small narrow lots along a feeder highway. An increasing number of states are zoning suburban areas to limit the density of houses using septic tanks (4). Community sewer systems are used in some areas to substitute for septic systems.

Some banks require the prospective seller of rural property to provide proof of a bacteria-free water supply. Some sellers chlorinate the water to destroy the bacteria in the well. The bacterial contamination is in the aquifer, so this treatment lasts only a short time (5). The homeowner should have the well water analyzed at least once year for bacteria.

A buyer of rural property should determine the location of the well and the septic system. The buyer should also determine the age, maintenance, distance to the drinking water supply well, and depth to water at the septic system site. A wet area, lush vegetation over the leaching field, or odor of sewage is cause for further investigation. A water sample from the well at a septic system site should be obtained and analyzed for fecal coliform bacteria.

BIBLIOGRAPHY

1. Wilson, W. and Moore, J.E. (1998). *Glossary of Hydrology*. American Geological Institute, Denver, CO.
2. Bedinger, M.S. Fleming, J.S., and Johnson, A.I. (Eds.). (1997). *Site Characterization and Design of On-Site Septic Systems*. ASTM, Philadelphia, PA.
3. Waller, R.W. (1959). *Ground Water and the Rural Homeowner*. U.S. Geological Survey, General Interest Publication, Washington, DC.
4. Wyoming Department of Environmental Regulations (DER). (1998). *Septic Systems: Rural Wellhead Protection Fact Sheet*. Cheyenne, WY.
5. Covin, C. (1999). *Handbook of Groundwater Quality Protection for Farmers*. CSIR, South Africa.

DOMESTIC SOLAR WATER HEATERS

Mervyn Smyth
Centre for Sustainable Technologies
Newtownabbey, United Kingdom

INTRODUCTION

Solar water heating systems convert solar radiation into useable thermal energy in the form of hot water. Domestic solar water heaters can provide households with a large proportion of their hot water needs while reducing the amount of conventional fuel used and hence reducing home energy costs. The amount of hot water produced will depend on the type and size of the system, the climate, and location for solar access. Over the years, a variety of system designs have been developed and tested to meet specific consumer needs and environmental conditions. The following article is a brief description of the many types of system in common use today.

The first solar water heaters consisted of exposed tanks of water left out to warm in the sun. Used on a few farms and ranches in the Southwestern United States in the late 1800s, they were reportedly capable of producing water hot enough for showering by the late afternoon on clear days (1). The first solar water heater, manufactured commercially under the trade name Climax Solar-Water Heater, was patented in 1891 (2). Figure 1 illustrates a reproduction of an advertisement for the Climax Solar-Water Heater. This water heater could be used from April to October in the State of Maryland in the eastern United States. It claimed to produce water hotter than 38 °C on sunny days even during early spring and in late autumn when daytime temperatures sometimes approached freezing.

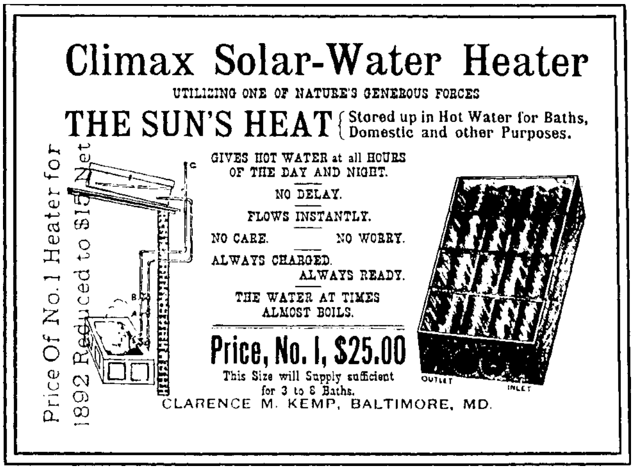

Climax Solar-Water Heater

UTILIZING ONE OF NATURE'S GENEROUS FORCES

THE SUN'S HEAT { Stored up in Hot Water for Baths, Domestic and other Purposes.

GIVES HOT WATER at all HOURS OF THE DAY AND NIGHT.

NO DELAY.

FLOWS INSTANTLY.

NO CARE. NO WORRY.

ALWAYS CHARGED. ALWAYS READY.

THE WATER AT TIMES ALMOST BOILS.

Price, No. 1, $25.00

This Size will Supply sufficient for 3 to 6 Baths.

CLARENCE M. KEMP, BALTIMORE, MD.

Price Of No.1 Heater for 1892 Reduced to $15 Net

Figure 1. Advertisement for the climax solar-water heater, 1892 (1).

Domestic solar water heaters can be categorized as being either active or passive and can be further grouped according to the configuration of the main solar water heating components: integral or distributed. Integrated systems combine the collector and storage functions in a single unit, whereas distributed systems have a separate solar collector and hot water store connected by a piping network. Distributed systems can be either active or passive. In active systems, a pump circulates the transfer fluid between the collector and the store. Integrated systems are almost always passive as they do not require external power.

THE INTEGRATED COLLECTOR/STORAGE SOLAR WATER HEATER

The most basic of solar water heaters is the integrated collector/storage solar water heating (ICSSWH) system or the integral passive solar water heater (IPSWH), commonly referred to as breadbox or batch water heaters. Kemp's early Climax Solar-Water Heater was an integrated system. A simplified diagram of a typical ICS solar water heating installation is shown in Fig. 2.

In its simplest form, the ICSSWH is a water tank painted black to absorb insolation (*in*cident *sol*ar radi*ation*). Variations consist of one or more tanks, painted black or coated with a selective absorbing surface, within a well-insulated box, possibly with reflectors and covered with single, double, or even triple layers of glass, plastic, or a combination of the two. Because of its simplicity, an integrated collector/storage system is easier to construct and install, which reduces maintenance and capital costs. In most climates, the large thermal mass of the store provides inherent resistance to freezing. However, the integrated unit has a significant problem because of its unique mode of operation.

The earliest systems suffered substantially from heat losses to ambient, especially at night and at noncollection periods, which meant no matter how effective the unit was in collecting solar energy, unless the hot water was fully withdrawn at the end of the collection period, losses

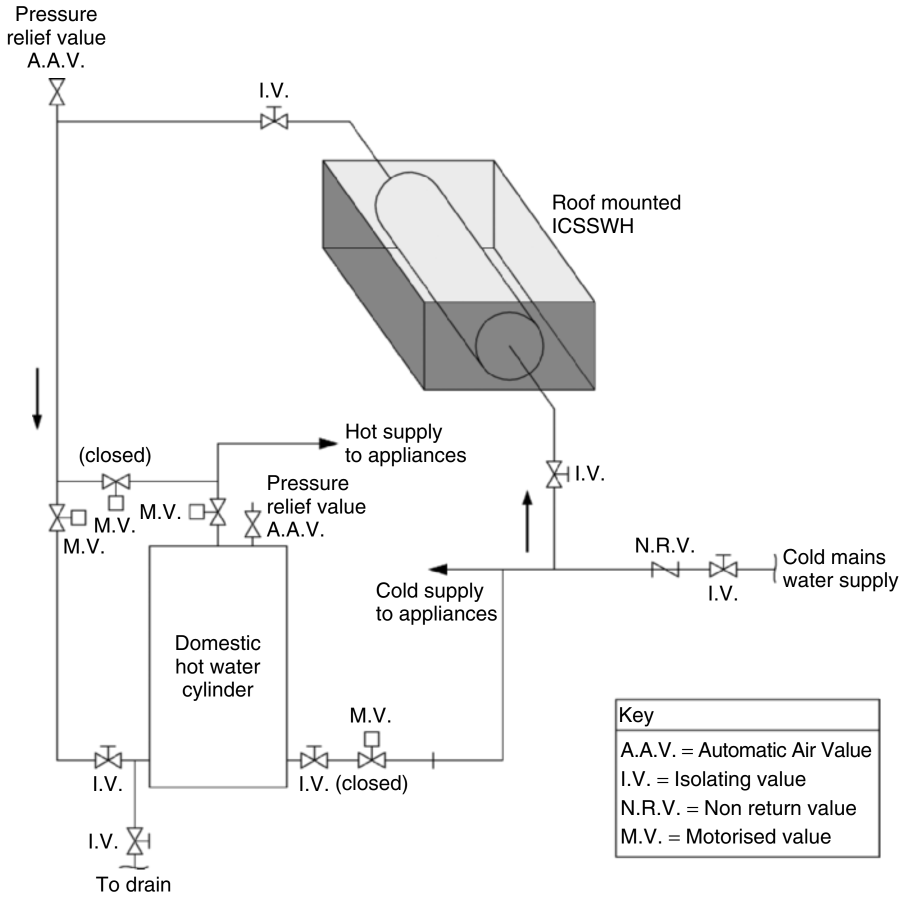

Figure 2. A simplified diagram of a typical roof-mounted ICS solar water heating installation.

to ambient led to only lukewarm water being available early the next day. This process reduced the overall solar fraction, which renders it less viable economically. Indeed this deficiency in the late nineteenth century led to the prominence of thermosyphon solar water heaters with diurnal heat storage to the detriment of the ICSSWH system. To overcome excessive heat loss and be in a position to compete with the more established distributed solar water heater systems, the ICSSWH design has had to evolve and incorporate new and novel methods of improving performance.

DISTRIBUTED SOLAR WATER HEATERS

Distributed systems consist of a separate solar collector and water store, with pipes connecting the collector(s) to and from store(s). As previously mentioned, these systems can be either active or passive, with the active system using an electric pump, and the passive system relying on buoyancy forces in the form of thermosiphonic action. Active systems also require more valves and control systems, which tend to make them more expensive than passive systems but generally more efficient. Figure 3 shows a simplified diagram of a typical roof-mounted distributed (flat-plate) solar water heating installation. Active systems are often easier to retrofit than are passive systems because their storage vessels do not need to be installed above or close to the collectors. In addition, a photovoltaic panel could power the pump, which results in stand-alone, proportional pump operation with reduced running costs.

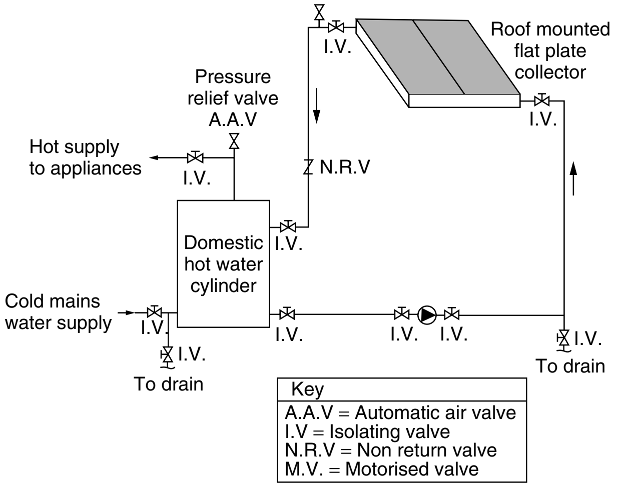

Figure 3. A simplified diagram of a typical roof-mounted distributed (flat-plate) solar water heating installation.

Distributed solar water heaters can also be characterized as being direct (open loop) or indirect (closed loop). A direct system circulates incoming mains water through the collector and into the tank, whereas an indirect system transfers collected thermal energy via a heat exchanger to the domestic water. Indirect systems usually contain an aqueous antifreeze solution that flows through the heat exchanger immersed in the hot water store to provide protection from freezing. This process, however, results in reduced collection efficiencies over the direct system through lower specific heat capacities and losses during the heat exchange process.

Active Direct Systems

Active direct systems use pumps to circulate incoming mains water through the collector and back into the tank. This design is efficient and reduces operating costs but is not appropriate where water is hard or acidic because of scale buildup and corrosion. However, direct active systems are popular in regions that do not experience freezing temperatures (Fig. 4).

Active Indirect Systems

Active indirect systems pump the heat-transfer fluid (usually a glycol-water antifreeze mixture) through the collector and a heat exchanger transfers the heat from the fluid to the water that is stored in the tank. Heat exchangers can be double-walled vessels or have twin coil arrangements. Indirect glycol systems are popular in areas where temperatures regularly fall below zero because they offer good protection from freezing. However, antifreeze systems are more expensive to purchase and install and require regular checking and maintenance.

Indirect drainback systems do not use antifreeze mixtures, but they use pumped water as the heat-transfer fluid in the collector loop. When freezing conditions prevail or the system is not in use, the pump is switched off and the water in the collector is drained out, thus providing protection from freezing. The collector installation and plumbing arrangement must be carefully positioned to allow complete drainage and the pump must have sufficient head pressure to pump the water up to the collector each time the pump starts.

Thermosiphon Systems

A thermosiphon system relies on warm water rising, a phenomenon known as natural convection or buoyancy forces, to circulate water to and from the collector and tank. In this type of installation, the tank must be located above the collector. As water in the collector heats, it becomes less dense and naturally rises into the tank above. Meanwhile, cooler water in the tank flows downward into the collector, thus causing circulation throughout the system. Some forms of thermosiphon solar water heaters can be described as being compact. Compact systems are close-coupled thermosyphon flat-plate or evacuated-tube collector units fabricated and installed as a single item as opposed to a separate collector, store, and pipework. Thermosiphon systems are much cheaper than are active systems as no pump or controller is required and are ideal where a low-cost solar heater is required such as holiday houses and cabins, or countries where low-cost solar heating is required.

SOLAR WATER HEATING COLLECTORS

Basically three types of domestic solar collector are in common use today: flat-plate, evacuated-tube, and concentrating.

Flat-plate Solar Collector

The flat-plate system consists of a "flat" absorber panel through which water or conducting fluid passes. The panel may be of formed channels in a sandwich format or may be pipes connected to expanded absorber plates. Most

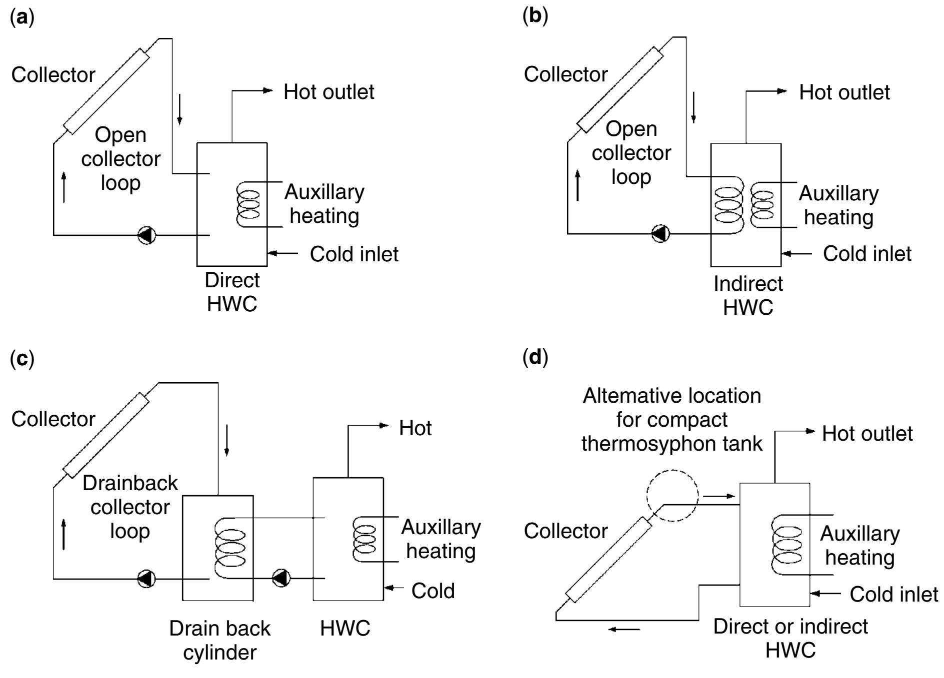

Figure 4. Schematic detail of common distributed solar water heating configurations: (**a**) active direct system, (**b**) active indirect system, (**c**) indirect drainback system, and (**d**) thermosiphon system.

absorbers are covered with a selective coating to improve solar radiation absorption and reduce long-wave radiative heat loss. As the fluid flows adjacent to the heated surface, it is heated. The absorber is mounted in an insulated, weatherproof unit, and the exposed collector aperture is covered with one or more transparent or translucent covers. The make-up of a typical flat-plate solar collector is shown in Fig. 5.

Evacuated-tube Collector

Evacuated-tube collectors are made up of rows of parallel, glass tubes, linked to a common flow (and return) manifold depending on the collector installed. Two types of evacuated-tube collector exist: glass/glass or metal/glass. The glass/glass collector consists of two concentric glass tubes. The inner tube is covered with a selective coating to improve solar radiation absorption and reduce long-wave radiative heat loss. The transparent outer tube forms a space between the two tubes that is evacuated to eliminate conductive and convective heat loss. The metal/glass collector consists of a copper plate attached to a heat pipe or water pipe mounted within a single evacuated glass tube. Again the absorber is coated with a selective coating to improve the collection performance. Figures 6 and 7 illustrate some common evacuated-tube collectors.

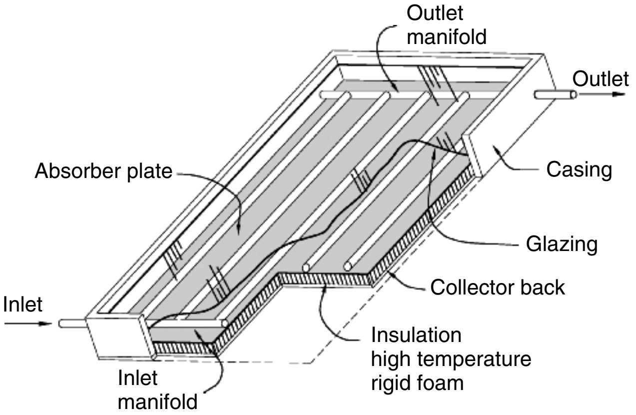

Figure 5. A typical flat-plate solar collector.

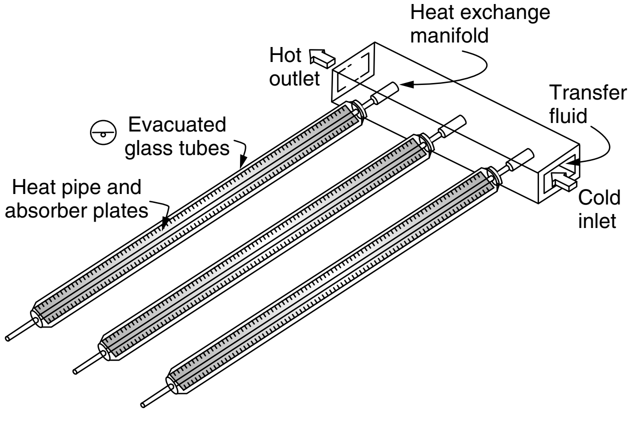

Figure 6. Metal/glass heat pipe evacuated-tube solar collector.

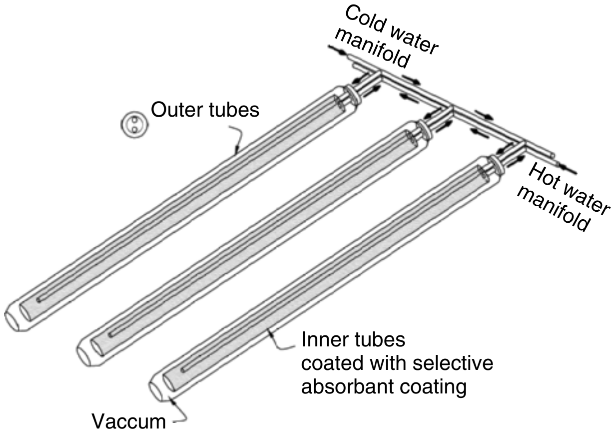

Figure 7. Metal/glass water pipe evacuated-tube solar collector.

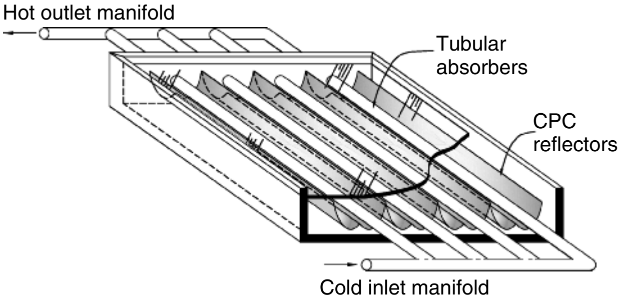

Figure 8. Diagram of a compound parabolic concentrating collector.

Concentrating collector

To increase the insolation on the absorber surface over that incident at the collector aperture, reflectors are employed in solar water heating systems. Concentrating reflectors can obtain higher temperatures on the absorbing surface than can those achievable by a flat absorber, and as the absorbing surface area is reduced relative to that of the aperture, a reduction in the overall heat loss from the system occurs, hence an improved thermal efficiency.

Internal reflectors are contained within the unit enclosure, whereas external reflectors are located outside the sealed casing. Reflecting concentrator designs for low-to-medium concentrations can be flat or curved, line-axis or line-focus (circular, parabolic or compound parabolic) reflectors, symmetrical or asymmetrical. The concentrating collector used for domestic applications usually incorporates a concentrating reflector in the form of parabolic trough or compound parabolic concentrating (CPC) collector, using highly reflective surfaces to concentrate the insolation onto the absorber. Most absorbers are tubular, although not exclusively. Figure 8 illustrates a CPC collector.

BIBLIOGRAPHY

1. Butti, K. and Perlin, J., (1981). *A Golden Thread*. Marion Boyars Publishers Ltd, London, UK.
2. Kemp, C.M. (1991). U.S. Patent 451384, April 28.

READING LIST

Duffie, J.A. and Beckman, W.A. (1991). *Solar Engineering of Thermal Processes*. John Wiley & Sons, New York.

Norton, B. (1992). *Solar Energy Thermal Technology*. Springer-Verlag London Ltd, London, UK.

HOUSEHOLD DRINKING WATER TREATMENT AND SAFE STORAGE

SUSAN MURCOTT
Massachusetts Institute of Technology
Cambridge, Massachusetts

Household water treatment is the decentralized treatment of drinking water in the home and safe storage is the protection of drinking water in specially designated household storage vessels prior to use. A safe water storage vessel is typically comprised of a container with a narrow mouth to prevent contact with potentially dipped cups or dirty hands, a lid, a spigot to access water and a flat base for easy water extraction. In many cultures and regions, household drinking water treatment and storage has been women's work based on traditional practices stretching back for millennia. For example, an ancient Indian medical text, the *Susruta Samhita*, compiled over several centuries and reaching its present form in about A.D. 300, includes the prescribed water treatment and handling practices as follows:

> Heat contaminated water by boiling on fire, heating in the sun, by dipping hot copper into it seven times, cooling in an earthen vessel and also scenting it with flowers of nagkesara, campaka, utpala, patala, etc. (Book 1, Chapter 45, Verse 12)

Moreover, home storage in various types of containers, including skin bags, ostrich eggs, vessels of wood, ceramic, metal, glass, or stone has been a traditional practice for hundreds or even thousands of years. Thus the twin concepts of household drinking water treatment and safe storage are not new. But there are new developments arising from a global need (Figs. 1–3).

Currently, about 50% of people worldwide are supplied with household connections that provide drinking water on tap in their homes. Sufficient, safe, acceptable, physically accessible, and affordable water for all is a fundamental human right essential to life and dignity. Tapped water for all is the long-term goal, but even among those with tap water today, the drinking water is not always considered safe, in terms of its water quality. In homes with a tapped water supply, household treatment devices typically provide a final "extra" step that begins with a well-protected source and includes a treatment process provided by a centralized water treatment system, administered by a municipal authority or private entity. In these cases, the purpose of the household treatment step is typically to improve the aesthetics of the water (e.g., chlorine odor or taste, hardness) and/or to remove certain harmful contaminants, including possible organic (e.g., benzene, toluene), inorganic (e.g., cadmium, lead), or microbiological (e.g., *Cryptosporidium, Giardia*) substances. These household drinking water

Figure 1. Ceramic filter—Nepal.

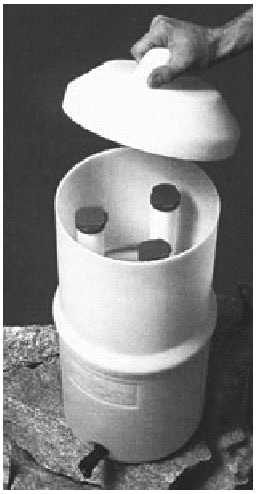

Figure 2. Ceramic candle filter by Katadyne, Switzerland.

Figure 3. Ceramic filter "filtron"—Ghana.

treatment devices take two forms—*point-of-entry* or *point-of-use*—depending on whether the device is installed at the point where the water main enters the home or whether the treatment unit is attached to or placed beside the kitchen faucet (i.e., at the point where drinking water is withdrawn). In such cases, household drinking water treatment in industrialized countries and regions is used to provide an additional barrier of safety to a water supply that has already received treatment upstream or is of known high quality.

Homes lacking a tapped drinking water supply via a household connection or lacking another form of "improved" water supply such as a public standpipe, a borehole, a protected dug well, a protected spring, or rainwater collection are more likely to bear the burden of water-related illnesses:

- 3.4 million deaths are water-related;
- 1.4 million children die annually of diarrhea, making this the third highest cause of illness and the sixth highest cause of mortality globally;
- 1.5–2 billion people are affected by intestinal parasites;
- 1.1 billion people lack access to safe drinking water;
- 2.6 billion people are without access to basic sanitation (1).

These combined conditions can be addressed and an improved quality of life can be realized by applying the same principles that brought about the industrialized world public health miracle of the nineteenth and twentieth centuries—a treated drinking water supply, sanitation, and good hygiene practices—to households globally.

Between 1990 and 2002, 1.1 billion more people worldwide gained access to improved water supplies. Yet that same number—1.1 billion or about one in six people—still lack access to improved water in 2004. Most of these people live in rural areas and urban and peri-urban slums. Their water needs are a focal point of international efforts to provide safe drinking water (2). For these 1.1 billion people, household water treatment and safe storage is not an additional barrier, post-treatment, as it is for those who purchase and use point-of-entry or point-of-use systems, but instead it may be their main barrier in the prevention of water-related illness. And these systems work! "There is now conclusive evidence that simple, acceptable, low-cost interventions at the household and community level are capable of dramatically improving the microbial quality of household stored water and reducing attendant risks of diarrheal disease and death" (3).

Moreover, we know that household drinking water treatment and safe storage, access to sanitation, and hygienic behavior are all interrelated activities. The combination of all these three main interventions will maximize health benefits to all. Household drinking water treatment and safe storage is one essential technology with a special role to play for households lacking a safe water supply. It was with this understanding that the World Health Organization formed the International Network to Promote Household Drinking Water Treatment and Safe Storage, a public–private partnership announced at the Kyoto World Water Forum in March 2003.

While we know that household water treatment and safe storage has been practiced locally and regionally, recognition of the role that household water treatment and safe storage can play globally in securing safe drinking water is a recent development dating to the 1990s. Research and development have been a process of adapting traditional wisdom and best engineering and public health practices, applied in settings that necessitate simple, low maintenance designs, use of local materials, applications under demanding local conditions, social acceptability, and economic sustainability. Research on cost effectiveness indicates that these household water treatment and safe storage practices can avert much of the burden associated with diarrheal disease at low cost(4).

Some of the treatment processes for household drinking water treatment and safe storage currently under investigation and/or in early stages of implementation include (see Figs. 4–6):

Sedimentation
Mechanical and/or biological filtration

Figure 4. Chlorine solution for household disinfection.

Figure 5. Safe water storage container.

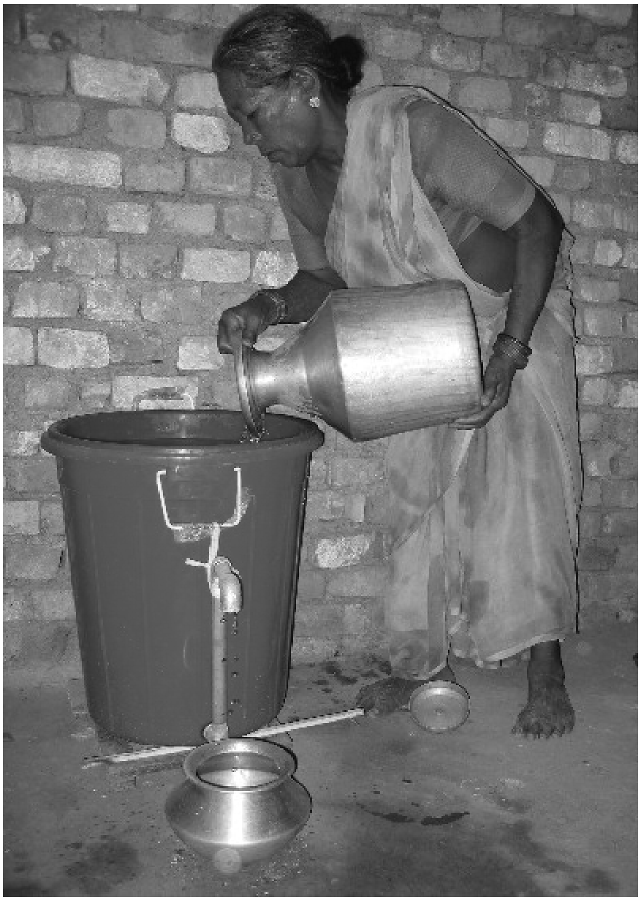

Figure 6. Household arsenic filter (Kanchan arsenic filter—Nepal).

Cloth filters
Ceramic water filters
Intermittent household slow sand filters

Coagulation/flocculation

Metal salts (e.g., alum, ferric chloride, ferric sulfate)
Natural polymers
Mixtures of coagulants/flocculants, weighting agents, calcium hypochlorite

Adsorption

Arsenic remediation household systems
Fluoride remediation household systems

Ion exchange processes
Membrane/reverse osmosis processes
Oxidation processes
Disinfection

Chlorine and the safe water system
Solar UV or UV lamp disinfection
Heat disinfection or pasteurization

Distillation
Combined (multiple barrier) household treatment systems

Sedimentation + solar UV disinfection
Pretreatment filters (strung-wound + granular activated carbon filter + chlorine disinfection)
Pretreatment cloth + sand + ceramic candles with colloidal silver
Coagulation/flocculation + filtration + chlorine disinfection
Other combinations

Beyond inactivation and/or removal of microbiological contamination—which is the major concern for those lacking access to safe drinking water—appropriately designed household drinking water treatment can effectively remove physical substances (e.g., turbidity) and/or toxic chemicals (e.g., arsenic, fluoride, pesticides) as well as microbiological contamination by one or several of the processes listed above.

On every continent, there are promising household drinking water treatment and safe water storage options available. In the decade to come, we will witness new research, innovation, and scale-up of these systems from hundreds to thousands to millions to meet the enormous global need for clean, safe drinking water.

BIBLIOGRAPHY

1. WHO/UNICEF Joint Monitoring Programme on Water Supply and Sanitation (2004). *Meeting the MDG Drinking Water and Sanitation Target: A Mid-Term Assessment of Progress*. Available at http://www.who.int/water_sanitation_health/monitoring/jmp2004/en/.
2. Millennium Development Goal #7: to reduce by half the proportion of people lacking access to safe water and sanitation by 2015. Available at www.developmentgoals.org.
3. Sobsey, M. (2003). Managing water in the home: accelerated health gains from improved water supply. Available at http://www.who.int/water_sanitation_health/dwq/wsh0207/en/index.htm.

4. Hutton, G. and Haller, L. (2004). Evaluation of the costs and benefits of water and sanitation improvements at the global level. In: *Water, Sanitation and Health Protection of the Human Environment*. World Health Organization, Geneva, WHO/SDE/WSH/04.04. Available at http://www.who.int/water_sanitation_health.

VIRUS TRANSPORT IN THE SUBSURFACE

Ann Azadpour-Keeley
Jack W. Keeley
United States Environmental Protection Agency
Ada, Okaholma

BACKGROUND

The 1986 Safe Drinking Water Act (SDWA) amendments directed the EPA to develop national requirements for drinking water disinfection. The legislation required all public water supply systems to disinfect unless they fulfill criteria ensuring equivalent protection. To provide direction for the regulations associated with "acceptable" health risks to the public (4), the EPA established goals for maximum contaminant levels (MCLGs) of pathogenic microorganisms in drinking water, set or setting a level of zero for viruses (5,6).

On June 29, 1989, a Surface Water Treatment Rule (SWTR) was published addressing microbial contamination of drinking water from surface sources or from groundwater sources directly influenced by surface water, that had strict provisions for filtration and disinfection (5). On January 14, 2002, a SWTR was promulgated with special emphasis on the protozoan *Cryptosporidium* (7).

The development of a corresponding rule for groundwater, Ground Water Disinfection Rule (GWDR, later designated as the Groundwater Rule), to meet SDWA requirements began in 1987 and led to a published discussion piece (8) and a deadline for the GWDR proposal upon completion of the status of public health with respect to the microbial contamination of groundwater by conducting studies to generate a more careful nationwide picture of the problem. On May 10, 2000, "US EPA proposed to require a targeted risk-based regulatory strategy for all groundwater systems addressing risks through a multiple barrier approach that relies on five major components: periodic sanitary surveys of groundwater systems requiring the evaluation of eight elements and the identification of significant deficiencies; hydrogeological assessments to identify wells sensitive to fecal contamination; source water monitoring for systems drawing from sensitive wells without treatment or with other indications of risk; a requirement for correction of significant deficiencies and fecal contamination (by eliminating the source of contamination, correcting the significant deficiency, providing an alternative source water, or providing a treatment which achieves at least 99.99 percent (4-log) inactivation or removal of viruses), and compliance monitoring to insure disinfection treatment is reliably operated where it is used (9). The Ground Water Rule will be issued in 2005.

INTRODUCTION

More than 97% of all freshwater on the earth is groundwater. Of more than 100 million Americans who rely on groundwater as their principal source of potable water, over 88 million are served by community water systems and 20 million by noncommunity water systems (9). Historically, groundwater has been considered a safe source of drinking water which required no treatment. It has long been believed that this valuable resource was protected from surface contamination because the upper soil mantle removed pollutants during percolation. It was also believed that, even if contaminated, groundwater would be purified through adsorption processes and metabolism of indigenous aquifer microflora.

In the United States alone, the estimated annual number of reported illnesses resulting from contact with waterborne pathogens was as low as one million and as high as seven million between 1971 and 1982, and 51% of all waterborne disease outbreaks due to the consumption of contaminated groundwater (1). It is estimated that approximately 20–25% of U.S. groundwater sources are contaminated with microbial pathogens, including more than 100 types of viruses. A literature review by Craun (2) indicated that approximately one-half of the surface water and groundwater sources tested contained enteric viruses. Even 9% of conventionally treated drinking water (coagulation, sedimentation, filtration, postfiltration disinfection using chlorine/ozone) tested positive for enteric viruses.

Although water-transmitted human pathogens include various bacteria, protozoa, helminths, and viruses, agents of major threat to human health are pathogenic protozoa (*Cryptosporidium* and *Giardia*) and enteroviruses. Despite ample information regarding the fate of viruses in the subsurface, research on the persistency of pathogenic protozoa through passage in soil and groundwater is just now emerging. In the past, it was generally believed that pathogenic protozoa are confined to surface water. Contrary to that expectation, recent monitoring results from 463 groundwater samples collected at 199 sites in 23 of the 48 contiguous states suggested that up to 50% of the groundwater sites were positive for *Cryptosporidium, Giardia*, or both, depending on the parasite and the type of groundwater source (vertical wells, springs, infiltration galleries, and horizontal wells) (3).

Viruses are small obligate intracellular parasites that infect and sometimes cause a variety of diseases in animals, plants, bacteria, fungi, and algae. Viruses are colloidal particles, negatively charged at high pH (pH >7), ranging in size from 20 to 350 nm. The smallest unit of a mature virus is composed of a core of nucleic acid (RNA or DNA) surrounded by a protein coat. Due to this unique feature of viral structure and colloidal physicochemical properties, the transport of viruses in soil and ground water can act with a combination of characteristics ranging from those of solutes, colloids, and microorganisms.

Enteroviruses are a particularly endemic class of waterborne microorganisms that cause a number of ubiquitous illnesses, including diarrhea, gastroenteritis, and meningitis, to name only a few. Included in this group are poliovirus, hepatitis type A (HAV), Coxsackie virus A and B, and rotavirus. Although gastroenteritis is the most common disease resulting from these microorganisms, other associated illnesses include hepatitis, typhoid fever, mycobacteriosis, pneumonia, and dermatitis (10).

SOURCES OF VIRUSES

A number of avenues are available for the introduction of viruses to the subsurface, including land disposal of untreated and treated wastewater, land spreading of sludge, septic tanks and sewer lines, and landfill leachates.

Among these, septic systems may pose a significant chemical as well as biological threat to surface and groundwaters. One trillion gallons of septic-tank waste are released into the subsurface annually. Although phosphate and bacteria are ordinarily removed by soil, nitrate and viruses may escape these processes and move through the soil into the groundwater. The presence of viral particles is even more significant in the light of studies that indicate they are not necessarily inactivated in septic tanks and may move into the groundwater where they may survive for long periods of time.

It is a general consensus that the transport of pathogens in the subsurface depends on the extent of their retention on soil particles and their survival. Among the major factors that affect viral transport characteristics in the subsurface are temperature, microbial activity, moisture content, and pH. Among all the factors, temperature appears to be the only well-defined parameter that causes a predictable effect on viral survival. A direct relationship between a rise in temperature and viral inactivation rates ($K = \log$ inactivated/h) among various viruses has been suggested. Badawy et al. (11) stated that during the winter (4–10 °C), viral inactivation rates for coliphage, poliovirus, and rotavirus were 0.17, 0.06, and 0.10 per hour, respectively. Whereas, during the summer (36–41 °C), the inactivation rates for MS-2, poliovirus, and rotavirus were 0.45, 0.37, and 0.20 per hour, respectively. It should be pointed out that this information is based on ambient air. A more direct comparison would be correlation with temperatures in the subsurface. In this regard, the inactivation rates for enteroviruses are 0.06 (10–15 °C), 0.08 (15–20 °C), and 0.19 (20–25 °C). This worker also indicated that viruses may remain viable for 3 to 5 weeks on crops irrigated with sewage effluent, polio and Coxsackie virus up to 4 months on vegetables during commercial and household storage; and up to 30 days on vegetables stored at 4 °C.

Microbial ecology may also play an important role in the inactivation of waterborne viruses. For example, microbial activity could affect viral survival by the action of proteolytic enzymes of some bacteria and protozoa in destroying the viral capsid protein.

As discussed earlier, viral transport through porous media is controlled by sorption and by inactivation. However, adsorption of viruses to soil should not be confused with their inactivation because adsorption is not permanent and can be reversed by the ionic characteristics of percolating water. Viruses can remain infective after a travel distance of 67 meters vertically and 408 meters horizontally (12). The various forces involved in attaching viruses to soil particles include hydrogen bonding, electrostatic attraction and repulsion, van der Wals forces, and covalent ionic interaction.

EFFECT OF HYDROGEOLOGIC SETTINGS ON VIRAL MOVEMENT

The concentration and loading of viruses and the hydrogeologic setting through which they move will control the potential for viral migration to wells to a much greater extent than biological survivability. A hydrogeologic setting often consists of a soil underlain by unconsolidated deposits of sand, silt, and clay mixtures over rock. The setting further incorporates unsaturated and saturated zones.

All other factors being equal, the persistence of viruses at a well or other source of water is most likely where saturated flow transports large concentrations of the particles along short flow paths through media that contribute little to attenuation. Although the interrelated processes that control viral movement and persistence in the subsurface are not completely understood, some of the major hydrogeological factors that can be used to evaluate the potential for viral presence in groundwater wells include

- transport mechanisms (unsaturated versus saturated flow conditions),
- type of media through which the virus will travel (clays versus sands versus fractured media),
- length of the flow path to the extraction point (well), and
- time of travel.

Hydrogeologic settings that have shallow water tables are more susceptible to viral transport. Viruses are attenuated or immobilized by processes such as dessication, microbial activity, and stagnation. Further, viruses commonly bind to soil particles, fine-grained materials, and organic matter. The lower transport velocities associated with unsaturated conditions (e.g., move, stop, move cycle) allow these processes more time to occur. If viruses are introduced directly into the water table (such as from leaching tile fields associated with on-site sewage disposal) or if the volume of contaminants can maintain saturated flow conditions (such as in some artificial recharge situations), the potential for contamination increases. Where the viral concentration is high, the probability of contaminant migration increases regardless of the hydrogeologic setting. Therefore, in hydrogeologic settings that have deeper water tables and where contaminants are not introduced into the aquifer through saturated flow conditions, viruses are much less likely to survive transport to a well.

Hydrogeologic settings that have interconnected fractures or large interconnected void spaces that lack fine-grained materials have a greater potential for viral transport and well contamination. Karst aquifers, fractured bedrock, and gravel aquifers have been identified in the proposed GWDR as sensitive hydrogeologic settings (9).

In these settings, fractures and large void spaces allow rapid transport through the aquifer, thereby reducing the amount of time and particulate contact available for attenuation. Potential interaction with rock walls along fractures is reduced, and contact with fine-grained materials for potential sorption sites is minimal.

Similar to fractured rock aquifers, gravel aquifers that have only a small fine-grained fraction have little potential for viral sorption. However, as the amount of fine-grained material increases, effective grain size decreases, the potential for sorption increases, and travel times decrease. Finer grained aquifers and aquifers where void spaces are less interconnected or smaller are therefore less likely to transport viruses significant distances.

The potential for physical viral removal by filtration also appears to increase as grain size becomes smaller, although the filtration processes are not well understood due to their size. However, filtration of bacteria, which are larger than viruses, it has been shown is an effective removal mechanism.

Hydrogeologic settings, where fractures are not as interconnected or where more tortuous flow paths must be followed to reach a well, also allow greater viral removal. For example, in many rock aquifers, groundwater flow follows bedding planes that may result in an elongated indirect pathway to a well. In other rock aquifers, flow must travel around and through cemented portions of the matrix thereby increasing the flow path. Similarly, sand and gravel aquifers that have fine-grained materials in the matrix will have less direct flow paths as the water flows around the finer grained materials. Generally, it can be stated that tortuosity increases the length of the flow path and decreases the hydraulic conductivity, thus decreasing viral survival. Where finer grained materials are present or fractures are less interconnected, flow paths are also longer, thereby offering some protection to wells in more permeable units.

Hydrogeologic settings where time of travel is short have a greater potential for viral contamination. Where less permeable units (called aquitards) restrict or reduce vertical flow to underlying aquifers, time of travel is increased. Although inactivation rates, it has been shown, are extremely variable, time is a major factor affecting virus viability.

Due to the importance of hydrogeologic settings, the Proposed Ground Water Rule thoroughly addresses this issue to identify wells that are sensitive to fecal contamination. A component of the Proposed Ground Water Rule requires states to perform hydrogeologic assessments for the systems that distribute groundwater that are not disinfected (source waters that are not treated to provide 99.99% removal or inactivation of viruses). The states are required to identify sensitive hydrogeologic settings and to monitor for indicators of fecal contamination from sensitive hydrogeologic settings (see Ref. 9 for the complete proposed strategy).

VIRUS TRANSPORT MODELING

One method of addressing regulations for viral exposure, such as groundwater disinfection, the application of liquid and solid waste to the land, and wellhead protection zones, is using predictive viral transport models. Like most predictive modeling efforts, the results depend on the conceptual basis of the model as well as the quality and availability of input data. Clearly, a thorough understanding of the processes and parameters of viral transport are essential elements in their application.

Some of the more important subsurface viral transport factors include, soil water content and temperature, sorption and desorption, pH, salt content, organic content of the soil and groundwater matrix, virus type and activity, and hydraulic stresses. Berger (14) indicated that the inactivation rate of viruses is probably the single most important parameter governing viral fate and transport in ground water.

Some of the existing models require only a few of these parameters which limit their use to screening level activities, whereas others require input information which is rarely available at field scale and is usually applied in a research setting. One limitation of most models is that they have been developed for use in the saturated zone. It has been shown, however, that the potential for viral removal is greater in the unsaturated zone than in groundwater.

Setback Distances

Traditionally, state and county regulators have established fixed setback distances for all geologic settings in their jurisdictions. For example, the distance between a septic tank and a private well would, in many instances, be as little as 50 feet and would apply for tight clays as well as fractured rock. It would apply to areas where the water table is near the surface as well as at considerable depth. As discussed in this document, the travel time or transport distance of viral particles depends on a number of factors, including moisture content, geologic setting, type and depth of the soil overburden, and source loading, to name only a few.

Frequently, guidelines established as minimum distances became so standard that a well was often positioned precisely 50 feet from the septic tank. In the survey conducted as part of the proposed Ground-Water Treatment Rule, setback distances were quite variable (9). Some of the distances were presumably based on scientific principles, while others were holdovers from past practices.

One approach in determining setback distances for septic tanks in wellhead protection areas and bank filtration sites is to determine travel times using groundwater flow characteristics. This approach has been implemented in the Federal Republic of Germany, for example, where three concentric zones protect each drinking-water well. The zone immediately surrounding the well is faced with the most restrictive regulations which are founded on the belief that 50-day residence time is adequate for inactivation of any pathogen in contaminated water. However, a comprehensive study by Matthess et al. (15), involving the evaluation of the "50-day zone," concluded that the reduction of viruses by 7 log units (current regulations) requires a much longer residence time. Matthess et al. indicated that a reduction of 7 log units occurred in about 270 days (Haltern and

Segeberger Forest) in one study and about 160–170 days (Dornach) would be required, according to another study.

Another approach to this important issue is to consider the vulnerability to viral transport in the subsurface of portions of a state or county or of individual aquifers. Although there are a number of approaches to rank vulnerability, DRASTIC is one assessment methodology that uses hydrogeologic setting descriptions and a numerical ranking system to evaluate groundwater pollution potential (13). DRASTIC assumes that a potential contaminant will be introduced at the ground surface, have the mobility of water, and be flushed toward the aquifer by infiltration. Using existing information on variable scales, the methodology was designed to evaluate areas of 100 acres or larger.

DRASTIC is an acronym representing seven reasonably available factors that are used to develop a numerical score. They are **D**epth to water, net **R**echarge, **A**quifer media, **S**oil, **T**opography (slope), **I**mpact of the vadose zone media, and hydraulic **C**onductivity of the aquifer. DRASTIC uses a weighting system to create a relative pollution potential index that varies between 65 and 223; the higher numbers express greater vulnerability.

Although DRASTIC was not designed specifically to evaluate the movement of viruses in the subsurface, the major transport mechanisms and flow paths for viral transport are considered, and the flexibility of the systems' rating scheme allows many of these factors to be taken into account. For example, depth to water addresses saturated versus unsaturated flow conditions and their importance. Aquifer media, soil, and impact of the vadose zone media all are based on descriptive soil and rock terms that allow variation due to fracturing, grain size, attenuation mechanisms, and overall characteristics that affect flow. Topography addresses the tendency of viruses to be introduced into the subsurface or to be carried away by runoff. Hydraulic conductivity addresses the relative ease of a contaminant to move with the velocity of water through the aquifer.

Clearly, meaningful setback distances can be developed only by using scientific principles that allow the use of available knowledge. The establishment of setback distances from sources of viral contamination to points of extraction (wells) can be established using DRASTIC if both the hydrogeologic setting and sensitivity rankings are considered. For example, high pollution potential index signal the need for greater setback distances. However, the hydrogeologic factors that control viral movement must be evaluated within this context to establish reasonable numbers for setback distances. A matrix that incorporates the important DRASTIC factors can be used to establish setback distances that include the vulnerability concept. Setback distances must incorporate the knowledge of saturated flow, transport pathway length, transport velocities, media interaction, and potential attenuation mechanisms. These setback distances can be used on a regional scale but can be modified if site-specific information is available. The beauty of DRASTIC is that its rationale and sensitivity factors are easily displayed, so that it can be readily modified.

BIBLIOGRAPHY

1. CDC. (1991). Waterborne-disease outbreaks, 1989–1990 *MMWR* **40**(SS-3): 1–21.
2. Craun, G.F. (1989). Causes of waterborne outbreaks in the United States. *Water Sci Technol.* **24**: 17–20.
3. Hancock, C.M., Rose, J.B., and Callahan, M. (1998). Crypto and Giardia in US groundwater. *J. Am. Water Works Assoc.* **90**(3): 58–61.
4. Macler, B.A. (1996). Developing the ground water disinfection rule. *J. Am. Water Works Assoc.* **3**: 47–55.
5. U.S. EPA. (June 29, 1989). Drinking water; national primary drinking water regulations; filtration, disinfection; turbidity; *Giardia lamblia*, viruses, *Legionella*, and heterotrophic bacteria: final rule. *Fed. Reg.* **54**: 27486.
6. U.S. EPA. (June 29, 1989). Drinking water; national primary drinking water regulations; total coliforms (including fecal coliforms and *E. coli*): final rule. *Fed. Reg.* **54**: 27544.
7. U.S. EPA. (January 14, 2002). National primary drinking water regulations: long term 1 enhanced surface water treatment rule; final rule. *Fed. Reg.* **67**(9): 1844.
8. U.S. EPA. (1992). Draft Ground-Water Disinfection Rule. U.S. EPA Office of Ground Water and Drinking Water, Washington, DC, EPA 811/P-92-001.
9. U.S. EPA. (May 10, 2000). National primary drinking water regulations: ground water rule; proposed rules. *Fed. Reg.* **65**(91): 30202.
10. Bull, R.J., Gerba, C.P., and Trussell, R.R. (1990). Evaluation of the health risks associated with disinfection. *CRC Crit. Rev. Environ. Contr.* **20**: 77–113.
11. Badawy, A.S., Rose, J.B., and Gerba, C.P. (1990). Comparative survival of enteric viruses and coliphage on sewage irrigated grass. *J Environ Sci Health.* **A25**(8): 937–952.
12. Keswick, B.H. and Gerba, C.P. (1980). Viruses in groundwater. *Environ. Sci. Technol.* **14**(11): 1290–1297.
13. Aller, L., Bennett, T., Lehr, J.H., Petty, R.J., and Hackett, G. (1987). *DRASTIC: A Standardized System for Evaluating Ground-Water Pollution Potential Using Hydrogeologic Settings*. EPA/600/2-87/035.
14. Berger, P. (1994). Regulation related to groundwater contamination: The draft groundwater disinfection rule. In: *Groundwater Contamination and Control*. U. Zoller (Ed.). Marcel Dekker Inc., New York, NY.
15. Matthess, G., Pekdeger, A., Zoller, U., and Schroeter, J. (1988). Persistence and transport of bacteria and viruses in groundwater—a conceptual evaluation. *J. Contam. Hydrol.* **2**: 171–188.

WINDMILLS

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

Windmills are machines that convert the force of the wind into energy that is applicable to various tasks, like pumping water. Windmills that are used to pump water use the energy generated by these mills to turn the gears that propel the pump. These types of windmills, sometimes referred to as wind pumps, have been used for centuries and continue in use around the world.

A BRIEF HISTORY

The first documented use of windmills was in Persia around 500–900 A.D., however, it is widely accepted that they were invented in China more than 2000 years earlier (1). In these early windmills, vertical-axis systems were used for grinding grain and pumping water. In the vertical-axis system, the wind would have to hit the mill from a specific angle to get the desired effect; therefore, most of the area around the mill had to be shielded. Later, horizontal-axis systems were developed that proved to have greater structural efficiency. By tilting the blade to a certain degree, the system eliminated the dependence on the direction of the wind and harnessed the wind energy lost in the vertical-axis systems by the areas shielded. Horizontal systems were later used throughout Europe.

Windmill technology, it is believed was introduced in Europe in the eleventh century (2). Aiming to their lowlands, the Dutch set out to develop a more efficient windmill and in the process became the driving force behind wind-machine development. Although many innovations were made, one of the fundamental improvements was designing sails that allowed optimum aerodynamic lift. This improvement made the sails rotate faster, greatly increasing their efficiency and speed in completing the task at hand. Wind energy was applied to irrigation, pumping from local wells, and drainage pumping; this in turn made areas habitable and liberated workers from these labor-intensive jobs.

As Europeans sought to expand and colonize, they brought the windmill technology with them. The colonization of the Americas is a prime example. Without this technology, it would have been impossible for immigrants to settle in areas that lacked a constant water supply. Certain areas in Texas, for example, lacked the needed water supply to sustain life and allow cultivating of the land. The drive to expand into these areas helped to stimulate the need to refine the windmill to solve this problem. Daniel Halladay addressed this problem in 1854 by changing the European windmill so it could operate unattended and more efficiently (3). The changes that he made appealed to many companies and small communities without water systems by providing an inexpensive way to get the water they needed (Fig. 1).

Figure 1. Multi-sailed windmill.

By the early twentieth century, windmills were being mass-produced, and millions of them were being used around the world. However, in the 1930s, as other fuel resources such as oil were demanded, wind pumping systems were not as desirable because they were more expensive and not as reliable as these other fuels. This was true until the 1970s when a shortage of oil prompted communities to revisit the idea of wind energy.

Wind pumps are currently being used for crop irrigation, drinking water supply for communities, and even individual household water supply. Between 5,000 to 10,000 wind pumps are being installed worldwide each year, and the market for these is expected to increase as wind technology advances and becomes less expensive (4).

THE WIND PUMP: HOW IT WORKS

The wind pump has four main features the wind turbine, the tower, the actual pumping equipment, and the storage basin. These parts can be found in almost all wind pumps. They vary in design depending on the wind conditions in which they will be used. If the wrong design is used in strong wind conditions, the pump may move too fast and malfunction, and if the wind speed is low, it may not be able to function at all.

The horizontal-axis is the typical system used for wind turbines to pump water. The rotor does not have to follow any specific design but instead should be designed for the wind conditions in which it will be used. However, the pump tends to have more force when there are more blades. The next main part of the wind turbine is its transmission. The transmission of a wind turbine converts each rotation of the rotor into an up and down motion, driving the pump rod in and out of the well. The tail of the wind turbine is a piece that was added during the westward expansion in the United States. It allows the wind turbine to work without constant supervision by changing the direction of the rotor to keep it facing the wind.

A tower holds the wind turbine generally between 10 m and 15 m high (5). It holds up the wind turbine and stabilize its connection with the pumping parts of the machine. The tower has either a square or triangular base, and the pump rod that moves in and out of the well is positioned in the center of the tower. The well that the pump rod enters can either be a shallow hand-dug well or a deep-drilled well. If the latter is the case, the walls of the well should be lined with a water permeable material to prevent them from caving in. A pipe called the rising main lies in the center of the pump rod. Its function is to carry the water pumped up to the surface. The actual pump is at the bottom of the pump rod submerged in the water and is attached to the rising main. The pump fills with water during the downward motion of the pump rod and pushes the water into the rising main and up to the surface as the pump rod moves up.

Finally, all wind pumps should have a storage basin that can hold the excess water that is pumped. When the water moves up into the rising main, it is then redirected into the storage tank. This stored water is also essential to water pumping systems because wind energy cannot always be relied on, and having a surplus at hand is useful.

THE ADVANTAGES OF WIND PUMPS

There are many advantages to using wind energy for pumping water. First, wind pumps are environmentally friendly. As with the atmosphere, wind turbines will not contaminate the land, and in the case of water pumps, there is no chance the water will be contaminated as a result of a malfunctioning wind pump.

Wind turbines generally do not affect the wildlife that inhabits the area. Sheep, cattle, deer, and other wildlife are not bothered by the turbines, and in fact have been known to graze under them. The only argument that has been raised about this issue is the tendency for birds to collide with them. However, several studies suggest that the impact of the wind turbines on birds does not compare to that of other things, such as electrical lines and buildings.

Wind pumps are cost-efficient. After the initial cost of installation, the owner basically has an infinite source of energy for just the cost of maintenance. Whereas using a fossil fuel to work the pump would make the owner subject to its cost.

THE DISADVANTAGES OF WIND PUMPS

There are, however, a few setbacks in using wind energy for pumping water. Because the wind is not reliable, if water is needed on a day where there is little to no wind, obviously the wind pump will not be able to pump water. This problem can be solved by a storage basin, as explained earlier. Many people who inhabit areas around wind turbines complain about the noise that is created by the rotation of the rotor, but this generally applies only to many wind turbines grouped together. As the rotors are improved upon and become more aerodynamic, the noise level will decrease greatly.

BIBLIOGRAPHY

1. *Illustrated History of Wind Power Development*. (7 June 2001). http://relosnet.com/wind/early.html.
2. World Meteorological Organization. (1981). *Meteorological Aspects of the Utilization of Wind As An Energy Source*. No. 575, p. 44.
3. *The Handbook of Texas Online*. (6 June 2001). http://www.tsha.utexas.edu/handbook/online/articles/view/WW/aow1.html.
4. Wind Energy & Atmospheric Physics. *Review of Historical and Modern Utilization of Wind Power*. http://www.risoe.dk/vea-wind/history.htm.
5. Jenkins, N. et al. (1997). *Wind Energy Technology*. John Wiley & Sons, New York, pp. 116–117.

READING LIST

Andersen, P.D. (9 July 2001). Review of historical and modern utilization of wind power. *Wind Energy and Atmospheric Physics*. http://www.risoe.dk/vea-wind/history.htm.

British Wind Energy Association. (28 June 2001). *Wind Energy in Agriculture*. http://www.britishwindenergy.co.uk/you/agric.html.

Danish Wind Turbine Manufacturers Association. (28 June 2001). *21 Frequently Asked Questions About Wind Energy*. http://www.windpower.dk/faqs.htm.

Gipe, P. (1995). *Wind Energy Comes of Age*. John Wiley & Sons, New York.

Illustrated History of Wind Power Development. (7 June 2001). http://relosnet.com/wind/early.html.

Jenkins, N. and Walker, J.F. (1997). *Wind Energy Technology*. John Wiley & Sons, New York.

Larkin, D. (2000). *Mill: The History and Future of Naturally Powered Buildings*. Universe, New York.

Look Learn & Do. *A History of Windmills*. (7 June 2001). http://looklearnanddo.com/documents/history_winmills.html.

Windmills. *The Handbook of Texas Online*. (6 June 2001). http://www.tsha.utexas.edu/handbook/online/articles/view/WW/aow1.html.

World Meteorological Organization. (1981). *Meteorological Aspects of the Utilization of Wind as an Energy Source*, No. 575.

MIXING AND AGITATION IN WATER TREATMENT SYSTEMS

Paul Mavros
Aristotle University
Thessaloniki, Greece

Mixing is one of the primary processes involved in water and wastewater treatment. In this article, the state of the art in mixing and impeller design is presented.

INTRODUCTION

Stirring is provided in a wide variety of processes to blend constituents or to disperse one phase into another or several other phases. In a blend-type operation, the purpose is to obtain a homogeneous mixture, whereas in the dispersion process, the goals vary widely, depending also on the nature of the phases involved:

- in gas–liquid dispersions, gas is dispersed into fine bubbles which must be distributed as evenly as possible in the vessel to take part in a subsidiary process, for example, absorption and/or reaction with a dissolved component, as in water and wastewater treatment, or flotation of hydrophobic particles, among others;
- in solid–liquid distributions, it is necessary to provide the appropriate conditions for entraining all solid particles inside the bulk of the liquid, either from the bottom of the vessel or from the free surface of the liquid;
- in liquid–liquid dispersions, fine droplets of one of the liquids have to be dispersed inside the other liquid to produce an emulsion or for a polymerization, among others.

For each of these processes, a particular type of agitator is appropriate. These have evolved from the simple paddles used during the past centuries; modern flow visualization techniques (1) have helped in designing agitator blade shapes optimized for specific processes.

In the following sections, the main types of agitators are presented, according to the processes for which they are intended; this presentation is limited to turbulent flow, which is typical in water and wastewater treatment processes, and does not describe agitators designed for viscous liquids (anchors, gates, etc.).

TYPES OF IMPELLERS

An impeller is a pump; by its rotation, it draws liquid from its neighborhood and then ejects it at a relatively high speed. It is typically mounted on a shaft connected to a motor, and the shaft–impeller structure is inserted in the stirred tank either axisymmetrically or sideways. Close to the impeller blades, the rotation induces a tangential flow. Inside the bulk of the vessel, on the other hand, the action of the impeller induces flow circulation, which follows a pattern typical of the impeller type. Thus, we distinguish mainly radial and axial impellers, depending on the direction of the flow that emerges from the impeller-swept region.

Radial impellers eject a liquid stream radially. In a typical stirred vessel where the impeller is mounted on a shaft, is vertical and is usually centrally located, the ejected liquid flows from the edge of the impeller blades toward the vessel walls. There, it separates into two streams; one flows in the upper part of the vessel and one in the lower part of the vessel, thus forming two flow loops. The liquid from these two streams circulates in the upper and the lower parts of the vessel and eventually is drawn back into the agitator-swept region; two primary circulation loops are established inside the stirred tank.

Figure 1 presents some typical radial impellers. The Rushton turbine (RT; Fig. 1a) is one of the most widely used impellers due to its efficiency in gas–liquid and liquid–liquid mixing. Its construction is simple; usually it has six flat blades mounted on a flat disk. Figures 1b (SCABA 6SRGT turbine or Chemineer CD-6) and 1c (Chemineer BT-6) present two variants of the Rushton turbine, where blades have a parabolic shape, which is even more efficient than the RT, especially in dispersing gas inside a stirred vessel. Finally, the Narcissus (NS) impeller (2) produces an inverse radial flow; liquid is drawn in from its side and pumped out from its upper and lower parts.

Figure 2, which is a 2-D plot of composite radial and axial velocities U_{RZ}—which are obtained from the vector sum of the radial (U_R) and axial (U_Z) components of the local velocity vector—illustrates the typical radial flow patterns of the Rushton turbine—radial flow directed from the impeller toward the vessel walls—and of the Narcissus (NS)—radial flow directed toward the impeller. A similar double-loop circulatory flow pattern is induced by the SCABA turbine (3).

Axial impellers draw liquid mainly from one of their sides, top or bottom, and eject it from the opposite site; when liquid is ejected toward the bottom of the vessel, the impeller is said to work in the "down-pumping" mode, whereas when the liquid is ejected toward the surface of the liquid, this corresponds to the "up-pumping" mode. Often, liquid is also drawn from the side of the rotating impeller. Note that axial impellers are sometimes called "mixed-flow" impellers, too: in some cases, part of the ejected flow is directed sideways; this becomes more pronounced when the viscosity of the liquid increases (4).

Figure 3 presents some typical axial-flow impellers. The marine propeller (Fig. 3a) has been used for the propulsion of boats, and nowadays it is its sole application; far more efficient agitators have been designed for mixing liquids. Figure 3b shows the widely used pitched-blade turbine (PBT); the number of blades and their inclination usually characterizes the PBT more specifically, for example, the PBT in the illustration is referred to as a "4-45-PBT." The Mixel TT has blades, which are wider than those of

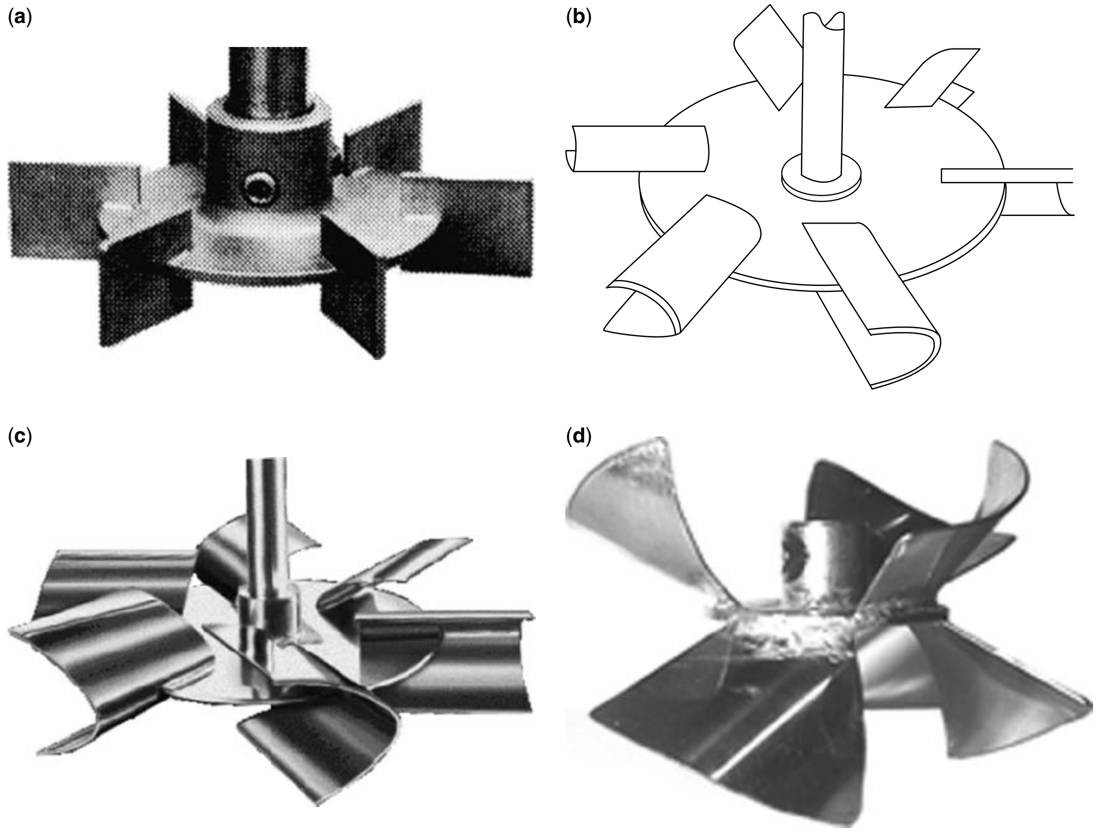

Figure 1. Radial agitators: (**a**) Rushton turbine; (**b**) SCABA 6SRGT (or Chemineer CD-6); (**c**) Chemineer BT-6; (**d**) Narcissus.

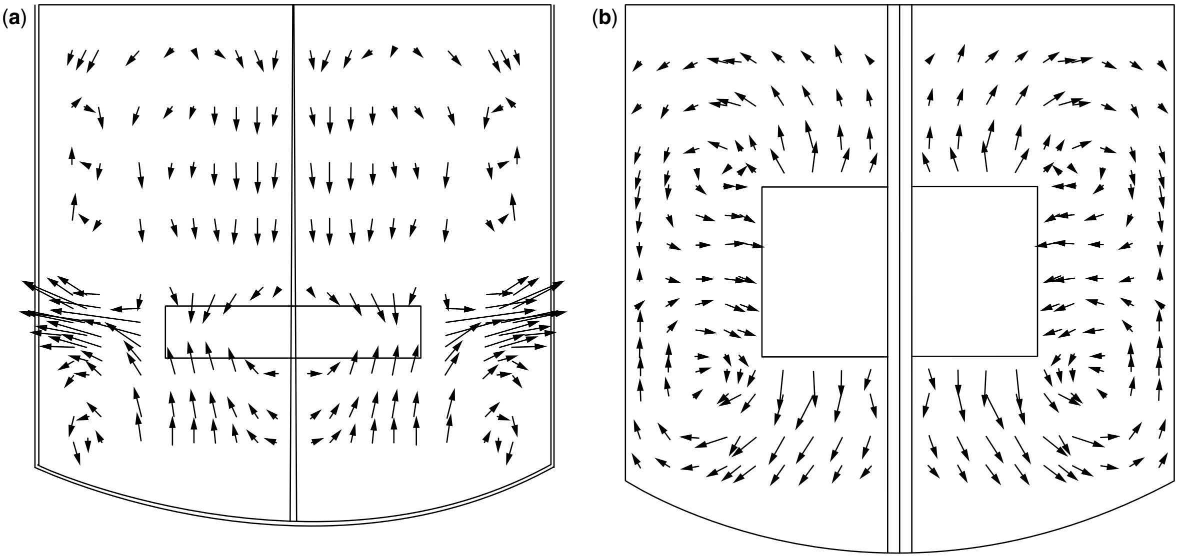

Figure 2. Flow patterns induced by radial impellers in a stirred tank: (**a**) Rushton turbine (4); (**b**) Narcissus (NS) impeller (2).

the PBT and are profiled to be more efficient in energy consumption.

The typical 2-D flow pattern induced by all axial-flow impellers in their usual configuration—“downpumping”—is illustrated in (Fig. 4a). As already stated, liquid is drawn from the upper part and the side of the impeller and is ejected downward. A single circulation loop is established in all cases; liquid flows upward close to the vessel walls and returns toward the impeller. The velocities in the upper part of the vessel are typically rather slow: the U_{RZ} vectors are much shorter than those close to the upper and lower sides of the impeller.

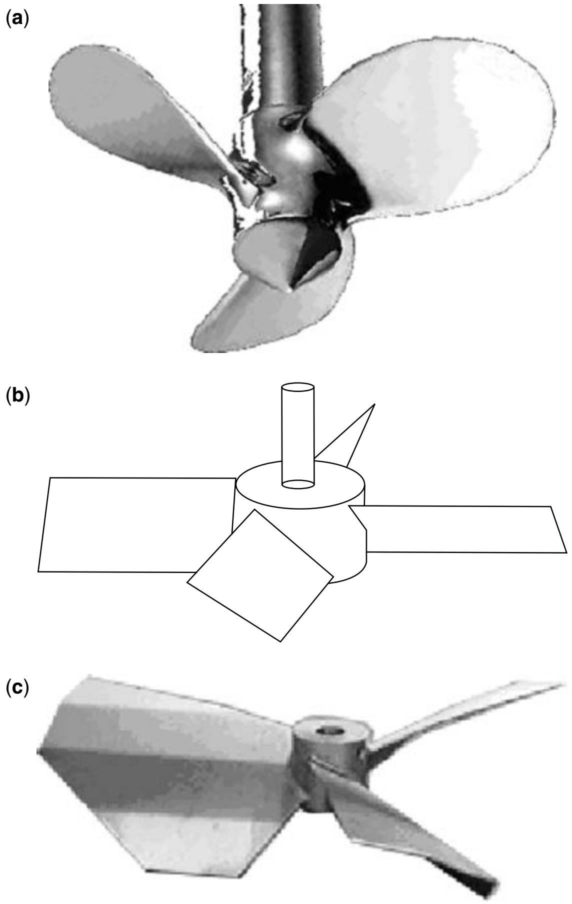

Figure 3. Typical axial-flow impellers; (**a**) marine propeller; (**b**) pitched-blade turbine; (**c**) Mixel TT.

Therefore, the liquid in the stirred tank may be divided into two regions: the first corresponds to the primary circulation loop, which is established around the impeller; liquid flows fast and results in an intensive mixing process. The liquid in the second region, located mainly in the upper part of the vessel, circulates slowly; therefore the mixing process is less intense and effective there; it is often necessary to add a second impeller on the same shaft, to enhance circulation and mixing in the upper part of the vessel.

The inverse configuration—“up-pumping”—again yields a single primary circulation loop, located around the impeller (Fig. 4b). A smaller, secondary circulation loop is established in the upper part of the vessel, achieving better overall circulation and mixing than the “down-pumping” mode.

Several other impellers have been tested and/or marketed, based on extensive hydrodynamic performance measurements, taking into consideration some optimization criterion; some of them are variants of the pitched-blade turbine; others have blade shapes originating from hydrofoils. Figure 5 illustrates some of these impellers.

The size and location of the impeller inside the stirred vessel are dictated by the process needs and affect its performance, for example,

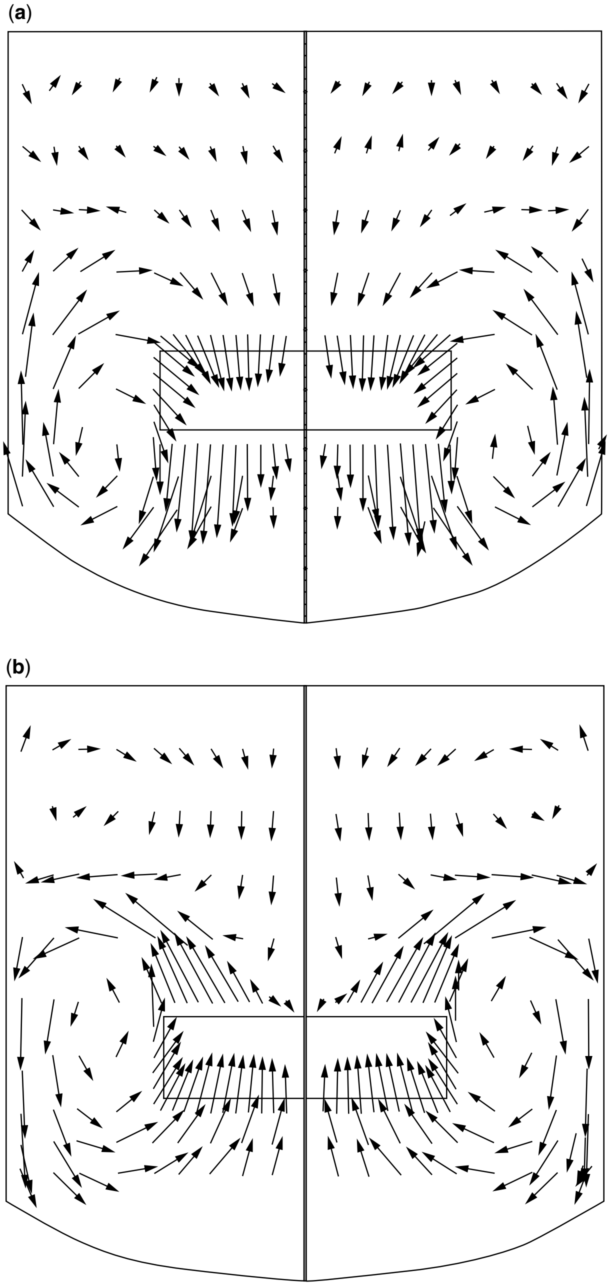

Figure 4. Flow patterns induced by axial impellers in a stirred tank: (**a**) Mixel TT in down-pumping mode [4]; (**b**) Mixel TT in up-pumping mode (5).

- to disperse gas effectively inside a stirred tank, it is necessary to use a radial agitator that has a large impeller diameter (D) to tank diameter (T) ratio, for example, $D/T = 1/2$, and to provide high rotational speed;
- if it is required to provide surface aeration to the stirred tank, the impeller is located close to the free liquid surface;
- if it is necessary to achieve an effective distribution of solid particles, an axial impeller having a reduced size ($D/T = 1/3$) should be used, located closer to the bottom of the vessel, having clearance (C), that is,

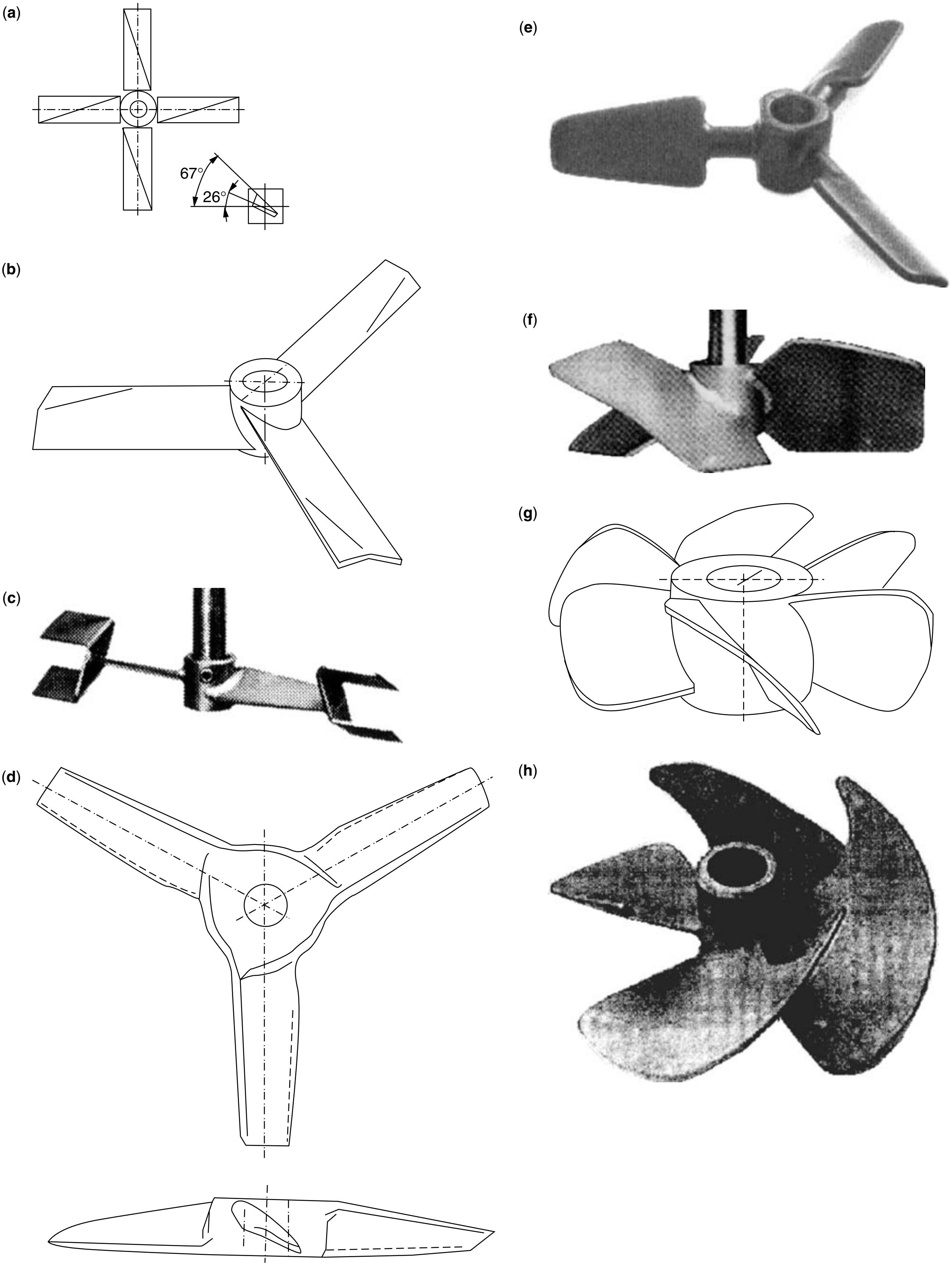

Figure 5. Examples of advanced impellers: (**a**) "Medek" PBT [6,7]; (**b**) Chemineer HE-3 [8]; (**c**) Ekato MIG; (**d**) Lightnin A-310; (**e**) DeDietrich hydrofoil; (**f**) Lightnin A-320; (**g**) Prochem Maxflo; (**h**) APV B2.

the distance from the bottom of the vessel, close to $T/3$ or even $T/4$.

When mixing is applied to rectangular troughs, the axis of the impeller is horizontally located at one end of the trough, and an axial-flow agitator with hydrofoil blades is used to induce longitudinal motion of the liquid in the trough.

PERFORMANCE DATA

The performance of the various impellers is characterized by quantitative criteria; some of these are power consumption, the amount of flow circulation caused by the pumping action of the impeller, the ability of the impeller to cause intense circulation in the stirred tank, and the time necessary to achieve homogeneity of the tank contents, among others.

The power consumption depends upon the impeller type; it has been found that in turbulent conditions, where the dimensionless Reynolds number (Re),

$$Re = \frac{\rho N D^2}{\mu} \tag{1}$$

is larger than about 4000, the dimensionless power number (Po),

$$Po = \frac{P}{\rho N^3 D^5} \tag{2}$$

is approximately constant and characterizes each impeller. Table 1 presents power numbers for a variety of commonly used impellers.

Another feature of impellers is the amount of fluid being "pumped out" of the agitator-swept region; from the flow rate of this stream (Q_p), another dimensionless number, the flow number, (Fl), which also characterizes impellers may be obtained:

$$Fl = \frac{Q_\mathrm{P}}{ND^3} \tag{3}$$

Table 1 presents typical values of flow numbers for the most common types of impellers.

Table 1. Characteristics of Various Impellers

Type	D/T	C/T	Power Number (Po)	Flow Number (Fl)
A310 (Lightnin)	1/2	1/3	0.56	0.62
A315 (Lightnin)	NA[a]	NA[a]	0.75–0.80	0.73
A320 (Lightnin)	0.40	0.39	0.64	0.64
A410 (Lightnin)	0.40	0.40	0.32	0.62
4-45-PBT	1/3	1/3	1.25	0.77
6-45-PBT (down-pumping)	1/3	1/3	1.93	0.75
6-45-PBT (up-pumping)	1/3	1/3	2.58	0.68
6SRGT (SCABA)	1/3	1/3	2.8–3.0	NA[a]
BT-6 (Chemineer)	1/3	1/3	2.1	NA[a]
CD-6 (Chemineer)	1/3	1/3	2.8–3.0	NA[a]
HE-3 (Chemineer)	1/2	1/4	0.31	0.41
Marine propeller	1/3	NA[a]	0.89	0.79
Medek PBT	1/3	1/3	0.41	0.60
Mixel TT (down pumping)	1/2	1/3	0.74	0.67
Mixel TT (up-pumping)	1/2	1/3	0.67	0.61
Narcissus	1/3	1/3	1.14	0.31
MaxFloT (Prochem)	0.35	0.45	1.58	0.82
Rushton turbine	1/3	1/3	4.9–5.2	0.78

[a] NA: not available.

One of the purposes of an impeller is to create circulation inside a stirred vessel, so one quantitative characteristic of its efficiency is the spatial mean velocity achieved in the vessel. This mean velocity, compared to the velocity at the tip of the blades (V_TIP), yields the "agitation efficiency" (I_G) of each particular impeller (9).

Finally, the time to obtain vessel homogeneity is termed "mixing time" (t_MIX); it has been found that for a wide variety of impellers it may be correlated to the power number and to the impeller-to-vessel diameter ratio (10):

$$N\, t_\mathrm{MIX} = 5.3 (Po)^{-1/3} \left(\frac{T}{D}\right)^2 \tag{4}$$

CONCLUSIONS

Mixing is used in a multitude of processes, including water and wastewater treatment, to achieve several goals: to disperse another phase—gas, liquid, or solid—into the bulk of the liquid; to homogenize the stirred tank contents; and to assist and promote a reaction between some of the dissolved and/or dispersed species, among others. This is usually achieved by using rotating impellers, whose blade design has been often optimized for particular processes. Radial impellers, such as the Rushton turbine, are more suitable for homogenization and for dispersing a second phase in liquids; however, they generate high-shear flows. Axial-flow impellers are more suitable for solids dispersion and for cases where shear-sensitive material exists in the liquid, requiring benign mixing conditions.

NOTATION

C: clearance of impeller (from midplane) to vessel bottom (m)
D: impeller diameter (m)
Fl: dimensionless flow number (-)
I_G: dimensionless agitation index (-)
N: impeller rotational frequency (Hz)
Po: dimensionless power number (-)
Re: dimensionless Reynolds number (-)
T: vessel diameter (m)
t_MIX: mixing time (s)
U: liquid velocity (m/s)
V_TIP: liquid velocity at the tip of the impeller blades ($= \pi$ND) (m/s)

GREEK LETTERS

μ: viscosity of liquid (Pa.s)
ρ: density of liquid (kg/m^3)

INDEXES

R: radial
RZ: composite radial-axial
Z: axial

BIBLIOGRAPHY

1. Mavros, P. (2001). Flow visualisation in stirred vessels. Review of experimental techniques. *Chem. Eng. Res. Des.* **79A**(2): 113–127.
2. Hristov, H.V., Mann, R., Lossev, V., and Vlaev, S.D. (2004). A simplified CFD for three-dimensional analysis of fluid mixing, mass transfer and bioreaction in a fermenter equipped with triple novel geometry impellers. *Food and Bioproducts Processing* **82**(C1): 21–34.
3. Khopkar, A., et al. (2004). Flow generated by radial flow impellers: PIV measurements and CFD simulations. *Int. J. Chem. Reactor Eng.* **2**: A18.
4. Mavros, P., Xuereb, C., and Bertrand, J. (1996). Determination of 3-D flow fields in agitated vessels by laser-Doppler velocimetry. Effect of impeller type and liquid viscosity on liquid flow patterns. *Chem. Eng. Res. Design* **74A**: 658–668.
5. Aubin, J., Mavros, P., Fletcher, D.F., Bertrand, J., and Xuereb, C. (2001). Effect of axial agitator configuration (up-pumping, down-pumping, reverse rotation) on flow patterns generated in stirred vessels. *Chem. Eng. Res. Des.* **79A**(8): 845–856.
6. Medek, J. and Fořt, I. (1991). *Relation between designs of impeller with inclined blades and their energetic efficiency. Proc. 7th Eur. Conf. Mixing*, Brugge, Belgium, Sept. 18–20, KVIV, vol. I, pp. 95–102.
7. Medek, J., Seichter, P., Rybníček, Z., and Chaloupka, Z. (1987). Czechoslovak patent no. 256630.
8. Ibrahim, S. and Nienow, A.W. (1995). Power curves and flow patterns for a range of impellers in Newtonian fluids $-40 < Re < 5 \times 10^5$. *Chem. Eng. Res. Design* **73**(A5): 485–491.
9. Mavros, P. and Baudou, C. (1997). Quantification of the performance of agitators in stirred vessels. Definition and use of an agitation index. *Chem. Eng. Res. Des.* **75A**: 737–745.
10. Ruszkowski, S. (1994). A rational method for measuring blending performance and comparison of different impeller types. *Proc. 8th Eur. Conf. Mixing*. Cambridge, UK, Sept. 21–23, IChemE, Rugby, pp. 283–291.

ARSENIC IN NATURAL WATERS

Robert Y. Ning
King Lee Technologies
San Diego, California

INTRODUCTION

Arsenic is widely distributed in nature in air, water, and soil as a metalloid and as chemical compounds, both inorganic and organic (1). This class of compounds was known to the ancient Greeks and Romans both as therapeutic agents as well as poisons. This dual nature as useful substances as well as toxic matter to be controlled has grown over the centuries. Arsenic and arsenicals have widened use in commerce, but so have the recognition that their presence in drinking water, largely from natural sources, is a major public health problem around the world.

Acute and chronic arsenic exposure via drinking water has been reported in many countries, especially Argentina, Bangladesh, India, Mexico, Mongolia, Thailand, and Taiwan, where ground (well) water is contaminated with a high concentration of arsenic of 100 to more than 2000 μg/liter (ppb) (2). Studies have linked long-term exposure to arsenic in drinking water to cancer of the bladder, lungs, skin, kidney, nasal passages, liver, and prostate. Noncancer effects of ingesting arsenic include cardiovascular, pulmonary, immunological, neurological, and endocrine (e.g., diabetes) disorders (3). Besides its tumorigenic potential, it has been shown that arsenic is genotoxic (4,5). Given the importance of arsenic as a global environmental toxicant, we will summarize the geochemistry, natural distribution, regulation, anthropogenic sources, and removal mechanisms.

GEOCHEMISTRY

Average concentrations of arsenic in the earth's crust reportedly range from 1.5 to 5 mg/kg. Higher concentrations are found in some igneous and sedimentary rocks, particularly in iron and manganese ores. Common minerals containing arsenic are shown in Table 1. Arsenopyrite, realgar, and orpiment are the most important of these minerals, and they are commonly present in the sulfide ores of other metals, including copper, lead, silver, and gold. Arsenic may be released from these ores to the soil, surface water, groundwater, and the atmosphere.

Natural concentrations of arsenic in soil typically range from 0.1 to 40 mg/kg; an average concentration is 5 to 6 mg/kg. Arsenic can be released to ground or surface water by erosion, dissolution, and weathering. Geothermal waters can be sources of arsenic in groundwater.

In Yellowstone National Park, the arsenic concentrations in geysers and hot springs range from 900 to 3,560 ppb. Waters from these sources cause elevated arsenic levels in rivers downstream. Other natural sources include volcanism and forest fires. Volcanic activity appears to be the largest natural source of arsenic emissions to the atmosphere, estimated variously between 2,800 to 44,000 metric tons annually. The relative contributions of volcanic sources, other natural sources (Table 1), and anthropogenic sources to the atmosphere have not been definitively established.

The predominant forms of arsenic in groundwater and surface water are arsenate (V) and arsenite (III). Examples

Table 1. Common Minerals of Arsenic[a]

Arsenopyrite, $FeAsS$	Smalite, $CoAs_2$
Lollingite, $FeAs_2$	Cobaltite, $CoAsS$
Orpiment, As_2S_3	Gersdorffite, $NiAsS$
Realgar, As_4S_4	Tennantite, $4Cu_2SAs_2S_3$
Chloanthite, $NiAs_2$	Proustite, $3Ag_2SAs_2S_3$
Niciolite, $NiAs$	Enargite, $3Cu_2SAs_2S_5$

[a] Reference 3.

of inorganic arsenic compounds found in the environment include oxides (As_2O_3, As_2O_5), and sulfides (As_2S_3, AsS, $HAsS_2$, $HAsS_3^{3-}$). Inorganic arsenic species that are stable in oxygenated waters include arsenic acid [As(V)] species (H_3AsO_4, $H_2AsO_4^-$, $HAsO_4^{2-}$ and AsO_4^{3-}). Arsenous acid [As(III)] is also stable as H_3AsO_3 and $H_2AsO_3^-$ under slightly reducing aqueous conditions. Arsenite is generally associated with anaerobic conditions. Oxidation state, oxidation–reduction potential, pH, iron concentrations, metal sulfide and sulfide concentrations, temperature, salinity, and distribution and composition of biota appear to be the significant factors that determine the fate and transport of arsenic. In surface waters, additional factors include total suspended sediment, seasonal water flow volumes and rates, and time of day.

Sorption of arsenic to suspended sediment may strongly affect the fate and transport of arsenic in surface water systems (6). Where pH and arsenic concentrations are relatively high and total suspended sediment levels are relatively low, sorption processes may be less important. However, where suspended sediment loads are higher, arsenic concentrations are lower, and pH levels are lower, arsenic is more likely to be present in the suspended particulate phase rather than in the dissolved phase. Particulate phase arsenic may settle to bottom sediment in reservoirs and areas of low flow levels. In deeper lakes, remobilization of arsenic from sediment may be minimal, whereas in shallower lakes, arsenic may be remobilized faster from wind-induced wave action and high-flow scouring. Diurnal changes of as much as 21% in arsenic concentrations have been observed in rivers, attributable to pH changes due to sunlight and photosynthesis.

NATURAL DISTRIBUTION

A survey of arsenic concentration in natural waters is of importance relative to the desirable maximum limit of 10 ppb or less for human consumption. An attempt has been made to quantify the global element cycle for arsenic, based on published data (1). Arsenic concentrations in environmental media are presented in Table 2.

Table 2. Arsenic Concentrations in Environmental Media[a]

Environmental Media	Arsenic Concentration Range	Units
Air	1.5–53	ng/m^3
Rain from unpolluted ocean air	0.019	μg/L (ppb)
Rain from terrestrial air	0.46	μg/L
Rivers	0.20–264	μg/L
Lakes	0.38–1,000	μg/L
Ground (well) water	<1.0–>1, 000	μg/L
Seawater	0.15–6.0	μg/L
Soil	0.1–1,000	mg/kg (ppm)
Stream/river sediment	5.0–4,000	mg/kg
Lake sediment	2.0–300	mg/kg
Igneous rock	0.3–113	mg/kg
Metamorphic rock	0.0–143	mg/kg
Sedimentary rock	0.1–490	mg/kg
Biota—green algae	0.5–5.0	mg/kg
Biota—brown algae	30	mg/kg

[a] Reference 7.

In addition to geochemical factors, microbial agents can influence the oxidation state of arsenic in water and can mediate the methylation of inorganic arsenic to form organic arsenic compounds (8). Microorganisms can oxidize arsenite to arsonate and reduce arsenate to arsenite or even to arsine (AsH_3). Bacteria and fungi can reduce arsenate to volatile methylarsines. Marine algae transform arsenate into nonvolatile methylated arsenic compounds such as methylarsonic acid [$CH_3AsO(OH)_2$] and dimethylarsinic acid [$(CH_3)_2AsO(OH)$] in seawater. Freshwater and marine algae and aquatic plants synthesize complex lipid-soluble arsenic compounds (9). Organic arsenical compounds were reportedly detected in surface water more often than in groundwater. Surface water samples reportedly contain low but detectable concentrations of arsenic species, including methylarsonic acid and dimethylarsinic acid. Methylarsenicals reportedly comprise as much as 59% of total arsenic in lake water. In some lakes, dimethylarsinic acid has been reported as the dominant species, and concentrations appear to vary seasonally from biological activity within waters.

REGULATIONS

In the United States, the Safe Drinking Water Act (SDWA) of 1974 called for establishing Maximum Contaminant Levels (MCL) as national drinking water standards and required the Environmental Protection Agency (EPA) to revise the standard periodically. Based on a Public Health Service standard established in 1942, the EPA established a standard of 50 μg/liter (50 ppb) as the maximum arsenic level in drinking water in 1975. In 1984, the World Health Organization (WHO) followed with the same 50 ppb recommendation. Since that time, rapidly accumulated toxicity information prompted a revision of the standard, and a provisional guideline of 10 ppb was recommended by WHO in 1993. In January 2001, EPA published a revised standard that would require public water supplies to reduce arsenic to 10 ppb by 2006. Perceived hardships in implementation and uncertainty in setting the standard at 3, 5, 10, or 20 ppb has led the EPA to announce temporary delays in the effective date for the January 2001 rule to allow for further cost–benefit analysis and public input. The rule is significant because it is the second drinking water regulation for which the EPA has used its discretionary authority under the SDWA to set the MCL higher than the technically feasible level, which is 3 ppb for arsenic, based on the determination that the costs would not justify the benefits at this level.

ANTHROPOGENIC SOURCES

Arsenic is released from a variety of anthropogenic sources, including metal and alloy manufacturing, petroleum refining, pharmaceutical manufacturing, pesticide manufacturing and application, chemical manufacturing, burning of fossil fuels, and waste incineration.

Most agricultural uses of arsenic are banned in the United States. However, sodium salts of methylarsonic acid are used in cotton fields as herbicides. Organic arsenic is also a constituent of feed additives for poultry and swine and appears to concentrate in the resultant animal wastes. About 90% of the arsenic used in the United States is for the production of chromated copper arsenate (CCA), the wood preservative. CCA is used to pressure treat lumber and is classified as a restricted use pesticide by the EPA. A significant industrial use of arsenic is in the production of lead-acid batteries; small amounts of very pure arsenic metal are used to produce gallium arsenide, which is a semiconductor used in computers and other electronic applications. The U.S. Toxics Release Inventory data indicated that 7,947,012 pounds of arsenic and arsenic-containing compounds were released to the environment in 1997; most of that came from metal smelting. The data did not include some potentially significant arsonic sources associated with herbicides, fertilizers, other mining facilities, and electric utilities.

REMOVAL MECHANISMS

At the regulated maximum arsenic level of 10 ppb, the U.S. EPA estimated that 5% of all U.S. community water systems would have to take corrective action to lower the current levels of arsenic in their drinking water. In high arsenic areas of the world, the need for removal from water supplies is even more acute.

Due to their predominance in natural waters, arsenic(V) acid (H_3AsO_4) and arsenous(III) acid ($HAsO_2$) and their salts can serve as the model for these and alkylated species for consideration of removal mechanisms. The pK values of arsenic acid = 2.26, 6.76, 11.29 (10) and arsenious acid = 9.29 (10) or 8.85 (11) are of prime importance in determining the degree of ionization at the pHs of the water from which removal strategies are considered. It is readily apparent that at a natural pH of 7 to 8, arsenic acid is extensively ionized as the divalent ion; arsenious acid remains largely un-ionized.

Due to the ionic charge, arsenate(V) is more easily removed from source waters than arsenite(III). In particular, activated alumina, ion exchange, and reverse osmosis may achieve relatively high arsenate removal rates, but they show lower treatment efficiencies for arsenite. Elevating the pH such as by caustic injection into reverse osmosis system feedwater would be one approach to greater removal of arsenite(III) compounds. Arsenite can also be oxidized to arsenate to improve removal efficiencies. In water that contains no ammonia or total organic carbon, chlorine rapidly (in less than 5 seconds at chlorine concentrations of 1.0 mg/L) oxidizes approximately 95% of arsenite to arsenate. Monochloramine at a concentration of 1.0 mg/L oxidized 45% of arsenite to arsenate. Potassium permanganate performs this oxidation rapidly; oxygen does so slowly unless activated by light and sensitizer.

In contrast to other heavy metals, As^{3+} and As^{5+} are not precipitated as hydroxides, only as sulfides. Alkyl and arylarsonic acids are precipitated by quadrivalent metals such as tin, thorium, titanium, and zirconium. The tendencies of dissolved arsenic species to adsorb on inorganic particle surfaces such as iron, ferric hydroxide, iron oxide, alumina, sulfur, and sulfides allow for removal strategies involving fixed-bed reactors or adsorption/coagulation/filtration schemes. Sulfate, fluoride, and phosphate ions are known to be strong competitors of arsenic adsorption in some systems.

BIBLIOGRAPHY

1. Matschullat, J. (2000). Arsenic in the geosphere—a review. *Sci. Total Environ.* **249**(1–3): 297–312.
2. Tchounwou, P.B., Wilson, B., and Ishaque, A. (1999). *Rev. Environ. Health* **14**(4): 211–229.
3. Various reports available from the US-Environmental Protection Agency website: www.epa.gov/safewater/arsenic.html and the World Health Organization website: www.who.int/water_sanitation_health/water_quality/arsenic.htm.
4. Basu, A., Mahata, J., Gupta S., and Giri, A.K. (2001). *Mutat. Res.* **488**(2): 171–194.
5. Gebel, T.W. (2001). *Int. J. Hyg. Environ. Health* **203**(3): 249–262.
6. Nimick, D.A., Moore, J.N., Dalby, C.E., and Savka, M.W. (1998). *Wat. Res. Research* **34**(11): 3051–3067.
7. US-Environmental Protection Agency. (2000). *Arsenic Occurrence in Public Drinking Water Supplies*, EPA-815-R-00-023, December 2000, available from EPA Website (Ref. 3).
8. Tamaki, S. and Frankenberger, W.T., Jr. (1992). *Rev. Environ. Contam. Toxicol.* **124**: 79–110.
9. Shibata, Y., Morita, M., and Fuwa, K. (1992). *Adv. Biophys.* **28**: 31–80.
10. Lide, D.R. (Ed.). (1995). *CRC Handbook of Chemistry and Physics*, 76th Edn. CRC Press, New York, pp. 8–43.
11. Meites, L. (Ed.). (1963). *Handbook of Analytical Chemistry*. McGraw-Hill, New York, pp. 1–29.

EVALUATION OF MICROBIAL COMPONENTS OF BIOFOULING

STUART A. SMITH
Smith-Comeskey Ground Water
Science LLC
Upper Sandusky, Ohio

Detecting the occurrence of biofouling and assessing its impact involve a range of analytical techniques, including informed observation and inspection of well components, interpretation of hydraulic and water quality testing, and direct analysis of microbial components of biofouling. The latter is the subject of this article.

SYMPTOMS: INDIRECT ANALYSIS

Symptoms can be used as qualitative indicators of biofouling but not specifically of the microbial component. Observable symptoms include the following:

- clogging (both formation/well and pump/discharge systems).

- corrosion
- alteration of water quality in pumped samples
- distinct coating on surfaces visible during inspections, such as borehole television surveys.

Such symptoms are indications of biofouling and should trigger analysis to determine the nature of the biofouling.

DIRECT ANALYSIS OF MICROBIAL BIOFOULING AND COMPONENTS

Standard Methods

Some existing "standard methods" for analyzing aspects of biofouling are described in:

- ASTM Test Method for Iron Bacteria in Water and Water-Formed Deposits (D 932) (1)
- Section 9240—Iron and Sulfur Bacteria, *Standard Methods for the Examination of Water and Wastewater* (APHA-AWWA-WEF, currently 20th ed. Supplement) (2).
- Additional *Standard Methods* (Part 9000) microbiological methods (and others accepted by technical and regulatory bodies) for analysis of heterotrophic bacteria and specific groups of interest such as the total coliform group. All relevant aquatic and public health microbiological methods are applicable in identifying microbiological components of biofouling.
- Microscopic particulate analysis (MPA) ((2), consensus method) can identify biofilm components as part of its larger scope and represents a systematic approach.

Microscopic Examination and Analysis. The presence of filamentous (e.g., *Leptothrix, Thiothrix,* or *Crenothrix*) or stalked (*Gallionella*) iron or sulfur bacterial forms is accepted as a positive indicator of biofouling. Examination by light microscopy has traditionally been the method of choice for confirming and identifying these "iron bacteria" or "sulfur bacteria" (1,2). However, the absence of such visible structures does not necessarily mean the absence of biofouling:

1. Samples may not include enough recognizable materials to provide the basis for a diagnosis of biofouling.
2. Samples examined may not include the filamentous or stalked bacteria normally searched for in such analyses. It is generally understood that the morphologically distinct types are only part of the biofouling present.
3. Analysts vary in their skill and opinions in interpreting what they see.

In addition, the existing standard microscopic tests (1,2) are specified as qualitative. Attempts have been made to quantify degrees of biofouling by microscopy. However, spatial variations and the pulsating, three-dimensional nature of active biofilms make them unreliable (4). An available semiquantitative method using microscopy is described by Barbic et al. (5).

Culturing Methods for Detection. Culturing enriches biofouling microflora that cannot be identified microscopically and helps provide more complete information on the nature of biofilm samples. It can also be used to draw reasonable conclusions about biofouling in the absence of microscopy. *Standard Methods* (2) describes a range of media formulations, and more for other purposes are found in other sections of Part 9000. With regard to Section 9240 formulations, a number of limitations have been identified. These weaknesses have all been recognized by the *Standard Methods* Section 9240 joint technical group:

- No reported effort has been made to standardize these media with reference cultures from well water. Thus the efficiency of recovery of iron-precipitating bacteria from groundwater samples remains unknown at present (4) but is probably very low.
- Except for the modified Wolfe's medium for *Gallionella* enrichment (6), the environmental conditions for growth in these media do not seem to match well with groundwater environmental conditions (7,8).
- Useful Mn-precipitation media have been one area of weakness in common practice (7), although they are being refined (9).
- The available sulfur oxidizer media are still nonisolating, enrichment media.
- Employment of all of these culture methods requires preparation from raw materials (no packaged agar or broth media specific to IRB are available), sterilization, and maintenance in a microbiological laboratory by a skilled person. Thus, they are rarely used in operational practice.

Cultural Media Improvements. An important innovation of the last 20 years in biofouling cultural recovery has been the development of prepackaged cultural media that permit (1) practical use in operational monitoring in addition to more academic analyses and (2) recovery of a range of microflora, resulting in a more complete understanding of the microbial ecology of wells and more refined maintenance and management of water quality (see entry GW-1311 Well Maintenance). Two groups using slightly different approaches developed these methods independently:

MAG Method (MAG Laboratorio Ambiental, La Plata, Argentina). MAG tests for heterotrophic iron-related bacteria (IRB) and sulfate-reducing bacteria (SRB) consist of a prepared liquid medium contained in small septum bottles (10,11). The MAG medium for iron-related heterotrophic bacteria (BPNM-MAG) uses ferric ammonium citrate [like W-R and R2A + FAC (7)] and the SRB medium (BRS-MAG) uses Postgate C medium under a reducing atmosphere, supplemented with iron filings (11). Inoculation of a single bottle provides a presence-absence (P-A) result. Dilution to extinction provides a semiquantitative (MPN) result (11).

BART Method. Currently, the most commonly used cultural approach for routine biofouling monitoring is the Biological Activity Reaction Test (BART^{TM}) Method (Droycon Bioconcepts Inc., Regina, Saskatchewan).

- BART tubes contain dehydrated media formulations and a floating barrier device, which is a ball that floats on the hydrated medium of the sample.
- These devices and their proposed use are described in detail in Cullimore (12). They can be used as an enrichment method to provide a presence-absence (P-A) or semiquantitative (MPN) detection of biofouling factors (7,12–14).

BART tubes are available in a variety of media mixtures. The IRB-BART^{TM} test, for example, is designed to recover microaerophilic heterotrophic Fe- and Mn-precipitating microorganisms and is (like the BPNM-MAG test) derived from the W-R iron bacteria medium (15). This method, which is gaining wide operational acceptance as a means of detecting and characterizing biofouling symptoms, has provided useful qualitative information in well biofouling in various field trials (7) and has proven useful in a range of applications.

These methods have proved to be significant advances in making microbial ecology an important factor in routine operational monitoring as they

1. Reliably provide results to the level of detection needed to make operational decisions.
2. Are relatively easy to use and interpret.
3. Offer a means for precise scientific characterization of the microbial system if properly used and interpreted within their limitations.

At present, neither of these systems is included in *Standard Methods*, although they have become operationally de facto standard methods within the water operational and hydrogeologic communities (16). A definitive comparability test among these and various *Standard Methods* media is yet to be conducted.

Hybrid Methods. A subcategory of cultural methods consists of field-usable enrichment procedures to increase the potential for successful detection by microscopy. For example, Alcalde and co-workers (10,17) describe a simple enrichment and staining technique to enhance the numbers and visibility of filamentous bacteria on glass slides.

Nonculturing Biofouling Analytical Methods. All cultural enrichment methods recover only a fraction (1% is an often cited value) of microflora in environmental samples. Employing a range of cultural media and using media that better approximate ideal growth conditions (and making them easier to use) helps to overcome these problems; however, many microflora types may remain unrecoverable. Nonculturing analytical methods that directly detect interpretable evidence of microflora bypass these limitations.

Phospholipid Fatty Acid (PLFA) or Phospholipid Fatty Acid Methyl Ester (PL-FAME) Analysis. The method described by White and Ringelberg (18,19) and its variations provide a nonculturing method of characterizing the microbial components of environmental samples (solids such as sediment cores as well as fluids). In PLFA or PL-FAME, "signature" lipid biomarkers from the cell membranes and walls of microorganisms are extracted from the sample. Particular types of biomarkers are linked to groups of microorganisms. Amounts of biomarkers can be reliably associated with viable biomass, and compound ratios can be linked to community nutritional status. Once extracted from cells with organic solvents, the lipids are concentrated and fractionated using gas chromatography/mass spectrometry (GC/MS). A profile of the fatty acids and other lipids is then used to determine the characteristics of the microbial community. This "fingerprint" represents the living portion of the microbial community because phospholipids degrade rapidly following cell death (20).

A variation is TC-FAME (total-cell FAME), in which the lipids of whole cells are extracted and characterized from microbial isolates. These fatty acid profiles are compared to libraries of known isolates to identify unknowns in environmental samples or pure cultures.

Adenosine Triphosphate (ATP) and Adenylate Kinase (AK) Detection. ATP is ubiquitous in cellular life and can be detected rapidly using bioluminescence methods. The ATP bioluminescence test (21), detects ATP, which is present only in living cells. The amount of light emitted in the reaction can be correlated with the amount of ATP that can be extracted from a known number of bacteria. The test is rapid (<1 minute) and can be conducted with convenient test kits. Differences in ATP amounts occur by cell type and by growth phase and nutrient availability (22), and 10^5 cells/mL is the lower detection limit, requiring concentration for lower concentrations (*Standard Methods* Section 9211 C). The AK test detects a cellular enzyme (adenylate kinase-catalyzing ATP) that is present at a fairly consistent level in cells of all growth phases except spores, and the process (also rapid) can detect cells at 10 cells/mL (22). These tests are primarily useful in determining the total numbers of cells in a sample. What this means depends on the nature of the sample, but the ATP or AK result can be used for comparisons and in making operational or treatment decisions.

Nucleic Acid Testing. To identify biofouling microbial constituents definitively, genetic testing is used. Nucleic acids extracted are compared to libraries of sequences. Among the most useful procedures in groundwater microbial analysis is 16S rRNA community structure analysis (20). This method permits identifying microbes within complex environmental samples (such as biofilms), typically expressed as percent homology with known groups, but it is possible to identify isolates to the strain level.

Analytical Choice Considerations

All available analytical methods have their limitations. Fortunately, the strengths and weakness of microscopy and cultural enrichment appear to be complementary (7), so they are best used in tandem. The best use of BART or MAG methods is as indications of:

1. early detection or confirmation of probable biofouling and MIC conditions,
2. the presence of viable bacterial types in both liquid and extracted or eluted solids samples, and
3. relative "aggressivity" or activity of biofouling and corrosion-inducing communities.

Standard or heterotrophic plate counts (HPC) have often been used in water analysis for quantitative results. The validity of HPC has been called into question on the viability issue and because agar plates are an osmotically hostile environment. The actual organisms recovered in HPC testing can also vary widely among locations, seasons, and consecutive samples at a single location (23). Also, unamended HPC media tell little about what the recovered bacteria are doing in the environment. For these reasons, cultural methods developed for biofouling analysis are preferred, but HPC can be used if calibrated for specific applications and as a gauge to interpret some BART or MAG results.

PL-FAME provides a nonculturing method that can reliably assay the types of microflora (some of which are nonviable or dormant) in a range of samples, including biofilm samples, and their relative health and vigor. ATP/AK methods provide rapid cells-per-unit counts of living microflora reliable within a margin of error, which can also be conducted under field conditions in certain circumstances.

Where genetic identification is important (as when pathogen presence is suspected or to be ruled out), 16S rRNA analysis identifies all bacterial types. Both PL-FAME and 16S rRNA analyses are relatively expensive, although information-rich.

Sampling Methods

A serious limitation in biofouling monitoring is that recovering "representative" samples of biofouling microflora is difficult. Pumped (grab) sampling is the easiest way to obtain samples for analysis from wells or sample taps in pipelines. However, if pumping fails to detach and suspend biofilm particles, they will not be available for collection and analysis. Thus, a presumption of "probably absent" based on lack of recovery of microscopically visible particles (per Section 9240 or D 932) or cultural recovery may be a false negative. To improve the odds in recovery, it is useful to employ one or more of the following:

1. Filtration or centrifugation as recommended in *Standard Methods* increases the odds of recovering material useful for microscopic identification.
2. Time-series sampling: Cullimore (12) describes a time-series pumped-sampling procedure (similar to familiar procedures for groundwater quality analysis) that attempts to overcome the randomness of grab sampling. The time-series sampling approach, which includes taking replicates of samples at each sampling step, helps to overcome the statistical limitations of random, pumped, grab sampling for culture analysis.
   - This information can be used in assessing the microbial ecology in the well and the aquifer adjacent to the well (12). This procedure can be incorporated into a step-drawdown pumping test procedure.
   - Analysis of samples taken at various points in a groundwater-source system (flow-series instead of time-series) permits developing a profile of the microbial ecology of the system.
3. Collection of a biofilm on immersed surfaces can provide essentially intact biofilms for analysis. These methods are also adaptable for collecting samples of inorganic encrustations.

Several experimental designs for collecting biofilm organisms have been presented in the literature since at least the 1920s. Recent methods and equipment for biofilm coupon collection are summarized in Smith (7).

Sampling methods are closely linked to analytical methods:

- Pumped sampling is used for cultural enrichment (MAG or BART) or noncultural analysis of water (e.g., PL-FAME or ATP/AK).
- Time-series or flow-series sampling permits spatial analysis of results.
- Collection on surfaces is preferred for microscopy samples.
- Collection of solids provides samples that can be examined microscopically or extracted for culturing (qualitative) and nonculturing assay.
- All are practical in the operational setting (7,24).

PRACTICAL ISSUES IN MAINTENANCE MONITORING IMPLEMENTATION AND INTERPRETATION

Successful control of biofouling and associated changes in water quality, such as increased turbidity or alteration of metals content, depends very much on preventive or early warning monitoring (see MW-76, Well Biofouling and GW-1311, Well Maintenance). Among applications of microbiological analysis, biofouling monitoring of groundwater systems is a case study of methods in transition attempting to meet user needs. In contrast to available detection methods for microflora of concern to human health, convenient methods yielding information useful to formulate a response to biofouling have not been readily available to operators of groundwater systems until very recently.

At present, biofouling analytical methods remain at best semiquantitative and offer results that require expert interpretation. Consistency in usage and cross-correlation with other information (e.g., biochemical, physicochemical water quality and mineralogy) are keys to extracting the maximum information from biofouling analysis for a particular situation.

BIBLIOGRAPHY

1. ASTM D 932. *Test Method for Iron Bacteria in Water and Water-Formed Deposits*. American Society for Testing and Materials, West Conshohocken, PA.
2. APHA-AWWA-WEF. (1998). Section 9240. *Standard Methods for the Examination of Water and Wastewater*. 20th ed. American Public Health Assn., Washington, DC.
3. U.S. EPA, 1992. *Consensus Method for Determining Groundwaters Under the Direct Influence of Surface Water Using Microscopic Particulate Analysis*, Method 910992029. Office of Drinking Water, Washington, DC.
4. Tuhela, L., Smith, S.A., and Tuovinen, O.H. (1993). Flowcell apparatus for monitoring iron biofouling in water wells. *Ground Water* **31**: 982–988.
5. Barbic, F., Bracilovic, D., Djorelijevski, S., Zivkovic, J., and Krajincanic, B. (1974). Iron and manganese bacteria in Ranney wells. *Water Res.* **8**: 895–898.
6. Hanert, H.H. (1981). The genus *Gallionella*. In: *The Prokaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria*, M.P. Starr et al. (Eds.). Springer-Verlag, Berlin, Chap. 40, pp. 509–521.
7. Smith, S.A. (1992). *Methods for Monitoring Iron and Manganese Biofouling in Water Supply Wells*. AWWA Research Foundation, Denver, CO.
8. Hallbeck, L. (1993). *On the Biology of Iron-Oxidizing and Stalk-Forming Bacterium Gallionella ferruginea*. Department of General and Marine Microbiology, University of Göteborg, Göteborg, Sweden.
9. Emerson, D. and Ghiorse, W.C. (1992). Isolation, cultural maintenance, and taxonomy of a sheath-forming strain of *Leptothrix discophora* and characterization of manganese-oxidizing activity associated with the sheath. *Appl. Environ. Microbiol.* **58**: 4001–4010.
10. Alcalde, R.E. and Gariboglio, M.A. (1990). Biofouling in Sierra Colorado water supply: A case study. *Microbiol. Civ. Eng., FEMS Symp. No. 59*, P. Howsam (Ed.). E.& F.N. Spon, London, pp. 183–191.
11. Gariboglio, M.A. and Smith, S.A. (1993). *Corrosión e encrustación microbiológica en sistemas de captación y conducción de agua - Aspectos teóricos y aplicados*. Consejo Federal de Inversiones, San Martin, C.F., Argentina.
12. Cullimore, D.R. (1993). *Practical Ground Water Microbiology*. CRC Lewis, Boca Raton, FL.
13. Mansuy, N., Nuzman, C., and Cullimore, D.R. (1990). Well problem identification and its importance in well rehabilitation. In: *Water Wells Monitoring, Maintenance, and Rehabilitation*, P. Howsam (Ed.). E.&F.N. Spon, London, pp. 87–99.
14. Walter, D.A. (1997). *Geochemistry and Microbiology of Iron-Related Well-Screen Encrustation and Aquifer Biofouling in Suffolk County*, Long Island, New York, Water Resources Investigation Report 97-4032. U.S Geological Survey, Coram, NY.
15. Cullimore, D.R. and McCann, A.E. (1977). The identification, cultivation, and control of iron bacteria in ground water. In: *Aquatic Microbiology*, F.A. Skinner and J.W. Sewan (Eds.). Academic Press, New York, pp. 219–261.
16. Smith, S.A. (1996). Monitoring biofouling in source and treated waters: Status of available methods and recommendation for standard guide. In: *Sampling Environmental Media*, ASTM STP 1282, J.H. Morgan, (Ed.), American Society for Testing and Materials, West Conshohocken, PA, pp. 158–175.
17. Alcalde, R.E. and Castronovo de Knott, E. (1986). *Occurrence of iron bacteria in wells in Rio Negro (Argentina)*, *Proc. Int. Symp. Biofouled Aquifers: Prevention and Restoration*, D.R. Cullimore (Ed.), American Water Resources Association, Bethesda, MD, pp. 127–136.
18. White, C.D. and Ringelberg, B.D. (1992). *Signature Lipid Biomarker Analysis*. Center for Environmental Biotechnology Standard Operating Procedures. University of Tennessee, Knoxville, TN
19. White, D.C. and Ringelberg, D.B. (1997). Utility of the signature lipid biomarker analysis in determining the *In Situ* viable biomass, community structure, and nutritional/physiological status of deep subsurface microbiota. In: *The Microbiology of the Deep Terrestrial Subsurface*, P.S. Amy and D.L. Haldeman (Eds.). CRC Press Lewis, Boca Raton, FL, pp. 119–136.
20. USEPA NRMRL undated. *Innovative measures of treatment performance*. National Risk Management Research Laboratory, U.S. EPA, Cincinnati, OH.
21. Picciolo, G.L., Chapelle, E.W., Deming, J.W., Thomas, R.R., Nible, D.A., and Okrend, H. (1981). *Firefly luciferase ATP Assay Development for Monitoring Bacterial Concentration in Water Supplies*, EPA-600/S2:81-04. U.S. Environmental Protectional Agency, Cincinnati, OH.
22. Biotrace International (2003). Biochemical Indexes, Issue Ref MLM177 TEP004 Issue no. 002, Biotrace International. http://www.biotrace.com.
23. WHO (2002). *Heterotrophic Plate Count Measurement in Drinking Water Safety Management*, WHO/SDE/WSH/02.10. Report of an expert meeting, Geneva, April 24–25, 2002, World Health Organization, Geneva.
24. Smith, S.A. (1995). *Monitoring and Remediation Wells: Problems Prevention and Cures*. CRC Lewis, Boca Raton, FL.

THREAT AGENTS AND WATER BIOSECURITY

JAMES HIGGINS
USDA-ARS
Beltsville, Maryland

THE NATURE OF THE THREAT

This article focuses on biological agents that can be maliciously introduced into potable water supplies and emphasizes those that would most likely be used, what the state of knowledge is on disinfection and/or inactivation protocols, and existing methods for detecting these agents.

The anthrax bioterror event that occurred in the eastern United States in the fall of 2001 provided graphic evidence of the considerable economic and public health impact that deliberate release of a pathogenic agent could engender, even in a technologically advanced nation. The event brought increased awareness of the perils of threat agents to public health and security personnel and led to the implementation of methods to determine when these agents may be present and to ameliorate their effects. The water industry (here used to refer principally to potable water suppliers) has been no exception; as part of the Public Health Security and Bioterrorism and Response Act (H.R. 3448) of 2002, all utilities providing water to 3300 people or more are required to have an emergency response plan in the event of a terror attack on their infrastructure.

In the light of the increased demands these new regulations make on the water industry, it is reasonable to ask, how real is the threat? Undoubtedly, some information on this topic is classified and not available to the general public, but there are indications that militant Islamic groups, such as Al Qaeda, are contemplating the use of threat agents and/or chemicals to compromise potable water supplies in the United States. Accordingly, managers and supervisors of water utilities may find it prudent to have some knowledge of potential threat agents, how they can be detected, and how they may be removed or inactivated when found.

THREAT AGENTS THAT MAY BE INTRODUCED INTO WATER SUPPLIES

Traditional lists of biological threat agents include both viruses and bacteria and the toxins that may be elaborated by some of these organisms. The list of viruses usually includes hemorrhagic fever- and encephalitis-causing viruses such as Venezuelan equine encephalitis (VEE), Ebola virus (EB), and Rift Valley fever virus (RVF). Perhaps the most worrisome virus is smallpox (Variola), enormous quantities of which were created by a clandestine bioweapons program in the former Soviet Union. Most of the world's population lacks immunity to smallpox and would thus be exceptionally vulnerable to infection (1).

Bacterial agents (and their diseases) include *Yersinia pestis* (plague), *Bacillus anthracis* (anthrax), *Francisella tularensis* (tularemia), *Coxiella burnetii* (Q fever), and *Burkholderia pseudomallei* (glanders or melioidosis). The category of toxins includes staphylococcal enterotoxins A and B; the botulinum toxin of *Clostridium botulinum*; ricin; which is produced by the castor bean; and trichothecene mycotoxins, which are extracted from fungi and gained notoriety in the late 1970s as the "yellow rain" used by Soviet-supplied forces in battlefields in Cambodia and Laos (2).

SURVIVAL OF THREAT AGENTS IN WATER

Generally, little is known about the survivability of threat agents in chlorinated or nonchlorinated water, how readily they disperse in water, and whether exposure to water alters their infectivity for the worse (or better, as the case may be). Conducting experiments with virulent strains of threat agents almost always requires use of dedicated Biosafety Level 3/4 laboratory facilities, which places sizeable constraints on the ability to carry out assays to answer these questions. Therefore, it often is necessary to extrapolate data derived from work done with related species, or avirulent strains, of these threat agents.

Some bacterial species that can be used as threat agents are "naturally" encountered in water and have been implicated in small but well-studied disease outbreaks. A waterborne outbreak of tularemia occurred in Chlumcany, Czech Republic, in 2000; 96 people became ill and contaminated water, used for drinking and food preparation, was the vehicle of infection (3). Other documented cases of tularemia linked to contaminated water were observed in Kosovo (4) and Norway (5), and it was hypothesized that an outbreak of oropharyngeal tularemia in Turkey in 2001 was caused by ingestion of nonchlorinated reservoir water (6). Laboratory experiments conducted by Forsman et al. (7) found that cells could be cultured in water devoid of carbon sources and maintained at 8 °C for up to 70 days postinoculation, indicating that *F. tularensis* can survive immersion in an aquatic environment for an appreciable length of time.

Outbreaks of melioidosis have been linked to the ingestion of water contaminated with *Burkholderia pseudomallei*; in Australia in 1997, seven culture-confirmed cases were identified in a small (300 people) rural community served by capped borehole wells 2 km distant (8). Chlorine levels in potable water samples were undetectable prior to the presumed exposure period, reflecting failure to treat water adequately before consumption.

A helpful overview of the ability of conventional water chlorination/disinfection protocols to inactivate threat agents is summarized in Burrows and Renner (9). The ability of some pathogens, such as spores of *B. anthracis*, to remain viable following typical chlorination regimens is a troubling but unavoidable facet of dealing with incidents of bioterorrorism aimed at potable water distribution networks (Fig. 1).

Note that for some agents, exposure to chlorine is bacteriostatic rather than lethal; though log reductions in the numbers of viable bacteria may be achieved, a remnant population may survive and can grow when chlorine concentrations decrease. For example, after 30 minutes of exposure, two strains of *B. pseudomallei* survived exposure to 1000 ppm chlorine; another could survive 300 ppm, but not 500 ppm (10). The utility of alternative methods, for such as chloramines, ozone, peroxides, UV radiation, metal ions, and electrochemistry, water disinfection may offer distinct advantages over chlorine (11). However, the paucity of experiments using actual threat agents requires us to extrapolate results achieved with conventional waterborne pathogens to gain insight as to how well these methods would inactivate species such as *B. anthracis, F. tularensis*, and poxviruses.

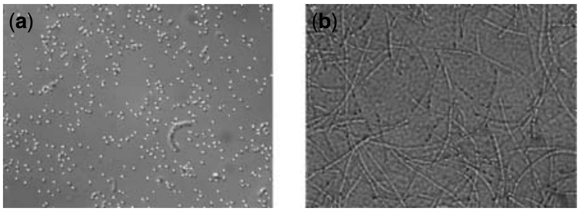

Figure 1. Nomarski interference contrast microscopy images of *Bacillus anthracis* Sterne strain spores: panel A, the comma-shaped organism in the center of the image is a vegetative cell, and vegetative cells (panel B), 1000X magnification (panel A photograph, courtesy of Jeffrey Karns).

MONITORING WATER FOR THREAT AGENTS: RECOVERY FROM SOURCE AND FINISHED WATERS

Successful monitoring of source and finished waters for threat agents is a two-component process: first, the organisms must be removed or isolated from the water samples, usually via filtration or other selective separation protocols. Second, the recovered organisms must be detected and their identity confirmed with a sufficient degree of reliability and accuracy to allow water utility managers to make decisions with confidence.

Approaches to establishing potable water security must deal with three segments of the water delivery system: the source water supply, the water purification/treatment facility, and the distribution network that brings water to individual buildings and homes. The water industry has considerable expertise in detecting select pathogens in source water, such as *Giardia* and *Cryptosporidium*, via protocols such as the U.S. EPA Method 1623, which involves capsule filter-mediated recovery of oocysts from 10 L volumes of water. However, source waters may contain appreciable numbers of naturally occurring flora that can confound detection efforts; they may also contain high concentrations of particulate matter that can impede filtration protocols for pathogen recovery from large volumes of water.

Recovery of viral threat agents presents greater difficulties than for bacterial agents because they are considerably smaller (except for poxviruses, which may approach 0.2–0.3 μm in length) and require culture in living cells to reproduce. As with protozoa, the water industry, private and academic laboratories, and agencies such as the U.S. EPA, have made considerable efforts to develop protocols for isolating enteric viruses from large volumes of water, and these protocols may be useful in dealing with agents such as poxviruses and hemorrhagic fever viruses.

Little is known about the behavior of biological toxins in source waters, particularly when large quantities of the toxins may be deliberately introduced into reservoirs close to intake pipes. The stability of toxins in such an environment is poorly understood. However, larger utilities continuously monitor source water for chemical contaminants (metals, organics, solvents, etc.) and provide some measure of confidence in the industry's ability to adapt preexisting assays for detecting threat toxins. Analytical chemistry protocols, such as gas or liquid chromatography and mass spectrometry, as well as immunoassays and bioassays, can be used in a comprehensive monitoring strategy (12).

Monitoring operations at the second and third segments of the water delivery system—the purification facility and the distribution network—are more difficult targets for bioterror attacks, because they have a more highly supervised structure and offer fewer opportunities for unobserved, deliberate introduction of pathogens. However, deliberate contamination at these segments of the delivery system has far more troubling aspects than for source waters because the infectious dose is comparatively more concentrated in finished water and this water is destined for direct (and indirect) consumption by the public. However, by containing disinfecting agents, "in plant" finished water is less likely to permit survival of infectious agents; by the time it reaches the household tap, potable water may support pathogen viability. For example, in experiments conducted in collaboration with Dr. Kalmia Kniel at the University of Delaware, we have found that raccoonpox virus can survive over two month's exposure in tap water, as measured by its ability to cause lytic infection of Vero cells. These experiments are preliminary, but they indicate a need to learn more about poxvirus survival in finished water (Fig. 2).

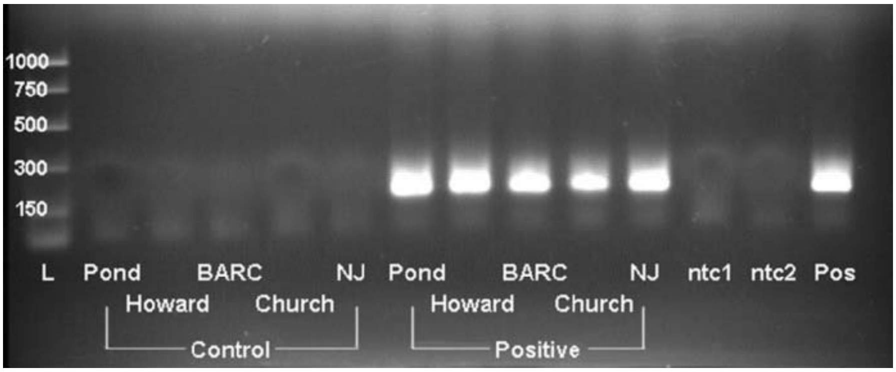

Figure 2. Polymerase chain reaction (PCR) assay conducted on Vero cell supernatant following infection with raccoonpox virus (RPV). RPV ($\sim 1 \times 10^6$ plaque-forming units) was incubated in 2 mL volumes of water from various sources: pond water and water from taps in Howard County, Maryland; the Beltsville Agricultural Research Center (BARC), Maryland; and a home in New Jersey, for 70 days at room temperature. Two hundred μL portions of the water were then mixed with a cell culture medium and applied to Vero cells for 1 hour; after that the medium was replaced and the Vero cells maintained at 37 °C and 5% CO_2. The supernatant (200 μL) was collected after 5 days and subjected to DNA extraction and nested PCR using the hemagglutinin primers of Ropp et al. (13). Control samples represent unspiked tap water samples. Lanes ntc1 and ntc2 represent PCR negative controls and lane Pos is the PCR positive control. Lane L is the DNA ladder, with rung sizes indicated.

MONITORING WATER FOR THREAT AGENTS: IDENTIFYING THREAT AGENTS

Once threat agents have been recovered from source or potable waters, they must be identified. Note Despite significant advances in detection technologies, sponsored in large part by U.S. government Department of Defense research initiatives, there is no single method currently that can reliably and reproducibly detect all potential threat agents in water samples in real time with any appreciable degree of sensitivity and specificity. The lack of such a method obviously presents the water industry with major challenges in monitoring potable water supplies for potentially harmful organisms and/or toxins.

Detection methods can be classified into three broad categories; the first includes traditional detection methods, which offer some advantages: they are often relatively inexpensive to perform, require modest training of the individuals performing the tests, and have standardized methods that have been reproduced by many laboratories over lengthy periods of time. For bacterial pathogens such as anthrax, traditional methods involving culture of vegetative cells on agar, the appearance and morphology of colonies on the agar, and the light microscopy features of the cells comprising the colonies offer a convenient and affordable means to monitor water for this agent (Fig. 3).

The same approach can be used with some degree of success for *Francisella and Burkholderia*, among others. As always, the type of water sampled can influence the sensitivity and specificity of the detection method; when closely related waterborne flora can be mistaken for threat agents, traditional methods must be augmented by other detection protocols to avoid generating false-positive results.

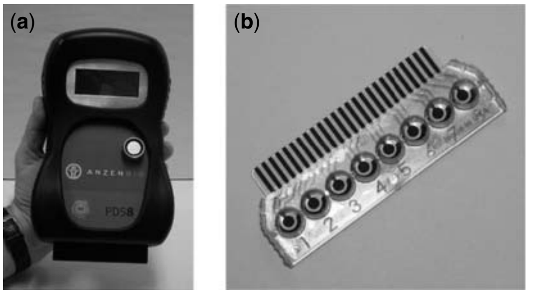

Figure 4. The Anzenbio PDS-8 instrument is a novel platform for field-based detection of toxins and pathogens. The detector (panel A) can be held in one hand and presents data to the end user via the inset LCD screen. The biosensor chip (panel B) is part of a disposable cartridge that is inserted into the detector. Up to six different samples (plus one positive and one negative control) can be analyzed on the chip, using a proprietary antibody-mediated, electrochemical format. The assay can be completed in as few as 2 hours (photo courtesy of Linda Williams, Anzenbio).

The second category of detection approaches involves immunologic methods, which are mediated by antibodies; the formats for such methods can be quite varied and include enzyme-linked immunosorbent assay (ELISA), immunochromatographic lateral flow assay, immunomagnetic separation-electrochemiluminescence (IM-ECL), and time-resolved fluorescence assay (14). Compared to traditional protocols, antibody-based detection methods provide greater specificity and sensitivity, timeliness, and ease of

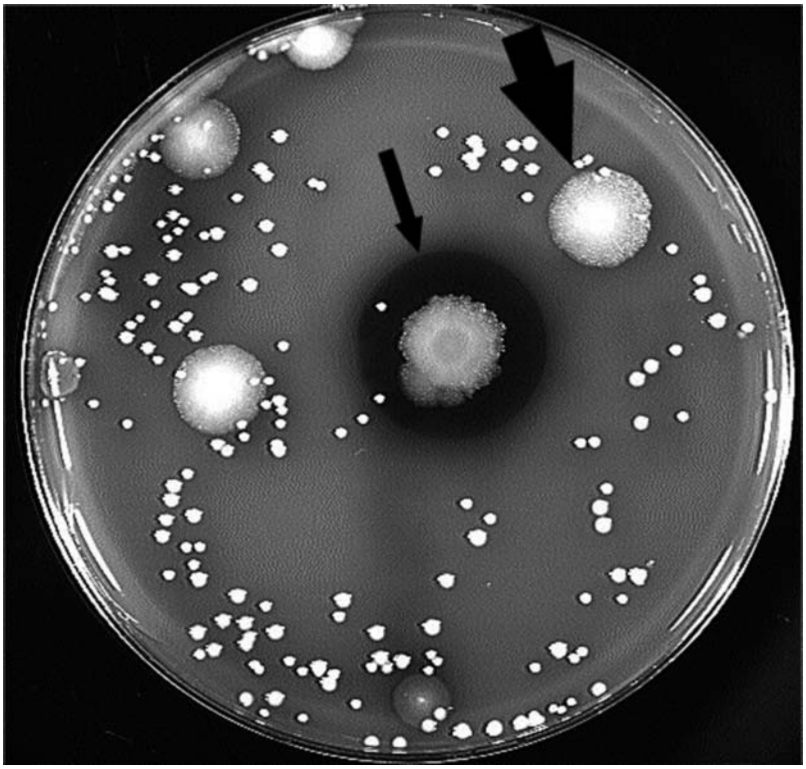

Figure 3. Photograph of bacterial colonies grown from stream water spiked with spores of *B. anthracis* Sterne strain on red blood cell agar plates. The large black arrow indicates a colony of *B. anthracis*; the smaller black arrow indicates a colony of a related *Bacillus* spp. Note the lytic zone surrounding this colony, a feature that can be used to differentiate between colonies of *anthracis* and those of other spore-forming members of waterborne flora.

use. Emerging technologies in this field seek to miniaturize the assay to the point where it can be done in the field or on site; one example of such an emergent technology is provided in Fig. 4.

Immunologic methods for the detecting bacterial, protozoal, and viral pathogens have been reported in the literature; in one of the more recent reports, Tims and Lim (15) detected purified *B. anthracis* spores in less than one hour, using a portable fiber-optic biosensor. Immunologic methods are particularly useful in detecting toxins, which are often proteins and thus not easily amenable to molecular biology-based detection (16).

Immunologic assays do possess some drawbacks. The specificity of the antibodies used to mediate detection is a critical factor; too poor a specificity means that false-positive signals can be generated from organisms that are closely related to the pathogen of interest. And, in general, antibody-based assays may not provide as great a degree of sensitivity as their molecular counterparts.

Molecular biology-based assays, such as PCR, have come into the forefront of pathogen detection technologies. By targeting regions of DNA intrinsic to particular species or strains of microorganisms, molecular methods can provide a high degree of specificity and sensitivity. There is a substantial and ever-increasing body of scientific literature on the use of molecular techniques for detecting and characterizing conventional waterborne pathogens. Many water quality laboratories in the United States, Europe, Australia, and Asia now use these methods routinely and readily adopt new instrumentation and procedures as their testing needs evolve.

Drawbacks of molecular detection techniques include high setup and maintenance costs, which translate into a higher cost per assay than traditional protocols, and the need for more training for laboratory personnel. Nonetheless, the enhanced detection capability of molecular protocols has led to their adoption by U.S. Department of Defense agencies as the first choice for threat agent identification. A detailed recitation of existing threat agent detection protocols is beyond the scope of this article; readers are directed to the paper by Firmani and Broussard (17) for a comprehensive overview of this evolving topic.

FUTURE DIRECTIONS FOR BIOSECURITY OF WATER RESOURCES

As this article is written in early fall 2004, there are encouraging signs that government agencies responsible for regulating water quality are moving to enhance the ability of the industry to respond to threats directed against water resources.

One of these efforts is being mounted by the U.S. EPA, which in July 2004 began testing three competing platforms for the rapid detection of threat agents in water under the aegis of the Environmental Technology Verification (ETV) program. All three platforms, the Idaho Technology, Inc. Ruggedized Advanced Pathogen Identification Device (RAPID), the Applied Biosystems 7000 Sequence Detection System, and the Agilent Bioanalyzer used by Invitrogen Corporation, rely on PCR to mediate amplification and identification of pathogen DNA in real time. These instruments will be tested by the Battelle laboratory's Advanced Monitoring Systems Center using the bacterial threat agents *Bacillus anthracis, Francisella tularensis, Brucella suis, E. coli*, and perhaps *Yersinia pestis*. The platform performing best in sensitivity and specificity may be designated the rapid detection method of choice for utilities and agencies contemplating enhancements of their biosecurity capabilities.

Other U.S. EPA initiatives, such as the Water Quality Protection Center and the Drinking Water Systems Center, are working to provide the industry with up-to-date technologies for providing safe and wholesome water to their customers. Industry organizations such as the International Water Association, the American Water Works Association, the Water Environment Foundation, the European Water Association, and the Australian Water Association are actively funding research and development to provide improved biosecurity resources to their subscribers. All of these developments should assist water quality managers as they deal with the challenge presented by this potential threat to human health.

BIBLIOGRAPHY

1. McClain, D.J. (1997). Smallpox. In: *Medical Aspects of Chemical and Biological Warfare*. F.R. Sidell, E.T. Takafuji, and D.R. Franz (Eds.). Borden Institute, Washington, DC, pp. 539–559.
2. Wannemacher, R.W., and Wiener, S.L. (1997). Trichothecene mycotoxins. In: *Medical Aspects of Chemical and Biological Warfare*. F.R. Sidell, E.T. Takafuji, and D.R. Franz (Eds.). Borden Institute, Washington, DC, pp. 665–676.
3. Pazdiora, P., Moravkova, I., Nocarova, D., Velkoborska M., and Valeckova, K. (2002). A waterborne epidemic of tularemia in Chlumcany. *Epidemiol. Mikrobiol. Immunol.* **51**: 23–25.
4. Reintjes, R. et al. (2002). Tularemia outbreak investigation in Kosovo: Case control and environmental studies. *Emerg. Infect. Dis.* **8**: 69–73.
5. Hoel, T., Scheel, O., Nordahl, S.H., and Sandvik, T. (1991). Water- and airborne *Francisella tularensis* biovar palaearctica isolated from human blood. *Infection* **5**: 348–350.
6. Gurcan, S., Tatman-Otkun, M., Otkun, M., Arikan, O., and Ozer, B. (2004). An outbreak of tularemia in Western Black Sea region of Turkey. *Yonsei Med. J.* **45**: 17–22.
7. Forsman, M., Henningson, E., Larsson, E., Johansson, T., and Sandstrom, G. (2000). *Francisella tularensis* does not manifest virulence in viable but non-culturable state. *FEMS Microbiol. Ecol.* **31**: 217–224.
8. Inglis, T. et al. (1999). Acute melioidosis outbreak in Western Australia. *Epidemiol. Infect.* **123**: 437–443.
9. Burrows, W.D. and Renner, S.E. (1999). Biological warfare agents as threats to potable water. *Environ. Health Perspect.* **107**: 975–984.
10. Howard, K. and Inglis, T. (2003). The effect of free chlorine on *Burkholderia pseudomallei* in potable water. *Water Res.* **37**: 4425–4432.
11. Morato, J., Mir, J., Codony, F., Mas, J., and Ribas, F. (2003). Microbial response to disinfectants. In: *The Handbook of Water and Wastewater Microbiology*. D. Mara and N. Horan (Eds.). Elsevier, New York, pp. 657–693.

12. Gullick, R.W., Grayman, W.M., Deininger, R.A., and Males, R.M. (2003). Design of early warning monitoring systems for source waters. *J. Am. Water Works Assoc.* **95**: 58–72.
13. Ropp, S.L., Jin, Q., Knight, J.C., Massung, R.F., and Esposito, J.J. (1995). PCR strategy for identification and differentiation of smallpox and other orthopoxviruses. *J. Clin. Microbiol.* **33**: 2069–2076.
14. Peruski, A.H. and Peruski, L.F., Jr. (2003). Immunological methods for detection and identification of infectious disease and biological warfare agents. *Clin. Diagn. Lab. Immunol.* **10**: 506–513.
15. Tims, T.B. and Lim, D.V. (2004). Rapid detection of *Bacillus anthracis* spores directly from powders with an evanescent wave fiber-optic biosensor. *J. Microbiol. Methods* **59**: 127–130.
16. Kijek, T.M., Rossi, C.A., Moss, D., Parke, R.W., and Henchal, E.A. (2000). Rapid and sensitive immunomagnetic-electrochemiluminescent detection of staphylococcal enterotoxin B. *J. Immunol. Methods* **236**: 9–17.
17. Firmani, M.A. and Broussard, L.A. (2003). Molecular diagnostic techniques for use in response to bioterrorism. *Expert Rev. Mol. Diagn.* **3**: 605–616.

defmw1702mw166

GRANULAR ACTIVATED CARBON

Dinesh Mohan
Kunwar P. Singh
Gomti Nagar, Lucknow, Uttar Pradesh, India

Activated carbon is the generic term used to describe a family of carbonaceous adsorbents with a highly crystalline form and extensively developed internal pore structure. Activated carbon is distinguished from elemental carbon by the removal of all noncarbon impurities and the oxidation of the carbon surface (1). Activated carbon has the highest volume of adsorbing porosity of any substance known to humans (5 grams of activated carbon can have the surface area of a football field). It can be defined as (2):

> Activated carbon is a crude form of graphite, with a random or amorphous structure, which is highly porous, over a broad range of pore sizes, from visible cracks and crevices, to cracks and crevices of molecular dimensions.

The use of activated carbon is not new. The Egyptians used carbonized wood around 1500 B.C. as an adsorbent for medicinal purposes as well as a purifying agent. The ancient Hindus in India used charcoal for filtration of drinking water. However, the basis for industrial production of active carbons was established in 1900–1901 in order to replace bone char in the sugar refining process (3).

Active carbons can be prepared from a wide range of carbonaceous materials, which include coconut shells, wood char, lignin, petroleum coke, bone char, peat, saw dust, carbon black, rice hulls, sugar, peach pits, fish, fertilizer waste, and waste rubber tire. (Table 1). The range of raw materials is diverse and widespread and greatly influenced by the need to produce low-cost carbon. Among the most commonly used raw materials, precursors for the production of commercially activated carbons are wood (130,000 tons/year), coal (100,000 tons/year), lignite (50,000 tons/year), coconut shell (35,000 tons/year), and peat (35,000 tons/year) (4,5).

TYPES OF ACTIVATED CARBONS

Based on the physical properties, activated carbons can be classified into the following broad categories:

Powdered Activated Carbon (PAC)

These activated carbons are made in the powders or fine granules according to the requirement. Powdered activated carbons have a diameter between 15 and 25 μm. These carbons provide a large internal surface with a small diffusion distance.

Granular Activated Carbon (GAC)

Granular activated carbon (GAC) is commonly used for the purification of liquids and gases. Granular activated carbon adsorbs a vast variety of dissolved organic materials, including many that are nonbiodegradable. GAC removes organic contaminants from water/wastewater by the adsorption process of the attraction and accumulation of one substance on the surface of another. Granular activated carbons typically have surface areas of 500–2000 m^2/g, with some reported as high as 3000 m^2/g. Much of the surface area available for the adsorption in granular activated carbon particles is found in the pores within the granular carbon particles created during

Table 1. Source Materials Used for the Preparation of Activated Carbons

Bagasse
Bark
Beat-sugar sludge
Blood
Blue dust
Bones
Carbohydrates
Cereals
Coal
Coconut coir
Coconut shell
Coffee beans
Corn Cobs and corn stalks
Cottonseed hulls
Distillery waste
Fertilizer waste slurry
Fish
Fruit pits
Fuller's earth
Graphite
Human hairs
Jute stick
Kelp and seaweed
Leather waste
Lignin
Lignite
Lampblack
Leather waste
Molasses
Municipal waste
News paper
Nut shells
Oil shale
Olive stones
Palm tree cobs
peat
Petroleum acid sludge
Petroleum coke
Potassium ferrocynide residue
Pulp-mill waste
Reffination earth
Refinery waste
Rice hulls
Rubber waste
Saw dust
Scrap tires
Spent Fuller's earth
Sunflower seeds
Sugar-beet sludge
Tea leaves
Rubber Tires
Wheat straw
Wood

the activation process. GAC have relatively larger particle sixes than powdered activated carbon and therefore provide a smaller external surface. These carbons are preferred for all the sorption of gases and vapors. GAC are also used in water/wastewater treatment, deodorization, decolorization, and separation of components in flow systems.

Spherical Activated Carbon (SAC)

These carbons are prepared from small spherical balls wherein pitch is melted in the presence of naphthalene or tetorlin and converted into spheres. These spheres are then contacted with naptha solution, which extracts naphthalene introduced into the porous structure. These porous spheres are then heated between 100 and 400 °C in the presence of an oxidizing agent. The oxidized spheres are then heated between 150 and 700 °C in the presence of ammonia to introduce nitrogen into spheres followed by activation in steam or CO_2.

Impregnated Activated Carbon (IAC)

In chemical activation, a catalyst may be impregnated into the feedstock. The most commonly used chemical activants include $ZnCl_2$, H_3PO_4, H_2SO_4, KOH, K_2S, and KCNS. In this process, a near-saturated solution of catalyst-impregnated feedstock is dried to influence pyrolysis in such a way that tar formation and volatilization can be kept at a minimum. The resulting product is then carbonized. Silver impregnated activated carbons are used for purification of domestic water.

Polymer Coated Activated Carbons (POAC)

In this process, porous carbon can be coated with biocompatible polymers resulting in a smooth and permeable coat without blocking the pores. It is well documented in literature that activated carbons possess a highly developed porous system. These pores are produced during the activation process of carbonized residue when spaces between elementary crystalline are cleared of carbonaceous compounds and nonorganized carbon.

The precursors used for the production of activated carbons have a large effect on the pore size distribution, surface area, and other physical and chemical properties. Table 2 sums up some of the basic differences between the raw materials used for the production of some important activated carbons, whereas the properties of some commercially available carbons with their sources as collected from literature are presented in Table 3.

METHODS FOR ACTIVATED CARBON DEVELOPMENT

The methods for the development of activated carbons are nearly as widespread as their potential uses and source materials (6). However, the basic steps most commonly used in the preparation of activated carbons are precursor material preparation, palletizing, low-temperature carbonization, followed by chemical or physical activation (Fig. 1). A number of methods were used for the preparation of activated carbons from waste materials from time to time using different activation parameters. Different steps/activation parameters used for the preparation of some of the activated carbons are presented in Table 4. Although the list is not complete, it will provide a general idea of the different methods used for the production of activated carbons.

Raw Materials

The selection of an appropriate raw material (Table 1) for the development of granular activated carbon is the most important aspect. The following points must be considered before selecting any raw material for the production of activated carbons:

1. Industrially inexpensive materials with high carbon and low inorganic content should always be preferred.
2. The impurities in raw materials should be kept at a minimum because after the activation process, many of these may be present in the carbon at higher concentrations than the precursors materials.
3. Importance should be given to the precursors having high density and sufficient volatile content. The volatile results in porous char, whereas high density favors the enhancement of structural strength of the carbon needed to withstand excessive particle crumble during use.
4. The raw material should be available in abundance locally.

The raw materials used for the preparation of activated carbons vary with their applications. A comparison of some of the raw materials is presented in Table 2.

Table 2. Basic Differences in Precursor Materials Used for the Production of Activated Carbons

Raw Materials	Percent Carbon	Percent Volatile Matter	Ash Percent	Density (kg/m3)	Activated Carbon Texture	Applications
Hard wood	40–42	55–60	0.25–1.2	0.50–0.8	Soft with large pore volume	Liquid phase adsorption
Soft wood	40–45	55–60	0.25–1.0	0.40–0.50	Soft with large pore volume	Liquid phase adsorption
Nut shells	40–45	55–60	0.40–0.60	1.4	Hard with large multi pore volume	Vapor phase adsorption
Lignite	50–70	25–40	5–6	1.0–1.40	Hard with small pore volume	Liquid phase adsorption
Soft coal	60–80	25–30	2–12	1.25–1.50	Medium hard with medium micropore volume	Liquid and vapor phase adsorption
Semihard coal	70–75	1–15	5–15	1.45	Hard with large pore volume	Vapor phase adsorption
Hard coal	85–95	5–10	2–15	1.50–2.0	Hard with large pore volume	Vapour phase adsorption

Table 3. Properties of Some Selected Activated Carbons Gathered from the Literature

Types of Carbons	Raw Material	Surface Area, BET (m^2/g)	Total Pore Volume (ml/g)	Iodine Number (mg/g)	Ash (%)	Uniformity Coefficient	pH_{zpc}	Moisture as Packed (%)	Effective Size, mm	Apparent Density
F-100	Bituminous coal	850–900	—	850		2.1	9.0	2	0.8–1.0	—
F-200	Bituminous coal	714	—	850		1.9	8.2	2	0.55–0.75	—
F-300	Bituminous coal	950–150	0.85	900	9	2.1	9.8	2	0.8–1.0	0.5 g/cc
F-400	Bituminous coal	1050–1200	0.94	1000	5.4	1.9	10.40	2	0.55–0.75	27
F-816	Bituminous coal	—	—	900	9	1.4	—	2	1.3–1.5	—
F-820	Bituminous coal	—	—	900	—	1.5	—	2	1.0–1.2	—
Centaur HSV	Bituminous coal	—	—	800	7	—	—	4	—	0.56 g/cc
Nuchar SN	Wood	1400–1800	—	900	3–6	—	4.0	10	—	337–369 (kg/m^3)
Nuchar SA	Wood	1400–1800	—	900	3–6	—	4.0	10	—	337–369 (kg/m^3)
Nuchar W	Wood	1400–1600	—	900	—	—	—	10	—	240–305 (kg/m^3)
HD-4000	Lignite coal	625	0.93	647	23	1.44	—	8	0.74	0.40 g/ml
Draco KB	Hard wood	1500	1.8	—	2	—	—		—	0.45 g/cc
Norit GAC 840 R	Reactivated Carbon	—	—	800	—	—	—	2	—	0.48 g/cc
Norit PAC 20 B	Coal	—	—	800	—	—	—	3	—	
HD-C	Lignite coal	556		500		—	—	4	—	0.51 g/cc
Norit GAC 1240	coal	1100	0.95	1020		1.8	—	2	—	0.50 g/cc
Barnebey & Sutcliffe SE	Coconut shell	1100–1200	—	1050	2–3	—	—	—	—	—
Barnebey & Sutcliffe PE	Coconut shell	1100–1200	—	1050	2–3	—	—		—	—
Barnebey & Sutcliffe KE	Coconut shell	1150	—	—	5	—	—		—	0.48 g/cc
Barnebey & Sutcliffe UU	Coconut shell	1150	—	—	5	—	—	5	—	0.4–0.8 g/cc
Pica POU/POE GX203	Coconut shell carbon	700–2200	—	1150	5	—	—	5	—	0.45–0.54 g/cc
Pica POU/POE NC 506	Coconut shell carbon	700–2200	—	1100	4	—	—	3	—	0.41–0.45 g/cc
Cameron PACarb	Resinous wood charcoal	—	—	500	8	—	—	3	—	0.45 g/cc
Selecto ABA 4000 LC		300–500	0.35–0.55	—	—	—	—		—	—
Witco 517	Petroleum	1050	—	1000	0.5	1.4	—	1	0.89	0.52 g/cc

Carbonization

Carbonization, sometimes called charring, converts the organic material into primary carbon, which is a mixture of ash, tars, amorphous carbon, and crystalline carbon.

In carbonization, the material is heated slowly in the absence of air. In this process, most of the noncarbon elements, hydrogen, and oxygen are first removed in gaseous form by pyrolytic decomposition of the starting materials. The important parameters, which determine the quality and quantity of the carbonized product, are (a) rate of heating, (b) final temperature, and (c) soaking time.

Activation

The activation is carried out basically to enlarge the diameters of the pores, which are created during the carbonization process, and to create some new porosity, which results in the formation of a well-defined and readily accessible pore structure with large internal surface area. During the activation process, the spaces between the elementary crystallites become cleared or less organized loosely bound carbonaceous material. The resulting channels through the graphitic regions, the spaces between the elementary crystallites, together with fissures within and parallel to the graphite planes constitute the porous structure, with large internal surface area (24). There are two types of activation, which are used to impart a porous structure within a starting material of relatively low surface area, namely thermal/physical or chemical activation.

Physical or Thermal Activation. Physical or thermal activation occurs after initial treatment and palletizing; it involves carbonization at 500–600 °C to eliminate the bulk of the volatile matter followed by partial gasification using mild oxidizing gas such as CO_2, steam, or fuel gas at 800–1000 °C to develop the porosity and surface area (25). An example includes the gasification of the carbonized material with steam, and carbon dioxide occurs by the following endothermic reactions:

$$C + H_2O \longrightarrow CO + H_2 \text{ (29 kcal)}$$

$$C + CO_2 \longrightarrow 2CO \text{ (39 kcal)}$$

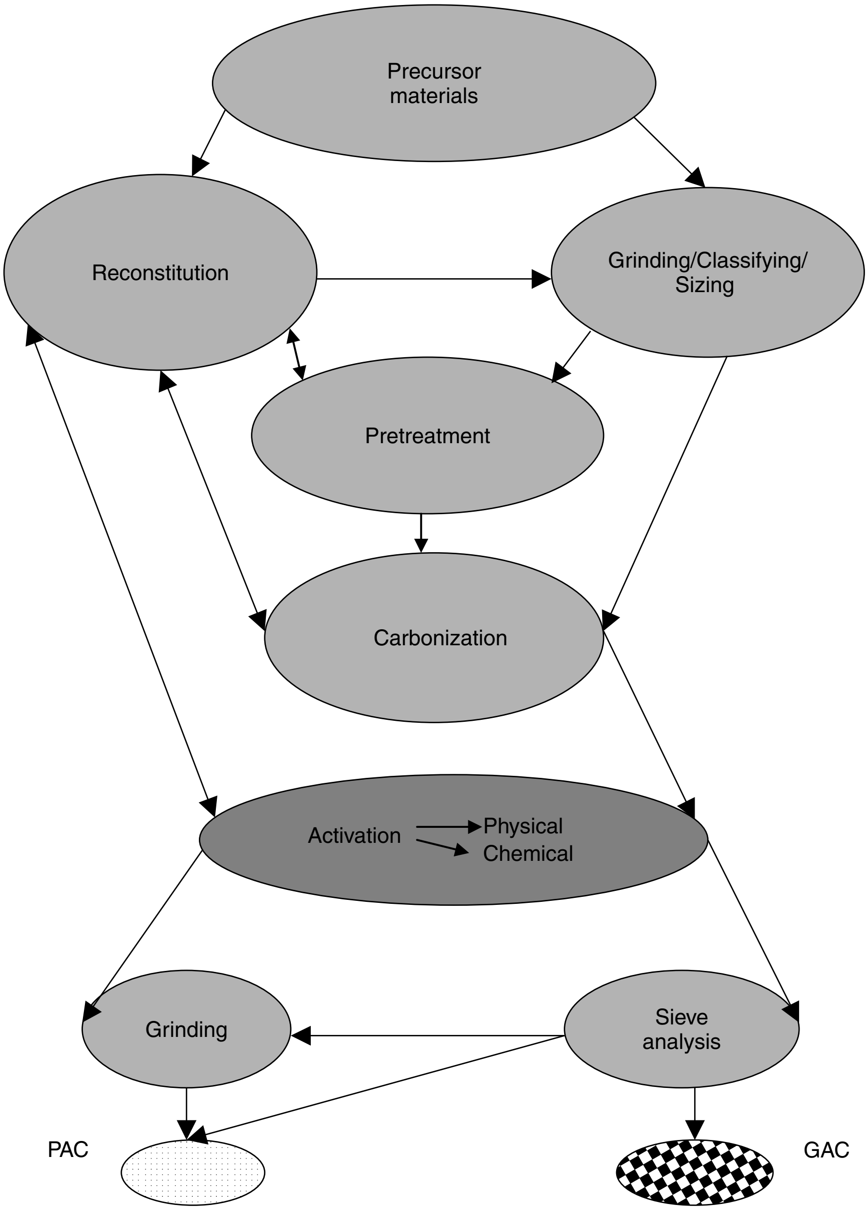

Figure 1. Various steps used in the activated carbon production.

$$CO + H_2O \longrightarrow CO_2 + H_2 \text{ (10 kcal)}$$

The H_2O molecule is smaller than the CO_2 molecule and thus diffuses faster into the pores of the carbon. Consequently, a reaction with steam is faster than that of CO_2. It has been reported in literature that a decrease in the reaction rate with CO_2 activation on carbon-containing wastes is nearly two times less than that of steam. When air or oxygen is used as an activating agent, problems develop because of the exothermic nature of the reactions of carbon with air (oxygen), and thus it is difficult to control. Despite these problems, several researchers have used the same for the activation of their products.

Chemical Activation. The second type of activation involves the incorporation of inorganic additives or metallic chlorides, such as zinc chloride or phosphoric acid, into the precursor before the carbonization (6). It has been reported that carbons with well-developed porous structure, mainly meso and microporous, can be produced by $ZnCl_2$ incorporation. KOH activation has also been shown to successfully increase the surface area and pore volume of active carbons (20,26)

Many other chemicals, such as ammonium salts, borates, calcium oxide, ferric and ferrous compounds, manganese dioxide, nickel salts, hydrochloric acid, nitric acid, and sulfuric acid, have also been used for the activation purpose.

The basic difference between physical and chemical activation is the number of stages required for activation and the temperature at which the activation takes place. Chemical activation is a one-step process, whereas physical activation is a two-step process, including carbonization and activation. The temperature required in physical activation (800–1000 °C) is higher than that of chemical activation (200–800 °C).

Table 4. Some Chemical Activant-Feedstock Couples [Extended Form of (4)]

Feedstock	Activant	Conditions	Reference
Coconut shell	Con. H_2SO_4	1.5 parts by weight H_2SO_4 for 24 hrs at 140–160 °C: steam activation at 1 Kg/m^2 pressure for 30 min.	7
Agricultural by products such as almond shell, olive stones and peach stones	—	Heating in CO_2 at 1123K	8
Coconut shell	Con. H_2SO_4	1.0 parts by weight H_2SO_4 for 24 h at 150 °C	9
Fertilizer slurry	H_2O_2/H_2O, N_2	450 °C, 1 h	10
Palm tree cobs	H_3PO_4/H_2SO_4	730 °C, 6 h	11
Coconut shell	H_3PO_4	450 °C	12
Petroleum coke	KOH/H_2O	700–850 °C, 4 h	3
Raffination earth	$H_2\ SO_4$	10% v/v,350 °C	13
Algerian coal	KOH/NaOH	930 °C	14
Pine saw dust	$Fe(NO_3)_3/CO_2$	850 °C, 1 h; 825 °C, 6 h	15
Almond and pecan shells	H_3PO_4	Chemical activation with H_3PO_4/Physical CO_2	16
Eucalytus woodchars	—	CO_2 activation, 400–800 °C	17
Bituminous coal	$ZnCl_2$	N_2/400–700 °C	18
Coal or coconut shell		Phosgene or chlorine gas at 180 °C	19
Petroleum coke	KOH	Dehydration at 400 °C followed by activation in 500–900 °C	20
Lignite	Na_2MoO_4/ $NaWO_4$/ NH_4VO3/ $(NH_4)_2MoO_4$/ $FeCl_3$/ $Fe(NO_3)_3$	Inert atmosphere/600–800 °C	21
Peanut hulls	H_2SO_4	150 °C, Sodium bicarbonate	22
Coconut shell and cocnut shell fibers	H_2SO_4	600 °C	23

Thus, the activated carbons can be prepared by either physical/chemical activation or a combination of both with well distributed porosity and high surface area.

Characterization of GAC

The effectiveness of activated carbon as a unit process in the treatment of water/wastewater/industrial effluents requires the best selection of an appropriate activated carbon. Following are some of the important properties used for the selection of suitable activated carbon for specific treatment:

Density. The density of carbon can be expressed in several different ways. The two most important among these are apparent and particle densities.

Apparent Density. Apparent density has little role in initial evaluation of an activated carbon, but it has an important role in the regeneration process. It is defined as the mass of carbon per unit volume that can be packed into an empty column. It is expressed in grams per cubic centimeter or pounds per cubic foot.

Particle Density. Particle density is defined as the ratio of the mass of dry and unloaded carbon particles to the total volume of the particles, including pore volume.

Higher density GAC is preferred because of the following reasons:

1. High density GAC has more carbon structure.
2. The high density GAC figured that, for each cubic meter of volume, more GAC could be installed.
3. Bituminous coal-based GAC provides a much denser material than lignite or subbituminous coal.

Moisture. It is important only for shipping and manufacturing purposes. It is defined as the percent by mass of water adsorbed on activated carbon.

Hardness and Attrition. It is an important factor in system design, filter life, and product handling. Large differences exist in the hardness of activated carbons, depending on the raw materials used and activity level (Tables 2 and 3). It is defined as the resistance of a granular activated carbon to the degradation action of steel balls in a Ro–Tap machine. It is calculated by using the mass of granular carbon retained on a particular sieve after the carbon has been with steel balls.

Sieve Analysis. It is very useful in carbon production evaluation. It is also important in the evaluation of purchased carbon to the specifications. The distribution of particle sizes in a given sample is obtained by mechanically shaking a weighed amount of activated carbon through a series of test sieves and determining the quantity retained by or passing given sieves.

Abrasion Number. The abrasion number is important in evaluating the ability of a carbon to withstand attrition. It is defined as a measure of the resistance of the particles to degrade on being mechanically abraded, which is measured by putting a carbon sample with steel balls in a pan on a Ro–Tap machine. In other words, it is defined as the ratio of the final average (mean) particle diameter to the original average (mean) particle diameter times 100.

Ash Percent. The ash percent is very important in the evaluation of the raw materials and the manufacturing process. It is the residue that remains when the carbonaceous portion is burned off. The ash contained mainly has silica, alumina, iron, magnesium, and calcium. Ash in activated carbon is not desirable and is considered to be an impurity. It can be measured by the change in weight by burning the carbon sample to constant weight at 800 °C.

pH. The effect of the carbon on pH of a volume of water is very much dependent on the relative quantities of both activated carbon as well as water. The pH effect can be studied by putting 1.0-gram carbon with 50 ml of de-ionized water and heating to 90 °C followed by cooling to 20 °C and measuring the pH of the supernatant.

Effective Size, Mean Particle Diameter, and Uniformity Coefficient. It is used to establish the hydraulic conditions of an adsorber column. Measurement of the gradation of carbon particles plays an important role in the evaluation of head loss in the flow through granular beds. The mean particle diameter can be calculated as follows:

First the average particle diameter (d) of each sieve used for the particle distribution is calculated by using the following equation:

$$d = w \times n$$

where w is the percent of the total weight retained by a particular sieve and n is the average of the mesh opening of the sieve that contained w and the sieve used immediately above.

By using this d, the mean particle diameter can be calculated as:

$$\text{Mean particle diameter} = \text{sum of all diameters divided by 100}$$

It is very important to note here that the smallest sieve is not considered in calculating the mean particle diameter.

The effective size can thus be calculated by using the cumulative percentage of carbon passing each size and plotting the sieve opening in millimeters (ordinate) versus the cumulative percentage (abscissa) on a semilogarithmic scale. Thus, the effective size is the opening in millimeters, which passes 10% of the total material. The 60% passing sieve size divided by the effective size gives the uniformity coefficient. The 50% passing sieve size is approximately the mean particle diameter. The uniformity coefficient is basically a dimensionless factor, which indicates the degree of uniformity of GAC. A value of one indicates that all parameters are identical in size, whereas greater values relate to a higher degree of variation.

Surface Area. Total surface area is very important to characterize porous solids. Large surface area is generally a requirement for a good adsorbent. However, the total surface area has to possess adequate pore size distribution and surface chemistry to adsorb the targeted species. Surface area determination relies on the accurate knowledge of the average area. It is determined by the sorption of nitrogen gas into the carbon and is expressed in square area per gram of carbon. The most widely accept method is the BET nitrogen adsorption method.

Pore Size Distribution. Activated carbon is a complex network of pores of different sizes and shapes. The shapes include cylinders, rectangular cross sections, and other irregular shapes and constrictions. The identification of different sizes is called the pore size distribution.

The pore size distribution is very much dependent on the source materials and method of activation. According to IUPAC accepted criteria (27), activated carbons have a trimolecular distribution of pore sizes (Fig. 2), as discussed below.

Macropores. These are the pores having diameter greater than 50 nm (Fig. 2). Their pore volume varies from 0.2 to 0.5 cm^3/g and surface area lies between 0.5 and 2 m^2/g. Such low value of surface area renders them of little use in adsorption except for large adsorbate molecules.

Mesopores. These pores have diameter ranging between 2 nm and 50 nm (Fig. 2). The pore volume of these pores varied from 0.02 to 0.01 cm^3/g. The surface area constituted by these pores lies between 10 and 100 m^2/g. The capillary condensation with the formation of a meniscus of adsorbate mainly takes place in these pores.

Micropores. These pores have diameter less than 2 nm (Fig. 2). Their size generally corresponds to that of molecules. The value of these pores ranges from 0.15 to 0.5 cm^3/g and surface area lies between 100 and 1000 m^2/g. They constitute about 95% of the total surface area. These pores are of great significance as far as adsorption studies are concerned.

Besides these, two different types of pores are exist in some of the carbons.

Ultramicropores. These pores have diameter <0.7 nm.

Supermicropores. These pores have diameter between 0.7 and 2 nm.

In practice, total pore volume is calculated by measuring three distinct volumes associated with one gram of the material at a constant temperature. These three distinct volumes are:

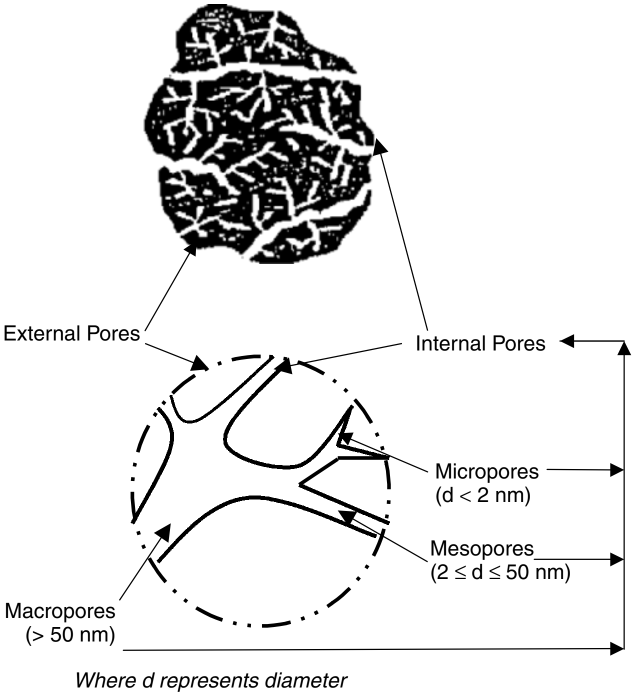

Figure 2. Schematic activated carbon structure.

(a) the volume of its purely solid structure
(b) the external geometrical volume of the solid
(c) the volume of a given fluid that it can displace

Thus, the total pore volume can be given by the difference between (b) and (a).

Iodine Number. It is defined as the number of milligrams of iodine adsorbed by one gram of activated carbon when the iodine concentration of the residual filtrate is 0.02N.

The iodine number provides an indication of the amount of small pores in carbon. It is correlated with the surface area in pores with diameters less than 10 Å. As iodine is a small molecule, it provides an indication of a particular carbon's capacity to adsorb smaller molecules.

Molasses Number. The molasses number represents the amount of large pores in carbon. It is calculated from the ratio of the optical densities of the filtrate of a molasses solution treated with a standard activated carbon and one in question. The molasses number can be correlated with the surface area in pores with diameters greater than 10 Å.

Both the iodine number and the molasses number decrease with time as adsorption occurs.

Methylene Blue Number. It is defined as the number of milligrams of methylene blue adsorbed by one gram of activated carbon in equilibrium with a solution of methylene blue dye having concentration 1.0 mg/L.

Phenol Number. Phenol number has been found to be a less sensitive test as it is very much effected by pH of the actual water used. It is very well documented in literature that at low pH, phenol adsorption is much greater than at neutral or alkali pH.

SURFACE CHEMISTRY OF ACTIVATED CARBON

The type of starting material and the method of activation used in producing activated carbon determines the nature of various surface functional groups. The surface chemistry of carbon has been studied extensively by several investigators from time to time. The nature of the carbon surface basically depends on the conditions and temperatures employed during the activation process. Activated carbon can be divided into two major types, namely acidic (L) and basic (H), according to Steenberg's classification (1) and can be defined as follows:

L Carbons

These carbons are prepared by heating the raw materials at about 200–400 °C in the presence of air. These carbons assume a negative charge (ionized) upon hydration and, thus, yielding acidic pH, are hydrophilic in nature and can neutralize a strong base. These carbons generally develop acidic surface oxides and lower the pH value of neutral or basic solutions. They primarily sorb bases from the solutions and exhibit a negative zeta potential. Wood-based activated carbons are usually L type in nature. The predominant surface functional groups are carboxyl, phenolic hydroxyl, carbonyl (quinone type), carboxylic acid, anhydrides, lactone, and cyclic peroxide (28,29). It has been reported that carboxylic and lactone groups tend to discourage the adsorpion of many aromatic compounds (29). Carbonyl groups in the form of quinone structures, and hydroquinone groups, however, can enhance the adsorption of aromatic compounds by the formation of an electron donor-acceptor complex between the aromatic ring and the surface carbonyl groups (30,31).

H Carbons

The activated carbons are prepared by heating the raw materials at about 800–1000 °C in the absence of air or in the presence of CO_2 followed by exposure to air at room temperature. These carbons assume a positive charge (protonated) upon introduction into water, thus yielding alkaline pH. They are hydrophobic in nature and can neutralize strong acids. These carbons generally develop basic surface oxides and increase the pH value of neutral or acidic solutions. They primarily sorb strong acids from the solutions and exhibit a positive zeta potential. Coconut shell-based and dust coal activated carbons are usually H type in nature. The predominant surface functional groups on the surface of the carbons are lactones, quinones, phenols, and carboxylates. More reports of the chromene (benzpyran) groups on the surface of the carbon exist (32). Boehm and Voll (33) suggested that basic surface oxide may be represented by pyrone-like structures.

It is interesting to note that H-type carbons can be converted into L-type carbons when they are oxides by chemical oxidants or aged in the atmosphere.

The acidic groups on activated carbons are believed to be one of the most important properties of activated carbons for metal ions adsorption (34). The L-type carbon is a stronger solid acid than the H-type carbon, therefore L-type carbon is more efficient for the adsorption of heavy metal ions. Surface area is not a primary factor for adsorption on activated carbon. High surface area does not mean high adsorption capacity, as reported by Perrich, (35), because of the following factors:

(a) In adsorption, only the wetted surface area is effective, which is never equal to the total surface area.
(b) In some adsorption processes, the material to be adsorbed is too large that it cannot enter the small pores where the bulk of the surface area exists.
(c) Data on the surface area, pore volume, and surface nature usually have not been correlated with data on the material to be adsorbed.

The physical and chemical properties of the carbons are very much dependent on the properties of raw materials, method of carbonization, and activation. The properties of some selected carbons developed from different sources are given in Table 3. These properties are collected from various sources including books, literature, and company brochures.

ADSORPTION EQUILIBRIUM

To determine the ultimate adsorption capacity, the solutes are brought in contact with a given amount of activated carbon or any other adsorbent in a closed system. If adsorption is the dominating removal mechanism, then the residual concentration will be reached that will remain unchanged with time, which is also know as equilibrium concentration. This process is known as adsorption equilibrium.

The adsorption process is essentially an attraction of gaseous or liquid adsorbate molecules onto a porous adsorbent surface. In gaseous separations, activated carbon is often used to remove odors and impurities from industrial gases, to recover valuable solvent vapor, and to dehumidify air and other gases. In liquid separations, activated carbon can be applied for removing the taste and odor from water, decolorizing, and treating industrial wastewater containing organics, dyes, and heavy metal ions.

There are two types of adsorption processes.

Physical Adsorption

Physical adsorption is a reversible phenomenon. It results from the action of van der Waals forces, comprised of London dispersion forces and classic electrostatic forces of attraction between molecules of the carbon and the substance adsorbed (adsorbate). Physical adsorption is usually dominant at low temperatures. This type of adsorption is usually multilayered; i.e., each molecular layer forms on top of the previous layer, with the number of layers being proportional to the contaminant concentration.

Chemisorption

It is the result of chemical interaction between the carbon surface and the adsorbed substance. It usually involves strong bonds, and is therefore irreversible.

A number of factors exist that affect adsorption, including chemical properties of adsorbate and activated carbon, pH, and temperature of the adsorbate. Chemical adsorption is usually dominant at high temperatures because chemical reactions proceed more rapidly at elevated temperatures than at low temperatures.

The following factors affect the sorption of organics on activated carbons:

Hydrocarbon Saturation: Double- or triple-carbon bond (unsaturation) organics are adsorbed more easily than single-carbon bond (saturated) organics.
Molecular Structure: Branch-chain organics are adsorbed more easily than straight-chain organics.
Molecular Weight: Larger molecules are generally adsorbed more easily than smaller molecules.
Polarity: Less polar (or weakly ionized) organics are adsorbed more easily than polar (or strongly ionized) organics.
Solubility: Less soluble compounds are adsorbed more easily than more soluble compounds.

ADSORPTION *ISOTHERMS*

An adsorption isotherm is the relationship between the amount of a substance adsorbed on the activated carbon surface and the equilibrium concentration of dissolved adsorbate at a constant temperature and other conditions. An adsorption isotherm is an expression of the principle of microscopic reversibility, although adsorption can be irreversible. The most common method for gathering isotherm data is the bottle point experiment. These equilibrium data are formulated into an adsorption isotherm model. Brunauer et al. (36) pointed out that although the isotherms are different for all sorbents and sorbates, more or less common shapes of isotherms are observed. Six different types of isotherms (37) exist, which are named as type 1, type II, type III, type IV, type V, and type VI. The pictorial representation of the isotherms are given in Fig. 3.

Type 1

This type of isotherm is obtained from carbons having micropores only, which corresponds to monolayer adsorption as postulated by Langmuir. The volume of the gas adsorbed approaches a limiting value, just enough to complete a monomolecular layer even when the gas pressure is rather low. Further increases in pressure hardly produces any further increase in the amount of adsorption. These types of isotherms are typical of a microporous solid where only monolayer adsorption occurs. Examples include the

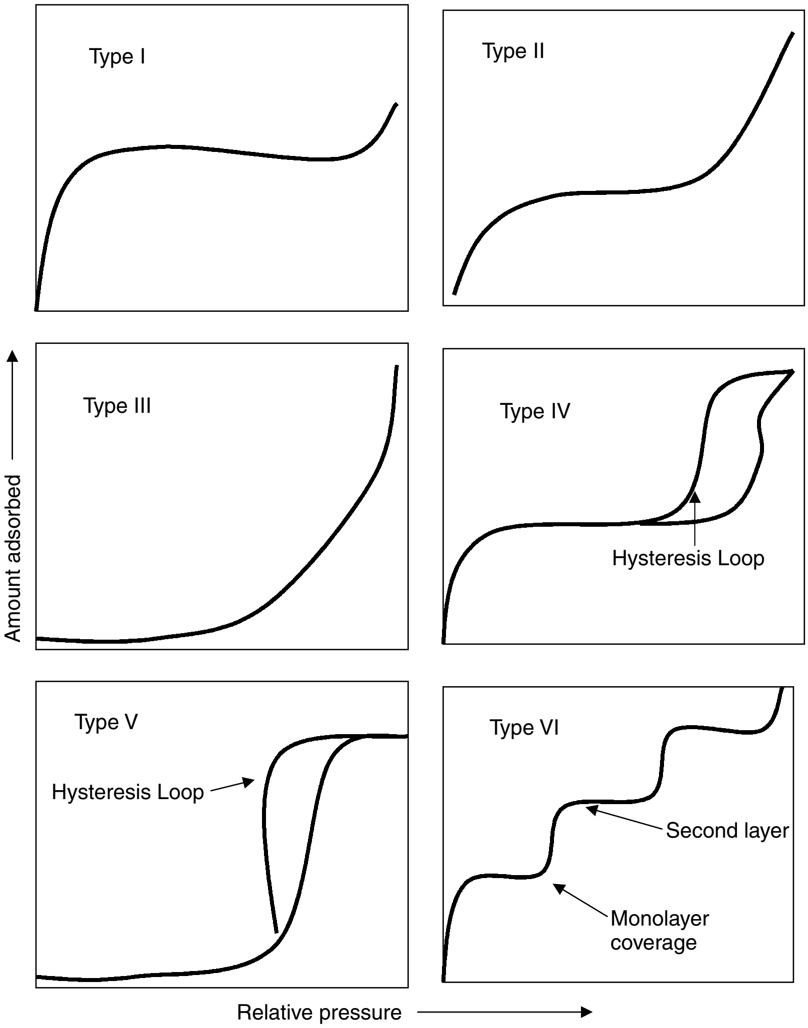

Figure 3. Type of isotherms.

adsorption of nitrogen or hydrogen on microporous carbon at temperatures close to −180 °C.

Type II

This type of isotherm describes the physical adsorption of gases by nonporous solids, and monolayer coverage is succeeded by multilayer adsorption at higher pressure. Examples include the adsorption of nitrogen on iron catalyst at −195 °C.

Type III

This type of isotherm is obtained from carbon having both micropores and mesopores. These types of isotherms show large deviations from the Langmuir model. The amount of adsorption keeps increasing in each case with an increase in pressure, which is attributed to the formation of additional layers of physically adsorbed gas molecules. It is considered that the gas molecules adsorbed in the first layer may hold by van der Waals forces a second layer of gas molecules, which, in turn, may hold a third layer and so on. The examples of type III isotherms include the adsorption of bromine on silica or alumina gel at 80 °C.

Type IV

This type of isotherm is obtained from both nonporous and mesoporous solids.

These are observed in the cases where a possibility of condensation of gases exists within the narrow capillary pores of the adsorbent. This phenomenon is also known as capillary condensation. Examples of type IV include adsorption of benzene on silica gel at 50 °C and that of water vapor on activated carbon at 100 °C.

Type V

This type of isotherm originates from microporous and mesoporous solids. Type V isotherms are basically related to type III isotherms and are very uncommon.

Type VI

This type of isotherm is obtained from uniform nonporous surfaces and represents stepwise multilayer adsorption. The sharpness of the steps depends on the system and the temperature. The step height represents the monolayer capacity for each adsorbed layer, and in the simplest case, it remains nearly constant for two or three adsorbed layers. Examples include the isotherms obtained with argon or krypton on graphitized carbon black at liquid nitrogen temperature.

APPLICATIONS OF ACTIVATED CARBONS

Activated carbons are produced in granular, powdered, and palletized forms and have a wide range of applications. The applications of activated carbon uses can broadly be divided into two categories.

Liquid-Phase Applications

Liquid-phase granular activated carbon adsorption (GACA) is an efficient, easy, and reliable treatment technology. It is very much different from gas-phase carbon as liquid-phase carbons have significantly more pore volume in the macropore range, which permits adsorbates to diffuse more rapidly into the micropores and mesopores. It is considered to be a best available control technology (BACT) by the U.S. Environmental Protection Agency (USEPA) and is a benchmark for other remediation technologies. In order for carbon adsorption to work well, it is important that the final design incorporate both the physical and adsorption process. In liquid-phase granular activated adsorption, activated carbon can be used either in powder, granular, or palletized forms. The average size of powder activated carbon ranges between 15 and 25 μm and are most frequently used in batch applications. On the other hand, the granular activated carbon particle size is normally 0.3–3.0 mm, and these are mostly used in continuous flow systems (fixed and moving bed). In this article, most of the discussion has been restricted to the liquid adsorption only.

Batch Systems. Generally, powdered activated carbons are used in batch systems. Batch system consists of contacting a whole volume of feed solution with a definite quantity of activated carbon in batch stirred vessels. The mixture is stirred or agitated to facilitate mass transfer. The important process design parameters can be calculated from laboratory batch adsorption isotherms, which precisely model the full-scale batch process. The batch adsorption processes are seldom used except in laboratories because they are highly inefficient compared with column adsorption processes and are therefore capital intensive and expensive to operate.

Various theoretical and empirical models have been proposed to describe the different types of adsorption isotherms in batch systems. The most commonly used models include Freundlich (38), Langmuir (39), and BET isotherms.

Freundlich Isotherm. The Freundlich adsorption isotherm model was given by Freundlich (38). This isotherm describes the equilibrium on heterogeneous surfaces and, hence, does not assume monolayer capacity

The Freundlich equation may be written as

$$q_e = K_F {C_e}^{1/n} \qquad \text{(Nonlinear form)}$$

$$\log q_e = \log K_F + \frac{1}{n \log C_e} \qquad \text{(Linear form)}$$

where q_e is the amount of solute adsorbed per unit weight of activated carbon (mg/g), C_e is the equilibrium concentration of solute in the bulk solution (mg/L), K_F is the constant indicative of the relative adsorption capacity of the adsorbent (mg/g), and 1/n is the constant indicative of the intensity of the adsorption.

Langmuir Isotherm. The Langmuir adsorption isotherm describes the surface to be homogeneous. The Langmuir adsorption isotherm assumed that all the adsorption sites have equal affinity for the molecules and that the adsorption at one site does not affect adsorption at an adjacent site (39,40).

The Langmuir equation may be written as

$$q_e = \frac{Q^0 b C_e}{1 + bC_e} \qquad \text{(Nonlinear form)}$$

$$\frac{C_e}{q_e} = \left(\frac{1}{Q^0 b}\right) + \left(\frac{1}{Q^0}\right) \cdot C_e \qquad \text{(Linear form)}$$

where q_e is the amount of solute adsorbed per unit weight of adsorbent (mg/g), C_e is the equilibrium concentration of solute in the bulk solution (mg/L), Q^0 is the monolayer adsorption capacity (mg/g), and b is the constant related to the free energy of adsorption. It is the value reciprocal of concentration at which half the saturation of the adsorbent is reached.

BET Isotherm. The Brunauer, Emmett, Teller (BET) isotherm assumes the pertaining of a compound between liquid and solid compartments or phases. This isotherm assumes the multilayers adsorption of solute on activated carbon (36,40)

$$q_e = \frac{BCQ^0}{(C_s - C)[1 + (B-1)(C/C_s)]} \qquad \text{Nonlinear Form}$$

$$q_e = \frac{C}{(C_s - C)q_e} = \frac{1}{BQ^0} + \frac{B-1}{BQ^0}\frac{C}{C_s} \qquad \text{Liner form}$$

where q_e is the amount of adsorbate adsorbed per unit weight of activated carbon, B is the constant related to the energy of interaction with the surface, C is the equilibrium concentration of adsorbate in solution (mg/L or mol/L), Q^0 is the number of moles of adsorbate per unit weight of carbon to form a complete monolayer, and C_s is the saturation concentration of the adsorbate.

The limitations of the adsorption isotherms are as follows:

1. Isotherms are equilibrium tests, and therefore, the time restrictions are not considered.
2. Isotherms are based on carbon exhaustion—granular systems do not totally exhaust—and the entire bed contents.
3. Long-term chemical and biological effects are not evident.

Thus, batch equilibrium adsorption isotherm tests cannot simulate or predict dynamic performance directly.

Continuous Systems. In continuous systems, mostly granular activated carbons (GAC) are used. The fixed bed adsorber systems are most widely used for conducting adsorption operations where the adsorbate to be treated is passed through a fixed bed. In the fixed bed adsorber operation, a degree of separation and removal is achieved that would require many steps in a batch system. The parameters, which are required to establish a fixed bed reactor, include the type of carbon, physical and chemical characteristics of the carbon, column diameter, water/wastewater/effluents flow rate, pH of the effluent, carbon bed depth, weight of carbon, contact time, concentration of the influent, concentration of the effluent, and desired effluent concentration

When contaminated solute (warer/wastewater/effluents) is passed through a bed of granular activated carbon, a wave front or a mass transfer zone (MTZ) is formed by continuous adsorption of solute in the carbon bed. Figure 4 shows the change in concentration of adsorbed species on the surface of activated carbon with time. The solute is rapidly adsorbed on the top layers of the bed until the amount adsorbed is in equilibrium with influent solute concentration. At this particular time, that portion of the bed is exhausted. Below this zone is a second zone where dynamic adsorption is taking place. The solute is being transferred from the liquid to the adsorbed phase. This zone is known as mass transfer zone, and the depth of the zone is controlled by many fcators depending on the solute concentration being adsorbed, characteristics of activated carbon, and hydraulic factors.

A plot between the concentration of the adsorbate exhibits an S-shaped curve in the adsorption zone with ends asymptotically approaching zero and the influent concentration C_0. This curve is known as a *break through curve*. An ideal break through plot obtained for a fixed bed adsorber is depicted in Fig. 4. The solute or impurity is adsorbed very rapidly by the few initial upper layers of the fresh granular carbon during the initial stages of operation, as shown in Fig. 4. These upper layers are in contact with the sorbate/impurity at its highest concentration level, C_0. The small amounts of solute/impurity that escape adsorption in initial stages are adsorbed in the lower stages, and no solute escapes from the fixed bed adsorber initially ($C_1 = 0$).

FIXED BED ADSORBER

An adsorption process in which liquid being treated is allowed to pass through a carbon column in that carbon becomes exhausted and the unit is removed from service and completely recharged with fresh carbon. The carbon remains fixed in the position during the whole adsorption process.

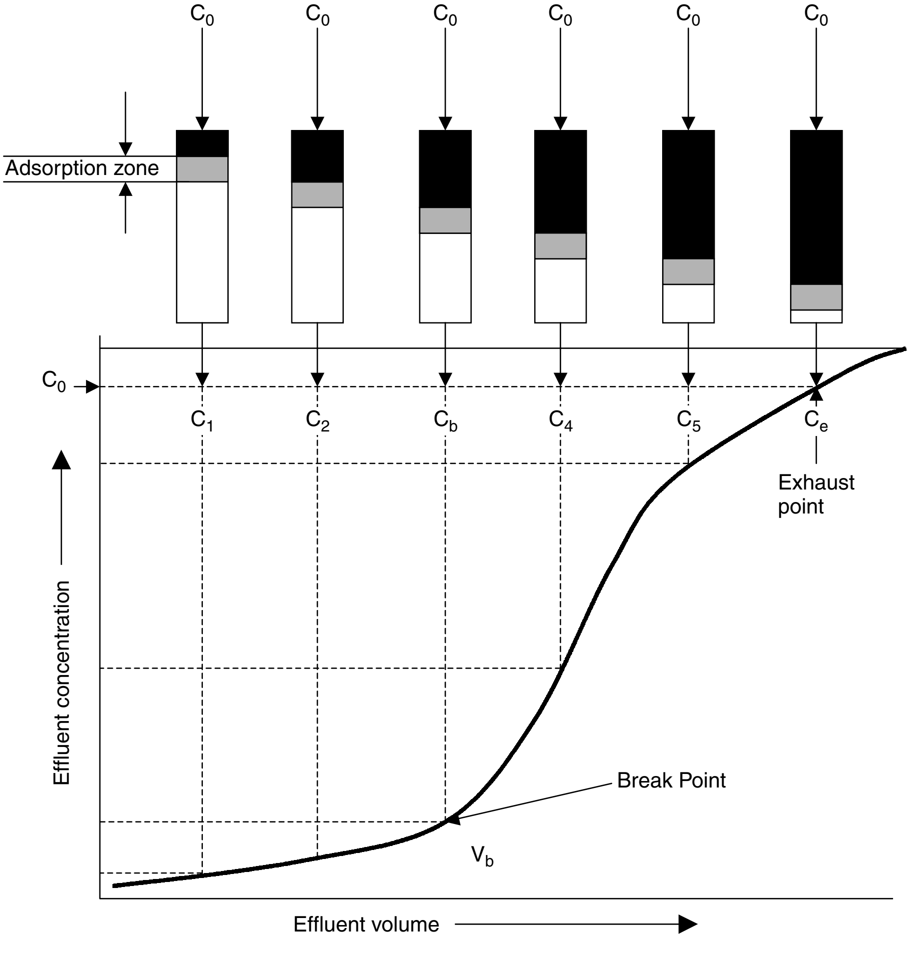

Figure 4. Typical adsorption zone movement in a fixed bed adsorber.

To construct a fixed bed adsorber, the following design parameters are used:

Volumetric Flow Rate (Q)

It is the quantity of solute fed per unit time. Flow rate influences the adsorption capacity of carbon in a dynamic system. If a 1-inch column is to be used, 20.6 ml/min flow rate, which is equal to 1 g/min/ft^2, is to be used.

The following values may further be considered for designing fixed bed reactors:

Column Diameter (cm)	Volumetric Flow Rate (cc/min)
2.56	41.6
5.08	165.2
7.62	374.0
10.16	661.6

Carbon Bed Volume (V_b)

This is the total volume of GAC packed bed, which accounts for both the activated carbon grains and the void volume.

Cross-Sectional Area (A_b)

It is simply the cross-sectional area.

Void Volume (α)

The volume between the carbon particles in a packed bed or column expressed as a percentage of the total bed volume, which corresponds to the part of the fixed bed volume that is not occupied by activated carbon particles

Filter or Linear Velocity (v_F)

It is also know as superficial linear velocity and surface loading rate, which is the velocity in an empty bed with the filter cross-sectional area.

It can be calculated from the following equation:

$$v_F = \frac{Q}{A_b}$$

Effective Contact Time, Resident Time, or Retention Time (t)

It is defined as the time within the GAC bed that is available for the mass transfer of the organic substances from bulk solution to the GAC particle. It is also defined as the theoretical length of time for a liquid to pass through a column assuming all the liquid moves through with the same uniform velocity. It is equal to the volume of liquid in the column divided by the rate of flow. The volume of liquid in a carbon column is simply the total volume of the column times the void fraction. It can be calculated by using the following equation:

$$t = \frac{V_F \alpha}{Q}$$

Empty Bed Contact Time (EBCT)

The time required for the liquid to pass through a carbon column assuming that all the liquid passes through at the same velocity. It is equal to the volume of the empty bed divided by the flow rate. Conventional water treatment plant adsorbers have EBCTs in the range of 7–10 min.

It can be calculated by using the equation

$$\mathrm{EBCT} = \frac{V_b}{Q} = \frac{A_b L}{V_F A_b} = \frac{L}{v_F} = \frac{t}{\alpha}$$

where V_b is the volume of GAC in contactor (m^3); Q is the volumetric flow rate (m^3/h); A_b is the cross-sectional area of GAC, m^2; L is the length of the GAC in contactor (m); V_F is the linear velocity or filter velocity (m/h); t is the retention time (h); and α is the Void volume (m^3).

Filter Operation Time (t_F)

It is the time period that a GAC bed has been in operation.

Throughput Volume (V_L)

It is the volume of solute that has passed through the filter at a time.

It is calculated by using the equation

$$V_L = t_F \times Q$$

Filter Density (ρ_{GAC})

It is calculated using the following equation in g/L:

$$\rho_{\mathrm{GAC}} = \frac{m_{\mathrm{GAC}}}{V_b}$$

where m_{GAC} is the mass of GAC (g) and V_{b} is the volume (L).

Specific Throughput Volume (V_{sp})

As the throughput volume depends on the filter size, it does not allow a direct comparison of different size of plants. If throughput volume is divided by the mass of activated carbon in the bed, the specific throughput volume is obtained. The mass of GAC is determined by multiplying filter density and bed volume.

It is calculated by using the equation

$$V_{sp} = \frac{V_L}{V_b \times \rho_{\mathrm{GAC}}} = \frac{Q \times t}{m_{\mathrm{GAC}}} = \frac{V_b \times t}{\mathrm{EBCT} \times m_{\mathrm{GAC}}} = \frac{V_b \times t}{\mathrm{EBCT} \times (\rho_{\mathrm{GAC}} \times V_b)} = \frac{t}{\mathrm{EBCT} \times \rho_{\mathrm{GAC}}}$$

Throughput Bed Volume (BV)

It is an another parameter used for the comparison of the removal efficiencies regardless of the bed size. It is the ratio of throughput volume and bed volume

$$\mathrm{BV} = \frac{V_L}{V_b} = \frac{t_F}{\mathrm{EBCT}}$$

Carbon Usage Rate (CUR)

It is expressed by the following expression:

$$\mathrm{CUR}, g/m^3 = \frac{m_{\mathrm{GAC}}}{Q \times t}$$

Bed Life

The volume of the water treated for a given EBCT, expressed in liters, and can be calculated as:

$$\text{Volume} = \frac{\text{mass} \cdot \text{of} \cdot \text{GAC} \cdot \text{for} \cdot \text{given} \cdot \text{EBCT}}{\text{GAC} \cdot \text{usage} \cdot \text{rate}}$$

and therefore, the bed life can be calculated as

$$\text{Bed} \cdot \text{Life} = \frac{\begin{array}{c}\text{volume} \cdot \text{of} \cdot \text{water} \cdot \text{treated} \\ \cdot \text{for} \cdot \text{given} \cdot \text{EBCT}\end{array}}{Q}$$

The typical values of the various important parameters as discussed above with their units presented in Table 5.

Carbon Dose

The amount of GAC required to fill each column is calculated by the expression:

$$\text{Weight of carbon} = \text{volume of column} \times \text{A.D} \times 0.85$$

where A.D is the apparent density and 0.85 is the factor to allow backwashed density.

This amount should be degassed and wetted prior to installation in the column. The degassing can be done by boiling the carbon in organic-free water for 2 h or soaking at room temperature for 24 h. The degassed carbon should be charged into the column in small increments as a slurry keeping a layer of organic-free water above the GAC during charging, which is best accomplished by filling the column one-third of organic-free water prior to charging the degassed-free water into the column. It is also noted that all connecting tube and other void space must be filled with liquid in order to avoid the formation of gas pockets in carbon bed.

For designing of fixed bed reactors, a number of theories have been proposed by various researchers. These theories include:

1. Length of unused bed (LUB) approach
2. Bed-depth-service time (BDST) approach
3. Empty bed residence time (EMRT) approach

Table 5

Parameter	Symbol	Units	Typical Values
Volumetric flow rate	Q	m^3/h	50–400
Bed volume	V_b	m^3	10–50
Cross-sectional area	A_b	m^2	5–30
Length	L	m	1.8–4.0
Void fraction	α	m^3/m^3	0.38–0.42
GAC density	ρ	kg/m^3	350–550
Filter velocity	v_F	m/h	5–15
Effective contact time	t	min	2–10
Empty bed contact time	EBCT	min	5–30
Operational time	t_F	days	100–600
Throughput volume	V_L	m3	10^4–10^5
Specific throughput	V_{sp}	m3/kg	50–200
Bed volume	BV	m3/m3	2,000–20,000

PULSED BED

In this type of bed, carbon is removed at intervals from the bottom of the column and replaced at top by fresh adsorbent.

MOVING BED SYSTEM

The most recent development of granular activated carbon has been the use of moving bed. In this system, the direction of the liquid flow is upward, whereas the carbon moves in a downward direction. The basic principle behind this technique is to have one column (or multiple columns running in parallel) packed completely with a carbon bed of sufficient height to have the adsorption–wave front and to provide some operating time with the effluent flow being within specification. As the adsorption wave front moves up the column, it is periodically displaced downward by the removal of a quantity of saturated carbon from the base of the column and the replacement of the same quantity with fresh or regenerated carbon at the top of the column. Although the principle of the moving bed can be applied in smaller units, it is most frequently used in larger units where the lower capital investment is important.

GAS-PHASE APPLICATIONS

Gas-phase applications of activated carbons include separation, gas storage, and catalysis. It is well documented in literature that only 20% of produced activated carbons are used for gas-phase applications. The carbons applied for gas-phase applications are mostly granular in shape.

REGENERATION

Once the granular activated carbons become saturated, it is necessary to change the carbon or to regenerate the fixed bed adsorber. Regeneration is the process of removing adsorbed compounds from the granular activated carbon surface. Here, the carbon surface includes external macropores and micropores.

Carbon regeneration frequently is a major part of total operating cost associated with granular activated carbon (GAC) adsorption systems. Following are the important and potential regeneration methods:

1. Thermal regeneration
2. Biological regeneration
3. Infrared regeneration
4. Supercritical fluids

Spent GAC used in water/wastewater, and other liquid-phase applications, is generally reactivated using a high-temperature thermal process where the GAC is heated to about 815 °C (1500 °F), allowing drying, baking, and gasification to occur. Various types of furnaces used for the regeneration purpose include multiple hearth furnace (MHF) or rotary kilns (RK), electric belt furnaces (EBF), and fluidized bed regenerators (FBR). Details about these furnaces and their operation are well documented in literature (41). The two major criteria involved in the regeneration of the granular activated carbon are (a) characteristic of the spent carbon and (b) choice of the furnace for that particular application. The spent carbon characteristics include the potential for char formation, corrosion, and slagging, whereas the furnace characteristics are assessed by determining the mass transfer efficiency, particle residence time, and temperature control. After reviewing the properties of furnaces and activated carbons, one can select the best possible option.

In thermal regeneration, 5–10% of granular activated carbon is lost as a result of oxidation and attrition, and by the cost of energy in heating, the carbon around 800–850 °C (42). An alternative technique is that of chemical regeneration in which chemical reagents are applied to the exhausted granular activated carbon. The chemical regeneration of exhausted GAC can be achieved by two main categories of substances: inorganic chemical regenerates with oxidizing powers and organic chemical regenerates with solubilizing powers. The efficiency of any regenerate is judged on the extent that it effects the recovery of the adsorptive powers of the granular activated carbons. The regeneration efficiency can be calculated as

$$\text{Regeneration efficiency (RE\%)} = (A_r/A_0) \times 100$$

where A_0 is the original capacity of GAC for a particular adsorbate and A_r is the capacity of regenerated carbon.

APPLICATIONS AND COMMON USES (BOTH LIQUID PHASE AND GAS PHASE)

Activated carbon has wide applications in both liquid and gas/solvent phase systems (43).

Groundwater Remediation

Granular activated carbon is the most common technology employed to pump and treat groundwater remediation systems. It is highly suited to this application and is often used as a single treatment step to remove compounds such as chlorinated hydrocarbons and aromatic compounds, including benzene, toluene, ethylbenzene, and xylene (BTEX). For more highly contaminated groundwater, two or more carbon units may be placed in series or carbon may be used in combination with other treatment technologies such as air stripping or advanced oxidation processes.

Soil Vapor Extraction and Air Sparring

Activated carbon can also be used for the removal of Volatile Organic Compounds (VOCs) from air streams resulting from *in situ* removal techniques such as soil vapor extraction.

Chemical and Pharmaceutical Industries

Chemicals. Activated carbon is suitable for the decolorization and purification of a wide range of organic and inorganic compounds, including amines, hydrochloric and other mineral acids, amino acids, glycols, and hydrocarbons.

Pharmaceuticals. Activated carbon is used to purify a wide range of pharmaceuticals and intermediates.

Granular activated carbon is installed to purify the recirculating amine to remove degradation products and dissolved hydrocarbons.

Military Use

Most of the world's armed forces are using activated carbons to protect against attack by toxic gases such as mustard gas. This is also used in military suiting where combat uniforms are coated by a layer of impregnated carbon under the outer cover.

Nuclear Reactors

Most of the nuclear reactors, especially in the western world, have activated carbon ventilators to protect against radioactive iodine leaks from the core or heat exchanger systems, if any. Special activated carbons impregnated with potassium iodide or potassium tri-iodide are commonly used for this purpose. Another application in nuclear technology is as a filter in emergency ventilation system for the reactor building, which is switched on automatically in case of breakdowns.

Landfill Leachate Treatment

Granular activated carbon in combination with biological pretreatment is the leading technology for the treatment of landfill leachate for the removal of Chemical Oxygen Demand (COD), Adsorbable Organic Halogens (AOX), and other toxic substances. Granular activated carbon is now used at over 50 sites in Europe for this application.

Catalysis

Many chemical reactions require a catalyst to improve efficiency, accordingly; in many cases, activated carbons provide large surface area, thus, further improving the efficiency.

Medicinal Activated Carbon

The activated carbon has been applied in medicine for a long time. The carbons originally used were prepared from waste materials of animal origin, especially blood (animal charcoal-carbo animalis). In *catarrhal* infections of the digestive system, the use of activated carbon serves primarily to remove bacterial toxins, which, being high-molecular-weight substances, are easily adsorbed on active carbon. Activated carbons are also prescribed in large doses in all the cases of acute gastritis and enteritis. It is also a very effective antidote in all the cases of poisoning.

Domestic Use

Activated carbons are used in various home appliances, including fridge deodorizers, air purifiers, and cooker hoods. The activated carbons are used in the removal of caffeine from coffee. Cigarettes are made that, in addition to an antismoke filter, contain finely granulated activated carbon.

Also, activated carbons are used in the coating and printing industries, degreasing and cleaning, solvent recovery, tank venting, ventilation, and air conditioning

The other uses are systems annihilation, condensate de-oiling, and gold recovery.

BIBLIOGRAPHY

1. Mattson, J.S. and Mark, H.B., Jr. (1971). *Activated Carbon*. Marcel Dekker, New York.
2. Hamerlinck, Y. and Mertens, D.H. (1994). *Activated Carbon Principles in Separation Technology*, E.F. Vansant (Ed.). Elsevier, New York.
3. Bansal, R.P., Donnet, J.-P., and Stoeckli, F. (1988). *Active Carbon*. Marcel Dekker, New York.
4. Pollard, S.J.T., Fowler, G.D., Sollars, C.J., and Perry, R. (1992). Low cost adsorbents for waste and wastewater treatment: A review. *The Science Total Environ.* **116**: 31–52.
5. Gupta, V.K. and Ali, I. (2002). Adsorbents for water treatment: low cost alternatives to carbon. In: *Encyclopedia of Surface and Colloid Science*. Marcel Dekker, New York, pp. 136–166.
6. Allen, S.J., Whitten, L., and Mckay, G. (1998). The production and characterization of activated carbons: A review. *Dev. Chem. Eng. Mineral Process* **6**(5): 231–261.
7. Arulanantham, A., Balasubramanian, N., and Ramakrishna, T.V. (1989). Coconut shell carbon for treatment of cadmium and lead containing wastewater. *Metal Finish* 51–55.
8. Ferro-Garcia, M.A., Rivera-Ultrilla, J., Ropdriquez-Gordillio, J., and Bautista-Tolego, I. (1988). Adsorption of zinc, cadmium, and copper on activated carbons obtained from agricultural by-products. *Carbon* **26**(3): 363–373.
9. Manju, G.N., Raji, C., and Anirudhan, T.S. (1998). Evaluation of coconut husk carbon for the removal of arsenic from water. *Water Res.* **32**(10): 3062–3070.
10. Srivastava et al. (1987).
11. Renouprez, A. and Avom, J. (1988). Characterization of active carbons from palm tree fibres using nitrogen adsorption and small angle X-ray scattering. In: *Characterization of Porous Solids*. K.K. Unger, J. Rouquerol, and K.S.W. Sing (Eds.). Elsevier Science, Amsterdam, the Netherlands, pp. 49–54.
12. Laine, J., Calafat, A., and Labady, M. (1989). Preparation and characterization of activated carbons from coconut shell impregnated with phosphoric acid. *Carbon* **27**(2): 191–195.
13. Korczak, M. and Kuribiel, J. (1989). New mineral-carbon sorbent: mechanism and effectiveness of sorption. *Water Res.* **23**(8): 937–946.
14. Ehrburger, P., Addoun, A., Addoun, F., and Donnet, J.-P. (1986). Carbonization of coals in the presence of alkaline hydroxides and carbonates: Formation of activated carbons. *Fuel* **65**(10): 1447–1449.
15. Alvim-Ferraz, M.C. (1988). Preparation of activated carbon for air pollution control. *Fuel* **67**(9): 1237–1241.
16. Toles, C.A., Marshall, W.E., and Johns, M.M. (1997). Granular activated carbons from nutshells for the uptake of metals and organic compounds. *Carbon* **35**(9): 1407–1414.
17. Tancredi, N., Cordero, T., Rodriguez-Mirasol, J., and Rodriguez, J.J. (1997). Activated carbons from eucalyptus wood. Influence of the carbonization temperature. *Sep. Sci. Technol.* **32**(6): 1115–1126.
18. Teng, H. and Yeh, T.-S. (1998). Preparation of activated carbons from bituminous coals with zinc chloride activation. *Ind. Eng. Chem. Res.* **37**(1): 58–65.
19. Hall, R.C. and Holmes, R.J. (1992). The preparation and properties of some activated carbons modified by treatment with phosgene or chlorine. *Carbon* **30**(2): 173–176.
20. Otowa, T., Nojima, Y., and Miyazaki, T. (1997). Development of KOH activated high surface area carbon and its application to drinking water purification. *Carbon* **35**(9): 1315–1319.
21. Duggan, O. and Allen, S.J. (1997). Study of the physical and chemical characterstics of a range of chemical treated lignite based carbons. *Wat. Sci. Tech.* **35**: 21–27.
22. Periasamy, K. and Namasivayam, C. (1995). Adsorption of Pb(II) by peanut hull carbon from aqueous solution. *Sep. Sci. Technol.* **30**(10): 2223–2237.
23. Mohan, D., Singh, K.P., Sinha, S., and Gosh, D. (2004). Removal of pyridine from aqueous solution using low cost activated carbon derived from agricultural waste materials. *Carbon* **42**: 2409–2421.
24. Rodriguez-Reinoso, R.F. and Solano, A.L. (1989). Microporous structure of activated carbons as revealed by adsorption methods. In: *Chemistry and Physics of Carbon*, Vol. 12, P.A. Thrower (Ed.). Marcel Dekker, New York.
25. Hassler, J.W. (1963). *Active Carbon*. Chemical Publishing Company, New York.
26. Hu et al. (1996).
27. Gregg, S.J. and Sing, K.S.W. (1982). *Adsorption, Surface Area and Porosity*. Academic Press, London.
28. Snoeyink, V.L. and Weber, W.J., Jr. (1967). The surface chemistry of active carbon;discussion of structure and surface functional groups. *Environ. Sci. Technol.* **1**(3): 228.
29. Cookson, J.T., Jr. (1975). The adsorption chemistry of organics from waters on activated carbon AIChE Synposium Series. *Water* **71**(151): 376.
30. Epstein, B.D.E., Dalle-Molle, J.S., and Mattson, J.S. (1971). Electrochemical investigation of durface functional groups of isotropic pyrolytic carbon. *Carbon* **9**: 609.
31. Mattson, J.S., Mark, Jr., H.B., Malbin, M.P., Weber, W.J., Jr., and Crittenden, J.C. (1969). Surface chemistry of active carbon:specific adsorption of phenols. *J. Colloid Interface Sci.* **31**(1): 116–130.
32. Garten, V.A. and Weiss, D.E. (1957). A new interpretation of the acidic and basic structures in carbon II: The chromene-carbonium ion couple in carbon. *Aust. J. Chem.* **10**: 309.
33. Boehm, H.P. and Voil, M. (1970). Basische Oberflachernoxide an Kohlenstoff-I: Adsorption von Sauren. *Carbon* **8**: 227.
34. Corapcioglu and Huang (1987).
35. Perrich, J.R. (1981). *Activated Carbon Adsorption for Wastewater Treatment*. CRC Press, Boca Raton, FL.
36. Brunauner, S., Emmett, S.P.H., and Teller, E. (1972). Adsorption of gases in multimolecular layers. *J. Am. Chem. Soc.* **60**: 309–319.
37. Sing et al. (1985). Physisorption data for gas/solid systems. *Pure Appl. Chem.* **57**: 603–619.
38. Freundlich, H. (1906). Uber die adsorption n losungen *Z. Phys. Chem.* **57**: 385–470.
39. Langmuir, I. (1918). The adsorption of gases on plane surfaces of glass, mica and platinum. *J. Am. Chem. Soc.* **40**: 1361–1367.
40. Weber, Jr., W.J. (1972). *Physicochemical Processes for Water Quality Control*. Wiley Interscience, New York.
41. Liu, P.K.T. and Wagner, N.J. (1985). Thermal regeneration of granular active carbon. *Environmental Progress* **4**(2): 136–140.
42. Gaymont, F.J. (1980). The effect of capital and operating costs on GAC adsorption system design. In: *Activated Carbon*

Adsorption of Organics from the Aqueous Phase, Vol. 2, M.J. McGuire and I.H. Suffet (Eds.). Cap 23. Ann Arbor Science, Ann Arbor, MI, pp. 531–538.

43. Mckay, G. (1995). *Use of Adsorbents for the Removal of Pollutants from Wastewaters*. CRC Press, Boca Raton, FL.

READING LIST

Mohan, D. and Chander, S. (2001). Kinetics of mercury adsorption from wastewater using activated carbon derived from fertilizer waste material. *Colloids and Surfaces A* **177**: 169–181.

COMPETITIVE ADSORPTION OF SEVERAL ORGANICS AND HEAVY METALS ON ACTIVATED CARBON IN WATER

Dinesh Mohan
Kunwar P. Singh
Gomti Nagar, Lucknow,
Uttar Pradesh, India

In practice, the water/wastewater/effluents to be treated are multicomponent mixtures. Therefore, other organic or inorganic substances influence the efficacy of activated carbons. In the recent past, most of the research work has been focused on the removal of single solute from its pure solution or a particular solute from a solution having ill-defined chemical character. The other reports available are addressed to the removal of gross organic/inorganic substances, which are measured in terms of BOD or COD from water/wastewater. In actual practice, the treatment of wastewater by activated carbon normally involves a variety of adsorbates competing for adsorption sites on the carbon surface. Thus, a thorough understanding of the competitive effect on various organic/inorganic compounds on carbon adsorption is a must. Unfortunately, very few reports/papers about the competitive adsorption on activated carbon in aqueous solution are available. Thus, a need exists to explore the current status of the competitive adsorption of various organic/inorganic substances from water/wastewater on activated carbons. Multiion systems received less attention than single ion systems.

The adsorption in multicomponent systems is complicated because of the solute–solute competition and the solute-surface interaction involved. Multicomponent interactions take place at the active adsorption sites where the solid-liquid phase equilibrium will emerge showing a different capacity of single solute with a new set of isotherm systems. The interpretation of the multicomponent systems has proved to be complex and may be the function of one or all of the following parameters: ionic radii, electronegativity, system pH, and the availability of the active sites on the adsorbent.

The adsorption capacity or the degree of removal of various organic and inorganic substances from water/wastewater on activated carbon (GAC) very much depends on the multicomponent competitive interactions of organic chemicals that are present in the system. In order to design a cost-effective system, it is necessary to know the adsorption capacity of a solute in the presence of others (1–4). To determine the effects of competing adsorption, mathematical models/equations must be developed to describe the adsorption equilibrium in multicomponent system. Several models are available in literature, but most of them suffer from one limitation or an other. In the paragraphs below, several models are discussed with a view that these are mostly used in multicomponent modeling in aqueous systems. Butler and Ockrent (5), in 1930, developed the first model for binary systems. Since then, a number of models have been developed that are shown to be valid for certain mixtures and conditions by comparing them to the experimental data.

MULTICOMPONENT ADSORPTION MODELS

Prediction of multicomponent adsorption is still a very challenging problem in the adsorption. Many models have been proposed and employed from time to time to obtain the multisolutes isotherms using single-component equilibrium data. However, most of the models are based on either unrealistic assumptions or an empirical equations with no apparent definitions (6,7). Some of the important and widely used models are discussed below to understand the multicomponent sorption of oraganic and inorganic substances on activated carbon.

Butler and Ockrent Model

This model was presented by Butler and Ockrent in 1930 (5). Basically, this model is an extended form of Langmuir model. According to this model, if two solutes are present in an adsorption system, the Langmuir model for the competitive adsorption can be given by Equations 1 and 2 as follows:

$$q_1 = \frac{Q_1^0 b_1 C_1}{1 + b_1 C_1 + b_2 C_2} \tag{1}$$

$$q_2 = \frac{Q_2^0 b_2 C_2}{1 + b_1 C_1 + b_2 C_2} \tag{2}$$

where Q_1^0, b_1 and Q_2^0, b_2 are the Langmuir constants determined from single-solute system with solutes 1 and 2, respectively, and C_1 and C_2 are the equilibrium concentrations of the solutes 1 and 2, respectively, in the mixture.

This model is valid only if the surface area available for adsorption is identical for both the solutes and Q_1^0 and Q_2^0 reflect the size difference for a mono-layer surface coverage of solutes. Further, the free energy change for adsorption is assumed to be independent of surface coverage.

For simultaneous adsorption of N components from a solution, the multicomponent Langmuir equation can be given as

$$q_i = \frac{Q_i^0 b_i C_i}{1 + \sum_{j=1}^{N} b_j C_j} \tag{3}$$

In spite of the limitations, this model was used by various researchers (8,9) to describe their multicomponent sorption on activated carbons.

Jain and Snoeyink Model

The model given by Butler and Ockrent (5) was modified by Jain and Snoeyink (10). Earlier models used for multicomponent systems are not able to explain the fact that in a bi-solute system, a portion of the adsorption often takes place without competition. To overcome this difficulty, Jain and Snoeyink proposed a model for binary systems, which takes into account the noncompetitive adsorption on activated carbon/adsorbent.

Based on this hypothesis, adsorption without competition occurs when $Q_1^0 \neq Q_2^0$ and the number of adsorption sites not being subjected to the competition is equal to the quantity of $(Q_1^0 - Q_2^0)$ where $Q_1^0 > Q_2^0$. Equations 4 and 5 are as follows:

$$q_1 = \frac{({Q_1}^0 - {Q_2}^0)b_1C_1}{1 + b_1C_1} + \frac{Q_2^0 b_1 C_1}{1 + b_1C_1 + b_2C_2} \tag{4}$$

$$q_2 = \frac{Q_2^0 b_2 C_2}{1 + b_1C_1 + b_2C_2} \tag{5}$$

where q_1 and q_2 are the amounts of solute 1 and 2, respectively, adsorbed per unit weight of activated carbon/adsorbent at equilibrium concentrations of C_1 and C_2, respectively. Q_1^0 and Q_2^0 are the maximum value of adsorption for solute 1 and 2, respectively, determined from the respective single-solute systems, whereas b_1 and b_2 are the constants related to the energy of adsorption from solute 1 and 2, respectively, in their pure solution systems.

This model also has some limitations. It is only appropriate to describe the competitive adsorption between the molecules having very different single-solute adsorption capacity. In spite of its limitations, a number of workers have applied this model in their multicomponent adsorption studies. In one of the studies, Huang and Steffens (11) applied this model to determine the competitive adsorption of organic materials by activated carbons. It was concluded that mutual suppression of equilibrium adsorption because of competition between acetic and butyric acids has shown that the observed data are somewhat closer to the values predicted by Jain and Snoeyink's model than by the original Langmuir equation. However, the actual degree of suppression is greater than the prediction for acetic acid and smaller for butyric acid.

Multicomponent Isotherm of Mathews and Weber

Mathew (12) proposed this model in 1975. This model is a modified and extended form of the Redlich–Peterson model (13), which can be given by the expressions below, which is basically three parameter, single-solute adsorption isotherm model

$$q_e = \frac{KC_e}{1 + a_R {C^b}_e} \tag{6}$$

where K, a_R, and b are the Redlich–Peterson constants.

For the N solute mixtures, this equation can be written as

$$q_i = \frac{K_j C_i}{1 + \sum_{j=1}^{N} a_j {C_j}^{b_j}} \tag{7}$$

The parameters K_j, a_j, and C_j can be determined from single-solute isotherm data. A new constant, η_i, has been introduced in this model, which has to be determined from adsorption data in the mixtures. With the addition of this constant, the equation may be written as

$$q_i = \frac{k_j \left[\dfrac{C_i}{\eta_i}\right]}{1 + \sum_{j=1}^{N} a_j \left[\dfrac{C_j}{\eta_i}\right]^{b_j}} \tag{8}$$

where η_I is the interaction parameter, which is constant. In practice, this parameter varies from different equilibrium compositions. Therefore, this model is not very successful in explaining the multicomponent systems.

Fritz and Schlunder Multicomponent Model

This model was given by Fritz and Schlunder in 1974 (14,15). For modeling multicomponent systems comprising species whose single-solute isotherms follow the Freundlich isotherm, a multicomponent Freundlich equation may be used. The first model of this type, as proposed by Fritz and Schlunder (14), can be expressed by Equations 9–12:

$$q_1 = \frac{K_1 {C_1}^{n_1 + \beta_{11}}}{C_1^{\beta_{11}} + \alpha_{12} {C_2}^{\beta_{21}}} \tag{9}$$

$$q_2 = \frac{K_2 {C_2}^{n_2 + \beta_{22}}}{C_2^{\beta_{22}} + \alpha_{21} C_1^{\beta_{21}}} \tag{10}$$

$$\alpha_{1,2} = \frac{\alpha_{1,2}}{\alpha_{1,1}} \tag{11}$$

$$\alpha_{2,1} = \frac{\alpha_{2,1}}{\alpha_{22}} \tag{12}$$

where q_1 and C_1 are the concentrations of solute 1 in the solid and liquid phase, respectively; q_2 and C_2 are the concentrations of the solute 2 in solid and liquid phase, respectively; and K_1, n_1, and K_2, n_2 are the Freundlich constants in single-solute 1 and solute 2 systems.

Equations 9 and 10 consist of ten adjustable variables; however, K_1 and K_2, n_1 and n_2 can be determined from single-solute isotherms using Freundlich model for single-solute systems.

Dastgheib and Rockstraw Model

Very recently Dastgheib and Rockstraw (16) proposed a multicomponent Freundlich equation for binary systems, as given by Equations 13–15:

$$q_1 = \left[\frac{K_1 {C_1}^{n_1}}{K_1 {C_1}^{n_1} + \alpha_{12} K_2 {C_2}^{n_2} + b_{12} {C_2}^{n_{12}}}\right] K_1 {C_1}^{n_1} \tag{13}$$

$$q_2 = \left[\frac{K_2 {C_2}^{n_2}}{K_2 {C_2}^{n_2} + \alpha_{21} K_1 {C_1}^{n_1} + b_{21} {C_1}^{n_{21}}}\right] K_2 {C_2}^{n_2} \tag{14}$$

where q_1 and C_1 are the concentrations of solute 1 in the solid and liquid phase, respectively; q_2 and C_2 are the concentrations of solute 2 in the solid and liquid phase, respectively; K_1, n_1, K_2, and n_2 are the Freundlich constants in single-solute system; and α_{12}, α_{21}, b_{12}, b_{21}, n_{12}, and n_{21} are the interaction constants obtained from a least-squares analysis of the binary data. The term in brackets on the right-hand side represents the overall competition and interaction factor and has a value of less than or equal to unity (where C_2 tends to zero, it is equal to 1). The terms $\alpha_{12}K_2$ and $\alpha_{21}K_1$ can be condensed to single terms and are considered as constants.

For the ith component systems, the general equation model can be written as

$$q_i = \frac{(K_i C_i^{n_i})^2}{K_i C_i^{n_i} + \sum_{j=1}^{N}(\alpha_{ij} K_j C_j^{n_j} + b_{ij} C_j^{n_{ij}})} \tag{15}$$

where q_i and C_1 are the concentrations of solute I in the solid and liquid phase, respectively; C_j is the concentration of other solutes in liquid phase; K_i and n_i, K_j and n_j are the single component Freundlich constants; α_{IJ}, b_{ij}, and n_{ij} are the binary interaction constants obtained from a least-squares analysis of the multicomponent data having $\alpha_{ii} = b_{ii} = 0$; and N is the number of solute. This model is different from the Fritz and Schlunder model in the sense that this demonstrates equal or stronger performance.

Sheindorf et al. Model

This model was given by Sheindorf et al. (17,18) for the multicomponent systems comprised of species whose single-solute isotherm obeys the Freundlich isotherm. This equation was based on the following assumptions:

1. each component in single system obeys the Freundlich model
2. each component in multicomponent system and the adsorption energies of different sites are disturbed exponentially, with the distribution function being identical to that for single-component systems.

The model equations can be given by Equations 16 and 17:

$$q_1 = K_1 C_1 (C_1 + \eta_{12} C_2)^{n_1 - 1} \tag{16}$$

$$q_2 = K_2 C_2 (C_2 + \eta_{21} C_1)^{n_2 - 1} \tag{17}$$

where q_1 and C_1 are the concentrations of solute 1 in the solid and liquid phase, respectively; q_2 and C_2 are the concentrations of solute 2 in the solid and liquid phase, respectively; K_1, n_1, K_2, and n_2 are the Freundlich constants in single-solute system; and η_{12} and η_{21} are the interaction constants.

Ideal Adsorbed Solution Theory (IAST)

The ideal adsorbed solution theory is based on the thermodynamics of adsorption, which is analogous to Roul's law in a liquid-gas system. The only difference is that it is applied to a solid-liqid system. Initially, this model was used to calculate multicomponent adsorption of gaseous mixtures (19) using the parameters calculated from single-solute systems. This model was later modified and applied to calculate various multicomponent adsorption parameters by Radke and Praunitz (20). In the IAST model, the following five basic equations (18–22) are used to predict multicomponent behavior from single-solute adsorption isotherms (21).

The total surface loading can be defined by Equation 18:

$$q_T = \sum_{i=1}^{N} q_i \tag{18}$$

where q_I is the single-solute solid phase concentration for component i, which is evaluated at spreading pressure of the mixture, and N is the number of components.

The mole fraction on the carbon surface for component i can be calculated by Equation 19:

$$Z_i = \frac{q_i}{q_T} \quad \text{and} \quad I = 1 \text{ to } N \tag{19}$$

$$C_i = Z_i C_i^0 \quad \text{and} \quad I = 1 \text{ to } N \tag{20}$$

where C_I is the single-solute liquid phase concentration for component i, which is evaluated at the spreading pressure of the mixture.

The single-solute liquid phase concentration in equilibrium with q_i^0 is

$$\frac{1}{q_T} = \sum_{i=1}^{N} Z_i / q_i^0 \tag{21}$$

$$\frac{\pi_m}{RT} = \int_0^{q_i^0} \frac{d \ln C_i^0}{d \ln q_1^0} dq_1^0 = \frac{\pi_1^0 A}{RT}$$
$$= \int_0^{q_j^0} \frac{d \ln C_j^0}{d \ln q_j^0} dq_j^0 = \frac{\pi_j^0 A}{RT} = \text{for} \cdot j = 2 \cdot \text{to} \cdot N \tag{22}$$

where A is the surface area of carbon per unit mass of adsorbent, R is the gas constant, T is the absolute temperature, π_i is the spreading pressure of the single solute i, and π_m is the spreading pressure of the mixture.

LeVan and Vermeulen Model

LeVan and Vermeulen (22) have modified the competitive Langmuir-like model. IAS theory was considered in modifying the model. This model predicts the equilibrium relationships of solute mixture only from the data derived from single adsorption isotherms. It is the simplest isotherms derived from the IAS model.

Statistical Design for Competitive Adsorption

2^4 factorial experimental design was used to study the competitive adsorption of Fe(II), Mn(II), Ca(II), and Zn(II) on selected activated carbons B3, W2, W3, and lignite by Mohan and Chander (23).

These designs are important for the following reasons:

1. They require relatively few runs per factor studied; and although they are unable to explore fully a

wide region in the factor space, they can indicate major trends and therefore determine a promising direction for further experimentation.

2. When a more thorough local exploration is needed, they can be suitably augmented to form composite designs.
3. These fractional designs are often of great value at an early stage of an investigation, when it is frequently good practice to use the preliminary experimental efforts to look at a large number of factors superficially rather than a small number.
4. These designs and the corresponding fractional designs may be used as building blocks so that the degree of complexity of the family-constructed design can match the sophistication of the problem.
5. The interpretation of the observations produced by the designs can proceed largely by using common sense and elementary arithmetic.

The authors have used a total run of $2^4 = 16$. The variables were the concentration levels of various metal ions, with high level (+) 100 ppm and low level (−) 0 ppm. The experiments were conducted at pH 3.5. The experiments were arranged as the design matrix.

General Factorial Design

To perform a general factorial design, a fixed number of "levels" or (versions) for each of the variables (factors) can be selected, and then experiment is run with all possible combinations. If l_1 levels exist for the first variable, l_2 for the second, ..., and l_k for the kth, the complete arrangement of $l_1 \times l_2 \times l_3 \times \ldots l_k$ experimental runs is called an $l_1 \times l_2 \times l_3 \times \ldots l_k$ factorial design, e.g., a $2 \times 3 \times 5$ factorial design requires $2 \times 3 \times 5 = 30$ runs and a $2 \times 2 \times 2 = 2^3$ factorial design.

COMPETITIVE SORPTION OF INORGANICS ON ACTIVATED CARBON

The adsorption of Pb^{2+}, Cu^{2+}, Zn^{2+}, and Cd^{2+} from aqueous solutions by three activated carbons in single and multicomponent systems were studied by Budinova et al. (24). These three activated carbons were obtained from apricot stones (*A*), coconut shells (*C*), and lignite coal (*L*). The results of the individual metal ions from an aqueous solution containing all four metal ions together in equal concentration are presented in Table 1.

It is clear from Table 1 that the presence of foreign ions diminishes the adsorption of each of the ions. The effect is greatest for the lead ions and smallest for the copper ions. The authors did not mention any mechanism for the multicomponent sorption of these ions. The only reason given was that, apart from the properties of the cations, the chemical nature of the metal ions is of great importance for the adsorption process. They also concluded that a selective adsorption of the metals is observed; the ones preferentially adsorbed do not completely prevent the adsorption of other ions.

Johns et al. (25) reported the sorption of cadmium, copper, lead, nickel, and zinc in single and multicomponent systems on various granular activated carbon developed from agricultural waste materials. A study on the competitive effect of metal ions was carried out from a solution having 2.5 mM of each metal at pH 5.0 and was also unbuffered to reduce solution species complexation. The uptake of various metals from a mixed solution is presented in Table 2.

Bansode et al. (26) evaluated the adsorption effectiveness of pecan shell-based granular activated carbons (GACs) in removing metal ions Cu(II), Pb(II), and Zn(II) commonly found in municipal and industrial wastewater. Pecan shells were activated by phosphoric acid, steam, or carbon dioxide activation methods. Metal ion adsorption

Table 1

	Adsorption from Solution Containing All the Four Ions (mol g^{-1})				Decrease of Ion Adsorption in the Presence of the ion (%)			
Carbon	Cu^{2+}	Pb^{2+}	Zn^{2+}	Cd^{2+}	Cu^{2+}	Pb^{2+}	Zn^{2+}	Cd^{2+}
A	434.4	355.4	410.0	385.0	11.6	28.3	11.0	19.9
C	430.7	354.0	390.0	360.0	11.6	28.1	17.2	19.4
L	403.5	328.7	390.0	360.3	15.4	32.1	17.0	21.0
OA	398.5	1550.4	73.6	47.6	12.1	37.3	27.1	32.2

Table 2

GAC	BET Surface Area (m^{2}g^{-1})	μ Moles of Metals Adsorbed per Gram of GAC					
		Ni(II)	Cu(II)	Zn(II)	Cd(II)	Pb(II)	Total
Calgon GAC	783	0	97	0	30	113	240
Norit RO3515	827	0	117	0	11	67	195
Norit vapure	876	0	98	0	4	66	168
Soybean hulls	479	14	127	29	36	190	396
Peanut shells	275	9	195	31	39	236	510
Sugar cane bagasse	162	7	132	21	29	206	395
Rice straw	460	2	144	24	32	174	376

of shell-based GACs was compared with the metal ion adsorption of a commercial carbon, namely Calgon's Filtrasorb 200. Adsorption experiments were conducted using solutions containing all three metal ions in order to investigate the competitive effects of the metal ions as would occur in contaminated wastewater. The results obtained from this study showed that acid-activated pecan shell carbon adsorbed more lead ion and zinc ion than any of the other carbons, especially at carbon doses of 0.2–1.0%. However, steam-activated pecan shell carbon adsorbed more copper ion than the other carbons, particularly using carbon doses above 0.2%. In general, Filtrasorb 200 and carbon dioxide-activated pecan shell carbons were poor metal ion adsorbents. The results indicate that acid- and steam-activated pecan shell-based GACs are effective metal ion adsorbents and can potentially replace typical coal-based GACs in treatment of metal contaminated wastewater. The surface complex formation model was used successfully to describe the surface change density, as well as the single and multispecies metal adsorption equilibrium by Chen and Lin (27).

Choi and Kim (28) studied the adsorption characteristics of zinc and cadmium ion on granular activated carbon in singular and binary systems. Features of binary adsorption were discussed for several influential parameters, and experimental observations for both ions were correlated with a predicted adsorption isotherm based on a Langmuir multicomponent model.

Yu and Kaewsarn (29) used the multicomponent model on the sorption of heavy metals on low-cost adsorbents. An equilibrium isotherm was predicted by the extended Langmuir model using Langmuir parameters as determined from a single component system. Copper ions were found to have adsorption affinity, and the separation factor $\alpha_{Cu/Cd}$ was determined as 3.05. Trujullo et al. (30) reported the competitive adsorption of six metal ions from a single solution, which led to a model applicable to their batch and semicontinuous packed beds. Binding capacity was highest for copper, independent of the other ions, and copper also exerted the largest competing effect.

Bunzle et al. (31) carried out the studies in the pH range of 3.5–4.5, and the order was found to be $Pb^{2+} > Cu^{2+} > Cd^{2+} \approx Zn^{2+} > Ca^{2+}$, whereas Masslenilov and Kiselva (32) reported the adsorption capacity order as $Cu^{2+} > Zn^{2+} > Fe^{3+} > Ca^{2}$.

Ho et al. (33) reported that competitive effect affected the sorption of three metals in the order $Ni^{2+} > Pb^{2+} > Cu^{2+}$. It was concluded that all the metals are not necessarily adsorbed by exactly similar mechanisms for all the biosorbents, and that each needs to be tested to determine its characteristics. Adsorption of lead is usually greater than of copper, although copper is a more aggressive competitor and the adsorption of nickel is usually weaker than that of others.

The influence of a range of commercially available, water-soluble surfactants on the uptake of heavy metal ions (Cu, Zn, Cd, and Pb) by three types of clay (kaolinite, illite, and a montmorillonite) was reported by Beveridge and Fickering (34). The adsorption of Cu, Pb, Cd, and Zn was significantly reduced in the presence of small amounts of cationic surfactants, particularly with montmorillonite suspensions. The addition of anionic surfactants led to increase metal loss from the solution. Studying multicomponent adsorption system equilibria must commence with an accurate description of each component in its single (or pure) component equilibrium state.

Allen and Brown (35) studied the single component and multicomponent metal sorption onto lignite. A comparison was made between the single component saturation uptake and multicomponent uptakes The multicomponent systems were equimolar binary solutions solutions Cu-Cd, Cu-Zn, and Cd-Zn and a ternary mixture of equimolar Cu-Cd-Zn. In single component systems, the adsorption capacity followed the order Cu > Zn > Cd. These capacities were reported on a molar basis. Despite the competition, the total sorption capacity was found to increase even though the adsorption capacity of a single ion may be less than if it were to present alone. It has been pointed out that a substantial effect of multicomponent mixtures was observed on the capacity of lignite for cadmium and zinc. There appears to be slight increases in capacity in binary Cu-Cd, Cd-Zn, and ternary mixtures and a decrease in the capacity of Cu-Zn mixture compared with single component data. The order of sorption of metals in multicomponent systems is as follows: Cu > Cd > Zn. The preference of the sorbents for metals uptake is related to the electronegativity of the ions. Copper possessing the greatest ionic potential has the strongest attraction to the adsorbent, followed by cadmium then closely by zinc. The sorption capacities for single as well as multicomponent systems as reported by Allen and Brown are presented in Table 3.

Tan et al. (36) reported the uptake of metal ions in single and multicomponent systems by chemically treated human hairs. Various suppressors and promoters were identified and given in Table 4.

Beveridge and Pickering (34) reported the effect of various water-soluble surfactants on the uptake of Cu, Zn, Cd, and Pb ions by three types of clays, kaolinite, illite, and montmorillonite, over the pH range 3–10. Adsorption of Cu(II), Pb(II), Cd(II), and Zn(II) was significantly reduced in the presence of small amounts of cationic surfactants, particularly with montmorillonite suspensions. No multicomponent sorption modeling was reported.

The effect of Cu(II), Hg(II), and Pb(II) on the uptake of Cd(II) by activated carbon was investigated by Krishnan

Table 3

Metal Systems	μmol/gram
Copper alone	440
Cu in Cu-Cd	350
Cu in Cu-Zn	370
Cu in Cu-Cd-Zn	360
Cadmium alone	360
Cadmium in Cu-Cd	90
Cadmium in Cd-Zn	250
Cadmium in Cu-Cd-Zn	85
Zinc alone	375
Zinc in Cu-Zn	50
Zinc in Cd-Zn	130
Zinc in Cu-Cd-Zn	70

Table 4

Metal Ions	Promoters	Suppressor
Hg(II)	Ag(I)>Pb(II)	Cu(II)
Ag(I)	Cd(II)>Cu(II)	Hg(II)
Pb(II)	Cd(II)>Cu(II)	Hg(II)>Cu(II)
Cd(II)	–	Cu(II)>Ag(I)>Ni(II)
Cu(II)	Ni(II)~Cd(II)	Ag(I)>Hg(II)>Cr(VI)>Pb(II)
Cr(VI)	Cu(II)	–
Ni(II)	–	Ch(II)>Cd(II)

and Anirudhan (37). The removal of Cd(II) was reported to be 98.8% in absence of any co-ions. The same decreases to 83.3%, 79.1%, and 72.1%, respectively, when Cu(II), Hg(II), and Pb(II) ions are present in a 1:1 ratio. The results further showed that a 72.2%, 70.5%, and 60.6% reduction in Cd(II) removal was observed when Cu(II), Hg(II), Pb(II) ions were present at a molar ratio of 1:2. The reduction may be because of the competitive ion effect between Cd(II) and co-ions for the adsorption sites available on the carbon surface. Based on these results, it was concluded that Pb(II) ions may be stronger competitive ions than Hg(II) and Cu(II) removal by SA-S-C. The results can also be explained by the selectivity sequence of the most common cations on the adsorbent surface. It was also observed that, among the cations used, interference of Pb(II) ion is highest, followed by Hg(II) and Cu(II). The observed order of interference was the same as that of their increasing ionic radii, i.e., their decreasing hydrated ionic radii. The smaller the hydrated ionic radii, the greater its efficiency to active groups of the adsorbent, which suggests that the energy required in the dehydration of the metal ions, in order that they could occupy a site in the adsorbent, plays an important role in determining the selectivity series for the metal ions.

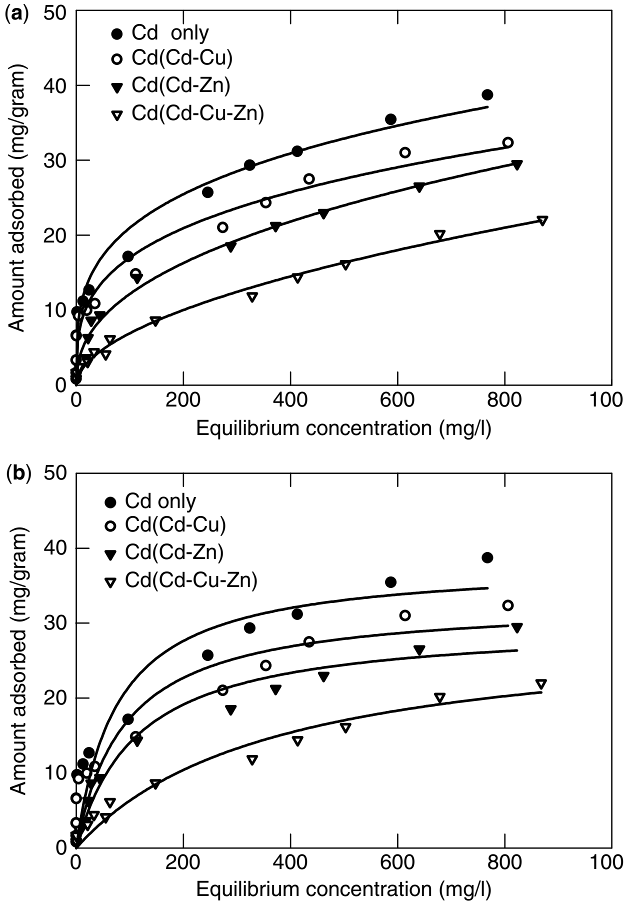

Figure 1. Multicomponent adsorption of Cd(II) on activated carbon developed from bagasse. Solid lines represent the fitting of data by (**a**) Freundlich and (**b**) Langmuir isotherms.

Recently, Mohan and Singh (38) reported the batch sorption isotherm studies to obtain the data required in the design and operation of column reactors for treatment of cadmium- and zinc-bearing wastewater both in single and multicomponent systems. The metals chosen for the investigation in single component studies were Cd (II) and Zn(II). In multicomponent system investigations, four binary systems, Cd(Cd-Cu), Cd(Cd-Zn), Zn(Zn-Cu), and Zn(Zn-Cd), and two ternary systems, Cd(Cd-Cu-Zn) and Zn(Zn-Cu-Cd), were selected. The adsorption isotherms for binary, ternary, and multicomponent systems were obtained at pH 4.5. A 1:1 ratio was used to determine the effect of other metal ions on the adsorption of Cd(II) and Zn(II) on the prepared carbon. The Freundlich and Langmuir adsorption isotherms for Cd(II) and Zn(II) in binary and ternary systems are presented in (Figs. 1 and 2), respectively (38). The results clearly revealed that the presence of other metal ions compete with Cd(II) and Zn(II) ions. It was observed that Cu(II) had the least interfering capacity among Cd(II), Zn(II), and Cu(II) ions in binary systems. Both Langmuir and Freundlich isotherms adequately described the data over the entire range of concentration, and corresponding parameters are presented in Table 5. The effect of ionic interaction (36,39) on the sorption process may also be represented by the ratio of the sorption capacity for one metal ion in the presence of the other metal ions, Q^{mix}, to the sorption capacity for the same metal when it is present alone in the solution, Q^0, such that when:

$$\frac{Q^{\mathrm{mix}}}{Q^0} > 1, \text{ the sorption is promoted by the presence of other metal ions}$$

$$\frac{Q^{\mathrm{mix}}}{Q^0} = 1, \text{ no observable net interaction exists}$$

$$\frac{Q^{\mathrm{mix}}}{Q^0} < 1, \text{ sorption is suppressed by the presence of other metal ions}$$

The values of $\frac{Q^{\mathrm{mix}}}{Q^0}$ were found to be less than 1, as presented in Table 5. The prepared activated carbon followed the same trend, that is, Q^{mix} decreased in the following order for the adsorption of Cd(II) and Zn(II) in multicomponent systems:

$$\mathrm{Cd(II)} < \mathrm{Cd\text{-}Cu} < \mathrm{Cd\text{-}Zn} < \mathrm{Cd\text{-}Cu\text{-}Zn} \qquad \text{for Cd(II)}$$

$$\mathrm{Zn(II)} < \mathrm{Zn\text{-}Cu} < \mathrm{Zn\text{-}Cd} < \mathrm{Cd\text{-}Cu\text{-}Zn} \qquad \text{for Zn(II)}$$

Table 5. Freundlich and Langmuir Isotherm Constants for Singe and Multicomponent Adsorption of Cd(II) and Zn(II) on Activated Carbon Developed from Bagasse Carbon

Metal Ions	System	Freundlich Constants				Langmuir Constants			
		K_F	1/n	R^2	K_F^{mix}/K_F	Q^0	$b \times 10^{-3}$	R^2	Q^{mix}/Q^0
Cd	Cd alone	5.78	0.28	0.9760	–	38.03	13.2	0.8886	–
Cd	Cd + Cu	4.30	0.29	0.9706	0.74	33.11	11.0	0.8829	0.87
Cd	Cd + Zn	1.74	0.02	0.9864	0.30	30.02	8.8	0.9678	0.79
Cd	Cd + Cu + Zn	0.59	0.03	0.9864	0.10	29.77	2.7	0.9583	0.78
Zn	Zn alone	5.62	0.25	0.9659	–	31.11	14.2	0.8683	–
Zn	Zn + Cu	3.96	0.27	0.9674	0.70	26.00	13.5	0.8723	0.84
Zn	Zn + Cd	1.42	0.41	0.9868	0.25	23.09	8.00	0.9794	0.74
Zn	Zn + Cu + Cd	0.79	0.45	0.9792	0.14	19.02	5.4	0.9804	0.61

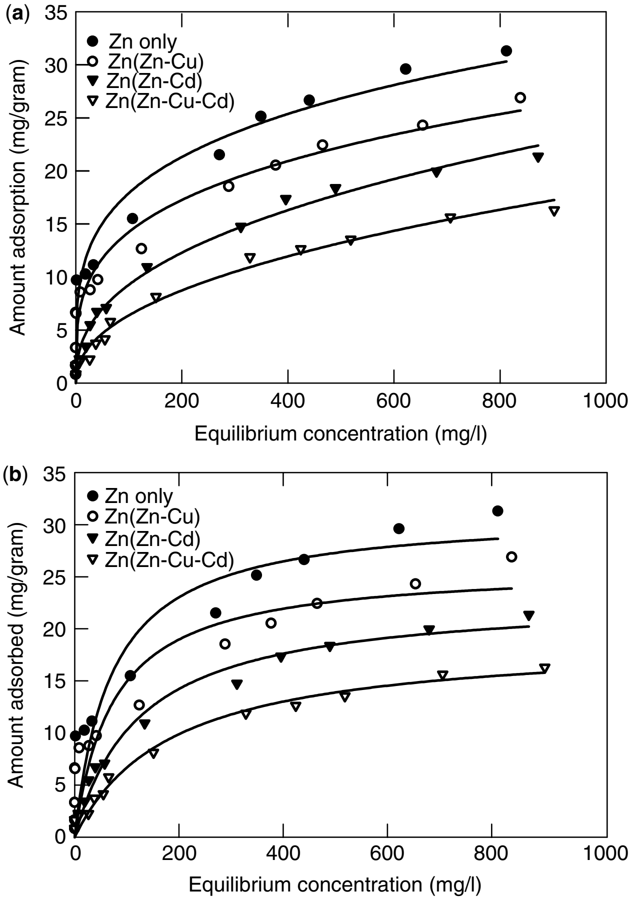

Figure 2. Multicomponent adsorption of Zn(II) on activated carbon developed from bagasse. Solid lines represent the fitting of data by (**a**) Freundlich and (**b**) Langmuir isotherms.

Overall, it was concluded that the adsorption capacity of activated carbon for Cd(II) and Zn(II) decreased more in ternary systems as compared with binary systems.

In acid mine wastewater, some other metal ions are always present besides iron and manganese; therefore, it is desirable to see the effect of other metal ions on the adsorption capacity of different activated carbons. A very important study in this regard was carried out by Mohan and Chander (39) where the adsorption of four metal ions, i.e., Mn(II), Fe(II), Zn(II), and Ca(II), were conducted in binary, ternary, and multicomponent systems on different types of activated carbons. The adsorption isotherms for binary, ternary, and multicomponent systems were obtained at pH 3.5 and 25 °C. The concentration range of 5.0×10^{-5} to 9.0×10^{-3} M was investigated, and a 1:1 ratio was used to determine the effect of Mn(II), Ca(II), and Zn(II) on the adsorption of Fe(II) on carbons. The Freundlich and Langmuir adsorption isotherms for Fe(II) in the absence and presence of interfering metal ions were determined.

Both Langmuir and Freundlich adsorption isotherms were found to adequately describe the data over the entire range of concentration, and the Langmuir and Freundlich isotherm parameters are presented in Tables 6 and 7, respectively. The detailed analysis of the regression coefficients showed that the data was slightly better fitted by Freundlich adsorption isotherm for multicomponent systems. The adsorption isotherms for different carbons revealed that, when Ca(II), Mn(II), and Zn(II) were present in the system with Fe(II), the interference did not change the adsorption of Fe(II) in the low concentration range, whereas a competitive uptake, with Fe(II) being preferentially adsorbed by carbons, took place at the higher concentrations. Carbon B0, which showed abnormal behavior and adsorption, was found to increase in the presence of other metal ions.

It was further concluded that the presence of manganese and/or zinc had limited effect on the capacity of carbons for Fe(II) in comparison with calcium. Thus, overall it was found that Ca(II) had the highest interfering capacity. The adsorption capacity of various activated carbons for Fe(II) in the presence and absence of Mn(II), Zn(II), and Ca(II) are presented in Table 6, whereas the Freundlich constants are presented in Table 7. Thus, when two or more metal ions are present in the solution, they seem to compete for the adsorption sites as the metal ions are adsorbed on the same sides.

The values of $\frac{Q^{\text{mix}}}{Q^0}$ are found to be less than 1, as presented in Table 6, except for carbon B0, thereby confirming the suppression in the adsorption of Fe(II) by the presence of other metal ions. These results are consistent with the adsorption isotherms obtained for Fe(II) in the absence and presence of various metal ions. It is clear from Table 6 that carbons can be divided into two different categories, i.e., wood-based activated

Table 6. Langmuir Isotherm Constants for Multicomponent Metal Ion Adsorption on Different Types of Activated Carbons

Activated Carbons	Parameters	Fe(II)	Fe(Fe-Ca)	Fe(Fe-Mn)	Fe(Fe-Zn)	Fe(Fe-Mn-Zn)	Fe(Fe-Mn-Zn-Ca)
W1	Q^0	22.27	15.35	17.65	16.26	13.79	12.36
	$b \times 10^{-3}$	38.63	42.14	38.67	53.86	30.36	44.75
	R^2	0.9606	0.8615	0.7013	0.7223	0.9609	0.7581
	Q^{mix}/Q^0	–	0.69	0.79	0.73	0.62	0.56
W2	Q^0	25.60	18.16	19.73	18.78	17.16	13.26
	$b \times 10^{-3}$	70.74	37.59	117.37	90.70	34.22	81.73
	R^2	0.9409	0.7529	0.7474	0.7643	0.9491	0.7516
	Q^{mix}/Q^0	–	0.71	0.77	0.73	0.67	0.52
W3	Q^0	21.67	14.58	16.86	16.10	13.64	13.06
	$b \times 10^{-3}$	53.97	82.39	78.06	86.71	34.91	21.71
	R^2	0.9302	0.8054	0.6983	0.7206	0.9735	0.6746
	Q^{mix}/Q^0	–	0.67	0.78	0.74	0.63	0.60
B0	Q^0	14.59	15.2	22.98	22.02	18.78	14.44
	$b \times 10^{-3}$	2.643	21.485	33.22	32.558	75.111	32.318
	R^2	0.8334	0.8525	0.6847	0.6700	0.8929	0.7157
	Q^{mix}/Q^0	–	1.0418	1.57505	1.509253	1.287183	0.989719
B4	Q^0	28.78	16.58	22.32	20.84	18.21	15.3184
	$b \times 10^{-3}$	27.560	14.609	35.662	34.002	23.964	18.943
	R^2	0.8773	0.8722	0.8257	0.7682	0.9549	0.7780
	Q^{mix}/Q^0	–	0.58	0.7241	0.7755	0.63	0.53
B3	Q^0	25.61	12.05	16.90	15.80	16.47	9.74
	$b \times 10^{-3}$	2.28	1.69	3.73	4.46	1.80	2.98
	R^2	0.9287	0.9809	0.9498	0.9672	0.9858	0.9728
	Q^{mix}/Q^0	–	0.47	0.66	0.61	0.64	0.38
C1	Q^0	46.35	21.77	27.40	26.05	23.85	18.31
	$b \times 10^{-3}$	36.85	34.77	77.15	31.73	23.72	43.12
	R^2	0.92668	0.619730	0.60214	0.5852	0.6432	0.7003
	Q^{mix}/Q^0	–	0.47	0.59	0.56	0.51	0.40
Lignite	Q^0	34.22	19.59	20.88	19.96	12.39	11.23
	$b \times 10^{-3}$	28.98	32.87	68.67	33.50	55.21	53.12
	R^2	0.8768	0.8912	0.8714	0.8986	0.9662	0.9238
	Q^{mix}/Q^0	–	0.5725	0.6101	0.5832	0.3621	0.3282

Table 7. Freundlich Isotherm Constants for Multicomponent Metal Ion Adsorption on Different Types of Activated Carbons

Activated carbons	Parameters	Fe(II)	Fe(Fe-Ca)	Fe(Fe-Mn)	Fe(Fe-Zn)	Fe(Fe-Mn-Zn)	Fe(Fe-Mn-Zn-Ca)
W1	K_F	90.70	36.90	34.78	27.00	44.30	34.12
	$1/n$	0.21	0.13	0.10	0.10	0.17	0.14
	R^2	0.9299	0.8715	0.7857	0.7957	0.9173	0.7240
W2	K_F	90.34	33.67	38.77	39.91	57.00	39.37
	$1/n$	0.18	0.09	0.09	0.10	0.18	0.52
	R^2	0.9452	0.8170	0.8235	0.8343	0.8712	0.7423
W3	K_F	80.67	29.37	28.80	25.42	51.54	27.83
	$1/n$	0.19	0.10	0.08	0.07	0.19	0.12
	R^2	0.9163	0.8653	0.7800	0.7953	0.8671	0.7989
B0	K_F	57.16	29.04	67.88	57.86	63.04	27.60
	$1/n$	0.25	0.10	0.15	0.14	0.16	0.097
	R^2	0.8257	0.84426	0.9054	0.9787	0.9938	0.7268
B4	K_F	281.88	123.33	112.86	109.11	265.72	100.18
	$1/n$	0.41	0.42	0.32	0.32	0.47	0.39
	R^2	0.9447	0.9667	0.9726	0.9725	0.9656	0.9614
B3	K_F	1.87	41.94	70.03	53.18	101.13	44.18
	$1/n$	0.87	0.14	0.16	0.13	0.24	0.16
	R^2	0.3759	0.8618	0.8376	0.8419	0.9675	0.7561
C1	K_F	444.84	37.36	–	–	–	38.1925
	$1/n$	0.28	0.08	–	–	–	0.11
	R^2	0.9325	0.6869	–	–	–	0.7738
Lignite	K_F	214.24	66.190	69.79	60.358	8.345	9.5887
	$1/n$	0.2455	0.1732	0.1618	0.1580	0.1367	0.1396
	R^2	0.9215	0.9061	0.9030	0.8971	0.9070	0.9067

carbons follow the same trend (Fe-Mn<Fe-Zn<Fe-Ca<Fe-Mn-Zn<Fe-Mn-Zn-Ca) whereas the other carbons behave in a similar fashion (Fe-Mn<Fe-Zn< Fe-Mn-Zn<Fe-Ca < Fe-Mn-Zn-Ca). Also, the effect of other interfering metal ions on the adsorption of Fe(II) was found to be less on coconut-based activated carbon followed by coal-based and wood-based activated carbon, respectively. Overall, it was concluded that the adsorption capacity of different carbons decreased more in ternary and quaternary systems as compared with binary systems.

Fixed bed studies were also conducted on solution containing a mixture of several metal ions. For these studies, a bed filled with activated carbon was treated with a solution containing 60 mg/L Fe (II) and 50 mg/L each of Mn (II), Zn (II), and Ca (II). The hydraulic flow rate was 1.7 mL/min. The breakthrough curves are presented in (Figs. 3 and 4). The results are plotted as dimensionless concentration (C_e/C_0) vs. effluent volume. Initially, all the metal ions were adsorbed nonselectively and a metal ion-free effluent was produced. With continued treatment, manganese began to escape the column, followed by zinc and then iron. One can see that manganese and zinc, which initially adsorbed nonselectively, are released with continued passage of iron-bearing solutions. Iron could be adsorbed selectively during the period of testing. Similar results were obtained with other carbons. These results demonstrate that metal ions can be separated and possibly recovered in useful form by treating metal-bearing wastewaters with activated carbons. The studies were also performed in the single column as well as when columns were connected in series. The difference in various parameters obtained are presented in Table 8.

The uptake of Pb(II), Hg(II), and Cr(VI) in a multicomponent system was studied by Srivastava et al. (40) on the activated carbon developed from fertilizer waste material. The sorption efficiency of the adsorbent was found to decrease by 15% and 7%, respectively, for lead and mercury in the presence of each other. No decrease, however, was reported for chromium in the presence of mercury, but the uptake of mercury reduced by 8–10% in the presence of chromium. The reduction in the sorption of metal ions from the mixtures was observed to be almost the same for different levels of interference. Competitive adsorption of various metal ions in the presence of each other is presented (40,41) in Tables 9–12.

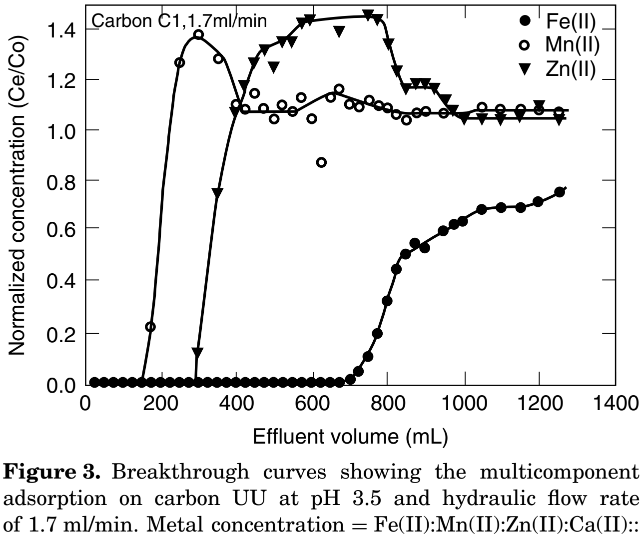

Figure 3. Breakthrough curves showing the multicomponent adsorption on carbon UU at pH 3.5 and hydraulic flow rate of 1.7 ml/min. Metal concentration = Fe(II):Mn(II):Zn(II):Ca(II):: 60:50:50:50 ppm.

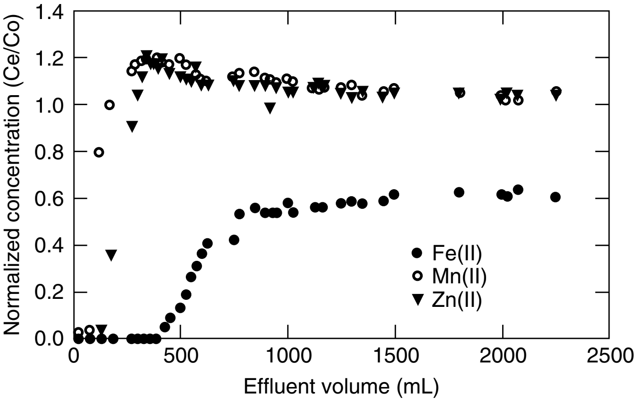

Figure 4. Breakthrough curves showing the multicomponent adsorption on carbon B3 at pH 3.5 and hydraulic flow rate of 1.7 ml/min. Metal concentration = Fe(II):Mn(II):Zn(II):Ca(II):: 60:50:50:50 ppm.

Table 8. Comparison Between Single- and Three-Column System for Fe(II) Adsorption

Types of Column	Number of Bed Volume	EBCT* (min)	$Q^\dagger$ (ml/min)	V_b	BV_b	Column Capacity (mg/g)
Single column	22	9	2.7	85 ml	4	3.4
Three column system	235	24	10	15 liter	63	18.7

*Empty bed contact time or residence time.

†Volumetric flow rate.

V_b = Breakthrough volume.

BV_b = Bed volumes at breakthrough.

[Concentration of metal ions in the inlet: 60 ppm for Fe(II); 50 ppm each for Mn(II), Zn(II), and Ca(II).]

Parameter	Value
Column diameter (cm)	2.5
Column bed height (cm)	16
Bed volume (ml)	78 each, i.e., 235
Flow rate (ml/min)	10
Residence time (min)	24
Particle size (mesh)	20 × 40
Amount of carbon	50 gm in each column, i.e., 150 gm
Concentration of metal ions (ppm)	Fe^{2+}: Mn^{2+}: Zn^{2+}: Ca^{2+}:: 60:50:50:50
pH of the solution	3.5

COMPETITIVE SORPTION OF ORGANICS

Yen and Singer (42) studied the adsorption of phenols and substituted phenols on activated carbon. Mixtures of phenolic compounds were taken to explore the Ideal Adsorbed Solution (IAS) theory using an improved method of calculation that had been developed to describe the multicomponent adsorption. The IAS model with the

Table 9. Competitive Adsorption of Pb(II) in Presence of Hg(II)

Pb(II) Concentration (mol/l)	Hg(II) Concentration (mol/l)	Amount of Pb(II) Adsorbed in Absence of Hg(II), (mol/g)	Amount of Pb(II) Adsorbed in Presence of Hg(II) (mol/g)
5×10^{-3}(fixed)	1×10^{-4}	4.7×10^{-3}	4.2×10^{-3}
	5×10^{-4}	4.7×10^{-3}	4.0×10^{-3}
	1×10^{-3}	4.7×10^{-3}	4.0×10^{-3}

Table 10. Competitive Adsorption of Hg(II) in Presence of Pb(II)

Hg(II) Concentration (mol/l)	Pb(II) Concentration (mol/l)	Amount of Hg(II) Adsorbed in Absence of Pb(II), (mol/g)	Amount of Hg(II) Adsorbed in Presence of Pb(II) (mol/g)
5×10^{-3}(fixed)	1×10^{-4}	4.0×10^{-3}	2.3×10^{-3}
	5×10^{-4}	2.7×10^{-3}	2.6×10^{-3}
	1×10^{-3}	2.7×10^{-3}	2.6×10^{-3}

Table 11. Competitive Adsorption of Cr(VI) in Presence of Hg(II)

Cr(VI) Concentration (mol/l)	Hg(II) Concentration (mol/l)	Amount of Cr(VI) Adsorbed in Absence of Hg(II), (mol/g)	Amount of Cr(VI) Adsorbed in Presence of Hg(II); (mol/g)
5×10^{-3}(fixed)	1×10^{-4}	4.0×10^{-3}	4.0×10^{-3}
	5×10^{-4}	4.0×10^{-3}	4.0×10^{-3}
	1×10^{-3}	4.0×10^{-3}	3.9×10^{-3}

Table 12. Competitive Adsorption of Hg(II) in Presence of Cr(VI)

Hg(II) Concentration (mol/l)	Cr(VI) Concentration (mol/l)	Amount of Hg(II) Adsorbed in Absence of Cr(VI), (mol/g)	Amount of Hg(II) Adsorbed in Presence of Cr(VI) (mol/g)
5×10^{-3}(fixed)	1×10^{-4}	3.0×10^{-3}	2.7×10^{-3}
	5×10^{-4}	2.7×10^{-3}	2.5×10^{-3}
	1×10^{-3}	2.7×10^{-3}	2.5×10^{-3}

modified calculation method was tested for its performance on ten sets of binary and ternary phenolic mixtures. The Langmuir model was also taken just to compare the results, but the IAS model was found to be successful in precisely describing the competitive adsorption behavior of phenolic mixtures.

The impact of the presence of molecular oxygen on multisolute adsorption of phenols on granular activated carbon was evaluated by Sorial et al. (43). Adsorption equilibrium for binary mixtures of phenol/O-cresol and ternary mixtures of phenol/O-cresol/3-ethylphenol was carried out at 23 °C using three different initial concentration combinations. Adsorption was carried out under oxic and anoxic conditions. The ideal adsorbed solution theory, using Myers equation for correlating the single-solute anoxic isotherms, was found to accurately describe the competitive adsorption behavior of these phenolic mixtures under anoxic conditions. When the Freundlich model was used to describe the single-solute sorption, the deviations were found to increase. It was concluded that poor model predictions for the oxic isotherms were attributed to the presence of molecular oxygen, which promotes the polymerization of solutes on the surface of granular activated carbon. In continuation of this study, the same authors (44) conducted the adsorption of these phenolic compounds in fixed bed GAC adsorbers. The adsorption breakthrough curves were obtained for a single-solute system, i.e., phenol, O-cresol, and 3-ethylphenol, as well as for multisolute systems, i.e., phenol/O-cresol and phenol/O-cresol/3-ethylphenol. The plug through homogenous surface diffusion model was evaluated as a predictor of GAS adsorber, whereas the binary and ternary solute calculations were performed using kinetic parameters determined from a single-solute system. The ideal adsorbed solution theory was implemented. The model for binary and ternary solute systems agrees well with the experimental data collected under anoxic conditions.

The adsorption of 2,4,6-trichlorophenol (TCP) and N-[2-(2,4,6-trichlorophenoxy)propyl] amine(BTS40348) in single and binary systems was carried out on four different activated carbon at pH 4.0 and 9.0 by Garner et al. (45). Competitive adsorption was observed, and at respective pH optima, reduction in adsorption efficiencies were found to be 10% for TCP and 43% for BTS40348.

The interference of catioinc, nonionic, and anionic detergents on the adsorption of phenols was studied as a function of the concentration of detergents (46). As all types of surfactants are found to interact to a varying degree with activated carbons, the effect of anionic detergent (Manoxol 1B), nonionic detergent (triton), and cationic detergent (cetylpyridinium chloride) on the uptake of 2,4,6-trinitrophenol, 4-nitrophenol, 4-chlorophenol, and 1,3-dihydroxybenzene by the activated carbons was studied. The adsorption of 2,4,6-trinitrophenol was found to decreased by 1.3–3.3% in the presence of increasing anionic detergent concentration. A decrease in adsorption efficiency by 3.0–6.0% was observed in the case of 4-nitrophenol. In the case of 4-chlorophemol, the uptake decreased by 5–7%, whereas a 2–5% decrease was observed for 1,3 dihydroxybenzene. In the case of nonionic surfactant, the removal efficiencies of activated carbon was found to decrease by 2–4%, 3–8%, 3.7–7.5%, and 3.3–6.7% for 2,4,6-trinitrophenol, 4-nitro phenol, 4-chloro phenol, and 1,3-dihydroxybenzene, respectively. In the case of cationic detergent, a decrease in the scavenging efficiency of activated carbon by 2.7–6.7% for 2,4,6-trinitrophenol and 4.0–10.0% for 1,3-dihydroxybenzene was observed. Thus, the maximum effect was found with cationic detergent and the minimum with anionic detergent. Nonionic detergent falls in between the two. Authors explained that the decrease in the scavenging efficiency of activated carbon for phenols in the presence of surfactants is consistent with the act that the particles of the carbon surface are negatively charged.

Jossens et al. (47) studied the sorption data at 20 °C for six dilute aqueous bi-solute systems in activated carbon. The six systems were phenol/p-nitrophenol, p-nitrophenol/p-chlorophenol, p-nitrophenol/benzoic acid, p-chlorophenol/phenyl acetic acid, p-nitrophenol/o-phenyl phenol, and 2,4-dichlorophenol/dodecyl-benzol sulfonic acid. A new three-parameter adsorption isotherm was presented that represents the single-solute data very well. For bi-solute systems where dissociation is negligible, the calculated individual adsorption agrees with experimental data within 2%.

The adsorption of acetaldehyde, acetone, 2-butanone, pyridine, and phenol from binary aqueous solutions on activated carbon was reported by Barton (48). The author has applied the Dubbinin–Radushkevich (DR) equation with some slight modifications. The authors explained that the displacement adsorption of 2-butanone and pyridine is driven mainly by the displacement enthalpy. As desorption of water involves an increase in motional entropy, it was therefore concluded that strong adsorbed solutes 2-butanone and pyridine produce a large compensating entropy decrease. On the other hand, more weakly adsorbed substances, such as acetone and acetaldehyde, were not able to compensate for the entropy increase associated with the desorbed water on adsorption. Single and multisolute adsorption isotherms studies of three phenolic compounds, i.e., gallic acid, *p*-hydroxybenzoic acid, and syringic acid, were investigated at 20 °C, 30 °C, and 40 °C using a bituminous coal-based activated carbon (49). The capacity of the activated carbon used to adsorb these compounds follows the order: syringic acid > *p*-hydroxybenzoic acid > gallic acid. In binary and ternary components, experimental data suggested that interactions between adsorbates improve the adsorption capacity of some of the phenolic acid compounds. On the contrary, at high organic concentrations, adsorbed gallic acid was partially removed from the activated carbon surface because of the presence of the other components. A two-component isotherm (phenol/2,6-dichlorophenol) adsorption was also studied by Mameli et al. (50) in order to test activated carbon behavior during competitive adsorption.

Kim and Lordgooei (51,52) reported the adsorption modeling of various volatile organic compounds in single and multicomponent systems. A Dubinin–Astakhov thermal equation of equilibrium adsorption (DA-TEEA) for single component and ideal/real adsorbed solution theories (IAST/RAST) for multicomponent systems were presented successfully. In an another study, Semmens et al. (53,54) studied the influence of pH and coagulation on the removal of organics by granular activated carbon. Okazaki et al. (55,56) reported the multicomponent adsorption of organics from water, whereas Ha et al. (57) developed a predictive isotherm model to evaluate the extent of bioregeneration of granular activated carbon loaded with phenol and 2,4-dichlorophenol (2,4-DCP) in single and bisolute systems. A bisolute system was taken up for assessing the competitive adsorption. The effect of byproducts, which were generated during biodegradation of substrate and measured as COD, on bioregeneration in the bisolute was also investigated. Freundlich adsorption capacity of 2,4-DCP was found to be more as compared with phenol in both single and bisolute systems. Byproducts in the bulk solution brought an adverse effect on adsorption capacity of GAC in all cases. By taking into account the byproduct effect on adsorption, the Freundlich isotherms were used to formulate a predictive model of bioregeneration.

The adsorption of benzoic acid and p-nitrophenol (PNP) at 25 °C was performed by Chern and Chien (58) in a binary system on GAC. The sorption experimental data were fitted to the extended Langmuir isotherm model successfully. The experimental data and the isotherm model parameters showed that the GAC used in this study had a higher affinity to PNP than benzoic acid. Three-column tests were performed to determine the breakthrough curves and effluent solution pH with varying feed compositions. The weakly adsorbed BA exhibited an intermediate zone of effluent concentration higher than its feed one. The authors predicted that the breakthrough curves with varying feed compositions could be predicted by the nonlinear wave propagation theory satisfactorily, only the adsorption isotherm models were required to construct the composition path diagram with which the breakthrough curves could be predicted. In an important study, Bulloch et al. (59) developed a thermodynamic model to predict adsorption equilibrium in the international space station water processor's multifiltration beds. The model was able to predict the multicomponent adsorption equilibrium behavior using single component isotherm parameters and fictitious components representing the background matrix. The fictitious components

Table 13. Competitive Adsorption of Phenols in Presence of Each Other

Adsorbate	Adsorbate Concentration mol/l	Amount Adsorbed mol/g	Amount Adsorbed in Presence of 2,4,6-Trinitrophenol (1×10^{-5} M) mol/g	Amount Adsorbed in Presence of 4-Nitrophenol (1×10^{-5} M) mol/g	Amount Adsorbed in Presence of 4-Chlorophenol (1×10^{-5} M) mol/g	Amount Adsorbed in Presence of 1,3 Dihydroxy Benzene (1×10^{-5} M) mol/g
2,4,6-trinitro phenol	1×10^{-4}	1×10^{-4}	—	—	—	X
2,4,6-trinitro phenol	1×10^{-3}	7.50×10^{-4}	—	—	—	X
4-nitrophenol	1×10^{-4}	0.9×10^{-4}	0.84×10^{-4}	—	—	X
4-nitrophenol	1×10^{-3}	5.00×10^{-4}	4.80×10^{-4}	—	—	X
4-chlorophenol	1×10^{-4}	0.73×10^{-4}	0.65×10^{-4}	0.70×10^{-4}	—	X
4-chlorophenol	1×10^{-3}	4.00×10^{-4}	3.50×10^{-4}	3.80×10^{-4}	—	X
1,3 dihydroxy benzene	1×10^{-4}	0.70×10^{-4}	0.55×10^{-4}	0.65×10^{-4}	0.70×10^{-4}	X
1,3 dihydroxy benzene	1×10^{-3}	3.00×10^{-4}	2.50×10^{-4}	2.75×10^{-4}	2.90×10^{-4}	X

were determined by fitting total organic carbon and tracer isotherms with the ideal adsorbed solution theory. Multicomponent isotherms using a wastewater with high surfactant and organic compound concentrations were used to validate the equilibrium description on a coconut-shell-based granular activated carbon (GAC), coal-based GAC, and a polymeric adsorbent.

The adsorption of three barbiturates—phenobarbital, mephobarbital, and primidone—from simulated intestinal fluid (SIF), without pancreatin, by activated carbon was studied by Wurster et al. (6). The competitive Langmuir-like model, the modified competitive Langmuir-like model, and the LeVan–Vermeulen model were each fit to the data. Excellent agreement was obtained between the experimental and predicted data using the modified competitive Langmuir-like model and the LeVan–Vermeulen model. The agreement obtained from the original competitive Langmuir-like model was less satisfactory. The results of these studies indicate that the adsorbates were found to compete for the same binding sites on the activated carbon surface. The results demonstrated that it is possible to accurately predict multicomponent adsorption isotherms using only single-solute isotherm parameters.

Rozada et al. (60) studied the adsorption of methylene blue and saphranine from single and bisolute systems using the activated carbons developed by chemical activation and pyrolysis of sewage sludges.

The effect of presence of an anionic (Manoxol 1B), nonionic (Triton), and cetyl pyridinium chloride detergent on the uptake of dinitrophenol by activated carbon developed from fertilizer waste slurry was studied by Srivastava et al. (61). The dinitrophenol adsorption showed a decreased (2–8%) w/w with increasing concentrations of anionic, nonionic, and cationic surfactants. The uptake of DNP was also observed in the presence of NaCl, $BaCl_2$, and $AlCl_3$ at a fixed adsorbent concentration. No effect of NaCl was observed on DNP uptake between pH 2 to 4, but at pH 10, adsorption increases by 66% w/w. The presence of $BaCl_2$ and $AlCl_3$ did not affect the uptake of DNP to a significant extent ($2 < 2\%$ w/w). The influence of anions (ClO_4^-, PO_4^{3-}, SO_4^{2-}, and NO_3^-) on the uptake of dinitrophenol was also reported. It was observed that anions have a negligible effect on the sorption of DNP. Various explanations were given for the competitive adsorption, some of which are listed below:

1. Interaction in solution between salts and organics to produce a change in the distribution of the organic species present, thereby influencing the rate or the extent of adsorption.
2. Interactions between salts and the adsorbed organics, resulting in the alteration of the packing, spacing, or alignment of the adsorbed molecules.
3. Interactions between salts and the adsorbate and adsorbent to perhaps create new or particularly favorable adsorption sites.

The enhanced adsorption of DNP in the presence of slats was thus attributed to the interaction between cations and organics in the solution or at a solid surface. Such interaction influences adsorption through the alteration of solubility or the degree of ionization of the organic molecule via the common ion effect, ion pairing, or complexation. The enhance in the uptake of DNP was attributed to ion pair formation

The adsorption of malachite in presence of Manoxol 1B detergent was reported by Gupta et al. (62). The removal was found to decrease by 1.5–2.0%.

Sung-Ryong and Vinitnantharat (63) studied the sorption of phenol and 2,4 dichlorophenol in single and bisolute systems. It was concluded that 2,4 DCP was a stronger adsorbate than phenol in both single and bisolute systems. On desorption of 2,4 DCP, the small fraction of sorbed compounds was reversible, but phenol has comparatively high reversibility.

The competitive adsorption of three cationic polymers, namely JR125, JR400, and JR30M, and a cationic surfactant (cetyl trimethyl ammonium bromide) onto a silica surface from a low ionic strength medium was investigated by Harrison et al. (64). It was reported that competition between polymers showed that smaller molecules were adsorbed, initially preventing the subsequently adsorption on larger polymers. Adsorption from the combined polymer and the smallest polymer (JR 125) was excluded from the surface; the intermediate molecular mass polymer (JR

400) was almost unaffected; the surfactant and the largest molecular mass polymer (JR 30M) were partially excluded from the surface.

Competitive adsorption of substituted phenols by activated carbon developed from fertilizer waste slurry was investigated by Srivastava and Tyagi (65). The uptake of 4-chlorophenol, 4-nitophenol, 2,4,6-trinitrophenol, and 1,3-dihydroxybenzene in the presence of each other provided some interesting data, which is presented in Table 13. Surprisingly, the adsorption of 2,4,6-trinitrophenol did not get at all affected by 4-nitrophenol, 4-chlorophenol, and 1,3-dihydroxybenzene. However, the uptake of 4-nitophenol went down by 6.0% in the presence of 2,4,6-trinitrophenol. No decrease, however, was observed in the presence of 4-chlorophenol and 1,3-dihydroxybenzene. Similarly, the adsorption of 4-chlorophenol was found to reduced by 5.0% and 12.5% in the presence of 4-nitophenol and 2,4,6-trinitrophenol. 4-chlorophenol was not affected by 1,3-dihydroxybenzene. The sorption capacity of 1,3-dihydroxybenzene decresed by 3.3%, 8.3%, and 16.7% in presence of 4-chlorophenol, 4-nitrophenol, and 2,4,6-trinitrophenol, respectively. The electron withdrawing nature of nitro as well as chloro reduced the electron density in the ring in phenols. The effect was highest in 2,4,6-trinitrophenol and lowest in 1,3-dihydroxybenzene. 4-nitophenol and 4-chlorophenol occupy the second and third position if a gradation of all four adsorbates is made vis-a-vis the reduction in electron density of the pi system of the ring, which explained that the removal of 2,4,6-trinitrophenol did not get reduced in the presence of the other three phenols. The observed fact that 1,3-dihydroxybenzene also did not reduce the uptake of other phenols could also be attributed to the above-mentioned reason.

MULTICOMPONENT KINETICS

Similar to equilibrium studies, multicomponent kinetics is also a very important aspect of any adsorption study. It is based on the single component kinetics model. Out of the various kinetic models, the film-solid diffusion model is the best one as far as theoretical background and optimum mathematical conveyance are concerned, and therefore opted by a number of researchers from time to time in multicomponent systems. The details about this model are also very much documented in various articles (12,15,66–68).

The kinetic adsorption of mixtures of phenolic compounds onto a polymeric adsorbent from aqueous solution was studied by Mijangos et al. (69). Van Laar's equation was applied to evaluate the influence of concentration on diffusion. The adsorption kinetics for phenol and p-cresol mixtures at different initial concentration ratios were studied and adjusted to the mentioned kinetic model. They concluded that adsorption from multicomponent aqueous solution is a surface diffusion-controlled process.

The binary soprtion of Cu-Cd, Cd-Zn, and Cu-Zn onto bone char has been studied using an equilibrium and batch agitation system (70). The sorption capacities and selectivity of metal ions follows the order Cu(II)>Cd(II)>Zn(II), which is a reverse order of the hydrated ionic radii. The binary sorption equilibria were predicted by the ideal adsorbed solution theory (IAST) on the basis of single component isotherm data using a Langmuir or Langmuir–Freundlich isotherm. The overall performance of IAST provided a reasonable curve fitting to the experimental data. The single component film-pore diffusion model was extended to the multicomponent systems to correlate the batch kinetic data by incorporating the shrinking core model and IAST. It was found that all the diffusivities in the binary systems are similar to or less than the pore diffusitivites in single component systems.

Table 14

Metal Ions	Single Component D_p (cm^2/s)	Cd-Cu	Cd-Zn	Cu-Zn
Cd	1.14×10^{-6}	1.20×10^{-6}	2.25×10^{-7}	—
Cu	1.59×10^{-6}	1.30×10^{-6}	—	1.50×10^{-6}
Zn	1.21×10^{-6}	—	1.10×10^{-6}	1.00×10^{-6}

The pore diffusivities of the multicomponent systems using the film-pore diffusion model and IAST are given in Table 14. Although the equal molar and unequal molar equilibrium data can generally be predicted by IAST, the results for the mole ratios 3:7 and 7:3 for the Cd-Cu systems cannot be predicted very well.

BIBLIOGRAPHY

1. Crittenden, J.C. and Weber, W.J., Jr. (1978). Predictive model for design of fixed bed adsorbers; Single-component model verification. *J. Environ. Eng. Div.* **104**(EE3): 433–443.
2. Crittenden, J.C. and Weber, W.J., Jr. (1978). Model for design of multi-component adsorption systems. *J. Environ. Eng. Div.* **104**(EE6): 1175–1195.
3. Crittenden, J.C., Wong, B.W.C., Thacker, W.E., Snoeyink, V.L., and Hinrichs, R.L. (1980). Mathematical modeling of sequential loading in fixed-bed adsorbers. *J. Water Pollution Control Fed.* **51**(11): 2780–2795.
4. Hand, D., Crittenden, J.C., and Thacker, W.E. (1984). Simplified models for design of fixed bed adsorption systems. *J. Environ. Eng. Div.* **110**(2): 441–456.
5. Butler, J.A.V. and Ockrent, C. (1930). Studies in electrocapillarity Part 3. The surface tensions of solutions containing two surface-active solutes. *J. Phys. Chem.* **34**: 2841–2859.
6. Wurster, D.E., Alkhamis, K.A., and Matheson, L.E. (2000). Prediction of adsorption from multicomponent solutions by activated carbon using single-solute parameters. Part 1. *AAPS Pharm. Sci. Tech.* **1**(3): E25.
7. Alkhamis, K.A. and Wurster, D.E. (2002). Prediction of adsorption from multicomponent solutions by activated carbon using single solute parameters. Part II. *AAPS Pharm. Sci. Tech.* **3**(3): 23.
8. Hsieh, C.P., Turian, R.M., and Tein, C. (1977). Multicomponent liquid phase adsorption in fixed bed. *Amer Inst. Chem. Eng. J.* **23**(3): 263–275.
9. Hsieh, C.P., Tsang, M.W., and Hsieh, Y.S. (1985). The removal of cobalt(II) from water by activated carbon. *Amer. Inst. Chem. Eng. Symp. Series.* **81**(243): 85–98.
10. Jain, J.S. and Snoeyink, V.L. (1973). Adsorption from bisolute systems on activated carbon. *J. Water. Pollut. Control. Fed.* **45**(12): 2463–2479.

11. Huang, J.-C. and Steffens, C.T. (1976). Competitive adsorption of organic materials by activated carbon. *Proceedings of the 31st Industrial Waste Conference*. Purdue University.
12. Mathew, A.P. (1975). *Mathematical Modeling of Multicomponent Adsorption in Batch Reactors*. Ph.D. Dissertation. University of Michigan, Ann Arbor, MI.
13. Redlich, O. and Peterson, D.L. (1959). A useful adsorption isotherm. *J. Phys. Chem*. **63**: 1024.
14. Fritz, W. and Schlunder, E.U. (1974). Simultaneous adsorption equilibria of organic solutes in dilute aqueous solutions on activated carbon. *Chem. Eng. Sci.* **29**: 1279–1282.
15. Fritz, W., Merk, W., and Schlunder, E.U. (1981). Competitive adsorption of two dissolved organics onto activated carbon, II Adsorption kinetics in batch reactors. *Chem. Eng. Sci.* **36**: 731–741.
16. Dastgheib, S.A. and Rockstraw, D.A. (2002). A systematic study and proposed model of the adsorption of binary metal ion solutes in aqueous solution onto activated carbon produced from pecan shells. *Carbon* **40**: 1853–1861.
17. Sheindorf, C.H., Rebhun, M., and Sheintuch, M. (1981). *J. Colloid. Interface Sci.* **79**: 136.
18. Sheindorf, C.H., Rebhun, M., and Sheintuch, M. (1982). *Water Res.* **16**: 357.
19. Myers, A.L. and Prausnitz, J.M. (1965). Thermodynamics of mixed-gas adsorption. *Amer. Inst. Chem. Eng. J.* **11**(1): 121–127.
20. Radke, C.J. and Prausnitz, J.M. (1972). Thermodynamics of multisolute adsorption from dilute liquid solutions. *Amer. Inst. Chem. Eng. J*. **18**(4): 761–768.
21. Crittenden, J.C. et al. (1985). Prediction of multi-component adsorption equilibria using ideal adsorbed solution theory. *Environ. Sci. Technol.* **19**: 1030–1043.
22. LeVan, M.D. and Vermeculen, T. (1981). Binary Langmuir like and Freundlich isotherms for ideal adsorbed solutions. *J. Phys. Chem.* **85**: 3247–3250.
23. Mohan, D. and Chander, S. (1999). *Recovery of Valuable Materials from Mine Drainage The Ben Franklin Partnership Program*, Project Number: 976C.1033R-1, Penn State University, University Park, State College, PA.
24. Budinova, T.K., Gergova, K.M., Petrov, N.V., and Minkova (1994). Removal of metal ions from aqueous solution by activated carbons obtained from different raw materials. *J. Chem. Technol. Biotechnol*. **60**: 177–182.
25. Johns, M.M., Marshall, W.E., and Toles, C.A. (1998). Agricultural by-products as granular activated carbons by adsorbing dissolved metals and organic. *J. Chem. Technol. Biotechnol*. **71**: 131–140.
26. Bansode, R.R., Losso, J.N., Marshall, W.E., Rao, R.M., and Portier, R.J. (2003). Adsorption of metal ions by pecan shell-based granular activated carbons. *Bioresour. Technol.* **89**(2): 115–119.
27. Chen, J.P. and Lin, M. (2001). Equilibrium and kinetics of metal ion adsorption onto a commercial H-type granular activated carbon: experimental and modeling studies. *Water Res.* **35**(10): 2385–2394.
28. Choi, J.Y. and Kim, D.S. (2002). Adsorption behavior of zinc and cadmium ion on granular activated carbon in singular and binary systems and the influence of nitrilotricetic acid as a complexing agent. *J. Environ. Sci. Health Part A Tox Hazard Subst. Environ. Eng.* **37**(9): 1701–1719.
29. Yu, Q. and Kaewsarn, P. (1999). Binary adsorption of copper(II) and cadmium(II) from aqueous solutions by biomass of marine alga Durvillaea Potatorum. *Separation Sci. Technol.* **34**(8): 1595–1605.
30. Trujullo, E.M., Jeffers, T.H., Ferguson, C., and Stevenson, Q. (1991). Mathematically modeling the removal of heavy metals from a wastewater using immobilized biomass. *Environ. Sci. Technol.* **25**: 1559–1664.
31. Bunzl, K., Schmidt, W., and Sansoni, B. (1976). Kinetic of ion exchange in soil organic metter-IV. Adsorption and desorption of Pb^{2+}, Cu^{2+}, Cd^{2+}, Zn^{2+}, Ca^{2+}. *J. Soil. Sci.* **27**: 32–41.
32. Masslenilov, B.I. and Kiselva, S.A. (1989). Physicochemical basis for the use of peat in ion-excahnge technology and adsorption processes. *Torfianaia Promyshlennost* **5**: 23–25.
33. Ho, Y.-S., Wase, D.A.J., and Forster, C.F. (1996). Removal of lead ions from aqueous solution using sphagnum moss peat as adsorbent. *Water SA* **22**: 219–224.
34. Beveridge, A. and Fickering, W.F. (1983). The influence of surfactants on the adsorption of heavy metal ions by clays. *Water Res.* **17**: 215–255.
35. Allen, S.J. and Brown, P.A. (1995). Isotherm analyses for single component and multicomponent metal sorption onto lignite. *J. Chem. Tech. Biotechnol*. **62**: 17–24.
36. Tan, T.C., Chia, C.K., and Teo, C.K. (1985). Uptake of metal ions by chemically treated human hairs. *Water Res.* **19**(2): 157–162.
37. Krishnan, K.A. and Anirudhan, T.S. (2003). Removal of Cd(II) from aqueous solutions by steam activated sulphurised carbon prepared from sugar bagasse-pitch: kinetic and equilibrium studies. *Water SA* **29**(2): 147–156.
38. Mohan, D. and Singh, K.P. (2002). Single and multicomponent adsorption of zinc and cadmium from wastewater using activated carbon derived from bagasse-an agricultural waste material. *Water Res.* **36**(9): 2304–2318.
39. Mohan, D. and Chander, S. (2001). Single component and multicomponent metal ions adsorption by activated carbons. *Colloids Surfaces A* **177**: 183–196.
40. Srivastara, S.K., Tyagi, R., and Pant, N. (1989). Adsorption of heavy metal ions on carbonaceous material developed from the waste slurry generated in local fertilizer plant. *Water Res.* **23**(9): 1161–1165.
41. Pant, N. (1988). *Studies on the Use of Activated Carbon Developed from a Fertilizer Waste for the Removal of Organic and Inorganic Pollutants*, Ph.D. Thesis, Indian Institute of Technology Roorkee, India.
42. Yen C-Y. and Singer, P.C. (1984). Competitive adsorption of phenols on activated carbon. *J. Environ. Engineer.* **119**(5): 976–989.
43. Sorial, G.A., Suidan, M.K., and Vidic, R.D. (1993). Competitive adsorption of phenols on GAC I: Adsorption equilibrium. *J. Environ. Engineer.* **119**(6): 1026–1043.
44. Sorial, G.A., Suidan, M.K., Vidic, R.D., and Maloney, S.W. (1993). Competitive adsorption of phenols on GAC I: Adsorption dynamics under anoxic conditions. *J. Environ. Engineer.* **119**(6): 1044–1058.
45. Garner, I.A., Watson-Craik, I.A., Kirkwood, R., and Senior, E. (2001). Dual solute adsorption of 2,4,6-trichlorophenol and N-[2-(2,4,6-trichlorophenoxy)propyl]amine onto activated carbon. *J. Chem. Tech. Biotechnol*. **76**: 932–940.
46. Srivastava, S.K. and Tyagi, R. (1993). Competitive effect of surfactants in the adsorption of phenols by activated carbon developed from fertilizer waste slurry. *Chem. Engineer. World* **28**(11): 83–86.
47. Jossens, L., Prausnitz, J.M., Fritz, W., Schlunder, E.U., and Myers, A.L. (1978). Thermodynamics of multicomponent adsorption from dilute aqueous solutions. *Chem. Engineer. Sci.* **33**: 1097–1106.

48. Barton, S.S. (1993). Adsorption from dilute, binary, aqueous solutions. *J. Colloid. Interface Sci.* **158**: 64–70.
49. García-Araya, J.F., Beltrán, F.J., Álvarez, P., and Masa, F.J. (2003). Activated carbon adsorption of some phenolic compounds present in agroindustrial wastewater. *Adsorption* **9**(2): 107–115.
50. Mameli, A., Cincotti, A., Lai, N., Crisafulli, C., Scire, S., and Cao, G. (2004). Adsorption of organic compounds onto activated carbons from recycled vegetables biomass. *Ann. Chim.* **94**(7–8): 547–554.
51. Kim, M-S. and Lordgooei, M. (2004). Modeling volatile organic compound sorption in activated carbon II. Multicomponent equilibrium. *J. Environ. Eng.* **130**(3): 223–230.
52. Kim, M-S. and Lodgooei, M. (2004). Modelling volatile organic compound sorption in activated carbon I: Dynamics and single component equilibrium. *J. Environ. Eng.* **130**(3): 212–222.
53. Semmens, M.J., Staples, A.B., Hohenstein, G., and Norgaard, G.E. (1986). Influence of coagulation on removal of organics by granular activated carbon. *J. Amer. Water Works Assoc.* **78**(8): 80–84.
54. Semmens, M.J., Norgaard, G.E., Hohenstein, G., and Staples, A.B. (1986). Influence of pH on the removal of organics by granular activated carbon. *J. Amer. Water Works Assoc.* **78**(5): 89–93.
55. Okazaki, M., Kage, H., Iijima, F., and Toei, R. (1981). Approximate description of multi solute adsorption equilibrium in organic aqueous solution. *J. Chem. Eng. Japan* **14**(1): 26–31.
56. Okazaki, M., Kage, H., and Toei, R. (1980). Prediction of liquid phase adsorption equilibria in muti-solutes in water. *J. Chem. Eng. Japan* **13**(4): 286–291.
57. Ha, S.R., Vinitnantharat, S., and Ishibashi, Y. (2001). A modeling approach to bioregeneration of granular activated carbon loaded with phenol and 2,4-dichlorophenol. *J. Environ. Sci. Health A Tox Hazard Subs. Environ. Eng.* **36**(3): 275–292.
58. Chern, J.M. and Chien, Y.W. (2003). Competitive adsorption of benzoic acid and p-nitrophenol onto activated carbon: isotherm and breakthrough curves. *Water Res.* **37**(10): 2347–2356.
59. Bulloch, J.L., Hand, D.W., and Crittenden, J.C. (1998). A model for predicting contaminant removal by adsorption within the International Space Station water processor: 1. Multicomponent equilibrium modeling. *Water Environ. Res.* **70**(1): 14–26.
60. Rozada, F., Calvo, L.F., Garcia, A.I., Martin-Villacorta, J., and Otero, M. (2003). Dye adsorption by sewage sludge-based activated carbons in batch and fixed-bed systems. *Bioresour. Technol.* **87**(3): 221–230.
61. Srivastava, S.K., Tyagi, R., Pal, N., and Mohan, D. (1997). Process development for the removal and recovery of substituted phenol from wastewater by a carbonaceous adsorbent developed from fertilizer waste material. *J. Environ. Engineer (ASCE)* **123**(9): 842–851.
62. Gupta, V.K., Srivastava, S.K., and Mohan, D. (1997). Equilibrium uptake, sorption dynamics, process optimization & column operation for the removal & recovery of malachine green from wastewater using activated carbon & activated slag. *Industr. Engineer. Chemistry Res. (ACS)* **36**(6): 2207–2221.
63. Sung-Ryong, H.A. and Vinitnantharat, S. (2000). Competitive removal of phenol and 2,4 dchlorophenol in biological activated carbon system. *Environ. Technol.* **12**(4): 387–396.
64. Harrison, I.M., Meadows, J., Robb, I.D., and Williams, P.A. (1995). Competitive adsorption of polymers and surfactants at the solid/liqud interface. *J. Chem. Soc. Farady Trans.* **91**(21): 3919–3923.
65. Srivastava, S.K. and Tyagi, R. (1995). Competitive adsorption of substituted phenols by activated carbon developed from fertilizer waste slurry. *Water Res.* **29**(2): 483–488.
66. Mckay, G. (1995). *Use of Adsorbents for the Removal of Pollutants from Wastewaters*. CRC Press, Boca Raton, FL.
67. Mackay, G. and Al-Duri, B. (1991). Multicomponent dye adsorption onto carbon using a solid diffusion mass transfer model. *Indust. Engineer. Chem. Research* **30**: 385.
68. Liapis, A.I. and Rippin, D.W.T. (1977). A general model for the simulation of multicomponent adsorption from a finite bath. *Chem. Eng. Sci.* **32**: 619.
69. Mijangos, F., Navarro, A., and Jodra, Y. (2001). Kinetic analysis of phenol adsorption from aqueous systems. *Can. J. Chem. Engineer.* **79**.
70. Cheung, C.W., Ko, D.C.K., Porter, J.F., and Mckay, G. (2003). Binary metal sorption on bone char mass transport model using IAST. *Langmuir* **19**: 4144–4153.

A REAL-TIME HYDROLOGICAL INFORMATION SYSTEM FOR CITIES

M. Eng Nguyen Quang Hung
Sutat Weesakul
Asian Institute of Technology
Pathumthani, Thailand

Uruya Weesakul
Chavalit Chaliraktrakul
Thammasat University
Pathumthani, Thailand

Ole Mark
Lars Chr Larsen
DHI Water and Environment
Hørsholm, Denmark

Nowadays, many organizations collect hydrologic information for various purposes. But still, little information from the huge sources of data is made public in real time, and only a small fraction of the data is applied for real-time decision making. Once the hydrologic data have become "historical" (the data are no longer applicable for real-time decision making), the data are still very valuable for design and for evaluating and understanding the hydrologic environment. But if hydrological data are used only after they have become "history," the value of the data collection is not fully used.

Before starting to disseminate hydrologic information it is important to address the following questions:

1. What kind of real-time information should be shown to the public?
2. How and where should the real-time information be presented to the public?
3. When do hydrologic data change from valuable real-time information to less interesting (from a public point of view) historical hydrologic data?
4. What does it take in knowledge, technology, and hardware to provide real-time hydrologic information?

This article discusses experience gathered from a research project in Bangkok concerning provision of real-time rainfall data to the public through the Internet, handheld computers, and mobile phones. It also provides an outline of the future use of hydrologic information in real time. The potential benefits from the described framework for using the information are

- a public rainfall and flood information service, like the daily weather forecast, and flood warning system
- traffic information about streets that have a potential risk for flooding,
- a decision support system for reducing flooding in the Bangkok area.

INTRODUCTION

Water is a basic necessity for sustaining life and developing society. Proper management, protection and development of water resources are challenges imposed by population growth, increasing pressure on water and land resources by competing usage, and degradation of scarce water resources in many parts of the world.

The cities in Asia are growing rapidly, so it is important to pay attention to the role of cities in integrated water resources management (IWRM). The infrastructure for water in cities consists of water supply networks, sewer systems, and purification and wastewater facilities. Many of the cities are old, and they have developed according to varying historical needs and visions. Hence, the layout and design of the infrastructure has gradually developed into rather complex systems covering large areas, and the systems often have inadequate capacities. The cities are located in the monsoon area, so they have to cope with very high rainfalls every year. Regular flooding during the rainy season is the rule rather than the exception in most areas. For smaller floods, the people seem to have adjusted their daily lives, although there is always a loss of income and property damage. However, when the floods are heavy, the socioeconomic and health impacts on the population are enormous. This requires information and tools that can handle such situations and describe flooding from local rainfall in cities.

An understanding of the physical system and its interaction with the environment is a prerequisite for effective planning and management of urban water resources. It is too expensive to eliminate flooding in urban areas by using only structural measures, such as building new sewers, embankments, and installation of pumps. To manage the existing infrastructure better, nonstructural measures, like real-time rain fall and flood risk information can minimize flooding and the impact of flooding. In addition, provision of flood risk information can help people make more intelligent decisions during floods, and hence reduce the stress and costs of floods.

Today's advances in computer technology can help many cities in the world manage local and minor flooding problems using computer-based solutions. This involves building computer models of the drainage/sewer system. These models are then used to understand the often rather complex interaction between rainfall and local flooding. Computer models provide the opportunity for well-structured analyses of rainfall/runoff/flooding, water availability, water demands, and wastewater disposal, and they offer a sound scientific framework for coordinated management and planning. Once the existing conditions have been analyzed and understood, alleviation schemes can be evaluated and the optimal scheme implemented. In addition to real-time information, modeling increases the value of the hydrologic data tremendously because the combination of models and real-time data provides a full cover of information over a catchment.

BACKGROUND

At present, the Asian Institute of Technology (AIT) is carrying out a research project to apply real-time rainfall information for Bangkok, Thailand. The Greater Bangkok area has a very high level of activity. Millions of people live in the area: On the outskirts of the city, there are a lot of big factories, and many national as well as international offices are situated in the center of town. When heavy rainfall occurs in the Bangkok area, some of the consequences are heavy traffic jams, waterlogging, blackouts, and property damage. If flooding follows, there is often great social impact as well. Many people might lose their jobs for a shorter or longer period due to temporary closing of businesses, and schools have to close because the streets are flooded.

A typical picture of flooding in the city after a minor heavy rain in Bangkok can be seen in Fig. 1. Knowing the condition of rainfall in Bangkok in advance can help in managing and dealing with these problems. Therefore, hydrometeorological forecasts and warnings are effective tools for preventing property damage caused by rainfall and subsequent flooding.

The objectives of the current project are to

1. provide a real-time information system concerning rainfall and flood risk, and

Figure 1. The Nontaburi area in Bangkok after a “minor” heavy rain on the April 27, 2002.

2. explore the possibilities of predicting rainfall and flooding in Bangkok.

The outcome of the project in economic and social impacts depends on the dissemination of the information to the people and the accuracy of the information supplied. The information system is designed to be generic, so that the concepts can be transferred easily and implemented at other Asian locations.

PROVISION OF REAL-TIME HYDROLOGICAL INFORMATION TO THE PUBLIC

Before embarking on a project to provide real-time hydrologic information to the public, it is important to assess which information the public would like to have as individuals and which information local authorities would like to have to manage the assets of the city. For example, a private person driving a car to work may be interested in knowing the approximate duration of the rain and the strength of the rain plus an assessment of the risk of flooding like "Probably small flooding" or "Flooding—Don't drive to zone A, B, etc.," whereas the local authority may want information about an accumulated rainfall of 60 mm during the past 60 minutes (which, by experience, provokes flooding) or other more specific numbers concerning the rainfall. The local authority may then want to process the rainfall information further and pass it to public information sources such as traffic radio and local news broadcasts.

In modern society, much information (maybe too much) surrounds us, and we must choose which information we would like to have. However, one of the basic principles is that information should be available at the time we need it and in a form that fits our demands. Providing rainfall information through a computer requires that we are in front of a computer, which rarely happens when we are walking or driving in the rain. We could stop and visit an Internet café to get the latest rainfall information, but that would be rather unrealistic. You would like rainfall information to be available wherever you are. Mobile phones provide the means of bringing the rainfall information directly to you wherever you are.

Technical Requirements to Transform Hydrologic Data to Real-Time Information

A prerequisite for turning hydrologic data into real-time hydrologic information is the basic infrastructure in terms of rain gauges and stable Internet connections. In addition, the communication between the rain gauges and the Internet service should be available at all times. If these components are available, it is possible to send the hydrologic data to a central place where the data can be processed into information. There are several ways of setting up such a communication system; it may be tailored depending on individual need and available hardware. Figure 2 shows the configuration of the real-time application that is under development at AIT.

Data from rain gauge stations and radar stations were sent to the center by radio and the Internet. After generating the data, Bangkok Metropolitan Administration (BMA) and Thai Meteorological Department (TMD) center sent data packages to AIT by the Internet. The server at AIT was set up to carry out all the main tasks: receiving, generating, and storing online data; creating applications based on data; and answering requests from users. The data are updated using a real-time Internet database (DIMS^{TM}). In DIMS^{TM}, macros have been developed to create the images and other information automatically for the web sites and then upload the results to the web sites, including WAP sites available to the public (Fig. 4). The forecasts of rainfall and flooding are at present based on 53 on-line rain gauges across Bangkok (Fig. 3) and a weather radar located in downtown Bangkok.

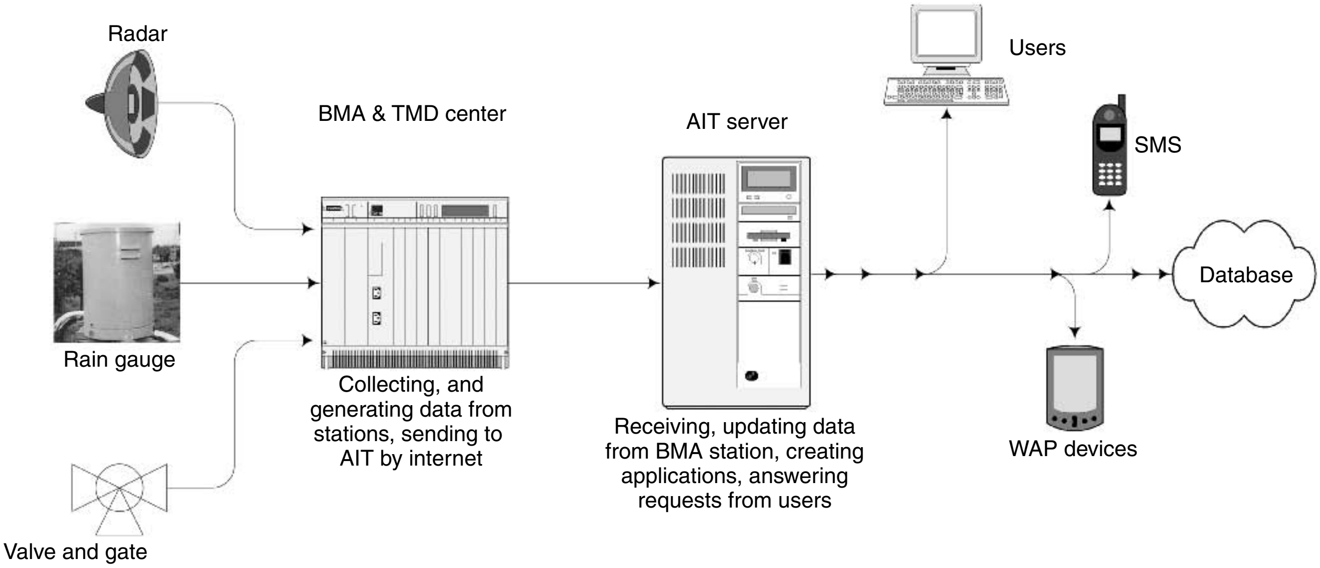

Figure 2. The layout of the real-time rainfall information application at AIT.

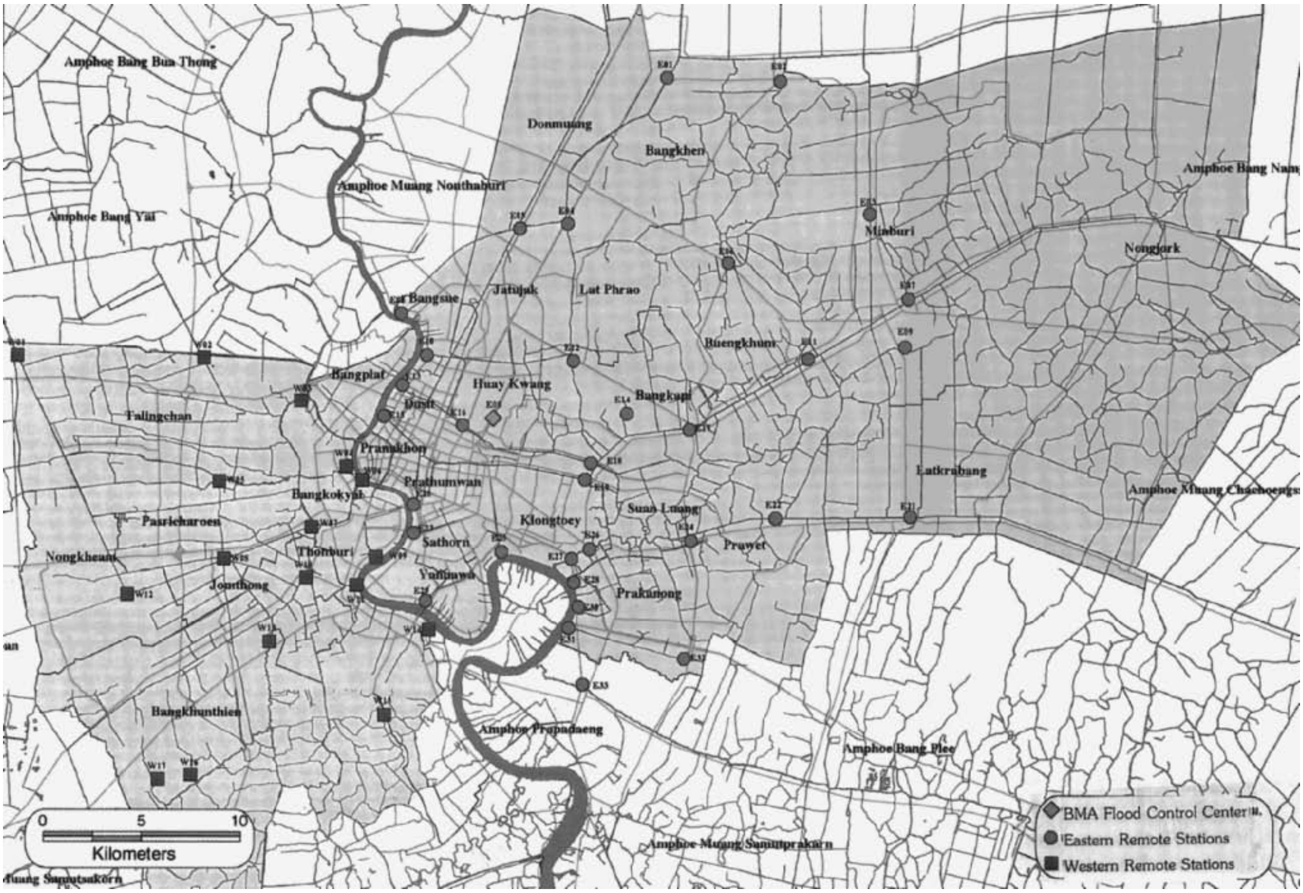

Figure 3. The locations of 47 Bangkok Metropolitan Administration rain gauges in Bangkok (source BMA).

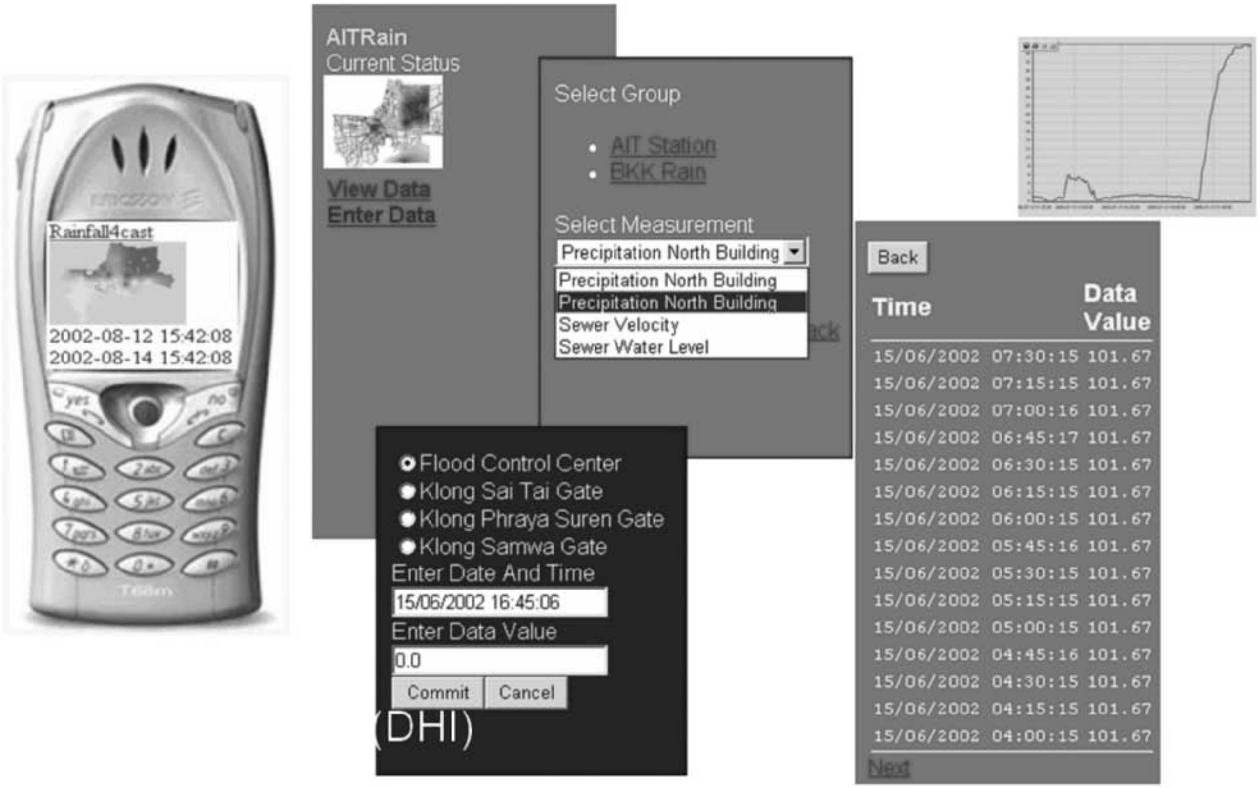

2002-08-12 15:42:08
2002-08-14 15:42:08

- ● Flood Control Center
- ● Klong Sai Tai Gate
- ● Klong Phraya Suren Gate
- ● Klong Samwa Gate

Enter Date And Time: 15/06/2002 16:45:06

Enter Data Value: 0.0

Time	Data Value
15/06/2002 07:30:15	101.67
15/06/2002 07:15:15	101.67
15/06/2002 07:00:16	101.67
15/06/2002 06:45:17	101.67
15/06/2002 06:30:15	101.67
15/06/2002 06:15:15	101.67
15/06/2002 06:00:15	101.67
15/06/2002 05:45:16	101.67
15/06/2002 05:30:15	101.67
15/06/2002 05:15:15	101.67
15/06/2002 05:00:15	101.67
15/06/2002 04:45:16	101.67
15/06/2002 04:30:15	101.67
15/06/2002 04:15:15	101.67
15/06/2002 04:00:15	101.67

Figure 4. Information about the rainfall over Bangkok as presented on a mobile phone.

How Close to Real Time must Hydrologic Information be to Consider It Valuable Real-Time Information?

This is a question, which is extremely difficult to answer. So far, this article has used the term "real time" to describe the measurements, which were made, sent to the AIT computers, processed, and presented to the public. But the information presented to the public will always have a time delay because of the time it takes for measurements, data transmission, and processing. These time delays are discussed in detail below:

1. The rain gauges run by Bangkok Metropolitan Administration (BMA) measure the rainfall at 15 minute intervals. Thus, it may only rain the first 5 minutes of the sampling interval and the rainfall intensities from the first 5 minutes will be averaged over a period of 15 minutes. Hence, the sample over 15 minutes produces a significantly lower peak intensity, which according to the measurement, lasts 10 minutes longer than the actual rain.
2. The data transmission from the rain gauges to BMA and further to AIT is rather swift and takes only 1 to 2 minutes. But the AIT rain data server looks only for new data from BMA every 5 min, which on average means a time delay of 2.5 minutes.
3. The data processing at AIT and the transmission to the Internet, mobile phones, etc. takes only another minute, but then the hydrologic information stays on the Internet until new information has been measured, transmitted and processed through the loop described before.

The "real-time" information available to the public is updated only after the repetition of a new measurement cycle, as described above. Hence, the information presented as real-time information may in the worst case be around 20–30 minutes old!!! Is this acceptable? This is a question that is difficult to answer and depends on may local conditions such as the speed of the rain front, the uniformity of the rain; the density and number of rain gauges, and the use of the real-time hydrologic rainfall information. For example, if a rain cloud moves at 5 m/s, then it will travel 6 km in 20 minutes. This means that details within an area of travel distance will be difficult to describe accurately. Means to achieve a higher level of accuracy are to reduce the sampling interval from 15 minutes to 5 minutes or even to 1 minute to be able to present information to the public closer to real time. Alternatively, more rain gauges can be installed. One can imagine that the public will not gain faith in an information system, which tells them it is still raining where they are 20 min. after they have felt the last raindrop.

APPLICATION OF REAL-TIME HYDROLOGIC INFORMATION IN CONJUNCTION WITH URBAN DRAINAGE MODELS

Applying a model in conjunction with real-time rain data provides means for improving information about the current status of the hydrologic system. In general, the rain gauges provide information only at specific locations. The application of a model provides the possibility for gap-filling of information between rain gauges. The simplest kind of modeling is an inverse-distance-based interpolation between the rainfall measurements in real time. At present, that is the method applied at AIT for generating hyetographs. In the near future, more interpolative methods will be evaluated and compared.

In addition to interpolation of the rain intensity between the spatial network of rain gauges, surface runoff and urban drainage models can easily be added to the hydrologic information system. The models are automatically executed when the rainfall at specified locations exceeds preset threshold values. Based on a forecast by the hydrologic model or the urban drainage model, the level of information about the hydrologic system can be extended into the future. The general approach behind the level of information about the hydrologic system as a function of the availability of rain gauges, models, and a forecast is shown in Fig. 5. The figure shows how the information level increases and extends in time when more rain gauges and a model are applied.

The present status of the research project in Bangkok, is that rain gauges reporting in real time together with a simple model predict the spatial variation of the rainfall. In addition, forecast tests in real time have been carried out on a small scale. First, the procedures were tested with a small urban drainage model covering AIT only where all the real-time information is measured directly at AIT. Second, the urban flood model for Bangkok was set up and made operational based on real-time data from BMA and TMD. The forecast has shown promising results and will be made operational during November–December 2003.

The ongoing work on flood forecasting for specific areas of Bangkok is based on real-time rainfall data and a deterministic hydrodynamic urban flood model. The urban flood model has already been established during 2001–2002 as part of the AIT research in modeling urban flooding (1,2). The urban flood model builds on a 1-D hydrodynamic urban drainage modelling package, MOUSE (3) from DHI—Water & Environment. Based on input in terms of time series of rainfall, the MOUSE model for Bangkok produces flood inundation maps, showing flooded areas and depths for the Sukhumvit business area in Bangkok. Layout of the system is shown in Fig. 6. The model has been successfully set up and applied to off-line studies of flooding in Bangkok.

WARNING SYSTEMS FOR FLOODING, FLASH FLOODS, AND LANDSLIDES

Real-time hydrologic information makes it possible to detect potential hazards shortly before they happen, for example, heavy rainfall recorded in certain parts of the catchment may indicate potential flooding, a potential flash flood, or a landslide in hilly areas. However, the level of data and information from a real-time hydrologic system is overwhelming, and it is not really suitable for decision support until targeted data processing has taken place. By keeping the rainfall information in a database, it is easy to write scripts, which automatically keeps track of

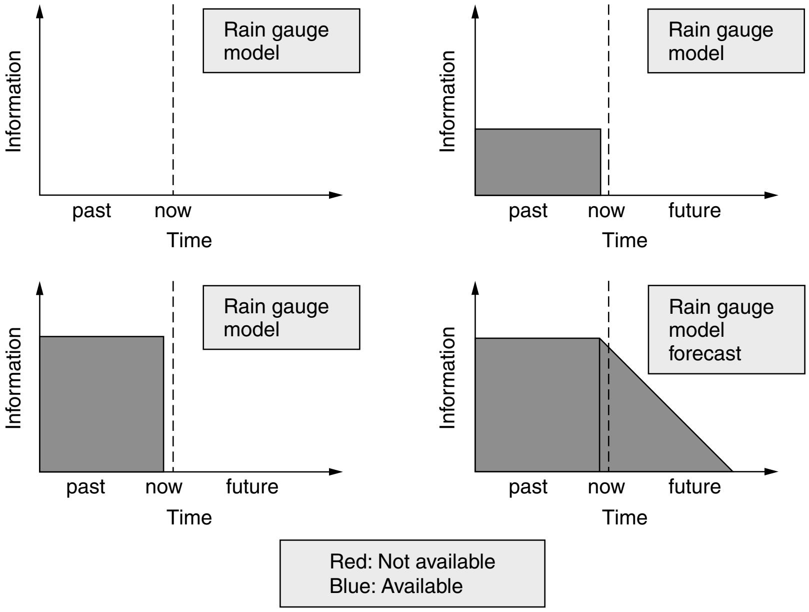

Figure 5. The level of information as a function of the application of (1) no rain gauges and no model available; (2) only rain gauges available; (3) rain gauges and a model available; and (4) rain gauges, a model, and a forecast available.

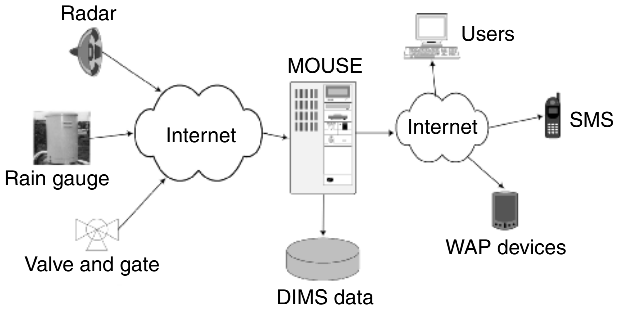

Figure 6. The layout of the real-time data transmission and real-time model.

the accumulated rainfall within the last hour. In Bangkok, experience says that if the accumulated rain within 1 hour is more than 60 mm, then flooding occurs. This criterion has been implemented in the rainfall information system, and warnings are automatically sent to BMA officers, whenever this criterion is exceeded. The warning system is built so that it both sends information about the rain and a personalized message in an e-mail and a SMS message to mobile phones when the criterion is exceeded. Whoever the BMA officer on duty is, he will be alerted about the rainfall conditions either by his email system or by a ringing phone. It is possible to let the information system automatically control gates, weirs, pumps, or other devices in the drainage system, but that is not considered for the moment.

APPLICATION OF WEATHER RADAR FOR HYDROLOGIC REAL-TIME INFORMATION

During the last decade, the use of weather radar has emerged as a tool for rainfall forecasting and at present, research on the use of weather radar for estimating rainfall is carried out at several locations (4–6). A weather radar measures the reflectivity from raindrops in the sky. It does not measure the actual rainfall, and a calibration between the reflectivity and the rainfall is always required for each individual weather radar. A radar provides the potential for getting the spatial variation of rainfall over a large catchment area. For example, the weather radar location in Bang Na, Bangkok, has a range of 60 km, and radar pictures from the Bang Na radar are presently updated every 15 minutes. However, weather radars are still a rather new topic for research in forecasting rainfall. Until now, it has not really been possible systematically to forecast rainfall dynamics (convective effects) with notable success, and operational application of weather radar is rare. However, even though the radar information at present is not highly accurate, radar can still be part of a warning system, where a human is automatically alerted (through the e-mail and SMS warning system described above) when very high rainfall intensities are measured by the weather radar. It will then be up to the duty officer to assess the flood/flash flood risk based on local data, guidelines, and experience. At present, a number of weather radars are already available in Thailand, so there is a good basis for making a warning system that covers the country.

FUTURE PERSPECTIVES FOR HYDROLOGIC DATA IN REAL TIME

At present, hydrologists working in urban areas are facing many new challenges imposed by the ever changing hydrologic environment in cities. Emphasis should be put on managing the urban systems as well as possible by applying currently available information and technology, for example, by implementing real-time hydrologic information systems, like that outlined above. Apart from managing urban hydrology in real time, many other challenges have to be addressed in the near future. An example is Dhaka city, which relies heavily (up to 97%) on groundwater for its water supply (7). During the last 25 years, the groundwater table has dropped by about 25 m. If this drop in the water table continues, it may generate problems for the city's water supply, and surface waters may be needed as additional resources. This is complicated by the growth of Dhaka city. Areas that used to be permeable are being transformed into hard, impermeable surfaces. Such impermeable surfaces prevent replenishing groundwater storage and further aggravate groundwater problems. In turn, the runoff from the new impermeable surfaces generates additional surface runoff, which again increases the flooding in Dhaka city. If the water supply pipes are under low pressure during a period of flooding, polluted floodwater may enter the water supply network. This poses an additional health risk to the population on top of the diseases spread by the floodwater. In addition, there are considerable losses from the water supply network, so quite a large amount of drinking water is lost. This means that being an urban hydrologist does not only involve one problem, but a group of strongly interrelated problems, that have major impacts on the people living in the cities.

The application of real-time hydrologic information is not the solution to all problems mentioned above, but it may mitigate flood problems and in this way reduce the economic losses and health problems the public faces from urban flooding. Finally, the information generated by a real-time hydrologic information system can be applied by using the historical data for design and maintenance analyses to achieve better functionality of the urban hydrologic system before the next heavy rain arrives.

SUMMARY

Making hydrological information publicly available on the Internet and mobile phone is at present a feasible task, which makes data collection more valuable and the work of hydrologists more highly appreciated by the public. In this project, the essential backbone for forecasting rain and urban flooding is established, and it can easily be extended to achieve more accurate forecasts, by adding online rain gauges to the system or by changing the sampling time for the rain gauges to a shorter period. In addition, the hydrologic information can be applied in conjunction with real-time hydrologic and urban drainage models providing decision support and warning systems to deal with urban flooding and flash floods. The authors believe that hydrologic information systems like that described in this article, in the future, may be part of the infrastructure of any major city around the world.

Acknowledgments
The authors sincerely thank the Danish International Development Assistance (DANIDA) for financial support of the current study at the Asian Institute of Technology and the Bangkok Metropolitan Administration and Thai Meteorological Department for providing data.

BIBLIOGRAPHY

1. Mark, O., Apirumanekul, C., Kamal, M., and Praydal, P. (2001). *Modelling of Urban Flooding in Dhaka City*, *UDM*, Orlando, FL, May 20–24.
2. Boonya-Aroonnet, S., Weesakul, S., and Mark, O. (2002). *Modelling of Urban Flooding in Bangkok*. *9th Int. Conf. Urban Drainage*, Portland, OR, September 8–13.
3. Lindberg, S., Nielsen, J.B., and Carr, R. (1989). *An Integrated PC-Modelling System for Hydraulic Analysis of Drainage Systems*. *1 Aust. Conf. Tech. Comp. Water Ind.: Watercomp '89*, Melbourne, Australia.
4. Braga, B. and Massambani, O. (1997). *Weather Radar Technology for Water Resources Management*. UNESCO, Unesco Press, Montevideo, Uruguay, p. 516.
5. Einfalt, T., Krejci, V., and Schilling, W. (1998). Rainfall data in urban hydrology. In: *Hydroinformatics Tools for Planning, Design, Operation and Rehabilitation of Sewer Systems*. J. Marsalek, C. Maksimovics, E. Zeman, and R. Price (Eds.). *NATO ASI Series 2: Environment*, Vol. 44. Kluwer, Dordrecht, pp. 129–168.
6. Collier, C.G. (1989). *Application of Weather Radar Systems: A Guide to Uses of Radar Data in Meteorology and Hydrology*. Ellis Harwood, Chichester, England.
7. Khan, H.R. and Siddique, Q.I. (1999). Urban water management problems in developing countries with particular reference to Bangladesh. *Int. J. Water Res. Dev.*, special issue.

READING LIST

Mark, O., Boonya-Aroonnet, S., Hung, N.Q., Buranautama, V., Weesakul, U., Chaliraktraku, C., and Larsen, L.C. (2002). *A Real-Time Hydrological Information System for Bangkok*. *Int. Conf. Urban Hydrol. 21st Century*. Kuala Lumpur, Malaysia, October 14–18.

Mark, O., Weesakul, S., and Hung, N.Q. (2002). *Modelling the Interaction Between Drainage System, Wastewater Treatment Plant and Receiver Water in Pattaya Beach*. *9th Int. Conf. Urban Drainage*, Portland, OR, September 8–13.

Mark, O., Hung, N.Q., and Helwigh, A.B. (2002). Waiting for a swim—the handling of wastewater in Asia. *WaterLines* **20**(4).

Moraes, L.R.S. (1996). *Health Impacts of Drainage and Sewerage in Poor Urban Areas in Salvador, Brazil*. Ph.D. Thesis. London School of Hygiene and Tropical Medicine, London.

CHLORINE AND CHLORINE RESIDUALS

S. Rajagopal
G. van der Velde
Radboud University Nijmegen
Nijmegen, The Netherlands

V.P. Venugopalan
BARC Facilities
Kalpakkam, India

H.A. Jenner
KEMA Power Generation and Sustainables
Arnhem, The Netherlands

INTRODUCTION

Chlorine, being a powerful oxidant, is capable of destroying biological molecules (1). It is universally used as a disinfectant for drinking water treatment and as a biocide for cooling water treatment (2). However, the very oxidizing nature of chlorine involves it in a number of side reactions with organic and inorganic substances present in water. A certain fraction of the chlorine actually dosed into water may therefore disappear and not be available for estimation by the normal methods employed. These methods actually estimate chlorine concentration by measuring the amount of oxidants present. The fraction of chlorine that is used in such side reactions is called chlorine demand. Chlorine residual refers to the amount of chlorine (as well as its reaction products, which retain some oxidation potential) that remains after such side reactions.

Chlorine, when dissolved in natural waters, gives rise to various oxidizing compounds depending on the reaction of hydrolysis and oxidation of ammonia, which leads to the production of free chlorine (as hypochlorous acid or its dissociated form) as well as various chloramines, all of which retain oxidant property. The oxidants also react with organic matter to produce halogenated organics. Therefore, chemistry of water chlorination is complex and involves many molecular and ionic species, with often confusing terminology. In literature, chlorine dissolved in water may be described as “free,” “active,” “available,” “combined,” or “residual”—or a combination of the above. A brief explanation of the nomenclature associated with water chlorination is given below.

Free Chlorine/Free Available Chlorine (FC/FAC)

This refers to the amount of hypochlorous acid present in water in its dissociated or undissociated form:

$$Cl_2 + H_2O \longrightarrow HOCl + HCl \tag{1}$$

$$HOCl \rightleftharpoons H^+ + OCl^- \tag{2}$$

In seawater (which contains about 65 mg/L bromide), the following reactions will also take place:

$$HOCl + Br^- \longrightarrow HOBr + Cl^- \tag{3}$$

Therefore, in seawater chlorination, HOBr (as well as hypobromite ion, OBr^-) is also categorized into FC/FAC.

Combined Chlorine/Combined Available Chlorine

Combined chlorine refers to chloramines, which result from the reaction of free chlorine with ammonia (or other nitrogenous compounds with an N-C link) present in water:

$$HOCl + NH_3 \longrightarrow NH_2Cl + H_2O \tag{4}$$

$$NH_2Cl + HOCl \longrightarrow NHCl_2 + H_2O \tag{5}$$

Additionally, the following reactions are also possible in seawater environments:

$$NH_3 + HOBr \longrightarrow NH_2Br + H_2O \tag{6}$$

$$NH_2Br + HOBr \longrightarrow NHBr_2 + H_2O \tag{7}$$

$$2NH_2Cl + Br^- \longrightarrow NHBrCl + NH_3 + Cl^- \tag{8}$$

Chloramines are much less reactive (that is, less effective as biocides) when compared with bromamines. Combined forms of chlorine are, in general, less efficient biocides than free chlorine, but are more persistent than it. Hence, they are very important from the environmental point of view.

Total Available Chlorine (TAC)

This refers to the sum of the two terms just defined and represents a major part of the biocidal capacity of the chlorinated water.

Residual Chlorine

This term is analogous to TAC and is often used to represent the oxydisinfectant capacity of water (consisting of free and combined oxidants) at that point in time (Fig. 1). It must be kept in mind that this capacity goes on reducing as a function of time because of what is known as chlorine decay. The dosed chlorine continuously engages in a series of reactions with substances present in water, which in due time will result in complete disappearance of all measurable chlorine.

Chlorine Demand (CD)

It is defined as the difference between the amount of chlorine added and the useful residual chlorine that remains at the end of a specified contact time. It refers to the amount of chlorine “lost” in side reactions referred to earlier. As the CD of a given sample of water varies with the chlorine dose applied and the contact time (that is, the interval between chlorine dosing and chlorine measurement), it is always denoted with reference to the dose and the contact time. For example, for a chlorine dose of 1 mg/L, a sample of water may return a CD value of 0.6 mg/L after a contact time of 30 minutes. In natural sea or fresh water, chlorine demand also varies as a function of season (Fig. 2) because of the seasonal changes in the concentration of oxidizable substances present in natural waters. The dosed chlorine continuously engages in a series of reactions with substances present in water, which in due time will result in complete disappearance of all measurable chlorine.

Total Residual Chlorine/Total Residual Oxidant (TRC/TRO)

Chlorine estimation in water is carried out indirectly by measuring the oxidant capacity via the stoichiometric iodide/iodine. As the method determines all oxidants produced in Equations 1–7, it is appropriate to use the term total residual oxidants (TRO) rather than total residual chlorine (TRC) in the case of seawater chlorination (Fig. 1). Although different in nomenclature, TRO numerically equals TRC.

Chlorine Produced Oxidants (CPO)

This term is used to distinguish oxidants produced in water after addition of chlorine from those naturally present in a given parcel of water. The latter appear as “background” in blank determination, but because their normal concentrations are very low, they can be generally ignored.

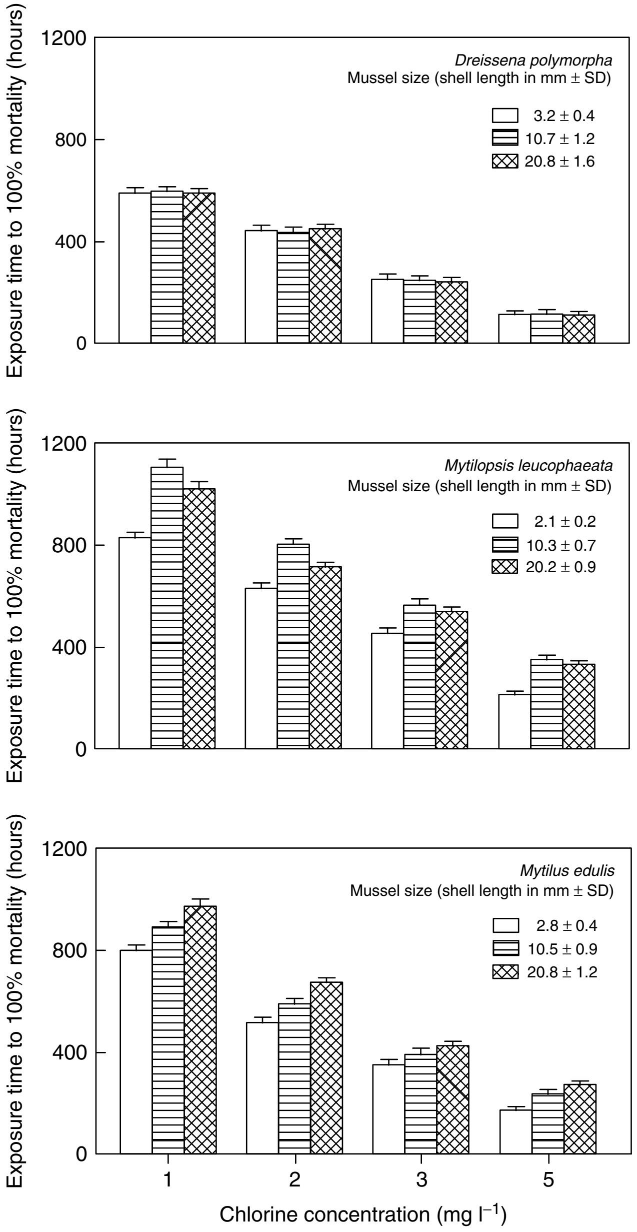

Figure 1. (**a**) Example of a chlorination curve in freshwater containing ammonia [modified after (2)]; (**b**) Experimental chlorination of seawater from intake canals of Gravelines power station. N ammonia = 0.21 mg/L; Contact time = 20 min [modified after (2,3)].

Nonoxidizing Chlorine Byproducts (CBP)

These are byproducts of chlorination that result from the reaction of chlorine with organic matter present (especially humic substances) in water. Several of these compounds exist, but the majority of them come under the category organohalogens called trihalomethanes (THM) (Fig. 3).

From the above discussion, it is clear that the most important term that one comes across in the context of chlorination for disinfection/biofouling control is “chlorine residual,” which denotes the biocidal capacity of water in terms of its oxidizing nature. Both HOCl and OCl$^-$ (hypochlorite ion), produced in Reactions 1 and 2 above, are oxidants and constitute free residuals. However, HOCl is more effective than its ionized form, because being neutrally charged, it can penetrate cells more easily than OCl$^-$ (Fig. 4). The dissociation of HOCl is a pH- and temperature-dependent reaction, higher pH and temperature favoring ionization of the acid. The ionization also increases with increasing total

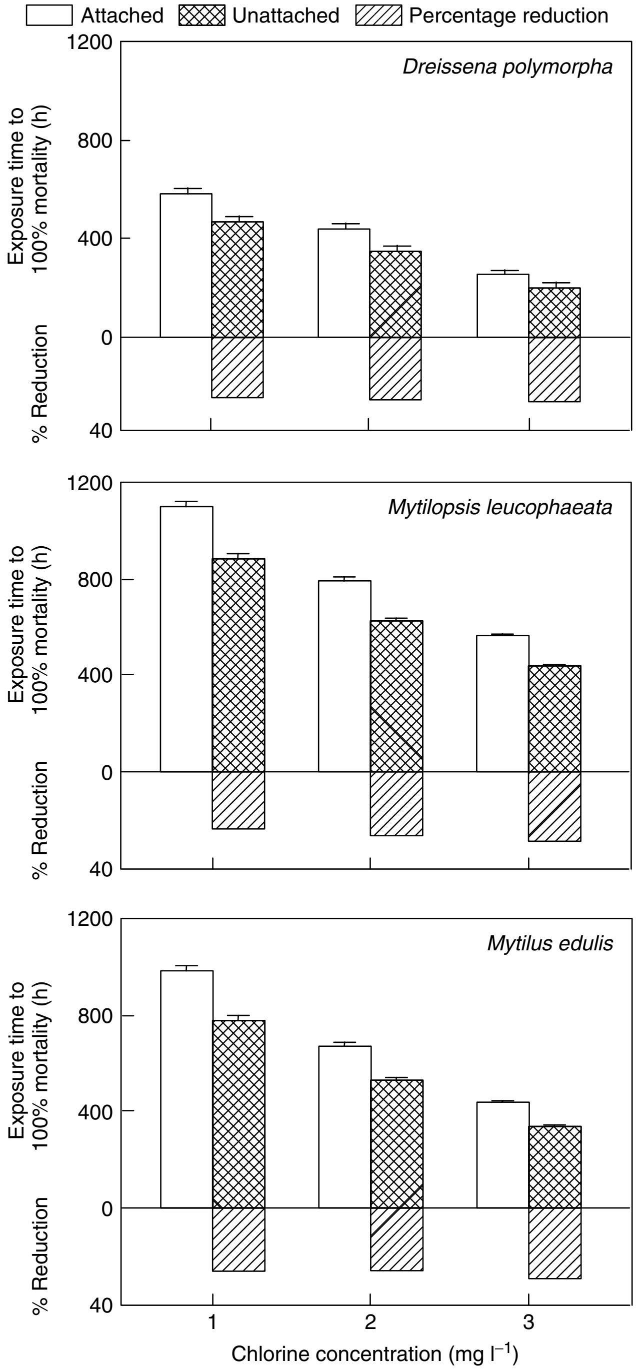

Figure 2. Seasonal variations of chlorine demand at three different chlorine doses in Kalpakkam coastal waters, east-coast of India from February 1989 to January 1990 [modified after (4)]. Total chlorine residuals were measured after 5 min for calculation of the chlorine demand.

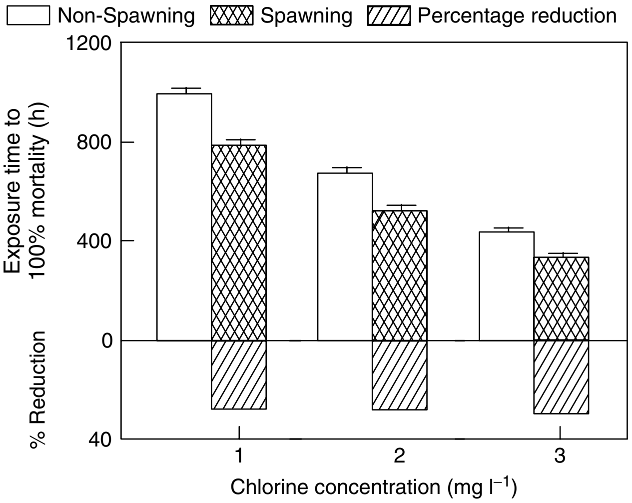

Figure 3. Average distribution of organohalogens measured by gas chromatography in seawater (chlorine dose = 0.8 mg/L) at the Gravelines power station [modified after (2,3)].

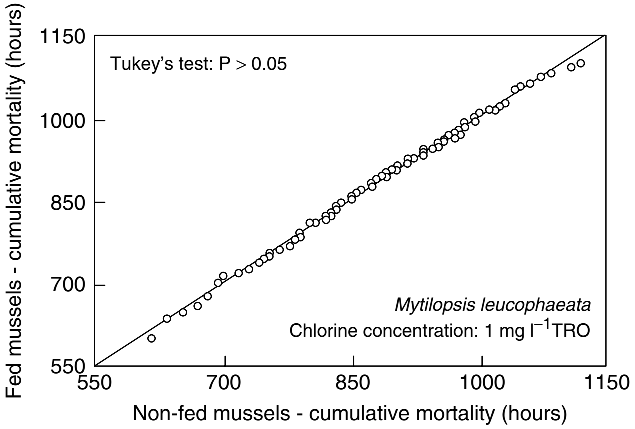

Figure 4. Comparison of the dissociation of hypochlorous acid and hypobromous acid with changing pH at a temperature of 20 °C [modified after (2)].

dissolved salts (TDS) or salinity, in the case of estuarine brackish water.

MEASUREMENT OF CHLORINE RESIDUALS

Several techniques are available for the measurement of chlorine residuals in water. Amperometric titration, potentiometric titration, or colorimetric titration using FAS-DPD (ferrous ammonium sulphate/N,N-diethyl-p-phenylene diamine) can be used. Simple colorimetry (using DPD) can also be used, especially under field sampling conditions, which has the ability to differentiate between free and combined forms of chlorine. Additionally, it allows rapid analysis after sample collection, reducing the chances of chorine loss through decay reactions. The required reagents are currently available commercially as portable kits, ready for field use. Although this method can theoretically detect a chlorine concentration of 0.01 mg/L, threshold detection level in practice is about 0.02 mg/L (2). Details of the various methods are available in APHA-AWWA-WPCF (5). Accurate and reliable instruments for continuous monitoring of residual chlorine levels in water are currently not available.

ENVIRONMENTAL DISCHARGE LIMITS

Several reports are available on the effects of residual chlorine in aquatic systems (6–9). Therefore, environmental release of chlorine residuals (for instance, in cooling water discharge from electric power plants) is tightly controlled in many countries through legislation. However, the permitted levels vary from country to country. For example, in the Netherlands, the permitted limit of chlorine residuals (0.2 mg/L for 2 h per day) is much less than that in India (0.5 mg/L).

BIBLIOGRAPHY

1. White, G.C. (1999). *Handbook of Chlorination and Alternative Disinfectants*. John Wiley & Sons, New York, p. 1569.
2. Jenner, H.A., Whitehouse, J.W., Taylor, C.J.L., and Khalanski, M. (1998). Cooling water management in European power stations: biology and control. *Hydroecologie Appliquée* **1–2**: 1–225.
3. Allonier, A.S. and Khalanski, M. (1997). *Chlorination By-Products in Sea Water*. EDF DER Report HE/31/97/027, Electricité de France, Chatou, Paris.
4. Rajagopal, S. (1997). *The Ecology of Tropical Marine Mussels and Their Control in Industrial Cooling Water Systems*, Ph.D. Thesis, University of Nijmegen, the Netherlands, p. 184.
5. APHA-AWWA-WPCF. (1992). *Standard Methods for the Examination of Water and Wastewater*, 17th Edn. Washington, DC.
6. Brook, A.J. and Baker, A.L. (1972). Chlorination at power plants: impact on phytoplankton productivity. *Science* **176**: 1414–1415.
7. Carpenter, E.J., Peck, B.B., and Anderson, S.J. (1972). Cooling water chlorination and productivity of entrained phytoplankton. *Mar Biol.* **16**: 37–40.
8. Eppley, R.W., Ringer, E.H., and Williams, P.M. (1976). Chlorine reaction with seawater constituents and the inhibition of photosynthesis of natural marine phytoplankton. *Estuar Coast Mar Sci.* **4**: 147–161.
9. Choi, D.H. et al. (2002). Effects of thermal effluents from a power station on bacteria and heterotrophic nanoflagellates in coastal waters. *Mar Ecol Prog Ser.* **229**: 1–10.

READING LIST

Allonier, A.S., Khalanski, M., Camel, V., and Bermond, A. (1999). Characterization of chlorination by-products in cooling effluents of coastal nuclear power stations. *Mar. Pollut. Bull.* **38**: 1232–1241.

Mattice, J.S. and Zittel, H.E. (1976). Site-specific evaluation of power plant chlorination. *J Water Pollut. Cont. Fed.* **48**: 2284–2308.

MODELING CHLORINE RESIDUALS IN URBAN WATER DISTRIBUTION SYSTEMS

M. S. Mohan Kumar
Indian Institute of Science
Bangalore, India

G. R. Munavalli
Walchand College of Engineering
Sangli, India

INTRODUCTION

The water to be supplied to a community should be adequate in quantity to satisfy their demands and potable in quality to safeguard their health. It is not enough to treat water in the treatment works for safeguarding the health of consumers without trying to take adequate steps in preventing deterioration of the water quality in a distribution system. Treatment can be adequate to provide an excellent product for the consumer at the plant outlet, but it is inadequate to overcome inferior protective measures in the distribution system (1). There are various causes for which the potable water may turn to an obnoxious liquid in the distribution system and become unfit for drinking as could be found after quality assessment. The various causes of deterioration of water quality are pipe materials, constituents of water, intermittent supply, cross connections, dead-end pipes, service reservoirs, commissioning of new distribution system without disinfection, and breakages (2). The significant constituents of water generally include residual chlorine, natural organic matter, nutrient concentrations, temperature, and pH of water. The distribution system plays an important role in assuring a good quality of water to the consumer.

It is well known that the quality of drinking water can change within a distribution system. The movement or lack of movement of water within the distribution system may have deleterious effects on a once acceptable supply. These quality changes may be associated with complex physical, chemical, and biological activities that take place during the transport process. Such activities can occur either in the bulk water column, the hydraulic infrastructure, or both, and they may be internally or externally generated (3). An ability to understand these reactions and model their impact throughout a distribution system will assist water suppliers in selecting operational strategies and capital investments to ensure delivery of high-quality drinking water (4). Water quality modeling provides an engineering insight into the distribution system activities.

The water quality in a distribution system can be represented in the form of a constituent, water age and source trace (in case of multiple sources) analysis. For constituent concentration, the residual chlorine is considered to be important to represent the overall quality of water. Chlorination is a form of disinfection that reduces the risk of infection to an acceptable level by controlling the number of bacteria. The alternative disinfection options available are ozonation and ultraviolet radiation. But these two have no residual action to protect the water during distribution and are expensive. Although the chlorine is not as powerful as ozone, it has a lasting residual effect, is relatively easy to use, and is cost effective. Hence, chlorine happens to be the most predominant water treatment disinfectant. Because of chlorine's oxidizing potential, minimum levels of chlorine residual must be maintained in the distribution system to preserve both chemical and microbial quality of treated water (5). It is well known that as chlorine travels through the distribution system, it reacts with different materials inside the pipe. Free chlorine is consumed on the one hand by water (oxidation of dissolved organic compounds) and on the other hand by the internal walls of pipes (6). For the age-old distribution systems made up of unlined cast iron pipes, the wall component of reaction plays an important role in chlorine reaction kinetics.

CHLORINE REACTION KINETICS

The chlorine disappearance in a drinking water distribution system is governed by both the quality of water (bulk decay) and the chlorine demand of the pipe (wall decay). The chlorine demand exerted in the bulk flow and at the pipe wall is caused by the reaction of chlorine with organic content and materials associated with the pipe wall, respectively. The reactions in bulk flow occur within the fluid volume and are a function of constituent concentration, reaction rate and order, and concentration of the formation products. The various reaction kinetic models that represent bulk decay of chlorine include first-order, second-order with respect to chlorine only, parallel first-order, nth order, limited first-order (7), and two component second-order (8). The variability of bulk reactions with the temperature and organic matter concentration was studied by Kiéné et al. (9). An empirical kinetic model predicted the bulk reaction parameter by taking into account the TOC and temperature. To summarize, the kinetic models are either a single-parameter or a multiple-parameter decay equation. The multiple-parameter decay equations are shown to have an edge over the single-parameter models in describing the chlorine decay in bulk water.

The applicable reaction kinetics for chlorine in the aqueous (bulk) phase can be determined experimentally by conducting the bottle tests, which consist of placing a sample of water in a series of nonreacting glass bottles and analyzing the contents of each bottle at different points in time. The results of these bottle tests are then used for selecting an appropriate kinetic model, and thereby the bulk reaction parameter(s) involved in the model are estimated.

The chlorine decay at the pipe wall is controlled by the mass transfer mechanism from bulk to the pipe wall. The mass transfer expressions are based on either a two-dimensional radial diffusion (10) or a lumped-mass transfer (4) approach. The radial diffusion model can produce more accurate results than those using a lumped-mass transfer coefficient, but it is less suitable for inclusion in network models applied to large systems (11). After the lumped-mass transfer approach, the wall reaction kinetics for chlorine with the materials associated with the pipe wall can be modeled by assuming an active reaction zone located on the wall or in an adjacent molecular layer. The transport of chlorine to the wall is represented by a film resistance model of mass transfer using a mass transfer coefficient suitable to the flow regime in the pipe (4). The reaction at the pipe wall can be modeled according to either first-order or zero-order kinetics. The reactivity of the pipe wall is a function of pipe radius, hydrodynamic conditions, nature of pipe materials, and amount and nature of deposits (12). The first-order wall model may best represent a process in which chlorine is the limiting reactant, as might be the case with reactions involving complex organic compounds that are found in the exocellular enzymes and metabolic products produced by the biofilm on the pipe wall. The zero-order model would better represent the case in which chlorine immediately oxidizes some reductant (such as ferrous compound), and the rate is dependent on how fast the reductant is produced by the pipe. This mechanism would more likely apply to corrosion-induced reactions. These two wall reaction models are further simplified by relating the wall reaction constant inversely with the pipe roughness and thereby converting the equations into a single unknown parameter (5).

MODELING CHLORINE RESIDUALS WITHIN A WATER DISTRIBUTION SYSTEM

The hydraulic analysis of the system is a prerequisite for attempting to model water quality in a distribution system. Steady-state network hydraulic analysis determines the operating behavior of the system under unchanging conditions. Dynamic-state network hydraulic analysis is usually carried out by an extended period simulation that determines the effects on the system over time and is more significant as possible flow reversals in the pipes may affect the quality of water drastically. Steady- or dynamic-state chlorine residual modeling is a direct extension of the corresponding hydraulic network modeling. Basically, the chlorine residual modeling is done in two environments, namely, steady and dynamic. In steady-state modeling, the external conditions of a distribution network are constant in time and the nodal concentrations of the constituents that will occur if the system is allowed to reach equilibrium are determined. The steady-state models determine the ultimate spatial distributions of chlorine concentrations throughout the network. These models (13–18) are based on the law of mass conservation. Dynamic modeling, on the other hand, represents both spatially and temporally varied chlorine concentrations in a distribution system. In dynamic models, the external conditions are temporally varied and the time-varying nodal concentrations of the constituents are determined. Dynamic models rely on a system-simulation approach to determine the movement and fate of chlorine under time-varying demand, supply, and hydraulic conditions. The dynamic-state algorithms developed include (3,9,19–26) models. Because water distribution systems and the processes within them are usually categorized as continuous systems, dynamic modeling provides a more accurate and realistic portrayal of the actual operation of the system.

Flow rates in pipes and the flow paths that define the movement of water through the network can determine mixing, residence times, and other hydraulic characteristics that affect the chlorine transport and reactions. The residence time of water within the system also has its impact on the chlorine concentration. In the chlorine transport models, the physical processes like transport and mixing, and chemical reactions like chlorine decay, are typically represented. The transport of chlorine in water along the length of a pipe takes place by advection and dispersion aided by the reactions involved. Although dispersion of chlorine in water supply pipes with very low flows has been shown to be important (25,27,28), the flows in the water distribution pipes are generally considered to be advection dominated and the dispersion term is negligibly small in most models. Hence, the water quality modeling is generally based on one-dimensional advective–reactive transport through a pipe. A water distribution system, typically represented as a network model, consists of links (pipes, pumps and control valves) interconnected by nodes (junctions, storage tanks and reservoirs) in some particular branched or looped configuration.

STEADY-STATE MODELING

The principle of mass conservation along with the suitable bulk and wall chlorine reaction kinetics are used in formulating the chlorine residual model. The basic governing equation required to determine the steady-state concentration of chlorine at any node j can be formulated as

$$C_j = \frac{\sum_{i=1}^{Ninp_j} C_{nu_i} Rec_i Q_i}{\sum_{i=1}^{Ninp_j} Q_i}; \quad j = 1, \ldots Njn \tag{1}$$

where C_j = chlorine concentration at node j (mg/L); C_{nu_i} = chlorine concentration at node nu_i (mg/L); $Ninp_j$ = number of incoming pipes at node j; nu_i = upstream node of incoming pipe i; Q_i = flow in pipe i (m^3/s); Njn = number of nodes in the network; and Rec_i = reaction coefficient for pipe i. The expression for the reaction coefficient depends on the kinetics used for bulk and wall reactions. For first-order bulk and wall reactions, it is given by

$$Rec_i = Exp - \left[k_{bi} + \frac{k_{wi} k_{f_i}}{r_{h_i}(k_{wi} + k_{f_i})}\right] ptt_i \tag{2}$$

where ptt_i = travel time in pipe i (s); r_{hi} = hydraulic radius of pipe i (m); k_{wi} = first-order wall reaction parameter of pipe i (m/s); k_{bi} = first-order bulk reaction parameter of pipe i (s^{-1}); k_{fi} = mass transfer coefficient of pipe i(m/s) = $(D/d_i)S_{hi}$, where, d_i = diameter of pipe i (m); D = molecular diffusivity of chlorine (m^2/s); and S_{hi} = Sherwood number of pipe i and is given by Rossman (29), for $R_{ei} \leq 2300$ (Laminar flow);

$$S_{hi} = 3.65 + \frac{0.0668\left(\frac{d_i}{L_i}\right)R_{ei}S_c}{1 + 0.04\left[\left(\frac{d_i}{L_i}\right)R_{ei}S_c\right]^{2/3}} \tag{3}$$

for $R_{ei} > 2300$;

$S_{hi} = 0.0149{R_{ei}}^{0.88}{S_c}^{1/3}$;

$S_c =$ Schmidt number $= \nu/D$; and $\nu =$ kinematic viscosity (m^2/s); $L_i =$ length of pipe i.

The reaction coefficients for other combinations of bulk and wall reactions are given in Ref. 25.

If Nsn is the number of source nodes, then Equation 1 with appropriate reaction kinetics results in $(Njn - Nsn)$ number of linear equations. An iterative procedure based on the Gauss–Siedel algorithm can be adopted to solve the equations. The converged solution gives steady-state chlorine concentrations at all nodes for given source chlorine concentrations.

DYNAMIC-STATE MODELING

The unsteady advection–reaction process for the transport of chlorine in a pipe flowing full is given by the following classic equation:

$$\frac{\partial C_i(x,t)}{\partial t} = -v_i\frac{\partial C_i(x,t)}{\partial x} - R[C_i(x,t)] \tag{4}$$

where $C_i(x,t) =$ chlorine concentration in pipe i (mg/L) as a function of distance x and time t; $v_i =$ mean flow velocity in pipe i (m/s); and $R[C_i(x,t)] =$ reaction rate expression. For the first-order bulk and wall chlorine reaction kinetics, it is given by the following equation:

$$R[C_i(x,t)] = k_{bi}C_i(x,t) + \frac{k_{wi}k_{f_i}}{r_{h_i}(k_{wi}+k_{f_i})}C_i(x,t) \tag{5}$$

k_{fi} can be calculated using expressions as described earlier.

Instantaneous and complete mixing of chlorine at the node j and time t is given by the following expression:

$$Cnc_{j,t} = \frac{\sum_{i=1}^{Ninp_j} Q_iC_i(L_i,t) + Q_EC_E}{\sum_{i=1}^{Ninp_j} Q_i + Q_E}; \quad j = 1,\cdots Njn \tag{6}$$

where $Q_i =$ flow in the pipe i (m^3/s); $Q_E =$ external source flow in to node j (m^3/s); $C_E =$ external source chlorine concentration into node j (mg/L); and $Njn =$ total number of nodes in the network.

The numerical approaches adopted to solve the above-formulated problem can be classified spatially as either Eulerian or Lagrangian and temporally as time driven or event driven (30). Eulerian models divide the pipe into a series of fixed, interconnected control volumes and record changes at the boundaries or within these volumes as water flows through them. Lagrangian models track changes in a series of discrete parcels of water as they travel through the pipe network. Event-driven simulation updates the state of the system only at times when a change actually occurs, such as when a new parcel of water reaches the end of a pipe and mixes with water from other connecting pipes, and at the output reporting time (30). The Eulerian methods include the finite difference method (FDM) (30) and the discrete volume element method (DVEM) (21), and the Lagrangian methods are the time-driven method (TDM) (19) and the event-driven method (EDM) (22). The Lagrangian methods are more efficient for simulating the chemical transport in a water distribution system (30).

The input to the dynamic simulation model essentially consists of the system demands, source chlorine concentrations, and reaction parameters (bulk and wall) of the network. The outputs of this model are the spatially and temporally varying nodal chlorine concentrations.

APPLICATION OF THE CHLORINE TRANSPORT MODEL

The Brushy plains zone of the South Central Connecticut Regional Water Authority is chosen to illustrate the application and is shown in Fig. 1. The pipe and node data of the network are taken from EPANET [distributed by USEPA, (29)] example networks. The first-order bulk

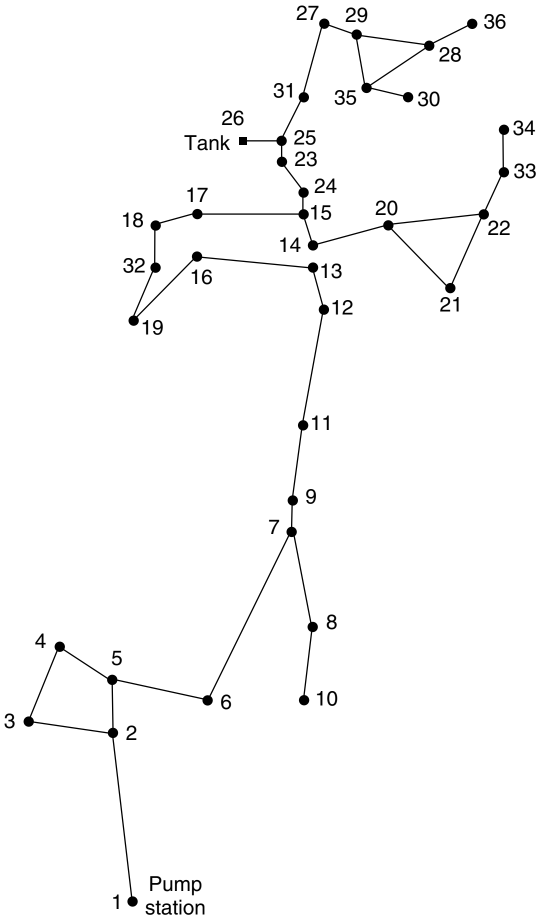

Figure 1. Brushy plains network.

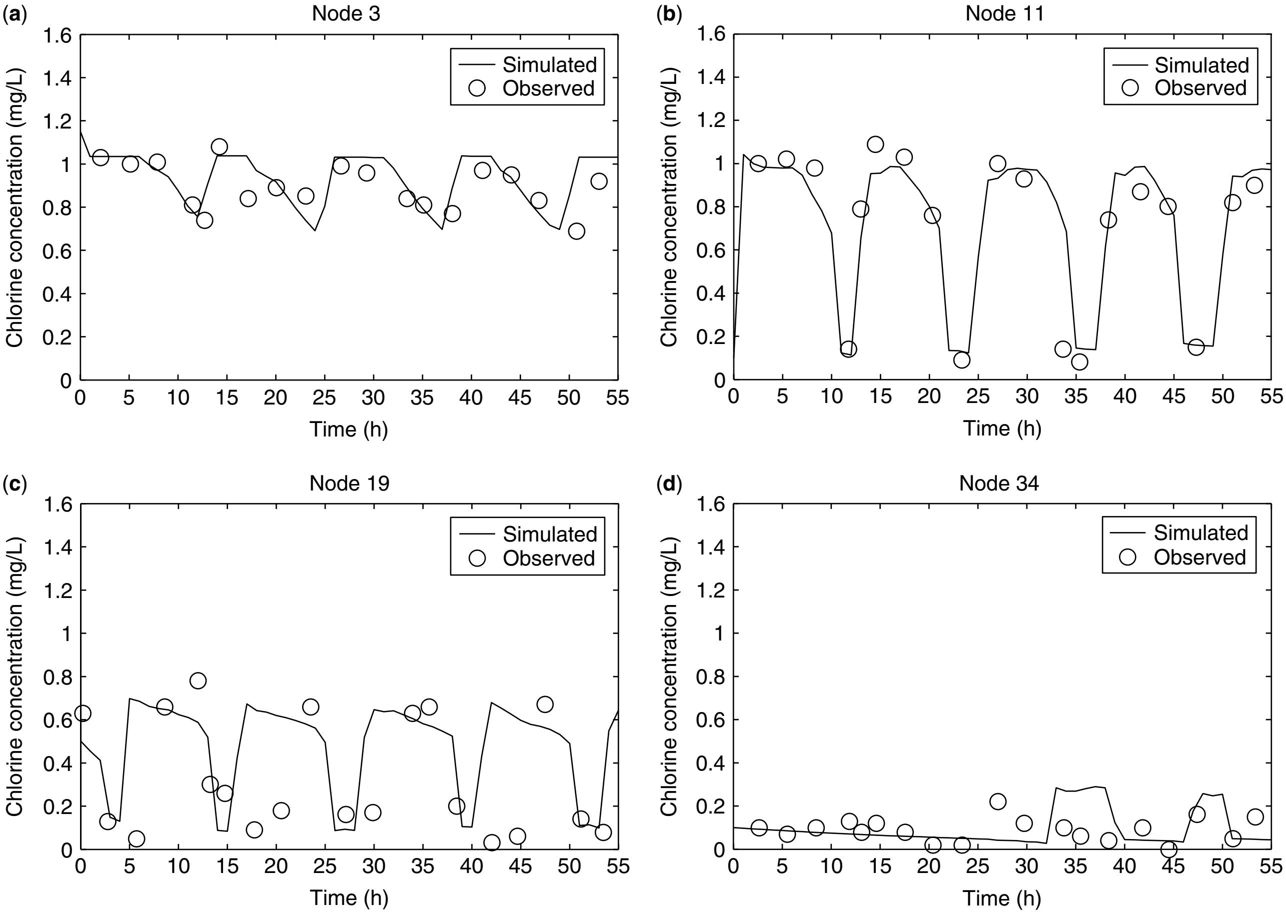

Figure 2. Simulated and observed chlorine concentration at nodes (**a**) 3, (**b**) 11, (**c**) 19, and (**d**) 34.

and wall reaction parameters used are 0.55 d^{-1} and 0.15 m/d, respectively. The system has been hydraulically well calibrated, with most pipes having been assigned roughness coefficients (HWC) of 100. The chlorine input at node 1 (Fig. 1) to the network has a constant value of 1.15 mg/L. The chlorine transport model is run using the time-driven method for a hydraulic time step of 1 hr and a water quality step of 3 min. The results are obtained for a simulation period of 55 hr. The chlorine concentrations at a few network nodes are represented in Fig. 2.

CHLORINE REACTION PARAMETER ESTIMATION

The chlorine transport model described earlier predicts the constituent concentrations throughout the distribution system under steady or dynamic state. The reliability of these predicted concentrations when compared with the field observations depends on the assigned parameter values involved in the type of reaction kinetics used in the model. The parameters that control the chlorine reaction kinetics within the system can be broadly classified into the bulk and wall reaction parameters. Bulk reaction parameters (first- or non-first-order) are associated with individual pipes and storage tanks, assigned to groups of pipes in an area, which is contributed more by a particular source, or applied globally. Wall reaction parameters (first- or zero-order) are associated with individual pipes, applied globally, or assigned to a group of pipes with similar material/age/roughness factors. The wall reaction parameter can also be related inversely to the Hazen–Williams roughness coefficient and represented in the wall reaction pipe-roughness parameter. The advantage of using this sort of representation is that it requires only a single parameter to allow wall reaction parameters to vary throughout the network in a physically meaningful way (5).

The parameters involved in the single-parameter or multiple-parameter bulk reaction expressions can be determined with the data sets observed by conducting the bottle tests on the water samples. By performing measurements in the distribution system, the researchers have calculated the decay rate constant (overall first-order) for site-specific tests (such as fixed pipe diameter, pipe material, or water source). This process may yield reasonable results; however, a wide range of values for this constant is obtained, thus severely limiting its use as a predictive tool (31). The determination of an overall reaction parameter (which represents the combined effect of bulk and wall reactions) and a wall reaction parameter is much more difficult than is establishing a bulk reaction parameter. Hence, these reaction parameters are more a product of calibration. Calibration is a process of adjusting a model so that the simulation reasonably predicts system behavior. The objective of water quality calibration is to capture the steady/transient dynamic behavior of the network. Wall reaction parameters are similar to pipe roughness coefficients in that they can and do vary from pipe to pipe. Unfortunately, unlike the head loss test for pipe roughness values, direct measurements of wall reaction parameters are extremely difficult to make (32). Because these parameter values are difficult to measure, they need to be estimated with the field measurements. The various techniques adopted to estimate these parameters include the trial-error (33), gradient-descent-based search technique (34), method of Lagrange multiplier (35), Gauss–Newton sensitivity analysis technique (36), and stochastic-based genetic algorithm (GA) technique (37). Gradient-based methods are generally faster, but they are more difficult to formulate because either an analytical expression must be derived or the gradient must be approximated. Stochastic search methods are more robust and simpler to formulate and use, but they are generally slower (32).

The parameter estimation can be formulated as an optimization problem so that the difference between the observed and computed chlorine concentrations at the monitoring nodes are minimized in the least-squares sense. Thus, the objective function is given by

$$Minimize\ E = \sum_{j=1}^{M}\sum_{k=1}^{N(j)}[Cno_{j,t_k} - Cnc_{j,t_k}]^2 \tag{7}$$

where M = number of monitoring nodes; $N(j)$ = number of monitoring times at node j; Cno_{j,t_k} = computed chlorine concentration at node j at time t_k (mg/L); and Cno_{j,t_k} = measured chlorine concentration at node j at time t_k (mg/L).

The simulation-optimization inverse modeling technique, which uses the field measurements and simulated chlorine concentrations at monitoring nodes, can be adopted to solve this unconstrained optimization problem. The flow diagram in Fig. 3 illustrates the simulation-optimization procedure if GA is used in its optimization module.

The various unknown reaction parameters constitute a set of decision variables to be evaluated by GA, and this can work by evaluating the fitness of each potential solution that consists of values for the set of unknown reaction parameters. Fitness is determined by comparing how well the simulated chlorine concentrations that result from the candidate solution match the measured values collected in the field. The computationally intensive step involved in the GA technique is the determination of fitness that is somehow related to the objective function value. The GA continues to spawn generations of potential solutions until comparison of solutions from successive generations no longer produces a significant improvement. In addition, the GA process eliminates most routine and tedious aspects of the calibration process. GA will generally achieve better fits to the available data if the correct set of variables is included in the solution and it can establish the correct range of possible solutions.

APPLICATION OF THE INVERSE MODEL FOR PARAMETER ESTIMATION

The network used earlier is chosen for applying the inverse model. The time-varying chlorine concentrations were observed for this network at the nodes 3, 6, 10, 11, 19, 25, 28, and 34 for a period of 55 hr (10). These measurements become input observed data to the inverse model, and the unknown global wall reaction parameter for the first-order reaction kinetics, which is assumed to be applicable for the system, is estimated. The inverse model is run with a zero concentration tolerance, quality time step of 3 min, and hydraulic step of 1 hr. The estimated value of the global first-order wall reaction parameter is found to be 0.3654 m/d, which results in the lowest possible RMS

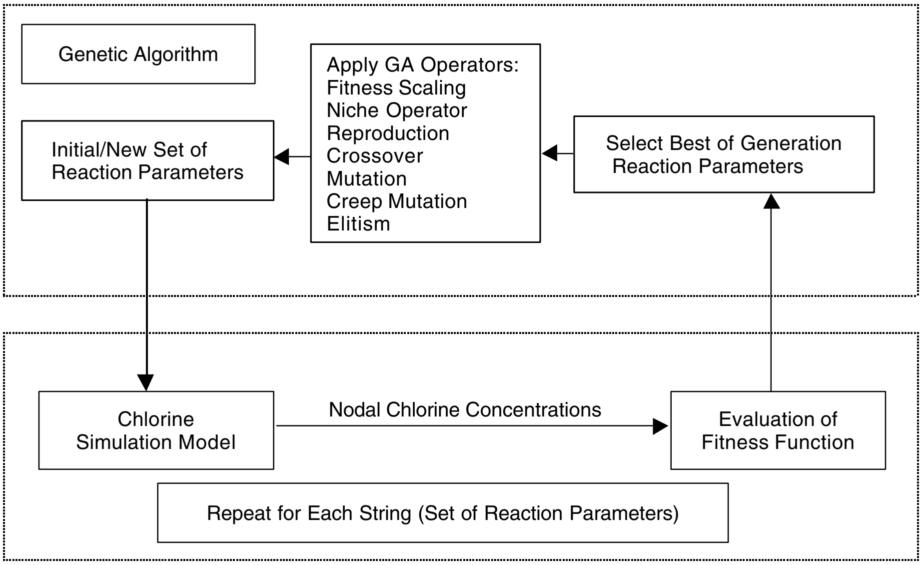

Figure 3. GA implementation.

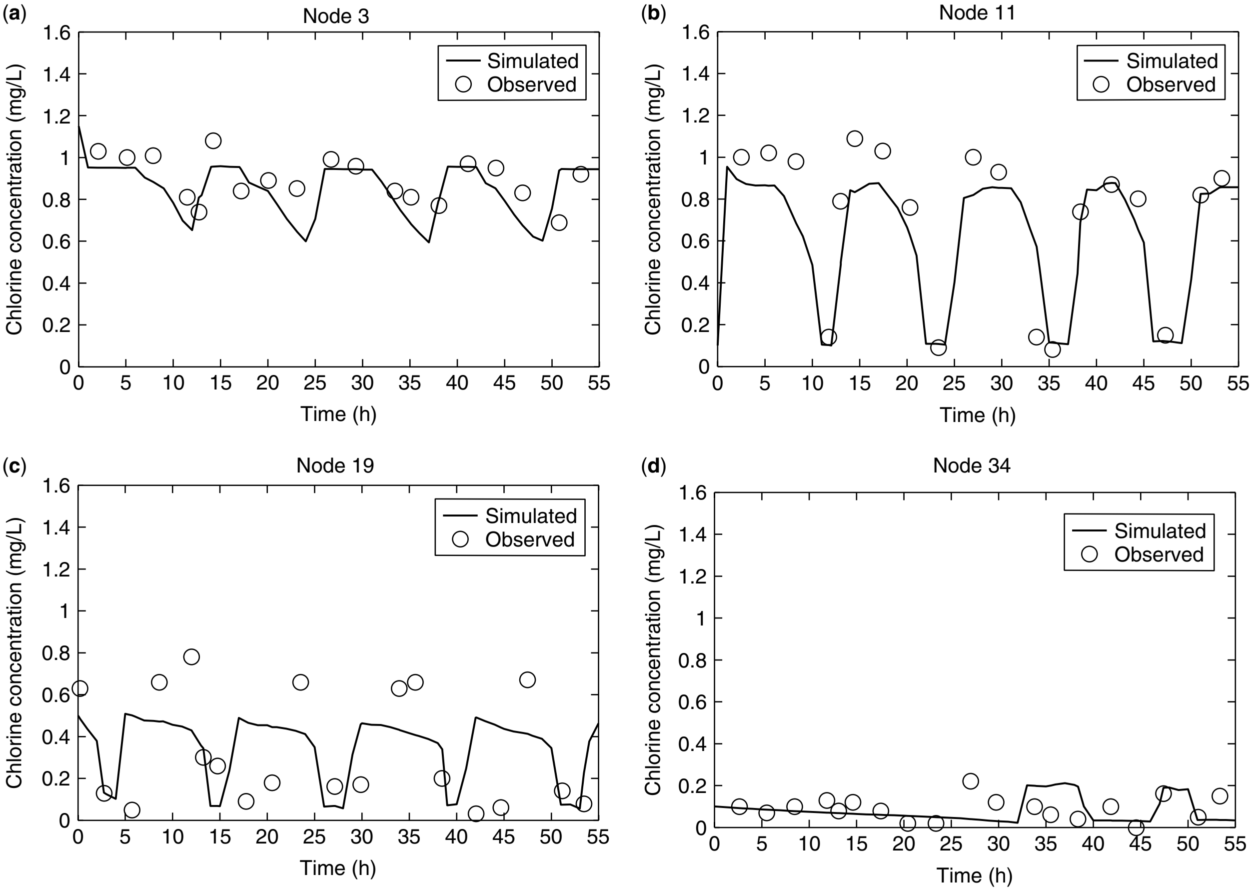

Figure 4. Simulated and observed chlorine concentration at nodes (**a**) 3, (**b**) 11, (**c**) 19, and (**d**) 34.

residual error of 0.172 mg/L without any tedious trial-and-error computations. The comparison of observed and simulated chlorine concentrations for this parameter estimated at node 3, 11, 19, and 34 are represented in Fig. 4.

CLOSURE

The maintenance of chlorine residual is needed at all points in the distribution system supplied with chlorine as a disinfectant. Chlorine is subjected to bulk flow and pipe wall reactions as it propagates through the pipes of the distribution system. The bulk flow reaction term depends on the organic content of the water, whereas the wall reaction term is related to the material and age of the distribution pipe. Because of these reactions, the loss of chlorine is significant between the outlet of the treatment plant and the consumer end. Thus, the study of spatial and temporal distribution of chlorine forms an important aspect of modeling. The chlorine simulation model forms the base for the reaction parameter estimation. The application of the inverse model is more relevant as any combination of bulk and wall reaction kinetics is possible for best fit with field chlorine observations. The inverse model is significant in deciding the operational strategy for real life systems. The chlorine simulation together with the inverse model provides the water supply agencies a tool for better management of their systems.

BIBLIOGRAPHY

1. Larson, T.E. (1966). Deterioration water quality in distribution systems. *J. AWWA.* **58**(10): 1307–1316.
2. Pathak, P.N. (1969). Deterioration of potable water quality in distribution system and its remedial measures. *Environmental Health* **11**(10): 220–228.
3. Boulos, P.F., Altman, T., Jarrige, P.A., and Collevatti, F. (1995). Discrete simulation approach for network water quality models. *J. Water Resour. Plng. Mgmt., ASCE* **121**(1): 49–60.
4. Rossman, L.A., Clark, R.M., and Grayman, W.M. (1994). Modeling chlorine residuals in drinking water distribution systems. *J. Envir. Eng., ASCE* **120**(4): 803–820.
5. Vasconcelos, J.J. et al. (1997). Kinetics of chlorine decay. *J. AWWA.* **89**(7): 55–65.
6. Frateur, I. et al. (1999). Free chlorine consumption induced by cast iron corrosion in drinking water distribution systems. *Water Res.* **33**(8): 1781–1790.
7. Powell, J.C. et al. (2000). Performance of various kinetic models for chlorine decay. *J. Water Resour. Plng. Mgmt., ASCE* **126**(1): 13–20.
8. Clark, R.M. and Sivaganesan, M. (1998). Predicting chlorine residuals and the formation of TTHMs in drinking water. *J. Envir. Eng., ASCE* **124**(12): 1203–1210.
9. Kiéné, L., Lu, W., and Lévi, Y. (1998). Relative importance of the phenomena responsible for chlorine decay in drinking water distribution systems. *Water Res.* **38**(6): 219–227.

10. Ozdemir, O.N. and Ucak, A. (2002). Simulation of chlorine decay in drinking-water distribution systems. *J. Envir. Eng., ASCE* **128**(1): 31–39.
11. Vasconcelos, J.J., Boulos, P.F., Grayman, W.M., Kiéné, L., Wable, O., Biswas, P., Bhari, A., Rossman, L.A., Clark, R.M., and Goodrich, J.A. (1996). Characterization and modeling of chlorine decay in distribution systems. AWWA Research Foundation, Denver, CO.
12. Zhang, G.R., Kiéné, L., Wable, O., Chan, U.S., and Duguet, J.P. (1992). Modeling of chlorine residuals in the water distribution network of Macao. *Environ. Technol.* **13**: 937–946.
13. Chun, D.G. and Selznick, H.L. (1985). Computer modeling of distri-bution system water quality. *Proc. Comp. Appl. Water Resour., ASCE*, NewYork, pp. 448–456.
14. Males, R.M., Clark, R.M., Wehrman, P.J., and Gates, W.E. (1985). Algorithm for mixing problems in water systems. *J. Hydr. Eng., ASCE* **111**(2): 206–219.
15. Clark, R.M., Grayman, W.M., and Males, R.M. (1988). Contaminant propagation in distribution system. *J. Envir. Eng., ASCE* **114**(4): 929–943.
16. Shah, M. and Sinai, G. (1988). Steady state model for dilution in water networks. *J. Hydraulic Eng., ASCE* **114**(2): 192–206.
17. Wood, D.J. and Ormsbee, L.E. (1989). Supply identification for water distribution systems. *J. AWWA.* **81**(7): 74–80.
18. Boulos, P.F., Altman, T., and Sadhal, K. (1992). Computer modeling of water quality in large networks. *J. Appl. Math. Modeling* **16**(8): 439–445.
19. Liou, C.P. and Kroon, J.R. (1987). Modeling the propagation of waterborne substances in water distribution networks. *J. AWWA.* **79**(11): 54–58.
20. Grayman, W.M., Clark, R.M., and Males, R.M. (1988). Modeling distribution system water quality: Dynamic approach. *J. Water Resour. Plng. Mgmt., ASCE* **114**(3): 295–311.
21. Rossman, L.A., Boulos, P.F., and Altman, T. (1993). Discrete volume element method for network water quality models. *J. Water Resour. Plng. Mgmt., ASCE* **119**(5): 505–517.
22. Boulos, P.F., Altman, T., Jarrige, P.A., and Collevatti, F. (1994). An event-driven method for modeling contaminant propagation in water networks. *J. Appl. Math. Modeling* **18**(2): 84–92.
23. Islam, M.R. and Chaudhary, M.H. (1998). Modeling of constituent transport in unsteady flows in pipe networks. *J. Hydr. Eng., ASCE* **124**(11): 1115–1124.
24. Shang, F., Uber, J.G., and Polycarpou, M.M. (2002). Particle backtracking algorithm for water distribution system analysis. *J. Environ. Eng., ASCE* **128**(5): 441–450.
25. Tzatchkov, V.G., Aldama, A.A., and Arreguin, F.I. (2002). Advection-dispersion-reaction modeling in water distribution networks. *J. Water Resour. Plng. Mgmt.* **128**(5): 334–342.
26. Munavalli, G.R. and Mohan Kumar, M.S. (2004). Modified Lagrangian method for water quality in distribution systems. *Water Res.* **38**(8): 2973–2988.
27. Axworthy, D.H. and Karney, B.W. (1996). Modeling low velocity/high dispersion flow in water distribution systems. *J. Water Resour. Plng. Mgmt., ASCE* **122**(3): 218–221.
28. Ozdemir, O.N., and Ger, A.M. (1999). Unsteady 2-D chlorine transport in water supply pipes. *Water Res.* **33**(17): 3637–3645.
29. Rossman, L.A. (2000). EPANET 2 Users manual. Risk Reduction Engineering Laboratory, U.S. Environmental Protection Agency, Cincinnati, Ohio.
30. Rossman, L.A. and Boulos, P.F. (1996). Numerical methods for water quality in distribution systems: A comparison. *J. Water Resour. Plng. Mgmt., ASCE* **122**(2): 137–146.
31. Biswas, P., Lu, C.S., and Clark, R.M. (1993). Chlorine concentration decay in pipes. *Water Res.* **27**(12): 1715–1724.
32. Walski, T.M., Chase, D.V., and Savic, D.A. (2001). *Water Distribution Modeling*. Haestad Press, Baltimore, MD.
33. Clark, R.M., Rossman, L.A., and Wymer, L.J. (1995). Modeling distribution system water quality: Regulatory implications. *J. Water Resour. Plng Mgmt., ASCE* **121**(6): 423–428.
34. Al-Omari, A.S. and Chaudhry, M.H. (2001). Unsteady-state inverse chlorine modeling in pipe networks. *J. Hydrol. Eng., ASCE* **127**(8): 669–677.
35. Zeirolf, M.L., Polycarpou, M.M., and Uber, J.G. (1998). Development and autocalibration of an input-output model of chlorine transport in drinking water distribution systems. *IEEE Trans. Control Syst. Technol.* **6**(4): 543–553.
36. Munavalli, G.R. and Mohan Kumar, M.S. (2003). "Water quality parameter estimation in a steady state distribution system", *J. Water Resour. Plng and Mgmt., ASCE*, **129**(2), 124–134, 2003.
37. Mohan Kumar, M.S. and Munavalli, G.R. (2003). Autocalibration of chlorine transport model for steady state distribution system by genetic algorithm. *Seventh International Water Technology Conference IWTC-2003*, Egypt, May 1–3.

PARTICULATE MATTER REMOVAL BY COAGULATION

RASHEED AHMAD
Khafra Engineering
Consultants
Atlanta, Georgia

Particulate matter in natural water varies in size, concentration, and surface chemistry. The particle size may range from a few tens of nanometers to a few hundred micrometers. Discrete particles less than one micron in size are called colloidal. Colloidal particles have significantly higher external surface area per unit area and move in a random diffusional motion known as Brownian motion. In colloidal suspension, surface phenomena dominate over mass phenomena. The most important surface property is the accumulation of electrical charges at the particle surface. Loss of atoms due to abrasion, molecular arrangement within the crystal, and imperfections within the molecular structure may result in surfaces being charged. The colloidal particles in most surface water are negatively charged. Because of hydration and/or electrostatic surface charges, colloidal particles repel other material and thereby remain suspended. Surface waters that are turbid due to colloidal particles cannot be clarified without special treatment.

Coagulation is a process for enhancing the tendency of particulate matter in aqueous suspension to attach to one another and/or to attach to collector surfaces. Coagulation promotes destabilization of surface charges on colloidal particles. Destabilization and aggregation of particulate matter and precipitation or adsorption of NOM in subsequent solid–liquid separation processes are the primary functions of the coagulation process. The coagulation process involves two steps: (1) the addition of chemical coagulants to destabilize particulate matter and react with NOM and (2) the physical transport of collisions among particulate matter, resulting in aggregation or floc formation. In the water treatment literature, coagulation refers to all reactions and mechanisms that result in aggregation, and the physical transport step of producing interparticle aggregation is called flocculation. In a water treatment plant, coagulation is achieved by rapid or flash mixing of coagulants followed by flocculation.

The two most common types of coagulants are metallic salts and polymers; the most common metallic salt coagulants are aluminum sulfate (alum) and ferric chloride. The selection of a particular coagulant depends on the required level of effectiveness. A standard jar test is a recommended method for determining the relative effectiveness of coagulants for a particular raw water supply. The factors that are considered normally in selecting a coagulant include cost, availability, overall safety, ease of storage, handling, and application.

Alum is the most widely used coagulant because of its availability, low cost, ease of use, and ease of storage. Ferric chloride, other metallic salts, and polymers are less widely used. Alum's performance, however, is greatly affected by the pH of the influent. The commonly used dosage of alum ranges from 5 to 150 mg/L, but the problem of sludge disposal increases at higher alum dosages. Due to special raw water characteristics and because of health concerns about aluminum, some water utilities use ferric chloride. Although ferric chloride is not always as effective as alum in reducing trihalomethane formation potential (THMFP) and total organic carbon (TOC), it is more effective than alum for water that has high dissolved color, low turbidity, and a moderate pH.

Polymers are effective coagulants, coagulant aids, and filter aids. They consist of monomers and are classified according to their charge or lack of charge. A polymer that has a charge is an ionized polymer, or a polyelectrolyte. Polymers can be cationic, anionic, or nonionic. In applications where polymers are effective, dosages are generally lower than alum dosages for the same effect. Typical polymer dosages range from 1.5 to 10 mg/L. Consequently, polymer coagulants produce less residual sludge than alum.

Coagulant aids are added to the influent after or simultaneously with the primary coagulants to improve particle capture efficiency during flocculation, sedimentation, and filtration. Nonionic and anionic polymers are commonly used as coagulant aids. The ratio of alum to coagulant aid dosages ranges from 100:1 to 50:1. Standard jar tests are required to determine precise coagulant aid dosages.

There are four coagulation mechanisms that, it is thought, occur in destabilizing colloidal particles: double layer compression, surface charge neutralization, sweep coagulation, adsorption and interparticle bridging.

The double layer model is used to understand the ionic environment near a charged colloid particle. The surface charge on the colloid attracts ions of opposite charge and forms a dense layer adjacent to the particle known as the Stern layer. Excess positive ions are still attracted by the negatively charged colloids but are repelled by the Stern layer. This dynamic equilibrium results in creating a diffuse layer of counterions. The Stern and the diffuse layer in the interfacial region around colloidal particles are referred to collectively as the double layer. The electrical potential at the junction of the Stern layer and the diffuse layer, called the zeta potential, can be measured experimentally. It correlates with colloid particle stability. Highly stable colloidal systems are characterized by a high zeta potential, whereas lower zeta potentials reflect less stable systems. The DLVO theory (named after Derjaguin, Landau, Verwey, and Overbeek) governs the net interactive force between colloidal particles by combining the van der Waals attractive force and the electrostatic repulsion force. The double layer can be compressed by adding a coagulant that has a positive charge (to counteract negatively charged colloids). In water treatment practice, destabilization by double layer compression is not a dominant mechanism because it requires an extremely high salt concentration. This is an important destabilization mechanism in natural systems, for instance, delta formation in estuaries.

Destabilization by surface charge neutralization involves reducing the net charge of colloidal particles in the suspension. The net surface charge can be reduced by adjusting the solution chemistry. In other cases, colloidal particles can be destabilized by neutralizing using counterions of coagulants. In water treatment practice, a similar type of surface charge destabilization occurs that is called heterocoagulation. The distribution of charges on a colloidal surface is not uniform. Large particles that have high negative surface charges may come in contact with smaller particles that bear relatively low positive charges. These particles may be destabilized by simple electrostatic interaction.

Sweep coagulation or sweep-floc coagulation is also known as enmeshment in a precipitate. At higher coagulant doses, excess metal salts hydrolyze into metallic hydroxides. These hydroxides are extremely insoluble in water, amorphous, heavier than water, and gelatinous. As the hydroxide precipitate forms and accumulates, the colloidal particles are enmeshed or entrapped in the hydroxide floc. This destabilization mechanism is called sweep coagulation.

Interparticle bridging destabilization occurs when high-molecular-weight polymers are used as coagulants or coagulant aids. These polymers are highly surface-active, and their surface structure may be linear or branched. The polymers destabilize particles by first adsorbing at one or more sites on the colloidal particle surface and then extending the chain length into solution and attaching to other particles. This results in forming an interparticle bridge. Sometimes an excessive dosage of polymer may cause restabilization due to surface saturation or sterical stabilization.

READING LIST

ASCE and AWWA. (1998). *Water Treatment Plant Design*, 3rd Edn. McGraw-Hill, New York.

AWWARF. (1998). *Treatment Process Selection for Particle Removal*. American Water Works Association Research Foundation, Denver, CO.

AWWA. (1999). *Water Quality and Treatment*, 5th Edn. McGraw-Hill, New York.

Hudson, H.E. (1981). *Water Clarification Processes*. Van Nostrand Reinhold Company, New York.

Kawamura, S. (2000). *Integrated Design of Water Treatment Facilities*. John Wiley & Sons, New York.

Qasim, S.R., Motley, E.M., and Zhu, G. (2000). *Water Works Engineering*. Prentice-Hall, Upper Saddle River, NJ.

USEPA. (1990). *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*. Office of Drinking Water, Center for Research Information, Cincinnati, OH.

SELECTIVE COAGULANT RECOVERY FROM WATER TREATMENT PLANT RESIDUALS USING THE DOMAIN MEMBRANE PROCESS

Prakhar Prakash
Pennsylvania State University
University Park, Pennsylvania

Arup K. SenGupta
Lehigh University
Bethlehem, Pennsylvania

BACKGROUND

During the last two decades, pressure-driven membrane processes, namely, reverse osmosis (RO), nanofiltration (NF), and ultrafiltration (UF), have found increased applications in water utilities and chemical industries. Unlike RO, NF, and UF, Donnan membrane process (DMP) or Donnan dialysis is driven by an electrochemical potential gradient across an ion exchange membrane. Theoretically, the Donnan membrane process is not susceptible to fouling because particulate matter or large organic molecules do not concentrate on the membrane surface, as commonly observed with pressure-driven membrane processes. Although information on several applications of DMP is available in the open literature (1,2), no work is reported on the use of DMP to treat a sludge or slurry with a high concentration of suspended solids or large organic molecules. It was conceived that a single-step Donnan membrane process could selectively recover coagulant alum ($Al_2(SO_4)_3 \cdot 14H_2O$) (3,4) from water treatment plant sludge or water treatment plant residuals (WTR), which are the endproduct of coagulation. WTR contain insoluble aluminum hydroxide (50–75%) along with suspended inorganic particles, natural organic matter (NOM), and trace amounts of heavy metal precipitates (5).

Several efforts were made to recover alum from WTR. The acid digestion process is the most commonly tried process at the laboratory, pilot-scale, and plant level (6). In this process, WTR are sufficiently acidified with sulfuric acid, dissolving insoluble aluminum hydroxide in the form of alum up to aluminum concentration levels of 360–3700 mg/L. However, the process is nonselective; with the dissolution of aluminum hydroxide, NOM-like humates and fulvates get dissolved too, and the resulting dissolved organic carbon (DOC) concentration ranges from 326 to 1800 mg/L (7). This recovered alum, if reused as a coagulant, may impart a high trihalomethane formation potential (THMFP) during the chlorination stage of water treatment. The trihalomethanes are suspected carcinogens regulated by the USEPA (8). As an alternative to the acid digestion process; the amphoteric nature of aluminum oxide also permits alum recovery from the WTR under alkaline conditions. However, the alkali digestion process suffers from the same limitation as the acid digestion process; i.e., NOM concentration is very high in the recovered solution. Figure 1 shows both DOC and aluminum concentrations of the Allentown Water Treatment Plant (AWTP) in WTR at different pH levels. The Donnan membrane process is uniquely capable of recovering alum from WTR in a single-step process using sulfuric acid and a cation-exchange membrane.

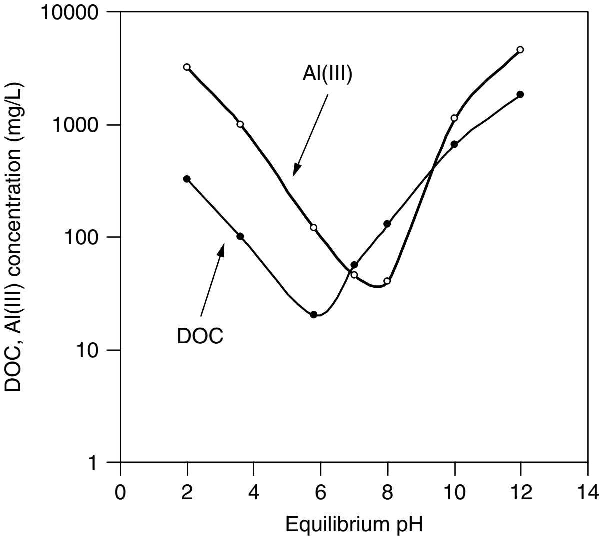

Figure 1. Variation of dissolved organic carbon (DOC) and aluminum concentration with pH for water treatment residuals (WTR) from Allentown Water Treatment Plant (AWTP).

THEORY

Let us consider solutions of aluminum sulfate (feed) and sulfuric acid (recovery) in a Donnan membrane cell divided into two chambers by a cation-exchange membrane that allows only cations to migrate from one side to the other but rejects any passage of anions according to Donnan's co-ion exclusion principle (9). At equilibrium, the electrochemical potential of aluminum ion Al^{3+} ion ($\overline{\mu}$) in the feed solution will be the same as that in the recovery solution for both aluminum and hydrogen ions, which corresponds to the following Donnan equilibrium condition:

$$\left(\frac{C_{Al}^{R}}{C_{Al}^{L}}\right) = \left(\frac{C_{H}^{R}}{C_{H}^{L}}\right)^{3}$$

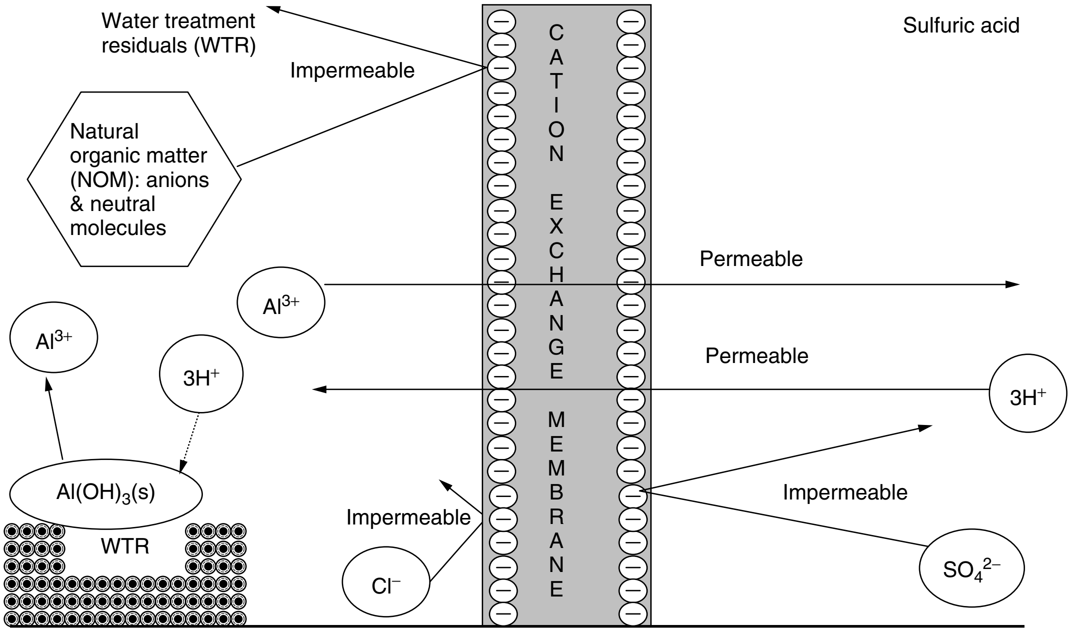

Figure 2. A schematic of Donnan membrane process illustrating selective alum recovery from WTR.

If the ratio $\frac{C_H^R}{C_H^L}$ is 10, it means C_{Al}^R is 1000 times greater than C_{Al}^L. Thus, by maintaining high hydrogen ion concentration in the recovery solution, aluminum ions can be driven from the feed to the recovery side even against a positive concentration gradient, i.e., from a lower concentration region to a higher concentration one. Figure 2 depicts the conceptualized selective alum recovery from WTR, highlighting the following.

KEY FINDINGS

In the Donnan membrane cell, the feed side of the membrane contained 6 L of the decanted and slightly acidified WTR collected from the AWTP, whereas the recovery side contained 1.5 L of 10% sulfuric acid solution, separated by a cation-exchange membrane Nafion 117. At the start, pH of the WTR side was between 3.0 and 3.5. With the progress of the run, aluminum ions from the WTR side moved to the recovery side through the cation-exchange membrane, whereas an equivalent amount of hydrogen ions permeated to the WTR side, thus further reducing the pH. Under the experimental conditions of the Donnan run, free aluminum ions, Al^{3+}, was the predominant aluminum species. Figure 3 shows the results of the process for a period of 24 hours; the percentage aluminum recovery and the concentration of aluminum in the two chambers were plotted against time. It can be seen that over 70% recovery (72%) was attained in 24 hours. The noteworthy observation is that the

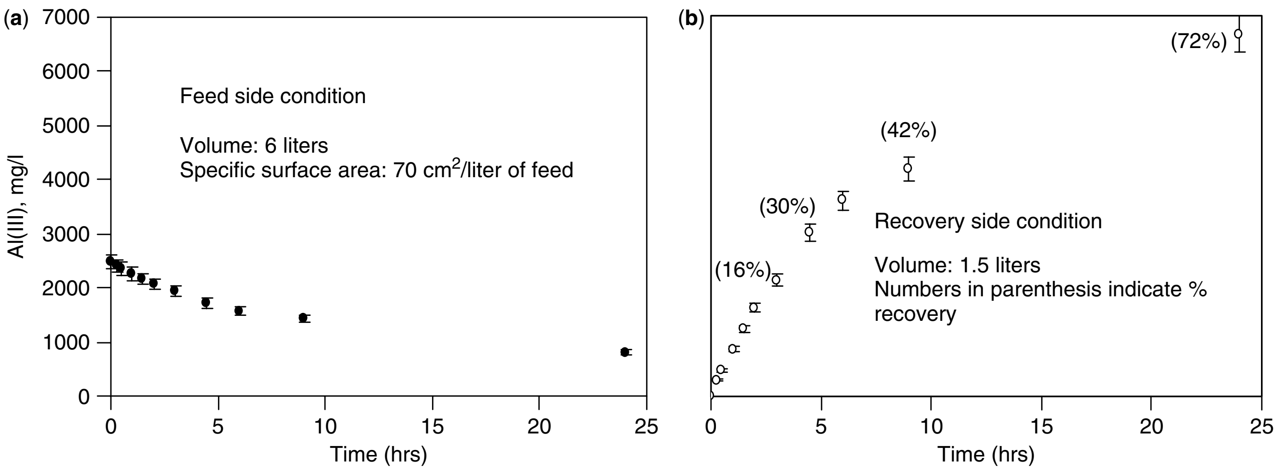

Figure 3. Aluminum recovery from AWTP residuals during Donnan membrane process: (**a**) decrease in Al concentration in feed; (**b**) percentage recovery and increase in Al concentration in recovery solution.

recovered aluminum concentration was 6,650 mg/L as Al, and it was significantly greater than the total aluminum concentration (2400 mg/L) present in the parent sludge. It was also noted that the recovery was selective with respect to trace heavy metal ions. The recovered alum did not contain any suspended solids, whereas NOM expressed as DOC was consistently less than 5 mg/L. The ratio of individual contaminants to aluminum in the recovered alum was comparable, and in some cases lower, than in the commercial alum currently being used in AWTP. Similar results were obtained with WTR received from the Baxter Plant (Philadelphia, PA), which used $FeCl_3$ as a coagulant, where over 75% recovery was made in 24 hours.

Figure 4(a,b) show the visual comparison of recovered coagulants, both alum and ferric sulfate, between the traditional acid digestion process and the Donnan membrane process. Higher transparency of the coagulants from AWTP and the Baxter Plant, recovered by Donnan membrane process, is readily noticeable because of the absence of turbidity and NOM.

CONCLUSIONS

In this work, it was worthy to note that (a) aluminum (ferric) hydroxide precipitates could be dissolved and coagulant ions concentrated in the recovery solution; (b) negatively charged NOM, sulfate, and chloride could not permeate the membrane because of Donnan exclusion; and (c) the recovered alum was sufficiently pure and reuseable in water treatment plants.

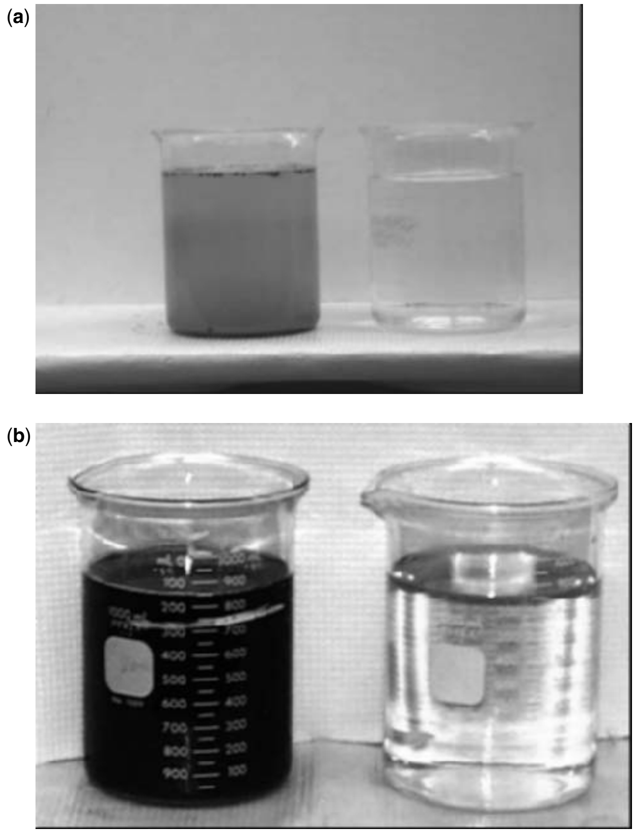

Figure 4. (**a**) Visual comparison of recovered alum coagulant from AWTP residuals by acid digestion process (left) and Donnan membrane process (right) (**b**) Visual comparison of recovered ferric coagulant from Baxter Plant residuals by acid digestion process (left) and Donnan membrane process (right).

BIBLIOGRAPHY

1. Wallace, R.M. (1967). *Ind. Eng. Chem. Process Des. Dev.* **6**(4): 423–431.
2. Kim, B.M. (1979). *AIChE Symposium Series* **76**(197): 184–192.
3. Prakash, P. and SenGupta, A.K. (2003). *Environ. Sci. Technol.* **37**: 4468–4474.
4. SenGupta, A.K. and Prakash, P. (2002). Process for selective coagulant recovery from water treatment plant sludge. U.S. Patent 6,495,047.
5. Cornwell, D.A and Westerhoff, G.P. (1981). Management of water treatment plant sludges. In: *Sludge and Its Ultimate Disposal*, J.A. Borchardt (Ed.). Ann Arbor Science, Ann Arbor, MI.
6. Bishop, M.M., Rolan, A.T., Bailey, T.L., and Cornwell, D.A. (1987). *J. AWWA.* **79**(6): 76–83.
7. Saunders, F.M. (1989). *Coagulant Recovery from Alum WTR at North Area Plant*. Final Report for Bureau of Water, City of Atlanta, GA.
8. Christman, R.F., Norwood, D.L., Millington, J.D.S., Stevens, J.D., and Stevens, A.A. (1983). *Environ. Sci. Technol.* **17**: 625–628.
9. Helfferich, F. (1995). *Ion Exchange*. Dover Publication, New York, pp. 134–135.

PHYSICAL WATER CONDITIONING

Charles H. Sanderson
Magnatech Corporation
Fort Wayne, Indiana

Magnetic water treatment technology that was used two centuries ago to control hard water scale is now being used in hundreds of different kinds of beneficial applications throughout the world. Permanent magnetic water conditioning (PMWC) does not change the chemistry of the water; only a *physical* change takes place. It is therefore referred to as physical water conditioning (Fig. 1) (1).

Physical water conditioning (PWC) has been highly controversial for many years; however, due to numerous successful installations on boilers, cooling towers, and other HVAC equipment, the technology has gained credibility throughout the industry in recent years. The ASHRAE Handbook (considered the "bible" of the industry) states, "Equipment based on magnetic, electromagnetic, or electrostatic technology has been used for scale control in boiler water, cooling water, and other process applications."

Several laboratory tests have been conducted to evaluate the results of magnetically treated water for scale and corrosion control in heat transfer equipment. The purpose of one such test was to determine the physical or chemical differences between treated and untreated residues. Emission spectrographic analyses of

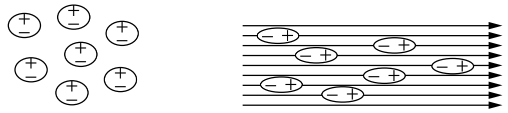

Arrangement of nonpolar molecules in the absence of a magnetic field

Arrangement of nonpolar molecules under the influence of a magnetic field

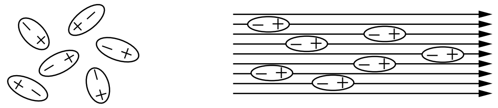

Arrangement of polar molecules in the absence of a magnetic field

Arrangement of polar molecules under the influence of a magnetic field

Figure 1. The effect of a magnetic field.

the two water residues, it was found, have the same chemical constituents; however, a distinctive difference was observed in the crystalline structure. The residue from the magnetically conditioned water was a soft powder (when dry), whereas the untreated water deposit was typical boiler scale (2).

Disassociated dissolved molecules of $CaCO_3$ in water have a tendency to recombine by forming scale that adheres to the inner walls of the piping system, containers, steam vessels, etc. When water flows through a magnetic field of relatively low intensity, the formation of scale in the treated water is prevented in many instances. Instead, aragonite is formed within the flowing bulk water (aragonite forms a dilute slurry in the water, and the sediment can be easily removed by blowdown or bleed-off). The magnetochemical reaction is only one of the many cross effect reactions that enable the transformation of calcite to aragonite. Other reactions include thermochemical and mechanochemical reactions (Fig. 2).

One of the benefits of PWC is energy savings from keeping the heat transfer area clean and free of scale, which prevents efficient transfer from the energy side to the water (see Fig. 3).

Billions of dollars are lost in the United States every year due to corrosion. Laboratory tests and field installations, using corrosion coupons, as a comparison, have proven that magnetic water treatment reduces corrosion rates in HVAC equipment. This is accomplished by eliminating the aggressive chemicals used to control scale, and aragonite talc, a by-product of calcium, which puts a microscopic film on all wetted parts and provides protection against oxygen pitting. The $CaCO_3$ with the PMWC does not build upon itself; however, a thin transparent coating deposits on the metal water side and dries as a fine white or gray powder.

The National Aeronautics and Space Administration (NASA), using their Dynamic Corrosion Test System in an evaluation study, compared magnetically conditioned water to chemical treatment. Corrosion coupon #42 in the magnetically conditioned water loop had a corrosion rate of 0.0 mils per year; #41 and #43, treated with chemical corrosion inhibitors in the loop, had 5 and 6 mils loss per year respectively (Fig. 4) (3).

(a)

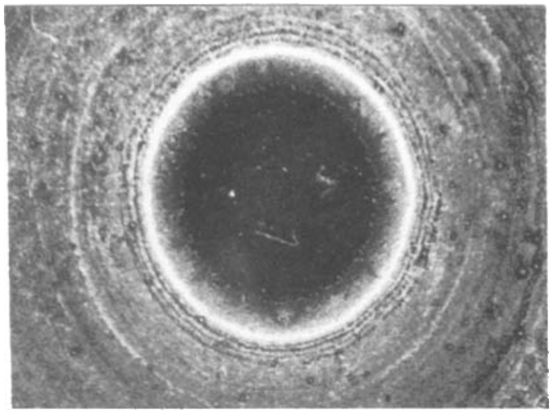

Untreated raw water

(b)

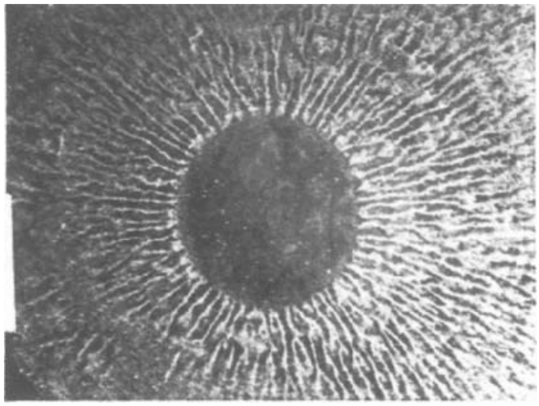

Magnetically treated water

Figure 2. Partially evaporated water droplets magnified to the 40th power.

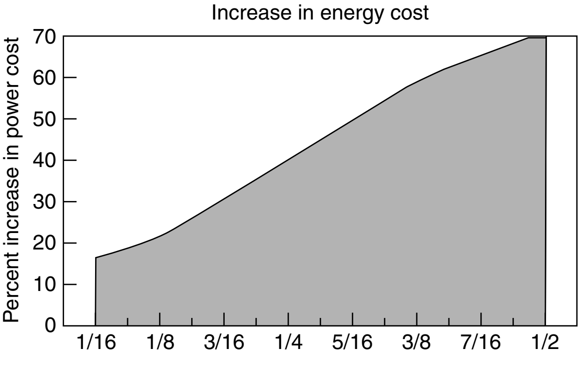

(Above data from the University of Illinois and the Bureau of Standards)

Figure 3. Additional fuel consumed with scale accumulation.

An independent laboratory used a test rig consisting of two 48 inch glass cylinders filled with steel wool and tapered at one end to evaluate the effect of magnetically treated water on corrosion. The flow rate of the water entering the top of each cylinder (one treated and one untreated) was adjusted to make the water head in the cylinders about 10 inches. After 72 hours, water in the untreated system overflowed the glass cylinder. The flow rate was then determined (4).

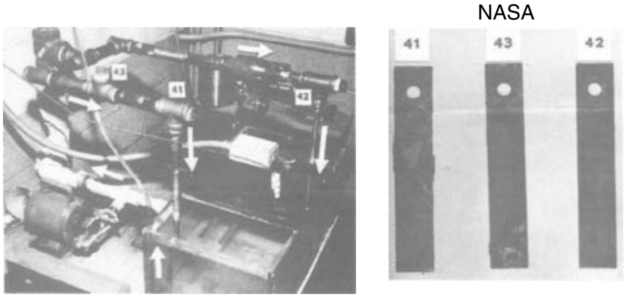

Figure 4. NASA dynamic corrosion test stand and # C-75-1495 photos of coupons.

	Untreated Tap Water	Treated Tap Water
Flow Rate	25 mL per 58.0 s	25 mL per 7.5 s
Water height	46 inches	20 inches

A PMWC system was installed on a 4-year-old, 1500-ton cooling tower and chillers that had moderate scale buildup in the tower and condenser tubes. There was also a considerable amount of corrosion and pitting on the inside of the chiller end caps. The chemical company that treated the system during the first four years of operation suggested that they monitor the results of the PMWC system.

Six months after the installation, the chillers were opened for inspection. The end caps did not require wire brushing to remove the hard brittle scale as previously found, and the small accumulation of mud in the tubes was easily flushed out of the system by a garden hose.

Prior to the PMWC installation, the average corrosion rate of the copper coupon ID # 10008 16-S was 0.01 MPY. After the PMWC installation, the copper coupon ID # 10007 A0882 corrosion rate was also 0.01 MPY; however, the loss was even and general and had an average penetration of 0.15 mils versus 0.50 mils penetration using the previous chemical treatment.

The prior steel coupon ID #10008 48 V had a corrosion rate of 0.36 MPY and an average pit depth of 0.500 mils. The PMWC steel coupon ID # 10007 A3379 had a corrosion rate of 0.22 MPY and an even and general average penetration of 0.150 mils. A 100% return on investment took only 27 months in chemical cost alone (Fig. 5)

In recent years, *pollution prevention* has taken on a new meaning. The transfer of polluted material from one place to another was, at one time, considered "prevention" if the contaminated water was prevented from entering our freshwater streams. However, true prevention is now understood as *eliminating the generation* of toxic waste. The 200-year-old technology is now being used to replace chemicals previously used for water treatment. Research has been spurred recently due to public awareness and concern for the environment. Several different types of applications have been discovered since the turn of the century. The science laid dormant for many years; however, it is now considered "state of the art technology" and is gaining momentum in numerous industrialized countries.

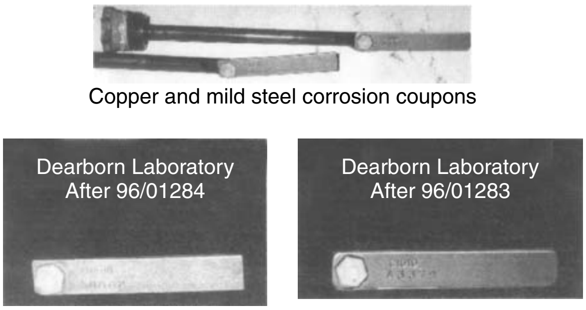

Figure 5. Nissan/Dearborn corrosion coupons.

The arrangement of the magnetic fields is a very important factor to provide adequate and effective treatment for most applications. To a multiple reversing polarity field that has a N–SS–NN–S arrangement is the most successful in controlling hard water scale deposition (Fig. 6) (5).

In addition to the multiple reversing fields, research has shown that scale can be best controlled when the water containing the minerals cuts through the magnetic lines of force at right angles. A steel pipe (magnetic material) surrounding the magnet pulls the magnetic lines of force at right angles through the water passageway. The pipe directs the water flow perpendicular to the magnetic fields (Figs. 7 and 8) (6).

Since the turn of the century, there have been dozens of new beneficial applications for magnetically treated water found through new research discoveries and hundreds more in the field. There are numerous case histories of PMWC installations on various types of equipment that have been documented by third parties as successful applications (7). Most convincing are the original equipment manufacturers (OEM) who install magnetic water conditioners on their equipment to extend the period of time between maintenance service calls.

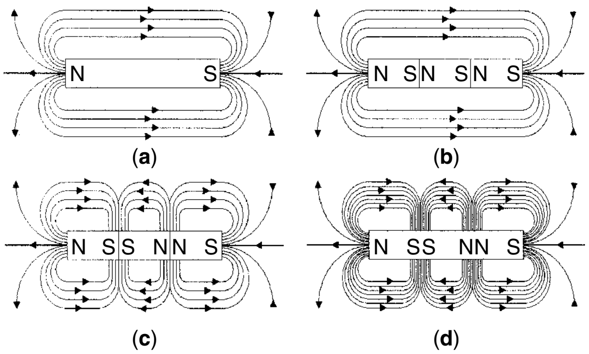

Figure 6. (**a**) One 2 pole, single field; (**b**) Three 2 pole, single field; (**c**) Three 2 pole, three separate fields; (**d**) One 6 pole, three dense fields.

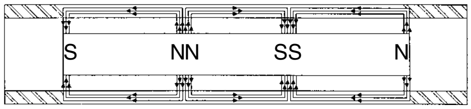

Figure 7. One 6 pole, three dense fields, with steel shield.

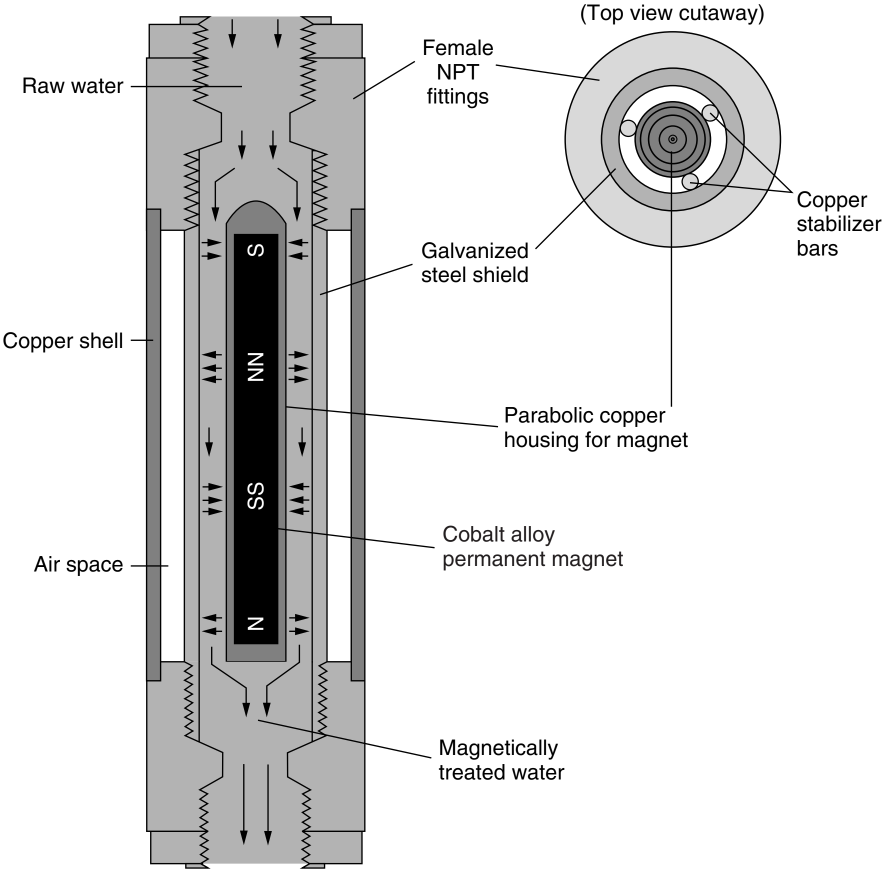

Figure 8. Cutaway view of a PMWC unit showing the water path perpendicular to the magnetic fields.

Water softeners that weighed over 100 pounds, plus bags of salt that were carried in carpet cleaning vans, are being replaced by small magnetic water conditioners (only $1\frac{1}{4}$ inches in diameter and 12 inches long), weighing less than 2 lb. Thousands of these small specially designed units for truck-mounted applications have been installed to prevent scale in steam boilers and related equipment by the manufacturer of a turnkey carpet steam cleaning equipment package.

Hard water is one of the most common causes of truck-mounted heating system failure. Untreated hard water causes loss of performance in heat exchangers and, many times, permanent damage to components (8). Several manufacturers of carpet cleaning systems throughout the United States have installed thousands of PMWC units on their truck-mounted systems during the last 5 years and reported excellent scale control and corrosion protection. Several manufacturers of the truck-mounts now encourage the use of PMWC by their customers who purchased their equipment with water softeners prior to the changeover.

Magnetic water conditioning is believed to be the wave of the future by many proponents, especially environmental groups. It is nonpolluting, does not require energy for continued operation, and needs very little, if any, maintenance. The total value of the overall benefits of PMWC is still not known. There are hundreds of different types of successful applications now in use, and there are dozens more evaluations of laboratory studies and field installation presently underway. The PMWC system has the ability to enhance our environment by reducing chemical use (true pollution prevention) and also helps save energy. The technology is cost-effective and plays a big role in water conservation. (9).

The steel industry is reportedly the single largest consumer of PMWC technology in *total gallons treated per day* due to the high demand for water needed to cool its furnaces. Units are presently available in sizes up to 50,000 gallon per minute with 72 inch flanges.

Economics is also one of the deciding factors in reducing chemical usage; the cost of preventing contaminated process wastewater from entering our freshwater rivers, lakes, and underground aquifers is very high. Transportation to another location can be as much as three times the cost of the chemical, depending on how hazardous the discharge water is.

The food service industry is the largest user of the PMWC in *total number of units*; installations are on ice machines, coffee makers, dish washers, proofers, steamers, and drink dispensers. Sizes usually range from 1 gpm up to 15 gpm (10). Original equipment manufacturers that supply the food service industry have found that adding a small specially designed unit to their water using product eliminates the majority of service calls and replacement of parts under warranty that are prone to scaling or corrosion (11).

As a result of millions of successful installations reported (many documented by qualified professionals), **the question of whether or not magnetic water**

conditioners work has been put to rest. However, the issue of how they work is still being debated. ASHRAE funded a comprehensive research study in 2000, titled *"Efficiency of Physical Water Treatment for the Control of Scale"* This report, published in 2003, confirms the results of several PMWC units performing under very rigid conditions and identifies the most effective velocity through each unit to control scale on heat transfer surfaces (12).

BIBLIOGRAPHY

1. Quinn, C.J. (1986). Professor Emeritus, Indiana University-Purdue University Chairman, Manufacturing Technology. *Magnetic Water Treatment For Heating, Refrigeration, and Air-Conditioning Systems*. DOE/CE/40568-T1 (DE86014306).
2. Schmutzer, M.A. and Hull, G.W. (1969). *Water Residues After Evaporation*. Report # 28258. United States Testing Company, Inc.
3. Kuivinen, D.E. (1975). *Dynamic Corrosion Test System*. Report # 70112. National Aeronautics and Space Administration, Lewis Research Center.
4. Levinsky, H.L. and Schmutzer, M.A. (1968). *Evaluation Of Steel Wool Corrosion With Magnetic Treated Water versus Untreated Water*. Report # 27421-2. United States Testing Company, Inc.
5. Lin, I.J. (1982). *Faculty of Civil Engineering*, Mineral Engineering Department, Technion Institute of Technology, Haifa, Israel.
6. Young, I.C. (2002). *Efficiency of Physical Water Treatments in Controlling Calcium Scale Accumulation in Recirculating Open Cooling Water System*. Department of Mechanical Engineering, and Mechanics, Drexel University, College of Engineering.
7. Craine, J.E. (1984). Professor Emeritus, Ohio State University, International Foodservice Consultants Society, *Control Lime/Scale Build-Up With Magnetic Water Treatment*, Volume XVII, Number 2.
8. HydraMaster Corporation, Manufacturers of Truck Mounted Steam Machine Carpet Cleaning Systems, *Instructions How To Retrofit Magnetic Water Conditioners* (1998).
9. Evers, D. (1998). *Technology for Improving Energy Efficiency Through the Removal or Prevention of Scale. DOE/EE-0162* Battelle Columbus Operations, United States Department of Energy—Federal Technology Alert.
10. Ferlin, J. (1982). *Magnetic Water Conditioning Will Do Many Thing For The Foodservice Operator*. President, Ferlin and Hopkins, USPWCA National Seminar Presentation.
11. Kappus, J. (1997). *Taylor Freezer Food Service*, Henny Penny Steamers.
12. ASHRAE Research Project 1155 TRP (2000).

CONSUMER CONFIDENCE REPORTS

Susan L. Franklin
Tetra Tech MPS
Ann Arbor, Michigan

On August 19, 1998, the U.S. Environmental Protection Agency (USEPA) promulgated a final rule (40 CFR Parts 141 and 142) requiring community water systems to prepare and provide annual water quality reports to their customers. Mandated by the 1996 amendments to the Safe Drinking Water Act, these "Consumer Confidence Reports" (CCRs) provide the missing link for consumers to have easy access to valuable water quality information. Just as food products have labeling requirements, CCRs are the public's "right-to-know" resource.

The federal rule outlines a number of required elements, including information on the source(s) of drinking water and any source water assessment that has been completed, plainly worded definitions of industry terminology, information on all regulated contaminants detected in the water, health effects language, notice for non-English-speaking residents, and specific language on vulnerable subpopulations and the reasonably expected presence of some contaminants. The CCR also gives the public direct access to their water system professionals by requiring the inclusion of information on public participation opportunities, local contacts, and additional water-related resources, such as USEPA's toll-free Safe Drinking Water Hotline. In addition to the federal requirements, some states mandate supplementary guidelines.

The first CCRs, due to customers on October 19, 1999, covered test results from January 1, 1998—December 31, 1998. Subsequent annual reports are due to consumers by July 1 each year and include data from the prior calendar year. Detections of contaminants that are tested for less frequently should also be included; however, no data more than 5 years old are permitted.

All community water systems that have at least 15 service connections serving residents year-round must prepare a CCR. Water wholesalers must provide monitoring data and other required information to their retail customers by April 1 of each year. The retail system is responsible for ensuring that its customers receive a CCR containing all required content, whether they produce the report themselves or contract with their wholesaler to provide the report. After completing the CCR, each utility must certify to the state that they have complied with the CCR rule; each state determines the method for certification. A copy of the CCR is shared with the local health department, and the utility must keep the report on file for 5 years.

CCR mailing and distribution requirements are based on the size of the water utility as outlined in the federal guidelines or as modified by the state. The governor of the state has the option to waive the mailing requirements for systems that serve fewer than 10,000 people; however, these systems must still prepare a CCR, inform customers that the report is available upon request, and publish the report in a local newspaper each year. Systems that receive the waiver and serve fewer than 500 customers are not required to publish the report in the newspaper. Systems that serve more than 10,000 people must mail one copy of the report to each customer must and make a "good faith effort" (mailing to all postal patrons in the service area, posting the CCR in public places, delivering multiple copies to apartment buildings or to large employers for distribution) to get reports to non-bill-paying customers. Systems that serve 100,000 or more people must also post their CCR on the Internet.

CCRs are promoted by industry organizations, such as the American Water Works Association (AWWA), as "a timely opportunity for utilities to deliver straight talk to the consumer on quality and other issues affecting drinking water" (1). The AWWA offered training on how to prepare a CCR when the rule was issued, and it commissioned a series of six qualitative focus groups in 1997 to assess public perception and reaction to prototype CCRs in anticipation of the final rule. Their findings suggested that CCRs be called "Water Quality Reports" to let customers decide for themselves whether or not they are confident in the water system; to keep the report basic, simple, and honest; to avoid self-serving or alarming statements; and to gear the materials toward the needs of the system's customers.

According to the USEPA (2), there are more than 54,000 community water systems in the United States. Most Americans (263 million) get their drinking water from one of these community water systems and use about 370 billion gallons of water daily. Approximately $22 billion is spent annually to make tap water fit to drink. CCRs are the first nationwide commitment to educate the public about these important drinking water quality issues.

BIBLIOGRAPHY

1. AWWA. (1998). Preparing the Consumer Confidence Reports—One Day Seminar, August.
2. USEPA Office of Water. (2001). Factoids: Drinking Water and Ground Water Statistics for 2000, EPA 816-K-01-004, June.

WATER CONSERVATION MEASURES

National Drinking Water
Clearinghouse

Water is a finite resource, and in many areas, future water supplies are uncertain. Individuals are usually aware when there is a drought; however, because water is inexpensive, there are often few incentives to reduce water loss. Water has no viable substitutes, and its depletion bodes profound economic and social impacts. Citizens and utilities need to consider water conservation programs.

This fact sheet considers the role of water conservation as an integral part of long-term resource planning. It might be more appropriate to use the term "water demand management." Traditional water supply management seeks to provide all the water the public wants, which, in some sections of the country, translates to a constant search for untapped sources.

WHAT METHODS CONSERVE WATER?

The water demand management methods described in this fact sheet incorporate the methods the August 1998 U.S. Environmental Protection Agency (EPA) *Water Conservation Plan Guidelines* recommend for water systems serving 10,000 or fewer people. EPA's Basic guidelines suggest (1) metering, (2) water accounting and loss control, (3) pricing and costing, and (4) education or information.

EPA's Guidelines are not regulations, but recommendations that suggest 11 different conservation methods. How appropriate and desirable any given method is must, in the end, be accepted by the individual community and utility. Pricing may be the primary way to encourage conservation, however, utilities should not automatically rely on any single method.

METER ALL WATER

Metering is a most important part of water demand management. In fact, unless a utility is 100 percent metered, it is difficult to enforce any conservation program. According to a U.S. Housing and Urban Development document, metered customers use an average of 13–45 percent less water than unmetered customers because they know they must pay for any misuse or negligence. A U.S. General Accounting Office report states that metering also assists in managing the overall water system, since it can help to:

- locate leaks in a utility's distribution system by identifying unaccounted-for blocks of water,
- identify high use customers, who can be given literature on opportunities for conserving, and
- identify areas where use is increasing, which is helpful in planning additions to the distribution system.

Once water meters are installed, equipment begins to deteriorate. Eventually meters will fail to measure flows accurately. The question of how long to leave a meter in service has long troubled the waterworks industry. According to a *Journal of the American Water Works Association* (AWWA) article by Tao and a Community Consultants report, average losses of accuracy, for periods greater than 10 years, range from 0.03–0.9 percent per year. To be fair to both customers and the utility, meters must be maintained at regular intervals.

ACCOUNT FOR WATER, REPAIR LEAKS

The EPA Guidelines recommend that all water systems—even smaller systems—implement a basic system of water accounting. The cost of water leakage can be measured in terms of the operating costs associated with water supply, treatment, and delivery. Water lost produces no revenues for the utility. Repairing larger leaks can be costly, but it also can produce substantial savings in water and expenditures over the long run.

Water accounting is less accurate and useful when a system lacks source and connection metering. Although the system should plan to meter sources, unmetered source water can be estimated by multiplying the pumping rate by the time of operation based on electric meter readings.

A utility may want to consider charging for water previously given away for public use or stepping up efforts to reduce illegal connections and other forms of theft.

Drinking water systems worldwide have begun to implement programs to address the problem of water loss. Utilities can no longer tolerate inefficiencies in water distribution systems and the resulting loss of revenue associated with underground leakage, water theft, and under registration. As pumping, treatment, and operational costs increase, these losses become more and more expensive.

If a utility does what it can to conserve water, customers will tend to be more cooperative in other water conservation programs, many of which require individual efforts. In *Economics of Leak Detection*, Moyer states that of the many options available for conserving water, leak detection is a logical first step. A highly visible leak detection program that identifies and locates water system leakage encourages people to think about water conservation before they are asked to take action to reduce their own water use. When leaks are repaired, water savings result in reduced power costs to deliver water, reduced chemicals to treat water, and reduced costs of wholesale supplies.

According to Le Moigne's technical paper *Using Water Efficiently: Technologies Options*, old and poorly constructed pipelines, inadequate corrosion protection, poorly maintained valves and mechanical damage are major factors contributing to leaks. In addition to loss of water, water leaks reduce pressure in the supply system. Raising pressure to compensate for such losses increases energy consumption and can make leaking worse, as well as causing adverse environmental impacts.

A World Bank technical paper by Okun and Ernst shows that, in general, it is normal to be unable to account for 10–20 percent of water. However a loss of more than 20 percent should raise a red flag. It should be noted that percentages are great for guidelines, but volume of water lost is probably more meaningful. According to AWWA's *Leak Detection and Water Loss Reduction*, once a utility knows the volume of water lost, it can determine revenue losses and decide the best way to correct the problem.

EPA's Guidelines recommend that each system institute a comprehensive leak detection and repair strategy. This strategy may include regular onsite testing using computer-assisted leak detection equipment, a sonic leak-detection survey, or another acceptable method for detecting leaks along water distribution mains, valves, services, and meters. Divers can inspect and clean storage tank interiors.

Increasingly, water systems are using remote sensor and telemetry technologies for ongoing monitoring and analysis of source, transmission, and distribution facilities. Remote sensors and monitoring software can alert operators to leaks, fluctuations in pressure, problems with equipment integrity, and other concerns.

Each system should institute a loss-prevention program, which may include pipe inspection, cleaning, lining, and other maintenance efforts to improve the distribution system and prevent leaks and ruptures. Whenever possible, utilities might also consider methods for minimizing water used in routine water system maintenance procedures.

COSTING AND PRICING

In a Journal of the *American Water Works Association* article "Long-Term Options for Municipal Water Conservation," Grisham and Fleming stress that water rates should reflect the real cost of water. Most water rates are based only on a portion of what it costs to obtain, develop, transport, treat, and deliver water to the consumer. Experts recommend that rates include not only current costs but those necessary for future water supply development. Only when rates include all costs can water users understand the real cost of water service and consequently, the need to conserve.

When utilities raise water rates, among other factors, they need to consider what members of the community can afford. According to Schiffler, the ability to pay for water depends on a number of variables, including its intended use. In households, the assumption is that if the share of water costs does not exceed 5 percent of total household revenue it can be considered as socially acceptable. This rule of thumb has no specific foundation, but is widely used.

Many utility managers argue, correctly, that an effective water conservation program will necessitate rate increases. In *Water Conservation*, Maddaus states that a reduction in water use by customers in response to a water conservation program can decrease a water utility's revenues, and the utility may need to re-examine the water rate structure needs and possibly raise rates to compensate for this effect.

Water charges have typically been looked at as a way of financing the operation and maintenance (O&M) costs of a water agency, rather than as a demand management measure to encourage water-use efficiency. As a World Bank document states, political objections and constraints to increasing water charges are often seen as insurmountable. However, low water charges encourage consumption and waste and can put pressure on O&M budgets, leading to poor water treatment and deterioration in water quality.

In *Water Strategies for the Next Century*, Rogers et al. advocate a positive price for water that is less than the cost of desalination, but not zero. Desalination presently costs about $2 a cubic meter. The ideal is to charge a reasonable amount that sends the message to the users.

EPA suggests that systems consider whether their current rate structures promote water usage over conservation. Nonpromotional rates should be implemented whenever possible.

Systems that want to encourage conservation through their rates should consider various issues, such as the allocation between fixed and variable charges, usage blocks and breakpoints, minimum bills and whether water is provided in the minimum bill, seasonal pricing options, and pricing by customer class.

Numerous sources recommend tying sewer prices to water prices. Billing for wastewater is not included in this analysis; however, it is expected to become a more significant motivation for reducing water use over the next 15 years.

INFORMATION AND EDUCATION

According to Maddaus, water conservation initiatives are more likely to succeed if they are socially acceptable. Measuring social acceptability, an exercise in anticipating public response to a potential water conservation measure, may be measured with a two-part survey technique. First, conduct interviews with community leaders to assess the political and social atmosphere. Second, assess the response to selected specific measures via a questionnaire mailed to a random sample of water customers.

The public tends to accept lawn watering restrictions, education, home water-saver kits, low-flush toilet rebates, and a low-flow fixtures ordinance for new construction. Overall acceptance of conservation is strongly related to attitudes about the importance of water conservation, as well as to age, income, and type of residence.

Howe and Dixon note that, “Public participation is now widely understood to be a necessary input for both efficiency and equity.” Public participation should be part of any long-term public education program, as well as an element of plan development. A plan responsive to public needs usually receives continuing support.

The EPA Guidelines state that water systems should be prepared to provide information pamphlets to customers on request. Consumers are often willing to participate in sound water management practices if provided with accurate information. An information and education program should explain to water users all of the costs involved in supplying drinking water and demonstrate how water conservation practices will provide water users with long term savings.

An informative water bill goes beyond the basic information used to calculate the bill based on usage and rates. Comparisons to previous bills and tips on water conservation can help consumers make informed choices about water use. Systems can include inserts in their customers’ water bills that provide information on water use and costs or tips for home water conservation.

School programs can be a great way to get information out. Systems can provide information on water conservation and encourage the use of water conservation practices through a variety of school programs. Contacts through schools can help socialize young people about the value of water and conservation techniques, as well as help systems communicate with parents.

Workshops and seminars can be used to solicit input, and water equipment manufacturers can be invited to these sessions to exhibit their equipment. Maddaus suggests that a number of groups may have a role in water conservation planning:

- Elected officials from all jurisdictions immediately affected by the process;
- Staff persons from private water companies, key personnel from local government agencies, and state agency people;
- Representatives of major local economic interest groups—major industries, chambers of commerce, builders’ associations, farm bureaus, boards of realtors, and landscape contractors;
- Representatives of major community forces, such as federated civic associations, neighborhood associations, school boards, local unions, churches, and local press and media owners;
- Representatives of local government interest groups;
- Local professionals, such as economists and engineers; and
- Representatives of major water users, for example, food processing plants and homeowners’ associations.

WHERE CAN I FIND INFORMATION?

Information in this fact sheet was obtained from the following sources:

American Water Works Association. 1986. *Leak Detection and Water Loss Reduction*. Distribution System Symposium Proceedings, Minneapolis, MN.

American Water Works Association Leak Detection and Water Accountability Committee, 1996. “Committee Report: Water Accountability.” *Journal of American Water Works Association*.

American Water Works Association. 1992. “Alternative Rates.” *Manual of Water Supply Practices*, Manual No. 34. Denver, CO.

Baumann D., J. Boland, and M. Hanemann. 1998. *Urban Water Demand Management and Planning*: McGraw Hill.

Community Consultants. 1986. “Testing of Residential Meters.” *Consultants Report for the City of Tempe*, Arizona.

Grisham, A., and M. Fleming. 1989. “Long-Term Options for Municipal Water Conservation.” *Journal of American Water Works Association*.

Howe, C. and J. Dixon. 1993. “Inefficiencies in Water Project Design and Operation in the Third World: An Economic Perspective.” *Water Resources Research* 29: 1889–1894.

Le Moigne, G., U. Kuffner, M. Xie, et. al. 1993. *“Using Water Efficiently: Technological Options.”* Technical Paper 205, World Bank.

Maddaus, W. 1987. Water Conservation. Denver: American Water Works Association.

Moyer, E. E. 1985. *Economics of Leak Detection: A Case Study Approach*. Denver: American Water Works Association.

Okun, D. A., W. Ernst. 1987. *Community Piped Water Supply Systems in Developing Countries: A Planning Manual*. World Bank Technical Paper 60.

Rogers, P., K. Frederick, G. Le Moigne, D. Seckler, and J. Keller. 1994. *Water Strategies for the Next Century: Supply Augmentation vs. Demand Management*. A debate sponsored by the U.S.

Agency for International Development and ISPAN. Washington, DC.: U.S. Department of State.

Schiffler, M. 1995. “Sustainable Development of Water Resources in Jordan: Ecological and Economic Aspects in a Long-Term Perspective.” in J. A. Allan and C. Mallat, eds. *Water in the Middle East: Legal, Political and Commercial Implications*. New York: I.B. Tauris Publishers.

Tao, P., 1982. “Statistical Sampling Technique for Controlling the Accuracy of Small Water Meters.” *Journal of American Water Works Association*.

U.S. Department of Housing and Urban Development. 1984. *Residential Water Conservation Projects-Summary Report*. Report No. HUD-PDR-903, Prepared by Brown and Caldwell Consulting Engineers for the Office of Policy Development and Research, Washington, DC.

U.S. Environmental Protection Agency. 1998. *Water Conservation Plan Guidelines*. EPA-832-D-001. Office of Water. Washington, DC.

U.S. General Accounting Office. 1978. *Municipal and Industrial Water Conservation—The Federal Government Could Do More*. Report CED-78-66; B-114885. Report to the Congress of the U.S. by the Comptroller General, Washington, DC.

Vickers, A. 1990. “Water Use Efficiency Standards for Plumbing Fixtures: Benefits of National Legislation.” Journal of American Water Works Association 82: 53.

World Bank. 1994. A Strategy for Managing Water in the Middle East and North Africa. Washington, DC.

For further information or comments about this fact sheet, call the National Drinking Water Clearinghouse (NDWC) at (800) 624-8301 or (304) 293-4191. Additional copies of the Water Conservation Measures fact sheets are free; however, postal charges are added to orders. To order, call one of the above numbers. You may also order online at ndwc_orders@estd.wvu.edu, *or download it from our Web site at* http://www.ndwc.wvu.edu *where the fact sheet is available in the Products section.*

PREVENTING WELL CONTAMINATION

VIPIN BHARDWAJ
NDWC Engineering Scientist

Nearly 80% of communities rely on groundwater as their primary drinking water source. Wells extract groundwater for use in homes and businesses. In addition, about 42 million Americans use private wells for drinking water. In light of this information, preventing groundwater contamination is of utmost importance, especially since a number of factors can contribute to groundwater contamination. To prevent well contamination, one of the first steps is to construct it properly. This Tech Brief presents tips about how to site a well and includes information about design issues; material selection and location, such as screens and filter pack; appropriate well sealing methods; and the use of pitless adaptors to prevent contamination

INTRODUCTION

To prevent well contamination, one of the first steps is to construct it properly. This tech brief presents tips on siting a well, its design, choosing proper materials, proper location of screens, filter pack and appropriate method of sealing a well and use of pitless adaptors to prevent contamination.

SITE SELECTION

To prevent groundwater contamination, the first step is to locate the well so that surface water and contaminants cannot flow into it. Site engineers try to install the well uphill from any potential contamination source. This means avoiding potential pollution sources, such as industrial plants, home septic systems, landfills, and underground storage tanks. Hiring a qualified hydrogeologist to investigate potential contaminant sources and likely subsurface conditions makes locating a well easier.

For most private wells, the primary contaminant source is the owner’s septic system. The best protection practice is to locate the well above the area where contaminants can enter it, usually about 50 to 100 feet away. In addition, install a surface seal into a fine-grained layer or non-fractured zone above the aquifer.

To prevent water from collecting near the casing, the ground surrounding the well should slope away from the wellhead on all sides. In addition, most states regulate how far a well must be located from potential contamination sources. For instance, most states require that wells be a minimum of 50 feet away from a septic system.

WELL DESIGN

Proper selection of well casings, seals, screens, filter packs, and pump chamber casings are important factors that determine the efficiency of the well and prevent contamination. Figure 1 shows the components of a well that prevent pollutants from entering the well. Most states have well construction standards and permitting processes that must be followed. The American Water Works Association has a standard A100-90 that deals with construction design.

CASING

A casing is a pipe that is usually made of steel or plastic. It lines the borehole dug in the earth and keeps the well from caving in and prevents runoff and other material from getting into the well.

When contactors select casing, they must take into account the forces that are exerted while installing. In addition, the surrounding materials, such as soil and rocks, tend to collapse into the hole. If possible, the driller should

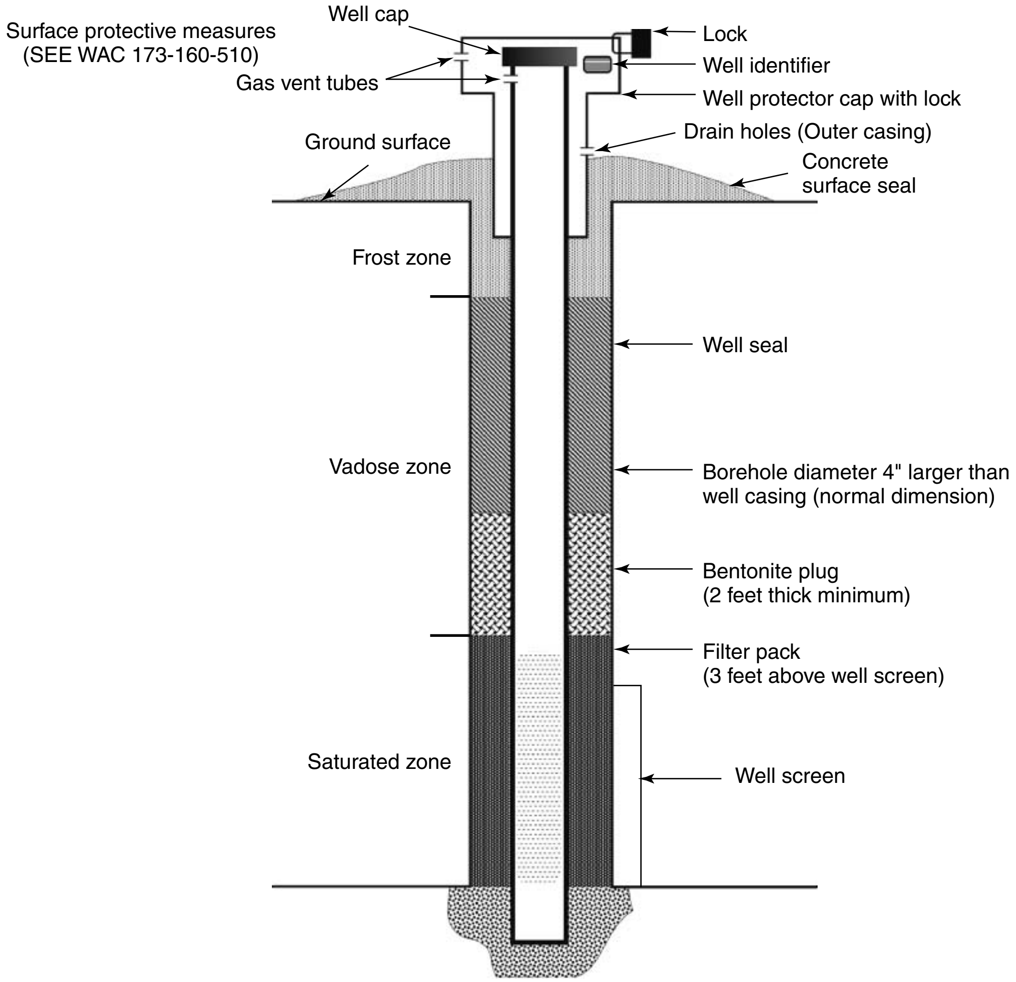

Figure 1. General resource protection well—cross section

Definitions

Well Seal: A seal is a cylindrical layer of material, usually cement, bentonite, or clay, that surrounds the casing up to a certain depth in the well. It prevents runoff or other contaminants from entering the well, and serves to further protect the casing.

Well Screen: Well screen is a cylindrical sieve-like structure that serves as the intake portion of the well. It is a metallic pipe that has holes or perforated sections or slotted sections that is placed on the water- carrying zones of the aquifer.

Filter Pack: A filter pack is made up of sand or gravel that is smooth, uniform, clean, well-rounded, and siliceous. It is placed in the annulus of the well between the bore-hole wall and the well screen to prevent formation material from entering the screen.

Vadose Zone: This is the zone that contains water under pressure less than that of the atmospheric pressure. It is the layer of soil between the water table and the ground surface.

Potentiometric Surface: This is an imaginary surface representing the total head of groundwater in a confined aquifer that is defined by a level to which water will rise in the well.

use a temporary casing for the borehole. The temporary casing diameter must be at least four inches larger than the permanent casing to provide sufficient space for a good well seal.

The American Society for Testing and Materials, the American Petroleum Institute, and the American Iron and Steel Institute have specifications for casings. Most state standards require steel casing of a specified wall thickness for wells, whether for a community or private individual.

The diameter of the casing must leave enough room to install the submersible pump and still have space for maintenance. The size of the pump depends upon the desired well yield.

Casing depth also helps prevent well contamination. Logs of any other nearby wells and the local geology can help determine how deep the casing should go. The casing should extend at least 12 inches above the ground for sanitary protection. Reducing the casing's diameter requires a minimum of eight feet of casing overlap. A watertight well cap should be placed on top of the casing. The Water Systems Council (WSC) has standards for well caps and other well components.

WELL SCREEN

A well screen is a cylindrical sieve-like structure that serves as the intake portion of the well. It is a metallic pipe that has holes or perforated sections or slotted sections that is placed on the water carrying zones of the aquifer. Proper selection, design, placement, and development of the screened section are very important and determine the well's efficiency and yield.

Since certain sections of the ground are more porous than others and, hence, carry more water, placing the screens in these sections will yield higher flow rates. By looking at the data collected during drilling, a good well driller can locate and place the screen in the proper zones.

To better understand conditions at the site, use borehole geophysical logs to grasp the subsurface conditions. In addition, visual inspection of the cuttings or samples can show if the layers of earth are sandy, coarse, or clayey. And to help determine well yield, use sieve analysis and hydraulic conductivity tests.

FILTER PACK

A filter pack is typically made up of sand or gravel that is smooth, uniform, clean, well rounded. It is placed in the area between the borehole wall and the well screen to prevent formation material from entering the screen.

To enhance the permeability of the zone surrounding the screen, place a filter pack around it. A good filter pack keeps sediment out and decreases friction losses around the screen and is especially important if the aquifer consists of uniform fine sand. A filter pack allows for larger openings in the screen and improves well yield. To install a filter pack, start from the bottom of the screen, filling in to at least three feet above the top of the screen. Domestic wells do not require a filter pack.

WELL SEALS

The most important components that prevent contaminants from entering the well are well seals. A seal is a cylindrical layer of material, usually cement, bentonite, or clay, that surrounds the casing up to a certain well depth. It prevents runoff or other contaminants from entering the well and serves to further protect the casing. The drilled hole must be four inches larger in diameter than the outer diameter of the casing so that the seal can be placed in the space between casing and the hole.

Well construction standards specify the material that well installers must use to seal the well, as well as the depth to which the well is grouted. Typically, public water supply wells are grouted to a depth of 50 feet. A cement slurry is pumped in the ring-shaped space between casing and hole and the well is sealed from the bottom up. Grout is placed using a small diameter pipe called a tremie. A layer of bentonite two feet thick should be placed on top of the filter pack.

PITLESS ADAPTORS

Pitless adapters and pitless units are devices that attach to the well casing below the frost line and provide sanitary connections. They prevent entry of contaminants into the well near the surface. These devices provide access to the well for servicing. The adapter connects the casing with a horizontal line that supplies water through a removable seal joint. This connection allows the drop pipe and pumping equipment in the well to be easily removed for repair or maintenance work without digging the ground around the well.

WSC has performance standards for pitless adapters, pitless units, and watertight well caps. A list of manufacturers that meet those standards can be obtained from the WSC.

DISINFECTION PROCEDURES

Well installers must disinfect all equipment and tools using a chlorine solution before any drilling operation to prevent bacterial contamination. The well must be disinfected after it's completed. Some types of bacteria, such as *E. coli*, are found in soils and can contaminate the well. By dissolving calcium hypochlorite or sodium hypochlorite, installers can make a chlorinated water solution. The strength of the solution can range from 50–200 milligrams per liter of available chlorine.

WHERE CAN I FIND MORE INFORMATION?

American Water Works Association. 1999. *Design and Construction of Water Systems, An AWWA Small System Resource Book, Second Edition*. Denver, Colorado: AWWA.

Driscoll, FG. 1995. *Groundwater and Wells*. St. Paul, MN: U.S. Filter/Johnson Screens.

U.S. Environmental Protection Agency. 1991. *Manual of Individual and Non-Public Water Supply Systems*. Washington, DC: EPA. (Available from the National Drinking Water Clearinghouse, order product #DWBKDM06).

U.S. Environmental Protection Agency. 1975. *Manual of Water Well Construction Practices*. Washington, DC: EPA. (Available from the National Drinking Water Clearinghouse, order product #DWBKDM01).

National Ground Water Association, Westerville, OH: NGWA. (www.ngwa.org)

Water Systems Council, National Programs Office, Washington, DC: WSC. (www.wellcarehotline.org)

Acknowledgments

The author wishes to thank **Dr. Dale Ralston**, Ralston Hydrogeologic Services, for his comments and corrections to this *Tech Brief*.

NDWC Engineering Scientist **Vipin Bhardwaj** has a BS in Chemical Engineering, and Masters degrees in Environmental engineering and agriculture from West Virginia University.

HAVE YOU READ ALL OF OUR TECH BRIEFS?

Tech Briefs, drinking water treatment and supply fact sheets, have been a regular feature in the National Drinking Water Clearinghouse (NDWC) publication *On Tap* for more than seven years.

A package of *Tech Briefs* is now available as a product. A three-ring binder holds all the current *Tech Briefs* in print. New selections can easily be added to the package as they become available. To order this free product, call the NDWC at (800) 624-8301 or (304) 293-4191 and ask for item #DWPKPE71.

(Additional copies of fact sheets are free; however, postal charges may be added.) You can also order copies of one or all of the free *Tech Briefs* listed below.

Tech Brief: **Organics Removal**, item #DWBLPE59;

Tech Brief: **Package Plants**, item #DWBLPE63;

Tech Brief: **Water Treatment Plant Residuals Management**, item #DWBLPE65;

Tech Brief: **Lime Softening**, item #DWBLPE67;

Tech Brief: **Iron and Manganese Removal**, item #DWBLPE70;

Water **Conservation Measures Fact Sheet**, item #DWBLPE74;

Tech Brief: **Membrane Filtration**, item #DWBLPE81;

Tech Brief: **Treatment Technologies for Small Drinking Water Systems**, item #DWPSPE82;

Tech Brief: **Ozone**, item #DWBLPE84;

Tech Brief: **Radionuclides**, item #DWBLPE84;

Tech Brief: **Slow Sand Filtration**, item #DWBLPE99;

Tech Brief: **Ultraviolet Disinfection**, item #DWBLPE101;

Tech Brief: **Leak Detection and Water Loss Control**, item #DWBLPE102;

Tech Brief: **Diatomaceous Earth Filtration for Drinking Water**, item #DWBLPE108.

Tech Brief: **Reservoirs, Towers, and Tanks**, item #DWFSOM15

Tech Brief: **System Control and Data Acquisition (SCADA)**, item #DWFSOM20

Tech Brief: **Valves**, item #DWFSOM21

Tech Brief: **Water Quality in Distribution Systems**, item #DWFSOM25

Tech Brief: **Water Treatment Plant Residuals Management**, item #DWBLPE65

CORROSION CONTROL

National Drinking Water
Clearinghouse

Corrosion occurs because metals tend to oxidize when they come in contact with water, resulting in the formation of stable solids. Corrosion in water distribution systems can impact consumers' health, water treatment costs, and the aesthetics of finished water.

Various methods can be used to diagnose, evaluate, and control corrosion problems. Techniques for controlling it include distribution and plumbing system design considerations, water quality modifications, corrosion inhibitors, cathodic protection, and coatings and linings.

CORROSION CAN CAUSE SYSTEM PROBLEMS

What Problems Does Corrosion Cause?

Corrosion can cause higher costs for a water system due to problems with:

- decreased pumping capacity, caused by narrowed pipe diameters resulting from corrosion deposits;
- decreased water production, caused by corrosion holes in the system, which reduce water pressure and increase the amount of finished water required to deliver a gallon of water to the point of consumption;
- water damage to the system, caused by corrosion-related leaks;
- high replacement frequency of water heaters, radiators, valves, pipes, and meters because of corrosion damage; and
- customer complaints of water color, staining, and taste problems.

How is Corrosion Diagnosed and Evaluated?

The following events and measurements can indicate potential corrosion problems in a water system:

Consumer Complaints: Many times a consumer complaint about the taste or odor of water is the first indication of a corrosion problem. Investigators need to examine the construction materials used in the water distribution system and in the plumbing of the complainants' areas (See Table 1).

Corrosion Indices: Corrosion caused by a lack of calcium carbonate deposition in the system can be estimated using

Table 1. Typical Water Quality Complaints That Might Be Due to Corrosion

Customer Complaint	Possible Cause
Red water or reddish-brown staining of fixtures and laundry	Corrosion of iron pipes or presence of natural iron in raw water
Bluish stains on fixtures	Corrosion of copper lines
Black water	Sulfide corrosion of copper or iron lines or precipitations of natural manganese
Foul taste and/or odors	Byproducts from microbial activity
Loss of pressure	Excessive scaling, tubercle (buildup from pitting corrosion), leak in system from pitting or other type of corrosion
Lack of hot water	Buildup of mineral deposits in hot water system (can be reduced by setting thermostats to under 60 degrees C [140 degrees F])
Short service life of household plumbing	Rapid deterioration of pipes from pitting or other types of corrosion

Source: U.S. Environmental Protection Agency indices derived from common water quality measures. The Langelier Saturation Index (LSI) is the most commonly used measure and is equal to the water pH minus the saturation pH (LSI = pH water − pH saturation). The saturation pH refers to the pH at the water's calcium carbonate saturation point (i.e., the point where calcium carbonate is neither deposited nor dissolved). The saturation pH is dependent upon several factors, such as the water's calcium ion concentration, alkalinity, temperature, pH, and presence of other dissolved solids, such as chlorides and sulfates. A negative LSI value indicates potential corrosion problems.

Sampling and Chemical Analysis: The potential for corrosion can also be assessed by conducting a chemical sampling program. Water with a low pH (less than 6.0) tends to be more corrosive. Higher water temperature and total dissolved solids also can indicate corrosivity.

Pipe Examination: The presence of protective pipe scale (coating) and the condition of pipes' inner surfaces can be assessed by simple observation. Chemical examinations can determine the composition of pipe scale, such as the proportion of calcium carbonate, which shields pipes from dissolved oxygen and thus reduces corrosion.

Can System Design Affect the Potential for Corrosion?

In many cases, corrosion can be reduced by properly selecting distribution and plumbing system materials and by having a good engineering design. For example, water distribution systems designed to operate with lower flow rates will have reduced turbulence and, therefore, decreased erosion of protective layers. In addition, some piping materials are more resistant to corrosion in a specific environment than others. Finally, compatible piping materials should be used throughout the system to avoid electrolytic corrosion.

Other measures that help minimize system corrosion include:

- using only lead-free pipes, fittings, and components;
- selecting an appropriate system shape and geometry to avoid dead ends and stagnant areas;
- avoiding sharp turns and elbows in the distribution and plumbing systems;
- providing adequate drainage (flushing) of the system;
- selecting the appropriate metal thickness of piping, based on system flow and design parameters;
- avoiding the use of site welding without replacing the pipe lining;
- reducing mechanical stresses, such as flexing of pipes and "water hammer" (hydraulic pressure surges);
- avoiding uneven heat distribution in the system by providing adequate coating and insulation of pipes;
- providing easy access for inspection, maintenance, and replacement of system parts; and
- eliminating the grounding of electrical circuits to the system, which increases the potential for corrosion.

How Can System Corrosion be Reduced?

Corrosion in a system can be reduced by changing the water's characteristics, such as adjusting pH and alkalinity; softening the water with lime; and changing the level of dissolved oxygen (although this is not a common method of control). Any corrosion adjustment program should include monitoring. This allows for dosage modification, as water characteristics change over time.

pH Adjustment: Operators can promote the formation of a protective calcium carbonate coating (scale) on the metal surface of plumbing by adjusting pH, alkalinity, and calcium levels. Calcium carbonate scaling occurs when water is oversaturated with calcium carbonate. (Below the saturation point, calcium carbonate will redissolve: at the saturation point, calcium carbonate is neither precipitated nor dissolved. See the section on "corrosion indices,".) The saturation point of any particular water source depends on the concentration of calcium ions, alkalinity, temperature, and pH, and the presence of other dissolved materials, such as phosphates, sulfates, and some trace metals.

It is important to note that pH levels well suited for corrosion control may not be optimal for other water treatment processes, such as coagulation and disinfection. To avoid this conflict, the pH level should be adjusted for corrosion control immediately prior to water distribution, and after the other water treatment requirements have been satisfied.

Lime Softening: Lime softening (which, when soda ash is required in addition to lime, is sometimes known as lime-soda softening) affects lead's solubility by changing the water's pH and carbonate levels. Hydroxide ions are then present, and they decrease metal solubility by promoting the formation of solid basic carbonates that "passivate," or protect, the surface of the pipe.

Using lime softening to adjust pH and alkalinity is an effective method for controlling lead corrosion. However, optimum water quality for corrosion control may not coincide with optimum reduction of water hardness. Therefore, to achieve sound, comprehensive water treatment, an operator must balance water hardness, carbonate levels, pH and alkalinity, as well as the potential for corrosion.

Dissolved Oxygen Levels: The presence of excessive dissolved oxygen increases water's corrosive activity. The optimal level of dissolved oxygen for corrosion control is 0.5 to 2.0 parts per million. However, removing oxygen from water is not practical because of the expense. Therefore, the most reasonable strategy to minimize the presence of oxygen is to:

- exclude the aeration process in the treatment of groundwater,
- increase lime softening,
- extend the detention periods for treated water in reservoirs, and
- use the correct size water pumps in the treatment plant to minimize the introduction of air during pumping.

What About the Use of Corrosion Inhibitors?

Corrosion inhibitors cause protective coatings to form on pipes. Although they reduce corrosion, they may not totally arrest it. Therefore, the success of any corrosion inhibitor hinges upon the water operator's ability to:

- apply double and triple the design doses of inhibitor during initial applications to build a protective base coat that will prevent pitting; (Note that initial coatings typically take several weeks to form.)
- maintain continuous and sufficiently high inhibitor doses to prevent redissolving of the protective layer; and
- attain a steady water flow over the system's metal surfaces to allow a continuous application of the inhibitor.

There are several commercially available corrosion inhibitors that can be applied with normal chemical feed systems. Among the most commonly used for potable water supplies are inorganic phosphates, sodium silicates, and mixtures of phosphates and silicates.

Inorganic Phosphates: Inorganic phosphate corrosion inhibitors include polyphosphates, orthophosphates, glassy phosphates, and bimetallic phosphates. Zinc, added in conjunction with polyphosphates, orthophosphates, or glassy phosphates, may help to inhibit corrosion in some cases.

Silicates: The effectiveness of sodium silicates depends on both pH and carbonate concentrations. Sodium silicates are particularly effective for systems with high water velocities, low hardness, low alkalinity, and pH of less than 8.4. Typical coating maintenance doses of sodium silicate range from 2 to 12 milligrams per liter. They offer advantages in hot-water systems because of their chemical stability, unlike many phosphates.

Before installing any technology for delivering corrosion inhibitors, several methods or agents first should be tested in a laboratory environment to determine the best inhibitor and concentration for each water system.

Is Cathodic Protection an Option?

Cathodic protection is an electrical method for preventing corrosion of metallic structures. However, this expensive corrosion control method is not practical or effective for protecting entire water systems. It is used primarily to protect water storage tanks. A limitation of cathodic protection is that it is almost impossible for cathodic protection to reach down into holes, crevices, or internal corners.

Metallic corrosion occurs when contact between a metal and an electrically conductive solution produces a flow of electrons (or current) from the metal to the solution. The electrons given up by the metal cause the metal to corrode rather than remain in its pure metallic form. Cathodic protection stops this current by overpowering it with a stronger, external power source. The electrons provided by the external power source prevent the metal from losing electrons, forcing it to be a "cathode," which will then resist corrosion, as opposed to an "anode," which will not.

There are two basic methods of applying cathodic protection. One method uses inert electrodes, such as high-silicon cast iron or graphite, which are powered by an external source of direct current. The current impressed on the inert electrodes forces them to act as anodes, thus minimizing the possibility that the metal surface being protected will likewise become an anode and corrode. The second method uses a sacrificial anode. Magnesium or zinc anodes produce a galvanic action with iron, so that the anodes are sacrificed (or suffer corrosion), while the iron structure they are connected to is protected.

Are Commercial Pipe Coatings and Linings Effective?

The nearly universal method of reducing pipe corrosion involves lining the pipe walls with a protective coating. These linings are usually mechanically applied, either when the pipe is manufactured or in the field before it is installed. Some linings can be applied even after the pipe is in service, but this method is much more expensive.

Mechanically applied coatings and linings differ for pipes and water storage tanks. The most common types of pipe linings include coal-tar enamels, epoxy paints, cement mortar, and polyethylene.

Water storage tanks are most commonly lined to protect the inner tank walls from corrosion. The most common types of water storage tank coatings and linings include coal-tar paints and enamels, vinyls, and epoxy.

Where Can I Find More Information?

Information for this fact sheet was obtained from three primary sources: *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities, EPA/625/4-89/023; Corrosion Manual for Internal Corrosion of Water Distribution Systems, EPA/570/9-84/001; and Corrosion in Potable Water Supplies*, EPA/570/9-83/013. All of these documents are free and may be ordered from the U.S. Environmental Protection Agency (EPA) Office of Research and Development by calling (513) 569-7562.

If these publications are no longer available from the EPA, call the National Drinking Water Clearinghouse (NDWC) at (800) 624-8301. A photocopied version of the 209-page document Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities, item #DWBKDM04, costs $30.05. There is no charge for the other two documents listed above; however, postage charges apply to all orders.

Also, the NDWC's Registry of Equipment Suppliers of Treatment Technologies for Small Systems (RESULTS), version 2.0, is a public reference database that contains information about technologies—including those related to corrosion—in use at small water systems around the country. For further information about accessing or ordering RESULTS, call the NDWC at (800) 624-8301 or (304) 293-4191. You may also obtain more information from the NDWC's World Wide Web site at www.ndwc.wvu.edu.

For additional copies of "Tech Brief: Corrosion Control," item #DWBRPE52, or for a copy of the previously published "Tech Brief: Filtration," item #DWBRPE50, or "Tech Brief: Disinfection," item #DWBRPE47, call the NDWC at the number printed above.

CROSS CONNECTION AND BACKFLOW PREVENTION

Vipin Bhardwaj
NDWC Engineering Scientist

Bethany Reed
NESC Graphic Designer

When drinking water is transported to a consumer, it is possible for contaminants to be introduced in the distribution system. This situation may occur due to connections between potable water lines and non-potable water sources or by a water flow reversal, resulting in contaminated water. This Tech Brief, discusses cross connections and backflow, and explores ways to prevent these situations.

WHAT IS A CROSS CONNECTION?

A cross connection is a link or structural arrangement where potable water in a distribution system can be exposed to unwanted contaminants. It is the point at which it is possible for a non-potable substance to come in contact with the drinking water system. Cross connections are generally unintentional and can happen anywhere pipes supply water.

WHAT IS A BACKFLOW?

Backflow is the reverse flow of undesirable materials and contaminants into the water mains. Backflow can happen because of two conditions: backpressure and backsiphonage.

Backpressure occurs when pressure in a pipe connected to a main pipe in the distribution system becomes greater than the pressure in the main pipe itself. When this happens, a net force acts on the volume of liquid in the connecting pipe, allowing unwanted material to enter the main pipe.

Backsiphonage refers to a situation where the pressure in a service pipe is less than the atmospheric pressure. If water in a supply line is turned off, such as when a pump fails, backsiphonage can cause contamination to be sucked into the system due to a vacuum in the service line.

If a cross connection exists in a system, it does not mean that there will be a backflow every time. But, where cross connections exist, there is always the possibility.

BACKFLOW CONTROL METHODS AND DEVICES

If possible, cross connections must first be eliminated before installing any backflow prevention devices. The device chosen depends on the degree of hazard involved, accessibility to the location of the device, and whether the backflow is due to backpressure or backsiphonage. Basic types of backflow prevention devices are:

- air gaps,
- reduced pressure principle devices,
- double check valves,
- vacuum breakers, and
- barometric loops

Air Gap

Air gaps are one of the most effective ways to prevent backflow and backsiphonage. An air gap is a vertical separation between a water outlet and the highest level of a potential fluid contamination source. However, because of air gaps, flow of water is interrupted and loss of pressure occurs. Because of this, air gaps are used at the end of a pipe. Air gaps should be twice the size of the supply pipe diameter or at least one inch in length, whichever is greater (see Fig. 1).

Reduced Pressure Principle Backflow Preventer

The reduced pressure zone backflow-preventing (RPBP) device has two spring check valves with a pressure-relief valve located between them that can be vented to the atmosphere.

During normal flow of water through this arrangement, the water flows through the two valves (see Fig. 2). The spring action of the first valve opposes the pressure of water as the water flows from left to right and enters the central chamber. Pressure in the central chamber is maintained lower than that in the incoming line by the operation of the relief valve.

The second check valve to the right is designed to open with a pressure drop of one pound per square inch (psi) in the direction of flow and is independent of the pressure required to open the relief valve. If the pressure downstream from the device increases for some reason (backpressure), the second check valve will close because of the spring action. Reverse flow of water or backflow is thus prevented. In case the pressure in the supply line on the left decreases abruptly or if there is a vacuum in the supply line (backsiphonage), the check valves close because of spring action, and backflow is prevented.

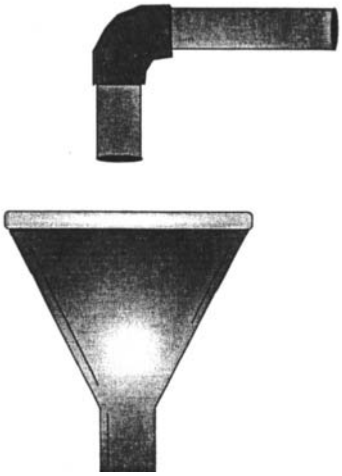

Figure 1. Air gap.

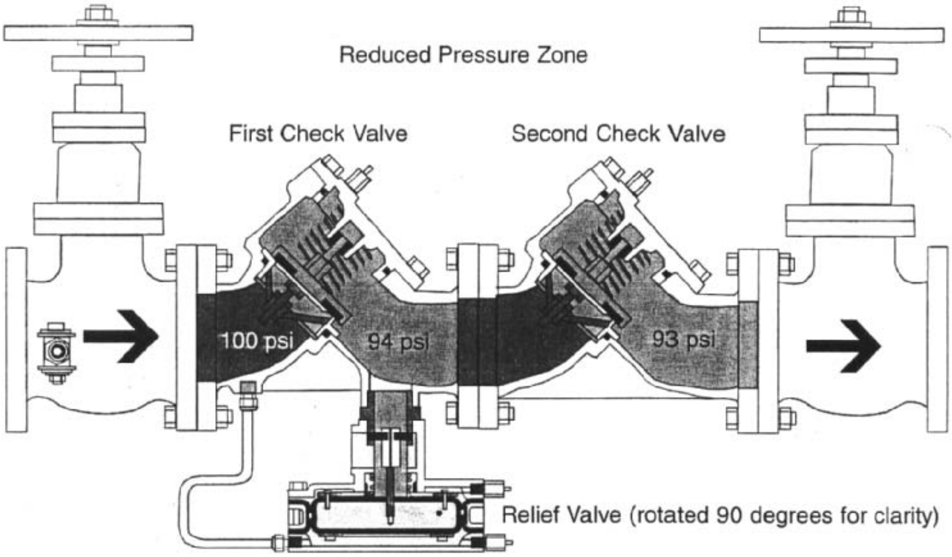

Figure 2. Reduced pressure principle backflow preventer. *Source*: U.S. Environmental Protection Agency. 2003. *Cross-Connection Control Manual*, Washington DC: EPA.

In case the valve leaks and the second check valve on the right does not close fully, water will leak back into the central chamber and increase the pressure in the chamber. The relief valve then opens and discharges water to the atmosphere. Keep in mind, if you see water coming from the relief valve on an RPBP, don't panic. It's actually working as designed. The valve assembly should be checked by a certified tester.

This type of device is usually installed on high hazard locations, such as hospitals, plating plants, and car washes.

Double Check Valve

The double check valve has two single check valves coupled within one body, and has test cocks (to determine if there's any leakage) and two closing gate valves (to isolate each section). It is essentially the same reduced pressure zone backflow-preventing (RPBP) device but without the relief valve (see Fig. 3).

The absence of the relief valve reduces the effectiveness of the device. Double check valves are used in low- to medium-level hazard installations. The check valves are spring-loaded and require one pound of pressure to open.

Vacuum Breakers

Vacuum breakers provide protection against back-siphonage. When the pressure in a service pipe is less than the atmospheric pressure, a vacuum can form in the pipe and cause contamination to be sucked into the system. Vacuum breakers have an element, such as a check valve, that glides on a supporting shaft and seals in the

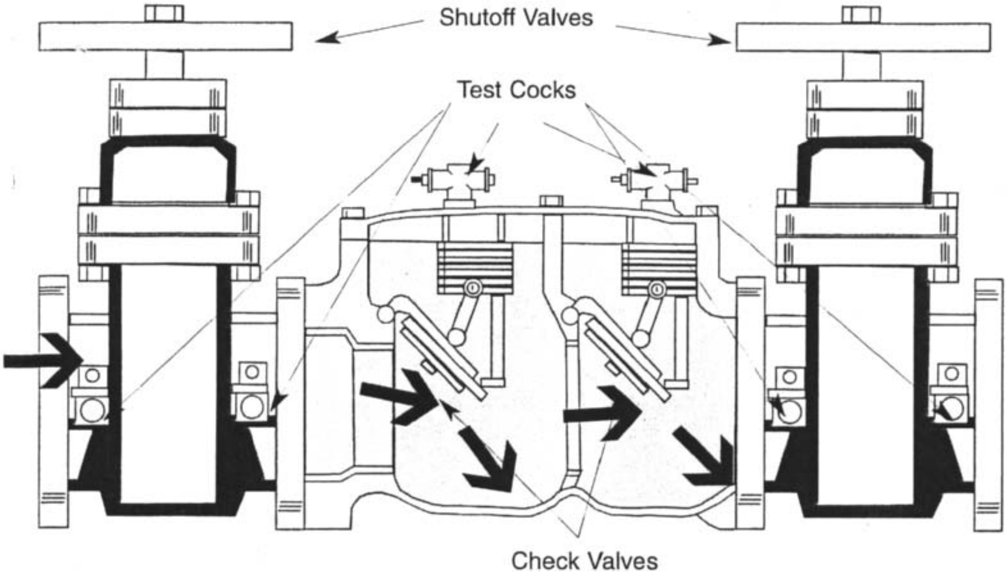

Figure 3. Double check valve. *Source*: AWWA, *Water Transmission and Distribution*, Second Edition.

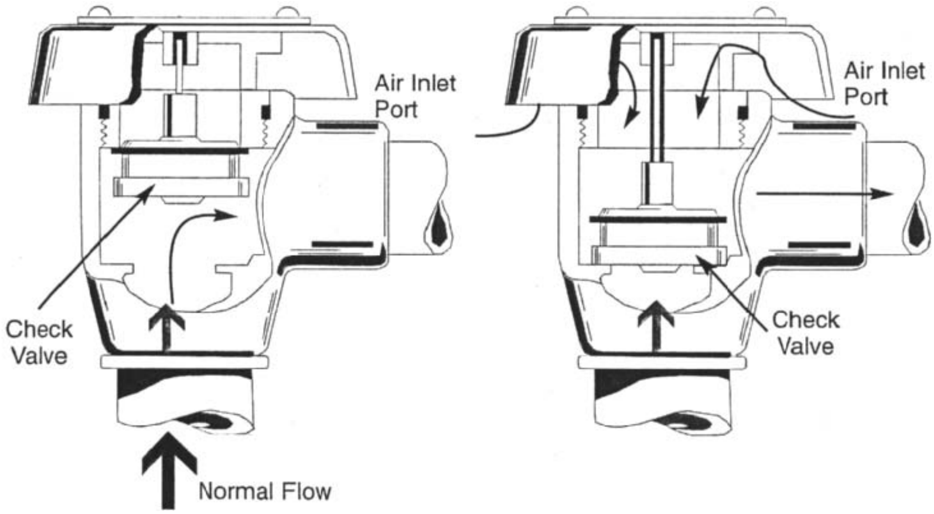

Figure 4. Vacuum breakers. *Source*: AWWA, *Water Transmission and Distribution*, Second Edition.

uppermost position from the push of water pressure (see Fig. 4).

If the flow in the pipe is stopped, the valve drops down, closes the water supply entry, and opens an air vent. This opens up the downstream piping to atmospheric pressure and prevents backsiphonage. Vacuum breakers do not protect against backpressure.

Barometric Loop

A barometric loop is formed by having a section of the pipe in the shape of an inverted "U" upstream of a cross connection (see Fig. 5). Based on a physics principle, the height of a water column open to the atmosphere at the bottom will not be greater than 33.9 feet at sea level pressure.

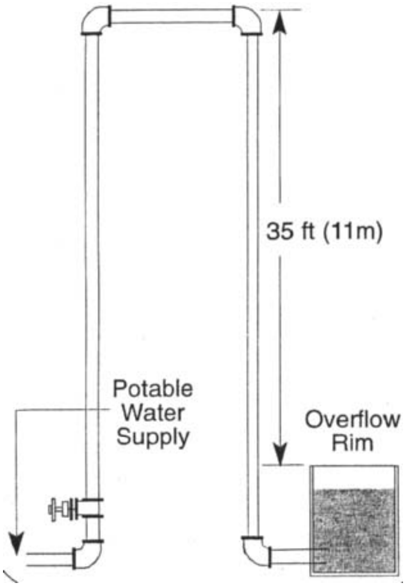

Figure 5. Barometric loop. *Source*: AWWA, *Water Transmission and Distribution*, Second Edition.

If the loop is greater than that height, backsiphonage cannot occur through it. However, the barometric loop is not effective against backpressure.

"YELLOW GUSHY STUFF" FROM FAUCETS

A small town in Maryland provides a dramatic example of what can happen because of a cross connection. One fateful day, yellow gushy stuff poured from some faucets there and the state warned residents from using the water for cooking, drinking, or bathing. The incident drew widespread attention and was a big story in the local media.

An investigation revealed that water pressure in the town mains was reduced temporarily due to a water pump failure in the distribution system. A gate valve between a herbicide chemical holding tank and the town water supply mains had been left open. A lethal cross connection was created that allowed the herbicide to flow into the water supply. When the normal water supply pressure returned, water containing the herbicide was pumped into many faucets.

Door to door public notification, extensive flushing, sampling, and other measures were taken to get the situation under control. Fortunately, no one was seriously injured in this incident.

Source: U.S. Environmental Protection Agency. 2003. *Cross-Connection Control Manual*. Washington DC: EPA.

CROSS CONNECTION CONTROL PROGRAMS

Numerous, well-documented cases about illnesses and other hazards posed by cross connections have been documented. (See the sidebar above.) More information about the health risks cross connections may present and methods to prevent them is needed. Water utility personnel (managers, operators, local officials), plumbers, public health officials, and consumers need to be aware of the risks and understand prevention methods.

As mandated by the Safe Drinking Water Act, water suppliers are responsible for ensuring that the water they supply meets federal primary drinking water regulations and is delivered to consumers without compromising water quality due to its distribution system. Water utilities may want to implement a cross connection program to stave off any problems that could occur. Such a program would include informing consumers, conducting inspections of the distribution system, analyzing and recognizing potential cross connections, and installing backflow prevention devices where needed.

WHAT ABOUT TERRORISM AND WATER SECURITY?

Concern that U.S. water supplies could be vulnerable to terrorist attacks has increased in the post-9/11 era. Because they are a place where lethal substances could be introduced into the water distribution system, cross connections should be identified and corrected.

The Public Health Security and Bioterrorism Preparedness and Response Act of 2002 (Section 1433) requires systems that serve more than 3,300 people to assess their vulnerability to a terrorist attack. (See the article "Security and Emergency Planning: Community-Wide Efforts Require Preparation" in the Winter 2003 *On Tap*.)

The act states that the vulnerability assessment shall include but not be limited to "a review of pipes and constructed conveyances, water collection, pretreatment, storage and distribution facilities, electronic, computer or other automated systems."

WHERE CAN I FIND MORE INFORMATION?

American Water Works Association. 1996. *Water Transmission and Distribution, Principles and Practices of Water Supply Operations*, 2nd Edn. Denver: AWWA.

U.S. Environmental Protection Agency. 2003. *Cross-Connection Control Manual*. Washington, DC: EPA. (Available as a product from NDWC, item #DWBLDM03)

Montana Water Center. 2002. *Sanitary Survey Fundamentals Preparation Course CD*. Bozeman, MT: Montana Water Center. (Available as a product from NDWC, item #DWCDTR19)

NDWC Engineering Scientist **Vipin Bhardwaj** has a bachelor's degree in chemical engineering, and master's degrees in environmental engineering and agriculture from West Virginia University.

HAVE YOU READ ALL OF OUR TECH BRIEFS?

Tech Briefs, drinking water treatment and supply fact sheets, have been a regular feature in the National Drinking Water Clearinghouse (NDWC) publication *On Tap* for more than seven years.

A free package of *Tech Briefs* is available as a product. A three-ring binder holds all the current *Tech Briefs* in print. New selections can easily be added to the package as they become available. To order, call the NDWC at (800) 624-8301 or (304) 293-4191 and ask for item #DWPKPE71.

(Additional copies of fact sheets are free; however, postal charges may be added.) You can also order copies of one or all of the free *Tech Briefs* listed below.

Tech Brief: **Organics Removal**, item #DWBLPE59
Tech Brief: **Package Plants**, item #DWBLPE63
Tech Brief: **Water Treatment Plant Residuals Management**, item #DWBLPE65
Tech Brief: **Lime Softening**, item #DWBLPE67
Tech Brief: **Iron and Manganese Removal**, item #DWBLPE70
Water **Conservation Measures Fact Sheet**, item #DWBLPE74
Tech Brief: **Membrane Filtration**, item #DWBLPE81
Tech Brief: **Treatment Technologies for Small Drinking Water Systems**, item #DWPSPE82
Tech Brief: **Ozone**, item #DWBLPE84
Tech Brief: **Radionuclides**, item #DWBLPE84
Tech Brief: **Slow Sand Filtration**, item #DWBLPE99
Tech Brief: **Ultraviolet Disinfection**, item #DWBLPE101
Tech Brief: **Leak Detection and Water Loss Control**, item #DWBLPE102
Tech Brief: **Diatomaceous Earth Filtration for Drinking Water**, item #DWBLPE108
Tech Brief: **Reservoirs, Towers, and Tanks**, item #DWFSOM15
Tech Brief: **System Control and Data Acquisition (SCADA)**, item #DWFSOM20
Tech Brief: **Valves**, item #DWFSOM21
Tech Brief: **Water Quality in Distribution Systems**, item #DWFSOM25
Tech Brief: **Water Treatment Plant Residuals Management**, item #DWBLPE65

MOLECULAR-BASED DETECTION OF *CRYPTOSPORIDIUM PARVUM* IN WATER

Nicholas J. Pokorny
Christine M. Carey
Jeanine L. Boulter-Bitzer
Hung Lee
Jack T. Trevors
University of Guelph
Guelph, Ontario, Canada

INTRODUCTION

Cryptosporidium is an apicomplexan protozoan parasite that is becoming increasingly recognized as an important infectious pathogen in water. The parasite is transmitted by the fecal–oral route, usually by ingestion of oocyst-contaminated water or food (1–3). A *C. parvum* oocyst is small, averaging 5 μm × 4.5 μm, and its coat is bilayered and highly proteinaceous, allowing it to resist chlorine-based disinfection in water treatment facilities (4). Within the oocyst are four motile sporozoites (Fig. 1), which are the infectious forms of *Cryptosporidium*. After passing through the stomach, the oocyst releases the infectious sporozoites, which attach to epithelial cells of the small intestine. Asexual and sexual reproduction of *Cryptosporidium* within the infected host results in autoinfection and self-perpetuation of the disease. Sexual reproduction also results in the production of infectious oocysts that are excreted along with fecal matter into the environment.

Cryptosporidiosis, the globally occurring disease caused by *Cryptosporidium* infection, can result in gastrointestinal distress with copious amounts of diarrhea, abdominal cramping, and fever. In healthy individuals, infection is an acute but self-limiting illness that generally lasts 1 to 2 weeks. However, in immunocompromised individuals, symptoms are more severe, and the illness may become chronic and can result in death. It has been recognized that two distinct *C. parvum* genotypes, referred to as the human genotype (genotype I or genotype H) and the cattle genotype (genotype II or genotype C), are responsible for human cryptosporidiosis. Morgan-Ryan et al. (5) proposed a new species, *Cryptosporidium hominis*, to denote the human genotype. *Cryptosporidium hominis* infects humans primarily, whereas *C. parvum* infects both humans and numerous other mammalian hosts such as mice, dogs, cattle, sheep, goats, and pigs. There is no cure therapy currently available for *Cryptosporidium* infection, so prevention is the key to containing and managing this disease. Prevention requires accurate monitoring and detection of the parasites in water and food samples.

As intestinal parasites, *C. parvum* cannot be cultured *in vitro*. Therefore, traditional culture methods used in routine microbiology labs are not suitable to detect and enumerate *C. parvum* oocysts. Classical methods to detect *Cryptosporidium* are based on microscopy of clinical samples that can be combined with different stains for improved visualization. However, microscopy requires training and is subject to human error in both sample preparation and viewing. The use of fluorogenic dyes can make detection easier, but microscopy cannot indicate oocyst viability or infectivity. Rather than relying on the results of clinical samples to identify a *Cryptosporidium* outbreak, recent research has focused on analyzing water samples for oocysts. The U.S. EPA has developed a method that is recommended for detecting *Cryptosporidium* in raw and treated waters. This method requires the concentration of a water sample, immunomagnetic separation (IMS) of the oocysts from the concentrated debris, and determination of oocyst concentrations by immunofluorescence assay (IFA). Potential oocysts are further stained with vital dyes such as DAPI and propidium iodide, followed by differential interference contrast microscopy.

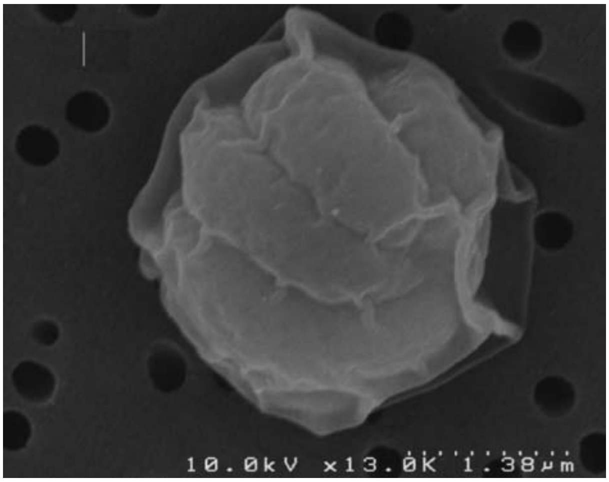

Figure 1. A scanning electron micrograph of a *C. parvum* oocyst. The four crescent-shaped sporozoites are visible within the oocyst. The deflated appearance of the oocyst is the result of dehydration required for sample preparation.

The number of oocysts within water sources is generally low and variable, so large volumes of water need to be concentrated for quantitative analysis. This can be accomplished by filtration, flocculation, or centrifugation. Oocysts are then purified from background detrital material by flotation, immunomagnetic separation, or flow cytometry. The concentrated water samples can be screened by molecular techniques to determine the presence of oocysts. A desired detection method should differentiate between dead or nonviable oocysts, which pose no threat to public health, and those oocysts that are viable and infective. Because only a few oocysts are needed to cause disease, efficient detection methods should be sensitive and amenable to quantification. Additionally, the detection method should differentiate between the *Cryptosporidium* species because not all species are harmful to humans. Molecular detection techniques have the advantage of being more specific, sensitive, reproducible, and often more rapid.

DETECTION OF *CRYPTOSPORIDIUM* INFECTIVITY

Although human infectivity assays are true representations of the disease, they are not practical for routine evaluations of oocyst infectivity. The most direct method for assessing oocyst viability and infectivity is to administer the oocysts experimentally to an animal (typically a neonatal mouse). Following a period of incubation, the intestine is sectioned longitudinally and the cells are examined for histological evidence of infection. Commonly used animal surrogates are neonatal or immunosuppressed rodents, typically CD-1 or BALB/c strains. In addition to ethical concerns, animal infectivity experiments are time-consuming, expensive, and not amenable to routine environmental testing.

Alternatively, *in vitro* infectivity involves exposing oocysts to excystation stimuli followed by inoculating them into a cultured adherent mammalian intestinal cell line, which supports infection and asexual development. Following a suitable infection period, samples are examined by fluorogenically labeled antibodies or nucleic acid sequences specific to *Cryptosporidium*. Based on the presence or absence of infection, the number of infective oocysts can be statistically determined according to a standardized most-probable number (MPN) table or MPN calculator (6).

The success of an *in vitro* cell culture assay depends on the type of cell line, the use of oocysts or sporozoites, preincubation excystation treatment, and centrifugation of the sample. Although cell culture requires intensive labor to grow and maintain the cell monolayer and a prolonged incubation period, the technique remains a good substitute for animal infectivity models. A number of cell lines can support the asexual development of *C. parvum*, but human ileocecal adenocarcinoma (HCT-8) cells that have been found superior in their ability to support parasite growth and represent a practical and accurate alternative to animal infectivity models, yield results similar to those with CD-1 mice in assessing *C. parvum* infectivity (7).

ANTIBODY TECHNOLOGY

Antibodies are produced by immune systems that bind to foreign antigens within the host organism. Strong target affinity, specificity, and sensitivity of antibodies make them ideal for targeting and identifying *Cryptosporidium*. Various types of antibodies can be used for molecular diagnostic methods: monoclonal antibodies (mAbs), polyclonal antibodies (pAbs), and recombinant antibodies (rAbs).

Monoclonal antibodies are identical because they were produced by one type of immune cell, all clones of a single parent cell. In contrast, pAbs are produced by different immune cells and differ from one another. Recombinant antibodies are constructed within a laboratory and are based on the artificial random recombination of antibody genes. This may yield antibodies that may not occur naturally.

Conjugation of antibodies to fluorophores or enzymes allow for the production of a visual signal. Many immunologic methods are available for detecting *C. parvum*, including flow cytometry, immunofluorescence assays (IFA), and enzyme-linked immunosorbent assays (ELISA).

Immunofluorescence Assay (IFA)

Routine detection of *Cryptosporidium* in environmental samples is most commonly performed using the IFA. In detecting the parasite, the characteristics used for identification include size, shape, and fluorescence. The cells become fluorescent via antibody staining, which is carried out either in suspension, on a microscope slide, or on a membrane prior to microscopic examination. Both fluorophore-conjugated pAbs and mAbs are used to identify purified oocysts. Nonspecific fluorescence is a problem with IFAs, along with background fluorescence of detritus, algae, and some freshwater diatoms and cyanobacteria.

Foci Detection Method (FDM)

An *in vitro* cell culture infectivity assay involves exposing oocysts to excystation stimuli followed by inoculating them into a cultured mammalian cell line, which supports infection and asexual development. An infectious focus is the site of infection of at least one sporozoite in the cell culture. After 24–48 h, samples are examined for the presence of infection. Foci are labeled with fluorophore-conjugated antibodies specific to the various life stages of *Cryptosporidium* and are visualized by epifluoresence microscopy.

Enzyme-Linked Immunosorbent Assay (ELISA)

ELISA has been developed to detect and evaluate the growth of *C. parvum* in culture systems using antibodies developed against its various life-cycle stages. ELISA can be used to test multiple samples simultaneously and does not require a high level of technical skill compared with that for identifying parasites based on morphological and staining characteristics by microscopic examination. Tests commonly use 96-well microtitration plates coated by antibodies specific for *Cryptosporidium*. Tests samples are then added to the plate. After washing away nonbinding materials, a second enzyme-conjugated antibody specific to *Cryptosporidium* is added to the sample wells. Addition of the enzyme substrate results in a color reaction to indicate the presence of the parasite. There may be problems of cross-reactivity with algae and turbidity interfering with results, but this can be overcome by the development and use of more specific antibodies.

Flow Cytometry

A flow cytometer analyzes particles in a suspension as they pass by a laser beam. The light scattering pattern is analyzed and correlated to size and internal complexity along with the fluorescent light emitted by each particle. Flow cytometer cell sorters can also sort particles of interest from unwanted particles by using the binding of a fluorescein-conjugated antibody to antigens present on *Cryptosporidium* oocysts. Particles within the sample that do not match the criteria for *Cryptosporidium* are filtered out.

POLYMERASE CHAIN REACTION (PCR) DETECTION METHODS

PCR involves using oligonucleotide primers to amplify a DNA fragment specific to the target organism. The desired fragment is identified by gel electrophoresis and can be subsequently sequenced. PCR is reproducible, cost-effective, sensitive and amenable to quantitation, and capable of differentiating among *Cryptosporidium* species that infect humans. Common gene targets for PCR detection include 18S ribosomal DNA, heat-shock proteins (i.e., *hsp*70), and oocyst wall proteins.

Real-Time PCR

Real-time PCR involves virtual real-time visualization of fluorescence emitted by a fluorogenic probe accumulated during PCR. Quantification of amplified DNA during the exponential phase of the reaction, when reagents are not limiting, allows precise determination of the initial quantity of the target sequences. Real-time PCR offers higher throughput, reduced turnaround time, and minimal amplicon contamination due to a closed-vessel system. The quantification range of real-time PCR methods is greater (5–6 log units) than conventional PCR (2–3 log units).

Cryptosporidium species and genotypes can be differentiated through melt curve analysis, which is based on melting temperature differences of PCR-probe complexes, and reflects the extent of complementation of the probes to the amplified PCR fragments. Real-time PCR investigations for environmental detection have demonstrated that the method is rapid, sensitive, and specific.

Reverse Transcription-PCR (RT-PCR)

Only viable organisms can produce messenger RNA (mRNA), so RT-PCR selectively detects viable organisms. In this method, a specific fragment of the complementary DNA (cDNA), which is produced from an mRNA template by reverse transcriptase, is amplified by PCR. Reverse transcriptase is extremely sensitive to chemical contamination, so the RT step is the source of greatest variability. Common gene targets for RT-PCR analysis of *C. parvum* viability include *hsp70*, amyloglucosidase, and *B*-tubulin.

Cell Culture-PCR (CC-PCR)

Comparable with FDM in detecting oocyst infectivity, a CC-PCR assay involves the extraction of genomic DNA from an infected cell culture and amplification of specific DNA fragments by PCR. A cell culture infectivity assay, coupled with either real-time PCR or an MPN assay, can be used to quantify infective oocysts. Alternatively, mRNA can be extracted from the infected cell culture and subjected to RT-PCR to detect infectious oocysts.

Most Probable Number-PCR (MPN-PCR)

In a standard MPN assay, replicate samples are serially diluted and evaluated by PCR for the presence of *Cryptosporidium*-specific nucleic acid sequences. Used in conjunction with a cell culture infectivity assay, MPN-PCR could provide a good alternative to using FDM alone for enumerating infective oocysts. Because MPN assays are costly, labor-intensive, and require a large volume of samples with multiple preparation steps, alternative detection strategies may be more suitable for quantitative analyses.

NUCLEIC ACID PROBES

Small nucleotide chains (oligonucleotides) can be constructed to a desired DNA sequence. The binding nature of DNA is such that complementary single-stranded DNA (ssDNA) fragments strongly bind to one another, whereas a single nucleotide mismatch drastically reduces the binding strength of the two fragments.

Fluorescent *In Situ* Hybridization (FISH)

FISH uses fluorescently labeled oligonucleotide probes targeted to *C. parvum*-specific sequences, such as ribosomal RNA (rRNA). Ribosomal RNA is used as a target because it should be present in high quantities in viable and potentially infective oocysts, but due to its short half-life, it would be degraded in nonviable oocysts. Oocysts are collected and made permeable to the labeled probes by the addition of organic solvents. Species-specific probes hybridize with target rRNA within the permeabilized oocysts and are detected by epifluorescence microscopy or flow cytometry.

Microarray (DNA Chip)

Microarray technology combines PCR with hybridization of oligonucleotide probes. An array of oligonucleotide probes is synthesized and fixed to a chip. Oocyst genomic DNA is extracted and highly specific, and conserved target genes, such as rRNA and *hsp70*, are amplified by PCR and then fluorescently labeled. For

Table 1. Summary of the Molecular Detection Methods That Can Be Used for Detecting *Cryptosporidium parvum* in Environmental Samples

Detection Method	Differentiate Among *Cryptosporidium* Species	Determine Infectivity	Quantifiable	Detection Limit[a]	Reference
IFA	−	−	+	10 oocysts	(8)
FDM—animal	−	+	+	60 oocysts	(10)
FDM—cell culture	−	+	+	5–10 oocysts	(10)
FDM-MPN	−	+	+	2–10 oocysts	(6,10)
PCR	+	−	−	1–10 oocysts	(18)
Real-time PCR	+	−	+	1–5 oocysts	(11)
Real-time CC-PCR	+	+	+	1 oocyst	(12)
RT-PCR	+	+	−	1 oocyst	(13)
CC-RT-PCR	+	+	−	10 oocysts	(14)
CC-PCR	+	+	−	10 oocysts	(15,16)
MPN-PCR	+	−	+	10 oocysts	(17)
Flow cytometry	−	−	+	10 oocysts	(18)
FISH	+	−	+	100 oocysts	(8)
Microarray	+	−	+	Not available	(19,20)

[a] Detection limits are not directly comparable as the conditions used by different researchers may not be similar.

identification at the species level, highly variable or species-specific genes are used. PCR products hybridize with their complementary probe sequences, and the presence or absence of *Cryptosporidium* can be determined by computerized detection of fluorescent emissions. The high throughout capability of microarrays can allow simultaneous detection of numerous different genera of pathogens or species within a genus.

A complete list of the methods described is presented in Table 1. The table indicates whether each method can distinguish among *Cryptosporidium* species, detect oocyst infectivity, quantify detected oocysts, and has a limitation on detection. A reference for each of the methods is also included in the table. A recent review article by Carey et al. (22) provides a more detailed description of *Cryptosporidium* detection methods.

Acknowledgment
This research was supported by the Canadian Water Network (CWN) Centre of Excellence.

BIBLIOGRAPHY

1. Carreno, R.A. et al. (2001). Phenotypic and genotypic characterization of *Cryptosporidium* species and isolate. *J. Ind. Microbiol. Biotechnol.* **226**: 95–106.
2. Millard, P.S. et al. (1994). An outbreak of cryptosporidiosis from fresh-pressed apple cider. *JAMA* **272**(20): 1592–1596.
3. Smith, H.V. (1998). Detection of parasites in the environment. *Parasitology* **117**(7): S113–S141.
4. Korich, D.G. et al. (1990). Effects of ozone, chlorine dioxide, chlorine and monochloramine on *Cryptosporidium parvum* oocyst viability. *Appl. Environ. Microbiol.* **56**(5): 1423–1428.
5. Morgan-Ryan et al.
6. Weir, S.C. et al. (2001). Improving the rate of infectivity of *Cryptosporidium parvum* oocysts in cell culture using centrifugation. *J. Parasitol.* **87**(6): 1502–1504.
7. Higgins, J.A. et al. (2003). Recovery and detection of *Cryptosporidium parvum* oocysts from water samples using continuous flow centrifugation. *Water Res.* **37**(15): 3551–3560.
8. Rochelle, P.A. et al. (2002). Comparison of *in vitro* cell culture and a mouse assay for measuring infectivity of *Cryptosporidium parvum*. *Appl. Environ. Microbiol.* **68**(8): 3809–3817.
9. U.S. Environmental Protection Agency. (1999). *Method 1622: Cryptosporidium in water by filtration/IMS/FA*. Office of Water EPA-821-R-99-061.
10. Slifko, T.R. et al. (2002). Comparison of tissue culture and animal models for assessment of *Cryptosporidium parvum* infection. *Exp. Parasitol.* **101**(2–3): 97–106.
11. Johnson, D.W. et al. (1995). Development of a PCR protocol for sensitive detection of *Cryptosporidium* oocysts in water samples. *Appl. Environ. Microbiol.* **61**(11): 3849–3855.
12. Limor, J.R., Lal, A.A., and Xiao, L. (2002). Detection and differentiation of *Cryptosporidium* parasites that are pathogenic for humans by real-time PCR. *J. Clin. Microbiol.* **40**(7): 2335–2338.
13. MacDonald, L.M. et al. (2002). The development of a real-time quantitative-PCR method for characterisation of a *Cryptosporidium parvum in vitro* culturing system and assessment of drug efficacy. *Mol. Biochem. Parasitol.* **121**: 279–282.
14. Stinear, T., Matusan, A., Hines, K., and Sandery, M. (1996). Detection of a single viable *Cryptosporidium parvum* oocyst in environmental water concentrates by reverse transcription-PCR. *Appl. Environ. Microbiol.* **62**(9): 3385–3390.
15. Rochelle, P.A., De Leon, R., Stewart, M.H., and Wolfe, R.L. (1997). Comparison of primers and optimization of PCR conditions for detection of *Cryptosporidium parvum* and *Giardia lamblia* in water. *Appl. Environ. Microbiol.* **63**(1): 106–114.
16. Di Giovanni, G.D. et al. (1999). Detection of infectious *Cryptosporidium parvum* oocysts in surface and filter backwash water samples by immunomagnetic separation and integrated cell culture-PCR. *Appl. Environ. Microbiol.* **65**(8): 3427–3432.
17. LeChevallier, M.W., Abbaszadegan, M., and Di Giovanni, G.D. (2000). Detection of infectious *Cryptosporidium parvum* oocysts in environmental water samples using an integrated cell culture-PCR (CC-PCR) system. *Water Air Soil Pollut.* **123**(1/4): 53–65.
18. Tsuchihashi, R., Loge, F.J., and Darby, J.L. (2003). Detection of *Cryptosporidium parvum* in secondary effluents using a most probable number-polymerase chain reaction assay. *Water Environ. Res.* **75**(4): 292–299.
19. Chung, J., Vesey, G., Gauci, M., and Ashbolt, N.J. (2004). Fluorescence resonance energy transfer (FRET)-based specific labeling of *Cryptosporidium* oocysts for detection in environmental samples. *Cytometry* **60**(A): 97–106.
20. Straub, T.M. et al. (2002). Genotyping *Cryptosporidium parvum* with an *hsp70* single-nucleotide polymorphism microarray. *Appl. Environ. Microbiol.* **68**(4): 1817–1826.
21. Wang, Z., Vora, G.J., and Stenger, D.A. (2004). Detection and genotyping of *Entamoeba histolytica, Entamoeba dispar, Giardia lamblia*, and *Cryptosporidium parvum* by oligonucleotide microarray. *J. Clin. Microbiol.* **42**(7): 3262–3271.
22. Carey, C.M., Lee, H., and Trevors, J.T. (2004). Biology, persistence and detection of *Cryptosporidium parvum* and *Cryptosporidium hominis* oocyst. *Water Res.* **38**: 818–862.

CRYPTOSPORIDIUM

Jacqueline Brabants
University of New Hampshire
Durham, New Hampshire

Cryptosporidia are small, spherical, obligate, intracellular sporozoan parasites that infect the intestinal tract of a wide range of mammals, including humans. *Cryptosporidium* oocysts range in size from 2 to 8 μm, depending on the species and the stage of the life cycle. Oocysts are commonly found in many of the lakes and rivers that supply public drinking water as a result of runoff from sewage and animal wastes applied to nearby fields and pastures or from areas of wildlife or livestock activity (4,7,10). Infection, resulting from ingesting oocysts, manifests as cryptosporidiosis. Although *Cryptosporidium* was first described in 1907, human infection was not reported until 1976. Today, *Cryptosporidium parvum* is well recognized as the cause of the disease cryptosporidiosis in humans and cattle and is one of the more opportunistic agents seen in patients with AIDS that occurs primarily in individuals with compromised immune systems. Currently, there is no totally effective therapy for cryptosporidiosis (3,8,11,12).

LIFE CYCLE AND MORPHOLOGY

Cryptosporidia undergo alternating life cycles of sexual and asexual reproduction that are completed within the gastrointestinal tract of a single host. The developmental stages of the life cycle occur intracellularly and extracytoplasmically and include schizogony, gametogony, fertilization, and sporogony (1). The developmental stages of the organism are contained within a host cell parasitophorous vacuole, located at the microvillous surface of the host cell. The cycle begins when infectious oocysts containing four sporozoites are discharged in the feces of a parasitized animal. These thick-walled oocysts remain viable for months unless exposed to extremes of temperature, desiccation, or concentrated disinfectants (1). Following ingestion by another animal, most likely from food or water that has become fecally contaminated, the oocyst excysts and releases sporozoites that attach to the microvilli of the small bowel epithelial cells, where they develop into trophozoites (2). Trophozoites divide asexually (schizogony) to form schizonts that contain eight daughter cells known as type I merozoites. Upon release from the schizont, these cells attach to another epithelial cell, and the schizogony cycle is repeated to produce schizonts that contain four type II merozoites. Type II merozoites develop into male (microgametocyte) and female (macrogametocyte) sexual forms. Fertilization results in a zygote (oocyte) that develops into an oocyst, which is ultimately shed into the lumen of the bowel. The oocysts undergo sporulation to the infective stage within the brush border of the enterocytes and are excreted as infectious oocysts in the stool (3–5).

The majority of oocysts generated possess a thick, protective cell wall that ensures their intact passage in the feces and survival in the environment; however, approximately 20% of the oocysts generated fail to develop the thick wall, and following release from a host cell, the thin cell membrane on these oocysts ruptures and releases four infectious sporozoites. These sporozoites penetrate the intestinal lumen and initiate a new autoinfective cycle within the original host. The presence of this thin-walled, autoinfective oocyst can lead to an overwhelming infection that creates a persistent, life-threatening infection in an immunocompromised individual (4).

CLINICAL DISEASE

The pathogenesis of *Cryptosporidia* is not completely understood; age and immune status at the time of primary exposure do not appear to influence susceptibility to infection. However, once the primary infection has been established, the immune status of the host plays an extremely important role in determining the length and severity of the illness (4).

Cryptosporidia undergo their life cycle in the enteric epithelial cells and also in the gallbladder, respiratory, and renal epithelium, especially in immunocompromised hosts (1). The symptom found in all reported cases is acute diarrhea (2). Clinical symptoms found in immunocompetent individuals include nausea; low-grade fever; abdominal cramps; anorexia; and profuse watery, frothy bowel movements that may be followed by constipation. Other individuals may be asymptomatic, particularly later in the course of the infection. In patients who have the typical watery diarrhea, the stool specimen will contain very little fecal material, consisting mainly of water and mucus flecks, and the organisms are trapped in the mucus. In most cases, a patient who has a normal immune system will have a self-limited infection lasting 1 to 2 weeks, whereas a patient who is immunocompromised may have a chronic infection with symptoms ranging from asymptomatic to severe (2,4). Studies to examine susceptibility and serologic responses to reinfection have demonstrated that previous exposure of immunocompetent adults to *Cryptosporidium* is not entirely protective but may decrease the severity of disease and the number of oocysts shed (6).

Cryptosporidium can generate life-threatening infections in immunocompromised individuals, particularly in human AIDS patients (1). In these individuals, the duration and severity of diarrheal illness will depend on the immune status of the patient. It is believed that cryptosporidiosis in AIDS patients causes malabsorption and intestinal injury in proportion to the number of organisms that infect the intestine. Most severely immunocompromised patients cannot overcome the infection; the illness becomes progressively worse with time and leads to death. The disease is prolonged; profuse, watery diarrhea persists from several weeks to months or years, as a result of the autoinfective nature of the organism, reportedly resulting in fluid losses as high as 25 L/day (7). In such patients, infections in areas other than the gastrointestinal tract may cause additional symptoms such as respiratory problems, cholecystitis, hepatitis, and pancreatitis (4).

Diagnosis

Oocysts recovered in clinical specimens usually represent the 80% that are thick-walled. The oocysts are difficult to visualize because of their small, colorless, transparent appearance, and may be confused with yeast cells. In the past, cryptosporidiosis has been diagnosed following examination of small or large bowel biopsy material, under both light and electron microscopy; however, in *Cryptosporidium* infections, the entire mucosa may not be infected uniformly; therefore, biopsy specimens may miss the infected area. As a result, cases have recently been diagnosed by recovering the oocysts from fecal material using flotation or fecal concentration techniques. Diagnosis is achieved by demonstrating the oocysts in feces, sputum, or possibly respiratory secretions (3). Special staining techniques such as the modified acid-fast, Kinyoun's, and Giemsa methods may be employed to enhance visualization, along with the direct fluorescent-antibody (FA) or enzyme-linked immunosorbent assay (ELISA) techniques that incorporate monoclonal antibody reagents (4).

Recently, the issues of water quality and water testing have become more important and controversial. Today, most water utilities have developed their own water quality testing laboratories or contracted with commercial water laboratories for the recovery and identification of *Cryptosporidium*. Currently, the most common methods for capturing and recovering oocysts from water employ polypropylene cartridge filters or membrane filtration. The oocysts are subsequently eluted and may be concentrated on a percoll-sucrose gradient by flotation or by immunomagnetic separation (IMS) and visualized with commercially available immunofluorescence assay kits. Oocysts are often present in small numbers in environmental samples; therefore, molecular techniques involving the polymerase chain reaction (PCR) may also be used to detect oocysts in water samples (2,8). Integrated cell culture-PCR infectivity assays address the drawbacks of alternative methods such as vital dye staining by allowing for both detection of organisms and the determination of viability and infectivity (9). Considering that, under favorable conditions of moisture and moderate temperature, oocysts can remain viable and infectious for a relatively long time and have been reported viable after storage for 12 months, a great need still exists for simple, efficient, and reliable procedures for capturing and recovering *Cryptosporidium* oocysts from water (10) (Fig. 1).

Treatment

There is no effective specific treatment for *Cryptosporidium* infection, despite testing of hundreds of compounds (1,6). Cryptosporidiosis tends to be self-limiting in patients who have an intact immune system; the clinical course of infection varies, depending on the immune status of the host. Treatment with antidiarrheal drugs along with rehydration therapy may reduce the severity of acute cryptosporidiosis but is less effective for chronic cryptosporidiosis that involves the colon and extraintestinal tissues (4). The antibiotic paromomycin, it has been shown, slightly reduces parasite numbers and stool frequency and may be combined with azithromycin as a course of treatment (6,11). The establishment of a parasitophorous vacuole within the host cell may somehow protect the parasite from antimicrobial drugs (6).

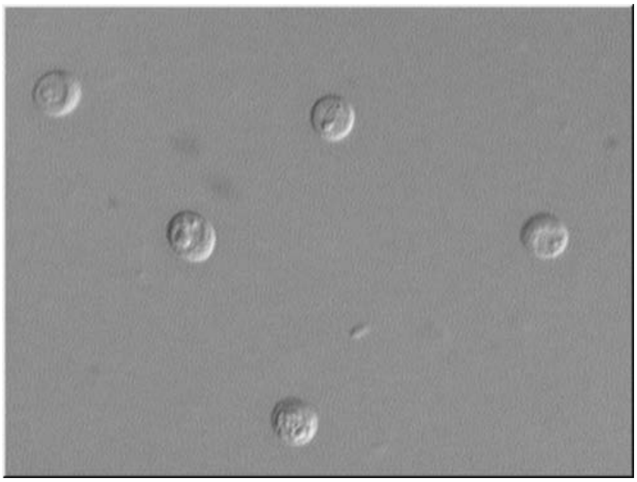

Figure 1. *Cryptosporidium parvum* oocysts viewed at 1000× magnification by Nomarski differential interference contrast (DIC) magnification.

Epidemiology and Prevention

Cryptosporidium is transmitted by oocysts that are usually fully sporulated and infective when they are passed in stool. The principal transmission route is direct fecal–oral spread and transmission by contaminated water. Calves and other animals such as livestock, dogs, cats, and wild mammals are potential sources of human infections, and contact with these animals or their feces may be an unrecognized cause of gastroenteritis in humans (1). Generally, young children tend to have higher infection rates, and there is a high prevalence of cryptosporidiosis in children in areas where sanitation and nutrition are poor (10). Direct person-to-person transmission is likely and may occur through direct or indirect contact with stool material. Outbreaks of human disease in day-care centers, hospitals, and urban family groups indicate that most human infections result from person-to-person contact (5). Indirect transmission may occur from exposure to positive specimens in a laboratory, from contaminated surfaces, or from consuming contaminated food or water. In healthy adults who have no serologic evidence of past infection by *Cryptosporidium parvum*, as few as thirty *Cryptosporidium parvum* oocysts is sufficient to cause infection, and in some cases, infection has occurred from just one oocyst (11). In the United States, the parasite has been identified in 15% of patients who have AIDS and diarrhea (5).

The potential contamination of water supplies by *Cryptosporidium* oocysts is a considerable issue for the drinking water industry. Oocysts can penetrate physical barriers and withstand the conventional disinfection processes used for drinking water treatment. Waterborne outbreaks of *Cryptosporidium* are an increasing public health problem and have resulted from untreated surface water, filtered public water supplies, and contaminated well water (3,12). Large-scale outbreaks of cryptosporidiosis in industrialized countries have been associated with contamination of community drinking water (9). Disease transmission through the waterborne route is especially important because of the capacity for affecting large communities of susceptible individuals. A massive waterborne outbreak was reported in Milwaukee, Wisconsin, where contamination of the public water supply during March and April of 1993 resulted in more than 400,000 infections and about 50 deaths (12,13). Additional outbreaks involving public swimming pools and wading pools demonstrate the ability of *Cryptosporidium* to cause infection even when ingested in small amounts of fully chlorinated water (11).

The increase in the number of reported waterborne disease outbreaks associated with the *Cryptosporidium* species can be attributed to improved techniques for oocyst recovery and identification resulting in the demonstration of oocysts in surface and drinking water and in sewage effluents. It is very likely that cryptosporidiosis is underdiagnosed, especially in immunocompetent adults and children, as analysis for the oocysts is not normally included in a routine stool analysis (6). The importance of agricultural wastewater and runoff, particularly from lambs and calves, is also now recognized as a potential source of infective *Cryptosporidium* oocysts (4).

Prevention involves taking proper steps to reduce the likelihood of waterborne contamination. Properly drilled and maintained wells that tap into groundwater are unlikely to contain pathogens because of the natural filtration that takes place as water passes through the soil; however, contamination may still occur if surface water can move through coarse soils or fractured bedrock into groundwater aquifers. Shallow or poorly constructed wells and springs are at risk of contamination from surface water runoff; therefore, wells should be protected from surface contamination by an intact well casing, proper seals, and a cap above ground (14). Human and animal waste contamination are minimized by protecting the watershed, controlling land use, creating and enforcing septic system regulations, and best management practices in an effort to control runoff (14).

Cryptosporidium oocysts are susceptible to ammonia, 10% formalin in saline, freeze-drying, exposure to temperatures below freezing or above 65 °C for 30 minutes, and 50% commercial bleach (4). For individuals who wish to take extra measures to avoid waterborne cryptosporidiosis in their drinking water, according to the EPA and the CDC, boiling the water is the most effective way of killing the organism (15). In addition to boiling water, oocysts can be removed by certain types of filters to ensure that drinking water is safe (14).

BIBLIOGRAPHY

1. Bowman, D. (1999). *Cryptosporidium, Georgis' Parasitology for Veterinarians*, 7th Edn. W.B. Saunders Company, New York, pp. 99–100.
2. Leventhal, R. and Cheadle, R. (1989). *Protozoa, Medical Parasitology: A Self Instructional Text*, 3rd Edn. F.A. Davis Company, Philadelphia, PA, pp. 96–97.
3. Ash, L. and Orihel, T. (1997). *Cryptosporidium parvum*. In: *Atlas of Human Parasitology*, 4th Edn. American Society of Clinical Pathologists, Chicago, IL, pp. 124–125.
4. Garcia, L. and Bruckner, D. (1997). *Intestinal Protozoa (Coccidia and Microsporidia) and Algae*. In: *Diagnostic Medical Parasitology*, 3rd Edn. ASM Press, Washington, DC, pp. 54–66.
5. Ryan, K.J. (1994). *Cryptosporidia*. In: *Sherris Medical Microbiology: An Introduction to Infectious Diseases*, 3rd Edn. Appleton & Lange, Norwalk, CT, pp. 655–657.
6. Clark, D. (1999). New insights into human cryptosporidiosis. *Clin. Microbiol. Rev.* **12**(4): 554–563.
7. Balows, A., Hausler, W., Herrmann, Jr., K., Isenberg, H., and Shadomy, H. (1991). *Cryptosporidium* species. *Manual of Clinical Microbiology*, 5th Edn. ASM, Washington, DC, pp. 764–766.
8. Widmer, G., Orbacz, E., and Tzipori, S. (1999). β-Tubulin mRNA as a marker of *Cryptosporidium parvum* oocyst viability. *Appl. Environ. Microbiol.* **65**(4): 1584–1588.
9. DiGiovanni, G. et al. (1999). Detection of infectious *Cryptosporidium parvum* oocysts in surface and filter backwash water samples by immunomagnetic separation and integrated cell culture-PCR. *Appl. Environ. Microbiol.* **65**(8): 3427–3432.
10. Dubey, J.P., Speer, C.A., and Fayer, R. (1990). *Cryptosporidiosis of Man and Animals*. CRC Press, Boca Raton, FL, pp. 1–82.
11. Guerrant, R. (1997). Cryptosporidiosis: an emerging, highly infectious threat. *Emerging Infect. Dis.* **3**(1): 51–57.
12. Sartory, D.P. et al. (1998). Recovery of *Cryptosporidium* oocysts from small and large volume water samples using a compressed foam filter system. *Lett. Appl. Microbiol.* **27**: 318–322.
13. Sreter, T. et al. (2000). Morphologic, host specificity, and molecular characterization of a Hungarian *Cryptosporidium meleagridis* isolate. *Appl. Environ. Microbiol.* **66**(2): 735–738.
14. Avery, B.K. and Lemley, A. (1996). *Cryptosporidium: A Waterborne Pathogen*. U.S. Department of Agriculture Water Quality Program. 329WQFS6.
15. *Guidance for people with severely weakened immune systems*, United States Environmental Protection Agency, Office of Water, 1999, EPA 816-F-99-005.

MEASURING *CRYPTOSPORIDIUM PARVUM* OOCYST INACTIVATION FOLLOWING DISINFECTION WITH ULTRAVIOLET LIGHT

Zia Bukhari
American Water
Voorhees, New Jersey

Historically, ultraviolet light inactivation of *Cryptosporidium parvum* oocysts was considered ineffective; however, recently it was demonstrated that the methods for measuring oocyst inactivation can yield erroneous results and that neonatal mouse infectivity assays, which indicated that very low UV doses are highly effective for oocyst inactivation, are needed for determining inactivation of UV-treated oocysts. The moral, ethical, and financial constraints of using mouse infectivity has generated the need for more user-friendly alternative methods for measuring oocyst inactivation. Cell culture infectivity assays are considered promising alternatives, and this article discusses a cell culture-immunofluorescence (IFA) procedure, which following optimization, yielded results similar to those expected from mouse infectivity assays. This cell culture-IFA procedure will be an invaluable analytical tool for control or bench scale studies using *C. parvum* oocysts and the same water matrices as those intended for use in UV reactor validation studies.

HISTORICAL PERSPECTIVE

A number of investigators have examined the effect of low-pressure UV light on viruses and bacteria (1,2) and found that MS2 bacteriophage requires approximately 70 mJ/cm^2 of UV light to render 4-log inactivation and that hepatitis A virus requires fourfold lower UV doses to yield similar levels of inactivation. From the studies of Wilson et al. (1), MS2 bacteriophage was approximately two times more resistant than viruses and three to ten times more resistant than bacteria. Protozoan parasites, which display greater resistance than bacteria or viruses to chemical disinfectants, were examined for their susceptibility to UV light by Karanis et al. (3). Reportedly delivery of 400 mJ/cm^2 yielded 3-log inactivation of *Trichomonas vaginalis*, and *Giardia lamblia* required 180 mJ/cm^2 to yield 2-log inactivation. In 1993, Lorenzo-Lorenzo et al. (4) examined the impact of UV light on *Cryptosporidium* oocysts; however, due to an inadequate description of their disinfection experiments, the effectiveness of UV light for *Cryptosporidium* oocyst inactivation was not recognized. Two years later, a unit known as the *Cryptosporidium* inactivation device (CID), which delivered a total UV dose of 8748 mJ/cm^2 from low-pressure lamps, was examined and 2- to 3-log inactivation of oocysts was reported using the fluorogenic vital dye assay (4′-6′-diamidino-2-phenylindole and propidium iodide) and *in vitro* excystation (5). Following this, Clancy et al. (6) confirmed the findings of Campbell et al. (5); however, the apparent requirement for high UV doses for significant oocyst inactivation continued to present an obstacle to the regulatory and water industry in supporting implementation of this technology.

From 1998 to 1999, comparative bench scale and demonstration scale (215 gpm) *Cryptosporidium* oocyst inactivation studies were performed that used medium-pressure UV lamps and examined oocyst inactivation using *in vitro* viability assays (DAPI/PI and *in vitro* excystation) as well as mouse infectivity assays (7). These studies conclusively demonstrated the effectiveness of low UV doses for inactivating *Cryptosporidium* in finished water. These data rapidly gained attention from both regulatory agencies and the water industry and instigated the revolution in the U.S. water industry for UV disinfection to inactivate *Cryptosporidium* oocysts effectively. In addition to demonstrating the effectiveness of low levels of UV light for oocyst inactivation, Bukhari et al. (7) also found that *in vitro* viability assays underestimated oocyst inactivation following their exposure to either UV light or ozonation (8) and that mouse infectivity assays, albeit cumbersome, needed to be the methods of choice for future studies of this nature. Following these initial studies, a number of investigators confirmed the effectiveness of UV light for inactivating *Cryptosporidium* oocysts as well as *Giardia* cysts (9) and Microsporidia spores (10).

MEASURING *C. PARVUM* OOCYST INACTIVATION FOLLOWING EXPOSURE TO UV LIGHT

In vitro assays for determining the viability of *C. parvum* oocysts offer several advantages over the traditional animal infectivity assays in that results can be generated in a short time. Although these assays are relatively simple to use and relatively inexpensive, recently it was shown that they do not accurately demonstrate whether oocysts are capable of infectivity in neonatal mice. For example, using fluorogenic vital dyes (i.e., DAPI/PI, SYTO-9, and SYTO-59) or *in vitro* excystation, the viability information for oocysts subjected to ultraviolet light (7) or ozone (8) disinfection was grossly overestimated compared with infectivity data using neonatal mice. Unfortunately the "gold" standard mouse infectivity assays have various limitations, which in addition to the moral, ethical, and financial constraints of animal experimentation, include the high degree of variability of using outbred strains of neonatal mice. As a result, the water industry has been seeking alternative, more user-friendly procedures for measuring oocyst viability/infectivity.

In vitro infectivity assays for determining *C. parvum* oocyst inactivation have the potential to fill this void. One such *in vitro* infectivity assay uses human ileocecal adenocarcinoma (HCT-8) cells in conjunction with quantitative polymerase chain reaction (q-PCR) and has been used previously to determine the infectivity of environmentally derived oocysts (11). This assay has the advantage of providing a direct indication of the amount of amplifiable DNA with a specific set of primers. Furthermore, this assay also can help avoid the high variability of using the most probable number format of mouse infectivity assays. Using this assay to measure inactivation of bench-scale, UV-treated *C. parvum* oocysts indicated log inactivation values of 1.16, 1.24, and 1.84 logs for 10, 20, and 40 mJ/cm^2, respectively. These inactivation values for *C. parvum* oocysts were considerably lower than those reported in previous studies using mouse infectivity assays (7,12). Previous research has indicated that oocysts exposed to UV doses ranging between 10 and 40 mJ/cm^2 continue to respond to excystation stimuli and release their sporozoites. This suggests that UV-treated oocysts inoculated onto cell monolayers could potentially excyst, and then released sporozoites could invade cell monolayers. Examination of HCT-8 monolayers confirmed the presence of pinpoints of invasion, which probably originated from invasive sporozoites (Fig. 1). A number of different oocyst pretreatment steps were used to promote differentiation between UV affected and unaffected sporozoites; however, these pretreatments did not nullify the background signal detected by the quantitative PCR procedures.

Figure 1. UV-treated oocysts yielding pinpoints of invasion.

CELL CULTURE-IFA FOR MEASURING INACTIVATION OF UV-TREATED *C. PARVUM* OOCYSTS

Bukhari and LeChevallier (13) used an infectivity-enhancing oocyst pretreatment step consisting of preacidification and 0.05% bile treatment followed by inoculation onto HCT-8 monolayers, incubation at 37 °C for 72 h, and quantitative detection of infection by immunofluorescence microscopy. The rationale behind this assay was that UV-treated oocysts would undergo excystation and subsequently invade HCT-8 cells, leading to discrete pinpoints of invasion. In contrast, untreated organisms would invade the HCT-8 cells but would continue to differentiate further to generate clusters of secondary infection (Fig. 2).

Enumerating secondary clusters of infection for various inocula of oocysts (i.e., 10,100, and 1000 oocysts) has enabled development of a dose response curve (Fig. 3). This curve, which was generated from multiple trials ($n = 75$ to 115) using predetermined oocyst inocula, was analyzed by linear regression to derive an equation for calculating the number of infectious organisms present in an inoculum of UV-treated oocysts after the number of clusters had been determined by immunofluorescence microscopy.

UV disinfection experiments using the Iowa isolate of *C. parvum* oocysts were conducted, and following cell culture infectivity, the infectious clusters were extrapolated from the dose response curve generated for the untreated oocysts to calculate levels of inactivation (Fig. 4). A UV dose of 1 mJ/cm^2 rendered 0.44 log inactivation of oocysts, and UV doses between 1 and 4 mJ/cm^2 yielded a linear decline in oocyst inactivation; 3 mJ/cm^2 rendered 2.79–2.84 log inactivation, and 4-log inactivation occurred at 4 mJ/cm^2. Between 4 and 20 mJ/cm^2, measurements by the cell culture procedure continued to indicate oocyst inactivation levels of 4 logs. It is highly probable that the actual levels of oocyst inactivation after UV doses between 4 and 20 mJ/cm^2 were in excess of 4 logs; however, as the determinable oocyst inactivation by the cell culture procedure is a factor of the highest original inoculum of UV-treated oocysts applied onto the monolayers (i.e., 1×10^5 oocysts), the maximum measurable levels of inactivation by this cell culture procedure were limited to 4 logs or lower in these experiments. Although theoretically it would be possible to use higher oocyst inocula (i.e., 1×10^6–1×10^8 oocysts per monolayer) to accurately determine the levels of oocyst inactivation at individual UV doses between 4 and 20 mJ/cm^2, this would be of little empirical value from the perspective of the disinfection needs of the water industry or the logistics of conducting experiments with adequate quantities of infectious oocysts.

As a result of the investigations by Bukhari and LeChevallier (13), it has been demonstrated that UV-treated *C. parvum* oocysts can undergo excystation and that the sporozoites from these inactivated oocysts can invade monolayers of HCT-8 cells to generate pinpoints of invasion. Optimization of excystation triggers and cell culture incubation periods led to development of a cell culture-IFA procedure that enabled detecting as few as 10 infectious oocysts. Using this cell culture-IFA allowed discriminating pinpoints (generated from noninfectious but invasive sporozoites) from secondary structures (generated from infectious sporozoites). This phenomenon is the primary reason that the cell culture quantitative PCR procedure described by DiGiovanni et al. (11) cannot be used to measure inactivation of UV-treated oocystes because the assay cannot discriminate between DNA originating from invasive sporozoites and that from infectious sporozoites.

Using the optimized cell culture-IFA procedure, it was confirmed that low UV doses (i.e., 2–5 mJ/cm^2) can be very effective for inactivating *C. parvum* oocysts. Comparison of data from this current study with previously published reports, using either tissue culture infectivity or mouse

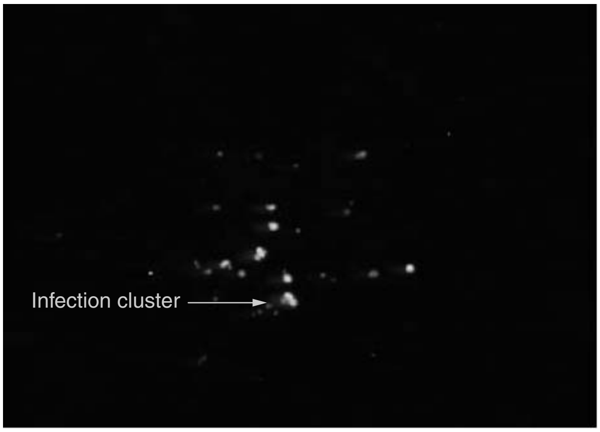

Figure 2. Infection clusters from infectious oocysts.

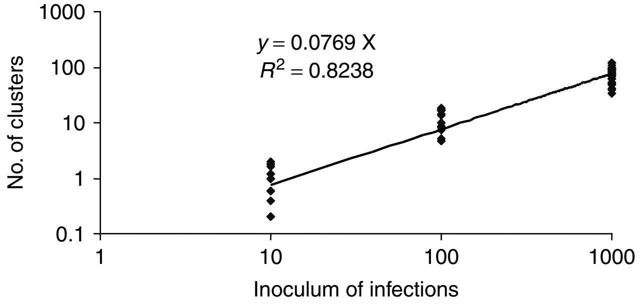

Figure 3. *Cryptosporidium parvum* oocyst infectivity in HCT-8 monolayers, as detected by immunofluorescence microscopy.

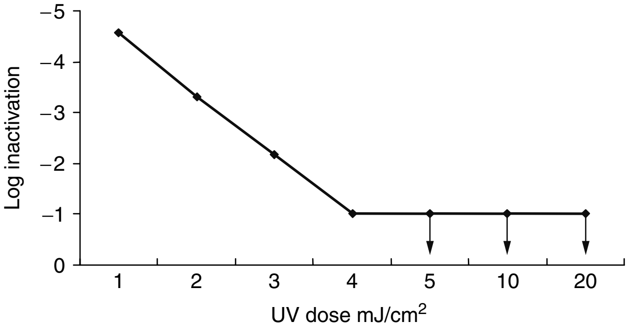

Figure 4. Inactivation of *C. parvum* oocysts by UV light, as determined by the cell culture-IFA procedure.

infectivity assays, also demonstrates excellent agreement. Numerous studies (9,12,14,15) have assessed levels of oocyst inactivation from UV doses ranging between 1 and <40 mJ/cm^2. Not surprisingly, there is wide variability in the data generated from these various studies, which may result from differences in oocyst preparation, disinfection conditions, and methods for infectivity measurements. Nonetheless, using UV doses between 1 and 3 mJ/cm^2, the oocyst inactivation data generated in our current study were within the range of inactivation data generated in the studies cited (Fig. 5).

Data from the previous disinfection studies have also been used in the USEPA's Ultraviolet Disinfection Guidance Manual, to predict the probable levels of oocyst inactivation for a given UV dose. According to these probability tables, the UV dose required for 0.5-log inactivation of oocysts ranged from 0.7 to 3.0 mJ/cm^2, whereas in our current study, 0.44-log inactivation of *C. parvum* oocysts was noted at delivery of 1 mJ/cm^2. The probability tables also predicted 50% likelihood that a UV dose of 2.9 mJ/cm^2 would lead to 2.5-log oocyst inactivation, whereas in our study, delivery of 3 mJ/cm^2 led to approximately 2.8-log inactivation. Although our data indicated 3- and 4-log oocyst inactivation from UV doses as low as 4 mJ/cm^2 and consistently >4-log inactivation of oocysts from UV doses ranging between 5 and 20 mJ/cm^2, >4-log oocyst inactivation was noted only when UV doses exceeded 10 mJ/cm^2 in the previous studies cited. At 5 mJ/cm^2, although our data indicated >4-log inactivation, the previous studies indicated >1-log and approximately 3-log inactivation of oocysts. These differences need to be interpreted with caution and do not necessarily imply differences in susceptibility between the isolates of *C. parvum* used. As an example, it is known that the oocyst inactivation levels can be censored as a result of the original oocyst inoculum administered in the infectivity assay, which in turn influences the outcome of the calculated levels of oocyst inactivation. In (Fig. 5), examining the inactivation data at 5 mJ/cm^2 suggests that a single oocyst inactivation value (>1 log) may be an outlier. Perhaps this inactivation level was derived from administering a lower inoculum of oocysts (i.e., 1×10^3 oocysts) than the inoculum size (i.e., 1×10^5 oocysts) used in our study. Should this be the case, then it would argue that our data provide a more robust indication of the levels of oocyst inactivation than those generated from censored data. Furthermore, examining the general oocyst inactivation patterns for increasing UV doses in (Fig. 5) adds further credibility to the previous discussion, as data accrued during the course of our study both overlap and extend the general trend in oocyst inactivation provided by previous UV disinfection studies. This would further support the finding that oocyst inactivation levels at 5 mJ/cm^2 are more likely to be greater than 3 or 4 logs.

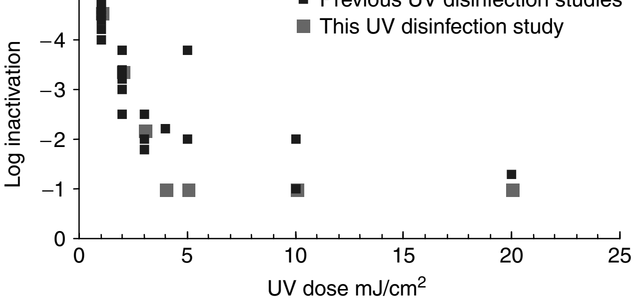

Figure 5. *Cryptosporidium parvum* oocyst inactivation in various UV disinfection studies.

In conclusion, UV disinfection is a promising technology for rapid and effective inactivation of waterborne *C. parvum* oocysts. Numerous manufacturers can supply equipment using medium-pressure, low-pressure, or low-pressure, high-output lamps within their UV reactors. For water utilities to select the appropriate UV reactors for their disinfection requirements, it is imperative that reactors are validated on-site or off-site using the matrices intended for disinfection. Although the validation studies are likely to be conducted using surrogate organisms such as the MS2 bacteriophage, it is imperative to do control- or bench-scale studies also using the same matrix and *C. parvum* oocysts. For the latter, the cell culture-IFA procedure described in this article is an invaluable analytical tool.

Acknowledgments

American Water and its subsidiary utilities funded this research. The author gratefully acknowledges the technical assistance of Felicia Abrams, Janice Weihe, and Rita Scanlon from the American Water Quality Control and Research Laboratory. Dr. Mark LeChevallier, American Water, Voorhees, NJ, is thanked for tirelessly sharing insights into water industry needs.

BIBLIOGRAPHY

1. Wilson, B.R., Roessler, P.F., Van Dellen, E., Abbaszadegan, M., and Gerba, C.P. (1992). Coliphage MS 2 as a UV water disinfection efficacy test surrogate for bacterial and viral pathogens. *Proc. Water Qual. Technol. Conf.*, Nov. 15–19, Toronto.
2. Wiedenmann, A. et al. (1993). Disinfection of hepatitis A virus and MS 2 coliphage in water by ultraviolet irradiation: Comparison of UV susceptibility. *Water Sci. Technol.* **27**: 335–338.
3. Karanis, P. (1992). UV sensitivity of protozoan parasites. *J. Water Supply Res. Technol. Aqua* **41**: 95.
4. Lorenzo-Lorenzo, M.J., Ares-Mazas, M.E., Villacorta-Martinez de Maturana, I., and Duran-Oreiro, D. (1993). Effect of ultraviolet disinfection of drinking water on the viability of *Cryptosporidium parvum* oocysts. *J. Parasitol.* **79**: 67–70.
5. Campbell, A.T., Robertson, L.J., Snowball, M.R., and Smith, H.V. (1995). Inactivation of oocysts of *Cryptosporidium parvum* by ultraviolet irradiation. *Water Res.* **29**: 2583–2586.
6. Clancy, J.L., Hargy, T.M., Marshall, M.M., and Dyksen, J.E. (1998). Inactivation of *Cryptosporidium parvum* oocysts in water using ultraviolet light. *J. Am. Water Works Assoc.* **90**: 92–102.
7. Bukhari, Z., Hargy, T.M., Bolton, J.R., Dussert, B., and Clancy, J.L. (1999). Inactivation of *Cryptosporidium parvum* oocysts using medium-pressure ultraviolet Light. *J. Am. Water Works Assoc.* **91**: 86–94.
8. Bukhari, Z. et al. (2000). Comparison of *Cryptosporidium parvum* viability and infectivity assays following ozone treatment of oocysts. *Appl. Environ. Microbiol.* **66**: 2972–2980.

9. Craik, S.A., Finch, G.R., Bolton, J.R., and Belosevic, M. (2000). Inactivation of *Giardia muris* cysts using medium pressure ultraviolet radiation in filtered drinking water. *Water Res.* **34**: 4325–4332.
10. John, D.E., Nwachuku, N., Pepper, I.L., and Gerba, C.P. (2003). Development and optimization of a quantitative cell culture infectivity assay for the microsporidium *Encephalitozoon intestinalis* and application to ultraviolet light inactivation. *J. Microbiological Methods* **52**: 183–196.
11. Di Giovanni, D.D., Hashemi, F.H., Shaw, N.J., Abrams, F.A., LeChevallier, M.W., and Abbaszadegan, M. (1999). Detection of infectious *Cryptosporidium parvum* oocysts in surface and filter backwash water samples by immunomagnetic separation and integrated cell culture-PCR. *Appl. Environ. Microbiol.* **65**: 3427–3432.
12. Clancy, J.L., Bukhari, Z., Hargy, T.M., Bolton, J.R., Dussert, B., and Marshall, M.M. (2000). Using UV to inactivate *Cryptosporidium*. *J. Am. Water Works Assoc.* **92**: 97–104.
13. Bukhari, Z. and LeChevallier, M.W. (2004). Finished water disinfection with UV light: Overview of validation studies at American Water. *IUVA News* **6**: 15–20.
14. Landis, H.E., Thompson, J.E., Robinson, J.P., Blatchley, E.R. (2000). Inactivation responses of Cryptosporidium parvum to UV radiation and gamma radiation. *Proceedings of the Water Quality Technology Conference*, Salt lake City, November 5–9.
15. Shin, G.A., Linden, K.G., Arrowood, M., and Sobsey, M. (2001). Low pressure UV inactivation and subsequent repair potential of Cryptosporidium parvum oocysts. *Appl. Environ. Microbiol.* **67**: 3029–3032.

DECHLORINATION

Rasheed Ahmad
Khafra Engineering Consultants
Lilburn, Georgia

Dechlorination is the practice of removing all or a specified fraction of total residual chlorine. In potable water practice, dechlorination is used to reduce the residual chlorine to a specified level at a point where the water enters the distribution system. Dechlorination has been beneficial for waters that are burdened with high concentrations of ammonia nitrogen and organic nitrogen. In some cases where taste and odor control is a severe problem, control is achieved by complete dechlorination, followed by rechlorination. This removes the taste-producing nuisance residuals and prevents the formation of nitrogen trichloride (NCl_3) in the distribution systems. Dechlorination of wastewater and power plant cooling water is required to eliminate residual chlorine toxicity, which is harmful to the aquatic life in the receiving waters. Other special applications requiring dechlorination are ahead of demineralizers, boiler makeup water, certain food plant operations, bottled water, and the beverage industry. In these cases, the dechlorination process is arranged to remove all residual chlorine.

The most practical method of dechlorination is by sulfur dioxide and/or aqueous solutions of sulfite compounds. Among sulfur/sulfite compounds, sulfur dioxide (SO_2) is the most widely used chemical for dechlorination. Other compounds such as sodium bisulfite ($NaHSO_3$) and sodium metabisulfite ($Na_2S_2O_5$) can be used as practical alternatives to sulfur dioxide. Sodium thiosulfate ($Na_2S_2O_3$) is another chemical that is used entirely as a laboratory dechlorinating chemical. Other methods used for dechlorination are granular-activated carbon and aeration.

Free and combined residual chlorine can be effectively reduced by sulfur dioxide and sulfite salts. Free chlorine is the amount of chlorine available as dissolved chlorine gas (Cl_2), hypochlorous acid (HOCl), and hypochlorite ion (OCl^-) that is not combined with ammonia (NH_3) or any other compounds in water. Combined chlorine is the sum of species composed of free chlorine and ammonia, including monochloramine (NH_2Cl), dichloramine ($NHCl_2$), and trichloramine or nitrogen trichloride (NCl_3). The sulfite ion is the active agent when sulfur dioxide or sulfite salts are dissolved in water. Their dechlorination reactions are identical. Sulfite reacts instantaneously with free and combined chlorine. Reactions yield small amounts of acidity, which is neutralized by the alkalinity of water (2.8 milligrams of alkalinity as calcium carbonate is consumed per milligram of chlorine reduced). Most potable waters and wastewaters have sufficient alkalinity buffering power that there is no cause for concern about lowering the pH by sulfur dioxide addition. The amount of sulfur dioxide required per part of chlorine is 0.9, but in actual practice, this ratio can be as high as 1.05. Owing to the low vapor pressure of sulfur dioxide, special precaution must be taken when using ton containers to prevent reliquefication. Unlike chlorination, dechlorination with sulfur dioxide does not require any contact chamber as the reaction occurs in a matter of seconds, probably 15–20 s at the most. There has been some apprehension about the possibility that excess sulfur dioxide might consume a significant amount of dissolved oxygen in the receiving waters downstream from a dechlorinated water discharge. In properly controlled systems, this reaction does not have sufficient time for completion. Hence, little effect on dissolved oxygen concentration has been reported.

Granular and powdered carbon may be used to dechlorinate free, and some combined, chlorine residuals. The carbon requirements for dechlorination are typically determined by on-site pilot testing. The parameters of significance include mean particle diameter of carbon (pressure drop within the carbon contactor) and influent quality (pH, organics, and colloids). Carbon doses in the range of 30 to 40 mg/L have been reported. Granular-activated carbon (GAC) has proved effective and reliable as a dechlorination agent in potable water treatment. In addition, carbon provides filtration that removes other undesirable materials. In wastewater treatment, however, GAC has not been successful as a dechlorinating agent, possibly, because GAC is poor at removing organochloramines that form when significant concentrations of organic nitrogen are present. Because of the higher cost of carbon systems, their use is typically limited to specific sites or effluent with special discharge limitations. An aeration process can be used for dechlorination. Chlorine, hypochlorous acid, chlorine dioxide, and nitrogen trichloride are sufficiently volatile to be removed by aeration.

READING LIST

AWWA. (1999). *Water Quality and Treatment*, 5th Edn. McGraw-Hill, New York.

Fair, G.M., Geyer, J.C., and Okun, D.A. (1968). *Waste and Wastewater Engineering*. John Wiley & Sons, New York.

WEF. (2002). *Chlorination/Dechlorination Handbook*. Water Environment Federation, Alexandria, VA.

WEF/ASCE. (1998). *Design of Municipal Wastewater Treatment Plants*, 4th Edn. Water Environment Federation, Alexandria, VA.

White, G.C. (1999). *Handbook of Chlorination and Alternative Disinfectants*, 4th Edn. John Wiley & Sons, New York.

DESALINATION

NIKOLAY VOUTCHKOV
Poseidon Resources Corporation
Stamford, Connecticut

INTRODUCTION

Desalination is the production of fresh, low-salinity potable water from a saline water source (seawater or brackish water) via membrane separation or evaporation. The mineral/salt content of the water is usually measured by the water quality parameter, total dissolved solids (TDS), in milligrams per liter (mg/L) or parts per thousand (ppt). The World Health Organization and the United Sates Environmental Protection Agency (EPA), under the Safe Drinking Water Act, have established a maximum TDS concentration of 500 mg/L as a potable water standard. This TDS level can be used as a classification limit to define potable (fresh) water. Typically, water of TDS concentration higher than 500 mg/L and lower or equal to 15,000 mg/L is classified as brackish. Natural water sources such as sea, bay, and ocean waters that usually have TDS concentrations higher than 15,000 mg/L are generally classified as seawater. For example, Pacific Ocean seawater along the U.S. West Coast has a TDS concentration of 35,000 mg/L, of which approximately 75% is sodium chloride. Approximately 97.5% of the water on our planet is located in the oceans and therefore is classified as seawater. Of the 2.5% of the planet's fresh water, approximately 70% is polar ice and snow, and 30% is groundwater, river and lake water, and air moisture. So even though the volume of the earth's water is vast, less than 10 million of the 1400 million cubic meters of water on the planet are of low salinity and are suitable for use after using only conventional water treatment. Desalination provides a means for tapping the world's main water resource—the ocean.

During the past 30 years, desalination technology has made great strides in many arid regions of the world such as the Middle East and the Mediterranean. Today, desalination plants operate in more than 120 countries worldwide, and some desert states, such as Saudi Arabia and the United Arab Emirates, rely on desalinated water for more than 70% of their water supply. According to the 2003 desalination plant inventory report prepared by the International Desalination Association (1), by the end of 2001, worldwide there were 15,233 desalination plants whose total installed treatment capacity is 32.4 million cubic meters per day (8560 MGD).

Seawater or brackish waters are typically desalinated using two general types of water treatment technologies—thermal evaporation (distillation) and membrane separation. Currently, approximately 43.5% of the world's desalination systems use thermal evaporation technologies. This percentage has been decreasing steadily during the past 10 years due to the increasing popularity of membrane desalination, which is driven by remarkable advances in membrane separation and energy recovery technologies and associated reduction of overall water production costs.

THERMAL DESALINATION

All thermal desalination technologies use distillation (heating of source water) to produce water vapor that is then condensed into low-salinity potable water. Thermal desalination is most popular in the Middle East, where seawater desalination is typically combined with power generation, which provides low-cost steam for distillation. Thermal desalination requires a large quantity of steam. The ratio of the mass of potable water produced to the mass of heating steam used to produce this water is commonly referred to as a gain output ratio (GOR) or performance ratio. Depending on the thermal desalination technology used, site-specific conditions, and the source water quality, the GOR varies between 2 and 24. The thermal desalination technologies most widely used today are multistage flash distillation (MSF), multiple effect distillation (MED), and vapor compression (VC).

Multistage Flash Distillation

In the MSF evaporator vessels (flash stages or effects), high-salinity source water is heated, whereas the vessel pressure is reduced to a level at which the water vapor "flashes" into steam. Each flash stage (effect) has a condenser to turn the steam into distillate. The condenser is equipped with heat exchanger tubes that are cooled by the source water fed to the condensers. An entrainment separator removes high-salinity mist from the low-salinity rising steam. This steam condenses into pure water (distillate) on the heat exchanger tubes and is collected in distillate trays from where it is conveyed to a product water tank. Historically, MSF is the first commercially available thermal desalination technology used to produce potable water on a large scale, which explains its popularity. More than 80% of the thermally desalinated water today is produced in MSF plants. The GOR for the MSF systems is typically between 2 and 8. The pumping power required to operate MSF systems is 2.0 to 3.5 kWh per cubic meter of product water.

Multiple Effect Distillation

In the MED process, the source water passes through a number of evaporators (effects or chambers) connected in series and operating at progressively lower pressures. In MED systems, the steam vapor from one evaporator (effect) is used to evaporate water from the next effect. MED desalination systems typically operate at lower temperatures than MSF plants (maximum brine concentrate temperature of 75 °C vs. 115 °C) and yield higher GORs. The newest MED technologies, which include vertically positioned effects (vertical tube evaporators or VTEs), may yield GORs up to 24 kilograms of potable water per kilogram of steam. The pumping power required to operate MED systems is also lower than that typically needed for MSF plants (0.8 to 1.4 kWh per cubic meter of product water). Therefore, MED is now increasingly gaining ground over MSF desalination, especially in the Middle East where thermal desalination is still the predominant method for potable water production from seawater.

Vacuum Compression

The heat source for VC systems is compressed vapor produced by a mechanical compressor or a steam jet ejector rather than a direct exchange of heat from steam. In these systems, the source water is evaporated and the vapor is conveyed to a compressor. The vapor is then compressed to increase its temperature to a point adequate to evaporate source water sprayed over a tube bundle through which the vapor is conveyed. As the compressed vapor exchanges its heat with the new source water that is being evaporated, it condenses into pure water. VC desalination has been used mostly for small municipal and resort water supply systems and industrial applications. The power required to operate mechanical VC systems is typically between 10 and 20 kWh per cubic meter of product water. Further discussion of the applicability of the thermal desalination technologies described is presented elsewhere (2,3).

MEMBRANE DESALINATION

Membrane desalination is a process of separating minerals from source water using semipermeable membranes. Two general types of technologies are currently used for membrane desalination—reverse osmosis (RO) and electrodialysis (ED). In reverse osmosis, the product water (permeate) is separated from the salts in the source water by pressure-driven transport through a membrane. By the RO process, desalinated water is transported under pressure through the membrane while the minerals of the source water are concentrated and retained by the membrane. Applying high pressure for desalination is mainly needed to overcome the naturally occurring process of osmosis, which drives the desalinated water back through the membrane into the water of more concentrated mineral content. Nanofiltration (NF) is a process similar to RO, where membranes whose order-of-magnitude larger pore size is used to remove high-molecular-weight compounds that make water hard (i.e., calcium and magnesium).

Desalination by Electrodialysis

In ED-based treatment systems, the mineral–product water separation is achieved by applying electrical direct current (DC) to the source water, which drives the mineral ions in the source water through membranes to a pair of electrodes of opposite charge. A commonly used desalination technology that applies the ED principle is electrodialysis reversal (EDR). In EDR systems, the polarity of the electrodes is reversed periodically during the treatment process.

The energy for ED desalination is proportional to the amount of salt removed from the source water. TDS concentration and source water quality determine to a great extent which of the two membrane separation technologies (RO or ED) is more suitable and cost-effective for a given application. Typically, ED membrane separation is cost-competitive for source waters whose TDS concentration is lower than 2000 mg/L. The TDS removal efficiency of ED desalination systems is not affected by nonionized compounds or objects of weak ion charge (i.e., solids particles, organics, and microorganisms). Therefore, the ED membrane desalination process can treat source waters of higher turbidity, biofouling, and scaling potential than RO systems. However, the TDS removal efficiency of ED systems is typically lower than that of RO systems (15 to 90% vs. 99.0% to 99.8%), which is one key reason why they are used mainly for brackish water desalination.

Reverse Osmosis Desalination

Reverse osmosis desalination is the most widely used membrane separation process today. Currently, there are more than 2000 RO membrane seawater desalination plants worldwide whose total production capacity is in excess of 3 million cubic meters per day (800 MGD). For comparison, the number of ED plants in operation is less than 300, and their total production capacity is approximately 0.15 million cubic meters per day (40 MGD).

RO membrane desalination plants include the following key components: a source water intake system, pretreatment facilities, high-pressure feed pumps, RO membrane trains, and a desalinated water conditioning system. The source water intake system could be an open surface water intake or a series of seawater beach wells or brackish groundwater wells. Depending on the source water quality, the pretreatment system may include one or more of the following processes: screening, chemical conditioning, sedimentation, and filtration. Figure 1 shows a typical configuration of a seawater RO membrane system. The filtered water produced by the plant's pretreatment system is conveyed by transfer pumps from a filtrate water storage tank through cartridge filters and into the suction pipe of the high-pressure RO feed pumps. The cartridge filters are designed to retain particles of 1 to 20 microns that have remained in the source water after pretreatment. The main purpose of the cartridge filters is to protect the RO membranes from damage. The high-pressure feed pumps are designed to deliver the source water to the RO membranes at a pressure required for membrane separation of the freshwater from the salts, typically 15 to 35 bars (200 to 500 psi) for brackish source water and 55 to 70 bars (800 to 1,000 psi) for seawater. The feed pressure required is site-specific and is mainly determined by the source water salinity and the configuration of the RO system.

The "engine" of every desalination plant that turns seawater into fresh potable water is the RO membrane element (Fig. 2). The most widely used RO membrane element consists of two membrane sheets glued together

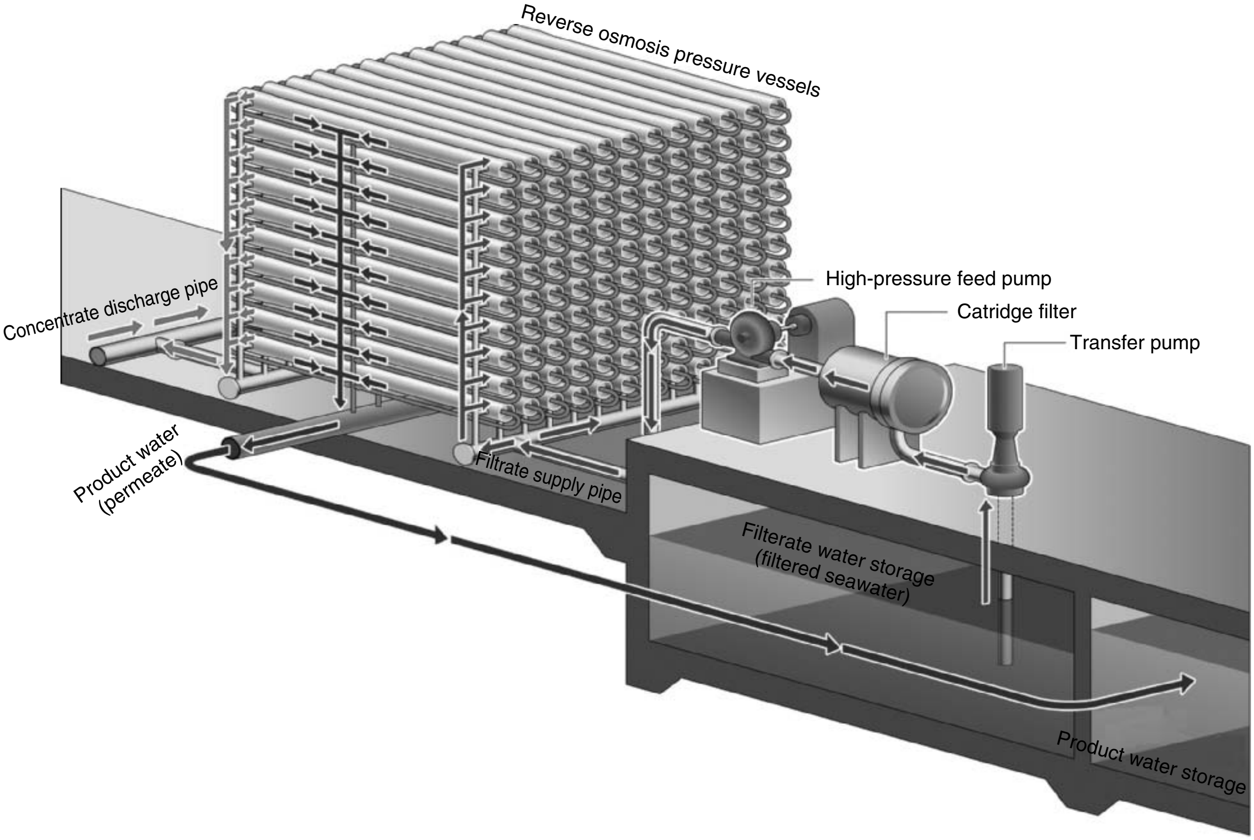

Figure 1. RO membrane train.

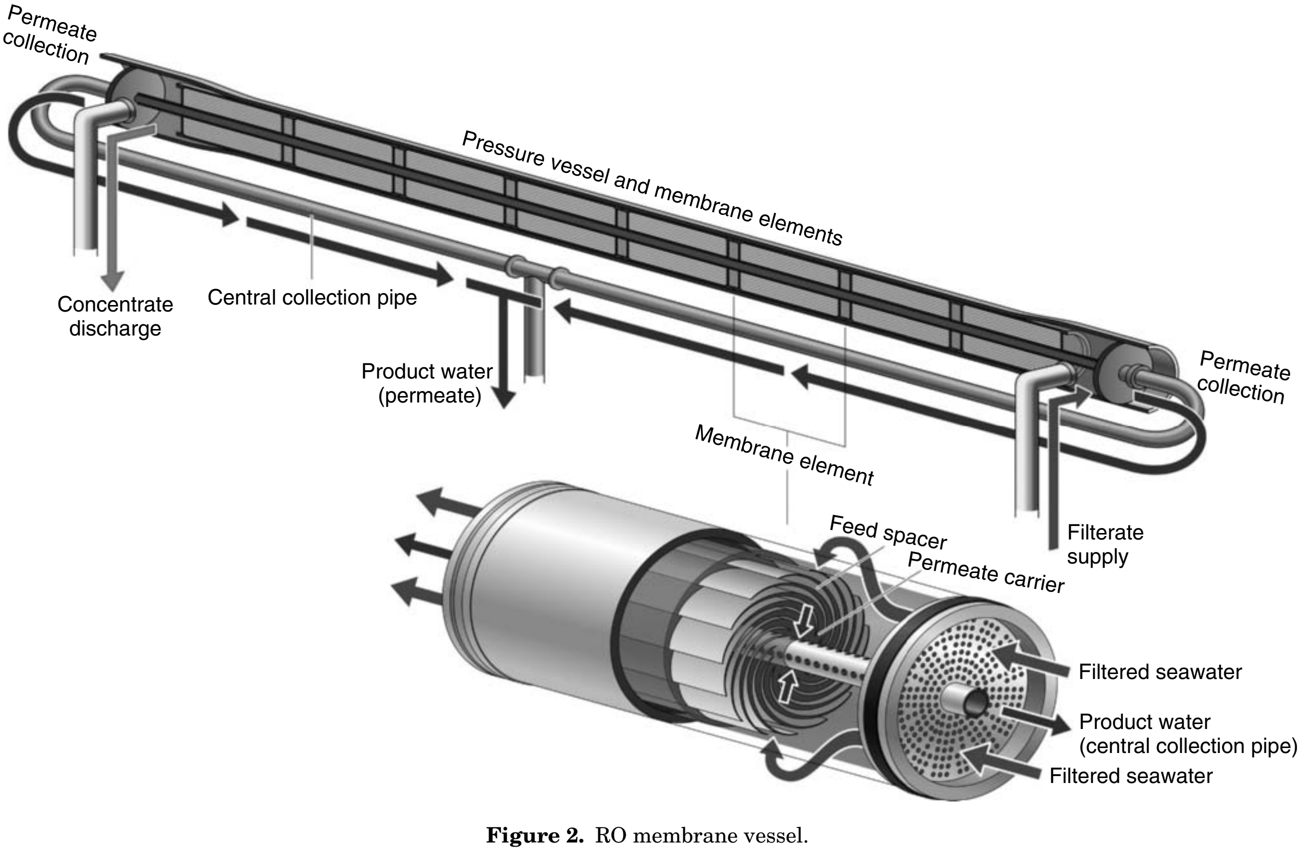

Figure 2. RO membrane vessel.

and spirally wound around a perforated central tube through which the desalinated water exits the membrane element. The first membrane sheet, which actually retains the source water minerals on one side of the membrane surface, is typically made of thin-film composite polyamide and has microscopic pores that can retain compounds smaller than 200 daltons. This sheet, however, is usually less than 0.2 microns thick, and to withstand the high pressure required for salt separation, it is supported by a second thicker membrane sheet, which is typically made of higher porosity polysulfone that has several orders-of-magnitude larger pore openings. The commercially available membrane RO elements today are of standardized diameter, length, and salt rejection efficiency. For example, the RO membrane elements most commonly used for potable water production in large-scale plants are 8 in. in diameter, 40 in. long, and can reject 99.5% or more of the TDS in the source water. Standard membrane elements have limitations with respect to a number of performance parameters such as feed water temperature (45 °C), pH (minimum of 2 and maximum of 10), silt density index (less than 4), chlorine content (not tolerant to chlorine in measurable amounts), and feed water pressure [maximum of 80 to 100 bars (1100 to 1400 psi)].

During reverse osmosis, the water molecules move through the RO membranes at a rate commonly referred to as flux. Membrane flux is expressed in cubic meters per second per square meter (m^3/sm^2) or gallons per day per square foot (gfd) of active membrane surface area. For example, a typical seawater membrane RO element is operated at 8 to 10 gfd.

Membrane performance tends to deteriorate over time due to a combination of wear and tear and irreversible fouling of the membrane elements. Typically, membrane elements have to be replaced every 3 to 5 years to maintain their performance in water quality and power demand for salt separation. Improvements in membrane element polymer chemistry and production processes have made the membranes more durable and have extended their useful life. Elaborate conventional media pretreatment technologies and ultra- and microfiltration membrane pretreatment systems prior to RO desalination are expected to extend membrane useful life to 7 years and beyond, thereby reducing the costs for their replacement and the overall cost of water.

RO membrane elements are installed in pressure vessels that usually house six to eight elements per vessel (see Fig. 2). Multiple pressure vessels are arranged on support structures (racks) that form RO trains. Each RO train is typically designed to produce between 10% and 20% of the total amount of the membrane desalination product water flow. Figure 1 depicts one RO train.

After the RO salt/water separation is complete, a great portion of the feed water energy applied through the high-pressure RO pumps stays with the more concentrated seawater and can be recovered and reused to minimize the overall energy cost of seawater desalination. Dramatic improvements in membrane elements and energy recovery equipment during the last 20 years coupled with enhancements in the efficiency of RO feed pumps and reduction of the pressure losses through the membrane elements have allowed reducing the power to desalinate seawater to less than 3.5 kWh/m^3 (13.5 kWh/1000 gallons) of freshwater produced today. Considering that the cost of power is typically 20% to 30% of the total cost of desalinated water, these technological innovations have contributed greatly to reducing the overall cost of seawater desalination. Novel energy recovery systems working on the pressure exchange principle (pressure exchangers) are currently available in the market, and use of these systems is expected to reduce further desalination power costs by approximately 10% to 15%. The pressure exchangers transfer the high pressure of the concentrated seawater directly into the RO feed water at an efficiency exceeding 95%. Future lower energy RO membrane elements are expected to operate at even lower pressures and to continue to yield further reduction in the cost of desalinated water.

The ratio between the volume of the product water produced by the membrane desalination system and the volume of the source water used for its production is commonly called recovery and is presented as a percentage of the plant RO system feed water volume. The maximum recovery that can be achieved by a given pressure-driven membrane desalination system depends mainly on the source water salinity and is limited by the magnitude of the osmotic pressure to be overcome by the RO system high-pressure feed pumps and by the scaling potential of the source water. Scaling occurs when the minerals left behind on the rejection side of the RO membrane are concentrated to a level at which they begin to form precipitates (crystalline compounds), which in turn plug the membrane pores and interfere with fresh water transport through the membrane. Typically, desalination plants using brackish source water can achieve 65% to 85% recovery. Seawater desalination plants can turn only 40% to 60% of source water into potable water because seawater typically has an order of magnitude higher salinity than brackish water. Detailed guidelines for designing membrane desalination plants are provided elsewhere (4,5).

DESALINATED WATER QUALITY

When membrane desalination is used to produce potable water, the product water quality depends mainly on the concentration of salts in the source water, the salt-rejection capability of the RO membranes, and the applied membrane feed pressure. In thermal desalination, the product water quality depends mostly on the heating system and temperature, the pressure applied, and the type of distillation technology. The product water quality of thermal desalination systems is significantly less dependent on source water salinity than that produced by a membrane system. This makes thermal distillation processes an attractive alternative for parts of the world, such as the Middle East, where source water TDS is very high (40,000 to 46,000 mg/L) and low-cost steam is readily available.

Typically, thermal desalination processes produce water of very low TDS concentration (1 to 60 mg/L). A single-stage, reverse osmosis system for seawater

Next Page

desalination produces permeate of 250 to 350 mg/L. More elaborate RO systems can produce water of quality matching that of a thermal desalination plant. Usually, the temperature of thermally desalinated water is 2 °C to 5 °C higher than the temperature of water produced by membrane separation.

Desalinated water (permeate) produced by both thermal and membrane desalination is highly corrosive and has to be chemically conditioned (typically by lime addition) to increase product water alkalinity and to adjust pH (usually by addition of chlorine dioxide and/or acid to the lime conditioned permeate) to meet potable water quality regulations.

CONCENTRATE MANAGEMENT

The two main streams produced by every desalination plant are low-salinity freshwater and high-salinity concentrate. The results of a recent study completed by the U.S. Bureau of Reclamation (6) on the concentrate disposal methods most widely used in the United States (in order of decreasing frequency) are shown in Table 1.

Sanitary sewer and surface water discharge are the two most popular and cost-effective methods for concentrate disposal. Depending on the site-specific conditions, deep well injection, evaporation ponds, and spray irrigation can also be competitive concentrate disposal alternatives. The zero liquid discharge system typically has the highest construction and operating costs. However, under specific circumstances, such as cold climate, low evaporation and soil uptake rates, high land costs, and low power costs, the zero liquid discharge system can be cost-competitive with evaporation pond and spray irrigation disposal.

Table 1. Concentrate Disposal Methods and Their Frequency of Use in the United States

Concentrate Disposal Method	Frequency of Use, % of Plants Surveyed
Surface water discharge	45
Sanitary sewer discharge	42
Deep well injection	9
Evaporation ponds	2
Spray irrigation	2
Zero liquid discharge	0

COST OF DESALINATED WATER

Desalination water costs depend on many factors, including the type of treatment technology; the source water quality; the target product water quality; the size of the desalination plant; and the costs of energy, chemicals, labor, and membranes. The all-inclusive cost of potable water produced using a brackish water source treated by membrane separation usually varies between US$0.25 and $0.65/m^3 (US$1.0 to $2.5/1000 gallons) of product water. Seawater desalination cost is estimated at US$0.65 to $1.20/m^3 (US$2.5 to $4.5/1000 gallons). For very small plants located in isolated areas, this cost could be significantly higher. For comparison, the cost of water produced from fresh natural water sources (i.e., low-salinity groundwater, lake, and river water) is typically between US$0.15 and $0.40/m^3 ($0.6 and $1.5/1000 gallons). This cost is driven mainly by the source water quality, the type of treatment technologies used, and the size of the water treatment plant. Developments in seawater desalination technology during the past two decades combined with the transition to construction of large capacity plants, colocation with power plant generation facilities, and enhanced competition by using the build-own-operate-transfer (BOOT) method of project delivery have resulted in a dramatic decrease in the cost of desalinated water (7). These factors are expected to continue to play a key role in further reducing the production cost of desalinated water.

BIBLIOGRAPHY

1. Wagnick Consulting. (2002). IDA *Worldwide Desalination Plant Inventory*. International Desalination Association, pp. 5–1.
2. Water Desalting Committee of AWWA. (2004). *Water Desalting—Planning Guide for Water Utilities*. John Wiley & Sons, New York, p. 14.
3. Pankratz, T. and Tonner, J. (2003). *Desalination.com—An Environmental Primer*, Lone Oak, Houston, TX, p. 29.
4. AWWA. (1999). *Manual of Water Supply Practices—M46. Reverse Osmosis and Nanofiltration*, 1st Edn. American Water Works Association, Denver, CO, p. 21.
5. Watson, I.C. et al. (2003). *Desalting Handbook for Planners*. 3rd ed. Desalination and Water Purification Research Program Report No. 72, U.S. Bureau of Reclamation, p. 85.
6. Mickley, M.C. (2001). *Membrane concentrate disposal: Practices and regulation*. Desalination and Water Purification Research Program Report, U.S. Bureau of Reclamation, p. 229.
7. Voutchkov, N. (2004). The ocean—a new resource for drinking water. *Public Works*, June: 30.

DIATOMACEOUS EARTH FILTRATION FOR DRINKING WATER

Vipin Bhardwaj
NDWC Engineering Scientist

Mel J. Mirliss
International Diatomite
Producers Association

Diatomaceous Earth (DE) filtration is a process that uses diatoms or diatomaceous earth—the skeletal remains of small, single-celled organisms—as the filter media. DE filtration relies upon a layer of diatomaceous earth placed on a filter element or septum and is frequently referred to as pre-coat filtration. DE filters are simple to operate and are effective in removing cysts, algae, and asbestos from water. DE has been employed in many food and beverage applications for more than 70 years and was used specifically to filter potable water during WWII. Since then, it has been used to produce high-quality, low-cost drinking water. DE filtration is currently one of the

Previous Page

desalination produces permeate of 250 to 350 mg/L. More elaborate RO systems can produce water of quality matching that of a thermal desalination plant. Usually, the temperature of thermally desalinated water is 2 °C to 5 °C higher than the temperature of water produced by membrane separation.

Desalinated water (permeate) produced by both thermal and membrane desalination is highly corrosive and has to be chemically conditioned (typically by lime addition) to increase product water alkalinity and to adjust pH (usually by addition of chlorine dioxide and/or acid to the lime conditioned permeate) to meet potable water quality regulations.

CONCENTRATE MANAGEMENT

The two main streams produced by every desalination plant are low-salinity freshwater and high-salinity concentrate. The results of a recent study completed by the U.S. Bureau of Reclamation (6) on the concentrate disposal methods most widely used in the United States (in order of decreasing frequency) are shown in Table 1.

Sanitary sewer and surface water discharge are the two most popular and cost-effective methods for concentrate disposal. Depending on the site-specific conditions, deep well injection, evaporation ponds, and spray irrigation can also be competitive concentrate disposal alternatives. The zero liquid discharge system typically has the highest construction and operating costs. However, under specific circumstances, such as cold climate, low evaporation and soil uptake rates, high land costs, and low power costs, the zero liquid discharge system can be cost-competitive with evaporation pond and spray irrigation disposal.

Table 1. Concentrate Disposal Methods and Their Frequency of Use in the United States

Concentrate Disposal Method	Frequency of Use, % of Plants Surveyed
Surface water discharge	45
Sanitary sewer discharge	42
Deep well injection	9
Evaporation ponds	2
Spray irrigation	2
Zero liquid discharge	0

COST OF DESALINATED WATER

Desalination water costs depend on many factors, including the type of treatment technology; the source water quality; the target product water quality; the size of the desalination plant; and the costs of energy, chemicals, labor, and membranes. The all-inclusive cost of potable water produced using a brackish water source treated by membrane separation usually varies between US\$0.25 and \$0.65/m^3 (US\$1.0 to \$2.5/1000 gallons) of product water. Seawater desalination cost is estimated at US\$0.65 to \$1.20/m^3 (US\$2.5 to \$4.5/1000 gallons). For very small plants located in isolated areas, this cost could be significantly higher. For comparison, the cost of water produced from fresh natural water sources (i.e., low-salinity groundwater, lake, and river water) is typically between US\$0.15 and \$0.40/m^3 (\$0.6 and \$1.5/1000 gallons). This cost is driven mainly by the source water quality, the type of treatment technologies used, and the size of the water treatment plant. Developments in seawater desalination technology during the past two decades combined with the transition to construction of large capacity plants, colocation with power plant generation facilities, and enhanced competition by using the build-own-operate-transfer (BOOT) method of project delivery have resulted in a dramatic decrease in the cost of desalinated water (7). These factors are expected to continue to play a key role in further reducing the production cost of desalinated water.

BIBLIOGRAPHY

1. Wagnick Consulting. (2002). IDA *Worldwide Desalination Plant Inventory*. International Desalination Association, pp. 5–1.
2. Water Desalting Committee of AWWA. (2004). *Water Desalting—Planning Guide for Water Utilities*. John Wiley & Sons, New York, p. 14.
3. Pankratz, T. and Tonner, J. (2003). *Desalination.com—An Environmental Primer*, Lone Oak, Houston, TX, p. 29.
4. AWWA. (1999). *Manual of Water Supply Practices—M46. Reverse Osmosis and Nanofiltration*, 1st Edn. American Water Works Association, Denver, CO, p. 21.
5. Watson, I.C. et al. (2003). *Desalting Handbook for Planners*. 3rd ed. Desalination and Water Purification Research Program Report No. 72, U.S. Bureau of Reclamation, p. 85.
6. Mickley, M.C. (2001). *Membrane concentrate disposal: Practices and regulation*. Desalination and Water Purification Research Program Report, U.S. Bureau of Reclamation, p. 229.
7. Voutchkov, N. (2004). The ocean—a new resource for drinking water. *Public Works*, June: 30.

DIATOMACEOUS EARTH FILTRATION FOR DRINKING WATER

VIPIN BHARDWAJ
NDWC Engineering Scientist

MEL J. MIRLISS
International Diatomite
Producers Association

Diatomaceous Earth (DE) filtration is a process that uses diatoms or diatomaceous earth—the skeletal remains of small, single-celled organisms—as the filter media. DE filtration relies upon a layer of diatomaceous earth placed on a filter element or septum and is frequently referred to as pre-coat filtration. DE filters are simple to operate and are effective in removing cysts, algae, and asbestos from water. DE has been employed in many food and beverage applications for more than 70 years and was used specifically to filter potable water during WWII. Since then, it has been used to produce high-quality, low-cost drinking water. DE filtration is currently one of the

U.S. Environmental Protection Agency's (EPA) approved technologies for meeting the requirements of the Surface Water Treatment Rule (SWTR) and is most suitable for small communities that need to comply with the rule.

WHAT IS DE FILTRATION?

DE contains fossil-like skeletons of microscopic water plants called diatoms, which are a type of algae. These diatoms range in size from less than 5 micrometers to more than 100 micrometers, and have a unique capability of extracting silica from water to produce their skeletal structure. When diatoms die, their skeletons form a diatomite deposit. In its natural state, diatomite is 85 percent inert silica. The soluble portion of diatomite is extremely low (less than 1 percent). The odorless, tasteless, and chemically inert characteristics make DE safe for filtering water or other liquids intended for human consumption.

APPLICATION AND HISTORICAL BACKGROUND

During WWII, the U.S. Army needed a new type of water filter suitable for rapid, mobile military operations. The U.S. Army Engineer Research and Development Laboratories (ERDL) developed a DE filter unit that was lightweight, easily transported, and able to produce pure drinking water. Later, DE filtration technology was applied to filtering swimming pool water and more gradually to producing drinking water.

The earliest municipal DE filter installation was a 75,000 gallons per day (gpd) system in Campbell Hills, Illinois, that began operating in 1948. By 1977, municipalities had constructed more than 145 plants. Today, nearly 200 DE plants are successfully operating.

HOW DOES DE FILTRATION WORK?

DE filtration strains particulate matter from water, and the process rarely uses coagulant chemicals. First, a cake of DE is placed on filter leaves. A thin protective layer of diatomaceous earth builds up, or accumulates, on a porous filter septum (a permeable cover over interior collection channels) or membrane. Recirculating DE slurry through the filter septum establishes this layer. The septum is most often plastic or metallic cloth mounted on a wire mesh-covered steel frame. The DE process is also called pre-coat filtration because the solids separation at the start of a run takes place on the built-up pre-coat layer of DE.

After the pre-coat forms on the filter leaves (usually 1/8 inch thick) raw water containing a low dose of DE, which is called body feed, is fed through the filter. Particulate solids in the product flow are separated on the pre-coat surface. With such separation, the unwanted particulate matter actually becomes part of the filter media. During a filter run, removing particulate matter from raw water causes head loss to gradually build up in the filter. The accumulation of DE body feed on the filter reduces the rate of head loss. When maximum head loss is reached, the flow of water into the filter is stopped and the filter cake is cleaned.

High-pressure sprays, directed at the accumulated cake, detach the cake and provide dilution for draining the slurry suspension from the filter vessel. When cleaned, the filtration operation is repeated, beginning with the pre-coat cycle (see Fig. 1). Operators typically discard the DE removed from the filter leaves.

APPROPRIATE FEED WATER QUALITY AND PERFORMANCE CAPABILITIES

The use of DE filters is limited to treating source waters with an upper limit of turbidity at 10 NTU. Also, filtration rates range from 0.5 to 2 gallons per minute per square foot (gpm/ft^2). The particle size that DE filtration removes relies upon the size distribution of the DE particles used for the pre-coat and body feed. DE filters are very effective for removing *Giardia* and *Cryptosporidium* cysts. In some cases, studies have reported up to a 6-log reduction of these cysts under routine operating conditions. Because DE filtration usually does not involve coagulation, its potential for removing dissolved constituents, such as color, is low. Therefore, the utility or its engineer must determine raw water quality before considering DE filteration.

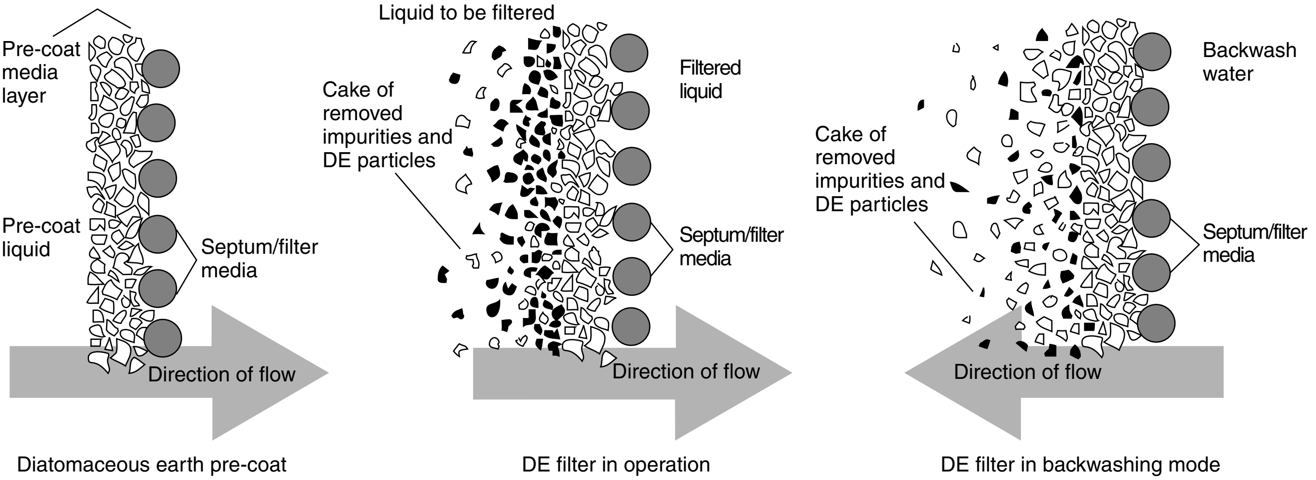

Figure 1. Diatomaceous earth filter. *Source*: Fulton, George P. 2000. *Diatomaceous Earth Filtration for Safe Drinking Water*.

WHAT ARE THE MONITORING AND OPERATING REQUIREMENTS?

Monitoring requirements for DE filtration are simpler than requirements for coagulation and filtration because operators rarely ever use coagulant chemicals for DE filtration. However, operators must continuously monitor raw and filtered water turbidity. Operators also must monitor filter head loss so that they can determine when to backwash the filter.

In general, DE filter plant operators need mechanical skills to operate the body feed pumps, pre-coat pumps, mixers, pipes, and valves. They also must be skilled in preparing the body feed and precoat slurries. Also, keeping DE filter leaves clean is of primary importance. A leaf filter that is not properly cleaned at the end of a filter run can accumulate dirt and slime on the filter cloth, which prevents a uniform pre-coat from forming when the filter is restored to service.

ELEMENTS OF A DE FILTER

Figure 2 shows the common elements in the manufacture of any flat leaf filter used in treating drinking water. The principal elements of a DE filter include the following:

- containment vessel,
- baffled inlet,
- filter leaves mounted on an effluent manifold,
- a method of cleaning the filter leaves at the end of a run,
- a drain to receive the backwash water,
- open top or access mode, and
- DE slurry preparation tank and pump feed.

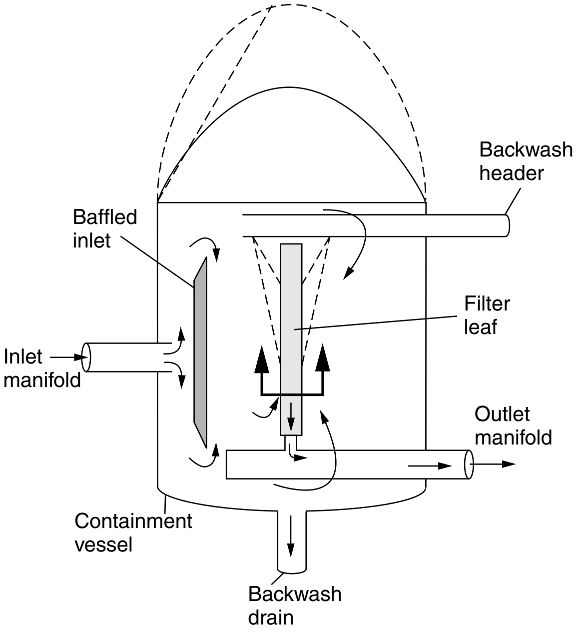

Figure 2. Elements of a flat leaf filter. *Source*: Fulton, George P. 2000. *Diatomaceous Earth Filtration for Safe Drinking Water*.

TYPES OF DE FILTRATION

Two types of DE filters exist: (1) *pressure filters*, which have a pump or high-pressure water source on the influent side and (2) *vacuum filters*, which have a pump on the effluent side. Vacuum filters are open to the atmosphere. Pressure filters are enclosed within pressure vessels.

The two basic groupings of DE filter designs are essentially defined by the hydraulic mode of operation, and are shown in (Fig. 3a and 3b).

The principal advantages of pressure filters over the vacuum filters are related to the significantly higher differential head available.

Table 1. Pressure and Vacuum Filters

Pressure Filters	Vacuum Filters
• Operates at higher flowrates, resulting in smaller, more compact filter units.	• Lower capital fabrication cost.
• Longer filter runs, reducing the use of pre-coat material and backwash water because of less frequent cleaning cycles.	• Lower maintenance costs.
• Less likelihood that gas bubbles will disrupt the media.	• Tanks are open at the top, making access and observation easy.

Source: Mel J. Mirliss, Vipin Bhardwaj, and the National Drinking Water Clearinghouse.

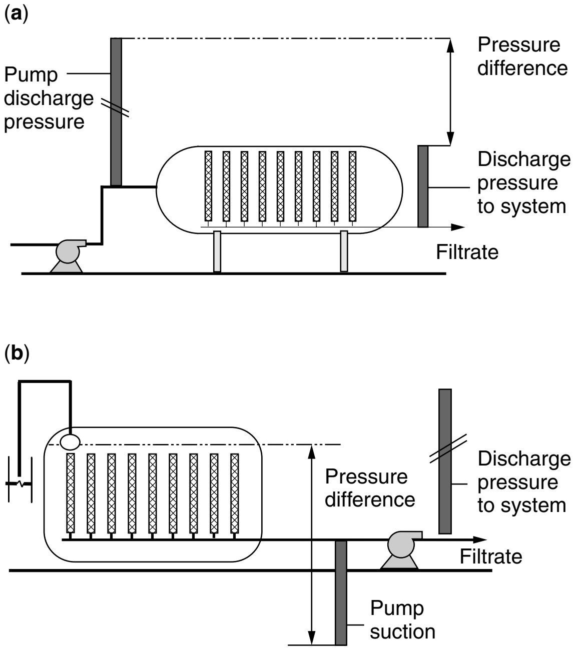

Figure 3. (**a**) Pressure filter. (**b**) Vacuum filter. *Source*: Fulton, George P. 2000. *Diatomaceous Earth Filtration for Safe Drinking Water*.

IS DE SUITABLE FOR SMALL SYSTEMS?

DE filtration is well-suited to small systems, because it does not require chemical coagulation, so operators do not need to learn about this complex aspect of water treatment. In addition, installation costs for DE systems are less than those for other technologies, such as membranes. DE filtration is currently one of the EPA's approved technologies for meeting SWTR requirements.

An ideal, cost-effective DE filtration application is for well water supplies under the influence of surface waters, but that are otherwise acceptable in quality. Superior cyst removal capability makes the DE filter more advantageous than other alternatives. However, there is a potential difficulty in maintaining complete and uniform thickness of DE on the filter septum.

WHERE CAN I FIND MORE INFORMATION?

- Fulton, George P., P.E. (2000). *Diatomaceous Earth Filtration for Safe Drinking Water*. American Society of Civil Engineers, Reston, VA.
- "Precoat Filtration." 1988. AWWA M30, *Manual of Water Supply Practices*. American Water Works Association, Denver, CO.
- *Technologies for Upgrading or Designing New Drinking Water Treatment Facilities*, EPA/625/4-89/023. (Available from the EPA) or contact the National Drinking Water Clearinghouse.

HAVE YOU READ ALL OUR TECH BRIEFS?

Tech Briefs drinking water treatment fact sheets have been a regular feature in the National Drinking Water Clearinghouse (NDWC) newsletter *On Tap* for more than five years.

A package of *Tech Briefs* is now available as a product. A three-ring binder holds all the current *Tech Briefs* in print. New selections can be easily added to the package as they become available. To order this free product, call the NDWC at the numbers listed below and ask for item #DWPKPE71.

Additional copies of fact sheets are also free; however, postal charges may be added.

To order, call the NDWC at (800) 624-8301 or (304) 293-4191. You also may order online at ndwc_order@mail.nesc.wvu.edu *or download fact sheets from our Web site at* www.ndwc.wvu.edu.

Also, the NDWC's Registry of Equipment Suppliers of Treatment Technologies for Small Systems (RESULTS) is a public reference database that contains information about technologies used by small waters systems around the country. For further information about accessing or ordering RESULTS, call NDWC.

READING LIST

Fulton, George P. 2000. *Diatomaceous Earth Filtration for Safe Drinking Water*. American Society of Civil Engineers, Reston, VA.

Safe Water from Every Tap, 1997. *Improving water service to small communities.* National Research Council, National Academy Press, Washington, DC.

Schuler, P.F., Gosh M.M., and Boutros S.N. 1988. *Comparing the removal of Giardia and Cryptosporidium using slow sand and diatomaceous earth filtration.* pp. 789–805, *Proceedings of 1988 AWWA Conference*, Denver, CO.

Ongerth, J.E., Hutton, PE December 1997. DE filtration to remove *Cryptosporidium*. *Journal of American Water Works Association*. American Water Works Association, Denver, CO.

Small System Compliance Technology List for the Surface Water Treatment Rule and Total Coliform Rule, 1998. U.S. Environmental Protection Agency, 815-R-98-001.

Vipin Bhardwaj, technical assistance specialist, has a B.S. in chemical engineering and has recently finished his master's degree in agriculture at West Virginia University. He expects to receive a master's degree in civil engineering this fall.

Mel Mirliss has been associated with the diatomaceous earth industry for the past 28 years. For the past nine years, he has served as the executive director of the International Diatomite Producers Association (IDPA) where he has been a director since the association's founding in 1987. He holds a B.S. in chemistry from the University of California, Berkeley.

EMERGING WATERBORNE INFECTIOUS DISEASES

Louis H. Nel
University of Pretoria
Pretoria, Gauteng, South Africa

Humanity is plagued at present by at least 1709 different infectious diseases. These diseases are caused by pathogens, which are microscopic and parasitic organisms of diverse natures, including infectious proteins (prions), viruses, bacteria, fungi, and protozoa. Compared to diseases like cancer and metabolic diseases that have genotypic roots, including heart disease, diabetes, and the like, progress in the fight against infectious diseases has been rapid during the last half century. Whereas infectious diseases such as smallpox, measles, polio, rabies, plague and numerous other bacterial diseases have once been a global scourge, many of these diseases have now become controllable due to spectacular advances in public health practices, including improvements in the quality of water and sanitation and by immunization, education, early diagnosis, and the use of antibiotics/antimicrobials and other drug therapies. However, globally, 45% of all deaths and 63% of early childhood deaths are still caused by infectious diseases.

INFECTIOUS DISEASES, EMERGING DISEASES, AND ZOONOSES

New, emerging, reemerging, and resurgent infections contribute significantly to the infectious disease problems that the world is experiencing now. At least 156 such diseases have now been identified. Taken from a report by the Institute of Medicine (1), emerging infections may be defined as follows: "emerging infections are those whose incidence in humans has increased within the past two decades or threatens to increase in the near future. Emergence may be due to the spread of a new agent, to the recognition of an infection that has been present in the population but has gone undetected, or to the realization that an established disease has an infectious origin. Emergence may also be used to describe the reappearance (or "reemergence") of a known infection after a decline in incidence." Generally, new diseases are synonymous with emerging diseases and can be regarded as diseases first described in the last 10 to 30 years. Most of these diseases are caused by specific modifications (mutation/evolution, species jump, etc.) of agents that are already in the environment. Reemerging diseases are those that have persisted at a subdued level in the population and recur as a result of antimicrobial drug resistance or other changes that might favor marked increases in disease incidence. Reemerging diseases can also be described as resurgent, pertinently referring to an abrupt increase in incidence or geographic distribution of the particular disease. The emergence and reemergence of diseases are clearly related to changes in the infectious pathogen, the vector or transmission system, and the host population. Such epidemiologically important changes may include drug resistance and mutations resulting in increased virulence, changes in the distribution or activity of vectors, globalization and increased travel, war, population explosions, climatic and ecological changes, geographical displacement of species, movement into previously uninhabited areas, poverty and breakdown of health care systems, and changes in agriculture and industrialization. Six infectious diseases can be identified as current, leading, high mortality rate diseases, and all of them conform to the criteria for emergence described before: AIDS, acute respiratory infection, diarrheal diseases, malaria, measles, and tuberculosis.

A great many infectious diseases are zoonoses: These are diseases that are naturally transmitted between vertebrate animals and man—49% or 832 of all known infectious diseases can be considered zoonoses. A fair number of infectious diseases of man started out as zoonotic events but have since established a much more important human-to-human epidemiological cycle (e.g., HIV-AIDS, dengue hemorrhagic fever). Strikingly, however, 73% of all the emerging diseases considered (114 of 156) are zoonoses. Infectious diseases are typically transmitted through direct contact, insect vectors, sexual contact, respiratory tract (aerosolized microbes), and by contaminated food and water. Here we are concerned with waterborne infection, with pathogen entry through the mouth and alimentary track or, in some instances, through epithelial cells of the respiratory system.

WATERBORNE INFECTIOUS DISEASES: THE SPECTRUM OF PATHOGENS

Human and animal fecal pollution of water sources is the leading cause of waterborne infections. This fecal–oral route of infection contributes to hundreds of millions of cases of diarrhea and millions of deaths (particularly among children) every year. Clearly, improved water sanitation that leads to safe drinking water for all of humanity would be the ultimate solution. Some emerging waterborne infections do not enter through the oral route but through the respiratory route. These infections are attributed to specific pathogens that have been made airborne through water spraying in nurseries, flower and vegetable markets, personal showers, and the like.

Viral Agents

Many waterborne diseases are caused by viruses, which include adenoviruses (types 40 and 41), astroviruses, caliciviruses (including Norwalk and hepatitis E viruses), Enteroviruses (picornaviruses), and reoviruses (including rotaviruses). By far the most common medical condition associated with waterborne viral infection is diarrhea. The medical and economic importance of viral diarrhea should not be underestimated, millions of deaths (particularly, of children) are caused by these infections. At present, only a handful of these viral diseases is considered emerging and the most important of them, astrovirus enteritis, calicivirus enteritis, and hepatitis (E), are specifically discussed here.

Bacterial Agents

Increased resistance to antibiotics is a major factor in the resurgence of common bacterial infectious diseases and epidemic bacterial diarrhea such as those caused by waterborne *Escherichia coli, Shigella*, and *Vibrio*. The use of antimicrobials in agriculture has significantly contributed to this phenomenon, although evolution and adaptation of organisms have also led to converting nonpathogens into pathogens by adding of toxin producing capability. In addition, new bacterial zoonoses emerge, and some human pathogens are newly recognized, thanks to rapidly improving diagnostic techniques. This, as well as the reemergence of pathogens due to immune deficiencies related to HIV/AIDS, applies to the whole spectrum of infectious disease agents.

Protozoal Agents

These parasites typically are contaminants of potable water supplies. Whereas inadequate sanitation and chlorination of water supplies should prevent most of the waterborne bacterial diseases, large outbreaks of protozoan enteritis are usually associated with surface water supplies that are inadequately flocculated and filtered or not treated at all (many of these pathogens are chlorine resistant). In the last decade, enteric protozoa have become the leading cause of waterborne disease outbreaks for which an etiologic agent could be identified. The most common of these infections are caused by *Cryptosporidium* and *Giardia*, *Cyclospora* has also recently emerged as an enteric protozoan.

SPECIFIC EMERGING WATERBORNE INFECTIOUS DISEASES

Aseptic Meningitis and Various Other Syndromes

These diseases are caused by enteroviruses, viruses that—as their name suggests, replicate in the intestinal tract. Enteroviral infection is one of the most common types of viral infections in humans, but they do not often result in serious disease, although some serotypes of enteroviruses may cause serious clinical syndromes. Such syndromes may include acute paralysis, encephalitis, meningitis, myocarditis, hepatitis, and chronic infection (particularly in immunocompromised individuals). Enteroviruses belong to the *Enterovirus* genus of the *Picornaviridae* family and include the polioviruses, group A Coxsackieviruses, group B Coxsackieviruses, echoviruses, and (newer) enteroviruses. There are 3 different polioviruses, 61 nonpolio enteroviruses, 23 Coxsackie A viruses, 6 Coxsackie B viruses, 28 echoviruses, and 4 other enteroviruses. Enteroviruses evolve rapidly and emerge as new strains, but others are of emerging importance due to drug resistance. Some enterovirus serotypes have been responsible for serious, large epidemics throughout the world in recent years. Echovirus infections have been responsible for several large epidemics of aseptic meningitis in Japan, Europe, and the Middle East in the past years. These epidemics involved different highly infectious serotypes of the virus, types 4, 9, and 30, but also of new genetic variants of these viral serotypes. Enterovirus 71 (EV71) has caused major disease outbreaks in North America, Europe, Malaysia, Japan, and Australia since 1995. In 1998, this virus infected at least a million individuals in Taiwan and caused a 20% case fatality among children under 5 years of age.

Astrovirus Enteritis

Today, astroviruses are considered much more prevalent than previously thought (e.g., 75% of 5 to 10-year-old children in Britain demonstrate antibodies against astroviruses). These are small, round viruses that have a starlike appearance under the electron microscope. Following a 1 to 4-day incubation period, the clinical symptoms present as watery diarrhea that lasts 2–3 days. These gastrointestinal disorders are usually not serious, but astroviruses are a leading cause of childhood diarrhea, and dehydration may be especially severe in immunocompromised individuals.

Calicivirus Enteritis

This disease, caused by a group of small cuplike viruses, is very common—serological evidence indicates that most people are infected by age 12. The symptoms are diarrhea with nausea and vomiting. Fever and respiratory infection occur in a small number of cases. It is thought that these viruses emerged from ocean reservoirs, with subsequent zoonotic and interspecies movement. Among the known calicivirus pathogens of humans are the Norwalk and Norwalk-like viruses and the Sapporo and Sapporo-like viruses. Although it is thought that pathogenic caliciviruses are likely to continue emerging from the world's oceans in various forms, only the Norwalk-like viruses are now and from a public health point of view, of emerging importance. Norwalk-like viruses cause sporadic and epidemic gastroenteritis in all age groups of humans, but losses of fluids and electrolytes are particularly serious in the very young and elderly. Transmission is by the fecal–oral route, and shellfish-related outbreaks frequently occur where appropriate standards for proper disposal of human sewage are not applied. Contaminated water, ice, eggs, salad ingredients, and ready-to-eat foods are other sources of infection.

Campylobacteriosis

This diarrheal disease, caused by *Campylobacter jejuni*, occurs worldwide. Infection by *Campylobacter* species has been known in livestock for many decades, but it has also been recognized as an important pathogen in humans since the middle 1970s. The development of selective media contributed to this recognition, and drug resistance and the emergence of toxin producing strains have recently become of considerable importance. Although this disease is usually self-limiting, treatment with antibiotics may reduce bacterial numbers and the duration of the infection and may be of particular value in cases involving immunocompromised individuals or where complications are evident. From this perspective, the emergence of antibiotic resistant strains of *Campylobacter* all over the world is disconcerting. *Campylobacter jejuni* infections, it is thought are more common than infections by other enteric bacteria such as *Shigella, Salmonella* and *Escherichia coli* O157:H7. Whereas other *Campylobacter* species (e.g., *C. laridis, C. hyointestinalis*) may also be involved in disease, *C. jejuni* is responsible for 99% of the cases. The bacterium is gram-negative, has a curved rod-shaped morphology and is motile. It is highly infectious and symptoms are diarrhea, fever, abdominal pain, nausea, headache, and myalgia. Complications are rare, but the infection may progress toward reactive arthritis, hemolytic uremic syndrome, and septicemia. Guillain-Barre syndrome is also recognized as a rare (1 per 1000 infected individuals) complication leading to disease of the peripheral nervous system.

Cholera

Cholera is caused by the gram-negative, motile, rod-shaped bacterium, *Vibrio cholerae*. The disease, essentially an acute and serious infection of the small intestine with diarrhea that can be severe and is accompanied by dehydration and shock, occurs worldwide. The symptoms associated with cholera are caused by the enterotoxins produced by V*ibrio cholerae* during the infection. Accurate records of cholera date back almost a century, but cholera may have been known in India and elsewhere for thousands of years. This disease may be considered emerging due to new pandemics (serotype O1) and the appearance of pathogenic strains that have evolved by genetic recombination (serotype 0139). The emerging serotype O139 appeared in 1992 and has as yet only been reported from Southeast Asia where it causes epidemics in populations that were previously exposed to other serotypes of the same pathogen. In similar evolutionary fashion, *Vibrio cholerae* can become resistant to antibiotics by acquiring a transposon element and/or a plasmid that confers resistance to multiple antibiotics.

Vibrio cholerae O1, however, caused a pandemic spread over six continents during the past 40 years. The bacterium initially (1960s) spread from Indonesia to Eastern Asia, India, the former USSR, Iran, and Iraq.

Thereafter, in 1970, serotype O1 took hold in West Africa and then became endemic in most of the African continent. In 1991, it spread throughout South and Central America, and involved more than a million known cases at a mortality rate of about 1%. Much higher mortality rates (33%) were observed in 1994 in a large epidemic among Rwandan refugees in Zaire (33%). In recent times, Africa continued to host large cholera epidemics. These included a 1997 outbreak in Kenya, Tanzania, and Mozambique, a 1998 outbreak in Congo and Zaire, and in 2000/2001, a large epidemic in South Africa.

Cryptosporidiosis

The etiological agent of this persistent diarrheal disease is the protozoal parasite *Cryptosporidium parvum*. Together with the AIDS epidemic, cryptosporidiosis emerged in the early 1980s as a zoonosis, after having been known as a veterinareal disease for more than a century. Apart from persistent diarrhea, symptoms may include abdominal pain, nausea, and low-grade fever, and significant weight loss. Infectious oocysts may still be excreted long after diarrheal disease has ended (up to 5 weeks). In the growing number of immunocompromised individuals worldwide, this has become a life-threatening and highly infectious enteric disease. Compounding the threat of this disease is the fact that the pathogen is difficult to filter from water resources, is resistant to chlorine, is ubiquitous in many animals, can survive in infected water for long periods of time, and is highly infectious. It is therefore not surprising that *C. parvum* has now become the major protozoan pathogen of humans. A wide spectrum of water supplies has been implicated and these include contaminated and inefficiently treated/flocculated/filtered drinking water such as from water plants and boreholes or from swimming pools, wading pools, hot tubs, jacuzzis, fountains, lakes, rivers, springs, ponds, and streams. A high rate of secondary person-to-person transmissions in households and institutions has also been documented.

Cyclosporiasis

This protracted, relapsing gastroenteritic disease is similar to cryptosporidiosis and is also caused by a coccidian parasite, in this case, *Cyclospora cayetanensis.* This organism is food- and waterborne, and it emerged in the 1990s as a serious and widespread gastrointestinal parasite. In contrast to the zoonotic *C. parvum*, *C. cayetanensis* is specific to human hosts and is not a known pathogen of vertebrate animals. Cyclosporiasis is most common in tropical and subtropical regions and endemic in Central and South America, Southeast Asia, the Caribbean islands, and parts of Eastern Europe, but large outbreaks have recently been reported in the United States and Canada, reports also came from the United Kingdom and various African countries. Apart from fecally contaminated water, various types of fresh produce such as raspberries, basil, field greens, and salad mixes have been indicated as sources of the pathogen. Although *Cyclospora* oocysts, like *Cryptosporidium* oocysts, are resistant to chlorine, they are double the size of *Cryptosporidium* oocysts and may be more easily removed by flocculation and filtration methods used in water plants. Further, due to the noninfectious nature of newly formed oocysts, secondary person-to-person spread of *Cyclospora cayetanensis* is a much less likely route of transmission than that for cryptosporidiosis.

Dermatitis

This is an inflammation of the skin marked by redness, pain, and itching (skin rashes) that is caused by the gram-negative and rod-shaped bacterium, *Pseudomonas aeruginosa*. Until the recent recognition of waterborne dermatitis caused by *Pseudomonas aeruginosa*, this organism was rarely implicated in disease. The bacterium typically survives in biofilms and antibiotic resistant strains are also appearing, but effective control should be possible by proper water treatment.

Gastritis

Helicobacter pylori, a spiral gram-negative bacterium was first recognized in 1982 as a cause of gastritis. It appears that the human stomach is the only recognized reservoir for this bacterium and infection is common in the general population. Although most infected individuals are asymptomatic and are unlikely to develop a serious medical problem from the infection, *H. pylori* causes 90% of duodenal ulcers and up to 80% of gastric ulcers, and infected persons have a twofold to sixfold increased risk of developing gastric cancer. In fact, *H. pylori* is classified as a group I (or definite) carcinogen by the World Health Organization's International Agency for Research on Cancer. It is not known why some patients become symptomatic and others do not. The transmission of *H. pylori* also remains unclear, although the bacteria are most likely to spread from person to person through fecal–oral or oral–oral routes.

Giardiosis

This disease, one of the most common diarheal diseases spread by drinking and recreational water in the United States and probably worldwide, is caused by the intestinal protozoan parasite *Giardia intestinalis*, also known as *G. lamblia* or *G. duodenalis*. This agent also infects domestic and wild animals (e.g., cats, dogs, cattle, deer, and beavers). Emerging as a widespread and common disease during the past two decades, giardiasis has been traced back to contaminated swimming pools, hot tubs, fountains, lakes, rivers, springs, streams and ponds. As for the other parasitic diseases discussed here, the life cycle begins when cysts are ingested through person-to-person transmission or ingestion of fecally contaminated food or water. The ingested cysts release trophozoites in the duodenum where they attach to the surface of the intestinal epithelium. *Giardia* cysts can be excreted in fecal stools intermittently for weeks or months, and infection of a new host can result from ingestion of as few as 10 cysts.

Hemorrhagic Colitis and Complications

Enterohemorrhagic *Escherichia coli* (EHEC) is another bacterial pathogen that has recently emerged as a major waterborne enteric pathogen. After recognition of waterborne infection by this agent in 1985, the disease has been reported worldwide, and case numbers are ever increasing. Whereas *E. coli* bacteria are usually symbiotic as part of the normal intestinal flora of animals and humans, some strains of the organism can cause disease by producing large quantities of Vero toxins. It is a bacterial virus (or bacteriophage) that carries the toxin gene. When infecting the bacteria, the bacteriophage, integrates into the bacterial genome and from there provides the toxin producing capability. Today, a large number of *E. coli* serotypes that can produce Vero toxins are recognized, but the serotype O157:H7 is the predominant pathogen that is most frequently associated with hemorrhagic colitis in humans. The disease is serious, usually characterized by bloody diarrhea, abdominal cramps and nausea, but more severe complications, known as hemolytic uremic syndrome (HUS)—leading to complete renal failure, may follow. Case fatality rates for *E. coli* O157:H7 range from 3–36%, the fate of patients depends on age and immunocompetence.

Waterborne transmissibility through both drinking and recreational waters together with antibiotic and chlorine resistance and an ability to survive in water for long periods of time, are major factors in the widespread emergence of EHEC. In addition, these organisms are highly infectious and are often spread to humans as zoonoses, EHEC has a very wide host range among domestic animals and wildlife. Outbreaks continue to be recognized worldwide in the United States, Japan (1996—6000 cases), United Kingdom, Australia, Europe, Argentina, Chile, and throughout Africa (e.g., South Africa, Swaziland, Kenya, Nigeria, Cote d'Ivoire, Central Africa Republic, and Egypt). In recent years, a number of efficient diagnostic techniques have been developed for accurate detection of isolate *E. coli* O157:H7, which also allows for distinction from other bacterial enteropathogens (e.g., *Shigella*) with which it may easily be confused due to clinical similarities of disease manifestation.

Hepatitis (Hepatitis E-Like Viruses)

Hepatis virus E (HEV) was isolated for the first time in 1988. This virus was then characterized in subsequent years, and it has recently been placed in its own taxonomic group, "hepatitis E-like viruses", within the class IV (+) sense RNA viruses. The disease is found most frequently in the developing world (e.g., South East Asia, India, Middle East, Africa, and Central America) and is transmitted primarily through the fecal–oral route in contaminated drinking water. Person-to-person transmission of HEV may occur during large epidemics, but this route of transmission appears to be uncommon. The incubation period following exposure to HEV is, on average a lengthy 40 days but may range from 15 to 60 days (mean, 40 days). Diarrhea is a less common symptom in this case, malaise, anorexia, abdominal pain, arthralgia, fever, anorexia, hepatomegaly, jaundice, nausea, and vomiting are more common. The disease is most often seen in young to middle-aged adults (15–40 years old) where the fatality rate is relatively low at 0.1–1%. However, pregnant women appear to be exceptionally susceptible to severe disease, and excessive mortality has been reported in this group (up to 30%). Recent genetic and serological evidence suggests frequent transmission of HEV between animals and humans, and zoonoses may explain the mechanism of HEV maintenance in populations during the periods between epidemic outbreaks.

Legionnaires' Disease

Legionnaires' disease, or legionnellosis, is an acute pneumonia caused by the bacterium *Legionnella pneumophila*. This is a flu-like illness, typically occurring 3 days after exposure and followed within a week by high fever, chills, dry cough, muscle aches and headache. The disease is usually self-limited but can lead to higher mortality in the presence of various risk factors such as immunodeficiency, cigarette smoking, chronic lung diseases, lung malignancies, heart disease, and old age. When this disease was first recognized in the late 1970s, the associated etiological agent was hard to isolate, given that *Legionnella* is a particularly fastidious organism that does not grow on typical culture media. The importance of these waterborne gram-negative, rod-shaped bacilli has since become clear, following the development of suitable diagnostic techniques. Of importance is the fact that these bacteria are parasites of protozoa (e.g., *Amoeba* sp.) that are common in rivers, lakes, and streams. Within these hosts, *Legionnella* will proliferate and be protected from environmental hazards, including chlorination. In addition, following intracellular replication in protozoal cells, *L. pneumophila* converts to a virulent form and expresses a number of traits (including enhanced motility) that will assist in extracellular survival and transmission to new cells. This phenomenonant is thought, explains the virulence of *L. pneumophila* for the macrophages in the human lung, given the similarities in cell biology of these cells and protozoal cells. If the virulent form of the bacteria is inhaled, replication may occur within the alveolar macrophages, and disease will ensue. This transmission occurs most easily in the presence of airborne water droplets, so Legionnellosis has become particularly important and common following the various practices leading to aerosolization of water (e.g., showers, spraying of produce in large markets, and air conditioning).

Shigellosis Colitis

Shigella is a highly infectious, gram-negative, rod-shaped bacterium, most species are implicated in waterborne infections. Although not new, it is the virulence and ability to develop drug resistance through plasmids that confers resistance to a large spectrum of antibiotics, which characterize a resurgence of shigellosis colitis. Disease symptoms include abdominal pains, fever, and rectal pain, and complications may include sepsis, seizure, renal failure, and hemolytic uremia syndrome. *Shigella dysenteriae type 1* (Sd1) is the cause of epidemic dysentery that usually originates from water polluted with human feces. This type is a major cause of dysentery in Asia and Central and South America and has only recently arrived in Africa. After major pandemics in southern Africa in the early and middle 1990s, *Shigella dysenteriae type 1* has now become endemic in many countries in Africa. Other species of *Shigella* that are frequently responsible for waterborne infections are *Shigella sonnei* and *Shigella flexneri.*

Toxoplasmosis

This infection, which is symptomless in most cases, is caused by the coccidian protozoan *Toxoplasma gondii*. However, the parasite is considered an important emerging pathogen in the AIDS patient group worldwide, where infection can cause lymphadenopathy, central nervous system disorders, myocarditis, pneumonitis, and cranial lesions. Domestic cats are the definitive host for *Toxoplasma* but, apart from humans, there are numerous additional intermediate hosts in the animal kingdom. Infection takes place by ingesting contaminated food and water or by inhaling airborne *Toxoplasma* oocysts.

Tuberculosis (TB)

Although most individuals infected with *Mycobacterium tuberculosis* are asymptomatic carriers of TB, this should be considered a serious long-term pulmonary disease. In symptomatic cases, listlessness, chest pain, loss of appetite, fever, and weight loss are early symptoms of pulmonary tuberculosis. This may progress into night sweats, bleeding in the lungs, coughing up of sputum with pus, and shortness of breath. Exposure usually occurs through inhalation of infected droplets and thus, similar to legionnellosis, the aerosolization of water contributes to the transmissibility of *Mycobacterium* where water supplies are contaminated with the organism (often in hospitals). Tuberculosis is not a new infection, but multidrug-resistant tuberculosis (MDRTB) has become a major public health problem during the last 10 years, particularly as an opportunistic infection in concert with the AIDS epidemic.

CONCLUDING REMARKS

In considering the reasons for the emergence of the specific waterborne infectious diseases described, it is easy to identify a number of common and very specific causal elements. These include (1) the rising incidence of diseases or other factors that affect immune competence in groups or populations, including HIV/AIDS, malnutrition, and the like; (2) genetic adaptation of pathogens, including the acquisition of resistance plasmids, toxin genes, and other virulent factors; (3) overcrowding combined with poor sanitation and deteriorating or inadequate public health infrastructure; (4) increased exposure to animal reservoirs (as a result of geographical expansion, overpopulation, and competition for dwindling water resources) and continual zoonoses. Based on these factors, the potential for the continued emergence and reemergence of infectious diseases seems inexhaustible. Advances in vaccination, therapeutic intervention, and surveillance are crucial elements in the fight against these diseases, but even so, much may be achieved through education, improved hygiene, and proper water sanitation.

BIBLIOGRAPHY

1. Ackelsberg, J., Sivapalasingam, S., Olsen, S.J. et al. (1999). *A Large Waterborne Outbreak of* Escherichia coli *O157:H7 and* Campylobacter *Infections in County Fair Attendees*. NY. Presented at the 37th Annual meeting of the Infectious Disease Society of America, Philadelphia, PA, November 18–21, 1999. Abstract 728.
2. Allos, B.M. (2001). *Campylobacter jejuni* infections: update on emerging issues and trends. *CID* **32**: 1201–1206.
3. Barrett, T.J. et al. (1994). Laboratory investigation of a multistate food-borne outbreak of *Escherichia coli* O157:H7 by using pulsed Field gel electrophoresis and phage typing. *J. Clin. Microbiol.* **32**: 3013–3017.
4. Basu, A. et al. (2000). *Vibrio cholerae* O139 in Calcutta, 1992—1998. *Emerging Infect. Dis.* **6**(2): 139–147.
5. Berke, T. et al. (1997). Phylogenetic analysis of the Caliciviruses. *J. Med. Virol.* **52**: 419–424.
6. Chalmers, R.M., Aird, H., and Bolton, F.J. (2000). Waterborne *Escherichia coli* 0157. *J. Appl. Microbiol. Symp. Supple.* **86**: 124S–132S.
7. Colwell, R.R. (1996). Global Climate and infectious disease: The Cholera paradigm. *Science* **274**: 2025–2031.
8. DaSilva, E. and Iaccarino, M. (1999). Emerging diseases: A global threat. *Biotechnol. Adv.* **17**: 363–384.
9. Emmerson, A.M. (2001). Emerging waterborne infections in health-care settings. *Emerging Infect. Dis.* **7**(2): 272–276.
10. Enserink, M. (2000). Malaysian Researchers trace Nipah virus outbreaks to bats. *Science* **289**: 518–519.
11. Ewald, P.W. (1991). Waterborne transmission and the evolution of virulence among gastrointestinal bacteria. *Epidemiol. Infect.* **106**: 83–119.
12. Ewald, P.W. (1998). The evolution of virulence and emerging diseases. *J. Urban Health* **75**(3): 480–492.
13. Farmer, J.J. and Davis, B.R. (1985). H7 antiserum-sorbitol fermentation medium: A single-tube screening medium for detecting *Escherichia coli* O157:H7 associated with hemorrhagic colitis. *J. Clin. Microbiol.* **22**: 620–625.
14. Feng, P. (1995). *Escherichia coli* serotype O157:H7: Novel vehicles of infection and emergence of phenotypic variants. *Emerging Infect. Dis.* **1**(2): 47–52.
15. Fields, P.I. et al. (1997). Molecular characterization of the gene encoding H antigen in *Escherichia coli* and development of a PCR restriction fragment length polymorphism test for identification of *E. coli* O157:H7 and O157: NM. *J. Clin. Microbiol.* **35**: 1066–1070.
16. Fratamico, P.M., Buchanan, R.L., and Cooke, P.H. (1993). Virulence of an *Escherichia coli* O157:H7 sorbitol positive mutant. *Appl. Environ. Microbiol.* **59**: 4245–52.
17. Goma Epidemiology Group. (1995). Public health impact of Rwanda refugee crisis: what happened in Goma, Zaire, in July 1994, *Lancet* **345**: 339–344.
18. Gomez-Lus, R., Clavel, A., Castillo, J., Seral, C., and Rubio, C. (2000). Emerging and re-emerging pathogens. *Int. J. Antimicrobial Agents* **16**: 335–339.
19. Beckett, G. et al. (2000). Pseudomonas Dermatitis/Folliculitis Associated with Pools and Hot Tubs—Colorado and Maine, 1999—2000. Morbidity and Mortality Weekly Report. *Centres for Disease Control and Prevention MMWR* **49**(48): 1087–1091.
20. Besser, R.E., Griffin, P.M., and Slutsker, L. (1999). *Escherichia coli* 0157:H7 gastroenteritis and the haemolytic uremic syndrome: An emerging infectious disease. *Annu. Rev. Med.* **50**: 355–367.
21. Herwaldt, B.L. (2000). *Cyclospora cayatanensis*: A review, focusing on the outbreaks of Cyclosporiasis in the 1990s. *Clin. Infect. Dis.* **31**: 1040–57.

22. Hunter, P.R. et al. (2001). Waterborne diseases. *Emerging Infect. Dis. Supple.* **7**(3): 544–545.
23. International Agency for Cancer Research. (1994). *Shistosomes, liver flukes and Helicobacter pylori. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*. IARC, Lyon, 61.
24. Iverson, E.R. et al. (1998). Epidemic *Shigella dysenteriae* in Mumiao, western Kenya. *Trans. Roy. Soc. Trop. Med. Hyg.* **92**: 30–31.
25. Jaykus, L. (1997). Epidemiology and detection as options for control of viral and parasitic foodborne disease. *Emerging Infect. Dis.* **3**(4): 529–539.
26. Karas, J.A., Pillay, D.G., Naicker, T., and Storm, A.W. (1999). Laboratory surveillance of *Shigella dysenteriae* type 1 in Kwa Zulu-Natal. *S. Afr. Med. J.* **89**: 59–63.
27. Karmali, M.A. (1989). Infection by Verocytotoxin-producing *Escherichia coli*. *Clin. Microbiol. Rev.* **2**: 15–38
28. Keene, W.E. et al. (1994). A swimming associated outbreak of hemorrhagic colitis caused by *Escherichia coli* O157:H7 and *Shigella sonnei*. *N. Engl. J. Med.* **331**: 579–84.
29. Lederberg, J. (1997). Infectious disease as evolutionary paradigm. *Emerging Infect. Dis.* **3**(4): 417–423.
30. Minshew, P. et al. (2000). Outbreak of gastroenteritis associated with an interactive water fountain at a Beachside Park—Florida, 1999. Morbidity and Mortality Weekly Report. *Centres for Disease Control and Prevention MMWR* **49**(25): 565–568.
31. Morse, S.S. (1995). Factors in the emergence of infectious diseases. *Emerging Infect. Dis.* **1**(1): 7–15.
32. Olsvik, O. et al. (1991). A nested PCR followed by magnetic separation of amplified fragments for detection of *Escherichia coli* Shiga like toxins genes. *Mol. Cell Probes* **5**: 429–35.
33. Pollard, A.J. and Dobson, S.R. (2000). Emerging infectious diseases in the 21st century. *Curr. Opin. Infect. Dis.* **13**: 265–275.
34. Satcher, D. (1995). Emerging infections: getting ahead of the curve. *Emerging Infect. Dis.* **1**(1): 2–6.
35. Shears, P. (2000). Emerging and re-emerging infections in Africa: the need for improved laboratory services and disease surveillance. *Microbes Infection* **2**: 489–495.
36. Smith, A.W. et al. (1998). Calicivirus emergence from ocean reservoirs: zoonotic and interspecies movements. *Emerging Infect. Dis.* **4**(1): 13–20.
37. Stehr-Green, J. et al. (1991). International notes campylobacter enteritis—New Zealand, 1990. Morbidity and Mortality Weekly Report. *Centres for Disease Control and Prevention MMWR* **40**(7): 116–117.
38. Swanson, M.S. and Hammer, B.K. (2000). *Legionella pneumophila* pathogenesis: fateful journey from Amoebae to macrophages. *Annu. Rev. Microbiol.* **54**: 567–613.
39. Swerdlow, D.L. et al. (1992). A waterborne outbreak in Missouri of *Escherichia coli* O157:H7 associated with bloody diarrhoea and death. *Ann. Intern. Med.* **117**: 812–819.
40. Szewzyk, U., Szewzyk, R., Manz, W., and Schleifer, K.H. (2000). Microbiological safety of drinking water. *Annu. Rev. Microbiol.* **54**: 81–127.
41. Vogt, R.L. et al. (1982). *Campylobacter enteritis* associated with contaminated water. *Ann. Intern. Med.* **96**: 292–296.
42. Walford, D. and Noah, N. (1999). Emerging infectious diseases—United Kingdom. *Emerging Infect. Dis.* **5**(2): 189–194.
43. Willshaw, G.A. et al. (1993). Examination of raw beef products for the presence of Vero cytotoxin producing *Escherichia coli*, particularly those of serogroup O157. *J. Appl. Bacteriol.* **75**: 420–426.
44. World Health Organisation. (1995). *Guidelines for the Control of Epidemics Due to* Shigella dysenteriae *Type 1 (WHO/CDR/95.4).* Pp. 1–47.
45. Yamai, S., Okitsu, T., Shimada, T., and Katsube, Y. (1997). Distribution of serogroups of *Vibrio cholera* non O1 non O139 with specific reference to their ability to produce cholera toxin and addition of novel serogroups. *J. Japanese Assoc. Infect. Dis.* **71**: 1037–1045.
46. Yamasaki, S. et al. (1997). Cryptic appearance of a new clone of *Vibrio cholerae* O1 biotype El Tor in Calcutta, India. *Microbiol. Immunol.* **41**: 1–6.
47. http://www.cdc.gov/.
48. http://www.who.int/.
49. http://www.cfsan.fda.gov/.
50. http://www.fsis.usda.gov/.
51. http://www.fas.org/.
52. http://www.doh.gov/.
53. http://www.pedianet.com/.
54. http://www.ncbi.nlm.nih.gov/ICTV/.

IMPROVING WATERBORNE DISEASE SURVEILLANCE

Floyd J. Frost, Ph.D.
The Lovelace Institutes
Albuquerque, New Mexico

Rebecca L. Calderon, Ph.D.
National Health and
Environmental Effects
Laboratory Research Triangle
Park, North Carolina

Gunther F. Craun, P.E.,
M.P.H., D.E.E.
Gunther F. Craun and
Associates
Staunton, Virginia

INTRODUCTION

Public health surveillance has played a key role in controlling the spread of communicable disease and identifying the need for specific public health practices, such as the filtration and chlorination of drinking water supplies. However, the characteristics of waterborne outbreaks since the early 1990s have raised questions about whether current water treatment practices can prevent transmission of some enteric pathogens (1–5). In addition, one analysis suggested that a significant fraction of all enteric disease in the United States may be due to drinking water (3). Another study found evidence that consuming surface-derived drinking water which meets current U.S. Environmental Protection Agency (USEPA) drinking water standards may significantly increase the risk of enteric illness (6). These concerns have motivated the U.S. Congress to require USEPA to prepare a report on the magnitude of epidemic and endemic waterborne disease in the United States.

Even as the needs increase for better information about waterborne disease occurrence and causes, some have suggested that our disease surveillance system is in a state of crisis and may possibly collapse (7). Another study revealed that state health departments often cannot dedicate any staff to enteric disease surveillance (8). Current concerns over the preparedness for detecting and controlling bioterrorism attacks have also motivated interest in the adequacy of waterborne disease surveillance. In this chapter, issues relating to disease surveillance and outbreak investigations are presented to assist readers in understanding the strengths and weaknesses of current waterborne disease surveillance and outbreak detection programs and to suggest additional steps to strengthen the system. With limited public health resources available, it is important to carefully consider the goals and approaches to waterborne disease surveillance. In addition to addressing the information needs of governmental disease control programs, it is essential to ensure that the information needs of the drinking water industry, the regulatory agencies, and the public are best served. It may also be essential for drinking water utilities to participate in and, perhaps, help fund these surveillance systems.

BACKGROUND

It is increasingly accepted that additional information is needed about the occurrence and causes of waterborne disease, both epidemic and endemic. The Centers for Disease Control (CDC) funded "emerging pathogen" surveillance projects in selected state health departments, in part to improve surveillance for several important waterborne agents. In New York City (NYC), the Department of the Environment (DEP), responsible for drinking water treatment and delivery, convened a panel of public health experts in 1994 to evaluate current health department disease surveillance programs. The panel recommended specific waterborne disease surveillance activities and epidemiologic studies to determine endemic waterborne disease risks associated with use of unfiltered surface water sources (Table 1) (9). Efforts to improve NYC waterborne disease surveillance are funded by the NYC DEP, the first time this has occurred for a drinking water utility in the United States.

Table 1. New York City Panel Recommendations on Waterborne Disease Surveillance

Designate an individual who is specifically responsible for coordinating waterborne disease surveillance
Conduct special surveillance studies of nursing and retirement home populations
Conduct surveillance in managed care populations
Monitor visits to emergency rooms
Conduct surveillance of high-risk populations
Monitor sales of prescription and nonprescription medications

An option for improving waterborne disease surveillance is to build on the current surveillance programs in place in most state and local health departments. This system is based on voluntary disease reporting by healthcare providers and clinical laboratories. However, a number of limitations of the system have been identified, and other factors may have already significantly reduced the effectiveness of traditional disease surveillance programs. Some pathogens, such as *Cryptosporidium*, are often difficult to diagnose, and other pathogens may exist for which there are no known diagnostic tests or no tests available for routine use. Changes in healthcare access and delivery practices may reduce the number of patients seeking healthcare and, also, the chances that medically attended diseases are confirmed by laboratory tests.

An outbreak resulting in many medically attended illnesses in a large city could be unrecognized, as almost happened in the Milwaukee outbreak. In that outbreak, a large increase in the occurrence of diarrheal illness occurred around March 30–31, 1993. On Thursday, April 1, 1993 a pharmacist noted a dramatic increase in sales of over-the-counter antidiarrheal and anticramping medications. Normally his pharmacy sold $30 a day of such medications. Starting that Thursday, drug sales increased to approximately $500–$600 a day, or 17–20 times the normal sales. The increased sales continued on Friday, as a result of which the pharmacy sold most of its supply of antidiarrheal medications. The pharmacist called the health department to inquire about excessive reports of diarrhea or intestinal illness. The health department was unaware of any outbreak. On Saturday the increased sales continued so the pharmacist contacted the three local television stations to report what he believed to be a major occurrence of diarrheal disease in the city. On Sunday night his report was carried on the evening news for one station and by Wednesday, April, 7, the outbreak was confirmed by the Milwaukee Health Department.

In the case of the Milwaukee outbreak, few of the people sought medical care for their diarrhea. However, even in situations where care was sought, it is possible that no one physician would notice an outbreak. For example, if many different healthcare providers treated the patients, it is possible that no one provider would recognize excess occurrences of illness. In addition, the existence of health effects in a small but extremely susceptible subpopulation might be difficult to detect because of the small number of people at risk.

As some changes have made it more difficult to detect outbreaks, other changes present new disease surveillance opportunities. Computerization of patient records, healthcare and laboratory workloads, prescription and nonprescription pharmaceutical sales, and calls to nurse hotlines are potential new tools for more effective and less costly disease surveillance. Technological advancements, such as detection of antigen or antibodies specific to a pathogen in sera, stools, and other secretions, may improve detection of etiological agents. These may also allow detection of infections in the absence of disease.

To better evaluate the current and alternative surveillance opportunities, five questions have been selected for discussion in this chapter:

1. What are the limitations of our current disease surveillance systems?
2. Should the early detection of outbreaks be the primary goal of a surveillance system and, if so, how can it be best achieved?
3. What is meant by endemic or background rates of disease, can some of this endemic disease be attributable to drinking water, and what should water utilities do to better understand these risks?
4. Can findings from outbreak investigations be used to estimate the unreported burden of enteric disease attributable to drinking water?
5. Since only a fraction of infected persons become ill from most enteric infections, should expanded surveillance programs monitor infection rather than illness?

LIMITATIONS OF THE CURRENT DISEASE SURVEILLANCE SYSTEMS

What are the limitations of our current disease surveillance systems? Detection of waterborne disease outbreaks depends, in part, on a state-federal system of notifiable or reportable diseases. Disease reporting is primarily the responsibility of healthcare providers and diagnostic laboratories. State or local laws require the reporting of certain diseases. Primary responsibility for disease surveillance rests with the state or local public health authorities. Most state surveillance systems are “passive,” in that reports are sent to the state or local health department by cooperative health care providers or laboratories. Providers and laboratories usually receive little encouragement from the health department to report illnesses. Government enforcement of reporting requirements is minimal. An “active” system will routinely contact some or all healthcare providers and laboratories, asking for illness reports (Table 2) (10).

It has long been recognized that both passive and active disease reporting incompletely ascertain the level of disease in the community. The level of completeness varies by disease, by state, and by areas or populations within a state (11). For example, reporting is likely to be more complete for severe diseases such as hemorrhagic *E. coli* than for milder infections, such as Norwalk virus gastroenteritis. Laboratories tend to be much better at reporting their findings than are physicians (10). Even within an area, there can be great variations in reporting, depending on the interest of clinical laboratories and the dedication of diagnosing physicians (11). For example, for pathogens that are new or where there are questions about the mode of transmission, reporting may be more complete than for agents that are common, where the mode of transmission is well known and where public health intervention is less necessary.

Table 2. Surveillance System Definitions

Mandatory reporting	A diagnosed case of disease is required, by law, to be reported; for example, in the case of cryptosporidiosis, all diagnosed cases are to be reported
Passive	Disease reports are submitted by providers and/or laboratories without specific follow-up by the health department
Active	Providers and/or laboratories are contacted to encourage diseases reporting; because of resource requirements, this is usually done as a special project for a limited duration of time
Enhanced	Special additional efforts are made to encourage disease reporting; this might include news releases, posters at strategic locations, presentation to special populations, or health surveys in communities with water quality problems

In addition to incomplete reporting of diagnosed illnesses, only a portion of all infections will ever be medically attended. As illustrated in (Fig. 1), only a fraction of infections will lead to illness. These infected persons may be unaware of their infection. In other cases, such as sometimes occurs as a result of childhood *Giardia* infection, the child fails to thrive but experiences none of the classic symptoms of giardiasis. When symptoms occur, they may be mild and/or may resolve in a short period of time. In this case, the person may not seek medical care or may simply visit a pharmacy to obtain medication to alleviate their symptoms. In the case of Milwaukee, despite the large number of reported cases of cryptosporidiosis, very few people visited their physician and few stool specimens were positive for *Cryptosporidium* oocysts.

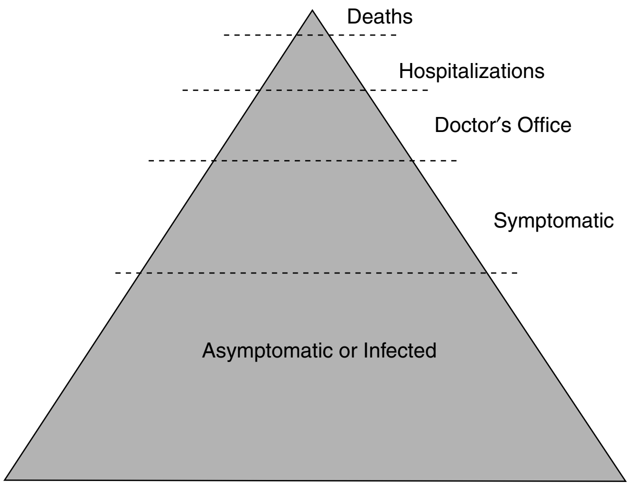

Figure 1. Disease pyramid.

If the person seeks healthcare, the physician may fail to correctly diagnose the infection, since in many cases symptoms are not sufficiently specific to accurately identify the pathogen. If misdiagnosed and the infection resolves itself, the patient may not seek additional healthcare and no report of an infection will be generated. Even when the physician correctly diagnoses the illness and prescribes the appropriate medication, a confirmatory laboratory test may not be ordered. If ordered, the patient may not submit the sample to the laboratory, since many patients are unwilling to submit stool specimens for laboratory analysis. Since laboratories are the primary source of disease reports for surveillance systems, without a laboratory-confirmed diagnosis, a report may never be filed. When a stool or blood sample is submitted for laboratory analysis, it can also test negative because of analytical or specimen collection error, untimely collection or because the material submitted was, by chance, free of the pathogen (12). Laboratory proficiency can vary considerably. This may be more of a problem for laboratories that run only a small number of the ordered test. For persons infected with enteric parasites, single stools may often be free of the parasite or have insufficient numbers of parasites to assure laboratory detection. In some cases, even multiple stools may be pathogen-negative.

If a sufficient number of cases of illness from the same pathogen are reported to the health department at about the same time and if the epidemiologist is alert to an increase in case reports, an outbreak may be identified. Because of the time required to perform the diagnostic tests and to report the results, outbreak recognitions may occur weeks after the onset of the actual outbreak.

Many outbreaks are first detected by an alert clinician. For example, in 1976, a Camas, Washington physician's son had returned from Russia with giardiasis. The physician later recognized that several of his patients had similar symptoms. This lead to the identification of a waterborne giardiasis outbreak (13). As mentioned earlier, in Milwaukee, Wisconsin a pharmacist noted a dramatic increase in sales of antidiarrheal medication. In California and Arizona, diarrheal illnesses reported to health agencies by 65 campers who had visited an Arizona park initiated an investigation that implicated contaminated water as the source of an outbreak that affected 1850 people (14). The fortuitous circumstances surrounding the detection of many outbreaks raises concerns about how many medium to large outbreaks are never detected. Small outbreaks may seldom be detected, especially among travelers who consume water from noncommunity systems or who swim in multiple locations.

Limitations of the current disease surveillance systems prompted a series of studies in the early 1980s to evaluate potential improvements in disease reporting and to evaluate the efficacy of active surveillance programs. A three state study of various approaches to active disease surveillance, funded by USEPA, detected no additional waterborne disease outbreaks in two states (Washington and Vermont) (15). However, in one state (Colorado) a greater than threefold increase in the number of detected waterborne outbreaks occurred (16). The reasons why Colorado was able to identify so many more outbreaks than either Washington State or Vermont are unclear. An intense effort was made to increase disease reporting in all states and dramatic increases in reports of enteric diseases were observed in all three states. It is possible that a combination of poor quality water supplies plus an exposed tourist population, without protective immunity, may have allowed Colorado to identify more outbreaks than the other two states.

In summary, active disease reporting can increase reporting of diagnosed illnesses only from providers and laboratories. All the other barriers to disease identification and reporting will still remain (Fig. 1). If healthcare access declines over time or, to reduce healthcare costs, physicians use fewer laboratory diagnostic services, then the number of diagnosed reportable illnesses will decline. This will occur despite the efforts of health departments to insure that most diagnosed illnesses are reported.

EARLY DETECTION OF OUTBREAKS

Should the early detection of outbreaks be the primary goal of a surveillance system, and, if so, how can it be best achieved? The occurrence of a waterborne disease outbreak is an exciting, newsworthy, and politically important event. Affected populations may experience severe illness and a large number of people may become ill. As a result of the investigation, much is often learned about the cause of major failures in water treatment or distribution. However, when the excitement has subsided, water system deficiencies have been corrected and the outbreak is officially said to be over, has the problem been solved or is disease continuing to occur but at a reduced level, below what is detectable by traditional surveillance activities?

For example, a waterborne disease outbreak investigation detected major problems with the filtration system of an anonymous small community water supply. The system was, at the time of its installation, considered adequate. However, high turbidity levels were observed in treated water at the time of the outbreak, suggesting poor operation of the filtration facility. Optimization of treatment by consulting engineers allowed the plant to dramatically improve pathogen removal. This improvement reduced the number of new cases of disease, and the outbreak officially ended. However, 2 years later a serological survey of the town's residents revealed the continued occurrence of infection by the same etiologic agent responsible for the earlier outbreak. These new data presented both philosophical and technical problems. Should all outbreaks be followed by such a survey? Is evidence of continuing infection sufficient reason for further intervention? If the serological survey were not conducted, there would be no evidence of increase risk of infection. If the plant was already optimized, what are the remaining intervention options without new filtration or disinfection technology?

This scenario assumes that the continued high serological levels resulted from waterborne transmission. In fact, without a follow-up epidemiologic investigation, it is not possible to distinguish waterborne from other routes of transmission. In addition, without improved surveillance activities, we know little about the absence of symptomatic disease. Low levels of disease from exposure to waterborne microbes over a period of many years can result in a much larger health burden for a community than the number of disease cases that might occur during a detected outbreak. However, exposure to some waterborne pathogens at levels that boost the immune response may prevent symptomatic illness. These concerns must all be considered when developing a surveillance system. Without clear goals and a commitment to conduct epidemiologic investigations and take appropriate actions, a better surveillance system will not improve public health.

Failure to detect low levels of disease transmission may provide a false sense of security. For example, why should an outbreak such as occurred in Milwaukee not have been preceded by many smaller outbreaks? Is it possible that in each of the cities experiencing a large waterborne cryptosporidiosis outbreak, prior undetected smaller outbreaks occurred? In fact, is it possible that lower levels of waterborne *Cryptosporidium* infection had occurred years prior to the outbreak? At the time of the detected outbreak, a higher number of oocysts may have passed through the treatment system or a more virulent strain of the pathogen emerged. If so, relying on disease surveillance systems that can only detect large outbreaks will seldom provide public health officials and the industry early warnings of emerging new diseases. This may be equivalent to basing the science of meteorology only on the study of hurricanes.

The detection of an outbreak can also affect future disease reports in an area. For example, it is possible that overreporting of symptoms consistent with the disease of interest could occur. If so, similar outbreaks may be detected in neighboring areas. Given the increased popularity of bottled water use, it is possible that the at-risk population could change following an outbreak if a significant fraction of the population discontinued drinking tapwater. Therefore, decreases in the occurrence of reported waterborne disease may not reflect better control of the contamination but a reduction in the number of exposed individuals.

ENDEMIC DISEASE

What is meant by endemic or background rates of disease and can some of this endemic disease be attributed to drinking water? Endemic level of disease is defined by the CDC as a persistent low to moderate level of disease occurrence. A persistently high level of occurrence is called *hyperendemic* while an irregular pattern of occurrence is called *sporadic* (Fig. 2). For most enteric infections, endemic disease results from a statistical averaging of small to moderate-sized undetected outbreaks or clusters of infection. There is little information to suggest that endemic levels of disease remain constant over time or across geographic areas, nor is there reason to believe the endemic level of disease is unimportant.

Over the past century, the importance of endemic disease has become increasingly recognized. Following World War I, an attempt was made to estimate the prevalence of parasite infections in both the returning British soldiers and the British population who remained at home (17). To the surprise of the researchers, a high prevalence of asymptomatic infection was found among persons who had never left Britain. Later, a survey of Wise County, Virginia in 1930 revealed that half of the population carried *Entamoeba histolytica* and that 38% carried *Giardia lamblia* (18). A study to determine the incidence of *Cryptosporidium* infection among Peace Corp workers to be sent overseas revealed that almost 30% had possibly experienced infection prior to leaving the United States (19). More recent work we conducted suggests that endemic rates of *Cryptosporidium* infection may be very high, but that rates of cryptosporidiosis may be low (20,21).

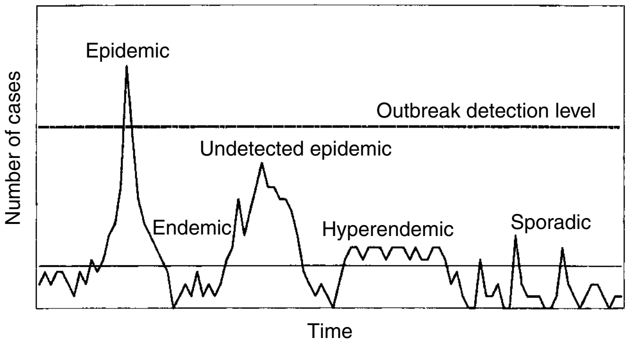

Figure 2. Epidemic versus endemic disease.

Data derived from disease surveillance systems cannot be used to compare endemic disease levels between areas or populations with different water systems. Whether observed differences in disease reports are due to the differences in the completeness of reporting or to differences in the occurrence of the disease or the infection cannot be answered, even with improved surveillance systems. In addition, it has become increasingly recognized that populations can develop protective immunity to infectious agents. If so, rates of infection may remain high while rates of illness remain low (21). The absence of disease in a population may, therefore, not mean that there is an absence of infections. Epidemiologic studies must be specifically designed and conducted to address the association of endemic disease with water system type or quality.

Several epidemiologic studies have reported waterborne disease associated with public water systems in the absence of a reported waterborne outbreak. In New Zealand, the incidence of laboratory-confirmed giardiasis was found to be higher in a part of the city receiving chlorinated, unfiltered surface water compared to the part where surface water was treated by coagulation, flocculation, granular filtration, and chlorination (22). In Vermont, a higher incidence of endemic giardiasis was found in municipalities using unfiltered surface water or wells than in municipalities with filtered surface water (23). A Canadian study attempted to estimate how much endemic enteric illness was due to drinking water (6). The fraction of illness attributable to drinking water was estimated by comparing rates of reports of "highly credible gastrointestinal illnesses" among persons drinking tapwater with rates among people drinking water from reverse osmosis filtration units. Although different rates of illness could have resulted from reporting biases, if the findings are confirmed by future studies, then drinking water may significantly contribute endemic disease in at least one community. Unfortunately, a study using a similar design conducted in Melbourne, Australia, did not provide evidence of endemic waterborne disease (24).

A variety of approaches have been proposed for estimating the burden of endemic diarrheal disease from drinking water sources. In addition to the Australian replication of the Payment design, a small pilot household intervention study in California has recently been completed (25). That study concluded that it was possible to blind families as to the type of treatment device they had, and although the study was not powered to examine illness rates, the families with true home treatment devices reported a lower rate of illness. A larger randomized household intervention study is under way in the United States. The advantage of the randomized household interventions is that the design precludes reporting biases and assignment biases, assuming that people do not know whether they are in the intervention or the control group. A major disadvantage of this approach is that only household drinking water quality is altered. Drinking water from other sources, such as work or at restaurants, is not altered. Another limitation is that long-term healthy residents are usually recruited and these people may have the lowest risk of suffering illness from waterborne infections. Therefore, negative results are difficult to interpret. Household intervention studies are limited in generalizability because they are conducted in single communities, although the study design would be amenable to national randomized trial.

Another proposed approach is to relate variations in the occurrence of health events, such as emergency room visits and hospitalization, with variation in drinking water turbidity levels (26,27). This approach has some merit; however, the results are difficult to interpret since no causal agents are identified. There are also concerns that the optimized statistical modeling cannot be statistically evaluated. Therefore, many of the claimed associations may be spurious.

Another approach uses planned changes in drinking water treatment and then evaluates the occurrence of potentially waterborne disease before and after intervention. The advantage of this approach is that most or all drinking water from an area is changed. This avoids one of the problems with household interventions. One disadvantage of this approach is that the sites receiving new water treatment technologies are not randomly assigned. For example, most unfiltered drinking water systems in the United States use high-quality source water. Adding filtration may not dramatically change the health risks from the drinking water. Another is that the community intervention looks at only one city or one pair of cities, so the sample size is restricted.

APPLICABILITY OF OUTBREAK INVESTIGATIONS

Can findings from outbreak investigations be used to estimate the burden of enteric disease attributable to drinking water? Epidemic disease is defined as an unusual occurrence or clustering of a specific illness. Between 1971 and 1994 there were 737 documented waterborne disease outbreaks (28–30). Almost half of these were due to unknown etiological agents that caused acute gastrointestinal illness. Among these outbreaks, the relative importance of different etiologic agents (viruses, bacteria, protozoa, and chemicals) can be estimated. For example, the etiologic agents most commonly associated with waterborne disease in the United States include, in descending order, undefined gastroenteritis, giardiasis, shigellosis, viral gastroenteritis, and hepatitis A. This ranking is based on outbreaks and may or may not reflect the relative importance of these etiologic agents for all waterborne disease.

For diseases where outbreaks account for the majority of illnesses, the outbreak is of primary interest. However, for many waterborne pathogens, outbreaks account for only a small fraction of all illnesses. For example, in a 1.5-year period during the late 1970s in Washington State, 1347 laboratory confirmed cases of giardiasis were reported to the state health department (31). Extensive follow-up of these cases (Table 3) revealed that clusters or possible small outbreaks accounted for only 16% of all cases of giardiasis reported during this time period. These data suggest that "endemic giardiasis" was overwhelmingly more abundant than "epidemic giardiasis" in Washington State during this time period.

There are a number of problems with extrapolating the characteristics of cases involved in outbreaks to revise all cases of illness, including the following:

1. If there is variation in the virulence of a pathogen, then detected outbreaks may predominantly be caused by the more virulent strains of the pathogen. This may overestimate the severe morbidity or mortality associated with the pathogen.
2. By examining only detected outbreaks, one may overestimate the importance of drinking water as a

Table 3. Case Clusters of Giardiasis in Washington State 1977–1978

Number of Cases	Etiology
10	Untreated streamwater consumption
14	Untreated water consumption at a work camp
11	One small community water system
12	Tourists returning from a resort in Mexico
17	One-daycare center outbreak
8	One-daycare center outbreak
24	Among 10 different daycare centers
73	Multiple cases among 21 families
51	Nonfamily association with another case
220	Total in all clusters

route of transmission. Because of the large number of cases often involved, waterborne outbreaks may be more detectable than outbreaks from other routes of transmission. Even a severe day care outbreak would involve only a few cases. Within family clusters usually involve too few cases to be a detectable outbreak.

3. Outbreak detection is often more difficult for common or endemic diseases than for uncommon diseases. For example, two cases of cholera anywhere in the United States might be considered an outbreak whereas 50 cases of cryptosporidiosis widely dispersed in a large U.S. city during a week might easily be absorbed as expected background cases of diarrhea and not recognized as an outbreak (9).

Outbreaks of short duration of illnesses (e.g., some viruses) are more difficult to detect and study than are outbreaks of long duration illnesses (e.g., giardiasis, shigellosis, hepatitis A). Therefore the importance of acute, self-limited gastrointestinal illness of undetermined etiology and short duration may be underestimated relative to outbreaks of parasitic infections and some bacterial or viral pathogens with a longer duration of symptoms.

Pathogens with long incubation periods are difficult to investigate since the conditions that allowed transmission of the pathogen may have changed between the time of infection and the time when the outbreak was detected. Underascertaining waterborne sources for disease outbreaks caused by these agents is likely.

MONITORING INFECTION VERSUS DISEASE

Only a fraction of infected persons become ill from the most commonly occurring enteric infections. Of the people that become ill, only a fraction of cases will be reposted (Fig. 3). Should expanded surveillance programs attempt to monitor infection rather than disease? The existence of asymptomatic carriers of infections has been known for some time (e.g., Typhoid Mary). However, the number of asymptomatic carriers for many infections has only relatively recently been appreciated. The parasite prevalence surveys in Britain (17) and in Virginia (18) found more asymptomatic infected persons than expected. Even as late as 1952, in New Hope, Tennessee, 10.6% of the general population was infected with *Giardia lamblia* (32). Following a 1966 giardiasis outbreak in Aspen, Colorado, a stool survey found that 5% of the population was infected with *Giardia* (33). A survey of Boulder, Colorado, also conducted following an outbreak, found a prevalence of 5% (34). Most of the individuals participating in these surveys were asymptomatic. A stool survey of one to 3-year-old Washington State children was conducted in 1980 (35). This survey found that 7% of the children were infected with *Giardia lamblia*. All participating children were reported as healthy at the time of the survey. The Seattle Virus Watch program, conducted during the 1960s and early 1970s monitored virus infections among a sample of people in selected U.S. cities. This study found that illness was reported in less than half of all enterovirus infections (36).

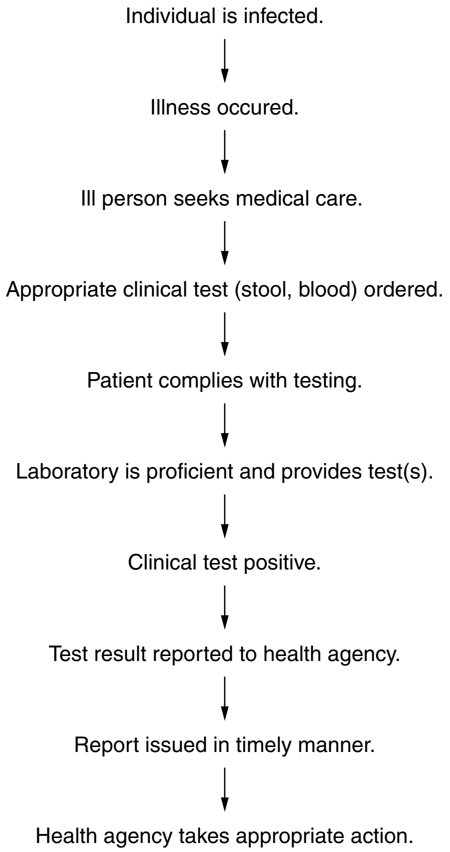

Figure 3. Events in reporting an individual infection.

New serological tools have been developed since the early 1980s to better monitor the prevalence of prior infections among the population. Even though infection may not result in moderate or severe illness, there are several reasons for considering infection rather than disease, including the following:

1. Information on infections can provide a much expanded understanding of the relative importance of various routes of transmission and provide an early warning for risks of outbreaks.
2. Serological epidemiologic studies of infection can better estimate the extent of endemic waterborne disease. These studies are statistically more powerful to detect low risks in moderate-size populations.
3. Just as the occurrence of a coliform test indicates the potential of disease risk for a drinking water source, the waterborne transmission of pathogens, even when infection is predominantly asymptomatic, can provide critical information for evaluating water treatment systems and may help identify correctable problems in water source protection and/or treatment.
4. Widespread, unrecognized transmission of infection in the general population may indicate a devastating outbreak for a susceptible subpopulation.

Another advantage of serological surveillance occurs during an outbreak. An outbreak of cryptosporidiosis was detected in Las Vegas, Nevada in Spring 1994. Although this was clearly a cryptosporidiosis outbreak, the inability to detect problems with the water treatment system and publicity prior to the investigation that suggested the outbreak was waterborne raised questions over whether the outbreak could be classified as waterborne (37–39). Since the majority of the diagnosed cases also suffered from HIV or AIDS, the extent of the outbreak was unclear. Had asymptomatic infected persons been identified serologically, the effects of reporting bias would be reduced since asymptomatic cases would have no motivation to explain an asymptomatic infection.

IMPROVING DISEASE SURVEILLANCE

Several options are available for enhanced waterborne disease surveillance. The option or combination of options selected will depend on the specific goals for disease surveillance. The currently used national system of surveillance, based on diagnosed illness, has a long-established record of both performance and nonperformance for detecting outbreaks (Table 4). Because the current system is both inexpensive to maintain and currently operational, it has considerable appeal among public health practitioners. However, monitoring pharmaceutical sales, nurse hotline calls, or physician visits is a potential enhancement to the traditional disease surveillance programs (Table 5) (39,40).

The goal of our current disease surveillance system is outbreak detection. Unfortunately, there is little rigorous evaluation of its capability to detect outbreaks. Furthermore, the common occurrence of fortuitous situations that lead to the outbreak detection raise questions about the sensitivity of the system. To improve the sensitivity to detect small to medium-size outbreaks or to provide early information on the occurrence of an outbreak, these alternative approaches mentioned have promise. Over-the-counter pharmaceutical sales may be useful, but it has some significant limitations (40). The use of nurse hotline calls to continuously monitor the occurrence of infectious disease has tremendous promise, but no efforts have been made to use this surveillance tool (39). Better linkages with infectious disease specialists in healthcare organizations may also improve disease surveillance.

Table 4. Advantages and Disadvantages of the Current Waterborne Disease Surveillance System

Advantages

In-place and operational across the nation

Extensive health department experience using the system

Inexpensive to maintain

An operational nationwide network, operated by the Centers for Disease Control (CDC), for summarizing and reporting findings

Methodological development of algorithms for detecting excess occurrences of disease

Disadvantages

Inability to detect outbreaks when diagnosed cases are not reported to the health department

Delays in detecting outbreaks due to the time required for laboratory testing and for reporting of findings

Undetected outbreaks where health problems are not medically treated or where infection results in only mild or no illness

Limited opportunities for system improvement

Possible long-term trend in healthcare delivery that may reduce its efficacy

Table 5. Advantages and Disadvantages of New Waterborne Surveillance Systems

Advantages

They may detect outbreaks where few patients seek healthcare or where the illness is of sufficiently short duration that healthcare is unimportant

They are relatively fast in reporting outbreaks since the time delay between the onset of symptoms and the purchase of drugs or calls to nurses is likely to be short

They are relatively inexpensive to maintain, especially if nationwide retail pharmacies are involved or common nurse hotline software is programmed for reporting

Disadvantages

Since only symptoms are ascertained, they will not usually identify an etiologic agent

Although inexpensive to maintain, initial computer programming and establishing data sharing agreements would require some investment

The specificity of the system for outbreak detection (e.g., number of false leads) is untested

None of the traditional or enhanced surveillance tools will provide much useful information on low-level or endemic risk of enteric pathogen infection. However, new serological tests have increased the feasibility of studies to estimate the incidence of new infections or the prevalence of antibody response to pathogens and to relate this information with modes of transmission. In the early 1970s, the Seattle Virus Watch program examined occurrences of viral infections among volunteers in selected communities (33). Similar approaches to monitoring the occurrence of *Giardia* (41) and *Cryptosporidium* (42) infections have been developed since then. More work is needed to evaluate these new tools as well as to develop other tests. We also need to design cost-effective approaches to their widespread implementation. These tools may give us an opportunity to greatly improve our understanding of the importance of various modes of transmission and identify reasons why one population group has a higher endemic level of disease than another. It is likely that as more is known about the modes of transmission, a better understanding will emerge of both drinking water and nondrinking water routes of pathogen transmission.

Healthcare reforms may reduce the use of diagnostic laboratory services, reducing the value of laboratory-based disease surveillance. However, new opportunities for improved disease surveillance, including both individual and community disease reporting and surveillance of endemic infections, may also result. To fully exploit these opportunities, a new public health partnership with distributed responsibilities may be needed between healthcare providers, health maintenance organizations (HMOs), pharmacies, and the traditional public health agencies.

The increasing age of our population has resulted in increases in the number of immunosuppressed persons. Some of this immunosuppression may result from chronic diseases, while some may result from medically induced immunosuppression following treatment for other conditions. For example, many cancer patients have temporary periods of immunosuppression following treatment. These populations may be at especially high risk of adverse consequences of infection. Since diarrheal disease in this population is also relatively common, many infections may not be detected.

Infectious disease surveillance systems are operated by state and local public health agencies with little or no direct contact with healthcare providers. To improve disease surveillance system, it will likely be necessary to better integrate healthcare delivery systems with those disease surveillance programs. This integration can only occur if both the state public health agencies and the healthcare providers recognize benefits from this cooperation and barriers to data sharing are reduced.

DISCLAIMER

The views expressed in this chapter are those of the individual authors and do not necessarily reflect the views and policies of the USEPA. The chapter has been subject to the Agency's peer and administrative review and approved for publication.

BIBLIOGRAPHY

1. D'Antonio, R.G., Winn, R.E., Taylor, J.P., Gustafson, T.L., G.W. Gray, Current, W.L., Zajac, R.A., and Rhodes, M.M. (1985). A waterborne outbreak of cryptosporidiosis in normal hosts. *Annals Int. Med.* **103**: 886–888.
2. Bennett J.V., Holmberg, S.D., and Rogers, M.F. (1987). Infectious and parasitic diseases. *Closing the Gap: The Burden of Unnecessary Illness*. R.W. Amler and H.B. Dull (Eds.). Oxford University Press, New York.
3. Hayes, E.P. et al. (1989). Large community outbreak of cryptosporidiosis due to contamination of a filtered public water supply. *New Engl. J. Med.* **320**: 372–376.
4. Leland, D., McAnulty, J., Keene, W., and Terens, G. (1993). A cryptosporidiosis outbreak in a filtered water supply. *J. Am. Water Works Assoc.* **85**: 34–42.
5. MacKenzie, W.R. et al. (1994). A massive outbreak in Milwaukee of *Cryptosporidium* infection transmitted through the public drinking water supply. *New Engl. J. Med.* **331**(3): 161–167.
6. Payment, P., Richardson, L., Siemiatycki, J., et al. (1991). A randomized trial to evaluate the risk of gastrointestinal disease due to consumption of drinking water meeting current microbiologic standards. *Am. J. Public Health* **81**(6): 703–708.
7. Berkelman, R.L., Bryan, R.T., Osterholm, M.T., LeDuc, J.W., and Hughes, J.M.. (1994). Infectious disease surveillance: A crumbling foundation. *Science* **264**: 368–370.
8. Frost, F.J., Calderon, R.L., R. L., and Craun, G.L. (1995). Waterborne disease surveillance: Findings of a survey of state and territorial epidemiology programs. *J. Environ. Health.* **58**(5): 6–11.
9. Craun, C.L., Birkhead, G., Erlandsen, S., et al. (1994). *Report of New York City's Advisory Panel on Waterborne Disease Assessment*. The New York City Department of Environmental Protection, New York.
10. Foster, L.R. (1990). Surveillance for waterborne illness and disease reporting: State and local responsibilities. In: *Methods for Investigation and Prevention of Waterborne Disease Outbreaks*. G.F. Craun (Ed.). EPA/600/1-90/005a. USEPA Office of Research and Development, Cincinnati.
11. Chorba, T.L., Berkelman, R.L., Safford, S.K., Gibbs, N.P., and Hull, P.E. (1989). Mandatory reporting of infectious diseases by clinicians. *J. Am. Med. Assoc.* **262**: 3018–3026.
12. Chappell, C.L., Okhuysen, P.C., Sterling, C.R., and DuPont, H.L. (1996). *Cryptosporidium parvum*: Intensity of infection and oocyst excretion patterns in healthy volunteers. *J. Infect. Diseases* **173**: 232–236.
13. Kirner, J.C., Littler, J.D., and Angelo, L.A. (1978). A waterborne outbreak of giardiasis in Camas. *J. Am. Water Works Assoc.* **70**: 35–40.
14. Starko, K.M., Lippy, E.C., Dominguez, L.B., Haley, C.E. and Fisher, H.J. (1986). Campers' diarrhea traced to water-sewage link. *Public Health Reports* **101**: 527–531.
15. Harter, L., Frost, F., Vogt, R., Little, A., Hopkins, R., Gaspard, B., and Lippy, E. (1985). A three-state study of waterborne disease surveillance techniques. *Am. J. Public Health* **75**: 1327–1328.
16. Hopkins, R.S., Shillam, P., Gaspard, B., Eisnach, L., and Karlin, R.S. (1985). Waterborne disease in Colorado: three years surveillance and 18 waterborne outbreaks. *Am. J. Public Health* **75**: 254–257.
17. Smith, A.M. and Matthews, J.R. (1917). The intestinal protozoa of non-dysenteric cases. *Annals Trop. Med. Parasitol.* **10**: 361–390.
18. Faust, E.C. (1930). A study of the intestinal protozoa of a representative sampling of the population of Wise County, southwestern Virginia. *Am. J. Hygiene* **11**: 371–384.
19. Ungar, B.L., Milligan, M., and Nutman, T.B. (1989). Serologic evidence of *Cryptosporidium* infection in U.S. volunteers before and during Peace Corps service in Africa. *Arch. Int. Med.* **149**: 894–897.
20. Frost, F. (1998). Two-city *Cryptosporidium* study. *Am. Water Works Assoc. Research Found.—Drink. Water Research* **8**(6): 2–5.
21. Frost, F.J., Muller, T., Craun, G.F., Calderon, R.L., and P.A. Roeffer. (2001). Paired city *Cryptosporidium* serosurvey in the southwest USA. *Epidemiol. Infect.* **126**: 301–307.
22. Frasher, G.G. and Cooke, K.R. (1989). Endemic giardiasis and municipal water supply. *Am. J. Public Health* **79**: 39–41.
23. Birkhead, G. and Vogt, R.L. (1989). Epidemiologic surveillance for endemic *Giardia lamblia* infection in Vermont. *Am. J. Epidemiol.* **129**: 762–768.
24. Hellard, M.E., Sinclair, M.I., Forbes, A.B., and Fairley, C.K. (2001). A randomized blinded controlled trial investigating the gastrointestinal health effects of drinking water quality. *Environ. Health Perspect.* **109**: 773–778.

25. Colford, J.M., Rees, J.R., Wade, T.J., Khalakdina, A., Hilton, J.F., Ergas, I.J., Burns, S., Benker, A., Ma, C., Bowen, C., Mills, D.C., Vugia, D.J., Juranek, D.D., and Levy, D.A. (2001). Participant blinding and gastrointestinal illness in a randomized, controlled trial of an in-home drinking water intervention. *Emerging Infect. Diseases* **8**(1): 29–36.
26. Schwarz, J., Levin, R., and Hodge, K. (1997). Drinking water turbidity and pediatric hospital use for gastrointestinal illness in Philadelphia. *Epidemiology* **8**: 615–620.
27. Morris, R.D.F., Naumova, E.N., and Griffiths, J.K. (1998). Did Milwaukee experience waterborne cryptosporidiosis before the large documented outbreak in 1993? *Epidemiology* **9**: 264–270.
28. Craun, G.F. (1992). Waterborne disease outbreaks in the United States of America: Causes and prevention. *World Health Statistics Quart.* **45**: 192–196.
29. Moore, A.C., Herwaldt, B.L., Craun, G.F., Calderon, R.L., Highsmith, A.K., and Juranek, D.D. (1994). Waterborne disease in the United States, 1991 and 1992. *J. Am. Water Works Assoc.* **84**(2): 87–99.
30. Kramer, M.H., Herwaldt, B.L., Craun, G.F., Calderon, R.L., and Juranek, D.D. (1996). Surveillance for waterborne disease outbreaks—United States, 1993–1994. *J. Am. Water Works Assoc.* **88**: 66–80.
31. Frost, F., Harter, L., Plan, B., Fukutaki, K., and Holman, B. (1983). *Giardiasis in Washington State*. USEPA Report 83-134-882. National Technical Information Service, Springfield, VA.
32. Eyles, D.E., Jones, F.E., and Smith, S.C. (1953). A study of *Entamoeba histolytica* and other intestinal parasites in a rural west Tennessee community. *Am. J. Trop. Med.* **2**: 173–190.
33. Gleason, N.N., Horwitz, M.S., Newton, L.H., and Moore, G.T. (1970). A stool survey for enteric organisms in Aspen, Colorado. *Am. J. Trop. Med. Hygiene* **19**: 480–484.
34. Wright, R.A., Spender, H.C., Brodsky, R.E., and Vernon, T.M. (1977). Giardiasis in Colorado: An epidemiologic study. *Am. J. Epidemiol.* **105**: 330–336.
35. Harter, L., Frost, F., and Jakubowski, W. (1982). *Giardia* Prevalence among 1-to-3 year-old children in two Washington State counties. *Am. J. Public Health* **72**: 386–388.
36. Elveback, L.R., Fox, J.P., Ketler, A., Brandt, C.D., F.E. Wassermann, and Hall, C.E. (1966). The Virus Watch program; a continuing surveillance of viral infections in metropolitan New York families. 3. Preliminary report on association of infections with disease. *Am. J. Epidemiol.* **83**: 436–454.
37. Craun, G.F. and Frost, F.J. (2002). Possible information bias in a waterborne outbreak investigation. *Int. J. Environ. Health Research* **12**: 5–15.
38. Craun, G.F., Frost, F.J., Calderon, R.L., Hilborne, H., Fox, K.R., Reasoner, D.J., Poole, C., Rexing, D.J., Hubbs, S.A., and A.P. Dufour. (2001). Improving waterborne disease outbreak investigations. *Int. J. Environ. Health Research* **11**: 229–243.
39. Rodman, J.S., Frost, F.J., and Jakubowski, W. (1998). Using nurse hot line calls for disease surveillance. *Emerg. Infect. Diseases* **4**: 1–4.
40. Rodman, J.S., Frost, F., Burchat, I.D., Fraser, D., Langer, J., and Jakubowski, W. (1997). Pharmacy sales—a method of disease surveillance. *J. Environ. Health* **60**(4): 8–14.
41. Nulsen, M.F., Tilley, P.G., Lewis, L., Zhang, H.Z., and Isaac-Renton, J.L. (1994). The humeral and cellular host immune responses in an outbreak of giardiasis. *Immunol. Infect. Diseases* **4**: 100–105.
42. Moss, D.M. and Lammie, P.J. (1993). *J. Am. Soc. Trop. Med. Hygiene* **49**: 393.

DISINFECTANTS

Raghuraman Venkatapathy
Oak Ridge Institute for Science
and Education
Cincinnati, Ohio

June Weintraub
City and County of
San Francisco
Department of Public Health
San Francisco, California

J. Michael Wright
Harvard School of Public Health
Boston, Massachusetts

Disinfection of drinking water was instituted at a time when waterborne infectious diseases were the primary focus in public health. The use of disinfectants such as chlorine, chloramine, chlorine dioxide, ozone, and ultraviolet radiation has resulted in a reduction in the outbreak of diseases such as cholera and polio in many parts of the world. The choice of disinfectant is dependent on their availability, ease of use, cost, and their efficacy in disinfecting water. Some disinfectants commonly used in water treatment plants are not stable in water for long periods of time, which reduces their disinfection efficacy. Current practices to control pathogens in drinking water include using a combination of disinfectants, coagulation, and filtration. A main disadvantage of using disinfectants in drinking water is their ability to react with organic and inorganic matter in water to form disinfection byproducts that may be of health concern. The disinfectants can also be hazardous at high concentrations, but toxicological studies suggest that their toxicity is likely not relevant at the low concentrations encountered in drinking water.

PURPOSE

Contaminated, untreated, or inadequately treated water is known to transmit disease-causing pathogens such as bacteria, viruses, fungi, and protozoa (1). The size of these pathogens varies from a few hundred microns for bacterial clusters to around 0.01 μm for viruses (2). Although physical water treatment processes such as sedimentation and filtration may remove pathogens that are greater than 10 μm in size, ultra- or nanofiltration may be required to remove disease-causing pathogens (3). Conventional filtration is usually not practical for smaller systems, and, ultrafiltration may not be economically viable to treat large quantities of water; therefore, chemical treatment may have to be employed to treat water. The process of reducing the number of pathogenic microorganisms through the addition of chemicals (i.e., disinfectants) into water is called disinfection.

A variety of disinfectants may be used in water treatment plants, with the choice dependent on the location and size of the plant, the cost effectiveness, the pathogen(s) of concern, the water source, the characteristics of the pretreated water, and the disinfectant characteristics (e.g., solubility, stability, disinfecting and deodorizing ability, corrosiveness, and the time required to disinfect). Other factors that are considered include the effectiveness of the disinfectant in killing a range of microorganisms, the potential to form disinfection byproducts (DBPs) that may cause health effects in humans and animals, and the ability of the disinfecting agents to remain effective in water throughout the water distribution system.

HISTORY

Prior to A.D. 1600, civilizations consumed water based on visual clarity, or after treatments such as exposing the water to sunlight, dipping heated copper or other metals in water, boiling, and filtering water through a cloth (4). Other types of water treatment such as exposure to germicidal metals (e.g., silver and copper), sand filtration, distillation, coagulation, and adsorption with alum, lime, plant extracts, charcoal, clay, or plant materials have been employed since the 1600s. Modern methods of water disinfection were first used in Europe in the mid-to-late 1800s. One of the first known uses of chlorine as a germicide was Ignac Semmelweis' introduction of chlorine water for hand cleansing in the Vienna General Hospital maternity wards in 1846 (5).

Following research advances on the effect of microorganisms on human health in the 1870s, and improvements to physical water treatment technologies such as slow sand filtration in the 1880s, chlorine (as chloride of lime) and ozone were first used as drinking water disinfectants on a plant-scale basis at water treatment facilities in Hamburg, Germany and Oudshoorn, Holland, respectively, in 1893 (5,6). In 1897, Sims Woodhead used bleach to sterilize potable water at Maidstone, England after a typhoid outbreak. In 1903, a water treatment facility in Middlekerke, Belgium was the first to use chlorine gas as a disinfectant (4,7).

In North America, the first use of chemical disinfection at water treatment plants began in 1908 when chlorine was used as a disinfectant in Jersey City, New Jersey (as sodium hypochlorite) and Chicago, Illinois (as chloride of lime). Chloramination was first used in Ottawa, Canada and Denver, Colorado in 1917, whereas ozonation was first used in the United States in the 1940s (4). Today, chemical disinfection of drinking water is widely recognized as a necessity and is widely practiced at city, community, and point-of-use levels throughout the United States.

TYPES OF DISINFECTANTS

Chemical disinfection is considered the most effective treatment to inactivate pathogens in drinking water. A majority of the disinfectants that are in use in water treatment plants today are oxidants such as chloramine, chlorine dioxide, chlorine gas, electrochemically generated oxidant from sodium chloride (NaCl), hypochlorite, ozone, and ultraviolet (UV) light. Additional chemical disinfectants used to treat drinking water, especially in households, include acids and bases such as citric juices, lime, mineral acids and hydroxide salts, metals such as copper or silver, surfactants, and permanganate (8). Since each disinfectant has its own advantages and disadvantages (summarized in CHLORINATION BYPRODUCTS and the ALTERNATIVE DISINFECTANTS), a combination of multiple treatment processes including sedimentation, coagulation, flocculation, filtration, and disinfection is used in the majority of the water treatment plants in the world today. Table 1 provides a summary of the water treatment methods used, their availability, ease of use, cost, and efficacy in neutralizing the disease-causing pathogens.

DISINFECTION EFFICACY

Disinfection kinetics is usually expressed by Chick's law (also known as the CT law), which relates the activity [e.g., 3-log (99.9%) or 5-log (99.999%) reduction] to the product of disinfectant concentration (C) and contact time (T) (12). For example, to provide a given degree of disinfection, a low concentration and high contact time may be maintained or vice versa (13). The CT law may not provide adequate disinfection due to various factors that may affect the efficacy of the disinfectant, such as the type of pathogen; type of disinfectant; chemical factors such as pH, dissolved organic matter, salts, and ions; and particulate matter (14). Among the pathogens, vegetative bacteria are generally the easiest to disinfect followed by viruses, bacterial spores, fungal spores, and protozoan parasites (8). The formation of aggregates with other microorganisms or particulates in water may reduce the disinfectant's efficacy by preventing access to the pathogens. Among the most commonly used disinfectants, ozone's ability to disinfect is generally the highest followed by chlorine dioxide, electrochemically generated chlorine, and chloramine. However, ozone is the least stable disinfectant in water and is hence unable to provide a stable disinfectant residual in treated water that is necessary to prevent regrowth of the disease-causing pathogens (15). The chlorine-based disinfectants on the other hand are relatively stable in water in the absence of pathogens. At neutral pH, the half-life of chlorine dioxide is 30 min and 14 h for a 0.01 M and 0.0001 M solution, respectively (15). At ambient temperatures, the half-life of chloramines is approximately 100 h (16), whereas the half-life of sodium hypochlorite varies between 60 and 1700 d for water containing 18% and 3% available chlorine, respectively (17).

The characteristics of the source water also influence the efficacy of the disinfectants. Certain disinfectants such as hypochlorite are more effective at low pH, whereas chloramines are effective at a higher pH (5,15). The disinfectants may be consumed through reactions with other dissolved constituents in the water such as dissolved

Table 1. Advantages of Different Water Treatment Methods and Their Efficacy

Method	Availability	Difficulty	Cost	Pathogen Efficacy	Comments
Acids and bases	☺☺☺☺	×	$$	⊖⊖	Mostly used for pH control in water treatment plants; some microbial deactivation reported.
Adsorption	☺☺☺	×××	$$$$	Variable	Some bioreactors remove pathogens and organics.
Aeration	☺☺☺	×	$	⊖	Has to be used with other methods to be practical.
Boiling	☺☺☺☺	×	$$	⊖⊖⊖⊖	Used mostly at the point-of-use level.
Chloramines	☺☺☺☺	××	$$$	⊖⊖⊖	Less effective than free chlorine when used as a primary disinfectant; more practical as a secondary disinfectant.
Chlorine gas	☺☺☺☺	×××	$$$	⊖⊖⊖⊖	Widely used disinfectant.
Chlorine generated electrochemically from NaCl	☺☺☺☺	××	$$$	⊖⊖⊖⊖	Can be generated on site.
Chlorine dioxide	☺☺	×××	$$$$	⊖⊖⊖⊖	Highly effective as a primary disinfectant but is a poor secondary disinfectant.
Coagulation	☺☺☺	×××	$$$	⊖⊖	Useful in settling large flocculants of bacteria.
Combination of disinfection, coagulation, and filtration	☺	××××	$$$$	⊖⊖⊖⊖⊖	Best method, but is accompanied by high costs.
Filtration	☺☺☺	×××	$$$$	⊖⊖⊖	Ultrafiltration needed to remove viruses.
Hypochlorite	☺☺☺☺	××	$$$	⊖⊖⊖⊖	Most widely used disinfectant; highly unstable chemical.
Iodine	☺	××	$$$$	⊖⊖⊖⊖	Mostly used in tablet form at the point-of-use.
Ion exchange	☺	××××	$$$$	⊖	Can remove salts but not pathogens.
Ozone	☺☺	××××	$$$$	⊖⊖⊖⊖	Most effective as a primary disinfectant; cannot be used as a secondary disinfectant.
Sedimentation	☺☺☺☺	××	$$	⊖	Useful for removing adhering bacteria.
Silver or copper	☺☺☺☺	×	$$	⊖	Heated metal dipped into water at the point-of-use level in lieu of boiling.
Sunlight	☺☺☺☺	×	$	⊖⊖	Large open spaces and constant sunlight required; impractical for everyday use.
UV lamps	☺	×××	$$$	⊖⊖⊖	Effective as a primary disinfectant against pathogens but not as secondary disinfectant; space requirements per volume of water disinfected are impractical for larger systems.

Source: References 4, 5, 8–11.

organic matter and other salts to form DBPs, thereby limiting the amount of disinfectant available for pathogen inactivation.

CHEMISTRY OF DISINFECTANTS

The physical and chemical properties of disinfectants can affect their behavior in drinking water as well as their toxicity. Apart from reacting with natural organic matter or other solutes to form DBPs, disinfectants undergo a number of reactions to form products that may be toxic to humans and animals.

Free Chlorine

Chlorine gas (Cl_2) when exposed to water forms hypochlorous acid (HOCl) and hydrochloric acid ($HCl \rightleftharpoons H^+$ and Cl^-).

$$Cl_2 + H_2O \rightleftharpoons HOCl + H^+ + Cl^- \tag{1}$$

At $pH > 7.5$, HOCl will dissociate to form hypochlorite ion (OCl^-) (18). In addition, calcium hypochlorite [$Ca(OCl)_2$] and sodium hypochlorite (NaOCl) also immediately dissociate in water to form OCl^-. HOCl is a better disinfectant than OCl^- due to the relative ease with which it can penetrate cell walls (5). Hence, the pH of the water has to be lowered to improve the disinfection efficacy.

$$Ca(OCl)_2 \rightleftharpoons Ca^{2+} + 2OCl^- \tag{2}$$

$$NaOCl \rightleftharpoons Na^+ + OCl^- \tag{3}$$

$$H^+ + OCl^- \rightleftharpoons HOCl \tag{4}$$

At $pH > 12$, hypochlorite ions can react to form chlorite ($ClO_2{}^-$) and chlorate ($ClO_3{}^-$), which are of major health concern (19,20).

$$OCl^- + OCl^- \rightarrow ClO_2{}^- + Cl^- \tag{5}$$

$$OCl^- + ClO_2{}^- \rightarrow ClO_3{}^- + Cl^- \tag{6}$$

In waters containing bromides, hypochlorite can react with bromide (Br^-) to form hypobromous acid (HOBr), thereby consuming the available free chlorine meant for disinfection (21).

$$HOCl + Br^- \rightarrow HOBr + Cl^- \tag{7}$$

Chlorine Dioxide

Chlorine dioxide (ClO_2) dissolves in water under alkaline conditions to form chlorite ($ClO_2{}^-$) and chlorate

Table 2. Disinfectant Use in the United States

Disinfectant	Medium and Large Systems (>10,000 People)	Small Systems (<10,000 People) Using Groundwater	Small Systems (<10,000 People) Using Surface Water
Chloramines	29%	—	2%
Chlorine dioxide	8%	—	6%
Chlorine gas	84%	61%	82%
Hypochlorite	20%	34%	26%
Ozone	6%	—	—

Source: References 36 and 37.

(ClO_3^-)(22,23). However, unlike free chlorine, hypobromous acid and organohalogen DBPs are not produced.

$$2ClO_2 + 2OH^- \rightleftharpoons ClO_2^- + ClO_3^- + H_2O \quad (8)$$

The half-life of aqueous ClO_2 decreases substantially with increasing ClO_2 molar concentration and with pH values above 9 (22,24). Even at neutral pH values and low ClO_2 concentrations, the half-life of ClO_2 is usually less than 24 h. Hence, the storage of stock solutions of ClO_2 for even a few hours is impractical.

Ozone

Dissolved ozone (O_3) either reacts with other dissolved solutes in water to form DBPs or decomposes in water to form highly reactive hydroxyl radicals ($OH^\bullet$) (25). Both ozone and hydroxyl radicals are effective disinfectants. Unlike ozone, hydroxyl radicals are highly nonselective in reactions with organic and inorganic solutes and are easily consumed, thereby limiting the amount available for disinfection.

$$O_3 + H_2O \rightarrow O_2 + OH^\bullet \quad (9)$$

The kinetics of ozone reactions are favorable for disinfection and oxidation of many organic and inorganic contaminants in drinking water. However, for many difficult-to-oxidize organic compounds such as chloroform, some alcohols, and saturated alkyl moieties, the kinetics of ozone oxidation may be very slow (26–28).

Chloramines

When ammonia is added to drinking water containing free chlorine, hypochlorous acid will react with ammonia to form one of three types of chloramines depending on temperature, pH, and reaction time (29). Monochloramine (NH_2Cl) and dichloramine ($NHCl_2$) are formed between pH 4.5 and 8.5 and are commonly found at pH typical of drinking water (pH 6–8.5). Monochloramine is most common when the pH is above 8. When the pH is below 4.5, the most common form of chloramine is trichloroamine (NCl_3), which produces a very foul odor and is sometimes found in swimming pools. Chloramines are effective disinfectants against bacteria but not viruses (30,31).

$$NH_3 + HOCl \rightarrow NH_2Cl + H_2O \quad (10)$$

$$NH_2Cl + 2HOCl \rightarrow NHCl_2 + 2H_2O \quad (11)$$

$$NHCl_2 + 3HOCl \rightarrow NCl_3 + 3H_2O \quad (12)$$

TOXICOLOGY OF DISINFECTANTS

At high concentrations, chlorine gas, chloramine, chlorine dioxide, dichloramine, monochloramine, and ozone have been shown to be strong respiratory irritants in humans (32). Hypochlorite is highly caustic and causes skin and esophagus irritation even at low concentrations (33). When household cleaning agents containing hypochlorite and ammonia are mixed, chloramine gas is formed, which can cause severe lung injury (34). The limited toxicological studies on the above-mentioned disinfectants suggest that their toxicity is likely not relevant at the low concentrations encountered in drinking water (35).

TYPICAL USE IN THE UNITED STATES

Chlorine-based disinfectants are used in more than 95% of the water treatment systems across the United States. Table 2 provides a breakdown of the percentages of water treatment systems that use the various disinfectants. Some water treatment systems use multiple disinfectants due to different primary and secondary disinfection practices. Additional information on the use of various disinfectants in other countries, their advantages and disadvantages, including DBP formation, are available in CHLORINATION BYPRODUCTS and ALTERNATIVE DISINFECTANTS.

BIBLIOGRAPHY

1. CDC (2000). *Safe Water Systems for the Developing World: A Handbook for Implementing Household-Based Water Treatment and Safe Storage Projects*. Department of Health and Human Services, Centers for Disease Control and Prevention, Atlanta.
2. American Water Works Association (1999). *Water Quality and Treatment*, 5th Edn. McGraw Hill, New York.
3. Najm, I. and Trussell, R.R. (1999). New and emerging drinking water treatment technologies. In: *Identifying Future Drinking Water Contaminants*. National Research Council, National Academy Press, Washington, DC.
4. Hall, E.L. and Dietrich, A.M. (2000). A brief history of drinking water. *Opflow* **26** (June).
5. White, G.C. (1972). *Handbook of Chlorination. For Potable Water, Wastewater, Cooling Water, Industrial Processes, and Swimming Pools*. Van Nostrand Reinhold, New York.
6. Langlais, B., Reckhow, D.A., and Brink, D.R. (1991). *Ozone in Water Treatment: Application and Engineering*. Lewis Publishers, Chelsea, MI.

7. Pontius, F. (1990). *Water Quality and Treatment*, 4th Edn. American Water Works Association/McGraw-Hill, New York.
8. Sobsey, M.D. (2002). *Managing Water in the Home: Accelerated Health Gains from Improved Water Supply*. World Health Organization, Geneva. Report WHO/SDE/WSH/02.07.
9. Trussell, R.R. (1991). Control strategy 1: alternative oxidants and disinfectants. Presentation at the 98th Annual American Water Works Association Conference.
10. Craun, G.F. et al. (1994). Balancing chemical and microbial risks of drinking water disinfection, Part I. Benefits and potential risks. *J. Water Res. Sci. Technol. Aqua.* **43**(4): 192–199.
11. Craun, G.F. et al. (1994). Balancing chemical and microbial risks of drinking water disinfection, Part II. Managing the risks. *J. Water Res. Sci. Technol. Aqua* **43**(5): 207–218.
12. Chick, H. (1908). An investigation of the laws of disinfection. *J. Hygiene* **8**: 92–157.
13. Haas, C.N. and Karra, S.B. (1984). Kinetics of microbial inactivation by chlorine-1. *Water Res.* **18**(11): 1443–1449.
14. Sobsey, M.D. (1989). Inactivation of health-related microorganisms in water by disinfection processes. *Water Sci. Technol.* **21**(3): 179–195.
15. Haas, C.N. (1990). Disinfection. In: *Water Quality and Treatment: A Handbook of Community Water Supplies*, F. Pontius (Ed.). American Water Works Association/McGraw-Hill, Washington, DC.
16. Kinman, R.N. and Layton, R.F. (1976). New method for water disinfection. *J. Am. Water Works Assoc.* **68**(6): 298–302.
17. Baker, R.J. (1969). Characteristics of chlorine compounds. *J. Water Pollut. Control Fed.* **41**: 482–485.
18. Morris, J.C. (1966). The acid ionization constant of HOCl from 5 °C to 35 °C. *J. Phys. Chem.* **70**: 3798–3805.
19. Gordon, G., Adam, L., and Bubnis, B. (1995). Minimizing chlorate ion formation. *J. Am. Water Works Assoc.* **87**(6): 97–106.
20. Gordon, G. (1997). Predicting liquid bleach decomposition. *J. Am. Water Works Assoc.* **89**(4): 142–149.
21. Farkas, L., Lewin, M., and Bloch, R. (1949). The reaction between hypochlorites and bromides. *J. Am. Chem. Soc.* **71**: 1988–1991.
22. Gordon, G., Kieffer, R.G., and Rosenblatt, D.H. (1972). The chemistry of chlorine dioxide. *Prog. Inorg. Chem.* **15**: 201–286.
23. Gordon, G. and Rosenblatt, A. (1996). Gaseous, chlorine-free chlorine dioxide for drinking water. In: *Proceedings of the Water Quality Technology Conference*. New Orleans, LA. American Water Works Association, Denver, CO.
24. Granstrom, M.L. and Lee, G.F. (1957). Rates and mechanisms of reactions involving oxychloro compounds. *Public Works* **88**(12): 90–92.
25. Reckhow, D.A. (1999). Control of disinfection by-product formation using ozone. In: P.C. Singer (ed.), *Formation and Control of Disinfection By-Products in Drinking Water*, AWWA, Denver, CO.
26. Hoigne, J. and Bader, H. (1979). Ozonation of water: selectivity and rate of oxidation of solutes. *Ozone Sci. Eng.* **1**(1): 73–85.
27. Staehelin, J. and Hoigne, J. (1983). Mechanism and kinetics of decomposition of ozone in water in the presence of organic solutes. *Vom Wasser* **61**: 337–348.
28. Hoigne, J., Bader, H., Haag, W.R., and Staehelin, J. (1985). Rate constants of reactions of ozone with organic and inorganic compounds in water: III. Inorganic compounds and radicals. *Water Res.* **19**(8): 993–1004.
29. Morris, J.C. (1965). Kinetics reactions between aqueous chlorine and nitrogen compounds. Fourth Rudolphs Research Conference, Rutgers University, June 15–18, 1965.
30. Chang, S.L. (1971). Modern concept of disinfection. *ASCE J. Sanitary Eng. Div.* **97**: 689–707.
31. Weber, G.R., Bender, R., and Levine, M. (1940). Effect of ammonia on the germicidal efficiency of chlorine in neutral solutions. *J. AWWA* **32**(11): 1904–1912.
32. MMWR (1991). Chlorine gas toxicity from mixture of bleach with other cleaning products—California. *MMWR* **40**(36): 619–621.
33. Cotter, J.L., Fader, R.C., Lilley, C., and Herndon, D.N. (1985). Chemical parameters, antimicrobial activities, and tissue toxicity of 0.1 and 0.5% sodium hypochlorite solutions. *Antimicrob. Agents Chemother.* **28**: 118–122.
34. Tanen, D.A., Graeme, K.A., and Raschke, R. (1999). Severe lung injury after exposure to chloramine gas from household cleaners. *N. Engl. J. Med.* **341**: 848–849.
35. WHO (2000). *Environmental Health Criteria 216: Disinfectants and Disinfectant By-products*, World Health Organization, Geneva.
36. American Water Works Association (2000). Committee report: disinfection at small systems. *J. Am. Water Works Assoc.* **92**(5): 24–31.
37. American Water Works Association (2000). Committee report: disinfection at large and medium-sized systems. *J. Am. Water Works Assoc.* **92**(5): 32–43.

DISINFECTION

National Drinking Water
Clearinghouse

Disinfection is an important step in ensuring that water is safe to drink. Water systems add disinfectants to destroy microorganisms that can cause disease in humans. The Surface Water Treatment Rule requires public water systems to disinfect water obtained from surface water supplies or groundwater sources under the influence of surface water.

Primary methods of disinfection are chlorination, chloramines, ozone, and ultraviolet light. Other disinfection methods include chlorine dioxide, potassium permanganate, and nanofiltration. Since certain forms of chlorine react with organic material naturally present in many water sources to form harmful chemical byproducts, the U.S. Environmental Protection Agency has proposed maximum levels for these contaminants.

DISINFECTION KEEPS WATER SAFE

Why Disinfect Drinking Water?

Disinfection kills or inactivates disease-causing organisms in a water supply and must provide a 99.9 percent inactivation of *Giardia lamblia* cysts and enteric viruses to protect health and to comply with the U.S. Environmental Protection Agency (EPA) regulations. There are two kinds of disinfection: primary disinfection achieves the desired level of microorganism kill or inactivation, while secondary disinfection maintains a disinfectant residual in the finished water that prevents the regrowth of microorganisms.

What Regulations Govern It?

The EPA Surface Water Treatment Rule (SWTR) requires systems using public water supplies from either surface water or groundwater under the direct influence of surface water to disinfect.

Also, since some disinfectants produce chemical by-products, the dual objective of disinfection is to provide the required level of organism destruction and remain within the maximum contaminant level (MCL) for the SWTR disinfection set by EPA. At this time, an MCL is set for only Total Trihalomethanes, and proposed for additional disinfection by-products.

How is Disinfection Achieved?

Our natural environment contains numerous microorganisms. Most of these present no concerns. However, some—such as *Giardia lamblia* and various viruses, which can be present in water supplies—are extremely harmful and can cause disease in humans. These disease-causing organisms are known as pathogens.

Because pathogens can be present in drinking water supplies, disinfection is very important—the EPA requires it for surface water and groundwater under the influence of surface water. Disinfection treatment methods include chlorination, chlorine dioxide, chloramines, ozone, and ultraviolet light.

When combined with conventional treatment, such as coagulation, flocculation, sedimentation, and filtration, good results have been obtained. Direct filtration, slow sand filtration, and diatomaceous earth filtration, along with disinfection, have been just as successful.

Groundwater systems that disinfect may have to add filtration if the water contains iron and manganese. In fact, insoluble oxides form when chlorine, chlorine dioxide, or ozone are added to these systems. Both ozonation and chlorination may cause flocculation of dissolved organics, thus increasing turbidity and necessitating filtration. The effectiveness of disinfection is judged by analyzing for an indicator organism (total coliform bacteria). This organism is considered harmless, but its presence indicates that pathogens may also have survived.

COMPARING DISINFECTANTS

Chlorination (Gas)

At normal pressures, elemental chlorine is a toxic, yellow-green gas, and is liquid at high pressures.

Advantages. Chlorine is very effective for removing almost all microbial pathogens and is appropriate as both a primary and secondary disinfectant.

Limitations. Chlorine is a dangerous gas that is lethal at concentrations as low as 0.1 percent air by volume.

Process. Chlorine gas is released from a liquid chlorine cylinder by a pressure reducing and flow control valve operating at a pressure less than atmospheric. The gas is led to an injector in the water supply pipe where highly pressurized water is passed through a venturi orifice creating a vacuum that draws the chlorine into the water stream. Adequate mixing and contact time must be provided after injection to ensure complete disinfection of pathogens. It may be necessary to control the pH of the water.

Equipment. A basic system consists of a chlorine cylinder, a cylinder-mounted chlorine gas vacuum regulator, a chlorine gas injector, and a contact tank or pipe (see Fig. 1). Prudence and/or state regulations would require that a second cylinder and gas regulator be provided with

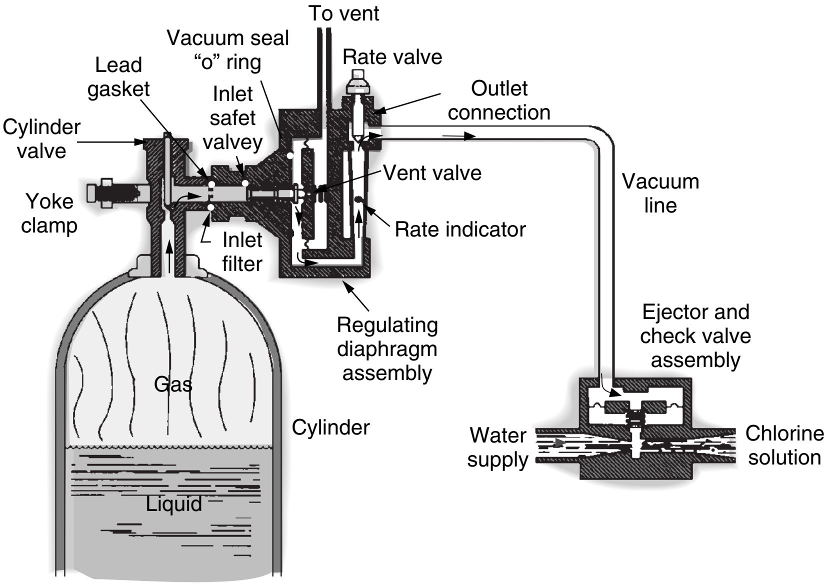

Figure 1. Cylinder-mounted chlorinator. Reprinted with permission from Capital Controls Company, Inc.

a change-over valve to ensure continuity of disinfection. Additional safety and control features may be required.

A gas chlorinator should be installed in a room or chamber with direct emergency access to outside air and fitted with an exhaust fan ventilation system.

Federal and state safety regulations must be observed. If not onsite, self-contained breathing apparatus and a chlorine cylinder repair kit should be available within a reasonable time frame and/or distance.

Chemicals. Chlorine gas is supplied as liquid in high pressure cylinders.

Chlorination (Sodium Hypochlorite Solution)

Sodium hypochlorite is available as a solution in concentrations of 5 to 15 percent chlorine, but is more expensive than chlorine gas (as available chlorine).

Advantages. Sodium hypochlorite is easier to handle than gaseous chlorine or calcium hypochlorite.

Limitations. Sodium hypochlorite is very corrosive and should be stored with care and kept away from equipment that can be damaged by corrosion. Hypochlorite solutions decompose and should not be stored for more than one month. It must be stored in a cool, dark, dry area.

Process. Sodium hypochlorite solution is diluted with water in a mixing/holding tank. The diluted solution is injected by a chemical pump into the water supply pipe at a controlled rate. Adequate mixing and contact time must be provided.

Equipment. A basic liquid chlorination system, or hypochlorinator, includes two metering pumps (one serving as a standby), a solution tank, a diffuser (to inject the solution into the water), and tubing.

Chemicals. Sodium hypochlorite solution is readily available.

Sodium hypochlorite can also be generated onsite by electrolysis of sodium chloride solution in specialized proprietary equipment. The only supplies required are common salt and electricity. Hydrogen is given off as a by-product and must be safely dispersed.

Chlorination (Solid Calcium Hypochlorite)

Calcium hypochlorite is a white solid that contains 65 percent available chlorine and dissolves easily in water.

Advantages. When packaged, calcium hypochlorite is very stable, allowing a year's supply to be bought at one time.

Limitations. Calcium hypochlorite is a corrosive material with a strong odor that requires proper handling. It must be kept away from organic materials such as wood, cloth, and petroleum products. Reactions between calcium hypochlorite and organic material can generate enough heat to cause a fire or explosion. Calcium hypochlorite readily absorbs moisture, forming chlorine gas. Therefore, shipping containers must be emptied completely or carefully resealed.

Process. Calcium hypochlorite may be dissolved in a mixing/holding tank and injected in the same manner as sodium hypochlorite. Alternatively, where the pressure can be lowered to atmospheric, such as at a storage tank, tablets of hypochlorite can be directly dissolved in the free flowing water by a proprietary device that provides flow-proportional chlorination with gravity feed of the tablets.

Equipment. The equipment used to mix the solution and inject it into the water is the same as that for sodium hypochlorite. Solutions of 1 or 2 percent available chlorine can be delivered by a diaphragm-type, chemical feed/metering pump or by tablet chlorinator.

Chemicals. Calcium hypochlorite can be purchased in granular, powdered, or tablet form.

Chloramine

Chloramines are formed when water containing ammonia is chlorinated or when ammonia is added to water containing chlorine (hypochlorite or hypochlorous acid).

Advantages. An effective bactericide that produces fewer disinfection by-products, chloramine is generated onsite. Usually, chloramine-forming reactions are 99 percent complete within a few minutes.

Limitations. Chloramine is a weak disinfectant. It is much less effective against viruses or protozoa than free chlorine. Chloramine is appropriate for use as a secondary disinfectant to prevent bacterial regrowth in a distribution system. Nitrogen trichloride appears to be the only detrimental reaction. It may be harmful to humans and imparts a disagreeable taste and odor to the water. The use of the proper amounts of each chemical reactant will avoid its production.

Process. Chlorine (gaseous solution or sodium hypochlorite) is injected into the supply main followed immediately by injection of ammonia (gaseous solution or as ammonium hydroxide). As before, adequate mixing and contact time must be provided. The mix of products produced when water, chlorine, and ammonia are combined depends on the ratio of chlorine to ammonia and the pH of the water. Chlorine-to-ammonia ratios of 5:1 should not be exceeded. If the pH drops below 5, some nitrogen trichloride may be formed.

Equipment. The generation of chloramines requires the same equipment as chlorination (gaseous or aqueous hypochlorination), plus equipment for adding ammonia (gaseous or aqueous).

All chlorine added to drinking water must meet American National Standards Institute (ANSI), and NSF *International*, formerly the National Sanitation Foundation (NSF) standards. *ANSI/NSF Standard 60: Drinking Water Chemicals—Health Effects* covers water treatment chemicals.

Chemicals. Chemicals used to generate chloramine from ammonia and chlorine gas depend on the ammonia-based chemical used. Anhydrous ammonia is the least expensive, while ammonium sulfate is the most expensive.

Ozonation

Ozone, an allotrope of oxygen having 3 atoms to each molecule, is a powerful oxidizing and disinfecting agent. It is formed by passing dry air through a system of high voltage electrodes.

Advantages. Requiring shorter contact time and dosage than chlorine, ozone is widely used as a primary disinfectant in many parts of the world—but is relatively new to the U.S. Ozone does not directly produce halogenated organic materials unless a bromide ion is present.

Limitations. Ozone gas is unstable and must be generated onsite. A secondary disinfectant, usually chlorine, is required because ozone does not maintain an adequate residual in water.

Process. The five major elements of an ozonation system are:

- air preparation or oxygen feed;
- electrical power supply;
- ozone generation—usually using a corona discharge cell consisting of two electrodes;
- ozone contact chamber; and
- ozone exhaust gas destruction.

Equipment. Ozonation equipment includes air preparation equipment; an ozone generator, contactor, destruction unit; and instrumentation and controls. The capital costs of ozonation systems are relatively high. Operation and maintenance are relatively complex. Electricity represents 26 to 43 percent of total operating and maintenance costs for small systems.

Chemicals. For many applications, pure oxygen is a more attractive ozone feed gas than air because:

- it has a higher production density,
- it requires lower energy consumption,
- it doubles the amount of ozone that can be generated per unit, and
- it requires smaller gas volumes for the same ozone output, thus lowering costs for ancillary equipment.

Ultraviolet Light (UV)

Ultraviolet (UV) radiation is generated by a special lamp. When it penetrates the cell wall of an organism, the cell's genetic material is disrupted and the cell is unable to reproduce.

Advantages. UV radiation effectively destroys bacteria and viruses. As with ozone, a secondary disinfectant must be used to prevent regrowth of micro-organisms. UV radiation can be attractive as a primary disinfectant for small systems because:

- it is readily available,
- it produces no known toxic residuals,
- it requires short contact times, and
- the equipment is easy to operate and maintain.

Limitations. UV radiation may not inactivate *Giardia lamblia* or *Cryptosporidium* cysts, and should be used only by groundwater systems not directly influenced by surface water—where there is virtually no risk of protozoan cyst contamination. UV radiation is unsuitable for water with high levels of suspended solids, turbidity, color, or soluble organic matter. These materials can react with or absorb the UV radiation, reducing the disinfection performance.

Process. The effectiveness of UV radiation disinfection depends on the energy dose absorbed by the organism, measured as the product of the lamp's intensity (the rate at which photons are delivered to the target) and the time of exposure. If the energy dosage is not high enough, the organism's genetic material might only be damaged instead of destroyed. To provide a safety factor, the dosage should be higher than needed to meet disinfection requirements.

Equipment. UV lamps and a reactor (see Fig. 2).

Chemicals. No chemical oxidant required; therefore, microorganisms can be killed without generating byproducts of chemical oxidation or halogenation.

HOW DO YOU CONTROL DISINFECTION BYPRODUCTS?

A number of factors can affect the formation of disinfection byproducts. These include the types and concentrations of organic materials present when chlorine is added, the dosage of chlorine, the temperature and pH of the water, and the reaction time.

To control the formation of halogenated byproducts (compounds formed by the reaction of a disinfectant, such as chlorine with organic material in the water supply) during chlorination, EPA has identified these three strategies:

1. Remove the byproducts after they are formed, which can be difficult and costly.
2. Use alternative disinfectants that do not produce undesirable byproducts, which is often the most cost-effective strategy.
3. Reduce the concentration of organics in the water before oxidation or chlorination to minimize the formation of byproducts. This will provide the highest quality finished water.

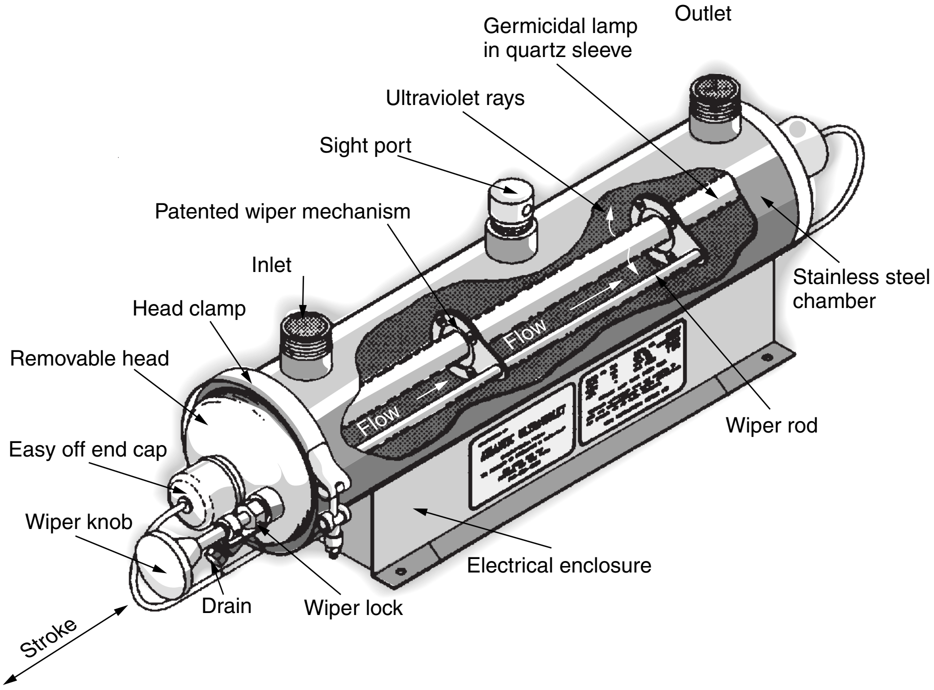

Figure 2. Ultraviolet water purifier. Reprinted with permission from Atlantic Ultraviolet Corporation.

WHERE CAN I FIND MORE INFORMATION

Information on disinfection was primarily obtained from two sources: Environmental Pollution Control Alternatives: Drinking Water Treatment for Small Communities, *EPA/625/5-90/025*; *and* Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities, *EPA/625/4-89/023. Both can be ordered free from the EPA Office of Research and Development at (513) 569-7562.*

These publications also can be ordered from the National Drinking Water Clearinghouse (NDWC); however, copying costs apply. The first book, item #DWBKGN09, an 82-page publication, costs $11.82; and the second, item #DWBKDM04, a 209-page book, costs $30.05. Shipping and handling charges apply.

For further information or to order additional copies of "Tech Brief: Disinfection," item #DWBRPE47, or the above publications call the NDWC at (800) 624-8301.

WATER DISTRIBUTION SYSTEM OPERATION

Mario O. Buenfil-Rodriguez
National University of Mexico &
Mexican Institute of Water
Technology
Cuernavaca, Morelos, Mexico

OPERATION, MAINTENANCE, AND DESIGN OF WATER SERVICES

Contemporary cities and their inhabitants depend on a reliable and generous public water supply system for their functioning and to protect and promote health, productivity, and well-being. The possibility of having piped water and sewerage systems at home explains why cities can support such huge human concentrations. Some major technological and cultural changes to make cities less dependent on such public services are needed but still far in the future. These vast systems, almost invisible to many, as they are underground networks, are complex and delicate arteries that require daily operation and care by professional people devoted exclusively to such tasks.

Various and ample are the duties and responsibilities of any water utility, although there are some variants around the world. Some utilities deal exclusively with clean water distribution; others have broader functions such as extraction from a remote site, conduction, purification, distribution, sewerage, waste treatment, and even tasks such as operation of pipe networks for fire combat, or selling treated wastewater, pluvial drainage, solid waste collection, or energy supply.

The operational tasks for distributing clean water in a city involve monitoring flows, pressures, storage levels, and water quality at different network locations; controlling and moving valves; attending to reports and enquiries from the public; billing and charging for consumption; keeping updated infrastructure maps and consumer records and census; repairing leaks; and replacing or improving pipes, hydrants, pumps, and other water distribution networks components.

There is no clear or definite boundary between operational and maintenance tasks, and their distinction is merely the intensity or frequency of a given job. Sometimes, the same employee team performs operational, preventive, or corrective maintenance tasks. In contrast, other rather specialized tasks can require a contract with a private service provider (*as analyzing water samples to detect a specific and difficult pollutant or controlling and supervising a sophisticated automated pump*).

There is a strong bond between the system's "hydraulic design" and its "operation," and often they can be synonymous. For instance, pressure may be controlled automatically by properly locating storage tanks or automatic valves. Although in some not-so-well-designed systems, frequent manual movements (*operation*) of valves may be needed. The same can be said for filling and emptying water volumes in regulating tanks to meet differences in demand and supply flows. When storage tanks are properly sized to accommodate daily variations of demand, with timing of supply pumps, they will function correctly, but if not, a lot of work is required from "operations personnel" (*usually accompanied by uneasiness and complaints from consumers*).

One practical definition of our topic, although not accepted everywhere (*and not even here*), is this: *Operation* is the organized procedure for causing a piece of equipment, a treatment plant, or other facility or system to perform its intended function, but not including the initial building or installation of the unit. Usually operation and maintenance are referred simultaneously, as *O & M*, meaning the management of a facility involving operating, repair, and replacement.

Another more general definition (*World Bank*) states that *operation* includes the planning and control of the extraction/collection, treatment, conveyance, and delivery of water, and/or the collection, treatment, and disposal of effluent. It also covers the management of client and public relations, legal, personnel, commercial, and accounting functions.

DESIRABLE OPERATING CONDITIONS

Not all cities offer the same service standards of water quality, pressure, and reliability, although most of them guide their performance and duties by a certain set of local goals or regulations. Sometimes, an external autonomous supervisory board exists, sponsored by consumers or by an international agency, to guarantee the provider's accountability in protecting people's health and economy.

Some of these operational goals may be

— Continuity of supply (*24 hours, 365 days per year*).

— Keep low water tariffs (*in balance with expenditures and investments*), and in an increasing block fashion (*the more you consume, the more you pay per water unit to keep the water demand as small as possible and preserve the resource*).

— Promote and advise customers to keep their water consumption low (*demand management*).

— Keep pressure not too weak, nor too strong at each house connection (*around 10 m to 30 m of water column, that is, between 1 to 3 bars*).

— Curb water leaks (*pipes burst*), and when they occur, repair them within a few hours.

— Satisfy daily and seasonal variations in water demand without spillage or damage to infrastructure or inconvenience to consumers.

— Use the least possible energy for pumping and other processes (*low operating costs*).

— Courteous, fair, and undiscriminatory treatment to all consumers.

— Water quality in all points of the network must meet established standards.

— Optimize the use of the system's installed capacity for water storage, pumping, and conduction (*profit from and use the existing infrastructure, keep it in good shape, and avoid the need for building more of it*).

ELEMENTS OF WATER DISTRIBUTION SYSTEM

It is convenient that any water supply system operator be simultaneously involved in, or at least aware of, the sewerage system and the treatment and disposal of wastewater. However, focusing only on the supply side, a water network may depend on the water coming from remote sources, purification plants, aqueducts and pumping stations. All these installations rely on various minor components and devices such as gates, valves, cisterns, and filters. All such items require careful operation and maintenance by the same institution attending to the distribution network in the city. Sometimes, the water sources are within the urban area, as in the case of pumping wells intermingled with the network. In that case, the water quality and pressure can be assured merely by local additions of chlorine or equivalent, and through booster pumps or pressure-breaking valves or tanks.

Some of the elements in a water distribution network are elevated tanks, underground reservoirs or cisterns, booster stations, valves of different types, meters (*for water mains and for customer consumption*), pipes, house connections, various types of fittings and protection cages, fire and public hydrants, local chlorination equipment, monitoring and sampling devices, and telemetry and SCADA systems (*supervisory control and data acquisition*) and all their electronic and electric apparatuses and connections. Besides those previously mentioned "fixed" (*on-site*) elements, there is a wide range of necessary mobile equipment such as vehicles, cranes, winches, mobile laboratories, and maintenance and repair equipment.

Water distribution pipes are normally arranged in grid or loop designs to provide easy flow of water by different routes (*particularly during hours of high demand*) and to avoid lines that dead-end and may cause water quality problems. All distribution systems should contain sufficient valving so that any area can be isolated for repair or for scrutiny in case of contamination risk, and only a minimum number of customers should be affected during the disruption.

PRINCIPAL OPERATIONS

Public water systems are designed to provide and maintain a reliable, high-quality water source (*e.g., groundwater or surface water*). Although one important operation to meet drinking water standards is water treatment, it will not be covered here to concentrate on what happens and is required in the water distribution network. Operations there deal with different cares and purposes, among which is preserving the water quality in the network.

The operation of a water distribution system (*inevitably linked to its maintenance and design*) includes attending to all kinds of requests or complaints from customers; meeting local standards and federal regulations; and the upkeep of pipes, storage tanks, and pumps that convey the water. Some of the activities are

- Hydraulic performance to meet water demands
- Map control and updating
- Modeling and corroboration of system's hydraulic behavior
- Records (*telemetry, samplings*), control, and analysis
- Leak detection and repair
- Cross connection control and backflow prevention
- Gauging flow (*district metering areas, DMA*) and consumption (*meters*)
- Billing consumers and collecting money from them
- Exercising of valves
- Water pipe flushing and internal cleaning
- Water quality monitoring
- Chlorination or dechlorination at local points
- Water main repair/replacement
- Electric and electronic equipment conservation and upgrading
- Inspecting quality of materials (*pipes, joints, chlorine*) and performance of equipment (*water meters, valves, etc.*)
- Storage tank maintenance
- Pump maintenance
- Prevention and training against vandalism, terrorism, and emergencies (*drought, flooding, earthquake, etc.*)

It would take too long to give details of all these operations, so only brief comments for some of them will be given in the following sections:

WATER DELIVERY AND MONITORING

The most important operation of any water distributing system is ensuring that enough and timely water arrives at each house. Therefore, one basic operation is the regulation of flows to adapt them to the hours of peak demand (*e.g., from 6 AM to 8 AM*), and to those of low demand (*e.g., during the night*). Usually, in well-designed and developed countries' systems, these maneuvers are performed "automatically" through "*operational storage*" from regulating tanks around the city. These tanks have float or diaphragm actuated valves that close the inflow when full, preventing spillage, and besides can send a signal to halt a remote pump. This exemplifies how operation depends on design, which in poor or simple systems, or in emergencies, has to be done manually.

Pumps and valves can be preprogrammed to start or stop on given schedules and can be monitored and acted-on through automatic or semiautomatic computerized *supervisory control* systems (*SCADA*) that can check pressure and flow at selected network points. On the other hand, they could simply be controlled manually, according to certain seasonal rules (*e.g., rationing of water in drought times*).

Telemetry and the ability to *communicate* efficiently are important in the successful operation and maintenance of a complex and expanded water supply system. One option is the use of mobile radiotelephones. Telemetry can provide a complete updated picture of the status of all components in the network, such as reservoir contents, water movements and trends, and pump stations. Condition monitoring is part of preventive maintenance and will help to minimize unscheduled shutdowns and improve reliability of service. Sites operating as unmanned installations require that intrusion and hazard *alarms* be reported as soon as possible, so that steps can be taken to limit possible damage. Remote operation of pumps, gates, etc. is required to ensure coordinated operation of tandem plants and to alleviate staff shortages.

One important and routine monitoring task is *reading the house meters* of all consumers. This will not only give data for billing according to consumption, but also gives clues about possible leaks in the system or about clandestine connections, when combined with information of *DMA* (*district metering areas*), which is another important monitoring operation. The purpose of these is to have knowledge about volumes supplied and consumed to detect possible money losses, leaks, bursts and waste, and to correct them as soon as possible.

Water balances (*audits*) and statistics will give information to decide if, instead of a local repair, full pipe replacement is more convenient. Water *leaks* can be visible, but often they are hidden, requiring specialized equipment and personnel to locate them.

Continuous *surveying and sampling* of water quality and hydraulic parameters (*pressure, flow direction, volume supplied*) at different locations is an important task. Those records must be compared to written standards and particular utility goals and performance rules. The purpose of that would be to detect failures and correct them, as well as to find opportunities to improve service quality and efficiency, besides the utility's revenues, through the design of better operating strategies. In this regard, an important tool is *hydraulic modeling* of the distribution system.

Hydraulic transients, or *surges*, occur during rapid flow changes in pressurized water conveying pipelines. Normally, such transients (*water hammer*) are considered a problem as they might damage the pipeline due to strong pressure peaks, subatmospheric pressures, or fatigue. By using simulating models incorporated into hydroinformatic tools, it is possible to assess existing water distribution systems, to suggest more efficient hydraulic performance and propose rehabilitative strategies.

PREVENTION AND SAFETY

Many operational tasks could be classed as related to "prevention and safety." Some of them, such as leak detection and control and water quality sampling, have already been mentioned. Some others are cross connection control, backflow prevention, and protection and vigilance against vandalism or emergencies.

Continuous positive pressure is necessary in all distribution systems to prevent a vacuum in the water-supply main. These vacuums can lead to backsiphonage and possible contamination or even collapse of pipe sections.

Cross connection control and backflow prevention are operational programs that both the utility and its customers must implement to prevent contaminants and nonpotable water (*e.g., wastewater, storm water, process water*) from being drawn into the drinking water system.

Backflow typically occurs when the distribution system's pressure drops, due to a water main break or due to firefighting demands. It has two interpretations: (1) a flow condition, induced by a differential in pressure, which causes the flow of water or other liquid into the distribution pipes of a potable water supply from any source or sources other than its intended source; (2) back up of water through a conduit or channel in the direction opposite to normal flow.

Cross connections are physical piped connections between potable water and an unsafe or polluted water source. They threaten water quality and public health through the backflow of such hazardous substances as antifreeze, boiler water, and sewage. Cross connection control operation, besides building codes and regulations to avoid them, consists of inspection routines and checking backflow prevention devices on particularly high-risk service connections (*e.g., a wastewater treatment plant*).

Numerous recommendations, concerns, and routines can be implemented for *general protection and vigilance* against vandalism or emergencies. Their selection and adoption depend on the risks, the characteristics of the utility and its infrastructure, and even the present or expected climate and political and socioeconomic atmosphere. These are some examples: *Lock all doors and set alarms at offices and main installations. Limit access to facilities, especially to water supply reservoirs. Secure hatches, meter boxes, hydrants, manholes, and other access points to the water distribution system. Control access to computer networks and control systems. Write and periodically review security and emergency plans, and train employees frequently. Assess the vulnerability of water sources, drinking water treatment plants, distribution networks, and other key infrastructure elements. Improve the computer system and remote operational security.*

MAINTENANCE AND REPAIR

It is important to operate all network valves periodically, merely to ensure their proper performance when really needed. A dead-end line must be flushed frequently or the customer on that line will receive poor quality water.

Water pipe flushing is a preventive maintenance task intended to preserve the hydraulic characteristics of the pipe, besides the quality of the water flowing through it. It removes any accumulated sediments or other impurities deposited in the pipe. Flushing is performed by isolating sections of the distribution system and opening flushing valves or, more commonly, fire hydrants to cause a large volume of flow to pass through the isolated pipeline and suspend the settled sediment.

Water *mains* may also be *mechanically cleaned* by using swabs or pigs which are devices that are pulled through a section of line to scrape the accumulated debris off the inside of the pipe. The major environmental concern of water pipe flushing is the discharge of flushed water, which may be high in suspended solids and other contaminants that can harm water bodies. The negative impacts of the discharge may be minimized by discharging the flush water into a sanitary sewer of adequate capacity or by discharging the flush water into a separate storm sewer system, preferably using management measures such as a detention pond, where solids can settle before the water is discharged.

Water *mains* must be *repaired* or *replaced* frequently to correct or substitute broken, corroded, or leaking sections of pipe. Either the broken pipe section is replaced or, as is often the case, a repair sleeve is placed around the outside of the broken pipe section and clamped into place. Following the pipe repair, the line is typically flushed and then disinfected with a chlorine solution. The chlorine solution is usually mixed on site using powdered calcium hypochlorite or sodium hypochlorite. Pipe repair and replacement could affect the environment (*erosion of soil, sediments, high concentrations of chlorine, etc.*), and such impacts should be minimized by control measures.

Storage tanks require frequent inspection and maybe occasional repairs. The most frequent repairs are repainting of walls and replacing screens over vents and other points of access to insects, birds, and rodents. Steel tanks are subject to corrosion, which is prevented by painting them regularly.

Pumps must be maintained to ensure that booster and other distribution pumps stay in working order. Their maintenance involves frequent inspection and tests for excessive vibration or noise, providing grease and lubrication regularly, and checking the pump bearings and packing glands.

PUBLIC RELATIONS AND MANAGEMENT

Operation of a water distribution system has much to do with the management of client and public relations, legal, personnel, commercial, and accounting functions.

To be sustainable, an operation must be *financially viable*. Because of the pressure to expand the area served, viability generally implies recovering the costs of operation and maintenance, as well as capital. Public utilities often have difficulty getting approval for increasing their charges to levels that are financially and economically adequate. Sometimes, this is for political reasons, but often it is also because the requests are poorly prepared and their urgency is not well perceived. Higher charges must be justified not only to the parent entity but also to consumers. Public relations campaigns can be helpful here, but quality service is often crucial to securing consumers' acceptance of increases.

Demand management to keep water needs low and congruent with natural local resources is a vital task. For this purpose, education campaigns directed to the public, installation and reading of water meters, appropriate and fair setting up of water tariffs, proper operation and location of consumer information and attention modules (*telephone, internet, direct presence*) are basic. Employees must be trained and aware of their role in transmitting correct information and messages to the public.

READING LIST

ABPA. *Recent Backflow Incidents and Articles*. American Backflow Prevention Association, http://www.abpa.org/, http://www.abpa.org/incidents.htm.

AWWA. (1976). *Water Distribution Operator Training Handbook*. American Water Works Association, Denver, USA: AWWA, http://www.awwa.org/wdoth.htm.

Chase, D., Savic, D., and Walski, T. *Water Distribution Modeling*. Haestad Methods, Inc. http://www.haestad.com/library/books/wdm/authors.asp.

CIWEM. (1984). *Water Distribution Systems. Water practice manual 4*. The Chartered Institution of Water and Environmental Management, CIWEM (formerly Institution of Water Engineers and Scientists, IWES), http://www.ciwem.org.uk/publications/manuals/index.asp.

Environment Canada. (1998). *The management of water*. National Water Issues Branch, Environment Canada, http://www.ec.gc.ca/water/en/manage/e_manag.htm.

EPA, Environmental Protection Agency, USA. (2001). *Recommendations to Improve Security at Drinking Water Facilities*, http://www.lgean.org/html/whatsnew.cfm?id=259.

Greater Vancouver Regional District Drinking Water Treatment Program. (1997). *Best Management Practice for Pigging and Flushing Water Mains*. http://www.gvrd.bc.ca/services/water/chlorlin/flushing.pdf.

Hawken, P., Lovins, A., and Lovins, L.H. *Aqueous Solutions*. Chapter 11 of book *"Natural Capitalism: Creating the next industrial revolution"*, Rocky Mountain Institute. http://www.natcap.org, http://www.natcap.org/images/other/NCchapter11.pdf http://www.rmi.org/sitepages/pid172.php, http://www.rmi.org/images/other/W-AqueousSol.pdf.

Havlik, V. *Solution of water supply and distribution networks*. Danish Hydraulic Institute (DHI). http://www.dhisoftware.com/mikenet/Download/Papers/hague97.pdf.

James, W. (1999). *Urban Water Systems*. Guelph University, Canada. http://www.eos.uoguelph.ca/webfiles/james/homepage/Teaching/661/wj661Modules.html.

James, W. (1996). *Hydraulics and hydrology vocabulary*. Web Manual. U. of Guelph, Sch of Eng'rg. http://www.eos.uoguelph.ca/webfiles/james/homepage/stillunclassified/wjvocab.html.

Jönsson, L. *Hydraulic transients as a monitoring device*. Department of Water Resources Engineering, University of Lund, Sweden, IAHR, International Association of Hydraulic Engineering and Research. http://www.iahr.org/membersonly/grazproceedings99/doc/000/000/110.htm.

Kenneth, K. (1994). *Water Distribution System Operation and Maintenance*.—3a. Ed.—Sacramento: California Department of Health Services, EPA. http://www.owp.csus.edu/WDS.html.

LGEAN, Local Government Environmental Assistance Network. *Water Distribution System Operation and Maintenance*. http://www.lgean.org/html/lgo/WTRSPPLY-01.html.

Mays, M.W. (1999). *Water Distribution System Handbook*. AWWA American Water Works Association, McGraw-Hill Professional Publishing. http://www.amazon.com/exec/obidos/ASIN/0071342133/002-3075060-0220023.

OIEAU, Office International de l' Eau (International Office for Water). (1998). *Overview of Training Services*. France. http://www.oieau.fr/espagnol/stages/smt97.htm#1997_4 or http://www.oieau.fr/english/stages/smt97.htm#1997_4.

University of Florida. *Backflow / Cross-Connection Control Program*. http://www.dep.state.fl.us/water/drinkingwater/bfp.htm, http://www.TREEO.ufl.edu/backflow/.

World Bank. *Managing Urban Water Supply and Sanitation: Operation and Maintenance*. GDRC, Global Development Research Center. http://www.gdrc.org/uem/water/wb-urbanwater.html.

WATER QUALITY IN DISTRIBUTION SYSTEMS

Z. Michael Lahlou
Technical Assistance
Consultant

A drinking water system's water quality may be acceptable when the water leaves a treatment plant. However, a variety of physical, chemical, and biological transformations can happen once the water enters and travels through a distribution system. Water producers need to understand the sources of water quality degradation during the distribution process because, in addition to taste and odor problems that can occur, research also suggests that degraded water quality increases the risk of gastrointestinal illnesses

WATER QUALITY DETERIORATION FACTORS

A distribution system's pipes and storage facilities constitute a complex network of uncontrolled physical, chemical, and biological reactors that can produce significant variations in water quality. The principal factors that affect water degradation during distribution are the system's structure, its operation, and a number of water quality factors.

STRUCTURAL FACTORS

Drinking water distribution systems are typically thought of as the underground network of interconnecting mains or pipes. They also can include storage facilities, valves, fire hydrants, service connections, and pumping stations.

Historically, water system designers tended to create oversized pipelines and storage facilities. While system designers may be considering an area's future drinking water needs, oversized facilities result in long detention times, loss of chlorine residual, taste and odor concerns, and other water quality problems.

Furthermore, some of the materials designers choose to install in distribution systems create suitable environments for microorganism growth. Materials, such as cast or ductile iron, asbestos-cement, or pressurized concrete, can pit and make way for microorganisms to colonize.

In addition, oxidant-resistant microorganisms settle on pipe surfaces and produce a complex microenvironment known as biofilm. Biofilms form when organisms enter the distribution system and become entrapped in slow-flow areas, line obstructions, or dead-end sections. They usually appear as a patchy mass in pipe sections or as a uniform layer along the inner walls of a storage tank.

While not all biofilm is bad, researchers are currently unsure of its exact effect. Coliform bacteria may colonize in it, and biofilm may interfere with coliform detection. It may also cause taste and odor problems.

Designers now theorize that the material manufacturers use in pipes, as well as the condition of pipes, valves, and storage facilities, may exert a high-chlorine demand.

Pipe materials can cause water quality to deteriorate in other ways. Iron pipes can corrode, and lead and copper from pipe walls can dissolve. For example, unlined or exposed ferrous materials in pipelines can corrode and cause red or rusty-colored water. To avoid corrosion problems, systems are turning to plastic materials, such as polyethylene and polyvinyl chloride (PVC).

Finally, contamination via cross-connection, leaky pipe joints, or pipe breaks may influence water quality. Pathogens, such as *Cryptosporidium* and *Giardia lamblia*, may enter the system through contaminated raw water, in-line reservoirs, or breaks in pipelines. System personnel need to carefully and thoroughly perform flushing and disinfection procedures following repairs.

OPERATIONAL FACTORS

From an operations standpoint, network operating conditions—such as slow water velocities, supply sources going on and off-line, and the amount of time that systems store water—greatly affect water quality. Any of these factors can cause chlorine residual to fade, and, thus, allow microbial growth in the network. Further, hydraulic conditions can cause sediment to deposit, accumulate, and serve as both habitat and protection from disinfectants for microbial growth.

What's more, many storage facilities are kept full so that the system can be better prepared for emergency conditions. However, the long detention times result in degraded water quality.

WATER QUALITY FACTORS

Some of the factors that provide optimal conditions for microorganisms to multiply include long water-detention times in tanks and pipes, adequate nutrient levels, and warm temperatures.

In addition, research has shown that the level of biodegradable organic matter in the distribution system strongly affects bacterial re-growth and harbors opportunistic pathogens. An opportunistic pathogen can be any disease-causing organism, bacterium, virus, helminth, or protozoan that slips through the treatment processes or enters the distribution system during pressure loss and finds the opportunity or favorable circumstances to lodge or reproduce in organic material, bacterial slime, or other material that it finds attractive.

A number of other conditions also can affect water quality. For example, disinfectants may react with organic and inorganic compounds and cause taste and odor problems or form disinfection by-products. Also, particulate re-suspension may cause increased turbidity.

CONTAMINATION PREVENTION AND CONTROL

Drinking water systems can improve water quality or prevent its deterioration in the distribution system. They can modify system operations and maintenance alternatives, make changes in treatment practices, and improve water quality monitoring and modeling. Generally, systems need to find an optimal combination of these actions, which can involve trade-offs between cost, water supply needs, and water quality considerations.

MODIFICATIONS TO SYSTEM OPERATION

Systems can use five primary operation procedures to maintain water quality:

1. minimize bulk water detention time,
2. maintain positive pressure,
3. control the direction and velocity of the bulk water (see Fig. 1),
4. maintain a disinfectant residual in the distribution system (Disinfectant residual, usually chlorine, provides a relatively effective barrier to the growth of microorganisms in bulk water and biofilm.), and
5. prevent cross-connections and backflow.

Utilities should minimize bulk water detention time because the interactions between the pipe walls and the bulk water result in water quality deterioration. Furthermore, stored finished water should be turned over frequently because the stored water's age contributes to the overall water age in the distribution system.

To reduce possible pathogen intrusions, drinking water utilities should maintain minimal water pressure in the distribution network, particularly if cross-connections are present. Further, the utilities should maintain a positive pressure throughout the distribution system to minimize the potential for back siphonage or backflow of contaminants to occur. However, excessive water pressure may cause pipe leaks or even breaks.

Utilities should minimize rapid or extreme fluctuations in flow velocities and should minimize the frequency of reversals. Activities that may affect flow velocities include rapidly opening or closing a valve, a power loss, and hydrant flushing. Changes in flow velocity can scour

Distribution Material	Corrosion Resistance	Potential Contaminants
Copper	Resists corrosion well, but is subject to corrosive attack from high velocities, soft water, chlorine, dissolved oxygen, and low pH.	Copper
Lead	Corrodes in soft water with low pH	Lead, arsenic, and cadmium
Mild steel	Subject to uniform corrosion, particularly sensitive to high dissolved oxygen levels	Iron, resulting in turbidity and red-water complaints
Cast or ductile iron	Aggressive waters can cause surface erosion	Iron, resulting in turbidity and red-water complaints
Galvanized iron	Aggressive waters can cause galvanic corrosion of zinc	Zinc and iron
Asbestos-cement	Good corrosion resistance; aggressive waters can leach calcium from cement	Asbestos fibers
Plastic	Resistant to corrosion	

Figure 1. Corrosion properties of different materials used in distribution systems. *Source*: Larry Mays, 1999.

sediments, tubercles, and deposits from interior pipe surfaces and degrade water quality.

MAINTENANCE ALTERNATIVES

Distribution system flushing is an important tool to keep the water system clean and free of sediment, remove stagnant water, and remove unwanted contaminants that may have inadvertently entered the system. (See summer 2002 *On Tap* "How to Flush a Distribution System.")

Drinking water systems can use a variety of pipeline cleaning techniques. These techniques include mechanical scraping, pigging, swabbing, chemical cleaning, and flow jetting. Utility maintenance also includes emergency pipe repairs with sanitary precautions in place. Utilities should:

- try to keep contaminated water out of the trench and pipe;
- flush the line in the vicinity of the break;
- apply disinfectant to potentially contaminated components;
- disinfect new mains;
- disinfect storage tanks after construction, inspection, or maintenance; and
- conduct bacteriological testing to confirm the absence of contaminates.

Other maintenance activities that utilities can use to minimize water quality degradation:

- prevent and eliminate cross-connections;
- cover and vent storage tanks;
- maintain an adequate separation from sewers; and
- enforce applicable building plumbing codes.

CHANGES IN TREATMENT PRACTICES

Disinfection means using chemicals to inactivate harmful microorganisms that might be present in water. This practice protects distributed water from pathogen regrowth or recontamination. Water systems treating surface water supplies maintain a level of residual chemical disinfectant throughout the distribution system. They also usually use some kind of booster disinfection or a more stable secondary disinfecting chemical, such as chloramines.

On the other hand, excessive chlorine levels will produce taste and odor problems, may accelerate pipe corrosion, may enhance formation of harmful disinfection by-products, or produce health concerns for the system's customers.

Pipes can and do corrode internally, reacting with the water and deteriorating. Internal corrosion can cause toxic metals, such as lead and copper, to leach into water, impart a metallic taste to water, stain plumbing fixtures, harbor nuisance and pathogenic microorganisms, reduce a pipe's hydraulic carrying capacity, and ultimately result in leaks and clogs.

Numerous physical, chemical, and biological factors can increase corrosion's rate and occurrence. Some individual factors can either promote or inhibit corrosion, depending upon other conditions. The most significant factors include temperature, pH, alkalinity, dissolved oxygen, total dissolved solids, hardness, and bacteria.

THREE APPROACHES TO CONTROL CORROSION

1. Modify the water quality to make the water less corrosive, such as adjusting the pH.
2. Lay down a protective lining between the water and the pipe, such as using chemical inhibitors.
3. Switch to plastic pipe, which is less prone to corrosion.

Attempts to control biofilm in the distribution system have taken several directions. To avoid releases of biofilm, utilities can manipulate their water chemistry, such as adjusting pH, alkalinity, or the Langlier Index (an indicator of corrosiveness). They also can apply corrosion inhibitors not only to protect pipe materials but also to firm up the coating of the sediments that harbor microbial communities.

WATER QUALITY MONITORING AND MODELING

Utilities should develop a sampling plan to address monitoring and modeling issues, such as selecting sampling sites, establishing test parameters and monitoring frequencies, establishing field-monitoring protocols, and addressing laboratory considerations. The most commonly monitored test parameters to determine general distribution system water quality include coliform bacteria, heterotrophic plate count (HPC) bacteria, disinfectant residual, temperature, turbidity, pH, and color.

Utilities may find that it is difficult to use monitoring data alone to understand all the possible interactions among the multiple parameters that affect water quality between the treatment plant and the user's tap. The flow pathways and travel times of water through these systems are highly variable because of the looped layout of the pipe network and the continuous changes in water usage over space and time. In addition, utilities commonly use storage facilities that are not part of the distribution system, making matters even more variable.

For these reasons, water utility managers are turning to hydraulic and water quality models as attractive monitoring supplements. They can use these models to perform a variety of water quality-related studies. Systems can:

- use chemical tracers to calibrate and test hydraulic models of the system,
- locate and size storage facilities,
- modify system operations to reduce the age of the water,
- modify the design and operation of the system to provide a desired blend of waters from different sources,
- find the best combination of pipe replacement, pipe relining, pipe cleaning, reduction in storage holding time, and location and injection rate at booster stations to maintain desired disinfectant levels throughout the system,
- assess and minimize the risk of consumer exposure to disinfectant by-products, and
- assess the system's vulnerability to incidents of external contamination.

Furthermore, as water utilities make more use of such tools as geographic information systems (GIS) and supervisory control and acquisition (SCADA) technologies, they can achieve a degree of data integration that enables more reliable network analysis and management.

WHERE CAN I FIND MORE INFORMATION?

American Water Works Research Foundation. 2000. *Guidance Manual for Maintaining Distribution System Water Quality*. Denver, CO.

American Water Works Association. 1999. *Water Distribution Operator Training Handbook*. Denver, CO.

American Water Works Association. 2001. *Rehabilitation of Water Mains (M28)*. Denver, CO.

Besner et al. 2001. “Understanding Distribution System Water Quality.” *Journal of the American Water Works Association*. July, pp: 101–114.

Kirmeyer et al. 2001. “Practical Guidelines for Maintaining Distribution System Water Quality.” *Journal of the American Water Works Association*. July, pp: 62–73.

Lei, J. and S. Saegrov. 1998. “Statistical Approach for Describing Failures and Lifetimes of Water Mains.” *Water Science and Technology*. 38:6:209.

Mays, L. W. 2000. *Water Distribution Systems Handbook*. American Water Works Association. McGraw-Hill: New York.

Z. Michael Lahlou holds a doctorate in environmental and natural resource economics, a masters in civil and environmental engineering, and an MBA. Formerly the technical assistance coordinator for the National Drinking Water Clearinghouse, Lahlou now resides in Huntington Beach, CA.

HAVE YOU READ ALL OUR TECH BRIEFS?

Tech Briefs, drinking water treatment and supply fact sheets, have been a regular feature in the National Drinking Water Clearinghouse (NDWC) publication *On Tap* for more than six years.

A package of *Tech Briefs* is available in a three-ring binder that holds all the *Tech Briefs* currently in print. New selections can easily be added to the package as they become available. To order this free product, call the NDWC and ask for item #DWPKPE71.

Additional copies of the *Tech Brief* fact sheets are also free; however, postal charges may be added.

DESIGN OF WATER DISTRIBUTION SYSTEMS

Srinivasa Lingireddy
Lindell E. Ormsbee
Don J. Wood
D. Ramalingam
University of Kentucky
Lexington, Kentucky

INTRODUCTION

Providing quality service at minimum possible cost is the ultimate goal of drinking water utilities. Quality of service includes delivering adequate and reliable quantities of water at sufficient pressures, meeting drinking water quality standards throughout the distribution system, and providing for adequate fire protection within the service area. Water distribution systems that help achieve this goal are generally designed in three stages: planning, preliminary design, and detailed design. The planning stage should properly identify the service-related constraints (e.g., limits on pressures, flows, and water quality parameters) along with the right set of decision variables (i.e., distribution system components).

The preliminary design should identify the alignment of pipelines, locations for storage and pumping facilities, availability of easement or right-of-way, and so on. Precise quantification of the decision variables takes place in the detailed design phase. The following sections describe various components of water distribution systems, the constraints associated with quality of service, and the steps involved in detailed design.

COMPONENTS OF A WATER DISTRIBUTION SYSTEM

Depending on the size of the population served, a water distribution system (WDS) could be a small network of branching pipes gravity fed by an elevated source to a highly complex web of pipe loops, isolation and control valves, multiple supply sources and associated pumping facilities, elevated/ground/pumped storage tanks, hydrants and blow-off valves, and so on. The following section describes various groups of components and their importance in the design of a WDS.

Pipes and Joints

Pipes and joints constitute the most abundant group of components in a WDS. Pipes owned and maintained by a water utility may generally range from 2 inches (50 mm) to several feet (or meters) in diameter, although for those systems that provide fire protection, a minimum diameter of 6 inches is usually specified (1,2). Pipes for water distribution are available in a variety of material—cast iron, ductile iron, steel, concrete, asbestos cement, polyvinylchloride, polyethylene, and so on. Selection of an appropriate pipe material is an important step in the design of a WDS as it influences initial and maintenance costs, type and number of joints, and hence leakage potential, load bearing capacity, capacity to withstand static and transient (water hammer) pressures, corrosion/tuberculation and consequent loss of capacity, pipe roughness characteristics, and so on. The relative advantages of one material over the other, along with many other relevant data needed for the design, may be obtained from the handbooks prepared by the respective manufacturers associations (3–5). Pipes should be buried to protect them against various natural and other forces, and care must be taken to provide minimum cover based on maximum frost-depth in cold climate regions.

Pumps

Although some water distribution systems rely completely on gravity for supplying water to the consumers, most systems employ pumps in some shape or form. Pumps in a distribution system may appear near clear wells or other sources, at pressure boosting stations, and at pumped storage tanks. Power requirements or the head-flow characteristic curve may define the pump capacity. It is common practice to use system head curves in sizing the pumps. The net positive suction head (NPSH) requirements play a crucial role in locating and sizing the pumps. Variable-speed pumps are becoming more common as they offer a wide range of operating conditions. For detailed description on different types of pumps, their operating characteristics, along with their relative advantages and disadvantages, the readers should refer to ANSI/HI (6) and Sanks et al. (7). Irrespective of the size of the pumps employed, the use of pumps in a WDS adds additional operating costs that should not be overlooked at the design stage. The use of pumps, however, allows for smaller transmission mains, thereby reducing the capital costs. From a design perspective, there will be an optimal tradeoff between the capital cost associated with the pipes and the operating costs associated with the pumps.

Tanks

Tanks in a WDS are used for a variety of reasons. Treated water may be stored in ground storage tanks or clear wells. Elevated storage tanks are generally used as balancing tanks that are filled during slack demand periods so they can feed the distribution system during peak demand periods. Such an operation minimizes frequent cycling (on/off status or pump speed changes) of pumps, thereby permitting efficient operation. Small distribution systems employ hydropneumatic tanks to maintain specified pressure ranges in an effort to minimize or eliminate continuous pump operation. Tank capacities are dictated by the need for emergency storage requirements (e.g., to temporarily offset treatment plant failures), for fire fighting requirements over extended periods, and to cater for peak-hour demands where the treatment plant capacities are limited to maximum-day demands. However, larger tank capacities—beyond the need to cater for peak-hour demands—could lead to water quality problems as water in storage for extended periods tends to lose disinfection potential. As water quality requirements on tank capacities are in direct conflict with other requirements for storage (e.g., fire fighting needs), a careful tradeoff must be arrived at when sizing the tanks. It is not uncommon for most systems to be designed to provide excess storage in the system equivalent to meet an average-day demand.

Valves

Valves in a WDS may be classified into isolation valves and control valves. Isolation valves are needed to isolate portions of WDS during emergency situations such as pipe breaks, or for routine maintenance purposes such as pipe rehabilitation, and flushing operations. Control valves, such as pressure-reducing valves and pressure-sustaining valves, throttle or maintain specified pressures in certain portions of the distribution system. They are also employed to isolate multiple pressure zones within a WDS. Flow control valves may regulate flowrates to certain processes that draw water directly from the WDS. Other types of control valves that are frequently used in a WDS include check valves, pump operational control valves, throttle control valves, and so on. Control of transient pressures might warrant the use of special types of valves such as pressure relief valves and air relief/vacuum valves.

Hydrants

Most water utilities are mandated to provide water for adequate fire protection within the service area. Hydrants connected to the pipes provide easy access for water for fighting fires. Insurance companies in the United States make use of hydrant ratings (hydrant flowrates without dropping the pressures below 20 psi during peak demand conditions) in setting the property insurance premiums. Hydrants are also used for flushing the pipelines as part of routine maintenance.

DESIGN CONSTRAINTS

Efficient design of a WDS must satisfy several constraints. These constraints include water quantity requirements, limits on pressures and velocities, limits on water quality parameters, and fire protection requirements, besides the usual constraints on overall system costs.

Constraints on Water Quantity

A well-designed WDS should supply adequate quantities of water to all consumers, including fire demands. Therefore, estimating the desired quantities of water by various consumers becomes an important task. Normal water consumption originates from three distinct groups of consumers: residential, commercial, and industrial. State and local guidelines on per-capita consumptions would be helpful in estimating the commercial and residential demands (8). Industrial demands should be estimated in consultation with the industries being served. Demands may also be estimated based on land-use and associated consumption data from neighboring systems of similar characteristics (population and commercial activities). Demand calculations for the existing WDS should come from monthly meter records. Peaking factors must be applied to the average daily demands to account for maximum-day demand as well as maximum-hour (peak–hour) demand (9). Most water utilities ensure that the supplies (treatment plant and pumping capacities) are adequate to meet maximum-day demand and rely on the floating storage tank capacities to meet the maximum-hour demand.

Constraints on Pressures and Velocities

It is not only important to deliver adequate quantities of water to the consumers but also to ensure that the water is delivered at adequate pressures. Pressures less than 30 psi under maximum-hour demand conditions are not acceptable for most water systems. Although the upper bound on pressure would depend on the pressure rating of the pipe material or the pipe joints, many water utilities are mandated to maintain pressures well below the rated capacities of pipes (e.g., 70–80 psi) to safeguard the appliances that derive water directly from the distribution network. Limiting the upper bound on pressures would also reduce the leakage losses as well as the pump operating costs (10). Lower and upper bounds on velocities may also be imposed when designing a WDS. Lower bounds on velocity ensure adequate scouring, thereby reducing sedimentation problems. Upper bounds on velocities minimize the transient pressure-related problems.

Constraints on Water Quality

With the new safe drinking water regulations, the utilities are required to ensure delivery of water that meets or exceeds all drinking water standards (limits on physical, chemical, and microbial parameters) right at the consumer connection. In addition, the water supplied should also carry specified minimum levels of residual disinfectant (e.g., chlorine or chloramines). As most disinfectants are nonconservative chemicals (i.e., concentrations drop with time), concentration levels higher than the required minimum values are necessary near the supply sources (11). Very high disinfectant concentration levels may not be acceptable either because of their perceived or real carcinogenicity or simply because of the resulting taste and odor problems. As a last resort, supplemental or chlorine booster stations must sometimes be installed at more remote or isolated sections of the system to ensure residual protection.

Constraints on Fire Demands

Except for certain small systems serving scattered rural communities, most WDS in the United States are expected to provide water for adequate fire protection. Hydrants are expected to provide water at a minimum of 500 gpm without dropping the pressures in the system below 20 psi under maximum-hour conditions. Many hydrants in a WDS are rated and are color coded based on the flowrates they can deliver while maintaining a minimum of 20-psi pressure (12). Hydrants that are capable of supplying 1500 gpm or more are painted light blue-green for 1000–1499-gpm capacity, orange for 500–999-gpm capacity, and red for those that cannot supply 500 gpm. Many utilities are not only mandated to ensure minimum specified instantaneous (or static) hydrant flowrates (e.g., 500 gpm) while maintaining the specified minimum pressures (e.g., 20 psi), but they are also required to ensure adequate hydrant flowrates over extended periods (e.g., 2 hours). The latter requirement would influence the capacity requirements of storage facilities.

DESIGN OF WATER DISTRIBUTION SYSTEMS

The need for designing an entirely new drinking water distribution system develops less frequently in modern times. Most current-day designs involve enhancements to existing systems, such as adding pipelines to serve areas of new development, strengthening the existing pipelines to improve the reliability of service, and/or replacing old and deteriorated pipelines. Current-day designs also include sizing new transmission mains, pumping stations, storage facilities, and so on. Whether a WDS is being designed from the ground-up or a few of its components are being strengthened, the use of mathematical models representing the WDS would greatly simplify the design process. The debate on the use of mathematical models in designing WDS dates back to the days of wood-stove pipes (13). Recent advances in computing power, both in terms of hardware and software, resulted in routine use of mathematical models in designing WDS. In addition, scores of optimization techniques have been proposed in the literature to help with the design process. An optimization technique involves formulating an objective function and the associated constraints as functions of decision variables and solving for the decision variables that optimize the objective function while satisfying all the constraints. Most optimization techniques for WDS design use minimization of annualized total costs as the objective function. Lansey (14) summarizes various methods proposed in literature for optimal design of WDS.

A robust and efficient optimization technique is supposed to provide an optimal set of decision variables that result in a minimal cost design while satisfying all constraints. However, design of water distribution systems involves an extensive and diverse set of decision variables and a complex set of constraints in addition to an elaborate cost function. Although a particular design may not involve all decision variables and constraints outlined in the previous sections, a robust optimization platform should allow the designer to choose any combination of constraints and decision variables besides allowing specification of a desired cost function. In addition, such optimization platforms could also work with existing mathematical models of the water distribution systems. Lack of such robust optimization platforms within the popular WDS modeling software packages prevented the widespread use of seemingly powerful optimization techniques in designing water distribution systems. Most designers continue to work with trial-and-error methods for sizing various components of a WDS with the help of a hydraulic network model representing the WDS being designed. A comprehensive iterative method for the design of a WDS is presented in the following. Although the method presented is meant for the design of new water distribution systems, the method can be easily adapted for designing enhancements to existing systems.

Iterative Method for Designing Water Distribution Systems

- Identify the supply sources: clear wells, natural springs, groundwater wells, and so on. Choose a layout for the distribution network. Start with trunk mains; other mains and laterals may be added later. The points to be connected by trunk mains include supply sources, large consumers, and storage tanks. Use a skeleton layout for preliminary designs and a complete layout for detailed designs. Most systems serving 100,000 populations or more do not model pipes less than 4 inches (100 mm) in diameter to simplify the design process. Most systems exclude all branching pipelines from the design process as well, unless the design is specifically aimed at sizing the branch lines. In addition, the hydraulic modeling software package might limit the number of pipes in the model, which in turn limits the ability to model all pipes in the network.
- Estimate nodal demands, taking future growth into account. Refer to the subsection *Constraints on Water Quantities* for details. If the WDS being designed is for serving a rural community with no mandatory provision for fire flows, special care must be taken in estimating the nodal demands. Refer to the section on *Rural Water Systems* in Appendix A for more details.
- Develop a hydraulic network model for the chosen layout, and provide all known data—supply sources, pump nodes, nodal elevations, demands and pipe lengths, and so on.
- Assign pipe diameters and roughness values for all pipes. Use larger diameters for pipes closer to the supply points and smaller diameters for pipes at the extremities of the system. Estimates of pipe roughness values may be obtained from manufacturer handbooks (3–5).
- If the WDS design involves pump stations, provide an estimated useful power (horsepower or kilowatt units) data for the pumps. Choose an arbitrary pump power that can be adjusted later to satisfy the pressure and flow constraints. If the supply source is a pressure main, collect head-flow data for the location where the distribution system will be connected to the pressure main. Refer to the section on *Pressure Supplies* in Appendix A for more details.
- Identify all constraints on the proposed design. Refer to the section on *Design Constraints* for details.
- Perform hydraulic analysis of the network model, and review the pressures and flowrates throughout the distribution system.
- Adjust pipe and/or pump sizes until pressures and velocities in the distribution system are well within the required range. Make sure the minimum pressures are at least 20% greater than the lower limit and the maximum pressures are at least 20% lesser than the upper limit to allow for eventual variation in pressures with diurnal demand changes.
- Replace the constant power data with suitable head-flow characteristic data obtained from the pump manufacturers. If required, use multiple pumps in parallel and/or series. Satisfactory design should include a standby pump in case of mechanical failure of a duty pump.
- Analyze the network model with maximum-day peak-hour and average-day slack-hour demand conditions, and review the pressures and velocities. Adjust the pipe and/or pump sizes if needed.
- Note the lowest and highest hydraulic gradeline (HGL) values at locations near the elevated storage tanks.
- Update the model to include the elevated storage tanks. Use an overflow elevation slightly below the highest HGL and a tank bowl elevation slightly above the minimum HGL. Choose a tank diameter large enough to hold a predetermined quantity of water (1/4 to 1/2 of the daily average demand or the specified minimum required storage, whichever is greater).
- Run an extended period simulation (EPS) with an appropriate diurnal demand curve. Use an AWWA-recommended diurnal demand curve (9), if no other information on the demand variation is available.
- Adjust the pump/pipe/tank characteristics, and rerun the analysis, if needed.
- Simulate hydrant flows, and verify if the flows to maintain the required residual pressure (e.g., 20 psi) are satisfactory. Revise network data, and rerun the hydrant flow simulation, if needed.
- Perform system reliability studies. Remove (or close) one pipeline at a time from the network model,

and quantify the reduction in service—pressures less than required values (15,16). Provide necessary improvements to the model to enhance the system performance under crippled conditions.

- Run water quality simulations with appropriate bulk and wall reaction rates and supply concentrations. Adjust network characteristics if the residual disinfectant concentrations are not within the specified bounds. Consider booster chlorination stations. Security and easement issues are of biggest concern when designing booster chlorination stations.
- Perform transient analysis on the model with several different operating scenarios to ensure no unwanted low/high pressures are generated. Modify the model characteristics or provide adequate surge protection, if needed. Refer to the section on *Transient Analysis* in Appendix A for more details on the significance of transient modeling studies.

The procedure outlined in this section ensures satisfactory design but not necessarily an optimal design. One could repeat this process by adjusting design variables, each time quantifying the total cost of the system. The readers should refer to the section on *Cost Calculations* in Appendix A for a discussion on various costs associated with WDS design and the use of annualized costs to compare various design alternatives. Figure 1 shows the design procedure in a flowchart format.

SUMMARY

The design of water distribution systems is a complex and tedious process. Several optimization techniques have been reported in literature to automate and reduce the complexity of the design process. However, lack of robust optimization platforms that allow specification of preferred objective functions and desired combinations of constraints and decision variables have seriously limited the use of available optimization technology for water distribution systems design. This article presents a comprehensive iterative method that makes use of a hydraulic network model for design of water distribution systems. Although the procedure presented is for design of new water distribution systems, the method can be easily adapted for sizing enhancements to existing systems.

BIBLIOGRAPHY

1. TSS (2003). *Recommended Standards for Water Works (Ten State Standards), Great Lakes-Upper Mississippi River Board of State Sanitary Engineers*. Health Education Services, Albany, NY.
2. AWWA (2002). *ANSI/AWWA C151/A21.51—Specification for Ductile Iron Pipe and Fittings for Water and Other Liquids*. American Water Works Association.
3. ACPA (2000). *Concrete Pipe Design Manual*. American Concrete Pipe Association, Texas.
4. CISPI (1994). *Cast Iron Soil Pipe and Fittings Handbook*. Cast Iron Soil Pipe Institute, Tennessee.
5. UBPPA (1991). *Handbook of PVC Pipe—Design and Construction*. Uni-Bell Pipe Association, Texas.
6. ANSI/HI (2002). ANSI/HI Pump Standards Set. Hydraulic Institute, Parsippany, NJ.
7. Sanks, R.L., Tchobanoglous, G., Bosserman, B.E., II, and Jones, G.M. (1998). Pumping *Station Design*. Butterworth-Heinemann, Massachusetts.
8. VAC (2002). *Chapter 590—Water Works Regulations*. Virginia Administrative Code.
9. AWWA. M32 (1989). *Manual 32—Distribution Network Analysis for Water Utilities*. American Water Works Association.
10. Lingireddy, S. and Wood, D.J. (1998). Improved operation of water distribution systems using variable speed pumps. *J. Energy Eng. Am. Soc. Civil Eng.* **124**(3).
11. Lingireddy, S., Ormsbee, L.E., and Wood, D.J. (1995). Time averaging water quality assessment for water distribution systems. *J. Am. Water Works Assoc.*
12. NFPA (2002). NFPA 291—Recommended practice for fire flow testing and marking of hydrants. National Fire Protection Association, Massachusetts.
13. Ramalingam, D., Lingireddy, S., and Ormsbee, L.E. (2002). History of water distribution network analysis: Over 100 years of progress. *Proceedings of the ASCE 150th Anniversary Conference on Environmental and Water Resources History*, EWRI of ASCE, Washington, DC.
14. Lansey, K.E. (2000). Optimal design of water distribution systems. In: *Water Distribution Systems Handbook*, Mays, L.W. (Ed.). McGraw-Hill, New York.
15. Lingireddy, S., Ormsbee, L.E., and Neelakantan, T.R. (2000). Application of genetic optimization in maximizing hydraulic network redundancy. *Proceedings of the 2000 Joint Conference on Water Resource Engineering and Water Resources Planning and Management*, ASCE, Section 51, Chapter 4.
16. Ormsbee, L. and Kessler, A. (1990). Optimal upgrading of hydraulic-network reliability. *J. Water Resour. Planning Manage., ASCE* **116**(6): 784–802.
17. Means, R.S. (2004). *Site Work & Landscape Cost Data*. R.S. Means Company, Inc. Kingston, MA.
18. LeChevallier, M.W., Gullick, R.W., and Karim, M. (2002). The potential for health risks from intrusion of contaminants into the distribution system from pressure transients. American Water Works Service Company, Inc., Voorhees, NJ. Distribution System White Paper prepared for the U.S. Environmental Protection Agency.
19. Wood, D.J. and Lingireddy, S. (2004). *Surge2000 Short Course Training Manual*. University of Kentucky, Lexington, KY.
20. Lingireddy, S., Wood, D.J., and Nelson, A. (1998). Modified network model for rural water systems. *J. Water Resour. Planning Manage., ASCE*, **124**(5): 296–299.

APPENDIX A

Cost Calculations

Cost is a major factor in designing a water distribution system. Two types of costs influence the decision-making process: capital costs and operating costs. One way of comparing competing designs is to annualize the capital costs and add the annual operating costs to arrive at annualized total costs. If C_P is the total capital cost, i is the annual interest rate for a period of N years and then the annualized capital cost C_A is given by

$$C_A = C_P\left[\frac{i}{(i+1)^N - 1} + i\right]$$

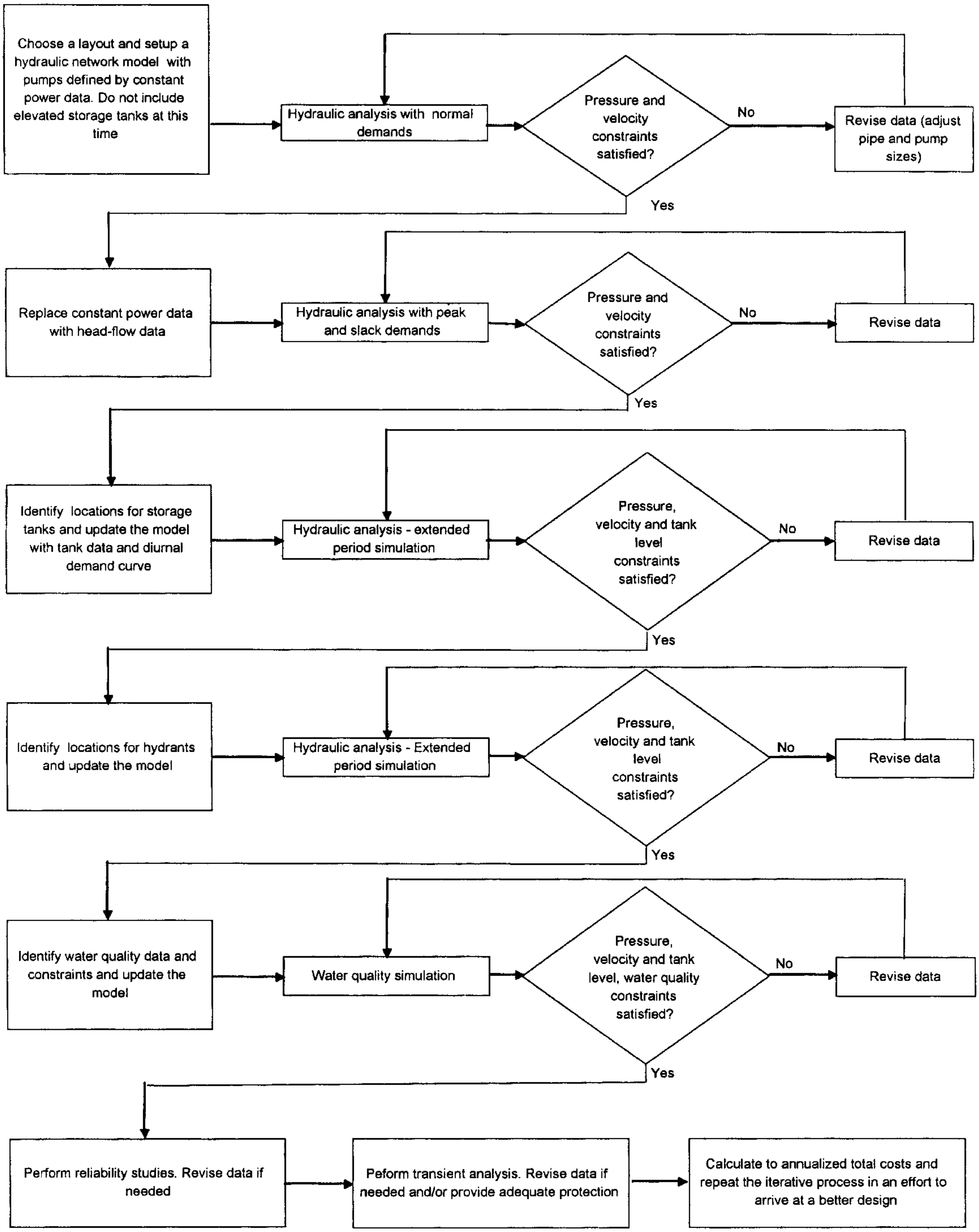

Figure 1. Iterative method for designing water distribution systems.

The total capital costs comprise primarily cost of easement or right-of-way, cost of piping including joints, excavation and backfill costs, cost of pumps and pumping station, cost of storage facilities, and cost of valves. Means' Construction Data Handbooks, in particular, the *Site Work and Landscape Cost Data Handbook* are excellent sources of information on capital costs (17). Operating costs are primarily the energy costs resulting from pumping operation. The energy costs vary inversely and nonlinearly with the piping costs. That is, the pressure constraints in the system could be met by using small-diameter pipes that require large-pump energy requirements or by using large-diameter pipes that require small-pump energy requirements. Figure A1 shows a typical capital and energy costs tradeoff curve.

Pressure Supplies

When designing pipelines serving a new subdivision, there are two ways to approach the problem (1). Append calibrated model of the existing distribution system with

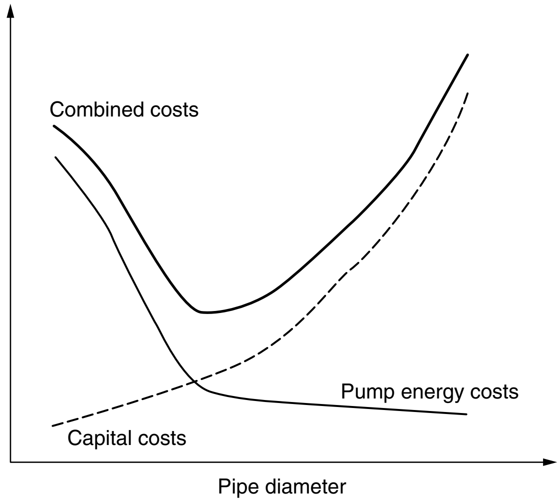

Figure A1. Typical capital and energy costs tradeoff curve.

new pipes and other network elements, and use the complete model in the iterative design process of sizing the new network elements (3). Create and use a network model pertaining to the proposed distribution system only, and use that model in the iterative design process. Although the first approach is more accurate, it requires a well-calibrated model of the existing distribution system along with a right set of demands and other boundary conditions. On the other hand, the use of the second approach requires head-flow characteristics at the location where the new piping will be tapped into the existing system. The head-flow characteristic curve at this location may be obtained by performing a hydrant flow test at or near the location where the new system will be tapped into the main distribution system. The hydrant flow test should record the static pressure as well as two other head-flow values to generate a head-flow characteristic curve. Figure A2 shows a typical head-flow characteristic curve generated from a hydrant flow test. This approach of using a head-flow characteristic curve in lieu of using the entire network model may not work well if a significant diurnal variation in static head is expected at the test location.

Transient Analysis. Rapid flow changes in a WDS can result in unexpectedly high or low pressures that can damage pipes or other components of the distribution system. Rapid flow changes are not uncommon in WDS. Routine pump startup and shutdown and pump trip resulting from power failure situations can result in rapid flow variations. High transient pressures might result in pipe breakages or loosening of pipe joints. Pressures below the vapor pressure of water lead to formation of vapor cavities. Subsequent high pressures can collapse the vapor cavities producing high-pressure spikes. Recent studies have shown that pressures below atmospheric levels can lead to pathogen intrusion problems at certain vulnerable points such as leaky pipe joints (18). Therefore, it would be prudent to perform a transient analysis of potential rapid flow variation scenarios and provide adequate protection to the WDS. The protection methods include use of pressure vessels (surge tanks), relief valves, rupture disks, and air release/vacuum valves (19).

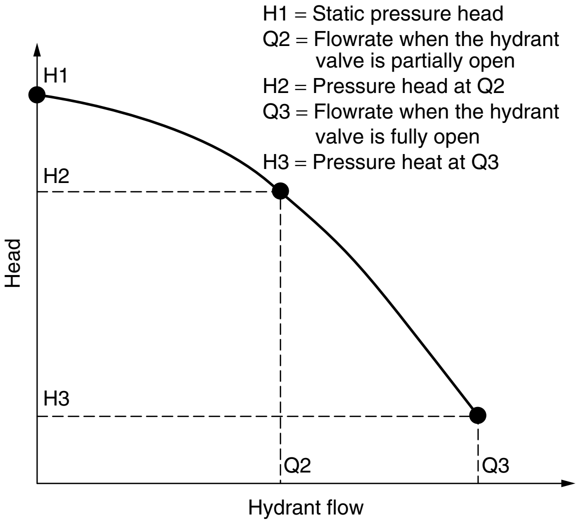

Figure A2. Example head-flow characteristic curve for a pressure supply.

Rural Water Systems

If rural water systems are not designed to provide fire flows, then special handling of long branch-lines serving few customers may be necessary. Requirements for residential water delivered through branched lines depend on the number of residential (domestic) connections served by each branch. These requirements should be based on probability considerations, and the requirement per connection served decreases as the number of connections increase. Relationships called Instantaneous peak demand curves (IPD) are available to calculate these requirements as a function of the number of connections served by each branch line. Additional conventional demands can be added at junctions. Branch line pipes should be sized to accommodate residential requirements calculated in this manner (20).

WHAT IS IN OUR DRINKING WATER?

Susan Richardson
U.S. Environmental Protection Agency

WHAT IS A DBP?

A drinking water disinfection by-product (DBP) is formed when the chemical used for disinfecting the drinking water reacts with natural organic matter in the source water. Popular disinfectants include chlorine, ozone, chlorine dioxide, and chloramine. Source waters include rivers, lakes, streams, groundwater, and sometimes seawater. DBPs have only been known since 1974, when chloroform was identified as a DBP resulting from the chlorination

---

Figure 1. Chloroform and trichloroacetic acid, two common DBPs, in chlorinated tap water.

of tap water. Since then, hundreds of DBPs have been identified in drinking water.

WHY IT'S IMPORTANT

Millions of people in the U.S. are exposed to these drinking water DBPs every day. While it is vitally important to disinfect drinking water, as thousands of people died from waterborne illnesses before we started disinfection practices in the early 1900s, it is also important to minimize the chemical DBPs formed. Several DBPs have been linked to cancer in laboratory animals, and as a result, the U.S. EPA has some of these DBPs regulated (Fig. 1). However, there are many more DBPs that have still not been identified and tested for toxicity or cancer effects. Currently, less than 50% of the total organic halide that is formed in chlorinated drinking water has been identified. Total organic halide is a measure of all of the organic compounds that contain a halogen atom (chlorine, bromine, iodine), and includes DBPs such as chloroform, bromoform, and trichloroacetic acid. There is much less known about DBPs from the newer alternative disinfectants, such as ozone, chlorine dioxide, and chloramine, which are gaining in popularity in the U.S. Are these alternative disinfectants safer than chlorine? Or do they produce more harmful by-products? And, what about the unidentified chlorine DBPs that people are exposed to through their drinking water—both from drinking and showering/bathing? The objective of our research is to find out what these DBPs are—to thoroughly characterize the chemicals formed in drinking water treatment—and to ultimately minimize any harmful ones that are formed.

OUR RESEARCH APPROACH

- Gas chromatography/mass spectrometry (GC/MS), liquid chromatography/mass spectrometry (LC/MS), and gas chromatography/infrared spectroscopy (GC/IR) techniques are used to identify the unknown by-products
- NIST and Wiley mass spectral databases are used first to identify any DBPs that happen to be present in these databases
- Because many DBPs are not in these databases, most of our work involves unconventional MS and IR techniques, as well as a great deal of scientific interpretation of the spectra
  - High resolution MS provides empirical formula information for the unknown chemical (e.g., how many carbons, hydrogens, oxygens, nitrogens, etc. are in the chemical's structure)
  - Chemical ionization MS provides molecular weight information when this is not provided in conventional electron ionization mass spectra
  - IR spectroscopy provides functional group information (e.g., whether the oxygens are due to a carboxylic acid group, a ketone, an alcohol, or an aldehyde)
  - LC/MS is used to identify compounds that cannot be extracted from water (the highly polar, hydrophilic ones). This is a major missing gap in our knowledge about DBPs—so far, most DBPs identified have been those that are easily extracted from water
  - Novel derivatization techniques are also applied to aid in the identification of highly polar DBPs
- Once DBPs are identified, ones that are predicted to have adverse health effects are studied in order to determine how they are formed (so that the treatment can be modified to ultimately minimize their presence in drinking water)

CURRENTLY

We have a major nationwide DBP occurrence study underway, where we are sampling drinking water across the U.S. (disinfected with the different disinfectants and with different water quality, including elevated levels of bromide in the source water).

SOME RESULTS

More than 180 previously unidentified DBPs have been identified for the first time.

Ozone is tending to produce oxygen-containing DBPs, with carboxylic acids, aldehydes, ketones, and di-carbonyl compounds the most prominent.

The presence of natural bromide in the source water is resulting in a tremendous shift from chlorine-containing DBPs to bromine-containing DBPs when chlorine or chloramine is used as a disinfectant (even in combination with ozone).

Chlorine dioxide is producing very few halogenated DBPs by itself, but when chlorine impurities are present, and natural levels of bromide or iodide are present in the source water, many bromo- and iodo-DBPs are formed.

New analytical methods have been developed (and are continuing to be developed) for the analysis of highly polar DBPs.

Collaborations have been forged with health effects researchers to study selected DBPs for potential adverse health effects.

BIBLIOGRAPHY

1. Richardson, S.D., Thruston, Jr., A.D., Caughran, T.V., Chen, P.H., Collette, T.W., Floyd, T.L., Schenck, K.M., Lykins, Jr., B.W., Sun, G.-R., Majetich, G. Identification of New Ozone Disinfection By-products in Drinking Water. *Environmental Science & Technology*, 1999, **33**, 3368–3377.
2. Richardson, S.D., Thruston, Jr., A.D., Caughran, T.V., Chen, P.H., Collette, T.W., Floyd, T.L., Schenck, K.M., Lykins, Jr., B.W., Sun, G.-R., Majetich, G. Identification of New Drinking Water Disinfection By-products Formed in the Presence of Bromide. *Environmental Science & Technology*, 1999, **33**, 3378–3383.
3. Richardson, S.D., Thruston, Jr., A.D., Caughran, T.V., Chen, P.H., Collette, Schenck, K.M., Lykins, Jr., B.W., Rav-Acha, C., Glezer, V. Identification of New Drinking Water Disinfection By-products from Ozone, Chlorine Dioxide, Chloramine, and Chlorine. *Water, Air, and Soil Pollution*, 2000, **123**, 95–102.
4. Richardson, S.D., Caughran, T.V., Poiger, T., Guo, Y., Crumley, F.G. Application of DNPH Derivatization with LC/MS to the Identification of Polar Carbonyl Disinfection By-products in Drinking Water. *Ozone: Science & Engineering*, 2000, **22**, 653–675.

THE ECONOMICS OF WATER RESOURCES ALLOCATION

PATRICK BOND
Kensington, South Africa

No one would doubt "the principle that water is a scarce good with dimensions of economic efficiency, social equity, and environmental sustainability." Yet as articulated in the World Bank's (1) *African Water Resources* document, the principle disguises two polar-opposite positions about pricing water.

On the one hand, *The Economist* magazine's (2; pp. 1–5) survey on water declared the central dilemma: "Throughout history, and especially over the past century, it has been ill-governed and, above all, collossally underpriced." Identifying this problem, naturally begets this solution: "The best way to deal with water is to price it more sensibly," for "although water is special, both its provision and its use will respond to market signals." In rural areas where there is competition among farmers for irrigation water, "The best solution is water trading." As for the problem of allocating and delivering water to the poor, "The best way of solving it is to treat water pretty much as a business like any other." We can call this the "neoliberal" position.

Six months later, on the other hand, an international movement with "progressive" politics, the People's World Water Forum (PWWF), emerged in Delhi, India, to promote the "decommodification" of water, based on feeder social movements from around the world in both cities and rural sites of struggle over water access. The PWWF (3; p. 1) founding statement declared that "Water is a human right; that corporations have no business profiting from peoples' need for water; and that governments are failing in their responsibilities to their citizens and nature."

Because there are many different ways that the pricing of water can be considered in bulk and retail settings, it makes sense first to outline the positions using a simple graph that distinguishes between neoliberal and progressive values. At the extreme where maximizing the production of economic goods is the prime objective, a leaked 1991 World Bank memo by then chief economist Lawrence Summers proposed an environmentally and socially insensitive allocation of resources: "I think the economic logic behind dumping a load of toxic waste in the lowest-wage country is impeccable and we should face up to that" (http://www.whirledbank.org). Rather than "internalize the externalities" of social impact or ecological damage, Summers' solution was simply to displace these to someplace where political power was negligible and the immediate environmental implications were less visible, in the name of overall economic growth.

And yet there are well-established critiques of gross domestic product (GDP), the most common measure of the aggregation of economic goods. Contrary to a more nuanced "economic logic," GDP treats the depletion of natural capital (including water resources) as income instead of depreciation of an asset; ignores the nonmarket economy of household and community, in turn devaluing child care, elder care, other home-based tasks and volunteer work (all of which rely on access to clean water); and considers natural disasters and pollution as economic gain because of associated service sector employment and repair/clean-up/replacement opportunities, not as a debit from social welfare or, with respect to drought and flooding, from ecosystem integrity.

Even as a caricature of neoclassical economics in the memo cited, the promotion of commodifed nature and society struck a chord during the neoliberal epoch of the late twentieth century. Throughout, those lobbying to treat water "much as any other business" grew in strength, and privatization and commercialization of water supplies expanded in many parts of the world. The intense conflict over the economics of water resources allocation was prefigured by the 1992 International Conference on Water and the Environment in Dublin, where water was formally declared an "economic good." Four years later, the formation of the Global Water Partnership and World Water Council advanced the position that commodification of water would lead to both private sector investments and more efficient use. In the same spirit, 1997 witnessed the first World Water Forum in Marrakesh, the founding of the World Commission for Water in the 21st Century, and an emblematic statement by the Swedish International Development Agency (4; pp. 11–13): "As the realisation increases that fresh water of satisfactory quality is a scarce and limited resource, matters related to management of the water resources have become more into focus .... At least four conditions need to be fulfilled to carry through efficient water allocation: (1) well defined user rights; (2) pricing at its marginal cost; (3) information related to availability, value, quality, delivery times; and (4) flexibility in allocation responding to technologic, economic and institutional changes." The United Nations Panel on Water declared in 1998 that "water should be paid for as a commodity rather than be treated as an essential staple to be provided free of cost" (5).

At the same time, the International Monetary Fund (IMF) and World Bank became much more explicit in promoting water commodification through what were once mainly macro-oriented structural adjustment programs, whether the Enhanced Structural Adjustment Facility, Poverty Reduction and Growth Facility, or Poverty Reduction Strategy Programme (6). According to one NGO critique by the Globalization Challenge Initiative (http://www.challengeglobalization.org), "A review of IMF loan policies in 40 random countries reveals that, during 2000, IMF loan agreements in 12 countries included conditions imposing water privatization or full cost recovery. In general, it is African countries, and the smallest, poorest and most debt-ridden countries that are being subjected to IMF conditions on water privatization and full cost recovery." When the World Bank (7, Annex 2) instructed its field staff on how to handle water pricing in even impoverished rural Africa, the mandate was explicit: "Work is still needed with political leaders in some national governments to move away from the concept of free water for all.... Promote increased capital cost recovery from users. An upfront cash contribution based on their willingness-to-pay is required from users to demonstrate demand and develop community capacity to administer funds and tariffs. Ensure 100% recovery of operation and maintenance costs." It was not long before a bureaucratic class emerged within Africa to work on cost recovery, for according to the 2001 Kampala Statement coauthored by the World Bank and the African Utility Partnership (8; p. 4), "The poor performance of a number of public utilities is rooted in a policy of repressed tariffs."

But a debate quickly followed over the merits of full cost recovery of operating and maintenance costs. In most urban systems, the cost of supplying an additional drop of water—the "short-run marginal cost curve" (Line A in Fig. 1)—tends to fall as users increase their consumption, because it is cheaper to provide the next unit to a large consumer than to a small consumer. Reasons for this include the large-volume consumers' economies of scale (i.e., bulk sales), their smaller per unit costs of maintenance, the lower administrative costs of billing one large-volume consumer instead of many small ones, and the ability of larger consumers to buy water at a time when it is not in demand—for example, during the middle of the night—and store it for use during peak demand periods. The premise here is that the pricing of water should correspond directly to the cost of the service all the way along the supply curve. Such a system might then include a profit markup across the board (Line B) that ensures the proper functioning of the market and an incentive for contracting or even full privatization by private suppliers.

The progressive principle of cross subsidization, in contrast, violates the logic of the market. By imposing a block tariff that rises for larger consumers (Line C), the state would consciously distort the relationship of cost to price and hence send economically "inefficient" pricing signals to consumers. In turn, argue neoliberal critics of progressive block tariffs, such distortions of market logic introduce a disincentive to supply low-volume users. For example, in advocating against South Africa's subsequent move toward a free lifeline and rising block tariff, the World Bank advised that water privatization contracts "would be much harder to establish" if poor consumers had the expectation of getting something for nothing. If consumers did not pay, the Bank suggested, South African authorities required a "credible threat of cutting service" (9; pp. 49–53).

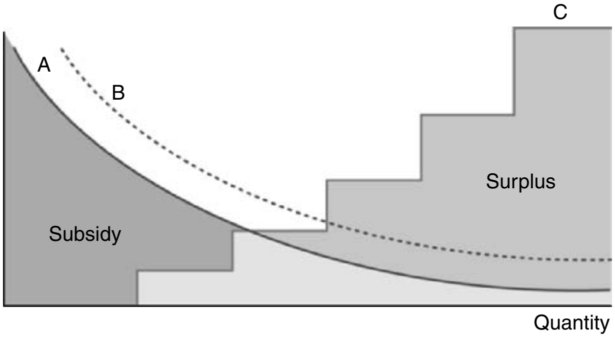

Figure 1. Three ways to price water: marginal cost (A), for-profit (B), and cross-subsidized lifeline plus rising block tariff (C).

The progressive rebuttal is that the difference between Lines A and C allows free universal lifeline services and a cross-subsidy from hedonistic users to low-volume users. There are also two additional benefits of providing free water services to some and extremely expensive services to those who have hedonistic consumption habits:

- Higher prices for high-volume consumption should encourage conservation, which would keep the longer run costs of supply down (i.e., by delaying the construction of new dams or supply-side enhancements).
- Benefits accrue to society from the "merit goods" and "public goods" associated with free provision of services, such as improved public health, gender equity, environmental protection, economic spinoffs, and the possibility of desegregating residential areas by class.

By way of definition, public goods can be observed and measured, for underlying their existence are two characteristics: "nonrival consumption" and "nonexclusion" from consumption. Nonrival consumption means that the consumption of a public good/service by one person need not diminish the quantity consumed by anyone else. A classic example is a national defense system, which is "consumed" by all citizens in a quantity that is not affected by the consumption of defense benefits by fellow citizens. Likewise, the benefits of a clean environment and hygienic public water system—reflecting a strong municipal water system and lifeline access to all—are enjoyed by all municipal consumers, regardless of how much water is consumed by a particular individual, although a minimum consumption level is required for all citizens to prevent the spread of infectious diseases. The principle of nonexclusion simply means that it is impossible to prevent other citizens from enjoying the benefits of public goods, regardless of whether they are paid for. This is important, as a state determines the detailed character of water pricing policy and distinguishes between necessities guaranteed by the state versus luxuries that people must pay for. A simpler way of putting it is that where the net benefits to society outweigh the costs of consumption of a good/service, the result is a "merit good." When the merit good benefits apply universally so that no one can be excluded from their positive effects, the result is a "public good." Markets usually underprovide for public goods, so the "Post-Washington Consensus" style of economics attempts to correct this market imperfection, even if it might mean introducing other market distortions, such as a free lifeline supply of water.

Another progressive critique of private suppliers who require tariffs reflecting marginal cost plus profit is that water infrastructure is a classic natural monopoly. The large investments in pipes, treatment centers, and sewage plants are "lumpy" insofar as they often require extensive financing and a long-term commitment, of which the state is more capable. To the argument that a progressive tariff could still coincide with a private sector supplier through a strong state regulator, progressives mistrust "captive regulatory" relations, given the long history of corruption in the water sector. Rebutting those who argue that Third World states are intrinsically incapable of providing water services, progressives cite more proximate reasons for the recent degeneration of state water sectors: 1980s–1990s structural adjustment programs that decapacitated most states, corrupt state bureaucrats, weak trade unions, and disempowered consumers/communities.

Finally, the progressive argument for making a water subsidy universal—not means-tested for only "indigent" people—is both practical and deeply political. If the service is means-tested, it invariably leads to state coercion and stigmatization of low-income people by bureaucrats. Furthermore, it is an administrative nightmare to sort out who qualifies because so many people depend on informal and erratic sources of income. More philosophically, though, it is a premise of most human rights discourse that socioeconomic rights such as water access are *universally* granted, not judged on the basis of a subjective income cutoff line, especially given the differences in household size for which different low-income people are responsible. This is partly because international experience shows that defense of a social welfare policy requires universality, so that the alliance of poor, working-class and middle-class people that usually win such concessions from the state can be kept intact (10).

As *The Economist* observed in mid-2003, one of the most important sites for considering the economics of water resources allocation is South Africa. One reason is that because of the international drive to commercialize water, even postapartheid South African citizens were subject to neoliberal cost recovery and disconnection regimes. This affected many who simply could not pay their bills. From the late 1990s through 2002, as a result, approximately 10 million people suffered water disconnections. Africa's worst-ever recorded cholera outbreak—affecting more than 150,000 people—can be traced to an August 2000 decision to cut water to people who were not paying a South African regional water board.

After the ruling, the African National Congress promised free basic water supplies in December 2000 during a municipal election campaign; the same bureaucrats responsible for water disconnections began redesigning the water tariffs. In July 2001, revised price schedules provided a very small free lifeline: 6000 liters per household per month, followed by a very steep, convex curve (see Fig. 2). But the next consumption block was unaffordable, leading to even higher rates of water disconnections in poor areas. The 6000 liters represent just two toilet flushes a day per person for a household of eight, for those lucky enough to have flush toilets. It left no additional water to drink, wash with, clean clothes, or for any other household purposes. In contrast, from the progressive point of view, an optimal strategy would provide a larger free lifeline tariff, ideally on a per-person, not a per-household basis, and then rise in a *concave* manner to penalize luxury consumption.

Johannesburg's tariff was set by the council with help from Suez Lyonnaise des Eaux, a Paris-based conglomerate, and began in July 2001 with a high price increase for the second block of consumption. Two years later, the price of that second block was raised 32%, with a 10% overall increase, putting an enormous burden on poor households that used more than 6000 liters each month. The rich got off with relatively small increases and a flat tariff after 40 kiloliters/household/month, which did nothing to encourage water conservation and hence did not mitigate the need for further construction of large dams, which in turn would drive up the long-run marginal cost curve and further penalize low-income Johannesburg townships residents.

This discussion of the economics of water resources allocation shows, simply, that pricing is political, and indeed the pursuit of "impeccable economic logic" in the water sector has generated some of the most intense struggles in the world today, calling into question the

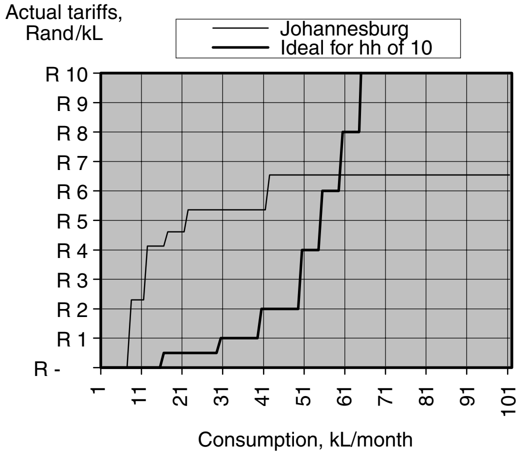

Figure 2. Divergent water pricing strategies—Johannesburg (2001) vs. ideal tariff for large household. *Source*: Johannesburg Water (thin) and own projection (thick).

very tenets of neoliberalism. The economics of privatized or commercialized urban water services have been challenged from Cochabamba, Bolivia—where the U.S. firm Bechtel tried to take ownership of rainwater collected by poor residents in the context of huge price increases—to Accra, Ghana, to most Argentine cities, to Manila and Jakarta, to Atlanta and Johannesburg, and to many other sites in between. Working out the contrasting discourses in political-economic analysis is crucial to any resolution of the problem in public policy.

BIBLIOGRAPHY

1. World Bank. (1996). *African Water Resources*, Technical Paper 331. Washington, DC.
2. *The Economist*. Survey of water. July 19, 2003.
3. People's World Water Forum. (2004). Statement. http://www.pwwf.org, January.
4. Swedish International Development Agency. (1997). *Environmental security and water management in Southern Africa*. Stockholm.
5. *New York Times*, Mar 22, 1998.
6. Hennig. 2001.
7. World Bank. (2000). *Sourcebook on Community Driven Development in the Africa Region*. Washington, DC, Annex 2.
8. World Bank and African Utilities Partnership. (2001). *The Kampala Statement*. Washington, DC, March 14.
9. Roome, J. (1995). *Water pricing and management*. World Bank Presentation to the *SA Water Conserv. Conf.*, unpublished paper, South Africa, October 2.
10. Esping-Andersen, G. (1990). *The Three Worlds of Welfare Capitalism*. Princeton University Press, Princeton.

ANSWERING THE CHALLENGE

EUGENE R. REAHL
Ionics, Inc.
Watertown, Massachusetts

The City of San Diego, California shares a problem in common with many other western cities. The problem is how to meet the water supply demands of a rapidly growing region when new sources are not readily available, at least not without spending huge sums of money.

In the late 1990s, San Diego (Fig. 1) took a major step to help solve this problem by equipping the brand new North City Reclamation Plant with an EDR (electrodialysis reversal) system, from Ionics, Incorporated. The system would desalinate tertiary treated wastewater and provide a new source of high-quality irrigation water, which would then reduce demands on the freshwater supply. Wastewater to be treated had salinity levels up to 1300-ppm TDS during the summer and early fall months.

With desalinated and blended reclaimed water having lower sodium levels (and less than 1000-ppm total TDS), San Diego used an existing 47-mile-long pipeline system to deliver high-quality irrigation water to golf courses, plant nurseries, parks, highway green belts, and homeowner associations. The strategy of selling this water as an attractive alternative to the use of hard-to-replace drinking water supplies caught on quickly.

In 1998, the EDR plant produced 2.2 mgd—with blending, 3.3 mgd was available. After two expansions, the latest of which will be completed in early 2005, EDR will produce 5.5 mgd of approximately 300-ppm TDS water. With blending, up to 12 million gallons per day of irrigation water can be supplied through a recently lengthened 67-mile-long pipeline. Future EDR expansion will increase this blended flow up to 15 mgd.

Originally, the city looked into desalination through the conventional use of microfiltration and reverse osmosis (MF-RO). This approach had already been installed at several locations in the Los Angeles area. Before making a final commitment, the city investigated to see if another alternative was possible (Fig. 2).

They found that electrodialysis reversal (EDR) had been successfully used at a number of wastewater locations to desalinate reclaimed irrigation water.

EDR offered the opportunity to use a less expensive single treatment membrane process, with a higher water recovery to make better use of available wastewater supplies. Overall, MF-RO water recovery would be about

Figure 1. San Diego uses EDR technology to produce low salinity irrigation water from reclaimed municipal wastewater.

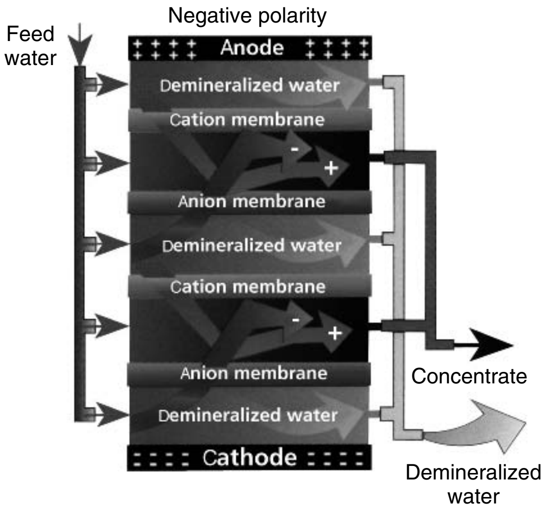

Figure 2. Electrodialysis reversal system.

80%. EDR would operate at 85% recovery. MF-RO required chemical pretreatment to control salt scale buildup on the RO membranes. EDR would operate with no chemical addition to the feedwater. One more major advantage to using EDR was that future capacity could be added by installing fully preassembled EDR systems within 53-foot-long enclosed trailers for up to 1.2 mgd flow per trailer (Fig. 3).

In 1997, city officials elected to use an EDR 2020 system for the North City irrigation water project after receiving bids from several general contractors. When all bids were opened, the least expensive MF-RO option was priced at US $4,775,975. The EDR cost was $3,569,000, or 25% less than MF-RO. This price was for the initial 2.2-mgd capacity EDR system.

With sequential use of MF and RO, water is forced through a membrane under pressure. MF membranes are used to pretreat RO feedwater, removing suspended solids materials, some organics, and other fouling-type materials that will affect the salt-removing RO membranes. With RO, water is pressure-driven through the membrane leaving unwanted salts behind. Chemicals (acid and antiscalants) are added to control scale buildup on the brine side of the membrane from concentrated wastewater constituents such as calcium phosphate, calcium carbonate, calcium sulphate, and others. One problem with today's high-performance RO membranes is that they cannot treat water containing chlorine, an essential ingredient in helping maintain bacterial control with wastewater. Chlorine will destroy RO membranes. It must be removed prior to the RO membranes.

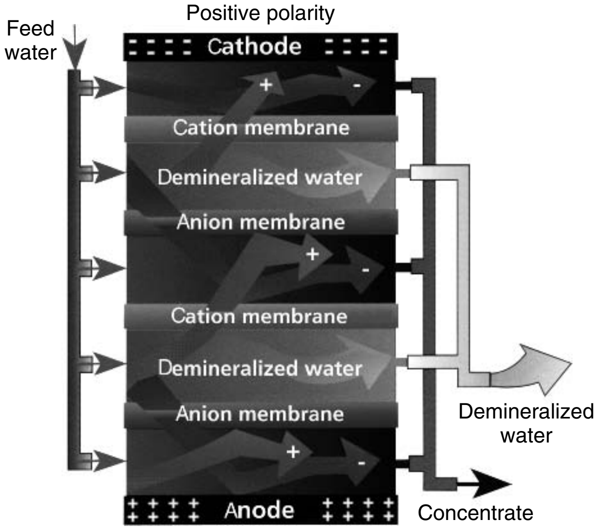

Figure 3. Electrodialysis reversal system.

ED (electrodialysis) was first commercially applied in 1954 for desalting brackish groundwater in a remote area of Saudi Arabia. ED is an electrically driven process in which a DC field generated between electrode plates pulls charged ions making up the water salinity through flat sheet cation and anion membranes. Water is not transported through the membranes—it continues on through the process. The ions (depending on their charge) are attracted to the anode or cathode electrode. Alternating cation and anion membranes generate a segregated product stream and a recirculating concentrated wastewater loop from which waste is eventually bled off. A single "ED stage" will remove up to 50–60% TDS. For still greater salt reduction, membrane and electrode plate stacks are staged in series one after another. Typically, up to 200,000 gpd can be treated in a single line of stacks. For higher flow applications, multiple lines of ED stacks are installed in parallel (Fig. 4).

By automatically switching the membrane stack DC field orientation every 15 to 30 minutes, the EDR system keeps ED membranes free of salt scale buildup without having to add chemicals in the feedwater. Even at 85% water recovery, with highly concentrated levels of salts in the EDR brine, the EDR systems in San Diego operate without chemical feed to control salt scale buildup. By choice, one chemical that is added is a 0.5 ppm of free residual chlorine to keep EDR membrane stacks free from unwanted bacterial growth. Compared with MF-RO, the only pretreatment used with EDR are 10-micron disposable cartridge filters to remove solid particulate matter prior to the membrane stacks.

PAST RESULTS AND TODAY'S LATEST DEVELOPMENTS

When the initial system started operation in 1998, it was supplied as two parallel, land-based 1.1-million gallon per day units located under a sunshade. The systems had a single-stage membrane stack design. Approximately 55% of the salt was removed in this process, and then ED product water was blended up to 3.3 mgd at less than 1000-ppm TDS.

The demand for EDR product flow increased as the city continued to sell more reclaimed irrigation water. In 1999, the EDR system was expanded with the addition of a third land-based unit (a two-stage, 1.1-mgd EDR unit), and with the retrofitting of second-stage membrane stacks to the original EDR equipment. With EDR product flow of 3.3 mgd now having still better quality (less than

Figure 4. San Diego took a major step in the late 1990s by equipping a new high-technology wastewater reclamation plant with EDR (electrodialysis reversal) system from Ionics, Incorporated, to desalinate tertiary treated wastewater.

300 ppm), more non-desalted water is blended with EDR product. The plant is now rated to deliver over 7 mgd of <1000-ppm irrigation water.

To continue development of this alternative irrigation water supply, city officials purchased still more EDR capacity in the fall of 2003. Two options were available. The first used the conventional approach of adding more "land-based" EDR systems. This option required a major rework of the site, with more conventional building of sunshades, concrete pads, and extensive site installation work. The second approach called for the installation of fully preassembled EDR systems within 53-ft-long over-the-road enclosed trailers. On arrival at site, a simpler retrofit using preassembled trailers allows for the quick installation of two additional 1.1-mgd EDR systems (Fig. 5).

Total EDR production will be 5.5 mgd when the latest EDR units go online in January 2005. With blending, up to 12 mgd of irrigation water will be supplied to a now further expanded irrigation water pipeline infrastructure system. This water will go to additional golf courses and new home developments. The latest expansion also includes adding enough concrete pad area and water treatment site pipeline capacity to bring in a third EDR trailer in the future. EDR's ultimate plant will then have 6.6 mgd of EDR capacity and up to 15 mgd + total blended flow. Results at North City have proven very positive for San Diego and for the EDR process

FUTURE OUTLOOK

Based on San Diego's success, other municipal agencies in California are looking at EDR to treat their municipal wastewaters for irrigation reuse. In a recent tertiary-treated wastewater pilot test performed in California, EDR and MF-RO were run side-by-side for 9 months. Based on a present worth analysis of capital cost, installation cost, and 20 years O&M, EDR was rated as 40% less expensive than MF-RO.

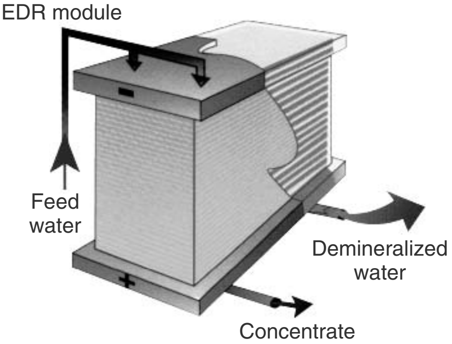

Figure 5. Preassembled EDR module.

In another case, where EDR or MF-RO could be used, highly concentrated brines have to be trucked 15 miles to a sewer line disposal site. Transportation costs are very high. The ability of EDR to operate at over 90% water recovery makes treatment practical. The greater volume of MF-RO treatment does not pencil out for this project.

AUTHOR'S NOTE

Eugene R. Reahl is the Western Regional Sales Manager for municipal systems sales at Ionics, Incorporated, based in Watertown, Massachusetts. Reahl sold the second and third EDR system expansions at North City and has been involved with many municipal EDR and RO systems since 1981. Contact the author at ereahl@ionics.com.

KEY CAUSES OF DRINKING WATER QUALITY FAILURE IN A RURAL SMALL WATER SUPPLY OF SOUTH AFRICA

M.N.B. Momba
N. Makala
B. Zani
University of Fort Hare
Alice, South Africa

B.M. Brouckaert
University of Natal
Durban, South Africa

INTRODUCTION

Of the advances made possible through science and technology, the treatment and distribution of safe water use is truly one of the greatest. Abundant, clean water is essential for good public health. It has been increasingly recognized by world leaders that safe drinking water is a critical building block of sustainable development. Disinfection, a chemical process whose objective is to control disease-causing micro-organisms by killing or inactivating them, is critically important in drinking water treatment. However, research has shown that about 1.2 billion people lack access to safe drinking water worldwide (1). As a result, the World Health Organization estimates that 3.4 million people, mostly children, die every year from water-related diseases (2).

The provision of adequate water services was one of the more difficult and pressing challenges inherited by the new South Africa. Prior to the change of government in 1994, an estimated 30–40% of South Africa's population (14–18 million people) were without adequate water supply services and some 21 million people were without adequate sanitation (3). As of 2004, some 10 million additional people have been supplied with drinking water, thereby reducing the backlog in 2003 to some 4 million (4). Provision of access to a basic level of water can be regarded as the first step up the water ladder, and progress achieved to date is indeed impressive. Although great strides have been made in the effort to provide safe and clean water to all South Africans, studies have shown that in small rural towns and small remote villages with adequate water supply services, the drinking water quality is generally poor and often not fit for human consumption at the point of use (5–7). A study aimed at determining the upgrading needs of small and rural surface water treatment plants in South Africa and the subsequent process of drawing up guidelines on how to address these needs also revealed that most small water treatment plants in the country do not produce safe, clean potable water (8).

The study presented here was undertaken in the town of Alice, a small rural town in the Eastern Cape, one of South Africa's poorest provinces. The Alice Water Treatment Plant (AWTP) is one of many small water treatment plants in the Eastern Cape Province that has difficulty providing safe drinking water. Previous studies (6,7,9) on the microbiological quality of Alice drinking water indicated a high number of fecal and total coliforms and heterotrophic bacteria far exceeding the limit allowed for potable water (10). *Aeromonas hydrophila, Escherichia coli, Serratia odorifera, Salmonella arizonae, Pseudomonas aeroginosa, Pseudomonas fluorescens, Serratia liquefaciens*, and *Serratia marcescens* were present in all water samples collected from various sites of the distribution system. The possibility of *Vibrio* species existed in one of these sites (7). This situation, therefore, places the Alice community at risk of disease outbreaks. A crucial need exists for scientifically sound answers to the problems of the inadequate treatment of drinking water in small water treatment plants in the country in general and the Alice Water Treatment Plant (AWTP) in particular.

The present study aimed at assessing the causes of the drinking water quality failure within the Alice Water Treatment Plant from the raw water intake to the finished water in the in-plant reservoir and at identifying the key causes affecting the performance of the plant including the effectiveness of the disinfection process for the removal of indicator bacteria.

METHODOLOGY

Plant Description

The Alice plant is a conventional water treatment plant, which includes coagulation, flocculation, sedimentation, rapid sand filtration, and chlorination. Water drawn from the Binfield Dam on the Tyume River (approximately 12 km from Alice) is fed under gravity to the head of the works. The design capacity of the plant is 7 mL/d; however, it is currently operating between 3 and 4 mL/d. A number of surrounding villages currently without water services are to be connected to the Alice system in the near future.

At the beginning of the investigation (June 2002) until November 2002, lime and alum were used for the pretreatment of the raw water. The coagulants are dosed at the same point, just ahead of a 90° V-notch weir and hydraulic jump. Powdered hydrated lime and granular alum were mixed with domestic water in separate dissolving/slurry tanks and then fed under gravity to the dosing point. Flash mixing of the chemicals with the raw water results from the turbulence generated over the hydraulic jump. Granular alum and hydrated lime were replaced by ultrafloc (a polyaluminium chloride and polyamide mix supplied by Sudchemie) at the beginning of December 2002. At the foot of the jump, the water enters a relatively short hydraulic flocculator before flowing under gravity to a three-way flow splitter via an underground pipe. From the splitter, the water flows through short intermediate channels into each of the three-clarifier inlet channels. From the influent channels, the water flows into the clarifier inlet chambers via pipes passing through the clarifier walls (nine per clarifier). The discharge into the inlet chambers induces gentle hydraulic mixing, which promotes the formation of large settleable floc. The water passes over a vertical baffle wall into the main body of each of the three rectangular horizontal clarifiers. Settled water overflows into a common channel, which feeds into the filter influent pipe. The filter influent is split among three 4.3-m-diameter self-backwashing valveless sand filters. Chlorine gas is added to the combined filtrate just before the onsite reservoir from where it is fed into the bulk distribution system.

Identification of the Problems Affecting the Performance of the AWTP

The following activities were undertaken from June 2002 to January 2003, with the exception of the experimental study on the dissolved organic carbon, biodegradable dissolved organic carbon, the bacterial regrowth, and the growth factor, which was performed from August to October 2003. The chlorine residual was determined during both periods of the study.

- A tour of the plant and a visual inspection of the various unit processes.
- Interviewing of the operators and plant superintendent about operating practices and the monitoring program.
- Measurement of the raw water flow rate using the V-notch weir.
- Determination of the optimum doses of alum and lime using Jar tests.
- Measurement and interpretation of the turbidity at each step in the treatment process.
- Opening and inspecting one of the valveless filters.
- Measurement of the initial chlorine dose and free chlorine residual concentration in the onsite reservoirs.
- Monitoring of the bacterial quality of the finished water.
- Investigation of the removal of dissolved organic carbon and assessment of the impact of the biodegradable dissolved organic carbon on the bacterial regrowth in the finished water. Details of these activities are given below.

Flow Rate Measurement and Control. The plant was equipped with a flow meter; however, this was located in a sump, which was flooded and the meter had not been read for several years. The flow could also be measured at the V-notch weir at the hydraulic jump; however, the plant operating personnel were unaware of this prior to the project team's visit. During the investigation, the flow of the raw water was calculated using the existing 90° V-notch weir according to Kawamura (11).

For a 90° V-notch weir, the total flow rate, Q, is related to the height of the crest over the weir, H, as follows:

$$Q(\mathrm{m}^3/\mathrm{s}) = 1.40\ \mathrm{H}^{2.5}$$

where H (m) is the height of water above the weir crest.

Table 1 shows H values with corresponding flows as calculated using the hydraulic equation given above.

Alum and Lime Consumption and Cost Estimation. The consumption of alum and lime onsite was estimated by collecting and weighing the amount of dry chemical fed to the mixing tanks during a given time interval. Using the flow rate measured on each day, the corresponding dosages in mg/L and the cost were estimated, which were compared with the optimum doses determined from standard jar tests (11). The costs of chemicals corresponding to the various dose estimates were then calculated taking into account the cost of a 50-kg bag of alum, which is R 96 (US$ 14.9), and a 25-kg bag of lime, which is R 30 (US$ 467).

Clarifier and Filter Performance. The turbidity of the clarified and filtered waters was measured for each settling tank and filter using a Microprocessor Turbidity Meter (HACH Co., Model 2100P). The filtrate was allowed to run out of a sampling tap on the filter under the drain until the measured turbidity stabilized. One filter was opened directly after it had backwashed to assess the state of the filter media. The impact of backwashing on filtrate turbidity was investigated by measuring turbidity before, immediately after, and 4 h and 6 h after backwashing.

Disinfection. The chlorine residual and the turbidity of the treated water were measured in the in-plant reservoir using a multiparameter ion-specific meter (Hanna-BDH laboratory supplies) and Microprocessor Turbidity Meter (HACH company, Model 2100P), respectively.

Monitoring of Bacterial Quality of the Finished Water. To assess the performance of the AWTP for the removal of bacterial indicators, the intake raw water and the chlorinated water samples from the in-plant reservoir were collected twice a week. Treated waters for microbiological analysis were collected in sterile 1-L glass bottles, which contained sodium thiosulphate (ca 17.5 mg/L), and placed in an ice bag. Microbiological analyses were conducted within 1 h after the collection of the water samples. Total coliforms (TC), fecal coliforms (FC), presumptive Salmonella (PS) spp., and heterotrophic plate count bacteria (HPC) were used as bacterial indicators. The membrane filtration technique and chromocult coliform agar (Merck) were used for the enumeration of coliforms and PS. Standard spread plate procedure and R2A agar (Oxoid) were used for the

Table 1. Examples of Flow Rate Measurement Considering Height (*H*) of Water

H (cm)	20	21	22	23	24	25	26	27	28	29	30
Q (m3/s) ⇒ 1.4$H^{2.5}$	0.025	0.028	0.032	0.034	0.040	0.044	0.048	0.053	0.058	0.063	0.069
Q (Ml/d) ⇒ 121$H^{2.5}$	2.16	2.45	2.75	3.07	3.41	3.78	4.17	4.58	5.02	5.48	5.96

enumeration of HPC bacteria. Water samples were analyzed for the above micro-organisms using internationally accepted techniques and principles (12). Bacterial colonies differing in size, shape, and color were randomly selected from the different plates and further isolated on MacConkey agar (Biolab) by the streak plate technique and incubated at 37 °C for 24 h. These were further purified by the same method at least three times using nutrient agar (Biolab) before Gram staining was done. Oxidase tests were then done on those colonies that were gram negative. The API 20E kit was used for the oxidase-negative colonies, and the strips were incubated at 35 ± 2 °C for 24 h. The strips were then read, and the final identification was done using API LAB PLUS computer software (BioMérieux).

Dissolved Organic Carbon (DOC), BDOC, and Bacterial Regrowth Analyses. Water samples (raw water, water after filtration, and water after chlorination) for the dissolved organic carbon (DOC), biodegradable dissolved organic carbon, bacterial regrowth, and growth factor were collected weekly using sterile glass bottles (brown sterile bottles for raw water and water after filtration). Prior to use, all glassware (except the filter frit) was pretreated according to Servais et al.'s (13) methods, which were modified by Grundlingh et al. (14) to make sure that there were no traces of organic matter left in the glassware, as this might lead to false results. Water samples for DOC analysis were prepared according to Mathieu et al. (15). The DOC concentrations were determined using an AQUADOC™ TOC Analyzer.

The preparation of the inoculums was performed according to Grundlingh et al. (14). A 50-mL raw water sample was filtered through a 2.0-μm nucleopore polycarbonate membrane, 47-mm diameter (Merck, Cat. No. 111111), and was stored in a sealed glass bottle. This water sample was used for the analysis of BDOC and for the growth factor. The BDOC was determined according to the method previously used by Servais et al. (13) modified by Grundlingh et al. (14). Briefly, a water sample was vacuum-filtered through a 0.2-μm nucleopore polycarbonate membrane, 47-mm diameter (Merck, Cat. No. 111106), and 200 mL of this water sample was measured using a glass measuring cylinder into a brown bottle. Two milliliters of the innoculum was added. The water sample was then incubated for 28 days in a dark cupboard to prevent the breakdown of DOC by sunlight. After every 7 days, 20 mL was removed and prepared for the DOC analysis. The concentration of BDOC was calculated as the difference in the DOC concentration before incubation and on the last day of incubation. Bacterial regrowth was then determined using total coliforms (TC), fecal coliforms (FC), and heterotrophic plate count (HPC) as main parameters, which were enumerated as described for the monitoring of the bacterial quality of the finished water using standard methods (12).

Growth Factor. The growth factor was analyzed according to Page et al. (16) to evaluate the potential of the water to support the regrowth of bacteria. Each water sample was allowed to warm to room temperature after collection and was then vacuum-filtered through a 2.0-μm nucleopore polycarbonate membrane, 47-mm diameter (Merck, Cat. No. 111111), and 200 mL of this water sample were measured into a brown bottle, 20 mL of sterile nutrient broth were added to the sample. This water sample was then vacuum-filtered using a 0.2-μm nucleopore polycarbonate membrane, 47-mm diameter (Merck, Cat. No. 111106), and then poured into three sterile turbidity cuvettes. A Microprocessor Turbidity Meter (HACH Co., model 2100P) was zeroed using this sample. The inoculum (bacteria from raw water) was added until the turbidity readings were between 0.2 and 0.4 NTU. The turbidity meter was zeroed again to ensure that only the growth of new bacteria was enumerated. Turbidity readings were taken after every 30 minutes until a plateau was reached. The growth factor was calculated as follows:

$$\frac{\text{Turbidity at the plateau}}{\text{Turbidity at the lag phase}}$$

Results and Discussion

To ensure effective water treatment processes, which will result in the production of safe and clean potable water, trained and skilled personnel remain a prerequisite. However, this is not the case in most small water supplies in rural developing communities. Lack of trained operators at the AWTP resulted in the malfunctioning of the plant.

Coagulant Dosing and Control. Two major operating problems identified during this investigation were the lack of flow rate measurement and chemical dosing control. Raw water flow rates were adjusted based on the level of the in-plant reservoir, but coagulant chemical feed rates were not automatically adjusted in proportion to the raw water flow rate, resulting in large fluctuations in coagulant doses. The operating staff apparently sometimes adjusted the alum and lime doses based on the taste and appearance of the final water and the formation of floc in the clarifiers; however, the coagulant demand was not determined. Table 2 summarizes the raw water flow rates, coagulant feed rates and doses for 4 different days, and an estimate of the annual cost of alum and lime used. Jar tests on the raw water indicated that the optimum alum dose was approximately 15 mg/L, three to five times smaller than the doses used by the plant operators during the study period. With an average flow of 3.27 mL/d obtained during the investigation, the cost of alum per year at a dosage rate of 15 mg/L was calculated at R 33,323 (US$ 5182.43), which represents one quarter of the cost of the dose currently used by the plant. The jar tests conducted also indicated that lime addition was not necessary for coagulation at the time the tests were undertaken and was probably detrimental to the process. Theoretically, lime addition should only be necessary if insufficient alkalinity exists in the raw water to prevent the pH dropping below 6 or the colloidal particles becoming restabilized. If the coagulated water pH is above 6, then the addition of lime increases the solubility of aluminium and the amount of alum required. Lime may have been necessary at other times of the year. Overdosing lime and alum appeared to have a negative impact on plant performance. There was a large amount of floc deposited throughout the plant, including in the in-plant reservoir. Post-precipitation of floc after the filters because of alum overdosing was evident on at least one occasion.

Clarification. The general performance goal of a settling tank preceded by chemical coagulation is that its effluent (settled water) must have turbidity of less than 5 NTU (17). This protects subsequent sand filtration processes against high loads of suspended solids, preventing turbidity breakthrough and/or short filter runs. However, the clarified water turbidities measured in the AWTP during the study period were found to range between 8.5 and 17.9 NTU. During the plant inspections, it was noted that the flow splitter was not being operated correctly: Gate valves downstream of the overflow weirs had been set so that one of the three weirs was submerged. An uneven flow split may have contributed to poor clarifier performance; however, the major problem appeared to be overdosing with lime and alum. Clarifiers were usually manually desludged twice a week. No written procedures or records for desludging existed.

Filtration. For efficient disinfection, WHO guidelines (17) for potable water production state that the filtrate must have turbidity of <1 NTU because high turbidities exert high disinfectant demands, and particles may shield micro-organisms from disinfectants. However, the average filtered water turbidity recorded from October 2002 to January 2003 was greater than 1 NTU (Fig. 1). When alum and lime were used, poor filter performance was sometimes linked to post-precipitation of alum floc. However, switching to Ultra-floc did not have a positive effect on filtrate quality. The small improvement in filtrate turbidity in December and January is not necessarily attributable to the change in coagulant because the raw water turbidity decreased at the same time. Filtration rates of ~ 2–4 m/h were recorded during the evaluation of the AWTP, which are conservative rates for rapid sand filtration and should not have unduly challenged the filter performance if pretreatment had been adequate. Apart from inadequate pretreatment, the most common source of problems in rapid filtration is inadequate backwashing.

Table 2. Estimating Alum and Lime Consumption and Costs According to Onsite Measurements

	Alum		Lime	
Flow (mL/d)	Consumption (Kg/d)	Dosage (mg/L)	Consumption (Kg/d)	Dosage (mg/L)
2.75	172	63	42	15
2.75	224	82	27	10
3.41	184	54	42	12
4.17	192	46	45	11
Average Consumption (kg/d)	**193**	**61**	**39**	**12**
Average cost per year for alum and lime				
	R 135 255.00		**17 082.00**	

Valveless filters are designed to backwash automatically when a specific headloss across the filter media is developed. Their chief advantage and the reason they are a popular choice for small treatment plants is that they backwash without any operator intervention. Important disadvantages include the following (18): (a) no auxiliary backwash is used, so cleaning is relatively inefficient; and (b) slow headloss development may result in excessively long run times, which could, in turn, lead to mudballing problems. The recommended maximum run length is 24 hours. No records of filter backwash were kept at the AWTP, but the filters appeared to go for days without backwashing. This may have been attributable to low filtration rates and poor coagulation.

Inspection of a filter bed directly after manual backwash revealed that the filter media was in a fairly good state, except at the edges of the filter, where it had built up against the filter walls and where a number of mudballs were found, which indicates poor flow distribution adjacent to the walls. The media away from the walls appeared to be clean except for a few millimeters of sludge on the top surface of the bed (typical of filters backwashed without auxiliary backwash). An analysis of the filter effluent turbidity immediately after a manual backwashing revealed very high values (27.5–39 NTU); thereafter a gradual decrease in the turbidity was noted within 4 h (2.33–3.44 NTU) and 6 h (1.98–2.02 NTU). These high turbidities indicated that the volume of backwash was insufficient to flush out all of the floc detached during backwashing.

Disinfection. The chlorine dose rate was manually controlled and set using a rotameter housed in the chlorination room. Like the coagulant doses, the chlorine dose was not ratioed to the plant flow rate. Furthermore, no attempt was made to match the dose to the chlorine demand. During the preliminary plant visits, the dosing rate was set at 100 g/h Cl_2, which translated into applied chlorine doses of 0.6 to 1.3 mg/L at plant flow rates of 2.75–3.97 mL/d. Given the poor quality of the filtrate, these doses were insufficient to establish an adequate residual. It was also noted that a substantial amount of floc had settled in the contact chambers, which could also lead to high consumption of the disinfectant. The dosing rate was subsequently increased to 400 g/h without any improvement in dosing control. Water samples collected at the point of treatment (in-plant reservoir) showed that the free chlorine residual values ranged from 0 to 1.14 mg/L for the period of 5 months. Figure 2 shows the average

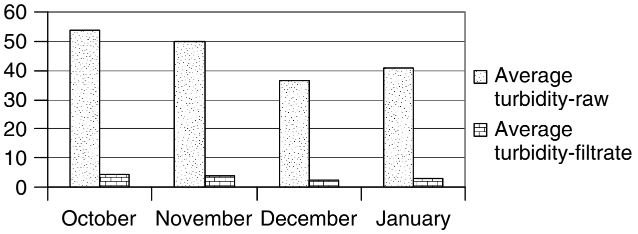

Figure 1. Performance of AWTP for the removal of the turbidity (NTU).

free chlorine residual for each month. It can also be seen that chlorine residual increased as the filtrate turbidity decreased. Alice drinking water at the point of treatment contained less chlorine than the target chlorine residual concentration, i.e., 1 mg/L. The present results confirm those of previous investigators who also observed a low level of chlorine residual in the AWTP's reservoirs (9).

Microbiological quality of drinking water at the point of treatment. The bacterial quality of the finished water as suggested by the indicator organisms used (Figs. 3 and 4) did not meet the standards for potable water. Among 14 species of potential pathogenic bacteria identified from the raw water (*Aeromonas hydrophila, Aeromonas caviae, Aeromonas sobria, Escherichia coli, Serratia odorifera, Serratia liquefaciens, Serratia marcescens, Salmonella arizonae, Pseudomonas aeruginosa, Pseudomonas fluorescens, Vibrio fluvialis, Chryseomonas luteola, Enterobacter aerogenus, Enterobacter sakazakii),* 8 species were still prevalent in the chlorinated water from the reservoirs (*Aeromonas hydrophila, Escherichia coli, Serratia odorifera, Serratia liquefaciens, Serratia marcescens, Salmonella arizonae, Pseudomonas aeruginosa, Pseudomonas fluorescens*). In general, the result suggested that the performance of the AWTP for the removal of indicator bacteria (HPC, TC, FC, and PS) was not adequate. The finished water produced by this plant places the consumers at a serious health risk. These results confirm our recent study, which also showed the failure of the AWTP to provide potable water to its consumers at the point of use (7).

The removal of DOC and the impact of BDOC on bacterial regrowth. Sand filtration processes are performed to decrease the DOC concentration from the raw water to the filtered water. Overall, DOC removal at the AWTP appeared to have been very poor for the entire monitoring period, as indicated in Table 3. The AWTP's filtration system did not decrease the DOC concentration to the level recommended by the water industries, which ranges between 0 and 5 mg/L (10). The DOC concentrations in the filtered water samples were above these limits (Table 3). The highest DOC concentration in raw water was found in October and the removal was at 59.57%, whereas the lowest concentration of the DOC in raw water was recorded in August. No removal of the DOC was noted in August and in September; in contrast, an increase in DOC concentration occurred at rates of 13.64% and 35.34%, respectively. The inefficiency of the filtration process could be attributed to the lack of backwashing of the sand filters in the AWTP. It was noted that the filters went for days without backwashing, which impacted the quality of the filtrate in terms of the DOC removal.

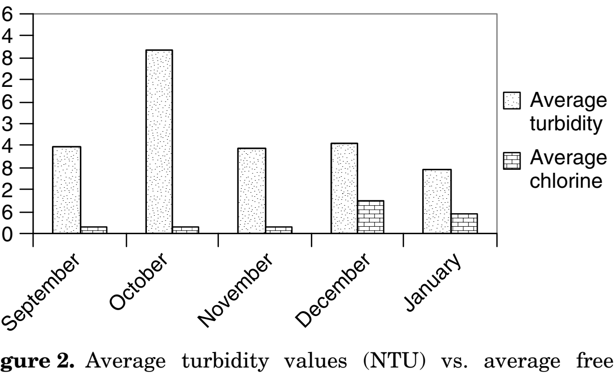

Figure 2. Average turbidity values (NTU) vs. average free chlorine residual concentration (mg/L) in the in-plant reservoir.

The concentration of the DOC in the filtered water accounted for the presence of that in the finished water

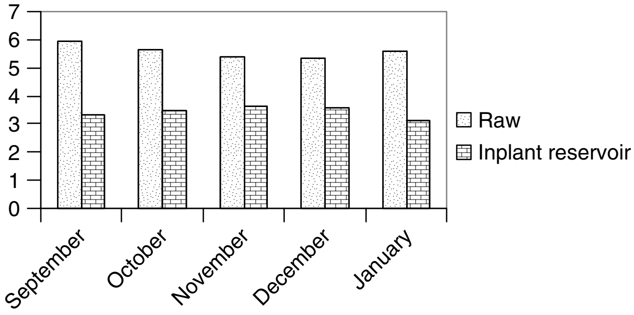

Figure 3. Average log HPC bacterial counts (cfu/mL) in raw water and the in-plant reservoir.

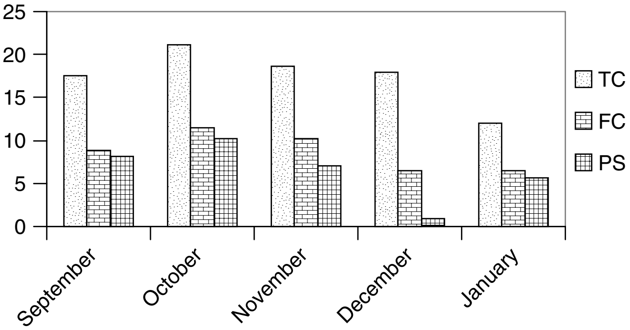

Figure 4. Average counts (cfu/100 mL) for total coliforms (TC), fecal coliforms (FC), and presumptive *Salmonella* (PS) in the in-plant reservoir.

Table 3. The Performance of AWTP in Terms of the DOC (mg/L) and BDOC (mg/L) Concentrations, the Growth Factor, and the Chlorine Residual Concentration

		DOC		BDOC	GF	CR
MONTH		FW	CW	CW	CW	CW
	MIN	13.00	15.00	1.00	2.90	0.00
	MAX	12.00	16.00	6.70	3.04	0.00
AUGUST	$\bar{x}$	12.50	15.50	3.85	2.97	0.00
	Sx	0.71	0.71	4.03	2.05	0.00
	MIN	13.00	13.00	4.20	3.29	0.00
	MAX	18.00	33.00	14.80	6.54	2.77
SEPTEMBER	$\bar{x}$	15.70	23.00	14.00	5.76	1.19
	Sx	2.52	10.00	9.43	0.74	1.42
	MIN	9.00	27.00	4.00	4.41	2.52
	MAX	133.00	50.00	40.00	7.12	3.58
OCTOBER	$\bar{x}$	37.20	31.00	18.80	6.86	2.82
	Sx	57.30	15.79	16.45	1.34	0.44

DOC = dissolved organic carbon, BDOC = biodegradable dissolved organic carbon, GF = growth factor, MIN = minimum value, MAX = maximum value, Sx = standard deviation, $\bar{x}$ = average value, CR = chlorine residual.

samples at the point of chlorination. Although no free chlorine residual was noted in August and during the first week of September (Table 3), the DOC concentrations increased by 24% and by 46.5%, respectively, during the monitoring period. In October, the free chlorine residual concentrations exceeded the limits recommended for potable water, which are in ranges of 0.3–0.6 mg/L as ideal free chlorine residual concentration and 0.6–0.8 mg/L as good free chlorine residual concentration for insignificant risk of health effects (19). During this month, the DOC concentration decreased by 16.67%. As indicated in Table 3, high concentrations of the DOC in the Alice finished water contributed to the biodegradability of the organic matter in the drinking water, which followed the trend of chlorine residual as it increased each month.

Biologically stable water is free of biologically available substances. In addition, it is well known that the degree of bacterial regrowth is determined by the overall content of biodegradable organic carbon present in the water and is largely determined by the origin of the drinking water produced (20). The results of the present experimental study showed close relationship between the DOC concentrations and both BDOC concentrations and the growth factors (Table 3). The concentration of BDOC and the growth factor revealed that water distributed by the AWTP was biologically unstable. Water with growth factors below 5 is considered biologically stable (16). In the AWTP, the average values of growth factors were found to be high in September (7.72), and October (5.97), exceeding the above limit. The lowest values of the growth factor were found in August (4.37).

Removal of BDOC to the level that limits microbial regrowth provides not only a direct control of bacterial population but also an indirect control of protozoan population through a trophic food web (21). A high BDOC concentration in finished water indicates poor quality and a potential for micro-organisms to multiply in the water (22). It has been reported that treated water containing a BDOC concentration of less than 1 mg/L is not prone to regrowth (23). However, treated water supplied by the AWTP had very high concentrations of BDOC, up to 40 mg/L (Table 3), which resulted in high bacterial regrowth during the study period (Fig. 5). The increase of the bacterial counts was parallel to the increase in the BDOC concentration from August to September, but bacterial regrowth decreased in October contrary to the increase in the BDOC concentration. This decrease in bacterial regrowth could be related to a high concentration of free chlorine residual, which was up to 3.58 mg/L (Table 3). The variation of the free chlorine residual concentration was associated with the fact that the operators in the AWTP have little understanding of the chlorination process in the purification system. Chlorination in the AWTP was done without considering the chlorine demand of the water.

CONCLUSION AND RECOMMENDATIONS

The Alice Water Treatment Plant should be operated to consistently produce a combined filtrate turbidity (measured at the first clear well) of less than 0.5 NTU and a free chlorine residual of at least 0.5 mg/L in the main storage reservoirs for the town and 0.2 mg/L at the end of the distribution system. If the water is to be stored for considerable periods of time at the point of use, then higher residuals may be required.

Coagulation should ideally also be optimized for maximum DOC removal prior to disinfection. The operating staff should aim to meet turbidity removal goals with the minimum amount of coagulants necessary. Jar tests should be used to determine what these optimum doses are, which will result in a reduction of sludge production and disposal problems, lower solids loading, and less risk of post-precipitation in the filters, as well as cost savings for the plant. Filters should backwash at least once a day. If necessary, manual backwash should be initiated to prevent filter run times of greater than 24 to 36 h.

The success of any attempt to optimize the plant operation requires the full cooperation of the operating staff, especially because they will be required to undertake additional duties. All decisions and strategies should be discussed with the operators, and they should be encouraged to take a personal interest and pride in any improvements achieved.

Acknowledgments
The authors wish to express their thanks to Water Research Commission, South Africa who sponsored this investigation.

BIBLIOGRAPHY

1. Chlorine Chemistry Council. (2003). Drinking water chlorination, a review of disinfection practices and issues. Available: http://c3.org/chlorine_issues/disinfection/c3white2003.html.
2. World Health Organization. (2002). *Water and Sanitation: Facts and Figures*. Available: http://www.who.int/water_sanitation_health/General/facts&fig.htm.
3. Van der Merwe, L.H. (2003). *Water Sewage Effluent*. **23**(5): 50–51.
4. Kasrils, R. (2003). *Media Release. Statement by Minister Ronnie Kasrils on the Celebration of the 9 Millionth Person to Receive Safe Water Since 1994*. July 5.
5. Pearson, I. and Idema, G. (1998). An assessment of common problems associated with drinking water disinfection in developing areas. *WRC Report No.* 649/1/98: 1–17.
6. Momba, M.N.B. and Kaleni, P. (2002). Regrowth and survival of indicator microorganisms on the surfaces of household containers used for the storage of drinking water in rural communities of South Africa. *Water Res.* **36**: 3023–3028.
7. Momba, M.N.B., Makala, N., Tyafa, Z., Brouckaert, B.M., Buckly, C.A., and Thompson, P.A. (2004). *Improving the Efficiency and Sustainability of Disinfection at a Small Rural Water Treatment Plant. Presented at the WISA Biennial Conference held at the Cape Town International Conference Centre*. May 2–6.
8. Swartz, C.D. (2000). Guidelines for the upgrading of existing small water treatment plants. *WRC Report No.* 730/1/00.
9. Muyima, N.Y.O. and Ngcakani, F. (1998). Indicator bacteria and regrowth potential of the drinking water in Alice, Eastern Cape. *Water SA*. **24**: 29–34.
10. Department of Water Affairs and Forestry. (1996). *South African Water Quality Guideline for Domestic Use*. 2nd ed. Pretoria.

11. Kawamura, S. (1991). *Integrated Design of Water Treatment Facilities*. John Wiley & Sons, New York.
12. American Public Health Association. (1998). *Standard Methods for the Examination of Water and Wastewater*. 20th Edn. AWWA. Water Environmental Federation, Washington, DC.
13. Servais, P., Billen, G., and Harscoet, M.C. (1987). Determination of the biodegradable fraction of dissolved organic matter in water. *Wat. Res. (G.B.)* **21**(4): 445–450.
14. Grundlingh, J.A., Nel, C., Kotze, E., and De Wet, C.M.E. (1999). Biodegradable compounds and microbial regrowth in water. *WRC Report No.* 541/1/99.
15. Matthieu, L., Block, J.C., Dutang, M., Mailard, J., and Reasoner, D.J. (1993). Control of biofilm accumulation in drinking water distribution systems. *Wat. Suppl.* **11**(3–4): 365–376.
16. Page, D.W. et al. (2002). Effect of alum treatment on the trihalomethane formation and bacterial regrowth potential of natural and synthetic waters. *WRC.* **36**(12): 4884–4892.
17. World Health Organization. (1993). *Guidelines for Drinking Water Quality*, Vol. 1, 2nd Edn. Recommendations, Geneva.
18. Brouckaert, B.M., Pryor, M.J., Pillay, P., Zondi, T.A., and Amirtharajah. (2002). Assessing the use of autonomous valveless filters for turbidity removalin rural applications. *Proceedings of IWA Biennal Conference and Exhibition*. Durban, May 19–23.
19. Department of Water Affairs and Forestry, Department of Health, Water Research Commission. (1998). *Quality of Domestic Water Supplies*, Vol. 1: Assessment Guide, 2nd Edn. Water Research Commision No: TT 101/98.
20. Momba, M.N.B., Cloete, T.E., Venter, S.N., and Kfir, R. (2000). Influence of disinfection processes on the microbial quality of potable groundwater in a laboratory-scale system model. *J. Water SRT—Aqua.* **49**: 23–33.
21. Servais, P., Anzil, A., and Ventersque, C. (1989). Simple method for determination of biodegradable dissolved organic carbon in water. *Appl. Environ. Microbiol.* **55**: 2732.
22. Lechevallier, M.W., Cawthon, C.D., and Lee, R.G. (1988). Mechanisms of bacterial survival in chlorinated drinking water. *Water Sci. Technol.* **20**(11/12): 145–151.
23. Bernhardt, H. and Wilhelms, A. (1985). Erfahrungen mit der desinfektion von trinkwasser durch chlordioxid. *38 Int. CFBEDEAU Konf.—CEBEDOC*, Liège, pp. 193–231.

FILTRATION

National Drinking Water
Clearinghouse

Filtration is the process of removing suspended solids from water by passing the water through a permeable fabric or porous bed of materials. Groundwater is naturally filtered as it flows through porous layers of soil. However, surface water and groundwater under the influence of surface water is subject to contamination from many sources. Some contaminants pose a threat to human health, and filtration is one of the oldest and simplest methods of removing them. Federal and state laws require many water systems to filter their water. Filtration methods include slow and rapid sand filtration, diatomaceous earth filtration, direct filtration, packaged filtration, membrane filtration, and cartridge filtration.

FILTRATION KEEPS WATER SAFE

Why Filter Drinking Water?

Natural filtration removes most suspended matter from groundwater as the water passes through porous layers of soil into aquifers. Surface water, however, may be subject to direct animal, human, and industrial contamination that can cause disease or illness in humans, so they must be filtered by a constructed treatment system.

What Regulations Govern Filtration?

The Surface Water Treatment Rule under the 1986 Safe Drinking Water Act (SDWA) Amendments requires that many surface water supply systems and groundwater under the influence of surface water filter their water supplies.

What Processes Precede Filtration?

Conventional filtration processes are normally preceded by coagulation, flocculation, and sedimentation. Direct filtration processes are preceded by coagulation and flocculation only; the floc is removed directly by the filters.

Filtration processes may include one or all of the following pretreatment procedures:

Chemical Feed and Rapid Mix: Chemicals may be added to the water to improve the treatment processes that occur later. These may include pH adjusters and coagulants. A variety of devices, such as baffles, and static mixers can be used to mix the water and distribute the chemicals evenly.

Flocculation: The chemically treated water is sent into a basin where the suspended particles can collide and form heavier particles called floc. Gentle agitation and appropriate detention times facilitate this process.

Sedimentation: The velocity of water is decreased so that suspended material (including flocculated particles) can settle out of the water stream by gravity. Once settled, the particles combine to form a sludge that is later removed from the clarified supernatant (the liquid removed from settled sludge) water.

How is Filtration Achieved?

Filtration is usually a combination of physical and chemical processes. Mechanical straining removes some particles by trapping them between the grains of the filter medium (such as sand). Adhesion is an equally important process by which suspended particles stick to the surface of filter grains or previously deposited material. Biological processes are also important in slow sand filters. These filters form a filter skin containing microorganisms that trap and break down algae, bacteria, and other organic matter before the water reaches the filter medium itself.

COMPARING FILTRATION SYSTEMS

Slow Sand Filtration

The filter consists of a bed of fine sand approximately 3 to 4 feet deep supported by a 1-foot layer of gravel and an underdrain system (Fig. 1).

Advantages. Low cost, simplex2 operation, reliable, and able to achieve greater than 99.9 percent *Giardia* cyst removal. It also does not require extensive active control by an operator.

Limitations. It is not suitable for water with high turbidity. Filter surface requires maintenance. Extensive land is required due to low-flow operation (0.03 to 0.10 gallons per minute per square foot [gal/min/ft^2] of filter bed area).

Process. Filters are operated under continuous, submerged conditions maintained by adjusting a control valve located on the discharge line from the underdrain system. Biological processes and chemical/physical processes common to various types of filters occur at the surface of the filter bed. A biological slime or mat referred to as "schmutzdecke" forms on the surface of the bed, which traps small particles and degrades organic material present in the raw water. Slow sand filters do not require coagulation/flocculation and may not require sedimentation.

Equipment. Small plants are typically designed with cast-in-place concrete structures with wood or concrete structures with wood or concrete slab covers. Piping is either cast iron or polyvinyl chloride (PVC). Flow meters are used to monitor the output for each filter. In climates subject to freezing temperatures, filters usually must be housed, and may require heating, lighting, and ventilation. Unhoused filters in cold climates develop an ice layer that prevents cleaning during winter months.

Chemicals. Water applied to slow sand filters is not prechlorinated because the chlorine can destroy organisms in the schmutzdecke.

Diatomaceous Earth Filtration

Diatomaceous earth filtration, also known as precoat or diatomite filtration, relies on a layer of diatomaceous earth approximately 1/8-inch thick placed on a septum or filter element. Septums may be placed in pressure vessels or operated under a vacuum in open vessels.

Advantages. The filters are simple to operate and effective in removing cysts, algae, and asbestos. They have been chosen for projects with limited initial capital, and for emergency or standby capacity to service large seasonal increases in demand.

Limitations. This filter is most suitable for water with low bacterial counts and low turbidity (less than 10 nephelometric turbidity units [NTU]). Coagulant and filter aids are required for effective virus removal. There is potential difficulty in maintaining complete and uniform thickness of diatomaceous earth on the filter septum.

Process. Operation and maintenance of diatomaceous earth filters require: preparing slurries of filter body feed (diatomaceous earth) and precoat diatomaceous earth; adjusting body feed dosages for effective turbidity removal; periodic backwashing—every 1 to 4 days; disposing of spent filter cake; periodically inspecting the filter septum for cleanliness and damage; and verifying effluent quality.

Equipment. The minimum amount of filter precoat should be 0.2 pounds per square foot (lb/ft^2); and minimum thickness of the precoat should be increased from 1/8 to 1/5 inch to enhance *Giardia* cyst removal. In addition, minimum design criteria outlined in the *Recommended Standards for Water Works* (better known as 10 State Standards) should be met.

Chemicals. Use coagulant to coat the body feed to improve removal rates for viruses, bacteria, and turbidity.

Direct Filtration

Direct filtration systems are similar to conventional systems, but omit sedimentation.

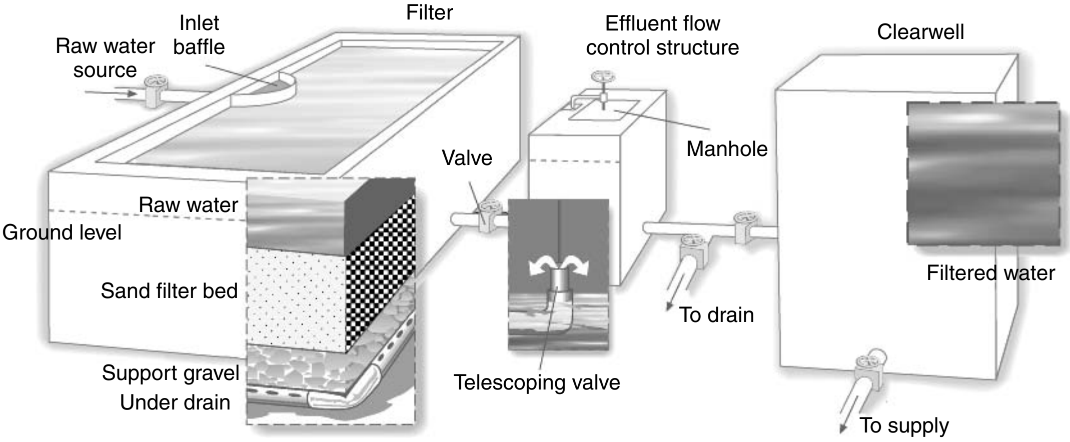

Figure 1. Slow sand filter.

Advantages. Effective direct filtration performance ranges from 90 to 99 percent for virus removal and from 10 to 99.99 percent for *Giardia* removal. The most effective direct filtration configurations for *Giardia* removal must include coagulation. Often used with steel pressure vessels to maintain the pressure in a water line to avoid repumping after filtration.

Limitations. Direct filtration is only applicable for systems with high quality and seasonally consistent influent supplies. The influent generally should have turbidity of less than 5 to 10 NTU and color of less than 20 to 30 units. (Water with 15 or more units of color causes aesthetic problems, such as staining.)

Process. Direct filtration consists of several combinations of treatment processes. It always includes coagulation and filtration, and may require a flocculation tank or a pressure vessel after the coagulation addition.

Equipment. Dual- and mixed-media filters are used to effectively process higher influent turbidities without the use of sedimentation.

Chemicals. Typical coagulant dosages range from less than 1 to 30 milligrams per liter. Cationic polymers often successfully coagulate water supplies and assist direct filtration. Nonionic polymers are sometimes added to the filtration step to increase filter efficiency.

Packaged Filtration

Packaged filtration is simply all of the features of filtration—chemical addition, flocculation, sedimentation, filtration—mounted as a unit on a frame for simple hookup of pipes and services. It is most widely used to treat surface water supplies for removal of turbidity, color, and coliform organisms with filtration processes. Packaged filtration is often used to treat small community water supplies, as well as supplies in recreational areas, state parks, construction sites, ski areas, and military installations, among others.

Advantages. The four major advantages of package plants are their compact size, cost effectiveness, relative ease of operation, and design for unattended operation. (Some states require that an operator be in attendance at all times. Check your state regulations.)

Limitations. When the turbidity of the raw water varies a great deal, these package plants require a high level of operator attention and skill.

Process. Package plants are most appropriate for plant sizes ranging from 0.025 to 6 million gallons per day. The most important factor to consider in selecting a package plant is the influent characteristics, such as temperature, turbidity, and color levels. Pilot tests might be necessary before a final system can be selected.

Equipment. Package plants are assembled in a factory, skid mounted, and transported to the treatment site, or they are transported as component units to the site and then assembled.

Chemicals. Chemical feed controls are especially important for plants without full-time operators or with variable influent characteristics. Even with these automated devices, however, the operator needs to be properly trained and well acquainted with the process and control system.

Membrane Filtration

A membrane is a thin layer of material capable of separating substances when a driving force is applied across the membrane.

Advantages. Membrane filtration can be an attractive option for small systems because of its small size and automated operation. Membrane processes are increasingly employed for removal of bacteria and other microorganisms, particulate material, and natural organic material, which can impart color, tastes, and odors to water.

Limitations. Fouling of the membranes is the major problem preventing widespread application of this technology.

Process. Membrane filtration works by passing water at high pressure through a thin membrane in the form of hollow fiber or spiral-wound composite sheets. Organic and other contaminants are retained on the high-pressure side and frequently must be removed by reversing the flow and flushing the waste. Periodic chemical cleaning may be required to remove persistent contaminants. Membrane assemblies are contained in pressure vessels or cartridges.

Equipment. The membrane technologies are relatively simple to install and for groundwater sources that do not need pretreatment, the systems require little more than a feed pump, a cleaning pump, the membrane modules, and some holding tanks.

Chemicals. Periodic backflushing and occasional chemical cleaning is necessary to maintain the membrane or fibers.

Cartridge Filtration

Cartridge filters are considered an emerging technology suitable for removing microbes and turbidity in small systems (Fig. 2).

Advantages. Cartridge filters are easy to operate and maintain, making them suitable for small systems with low-turbidity influent.

Limitations. Cartridge filtration systems require raw water with low turbidity. Polypropylene cartridges become fouled relatively quickly and must be replaced with new units. Although these filter systems are operationally simple, they are not automated and can require relatively large operating budgets.

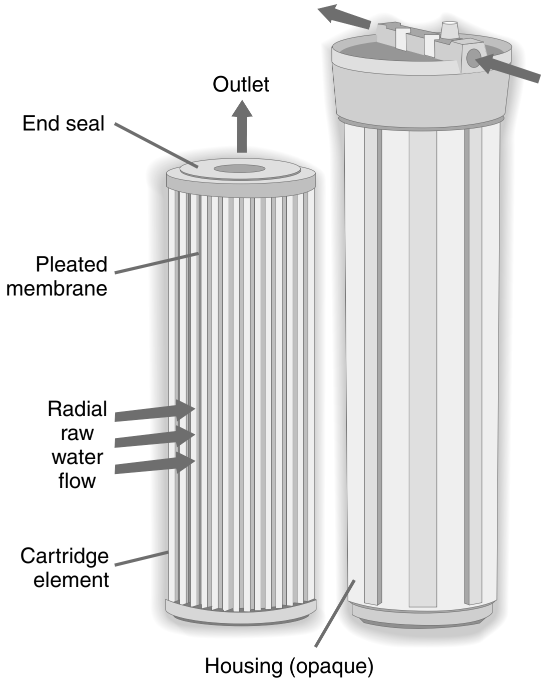

Figure 2. Cartridge filter with replaceable inner package.

Process. Cartridge filtration uses a physical process—straining water through porous media. It can exclude particles 0.2 micrometers (μm) or smaller. The pore sizes that are suitable for producing potable water range from 0.2 to 1.0 μm. Roughing filters, for pretreatment prior to cartridge filtration, are sometimes necessary to remove larger suspended solids and prevent the rapid fouling of the cartridges. Roughing filters can be rapid sand filters, multimedia filters, fine mesh screens, or bag filters.

Equipment. A cartridge consists of ceramic or polypropylene filter elements fitted into pressurized housings.

Chemicals. A disinfectant is recommended to prevent surface-fouling microbial growth on the cartridge filters and to reduce microbial pass-through. Except for a disinfectant, no chemical additions are necessary. However, corrosive chemicals may be required for the periodic membrane cleaning process.

How Does One Select the Appropriate Filtration System?

First, review all raw water quality data to establish the requirements for the alternatives. Once the potential alternatives are selected, determine the necessity of pilot or bench-scale tests. If the desired performance of one or more of the alternatives is in doubt, testing is appropriate. (Testing is always useful if time and budget allow).

Otherwise, literature surveys, bench-scale studies, or pilot-test results can be used to derive each alternative's performance characteristics and design considerations. Following this initial selection, the basic process concerns for the various alternatives should be identified and evaluated, including:

- turbidity removal performance,
- *Giardia* removal performance,
- color removal performance,
- cleaning cycle frequency,
- necessary chemicals/chemical dosages,
- applicable regulatory standards,
- required operational skills, and
- necessary sludge management.

Where Can I Find More Information?

Information in this fact sheet was primarily obtained from two sources: Environmental Pollution Control Alternatives: Drinking Water Treatment for Small Communities, *EPA/625/5-90/025; and* Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities, *EPA/625/4-89/023. Both can be ordered free from the U.S. Environmental Protection Agency Office of Research and Development at (513) 569-7562.*

If these publications are no longer available from the EPA, call the National Drinking Water Clearinghouse (NDWC). Environmental Pollution Control Alternatives: Drinking Water Treatment for Small Communities, *item #DWBKGN09, an 82-page publication, costs $11.82.* Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities, *item #DWBKDM04, a 209-page book, costs $30.05. Shipping and handling charges apply.*

Also, the NDWC's Registry of Equipment Suppliers of Treatment Technologies for Small Systems (RESULTS) is a public reference database that contains information about technologies in use at small water systems around the country. For further information on accessing or ordering RESULTS, call the NDWC at (800) 624-8301.

For additional copies of "Tech Brief: Filtration," item #DWBRPE5O, or a copy of "Tech Brief: Disinfection," item #DWBRPE47, call NDWC at the number above.

WATER FILTRATION

Rasheed Ahmad
Khafra Engineering Consultants
Atlanta, Georgia

Particulates that are removed by a water filtration process include microorganisms (bacteria, viruses, and protozoan cysts), clay and silt particles, colloidal and precipitated humic substances and other organic particulates from natural decay of vegetation, precipitates of an aluminum or iron coagulation process, calcium carbonate and magnesium hydroxide precipitates from lime softening, and iron and manganese precipitates.

Granular medium filtration is a water treatment process that uses a porous medium through which water passes to remove particulates or suspended solids (Fig. 1). There are different types of filters classified by various schemes. Filters can be classified based on the type of medium (single medium, dual medium), hydraulic arrangement (gravity or pressure), rate of filtration (rapid

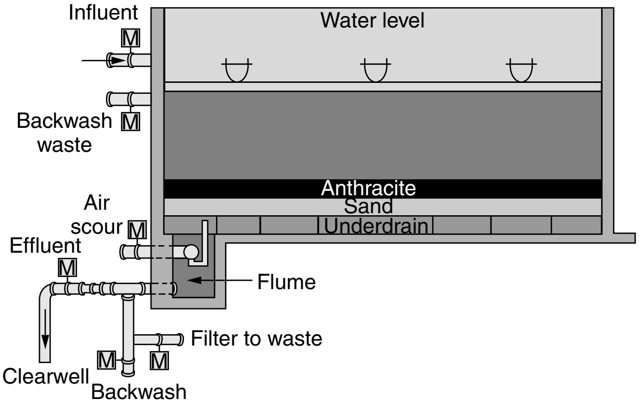

Figure 1. Rapid gravity filter (*Source*: F.B. Leopold Company).

or slow), and depth of solids removal (deep or cake). The most common types of filter media used in granular bed filters are silica sand, anthracite coal, and garnet. These may be used alone or in dual- or triple-media combinations. Granular-activated carbon (GAC), another type of medium, has been used to reduce taste and odor in granular beds that serve for both filtration and adsorption.

Rapid granular bed filtration consists of passing pretreated water through a granular bed at rates ranging from 2 to 10 gpm/sq ft (5 to 25 m/h). Flow is typically downward through the bed. During the filtration cycle, solids are removed from the water and accumulate within the voids and on the top surface of the filter medium. This accumulation results in clogging and a gradual increase in head loss. When the head loss becomes excessive (generally 8–9 ft), the filtrate quality begins to deteriorate, and/or the filter cycle reaches a predetermined time limit (usually 3 or 4 days), then rapid filters need cleaning.

Dirty/clogged filters may be cleaned by scouring the clogged portion or by reversing the flow through the bed. This application of expanding and washing out trapped particles is called backwashing. Backwashing by water fluidization is frequently assisted by a surface wash or an air scour. The most effective backwash is achieved by simultaneous air scour and subfluidization water backwash. Typical backwash rates range from 15 to 23 gpm/sq ft depending on media sizes. The bed expansion varies from 20% to 50%. The backwash usually lasts from 5 to 15 minutes. The air-scour rate varies from 2 to 4 cu ft/min/sq ft (0.6 to 1.2 cu meter/min/sq meter). For effective rapid granular bed filtration, source water must be pretreated.

Chemical destabilization is an essential prerequisite for effective filtration. Chemicals used for particle destabilization are limited primarily to metal salts or cationic polymers as primary coagulants. Pretreatment may also include aeration or introducing an oxidant if water treatment aims to remove iron or manganese. Sometimes, a filter aid polymer is added in the influent to the filter to improve particle capture efficiency.

Granular bed filtration consists of three principal mechanisms: (1) transport, (2) attachment, and (3) detachment. *Transport* mechanisms move a particle into and through a filter pore so that it comes very close to the surface of the filter medium or existing deposits where *attachment* mechanisms retain the suspended particle in contact with the medium's surface or with previously deposited solids. *Detachment* mechanisms result from the hydrodynamic forces of the flow acting so that a certain portion of the previously attached particles, less strongly adhered to others, is detached from the filter medium or previous deposits and carried further, deep into or through the filter. Important transport mechanisms include screening, interception, inertial forces, sedimentation, diffusion, and hydrodynamic forces. Attachment of particles to media surfaces is generally governed by physicochemical and molecular forces. Detachment is caused by the impact of arriving particles on unstable deposits and by hydraulic shear stresses without the influence of arriving particles. A number of properties of a filter medium are important for filtration performance and for characterizing the medium. Important media properties include size, shape, density, and hardness. The efficiency of a filter is more sensitive to changes in the filtration rate rather than in the actual rate. Therefore, filtration process control is critical to successful operation.

In recent years, direct filtration has received considerable attention and application in treating of drinking water. Direct filtration is not preceded by sedimentation. It offers several advantages over conventional treatment of "good-quality" surface water. Because there is no sedimentation process and a lower coagulant dose, the capital and operating costs are lower compared with conventional treatment. Sludge volumes are lower, which results in lower chemical costs for sludge treatment and disposal. Because of the increased interest in ozone application in potable water, biological filtration has received noticeable attention and application in recent years. Granular filters become biologically active when ozone is used as a preoxidant. Ozone reacts with organic compounds in water and forms several ozonation byproducts such as aldehydes, carboxylic acids, and keto acids, which are relatively easy to biodegrade. By operating filters in a biological mode, these ozonation byproducts can be effectively removed and in turn reduce the potential for bacterial regrowth in a distribution system and formation of chlorination byproducts.

A slow sand filter generally consists of a watertight basin containing a layer of sand over a layer of gravel (Fig. 2). It is operated at very low filtration rates without coagulation in pretreatment. A smaller grain size and lower filtration rate result in removing the solids almost entirely in a thin layer on the top of the sand bed. This layer, composed of dirt and living and dead micro- and macro-organisms from the water, is called *schmutzdecke* or dirty skin. Most of the solid removals in a slow sand filter take place in the *schmutzdecke* layer. The dominant removal mechanisms are both physical and biological. A typical filter cycle may vary from 1 to 6 months (or longer) depending on the source water quality and the filtration rate. Filtration rates range from 0.016 to 0.16 gpm/sq ft (0.04 to 0.40 m/h). Effective sand sizes ranged from 0.15 to 0.40 mm, sand uniformity coefficients from 1.5 to 3.6, and initial bed depths from 1.5 to 5.0 ft. The sand is supported on graded gravel 6 to 36 inches deep. Slow sand filters

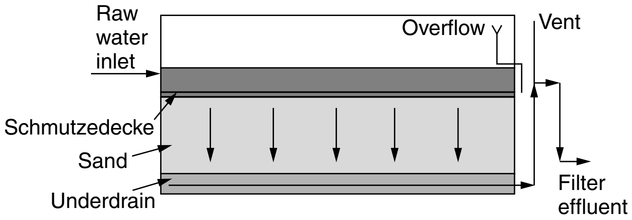

Figure 2. Conventional slow sand filter.

are cleaned by scraping *schmutzdecke* and a small amount of sand depth. Usually scraping is done manually but, in some cases, mechanically. The scraped sand is cleaned hydraulically and stockpiled for later reuse.

Diatomaceous earth filtration, also called precoat or diatomite filtration, uses a layer of diatomaceous earth approximately one-eighth inch thick placed on a septum or filter element as a filter medium. The septum may be placed in a pressure vessel or operated under a vacuum in an open vessel. As the water passes through the filter medium and septum, the suspended particles are captured and deposited. The majority of particles removed by the filters are strained at the surface layer of the filter medium, and some are trapped within the layer. As the filter cycle proceeds, additional filter medium called *body feed* is regularly metered into the influent water in proportion to the solids removed. Ultimately, a gradually increasing pressure drop through the filter system reaches a point where continued filtration is impractical. The forward filtration process is stopped, the filter medium and collected dirt are washed off the septum, a new precoat of filter medium is applied, and the filtration process continues. Diatomaceous earth filters are widely used in industrial filtration applications and in swimming pool filtration. They have also been used in municipal potable water treatment, primarily in direct, in-line filtration of high-quality surface water (turbidity 10 NTU or less and acceptable color), and in filtering iron and manganese from groundwater after appropriate pretreatment to precipitate these contaminants. Precoat filters are simple to operate and are effective in removing cysts, algae, and asbestos.

Membrane filtration represents an important set of processes for drinking water treatment. A membrane is a thin layer of natural or synthetic material that can separate substances when a driving force is applied across the membrane. Membranes used for water treatment are commonly made of a synthetic organic polymer. Membrane processes applicable to potable water treatment are reverse osmosis (RO), nanofiltration (NF), electrodialysis (ED), ultrafiltration (UF), and microfiltration (MF). Reverse osmosis is primarily used to remove salts from brackish water or seawater. Nanofiltration is used to soften freshwaters and remove disinfection byproduct (DBP) precursors. Electrodialysis is used to demineralize brackish water and seawater and to soften freshwater. Ultrafiltration and microfiltration are used to remove turbidity, pathogens, and other particulates from freshwater (Fig. 3).

Membranes are normally classified by solute exclusion size, which is sometimes referred to as *pore size*. Membrane filtration works by passing water at high pressure through a thin membrane in the form of hollow fiber or spiral-wound composite sheets. The contaminants are retained on the high-pressure side and frequently must be cleaned by reversing the flow and flushing the waste. Periodic chemical cleaning may be required to remove persistent contaminants. Membrane assemblies are contained in pressure vessels or cartridges. Low-pressure membranes, in the form of either ultrafiltration (UF) or microfiltration (MF), have become more economical in both capital and operating costs and have received increased attention in drinking water applications. Fouling of the membrane is the major problem preventing widespread use of this technology.

Figure 3. Microfiltration (MF) system (*Source*: U.S. Filter).

Cartridge filters traditionally have been applied to point-of-use systems and for pretreatment prior to membrane treatment systems. Composed of membrane, fabric, or string filter media, the filter material is supported by a filter element and housed in a pressure vessel. The application of cartridge filters using either a cleanable ceramic or disposable polypropylene cartridge seems to be a feasible method for removing modest levels of turbidity, algae, and microbiological contaminants. As water is filtered through a cartridge filter, the pressure drop increases, which necessitates terminating the filter run. When this process is done, the filter is not backwashed, but rather the cartridge is thrown away and replaced by a clean filter.

The desire to reduce costs or to treat some waters more effectively is driving the development of new technologies involving filtration. Several examples of such proprietary filters include low-head continuous backwash filters, two-stage filtration systems, bag filters, moving media filters, and flotation and filtration systems.

READING LIST

ASCE and AWWA. (1998). *Water Treatment Plant Design*, 3rd Edn. McGraw-Hill, Inc., New York, NY.

AWWARF. (1998). *Treatment Process Selection for Particle Removal*. American Water Works Association Research Foundation, Denver, CO.

AWWA. (1999). *Water Quality and Treatment*, 5th Edn. McGraw-Hill, Inc., New York, NY.

HDR, Inc. (2001). *Handbook of Public Water System*, 2nd Edn. John Wiley & Sons, New York, NY.

Hudson, H.E. (1981). *Water Clarification Processes*. Van Nostrand Reinhold Company, New York, NY.

Kawamura, S. (2000) *Integrated Design of Water Treatment Facilities*. John Wiley & Sons, New York, NY.

Qasim, S.R., Motley, E.M., and Zhu, G. (2000). *Water Works Engineering*. Prentice-Hall, Upper Saddle River, NJ.

USEPA. (1990). *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*. Office of Drinking Water, Center for Research Information, Cincinnati, OH.

FILTRATION WITH GRANULAR MEDIA

Rasheed Ahmad
Khafra Engineering Consultants
Atlanta, Georgia

REGULATORY REQUIREMENTS FOR DRINKING WATER FILTRATION

The U.S. Environmental Protection Agency's (USEPA) Surface Water Treatment Rule (SWTR), promulgated on June 29, 1989 (Federal Register 40 CFR Parts 140 and 141), requires community water systems to disinfect all surface waters and requires filtration for most surface water sources. The 1998 Interim Enhanced Surface Water Treatment Rule (IESWTR) requires that surface water or groundwater systems under the influence of surface waters must produce a turbidity of less than or equal to 0.3 NTU in 95% of measurements taken each month. This rule was later extended to smaller systems as part of the Long Term Enhanced Surface Water Treatment Rule (LTESWTR).

FILTRATION HYDRAULICS

Pressure drop (i.e., head loss) occurs when water flows through a bed of a clean medium. The head loss caused by water flowing through a filter medium can be calculated from well-known equations such as the Kozeny equation. At the velocities commonly employed in granular water filters, flow is normally laminar. Initial head losses in clean filters commonly range from less than 1 to 2.5 ft, depending on the particle size distribution of the medium and the overflow rate (flow/filter surface area). The error in the head loss equation increases at flow rates used for filter backwashing where calculations show that the Reynolds number of particles lies in the transitional range. The most commonly used expression to calculate these head losses is the two-term Ergun equation, which applies to all regimes of flow. The flow through the underdrain system and effluent pipes is frequently turbulent. Negative head (less than atmospheric pressure) can occur in a gravity filter when the summation of head loss from the filter medium surface downward exceeds the pressure available.

PARTICLE REMOVAL MECHANISMS DURING FILTRATION

Filtration consists of three principal mechanisms: (1) transport, (2) attachment, and (3) detachment. *Transport* mechanisms move a particle into and through a filter pore so that it comes very close to the surface of the filter medium or existing deposits where *attachment* mechanisms retain the suspended particle in contact with the medium's surface or with previously deposited solids. *Detachment* mechanisms result from the hydrodynamic forces of the flow acting so that a certain portion of the previously attached particles, less strongly adhered to others, is detached from the filter medium or previous deposits and is carried further, deep into or through the filter. Important transport mechanisms include screening, interception, inertial forces, sedimentation, diffusion, and hydrodynamic forces. Physicochemical and molecular forces generally govern attachment of particles to media surfaces. Detachment is caused by the impact of arriving particles on unstable deposits and by hydraulic shear stresses without the influence of arriving particles.

TYPES OF FILTER MEDIA AND PROPERTIES

The most common types of filter media used in granular bed filters are silica sand, anthracite coal, and garnet. These media may be used alone or in dual- or triple-media combinations. Granular-activated carbon (GAC), another type of medium, has been used in granular beds for both filtration and adsorption to reduce taste and odor. A number of properties of a filter medium are important in filtration performance and in characterizing the medium. Important media properties include size, shape, density, and hardness. The porosity of the granular bed formed by the grains is also an important governing factor in filtration performance. Selection of filter media type and characteristics is based on a number of design decisions concerning source water quality, pretreatment, and desired filtered water quality. Filter media cleaning requirements and underdrain system options depend on the filter configuration and filter medium selected. Pilot plant studies using alternative filter media and filtration rates can determine the most effective and efficient medium for a particular water.

FILTER CLASSIFICATION

Several different types of filters are based on various classification schemes. Filters can be classified by type of media (single media, dual media), hydraulic arrangement (gravity or pressure), rate of filtration (rapid or slow), or depth of solids removal (deep or cake).

Slow Sand Filters

A slow sand filter generally consists of a watertight basin containing a layer of sand over a layer of gravel. It is operated at very low filtration rates without coagulation in pretreatment. Smaller grain size and lower filtration rate remove the solids almost entirely in a thin layer on top of the sand bed. This layer, composed of dirt and living and dead micro- and macro-organisms from the water, is called *schmutzdecke* or dirty skin. Most solid removals in a slow sand filter take place in the *schmutzdecke* layer. The dominant removal mechanisms are both physical and biological. A typical filter cycle may vary from 1 to 6 months (or longer) depending on the source water quality and the filtration rate. Filtration rates range from 0.016 to 0.16 gpm/sq ft (0.04 to 0.40 m/h). Effective sand sizes range from 0.15 to 0.40 mm, sand uniformity coefficients from 1.5 to 3.6, and initial bed depths from 1.5 to 5.0 ft. The sand is supported on graded gravel 6 to 36 inches deep. Slow sand filters are cleaned by scraping *schmutzdecke* and a small amount of sand depth. Usually scraping is done manually but, in some cases, mechanically. The scraped sand is cleaned hydraulically and stockpiled for later reuse.

Rapid Granular Bed Filters

Rapid granular bed filtration consists of passing of pretreated water through a granular bed at rates ranging from 2 to 10 gpm/sq ft (5 to 25 m/h). Flow is typically downward through the bed. During the filtration cycle, solids are removed from the water and accumulate within the voids and on the top surface of the filter medium. This accumulation results in clogging and a gradual increase in head loss. When head loss becomes excessive (generally 8–9 ft), filtrate quality begins to deteriorate, and/or filter cycle reaches a predetermined time limit (usually 3 or 4 days), then rapid filters need cleaning. Dirty/clogged filters may be cleaned by scouring the clogged portion or by reversing the flow through the bed. This application of expanding and washing out trapped particles is called backwashing. Backwashing by water fluidization is frequently assisted by a surface wash or an air scour. The most effective backwash is achieved by simultaneous air scour and subfluidization water backwash. Typical backwash rates range from 15 to 23 gpm/sq ft depending on media sizes. The bed expansion varies from 20% to 50%. The backwash usually lasts from 5 to 15 minutes. The air-scour rate varies from 2 to 4 cu ft/min/sq ft (0.6 to 1.2 cu meter/min/sq meter) (Fig. 1).

Filter Underdrain Systems

Underdrain systems are used to support the filter medium, collect filtered water (in downflow filters), and distribute backwash water (and air, if employed). Four basic types of underdrain systems are commonly used: pipe laterals, blocks, false bottom, and porous bottom. Pipe underdrain systems generally consist of a centrally located pipe to which are attached smaller, equally spaced laterals. The lateral pipes usually have one or two rows of less than one-inch diameter perforations on their bottom sides. The lateral pipes may be fitted with nozzles. The blocks are self-supporting type underdrains. The block underdrain consists of vitrified clay blocks with one-fourth inch diameter dispersion orifices across the top of each block. Support gravel is typically used with this type of underdrain. As a replacement for support gravel, an integral media support (IMS) cap made of plastic beads sintered together may be installed. False-bottom underdrains are made of precast or cast-in-place reinforced concrete supported on concrete sills. This underdrain system consists of uniformly spaced inverted pyramidal depressions. Unglazed porcelain spheres are placed in the depressions to distribute flow. Porous-bottom underdrains constructed of porous aluminum oxide plates have been used in both block and false-bottom configurations. These types of underdrains are constructed of plates mounted on

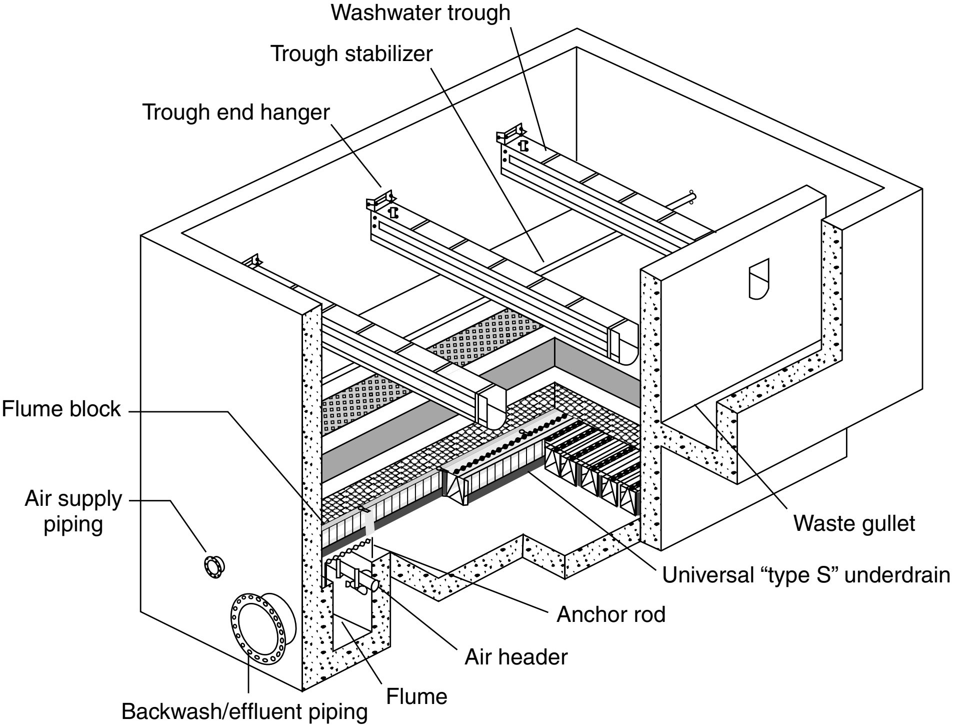

Figure 1. Rapid gravity filter box (*Source*: F.B. Leopold Co.).

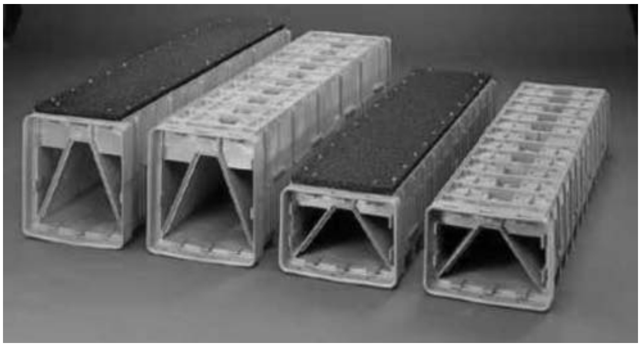

Figure 2. Filter underdrain (*Source*: F. B. Leopold Co.).

concrete or steel piers or on clay tile saddles to form blocks (Fig. 2).

FILTER OPERATION AND CONTROL

The efficiency of a filter is more sensitive to changes in the filtration rate rather than in the actual rate. Therefore, filtration process control is critical to successful operation. Two basic modes of gravity filter control are commonly found: constant rate and declining rate. Constant-rate mode operation of a filter can be achieved in three ways: (1) use of a flow control valve in the filtered water piping; (2) influent flow splitting where the water level over the filter is maintained at a constant level; and (3) influent flow splitting where the water level varies during the filter run. Declining-rate filters, which are most widely designed, are equipped with effluent weirs rather than with rate controllers. Flow is distributed on the basis of the relative conditions of the beds. Common filter problems include inadequate pretreatment or filter washing, gravel bed upset, air binding, restart after shut down, and filter media replacement. These problems lead to operational difficulties like dirty filter media, mudballs, and mineral deposits.

INNOVATIONS IN GRANULAR FILTRATION TECHNOLOGIES

In recent years, direct filtration has received considerable attention and application in treating of drinking water. Direct filtration is not preceded by sedimentation. It offers several advantages over conventional treatment of “good-quality” surface water. Because there is no sedimentation process and a lower coagulant dose, the capital and operating costs are lower compared with those of conventional treatment. Sludge volumes are lower, which results in lower chemical costs for sludge treatment and disposal. Because of increased interest in ozone application in potable water, biological filtration has received noticeable attention and application in recent years. Granular filters become biologically active when ozone is used as a preoxidant. Ozone reacts with organic compounds in water and forms several ozonation byproducts such as aldehydes, carboxylic acids, and keto acids, which are relatively easy to biodegrade. By operating filters in a biological mode, these ozonation byproducts can be effectively removed and in turn reduce the potential for bacterial regrowth in distribution systems and formation of chlorination byproducts. The desire to reduce costs or to treat some waters more effectively is driving the development of new technologies for filtration. Several examples of such proprietary filters include low-head continuous backwash filters, two-stage filtration systems, bag and cartridge filters, moving media filters, and flotation and filtration systems.

SLOW SAND FILTRATION AND THE IMPACT OF *SCHMUTZDECKE*

Stephen J. Rooklidge
Aurora, Oregon

INTRODUCTION

The use of sand media to filter drinking water dates to the early nineteenth century; installations in Paisley and Greenock, Scotland are the earliest known successful uses of the technology. The European cholera epidemic of the early 1890s proved conclusively the benefits of filtered water supplies to the health of urban populations (1). Current construction of slow sand filter systems are based on the design refined by Simpson for the London-area Chelsea Water Company in 1829 (2).

Rural areas are the primary benefactors of slow sand technology due to the passive nature of the treatment process and the availability of large tracts of land necessary for filter construction. The lack of a need for constant pumping to maintain the filtration rate through the filter, the ease of operation, and the possibility of using nonmechanical maintenance procedures make slow sand filtration an appropriate technology for small and rural communities (3).

SLOW SAND FILTERS

The source water flows by gravity through the filter column. The supernatant level can be kept constant as the effluent flow rate declines with head loss in the sand media, or the influent may be adjusted for constant flow while the supernatant level rises. The water flows through a layer of sand (0.5–1.0 m), travels through a layer of larger gravel (0.3–0.6 m), and flows out from a perforated pipe or tile underdrain collection grid to the treated water storage and distribution system (Fig. 1).

During slow sand filtration, biomass forms at the sand/water interface (*schmutzdecke*) that provides an effective layer for removing pathogenic micro-organisms. The biomass is formed from populations of algae, diatoms, protozoa, rotifers, and bacteria (4). Also present are extracellular polysaccharides exuded by microbial populations inherent in surface water sources (5), and this polymer formation has been suggested as a source of destabilization for bacteria and clay entering the filter (6).

The importance of *schmutzdecke* on overall filter performance is illustrated by removing it when scraping

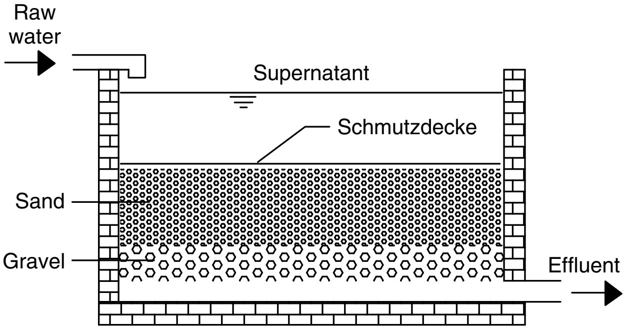

Figure 1. Schematic of a common slow sand filter design.

the top centimeter of media during maintenance; this immediately lessens the head loss caused by biomass growth and returns the flow rate to design levels ($0.1-0.2\ \text{m}^3/\text{m}^2/\text{h}$). After removing the *schmutzdecke*, a ripening period is required to allow new biomass to mature. Effluent is wasted until turbidity and biological criteria are within established limits because influent contaminant removal is diminished during the filter ripening period. Efficient filter ripening can be altered by factors such as ambient and source water temperature, aquatic bacterial populations, and available source water nutrient supply (4).

The combination of biological and physical removal mechanisms in slow sand filters, uninterrupted use of slow sand filters during the 2–18-month filtration cycle, proper filter design, and the diligence of operating personnel produces drinking water of excellent quality (Table 1) from raw water sources with low turbidity (7). Source water quality is an important design consideration when communities consider slow sand filtration as a treatment method. The recommended influent turbidity limit for optimum slow sand filter operation is less than 10 nephelometric turbidity units (NTU), and many designers consider source water that regularly exceeds this limit unacceptable for slow sand filter treatment (8). Raw water sources that occasionally exceed turbidity limits are also regarded with apprehension unless the slow sand filter is augmented by a pretreatment method that reduces influent turbidity. Passive pretreatment systems (multistage filtration) that are effective and inexpensive to operate have been developed to mitigate many slow sand filtration deficiencies.

Table 1. Common Treatment Performance of Slow Sand Filters

Parameter	Removal
Color	30–100%
Turbidity	>90%
Fecal coliform	1–3 log units
Total organic carbon	15–25%
Trihalomethane precursors	<25%
Enteric viruses	2–4 log units
Giardia cysts	2–4 log units
Organic matter	60–75% as COD
Heavy metals	30–95%

Research continues toward better understanding of the biological and physical treatment mechanisms of slow sand filtration (9). Combining basic research on solids penetration (10), pathogen removal (11), and computer modeling (12,13) with operational modifications of sand media (14), filtration rates (15), maintenance and cleaning (16), and corrosion control amendments (17) have led to a substantial increase in the understanding and application of slow sand filtration.

SUMMARY

Slow sand filtration is one of the oldest drinking water treatment techniques and is still a viable method of biological water treatment best suited for raw water sources low in turbidity and suspended solids.

BIBLIOGRAPHY

1. Hazen, A. (1913). *The Filtration of Public Water Supplies*, 3rd Edn. John Wiley & Sons, New York.
2. Baker, M. (1948). *The Quest for Pure Water: The History of Water Purification From the Earliest Records to the Twentieth Century*. AWWA, Denver, CO.
3. Visscher, J., Paramasivam, R., Raman, A., and Heijen, H. (1987). *Slow Sand Filtration for Community Water Supply, Planning, Design, Construction, Operation, and Maintenance.* Technical Paper 24, International Water Centre, The Hague, the Netherlands.
4. Huisman, L. and Wood, W. (1974). *Slow Sand Filtration*. World Health Organization, Geneva, Switzerland.
5. Cunningham, A., Characklis, W., Aberdeen, F., and Crawford, D. (1991). Influence of biofilm accumulation on porous media and hydrodynamics. *Environ. Sci. Technol.* **25**: 1305–1311.
6. Bellamy, W., Silverman, G., and Hendricks, D. (1985). Filtration of *giardia* cysts and other substances. *Slow Sand Filtration*, Vol. 2, Report EPA/600/s2-85/026, USEPA.
7. AWWA/ASCE. (1998). *Water Treament Plant Design*, 3rd Edn. McGraw-Hill, New York.
8. AWWARF. (1991). *Manual of Design for Slow Sand Filtration*, D. Hendricks (Ed.). AWWARF and AWWA, Denver, CO.
9. Logsdon, G., Kohne, R., Abel, S., and LaBlonde, S. (2002). Slow sand filtration for small water systems. *J. Environ. Eng. Sci.* **1**: 339–348.
10. Ellis, K. and Aydin, M. (1995). Penetration of solids and biological activity into slow sand filters. *Water Res.* **29**(5): 1333–1341.
11. Calvo-Bado, L. et al. (2003). Spatial and temporal analysis of the microbial community in slow sand filters used for treating horticultural irrigation water. *Appl. Environ. Microbiol.* **69**(4): 2116–2125.
12. Ojha, C. and Graham, N. (1994). Computer-aided simulation of slow sand filter performance. *Water Res.* **28**(5): 1025–1030.
13. Rooklidge, S., Burns, E., and Bolte, J. (in press). Modeling antimicrobial contaminant removal in slow sand filtration. *Water Res.*
14. McMeen, C. and Benjamin, M. (1996). Removal of natural organic matter by slow sand filtration through iron-oxide-coated-olivine. In: *Advances in Slow Sand and Alternative Biological Filtration.* N. Graham and R. Collins (Eds.). John Wiley & Sons, West Sussex, England.

15. Van der Hoek, J., Bonne, P., Kors, L., and Te Welscer, R. (1996). Slow sand filtration: Effects of grain size and filtration rate on operation and performance. In: *Advances in Slow Sand and Alternative Biological Filtration*, N. Graham and R. Collins (Eds.). John Wiley & Sons, West Sussex, England.
16. Kors, L., Wind, A., and van der Hoek, J. (1996). Hydraulic and bacteriological performance affected by resanding, filtration rate and pretreatment. In: *Advances in Slow Sand and Alternative Biological Filtration*, N. Graham and R. Collins (Eds.). John Wiley & Sons, West Sussex, England.
17. Rooklidge, S. and Ketchum, L. (2002). Corrosion control enhancement from a dolomite-amended slow sand filter. *Water Res.* **36**(11): 2689–2694.

MULTISTAGE DRINKING WATER FILTRATION

Stephen J. Rooklidge
Aurora, Oregon

INTRODUCTION

Multistage drinking water filtration is a combination of roughing filter and slow sand filter unit processes that broadens the use of these passive treatment methods to include source water whose turbidity is higher than that normally suitable for biological filtration. Slow sand filtration is one of the oldest surface water treatment techniques, but it is easily compromised by source water with high turbidity caused by erosional or meteorologic events in the upstream watershed. By providing roughing filter pretreatment, suspended solids are decreased, and the biological integrity of the slow sand filter is maintained with a concomitant extension of the filtration period.

ROUGHING FILTER PRETREATMENT

Roughing filters are generally designed similarly to slow sand filters, except for larger media grain size and faster filtration rates. Roughing filters are usually composed of individual flow-through filter compartments that contain decreasing sizes of gravel in the direction of flow. The crushed gravel media in the roughing filter provide a large surface area for the flowing water to contact, allowing flocculation from particle collision in the circuitous route through the rock and sedimentation within the interstitial media voids. The addition of roughing filter pretreatment maintains the simplicity of slow sand filtration and achieves suspended solids and turbidity removal acceptable for drinking water.

DYNAMIC ROUGHING FILTERS

Passive filter systems in common use near rivers or canals include dynamic roughing filtration (Fig. 1). This design allows gravity flow from source water through a bed of crushed gravel (6 mm average diameter), and the overflow is wasted back to the river channel. A drawback of the design is the need to remove and clean the gravel media when the filter becomes clogged during high turbidity, and this limits dynamic filtration to raw water sources that experience turbidity peaks of less than a few hours in duration. The substantial filtration rate of 5 m/h in these filters is not designed to alter water quality of low turbidity, but instead it is intended to protect downstream filter systems from intermittent peaks of high solids concentrations, effectively shutting off the water supply as the filter medium clogs. These systems require immediate attention by operating personnel or the flow of water to the final filters will cease, and the resulting loss of supernatant in a slow sand filter may dry out the *schmutzdecke* and drastically reduce its filtration cycle.

VERTICAL ROUGHING FILTERS

Vertical-flow filtration can be used downstream of a dynamic filter to further increase solids removal and protect the slow sand filter from clogging. Modern designs use a minimum of three filter compartments; each subsequent compartment contains smaller sizes of gravel. Media size range from average diameters of 20 mm in the first compartment to 4 mm in the last, and they can operate in either upflow or downflow configurations (Fig. 2). The gravel media are submerged under a supernatant water level, and filtration rates vary from 0.3 to 1.0 m/h. Filter velocities between 0.1 and 0.8 m/h were tested on natural river water in Colombia with an average of 50 nephelometric turbidity units (NTU), and removal efficiencies were between 55% and 45%. Challenge experiments were conducted with turbidities greater than 200 NTU, and removal efficiencies ranged from 70% to 90% (1).

Upflow roughing filters are also equipped with decreasing sizes of gravel in the direction of water flow and have drainage chambers in each filter compartment to facilitate regular cleaning. Filter freeboard must be sufficient to produce gravity hydraulic flushing capable of flow velocities greater than 90 m/h (2). Galvis et al. (3) reported 97% suspended solids removal and 80% turbidity removal for raw water with average suspended solids of 100 mg/L and turbidity of 95 NTU.

HORIZONTAL ROUGHING FILTERS

Horizontal-flow roughing filtration has been a particularly successful pretreatment method for suspended solids removal in rural areas of developing nations that have limited availability of skilled operators (4). Raw water enters the filter by an intake weir and flows horizontally through decreasing sizes of media contained in separate

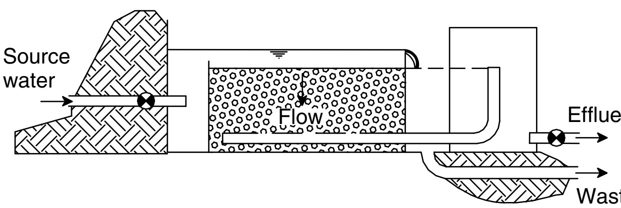

Figure 1. Example of a dynamic roughing filter design.

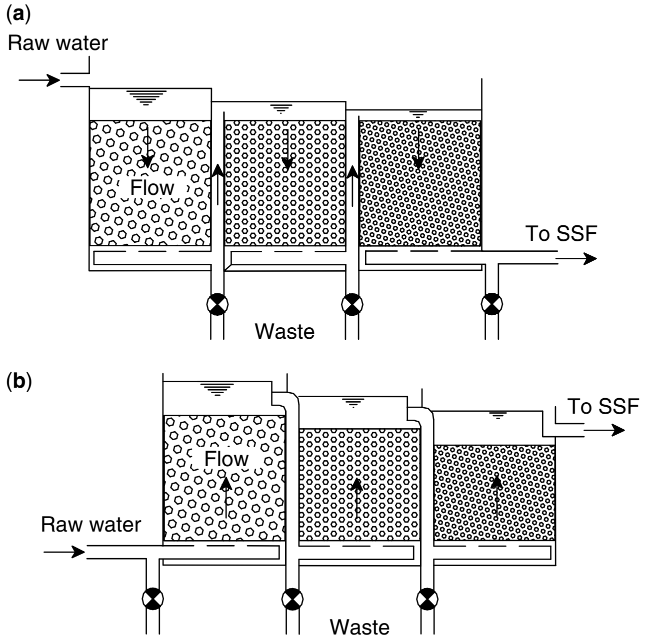

Figure 2. Schematic of downflow (**a**) and upflow (**b**) vertical roughing filters.

compartments (Fig. 3). Flow velocities range from 0.3 to 1.5 m/h, and the media size range and removal efficiencies are similar to vertical-flow filtration. Overall filter length ranges from 5 to 7 m with a filter compartment length ratio of 3:2:1. Typical media depth designs are less than 1 m to facilitate ease of media replacement, and drainage chambers are provided with the same freeboard and flushing velocities as vertical filters. The solids storage capacity for horizontal roughing filters is very large, and adequate filter service life can be years before clogging necessitates media removal (5).

DIRECT ROUGHING FILTRATION

The addition of coagulants to roughing filter influent is regarded as direct roughing filtration. Horizontal roughing filter coagulant injection of 1 mg (Al^{3+})/L as alum decreased raw water influent from 200 NTU to less than 3 NTU at filtration rates of 5–7 m/h (6). Coagulants cause roughing filters to clog much faster than normal (3–5 days), and cleaning process modifications must be designed before this type of treatment can be used reliably in the field. The addition of calcite limestone media in the first compartment of a horizontal roughing filter also aids coagulation, while maintaining the passive nature of the filter system. Source water with 150 NTU clay turbidity was reduced to less than 1 NTU by a limestone-amended multistage filter, and the effluent corrosion control characteristics of neutral pH source water were also enhanced by limestone roughing filter amendments (7).

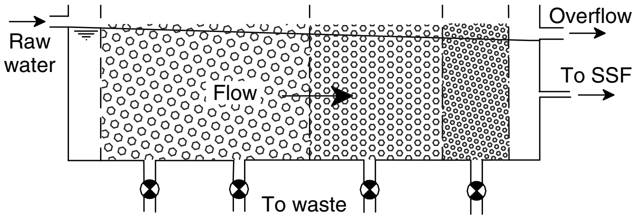

Figure 3. Horizontal flow roughing filter.

SUMMARY

Slow sand filtration is still a viable method of water treatment most suitable for raw water sources low in turbidity and suspended solids. Roughing filter pretreatment prior to slow sand filtration (multistage filtration) has been shown to be an efficient and effective drinking water treatment technique for source water with high turbidity, organic matter, and suspended solids.

BIBLIOGRAPHY

1. Galvis, G., Fernandez, J., and Visscher, J. (1993). Comparative study of different pre-treatment alternatives. *Aqua* **42**(6): 337–346.
2. Wolters, H., Smet, J., and Galvis, G. (1989). Upflow roughing filtration. *Pretreatment Methods for Community Water Supply: An Overview of Techniques and Present Experience*. IRC International Water Centre, The Hague, the Netherlands.
3. Galvis, G., Visscher, J., Fernandez, J., and Beron, F. (1996). *Pre-Treatment Alternatives for Drinking Water Supply Systems*. IRC International Water and Sanitation Centre, The Hague, Netherlands.
4. Wegelin, M. (1996). *Surface Water Treatment by Roughing Filters*. Swiss Centre for Development Cooperation in Technology and Management, Gallen, Switzerland.
5. Collins, M., Westersund, C., Cole, J., and Roccaro, J. (1994). *Evaluation of Roughing Filtration Design Variables*. AWWARF and AWWA, Denver, CO, pp. 77–88.
6. Ahsan, T., Alaerts, G., and Buiteman, J. (1996). Direct horizontal-flow roughing filtration. Part 1: Optimization of process parameters. *Aqua* **45**(5): 262–271.
7. Rooklidge, S. and Ketchum, L. (2002). Calcite-amended horizontal roughing filtration for clay turbidity removal. *J. Water Supply Res. Technol.-Aqua* **51**(6): 333–342.

MULTISTAGE FILTRATION: AN INNOVATIVE WATER TREATMENT TECHNOLOGY

Dick de Jong
Jo Smet
IRC International Water and
Sanitation Centre
Delft, The Netherlands

The technology of multiple stage filtration (MSF) presented here is a combination of coarse gravel prefiltration and slow sand filtration (SSF). This combination allows treating of surface water that has considerable levels of contamination, well above the levels that can be treated by SSF alone. To improve the performance of MSF, other treatment processes such as simple sedimentation, sand traps, and screens can precede MSF technology. Wherever possible, terminal disinfection needs to be included as a safety barrier after MSF.

MSF retains the advantages of SSF in that it is a robust and reliable treatment method that can be maintained by operators who have low levels of formal education. It is much better suited than conventional chemical water treatment to conditions in rural communities and small and medium size municipalities in the South as well as in more remote areas in the North. The article provides a summary description of the components of MSF systems and the findings of recent research. It gives an overview of cost implications and ends with a selection guide.

SELECTION OF MULTISTAGE FILTRATION CRITERIA

Factors for selecting the *best* surface water treatment include the quality of the water sources, the availability of resources (construction material, sand, gravel, chemicals, financial resources, and land), the socioeconomic and cultural conditions of the community, and the level of institutional support. Careful selection of the water treatment technology is crucial to ensuring sustained good quality water at an acceptable cost to users.

MSF alternatives include the dynamic roughing filtration (DyRF) step and SSF alone or in combination with other pretreatment steps. The most common are up-flow roughing filters in layers (URFL) or in series of two or three units (URFS2, URFS3) (see Fig. 1), down-flow roughing filters (DRF), or horizontal-flow roughing filters (HRF). A comparative study showed that URF is technically and economically preferable to DRF and HRF, although these also present good removal efficiencies (1).

The MSF alternatives are similar in their environmental impact and complexity of operation and maintenance. This implies that one can choose among them using three main selection criteria: the sanitary risk of the water source, the treatment efficiency, and the overall cost. Conventional treatment options such as rapid sand filtration of chemically coagulated water often face problems in managing the systems and availability of chemical products. MSF technology, on the other hand, is well received by communities because it is much simpler to administer, operate, and maintain compared to other treatment systems.

Slow Sand Filtration, the Heart of MSF

Slow sand filtration is the heart of MSF technology. It has been recognized as a simple, reliable, and efficient treatment technology and the most effective as a single unit treatment process. An SSF produces an effluent low in turbidity, free of impurities, and even more important, virtually free of bacteria, enteroviruses, and protozoa. Table 1 presents an overview of the removal efficiencies that may be obtained from SSF technology for different parameters of sanitary importance. These efficiencies have been reported for SSF units operating in temperature zones above 5 °C, at flow velocities between 0.04 and 0.20 m^3/m^2h, filter medium depths above 0.5 m, and effective sand grain sizes between 0.15 and 0.30 mm.

The use of SSF technology often has been determined by the availability of good raw quality water resources, as in different areas of Europe and the United State. However, SSF alone is not a feasible solution for more contaminated surface water sources. It is not a panacea for all water quality problems under all circumstances. In general, two situations can be distinguished under which SSF presents limitations.

- Levels of contamination in the raw water that supersede the treatment capacity to comply with the existing norms or that may result in short filter runs. This relates particularly to high quantities of suspended solids, high coliform counts, and large quantities of algae.
- Conditions that inhibit or reduce the efficiency of the treatment process such as low temperature, low nutrient content, and low oxygen level.

Coarse Gravel Filtration

For more contaminated surface water (i.e., turbidity levels higher then 20NTU), it is necessary to apply good pretreatment modalities to condition the water before it enters the SSF units. Commonly used pretreatment systems include simple sedimentation, microscreening, roughing filtration in gravel filters, and, more recently in Europe, also ozone and activated carbon treatment to reduce the levels of organic material.

In the last few decades, the prospect of using SSF technology in rural communities and small/medium size municipalities has increased because of good research

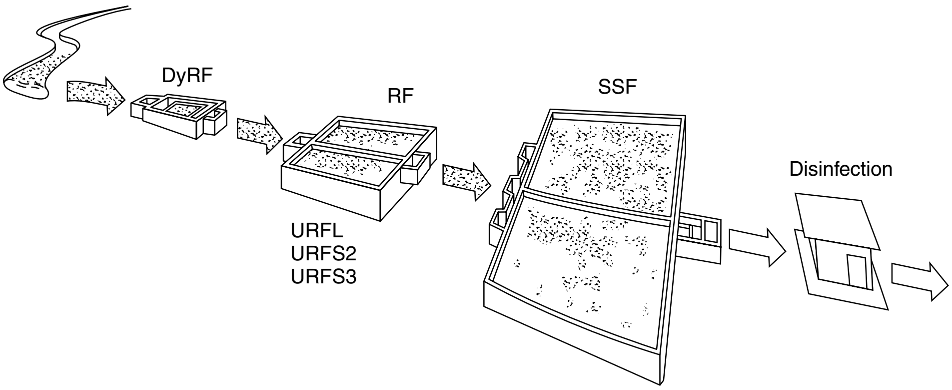

Figure 1. Components of a multi-stage filtration system.

Table 1. Typical Removal Efficiencies of Conventional SSF Units[a]

Water Quality Parameter	Effluent or Removal Efficiency	Comments
Turbidity	<1 NTU	The level of turbidity and the nature and distribution of particles affect treatment capacity and efficiency
Enterobacteria	90 to 99.9%	Affected by temperature, filtration rate, size, uniformity and depth of sand bed, cleaning operation
Enteroviruses and Giardia/Amoeba cysts	99 to 99.99%	High removal efficiencies, even directly after cleaning (removal of the biomembrane (schmutzdecke))
Cercariae	100%	In good operation and maintenance conditions, virtually complete removal is obtained
True color	25 to 30%	Color associated with organic material and humic acids
TOC	<15–25%	TOC, total organic carbon
AOC	<50%	AOC, biodegradable organic carbon
THM precursors	<25%	Precursors of trihalomethanes
Iron, manganese	30 to 90%	Fe^{2+} levels >1 mg/L reduce filter runs

[a] Reference 1.

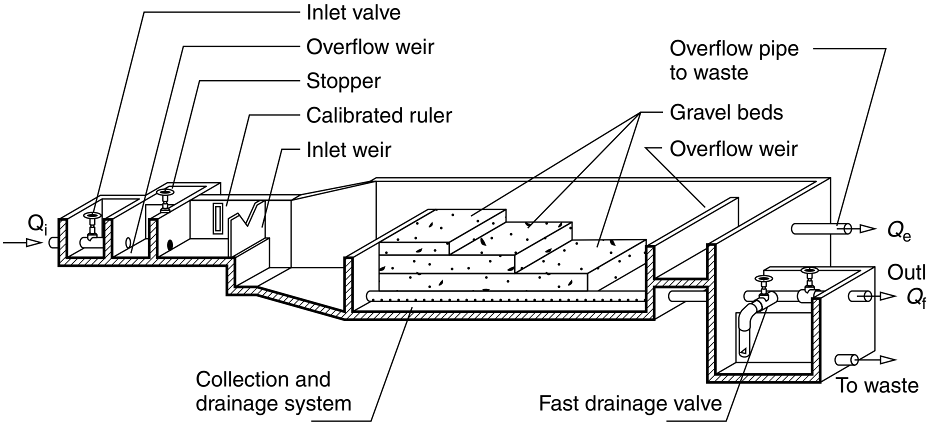

Figure 2. Dynamic roughing filter.

results in roughing filtration (coarse gravel filtration). An important breakthrough in research on pretreatment technology has come from comparative studies of different roughing filtration systems in Colombia (1). Both field testing and full-scale plants were included. The research concluded that for the conditions in Colombia, the combination of DyRF and URF proved to be the best pretreatment option prior to SSF.

PUTTING MSF IN PERSPECTIVE

MSF has advantages similar to SSF in ease of operation and maintenance and production of good water quality. Many rural communities and small/medium size municipalities face problems with conventional surface water treatment using chemicals. MSF technology is then a sustainable and effective treatment alternative (Table 2).

DESCRIPTION OF THE COMPONENTS OF ROUGHING FILTERS

MSF may consist of two or three main filtration stages, depending on the level of contamination of the source. The option with three components, DyRF, URF, and SSF, is indicated in (Fig. 1). In principle, DyRF focuses on reducing suspended solids, particularly removing larger particles and on protecting the subsequent units from high loads of suspended solids. The removal of smaller particles and microorganisms takes place primarily in the next treatment steps. URF can be avoided if good quality surface water is available.

Dynamic Roughing Filters (DyRF) contain a shallow layer of fine gravel in their upper part on top of a layer of coarse gravel that covers the underdrains (total height of the filter bed is about 0.6 m) (see Fig. 2). Water enters the unit and passes through a bed of fine gravel on top of coarse gravel to the drainage system at a filtration rate of some 2 to 3 m^3/m^2h. Moderate levels of suspended solids in the source gradually clog the DyRF. For high suspended solids, clogging happens very quickly. When the gravel bed is completely blocked, raw water will flow over the clogged surface area to waste. In this way, the DyRF acts as a safety valve, and subsequent treatment steps are protected from receiving high suspended solid loads.

Up-flow roughing filters consist of one or more compartments filled with gravel (diameter from 3–24 mm) that reduces in size in the direction of flow. Filtration rates usually range between 0.3 and 0.6 m^3/m^2h (see Fig. 3). A drainage system placed on the bottom of the structure distributes the flow during the filtration period or drains the gravel layers during periods of cleaning and discharges

Table 2. Summary of Considerations for MSF Treatment[a]

Issue	Comment Concerning MSF Treatment
Quality of treated water	It is a good alternative for improving the physical, chemical, and bacteriological quality of the water. In many areas and particularly those that have a less developed infrastructure, MSF may be the only feasible treatment option.
Ease of construction	The relatively simple design facilitates the use of local materials and local skilled manpower. There is no need for special equipment.
Construction cost	Construction in local materials and with local labor reduces the cost. Usually, there is no need for imported materials.
Ease of operation and maintenance	After a short period of training, local operators who have a minimum of formal education can operate and maintain the system.
Cost of operation and maintenance	The cost of operation and maintenance and the requirements in electrical energy are much less than for other systems. There is no need for chemical coagulation products.
Reliability	A low risk of mechanical problems or problems related to changes in raw water quality; in the majority of cases, these can be accommodated without interrupting service.
Cleaning	The cleaning process is simple although laborious but almost always involves low cost, as in most 'potential' countries, labor is relatively cheap.
Requirements of surface area	A conventional chemical and RSF water treatment may require surface areas comparable to an MSF treatment plant.
MSF, the panacea?	There are levels of contamination that interfere with the treatment or that surpass the capacity of MSF.

[a] Reference 1.

Table 3. Performance of MSF Systems in Colombia

Performance Indicator	Raw Water Quality	Efficiency Range	Remarks
Turbidity (NTU)	2–15 NTU and peaks over 180 NTU	83–97% <1 NTU in 65–98% of all samples <5 NTU in 98% of all samples	Lower figure was measured in older MSF system.
Suspended solids, mg/L	2–23 and peaks over 390	0.2–2.8 mg/L	
Fecal coliform density, CFU/100 mL	300–60,000	<25 CFU/100 mL in 97% of all samples	
True color, TCU	5–24 and peaks over 180	<15 TCU in 98% of all samples	
Filter run periods		46–178 days	A short filter run due to algae

the water through the drainage system. The impurities are retained at the bottom, and therefore, it is easy to drain. Two alternatives can be identified: up-flow roughing filters in layers (URFL) when the gravel layers of different size are installed in the same unit (total filter bed height 1.2 m) and up-flow roughing filters in series (URFS) when the gravel layers are installed in two or three different units (each filter bed is about 1.2 m high). (See Table 4.)

Downflow roughing filters in series (DRFS) consist of three compartments of gravel reducing in size (2). The performance of the DRFS is similar to that of the URFS in removal efficiency. Maintenance and cleaning are more difficult than for the URF because the sludge tends to accumulate on the surface and in the deeper layers of the bed.

Horizontal roughing filters (HRF) consist of three compartments separated by perforated walls (3). Their performance is similar to that of the URF, but construction costs are higher (1).

PERFORMANCE OF MSF SYSTEMS

MSF performance data (Table 3) are based on seven community-managed MSF systems in Colombia that were monitored for a 7 years. The systems receive water from catchment areas of low or moderate levels of human intervention.

Based on these effluent water qualities, constant dose disinfection with chlorine, as suggested by WHO (4), functions effectively as a final safety barrier.

The findings show that MSF treatment can adapt itself to the type of raw water and the concentration of contamination. The systems have higher removal efficiencies for water that is higher in contamination. This implies that the barriers become more effective if the water to be treated has a higher risk and still can produce a water at a low sanitary risk level. The results on both pilot and full scale clearly indicate the great potential of MSF technology in reducing the physicochemical and bacteriological risk from surface water sources. Nevertheless, it also shows that the technology is not a panacea and has its limitations. Particularly, for high levels of contamination, MSF systems do not always produce water that meets the required quality, and then extra pretreatment steps may be needed.

MSF systems comply with the multibarrier concept: in general, all treatment steps (barriers) together achieve a removal efficiency that is sufficient to ensure low dose disinfection as a final and efficient safety barrier.

The technology is a reliable alternative for treating water of a low or moderate contamination level such

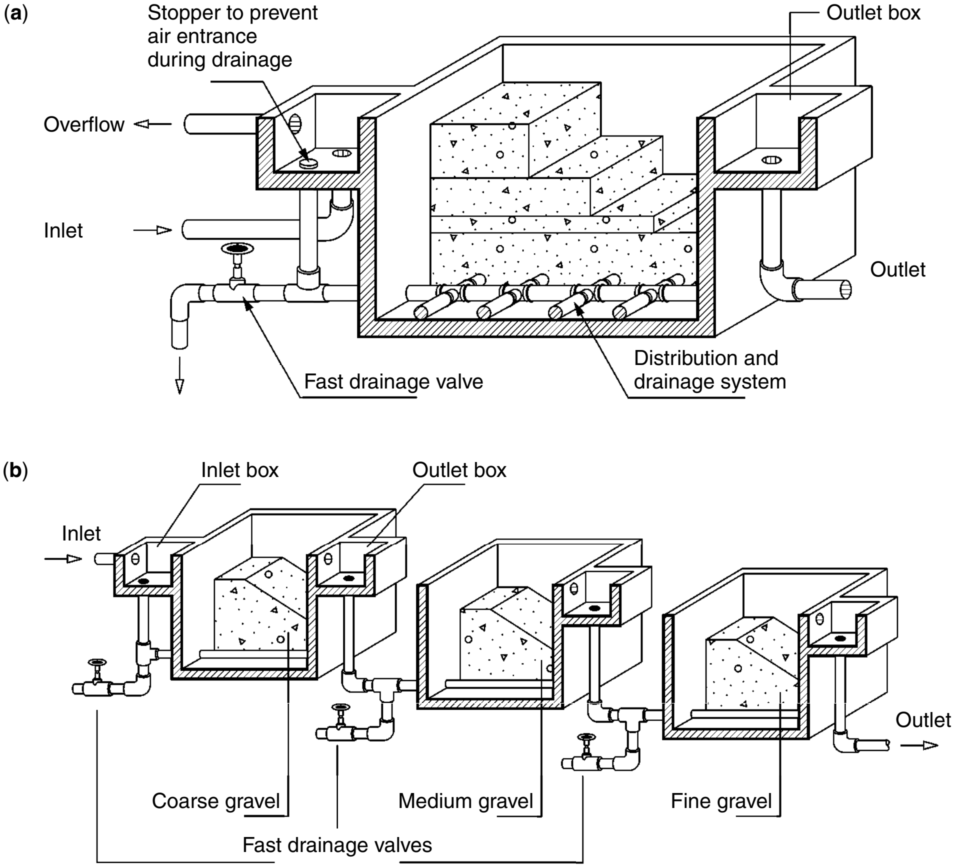

Figure 3. (**a**) Schematic design of an up-flow roughing filter in layers (URFL). (**b**) Schematic design of an up-flow filter in series (URFS).

Table 4. Design Criteria for the MSF Type Designs Used to Establish Construction Volumes for the Cost Model in (Fig. 3)

MSF Component Design Criteria	DyRF	URFL	URFS2	URFS3	SSF
Filtration rate (m^3/m^2h)	2	0.6	0.6	0.6	0.15
Depth of filter medium (m)	0.60	1.2	1.2[a]	1.2[a]	1.05
Filter box height (m)	0.80	1.4	1.4	1.4	2.15
Number of units in parallel	2	2–4	2–4	2–6	4–8
Number of units in series	1	1	2	3	1

[a] Per filter unit.

as received by full-scale plants. However, river water of higher contamination levels can present difficulties. For those conditions, it is essential to study the performance of the MSF system in a pilot plant, before deciding to construct a full-scale plant.

COST ISSUES

Construction costs vary by country. The economy of scale is limited, thus favoring relatively short design periods of some 10 years. Operating and maintenance costs have been estimated on the basis of experience with MSF systems in Colombia. In these plants that are operated under gravity supply, it was found that staff makes up 85% of the cost. Thus, the time involved in the different operating and maintenance tasks are a good basis for assessing the overall cost. The labor requirements for daily and periodic operation and maintenance of MSF systems range between 0.2 and 1.5 equivalent person days per day, depending on the size and composition of the systems. In addition 0.3 equivalent person days are needed for administration, including supervision of the plant operator, planning, customer relations, purchasing, payments, and meetings with sector organizations and local government.

For example, for a system of 10 L/s in Colombia that includes DyRF + URFS2 + SSF, the construction cost would amount to US$ 125,000. The total time required for operation and maintenance of this plant amounts to 1.6 PE (person equivalents per year), 1.0 PE for daily and periodic activities, 0.3 PE for exceptional activities, and 0.3 PE for administration. For 1996, at an average salary of US$150/ month and taking into account the 85% factor, operating and maintenance cost amounted to $(1.6\times150)/0.85 =$ US$ 276/ month.

TECHNOLOGY SELECTION GUIDE

On the basis of the previous considerations, the minimum treatment for surface water of *low sanitary risk*—but requiring more than just disinfection—is a combination of DyRF and SSF. Surface water of a *moderate risk level* requires a URFL or one of the URFS systems in combination with DyRF and SSF. For surface water sources of *high risk levels*, the options of URFS in two or three steps need to be considered in combination with DyRF and SSF.

The findings and design considerations are mainly based on research from Colombia. The biological treatment process involved in MSF technology is important but not easy to predict. In low temperature zones, the efficiencies are lower than indicated here. So, designing MSF under different climatic conditions needs a critical attitude toward design, function and performance. In regions with little experience with MSF technology and/or important uncertainties about the quality of the water sources, it is important to adopt conservative design parameters in the first MSF plants and preferably start with pilot plant experiments to review the functioning of MSF systems under local conditions. A joint learning program, as described by Visscher (5) is recommended. This type of program brings together communities, sector professionals, sector authorities, and other important institutions in the sector, including local universities and research institutions. Together, they review the situation and ensure that the technology can find roots in the new situation.

BIBLIOGRAPHY

1. Galvis, G., Latorre, J., and Visscher, J.T. (1998). *Multi-Stage Filtration, an Innovative Water Treatment Technology*. IRC Technical Paper Series no. 34, The Hague, The Netherlands, IRC International Water and Sanitation Centre; CINARA—Universidad del Valle, Cali, Colombia.
2. Pardon, M. (1987). *Consideraciones, Desarrollo y Evaluación de un Sistema de Tratamiento que Implementa la Filtración Gruesa de Flujo Vertical en Gravas*. Lima, Perú.
3. Wegelin, M. (1986). *Horizontal Flow Roughing Filtration (HRF) Design, Construction and Operation Manual IRCWD. Report No. 06/86.* Duebendorf, Switzerland.
4. WHO. (1996). *Guidelines for Drinking Water Quality: Health Criteria and other Supporting Information*, 2nd Edn. Vol. II. World Health Organization, Geneva, Switzerland.
5. Visscher, J.T. (Ed.). (1997). *Technology Transfer in the Water and Sanitation Sector: A Learning Experience from Colombia*. Technical Paper Series no. 32, The Hague, The Netherlands and CINARA, Cali, Colombia.

PARTICULATE MATTER REMOVAL BY FILTRATION AND SEDIMENTATION

RASHEED AHMAD
Khafra Engineering
Consultants
Atlanta, Georgia

The settling efficiency for an ideal condition is independent of the depth and dependent on the tank plan or surface area. In reality, depth is important because it can affect flow stability if it is large and scouring if it is small. The basin depth affects flow-through velocity and allows greater opportunity for flocculant particle contact. Conventional settling facilities at larger water treatment plants are often long, narrow (4 or 5 to 1 length to width ratio) rectangular basins that have theoretical detention times in the range of 1.5 to 3 hours at the design flow rate. Most regulatory agencies specify a minimum detention time and a maximum surface overflow rate. Typical sedimentation surface loading rates for long, rectangular tanks and circular tanks using alum coagulation vary from 500 to 1200 gallon per day per square feet (20 to 48 cubic meters per day per square meter). Regulatory agencies sometimes stipulate that their weir rates should not exceed around 20,000 gallons per day per linear foot (250 cubic meters per day per meter) of weir. Flow-through velocities of 2 to 4 ft/min (0.6 to 1.2 m/min) are usually acceptable for basin depths of 7 to 14 ft (2.1 to 4.3 m).

Particulate matter will settle out of a suspension in one of four different ways, depending on the concentration of the suspension and the flocculating properties of the particles. The various settling regimes of particles are commonly referred to as types 1 to 4. In type 1 settling, also called discrete particle settling, particles have little tendency to flocculate upon contact with each other in a dilute suspension. If the particles flocculate, the settling regime in a dilute suspension is identified as type 2 or flocculant settling. In hindered, zone, or type 3 settling, the particle concentration causes interparticle effects, which might include flocculation, to the extent that the rate of settling is a function of solids concentration. Compression or type 4 settling develops under the layers of zone settling. The rate of compression depends on time and the force caused by the weight of solids above. In sedimentation, particulate matter removal is largely governed by types 1 and 2 settling. However, types 3 and 4 regimes are dominant in clarification and thickening processes.

High-rate settlers such as tube or plate settlers are often added to settling basins to improve their efficiency, especially if flows are to be increased beyond original design conditions. The shallow settling depths and the large surface area provided by tube or plate settlers permit effective sedimentation at detention times of several minutes versus several hours in conventional sedimentation basins. The use of tube or plate settlers in new plant construction minimizes settling basin costs by minimizing the basin size necessary to attain a desired level of treatment. They are generally designed to accept overflow rates ranging from 1400 to 4000 gallons per day per sq ft (55 to 160 cubic meters per day per square meter).

In a conventional water treatment process train, filtration follows sedimentation. Filtration is one of the unit processes used in the production of potable water. Particulate matter that is removed by this process includes micro-organisms (bacteria, viruses, and protozoan cysts), clay and silt particles, colloidal and precipitated humic substances and other organic particulates from natural decay of vegetation, precipitates of aluminum or iron coagulation processes, calcium carbonate and magnesium hydroxide precipitates from lime softening, and iron and manganese precipitates.

A number of different types of filters are used in water filtration, and they are described by various classification schemes. Most common filtration technologies used in water supply systems are rapid sand filtration, slow sand filtration, package plants, diatomaceous earth filtration (precoat filtration), membrane filters, and cartridge filters. High-rate granular filters or rapid sand filters are most widely used in conventional treatment and direct filtration. Slow sand filters have been used in water treatment for several decades, but their use in recent decades has declined because of the inherent low filtration rate, large area requirement, and difficulties in filter cleaning. Membrane and cartridge filtration systems are considered emerging technologies. In recent years, some water treatment plants have replaced granular filters with membrane filters. Package plants, slow sand, diatomaceous earth, membrane, and cartridge filters are considered best suited for small water systems [less than 0.44 cubic meter per s (10 MGD)].

Granular medium filtration is a water treatment process that uses a porous medium through which water passes to remove particulates or suspended solids. For granular medium filtration to be effective, pretreatment of source water is required. Chemical destabilization is an essential prerequisite for effective filtration. Chemicals used for particle destabilization are limited primarily to metal salts or cationic polymers as primary coagulants. Pretreatment may also include aeration or introducing an oxidant if water treatment aims to remove iron or manganese. Sometimes a filter aid polymer is added in the influent to the filter to improve particle capture efficiency.

The U.S. Environmental Protection Agency's (USEPA) Surface Water Treatment Rule (SWTR), promulgated on June 29, 1989 (Federal Register 40 CFR Parts 140 and 141), requires community water systems to disinfect all surface waters and requires filtration for most surface water sources. The 1998 Interim Enhanced Surface Water Treatment Rule (IESWTR) requires that surface water or groundwater systems under the influence of surface waters must produce a turbidity less than or equal to 0.3 NTU in 95% of the measurements taken each month. This rule was later extended to smaller systems as part of the Long Term Enhanced Surface Water Treatment Rule (LTESWTR).

Filtration by granular media consists of three principal mechanisms: (1) transport, (2) attachment, and (3) detachment. *Transport* mechanisms move a particle into and through a filter pore so that it comes very close to the surface of the filter medium or existing deposits where *attachment* mechanisms retain the suspended particle in contact with the medium's surface or with previously deposited solids. *Detachment* mechanisms result from the hydrodynamic forces of flow acting so that a certain portion of the previously attached particles, less strongly adhered to others, is detached from the filter medium or previous deposits and is carried further, deep into or through the filter. Important transport mechanisms include screening, interception, inertial forces, sedimentation, diffusion, and hydrodynamic forces. Attachment of particles to the media surfaces is generally governed by physicochemical and molecular forces. Detachment is caused by the impact of arriving particles on unstable deposits and by hydraulic shear stresses without the influence of arriving particles.

The most common types of filter media used in granular bed filters are silica sand, anthracite coal, and garnet. These may be used alone or in dual- or triple-media combinations. Granular-activated carbon (GAC), another type of medium, has been used to reduce taste and odor in granular beds that serve for both filtration and adsorption. A number of properties of a filter medium affect filtration performance and are important in characterizing the medium. Important media properties include size, shape, density, and hardness. The porosity of the granular bed formed by the grains is also an important governing factor in filtration performance. Selection of filter media type and characteristics is based on a number of design decisions concerning source water quality, pretreatment, and desired filtered water quality. Filter media cleaning requirements and underdrain system options depend on filter configuration and filter media selected. Pilot plant studies using alternative filter media and filtration rates can determine the most effective and efficient medium for a particular water.

A slow sand filter generally consists of a watertight basin containing a layer of sand over a layer of gravel. It is operated at very low filtration rates without coagulation in pretreatment. Smaller grain size and lower filtration rate result in removing the solids almost entirely in a thin layer on the top of the sand bed. This layer, composed of dirt and living and dead micro- and macro-organisms from the water, is called *schmutzdecke* or dirty skin. Most of the solid removal in a slow sand filter takes place in the *schmutzdecke* layer. The dominant removal mechanisms are both physical and biological. A typical filter cycle length may vary from 1 to 6 months (or longer) depending on source water quality and filtration rate. Filtration rates range from 0.016 to 0.16 gpm/sq ft (0.04 to 0.40 m/h). Effective sand sizes range from 0.15 to 0.40 mm, sand uniformity coefficients from 1.5 to 3.6, and initial bed depths from 1.5 to 5.0 ft. The sand is supported on graded gravel 6 to 36 inches deep. Slow sand filters are cleaned by scraping the *schmutzdecke* and a small amount of sand depth. Usually, scraping is done manually but, in some cases, mechanically. The scraped sand is cleaned hydraulically and stockpiled for later reuse. Slow sand filters are very effective in removing turbidity, viruses, and bacteria.

Rapid granular bed filtration consists of passing pretreated water through a granular bed at rates ranging from 2 to 10 gpm/sq ft (5 to 25 m/h). Flow is typically downward through the bed. During the filtration cycle, solids are removed from the water and accumulate within the voids and on the top surface of the filter medium. This accumulation results in clogging and a gradual increase in head loss. When head loss becomes excessive (generally 8–9 ft), filtrate quality begins to deteriorate, and/or the filter cycle reaches a predetermined time limit (usually 3 or 4 days), then rapid filters need cleaning. Dirty/clogged filters may be cleared by scouring the clogged portion or by reversing the flow through the bed. This application of expanding and washing out trapped particles is called backwashing. Backwashing by water fluidization is frequently assisted by a surface wash or an air scour. The most effective backwash is achieved by simultaneous air scour and subfluidization water backwash. Typical backwash rates range from 15 to 23 gpm/sq ft depending on media sizes. The bed expansion varies from 20% to 50%. The backwash usually lasts from 5 to 15 minutes. The air-scour rate varies from 2 to 4 cu ft/min/sq ft (0.6 to 1.2 cu meter/min/sq meter). Underdrain systems are used to support the filter medium, collect filtered water (in downflow filters), and distribute backwash water (and air, if employed). The efficiency of a filter is more sensitive to changes in the filtration rate rather than in the actual rate. Therefore, filtration process control is critical to successful operation. Two basic modes of gravity filter control are commonly found, constant rate and declining rate. Common filter problems include inadequate pretreatment or filter washing, gravel bed upset, air binding, restart after shut down, and filter media replacement. These problems lead to operational difficulties like dirty filter media, mud balls, and mineral deposits. Rapid gravity filters that use coagulation and sedimentation normally have 96.9% to 99.9% bacterial and 99.0% virus removal efficiencies. It is not uncommon for a well-operated filter to produce an effluent turbidity of 0.1 NTU (nephelometric turbidity units).

Precoat filters use a thin layer of very fine material such as diatomaceous earth as a filter medium. In precoat filtration, water to be filtered is passed through a uniform layer of the filter medium that has been deposited (precoated) on a *septum*, a permeable material that supports the filter medium. The filter operation is composed of three cycles: (1) precoating, (2) filtering, and (3) cleaning. When the pressure drop through the filter system reaches a point where continued filtration is impractical, the filtration process is stopped. The filter medium and collected dirt are washed off the septum, a new precoat of filter medium is applied, and filtration continues. Diatomaceous filters are effective in removing cysts, algae, and asbestos. These filters are especially suited for low influent turbidity water and are very effective against *Giardia* cysts.

READING LIST

ASCE and AWWA. (1998). *Water Treatment Plant Design*, 3rd Edn. McGraw-Hill, Inc., New York, NY.

AWWARF. (1998). *Treatment Process Selection for Particle Removal*, American Water Works Association Research Foundation, Denver, CO.

AWWA. (1999). *Water Quality and Treatment*, 5th Edn. McGraw-Hill, Inc., New York, NY.

Hudson, H.E. (1981). *Water Clarification Processes*. Van Nostrand Reinhold Company, New York, NY.

Kawamura, S. (2000). *Integrated Design of Water Treatment Facilities*, John Wiley & Sons, New York, NY.

Qasim, S.R., Motley, E.M., and Zhu, G. (2000). *Water Works Engineering*. Prentice-Hall, Upper Saddle River, NJ.

USEPA. (1990). *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*. Office of Drinking Water, Center for Research Information, Cincinnati, OH.

FILTRATION WATER TREATMENT

Rasheed Ahmad
Khafra Engineering
Consultants
Atlanta, Georgia

All filtration technologies have the same basic objectives: to produce a filtrate that is clear and contains very few particles and to provide a medium-cleaning process so that operation may be continued with the same medium or by replacing the old medium. Particulates that are removed by water filtration process include micro-organisms (bacteria, viruses, and protozoan cysts), clay and silt particles, colloidal and precipitated humic substances and other organic particulates from natural decay of vegetation, precipitates of an aluminum or iron coagulation process, calcium carbonate and magnesium hydroxide precipitates from lime softening, and iron and manganese precipitates.

Rapid gravity filtration consists of passing pretreated water through a granular bed at rates ranging from 2 to 10 gpm/sq ft (5 to 25 m/h) (Fig. 1). Flow is typically downward through the bed. During the filtration cycle, solids are removed from the water and accumulate within the voids and on the top surface of the filter medium. This accumulation results in clogging and a gradual increase in head loss. When head loss becomes excessive (generally 8–9 ft), filtrate quality begins to deteriorate,

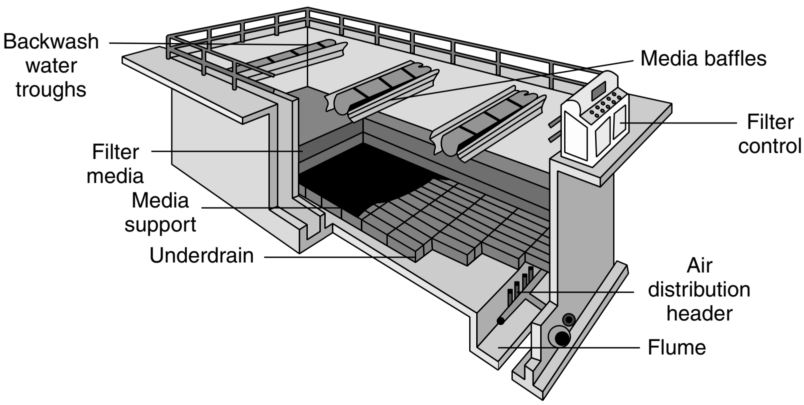

Figure 1. Rapid gravity filter (*Source*: F.B. Leopold Company).

and/or the filter cycle reaches a predetermined time limit (usually 3 or 4 days), then rapid filters need cleaning. Dirty/clogged filters may be cleaned by scouring the clogged portion or by reversing the flow through the bed. This application of expanding and washing out trapped particles is called backwashing. Backwashing by water fluidization is frequently assisted by a surface wash or an air scour. The most effective backwash is achieved by a simultaneous air scour and subfluidization water backwash. Typical backwash rates range from 15 to 23 gpm/sq ft depending on media sizes. The bed expansion varies from 20% to 50%. The backwash duration usually lasts from 5 to 15 minutes. The air-scour rate varies from 2 to 4 cu ft/min/sq ft (0.6 to 1.2 cu meter/min/sq meter).

For rapid gravity filtration to be effective, the source water must be pretreated. Chemical destabilization is an essential prerequisite for effective filtration. Chemicals used for particle destabilization are limited primarily to metal salts or cationic polymers as primary coagulants. Pretreatment may also include aeration or introducing an oxidant if water treatment aims to remove iron or manganese. Sometimes a filter aid polymer is added in the influent to the filter to improve particle capture efficiency. Granular bed filtration consists of three principal mechanisms: (1) transport, (2) attachment, and (3) detachment. *Transport* mechanisms move a particle into and through a filter pore so that it comes very close to the surface of the filter medium or existing deposits where *attachment* mechanisms retain the suspended particle in contact with the medium's surface or with previously deposited solids.

Detachment mechanisms result from the hydrodynamic forces of flow acting so that a certain portion of the previously attached particles, less strongly adhered to others, is detached from the filter medium or previous deposits and carried further, deep into or through the filter.

Important transport mechanisms include screening, interception, inertial forces, sedimentation, diffusion, and hydrodynamic forces. Physicochemical and molecular forces generally govern attachment of particles to media surfaces. Detachment is caused by the impact of arriving particles on unstable deposits and by hydraulic shear stresses without the influence of arriving particles. A number of properties of a filter medium are important in filtration performance and in characterizing the medium. Important media properties include size, shape, density, and hardness. The efficiency of a filter is more sensitive to changes in the filtration rate rather than in the actual rate. Therefore, filtration process control is critical to successful operation.

Several different granular types of filters are classified by various schemes. Filters can be classified by the type of medium (single media, dual media), hydraulic arrangement (gravity or pressure), rate of filtration (rapid or slow), and depth of solids removal (deep or cake). The most common types of filter media used in granular bed filters are silica sand, anthracite coal, and garnet. These may be used alone or in dual- or triple-media combinations. Granular-activated carbon (GAC), another type of medium, has been used in granular beds that serve for both filtration and adsorption to reduce taste and odor.

In recent years, direct filtration has received considerable attention and application in the treating of drinking water. Direct filtration is not preceded by sedimentation. It offers several advantages over conventional treatment of "good-quality" surface water. Because there is no sedimentation process and a lower coagulant dose, the capital and operating costs are lower compared with those of conventional treatment. Sludge volumes are lower, which results in lower chemical costs for sludge treatment and disposal. Because of increased interest in ozone application in potable water, biological filtration has received noticeable attention and application in recent years. Granular filters become biologically active when ozone is used as a preoxidant. Ozone reacts with organic compounds in water and forms several ozonation byproducts such as aldehydes, carboxylic acids, and keto acids, which are relatively easy to biodegrade. By operating filters in the biological mode, these ozonation byproducts can be effectively removed and in turn reduce the potential for bacterial regrowth in distribution systems and formation of chlorination byproducts.

A slow sand filter generally consists of a watertight basin containing a layer of sand over a layer of gravel (Fig. 2). It is operated at very low filtration rates without coagulation in pretreatment. A smaller grain size and lower filtration rate result in solids removal almost entirely in a thin layer on the top of the sand bed. This layer, composed of dirt and living and dead micro- and macro-organisms from the water, is called *schmutzdecke* or dirty skin. Most of the solid removals in slow sand filter take place in the *schmutzdecke* layer. The dominant removal mechanisms are both physical and biological. A typical filter cycle may vary from 1 to 6 months (or longer) depending on the source water quality and the filtration rate. Filtration rates range from 0.016 to 0.16 gpm/sq ft (0.04 to 0.40 m/h). Effective sand sizes range from 0.15

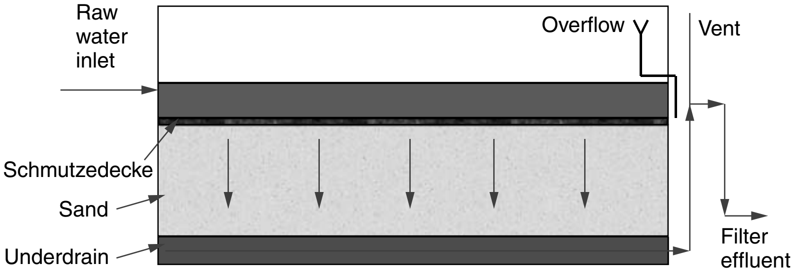

Figure 2. Conventional slow sand filter.

to 0.40 mm, sand uniformity coefficients from 1.5 to 3.6, and initial bed depths from 1.5 to 5.0 ft. The sand is supported on graded gravel 6 to 36 inches deep. Slow sand filters are cleaned by scraping the *schmutzdecke* and a small amount of sand depth. Usually, scraping is done manually but, in some cases, mechanically. The scraped sand is cleaned hydraulically and stockpiled for later reuse.

Rapid rate pressure filters are similar to gravity rapid rate filters, except that the flow enters and exits the filter under pressure (Fig. 3). The filter medium is contained in a steel pressure vessel. The pressure vessel can be vertical or horizontal. Proper backwashing of a pressure filter is more difficult compared with a gravity filter because the filter medium is not conveniently visible to the operator during the backwash operation. Pressure filters tend to be used in small water systems. Many pressure filters are used in industrial water and wastewater filtration. They are also used widely in swimming pool filtration.

Diatomaceous earth filtration, precoat or diatomite, filtration is another water filtration process that works under pressure. It consists of a layer of diatomaceous earth used as a filter medium approximately one-eighth inch thick placed on a septum or filter element. The septum may be placed in a pressure vessel or operated under a vacuum in an open vessel. As the water passes through the filter medium and septum, the suspended particles are captured and deposited. The majority of particles removed by the filters are strained at the surface layer of the filter medium, and some are trapped within the layer. As the filter cycle proceeds, additional filter medium called *body feed* is regularly metered into the influent water in proportion to the solids being removed. Ultimately, a gradually increasing pressure drop through the filter system reaches the point where continued filtration is impractical. The forward filtration process is stopped, the filter medium and collected dirt are washed off the septum, a new precoat of filter medium is applied, and the filtration process continues. Diatomaceous earth filters are widely used in industrial filtration and in swimming pool filtration. They have also been used in municipal potable water treatment, primarily in direct, in-line filtration of high-quality surface water (turbidity 10 NTU or less and acceptable color), and in filtering iron and manganese from groundwater after appropriate pretreatment to precipitate these contaminants. Precoat filters are simple to operate and are effective in removing cysts, algae, and asbestos.

Figure 3. Pressure filter (*Source*: U.S. Filter).

Membrane filtration represents an important set of processes for drinking water treatment. A membrane is a thin layer of natural or synthetic material that can separate substances when a driving force is applied across the membrane. Membranes used for water treatment are commonly made of synthetic organic polymers. Membrane processes for potable water treatment are reverse osmosis (RO), nanofiltration (NF), electrodialysis (ED), ultrafiltration (UF), and microfiltration (MF). Reverse osmosis is used primarily to remove salts from brackish water or seawater. Nanofiltration is used to soften freshwaters and remove disinfection byproduct (DBP) precursors. Electrodialysis is used to demineralize brackish water and seawater and to soften freshwater. Ultrafiltration and microfiltration are used to remove turbidity, pathogens, and other particulates from freshwater (Fig. 4). Membranes are normally classified by solute exclusion size, which is sometimes referred to as *pore size*. Membrane filtration works by passing water at high-pressure through a thin membrane in the form of hollow-fiber or spiral-wound composite sheets. The contaminants are retained on the high pressure side and frequently must be cleaned by reversing the flow and flushing the waste. Periodic chemical cleaning may be required to remove persistent contaminants. Membrane assemblies are contained in pressure vessels or cartridges. Low-pressure membranes in the form of either ultrafiltration (UF) or microfiltration (MF) have become more economical in capital and operating costs and have received increased attention in drinking water application. Fouling of the membrane is the major problem preventing widespread application of this technology.

Cartridge filters traditionally have been applied for point-of-use systems and for pretreatment prior to membrane treatment systems. Composed of membrane, fabric, or string filter media, the filter material is supported by a filter element and housed in a pressure vessel. The application of cartridge filters using either cleanable ceramic or disposable polypropylene cartridge

Figure 4. Microfiltration (MF) system (*Source*: U.S. Filter).

seems to be a feasible method for removing modest levels of turbidity, algae, and microbiological contaminants. As water is filtered through a cartridge filter, the pressure drop increases, which necessitates terminating the filter run. When this process is done, the filter is not backwashed, but the cartridge is thrown away and replaced by a clean filter.

The desire to reduce costs or to treat some waters more effectively is driving the development of new technologies for filtration. Several examples of such proprietary filters include low-head continuous backwash filters, two-stage filtration systems, bag filters, moving media filters, and flotation and filtration systems. A variety of modes of operation are available such as downflow, up-flow, multiple layer, biflow, radial flow, semicontinuous (moving bridge), continuous (media recycle), variable voidage, and pebble matrix.

READING LIST

ASCE and AWWA. (1998). *Water Treatment Plant Design*, 3rd Edn. McGraw-Hill, Inc., New York, NY.

AWWARF. (1998). *Treatment Process Selection for Particle Removal*. American Water Works Association Research Foundation, Denver, CO.

AWWA. (1999). *Water Quality and Treatment*, 5th Edn. McGraw-Hill, Inc., New York, NY.

HDR, Inc. (2001). *Handbook of Public Water System*, 2nd Edn. John Wiley & Sons, New York, NY.

Hudson, H.E. (1981). *Water Clarification Processes*. Van Nostrand Reinhold Company, New York, NY.

Kawamura, S. (2000). *Integrated Design of Water Treatment Facilities*. John Wiley & Sons, New York, NY.

Qasim, S.R., Motley, E.M., and Zhu, G. (2000) *Water Works Engineering*. Prentice-Hall, Upper Saddle River, NJ.

USEPA. (1990). *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*. Office of Drinking Water, Center for Research Information, Cincinnati, OH.

SYNTHETIC AND NATURAL ORGANIC REMOVAL BY BIOLOGICAL FILTRATION

Rasheed Ahmad
Khafra Engineering Consultants
Atlanta, Georgia

Oligotrophy is important because organic substrates are present in drinking water supplies at microgram per liter levels. Most biological processes are aerobic, which means that dissolved oxygen is present and used as the electron acceptor by the bacteria.

The benefits of biological filtration include decreased potential for bacterial regrowth in the water distribution system, reduced chlorination disinfection byproducts (DBPs), reduced chlorine demand, and decreased corrosion potential. So biological filtration is used to achieve three broad goals: (1) biologically oxidize biodegradable components, making the water biologically stable and reducing the need for excess chlorination; this will in turn reduce the formation of DBP compounds, which are known human carcinogens; (2) biodegrade synthetic organic micropollutants that are harmful to human health; and (3) remove nitrate and nitrite via denitrification. In most drinking water treatment, however, the first two goals are the main objectives of biological filtration. Particulate (silt, clay, precipitates) removal does occur in biological filters even though it may not be an intentional goal.

It is well established that ozonation increases the fraction of natural organic material (NOM) that is biodegradable. The effects of ozonation on NOM include formation of hydroxyl, carbonyl, and carboxyl groups; increased polarity and hydrophilicity; loss of double bonds and aromaticity; and a shift in molecular weight distribution toward lower molecular weight compounds. Thus, an increase in biodegradable organic matter (BOM) on ozonation generally enhances biological activity considerably in filters, after ozonation. Often, biological filtration can reduce BOM concentration to approximately preozonation levels, although this depends on the specifics of biological filters and water quality parameters, and the composition of BOM may be different after biological filtration. Currently identified ozonation byproducts such as aldehydes, carboxylic acids, and keto acids are biodegraded by biological filters that have more than 75% removal efficiency. Several synthetic organic compounds (SOCs) are also substantially biodegraded by the biological filtration process. In particular, phenol, chlorinated phenols, and chlorinated benzenes show significant percentage removals immediately or after a short acclimation period. In general, SOCs that are built on an aromatic structure should be most susceptible to biodegradation, and halogenated aliphatics should be more resistant to biodegradation.

READING LIST

Ahmad, R., Amirtharajah, A., Al-Shawwa, A., and Huck, P. (1998). Effects of backwashing on biological filters. *J. American Water Works Assoc.* **12**: 62–73.

AWWA. (1999). *Water Quality and Treatment*, 5th Edn. McGraw-Hill, Inc., New York, NY.

Bouwer, E.J. and Crowe, P.B. (1988). Biological processes in drinking water treatment. *J. American Water Works Assoc.* **9**: 82–91.

Hozalski, R.M., Goel, S., and Bouwer, E.J. (1995). TOC removal in biological filters. *J. American Water Works Assoc.* **12**: 40–54.

Mallevialle, J., Suffet, I.H., and Chan, U.S. (Eds.). (1992). *Influence and Removal of Organics in Drinking Water*. CRC Press, Inc., Boca Raton, FL.

Manem, J.A. and Rittman, B.E. (1992). Removing trace-level organic pollutants in a biological filter. *J. American Water Works Assoc.* **4**: 152.

Rittman, B.E. and Snoeyink, V.L. (1984). Achieving biologically stable drinking water. *J. American Water Works Assoc.* **10**: 106.

Rittman, B.E. and Huck, P.M. (1989). Biological treatment of public water supplies. *Crit. Rev. Environ. Cont.* **19**: 119–184.

Urfer, D., Huck, P.M., Booth, S.D.J., and Coffey, B.M. (1997). Biological filtration for BOM and particle removal: A critical review. *J. American Water Works Assoc.* **12**: 83–98.

GRANULAR BED AND PRECOAT FILTRATION

Kwok-Keung Au
Greeley and Hansen
Chicago, Illinois

Various filtration processes are used in potable water treatment. Granular bed and precoat filtration processes are described in this article, and focus is on descriptions of these processes and their particle removal efficiencies. Filtration processes are some of the major tools that are used to minimize microbial pathogens in drinking water, so the descriptions of their removal efficiencies contained herein emphasize their removal of microbes. This type of efficiency is characterized by a term called log removal, which is defined as log C_0/C, where C_0 and C are the pathogen concentrations in a filter's inlet and outlet flows, respectively. Thus, a 1-log removal is equivalent to 90% removal. Current drinking water regulations target log removals of pathogens. Filter performance is also based on the amount of turbidity in the filtered water. Turbidity is a measure of water clarity. Turbidity reduces the aesthetic acceptability of drinking water and also shields pathogens from disinfectants that are added downstream of filters.

GRANULAR BED FILTRATION

In this process, water passes through a filter containing a bed of granular material. Particles are removed by different mechanisms, depending, primarily, on the filter medium used and the hydraulic filtration rate. The term "particles," used herein, refers to any particulate contaminants in water. These include very fine inanimate and biological objects and microbial pathogens. Granular filtration can be classified as slow sand filtration or high-rate granular filtration. Slow sand filters are operated at filtration rates of less than 0.4 m/h; the filtration rates for high-rate filters are typically 5 to 25 m/h. The design/operating criteria and particle removal mechanisms for slow sand filtration and high-rate filtration are considerably different, so each process is described in a separate subsection.

Slow Sand Filtration

Slow sand filtration was first developed in England about 200 years ago to protect drinking water consumers from microbe-related health risks. Detailed design guidelines for slow sand filters are available (1). A slow sand filter consists of a layer of sand supported on a graded gravel layer. The gravel layer surrounds a system of perforated pipes that collect the filtered water. The gravel keeps the sand from entering the pipe perforations and, thereby, being lost from the filter. Water passes through the sand and gravel by gravity, normally without being pretreated, and as it does, the sand removes particles. Biological and physicochemical removal mechanisms are believed to be involved (2,3). Specifically, these may include biological action (e.g., ciliate protozoa on the sand grains acting as bacterial predators), particle attachment to the sand grains (e.g., the extracellular macromolecules of microbial particles attaching to the sand grains) and physical straining between sand grains.

The predominant removal location is a thin layer at the top of the sand bed. This biologically active layer is termed the *schmutzdecke*. It contains living and dead micro- and macroorganisms. As filtration progresses, materials removed and biological growths on the sand grains increase the loss of hydraulic head as the water passes from above to below the sand layer. When the head loss reaches a predetermined limit, normally from 1 to 2 m, the filter is removed from service. Before being returned to service, the filter is cleaned usually by scraping about 2 cm of accumulated material and sand from the top layer of the sand bed, which is nominally the *schmutzdecke*. The typical time period between cleanings (called a filter run) varies from 1 to 6 months, depending on the inlet water characteristics and filtration rate. After the sand bed depth is reduced to the lowest acceptable value by multiple scrapings, replacement of the entire sand bed is necessary.

Pretreatment is normally not required for slow sand filtration. However, if the source water has a high concentration of suspended particles, physical pretreatment processes, such as roughing filters or microstrainers, may be employed to prevent rapid clogging of the sand and, thus, maintain reasonable filter run periods.

Virus removal by slow sand filters can range from slightly better than 1 log to as much as 5 logs (4). Fairly good removal of bacteria and protozoa can also be achieved by slow sand filtration. During a pilot-scale study, a new filter removed 0.82 logs of total coliform bacteria and greater than 1.7 logs of *Giardia* (5). After 2 weeks, a microbiological population was established within the sand bed, the removal of total coliforms increased to 4 logs, and no *Giardia* were detected in the filtered water. The calculated reduction of *Giardia* was greater than 2.6 logs.

In a full-scale study at three slow sand filtration plants in Idaho, no *Giardia* were detected in the filtered water at two of the three plants (6). For the one positive sample found at the third plant, 1-log removal of *Giardia* was achieved. In the same study, removal of total coliforms and fecal coliforms varied from 0.81 to 2.30 logs and from 0.29 to 0.52 logs, respectively. These ranges in removal efficiency were influenced by inlet water quality, filtration rate, and media size and depth. Removing *Cryptosporidium* by slow sand filtration is often more difficult than removing *Giardia*. In a full-scale study conducted in British Columbia, the average removal of *Giardia* was 1.16 logs, whereas it was only 0.28 logs for *Cryptosporidium* (7).

The reduction of turbidity by slow sand filtration may be site-specific. Turbidity removal in a pilot study was 1.66 logs or better after the filter was operated for 2 days (8). However, only 0.14 to 0.22 log removal of turbidity was found in another study, even though the reduction of *Giardia* in the same study was as high as 4 logs (5). The authors concluded that this was due to the fine clay particles in the inlet water to the filter. In a full-scale study, between 0 to 0.43 log removal of turbidity was reported (6). This low removal was due to the fine particles in the inlet water and to the large fraction (4% by weight) of fines in the sand medium used in the study.

Under proper design and operating conditions, slow sand filtration can provide a high degree of protection against microbial pathogens. This includes effective removal of viruses, bacteria, and protozoa. If the inlet water is of high quality, pretreatment is normally not required. The construction, operating, and maintenance costs of slow sand filtration are low. However, land requirements are high. This technology is therefore attractive for small systems that treat high-quality surface water.

High-Rate Filtration

High-rate granular filtration is the most widely used filtration process in drinking water treatment. A comprehensive review of this process is available (9). Filter media types include monomedia (typically silica sand but also anthracite coal), dual media (usually anthracite coal over silica sand), and trimedia (commonly anthracite coal over silica sand over garnet). Granular activated carbon is used when both particle removal and adsorption of dissolved organic compounds are desired.

Particle removal by high-rate granular filtration is a physicochemical process that involves two sequential steps, transport and attachment. Particles are initially transported from suspension to within a very close distance of filter media grains, and then attachment of particles to the media grains occurs (10). The transport step depends on the physical–hydrodynamic properties of the system. Process variables such as the size and density of particles, size and depth of filter media, and filtration rate affect transport efficiency. Attachment is controlled by the surface and solution chemistry of the filter media and particle suspension (11). Unfavorable interactions between the particles and the filter media must be avoided so that the particles can attach to the media. Chemical coagulation prior to filtration is used to overcome any unfavorable interactions between the particles and the filter media, and it is the single most important factor in determining filtration efficiency. Optimum coagulation, which is typically defined as the minimum coagulant dose required to achieve the best filtered water quality, should be provided. Without proper chemical coagulation, high-rate filtration works as a simple straining process and is not effective for particle removal.

Depending on the degree of pretreatment of the water before it is applied to the filters, granular filtration can be operated in three different treatment modes: conventional, direct, and in-line. Conventional treatment, appropriate for most source waters, consists of adding coagulant chemicals in rapid mixing basins, flocculation in slow mixing basins, and particle removal in sedimentation basins and filters. When sedimentation is omitted, the treatment mode is termed “direct filtration.” If both flocculation and sedimentation are omitted, the term “in-line filtration” is used. Direct or in-line filtration is used for source waters of consistently good quality, such as those of low turbidity and low color.

Particle removal by high-rate granular filtration occurs within the depth of the granular media (depth filtration) rather than only at the top of the filter media (cake filtration). A filter is cleaned when, during operation, the hydraulic head lost across the filter medium or the filtered water quality reaches unacceptable levels. The cleaning consists of surface washing and backwashing the filter medium. Surface washing is used to break up the mat of particles that has accumulated near the surface of the medium. This is achieved by directing water jets downward toward the filter medium’s surface or causing air bubbles to rise upward through the medium. Backwashing is achieved by a reverse flow of water applied below the medium and rising through the medium bed to flush the particles that have accumulated at the surface and within the filter bed out of the medium and the filter.

After backwashing, filters exhibit an initial period of poor performance (termed the ripening period) before achieving good filtered water quality. The passage of particles during the ripening period can be formidable. The ripening period can last from less than an hour to several hours. Different methods are used to minimize this impact, such as wasting the initial filtered water (filter -to-waste), limiting the initial filtration rate until the filtrate quality is acceptable (slow start), placing the filter in operation after a period of inactivity following backwashing (delayed start), or adding coagulant or polymer chemicals to the backwash water supply.

The importance of chemical pretreatment to removal efficiency by high-rate granular filtration has been emphasized in numerous studies. Without adding any chemicals, the average log removal of *Giardia* was 0.60 logs for conventional treatment and 0.44 logs for in-line filtration in a pilot study (12). In optimum chemical pretreatment, the removal increased to 1.70 logs for conventional treatment and 1.19 logs for in-line filtration. Results from a 2-year study indicated that maintaining optimum coagulation conditions to produce low filtered water turbidity (0.1 to 0.2 NTU) provided effective removal of *Giardia* and *Cryptosporidium* (13). Under optimum coagulation conditions, the average removal of *Giardia* was 3.3 logs or better, and the average removal of *Cryptosporidium* was 2.3 logs or better.

Results from a full-scale study confirmed that a properly operated conventional treatment plant provided a substantial barrier to microbial pathogens (14). Dual media filters were used in this plant. Coagulation was provided using alum and activated silica. Prechlorination was applied at a dosage of 1 mg/L. *Giardia* was detected in only 1 of 32 filtered water samples, and the mean removal was 3.6 logs by sedimentation and filtration. *Cryptosporidium* was detected in 7 of 32 filtered water samples, and the mean removal was 2 logs. *Clostridium perfringens* was detected in 9 of 33 filtered water samples, and the mean removal was 4.4 logs. No human enteric viruses were detected in 32 filtered water samples where the mean removal was 3.1 logs. *Somatic coliphage* was detected in 24 of 32 filtered water samples whose a mean removal was 3.5 logs.

Under proper design and operating conditions, high-rate granular filtration can act as a consistent and effective barrier to pathogenic microbes (15). Providing effective chemical pretreatment is the single most important factor in enhancing treatment efficiency. It is expensive, if not impossible, to remove microbes effectively without the proper chemistry provided by optimum coagulation.

Precoat Filtration

Precoat filtration was originally developed by the United States Army during World War II. The objective was a portable device to remove *Entamoeba histolytica*, a protozoan parasite prevalent in the Pacific war zone, from potable water. Precoat filtration is now accepted by the U.S. Environmental Protection Agency as a filtration technique for potable water treatment. A detailed design and operation manual for precoat filtration has been published (16).

In precoat filtration, water is forced by pressure or pulled by vacuum through a uniform, thin layer of filtering material precoated onto a permeable, rigid supporting structure (septum). As water passes through the precoat filter medium (filter cake) and septum, particles whose sizes are larger than the pore size of the filter medium are strained and removed. Typical filtration rates for this process are about 2.5 to 7 m/h.

During filtration, a slurry (called the body-feed solution) containing the coating material is often continuously added at the filter inlet to maintain filter cake permeability. As the filter cake thickness builds due to captured particles and the added precoat, the hydraulic head lost across the filter cake increases to a level impractical for further filtration. The filter cake is then removed from the support septum and disposed of. The filter is then cleaned and precoated with a new layer of coating materials, and a new filter run is started.

Precoat materials include diatomaceous earth (DE) and perlite; DE is more commonly used in drinking water treatment. DE is composed of fossilized skeletons of microscopic water plants called diatoms and is almost pure silica. Perlite originates as rock formed by the solidification of magma, is composed primarily of alumina silicate, and processed into a fine granular material for use as precoat. The primary removal mechanism of precoat filtration is straining, so its removal efficiency depends to a great extent on the pore size of the coating materials. The median pore size is called the media grade. Diatomite grades used for drinking water treatment have a median pore diameter of about 3 to 17 μm. Other important factors that influence the removal efficiency are chemical pretreatment of the filter medium, filtration rate, and body-feed rate. To maintain reasonable filter run time, the inlet water must be of high quality. Inlet water turbidity levels of less than 10 NTU are desirable, or pretreatment of the source water should be considered.

A pilot study showed complete removal of *Giardia* by both coarse and fine grades of DE in filtration across a wide range of operating conditions (17). Removal of *Cryptosporidium* by a bench-scale DE filter ranged from 3.60 to 6.68 logs, depending on the media grade and filtration rate (18). In a pilot-plant study, complete removal of *Giardia* and 3-log removal of *Cryptosporidium* by DE filtration were reported (19). Based on the results from another pilot-scale study, 6-log removal of *Cryptosporidium* was achieved by DE filtration at rates from 2.5 to 5 m/h (20). In most of the runs, the turbidity level decreased from about 1 NTU in the inlet water to less than 0.1 NTU in the outlet water.

Precoat filtration can remove protozoan parasites very effectively, but chemical pretreatment of coating media with aluminum or iron coagulants or cationic polymers is required to remove submicron-sized microbes such as viruses and some bacteria. In a pilot study, removal of coliforms by untreated DE was about 0.36 logs (19). The removal increased to 0.82 logs for DE coated with an alum dosage of 0.001 g/g DE and to 2 logs with an alum dosage of 0.003 g/g DE. This enhancement was probably due to the enmeshment of the bacteria on the aluminum precipitates. A similar beneficial effect was observed by coating the DE with cationic polymers. When the DE was coated with a polymer dosage of 0.0035 g/g DE, the removal of coliforms was greater than 3.3 logs. The authors concluded that this could be due to increased density of positively charged sites on the polymer-coated DE for adsorption of negatively charged coliforms. An enhancement in virus removal by chemical pretreatment of filter cake has been reported (21). The removal of bacteriophage T2 and poliovirus was about 1 log for an uncoated filter cake. When the filter cake was coated with ferric hydrate or polyelectrolytes, the removal increased to more than 1.7 logs.

Under proper design and routine operating conditions, precoat filtration can remove protozoan parasites such as *Cryptosporidium* and *Giardia* very effectively. Precoat filters can also be effective in removing submicron-sized microbes including viruses and some bacteria, provided that chemical pretreatment of the coating materials is employed.

BIBLIOGRAPHY

1. Hendricks, D. (1991). *Manual of Design for Slow Sand Filtration*. American Water Works Association Research Foundation, Denver, CO.

2. Weber-Shirk, M.L. and Dick, R.I. (1997). Biological mechanisms in slow sand filters. *Journal of American Water Works Association* **89**(2): 72–83.
3. Weber-Shirk, M.L. and Dick, R.I. (1997). Physical-chemical mechanisms in slow sand filters. *Journal of American Water Works Association* **89**(2): 87–100.
4. Ellis, K.V. (1985). Slow sand filtration. *Critical Reviews in Environmental Control* **15**(4): 315–354.
5. Bellamy, W.D., Silverman, G.P., Hendricks, D.W., and Logsdon, G.S. (1985). Removing *Giardia* cysts with slow sand filtration. *Journal of American Water Works Association* **77**(2): 52–60.
6. Tanner, S.A. and Ongerth, J.E. (1990). Evaluating the performance of slow sand filters in northern idaho. *Journal of American Water Works Association* **82**(12): 51–61.
7. Fogel, D. et al. (1993). Removing *Giardia* and *Cryptosporidium* by slow sand filtration. *Journal of American Water Works Association* **85**(11): 77–84.
8. Cleasby, J.L., Hilmoe, D.J. and Dimitracopoulos, C.J. (1984). Slow sand and direct in-line filtration of a surface water. *Journal of American Water Works Association* **76**(12): 44–55.
9. Cleasby, J.L. and Logsdon, G.S. (1999). Granular bed and precoat filtration. In: *Water Quality & Treatment*, R.D. Letterman (Ed.). McGraw Hill, Inc., New York, pp. 8.1–8.99.
10. Yao, K.M., Habibian, M.T., and O'Melia, C.R. (1971). Water and waste water filtration: Concepts and applications. *Environmental Science and Technology* **5**(11): 1105–1112.
11. Tobiason, J.E. and O'Melia, C.R. (1988). Physicochemical aspects of particle removal in depth filtration. *Journal of American Water Works Association* **80**(12): 54–64.
12. Ongerth, J.E. (1990). Evaluation of treatment for removing *Giardia* cysts. *Journal of American Water Works Association* **82**(6): 85–96.
13. Nieminski, E.C. and Ongerth, J.E. (1995). Removing *Giardia* and *Cryptosporidium* by conventional treatment and direct filtration. *Journal of American Water Works Association* **87**(9): 96–106.
14. Payment, P. et al. (2000). *An Epidemiological Study of Gastrointestinal Health Effects of Drinking Water*. AWWA Research Foundation and American Water Works Association, Denver, CO.
15. Au, K.-K. and LeChevallier, M.W. (2002). Removal of pathogenic microbes by granular high-rate filtration. In: *Encyclopedia of Environmental Microbiology*. G. Bitton (Ed.). John Wiley & Sons, New York, pp. 2707–2713.
16. *American Water Works Association*, Precoat Filtration. AWWA Manual M30, Denver, CO, (1995).
17. Langé, K.P., Bellamy, W.D., Hendricks, D.W., and Logsdon, G.S. (1986). Diatomaceous earth filtration of *Giardia* cysts and other substances. *Journal of American Water Works Association* **78**(1): 76–84.
18. Ongerth, J.E. and Hutton, P.E. (1997). DE filtration to remove *Cryptosporidium*. *Journal of American Water Works Association* **89**(12): 39–46.
19. Schuler, P.F. and Ghosh, M.M. (1990). Diatomaceous earth filtration of cysts and other particulates using chemical additives. *Journal of American Water Works Association* **82**(12): 67–75.
20. Ongerth, J.E. and Hutton, P.E. (2001). Testing of diatomaceous earth filtration for removal of *Cryptosporidium* oocysts. *Journal of American Water Works Association* **93**(12): 54–63.
21. Brown, T.S., Malina, J.F., Jr., and Moore, B.D. (1974). Virus removal by diatomaceous-earth filtration-part 2. *Journal of American Water Works Association* **66**(12): 735–738.

FLOCCULATION

Rasheed Ahmad
Khafra Engineering
Consultants
Atlanta, Georgia

The flocculation process aggregates destabilized particles into larger and more easily settleable flocs. The coagulation process destabilizes particles through chemical reactions between the coagulant and the suspended colloids, but flocculation is the transport step that causes the necessary collision between destabilized particles. The purpose of the flocculation process is to promote the interaction of particles and to form aggregates that can be efficiently removed in the subsequent separation processes such as sedimentation, flotation, and coarse bed filtration. In practice, flocculation is accomplished by providing a gentle mixing phase that follows the rapid dispersion of coagulants.

There are three major mechanisms of flocculation: perikinetic, orthokinetic, and differential settling. In perikinetic flocculation, the aggregation of particles occurs as a result of random thermal motion (Brownian diffusion) caused by continuous bombardment by surrounding fluid molecules. The driving force for this type of particle movement is the thermal energy of the fluid. This is significant for particles that are 1 to 2 microns in size. In orthokinetic flocculation, the aggregation of particles is caused by induced velocity gradients in the fluid. The suspended particles follow the streamlines at different velocities and eventually lead to interparticle contacts. The velocity gradient is related to the energy dissipated into the water (via mixing). Flocculation by differential settling occurs when particles have unequal settling velocities, and their alignment in the vertical direction makes them tend to collide when one overtakes on other. This collision occurs in sludge blanket or solids contact clarifiers where differential and fluctuating velocities can lead to particle collision and aggregation. In water treatment, the predominant mechanism is orthokinetic flocculation.

In systems that are mixed (velocity gradients are induced), the velocity of the fluid varies both spatially and temporally. The spatial changes in the velocity are termed the *velocity gradient*, G. The unit of velocity gradient is $\sec^{-1}$. In water treatment plants, mean velocity gradients of 10 to 100 $\sec^{-1}$ are typical for flocculation. Flocculation usually follows a rapid mixing process in conventional treatment plants. Flocculation is a time-dependent process that directly affects clarification efficiency by providing multiple opportunities for particles suspended in water to collide through gentle and prolonged agitation. Most water treatment plants provide 20 to 30 min of flocculation time (at 20 °C). An increase in flocculation time does not improve flocculation significantly. The optimum flocculation conditions are those that rapidly form flocs large enough to be separated by settling, flotation, or filtration leaving no residual primary particles or small aggregates. However, rates of flocculation that are too high cause floc breakup. Thus, a balance must be achieved between the velocity gradient G and the flocculation time t, one compensating for the other. In some cases, flocculation is characterized by the dimensionless product Gt (sometimes called the Camp number). Based on observational data from water utilities in the United States, the optimal value of Gt is set between the limits of 10^4 and 10^5. It is desirable to compartmentalize the flocculation process by dividing the basin into two or more defined compartments or stages. Compartments prevent short-circuiting and permit defined zones of reduced energy input or tapered flocculation. Studies have shown that tapered flocculation at a diminishing velocity gradient is more efficient than uniform velocity flocculation.

Flocculation can be achieved by hydraulic methods or mechanical devices. Hydraulic methods are used most often in small plants. Mechanical flocculators cover a broad range of configurations. Mechanical flocculators are preferred by most design engineers in the United States because of their flexibility in varying G values and because they cause low head loss. Mechanical flocculators can be divided into two major types: (1) shaft with turbine or propeller type blades and (2) paddle type with either horizontal or vertical shafts. Turbines can have flat or curved blades that are connected to a disk or shaft (Fig. 1). Propellers are shaped like ships' screws. The blades are mounted on a vertical or inclined shaft. The paddle type consists of blades attached directly to vertical or horizontal shafts (Figs. 2 and 3). In some flocculators, moving blades (rotors) may be complemented by stationary blades (stators) that oppose rotational movement of the entire mass of water within the treatment unit and help suppress vortex formation. Paddles are rotated at slow to moderate speeds of 2 to 15 rpm. The currents generated by them are both radial and tangential. The type of mechanical flocculators influences the shape of flocculation compartments. Vertical flocculators are often associated with square compartments, whereas the horizontal shaft, reel, or paddle flocculators are associated with rectangular compartments. Between each zone or stage of mechanical flocculation, baffles are designed to prevent short-circuiting. Vertical flocculators are more applicable to high-energy flocculation such as direct filtration.

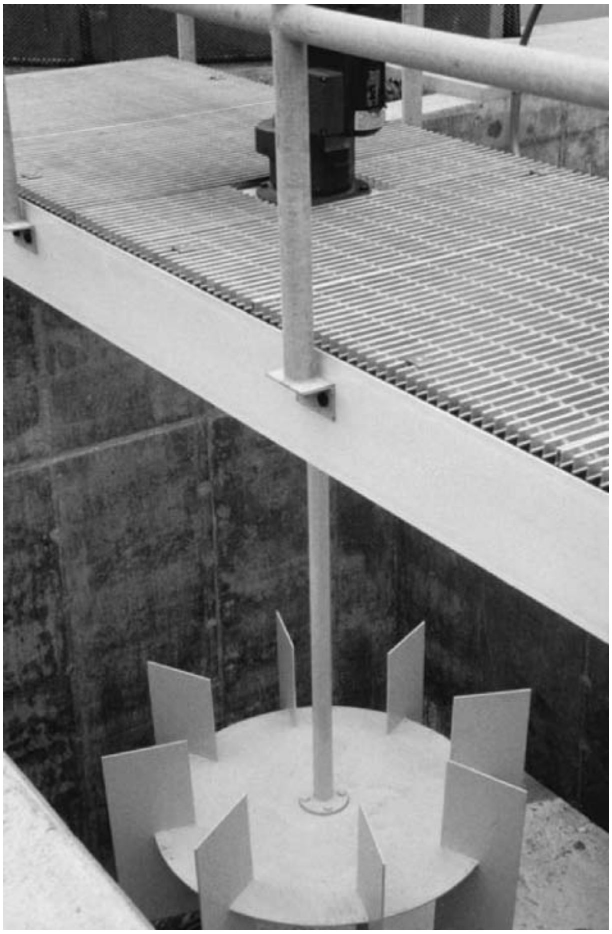

Figure 1. Flat blade turbine type flocculator (*Source*: Dorr Oliver Eimco).

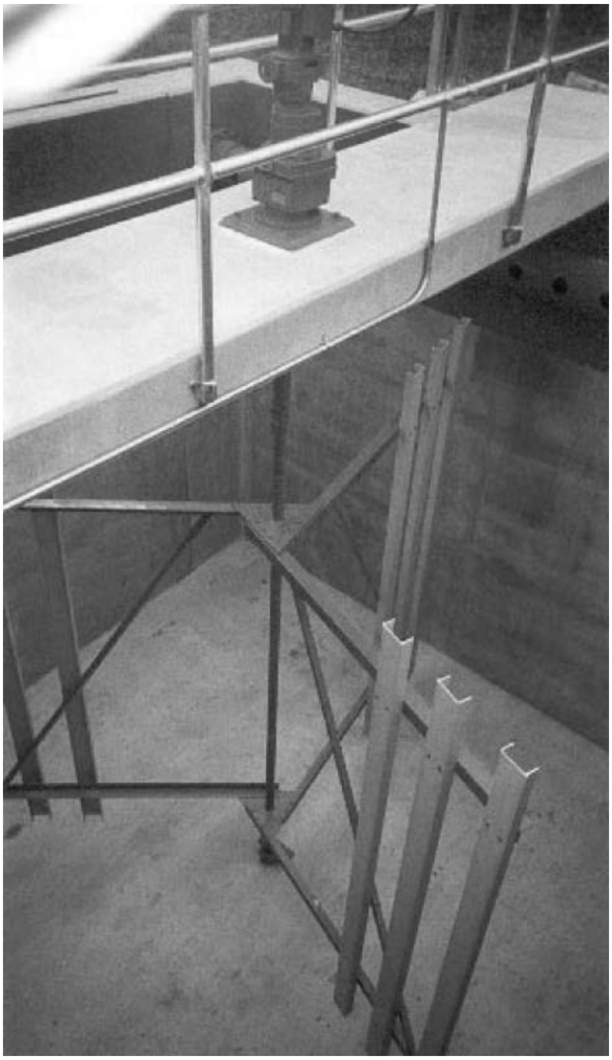

Figure 2. Vertical shaft paddle flocculator (*Source*: Walker Process Equipment, a Division of McNish Corporation).

Hydraulic flocculation methods are simple and effective, especially if flows are relatively constant. Hydraulic flocculation is achieved in baffled channel basins. They can be horizontally baffled (around-the-end flow) or

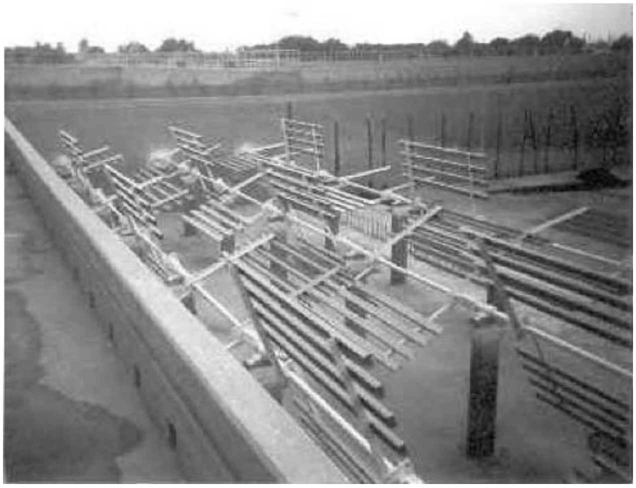

Figure 3. Horizontal shaft paddle flocculator (*Source*: Walker Process Equipment, a Division of McNish Corporation).

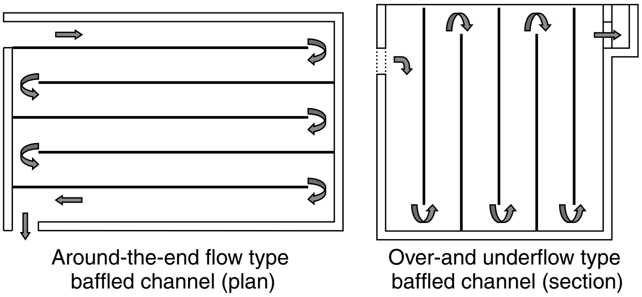

Figure 4. Hydraulic or baffled channel flocculation.

vertically baffled (over- and underflow) channels (Fig. 4). The velocity gradients are purposely intensified in these channels by enforced changes in the direction of flow. When designed properly, baffle channels perform well and exhibit good plug flow characteristics. However, they have two disadvantages: there is a significant head loss across the tank, and the mixing intensity is a function of plant flow rate. Some new water treatment plants in the United States have used a combined system of hydraulic and mechanical flocculation. It offers a large reduction in mechanical flocculator units, so there is a saving in capital and maintenance costs and increased reliability.

Some other types of flocculators are not commonly used in water treatment. Contact flocculation (gravel-packed filter) requires minimal maintenance because it has no moving parts. This type of flocculation system depends on plant flow rate and requires minimal variation in flow rate and water temperature for adequate mixing intensity. Contact flocculation also requires frequent backwashing. Contact flocculation may be applicable for difficult high-suspended-solids water or water with low total dissolved solids that may not respond readily to metal ion coagulants. Diffused air or water jet agitation is used as an auxiliary or temporary flocculation. This system is characterized by a high rate of energy consumption and is an inefficient flocculation process.

Several manufacturers provide proprietary designs that incorporate rapid mixing, flocculation, and settling in one unit, which is sometimes called a solids contact reactor unit. These units perform best in waters that can develop a dense, fast-settling floc, but the process may be difficult to control in some waters. These units have an option for sludge recirculation. Recirculating water treatment sludge may improve efficiency and reduce chemical requirements. The walking beam is another kind of proprietary flocculator. Walking beam flocculators are driven in the vertical direction in reciprocating fashion. The unit contains a series of cone-shaped devices on a vertical rod. The cone devices impart energy to the water as they move up and down, thereby creating velocity gradients.

READING LIST

ASCE and AWWA. (1998). *Water Treatment Plant Design*, 3rd Edn. McGraw-Hill, New York.

AWWARF. (1998). *Treatment Process Selection for Particle Removal*. American Water Works Association Research Foundation, Denver, CO.

AWWA. (1999). *Water Quality and Treatment*, 5th Edn. McGraw-Hill, New York.

Fair, G.M., Geyer, J.C., and Okun, D.A. (1968). *Waste and Wastewater Engineering*. John Wiley & Sons, New York.

Hudson, H.E. (1981). *Water Clarification Processes*. Van Nostrand Reinhold, New York.

HDR Engineering, Inc. (2001). *Handbook of Public Water Systems*, 2nd Edn. John Wiley & Sons, New York.

Kawamura, S. (2000). *Integrated Design of Water Treatment Facilities*. John Wiley & Sons, New York.

USEPA. (1990). *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*. Office of Drinking Water, Center for Research Information, Cincinnati, OH.

FLUORIDATION

John E. Dodes
Forest Hills, New York

Fluoridation of water supplies has been a controversial subject for more than 60 years. Much of the debate is concerned with questions of safety, but recently questions have been also raised about the efficacy of fluoridation. More than 35,000 papers on fluoridation have been published in the years since fluoride supplementation was first proposed as a safe and effective way of decreasing tooth decay. The overwhelming conclusion is that fluoride is safe and cost-effective.

TOOTH DECAY

Tooth decay is the most common disease of all humanity. It is the principal cause of tooth loss from early childhood through middle age. And decay continues to be a major problem as we age, particularly root decay which seriously affects the elderly due to receded gums and a general decrease in the amount of saliva. Dental decay not only affects the general health by interfering with the ability to eat and digest certain foods, but it also impacts an individual's emotional and social well-being. Recent surveys have found that more than 94% of adults have had decayed teeth, and 22.5% had root surface decay. In 1997, the United States spent more than $50.6 billion on dental care.

HISTORY OF WATER FLUORIDATION

In 1892, Sir Crichton-Browne advocated augmenting the common diet of his era with fluoride to reduce decay (1). Not until the 1950s, however, did years of research pay off, and it was unequivocally determined that when water contains approximately 1 part of fluoride per million parts of water (1 ppm) which is equivalent to 1 milligram of fluoride per liter of water (1 mg/L), decay rates are reduced by up to 60%. In the 1940s, there was substantial resistance to adding fluoride to community water supplies, even though people had been drinking naturally occurring fluoridated water containing several times the 1-ppm level of fluoride for a lifetime without any negative side effects (2).

The most famous study of the effects of fluoridation was done in the 1940s in Newburgh and Kingston, New York. These cities are located 35 miles apart near the Hudson River; both had populations of about 30,000 and were also similar in their demographics. Newburgh's water supply was fluoridated, and Kingston's was not. After 10 years, the study found that there were no medical differences between the two groups except for the fact that Newburgh's children had almost 60% fewer cavities (2). Many studies have confirmed these findings in the years since.

FLUORIDE

Fluoride is one of the earth's most common elements and is therefore present in variable amounts in all water supplies. It is also found in most plants and animals that we eat. The fluoride concentrations in water supplies within the United States vary from 0.1 to 10 mg/L (3). Ocean water also contains fluoride at concentrations of 1.0 to 1.5 mg/L. This causes a fairly uniform level of fluoride in all seafood.

Fluoride is classified by the National Academy of Sciences as an essential nutrient (4). Unlike many other essential elements that are found in food, water consumption is the most practical, consistent and effective method of fluoride application to the teeth. In growing children, fluoride will be incorporated throughout the entire hard structure (enamel and dentin) of the teeth. This continues until around age 50 when it appears that the spaces available to fluoride in the tooth's structure will be filled. In adults, fluoride will continue to be absorbed by the enamel surface, lending the teeth temporary but substantial resistance to decay. Fluoride is easily absorbed into the blood stream from the gastrointestinal tract and reaches a peak concentration within 20–60 minutes. This level declines rapidly due to the uptake of fluoride by the hard tissues and the removal of fluoride by the kidneys. Approximately 50% of the fluoride that is absorbed is incorporated in the body's teeth and bones within 24 hours.

The amount of decay reduction caused by fluoridation of local water supplies has decreased during the last 40 years probably because of improved dental hygiene and widespread use of fluoride toothpaste. This increase in fluoride availability has led to a reduction in the dose of fluoride supplementation recommended for children living in non-optimally fluoridated communities (Table 1). In 1991, it was found that fluoride reduces the incidence of cavities 20% to 40% in children and 15% to 35% in adults (5).

Table 1. Supplemental Fluoride Dosage (Milligrams of Fluoride per Day)[a]

	Concentration of Fluoride in Water (parts/million)		
Age (years)	0.0 to 0.3	0.3 to 0.6	Over 0.6
Birth to six months	None	None	None
Six months to three years	0.25	None	None
Three to six years	0.50	0.25	None
Six to sixteen years	1.0	0.50	None

[a] Reference 6.

THE SAFETY OF FLUORIDE

In 1970, the World Health Organization (WHO) issued a report, "Fluorides and Human Health," that had taken years of research to compile. The WHO wanted to evaluate impartially the vast number of scientific studies of fluoridation that had been published. These papers included population studies, experimental research, animal studies, human autopsy studies, clinical trials, and X-ray research. The WHO expert panel concluded that there was no reliable evidence that drinking water fluoridated at the recommended levels caused any ill health effects. In 1975, the WHO stated, "The only sign of physiological or pathological change in life-long users of optimally fluoridated water supplies .... is that they suffer less from tooth decay" (3).

Fluoride ingested through community water systems has a large margin of safety. It has never been shown that fluoride intake at 1 ppm has any negative effect on disease or death rates. Numerous studies performed before and after supplemental fluoridation have shown no changes in death rates from cancer, heart disease, intracranial lesions, nephritis, cirrhosis, or from any other cause. In addition, the normal disease and death rates of more than 7 million Americans who have lived for generations where the natural fluoride concentration was 2 to 10 mg/L (1 mg/L is the recommended dose), is compelling evidence of fluoridation's safety. Two extensive studies have established that there is no link between fluoridation and Down's syndrome, cleft palate, heart abnormalities, clubfoot, and other common birth defects.

Antifluoridationists have long claimed that fluoride use leads to an increase in cancer rates. Consumer's Union characterized this accusation as “absurd.” It has been shown that fluoride has no mutagenic effect in studies of cattle (7) or mice (8). This makes sense because fluoride is not in the class of electrophilic compounds that can interact with DNA, nor is it likely that the small tissue levels of fluoride present due to fluoridated water supplies could interfere with DNA replication.

It has also been proven that fluoride does not cause allergic reactions. The executive committee of the American Academy of Allergy has stated, “There is no evidence of allergy or intolerance to fluorides as used in fluoridation of community water supplies.”

The effect of fluoridated water on kidney function has been thoroughly investigated, and here again no ill effects have been shown. No kidney changes were detected in a population exposed for a lifetime to water supplies that have fluoride levels of 8 mg/L (1).

Opponents of fluoridation also routinely claim that it causes coronary artery disease. They use data on heart disease rates from Antigo, Wisconsin, to support this claim. Antigo did show increased death rates from heart disease in the period since fluoride was introduced to its water supply in 1949. But, in that same period, the percentage of elderly people living in Antigo doubled, due to longer life spans. Actually, the segment of the population 75 years old or older increased 106% (2). The higher death rates from heart disease amongst the elderly were never factored into the interpretation of the Antigo death-rate data. When this factor is taken into account, the alleged deleterious effect of fluoride disappears. Unfortunately, the population of Antigo believed the scare tactics and false claims of the antifluoridationist movement and voted to end the fluoridation of their water supply. After only 4 years, the decay rates in permanent teeth of second graders rose 183%. A year later, Antigo voted to reinstate fluoridation (2).

FLUORIDE SAFETY MARGINS

Food and water account for 1.2 to 2.6 mg of fluoride ingestion per day. Urban air usually contains less than 1 μg/m^3, an insignificant amount. The margin of safety for fluoride is very large, and toxic effects have not been demonstrated at levels far higher than one could receive in a lifetime of exposure to drinking water containing 8 times the recommended level of 1 ppm. The acute lethal dose for a 150 lb. (70 kg) man is 5 to 10 g of sodium fluoride or 2.3 to 4.5 g of fluoride (9). Chronic overexposure to fluoride at levels over 1 ppm before age 8 can lead to discoloration of the enamel, ranging from barely detectable white flecks to large brown areas. Known as enamel fluorosis, this solely cosmetic problem can be achieved only by ingesting high amounts of fluoride, not by topical application.

THE COST OF FLUORIDATING WATER SYSTEMS

Fluoride protection obtained through community water treatment costs approximately 50 ¢ per person per year (10). This saves an estimated $75 in dental treatment costs (11). The National Preventive Demonstration Program monitored nearly 30,000 children, ages 5 to 14, for 4 years and found that the most cost-effective method of decay prevention was to drink fluoridated water from birth and have sealants applied as needed (12).

THE ANTIFLUORIDATIONIST MOVEMENT

The water supply of the developed world is slowly becoming entirely fluoridated, although overcoming the misinformation promoted by those opposed to fluoridation is a long and difficult process. The resistance to this overwhelmingly endorsed public health benefit tells us a great deal about the psychology of fear, the strength of rumor, and the loss of trust in government and industry. William T. Jarvis, Ph.D., the Executive Director of the National Council Against Health Fraud, has stated,

> These charges seem to grow out of a mentality of distrust. Antifluoridation groups are led by many of the same people who oppose immunization, pasteurization, sex education, mental health programs, and other public health advances. Most are closely connected with sellers of alternatives to medically accepted products and services. The so-called “health food” industry justifies its existence by declaring that our conventional sources of food, water, and health care are misguided.

No studies will satisfy those who are opposed to fluoridation. A vociferous minority, made up mostly of food faddists, cultists, chiropractors, and people who misunderstand what fluoridation is, has developed effective ways of stopping fluoride from being adjusted to optimum levels in water. “Lifesavers Guide to Fluoridation,” a pamphlet by John Yiamouyiannis, Ph.D., is often distributed in communities that are considering fluoridation. It cites 250 references as proof that fluoride is dangerous. However, in 1988, experts from the Ohio Department of Health published “Abuse of the Scientific Literature in an Antifluoridation Pamphlet.” This review traced the references and found that almost half had no relevance to community water fluoridation and that many others actually supported fluoridation but were selectively misquoted and misrepresented.

In 1990, an article in *Newsweek* magazine implied that fluoridation was ineffective and unsafe. The article was a response to the unauthorized and premature release of data from an experiment at the National Institute of Environmental Health Science. The experiment exposed rats and mice to high doses of fluoride. A thorough review of the experiments by a U.S. Public Health Service expert panel concluded that the data were insignificant and that fluoridation posed no risk of cancer or any other disease. Dr. Stephen Barrett, a leading consumer health advocate, called the *Newsweek* article “the most irresponsible analysis of a public health topic ever published by a major national news outlet.”

Fluoride has been a favorite target of those who want to frighten the public into believing that our health system is dangerous and uncaring. They continue to undermine the significant health advances that have been made by public health measures, modern agriculture, and industry. As Consumers Union has concluded,

> The simple truth is that there's no "scientific controversy" over the safety of fluoridation. The practice is safe, economical, and beneficial. The survival of this fake controversy represents one of the major triumphs of quackery over science in our generation (2).

FLUORIDATION RATES AROUND THE WORLD

Tooth decay, or caries, is the most prevalent disease of all humanity, more common than the common cold (13). The cost of tooth decay in human suffering as well as economic expenditures is also enormous, accounting for many billions of dollars. Yet, as we enter the twenty-first century, only a small percentage of the world's population can easily drink fluoridated water. In addition, the majority of bottled water is not fluoridated, and many types of home water filtering devices actually remove the fluoride that may be in the water.

In the United States, 62.2% of the population has access to public water supplies that are optimally fluoridated (14). More than 360 million people worldwide in approximately 60 countries also drink fluoridated water. Dr. C. Everett Kopp, the former Surgeon General of the United States has stated, "Fluoridation is the single most important commitment that a community can make to the oral health of its citizens." It is imperative to continue to fluoridate water systems throughout the world.

BIBLIOGRAPHY

1. Richmond, V.L. (1985). Thirty years of fluoridation. *Amer. Jour. Clinical Nutrition* **41**: 129–138.
2. Anon. (July, August, 1978). Fluoridation: a two-part report. *Consumers Reports*.
3. World Health Organization. (1970). *Fluorides and Human Health*. Monograph 59, Geneva. World Health Organization.
4. National Academy of Sciences. (1980). *National Research Council Food and Nutrition Board. Recommended Daily Allowances*, 9th Edn. Washington, DC, pp. 156–59.
5. *American Dental Association* (1991). Facts about Fluoride.
6. Barrett, S., Jarvis, W.T., Kroger, M., and London, W.M. (1997). *Consumer Health: A Guide to Intelligent Decisions*, 6th Edn. Brown & Benchmark, p. 121.
7. Leonard, A., Deknudt, G.H., Decat, G., and Leonard, E.D. (1977). Cytogenic investigations of leucocytes of cattle intoxicated with fluoride. *Toxicology* **7**: 239–42.
8. Japello, G. and Lia, J.S. (1974). Sodium fluoride as potential mutagen in mammalian egg. *Arch. Environ. Health* **29**: 230–35.
9. Ripa, L., DePaola, P., Horowitz, H., Nowak, A. et al. (1986). *A Guide to the Use of Fluorides for the Prevention of Dental Caries*. American Dental Assoc.
10. Barrett, S., Jarvis, W.T., Kroger, M., and London, W.M. (1997). *Consumer Health: A guide to intelligent Decisions*, 6th Edn. Brown & Benchmark, Madison, WI.
11. Centers for Disease Contol. (1992). Fluoridation of community water systems. *JAMA* **267**: 3264–3265.
12. Weisfeld, V.D. (Ed.). *Preventing tooth decay: Results from a four-year national study. Special Report Number Two, 1983*. Robert Wood Johnson Foundation, Princeton, NJ.
13. Schissel, M.J. and Dodes, J.E. (1999). *Healthy Teeth: A User's Manual*. St. Martin's/Griffin, New York.
14. *American Dental Association*. (1999). Fluoridation Facts. Council on Access, Prevention and Interprofessional Relations, Chicago, IL.

GIARDIASIS

Christine L. Bean
University of New Hampshire
Durham, New Hampshire

GIARDIA LAMBLIA

Giardia lamblia is an intestinal flagellated protozoan, which infects the intestinal tract of a wide range of mammals, including humans. Two forms of this protozoan exist, a highly resistant cyst and a motile, feeding trophozoite. Cysts may be found in freshwater surface water supplies as a result of contamination from sewage and animal waste runoff. Trophozoites of this obligate parasite do not typically survive outside the host and are therefore not the major mode of transmission. The cyst form is capable of surviving for long periods of time in cool water supplies and has been responsible for major waterborne outbreaks in the United States.

LIFE CYCLE AND MORPHOLOGY

The life cycle of *Giardia lamblia* begins with the ingestion of the cyst form, which reaches the small bowel, and in response to the low pH, encysts into the motile, feeding trophozoite form. Using flagella and the ventral disk, the trophozoite moves from the small intestine to the duodenum and jejunum where it attaches to the intestinal epithelium. Trophozoites do not seem to penetrate the epithelium and may remain attached to the mucosal surface until it sloughs off the tip of the villus approximately every 72 hours (1). The trophozoites are pear-shaped and have a concave ventral disk, which they use for attachment. They have four pairs of flagella, two nuclei, two axonemes, and two median bodies and are 10 to 20 μm long and 5 to 15 μm wide (see Fig. 1).

Trophozoites encyst as they move down through the colon. During encystation, flagella are retracted, the cytoplasm becomes condensed, and the cyst wall is secreted. This process of encystation is a key step in the life cycle of *Giardia* that allows the organism to survive between hosts outside the human body. The cysts can survive for 3 months in water at 4 °C (2). They are usually oval-shaped and contain four nuclei, axonemes, and median bodies. They measure 11 to 14 μm in length and 7 to 10 μm wide (see Fig. 2).

CLINICAL DISEASE

The cyst form is infectious, and an inoculum of 10 to 100 cysts is required to infect humans (1,3). The incubation time to infection ranges from 12 to 20 days before the acute stage begins, which lasts only a few days. The organisms are present in the duodenal mucosa, but they do not cause

Previous Page

> The simple truth is that there's no "scientific controversy" over the safety of fluoridation. The practice is safe, economical, and beneficial. The survival of this fake controversy represents one of the major triumphs of quackery over science in our generation (2).

FLUORIDATION RATES AROUND THE WORLD

Tooth decay, or caries, is the most prevalent disease of all humanity, more common than the common cold (13). The cost of tooth decay in human suffering as well as economic expenditures is also enormous, accounting for many billions of dollars. Yet, as we enter the twenty-first century, only a small percentage of the world's population can easily drink fluoridated water. In addition, the majority of bottled water is not fluoridated, and many types of home water filtering devices actually remove the fluoride that may be in the water.

In the United States, 62.2% of the population has access to public water supplies that are optimally fluoridated (14). More than 360 million people worldwide in approximately 60 countries also drink fluoridated water. Dr. C. Everett Kopp, the former Surgeon General of the United States has stated, "Fluoridation is the single most important commitment that a community can make to the oral health of its citizens." It is imperative to continue to fluoridate water systems throughout the world.

BIBLIOGRAPHY

1. Richmond, V.L. (1985). Thirty years of fluoridation. *Amer. Jour. Clinical Nutrition* **41**: 129–138.
2. Anon. (July, August, 1978). Fluoridation: a two-part report. *Consumers Reports*.
3. World Health Organization. (1970). *Fluorides and Human Health*. Monograph 59, Geneva. World Health Organization.
4. National Academy of Sciences. (1980). *National Research Council Food and Nutrition Board. Recommended Daily Allowances*, 9th Edn. Washington, DC, pp. 156–59.
5. *American Dental Association* (1991). Facts about Fluoride.
6. Barrett, S., Jarvis, W.T., Kroger, M., and London, W.M. (1997). *Consumer Health: A Guide to Intelligent Decisions*, 6th Edn. Brown & Benchmark, p. 121.
7. Leonard, A., Deknudt, G.H., Decat, G., and Leonard, E.D. (1977). Cytogenic investigations of leucocytes of cattle intoxicated with fluoride. *Toxicology* **7**: 239–42.
8. Japello, G. and Lia, J.S. (1974). Sodium fluoride as potential mutagen in mammalian egg. *Arch. Environ. Health* **29**: 230–35.
9. Ripa, L., DePaola, P., Horowitz, H., Nowak, A. et al. (1986). *A Guide to the Use of Fluorides for the Prevention of Dental Caries*. American Dental Assoc.
10. Barrett, S., Jarvis, W.T., Kroger, M., and London, W.M. (1997). *Consumer Health: A guide to intelligent Decisions*, 6th Edn. Brown & Benchmark, Madison, WI.
11. Centers for Disease Contol. (1992). Fluoridation of community water systems. *JAMA* **267**: 3264–3265.
12. Weisfeld, V.D. (Ed.). *Preventing tooth decay: Results from a four-year national study. Special Report Number Two, 1983*. Robert Wood Johnson Foundation, Princeton, NJ.
13. Schissel, M.J. and Dodes, J.E. (1999). *Healthy Teeth: A User's Manual*. St. Martin's/Griffin, New York.
14. *American Dental Association*. (1999). Fluoridation Facts. Council on Access, Prevention and Interprofessional Relations, Chicago, IL.

GIARDIASIS

Christine L. Bean
University of New Hampshire
Durham, New Hampshire

GIARDIA LAMBLIA

Giardia lamblia is an intestinal flagellated protozoan, which infects the intestinal tract of a wide range of mammals, including humans. Two forms of this protozoan exist, a highly resistant cyst and a motile, feeding trophozoite. Cysts may be found in freshwater surface water supplies as a result of contamination from sewage and animal waste runoff. Trophozoites of this obligate parasite do not typically survive outside the host and are therefore not the major mode of transmission. The cyst form is capable of surviving for long periods of time in cool water supplies and has been responsible for major waterborne outbreaks in the United States.

LIFE CYCLE AND MORPHOLOGY

The life cycle of *Giardia lamblia* begins with the ingestion of the cyst form, which reaches the small bowel, and in response to the low pH, encysts into the motile, feeding trophozoite form. Using flagella and the ventral disk, the trophozoite moves from the small intestine to the duodenum and jejunum where it attaches to the intestinal epithelium. Trophozoites do not seem to penetrate the epithelium and may remain attached to the mucosal surface until it sloughs off the tip of the villus approximately every 72 hours (1). The trophozoites are pear-shaped and have a concave ventral disk, which they use for attachment. They have four pairs of flagella, two nuclei, two axonemes, and two median bodies and are 10 to 20 μm long and 5 to 15 μm wide (see Fig. 1).

Trophozoites encyst as they move down through the colon. During encystation, flagella are retracted, the cytoplasm becomes condensed, and the cyst wall is secreted. This process of encystation is a key step in the life cycle of *Giardia* that allows the organism to survive between hosts outside the human body. The cysts can survive for 3 months in water at 4 °C (2). They are usually oval-shaped and contain four nuclei, axonemes, and median bodies. They measure 11 to 14 μm in length and 7 to 10 μm wide (see Fig. 2).

CLINICAL DISEASE

The cyst form is infectious, and an inoculum of 10 to 100 cysts is required to infect humans (1,3). The incubation time to infection ranges from 12 to 20 days before the acute stage begins, which lasts only a few days. The organisms are present in the duodenal mucosa, but they do not cause

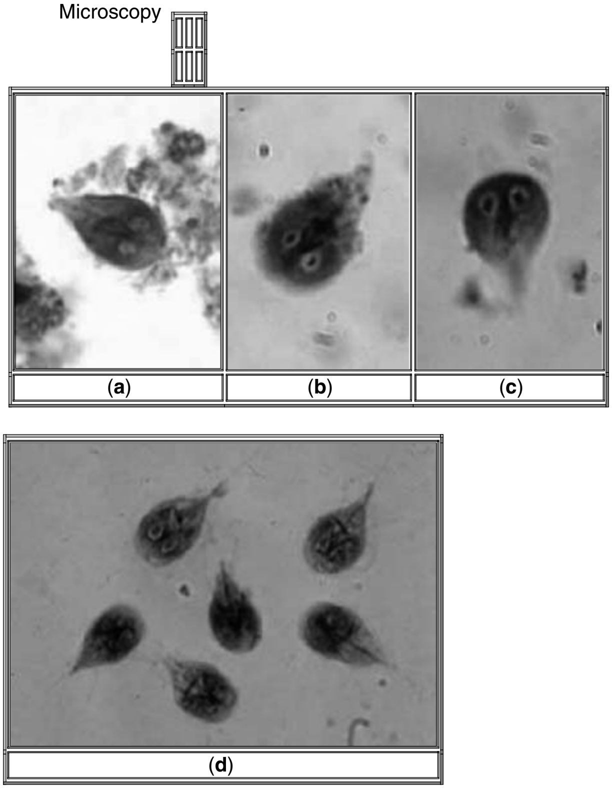

Figure 1. Three trophozoites of *Giardia intestinalis*, stained with trichrome (**a**) and stained with iron hematoxylin (**b** and **c**). Each cell has two nuclei with a large, central karyosome. Cell size: 9 to 21 μm in length. Trophozoites in culture (**d**). Photographs are courtesy of Centers for Disease Control & Prevention National Center for Infectious Diseases Division of Parasitic Diseases DPDx Laboratory Identification of Parasites of Public Health Concern.

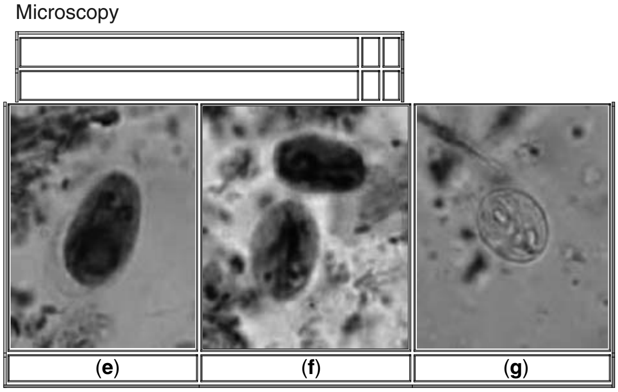

Figure 2. Cysts of *Giardia intestinalis*, stained with iron-hematoxylin (**e, f**) and in a wet mount (**g**) from a patient seen in Haiti). Size: 8 to 12 μm in length. These cysts have two nuclei each (more mature ones will have four). Photographs are courtesy of Centers for Disease Control & Prevention National Center for Infectious Diseases Division of Parasitic Diseases DPDx Laboratory Identification of Parasites of Public Health Concern.

pathology, as they do not penetrate the mucosa. Symptoms of the acute phase can include nausea, anorexia, fever, chills, and a sudden onset of explosive, foul-smelling diarrhea. Stools may have increased amounts of fecal fat and mucus, but no blood is present. Fat absorption may be disrupted because of the presence of trophozoites coating the mucosal lining. Rapid multiplication of trophozoites by binary fission creates a barrier between the intestinal epithelial cells and the lumen causing interference with nutrient absorption (1).

A chronic phase may follow the acute phase and induces symptoms of brief, recurrent diarrhea followed by normal stools or even constipation. Antigenic variation is a factor in the development of this chronic stage as certain surface antigens may allow the organism to survive better in the intestinal tract. Some other patients are asymptomatic and continue to pass cysts in stools.

Giardiasis occurs often in patients with hypogammaglobulinemia, but it has not become the problem it was expected to in AIDS patients. AIDS patients do not seem to be more susceptible to *Giardia* than do healthy people, and when cases occur, treatment is available and effective for this protozoan (1).

In summary, giardiasis causes infection ranging from an asymptomatic carrier state to severe malabsorption syndrome. Factors contributing to the variation in the effects of exposure include the strain of *Giardia,* dose, age of the host, and immune system function of the host.

DIAGNOSIS

Cysts and trophozoite forms are recovered by performing either a flotation or sedimentation procedure to concentrate them from fecal samples. Diagnosis is achieved by demonstrating either form in feces. Once the fecal sample is concentrated, a trichrome stain may be performed to enhance visualization of trophozoite and cyst forms. Fecal concentrates may be examined directly for cyst forms with iodine to enhance visualization. Direct fluorescent antibody (FA) or enzyme-linked immunosorbent assay (ELISA) techniques that incorporate monoclonal antibody-based reagents are also instruments for rapid diagnosis (1).

TREATMENT

Once an infection is diagnosed, treatment is recommended for all cases of *Giardia* for several reasons. First, the infection may cause subclinical malabsorption. The symptoms may be periodic, and if a carrier state develops, asymptomatic carriers are potential sources of infection for others with whom they may come in contact. The treatment of choice is metronidazole for all cases except during pregnancy. Other drugs are available and have been used by infected persons successfully, including paromomycin and tinidazole.

EPIDEMIOLOGY AND PREVENTION

Transmission of viable cyst forms can occur through many modes, including the ingestion of contaminated food or water, direct contact with an infected person such as children in day care centers, any group of people living in close quarters, or through oral/anal sexual practices, especially in male homosexuals. Areas with poor sanitation, such as those that travelers and/or campers may encounter, have resulted in giardiasis being referred to as "traveler's diarrhea." Also, some occupations increase the likelihood of exposure, including sewage and irrigation workers and clinical laboratory workers handling fecal samples. Waterborne outbreaks have occurred in the United States and have been associated with hikers and campers drinking stream water that was contaminated with fecal material from animal reservoirs, such as beavers.

Increased susceptibility has been demonstrated in patients with decreased gastric acidity and with malnutrition in young children in developing countries. The disease is more prevalent in children than in adults and is now reportable to the State Public Health Department and the Centers for Disease Control only in those patients aged 5 years and younger in the United States. The total cases reported in 2002 in the United States was 16,124 (4). Epidemics among adults in this country are seen in association with fecally contaminated food or water. Dogs, cats, and other animals can also carry *Giardia* infection.

To prevent transmission of this disease, personal hygiene must be addressed in all patient populations. Improved sanitation measures can be effective in controlling the person-to-person spread of infection. Travelers are advised to avoid drinking from local water supplies when traveling to foreign countries and to limit intake to bottled water. Iodine disinfection of drinking water as well as the use of filtration systems may serve to allow the decontamination of water sources for campers and travelers.

BIBLIOGRAPHY

1. Garcia, L. S. and Bruckner, D. (2002). *Diagnostic Medical Parasitology*, 4th Edn. American Society for Microbiology (ASM) Press, Washington, DC.
2. Olson, M.E., Goh, J., Phillips, M., Guselle, N., and McAllister, T. (1999). *Giardia* cyst and *Cryptosporidium* oocyst survival in water, soil and cattle feces. *J. Environ. Qual.* **28**: 1991–1996.
3. AWWA (American Water Works Association). (1999). *Waterborne Pathogens, Manual of Water Supply Practices*. AWWA M48 177–182.
4. MMWR (Morbidity and Mortality Weekly Report). (2003). Selected notifiable disease reports, United States, comparison of provisional 4-week totals October 11, 2003, with historical data. **52**(41): 998–1005.

GRAVITY SEPARATION/SEDIMENTATION

RASHEED AHMAD
Khafra Engineering Consultants
Atlanta, Georgia

Gravity separation or sedimentation is the step in conventional water treatment systems that follows flocculation and precedes filtration. Its purpose is to enhance the filtration process by removing particulate matter. Sedimentation requires that the water flow through a basin at a slow enough velocity to permit the particulate matter to settle to the bottom of the basin before the water exits the basin. The equipment required for this process includes a settling basin of rectangular, square, or circular configuration. The basin includes provisions for inlet and outlet structures and a sludge collection system. In addition, sedimentation systems are optionally equipped with tube or plate settlers to improve performance.

The settling velocity of particles or particulate matter is governed by particle size, shape, density, and water viscosity (which varies with temperature). Stokes' law describes the terminal velocity of discrete spherical particles in a laminar flow regime. In an ideal upflow rectangular settling tank (Figs. 1 and 2), all particles that have a settling velocity greater than the liquid upflow velocity will be captured. This upflow velocity is more commonly called the hydraulic overflow rate or surface loading rate. The surface overflow rate is the primary design parameter for sizing sedimentation basins. This rate is defined as the rate of inflow (Q) divided by the tank surface or floor area (A). Units are typically rated in gallons per day per square foot, gallons per minute per square foot, or cubic meter per hour per square meter. Design hydraulic overflow rates vary with the nature of the settling solids, water temperature, and the hydraulic characteristics of the sedimentation basin.

The settling efficiency for an ideal condition is independent of the depth and dependent on the tank plan

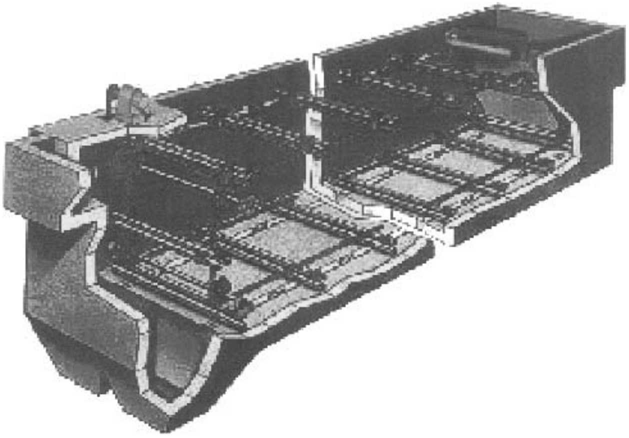

Figure 1. Rectangular sedimentation tank (*Source*: U.S. Filter).

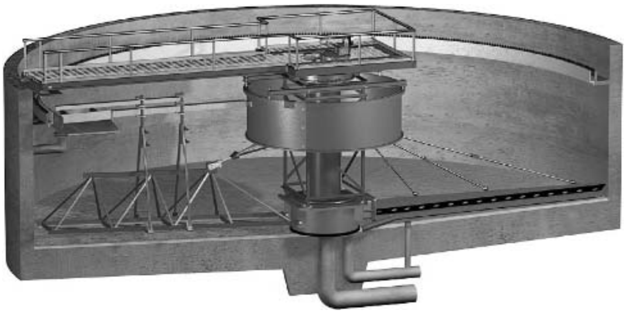

Figure 2. Circular sedimentation tank (*Source*: U.S. Filter).

or surface area. In reality, depth is important because it can affect flow stability if it is large and scouring if small. The basin depth affects flow-through velocity and plays a role in allowing greater opportunity for flocculant particle contact. Conventional settling facilities provided at larger water treatment plants are often long, narrow (4 or 5 to 1 length to width ratio) rectangular basins whose theoretical detention times are in the range of 1.5 to 3 hours at the design flow rate. Most regulatory agencies specify a minimum detention time and a maximum surface overflow rate. Typical sedimentation surface loading rates for long, rectangular tanks and circular tanks using alum coagulation vary from 500 to 1200 gallons per day per square feet (20 to 48 cubic meters per day per square meter). Regulatory agencies sometimes stipulate that weir rates should not exceed around 20,000 gallons per day per linear foot (250 cubic meters per day per meter) of weir. Flow-through velocities of 2 to 4 ft/min (0.6 to 1.2 m/min) are usually acceptable for basin depths of 7 to 14 ft (2.1 to 4.3 m).

Particulate matter settles out of a suspension in one of four different ways, depending on the concentration of the suspension and flocculating properties of the particles. The various settling regimes of particles are commonly referred to as types 1 to 4. In type 1 settling, also called discrete particle settling, the particles have little tendency to flocculate on contact with each other in a dilute suspension when settling. If the particles flocculate, the settling regime in a dilute suspension is identified as type 2 or flocculant settling. In hindered zone, or type 3 settling, the particle concentration causes interparticle effects, which might include flocculation, to the extent that the rate of settling is a function of solids concentration. Compression or type 4 settling develops under the layers of zone settling. The rate of compression depends on time and the force caused by the weight of solids above. In sedimentation, particulate matter removal is largely governed by types 1 and 2 settling. However, types 3 and 4 regimes are dominant in clarification and thickening processes.

High-rate settlers such as tube or plate settlers are often added to settling basins to improve their efficiency, especially if flows are to be increased beyond original design conditions (Figs. 3 and 4). The shallow settling depths and the large surface area provided by tube or plate settlers permit effective sedimentation at detention times of several minutes versus several hours in conventional sedimentation basins. The use of tube or plate settlers in new plant construction minimizes settling basin costs by minimizing the basin size necessary to attain a desired level of treatment. They are generally designed to accept overflow rates ranging from 1400 to 4000 gallons per day per sq foot (55 to 160 cubic meters per day per square meter).

Figure 3. Plate settlers for installation (*Source*: U.S. Filter).

Figure 4. Installed plate settlers (*Source*: U.S. Filter).

In wastewater practice, sedimentation tanks are used in primary and in secondary treatment. They are called primary and secondary clarifiers. For sludge treatment, thickeners are used whose design is also based on sedimentation principles. However, the dominant settling regimes in wastewater treatment are clarification and thickening.

READING LIST

ASCE and AWWA. (1998). *Water Treatment Plant Design*, 3rd Edn. McGraw-Hill, New York.

AWWARF. (1998). *Treatment Process Selection for Particle Removal*. American Water Works Association Research Foundation, Denver, CO.

AWWA. (1999). *Water Quality and Treatment*, 5th Edn. McGraw-Hill, New York.

Hudson, H.E. (1981). *Water Clarification Processes*. Van Nostrand Reinhold, New York.

Kawamura, S. (2000). *Integrated Design of Water Treatment Facilities*. John Wiley & Sons, New York.

Qasim, S.R., Motley, E.M., and Zhu, G. (2000). *Water Works Engineering*. Prentice-Hall, Upper Saddle River, NJ.

USEPA. (1990). *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*. Office of Drinking Water, Center for Research Information, Cincinnati, OH.

WATER HAMMER

Zacharia Michael Lahlou
Technical Assistance Consultant

Water hammer refers to fluctuations caused by a sudden increase or decrease in flow velocity. These pressure fluctuations can be severe enough to rupture a water main. Potential water hammer problems should be considered when pipeline design is evaluated, and a thorough surge analysis should be undertaken, in many instances, to avoid costly malfunctions in a distribution system. Every major system design change or operation change—such as the demand for higher flow rates—should include consideration of potential water hammer problems. This phenomenon and its significance to both the design and operation of water systems is not widely understood, as evidenced by the number and frequency of failures caused by water hammer.

WHAT IS WATER HAMMER?

Water hammer (or hydraulic shock) is the momentary increase in pressure, which occurs in a water system when there is a sudden change of direction or velocity of the water. When a rapidly closed valve suddenly stops water flowing in a pipeline, pressure energy is transferred to the valve and pipe wall. Shock waves are set up within the system (Fig. 1). Pressure waves travel backward until encountering the next solid obstacle, then forward, then back again (Fig. 2). The pressure wave's velocity is equal to the speed of the sound; therefore it "bangs" as it travels back and forth, until dissipated by friction losses. Anyone who has lived in an older house is familiar with the "bang" that resounds through the pipes when a faucet is suddenly closed. This is an effect of water hammer.

A less severe form of hammer is called surge, a slow motion mass oscillation of water caused by internal pressure fluctuations in the system. This can be pictured as a slower "wave" of pressure building within the system. Both water hammer and surge are referred to as transient pressures. If not controlled, they both yield the same results: damage to pipes, fittings, and valves, causing leaks and shortening the life of the system. Neither the pipe nor the water will compress to absorb the shock.

INVESTIGATING THE CAUSES OF WATER HAMMER

A water transport system's operating conditions are almost never at a steady state. Pressures and flows change continually as pumps start and stop, demand fluctuates, and tank levels change. In addition to these normal events, unforeseen events, such as power outages and equipment malfunctions, can sharply change the operating conditions of a system. Any change in liquid flow rate, regardless of the rate or magnitude of change, requires that the liquid be accelerated or decelerated from its initial flow velocity. Rapid changes in flow rate require large forces that are seen as large pressures, which cause water hammer.

Entrained air or temperature changes of the water also can cause excess pressure in the water lines. Air trapped in the line will compress and will exert extra pressure on the water. Temperature changes will actually cause the water to expand or contract, also affecting pressure. The maximum pressures experienced in a piping system are frequently the result of vapor column separation, which is caused by the formation of void packets of vapor when pressure drops so low that the liquid boils or vaporizes. Damaging pressures can occur when these cavities collapse.

The causes of water hammer are varied. There are, however, four common events that typically induce large changes in pressure:

1. Pump startup can induce the rapid collapse of a void space that exists downstream from a starting pump. This generates high pressures.
2. Pump power failure can create a rapid change in flow, which causes a pressure upsurge on the suction side and a pressure downsurge on the discharge side. The downsurge is usually the major problem. The pressure on the discharge side reaches vapor pressure, resulting in vapor column separation.
3. Valve opening and closing is fundamental to safe pipeline operation. Closing a valve at the downstream end of a pipeline creates a pressure wave that moves toward the reservoir. Closing a valve in less time than it takes for the pressure surge to travel to the end of the pipeline and back is called "sudden valve closure." Sudden valve closure will change velocity quickly and can result in a pressure surge. The pressure surge resulting from a sudden valve opening is usually not as excessive.
4. Improper operation or incorporation of surge protection devices can do more harm than good. An example is oversizing the surge relief valve or improperly selecting the vacuum breaker-air relief valve. Another example is to try to incorporate some means of preventing water hammer when it may not be a problem.

FINDING PRACTICAL SOLUTIONS

The surge pressure must be incorporated with the operating pressure in the design of the pipe. The recommendations and requirements regarding allowances for surge pressure are given in the American Water Works (AWWA) standards and manuals for water supply practice, and vary depending on the type of pipe used. The following tools reduce the effects of water hammer.

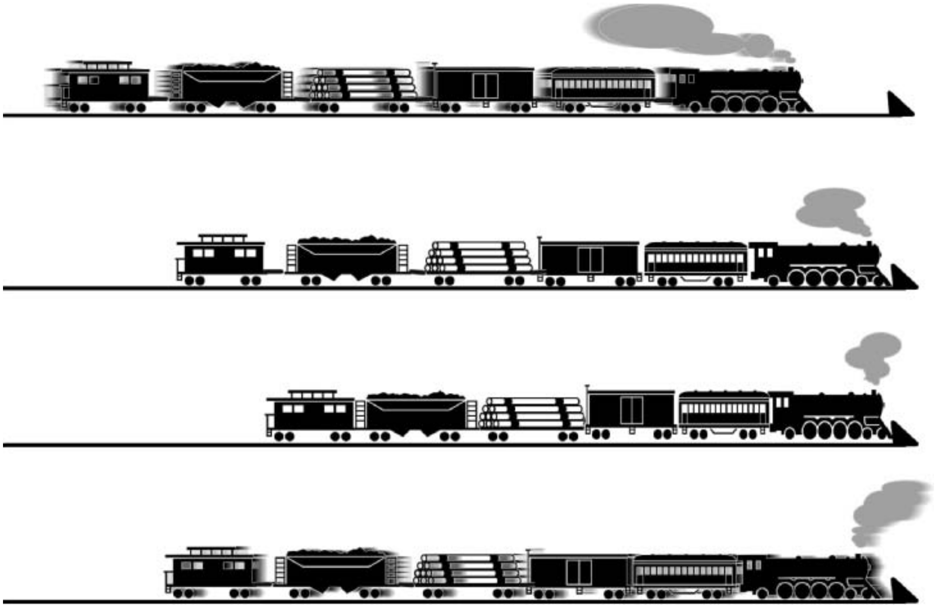

Figure 1. Illustration of water hammer. A column of water acts like a freight train suddenly stopping when an outlet valve is suddenly closed. *Source*: Pickford, John. 1969. *Analysis of Water Surge.* Gordon and Breach Science Publishers.

VALVES

Water hammer often damages centrifugal pumps when electrical power fails. In this situation, the best form of prevention is to have automatically-controlled valves, which close slowly. (These valves do the job without electricity or batteries. The direction of the flow controls them.) Closing the valve slowly can moderate the rise in the pressure when the downsurge wave—resulting from the valve closing—returns from the reservoir.

Entrained air or temperature changes of the water can be controlled by pressure relief valves, which are set to open with excess pressure in the line and then closed when pressure drops. Relief valves are commonly used in pump stations to control pressure surges and to protect the pump station. These valves can be an effective method of controlling transients. However, they must be properly sized and selected to perform the task for which they are intended without producing side effects.

If pressure may drop at high points, an air and vacuum relief valve should be used. All downhill runs where pressure may fall very low should be protected with vacuum relief valves. Vacuum breaker-air release valves, if properly sized and selected, can be the least expensive means of protecting a piping system. A vacuum breaker valve should be large enough to admit sufficient quantities of air during a downsurge so that the pressure in the pipeline does not drop too low. However, it should not be so large that it contains an unnecessarily large volume of air, because this air will have to be vented slowly, increasing the downtime of the system. The sizing of air release valves is, as mentioned, critical.

PUMP

Pump startup problems can usually be avoided by increasing the flow slowly to collapse or flush out the voids gently. Also, a simple means of reducing hydraulic surge pressure is to keep pipeline velocities low. This not only results in lower surge pressures, but results in lower drive horsepower and, thus, maximum operating economy.

SURGE TANK

In long pipelines, surge can be relieved with a tank of water directly connected to the pipeline called a "surge tank." When surge is encountered, the tank will act to relieve the pressure, and can store excess liquid, giving the flow alternative storage better than that provided by expansion of the pipe wall and compression of the fluid. Surge tanks can serve for both positive and negative pressure fluctuations. These surge tanks can also be designed to supply fluid to the system during a downsurge, thereby preventing or minimizing vapor column separation. However, surge tanks may be an expensive surge control device.

AIR CHAMBER

Air chambers are installed in areas where water hammer is encountered frequently, and are typically seen behind sink and tub fixtures. Shaped like thin, upside-down bottles with a small orifice connection to the pipe, they are air-filled. The air compresses to absorb the shock, protecting the fixture and piping.

CONCLUSION

Water hammer will continue to challenge engineers, operators, and managers of water systems because it is associated with systems that cannot be exactly defined due to the size and length of the water distribution system with ondulating profile or the lack of definition of the system components such as valves or pumps. There is a need for a more practical approach while research continues

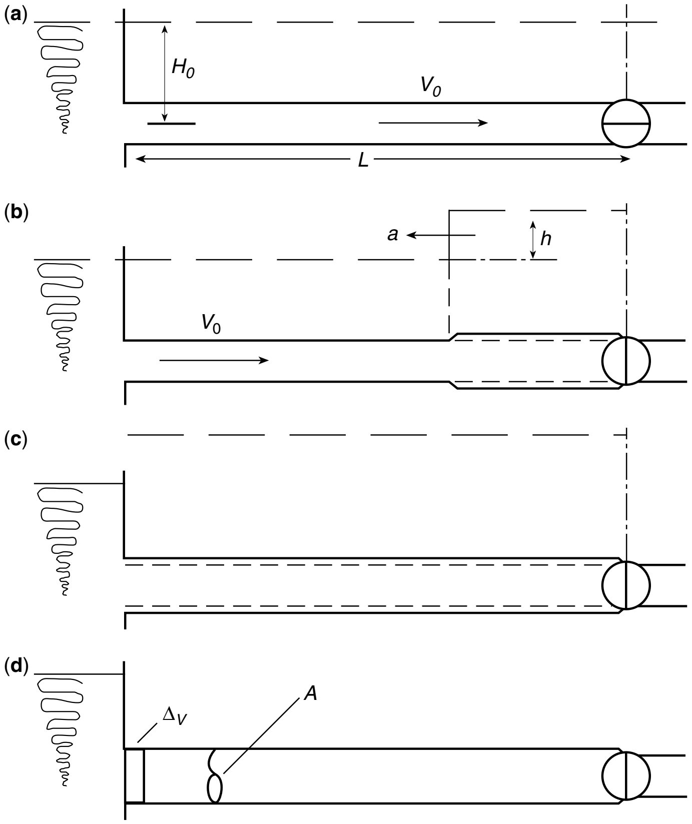

(a)
Valve in open position.

(b)
Valve is shut. A pressure wave moves upstream with velocity "a". At the same time water still enters the pipe with velocity V_0.

(c)
Wave front continues upstream until it reaches the end taking time L/a to reach there. The time 2L/a is known as the "period" of the pipe μ.

(d)
If the total quantity of water that enters the pipe during this time 1/2 μ is ΔV then because it is moving with velocity V_0, $\Delta V = 1/2\ V_0 A\mu$. This whole extra volume of water is occupying the space ($A \times L$) where A is the cross section of the pipe. The resulting increase in pressure or "water hammer" pressure is $P = \mathrm{K}\,\frac{\Delta V}{\Delta L}$.

Figure 2. Sudden valve closing. *Source*: Pickford, John. 1969. *Analysis of Water Surge.* Gordon and Breach Science Publishers.

to provide better descriptions of the physics of water hammer and for useful computational solutions including those basics.

WHERE CAN I FIND MORE INFORMATION?

Kroon, J. R., M. A. Stoner, and W. A. Hunt. 1984. “Water Hammer: Causes and Effects.” *Journal of the American Water Works Association.* 76: 39–45.

National Drinking Water Clearinghouse. 2001. “Ask the Experts.” *On Tap.* Vol. 1, Issue 3: 10–11.

Parmakian, J. 1963. *Waterhammer Analysis.* Dover Publications.

Sharp, B.B. and D. B. Sharp. 1996. *Water Hammer: Practical Solutions.* New York: Halsted Press.

Weis, F. 1996. “Dispelling Common Misconceptions about Water Hammer.” *Water Engineering and Management.* 143: 24–30.

Wood, D. J. 2002. SURGE2000 Software. (Modeling water hammer in pipes and a wide range of hydraulic and surge protection devices are addressed). Civil Engineering Software Center, University of Kentucky Lexington, KY.

Z. Michael Lahlou holds a doctorate in environmental and natural resource economics and a master’s in civil and environmental engineering. Formerly the technical assistance coordinator for the National Drinking Water Clearinghouse, Lahlou now resides in Huntington Beach, California.

HAVE YOU READ ALL OUR TECH BRIEFS?

Tech Briefs, drinking water treatment and supply fact sheets, have been a regular feature in the National Drinking Water Clearinghouse (NDWC) publication *On Tap* for more than six years.

A package of *Tech Briefs* is now available as a product. A three-ring binder holds all the current *Tech Briefs* in print. New selections can easily be added to the package as they become available.

To order this free product, call the NDWC at the numbers listed at the bottom of this ad and ask for item #DWPKPE71. Additional copies of fact sheets are also free; however, postal charges may be added. You can also order copies of one or all of the free *Tech Briefs* listed below.

Tech Brief: Organics Removal, item #DWBLPE59;
Tech Brief: Package Plants, item #DWBLPE63;
Tech Brief: Water Treatment Plant Residuals Management, item #DWBLPE65;
Tech Brief: Lime Softening, item #DWBLPE67;
Tech Brief: Iron and Manganese Removal, item #DWBLPE70;

Water Conservation Measures Fact Sheet, item #DWBLPE74;

Tech Brief: Membrane Filtration, item #DWBLPE81;

Tech Brief: Treatment Technologies for Small Drinking Water Systems, item #DWPSPE82;

Tech Brief: Ozone, item #DWBLPE84;

Tech Brief: Radionuclides, item #DWBLPE84;

Tech Brief: Slow Sand Filtration, item #DWBLPE99;

Tech Brief: Ultraviolet Disinfection, item #DWBLPE101;

Tech Brief: Leak Detection and Water Loss Control, item #DWBLPE102;

Tech Brief: Diatomaceous Earth Filtration for Drinking Water, item #DWBLPE108.

Tech Brief: Reservoirs, Towers, and Tanks, item #DWFSOM15

Tech Brief: System Control and Data Acquisition (SCADA), item #DWFSOM20

Tech Brief: Valves, item #DWFSOM21

Tech Brief: Water Quality in Distribution Systems, item #DWFSOM25

HEALTH EFFECTS OF COMMONLY OCCURRING DISINFECTION BYPRODUCTS IN MUNICIPAL WATER SUPPLIES

Raghuraman Venkatapathy
Oak Ridge Institute for Science and Education
Cincinnati, Ohio

Chandrika Moudgal
Brenda Boutin
Robert Bruce
National Center for Environmental Assessment
Cincinnati, Ohio

Disinfectants, such as chlorine, chlorine dioxide, chloramine, and ozone, have been used in drinking water treatment plants throughout the world to disinfect source waters containing disease-causing pathogens. The noteworthy biocidal properties of the disinfectants are somewhat offset by the formation of disinfection byproducts (DBPs) through reactions between the disinfectants and the dissolved organic and inorganic chemicals in raw water.

From the present knowledge of occurrence and health effects, the DBPs of most interest are halomethanes, including trihalomethanes (THMs), haloacetic acids (HAAs), bromate, chlorate, and chlorite, followed by haloacetonitriles, haloacetaldehydes, haloketones, and halonitromethanes (1). In addition, a number of epidemiological studies exist that indicate an association between chronic ingestion of drinking water containing DBPs and increased risk of rectal, bladder, or colon cancer in humans (2–5), but these studies cannot provide information on whether the observed health effects are because of one or more of the hundreds of other DBPs that are also present in drinking water.

Health effects on humans (if available) and animals including general, immunological, neurological, reproductive, and developmental toxicity; mutagenicity [based on the *Salmonella typhimurium* assay (6)]; and carcinogenicity associated with oral exposure to some commonly occurring DBPs are summarized in the sections below. A majority of the health effects observed in these studies were because of exposure to single chemicals at dose levels much above those typically found in drinking water. Hence, health effects associated with drinking water, which typically contains mixtures of DBPs, may be different from those mentioned below.

HALOMETHANES

Data on the health effects of non-THMs are available to a limited extent in the literature. However, THMs, which have been fairly well studied in the literature, are known to induce cytotoxicity in the liver and kidneys of rodents, and to cause cancer in the liver, kidneys, and the large intestine. Most THMs do not induce reproductive or developmental toxicity. In addition, halomethanes are considered to be weak mutagens. Health effects of halomethanes are summarized below.

Chloromethane

Although inhalation of chloromethane may cause damage to the lung and kidneys, no data are currently available for carcinogenic, reproductive, developmental, neurological, and immunotoxicity endpoints because of oral exposures in humans or animals. Chloromethane tested positive for mutagenicity in the *Salmonella typhimurium* assay (7).

Bromomethane

As in the case of chloromethane, direct inhalation of bromomethane may cause seizures and kidney and nerve damage (8). No data exist on health effects because of human oral exposure in the literature. However, limited evidence exists that chronic exposure to bromomethane may occasionally lead to neoplasia in the forestomach at high concentrations in rats (9). At concentrations of 0.4 mg/kg/day (∼40 mg/L) or less, no evidence existed of significant adverse effects in rats (9). Bromomethane tested positive in the *Salmonella* mutagenicity assay (10).

Bromochloromethane

No data on the health effects of oral exposure to bromochloromethane in humans or animals are available in the literature. The only studies that were available in the literature were the *Salmonella* mutagenicity results, which were consistently positive for bromochloromethane (11).

Dibromomethane

No data on the health effects of oral exposure to dibromomethane in humans or animals are available in the literature. Various studies showed a positive response for mutagenicity as measured by the *Salmonella typhimurium* assay (12).

Chloroform

Limited data are available in the literature on death because of oral exposure to pure chloroform in both humans (13) and animals (14–17). Various studies have indicated the liver and kidney as the primary target organs of chloroform toxicity and have observed reversibility in toxic effects on cessation of exposure in both humans (13,18,19) and animals (20,21). However, the doses required to cause toxicity in the liver and kidney were high for both humans and animals. No indications of renal effects were observed in humans (22) and animals (21) because of ingestion of low doses of chloroform. No studies existed on the effect of chloroform exposure on immunological effects in the humans, whereas a few studies documented a reduction in immunity levels in animals exposed to chloroform (23,24). With respect to neurological effects, humans exposed to high concentrations of chloroform immediately fell into a deep coma (13,19), whereas the central nervous system was affected in animals (25,26). An epidemiologic study on drinking water consumption by the population of a small town in northern New Jersey between 1985 and 1988 noticed low birth weight among term babies (27). However, in addition to predominant amounts of chloroform in the drinking water, other THMs were also present in the town's water supply. An equivalent study on animals demonstrated increased resorptions among fetuses at high doses of chloroform in drinking water (20). Other epidemiologic studies have suggested an association between cancer in humans and consumption of chlorinated drinking water predominantly containing chloroform (28–31). However, no concrete evidence exists linking chloroform specifically to the increased incidence of cancer in those studies. In experiments with animals, chloroform has been reported to be carcinogenic in several chronic animal bioassays, with significant increases in the incidence of liver tumors in male and female mice and significant increases in the incidence of kidney tumors in male rats and mice (32,33).

Bromodichloromethane

No studies on the human health effects of bromodichloromethane because of oral exposure were available in the literature. In animals with acute exposures to bromodichloromethane, typical pathological changes observed include fatty infiltration of liver and hemorrhagic lesions in kidney, adrenals, lung, and brain (26). Animals orally exposed to bromodichloromethane show liver and kidney damage followed by death within 2 weeks. Typical signs include increased liver weight, pale discoloration of liver tissue, increased levels of hepatic tissue enzymes in serum, decreased levels of secreted hepatic proteins in blood, and focal areas of inflammation or degeneration (24,34–36). Rats and mice administered oral doses of 150 to 600 mg/kg often displayed acute signs of central nervous system depression, including lethargy, labored breathing, sedation, and flaccid muscle tone, but the signs were reversible several hours after cessation of exposure (25,34,37,38). Ruddick et al. (36) reported fetal and maternal toxicity effects on rats exposed to bromodichloromethane during gestation. With respect to cancer, several epidemiological studies have indicated that an association may exist between ingestion of chlorinated drinking water containing bromodichloromethane and increased risk of cancer in humans (4,5,39), but those studies could not assign the cause of cancer specifically to bromodichloromethane. However, chronic oral studies in animals provide evidence of liver and kidney tumors because of bromodichloromethane exposure (34,40,41).

Dibromochloromethane

No studies exist on human health effects because of oral exposure of dibromochloromethane in the literature. Significant increases in animal mortality have been reported because of acute, subchronic, and chronic oral exposure to dibromochloromethane in the literature (16,26,38,42,43). Typical sites of toxicity in animals exposed to dibromochloromethane include the central nervous system, liver, and kidney (26,43). Oral exposure to dibromochloromethane results in an accumulation of fat in the liver, hepatocyte vacuoles, alterations in serum cholesterol levels, and decreases in serum triglyceride levels followed by liver necrosis at higher doses (16,24,43,44). In addition, tubular regeneration and mineralization leading to nephrosis has been observed in both rats and mice (43). A few studies have examined the potential association between adverse human reproductive outcomes, including spontaneous abortions, stillbirths, and preterm delivery, and the consumption of municipal drinking water containing a mixture of THMs (27,45–53), but they have not been able to implicate dibromochloromethane specifically to the reproductive effects. However, animal studies show an association between reproductive effects and dibromochloromethane in drinking water at high doses. For example, female mice exposed to dibromochloromethane experienced a marked reduction in fertility, with significant decreases in litter size, body weight of the dam and the litter, gestational survival, and postnatal survival (54). Various studies have reported increased incidences of hepatocellular adenomas and carcinomas in mice exposed to dibromochloromethane [summarized in (43)], leading the EPA to classify the chemical as a possible human carcinogen (55).

Bromoform

Several studies in the literature have documented deaths of children because of accidental overdoses of bromoform (when used as a sedative) leading to severe central nervous system depression and respiratory failure (56–58). As in the case of dibromochloromethane, the central nervous system and the liver seem to be the primary sites for bromoform-induced toxicity in animals, leading to liver necrosis and slight nephrosis at higher doses (16,24,43,44). Animal studies that examined the potential of bromoform to induce histological alterations and impair reproductive function indicate that the chemical does not induce any reproductive or developmental effects (59,60). With respect to cancer, an increased incidence of tumors and neoplastic lesions in the large intestines after oral administration of bromoform in rats and intraperitoneal administration in mice was observed, leading the EPA to classify the chemical as a probable human carcinogen (61).

Iodinated THMs

Recently, several iodinated THMs, such as dichloroiodomethane, bromochloroiodomethane, dibromoiodomethane, chlorodiiodomethane, bromodiiodomethane, and iodoform, have been identified in finished water at a few water treatment plants across the United States (1). No human or animal toxicity information on any of the iodinated THMs are available in the literature except for a positive response in the *Salmonella* mutagenicity assay and a negative response for carcinogenicity in mice for iodoform (62).

HALOACETIC ACIDS

Most haloacetic acids of concern have carcinogenic, reproductive, and developmental effects, whereas neurotoxic effects are significant when high doses are used for therapeutic purposes. Carcinogenic effects seem to be limited to the liver at high doses. The health effects of the HAAs are summarized below.

Chloroacetic Acid

Acute exposure of chloroacetic acid by humans may lead to interference with essential enzyme systems in the body, leading to intestinal perforation and peritonitis (63). Rodents exposed to chloroacetic acid orally or by gavage exhibited neurological dysfunction, damage to the respiratory tract including inflammation of the lung and nasal mucosa, metaplasia of the olfactory epithelium, and squamous cell hyperplasia of the forestomach (33,64). In addition, chronic exposure to chloroacetic acid via gavage resulted in myocarditis and mortality because of myocardial failure in rats and hepatic vacuolar degeneration in mice (33,64). No information on the reproductive or developmental effects of chloroacetic acid in animals or humans is available in the literature. With respect to cancer, no significant increase in tumor incidences were reported in studies involving rats and mice exposed to chloroacetic acid (64). In addition, chloroacetic acid was not mutagenic in the *Salmonella typhimurium* assay (64).

Bromoacetic Acid

No studies exist on human health effects because of oral exposure to bromoacetic acid in the literature. Acute exposure of rats to bromoacetic acid lead to various clinical symptoms including excess drinking water intake, hypomobility, labored breathing, and diarrhea (65). No adverse reproductive effects were observed in a spermatotoxicity study involving male rats (65), but developmental effects were observed at high doses in rats (66). Carcinogenic effects of bromoacetic acid have not been evaluated, but the chemical was found to be positive for mutagenicity in the *Salmonella typhimurium* assay (67).

Dichloroacetic Acid

Dichloroacetic acid has been shown to have some beneficial effects in curing a variety of metabolic diseases in humans (68–72). No evidence exists of adverse effects because of acute exposures or exposures at low concentrations in humans. However, limited human experiments coupled with animal studies have indicated that dichloroacetic acid is neurotoxic to humans (70,73). Animals exposed to dichloroacetic acid have been shown to produce developmental, reproductive, neural, and hepatic effects at high doses. Short-term exposure to dichloroacetic acid produced hind-limb paralysis at the highest dose in rats, whereas a significant increase occurred in relative liver and kidney weights and histopathological changes in the brain and testes of dogs (74,75). Reproductive effects include testicular toxicity manifested by distorted sperm heads, fused sperm, and reduced sperm counts in rats exposed to high doses of dichloroacetic acid (75,76), whereas a degeneration of the testicular epithelium and syncytial giant cell formation was noticed in dogs, even at low doses (74). Developmental effects include soft tissue abnormalities in the fetus, interventricular septal defects between the aorta and right ventricle, and urogenital defects in rats (77,78). With respect to cancer, dichloroacetic acid has been shown to be a very effective inducer of hepatic adenomas and carcinomas in both mice and rats at high doses (79–83), thus prompting the EPA to classify the chemical as possibly carcinogenic to humans (84).

Bromochloroacetic Acid

No studies exist on human health effects because of oral exposure to bromochloroacetic acid in the literature. No treatment-related differences in body weight, food and water consumption, clinical observations, or mortality were reported in rats exposed to bromochloroacetic acid for 14 days (85). However, mice showed liver effects including hepatomegaly, glycogen accumulation, and cytomegaly (86). No reproductive, developmental, or embryo toxicity effects were observed in rats treated with low doses of bromochloroacetic acid in a 35-day study (85). However, bromochloroacetic acid at high doses (>600 ppm) produced a consistent decrease in water consumption in both sexes of rats and affected reproductive function in females of both rats and mice (87,88). No data on carcinogenicity were available in the literature; however, bromochloroacetic acid was shown to be positive for mutagenicity in the *Salmonella typhimurium* assay (89).

Dibromoacetic Acid

Acute exposures of high doses of dibromoacetic acid by rats lead to several clinical symptoms including excess drinking water intake, hypomobility, labored breathing, diarrhea, and ataxia (65). In addition, histopathological examination of epididymal sperm in surviving animals showed the presence of misshapen and degenerating sperm, and abnormal retention of spermatids (65,90). Male mice exposed to dibromoacetic acid showed an increase in absolute and relative liver weights at doses greater than 125 mg/kg/d (91), whereas female mice showed body weight gain, decrease in thymus weights at 236 mg/kg/d, and increased liver and kidney weights at all tested doses (92). Reproductive effects because of dibromoacetic acid exposure in male rats include adverse effects on caput-sperm count, sperm morphology, and sperm motility, but only marginal changes in reproductive organ weights (90) and decreased fertility as measured by the number of copulations, number of litters, and number of implants and fetuses per litter (93). Immunological effects because of dibromoacetic acid exposure in animals include changes in spleen cell numbers and spleen cell populations, as well as alteration in natural killer cell activity and an antibody-forming cell response to sheep erythrocytes (92). No data on carcinogenicity were available in the literature; however, dibromoacetic acid was shown to be positive for mutagenicity in the *Salmonella typhimurium* assay (93).

Trichloroacetic Acid

Acute exposures to trichloroacetic acid by humans can cause severe burning in the mouth, pharynx, and abdomen, and low blood pressure followed by vomiting and diarrhea (94). Animals exposed to high doses of trichloroacetic acid showed liver and kidney effects, including significant increases in relative kidney weight and hepatic peroxisomal beta oxidation activity, and focal areas of hepatocellular hypertrophy and narcosis (95). At lower doses typical of concentrations in drinking water, trichloroacetic acid has been shown to seriously impair water and food consumption (79,81,96,97) and cause liver effects (79,97–99) in experimental animals. Trichloroacetic acid has also been shown to produce reproductive and developmental effects at very high doses in experimental animals, but no effects have been observed at drinking water concentration levels (100–102). With respect to cancer, trichloroacetic acid induces hepatocellular carcinomas, dose-related increases in the incidence of malignant tumors, and precancerous lesions when administered in drinking water to male B6C3F1 mice (79,82,103), but not in rats (104). In addition, trichloroacetic acid has been shown to be negative for mutagenicity in the *Salmonella typhimurium* assay (105).

Bromodichloroacetic Acid

No studies exist on the health effects on humans, and limited studies exist on the health effects on animals, because of oral exposure to bromodichloroacetic acid in the literature. Narotsky et al. (87) reported no maternal toxicity but some developmental toxicity (reduced pup weights) in mice at the highest dose tested. In the only other study (unpublished) available, Bull reported the observance of liver and lung tumors in mice exposed to bromodichloroacetic acid (106). In addition, bromodichloroacetic acid was shown to be weakly positive for mutagenicity in the *Salmonella typhimurium* assay (107).

Dibromochloroacetic Acid

No studies exist on the health effects on humans, and one short-term reproductive and developmental study exists on animals, because of oral exposure to dibromochloroacetic acid in the literature. Results from the reproductive study on rats indicate that although dibromochloroacetic acid produced consistent decreases in food and water consumption in both sexes at doses above 1000 ppm, it did not result in any female reproductive toxicity or any visceral malformation or variations in any pups (108).

Tribromoacetic Acid

No studies exist on the health effects on humans, and one short-term reproductive and developmental study exists on animals, because of oral exposure to tribromoacetic acid in the literature. Results from this reproductive study indicate that tribromoacetic acid at up to 400 ppm marginally reduced water consumption, but it did not affect reproductive function or produce general toxicity in rats of either sex (109). In addition, tribromoacetic acid was shown to be positive for mutagenicity in the *Salmonella typhimurium* assay (110).

HALOACETONITRILES

The toxicological effects of haloacetonitriles have not been well studied to date. Although very limited studies exist on the carcinogenicity of this group of chemicals, results of *Salmonella typhimurium* mutagenicity assays are available in the literature. The known health effects of haloacetonitriles are summarized below.

Chloroacetonitrile

No human health studies because of oral chloroacetonitrile exposure are available in the literature. Among the haloacetonitriles, chloroacetonitrile was shown to be metabolized to cyanide in rats and excreted as thiocyanate to the greatest extent (111). Chloroacetonitrile was shown to induce cytotoxic effects and oxidative stress in cultured rat gastric epithelial cells *in vitro* (112). In rat studies, slight decreases in maternal weight gain occurred during the treatment period and in the birth weights of pups born to rats given 55-mg/kg chloroacetonitrile orally on gestation days 7–21 (113). No other evidence of developmental or reproductive effects was reported in the literature. The IARC evaluated the carcinogenic potential of chloroacetonitrile and concluded that inadequate evidence exists for classification purposes, and they were hence assigned to group 3: The agent is not classifiable as to its carcinogenicity to humans (113,114). In addition, chloroacetonitrile was shown to be negative for mutagenicity in the *Salmonella typhimurium* assay (115).

Bromoacetonitrile

No human health studies because of oral bromoacetonitrile exposure are available in the literature. Rats exposed to high doses of bromoacetonitrile (>500 ppm) in drinking water resulted in significant reductions in body weight and reductions in feed and water consumption (116). At lower doses, study results indicated that a significant aversion to bromoacetonitrile in the water existed and suggested that the chemical may be a possible mild renal toxicant at 100 ppm as evidenced by an increase in blood urea nitrogen, and a potential reproductive toxicant at the same dose as evidenced by increased post-implantation loss (116). No studies on the carcinogenicity potential of bromoacetonitrile are available in the literature. However, bromoacetonitrile was shown to be negative for mutagenicity in the *Salmonella typhimurium* assay (117).

Dichloroacetonitrile

No human health studies because of oral dichloroacetonitrile exposure are available in the literature. In rats exposed to dichloroacetonitrile in corn oil by gavage, increased mortality was observed at 33-mg/kg/d dose, whereas decreased body weight, lower organ weight, and organ-to-body weight ratios were observed for spleen and gonads at higher doses (118). However, no serum chemistry changes indicative of adverse effects were seen at sublethal doses (118). Reproductive effects in rats because of dichloroacetonitrile exposure by gavage during gestation include significantly reduced percentage of females delivering litters, increased resorption rates, reduced mean birth weight of pups, and reduced maternal weight gain (119). Developmental effects in rats exposed to dichloroacetonitrile resulted in soft tissue anomalies, including an intraventricular septal defect in the heart, hydronephrosis, fused ureters and cryptorchidism, and skeletal abnormalities such as fused and cervical ribs (120). The IARC evaluated the carcinogenic potential of dichloroacetonitrile and concluded that inadequate evidence exists for classification purposes, and they were hence assigned to group 3: The agent is not classifiable as to its carcinogenicity to humans (113,114). However, dichloroacetonitrile was shown to be positive for mutagenicity in the *Salmonella typhimurium* assay (121).

Bromochloroacetonitrile

No human health effect studies because of oral bromochloroacetonitrile exposure are available in the literature. In rats exposed to bromochloroacetonitrile in corn oil by gavage, reduced mean birth weight of pups and reduced postnatal weight gain were observed to a much lesser extent than rats exposed to dichloroacetonitrile (119). Christ et al. (122) reported maternal toxicity and urogenital, skeletal malformations at doses above 65 mg/kg/d and total litter loss at 45 mg/kg/d in rats, whereas lower doses resulted in shorter fetal crown-rump lengths, reduced fetal weights, and increased frequency of cardiovascular malformations. Female mice treated with bromochloroacetonitrile showed a significant increase in the incidence of lung adenomas; however, the authors cautioned that a large variation in background incidence of lung tumors was also observed in the same species (123). However, bromochloroacetonitrile was shown to be positive for mutagenicity in the *Salmonella typhimurium* assay (124).

Dibromoacetonitrile

No human health effect studies because of oral dibromoacetonitrile exposure are available in the literature. In rats exposed to dibromoacetonitrile in corn oil by gavage, increased mortality was observed at 45-mg/kg/d dose (118). However, no serum chemistry changes indicative of adverse effects were seen at sublethal doses (118). Reproductive effects in rats exposed to dibromoacetonitrile were similar to effects observed in rats exposed to bromochloroacetonitrile (119). In addition, exposure to dibromoacetic acid reduced water consumption in male and female rats exposed to doses greater than 50 ppm; however, no other reproductive effects were observed in the animals (125). Female mice treated with dibromoacetonitrile showed only a marginal increase in the incidence of lung adenomas (123). In addition, dibromoacetonitrile was shown to be weakly positive for mutagenicity in the *Salmonella typhimurium* assay (126).

Trichloroacetonitrile

No human health effect studies because of oral trichloroacetonitrile exposure are available in the literature. Reproductive effects in rats because of trichloroacetonitrile exposure by gavage during gestation include a significantly reduced percentage of females delivering litters, increased resorption rates, reduced mean birth weight of pups, and reduced maternal weight gain (119). Developmental effects observed in rats include embryolethality with doses as low as 7.5 mg/kg/d and soft tissue abnormalities, including fetal cardiovascular anomalies and cardiovascular and urogenital malformations with a 15-mg/kg/d dose (127). Female mice treated with trichloroacetonitrile showed a significant increase in the incidence of lung adenomas; however, the authors cautioned that a large variation in background incidence of lung tumors was also observed in the same species (123). In addition, trichloroacetonitrile was shown to be weakly positive for mutagenicity in the *Salmonella typhimurium* assay (128).

HALOACETALDEHYDES AND HALOKETONES

Trichloroacetaldehyde is the most common haloacetaldehyde found in drinking water. Other acetaldehydes and haloketones found at low concentrations include dichloroacetaldehyde, bromochloroacetaldehyde and tribromoacetaldehyde, chloropropanone, and di-, tri-, and tetrahalopropanones. Limited toxicity data are available for the halogenated aldehydes and ketones other than trichloroacetaldehyde. Chloroacetaldehyde exposure causes hematological effects (129) and liver tumors (130) in rats. Exposure of mice to 1,1-dichloropropanone resulted in liver toxicity (131). Some halogenated aldehydes and ketones are potent inducers of mutations in bacteria (132–139). Other halogenated aldehydes, e.g., 2-chloropropenal, have been identified as tumor initiators in the skin of mice (131). The haloketones have not been tested for carcinogenicity in drinking water. However, 1,3-DCPN acted as a tumor initiator in a skin carcinogenicity study in mice (131). The health effects of trichloroacetaldehyde are summarized below.

Trichloroacetaldehyde Monohydrate

The primary effect seen with ingestion of trichloroacetaldehyde monohydrate or chloral hydrate in humans is central nervous system depression, the basis of its use in therapeutics (140–142). At high doses, chloral hydrate produces cardiac arrhythmias (141–144) and liver damage (140,145), whereas at lower doses, high incidences of direct hyperbilirubinemia were observed in newborns (146). In animals, hepatocellular necrosis has been observed in rats dosed with 1200-mg/L chloral hydrate (147), whereas mice displayed hepatomegaly at a dose of 144-mg/kg/d chloral hydrate (148). In a reproductive study, Klinefelter et al. (149) observed a significant decrease in the percentage of motile sperm in rats exposed to 188-mg/kg/d chloral hydrate, whereas Saillenfait et al. (101) observed developmental effects *in vitro* using a rat whole-embryo culture system. A chronic bioassay in female mice showed a slight increase in the severity grade of hyperplasia and a slight increase in the incidence of adenoma in the pituitary gland pars distalis at the highest exposure tested, whereas some evidence existed that chloral hydrate causes hepatocellular tumors in male mice (150). As a result of a lack of human carcinogenic studies and inadequate animal studies, the EPA has classified chloral hydrate as a Group C or possible human carcinogen (151). In addition, chloral hydrate was shown to be positive for mutagenicity in the *Salmonella typhimurium* assay (152).

MISCELLANEOUS CHEMICALS

Many other DBPs that are found in low concentrations can be found in drinking water. DBPs whose health effects have been summarized below include MX, chloropicrin, and the inorganic chemicals bromate, chlorate, and chlorite.

MX (3-Chloro-4-(Dichloromethyl)-5-Hydroxy-2(5H)-Furanone)

No human health effect studies because of oral MX exposure are available in the literature. In acute animal studies, mice that died within 24 hours of receiving a dose of 128 mg/kg MX were found to have enlarged stomachs with moderate hemorrhagic areas in the forestomach (153). In a second study, very limited mortality was observed in mice administered a dose of 144 mg/kg (154). Other effects observed in the study included focal epithelial hyperplasia in the stomach, vacuolation of the superficial villus epithelium in the duodenum and jejunum, increased numbers of mitotic figures in the liver, and some cytotoxicity in the urinary bladder (154). Meier et al. (155) observed reduced hepatic levels of catalase, cytochrome P450 reductase, aminopyrine demethylase, and aromatic hydrocarbon hydroxylase, thereby leading to potential modifications of metabolism of various xenobiotics and endogenous biochemicals in rats exposed to 64-mg/kg/d MX for 14 days. Vaittinen et al. (156) observed some lethality, hypersalivation, wheezing respiration, emaciation, tangled fur, increased water consumption, decreased body weights and food consumption, elevated plasma cholesterol, increased relative weights in the liver and kidneys, and increased urine excretion in rats dosed with 75-mg/kg/d MX for 14 weeks. MX administered to rats in drinking water for 104 weeks showed dose-dependent increases in tumors of the lung, mammary gland, hematopoietic system, liver, pancreas, adrenal gland, and thyroid (157). In addition, chloral hydrate was shown to be positive for mutagenicity in the *Salmonella typhimurium* assay (158).

Chloropicrin

Humans exposed orally to chloropicrin produced severe nausea, vomiting, colic, diarrhea, respiratory problems, and sometimes death (159,160). Health effects in rats orally exposed to chloropicrin include significantly reduced mean body and thymus weights and death because of pulmonary complications at high doses and dose-related histological changes in the stomach including forestomach inflammation, necrosis, acantholysis, hyperkeratosis, and ulceration (161). Oral developmental and reproductive animal studies were not available in the literature, but an inhalational study on rats and rabbits reported developmental toxicity in both species manifested as reduced fetal body weight, and maternal toxicity manifested in rats by mortality, reduction in body weight gain, and, in rabbits, by mortality, abortion, and inhibition of body weight gain and food consumption (162). In cancer studies involving rats and mice, proliferative lesions of the squamous epithelium of the forestomach, including carcinoma, papilloma, acanthosis, and hyperkeratosis, were observed in the dosed animals (163). However, owing to the short survival time of the dosed animals, conclusive evidence of carcinogenicity could not be established (163). However, chloropicrin was shown to be positive for mutagenicity in the *Salmonella typhimurium* assay (164).

Chlorite

No treatment-related effects were observed in humans exposed to rising doses of chlorite in drinking water (165). Short-term toxicity studies in rats, mice, cats, and monkeys indicated oxidative stress and signs of hemolytic anemia with decreases in red blood cell count, hemoglobin concentration, and hematocrit (166–171). The study results also indicated that although the changes in the blood parameters decreased in severity with duration of exposure, signs of oxidative stress remained elevated. Reproductive and developmental effects because of 100-mg/kg/d chlorite exposure include reduced conception rate and number of pups alive at weaning in mice (172) and vaginal and urethral bleeding, decreased body weight and food consumption, and mortality in rats (173,174). However, no effects were observed in animals dosed with chlorite at drinking water concentration levels. Chronic oral studies in rats showed no evidence of carcinogenicity (175), whereas mouse studies showed an increase in liver and lung tumors in male mice (176,177). However, the relatively short study duration and high mortality in the controls made the study interpretation difficult, leading the EPA to classify chlorite as a Group D chemical: not classifiable as to human carcinogenicity because of inadequate data (151). However, chlorite was shown to be positive for mutagenicity in the *Salmonella typhimurium* assay (178).

Chlorate

There have been a few reports of chlorate poisoning with evidence of oxidative damage to erythrocytes, methemoglobin formation, and hemolytic anemia in humans (179–182). In animals, high acute doses of 1000 mg/kg/d or higher induced methemoglobinemia and discoloration of the kidneys in dogs (183,184), but no consistent effects were noted in rats exposed to 10-mg/kg/d chlorate for 12 months (168). In an NTP study on reproductive and developmental effects because of chlorate exposure, transient changes in maternal food intake, urinary color, and/or output were noted at >100 mg/kg/d, but clear evidence of maternal toxicity was observed only at doses greater than 475 mg/kg/d in the screening study (185). In addition, chlorate did not cause any significant treatment-related developmental toxicity under the conditions of this study. In the lone cancer study available in the literature, a statistically insignificant increase in incidence of renal cell tumors in the N-ethyl-N-hydroxyethylnitrosamine-initiated rats treated with sodium chlorate was observed, but no effect was observed with potassium chlorate (186). In addition, chlorite was shown to be negative for mutagenicity in the *Salmonella typhimurium* assay (187).

Bromate

Clinical signs of bromate poisoning in humans include anemia, hemolysis, renal failure, and hearing loss (188,189). Acute toxic effects of bromate in rats, mice, and hamsters include suppressed locomotor activity, ataxia, tachypnoea, hypothermia, diarrhea, lacrimation and piloerection, hyperemia of the stomach, congestion of lungs, and damage to renal tubules including necrosis in the proximal tubular epithelium (189). Elevated levels of glutamate-oxalate transaminase, glutamate-pyruvate transaminase, lactate dehydrogenase, alkaline phosphatase, and blood urea nitrogen were observed in blood samples of rats exposed to bromate for 13 weeks (189,190). An NTP study on reproductive and developmental toxicity in rats indicated no reproductive effects in females but suggested that bromate may be a selective male reproductive toxicant at 250-ppm dose or higher (191). Cancer studies on rats indicated tumors at multiple sites including the kidney, thyroid, and the peritoneum (175,176,192), leading the EPA to classify bromate as a probable human carcinogen (193). However, bromate was shown to be weakly positive for mutagenicity in the *Salmonella typhimurium* assay (178).

MODELED TOXICITY ESTIMATES

Quantitative structure activity relationships (QSARs) may be used to estimate toxicities of many DBPs that lack experimental data in the literature. Commercial QSAR software packages such as DEREK (LHASA Ltd., Leeds, UK), CASE and MULTICASE (MultiCase Inc., Cleveland, OH), COMPACT (School of Biological Sciences, University of Surrey, UK), HazardExpert (Compudrug Ltd., Budapest, Hungary), ONCOLOGIC (Logichem Inc., Boyertown, PA), and TOPKAT (Accelrys Inc., Birmingham, MA) are able to predict the toxicities of a wide variety of chemicals for a number of health endpoints. In this section, TOPKAT has been used to estimate various toxicological endpoints for DBPs of health concern. Table 1 lists the rodent carcinogenicity based on the weight of evidence, Salmonella mutagenicity, rat developmental toxicity, lethal dose that causes 50% mortality in a rat study population, and the lowest dose that causes an adverse effect in rats. The procedure for estimating the endpoints using TOPKAT and its potential limitations are listed in Venkatapathy et al. (195) and Moudgal et al. (196,197).

BIBLIOGRAPHY

1. Weinberg, H.S., Krasner, S.W., Richardson, S.D., and Thruston, Jr., A.D. (2002). The Occurrence of Disinfection By-Products (DBPs) of Health Concern in Drinking Water: Results of a Nationwide DBP Occurrence Study. National Exposure Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Athens, GA. EPA Report EPA/600/R-02/068.
2. Cantor, K.P., Hoover, R., Hartge, P. et al. (1987). Bladder cancer, drinking water source, and tap water consumption: A case-control study. *J. Natl. Cancer Inst.* **79**: 1269–1279.
3. Crump, K.S. (1983). Chlorinated drinking water and cancer: The strength of the epidemiologic evidence. In: *Water Chlorination: Environmental Impact and Health Effects*. Jolley, R.L., Brungs, W.A., Cotruvo, J.A. et al. (Eds.). *Vol. 4, Book 2. Environment, Health, and Risk*. Ann Arbor Science Publishers, Ann Arbor, MI, pp. 1481–1491.
4. Kanarek, M.S. and Young, T.B. (1982). Drinking water treatment and risk of cancer death in Wisconsin. *Environ. Health Perspect.* **46**: 179–186.
5. Marienfeld, C.J., Collins, M., Wright, H. et al. (1986). Cancer mortality and the method of chlorination of public drinking water: St. Louis City and St. Louis County, Missouri. *JEPTO* **7**: 141–158.
6. Mortelmans, K. and Zeiger, E. (2000). The Ames Salmonella/microsome mutagenicity assay. *Mutat. Res.* **455**(1–2): 29–60.
7. du Pont de Nemours. (1977). Mutagenic activity of methane, chloro- in the Salmonella/microsome assay. E.I. du Pont de Nemours and Company, Haskell Laboratory Report, Wilmington, DE.
8. ATSDR. (1992). *Toxicological Profile for Bromomethane*. U.S. Department of Health and Human Services, Public Health Service, Atlanta, GA.
9. Danse, L.H., van Velson, F.L., and van der Heljden, C.A. (1984). Methylbromide: Carcinogenic effects in the rat forestomach. *Toxicol. Appl. Pharmacol.* **72**: 262–271.
10. Voogd, C.E., Knaap, A.G.A.C., Van der Heijden, C.A., and Kramers, P.G. (1982). Genotoxicity of methylbromide in short-term assay systems. *Mutat. Res.* **97**: 233.
11. Simmon, V.F. and Tardiff, R.G. (1978). Mutagenic activity of halogenated compounds found in chlorinated drinking water. *Water Chlorination: Environ. Health Eff. Proc. Conf.* **2**: 417–431.
12. Thier, R. et al. (1996). Human glutathione S-transferase T1-1 enhances mutagenicity of 1,2-dibromoethane, dibromomethane and 1,2,3,4-diepoxybutane in Salmonella typhimurium. *Carcinogenesis* **17**(1): 163–166.

Table 1. Toxicity Estimate of DBPs of Health Concern Using TOPKAT

DBP	Carcinogenicity[a]	Mutagenicity[b]	DTP[c]	LD$_{50}$[d]	LOAEL[e]
Halomethanes					
Chloromethane	Positive	Positive	Negative	1500	65.4
Bromomethane	Positive	Positive	Negative	622.9	71.2
Bromochloromethane	Negative	Positive	Ind*	1300	26
Dibromomethane	Positive	Positive	Positive	465.2	89.1
Chloroform	Positive	Negative	Negative	536.5	57.1
Bromodichloromethane	Positive	Negative	Negative	954.3	14
Dibromochloromethane	Positive	Negative	Negative	914.5	12.1
Bromoform	Positive	Positive	Negative	308.9	99.4
Dichloroiodomethane	Positive	Negative	Negative	1200	23.8
Bromochloroiodomethane	Negative	Negative	Negative	1600	4.7
Dibromoiodomethane	Positive	Positive	Negative	521.8	38.9
Chlorodiiodomethane	Positive	Negative	Negative	1300	34.1
Bromodiiodomethane	Positive	Positive	Negative	603.6	66.3
Iodoform	Negative	Positive	Negative	481.2	496.7
Carbon tetrachloride	Positive	Negative	Ind	2700	820.7
Tribromochloromethane	Positive	Negative	Negative	4500	158.5
Haloacetic acids					
Monochloroacetic acid	Negative	Negative	Positive	195.6	14.3
Monobromoacetic acid	Negative	Positive	Positive	1400	15.3
Dichloroacetic acid	Negative	Negative	Negative	1100	5.9
Bromochloroacetic acid	Negative	Positive	Negative	652.8	2.7
Dibromoacetic acid	Negative	Positive	Negative	867.3	19.9
Trichloroacetic acid	Ind	Negative	Positive	263.5	60.3
Bromodichloroacetic acid	Negative	Negative	Positive	164	28.8
Dibromochloroacetic acid	Negative	Ind	Positive	134.4	29
Tribromoacetic acid	Ind	Negative	Positive	162.9	62.3
Haloacetonitriles					
Chloroacetonitrile	Negative	Positive	Positive	194.5	0.645
Bromoacetonitrile	Negative	Positive	Positive	1100	0.6989
Dichloroacetonitrile	Positive	Positive	Positive	526.9	0.287
Bromochloroacetonitrile	Negative	Positive	Negative	305.1	0.0932
Dibromoacetonitrile	Positive	Positive	Positive	425.8	0.9403
Trichloroacetonitrile	Positive	Positive	Positive	757.7	4.5
Haloacetaldehydes					
Dichloroacetaldehyde	Negative	Positive	Negative	685.6	5.8
Bromochloroacetaldehyde	Negative	Positive	Negative	399	2.2
Trichloroacetaldehyde	Positive	Positive	Positive	966.1	72.2
Tribromoacetaldehyde	Positive	Ind	Positive	589.5	63.4
Haloketones					
Chloropropanone	Positive	Positive	Positive	283.6	47.5
1,1-Dichloropropanone	Positive	Positive	Negative	886.2	17.7
1,3-Dichloropropanone	Negative	Positive	Positive	70.5	77.3
1,1-Dibromopropanone	Positive	Positive	Negative	669	57.2
1,1,1-Trichloropropanone	Positive	Positive	Positive	511.5	169.2
1,1,3-Trichloropropanone	Negative	Positive	Negative	323.6	11.5
1-Bromo-1,1-dichloropropanone	Negative	Positive	Positive	312.1	89.7
1,1,1-Tribromopropanone	Positive	Positive	Positive	286.1	138.4
1,1,3-Tribromopropanone	Negative	Positive	Negative	2200	25.6
1,1,3,3-Tetrachloropropanone	Negative	Negative	Negative	1600	36.3
1,1,3,3-Tetrabromopropanone	Negative	Positive	Negative	596.5	62.6
Halonitromethanes					
Bromonitromethane	Negative	Positive	Positive	1500	0.8421
Dichloronitromethane	Negative	Negative	Negative	323.1	0.647
Dibromonitromethane	Negative	Positive	Negative	258.4	2.2
Chloropicrin	Ind	Ind	Positive	345.6	14.3

[a] Carcinogenicity based on Weight of Evidence characterization (194).

[b] Ames Mutagenicity.

[c] Developmental Toxicity Potential.

[d] Rat oral lethal dose (LD50) that causes mortality in 50% of an exposed population.

[e] Rat oral lowest observed adverse effect level—lowest dose that causes an adverse effect in an exposed population.

*Model was unable to determine positive or negative response unequivocally.

13. Schroeder, H.G. (1965). Acute and delayed chloroform poisoning. *Br. J. Anaesth.* **37**: 972–975.
14. Torkelson, T.R., Oyen, F., and Rowe, V.K. (1976). The toxicity of chloroform as determined by single and repeated exposure of laboratory animals. *Am. Ind. Hyg. Assoc. J.* **37**: 697–705.
15. Smyth, H.F., Jr., Carpenter, C.P., Weil, C.S. et al. (1962). Range-finding toxicity data: List VI. *Am. Ind. Hyg. Assoc. J.* **23**: 95–107.
16. Chu, I., Villeneuve, D.C., Secours, V.E. et al. (1982). Trihalomethanes: II. Reversibility of toxicological changes produced by chloroform, bromodichloromethane, chlorodibromomethane and bromoform in rats. *J. Environ. Sci. Health B.* **17**: 225–240.
17. NCI. (1976). Report on carcinogenesis bioassay of chloroform. Carcinogenesis Program, National Cancer Institute, Bethesda, MD.
18. Wallace, C.J. (1950). Hepatitis and nephrosis due to cough syrup containing chloroform. *Calif. Med.* **731**: 442–443.
19. Piersol, G.M., Tumen, H.J., and Kau, L.S. (1933). Fatal poisoning following the ingestion of chloroform. *Med. Clin. N. Am.* **17**: 587–601.
20. Thompson, D.J., Warner, S.D., and Robinson, V.B. (1974). Teratology studies on orally administered chloroform in the rat and rabbit. *Toxicol. Appl. Pharmacol.* **29**: 348–357.
21. Larson, J.L., Wolf, D.C., and Butterworth, B.E. (1993). Acute hepatotoxic and nephrotoxic effects of chloroform in male F-344 rats and female B6C3Fl mice. *Fundam. Appl. Toxicol.* **20**(3): 302–315.
22. De Salva, S., Volpe, A., Leigh, G. et al. (1975). Long-term safety studies of a chloroform-containing dentifrice and mouth-rinse in man. *Food Cosmet. Toxicol.* **13**: 529–532.
23. Chu, I., Villeneuve, D.C., Secours, V.E. et al. (1982). Toxicity of trihalomethanes: I. The acute and subacute toxicity of chloroform, bromodichloromethane, chlorodibromomethane and bromoform in rats. *J. Environ. Sci. Health B.* **17**: 205–224.
24. Munson, A.E., Sain, L.E., Sanders, V.M. et al. (1982). Toxicology of organic drinking water contaminants: Trichloromethane, bromodichloromethane, dibromochloromethane and tribromomethane. *Environ. Health Perspect.* **46**: 117–126.
25. Balster, R.L. and Borzelleca, J.F. (1982). Behavioral toxicity of trihalomethane contaminants of drinking water in mice. *Environ. Health Perspect.* **46**: 127–136.
26. Bowman, F.J., Borzelleca, J.F., and Munson, A.E. (1978). The toxicity of some halomethanes in mice. *Toxicol. Appl. Pharmacol.* **44**: 213–215.
27. Bove, F.J., Fulcomer, M.C., Klotz, J.B. et al. (1995). Public drinking water contamination and birth outcomes. *Amer. J. Epidemiol.* **141**(9): 850–862.
28. Alavanja, M., Goldstein, I., and Susser, M. (1978). A case control study of gastrointestinal and urinary tract cancer mortality and drinking water chlorination. In: *Water Chlorination, Environmental Impact and Health Effects*. Jolley, R.J., Gorchen, H., Hamilton, D.H., Jr. (Eds.). Ann Arbor Science Publications, Ann Arbor, MI, 395–409.
29. Cantor, K.P., Hoover, R., Mason, T.J. et al. (1978). Associations of cancer mortality with halomethanes in drinking water. *J. Natl. Cancer Inst.* **61**: 979–985.
30. Suarez-Varela, M.M. and Gonzalez, A.L. (1994). Chlorination of drinking water and cancer incidence. *J. Environ. Path. Toxicol. Onocol.* **13**(1): 39–41.
31. Young, T.B., Kanarek, M.S., and Tsiatis, A.A. (1981). Epidemiologic study of drinking water chlorination and Wisconsin female cancer mortality. *J. Natl. Cancer Inst.* **67**: 1191–1198.
32. U.S. EPA. (1994). Final Draft for the Drinking Water Criteria Document on Trihalomethanes. Prepared for Health and Ecological Criteria Division, Office of Science and Technology, Washington, DC. Under EPA Contract No. 68-C2-0139 by Clement International Corporation.
33. U.S. EPA. (1998). National primary drinking water regulations: Disinfectants and disinfection byproducts. Final Rule. *Fed. Reg.* **63**(241): 69406–69407.
34. NTP. National Toxicology Program. (1987). Toxicology and Carcinogenesis Studies of Bromodichloromethane (CAS No. 75-27-4) in F344/N Rats and B6C3F, Mice (Gavage Studies). NIH Publication No. 88–2537, TR-321, Research Triangle Park, NC.
35. Condie, L.W., Smallwood, C.L., and Laurie, R.D. (1983). Comparative renal and hepatotoxicity of halomethanes: ‘bromodichloromethane, bromoform, chloroform, dibromochloromethane and methylene chloride. *Drug Chem. Toxicol.* **6**: 563–578.
36. Ruddick, J.A., Villenueve, D.C., and Chu, I. (1983). A teratological assessment of four trihalomethanes in the rat. *J. Environ. Sci. Health B.* **18**: 333–349.
37. Aida, Y., Takada, K., Momma, J. et al. (1987). Acute and subacute toxicities of three trihalomethanes in rats (I). *J. Toxicol. Sci.* **12**: 585.
38. Chu, I., Secours, V., Marino, I. et al. (1980). The acute toxicity of trihalomethanes in male and female rats. *Toxicol. Appl. Pharmacol.* **52**: 351–353.
39. Gottlieb, M.S., Carr, J.K., and Morris, D.T. (1981). Cancer and drinking water in Louisiana: Colon and rectum. *Int. J. Epidemiol.* **10**: 117–125.
40. Tumasonis, C.F., McMartin, D.N., and Bush, B. (1985). Lifetime toxicity of chloroform and bromodichloromethane when administered over a lifetime in rats. *Ecotoxicol. Environ. Safety.* **9**: 233–240.
41. Dunnick, J.K., Eustis, S.L., and Lilja, H.S. (1987). Bromodichloromethane, a trihalomethane that produces neoplasms in rodents. *Cancer Res.* **47**: 5189–5193.
42. Hewitt, W.R., Brown, E.M., and Plaa, G.L. (1983). Acetone-induced potentiation of trihalomethane toxicity in male rats. *Toxicol. Lett.* **16**: 285–296.
43. NTP. (1985). Toxicology and Carcinogenesis Studies of Chlorodibromomethane (CAS No. 124-48-1) in F344/N Rats and B6C3F, Mice (Gavage Studies). Technical report series no. 282. NIH Publication No. 88–2538, Research Triangle Park, NC.
44. Aida, Y., Takada, K., Uchida, O. et al. (1992). Toxicities of microencapsulated tribromomethane, dibromochloromethane and bromodichloromethane administered in the diet to Wister rats for one month. *J. Toxicol. Sci.* **17**(3): 119–133.
45. Bove, F.J., Fulcomer, M., Klotz, J.B. et al. (1992). *Population-Based Surveillance and Etiological Research of Adverse Reproductive Outcomes and Toxic Wastes*. Report on phase 1V-B: Public drinking water contamination and birthweight, and selected birth defects. A case control study. New Jersey Department of Health, Trenton, NJ.
46. Bove, F., Shim, Y., and Zeitz, P. (2002). Drinking water contaminants and adverse pregnancy outcomes: A review. *Environ. Health Perspect.* **110**(Suppl 1): 61–74.

47. Dodds, L., King, W., Woolcott, C. et al. (1999). Trihalomethanes in public water supplies and adverse birth outcomes. *Epidemiology*. **10**(3): 233–237.
48. Kramer, M.D., Lynch, C.F., Isacson, P. et al. (1992). The association of waterborne chloroform with intrauterine growth retardation. *Epidemiology*. **3**: 407–413.
49. Mills, C.J. et al. (1998). Workshop report on health risks of drinking water chlorination by-products: Report of an expert working group, *Chronic Diseases in Canada.* **19**(3).
50. Nieuwenhuijsen, M.J., Toledano, M.B., Eaton, N. et al. (2000). Chlorination disinfection byproducts in water and their association with adverse reproductive outcomes: A review. *Occup. Environ. Med.* **57**(2): 73–85.
51. Savitz, D.A., Andrews, K.W., and Pastore, L.M. (1995). Drinking water and pregnancy outcome in central North Carolina: Source, amount, and trihalomethane levels. *Environ. Health Perspect.* **103**(6): 592–596.
52. Waller, K., Swan, S.H., Delorenze, G. et al. (1999). Trihalomethanes in drinking water and spontaneous abortion. *Epidemiology* **10**(2): 203–204.
53. Mills, W.B., Lew, C.S., and Loh, J.Y. (1998). Predictions of potential human health and ecological risks from power plant discharges of total residual chlorine and chloroform into rivers. *Environ. Sci. Technol.* **32**: 2162–2171.
54. Borzelleca, J.F. and Carchman, R.A. (1982). Effects of Selected Organic Drinking Water Contaminants on Male Reproduction. U.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park, NC. EPA 600/l-82-009. NTIS No. PB82-259847.
55. IRIS. (2004). Dibromochloromethane. U. S. Environmental Protection Agency, Washington, DC. Available at http://www.epa.gov/iris/subst/0222.htm.
56. Dwelle, E.H. (1903). Fatal bromoform poisoning. *JAMA* **41**: 1540.
57. Kobert, R. (1906). *Lehrbuch der Intaxikationen. II band. Spezieller teil*. Verlag Von Ferdinand Enke, Stuttgart, Germany, pp. 914–916.
58. von Oettingen, W.F. (1955). The Halogenated Aliphatic, Olefinic, Cyclic, Aromatic, and Aliphatic-aromatic Hydrocarbons Including the Halogenated Insecticides, their Toxicity and Potential Dangers. U.S. Department of Health, Education and Welfare, Public Health Service, Washington, DC, pp. 65–67.
59. NTP. (1989). Toxicology and Carcinogenesis Studies of Tribromomethane (Bromoform) (CAS No. 75-25-2) in F344/N Rats and B6C3F1 Mice (Gavage Studies). Technical report series no. 350. NIH Publication No. 88–2805, Research Triangle Park, NC.
60. NTP. (1989). Bromoform: Reproduction and Fertility Assessment in Swiss CD-1 Mice When Administered by Gavage. Report to National Toxicology Program. NTP-89-068. NTIS No. PB89-169254, Research Triangle Park, NC.
61. IRIS. (2004). Bromoform. U. S. Environmental Protection Agency, Washington, DC. Available at http://www.epa.gov/iris/subst/0214.htm.
62. NCI. (1978). Bioassay of Iodoform for Possible Carcinogenicity. Carcinogenesis Program, National Institute of Health, Bethesda, MD, NCI-CG-TR-110.
63. HSDB. Hazardous Substances Data Bank. (2004). U.S. Department of Health and Human Services. National Toxicology Information Program, National Library of Medicine, Bethesda, MD. Available at http://toxnet.nlm.nih.gov/.
64. NTP. (1992). Toxicology and Carcinogenesis Studies of Monochloroacetic Acid. (CAS no. 79-11-8) in F344/N Rats and B6C3F1 Mice (Gavage Studies). Department of Health and Human Services, Public Health Service, National Institutes of Health. NTP TR 396, Beltsville, MD.
65. Linder, R.E. et al. (1994). Acute spermatogenic effects of bromoacetic acids. *Fundam. Appl. Toxicol.* **22**: 422–430.
66. Randall, J.L. et al. (1991). Developmental effects of 2-bromoacetic acid in the Long Evans rat. *Teratology* **43**(5): 454.
67. NTP. National Toxicology Program. (2000). Water disinfection byproducts (bromoacetic acid). Available at http://ntpserver.niehs.nih.gov/htdocs/Results_Status/Resstatb/M920034.html.
68. Crabb, D.W., Yount, E.A., and Harris, R.A. (1981). The metabolic effects of dichloroacetate. *Metabolism* **30**: 1024–1039.
69. Coude, F.X. et al. (1978). Dichloroacetate as treatment for congenital lactic acidosis. *N. Engl. J. Med.* **299**: 1365–1366.
70. Stacpoole, P.W. (1989). The pharmacology of dichloroacetate. *Metabolism* **38**: 1124–1144.
71. Stacpoole, P.W. and Greene, Y.J. (1992). Dichloroacetate. *Diabetes Care* **15**: 785–791.
72. Toth, P.P. et al. (1993). Transient improvement of congenital lactic acidosis in a male infant with pyruvate decarboxylase deficiency treated with dichloroacetate. *J. Pediatr.* **123**: 427–430.
73. Moore, G.W. et al. (1979). Reduction of serum cholesterol in two patients with homozygous familial hypercholesterolemia by dichloroacetate. *Atherosclerosis* **33**: 285–293.
74. Cicmanec, J.L., Condie, L.W., Olson, G.R., and Wang, S.R. (1991). 90-day toxicity study of dichloroacetate in dogs. *Fundam. Appl. Toxicol.* **17**: 376–389.
75. Katz, R. et al. (1981). Dichloroacetate, sodium: 3-month oral toxicity studies in rats and dogs. *Toxicol. Appl. Pharmacol.* **57**: 273–287.
76. Toth, G.P. et al. (1992). Adverse male reproductive effects following subchronic exposure of rats to sodium dichloroacetate. *Fundam. Appl. Toxicol.* **19**: 57–63.
77. Epstein, D.L. et al. (1992). Cardiopathic effects of dichloroacetate in the fetal Long-Evans rat. *Teratology*. **46**: 225–235.
78. Smith, M.K., Randall, J.L., Read, E.J., and Stober, J.A. (1992). Developmental toxicity of dichloroacetate in the rat. *Teratology.* **46**: 217–223.
79. Bull, R.J. et al. (1990). Liver tumor induction in B6C3F1 mice by dichloroacetate and trichloroacetate. *Toxicology*. **63**: 341–359.
80. Daniel, F.B. et al. (1992). Hepatocarcinogenicity of chloral hydrate, 2-chloroacetaldehyde, and dichloroacetic acid in male B6C3F1 mouse. *Fundam. Appl. Toxicol.* **19**: 159–168.
81. DeAngelo, A.B., Daniel, F.B., Stober, J.A., and Olson, G.R. (1991). The carcinogenicity of dichloroacetic acid in the male B6C3F1 mouse. *Fundam. Appl. Toxicol.* **16**: 337–347.
82. Herren-Freund, S.L., Pereira, M.A., Khoury, M.D., and Olson, G. (1987). The carcinogenicity of trichloroethylene and its metabolites, trichloroacetic acid and dichloroacetic acid, in mouse liver. *Toxicol. Appl. Pharmacol.* **90**: 183–189.
83. Pereira, M.A. (1996). Carcinogenic activity of dichloroacetic acid and trichloroacetic acid in the liver of female B6C3F1 mice. *Fundam. Appl. Toxicol.* **31**: 192–199.
84. IRIS. (2004). Dichloroacetic acid. U. S. Environmental Protection Agency, Washington, DC. Available at http://www.epa.gov/iris/subst/0654.htm.
85. NTP. National Toxicology Program. (1998). Bromochloroacetic Acid: Short Term Reproductive and Developmental

Toxicity Study When Administered to Sprague-Dawley Rats in the Drinking Water. Final report. R.O.W. Sciences, Inc. R.O.W. Sciences Study No. 7244-301; NTP/NIEHS No.: NO1-ES-75409; NTP-RDGT No.:1998, 96-001, Gaithersburg, MD.

86. Bull, R.J. and DeAngelo, A.B. (1995). *Carcinogenic Properties of Brominated Haloacetates—Disinfection by-products in Drinking Water: Critical Issues in Health Effects Research*. International Life Sciences Institute, Washington, DC, p. 29.
87. Narotsky, M.G., Best, D.S., Rogers, E.H., and Hunter, E.S. (2001). Developmental effects of disinfectant by-products bromoform, bromodichloromethane, bromodichloroacetic acid, and bromochloroacetic acid in mice. *Teratology* **63**(6): 279.
88. Wolfe, G.W., Kaiser, L.B., Mowry, R.N., Klinefelter, G., Hunter, E.S., and Chapin, R.E. (1998). Final Report on the Short Term Reproductive and Developmental Toxicity of Bromochloroacetic Acid (CAS #5589-96-8) Administered in Drinking Water to Sprague-Dawley Rats. NTIS Technical Report (NTIS/PB98-172414) (NTP/RDGT-96001), p. 222.
89. NTP. National Toxicology Program. (2000). Water disinfection byproducts (bromochloroacetic acid). Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/Resstatw/M980085.html.
90. Linder, R.E. et al. (1994). Spermatotoxicity of dibromoacetic acid in rats after 14 daily exposures. *Reprod. Toxicol.* **8**(3): 251–259.
91. Parrish, J.M. et al. (1996). Haloacetate-induced oxidative damage to DNA in the liver of male B6C3F1 mice. *Toxicology*. **110**: 103–111.
92. NTP. National Toxicology Program. (White, K.L. and Munson, J.A.). (1999). Immunotoxicity of Dibromoacetic Acid in Female B6C3F1 Mice. Final report, Medical College of Virginia, Richmond, VA.
93. NTP. National Toxicology Program. (2000). Water disinfection byproducts (dibromoacetic acid). Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/Resstatw/M960093.html.
93. Linder, R.E. et al. (1995). Dibromoacetic acid affects reproductive competence and sperm quality in the male rat. *Fundam. Appl. Toxicol.* **28**: 9–17.
94. Dreisbach, R.H. (1977). *Handbook of Poisoning*, 9th Edn. Lange Medical Publications, Los Altos, CA.
95. ACGIH. (1991). *Documentation of the Threshold Limit Values and Biological Exposure Indices*, 6th Edn. American Conference of Governmental Industrial Hygienists, Cincinnati, OH.
96. Davis, M.E. (1990). Subacute toxicity of trichloroacetic acid in male and female rats. *Toxicology* **63**: 63–72.
97. Mather, G.G., Exon, J.H., and Koller, L.D. (1990). Subchronic 90 day toxicity of dichloroacetic and trichloroacetic acid in rats. *Toxicology* **64**: 71–80.
98. Goldsworthy, T.L. and Popp, J.A. (1987). Chlorinated hydrocarbon-induced peroxisomal enzyme activity in relation to species and organ carcinogenicity. *Toxicol. Appl. Pharmacol.* **88**: 225–233.
99. Sanchez, I.M. and Bull, R.J. (1990). Early induction of reparative hyperplasia in the liver of B6C3F1 mice treated with dichloroacetate and trichloroacetate. *Toxicology* **64**: 33–46.
100. Cosby, N.C. and Dukelow, W.R. (1992). Toxicology of maternally ingested trichloroethylene (TCE) on embryonal and fetal development in mice and of TCE metabolites on *in vitro* fertilization. *Fundam. Appl. Toxicol.* **19**: 268–274.
101. Saillenfait, A.M., Langonne, I., and Sabate, J.P. (1995). Developmental toxicity of trichloroethylene, tetrachloroethylene and four of their metabolites in rat whole embryo culture. *Arch. Toxicol.* **70**: 71–82.
102. Smith, M.K., Randall, J.L., Read, E.J., and Stober, J.A. (1989). Teratogenic activity of trichloroacetic acid in the rat. *Teratology* **40**: 445–451.
103. Daniel, F.B. et al. (1993). Toxicity of 1,1,1-trichloro-2-propanone in Sprague-Dawley rats. *J. Toxicol. Environ. Health.* **39**: 383–393.
104. DeAngelo, A.B., Daniel, F.B., Most, B.M., and Olson, G.R. (1997). Failure of monochloroacetic acid and trichloroacetic acid administered in the drinking water to produce liver cancer in male F344/N rats. *J. Toxicol. Environ. Health.* **52**: 425–445.
105. IARC. (1995). Dry Cleaning, Some Chlorinated Solvents and Other Industrial Chemicals. Lyon, International Agency for Research on Cancer (IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Volume 63).
106. Mills et al. (2000).
107. NTP. National Toxicology program. (1998). Water disinfection byproducts (bromodichloroacetic acid). Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/ResstatB/M970039.html.
108. NTP. (2000). Short Term Reproductive and Developmental Toxicity Study of Dibromochloroacetic Acid (CAS No. 5278-95-5) Administered to Sprague-Dawley Rats in the Drinking Water NTP Study Number: RDGT98001. NTIS # PB2000-103420.
109. NTP. (1998). Short Term Reproductive and Developmental Toxicity of Tribromoacetic Acid (CAS No. 75-96-7) Administered in Drinking Water to Sprague-Dawley Rats. NTP Study Number: RDGT94009. NTIS#: PB98-165111.
110. NTP. National Toxicology Program. (1997). Water disinfection byproducts (Tribromoacetic acid). Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/Resstatt/M940117.html.
111. Lin, E.L., Daniel, F.B., Herren-Freund, S.L., and Pereira, M.A. (1986). Haloacetonitriles: metabolism, genotoxicity, and tumor-initiating activity. *Environ. Health Perspect.* **69**: 67–71.
112. Mohamadin, A.M. and Abdel-Naim, A.B. (1999). Chloroacetonitrile-induced toxicity and oxidative stress in rat gastric epithelial cells. *Pharmacol. Res.* **40**(4): 377–383.
113. IARC. (1999). Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Man. World Health Organization, International Agency for Research on Cancer, Geneva, 1972-PRESENT. V71, p. 1327.
114. IARC. (1991). Chlorinated Drinking-water; Chlorination By-products; Some Other Halogenated Compounds; Cobalt and Cobalt Compounds. Lyon, International Agency for Research on Cancer (IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Volume 52).
115. NTP. National Toxicology program. (1981). Chloroacetonitrile. Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/ResstatC/10979-X.html.
116. NTP. (1997). Short Term Reproductive and Developmental Toxicity of Bromoacetonitrile (CAS No. 590-17-0) Administered in Drinking Water to Sprague-Dawley Rats. NTP Study Number: RDGT94013. NTIS # PB97-167084.
117. NTP. National Toxicology program. (1982). Bromoacetonitrile. Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/Resstatb/10972-V.html.

118. Hayes, J.R., Condie, L.W., and Borzelleca, J.F. (1986). Toxicology of haloacetonitriles. *Environ. Health Perspect.* **69**: 183–202.
119. Smith, M.K. et al. (1987). Developmental toxicity of halogenated acetonitriles: Drinking water by-products of chlorine disinfection. *Toxicology* **46**: 83–93.
120. Smith, M.K., Randall, J.L., Stober, J.A., and Read, E.J. (1989). Developmental toxicity of dichloroacetonitrile: A by-product of drinking water disinfection. *Fundam. Appl. Toxicol.* **12**: 765–772.
121. NTP. National Toxicology program. (1981). Dichloroacetonitrile. Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/Resstatd/10989-F.html.
122. Christ, S.A., Read, E.J., Stober, J.A., and Smith, M.K. (1995). The developmental toxicity of bromochloroacetonitrile in pregnant Long-Evans rats. *Int. J. Environ. Health Res.* **5**(2): 175–188.
123. Bull, R.J. and Robinson, M. (1985). Carcinogenic activity of haloacetonitrile and haloacetone derivatives in the mouse skin and lung. In: *Water Chlorination: Chemistry, Environmental Impact and Health Effects*, Jolley, R.L. et al. (Eds.), Vol. 5. Lewis Publishers, Chelsea, MI, pp. 221–227.
124. Bull, R.J. et al. (1985). Evaluation of mutagenic and carcinogenic properties of brominated and chlorinated acetonitriles byproducts of chlorination. *Fundam. Appl. Toxicol.* **5**(6 Part 1): 1065–1074.
125. NTP. (1997). Reproductive Toxicity of Dibromoacetonitrile (CAS No. 3252-43-5) Administered to Sprague-Dawley Rats in the Drinking Water. NTP Study Number: RDGT94014. NTIS # PB97-143127.
126. NTP. National Toxicology program. (1982). Dibromoacetonitrile. Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/Resstatd/10985-P.html.
127. Smith et al. (1988).
128. NTP. National Toxicology program. (1981). Trichloroacetonitrile. Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/Resstatt/11036-S.html.
129. Lawrence, W.H., Dillingham, E.O., Turner, J.E., and Autian, J. (1972). Toxicity profile of chloroacetaldehyde. *J. Pharm. Sci.* **61**: 19–25.
130. van Duuren, B.L. et al. (1979). Carcinogenicity of halogenated olefinic and aliphatic hydrocarbons in mice. *J. Natl. Cancer Inst.* **63**: 1433–1439.
131. Robinson, M., Bull, R.J., Olson, G.R., and Stober, J. (1989). Carcinogenic activity associated with halogenated acetones and acroleins in the mouse skin assay. *Cancer Lett.* **48**: 197–203.
132. Bignami, M. et al. (1980). Mutagenicity of halogenated aliphatic hydrocarbons in Salmonella typhimurium, Streptomyces coelicolor and Aspergillus nidulans. *Chem-Biol. Interact.* **30**: 9–23.
133. Douglas, G.R., Nestmann, E.R., McKague, A.B., San, R.H.C., Lee, E.G., Liu-Lee, V.W., and Kobel, D.J. (1985). Determination of potential hazards from pulp and paper mills: Mutagenicity and chemical analysis. In: *Carcinogens and Mutagens in the Environment—Volume 5: The Workplace*. Stich, H.F. (Ed.). CRC Press, Boca Raton, FL, pp. 151–166.
134. Gordon, W.P. et al. (1985). The genotoxicity of 2-bromoacrolein and 2,3-dibromopropanal. *Carcinogenesis* **6**: 705–709.
135. Meier, J.R. et al. (1985). Identification of mutagenic compounds formed during chlorination of humic acid. *Mutat. Res.* **157**: 111–122.
136. Meier, J.R., Bull, R.J., Stober, J.A., and Cimino, M.C. (1985). Evaluation of chemicals used for drinking water disinfection for production of chromosomal damage and sperm-head abnormalities in mice. *Environ. Mutagen.* **7**: 201–211.
137. Merrick, B.A. et al. (1987). Chemical reactivity, cytotoxicity, and mutagenicity of chloropropanones. *Toxicol. Appl. Pharmacol.* **91**: 46–54.
138. Rosen, J.D., Segall, Y., and Casida, J.E. (1980). Mutagenic potency of haloacroleins and related compounds. *Mutat. Res.* **78**: 113–119.
139. Segall, Y., Kimmel, E.C., Dohn, D.R., and Casida, J.E. (1985). 3-Substituted 2-halopropenals: mutagenicity, detoxification and formation from 3-substituted 2,3-dihalopropenal promutagens. *Mutat. Res.* **158**: 61–68.
140. Gilman, A.G., Rall, T.W., Neis, A.S., and Taylor, P. (1991). *Goodman and Gilman's Pharmacological Basis of Therapeutics*, 8th Edn. Pergamon Press, New York, pp. 364–365.
141. Farber, B. and Abramow, A. (1985). Acute laryngeal edema due to chloral hydrate. *Isr. J. Med. Sci.* **21**: 858–859.
142. Silver, W. and Stier, M. (1971). Cardiac arrhythmias from chloral hydrate. *Pediatrics* **48**(2): 332–333.
143. Hirsch, I.A. and Zauder, H.L. (1986). Chloral hydrate: A potential cause of arrhythmias. *Anesth. Analg.* **65**: 691–692.
144. Nordenberg, A., Delisle, G., and Izukawa, T. (1971). Cardiac arrhythmia in a child due to chloral hydrate ingestion. *Pediatrics* **47**: 134–135.
145. van Heijst, A.N., Zimmerman, A.N., and Pikaar, S.A. (1977). Chloral hydrate—the forgotten poison. *Ned. Tijdschr. Geneesk.* **121**(40): 1537–1539.
146. Lambert, G.H., Muraskas, J., Anderson, C.L., and Myers, T.F. (1990). Direct hyperbilirubinemia associated with chloral hydrate administration in the newborn. *Pediatrics* **86**: 277–281.
147. Daniel, F.B. et al. (1992). Ninety-day toxicity study of chloral hydrate in the Sprague-Dawley rat. *Drug. Chem. Toxicol.* **15**: 217–232.
148. Sanders, V.M. et al. (1982). Toxicology of chloral hydrate in the mouse. *Environ. Health Perspect.* **44**: 137–146.
149. Klinefelter, G.R., Suarez, J.D., and Roberts, N.L. (1995). Preliminary screening test for the potential of drinking water disinfectant by-products to alter male reproduction. *Reprod. Toxicol.* **9**: 571–578.
150. National Toxicology Program (NTP). (2000). Toxicology and Carcinogenesis Studies of Chloral Hydrate in B6C3F1 Mice (Gavage Studies). NTP TR 502, Silver Springs, MD.
151. IRIS. (2000). Chloral Hydrate. U. S. Environmental Protection Agency. Washington, DC. Available at http://www.epa.gov/iris/subst/0304.htm.
152. NTP. National Toxicology program. (1980). Chloral Hydrate. Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/Resstatc/10609-C.html.
153. Meier, J.R., Blazek, W.F., and Knohl, R.B. (1987). Mutagenic and clastogenic properties of 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5H)-furanone: A potent bacterial mutagen in drinking water. *Environ. Mol. Mutagen.* **10**: 411–424.
154. Mullins, P.A. and Proudlock, R.J. (1990). *Assessment of Nuclear Anomalies in Mice After Administration of MX*. Water Research Centre, Medmenham, UK.
155. Meier, J.R. et al. (1996). Urine mutagenicity and biochemical effects of the drinking water mutagen, 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5H)-furanone (MX) following repeated oral administration to mice and rats. *Toxicology* **110**: 59–70.

156. Vaittinen, S.L. et al. (1995). Subchronic toxicity of 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5H)-furanone (MX) in Wistar rats. *Food Chem. Toxicol.* **33**(12): 1027–1037.
157. Komulainen, H., Kosma, V.M., and Vaittinen, S. (1997). Carcinogenicity of the drinking water mutagen 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5H)-furanone (MX). *J. Natl. Cancer Inst.* **89**: 848–856.
158. Hyttinen, J.M., Myohanen, S., and Jansson, K. (1996). Kinds of mutations induced by 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5H)-furanone (MX) in the hprt gene of Chinese hamster ovary cells. *Carcinogenesis* **17**: 1179–1181.
159. Clayton, G.D. and Clayton, F.E. (Eds.). (1981). *Patty's Industrial Hygiene and Toxicology: Volume 2A, 2B, 2C: Toxicology*, 3rd Edn. John Wiley & Sons, New York.
160. *Merck Index*. (1983), 10th Edn. Merck Co., Rahway, NJ.
161. Condie, L.W., Daniel, F.B., Olson, G.R., and Robinson, M. (1994). Ten and ninety-day toxicity studies of chloropicrin in Sprague-Dawley rats. *Drug Chem. Toxicol.* **17**(2): 125–137.
162. York, R.G., Butala, J.H., Ulrich, C.E., and Schardein, J.L. (1994). Inhalation developmental toxicities of chloropicrin in rats and rabbits. *Teratology* **49**(5): 419.
163. NCI. (1978). Bioassay of Chloropicrin (CAS No. 76-06-2) for Possible Carcinogenicity. Carcinogenesis Program, National Institute of Health, Bethesda, MD, NCI-CG-TR-65.
164. NTP. National Toxicology program. (1980). Chloropicrin. Available at http://ntp-server.niehs.nih.gov/htdocs/Results_Status/ResstatC/10848-P.html.
165. Lubbers, J.R. and Bianchine, J.R. (1984). Effects of the acute rising dose administration of chlorine dioxide, chlorate and chlorite to normal healthy adult male volunteers. *J. Exp. Pathol. Toxicol. Oncol.* **5**: 215–228.
166. Abdel-Rahman, M.S., Couri, D., and Bull, R.J. (1980). Kinetics of ClO_2 and effects of ClO_2, ClO_2^-, and ClO_3^- in drinking water on blood glutathione and hemolysis in rat and chicken. *J. Environ. Pathol. Toxicol.* **3**: 431–449.
167. Bercz, J.P. et al. (1982). Subchronic toxicity of chlorine dioxide and related compounds in drinking water in the nonhuman primate. *Environ. Health Perspect.* **46**: 47–55.
168. Couri, D. and Abdel-Rahman, M.S. (1980). Effect of chlorine dioxide and metabolites on glutathione dependent system in rat, mouse and chicken blood. *J. Environ. Pathol. Toxicol.* **3**: 451–460.
169. Harrington, R.M., Romano, R.R., Gates, D., and Ridgway, P. (1995). Subchronic toxicity of sodium chlorite in the rat. *J. Am. Coll. Toxicol.* **14**: 21–33.
170. Heffernan, W.P., Guion, C., and Bull, R.J. (1979). Oxidative damage to the erythrocyte induced by sodium chlorite, *in vivo*. *J. Environ. Pathol. Toxicol.* **2**: 1487–1499.
171. Moore, G.S. and Calabrese, E.J. (1982). Toxicological effects of chlorite in the mouse. *Environ. Health Perspect.* **46**: 31–37.
172. Moore, G.S., Calabrese, E.J., and Leonard, D.A. (1980). Effects of chlorite exposure on conception rate and litters of A/J mice. *Bull. Environ. Contam. Toxicol.* **25**: 689–696.
173. Couri, D. et al. (1982). Assessment of maternal toxicity, embryotoxicity and teratogenic potential of sodium chlorite in Sprague-Dawley rats. *Environ. Health Perspect.* **46**: 25–29.
174. Couri, D., Abdel-Rahman, M.S., and Bull, R.J. (1982). Toxicological effects of chlorine dioxide, chlorite and chlorate. *Environ. Health Perspect.* **46**: 13–17.
175. Kurokawa, Y. et al. (1986). Long-term *in vivo* carcinogenicity tests of potassium bromate, sodium hypochlorite, and sodium chlorite conducted in Japan. *Environ. Health Perspect.* **69**: 221–235.
176. Kurokawa, Y. et al. (1986). Dose-response studies on the carcinogenicity of potassium bromate in F344 rats after long-term oral exposure. *J. Natl. Cancer Inst.* **77**: 977–982.
177. Yokose, Y. et al. (1987). Studies of carcinogenicity of sodium chlorite in B6C3F1 mice. *Environ. Health Perspect.* **76**: 205–210.
178. Ishidate, M. et al. (1984). Primary mutagenicity screening of food additives currently used in Japan. *Food Chem. Toxicol.* **22**: 623–636.
179. Bloxham, C.A., Wright, N., and Hoult, J.G. (1979). Self-poisoning by sodium chlorate: Some unusual features. *Clin. Toxicol.* **15**: 185–188.
180. Helliwell, M. and Nunn, J. (1979). Mortality in sodium chlorate poisoning. *Br. Med. J.* **6171**: 1119.
181. Steffen, C. and Seitz, R. (1981). Severe chlorate poisoning: Report of a case. *Arch. Toxicol.* **48**: 281–288.
182. Yoshida, Y. et al. (1977). A cytological study of Heinz body-hemolytic anemia. Report of a case of sodium chlorate poisoning complicated by methemoglobinemia and acute renal failure. *Acta. Haematol. Jpn.* **40**: 147–151.
183. Heywood, R., Sortwell, R.J., Kelly, P.J., and Street, A.E. (1972). Toxicity of sodium chlorate to the dog. *Vet. Rec.* **90**: 416–418.
184. Sheahan, B.J., Pugh, D.M., and Winstanley, E.W. (1971). Experimental sodium chlorate poisoning in dogs. *Res. Vet. Sci.* **12**: 387–389.
185. NTP. (1998). Developmental Toxicity Evaluation for Sodium Chlorate (CAS No. 7775-09-9) Administered by Gavage to New Zealand White Rabbits on Gestational Days 6 through 29. NTP Study: TER97005. NTIS# PB2003-101297.
186. Kurokawa, Y. et al. (1985). Dose-related enhancing effect of potassium bromate on renal tumorigenesis in rats initiated with N-ethyl- N-hydroxyethyl-nitrosamine. *Jpn. J. Cancer Res. (Gann).* **76**: 583–589.
187. Prieto, R. and Fernandez, E. (1993). Toxicity and mutagenesis by chlorate are independent of nitrate reductase in Chlamydomonas reinhardtii. *Mol. Gen. Genet.* **237**: 429–438.
188. Kuwahara, T. et al. (1984). Two cases of potassium bromate poisoning requiring long-term hemodialysis therapy for irreversible tubular damage. *Nephron.* **37**: 278–280.
189. Kurokawa, Y., Maekawa, A., Takahashi, M., and Hayashi, Y. (1990). Toxicity and carcinogenicity of potassium bromate - A new renal carcinogen. *Environ. Health Perspect.* **87**: 309–335.
190. Onodera, H. et al. (1986). Eosinophilic bodies in the proximal renal tubules of rats given potassium bromate. *Bull. Natl. Inst. Hyg. Sci.* **103**: 15–20.
191. NTP. (1996). Sodium Bromate (CAS No. 7789-38-0): Short Term Reproductive and Developmental Toxicity Study When Administered to Sprague-Dawley Rats in the Drinking Water. NTP Study Number: RDGT94007. NTIS# PB96-190640.
192. DeAngelo, A.B. et al. (1998). Carcinogenicity of potassium bromate administered in the drinking water to male B6C3F1 mice and F344/N rats. *Toxicol. Pathol.* **26**(5): 587–594.
193. IRIS. (2001). Bromate (CASRN 15541-45-4). U. S. Environmental Protection Agency, Washington, DC. Available at http://www.epa.gov/iris/subst/1002.htm.
194. U.S. EPA. (1996). Proposed guidelines for carcinogen risk assessment; notice. April 23, 1996. *Fed. Reg.* **61**(79): 17960–18011.
195. Venkatapathy, R., Moudgal, C.J., and Bruce, R.M. (2004). Assessment of the oral rat chronic lowest observed adverse

effect level model in TOPKAT, a QSAR software package for toxicity prediction. *JCICS* **44**(5): 1623–1629.

196. Moudgal, C.J., Lipscomb, J.C., and Bruce, R.M. (2000). Potential health effects of drinking water disinfection by-products using quantitative structure-toxicity relationship. *Toxicology* **147**: 109–131.

197. U.S. EPA. (1988). Health and Environmental Effects Document for Chloroacetic Acid. ECAO-CIN-G038. Environmental Criteria and Assessment Office, Office of Health and Environmental Assessment, Office of Research and Development, Cincinnati, OH.

HEALTH EFFECTS OF MICROBIAL CONTAMINANTS AND BIOTOXINS IN DRINKING WATER

Raghuraman Venkatapathy
Oak Ridge Institute for Science and Education
Cincinnati, Ohio

Brenda Boutin
Chandrika Moudgal
Robert Bruce
National Center for Environmental Assessment
Cincinnati, Ohio

Although drinking water quality has increased dramatically over the last century because of better treatment of raw water through processes such as disinfection, protection of source waters, and better sewage disposal practices, increasing pressure on available water sources through factors such as increased population, increased consumption, inadequate sewage treatment, runoffs, confined feedlot operations, and floods have resulted in a heavier load of contaminating micro-organisms into many water treatment plants. Established disinfection and filtration techniques may not always remove such microbial contaminants from drinking water. Although a majority of these contaminants are harmless, a few of them are pathogens and may cause diseases in populations exposed to the contaminated drinking water, especially in sensitive individuals and in the immunocompromised.

Drinking water pathogens that have been implicated in waterborne diseases include algae, bacteria, fungi, protozoa, and viruses. Waterborne diseases are usually acute in nature, are rarely fatal, and are often characterized by gastrointestinal symptoms (1). The severity and duration of illness is generally greater in weakened immune systems such as in children, the elderly, people on chemotherapeutic drugs, and those with AIDS (1).

ALGAE

Unlike other waterborne micro-organisms, such as bacteria, viruses, protozoa, and fungi, algae use photosynthesis as their primary mode of nutrition. Algae are not pathogenic to humans, but some species produce biotoxins that may be harmful to humans at high enough concentrations. In particular, algae produce toxins that affect the liver (hepatotoxins) and the nervous system (neurotoxins). The most environmentally relevant toxins are produced by cyanobacteria.

Cyanobacteria

Cyanobacteria, or blue-green algae, have characteristics of both algae and bacteria, and they generally occur in calm, nutrient-rich waters. People may be exposed to cyanobacterial toxins by drinking or bathing in contaminated water. Symptoms include skin irritation, stomach cramps, vomiting, diarrhea, fever, sore throat, headache, muscle and joint pain, blisters, and liver damage (2). Species that produce hepatotoxins include some strains of *Microcystis, Anabaena, Oscillatoria, Nodularia, Nostoc, Cylindrospermosis*, and *Umezakia*, whereas some strains of *Aphanizomenon* and *Oscillatoria* produce neurotoxins (2). Although physical methods of coagulation and filtration may remove cyanobacteria from water, ozonation and treatment with permanganate are necessary to destroy the toxins (2).

BACTERIA

Bacteria are cylindrical or spherical single-celled microorganisms that possess no nucleus. Common types of pathogenic bacteria that are found in drinking water include *Salmonella, Shigella, Campylobacter, Legionella, Escherichia coli, Vibrio cholerae*, and *Pseudomonas*. The health effects of common bacterial pathogens are described below.

Aeromonas

Aeromonas is an opportunistic bacterium that is commonly found in ground and surface waters, sewage, biofilms, and filtration units, and it may occasionally cause enteritis, septicemia, colitis, and puncture wound infections, particularly in the immunocompromised (3–6). Several species of *Aeromonas* may exist in the source waters including *A. bestiarum, A. caviae, A. eucrenophilia, A. hydrophilia, A. media, A. popoffii, A. salmonicida, A. schubertii*, and *A. veronii* (7). *Aeromonas* in water may be controlled through filtration and disinfection with chlorine, whereas *Aeromonas* in biofilms are relatively resistant to chlorination (8).

Campylobacter

C. coli and *C. jejuni* are found in sewage, ground and surface waters, and biofilms and are behind one of the most common bacterial causes of enteritis (3,9,10). *Campylobacter* may be controlled through chlorination (11).

Escherichia Coli

Of the more than 140 known serotypes of *E. coli*, 11 serotypes are known to cause gastrointestinal and other diseases in humans, including enteritis, colitis, kidney disease, and hemolytic uremic syndrome (3,9,12). The serotype *E. coli* O157:H7 is known to cause bloody diarrhea in infants (13). *E. coli* is generally transmitted through sewage and ground and surface waters, and it may be effectively controlled through disinfection (14).

Legionella

Of the 25 identified species of *Legionella*, a few may cause a type of pneumonia called Legionnaire's disease (13). The majority (90%) of the cases of Legionnaire's disease in the United States are caused by *L. pneumophila*, which is transported through ground and surface waters and aerosolized water, such as in hot water cooling towers, and is known to cause legionellosis, nosocomical and respiratory infections, pneumonia, and Pontiac fever (3,9,15,16). Free-floating *Legionella* may be effectively controlled through treatment with ozone, chlorine dioxide, and UV light, but *Legionella* inside suspended particles or free-living amoeba may be immune to disinfection (17).

Mycobacterium

Various species of NTM or nontuberculosis causing *Mycobacterium*, including *M. avium complex* (MAC), *M. chelonei, M. fortuitum, M. gordonae, M. kansasii, M. marinum, M. paratuberculosis, M. phlei*, and *M. xenopi*, are disseminated through biofilms, sewage, ground and surface waters, air, biofilms, and aerosolized water and can cause nosocomial infections, septicemia, skin lesions, various pulmonary diseases, and lymphadenitis, particularly in the immunocompromised (3). In addition, MAC colonizes water distribution systems, thereby recontaminating treated drinking water and generally infects the immunocompromised (3). *Mycobacterium* is relatively resistant to disinfection and can survive in hot water (18).

Pseudomonas

P. aeruginosa and *P. pseudomallei* are commonly found in ground and surface waters, sewage, and biofilms, and they cause various gastrointestinal, nosocomial, and respiratory infections in the immunocompromised (3,9). *Pseudomonas* may be disinfected through chlorination or UV treatment (19).

Salmonella

A majority of the waterborne *Salmonella* species are commonly found in sewage and surface waters and are pathogenic to humans. *S. enteritidis, S. montevideo B*, and *S. typhimurium* cause gastroenteritis and salmonellosis, whereas *S. typhi* and *S. paratyphi* (A & B) cause typhoid and paratyphoid fevers, respectively (9,20). *Salmonella* may be controlled by disinfection (21,22).

Shigella

S. sonnei, S. flexneri, S. boydii, and *S. dysenteriae* are found in sewage and ground and surface waters and may cause dysentery in exposed individuals (9,23). *S. sonnei* causes the bulk of waterborne infections among the four *Shigella* species listed above. Common disinfection processes are effective against controlling *Shigella* in drinking water.

Vibrio

V. cholerae is the most well known of the *Vibrio* species. It causes cholera, an acute intestinal disease with diarrhea, vomiting, and dehydration, and it is found in sewage and ground and surface waters (9,20). Severe cases of cholera may be fatal if left untreated for a few hours. *V. fluvialis, V. mimicus, V. parahaemolyticus*, and *V. vulnificus* are transmitted though surface waters and marine life and may cause enteritis and gastroenteritis (9). *Vibrio* may effectively be disinfected through treatment with chlorine (24).

Yersinia Enterocolitica

Y. enterocolitica is found in sewage and surface water and is carried by animals and humans (9). It can cause gastrointestinal illness including enteritis (9). *Y. enterocolitica* can be controlled through the use of chlorine (25).

FUNGI

Species such as *Aspergillus, Penicillium, Alternaria, Candida, Trichophyton*, and *Cladosporium* have been isolated from chlorinated water and service mains (26,27). Inhalation of large numbers of fungal spores can cause respiratory diseases, including pneumonia, fever, and meningoencephalitis, especially in the immunocompromised (28,29). Fungal spores are relatively resistant to chlorine (30), but a combination of coagulation, slow sand filtration, and disinfection can reduce the number of spores to a limited extent (31).

PROTOZOA

Protozoa are single-celled nucleated organisms that are typically larger than bacteria and lack a cell wall. Protozoa that are waterborne and pathogenic to humans are often found as resistant forms that protect themselves from environmental stresses (13). Common protozoa found in drinking water include *Giardia, Entamoeba*, and *Cryptosporidium*. The health effects associated with exposure to protozoa are described below.

Acanthamoeba

Acanthamoeba is a group of amoebae found in sewage and drinking water that can cause meningoencephalitis and amebic keratitis (9,32,33) and have been associated with causing cutaneous lesions and sinusitis in the immunocompromised (33). *Acanthamoeba* are relatively resistant to disinfection and have to be removed through filtration (34).

Cryptosporidium

Cryptosporidium, especially *C. parvum*, is commonly found in ground and surface waters, sewage, and springs and causes cryptosporidiosis in humans (9,35,36). Cryptosporidiosis may be severe and chronic in the immunocompromised and may hasten death (37). *Cryptosporidium* is resistant to disinfection by chlorine and chloramines, although ozone, chlorine dioxide, and filtration are more effective (38,39).

Cyclospora

C. cayetanensis is found in sewage and surface waters and causes gastroenteritis including diarrhea, abdominal cramping, decreased appetite, and fever (40). The symptoms may be chronic and unremitting in the immunocompromised (41). *Cyclospora* may be resistant to chlorine and have to be removed through filtration (42).

Entamoeba

E. histolytica is found in drinking water contaminated by sewage and causes amoebic dysentery with symptoms ranging from acute bloody diarrhea and fever to mild gastroenteritis (9,13). The effects of disinfection on *Entamoeba* are unknown. However, *Entamoeba* cysts are less robust than those of some other types of protozoa.

Giardia

G. lamblia are found in ground and surface waters, sewage, and springs and cause giardiasis, a gastrointestinal disease manifested by diarrhea, fatigue, and cramps (9,43). *Giardia* cysts can remain infective in water for 1 to 3 months, depending on the temperature (13). *Giardia* cysts are resistant to chlorine disinfection but may be removed through filtration (44).

Microspora

Microsporidia, especially *Enterocytozoon bieneusi* and *Encephalitozoon intestinalis*, are fairly common in the environment and may cause microsporidiosis in humans, especially the immunocompromised (45,46). Symptoms include diarrhea and illnesses of the respiratory tract, urogenital tract, eyes, kidneys, liver, or muscles (45–47). *Microsporidia* spores are more susceptible to disinfection using chlorine than *Cryptosporidia* (48).

Naegleria

N. fowleri is an opportunistic amoeba that is primarily found in ambient waters, drinking water, and recreational facilities that contain warm water, such as swimming pools, and causes primary amoebic meningoencephalitis (PAM), a fatal disease that leads to death within 72 hours after the appearance of symptoms (9,49). However, no reported occurrences of *Naegleria* outbreaks through drinking water in the United States exist, although there have been several fatalities from *N. fowleri* contacted through recreational water (49). Data suggest that disinfection through chlorination may not be effective in removing *Naegleria* at the doses typically used in treatment plants (50). However, exposure to *Naegleria* may be reduced by avoiding contact with warm waters.

Toxoplasma

T. gondii is found in surface waters and may cause toxoplasmosis in fetuses and the immunocompromised (51). Although *Toxoplasma* oocysts originate from cat feces, transmission may also come from tissue or meat of other infected animals (51). Symptoms include fever, swelling of the lymph glands, blindness, mental retardation, and encephalitis (52). Although chlorination is not effective against *Toxoplasma*, filtration may be used in removing the organism (53).

VIRUS

Viruses are a large group of infectious agents that consist of a protein coat and a nucleic acid core. Most pathogenic waterborne viruses cause gastrointestinal diseases, whereas others may cause serious sequelae such as diabetes. Pathogenic viruses found in drinking water include adenoviruses, caliciviruses, Norwalk virus, coxsackieviruses, echovirus, hepatitis viruses, poliovirus, and rotavirus. The health effects of common viruses are described below.

Adenovirus

Adenoviruses are transported through ground and surface waters, sewage, and mucus and cause gastroenteritis, pharyngitis, conjunctivitis, and the common cold, especially in children (9). Adenoviruses are relatively resistant to disinfectants (13).

Astrovirus

Astroviruses are small, star-like viruses that are transported through ground and surface waters and sewage and cause gastroenteritis, especially in very young children and in the immunocompromised (9,54). Astroviruses are resistant to disinfection, but infection may be prevented through frequent handwashing (55).

Calicivirus

The calicivirus, including the Norwalk virus, is a common cause of gastroenteritis and is transported through ground and surface waters and sewage (9). Illness is generally mild, although gastroenteritis-induced dehydration may cause death in infants and young children (9). Ozone disinfection is effective against the caliciviruses (56).

Enterovirus

The most common enteroviruses (viruses that are found in the intestine) include the coxsackieviruses, echoviruses, and the polioviruses. Enteroviruses are generally found in ground and surface waters and sewage (9). Coxsackieviruses and echoviruses produce a variety of illnesses, including meningitis, enteritis, cold, fever, and heart disease (9). The polioviruses cause poliomyelitis, a paralytic disease (9). Enteroviruses may be resistant to chlorination (57).

Hepatitis Virus

Hepatitis A and E viruses cause infectious hepatitis, an illness characterized by inflammation of liver, fever, abdominal pain, weakness, and a lack of appetite (3,9). Hepatitis E infections are mild except for pregnant women, who may have a fatality rate of about 39% (13). Hepatitis A virus may be more resistant to disinfection than other enteric viruses (58).

Rotavirus

Rotaviruses are found in sewage and ground and surface waters and cause acute gastroenteritis, especially in children (3,9). Rotaviruses may be removed through filtration or inactivated by chlorine, chlorine dioxide, and ozone (59–61).

BIBLIOGRAPHY

1. CDC. (2000). *Safe Water Systems for the Developing World: A Handbook for Implementing Household-Based Water*

Treatment and Safe Storage Projects. Department of Health and Human Services, Centers for Disease Control and Prevention, Atlanta, GA.

2. World Health Organization. (1999). *Toxic Cyanobacteria in Water: A Guide to their Public Health Consequences, Monitoring and Management*. J. Bartram and I. Chorus (Eds.). WHO, Geneva, Switzerland.
3. Reasoner, D.J. (1991). Pathogens in drinking water—are there any new ones? *Proc. AWWA Water Quality Technol. Conf.* Denver, CO.
4. Moyer, N.P. (1989). *Aeromonas* gastroenteritis: another waterborne disease? *Proc. AWWA Water Quality Tech. Conf.* Denver, CO.
5. van der Kooij, D. (1991). Nutritional requirements of Aeromonads and their multiplication in drinking water. *Experientia.* 47, Birkauser Verlag, Basel, Switzerland.
6. Havelaar, A.H., Versteegh, J.F.M., and During, M. (1990). The presence of *Aeromonas* in drinking water supplies in the Netherlands. *Zbl. Hyg.* **190**: 236–256.
7. Szczuka, E. and Kaznowski, A. (2004). Typing of clinical and environmental *Aeromonas* sp. strains by random amplified polymorphic DNA PCR, repetitive extragenic palindromic PCR, and enterobacterial repetitive intergenic consensus sequence PCR. *J. Clin. Microbiol.* **42**(1): 220–228.
8. Meheus, J. and Peeters, P. (1989). Preventive and corrective actions to cope with *Aeromonas* growth in water treatment. *Water Supply* **7**: 10-1–10-4.
9. Hazen, T.C. and Toranzos, G.A. (1990). Tropical source waters. In: *Drinking Water Microbiology*, G.A. McFeters (Ed.). Springer-Verlag, New York, pp. 32–53.
10. Fung, D.Y.C. (1992). Foodborne illness. In: *Encyclopedia of Microbiology*, Vol. 2, J. Lederberg (Ed.). Academic Press, San Diego, CA, pp. 209–218.
11. Blaser, M.J., Smith, P.F., Wang, W-L.L., and Hoff, J.C. (1986). Inactivation of *Campylobacter jejuni* by chlorine and monochloramine. *Appl. Environ. Microbiol.* **51**: 307–311.
12. Principie, M.A., Faber, W.W., Freud, S., and Wang, X.P. (1995). *Escherichia coli* occurrences in the New York City distribution system. *Proc. 1995 AWWA Annu. Conf.* Denver, CO.
13. Cohn, P.D., Cox, M., and Berger, P.S. (1995). Health and aesthetic aspects of water quality. In: *Water Quality and Treatment*, 5th Edn. R.D. Letterman (Ed.). McGraw-Hill, New York.
14. Mitchell, R. (1972). *Water Pollution Microbiology*. John Wiley & Sons, New York.
15. States, S.J., Conley, L.F., Knezevich, C.R., Keleti, G., Sykora, J.L., Wadowsky, R.M., and Yee, R.B. (1989). Free-living amoebae in public water supplies: Implications for *Legionella, Giardia*, and *Cryptosporidia*. *Proc. Annu. AWWA Water Quality Technol. Conf.* Denver, CO.
16. Harf, C. and Monteil, H. (1988). Interactions between free-living Amoebae and *Legionella* in the environment. *WST* **20**(11–12): 235–239.
17. States, S.J., Wadowsky, R.M., Kuchta, J.M., Wolford, R.S., Conley, L.F., and Yee, R.B. (1990). *Legionella* in drinking water. In: *Drinking Water Microbiology*, G.A. McFeters (Ed.). Springer-Verlag, New York.
18. Pelletier, P.A., du Moulin, G.C., and Stottmeier, K.D. (1988). Mycobacteria in public water supplies: Comparative resistance to chlorine. *Microbiol. Sci.* **5**: 147–148.
19. Homberger, F.R., Pataki, Z., and Thomann, P.E. (1993). Control of *Pseudomonas aeruginosa* infection in mice by chlorine treatment of drinking water. *Lab Anim. Sci.* **43**(6): 635–637.
20. Brock, T.D. (1979). *Biology of Microorganisms*, 3rd Edn. Prentice-Hall, Englewood Cliffs, NJ.
21. Engelbrecht, R.S., Severin, B.F., Masarik, M.T., Farooq, S., and Lee, S.H. (1975). New Microbial indicators of disinfection efficiency. U.S. National Technical Information Service, AD Report, AD-A03057.
22. Butterfield, C.T., Wattie, E., Megregian, S., and Chambers, C.W. (1943). Influence of pH and temperature on the survival of coliforms and enteric pathogens when exposed to free chlorine. *Public Health Rep.* **58**: 1837–1866.
23. Metcalf and Eddy, Inc. (1979). *Wastewater Engineering: Treatment, Disposal, Reuse*. McGraw-Hill, New York.
24. Rice, E.W. et al. (1993). *Vibrio cholerae* 01 can assume a 'rugose' survival form that resists killing by chlorine, yet retains virulence. *Int. J. Environ. Health Res.* **3**: 89–98.
25. Lund, V. (1996). Evaluation of *E. coli* as an indicator for the presence of *Campylobacter jejuni* and *Yersinia enterocolitica* in chlorinated and untreated oligotrophic lake water. *Water Res.* **30**: 1528–1534.
26. Rosenzweig, W.D., Minnigh, H.A., and Pipes, W.O. (1986). Fungi in potable water distribution systems. *J. AWWA.* **78**: 53–55.
27. Frankova, E. and Horecka, M. (1995). Filamentous soil fungi and unidentified bacteria in drinking water from wells and water mains near Brastislava. *Microbial. Res.* **150**: 311–313.
28. Bennett, J.E. (1994). Fungal Infections. In: *Harrison's Principles of Internal Medicine*, 13th Edn., K.J. Isselbacher, E. Braunwald, J.D. Wilson, J.B. Martin, A.S. Fauci, and D.L. Kasper (Eds.). McGraw-Hill, New York, pp. 854–865.
29. Geldreich, E.E. (1990). Microbial quality control in distribution systems. In: *Water Quality and Treatment*, F.W. Pontius (Ed.). American Water Works Association, McGraw-Hill, New York.
30. Rosenzweig, W.D., Minnigh, H.A., and Pipes, W.O. (1983). Chlorine demand and inactivation of fungal propagules. *Appl. Environ. Microbiol.* **45**: 182–186.
31. Niemi, R.M., Knuth, S., and Lundstrom, K. (1982). Actinomycetes and fungi in surface waters and in potable water. *Appl. Environ. Microbiol.* **43**: 378–388.
32. Sawyer, T.K. (1989). Free-living pathogenic and nonpathogenic Amoebae in Maryland soils. *Appl. Environ. Microbiol.* **55**: 1074–1077.
33. Marciano-Cabral, F. and Cabral, G. (2003). *Acanthamoeba* spp. as agents of disease in humans. *Clin. Microbiol. Rev.* **16**(2): 273–307.
34. de Jonchkheere, J. and van de Voorde, H. (1976). Differences in destruction of cysts of pathogenic and nonpathogenic *Naegleria* and *Acanthamoeba* by chlorine. *Appl. Environ. Microbiol.* **31**: 294–297.
35. Pontius, F.W. (1993). Protecting the public against *Cryptosporidium*. *J. AWWA.* **85**(8): 18, 22, 122–123.
36. Craun, G.F. et al. (1998). Waterborne outbreaks of Cryptosporidiosis. *J. AWWA.* **90**(9): 81–91.
37. Janoff, E.N. and Reller, L.B. (1987). *Cryptosporidium* species, a protean Protozoan. *J. Clin. Microbiol.* **25**: 967–975.
38. Korich, D.G. et al. (1990). Effects of ozone, chlorine dioxide, chlorine, and monochloramine on *Cryptosporidium parvum* oocyst viability. *Appl. Environ. Microbiol.* **56**: 1423–1428.
39. Peeters, J.E. et al. Effect of disinfection of drinking water with ozone or chlorine dioxide on survival of *Cryptosporidium parvum* oocysts. *Appl. Environ. Microbiol.* **55**: 1519–1522.

40. Huang, P. et al. (1995). The first reported outbreak of diarrheal illness associated with *Cyclospora* in the United States. *Ann. Intern. Med.* **123**(6): 409–414.
41. Soave, R. and Johnson, W.D. (1995). *Cyclospora*: Conquest of an emerging pathogen. *Lancet* **345**: 667–678.
42. Rabold, J.G. et al. (1994). *Cyclospora* outbreak associated with chlorinated drinking water. *Lancet* **244**: 1360–1361.
43. Crockett, C.S. and Haas, C.N. (1997). Understanding Protozoa in your watershed. *J. AWWA*. **89**(9): 62–73.
44. Logsdon, G.S., Dewalle, F.B., and Hendricks, D.W. (1984). Filtration as a barrier to passage of cysts in drinking water. In: *Giardia and Giardiasis*, S.L. Erlandsen and E.A. Meyer (Eds.). Plenum Press, New York.
45. Goodgame, R.W. (1996). Understanding intestinal spore-forming Protozoa: *Cryptosporidia, Microsporidia, Isospora*, and *Cyclospora*. *Ann. Intern. Med.* **1214**: 429–441.
46. Bryan, R.T. (1995). Microsporidiosis as an AIDS-related opportunistic infection. *Clin. Infect. Dis.* **21**(Suppl. 1): 62–65.
47. Cali, A. (1991). General *Microsporidium* features and recent findings on AIDS isolates. *J. Protozool.* **38**: 625–630.
48. Rice, E.W., Johnson, C.H., Naumovitz, D.W., Marshall, M.M., Plummer, C.B., and Sterling, C.R. (1999). Chlorine disinfection studies of *Encephalitozoon* (*Septata*) *Intestinalis*. Abstracts, 99th General Meeting, American Society for Microbiology, Q-130.
49. CDC. (1992). Primary Amoebic Meningoencephalitis—North Carolina, 1991. *Morbidity and Mortality Weekly Report 41*, 437–440.
50. Chang, S-L. (1978). Resistance of pathogenic *Naegleria* to some common physical and chemical agents. *Appl. Environ. Microbiol.* **35**: 368–375.
51. Sinclair, M. (1997). Health Stream. Cooperative Research Centre for Water Quality and Treatment. **8**: 1.
52. Fishback, J.L. (1992). Toxoplasmosis. In: *Encyclopedia of Microbiology*. Vol. 4. J. Lederberg (Ed.). Academic Press, San Diego, CA, pp. 255–264.
53. Girdwood, R.W.A. (1995). Some clinical perspectives on waterborne parasitic Protozoa. In: *Proceedings, Protozoan Parasites and Water*. Symposium held at York, England, September 1994. W.B. Betts, D. Casemore, C. Fricker, H. Smith, and J. Watkins (Eds.). Royal Society of Chemistry, London.
54. Kurtz, J.B. and Lee, T.W. (1987). Astroviruses: Human and animal. In: *Novel Diarrhea Viruses*. John Wiley & Sons, New York.
55. White and Fenner. (1994). *Med. Virol*. 415–416.
56. Shin, G-A. and Sobsey, M.D. (2003). Reduction of Norwalk Virus, Poliovirus 1, and Bacteriophage MS2 by ozone disinfection of water. *Appl. Environ. Microbiol.* **69**(7): 3975–3978.
57. Abad, F.X., Pinto, R.M., and Bosch, A. (1997). Disinfection of human enteric viruses on fomites. *FEMS Microbiol Lett.* **156**(1): 107–111.
58. Peterson, D.A., Hurley, T.R., Hoff, J.C., and Wolfe, L.G. (1983). Effect of chlorine treatment on infectivity of Hepatitis A virus. *Appl. Environ. Microbiol.* **45**: 223–227.
59. Rao, V.C. et al. (1988). Removal of Hepatitis A virus and Rotavirus by drinking water treatment. *J. AWWA*. **80**: 59–67.
60. Berman, D. and Hoff, J.C. (1984). Inactivation of simian rotavirus SA11 by chlorine, chlorine dioxide, and monochloramine. *Appl. Environ. Microbiol.* **48** 317–323.
61. Chen, Y-S. and Vaughn, J.M. (1990). Inactivation of human and simian rotaviruses by chlorine dioxide. *Appl. Environ. Microbiol.* **56**: 1363–1366.

DRINKING WATER AND PUBLIC HEALTH PROTECTION

Daniel A. Okun
(from *Drinking Water Regulation and Health*, Wiley 2003)
University of North Carolina
Chapel Hill, North Carolina

INTRODUCTION

The provision of drinking water for communities is an urban utility, but a utility with a difference. As with other urban utilities, such as electricity and gas, water for household use is a necessity that cannot readily be obtained by urban householders for themselves. The difference is that, while water may satisfy many household needs, including drinking, it has the potential of spreading disease, often without the knowledge of the consumer. As a result, water supplies have become subject to regulations for assuring adequate quality, regulations that are not faced by other municipal public utilities.

Beginning with the water supply for Rome some 2000 years ago, the responsibility for water supply and its quality rested with the community. During the nineteenth century, with the beginning of the industrial era and the rapid growth of cities, public water supplies began to be provided by private entrepreneurs who sought profit in providing an essential service, frequently in competition with others. In the interest of getting a larger share of the market, they might provide a water of better quality than a competitor. The experiences with the provision of water for London from the Thames in the 1850s illustrate that the selection of the source of a water supply is important. Then Dr. John Snow took advantage of the competition between two water suppliers to prove that water was responsible for the transmission of cholera.

As cities grew, the need for large capital investments to provide adequate water supplies of high quality resulted in most cities abandoning private utilities when it became clear that they did not have the financial resources for the construction of the reservoirs, the long transmission lines, and the treatment plants. Decisions for selection of sources and treatment, which were introduced in the late nineteenth—early twentieth century, became the responsibility of the community, and not a regulatory body.

Treatment in the form of filtration and then chlorination was widely introduced, although not primarily through regulation. City officials recognized that they had an obligation to their constituencies to provide water that would not spread typhoid and cholera. Some cities were slow to assume this responsibility, and, in the United States, some newly organized state health agencies began to institute regulations. The choice of sources between a costly high-quality upland supply and a polluted source and the treatment to be provided was local.

The first nationwide water quality regulations in the United States were introduced by the federal government in 1909 to assure the safety of water to which the public was exposed in interstate and international traffic. Many states adopted these regulations even for smaller cities that did not have train or bus service. These federal regulations were upgraded over the years, and the bulk of this chapter is devoted to the nature of these regulations at the federal level until the passage of the Safe Drinking Water Act (SDWA) in 1974, after the U.S. Environmental Protection Agency (USEPA) became responsible for ensuring the safety of all public water supplies.

This article recounts the high points in the history of the role that urban water supplies play in the health of those who are obliged to drink from public supplies, beginning with concerns with the water supply for Rome, followed by the story of the cholera outbreaks in London that led to the recognition that water was responsible for the spread of infectious disease, the introduction of successful public health measures to control infectious disease, and the explosion of the chemical revolution that became responsible for the spread of chronic disease through ingestion of public water supplies (1).

WATER SUPPLY FOR THE CITY OF ROME

Among the major ancient cities of the world, none was better provided with water for its citizens than Rome. Initially, the city obtained its water from the Tiber River, which ran through the city. When it was apparent that the water had become heavily polluted, Appius Claudius built an aqueduct, the Aqua Appia, in 312 B.C. from the Tiber, about 11 miles upstream. Some 40 years later, the need was so great that a second aqueduct, 40 miles long, the Anio Novis, was built. Sextus Julius Frontinus, the water commissioner of Rome, wrote two books describing the water works of the city and their management (2). By A.D. 305, 14 aqueducts were serving the city.

The aqueducts fed the city by gravity with relatively short sections passing over valleys on stone arches, some three tiers high. Many of them carried water into the twentieth century. Such aqueducts remain throughout Europe and the Middle East as monuments to the early Romans.

The water from the aqueducts passed through large cisterns and from these was distributed through lead pipes to other cisterns, to public buildings, baths, and fountains, and to a relatively small number of private residences. Incidentally, they also built stone sewers to carry off wastewater from bathtubs and toilets in the larger buildings.

Frontinus questioned the wisdom of Augustus, whom he considered a most cautious ruler, in building one of the aqueducts, the Alsietinian, because the quality of its water was very poor and not suitable for the people. He speculated that Augustus built the aqueduct to serve nonpotable purposes and thereby "to avoid drawing on better sources of supply." The most important nonpotable use was for a *naumachia*, an artificial lake that was used for exhibitions of sham naval battles (Fig. 1). This is also current practice in American cities that erect stadia for baseball, football, and basketball on behalf of the team owners. The surplus nonpotable water was used for landscape irrigation and fountains. Words from an inscription state: "I gave the people the spectacle of a naval combat.... Besides the rowers, three thousand men fought in these fleets."

Thus, Rome can claim to be the first city to employ a dual distribution system and to base the use of its water supply on its quality. The water quality from the aqueducts was variable, and the Marcia aqueduct carried the best water. Frontinus points out that it was "determined to separate (the aqueducts) and then to arrange that the Marcia should serve wholly for drinking purposes, and that the others should be used for purposes adapted to their special qualities."

It is interesting to note that, in 1958, some 2000 years later, the United Nations Economic and Social Council enunciated a policy (3): "No higher quality water, unless there is a surplus of it, should be used for a purpose that can tolerate a lower grade."

THE MIDDLE AGES AND THE INDUSTRIAL REVOLUTION

Beginning in the sixth century, the Roman Empire began to disintegrate and, up to the fourteenth century, infectious diseases rode rampant throughout Europe. Leprosy, bubonic plague, smallpox, diphtheria, measles, influenza, and countless other afflictions were epidemic, particularly

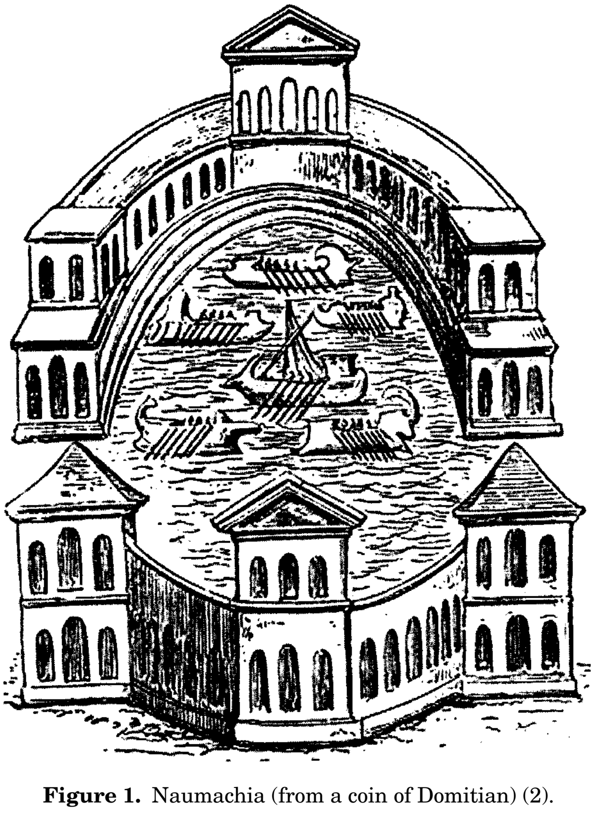

Figure 1. Naumachia (from a coin of Domitian) (2).

in the cities. Water was only one of the many vectors for the spread of disease. Knowledge of the specific vectors was limited, and food received the most attention. Quarantine was the principal approach to control of the spread of disease. The lack of proper sanitation and the dense urban populations were largely responsible for the epidemics and there was little focus on water quality and its availability.

The major accomplishment toward the end of the Middle Ages was the establishment of hospitals, often for specific diseases, by local governments and workers' guilds. Little of importance with regard to the water environment and the public health emerged during that period.

THE GREAT SANITARY AWAKENING

In the middle of the nineteenth Century, the causes of the many common diseases of the day that afflicted the growing urban populations that accompanied the Industrial Revolution were still unknown. Water was beginning to be piped to houses of the well-to-do while the poor either carried their water from wells or bought water from purveyors who obtained the water at the most convenient sources. When the water was contaminated, which was its general condition in urban areas, the spread of disease was inevitable.

A more significant and serious situation resulted from the growing installation rate of piping and then water closets in homes and commercial establishments. In addition to the impact of a poor quality water for drinking was the necessity for disposing of the discharges of these new flush toilets. London had found it necessary to construct storm sewers to drain the streets to permit the conduct of commerce. The obvious solution was to discharge the household wastes from the toilets to the storm sewers, which, in turn, discharged directly into the Thames River, which ran through London and served as a source of water for several private companies that distributed the water to households.

London was exemplary of the unsavory and squalid conditions in all cities in the early years of the century. The medical fraternity believed that the diseases were spread by poisons in the miasmatic air emanating from the "bowels of the earth." The Thames at London at that time was a tidal river and the heavily polluted waters would flow very slowly to sea. In warm periods, Londoners avoided crossing London Bridge because the air was so foul. A headline of the period read "India is in Revolt and the Thames Stinks." The drapery in the Houses of Parliament, located on bank of the Thames, needed to be soaked in chloride of lime to make the meeting room tolerable, and stirred the Parliament to establish the first of many committees to see to alleviating the situation.

Two cholera outbreaks in the summer of 1854 were the greatest in London's history. The first developed in Soho, a densely populated section in the heart of the city. Dr. John Snow, then physician to Queen Victoria Hospital, and reasonably the first epidemiologist, undertook to mark the deaths in the summer of 1953. In 2 days, 197 people died, and after 10 days more than 500 people died in an area only 250 yards across (4). Plotting the deaths on a map of the area (Fig. 2), the result resembled a target, with the greatest concentration of hits at the center. A church-owned well on Broad Street was identified at the site as being the source of the water ingested by the victims. The water had appeared to be of excellent quality. A woman living about a mile away regularly sent a cart to carry water to her home; she and a guest from outside London died of cholera in that epidemic.

Dr. Snow examined the well site and concluded that a tannery on property owned by the church had a cesspool for discharge of its wastewaters. He ordered the church to remove the handle on the pump, ending the epidemic, but, by that time, the epidemic might well have been spent. At any rate, this demonstration was the first to suggest that drinking water was the source of the cholera. This was generations before the germ theory of disease had been elucidated, and Snow's other studies in London were even more convincing. The John Snow Pub is on the site of the Broad Street pump, and these data decorate its walls.

Annual death rates from cholera among households using Thames River water ranged from 10 to 110 per 10,000 households in 1832, increasing to 200 per 10,000 among those taking water from the downstream reaches of the river. While this justified the inference that water was responsible, Dr. Snow found a more definitive proof during the 1854 epidemic. Two private water companies, the Southwark and Vauxhall Company and the Lambeth Company were in direct competition, serving piped water to the same area near the center of London but on the south side of the river. These water companies were characterized as "by far the worst of all those that take their water from the Thames, with 120 to 180 deaths per 10,000 households in 1849 for each of the two companies" (6). In 1852, however, the Lambeth Company, to attract more customers, improved the aesthetic quality of the Thames River water by moving its intake upstream above the heaviest pollution from London.

Snow's data showed that, in the 1854 epidemic, the death rate among those using Lambeth water was 37 deaths per 10,000 households as compared with 315 per 10,000 households for those using the intake downstream. During that period, the death rate in all of London was 59 per 10,000 households (256,423 deaths) among those taking water from all sources in London.

In addition to establishing that the cholera outbreaks were caused by drinking water, Snow demonstrated the importance of source selection. As is pointed out below, almost a century later, some cities still chose to take water from run-of-river sources when better sources were available primarily because it was less costly. Professor Fair, in presenting his philosophy about water supply, characterized the issue by declaring that he "preferred the virginal to the repentant," a paraphrase of the philosophy of Allen Hazen, possibly the most important engineer in the early history of water supply in the United States, who put it: "Innocence is better than repentance." (7).

THE EMERGENCE OF WATER AS A PUBLIC HEALTH ISSUE

The Industrial Revolution, beginning in the late eighteenth century in Britain and extending to Europe and the United States, was responsible for an explosive increase in

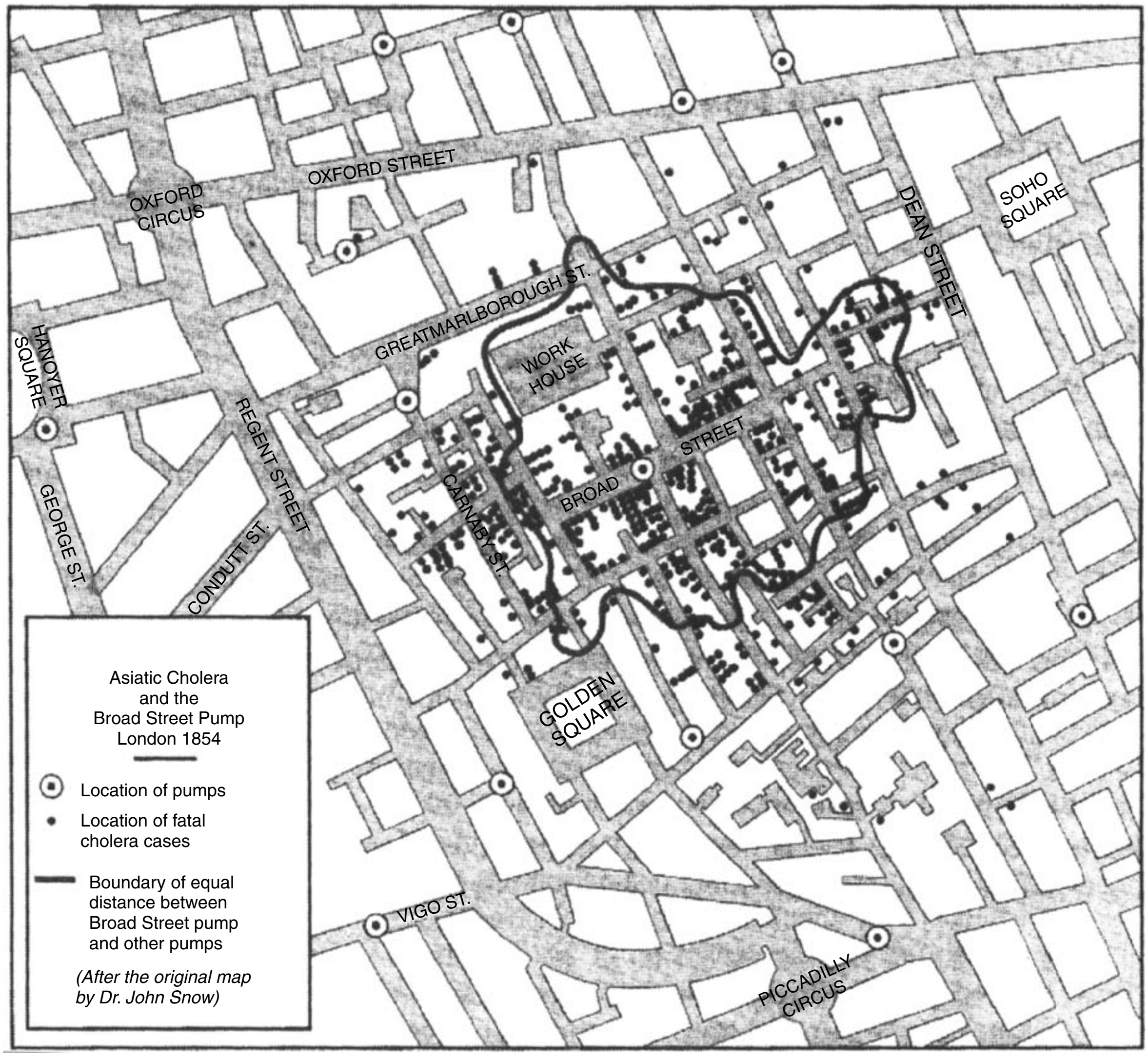

Figure 2. Map of Soho showing the location of those who died from cholera within the vicinity of the Broad Street pump in London 1854 (5).

urbanization with the development of the slums so "celebrated" by Dickens. It did eventually result in the English government and the northeastern states in the United States establishing agencies for addressing the terrible health conditions that emerged. Massachusetts, Pennsylvania, and New York established health boards to improve housing conditions; this resulted in the establishment of regulations for water supply and disposal of household wastes (Fig. 3).

These efforts at regulating activities that might damage the environment led to the establishment of the public health movement. Two figures of lasting fame: Sir Edwin Chadwick, a lawyer, in England (8), and Lemuel Shattuck, a physician (9) in Massachusetts, who was inspired by Chadwick, were responsible for the creation of regulatory agencies and laws protecting the public from the wide range of microbial and chemical contaminants that inevitably found their way to the nearby streams and rivers that were drawn upon for water supply. Shattuck's plan for the board of health for Massachusetts called for its membership to be composed of two physicians, one counselor-at-law, one chemist or natural philosopher, one civil engineer, and two persons of other professions or occupations. This comprehensive view of the needs for an agency for the protection of the public health was the springboard for the establishment of a sanitary engineering specialty within civil engineering. Shattuck had pointed out that competence in "planning and constructing public works" was essential to the provision of water supply and the disposal of household wastes.

In 1886, the Massachusetts legislature passed "An Act to Protect the Purity of Inland Waters" and, to implement the Act, it called for the establishment of an engineering department in the State Board of Health. Among its activities was the establishment of the Lawrence Experiment Station, the first of its kind, which was instrumental in attracting engineers, chemists, and biologists from the Massachusetts Institute of Technology, many of whom were responsible not only for spreading the study of water-related diseases and their control but also

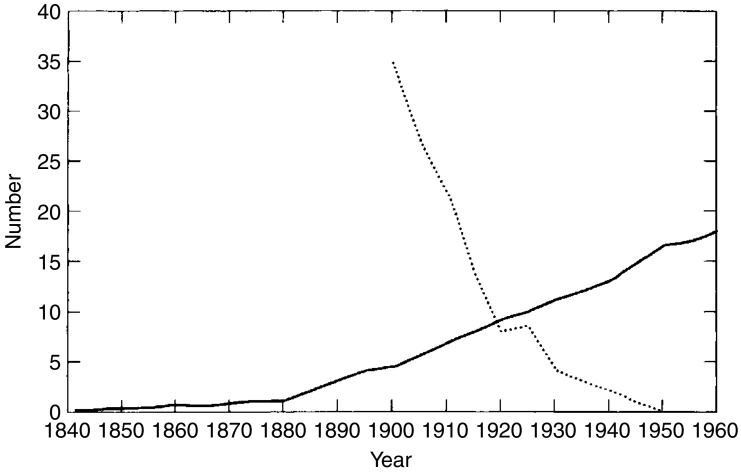

Figure 3. Simultaneous decline in typhoid fever death rate and rise in number of community water supplies in the United States (——deaths per 100,000 population; water supplies: 1000s) (*Source*: F. W. Pontius).

in the introduction of community water supply systems. From a total of only 17 water supply systems in the state in 1869, the number grew to 138 in 1890 while the annual death rate from typhoid fever in the state dropped from 89 per 100,000 in 1873 to 37 in 1890, and by 1940 to 0.2 (Fair 1945).

Despite the appearance of regulatory agencies, many years passed before they played a significant role in the monitoring of municipal water supply and wastewater collection, treatment, and disposal systems. Actually, there was, is, and should be far less need for regulation of drinking water quality than for regulation of wastewater discharges.

In the early days of public water supplies, most were privately owned and needed to meet the requirements of the communities they served. When they were inadequate to the task, sometimes because they failed to satisfy the communities they served, but more generally because the rapid growth of the cities called for capital investments beyond the capacity of the private purveyors to meet, the community government became responsible for the water supply. When the community government itself was providing the water, there seemed to be little need for regulating the performance of their own utility as its objectives would naturally be to protect its citizenry from public health risks.

A good example of this was the early history of the water supply for London, as mentioned above, where the private companies were generally loathe to invest in improvements. At the end of the nineteenth century, a Metropolitan Water Board (MWB) was created to take over responsibility for the water supply of London from eight private companies. In some other cities in England, private water companies continued to serve satisfactorily and continue to this day. The MWB established new technology and were seen to be at the leading edge of water supply technology and they set their own standards which were emulated by other communities.

In the case of New York City, its early private purveyors also were inadequate to their responsibilities. The driving forces were the need to have water to prevent epidemics of yellow fever (which were not related to water) and to fight fires. One of the last private efforts was that inspired by Aaron Burr, who promised a water supply as a condition of establishing the Bank of the Manhattan Company, the predecessor of the Chase Manhattan Bank. He had little interest in providing water and “this brilliant and unprincipled man suffered a series of political disasters that plunged him ... to ruin and exile.” (10). Burr’s plans were doomed.

The city finally decided to develop its own supply and, after extensive study had to choose between two possible sources: the Bronx River very near the city and the Croton River some 40 miles distant. The former was considerably lower in cost but the latter promised a much better quality of water and a greater quantity for the future. The City Board of Water Commissioners Committee on Fire and Water, addressing this question in 1835 opted for the Croton and for public ownership in this language (10):

> The question remains, ought the Corporation of the City of New York Committee to embark on this great work? The Committee are firmly of the opinion, that it ought to be done by no other body, corporate or personal. ... The control of the water of the City should be in the hands of this Corporation, or in other words, in the hands of the *people*.

The City celebrated the delivery of high-quality water from the Croton Aqueduct to New York City by gravity at high pressure in ample quantity in 1842, then one of the largest water supplies in the world. It still provides about 15% of the water that the City uses. This costly choice was made by the city officials not to meet a regulation but to serve their constituency well. Another example is the city of Cincinnati, Ohio, which installed granular activated carbon (GAC) filtration in the 1980s though it is not required by regulations. Many cities do more than the existing regulations require because the regulations tend to be years behind our knowledge. Water officials desiring to serve their community best may find it wise to anticipate quality problems that will not be addressed by regulations for years. Unfortunately, the reception given new regulations is not always one of appreciation by many water officials but of concern for the costs that may be involved. Industry groups such as the American Water Works Association (AWWA) often challenge regulations that are in the process of being promulgated to reduce public health risks because it would increase costs and water rates.

On the other hand, regulations for the quality and quantity of discharges of wastewaters to receiving waters are necessary, because the cost burden falls on the community while those who benefit are generally residents of other communities and not liable for the costs. This is also one of the reasons why the Clean Water Act (CWA) and similar earlier programs have been obliged to meet a significant share of the costs.

THE BEGINNING OF WATER TREATMENT

The relationship between source, water quality, and disease was demonstrated in the United States but much later than cholera in England and with much lower typhoid fever rates. Kober (11) made a study of typhoid rates in American cities from 1902 through 1908, summarized in Table 1. New York City, with its upland supply, had the lowest rate of the 61 cities with 15 typhoid deaths per 100,000 while Pittsburgh, with its run-of-river supply, suffered the highest death rate, 120 per 100,000.

Filtration of water was introduced well before the turn of the nineteenth century in Europe, where run-of-river supplies were more common. An eightfold increase of filtration in the United States reduced the typhoid death rate from water supply from 1900 to 1913 by 55% (12).

The availability of filtration mistakenly seemed to make the need for selecting better sources unnecessary. Philadelphia, which had been taking water from run-of-river sources and had been one of the last of the large U.S. cities to adopt filtration, was suffering a typhoid death rate of 75 per 100,000 into the twentieth century. The city officials had contended that filtration was not as effective as boiling the water. In 1900, a reform mayor was determined to address the issue. A panel of distinguished engineers prepared “a report that was characterized as not having any surprises.” (13). It recommended filtration and continued use of water from the lower Delaware and Schuylkill Rivers. The report stated that “Water from upcountry sources might be preferable but the great cost of building aqueducts and reservoirs made that option very expensive and really unnecessary since filtration would provide safe water.”

In 1911, before the filters were operational, a typhoid outbreak in Philadelphia resulted in 1063 deaths. After filtration, the death rate dropped to 13 per 100,000, still a relatively high figure.

Philadelphia still takes most of its water from the “mouth” (more properly, the “anus”) of the Delaware River and has had to adopt Herculean methods to deliver water of good quality ever since. A multimedia study of environmental health risks in Philadelphia in the 1980s determined that the water supply posed the highest risk of all sources of pollution in the city. Since then improved treatment processes along with stricter USEPA standards have been introduced.

Many cities have no alternatives and are obliged to draw from run-of-river sources. Slow sand filtration became the treatment of choice in Massachusetts in the 1870s. In the 1890s, the Louisville Water Company, which took water from the Ohio River, introduced sedimentation of the water prior to filtration. For better removal of turbidity, they introduced chemical coagulation and rapid sand filtration.

The introduction of chlorination for disinfection of water for municipal water supply took place in Boonton, New Jersey, in 1908 following decades of study of the use of chlorine in Europe and the United States (14). It clearly was the greatest step in the reduction of the transmission of infectious diseases via water supply.

An example of the role of chlorine was the effect it had in a city drawing water from a clear lake. Chlorine reduced the annual typhoid death rate from about 20 to 2 per 100,000 population, which was then reduced to virtually zero with the addition of filtration (Fair et al. 1968). Together with pasteurization of milk and better handling of human wastes, typhoid virtually disappeared in the United States by the middle of the twentieth century.

Table 1. Typhoid Rates in American Cities, 1902 Through 1908

Source	Number of Cities	Death Rate per 100,000
Groundwater	4	18.1
Impoundments and protected watersheds	18	18.5
Small lakes	8	19.3
Great lakes	7	33.1
Mixed surface and groundwater	5	45.7
Run-of-river supplies	19	61.6

Source: (12).

THE CHEMICAL REVOLUTION

While infectious disease was brought under control, although other diseases emerged later, two other problems arose. The first was that water treatment tools were believed to be so effective, engineers became sanguine about the need to seek waters of high quality; treatment would make it safe. The conventional treatment of the midtwentieth century, which remains the conventional treatment now, at the beginning of the twenty-first century—chemical coagulation, rapid sand filtration, and chlorination—does little to remove the trace synthetic organic chemicals in ambient water resulting from the post World War II surge in industrial development what has been labeled the “chemical revolution” (1).

The second problem is truly ironic—the life-saving treatment, chlorination, increases the risk from synthetic organic chemicals created by the chlorine itself. Other disinfection byproducts have surfaced and added to the problem of the trace synthetic organic chemicals discharged from industry and households using a wide range of such chemicals for house and garden.

The first published material about disinfection byproducts (DBPs) emanated from Rook's work at the Rotterdam water treatment plant, which drew water from the mouth of the Rhine River (15). While it had been picked up quickly by USEPA, the potential had been recognized 5 years earlier. Dr. Joshua Lederberg, a Nobel Prize geneticist, who had been somewhat active in drinking water issues, wrote a 1969 syndicated column in the Washington Post. One column was headlined "We're so accustomed to using chlorine that we tend to overlook its toxicity" (16):

> What little we do know of the chemistry of chlorine reactions is portentous. It should sometimes react . . . to form substances that may eventually reach and react with genetic material, DNA, of body cells . . . That chlorine is also intended to inactivate viruses should provoke questions about the production of mutagens in view of the close similarity between viruses and genes.

USEPA was created in 1970, but Lederberg failed in attempts to attract funds to follow this up with his research team. The discovery of trihalomethanes and other disinfection byproducts and concerns regarding the potential cancer risks associated with chloroform would be a major driving force behind passage of the 1974 SDWA.

THE INTRODUCTION OF REGULATIONS

In the absence of regulations, many cities adopted practices that were believed to be the most appropriate for their own conditions on the recommendations of professional engineers and water scientists. As noted above, the spread of disease had been controlled in large measure by the quarantine of the sick. It was not unreasonable, therefore, for federal authority over the control of the spread of disease via water to be initially addressed by the U.S. Congress in the Interstate Quarantine Act of 1893 (17).

Under the Act, the surgeon general of the U.S. Public Health Service (USPHS) was empowered "to make and enforce such regulations as are necessary to prevent the introduction, transmission, or spread of communicable disease from foreign countries into the states or possessions, or from one state or possession into any other state or possession."

Interstate regulations were first promulgated in 1894 with the first water-related regulation adopted in 1912, which prohibited the use of the common cup on carriers in interstate commerce (18).

The first federal drinking water standards were adopted in 1914. The USPHS was then part of the U.S. Treasury Department and was charged with responsibility for the health care of the sailors of the Merchant Marine. The surgeon general of the USPHS recommended and the Treasury Department adopted standards for drinking water to be supplied to the public on interstate carriers, then called "Treasury Standards." Because the group that was charged with developing the standards could not agree on physical and chemical parameters, only a bacterial standard of 100 microorganisms per milliliter was adopted. The organism adopted was *Bacteria coli*, now known as *Escherichia coli*. It was further stipulated that not more than one of five 10-mL portions (2 *Bacteria coli* per 100 mL) would be permitted (19). These coliform organisms were not themselves pathogenic but, originating in large numbers in the human colon and found in feces, they served as a surrogate for enteric pathogens because they were more resistant to removal and were present in large numbers and, if they were not present, it could be inferred that the enteric pathogens likely would not be present. Many local and state officials adopted the standard and monitored the water systems that served interstate carriers for themselves and on behalf of the Treasury Department. A federal commitment was made in 1915 to review the regulations on a regular basis.

By 1925, most large cities drawing water from run-of-river sources were already using filtration and chlorination and having little difficulty in meeting the 1914 coliform standard of 2 $100\,\mathrm{mL}^{-1}$ (2 coliforms per milliliter). Following a principle of attainability, the standard was tightened to 1 $100\,\mathrm{mL}^{-1}$. In addition, standards were established for physical and some chemical constituents, including lead, copper, zinc, and dissolved solids (20).

The 1925 standards introduced the concept of relative risk. The preamble stated in part:

> The first step toward the establishment of standards which will ensure safety of water supplies conforming to them is to agree upon some criterion of safety. This is necessary because "safety" in water supplies, as they are actually produced, is relative and quantitative, not absolute. Thus, to state that a water supply is "safe" does not necessarily signify that no risk is ever incurred in drinking it. What is usually meant, and all that can be asserted from any evidence at hand is that the danger, if any, is so small that it cannot be discovered by available means of observation.

In 1941, an advisory committee for revision of the 1925 standards was assembled by the USPHS, composed of representatives of federal and state agencies, scientific associations, and members at large, which produced the 1942 standards (21). One new initiative was the introduction of requirements for monitoring microbial water quality in the distribution system, with specifications for the minimum number of samples to be collected each month according to the size of the community. Specifications for the laboratories and procedures involved were provided.

Maximum permissible concentrations were established for lead, fluoride, arsenic, and selenium as well as for salts of barium, hexavalent chromium, heavy metals, and other substances having deleterious physiological effects. Maximum concentrations where other alternative sources were not available were set for copper, the total of iron and manganese, zinc, chlorides, sulfates, phenolic compounds, total solids, and alkalinity.

Only minor changes were introduced in 1946 (22). Publication in the *Federal Register* was introduced, assuring more rapid dissemination of changes that might be made, one of which was the authorization in March 1957 of the use of the membrane filter procedure for the bacteriological examination of water samples.

World War II (for the United States, 1942–1946) was the first war where deaths of American troops by infectious disease did not exceed deaths in combat. Steps had been introduced to reduce exposure to mosquitoes that were responsible for malaria and other related diseases in the tropics and facilities were provided to assure chlorination of the drinking water. In the postwar period, and driven by the need to make up for years during which the construction of state-side water-related civilian infrastructure had been dormant, attention was turned to making heavy investments for urban water supply.

The need for standards was apparent. Dr. Abel Wolman (23) addressed this issue thus:

> From its beginning, society by one means or another, has surrounded itself with restraints. These have had, for the most part, empiric origins—moral, ethical, economic, or spiritual. All the restraints have had the common basis of an assumed benefit to the particular society establishing them. As societies became more complex and more sophisticated, efforts towards both standardization and restraint became more frequent, more necessary, and presumably more empiric, although examples of the last are not as numerous as one might expect.

He then went on to characterize the types of standards that are necessary:

- Regularization of techniques of measurement;
- Establishment of limits of concentration or density of biologic life and physical and chemical constituents;
- Regularization of administrative practice;
- Regularization of legislative fiat; and
- Specification of materials.

The increasing complexity of the issues is exemplified in all that follows, including not only in the specific regulations required but also in the methodologies of reaching consensus among the many stakeholders involved. The beginning of the "chemical revolution" and regulating the thousands of synthetic organic compounds (SOCs) that are being invented annually and that find their way into the environment and into waters drawn on for drinking began with the 1962 update of the federal Drinking Water Standards.

The establishment of the 1962 USPHS standards involved examining many new issues, including two important problems not previously addressed: radioactivity and SOCs. A new 18-member Advisory Committee was established representing 13 professional and scientific organizations that included consulting engineers, state officials, industry, academics, and water utility executives as well as personnel from the Food and Drug Administration and the U.S. Geological Survey. In addition, 10 officers of the USPHS formed a Technical Subcommittee that, with a six-member Task Force on Toxicology, were advisory to the main Committee (24).

The 1962 USPHS standards were by far the most comprehensive to that date. They included three physical characteristics, odor, color, and turbidity; the last was the most controversial. The turbidity was established at 5 units over the objections of many on the committee from communities that were filtering their waters and who recommended 1 unit, which they could easily meet. Representatives from the northeast, where impounded surface sources were used without filtration, would have had to provide filtration, a measure they believed unnecessary. The bacteriological quality requirement was modified, essentially allowing no more than a monthly average of one coliform per milliliter when the membrane filter technique is used.

The chemical standards were the most difficult to address. Fourteen parameters were listed, but the SOC problem was resolved with the introduction of a Carbon Chloroform Extract (CCE) standard of $0.2\,\mathrm{mg\,L^{-1}}$. A manual was prepared describing the procedure to be used; adsorption of organics by passing a sample of the water through a granular activated carbon (GAC) filter and then desorbing the filter with chloroform (25). The standard was meaningless as a measure of public health risk, because SOCs could not be distinguished from natural organics that are generally of little health consequence, except when they are precursors for chlorination and the creation of trihalomethanes (THMs). But the CCE standard was an attempt to address the SOC problem. The treatment to be provided to remove SOCs was the installation of GAC filters in the treatment train. Forty years later, only a handful of GAC filter plants are being used for treating the most vulnerable public water supplies, those drawing from run-of-river sources. It can be assumed that, at this writing, few supplies that draw from large rivers are removing SOCs that may be present.

The 1962 Standards did introduce two principles that had not been incorporated in previous standards. The first was that "The water supply should be taken from the most desirable source which is feasible, and effort should be made to prevent or control pollution of the source." The second issue was the absence of regulations related to availability of service. A community might be found to be violating the standards if one of the standards is not met but no violation is involved if water service is curtailed because of drought or mechanical failure. The 1962 Standards state "Approval of water supplies shall be dependent in part on: . . . adequate capacity to meet peak demands without development of low pressures or other health hazards."

The 1962 Standards were accepted by all the states, with minor modifications either as regulations or guidelines, but were binding only on about two percent of the communities, those that served interstate carriers (26).

PRELUDE TO THE 1974 SAFE DRINKING WATER ACT

On June 3, 1968, the keynote speaker at the Annual Conference of the AWWA quoted from a report of the Secretary of the U.S. Department of Health Education and Welfare (27): "Fifty million Americans drink water that does not meet Public Health Service drinking water standards. Another 45 million Americans drink water that has not been tested by the Public Health Service."

The AWWA officials were reluctant to publish the paper because it appeared to be too critical of the water supply industry. They acceded only when the author happily agreed to allow rebuttals (28). The task force that prepared the report was not satisfied that the USPHS drinking water standards adequately reflect the health need of the people of the United States. Several issues troubled them.

Little information is available on the health implications of trace substances that may produce disease after exposure over long periods of time. Health experts have repeatedly pointed out that grave, delayed physical manifestations can result from repeated exposure to concentrations of environmental pollutants so small that victims do not report symptoms to a physician.

Furthermore, an individually acceptable amount of water pollution, added to a bearable amount of air pollution, plus nuisances from noise and congestion, can produce a totally unacceptable health environment. It is entirely possible that the biological effects of these environmental hazards, some of which reach individuals slowly and silently over decades or generations, will first begin to reveal themselves after their impact has become irreversible.

In a prescient paper on cancer hazards, Hueper (29) stated:

> It is obvious that with the rapidly increasing urbanization and industrialization of the country and the greatly increased demand on the present resources of water from lakes, rivers, and underground reservoirs, the danger of cancer hazards will grow considerably within the foreseeable future.

Hueper (29) went on to report that studies in Holland revealed that cities drawing water from polluted rivers had higher cancer death rates than those taking water from higher-quality underground sources. At about the same time, the Genetic Study Section of the National Institutes of Health (30) reported that a number of widely used chemicals are known to induce genetic damage in some organisms and that chemicals mutagenic to one species are likely to be mutagenic to others. They believed that when large populations are exposed to highly mutagenic compounds, and they are not demonstrably mutagenic to individuals, the total number of deleterious mutations in the whole population over an extended period of time could be significant.

In 1969, at the beginning of a review of the 1962 standards, the USPHS Bureau of Water Hygiene undertook a comprehensive survey of water supplies in the United States, known as the Community Water Supply Study (CWSS) (31). A total of 969 public water systems, representing about five percent of the total number of systems in the United States serving 18 million people, about 12% of the population being served, were tested (32). About 41% of the systems served did not meet the guidelines in the 1962 Standards. Deficiencies were found in source protection, disinfection, clarification, pressure in the distribution systems, and combinations of these. The small systems, mainly those serving fewer than 500 people, had the greatest difficulty in maintaining water quality. The study revealed that several million people were being supplied with water of inadequate quality and about 360,000 people were being supplied with potentially dangerous drinking water.

The results of the CWSS generated interest in federal legislation that would bring all community water systems under the purview of federal regulations. In 1972, a report of an investigation of the quality of Mississippi River water, as withdrawn from the Carrolton filtration plant in New Orleans, extracted by GAC filtration and a solvent, revealed 36 organic chemicals in the finished water (33). Later, the U.S. General Accounting Office, an agency of the Congress, released a report of the results of an investigation of 446 community water supply systems in six states around the country and found that only 60 of them fully complied with the bacterial and sampling requirements of the 1962 Standards (34). Bacteriological and chemical monitoring were inadequate in five of the states.

In addition to government concern, public organizations and the press had begun to give attention to water supply issues. A three-part series in *Consumer Reports* drew attention to the organic contaminants in New Orleans drinking water (35) Several points were made at the outset of the series that are appropriate today:

> New Orleans, like many other American cities gets its drinking water from a heavily polluted source.... Many industries discharge their wastes into the river and many upriver cities discharge their sewage into it ... runoff from farmland carries a wide variety of pesticides, herbicides, fertilizers, and other agricultural chemicals that swell the Mississippi's pollution burden.
>
> Few New Orleans residents are alarmed. They have been repeatedly assured by city officials that their water, processed according to established water-treatment principles, meets the drinking water standards of the US Public Health Sevice and is "safe." And so it probably is, if one takes "safe" to mean that the water won't cause typhoid, cholera, or other bacterial diseases—the diseases that the standard water treatment is designed to prevent.
>
> In 1969, the Federal Water Pollution Control Administration sampled New Orleans drinking water.... Thirty six (organic compounds) were identified; others were found but could not be identified.
>
> Three of the organic chemicals (chloroform, benzene, and bis-chloroethyl ether) were carcinogens, shown to cause cancer in animal experiments. Three others were toxic, producing liver damage in animals when consumed even in small quantities for long periods. The long-term effects ... are unknown.

The Environmental Defense Fund (EDF) conducted an epidemiologic study in the New Orleans area that compared cancer death rates from communities using lower Mississippi River water as a source with those from nearby communities that were using groundwater sources. The report indicating higher cancer rates among those using the Mississippi River Water was released to the press on November 7, 1974 (36–38). Further publicity followed on December 5, when Dan Rather on CBS aired nationally a program titled "Caution, drinking water may be dangerous to your health." It is interesting to note that upon learning of this situation and the passage of the SDWA, the City of Vicksburg, which had been drawing its water from the Mississippi River, shifted its source to groundwater.

These events, together with the revelation at the time that the chlorine used to make water microbiologically safe would create a family of compounds, trihalomethanes, that were themselves believed to be carcinogenic, led to the passage of the 1974 SDWA.

DRINKING WATER IN DEVELOPING COUNTRIES

The safety of drinking water cannot be examined without considering the problems of drinking water supply and safety in the countries of Asia, Africa, and Latin America. In the industrialized world, attempts are being made to eliminate the use of chlorine for disinfection. Several cities in the Netherlands have abandoned chlorine and other disinfectants entirely because of their concern for DBPs. On the other hand, the situation in the developing world is so serious that the availability of chlorine for every water supply would reduce infant mortality by about 90%.

In 1991, cholera broke out in the Pacific coast of Peru, most probably introduced by maritime traffic from Asia by the discharge of ballast water into the coastal zone from which fish are taken for food, often eaten uncooked. Within two weeks, most of the Peruvian coast, where half of the 22 million Peruvians reside, was host to the disease. Of the some 322,000 cases reported for the year, 55% occurred in the first 12 weeks of the epidemic. The case fatality rate was 0.9% signifying about 30,000 deaths in 1991. By the end of the year, 15 other countries in the Americas, including the United States and Canada, had reported outbreaks caused by the same strain of cholera (39).

Because of its explosive and urban character, contaminated water was identified as the medium for the rapid spread and the intensity of the disease in the cities. Most of the cities had conventional water treatment plants with filtration for water drawn from surface sources. Investigation revealed that chlorination was curtailed and often entirely absent when well water was used. Some Peruvian officials blamed the USEPA for the failure to use chlorine because it had been trumpeting the cancer risks associated with chlorine in water supplies (40).

A serious cholera outbreak occurred in early 2001 in KwaZulu-Natal in the Republic of South Africa, with more than 30,000 cases and some 100 deaths (41). At the height of the outbreak, more than 1000 cases were being reported daily. The reason stated was that the people do not have access to tapwater and are obliged to rely on water from very polluted streams. Even where “bleach” is available, it is not used because it is believed to interfere with fertility. Boiling is not feasible, as firewood is scarce.

In 1980, only 44% of the total population of the developing countries was being served with water by any means, including carrying water of questionable quality long distances from standposts. In urban areas, 69% of the population was being served and very little of that can be considered safe because few cities maintained 24-hour service. When water pressure in distribution pipes is absent, which is most of the day, treated drinking water inevitably becomes contaminated from infiltration of groundwater that is highly contaminated because sewerage systems are absent or in poor condition.

International agencies such as the World Health Organization, the World Bank, the regional development banks, and the developed countries along with the developing countries designated the 1980s the “international drinking water supply and sanitation decade,” during which special efforts were to be made to bring water to the people of the developing world. Ten years later, the population in the developing countries with water supplies had increased to 69%, but the number of people unserved in urban areas had increased by 31 million (42). The rate of urbanization in Asia Africa and Latin America is so great that, even with intensified financial support in grants and loans, the number of urban residents without water service is growing. More important is that those who are counted as having water service do not have safe water by any standard. All that is required to reduce the infant death rate is the type of treatment facilities and their operation and maintenance that was conventional in the industrial world almost a hundred years ago.

Given the nature of world travel today, it is clearly in the self-interest of the industrialized countries to help the developing countries provide water that at least meets 1925 U.S. standards. This would reduce infectious disease that is the major health risk to people and visitors in these countries.

THE FUTURE OF PUBLIC WATER SUPPLY

The history of the monitoring and control of drinking water quality from its earliest days through the present has lessons for those charged with protecting the public health, particularly for those responsible for providing the drinking water to their constituents. This volume demonstrates, if nothing else, that setting standards is a difficult and lengthy procedure. It may be many years, even decades between the time a new risk surfaces and regulations for its control are established and many years more before they are published. Also, years must be allowed for constructing the necessary facilities for eliminating the risks. It behooves the professionals in water utility leadership to educate themselves concerning new risks and prepare to address them before the standards appear in the *Federal Register*. The object is to minimize health risks to the public. Failure, or the perception of failure, drives the public to bottled water with its own risks and costs that are a hardship for a sizable fraction of the population.

A not unrelated issue that is growing in importance as our population ages is the significant percent of the population that is more vulnerable to contaminants by virtue of compromised immune systems. Standards for this population may need to be promulgated.

A similar solution is now being proposed in addressing the quality of water suitable for the potable reuse of wastewaters. Wastewaters contain a large number and a great variety of SOCs. The California Department of Health Services is proposing for the regulation of water quality for groundwater recharge with reclaimed wastewater to potable water aquifers drawn on for drinking water that total organic carbon (TOC) limits be set (43). Again, the carbon compounds may be innocuous or toxic, but in any case a maximum contaminant level (MCL) for TOC of wastewaters is hardly appropriate to assure drinking water safety.

This principle carried over to the 1976 USEPA National Interim Primary Drinking Water Regulations, referred to by this language in Appendix A as "background used in developing the national interim primary drinking water regulations":

> Protection of water that poses no threat to the consumer's health depends on continuous protection. Because of human frailties associated with protection, priority should be given to selection of the purest source. Polluted sources should not be used unless other sources are economically unavailable, and then only when personnel, equipment, and operating procedures can be depended upon to purify and otherwise continually protect the drinking water supply.

This principle is being ignored today, in part because of our faith in treatment technology. Reclaimed wastewater is being proposed as a source for drinking water supplies. Wastewater is hardly likely to be the purest source, and its use for potable reuse is resisted by consumers. Use of reclaimed wastewater for nonpotable purposes is currently being practiced in many hundreds of communities in the United States (44,45), and will be increasingly considered for relieving the pressure on limited high-quality resources.

BIBLIOGRAPHY

1. Okun, D.A. (1996). From cholera to cancer to cryptosporidiosis. *J. Environ. Eng.* **122**: 453–458.
2. Frontinus, S.J. (A.D. 97). *The Two Books on the Water Supply of the City of Rome.*, transl. Clemens Herschel, 1899. Boston: Dana Estes and Company.
3. United Nations. (1958). *Water for Industrial Use*. UN Report E/3058ST/ECA/50. Economic and Social Council, New York.
4. Longmate, N. (1966). *King Cholera: The Biography of a Disease*. Hamish Hamilton, London.
5. Cosgrove, J.J. (1909). *History of Sanitation*. Standard Sanitary Manufacturing Co, Pittsburgh, PA.
6. Snow, J. (1936). *Snow on Cholera*. Oxford Univ. Press, England.
7. Okun, D.A. (1991). Clean water and how to get it. *J. New Engl. Water Works Assoc.* **105**(1): 110.
8. Ives, K.J. (1990). The Chadwick Centenary. *The Life and Times of Sir Edwin Chadwick: 1800–1890*. University College, London.
9. Fair, G.M. (1945). Engineers and engineering in the Massachusetts State Board of Health. *New Engl. J. Med.* **232**: 443–446.
10. Blake, N.M. (1956). *Water for the Cities*. Syracuse University Press, Syracuse, NY.
11. Kober, G.M. (1908). *Conservation of life and health by improved water supply*. *Engineering Record* 57.
12. Ellms, J.W. (1928). *Water Purification*. McGraw-Hill, New York.
13. McCarthy, M.P. (1987). *Typhoid and the Politics of Public Health in 19th-Century Philadelphia*. American Philosophical Society, Philadelphia.
14. Baker, M.N. (1948). *The Quest for Pure Water*. American Water Works Association, Denver.
15. Bellar, T.A., Lichtenberg, J.J., and Kroner, R.C. (1974). The occurrence of organohalides in chlorinated drinking water. *J. Am. Water Works Assoc.* **66**: 703.
16. Lederberg, J. (1969). *We're so accustomed to using chlorine that we tend to overlook its toxicity*. *The Washington Post* May 3, p. A15.
17. United States Statutes. (1893). Interstate Quarantine Act of 1893. *U.S. Statutes at Large*. Chap. 114, Vol. 27, p. 449, Feb. 15.
18. McDermott, J.H. (1973). Federal drinking water standards—past, present, and future. *J. Environ. Eng. Div. Am. Soc. Civil Eng.* **EE4**(99): 469.
19. Borchardt, J.A. and Walton, G. (1971). *Water Quality and Treatment*, 3rd Edn. McGraw-Hill, New York.
20. USPHS. (1925). Report of the Advisory Committee on Official Water Standards. *Public Health Reports* **40**: 693.
21. USPHS. (1943). Public Health Service Drinking Water Standards and Manual of Recommended Water Sanitation Practice. *Public Health Reports* **56**: 69.
22. USPHS. (1946). Public Health Service Drinking Water Standards. *Public Health Reports* **61**: 371.
23. Wolman, A. (1960). Concepts of policy in the formulation of so-called standards of health and safety. *J. Am. Water Works Assoc.* **52**: 11.
24. USPHS. (1962). *Public Health Service Drinking Water Standards*. HEW, Washington, DC.
25. Middleton, F.M., Rosen, A.A., and Burttschell. (1962). Manual for recovery and identification of organic chemicals in water. *J. Am. Water Works Assoc.* **54**: 223–227.
26. Train, R.S. (1974). Facing the real cost of clean water. *J. Am. Water Works Assoc.* **66**: 562.
27. USDHEW (U.S. Department of Health, Education and Welfare). (1967). *A Strategy for a Livable Environment*. HEW, Washington, DC.
28. Okun, D.A. (1969). Alternatives in water supply. *J. Am. Water Works Assoc.* **61**(5): 215–224.
29. Hueper, W.C. (1960). Cancer hazards from natural and artificial water pollutants. *Proc. Conf. Physiol. Aspects Water Quality*. USPHS (U.S. Public Health Service), Washington, DC.
30. National Institutes of Health. Undated. *Report of Chemical Mutagens as a Possible Health Hazard*. Bethesda, MD: NIH Genetics Study Section.
31. USPHS. (1970). *Community Water Supply Study: Analysis of National Survey Findings*. PB214982. National Technical Information Service, Springfield, VA.
32. USPHS. (1970). *Community Water Supply Study: Significance of National Findings*. PB215198/BE. National Technical Information Service, Springfield, VA.
33. USEPA. (1972). *Industrial Pollution of the Lower Mississippi River in Louisiana*. USEPA Region VI, Dallas.
34. Symons, G.E. (1974). That GAO Report. *J. Am. Water Works Assoc.* **66**: 275.
35. Harris, R.H. and Brecher, E.M. (1974). *Is the water safe to drink? Part I. The problem. Part II. How to make it safe.*

Part III. What you can do. Consumer Reports 436 (June), 538 (July), 623 (August).

36. *The States-Item*. (1974). Cancer victims could be reduced—deaths tied to New Orleans water **98**(129)1 (Nov. 7).
37. Page, T., Talbot, E., and Harris, R.H. (1974). *The Implication of Cancer-Causing Substances in Mississippi River Water*. Environmental Defense Fund, Washington, DC.
38. Page, T., Harris, R.H., and Epstein, S.S. (1976). Drinking water and cancer mortality in Louisiana. *Science* **193**: 55.
39. Salazar-Lindo, E. and Alegre, M. (1993). The Peruvian cholera epidemic and the role of chlorination in its control and prevention. In: *Safety of Water Disinfection; Balancing Chemical and Microbial Risks*. G. Craun (Ed.). International Life Sciences Institute, Washington, DC.
40. Anderson, C. (1991). Cholera epidemic traced to risk miscalculation. *Nature* **354**: 255.
41. *Yahoo! Asia—News*. (2001). Cholera infections sky-rocket in South Africa (Feb. 1).
42. Okun, D.A. (1991). A water and sanitation strategy for the developing world. *Environment* **33**(8): 16. Washington, DC: Heldref Publications.
43. California Code of Regulations. (2001). *Title 22. Draft Recycling Criteria*. Sacramento: State Department of Health Services.
44. Okun, D.A. (1997). Distributing reclaimed water through dual systems. *J. Am. Water Works Assoc.* **89**(11): 52–64.
45. Okun, D.A. (2000). Water reclamation and unrestricted nonpotable reuse: A new tool in urban water management. *Annual Rev. Public Health* **21**: 223–245.

1962 U.S. PUBLIC HEALTH SERVICE STANDARDS

From *Drinking Water Regulation and Health*, Wiley 2003

Adoption of the 1962 U.S. Public Health Service standards was completed by publication in the Federal Register, March 6, 1962 (pp. 2152–2155). The full text of these standards is provided below.

PREFACE BY THE ADVISORY COMMITTEE

Domestic water supplies should protect the health and promote the well-being of individuals and the community. In this report on the revision of the 1946 edition of the USPHS drinking water standards, the objective of the committee is to recommend minimum requirements for reaching these goals. The USPHS drinking water standards were first adopted in 1914 to protect the health of the traveling public. The general and widespread use of these standards since that time has led to a series of revisions that have been applicable to water supplies generally. The development of atomic energy and other technological advances require that these standards again be revised. To carry out this revision, the chief sanitary engineer of USPHS appointed the undersigned advisory committee. A technical subcommittee of USPHS officers and a toxicological task force were established to collect information and prepare suggestions for the consideration of the advisory committee.

In preparing this report on the revision of the standards, the committee established the following guidelines:

1. The proposed standards should be discussed widely, and due cognizance should be given to international and other standards of water quality before a final report is submitted.
2. A new section on radioactivity should be added.
3. Greater attention should be given to the chemical substances being encountered increasingly, in both variety and quantity, in water sources.
4. In establishing limits for toxic substances, intake from food and air should be considered.
5. The rationale employed in determining the various limits should be included in an appendix.
6. The proposed format, with the exceptions noted above, should not differ greatly from the present standards.
7. The standards should be generally acceptable and should be applicable to all public water supplies in the United States, as well as to those supplies used by carriers subject to USPHS regulations.
8. The following two types of limits used in previous editions should be continued:
    a. Limits that, if exceeded, shall be grounds for rejection of the supply. Substances in this category may have adverse effects on health when present in concentrations above the limit.
    b. Limits that should not be exceeded whenever more suitable supplies are, or can be made, available at reasonable cost. Substances in this category, when present in concentrations above the limit are either objectionable to an appreciable number of people or exceed the levels required by good water quality control practices.
9. These limits should apply to the water at the free-flowing outlet of the ultimate consumer.

This revision of the drinking water standards includes, for the first time, limiting concentrations of radioactivity in water. The effects on large population groups of chronic exposure to low levels of radioactivity are not yet well defined. The limits presented herein are an effort to derive conservative values from the best information now available and may be adjusted upward or downward as new and better data become available.

The committee has taken cognizance of the growing problem of potentially harmful chemicals in sources of drinking water. Limits for several new chemicals have been added, including a gross limit for the concentration of some types of synthetic chemicals. It was not feasible, however, to include limits for all the many chemicals that have varying degrees of toxic potential. Consideration was given to the more common chlorinated hydrocarbon and organophosphate insecticides, but the information available was not sufficient to establish specific limits for these chemicals. Moreover, the concentrations of these chemicals, where tested, have been below those that would constitute a known health hazard. The committee believes that pollution of water supplies with such contaminants can become significant and urges that the problem be kept under closer surveillance. Further, the committee recommends that regulatory actions be taken to minimize concentrations of such chemicals in drinking water.

In view of the accelerating pace of new developments affecting water quality, the committee recommends that a mechanism be established for continual appraisal and appropriate revision of the standards. It also recommends that USPHS intensify its continuing studies toward the development of basic information on the relationship of the biological, chemical, physical, and radiological aspects of water quality to health.

Membership of the Advisory Committee

U.S. Public Health Service: O. C. Hopkins, Chairman
U.S. Public Health Service: George W. Burke, Jr., Secretary
Federal Food and Drug Administration: L. M. Beacham, Jr.
U.S. Geological Survey: S. K. Love
Air Transport Association of America: K. L. Stratton
American Chemical Society: T. E. Larson
American Dental Association: Robert A. Downs
American Medical Association: W. D. Stovall
American Public Health Association: Daniel A. Okun
American Society of Civil Engineers: Thomas R. Camp
American Water Works Association: Oscar Gullans
Association of American Railroads: R. S. Glynn
Association of State and Territorial Public Health Laboratory Directors: F. R. Hassler
Conference of State Sanitary Engineers: E. C. Jensen
National Committee on Radiation Protection: John B. Hursh
Society of American Bacteriologists: Charles C. Croft
Water Pollution Control Federation: F. W. Gilcreas
Member at Large: Henry J. Ongerth
Technical Subcommittee, Officers of the USPHS
George W. Burke, Jr
M. B. Ettinger
Malcolm Hope
O. C. Hopkins, Cochairman
Paul Kabler
H. G. Magnuson
H. E. Stokinger
Floyd Taylor
James G. Terrill
Richard Woodward, Cochairman
Task Force on Toxicology
Kenneth P. DuBois
Harvey Haag
Wayland J. Hayes, Jr.
Harry Hays
Arnold J. Lehman
H. E. Stokinger

Drinking Water Standards

Definition of Terms. The terms used in these standards are as follows:

Adequate protection by natural means involves one or more of the following processes of nature that produce water consistently meeting the requirements of these standards: dilution, storage, sedimentation, sunlight, aeration, and the associated physical and biological processes that tend to accomplish natural purification in surface waters and, in the case of groundwaters, the natural purification of water by infiltration through soil and percolation through underlying material and storage below the ground water table.

Adequate protection by treatment is anyone or any combination of the controlled processes of coagulation, sedimentation, sorption, filtration, disinfection, or other processes that produce a water consistently meeting the requirements of these standards. This protection also includes processes that are appropriate to the source of supply; works that are located, designed, and constructed to eliminate or prevent pollution; and conscientious operation by well-trained and competent personnel whose qualifications are commensurate with the responsibilities of the position and acceptable to the reporting agency and the certifying authority.

Certifying authority is the surgeon general of USPHS or his duly authorized representatives. (Reference to the certifying authority is applicable only for those water supplies to be certified for use on carriers subject to the USPHS regulations—42 CFR Part 72.)

The coliform group includes all organisms considered in the coliform group as set forth in *Standard Methods for the Examination of Water and Wastewater*, current edition [11th ed., 1960], prepared and published jointly by the American Public Health Association, American Water Works Association, and Water Pollution Control Federation.

Health hazards are conditions, devices, or practices in the water supply system and its operation that create, or may create, a danger to the health and well-being of the water consumer. An example of a health hazard is a structural defect on the water supply system, whether of location, design, or construction, that may regularly or occasionally prevent satisfactory purification of the water supply or cause it to be polluted from extraneous sources.

Pollution, as used in these standards, is defined as the presence of any foreign substance (organic, inorganic, radiological, or biological) in water which tends to degrade its quality so as to constitute a hazard or impair the usefulness of the water.

Reporting agencies are the respective official state health agencies or their designated representatives.

The standard sample for the bacteriological test shall consist of

1. For the bacteriological fermentation tube test, five standard portions of either: (a) 10 mL or (b) 100 mL
2. For the membrane filter technique, not less than 50 mL

Water supply system includes the works and auxiliaries for collection, treatment, storage, and distribution of the water from the sources of supply to the free-flowing outlet of the ultimate consumer.

Source and Protection

The water supply should be obtained from the most desirable source feasible, and effort should be made to prevent or control pollution of the source. If the source is not adequately protected by natural means, the supply shall be adequately protected by treatment.

Frequent sanitary surveys shall be made of the water supply system to locate and identify health hazards that might exist in the system. The manner and frequency of making these surveys, and the rate at which discovered health hazards are to be removed, shall be in accordance with a program approved by the reporting agency and the certifying authority.

Approval of water supplies shall be dependent in part on

a. Enforcement of rules and regulations to prevent development of health hazards
b. Adequate protection of the water quality throughout all parts of the system, as demonstrated by frequent surveys
c. Proper operation of the water supply system under the responsible charge of personnel whose qualifications are acceptable to the reporting agency and the certifying authority
d. Adequate capacity to meet peak demands without development of low pressures or other health hazards
e. Record of laboratory examinations showing consistent compliance with the water quality requirements of these standards

For the purpose of application of these standards, responsibility for the conditions in the water supply system shall be considered to be held by

a. The water purveyor, from the source of supply to the connection to the customer's service piping.
b. The owner of the property served and the municipal, county, or other authority having legal jurisdiction, from the point of connection to the customer's service piping to the free-flowing outlet of the ultimate consumer.

Bacteriological Quality

Sampling. Compliance with the bacteriological requirements of these standards shall be based on examinations of samples collected at representative points throughout the distribution system. The frequency of sampling and the location of sampling points shall be established jointly by the reporting agency and the certifying authority after investigation by either agency, or both, of the source, method of treatment, and protection of the water concerned.

The minimum number of samples to be collected from the distribution system and examined each month should be in accordance with the number in (Fig. 1) for the population served by the system. For the purpose of uniformity and simplicity in application, the number determined from (Fig. 1) should be in accordance with the following: for a population of 25,000 or less, to the nearest 1; 25,001–100,000, to the nearest 5: and more than 100,000 to the nearest 10.

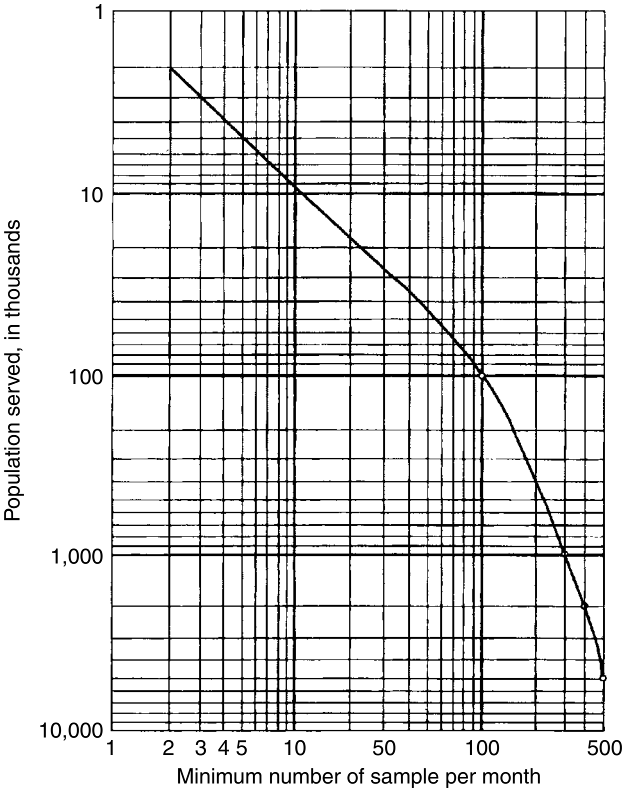

Figure 1. Relation between minimum number of samples to be collected per month and population served.

In determining the number of samples examined monthly, the following samples may be included, provided all results are assembled and available for inspection, and the laboratory methods and technical competence of the laboratory personnel are approved by the reporting agency and the certifying authority:

a. Samples examined by the reporting agency
b. Samples examined by local government laboratories
c. Samples examined by the water works authority
d. Samples examined by commercial laboratories

The laboratories in which these examinations are made and the methods used in making them shall be subject to inspection at any time by the designated representatives of the certifying authority and the reporting agency. Compliance with the specified procedures and the results obtained shall be used as a basis for certification of the supply.

Daily samples collected following a bacteriologically unsatisfactory sample as provided in Secs. 3.21, 3.22, and 3.23 shall be considered as special samples and shall not be included in the total number of samples examined. Neither shall such special samples be used as a basis for prohibiting the supply, provided that

1. When waters of unknown quality are being examined, simultaneous tests are made on multiple portions of a geometric series to determine a definitive coliform content.
2. Immediate and active efforts are made to locate the cause of pollution.
3. Immediate action is taken to eliminate the cause.
4. Samples taken following such remedial action are satisfactory.

Limits. The presence of organisms of the coliform group as indicated by samples examined shall not exceed the following limits:

When 10-mL standard portions are examined, not more than 10% in any month shall show the presence of the coliform group. The presence of the coliform group in three or more 10-mL portions of a standard sample shall not be allowable if this occurs

a. In two consecutive samples.
b. In more than one sample per month when less than twenty are examined per month.
c. In more than 5 per cent of the samples when twenty or more are examined per month.

When organisms of the coliform group occur in three or more of the 10-mL portions of a single standard sample, daily samples from the same sampling point shall be collected promptly and examined until the results obtained from at least two consecutive samples show the water to be of satisfactory quality.

When 100-mL standard portions are examined, not more than 60% in any month shall show the presence of the coliform group. The presence of the coliform group in all five of the 100-mL portions of a standard sample shall not be allowable if this occurs:

a. In two consecutive samples.
b. In more than one sample per month when less than five are examined per month.
c. In more than 20 percent of the samples when five or more are examined per month.

When organisms of the coliform group occur in all five of the 100-mL portions of a single standard sample, daily samples from the same sampling point shall be collected promptly and examined until the results obtained from at least two consecutive samples show the water to be of satisfactory quality.

When the membrane filter technique is used, the arithmetic mean coliform density of 0 standard samples examined per month shall not exceed 1 per 100 mL. Coliform colonies per standard sample shall not exceed 3 per 50 ml, 4 per 100 mL, 7 per 200 mL, or 13 per 500 mL in

a. Two consecutive samples.
b. More than one standard sample when less than 20 are examined per month.
c. More than 5% of the standard samples when 20 or more are examined per month.

When coliform colonies in a single standard sample exceed the above values, daily samples from the same sampling point shall be collected promptly and examined until the results obtained from at least two consecutive samples show the water to be of satisfactory quality.

Physical Characteristics

Sampling. The frequency and manner of sampling shall be determined by the reporting agency and the certifying authority. Under normal circumstances, samples should be collected one or more times per week from representative points in the distribution system and examined for turbidity, color, threshold odor, and taste.

Limits. Drinking water should contain no impurity that would cause offense to the sense of sight, taste, or smell. Under general use, the following limits should not be exceeded: turbidity, 5 units; color, 15 units; and threshold odor number, 3 units.

Chemical Characteristics

Sampling. The frequency and manner of sampling shall be determined by the reporting agency and the certifying authority. Under normal circumstances, analyses for substances listed below need be made only semiannually. If, however, there is some presumption of unfitness because of the presence of undesirable elements, compounds, or materials, periodic determinations for the suspected toxicant or material should be made more frequently, and an exhaustive sanitary survey should be made to determine the source of the pollution. Where the concentration of a substance is not expected to increase in processing and distribution, available and acceptable source water analyses performed in accordance with standard methods may be used as evidence of compliance with these standards.

Where experience, examination, and available evidence indicate that particular substances are consistently absent from a water supply or below levels of concern, semiannual examinations for those substances may be omitted when approved by the reporting agency and the certifying authority.

The burden of analysis may be reduced in many cases by using data from acceptable sources. Judgment concerning the quality of water supply and the need for performing specific local analyses may depend in part on information produced by such agencies as (1) USGS, which determines chemical quality of surface water and groundwater of the United States and publishes these data in *Water Supply Papers* and other reports, and (2) USPHS, which determines water quality related to pollution (or the absence of pollution) in the principal rivers of the United States and publishes these data annually in *National Water Quality Network*. Data on pollution of waters as measured by carbon chloroform extracts (CCES) may be found in the latter publication.

Limits. Drinking water shall not contain impurities in concentrations that may be hazardous to the health of the consumers. It should not be excessively corrosive to the water supply system. Substances used in its treatment shall not remain in the water in concentrations greater than required by good practice. Substances that may have deleterious physiological effect, or substances for which physiological effects are not known, shall not be introduced into the system in a manner that would permit them to reach the consumer.

The chemical substances shown in Table 1 should not be present in a water supply in excess of the listed concentrations where, in the judgment of the reporting agency and the certifying authority, other more suitable supplies are or can be made available.

The presence of substances in excess of the concentrations listed in Table 2 shall constitute grounds for rejection of the supply.

Fluoride. When fluoride is naturally present in drinking water, the concentration should not average more than the appropriate upper limit shown in Table 3. Presence of fluoride in average concentrations greater than two times the optimum values in Table 3 shall constitute grounds for rejection of the supply. Where fluoridation (supplementation of fluoride in drinking water) is practiced, the average fluoride concentration shall be kept within the upper and lower control limits shown in Table 3. In addition to the sampling required by Sec. 5.1, fluoridated and defluoridated supplies shall be sampled with sufficient frequency to determine that the desired fluoride concentration is maintained.

Table 1. A Water Supply with Concentrations Exceeding These Limits Should Be Avoided If Possible[a]

Substance	Concentration (mg/L)
Alkyl benzene sulfonate (ABS)	0.5
Arsenic (As)	0.01
Chloride (Cl)	250.0
Copper (Cu)	1.0
Carbon Chloroform extract (CCE)	0.2
Cyanide (CN)	0.01
Fluoride (F)	—[b]
Iron (Fe)	0.3
Manganese (Mn)	0.05
Nitrate (NO_3)[c]	45.0
Phenols	0.001
Sulfate (SO_4)	250.0
Zinc (Zn)	5.0

[a] Refer to Section 5.2.1 (in this Appendix).
[b] See Section 5.2.3.
[c] In areas in which the nitrate content of water is known to be in excess of the listed concentration, the public should be warned of the potential dangers of using the water for infant feeding.

Table 2. A Water Supply with Concentrations Exceeding These Limits Should Be Rejected[a]

Substance	Concentration (mg/L)
Arsenic (As)	0.05
Barium (Ba)	1.0
Cadmium (Cd)	0.01
Chromium (Cr^{6+})	0.05
Cyanide (CN)	0.2
Fluoride (F)	—[b]
Lead (Pb)	0.05
Selenium (Se)	0.01
Silver (Ag)	0.05

[a] Refer to Section 5.2.2.
[b] See Section 5.2.3.

Radioactivity

Sampling. The frequency of sampling and analysis for radioactivity shall be determined by the reporting agency and the certifying authority after consideration of the likelihood of significant amounts being present. Where concentrations of ^{226}Ra or ^{90}Sr may vary considerably, quarterly samples composited over a period of 3 months are recommended. Samples for determination of gross activity should be taken and analyzed more frequently. As indicated in Sec. 5.1, data from acceptable sources may be used to indicate compliance with these requirements.

Limits. The effects of human radiation exposure are viewed as harmful and any unnecessary exposure to ionizing radiation should be avoided. Approval of water supplies containing radioactive materials shall be based on the judgment that the radioactivity intake from such water supplies when added to that from all other sources will not result in an intake greater than the radiation protection guidance recommended by the Federal Radiation Council and approved by the President. Water supplies shall be approved without further consideration of other sources of radioactivity intake of ^{226}Ra and ^{90}Sr when the water contains these substances in amounts not exceeding 3 and 10 $\mu\mu$Ci/L, respectively. When these concentrations are exceeded, a water supply shall be approved by the certifying authority if surveillance of total intakes of radioactivity from all sources indicates that such intakes are within the limits recommended by the Federal Radiation Council for control action.

The Federal Radiation Council, in its Memorandum for the President, Sept. 13, 1961, recommended that "routine control of useful applications of radiation and atomic energy should be such that expected average exposures of suitable samples of an exposed population group will not exceed the upper value of Range II (20 $\mu\mu$Ci/day of ^{226}Ra and 200 $\mu\mu$Ci/day of ^{90}Sr).

Table 3. Recommended Fluoride Control Limits[a]

Annual Average of Maximum Daily Air Temperature[b], °F	Recommended Control Limits for Fluoride Concentration, mg/L		
	Lower	Optimum	Upper
50.0–53.7	0.9	1.2	1.7
53.8–58.3	0.8	1.1	1.5
58.4–63.8	0.8	1.0	1.3
63.9–70.6	0.7	0.9	1.2
70.7–79.2	0.7	0.8	1.0
79.3–90.5	0.6	0.7	0.8

[a] Refer to Section 5.2.3.
[b] Based on temperature data obtained for a minimum of 5 years.

In the known absence[1] of ^{90}Sr and alpha emitters, the water supply is acceptable when the gross beta concentrations do not exceed 1000 $\mu\mu$Ci/L. Gross beta concentrations in excess of 1000 $\mu\mu$Ci/L shall be grounds for rejection of supply except when more complete analyses indicate that concentrations of nuclides are not likely to cause exposures greater than the Radiation Protection Guides as approved by the President on recommendation of the Federal Radiation Council.

Recommended Analytical Methods

Analytical methods to determine compliance with the requirements of these standards shall be those specified in *Standard Methods for the Examination of Water and Wastewater*, APHA, AWWA, and WPCF, New York, current edition [11th ed., 1960], and those specified as follows:

Barium. Rainwater, F. H., and L. L. Thatcher: *Methods for the Collection and Analysis of Water Samples*, USGS, Water Supply Papers, 1454, Govt. Printing Office, Washington, DC.

Carbon Chloroform Extract (CCE). *Method for Determining the Carbon Chloroform Extract (CCE) in Drinking Water*, R. A. Taft San. Eng. Center, USPHS, Cincinnati (1961).

Radioactivity. *Laboratory Manual of Methodology; Radionuclide Analysis of Environmental Samples*, Tech. Rept. R59-6, R. A. Taft San. Eng. Center, USPHS, Cincinnati; and Methods of Radiochemical Analysis, Tech. Rept. 173, Joint WHO-FAO Committee, World Health Organization (1959).

Selenium. Magin, C. B., et al.: Suggested Modified Method for Colorimetric Determination of Selenium in Natural Water, J. Am. Water Workd Assoc. 52:119 (Sept. 1960).

Organisms of the Coliform Group. All of the details of techniques in the determination of bacteria of this group, including the selection and preparation of apparatus and media, the collection and handling of samples, and the intervals and conditions of storage allowable between collection and examination of the water sample, shall be in accordance with *Standard Methods for the Examination of Water and Wastewater*, current edition, and the procedures shall be those specified therein for

1. Membrane Filter Technique, standard test, or
2. Completed Test, or
3. Confirmed Test, procedure with brilliant green lactose bile broth.
4. Confirmed Test, procedure with Endo or eosin methylene blue agar plates.[2]

---

[1] Absence is taken here to mean a negligibly small fraction of the above specific limits where the limit for the unidentified alpha emitters is taken as the listed limit for ^{226}Ra.

[2] The Confirmed Test is allowed, provided the value of this test to determine the sanitary quality of the specific water supply being examined is established beyond reasonable doubt by comparisons with Completed Tests performed on the same water supply.

ION EXCHANGE AND DEMINERALIZATION

National Drinking Water
Clearinghouse

Ion exchange and membrane processes are becoming used extensively in water and wastewater treatment. Ion exchange is primarily used for the removal of hardness ions, such as magnesium and calcium, and for water demineralization. Reverse osmosis (RO) and electrodialysis, both membrane processes, remove dissolved solids from water using membranes.

ION EXCHANGE AND DEMINERALIZATION ARE BECOMING WIDELY USED

Ion exchange units can be used to remove any charged (ionic) substance from water, but are usually used to remove hardness and nitrate from groundwater.

Water is pretreated to reduce the suspended solids and total dissolved solids (TDS) load to the ion-exchange unit. Methods of pretreatment include:

- filtration,
- coagulation and filtration,
- cold lime with or without soda ash,
- hot lime with or without soda ash,
- evaporation or distillation,
- electrodialysis,
- RO,
- continuous deionization,
- ultrafiltration,
- degasification, or
- combinations of the above.

(Source: Owens, 1995)

RO systems are compact, simple to operate, and require minimal labor, making them suitable for small systems. They are also suitable for systems where there is a high degree of seasonal fluctuation in water demand.

Electrodialysis is a process that also uses membranes. However, in electrodialysis, direct electrical current is used to attract ions to one side of the treatment chamber. Electrodialysis systems include a source of pressurized water, a direct current power supply, and a pair of selective membranes.

ION EXCHANGE

Ion exchange effectively removes more than 90 percent of barium, cadmium, chromium (III), silver, radium, nitrites, selenium, arsenic (V), chromium (VI), and nitrate. Ion exchange is usually the best choice for small systems that need to remove radionuclides.

Advantages

- Ion exchange process, like reverse osmosis, can be used with fluctuating flow rates.

- Effluent contamination is virtually impossible.
- Large variety of specific resins are available from suppliers. Each resin is effective in removing specific contaminants.

Limitations

- Ion exchange waste is highly concentrated and requires careful disposal.
- Potential for unacceptable levels (peaks) of contamination in effluent.
- Usually not feasible with high levels of TDS.
- Pretreatment required for most surface waters.
- Ion exchange units also are sensitive to the presence of competing ions. For example, influent with high levels of hardness will compete with other cations (positive ions) for space on the exchange medium, and the exchange medium must be regenerated more frequently.

Process

Inorganics removal is accomplished through adsorption of contaminant ions onto a resin exchange medium. As the name implies, one ion is substituted for another on the charged surface of the medium, which is usually a synthetic plastic resin. This resin surface is designed as either cationic or anionic (negatively charged). The exchange medium is saturated with the exchangeable ion before treatment operations.

During ion exchange, the contaminant ions replace the regenerant ions because they are preferred by the exchange medium. When there are no ions left to replace the contaminant ions, the medium is regenerated with a suitable solution, which resaturates the medium with the appropriate ions. Because of the required "down time," the shortest economical regeneration cycles are once per day.

The resin exchange capacity is expressed in terms of weight per unit volume of the resin. The calculation of the breakthrough time for an ion exchange unit requires knowledge of the resin exchange capacity, the influent contaminant concentration, and the desired effluent quality.

Equipment

Typical ion exchange units consist of prefiltration, ion exchange, disinfection, storage, and distribution elements (see Fig. 1).

Chemicals

Sodium chloride is often used to regenerate the exchange medium in ion exchangers because of the low cost of the chemical. However, this can result in a high sodium residual in the finished water, which may be unacceptable for individuals with salt restricted diets. This problem can be avoided by using other regenerant materials, such as potassium chloride.

REVERSE OSMOSIS (RO)

RO can effectively remove nearly all inorganic contaminants from water. It removes more than 70 percent of arsenic (III), arsenic (IV), barium, cadmium, chromium (III), chromium (VI), fluoride, lead, mercury, nitrite, selenium (IV), selenium (VI), and silver. Properly operated units will attain 96 percent removal rates. RO can also effectively remove radium, natural organic substances, pesticides, and microbiological contaminants. RO is particularly effective when used in series. Water passing through multiple units can achieve near zero effluent contaminant concentrations.

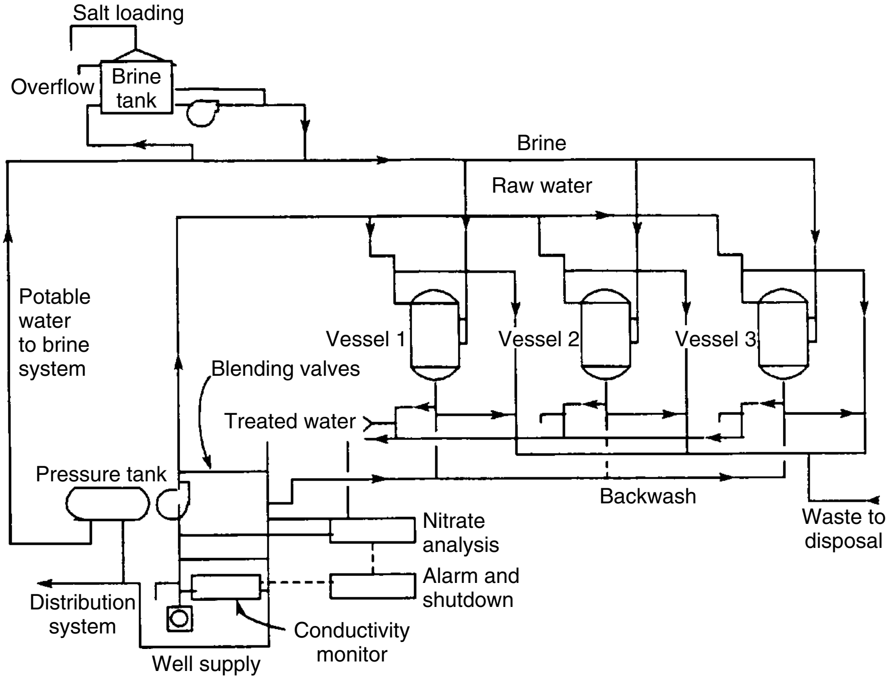

Figure 1. Ion exchange treatment system. *Source*: U.S. Environmental Protection Agency, 1989.

Advantages

- Removes nearly all contaminant ions and most dissolved non-ions.
- Relatively insensitive to flow and TDS level, and thus suitable for small systems with a high degree of seasonal fluctuation in water demand.
- RO operates immediately, without any minimum break-in period.
- Low effluent concentration possible.
- Bacteria and particles are also removed.
- Operational simplicity and automation allow for less operator attention and make RO suitable for small system applications.

Limitations

- High capital and operating costs.
- Managing the wastewater (brine solution) is a potential problem.
- High level of pretreatment is required in some cases.
- Membranes are prone to fouling.

Process

RO removes contaminants from water using a semipermeable membrane that permits only water, and not dissolved ions (such as sodium and chloride), to pass through its pores. Contaminated water is subject to a high pressure that forces pure water through the membrane, leaving contaminants behind in a brine solution. Membranes are available with a variety of pore sizes and characteristics.

Equipment

Typical RO units include raw water pumps, pretreatment, membranes, disinfection, storage, and distribution elements (see Fig. 2). These units are able to process virtually any desired quantity or quality of water by configuring units sequentially to reprocess waste brine from the earlier stages of the process. The principal design considerations for reverse osmosis units are:

- operating pressure,
- membrane type and pore size,
- pretreatment requirements, and

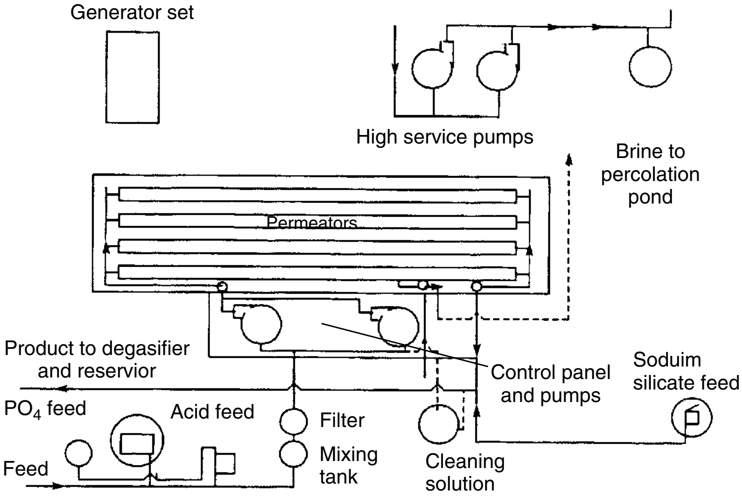

Figure 2. Schematic of a reverse osmosis system. *Source: U.S. Environmental Protection Agency, 1989.*

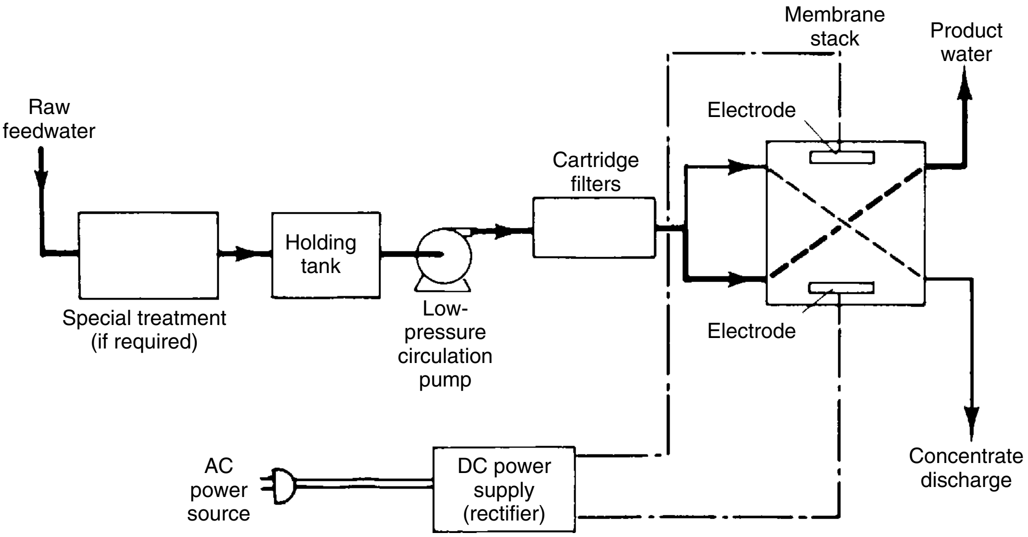

Figure 3. Basic components of an electrodialysis unit. *Source: U.S. Agency for International Development, 1980.*

- product conversion rate (the ratio of the influent recovered as waste brine water to the finished water).

Electrodialysis

Electodialysis is very effective in removing fluoride and nitrate, and can also remove barium, cadmium, and selenium (see Fig. 3).

Advantages

- All contaminant ions and most dissolved non-ions are removed.
- Relatively insensitive to flow and TDS level.
- Low effluent concentration possible.

Limitations

- High capital and operating costs.
- High level of pretreatment required.
- Reject stream is 20–90 percent of feed flow.
- Electrodes require replacement.

Process

The membranes adjacent to the influent stream are charged either positively or negatively, and this charge attracts counter-ions toward the membrane. The membranes are designed to allow either positively or negatively charged ions to pass through the membrane, thus ions

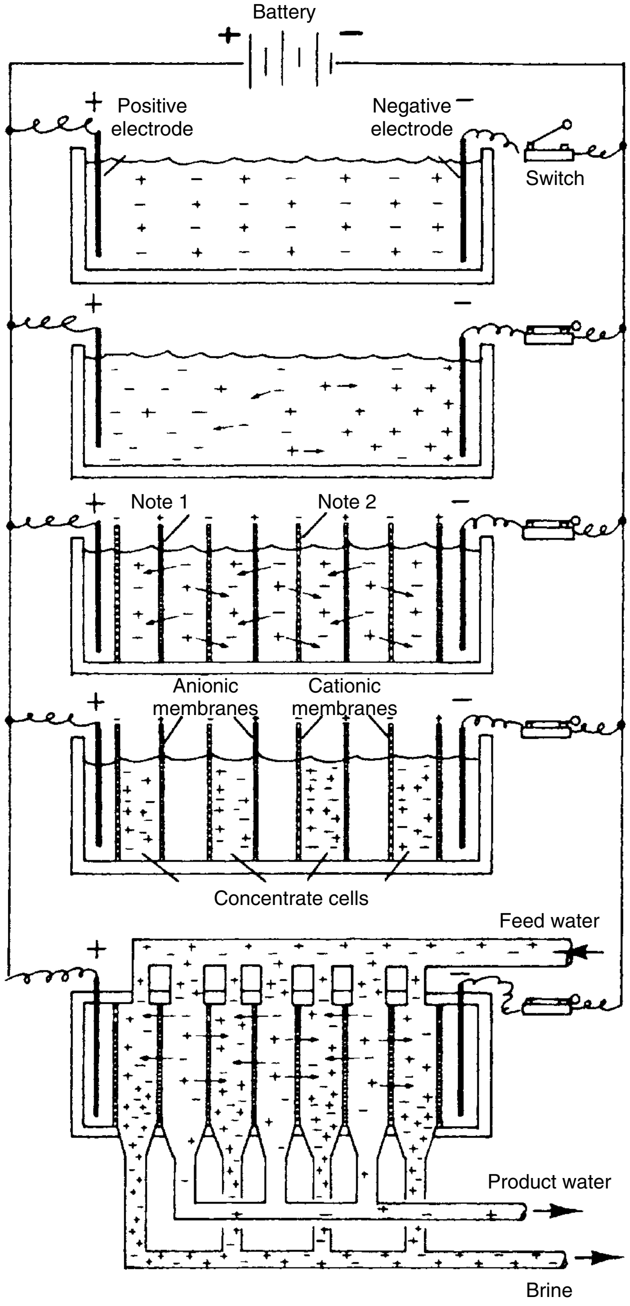

Many of the substances which make up the total dissolved solids (TDS) in brackish water are strong electrolytes. When dissolved in water they ionize; that is, the compounds dissociate into ions which carry an electric charge. Typical of the ions in brackish water are Cl^{-1}, Na^{+1}, HCO_3^{-1}, Mg^{+2},$SO4^{-2}$, and Ca^{+2}. These ions tend to attract the dipolar water molecules and to be diffused in times, fairly evenly throughout a solution.

If two electrodes are placed in a solution of ions and energized by a battery orother direct current source, the current is carried through the solution by the charged particles and the ions tend to migrate to the electrode of the opposite charge.

If alternately fixed charged membranes (which are selectively permeable to ions of the opposite charge) are placed in the path of the migrating ions, the ions will be trapped between the alternate cells formed.

Note 1: A positively fixed charge (anionic) membrane will allow negative ions to pass, but will repel positive ions.

Note 2: A negatively fixed charge (cationic) membrane will allow positive ions to pass, but will repel negative ions.

If this continued, almost all the ions wiould become trapped in the alternate cells (concentrate cells). The other cells, which lack ions, would have a lower level of dissolved constituents and would have a high resistance to current flows.

The phenomenon illustrated above is used in electrodialysis to remove ions from incoming saline water on a continuous basis. Feedwater enters both the concentrate and product cells. Up to about half of the ions in the product cells migrate and are trapped in the concentrate cells. Two streams emerge form the device: One of concentrated brine and the other with a much lower concentration of TDS (product water).

Figure 4. Movement of ions in the electrodialysis process. *Source: U.S. Agency for International Development, 1980.*

move from the product water stream through a membrane to the two reject water streams (see Fig. 4).

Equipment

The three essential elements of the system are (1) a source of pressurized water, (2) a direct current power supply, and (3) a pair of selective membranes. The average ion removal varies from 25 to 60 percent per stage. Multistage units can increase the efficiency of removal. Many membrane pairs are "stacked" in the treatment vessel.

Chemicals

Fouling of membranes may limit the amount of water treated. Fouling is caused when membrane pores are clogged by salt precipitation or by physical obstruction of suspended particulates. Particulates, suspended in water, can be removed in pretreatment but salts that exceed their solubility product at the membrane surface must be controlled chemically by pH reduction (to reduce carbonate concentration) or chelation of metal ions (by use of phosphate, for example). A reversal of the charge on the membranes, a process called electrodialysis reversal (EDR), helps to flush the attached ions from the membrane surface, thus extending the time between cleanings.

Where Can I Find More Information?

Information on ion exchange/demineralization was primarily obtained from two sources: Environmental Pollution Control Alternatives: Drinking Water Treatment for Small Communities, *EPA/625/5-90/025; and* Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities, *EPA/625/4-89/023. Both may be ordered free from the U.S. Environmental Protection Agency Office of Research and Development at (513) 569-7562.*

These publications also may be ordered from the National Drinking Water Clearinghouse (NDWC); however, copying costs apply. The first book, item #DWBKGN09, an 82-page publication, costs $11.82; and the second, item #DWBKDM04, a 209-page book, costs $30.05. Shipping and handling charges also apply.

For further information or to order copies of "Tech Brief: Ion Exchange and Demineralization," item #DWBLPE56, "Tech Brief: Filtration," item #DWBLPE50, "Tech Brief: Disinfection," item #DWBLPE47, or "Tech Brief: Corrosion Control," item #DWBLPE52 call the NDWC at (800) 624-8301 or (304) 293-4191. These fact sheets are available at no cost, except for shipping and handling charges.

THE STATE OF THE WATER INDUSTRY—2004

STEVE MAXWELL
TechKNOWLEDGEy Strategic Group
Boulder, Colorado

The water and wastewater treatment industry continues to be an evolving and often unpredictable business, and the year 2003 did not disappoint. The ongoing corporate consolidation in the industry produced surprises and new directions, whereas the headlong rush toward privatization seemed to falter, or at least slow a bit. Water resource issues and water management controversies continued to climb onto the front pages of the popular media; water quality scares, water rights battles, and water security issues gained significant public exposure. Wide-ranging debate over how to best fund the vast future capital needs of the water business began to take center stage. The extensive regional droughts of 2002 are still fresh in the minds of many people. And, although water stocks outperformed the general stock market this year, there still seem to be no big winners emerging. Below, we highlight the key industry developments of the past year, review the current characteristics of the market, and discuss the evolving drivers behind this business.

OVERVIEW

Studies and reports announcing the certainty and severity of future water challenges continue to be issued from the federal government, the various water-related trade associations, and environmental think tanks around the country. Earlier in the year, and to much fanfare, the American Council of Engineering Companies (ACEC) gave the water infrastructure system of the country a general grade of "D," and—supplementing numerous prior studies—the Congressional Budget Office predicted that we will need to spend more than $800 billion over the next 20 years to fix this problem. The increasing focus of concern among water industry and municipal officials is how to reconcile the difference between *current* spending rates and projected *future* needs—how to finance the "gap" that is becoming increasingly clear. As one observer put it, sometimes it seems that the fluid of most concern in this industry is not water, but red ink.

Although the trend to greater consolidation continued, 2003 may be remembered as the year in which the much-ballyhooed "foreign invasion" of recent years began to reverse directions. After practically tripping over each other in the rush to acquire assets during the late 1990s, many major European water companies began to shed major U.S. water businesses during the year. Although perhaps not totally unexpected, these divestitures constituted a major shift in the competitive dynamics and transactional trends in the business. In sum, several major players that have long been viewed as the prominent buyers in the industry, suddenly appeared to turn (at least partially) into sellers.

The reasons behind this shift are varied and complex (see discussion below), but it seems clear that a large-scale game of "musical chairs" is starting to happen in this industry. A considerable rearrangement of key assets

---

The information contained in this report has been obtained from sources believed to be reliable: however, their accuracy or completeness cannot be guaranteed. The opinions expressed herein are solely those of *The Environmental Benchmarker and Strategist*. Neither this report nor any opinion expressed herein should be construed as a solicitation to sell or acquire any of the securities or other investments mentioned herein.

is already starting to happen—in general, away from foreign companies and back toward domestic ownership. And, unfortunately for these foreign companies who were buying several years ago, average valuations in the industry are now lower (see Table 1), and the dollar has slid significantly against the Euro during the past couple of years.

Another longer term trend in the market—greater municipal outsourcing and private contract operation—also seemed to begin to shift direction or become less clear this year. The pace of privatization has been rapid during the past decade or so, and many analysts believe it was the perceived privatization opportunity that effectively drove much of the acquisition frenzy of the late 1990s. The last few years, however, have seen a gradual slowdown in this trend and an increasingly widespread reconsideration of the benefits of privatization of public water treatment operations. Starting with the much-publicized cancellation of Atlanta's water privatization plan and the postponement of programs in cities such as New Orleans and Stockton, the whole concept of water privatization has taken a major hit in public opinion circles during the past couple of years. For the first time in several years, the volume of outsourcing contracts in the country was actually down, and fairly sharply down, in 2003.

Nonetheless, the business of private contract operation is still projected to be one of the fastest growing market niches in the whole water industry. The forces behind privatization remain strong primarily because of the difficult position of municipalities. Public works managers are between a rock and a hard place—costs, technical requirements, and regulatory complexities continue to increase, but the general public remains resistant to increasing taxes and user's fees. As the contract operations firms point out, often the best solution to this dilemma

Table 1. 40 Publicly Traded Companies with Interests in the Water Industry

Company	Symbol	Revenues $ mils.	Income $ mils.	Market Cap. $ mils.
American States	awr	213	17.2	388.36
Ameron	amn	576	27.7	316.224
Aqua America	wtr	367	71	1940.4
Artesian Resources	arina	36	4.2	110.175
Badger Meter	bmi	184	7.6	123.222
Calgon Carbon	ccc	278	4.5	272.22
California Water	cwt	277	19.3	473.2
Clarcor	clc	741	55	1068.925
Clorox	clx	4140	486	10478.26
Cuno	cuno	288	26.8	688.708
Danaher	dhr	5290	537	14579.18
Dionex	dnex	232	34.6	1158.792
FlowServe	fls	2370	49.2	1103.448
Fluor	flr	8810	179	3348.06
Franklin Electric	fele	360	34.5	652.365
Glacier Water	hoo	71.2	−1.5	42.21
Great Lakes Chemical	glk	1470	−33	1277.144
Insituform Tech.	insu	491	23.5	424
Ionics	ion	349	−19.6	507.656
Isco	isko	62.5	1.46	48.8196
ITT Industries	itt	5630	391	7088.6
Layne Christenson	layn	278	2	164.3
Lindsay Mfg.	lnn	166	12.8	292.994
Metpro	mpr	74	6.4	141.681
Middlesex Water	msex	63.8	7.2	216.452
Millipore	mil	800	101	2434.144
Pall	pll	1660	151	3311.28
Pentair	pnr	2720	144	2637.96
Robbins & Myers	rbn	568	14.5	286.375
SJW Corporation	sjw	150	18.7	292.89
Southwest Water	swwc	164	6.1	205.506
Suez SA	sze	55850	−3360	21906.9
Tetra Tech	ttek	922	54	1188.096
Thermo Electron	trno	2100	173	4619.42
URS	urs	3190	58	978.51
Valmont Industries	vmi	826	26	522.41
Vermont Pure Hldgs.	vps	76	1.35	72.846
Veolia Environne.	ve	36740	−2490	11736.9
Waterpik	pik	305	11	156.86
Watts Water Tech.	wts	676	34	623.76

may be to turn to private companies to finance, build, and operate their water or wastewater systems.

From the perspective of Wall Street, the water industry was relatively stronger this year, and most water funds and indices outperformed in generally stronger markets. But even though performance was stronger, we still have not seen any really "big winners" in the industry—stocks that have returned investors many times their original investment. However, the professional stock-pickers assure us that there *will* be such companies emerging in the near future (Table 1).

An emerging issue in the water industry—one that is likely to loom considerably larger in the future—is the appropriate balance of the federal government and local authorities in the financing and management of water resources in this country. Groups like AMSA (the Association of Metropolitan Sewerage Agencies) claim that the federal government is sidestepping the whole issue of clean water. In an aggressive public information campaign that the group began to mount in 2003, they pointed out that in 1980, over half of the total capital investment in clean water came from federal funds, whereas today that figure is less than 10%.

Certainly, watershed management and water quality are interstate issues—water quality and management issues rarely follow state boundaries. Another lobbying group, the Water Infrastructure Network—a coalition of various elected officials, water utilities, and water resource groups—points out that clean water supports $50 billion per year in water-based recreation, $300 billion per year of coastal tourism, $45 billion in commercial fishing, and of course, hundreds of billions in basic manufacturing and industry. Despite the obvious criticality of water to the overall economy, so far there are few indications that the federal government plans to address the complex funding issues facing the business.

And other new challenges continue to arise and bedevil the U.S. water industry. Following the events of 9/11 and the discovery of potential terrorist interest in public drinking water supplies, security quickly moved to the forefront of immediate challenges facing the industry. Although there has been a lot of discussion about security threats, most of the work remains to be done, and most of the dollars still need to be spent—one more increment to the total cost of providing the public with acceptable drinking water.

A related problem also received new visibility this year—the issue of backup electrical power in the water system. During the massive northeastern blackout of August 14, several major cities that rely primarily on pumping for water distribution (as opposed to cities like New York with largely gravity-fed systems) were without sufficient water for several hours. An outbreak of fire during that particular day, for example, in downtown Cleveland or Detroit, could have been catastrophic. Installation of sufficient backup power for pumping and distribution has quickly become a high priority at many water systems across the country.

Technology development and implementation continues to be an active area, attracting many new players, and holding out hope for future solutions. The full-scale commercialization of membrane filtration is a good example—more and more water and wastewater utilities are now examining this technology, as its costs have come down dramatically in recent years. Although the much-discussed Tampa desalination plant suffered extensive operating problems this year, including the bankruptcy filing of its operator, the market for this particular technology—particularly in desalination and water reuse applications—remains strong. Other areas include *in situ* repair and relining of aging distribution pipelines; "smart" monitoring, metering, and system control software; and a wide range of other new treatment, distribution, and conservation technologies.

In response to these changing social, political, and economic drivers, the water industry continues to undergo gradual evolution and coalescence. With such dramatic changes occurring in technology, asset ownership, regulation, and public concern, the water industry seems certain to remain in a constant state of change for the foreseeable future.

Another way to gauge the impact of these changes on the commercial industry is to look back at our list of public companies in *The Environmental Benchmarker and Strategist* from just 5 short years ago, in late 1998. The experiences of these companies—most of which have now disappeared from the ranks of the public companies—illustrates some general trends in the industry since that time:

- Air and Water Technologies—sold by Vivendi in 2000
- American Water Works—purchased by Thames Water/RWE in 2001
- Aquarion—purchased in 1999 by Yorkshire Water (now Kelda)
- Azurix—Enron's heralded entry into the water industry; dead three years later
- BetzDearborn—purchased by Hercules, later sold to General Electric in 2002
- Cadiz—company virtually disappeared after major water storage program was nixed by the state of California in 2002
- Culligan—acquired by U.S. Filter in 1998, and now for sale again
- E-One—sold to Precision Castparts in the late 1990s
- Hach—acquired by Danaher Corporation in 1998
- Nalco—bought by Suez in 1999 and sold to private investors in 2003
- Osmonics—acquired by General Electric in 2003
- Recovery Engineering—acquired by Procter and Gamble in 1998
- Stone and Webster—filed for bankruptcy, assets acquired by Shaw Group
- Thermo Instrument Systems—collapsed back into parent Thermo Electron
- US Filter—acquired by Vivendi, spun into Veolia subsidiary, most of its noncontract services businesses now for sale again

- U.S. Liquids—suffered through extreme financial difficulties, selling off assets, and now trading at a few cents per share
- United Water—remaining traded stock acquired by Suez in 2001
- Western Water—like Cadiz, has shrunk down considerably, perhaps a company before its time in water rights trading
- Zurn—company was sold to U.S. Industries in 1998

MARKET SIZE AND GROWTH CHARACTERISTICS

According to recent Department of Commerce statistics, the water and environmental business in this country comprises more than $200 billion of revenues per year, encompasses some 115,000 companies and organizations, and employs more than 1.4 million workers. More specifically, the water and wastewater industry is generally estimated at around $90 to $100 billion per year; the comparable world market is about five times as large, or around $500 billion.

Estimates of the true size of this business vary, largely due to one key reason—it is a tough business to define and "encircle"—definitions differ, and even where they do not differ, market estimates still vary widely. Perhaps the most thoroughly researched and widely cited numbers on the overall domestic water industry are those published annually by Environmental Business International, as shown in Table 2.

Because it is such a broad and varied business, it is not meaningful to try to capture the overall industry's growth rate in a single figure either; different sectors of the business are growing at considerably different rates. For example, the more mature chemicals sector and infrastructural supply businesses (pumps, pipes, tanking, and so on) are probably growing in the 2% to 3% range. The water and wastewater utility sectors, which comprise the largest individual source of revenues in the industry, largely reflect population and GNP growth and show growth in the 3% to 4% range. Other businesses, tied to the upgrading and replacement of our vast water infrastructure, such as consulting engineering and instrumentation/monitoring, show slightly higher growth in the 6% to 8% range.

On the other hand, other sectors of the business boast growth well above these rates. The contract operations and outsourcing sector of the business has demonstrated growth in the 15% to 20% range for several years and even given the recent slowdown is still likely to average something in the 8% to 10% range. And although the overall treatment technology sector shows average growth of around 5%, there are certain niches—such as the membrane treatment and advanced oxidation technologies—which are currently enjoying growth more in the 10% to 15% range.

Table 2. Size of the Water Industry (Revenues in Millions)

Business Segment	2003 Revenue	'04–'06 Growth
Water treatment equipment	$8,860	4–6%
Delivery equipment	$8,880	2–3%
Chemicals	$3,660	0–1%
Contract operations	$2,290	6–10%
Consulting/engineering	$6,090	5–6%
Maintenance services	$1,640	3–5%
Instruments/monitoring	$800	5–7%
Analytical testing	$480	2–4%
Wastewater utilities	$30,780	3–4%
Drinking water utilities	$32,650	3–4%
Total U.S. water industry	$96,130	

Source: *Environmental Business Journal*, 2003

KEY MARKET DRIVERS

As we have discussed, a handful of key factors—economic and social demands and political realities—are driving the overall development of the water industry. These drivers, in turn, are giving rise to various trends and effects that will likely dominate the water industry for years to come. Some of the key factors are listed below.

Water Scarcity and Water Quality Problems Are Growing

This is clearly the critical core issue behind this entire industry and the ultimate driver behind the challenges and growth projected for the water business over the coming decades. The gradual development of this situation has resulted from decades (indeed, centuries) of unfettered industrial expansion, continuing population growth, and a careless and uninformed belief that the environment would take care of itself. We do not need to recite the statistics again here, as they are well known—the millions of children who die each year because of unsanitary water conditions, the major cities of the world that still discharge their untreated wastes directly into the natural waterways, the number of people even in the United States who still drink water that is out of compliance with the Safe Drinking Water Act, and so on. Hundreds of studies have pinpointed the lack of sufficient clean water as one of the most serious threats facing humankind.

Public Awareness and Concerns Are Growing

As these water scarcity and quality problems have become more serious and more apparent, the public has become better informed and more concerned about the water problems that their children and grandchildren may inherit. One needs only to look at the pages of the popular press to see how broad and widespread this recognition is becoming. Stories about water scarcity or pollution problems are common headlines and have even become the stuff of popular Hollywood entertainment. As the general public becomes more aware and concerned about water, peoples' demands and perceptions will become a more important driver in determining the shape of the business. (One need only look at the explosive growth of the bottled water industry over the past few years to see how significantly customer perceptions—rightly or wrongly—can create and drive new markets.)

Regulatory Controls and Enforcement Are Growing

In turn, as public awareness and concern have grown, it has translated inexorably into greater government review, legislation, and regulatory control. Although environmental regulation and enforcement have waxed and waned during the past three decades, when it comes to drinking water, the public is insistent on ever-stronger regulatory protection of our water resources. For example, there were just nine new bodies of regulations, or rules, implemented between 1974 and 1996. Since the 1996 reauthorization of the Safe Drinking Water Act (SDWA), there have been more than 10 new rules implemented.

Under the SDWA, some key ongoing areas of regulatory development include the disinfection byproduct rule, the interim enhanced surface water treatment rule, the arsenic rule, and the groundwater rule. On the *waste*water side of the business, the primary governing legislation is the Clean Water Act, where key trends have included a more intense focus on *nonpoint* source runoff, continuing implementation of the total maximum daily load (TMDL) regulations, and programs geared toward combined and sanitary sewer overflow regulations to control wet weather runoff in metropolitan areas. Overlying these strictly regulated areas is the growing concern about potentially "introduced" compounds—terrorism—and all of the security concerns regarding primary drinking water supplies that have arisen as a result of 9/11. And finally, although the pace and intensity vary, regulatory controls are also becoming stronger in most other regions of the globe.

Meeting These Needs Is Going to Require Huge Expenditures

Finally, for municipalities and industry to comply with these regulations and to maintain and expand water infrastructure, huge capital expenditures will be required over the coming decades. As we have mentioned, numerous studies have predicted how large this expenditure must be; the estimates vary, but it is definitely in the hundreds and hundreds of billions of dollars. These dollars represent a huge challenge to this country, but they also constitute a huge opportunity for firms that serve the water and wastewater treatment industry.

KEY INDUSTRY TRENDS

These drivers, in turn, are leading to various trends and developments, both in terms of how the economy uses water and in terms of supply and demand within the commercial water technologies and services industry.

Continuing Consolidation and Ownership Changes

The water industry has been experiencing a dramatic rearrangement of ownership and increasing consolidation, as firms strategically position to address these opportunities. The dizzying pace of transactions the last few years—particularly those involving the larger domestic players and major foreign acquirers—has dramatically altered the face of this industry.

On the other hand, it would be easy to get the impression—particularly in the last few months—that events may actually be headed in the other direction. Quite abruptly, several key buyers (who have pumped hundreds and hundreds of millions of dollars into U.S. water acquisitions in recent years) seem to be reversing direction and deciding to spin off key businesses. In September, Suez announced that it had sold its Nalco water treatment unit to a group of private equity investors for $4.2 billion—roughly the same price for which it had acquired the business 4 years earlier. In its announcement, the company cited a narrowing strategic focus and the need of the parent organization to reduce its debt.

But the real surprise came in late September, when Veolia's U.S. Filter unit announced plans to sell off many businesses it had acquired through the 1990s. U.S. Filter has been the "king" of the emerging water industry for more than a decade, having conducted dozens of consolidating acquisitions during the 1990s, before being acquired by Veolia (Vivendi) in 1999. Although many had expected the company to spin off a few units here and there as it continued to focus its business, few expected a divestiture of this magnitude. Veolia cited considerations similar to those of Suez—the need to reduce high debt at the parent company level and further strategic concentration on its historical core business in contract operations and outsourcing services.

As of early February 2004, the sale of U.S. Filter's businesses is still ongoing. Everpure, the firm's water filtration products business, recently sold to Pentair Corporation at a very high price–$215 million for a business doing $60 million in revenue per year and producing some $20 million of EBITDA per year. The well-known Culligan subsidiary has attracted interest, but to date, it has not been sold. The remaining bulk of U.S. Filter's equipment business, which represents $1.2 billion in revenues, also remains on the block. U.S. Filter's preference that the business be sold as one piece, and the likelihood that it will trade at a multiple of greater than one times revenues, has obviously restricted the pool of capable and interested buyers. There simply are not many buyers financially capable of undertaking that large a purchase.

Some of the major British water companies also have begun to divest various units, again, primarily equipment and product businesses. In other recent transaction news:

- Ionics, one of the larger remaining independent players in the water industry, acquired Ecolochem, a private water treatment company, at a value of more than three times revenues. This was a large deal for Ionics, which has been struggling financially during the last 2 years.
- ITT Industries maintained its rapid pace of expansion in the industry, buying several more small players, but also commencing a bid for the major German ultraviolet treatment technology firm Wedeco. Wedeco had approximately $150 million of revenue in its latest fiscal year. ITT remains one of the most active buyers in the U.S. water market.

- Pentair acquired WICOR, a unit of Wisconsin Energy that produces water systems, filtration products, and swimming pool products and services. The deal was priced at $850 million in cash for a company generating approximately $750 million in revenues. At the same time, Pentair announced the divestiture of some of its other nonwater-related businesses, culminating a dramatic strategic repositioning of the firm into almost a water industry "pure play."
- Philadelphia Suburban Water company, now the largest independent investor-owned utility in the country, continues to make additional acquisitions. Most recently, it acquired several businesses from Allete Water in North Carolina. Philadelphia Suburban also changed its name to Aqua America to reflect its new and broader strategy.
- Finally, as always, there are also numerous smaller deals occurring, including the acquisition of Flowmatic, a reverse osmosis component manufacturer, by Watts Water, and the sale of Waterlink's remaining Barneby Sutcliffe operation to Calgon Carbon Corporation.

So many major assets are up for sale at the same time that the competitive situation in the water treatment equipment industry has been turned upside down. It will be interesting to watch where these various assets end up and to see which companies step forward to be major players in the next generation of this industry. Most observers are betting on the various diversified U.S. companies mentioned above—ITT Industries, GE Water, Pentair, and perhaps several others who have not yet made their first move. But the big questions remain—how can companies that were committed buyers turn into sellers in just a few short years? Was something wrong with their business strategies? Will the new owners of these assets be more logical owners than the previous ones? And most importantly, what will be the ultimate impact of this massive ownership rearrangement on employees, shareholders, and finally, the customer?

The Pace of Privatization and Outsourcing Begins to Slow

As mentioned before, one of the most controversial aspects of the water industry today involves the transfer of water management, treatment, and distribution operations from public to private control. Although private operation of water and wastewater utilities has been common in parts of Europe for decades, increasing privatization in North America and other parts of the world seems to be generating increasingly bitter political debate. During the past couple of years, it has become considerably less clear how far privatization will actually proceed in this country. The highly publicized misfortunes of several large projects like Atlanta, combined with an active and organized opposition movement, has forced a wholesale reevaluation of water and wastewater privatization. A number of large projects have been shelved or put on hold, and the whole nature of private contract operations is undergoing a major shift.

As a result, although the industry had been growing at rates of as high as 25% per year, considerably lower growth is now expected. As an example of how tough the last year was, one industry publication recently indicated that the U.S. contract operations business had seen the number of contracts drop off by 20%, whereas the *value* of the contracts declined by almost 73% during the year. OMI and Thames stand to lose as much as $300 million if their project in Stockton, California, is unwound. United Water was a big loser with the Atlanta job and with PRASA in Puerto Rico, which would have been one of the world's largest privatization projects.

The contract operations business also seems to be experiencing somewhat of a shift from a major "big-city" project orientation to smaller and shorter term contracts for smaller municipalities. The firms that have focused on this latter type of business seem to be doing reasonably well, whereas the firms that have focused on the huge projects seem to be having a more difficult time. Nonetheless, many firms continue to believe that contract operators have a very important and growing role to play in the industry. Veolia/U.S. Filter is obviously still very optimistic about the potential to manage the larger privatization jobs successfully and profitably, as indicated by its current effort to strategically reposition the firm primarily as a contract services provider.

And, in the bigger scheme of things, it is important to note that the simple threat of privatization has already forced widespread efficiencies by itself—a sort of "de facto" privatization. In summary, despite the concerns of labor organizations and various public interest groups, the urgency of infrastructural needs and the political barriers against major tax increases make it seem likely that privatization will continue to grow. At the same time, it is clear that private operators are going to be judged by a very demanding and critical public. A few other key trends are also important to note in any broad review of this industry.

Increased Focus on Water Recycling and Reuse. There are continuing and inexorable pressures toward greater water reuse and recycling systems—an obvious but as yet largely untapped means of addressing water shortages. With technologies readily available today, wastewater can easily be cleaned to levels where it can be recycled back into primary usage and at steadily declining costs. Examples of new wastewater reuse projects are coming out regularly now—typically for irrigation projects, or for more innovative applications, such as one recent project where treated wastewater was injected into the ground to act as a barrier against seawater intrusion to protect underground freshwater aquifers in Southern California.

Direct reuse of wastewater, particularly for drinking, is still a bit of a stretch for most Americans (and indeed is only commercially practiced in a small handful of locations around the world). However, this resistance is primarily due to poor understanding of the hydrologic cycle and our nation's current water usage patterns. For example, on some major river systems in the United States, water is used and reused up to 20 times as it travels to the sea. The discharge water from one wastewater treatment plant contributes to the raw water intake for a primary drinking water plant a few miles downstream. As a result of 30 years of steady progress under the Clean Water Act, the discharged waters from wastewater treatment plants are often cleaner than the rivers and streams they flow into. This type of *indirect* reuse of wastewater for drinking (after it has flowed in and out of a river, or into and out of an underground aquifer, and then through a treatment process) is clearly widespread and is obviously "acceptable" to most Americans.

An interesting statistic to remember when evaluating the long-term impact of wastewater reuse to extend our primary water resources is that only a tiny percentage of our primary water supply is actually used for drinking. Compared with the roughly 1200 gallons of water per capita per day that it currently takes to power the U.S. economy, an individual typically drinks less than a gallon a day. This leaves 1199 gallons per person per day that could be recovered without anyone ever having to drink "recycled" wastewater. Hence, even if only small incremental gains could be made in nonpotable water reuse, overall water availability could be substantially impacted. Over the longer term, we will move beyond our "linear" thinking of today and develop a more "circular" philosophy of water usage.

Conservation and Efficient Water Use

Closely related to improving our water reuse habits as a nation is the whole area of conservation of water and more efficient usage of water. During the droughts of 2002, many of us were forced to see how efficient we could become—and most of us found that it was not that difficult to save quite a bit of water. Indeed, water conservation ideas are only beginning to take advantage of the "low-hanging fruit." As water prices rise, there will be increasing incentives for people to use water more carefully and for industrial companies to retool their manufacturing systems to use less water.

OUTLOOK FOR THE FUTURE

As we have said many times, the challenges and requirements of the water industry are likely to be one of the most pressing problems facing humankind over the next century. The primary and overriding conclusion that falls out of the foregoing discussion is the inevitability of continuously rising water prices over the longer term future; as water becomes more scarce, there seems to be no alternative. Water is still very inexpensive—ridiculously cheap in many ways. And, at least in the United States, we are still on the highly inelastic portion of the demand curve for water, where increases in price have relatively little impact on usage. This inescapable conclusion and the vast scale of the world's water problems were highlighted in a special report in the July 2003 issue of the respected British newsweekly *The Economist*, which concisely concluded that water is "ill-governed and colossally under-priced." Discouragingly, the study also reported that the United States is the most wasteful nation on the earth in water usage.

As prices continue to increase, decisions about water usage will necessarily begin to take on greater economic significance. This will force us to focus on more efficient water usage and demand management techniques, technologies such as membrane desalination, and practices such as improved water conservation and recycling. All of these are already happening, but in the future, water expenditures will make up a larger and larger share of the GNP.

Water will increasingly be recognized as an economic good. However, exactly what that means is subject to varying interpretation. Some argue that water should be treated as a fully tradeable market commodity, subject to the general forces of supply and demand in an unregulated market—that water's value is essentially the same as its free market price. On the other hand, there is the antimarket argument—that water should be exempt from market forces because it is an essential prerequisite to life. Adherents to this philosophy argue that water is a basic human right and that forces greater than the free market are required to ensure that everyone has enough clean water to live.

Over the long term, some sort of intermediate position is likely to prevail—water should be treated as a scarce resource, which means that we have to balance economic and social objectives and carefully allocate water to its myriad uses. Market incentives can be a powerful means of properly managing and allocating a scare resource, but in the case of water, we must ensure that such markets are sufficiently regulated to protect social equity as well.

In closing, consider the following "big questions"—many more of which will continue to emerge in the future:

- Water is being transferred from American agriculture to American cities at an alarming rate. Farmers in the southwest are increasingly finding that current economics suggest that they stop farming and sell their underlying water rights to neighboring cities. What are the long-term implications of allowing our farmlands to dry up to slake the thirst of our ever-expanding cities?
- Why do we as a society treat such vast amounts of water to drinking quality standards when less than 1% is actually used for drinking? Should we really be incurring the capital costs of providing drinking quality water for fire fighting, mixing cement, washing cars, or watering yards?
- As sedimentary deposits fill the massive reservoirs that we have constructed during the last 60 to 80 years, how will we continue to provide controlled water supplies to the arid southwestern part of the country? We have become experts at building dams. What do we do with them when they no longer function?

So ... the water industry is full of challenges, but it is also full of opportunities. Water is an essential prerequisite of life, and we are not going to find any substitutes for it. The amount of freshwater on this earth is relatively fixed, and we need to become much smarter and more efficient in our usage of that scarce resource. For creative, innovative, and well-managed firms, the water industry offers unrivaled opportunities.

THE GLOBAL DESALINATION MARKET: TRENDS AND OUTLOOK FOR THE FUTURE

Lisa Henthorne and Eric Jankel

The desalination market has reached an all-time high in terms of growth in capacity, primarily due to the fact that the cost to desalinate water continues to decrease and now stands at less than $2.00 per thousand gallons of produced clean drinking water. Remarkably, in 2001, over 845 million gallons per day (gpd) of new capacity was contracted for installation worldwide, making it the highest recorded year in contracts for new installations. The final numbers for 2002 are expected to be even higher. Prior to 1996, an average of about 265 million gpd of new capacity was contracted per year, which rose to about 500 million gpd average per year in the years between 1996 and 2000. As a result, this is a very exciting time to be involved in the desalination market. Presently, approximately 8.5 billion gpd of desalination capacity has been installed or contracted through the conclusion of calendar year 2001 in 9500 plants worldwide.

The desalination market is in a state of rapid change. Membrane plants using seawater as a source and reverse osmosis (RO) as a process are being designed and constructed with capacities that were unheard of as recently as 1-5 years ago. Examples include the Ashkelon, Israel, and Fujuirah, U.A.E., plants of 87 million gpd and 45 million gpd, respectively. Unit capacities for multiple stage flash (MSF) distillation units are approaching 20 million gpd each. The Middle East has traditionally been the focus of desalination, predominantly using thermal technology such as MSF, but today large-scale desalination plants are being built throughout the world; plants outside the Middle East prefer RO technology. As a result, there is a broad-based shift occurring, from thermal technologies to membrane technologies for desalination applications. In 1990, 60% of the world's existing desalination capacity used thermal technologies, whereas at the end of 2001, this had dropped to 47%; the remaining capacity uses membrane technology.

Worldwide, desalination installed capacity has grown at an average rate of 10.6% per year over the last 30 years. In more recent few years, the growth rate has begun to increase. The key drivers behind the tremendous growth in the desalination market can be summarized as follows:

1. Significant reductions in the cost of desalting water due to
   - improved productivity and reduced cost of membrane elements in reverse osmosis;
   - increased efficiencies and economies of scale improvements in thermal processes;
   - the global trend toward privatization of water/power projects; and
   - improved energy recovery devices, which reduce the net energy requirement for reverse osmosis.
2. Changing demographics, specifically, population growth in arid, semiarid areas, and water-deficient areas.
3. Continuing increased standards of living in water-deficient areas.
4. Environmental concerns and increasing regulatory requirements that are forcing water suppliers to examine and use alternative sources of raw water.
5. Concerns and reaction to sustained drought conditions; seawater desalination is a drought-proof source of raw water.
6. Increased demand for high-grade water for industrial and commercial applications.

There are several recent desalination activities of particular note going on around the world. Due to increasing revenues as the price of crude oil has risen, construction of new desalination capacity and refurbishment of installed desalination capacity in the Gulf States is taking place at a rapid pace, especially in the U.A.E. The first large-scale (90 million gpd) privatized wastewater reuse RO plant is under construction in Kuwait using ultrafiltration (UF) as a pretreatment.

For the first time, a large-scale seawater RO plant (25 million gpd) has begun operating in North America, in Tampa Bay, Florida. At least 10 more large-scale seawater RO plants are in the planning stages in California, Texas, and Florida. Large-scale (40 million gpd) nanofiltration (NF) plants are also pending startup of operations in the United States.

The growth of both thermal and membrane technologies is shown graphically in (Fig. 1).

The primary factor behind the increasing growth of membrane desalination is the rapidly declining cost of membranes. A U.S. government-funded research and development program in the early 1960s developed and commercialized the first membrane-based RO process. The membranes that resulted from this effort became commercially available in the 1970s. Continued development and improvements in membrane productivity and reduction in pricing of the membranes due to competition among key manufacturers has continued through today. Now, over 200 hundred RO and nanofiltration products are available. Today, membrane-based reverse osmosis has become an extremely competitive and cost-effective means of producing large volumes of very high quality water.

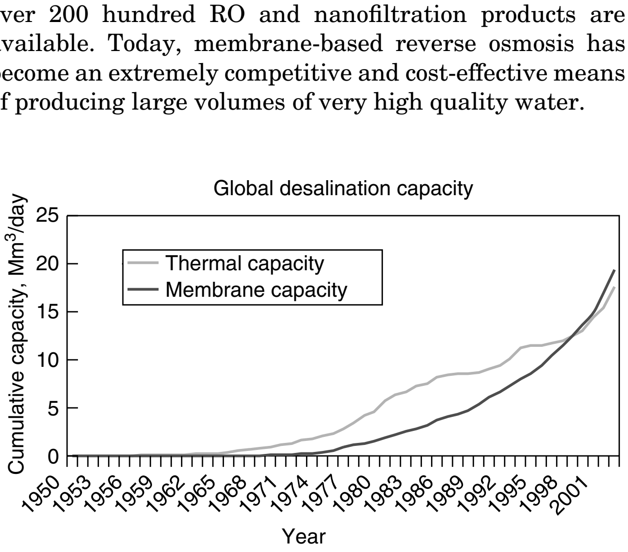

Figure 1. Cumulative capacity of all desalting technologies installed worldwide between 1950 and 2001 with projected plants in 2002.

Another way to evaluate the impact of the improvement in RO membrane properties on the cost of membrane technology is to consider the amount of water produced per unit capital cost. Figure 2 illustrates the increase in water produced using seawater desalination for the same capital investment, compared over time since 1980. These data indicate that for the same capital investment, approximately 27 times more water can be produced today using RO membranes than was possible in 1980.

Capital costs for membrane technology vary considerably based on the feedwater source and technology used. Table 3 provides a range of typical capital costs based on facilities contracted over the last few years and over a range of throughput capacities.

The privatized total water cost for seawater desalination using RO technology in large-scale facilities is generally in the $2.00 per thousand gpd price range.

A key recent project, which may portend future market trends and innovative approaches to membrane desalination in the United States, is the Tampa Bay Water program. This regional water utility, which supplies 225 million gallons per day of water to a five-county area, was required by regulators to reduce the pumping of groundwater at the same time that it faced a sustained period of drought. The utility responded by asking the industry to propose new alternatives for operating on a privatized basis. Through a long and arduous process, four bidders submitted tenders to supply a 25-million gpd desalination system. There are unique site-specific factors that contribute to the $2 per 1000 gpd price, including salinity fluctuations below average seawater levels (from 16,000 to 32,000 mg/L) and colocation; the adjacent power plant provides warmed cooling water and shared intake/outfall. The facility was developed by a privatized developer, Poseidon Resources, but ownership was transferred to Tampa Bay Water prior to the plant commissioning in May 2003.

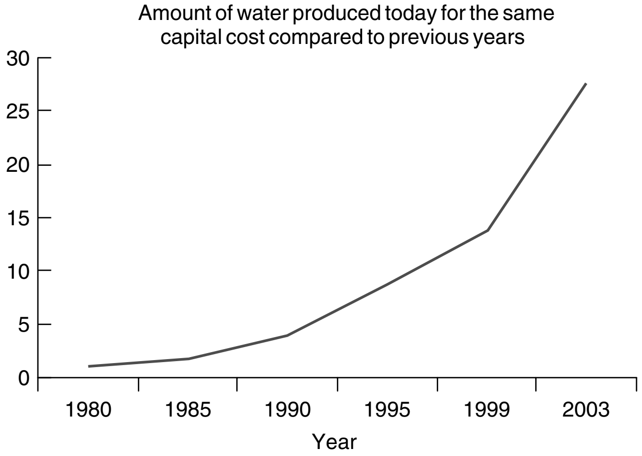

Figure 2. Change in water produced using seawater RO technology for the same capital investment over time.

Table 3. Range of Capital Costs for Membrane Desalination Plants (2001–2003)

Technology	Feedwater Source	Cost, $/gpd
NF	Wells with hardness and salinity	$1.13–$1.51
RO	Brackish wells	$1.13–$1.89
RO	Brackish surface	$1.51–$2.65
RO	Seawater	$2.84–$4.54
RO	Municipal wastewater	$2.27–$4.16

Covanta Water constructed the plant and would have operated the facility under a 25-year agreement with Tampa Bay Water. However, under a settlement agreement recently executed, Covanta has been expelled as the operator due to problems with the performance testing. Hydranautics, the RO membrane supplier used some unique design features for the facility, including a partial second pass to optimize product water quality. There are over 10,000 RO membranes installed at the facility, and their total energy demand is about 14 MW. Presently, the plant is experiencing difficulty in achieving successful pretreatment using the Covanta proprietary dual-sand filtration process.

The desalination market has a very bright future and should offer vast opportunities for private companies. The Middle East is still the largest market for desalination systems, maintaining approximately 49% of the worldwide contracted capacity. Figure 3 shows the percentage that each of the key regions of the world comprise, based on their desalination capacity contracted or installed through 2001.

Prospects for the continuing growth of the use of desalination technology are very optimistic. In the big picture, the availability of freshwater in the earth's hydrologic cycle is fixed; at the same time, the global population is growing, especially in arid and semiarid areas; standards of living are generally rising; existing sources of fossil groundwater are being depleted; surface supplies are being more stringently regulated; and the economic, political, and environmental costs of developing new sources of surface supply are increasing.

Reverse osmosis applications are anticipated to continue growing at roughly 11% to 12% in the future, and nanofiltration capacity is expected to grow at about 16%

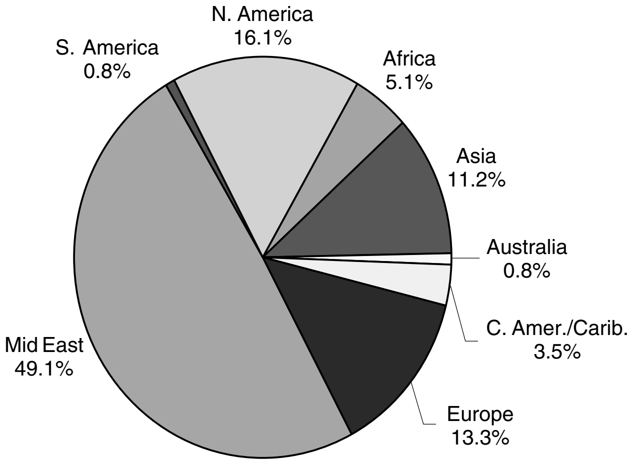

Figure 3. Distribution of desalination capacity by region based on installed capacity.

Table 4. Profitability and Performance[a,b]

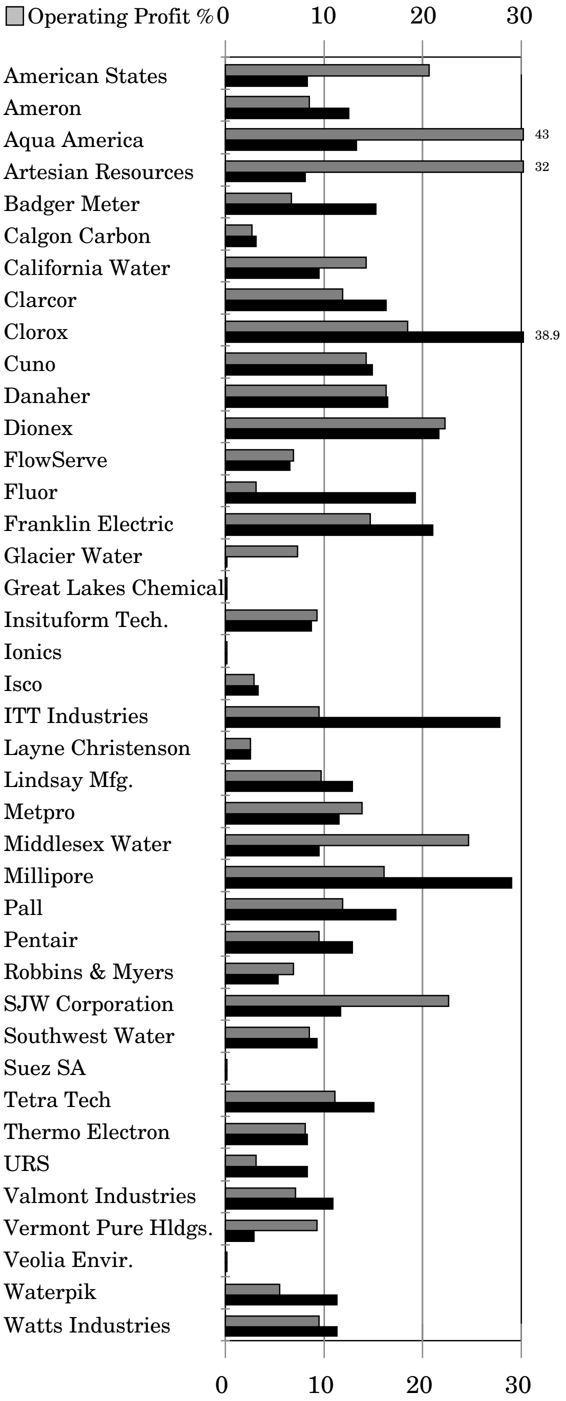

[a] ■ Return on equity, %.

[b] Where there is no bar, the number is zero or has a negative value.

Table 5. Balance Sheet Information[a]

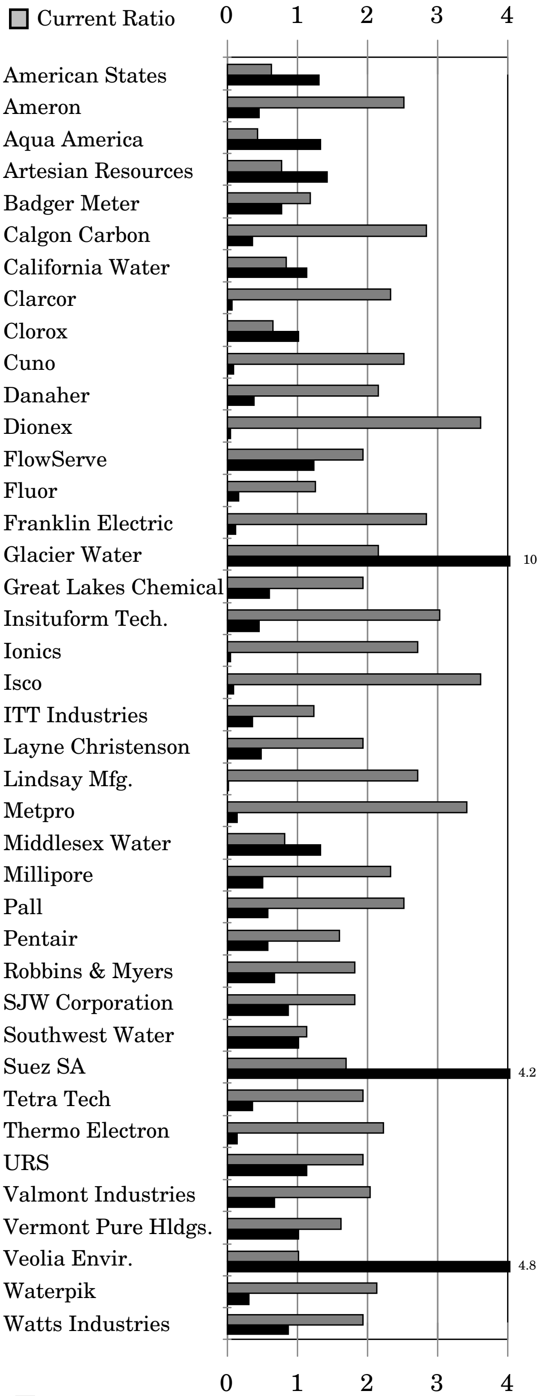

[a] ■ Debt/equity ratio.

per year. Thermal technologies and electrodialysis are anticipated to grow at lesser rates in the future. Over the period from 2004 to 2009, there will be over six billion gpd of new desalination capacity contracted at a capital investment of between $15 and $25 billion, as determined by our current evaluation and projections for the overall world desalination market.

FINANCIAL PERFORMANCE OF WATER COMPANIES

(*Note*: The data used to construct the charts and analysis in Tables 4–8 are from early February 2004.)

Table 6. Productivity Measures[a,b]

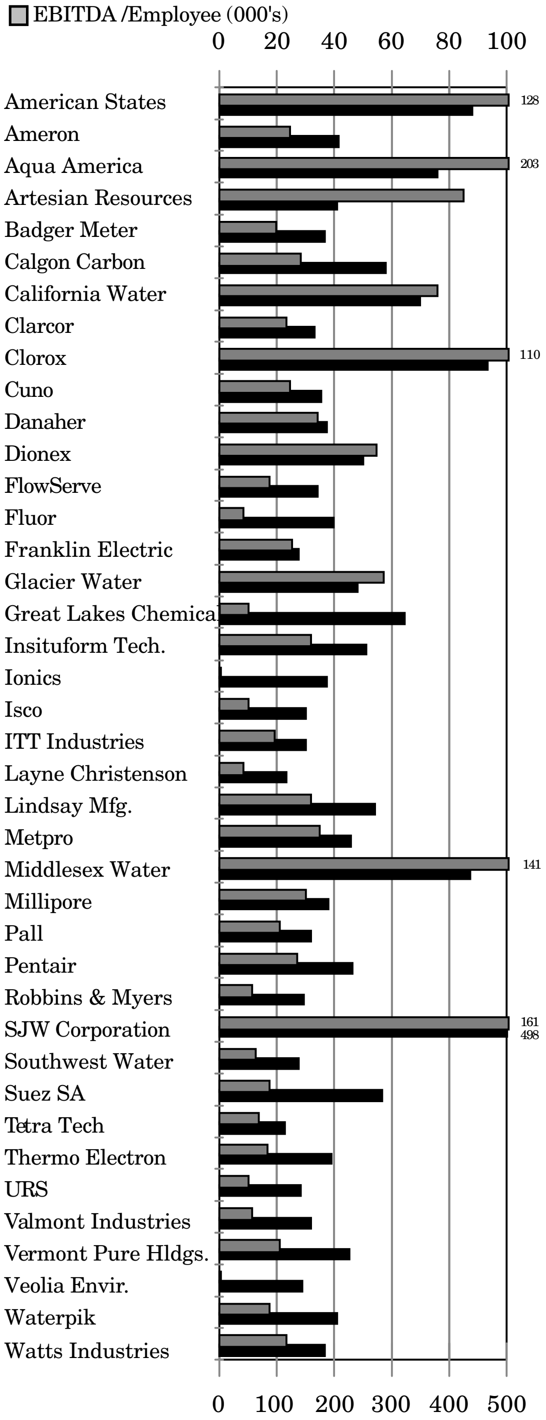

[a] ■ Revenues /employee (000's)

[b] Where there is no bar, the number is zero or has a negative value.

Table 7. Equity Valuations[a,b]

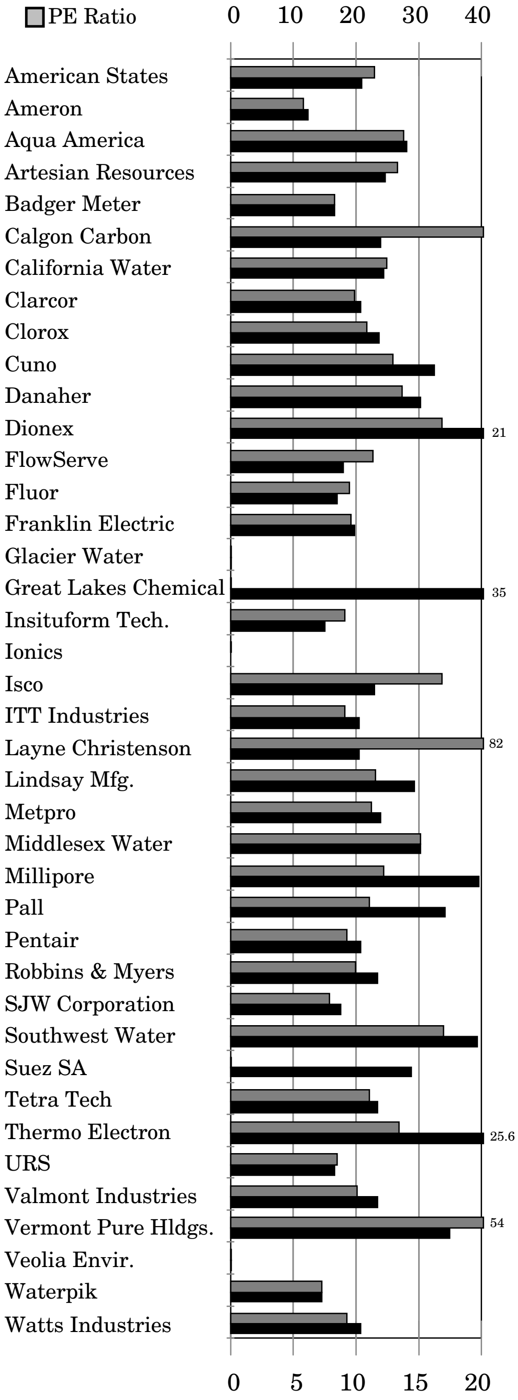

[a] ■ Enterprise value/EBITDA.

[b] Where there is no bar, the number is zero or has a negative value.

Table 8 demonstrates the recent strength in the broader stock market; most of the water firms traded at or near their 52-week highs, as was the overall market in early February. Table 7 is also a reflection of the price of the individual company's stock; most water firms show a P/E ratio between 15 and 20. The exceptions are a handful of firms, like Glacier Water and Great Lakes Chemicals, who have negative earnings for the trailing 12-month period used in these calculations, and for whom this is hence not a relevant measure.

Table 8. Stock Price Performance

Company	Price
American States	$25.55
Ameron	$39.04
Aqua America	$21.00
Artesian Resources	$28.25
Badger Meter	$37.34
Calgon Carbon	$19.90
California Water	$28.00
Clarcor	$42.25
Clorox	$49.66
Cuno	$41.24
Danaher	$94.67
Dionex	$54.66
FlowServe	$19.99
Fluor	$40.83
Franklin Electric	$59.85
Glacier Water	$20.10
Great Lakes Chemical	$25.24
Insituform Tech.	$16.00
Ionics	$28.52
Isco	$ 8.52
ITT Industries	$77.05
Layne Christenson	$13.25
Lindsay Mfg.	$24.83
Metpro	$17.07
Middlesex Water	$20.42
Millipore	$49.88
Pall	$26.28
Pentair	$53.40
Robbins & Myers	$19.75
SJW Corporation	$97.63
Southwest Water	$13.98
Suez SA	$21.69
Tetra Tech	$21.84
Thermo Electron	$28.34
URS	$28.95
Valmont Industries	$21.95
Vermont Pure Hldgs.	$ 3.42
Veolia Envir.	$28.98
Waterpik	$12.65
Watts Industries	$22.60

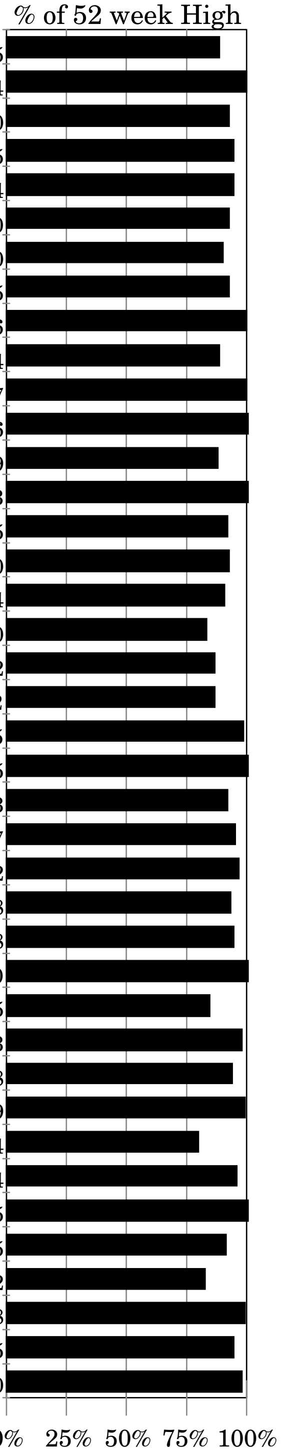

There is one important note about the Equity valuations in Table 7—the sometimes overlooked effect of simple arithmetic in ratio calculations. One of the inherent difficulties in interpreting either P/E ratios (or EBITDA multiples) is the effect of a very small value in the denominator. Generally speaking, companies with higher P/Es are perceived to be more valuable; however, this applies only within given ranges of "reasonableness." For example, the very high estimated P/E ratios for Layne Christenson or Vermont Pure are more the result of tiny earnings than intense interest in and demand for the stock. Detailed understanding of the company's individual economic circumstances and some judgment are needed to interpret what an individual P/E ratio really means.

Two different measures of profitability are shown in Table 4, operating profit as a percentage of revenue and return on equity. As can be seen, many companies in the water industry generate operating profits between 10% and 15%, which is higher than most other environmental companies. It is also clear that the water utility companies, as regulated local monopolies, have uniformly higher operating profits; see Artesian Resources, Aqua America, American States, and SJW. Return on equity measures not only general profitability, but also the relative balance of debt and equity in financing the company. Table 5 shows two key figures from a company's balance sheet, the current ratio, or ratio of currents assets to current liabilities, and the longer term debt to equity ratio. A higher current ratio usually implies a stronger short-term financial situation. As detailed in the cover story, note the relatively much higher debt levels of both key French companies. Veolia and Suez.

Table 6 illustrates two rough productivity measures—the amount of earnings (actually EBITDA is used here) and revenue that a company manages to generate per employee. Interestingly, when compared in this manner, water companies show a wide range of variation. Consulting and engineering firms, for example, typically show revenues per employee in the $100,000 to $150,000 range, whereas equipment manufacturers may show considerably higher figures. Note again the relatively high earnings per employee that the drinking water utilities are able to produce.

IRON AND MANGANESE REMOVAL

National Drinking Water
Clearinghouse

Iron and manganese are common in groundwater supplies used by many small water systems. Exceeding the suggested maximum contaminant levels (MCL) usually results in discolored water, laundry, and plumbing fixtures. This, in turn, results in consumer complaints and a general dissatisfaction with the water utility.

There are secondary standards set for iron and manganese, but these are not health related and are not enforceable. The secondary (aesthetic) MCLs for iron and manganese are 0.3 milligrams per liter (mg/L) and 0.05 mg/L, respectively.

Small water plants may choose to either sequestrate or remove iron and manganese. Sequestration only works for combined iron and manganese concentrations up to 1.0 mg/L and only in cases where the treatment is not permanent. Removal is usually achieved through ion exchange or oxidation/filtration. There are a number of chemical oxidants and filtration media available that can be used in various combinations

WHAT PROBLEMS ARE CAUSED BY IRON AND MANGANESE?

Small amounts of iron are often found in water because of the large amount of iron present in the soil and because corrosive water will pick up iron from pipes. Clothing washed in water containing excessive iron may become stained a brownish color. The taste of beverages, such as tea and coffee, may also be affected by iron. Manganese produces a brownish color in laundered clothing, leaves black particles on fixtures, and—as with iron—affects the taste of beverages, including coffee and tea.

Well water from the faucet or tap is usually clear and colorless. However, when water containing colorless, dissolved iron is allowed to stand in a cooking container or comes in contact with a sink or bathtub, the iron combines with oxygen from the air to form reddish-brown particles (commonly called rust). Manganese forms brownish-black particles. These impurities can give a metallic taste to water or to food.

The rusty or brown stains on plumbing fixtures, fabrics, dishes, and utensils cannot be removed by soaps or detergents. Bleaches and alkaline builders (often sodium phosphate) can make the stains worse. Over time, iron deposits can build up in pressure tanks, water heaters, and pipelines, reducing the quantity and pressure of the water supply.

Iron and/or manganese in water creates problems common to many water supply systems. When both are present beyond recommended levels, special attention should be paid to the problem. How iron and manganese are removed depends on the type and concentration and this helps determine the best procedure and (possible) equipment to use.

WHAT IS THE CHEMISTRY OF IRON AND MANGANESE IN WATER SYSTEMS?

Iron (Fe) and manganese (Mn) can be present in water in one of three basic forms:

1. Dissolved: ferrous (Fe^{2+}) and manganous (Mn^{2+})
2. Particulate: ferric (Fe^{3+}) and manganic (Mn^{4+}) states
3. Colloidal: very small particles (difficult to settle and filter).

The predominance of one form over another is dependent on the pH, Eh (redox potential), and temperature of the water. Knowledge of the forms or states of iron and manganese can help finetune a given treatment practice for these metals.

WHAT ARE THE MOST COMMON TREATMENT PROCESSES?

The majority of iron and manganese treatment systems employ the processes of oxidation/filtration. The oxidant chemically oxidizes the iron or manganese (forming a particle), and kills iron bacteria and any other disease-causing bacteria that may be present. The filter then removes the iron or manganese particles (Fig. 1).

Oxidation followed by filtration is a relatively simple process. The source water must be monitored to determine proper oxidant dosage, and the treated water should be monitored to determine if the oxidation process was successful.

Oxidation

Before iron and manganese can be filtered, they need to be oxidized to a state in which they can form insoluble complexes. Oxidation involves the transfer of electrons from the iron, manganese, or other chemicals being treated to the oxidizing agent. Ferrous iron (Fe^{2+}) is oxidized to ferric iron (Fe^{3+}), which readily forms the insoluble iron hydroxide complex $Fe(OH)_3$. Reduced manganese (Mn^{2+}) is oxidized to (Mn^{4+}), which forms insoluble (MnO_2).

The most common chemical oxidants in water treatment are chlorine, chlorine dioxide, potassium permanganate, and ozone. Oxidation using chlorine or potassium permanganate is frequently applied in small groundwater systems. The dosing is relatively easy, requires simple equipment, and is fairly inexpensive.

Chlorination is widely used for oxidation of divalent iron and manganese. However, the formation of trihalomethanes (THMs) in highly colored waters may be a problem. Chlorine feed rates and contact time requirements can be determined by simple jar tests.

As an oxidant, potassium permanganate ($KMnO_4$) is normally more expensive than chlorine and ozone, but for iron and manganese removal, it has been reported to be as efficient and it requires considerably less equipment and capital investment. The dose of potassium permanganate, however, must be carefully controlled. Too little permanganate will not oxidize all the iron and manganese, and too much will allow permanganate to enter the distribution system and cause a pink color. Permanganate can also form precipitates that cause mudball formations on filters. These are difficult to remove and compromise filter performance.

Ozone may be used for iron and manganese oxidation. Ozone may not be effective for oxidation in the presence of humic or fulvic materials. If not dosed carefully, ozone can oxidize reduced manganese to permanganate and result in pink water formation as well. Manganese dioxide particles, also formed by oxidation of reduced manganese, must be carefully coagulated to ensure their removal.

A low-cost method of providing oxidation is to use the oxygen in air as the oxidizing agent in a tray aerator. Water is simply passed down a series of porous trays to provide contact between air and water. No chemical dosing is required, which allows for unattended operation. This method is not effective for water in which the iron is complexed with humic materials or other large organic molecules. Oxygen is not a strong enough oxidizing agent to break the strong complexes formed between iron and manganese and large organic molecules. Furthermore, the rate of reaction between oxygen and manganese is very slow below pH values of 9.5.

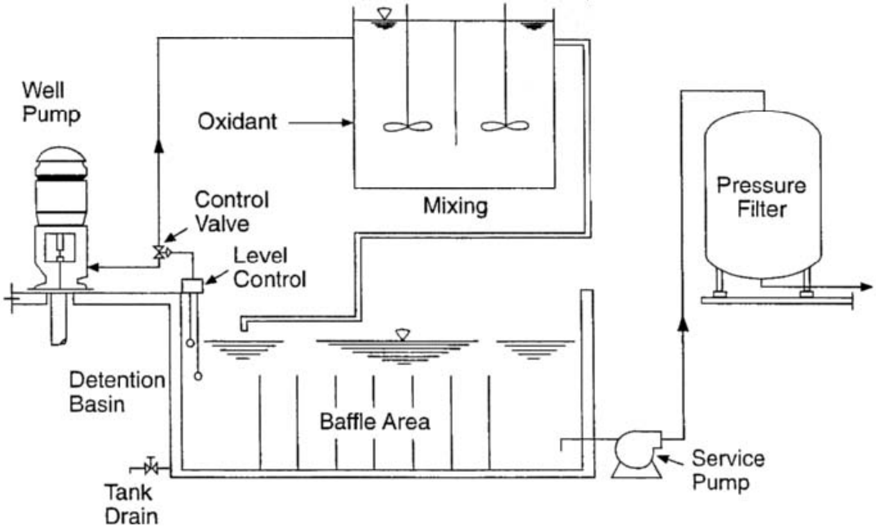

Figure 1. Chlorination, detention, and filtration (iron and manganese removal).

The presence of other oxidizable species in water hinders oxidation of the desired reduced compounds. Volatile organic chemicals, other organic compounds, or taste- and odor-causing compounds may result in an oxidant demand. This additional oxidant demand must be accounted for when dosing the oxidant. The expense of operation derives from the chemical use in most cases, and therefore is directly related to the source water quality.

Filtration

In general, manganese oxidation is more difficult than iron oxidation because the reaction rate is slower. A longer detention time (10 to 30 minutes) following chemical addition is needed prior to filtration to allow the reaction to take place.

There are different filtration media for the removal of iron and manganese, including manganese greensand, anthra/sand or iron-man sand, electromedia, and ceramic.

Manganese greensand is by far the most common medium in use for removal of iron and manganese through pressure filtration. Greensand is a processed material consisting of nodular grains of the zeolite mineral glauconite. The material is coated with manganese oxide. The ion exchange properties of the glauconite facilitates the bonding of the coating. This treatment gives the media a catalytic effect in the chemical oxidation-reduction reactions necessary for iron and manganese removal. This coating is maintained through either continuous or intermittent feed of potassium permanganate.

Anthra/sand (also iron-man sand) are other types of media available for removal of iron and manganese. They consist of select anthracite and sand with a chemically bonded manganese oxide coating. Unlike manganese greensand, these media are conditioned in the filter after media installation.

Electromedia provides a slightly different option from the manganese oxide coated media. This is a proprietary multi-media formulation which uses a naturally occurring zeolite and does not require potassium permanganate regeneration.

Finally, macrolite, unlike the other media discussed so far, is not a naturally occurring material which then undergoes processing for iron and manganese removal purposes. It is a manufactured ceramic material with a spherical shape and a rough, textured surface. The principal removal mechanism is physical straining rather than contact oxidation or adsorption.

Each medium has its advantages and disadvantages. Selection of a medium and oxidant should be based on pilot testing in which all necessary design criteria can be determined. Pressure filtration system manufacturers who offer the indicated media also offer fully automated systems.

ARE THERE ALTERNATIVE TREATMENTS?

Sequestration

Sequestration is the addition of chemicals to groundwater aimed at controlling problems caused by iron and manganese without removing them. These chemicals are added to groundwater at the well head or at the pump intake before the water has a chance to come in contact with air or chlorine. This ensures that the iron and manganese stays in a soluble form.

If the water contains less than 1.0 mg/L iron and less than 0.3 mg/L manganese, using polyphosphates followed by chlorination can be an effective and inexpensive method for mitigating iron and manganese problems. No sludge is generated in this method. Below these concentrations, the polyphosphates combine with the iron and manganese preventing them from being oxidized. Any of the three polyphosphates (pyrophosphate, tripolyphosphate, or metaphosphate) can be used.

To determine the best polyphosphate to use and the correct dosage, a series of samples at different concentrations may be prepared. Chlorine is added, and the samples are observed daily against a white background. The right polyphosphate dose is the lowest dose that does not noticeably discolor the water samples for four days.

Applying sodium silicate and chlorine simultaneously has also been used to sequester iron and manganese.

However, while this technique is reliable in the case of iron treatment, it has not been found to be effective in manganese control.

Ion Exchange

Ion exchange should be considered only for the removal of small quantities of iron and manganese because there is a risk of rapid clogging. Ion exchange involves the use of synthetic resins where a pre-saturant ion on the solid phase (the "adsorbent," usually sodium) is exchanged for the unwanted ions in the water (see Ion Exchange and Demineralization Tech Brief #DWBLPE56). One of the major difficulties in using this method for controlling iron and manganese is that if any oxidation occurs during the process, the resulting precipitate can coat and foul the media. Cleaning would then be required using acid or sodium bisulfate.

Other

Systems that have a lime-soda ash softening plant do not need a separate iron and manganese removal plant. The high pH during softening allows rapid oxidation and precipitation of the iron and manganese as well as incorporation in the calcium and magnesium precipitates. Similarly, surface water treatment plants using coagulation, flocculation, sedimentation, and filtration also will remove iron and manganese as long as they make certain the iron and manganese get oxidized. Oxidation is sometimes a problem because of the presence of organic matter.

Finally, biological treatment methods are being pilot tested at different locations. Biological treatment methods are used extensively in European countries, such as the Netherlands, France, and Germany, and are advantageous primarily when water simultaneously contains iron, manganese, and ammonia.

HOW CAN IRON AND MANGANESE PROBLEMS BE MINIMIZED IN DISTRIBUTION MAINS?

Problems due to iron and manganese in distribution mains may be minimized by:

- prior removal by appropriate treatment,
- protecting iron/steel mains with bituminous linings, or using noncorrosive materials,
- avoiding dead-end mains,
- avoiding disturbances in the water flow, and
- flushing periodically.

WHERE CAN I FIND MORE INFORMATION?

Information in this fact sheet was primarily obtained from the following sources:

(1) American Water Works Association. 1998. *Water Treatment Plant Design*. Third Edition.

(2) American Water Works Association. 1990. *Water Quality 2nd Treatment*. Fourth Edition.

(3) Metcalf, B. M. 1998. "Pressure Filtration for Iron & Manganese Removal," Proceedings of the New England Water Works Association Conference and Exhibition. Marlborough, MA.

(4) National Research Council. 1997. *Safe Water From Every Tap: Improving Water Service to Small Communities*. National Academy Press. Washington DC.

(5) Robinson, R. B. 1998. "State-of the-Art: Iron and Manganese Control," Proceedings of the New England Water Works Association Conference and Exhibition. Marlborough, MA.

(6) U.S. Environmental Protection Agency. 1991. *Manual of Small Public Water Supply Systems*. EPA 570/9-91-003. Office of Water, Washington, DC.

(7) Vigneswaran, S., C. Visvanathan. 1995. *Water Treatment Processes: Simple Options*. CRC Press. New York, NY.

National Drinking Water Clearinghouse (NDWC) offers a Registry of Equipment Suppliers of Treatment Technologies for Small Systems (RESULTS) database. This public reference database contains information about technologies in use at small water systems around the country. A database search could locate other small systems currently treating for iron and manganese.

For further information, call the NDWC at (800) 624-8301 or (304) 203-4191. Additional copies of Tech Brief fact sheets are free; however, postal charges are added to orders. To order, call the NDWC at (800) 624-8301 or (304) 293-4191. You may also order online at ndwc_orders@estd.wvu.edu, or download Tech Briefs from our Web site at http://www.ndwc.wvu.edu where they are available in the educational products section.

- Tech Brief: Disinfection, item #DWBLPE47;
- Tech Brief: Filtration, item #DWBLPE50;
- Tech Brief: Corrosion Control, item #DWBLPE52;
- Tech Brief: Ion Exchange and Demineralization, item #DWBLPE56;
- Tech Brief: Organics Removal, item #DWBLPE59;
- Tech Brief: Package Plants, item #DWBLPE63;
- Tech Brief: Water Treatment Plant Residuals Management, item #DWBLPE65
- Tech Brief: Lime Softening, item #DWBLPE67
- Tech Brief: Iron and Manganese Removal, item #DWBLPE70

EXTRATERRITORIAL LAND USE CONTROL TO PROTECT WATER SUPPLIES

LAUREL PHOENIX
Green Bay, Wisconsin

Cities continually need water to grow. If they are located near a river, lake, or groundwater source, they will use these closest sources first. Some cities, however, have needed to search for water far beyond their boundaries, typically looking for rivers, large lakes, or valleys suitable for damming to create reservoirs. Because these distant areas are in different political jurisdictions that may fight a city in its quest for water, a state may step in and give a city the power it needs to expropriate and subsequently control lands outside its municipal boundaries.

A city needing water from a new source may need extraterritorial control over land in that area in order to (1) establish the source, (2) maintain access to the source, and (3) maintain quality of the source. In establishing the source, a city can face numerous obstacles. There may be a small community using or living next to the water supply the city wants. It may not want to share that water or allow the city to put intake pipes into the water. It may worry that a drawdown of lake levels would change the shape of the shoreline, expose muddy littoral slopes (making it ugly), distance waterfront properties from the water, and make piers and docks useless. This would have a negative impact on property values around the lake. Future effects on property values would result if the city was given land use control authority in the watershed that would constrain the economic growth potential of the local community. Or what if the small community is in a valley bottom instead? It certainly will not want a dam and reservoir to drown its homes and businesses. Forcing an entire community to move not only destroys homes, businesses, farmlands, and forests, but also makes many local citizens lose their jobs and their social networks. Any community would naturally fight the forfeiture of its own potential economic growth to benefit a bigger city's economic growth.

The second step, maintaining access to the source, means that a city would need the right to travel over any properties necessary to maintain and repair miles of pipeline or to manage any dams.

It is the third step, maintaining quality of the source, which requires the most extraterritorial control of the source watershed. Different land uses create different nonpoint source pollution that can contaminate the source water. The city seeking the new water source will want to prevent watershed land uses that contaminate the source and require more money to treat the water. Inadequately treated municipal sewage, failing septic tanks, or farms with animals can all contribute *Escherichia coli (E. coli)* to the water source. Sediment from farm field erosion, steep building sites, or building too close to the water source can cause siltation of the reservoir and reduce the reservoir's storage capacity. Moreover, sediment adds to turbidity levels, and these total dissolved solids (TDS) can reach a high enough level that a drinking water treatment plant would have to shut down its intake pipes. Parasites such as *Cryptosporidium parvum* and *Giardia* spp., which cause cryptosporidiosis and giardiasis in humans, require extremely expensive filtration treatment, so cities that can keep a watershed very clean can obtain a filtration waiver from the Environmental Protection Agency (EPA) and avoid spending billions on filtration plants.

Because land use decisions most closely affect property values and the economic potential of a community, land use controls have historically been the right of local governments to decide. To give this decision-making power to another local government essentially gives the smaller community unequal rights in their state, a loss of control in determining their future, and a transfer of wealth to the larger city.

It is an extraordinary step for a state to give one local government extraterritorial control over another local government, as this gives the equivalent of limited state power to some local governments. Because of the critical nature of water supplies to cities, and because larger cities are a state's economic engines, a state will often pass legislation to allow establishment and access to a new water source. Some states have gone further to give extraterritorial land use control over communities in the source watershed to maintain the source water quality. To provide a few examples of extraterritorial land use controls, Boston, New York, and Syracuse are described here.

BOSTON

The Quabbin Reservoir is the largest and most recently developed water supply for the greater metropolitan area of Boston. The Swift River in western Massachusetts was impounded by a dam to flood a valley in the early 1930s. The Watershed Protection Act, 350 CMR 11 (to read the complete regulations, see http://www.mass.gov/mdc/350CMR11.html), delineates a variety of restrictions on land use in the four towns within the watershed in order to protect its water quality. Land has been categorized by proximity to a tributary or reservoir, presence of wetlands, areas overlying aquifers, and so on to distinguish the degree of regulation for each zone. The first buffer zone (within 400 ft of the reservoir or within 200 ft from a tributary or other surface water) is the Primary Protection Zone. No new construction, excavation, or grading is allowed. Dumping of any substance is illegal, and no new paving can be laid. Substances defined as pollutants cannot be manufactured, stored, applied, or dumped in this zone.

The second buffer zone is the Secondary Protection Zone. This includes vegetated wetlands, floodplain lands, land lying over selected aquifers, and any land between 200 and 400 ft of tributaries or surface waters. Development density is controlled (determined by X bedrooms per X acres), wetlands cannot be disturbed, and certain chemicals cannot be stored, used, or dumped. Sewage treatment facilities may not dump their effluent, leach fields must be at least 4 ft above the water table, liquid petroleum can only be stored in a structure that can hold 125% of the petroleum tank capacity, and this structure must be placed over a poured concrete floor. Storage or use of hazardous or solid waste beyond what a typical house would use is prohibited. Road salt, fertilizers, and pesticides must be stored indoors. No boat and car washing or servicing businesses are allowed, and no junkyards are allowed. The construction of impervious surfaces is limited to 10% of a lot. Recreational activities are limited in the land surrounding the reservoir. For example, hiking is allowed, but snowmobiles and off-road vehicles are not. Boats are not allowed on the lake, and no domestic animals or horseback riding are allowed.

NEW YORK

In 1905, the state legislated the creation of the New York City Board of Water Supply, which would be in charge of researching and developing bigger and cleaner water supplies for the city. New York City dammed several valleys in the Catskill mountains, inundating villages and farms. These dams were completed by 1927, and more dams were built in neighboring valleys that feed the Delaware River by the 1960s. All of these reservoirs were linked together by pipelines, and gravity flow delivered the water to the city boundaries. After years of contention between the watershed towns and New York City, the Final Rules and Regulations to protect the city's water supply became law on May 1, 1997 (to read the complete regulations, see http://www.ci.nyc.ny.us/html/dep/html/ruleregs/finalrandr.html). Because there is heavy pressure to develop land in these watersheds, these final rules and regulations list prohibited land uses, but the city agreed that it would never use the power of eminent domain to prevent new development in these areas. Rather, it would seek out willing sellers from which to purchase lands, and then it would leave these lands undeveloped. Development in these watersheds must be guided by proximity to water, slope of building lots, percentage of impervious area, and other requirements similar to that mentioned for the Quabbin watershed. Other pressures that New York City could have exerted were resolved by city-funded programs. For example, the city is paying to replace failing septics for houses and businesses. In addition, extra funding has been found to upgrade public sewage treatment plants to prevent their effluent from degrading the water quality of the reservoirs. Farms receive assistance on manure management and other farm activities that reduces their nutrient, sediment, and pesticide contributions to the reservoirs.

The New York City watershed agreement is distinct from how Boston's water authority controls the Quabbin watershed. The dynamics of the residents are different, as there are far more people living or wanting to live in the Catskill area. This means the Catskill residents have all the more potential money to lose if land use regulations in their watershed virtually eliminated the potential for economic growth.

SYRACUSE

The city of Syracuse, New York receives its water from Skaneateles Lake, one of the Finger Lakes of New York. Relative to Quabbin watershed or the Catskill watersheds, this lake lies in a much smaller watershed with a low-density population. There is a small but wealthy village that relies on tourism, a lake surrounded by large-lot estates, and a small amount of farmland. There are no industries or sewage treatment plants that discharge into this lake.

Syracuse has several tools to protect its water source. The watershed regulations (to read the complete regulations, see http://www.syracuse.ny.us/deptWater.asp), for example, allow Syracuse to test yearly for failing septic systems on every lakeshore property. The city also established two programs to protect water quality, the Skaneateles Lake Watershed Land Protection Program (SLWLPP) and the Skaneateles Lake Watershed Agricultural Program (SLWAP). The SLWLPP provides funds to buy permanent conservation easements from environmentally significant properties from willing sellers. The SLWAP helps farmers in the watershed create voluntary Whole Farm Plans that reduce nonpoint source pollution from these farms while saving farmers money.

THE CHOICES OF EXTRATERRITORIAL CONTROL

The ways in which these three cities exerted their extraterritorial control is influenced by how many citizens live in the watersheds now and how locals used their political voice to soften the heavy-handedness that total land use controls engender. Consequently, New York City and Syracuse had to develop more cooperative and voluntary solutions between municipalities to protect the water source without impoverishing the communities along their shores.

LEAK DETECTION AND WATER LOSS CONTROL

Zacharia Michael Lahlou
Technical Assistance
Consultant

Utilities can no longer tolerate inefficiencies in water distribution systems and the resulting loss of revenue associated with underground water system leakage. Increases in pumping, treatment and operational costs make these losses prohibitive. To combat water loss, many utilities are developing methods to detect, locate, and correct leaks.

Old and poorly constructed pipelines, inadequate corrosion protection, poorly maintained valves and mechanical damage are some of the factors contributing to leakage. One effect of water leakage, besides the loss of water resources, is reduced pressure in the supply system. Raising pressures to make up for such losses increases energy consumption. This rise in pressure makes leaking worse and has adverse environmental impacts.

Of the many options available for conserving water, leak detection is a logical first step. If a utility does what it can to conserve water, customers will tend to be more cooperative in other water conservation programs, many of which hinge on individual efforts. A leak detection program can be highly visible, encouraging people to think about water conservation before they are asked to take action to reduce their own water use. Leak detection is an opportunity to improve services to existing customers and to extend services to the population not served.

In general, a 10 to 20 percent allowance for unaccounted-for-water is normal. But a loss of more than 20 percent requires priority attention and corrective actions. However advances in technologies and expertise should make it possible to reduce losses and unaccounted-for-water to less than 10 percent. While percentages are great for guidelines, a more meaningful measure is volume of lost water. Once the volume is known, revenue losses can be determined and cost effectiveness of implementing corrective action can then be determined.

Figure 1. Shawn Menear, a graduate student in technology education at west virginia university, uses geophones to listen for water main leaks. Similar to a doctor or nurse's stethoscope, geophones are an inexpensive leak detection device used by water utilities.

BENEFITS OF LEAK DETECTION AND REPAIR

The economic benefits of leak detection and repair can be easily estimated. For an individual leak, the amount lost in a given period of time, multiplied by the retail value of that water will provide a dollar amount. Remember to factor in the costs of developing new water supplies and other "hidden" costs.

Some other potential benefits of leak detection and repair that are difficult to quantify include:

- increased knowledge about the distribution system, which can be used, for example, to respond more quickly to emergencies and to set priorities for replacement or rehabilitation programs;
- more efficient use of existing supplies and delayed capacity expansion;
- improved relations with both the public and utility employees;
- improved environmental quality;
- increased firefighting capability;
- reduced property damage, reduced legal liability, and reduced insurance because of the fewer main breaks; and
- reduced risk of contamination.

CAUSES OF LEAKS

Water produced and delivered to the distribution system is intended to be sold to the customer, not lost or siphoned from the distribution system without authorization. Not long ago, water companies sold water at a flat rate without metering. As water has become more valuable and metering technology has improved, more and more water systems in the U.S. meter their customers. Although all customers may be metered in a given utility, a fairly sizable portion of the water most utilities produce does not pass through customer meters. Unmetered water includes unauthorized uses, including losses from accounting errors, malfunctioning distribution system controls, thefts, inaccurate meters, or leaks. Some unauthorized uses may be identifiable. When they are not, these unauthorized uses constitute unaccounted-for water. Some unmetered water is taken for authorized purposes, such as fire fighting and flushing and blowoffs for water-quality reasons. These quantities are usually fairly small. The primary cause of excessive unaccounted-for water is often leaks.

CALCULATING UNACCOUNTED-FOR WATER

Unaccounted-for water is the difference between water produced (metered at the treatment facility) and metered use (i.e., sales plus non-revenue producing metered water). Unaccounted-for water can be expressed in millions of gallons per day (mgd) but is usually discussed as a percentage of water production:

$$\textbf{Unaccounted} - \textbf{for water}\ (\%) = \frac{(\text{Production} - \text{metered use}) \times 100\%}{(\text{Production})}$$

There are different types of leaks, including service line leaks, and valve leaks, but in most cases, the largest portion of unaccounted-for water is lost through leaks in the mains. There are many possible causes of leaks, and often a combination of factors leads to their occurrence. The material, composition, age, and joining methods of the distribution system components can influence leak occurrence. Another related factor is the quality of the initial installation of distribution system components. Water conditions are also a factor, including temperature, aggressiveness, and pressure. External conditions, such as stray electric current; contact with other structures; and stress from traffic vibrations, frost loads, and freezing soil around a pipe can also contribute to leaks. All water plants will benefit from a water accounting system that helps track water throughout the distribution system and identifies areas that may need attention, particularly large volumes of unaccounted-for water.

LEAK DETECTION AND REPAIR STRATEGY

There are various methods for detecting water distribution system leaks. These methods usually involve using sonic leak-detection equipment, which identifies the sound of water escaping a pipe. These devices can include pinpoint listening devices that make contact with valves and hydrants, and geophones that listen directly on the ground. In addition, correlator devices can listen at two points simultaneously to pinpoint the exact location of a leak. (See Figs. 1 and 2.)

Large leaks do not necessarily contribute to a greater volume of lost water, particularly if water reaches the surface; they are usually found quickly, isolated, and repaired. Undetected leaks, even small ones, can lead to large quantities of lost water since these leaks might exist for long periods of time. Ironically, small leaks are easier to detect because they are noisier and

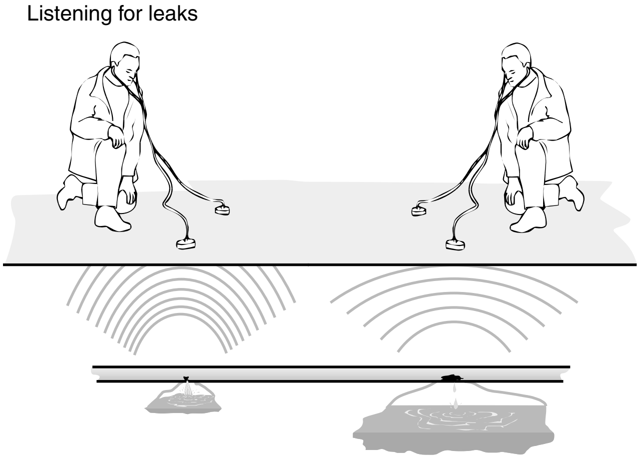

Figure 2. An important goal of leak detection is to find exactly where a leak is located. Typically, the louder the noise, the closer you are to the leak. Small leaks under high pressure usualy make more noise than larger leaks under low pressure. In fact, many large leaks make almost no sound whatsoever.

easier to hear using hydrophones. The most difficult leaks to detect and repair are usually those under stream crossings.

Leak detection efforts should focus on the portion of the distribution system with the greatest expected problems, including:

- areas with a history of excessive leak and break rates;
- areas where leaks and breaks can result in the heaviest property damage;
- areas where system pressure is high;
- areas exposed to stray electric current and traffic vibration;
- areas near stream crossings; and
- areas where loads on pipe exceed design loads.

Of course, detecting leaks is only the first step in eliminating leakage. Leak repair is the more costly step in the process. Repair clamps, or collars, are the preferred method for repairing small leaks, whereas larger leaks may require replacing one or more sections of pipe.

On average, the savings in water no longer lost to leakage outweigh the cost of leak detection and repair. In most systems, assuming detection is followed by repair, it is economical to completely survey the system every one to three years.

Instead of repairing leaking mains, some argue it is preferable to replace more leak-prone (generally older) pipes. Selecting a strategy depends upon the frequency of leaks in a given pipe and the relative costs to replace and repair them.

Deciding whether to emphasize detection and repair over replacement depends upon site-specific leakage rates and costs. In general, detection and repair result in an immediate reduction in lost water, whereas replacement will have a longer-lasting impact to the extent that it eliminates the root cause of leaks.

The most important factor in a leak detection and repair program is the need for accurate, detailed records that are consistent over time and easy to analyze. Records concerning water production and sales, and leak and break costs and benefits, will become increasingly important as water costs and leak and break damage costs increase and as leak detection and rehabilitation programs become more important. In order to optimize these programs by allocating funds in such a way that results in the greatest net benefits, adequate information is needed on which to base decisions and determine needs. Three sets of records should be kept: (1) monthly reports on unaccounted-for water comparing cumulative sales and production (for the last 12 months, to adjust discrepancies caused by the billing cycle); (2) leak-repair report forms; and (3) updated maps of the distribution system showing the location, type, and class of each leak.

COORDINATING LEAK DETECTION AND REPAIR WITH OTHER ACTIVITIES

In addition to assisting with decisions about rehabilitation and replacement, the leak detection and repair program can further other water utility activities, including:

- inspecting hydrants and valves in a distribution system;
- updating distribution system maps;
- using remote sensor and telemetry technologies for ongoing monitoring and analysis of source, transmission, and distribution facilities. Remote sensors and monitoring software can alert operators to leaks, fluctuations in pressure, problems with equipment integrity, and other concerns; and
- inspecting pipes, cleaning, lining, and other maintenance efforts to improve the distribution system and prevent leaks and ruptures from occurring. Utilities might also consider methods for minimizing water used in routine water system maintenance.

BEYOND LEAK DETECTION AND REPAIR

Detecting and repairing leaks is only one water conservation alternative; others include: meter testing and repair/replacement, rehabilitation and replacement programs, installing flow-reducing devices, corrosion control, water pricing policies that encourage conservation, public education programs, pressure reduction, requests for voluntary cutbacks or bans on certain water uses, and water recycling.

WHERE CAN I FIND MORE INFORMATION?

Jeffs, C., Lloyd C., and Pospishill D. 1989. *An Introduction to Water Loss and Leak Detection*. Duncan OK: National Rural Water Association.

Mays, W.L. 2000. *Water Distribution Systems Handbook*. American Water Works Association. New York: McGraw-Hill.

Moyer, E. M. 1985. *Economics of Leak Detection: A Case Study Approach*. Denver: American Water Works Association

Pask, David. 50 Percent Loss? How to Detect Small Utility Water Leaks. *On Tap*. Winter 1993. Morgantown WV: National Drinking Water Clearinghouse.

U.S. Environmental Protection Agency. 1998. *Water Conservation Plan Guidelines*. Washington, DC: Office of Water. EPA-832-D-98-001

For further information, comments about this fact sheet, or to suggest topics, contact Lahlou via e-mail at lahloum@hotmail.com.

Zacharia M. Lahlou, Ph.D., formerly technical assistance coordinator with the National Drinking Water Clearinghouse, is a civil and environmental engineer with Wiley & Wilson, Lynchburg, VA. Lahlou received a doctorate in Environmental and Natural Resources Economics, an MBA and a Master of Science in Civil and Environmental Engineering from West Virginia University. He may be reached by e-mail at *lahloum@hotmail.com*.

HAVE YOU READ ALL OUR TECH BRIEFS?

Tech Briefs, drinking water treatment fact sheets have been a regular feature in the National Drinking Water Clearinghouse (NDWC) newsletter *On Tap* for more than five years. Former NDWC Technical Assistance Coordinator Zacharia M. Lahlou, Ph.D, researches, compiles information, and writes these very popular items.

Tech Brief: Disinfection, item #DWBLPE47;
Tech Brief: Filtration, item #DWBLPE50;
Tech Brief: Corrosion Control, item #DWBLPE52;
Tech Brief: Ion Exchange and Demineralization, item #DWBLPE56;
Tech Brief: Organics Removal, item #DWBLPE59;
Tech Brief: Package Plants, item #DWBLPE63;
Tech Brief: Water Treatment Plant Residuals Management, item #DWBLPE65;
Tech Brief: Lime Softening, item #DWBLPE67;
Tech Brief: Iron and Manganese Removal, item #DWBLPE70;
Water Conservation Measures Fact Sheet, item #DWBLPE74;
Tech Brief: Membrane Filtration, item #DWBLPE81;
Tech Brief: Treatment Technologies for Small Drinking Water Systems, item #DWPSPE82;
Tech Brief: Ozone, item #DWBLPE84.
Tech Brief: Radionuclides, item #DWBLPE84.
Tech Brief: Slow Sand Filtration, item #DWBLPE99.
Tech Brief: Ultraviolet Disinfection, item #DWBLPE101.
Tech Brief: Leak Detection and Water Loss Control, item #DWBLPE102.

A package of the *Tech Briefs* is now available as a product. A three-ring binder holds all the current *Tech Briefs* in print. New selections can be easily added to the package as they become available. To order this product, call the NDWC at the numbers listed below and ask for item #DWPKPE71. The item is free.

Additional copies of fact sheets are also free; however, postal charges may be added.

To order, call the NDWC at (800) 624-8301 or (304) 293-4191. You also may order online at ndwc_order@mail.nesc.wvu.edu *or download fact sheets from our Web site at* www.ndwc.wvu.edu.

LIME–SODA ASH PROCESSES

Roger C. Viadero Jr.
West Virginia University
Morgantown, West Virginia

INTRODUCTION AND PROCESS OVERVIEW

As in other chemical water softening technologies, the objective of the lime–soda ash process is to precipitate polyvalent cations from solution by pH adjustment. In water treatment applications, Ca^{2+} and Mg^{2+} are the two major cations of concern, though Fe^{2+}, Fe^{3+}, Al^{3+}, and Mn^{2+} may also be present at problematic concentrations, depending on the properties of soil or geological material encountered by the surface or groundwater. Typically, a total hardness ranging from 75–120 mg/L as $CaCO_3$, is sought in domestic water softening.

Total [water] hardness (TH) is composed of carbonate and noncarbonate fractions, where carbonate hardness (CH) is the equivalent concentration (meq/L or mg/L as $CaCO_3$) of polyvalent cation species associated with carbonate ions [e.g., $Mg(HCO_3)_2$]. Similarly, noncarbonate hardness (NCH) is the equivalent concentration of all aqueous phase cations not associated with carbonate species, such as the "free" ions Ca^{2+}, Mg^{2+}, and Al^{3+}.

In this process, lime and soda ash are added to solution, Ca^{2+} precipitates as $CaCO_3(\downarrow)$, and Mg precipitates as $Mg(OH)_2(\downarrow)$. To precipitate $CaCO_3$, the pH of the water must be raised to approximately 10.3. Similarly, $Mg(OH)_2(\downarrow)$ forms when the pH is raised above 11. When there is not enough naturally occurring bicarbonate alkalinity ($HCO_3{}^-$) in water to form the $CaCO_3$ precipitate, alkalinity must be added (typically as Na_2CO_3).

CHEMICAL REACTIONS OF LIME–SODA ASH SOFTENING

Initially, any free acids in the water must be neutralized prior to raising the pH for cation precipitation. In natural waters, which are generally considered in equilibrium with the atmosphere, CO_2 and H_2CO_3 (carbonic acid) are the primary free acids of concern (1).

Typically, hydrated lime, $Ca(OH)_2$, is used to neutralize the weak acid, as presented in Equations 1 and 2. When added to water, lime initially dissociates into its constituent ions, Ca^{2+} and OH^-. Free calcium forms, so water hardness is actually increased. However, as the pH is raised, free calcium ions are used to form $CaCO_3(\downarrow)$, and the hardness imparted to the water by lime dissociation is removed.

$$Ca(OH)_2 + H_2O \longleftrightarrow Ca^{2+} + 2OH^- + H_2O \qquad (1)$$

$$CO_2 + Ca(OH)_2 \longleftrightarrow CaCO_3(\downarrow) + H_2O \qquad (2)$$

Following the neutralization of free acids, carbonate hardness due to calcium is removed, as presented in Equation 3. When the pH is less than 8.3, bicarbonate is the dominant carbonate species that contributes to alkalinity; thus, the dominant calcium carbonate species is $Ca(HCO_3)_2$.

$$Ca^{2+} + 2HCO_3{}^- + Ca(OH)_2 \longleftrightarrow 2CaCO_3(\downarrow) + 2H_2O \qquad (3)$$

By adding hydrated lime, the pH is raised to ~10.3, which results in converting bicarbonate to carbonate ions and facilitates calcium precipitation as $CaCO_3$ (limestone).

Similarly, carbonate hardness associated with magnesium is removed by raising the pH to ~11 and precipitating magnesium as a hydroxide; however, the process occurs in two steps because $MgCO_3$ is soluble in water, as in Equations 4a and 4b.

$$Mg^{2+} + 2HCO_3{}^- + Ca(OH)_2 \longleftrightarrow MgCO_3 + CaCO_3(\downarrow) + 2H_2O \qquad (4a)$$

$$Mg^{2+} + CO_3{}^{2-} + Ca(OH)_2 \longleftrightarrow Mg(OH)_2(\downarrow) + CaCO_3(\downarrow) \qquad (4b)$$

At this point, the pH is ~11. Consequently, the pH does not need to be increased any further to enable the removal of noncarbonate hardness associated with calcium, Ca^{2+}. However, carbonate must be added to the system to form the calcium carbonate precipitate. Carbonate is added as soda ash, and the following reaction occurs:

$$Ca^{2+} + Na_2CO_3 \longleftrightarrow CaCO_3(\downarrow) + 2Na^+ \qquad (5)$$

Note that "free" Na^+ ions do not contribute to water hardness.

When removing noncarbonate hardness due to magnesium, both lime and soda ash are required; lime is added to provide the OH^- ions necessary to precipitate Mg as $Mg(OH)_2(\downarrow)$, and soda ash is needed to precipitate calcium.

$$Mg^{2+} + Ca(OH)_2 \longleftrightarrow Mg(OH)_2(\downarrow) + Ca^{2+} \qquad (6a)$$

$$Ca^{2+} + Na_2CO_3 \longleftrightarrow CaCO_3(\downarrow) + 2Na^+ \qquad (6b)$$

After removing Ca^{2+} and Mg^{2+} from solution, the pH is near 11, and $CO_2(g)$ is added to reduce the pH to ~9.2–9.7 and thus, remove excess lime.

$$CO_2 + Ca(OH)_2 \longleftrightarrow CaCO_3(\downarrow) + H_2O \qquad (7a)$$

$$Mg(OH)_2(\downarrow) + CO_2 \longleftrightarrow MgCO_3 + H_2O \qquad (7b)$$

$CO_2(g)$ is then added to reduce the water to a finished pH of ~8.6. (2,3).

Based on the relationships in Equations 1–7, the equivalent concentrations of lime and soda ash required are given by the following (2):

$$[Ca(OH)_2] = [CO_2] + [HCO_3{}^-] + [Mg^{2+}] + \text{excess} \qquad (8)$$

$$[Na_2(CO_3)] = [Ca^{2+}] + [Mg^{2+}] - \text{alkalinity} \qquad (9a)$$

$$\text{Alkalinity} = [HCO_3{}^-] + [CO_3{}^{2-}] + [OH^-] - [H^+] \qquad (9b)$$

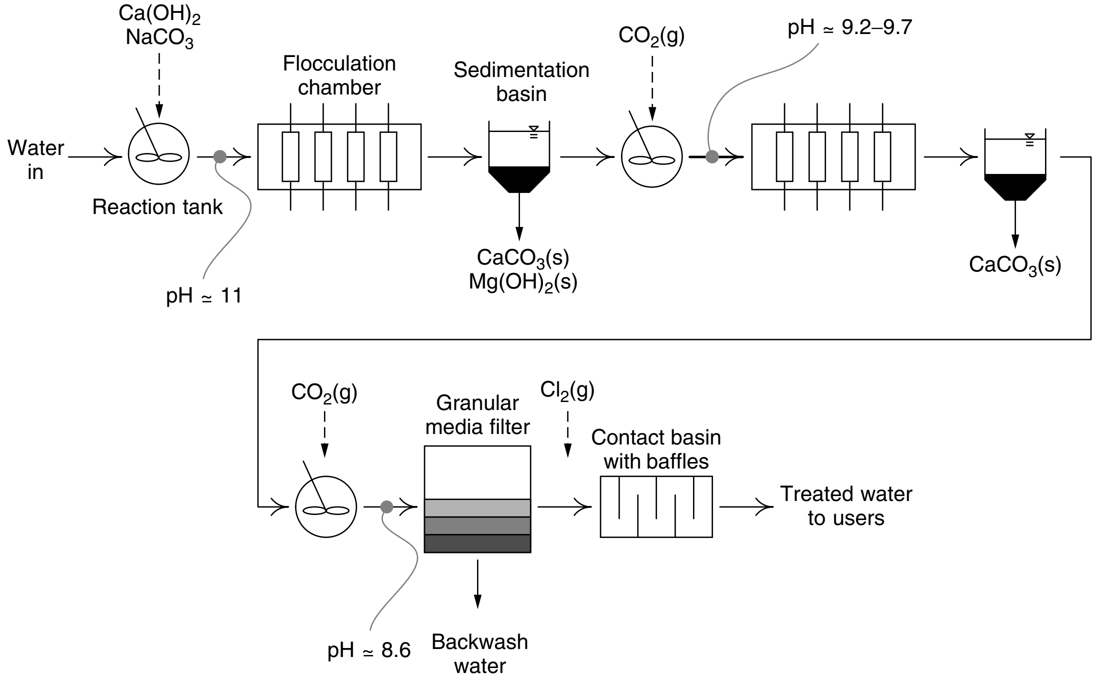

Figure 1. Schematic of typical layout of lime–soda ash softening process used in domestic water treatment.

PROCESS IMPLEMENTATION AND LIMITATIONS

The layout of a typical softening process in domestic water treatment is presented in (Fig. 1) (2). In this process, two flocculation basins and two sediment basins are required for the formation, precipitation, and subsequent removal of colloidal-particles. However, filtration is generally necessary after flocculation and sedimentation to remove the smallest particles that escape gravity settling. Due to the number of unit operations required in this process, capital costs can be high (2,3).

Further, Mg^{2+} is more expensive to remove than Ca^{2+}; thus, it is common to leave as much Mg in the water as feasible. Additionally, it is more costly to remove noncarbonate hardness than carbonate hardness because we must add $CO_3{}^{2-}$ (as Na_2CO_3; see Equation 6b) to facilitate the formation of the calcium carbonate precipitate. Consequently, as much noncarbonate hardness is left in the water as possible (2,3).

It is not possible to remove all Ca and Mg using lime–soda ash softening due to (1) the solubility of $CaCO_3(\downarrow)$ and $Mg(OH)_2(\downarrow)$, (2) limitations of mixing and contact, and (3) the lack of sufficient time for reaction kinetics to reach completion.

In general, the minimum calcium and magnesium hardnesses attainable in the lime–soda ash softening process are 30 and 10 mg/L as $CaCO_3$, respectively, though process economics is a major driving force in setting actual process end points (3). To attain treatment goals in a timely fashion, an excess of lime above stoichiometric requirements is supplied. Based on experience, a lime excess of ~20 mg/L, as $CaCO_3$, is typically recommended. It has been found that magnesium concentrations in excess of 40 mg/L, as $CaCO_3$, cause scaling in domestic hot water heaters. However, due to the expense of Mg removal, treatment is not generally carried out to achieve a Mg concentration much lower than 40 mg/L, as $CaCO_3$. When it is necessary to remove magnesium to between 20 and 40 mg/L, as $CaCO_3$, an additional lime excess must be added equivalent to the amount of magnesium removed. However, through experience, it has been found that the addition of lime in excess of 40 mg/L, as $CaCO_3$, does not provide appreciably improved reaction kinetics. As a result, the application of other advanced unit operations such as ion exchange would be needed to reduce water hardness further (2). Further detailed discussion of lime–soda ash water softening processes is presented by Reh (4).

BIBLIOGRAPHY

1. Stumm, W. and Morgan, J. (1996). *Aquatic Chemistry*, 3rd Edn. John Wiley & Sons, New York.
2. Tchobanoglous, G. and Schroeder, E. (1987). *Water Quality*. Addison-Wesley, Reading, MA.
3. Reynolds, T. and Richards, P. (1995). *Unit Operations and Processes In Environmental Engineering*, 2nd Edn. PSW Publishing, Boston, MA.
4. Reh, C. (1978). Lime soda softening processes. In: *Water Treatment Plant Design for the Practicing Engineer*. R. Sanks (Ed.). Ann Arbor Science, Ann Arbor, MI.

LIME SOFTENING

National Drinking Water
Clearinghouse

Hard water can cause scaling problems in water heaters and soap does not lather well in hard water. Therefore, some water utilities soften water to improve its quality for domestic use. Lime softening is best suited to groundwater sources, which have relatively stable water quality. The combination of variable source water quality and the complexity of the chemistry of lime softening may make lime softening too complex for small systems that use surface water sources.

Although lime softening has been used successfully by groundwater systems serving fewer than 3,000 people, it is unlikely to be suitable for treating groundwater in systems serving 500 or fewer people unless those systems have some form of contract or satellite operation that would enable a trained operator to monitor the treatment process. Prefabricated lime softening equipment is available for small systems. Also, there is an American Water Works Association Standard for quicklime and hydrated lime (ANSI/AWWA B202-93) that provides purchasers, manufacturers, and suppliers with the minimum requirements, including physical, chemical, packaging, shipping, and testing requirements.

Either hydrated lime [$Ca(OH)_2$] or quicklime (CaO) may be used in the softening process. The choice depends upon economic factors, such as the relative cost per ton of the two materials as well as the size and equipment of the softening plant. Hydrated lime is generally used more in smaller plants because it stores better and does not require slaking (producing a chemical change in lime by combining it with water) equipment. On the other hand, quick-lime costs less per ton of available calcium oxide and is thus more economical for use in large plants.

Softened water has high causticity and scale-formation potential; hence, recarbonation is employed to reduce pH and mitigate scaling of downstream processes and pipelines. Onsite combustion generation of carbon dioxide (CO_2) or liquid CO_2 is the most common source of carbon dioxide for recarbonation.

WHAT IS HARD WATER?

"Hardness" in water is primarily the result of concentrations of calcium and magnesium. Thus, some water utilities remove calcium and magnesium to soften the water and improve its quality for domestic use. Other ions that produce hardness include iron, manganese, strontium, barium, zinc, and aluminum, but these ions are generally not present in significant quantities. Therefore, total hardness is usually defined as the sum of magnesium and calcium hardness in milligrams per liter (mg/L), as calcium carbonate ($CaCO_3$). Total hardness can also be differentiated into carbonate and noncarbonate hardness. Carbonate hardness is the portion of total hardness present in the form of bicarbonate salts [$Ca(HCO_3)_2$ and $Mg(HCO_3)_2$] carbonate compounds ($CaCO_3$ and $MgCO_3$).

Noncarbonate hardness is the portion of calcium and magnesium present as noncarbonate salts, such as calcium sulfate ($CaSO_4$), calcium chloride ($CaCl_2$), magnesium sulfate ($MgSO_4$), and magnesium chloride (MgCl). The sum of carbonate and noncarbonate hardness equals total hardness.

What may be "hard" water in one area may be perfectly acceptable water in another area. In general, the degree of hardness is classified as follows:

Hardness	mg/L as $CaCO_3$
Soft	0 to 75
Moderate	75 to 150
Hard	150 to 300
Very hard	Above 300

Source: *Water Treatment Plant Design*, 1998.

For most applications, total hardness of 120 mg/L or less and magnesium hardness of 40 mg/L or less appear to be acceptable design criteria for softening facilities.

HOW DOES THE LIME-SOFTENING PROCESS WORK?

In the lime-softening process, the pH of the water being treated is raised sufficiently to precipitate calcium carbonate and, if necessary, magnesium hydroxide. The normal pH of water is between 6.5–8.5. In small systems, lime softening is typically practiced by adding hydrated lime to raw water to raise the pH to approximately 10. This removes calcium carbonate, essentially limestone. If magnesium removal is also required, the pH during softening needs to be closer to 11. Lime-softening systems need not be pilot tested for small systems using groundwater sources. Jar testing to determine appropriate process pH and chemical doses is sufficient. Doses of these chemicals should not change greatly over time unless the groundwater is subject to periodic infiltration by surface water that changes in quality. Lime-softening systems do need to be pilot tested if used on surface water sources with variable quality.

WHAT PRETREATMENT IS USED?

The principal types of pretreatment used before lime softening are aeration and presedimentation.

Aeration may be used to remove carbon dioxide from the source water before softening. This is only applicable to groundwaters where carbon dioxide concentrations are relatively high. Lime removal of carbon dioxide in source water adds to operation costs because of chemical expenses and increased calcium carbonate residuals.

Induced draft or open tray aeration is often used and may reduce the carbon dioxide level to 10 mg/L or less. Aeration also oxidizes iron and manganese; clogging of the aeration tray is a problem.

Presedimentation is used primarily by those plants treating high turbidity surface waters.

WHAT ARE THE MONITORING AND OPERATING REQUIREMENTS?

Regulatory monitoring requirements for lime softening plants depend on whether the source water is surface water or groundwater. Process monitoring requirements should focus on measurement of pH, hardness, and alkalinity for plants treating groundwater. In addition, filtered water turbidity monitoring is needed at the plants treating surface water for compliance purposes, as well as to manage filter operation.

One of the difficult aspects of lime softening is the operation and maintenance of lime feeders and lines carrying lime slurry to the point of application. In addition, plant operators must understand lime softening chemistry. Measurement of pH must be accurate, and the operator must know that the pH meter is properly calibrated. Failure to maintain the proper pH in softened water prior to filtration at a lime softening plant could result in precipitation or excess lime in the filter beds and formation of calcium carbonate deposits within the filters.

WHAT ARE THE CHEMICAL REQUIREMENTS?

The amount of lime required to remove carbonate hardness and magnesium can be calculated using in the following equation:

$$\mathrm{CaO(lb/mil\ gal)} = 10.6\ \mathrm{CO_2(mg/L)} + 4.7[\mathrm{alkalinity(mg/L)} + \mathrm{magnesium\ hardness(mg/L)} + X]$$

Where CaO is 100 percent pure, CO_2 is expressed as CO_2, alkalinity is expressed as $CaCO_3$, and X is the required excess hydroxide alkalinity in mg/L as $CaCO_3$. The magnesium hardness shown is the amount to be removed by softening, and not the amount present. Desired excess alkalinity can be determined from the magnesium hydroxide solubility relationship: it is typically in the range of 30 to 70 mg/l and is often estimated at 50 mg/L expressed as $CaCO_3$.

WHAT ARE OTHER SOFTENING ALTERNATIVES?

The selection of lime, lime-soda ash, or caustic soda softening is based on cost, total dissolved solids criteria, sludge production, carbonate and noncarbonate hardness, and chemical stability. Water containing little or no noncarbonate hardness can be softened with lime alone. However, water with high noncarbonate hardness may require both lime and soda ash to achieve the desired finished water hardness. Softening with lime or lime-soda ash is generally less expensive than caustic softening. Caustic soda softening increases the total dissolved solids of treated water, while lime and lime-soda ash softening often decrease total dissolved solids. Caustic soda softening produces less sludge than lime and lime-soda ash softening. Caustic soda does not deteriorate during storage, while hydrated lime may absorb carbon dioxide and water during storage, and quicklime may slake in storage causing feeding problems. The final selection is

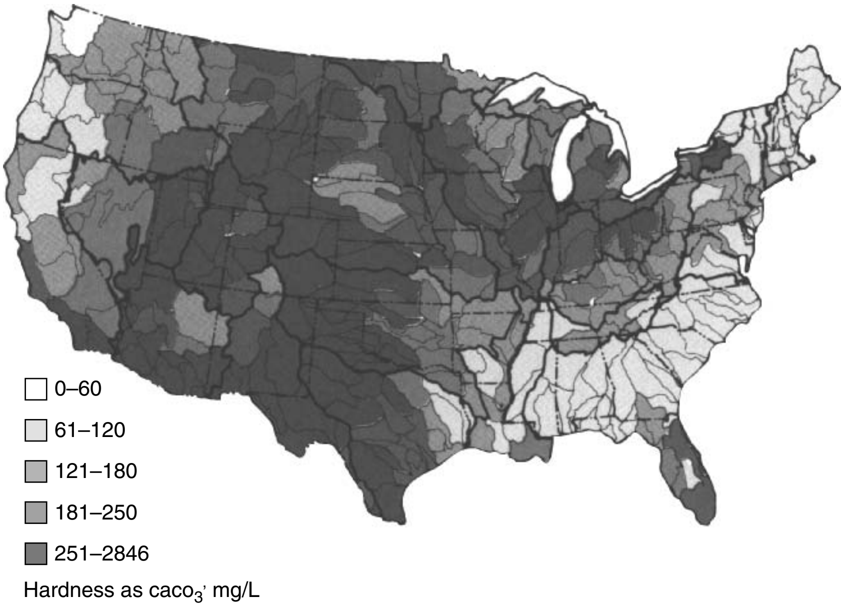

Figure 1. Average water hardness in the continental united states. Courtesy of the U.S. Geological Survey, Report 78-200.

generally based on cost, water quality, and owner and operator preference.

WHAT ARE THE BENEFITS OF SOFTENING?

Potential benefits of softening water at a central treatment plant include the following:

- reducing dissolved minerals and scale-forming tendencies,
- reducing consumption of household cleaning agents,
- removing radium 226 and 228,
- removing arsenic and uranium,
- removing heavy metals,
- supplementing disinfection and reducing algal growths in basins,
- removing certain organic compounds and reducing total organic carbon (TOC),
- removing silica and fluoride,
- removing iron and manganese,
- reducing turbidity of surface waters in conjunction with the hardness precipitation process,
- increasing the Langelier Saturation Index, useful for corrosion control under some conditions, and
- possibly removes *Giardia Lamblia* cysts.

ARE THERE CONCERNS ABOUT SOFT WATER?

In most cases, for consumer use, relatively soft water is preferable; however, it does have several disadvantages. Excessively soft water can cause corrosion in pipes. This corrosion can shorten the service life of pipes and household appliances and can result in toxic materials, such as lead and cadmium, being dissolved in drinking water. The decision to soften a water supply depends completely on the community. Drinking water regulations do not generally require softening.

WHAT ABOUT LIME SOFTENING RESIDUALS?

All water softening plants, large or small, are now required to refrain from directly discharging any wastes—liquid or solid—into rivers or streams. Where abandoned stone quarries, sand and gravel pits, or coal mines are available nearby, it may be possible to dispose of the sludge for many years without any serious difficulty. In areas where there are abandoned coal-strip mines, pumping the sludge into these areas may be a logical and convenient solution. If acid waters are leaching out of the mine into the watershed, as many such mines do, the waste sludge can help in abating serious stream pollution by neutralizing the acid water. Such neutralization may offer a great potential savings in material treatment cost to the state.

Selection of lime, lime-soda ash, or caustic soda chemical precipitation process must adequately address the disposal of generated sludges. Ultimate disposal of lime or caustic sludges now includes options, such as discharge to sanitary sewers, drying lagoons, and land application.

In some cases, sludge is discharged directly into a community's sewage system. Such a discharge must be done with the approval of the municipal sewage department, since not all systems can accommodate this alkaline waste, at least in the volume produced. In other instances, the discharge of such sludge could prove beneficial in neutralizing other acid trade wastes entering the same system. A few small plants are permitted to discharge directly into the community's storm sewer system, although this method has been phased out in most cases.

Lagooning is another practical solution. If the sludge can be dried to approximately 50 percent moisture content in lagoons, the annual requirement will be approximately

0.5 to 1.0 acre-feet per 100 mg/L hardness removed for every million gallons per day (MGD). Also, multiple lagoons used on a fill and let dry sequence, or filling from one end and withdrawing or decanting from the other, often result in better drying of the sludge.

Increasing ingenuity is necessary for sludge disposal, since lagooning, one of the most common practices, is becoming less frequent due to the dearth of available land and soaring land costs. Therefore, other methods are being increasingly employed.

An attractive method for a water plant located in a farming area is to use land application of the sludge on farm fields where the soil pH is too low for optimal plant growth. In addition to providing a desirable pH for plant growth, lime-softening sludge is generally a pure source of calcium carbonate with varying small amounts of magnesium hydroxide and provides an effective source of liming material for farmers. It has a neutralizing value for soil acids in excess of most agricultural liming materials.

WHERE CAN I FIND MORE INFORMATION?

Information in this fact sheet was primarily obtained from the following sources:

(1) Pizzi, Nicholas G. 1995. *Hoover's Water Supply and Treatment*. Twelfth Edition by Bulletin 211. National Lime Association, Kendall/Hunt Publishing Company, Arlington, VA 22203.

(2) American Water Works Association. 1984. *Introduction to Water Treatment. Principles and Practices of Water Supply Operations*.

(3) National Lime Association at http://www.lime.org/.

(4) National Research Council. 1997. *Safe Water From Every Tap: Improving Water Service to Small Communities*. National Academy Press. Washington D.C.

(5) *Technology Transfer Handbook: Management of Water Treatment Plant Residuals*. American Association of Civil Engineers Manuals and Reports on Engineering Practice No. 88, AWWA Technology Transfer Handbook, U.S. EPA/625/R-95/008.

(6) American Water Works Association. 1990. *Water Quality and Treatment: A Handbook of Community Water Supplies*. Fourth Edition.

(7) American Water Works Association and American Society of Civil Engineers. 1998. *Water Treatment Plant Design*. McGraw-Hill, New York, NY.

(8) Vigneswaran, S., C. Visvanathan. 1995. *Water Treatment Processes: Simple Options*. CRC Press. New York, NY.

Also, the National Drinking Water Clearinghouse (NDWC) Registry of Equipment Suppliers of Treatment Technologies for Small Systems (RESULTS) is a public reference database that contains information about technologies in use at small water systems around the country.

For further information, call the NDWC at (800) 624-8301 or (304) 293-4191. Additional free copies of Tech Brief fact sheets are available from the NDWC at the above numbers or you may download Tech Briefs from our Web site at http://www.ndwc.wvu.edu. You may also order via e-mail at ndwc_orders@ndwc.wvu.edu.

Tech Brief: Disinfection, *item #DWBLPE47*
Tech Brief: Filtration, *item #DWBLPE50*
Tech Brief: Corrosion Control, *item #DWBLPE52*
Tech Brief: Ion Exchange and Demineralization, *item #DWBLPE56*
Tech Brief: Organics Removal, *item #DWBLPE59*
Tech Brief: Package Plants, *item #DWBLPE63*
Tech Brief: Management of Water Treatment Plant Residuals, *item #DWBLPE65*
Tech Brief: Lime Softening, *item #DWBLPE67*

ION EXCHANGE—USE OF MAGNETIC ION EXCHANGE RESIN FOR DOC REMOVAL

MARIN SLUNJSKI
Orica Watercare
Regency Park, SA, Australia

BACKGROUND

Conventional treatment by flocculation has a low efficiency of around 30% for DOC removal, although this can be increased to 50% by enhanced coagulation. Alternative treatment processes more effective at DOC removal include granular-activated carbon (GAC) adsorption and membrane filtration technologies, specifically nanofiltration. The latter process requires a pretreatment stage to remove suspended solids prior to membrane filtration.

ION EXCHANGE FOR DOC REMOVAL

Much of the DOC is acidic and amenable to removal by ion exchange. The DOC molecules are large (apparent molecular weight from 0.5 kD to more than 10 kD) and have a high negative charge density (10–15 meq/g DOC).

MIEX® resin is very small, five times smaller than conventional resins. This allows it to be easily dispersed in water. It is a macroporous, strong base anion exchange resin that, together with its small size, facilitates rapid DOC exchange during both DOC removal from water and resin regeneration.

The reaction mechanism is shown in Fig. 1. The resin is suspended in water containing low concentrations of DOC and other inorganic anions. The resin's affinity for the DOC molecules shifts the equilibrium to DOC uptake. The reaction is very efficiently reversed by treating the resin with a solution of high chloride concentration, such as brine.

This resin is specially formulated to include a magnetic component that allows individual resin beads to interact like small magnets, a feature that is the basis for the unique handling characteristics of the MIEX® resin and for the innovative engineering of this ion exchange process.

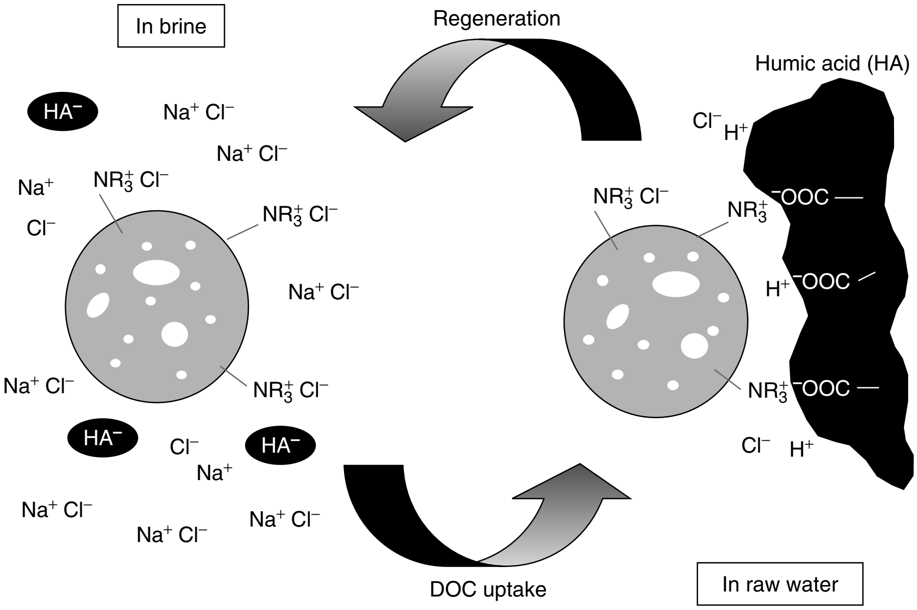

Figure 1. Standard DOC exchange mechanism.

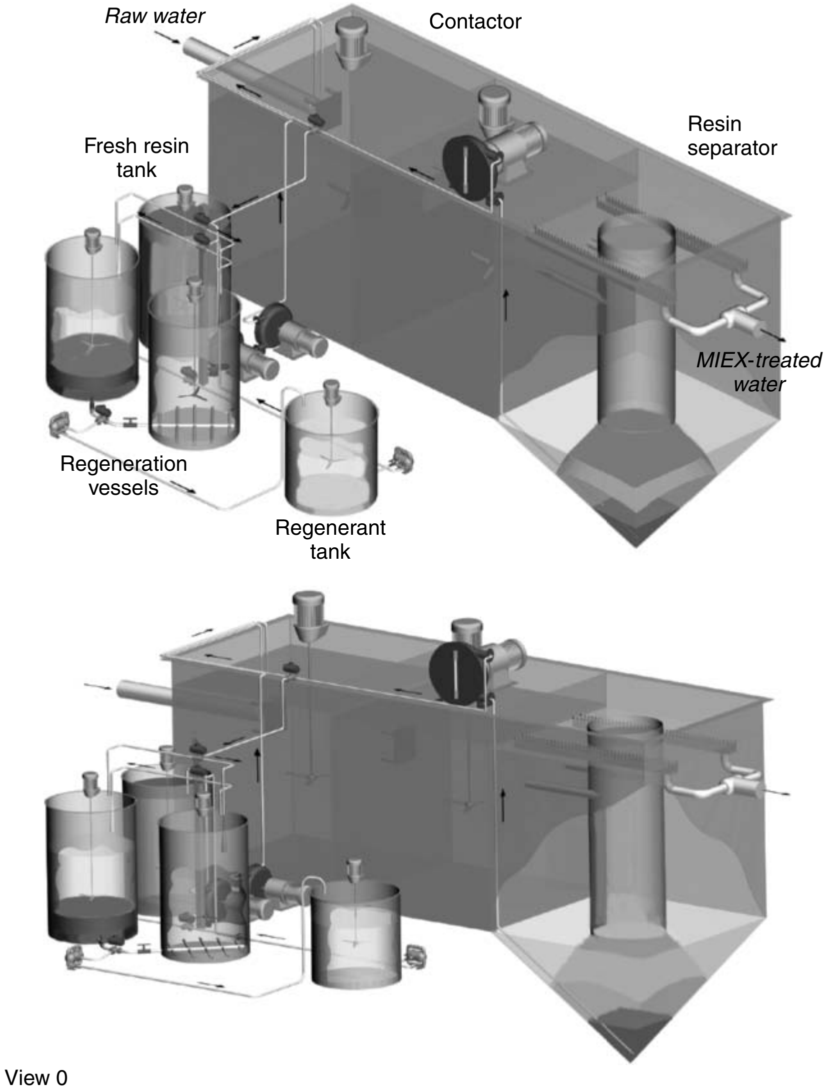

Figure 2. Diagram of 1-mL/day (0.265 mgd) MIEX® resin plant.

MIEX® RESIN IN DRINKING WATER TREATMENT

The MIEX® resin treatment process employs conventional water treatment plant facilities, including a stirred resin contact tank and gravity settlers, as shown in Fig. 2. The resin is dosed into raw/untreated water at a concentration required to achieve a specific water quality. Resin concentrations are in the range of 5–20-mL settled resin per L water (1.25–5.0-g dry weight resin per L). The retention time in the flow-through contacting tanks ranges from 10 to 30 minutes, which is sufficient to achieve more than 50% DOC removal.

The suspension of resin in water from the contacting tanks flows by gravity into a separator. The separator feed well is designed to provide conditions that promote agglomeration of individual resin beads under their magnetic forces. The agglomerates formed are heavy enough to settle to the bottom of the separator against water upflow rates between 5 and 15 $m^3/m^2/h$ (2–6 gpm/ft^2). In this process, more than 99.95% of the resin fed into the separator is recovered in the concentrated (~300 mL-settled resin per L) separator underflow referred to as resin recycle.

In addition to keeping the plant footprint small, the high upflow rate in the resin separator is required to flush out effectively turbidity introduced into the MIEX® plant with raw water. Because a small amount of the resin attrited by pumping and mixing is carried over from the separator with treated water, turbidity increases through the MIEX® process by about 1–2 NTU.

The resin recycle from the separator is directed back to the contacting tanks for another DOC uptake cycle. To maintain the resin's capacity for DOC removal, a small amount of the recycled resin, typically 10% of the total resin flow, is continuously diverted to resin regeneration and is replaced by the same amount of fresh (regenerated) resin. In this way, each resin bead is, on average, contacted by water 10–20 times before being diverted to regeneration.

The DOC-laden resin is stored in a regeneration vessel until a preset settled resin volume is reached. This triggers switching over the resin off-take to the second regeneration

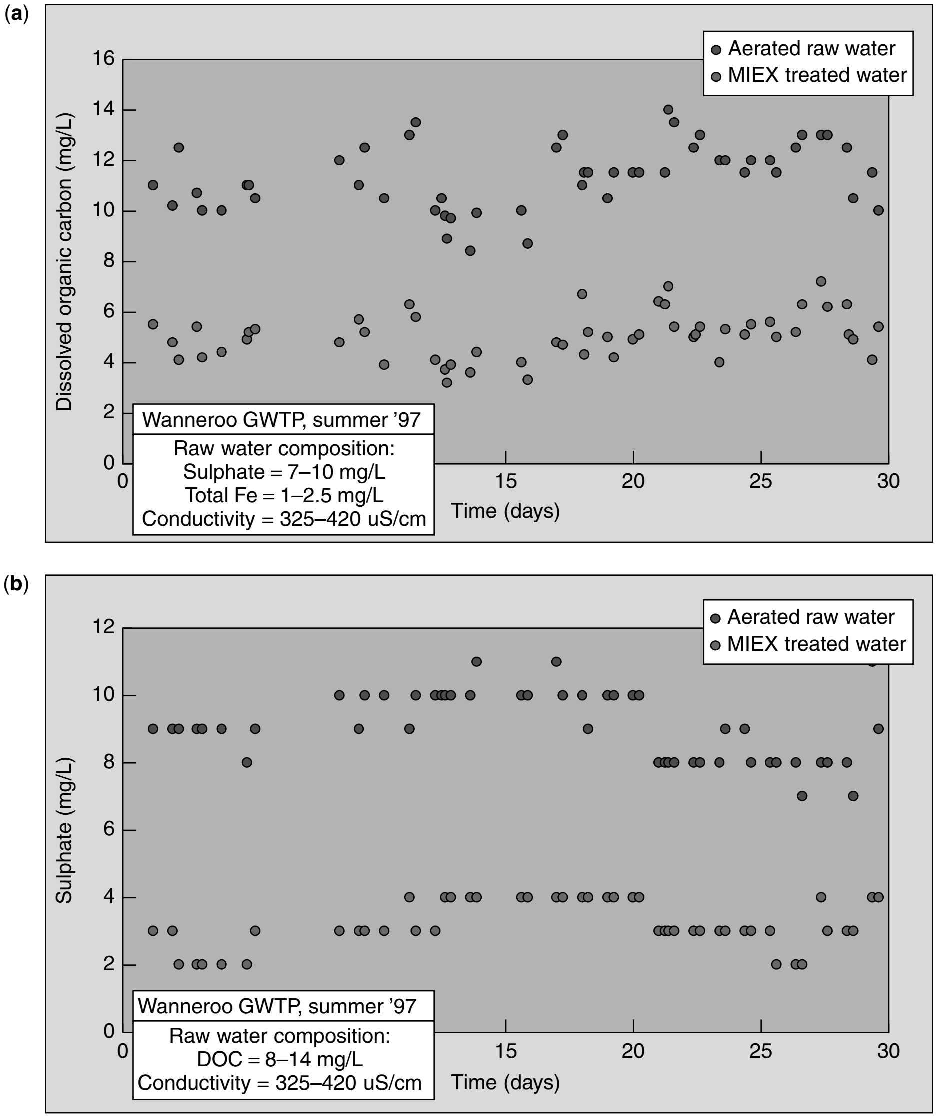

Figure 3. DOC and sulfate removal results.

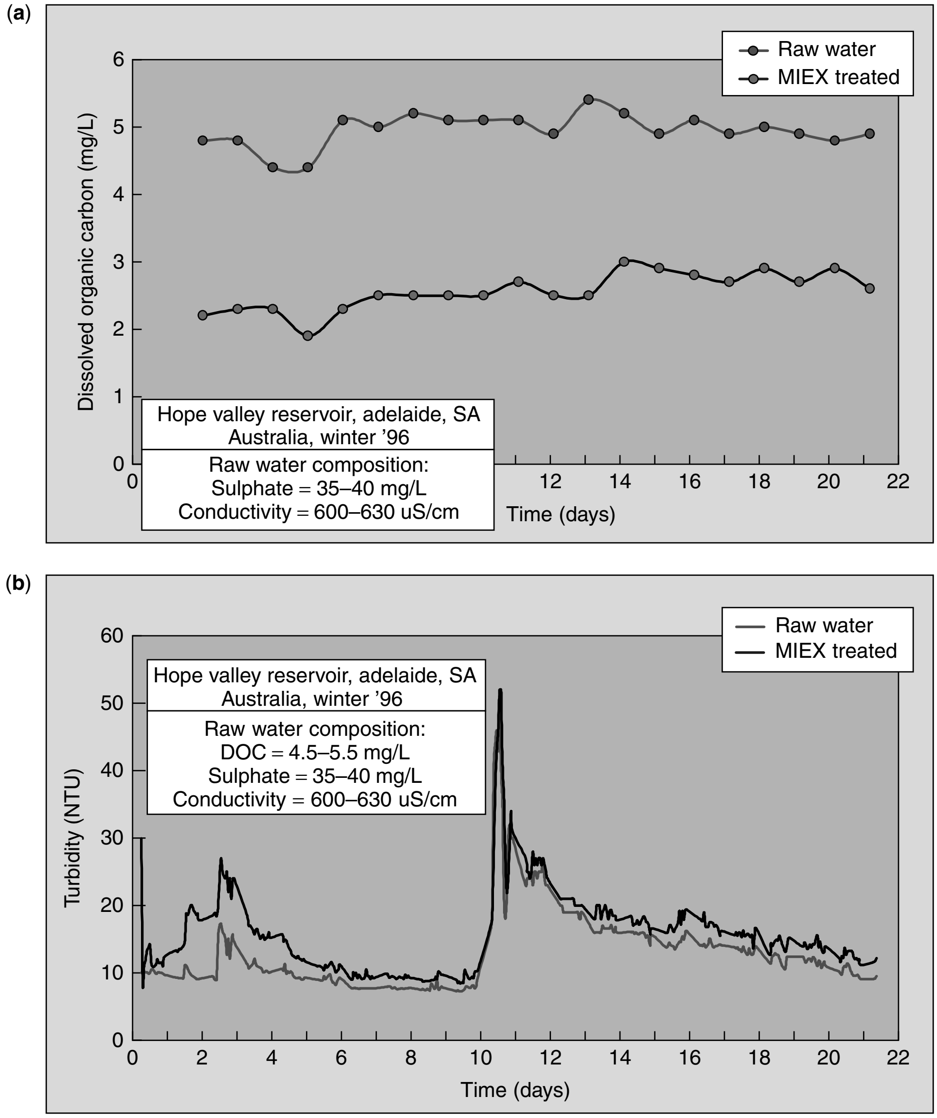

Figure 4. Effect of turbidity on MIEX® resin DOC removal.

vessel, and the start of regeneration in the full vessel. The regeneration is a sequential process comprising draining of carrier water, mixing resin with regenerant, draining of used regenerant, optional rinsing of regenerated resin with water, resuspending of regenerated resin in water, and transferring it into the fresh resin tank.

The basic regenerant is 120-g/L NaCl solution dosed at 3:1 brine to resin volume. Depending on the raw water quality, in some cases, pH modified regenerations using small additions of NaOH and HCl are beneficial. The consumption of salt for regeneration and waste regenerant production is minimized by reusing the regenerant a number of times. Salt usage is from 50-kg to 80-kg NaCl per mL water treated depending on conditions, and salt is adjusted after each use.

RESULTS OF MIEX® TREATMENT

Due to the steady-state nature of operation, this process produces water of constant quality. The high resin preference for DOC anions removes no other anions (bicarbonate, nitrate, etc.), except for sulfate removed in appreciable quantities (Fig. 3).

DOC removal is undisturbed by excursions of raw water turbidity, as shown in Fig. 4. This process is not designed to remove turbidity and may actually contribute 1–2 NTU due to a small amount of resin carryover (Fig. 4b).

Under typical process conditions, MIEX® resin preferentially removes the smaller molecular weight (MW) fraction of DOC. This makes it complementary to alum coagulation, which is known to be more efficient in removing the larger MW fraction of DOC, as shown in Fig. 5.

Pretreatment with MIEX® resin usually results in 50–60% DOC removal. This in turn results in even higher reductions in the alum dose required for turbidity removal. Alum used for turbidity removal will remove a further 20–30% DOC. The combined MIEX®/low-dose alum treatment can remove an overall 80–90% DOC. The low residual DOC levels result in a 50–60% reduction in chlorine demand compared with using

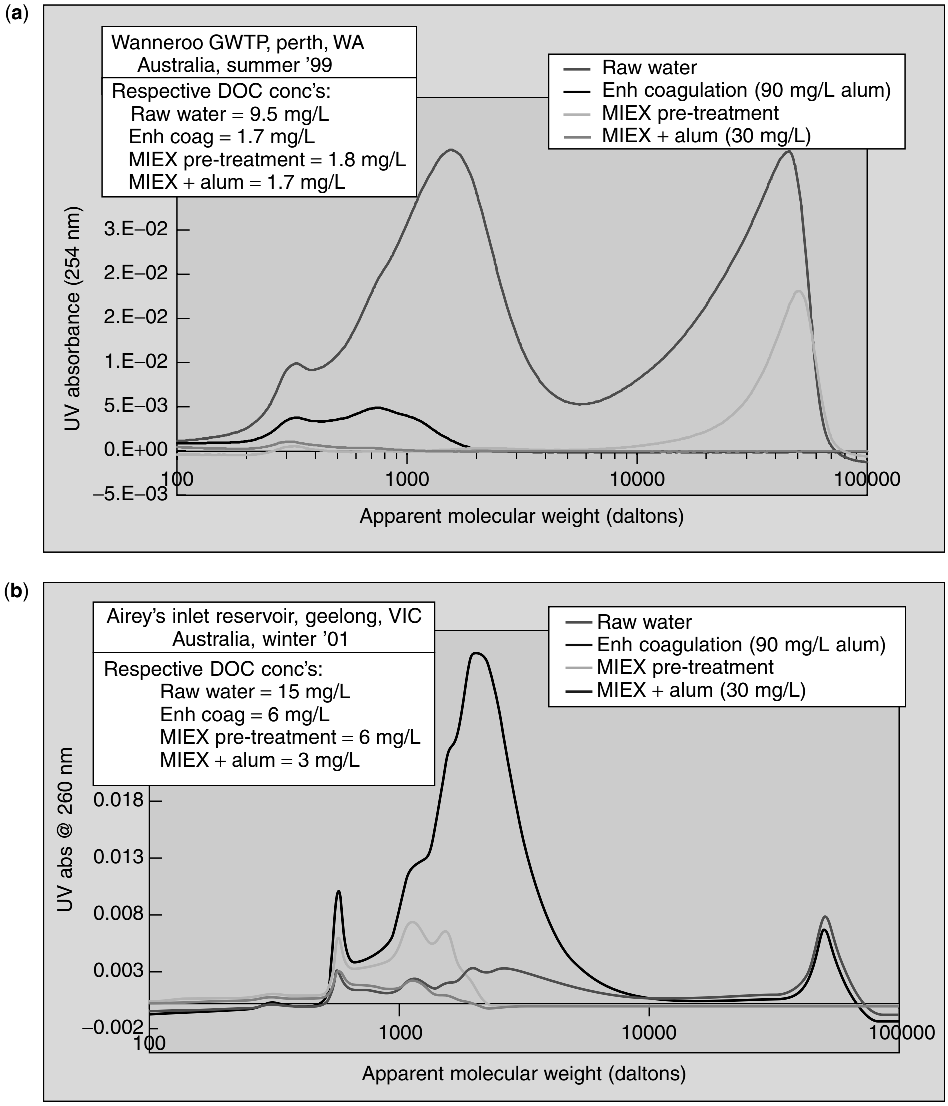

Figure 5. Effect of MIEX® on NOM's apparent molecular weight.

only conventional alum treatment. However, the main benefit of more efficient DOC removal from the combined MIEX®/coagulation process is the reduction in disinfection byproducts formed (Fig. 6).

PROCESS CAPITAL AND OPERATING COSTS

The MIEX® plant uses conventional water treatment infrastructure and equipment. There is no need for pressure columns or complicated piping/valve arrangements. Due to the high settler loading rate, the plant footprint is small, and the production rate high—the average water retention time in the plant is in the range of 40 min to 1 h.

Furthermore, the plant requires only 10% of the resin inventory compared with a conventional system based on the 10–15 min empty EBCT[3] generally recommended for DOC removal (1). This results in further reductions in plant capital costs. There is also no need for periodical replacement of the complete resin inventory, as it is continuously refreshed with small additions of makeup resin.

The main operating cost is resin makeup, the replacement of resin carried over from the separator. The actual resin loss is a physical parameter determined by the plant design and operating conditions. At a separator rise rate of 7 m/h, the loss rate is estimated at 4.5-L settled resin per mL treated water (4.5 gal/mg).

Another significant cost is salt for resin regeneration. Depending on the regeneration process design and raw water quality, the actual salt consumption ranges between 50- and 300-kg NaCl per mL treated water (420–2500 lb/mg). Because only small amounts of NaOH and HCl are used for occasional pH modified regenerations, these chemical costs are minor. The MIEX® plant power consumption is about 50 kW per mL treated water (190 kW/MG).

Finally, there may be some costs for disposal of spent regenerant. This waste stream can be as small as 50–200 L per mL water treated (gal/mg) based on a

---

[3] EBCT—empty bed contact time: time to treat the amount of water equal to the resin bed volume.

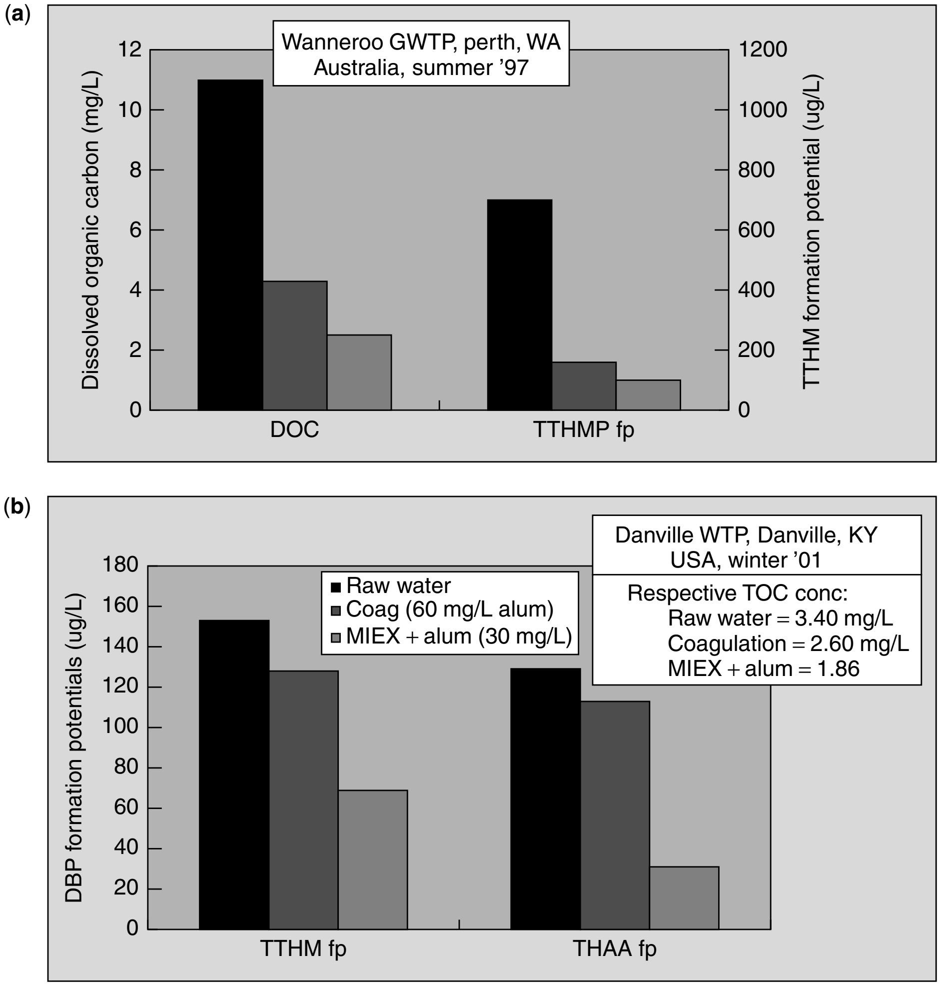

Figure 6. Effect of MIEX® on disinfection byproduct formed.

spent brine concentration of 30-g/L DOC and depending on the amount of DOC removed from raw water by the MIEX® resin. There are a number of disposal options available—sewering, deep well injection, deep ocean outfalls, and evaporation, depending on a particular plant site.

In conclusion, the process based on the MIEX® resin represents a simple, economical option for upgrading water treatment plants to improve removal of DOC and disinfection byproduct precursors. The process can easily be combined with an existing coagulation treatment plant either as an add-on greenfield plant or retrofitted into the existing infrastructure.

It is also possible to combine the MIEX® resin process as the single color/DOC removing step followed by turbidity removal using microfiltration (1,2). This solution is particularly attractive as a "chemical-free" water treatment process.

BIBLIOGRAPHY

1. Fettig, J. (1999). Removal of humic substances by adsorption/ion exchange. In: *IWSA/IAWQ Conference Proceedings: Removal of Humic substances from Water*, H. Ødegaard (Ed.). Trondheim, Norway, June 24–26, pp. 185–194.
2. Morran, J., Drikas, M., Hepplewhite, C., and Pelekani, C. (2001). *MIEX® and Microfiltration—A Winning Alliance*, in Conference Proceedings OzWater 2001, April 1–4, Canberra, Australia.
3. Bourke, M., Harrison, S., Long, B., and Lebeau, T. (2001). *MIEX® Resin Pretreatment Followed by Microfiltration as an Alternative to Nanofiltration for DBP Precursor Removal*, in AWWA 2001 Membrane Technology Conference Proceedings, March 4–7, San Antonio, TX.

READING LIST

Ødegaard, H., Eikebrokk, B., and Storhaug, R. (1999). Processes for the Removal of Humic Substances from Water—An Overview Based on Norwegian Experience. In: *IWSA/IAWQ Conference Proceedings: Removal of Humic substances from Water*. H. Ødegaard (Ed.). June 24–26, Trondheim, Norway, pp. 43–52.

Potts, J., Wessels, L.P., and Van Paassen, W.C. (2000). Ion exchange for colour and total carbon removal in combination with dead-end nano-filtration for reuse of regenerate. In: *Proceedings from the Conference on Membranes in Drinking and Industrial Water Production*, Oct. 3–6, Paris, France, published by Deasalination Publications, L'Aquila, Italy, Vol. 1, pp. 241–246.

MEMBRANE FILTRATION

National Drinking Water
Clearinghouse

A membrane or, more properly, a semipermeable membrane, is a thin layer of material capable of separating substances when a driving force is applied across the membrane.

Once considered a viable technology only for desalination, membrane processes are increasingly employed for removal of bacteria and other microorganisms, particulate material, and natural organic material, which can impart color, tastes, and odors to the water and react with disinfectants to form disinfection byproducts (DBP). As advancements are made in membrane production and module design, capital and operating costs continue to decline.

The pressure-driven membrane processes discussed in this fact sheet are microfiltration (MF), ultrafiltration (UF), nanofiltration (NF), and reverse osmosis (RO).

MEMBRANE FILTRATION: ALTERNATIVE TO CONVENTIONAL FILTRATION

Membrane filtration systems' capital costs, on a basis of dollars per volume of installed treatment capacity, do not escalate rapidly as plant size decreases. This factor makes membranes quite attractive for small systems. In addition, for groundwater sources that do not need pretreatment, membrane technologies are relatively simple to install, and the systems require little more than a feed pump, a cleaning pump, the membrane modules, and some holding tanks. According to a 1997 report by the National Research Council, most experts foresee that membrane filtration will be used with greater frequency in small systems as the complexity of conventional treatment processes for small systems increases.

NEW REGULATIONS FAVOR MEMBRANE TECHNOLOGIES

Membrane processes have become more attractive for potable water production in recent years due to the increased stringency of drinking water regulations. Membrane processes have excellent separation capabilities and show promise for meeting many of the existing and anticipated drinking water standards. The Surface Water Treatment Rule (SWTR) and the anticipated Groundwater Disinfection Rule have led to the investigation of UF and MF for turbidity and microbial removal. The new Disinfectants/Disinfection Byproduct (D/DBP) rules have increased interest in NF and UF membranes for DBP precursor removal.

Potable water treatment has traditionally focused on processes for liquid-solid separation rather than on processes for removing dissolved contaminants from water. Thus, the effect of the 1996 Safe Drinking Water Act (SDWA) amendments has been to encourage water treatment professionals to consider the more unconventional treatment processes, such as membrane technologies, alone, or in conjunction with liquid-solid separation, to meet current regulations.

COMPARING MEMBRANE FILTRATION SYSTEMS

While all types of membranes work well under proper conditions, choosing the most appropriate membrane for a given application still remains crucial (see Fig. 1). In many cases, selection is complicated by the availability of new types of membranes, applications, or by site-specific conditions. Bench and pilot tests are powerful tools for situations where process risks and uncertainties exist or the cost impacts from problems are potentially high.

Membrane classification standards vary considerably from one filter supplier to another. What one supplier sells as a UF product, another manufacturer calls a NF system. It is better to look directly at pore size, molecular weight cutoff (MWCO), and applied pressure needed when comparing two membrane systems.

MWCO, which can be regarded as a measure of membrane pore dimensions, is a specification used by membrane suppliers to describe a membrane's retention capabilities.

MICROFILTRATION (MF)

MF is loosely defined as a membrane separation process using membranes with a pore size of approximately 0.03 to 10 microns, a MWCO of greater than 100,000 daltons, and a relatively low feedwater operating pressure of approximately 100 to 400 kPa (15 to 60 psi). Representative materials removed by MF include sand, silt, clays, *Giardia lamblia* and *Cryptosporidium* cysts, algae, and some bacterial species (see Fig. 2 and Table 1). MF is not an absolute barrier to viruses; however, when used in combination with disinfection, MF appears to control these microorganisms in water.

The primary impetus for the more widespread use of MF has been the increasingly stringent requirements for removing particles and micro-organisms from drinking water supplies. Additionally, there is a growing emphasis on limiting the concentrations and number of chemicals that are applied during water treatment. By physically removing the pathogens, membrane filtration can significantly reduce chemical addition, such as chlorination.

Another application for the technology is for removal of natural or synthetic organic matter to reduce fouling potential. In its normal operation, MF removes little or no organic matter; however, when pretreatment is applied, in-creased removal of organic material, as well as a retardation of membrane fouling can be realized.

Two other applications involve using MF as a pretreatment to RO or NF to reduce fouling potential. Both RO and NF have been traditionally employed to desalt or remove hardness from groundwater.

Process

MF membranes provide absolute removal of particulate contaminants from a feed stream by separation based

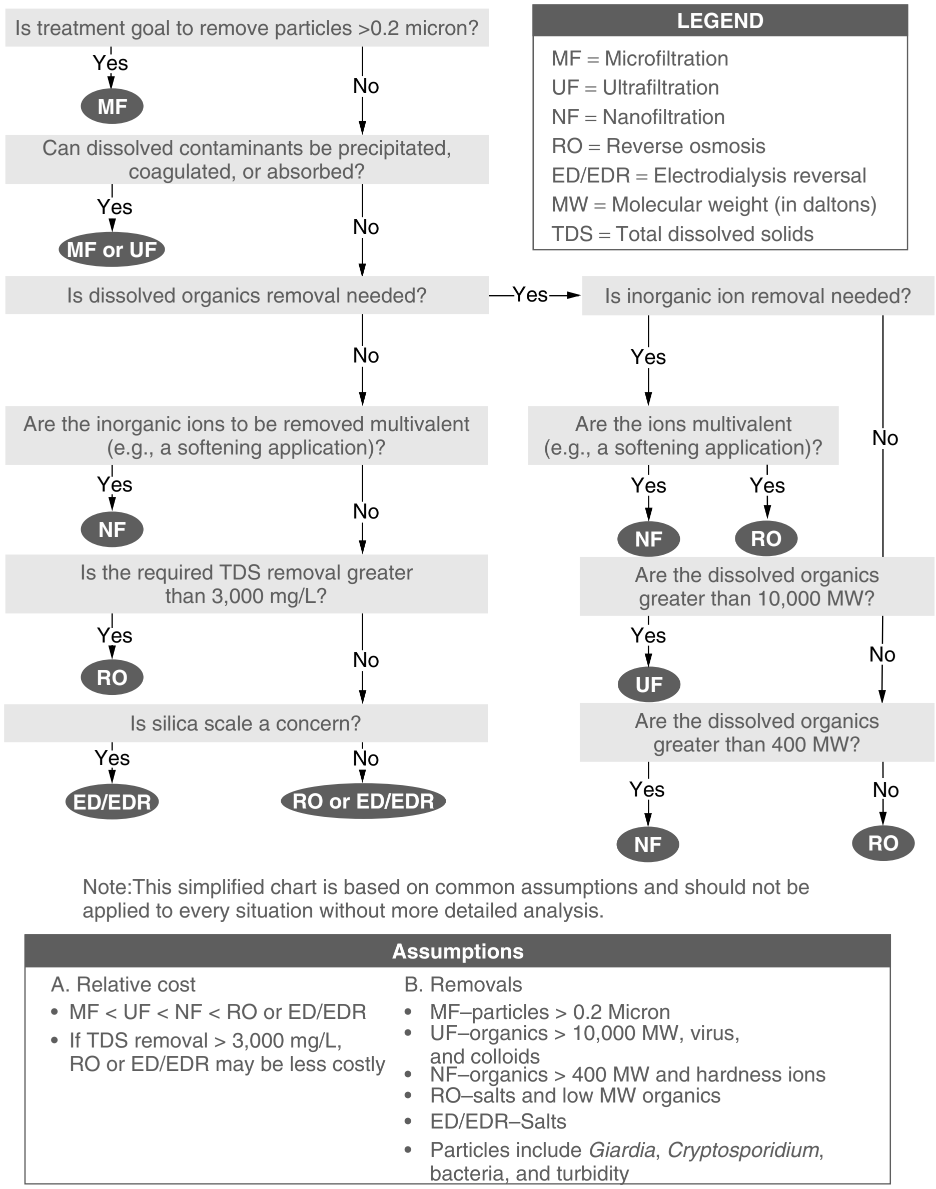

Figure 1. Generalized membrane process selection chart. *Reprinted from* Proceedings of the 1993 Membrane Technology Conference, *by permission.Copyright* © 1993, *American Water Works Association*.

on retention of contaminants on a membrane surface. It is the "loosest" of the membrane processes, and as a consequence of its large pore size, it is used primarily for removing particles and microbes and can be operated under ultralow pressure conditions.

In the simplest designs, the MF process involves prescreening raw water and pumping it under pressure onto a membrane. In comparison to conventional water clarification processes, where coagulants and other chemicals are added to the water before filtration, there are few pretreatment requirements for hollow-fiber systems when particles and microorganisms are the target contaminants.

Prefilters are necessary to remove large particles that may plug the inlet to the fibers within the membrane module. More complex pretreatment strategies are sometimes employed either to reduce fouling or enhance the removal of viruses and dissolved organic matter. In such cases, pretreatment by adding coagulants or powdered activated carbon (PAC), has been employed. In some cases, the cake layer built up on the membrane during the water production cycle can remove some organic materials.

It may be necessary to adjust the feedwater pH by chemical dosing prior to membrane filtration in order to maintain the pH within the recommended operating range for the membrane material employed. It should be noted that pH adjustment is not required for scaling control, since MF membranes do not remove uncomplexed dissolved ions.

MF membranes, under the most conservative conditions, appear to act as an absolute barrier to selected bacteria and protozoan cysts and oocysts. Unlike UF however, MF does not remove appreciable densities of viruses. Therefore, it is necessary to complement MF with a postmembrane disinfection process. Chemical disinfection may

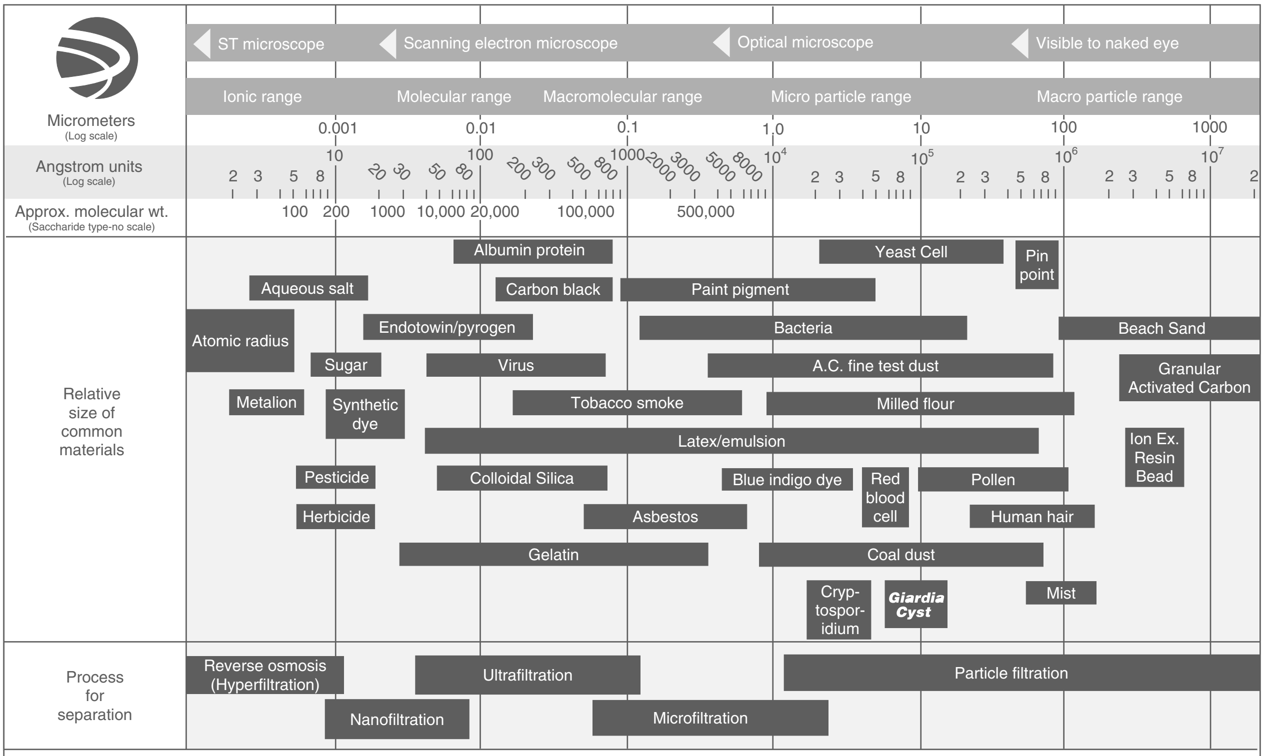

Note: I Micron (1×10^{-6} meters) ≈ 4×10^{-5} Inches (0.00004 inches)
I Angstrom unit = 10^{-10} Meters = 10^{-4} Micrometers (microns)

http://www.osmonics.com

© Copyright 1998, 1996, 1993, 1990, 1984 Osmonics, Inc., Minnetonka, *MN USA*

Figure 2. The filtration spectrum. *Reprinted courtesy of Osmonics, Inc., Minnetonka, MN.*

Table 1. Surface Water Treatment Compliance Technology: Membrane Filtration

Unit Technologies	Removals: Log Giardia & Log Virus	Raw Water, Pretreatment & Other Water Quality Issues
Microfiltration (MF)	Very effective Giardia, >5–6 log; Partial removal of viruses (disinfect for virus credit).	High quality or pretreatment required. Same note regarding TOC.
Ultrafiltration (UF)	Very effective Giardia, >5–6 log; Partial removal of viruses (disinfect for virus credit).	High quality or pretreatment required (e.g., MF). TOC rejection generally low, so if DBP precursors are a concern, NF may be preferable.
Nanofiltration (NF)	Very effective, absolute barrier (cysts and viruses).	Very high quality or pretreatment required (e.g., MF or UF to reduce fouling/extend cleaning intervals). See also RO pretreatments, below.
Reverse Osmosis (RO)	Very effective, absolute barrier (cysts and viruses).	May require conventional or other pretreatment for surface water to protect membrane surfaces: may include turbidity or Fe/Mn removal; stabilization to prevent scaling; reduction of dissolved solids or hardness; pH adjustment.

Source: U.S. Environmental Protection Agency, 1998.

be employed by applying chlorine, chlorine dioxide, or chloramines; however, long contact times are required to inactivate viruses.

Equipment

For municipal-scale drinking water applications, the commercially available membrane geometries that are the most commonly employed are spiral wound, tubular, and hollow capillary fiber. However, spiral-wound configurations are not normally employed for MF due to the flat-sheet nature of the membrane, which presents difficulties in keeping the membrane surface clean. Unlike spiral-wound membranes, hollow-fiber and tubular configurations allow the membrane to be backwashed, a process by which fouling due to particulate and organic materials is controlled.

Membrane "package" plants are normally employed for plants treating less than one million gallons per day (mgd). The components of the plant may include prescreens, a feed pump, a cleaning tank, an automatic gas backwash system, an air compressor, a membrane integrity monitor, a backwash water transfer tank, a pressure break reservoir, an air filter for the gas backwash, controls for the programmable logic controller, and a coalescer.

Operation and Maintenance

In MF, there are two methods for maintaining or re-establishing permeate flux after the membranes are fouled:

- Membrane backwashing: In order to prevent the continuous accumulation of solids on the membrane surface, the membrane is backwashed. Unlike backwashing for conventional media filtration, the backwashing cycle takes only a few minutes. Both liquid and gas backwashing are employed with MF technology. For most systems, backwashing is fully automatic. If backwashing is incapable of restoring the flux, then membranes are chemically cleaned. The variables that should be considered in cleaning MF membranes include: frequency and duration of cleaning, chemicals and their concentrations, cleaning and rinse volumes, temperature of cleaning, recovery and reuse of cleaning chemicals, neutralization and disposal of cleaning chemicals.
- Membrane pretreatment: Feedwater pretreatment can be employed to improve the level of removal of various natural water constituents. It is also used to increase or maintain transmembrane flux rates and/or to retard fouling. The two most common types of pretreatment are coagulant and PAC addition.

ULTRAFILTRATION (UF)

UF involves the pressure-driven separation of materials from water using a membrane pore size of approximately 0.002 to 0.1 microns, an MWCO of approximately 10,000 to 100,000 daltons, and an operating pressure of approximately 200 to 700 kPa (30 to 100 psi). UF will remove all microbiological species removed by MF (partial removal of bacteria), as well as some viruses (but not an absolute barrier to viruses) and humic materials (see Fig. 2 and Table 1). Disinfection can provide a second barrier to contamination and is therefore recommended.

The primary advantages of low-pressure UF membrane processes compared with conventional clarification and disinfection (postchlorination) processes are:

- No need for chemicals (coagulants, flocculants, disinfectants, pH adjustment);
- Size-exclusion filtration as opposed to media depth filtration;
- Good and constant quality of the treated water in terms of particle and microbial removal;
- Process and plant compactness; and
- Simple automation.

Fouling is the limiting phenomenon responsible for most difficulties encountered in membrane technology for water treatment. UF is certainly not exempt from this fouling control problem. Therefore, membrane productivity is still an important subject, which should be thoroughly researched in order to have a better understanding of this phenomenon and its mechanisms.

Process

UF is a pressure-driven process by which colloids, particulates, and high molecular mass soluble species are retained by a process of size exclusion, and, as such, provides means for concentrating, separating into parts, or filtering dissolved or suspended species. UF allows most ionic inorganic species to pass through the membrane and retains discrete particulate matter and nonionic and ionic organic species.

UF is a single process that removes many water-soluble organic materials, as well as microbiological contaminants. Since all UF membranes are capable of effectively straining protozoa, bacteria, and most viruses from water, the process offers a disinfected filtered product with little load on any post-treatment sterilization method, such as UV radiation, ozone treatment, or even chlorination.

Unlike RO, the pretreatment requirement for UF is normally quite low. Fortunately, due to the chemical and hydrolytic stability of UF membrane materials, some of the pretreatments essential for RO membranes, such as adjustment of pH or chlorine concentration levels, do not apply. However, it may be necessary to adjust the pH to decrease the solubility of a solute in the feed so that it may be filtered out.

UF is designed to remove suspended and dissolved macromolecular solids from fluids. The commercially available modules are therefore designed to accept feedwaters that carry high loads of solids. Because of the many uses for UF membranes, pilot studies are normally conducted to test how suitable a given stream is for direct UF.

Water containing dissolved or chelated iron and manganese ions needs to be treated by an adequate oxidation process in order to precipitate these ions prior to UF membrane filtration, as with all membrane processes. This is recommended to avoid precipitation of iron and manganese in the membrane, or even worse, on the permeate side of the membrane (membrane fouling during the backwash procedure). Preoxidation processes generally used include aeration, pH adjustment to a value greater than eight, or addition of strong oxidants, such as chlorine, chlorine dioxide, ozone, or potassium permanganate.

Natural organic matter (NOM) is of great importance in potential fouling of the UF membrane and, consequently, in permeate flux that can be used under normal operating conditions. Thus, it is an interesting design option to use PAC or coagulants to pretreat the water to remove NOM and, consequently, decrease the surface of membrane needed.

Equipment

UF membranes can be fabricated essentially in one of two forms: tubular or flat-sheet.

Package plants, skid-mounted standard units that allow significant cost savings, are usually employed for plants treating less than 1.5 mgd. The primary skid-mounted system components may include an auto-cleaning prefilter, raw water pump, recirculation pump, backwash pump, chlorine dosing pump for the backwash water, air compressor (valve actuation), chlorine tank, chemical tank (detergent), programmable logic controller with program and security sensor (high pressure, low level, etc.)

Operation and Maintenance

The UF membrane plant may be divided into several subcategories:

- Raw water intake and pressure pumps;
- Pretreatment, which includes prescreening, prefiltration, and pH adjustment (if required) or any of the needed pretreatments;
- UF units;
- Chemical cleaning station, backwash station (which uses chlorinated product water), chlorine station, conditioner/preservative station; and
- Line for discharging or treatment of back wash water.

Operation and performance of a UF membrane plant are greatly influenced by raw water quality variations. Turbidity as well as Total Organic Carbon (TOC) of the raw water are water quality parameters of major importance that drive operation mode and membrane flux for all the UF plants presently in operation worldwide.

NANOFILTRATION (NF)

NF membranes have a nominal pore size of approximately 0.001 microns and an MWCO of 1,000 to 100,000 daltons. Pushing water through these smaller membrane pores requires a higher operating pressure than either MF or UF. Operating pressures are usually near 600 kPa (90 psi) and can be as high as 1,000 kPa (150 psi). These systems can remove virtually all cysts, bacteria, viruses, and humic materials (see Fig. 2 and Table 1). They provide excellent protection from DBP formation if the disinfectant residual is added after the membrane filtration step. Because NF membranes also remove alkalinity, the product water can be corrosive, and measures, such as blending raw water and product water or adding alkalinity, may be needed to reduce corrosivity. NF also removes hardness from water, which accounts for NF membranes sometimes being called "softening membranes." Hard water treated by NF will need pretreatment to avoid precipitation of hardness ions on the membrane.

More energy is required for NF than MF or UF, which has hindered its advancement as a treatment alternative.

Process

NF membranes have been observed to operate on the principle of diffusion rather than sieving as with MF and UF membranes.

Operation and Maintenance

Operational parameters of membranes include the physical and chemical properties of the membrane, the pore size or molecular weight cut-off (MWCO), and configuration.

REVERSE OSMOSIS (RO)

RO systems are compact, simple to operate, and require minimal labor, making them suitable for small systems. They are also suitable for systems where there is a high degree of seasonal fluctuation in water demand.

RO can effectively remove nearly all inorganic contaminants from water. RO can also effectively remove radium, natural organic substances, pesticides, cysts, bacteria, and viruses (see Fig. 2 and Table 1). RO is particularly effective when used in series. Water passing through multiple units can achieve near zero effluent contaminant concentrations. Disinfection is also recommended to ensure the safety of water.

Some of the advantages of RO are:

- Removes nearly all contaminant ions and most dissolved non-ions,
- Relatively insensitive to flow and total dissolved solids (TDS) level, and thus suitable for small systems with a high degree of seasonal fluctuation in water demand,
- RO operates immediately, without any minimum break-in period,
- Low effluent concentration possible,
- Bacteria and particles are also removed, and
- Operational simplicity and automation allow for less operator attention and make RO suitable for small system applications.

Some of the limitations of RO are:

- High capital and operating costs,

- Managing the wastewater (brine solution) is a potential problem,
- High level of pretreatment is required in some cases,
- Membranes are prone to fouling, and
- Produces the most wastewater at between 25–50 percent of the feed.

Process

RO removes contaminants from water using a semipermeable membrane that permits only water, and not dissolved ions (such as sodium and chloride), to pass through its pores. Contaminated water is subject to a high pressure that forces pure water through the membrane, leaving contaminants behind in a brine solution. Membranes are available with a variety of pore sizes and characteristics.

Equipment

Typical RO units include raw water pumps, pretreatment, membranes, disinfection, storage, and distribution elements. These units are able to process virtually any desired quantity or quality of water by configuring units sequentially to reprocess waste brine from the earlier stages of the process. The principal design considerations for reverse osmosis units are:

- operating pressure,
- membrane type and pore size,
- pretreatment requirements, and
- product conversion rate (the ratio of the influent recovered as waste brine water to the finished water).

WASTE STREAM DISPOSAL

Waste stream disposal is a significant problem in many areas. Unlike conventional treatment processes, in which approximately 5 to 10 percent of the influent water is discharged as waste, membrane processes produce waste streams amounting to as much as 15 percent of the total treated water volume. (See Table 2.) Because little or no chemical treatment is used in a membrane system, the concentrate stream usually contains only the contaminants found in the source water (although at much higher concentrations), and for this reason the concentrate can sometimes be disposed of in the source water. Other alternatives include deep well injection, dilution and spray irrigation, or disposal in the municipal sewer. These alternatives are usually necessary for NF wastes, which usually contain concentrated organic and inorganic compounds. Regardless of the type of membrane, disposal must be carefully considered in decisions about the use of membrane technology. Applicable local discharge regulations must be respected.

MEMBRANE INTEGRITY TESTING

One of the most critical aspects of employing membrane technology is ensuring that the membranes are intact and

Table 2. Surface Water Treatment Compliance Technology: Membrane Filtration

Unit Technologies	Complexity: Ease of Operation (Operator Skill Level)	Secondary Waste Generation	Other Limitations/ Drawbacks
Microfiltration	Basic: increases with pre/post-treatment and membrane cleaning needs.	Low-volume waste may include sand, silt, clay, cysts, and algae.	Disinfection required for viral inactivation.
Ultrafiltration	Basic: increases with pre/post-treatment and membrane cleaning needs.	Concentrated waste: 5 to 20 percent volume. Waste may include sand, silt, clays, cysts, algae, viruses, and humic material	Disinfection required for for viral inactivation.
Nanofiltration	Intermediate: increases with pre/post-treatment and membrane cleaning needs.	Concentrated waste: 5 to 20 percent volume.	Disinfection required under regulation, and recommended as a safety measure and residual protection.
Reverse Osmosis	Intermediate: increases with pre/post-treatment and membrane cleaning needs.	Briney waste. High volume, e.g., 25 to 50 percent. May be toxic to some species.	Bypassing of water (to provide blended/stabilized distributed water) cannot be practiced at risk of increasing microbial concentrations in finished water. Post-disinfection required under regulation, is recommended as a safety measure and for residual maintenance. Other post-treatments may include degassing of CO_2 or H_2S, and pH adjustment.

Source: U.S. Environmental Protection Agency, 1998.

continuing to provide a barrier between the feedwater and the permeate or product water. There are several different methods that can be employed to monitor membrane integrity, including:

- Turbidity monitoring,
- Particle counting or monitoring,
- Air pressure testing,
- Bubble point testing,
- Sonic wave sensing, and
- Biological monitoring.

WHERE CAN I FIND MORE INFORMATION?

(1) American Water Works Association and American Society of Civil Engineers. 1998. *Water Treatment Plant Design*. New York: McGraw-Hill.

(2) Bergman, A.R. and J.C. Lozier, 1993. Membrane Process Selection and the Use of Bench and Pilot Tests. *Membrane Technology Conference Proceedings*, Baltimore: American Water Works Association.

(3) Jacangelo, J.G., J-M. Laine, E.W. Cummings, A. Deutschmann, J. Mallevialle, M.R. Wiesner. 1994. *Evaluation of Ultrafiltration Membrane Pretreatment and Nanofiltration of Surface Waters*. Denver: American Water Works Association and AWWA Research Foundation.

(4) Jacangelo, J.G., S. Adham, J-M. Laine. 1997. *Membrane Filtration for Microbial Removal*. Denver: American Water Works Association Research Foundation and American Water Works Association.

(5) Mallevialle, J., P.E. Odendaal, and M.R. Wiesner, 1996. *Water Treatment Membrane Processes*. New York: McGraw-Hill.

(6) National Research Council. 1997. *Safe Water From Every Tap*. Washington, D.C.: National Academy Press.

(7) U.S. Environmental Protection Agency. 1990. *Environmental Pollution Control Alternatives: Drinking Water Treatment for Small Communities*. Washington, D.C.: Office of Water. EPA/625/5-90/025.

(8) U.S. Environmental Protection Agency. 1989. *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*. Washington, D.C.: Office of Water. EPA/625/4-89/023.

(9) U.S. Environmental Protection Agency. 1998. *Small System Compliance Technology List for the Surface Water Treatment Rule and Total Coliform Rule*. Washington, D.C.: Office of Water. EPA/815/R/98/001.

HAVE YOU READ ALL OF OUR FACT SHEETS?

“Tech Briefs,” four-page drinking water treatment fact sheets have been a regular feature in *On Tap* for more than two years. National Drinking Water Clearinghouse (NDWC) Technical Assistance Specialist Mohamed Lahlou, Ph.D., researches, compiles information, and writes these very popular items. “Tech Brief: Membrane Filtration” is the NDWC’s first eight-page fact sheet.

- Tech Brief: Disinfection, item #DWBLPE47;
- Tech Brief: Filtration, item #DWBLPE50;
- Tech Brief: Corrosion Control, item #DWBLPE52;
- Tech Brief: Ion Exchange and Demineralization, item #DWBLPE56;
- Tech Brief: Organics Removal, item #DWBLPE59;
- Tech Brief: Package Plants, item #DWBLPE63;
- Tech Brief: Water Treatment Plant Residuals Management, item #DWBLPE65;
- Tech Brief: Lime Softening, item #DWBLPE67;
- Tech Brief: Iron and Manganese Removal, item #DWBLPE70;
- Water Conservation Measures Fact Sheet, item #DWBLPE74; and
- Tech Brief: Membrane Filtration, item #DWBLPE81.

Additional copies of fact sheets are free; however, postal charges may be added. To order, call the NDWC at (800) 624–8301 or (304) 293–4191. You may also order online at ndwc_orders@estd.wvu.edu, *or download fact sheets from our Web site at* http://www.ndwc.wvu.edu.

For further information, comments about this fact sheet, or to suggest topics, call Lahlou at one of the above numbers or contact him via e-mail at mlahlou2@wvu.edu.

WATER METERS

Zane Satterfield
Vipin Bhardwaj
NDWC Engineering Scientist

Any viable business must be able to determine how much product it is making and selling and if that product is profitable. Water is a business. And, the best way for a water utility to measure or account for the water produced and then sold is by using water meters. This Tech Brief, discusses the different types of meters, their applications, and their importance for a water utility business.

WHY ARE METERS IMPORTANT?

Water meters are important to a utility for several reasons:

1. They make it possible to charge customers in proportion to the amount of water they use.
2. They allow the system to demonstrate accountability.
3. They are fair for all customers because they record specific usage.
4. They encourage customers to conserve water (especially as compared to flat rates).
5. They allow a utility system to monitor the volume of finished water it puts out.
6. They aid in the detection of leaks and waterline breaks in the distribution system.

A system without meters is like a taxi without a fare counter. Without a meter, it costs the same to drive around the block as it does from New York to Los Angeles.

TYPES OF METERS

Meters are classified into two basic types: positive displacement and velocity. Each of these meter types has variations, leading to the perception that there are several different kinds. Meters that feature both positive displacement and velocity are known as compound meters. The unit of measurement is usually in gallons but sometimes in cubic feet.

POSITIVE DISPLACEMENT METERS

In this type of meter, a known volume of liquid in a tiny compartment moves with the flow of water. Positive displacement flow meters operate by repeatedly filling and emptying these compartments. The flow rate is calculated based on the number of times these compartments are filled and emptied. The movement of a disc or piston drives an arrangement of gears that registers and records the volume of liquid exiting the meter. There are two types of positive displacement meters: nutating disc and piston.

Nutating disc meters have a round disc that is located inside a cylindrical chamber. The disc is mounted on a spindle. The disk nutates, or wobbles, as it passes a known volume of liquid through the cylindrical chamber. The rotating motion of the disk is then transmitted to the register that records the volume of water that went through the meter (see Fig. 1 below).

Piston meters have a piston that oscillates back and forth as water flows through the meter. A known volume of water is measured for each rotation, and the motion is transmitted to a register through an arrangement of magnetic drive and gear assembly.

Positive displacement meters are sensitive to low flow rates and have high accuracy over a wide range of flow rates. Positive displacement meters are used in homes, small businesses, hotels, and apartment complexes. They are available in sizes from 5/8" to two inches.

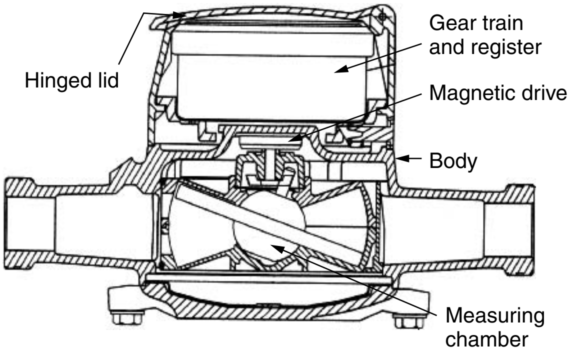

Figure 1. Nutating-disk meter with plastic housing. *Source*: AmericanWater Works Association. *Water Transmission and Distribution. Principles and Practices of Water Supply Operations*, 2nd Edition.

VELOCITY METERS

Velocity meters operate on the principle that water passing through a known cross-sectional area with a measured velocity can be equated into a volume of flow. Velocity meters are good for high flow applications.

Velocity meters come in different types, including turbine, multi-jet, propeller, ultrasonic, venture, and orifice meters. These meters are available in sizes of two inches and larger with the exception of multi-jet meters, which are between 5/8" and two inches.

Turbine meters have a rotating element that turns with the flow of water. Volume of water is measured by the number of revolutions by the rotor.

Venturi meters have a section that has a smaller diameter than the pipe on the upstream side. Based on a principle of hydraulics, as water flows through the pipe, its velocity is increased as it flows through a reduced cross-sectional area. Difference in pressure before water enters the smaller diameter section and at the smaller diameter "throat" is measured. The change in pressure is proportional to the square of velocity. Flow rate can be determined by measuring the difference in pressure. Venturi meters are suitable for large pipelines and do not require much maintenance.

Orifice meters work on the same principle as venturi meters, except that, instead of the decreasing cross-sectional area, there is a circular disk with a concentric hole. Flow rate is calculated similarly to the venturimeter by measuring the difference in pressures.

Ultrasonic meters send sound waves diagonally across the flow of water in the pipe. Changes in the velocity of water are converted electronically to change in flow rate.

Magnetic meters have an insulated section through which water flows. The flow of water induces an electrical current that is proportional to the velocity and hence the flow rate.

Propeller meters have a fan-shaped rotor that spins with the flow of water. A recorder is attached to the rotor to register the readings.

Multi-jet meters have tangential openings in a chamber to direct the water flow across a rotor with many vanes. Flow is measured proportional to the speed of the rotor.

COMPOUND METERS

In some cases, it is necessary to have a combination meter—both a positive displacement meter and velocity meter installed together—to be able to measure high and low flows. Low flows are measured through positive displacement while high flows are measured by velocity. A valve arrangement directs flows into each part of the meter (see Fig. 2).

METER INSTALLATION

Meters are installed either in outdoor meter pits (also known as meter wells) or inside the building served. Although it is common to have a meter installed in the basement, it is better to have the meter located at the curb or property line because of easy access for reading or maintenance. It is sometimes difficult to gain access to the

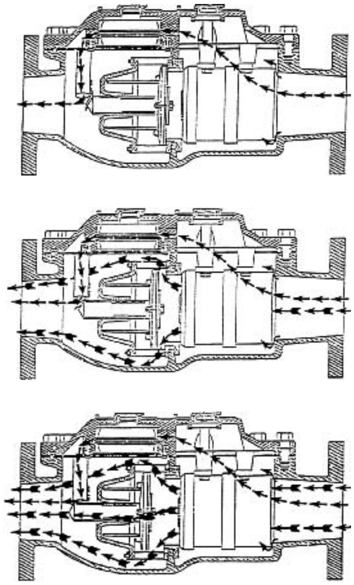

Low flow: All of the water passes through the nutating disk measuring element.

Crossover: As the control valve opens under higher flow rates, water passes through both measuring elements while the disk-side throttling begins.

Full flow: At high flow rates the control valve is fully open. The bulk of the water passes through the turbine measuring element, and the disk-side is throttled to a minimal amount.

Figure 2. Compound meter. *Source*: American Water Works Association, *Water Transmission and Distribution, Principles and Practices of Water Supply Operations*, 2nd Edition.

residence or building when no one is there. Nevertheless, it may be necessary due to extreme cold weather to have the meter located indoors to prevent frost damage. Large meters are usually installed in concrete or block vaults preferably out of traffic areas, but they also can be installed inside.

General guidelines for installing meters are:

- In outdoor meter pits, the face of the meter should be between 18 and 24 inches from the ground surface or top of the meter pit lid.
- Meter pits or wells should have six to 12 inches of gravel at the bottom to help with drainage.
- Large meter pits should have a drain or a sump pump if a drain is not possible. There should never be standing water in a meter pit or well.
- If at all possible, the meter pit or well should not be located in an area prone to flooding.
- The meter setting should have a shut-off valve on both sides of the meter (i.e., the customer's side and the water main side).
- If possible, the meter should be installed in a horizontal position.
- The meter should be easily accessible for service, inspection, and reading even if meter reading is via remote sensing equipment.
- Protect the meter from freezing.
- Always be mindful of public safety. Don't have the meter lid sticking up or have it too low so that it creates a hole.
- Install seals on the meter to reduce tampering and indicate when tampering does occur.

The following suggestions apply to larger meter vaults or pits:

- Large meter pits should also have a drain or a sump pump if a drain is not possible. There should never be standing water in a meter pit or well.
- To prevent disruption of service when replacing or repairing large meters, there should be a bypass that can also be metered. Having the bypass metered would be similar to a manifold set-up where you have two or more meters in parallel making service of one meter easy without service disruption or lost revenue.
- Large meter installations should have good structural support to prevent stress on the water line. Also there should be at least 10 times the pipe diameter of straight pipe before the meter and five times the pipe diameter of straight pipe after the meter.
- Some large meters recommend or require a strainer to be installed ahead of the meter.
- With large meters that are located in concrete vaults or pits, ideally, the meter face should be located over the hatchway to help with reading the meter, possibly reading the meter without even entering the pit. This will also help in lifting the meter from the pit. Remember, if the meter reader has to enter the meter vault or pit, this is considered a confined space entry and the proper safety procedures must be followed.

Small meter installation is easier with a meter yoke (also known as a meter setter). Meter yokes have different configurations and can have any combination of built-in check valves, regulators, and lockable shutoff valves. Utilities should have their own set specifications with illustrations depicting proper meter installations.

METER READING

There are two common types of water meter registers: circular and straight. Either of the meters can be manufactured to read in gallons or cubic feet. The circular register is somewhat more difficult to read.

The straight register is what is seen on newer meters and is read like a car's odometer. On some larger meters there will be a multiplier on the register face or the meter itself. The multiplier will be noted as 10X, 100X, or 1000X depending on the size of the meter. The multiplier indicates that the meter reader must multiply what is read by the number noted 10, 100, or 1,000. For example, if the register displays the number 975 and the multiplier is 10X, then the meter reading is 9750 gallons.

There are several different ways meter reading can be accomplished.

- Direct read is, as the name suggests, an individual directly reading the register.
- Remote reading is done through an electronic signal, usually with a wire to a separate station where the individual could either direct read or touch read.
- Touch read or plug-in readers use a handheld device that takes the electronic signal from plugging-in or touching the remote station and translates it into the number of gallons used. The remote station can even be on top of the meter pit lid. The readings are stored

on the hand-held device until the meter reader gets back to the office to be downloaded to a computer.

- A newer development is automatic meter reading, which enables the meter reader to drive down the street and obtain the readings from the meters usually through radio transmission.

SELECTING A METER

Meters are selected using several factors: flow rate, size of pipe, pressure loss and safety considerations, such as fire service regulations. For sizes of one inch and smaller and low flow rates, positive displacement types of meters are common. For residential uses, 5/8" or 3/4" meters are used.

For medium flows, such as in apartment buildings, businesses, and public buildings, positive displacement meters in sizes of 1", 11/2", or two inches are used. In sizes of two and three inches, either, displacement, multi-jet, or turbine types of meters can be used. In the three- to four-inch size range, the meter type depends on the average flow rate. If the flow rate is between five and 35 percent of maximum flow rate, the positive displacement type is better. If the flow rates are going to be 10 to 15 percent of the maximum capacity, a turbine type should be used. If close accuracy at low flows is important, but large flows also have to be measured, a compound meter is best.

For large flows, velocity meters are more appropriate. Turbine meters are suitable for large flows where minimum flow rate is above 10 to 20 percent of maximum rating. Turbine meters have low pressure loss at high flow rates. Propeller meters are suitable for large mainlines or for pump station discharge.

METER TESTING AND MAINTENANCE

For meters to be effective, they must accurately read the amount of water flowing through them. Proper testing and maintenance procedures should be in place to assure accuracy.

Master meters—meters at the treatment plant that measure the largest water volumes—are the highest priority in the meter testing and maintenance program. These meters determine how much water is produced or sold and should be tested once a year. (More often if large amounts of water are measured through the meter.)

Industrial and commercial meters are the second testing and maintenance priority in the system. "The most expensive water loss in the distribution system is associated with under registration of industrial and commercial water meters, rather than leaks," says Jeff Bennett utility inspector with the West Virginia Public Service Commission. "Lost water sales are lost revenue at the retail rate. Lost water due to leaks is only the loss of the cost to produce the water."

Bennett also points out, "a two-inch meter will, at a minimum, provide as much revenue as eight residential meters, while a six-inch meter will provide at least as much revenue as 50 residential meters. Since large meters provide considerably more revenue per meter, it only makes 'cents' to test them more frequently."

Meters larger then two inches are usually tested in-place using a calibrated field test meter. To be able to field test a meter, there must be a tap and an isolation valve immediately downstream from the meter. This temporary connection can be made with a fire hose. Even new meters should be tested.

Master meters in water plants are, in some cases, difficult or impossible to test using a field test meter. In these situations, a draw-down test can be conducted, which involves comparing a known volume of water pumped out of a tank (clearwell) to the volume recorded on the meter being tested.

Although residential meters may be the final priority in the testing and maintenance program, in small systems, they may be the most important. The residential meters can be broken down to the ones that have high usage and the ones that don't. It only makes sense to concentrate on residential meters that have high usage because of a big family or multiple occupants compared to residential meters that never reach the minimum bill because of single occupancy of an elderly person.

The majority of residential meters are the positive displacement type, which almost always slow down when they are worn or encrusted by minerals or debris. With residential meters, the testing program should either consist of periodic testing on a test bench or a complete change-out program.

Residential meters should be checked, cleaned, and calibrated every seven to 10 years or as indicated by your state primacy agency. If your system does not have a meter test bench, check with a neighboring system or shop around for a company that can do testing on a certified test bench.

Testing and maintenance of the meters depends on the quality and quantity of the water. If adverse conditions, such as high minerals or large flows are encountered, meters will require more frequent attention. Water that has high levels of minerals will affect the operation of a meter over time. This should be taken into account with a system's preventive maintenance program by cleaning the meter to extend its life and to improve the accuracy of the meter.

Maintenance requirements will vary depending upon the material used in the meter components. Although plastic meters are less expensive, more maintenance and more frequent replacement may be necessary.

If your water system does not have any meters for residential connections, and you want to start metering it, a typical cost would be $280 for connection (house). The price breaks down as follows:

• Direct read 5/8" meter	**$40**
• Tandem meter setter (yoke) able to accept regulator	**$90**
• Pressure regulator (3/4" typical size)	**$40**
• Meter well (pit)	**$25**
• Meter frame and lid	**$35**
• Corporation-stop	**$15**
• 6" service saddle	**$35**
• Total for each connection	**$280**

If the system has 200 residential connections, the total material cost would be $56,000. With a meter life expectance of 15 years, it would take less than $1.50 per connection per month to recoup the cost.

If your system already has meters, but doesn't have a good maintenance program, here is an example of revenue lost due to a slow residential meter (assuming a 4,500 gallons per month rate at a cost of $25.00):

Loss with a meter 3 percent slow = $0.75 per month or $9.00 per year **for one meter**.

Loss with a meter 5 percent slow = $1.25 per month or $15.00 per year **for one meter**.

In some small systems, city buildings, golf courses, municipal swimming pools, and even the water plant are not charged. These connections should be metered to account for all treated water.

Keep in mind that if your system has a sewer utility, the loss is compounded, because most sewer utilities rely on water meters to bill for the sewer service. In these cases having an inaccurate water meter will have a direct impact on sewer revenue.

READING LIST

American Water Works Association. 1996. *Water Transmission and Distribution, Principles and Practices of Water Supply Operations*, 2nd Edn. AWWA, Denver.

American Water Works Association. 1986. *Water Meters—Selection, Installation, Testing, and Maintenance*, 3rd Edn. AWWA, Denver.

Bennett, Jeff. Unpublished. *"Water Metering" and "Meter Testing Programs."* White papers for the Public Service Commission of West Virginia.

MICROBIOLOGICAL CONCERNS OF DRINKING WATER DISTRIBUTION SYSTEMS

Mohammad R. Karim
American Water, Quality Control and Research Laboratory
Belleville, Illinois

Mark W. LeChevallier
American Water
Voorhees, New Jersey

Drinking water in a distribution system may contain microbes that enter the distribution system through treatment breakthrough, regrowth, or contamination of water in the system. The majority of these organisms are harmless. However, pathogenic organisms may also be present. A number of waterborne disease outbreaks related to distribution systems have been reported. Many factors contribute to the regrowth and contamination of drinking water in distribution systems. This article discusses the factors that contribute to microbial regrowth and contamination of distribution systems. To ensure protection of public health, these factors should be adequately addressed to safeguard drinking water from contamination in distribution systems.

INTRODUCTION

The purpose of a drinking water distribution system is to deliver water that is safe and acceptable in odor, taste, and appearance. Most drinking water utilities rely on the "multiple-barrier concept" as the guiding principle for providing safe water. Traditionally, the barriers have included (1) source water protection; (2) coagulation, flocculation, and sedimentation; (3) filtration; (4) disinfection; and (5) protection of water in the distribution system. These barriers are used to duplicate microbial removal capabilities by succeeding process steps so that sufficient backup systems are available to permit continuous operation in the face of normal mechanical failures. The distribution system is considered the final and one of the most important barriers in the multiple-barrier concept for providing safe drinking water.

After the treatment process, the microbiological quality of drinking water is at its highest; however, as the treated water travels through the distribution system, the microbiological quality of the water may deteriorate through three basic mechanisms: (1) breakthrough from treatment plants, (2) microbial growth/regrowth, and (3) contamination within the distribution system. This article discusses the microbiological concerns of drinking water distribution systems.

OUTBREAKS FROM DISTRIBUTION SYSTEM DEFICIENCIES

From 1971 to 1998, 619 drinking waterborne disease outbreaks were reported in the United States due to microbiological, chemical, or unknown etiology (1). Of the 619 waterborne disease outbreaks, 113 outbreaks and 21,058 reported cases of illness or 13 deaths were from distribution system contamination. Increasingly, concern has been raised regarding the adequacy of distribution system barriers. Distribution system deficiencies were associated with 30.3% of the community outbreaks (1). Lippy and Waltrip (2) reported that nearly 37% of the community outbreaks from 1946 to 1980 were due to failures of a distribution system. Craun and McCabe (3) reported that distribution system deficiencies accounted for most of the waterborne hepatitis outbreaks (10 out of 17 outbreaks) in public water systems from 1946 to 1970. Contamination of mains during construction or repair and cross-connections were cited as some of the primary causes of distribution deficiencies. In recent years, waterborne disease outbreaks due to distribution system deficiencies have had disastrous consequences. For example, an outbreak of hemorrhagic *Escherichia coli* serotype 0157:H7 occurred in Cabool, MO during December 1989 and January 1990 and resulted in 243 cases of diarrhea and 4 deaths (4). It was concluded that the illness was caused by contaminants that entered the distribution system through two major pipe breaks and 43 service meter failures that occurred during unusually cold weather.

ETIOLOGICAL AGENTS OF OUTBREAKS RELATED TO DRINKING WATER DISTRIBUTION SYSTEMS

Microbiological and chemical contaminants were associated with 75 and 38 outbreaks, respectively, for the 113 reported outbreaks caused by distribution system deficiencies from 1971 to 1998 (1). Of the outbreaks caused by microbial contaminants, bacteria were responsible for 21%, protozoa for 17%, viruses for 8%, and unknown etiological agents for 54% of outbreaks (Fig. 1). The bacterial agents identified were *Salmonella, E. coli* 0157:H7, *Shigella, Campylobacter, Salmonella typhimurium,* and *Vibrio cholerae*. The protozoan agents identified were *Giardia* and *Cyclospora*. Viruses identified were norovirus and hepatitis A.

MICROBIAL OCCURRENCE IN THE DISTRIBUTION SYSTEM

Drinking water in a distribution system may contain many microbes that enter the system through treatment breakthrough, regrowth, or contamination of water in the system. The majority of these organisms are harmless and not many have the potential to be pathogenic. However, pathogens may also pass into drinking water in a distribution system. Two types of pathogens may be present in the drinking water—primary pathogens (those that inhabit the gastrointestinal tract) and opportunistic pathogens (microorganisms that may exist as part of a normal body microflora but under certain conditions may cause disease in compromised hosts such as immunocompromised individuals). Primary pathogens include *Salmonella, E. coli* 0157, *Shigella, Yersinia, Campylobacter*, enteric viruses, and protozoan parasites such as *Giardia* and *Cryptosporidium*. These organisms may cause a variety of diseases including gastroenteritis, typhoid, paratyphoid fever, dysentery, cholera, hepatitis, paralysis, and meningitis.

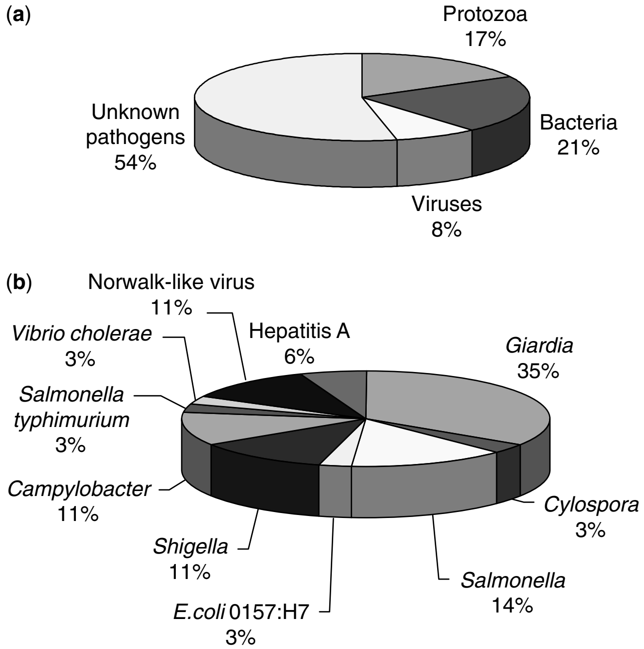

Figure 1. (**a**) Etiological agents for U.S. waterborne disease outbreaks caused by distribution system deficiencies from 1971–1998; (**b**) percent of outbreaks caused by the specific microbial etiology from 1971–1998. Data from Reference 1.

Opportunistic pathogens that may occur in a drinking water distribution system include *Klebsiella pneumoniae, Pseudomonas aeruginosa, Legionella pneumophila, Aeromonas hydrophila, Citrobacter spp., Enterobacter spp., Moraxella spp., Mycobacterium avium* complex, *Staphylococcus aureus*, fungi, and yeast (4,5). These organisms may cause diseases such as pneumonia, meningitis, septicemia, gastrointestinal illness, and respiratory tract infection.

Factors Contributing to Distribution System Contamination

Because of the expansive nature of a distribution system, that has many miles of pipe, storage tanks, and interconnections with industrial users, opportunities for contamination do exist. A list of possible sources of distribution system contamination is presented in Table 1. Cross-connections are recognized as a major risk to water quality, but their probability of occurrence is low due to vigilant cross-connection control programs. However, complete cross-connection control is difficult to achieve, and many systems face challenges in maintaining an effective cross-connection control program in the face of dwindling resources and personnel cutbacks.

Inadequate separation of water mains and sewers may contribute to distribution system contamination through broken or leaking water mains. Engineering standards call for a separation of 10 ft between drinking water and sewer pipelines, although separations can be as little as 18 inches if the pipelines are placed on different levels (6). In saturated soil, microbes can move several meters in short periods of time (7). This transport could be aided by water flowing out of the sewer line. Because it is not uncommon for water systems to lose more than 10% of total production through leaks in pipelines (8), opportunities exist for water to intrude through these portals.

Despite the best efforts to repair main breaks using good sanitary procedures, main breaks are an opportunity

Table 1. Sources of Distribution System Contamination

- Cross-connection
- Inadequate separation of water main and sewer
- Contamination of broken or leaking water main
- Contamination of mains during construction/repair
- Back-siphonage
- Contamination of distribution storage
- Intrusion of contaminants due to pressure transients

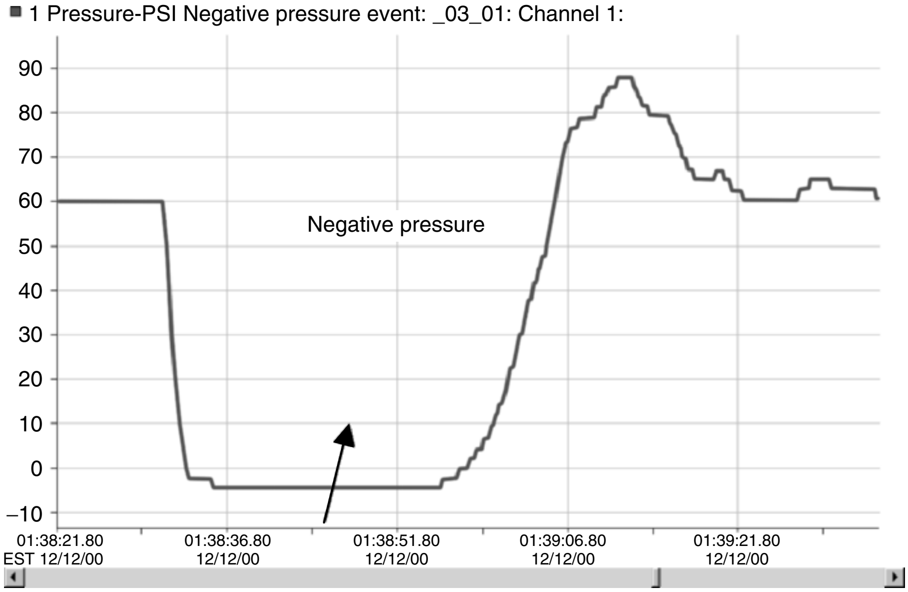

Figure 2. A negative pressure transient in a drinking water distribution system following a power outage. The negative pressure lasted approximately 22 seconds before static pressures were restored.

for contaminants to enter a distribution system. Utilities typically isolate the affected section, super chlorinate, and flush the repaired pipe. However, flushing velocities may not always be achievable to remove all contaminated debris, and microbiological tests to check the final water quality may not detect contaminating organisms. McFeters et al. (9) reported high levels of injured coliform bacteria, not detectable by standard coliform techniques, following the repair of a main break. Resampling at the site 1 week later showed high levels of the coliform bacteria, detectable only using m-T7 agar, a medium specially designed to recover chlorine-injured coliforms.

Installation of backflow devices to prevent the entry of contaminated water is an important distribution system barrier. Due to cost considerations, backflow devices are primarily installed on commercial service lines where the facility uses potentially hazardous substances (4). Examples of such facilities include hospitals, mortuaries, dry cleaners, and industrial users. It is not common that all service connections have backflow devices, so the possibility of back-siphonage exists at these points. In addition, installation of backflow devices for all service connections would make routine checking of the devices nearly impossible and without routine inspection, the proper functioning of the units cannot be determined. Even when backflow devices have been installed, contamination has occurred. The failure of a backflow check valve allowed water stored for fire protection to enter the Gideon, Missouri, distribution system (4). A broken vent in the storage tank allowed birds to enter and contaminate the water with *Salmonella*. Three people died due to *Salmonella* infection.

Pressure transients in drinking water pipelines are caused by an abrupt change in the velocity of water. These events are also termed “surges” or “water hammer.” These pressure transients are waves that have both positive and negative amplitude. They can draw transient negative pressures in a distribution system that last only for seconds and may not be observed by conventional pressure monitoring. Circumstances that produce these pressures waves may commonly occur in every water system. Because these waves travel through the distribution system, the transient negative pressure wave can momentarily draw water back into the pipe at any point where water is leaking out of the system. Depending on the size of the leaks, the volume of intrusion can range from one to hundreds of gallons (10,11). A negative pressure surge in a distribution system has been documented (12). A negative pressure surge wave was observed following a power outage that suddenly shut down a pumping station (Fig. 2). The rapid change in water velocity resulted in a negative pressure that lasted for approximately 22 seconds until static pressures were restored. Pressure transients can be caused by main breaks, sudden change in demand, uncontrolled pump starting or stopping, opening and closing of a fire hydrant, power failure, air valve slam, flushing operations, feed tank draining, and other conditions.

If pathogens are present at the exterior of the distribution pipelines, they may enter into the system due to hydraulic pressure gradients or during construction, repair, cross-connections, and conditions or activities where the system is open to the atmosphere or the environment. Karim et al. (12) reported the presence of microbial indicator and pathogens adjacent to distribution systems. Total coliform and fecal coliform bacteria were detected in water and soil samples adjacent to distribution systems in about half of the samples, indicating the presence of fecal contamination (Fig. 3). *Bacillus* was found in almost all the samples, which is not a surprise because it is a normal soil organism. Viruses were detected using culturable methods in 12% of the soil and water samples and by molecular methods in 19% of the soil samples and 47% of the water samples. When these data are combined, 56% of the samples were positive for viruses

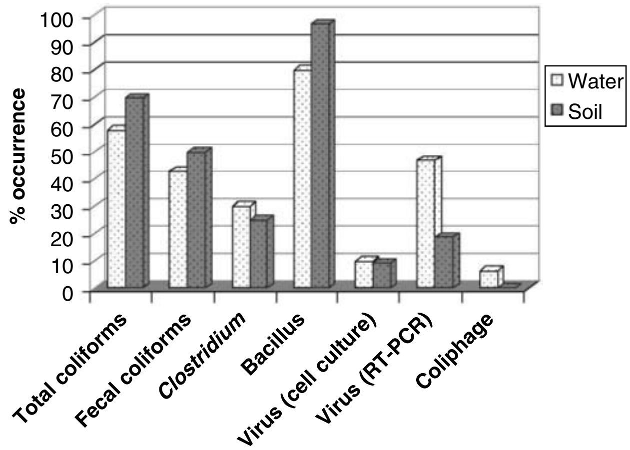

Figure 3. Occurrence of microorganisms in water and soil samples adjacent to distribution system pipelines.

either in the water or the soil. Sequence analysis showed that these viruses were predominantly enteroviruses (the vaccine strain of poliovirus), but Norwalk and hepatitis A viruses were also detected, providing clear evidence of human fecal contamination immediately exterior to the pipe. This data suggest that opportunities could exist for pathogens to intrude into the distribution system and emphasize the need to maintain an effective disinfectant residual in all parts of the distribution system (13).

MICROBIAL REGROWTH IN A DISTRIBUTION SYSTEM

Recent studies have examined data from more than 90 water systems to determine the factors that contribute to microbial growth in drinking water (14,15). These studies have shown that the occurrence of coliform bacteria can be related to the following factors: filtration, temperature, disinfectant type and residual, assimilable organic carbon (AOC) level, corrosion control, and pipe material selection.

Filtration

Four unfiltered surface water systems included in a study (14) accounted for 26.6% of the total number of bacterial samples collected but represented 64.3% (1013 of 1576) of the positive coliform samples. Although the results do not suggest that treatment was inadequate (e.g., coliforms were not related to breakthrough of treatment barriers), the data suggested that filtration may be an important factor in preventing coliform regrowth. Following the study, one of the systems installed filtration, and distribution system coliform levels were reduced by a factor of 3 during the following 18-month interval.

Temperature

Temperature is an important factor in microbial growth. On average, the occurrence of coliform bacteria was significantly higher when water temperatures were >15 °C. (14). However, the minimum temperature at which microbial activity was observed varied from system to system. Systems that typically experienced cold water had increases in coliform occurrences when water temperatures ranged near 10 °C. The strains of coliform bacteria in these systems may be better adapted to grow at lower temperatures (psychrophiles).

Disinfectant Residual and Disinfectant Level

Maintenance of a disinfectant residual throughout a distribution system is intended to produce conditions unfavorable for bacterial survival in drinking water. Choosing the appropriate disinfectant or improper application of disinfectant can affect bacterial growth conditions. For filtered systems, there was a difference between systems that maintained free chlorine residual and systems that used chloramines (14). For systems that used free chlorine, 0.97% of 33,196 samples contained coliform bacteria, whereas 0.51% of 35,159 samples from chloraminated systems contained coliform bacteria (statistically different at $p < .0001$). The average density of coliform bacteria was 35 times higher in free chlorinated systems compared to chloraminated water [0.60 colony forming units (cfu)/100 mL for free chlorinated water, compared to 0.017 cfu/100 mL for chloraminated water]. Previous research hypothesized that chloramines may be better able to penetrate into distribution system biofilms and inactivate attached bacteria (16,17).

In addition to the type of disinfectant used, the residual maintained at the end of the distribution system was also associated with coliform occurrences (14). Systems that maintained dead-end free chlorine levels <0.2 mg/L or monochloramine levels <0.5 mg/L had substantially more coliform occurrences than systems that maintained higher disinfectant residuals. However, systems that had high AOC levels needed to maintain high disinfectant residuals to control coliform occurrences. Therefore, maintenance of a disinfectant residual alone did not ensure that treated waters would be free of coliform bacteria.

AOC Level

The combined results from two surveys of AOC levels in North American drinking water systems are shown in (Fig. 4) (14,15). The levels (summarized as the geometric mean based on 12 to 36 samples) range from 20 to 214 µg/L. The results also indicate that the majority of the total AOC results from the growth of the test organism,

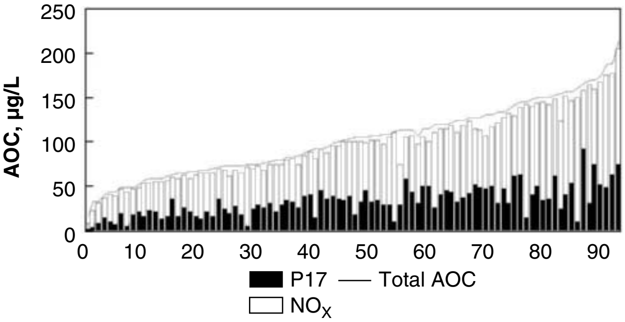

Figure 4. AOC levels in 94 North American water systems.

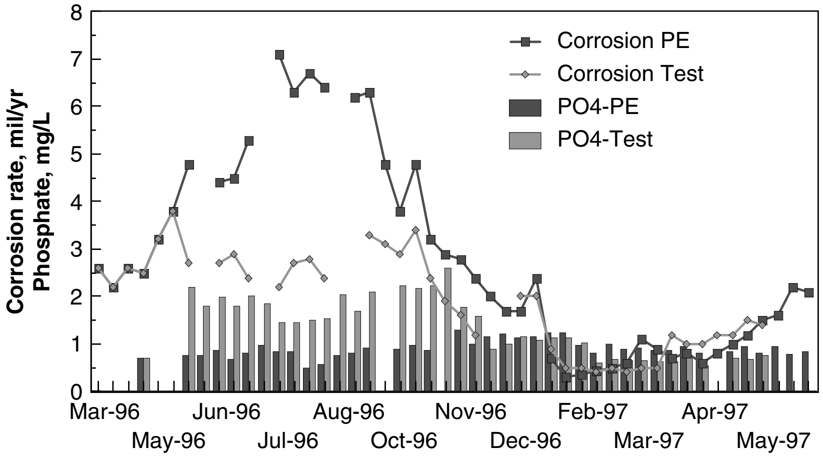

Figure 5. Increasing phosphate levels can reduce corrosion rates.

Spirillum sp., strain NOX. This $\mathrm{AOC}_{\mathrm{NOX}}$ fraction is influenced by disinfection practices (chlorine, ozone, etc.) and suggests that changes in these practices (i.e., the type of disinfectant, the point of application, dose) can impact AOC levels in finished drinking water.

High levels of AOC can stimulate bacterial growth in distribution system biofilms. On average, free chlorinated systems with AOC levels greater than 100 μg/L had 82% more coliform positive samples, and the coliform densities were 19 times higher than in free chlorinated systems that had average AOC levels less than 99 μg/L. However, high levels of AOC alone do not dictate the occurrence of coliform bacteria in drinking water; they are only one factor.

Corrosion Control and Pipe Materials

Most systems do not measure corrosion rates on a daily basis, so this parameter is difficult to evaluate on a full scale. However, recent research has demonstrated that corrosion of iron pipes can influence the effectiveness of chlorine-based disinfectants in inactivating biofilm bacteria (17,18). Therefore, the choice of pipe material and the accumulation of corrosion products can dramatically impact the ability to control the effects of biofilms in drinking water systems.

Figure 5 shows the average monthly corrosion rates [in mils (thousandths of an inch) per year] in a system in Illinois. The conventional plant effluent corrosion rate showed marked seasonal variations. Corrosion rates were highest during the summer months when, traditionally, coliform occurrences the highest. Similar seasonal variations have been observed in other systems (19). This variation in rates of corrosion is important because the corrosion products react with residual chlorine and prevent the biocide from penetrating the biofilm and controlling bacterial growth. Studies have shown that free chlorine is impacted to a greater extent than monochloramine, although the effectiveness of both disinfectants is impaired if corrosion rates are not controlled (17,18). Increasing the phosphate-based corrosion inhibitor dose, especially during the summer months, can help reduce corrosion rates. In full-scale studies, systems that used a phosphate-based corrosion inhibitor had lower coliform levels than systems that did not control corrosion (14).

CONCLUSION

Drinking water in a distribution system may contain microbes that enter the distribution system through treatment breakthrough, regrowth, or contamination of water in the system. The majority of these organisms are harmless. However, pathogenic organisms may also be present in drinking water. A number of waterborne disease outbreaks related to distribution systems have been reported. Factors that may contribute to distribution system contamination include cross-connections, inadequate separation of water mains and sewers, contamination of broken or leaking water main, contamination of mains during construction/repair, back-siphonage, contamination of distribution storage, and intrusion of contaminants due to pressure transients. Factors that may contribute to regrowth include filtration, temperature, disinfectant type and residual, assimilable organic carbon (AOC) level, corrosion control, and pipe material selection. To ensure the protection of public health, these factors should be adequately addressed to safeguard drinking water from contamination in a distribution system.

BIBLIOGRAPHY

1. Craun, G. and Calderon, R. (2001). Waterborne disease outbreaks caused by distribution system deficiencies. *JAWWA* **93**(9): 64–75.
2. Lippy, E.C. and Waltrip, S.C. (1984). Waterborne disease outbreaks—1946–1980: A thirty-five year perspective. *JAWWA* **76**: 60–67.
3. Craun, G.F. and McCabe, L.J. (1973). Review of the causes of waterborne disease outbreaks. *JAWWA* **65**: 74–84.
4. Geldreich, E.E. (1996). *Microbial Quality of Water Supply in Distribution Systems*. Lewis, Boca Raton, FL.
5. Geldreich, E.E. (1990). Microbiological quality control in distribution systems. In: *Water Quality and Treatment*. F.W. Pontius (Ed.). American Water Works Association. McGraw-Hill, New York, Chap. 18.
6. *Recommended Standards for Water Works*. (1992). *Great Lakes Upper Mississippi River Board of State Public Health*

& Environmental Managers. Health Education Services, Albany, NY.

7. Abu-Ashour, J., Joy, D.M., Lee, H., Whiteley, H.R., and Zelin, S. (1994). Transport of microorganisms through soil. *Water Air Soil Pollut.* **75**: 141–158.
8. American Water Works Association and American Water Works Association Research Foundation. (1992). *Water Industry Database: Utility Profiles*. AWWA, Denver, CO.
9. McFeters, G.A., Kippen, J.F., and LeChevallier, M.W. (1986). Injured coliforms in drinking water. *Appl. Environ. Microbiol.* **51**: 1–5.
10. Funk, J.E., Van Vuuren, S.J., Wood, D.J., Lecherallier, M., and Freidman, M. (1999). Pathogen intrusion into water distribution system due to transients. *Proceedings of the 3rd ASME/JSME Joint Fluid Engineering Conference*, July 18–22 San Francisco, CA.
11. LeChevallier, M.W. (1999). The case for maintaining a disinfectant residual. *JAWWA* **91**(6): 86–94.
12. Karim, M.R., Abbaszadegan, M., and LeChevallier, M. (2003). Potential for pathogen intrusion during pressure transients. *JAWWA* **95**(5): 134–146.
13. LeChevallier, M.W., Gullick, R.W., and Karim, M.R. (in press). The potential for health risks from intrusion of contaminants into the distribution system from pressure transients. *J. Water Health.*
14. LeChevallier, M.W., Welch, N.J., and Smith, D.B. (1996). Full scale studies of factors related to coliform regrowth in drinking water. *Appl. Environ. Microbiol.* **62**(7): 2201–2211.
15. Volk, C., Welch, N., and LeChevallier, M. (1996). *Survey of Assimilable Organic Carbon and Coliform Regrowth in the American Water System*. American Water Works Service Company, Inc., Voorhees, NJ.
16. LeChevallier, M.W., Lowry, C.D., and Lee, R.G. (1990). Disinfecting biofilms in a model distribution system. *JAWWA* **82**(7): 87–99.
17. LeChevallier, M.W. (1991). Biocides and the current status of biofouling control in water systems. *Proc. Int. Workshop Ind. Biofouling Biocorrosion*. Springer-Verlag, Germany, pp. 113–132.
18. LeChevallier, M.W., Lowry, C.D., Lee, R.G., and Gibbon, D.L. (1993). Relationship between iron corrosion and disinfection of biofilm bacteria. *JAWWA* **85**(7): 111–123.
19. Norton, C.D. and LeChevallier, M.W. (1997). Chloramination: Its effect on distribution system water quality. *JAWWA* **89**(7): 66–77.

NITRIFICATION OF POTABLE WATER USING TRICKLING FILTERS

D.V. Vayenas
University of Ioannina
Agrinio, Greece

Gerasimos Lyberatos
University of Patras
Patras, Greece

INTRODUCTION

It is well established that nitrogen in the form of ammonia must be removed from wastewater before it is disposed of, as well as from potable water before it can be consumed. When ammonia is discharged to an aquatic environment, it leads to oxygen consumption (nitrogenous oxygen demand) that causes fish death. Furthermore, ammonia must be removed before water is disinfected with chlorine because ammonia reacts with chlorine and produces chloramines, which have been found to be carcinogenic (1,2).

Ammonia may be removed chemically or biologically. Nitrification, the key biological process for ammonia oxidation, is carried out by two different nitrifying bacteria (aerobic, autotrophs), *Nitrosomonas* and *Nitrobacter*. The overall reaction may be written as:

$$NH_4{}^+ + 2O_2 \rightarrow NO_3{}^- + H_2O + 2H^+ \tag{1}$$

According to this reaction, it takes 4.57 g of oxygen to oxidize 1 g of $NH4^+{-}N$. The two bacterial groups are each responsible for a single step in the oxidation of ammonia to nitrate, according to the following reaction scheme:

Nitrosomonas:

$$NH_4{}^+ + \frac{3}{2}O_2 \rightarrow NO_2{}^- + 2H^+ + H_2O \tag{2}$$

Nitrobacter:

$$NO_2{}^- + \frac{1}{2}O_2 \rightarrow NO_3{}^- \tag{3}$$

Nitrification is typically used for tertiary treatment of wastewater and for pretreatment of potable water. Nitrifying bacteria are characterized by low specific growth rates. Trickling filters provide a support medium for biofilm growth, thus allowing the possibility of maintaining nitrifying bacteria at high hydraulic loadings. Wastewater treatment using biofilms grown on support media was the first continuous flow bioprocess employed by sanitary engineers nearly a century ago. As the wastewater flows over the biological slime in a nitrifying trickling filter, ammonia as well as dissolved oxygen diffuse into the biofilm where they are metabolized by the bacteria. The essential processes are mass transport and bioconversion. Modern trickling filters often use plastic support media that attain better performance but at a substantially higher cost than filters based on rock. Nitrification of potable water is a relatively new process, and only limited information is available. Pozzuolana and Biolite filters have been used to nitrify potable water (3).

Trickling filter operation is based on the following principles: The organic material present in the wastewater is degraded by a population of microorganisms attached to the filter medium (Fig. 1). Organic material from the liquid is adsorbed onto the biological film or slime layer. In the outer portions of the biological slime layer (0.1 to 0.2 mm), the organic material is degraded by aerobic microorganisms. As the microorganisms grow, the thickness of the slime layer increases, and the diffused oxygen is consumed before it can penetrate the full depth of the slime layer. Thus, an anaerobic environment is established near the surface of the medium. As the slime layer increases in thickness, the adsorbed organic matter is metabolized before it can reach the microorganisms near the medium face. As a result of having no external organic source available for cell carbon, the microorganisms near the medium face enter into an endogenous phase of growth and lose their ability to cling to the medium surface. The liquid then washes the slime off the medium, and a new slime layer starts to grow (Fig. 1). This phenomenon of losing the slime layer is called “sloughing” and is primarily a function of the organic and hydraulic loading on the filter. In modern trickling filters, the hydraulic loading rate is adjusted to maintain a slime layer of uniform thickness (4).

THE EFFECT OF SUPPORT MATERIAL ON NITRIFYING TRICKLING FILTERS FOR POTABLE WATER TREATMENT

The advantages of trickling filters include low maintenance, inexpensive installation, and great tolerance of variations in hydraulic and organic loads. As mentioned earlier, the bacteria responsible for oxidizing ammonia to nitrate are autotrophic organisms that have very slow specific growth rates. Thus, their stabilization on a support material and the development of biofilm structures enables the permanent existence of these organisms in the filter and the possibility of maintaining nitrifying bacteria at high hydraulic loadings. The filter medium is of such great importance for the function of biofilters that several experiments have been carried out to study and compare different filter media types (5,6). The support materials used in trickling filters are either granulated or fixed media. Granulated media are small separate elements (rock, gravel, sand, and plastic elements); fixed media consist of larger constructions with connected areas. Several factors are important in the function of trickling filters, but the most important is selection of the appropriate filter medium. Among the selection criteria for filter media in trickling filters are void ratio, specific surface area, homogeneous water flow, and cost.

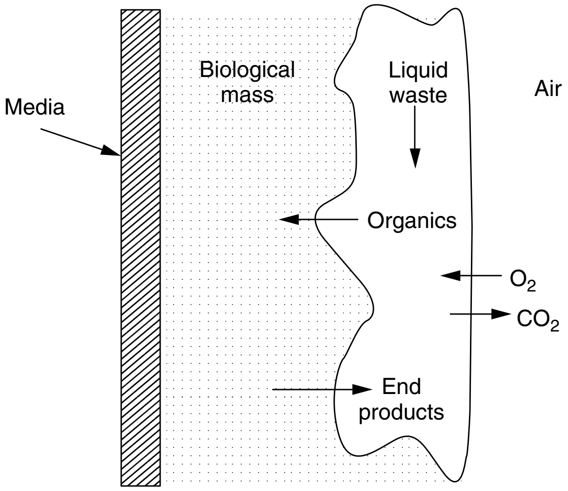

Figure 1. Schematic of the cross section of a biofilm in a trickling filter.

The use of plastic filter media drastically increases the installation cost of a nitrifying trickling filter. Plastic filter media are specially designed to provide high void space and high specific surface area. Thus, they are very attractive for wastewater treatment, where the high organic and ammonia load and the presence of suspended solids leads to the formation of very thick biofilm structures (150–2000 μm) (7, 8). To reduce the cost of the process in this case, we can use large diameter gravel instead of plastic media, resulting in low specific surface area and consequently huge constructions.

For potable water, the ammonia load is rather low (usually 0.5–2 mg/L) almost without any organic load and with a very low concentration of suspended solids. As a result, the biofilm remains very thin (50–200 μm) (9), and the use of small gravel becomes feasible thus resulting in higher specific surface areas and very compact constructions; the use of small gravel enables more area for biofilm growth per volume unit in a trickling filter. Hence, the use of small-sized gravel is very attractive due to the minimal cost of this material and its availability. On the other hand, small-sized gravel leads to small void ratios and all the problems mentioned before.

Quartz sand was the first material used for filtration and is still the basic material in many existing filters. Sand may be combined with anthracite, garnet, and schists of varying porosity, provided that these materials have low friability and low loss in acid (3). Gravels consist mainly of calcitic and silicic salts. For trickling filters, and sand filters, the use of silicic materials is the most common.

INFLUENCE OF RECIRCULATION ON THE EFFICIENCY OF A NITRIFYING TRICKLING FILTER FOR POTABLE WATER TREATMENT

Recirculation in biological filter process is usually considered to reduce nutrient loading on the filter, increase the wettability of the filter, and result in a more even bioconversion along the filter depth. Vayenas and Lyberatos (10) studied the influence of recirculation on the efficiency of a pilot-scale nitrifying trickling filter for potable water treatment. This filter exhibited excellent performance for the ammonia concentrations usually encountered in potable water even for high values of hydraulic loading (9). The pilot-scale filter has, of course, limitations on ammonia and hydraulic loading. For high hydraulic loading (680 m^3/m^2 d), the filter floods; for high ammonia concentration (above 5 mg/L), pore clogging due to bacterial mass becomes a serious problem and the filter requires frequent backwashing.

Experimental results and model predictions showed that for low hydraulic loading (110 m^3/m^2 d), recirculation leads to lower filter efficiency even for high inlet ammonia concentrations. For intermediate hydraulic loading (225 m^3/m^2 d) and for low inlet ammonia concentrations (2 mg/L), recirculation leads to inferior performance, whereas for high inlet ammonia concentration (15 mg/L), the filter performance is improved and the recycle ratio increases. Finally, for high hydraulic loading (450 m^3/m^2 d), recirculation improves the filter efficiency for any values of ammonia concentration and recycle ratio. The higher the recycle ratio, the better the filter performance.

These results are expected because essentially complete conversion is obtained for low loadings. Recirculation evens out ammonia concentrations along the filter depth and leads to higher effluent concentrations. For high loadings, lower overall conversions are expected in the absence of recirculation. In these instances, recirculation leads to better exploitation of the total filter depth. In conclusion, recirculation can improve the efficiency of the filter only for high or intermediate values of hydraulic loading and high ammonia concentrations.

IRON AND MANGANESE INHIBITION OF BIOLOGICAL AMMONIA REMOVAL FROM POTABLE WATER

Water sources that contain ammonia often may also contain iron and manganese. Ammonia, iron, and manganese may be removed chemically or biologically from a water supply. Biological removal of these pollutants is preferable because there is no need to add extra chemicals and the volume of the sludge generated is appreciably smaller and hence easier to handle (3). The simultaneous biological removal of the three elements is very difficult mainly due to the different redox potential values needed to oxidize them. Gouzinis et al. (11) studied the simultaneous removal of ammonia, iron, and manganese from potable water using a pilot-scale trickling filter and the influences these pollutants have on filter performance and efficiency. They concluded that iron has a strong negative effect on ammonia removal whereas ammonia has very low impact on iron removal. They also concluded that for low ammonia concentrations (1 mg/L), the presence of manganese does affect ammonia removal. The influence of manganese becomes sensible at an ammonia concentration of 2 mg/L; at higher ammonia concentrations, manganese strongly affects ammonia removal. The main conclusion of that work was that iron should be removed before ammonia and manganese oxidation. For low iron concentrations, oxidation by extended aeration can reduce iron, and subsequent biological ammonia and manganese removal can be carried out without any problems. For higher iron concentrations, an initial biological iron oxidation stage should be necessary.

MODELING AND DESIGN OF NITRIFYING TRICKLING FILTERS FOR POTABLE WATER TREATMENT

The design of trickling filters is much more difficult than design for homogeneous processes (e.g., activated sludge process) because design engineers must take extra care with respect to the attachment of microorganisms to the surface of the filter medium, the filter medium's influence on the process, nutrient and oxygen transport from the air and the liquid phase through the biofilm structure, biomass detachment, pore clogging and filter flooding, and filter backwash.

The design of biological trickling filters was initially based on empirical or semiempirical models. For many years, design engineers had limited success in applying fairly simple design equations to modeling trickling filter performance. The first approaches were based on NRC equations (12), Eckenfelder's equation (13), or Atkinson's model (14). Later, detailed mathematical steady-state models for describing the biofilm processes were proposed by Vaughan and Holder (15), Benefield and Molz (16), Logan et al. (17) and Gujer and Boller (18). These models were more accurate and offered a better understanding of the mechanisms of biofilm development, but there was still uncertainty in the model parameter values.

Vayenas and Lyberatos (9) developed a simple steady-state model for ammonia removal from potable water that predicts the ammonia concentration profile and the mean biofilm thickness profile along the filter depth. The model considers that the filter may be divided into a number of equal size compartments, for each of which homogeneous conditions may be assumed. The model also considers a somewhat different mechanism for biofilm development because ammonia concentrations in potable water are usually low and sloughing is not so intense as in wastewater treatment. It is worth noting that the model includes analytical recursive equations concerning ammonia concentration and predicts filter performance very well for all operating conditions that were tested. Figure 2 presents a typical ammonia concentration profile along the filter depth and the corresponding model predictions.

The dynamics of nitrifying trickling filters can be divided into fast and slow modes. Slow modes are caused by the growth and decay of the organisms

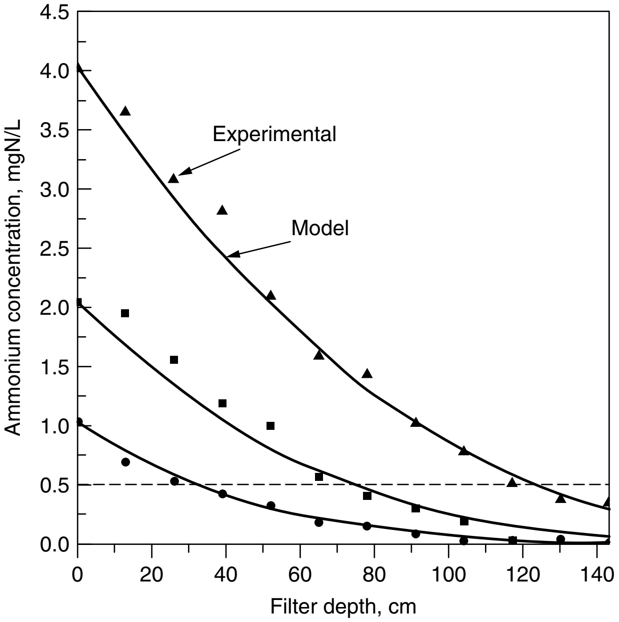

Figure 2. Experimental data and model predictions based on the model of Vayenas and Lyberatos (9) for an ammonia concentration profile in a pilot-scale nitrifying trickling filter for potable water treatment (hydraulic loading 225 m^3/m^2 d) (19).

in the biofilm; fast modes are caused mainly by the hydraulics and mass transfer resistances in the biofilm (20,21). The fast dynamics of nitrifying trickling filters for tertiary wastewater has been described by Kissel et al. (22), Wanner and Gujer (23), and Wik (20,21), and experimental evidence has been presented by Wik (20,21).

Vayenas et al. (24) developed a dynamic model to describe nitritification and nitratification in trickling filters. Their results showed that after an organic or hydraulic loading shock, nitrifying trickling filters reach a new steady state for ammonia-, nitrite-, and nitrate nitrogen concentrations in only a few minutes, whereas biofilm thickness requires a very long time to reach steady state (Fig. 3).

Additionally, Vayenas et al. (24) predicted the biofilm thickness and concentration profiles of ammonia, nitrite, and nitrate inside the biofilm structure (Fig. 4).

Recently, Tekerlekopoulou and Vayenas (25), using the steady-state version of the dynamic model of Vayenas et al. (24), constructed the operating diagram of a nitrifying trickling filter, thus defining the operating conditions for complete and safe nitrification (Fig. 5).

Figure 5 contains the three critical lines of 0.5, 0.1, and 10 mg/L for ammonia-, nitrite-, and nitrate-nitrogen (26). The diagram in (Fig. 5) is the operating diagram of the system and shows the effect of inlet ammonia concentration and hydraulic load on ammonia-, nitrite- and nitrate-nitrogen concentrations. The operating diagram is particularly useful when one wants to compare experimental observations made in a bioreactor with the predictions of a mathematical model describing the system. In the various regions of the operating diagram, different behaviors of the system are observed. In the diagram shown in (Fig. 5), there are six regions in each of which one or more of the ammonia-, nitrite-, and nitrate-nitrogen concentrations is below the maximum permitted limit, as shown below:

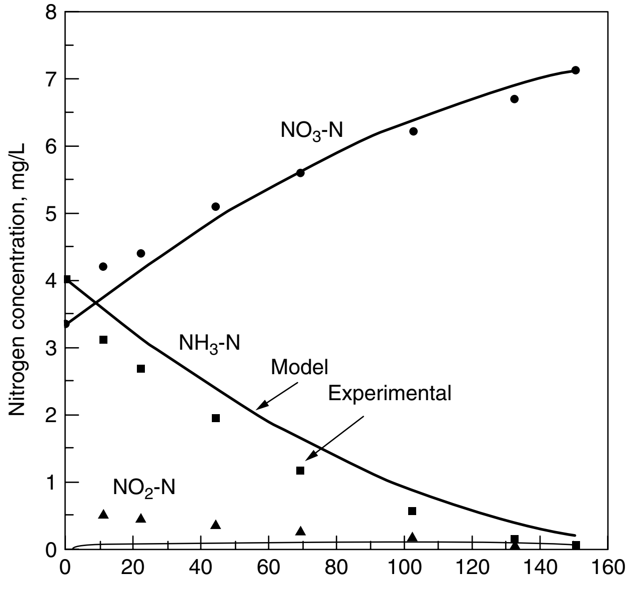

Figure 3. Predicted and experimental concentration profiles of ammonia-, nitrite- and nitrate-nitrogen (hydraulic loading: 225 m^3/m^2 d; feed concentrations: $NH_3 - N = 4$ mg/L, $NO_2 - N = 0$ mg/L, and $NO_3 - N = 3.36$ mg/L) (24).

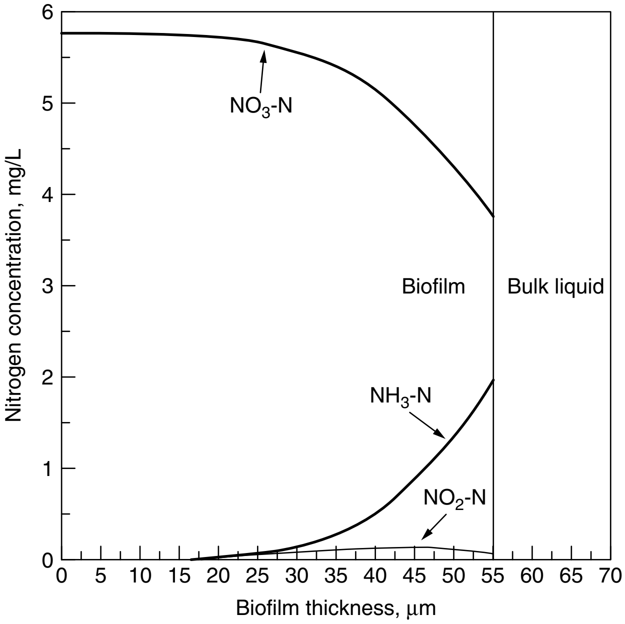

Figure 4. Ammonia-, nitrite-, and nitrate nitrogen concentration profiles along the biofilm thickness near the top of the filter (hydraulic loading: 225 m^3/m^2 d: feed concentrations: $NH_3 - N = 3$ mg/L, $NO_2 - N = 0$ mg/L, and $NO_3 - N = 3.36$ mg/L) (24).

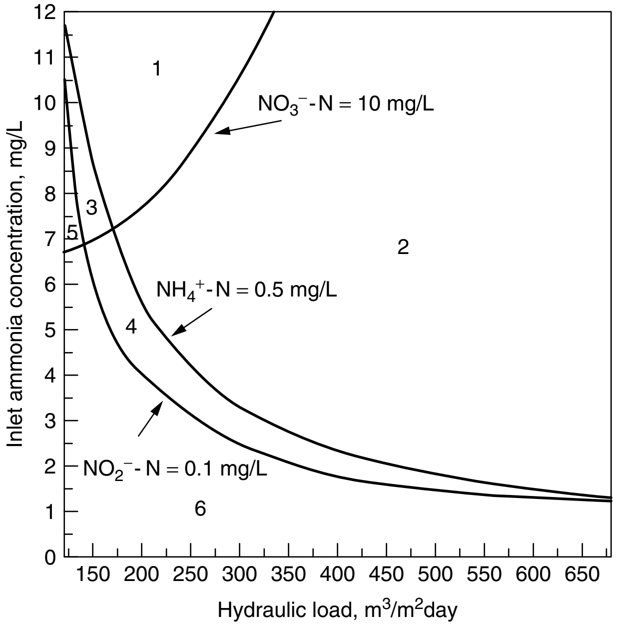

Figure 5. Operating diagram of a pilot-scale nitrifying trickling filter (25).

Region 1. All outlet concentrations are above the maximum permitted limits
Region 2. Outlet $NO^{-}{}_3$-N concentration below the maximum permitted limit
Region 3. Outlet $NH^{+}{}_4$-N concentration below the maximum permitted limit
Region 4. Outlet $NH^{+}{}_4$-N and $NO^{-}{}_3$-N below the maximum permitted limits
Region 5. Outlet $NH^{+}{}_4$-N and $NO^{-}{}_2$-N below the maximum permitted limits
Region 6. All outlet concentrations are below the maximum permitted limits

According to (Fig. 5), to achieve complete nitrification, one should combine the operating conditions so that the system would always remain in Region 6. The operating diagram of the system provides an overview of the process limitations where the areas of safe/complete nitrification operation are defined. According to the operating diagram, complete nitrification at low flow rates can be achieved even for ammonia feed concentrations up to about 6.5 mg/L, whereas at high flow rates, complete nitrification is possible only for ammonia feed concentrations up to about 2 mg/L.

BIBLIOGRAPHY

1. Duddles, A.G., Richardson, E.S., and Barth, I.E. (1974). Plastic medium trickling filters for biological nitrogen control. *J. Water Pollut. Control Fed.* **46**: 937–946.
2. Bull, R.J., Gerba, C., and Trussel, R.R. (1990). Evaluation of the health risks associated with disinfection. *Crit. Rev. Environ. Control* **20**: 77–114.
3. Degremont (1991). *Water Treatment Handbook*, 6th Edn. Lavoisier, Paris.
4. Metcalf and Eddy. (1991). *Wastewater Engineering, Treatment, Disposal, Reuse*, 3rd Edn. McGraw-Hill, New York, pp. 404–406.
5. Krüner, G. and Rosenthal, H. (1983). Efficiency of nitrification in trickling filters using different substrates. *Aquaculture Eng.* **2**: 49–67.
6. Lekang, O. and Kleppe, H. (2000). Efficiency of nitrification in trickling filters using different filter media. *Aquaculture Eng.* **21**: 181–199.
7. Gullicks, H.A. and Cleasby, J.L. (1986). Design of trickling filter nitrification towers. *J. Water Pollut. Control Fed.* **58**(1): 60–67.
8. Hao, O.J., et al. (1991). Biological fixed-film systems. *Res. J. Water Pollut. Control Fed.* **63**(4): 388–394.
9. Vayenas, D.V. and Lyberatos, G. (1994). A novel model for nitrifying trickling filters. *Water Res.* **28**(6): 1275–1284.
10. Vayenas, D.V. and Lyberatos, G. (1995). On the design of nitrifying trickling filters for potable water treatment. *Water Res.* **29**(4): 1079–1084.
11. Gouzinis, A., Kosmidis, N., Vayenas, D.V., and Lyberatos, G. (1998). Removal of Mn and simultaneous removal of NH_3, Fe and Mn from potable water using a trickling filter. *Water Res.* **32**(8): 2442–2450.
12. National Research Council. (1946). Sewage treatment at military installations. Report of the subcommittee on sanitary engineering. *Sewage Works J.* **18**(5): 787.
13. Eckenfelder, W.W. (1963). Trickling filtration design and performance. *Trans. Am. Soc. Civ. Eng.* **128**: 394–398.
14. Atkinson, B., Davies, I.J., and How, S.Y. (1974). The overall rate of substrate uptake by microbial films, parts I and II. *Trans. Inst. Chem. Eng.* **52**: 248–260.
15. Vaughan, G.M. and Holder, G.A. (1984). Substrate removal in the trickling filter process. *J. Water Pollution Control Fed.* **56**: 417–423.
16. Logan, E.B., Hermanowich, W.S., and Parker, S.D. (1987). A fundamental model for trickling filter process design. *J. Water Pollut. Control Fed.* **59**: 1029–1042.
17. Benefield, L. and Molz, F. (1984). Mathematical simulation of a biofilm process. *Biotechnol. Bioeng.* **27**: 921–931.
18. Gujer, W. and Boller, M. (1986). Design of a nitrifying tertiary trickling filter based on theoretical concepts. *Water Res.* **20**(11): 1353–1362.
19. Tekerlekopoulou, A.G., Dermou, E., and Vayenas, D.V. (2004). *The effect of support material and ammonia shock load on nitrifying trickling filters performance for potable water treatment* (submitted).
20. Wik, T. (1999). Adsorption and denitrification in nitrifying trickling filters. *Water Res.* **33**(6): 1500–1508.
21. Wik, T. (1999). Rational transfer function models for nitrifying trickling filters. *Water Sci. Technol.* **39**(4): 121–128.
22. Kissel, J.C., McCarty, P.L., and Street, R.L. (1984). Numerical simulation of mixed culture biofilm. *J. Environ. Eng.* **110**(2): 393–411.
23. Wanner, O. and Gujer, W. (1986). A multispecies biofilm model. *Biotechnol. Bioeng.* **28**: 314–328.
24. Vayenas, D.V., Pavlou, S., and Lyberatos, G. (1997). Development of a dynamic model describing nitrification and nitratification in trickling filters. *Water Res.* **31**(5): 1135–1147.
25. Tekerlekopoulou, A.G. and Vayenas, D.V. (2003). Operational and design considerations of a trickling filter for ammonia removal from potable water. *Environ. Modelling Assessment* **8**(2): 55–62.
26. *EEC-Official Journal of the European Communities*, No 80/779, August 30, 1980.

ORGANIC REMOVAL

National Drinking Water
Clearinghouse

Some small drinking water systems face contamination of raw water by natural or synthetic organic chemicals (SOCs). Natural organic materials might be present in water supplies, especially from surface water sources. Dissolved organics may cause taste, odor, or color problems in a community's drinking water, resulting in consumer complaints. Sources of SOCs include leaking underground gasoline/storage tanks, agricultural runoff containing herbicides or pesticides, solid waste or hazardous waste landfills, and improperly disposed chemical waste. The technologies most suitable for organic contaminant removal in drinking water systems are granular activated carbon (GAC) and aeration.

GAC has been designated by the U.S. Environmental Protection Agency (EPA) as the best available technology (BAT) for synthetic organic chemical removal. Various kinds of GAC are available for removing organics from drinking water. The most frequently used carbon in U.S. treatment plants is coal-based carbon because of its hardness, adsorption capacity, and availability. Some peat and lignite carbons have been used also.

Aeration systems that might be suitable for small drinking water systems include packed column aeration, diffused aeration, and multiple-tray aeration. Recent technologies that use aeration for organics removal include mechanical aeration, catenary grid, and Higee aeration.

Table 1 presents operational conditions for the organics treatment technologies most suitable for small systems.

SOME TECHNOLOGIES FOR ORGANIC REMOVAL IN SMALL SYSTEMS:

A. Activated Carbon

Description. Activated carbon is carbon that has been exposed to very high temperatures, creating a vast network of internal pores.

Two types of activated carbon, granular and powdered, have been used widely in drinking water treatment. Powdered activated carbon (PAC), which is most often used for taste and odor control, is added directly to the raw water and removed by settling in sedimentation basins.

GAC removes many organic contaminants as well as taste and odor from water supplies.

Performance/Advantages. Organics that are readily adsorbed by activated carbon include:

- aromatic solvents (benzene, toluene, nitrobenzenes);
- chlorinated aromatics (PCBs, chlorobenzenes, chloroaphthalene);
- phenol and chlorophenols;
- polynuclear aromatics (acenaphthene, benzopyrenes);
- pesticides and herbicides (DDT, aldrin, chlordane, heptachlor);
- chlorinated aliphatics (carbon tetrachloride, chloroalkyl ethers); and
- high molecular weight hydrocarbons (dyes, gasoline, amines, humics).

Table 1. Organic Treatment Technologies Suitable for Small Systems

Technology	Level of Operational Skill Required	Level of Maintenance Required	Energy Requirements
Granular Activated Carbon (GAC)	Medium	Low	Low
Packed Column Aeration (PCA)	Low	Low	Varies
Diffused Aeration	Low	Low	Varies
Multiple-Tray Aeration	Low	Low	Low
Mechanical Aeration	Low	Low	Low
Catenary Grid	Low	Low	High
Higee Aeration	Low	Medium	High

Source: U.S. Environmental Protection Agency, 1989.

Limitations. Organics that are poorly adsorbed by activated carbon include:

- alcohols;
- low molecular weight ketones, acids, and aldehydes;
- sugars and starches;
- very high molecular weight or colloidal organics; and
- low molecular weight aliphatics.

GAC is not effective in removing vinyl chloride from water. In addition, because of the long empty bed contact time (EBCT) required, radon removal at the treatment plant scale is not feasible. However, at the residential scale, GAC systems are cost-effective for radon removal.

Several operational and maintenance factors affect the performance of GAC. Contaminants in the water can occupy GAC adsorption sites, whether they are targeted for removal or not. Also, adsorbed contaminants can be replaced by other contaminants with which GAC has a greater affinity. Therefore, the presence of other contaminants might interfere with the removal of the contaminants of concern.

A significant drop in the contaminant level in influent water will cause a GAC filter to desorb, or slough off, adsorbed contaminants because GAC is an equilibrium process. As a result, raw water with frequently changing contaminant levels can result in treated water of unpredictable quality.

Bacterial growth on the carbon is another potential problem. Excessive bacterial growth may cause clogging and higher bacterial counts in the treated water. Bacterial levels in the treated water must be closely monitored, and the final disinfection process must be carefully controlled.

Process. Activated carbon removes contaminants through adsorption, primarily a physical process in which dissolved contaminants adhere to the porous surface of the carbon particles. The adsorption process can be reversed relatively easily. The ease of reversing adsorption is another key factor in activated carbon's usefulness because it facilitates the recycling or reuse of the carbon.

GAC can be used as a replacement for existing media (such as sand) in a conventional filter, or it can be used in a separate contactor (a vertical steel pressure vessel used to hold the activated carbon bed).

GAC contactors require monitoring to ensure that they work effectively. A GAC monitoring system should include:

- laboratory analysis of treated water to ensure that the system is removing organic contaminants,
- monitoring of headloss (the amount of energy used by water in moving from one point to another) through the contactors to ensure that backflushing (reversing

the flow to remove trapped material) is performed at appropriate times,

- bacteria monitoring of the contactor's effluent (since bacteria can grow rapidly within the activated carbon bed),
- turbidity monitoring of the contactor's effluent (to determine if suspended material is passing through GAC bed).

After a period of months or years, depending on the concentration of contaminants, the surface of the pores in the GAC can no longer adsorb contaminants. The carbon must then be replaced. The GAC vendor will be able to provide guidance concerning when to replace the GAC. Disposing of carbon with contaminants that are classified as hazardous wastes will dramatically increase disposal costs.

Equipment/Design. The typical GAC unit can be similar in design to either gravity or pressure filters. In some communities, the sand in existing filters has been either partially or completely replaced with GAC. Media depth of up to 10 feet is needed to ensure adequate removal of potentially harmful organic contaminants. Activated carbon filters can be designed to treat hydraulic loadings of 2 to 10 gallons per minute per square foot (gpm/ft^2). Sufficient detention time in the filter must be provided to achieve the desired level of the organic contaminant removal. The detention time is determined by the volume of the GAC filter divided by the flow rate. This is referred to as the EBCT since the volume of carbon in the bed is not considered. For adequate removal of most organic contaminants to occur, the EBCT should be about 10 minutes. EBCTs less than 7.5 minutes are generally ineffective.

GAC is available in different grades of effectiveness. Low-cost carbon requires a lower initial capital outlay but must be replaced more often, resulting in higher operating costs.

B. Aeration

Description. Aeration, also known as air stripping, mixes air with water to volatilize contaminants (turn them to vapor). The volatilized contaminants are either released directly to the atmosphere or treated and released. Aeration is used to remove volatile organic chemicals and can also remove radon.

Equipment. A small system might be able to use a simple aerator constructed from relatively common materials instead of a specially designed aerator system. Examples of simple aerators include:

- a system that cascades the water or passes it through a slotted container,
- a system that runs water over a corrugated surface, or
- an airlift pump that introduces oxygen as water is drawn from a well.

Other Aeration Types

Packed Column Aeration. Packed column aeration (PCA) or packed tower aeration (PTA) is a waterfall aeration process that drops water over a medium within a tower to mix the water with air. The medium is designed to break the water into tiny droplets and to maximize its contact with tiny air bubbles for removal of the contaminant. Air is also blown in from underneath the medium to enhance this process (see Fig. 1).

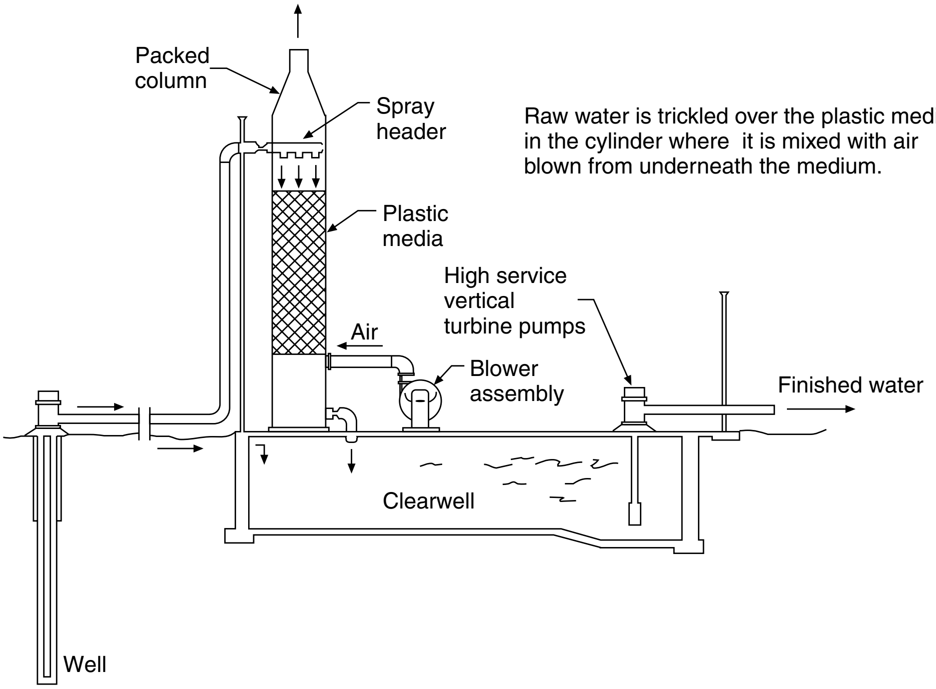

Figure 1. Packed tower aeration system.

Systems using PCA may need pretreatment to remove iron, solids, and biological growth to prevent clogging of the packing material. Post treatment such as the use of a corrosion inhibitor, may also be needed to reduce corrosive properties in water due to increased dissolved oxygen from the aeration process.

Packed columns usually operate automatically and need only daily visits to ensure that the equipment is running satisfactorily. Maintenance requirements include servicing pump and blower motors and replacing air filters on the blower, if necessary.

PCA exhaust gas may require treatment to meet air emissions regulations, which can significantly increase the costs of this technology.

Diffused Aeration. In a diffused aeration system, a diffuser bubbles air through a contact chamber for aeration. The diffuser is usually located near the bottom of the chamber. The air introduced through the diffuser, usually under pressure, produces fine bubbles that create water-air mixing turbulence as they rise through the chamber.

The main advantage of diffused aeration systems is that they can be created from existing structures, such as storage tanks. However, they are less effective than packed column aeration, and usually are employed only in systems with adaptable existing structures.

Multiple Tray Aeration. Multiple tray aeration directs water through a series of trays made of slats, perforations, or wire mesh. A blower introduces air from underneath the trays.

Multiple tray aeration units have less surface area than PCA units. This type of aeration is not as effective as PCA and can experience clogging from iron and manganese, biological growth, and corrosion problems.

Multiple tray aeration units are readily available from package plant manufacturers.

Mechanical Aeration. Mechanical aeration uses mechanical stirring mechanisms to mix air with the water. These systems can effectively remove volatile organic chemicals (VOCs).

Mechanical aeration units need large amounts of space because they demand long detention times for effective treatment. As a result, they often require open-air designs, which can freeze in cold climates. These units also can have high energy requirements. However, mechanical aeration systems are easy to operate and are less susceptible to clogging from biological growth than PCA systems.

Catenary Grid. Catenary grid systems are a variation of the packed column aeration process. The catenary grid directs water through a series of wire screens mounted within the column. The screens mix the air and water in the same way as packing materials in PCA systems.

These systems can effectively remove VOCs. They have higher energy requirements than PCA systems, but their more compact design lowers their capital cost relative to PCA.

Higee Aeration. Higee aeration is another variation of the PCA process. These systems pump water into the center of a spinning disc of packing material, where the water mixes with air.

Higee units require less packing material than PCA units to achieve the same removal efficiencies. Because of their compact size, they can be used in limited spaces and heights. Current Higee systems are best suited for a temporary application of less than 1 year with capacities up to 380 liters (100 gallons) per minute.

WHERE CAN I FIND MORE INFORMATION?

Information in this fact sheet was primarily obtained from three documents:

- *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*. EPA/625/4-89/023;
- *Small Community Water and Wastewater Treatment*, EPA/625/R-92/010; *and*
- *Environmental Pollution Control Alternatives: Drinking Water Treatment for Small Communities*, EPA/625/5-90/025.

All publications can be ordered free from the U.S. Environmental Protection Agency Office of Research and Development at (513) 569-7562.

The NDWC offers these documents as well, but at a cost to help recover photocopying and other expenses. *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*, item #DWBKDM04, a 209-page book, costs $30.05. *Small Community Water and Wastewater Treatment*, item #DWBKRE03, a 95-page book, costs $ 13.65. The third book, *Environmental Pollution Control Alternatives: Drinking Water Treatment for Small Communities*, item #DWBKGN09, an 82-page publication, costs $11.82. Shipping and handling charges apply.

Also, the NDWC's Registry of Equipment Suppliers of Treatment Technologies for Small Systems (RESULTS) is a public reference database that contains information about technologies in use at small water systems around the country. For further information on accessing or ordering RESULTS, call the NDWC at (800) 624-8301 or (304) 293-4191.

For additional free copies of Tech Brief fact sheets *call the NDWC at one of the numbers above. You may also download Tech Briefs from our Web site at* http://www.ndwc.wvu.edu.

- Tech Brief: Organic Removal, *item #DWBLPE59*;
- Tech Brief: Ion Exchange and Demineralization, *item #DWBLPE56*;
- Tech Brief: Corrosion Control, *item #DWBLPE52*;
- Tech Brief: Filtration, *item #DWBLPE50*; or
- Tech Brief: Disinfection, *item #DWBLPE47*.

OZONE

National Drinking Water
Clearinghouse

New water treatment goals for disinfection byproducts (DBP) and for microbial inactivation will increase the need to consider new disinfection technologies. Ozone is an attractive alternative. This technology has evolved and improved in recent years, thereby increasing its potential for successful application. It is important to note that ozone, like other technologies, has its own set of advantages and disadvantages that show up in differing degrees from one location to the next

WHAT IS OZONE?

Ozone is a colorless gas that has an odor most often described as the smell of air after a spring electrical thunderstorm. Some people also refer to the odor as similar to the smell of watermelons. Ozone (O_3) is an extremely unstable gas. Consequently, it must be manufactured and used onsite. It is the strongest oxidant of the common oxidizing agents. Ozone is manufactured by passing air or oxygen through two electrodes with high, alternating potential difference.

WHAT REGULATIONS GOVERN OZONE?

In the next century, small water systems will need to use disinfection methods that are effective for killing pathogens without forming excessive DBP. Disinfection byproduct compliance is more likely to be a problem for small water systems treating surface water than for those treating groundwater because surface water sources tend to contain more natural organic matter that forms byproducts when mixed with disinfectants. In August 1997, the U.S. Environmental Protection Agency (EPA) listed ozone as a “compliance” in the requirements of the Surface Water Treatment Rule for all three sizes of small drinking water systems.

ADVANTAGES

Using ozone to treat water has many advantages, including the following:

- Possesses strong oxidizing power and requires short reaction time, which enables the germs, including viruses, to be killed within a few seconds;
- Produces no taste or odor;
- Provides oxygen to the water after disinfecting;
- Requires no chemicals;
- Oxidizes iron and manganese;
- Destroys and removes algae;
- Reacts with and removes all organic matter;
- Decays rapidly in water, avoiding any undesir able residual effects;
- Removes color, taste, and odor; and
- Aids coagulation.

LIMITATIONS

The use of ozone to treat water has some limitations:

- Toxic (toxicity is proportional to concentration and exposure time);
- Cost of ozonation is high compared with chlorination;
- Installation can be complicated;
- Ozone-destroying device is needed at the exhaust of the ozone-reactor to prevent toxicity and fire hazards;
- May produce undesirable aldehydes and ketones by reacting with certain organics;
- No residual effect is present in the distribution system, thus postchlorination is required;
- Much less soluble in water than chlorine; thus special mixing devices are necessary; and
- It will not oxidize some refractory organics or will oxidize too slowly to be of practical significance.

DISINFECTION

Design of an ozone system as primary treatment should be based on simple criteria, including ozone contact concentrations, competing ozone demands, and a minimum contact time (CT) to meet the required cyst and viral inactivation requirements, in combination with EPA recommendations.

Systems that need to provide CT to comply with the Ground Water Disinfection Rule, but are also having problems with DBP or maintaining distribution system residuals, should consider using ozone as the primary disinfectant and then chloramines for distribution system protection.

Ozone has been observed to be capable of disinfecting *Cryptosporidium*, and there is significant interest in this aspect of its application. Available data indicate that a significant increase in ozone dose and CT may be required as compared with past practices. Therefore, these needs should be considered in planning.

Iron and Manganese Removal

The standard oxidation-reduction potential and reaction rate of ozone is such that it can readily oxidize iron and manganese in groundwater and in water with low organic content. Groundwater systems that have iron levels above 0.1 milligrams per liter (mg/L) may have iron complaints if ozonation or chlorination is added.

Excessive doses of ozone will lead to the formation of permanganate, which gives water a pinkish color. This soluble form of manganese (Mn) corresponds to a theoretical stoichiometry of 2.20 mg O_3/mg Mn. Stoichiometry is the determination of the proportions in which chemical elements combine or are produced and the weight relation in a chemical reaction.

Color Abatement

Because humic substances are the primary cause of color in natural waters, it is useful to review the reactions of ozone with humic and fulvic acids. According to different authors, ozone doses of 1 to 3 mg O_3/mg C lead to almost complete color removal. The ozone dosages to be applied in order to reach treatment goals for color can be very high. It is interesting to note that when the ozone dosage is sufficient, the organic structure is modified such that the final chlorine demand can decrease.

Control of Taste and Odor

The National Secondary Drinking Water Regulations recommend that the threshold odors number (TON) be 3 or less in finished water. It has been shown that ozone can be effective in treating water for taste and odor problems, especially when the water is relatively free from radical scavengers.

It has also been observed that ozone, in combination with other downstream treatment processes, especially granular activated carbon (GAC) filtration, can greatly increase taste and odor treatment efficiency and reliability. Again, the cause of taste and odor compounds, as well as the source water to be treated, need to be carefully considered prior to designing a treatment system. Analysis and possibly pilot-scale experimentation may be required to determine the optimum choice of ozone and downstream treatment.

Elimination of Synthetic Organic Chemicals

Ozone or advanced ozonation processes can remove many synthetic organic chemicals (SOC). This removal leads to the chemical transformation of these molecules into toxic or nontoxic byproducts. Such transformation can theoretically lead to complete oxidation into carbon dioxide (CO_2); however, this is rarely the case in water treatment. Any observable reduction in total organic carbon (TOC) is due either to a small degree of CO_2 formation (for example, decarboxylation of amino acids) or the formation and loss of volatile compounds through stripping.

Effects on Coagulation

It is important to understand that the coagulating effects of ozone go beyond any direct oxidative effects on organic macro-pollutants. For this reason, one must be wary of studies claiming improved removal of organic matter when the data are based solely on color removal or ultraviolet (UV) absorbtion. Also, when studying the removal of DBP (for example, trihalomethanes), one must be careful to incorporate controls permitting the separate evaluation of ozone's direct effects. Finally, the coagulating effects of ozone may not be observed with all water. Whenever considering the use of ozone as a coagulant aid, the pre-ozonation effects should be critically evaluated in pilot studies incorporating the proper controls.

Algae Removal

Ozone, like any other oxidant, such as chlorine or chlorine dioxide, has a lethal effect on some algae or limits its growth. Ozone is also capable of inactivating certain zooplankton, e.g., mobile organisms, *Notholca caudata.* Such organisms must first be inactivated before they are removed by flocculation and filtration.

Byproducts

The alternative use of ozonation has generated much interest because of its ability to avoid the formation of halogenated organics inherent in the practice of chlorine treatment. However, raw water quality significantly affects ozonation results and could lead to the formation of other undesirable byproducts. Brominated byproducts are a major concern in source waters containing bromide. Ozonation produces its own byproducts, such as aldehydes, ketones, and carboxylic acids.Assuming equivalent disinfection, benefit is achieved as long as the health concerns for the new products are less than those for the chlorine byproducts.

Personnel Requirements

Personnel time requirements for system cleaning may be fairly substantial. However, recent advancements in ozonation technology include use of high purity oxygen feed systems, rather than ambient air-feed systems. Ozonation treatment is therefore said to run cleaner and require less cleaning-related maintenance than had the earlier versions of this technology.

No Residual

Ozone will not provide a disinfecting residual that protects finished water in the distribution system. Therefore, the role of chlorine as a disinfecting agent is not entirely replaced, and its use in either the free chlorine or chloramine form will be required for this purpose in many locations.

Process and Equipment

The basic elements of an ozone system include ozone generation, feed gas preparation, ozone contacting, and ozone off-gas destruction components. While many of these components can involve a high degree of sophistication in large facilities, less complex alternatives are available for smaller systems. Figure 1 (see below) shows the five basic components of an ozonation system. To insure effectiveness and safety simultaneously, all components must be taken into account when designing/installing an ozone system. Central to the ozonation system is the ozone generator itself, which in turn is connected to an appropriate power supply. Instrumentation and controls for ensuring the effective and safe operation of the total system may be added to the five-component system shown in (Fig. 1).

Feed Gas Preparation

The feed gas preparation component is critical, as a high-quality gas stream is required for the generator to

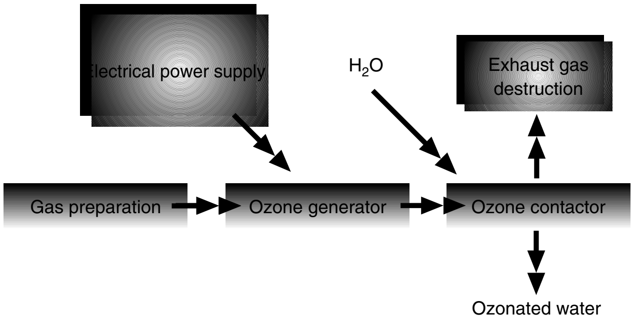

Figure 1. Basic components of an ozonation system. *Source*: "Ozone Treatment of Small Water Systems" 1999.

perform properly. This requires a gas stream that is low in moisture and particles. In older air-feed systems, the feed gas preparation systems for small systems tended to rely on high-pressure compressors that produce a pressurized feed stream, which is easier to dry and can accom- modate simpler, less maintenance-intensive drying devices than were typically applied in larger systems where lower-pressure compressors are typically used.

Alternatives for application of oxygen-rich feed gas streams have also emerged for small system applications in recent years. These include purchasing liquid oxygen and using small package oxygen-generation systems that are based on pressure swing adsorption. The use of these oxygen-enriched feed streams allows improved electrical efficiency in ozone generation and yields a product stream with a higher ozone content.

Ozone Contactors

Since ozone is only partially soluble in water, once it has been generated it must contact water to be treated. Many types of ozone contactors have been developed for this purpose.

Ozone contacting for disinfection has typically been accomplished in deep, multistage contactors that employ fine bubble diffusers. Newer alternatives have emerged that provide an option for small systems through the use of side-stream injection technologies that eliminate the need for fine bubble injection. Package units are available that include a gas separator that fuses and eliminates excess gas that results from ozone addition and a venturi jet that is used to inject and blend the ozone with a solution feed stream. These systems allow the alternative of injecting ozone into an enclosed vessel or a pipe. Note that materials need to be compatible with ozone. One possibility in this regard is the use of stainless steel. Several other contacting configurations, including turbine mixers, have been developed and may provide benefits as well.

Exhaust Gas Destruction

Ozone off-gas destruction is the final major component in the ozonation process. This system is required to remove ozone from spent off-gas streams, which are collected and treated prior to discharge into the atmosphere. Both catalytic and thermal destruction devices are used for this purpose or by passage through GAC.

How Safe is Ozone?

EPA notes that ozonation technology requires careful monitoring for ozone leaks, which pose a hazard in the work place. As with any other chemical, the Occupational Health and Safety Administration (OSHA) has established maximum contaminant inhalation guidelines for ozone in the work place. Ozone concentration of 0.1 part per million inhaled during an eight hour work period in a work area is the maximum limit set by OSHA regulations.

WHERE CAN I FIND MORE INFORMATION?

(1) American Water Works Association. 1993. *Controlling Disinfection By-Products*. Denver: American Water Works Association.

(2) Budd, G.C., G.S. Logdson, and B.W. Long. 1999. "Overview of Chlorine Dioxide, Ozone and Ultraviolet Irradiation." J.A. Cotruvo, G.F. Craun, and N. Hearne, eds. *Providing Safe Drinking Water in Small Systems: Technology, Operations, and Economics.* Boca Raton: CRC Press LLC.

(3) Collins, M.R. 1998. *Small Systems Water Treatment Technologies: State-of-the-Art Workshop*. Denver: American Water Works Association.

(4) International Ozone Association. 1999. Regional Conference on Ozonation and Advanced Oxidation Processes (AOPs) in Water Treatment. *Applications and Research in Poitiers, France. Stamford: International Ozone Association.*

(5) Langlais, B., D.A. Reckhow, and D.R. Brink. 1991. *Ozone in Water Treatment: Application and Engineering*. Denver: AWWA Research Foundation and Lewis Publishers.

(6) Rice, R.G., P.K. Overbeck, and K. Larson. 1999. "Ozone Treatment of Small Water Systems." J. A. Cotruvo, G. F. Craun, and N. Hearne, eds. *Providing Safe Drinking Water in Small Systems: Technology, Operations, and Economics*. Boca Raton: CRC Press LLC.

(7) Vigneswaran, S. and C.Visvanathan. 1995. *Water Treatment Processes: Simple Options*. Boca Raton: CRC Press, Inc.

(8) U.S. Environmental Protection Agency. 1998. *Small System Treatment Technologies for Surface Water and Total Coliform Rules.* Washington, DC: EPA Office of Ground Water and Drinking Water.

HAVE YOU READ ALL OUR TECH BRIEFS?

"Tech Briefs," drinking water treatment fact sheets have been a regular feature in the National Drinking Water Clearinghouse (NDWC) newsletter *On Tap* for more than four years. NDWC Technical Assistance Specialist Mohamed Lahlou, Ph.D.,researches, compiles information, and writes these very popular items.

- *Tech Brief:* Disinfection, item #DWBLPE47;
- *Tech Brief:* Filtration, item #DWBLPE50;

- *Tech Brief:* Corrosion Control, item #DWBLPE52;
- *Tech Brief:* Ion Exchange and Demineralization, item #DWBLPE56;
- *Tech Brief:* Organics Removal, item #DWBLPE59;
- *Tech Brief:* Package Plants, item #DWBLPE63;
- *Tech Brief:* Water Treatment Plant Residuals Management, item #DWBLPE65;
- *Tech Brief:* Lime Softening, item #DWBLPE67;
- *Tech Brief:* Iron and Manganese Removal, item #DWBLPE70;
- Water Conservation Measures Fact Sheet, item #DWBLPE74;
- *Tech Brief:* Membrane Filtration, item #DWBLPE81;
- *Tech Brief:* Treatment Technologies for Small Drinking Water Systems, item #DWPSPE82;
- *Tech Brief:* Ozone, item #DWBLPE84.

Additional copies of fact sheets are free; however, postal charges may be added.

To order, call the NDWC
You may also order online at
ndwc_orders@mail.estd.wvu.edu
at (800) 624-8301 or (304) 293-4191.

or download fact sheets from our Web site at www.ndwc.wvu.edu.

For further information, comments about this fact sheet, or to suggest topics, call Lahlou at one of the above numbers or contact him via e-mail at mlahlou2@wvu.edu.

OZONE WITH ACTIVATED CARBON FOR DRINKING WATER TREATMENT

Rasheed Ahmad
Khafra Engineering Consultants
Atlanta, Georgia

The removal of organic compounds by adsorption on activated carbon is very important in water purification. Removal of organic compounds from the drinking water supply results in reduction in taste and odor and in DBP precursors and removal of volatile organic compounds (VOCs) and synthetic organic compounds (SOCs). Activated carbon supports a higher rate of biodegradation than sand or nonactivated carbon. The advantages of activated carbon include (1) enrichment of oxygen by sorption, (2) enrichment of the substrate in a biofilm, and (3) extended adsorption resulting from bioregeneration and development of a biofilm that can degrade less biodegradable but adsorbable organics.

In recent years, there has been a lot of emphasis on finding alternative disinfectants to free chlorine because of the carcinogenic nature of the halogenated disinfection byproducts (DBPs). Ozone has come up as an alternative disinfectant because of its distinct beneficial effects in water purification. However, ozonation raises two design issues. First, ozone produces byproducts that may eventually be regulated by environmental agencies; these include aldehydes, organic acids, and peroxides. Second, ozone increases the biodegradability of natural organic matter (NOM), which can lead to bacterial regrowth in distribution systems. The engineering solution to both problems has been to recommend that ozonation precede filtration. In this way, the filter can be used to promote biodegradation of NOM and ozonation byproducts; the process has been referred to as biological filtration. The filter could be either an adsorbing medium, the most common of which is granular-activated carbon (GAC), or a nonadsorbing medium, typically a combination of anthracite and sand. If activated carbon is used, the interaction between adsorption and biodegradation becomes important. Activated carbon removes contaminants by adsorption, whereas a microbial population removes biodegradable components. It is observed frequently that adsorbability is affected negatively by ozonation but positively by biodegradation.

Activated carbon coupled with an ozone pretreatment can be used in a treatment process either in a filter–adsorber or postfilter–adsorber configuration. Filter-adsorbers basically integrate filtration and adsorption processes into a single process. In other words, a filter–adsorber is a dual medium filter in which the top layer is GAC and the bottom layer is sand. Postfilter-adsorbers are simply granular-activated carbon columns that are employed after a conventional filtration process. Both filter–adsorbers and postfilter–adsorbers become biologically active when ozone precedes the treatment process. An ozone-activated carbon filtration system (e.g., a filter-adsorber) will enable removal of turbidity, organics, pesticides, and organic micropollutants. It will be biologically stable and can help in preventing bacterial regrowth in distribution systems. Postfilter–adsorbers normally do not remove turbidity because they are designed for organics removal.

OZONE–BROMIDE INTERACTIONS

M. Siddiqui
University of Utah
Salt Lake City, Utah

As utilities consider changing their primary disinfection practices from free chlorine to ozone to provide pathogen control while minimizing chlorinated disinfection by-product (DBP) formation, the potential for producing ozonation DBPs like bromate ion must be addressed. This article focuses on the interactions, not only of molecular ozone and bromide ion, but also of radical species with bromide ion. The data evaluations include discussions of reaction products, potential for bromate ion formation, reaction kinetics, and the effect of various water quality parameters and treatment conditions on bromate ion formation. Based on evaluation of existing data, areas requiring new or intensified research efforts are identified.

An additional aspect of this article is a presentation of most available data on the minimization of bromate ion formation during ozonation. A summary of bromate ion removal processes after ozonation has been incorporated.

SOURCES AND OCCURRENCE OF BROMIDE ION

Bromide ion (Br^-) in drinking water supplies has no direct public health ramifications; however, its presence during water-treatment disinfection can lead to the formation of DBPs such as bromate ion (BrO_3^-). Hence, bromide ion is a precursor to the formation of bromate ion and other brominated oxidation/disinfection by-products. An understanding of the sources and levels of bromide ion in different source waters is crucial for an understanding of the bromate ion formation potential in drinking waters. If bromide ion is present in source waters, it is difficult to remove it economically during drinking water treatment.

Bromide ion can enter water sources from dissolution of geologic sources, from saltwater intrusion into surface water and aquifers, and from human activities. Methyl bromide and ethylene dibromide are used to fumigate crops and soil and as additives to leaded gasoline. Once in soil, they break down into inorganic forms that subsequently are leached by agricultural runoff into ground waters; bromide ion concentrations may attain 0.01–6 mg/L (1). Bromide ion is also associated with salt-spreading on roads during the winter, and some fertilizers are known to contain iodate and bromate ions. Other bromide ion containing compounds can enter water through sewage and industrial effluents (2). Electrolytic chlorine obtained from salt contains bromine estimated at a level of 2 mg/L (3).

Amy et al. (4) conducted a nationwide survey of bromide ion in source waters throughout the United States, and found an average concentration of almost 100 μg/L. In a study of more than 35 water treatment facilities that were selected to provide a broad range of source water qualities and treatment processes, influent bromide ion levels ranged from 10–3000 μg/L (5). Legube et al. (6) found bromide ion concentrations in the influents of 23 different European utilities ranging from 12–210 μg/L, and Kruithof et al. (7) found bromide ion concentrations in Dutch waters ranging from 10–240 μg/L. Considering that the average bromide ion concentration in U.S. waters is approximately 100 μg/L, it is expected that detectable bromate ion can form in a majority of waters that are subjected to ozonation. A Cl^-/Br^- ratio of about 300:1 is indicative of seawater influence.

MEASUREMENT OF BROMATE ION

Ion chromatography is the predominant analytical technique for bromate ion measurement at such low concentrations. Several techniques have been described to reach the required level of sensitivity for measuring bromate ion (8–11).

Siddiqui et al. (11), Legube et al. (6), and Kuo et al. (9) measured bromate ion with detection limits of 2 μg/L, 2 μg/L, and 5 μg/L, respectively, using a conductivity detector without preconcentration and employing a borate eluent (100–150 μL sample loop). They also employed Ag cartridges to remove chloride prior to injection. Legube et al. (6) followed a sequence of steps to measure bromate ion. Each sample was first filtered using membrane filters to remove suspended solids, then filtered through Ag filters to remove Cl^-, and finally was acidified and stripped of bicarbonate as carbon dioxide by helium to avoid interference from the bicarbonate peak.

Gordon et al. (12) developed a non-ion chromatographic method using chlorpromazine to determine bromate ion in ozone-treated waters. In the presence of bromate ion, chlorpromazine oxidizes forming a relatively stable colored product that can be monitored at 530 nm. Chemical masks are described to minimize potential interference from nitrite and chloride ions.

BROMATE ION FORMATION: PATHWAYS AND MECHANISMS

During the oxidation and/or chemical disinfection of natural waters containing bromide ion with ozone, bromate ion can be formed at concentrations ranging from 0 to 150 μg/L under normal water treatment conditions. During ozonation, bromide ion is first oxidized by dissolved ozone to hypobromite ion (OBr^-) which is then further oxidized to bromate ion. The reaction is pH-dependent because OBr^- is in equilibrium with hypobromous acid (HOBr). OBr^- builds up rapidly during ozonation and becomes the main reservoir for bromate. Hypobromite reacts further with ozone to form bromide ion (77%) and bromate ion (23%) in pure aqueous solutions.

The molecular ozone mechanism does not account for OH radicals formed as secondary oxidants from decomposed ozone during water treatment. Richardson et al. (13) and Yates and Stenstrom (14) indicate that there is a radical pathway which is influenced by both pH and alkalinity. The $OH\cdot$ radical and, to a lesser degree, the carbonate radical pathway ($CO_3\cdot^-$) may be more important than the molecular ozone pathway. Oxidants such as OH and CO_3^- radicals may interact with intermediate bromine species leading to the formation of BrO radicals which eventually undergo disproportionation to form hypobromite and bromite (BrO_2^-). Bromate ion is then formed through oxidation of bromite by ozone. The radical mechanism for the formation of bromate ion includes two decisive reaction steps still involving molecular ozone: the formation of hypobromite and oxidation of bromite (15). In contrast, Yates and Stenstrom (14) suggest that $BrO\cdot$ is formed primarily by the direct reaction of ozone with $Br\cdot$, which is formed by the oxidation of Br^- by $OH\cdot$. Siddiqui et al. (16) assumed that rate constants obtained by Haag and Hoigne (17) through molecular ozone pathways may intrinsically take into account bromate ion formation through both molecular ozone and radical pathways.

Bromate ion formed through reactions with molecular ozone contributes in the range of 30–80% to overall bromate ion formation in natural organic matter (NOM)-containing waters. Ozekin et al. (18) report up to 65% and 100% bromate formation through a radical pathway in NOM-free and NOM-contaning waters, respectively. Differences in NOM-containing waters can be attributed to differences in the characteristics of NOM. A change in mechanism as a function of pH and the competitive roles of the free radical (one-electron transfer) mechanism above pH 7 versus oxygen-atom (two-electron transfer) mechanisms help explain both the large variations in bromate ion yield and the sensitivity to reactor design, concentration of NOM, and ozone/bromide ion concentrations (19). In summary, bromate ion formation can occur through both radical and molecular ozone pathways depending on NOM concentration in source waters. This is an important finding that will have strong implications for bromate ion control strategies.

BROMATE FORMATION: EFFECT OF WATER QUALITY PARAMETERS AND TREATMENT CONDITIONS

Several studies have demonstrated that several water quality and ozonation parameters affect the formation of bromate ion. A better mechanistic understanding relating these parameters to bromate ion formation is critical to control its formation during ozonation.

Bromide Ion Concentration

An important difference from the chlorine system exists in that Br^-, unlike Cl^-, reacts with ozone at a significant rate and this reaction is independent of pH because Br^- is not protonated in water. At low pH (≤ 6), the bromine formed is relatively inert to ozonation ($k < 10^{-3}$ M^{-1} s^{-1}), and nearly all of the Br^- is present as HOBr. In the absence of NOM (or other substances that react rapidly with HOBr), all of the Br^- can eventually be recovered as bromate ion when ozonated. The ranges of bromate ion levels found by different investigators for various bromide ion-containing waters are summarized in Table 1.

Siddiqui and Amy (21) have shown that there is a threshold bromide ion concentration below which no detectable (<2 μg/L) bromate ion formation occurs. This threshold level is dependent on pH, ozone dose, DOC, and alkalinity of the water being ozonated and hence will vary from source to source. Siddiqui and Amy (21) have shown a bromide ion threshold concentration of 0.25 mg/L for an O_3/DOC ratio of 1.5:1 at pH 7.5 for groundwater with a DOC of approximately 4 mg/L. Krasner et al. (23) indicated that bromide ion levels of ≥0.18 mg/L yield measurable levels of bromate ion when a target ozone residual for disinfection is met.

Ozone Dose

The ozone dose plays a critical role in the formation of bromate ion. Increasing ozone dose increases the formation of bromate ion until all bromide ion is converted to bromate ion (21). If sufficient ozone is added to meet the ozone demand of a water, bromate ion will be produced if there is sufficient bromide ion (≥100 μg/L). However, according to Krasner et al. (23) and Siddiqui et al. (21), appropriate staging of ozone through two or three chambers can minimize ozone residual and bromate ion formation in full-scale ozone contactors. Siddiqui and Amy (24) and Kruitof et al. (7) have shown that there is an ozone threshold concentration below which no bromate ion formation occurs. For pretreated Meuse River water with a DOC content of 2.4 mg/L, the lowest ozone dose at which bromate ion could be detected was 1.0 mg/L over a pH range of 6.8–7.8 (7, Table 2.4). Optimization of ozone dose for preoxidation or disinfection should take into account the formation potential of bromate ion.

Alkalinity

An important factor controlling the half-life of ozone in natural waters is alkalinity. Addition of alkalinity increases the formation of bromate ion (16). Their research has shown that the contribution to bromate ion formation due to carbonate radicals at pH 8.5 was much more predominant than at pH 6.5 (30% vs. 5%). This is presumably due to a shift toward HCO_3^- ions from CO_3^{2-} ions and higher production of OH radicals. CO_3^{2-} ions react about 30 times faster with OH radicals than HCO_3^- ions and therefore have a larger scavenging effect.

In summary, during ozonation, part of the ozone added reacts directly with bromide ion, and a part of it decomposes to radicals which further react with carbonate alkalinity to produce carbonate radicals. The higher the pH, the faster the decomposition rate of ozone (initiated by OH^-), and the faster the formation of bromate ion. Additionally, it is accelerated by a chain reaction for which the bromine species radicals formed act as carriers. Because of this chain reaction, dissolved ozone decay depends on bromate ion yielding radical intermediates which additionally catalyze the decomposition of ozone and alkalinity which scavenge the OH radicals. Alkalinity quenches the chain reaction resulting in an increase in ozone half-life.

Natural Organic Matter

The reaction of ozone with NOM can occur by direct reaction or by radical processes. The ozone decomposition and OH radical yield from NOM is influenced by the type and concentration of NOM in natural waters. Direct consumption of ozone is greater when the UV absorbance

Table 1. Summary of Bromate Ion Formation Potential in Different Source Waters

No. Samples	Br^-, μg/L	O_3, mg/L	pH	Alkalinity, mg/L	DOC, mg/L	BrO_3^-, μg/L	Reference
18	10–800	1–9.3	5.6–9.4	20–132	2.2–8.2	<5–60	(20)
4	60–340	3–12	6.5–8.5	90–230	3–7	<5–40	(21)
28	10–100	2–4	6.8–8.8	20–120	0.3–11	<5–100	(22)
4	12–37	0–3.97	7.8	N/A	N/A	<7–35	(8)
1	500	2.3–9.5	7.2–8.3	N/A	N/A	13–293	(14)
23	12–207	0.3–4.3	5.7–8.2	14–246	0.5–6.8	<2–16	(6)
8	107–237	1–5	6.8–8.0	N/A	2–5	<5–50	(7)

(due to electrophilic and nucleophilic sites) of the source water is significant, resulting in decreased bromate ion. The radical process is initiated by active sites present in humic substances or formed during ozonation of nucleophilic sites and by hydroxide ions. Amy et al. (22) have shown that bromate ion formation, with all factors equal except NOM, varies from source to source. Amy et al. (22) studied more than seven different source waters containing a wide range of NOM characteristics. Krasner et al. (23) suggest that at fixed CT (Concentration of Disinfectant × Time of Contact) values, BrO_3^- increases with NOM.

In summary, the presence of NOM may accelerate ozone consumption and can reduce bromate ion formation potential. Besides pH, NOM concentration, measured as dissolved organic carbon (DOC), is influential; a higher Br^-/DOC ratio enhances bromate ion over organo-Br formation. The formation of bromate ion in NOM-containing waters is predominantly through a radical pathway rather than by the molecular ozone pathway.

Temperature

The effects of temperature are several: (1) dissolved ozone (DO_3) is more stable at lower temperatures, (2) increasing temperature increases the reaction rate constant, and (3) the pK_a of the HOBr/OBr^- system is temperature dependent. Siddiqui and Amy (21), Amy et al. (22), and Kruithof et al. (7) have observed an increase in bromate ion concentration as temperature increases. Kruitof et al. (7) studied bromate ion formation at two different temperatures (5 and 20 °C) in Meuse River water in bench-scale experiments and found a significant increase in bromate ion formation potential upon increasing the temperature from 5 to 20 °C. Siddiqui and Amy (21) indicated that increasing either ozonation temperature or incubation temperature had a positive effect on the formation of bromate ion. An increase in ozonation temperature produced more bromate ion than a corresponding increase in incubation temperature. This may be partially attributable to bromate ion formation occurring in less than 10 minutes.

In summary, increasing temperature generally increases the formation of bromate ion during drinking water treatment. Hence, more bromate ion formation may be expected in summer months than in winter months for the same applied ozone dose.

Reaction Time

Perhaps more than any other factor, reaction time is most important. Amy et al. (22) and Yamada (25) have shown that most bromate ion formation occurs within about 5 minutes, although its formation can continue over 30 minutes, suggesting that bromate ion forms mainly through the radical pathway. Control strategies to minimize bromate ion formation must take into account its kinetic time frame. Reaction rate increases upon the addition of H_2O_2 and at elevated pH levels, strengthening the argument that radicals are responsible for bromate ion formation. Bromate ion formation occurs only in the presence of DO_3, so maintaining an appropriate amount of DO_3 is critical to minimizing bromate ion formation. A small increase in DO_3 can result in a severalfold increase in bromate ion formation.

BROMATE ION MINIMIZATION STRATEGIES

Given preliminary health information and the potential of a low MCL for bromate ion, control options require careful optimization to achieve low concentrations in ozonated waters containing significant levels of bromide ion. A better understanding of the mechanisms of bromate ion minimization strategies is required to control its formation during ozonation. Bromate ion can be minimized by chemical factors (chemical addition) or physical/hydrodynamic factors (contactor design).

pH Depression

Bromate ion concentration increases as pH increases. Reducing the pH of water before ozonation impacts bromate formation minimization in a number of ways. At $pH < 7$, oxidized bromide ion is primarily in the form of HOBr, thus minimizing bromate ion formation, and at $pH>8.0$, the efficiency of OH radical generation from ozone decomposition increases. Siddiqui et al. (21) showed that decreasing the pH from 8 to 6 lowered the ozone dose required for disinfection by 33% for a source water containing 3.5 mg/L DOC. In addition, because ozone residuals are more stable at lower pH levels, a lower ozone dose is required to achieve the same CT value at a reduced pH, compared to ozonation at ambient pH.

Draft drinking water regulations in the United States will tentatively specify a best available treatment (BAT) of pH adjustment for bromate ion minimization. The major constraint to this BAT is the acid cost for high alkalinity waters, and the subsequent need to adjust pH after treatment for corrosion control. As an example, Krasner et al. (23) indicated that treating State Project Water (CA) at an average flow of 520 mgd could potentially reduce the ozone dose for minimum disinfection from 1.7 to 1.0 mg/L, resulting in an ozone cost savings of approximately $0.8 million per year. The estimated chemical costs, however, would be approximately $4 million and $2 million per year, respectively, for the acid to reduce the pH from 8 to 6 and the caustic to raise the pH back to 8.0 prior to distribution.

Ammonia Addition

Ammonia addition theoretically can tie up bromine (HOBr/OBr^-) as bromamine and can exert a free radical demand; however, the complexity of bromamine chemistry has yielded mixed BrO_3^- formation results in lab and pilot studies. Ammonia addition may result in a time-lag minimization of BrO_3^- formation in low DOC waters because of additional reactions with ammonia (reaction of ammonia with bromine proceeds faster than oxidation of bromamine and bromide ion by ozone). Glaze et al. (26) and Siddiqui and Amy (21) observed up to a 30% decrease in bromate ion formation upon the addition of ammonia at pH levels near 7.0. Siddiqui et al. (27) indicated that the effect of ammonia addition on bromate ion formation is more predominant in low

DOC groundwaters. They attribute this to a majority of bromate ion formation in NOM-containing waters proceeding through the radical pathway.

Hydrogen Peroxide

At low pH levels ($pH < 7$), production of OH radicals and subsequent consumption of H_2O_2 by OH radicals is slower, thereby decreasing the rate of decomposition of ozone and enhancing the rate of bromate ion formation. At high pH values ($pH > 9$), the degree of dissociation of H_2O_2 to HO_2^- becomes apparent, and the decomposition of H_2O_2 by the reaction of O_3 with HO_2^- proceeds fast. The amount of bromate ion formed increases as long as the synergistic effect of OH radicals and molecular ozone is maintained. If the H_2O_2 concentration is too high, the half-life for ozone is too small for the formation of bromine species and eventually bromate ion.

Siddiqui and Amy (21) and Krasner et al. (23) have observed an increase in bromate ion formation, whereas Daniel et al. (28) and Kruitof et al. (7) have observed a decrease in bromate ion formation upon the addition of H_2O_2.

The differences in H_2O_2 effects on bromate ion formation may be attributed to differences in ozonation pH, H_2O_2 concentration maintained during ozonation, and the DOC content of the source waters. Excess H_2O_2 decreases the DO_3 concentration (less CT value) resulting in reduced bromate ion formation. On the other hand, if an optimal amount of H_2O_2 with respect to O_3 is present, bromate ion formation increases. In some full-scale plants, the addition of ozone is automated based on ozone residual at the outlet of the ozone contactors. Maintaining a constant residual with and without the addition of peroxide implies an increase in applied ozone which leads to an increase in bromate ion concentration.

Contactor Hydrodynamic Factors

Gas–liquid contactor hydrodynamics plays a critical role in bromate ion formation; it influences mass transfer rates, peak ozone residuals, spatial-temporal concentration gradients, and liquid backmixing phenomena. A number of conditions may be imposed on the operation of a full-scale contactor to minimize bromate ion formation. These contactor operating conditions may include contactor operation at low residual ozone concentrations and minimal backmixing.

Krasner et al. (23) and Siddiqui et al. (24) demonstrated that staging ozone application within a full-scale contactor or splitting ozone application in a pilot-scale ozone contactor can optimize ozone residual throughout the contactor as well as result in less bromate ion formation. This strategy minimizes bromate ion formation but may require larger ozone doses to ensure the same disinfection.

The ozone-content of the carrier gas and ozonation contact time affect the formation of bromate ion in laboratory-scale reactors (24). In full-scale experiments, Gramith et al. (29) observed higher bromate ion formation in an eductor system (6% ozone) than in an air-fed system (1.5% ozone). These differences in bromate ion production by the two different configurations indicate the sensitivity of bromate ion formation to ozone contactors and differences in mass transfer of ozone.

BROMATE ION REMOVAL

Several different options to remove bromate after its formation, applicable to surface water treatment plants contemplating the use of ozone at various points of application, have been evaluated by Siddiqui et al. (11): ferrous iron reduction, granular activated carbon (GAC) surface reduction, ultraviolet irradiation (UV) and high energy electron beam irradiation. In all processes, chemical analysis of the treated water showed the formation of bromide ion, indicating that chemical reduction of bromate to bromide ion is the significant mechanism. Bromate removal by activated carbon has been shown to be carbon specific, and not all the carbons have shown the ability to reduce bromate to bromide. The use of ferrous iron appears more promising because it also acts as a coagulant for removing disinfectant by-product precursors.

SUMMARY

Bromate ion formation occurs through both a molecular ozone and a free radical pathway. Most authors have shown that pH, bromide ion concentration, temperature, and alkalinity favor bromate ion formation whereas DOC has been shown to decrease bromate ion formation. pH depression reduces the formation of bromate ion, but this technique may not be cost-effective for high alkalinity source waters and may favor organobromine compounds. Different types of NOM can exert different ozone demands and associated different bromate ion formation potentials. Addition of ammonia has produced mixed results. Clearly, optimization of pH to minimize bromate ion formation by ammonia addition requires further study because the stability of different bromamine species depends on the pH of the water. Bromate ion can be minimized by manipulating ozone contactor design and operation. Appropriate staging of ozone through two or more chambers and optimizing the hydrodynamics of the ozone contactor can minimize ozone residual and bromate formation.

BIBLIOGRAPHY

1. Wegman, R.C. and Greve, P.A. (1981). Methyl bromide ion and bromide ion in drainage water after leaching of glassware soils. *Water Air Soil Pollut.* **16**(3): 3–11.
2. Luong, T.V., Peters, C.J., and Perry, R. (1981). Influence of bromide ion upon THM formation during water chlorination. *Water Ind., 81, Int. Water Conf.*, Brighton, U.K.
3. Vogt (1980).
4. Amy, G., Siddiqui, M., Zhai, W., and Debroux, J. (1993). National survey of Br^- in drinking waters. *1993 AWWA Annu. Conf. Proc.*, San Antonio, TX.
5. Krasner, S.W., et al. (1989). The occurence of disinfection by-products in US drinking water. *JAWWA* **81**(8): 41.

6. Legube, B., Bourbigot, M., Bruchet, A., Deguin, A., Montiel, A., and Matia, L. (1998). Bromide ion/bromate ion survey on different European water utilities. *Proc. Int. IWSA Workshop*, Paris.
7. Kruithof (1993).
8. Hautman, D.P. and Bolyard, M. (1993). Using ion chromatography to analyze inorganic disinfection by-products. *JAWWA* **85**(10): 88–93.
9. Kuo, C., Krasner, S.W., Stalker, G.A., and Weinberg, H.S. (1990). Analysis of inorganic disinfection by-products in ozonation drinking water by ion chromatography. *AWWA WQTC Proc.*, San Diego, CA.
10. Jagt vander der, H., Noij, Th.H.M., and Ooms, P.C.A. (1993). Analysis and identification of bromate ion in water by ion chromatography and multiple detection at the low-ppb level. *IWSA Workshop*, Paris, pp. 25–32.
11. Siddiqui, M., Amy, G., Ozekin, K., Zhai, W., and Westerhoff, P. (1994). Alternative strategies for bromate ion removal from drinking water. *JAWWA* **86**(10): 81.
12. Gordon, G., Bubnis, B., and Kuo, C. (1994). A flow injection, non-ion chromatographic method for measuring low level bromate ion in ozone treated waters. *Ozone: Sci. Eng.* February.
13. Richardson, L.B., Burton, D.T., Helz, G.R., and Rhoderick, J.C. (1981). Residual oxidant decay and bromate formation in chlorinated and ozonated sea-water. *Water Res.* **15**: 1067–1074.
14. Yates, R.S. and Stenstrom, M.K. (1993). Bromate production in ozone contactors. *AWWA Annu. Proc.*, San Antonio, TX.
15. von Gunten, U., Hoigne, J., and Bruchet, A. (1993). Bromate formation during ozonation of bromide ion containing waters. *IWSA Workshop*, Paris, pp. 51–56.
16. Siddiqui, M.S., Amy, G.L., and Rice, R.G. (1995). Bromate ion formation: A critical review. *JAWWA* **87**(10): 58.
17. Haag, W.R. and Hoigne, J. (1983). Ozonation of bromide-contaning waters: Kinetics of formation of hypobromous acid and bromate ion. *Environ. Sci. Tech.* **17**: 261.
18. Ozekin, K., Siddiqui, M., Amy, G., and Haag, W. (1994). Pathways for bromate formation in NOM-Free water. *AWWA Annu. Conf.*, New York.
19. Gordon, G. (1993). The chemical aspects of bromate control in ozonated drinking water containing bromide ion. *IWSA Workshop*, Paris, pp. 41–49.
20. Krasner, S.W., Glaze, W.H., Weinberg, H.S., Daniel, P.A., and Najm, I. (1993). Formation and control of bromate during ozonation of waters containing bromide ion. *JAWWA* **85**(1), 73–81.
21. Siddiqui, M. and Amy, G. (1993). DBPs formed during ozone-bromide reactions in drinking water treatment. *JAWWA* **85**: 63–72.
22. Amy, G., Siddiqui, M., Ozekin, K., and Westerhoff, P. (**1993**). Threshold levels for bromate ion formation in drinking water. *Proc. IWSA Workshop*, Paris, pp. 169–180.
23. Krasner, S.W., Gramith, J.T., Coffey, B.M., and Yates, R.S. (**1993**). Impact of water quality and operational parameters on the formation and control of bromate during ozonation. *IWSA Workshop*, Paris, pp. 157–168.
24. Siddiqui, M., Amy, G., Ozekin, K., Miller, K., and Westerhoff, P. (1993). The role of relating pilot-scale tracer data with bench-scale data for bromate prediction. *11th Ozone World Congr.*, San Francisco, CA.
25. Yamada, H. (1993). By-products of ozonation of low bromide waters and reduction of the by-products by activated carbon. *11th IOA Proc.*, San Francisco, CA.
26. Glaze, W.H., Weinberg, H.S., and Cavanagh, J.E. (1993). Evaluating the formation of brominated DBPs during ozonation. *JAWWA* **85**(1): 96–103.
27. Siddiqui, M. and Amy, G. (1994). Removal of bromate using UV irradiation. *IOA Reg. Conf.*, Zurich, Switzerland.
28. Daniel, P.A., Zafer, M.A., and Meyerhofer, P.F. (1993). Bromate control: Water quality, engineering and operational considerations. *IWSA Workshop*, Paris, pp. 181–188.
29. Gramith, J.T., Coffey, B.M., Krasner, S.W., Kuo, C., Means, E.G. (1993). Demonstration-scale evaluation of bromate formation and control strategies. *AWWA Annu. Conf.*, San Antonio, TX.

MUNICIPAL WATER SUPPLY: OZONATION

JAEHONG KIM
Georgia Institute of Technology
Atlanta, Georgia

OZONE IN WATER TREATMENT

Since the first full-scale application of ozone for water treatment in late 1800s, ozone has been widely used in municipal water supplies in many countries. The primary goal of the ozonation process in water treatment is often disinfection, but ozonation is also used in many other applications, including taste and odor control, color removal, oxidation of manganese and iron, flocculation enhancement, oxidation of natural organic matter (NOM) and accompanying promotion of its biodegradability, and control of chlorination byproducts (1). Recently, the application of ozone as a disinfectant has been receiving more attention because of its effectiveness in treating pathogens, such as oocysts of the protozoan parasite *Cryptosporidium parvum*, which have shown strong resistance to chlorine-based disinfectants (2,3).

In water treatment, ozone can be applied either at the early stage of treatment (i.e., preozonation) or after other treatment processes, including conventional processes such as coagulation/flocculation and sedimentation. Ozone is sometimes used with hydrogen peroxide or UV to promote the formation of hydroxyl radicals in advanced oxidation processes (AOPs). Due to the increased biodegradability of organic matter after ozonation, ozonation is typically followed by biological filtration to remove biodegradable organic matter and to minimize microbial regrowth in the distribution system. Ozone decomposes relatively fast, so secondary disinfectants such as free chlorine and monochloramine are often applied after ozonation to provide residual disinfectant in the distribution system.

OZONE CHEMISTRY

Ozone is a strong oxidant with a standard redox potential of 2.07 V in aqueous solution (4). Ozone has four resonance structures thus allowing it to act as both an electrophile and as a nucleophile (5). Because ozone is a highly selective oxidant, its use in water treatment allows it to inactivate pathogenic microorganisms selectively, oxidize manganese and iron, and react with organic matter that contains olefins, amines, sulfides, and phenolic functional groups. Reactions with ozone are generally categorized into electron-transfer, oxygen atom transfer, or ozone addition processes (6). Inorganic compounds such as ferrous ion (Fe^{2+}) tend to react with ozone primarily through oxygen atom transfer, but ozone addition to a saturated bond followed by rearrangement and formation of carboxylic acid is typical of ozone reactions with organic matter. The rate constants of ozone reactions with various inorganic and organic compounds in the aqueous phase can be found in the literature (6–8). The reactions of ozone with NOM are complicated and not well understood primarily due to the heterogeneity of NOM. Ozone reactions with NOM are known to reduce aromatic carbon content, selectively reduce phenolic carbon content, and can lead to the formation of smaller molecular weight carboxylic acids and aldehydes (9–11).

Ozone is not stable in water and has a half-life ranging from seconds to hours depending on the water quality. In addition to the various reactions between ozone and organic and inorganic constituents in water, aqueous phase ozone undergoes a chain-type decomposition that results in the formation of secondary oxidants such as hydroxyl radicals ($\cdot OH$), ozonide radicals ($HO_3\cdot/O_3{}^-\cdot$), and superoxide radicals ($HO_2\cdot/O_2{}^-\cdot$). The mechanism of ozone decomposition has been studied for decades. Figure 1 shows one plausible mechanism, constructed from the literature, of ozone decomposition in the absence of organic matter (12–18). The rate of ozone decomposition in the aqueous phase depends strongly on pH because ozone decomposition is initiated by hydroxide ion. In the presence of alkalinity, carbonate radicals are formed (19–22), as shown in Fig. 1. In natural waters, ozone and secondary oxidants react with naturally occurring organic and inorganic matter. Such reactions may result in consumption of ozone and secondary oxidants as well as formation of other secondary oxidants such as hydrogen peroxide, ozonide, and superoxide radicals (23,24), therefore adding considerable complexity to the decomposition mechanism shown in Fig. 1. Overall, ozone decay in natural water is a combined effect of ozone decomposition and reaction with aqueous constituents. The kinetics of ozone decay in natural waters is often characterized as an initial fast decrease in ozone concentration followed by a slower phase where ozone follows a first-order decrease in concentration (25).

The contribution of hydroxyl radical, the major secondary oxidant, to the overall oxidation potential during ozone application has often been expressed as R_{ct}, the ratio between hydroxyl radicals and ozone (26). In the case of AOPs, the intent is to form hydroxyl radicals typically to react with inorganic and organic compounds that are resistant to ozone.

BY-PRODUCT FORMATION

A variety of organic byproducts such as aldehydes, ketones, keto aldehydes, carboxylic acids, keto acids, hydroxyl acids, alcohols, and esters are formed during ozonation of natural waters (27). In the presence of bromide, various brominated organic by-products are formed. Bromate ($BrO_3{}^-$) is one of the by-products formed during ozonation of bromide-containing waters. Bromate formation follows a complex set of reactions involving several bromine species such as $HOBr/OBr^-$, $BrO\cdot$, and $BrO_2{}^-$ as intermediates (28–30). Formation of bromate during ozonation has been one of the major concerns in using ozone in drinking water treatment. Bromate is currently regulated at 10 μg/L in the United States and the European Union (31,32) due to its potential carcinogenicity (33).

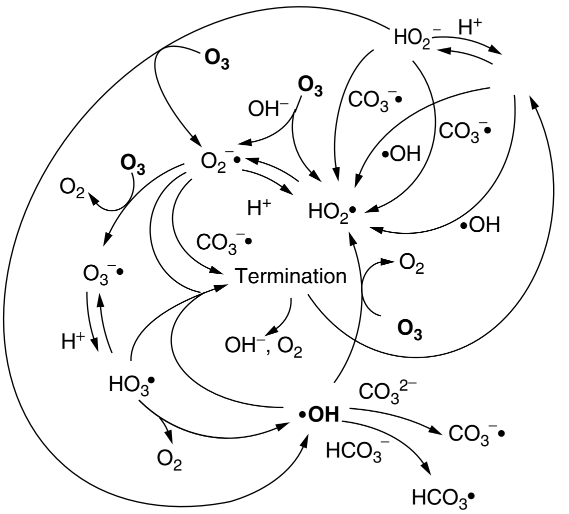

Figure 1. Mechanism of ozone decomposition in aqueous solution.

OZONE CONTACTOR

Ozone is generally produced on site from air or pure oxygen by electrical discharge using ozone generators. A typical ozone concentration in the gas phase is 15 g/m^3 when air is used, but concentrations as high as 150 g/m^3 can be achieved when pure oxygen is used. The gaseous phase ozone is then transferred into water in ozone contactors that typically provide 10–20 min of contact time. Ozone is relatively insoluble in water (e.g., Henry's law constant at 20 °C is 100 atm/M), so efficient transfer of ozone to water is an important factor in designing ozone contactors. There are several different types of ozone contactors used for disinfection, depending on the methods of ozone introduction and modes of contacting with water. The most typical configuration is a multichamber bubble-diffuser contactor, though other types of contactor configurations such as rapid mixers, submerged turbines, pipeline injectors, packed towers, and deep U-tubes are also used (1). In the multichamber bubble-diffuser contactor, one or more of the upstream chambers in these units are transfer chambers in which ozone gas is introduced through a set of ceramic or stainless steel diffusers installed at the bottom of the unit. The remaining downstream chambers are used as reactive chambers for additional contact with residual dissolved ozone. A transfer chamber can be designed for operation in either countercurrent flow (water and gas flowing in opposite

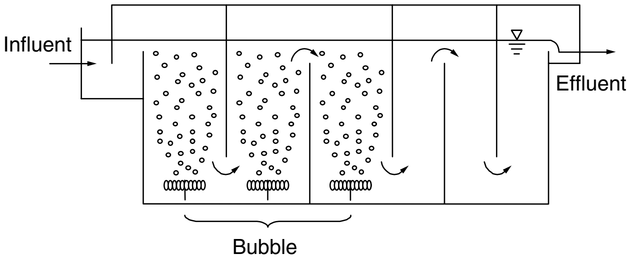

Figure 2. Schematic of full-scale ozone bubble-diffuser contactor located in Oakland, CA.

directions) or cocurrent flow (both water and gas flowing in the same direction) configuration. The transfer portion of a bubble diffuser contactor can consist of countercurrent and cocurrent flow chambers of similar size connected in series (34) or countercurrent flow chambers connected by small reactive upflow chambers (35). The residual ozone in the effluent of the contactor can be destroyed by hydrogen peroxide or activated carbon. A schematic of a typical ozone contactor is shown in Fig. 2. Off-gas from the contactor may contain residual ozone, which is subsequently treated by heat and/or catalyst for further destruction.

The design of an ozone bubble-diffuser contactor requires decisions such as selecting the number and distribution of transfer and reactive chambers, water column height, chamber cross-sectional area, number and type of diffusers, gas flow rate, and ozone gas concentration. The choice of contactor configuration and operating conditions can affect the hydrodynamics and mass transfer, which in turn can impact the ozone concentration throughout the ozone contactor. In addition, the performance of the contactor is also affected by certain water quality parameters, primarily those affecting the kinetics of ozone decomposition, such as temperature, pH, natural organic matter content, and alkalinity. Computer simulation has often been employed to model the complex phenomena involved in process design and optimization (25,36).

BIBLIOGRAPHY

1. Langlais, B., Reckhow, D.A., and Brink, D.R. (1991). *Ozone in Water Treatment Application and Engineering*. American Water Works Association Research Foundation, Denver, CO, and Lewis, Chelsea, MI.
2. Rennecker, J.L., Mariñas, B.J., Rice, E.W., and Owens, J.H. (1999). Inactivation of *Cryptosporidium parvum* oocyst with ozone. *Water Res.* **33**: 2481–2488.
3. Rennecker, J.L., Driedger, A.M., Rubin, S.A., and Mariñas, B.J. (2000). Synergy in sequential inactivation of *Cryptosporidium parvum* with ozone/free chlorine and ozone/monochloramine. *Water Res.* **34**: 4121–4130.
4. Bard, A.J., Parsons, R., and Jordan, J. (1985). *Standard Potentials in Aqueous Solution*. Marcel Dekker, New York.
5. Fessenden, R.J. and Fessenden, J.S. (1994). *Organic Chemistry*, 6th Edn. Brooks/Cole, New York.
6. Hoigné, J. (1998). Chemistry of aqueous ozone and transformation of pollutants by ozonation and advanced oxidation processes. *The Handbook of Environmental Chemistry*, Vol. 5, Part C. J. Hrubec (Ed.). Springer-Verlag, Berlin.
7. Neta, P., Huie, R.E., and Ross, A.B. (1988). Rate constants for reactions of inorganic radicals in aqueous solution. *J. Phys. Chem. Ref. Data* **17**: 1027–1284.
8. von Gunten, U. (2003). Ozonation of drinking water: Part I. Oxidation kinetics and product formation. *Water Res.* **37**: 1443–1467.
9. Killops, S.D. (1986). Volatile ozonation products of aqueous humic material. *Water Res.* **20**: 153–165.
10. Schechter, D.S. and Singer, P.C. (1995). Formation of aldehydes during ozonation. *Ozone Sci. Eng.* **17**: 52–69.
11. Westerhoff, P., Debroux, J., Aiken, G., and Amy, G.L. (1999). Ozone-induced changes in natural organic matter (NOM) structures. *Ozone Sci. Eng.* **21**: 551–570.
12. Staehelin, J. and Hoigné, J. (1982). Decomposition of ozone in water: Rate of initiation by hydroxide ions and hydrogen peroxide. *Environ. Sci. Technol.* **16**: 676–681.
13. Bühler, R.E., Staehelin, J., and Hoigné, J. (1984). Ozone decomposition in water studied by pulse radiolysis 1.HO_2/O_2^- and HO_3/O_3^- as intermediates. *J. Phys. Chem.* **88**: 2560–2564.
14. Staehelin, J., Bühler, R.E., and Hoigné, J. (1984). Ozone decomposition in water studied by pulse radiolysis. 2.OH and HO_4 as chain intermediates. *J. Phys. Chem.* **88**: 5999–6004.
15. Forni, L., Bahnemann, D., and Hart, E.J. (1982). Mechanism of the hydroxide ion initiated decomposition of ozone in aqueous solution. *J. Phys. Chem.* **86**: 255–259.
16. Sehested, K., Holcman, J., and Hart, E.J. (1983). Rate constants and products of the reactions of e_{aq}^-, O_2^-, and H with ozone in aqueous solution. *J. Phys. Chem.* **87**: 1951–1954.
17. Sehested, K., Holcman, J., Bjergbakke, E., and Hart, E.J. (1984). A pulse radiolytic study of the reaction $OH + O_3$ in aqueous medium. *J. Phys. Chem.* **88**(18): 4144–4147, additions and corrections (1985) **89**: 388.
18. Christensen, H., Sehested, K., and Corfitzen, H. (1982). Reactions of hydroxyl radicals with hydrogen peroxide at ambient and elevated temperatures. *J. Phys. Chem.* **86**: 1588–1590.
19. Buxton, G.V. and Elliot, A.J. (1986). Rate constant for reactions of hydroxyl radicals with bicarbonate ions. *Radiat. Phys. Chem.* **27**: 241–243.
20. Buxton, G.V., Greenstock, C.L., Helman, W.P., and Ross, A.B. (1988). Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms, and hydroxyl radicals in aqueous solution. *J. Phys. Chem. Ref. Data* **17**: 513–886.
21. Behar, D., Czapski, G., and Duchovny, I. (1970). Carbonate radical in flash photolysis and pulse radiolysis of aqueous carbonate solutions. *J. Phys. Chem.* **74**: 2206–2210.
22. Eriksen, T.E., Lind, J., and Merenyi, G. (1985). On the acid-base equilibrium of the carbonate radical. *Radiat. Phys. Chem.* **26**: 197–199.
23. Echigo, S. (2002). *Kinetics and Speciation of Brominated Disinfection By-Products During Ozonation*, Ph.D. Dissertation, University of Illinois at Urbana-Champaign, Urbana, IL.
24. Staehelin, J. and Hoigné, J. (1985). Decomposition of ozone in water in the presence of organic solutes acting as promoters and inhibitors of radical chain reactions. *Environ. Sci. Eng.* **19**: 1206–1213.
25. Kim, J.H., Tomiak, R.B., and Mariñas, B.J. (2002). Inactivation of *Cryptosporidium* in a pilot-scale ozone bubble-diffuser contactor. Part I: Model development. *ASCE J. Environ. Eng.* **128**: 514–521.

26. Elovitz, M.S. and von Gunten, U. (1999). Hydroxyl radical/ozone ratios during ozonation processes. I. The R_{ct} concept. *Ozone Sci. Eng.* **21**: 239–260.
27. von Gunten, U. (2003). Ozonation of drinking water: Part II. Disinfection and by-product formation in presence of bromide, iodide, or chlorine. *Water Res.* **37**: 1469–1487.
28. von Gunten, U. and Hoigné, J. (1994). Bromate formation during ozonation of bromide-containing waters: Interaction of ozone and hydroxyl radical reactions. *Environ. Sci. Technol.* **28**: 1234–1242.
29. von Gunten, U. and Oliveras, Y. (1998). Advanced oxidation of bromide-containing waters: Bromate formation mechanisms. *Environ. Sci. Technol.* **32**: 63–70.
30. Kim, J.H., von Gunten, U., and Mariñas, B.J. (2004). Modeling bromate formation and *Cryptosporidium parvum* inactivation in synthetic waters. *Environ. Sci. Technol.* **38**: 2232–2241.
31. *Federal Register* (1998). **63**: 241: 69389.
32. European Union. (1998). *Amtsblatt der Europäischen Gemeinschaften. Richtlinie 98/83/EG des Rates*. 3-12-1998, pp. 32–54.
33. Wilbourn, J. (1995). Toxicity of bromate and some other brominated compounds in drinking water. *Water Supply* **13**: 1–8.
34. Coffey, B.M., Graff, K.G., Mofidi, A.A., and Gramith, J.T. (1995). On-line monitoring of ozone disinfection effectiveness within an over/under baffled contactor. *Proc. 1995 Am. Water Works Assoc. Natl. Conf.*, Anaheim, CA, June 18–22, 1995, pp. 77–132.
35. Mariñas, B.J., Liang, S., and Aieta, M.E. (1993). Modeling hydrodynamics and ozone residual distribution in a pilot-scale ozone bubble-diffuser contactor. *J. Am. Water Works Assoc.* **85**: 90–99.
36. Kim, J.H. et al. (2002). Inactivation of *Cryptosporidium* in a pilot-scale ozone bubble-diffuser contactor. Part II: Model verification and application. *ASCE J. Environ. Eng.* **128**: 514–521.

REVIEW OF PARASITE FATE AND TRANSPORT IN KARSTIC AQUIFERS

William J. Blanford
Louisiana State University
Baton Rouge, Louisiana

Every year in the United States cases of waterborne illness are generally characterized by gastrointestinal problems (vomiting, diarrhea, etc.). However, these same symptoms can be more serious and even fatal in certain groups, such as the young, elderly, and those with compromised immune systems. Furthermore, *viral pathogens* found in groundwater are linked to chronic diseases such as adult onset diabetes and myocarditis. Recognizing that preventing contamination of water sources is the first step in the multiple barrier approach to drinking water protection, the Interim Enhanced Surface Water Treatment Rule (IESWTR) and the Disinfectants/Disinfection Byproducts Rule (DBPR) were released in 1999. The two rules will provide additional protection of public water supplies and address microbiological contamination. The rules address the risks of the trade-offs between disinfection necessary for *microbial inactivation* and disinfection byproducts. The requirements of 40 CFR 141.72(b) are consistent 99.99% removal and/or inactivation of viruses and 99% removal of *Cryptosporidium oocysts*. The code gives public water suppliers the option to demonstrate compliance through pilot plant studies or other means that combine filtration techniques with *disinfection* as in EPA 1999 Microbial and Disinfection Byproducts Rules Simultaneous Compliance as in Guidance Manual.

Presently, only surface water systems and systems using groundwater under the direct influence of surface water are required to disinfect their water supplies. Although historically considered free of microbial contamination, most recent research has shown that aquifers can be readily contaminated by pathogens. To lessen the requirement for disinfection and to protect water resources, water suppliers have sought to use natural systems to treat *Cryptosporidium*. Thus, there is a critical need for information regarding the transport and fate of pathogens during water infiltration into subsurface systems.

Highly porous limestone aquifers are the primary water supply for about 25% of the U.S. population, including several large municipalities such as Atlanta, Austin, Miami, Nashville, San Antonio, and Tampa. After discussion with the water utilities and regulators in some of these cities, we found that there was tremendous concern about microbial contamination from parasites such as *Cryptosporidium*. Their concern was from the lack of information needed to determine the potential for *Cryptosporidium* and other pathogens to travel through their limestone aquifers from septic fields, wastewater treatment plants, and agricultural feedlots to the wells and springs from which they derive their water supplies. Through our discussion, we determined that a contaminant transport model compatible with existing models for limestone systems is needed. To produce the model, extensive laboratory tests are needed to garner transport parameters, and field tests are needed to evaluate the performance of the model. A recent report by the National Research Council on the role of science and technology in countering terrorism reached similar conclusions (1). Section 8.15 of that report states,

> research should be undertaken on water sampling schemes to determine what types and population of data points are required for a spatiotemporal network and on intelligent decision processing to be able to reliably recognize the patterns of attack indicators vs. natural hazards. Such research would require that priority attention be given to the development of simulation models that would both analyze and simulate events and serve to train operators in systematic recovery, emergency response, and evacuation.

KARSTIC LIMESTONE AQUIFERS

Karstic landscapes cover approximately 15% of the earth's land surface and are characterized by caves, sinkholes, and springs. Karstic aquifer systems occur throughout the United States and are particularly significant for groundwater resources in Texas and Florida, the second

Previous Page

26. Elovitz, M.S. and von Gunten, U. (1999). Hydroxyl radical/ozone ratios during ozonation processes. I. The R_{ct} concept. *Ozone Sci. Eng.* **21**: 239–260.
27. von Gunten, U. (2003). Ozonation of drinking water: Part II. Disinfection and by-product formation in presence of bromide, iodide, or chlorine. *Water Res.* **37**: 1469–1487.
28. von Gunten, U. and Hoigné, J. (1994). Bromate formation during ozonation of bromide-containing waters: Interaction of ozone and hydroxyl radical reactions. *Environ. Sci. Technol.* **28**: 1234–1242.
29. von Gunten, U. and Oliveras, Y. (1998). Advanced oxidation of bromide-containing waters: Bromate formation mechanisms. *Environ. Sci. Technol.* **32**: 63–70.
30. Kim, J.H., von Gunten, U., and Mariñas, B.J. (2004). Modeling bromate formation and *Cryptosporidium parvum* inactivation in synthetic waters. *Environ. Sci. Technol.* **38**: 2232–2241.
31. *Federal Register* (1998). **63**: 241: 69389.
32. European Union. (1998). *Amtsblatt der Europäischen Gemeinschaften. Richtlinie 98/83/EG des Rates*. 3-12-1998, pp. 32–54.
33. Wilbourn, J. (1995). Toxicity of bromate and some other brominated compounds in drinking water. *Water Supply* **13**: 1–8.
34. Coffey, B.M., Graff, K.G., Mofidi, A.A., and Gramith, J.T. (1995). On-line monitoring of ozone disinfection effectiveness within an over/under baffled contactor. *Proc. 1995 Am. Water Works Assoc. Natl. Conf.*, Anaheim, CA, June 18–22, 1995, pp. 77–132.
35. Mariñas, B.J., Liang, S., and Aieta, M.E. (1993). Modeling hydrodynamics and ozone residual distribution in a pilot-scale ozone bubble-diffuser contactor. *J. Am. Water Works Assoc.* **85**: 90–99.
36. Kim, J.H. et al. (2002). Inactivation of *Cryptosporidium* in a pilot-scale ozone bubble-diffuser contactor. Part II: Model verification and application. *ASCE J. Environ. Eng.* **128**: 514–521.

REVIEW OF PARASITE FATE AND TRANSPORT IN KARSTIC AQUIFERS

William J. Blanford
Louisiana State University
Baton Rouge, Louisiana

Every year in the United States cases of waterborne illness are generally characterized by gastrointestinal problems (vomiting, diarrhea, etc.). However, these same symptoms can be more serious and even fatal in certain groups, such as the young, elderly, and those with compromised immune systems. Furthermore, *viral pathogens* found in groundwater are linked to chronic diseases such as adult onset diabetes and myocarditis. Recognizing that preventing contamination of water sources is the first step in the multiple barrier approach to drinking water protection, the Interim Enhanced Surface Water Treatment Rule (IESWTR) and the Disinfectants/Disinfection Byproducts Rule (DBPR) were released in 1999. The two rules will provide additional protection of public water supplies and address microbiological contamination. The rules address the risks of the trade-offs between disinfection necessary for *microbial inactivation* and disinfection byproducts. The requirements of 40 CFR 141.72(b) are consistent 99.99% removal and/or inactivation of viruses and 99% removal of *Cryptosporidium oocysts*. The code gives public water suppliers the option to demonstrate compliance through pilot plant studies or other means that combine filtration techniques with *disinfection* as in EPA 1999 Microbial and Disinfection Byproducts Rules Simultaneous Compliance as in Guidance Manual.

Presently, only surface water systems and systems using groundwater under the direct influence of surface water are required to disinfect their water supplies. Although historically considered free of microbial contamination, most recent research has shown that aquifers can be readily contaminated by pathogens. To lessen the requirement for disinfection and to protect water resources, water suppliers have sought to use natural systems to treat *Cryptosporidium*. Thus, there is a critical need for information regarding the transport and fate of pathogens during water infiltration into subsurface systems.

Highly porous limestone aquifers are the primary water supply for about 25% of the U.S. population, including several large municipalities such as Atlanta, Austin, Miami, Nashville, San Antonio, and Tampa. After discussion with the water utilities and regulators in some of these cities, we found that there was tremendous concern about microbial contamination from parasites such as *Cryptosporidium*. Their concern was from the lack of information needed to determine the potential for *Cryptosporidium* and other pathogens to travel through their limestone aquifers from septic fields, wastewater treatment plants, and agricultural feedlots to the wells and springs from which they derive their water supplies. Through our discussion, we determined that a contaminant transport model compatible with existing models for limestone systems is needed. To produce the model, extensive laboratory tests are needed to garner transport parameters, and field tests are needed to evaluate the performance of the model. A recent report by the National Research Council on the role of science and technology in countering terrorism reached similar conclusions (1). Section 8.15 of that report states,

> research should be undertaken on water sampling schemes to determine what types and population of data points are required for a spatiotemporal network and on intelligent decision processing to be able to reliably recognize the patterns of attack indicators vs. natural hazards. Such research would require that priority attention be given to the development of simulation models that would both analyze and simulate events and serve to train operators in systematic recovery, emergency response, and evacuation.

KARSTIC LIMESTONE AQUIFERS

Karstic landscapes cover approximately 15% of the earth's land surface and are characterized by caves, sinkholes, and springs. Karstic aquifer systems occur throughout the United States and are particularly significant for groundwater resources in Texas and Florida, the second and fourth most populous states. There is significant interchange between surface and subsurface features, fluids, and waterborne materials. In many unconfined karstic aquifers, groundwater and surface water behave as a single body of water (2). The Edwards and Trinity aquifers are the only significant water sources for semiarid south central Texas, including San Antonio, the eighth largest city in the United States. The Edwards aquifer is one of the most permeable and productive carbonate aquifers in the world. It provides a public water supply to more than 2 million people, water for agriculture and industry, and discharges into major springs. These springs support recreation and business, provide flow to downstream users, and are a habitat for several threatened and endangered species. Its high porosity and permeability result in part from the development of secondary porosity and fracturing. The Trinity aquifer, by contrast, is much less permeable or productive than the Edwards. Its lithologies contain significantly more clay minerals and other siliciclastics, as well as minor evaporites. Nevertheless, it supplies water to scattered communities, ranches, and individual wells north of the Edwards in an area of rapid urbanization and growth.

Karstic limestone aquifers are characterized by three primary types of porosity: intergranular matrix porosity, fracture porosity, and the development of cavern conduits (3). A conduit is defined as any interconnected pathway for water of sufficient size to permit turbulent flow. Resistance to flow in conduits is much less than in the adjacent matrix, and, as a consequence, most regional flow is concentrated in conduits. Similarly, most local matrix flow is toward the nearest conduit. Thus, hydrodynamic gradients can vary significantly between conduit recharge and regional flow.

Contaminant dispersion would take place by a variety of methods using this model. Contaminants within conduits travel much more rapidly than those within the matrix, which follow Darcy's law and the advection–dispersion equation. However, exchange of contaminants between the two can occur under a variety of naturally occurring cycles, such as floods versus droughts. Thus the system should be modeled as a network of conduits embedded in a porous matrix (4). The complex depositional environments that produce carbonate rocks, their pronounced digenetic susceptibilities, and the fabric controls related to tectonic alteration of large carbonate units result in groundwater systems that are highly heterogeneous and poorly understood. Regional bedrock karstic limestone aquifers cannot be understood without critical information on the geologic boundary conditions.

CRYPTOSPORIDIUM

Cryptosporidiosis has been recognized as a human disease since 1976 (5). The first diagnosed waterborne outbreak of cryptosporidiosis in the world occurred in Braun Station, Texas, in 1984 (6). The first report of the disease associated with a contaminated municipal water supply was in 1987 in Carrollton, Georgia, where 13,000 became ill (7). This water system met all state and federal drinking water standards. In Milwaukee, Wisconsin, in 1993, municipal drinking water infected 400,000 people with *Cryptosporidium* and resulted in approximately 50 deaths (8,9). Subsequently, the Texas Department of Health reported a more recent water outbreak in Brushy Creek, Texas, in 1998. Overall, there have been 12 documented waterborne outbreaks of cryptosporidiosis in North America between 1985 and 1997; in two of these (Milwaukee and Las Vegas), mortality rates in the immunocompromised ranged from 52% to 68% (10).

Cryptosporidium parvum is a protozoan tissue parasite and is an agent of enterocolitis in mammals (11). *Cryptosporidium* has a complicated and extensive life cycle. The environmental stage is an oocyst, which is a metabolically dormant protective phase (12). The oocysts are nearly spherical and have a diameter of 4.5 to 5.5 μm. Their surfaces are slightly negative to neutrally charged in natural waters (13,14), and their density is close to that of water at 1.025 to 1.070 (14). The oocyst encases four sporozoites, each capable of infecting a host cell. Ingestion of as few as 10 oocysts can lead to infection; the feces of infected mammals may contain as many as 10^7 oocysts/mL (15).

Cryptosporidium can enter the environment via human and animal wastes. It has been found in marine water and bathing beaches in the vicinity of a nearby sewage outfall (16). Cryptosporidiosis has been reported in many domestic animals, especially cattle. An infected calf can excrete 10^{10} oocyts per day. In a study of farm drains, Kemp et al. (17) found 0.06 to 19.4 oocysts per liter, which can result in contamination of surface waters. Typical concentrations of *Cryptosporidium* in untreated domestic wastewaters are between 1 and 10 oocysts/mL (18) and in polluted streams, are between 0.1 to 100 oocysts/mL (19). *C. parvum* forms a hardy oocyst that can survive chlorine disinfection as commonly practiced in conventional water treatment (20,21). Furthermore, Chauret et al. (22) concluded oocysts exposed to environmental conditions are as resistant to inactivation by chlorination as freshly shed oocysts. *C. parvum* oocysts have also survived for weeks in surface waters (23).

Mawdsley et al. (24) quantitatively monitored the movement of *Cryptosporidium parvum* oocysts from livestock waste through low-permeability silt clay loam soil in laboratory column and box studies. In their box studies, livestock waste was applied to a portion of the surface of 20 cm deep by 80 cm long blocks of undisturbed low-permeability silt clay loam soil. The soil was contained within a tilted box and water was applied for 70 days. Oocysts were found in leachate in numbers ranging from 10^4 to 10^6 oocysts. Postsoil core analysis found decreasing numbers of oocyst with distance from inoculation.

Brush et al. (25) and Harter et al. (11) performed laboratory column studies of *Cryptosporidium parvum* oocyst transport and fate. Brush et al. examined oocyst transport through columns containing glass beads, well-sorted sand, and shale aggregates. In these short-duration studies, approximately 50% of the oocysts were retained in the sand and roughly 40% in the glass beads and shale. Despite the losses within the column, oocysts were eluted during the same timeframe as the conservative tracer chloride in the sand and glass beads. The velocity of *C.*

parvum oocyst through the shale aggregate was slightly faster than chloride ions. Brush et al. (13) suggest that this was due to size exclusion and not charge exclusion because of the nearly neutral surfaces of the oocysts. Brush et al. modeled the movement of oocyst with a one-dimensional convective–dispersion transport equation. Sorption was described as instantaneous equilibrium sorption where the relationship between sorbed and aqueous concentrations was linear. All loss processes, including decay, sieving, impingement, and settling, were modeled as a single (i.e., lumped) first-order rate process. The results of their modeling efforts indicated that oocysts experienced less shear and turbulence than dissolved chloride ions and that the oocysts did not adhere to the porous media.

Harter et al. (11) investigated the influence of pore-water velocity and sand grain size on the transport and fate of *Cryptosporidium parvum* oocysts. They used 10 cm long columns, groundwater of medium ionic strength (100 to 150 mg/L TDS), and bovine *Cryptosporidium parvum* oocysts. Their results indicated no trend between sand size and oocyst recovery in the three studies. They found that oocysts arrived before chloride and the elution of oocysts continued after chloride for all column studies. In a limited number of extended studies, tailing of oocyst concentration was observed until the end of the experiment (250+ pore volumes). The early breakthrough, as with Brush et al., was attributed to size exclusion. The tailing in oocyst elution was attributed to reversible deposition.

Harter et al. (11) used a more complex one-dimensional transport model to simulate their data. Filtration was modeled as a first-order irreversible process calculated from the physical properties of the soil (grain size, porosity, and bulk density), water (density, viscosity, pore velocity), and microbial colloid (density, size, diffusion coefficient). Sorption was modeled as a first-order, rate-limited reversible process. Half of the loss was attributed to irreversible filtration, but rate-limited reversible sorption could account for the early breakthrough and partially account for the extended tailing. Harter et al. called for further experimental and theoretical research to measure and explain the long-term elution behavior of *Cryptosporidium parvum* oocysts.

GIARDIA

Giardia lamblia is the cause of the most frequently identified intestinal disease in the United States (26). Humans become infected by ingesting giardia cysts, which are the environmentally resistant stage of giardia (27). Giardia cysts can survive for prolonged periods. For example, Bingham et al. (28) documented cysts surviving in distilled water for 77 days at 8 °C. In another study, cysts of *Giardia muris* (a species that infects mice but is often used as analogue for *Giardia lamblia*), survived 28 and 56 days in lake water at depths of 4.5 meter (19 °C) and 9 m (6.6 °C), respectively. In cold river water (0–2 °C), cysts survived for 56 days (28). Typical concentrations of *Giardia* are between 1 and 100 cysts/mL in untreated domestic wastewater (18) and about 1 cyst/mL in polluted streams (19).

The *Giardia* cysts, which are 8–16 μm in diameter, are somewhat resistant to typical levels of wastewater treatment methods. During primary settling, only 0–53% of cysts are removed. Secondary treatment with clarification can remove 98.6–99.7% of cysts (28). Advanced tertiary treatment can further reduce the numbers of cysts by physical filtration and precipitation.

MICROSPORIDIUM

Microsporidia are unicellular protozoan parasites that infect a wide variety of animals, from insects and fish to every class of mammal, including humans. The vehicle for transmitting these organisms is its spores, which are shed from infected individuals (animal and human) via the urine, feces, respiratory sputum, and upon death and decay. Human infections are of concern in immunodeficient patients, especially those who are HIV+, and may infect a broad range of tissues. *Enterocytozoon bieneusi* and *Encephalitozoon intestinalis* are the most common species of microsporidia isolated from patients with chronic diarrhea attributed to microsporidiosis (29,30). Exact routes of transmission for human microsporidial infection have not been confirmed, but considerable evidence supports fecal–oral, sexual, respiratory, and waterborne routes (29–31).

The potential for waterborne transmission of spores is the focus of our research. At least two published studies have detected the presence of *E. bieneusi* (32) or *E. bieneusi, E. intestinalis*, and *Vittaforma corneae* (another human pathogenic microsporidium) (33,34) in surface water using the polymerase chain reaction (PCR). Hutin et al. reported in a case-controlled risk factor analysis of HIV+ individuals in France that the greatest risk for intestinal microsporidiosis was use of a swimming pool in the prior 12 months.

At least one outbreak of microsporidiosis has been described and attributed to contaminated drinking water originating from a particular treatment plant. According to Cotte et al. (31), an outbreak of intestinal microsporidiosis in the summer of 1995 affected 200 people (all apparently HIV+), or about 1% of the HIV+ population in the study area. Additionally, 15% of the 361 patients from whom microsporidia were identified during the 3-year study had no known immunodeficiency condition. The clustering of residences of infected individuals in the town of Lyon, France, led the authors to suggest that a water treatment facility serving the area may have been a contributing factor to these cases. This plant employs flocculation, ozoflotation, and filtration but not chlorination in the treatment process and uses surface water as a water source.

A few studies have described the susceptibility of human-pathogenic microsporidial spores to drinking water treatment methods and environmental conditions. Laboratory studies by Kucerova-Posisilova et al. (35) indicate that *E. intestinales* spores retain infectivity for at least 2 weeks at temperatures up to 33 °C. However, information on the viability of spores concentrated from the environment has yet to be published to our knowledge. A very recent study by Wolk et al. (36) determined that a chlorine concentration of 2 mg/L with exposure for at least 16 minutes resulted in a 99.9% reduction in the viability of *E. intestinalis* spores. Microsporidia were detected in four recreational water samples from Arizona; one was confirmed as *E. intestinalis*. In addition to the three recreational water samples, three irrigation water samples from Mexico and two from Arizona were positive for microsporidia. One of the samples from Mexico was confirmed as *E. intestinalis. V. corneae* spores were identified from irrigation water from Costa Rica and from secondary sewage from Tucson, Arizona (37).

In summary, mounting evidence demonstrates that microsporidia are found in water, may originate from human and animal reservoirs, and need to be considered as potential waterborne pathogens. As such, more information is needed on the incidence and survival of spores in the environment and through the treatment process.

BACTERIOPHAGE

Several bacteriophage transport studies have been conducted under well-characterized conditions [see recent review by Schijven and Hassanizadeh (38)]. The bacteriophages PRD-1 and MS-2 are often employed as surrogates for viruses of concern in human health (39,40) because of the hazards and costs associated with human viruses.

CALCULATING FLOW IN KARSTIC LIMESTONE AQUIFERS

In karstic limestone aquifers, groundwater often flows through highly permeable flow paths formed by dissolution along faults, fractures, bedding plane partings, or stratigraphic features. Compared with diffuse flow through granular aquifers, groundwater velocities in karstic systems can be very high, often in the range of miles per day. Consequently, karstic aquifers require a much larger wellhead protection area than common for wells in sand and gravel.

Multiple options exist for modeling flow and transport in karstic aquifers. For simple applications requiring only global water balances in steady-state conditions, existing modeling tools such as MODFLOW have served adequately. For modeling responses to storms or contaminant movement, more advanced models involving explicit flow features (40), irregular grids, and/or multiple interacting flow systems (41) are needed. Software tools necessary for explicit modeling of karst features have recently appeared (42).

TRANSPORT MODELS

Models for microbe transport and of colloid transport in general are analogous to solute transport models nonideal sorption terms account for rate limitations in the attachment/detachment processes. The relevant physical processes include advection, dispersion, attachment/detachment, physical filtration, inactivation, and advection facilitated by sorption on other types of colloids (38). Early models of microbe transport in saturated porous media used the equilibrium sorption assumption and an empirical distribution coefficient to model microbe attachment/release (43,44). These equilibrium sorption models performed poorly in case studies (45) and fail to reproduce the results of experiments that show unretarded breakthrough and slow (nonequilibrium) release (50). Modern models of microbe transport and of colloid transport in general use kinetic models for attachment/release. Single-site kinetic models are often used (46–49). Bales et al. (50) use a two-site model with one set of sites in equilibrium and the other kinetically controlled. Bhattacharjee et al. (55) and Schivjen et al. (56) use two-site models wherein attachment to both types of sites is kinetically controlled but with different rate constants. Schivjen et al. (56) analyze several laboratory experiments and clearly demonstrate that a two-site kinetic model is necessary to reproduce the measured breakthrough curves.

Most previous studies use the first-order rate law. Second-order rate laws have been used to model sorption of inorganic colloids (49) and microbes (55). In either case, attachment rates may be empirical or determined from mechanistic models of colloid filtration (46,49). Combinations of empirical rates and mechanistic rate models have also been used (48). When mechanistic models are used to calculate the attachment rate, models for colloid filtration in packed-bed reactors (50) are the usual choice.

Transport facilitated by other colloids is another process to be considered. Schivjen and Hassanizadeh (38) note that the removal rate for viruses declines with increasing travel distance, and that this nonlinear removal may be due to partial attachment to other colloidal particles. Colloid-facilitated transport is of particular concern in karst systems that often have large amounts of suspended sediments whose attributes are favorable for facilitating transport (51,52). Two studies (48) have addressed the effect of random spatial variability of hydraulic conductivity on microbe transport. These studies clearly provide important insights into the effect of spatial variability on microbe transport but are better suited to granular aquifers, where spatial variability can be more readily represented as a simple random space function.

BIBLIOGRAPHY

1. National Research Council: Committee on Science and Technology for Countering Terrorism. (2002). *Making the Nation Safer: The Role of Science and Technology in Countering Terrorism*. National Academy Press, Washington, DC, p. 251.
2. Katz, B.G., DeHan, R.S., Hirten, J.J., and Catches, J.S. (1997). Interactions between ground water and surface water in the Suwannee River Basin, Florida. *J. Am. Water Resour. Assoc.* **33**: 1237–1254.
3. Martin, J.B. and Screaton, E.J. (2001). *Exchange of matrix and conduit water with examples from the Floridian Aquifer*. United States Geological Survey Water-Resources Investigations Report 01-4011: 38–44.
4. Loper, D. (2001). *Steps Toward Better Models of Transport in Karstic Aquifers*. United States Geological Survey Water-Resources Investigations Report 01-4011: 56–57.
5. Meisel, J., Perra, R., and Meloigro, C. (1976). Overwhelming watery diarrhea associated with cryptosporidium in an immunosuppressed patient. *Gastroenterology* **70**: 1156–1160.

7. Avery, B.K. and Lemley, A. (1996). *Cryptosporidium: A Waterborne Pathogen*. Cornell University, Ithaca, NY, under the sponsorship of the U.S. Department of Agriculture's Working Group on Water Quality.
8. MacKenzie, W. et al. (1994). A massive outbreak in Milwaukee of Cryptosporidium infection transmitted through the public water supply. *N. Engl. J. Med.* **331**: 161–167.
9. Hoxie, N.J., Davis, J.P., Vergeront, J.M., Nashold, R.D., and Blair, K.A. (1997). Cryptosporidium-associated mortality following a massive waterborne outbreak in Milwaukee, Wisconsin. *Am. J. Public Health* **87**: 2032–2035.
10. Rose, J. (1997). Environmental ecology of Cryptosporidium and public health implications. *Annu. Rev. Public Health* **18**: 135–161.
11. Harter, T.W. and Atwill, E.R. (2000). Colloid transport and filtration of Cryptosporidium parvum in sandy soils and aquifer sediments. *Environ. Sci. Technol.* **34**: 62–70.
12. Reynolds, K.A. and Pepper, I.L. (2000). Microorganisms in the environment. In: *Environmental Microbiology*. R.M. Maier, I.L. Pepper, and C.P. Gerba (Eds.). Academic Press, San Diego, CA.
13. Brush, C., Walter, M., Anguish, L., and Ghiorse, W. (1998). Influence of pretreatment and experimental conditions on electrophoretic mobility and hydrophobicity of Cryptosporidium parvum oocysts. *Appl. Environ. Microbiol.* **64**: 4439–4445.
14. Medema, G.J., Bahar, M., and Schets, F.M. (1997). Survival of Cryptosporidium parvum, Escherichia coli, faecal enterococci and Clostridium perfringens in river water: Influence of temperature and autochthonous microorganisms. *Water Sci. Technol.* **35**: 249–252.
15. Casemore, D.P., Wright, S.E., and Coop, R.L. (1997). In: *Cryptosporidium and Cryptosporidiosis*. R. Fayer (Ed.). CRC Press, Boca Raton, FL, pp. 65–92.
16. Johnson, D.C., Reynolds, K.A., Gerba, C.P., Pepper, I.L., and Rose, J.B. (1995). Detection of Giardia and Cryptosporidium in marine waters. *Water Sci. Technol.* **31**: 439–442.
17. Kemp, J.S., Wright, S.E., and Bukhari, Z. (1995). On farm detection of Cryptosporidium parvum in cattle, calves and environmental samples. In: *Protozoan Parasites and Water*. W.B. Betts, D. Casemore, C. Fricker, H. Smith, and J. Watkins (Eds.). The Royal Society of Chemistry, Cambridge, UK.
18. Metcalf, R. and Eddy, I. (1991). *Wastewater Engineering*. McGraw-Hill, New York.
19. USEPA. (1998). *Comparative Health Risk Effects Assessment of Drinking Water*. United States Environmental Protection Agency, Washington, DC.
20. Kornich, D.G. et al. (1990). Effects of ozone, chlorine dioxide, chlorine, and monochloramine on *Cryptosporidium parvum* oocyst viability. *Appl. Environ. Microbiol.* **56**: 1423–1428.
21. Venczel, L.V., Arrowood, M., Hurd, M., and Sobsey, M.D. (1997). Inactivation of Cryptosporidium parvum oocysts and Clostridium perfringens spores by a mixed-oxidant disinfectant and free chlorine. *Appl. Environ. Microbiol.* **63**: 1598–1601.
22. Chauret, C., Nolan, K., Chen, P., Springthorpe, S., and Satter, S. (1998). Aging of Cryptosporidium parvum oocysts in river water and their suspectibility to disinfection by chlorination and monochloramine. *Can J. Microbiol.* **44**: 1154–1160.
23. Johnson, D.C. et al. (1997). Survival of Giardia, Cryptosporidium, poliovirus and Salmonella in marine waters. *Water Sci. Technol.* **35**: 261–268.
24. Mawdsley, J., Brooks, A., Merry, R., and Pain, B. (1996). Use of a novel soil tilting table apparatus to demonstrate the horizontal and vertical movement of the protozoan pathogen Cryptosporidium parvum in soil. *Biology and Fertility Soils* **23**: 215–220.
25. Brush, C.G., Anguish, L.J., Parlange, J., and Grimes, H.G. (1999). Transport of Cryptosporidium parvum oocysts through saturated columns. *J. Environ. Qual.* **28**: 809–815.
26. Adams, R. (1991). The biology of giardia. *Microbiol Rev.* **55**: 706–732.
27. Rusin, P., Enriquez, C.E., Johnson, D., and Gerba, C.P. (2000). Environmentally transmitted pathogens. In: *Environmental Microbiology*. R.M. Maier, I.L. Pepper, and C.P. Gerba (Eds.). Academic Press, San Diego, CA, pp. 447–489.
28. Deregnier, D., Cole, L., Schupp, D., and Erlandsen, S. (1989). Viability of Giardia cysts suspended in lake, river and tap water. *Appl. Environ. Microbiol.* **55**: 1223–1229.
29. Hutin, Y.J.F. et al. (1998). Risk factors for intestinal microsporidiosis in patients with human immunodeficiency virus infection: A case-control study. *J. Infect. Dis.* **178**: 904–907.
30. Franzen, C. and Muller, A. (1999). Molecular techniques for detection, species differentiation, and phylogenetic analysis of microsporidia. *Clin. Microbiol. Rev.* **12**: 243–285.
31. Cotte, L. et al. (1999). Waterborne outbreak of intestinal microsporidiosis in persons with and without human immunodeficiency virus infection. *J. Infect. Dis.* **180**: 2003–2008.
32. Sparfel, J.M. et al. (1997). Detection of microsporidia and identification of *Enterocytozoon bieneusi* in surface water by filtration followed by specific PCR. *J. Eukaryotic Microbiol.* **44**: 78S.
33. Dowd, S., Gerba, C., Kamper, M., and Pepper, I. (1998). Evaluation of methodologies including immunofluorescent assay (IFA) and the polymerase chain reaction (PCR) for detection of human pathogenic microsporidia in water. *Can. J. Microbiol.* **44**: 1154–1160.
34. Dowd, S.E., Gerba, C.P., and Pepper, I.L. (1998). Confirmation of the human-pathogenic microsporidia *Enterocytozoon bieneusi*, *Encephalitozoon intestinalis*, and *Vittaforma corneae* in water. *Appl. Environ. Microbiol.* **64**: 3332–3335.
35. Kucerova-Pospisilova, Z., Carr, D., Leitch, G., Scanlon, M., and Visvesvara, G. (1999). Environmental resistance of *Encephalitozoon* spores. *J. Eukaryotic Microbiol.* **46**: 11S–13S.
36. Wolk, D.M. et al. (2000). A spore counting method and cell culture model for chlorine disinfection studies of Encephalitozoon syn. Septata intestinalis. *Appl. Environ. Microbiol.* **66**: 1266–1273.
37. Gerba, C.P. (1999). Virus survival and transport in groundwater. *J. Ind. Microbiol. Biotechnol.* **24**: 247–251.
38. Schijven, J.F. and Hassanizadeh, S.M. (2000). Removal of viruses by soil passage: Overview of modeling, processes, and parameters. *Crit. Rev. Environ. Sci. Technol.* **30**: 49–127.
39. Gerba, C.P. (1984). Applied and theoretical aspects of virus adsorption to surfaces. *Adv. Appl. Microbiol.* **30**: 133–168.
40. Sudicky, E.A. and McLaren, R.G. (1992). The Laplace transform Galerkin technique for large-scale simulation of mass transport in discretely fractured porous formations. *Water Resour. Res.* **28**(2): 499–514.
41. Teutsch, G. (1993) An extended double-porosity concept as a practical modelling approach for a karstified terrain. In: *Hydrogeological Processes in Karst Terranes*. G. Gultekin,

A.I. Johnson, and W. Back (Eds.). *IAHS Publication 207*, Wallingford, UK, pp. 281–292.

42. Diersch, H-J.G. (2002). *Interactive, Graphics-Based Finite-Element Simulation System FEFLOW for Modeling Groundwater Flow, Contaminant Mass and Heat Transport Processes*. WASY Ltd., Berlin.
43. Park, N-S., Blanford, T.N., and Huyakorn, P.S. (1991). *VIRALT: A Model for Simulating Viral Transport in Groundwater, Documentation and User's Guide*. Version 2.0. Hydrogeol., Inc., Herndon, VA.
44. Tim, U.S. and Mostaghimi, S. (1991). Model for predicting virus movement through soils. *Ground Water* **29**(2): 251–259.
45. Yates, M.V. (1995). Field evaluation of the GWDR's natural disinfection criteria. *J. Am. Water Works Assoc.* **87**: 76–85
46. Dowd, S.E., Pillai, S.D., Wang, S.Y., and Corapcioglu, M.Y. (1998). Delineating the specific influence of virus iso-electric point and size on virus adsorption and transport through sandy soils. *Appl. Environ. Microbiol.* **64**(2): 405–410.
47. Sim, Y. and Chrysikopoulos, C.V. (1996). One-dimensional virus transport in porous media with time-dependent inactivation rate coefficients. *Water Resour. Res.* **32**(8): 2607–2611.
48. Rehmann, L.L.C., Welty, C., and Harvey, R.W. (1999). Stochastic analysis of virus transport in aquifers. *Water Resour. Res.* **35**(7): 1987–2006.
49. Saiers, J.E., Hornberger, G.M., and Liang, L. (1994). First- and second-order kinetics approaches for modeling the transport of colloidal particles in porous media. *Water Resour. Res.* **30**(9): 2499–2506.
50. Yao, K.M., Habibian, M.T., and O'Melia, C.R. (1971). Water and waste water filtration: Concepts and applications. *Environ. Sci. Technol.* **5**(11): 1105–1112.
51. Mahler, B.J., Lynch, L., and Bennett, P.C. (1999). Mobile sediment in an urbanizing karst aquifer: implications for contaminant transport. *Environ. Geol.* **39**(1): 25–38.
52. Mahler, B.J., Personne, J.C., Lods, G.F., and Drogue, C. (2000). Transport of free and particulate-associated bacteria in karst. *J. Hydrol.* **238**(3–4): 179–193.
53. Gerba, C.P. and Naranjo, J.E. (2000). Microbiological water purification without the use of chemical disinfection. *Wilderness Environ. Med.* **11**: 12–16.

PARTICULATE REMOVAL

Rasheed Ahmad
Khafra Engineering Consultants
Atlanta, Georgia

Coagulation is used in water treatment plants as a pretreatment step to remove particulate matter (such as clay and silt particles, bacteria, viruses, and protozoan cysts) and dissolved materials, especially natural organic material (NOM). Although the removal of particulate matter continues to be an important reason for using coagulation, recent concern about carcinogenic byproducts has emphasized that a coagulation process must also optimize the removal of NOM to reduce the formation of disinfection byproducts. Coagulation is a process for enhancing the tendency of particulate matter in aqueous suspension to attach to one another and/or to attach to collector surfaces. Coagulation promotes destabilization of surface charges on colloidal particles. Destabilization and aggregation of particulate matter and precipitation or adsorption of NOM in subsequent solid–liquid separation processes are the primary functions of a coagulation process. The coagulation process involves two steps: (1) the addition of chemical coagulants to destabilize particulate matter and react with NOM and (2) physical transport of collisions among particulate matter resulting in their aggregation or floc formation. In the water treatment literature, coagulation refers to all reactions and mechanisms that result in aggregation, and the physical transport step of producing interparticle aggregation is called flocculation. In water treatment plant operation, coagulation is achieved by rapid or flash mixing of coagulants followed by flocculation.

The other example of particle alteration and particle production process technology is chemical oxidation. The main application of this process is in iron and manganese removal. Iron and manganese are relatively soluble under reducing conditions, for example, in groundwater, stagnant surface water, and certain lakes. Dissolved iron and manganese are usually removed from water by oxidizing them under engineered conditions to their insoluble forms by adding an oxidant and removing the precipitated ferric hydroxide and manganese dioxide by sedimentation and filtration. The oxidants used most often for this are oxygen, chlorine, permanganate, chlorine dioxide, and ozone.

Sedimentation is a particle separation process (Fig. 1). In conventional water treatment systems, it follows flocculation and precedes filtration. Its purpose is to enhance the filtration process by removing particulate matter. Sedimentation requires that water flow through the basin at a slow enough velocity to permit the particulate matter to settle to the bottom of the basin before the water exits the basin. The equipment required for this process includes a rectangular, square, or circular settling basin. The basin includes provisions for inlet and outlet structures and a sludge collection system. In addition, sedimentation systems are optionally equipped with tube or plate settlers to improve performance. The settling velocity of the particles or particulate matter is governed by the particle size, shape, density, and water viscosity (which varies with temperature). The surface overflow rate is the primary design parameter for sizing sedimentation basins. This rate is defined as the rate of inflow (Q) divided by the tank surface or floor area (A). Units are typically rated in gallons per day per

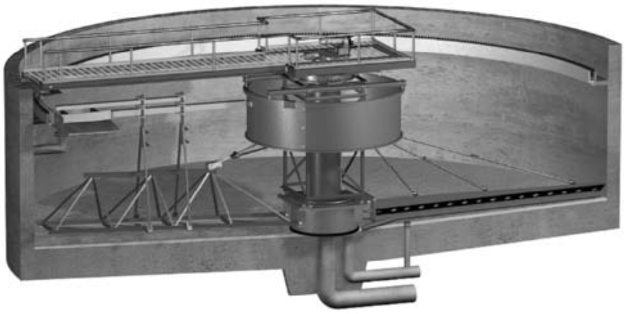

Figure 1. Circular sedimentation tank (*Source*: U.S. Filter).

square foot, gallons per minute per square foot, or cubic meters per hour per square meter. Design hydraulic overflow rates vary with the nature of the settling solids, water temperature, and hydraulic characteristics of the sedimentation basin. The settling efficiency for ideal conditions is independent of the depth and dependent on the tank plan or surface area. In reality, depth is important because it can affect flow stability if it is large and scouring if it is small.

In a conventional water treatment process train, filtration follows sedimentation. Particulate matter that is removed by this process includes microorganisms (bacteria, viruses, and protozoan cysts), clay and silt particles, colloidal and precipitated humic substances and other organic particulates from natural decay of vegetation, precipitates of aluminum or iron coagulation processes, calcium carbonate and magnesium hydroxide precipitates from lime softening, and iron and manganese precipitates. A number of different types of filters are used in water filtration, and they are described by various classification schemes. The most common filtration technologies used in water supply systems are rapid sand filtration, slow sand filtration, package plants, diatomaceous earth filtration (precoat filtration), membrane filters, and cartridge filters. High-rate granular filters or rapid sand filters are most widely used in conventional treatment and direct filtration (Fig. 2). Granular medium filtration is a water treatment process that uses a porous medium through which water passes to remove particulates or suspended solids. For granular medium filtration to be effective, source water must be pretreated. Chemical destabilization is an essential prerequisite for effective filtration. Chemicals used for particle destabilization are limited primarily to metal salts or cationic polymers as primary coagulants. Pretreatment may also include aeration or introducing an oxidant if water treatment aims to remove iron or manganese. Sometimes, a filter aid polymer is added in the influent to the filter to improve particle capture efficiency. Filtration by granular media consists of three principal mechanisms: (1) transport, (2) attachment, and (3) detachment. *Transport* mechanisms move a particle into and through a filter pore so that it comes very close to the surface of the filter medium or existing deposits where *attachment* mechanisms retain the suspended particle in contact with the medium's surface or with previously deposited solids. *Detachment* mechanisms result from the hydrodynamic forces of flow acting so that a certain portion of the previously attached particles, less strongly adhered to others, is detached from the filter medium or previous deposits and carried further, deep into or through the filter.

Flotation can be described as a particle separation process. It is a gravity separation process in which gas bubbles attach to solid particles to make the apparent density of the bubble–solid agglomerates less that of water, thereby allowing the agglomerate to float to the surface. The floated material (float) is removed from the surface, and clarified water is taken from the bottom of the flotation tank. Different methods of producing gas bubbles give rise to different types of flotation processes, which are electrolytic flotation, dispersed-air flotation, and dissolved air flotation (Fig. 3). Flotation is employed mainly for treating nutrient-rich reservoir waters that may contain heavy algal blooms and for low-turbidity, low-alkalinity waters.

In recent years, there has been considerable interest in using ultrafiltration (UF) or microfiltration (MF) membranes for particle separation in potable water treatment (Fig. 4). Similar to depth filtration, the performance of membrane filtration units is likely to be affected by the concentration and size of particles in raw water as well as the dissolved organic carbon

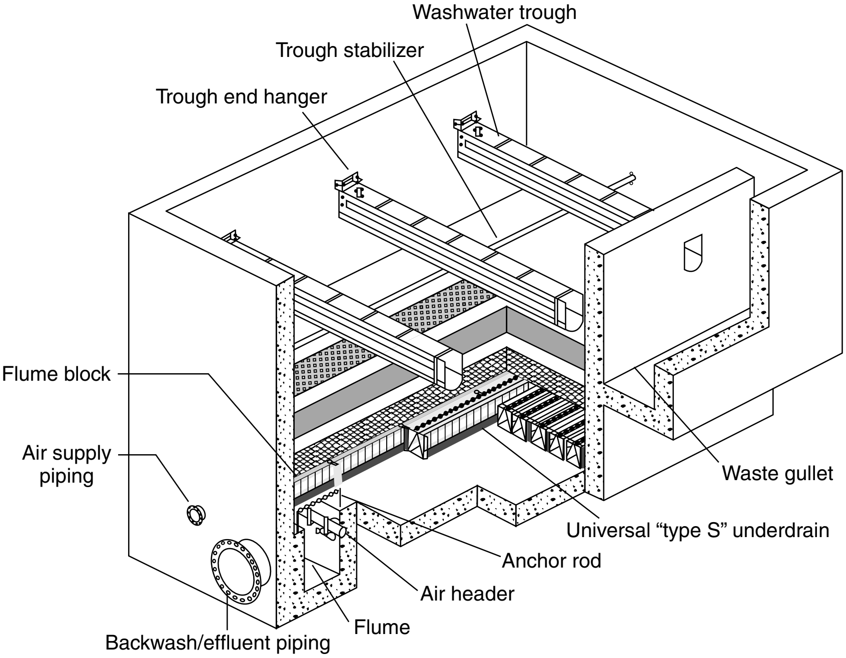

Figure 2. Rapid gravity filter box (*Source*: F.B. Leopold Co.).

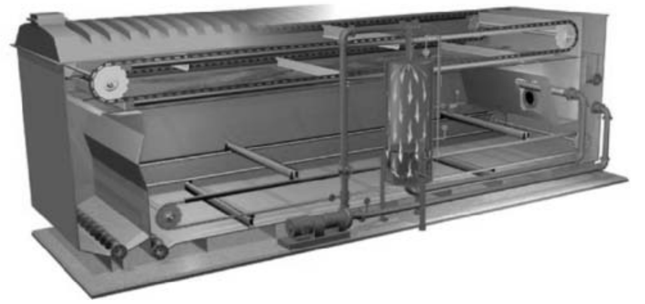

Figure 3. Dissolved air flotation (DAF) system (*Source*: U.S. Filter).

Figure 4. Membrane filtration system (*Source*: PALL).

(DOC). Particles of an effective size larger than the minimum pore size (or membrane cutoff) will be removed by the membrane. Permeate flux is closely related to the pressure drop across the membrane filter. Materials that accumulate on the membrane filter produce an additional layer of resistance to flow that is manifested as an increase in head loss or pressure drop across the membrane. A frequent backwash (every 15 min to 1 h) removes the cake formed on the membrane surface. In a conventional UF or MF process, the driving force to produce filtrate can work in two ways: Positive pressure moves fluid through the fibers; negative pressure moves fluid through fibers under vacuum pressure. The conditions under which membrane filtration processes might be preferable to granular medium filtration or conventional treatment are still largely unknown.

The selection of particulate removal technologies should be based on treatment objectives and control requirements. This process involves a number of basic data and other considerations, including effluent requirements, influent characteristics, existing system configuration, required investment, operation and maintenance (O&M), additional pretreatment and posttreatment processes, and waste management.

READING LIST

ASCE and AWWA. (1998). *Water Treatment Plant Design*. 3rd Edn. McGraw-Hill, New York.

AWWARF. (1998). *Treatment Process Selection for Particle Removal*. American Water Works Association Research Foundation, Denver, CO.

AWWA. (1999). *Water Quality and Treatment*. 5th Edn. McGraw-Hill, New York.

Hudson, H.E. (1981). *Water Clarification Processes*. Van Nostrand Reinhold Company, New York.

Kawamura, S. (2000). *Integrated Design of Water Treatment Facilities*. John Wiley & Sons, New York.

USEPA. (1990). *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*. Office of Drinking Water, Center for Research Information, Cincinnati, OH.

PHARMACEUTICALS IN WATER SYSTEMS

Kelly Pollack
Oladele Ogunseitan
University of California
Irvine, California

THE EMERGING PROBLEM OF PHARMACEUTICALS IN THE ENVIRONMENT

As the demographic pattern in most regions of the world changes in response to epidemiological transition, life expectancy is lengthened, and human reliance on medications to maintain the quality of life intensifies. It is now indisputable that widespread use of pharmaceutical products results in their release into the environment. This occurs primarily through the excretion of partially metabolized doses and the disposal of expired or unused medication (1–3). It is inevitable that that these fugitive releases lead to exposures of aquatic organisms and other natural ecosystem components. However, the driving force behind recent concerns is the potential for large-scale human exposures through natural drinking water recycling processes. Historically, most investigations of anthropogenic chemicals in the environment have focused on acutely toxic industrial chemicals from mining, energy, and agricultural industries. The environmental impacts of products emanating from the pharmaceutical industry have largely been excluded from scrutiny. The U. S. Geological Survey conducted the first national survey of pharmaceutical products in natural water systems in 2000. The results of that study confirmed prior suspicions about the widespread contamination of aquatic systems by these products (4).

The conclusions of several investigations of pharmaceuticals in the environment have now demonstrated an urgent need for rigorous ecological evaluation of potentially harmful prescription and over-the-counter products. These include hormones, antibiotics, analgesics, steroids, and their bioactive metabolites that are continuously introduced into the environment from the disposal of partially treated domestic wastewater. The fate of these compounds in the aquatic ecosystem depends in part on their chemical structure and physicochemical characteristics, including solubility in water and lipids, adsorption to soil particles, volatility, and potential biodegradability. These characteristics determine whether the chemicals will bioaccumulate through the foodweb, partition into sediments, mobilize through groundwater, or impact drinking water resources (5).

Information is sparse on the ecological effects of pharmaceutical products in the environment because the mode of action, dose response, and pharmacokinetic metabolism in humans are not fully understood for many medications. Therefore, it is difficult to predict with reasonable certainty, how specific pharmaceutical products will affect the myriad species that may become exposed following the release of pharmaceutical chemicals into the environment. However, there is a growing body of evidence that subtle, chronic effects from low-level exposures are occurring. These effects are occurring, in particular, from hormonal products and from chemicals that are known to induce the expression of phylogenetically conserved genes and the production of key metabolic enzymes (3).

For nontarget (e.g., aquatic) organisms whose chemical receptors are similar to those found in humans, it is possible to predict unintentional physiological effects of exposure to pharmaceutical products such as growth and reproductive hormones. In fact, estrogen receptors have been designed to monitor the occurrence and distribution of estrogenic pharmaceutical compounds in the environment. However, the sensitivity of such monitoring systems means that they can recognize chemicals from sources other than the pharmaceutical industry (6). Although understanding chemical interactions with biological receptors has been very useful for predictive modeling of the ecosystemic effects of pharmaceutical products, there is only rudimentary knowledge of potential synergistic and additive effects due to human or wildlife exposure to multiple pharmaceutical products found in many contaminated systems. In addition, unlike persistent organic pollutants from other industrial sectors, pharmaceutical contaminants do not necessarily have to be persistent to exert major physiological effects. This is due in part to the trained potency of their actions and the fact that concentrations in receiving waters are constantly replenished through the disposal of contaminated wastewater effluents (7). To initiate a thorough understanding of which pharmaceutical products are most likely to pollute water systems and which products are likely to provoke disturbances to ecosystemic and public health, either alone or in interactions with other chemicals, it is important to evaluate the societal distribution of pharmaceuticals with respect to prescription volume, doses, and disposal.

DISTRIBUTION OF PHARMACEUTICALS IN THE UNITED STATES

The pharmaceutical industry is a major contributor to the global economy; annual commercial transactions in prescription drugs in the United States approach $300 billion. In the year 2001, approximately 3.1 billion pharmaceutical prescriptions were filled for several thousand different medicinal chemicals in the United States alone (8). The top-ranked pharmaceuticals products prescribed in 2002 are presented in Table 1. The identity of popular prescriptions and their associated medical functions provides a reasonable summary of the societal burden of diseases. Attention to the quantity of drug doses multiplied by the number of each prescription is also a reasonable estimate of the quantity of potent drugs that enters the domestic sector annually. Very little attention has been paid to the question of what happens to the drugs after they leave the human body. In general, many drugs are prescribed at doses that can be supported by human metabolic capacity. Quite frequently, administered drugs and their partial metabolites leave the human body through excretion in urine and feces soon after ingestion. The period of excretion for some drugs may last for days (2,3). The list of products in Table 1 does not include those used in the agricultural industry, although they are important contributors to the distribution of pharmaceuticals into the environment, particularly antibiotics (9). Veterinary pharmaceutical products are routinely dispersed into the environment through the same pathways as human medications, but in many cases, there is no benefit of sewage treatment to reduce the concentration of influent pharmaceuticals because animal wastes can contaminate surface waters directly (2).

The top 20 pharmaceutical products found in streams during a reconnaissance project conducted by the U.S. Geological Survey between 1999 and 2000 are shown in Table 2. The data are limited by the narrow scope of chemicals investigated, compared to the main categories of prescriptions drugs presented in Table 1. Caffeine, the fourth ranked most commonly found organic chemical in surface waters, is the first in the list to be associated with pharmaceutical products, although its main source is dietary. Caffeine is an inducer of cytochrome P450 enzymatic activity in humans and in wildlife. Its presence in water systems is due to its resistance to biodegradation in many wastewater treatment processes. Therefore, caffeine has been used as a reliable tracer of wastewater effluent discharge, and it has been proposed as a monitoring sentinel for the distribution of pharmaceutical products in the environment (10,11).

FATE OF PHARMACEUTICALS IN SEWAGE TREATMENT FACILITIES

The capacity of human physiological systems to metabolize pharmaceutical products varies greatly from complete breakdown of a compound into various metabolites to excretion by the consumer in an essentially unaltered free form. The degree of metabolism is a function of chemical characteristics; the genetic, physiological, and dietary potential of the consumer; and the timing of the dose (2). Some pharmaceutical degradation products can be more bioactive than the parent compound, and the nontoxic conjugates can later be converted by microbial action to the original bioactive parent compound (4). It is also possible for microbial action to convert certain pharmaceutical products into chemicals that could resemble pollutants from other industries, thereby confusing monitoring programs and enforcement regulations. For example, (Fig. 1) shows the hypothetical pathway for the biodegradation of clofibrate, a hypolipidemic medicinal product commonly found in contaminated water systems when it enjoyed a

Table 1. Characteristics of Top Prescription Pharmaceutical Products in the United States[a]

Drug Category, 2001 Rank Order	Total Prescriptions for 2000, Millions	Leading Product	Manufacturer	Function	Prescriptions per Year (1999) (Millions)	Quantity per Dose, mg	Human Metabolite, % Excreted	Conc. in Surface Water, μg/L[b]
Codeine and combinations	124.2	Hydrocodone	Watson Laboratories	Pain reliever; antitussive and opioid analgesic	30.4	5–7.5	Codeine and morphine up to 10 μg/mL in urine	0.1–1.0 (codeine)
Cholesterol reducers	96.9	Lipitor (Atorvastatin)	Parke-Davis	Lipid-lowering	37.7	10–40	Varies	No data
Synthetic thyroid hormone	70.2	Synthroid (Levothyroxine)	Knoll	Thyroid hormone	41.1	50–200 μg	Varies	No data
Estrogen/Progesterone	79.1	Premarin (conjugated estrogens)	Wyeth-Ayerst	Reproductive hormone therapy	47.8	0.3–2.5 (tablet)	Varies	
						25 (injection)		
		Estrone						0.11
		Equilin						0.15
		17-a-Dihydroequilin				0.625/g (cream)		
		17-a-Estradiol						
		Equilenin						0.03
		17-a-Dihydroequilenin						0.14
Beta blockers	87.6	Atenolol (Tenormin)	Geneva Pharm	Antihypertensive; antianginal	12.4	50	Varies	No data
Aminopenicillins	62.3	Trimox	Apothecon	Antibiotic	24.8	125–500	Up to 0.73	No data
Diuretics	49.6	Furosemide (Lasix)		Hypertension	16.0	10	Varies	No data
		Hydrochlorothiazide	Parke-Davis					
Calcium blockers	94.1	Amlodipine (Norvasc)	Pfizer	Hypertension	27.1	5–10	Varies	No data
Steroid antibronchospasm	–	Albuterol Ventolin	Warrick	Asthma	–	0.5% or 200 μg	20–70% of absorbed dose in urine	No data
Benzodiazepines	66.6	Alprazolam (Xanax)	Pharmacia/Upjohn	Anxiety disorder	1.8	0.25–2	Varies	No data
Proton pump inhibitors	60.1	Prilosec	Astra	Proton pump inhibitor (heartburn)	31.1	20–40	Varies	No data
Propoxyphene compound	–	Propoxyphene HCl		Narcotic	–	65	Varies	No data
		Aspirin		Analgesic		389		
		Caffeine		Pain relief		32.4		
Selective serotonin reuptake Inhibitor/Selective Norepinephrine reuptake inhibitors (SSRI/SNRI)	96.4	Prozac	Lilly	Antidepressant	24.7	20–80	Varies	No data
		Zoloft	Pfizer	Antidepressant	23.1	50–200		
ACE inhibitors	94.1	Zestril (Lisinopril)	Zeneca	Hypertension	20.7	2.5–40	Varies	No data
Antiarthritic	61.3	Celebrex	Searle	Rheumatoid arthritis; osteoarthritis	17.5	100–400	Varies	No data
Antihistamines	59.1	Claritin	Schering	Environmental allergies	25.4	10	Varies	No data
Seizure disorders	57.0	Lamictal	Glaxo	Epilepsy and seizures	–	12.5–200	Varies	No data
		Dilantin	Parke-Davis			30–100		

[a] Reference 8.

[b] Reference 4.

high prescription rate in the 1980s. Human metabolism of clofibrate produces clofibric acid, which is excreted in feces and disposed of into the domestic sewer system. The structure of clofibric acid suggests that heterotrophic bacteria can cleave the side chain to produce chlorobenzoate and/or chlorobenzene, which are chemicals commonly associated with the petrochemical solvent industry.

Pharmaceutical products, their synthetic precursors, and biotransformation products are continuously released into the environment through consumer excretion and through the disposal of unused or expired medications. Although sewage treatment facilities are the most obvious sinks for pharmaceutical products, municipal landfills are also important sinks for solid products that are disposed of with domestic solid waste (3). In addition, episodic leakage of sewer lines and overflow of veterinary solid waste contaminated with pharmaceuticals may lead to the release of completely untreated medicines into the environment (5). In the United States, about a million homes do not have sewage treatment systems but instead rely on direct discharge of raw sewage into streams by “straight-piping” (3). Some Canadian cities reportedly discharge 3.25 billion liters per day of essentially untreated sewage into surface waters and the ocean (3).

Sewage treatment generally processes human waste in three steps: primary (clarifying), secondary (aerobic biodegradation of organic matter), and tertiary (finishing step to reduce the concentrations of specific pollutants such as nitrates and phosphorus to acceptable standards). The liquid effluent is generally discharged by release into surface freshwater systems or the ocean. Waste from biomass and other precipitated solids are treated as sludge, which is used in many different ways. Most domestic sewage treatment facilities are not designed to process specialty chemicals such as pharmaceutical products. It is also possible that certain pharmaceutical products may be sufficiently toxic to inhibit the action of activated sludge microorganisms that are responsible for degrading organic materials in sewage. If undegraded pharmaceutical products in sewage are precipitated, they may end up in the sludge material. This is sometimes used to fertilize soils, raising the complex issues of soil contamination, groundwater infiltration, and human exposure through agricultural products. In addition, many municipalities do not process sewage completely through the tertiary stage, and the effluent is used for landscape irrigation, again raising the possibility of human and ecosystemic exposures. Finally, the process of chlorinating completely treated sewage and recycling the water for drinking is increasingly attractive as many regions experience water shortages. Chlorination of pharmaceutical products may produce highly toxic chemicals. For example, there is some evidence that the chlorination of caffeine will result in the production of chlorocaffeine, a potential mutagen (10,11).

Many pharmaceutical compounds survive biodegradation in sewage treatment plants. Some products are as persistent as more widely studied organochlorine pollutants, but only few large-scale studies have been completed. The removal efficiencies of pharmaceuticals in

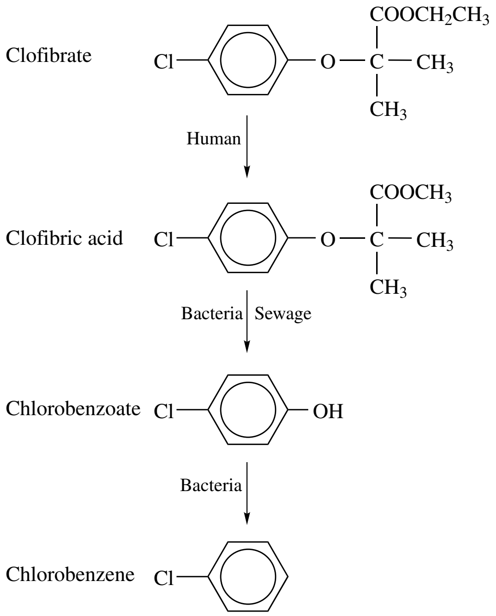

Figure 1. The production of chemical pollutants usually attributed to other industries may result from human and microbial metabolism of pharmaceutical products. The example shows the hypothetical degradation of clofibrate to chlorobenzene.

Table 2. Top 20 Organic Contaminants (Including Pharmaceuticals) Found in Streams in 39 States from 1999–2000 in the United States[a]

Rank	Chemical Compound	Source
1	Coprostanol	Fecal steroid/Hormone
2	Cholesterol	Human steroid
3	*N*-*N*-diethyltoluamide	Insect repellent
4	Caffeine	Dietary stimulant, pharmaceutical additive
5	Tri (2-chloroethyl) phosphate	Fire retardant
6	Triclosan	Antimicrobial disinfectant
7	4-Nonylphenol	Nonionic detergent metabolite
8	4-Nonylphenol monoethoxylate	Nonionic detergent metabolite
9	Ethanol, 2-butoxy-phosphate	Plasticizer
10	4-Octylphenol monoethoxylate	Nonionic detergent metabolite
11	Bisphenol A	Plasticizer
12	Cotinine	Tobacco cigarettes (nicotine metabolite)
13	4-Nonylphenol diethoxylate	Nonionic detergent metabolite
14	5-Methyl-1H-benzotriazole	Anticorrosive agent
15	Fluoranthene	Polyaromatic hydrocarbon oils
16	1,7-Dimethylxanthine	Caffeine metabolite
17	Pyrene	Polyaromatic hydrocarbon oils
18	Trimethoprim	Antibiotic
19	1,4-Dichlorobenzene	Deodorizer
20	Diazinon	Insecticide

[a] Reference 4.

treatment plants are still largely unknown (4). However, a study in Germany demonstrated that the removal efficiencies of pharmaceuticals by sewage treatment might be as low as 60% (2). There are many possible reasons for the low efficiency. Most pharmaceutical products are designed for metabolism at the human body temperature of 37 °C, which is not usual in sewage treatment facilities or anywhere else in nature. The identification of metabolic products is difficult because different metabolites can result under different environmental conditions and standard reference compounds are difficult to synthesize (2). The ability of treatment facilities to degrade pharmaceuticals can also change on the basis of operational state and seasons. Wet weather runoff and overflows from plant failure or overcapacity can also lead to direct release. Finally, influent concentrations of most pharmaceutical products may be so low that the microbial enzymes that degrade them are either not induced, or other more abundant substrates compete for their actions (2).

In the most comprehensive reconnaissance project that has been conducted to date, the U.S. Geological Survey investigated the occurrence of medicines, hormones, and other organic wastewater contaminants in a network of 139 streams across 39 states (4). Whereas previous research has shown that antibiotics, prescription drugs, and nonprescription drugs can be present in streams, this was the first study to examine their occurrence in a wide variety of hydrogeologic, climatic, and land-use settings across the United States. Although the data covered only a limited number of pharmaceutical compounds, it is reasonable to assume that many other compounds survive wastewater treatment and biodegradation and are ultimately released into water systems.

DETECTION OF PHARMACEUTICALS IN WATER SYSTEMS THAT RECEIVE EFFLUENT

There is a paucity of data on the distribution and fate of pharmaceuticals in natural water systems because, until recently, there have been few analytical methods that are sufficiently sensitive to detect these compounds at the expected low concentrations (4). In addition to sensitivity, analytic methods specifically developed for pharmaceutical products must also be able to discriminate among anthropogenic and natural constituents. For example, the use of artificial reproductive hormones as birth control pills and to control osteoporosis in postmenopausal women is widespread and has probably contributed significantly to the environmental distribution of these human hormones. However, it is sometimes difficult to discriminate between the artificial variety from the pharmaceutical industry and natural hormones excreted by humans, especially after partial metabolism by microbial action (2). The pending creation of mass spectral libraries for environmental pharmaceutical products and their degradation products will facilitate the development and reliability of methods in this direction. Ecotoxicological assessment of pharmaceutical products in the environment also requires developing sentinel and monitoring species that are expected to concentrate certain chemical moieties reliably or to respond to their effects at the physiological level (2).

In the USGS study, five analytical methods were used to measure concentrations of 95 organic wastewater contaminants (of which pharmaceuticals were a major part) in water samples. The methods used for measuring contaminants in the water included solid-phase extraction (SPE) with liquid chromatography/mass spectrometry positive-ion electrospray [LC/MS-ESI (+)] analysis and whole-water continuous liquid–liquid extraction (CLLE) with capillary gas chromatography/mass spectrometry (GC/MS) analysis (5). Although sampling locations were biased toward locations that were downstream of urban centers and livestock production, organic wastewater contaminants were still present in 80% of the streams that were sampled. There was a median of 7 and as many as 38 contaminants found in the water samples. Nonprescription drugs were found at the greatest frequency. Antibiotics, other prescription drugs, and reproductive hormones were found at relatively similar frequencies. It was hypothesized that the greater frequency of detection for nonprescription drugs might have been due in part to suspected greater annual use compared to other compounds. Measured concentrations rarely exceeded drinking-water guidelines, drinking-water health advisories, or aquatic-life criteria. However, most of the 95 contaminants do not have guidelines or established safe-level concentrations because not much is known about the ecotoxicological effects of the contaminants investigated (5).

ECOLOGICAL EFFECTS OF PHARMACEUTICALS IN WATER SYSTEMS

Acute toxicity is one of the many possible ecotoxicological end points of the exposure of nontarget species to pharmaceutical products synthesized strictly for human consumption. The low levels of pharmaceuticals found in polluted waters that contain many different species is not a reason for complacency in the expectation of toxicity (2). The low concentrations may not always pose acute risk, but the cumulative risks of prolonged exposure to low levels of chemicals is known to represent considerable hazard in animal studies and human epidemiology. Laboratory studies that do not simulate simultaneous exposures to multiple pharmaceutical products are likely to underestimate their ecotoxicological impacts. In most cases, there is limited information on which receptors in nontarget organisms are sensitive to chemical exposure. For these organisms, the most important threats are subtle behavioral modifications or genetic alterations that can lead to profound long-term changes in ecosystemic structure and function. Many such subtle effects have been demonstrated in wildlife as a result of chemical exposures. These include reversal of attraction, boldness, contact avoidance, feeding and mating disturbances, and confused directional sensing (2).

The environmental distribution of antidepressants and chemical regulators of obsessive–compulsive behavior can have subtle effects on nontarget organisms at chronic low doses. Serotonin reuptake inhibitors are a major component of widely prescribed antidepressants, including Prozac, Zoloft, Luvox, and Paxil. In addition to playing an important role in mammalian neurotransmission, serotonin is involved in a wide array of physiological regulatory functions in many species. In bivalves, serotonin regulates reproductive functions, including spawning, oocyte maturation, and parturition. At molar concentrations of 10^{-4} to 10^{-3}, serotonin induces spawning. Serotonin stimulates the release of neurohormones in crustaceans. Prozac and Luvox induce spawning and spawning behavior in zebra mussels at very low concentrations. In lobsters, serotonin causes behavioral reversal by stimulating subordinates to engage in fighting against dominants by reducing their propensity to retreat (2).

Kolpin and co-workers (4) detected compounds in water systems known or suspected to have weak hormonal activity and the potential to disrupt normal endocrine function. These compounds were detected in practically all streams investigated during the study. The expected low levels of exposure (<0.001 μg/l) to specific hormonally active compounds can adversely affect the reproduction of aquatic species. Other potential ecological problems of pharmaceutical exposure include increases in the incidence of cancers and the proliferation of antibiotic-resistant bacteria.

In addition to possible ecotoxicological effects, the prospects of human health effects from ingesting subtherapeutic doses of multiple pharmaceuticals from every intake of drinking water or through inhalation of volatilized products from every bath raises substantial concern. For many pharmaceutical products, the potential effects on humans and aquatic ecosystems are not clearly understood. Many drugs currently detected in aquatic waters already have known effects on nontarget organisms. Synthetic oral contraceptives (17α-ethynylestradiol) generally occur at concentrations lower than 7 μg/liter in effluents from publicly owned waste treatment facilities. In combination with steroidal estrogens 17β-estradiol and estrone, exposures to 17α-ethynylestradiol, it is believed, causes feminization in male fish, as observed in sewage treatment lagoons in the mid-1980s (2).

The consolidation and proliferation of resistance to multiple antibiotics in pathogenic microorganisms is an increasing threat to human reliance on antibiotics for treating dangerous infectious diseases. Acquired antibiotic resistance may remain stable indefinitely, even after the sources of exposures are removed. This means that the threats posed by microbial antibiotic resistance can only grow in the future, even if strict controls are placed now on the environmental release of antibiotics. For example, bacterial isolates from wild geese that have never been treated with antibiotics were resistant to ampicillin, tetracycline, penicillin, and erythromycin (2). Sufficiently high concentrations of multiple antibiotics can also have ecotoxicological effects on microbial communities that sustain global biogeochemical cycles and many local ecosystem functions, including the detection of indicators of fecal contamination in water systems.

REGULATORY STRATEGIES FOR CONTROLLING PHARMACEUTICAL PRODUCTS IN WATER SYSTEMS

Pharmaceutical products target more than 500 distinct biochemical receptors in the human body. Many of these receptors are also found in aquatic organisms (2). However, current guidelines for approving the effectiveness and metabolic fate of pharmaceutical products do not necessarily include testing across phylogenetic branches or documenting potential bioaccumulation and biodegradation. Even among humans, unrecognized genetic polymorphisms may make certain individuals much more sensitive to the effects of certain drugs, to the extent that a "zero tolerance" guiding principle for pharmaceuticals in domestic water supplies does not appear unreasonable. However, such broad regulatory goals are deemed too expensive to be implemented effectively.

Although the risks posed by inadvertent exposures to pharmaceutical products are not specifically controversial, only a limited number of federal agencies are responsible for managing the release of pharmaceuticals into the environment. In Europe, the Organization for Economic Cooperation and Development (OECD), an intergovernmental organization that has representatives in 29 countries, regularly publishes "Test Guidelines" for assessing the hazards of chemicals, including pharmaceuticals, in various environmental contexts (2). The problems of pharmaceutical products in the environment were first recognized widely in Europe, and the leading edges of regulatory issues appear to be evolving similarly from Europe. The European Union (EU) first expressed concern for the release of veterinary pharmaceuticals and their metabolites into the environment in the early 1980s because these products, it was perceived, have a much more direct route of introduction to the environment than medications prescribed for humans. The EU established two assessment levels to determine if a veterinary drug is dangerous to the environment. The first level considers the potential for release; the second level evaluates fate in various environmental compartments and investigates effects on specific biota that are likely to be exposed. All new applications for veterinary pharmaceuticals must include an environmental impact report based on the two levels of assessments. Regulatory guidelines for human pharmaceuticals are pending ratification by the EU (2).

In the United States, the Food and Drug Administration (FDA) and the U.S. Environmental Protection Agency (EPA) have separate roles regarding the distribution and release of pharmaceutical products in the environment. The FDA is responsible for assessing the risk to the environment caused by the manufacture, use, and disposal of human and animal drugs (5,12). FDA approval for new medications requires the submission of Environmental Assessment (EA) reports under the National Environmental Policy Act of 1969. As with the European Commission, the concern for new drugs rests primarily on acute and chronic effects, as measured by traditional toxicity tests (2). Much less concern has been given for more subtle behavioral effects. The FDA uses a tiered approach, similar to that of the EU, to determine if regulatory action is required. Generally, an Environmental Assessment report is required if the expected environmental concentration of the active ingredient of the drug in the aquatic environment exceeds 1 ppb (5,12). More recently, published literature and various conferences have indicated that in addition to the EA reports, pharmaceutical chemicals might have long-term effects on various wildlife species, but firm regulation will require additional research to evaluate these effects (12).

One of the top five goals of the EPA's Strategic Plan 2000 is identifying emerging risks of previously unknown, unrecognized, unanticipated, or unsuspected chemical pollutants in the environment (2). The EPA regulates emissions and effluent discharges from pharmaceutical manufacturing facilities, which are subjected to various restrictions or schedules of compliance regarding processed or raw effluent discharges (5,6). Despite considerable progress in recognizing and regulating risks posed by pharmaceuticals in water systems, important shortcomings remain in the current approaches used to determine ecological risk. Additive, synergistic, or antagonistic impacts of similar drugs or different classes of drugs that affect the same receptors are not currently being evaluated or considered part of regulatory principles. Additionally, the FDA's expected environmental concentration limit of 1 ppb for any given drug could easily be exceeded when the cumulative concentrations of similar-mode-of-action medications are considered. This regulatory loophole may allow potentially damaging chemicals to skip through the environmental assessment process.

In conclusion, the multidisciplinary demands of understanding and regulating the design, approval, consumption, metabolism, excretion, disposal, and environmental fate of pharmaceutical products in water systems requires developing innovative strategies in academic and regulatory institutions. These strategies may be consolidated under the umbrella of a new subdiscipline entitled "pharmecology" whose ultimate mission is protecting the environment from the chemical hazards associated with the pursuit of human health.

Acknowledgment
The Program in Industrial Ecology at the University of California, Irvine, supported this work.

BIBLIOGRAPHY

1. Buser, H-R., Poiger, T., and Müller, M.D. (1999). Occurrence and environmental behavior of the pharmaceutical drug ibuprofen in surface waters and in wastewater. *Environ. Sci. Technol.* **33**: 2529–2535.
2. Daughton, C.G. and Ternes, T.A. (1999). Pharmaceuticals and personal care products in the environment: Agents of subtle change? *Environ. Health Perspect.* **107**(Suppl. 6): 907–938.
3. Kummerer, K. (Ed.). (2001). *Pharmaceuticals in the Environment*. Springer, Berlin, p. 265.
4. Kolpin, D.W. et al. (2002). Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999–2000: A national reconnaissance. *Environ. Health Sci. Technol.* **36**(6): 1202–1211.
5. Daughton, C.G. and Jones-Lepp, T.L. (2001). *Pharmaceuticals and Personal Care Products in the Environment: Scientific and Regulatory Issues*. ACS Symposium Series 197, American Chemical Society, Washington, DC, pp. 2–53; 264–281.
6. Soto, A.M. et al. (1995). The E-SCREEN assay as a tool to identify estrogens: An update on estrogenic environmental pollutants. *Environ. Health Perspect.* **103**: 113–122.
7. Daughton, C.G. (2003). Cradle-to-cradle stewardship of drugs for minimizing their environmental disposition while promoting human health. 1. Rationale for and avenues toward a green pharmacy. *Environ. Health Perspect.* **111**(5): 757–774.
8. Rx List: The Internet Drug List. (2003). *The Top 200 Prescriptions for 2002 by Number of Prescriptions Dispensed*. ONLINE 2003. Available: http://www.rxlist.com/top200.htm.
9. Hirsch, R., Ternes, T., Haberer, K., and Kratz, K.L. (1999). Occurrence of antibiotics in the aquatic environment. *Sci. Total Environ.* **225**: 109–118.
10. Ogunseitan, O.A. (1996). Removal of caffeine in sewage by *Pseudomonas putida*: Implications for water pollution index. *World J. Microbiol. Biotechnol.* **12**: 251–256.
11. Ogunseitan, O.A. (2002). Caffeine-inducible enzyme activity in *Pseudomonas putida* ATCC 700097. *World J. Microbiol. Biotechnol.* **18**: 423–428.
12. Velagaleti, R., Burns, P.K., Gill, M., and Prothro, J. (2002). Impact of current good manufacturing practices and emission regulations and guidances on the discharge of pharmaceutical chemicals into the environment from manufacturing, use, and disposal. *Environ. Health Perspect.* **110**(3): 213–220.

POINT-OF-USE/POINT-OF-ENTRY SYSTEMS (POU/POE)

Zacharia Michael Lahlou
Technical Assistance Consultant

Numerous households use point-of-use/point-of-entry (POU/POE) systems primarily to deal with aesthetic concerns, such as taste and odor. These treatment devices are installed just as their name implies—at the point where water enters a household or where it is used, such as a faucet. In certain situations, however, using POU/POE systems to provide safe drinking water to a system's customers is not an individual's choice, but that of the water system cooperating with regulatory authorities. Therefore, this "Tech Brief" only discusses POU/POE treatment options that meet Safe Drinking Water Act (SDWA) water quality requirements.

WHAT IS THE ROLE OF POU/POE SYSTEMS?

Some small water systems find complying with U.S. Environmental Protection Agency (EPA) maximum contaminant level (MCL) requirements difficult. Installing expensive treatment technologies may be the largest obstacle for these systems. In many cases, obtaining water from some other source may not be an option. In these situations, POU/POE water treatment systems may provide a low-cost alternative to centralized water systems.

Water system personnel can install POE treatment units at individual households where water lines enter the home, alleviating the expense of treating large amounts of water at a central facility. An even less expensive alternative is the POU system, as these systems only treat the water at an individual tap.

POU/POE systems are used to control a wide variety of contaminants in drinking water and often use the same technology concepts employed in centralized treatment—but at a much smaller scale. This technology is applied to reduce levels of organic contaminants, control turbidity, fluoride, iron, radium, chlorine, arsenic, nitrate, ammonia, microorganisms (including cysts) and many other contaminants. Aesthetic factors, such as taste, odor, or color, can be improved with POU/POE treatment.

In addition to treating raw water, POU/POE systems can be used to treat finished water that may have degraded during distribution or storage. They insure that susceptible consumers, such as the very young or immunocompromised, receive safe drinking water.

POU/POE can save many small communities money where individual households have private wells. The community then does not have to build a treatment plant or install and maintain water distribution mains. Many states, however, have concerns about POU/POE treatment devices, such as:

- how well the units treat drinking water,
- the potential health risk posed by not treating all the water in the house, and
- water system officials' ability to properly monitor and maintain the equipment.

WHAT REGULATIONS AFFECT POU/POE SYSTEMS?

The 1996 SDWA lists POU/POE systems as options for compliance technologies. When a water system uses this technology to comply with a National Primary Drinking Water Regulation (NPDWR) the SDWA identifies requirements that must be met. The SDWA states the public water system or a person under contract with the public water system shall own, control, and maintain the POU/POE system to ensure proper operation, maintenance, and compliance. The act also states that the treatment mechanism should be equipped with mechanical warnings that automatically notify customers of operational problems.

Other conditions in the SDWA include: "If the American National Standards Institute has issued product standards applicable to a specific type of POE/POU treatment unit, individual units of that type shall not be accepted for compliance with a MCL or treatment technique unless they are independently certified in accordance with such standards."

POU devices are listed as compliance technologies for inorganic contaminants, synthetic organic contaminants, and radionuclides. POU devices are not listed for volatile organic contaminants because they do not address all routes of exposure.

WHAT ARE THE TYPES OF POU/POE SYSTEMS?

No single type of residential water treatment system is available to remove all water quality problems. Therefore, selection of one or more technologies may be necessary to solve multiple water quality problems. The following is a short summary of different types of POU/POE available (see also Table 1). Additional information and guidance is available from manufacturers, distributors, and public health agencies to help select the appropriate treatment technologies to remove specific water contaminants from water.

FILTERS

Water passes through the filter media (usually in a cartridge in smaller units), which either adsorbs or physically screens various contaminants. Common filter media include:

Table 1. Summary of POE/POU Systems and Costs (NSF, 1999)

Technology	Some Contaminants Removed	Initial Cost	Operating Cost	Operating & Maintenance Skills
Chlorine	Microbial	**$**	**$**	**$**
UV, Ozone	Microbial	**$$**	**$**	**$$**
Cartridge	Protozoa			
Filter	Bacteria	**$**	**$ to $$**	**$**
Reverse Osmosis	Microbial, Inorganic Chemicals and Metals Radium, Minerals, Some Organic Chemicals	**$$**	**$$$**	**$$$**
Distillation	Microbial, Inorganic Chemicals and Metals, Minerals, Some organic Chemicals, Radium, Uranium	**$$**	**$$**	**$**
Activated Carbon	Organic Chemicals, Radon, Odors (solid block can filter protozoa and some bacteria)	**$$**	**$$ to $$$**	**$**
Packed Tower Aeration	Radon, Volatile Organic Chemicals, Tastes, Odors	**$$**	**$**	**$$$**
Ion Exchange	Inorganic Chemicals, Radium, Nitrate	**$$**	**$$ to $$$**	**$$**
Activated Alumina	Arsenic, Selenium, Fluoride	**$$$**	**$$$**	**$$$**
		$ Low	**$$ Moderate**	**$$$ High**

- Granular activated carbon (GAC) is used for taste and odor control and to remove radon and regulated organic compounds.
- Solid block carbon treats the same contaminants as GAC, but also to removes lead, asbestos, various bacteria, cysts, and sediment particulates.
- Ceramic or synthetic fiber microfilters treat various bacteria, cysts, and sediment particulates.
- Activated alumina treatment is most often used for fluoride, selenium, silica, and arsenic removal.

REVERSE OSMOSIS

Water passes through a synthetic, semi-permeable membrane that filters all pathogens and most organic and inorganic contaminants. Reverse osmosis units must have a means of discharging filtered matter to a drain. The discharge line should be installed with an air gap so a cross-connection between wastewater and drinking water will not occur.

DISTILLATION

Distillers heat water in one chamber and turn it into steam. The steam then passes into another chamber where it is cooled and condensed to a liquid. Distillation can effectively remove microorganisms, dissolved minerals, metals, nitrates, and some organic contaminants. Distillation units require a dependable supply of electricity and usually produce only small amounts of drinking water.

ION EXCHANGE

Ion Exchange, commonly known as water softening, is used to treat all household potable drinking water. Ions of either sodium or potassium, stored in the softener's "resin bed" are exchanged for ions of the calcium and magnesium hardness minerals. Ion exchange can be used for dealkalization and to remove iron and manganese, heavy metals, some radioactivity, nitrates, arsenic, chromium, selenium, and sulfate.

DISINFECTION AND OXIDATION

Oxidizing chemicals, such as chlorine and ozone, are added to water through a feed system that controls the concentration and allows appropriate contact time. These chemicals break down organic contaminants and destroy pathogens.

Ultraviolet light (UV) is a popular disinfection method in combination with other treatment techniques. UV uses rays of ultraviolet light to deactivate pathogens. UV light damages a pathogen's DNA and prevents it from reproducing. One of the major advantages of UV disinfection is that it disinfects without the addition of chemicals; therefore, it does not generate taste, odor, or chemical by-products.

AIR STRIPPING OR AERATION

Air stripping has been used in POE systems to remove volatile organic compounds, hydrogen sulfide, and radon from water. Air stripping is a treatment method that exposes water to air. This treatment process removes or "strips" volatile organic contaminants from groundwater as air is forced through the water, causing the compounds to evaporate. GAC alone can remove volatile organics but can only be operated for short periods before the carbon has to be replaced. For these applications, it is important to vent gases adequately to avoid creating an air pollution hazard inside the home.

POU/POE OPERATION AND MAINTENANCE

Selecting POU/POE systems does not eliminate the need for evaluating treatment efficiency before the units are installed. For systems that employ cartridges (e.g., GAC columns or activated alumina), source water pilot testing may be necessary to develop valid estimates of the unit's service life.

Effective operation, maintenance, and monitoring programs are especially significant for POU/POE systems. Many homeowners assume their systems will perform properly once installed and do not understand the level of effort required to ensure proper operation. For this reason, when POU/POE systems are installed for regulatory purposes, water utilities or regulatory agencies must provide programs for long-term operation, maintenance, and monitoring.

Proper installation is the first step in effective long-term operation and maintenance (O&M) of POU/POE systems. Experienced contractors or installers whose products conform to applicable plumbing codes should be the only personnel who install the units. Qualified installers:

- carry liability insurance for property damage during installation,
- are accessible for service calls,
- accept responsibility for minor adjustments after installation, and
- give a valid estimate of the cost of installation.

After installation, POU/POE systems should have a well-defined program of O&M for continued production of high quality drinking water. The equipment manufacturer's recommended O&M requirements can serve as the basis for the O&M program. Equipment dealers may provide maintenance for a limited time period as part of an installation warranty. A local plumbing contractor, a POU/POE service representative or equipment dealer, a water service company, the local water utility, or a circuit rider may carry out a long-term maintenance program.

Monitoring programs need to be site specific and reflect the contaminant or contaminants being removed, the equipment used, the number of POE/POU units in service, and the logistics of the service area.

Minimum sampling frequencies and types of analyses should be established in cooperation with the local health department, the state regulatory agency, and the treatment system.

Monitoring programs generally include:

- raw and treated water sample collection,
- meter reading,
- field analyses (measuring pH, dissolved oxygen concentration, and other parameters),
- shipment of samples to a laboratory, and
- recordkeeping.

The use of state-approved sampling methods and certified laboratories is a requirement for regulatory compliance.

Remote monitoring and control is becoming a more practical option for small communities and could be part of an O&M agreement developed by a POU/POE vendor. Remotely monitoring operating parameters could provide an attractive alternative to fixed sampling and O&M schedules, thus, allowing customized customer service of the POU/POE device.

A POU/POE vendor could use remote telemetry equipment to review several installations and reduce the number of unnecessary system checks and visits. Remote telemetry also may be used to better schedule routine O&M, trouble-shooting problems, and emergency situations.

SELECTING AN APPROPRIATE POU/POE SYSTEM

The selection of POU/POE treatment units should be based on how well a particular type of unit removes specific contaminants from drinking water. Therefore, the selection process involves:

- an evaluation of the quality and type of source water,
- type and extent of contamination,
- treatment requirements, and
- waste disposal requirements.

The treatment requirements are usually compared based on operation and maintenance requirements, cost, and institutional requirements.

For a guarantee that a water treatment unit will perform and remove contaminants that the manufacturer claims, look for certification or registration labels on the treatment units. Two private organizations, the National Sanitation Foundation International (NSF) and the Water Quality Association (WQA), provide product testing. However, manufacturers are not required to test their products under these programs.

NSF INTERNATIONAL LISTING PROGRAM

NSF International is a nonprofit, independent testing and research group that provides standards for drinking water treatment components and tests home and personal water treatment devices to determine their ability to improve the aesthetic quality of water and remove health-related contaminants. NSF International has a certification laboratory that can conduct a full range of physical, microbiological, radiological, inorganic, and organic analyses.

NSF uses expert committees to develop its technology standards. The committee includes representation from industry, government, and consumer groups. The committee also receives input from a council of public health consultants and a certification council that has expertise in test methods.

Once an NSF committee develops a standard, the NSF applies to have it certified by the American National Standards Institute (ANSI). An ANSI designation means that only one standard exists for that type of product in the U.S. and that the standard follows all of ANSI's guidelines.

WQA Voluntary Product Validation Program and Voluntary Certification Program WQA, a nonprofit international trade association, awards its Gold Seal to water treatment equipment that passes testing under industry standards for performance, capacity, and durability. WQA does not evaluate claims for removing health-related contaminants. However, WQA has a certification program for POU/POE devices. But remember, WQA is a trade association for POU/POE equipment manufacturers, and although the association provides educational material to the consumer, they also promote the use of treatment equipment.

WHAT ABOUT SAFETY AND TERRORISM?

Recent events confirm that bioterrorism is no longer a threat, but a reality. POU filtration systems can reduce many biological warfare agents. POU systems, rated for their ability to remove waterborne biological agents, are available in a range of sizes from individual use to those capable of delivering thousands of liters of water per hour.

WHERE CAN I FIND MORE INFORMATION?

Lykins, B. W., R. M. Clark, and J. A. Goodrich. 1992. *Point-of-Use/Point-of-Entry for Drinking Water Treatment*. Boca Raton, FL: Lewis Publishers.

National Research Council. 1997. *Safe Water From Every Tap*. Washington, DC: National Academy Press

New England Water Works Association. 1998. *Small Systems Water Treatment Technologies: State-of-the-Art Workshop*. Joint Regional Operations Conference and Exhibition Proceedings, Marlborough, MA.

NSF International, World Health Organization, and Pan American Health Organization. 1999. *Providing Safe Drinking Water in Small Systems: Technology, Operations, and Economics*. Boca Raton, FL: Lewis Publishers.

Reynolds, K. 2001. "Point-of-Use Protection Against Bioterrorism." *Water Conditioning & Purification* 43:12.

U.S. Environmental Protection Agency. 1998. "Cost Evaluation of Small System Compliance Options: Point-of-Use and Point-of-Entry Treatment Units."

HAVE *YOU* READ ALL OUR *TECH BRIEFS?*

Tech Briefs, drinking water treatment and supply fact sheets, have been a regular feature in the National Drinking Water Clearinghouse (NDWC) publication *On Tap* for more than seven years.

A package of *Tech Briefs* is now available as a product. A three-ring binder holds all the current *Tech Briefs* in print. New selections can easily be added to the package as they become available. To order this free product, call the NDWC at (800) 624-8301 or (304) 293-4191 and ask for item #DWPKPE71.

You also may order online at ndwc_orders@mail.nesc.wvu.edu or download fact sheets from our Web site at www.ndwc.wvu.edu. (Additional copies of fact sheets are free; however, postal charges may be added.)

Z. Michael Lahlou holds a doctorate in environmental and natural resource economics and a master's in civil and environmental engineering. Formerly the technical assistance coordinator for the National Drinking Water Clearinghouse, Lahlou now resides in Huntington Beach, California.

ASSESSING THE BACTERICIDAL EFFICIENCY OF POLYDEX FOR THE DISINFECTION OF DRINKING WATER IN RURAL AREAS OF SOUTH AFRICA

M. N. B. Momba
N. Makala
A. Okeyo
University of Fort Hare
Alice, South Africa

INTRODUCTION

In developing countries, a large number of people lack access to adequate water supply. They, therefore, are forced to use surface water drawn from polluted rivers, irrigation canals, ponds, and lakes without prior treatment. For small rural communities, groundwater remains the main water supply source. Some communities receive their drinking water directly from uncovered or covered boreholes and wells, whereas for others, the water is drawn from the boreholes (using an engine) to a reservoir, and from the reservoir, it is then delivered to the people through a public standpipe system. In both cases, groundwater is distributed to the community without any purification. However, inadequate design, construction, operation, and maintenance of wells and boreholes can lead to quality problems related to groundwater. Poor sanitary sealing is more generally a common cause of microbiological quality deterioration (1). The failure to periodically drain and clean water supply holding tanks to remove sediments provides opportunities for heterotrophic bacterial colonization and biofilm development (2).

Polluted waters are an important vehicle for the spread of disease as they are carriers of disease-causing organisms such as pathogenic bacteria and viruses. Thousands of people in developing countries become ill each year through contact with contaminated water. The full impact of contaminated water on health is difficult to determine because of poor recording of diseases. A crucial need exists for scientifically sound answers to the problems of the contamination of water sources with potentially hazardous microbial organisms and the assessment of the risks posed by such pollution. Where public health is concerned, disinfection is the most important step in any process involving drinking water treatment.

Disinfection—by application of chlorine—is the most widely used treatment for bacteriological water quality improvement. However, its efficiency is limited to the presence of an adequate chlorine residual in potable water. Once the latter depletes, any drinking water system will experience a correlation between increase in bacterial numbers and the distance away from the point of treatment. This increase has been termed regrowth and is recognized as a major problem within many water systems (3). The identification of other disinfectants and the understanding of their modes of action should be stepped up in order to improve the microbial quality of drinking water from the point of treatment to the consumers.

Polydex is the result of extensive research and development in the search for a healthier microbial control and protection of potable water and wastewater. Polydex is a Pest Management Regulatory Agency (PCPA) registered product, which is designed for use in ponds, dugouts, and potable water tanks for the control of algae and bacteria (4). Polydex is also used in agriculture for the control of noxious gases and odors in animal waste and in slaughterhouses for reducing the risks of *E. coli* contamination, such as the one which claimed several lives in Walkerton, Ontario (5). The active ingredient in Polydex is a biologically active form of the copper ion (Cu^{2+}). Polydex contains a unique hydrotropic and colloidal mineral matrix, which effectively captures and carries copper ions (Cu^{2+}) evenly throughout water in high concentrations under a wide variety of water conditions.

The biocidal effects of copper have been used for centuries. The early Greeks and Romans made water storage and drinking vessels out of this metal. More recently, copper has been used in hospital, recreational, drinking, and industrial water systems. Unlike chlorine, copper does not result in dangerous halogenated organic byproducts, such as trihalomethane (THM), chloramines, and chloroform, and this ion is stable, making it easier to maintain an effective residual (6).

The disinfection action is attributed to the positive charge of copper ion. Positively charged copper ion has an affinity for electrons and, when introduced into the interior of a bacterial cell, they interfere with electron transport in cellular respiration systems. Metal ions will bind to the sulfhydryl, amino, and carboxyl groups of amino acids, thereby denaturing the proteins, which renders enzymes and other proteins ineffective, compromising the biochemical process they control. Cell surface proteins necessary for transport of materials across cell membranes also are inactivated as they are denatured. Finally, copper will bind with the phosphate groups that are part of the structural backbone of DNA molecules, which results in unraveling of the double helix and consequent destruction of the molecule (6,7).

It is also important to consider how copper in water can be advantageous to humans, animals, and plants. Our daily diet must provide specific trace amounts of copper for a number of reasons in order to maintain human health. The essential role of copper in maintaining normal health in both animal and humans has been recognized for many years. Normally, copper is readily available in a range of foods, and normal balanced diets should provide adequate daily amounts without the need for additional supplements. Dietary copper intake will vary considerably with the type of food consumed, the condition of the soils (e.g., copper content, pH, etc.) from which certain foods are produced, and drinking water characteristics. The average daily dietary requirement for copper in the adult human has been estimated at 2 mg and for infants and children at 0.05 mg/kg bw (8–10). The NRC (11) reported "estimated safe and adequate" daily dietary intakes of copper ranging from 0.5 to 0.7 mg/for infants 6 months of age or less and up to 2–3 mg/day for adults. In cases where individuals have a change in their everyday eating habits or decide to indulge in a limited medically controlled diet that may result in an inadequate intake of copper, water treated with Polydex may come in handy in supplying the body with needed copper.

Although several tests have shown the effectiveness of copper in killing bacteria, our preliminary experiments revealed that its effectiveness was limited in raw water with high turbidity when considering the initial dose recommended by the suppliers (1 part of polydex into 60,000 parts), which automatically showed that the quality of raw water remains one of the most important parameters in determining the efficiency of polydex. It was therefore vital to re-evaluate the bactericidal efficacy of polydex by considering various types of water sources and determining the disinfectant (polydex) demand of the water. Coliform bacteria were used as the main parameters with references to South African Water Quality Guidelines–Domestic use (12,13). Total coliforms are frequently used to assess the general hygienic quality of water and to evaluate the efficiency of drinking water treatment and the integrity of the distribution system. They should not be detectable in treated water. When found, they suggest inadequate treatment, post-treatment contamination and/or aftergrowth, and an excessive concentration of nutrients. In some instances, they may indicate the presence of pathogens responsible for the transmission of infectious diseases. The total coliform group includes bacteria of fecal origin and indicates the possible presence of bacterial pathogens such as *Salmonella spp., Shigell spp., Vibrio cholerae, Campylobacter jejuni, C. coli, Yersinia enterocolitica*, and pathogenic *Escherichia coli*, especially when detected in conjunction with other fecal coliform (13). The study aimed at generating valuable information in regard to the bactericidal efficiency of polydex for the disinfection of drinking water in rural communities.

MATERIALS AND METHODS

The experimental study was conducted on the basis of five replicates for each type of test water, and the effectiveness of polydex was evaluated every hour for the period of 7 h. Microbiological analyses were conducted in aseptic conditions under a laminar flow cabinet in the laboratory. Turbidity and pH were also considered, as both factors play a significant role in the disinfection process. All microbial and physicochemical tests were performed before and after disinfection.

Sampling Sites

Raw, filtered (water after flocculation, sedimentation, and filtration), and surface water samples were collected at the Alice water purification system (Alice, Eastern Cape), Umgeni water (Durban, Natal), Rand water, (Vereeniging, Gauteng), and in Thohoyandou water purification systems (Venda, Limpopo). Groundwater samples were also collected in Alice, Vereeniging, and Thohoyandou. Test waters were collected in clean sterile 5l polyethylene bottles.

Disinfection of Test Waters

The bactericidal effect of polydex was determined using two different initial doses:

1. Dose indicated by the manufacturer: 1 part of polydex into 60,000 parts of test water to give one part per million biologically active copper ions. Therefore, 0.083 ml polydex (approximately 3.9 mg/l copper) was used to disinfect 5l of test water with turbidities ranging between 0.59 NTU and 8 NTU.
2. 5.4 mg/l copper (corresponding to 2 × 0.083 ml) was used for the disinfection of surface water with higher turbidity (>9NTU).

Physicochemical Analyses

Turbidity and pH were measured using the microprocessor Turbidity Meter (HACH Co., Model 2100P) and pH Meter, respectively, and copper (Cu) concentrations were determined according to the spectroquant NOVA 60 manual (1998) using photometric test kits (Merck).

Microbiological Analyses

Total coliforms, presumptive *E. coli*, and *Salmonella* species were detected by the membrane filtration technique using filters with 0.45 μm pore size and 47 mm diameter (Millipore). Different volumes (10 and 100 ml) were filtered depending on the type of water used. Saline water was used as a diluent for the 10 ml volumes to spread the bacteria evenly over the filter membrane. The membrane filters were placed on Chromocult agar (Merck) plates and incubated for 24 h at 37 °C. Analyses were carried out in triplicates. Water samples were analyzed for the above micro-organisms using internationally accepted techniques (14).

RESULTS AND DISCUSSION

Characteristic of Raw Surface Water Before and After Disinfection with An Initial Dose of 3.9 mg/l Polydex

Figure 1 summarizes the counts of indicator bacteria before and after disinfection and the impact of turbidity and pH on the effectiveness of this disinfectant. There was a gradual decrease in all indicator micro-organisms after an hour of disinfection, which continued throughout the 7 h of the study period. Although presumptive *Salmonella* and presumptive *E.coli* were removed at a percentage of 100% within 1 h and 5 h, respectively, in all raw water samples, the removal of total coliforms in water samples from Venda, Rand water, and Alice ranged between 90–91% throughout the 7 h of the study period. However, complete removal of total coliforms occurred 5 h after disinfection of Umgeni water, which appeared to have lesser turbidity compared with other test waters. Although all raw water samples were disinfected with the same initial copper concentration of 3.9 mg/l, copper residual was more slowly depleted in Umgeni water than in the other raw water samples. The average copper residuals in Venda, Rand, and Alice waters after 7 h were 1.02 mg/l; 1.20 mg/l, and 1.12 mg/l, respectively, whereas samples from Umgeni water still maintained the copper residual concentration of 1.71 mg/l, which gave a clear indication that higher turbidities hinder the effectiveness of polydex as a disinfectant.

Characteristics of Filtered Surface Water Before and After Disinfection with an Initial Dose of 3.9 mg/l Polydex

The effectiveness of polydex was further proved by the dramatic decrease in numbers of all indicator micro-organisms in filtered water samples with reasonably low turbidities after disinfection (Fig. 2). Complete removal of presumptive *Salmonella,* presumptive *E. coli*, and total coliform bacteria occurred after 1 h, 2 h, and 4 h of disinfection in all the water samples, respectively. Although all filtered water samples maintained copper residuals ranging between 2.06 mg/l and 2.44 mg/l after the 7 h of the study period, the water samples from Umgeni water purification system had a higher residual (2.44 mg/l) than those from Venda, Rand, and Alice water supplies, which had residuals of 2.11 mg/l, 2.20 mg/l, and 2.06 mg/l, respectively. There was an indication of a slight decrease in pH after disinfection in all samples (raw surface water, raw groundwater, filtered surface water); however, this appeared not to have any effect on the effectiveness of polydex—or public health—as the pH values obtained were within the limits recommended for potable water, which are 5–9.5 (12).

Characteristic of Raw Groundwater Before and After Disinfection with An Initial Dose of 3.9 mg/l Polydex

Initial total coliform bacterial counts of 68 cfu/100 ml, 46 cfu/100 ml, and 82 cfu/100 ml were noted in Venda, Rand, and Alice groundwaters, respectively. These bacterial counts gradually decreased to 0 cfu/100 ml after 4 h in Venda and Rand water when the copper residuals were 2.43 mg/l and 2.96 mg/l, respectively. Although this was not the case with Alice water samples, the total coliform bacterial counts were found to be within the South African recommended limits for no risk (0–5 cfu/100 ml) (13). Although a complete removal of presumptive *E. coli* occurred 3 h after disinfection in all types of water, no presumptive *Salmonella* was recorded in Venda and Alice water samples within 1 h (Fig. 3).

Characteristic of Raw Groundwater Before and After Disinfection with An Initial Dose of 5.4 mg/l Polydex

There was a notable decrease in coliform bacterial counts when the initial dose of 5.4 mg/l polydex was used for

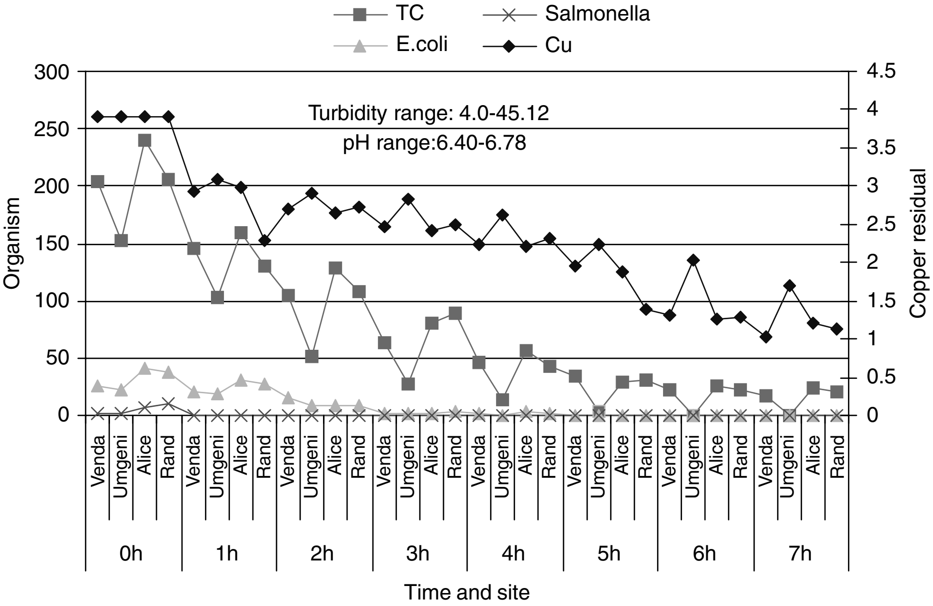

Figure 1. Average counts of indicator bacteria in raw surface water samples before and after disinfection with an initial dose of 3.9mg/l copper.

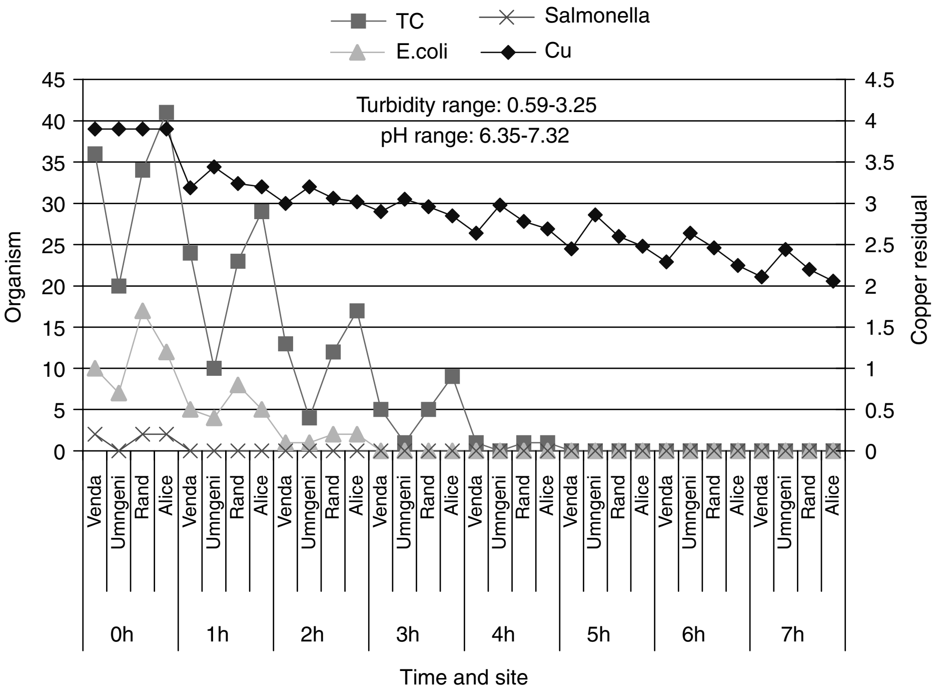

Figure 2. Average counts of indicator bacteria in filtered water samples before and after disinfection with an initial dose of 3.9 mg/l copper.

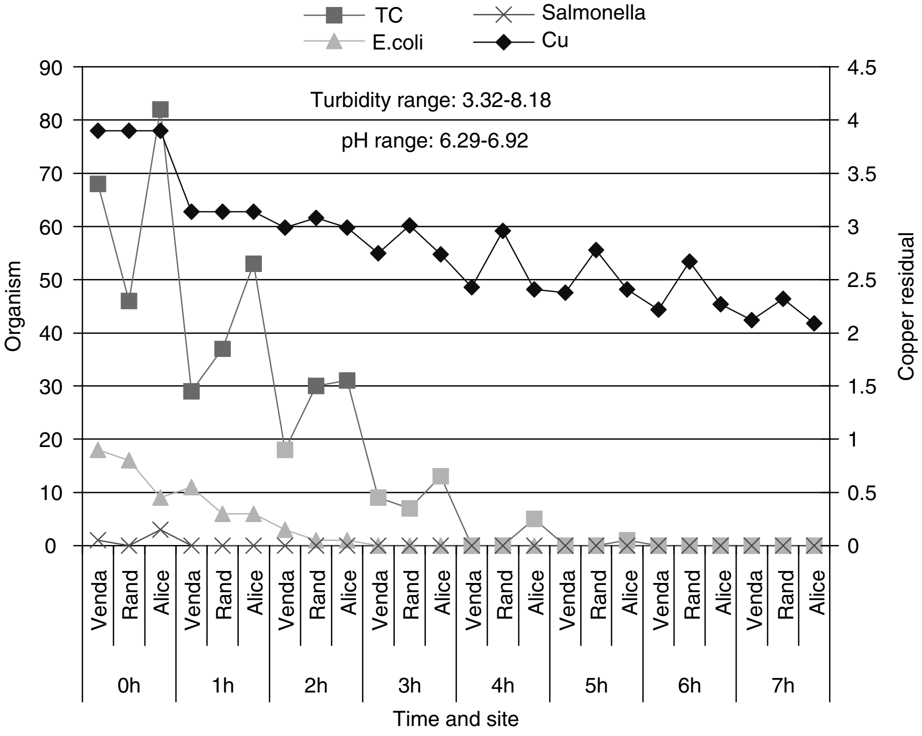

Figure 3. Average counts of indicator bacteria in groundwater samples before and after disinfection with an initial dose of 3.9mg/l copper.

the disinfection of water samples with higher turbidities. Complete removal of total coliforms occurred 6 h after disinfection of Venda and Alice water samples. Although the removal of bacteria was not complete in Rand water samples within this time, the number of total coliforms was within the limits of South African Water Quality Guidelines for no risk (0–5 cfu/100 ml) (12,13), after which the complete removal of coliforms occurred 7 h after disinfection (Fig. 4). The above observations further confirmed that polydex is highly effective in less turbid waters and, therefore, calls for an increase in dosage in highly turbid waters.

As indicated in Fig. 4, the total removal of presumptive *E. coli* in all test waters was recorded 3 h after disinfection.

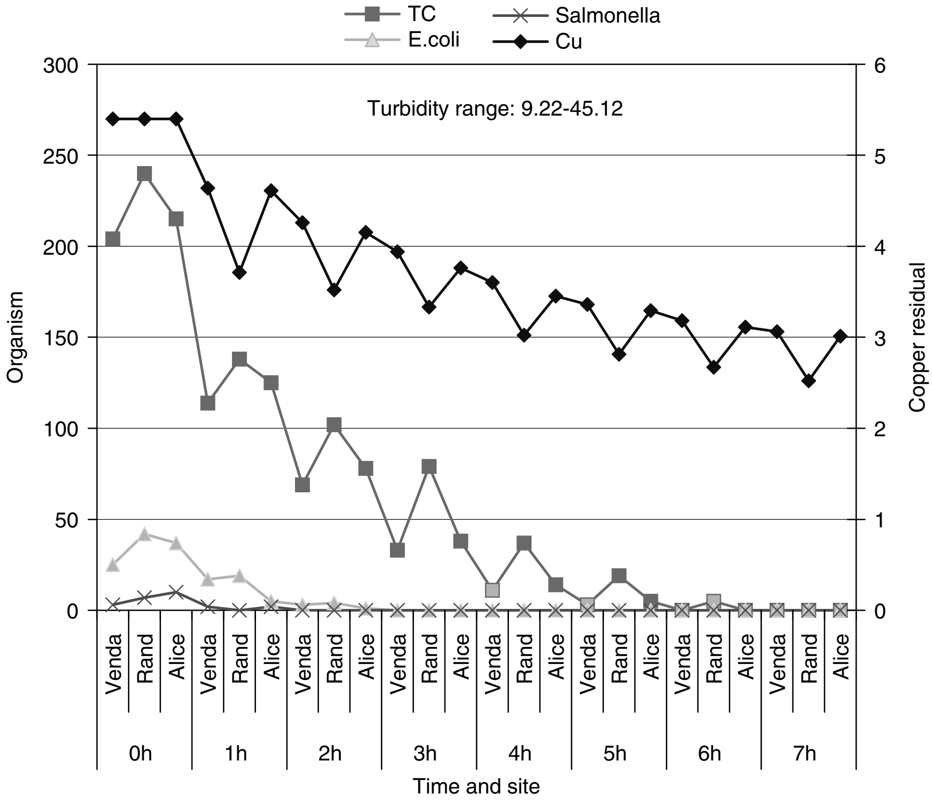

Figure 4. Average counts of indicator bacteria in water samples with high turbidity before and after disinfection with an initial copper dose of 5.4 mg/l.

In Rand water samples, the removal of presumptive S*almonella* was observed within 1 h after disinfection, and in Venda and Alice water samples, complete removal of this bacteria species occurred after 2 h.

CONCLUSIONS AND RECOMMENDATIONS

The study has revealed that polydex was effective for the removal of indicator micro-organisms. However, the turbidity of water remains one of the most important factors that can have a negative effect on its bactericidal efficiency. The dose of 3.9 mg/l appeared to be suitable for the disinfection of water samples with turbidity values ranging between 0.59 NTU and 8.18 NTU; however, this dose is not sufficient for the disinfection of highly turbid waters (>9 NTU). A dose of 5.4 mg/l is recommended for these waters as it completely removes the indicator micro-organisms thereafter leaving residuals of 2.52–3.06 mg/l after 7 h, which are within the recommended limits for potable water.

Based on the present investigations, it is important to filter highly turbid waters or increase the dose in order to increase the effectiveness of polydex. Consequently, this study suggests the followings:

- Polydex can be recommended as a disinfectant for drinking water as limits allowed by South African Water Quality guidelines are reached when this disinfectant is used: (0–5 counts/100 ml total Coliforms, 0 counts/100 ml Faecal Coliforms) (13).
- The dose of 3.9 mg/l is recommended for water with turbidity values ranging between 0.59 NTU and 8.18 NTU, and 5.4 mg/l is recommended for highly turbid waters (>9 NTU).
- The product is also recommended because the residual disinfectant ranges between limits allowed by South African Water Quality Standards (target water quality range 0–3 mg.l^{-1}).

Acknowledgment
The authors wish to express their thanks to Sibambene Trade & Services (Pty), Ltd, who sponsored this investigation. The effort of some personnel of the University of Venda, Umgeni Water, and Rand Water, especially for the collection of water samples, are immensely appreciated.

BIBLIOGRAPHY

1. Momba, M.N.B. and Notshe, T.L. (2003). The microbiological quality of groundwater-derived drinking water after long storage in household containers in a rural community of South Africa. *J. Wat. Suppl: Ers. Technol. AQUA*. **52**(1): 67–77.
2. Geldreich, E.E., Nash, D.H., Reasoner, D.J., and Taylor, R.H. (1972). The necessity of controlling bacterial populations in potable water: community water supply. *J. Am. Wat. Wks. Assoc.* **64**: 596–602.
3. Momba, M.N.B. (2000). Influence of Chloramination on Bacterial Regrowth in a chlorinated surface water laboratory scale system. *Adv. In Wat. Res.* **1**: 51–62.
4. Water Technology. (2004). *The Future of Drinking Water*. Available: http://www.watertechonline.com.

5. Mogollón, C.D. (2000). Eau Canada-Vancouver. WCP Online at: http://www.wcp.net/archive/sep00dealer.htm.
6. Meyer, W.C. (2001). Coping with resistance to copper/silver disinfection—Terms & Condition of use. *Wat Engin Manag.* **148**(11): 1–4.
7. Yayha, M., Landeen, K.L., Kutz, S.M., and Gerba, C.P. (1989). Swimming pool disinfection: an Evaluation of the efficiency of copper-silver ions. *J Environ Health*. **51**: 282–285.
8. Food Standard Committee. (1956). *Report on Copper*. Her Majesty's Stationary Office, London.
9. Browning, E. (1969). *Toxicity of Industrial Metals*. 2nd Edn. Buttersworth, London.
10. World Health Organization. (1974). *Toxicological Evaluation of some Food Additives Including Anticaking Agents, Antimicrobials, Antioxidants, Emulsifiers and Thickening Agents: Cupric sulphate*. WHO Food Additives Series No. 5.
11. NRC. (1980). *Recommended Dietary Allowances*. Food and Nutrition Board, National Research Council and National Academy of Science, Washington, DC.
12. Department of Water Affairs and Forestry, Department of Health and Water Research Commission. (1998). *Quality of Domestic Water Supplies*. Vol. 1: Assessment Guide. Water research Commission No: TT101/98. ISBN No: 1 86845 416 9.
13. Department of Water Affairs and Forestry (DWAF). (1996). *South African Water Quality Guidelines*. Vol. 1: Domestic Use, 2nd Edn. Pretoria.
14. Standard Methods for the Examination of Water and Wastewater. (1998).

PRIVATE SECTOR PARTICIPATION, MARKETING AND CORPORATE STRATEGIES IN MUNICIPAL WATER SUPPLY AND SEWERAGE

David Lloyd Owen
Envinsager
Llangoedmor, Ceredigion,
United Kingdom

INTRODUCTION—PRIVATE SECTOR PARTICIPATION: DEFINITION AND DRIVERS

The entry will describe how financial, commercial, and technical services are being mobilized to enable economic and social services to be provided and improved by water sector organizations in the North (the developed economies) and the South (developing economies). These needs are different because the policy contexts (politics and governance), standards, and risks experienced by entities installing and improving water services in the North and the South are different. It will be shown that the international water sector companies have evolved a diverse suite of financial and contractual arrangements often in association with international financial institutions (e.g., the World Bank). It will also be shown that the water sector is evolving as new companies shape their financial and technical competence to operate and compete internationally.

Private sector participation (PSP) refers to the formal provision of municipal water and sewerage services by the private sector. This involves providing these services commercially, where infrastructure work is funded by private capital or through a combination of finance from private investors and multilateral agencies.

There are two principal drivers for PSP. In the developed economies of North America, Western Europe, and Southeast Asia, PSP is primarily used to shift the financial burden of upgrading and extending municipal water and sewerage services from central and local government to the private sector. In developing economies, PSP is being employed to finance and manage the development and operation of the water and sewerage infrastructure on commercial lines to mobilize funding from a variety of sources.

The World Water Council's 'World Water Vision for 2025' seeks to address the current lack of access to water provision, sewerage, and sewage treatment by 2025 (1). To provide universal access to water and sewerage services, Vision 2025 calls for investment in new water and sanitation assets to rise from $30 billion in 1995 to $75 billion pa and for investment by industry and for environmental protection to increase from $10–15 billion in 1995 to $75 billion pa between 2000 and 2025 (2). Vision 2025 anticipates that national private companies will contribute 45% of this investment, against 15–21% in 1995; international finance will increase from 5–6% to 24% (3). Such a financial commitment will not take place unless adequate investment conditions exist, and these require private sector participation to manage these services (4).

HISTORY OF PSP

The first formal example of private sector participation in water provision—as opposed to municipal corporations and religious foundations—dates from 1582 when Peter Morris was granted a 500-year lease to pump water from a station underneath London Bridge (5). Early private sector water companies were concerned with providing water to individual buildings not directly served by watercourses and wells as urbanization and industrial development emerged. London's New River Water Company, incorporated in 1619 (6) was the earliest example of a corporate entity, surviving today as Thames Water, part of Germany's RWE. Typically, these companies were taken over as European municipalities developed and were given statutory obligations regarding water provision and sewerage. In the United Kingdom, a series of Acts of Parliament passed between 1867 and 1894 brought the majority of these entities under municipal control (7). Between 1850 and 1914, economic development and urbanization in Europe and the Americas brought about a new wave of water company formation (Table 1). In Europe and Latin America, these companies were usually set up to integrate water provision and distribution contracts for an entire city as opposed to localized undertakings in the United States. York Water Company, the earliest extant company in the United States was incorporated in Pennsylvania in 1816.

Economic and political considerations between 1914 and 1945 caused many European companies to be nationalized or concessions to be revoked. In the United States, the market ceased to expand, but there were a

Table 1. Early European Water Companies and Corporations (Current Owners in Brackets)[a]

Company	Country	Formed	Comments
New River Water Co (Thames Water/RWE)	U.K.	1619	Nationalized, privatized, acquired
Acque Potabili (Italgas)	Italy	1852	Private sector since foundation
Generale des Eaux (Veolia Environment)	France	1853	Private sector since foundation
Berlin Wasser Betriebe	Germany	1856	Privatized in 1999
Lyonnaise des Eaux (Suez Ondeo)	France	1880	Private sector since foundation
Aguas de Barcelona	Spain	1882	Private sector since foundation
Cia. Generale dell Acque	Italy	1884	Municipally owned since 1918
Gelsenwasser	Germany	1887	Acquired by E.ON in 2000
Aguas de Valencia	Spain	1890	Privatized in 1976

[a] Reference 8.

number of concession awards in Mexico. Except for France and Spain, where lease and concession awards were made from the 1950s, PSP continued to play a peripheral role, if at all, until the late 1980s. Between 1988 and 2002, there has been a fourfold increase in PSP and a reappraisal of its role in various economies.

FORMS OF PSP

PSP contracts can be gained outright through a bidding process, or they can evolve from contacts established through private sector consulting, construction, or engineering activities. The characteristics of the main types of water and wastewater privatization contracts are outlined in Tables 2 and 3.

These two types of contracts do not delegate full financial responsibility to the private operator, especially with regard to private capital investments. Operations and maintenance (O & M) contracts operate on a fixed fee basis and cannot address problems of municipal inefficiency.

The municipality controls the assets, and the private sector controls their operation. Risk elements start emerging because the private sector now deals directly with customers.

Table 2. Operations and Management (O & M) and Lease Contracts

Time horizon	2–5 years, up to 10	Ownership	Public
Customer	Government/ Municipality	Investment	Public
Cash flow profile	Fixed fee for service	Operation	Public
Construction risk	None	Tariff collection	Public/ Private
Regulatory risk	None		

Table 3. Lease Contract

Time horizon	10–15 years, up to 25	Ownership	Public
Customer	Retail customer	Investment	Public
Cash flow profile	Subject to market risk	Operation	Private
Construction risk	None	Tariff collection	Private
Regulatory risk	Medium		

Concessions

Concessions involve the private sector operation of assets to pay for new or upgraded facilities and upgrading work. Build-own-operate (BOO) and build-operate-transfer (BOT) contracts involve specific services to the municipality in relation to a specific program of capital improvements, whereas the full utility concession contract embraces all aspects of service provision and capital spending (Table 4). Concessions require a much more specific regulatory environment to account for the elements of risk involved. The concession company controls the service provision entity, but the municipality retains control of the asset owning entity. The latter entity is subsequently responsible for the extant assets, and new assets are vested into this entity at an agreed date. Dalton (9) discusses the political risk involved in developing concessions in developing economies.

A BOO/BOT project's cash flows are usually contractually predetermined and often have government backing. BOT/BOO projects are an effective means of rapidly organizing private capital and management toward a narrow range of services. However, some of the simpler project-oriented contracts do not affect the utility's management and operation; thus underlying problems such as leakage (and illegal interception), overstaffing, and tariff collection may not be addressed.

In full utility concessions (Table 5), existing revenues can be used immediately to service debt, thereby mitigating construction risk. During a period of time, a utility can benefit from a steady flow of revenues from a diversified customer base. A more robust balance sheet can be created, allowing for internal finance as well as the use of capital markets to sell long term debt. The operator is responsible for upgrading and operating

Table 4. BOOT/BOT/BOO Concession

Time horizon	10–30 years, up to 95	Ownership	Public
Customer	Govt./Municipal	Investment	Private
Cash flow profile	Pay on completion	Operation	Private
Construction risk	High	Tariff collection	Public
Regulatory risk	Low		

Table 5. Full Utility Concession

Time horizon	20–30 years	Ownership	Public
Customer	Retail Customer	Investment	Private
Cash flow profile	Subject to market risk	Operation	Private
Construction risk	Low	Tariff collection	Private
Regulatory risk	High if politics volatile		

Table 6. Asset Sale/Asset Ownership

Time Horizon	In Perpetuity	Ownership	Private
Customer	Retail Customer	Investment	Private
Cash flow profile	Subject to market risk	Operation	Private
Construction risk	Very low	Tariff collection	Private
Regulatory risk	Very high		

the services, while developing new assets to hand over to the municipalities in the longer term.

Asset sale is the most dramatic and politically contentious form of privatization (Table 6). To date, it has been used in the 1989 sale of the English and Welsh water and sewage companies (WASCs) and in Chile. The assets are in private hands, but the license to operate them can be subject to renewal. In the case of the U.K. WASCs, a 30-year operating license was awarded to each entity in 1989. In the United States, companies developed the assets in the first place (Table 7).

THE ‘BRITISH’ AND ‘FRENCH’ MODELS

The World Bank calls delegated water management through concession awards the ‘French model.’ The ‘French model’ is typically used to contrast it with the ‘British model’ of asset sales. The real ‘French model’ is the Affermage lease as traditionally used in private sector contracts in France.

Initial Public Offerings

Initial public offerings (IPOs) of a corporatized utility take place when all or part of the shares of a water or multiutility company are listed on the local stock exchange. Since 1976, 48 municipal entities have been privatized by this method; all but nine took place since 1989. Twelve involved outright share sales, but municipalities retain a majority holding in the other 36. In addition, seven privately held water utilities have had IPOs since 1991, along with the water activities of two larger conglomerates. Prime Utilities (Malaysia, 1994) has subsequently left the sector; Azurix was bought back by Enron, its majority holder in 2001, and its main asset, Wessex Water, was in turn acquired by YTL of Malaysia in 2002. One, Northumbrian Water of the United Kingdom was privatized in 1989, acquired by Suez in 1996, and refloated in 2003.

EXTENT OF PSP

In 1988, PSP was restricted to its ‘traditional’ markets in the United States (asset owning companies); England and Wales (statutory water companies); and France, Italy, and Spain (concessions and lease contracts). The World Bank noted eight PSP water and sewerage contracts in developing economies between 1984 and 1989 (10) against 97 between 1990 and 1997 (Table 8) (11).

The total number served by the private sector in 1998 is estimated at 93 million. Since 1988, it is estimated that PSP has reached a further 302 million people through privatizations; 64 million were added through service extension and population growth within contracts, along with minor contract awards, to cover a total of 460 million people, or 8% of the global population. This includes 7.2 million people in six contracts, which subsequently have been returned to municipal ownership (Table 9).

Table 10 is based on 97 water and sewerage privatization awards identified by the World Bank during the first 8 years of the 1990s involving a total investment of $24.95 billion (12) from banks and multilateral agencies.

O&M contracts do not mobilize new sources of private sector investment. Greenfield operations are typically site specific, involving the construction of a water or sewage treatment facility, as seen in the Scottish sewage treatment construction in the United Kingdom. In recent years, a number of greenfield contracts have been awarded in areas earmarked to become new housing or industrial zones. This approach has had some popularity in the Philippines. Divestitures have been seen in Chile. The concession approach, allied with the splitting of water and sewerage entities into operating and asset holding companies has become the favored approach toward water privatization in many countries (Table 11).

CORPORATE STRATEGIES

A variety of relationships have developed between companies and within companies with regard to global

Table 7. Examples of PSP Contracts by Size and Type

	O&M	BOT	Full Concession	Asset Ownership
Local/site	USA	France	Scottish PFI	USA
Town	Kazakhstan	Germany	Germany	Czech Republic
City	Mexico City	Budapest	Manila	Chile
Region	Greater Amman	Czech Regions	Argentina	UK WASCs
Country	Chad (Phase 1)	Ghana (urban)	Chad (Phase 2)	—

Table 8. Water or Wastewater Services Privatized Each Year, 1988–2002 (Million People)

Year	Running Total	Major Awards	Organic Growth	Minor Gains	Contract Losses	Year End Total	Y-O-Y Increase
1988	93	0.6	1	0	0	95	2
1989	95	45.6	1	1	0	142	48
1990	142	4.6	1	0	0	148	6
1991	148	0.0	1	0	0	149	1
1992	149	0.8	1	0	0	151	2
1993	151	32.3	2	1	0	186	35
1994	186	3.6	2	0	0	192	6
1995	192	16.4	3	1	0	212	20
1996	212	22.6	3	1	0	239	27
1997	239	42.1	3	2	0	286	47
1998	286	16.0	3	2	−1.5	305	20
1999	305	36.1	4	6	−3.6	348	43
2000	348	35.3	5	3	−0.6	390	43
2001	390	32.6	5	4	0	432	42
2002	432	22.3	5	3	−1.5	460	30
Total		310.9	40.0	24.0	−7.2		369

Table 9. People Served Through Private Sector Participation (Start of 2003)

	PSP	Population[a]	% PSP
Western Europe	161.5	388	42
C&E Europe	14.1	341	4
ME & Africa	37.6	1,041	4
Central & South Asia	1.6	1,407	0
South East Asia & Oceania	95.7	2,049	5
North America	65.2	310	21
Latin America	84.0	519	16
World total	459.7	6,055	8

[a] Million people.

Table 10. Private Participation in Water and Sewerage in Developing Countries by Contract Type, 1990–1997

% of Total	Projects	Total Investment
Concessions	50%	80%
Divestiture	6%	4%
Greenfield projects	31%	16%
Operations and management and leases	13%	0%

water and sewerage contracts. None of the water company linkages existed before 1992; only Veolia (then called Generale des Eaux) and Suez (then called Lyonnaise des Eaux) were operating as private sector multiutility companies at that time

Power and Gas Utilities Entering the Sector

The attraction of the water market for these companies is that it represents an extension of their abilities in client management, asset maintenance, and development and in dealing with regulation and politics. Many of these companies have been publicly owned until relatively recently. Power companies across Europe have entered the sector either by acquiring companies in their home markets or internationally.

Table 11. Investment in Water and Sanitation with PSP in Developing Economies ($million)[a]

Region	1990–1994	1995–1999
East Asia & Pacific	4,023	8,631
Europe & Central Asia	16	1,539
Latin America & Caribbean	4,732	8,965
Middle East & North Africa	0	4,106
South Asia	0	0
Sub-Saharan Africa	23	1,054

[a] Reference 13.

Municipal Multiutilities

To date, one German and six Italian municipal multiutilities have been partially floated. In these cases, the water activities account for a minority of group turnover, which tends to be dominated by power activities.

Multiutilities active in seeking water contracts or investments include United Utilities (UK: North West Water/NORWEB), VE (France: Veolia Water/Cofreth & Esys Montenay), Suez (France: Ondeo/Tractabel & Elyo), RWE (Germany: Thames Water/RWE), ACEA (Italy: Acqua Italia/ACEA), and Amga (Italy: Acque Genova/Amga).

From $4.8 billion in investment in 1990–2004 against $24.3 billion in 1995–1999, it is evident that the pace of investment increased in the latter half of the decade, although only $5.1 billion was invested in 1998–1999, as a result of economic problems in Asia during those 2 years. Typically, equity investment provided by the companies involved in operating concessions is equivalent to 30% of the total investment, indicating that the private sector invested $12.2 billion during the period 1990–1999, a total investment of $41.3 billion.

CONSTRUCTION AND WATER

Constructing water and wastewater facilities and their operation have always been closely interlinked. In Vivendi Environment and Suez, the water companies have branched into construction, as has been more erratically seen in the English and Welsh water companies during the past 13 years. Usually, the construction company is attracted to the services concessions, having built the facilities. Bouygues, one of the leading French construction companies, acquired Société d'Aménagement Urbain et Rural (SAUR, founded in 1933) in 1984. FCC (originally Focsa) was founded as a construction company in 1900 and gained the Barcelona sewerage contract in 1911. Dragados, a Spanish construction company, has extensive experience in water supply and sewerage projects. Dycagua, Urbaser's water tendering arm, was set up in October 1994 and has won a number of concessions. Ferrovial has adopted a similar approach through its Ferroser subsidiary.

JOINT VENTURES

Joint ventures for international water/sewerage contracts typically include a company that has experience in the water sector and has a significant presence in the target market(s). Examples include International Water (Edison/Bechtel/United Utilities—all non-U.S. markets), Sino-French Holdings (Suez/New World Developments - China, HK & Macao), VE/Hyundai (South Korea), and Proactiva (VE/FCC-Latin America) (Table 12).

It is evident that the two pioneering companies, Suez and Veolia Environment, continue to lead this market. That dominance is now being challenged, especially by Germany's RWE and a number of British companies.

BIBLIOGRAPHY

Except where a citation has been employed, data and their interpretation are based on the author's private databases and information published by the companies cited.

1. Cosgrove, W.J. and Rijsberman, F.R. (2000). *World Water Vision*. Earthscan, London, p. 53.
2. Cosgrove, W.J. and Rijsberman, F.R. (2000). *World Water Vision*. Earthscan, London, p. 60.
3. Cosgrove, W.J. and Rijsberman, F.R. (2000). *World Water Vision*. Earthscan, London, p. 64.
4. Cosgrove, W.J. and Rijsberman, F.R. (2000). *World Water Vision*. Earthscan, London, p. 61.
5. Barty-King, H. (1992). *Water, The Book*. Quiller Press, London, pp. 45–46.
6. Barty-King, H. (1992). *Water, The Book*. Quiller Press, London, p. 51.
7. Barty-King, H. (1992). *Water, The Book*. Quiller Press, London, p. 138.
8. Owen, D.A.L. (1998). The European Water Industry, A country-by-country analysis. *Financial Times Energy*. London, p. 46.
9. Dalton, G. (2001). *Private Sector Finance for Water Sector Infrastructure: What Does Cochabamba Tell Us about Using this Instrument?* Occasional Paper No 37, Water Issues Study Group, School of Oriental and African Studies (SOAS), University of London.
10. Silva, G., Tynan, N., and Yilmaz, Y. (1998). *Private Participation in the Water and Sewerage Sector—Recent Trends*. Public policy for the private sector note 147, World Bank, Washington, DC, p. 1.
11. Silva, G., Tynan, N., and Yilmaz, Y. (1998). *Private Participation in the Water and Sewerage Sector—Recent Trends*. Public policy for the private sector note 147, World Bank, Washington, DC, p. 2.
12. Silva, G., Tynan, N., and Yilmaz, Y. (1998). *Private Participation in the Water and Sewerage Sector—Recent Trends*. Public policy for the private sector note 147, World Bank, Washington, DC, p. 5.
13. World Bank. (2001). *World Development Indicators 2001*. The World Bank, Washington, DC, p. 272.

Table 12. The Leading International Players (by People Served)

Company (Home Country)	Home	International	Total	% Home
Suez—Ondeo (France)	17,000,000	114,100,000	131,100,000	13
Veolia Environment (France)	26,000,000	87,800,000	113,800,000	23
RWE (U.K. & Germany)	27,600,000	41,500,000	69,100,000	40
Bouygues (France)	6,000,000	24,500,000	30,500,000	20
SABESP (Brazil)	19,100,000	0	19,100,000	100
United Utilities (U.K.)	10,300,000	10,200,000	20,500,000	50
Awg (U.K.)	5,800,000	10,600,000	16,200,000	36
Severn Trent (U.K.)	8,300,000	6,300,000	14,600,000	57
FCC (Spain)	6,100,000	5,900,000	12,000,000	51
Bechtel (U.S.)	0	10,200,000	10,200,000	0

PUMPS

VIPIN BHARDWAJ
NDWC Engineering Scientist

BETHANY REED
NESC Graphic Designer

From start to finish, pumps play an important role in all water distribution systems. They are used to transfer raw water to the treatment plant; they supply water to sections of a distribution system where it is not possible to supply it by gravity; they add appropriate chemical doses during treatment; and they transfer sludge from settling chambers for further treatment and disposal. Pumps will provide long-term efficient service if they are properly operated and maintained

PUMPS IN WATER DISTRIBUTION SYSTEMS

Pumps are an important part of any water distribution system. At the source, they help deliver raw water to the treatment plant. In the plant, pumps add chemical solutions at desired dosages for treatment processes, and other pumps remove sludge from sedimentation facilities for further treatment and disposal. Within distribution systems, pumps discharge water under pressure to the pipe network and lift water where it cannot go by gravity, especially to water towers.

PUMP APPLICATIONS

Different types of applications require different types of pumps. Pumps are selected based on system requirements, discharge pressure required, flow capacity required, and availability of space. The two most common pumps in the water industry are: centrifugal pumps, used most often in water distribution, and positive-displacement pumps, most commonly used in treatment plants for chemical dosing.

TYPES OF PUMPS USED IN WATER SUPPLY

Three kinds of pumps are most often found in water distribution systems: the aforementioned centrifugal pumps, used to move water, vertical turbine pumps, used at the intake, and submersible pumps, used in wells.

Centrifugal Pumps

Centrifugal pumps are the most common type used in water distribution. Centrifugal pumps have a circular "fan/turbine-shaped" structure called an impeller that is mounted on a centrally supporting structure called the shaft. The motor rotates the shaft and can be powered by electricity or diesel fuel. Water enters at an opening in the center called the suction. The rotating impeller imparts a high velocity to the water, and it is circulated and thrown outward (See Fig. 1). A circular-shaped covering called casing surrounds the impeller. The casing is shaped like a spiral so that the water slows down, and the velocity head is converted to pressure head as it flows out of the casing. At the junction of the casing and shaft, a seal or packing rings are provided to prevent leakage. A packing gland presses against the rings to maintain a tight seal.

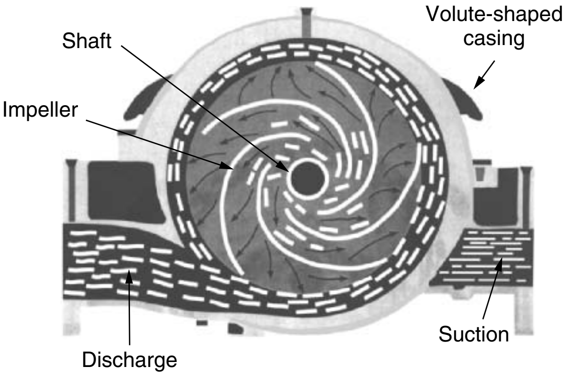

Figure 1. Volute centrifugal pump schematic. Adapted from AWWA *Water Transmission and Distribution: Principles and Practices of Water Supply Operations,* 2nd Edition.

GLOSSARY OF PUMP TERMS

Head—(1) A measure of the energy possessed by water at a given location in the water system expressed in feet; (2) a measure of the pressure or force exerted by water expressed in feet.

Velocity Head— A measurement of the amount of energy in water due to its velocity or motion.

Pressure Head—A measurement of the amount of energy in water due to water pressure.

Impeller—The moving element in a pump that drives the fluid.

Efficiency—A ratio of total energy output to the total energy input expressed as a percent.

Priming—The action of starting the flow in a pump or siphon. With a centrifugal pump, this involves filling the pump casing and suction pipe with water.

Source: American Water Works Association. 1996. *Water Transmission and Distribution: Principles and Practices of Water Supply Operations*, 2nd Edition. Denver, CO: AWWA.

Different flow rates and outlet pressures can be generated by this type of pump, depending on the size of the pump and the space between the casing and the impeller. Impellers can also have varying shapes. These pumps can be used either as a single unit or can be coupled with another similar pump in series to generate more output pressure. When more than one pump is used, it becomes a multistage pump. The pressure desired in the distribution system determines the number of stages required. If a single stage pump is not enough to supply water at adequate pressure, multiple stages are used. Centrifugal pumps can achieve an output pressure of up to 250 feet for each stage. Centrifugal pumps are used to supply large volumes of water at a constant output pressure. The pressure generated by a centrifugal pump is less compared to other types of pumps, such as a reciprocating (positive-displacement) pump.

According to the American Water Works Association's *Water Transmission and Distribution: Principles and Practices of Water Supply Operations*, some of the advantages of centrifugal pumps are:

- wide range of capacities that range from a few gallons per minute [gpm] to 50,000 gpm. Heads of five to 700 feet are generally available;
- uniform flow at constant speed and head;
- simple construction (small amounts of suspended matter in the water will not jam the pump);
- low to moderate initial cost for a given size;
- ability to adapt to several drive types—motor, engine, or turbine;
- moderate to high efficiency at optimal operation;
- no need for internal lubrication;
- little space required for a given capacity;
- relatively low noise level; and
- ability to operate against a closed discharge valve for short periods without damage.

Some of the disadvantages are:

- an efficiency that is limited to a narrow range of discharge flows and heads;
- low capacity that is greatly dependent on discharge pressure;
- generally no self-priming ability;
- potential for running backward if stopped with the discharge valve open; and
- potential for impeller to be damaged by abrasive matter in water, or clogged by large quantities of particulate matter.

During startup, centrifugal pumps require a procedure called “priming” in which the pump is filled with water before turning the switch on. This insures that when the impeller starts rotating it starts pumping water, and energy imparted to the impeller is not lost. If a centrifugal pump is not primed, it does not operate efficiently (i.e., it does not pump water even when it is turned on). Generally, pumps have an adjacent chamber, called the priming chamber, that sucks in water when the pump is turned on and keeps the impeller submerged.

Vertical Turbine Pumps

In vertical turbine pumps, the water flows vertically through a channel of uniform cross-sectional area. The impeller is positioned in the center along the axis in the channel. The blades of the impeller are shaped so that the water flows in a radial direction (See Fig. 2). The casing has diffuser vanes, shaped so that they guide water and make it flow either into the discharge or through diffuser bowls into inlets of succeeding stages. Vertical turbine pumps are used most often at raw water intakes and at booster stations in the distribution system to augment the pressure required for service. Vertical turbine pumps are relatively more expensive than centrifugal pumps for the same capacity and require more maintenance.

According to the American Water Works Association’s *Water Transmission and Distribution: Principles and Practices of Water Supply Operations,* some of the advantages of vertical turbine pumps are:

- uniform flow at constant speed and head;
- simple construction;
- individual stages capable of being connected in series, thereby increasing the head capacity of the pump;
- adaptability to several drive types—motor, engine, or turbine;
- moderate to high efficiency under the proper head conditions;
- little space occupied for a given capacity; and
- low noise level.

The main disadvantages are:

- high initial cost;
- high repair costs;
- the need to lubricate support bearings located within the casing;
- inability to pump water containing any suspended matter; and
- an efficiency that is limited to a very narrow range of discharge flow and head conditions.

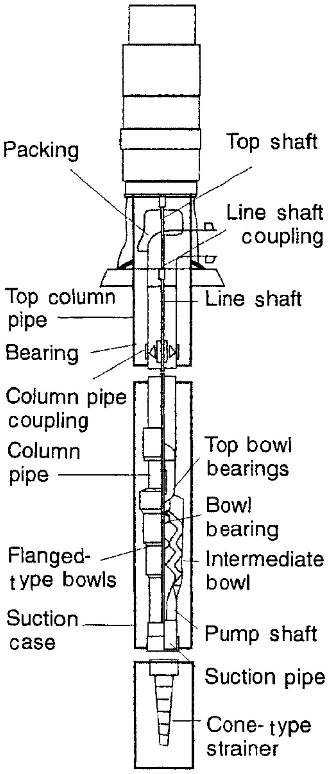

Figure 2. Verticle turbine pump. Adapted from AWWA *Water Transmission and Distribution: Principles and Practices of Water Supply Operations,* 2nd Edition.

Submersible Pumps

Submersible pumps, as the name implies, are placed below the water level (See Fig. 3). They are used mostly for pumping groundwater from wells. The pump is basically a multi-stage centrifugal pump. The impellers of the pump are mounted on a vertical shaft. Each impeller passes water to the next stage through a diffuser shaped so that the water flows vertically to the next diffuser. Each impeller and diffuser is called a stage. The desired output pressure determines the number of stages.

The pump is driven by an electric motor placed adjacent to the pump and is constructed for submerged operation. The wiring of the electric motor is such that it is waterproof. The motors are manufactured to ensure water does not cause it to fail by grounding or shorting out. However, most failures of submersible pumps are due to electrical problems with the motors.

Positive-Displacement Pumps. Positive-displacement pumps are used in water supply operations for feeding chemicals at various stages of the treatment process. They displace a certain volume of water in each stroke

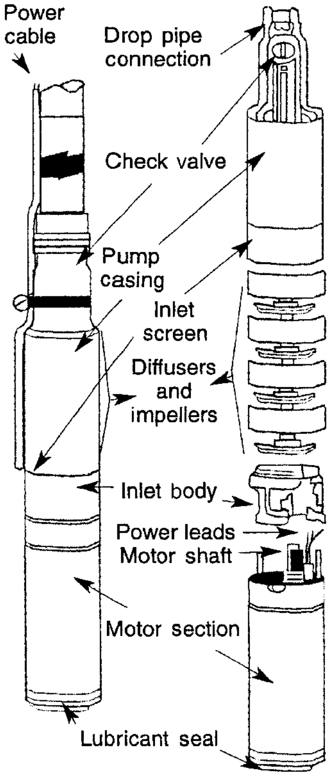

Figure 3. Submersible pumps. Adapted from AWWA *Water Transmission and Distribution: Principles and Practices of Water Supply Operations,* 2nd Edition.

as they operate. These pumps are not suitable for pumping large volumes of water, they are more suited for high pressure and low flow service. There are two types of positive-displacement pumps: reciprocating pumps and rotary pumps.

Reciprocating Pumps

Reciprocating pumps have a piston that moves back and forth in an enclosing cylinder. The arrangement of rotating shaft, connecting rod, and joint pushes the volume of the fluid in the cylinder through an outlet valve. The liquid enters and leaves the cylinder through check valves. Reciprocating pumps are suitable for applications where very high pressures are required, or where abrasive fluids, such as acids or viscous liquids, have to be pumped.

Rotary Pumps

In rotary pumps, the impeller rotates within an enclosing structure and imparts energy to the water. The impeller can be in the shape of a gear, screw, or lobes. For rotary pumps, the output is continuous and smoother compared to reciprocating pumps. Rotary pumps are better suited for lower pressures than reciprocating pumps because of slippage at higher pressures. Rotary pumps can jam if any solids enter the small clearance spaces between the impeller and casing. Rotary pumps are suitable for handling fluids of varying viscosities and are used in water treatment for chemical dosing.

OPERATION OF CENTRIFUGAL PUMPS

Because centrifugal pumps are so widely used in water distribution and because so many brands are available, the individual procedures for proper operation vary. Most manufacturers supply instructions, guidelines, or manuals with a new pump. Users must follow these instructions while operating pumps.

Pump Starting and Stopping

During pump startup, check the lubrication of the motor bearings and prime the pump. After priming, the pump must be started with the discharge valve fully closed. The discharge valve must then be opened slowly to allow any air in the system to escape. The pump discharge valve must be opened and closed slowly to prevent water hammer in the system. (For more information about water hammer, see the *Tech Brief* in the Winter 2003 *On Tap*.)

During power failure (when the motor stops running), the discharge valve can remain open. This might be a problem because the water can actually flow backward. On large pumps, emergency power operated valves are provided to prevent the impellers from running in reverse. On smaller pumps, check valves are provided.

If the pump will be shut down for longer periods, additional steps should be taken. In cold weather, for example, all water must be drained out from the pump to prevent freezing. The pump packing or seal must be removed if the pump is to be stored idle for longer periods.

Centrifugal Pump Maintenance

Centrifugal pumps require regular inspection and maintenance. Bearings on the motor may become worn and must be checked and kept well lubricated. The packing or seal can wear out due to friction resulting in leakage from the pump. These should be checked regularly. Bearing and motor temperatures must be monitored. If a surface is substantially hotter than normal, the unit should be shut down and the cause of excessive heat investigated. In case of any unusual noises or vibrations from the pump, it must be stopped and the cause investigated.

WHAT ABOUT SAFETY?

Safety instructions provided by the manufacturer must be followed. Pumps should always be switched off before any maintenance. All liquids must be drained out before servicing. People must stand away from any moving parts. Safety precautions must be taken regarding electrical motors. Maintenance personnel must wear rubber gloves and use insulated tools while servicing electrical motors. A record of all maintenance work must be maintained.

Concern that the water supply could be vulnerable to terrorist attacks has understandably been heightened by recent world events. Knowledge, planning, and preparedness helps relieve this concern. It is crucial now to regularly inspect the location, accessibility, and operation of all the pumps in the distribution system. A large-scale disruption in water supply for a large water system can happen because of pump failure at a key location. After gathering complete and accurate information on all the pumps in the system, the utility manager can run through "what if" scenarios to determine where a contaminant is likely to move. Workers can also learn how water movement can be controlled by running or shutting down pumps or other utility operations.

HAVE YOU READ ALL OUR TECH BRIEFS?

Tech Briefs, drinking water treatment and supply fact sheets, have been a regular feature in the National Drinking Water Clearinghouse (NDWC) publication *On Tap* for more than seven years.

A package of *Tech Briefs* is now available as a product. A three-ring binder holds all the current *Tech Briefs* in print. New selections can easily be added to the package as they become available. To order this free product, call the NDWC at (800) 624–8301 or (304) 293–4191 and ask for item #DWPKPE71.

You also may order online at *ndwc_orders@mail.nesc.wvu.edu* or download fact sheets from our Web site at www.ndwc.wvu.edu. (Additional copies of fact sheets are free; however, postal charges may be added.)

WHERE CAN I FIND MORE INFORMATION?

American Water Works Association. 1996. *Water Transmission and Distribution: Principles and Practices of Water Supply Operations,* 2nd Edition. Denver, CO: AWWA.

American Water Works Association. 1999. *Design and Construction of Water Systems: An AWWA Small System Resource Book*, 2nd Edition. Denver, CO: AWWA.

Dickenson, T.C. 1995. Pumping Manual. Oxford, UK: Elsevier Advanced Technology.

HDR Engineering, Inc. 2001. *Handbook of Public Water Systems*. New York: John Wiley and Sons, Inc.

For further information, to comment on this *Tech Brief*, or to suggest topics, contact Vipin Bhardwaj via e-mail at vbhardw2@wvu.edu.

NDWC Engineering Scientist, **Vipin Bhardwaj**, has master's degrees in environmental engineering and agriculture from West Virginia University.

RADIONUCLIDES

National Drinking Water
Clearinghouse

Radionuclide contamination of drinking water is a significant, emerging issue. Until now, manmade radioactivity in drinking water has not been a major problem. Natural sources have been the primary cause of contamination. However, the potential for contamination exists throughout the country as releases from medical facilities or nuclear power plants may wind up in drinking water. Because of their potential health effects and widespread occurrence, natural radionuclides—including radon, radium, and uranium—cause much concern.

WHERE DO RADIONUCLIDES OCCUR, AND WHAT ARE THE PUBLIC HEALTH RISKS?

Radionuclides occur naturally as trace elements in rocks and soils as a consequence of the "radioactive decay" of uranium-238 (U-238) and thorium-232 (Th-232). This decay happens because radioactive atoms have too much energy. When radioactive atoms release or transfer their extra energy, it is called decay. The energy they release is called ionizing radiation, which may be alpha particles, beta particles, or gamma rays. This energy is transmitted through space or another medium in waves (e.g., x-rays or gamma rays) or particles (e.g., electrons or neutrons) and is capable of either directly or indirectly removing electrons from atoms, thereby creating ions, which are electrically charged atoms.

Radon-222, radium-226, radium-228, uranium-238, and uranium-234 are ions of the U-238 and Th-232 decay series. They are the most common radionuclides found in groundwater. Other naturally occurring radionuclides tend to be environmentally immobile or have short half-lives, meaning they are far less likely to be found in significant amounts in groundwater.

When ionizing radiation strikes a living organism's cells, it may injure the organism's cells. If radiation affects a significant number of cells, the organism may eventually develop cancer. Or at extremely high doses, it may even die.

Radon. Radon is a naturally occurring radioactive gas that emits ionizing radiation. National and international scientific organizations have concluded that radon causes lung cancer in humans. Ingesting drinking water that contains radon also presents a risk of internal organ cancers, primarily stomach cancer. The U.S. Environmental Protection Agency (EPA) and the U.S. Surgeon General recommend testing indoor air for radon in all homes and apartments located below the third floor. If you smoke and your home has high indoor radon levels, your risk of lung cancer is especially high.

Tap water only emits approximately 1 to 2 percent of the radon found in indoor air. However, breathing radon from this source increases the risk of lung cancer over the course of a lifetime.

Radium. Radium-226 and radium-228 are natural groundwater contaminants that usually occur at trace levels. At high exposure levels, radium-226 and radium-228 can cause bone cancer in humans and are believed to cause stomach, lung, and other cancers as well.

Uranium. Uranium is a naturally occurring radioactive contaminant that is found in both groundwater and surface water. At high exposure levels, uranium is believed to cause bone cancer and other cancers in humans. EPA also believes that uranium can be toxic to the kidneys.

Table 1. What Regulations Govern Radionuclides?

Radionuclide National Primary Drinking Water Regulations						
		EPA Standards				
Contaminants	Drinking Water Health Effects	Proposed Maximum Contaminant Level Goal MCLG	Proposed Maximum Contaminant Level (MCL)	Current MCL	Sources	Best Available Technology (BAT)[2]
Radium-226	Cancer	Zero	20 pCi/L	5 pCi/L combined with radium-228	Naturally occurring	Ion Exchange (IE); Lime Softening (LS); Reverse Osmosis (R0)
Radium-228	Cancer	Zero	20 pCi/L	5 pCi/L combined with radium-226	Naturally occurring	IE; LS; RO
Radon-222	Cancer	Zero	300 pCi/L	–	Naturally occurring	Aeration
Uranium	Kidney toxicity, Cancer	Zero	20 ug/L[3]	–	Naturally occurring	Coagulation/ Filtration; Anion exchange; LS; RO[4]
Adjusted gross alpha emitters	Cancer	Zero	15 pCi/L	15 pCi/L	Naturally occurring and manmade	RO
Gross beta and photon emitters	Cancer	Zero	4 mrem ede/yr	4 mrem ede/yr any organ or whole body	Naturally occurring	IE; RO

1. *PicoCuries per liter* (pCi/L) is an activity measurement of radioactive decay (1 pCi/L = 2.2 disintegrations per minute); *micrograms per liter* (μg/L) is a mass measurement; *mrem* is measurement of effective radiation dose to organs.
2. Except as noted, BAT for the purpose of issuing variances is the same as BAT for compliance.
3. 20 μg/L is based on kidney toxicity. 20 μg/L is the equivalent of 30 pCi/L.
4. *Coagulation/Filtration* and *Lime Softening* are not BAT for small systems (those with fewer than 500 connections) for the purpose of granting variances.

Note: EPA recognizes that most radionuclides emit more than one type of radiation as they decay. The lists of compounds labeled “alpha” or “beta” emitters identify the predominant decay mode.
Note: In this document the unit mrem ede/yr refers to the dose ingested over 50 years at the rate of 2 liters of drinking water per day.
Source: U.S. Environmental Protection Agency, 1991.

Gross Alpha. Alpha emitters naturally occur as radioactive contaminants, but several come from manmade sources. They may occur in either groundwater or surface water. At high exposure levels, alpha emitters are believed to cause cancer in humans.

Beta and Photon Emitters. Beta and photon emitters are primarily manmade radioactive contaminants associated with operating nuclear power plants, facilities that use radioactive material for research or manufacturing, or facilities that dispose of radioactive material. Some beta emitters occur naturally. Beta and photon emitters primarily occur in surface water. At high exposure levels, beta and photon emitters are believed to cause cancer in humans.

WHAT ARE THE TREATMENT TECHNOLOGIES FOR RADIONUCLIDES?

Whether or not a particular treatment technology effectively removes radionuclides from drinking water depends on the contaminant’s chemical and physical characteristics as well as the water system’s characteristics (e.g., source water quality and water system size). Other considerations include cost, service life, and co-treatment compatibility.

Table 2. Technologies for Radionuclides

Unit Technology	Limitations (See Footnotes)	Operator Skill Level Required[1]	Raw Water Quality Range and Considerations[1]
1. Ion Exchange (IO)	(a)	Intermediate	All groundwater
2. Point of Use (POU) IO	(b)	Basic	All groundwater
3. Reverse Osmosis (RO)	(c)	Advanced	Surface water usually requires prefiltration
4. POU RO	(b)	Basic	Surface water usually requires prefiltration
5. Lime Softening	(d)	Advanced	All water
6. Green Sand Filtration	(e)	Basic	
7. Co-precipitation with Barium Sulfate	(f)	Intermediate to Advanced	Groundwater with suitable water quality
8. Electrodialysis/Electrodialysis Reversal		Basic to Intermediate	All groundwater
9. Pre-formed Hydrous Manganese Oxide Filtration	(g)	Intermediate	All groundwater

[1] National Research Council (NRC). "Safe Water from Every Tap: Improving Water Service to Small Communities." National Academy Press. Washington, DC. 1997.

Limitations Footnotes

a. The regeneration solution contains high concentrations of the contaminant ions. Disposal options should be carefully considered before choosing this technology.

b. When POU devices are used for compliance, programs for long-term operation, maintenance, and monitoring must be provided by the water utility to ensure proper performance.

c. Reject water disposal options should be carefully considered before choosing this technology.

d. The combination of variable source water quality and the complexity of the chemistry involved in lime softening may make this technology too complex for small surface water systems.

e. Removal efficiencies can vary depending on water quality.

f. This technology may be very limited in application to small systems. Since the process requires static mixing, detention basins, and filtration; it is most applicable to systems with sufficiently high sulfate levels that already have a suitable filtration treatment train in place.

g. This technology is most applicable to small systems that already have filtration in place.

Source: Environmental Protection Agency, 1998.

EVALUATION OF TECHNOLOGIES

The following treatment technologies were evaluated for their ability to remove radionuclides from water:

- ion exchange (IE);
- point-of-use (POU) IE;
- point-of-entry (POE) IE;
- reverse osmosis (RO);
- POU RO; POE RO;
- lime softening (LS);
- greensand filtration;
- co-precipitation with barium sulfate;
- selective sorbents;
- electrodialysis/electrodialysis reversal (ED/EDR); and
- preformed hydrous manganese oxides (HMOs).

Ion Exchange. Small systems may readily use IE treatment, which removes approximately 90 percent of radionuclides. The effluent must be regularly monitored and the resin must be frequently regenerated to ensure that breakthrough does not occur. Ion exchange units may be controlled automatically, requiring less of the operator's time. However, it is necessary to employ a skilled operator to determine when regeneration is needed and to troubleshoot. Also, disposal of concentrated radionuclides can be expensive.

Ion exchange treatment generates wastes that include rinse and backwash water, and the resin. The rinse and backwash liquid waste includes brine, radium, and any other contaminants that the process removes.

Cation Exchange. A cation is a positively charged ion. Cation exchange resins exchange like-charged ions equally with protons—sodium ions (Na+), or in sodium-restriction cases, potassium ions (K+)—to remove undesirable cations from water. Cation exchange is often used to remove calcium and magnesium cations, and to treat hard water.

The amount of waste (rinse and backwash) that cation exchange typically generates ranges between 2 and 10 percent of the treated water.

Lime Softening. Lime softening can be used to remove radium from drinking water with 80 to 95 percent efficiency. Also, adding lime or lime-soda ash to water increases the pH of the water and induces calcium carbonate and magnesium hydroxide precipitation. Lime softening also is used to treat hard water.

Lime softening generates wastes that include lime sludge, filter backwash liquid and sludge, and sludge supernatant.

Reverse Osmosis (RO). Reverse osmosis effectively removes many inorganic contaminants, including heavy metals and radionuclides, such as radium and uranium. RO can remove 87 to 98 percent of radium from drinking water. Similar elimination can be achieved for alpha particle activity and total beta and photon emitter activity.

When using an RO system to remove radionuclides, performance depends on a number of factors, including pH, turbidity, iron/manganese content of the raw water, and membrane type. The pretreatment design depends on the quality and quantity of the source water. Existing treatment plants may already provide much of the required pretreatment—for example, coagulation/filtration of highly turbid surface water or iron removal for well waters. RO can be cost effective for small systems.

TECHNOLOGIES FOR RADIONUCLIDE REMOVAL NOT PROPOSED AS BAT IN 1991

Greensand Filtration for Radium Removal

Greensand filtration technology consists of a conventional filter box with manganese greensand replacing the traditional filtration medium. Studies indicate that greensand filtration removes up to 56 percent of radium.

This process generates wastes that include sludge and supernatant from the filter backwash, and eventually the greensand media must be disposed of.

PREFORMED HYDROUS MANGANESE OXIDE (HMO) FILTRATION

Costs for HMO treatment can be quite low, if filtration is already in place.

HMO filtration is similar to oxidation/filtration in its complexity and the operator skill it requires. Proper dosages must be determined, and if water quality varies, the dosage must be recalibrated. Once the proper dose is determined, dosing is relatively easy. HMO filtration requires simple equipment and is fairly inexpensive. Filters must be backwashed, which may require intermediate operator skill. Radium containing wastes include HMO sludge, filter backwash, and sludge supernatant.

CO-PRECIPITATION OF RADIUM WITH BARIUM SULFATE

Adding a soluble barium salt—such as barium chloride—to radium and sulfate contaminated water causes co-precipitation of a highly insoluble radium-containing barium sulfate sludge. This process has primarily been used for wastewater treatment. Mine wastewater treatment data indicates that this process removes up to 95 percent of radium.

This process generates wastes that include the barium sulfate precipitate sludge, filter backwash, and sludge supernatant.

OTHER TECHNOLOGIES

Some other technologies also are available that may remove radionuclides from drinking water. However, these technologies have not been fully tested for drinking water treatment or have only been used in industrial or experimental situations. Examples of processes that remove radium include selective sorbents (e.g., acrylic fibers or resins impregnated with manganese dioxide) and non-sodium cation exchangers (e.g., hydrogen ions and calcium ions).

WHERE CAN I FIND MORE INFORMATION?

(1) American Water Works Association & American Society of Civil Engineers. 1998. *Water Treatment Plant Design*; 3rd ed. The McGraw-Hill Companies, Inc.

(2) U. S. Environmental Protection Agency. September 1998. *Small System Compliance Technology List for the Non-Microbial Contaminants Regulated Before 1996*. EPA 815-R-98-002.

(3) U. S. Environmental Protection Agency. August 1998. *Federal Register / Notices*. Vol.63, No. 151.

(4) U. S. Environmental Protection Agency. October 1999. Office of Groundwater and Drinking Water. *Proposed Radon in Drinking Water Rule: Technical Fact Sheet* EPA 815-F-99-006. www.epa.gov/safewater/radon/fact.html

(5) U. S. Environmental Protection Agency. June 1991. Office of Ground Water and Drinking Water. *Radionuclides in Drinking Water - Fact Sheet* EPA 570/9-91-700.

(6) U. S. Environmental Protection Agency. July 1991. "National Primary Drinking Water Regulations; Radionuclides; Proposed Rule" *Federal Register*, Vol.56, No. 138.

(7) U. S. Environmental Protection Agency. March 1997. *Federal Register*, Vol.62, No. 43.

HAVE YOU READ ALL OUR TECH BRIEFS?

"Tech Briefs," drinking water treatment fact sheets have been a regular feature in the National Drinking Water Clearinghouse (NDWC) newsletter *On Tap* for more than four years. NDWC Technical Assistance Specialist Mohamed Lahlou, Ph.D., researches, compiles information, and writes these very popular items.

To order, call the NDWC at (800) 624-8301 or (304) 293-4191. You may also order online at ndwc_orders@mail.estd.wvu.edu or download fact sheets from our Web site at *www.ndwc.wvu.edu.*

Additional copies of fact sheets are free; however, postal charges may be added.

For further information, comments about this fact sheet, or to suggest topics, call Lahlou at one of the above numbers or contact him via e-mail at mlahlou2@wvu.edu.

USE OF REDOX POTENTIALS IN WASTEWATER TREATMENT

Enos C. Inniss
University of Texas
San Antonio, Texas

INTRODUCTION

Effective process control requires an understanding of where the process is with respect to reaction potential and completion. Online measurements allow for greater monitoring of the state of the process from flowrates to chlorine residuals. A measure of dissolved oxygen in the system is useful until anoxic and anaerobic conditions are encountered. Once those environmental conditions are achieved, as is the case with biological nutrient removal (BNR), the ability to measure a wider range of conditions may be achieved with oxidation–reduction potential (ORP).

BASICS OF OXIDATION–REDUCTION REACTIONS

Oxidation–reduction reactions also referred to as redox or oxido reduction involves the transfer of electrons and provides microorganisms with their free energy during transformations. The energy exchange is usually in the form of ATP (adenosine triphosphate), where the energy is in the bonds between the phosphoryl groups. Both aerobic and anaerobic metabolism produce ATP, but the anaerobic rate of production is much lower than that of the aerobic process.

The redox reaction typically transfers electrons from an electron donor (ED) to an electron acceptor (EA), where two half-reactions (or redox couples) can be derived to show the movement of the electron (Fig. 1). By taking the electron, the EA gains a net negative charge, while the ED loses an electron and gains a net positive charge.

THE NERNST EQUATION

The redox potential (ΔE) can be viewed as the electron pressure or the degree of tendency for electrons to be transferred and is defined by the Nernst equation (Fig. 2). The standard redox potential, ΔE°, is defined as the difference between the reduction potentials of the electron acceptor and donor when all components are in their standard state ($\Delta E^\circ = E^\circ_{\mathrm{EA}} - E^\circ_{\mathrm{ED}}$). This reduction potential is measured with respect to the hydrogen half-reaction.

Biochemical standard state (E° at pH 7) for the hydrogen half-reaction is -427 mV (negative 427 millivolts).

$$\mathrm{EA} + e^- \rightleftharpoons \mathrm{EA}^{(-)} \quad \text{reduction}$$

$$\mathrm{ED} \rightleftharpoons \mathrm{ED}^{(+)} + e^- \quad \text{oxidation}$$

Figure 1. Generic reduction and oxidation half-reactions (1).

$$\Delta\varepsilon = \Delta\varepsilon^\circ - \frac{RT}{n\mathscr{F}} \ln\left(\frac{[\mathrm{ED}^{(+)}][\mathrm{EA}^{(-)}]}{[\mathrm{ED}][\mathrm{EA}]}\right)$$

where
R = universal gas constant,
T = temperature,
n = 1 mole of electrons, and
$\mathscr{F}$ = faraday, the charge of 1 mole of ?

Figure 2. The Nernst equation for redox potential (1).

The more positive the half-reaction, the greater the tendency for the oxidized form to be reduced and therefore the electron acceptor is defined (1).

TYPICAL RANGES FOR REDOX POTENTIAL IN BIOLOGICAL SYSTEMS

The wastewater treatment process involves a multitude of environmental conditions in order to achieve complete treatment of influent constituents. Each of these general environmental conditions may be defined by several ranges of redox potentials (Fig. 3). As a general rule of thumb, readings taken with an ORP electrode that are less than -200 mV are indicative of anaerobic or reducing conditions. ORP readings between approximately -200 and $+200$ mV are for anoxic conditions. Aerobic or oxidative conditions are present above $+200$ mV.

The typical wastewater processes for the anaerobic range include sulfate reduction, occurring primarily in the collection system, and methanogenesis and fermentation, which are more typical of the anaerobic digestion process. This range is defined by sulfate as the primary terminal electron acceptor. The anoxic range is useful for biological nutrient removal, where nitrate serves as the dominant terminal electron acceptor. The aerobic range involves the activated sludge process and disinfection, where oxygen or the disinfectant serves as the terminal electron acceptor (2,3).

NUTRIENT REMOVAL

For biological nutrient removal (BNR), there is a need to switch between anaerobic/anoxic conditions and aerobic

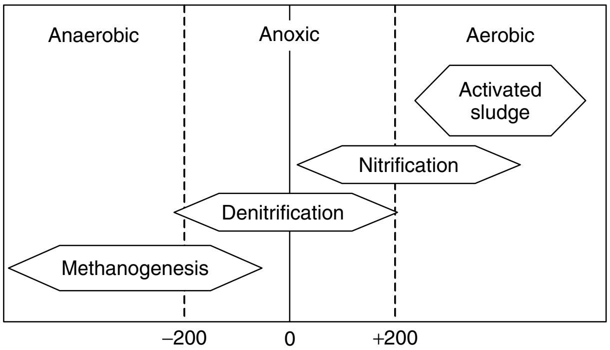

Figure 3. Redox potential ranges (mV) for typical wastewater processes. (Modified from References 2 and 3).

conditions to achieve complete removal of nitrogen and phosphorus compounds from the water. For phosphorus removal, anaerobic conditions result in a release of orthophosphate to gain energy by using up ATP. The higher energetics of aerobic metabolism then causes phosphorus uptake, so that the microorganisms can store the energy. Similarly, anoxic/anaerobic conditions allow for the conversion of nitrate and nitrite to nitrogen gas, while aerobic conditions allow for the oxidation of ammonia to nitrite and nitrate. Controlling these two processes, denitrification and nitrification, respectively, requires the ability to switch and establish the proper environmental conditions. ORP readings afford the operator greater control over these conditions (4–6).

OTHER POTENTIAL APPLICATIONS

Other applications for ORP monitoring and control are currently being researched and implemented. Waste-activated sludge hydrolysis may be enhanced to increase its bioavailability during sludge handling (7–9). Dosing may be controlled in the disinfection process (10). Aeration may be controlled and thus the organism growth influenced in the activated sludge process (11).

CONCLUSION

With increasingly better electrode technology and online monitoring and data storage capability, process control using ORP has a future in reducing operational cost and improving operational efficiency.

BIBLIOGRAPHY

1. Voet, D. and Voet, J. (1995). *Biochemistry*, 2nd Edn. Wiley, Hoboken, NJ, pp. 428–437.
2. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters*, 3rd Edn. Wiley, Hoboken, NJ, p. 477.
3. Byl, T.D. and Williams, S.D. (2000). *Biodegradation of Chlorinated Ethenes at a Karst Site in Middle Tennessee*. U.S. Geological Survey, Nashville, TN, Water Resources Investigations Report 99–4285, p. 4.
4. Hidaka, T., Tsuno, H., and Kishimoto, N. (2003). Advanced treatment of sewage by precoagulation and biological filtration process. *Water Res.* **37**: 42559–4269.
5. Lee, D.S., Jeon, C.O., and Park, J.M. (2001). Biological nitrogen removal with enhanced phosphate uptake in a sequencing batch reactor using single sludge system. *Water Res.* **35**(16): 3968–3976.
6. Metcalf & Eddy, Inc. (2003). *Wastewater Engineering: Treatment and Resuse*, 4th Edn. McGraw-Hill, Boston, pp. 611–629.
7. Chang, C.-N., Ma, Y.-S., and Lo, C.-W. (2002). Application of oxidation–reduction potential as a controlling parameter in waste activated sludge hydrolysis. *Chem. Eng. J.* **90**: 273–281.
8. Chen, G.-H. et al. (2003). Possible cause of excess sludge reduction in an oxic-settling-anaerobic activated sludge process (OSA process). *Water Res.* **37**: 3855–3866.
9. Filani-Meknassi, Y., Tyagi, R.D., and Narasiah, K.S. (2000). Simultaneous sewage sludge digestion and metal leaching: effect of aeration. *Proc. Biochem.* **36**: 263–273.
10. Yu, R.-F. (2004). Feed-forward dose control of wastewater chlorination using on-line pH and ORP titration. *Chemosphere* **56**: 973–980.
11. Chen, G.-H., Yip, W.-K., Mo, H.-K., and Liu, Y. (2001). Effect of sludge fasting/feasting on growth of activated sludge cultures. *Water Res.* **35**(4): 1029–1037.

REPAIRING DISTRIBUTION LINE BREAKS

VIPIN BHARDWAJ
NDWC Engineering Scientist

Occasionally, water systems encounter situations where they must repair distribution system pipes. Corrosion and tuberculation (i.e., buildup of sediment, dirt, or rust) may have caused small leaks in the lines, or worse, a major line break may occur, creating an emergency situation. This Tech Brief discusses ways to manage these situations and outlines steps a utility may take to repair the distribution system.

WHAT CAUSES PIPES TO BREAK?

Distribution lines can break for a variety of reasons. Excessive weight, such as increased traffic continuously running over a buried pipe, can trigger a line break. Also, extremely cold temperatures can cause breaks because when water freezes, it expands. But environmental conditions are not the only reason a line may break.

Sometimes utility workers may unknowingly install pipes that have defects from the manufacturing process, and they are not strong enough to handle high-pressure surges. Consequently, the pipes may split or crack. Other times pipes may not have been properly installed into the trench, creating a situation where it's only a matter of time before a line bursts.

Corrosion and tuberculation are two more reasons that pipe can rupture. Corrosion may cause breaks or leaks because acidic conditions can cause pitting or holes in a metallic pipe. In addition, tuberculation can cause high-pressure pockets in some areas of pipe, because water that's under pressure may not be able to easily move beyond the area of buildup.

WHEN ARE REPAIRS NEEDED?

Over time, even small leaks can waste a substantial amount of water, which is expensive. When water system personnel detect a leak, they should repair it immediately—no matter how small it is.

When water comes out of the ground, it is obvious a leak exists nearby and its location is fairly easy to determine. However, scheduling periodic leak detection exercises helps water systems determine where and when they should make repairs in situations that aren't so obvious.

Because most leaks are not visible, all a utility worker may know is that the system is losing water. Someone then has to find the leak. This involves using listening devices, such as Geophones® or other electronic equipment. (For more information about leak detection, see the *Tech Brief* in the Spring 2001 *On Tap*, and the article, “On the Trail of the Elusive Water Leak” in the Summer 2003 *On Tap*.)

One of the best ways to trace water loss is to conduct a water audit. A water audit helps systems keep unaccounted for water loss to the recommended 15 percent or less of the total water a system produces. Conducting a water audit and following up on the results can help the utility control its water losses. (For more information about water audits, see the Fall 2002 *On Tap* article, “The Economics of Water Loss: What is unaccounted for water?”).

To conduct a water audit, utility workers will need to take flow measurements over a 24-hour period. This task will require pressure gauges. Utility workers also should make sure to check meters on main lines first to see if they are accurate and calibrated. As water flows from main lines into smaller lines, utility workers take flow measurements at a number of points in the smaller lines. Large, unaccounted for nighttime flows indicate a leak.

EMERGENCY REPAIRS

Sometimes leaks aren’t small. Sometimes they gush. If a main breaks, utility workers must repair it immediately. To avoid too much stress, utilities need to have a plan of action ready that details the necessary equipment, personnel, and procedures. This plan is usually referred to as an emergency response plan (ERP) and should involve law enforcement, fire protection, and city officials because utility personnel must first figure out if the leak poses a hazard to life or property.

As soon as a major break is detected, utility workers should notify customers that their water will be turned off while workers make repairs. Valves must first be shut off in the area to isolate the break and to prevent further water loss.

More often than not, valves can become hard to close or may even break if they are not used on a regular basis. Systems should have regular valve-exercising and hydrant-flushing programs to alleviate this problem. (For tips about developing a valve-exercising program, see the article, “Why bother with a valve-exercising program?” in the Winter 2004 *On Tap*. For more information about hydrant flushing, see “Fire Hydrant Operation and Maintenance,” in the Fall 2002 *On Tap*).

If possible, repair the leak without shutting off the entire water supply. Figure 1 shows a section of a distribution system with an arrangement of four valves. The leak is near a corner before valve 2. Water flow is from left to right through the loop. When making repairs, close valve 4, valve 3, and valve 2, keeping valve 1 open till the very last minute (Fig. 3). Valve 1 is then closed slowly. Utility workers then make repairs on the leak and turn the water back on immediately. Closing the valve nearest the leak at the last minute allows the line to remain under pressure, preventing back siphoning and back pressure that can cause contaminants to get sucked in through the leak.

If, however, the valves cannot be opened and, consequently, the water cannot be turned off, one practice is to open several fire hydrants in the area. While this method will not stop the water, it will lower the water pressure enough that utility workers can repair the leak. Once the repairs have been completed, workers then slowly turn the water back on. Turning the water back on slowly ensures that water hammer will not become a problem.

Water hammer is a hydraulic shock that happens when a sudden change in velocity occurs, such as quickly closing a valve. Shock waves occur within the pipes, which travel back and forth and cause a “bang” within the pipe. This bang is the hammer that can cause pipes to burst. (See the *Tech Brief* about water hammer in the Winter 2003 *On Tap*.)

When a line is under pressure, less probability exists for contaminants to get in through the break. But if contamination does occur, the utility should take the appropriate actions to inform the public through local radio, TV, or other new service. Notification should outline

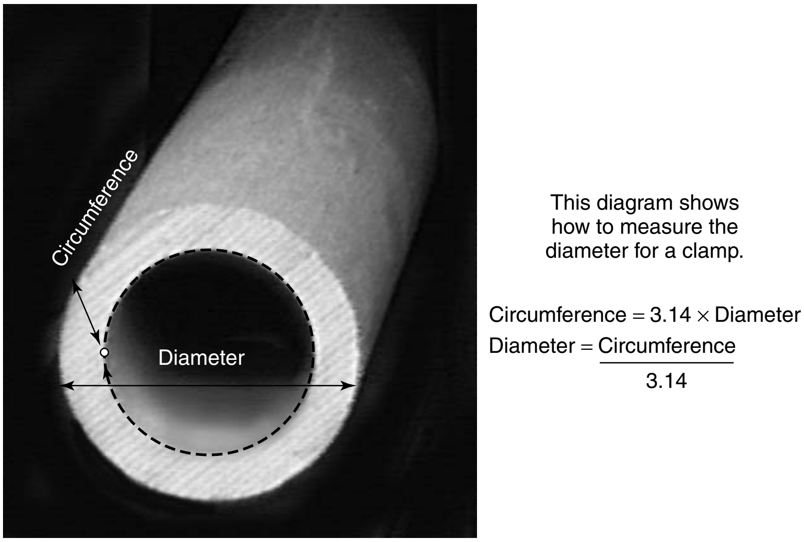

Figure 1. Calculating diameter from circumference. *Source*: *Water Transmission and Distribution*, AWWA, 2nd Edition.

steps customers can take to be sure their water is safe. In case of an emergency, utilities should refer to their ERP that should specifically outline what they need to do. For example, they also may need to notify state primacy agencies, depending upon what level of action they need to take.

DIGGING UP THE PIPE

Once utility workers detect the leak, they have to dig up the pipe. The trench must be parallel to the pipe on both sides and be deep enough that a person can work around the pipe to repair it. Workers must calculate the trench's depth and width based on pipe size. Water will collect in the pit and will have to be pumped out.

Utility workers also should remember that it's an Occupational Safety and Health Administration (OSHA) rule to shore the trench if it's five feet or deeper—or less than five feet if the conditions warrant it, such as extremely sandy or muddy soil.

REPAIR METHOD

If the break in the pipe is very large, that portion of the pipe will have to be cut off and a new section put in place. Cutting a pipe requires saws, pipe cutters, welding equipment (if the pipes are made of steel), pipe wrenches, couplings, and other hand tools.

If the break in the pipe is small, such as a rupture across the pipe, utility workers can wrap a flexible clamp or sleeve on it, which acts like a bandage for the pipe. Wrapping a coupling around the pipe, which is a device that holds two separate pieces of the pipe together in a "covering" device, is another way to repair a small break (see Fig. 2).

A clamp installed on the pipe has a gasket (usually made of rubber) that covers the ruptured section and helps to maintain a secure fitting. There is an arrangement of screws (preferably stainless steel) that are tightened around the clamp. Utility workers will need to clean and scrape corrosion and dirt from the pipe surface and then disinfect it with bleach.

In addition, they should measure the outside circumference of the pipe to determine the exact diameter of a coupling. (The diameter is equal to the circumference divided by 3.14159.)

After the pipe is repaired, workers should turn on the water slowly to check if there are any remaining leaks. The trench is typically covered with filling material, such as sand, crushed stone, or processed material, and compacted. The workers should ask system customers to turn on their taps and allow the water to run for some time because the water may or may not have sediment in it.

BEING PREPARED

A water utility should have an emergency vehicle, such as a pickup truck, stocked with essential tools, including pipe cutters, wrenches, shovels, traffic control equipment, barricades, flashers, cones, and flashlights. This preparation will save time in an emergency. They should keep the truck's gas tank full and not loan out tools or equipment. Qualified individuals should be on call for emergency response at all times, including scheduled holidays.

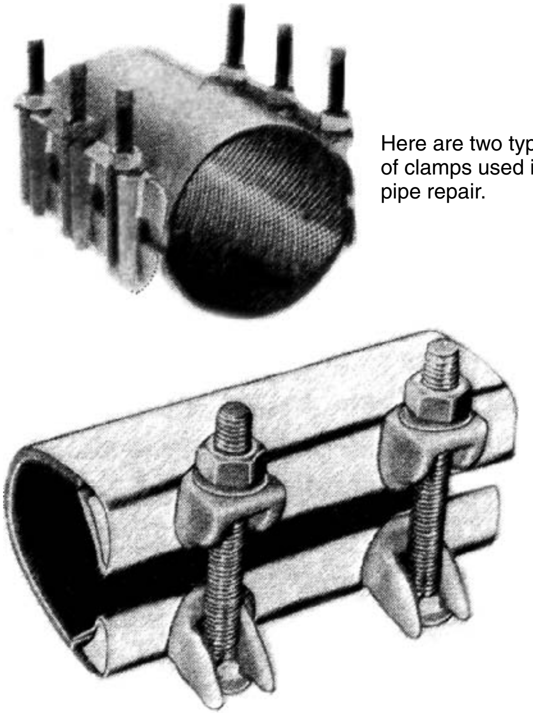

Here are two types of clamps used in pipe repair.

Figure 2. Types of clamps. *Source*: *Water Transmission and Distribution*, AWWA, 2nd Edition.

LAYING NEW PIPES

In some situations, a utility may have to replace an entire pipeline. New pipes may have to be hauled in and installed. Care must be taken to prevent pipes from breaking during shipping or storage, and the pipes must be installed correctly. Approvals or permits from the relevant state or local agency must be obtained before starting a new project.

Pipes should be placed where they will be installed before digging out the trench. One end of a pipe is shaped like a bell, and the other end is straight. Pipes are joined together straight end into bell end and secured with a gasket. Pipes should have their bell-shaped ends placed in the direction in which the installation of the pipe will proceed.

Excavating the earth to install pipes is the most expensive part of many replacement projects. Before beginning the repair project, other utility installations, such as sewer and gas lines, must be located. In addition, utility workers must notify nearby property owners and the general public.

Once the project is underway, utility workers must:

- determine the trench's width according to the pipe size required,

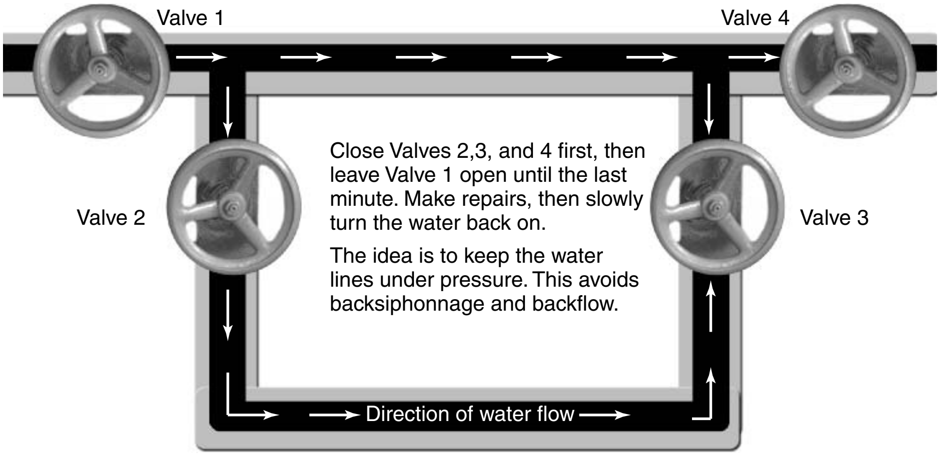

Figure 3. Overhead view of a distribution system. *Source*: *Operator Basics* CD, Montana Water Center.

- lower pipes carefully into the trench using ropes,
- dig out additional soil at the bell ends so that the pipe is supported along its entire length, and
- add some backfill material or bedding to the trench to support the pipe.

Pipes have gaskets on the bell-shaped ends. When laying new pipelines:

- the gasket must be clean, and
- pipes must be pushed all the way into the bell of an adjoining pipe to make a tight fit.

COMPLETING REPAIRS

Once workers have completed repairs, they should turn the water back on and check for leaks before backfilling the trench. The new lines should be disinfected and thoroughly flushed to remove all sediment and dirt.

According to the American Water Works Association's standards, water samples should be taken every 1,200 feet following a new installation. Samples must be bacteria free before the line can be put back into service. Workers should be sure to keep a record of where the repair needed to be made and its cause.

WHERE CAN I FIND MORE INFORMATION?

American Water Works Association. 1996. *Water Transmission and Distribution*, Second Edition. Denver, Colorado: AWWA.

American Water Works Association. 1999. *Water Distribution Operator Training Handbook*, Second Edition. Denver, Colorado: AWWA.

Male, James T., Walski, Thomas M. 1991. *Water Distribution Systems, A Troubleshooting Manual*, Second Edition. Chelsea, MI: Lewis Publishers, Inc.

OSHA Rules and Regulations, CFR 1926.652 (a) at www.osha.gov/doc/outreachtraining/htmlfiles/exacavate.html

NDWC Engineering Scientist **Vipin Bhardwaj** has a bachelor's degree in chemical engineering, and master's degrees in environmental engineering and agriculture from West Virginia University.

ROLE OF SMALL WATER RESERVOIRS IN ENVIRONMENT

Janusz Guziur
University of Warmia and
Mazury in Olsztyn
Olsztyn, Poland

> Motto:
> 'Water ... you are the greatest treasure ever existing in all the world'.
>
> (Antoine Saint-Exupery)

The earth's fresh water supplies are diminishing at a terrifying rate. The loss of fresh-water supply per capita has reached 40% since 1970.

The annual gross supply of water in the world per person is 10,000 m^3, 1500 m^3 in Poland, 2400 m^3 in Germany, and 2800 m^3 in France. These data converted into the water availability index in cubic meters daily per capita in some European countries are presented in Fig. 1 (1,2).

Melioration works (land improvement) completed in the past 50 years in many countries in Europe, Asia, or America have focused mainly on enlarging areas of farmland (increasing food production). They were intended to drain off water rapidly and dry and reclaim land. As a result, many natural watercourses have been replaced by drainpipes (ditches) and simple canals. Other water features have also disappeared, for example, small natural lakes and other small waterbodies, natural flood waters, midfield trees, and coppices. In Poland, 80% of small ponds and water mill dams have also been liquidated (3). Natural water retention potential has been distorted. As a result of declining available freshwater supplies in many areas of the world, interest is increasing in expanding the use of small water reservoirs.

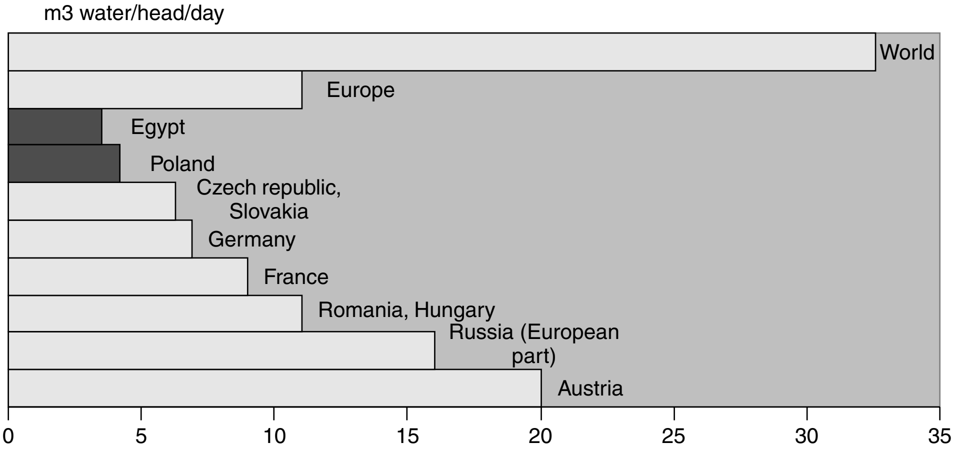

Figure 1. Water availability index in different European countries in cubic meters water per person daily (1).

The present situation calls for urgent action to reclaim lost properties in the natural environment. Natural landscape reclamation can be attained in two ways:

- *Natural Measures (Proecological, Nontechnical)*: by preserving small water reservoirs and cave-in lakes, natural wetlands and floodplains, village ponds and alike, as well as by planting forests, trees, and shelter belts in degraded areas.
- *Technical Measures*: by water retention, especially so-called small retention objects.

Natural retention means

- increased potential for water self-purification, accompanied by enriched landscape values.
- enhanced landscape attractiveness, which stimulates the development of ecotourism and also encourages the local population to enjoy living in more comfortable surroundings.

Of about 28 types of small natural inland water reservoirs and water ecosystems (4,5), wetlands known as marshes (swamps, bogs, wetlands, fens, mires) play a key role in the environment and water management. They are structural and functional waterlogged ecosystems that are transitional forms between typical land and typical water ecosystems. Features common to all marshy ecosystems are permanent or temporary waterlogging of the substrate, the presence of hydrophilous plants (hydrophytes), and specific soils. Such ecosystems are inhabited by rare plant and animal species and function as important flight paths for seasonally migrating birds as well as water plant habitats for nomadic species (6). They are also nesting and breeding areas for some rare species, frequently those on the verge of extinction (including endemic species). Owing to their specific soils and flora, marshes are often nicknamed 'kidneys of the landscape' or small wastewater purification plants (7).

The sanitary impact of wetlands on the environment is gaining in importance due to considerable contamination of surface and underground waters. A waterlogged area can function as a trap, source, or transformer of matter reaching it from the outside; each role depends on the type, age, and use of the ecosystem. Wetlands become traps mainly in spring and early summer. They are a source of phosphorus (P) when the vegetative season is ended (disappearing assimilation). Depending on the hydrologic type of supply, waterlogged areas are divided into four types (7):

- *Ombrogenic*: present in watershed zones, at the highest points, represented by high and transitional moors;
- *Topogenic*: present on lowlands and near lake areas as low, oligotrophic moors;
- *Soligenic*: marshy grounds in valleys near terraced parts of river valleys;
- *Fluviogenic*: riparian marshes, subject to the erosive impact of rivers, heavily bogged, with highly trophic (fertile) river sediments.

These types of wetlands have an exceptionally beneficial influence on the hydrologic regime of the terrain. In particular, they help retain water in the catchment, improve its water balance, and retain water in the countryside, which involves preventing water outflow from the catchment and rational time distribution of water (2,8). Wetlands also have an indirect positive effect on agricultural production (plant production) and pond fish culture (9).

In some countries, marshes and bogs cover large areas. In Poland, for example, natural and transformed marshes cover 13.1% of the country's total area (more than 4.3 million ha), of which 10.2% belongs to open (not forested) marshes, preserved in a nearly natural state.

More than 56% of Poland's natural and transformed marshes are contained in the European system ECONET Poland (European Ecological Network), where they constitute 16.7% of the total network area. The ECONET-PL network is composed of international and national core areas (that have a wealth of habitats and species) and ecological corridors. The latter are situated mainly in river valleys, providing links between core areas (7).

The most effective and currently desirable is the type of water retention that involves retaining the largest

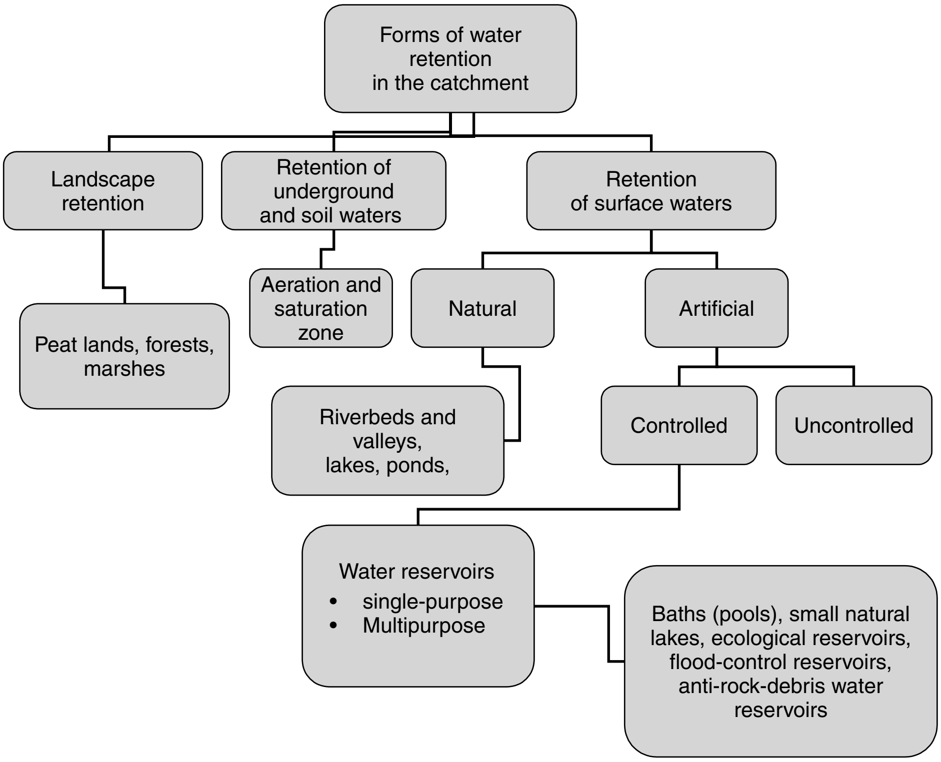

Figure 2. Basic forms of small retention in a catchment (10).

amounts of water possible in surface and near-surface water circulation systems using different measures. In short, the objective is to prevent unproductive water flow to the sea.

The total retention capacity of a catchment is the sum of individual elements that comprise many constituent forms of the catchment (Fig. 2). Economically, retention can be divided into natural and artificial forms (10). The major forms of natural retention include

–retention by river beds and valleys,

–retention by peatlands, marshes, and wetlands,

–retention by soil, underground, and surface waters,

–retention by stagnant waters—especially forest and soil retention.

Forest and soil retention are difficult to quantify. Estimates suggest that an increase in soil retention by 20 mm across 50% of the agricultural lands in Poland amounts to an additional 2 billion m^3 of water (11,12).

Artificial water retention consists mainly of retaining water in artificial reservoirs (e.g., dammed reservoirs in lowlands and mountains) and less often of lake retention obtained by artificial water lifting in natural lakes. Two types of retention are distinguished—uncontrolled and controlled. The latter is represented by artificial reservoirs of constant water lifting equipped with regulation devices, which enable water circulation control (10).

Small water retention relies mainly on waterbodies of different types. Waterbodies improve the water balance of a catchment; depress the level of flood waves, high water (after torrential rains), and thaw waves (in spring); increase low flows (summer drought seasons); and increase air humidity and soil moisture, thus improving the microclimate and conditions for arable crops.

Water reservoirs are objects that can be described by using measurable and nonmeasurable parameters. The following are most often used to compare waterbodies: size of retention (ha, m^3), runoff compensation ratio, type of regulation and discharge facilities, and hydraulic and economic functions (Table 1).

Development of water retention is an important element of integrated water management, particularly in terms of maintaining quality of water supplies and producing a positive effect on the outflow and levels of water from its surface, which in turn has some influence on the formation and shape of flood tides.

In addition, retention objects perform a variety of other tasks and functions, of which the most important are

- industrial water supply (e.g., water intake),
- fire protection
- relaxation, sports, and recreation,
- fisheries (professional and amateur),
- landscape and environmental management.

Surface water retention is attained through the following activities:

- construction and reconstruction of water reservoirs,
- construction of water lifting installations in rivers,
- filling fish ponds (especially larger pond complexes),
- water retention in amelioration installations and systems,

Table 1. Classification of Storage Reservoirs[a]

Classification Criteria	Size and Type of Reservoirs			
Retention size	**Small**		Medium	Large
Height of a dam (m)	From **5**		5 to 15	Over 15
Water capacity (m^3)	to **100,000**		to 500,000	Over 500,000
Runoff compensation (below the reservoir)	Long-term	Seasonal	Annual	Daily
Number of functions	Single-purpose[b]		Multipurpose[c]	
Hydraulic devices (discharge)	Uncontrolled : dry		Controlled: of a defined (set) water lifting level	

[a] References 10 and 13.

[b] Retention function (general).

[c] Potable water retention, flood control, irrigation of arable crops, recreational and water sport reservoirs, land melioration, electric power, sailing reservoirs, etc.

- water lifting in natural waterbodies (lakes),
- protection and management of small water reservoirs and cave-in lakes,
- flooding of valley depressions (bogs and marshes, cutoff meander lakes, old riverbeds).

The following water retention methods seem best for the natural environment:

Higher Water Retention on Areas Previously Ameliorated or Planned To Be Ameliorated

The primary objective is reconstruction of weirs on drainage ditches, which brings large contributions to water supplies [in Poland, about 500 million m^3: (11)]. The second stage comprises construction of small retention reservoirs at the end points of amelioration systems, that is, before water outflow to a river or sea. Some experts on hydrology and hydraulic engineering, however, claim that large retention reservoirs should not be built as long as 'all possibilities of constructing small retention reservoirs have not been exhausted.' It is also important to abandon amelioration of peat bogs or waterlogged meadows and pastures, which further improves water retention potential.

The most superior natural values are attributed to nonregulated rivers in undeveloped valleys. Such valleys contain a mosaic of different water ecosystems: cutoff meander lakes (old riverbeds), peat bogs, dry and moist meadows, and the so-called alder carrs (8,9). Flood control involves several methods that help preserve these unique natural water ecosystems. Of the two possible flood protection strategies—increasing water flow capacity of river valleys (hydraulic installations) or increasing water retention in the whole watershed—most specialists opt for the latter solution as more favorable for the natural environment (12,14,15). Of all methods for increasing water retention, the following seem recommendable from the viewpoint of preserving the natural environment:

Maintaining and Increasing Water Retention on Lands Used for Agricultural Production

The most important activities include doing without heavy agricultural machinery (which compacts arable layers of soil), plowing across the slope, creating terraces, not burning grass and stubble, and re-creating midfield water ponds.

Increasing Water Retention in Forested Areas

There is an array of activities available, including afforestation of lands with forest polycultures (monocultures have inferior water retaining capacity), safe transportation of felled trees (soil erosion risk), use of cable trains for transporting felled trees in mountainous areas, biological consolidation of streams, and no planning or building of main roads near streams and rivers. Despite numerous controversies, forests in river valleys are largely categorized as water protecting forests and therefore subject to less intense forest management practices.

Maintaining and Increasing Water Retention in River Valleys

It should be remembered that nonregulated rivers that have undeveloped valleys provide the highest natural values. Water retention measures undertaken in river valleys should be integrated with flood control. A partial solution that could reconcile flood control requirements with natural protection is the construction of so-called polders, waterlogged and nonreclaimable depressions. There are two types of polders: overflow polders that are flooded each time the water level rises (more favorable for protecting nature), and control polders (better for flood protection). Such solutions are used, for example, in Germany in the Rhine river valley near the water step in *Iffezheim* (16). Of all types of dammed reservoirs, dry reservoirs, which cause minimal changes in the hydrologic regime, are responsible for the smallest losses in river valley ecosystems. As far as wet reservoirs are concerned, small retention reservoirs are least damaging. According to Lewis and Williams (17), the least damage to the environment is caused by those wet reservoirs that are located in a river valley and are connected to the river by an inflow and an output canal.

To preserve ecosystems below a large reservoir, it is necessary to cause artificial floods (lasting a few days) that discharge large amounts of water in the spring. Natural and hydrologic effects of an artificial flood on the Colorado River in the United States (1996) are now a subject of wide research, so far yielding positive and promising results (18).

The water retention effects of the aforementioned measures are quite varied and, apart from typical retention reservoirs, rather limited. Their capacity for storing water (known as retention capacity) is usually measured in hundreds of thousands cubic meters of water (less frequently in millions of m^3). In Poland, objects whose capacity is less than 5 million m^3 are referred to as small retention objects (3,19).

Fish ponds—An Undervalued Element of Small Retention

Ponds are shallow artificial retention reservoirs, built in less fertile areas, usually waterlogged and agriculturally useless. Fish ponds make aquaculture possible, and they also facilitate ecological use of water and systematic water alimentation (permeation, leakage), more intense alimentation occurs when water is discharged in spring and summer and during the autumn fish catch. Pond objects are localized in the upper parts of a catchment and are supplied with water from small streams and by precipitation. Ponds significantly delay water outflow and make water flow more uniform in the whole watercourse (9,20) (Fig. 3).

Large fish pond objects and ponds of higher capacity (more than 1–2 thousand ha) located in lower parts of a catchment make it possible to store water and summer flood tides (so-called St John's floods—rains in midsummer), thus successfully alleviating the course of floods. In Europe, such pond complexes are found in the region of Milicz near Wrocław (5 thousand ha), and near Čs. Budějovic (22 thousand ha) in the south of the Czech Republic. The largest ponds in those complexes (regions) cover 100 to 700 ha (2,4,5,21,22).

The water retention potential of such pond, lies in the fact that they store water in those periods when it would otherwise flow away fruitlessly. Ponds are filled with water when the water level in rivers is the highest, meaning that it does not collide with the water uptake by other users (consumers). On the contrary, it has a positive effect on the whole water management in the catchment (23), and, particularly, on flood control. A clear example of the beneficial influence and protection against floods is the Czech pond Rožemberk (711 ha), built at the end of the sixteenth century and used for fish farming until today. In 1889 (and more recently in August 2002), it protected the town of Třeboň from a devastating flood and complete damage by spilling its waters over an area of 2,000 ha and storing more than 50 million m^3 water (3).

Figure 3. Small retention—fish ponds in Ostróda, Poland.

In Poland, for instance, there are about 71 thousand ha of fish ponds, including small ponds at farmsteads, which store about 900–950 million m^3 water, one-third of the country's total water retention (2). Regardless of these figures, many countries of central Europe pay little attention to the role of ponds and possible inclusion of fish ponds in water and amelioration systems when planning water management, and local water retention projects, in particular (4). Such notorious underestimation of the nonproductive value of fish ponds results from an erroneous assumption that fish ponds are exclusively fish production reservoirs. The truth is that, apart from their fish breeding function, they have more than 50 positive characteristics in terms of water management, agriculture, environmental protection, landscape values, water sports, angling, and professional or economic stimulation of the local population (5,22,24–27). Therefore, such merits can often exceed the primary fish breeding value of ponds in their economic assessment (28).

In general, dammed retention reservoirs significantly alter the river valley landscape, although they can also create suitable nesting and feeding areas for fish and birds, once islets and shallow shores are formed. This effect can partly compensate for the losses in the natural environment caused by construction of dammed reservoirs.

It should be added that the European Union has accepted the 'Habitat Directive' to preserve, nearly unchanged, valleys of large rivers and adjacent small reservoirs and floodplains with their wealth of ecosystems. The high diversity of species in such areas is further protected by the Convention on Biodiversity Protection, signed in 1992 by many European countries, including Poland (7,15).

BIBLIOGRAPHY

1. Szczerbowski, J.A. (1998). *Fishing and Angling Encyclopedia*. (In Polish). IRS, Olsztyn.
2. Guziur, J., Białowas, H., and Milczarzewicz, W. (2003). *Pond Aquaculture*. (In Polish). Oficyna Wyd, Hoża, Warszawa.
3. Guziur, J. (1991). *Fisheries in Small Inland Reservoirs*. (In Polish). PWRiL, Warszawa.
4. Guziur, J. (1988). Characteristics of small water reservoirs and their potential use for aquaculture Part, I-VI. (In Polish). *Acta Acad. Agric. Tech. Olst.* **16**: 15–66.
5. Guziur, J. (2000). Significance of non-productive (external) value of small inland waterbodies. *Zesz. Nauk. Tow. Inżyn. Ekolog.Polit., Rzeszów*. **1**: 62–73.
6. Kabisch, K. and Hemmerling, J. (1981). *Tümpel. Teiche und Weiher Oasen in Unsere Landschaft*. Edition Leipzig.

7. Raport SGGW Warszawa. (2002). *Use and Management of a Catchment Area—Part. 6. Small Waters Retention*. (In Polish). Available: http://levis.sggw.waw.pl/ ozw1/zintegrowgospwod/ZintegrowanagospwodREW20/U.
8. Starmach, K., Wróbel, S., and Pasternak, K. (1976). *Hydrobiology. Limnology*. (In Polish). PWN, Warszawa.
9. Hynes, H.B. (1970). *The Ecology of Running Waters*. Liverpool University Press.
10. Radczuk, L. and Olearczyk, D. (2002). Small storage reservoirs as an element of water balance improvement in the agricultural catchment. (In Polish). *Zesz. Nauk. AR w Krakowie* **393**(23): 139–149.
11. Mioduszewski, W. (1994). Water Supply Conservation and Shaping in Small Agricultural River Catchments. (In Polish). *Mater. Inform*. 25, Falenty IMUZ.
12. Mioduszewski, W. (2002). Recreation of the retention of small river catchment as an element of flood and drought control. (In Polish). *Gospod. wodna, Warszawa* **11**: 12–16.
13. Dziewoński, Z. (1972). *Storage reservoirs of agriculture*. (In Polish). PWN, Warszawa.
14. Ryszkowski, I. and Kędziora, A. (1996). Small water retention in the rural landscape. (In Polish). *Zesz. Nauk. AR Wrocław*. **289**: 217–225.
15. Jankowski, W. (2002). Ecologic flood control methods. (In Polish). Available: http://www.tnz.most.org.pl/pow/ios_2.htm.
16. Nieznański, P. (1999). The upper course of the Rhine river as an example of possible combination of flood control and nature conservation. (In Polish). *Polder*. **15**: 16–21.
17. Lewis, G. and Williams, G. (1984). *Rivers and Wildlife Handbook: A Guide to Practices Which Further the Conservation of Wildlife on Rivers*. RSPB.
18. Collier, M.P., Webb, R.H., and Andrews, E.D. (1997). Experimental flood in the Grand Canyon, Colorado, the USA. (In Polish). *Świat Nauki*. **3**(967): 62–69.
19. Hochman, L. (1988). Limnological charakteristics of irrigation reservoirs in southwestern Slovakia (Czech Rep.). *Prace Ustavu Ryb. Hydrob. Bratislava* **6**: 139–180.
20. Manson, P.W. and Schwartz, G.W. (1968). Some aspects of the hydrology of ponds and small lakes. *Agricul. Exper. Station—Univ. of Minnesota, Tech. Bull.* **25**(7): 88.
21. Lavický, K. (1978). Function of fish ponds in a creation of life environment. In: *Current State and Perspectives of Czech Aquaculture*. CSAZ Praha, pp. 43–50.
22. Guziur, J. and Stafiniak, J. (2000). Polish experience with culture of common carp in single ponds and small water reservoirs. *Bull. VÚRH-Ju Vodňany*. **4**: 174–179.
23. Boyd, C.E. (1985). Pond evaporation. *Trans. Am. Fish. Soc.* **144**(2): 299–305.
24. Welsh, G. and Neils, K. (1978). Fish farming on the home front. *Fish Farm Aquacult. News*. **4**(6): 19–23.
25. Gergel, J. and Kratochvill, A. (1984). Teiche dienen nicht nur der Fischproduktion. *Z.Binnenfisch. DDR*. **3**(5): 178–182.
26. Szumiec, M.A. (1990). Role of ponds in landscape. (In Polish). *Biul. Inform. IRS Żabieniec*. **16**: 1–7.
27. Rożański, S. (1996). The additional benefits of fish farming in small water bodies. *Acta Acad. Agric. Ac Tech. Olst. Protectio Aquar. et Piscat.* **22**: 83–94.
28. Turkowski, K. and Guziur, J. (1999). *Small water reservoirs as an element of rural area management (case study of region of Warmia and Mazury, North. Poland)*. (In Polish). Mater. XII Sympoz. Nauk., Instrumentalizm rozwoju obszarów wiejskich. AGH, Kraków, pp. 48–56.

RESERVOIRS, TOWERS, AND TANKS DRINKING WATER STORAGE FACILITIES

VIPIN BHARDWAJ
NDWC Engineering Scientist

CHRIS METZGAR
Graphic Designer

After water leaves the treatment plant but before it reaches the customer, it must be adequately and safely stored. This Tech Brief explores the various aspects of water storage.

A water distribution system transports water from the treatment facility to the user. The distribution system should supply water, without impairing its quality, in adequate quantities and at sufficient pressures to meet system requirements.

The facilities that make up the distribution system include finished water storage; pumping, transmission and distribution piping supply mains; and valves.

Storage facilities—such as reservoirs, towers, and tanks—provide storage for treated water before it is distributed. The water distribution system should have storage so that it is capable for basic domestic purposes, commercial and industrial uses, and to accommodate the flows necessary for emergencies such as fire fighting.

STORAGE RESERVOIR FUNCTIONS

Service reservoirs provide the following functions:

- provide a reserve of treated water that will minimize interruptions of supply due to failures of mains, pumps, or other plant equipment;
- help maintain uniform pressure;
- provide a reserve of water for fire fighting and other emergencies;
- act as a relief valve on a system of mains supplied by pumping;
- permit a reduction in the size of distribution mains below that which would be required in the absence of a reservoir; and
- allow pumping at the average rather than peak flow rate.

CLASSIFICATION OF STORAGE REQUIREMENTS

Storage volume requirements are classified by function: operating, equalizing, fire and/or emergency, and dead-storage volumes. Engineers must consider these individual volume components in combination to determine the total volume of storage capacity that is required for any system. The total storage required is typically the sum of all these functions.

Rather than requiring both fire and emergency storage, some local fire and state agencies allow systems to use the larger of either fire or emergency storage volumes.

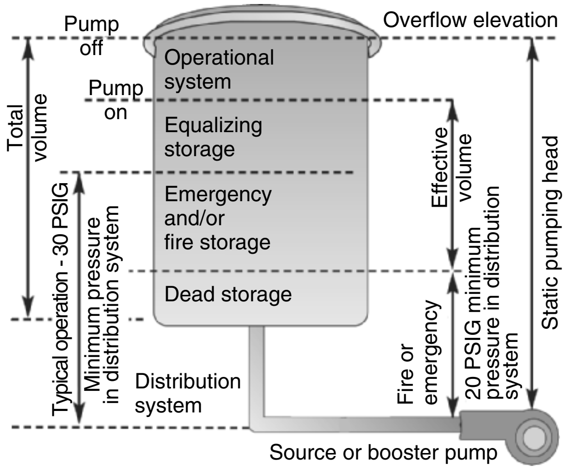

Figure 1. Storage volume classification. Source or booster pump.

These local agency requirements need to be determined on a case-by-case basis before designing a storage reservoir. Figure 1 illustrates the different components of the total storage volume.

Operating Storage—This is the difference in volume between the "pump on" and "pump off" levels when the tank is normally being used and the sources of supply pumps to the storage tank are off.

Equalizing Storage—This storage component is used when the source pump capacity is less than the peak system demands. The storage is needed so that water production facilities can operate at a relatively constant rate. Daily peak rates determine the volume, compared to the average daily demand and source capacity.

Fire Storage—The volume of water stored within the water system for fighting fires is known as "fire storage." The storage volume required varies with the size of city and with the size, type, and classification of construction within the service area. Storage volume requirements for fire fighting should be determined based upon state and local fire districts and municipalities. Fire authorities often refer to the latest edition of the *Guide for Determination of Required Fire Flow*, published by the Insurance Services Office to determine local fire flow requirements.

The typical minimum municipal fire flow requirement is 500 to 1,000 gallons per minute (gpm) for two hours for single family residential areas. Commercial and industrial areas fire flows can be as high as 8,000 gpm or more for many hours. Typical fire storage requirements are shown in Table 1. Engineers should base the actual capacity needs upon local fire flow requirements.

Emergency Storage—This storage is used to provide water to the system during other unusual or emergency conditions. Emergency storage volume depends upon the likelihood of supply interruption and the time required to make repairs or arrange for an alternative water supply.

Dead Storage—Storage in tanks or reservoirs that cannot be drawn out or used beneficially because of piping elevations or low pressures is known as dead storage. Dead storage is typically most significant in tall standpipe-type tanks where water in the bottom of the tank cannot be used because of low system pressure.

Table 1. Typical Range of Storage Requirements for Fire Protection

Type of Development	Storage Volume, Gallons	Storage Volume, ML
Low-density residential, 2 hr at 500 gpm	60,000	0.23
Built-up residential, 2 hr at 1,000 gpm	120,000	0.45
Light commercial, 4 hr at 2,000 gpm	480,000	1.8
Commercial, 4 hr at 4,000 gpm	960,000	3.6

Storage Volume Dedicated to Contact Time—Finished water stored in clearwells at water treatment plants is sometimes used to meet the disinfection contact time. In these cases, the amount of fixed volume used to meet the disinfection requirement should also be considered in the total volume and operational limits of the reservoir.

DAILY STORAGE VOLUME USE VARIES

Water use is greater during daylight hours—typically peaking in the mid-morning and early-evening hours. Stored water is withdrawn during these peak demand hours of the day and is replenished during minimum-demand times in the late-night and early-morning hours.

Figure 2 illustrates the hourly variation in daily water use (diurnal variation) that might occur in a typical residential community on the day of maximum water use for the year.

The shape of the diurnal curve of water demand will vary significantly between different cities because of differences in climates and local economies. Local design data should be obtained for each water system to determine storage needs. However, with an adequate source of supply capacity, equalizing storage of approximately 22 percent of the maximum daily demand is typical for small residential areas.

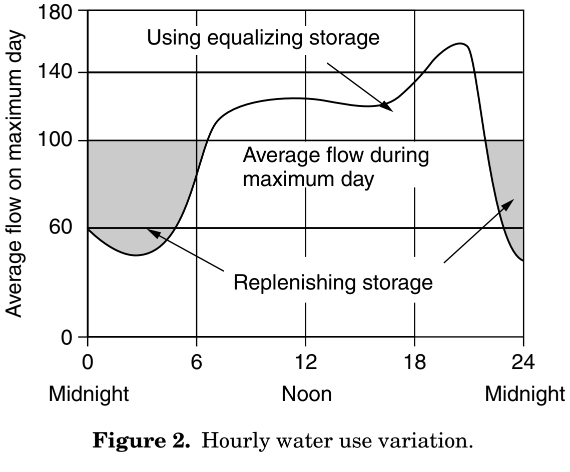

Figure 2. Hourly water use variation.

STORAGE TANK SHAPE AND VOLUME

Water towers can be made of concrete or steel and can take various forms. The most suitable form for concrete towers is a cylinder with a curved shaped bottom or with a flat bottom. Steel tanks may have a spherical or dome shaped bottom. The shape chosen is usually a compromise between function, construction and maintenance costs, and aesthetics.

The lowest water level in the tank is determined according to the pressure requirements in the pipeline. The pressure in the pipelines may vary depending on the type of community and pressure needs of different areas in a city. Typically, minimum acceptable water system pressures are 35 to 40 pounds per square inch (psi) and maximum pressures are 100 to 120 psi.

To keep pumping costs low, water depth in the tank is generally kept small. Due to structural considerations, the depth is kept equal to the diameter.

LOCATING STORAGE RESERVOIRS

A service reservoir stores the water and supplies it at the required pressure to the farthest point in the area. In view of the cost of pipelines and uniform pressure distribution, the reservoir should be located near the center of the service area.

In flat areas, it is relatively easy to build the water tower at the center. In hilly areas, however, it may be more advantageous to select the highest point for the construction of an elevated tank, which may lie at one end of the area instead of the center.

Apart from the center, the tank or tower can be situated between the area and the source of supply (pumping or gravity flow). When the service reservoir lies between the area and the source, all the water must pass through the elevated tank before flowing through the area (see Fig. 3a).

The pressure in the water supply system depends upon the water level in the service reservoir. A water supply system needs to guarantee a minimum pressure even at the most remote point in the area. Therefore, it is essential that the hydraulic gradient line always be above the required pressure.

When water is supplied from an impounded high-level reservoir, the service reservoir may function as a pressure-reducing device (see Fig. 3b). This reduces the possibility of damage to the pipes due to high hydrostatic pressure.

When the area lies between the source and the service reservoir, then most of the requirements are met by direct pumping and the excess water flows to the service reservoir (see Figs. 3c and 3d). In this system there may be larger fluctuations in the supply pressure.

CONSTRUCTION MATERIALS

Most states permit steel and concrete construction materials. All piping, joints, and fittings should conform to American Water Works Association (AWWA) specifications. Welded steel water tanks should confirm to American National Standards Institute (ANSI)/AWWA Standard D100. Factory-coated bolted steel tanks should confirm to ANSI/AWWA D103. Wire and strand wound, circular, prestressed concrete tanks should confirm to ANSI/AWWA Standard D110.

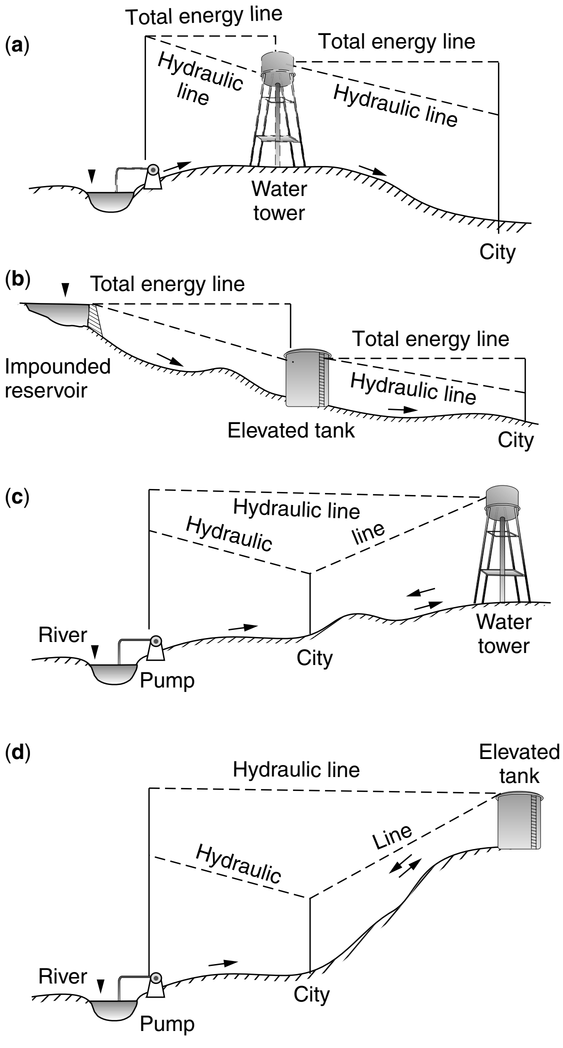

Figure 3. Service reservoir.

The storage tanks should be painted or have cathodic protection. The AWWA standards for painting exclude the use of paints that might add toxic materials to the stored water. The paint, both external and internal, should comply with the standards prescribed by AWWA D101 and D102.

OTHER CONSIDERATIONS

All water reservoirs should be covered to protect the stored water against contamination. Overflow pipes should be brought down near the ground surface and discharged to minimize erosion. The storage structure should be designed so that there is water circulates. There should be a convenient access to the interior for cleaning, maintenance, and sampling. Rigid storage reservoirs should be vented.

READING LIST

Pizzi, NicholasG 1995. *Hoover's Water Supply and Treatment*, 12th Edn. National Lime Association: Arlington, VA.

HDR Engineering, Inc. 2001. *Handbook of Public Water Systems*. 2nd Edn. John Wiley & Sons, Inc.: New York, New York.

Hobbs, AubreyThomas (ed.) 1969. *Manual of British Water Engineering Practice*. Vol 2. W. Heffer and Sons, Ltd: Cambridge, England.

Middle Brooks, EJoe, et al. 1977. *Water Supply Engineering Design*. Ann Arbor Science Publishers: Ann Arbor, MI.

For more information, comments about this fact sheet, or to suggest topics for future *Tech Briefs*, contact Bhardwaj at (800) 624-8301 or via e-mail at vbhardw2@wvu.edu.

Vipin Bhardwaj, technical assistance specialist, received his bachelor's degree in chemical engineering and has recently finished his master's degree in agriculture at West Virginia University.

HAVE YOU READ ALL OUR TECH BRIEFS?

Tech Briefs, *drinking water treatment fact sheets have been a regular feature in the National Drinking Water Clearinghouse (NDWC) newsletter* On Tap *for more than five years*.

Tech Brief: Disinfection, item #DWBLPE47;
Tech Brief: Filtration, item #DWBLPE50;
Tech Brief: Corrosion Control, item #DWBLPE52;
Tech Brief: Ion Exchange and Demineralization, item #DWBLPE56;
Tech Brief: Organics Removal, item #DWBLPE59;
Tech Brief: Package Plants, item #DWBLPE63;
Tech Brief: Water Treatment Plant Residuals Management, item #DWBLPE65;
Tech Brief: Lime Softening, item #DWBLPE67;
Tech Brief: Iron and Manganese Removal, item #DWBLPE70;
Water Conservation Measures Fact Sheet, item #DWBLPE74;
Tech Brief: Membrane Filtration, item #DWBLPE81;
Tech Brief: Treatment Technologies for Small Drinking Water Systems, item #DWPSPE82;
Tech Brief: Ozone, item #DWBLPE84;
Tech Brief: Radionuclides, item #DWBLPE84;
Tech Brief: Slow Sand Filtration, item #DWBLPE99;
Tech Brief: Ultraviolet Disinfection, item #DWBLPE101;
Tech Brief: Leak Detection and Water Loss Control, item #DWBLPE102;
Tech Brief: Diatomaceous Earth Filtration for Drinking Water, item #DWBLPE108.

A package of the *Tech Briefs* is now available as a product. A three-ring binder holds all the current *Tech Briefs* in print. New selections can be easily added to the package as they become available. To order this product, call the NDWC at the numbers listed below and ask for item #DWPKPE71.

The item is free.

Additional copies of fact sheets are also free; however, postal charges may be added.

WATER TREATMENT PLANT RESIDUALS MANAGEMENT

National Drinking Water
Clearinghouse

Water treatment plants produce a wide variety of waste products as well as safe drinking water. These residuals may be organic and inorganic compounds in liquid, solid, and gaseous forms depending on the source of raw water and the type of treatment processes, commonly, coagulation/filtration, precipitative softening plant, membrane separation, ion exchange, granular activated carbon.

The differences between the unit processes of the five plant types listed above characterize the type of residuals generated at a given facility. In the current regulatory climate, a complete management program for a water treatment facility should include the development of a cost-effective plan to remove and dispose of residuals. The following steps need to be considered when developing a comprehensive water treatment residuals management plan:

- Characterize form, quantity, and quality of the residuals;
- Determine appropriate regulatory requirements;
- Identify feasible disposal options;
- Select appropriate residuals processing/treatment technologies; and
- Develop a residuals management strategy that meets both the economic and noneconomic goals established for a water treatment facility.

WHAT ARE THE RESIDUALS CATEGORIES?

Water treatment plant residuals form when suspended solids in the raw water react with chemicals (e.g., coagulants) added in the treatment processes and associated process control chemicals (e.g., lime). Some potable water treatment processes generate residuals that are relatively easy to process and dispose of. For example, leaves, limbs, logs, plastic bottles, and other large floating debris separated from water during the initial screening process can be disposed of at conventional solid waste landfills. However, most other treatment processes produce more complex residual waste streams that may require advanced processing and disposal methods to protect human health and the environment.

The four major types of residuals produced from water treatment processes are:

Major Treatment Processes and Types of Residual Solids Treatment Processes (Robinson and Witko, 1991)

Typical Residual Waste Streams Generated	Typical Contaminant Categories	Typical Disposal Methods	Regulation Covering Disposal Method
Coagulation/Filtration			
Aluminum hydroxide, ferric hydroxide, or polyaluminum chloride sludge with raw water suspended solids, polymer and natural organic matter (sedimentation basin residuals)	Metals, suspended solids, organics, radionuclides, biological, inorganics	Landfilling Disposal to sanitary sewer/WWTP Land application Surface discharge	RCRA/CERCLA State and local regulations RCRA, DOT NPDES (CWA), state and local DOH
Spent backwash filter-to-waste	Metals, organics, suspended solids, biological, radionuclides, inorganics	Recycle Surface discharge (pumping, disinfection, dechlorination) Disposal to sanitary sewer/WWTP	State and local DOH NPDES (CWA), state and local regulations State and local regulations
Precipitative Softening			
Calcium carbonate and magnesium hydroxide sludge with raw water suspended solids and natural organic matter	Metals, suspended solids organics, unreacted lime, radionuclides	Landfilling Disposal to sanitary sewer/WWTP Land application	RCRA/CERCLA, state and local regulations State and local regulations RCRA, state and local regulations, DOT
Spent backwash filter-to-waste	Metals, organics, suspended solids, biological, radionuclides, inorganics	Recycle Surface Discharge (pumping, disinfection, dechlorination) Disposal to sanitary sewer/WWTP	State and local DOH NPDES (CWA), state and local regulations State and local regulations
Membrane Separation			
Reject streams containing raw water suspended solids (microfiltration), raw water natural organics (nanofiltration), and brine (hyperfiltration, RO)	Metals, radionuclides, TDS, high molecular weight contaminants, nitrates	Surface discharge (pumping, etc.) Deep well injection Discharge to sanitary sewer/WWTP Radioactive storage	RCRA, NPDES, state and local regulations RCRA, NPDES, state and local regulations State and local regulations RCRA, DOT, DOE
Ion Exchange			
Brine stream	Metals, TDS, hardness nitrates	Surface discharge Evaporation ponds Discharge to sanitary sewer/WWTP	RCRA, NPDES, state and local regulations RCRA, NPDES, state and local regulations State and local regulations
Granular Activated Carbon			
Spent GAC requiring disposal and/or reactivation, spent backwash, and gas-phase emissions in reactivation systems	VOCs, SOCs (nonvolatile pesticides), radionuclides, heavy metals	Landfill Regeneration (on/off site) Incineration Radioactive storage Return spent GAC to supplier	RCRA, CERCLA, DOT State and local air quality regulations (CAA) State and local air quality regulations (CAA) DOT, DOE
Stripping Process (Mechanical or Packed Tower)			
Gas phase emissions	VOCs, SOCs, radon	Discharge to atmosphere GAC adsorption of off-gas (contaminant type and concentration dependent)	State and local air quality regulations (CAA)
Spent GAC if used for gas-phase control	VOCs, SOCs radionuclides	GAC adsorption of off-gas (contaminant type and concentration dependent) Return spent to GAC to supplier	State and local air quality regulations (CAA)

Chart Key

CAA = Clean Air Act
CWA = Clean Water Act
CERCLA = Comprehensive Environmental Response, Compensation and Liability Act
DOE = Department of Energy
DOH = Department of Health
DOT = Department of Transportation
GAC = Granular Activated Carbon
NPDES = National Pollutant Discharge Elimination System
RCRA = Resource Conservation and Recovery Act
RO = Reverse Osmosis
SOC = Synthetic Organic Chemical
TDS = Total Dissolved Solids
VOC = Volatile Organic Compound
WWTP = Wastewater Treatment Plant

- *Sludges* (i.e., water that contains suspended solids from the source water and the reaction products of chemicals added in the treatment process). Presedimentation, coagulation, filter backwashing operations, lime soften-ing, iron and manganese removal, and slow sand and diato-maceous earth filtration all produce sludge.
- *Concentrate* (brines) from ion exchange regeneration and salt water conversion, membrane reject water and spent backwash, and activated alumina waste regenerant.
- *Ion exchange resins,* spent granular activated carbon (GAC), and spent filter media (including sand, coal, or diatomaceous earth from filtration plants).
- *Air emissions* (off-gases from air stripping, odor control units, or ozone destruction).

The chemical characteristics and contaminant concentration levels in these residual waste streams often impose the ultimate disposal options. Furthermore, it is reasonable to expect that as drinking water quality is increasingly regulated, higher removal efficiencies of more contaminants will be required. To achieve these higher efficiencies, water treatment plants (WTPs) will need to use more sophisticated treatment technologies. Of particular concern are cases in which residuals are characterized as either hazardous or radioactive waste. Depending on the raw water quality and treatment process removal efficiency, hazardous or radioactive characteristics could be exhibited in potentially any residual waste stream mentioned on page 1.

WHAT REGULATIONS GOVERN MANAGEMENT OF WATER TREATMENT PLANT RESIDUALS?

Identifying the regulations that affect various management practices may be difficult for water treatment utility managers. The difficulty is due to the many different types of wastes produced by WTPs and various types of waste disposal: direct discharge, discharge to wastewater treatment plants, disposal in landfills, land application, underground injection, disposal of radioactive waste, and treatment of air emissions.

At the federal level, the U.S. Environmental Protection Agency (EPA) has not established any regulations that are specifically directed at WTP residuals. Applicable regulations are those associated with the Clean Water Act (CWA); Criteria for Classification of Solid Waste Disposal Facilities and Practices (40 CFR, Part 257); the Resource Conservation and Recovery Act (RCRA); the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) or Superfund; and the Clean Air Act (CAA). The CWA limits direct discharges into a water course while the other legislation governs other methods of use and/or disposal of wastes. Most states are responsible for establishing and administering regulations that will meet the requirements of these acts. The regulation of wastes, therefore, is the responsibility of the states.

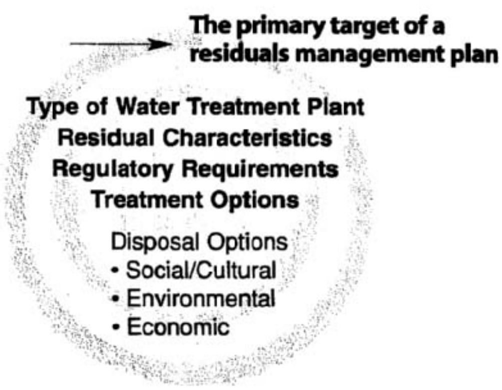

Figure 1. *Source*: Management of Water Plant Residuals.

HOW ARE RESIDUALS MANAGEMENT PLANS SELECTED?

To develop a residuals management plan, the WTP manager may start with a broad array of residuals processing and disposal options that can be narrowed through considering specific residuals characteristics and associated regulatory requirements. A focus on available disposal options further narrows the management alternatives.

Figure 1 illustrates the need for practical disposal options and treatment processes that will take into account economic and noneconomic factors of concerns to the community. The technical criteria used to select the final management plan differ from user to user; economic, cultural, social, and environmental factors are also site-specific, and are typically included in any final selection.

WHERE CAN I FIND MORE INFORMATION?

Information in this fact sheet was primarily obtained from:

(1) Management of Water Treatment Plant Residuals. American Society of Civil Engineers (ASCE) Manuals and Reports on Engineering Practice No. 88, American Water Works Association (AWWA) Technology Transfer Handbook, and U.S. EPA 625/R-95/008.

(2) Handbook of Practice: Water Treatment Plant Waste Management. American Water Works Association (1987). [This book is out of print but available in libraries.]

(3) Robinson, M.P., and J.B. Wiko. 1991.“Overview of Issues and Current State-of-the Art Water Treatment Plant Waste Management Programs.” 1991 Annual Conference Proceedings. AWWA Quality for the New Decade, Philadelphia, PA. June 23–27. [This book is out of print; however, copies of the article are avail able for a fee from AWWA.]

Management of Water Treatment Plant Residuals can be ordered from the AWWA. For more information, write the AWWA at 6666 W. Quincy Ave., Denver, CO 80235, or call (800) 926–7337. You may also view a selection of AWWA literature on their Web site at www.awwa.org/store.htm.

REVERSE OSMOSIS, PROCESS CHEMISTRY

Robert Y. Ning
King Lee Technologies
San Diego, California

INTRODUCTION

The application of membranes to purify water provides a powerful means of producing high-purity water. In the total treatment of raw water, reverse osmosis (RO) is an important midstream process for fractionating feedwater into high-purity water and brine. The brine may be used to feed cooling towers and evaporators.

The RO process is sensitive to upstream treatments and itself impacts downstream operations. The almost infinite variety of contaminants that enter the membrane systems with the feedwater and the economic necessity to expend minimal resources on raw water pretreatment pose a major challenge to RO process chemistry design and control. RO process chemistry itself is important in understanding and controlling the quality of the product permeate water which may be further polished to remove trace contaminants to meet the ultrapure specifications needed in industries such as pharmaceuticals, microelectronics, and power generation. Presented here are the chemistry of feedwater membrane fouling mechanisms, composition of foulants, and chemicals used to prevent fouling.

FEEDWATER CHEMISTRY

Water that enters reverse osmosis systems comes from all imaginable sources. Presented in Table 1 is a list of sources that are differentiable by their quality and the significant chemistry in these waters that impacts the performance of RO systems.

Major physical occurrences in natural waters are the processes of erosion of the rocks that generate sand (50 microns −2 mm), silt (5–10 microns), clay (<5 microns), and submicron colloidal particles. Just as active a physical process is the growth, excretion, death, and decay of animals and microorganisms.

Table 1. Sources of Feedwater and Significant Chemistry that Impacts RO

Source Water	Significant Chemistry
1. Rain	Oxidation–reduction (e.g., $O_2 + NO_x$), acid–base (e.g., $CO_2 + H_2O$)
2. Springs	Dissolution, acid–base (acid rain + carbonate rocks)
3. Ponds, Lakes	Complex interactions of living and nonliving matter
4. Wells, aquifers	Dissolution–precipitation (e.g., $CaCO_3$, $CaSO_4$, $SrSO_4$, $BaSO_4$, silicates, iron and aluminum salts, sulfides), relatively low in organic matter
5. Rivers, municipal and industrial wastewater	Complexation interactions, precipitation
6. Oceans (99.4% of the earth's water including 2% as ice)	Biotic life, complexation interactions, precipitation

Some significant chemical processes occurring in these waters can be enumerated. Rain extracts oxygen, oxides of carbon, sulfur and nitrogen, and organic aerosol particles from the atmosphere. Spring water dissolves soluble matter, and the acidity of rainwater is neutralized by alkaline rocks such as limestone. Ponds and lakes are biotic reactors in which biomass is produced and accumulated, and complexation interaction of organic matter with inorganic particles occurs. Much of this particulate matter is retained in upper zones of the ground as water enters aquifers. The large surface areas of contact and the long times available for dissolution and precipitation allow water in aquifers to be equilibrated and fully saturated with respect to the composition of the minerals they are in contact with. Thus it is from the deep well source water that the greatest scaling potentials from foulants such as calcium carbonate, calcium sulfate, strontium sulfate, barium sulfate, silica, iron, and aluminum exist. We can compare the water in the lower sections of rivers to municipal and industrial waste streams in which complexation interactions of organic and inorganic particles predominate.

The chemistry of seawater is particularly interesting to note. It reflects the equilibrated end point of the processes that began in the surface waters just mentioned. Interestingly, the chemistry in the water undergoing concentration within RO systems models the transition of surface waters to highly saline seawater with respect to coagulation of colloidal matter. Heavy precipitation of biomass coagulated with silt and clay occurs in the delta region of major rivers. The quantity of the different components in seawater is not proportional to the quantity of the components that river water pours into the sea but is inversely proportional to the facility with which the components in seawater are insolubilized by chemical actions in the sea (1). Complexation interactions, precipitation of particles, and crystallization of salts of low solubility that began in ponds and lakes continue in the ocean.

Some of the types of colloids that exist in natural water, especially in the sea, are listed in Table 2. Polyphenolic complexes such as humic acid, lignin, and tannin are decay

Table 2. Colloidal Matter in Natural Water

- Microorganisms
- Biological debris (plant and animal)
- Polysaccharides (gums, slime, plankton, fibrils)
- Lipoproteins (secretions)
- Clay (hydrous aluminum and iron silicates)
- Silt
- Oils
- Kerogen (aged polysaccharides, marine snow)
- Humic acid/lignins/tannins
- Iron and manganese oxides
- Calcium carbonate
- Sulfur and sulfides

products of woody tissues of plants that are particularly noticeable in surface waters due to their tendency to foul RO membranes by themselves or coagulate with other foulants (2). Polysaccharides that constitute cell walls of microorganisms and plants and are excreted variously as gums, slimes, and biofilms are prominent in membrane fouling. In the ocean, they are added to polysaccharides from sea organisms, especially planktons (3), which become kerogen upon aging. It is believed that white materials called "marine snow" that has been observed falling in the deep oceans are these aged and ubiquitous polysaccharides (4).

MEMBRANE FOULING MECHANISMS

Feedwaters to RO systems typically are concentrated by a factor of 2–10 (50–90% recovery) during production of permeate water. For simplicity in classification, three classes of fouling (5–11) can be said to occur.

First, there is scaling. The solubility limits of various dissolved salts in the brine stream may be exceeded, leading to deposition and growth of crystals in the flow channels and membrane surface of the RO elements. This type of fouling is referred to as scaling. Most common examples of scales are calcium carbonate; sulfates of calcium, strontium, and barium; and calcium fluoride. In the examination of foulants (2) under the microscope, crystals have well defined shapes. Inhibitors injected into feedwaters to suppress crystallization are called antiscalants.

Second, there is colloidal fouling. Particulate matter that preexists in feedwater can aggregate and adhere to the membrane and brine flow channels due to increased concentration, salinity, compaction, flocculation, surface interactions, and other physical and chemical factors. The particles, such as silt, carbon fines, coagula, and microbial clusters, may be large enough to be removed by 1 to 5 micron RO prefilters, or they may be colloidal particles that escape filtration even with the standard 0.45-micron filters used in silt density index measurements. The colloids can be organic or inorganic or composites. Ferric, aluminum, and calcium hydroxides and silicic acid grow to increasing particle sizes by polymerization. Cross-linking and complexation of organic and inorganic polymers produce gels and amorphous foulants commonly seen on membranes (2,12). Biotic debris such as polysaccharides and dead cellular matter contribute largely to this type of foulant. We have found that anticoagulants and antideposition agents show promise in inhibiting this fouling process. Antifoulants then can be a term applied specifically in this particulate fouling context. In a broader sense, the term antifoulant is used in the field to include antiscalants as well (6).

Third, biofouling is a prominent source of fouling. True of all water treatment and distribution systems is the growth and anchoring of microorganisms. At times, moderate temperatures and minimal nutrient levels in RO waters can support explosive growth of microorganisms. Bacteria capable of cell division in 20 minutes can grow from a normal count per unit volume of water to millions in an 8-hour shift. Due to the tendency of bacteria to secrete polymers that anchor themselves to surfaces to facilitate growth as a biofilm, this fouling mechanism is unique and poses a serious threat to operation of RO systems (13). This threat is compounded by the great difficulty of treating and completely removing a biofilm from membrane surfaces.

COMPOSITION OF FOULANTS

The chemical composition of foulants correlated with their sources are given in the Reverse Osmosis, Membrane Foulants.

ANTIFOULANT DESIGN AND APPLICATION

The term antifoulant is used here in its broadest meaning covering scaling, particulate fouling, and microbial fouling—the three classes of fouling mechanisms discussed before. Strategies aimed at controlling each type of fouling are summarized here.

For scale control, the development and application of antiscalants is well known and reviewed in the field of boiling water and cooling water chemistry (6,14–16) and applied to boilers, evaporators, cooling towers, and cooling systems. Anionic polymers, polyphosphates, and organophosphorous compounds, sometimes referred to as threshold inhibitors and dispersants, are used in substoichiometric amounts, usually in the range of 1–5 mg/L in RO systems (17). By binding to surfaces of growing crystal nuclei, the rates of crystallization from supersaturated solutions are retarded, and crystal packing orders are modified. By this mechanism, crystallization rates are so retarded that although supersaturation of solutes in the water will eventually equilibrate through crystallization within the residence time of the water in the system, there is little or no scale formation. The uniqueness of RO among water conditioning systems is that residence time is very short (a few seconds), concentration of seed crystals is low, and temperature is constant. For this reason, higher levels of supersaturation without crystallization are possible. On the other hand, the limits of saturation and rates of scaling are hard to model, measure, and predict. Interference comes from other solutes in the water, organic or inorganic. Assumptions of RO fouling limits vary considerably among practitioners.

Controlling fouling by preexisting particulate matter is much more challenging due to the variety of types of potential foulants and the complexity of their interactions (1,2,8,10,12), with each other in the same water and with the membrane. The stability and agglomeration of colloidal particles is a subject of major importance in natural water (1) as well as in treating process water (16). Drawing on the basic science of colloids and testing of model foulants suggested by RO foulant analysis data (2), progress is made steadily by the development of antifoulants.

The literature on preventing and managing biofouling in water treatment systems, is extensive. Much of the art and science found useful generally is applicable to RO systems as well. Several factors peculiar to the RO system can be mentioned. Chemicals used to sanitize and clean the system have to be chemically compatible with the thin, salt-rejecting, polyamide barrier membrane. It is of prime concern that accumulation and exponential growth of microorganisms should not be allowed to occur within the system. Pretreatment of feedwater, adequate maintenance of upstream unit operations, continuous flow of water through the RO unit, a good monitoring and sanitization program, and use of preservatives during downtime (18) are important to this end. Normalized flow and differential pressure in the system are sensitive indicators of biofouling.

To prevent irreversible fouling, trend charting of normalized permeate flow, differential pressure, and salt rejection, and readiness to perform adequate cleaning are important aspects of system operation (see Reverse Osmosis, Membrane Cleaning). Immediate assistance should be sought when cleaning has been inadequate. Powerful new cleaners are now available that provide alternatives to membrane replacement in the worst cases of fouling.

BIBLIOGRAPHY

1. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry*, 3rd Edn. John Wiley & Sons, New York, NY, p. 3.
2. Ning, R.Y. and Shen, P.T.L. (1998). Observations from analysis of reverse osmosis membrane foulants. *Ultrapure Water* **15**(4): 37–44.
3. Aluwihare, L.I., Repeta, D.J., and Chen, R.F. (1997). A major biopolymeric component to dissolved organic carbon in surface sea water. *Nature* **387**: 166–169.
4. Ritter, S.K. (1997). Chemistry on the high seas. *Chemical & Engineering News, Amer. Chem. Soc.* (October 27): 23–31.
5. Byrne, W. (1995). Chapter 3, *Reverse Osmosis*. Tall Oaks Publishing Inc., Littleton, CO, pp. 97–166.
6. Meltzer, T.H. (1993). *High Purity Water Preparation*. Tall Oaks Publishing Inc., Littleton, CO, pp. 485–487.
7. Potts, D.E., Ahlert, R.C., and Wang, S.S. (1981). A critical review of fouling of reverse osmosis membranes. *Desalination* **36**: 235–264.
8. Cohen, R.D. and Probstein, R.F. (1986). Colloidal fouling of reverse osmosis membranes. *J. Colloid Interface Sci.* **114**: 194–207.
9. Amjad, Z. (1987). Advances in scaling and deposit control for RO system. *Ultrapure Water* **4**(6): 34–38.
10. Zhu, X. and Elimelech, M. (1995). Fouling of reverse osmosis membranes by aluminum oxide colloids. *J. Environmental Engineering* **121**(12): 884–892.
11. Tracey, D. (1996). Membrane fouling—What is it? Where does it come from? and What does it mean? *Ultrapure Water* **13**(7): 47–53.
12. Ning, R.Y. and Stith, D. (1997). The iron, silica and organic polymer triangle. *Ultrapure Water* **14**(3): 30–33.
13. Ridgway, H.F., Rigby, M.G., and Argo, D.G. (1985). Bacterial adhesion and fouling of reverse osmosis membranes. *J. AWWA* **77**: 97–106.
14. Drew. (1994). *Principles of Industrial Water Treatment*. Ashland Chemical Co., Boonton, NJ, pp. 75–85; 183–186.
15. Betz. (1991). *Handbook of Industrial Water Conditioning*. Betz Laboratories, Trevose, PA, pp. 96–104; 181–185.
16. Kemmer, F.N. (Ed.). (1988). *The Nalco Handbook*, 2nd Edn. McGraw-Hill, New York, NY, pp. 3.14–3.15.
17. Darton, E.G. (2000). Membrane chemical research: centuries apart. *Desalination* **132**: 121–131.
18. Varnava, W., Silbernagel, M., Kuepper, T., and Miller, M. (1996). *Reverse osmosis element preservation study*, Proceedings of Biennial Conf., Amer. Desalting Assoc., pp. 308–327.

REVERSE OSMOSIS, MEMBRANE FOULANTS

Robert Y. Ning
King Lee Technologies
San Diego, California

INTRODUCTION

Periodic cleaning of reverse osmosis (RO) membranes is generally accepted as a necessary routine maintenance operation in running water purification plants (1–3). Common foulants found in RO membrane elements include calcium carbonate, calcium sulfate, strontium sulfate, barium sulfate, calcium fluoride, calcium phosphate, ferric hydroxide, aluminum hydroxide, silica, silt, polymers from natural sources, and synthetic polymers used in pretreatment. These foulants may occur singly or in most cases, in complex combinations. Single foulants are simpler to address in prevention and dissolving (except for barium sulfate) during cleaning. Complex foulants usually resulting from coagulation of colloidal particles, offer the greatest challenge in prevention and removal during cleaning.

A large body of data and observations from troubleshooting RO systems is summarized here. Complex foulants are extracted from membrane elements sent from plants across the world that need assistance. Qualitative tests and quantitative elemental analyses of these foulants, combined with a thorough review of the pretreatment chemistry, provide insights into the nature and the source of the foulants. Experience with the pattern of responses to various cleaners and successful resolution and elimination of fouling problems confirm the conclusions regarding the source of foulants and the mechanisms of formation in each specific situation. The following generalizations have been helpful in solving fouling problems and in selecting the most appropriate cleaning chemicals:

1. Organic matter usually contributes more than 10% of the mass of foulants; the majority of the samples contain 50–100% organic matter.
2. Biomass constitutes the major part of the organic content of the foulants.
3. Synthetic polymers used in water pretreatment can often be differentiated from biomass in foulants.
4. Anionic polymers used as antiscalants and dispersants can deposit on membranes and act to concentrate high valence metal ions (Fe, Al, Ca) on the membrane surface.
5. Colloidal silica and clays have complex and variable content of iron, aluminum, calcium, and magnesium.

6. Elemental phosphorous is associated with biomass, phosphate, and phosphonates.

The six generalizations separately presented below are derived from a synthesis of foulant analysis data and correlation with plant process chemistry and responses to cleaners. Positive outcomes in the plants have consistently been realized by applying them.

ORGANIC MATTER

Organic matter is defined as carbon-containing matter. Each foulant is washed by decantation with pure water, then dried at 120–130 °C. It is gravimetrically analyzed for carbon, hydrogen, and nitrogen (C,H,N) content by standard methods of pharmaceutical analysis, along with a quantitative ash test (furnace at 400–450 °C) to determine the percentage of inorganic matter. The percent total volatile or combustible elements (C,H,N and oxygen) from the ash test is considered organic matter, and it correlated well with C,H,N values along with the general oxygen atom content of biological and synthetic organic polymers commonly used in water pretreatment. The organic matter in 22 foulant samples analyzed by these methods, during a recent period has the values shown in Table 1. From the table, it is interesting to note that all samples contained at least 10% by weight of organic matter and that 73% of the samples contained 50–100% by weight of organic matter. Note that the foulants analyzed are representative of the more difficult to clean types of foulants. Foulants such as calcium carbonate, calcium sulfate, and ferric and aluminum hydroxides that are readily cleaned with generic acid cleaners are not represented. Organic matter, which includes biomass and polymers used in pretreatment, usually makes the foulant difficult to remove.

BIOMASS

Biomass refers to organic matter arising from living or dead organisms. In colloidal form, it can enter the RO system with the feedwater and then coagulates and accumulates on membrane elements. A more severe source of biomass is the exponential growth of viable organisms within the system (4) and the accumulation of the biofilm in the system due to the difficulties in removal (5). Such accumulation is usually promoted by inadequate maintenance sanitization procedures of the entire water processing system and stagnation during frequent or prolonged downtimes without preservative. Severe growths are often detectable by odor or slime on the internal surfaces in contact with water. Milder growths are detectable by microbial culturing and enzyme activity tests. Natural surface waters (lakes, rivers and sea), especially in warm regions, have very high levels of decaying or excreted plant and animal matter. High biomass content of foulants in the samples examined correlated with the factors mentioned, in the feedwater and in the plant.

Table 1. Organic Matter in RO Foulant Samples

Organic Matter (% by weight)	Number of Foulants Analyzed	
0–10%	0	
10–20	3	
20–30	2	
30–40	0	
40–50	1	27%
50–60	3	
60–70	3	
70–80	4	
80–90	5	
90–100	1	73%
Total:	22	100%

SYNTHETIC POLYMERS

Synthetic polymers used in water pretreatment can often be differentiated from biomass in foulants. The absence of the characteristics of biomass—odor, visual and microscopic appearance, color, slime, microbial count, enzyme activity, presence of phosphates, sodium and potassium in ash—and the presence usually of a thin and even coating of colorless to tan gel suggest synthetic polymers. Severe flux reduction accompanied by an almost invisible film of organic material is the hallmark of poisoning by cationic polymers. Using Fourier transform infrared spectroscopy (FT-IR), if the reference spectrum of the suspect pretreatment polymer is available and the foulant spectrum is uncluttered, positive identification through "fingerprinting" is possible. Carryover of high molecular weight synthetic cationic, anionic and nonionic polymers used as coagulants in pretreatment often is flocculated with colloidal particles of complex compositions, carries more mass, and is less readily identifiable. Lower molecular weight polymers used as antiscalants often participate in these complexes and are deposited as foulants. Even in the absence of polymeric coagulants, low molecular weight polymers, used as antiscalants and dispersants, can and will deposit on membranes at a slower rate (see following section).

ANIONIC POLYMERIC ANTISCALANTS

Anionic polymers used as antiscalants and dispersants are known to be quite intolerant of multivalent cations such as ferric and aluminum ions. These ions act as ionic cross-linkers, causing even polymers in the molecular weight range of 1,000 to 2,000 daltons to flocculate. By a combination of gradual deposition and accumulation due to incomplete removal during routine cleaning, such foulants build up over several years and gradually concentrate high valence cations as an ion exchange resin would. As an example, the elemental composition of such a foulant is given in Table 2. This sample contains 52% volatile organic matter (C,H,N,O) and 48% inorganic ash. The carbon and hydrogen content of the foulant (23.67% and 3.73%, respectively) are in reasonable agreement with the theoretical polyacrylic acid composition (at 52%) of 26.4% carbon, 2.9% hydrogen, and 23.4% oxygen. Smaller amounts of coagulated nitrogen-containing polymers or biomass can account for the nitrogen content. The inorganic composition of the foulant is revealed by scanning electron microscopy with energy diffusive X-ray analysis (SEM-EDX). The instrument is calibrated to quantitate elements between the atomic numbers of 12–42, which correspond to magnesium to molybdenum in the periodic table. The lighter elements such as B,C,H,N,O, and F are not measured, so the percentage by weight values are relative for the range of atoms measured. As shown, the predominant cation is iron (44% relative), followed by calcium, aluminum, chromium, and magnesium. Bridging to phosphate, silicate, and sulfate anions is apparent.

COLLOIDAL SILICA AND CLAYS

Colloidal silica and clays refer to preexisting siliceous particles in the feedwater, which coagulate during passage through the reverse osmosis system and become deposited on the membrane as a foulant (1,6,7). Colloidal silica, also known as nonreactive silica, results in natural waters from the polymerization of silicic acid and reactive oligomeric species, which as a population is termed reactive silica. Reactivity is measured by and often referred to as reactivity toward molybdate ions by the same type of condensation reaction. Condensation of reactive silica with aluminum, iron, and calcium hydroxides forms clays. Clays in feedwaters result mainly from the erosion of rocks, however. It is stated (8) that the chemical composition of most clays is similar to the composition of the earth's crust: oxygen 49.9%, silicon 26.0%, aluminum 7.3%, iron 4.1%, calcium 3.2%, sodium 2.3%, potassium 2.3%, magnesium 2.1%, and all other elements 2.8%. Following erosion is sand (50 microns −2 mm), silt (5–10 microns), and clay (<5 microns), differentiated by particle sizes.

Depending on the extent of clarification and prefiltration of feedwaters, silt and clays often end up on the membranes. The silt density index (SDI) is a means of measuring particle load in the feedwater by timing the fouling of 0.45-micron test filters. In practice, feedwater with an SDI value greater than 5 would lead to high fouling rates in plants. The 0.45-micron filters are nominal filters, so it is likely that a significant fraction of clay entering with the feedwater is not detected by SDI monitoring.

Table 2. Elemental Composition of an Iron Acrylate Foulant

Component	Percent by Weight	
By gravimetric elemental analysis:		
Carbon	23.67	(absolute)
Hydrogen	3.73	
Nitrogen	3.26	
Ash (inorganic)	48.08	
By SEM-EDX (at. no. 12–42, Mg–Mo):		
Iron	44.0	(relative)
Calcium	15.0	
Phosphorous	13.0	
Silica	9.9	
Aluminum	8.8	
Sulfur	4.2	
Chromium	1.8	
Magnesium	1.7	
Total:	98.4	

Table 3. Elemental Composition of a Colloidal Silica Foulant

Component	Percent by Weight	
By gravimetric elemental analysis:		
Carbon	1.15	(absolute)
Hydrogen	0.67	
Nitrogen	0.02	
Ash (inorganic)	92.46	
By SEM-EDX (at. no. 12–42, Mg–Mo):		
Silicon	90.0	(relative)
Iron	3.6	
Aluminum	2.7	
Potassium	1.3	
Sodium	0.8	
Calcium	0.8	
Magnesium	0.6	
Sulfur	0.5	
Chlorine	0.3	
Total:	100.6	

Colloidal silica and clays are often encountered in analyses of foulants. They frequently appear in the presence of organic matter and accumulate in the flow channels of the membrane elements in significant masses. Occasionally, colloidal silica can severely foul membranes in amounts almost too minute to sample physically. The composition of one such example is given in Table 3.

PHOSPHOROUS

Phosphorous is detected in foulant samples by SEM-EDX. Relative amounts in the inorganic portions of 20 recent foulants surveyed ranged from 0–58% of the elements measured in these samples (mean = 14.6%, SD = 15.7%). Sources of phosphorous-containing matter can be correlated with a number of sources. Biomass generally contains measurable amounts of phosphates. Algal biomass in natural waters reportedly contains almost 1% elemental phosphorous (9). Phosphates, except those of sodium, potassium, and ammonium, are generally only slightly soluble in water. If phosphates are present in a feedwater to any appreciable extent, phosphate scales will likely form in a reverse osmosis system unless the water is acidified (2). Calcium phosphate has limited solubility at neutral pH and an even lower solubility at higher pH. Phosphates are widely used in agriculture and in laundry products, so it is common to find phosphate in silt and agricultural runoff. Municipal wastewater has phosphate concentrations usually in the range of 15–30 mg/L of phosphate ion (about 5–10 mg/L as phosphorous) (11). In addition to calcium phosphate, ferric and aluminum phosphates are particularly insoluble and can enter the reverse osmosis system in clay or colloidal form.

Other sources of phosphorous are found in the environment of the membrane itself. High pH membrane cleaners widely used for removing organic foulants are often formulated with trisodium phosphate and sodium tripolyphosphate (10). Incomplete cleaning is likely to deposit insoluble phosphates with residual foulants which accumulate over time. Another source is the commonly used generic antiscalant sodium hexametaphosphate (SHMP). Solutions of SHMP are hydrolytically unstable and generate phosphate ions and phosphate foulants (2). Even phosphonates can precipitate with high levels of calcium, iron, and aluminum and can participate in fouling.

CONCLUSIONS

These conclusions regarding the formation of RO foulants have been the basis of resolving many operating problems in RO plants. Hopefully, they will serve as a framework for further insights into making RO systems more stable in operation.

BIBLIOGRAPHY

1. Ebrahim, S. (1993). *Cleaning and Regeneration of Membranes in Desalination and Wastewater Applications: State-of-the-Art*. Vol. 1. Proc. IDA and WRPC World Conference, International Desalination Association, pp. 197–208.
2. Byrne, W. (1995). *Reverse Osmosis*. Tall Oaks Publishing Inc., Littleton, CO.
3. Veza, J.M. and Sadhwani, J. (1997). *Cleaning Methods for Reverse Osmosis Membranes*. Proc. IDA World Congress on Desalination and Water Reuse, Vol. 2, pp. 393–407, and references cited therein.
4. Ridgway, H.F. et al. (1984). Biofilm fouling of RO membranes—its nature and effect on treatment of water for reuse. *J. AWWA* **76**: 94–102.
5. Whittaker, C., Ridgway, H., and Olson, B.H. (1984). Evaluation of cleaning strategies for removal of biofilms from reverse-osmosis membranes. *Appl. Environ. Microbiol.* **48**(3): 395–403.
6. Ning, R.Y. and Stith, D. (1997). The iron, silica and organic polymer triangle. *Ultrapure Water* **14**(3): 30–33.
7. Zhu, X. and Elimelech, M. (1995). Fouling of reverse osmosis membranes by aluminum oxide colloids. *J. Environ. Eng.* **121**(12): 884–892, and references cited therein.
8. Kemmer, F.N. (Ed.). (1988). *The Nalco Handbook*, 2nd Edn. McGraw-Hill, New York, Section 4.9.
9. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry*, 3rd Edn. John Wiley & Sons, New York, p. 887.
10. A survey of generic high pH cleaners recommended by all major membrane manufacturers that contain one or both of these reagents.
11. Reference 8, Chapter 36.

REVERSE OSMOSIS, MEMBRANE CLEANING

Robert Y. Ning
King Lee Technologies
San Diego, California

INTRODUCTION

Reverse osmosis membrane cleaning has been extensively addressed over the years (1–7). Severe concentration of dissolved and suspended matter occurs when water is purified by using reverse osmosis (RO) membranes. Permeate recoveries of 50–90% correspond to 2–10-fold concentrations in the reject brine. Precipitation or crystallization of foulants can occur in the narrow brine flow channels within the membrane elements, leading to reduced membrane flux or increased differential pressure across the system, or both. Crystallization of salts or precipitation of colloidal matter occur more toward the back end of the system, and preexisting silt and flocculated organic matter can clog the leading elements of the system.

LIMITATIONS OF CLEANING

A practical procedure for maintenance cleaning is limited to recirculating cleaning solutions through the membrane elements. By a patented method (8) of membrane reconditioning, spiral wound elements, whose hard casing is removed, are routinely used in selecting effective cleaners for cleaning by recirculation. This allows visual inspection of the membrane surface after each cleaning test. It is apparent that, except for easily soluble foulants such as calcium carbonate or ferric and aluminum hydroxides, nearly all RO foulants are only slightly dissolved in even the best matched cleaning solutions by extensive soaking. Removal requires high tangential flow velocities and is usually partially effective especially where flow channels are clogged and large patches within the elements are inaccessible to the recirculating solution. For this reason, the need for cleaning should be minimized or completely eliminated by the new antiscalants and dispersants now available and adequate pretreatment and pilot testing of process designs. When cleaning is necessary, it should be performed at the earliest stages of fouling.

WHEN TO CLEAN

It is generally agreed among membrane manufacturers and practitioners that RO systems should be cleaned before the following performance changes are reached:

1. Loss of 10 to 15% in normalized permeate flow rate.
2. Increase of 10 to 15% in differential pressure.
3. Decrease of 1 to 2% in salt rejection.

If a cleaning procedure fails to restore system performance fully to reference start-up values, it is certain that continued use of the same cleaning procedure will lead to an accelerating decline in system performance and increased cleaning frequency. For this reason, it is important to address two issues at this point: (1) find an improved cleaning procedure, or (2) investigate possible improvement of pretreatment to avoid membrane fouling. Continue cleaning and process improvement efforts until the RO performance is stable.

CHOOSING CLEANERS

Major membrane manufacturers generally define five types of foulants for which various generic chemicals

Table 1. Cleaning Solutions Formulated from Generic Chemicals in Concentrations Recommended by Major Membrane Manufacturers[a,b]

Types of Foulants-	Acid-Soluble Foulants	Biofilm, Bacterial Biological Matter	Oils and Organic Matter	Organic and Inorganic Colloids	Silica and Silicates
Fluid Systems	1% citric acid	1% STPP (TFC only) 1% EDTA.Na$^+$ 1% TSP 1% Borax (TFC only) 0.1% Triton X-100 (CA only)			–
Hydranautics	2% citric acid 0.1% Triton X-100 or Tergitol 8, etc.	2% STPP 0.26% SDDBS pH 7.5	0.5% Na Perborate 0.1% TritonX-100 or Tergitol 8 etc.	2% STPP 0.8% EDTA.Na$^+$ 0.1% Triton X-100 or Tergitol 8, etc.	–
Filmtec	0.2% HCl or 0.5% H_3PO_4 or 2.0% citric acid or 0.2% sulfamic acid	1% EDTA.Na 0.1% NaOH to pH = 12 or following pH = 12 solutions: 0.05% Na-DDS or 0.1% STPP + 1% EDTA.Na, or 0.1% TSP + 1% EDTA.Na		0.1% NaOH 0.05% Na-DDS pH = 12	–
Desal	HCl pH = 3 or citric acid pH = 2	1% TSP 1% STPP 0.1% Na-DDS 1% EDTA.Na pH = 8–11.5			–
Permasep	0.5% HCl (pH 2.3 min) or 0.5% H_3PO_4 (pH 2.3 min) or 0.2% sulfamic acid or 2% citric acid, pH = 4 or 2% citric acid 2% Na_2EDTA NH_4OH to pH = 4 or 1% $Na_2S_2O_4$ or 1% NaHMP	1% Na_2EDTA 0.1% NaOH pH = 11 max	0.3% NaPerborate 0.25% Na-DDBS pH = 10 or 1% Na_2EDTA 1% STPP 1% TSP, pH = 11 or 2% STPP 0.25% Na-DDBS pH = 10 or 1% NaHMP	0.3% NaPerborate 0.25% Na-DDBS pH = 10 or 2% STP 0.25% Na-DDBS pH = 10 or 1% NaHMP	0.5% NaOH, pH = 11 or 1% Na_2EDTA 0.1% NaOH, pH = 11 or 0.3% Naperborate 0.25% Na-DDBS pH = 10 or 2% STPP 0.25% Na-DDBS pH = 10

[a] Reference (10).

[b] STPP = sodium tripolyphosphate; EDTA.Na = sodium salt of ethylenediaminetetraacetic acid; TSP = trisodium phosphate; Triton X-100 and Tergitol 8 = nonionic surfactants; Na-DDS = sodium dodecylsulfate; Na-DDBS = sodium dodecylbenzenesulfonic acid; Na_2EDTA = disodium salt of ethylenediaminetetraacetic acid; NaHMP = sodium hexametaphosphate.

are recommended for blending at the site where cleaning solutions are prepared. The five types of foulants are (1) acid-soluble foulants, (2) biofilm/bacterial slime/biological matter, (3) carbon-containing oils/organic matter, (4) dual organic and inorganic coagulated colloids, and (5) silica and silicates. The recommended generic chemicals and concentrations in the cleaning solutions to be prepared are given in Table 1 for the five respective types of foulants. Proprietary booster cleaners are commercially available to fortify the effectiveness of these generic cleaners that are formulated at the site. For convenience and technical support, a large variety of proprietary RO membrane cleaners are available from chemical suppliers that specialize in RO operations.

CLEANING STRATEGIES

Experience has shown that within the same class of foulants, responses to the same cleaning solution can vary considerably. Elemental analyses of foulants and cleaning studies have shown that more than one type of foulant can be present on the membrane at the same time, requiring sequential cleaning with different cleaners. Sometimes even the order of cleaners used would make a significant difference. All this is to say that the choice of cleaners and the cleaning procedure to be used is an empirical science. For a given set of conditions in a plant, cleaning efficiencies are improved by trials over time. The progress of improvement can be greatly accelerated by conducting off-line cleaning studies on single elements.

When a better cleaning method is needed in the plant, the following are the alternative strategies:

Strategy 1: The plant has a history of using generic cleaning chemicals, and modest improvement in effectiveness is needed. Consider purchase of proprietary booster cleaners.

Strategy 2: Significant cleaning improvement is needed. Look for a proprietary cleaner supplier who has cleaning expertise. Option 1:

Based on prior knowledge of the characteristics of the foulant on hand and consultation with the supplier, select a combination of cleaners for trial in the plant. Option 2: Send one to three fouled elements to a specialist for a cleaning study, foulant analysis, and a review of plant performance history and pretreatment processing. Document the findings and pilot cleaning results using a recommended improved cleaning procedure. Simultaneously, address recovery of the plant and avoidance of repeated fouling. Option 3: Send all fouled elements for off-site cleaning by a specialist.

Strategy 3: All cleaning efforts by recirculation of cleaning solutions have failed. Consider nonroutine methods such as using proprietary membrane conditioning liquids or a patented membrane reconstruction process by which the membrane bundle is unrolled, cleaned leaf by leaf, then restored with a new hard casing. Both methods are offered by King Lee Technologies (9).

ON-LINE CLEANING PROCEDURE

There are six steps in cleaning membrane elements in place in RO systems:

1. Mix cleaning solution.
2. Low flow pumping. Pump preheated cleaning solution to the vessels at a low flow rate (about half of that shown in Table 2) and low pressure to displace the process water. With the RO concentrate throttling valve completely open to minimize pressure during cleaning, use only enough pressure to compensate for the pressure drop from feed to concentrate. The pressure should be low enough that essentially no permeate is produced. Low pressure minimizes redeposition of dirt on the membrane. Dump the concentrate, as necessary, to prevent dilution of the cleaning solution.
3. Recirculate. After the process water is displaced, cleaning solution will be present in the concentrate stream. Recirculate the concentrate to the cleaning solution tank, and allow the temperature to stabilize.
4. Soak. Turn the pump off, and allow the elements to soak. Sometimes, a soaking period of about 1 hour is sufficient. For difficult to clean foulants, an extended overnight soaking period of 10–15 hours is beneficial. To maintain a high temperature during an extended soaking period, use a slow recirculation rate (about 10% of that shown in Table 2).
5. High flow pumping. Feed the cleaning solution at the rates shown in Table 2 for 30–60 minutes. The high cross-flow rate flushes out the foulants removed from the membrane surface by the cleaning; minimal or no permeation through the membrane avoids compacting the foulant. If the elements are heavily fouled (which should not be a normal occurrence), a flow rate 50% higher than that shown in Table 2 may aid cleaning. At higher flow rates, excessive pressure drop may be a problem. The maximum recommended pressure drops are 20 psi per element or 60 psi per multielement vessel, whichever value is more limiting.

   Note: In this cleaning mode, foulants are generally partially dissolved in the cleaner and partially dislodged physically from the membrane and flow channels without dissolving. An in-line filter removes the recirculated particles and should be monitored for cartridge replacement.
6. Flush out. Prefiltered raw water can be used for flushing out the cleaning solution, unless there will be corrosion problems such as seawater corroding stainless steel piping. To prevent precipitation, the minimum flush temperature is 20°C.

   Additional Notes: The pH should be monitored during acid cleaning. The acid is consumed when it dissolves alkaline scales. If the pH increases more than 0.5 pH units, add more acid.

Table 2. Recommended High Recirculation Flow Rates During Cleaning

Feed Pressure,[a] psig	Element Diameter, inches	Feed Flow Rate per Vessel (gpm)
20–60	2.5	3–5
20–60	4	8–10
20–60	6	16–20
20–60	8	30–40

[a] Dependent on the number of elements in the pressure vessel.

MULTISTAGE SYSTEMS

For tapered multistage systems, the flushing and soaking steps can be performed simultaneously in the entire array. The high flow rate recirculation step, however, should be carried out separately for each stage, so that the flow rate is not too low in the first stage and too high in the last. This can be accomplished either by using one cleaning pump and operating one stage at a time or by using a separate cleaning pump for each stage.

CONTROL AND IMPROVEMENT OF CLEANING PROCESS

To ensure complete recovery of membrane performance by cleaning, the system performance should be adequately controlled by trending of normalized flux, differential pressure, and salt rejection to (1) trigger cleaning when any monitored parameters change from normal baseline by 10–15%, (2) record the trended parameters before and after each cleaning, and (3) initiate improvement actions for better cleaning if membrane performance does not fully recover. A change in responsiveness to a previously effective cleaning process signals a change in the fouling pattern that requires immediate attention.

If partial cleanings are allowed to continue, the system performance will decline at an increasing rate and will become increasingly difficult to recover.

In-place cleaning processes are improved primarily by the choice of cleaning chemicals and the order of the application sequence. Depending on the composition of the complex foulants, when two or more cleaners are found necessary, often the order in which they are used is important. Also critical, but to lesser extents, are the variables of time, temperature, and cross-flow rate.

Through a thorough review of the water and pretreatment chemistry, analyses of the foulant composition and source, and customized selection of antiscalants, dispersants, and high performance cleaners, both fouling avoidance and reliable plant performance can be attained.

BIBLIOGRAPHY

1. Veza, J.M. and Sadhwani, J. (1997). *Proceedings IDA World Congress* **2**: 393.
2. Amjad, Z., Pugh, J., and Zuhl, R.W. (1996). *Ultrapure Water* **13**(7): 27.
3. Byrne, W. (1995). *Reverse Osmosis*. Tall Oaks Publishing, Littleton, CO.
4. Ebrahim, S. (1994). *Desalination* **96**: 225.
5. Hickman, C.E. (1991). *Proceedings Membrane Proc. Conf., Amer. Water Works Assoc.*, p. 329.
6. Tragardh, G. (1989). *Desalination* **71**: 325.
7. Whittaker, C., Ridgway, H., and Olson, B.H. (1984). *Appl. Environ. Microbiol.* **48**: 395.
8. U.S. Patent No. 5,250,118 (1993). C.L. Netwig and D.L. Kronmiller. Method of Removing Foulants and Restoring Production of Spiral-Wound Reverse Osmosis Cartridges
9. *Membrane conditioners and the reconstruction service are unique from King Lee Technologies*, San Diego, CA (www.kingleetech.com).
10. *Technical Bulletins from several major membrane manufacturers provide guidance on membrane cleaning procedures and generic cleaning chemicals: Koch-Fluid Systems and Hydranautics*, San Diego, CA; Dow-Filmtec, Midland, MI; Desal/Osmonics, Minneapolis, MN; and DuPont-Permasep, Wilmington, DE.

APPLICATION OF RISK ASSESSMENTS IN CRAFTING DRINKING WATER REGULATIONS

Bruce A. Macler
Toxicologist, U.S.
Environmental Protection
Agency
San Francisco, California

INTRODUCTION

Assessments of adverse health outcomes and estimations of health risks from contaminants of drinking water are used at several points in the overall development of U.S. drinking water regulations. They are applied during the extensive discussions surrounding consideration of a contaminant's risk and approaches to risk mitigation, and in the required elements of National Primary Drinking Water Regulation (NPDWR) proposals, support documents, and final rules. Risk assessments are of several types with substantially different intents and characteristics. No one approach fits all applications. The type of assessment chosen depends on the application and the nature and availability of relevant data.

The regulatory process starts with the identification of a constituent as either a possible contaminant of drinking water or as having some adverse toxicological or pathogenic properties. Available health data are then collected and analyzed, leading to the possible development of a toxicological profile of adverse health outcomes and dose–response characteristics. This is then coupled with occurrence data and exposure estimates to provide an estimate of national public health impact. If a decision is made to move forward on a NPDWR, more refined assessments are then used to develop maximum contaminant level goals (MCLGs). Separate assessments are used to help identify possible maximum contaminant levels (MCLs) or treatment techniques consistent with public policies. In some cases, risk assessments may comprise regulatory elements themselves, such as sanitary surveys or comprehensive performance evaluations. Finally, risk assessments are necessary as formal components of regulatory supporting documents to help estimate risk reductions and quantify public health benefits. Some of these assessments, particularly those associated with public health policies and legal requirements, are highly constrained in form and content. Additionally, while the art and science of environmental risk assessment continue to evolve, we still have substantial uncertainty and imprecision in these estimates. Effective interpretation of these risk products requires some understanding of their purpose and form. This chapter details these applications and presents some examples.

Disclaimer: The views expressed in this chapter are those of the author and do not necessarily represent those of the USEPA.

RISK ASSESSMENT APPROACHES FOR DRINKING WATER REGULATIONS

In risk management activities to ensure the safety of drinking water, assessments of health risk are performed to answer questions posed in the management process (Fig. 1). Examples include broad questions such as "What is the nature and magnitude of waterborne disease in the United States?"—answers to which can help define where the safety of drinking water fits into the overall considerations of public health, or identify particular situations (e.g., undisinfected wells, crossconnections) that are associated with disease. More familiarly, risk assessments can address narrower questions such as "What is the risk that oral ingestion of hexavalent chromium will cause lung cancer?" or "What is the likelihood of *Cryptosporidium* illnesses from a turbidity spike in an unfiltered surface water system?" where the answers might help define regulatory actions. Specific questions, such as "What is the likely differential number of cancers prevented between a MCL for arsenic at

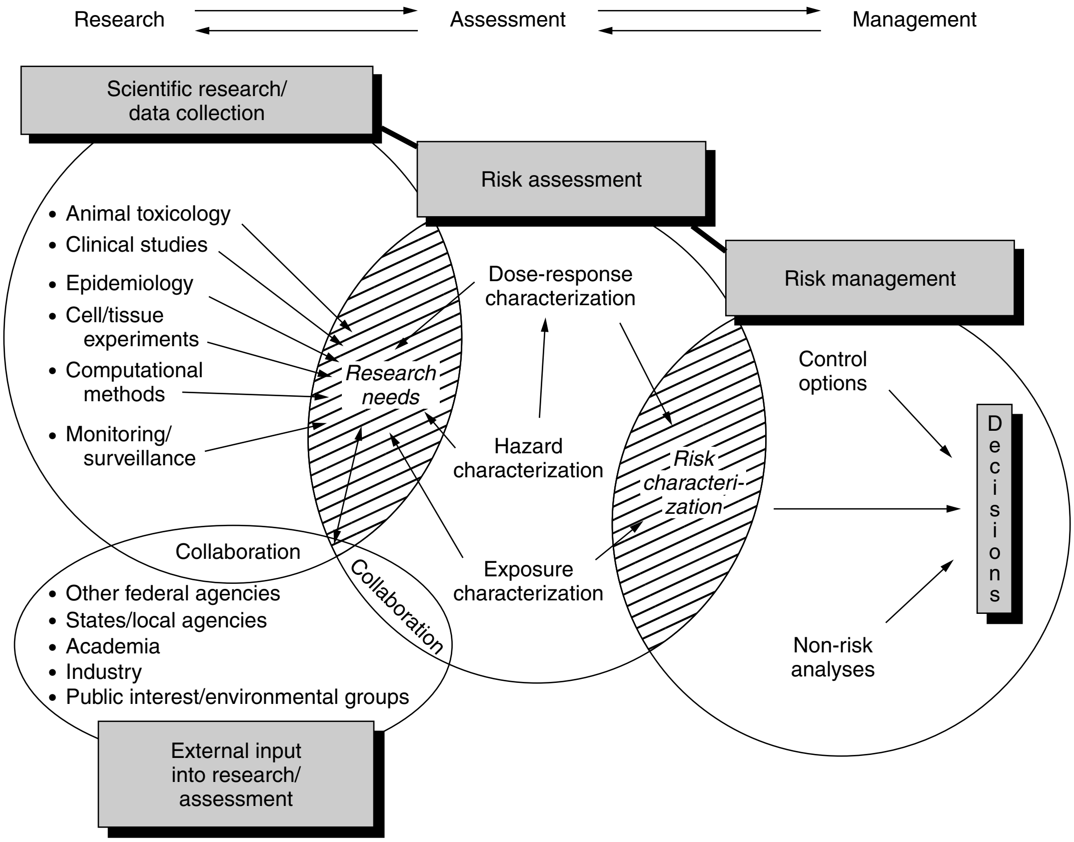

Figure 1. Integration of data collection, risk assessment, and risk management [*Source*: Farland (1)].

3 μg/L and one at 5 μg/L?” or “What are failure modes in the operation of an upflow clarifier?” may need to be answered to compare or determine possible solutions to the identified problems.

Risk assessment approaches for drinking water health and regulatory questions fall into three main types: those based on epidemiologic data, those calculated using mathematical risk models, and those based on analysis of systems and components. Epidemiologic and risk model approaches have been described more completely in earlier chapters, so only their applicability to regulatory development is addressed here. For this purpose, it is important to remember that there are substantial and important limitations to these approaches and that the conclusions from such assessments must be considered with due caution by those using the information. In addition, management decisions are often required in the face of uncertainty and lack of information. Risk assessments can provide useful information, but seldom give clear, unambiguous answers.

With respect to epidemiologic data, information from waterborne disease outbreaks, intervention studies, or controlled experiments on humans can be used directly to both quantitatively and qualitatively describe risks. However, the precision and accuracy of epidemiologic data typically limit these assessments to situations where excess risks (risks above background levels) are about 1% or greater. More often, epidemiologic studies of large groups or populations can describe only effects greater than 10–100% or more. In general, causal associations between contaminants and effects are considered significant only when effects are severalfold higher than background levels. Therefore, risk assessments done directly from these data can generally describe risks and answer risk questions only to this level of resolution. Because public health questions for drinking water often involve situations where risks are substantially below epidemiologic resolution, this approach is frequently inappropriate.

Perhaps the most familiar form of risk assessment is based on mathematical models used to extrapolate existing data to make quantitative estimates relevant to other situations. This well-known approach, first described by the National Academy of Sciences (NAS) in 1983 (2), organizes the process of human health risk assessment into four steps: hazard identification, dose–response assessment, exposure assessment, and risk characterization (Fig. 2). Hazard identification involves an evaluation of whether exposure to a substance would produce an adverse or otherwise undesirable effect. The data used to make such a determination usually come from animal studies. In some instances, human data may be available for the contaminant of interest. Dose–response assessment involves a more quantitative evaluation of the empirical evidence relating a specific exposure dose to the effect of interest. In particular, the available data are examined to determine the relationship between the magnitude of the exposure and the probability of the observed effect.

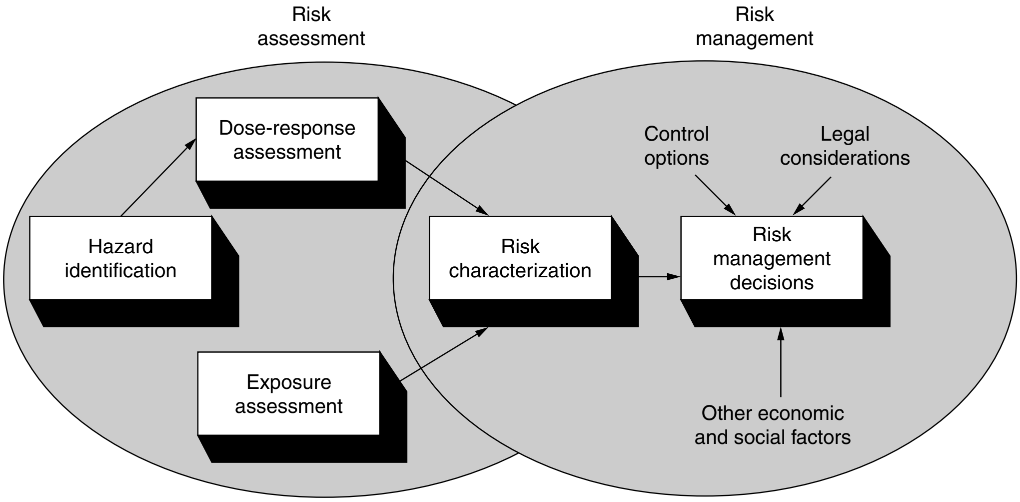

Figure 2. NAS risk assessment paradigm (left circle) and risk management (right circle) [*Source*: USEPA (3)].

Exposure assessments involve an evaluation of contaminant occurrence data, characterization of the environmental fate and transport of the contaminant from the source to the exposed population by different media (e.g., air, water, food), and physiological considerations of different exposure routes (e.g., ingestion, inhalation, dermal contact).

These toxicity and exposure products are combined in the risk characterization. The risk characterization describes the overall nature and magnitude of risk posed to human populations exposed to a particular contaminant. Included in the description is a discussion of what is known and not known about the hazards posed by the substance, what models were used to quantify the risk and why they were selected, assumptions and uncertainties associated with the qualitative and quantitative aspects of the assessment, and general level of confidence in the assessment.

A caution with these is that, while the calculated results are often presented as point value expressions of risk, it must be recognized that the farther the extrapolation from the original data, the more uncertainty and less precision in the results. The animal data themselves result from what are essentially small, highly controlled epidemiologic studies, and thus have the resolution limitations noted above. In addition, experimental variation may be 10% or greater. Because these studies are most often conducted with high exposures to the contaminants of interest and generate high risks, models are used to extrapolate the data to answer questions about lower environmental exposures and/or risks. These models have inherent limitations that magnify uncertainties. Exposure estimates likewise have substantial variation and uncertainty. The net result is that quantitative estimates of risk cannot describe a defined point risk for a certain exposure, but instead a range of possible risks. These ranges tend to increase with increasing model complexity to frequently span orders of magnitude. Most often, the range includes zero.

A third approach to risk assessment is based on analysis of an entire system or operation to identify vulnerabilities that could allow contaminants to reach the consumer. It is based on standard engineering design assessment approaches used to identify failure modes, judge probabilities of occurrence, and describe consequences. These have been adapted to focus on vulnerabilities in systems or operations that could result in human exposure to contaminants. This approach begins with a full description of the system or flow diagram of the process. Points where contamination can occur are identified. The likelihood and consequences of contamination at these points are described. From such an assessment, management actions can focus on controlling high-impact situations (Fig. 3). In the food industry, hazard assessment critical control point (HACCP) programs have been put into place to protect the public from, among other things, contaminated shellfish and other food items. Sanitary surveys, source water vulnerability assessments, and comprehensive performance evaluations (CPEs) are examples of this type seen in the drinking water industry. This is also the approach used for water system security and counterterrorism assessments. As currently practiced, these are qualitative, rather than quantitative in nature. Once the system is described and vulnerabilities identified, probabilities and consequences are typically rated on the basis of best professional judgment and using categories such as “high, medium, or low” or “minor, significant, or catastrophic.” The end result may be a list of vulnerabilities with some rankings for risks and consequences. Managers can use this information to identify problem areas and prioritize activities. Although not commonly practiced, it should be noted that such system analysis approaches are open to quantification of risks (4).

RISK MANDATES FROM THE SAFE DRINKING WATER ACT

The Safe Drinking Water Act (SDWA), as amended in 1986 and 1996 (5) has language that directs the U.S. Environmental Protection Agency (USEPA) to establish

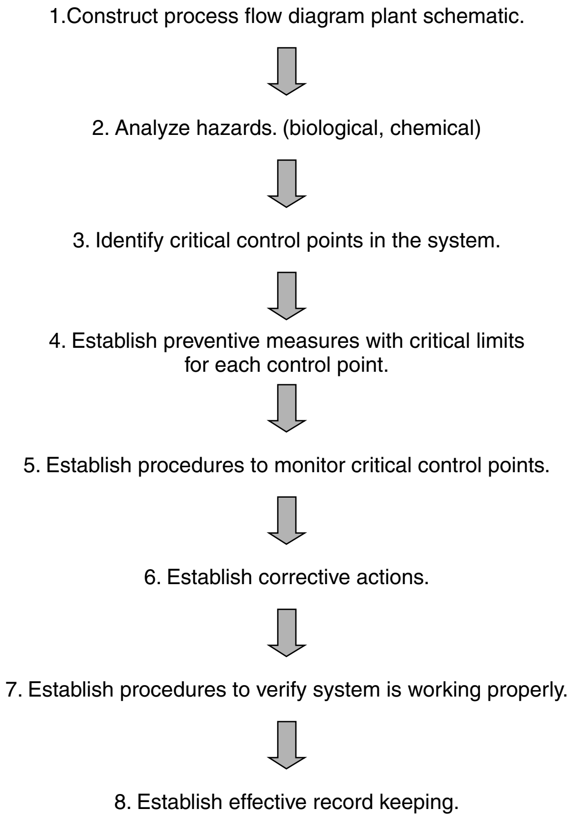

Figure 3. Hazard assessment critical control point (HACCP) risk assessment approach.

MCLGs for contaminants of public health concern for drinking water: “Each maximum contaminant level goal established under this subsection shall be set at the level at which no known or anticipated adverse effects on the health of persons occur and which allows an adequate margin of safety.” These goals, which are not enforceable themselves, are to be used to set the enforceable NPDWRs (5):

> Each national primary drinking water regulation for a contaminant for which a maximum contaminant level goal is established under this subsection shall specify a maximum contaminant level for such contaminant which is as close to the maximum contaminant level goal as is feasible. . . . For the purposes of this subsection, the term “feasible” means feasible with the use of the best technology, treatment techniques and other means which the Administrator finds, after examination for efficacy under field conditions and not solely under laboratory conditions, are available (taking cost into consideration).

In addition, a provision added to the Act in 1996 specifies priorities for selecting contaminants for rulemaking to

> take into consideration, among other factors of public health concern, the effect of such contaminants upon subgroups that comprise a meaningful portion of the general population (such as infants, children, pregnant women, the elderly, individuals with a history of serious illness, or other subpopulations) that are identifiable as being a greater risk of adverse health effects due to exposure to contaminants in drinking water than the general population.

In addition, the Act specifies that USEPA shall document health risks, opportunities for risk reductions, and benefits and costs of mandating these reductions and make these available for public comment. This health risk reduction and cost analysis (HRRCA) is a required component of the regulatory process.

By this and other language, Congress described both the necessary characteristics of risk assessments for drinking water regulations and the broad principles for managing these risks. It can be seen that by their words, Congress established a precautionary policy with regard to drinking water safety. The MCLG was to be set conservatively with respect to risk to more vulnerable individuals. The MCL was to be set to reflect the MCLG, with the additional considerations for technical feasibilities and costs.

These broadly stated goals set directions, but were not sufficiently described to be used for specific regulatory decisions. USEPA evolved operational interpretations of Congressional intentions for both the MCLG and for the acceptable public health risks associated with MCLs and treatment techniques following the 1986 amendments. These have been used consistently for NPDWRs from that time. Following the 1996 amendments, USEPA developed additional approaches for the benefit and cost analyses for the HRRCA.

DEVELOPING MCLS AND TREATMENT TECHNIQUES

The SDWA grants the USEPA Administrator the authority to publish a MCLG and promulgate a NPDWR for a contaminant if the contaminant may have an adverse effect on the health of persons, the contaminant is known or likely to occur in drinking water with a frequency or level of health concern, and there is a meaningful opportunity for health risk reduction. Development of a NPDWR normally begins with the identification of a drinking water contaminant. As provided in the SDWA, USEPA must list candidate contaminants for regulation on a periodic basis (6). Following listing, more detailed health, occurrence, exposure, and treatment technology information is gathered. When adequate information becomes available, a determination is made on whether to go forward with development of a NPDWR proposal (7). This determination uses a protocol developed and recommended to USEPA by the National Drinking Water Advisory Council (NDWAC). The health risk information is combined with occurrence data (levels, frequency, national distribution, persistence, etc.) and exposure estimates to predict the national number of individuals exposed above advisory levels. This risk assessment forms the basis for USEPA’s determination if regulation would provide a meaningful opportunity for health risk reduction.

Maximum Contaminant Level Goals

Once the decision is made to move forward, the MCLG is determined. MCLGs are risk assessment products developed by USEPA Office of Water and Office of

Research and Development. MCLGs are strictly health-based levels. They are developed and set at contaminant levels believed to be without appreciable health risk to individuals, to be consistent with the provision "set at the level at which no known or anticipated adverse effects on the health of persons occur." Additionally, they must address the concern for the protection of sensitive subpopulations. Therefore, risk assessments used to develop MCLGs must at a minimum provide an estimate for a zero-risk exposure level for humans who may be more sensitive to the contaminant. While these risk assessments are caveated to be upper bounds for estimated risks, such that the true risks may be less or even zero, the assessments do not have to, nor are they designed to, estimate the full range of true risks to the average individual.

MCLG risk assessments use the available toxicologic and epidemiologic health study data. The data are evaluated with respect to the nature of the adverse health effects from the contaminant, the strength of evidence for causal relationships, and their quality. Depending on the outcome, a dose–response estimate is made. As noted above, these data are almost always limited in quantity and quality, yielding substantial ranges for uncertainty. USEPA has chosen as a matter of policy to work to risks at the more conservative end of these ranges, in order to comply with the provision that the MCLG "allows an adequate margin of safety" in the face of these uncertainties. Therefore, poorer-quality data will lead to more stringent quantitative descriptions of risk.

For chemicals that produce adverse health effects and are not considered to be carcinogenic (noncarcinogens), the MCLG is based on the *reference dose* (RfD), which is defined as an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime. In practice, the RfD is set at a plausible zero-risk level. USEPA assumes that a physiological threshold exists for noncancer health effects from chemical contaminants, below which the effect will not occur. Thus the MCLG will be a nonzero number.

Depending on the quality of the available toxicity data, the RfD is usually derived from an experimental no-observed-adverse-effect level (NOAEL), identified as the highest dose in the most relevant study that did not result in a known adverse effect. The adverse effects chosen may themselves be mild and without clinical significance, but typically represent early stages in progression to more serious disease. In order to extrapolate from the data to human exposures protective of sensitive subpopulations, the NOAEL is divided by various uncertainty factors to derive the RfD. These uncertainty factors conservatively account for the variation in human response to the contaminant, extrapolation to human responses if animal data were used, the nature of the studies, data quality, and relevance. The result of this is that the RfD may differ from the NOAEL by as little as a factor of 3 (e.g., nitrate, arsenic) or as much as $\geq$1000 (e.g., methyl bromide, chlorobenzene). The RfD takes the form of dose ingested per unit body weight per day (μg kg^{-1} day^{-1}). RfDs that have been reviewed and agreed on by consensus within USEPA are listed in USEPA's Integrated Risk Information System (IRIS).

The RfD, which is based on the total daily amount of contaminant taken up by a person on a body weight basis, is converted to a drinking water equivalent level (DWEL) concentration and adjusted for the percentage contribution of other sources [relative source contribution (RSC)] of the contaminant besides drinking water (air, food, etc.) to arrive at the MCLG. This calculation traditionally assumes a lifetime consumption of 2 L of drinking water per day by a 70-kg adult, which is about the upper 90th percentile consumption level. More recent USEPA regulatory risk assessments, such as those for arsenic, have considered different drinking water consumption rates in addition to this default value to represent specific populations (infants, agricultural workers) and situations (8).

A different approach is taken for contaminants that may be carcinogenic. USEPA assumes as a default position that no toxicity threshold exists for induction of cancer and thus, there is no absolutely safe level of exposure. Until relatively recently, once it was determined that a contaminant is a known or probable human carcinogen, the MCLG was automatically set at zero. USEPA has now revised its guidelines for cancer risk assessments to reflect the increasing understanding of the several steps in the progression of cancer (9,10). Some contaminants have been reevaluated for their carcinogenicity using the current draft version of these guidelines. A good example is chloroform (6), which is now considered a carcinogen only at very high exposures associated with tissue damage, such that a threshold is indicated. This allows calculation of a nonzero MCLG using a margin of exposure (MoE) approach. As with RfDs, these determinations of carcinogenicity and their associated dose–response assessments that have been reviewed and agreed on by consensus within USEPA are listed in IRIS.

An alternative approach is used for situations where the data on carcinogenicity of a contaminant are equivocal or too scanty to make a clear judgment. These contaminants are termed "possible human carcinogens." For these, the MCLG may be derived from their relevant noncancer health effects as described above. The resulting RfD is divided by an additional uncertainty factor of 10 as a margin of safety for the possible carcinogenicity.

In a vein similar to that for carcinogens, microbial pathogens and indicator organisms are also assigned a MCLG of zero as a matter of policy, from the consideration that one infective unit (oocyst, cyst, virus particle, bacterium) could be sufficient to cause an infection. The available data on infectivity are supportive for this assumption for the pathogenic viruses studied and for the protozoa, *Giardia* and *Cryptosporidium*. While it is less clear that this is so for pathogenic bacteria, the data cannot exclude this possibility.

Identifying Candidate MCLs

Once the MCLG is established, it is combined with information on contaminant occurrence, treatment technologies, and analytical methods to suggest and evaluate possible regulatory criteria for further discussion. Although the MCLG is a regulatory value, it is not enforceable; the NPDWR is the enforceable regulatory element.

One element in the standard setting process is to determine whether the NPDWR should be a MCL or a treatment technique. An enforceable MCL can be established to control exposure to a contaminant when appropriate analytical methods exist to quantify the contaminant and determine compliance at the MCL. When methods are not available, as for certain microbial exposure situations, treatment techniques may be established that do not directly measure exposure, but use other indicators for compliance.

Two risk considerations come into play in the process of setting a MCL to reduce a contaminant as close as feasible to the MCLG. The first has to do with MCLGs of zero. It is impossible to quantify a contaminant or confirm treatment to a zero exposure. Therefore, MCLs must be above zero, and thus have some risk. A risk benchmark is used to identify appropriate safe drinking water exposures. These guide selection of analytical techniques, treatment approaches, and, ultimately, the MCL choices. This benchmark is based on a consideration of de minimus or “acceptable” risk. As a matter of policy, USEPA Office of Water has used an acceptable risk range for chemical carcinogens from one additional cancer per million people to one additional cancer per 10,000 people exposed to the contaminant over a lifetime (8,11). For pathogenic microorganisms, an acceptable risk of one additional infection per 10,000 people exposed per year has been used (12,13).

These allowable exposures are estimated from the associated dose–response curves. For carcinogens, this dose–response assessment uses models to extrapolate from the available data to zero exposure–zero risk, defining a curve that is essentially linear at low exposures. The resulting “cancer slope factor” allows for a convenient probability analysis of risks to individuals associated with different exposures. The exposures for the acceptable risk range are taken directly from the curve. However, as noted above, the uncertainties increase substantially in these extrapolations. USEPA traditionally uses the 90th percentile upper bound of the modeled results to minimize the possibility that risks in this exposure range are not greater than estimated. The end result of this is that the risk from a lifetime consumption of water at a given level is unlikely to be greater than estimated, is more likely substantially less, and may be zero.

A similar probabilistic risk approach is used for microbial contaminants. The dose–response models used for microbial risk assessments are somewhat more complex in that they must be selected to account for the particulate nature of the infective material in the environment. The modeled exposures are then used for further risk management.

For contaminants with noncancer health risks, the MCLG is used as the starting point for determining a MCL. Because this uses a “bright line” reference point (the DWEL), exposures need only be estimated and compared to the MCLG. However, since this is nonprobabilistic, alternative MCLs cannot be considered on the basis of risks. For most noncarcinogens, the MCL is set equal to the MCLG.

The estimated risks associated with different exposures are then matched against the existing environmental exposure levels to determine the magnitude of the public health problem to be solved. As discussed in other chapters, the additional considerations for treatment technologies and their feasibility and the availability of appropriate analytical methods are factored in at this point.

Health Risk Reduction and Cost Analysis

The second application of risk assessment in the development of a MCL or treatment technique is in the estimation of public health benefits to be gained by regulation. This is used in the management discussions leading up to a regulatory proposal. Under the SDWA of 1986, it was necessary to consider only analytical and treatment feasibility in establishing the MCL. However, a cost benefit assessment was produced as an element of the regulatory impact analysis (RIA). The 1996 revisions to the SDWA required USEPA to explicitly consider costs and benefits in determining the MCL. Therefore, an expanded HRRCA is now required as part of a regulatory proposal. To estimate benefits for different MCLs or treatment requirements, the risks to an individual at the resulting exposure levels must be multiplied by estimates of the number of individuals in the United States exposed to the different levels. In practice, a relationship between the number of individuals exposed versus exposure level is first produced, then further manipulated to account for existing and proposed treatment controls. This relationship may be calculated stepwise or by using Monte Carlo simulations based on exposure distributions. From this, the number of cancers or microbial illnesses avoided at a given regulatory level can be estimated. These “body counts” can be matched with information on the costs of treating the associated diseased and the dollar value of avoiding illness to give quantitative information on the monetary benefits of the regulation. These benefits are then matched against the implementation and compliance costs to utilities and oversight agencies (14). From a risk assessment perspective, it must be remembered that carcinogen dose–response curves represent upper-bound risks; thus estimates of the number of cancers based on these curves are also upper-bound values for any given exposure level.

This approach is most useful for benefits from reducing cancer or microbial illness risks. This is both because these disease endpoints are recognizable and definitive and because the impacts are quantifiable from their probabilities. This is not so for noncancer risks from chemicals. Because these are described by nonprobabilistic, zero-risk DWELs associated with subclinical health effects of indeterminate public health importance, it is difficult to quantify or assign monetary value to the benefits of reducing exposures to these chemicals.

Risk Assessments as Regulations

Risk assessments can comprise regulatory elements themselves. These may apply directly to utilities or secondarily through requirements on primacy agencies. The type of risk assessments currently required in NPDWRs are all qualitative system analyses. These include treatment system sanitary surveys found in the Total Coliform Rule (15) and Interim Enhanced Surface Water Treatment Rule (IESWTR) (16), the watershed sanitary surveys in the Surface Water Treatment Rule (SWTR) (12), and the CPEs, also in the IESWTR. These all require on-site evaluations to determine sources of contamination and vulnerabilities to failures that could compromise water quality. These analyses are similar to the HACCP process used by the National Aeronautics and Space Administration (NASA) and U.S. Department of Agriculture (USDA) to protect foods.

Australia has taken a more formal HACCP approach to protect their drinking water. Termed the "framework for management of drinking water quality," it is to be incorporated into the Australian Drinking Water Guidelines (17). The approach includes a systematic and comprehensive analysis of the entire route of water from source to tap (Table 1). This analysis leads to identification of hazards, sources of hazards, and associated risks. These are addressed in the subsequent institution of protective measures to yield multiple barriers to contamination. The emphasis of the HACCP approach is on proactive protection, rather than reactive responses.

Regulatory Reviews of NPDWRs

The 1996 SDWA Amendments required that USEPA review all existing NPDWRs every 6 years to determine if information available subsequent to promulgation would support regulatory revision. Congress stipulated that all such revisions must maintain, or provide for greater, protection of the health of persons. USEPA, working with the NDWAC, developed a protocol for these reviews, driven largely by reevaluations of health risk information (7). USEPA principally considered whether any new evaluation of oral ingestion risks could lead to revision of the MCLG. This approach allows the Agency to use up-to-date risk assessment approaches in its reevaluations, which will include a wider range of possible adverse health outcomes (reproductive and developmental) and better characterizations of carcinogenicity. The potential here is that some carcinogens would be reclassified in such a way that their MCLs could be relaxed while maintaining the same level of health protection to the public. The reclassification of chloroform as a threshold carcinogen is an example. This would recognize that improved risk assessments could reduce the scientific uncertainties that led to excessively stringent NPDWRs.

Table 1. Australian Framework for Management of Drinking Water Quality

First element
Commitment to drinking water quality management
System analysis and management
Assessment of the drinking water supply system
Planning preventive strategies for water quality management
Implementation of operational procedures and process control
Verification of drinking water quality
Incident and emergency response
Supporting requirements
Employee awareness and training
Community involvement and awareness
Research and development
Documentation and reporting
Review
Evaluation and audit
Review and continual improvement

FUTURE OUTLOOK

Risk assessments are used both formally and informally at several points within the regulatory process. They serve specific purposes and mandates and may have severe constraints on their representations of risk. While assessments may be qualitative or quantitative in nature, they are always limited by the available information and our current abilities to understand disease processes and predict outcomes. They will always be inexact. Recognizing and accepting their limitations is important for making sound regulatory decisions. An increased awareness of how risk assessments may be used will improve future regulatory discussions.

Acknowledgments
The author wishes to acknowledge the patient help of Fred Pontius, without which this article would have been the poorer.

BIBLIOGRAPHY

1. Farland, W.H. (2000). *Current and Proposed Approaches to Assessing Children's Cancer Risk*. USEPA/National Institute of Environmental Health Sciences Workshop: Information Needs to Address Children's Cancer Risk. March 30 and 31. USEPA Office of Research and Development, Washington, DC.
2. National Academy of Sciences. (1983). *Risk Assessment in the Federal Government: Managing the Process*. National Academy Press, Washington, DC.
3. USEPA. (1999c). *Research and Development: Fiscal Years 1997–1998 Research Accomplishments*. USEPA Office of Research and Development, Washington, DC.
4. Buchanan, R.L. and Whiting, R.C. (1998). Risk assessment: A means for linking HACCP plans and public health. *J. Food Protect.* **61**(11): 1531–1534.
5. *Title XIV—Safety of Public Water Systems*; Sec. 1412(b), National Drinking Water Regulations: Standards; 42 USC Sec. 300g.
6. USEPA. (1998a). Announcement of the drinking water candidate list; notice. *Fed. Reg.* **63**: 10273–10287.
7. USEPA. (2002a). National primary drinking water regulations; announcement of the results of EPA's review of existing drinking water standards and request for public comment. *Fed. Reg.* **67**: 19029–19090.
8. USEPA. (2001a). National primary drinking water regulations; arsenic and clarifications to compliance and new source contaminants monitoring; final rule. *Fed. Reg.* **66**: 6976–7066.
9. USEPA. (1999b). *Draft Guidelines for Carcinogen Risk Assessment*. NCEA-F-0644. USEPA, Washington, DC.

10. USEPA. (2001b). Notice of opportunity to provide additional information and comment. Draft revised guidelines for carcinogen risk assessment. *Fed. Reg.* **66**: 59593–59594.
11. USEPA. (1998d). National primary drinking water regulations: disinfectants and disinfection byproducts; final rule. *Fed. Reg.* **63**: 69390–69476.
12. USEPA. (1989a). Drinking water; national primary drinking water regulations; filtration, disinfection; turbidity, giardia lamblia, viruses, legionella, and heterotrophic bacteria; final rule. *Fed. Reg.* **54**: 27486–27541.
13. USEPA. (1998c). National primary drinking water regulations: interim enhanced surface water treatment; final rule. *Fed. Reg.* **63**: 69478–69521.
14. USEPA. (1999a). Health risk reduction and cost analysis for radon in drinking water: notice. *Fed. Reg.* **64**: 9560–9599.
15. USEPA. (1989b). Drinking water; national primary drinking water regulations; total coliforms (including fecal coliforms and *E. coli*); final rule. *Fed. Reg.* **54**: 27544–27568.
16. USEPA. (1998b). National primary drinking water regulations: disinfectants and disinfection byproducts notice of data availability; proposed rule. *Fed. Reg.* **63**: 15606–15692.
17. Australian Department of Health and Ageing. (2001). *Framework for Management of Drinking Water Quality: A Preventative Strategy from Catchment to Consumer*. National Health and Medical Research Council, Canberra, Australia.

READING LIST

USEPA. (1986). *The Risk Assessment Guidelines of 1986*. EPA/600/8-87/045. USEPA, Washington, DC.

USEPA. (2002b). Announcement of preliminary regulatory determination of priority contaminants on the drinking water contaminant candidate list. *Fed. Reg.* **67**: 38222–38244.

POTENTIAL RISKS OF WATERBORNE TRANSMISSION OF *ESCHERICHIA COLI* O157:H7

JANICE WEIHE
American Water, Quality Control and Research Laboratory
Belleville, Illinois

MARK LECHEVALLIER
American Water
Voorhees, New Jersey

ZIA BUKHARI
American Water, Quality Control and Research Laboratory
Belleville, Illinois

Escherichia coli O157:H7 can be transmitted by any route where fecal contaminated material is ingested. These routes may include foodborne, waterborne, person-to-person, and animal-to-person. Epidemiological investigations during outbreaks have identified that only a small number of organisms is needed to cause infection (i.e., 10–200 organisms) (3). This low infectious dose combined with the possibility of *E. coli* O157:H7 infections that do not show symptoms allows direct person-to-person transmission via poor hygiene.

The most well-documented route of *E. coli* O157:H7 transmission is foodborne. Several outbreaks have occurred from consumption of undercooked ground beef either in restaurants or within individual homes. In addition to meat, other foods have been implicated, including mayonnaise, unpasteurized apple juice, fermented hard salami, lettuce, vegetable sprouts, yogurt, and milk. Contamination of these foods may have occurred due to improper hygiene among food handlers or by contact between uncooked foods and raw meats contaminated with *E. coli* O157:H7. There is also the possibility of direct contamination of produce in the fields with manure from infected animals (4).

An increasingly common route of transmission is waterborne via recreational or municipal water sources. Several *E. coli* O157:H7 outbreaks have been linked to use of recreational waters. Typical sources of recreational water outbreaks may be fecal contaminated material entering a lake and lake water ingested by swimmers or by children with diarrhea contaminating pool water, usually with an insufficient level of chlorine. Outside Portland, Oregon, 59 individuals were infected after swimming in a recreational lake in 1991 (5). A water park outbreak in Georgia in 1998 resulted in dozens of infections and the death of one child. The outbreak was linked to contamination from a child with diarrhea combined with low chlorine levels in the pool water (6).

In addition to these recreational water outbreaks, several incidents have occurred with drinking water. A list of outbreaks linked to contaminated drinking water is shown in Table 1.

The first municipal waterborne *E. coli* O157:H7 outbreak occurred in Missouri in 1989 (7); more than 240 people were infected, and four individuals died. It was suspected that the outbreak occurred from backflow during two water main breaks, leading to intrusion of the pathogen. The most publicized waterborne outbreak of *E. coli* O157:H7 occurred in May 2000 in Walkerton, Ontario, Canada. Approximately 1350 cases of gastroenteritis were reported in individuals exposed to Walkerton municipal water, and overall, it was estimated that the number of cases associated with Walkerton exceeded 2300 individuals. Stool samples confirmed 167 cases of *E. coli* O157:H7 infection; 65 individuals were hospitalized, and 27 individuals developed hemolytic uremic syndrome. Four deaths were directly due to the *E. coli* outbreak in Walkerton, and *E. coli* O157:H7 was a contributing factor leading to three additional deaths. During the outbreak, a supply source well was contaminated with coliform bacteria and *E. coli* O157:H7. This well was prone to surface contamination, especially following flooding conditions, which is what preceded the Walkerton incident. Furthermore, environmental testing of livestock farms, especially a farm adjacent to this well, indicated the presence of *E. coli* O157:H7 infections in the livestock (16,17).

This most common drinking water treatment is disinfection with chlorine. Susceptibility of various strains of *E. coli* O157:H7 to chlorine disinfection has

Table 1. Outbreaks of *E. coli* O157:H7 from Drinking Water

Year, Location	# Infected	Description	Reference
1989, Missouri, U.S.	243	Two water main breaks may have led to backflow in an unchlorinated municipal water supply leading to four deaths	(7)
1990, Scotland	4	O157 contamination of reservoir water supply	(8)
1990, Japan	42	O157 cultured from tap and well water	(9)
1995, Scotland	6	Sewage contaminated a public water supply	(10)
1995, Minnesota, U.S.	33	Contaminated spring water used as drinking water source at campgrounds	(11)
1997, Washington, U.S.	4	Contaminated chlorinated groundwater supply at a trailer park	(12)
1998, Illinois, U.S.	3	Well water contaminated by cattle feces	(12)
1998, Wyoming, U.S.	157	Contamination of a spring and two wells that were municipal water sources	(13)
1999, Texas, U.S.	22	Inadequately chlorinated city well water	(14)
1999, New York, U.S.	127	Septic system contaminated a well used for drinking water at fairgrounds, resulting in 2 deaths	(15)
2000, Utah, U.S.	102	Illness among those drinking from an irrigation system	(14)
2000, Ohio, U.S.	29	Possible backflow at fairgrounds from animal barn to food vendor taps	(14)
2000, Walkerton, Ontario, Canada	167	Improper chlorination of municipal well water contaminated with O157 from surface water runoff resulted in seven deaths	(16,17)

been examined in several studies, and except for an occasional resistant strain (18), the organism can be rapidly inactivated by low levels of chlorine disinfectant and shows no unusual tolerance to chlorine.

According to a Water Quality Disinfection Committee survey (19), it was determined that United States water utilities must maintain a median chlorine residual of 1.1 mg/L and a median exposure time of 45 minutes before the point of first use in the distribution system. Based on these guidelines, it seems unlikely that *E. coli* O157:H7 would survive conventional water treatment practices in the United States at these levels of chlorination. Had appropriate chlorine residual levels been maintained, many of the aforementioned drinking water outbreaks, such as the Walkerton event, would have been avoided. In nondisinfected drinking water, *E. coli* O157:H7 demonstrates environmental survival and chlorine susceptibility similar to wild-type *E. coli* (20), which suggests that wild-type *E. coli* could be an adequate indicator organism for fecal contamination of water.

The United States Environmental Protection Agency regulates systems that operate at least 60 days per year and serve 25 people or more or have 15 or more service connections as public water systems under the Total Coliform Rule. This compliance monitoring of tap water for all public water systems can indicate whether the system is contaminated or vulnerable to fecal contamination, if it fails to identify specifically the presence of *E. coli* O157:H7. Existing detection methods for *E. coli* O157:H7 are intended for analyses of clinical specimens or food products, which are expected to contain large numbers of organisms. These methods lack sensitivity for detecting low numbers of contamination in environmental samples and typically require long (20–48 hour) incubation periods for analyzing small sample volumes. To ensure that public water systems can rapidly respond to contamination from this significant human pathogen, there is an urgent need for rapid detection methods for *E. coli* O157:H7 in water. The reader is referred to the article entitled "Understanding *Eschericia coli* O157:H7 and the need for rapid detection in water" by Bukhari, Weihe, and LeChevallier for further discussion of rapid methods for detecting *E. coli* O157:H7 in water.

BIBLIOGRAPHY

1. Boyce, T., Swerdlow, D., and Griffin, P. (1995). Current concepts: *Escherichia coli* O157:H7 and the hemolytic-uremic syndrome. *N. Engl. J. Med.* **333**: 364–368.
2. Shelton, D. and Karns, J. (2001). Quantitative detection of *Escherichia coli* O157 in surface waters by using immunomagnetic electrochemiluminescence. *Appl. Environ. Microbiol.* **67**: 2908–2915.
3. Willshaw, G.A., et al. (1994). Vero cytoxin-producing *Escherichia coli* O157 in beefburgers linked to an outbreak of diarrhea, haemorrhagic colitis and haemolytic uraemic syndrome in Britain. *Lett. Appl. Microbiol.* **19**: 304–307.
4. Mead, P. and Griffin, P. (1998). *Escherichia coli* O157:H7. *Lancet* **352**: 1207–1212.
5. Keene, W.E., et al. (1994). A swimming-associated outbreak of hemorrhagic colitis caused by *Escherichia coli* O157:H7 and *Shigella sonnei*. *N. Engl. J. Med.* **331**: 579–584.
6. Gilbert, L. and Blake, P. (1998). Outbreak of *Escherichia coli* O157:H7 infections associated with a waterpark. *Georgia Epidemiol. Rep.* **14**: 1–2.
7. Swerdlow, D.L., et al. (1992). A waterborne outbreak in Missouri of *Escherichia coli* O157:H7 associated with bloody diarrhea and death. *Ann. Intern. Med.* **117**: 812–819.

8. Dev, V., Main, M., and Gould, I. (1991). Waterborne outbreak of *Escherichia coli* O157. *Lancet* **337**: 1412.
9. Akashi, S., et al. (1994). A severe outbreak of haemorrhagic colitis and hemolytic uremic syndrome associated with *Escherichia coli* O157:H7 in Japan. *Eur. J. Pediatr.* **153**: 650–655.
10. Jones, I. and Roworth, M. (1996). An outbreak of *Escherichia coli* O157 and Campylobacteriosis associated with contamination of a drinking water supply. *Public Health* **110**: 277–282.
11. Levy, D., Bens, M., Craun, G., Calderon, R., and Herwaldt, B. (1998). *Surveillance for waterborne-disease outbreaks—United States, 1995–1996*. CDC Surveillance Summaries. MMWR 47 (No. SS-5): pp. 1–33.
12. Barwick, R., Levy, D., Craun, G., Beach, M., and Calderon, R. (2000). *Surveillance for waterborne-disease outbreaks—United States, 1997–1998*. CDC Surveillance Summaries. MMWR 49 (No. SS-4): pp 1–35.
13. Olsen et al. (2002).
14. Lee, S., Levy, D., Craun, G., Beach, M., and Calderon, R. (2002). *Surveillance for waterborne-disease outbreaks—United States, 1999–2000*. CDC Surveillance Summaries. MMWR 51 (No. SS-8): pp. 1–43.
15. New York Department of Heath. (2000). Health commissioner releases *E. coli* outbreak report. Available: <www.health.state.ny.us/nysdoh/commish/2000/ecoli.html> (cited November 2003).
16. Hrudey, S.E., Payment, P., Huck, P.M., Gillham, R.W., and Hrudey, E.J. (2003). A fatal waterborne disease epidemic in Walkerton, Ontario: Comparison with other waterborne outbreaks in the developed world. *Water Sci. Technol.* **47**: 7–14.
17. O'Conner, D.R. (2002). Report of Walkerton Inquiry: Part 1-The events of May 2000 and related issues. Available: <http://web.utk.edu/∼hydro/Geol685/Walkerton_Summary.pdf> (cited November 2003).
18. Zhao, T., Doyle, M.P., and Zhao, P. (2001). Chlorine inactivation of *Escherichia coli* O157:H7 in water. *J. Food Prot.* **64**: 1607–1609.
19. Water Quality Disinfection Committee. (1992). Survey of water utility disinfection practices. *J. Am. Water Works Assoc.* **84**: 121–128.
20. Rice et al. (1992).

SLOW SAND FILTRATION

National Drinking Water
Clearinghouse

First used in the U.S. in 1872, slow sand filters are the oldest type of municipal water filtration. Today, they remain a promising filtration method for small systems with low turbidity or algae-containing source waters. Slow sand filtration does not require pretreatment or extensive operator control—which can be important for a small system operator with several responsibilities

WHAT IS SLOW SAND FILTRATION?

Slow sand filtration is a simple and reliable process. Slow sand filters are relatively inexpensive to build, but they do require highly skilled operators (Fig. 1).

The process percolates untreated water slowly through a bed of porous sand, with the influent water introduced over the surface of the filter, and then drained from the bottom.

Properly constructed, the filter consists of a tank, a bed of fine sand, a layer of gravel to support the sand, a system of underdrains to collect the filtered water, and a flow regulator to control the filtration rate. No chemicals are added to aid the filtration process.

ADVANTAGES

Design and operation simplicity—as well as minimal power and chemical requirements—make the slow sand filter an appropriate technique for removing suspended organic and inorganic matter. These filters also may remove pathogenic organisms.

Slow sand filtration reduces bacteria, cloudiness, and organic levels—thus reducing the need for disinfection and, consequently, the presence of disinfection byproducts in the finished water. Other advantages include:

- Sludge handling problems are minimal.
- Close operator supervision is not necessary.
- Systems can make use of locally available materials and labor.

Slow sand filters also provide excellent treated- water quality (See Table 1). Slow sand filters consistently demonstrate their effectiveness in removing suspended particles with effluent turbidities below 1.0 nephelometric turbidity units (NTU), achieving 90 to 99 + percent reductions in bacteria and viruses, and providing virtually complete *Giardia lamblia* cyst and *Cryptosporidium* oocyst removal.

LIMITATIONS

Slow sand filters do have certain limitations. They require a large land area, large quantities of filter media, and manual labor for cleaning.

Table 1. Typical Treatment Performance of Conventional Slow Sand Filters

Water Quality Paramenter	Removal Capacity
Turbidity	<1.0 NTU
Coliforms	1–3 log units
Enteric Viruses	2–4 log units
Giardia Cysts	2–4 + log units
Cryptosporidium Oocysts	>4 log units
Dissolved Organic Carbon	<15–25%
Biodegradable Dissolved Organic Carbon	<50%
Trihalomethane Precursors	<20–30%
Heavy Metals	
Zn, Cu, Cd, Pb	>95–99%
Fe, Mn	>67%
As	<47%

Source: Adapted from Collins, M.R. 1998.

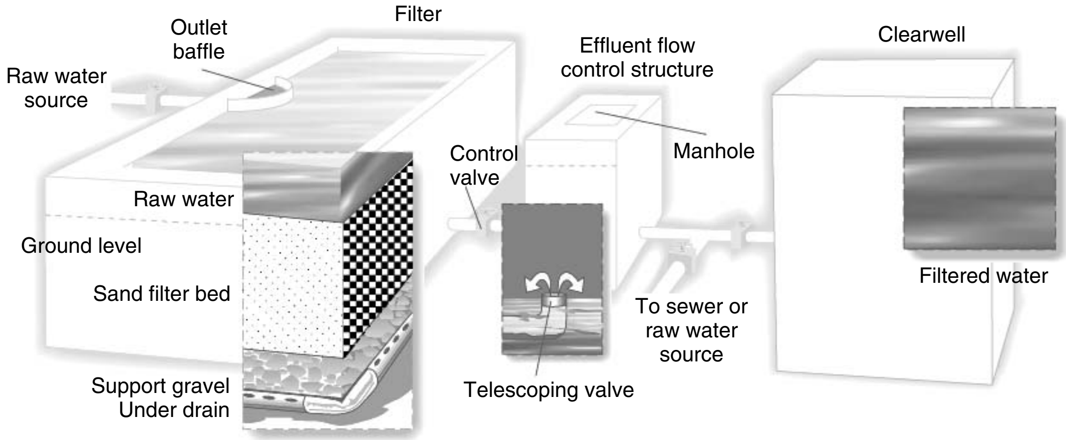

Figure 1. Slow sand filter.

Water with high turbidity levels can quickly clog the fine sand in these filters. Water is applied to slow sand filters without any pretreatment when it has turbidity levels lower than 10 NTU.

When slow sand filters are used with surface waters that have widely varying turbidity levels, infiltration galleries or rough filters—such as up- flow gravel filters—may be used to reduce turbidity.

Waters with a very low nutrient content may impair turbidity removal since some nutrients must be present that promote biological ecosystem growth within the filter bed.

Slow sand filters do not completely remove all organic chemicals, dissolved inorganic substances, such as heavy metals, or trihalomethane (THM) precursors—chemical compounds that may form THMs when mixed with chlorine. Also, waters with very fine clays are not easily treated using slow sand filters.

A granular activated carbon (GAC) sandwich filter is a modified slow sand filter that removes organic material. This filter uses a base sand layer that is approximately 1 foot deep, an intermediate GAC layer approximately 0.5 feet, and a top sand layer approximately 1.5 feet deep. This modified slow sand filter effectively removes pesticides, total organic carbon, and THM precursors.

Slow sand filters are less effective at removing microorganisms from cold water because as temperatures decrease, the biological activity within the filter bed declines.

PROCESS DESCRIPTION

Slow sand filters require a very low application or filtration rate (0.015 to 0.15 gallons per minute per square foot of bed area, depending on the gradation of the filter medium and the quality of the raw water). The removal action includes a biological process in addition to physical and chemical ones (Table 2).

A sticky mat of biological matter, called a "schmutzdecke," forms on the sand surface, where particles are trapped and organic matter is biologically degraded. Slow sand filters rely on this cake filtration at the surface of the filter for particulate straining. As the surface cake develops during the filtration cycle, the cake assumes the dominant role in filtration rather than the granular media.

Pilot testing is always necessary when designing slow sand filters. Currently, engineers are not able to predict the performance of a slow sand filter with a specific quality of raw water. Operation of a small pilot filter, preferably over several seasons of the year, will insure adequate performance of the full-scale plant.

Remember, after the designer sets the parameters—such as the plant filtration rate, bed depth, and sand size—there is little a plant operator can do to improve the performance of a slow sand filter that does not produce satisfactory water.

Slow sand filter pilot plant testing does not have to be expensive. Pilot plant testing has been done using manhole segments and other prefabricated cylindrical products, such as filter vessels.

Slow sand filter pilot facilities operate over long periods of time—up to a year—but the level of effort can be quite low, consisting of daily checks of head loss, flow rate, water temperature, and turbidity and taking coliform samples.

Since the purification mechanism in a slow sand filter is essentially a biological process, its efficiency depends upon a balanced biological community in the schmutzdecke. Therefore, filters should operate at a constant rate. When operation is stopped, the microorganisms causing bacteriological degradation of trapped impurities lose their effectiveness. Intermittent operation disturbs the continuity needed for efficient biological activity.

Allowing the filter to operate at a declining rate is one way of overcoming this problem. Declining rate filtration produces additional water, which is generally satisfactory. Moreover, the declining-rate mode may be applied during overnight operation, resulting in significant labor savings.

Storing filtered water is essential at a slow sand filter plant for two reasons. First, because of the importance of establishing biological activity, using chlorine ahead of the filter is inappropriate, and the operator must provide disinfectant contact time in a storage basin. Second, storage is needed for production equalization and demand.

Table 2. Design Summary of a Slow Sand Filter

Design Parameters	Recommended Range of Values
Filtration rate	$0.15\ \mathrm{m}^3/\mathrm{m}^2 \cdot \mathrm{h}$ ($0.1–0.2\ \mathrm{m}^3/\mathrm{m}^2 \cdot \mathrm{h}$)
Area per filter bed	Less than 200 m^2 (in small community water supplies to ease manual filter cleaning)
Number of filter beds	Minimum of two beds
Depth of filter bed	1 m (minmum of 0.7 m of sand depth)
Filter media	Effective size (ES) = 0.15–0.35 mm; uniformity coefficient (UC) = 2–3
Height of supernatant water	0.7–1 m (maximum 1.5 m)
Underdrain system	
Standard bricks	
Precast concrete slabs	Generally no need for further hydraulic calculations
Precast concrete blocks with holes on the top	
Porous concrete	
Perforated pipes (laterals and manifold type)	Maximum velocity in the manifolds and in laterals = .3 m/s Spacing between laterals = 1.5 m Spacing of holes in laterals = 0.15 m Size of holes in laterals = 3 mm

Source: Vigneswaran, S. and C. Visvanathan, 1995.

MONITORING AND OPERATION REQUIREMENTS

A slow sand filter must be cleaned when the fine sand becomes clogged, which is measured by the head loss. The length of time between cleanings can range from several weeks to a year, depending on the raw water quality. The operator cleans the filter by scraping off the top layer of the filter bed. A ripening period of one to two days is required for scraped sand to produce a functioning biological filter. The filtered water quality is poor during this time and should not be used.

In some small slow sand filters, geotextile filter material is placed in layers over the surface. In this cleaning method, the operator can remove a layer of filter cloth periodically so that the upper sand layer requires less frequent replacement.

In climates subject to below-freezing temperatures, slow sand filters usually must be housed. Uncovered filters operating in harsh climates develop an ice layer that prevents cleaning. Thus, they will operate effectively only if turbidity levels of the influent are low enough for the filter to operate through the winter months without cleaning. In warm climates, a cover over the slow sand filter may be needed to reduce algae growth within the filter.

Before cleaning a slow sand filter, the operator should remove floating matter, such as leaves and algae. When one unit is shut down for cleaning, the others are run at a slightly higher rate to maintain the plant output.

After cleaning, the unit is refilled with water through the underdrains. This water can be obtained from an overhead storage tank or by using water from an adjacent filter. When the clearwell is designed, the temporary reduction of plant output should be considered, ensuring that sufficient water is available for the users.

Once the filter is cleaned, the microorganisms usually re-establish and produce an acceptable effluent. In cooler areas, ripening may take a few days. Even then, if the effluent's turbidity is sufficiently low, the water supply can be resumed after one day with adequate chlorination.

Slow sand filter monitoring and operation is not complicated. Daily tasks include reading and recording head loss, raw and filtered water turbidity, flow rates, and disinfectant residual. If necessary, the operator should adjust the flow to bring water production in line with demand.

In addition, with the promulgation of the Surface Water Treatment Rule, each day the operator needs to use the flow data and disinfectant residual data to calculate contact time values and determine if disinfection is sufficiently rigorous. These duties may require one to two hours unless automated.

WHERE CAN I FIND MORE INFORMATION?

American Water Works Association. 1993. *Back to Basics Guide to Slow Sand Filtration*. Denver: American Water Works Association.

American Water Works Association. 1994. *Slow Sand Filtration: International Compilation on Recent Scientific and Operational Developments*. Denver: American Water Works Association.

Clark, R. M., and D. A. Clark. 1995. *Drinking Water Quality Management*. Lancaster, Pennsylvania: Technomic Publishing Company.

Collins, M. R. 1998. *"Assessing Slow Sand Filtration and Proven Modifications." In Small Systems Water Treatment Technologies: State of the Art Workshop*. NEWWA Joint Regional Operations Conference and Exhibition. Marlborough, Massachusetts.

National Research Council. 1997. *Safe Water from Every Tap: Improving Water Services to Small Communities*. Washington, DC: National Academy Press.

U.S. Environmental Protection Agency. 1990. *Environmental Pollution Control Alternatives: Drinking Water Treatment for Small Communities*. Washington, DC: Office of Water. EPA/625/5-90/025

U.S. Environmental Protection Agency. 1998. *Small System Compliance Technology List for the Surface Water Treatment Rule and Total Coliform Rule*. Washington, DC: Office of Water. EPA/815/R/98/001.

Vigneswaran, S. and C. Visvanathan. 1995. *Water Treatment Processes: Simple Options*.

HAVE YOU READ ALL OUR TECH BRIEFS?

"Tech Briefs," drinking water treatment fact sheets have been a regular feature in the National Drinking

Water Clearinghouse (NDWC) newsletter *On Tap* for more than four years. NDWC Technical Assistance Specialist Mohamed Lahlou, Ph.D., researches, compiles information, and writes these very popular items.

To order, call the NDWC at (800) 624–8301 or (304) 293–4191. You may also order online at *ndwc_orders@mail.estd.wvu.edu* or download fact sheets from our Web site at www.ndwc.wvu.edu.

Additional copies of fact sheets are free; however, postal charges may be added.

For further information, comments about this fact sheet, or to suggest topics, call Dr. Lahlou at one of the above numbers or contact him via e-mail at *mlahlou2@wvu.edu.*

APPROACHES FOR SECURING A WATER DISTRIBUTION SYSTEM

Marsha A. Hosner
DHI
Newtown, Pennsylvania

The world focus on water security has been extremely heightened since the attack on the World Trade Center on September 11, 2001. The thinking behind the new focus is in response to terrorist threats on our drinking water and infrastructure in general. This heightened focus is generally a public response—citizens wanting something to be done to make them safer than they are currently. But water related security issues are not new.

INHERENT RISKS

Catastrophic events involving water have been taking place regularly throughout history. On April 1, 1913, hundreds of homes in Portsmouth, Ohio, were washed away when a 69-foot wall of water washed through from a flood on the Great Miami River. More than 120 people were killed, and there was a subsequent smallpox outbreak. On Cape Cod, Massachusetts, treated sewage disposal has formed a contaminated plume of ground water that today is slowly spreading into drinking water wells through the shallow gravel and sand soil composition. Droughts take place on a regular cycle related to various weather systems (1).

The possibilities of things that could go wrong are impossible to list—a fact that points to a reality that has to be faced. There is no way to stop things from happening to your water systems. And though there are a number of things you can do to reduce the risk, the most important step to take is preparing in advance for the time when something does happen.

DRINKING WATER SPECIFICS

Water arrives at the consumer through a network of pipes that ultimately link to the water supply or treatment facility. However, to achieve maximum control over the system, for instance, in case of pipe failures or unusual demand patterns (such as fire flows), water supply networks are generally designed as complicated, looped systems, wherein each individual tap may be capable of receiving water from several sources and intermediate storage facilities. This means that the water from any given tap can arrive through several different routes and can be a mixture of water from several sources. The routes and sources for a given tap can vary over time, depending on the pattern of water use (2).

The Association of State Drinking Water Administrators devised a security vulnerability checklist for a water distributor to use as a guideline for security (3). It includes a list of critical components, such as shown in (Fig. 1): ground water, surface water, purchased water, buildings, pumps, treatment equipment and supplies, laboratories, chemicals, storage tanks, power supplies, auxiliary power supplies, pipes, valves, buildings, computers, files, work vehicles, telephones, and dispatch radios.

Citizens have always been concerned about who is watching the water and how and when they would be told if something went wrong. In 1993, there was a an outbreak of *Cryptosporidium*—an intestinal protozoan—that infested the Milwaukee water supply and sickened more than 400,000 people (4). The cause of the infestation was not determined, but it is thought that it related to storm water runoff across pasture land that fed cattle—known for carrying this protozoan. *Giardia* is another germ that can often be found in the water supply.

Except for freak accidents or intentional attacks on the system, U.S. drinking water is thoroughly tested and is safe to drink. The maximum contaminant levels are frequently tested on a prescribed schedule, and it is treated with chlorine and chloramines to kill germs. Of course, people not on the public water supply need to take steps on their own to ensure that there are no contaminants. There are many countries that do not treat water or test it—and in those countries, the risk of serious disease is much higher.

The Danish government hired the Danish Hydraulic Institute (DHI) to secure the water system for their country. The end result of that study, again, was that it cannot be done (Fig. 2). So steps were then taken to prepare responses in case of accidental or intentional contamination. One of the methods being used there is called Artificial Recharge Technology—a mixing of water sources to dilute the supply, so that it can be flushed to reduce contaminants quickly and eventually to remove them (5).

GOVERNMENT ACTIONS

The states have had money available for loan for many years through the EPA Superfund. This is a loan that has to be paid back over 20 years. This money can be used for fencing, cameras, and other physical security and also for software models and other emergency action planning tools. By the year 2004, all water distribution utilities serving approximately 3000 customers or more

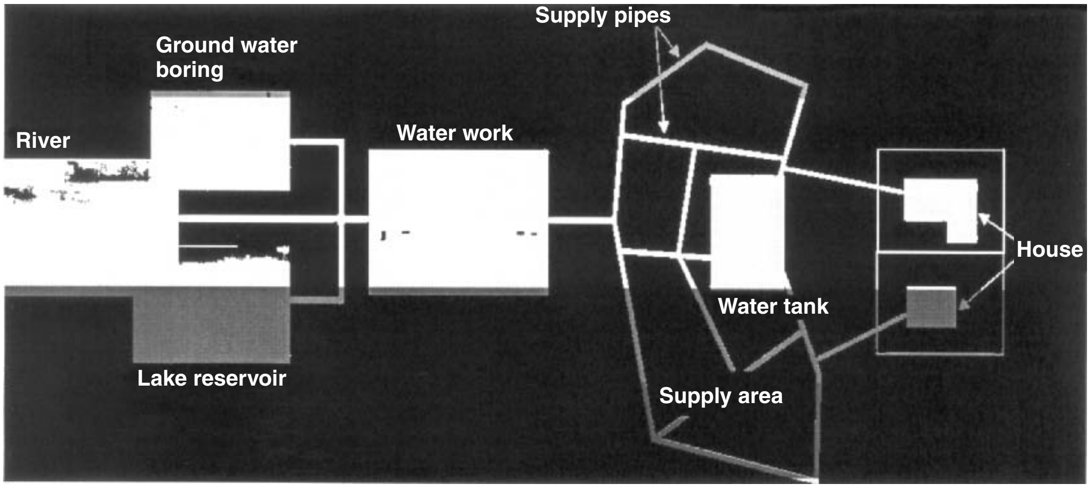

Figure 1. Components of a water system.

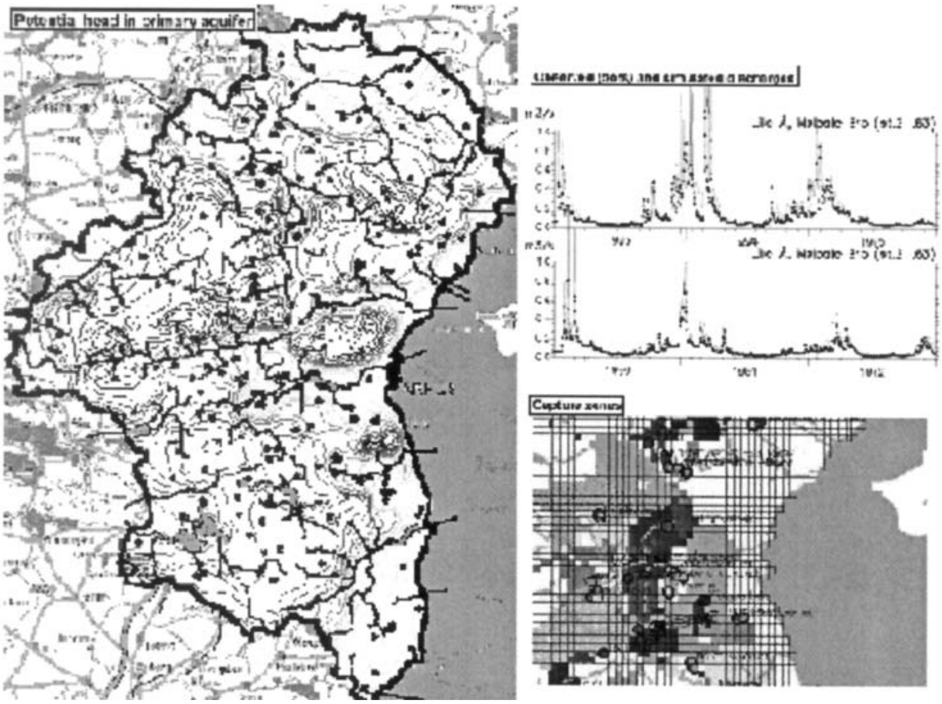

Figure 2. Aarhus, Denmark, water supply system data as seen in the emergency response planning software screen.

will be required to have written emergency action plans filed with the government.

These formal requirements for water utilities will be followed by formal requirements for wastewater utilities, and voluntary compliance by the wastewater utilities is expected to take place prior to formal implementation. In response to the need to establish these documents, many organizations have created written guidelines, as well as various training schemes. Sandia National Labs has been designated as the official resource for vulnerability assessment training (6). Many other private firms have been trained by Sandia to train others to perform these assessments.

Water security issues were addressed formally during the Clinton administration as just one part of a needed focus on all our infrastructure systems. In the White Paper printed by the U.S. Department of Justice, the Presidential directive states, “Critical infrastructures are those physical and cyber-based systems essential to the minimum operations of the economy and government . . . these infrastructures have become increasingly automated and interlinked. These same advances have created new vulnerabilities to equipment failures, human error, weather and other natural causes, and physical and cyber attacks. Addressing these vulnerabilities will necessarily require flexible, evolutionary approaches that span both the public and private sectors, and protect both domestic and international security. . .non-traditional attacks on our infrastructure and information systems may be capable of significantly harming both our military power and our economy" (7). Even in this document written in 1998, our government was requiring emergency action plans.

The Environmental Protection Agency (EPA) is, today, setting up the training, the grants, and the documentation for water agencies to meet their required vulnerability assessments and emergency action plans (8). Every water and wastewater agency is coming out with its own version of training aid documents, and commercial software packages are available from many companies to facilitate the planning process (Fig. 3). At least one commercial company, DHI, offers additional discounts to government buyers to hold down the cost of this compliance (www.dhigroup.com). There is also freeware available from the U.S. government. So lack of tools is not an issue for suppliers to prepare themselves.

TAKING ACTION

Whether it is a terrorist threat, flood, drought, earthquake, hurricane, or other natural catastrophe, there are logical steps to take to reduce the risk in any situation:

1. Identify the resources at risk and label them.
2. Determine who is responsible as a caretaker of each resource—and make sure they know it.
3. Measure the degree of risk.
4. Monitor the condition of the resource frequently.
5. Develop contingency plans.

First determine the water resources in your realm that are susceptible to loss or destruction—water supply, bridges, or treatment plants. Make an inclusive list. Figure 1 shows all the components of a water supply system—it is more than the waterworks plant. Be sure that the surrounding community is aware of the resources—posting signage, for example, that a water supply is not for dumping, and other things that might seem obvious to some are not so obvious to others. We are all caretakers of our environment. Many catastrophic events can be avoided simply by educating those who are in the area about proper care of their resources. Running out of water is a very scary possibility, and the public should be informed and trained in water conservation tactics—especially in time of droughts.

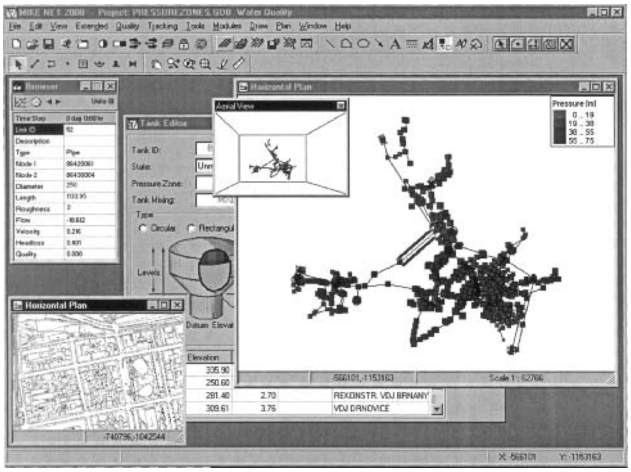

Figure 3. Use of a commercial software package to prepare action plans.

Second, identify the responsible person for each item you listed—it could be an organization, individual, or a department. Ensure that they know the full scope of their responsibility. They may know they are responsible, say, for water distribution—but have they been acknowledged as responsible for safety for that system? Ensure that they have been empowered to take needed steps to secure it. These are organizational issues generally that have to be met by some level of municipal government. The responsible parties must have the proper authority for exercising their duty.

Third, the responsible authority should measure the resource. They should get a firm grasp of the magnitude of their responsibility—both the physical size and the number of people involved. They need to know what condition it is in today. If steps can be taken to make them less vulnerable, they should be taken. It is impossible to improve the condition of anything if you do not know what condition it is in now. Measure it, and determine what shape it should be in to satisfy the community it serves (Fig. 4). Then take steps to get it in shape.

Fourth, monitoring the current condition can vary greatly depending on the vulnerability of the resource. If it is a condition that can be watched, then someone should be checking the vulnerability or watching for safety reasons. For instance, water quality samplers can be installed to check, at any interval desired, the condition of any water supply. Simple fences can keep out unwanted visitors. Periodic inspections to keep things at optimum condition are easily performed.

Fifth, contingency planning is multiphased. It should be determined if there are adequate resources, in the way of tools and manpower, to manage the system in the event of a catastrophe. Contingency plans should exist

Figure 4. Monitoring the data from a storm drain flow monitor.

that can be readily executed. Authorities should know of the contingency plans—how to locate them and how to execute them. Models can be developed of water supplies and wastewater collection systems that can then be used to show a multitude of "what if" scenarios. For water distribution systems, the use of a model is key to creating proper emergency response plans.

The American Water Works Association published the Manual of Water Supply Practices M19, which includes the section "Emergency Planning for Water Utility Management" (9). In chapter 5 of this document, it states that the basic principles of an emergency action plan are that it should "(1) use or reference existing resources, (2) be concise and logical, and (3) be coordinated with other agencies." There is a great push to implement SCADA (supervisory control and data acquisition) systems to watch the operation of a water system in real time—even monitoring the water quality, so it is important that these automated systems have built into them the ability to notify officials in case of sudden change and emergency. Many systems simply monitor and react without human notification. For introduction of toxins and protection of public health and safety, this would not be acceptable to the public, nor should it be.

In summary, I quote the AWWA from the *Security Analysis & Response for Water Utilities*: "A completed security assessment and response plan is useful only if the utility staff and supporting agencies maintain continued vigilance" (10). Ongoing training of staff, ongoing review of the emergency system, along with periodic drills of implementation of emergency actions are vital to successful performance of any emergency action planning.

BIBLIOGRAPHY

1. Hosner, M. (2002). Five steps to risk reduction. *Water 21 Mag.* 54–55.
2. Hosner, M. (2002). Use of a computer model to help secure a water system in Denmark. *cover, EH&S Software News Online* **XI**(5).
3. *Association of State Drinking Water Administrators, National Rural Water Association, "Security Vulnerability Self-Assessment Guide for Small Drinking Water Systems"*, May 30 2002.
4. Symons, J. M. (2001). *Plain Talk About Drinking Water*, 4th Edn. American Water Works Association, pp. 5, 25.
5. Brun, A. (2000). Artificial recharge technologies. *Programme for Part A: Environment and Sustainable Development*, Official Journal of the European Communities **C(2000)3118**.
6. AWWA. Mainstream—*AWWA seeks $2.5 million from Congress for security training*. September/October 2002, p. 5.
7. US Department of Justice *www.usdoj.gov/criminal/cybercrime/white_pr.htm*, White Paper, "The Clinton Administration's Policy on Critical Infrastructure Protection: Presidential Decision Directive 63," May 22, 1998.
8. US Environmental Protection Agency, *www.epa.gove/safewater/security*, "What is being done to protect the nation's water infrastructure?"
9. AWWA. (2001). *Emergency Planning for Water Utilities, a manual of water supply practices, M19*, 4th Edn.
10. AWWA. (2001). *Security Analysis & Response for Water Utilities*. p. iii.

WATER SECURITY: AN EMERGING ISSUE

Hemant J. Purohit
Sukumar Devotta
Environmental Genomics Unit,
National Environmental
Engineering, Research
Institute, CSIR
Nagpur, India

One of the most pressing global issues is the increase in world population and its impact on the availability of freshwater. Freshwater is generated, transported, and stored within confined river basins. In 1992, the Dublin conference, Assessment, CFWA, was considered, where water was regarded as an economic good, and in 1994, the United Nations Commission for Sustainable Development proposed that a comprehensive assessment of the freshwater resources of the world should be undertaken (1). Climate change adds further risks and uncertainties to the global picture, requiring adaptive management in water resources based on monitoring and re-evaluation (2). A global-scale assessment of water resources and their use was performed in the framework of the 1997 United Nations Comprehensive Assessment of the Freshwater Resources of the World (3). Since 1970, global water demand has risen at an estimated rate of 2.4% per annum, with much higher trends in developing countries because of increasing urbanization and water-intensive agricultural activities (4). Developing countries are home to the overwhelming proportion of the world's 2.6 billion people without access to clean water, a figure that is expected to significantly increase in the next decades (5). The relationships between water quantity and environment have been overlooked in the past. A need exists for a new dimension of the impending water crisis (6). For example, in India, it is expected that severe water crises will occur by the year 2025 (7).

MODELS FOR PREDICTION OF GLOBAL WATER AVAILABILITY

Identification of the links between rapidly growing populations and shrinking freshwater supplies is the essential first step in making water use sustainable. Various models are proposed for prediction of water usage based on climate data. Even water security has been correlated with food security by modeling (8). Yates (9) presented an integrated assessment model for continental scale runoffs where the limiting data for the spatial and temporal variability of freshwater resources is the constraint. An analysis of Western Europe's and Africa's freshwater runoff, which spans a range of climate variability, was performed at varying levels of spatial aggregation and at both monthly and annual time steps. Model results showed that regional runoff characteristics were lost beyond a data aggregation of 1° × 1° resolution. A macroscale hydrological model, designed to simulate stream flow from climatic inputs over a large geographic domain, has been proposed by Arnell (10). In the case of

Hydrological Model of WaterGAP 2 (WGHM), the model tuning has aimed at achieving a good representation of the long-term average discharges (11). It has a strong positive effect on the performance of all proposed water availability indicators.

WATER SOURCE AND RESOURCES

The prevention of water contamination is always a preferable scenario instead of attempting to remove contaminants once they have entered the aquatic environment. Although it is likely that some contamination events will always occur, a large proportion of water resources can be secured through protection and good water resource management; efficient design, operation, and management of water supplies; and regular and thorough surveillance activities. Water source protection is a mixture of case-specific measures of a particular area to protect individual sources. It also involves wider ranging measures, which are designed to protect the larger water resource bodies on a provincial, national, or regional (international) basis, which always requires guidelines such as state-level or provincial-level guidelines up to international treaties to ensure sustainable management of resources. Surface water sources and resources are far more open to contamination, and potential sources for contamination are generally far larger than for groundwater. Thus, any measures taken to protect surface water resources will generally encompass a far wider geographical region than measures designed to protect groundwater resources. At a global level, the advanced scientific methods, such as GIS applications and remote sensing techniques, quantify the existing water resources and help in formulating a sustainable management scheme by considering both water uses and environmental protection (12). Groundwater is an important source of drinking water. In its natural state, groundwater is generally of high microbiological quality with little or no contamination. The relative purity of groundwater in its natural state is largely a result of infiltration through the soil and unsaturated layers of rock, although some groundwaters do have high levels of harmful chemicals, such as fluoride and arsenic (7).

THREATS TO WATER RESOURCES

The water resources can be at threat because of natural geological disturbances or anthropological activities, which are making water unsafe for human consumption. Various anthropogenic activities exist that are making water resources unsafe, such as selenium pollution, which is a globally identified phenomenon and is associated with the most basic agricultural practices to the most high-tech industrial processes (13).

The prevention of contamination of groundwaters by persistent mobile contaminants is an essential element in the protection of groundwater resources. Generally, where aquifers are overlain by a substantial unsaturated zone and have high primary porosity and reasonable permeability, they tend to be less vulnerable to pollution. In industrial cities, initial development led to a decrease in the groundwater levels and subsequent accumulation of contaminants in the unsaturated zone. However, further developmental activities that have occurred there, or maybe elsewhere using different water sources, sometimes leads to a recovery of groundwater levels owing to desorption of contaminants and groundwater pollution, which has led the concept of “Groundwater Protection Zones” where acceptable land uses are defined in order to protect the underlying groundwater. Many factors exist that influence the shape of the zone: the nature of the aquifer (which are very rarely isotropic); the number of rivers in the zone; the condition of rivers (whether influent, effluent, perched, or changing); and the number and location of other abstraction points within the zone. The more rivers associated with the aquifer, the greater the distortion and extension of the zone.

In general, water can become contaminated at the original water source, during treatment, in the pipes that distribute water from a treatment facility to homes and businesses, or in containers. Surface water (rivers or lakes) can be exposed to acid rain, storm water runoff, sewage overflow, pesticide runoff, and industrial waste. This water is cleansed somewhat by exposure to sunlight, aeration, and micro-organisms in the water. Groundwater (aquifer) generally takes longer to become contaminated, but the natural cleansing process also may take much longer because it moves slowly and is not exposed to sunlight, aeration, or aerobic (requiring oxygen) micro-organisms. Groundwater can be contaminated by disease-producing pathogens, leachate from landfills and septic systems, careless disposal of hazardous household products, agricultural chemicals (fertilizers, pesticides, herbicides), industrial chemicals, and leaking underground storage tanks. Water distribution systems contain living micro-organisms and nutrients that enter a system with raw water during water treatment failures or from leaks, cross-connections, and back-flows. Bacterial growth may also occur at or near the pipe surfaces (biofilms), the interface with suspended particulates, and within the water itself. Besides microbes, other contaminates occurring in drinking water can include: (a) organics (trihalomethanes and other disinfection byproducts), which are formed when chlorine and other water disinfectants combine with naturally occurring organic matter; pesticides, including herbicides, insecticides, and fungicides; and volatile organic chemicals; and (b) inorganics (arsenic, barium, chromium, lead, mercury, and silver). The Centre for Science and Environment, New Delhi, India, has also expressed concern over the contamination of pesticides in bottled water.

WATER SECURITY AND ITS RELATION TO FOOD SECURITY AND AGRICULTURE

The International Food Policy Research Institute has developed a 2020 Vision of *a world where every person has economic and physical access to sufficient food to sustain a healthy and productive life, where malnutrition is absent, and where food originates from efficient, effective, and low-cost food and agricultural systems that are compatible with* *sustainable use and management of natural resources*. The major challenges observed to this Vision 2020 are ignorance to water security (14). The Indian national consultation program (15) also projected the World Water vision for the year 2025. Projections for world agriculture in the first half of the twenty-first century vary widely, largely depending on assumptions of yield growth. The pattern of yield growth for major crops is logistic, and not exponential. This pattern is consistent with ecological limits on soil fertility, water availability, and nutrient uptake (16).

Agriculture demands more water than any other activity. Although the total amounts of water made available by the hydrologic cycle provide enough freshwater for the world's current population, most of this water is concentrated in specific regions, leaving other areas water deficient. Therefore, the "Virtual water" concept has been introduced for water-short countries in recent years. These countries can minimize their use of water and achieve food security at the same time by importing a portion of their food requirements from other areas or countries where water resources are adequate and available at a lower cost (17). Similarly, a Water Poverty Index has been proposed, which integrates water stress and scarcity and physical estimates of water availability with socioeconomic variables (18). The effect of usage of virtual water has been analyzed nicely in a case study where, by taking six southern Mediterranean countries, the water-food challenges facing water-scarce countries and the implications for the world food economy have been discussed. By accounting for the volume of virtual water embedded in food imports into the countries concerned, a close relationship between water endowment and food import dependence has been elaborated. The results of this case study highlight that food imports are imperative for compensating water resource deficiency and that water scarcity-related food imports will continue to have an impact on food economy (19).

Globally, water is likely to become an increasingly critical resource issue in the developing world. In the coming time, the forces of globalization with technological revolution will influence the trans-boundary management of water as equity. The world is moving into a new kind of economy as well as into a new kind of society, where we need new mindsets and knowledge to resolve increasingly complex and interrelated continental issues. The water sector is no exception to this development.

BIBLIOGRAPHY

1. Björklund, G. and Kuylenstierna, J. (1998). The comprehensive freshwater assessment and how it relates to water policy world wide. *Water Policy* **1**: 267–282.
2. Sophocleous, M. (2004). Global and regional water availability and demand: Prospects for the future. *Natural Resources Res.* **13**: 61–75.
3. Raskin, P., Gleick, P., Kirshen, P., Pontius, G., and Strzepek, K. (1997). Water futures: Assessment of long-range patterns and problems, *Comprehensive Assessment of the Freshwater Resources of the World*. Stockholm Environment Institute, Stockholm, Sweden.
4. Clarke, R. (1993). *Water: The International Crisis*. MIT Press, Cambridge, MA.
5. Jong-Wook, L. and Bellamy, C. (2004). In: *WHO/UNICEF Joint Monitoring Program for Water Supply and Sanitation Report*.
6. Lundqvist, J. (2000). Physics and chemistry of the earth, Part B. *Hydrol. Oceans Atmosphere* **25**: 259–264.
7. Gordon, B., Mackay, R., and Rehfuess, E. (2004). *Inheriting the World: The Atlas of Children's Health and the Environment* WHO.
8. Kamara, A. and Sally, H. (2003). Water for food, livelihoods and nature: Simulations for policy dialogue in South Africa. *Phys Chem Earth, Parts A/B/C* **28**: 1085–1094.
9. Yates, D.N. (1997). Approaches to continental scale runoff for integrated assessment models. *J. Hydrol.* **291**: 289–310.
10. Arnell, N.W. (1999). A simple water balance model for the simulation of streamflow over a large geographic domain *J. Hydrol.* **217**: 314–335.
11. Döll, P., Kaspar, F., and Lehner, B. (2003). A global hydrological model for deriving water availability indicators: Model tuning and validation. *J. Hydrol.* **270**: 105–134.
12. Zacharias, I., Dimitriou, E., and Koussouris, T. (2003). Developing sustainable water management scenarios by using thorough hydrologic analysis and environmental criteria. *J. Environ. Manage.* **69**: 401–412.
13. Lemly, A.D. (2004). Aquatic selenium pollution is a global environmental safety issue. *Ecotoxicol. Environ. Safety* **59**: 44–56.
14. Pinstrup-Andersen, P. and Pandya-Lorch, R. (1998). Food security and sustainable use of natural resources: A 2020 Vision. *Ecologic Econom.* **26**: 1–10.
15. ICID. (2003). *Indian National Consultation Program of International Commission on Irrigation and Drainage and Country Policy Support Program Report*. Nov. 21–22, New Delhi, India.
16. Harris, J.M. and Kennedy, S. (1999). Carrying capacity in agriculture: global and regional issues. *Ecologic Econom.* **29**: 443–461.
17. Allan, J.A. (1998). Virtual water: a strategic resource, global solutions to regional deficits. *Ground Water* **36**: 545–546.
18. Sullivan, C. (2002). Calculating a water poverty index. *World Develop.* **30**: 1195–1210.
19. Yang, H. and Zehnder, A.J.B. (2002). Water scarcity and food import: a case study for southern mediterranean countries. *World Develop.* **30**: 1413–1430.

READING LIST

Rosegrant, M.W. and Perez, N.D. (1995). *Water Resources Development in Africa: A Review and Syntheses of Issues, Potentials and Strategies for the Future*. International Food Policy Research Institute (IFPRI), Washington, DC.

GUIDE TO SELECTION OF WATER TREATMENT PROCESSES

Michael Muntisov
GHD Pty Ltd.
Melbourne, Victoria, Australia

Selecting the appropriate treatment process is a critical step in providing safe, reliable, good quality drinking water at a cost-effective price.

This article provides guidance in selecting appropriate treatment processes taking into account raw water quality and costs.

RAW WATER QUALITY AND TESTING

There is a need for raw water quality data covering an extended period sufficient to show seasonal and extreme events to make a sound decision on appropriate treatment processes.

Ideally, the parameters listed in Table 1 should have data collected as set out.

In practice, the extent of data set out in Table 1 is rarely available. In these cases, immediate intensive periods of testing are essential to develop as much understanding of the raw water characteristics as possible in advance of selecting treatment processes. Adjacent similar water sources are possible sources of indicative data. In addition, risk assessment techniques can identify potential hazard parameters even where sampling data are limited. For some possible contaminants, it is appropriate to assess risk rather than sampling because measurements may not be representative. This approach involves examining possible sources of contaminants in the catchment and then calculating the risk or approximate concentration that might occur in the raw water.

Before any process is finally selected, it is important to carry out treatability testing on the actual source water. All waters have subtle differences, and these can have a significant effect on process selection and performance. As a minimum, coagulation jar tests should be performed.

Testing is particularly important for correct sizing and cost analysis of competing options to take account of actual conditions. For example the amount of coagulant and other chemical usage can have a significant influence on the cost-effectiveness of a coagulant-based conventional process versus, say, microfiltration. Conversely, although it can be assumed that microfiltration achieves excellent particle and protozoa removal reliably, its performance in terms of fouling and sustainable flux rate can only be determined realistically by pilot testing the water to be treated. Pilot testing is always recommended, if feasible, particularly for large projects.

THE BENCHMARK PROCESS

To allow comparison between process options, particularly on a cost basis, it is appropriate to define a "benchmark" process.

The benchmark process is based on the most commonly used treatment configuration in the world, coagulation/clarification/granular media filtration.

A typical process flow schematic is shown in Fig. 1.

Within the benchmark, there can also be considerable variation between clarifier and filtration types. An indication of types of clarifier available and their basic characteristics are set out in Table 2.

Modern granular media filters typically operate at 10 to 15 m/h in postclarifier configurations, although there are examples of deep-bed higher rated filters operating under suitable conditions. There is a wide variation in granular media filter designs. Table 3 sets out some examples at both ends of the spectrum.

Table 1. Minimum Sampling Regime for Raw Water

Surface Water	Groundwater (Not Significantly Influenced by Surface Water)
Initially and then annually • Complete scan of WQ parameters. From this, additional parameters may be added or removed from the weekly/monthly testing regime. *Weekly* • Turbidity • True colour • Alkalinity • pH (at sample point) • Temperature (at sample point) • Conductivity *Monthly* • *Cryptosporidium / Giardia* • Total coliforms • *E.coli* • Iron • Manganese • DOC • Calcium (component of hardness) • Magnesium (component of hardness) *Special* • Short-term intensive testing of certain 'suspect' parameters depending on circumstances for example hourly measurement of turbidity on flashy river; pesticides or *Cryptosporidium* after storm;	*Initially and then annually* • Complete scan of WQ parameters. From this, additional parameters may be added or removed from the monthly/quarterly testing regime. *Monthly* • Turbidity • pH (at sample point) • Calcium (component of hardness) • Magnesium (component of hardness) • Iron • Alkalinity • Temperature (at sample point) • Conductivity *Quarterly* • True color • Manganese • Total coliforms (monthly if detected) • *E.coli* (monthly if detected) *Potential Concerns* • Parameters sometimes found in groundwaters include — Nitrate — Radionuclides — Hydrogen sulfide — Arsenic — Other suspect parameters, for example, pesticides

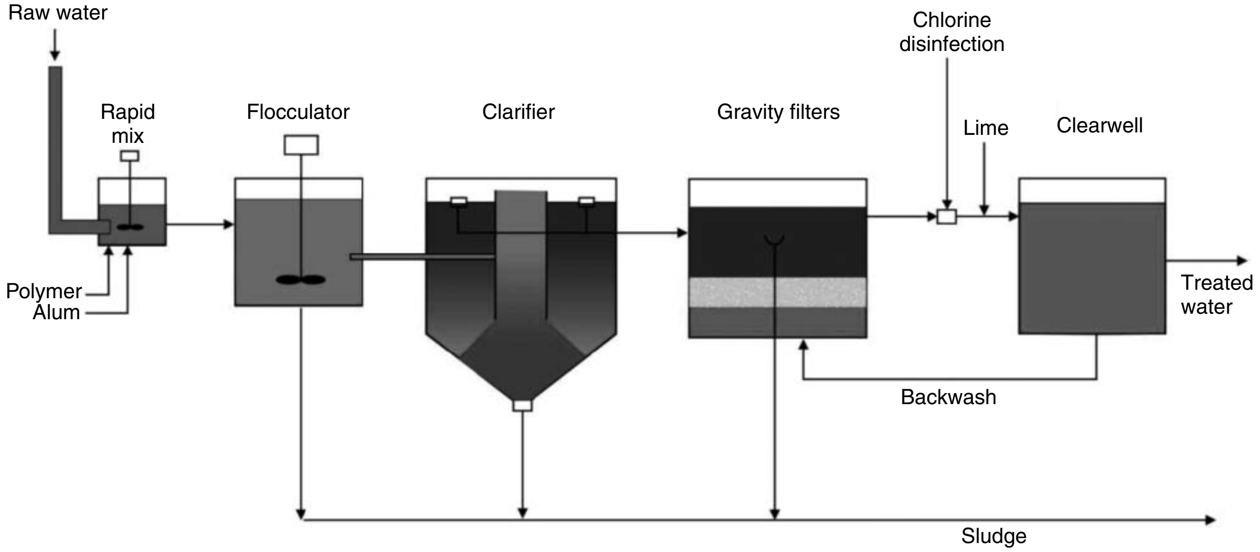

Figure 1. The "Benchmark" process.

Table 2. Clarifier Options in "Benchmark" Process

Clarifier Type	Typical Loading Rate m/hr	Minimum Flocculation	Comment
Horizontal flow sedimentation	1.5	2 stage, 15 minutes }	Sedimentation particularly suited to waters with high solids load
Radial flow sedimentation	1.5	2 stage, 15 minutes }	
Sludge blanket clarifier	2.5	NA }	
Sludge recirculation contact clarifier	3.0	15 minutes }	
Pulsation type	3.0	NA }	
Dissolved air flotation (DAF)	10	2 stages, 10 minutes	• Suited to colored, algal-laden, impounded raw water • Can be located in-filter
Tube settler	5	2 stages, 15 minutes	Loading rate depends on spacing and depth of tubes. Often retrofitted to increase plant capacity
Ballasted sand	30	1 stage, 10 minutes	Proprietary process
Sirofloc	10	2 stages, 10 minutes	Proprietary process

Table 3. Granular Media Filter Arrangements

Media	Typical Grading and Depth	Typical Loading Rate	Comment
Monomedia (sand)	600 mm of 0.6 m effective size	6 m/h	Typical old style rapid gravity filter designs
Monomedia (sand or anthracite)	2000 mm of 2 mm effective size	Up to 24 m/h	e.g., 3000 ML/day direct filtration plant at Prospect, Sydney, Australia
Dual media (typically anthracite/sand)	1500 mm of 1.5 mm anthracite on 300 mm of 0.8 mm sand	12–18 m/h	Also used in direct filtration configuration (i.e., no preclarification)
Dual media (typically anthracite/sand)	Wide range of anthracite sand combinations are in use	8–14 m/h	Finer grading used in postclarifier configurations

GUIDE TO PROCESS SELECTION

Table 4 provides a summary guide to selecting the most applicable treatment process, given particular water quality problems. The parameters covered focus on challenges that occur in raw waters. They do not include parameters that are created through treatment and distribution. Therefore, parameters such as disinfection by-products like trihalomethanes (THMs), lead (usually a corrosion product), and soluble aluminum (usually a residual of alum coagulation) are not included.

Table 4 focuses on primary treatment needs and does not consider residual disinfection requirements.

The "primary process candidates" focus on processes that can cost-effectively provide very good removal efficiency of the parameter in question, so that water

Table 4. Water Treatment Process Selection Guide

Water Quality Parameter	Primary Process Candidates	Relative Cost	Comments	Other Process Candidates
		Biological		
Bacteria	Chemical disinfection	0.1	Chlorine, chlorine dioxide, chloramine, and ozone are the main disinfectants. Cost is based on chlorine.	
	Ultraviolet disinfection	0.2	Good pretreatment often required.	
	Microfiltration	1.3		
Virus	Chemical disinfection	0.1		Reverse osmosis
	Ultraviolet disinfection	0.2	Some viruses can be relatively resistant	
Giardia	Benchmark	1.0	Filtered water turbidity of better than 0.3 NTU and preferably 0.1 NTU is required.	Reverse osmosis
	Chemical	0.1	Most practical choices are chlorine, chlorine dioxide, and ozone. Cost is based on chlorine.	
	Ultraviolet disinfection	0.2	Good pretreatment often required.	
	Microfiltration	1.3		
Cryptosporidium	Benchmark	1.0	Filtered water turbidity of better than 0.3 NTU and preferably 0.1 NTU is required.	Reverse osmosis
	Ultraviolet disinfection	0.2	Good pretreatment often required.	
	Microfiltration	1.3		
Algae	Benchmark	1.0	High algae levels needs a suitable clarifier process, e.g., DAF. See "Organics" for algal taste, odor and toxins.	
		Aesthetics		
Turbidity	Benchmark	1.0		Reverse osmosis
	Microfiltration	1.3		
Color	Benchmark	1.0		
	Ozonation	0.5	Practical limit to color removable	
	Nanofiltration	2.0	Pretreatment required	
Hardness	Lime softening	1.2		
	Ion exchange (cation)	0.7	Need also to consider pretreatment requirements, if any	
	Nanofiltration	2.0	Pretreatment required	
Iron	Benchmark	1.0		Lime softening
	Oxidation/filtration	0.8		Activated alumina
Manganese	Oxidation/benchmark	1.1	Potassium permanganate commonly used	Lime softening
	Catalytic filtration	0.8	Chlorination ahead of Mn-coated filter media	
	Benchmark/high pH	1.1	Coagulation at high pH	
Hydrogen sulfide	Aeration	0.3		
		Organics		
DOC	Benchmark	1.0		
	Ion exchange (anion)	0.9		
	Nanofiltration	2.0		
Algal tastes & odors	Granular activated carbon	0.6—1.0	Cost depends on bed life; bed life uncertain: testing required	
	Powdered activated carbon	0.2—0.4	Cost depends on dose and period of use	
	Ozonation	0.5	Sometimes used in conjunction with hydrogen peroxide	
Algal toxins	Chlorination	0.1	Effective on many hepatotoxins	Nanofiltration
	Ozonation	0.5	Sometimes used in conjunction with hydrogen peroxide	

(*continued overleaf*)

Table 4. (*Continued*)

Water Quality Parameter	Primary Process Candidates	Relative Cost	Comments	Other Process Candidates
	Ozone/biological activated carbon	1.0		
	Granular activated carbon	0.6—1.0	Bed life uncertain; testing required	
	Powdered activated carbon	0.2—0.4	Dose setting can be difficult	
Pesticides	Ozonation	0.5	Sometimes used in conjunction with hydrogen peroxide	Reverse osmosis
	Ozone/biological activated carbon	1.0		
	Granular activated carbon	0.5—2	Bed life uncertain	
	Powdered activated carbon	0.2—0.5	Dose setting can be difficult	
Organic chemicals generally	Aeration	0.3—1	Only for highly volatile compounds	
	Granular activated carbon	0.5—2	Cost highly sensitive to bed life; testing required	
		Inorganics		
Antimony	Benchmark	1.0		Reverse osmosis
Arsenic	Oxidation & benchmark	1.1		
	Oxidation & lime softening	1.3		
	Activated alumina	0.9		
	Ion exchange (anion)	0.9		
Barium	Lime softening	1.2		Reverse osmosis
	Ion exchange (cation)	0.7		
Beryllium	Benchmark	1.0		Reverse osmosis
	Lime softening	1.2		
	Activated alumina	0.9		
Boron	Ion exchange (anion)	0.9		
	Reverse osmosis	2.5		
Cadmium	Benchmark	1.1	Ferric coagulation at high pH	Reverse osmosis
	Lime softening	1.2		
Chromium (III)	Benchmark	1.0		
	Lime softening	1.2		
	Ion exchange (cation)	0.9	Anion for Cr(VI)	
	Reverse osmosis	2.5	OK for Cr(VI)	
Cyanide	Oxidation and benchmark	1.1		Reverse osmosis
Fluoride	Activated alumina	0.9		
	Reverse osmosis	2.5		
Mercury (inorganic)	Benchmark	1.0	Ferric coagulant preferred	Reverse osmosis
	Granular activated carbon	0.5—1.0	Cost depends on bed life	
	Lime softening	1.2		
Nitrate	Ion exchange (anion)	0.9		
	Reverse osmosis	2.5		
Sulfate	Ion exchange (anion)	0.9		
	Reverse osmosis	2.5		
Selenium	Ion exchange (anion)	0.9		Reverse osmosis
	Activated alumina	0.9		
Silver	Benchmark	1.0		
	Lime softening	1.2		
TDS	Reverse osmosis	2.5	Brackish water	
	Electrodialysis	3.0		
	Distillation	5.0	Very dependent on energy cost	
		Radiological		
Uranium	Benchmark	1.0	Waste disposal is an issue	
	Lime softening	1.2	Waste disposal is an issue	
	Ion exchange (cation)	0.9	Waste disposal is an issue	
Radon	Aeration	0.3		
	Granular activated carbon	0.5—1.0	Depends on bed life	
Radium	Lime softening	1.2	Waste disposal is an issue	
	Ion exchange (cation)	0.9	Waste disposal is an issue	
	Reverse osmosis	2.5	Waste disposal is an issue	

quality objectives can be met. Processes that provide only limited removal are generally not listed. The focus is also on the most readily available technologies.

The "relative cost" column indicates life-cycle cost in comparison with the "benchmark" treatment process. The cost relativity is indicative only and depends on plant scale, the site and local conditions, and costs. "Other process candidates" list potential processes which would not normally be selected for the water quality parameter in question but are listed to provide guidance if that particular process is also a good primary candidate for another water quality parameter of concern.

As a reference, the expected range of life-cycle cost for the benchmark process in developed countries is of the order of US$0.15 to US$0.40 per kL (2003); the variation depends mainly on scale—the larger scale plant has lower unit costs.

Examples of Application

Example 1: Braided River from Farming Catchment and High Base Flow. A source water drawn from a braided river suffers from elevated turbidity after storms and is suspected of *Cryptosporidium/Giardia* and bacteria/contamination from farming in the catchment.

From Table 4, microfiltration is suitable for all identified contaminants, as is the combination of the benchmark process plus disinfection (either chemical or UV). From this, the prime process candidates are therefore

- benchmark plus disinfection
- microfiltration

Example 2: Highly Turbid Tropical River in Unprotected Catchment. In this case, the water quality parameters of concern and corresponding process selections are

Cryptosporidium: benchmark, UV, MF
Turbidity: benchmark, MF
Color: benchmark, ozonation
Virus: disinfection, UV
Bacteria: disinfection, UV, MF
Iron: benchmark, oxidation/filtration.

Looking at these technology selections suggests selecting the benchmark process plus disinfection would most cost-effectively deal with all of the parameters of concern.

Example 3: Algal-Prone Surface Water Reservoir from a Compromised Catchment. In this case, the parameters of concern are turbidity, algae, DOC, color, *Cryptosporidium*, bacteria, virus, iron, manganese, taste and odor, and toxins.

Using the same procedure as previously highlights the potential to select a process train consisting of the benchmark process plus ozone/BAC.

Example 4: Groundwater. This groundwater has problems of elevated arsenic, iron, and manganese. From Table 4, the benchmark process with oxidation would be an appropriate selection.

Testing the most likely process selections obtained through the guidance process is always important to confirm achievable performance and likely actual costs. In evaluating the final short list of process options, factors like local skills, cost and availability of materials, level of support services, and waste management constraints, need to be taken into account.

READING LIST

EPA. (1990). *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Plants*. Office of Drinking Water, USEPA, Cincinnati, OH.

AWWA. (1994) *Water Quality and Treatment*, 4th Edn. McGraw-Hill, New York.

www.epa.gov

www.usbr.gov

Department of Natural Resources & Environment. (1997). *Low Cost Water Treatment for Small Towns*. Report No. WQ97.001. Melbourne, Australia.

SOURCE WATER ASSESSMENT

U.S. Geological Survey

FREQUENTLY ASKED QUESTIONS

What is a Source Water Assessment?

A source water assessment is a study and report, unique to each water system, that provides basic information about the water used to provide drinking water.

What Will My Assessment Tell Me?

- *Where My Drinking Water Comes From*. Source Water Assessments identify the area of land that most directly contributes the raw water used for drinking water.
- *What Could Pose a Threat to My Drinking Water Quality*. Assessments identify the major potential sources of contamination to drinking water supplies. This information is used to determine how susceptible the water system is to contamination.

Is This Information Available?

The results of an assessment will be provided to the public to help communities plan for protection activities. Information about the assessment will be included in the yearly consumer confidence report you receive from your water utility.

Does My Water System Have a Source Water Assessment?

Every state is moving forward to implement assessments of its public water systems, as required under a new

---

federal Safe Drinking Water Act program called the Source Water Assessment Program (SWAP). Assessments must be completed by 2003 for every public water system—for major metropolitan areas and the smallest towns, including schools, restaurants, and other public facilities that have wells or surface water supplies. Assessments will not be conducted for drinking water systems that have less than fifteen service connections or that regularly serve less than twenty-five individuals, since these are not considered public water systems.

Who is Paying for These Assessments?

Source water assessments and protection measures are eligible uses of the Drinking Water State Revolving Fund (DWSRF) set-asides. States may use the funds for a mixture of source water related local assistance activities. For example, funds are available for Land Acquisition and Conservation Easements.

SOURCE WATER ASSESSMENTS

The source water assessment programs created by states differ since they are tailored to each state's water resources and drinking water priorities. However, each assessment must include four major elements:

- delineating (or mapping) the source water assessment area,
- conducting an inventory of potential sources of contamination in the delineated area,
- determining the susceptibility of the water supply to those contamination sources, and
- releasing the results of the determinations to the public.

These steps are described in more detail below, with information on how citizens and organizations can join in the assessment process. More information about the SWAP program is available. Our SWAP Contact List has state specific contacts and links to State web sites.

Source Water Assessment Using Geographic Information Systems: This document provides guidance to states, municipalities, and public water utilities for assessing source waters using geographic information system (GIS) technology.

STEP 1: DELINEATE THE SOURCE WATER ASSESSMENT AREA

For each ground water well or surface water intake that supplies public drinking water, the land area that could contribute water and pollutants to the water supply must be delineated, or mapped. Significant potential sources of contamination will then be identified in this delineated area during Step 2 of the assessment process.

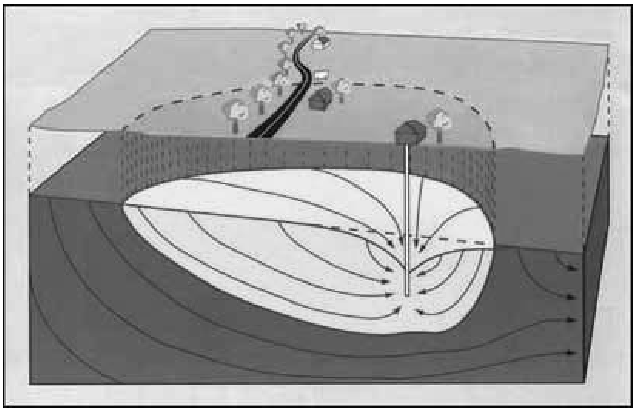

For ground water supplies, states commonly use information about the flow of underground water to delineate source water assessment boundaries. This results in a map of land areas where, if pollutants are spilled or discharged on the surface, they could filter through the soil to the ground water and be drawn into a particular well. Some states may use a simpler mapping approach, by drawing a circle of a certain radius around the well.

For a community that uses surface water from a stream, river, lake, or reservoir, the land area in the watershed upstream of the intake is identified on the map. A watershed boundary is drawn using a topographic map, and includes the land areas where rain or melted snow flows over or through the ground and eventually enters the water source upstream of the water system's intake.

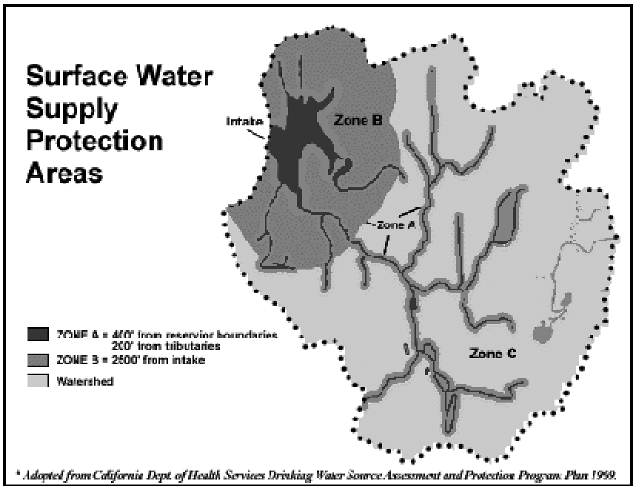

* Adopted from California Dept. of Health Services Drinking Water Source Assessment and Protection Program Plan 1999.

Diagram2

Some states plan to divide the watershed area into segments—areas closest to the intake where most types of contamination sources can impact the water supply, and other more distant areas. The entire watershed up to the state's boundaries is required to be delineated, but the inventory of potential pollution sources may be more detailed in segments that are closer to the intake.

After the state has completed its assessment for a water system, the community may decide to undertake protection efforts for targeted sources of contamination. An initial step could be to expand upon the state's delineation process. Particularly for smaller ground water systems, where states may not have the resources to conduct a detailed delineation, additional scientific methods can be used to more accurately delineate the area that contributes ground water to the well.

Community members can seek assistance from the environmental sciences, geology, or engineering departments of local colleges, or from environmental consulting firms to assist in creating more detailed delineations. Sometimes these services are provided by professors, graduate students, or local firms for a reduced fee or none at all. In addition, local water resource information is often available from other sources such as the federal Natural Resource Conservation Service, the United States Geological Survey, and the state's Cooperative Extension Service.

For more information on delineation see Reliable Sources—Delineation.

STEP 2: CONDUCT AN INVENTORY OF POTENTIAL SOURCES OF CONTAMINATION

Community groups can become especially involved in the second step of an assessment—identifying potential sources of pollutants that could contaminate the water supply. This inventory usually results in a list and a map of facilities and activities within the delineated area that may release contaminants into the ground water supply (for wells) or the watershed of the river or lake (for surface water sources).

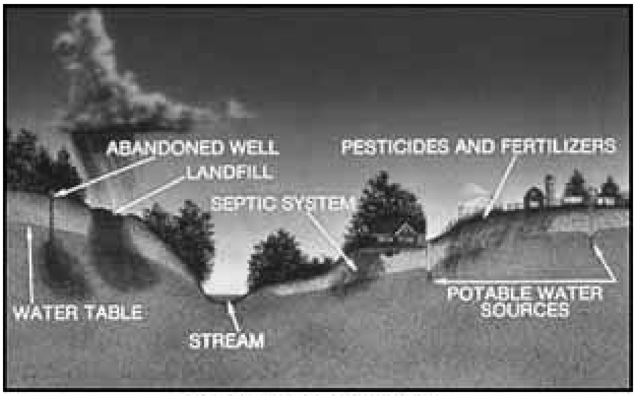

Contamination Sources

Some examples of the many different types of potential pollutant sources include landfills, underground or above-ground fuel storage tanks, residential or commercial septic systems, storm water runoff from streets and lawns, farms that apply pesticides and fertilizers, and sludge disposal sites.

Some states are asking communities to conduct the inventory themselves, in order to obtain detailed information about potential contaminant sources. Others will use computer databases and focus the inventory on land uses and activities that are currently mapped or regulated. Although this approach may not address sources of contaminants that are not currently regulated, such as smaller livestock areas or auto salvage yards, the database inventories could include industries and sewage treatment plants that discharge wastewater, hazardous waste sites, mining operations, particular land use categories (such as industrial, agricultural and urban areas), and various facilities that have environmental permits.

Community groups such as watershed organizations, local environmental committees or scout troops can enhance the state's assessment by conducting site-specific inventories of potential pollutant sources that may not be on state databases or maps. Local inventories may provide information on abandoned dump sites, businesses with septic tanks or floor drains such as dry cleaners or car repair shops, pesticide mixing and storage areas, golf courses, and other land uses that may release pollutants to ground water or surface water. Community groups can coordinate their local inventory with the state's assessment process or can enhance a completed assessment with a more detailed inventory.

A helpful document to aid community groups is EPA's "Drinking Water Contaminant Source Index" which is a list of potential contaminant sources and the pollutants they can release. Additional resources about contamination are available at Reliable Sources—Contaminant Source Inventory.

STEP 3: DETERMINE THE SUSCEPTIBILITY OF THE WATER SUPPLY TO CONTAMINATION

For the susceptibility analysis, the state combines the inventory results with other relevant information to decide how likely a water supply is to become contaminated by identified potential sources of contamination. This critical step makes the assessments useful for communities, since it provides information that local decision-makers may use to prioritize approaches for protecting the

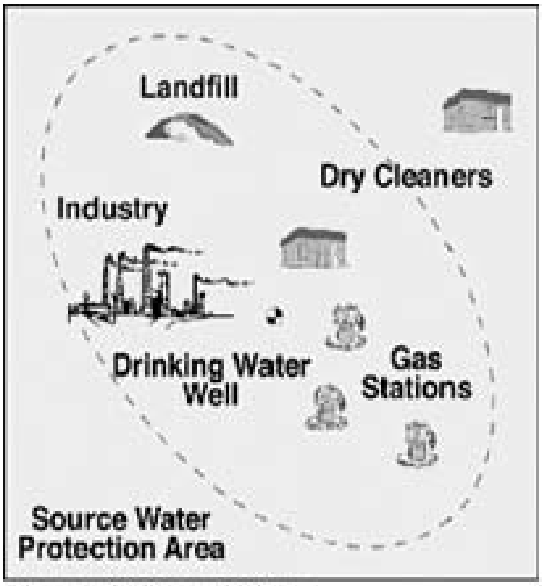

Diagram3: Ground Water

drinking water supply. Local information provided to the state by local community groups about contaminant sources, water resource characteristics, or environmental management practices may be used in the susceptibility determination process. Some states prioritize the potential for contamination from identified potential contamination sources or specific chemicals that could pollute the water. Other states assign susceptibility rankings of high, medium or low to the water sources.

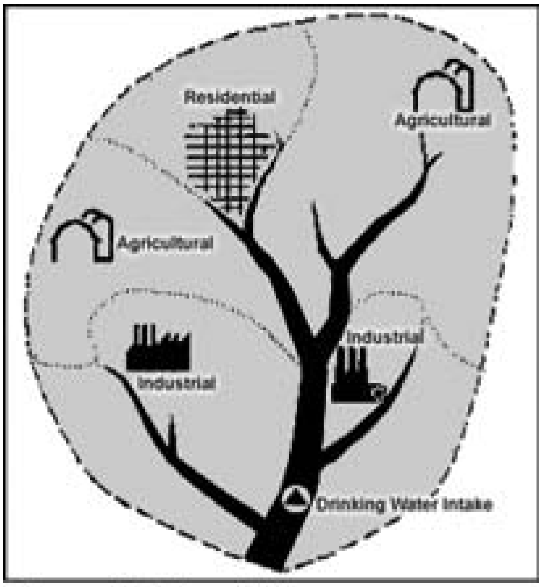

Diagram4: Surface Water

For more information on susceptibility see Reliable Sources—Susceptibility Determination.

STEP 4: RELEASE THE ASSESSMENT RESULTS TO THE PUBLIC

After a state completes the assessment of a particular water system, it will summarize the information for the public. Such summaries help communities understand the potential threats to their water supplies and identify priority needs for protecting the water from contamination. States will make the assessment summaries available to the public in a variety of ways. Some states plan to convene public workshops, while others will have copies available in public libraries and from local government offices or water suppliers. Many also plan to post the assessment summaries on the Internet. The results of the assessments will also be included in the annual water quality reports that community water systems are required to prepare for their customers. Community groups can convene local meetings to discuss the results and begin the process of protecting the drinking water source.

For more information on results see Reliable Sources—Public Availability.

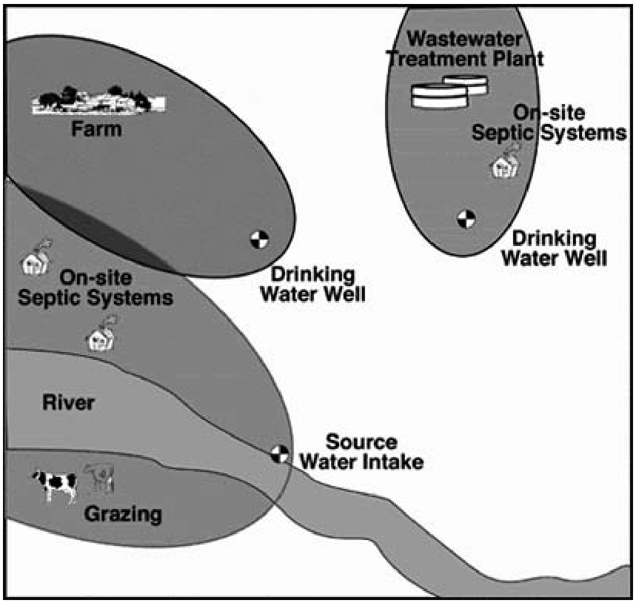

USING THE ASSESSMENT

Whether using the state's assessment or expanding it into a more detailed local assessment, communities can use information gathered through the assessment process to create a broader source water protection program. Community groups and local officials, working in cooperation with local, regional, and state government agencies can plan how to manage identified potential contamination sources and prevent new contaminant threats in the source water assessment area.

Communities use a wide array of different source water protection methods to prevent contamination of their drinking water supplies. One management option involves regulations, such as prohibiting or restricting land uses that may release contaminants in critical source water areas. Along with regulations, many communities hold local events and distribute information to educate and encourage citizens and businesses to recycle used oil, limit their use of pesticides, participate in watershed cleanup activities, and a multitude of other prevention activities. Another aspect of a source water protection program can be the purchase of land or creation of conservation easements to serve as a protection zone near the drinking water source. For an effective protection program, communities should consider using a variety of prevention measures.

EPA's fact sheet, Community Involvement in Source Water Assessments, explains the four steps of source water assessments; how communities can participate in the assessment process; and how communities can utilize assessment information for source water protection. A downloadable version (PDF, 180 kb) is available.

HELPFUL CONTACTS

EPA, State and other contacts are also available to help you.

Your local water supplier may also have more information about opportunities to become involved in the source water assessment process. EPA can provide some local information about your drinking water system. You can also call the number on your water bill or contact your local health department for information on your water supplier.

HYDRAULIC DESIGN OF WATER DISTRIBUTION STORAGE TANKS

Rasheed Ahmad
Khafra Engineering Consultants
Atlanta, Georgia

Water distribution storage ensures the reliability of supply, maintains pressure, equalizes pumping and treatment rates, reduces the size of transmission mains, and improves operational flexibility and efficiency. Numerous decisions must be made in designing a storage tank, including size, location, type, and expected operation. There are several key considerations in the hydraulic design of water storage tanks.

Equalization is the primary purpose for constructing water distribution storage tanks. Water utilities like to operate treatment plants at a relatively constant rate, and wells and pumping stations generally work best when pumped at a steady rate. However, water use in most utilities varies significantly during the course of the day. These variations in use can be met by continuously varying source production, continuously varying pumping rates, or filling and draining storage tanks.

Filling and draining storage tanks is much easier operationally and is generally less expensive than other methods. Facilities that serve portions of a distribution system that has storage tanks generally need to be sized only to meet maximum daily demands; storage tanks provide water during instantaneous peak demands. Maintaining adequate pressure in a water distribution system is very important. To a great extent, the elevation of water stored in a tank determines the pressure in the pipes directly connected to the tank. The larger the tank volume, the more stable the pressure in the distribution system, despite fluctuations in demand or changes in pump operation. If distribution storage were not used, larger water transmission mains and larger treatment plant capacity would be required by most utilities to provide water needed for fire fighting. Especially in smaller systems, storage tanks are a much more economical and operationally reliable means for meeting the short-term large demands placed on a water supply system during fire fighting. In addition to fires, emergencies such as power outages, breaks in large water mains, problems at treatment plants, and unexpected shutdowns of water supply facilities can cause failure of the water system if sufficient water is not available in storage. Water storage tanks meet demands during emergencies.

Figure 1. Ground-level water storage tank (*Source*: Columbian TecTank).

Figure 2. Traditional multicolumn elevated storage tank (*Source*: Caldwell Tank, Inc.)

Besides equalization, pressure maintenance, and fire and emergency storage, other factors such as energy consumption, water quality, hydraulic transient control, and aesthetics should be considered in designing water distribution storage tanks. Tanks enable utilities to store energy as well as water for later use. To the extent that equalization storage slows down the velocity (and friction losses) in large transmission mains, the energy used to pump water is reduced by having distribution storage tanks that equalize pumping. Tanks may affect water quality in two general ways: (1) through chemical, physical, and biological processes that occur as water ages while stored in the tank and (2) through external contamination of water in tanks.

Figure 3. Composite elevated storage tank (*Source*: Caldwell Tank, Inc.).

Figure 4. Pedesphere elevated storage tank (*Source*: Caldwell Tank, Inc.).

Some water quality problems can be eliminated by proper design and maintenance. In recent years, because of the concern for microbial regrowth and formation of disinfection byproducts, protecting the water quality in the storage tanks has become increasingly important. Another important consideration in designing water distribution storage tanks is controlling hydraulic transients. Changes in velocity in water mains can result in "water hammer." A water storage tank is effective in dampening the extreme high or low pressure caused by a transient. Sometimes the aesthetics of storage tanks becomes important. Tanks should be designed to satisfy the aesthetic considerations of stakeholders without sacrificing the purpose of the tank and the efficiency of the system's operation.

Depending on the topography and local environmental conditions, storage tanks may be located above, on, or below the ground. Underground tanks or reservoirs are constructed of reinforced concrete. Small ground-level tanks are usually earth-lined with gunite, asphalt, or some synthetic membrane. Large surface tanks are concrete-lined (Fig. 1). Most large surface tanks are covered to prevent contamination by bird, animal, and human wastes. Elevated water storage tanks are usually constructed of steel or concrete. Common shapes for elevated storage tanks are illustrated (Figs. 2–4).

READING LIST

HDR, Inc. (2001). *Handbook of Public Water System*, 2nd Edn. John Wiley & Sons, New York.

May, L.W. (Ed.). (1999). *Water Distribution System Handbook*. McGraw-Hill, New York.

Peavy, H.S., Rowe, D.R., and Tchobanoglous, G. (1985). *Environmental Engineering*. McGraw-Hill, New York.

Qasim, S.R., Motley, E.M., and Zhu, G. (2000). *Water Works Engineering*. Prentice-Hall, Upper Saddle River, NJ.

SYSTEM CONTROL AND DATA ACQUISITION (SCADA)

Zacharia Michael Lahlou
Technical Assistance
Consultant

As small water systems continue to struggle to become more efficient, technology provides good ways to improve operations. Nowhere is this more apparent than in the application of control and communications systems—namely SCADA [System Control and Data Acquisition]—and their ability to provide system and information management that can help utilities meet regulatory requirements, reduce operating costs, and improve customer service. A water distribution system transports water from the treatment facility to the user. The distribution system should supply water, without impairing its quality, in adequate quantities and at sufficient pressures to meet system requirements

WHAT IS SCADA?

A SCADA system consists of a computer placed at a central location, communications equipment, programmable logic controllers, sensors, and other devices that when put together, will monitor and control equipment and processes in a utility, such as a water system. Remote sites, equipped with remote telemetry units (RTUs), can be located as close as within the same building or as far away as across the country. RTUs measure a variety of conditions and parameters, including tank levels, temperature, voltage, current, volumes, and flow rates, and report back to a central processing unit (CPU). This technology is widely accepted in the water and wastewater industries as a reliable and efficient control system.

SCADA EQUIPMENT ENHANCES WATER OPERATIONS

Functions that the SCADA system can perform include: remote monitoring of well levels and control of their pumps, and monitoring flows, tank levels, or pressures in storage tanks. A SCADA system can also monitor water quality characteristics, such as pH, turbidity, and chlorine residual, and control the addition of chemicals. In the distribution system, SCADA can supervise and control the water pressure of networks, assure water pressure is uniformly distributed, lower the leakage rate, and store data for future analysis.

SCADA is not a new technology by any means, but significant innovations and improvements have been achieved since its introduction. By automating many routine tasks, a SCADA system frees the plant operator to perform other duties, such as addressing state and federal reporting requirements. In fact, SCADA can assist in mandated reporting because it can store various activities and information on the computer. Graphs and reports can be generated automatically using the data collected remotely from the field. These reports are important in inferring production and consumption patterns, data that help manage the water resources more efficiently. Stored information also proves invaluable when producing the annual Consumer Confidence Reports required under the 1996 Safe Drinking Water Act (SDWA) Amendments.

SCADA is useful in an emergency situation. Immediately following an incident, an operator can be notified via personal pager and increase response capability dramatically. SCADA provides multipurpose utility management, operating flexibility and more complex system control. SCADA keeps an eye on the entire system from one place.

SCADA PROVIDES COST-EFFECTIVE CONTROL AND MONITORING FOR SMALL WATER SYSTEMS

Because of its cost, smaller systems often viewed SCADA as a luxury item. But water system management has become complex and SCADA has become more advanced yet, paradoxically less expensive, making SCADA a viable option. Initial installation costs, which are often high, typically pay for themselves in a short time through direct labor and vehicle cost savings, as well as increased efficiency.

Due to a low tax base, many small communities have limited financial resources for drinking water system management and operation. Thus, it is difficult for operators to maintain extensive manual monitoring to ensure the system complies with complex regulations. These small systems, therefore, frequently violate (monitoring/reporting violations) the SDWA and its amendments, according to the National Research Council.

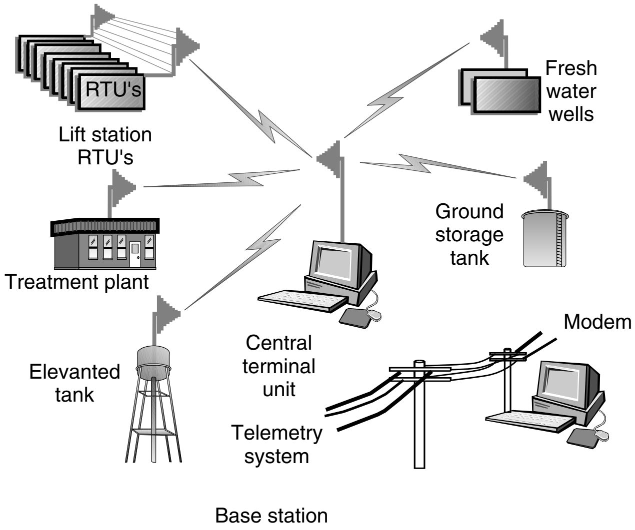

SCADA system

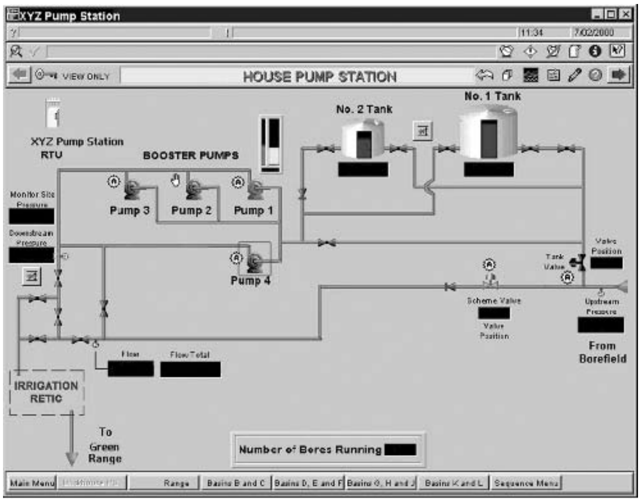

SCADA allows water operators to monitor the entire distribution system from one location

One solution to keeping up with ever more stringent requirements is for several small communities to pool their resources. The U.S. Environmental Protection Agency is evaluating the option of using remote telemetry, called an “electronic circuit-rider,” that allows one qualified operator to monitor and control the operation of several small treatment systems from a centrally located computer. Using such a system can optimize the time spent taking daily readings out in the field. For onsite inspection and maintenance efforts, the RTS allows the operator to visit only the problematic systems. The results expected from an appropriately designed and successfully deployed remote monitoring and control system include enhanced water quality, compliance with existing water quality regulations, and reduced operating and maintenance costs for small communities.

SECURITY MEASURES ARE A CONCERN

Experienced computer hackers can access SCADA systems that operate over the Internet and deactivate process alarms, change chemical parameters, start and stop equipment, and so on. The results of such an intrusion can be devastating to the general public and to plant employees.

Many system managers installed protection as part of their “Y2K” contingency plans. The likelihood of a terrorist cyber attack is less likely than a cyber attack from a disgruntled employee. Fortunately, there are steps a system can take to minimize a security breech.

SCADA COMPONENTS AND TERMINOLOGY

Distributed Control System (DCS): An integrated system made up of many subsystems that are remotely located. Each subsystem can operate independently.

Modem (MOdulator DEModulator): Used to convert signals in one form to another. This is generally used for communication between computers and other devices over telephone lines or radio.

Programmable Logic Controller (PLC): A microprocessor-based controller, usually with multiple inputs and outputs and a program to perform control functions

Remote Telemetry Unit or Remote Terminal Unit (RTU): A microprocessor device with multiple inputs and outputs connected to field instruments and devices. The RTUs translate these signals to digital form and transmit the same to central location by radio or telephone lines.

The National Infrastructure Protection Center (NIPC), located in the Federal Bureau of Investigations headquarters in Washington, D.C., shares information with public and private sector owners and operators of critical infrastructures. The NIPC monitors, warns, and investigates unlawful acts involving computer and information technologies. The agency manages computer intrusion investigations and supports law enforcement related to cyber crimes.

Best practices regarding the security of SCADA operation include:

- Internal threats are usually the main security challenge. The key to managing internal threats is understanding who might do what and why.
- Log-ins should be traceable and a strong password authentication process used.
- Suggest that a utility consider fiber optics to each of its remote facilities. This is the best solution in terms of bandwidth and security.
- Suggest taking a look at commercially available tools for monitoring computers and having a security team review activity weekly to verify that only authorized users are accessing the system.

WHAT SHOULD BE CONSIDERED WHEN BUYING A SCADA SYSTEM?

Selecting the appropriate SCADA system software is all-important. The software program should be capable of performing every task needed to operate and maintain water treatment and the distribution system. Its use in many similar applications should prove its reliability. Make sure the hardware supplier and the software developer can guarantee prompt, efficient, and cost effective support. The manufacturer should have experience in the water and wastewater industry and be able to provide service, replacement parts, and support for the system when needed.

John Barkey, plant operator for Martin’s Ferry, Ohio, inspects the water levels of the system’s storage tanks from his desk

To expedite the selection of specific devices/components for an RTS that monitors and controls a small drinking water facility, prepare a list of features associated with the particular water treatment facility. Identify the water quality parameters and types of monitoring that are key to the specific operation. For example, monitor residual chlorine to verify disinfection operations. Next, review the regulatory compliance requirements and consider them when reviewing manufacturers’ specifications and discussing the applicability of their device or system. Finally, select the monitoring device.

It is important to document the pump characteristics, operational functions, and the physical dimensions of the treatment system before selecting components to monitor flow rates, pressure, electrical usage, or other process functions. Know facility specifications when contacting technical representatives of process monitoring/control devices. Component selection must be based on the plant’s operational characteristics.

Data acquisition and telemetry components can be purchased as package items, but again, identify an inventory of what devices will be used on line and what options exist for transmitting data in the locality of the treatment plant before contacting suppliers.

FOR MORE INFORMATION

Booth, R., D. Ryan, and C. Hewell. 2001. *Technical Security and Countermeasures White Paper for Water Utilities*. Washington, DC: The National Council for Public-Private Partnerships.

Lebeau, L. W. 1997. “Automation Project Floods Water Department with Useful Data.” *Water Engineering* *and Management*. Des Plaines, IL: Scranton Gilette Communications.

Madabhushi, B. S. 2000. “Can SCADA benefit small systems?” *Water Sense*. 6, no. 2. Morgantown, WV: National Drinking Water Clearinghouse.

National Research Council. 1997. *Safe Water From Every Tap*. Washington, DC: National Academy Press.

Pollack, A. J., A. S. Chen, R. C. Haught, and J. A. Goodrich. 1999. *Options for Remote Monitoring and Control of Small Drinking Water Facilities*. Columbus, OH: Batelle Press.

Rossmiller, R., and M. C. Moaveni. 2000. “Water Utility Mixes Technology and Creativity for Control and Communications Solution.” Water Engineering and Management. Des Plaines, IL: Scranton Gilette Communications.

To learn more about the National Infrastructure Protection Center (NIPC), write to the NIPC, 935 Pennsylvania Avenue, NW, Washington, D.C. 20535-0001, call toll free (888) 585–9078, or e-mail nipc.watch@fbi.gov. *Information is also available on the NIPC Web site at* www.nipc.gov.

For further information, to comment on this fact sheet, or to suggest topics, contact Dr. Lahlou via e-mail at *lahloum@hotmail.com*.

Zacharia M. Lahlou holds a doctorate in Environmental and Natural Resource Economics, as well as an MBA and a master’s degree in Civil and Environ-mental Engineering from West Virginia University. Now a consultant, he formerly served as technical assistance coordinator for the National Drinking Water Clearinghouse.

HAVE YOU READ ALL OUR TECH BRIEFS?

Tech Briefs, drinking water treatment and supply fact sheets, have been a regular feature in the National Drinking Water Clearinghouse (NDWC) publication *On Tap* for more than six years.

A package of *Tech Briefs* is now available as a product. A three-ring binder holds all the current *Tech Briefs* in print. New selections can easily be added to the package as they become available.

To order this product, call the NDWC at the numbers listed at the bottom of this ad and ask for item #DWPKPE71. **The item is free.** Additional copies of fact sheets are also free; however, postal charges may be added. You can also order copies of one or all of the free *Tech Briefs* listed below.

Tech Brief: Disinfection, item #DWBLPE47;

Tech Brief: Filtration, item #DWBLPE50;

Tech Brief: Corrosion Control, item #DWBLPE52;

Tech Brief: Ion Exchange and Demineralization, item #DWBLPE56;

Tech Brief: Organics Removal, item #DWBLPE59;

Tech Brief: Package Plants, item #DWBLPE63;

Tech Brief: Water Treatment Plant Residuals Management, item #DWBLPE65;

Tech Brief: Lime Softening, item #DWBLPE67;

Tech Brief: Iron and Manganese Removal, item #DWBLPE70;

Water Conservation Measures Fact Sheet, item #DWBLPE74;

Tech Brief: Membrane Filtration, item #DWBLPE81;

Tech Brief: Treatment Technologies for Small Drinking Water Systems, item #DWPSPE82;

Tech Brief: Ozone, item #DWBLPE84;

Tech Brief: Radionuclides, item #DWBLPE84;

Tech Brief: Slow Sand Filtration, item #DWBLPE99;

Tech Brief: Ultraviolet Disinfection, item #DWBLPE101;

Tech Brief: Leak Detection and Water Loss Control, item #DWBLPE102;

Tech Brief: Diatomaceous Earth Filtration for Drinking Water, item #DWBLPE108.

SETTLING TANKS

Nikolay Voutchkov
Poseidon Resources Corporation
Stanford, Connecticut

INTRODUCTION

Settling tanks (clarifiers) are an integral part of every wastewater treatment plant. These treatment facilities are used to remove solids from the wastewater by gravity sedimentation in quiescent conditions. Each settling tank has two functional zones—a clarification zone, where the process of gravity sedimentation occurs, and a thickening zone where the settled solids are accumulated forming a dense layer of sludge (sludge blanket). Settling tank effluent of low solids concentration is collected from the top of the clarification zone over overflow weirs into collection channels which convey the effluent to the tank outlet. The sludge collected at the bottom of the settling tank is removed for further treatment at the wastewater treatment plant’s solids handling facilities. The depth of the clarification zone is commonly referred to as a clear water zone (CWZ) depth, while the depth of the zone of sludge accumulation is named sludge blanket depth (SBD). The sum of the CWZ depth and the SBD is typically defined as a sidewater depth (SWD).

SETTLING TANK TYPES

Depending on their function settling tanks are categorized as primary and secondary. Primary settling tanks are located downstream of the wastewater treatment plant headworks and their main purpose is to remove the settleable suspended solids in the plant influent. Typically, primary settling tanks are also equipped with devices for removal of the floatable compounds (i.e., scum, oil and grease) in the wastewater influent which accumulate on the surface of the tanks during the sedimentation process. Secondary tanks are located downstream of the biological (secondary) treatment facilities of the wastewater treatment plant (such as activated sludge aeration basins or trickling filters) and are used to separate the biomass generated during the secondary treatment process from the treated plant effluent.

Depending on their geometrical shape, both primary and secondary clarifiers are classified in two main categories: rectangular and circular. Clarifier shape most suitable for a given application depends on a number of factors and has to be selected based on a cost-benefit analysis. Table 1 summarizes key advantages and disadvantages of rectangular and circular settling tanks.

Rectangular Settling Tanks

Rectangular settling tanks are long concrete structures which consist of individual basins (units) with common inner walls, and inlet and outlet channels (Fig. 1). Each individual tank basin is equipped with a separate sludge collection mechanism which transports the solids settled in the tank into a hopper for withdrawal or has a sludge suction collection mechanism which sweeps and removes solids accumulated at the tank's bottom. The length-to-width ratio of the individual tank basins is usually 3:1 to 15:1 (Fig. 1). The minimum settling tank length from inlet to outlet is typically 3 meters (10 ft). Tank depth is most frequently between 2 to 6 meters (6.6 to 20 ft). Rectangular tank unit width is usually selected based on available standard sizes of sludge collection mechanisms and varies between 2 to 6 meters (6.6 to 20 feet).

Circular Settling Tanks

Circular settling tanks are round facilities which consist of inlet structure, cylindrical clarification zone, conical sludge accumulation zone and effluent weirs (Fig. 2). The effluent weirs of these tanks are placed near the facility perimeter to create radially-directed flow pattern from the tank center towards the walls. The slope of the bottom

Table 1. Comparison of Rectangular and Circular Settling Tanks

Item	Rectangular Settling Tanks	Circular Settling Tanks
Advantages	Less land required for construction of multiple units. Potential construction cost savings because of use of common walls between individual tanks. Longer flow path minimizing short-circuiting. Higher effluent weir loading rates acceptable. Better sludge thickening.	Shorter detention time for settling sludge favoring use as secondary clarifiers. More simple sludge collection system. Easier to accommodate in-tank flocculation chamber—a benefit for activated sludge settling. Overall, lower maintenance requirements. Easier to remove heavy sludge.
Disadvantages	Longer detention time of the settled sludge—not favorable for plants with septic wastewater influent. Less effective for high solids loading conditions.	Higher short-circuiting potential. Higher flow distribution headlosses. Small circular tanks require more yard piping than do rectangular tanks of similar size.

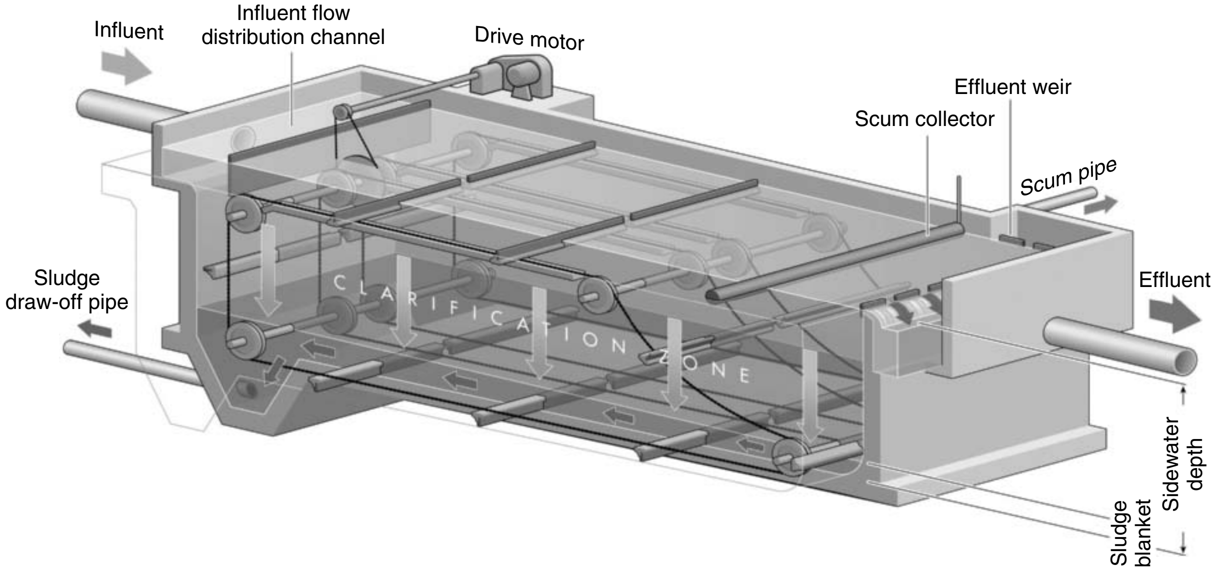

Figure 1. Rectangular settling tank.

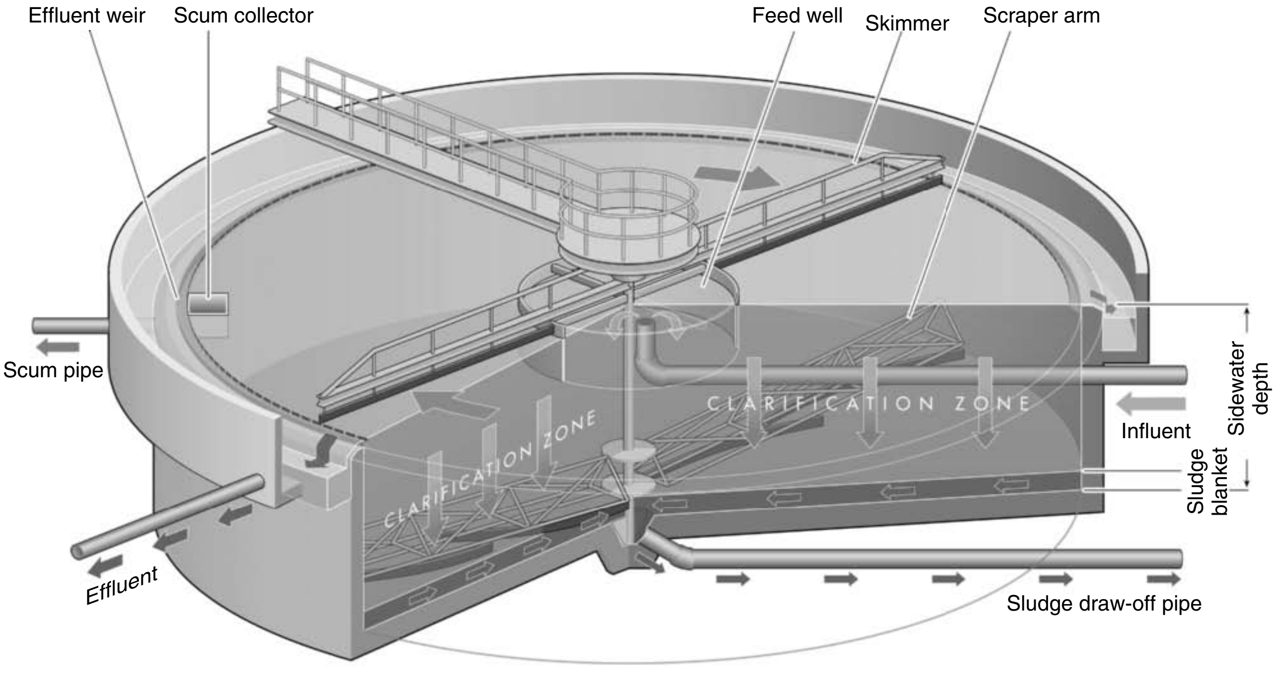

Figure 2. Circular settling tank.

conical floor is usually 1:10 to 1:12 and depends on the type of the sludge collection mechanism. Tank diameter ranges from 3 meters (10 ft) to over 100 meters (300 ft). Circular settling tanks are typically built in pairs of 2 or 4 to simplify the influent flow distribution between the individual units. Circular tank sidewater depth varies from 2.5 to 5 meters (8 to 16 feet).

Depending on the configuration of the tank inlet, circular settling thanks are classified as center feed and peripheral feed. Currently, the most widely used circular tanks are center feed type (see Fig. 2). In these tanks, influent flow enters though a feed pipe located in the center of the tank into a feedwell. The purpose of the feedwell is to provide uniform radial distribution of the tank influent and to dissipate the energy of the feed stream to a level adequate for efficient quiescent settling and uniform radial flow distribution. The conventional feedwells most widely used today are cylindrical metal structures with a diameter of 15 to 25% of the tank diameter which extend to 30 to 75% of the tank sidewater depth. Usually, conventional feedwells are designed for an average downflow velocity of 10 to 13 mm/s (2.0 to 2.5 ft/min) and maximum velocity of 25 to 30 mm/s (5.0 to 6.0 ft/min).

Settling Tank Enhancements

Inclined Plates and Ballasted Flocculation. Inclined plates and ballasted flocculation are used predominantly to enhance the performance of primary rectangular settling tanks. A typical inclined plate (lamella) system consists of bundles of parallel plastic tubes or metal plates inclined at 45 to 60° which are installed at the surface of the settling tank to a vertical depth of approximately 2 meters (6 feet). The distance between the individual plates is between 40 and 120 mm.

Ballasted flocculation combines the addition of coagulant and settling ballast (usually fine sand or sludge) to the tank influent with the installation of inclined plates in the tanks. A portion of the settled sludge or the recovered ballast is recycled to the primary clarifier influent to seed the influent. The addition of ballast increases the density of the influent floc particles by agglomeration. This enhancement typically yields three to five fold increase of the allowable clarifier surface overflow rate (SOR). Typically, conventional settling tanks are designed for SOR of 33 to 49 $m^3/m^2\cdot$ day (800 to 1,200 $gal/ft^2\cdot$ day). The use of high-rate ballasted solids separation technology allows increasing design clarifier SOR to at least 160 $m^3/m^2\cdot$ day (4,000 $gal/ft^2\cdot$ day). Because the ballast enhances solids removal, its use in primary clarification reduces the solids and organic loading of the downstream biological treatment processes.

Flocculating Center Feed Well. Flocculating center feed wells are used to enhance the performance of secondary settling tanks used for clarification of activated sludge. As compared to a conventional center feedwell which radius is approximately 10 to 13% of the tank radius, the flocculating feedwell's radius extends to 20 to 50% of the tank radius and the well size is designed to obtain a detention time of 20 to 30 minutes. The flocculating feedwell typically extends down to 40 to 50% of the tank depth. Some designs also include installation of mechanical mixers in the feedwell to enhance the flocculation process. The flocculating feedwell enhancement aims at creating optimum conditions for coagulation and flocculation of the incoming solids with the return activated sludge (RAS) recycled to the sedimentation tank. In the feedwell, the larger-size recycled RAS particles are given an ample time to attract and flocculate the smaller-size activated sludge particles conveyed form the aeration basins, thereby creating stronger and heavier solids particles that settle better and faster. More detailed design considerations for circular clarifiers with flocculating feedwells as well as a number of other available sedimentation tank process and equipment enhancements are presented elsewhere (1–3).

DESIGN CONSIDERATIONS FOR PRIMARY SETTLING TANKS

Performance efficiency of the primary settling tanks is affected by the upstream wastewater collection and treatment facilities and has a significant impact on downstream biological treatment and solids handling facilities. Primary settling tank performance is typically measured by tank's total suspended solids (TSS), biological oxygen demand (BOD), and phosphorus removal efficiencies and by the condition of the primary sludge (sludge septicity, concentration, and volume). Adequately designed and operated conventional primary settling tanks treating municipal wastewater typically remove 50 to 65% of the influent TSS; 25 to 35% of the influent BOD and 5 to 10% of the influent nitrogen and phosphorus. Clarifier TSS, BOD and nutrient removal efficiencies could be improved by chemical coagulation and flocculation of the influent wastewater solids prior to sedimentation (2–4).

Key design criteria used for sizing primary settling tanks are surface overflow rate and hydraulic detention time. Recommended values for these criteria according to various design guideline sources are presented in Table 2. Typically, primary sedimentation tanks are designed for effluent weir loading rates of less than 190 m^3/day per meter of length of the weir (5,000 gpd/ft) (Table 2).

Proper primary settling tank sludge collection, removal and withdrawal are of key importance for maintaining consistently high primary effluent quality and efficient and cost-effective solids handling. If primary settling tank sludge is retained for excessively long time in the tanks, the sludge could easily turn septic. Sludge septicity is accompanied with release of malodorous gases, which may disturb the normal sedimentation process as they travel from the tank bottom to the surface. Septic sludge is also more corrosive and more difficult to pump and dewater. Besides creating conditions for sludge septicity, maintaining relatively deep sludge blanket in the primary settling tanks may also make sludge collection and withdrawal more difficult and in extreme conditions, may cause damage of the sludge collection and withdrawal equipment (broken sludge collectors, plugged solids lines, and damaged pumps).

A widely accepted practice to prevent primary sludge septicity and its negative effect on settling tank performance is not to carry a sludge blanket, which is achieved by removing sludge continuously or very frequently from the settling tank's bottom. When not controlled appropriately, continuous sludge removal often results in pumping large quantities of diluted sludge or wastewater to the downstream solids handling facilities, which has a negative effect on their performance. In order to avoid over-pumping of diluted sludge to the downstream solids handling facilities and prevent the negative effects of excessively deep sludge blanket and associated sludge speticity, primary settling tank sludge blanket and concentration have to be maintained at optimum levels. The optimum primary sludge concentration is usually 3 to 5 percent and the most viable sludge blanket depth is typically between 1 and 3 feet. The optimum sludge blanket depth would vary seasonally and change during dry-weather and wet-weather conditions.

Table 2. Key Design Criteria for Primary Sedimentation Tanks

Design Guideline Source	Surface Overflow Rate (m^3/m^2/day)	Hydraulic Detention Time (hrs)
Metcalf and Eddy (5) (Primary Settling Followed by Secondary Treatment)	32–48 (at average flow) 80–120 (at peak hourly flow)	1.5–2.5
Randall et al. (4)	For SWD of 1.83–3.05 m: $\leq 2.184 \times \mathrm{SWD}^2$ (at average flow) $\leq 4.368 \times \mathrm{SWD}^2$ (at peak hourly flow). For SWD of 3.05–4.57 m: $\leq 6.672 \times \mathrm{SWD}$ (at average flow) $\leq 13.344 \times \mathrm{SWD}$ (at peak hourly flow)	NA
Great Lakes (6)	$\leq$ 40 (at average flow) $\leq$60 (at peak hourly flow) Tank surface area is determined based on the larger of the two SORs. Minimum SWD = 2.1 m	NA
Qasim (7)	30–50 (at average flow) 40 (typical at average flow) 70–130 (at peak hourly flow) 100 (typical at peak hourly flow)	1.0–2.0

Note: SWD—Sidewater Depth; 1 m^3/m^2/day = 24.542 gpd/ft^2.

DESIGN CONSIDERATIONS FOR SECONDARY SETTLING TANKS

The performance of the secondary settling tanks has a significant effect on the wastewater plant's effluent water quality, on the operational efficiency of the biological treatment system and on the solids handling facilities. The secondary clarifiers have two key functions: clarification of the biologically treated wastewater; and thickening and storage of the sludge from the biological treatment process. Main factors that impact secondary settling tank performance are: (1) the amount of solids retained in the tanks, which is determined based on the concentration of the solids removed from these tanks (return activated sludge (RAS)/waste activated sludge (WAS) concentration) and the sludge blanket depth; (2) the amount of solids in the aeration basins, which is established by measuring the MLSS concentration and the RAS flowrate; (3) the activated sludge settleability; and (4) the plant influent flow and waste load, significant fluctuations of which may result in shifting solids between the settling tank and the aeration basin, and ultimately in solids loss with the secondary settling tank effluent. The two key secondary settling tank design criteria are: the SOR; and the solids loading rate (SLR). Table 3 presents recommendations for determining secondary clarifier design SOR and SLR. The tank effluent weir loading rates are typically designed not to exceed 124 m^3/day per meter of length of the weir (10,000 gpd/ft).

The maximum allowable SLR of settling tanks for clarification of activated sludge could be determined for using solids flux analysis (2,3,5,8). This method is based on the fact that for an activated sludge of given settleablity, these is a maximum amount of solids that can be processed through the clarifier (limiting solids flux), above which the clarifier will not be able to operate in a steady-state condition in terms of sludge blanket elevation and effluent water quality. One of the main benefits of the solids analysis concept is that it allows to link the design and operation of the secondary clarifier and the aeration basin and to optimize their performance as one system.

The amount of solids retained in the sedimentation basins can be effectively monitored by frequent manual or automated measurements of the settling tank sludge blanket depth and the concentration of the sludge removed from the settling tanks. While keeping track of the sludge blanket and plant influent flow changes allows gaining a general understanding of the settling tank performance, it is also very advantageous to monitor sludge settleability as well. The Water Environment Research Foundation and the Clarifier Research Committee of the American Society of Civil Engineers have developed protocols for evaluating sludge settleablity and analyzing secondary clarifier performance (9). These protocols are suitable for operational assessment of existing secondary sedimentation tanks and for planning of new facilities.

BIBLIOGRAPHY

1. Parker, D.S. (1983). Assessment of secondary clarifier design concepts. *J. Water Pollution Control Federation* **55**: 349–356.

Table 3. Key Design Criteria for Secondary Sedimentation Tanks

Design Guideline Source	Surface Overflow Rate (m^3/m^2/day)	Solids Loading Rate (kg/m^2/h)
Metcalf and Eddy (5)		
Settling After Air-Activated Sludge (Excluding Extended Aeration).	16–32 (at average flow) 40–50 (at peak hourly flow)	4–6 (at average flow) ≤14 (at peak hourly flow)
Settling After Extended Aeration.	8–16 (at average flow) 24–32 (at peak hourly flow)	1–5 (at average flow) ≤7 (at peak hourly flow)
Settling After Trickling Filtration	16–24 (at average flow) 40–50 (at peak hourly flow)	3–5 (at average flow) ≤8 (at peak hourly flow)
Randall et al. (4)	For CWZ of 1.83–3.05 m: ≤2.184 × CWZ (at average flow) ≤6.672 × CWZ (at peak hourly flow). For CWZ of 3.05–4.57 m: ≤4.368 × CWZ (at average flow) ≤13.344 × CWZ (at peak hourly flow) Minimum SWD = 4.5 m	≤5(at average flow)
Great Lakes (6)		
Settling After Air-Activated Sludge (Excluding Extended Aeration).	≤49 (at peak hourly flow)	≤10 (at peak hourly flow)
Settling After Extended Aeration	≤41 (at peak hourly flow) Minimum SWD = 3.7 m	≤10 (at peak hourly flow)
Qasim (7)	≤15 (at average flow) ≤40 (at peak hourly flow)	≤2 (at average flow) ≤6 (at peak hourly flow)

Note: CWZ—Clear Water Zone Depth; 1 m^3/m^2/day = 24.542 gpd/ft^2; 1 kg/m^2/h = 0.2048 lb/ft^2/h.

2. Water Pollution Control Federation. (1985). *Manual of Practice FD-8, Clarifier Design*.
3. Water Environment Federation. (1998). *Manual of Practice 8, Design of Municipal Wastewater Treatment Plants*. Vol. 2, 4th Edn.
4. Randall, C.W., Barnard, J.L., and Stensel, H.D. (1992). *Design and Retrofit of Wastewater Treatment Plants for Biological Nutrient Removal, Water Quality Management Library*. Vol. 5, Technomic, Lancaster, PA.
5. Metcalf and Eddy. (1991). *Wastewater Engineering, Treatment Disposal and Reuse*. 3rd Edn. McGraw-Hill, New York.
6. Great Lakes—Upper Mississippi River Board of State Sanitary Engineering Health Education Services, Inc. (1978). *Recommended Standards for Sewage Works*. Albany ("Ten States Standards"), NY.
7. Qasim, S.R. (1985). *Wastewater Treatment Plants Planning, Design and Operation*. Holt, Rinehart and Wilson, New York.
8. Ekama, G.A., Barnard, J.L., Guntert, F.W., Krebs, P., McCorquodale, J.A., Parker, D.S., and Wahlberg, E.J. (1997). *Secondary Settling Tanks: Theory, Modeling, Design and Operation*. International Association on Water Quality. Scientific and Technical Report No. 6.
9. Wahlberg, E.J. (2001). *WERF/CRTC Protocols for Evaluating Secondary Clarifier Performance*. Water Environment Research Foundation.

TREATMENT FOR TECHNOLOGIES FOR SMALL DRINKING WATER SYSTEMS

National Drinking Water
Clearinghouse

INTRODUCTION

Small systems still face difficulties in meeting the requirements of the Safe Drinking Water Act (SDWA) because many technologies available to large systems may be too expensive or complicated for small systems to consider. Furthermore, trained operators and maintenance personnel may not always be available or affordable, leading to standards violations.

Overview of Some Treatment Technologies Used by Small Systems

When the SDWA was reauthorized in 1996, it addressed small system drinking water concerns and required the U.S. Environmental Protection Agency (EPA) to assess treatment technologies relevant to small systems serving fewer than 10,000 people. With this requirement, the SDWA also identified two classes of technologies:

- compliance technologies—which refer to affordable technologies or other treatment techniques (TT) that comply with the maximum contaminant level (MCL) and to technologies that satisfy a TT requirement. Options include package plants or modular systems, and point-of-entry (POE) or point-of-use (POU) treatment; and
- variance technologies—which refer to technologies that must reduce contaminants to levels that protect public health. These technologies may not achieve compliance with the MCL or TT requirement, but must achieve the maximum reduction or inactivation efficiency affordable to a system, considering its size and the quality of the source water.

With small systems' needs in mind, the National Research Council (NRC) recently published the results of a study—*Safe Water From Every Tap: Improving Water Service to Small Communities*—which found that continuous technical and financial assistance is still needed to help more than 54,000 small systems comply with changing regulations. In addition, the NRC study discussed some water treatment technologies that small systems may use to provide safe drinking water to their customers. These treatment technologies are also explained separately through Tech Briefs, four-page water treatment fact sheets, offered by the National Drinking Water Clearinghouse (NDWC). These fact sheets are available online at *www.ndwc.wvu.edu* or by calling (800) 624-8301.

DISINFECTION

The Surface Water Treatment Rule (SWTR) requires public water systems to disinfect water obtained from surface water supplies or groundwater sources under the influence of surface water. Primary methods of disinfection are chlorine gas, chloramines, ozone, ultraviolet light, chlorine dioxide, and hypochlorite.

Chlorine (Gas)

Chlorine gas removes almost all microbial pathogens and is appropriate as both a primary and secondary disinfectant. Chlorine is a dangerous gas that is lethal at concentrations as low as 0.1 percent air by volume. Adequate mixing and contact time must be provided after injection to ensure complete disinfection of pathogens.

Hypochlorites

Sodium Hypochlorite. is available as a solution in concentrations of five to 15 percent chlorine, but is more expensive than chlorine gas. Sodium hypochlorite is easier to handle than gaseous chlorine or calcium hypochlorite, but it is very corrosive and must be kept away from equipment that can be damaged by corrosion.

Calcium Hypochlorite. is a solid white substance, which is 65 percent available chlorine and dissolves easily in water. It is a corrosive material with a strong odor that must be kept away from organic materials, such as wood, cloth, and petroleum products because of the dangers of fire or explosion. Calcium hypochlorite readily absorbs moisture, forming chlorine gas so shipping containers must be emptied completely or carefully resealed.

Chloramines

Chloramines are formed when water containing ammonia is chlorinated or when ammonia is added to water containing chlorine. An effective bactericide that produces fewer disinfection byproducts, chloramine is generated onsite. It is a weak disinfectant and is much less effective against viruses or protozoa than free chlorine. Chloramine is appropriate for use as a secondary disinfectant to prevent bacterial regrowth in a distribution system. Nitrogen trichloride appears to be the only detrimental reaction. Adequate contact and mixing time must be provided.

Ozonation

Ozone is a powerful oxidizing and disinfecting agent formed by passing dry air through a system of high voltage electrodes. Requiring shorter contact time and a smaller dosage than chlorine, ozone is widely used as a primary disinfectant. Ozone does not directly produce halogenated organic materials unless a bromide ion is present. A secondary disinfectant, usually chlorine, is required because ozone does not maintain an adequate residual in water. The capital costs of ozonation systems may be high and operation and maintenance are relatively complex.

Ultraviolet Light

Ultraviolet (UV) radiation, which is generated by a special lamp, penetrates the cell wall of an organism, rendering it unable to reproduce. UV radiation effectively destroys bacteria and viruses. As with ozone, a secondary disinfectant must be used to prevent regrowth of microorganisms. UV radiation:

- is readily available,
- produces no known toxic residuals,
- requires short contact times, and
- is easy to operate and maintain.

Conventional UV radiation may not inactivate *Giardia lamblia* or *Cryptosporidium* cysts in a cost-effective way, and should be used only by groundwater systems not directly influenced by surface water and where there is virtually no risk of protozoan cyst contamination. UV radiation is unsuitable for water with high levels of suspended solids, turbidity, color, or soluble organic matter. However, microorganisms can be killed without generating byproducts of chemical oxidation or halogenation.

Chlorine Dioxide

Chlorine dioxide, although a powerful oxidant, may be more difficult to handle than other forms of chlorine. Chlorine dioxide requires trained staff to manage its use and is so reactive that it may not provide a residual disinfectant in the distribution system. Photochemical decomposition of chlorine dioxide in reservoirs may increase chlorate concentrations, and other factors, including the generation process used and water pH, can affect chlorate and chlorite levels.

FILTRATION

Federal and state laws require all surface water systems and systems under the influence of surface water to filter their water. Filtration methods include slow and rapid sand filtration, diatomaceous earth filtration, direct filtration, membrane filtration, and cartridge filtration.

Slow Sand Filtration

The filter consists of a bed of fine sand approximately three to four feet deep supported by a one-foot layer of gravel and an underdrain system. It is a low-cost, simple to operate, reliable technology, and it is able to achieve greater than 99.9 percent *Giardia* cyst removal. Slow sand filtration is not suitable for water with high turbidity. The filter surface requires maintenance. Extensive land is required due to low-flow operation. Biological processes and chemical/physical processes common to various types of filters occur on the surface of the filter bed. Slow sand filters do not require coagulation/flocculation and may not require sedimentation.

Diatomaceous Earth Filtration

Diatomaceous earth (DE) filtration, also known as precoat or diatomite filtration, relies on a layer of diatomaceous earth approximately 1/8-inch thick placed on a septum or filter element. Septums may be placed in pressure vessels or operated under a vacuum in open vessels. The filters are simple to operate and effective in removing cysts, algae, and asbestos. They have been chosen for projects with limited initial capital, and for emergency or standby capacity to service large seasonal increases in demand. This filter is most suitable for water with low bacterial counts and low turbidity. Coagulant and filter aids are required for effective virus removal. Since chemical coagulation is not required, small water systems have used DE filtration for many years.

Direct Filtration

Direct filtration systems are similar to conventional systems, but omit sedimentation. Effective direct filtration performance ranges from 90 to 99 percent for virus removal and from 10 to 99.99 percent for *Giardia* removal. Coagulation must be included for *Giardia* removal. Direct filtration is often used with steel pressure vessels to maintain the pressure in a water line to avoid repumping after filtration. Direct filtration is only applicable for systems with high quality and seasonally consistent influent supplies. Direct filtration requires advanced operator skill and has frequent monitoring requirements.

Membrane Filtration

More stringent water quality regulations and inadequate water resources are making membrane technology increasingly popular as an alternative treatment technology for drinking water. Capital, operation, and maintenance costs continue to decline, making membrane processes more viable.

Nanofiltration (NF). This membrane process employs pressures between 75 to 150 pounds per square inch (psi) for operation. While it provides removal of ions contributing to hardness (i.e., calcium and magnesium), the technology is also very effective for removing color and disinfection byproducts precursors.

Ultrafiltration (UF). Operational pressures range from 10 to 100 psi, depending upon the application. UF may be employed for removal of some organic materials from freshwater, and may be used for liquid/solid separation.

Microfiltration (MF). A major difference between MF and UF is membrane pore size. The primary applications for this membrane process are particulate and microbial removal.

Bag Filtration

Bag filtration systems are based on physical screening processes. If the pore size of the bag filter is small enough, parasite removal will occur. Unless the quality of the raw water precludes the need for pretreatment, EPA recommends pretreatment of the raw water using sand or multimedia filters, followed by preliminary bag or cartridge filtration, and the use of micron filters as final filters to increase particulate removal efficiencies and to extend the life of the filter.

Cartridge Filtration

Cartridge filters are an emerging technology suitable for removing microbes and turbidity. These filters are easy to operate and maintain, making them suitable for treating low-turbidity influent. They can become fouled relatively quickly and must be replaced with new units. Although these filter systems are operationally simple, they are not automated and can require relatively large operating budgets. A disinfectant is recommended to prevent surface-fouling microbial growth on the cartridge filters and to reduce microbial pass-through.

Backwashable Depth Filtration

Backwashable depth filters operate in part like cartridge filters. This method filters uncoagulated water and is designed to be backwashed when terminal head loss is attained or turbidity breakthrough occurs.

CORROSION CONTROL

Corrosion in a system can be reduced by adjusting pH and alkalinity, softening the water, and changing the level of dissolved oxygen. Any corrosion adjustment program should include monitoring as water characteristics change over time.

pH Adjustment. Operators can promote the formation of a protective calcium carbonate coating (scale) in water lines by adjusting pH, alkalinity, and calcium levels.

Lime Softening. Lime softening affects lead's solubility by changing the water's pH and carbonate levels. Hydroxide ions are then present, and they decrease metal solubility by promoting the formation of solids that protect the surface of the pipe.

Dissolved Oxygen Levels. The presence of excessive dissolved oxygen increases water's corrosive activity. However, removing oxygen from water is not practical because of the expense. The following strategies may be used to minimize the presence of oxygen:

- exclude the aeration process in groundwater treatment,
- increase lime softening,
- extend the detention periods for treated water in reservoirs, or
- use the correct size water pumps in the treatment plant to minimize the introduction of air during pumping.

ION EXCHANGE AND DEMINERALIZATION

Ion exchange and membrane processes are becoming used extensively in water and wastewater treatment. Ion exchange is primarily used to remove of hardness ions, such as magnesium and calcium, and for water demineralization. Reverse osmosis and electrodialysis, both membrane processes, remove dissolved solids from water using membranes.

Ion Exchange (IO)

IO units can be used to remove any charged (ionic) substance from water, but are usually used to remove hardness and nitrate from groundwater. Ion exchange effectively removes more than 90 percent of barium, cadmium, chromium, silver, radium, nitrites, selenium, arsenic, and nitrate. Ion exchange is usually the best choice for removing radionuclides.

Reverse Osmosis (RO)

RO systems are compact, simple to operate, and require minimal labor, making them suitable for small systems where there is a high degree of seasonal fluctuation in water demand. RO can effectively remove nearly all inorganic contaminants from water. Properly operated units will attain 96 percent removal rates. RO can also effectively remove radium, natural organic substances, pesticides, and microbiological contaminants. RO is particularly effective when used in series. Water passing through multiple units can achieve near zero effluent contaminant concentrations.

Electrodialysis

Electrodialysis is very effective in removing fluoride and nitrate and can also remove barium, cadmium, and selenium.

Some of the advantages are:

- all contaminant ions and most dissolved non-ions are removed,
- it is relatively insensitive to flow and total dissolved solids (TDS) level, and
- it may have low effluent concentration.

Some of the limitations are:

- high capital and operating costs,
- high level of pretreatment required,
- reject stream is 20 to 90 percent of feed flow, and
- electrodes require replacement.

Activated Alumina

Activated Alumina (AA) is a physical and chemical process in which ions in the feed water are sorbed to an oxidized AA surface. AA is used in packed beds to remove contaminants such as fluoride, arsenic, selenium, silica, and natural organic matter.

ORGANIC REMOVAL

The technologies most suitable for organic contaminant removal in drinking water systems are granular activated carbon (GAC) and aeration. GAC has been designated by the EPA as the best available technology (BAT) for synthetic organic chemical removal.

Granular Activated Carbon

Several operational and maintenance factors affect the performance of GAC. Contaminants in the water can occupy GAC adsorption sites, whether they are targeted for removal or not. Also, adsorbed contaminants can be replaced by other contaminants with which GAC has a greater affinity. Therefore, the presence of other contaminants might interfere with the removal of the contaminants of concern.

After a period of months or years, depending on the concentration of contaminants, the surface of the pores in the GAC can no longer adsorb contaminants. The carbon must then be replaced.

Aeration

Aeration, also known as air stripping, mixes air with water to volatilize contaminants (turn them to vapor), which are either released directly to the atmosphere or treated and released. Aeration is used to remove volatile organic chemicals (VOC) and can also remove radon. A small system might be able to use a simple aerator constructed from relatively common materials instead of a specially designed aerator system. Aerators include:

- a system that cascades the water or passes it through a slotted container,
- a system that runs water over a corrugated surface, or
- an airlift pump that introduces oxygen as water is drawn from a well.

Other Aeration Types

Packed Column Aeration (PCA). PCA or packed tower aeration (PTA) is a waterfall aeration process that drops water over a medium within a tower to mix the water with air. The medium is designed to break the water into tiny droplets and to maximize its contact with air bubbles for removal of the contaminant. Air is also blown in from underneath the medium to enhance this process. Packed columns usually operate automatically and need only daily visits to ensure that the equipment is running satisfactorily. Maintenance requirements include servicing pump and blower motors and replacing air filters on the blower.

Diffused Aeration. In a diffused aeration system, a diffuser bubbles air through a contact chamber for aeration. The diffuser is usually located near the bottom of the chamber where pressurized air is introduced. The main advantage of diffused aeration systems is that they can be created from existing structures, such as storage tanks. However, these systems are less effective than PCA and usually are employed only in systems with adaptable existing structures.

Multiple Tray Aeration. Multiple tray aeration directs water through a series of trays made of slats, perforations, or wire mesh. A blower introduces air from underneath the trays. Multiple tray aeration units have less surface area than PCA units and can experience clogging from iron and manganese, biological growth, and corrosion problems. Multiple tray aeration units are readily available from package plant manufacturers.

Shallow Tray Aeration (STA). STAs involve the use of shallow trays and are more efficient than multiple tray aerators. STAs increase the available area of mass transfer; thereby increasing the removal efficiency of most VOCs. However, because of the high air-to-water ratio, greater energy costs may be incurred.

Spray Aeration. Spray aeration is an accepted technology in which the contaminated water is sprayed through nozzles. The small droplets produced expose a large interfacial surface area through which VOCs can migrate from a liquid (water) phase to the gaseous (air) phase. Spray aerators have been used to effectively treat VOCs, but are not energy efficient and need a large operational area.

Mechanical Aeration. Mechanical aeration uses mechanical stirring mechanisms to mix air with the water. These systems can effectively remove VOCs. Mechanical aeration units need large amounts of space because they demand long detention times for effective treatment. As a result, they often require open-air designs, which can freeze in cold climates. However, mechanical aeration systems are easy to operate and are less susceptible to clogging from biological growth than PCA systems.

LIME SOFTENING

Lime softening is best suited to groundwater sources, which have relatively stable water quality. The combination of variable source water quality and the complexity of the chemistry of lime softening may make it too complicated for small systems that use surface water sources. Lime softening is unlikely to be suitable for treating

Surface Water Treatment Rule Compliance Technologies for Disinfection

Unit Technology	Limitations (see Footnotes)	Operator Skill Level Required	Raw Water Quality Range and Considerations[1]	Removals: Log *Giardia* & Log Virus w/CT's Indicated in ()[1]
Free Chlorine	(a, b)	Basic	Better with high quality. High iron or manganese may require sequestration or physical removal.	3 log(104) & 4 log(6).
Ozone	(c, d)	Intermediate	Better with high quality. High iron or manganese may require sequestration or physical removal.	3 log(1.43) & 4 log(1.0).
Chloramines	(e)	Intermediate	Better with high quality. Ammonia dose should be tempered by natural ammonia levels in water.	3 log(1850) & 4 log(1491).
Chlorine Dioxide	(f)	Intermediate	Better with high quality.	3 log(23) & 4 log(25).
Onsite Oxidant Generation	(g)	Basic	Better with high quality.	Research pending on CT values. Use free chlorine.
Ultraviolet (UV) Radiation	(h)	Basic	Relatively clean source water required. Iron, natural organic matter and turbidity affect UV dose.	1 log *Giardia* (80–120) & 4 log viruses (90–140) mWsec/cm2 doses in parentheses[2].

[1] CT (Concentration × Time), in mg-min/L, based upon 1989 Surface Water Treatment Rule Guidance Manual. Temp. 10 C, mid-pH range, unless otherwise indicated.

[2] UV dose is product of mW/cm2 (intensity) × sec (time); bases of viral inactivation ranges are rotavirus and MS–2 tests.

Limitations Footnotes

a. Providing adequate CT (contact time) may be a problem for some supplies.
b. Chlorine gas requires special caution in handling and storage, and operator training.
c. Ozone leaks represent hazard: air monitoring required.
d. Ozone used as primary disinfectant (i.e., no residual protection).
e. Long CT. Requires care in monitoring of ratio of added chlorine to ammonia.
f. Chlorine dioxide requires special storage and handling precautions.
g. Oxidants other than chlorine not detected in solution by significant research effort. CT should be based on free chlorine until new research determines appropriate CT values for electrolyzed salt brine.
h. No disinfectant residual protection for distributed water.

groundwater in systems serving 500 or fewer people unless those systems have access to a trained operator who can monitor the treatment process. Either hydrated lime or quicklime may be used in the softening process. The choice depends upon economic factors, such as the relative cost per ton of the two materials as well as the size and equipment of the softening plant.

What Are Other Softening Alternatives?

The selection of lime, lime-soda ash, or caustic soda softening is based on cost, TDS criteria, sludge production, carbonate and noncarbonate hardness, and chemical stability. Water containing little or no noncarbonate hardness can be softened with lime alone. Caustic soda softening increases the TDS of treated water, while lime and lime-soda ash softening often decrease TDS. Caustic soda softening produces less sludge than lime and lime-soda ash softening. Caustic soda does not deteriorate during storage, while hydrated lime may absorb carbon dioxide and water during storage, and quicklime may slake in storage causing feeding problems. The final selection is generally based on cost, water quality, and owner and operator preference.

FOR MORE INFORMATION

Small drinking water systems are more likely to violate SDWA regulations because when MCLs were set, they were based upon systems serving larger metropolitan areas. Thus, small systems must explore innovative technologies that they can afford. The NDWC's RESULTS (Registry of Equipment Suppliers of Treatment Technologies for Small Systems) database houses information related to small drinking water systems. The clearinghouse gathered this information from system operators, drinking water state offices, vendors, and others.

Database searches are available from the NDWC through combinations of site location, vendor name, type of technology, type of contaminant, and system size—and they include contact names and telephone numbers. Consulting engineers, local officials, private owners, and regulators may use RESULTS not only to understand technologies that small systems use, but

Surface Water Treatment Rule Compliance Technology for Filtration

Unit Technology	Limitations (see Footnotes)	Operator Skill Level Required	Raw Water Quality Range and Considerations[1]	Removals: Log *Giardia* & Log Virus
Conventional Filtration (includes dual-stage and dissolved air flotation)	(a)	Advanced	Wide range of water quality. Dissolved air flotation is more applicable for removing particulate matter that doesn't readily settle: algae, high color, low turbidity—up to 30–50 nephelometric turbidity units (NTU) and low-density turbidity.	2–3 log *Giardia* & 1 log viruses.
Direct Filtration (includes in-line filtration)	(a)	Advanced	High quality. Suggested limits: average turbidity 10 NTU; maximum turbidity 20 NTU; 40 color units; algae on a case-by-case basis.[1]	0.5 log *Giardia* & 1–2 log viruses (1.5–2 log *Giardia* w/coagulation).
Slow Sand Filtration	(b)	Basic	Very high quality or pretreatment. Pretreatment required if raw water is high in turbidity, color, and/or algae.	4 log *Giardia* & 1–6 log viruses.
Diatomaceous Earth Filtration	(c)	Intermediate	Very high quality or pretreatment. Pretreatment required if raw water is high in turbidity, color, and/or algae.	Very effective for *Giardia*; low bacteria and virus removal.
Reverse Osmosis	(d, e, f)	Advanced	Requires prefiltrations for surface water—may include removal of turbidity, iron, and/or manganese. Hardness and dissolved solids may also affect performance.	Very effective (cyst and viruses).
Nanofiltration	(e)	Intermediate	Very high quality of pretreatment. See reverse osmosis pretreatment.	Very effective (cyst and viruses).
Ultrafiltration	(g)	Basic	High quality or pretreatment.	Very effective *Giardia*, > 5–6.
Microfiltration	(g)	Basic	High quality or pretreatment required.	Very effective *Giardia*, > 5–6 log; Partial removal viruses.
Bag Filtration	(g, h, i)	Basic	Very high quality or pretreatment required, due to low particulate loading capacity. Pretreatment if high turbidity or algae.	Variable *Giardia* removals & disinfection required for virus credit.
Cartridge Filtration	(g, h, i)	Basic	Very high quality or pretreatment required, due to low particulate loading capacity. Pretreatment if high turbidity or algae.	Variable *Giardia* removals & disinfection required for virus credit.
Backwashable Depth Filtration	(g, h, i)	Basic	Very high quality or pretreatment required, due to low particulate loading capacity. Pretreatment if high turbidity or algae.	Variable *Giardia* removals & disinfection required for virus credit.

[1] National Research Council (NRC), Committee on Small Water Supply Systems. "Safe Water From Every Tap: Improving Water Service to Small Communities." National Academy Press, Washington, D.C. 1997.

[2] Adham, S.S., Jacangelo, J.G., and Laine, J.M. "Characteristics and Costs of MF and UF Plants." *Journal American Water Works Association*, May 1996.

Limitations Footnotes

a. Involves coagulation. Coagulation chemistry requires advanced operator skill and extensive monitoring. A system needs to have direct full-time access or full-time remote access to a skilled operator to use this technology properly.

b. Water service interruptions can occur during the periodic filter-to-waste cycle, which can last from six hours to two weeks.

c. Filter cake should be discarded if filtration is interrupted. For this reason, intermittent use is not practical. Recycling the filtered water can remove this potential problem.

d. Blending (combining treated water with untreated raw water) cannot be practiced at risk of increasing microbial concentration in finished water.

e. Post-disinfection recommended as a safety measure and for residual maintenance.

f. Post-treatment corrosion control will be needed prior to distribution.

g. Disinfection required for viral inactivation.

h. Site-specific pilot testing prior to installation likely to be needed to ensure adequate performance.

i. Technologies may be more applicable to system serving fewer than 3,300 people.

Compliance Technology for the Total Coliform Rule

40 CFR 141.63(d)–Best Technologies or Other Means to Comply (Complexity Level Indicated)	Comments/Water Quality Concerns
Protecting wells from contamination, i.e., placement and construction of well(s) (Basic).	Ten State Standards and other standards (AWWA A100–90) apply; interfacing with other programs essential (e.g., source water protection program).
Maintenance of a disinfection residual for distribution system protection (Intermediate).	Source water constituents may affect disinfection: iron, manganese, organics, ammonia, and other factors may affect dosage and water quality. Total Coliform Rule (TCR) remains unspecific on type/amount of disinfectant, as each type differs in concentration, time, temperature, pH, interaction with other constituents, etc.
Proper maintenance of distribution system: pipe repair/replacement, main flushing programs, storage/reservoir and operation and maintenance (O&M) programs (including cross-connection control/backflow prevention), and maintenance of positive pressure throughout (Intermediate).	O&M programs particularly important for smaller systems needing to maintain water purity. States may vary on distribution protection measures. See also EPA's Cross-Connection Control Manual (#EPA 570/9-89-077).
Filtration and/or disinfection of surface water or other groundwater under direct influence; or disinfection of groundwater (Basic thru Advanced).	Same issues as cited above under maintaining disinfection residual; pretreatment requirements affect complexity of operation. Refer to Surface Water Treatment Rule Compliance Technology List; and other regulations under development.
Groundwaters: Compliance with State Wellhead Protection Program (Intermediate).	EPA/State Wellhead Protection Program implementation (per §1428 SDWA): may be used to assess vulnerability to contamination, and in determination of sampling and sanitary survey frequencies.

Technologies for Inorganic Contaminants

Unit Technology	Limitations (see Footnotes)	Operator Skill Level Required	Raw Water Quality Range
1. Activated Alumina	(a)	Advanced	Groundwaters, competing anion concentrations will affect run length.
2. Ion Exchange (IO)		Intermediate	Groundwaters with low total dissolved solids, competing ion concentrations will affect run length.
3. Lime Softening	(b)	Advanced	Hard ground and surface waters.
4. Coagulation/Filtration	(c)	Advanced	Can treat wide range of water quality.
5. Reverse Osmosis (RO)	(d)	Advanced	Surface water usually require prefiltration.
6. Alkaline Chlorination	(e)	Basic	All groundwaters.
7. Ozone Oxidation		Intermediate	All groundwaters.
8. Direct Filtration		Advanced	Needs high raw water quality.
9. Diatomaceous earth filtration		Intermediate	Needs high raw water quality.
10. Granular Activated Carbon		Basic	Surface waters may require prefiltration.
11. Electrodialysis Reversal		Advanced	Requires prefiltration for surface water.
12. Point of Use (POU)-IO	(f)	Basic	Same as Technology #2.
13. POU-RO	(f)	Basic	Same as Technology #5.
14. Calcium Carbonate Precipitation	(g)	Basic	Waters with high levels of alkalinity and calcium.
15. pH and alkalinity adjustment (chemical feed)	(g)	Basic	All ranges.
16. pH and alkalinity adjustment (limestone contactor)	(h)	Basic	Waters that are low in iron and turbidity. Raw water should be soft and slightly acidic.
17. Inhibitors		Basic	All ranges.
18. Aeration	(i)	Basic	Waters with moderate to high carbon dioxide content.

Limitations Footnotes

a. Chemicals required during regeneration and pH adjustments may be difficult for small systems to handle.
b. Softening chemistry may be too complex for small systems.
c. It may not be advisable to install coagulation/filtration solely for inorganics removal.
d. If all of the influent water is treated, post-treatment corrosion control will be necessary.
e. pH must exceed pH 8.5 to ensure complete oxidation without build-up of cyanogen chloride.
f. When POU devices are used for compliance, programs for long-term operation, maintenance, and monitoring must be provided by water utility to ensure proper performance.
g. Some chemical feeds require high degree of operator attention to avoid plugging.
h. This technology is recommended primarily for the smallest size category.
i. Any of the first five aeration technologies listed for volatile organic contaminants can be used.

Technologies for Volatile Organic Contaminants

Unit Technology	Limitations (see Footnotes)	Operator Skill Level Required	Raw Water Quality Range[1]
1. Packed Tower Aeration (PTA)	(a)	Intermediate	All groundwaters.
2. Diffused Aeration	(a, b)	Basic	All groundwaters.
3. Multi-Stage Bubble Aerators	(a, c)	Basic	All groundwaters.
4. Tray Aeration	(a, d)	Basic	All groundwaters.
5. Shallow Tray Aeration	(a, e)	Basic	All groundwaters.
6. Spray Aeration	(a, f)	Basic	All groundwaters.
7. Mechanical Aeration	(a, g)	Basic	All groundwaters.
8. Granular Activated Carbon (GAC)	(h)	Basic	All groundwaters.

[1] National Research Council (NRC). "Safe Water from Every Tap: Improving Water Service to Small Communities." National Academy Press. Washington, DC. 1997.

Limitations Footnotes

a. Pretreatment for the removal of microorganisms, iron, manganese, and excessive particulate matter may be needed. Post-treatment disinfection may have to be used.

b. May not be as efficient as other aeration methods because it does not provide for convective movement of the water thus limiting air-water contact. It is generally used only to adapt existing plant equipment.

c. These units are highly efficient; however, the efficiency depends upon the air-to-water ratio.

d. Costs may increase if a forced draft is used. Slime and algae growth can be a problem but can be controlled with chemicals such as copper sulfate or chlorine.

e. These units require high air-to-water ratios (100–900 m3/m3).

f. For use only when low removal levels are needed to reach a maximum contaminant level (MCL) because these systems may not be as energy efficient as other aeration methods because of the contacting system.

g. For use only when low removal levels are needed to reach an MCL because these systems may not be as energy efficient as other aeration methods because of the contacting system. The units often require large basins, long residence times, and high energy inputs, which may increase costs.

h. See the Synthetic Organic Compounds (SOC) compliance technology table for limitation regarding these technologies.

Technologies for Synthetic Organic Compounds

Unit Technology	Limitations (see Footnotes)	Operator Skill Level Required[1]	Raw Water Quality Range and Considerations[1]
1. Granular Actived Carbon (GAC)		Basic	Surface water may require prefiltration.
2. Point of Use GAC	(a)	Basic	Surface water may require prefiltration.
3. Powdered Activated Carbon	(b)	Intermediate	All waters.
4. Chlorination	(c)	Basic	Better with high quality waters.
5. Ozonation	(c)	Basic	Better with high quality waters.
6. Packed Tower Aeration (PTA)	(d)	Intermediate	All groundwaters.
7. Diffused Aeration	(d, e)	Basic	All groundwaters.
8. Multi-Stage Bubble Aerators	(d, f)	Basic	All groundwaters.
9. Tray Aeration	(d, g)	Basic	All groundwaters.
10. Shallow Tray Aeration	(d, f)	Basic	All groundwaters.

[1] National Research Council (NRC). "Safe Water from Every Tap: Improving Water Service to Small Communities." National Academy Press. Washington, DC. 1997.

Limitations Footnotes

a. When POU devices are used for compliance, programs for long-term operation, maintenance, and monitoring must be provided by water utility to ensure proper performance.

b. Most applicable to small systems that already have a process train including basins mixing, precipitation or sedimentation, and filtration. Site specific design should be based on studies conducted on the system's particular water.

c. See the Surface Water Treatment Rule compliance technology tables for limitations associated with this technology.

d. Pretreatment for the removal of microorganisms, iron, manganese, and excessive particulate matter may be needed. Post-treatment disinfection may have to be used.

e. May not be as efficient as other aeration methods because it does not provide for convective movement of the water thus limiting air-water contact. It is generally used only to adapt existing plant equipment.

f. These units are highly efficient; however, the efficiency depends upon the air-to-water ratio.

g. Forces may increase if a forced draft is used.

technologies that are affordable, appropriate, and reliable. Information in RESULTS may be obtained three ways: access the database through the NDWC's Web site located at *www.ndwc.wvu.edu*; call the NDWC at (800) 624–8301 or (304) 293–4191 and ask a technical assistant to perform a search for you.

READING LIST

NDWC Report. 1998. *RESULTS Database: Small Water Systems Technologies*. Morgantown, WV, September.*

**NDWC* RESULTS Database: Small Water Systems Technologies *report and Tech Briefs are available online at www.ndwc.wvu.edu or by calling (800) 624-8301 or (304) 293-4191.*

Technologies for Radionuclides

Unit Technology	Limitations (see Footnotes)	Operator Skill Level Required[1]	Raw Water Quality Range and Considerations[1]
1. Ion Exchange (IO)	(a)	Intermediate	All groundwaters.
2. Point of Use (POU) IO	(b)	Basic	All groundwaters.
3. Reverse Osmosis (RO)	(c)	Advanced	Surface waters, usually require prefiltration.
4. POU RO	(b)	Basic	Surface waters, usually require prefiltration.
5. Lime Softening	(d)	Advanced	All waters.
6. Green Sand Filtration	(e)	Basic	
7. Co-precipitation with Barium Sulfate	(f)	Intermediate to Advanced	Groundwaters with suitable water quality.
8. Electrodialysis/Electrodialysis Reversal		Basic to Intermediate	All groundwaters
9. Pre-formed Hydrous Manganese Oxide Filtration	(g)	Intermediate	All groundwaters

[1] National Research Council (NRC). "Safe Water from Every Tap: Improving Water Service to Small Communities." National Academy Press. Washington, DC. 1997.

Limitations Footnotes

a. The regeneration solution contains high concentrations of the contaminant ions. Disposal options should be carefully considered before choosing this technology.

b. When POU devices are used for compliance, programs for long-term operation, maintenance, and monitoring must be provided by water utility to ensure proper performance.

c. Reject water disposal options should be carefully considered before choosing this technology. See other RO limitations described in the Surface Water Treatment Rule Compliance Technologies Table.

d. The combination of variable source water quality and the complexity of the chemistry involved in lime softening may make this technology too complex for small surface water systems.

e. Removal efficiencies can vary depending on water quality.

f. This technology may be very limited in application to small systems. Since the process requires static mixing, detention basins, and filtration; it is most applicable to systems with sufficiently high sulfate levels that already have a suitable filtration treatment train in place.

g. This technology is most applicable to small systems that already have filtration in place.

National Research Council. 1997. *Safe Water From Every Tap: Improving Water Service to Small Communities*. National Academy Press, Washington, DC.

NDWC Tech Brief: Disinfection, item #DWBLPE47.*

NDWC Tech Brief: Filtration, item #DWBLPE50.*

NDWC Tech Brief: Corrosion Control, item #DWBLPE52.*

NDWC Tech Brief: Ion Exchange and Demineralization, item #DWBLPE56.*

NDWC Tech Brief: Organics Removal, item #DWBLPE59.*

NDWC Tech Brief: Package Plants, item #DWBLPE63.*

NDWC Tech Brief: Lime Softening, item #DWBLPE67.*

NDWC Tech Brief: Iron and Manganese Removal, item #DWBLPE70.*

NDWC Tech Brief: Membrane Filtration, item #DWBLPE81.*

U.S. General Accounting Office. 1994. *Drinking Water: Stronger efforts Essential for Small Communities to Comply with Standards*. GAO/RCED-94-40. Washington, DC.

To ease many of the demands placed on small systems, the 1996 Safe Drinking Water Act amendments require the U.S. Environmental Protection Agency (EPA) to evaluate affordable technologies and address existing and future regulations, which establish a maximum contaminant level or treatment technique.

The following tables are taken from three EPA guidance documents: EPA-815-R-98-001, *Small System Compliance Technology List for the Surface Water Treatment Rule and Total Coliform Rule*; EPA-815-R-98-002, *Small System Compliance Technology List for the Non-Microbial Contaminants Regulated Before 1996*; and EPA-815-R-98-003, *Variance Technology Findings for Contaminants Regulated Before 1996*.

For information about the availability of these guidance and support documents, please contact the Safe Drinking Water Hotline: phone (800) 426–4791, fax (703) 285–1101, or e-mail *hotline-sdwa@epamail.epa.gov*.

The National Drinking Water Clearinghouse assists small communities by collecting, developing, and providing timely information relevant to drinking water issues.

(800) 624-8301/(304) 293–4191
www.ndwc.wvu.edu
National Drinking Water Clearinghouse
West Virginia University
P.O. Box 6064
Morgantown, WV 26506-6064

HAVE YOU READ ALL OF OUR FACT SHEETS?

Tech Briefs, drinking water treatment fact sheets, have been a regular feature in the National Drinking Water Clearinghouse (NDWC) magazine *On Tap* for several years. These very popular items are researched, complied, and written by NDWC staff.

For a complete listing of our Tech Briefs, visit our Web site at *www.ndwc.wvu.edu*. Additional copies of fact sheets are free; however, postal charges may be added. To order, call the NDWC at (800) 624–8301 or (304) 293–4191. You also may order online at *ndwc_orders@mail.nesc.wvu.edu*.

For further information, comments about this fact sheet, or to suggest topics, call the NDWC at one of the above numbers or contact Vipin Bhardwaj at *vbhardw2@wvu.edu*, Geri Ramiser at *gramiser@wvu.edu*, Kathy Jesperson at *kjespers@wvu.edu*, or Mark Kemp-Rye at *mkemp@wvu.edu*.

ULTRAVIOLET DISINFECTION

National Drinking Water
Clearinghouse

Using ultraviolet (UV) light for drinking water disinfection dates back to 1916 in the U.S. Over the years, UV costs have declined as researchers develop and use new UV methods to disinfect water and wastewater. Currently, several states have developed regulations that allow systems to disinfect their drinking water supplies with UV light. Running a UV light system requires a basic level of operator skill and relatively clean source water. On the down side, however, UV offers no residual disinfection within the distribution system.

WHAT IS UV DISINFECTION?

UV light, which continues to be a reliable means of disinfection, involves exposing contaminated water to radiation from UV light. The treatment works because UV light penetrates an organism's cell walls and disrupts the cell's genetic material, making reproduction impossible.

A special lamp generates the radiation that creates UV light by striking an electric arc through low-pressure mercury vapor. This lamp emits a broad spectrum of radiation with intense peaks at UV wavelengths of 253.7 nanometers (nm) and a lesser peak at 184.9 nm. Research has shown that the optimum UV wavelength range to destroy bacteria is between 250 nm and 270 nm. At shorter wavelengths (e.g., 185 nm), UV light is powerful enough to produce ozone, hydroxyl, and other free radicals that destroy bacteria.

The U.S. Department of Health, Education, and Welfare set guidelines for UV light disinfection in 1966. These guidelines require a minimum dose of 16 mWs/cm^2 [milliwatt seconds per square centimeter] at all points throughout the water disinfection unit.

However, the American National Standards Institute and the National Sanitation Foundation International set the minimum UV light requirement at 38 mWs/cm2 for class A point of use (POU) and point of entry (POE) devices that treat visually clear water.

The U.S. Environmental Protection Agency (EPA) lists UV disinfection as an approved technology for small public water systems. In addition, EPA is considering the following variations of conventional UV treatment as "emerging" technologies: pulsed UV, medium-pressure UV, and UV oxidation (i.e., used in combination with peroxide or ozone).

ADVANTAGES

Generally, UV is simple to install and requires little supervision, maintenance, or space. Improved safety, minimum service time, low operation and maintenance costs, and the absence of a chemical smell or taste in finished water are primary factors for selecting UV technology rather than traditional disinfection technologies.

UV treatment breaks down or removes some organic contaminants. UV achieves 1-log reduction of *Giardia lamblia* at an intensity of 80–120 mWs/cm^2, and 4-log reduction of viruses at an intensity of 90–140 mWs/cm^2. Only recently has the scientific community begun to accept UV as a highly effective tool for *Cryptosporidium* control.

UV light disinfection does not form any significant disinfection byproducts, nor does it cause any significant increase in assimilable organic carbon (AOC).

Research has confirmed that UV effectiveness is relatively insensitive to temperature and pH differences. In addition, researchers found that UV application does not convert nitrates to nitrites, or bromide to bromines or bromates.

Recent pilot studies show that UV-treated drinking water inhibits bacterial growth and replication in the distribution system; however, conditions within distribution systems, such as leaks, still require additional residual disinfection (e.g., free chlorine).

The advantages of using UV, rather than chemical disinfection, include:

- Has no known toxic or significant nontoxic byproducts;
- Has no danger of overdosing;
- Removes some organic contaminants;
- Has no volatile organic compound (VOC) emissions or toxic air emissions;
- Has no onsite smell and no smell in the final water product;
- Requires very little contact time (seconds versus minutes for chemical disinfection);
- Does not require storage of hazardous material;
- Requires minimal space for equipment and contact chamber;
- Improves the taste of water because of some organic contaminants and nuisance microorganisms are destroyed;
- Does not affect minerals in water; and
- Has little or no impact on the environment except for disposing of used lamps or obsolete equipment.

LIMITATIONS

Microbial and chemical characteristics are two major water quality factors that affect the UV unit performance. Microbial characteristics of water include type, source, age, and density. Chemical water characteristics include nitrites, sulfites, iron, hardness, and aromatic organic levels.

UV radiation is not suitable for water with high levels of suspended solids, turbidity, color, or soluble organic matter. These materials can react with UV radiation, and reduce disinfection performance. Turbidity makes it difficult for radiation to penetrate water.

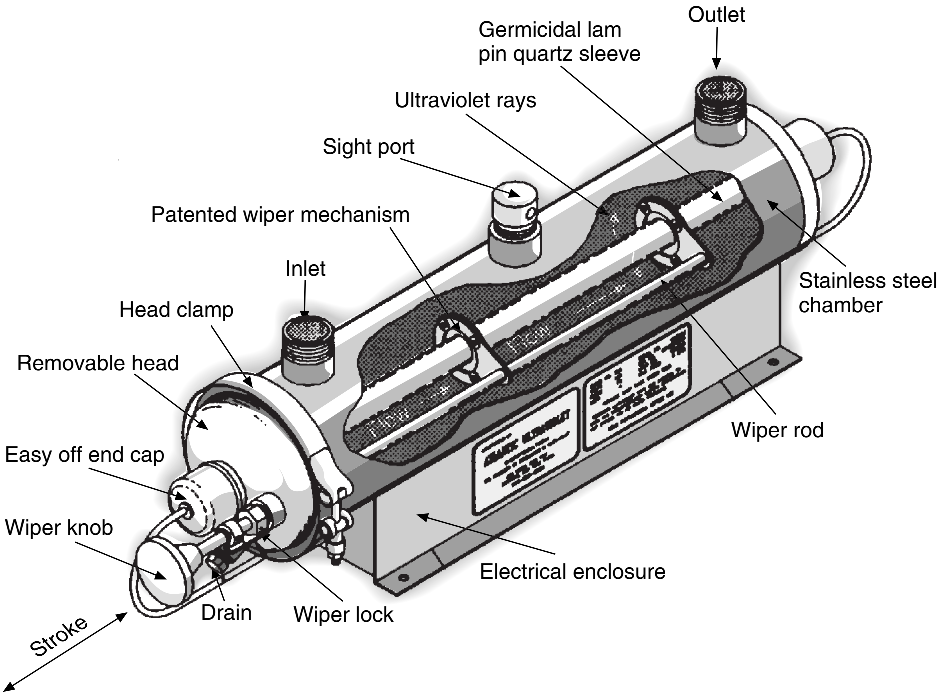

Closed vessel ultraviolet reactor. *Reprinted with permission from the Atlantic Ultraviolet Corporation.*

Disadvantages of UV disinfection include:

- No disinfection residual;
- No technical database exists on how well UV systems perform for various water quality conditions; and
- No standardized mechanism measures, calibrates, or certifies how well equipment works before or after installation.

Systems also should consider using different kinds of microbial testing. Laboratories typically test for total coliform to judge microbiological activity in drinking water—but coliforms are sensitive to UV light. Because of this sensitivity, microbial tests for UV treated finished water should include a Heterotrophic Plate Count (HPC) test. HPC microorganisms may provide a better disinfection assessment than the UV sensitive coliforms.

PROCESS DESCRIPTION

UV light effectively destroys bacteria and viruses. However, how well the UV system works depends on the energy dose that the organism absorbs. If the energy dose is not high enough, the organism's genetic material may only be damaged rather than disrupted.

An effective dose is measured as a product of the lamp's intensity (the rate at which photons are delivered to the target), including radiation concentration, proper wavelength, exposure time, water quality, flow rate, and the microorganism's type, and source, as well as its distance from the light source.

At a minimum, drinking water systems should install two UV units, which are both capable of carrying the amount of water the system was designed to handle. Having two units in place assures continuous disinfection when one unit is being serviced. Two units also can ensure operation during low-flow demand periods.

Modular units designed for small drinking water systems are easy to install and operate (two plumbing connections per unit and one electrical hook-up). They should be equipped with automatic cleaners and remote alarm systems. For systems in isolated areas, operators should maintain and store a set of spare parts onsite, and consider a telemetry system for monitoring treatment.

Typical UV light components include:

- A stable high-voltage source of electricity because low-line voltage would result in a lower UV dose;
- A chamber made of stainless steel or any other material that is opaque and will not corrode;
- UV lamps that are properly secured inside quartz sleeves, easing installation, replacement, and maintenance;
- Quartz sleeves with sufficiently high transmission rates to deliver the UV energy produced by UV lamps;
- Mechanical wipers to maintain optimum transmission between scheduled cleaning and maintenance work;
- Sensors to monitor the UV intensity passing through the water. These sensors need to be connected to alarm systems to alert the operator in case of low UV intensity. The operator must have easy access to these sensors for necessary installation, replacement, calibration, and maintenance;
- Safety control to shut off UV lamps in case of low-flow levels and elevated lamp temperature;
- Arc and lamp-out monitors to alert the operator of system failure; and
- Electronic ballasts.

UV units are currently used as stand-alone treatment systems or as part of a series of other drinking water treatment processes or multiple barrier system. A common treatment that uses UV light to remove and disinfect contaminants from groundwater sources involves a combined ozone or hydrogen-peroxide process along with UV application. So, it is common to find that manufacturers of UV equipment also manufacture ozone equipment.

Furthermore, the drinking water treatment industry provides UV equipment (mainly closed chamber units) for short-term uses. Rental units are used in cleanup and emergency situations, such as if groundwater is contaminated by spilled toxic organic compounds.

MONITORING AND OPERATION REQUIREMENTS

Factors that affect UV light system performance are: lamp output, lamp aging, and plating or fouling of unit surfaces. To better control these factors, operators must ensure continuous dose measurement (i.e., accurate intensity and flow-rate measurement) and proper maintenance (cleaning as well as lamp and sleeve replacement regimes).

Technological advances have eliminated many of the operation and maintenance problems that were associated with earlier UV applications. Current systems are equipped with mechanical cleaners, ultrasonic cleaners, or some self-cleaning mechanism (mandatory if water fouling agents, such as iron, are present in the water entering the unit); lamps that are easy to install and replace; and alarm systems that indicate minor and major failure.

To ensure continued system operation, a maintenance schedule needs to be in place. This schedule should include periodic site inspections; changing lamps annually or when light transmission efficiency has decreased to 70 percent; inspecting and cleaning surfaces; inspecting or cleaning the UV chamber interior every six months; and inspecting and replacing ballasts, O-rings, valves, and switches.

Furthermore, the operator should monitor water turbidity and color since they are natural barriers to UV light transmission. And some dissolved minerals, such as calcium, have a tremendous negative effect on UV absorbance.

Since it may not be practical to provide instantaneous stand-by power during power outages, the system should be designed to automatically stop water flow or provide an alternate means of disinfection as a backup. Where the system is dependent on electrically powered pumps, this measure may not be necessary because the pumps will shut off when the power goes out. However, gravity flow systems may be vulnerable.

UV disinfection should have the following minimum operational controls and procedures:

- A central display indicating alarms for power failure, lamp failure, hours of lamp operation, low UV dosage, high lamp temperature, high ballast temperature, and high system flows;
- Methods that monitor lamp temperature, ballast temperature, and system water flows;
- A minimum of two photodiode sensors per unit to monitor UV dosage at 254 nm. These sensors must be calibrated using approved standards each time the lamps are cleaned or replaced or the UV chamber is serviced;
- Automatic UV system by-pass or shutoffs, which are activated whenever the system exceeds peak design flow rates, when UV dosage is low, or when lamp or ballast temperatures are high; and
- Two UV units should be installed so flow is not interrupted when one unit is out of service.

WHERE CAN I FIND MORE INFORMATION?

Hargy, T.M., J.L. Clancy, and Z. Bukhari. 2000. “Shedding UV Light on the *Cryptosporidium* Threat.” In NSF Proceedings of the Small Drinking Water and Wastewater Systems. International Symposium and Technology Expo: Phoenix, Arizona.

Malley, J.P., G.A. Snicer, and A.M. Doucette. 1998. “Alternative Disinfection Strategies for Small Systems.” In *Small Systems Water Treatment Technologies: State of the Art Workshop*. NEWWA Joint Regional Operations Conference and Exhibition: Marlborough, Massachusetts.

Parrotta, M.J. and F. Bekdash. 1998. “UV Disinfection of Small Groundwater Supplies.” *Journal of the American Water Works* Association. AWWA: Denver.

U.S. Environmental Protection Agency. 1996. *Ultraviolet Light Disinfection Technology in Drinking Water Application—An Overview*. Office of Water: Washington, DC. EPA/811-R-96-002.

U.S. Environmental Protection Agency. 1998. *Small System Compliance Technology List for the Surface Water Treatment Rule and Total Coliform Rule*. Office of Water: Washington, DC. EPA/815/R/98/001.

HAVE YOU READ ALL OUR TECH BRIEFS?

“Tech Briefs” drinking water treatment fact sheets have been a regular feature in the National Drinking Water Clearinghouse (NDWC) newsletter *On Tap* for more than four years. NDWC Technical Assistance Specialist Mohamed Lahlou, Ph.D., researches, compiles, and writes these very popular items.

A package of the Tech Briefs is now available as a product. A three ring binder holds all the current Tech Briefs in print. New selections can be easily added to the package as they become available. To order this product, call the NDWC at the numbers listed below and ask for item #DWPKPE71. The item is free.

To order, call the NDWC at (800) 624-8301 or (304) 293-4191. You also may order online at ndwc_orders@mail.estd.wvu.edu or download fact sheets from our Web site at www.ndwc.wvu.edu.

(Additional copies of fact sheets are free; however, postal charges may be added.)

For further information, to comment on this fact sheet, or to suggest topics, call Dr. Lahlou at one of the above numbers or contact him via e-mail at mlahlou2@wvu.edu.

ULTRAVIOLET IRRADIATION

M. SIDDIQUI
University of Utah
Salt Lake City, Utah

Ultraviolet irradiation (UV) is a potential alternative to chlorination for drinking water disinfection. UV provides disinfection without producing the problematic disinfection by-products of chemical disinfectants such as chlorine. During retrofitting of UV to an existing potable water treatment process or during design of a new installation, there can be significant benefits in reexamining the overall disinfection strategy for the treatment plant and the water distribution system (1).

MECHANISM

Light whose wavelengths are longer than 700 nm is in the infrared and radio emission range. Light of wavelengths shorter than 400 nm is in the ultraviolet range. This range is subdivided into UVA (320–400 nm) light, which is responsible for sun tanning, UVB (280–320 nm) light, which causes sun burning and the danger of skin cancer, and UVC (200–280 nm), which is absorbed by DNA and causes genetic damage and inactivates bacteria and viruses (2). UVC light is often called “germicidal” radiation. Organic molecules absorb energy primarily in the 200 to 300 nm range, and this range evidences the most disinfectant or germicidal properties. Bacterial disinfection is greatest at 260 nm, although other wavelengths also affect cells. A given individual microorganism is almost transparent to UV; it absorbs only a small fraction of the incident UV irradiance. Thus, germicidal effectiveness is directly proportional to the absorbance of the organism, which often peaks around 260 nm and accounts for the maximum germicidal effectiveness at that wavelength.

DNA is a nucleic acid polymer in a double-stranded helix linked together by a sequence of four constituent bases (adenine, cytosine, guanine, and thymine), which constitute the genetic code (2). Of these four bases, thymine undergoes a unique photochemical reaction when exposed to germicidal wavelengths of UV light. If two thymine bases are located adjacent to each other, absorption of a UV photon by one of the thymines leads to formation of a chemical bond between the two thymines (called a thymine dimer). This disrupts the structure of the DNA, so that if enough thymine dimers are formed, the DNA cannot replicate in cell mitosis. This is the fundamental mechanism of UV disinfection or inactivation.

UV DOSAGE

The amount of damage created by UV radiation and hence, the effectiveness of the disinfection process are related to the intensity of light and the exposure time to that intensity. UV dose is defined as the product of the average intensity of light within the reactor and the duration of exposure to that intensity (UV dose = average UV intensity × exposure time). The units of intensity are milliwatts per square centimeter, and those of exposure time are seconds. Dose therefore has the units of milliwatt seconds per square centimeter ($\mathrm{mW{\cdot}s/cm^2}$). Typical UV doses for drinking water range between 16 and 40 $\mathrm{mW{\cdot}s/cm^2}$ depending on the application, water quality, target disinfection level, etc.

MEASUREMENT AND ESTIMATION OF UV DOSE

In a UV reactor, UV irradiance at a certain point can be measured with a radiometer; however, the UV irradiance detected is only measured at a specific point, and only those photons that enter the detector from a narrow acceptance angle can be detected (3,4). This makes it difficult to obtain a direct “measurement” of the UV dose in a reactor configuration (3). Actinometry is a direct method for estimating UV doses. Actinometry measures UV light through a photochemical reaction for which the quantum yield (molecules of product formed per photon absorbed) is well established. A chemical mixture sensitive to UV light at the wavelength of interest is exposed, and the resulting photochemical changes are determined analytically. A quartz cell containing the chemical mixture is inserted into the reactor and exposed to the UV light. The chemical change produces a product over the period of exposure. From the product yield and the quantum yield, one can obtain the total photons incident at that point (3).

Several manufacturers also recommend employing mathematical models such as a multiple point source summation (MPSS) model for estimating UV doses (3,5). Most UV light sources are long narrow lamps. The light output from such a lamp may be approximated by a large number (n) of “point sources” equally spaced along the lamp axis. The light from each point source is assumed to radiate equally in all directions, and the irradiance across a small volume element, a section of an area to be irradiated, in the reactor is then obtained by summing the irradiance at that volume element from all n point sources. Proper account has to be taken of reflection, refraction, and absorption of a beam as it propagates toward the target volume element (3). Once the irradiance is calculated for each volume element in a large three-dimensional grid filling the reactor, the average irradiance is obtained by averaging the irradiance of each volume element over the entire grid. If more than one lamp is in the reactor, the irradiance at a given point is the sum of the irradiance from each lamp.

EFFECTIVENESS

UV irradiation is effective against bacteria at UV doses of 3 to 30 $\mathrm{mW{\cdot}s/cm^2}$ and against viruses at 30 to 100 $\mathrm{mW{\cdot}s/cm^2}$ (2). However, it was thought that protozoa such as *Giardia* and *Cryptosporidium* were unsusceptible to this approach. This inaccurate perception was linked to assays that evaluated UV treatments based on the physical damage they caused to the organisms' outer surfaces. Studies measuring infectivity, however, confirm that UV radiation penetrates the outer walls of *Giardia* and *Cryptosporidium*, causing internal damage and eliminating their threat, even when they are exposed to doses of 30 mW·s/cm^2 or less.

Generally, it has been found that the more complex the microorganism, the more sensitive it is to UV inactivation. Thus, viruses are the least sensitive, then bacterial spores, and finally bacteria are very sensitive. Until recently, protozoa, such as *Cryptosporidium parvum* and *Giardia lamblia*, appeared to go against this trend as it was thought they were very insensitive to UV because of difficulty in penetrating the shell in their cyst or oocyst state.

TYPICAL TREATMENT SYSTEM

A typical UV treatment system consists of a reactor vessel, where a UV lamp irradiates the flowing water. Ideally, uneven distribution is prevented by radial mixing within the reactor, so that all objects moving through it receive equivalent irradiation.

UV SOURCES

Light sources can vary widely. They encompass high- and low-vapor- pressure lamps, broadband sources covering the germicidal region from 200 to 300 nm, and narrowband emitters targeting sensitive wavelengths. UV light is most commonly generated by low- and medium-pressure mercury vapor lamps. In low-pressure mercury lamps, the energy is emitted primarily at 254 nm. A 4-foot, 40 watt (W) fluorescent lamp is a low-pressure mercury lamp with a "fluorescing" pigment on the inner surface. In medium-pressure mercury lamps, the emission is more widely distributed across the 200 to 600 nm range, and the power density is much higher. For example, a 4-foot lamp can carry as much as 30 kilowatts (kW) of electrical power.

FACTORS AFFECTING UV

Dissolved substances in the water—especially organic carbon and nitrates—can absorb UV light. Turbidity also plays a role in the transmittance of light, but mostly due to scattering or reflecting properties. In water, the particles that make up turbidity can also block or hide microbes (some of which can attach to a particle) and then prevent exposure to UV irradiation. However, until the turbidity reaches 5 NTUs (nephelometric turbidity units) or greater, this effect is usually negligible. Other factors include average UV intensity within the reactor received by the microbes in the water, hydraulic behavior of the fluid as it flows through the reactor and the amount of turbulence created, short-circuiting and retention time distribution, pretreatment processes prior to UV disinfection, particulate count and size distribution, and the microbial characteristics of the water to be disinfected (1).

CONCLUSION

UV technology is just beginning to be applied on a large scale to drinking water. The cost of obtaining a total biological barrier with UV disinfection is estimated to be same or less than traditional disinfection methods such as chlorination or ozonation. Because of this, UV disinfection of drinking water could become the most significant advance in protecting drinking water sources since the introduction of chlorination.

BIBLIOGRAPHY

1. Cairns, W.L. (1994). Ultraviolet technology for water supply treatment. *20th Annu. Convention and Exhibition, Water Qual. Assoc.*, March 15–20, 1994, Phoenix, AZ.
2. Bolton, J.R., Dussert, B., Bukhari, Z., Hargy, T., and Clancy, J.L. (1998). Inactivation of *Cryptosporidium parvum* by ultraviolet light: Comparison of laboratory- and pilot-scale results on finished water. *AWWA Annu. Conf.*, Dallas, June 1998.
3. Bolton, J. (2000). Calculation of ultraviolet fluence rates distribution in an annular reactor: Significance of refraction and reflection. *Water Res.* **34**(13): 3315–3324.
4. Severin, B.F. and Roessler, P.F. (1998). Resolving UV photometer outputs with modeled intensity profiles. *Water Res.* **32**(5): 1718–1724.
5. Blatchley, E.R. (1997). Numerical modeling of UV intensity: Application to collimated beam reactors and continuous-flow systems. *Water Res.* **31**(9): 2205–2218.

READING LIST

American Public Health Association (APHA). (1995). *Standard Methods for the Examination of Water and Wastewater*. 19th Edn. AWWA, Denver, CO.

Anderson, W.B., Huck, P.M., Dixon, D.G., and Mayfield, C.I. (in press). Endotoxin inactivation in water using medium pressure UV lamps. *Appl. Environ. Microbiol.*

Hrudey, S.E., Payment, P., Huck, P.M., Gillham, R.W., and Hrudey, E.J. (2003). A fatal waterborne disease epidemic in Walkerton, Ontario: Comparison with other waterborne outbreaks in the developed world. *Water Sci. Technol.* **47**(3): 7–14.

Ollos, P.J., Huck, P.M., and Slawson, R.M. (2003). Factors affecting biofilm accumulation in model distribution systems. *JAWWA* **95**(1): 87–97.

Park, C.H. and Huck, P.M. (2003). A conceptual model for *Cryptosporidium* transport in watersheds. *Water Qual. Res. J. Can.* **38**(1): 77–113.

Parrotta, M.J. and Bekdash, F. (1998). UV disinfection of small Groundwater supplies. *JAWWA* **90**(2): 71–81.

US EPA. (1996). *Ultraviolet Light Disinfection Technology in Drinking Water Application—an Overview*. EPA 811-R-96-002.

Verkoeven, J.W. (1996). Glossary of terms used in photochemistry. *Pure Appl. Chem.* **68**(12): 2223–2286.

Zimmer, J.L., Slawson, R.M., and Huck, P.M. (2003). Inactivation and potential repair of *Cryptosporidium parvum* following low- and medium-pressure UV irradiation. Submitted to *Water Res.*

WATER DISINFECTION USING UV RADIATION—A SRI LANKAN EXPERIENCE

Niranjanie Ratnayake
University of Moratuwa
Moratuwa, Sri Lanka

BACKGROUND

Treated public water supplies are available in most urban areas in Sri Lanka. National Drinking Water Quality Standards have been established, based on WHO Guidelines for Drinking Water Quality, and usually the quality of water leaving the treatment plant is up to the Standards. However, for various reasons such as old pipelines, leaks in the distribution system, low pressure leading to back-siphonage and cross connections, the risk of subsequent contamination of the water is very high. In addition, many users depend on either overhead tanks, which are filled at night, or sump-pump-overhead tank arrangements to store the water, due to low pressure in the distribution system and unreliable supplies. Dual water supplies, where private surface or groundwater supplies are used to supplement the public water supply, are also common due to the inadequacy of the public water supply. Often, the same plumbing system is used for the two systems. The public water supply is treated, but the well water is usually untreated and is pumped directly into the sump or the overhead tank.

The risks of contaminating the drinking water in these situations are considerable. Residual chlorine doses applied at treatment plants are often insufficient to maintain the safety of the water, and many domestic consumers boil water before drinking it. Although boiling is a very reliable method of disinfection of small quantities of water, it is not practical for large-scale uses, such as hostels, canteens, hospitals, and factories. Even though bottled water is available in the market, it is expensive, and the reliability of the cheaper products is questionable. Therefore, it has become necessary to find an appropriate method for disinfecting such water supplies in the light of possible outbreaks of waterborne diseases among users, particularly where dual water supplies are used.

BREAK-POINT CHLORINATION

Chlorination is the most common disinfection method used in water treatment. Chlorine is used in the gaseous form, in solid form as calcium hypochlorite [Ca $(OCl)_2$], or in liquid form as sodium hypochlorite [NaOCl] solution.

Chlorine, a very strong oxidizing agent, reacts with many organic and inorganic substances in water, in addition to disinfecting water. The “chlorine demand” of these other substances has to be satisfied before excess free chlorine is available for disinfection.

Thus the total amount of chlorine added to water is not available for disinfection. Initially, the added chlorine is used immediately to oxidize inorganic and organic substances in reduced form. When this immediate chlorine demand is satisfied, the added chlorine combines with ammonia and similar compounds in the water to form chloramines. Chloramines in water are called “combined available chlorine,” because they are also effective in disinfection, although not to the same extent as free chlorine. Further addition of chlorine has the effect of oxidizing the chloramines formed, thus reducing the available chlorine. When all chloramines are oxidized, further addition of chlorine produces “free available chlorine”, which is the most effective form of chlorine for disinfection. The point at which all oxidation is completed and free available chlorine starts to appear is called the “break point,” and chlorination beyond this point is called “break-point chlorination.”

Thus, to produce a free chlorine residual which would be available to disinfect any contamination within the distribution system at the consumer’s premises, such as in the household plumbing system, overhead tank, or sump, sufficient chlorine should be added to achieve break-point chlorination.

UV RADIATION TECHNOLOGY

UV radiation has been established as an effective bactericide and viricide. UV radiation penetrates the cell wall and is absorbed by cellular nucleic acids. Radiation absorption prevents replication, thus killing the cell. However, the effectiveness of UV radiation in killing spore forming bacteria such as *Giardia lamblia* has not yet been established. UV radiation is not a chemical agent, so it produces no toxic residuals.

A special lamp is used to transfer electromagnetic energy to the target organism’s cells. Mercury arc lamps are the most commonly used type of lamp because about 85% of the energy output is of the 253.7 nm wavelength, which is within the optimum germicidal range of 250–270 nm.

The effectiveness of radiation is a direct function of the energy dose absorbed by the organism, which is measured as the product of the lamp’s intensity and the time of exposure. Intensity is the rate at which photons are delivered to the target organism. It is governed by the lamp’s power, the depth of water, and the presence of substances that absorb the energy, such as suspended solids, color, turbidity, and soluble organic matter. The radiation dose absorbed by the water is the water’s UV demand (which is similar to the chlorine demand) and is quantified as the absorption of UV energy (wavelength 253.7 nm) in a given depth of water. The energy dose also depends on the exposure time of the water to the UV radiation, although the required time is only seconds, rather than minutes, for effective disinfection. However, water that has high concentrations of substances like suspended solids, color, turbidity, and soluble organic matter may not receive an adequate radiation dose for effective disinfection.

If the radiation dose is not sufficient to destroy the target organism’s DNA macromolecules, but only damages them, disinfection would not be effective. Photoenzymatic repair, called “photoreactivation,” occurs from exposure to sunlight, incandescent, or fluorescent light for certain types of bacteria. Therefore, as a rule of thumb, the UV dosage is increased to more than that required to achieve the specified percentage of kill.

MAIN ADVANTAGES AND DISADVANTAGES OF UV RADIATION OVER CHEMICAL METHODS OF DISINFECTION

The main advantages of UV radiation over chemical methods of disinfection are the simplicity of the system for installation and operation, minimal space requirement, absence of any changes in the characteristics of the water and any residual by-products due to reactions in the water, and very short required contact time (seconds rather than minutes) for effective disinfection.

On the other hand, chlorination requires chemical handling and dosing equipment, and ozone and chlorine dioxide need to be produced at the site due to their chemical instability. All chemical disinfectants produce disinfection by-products, by oxidation, reduction, or disproportionation of the disinfectant itself and by reaction of the disinfectant with substances that are already in the raw water. For example, chlorination adds chloride ions to the water and produces trihalomethanes (THMs) and other halogenated organics, which are known to have harmful health effects, when the water contains organic matter. Chlorine dioxide produces chlorite and chlorate ions, which have been identified as harmful, and ozone is also known to produce some harmful by-products if certain organic substances, such as the pesticide heptachlor, are present in the raw water.

The major disadvantage of UV disinfection is the absence of a residual disinfecting mechanism left in the water to be used downstream, as would occur in chlorination and ozonation to some extent. Therefore, it is very important that the UV radiation is carried out as close to the point of consumption as practically possible to prevent recontamination of the treated water. Other disadvantages are the necessity of pretreatment to avoid shielding of microbiological cells by solids and the high cost of the equipment.

CASE STUDY

A UV water purification system that had a maximum flow rate of 60 L/h was used in the laboratory as a case study. The system consisted of an activated carbon filter unit and a UV light unit, as shown in (Fig. 1). Observations were also made on a system whose capacity was 300 L/h; it was installed at the University canteen.

The study was carried out for a period of 2 years. The main objectives of the study were

1. to find the potential of the UV water purification system to destroy pathogens at various flow rates for various sources of water;
2. to study the time taken by the UV equipment to produce disinfected water at the point of consumption; and
3. to compare the performance and cost of disinfection using UV radiation with that of chlorination using bleaching powder.

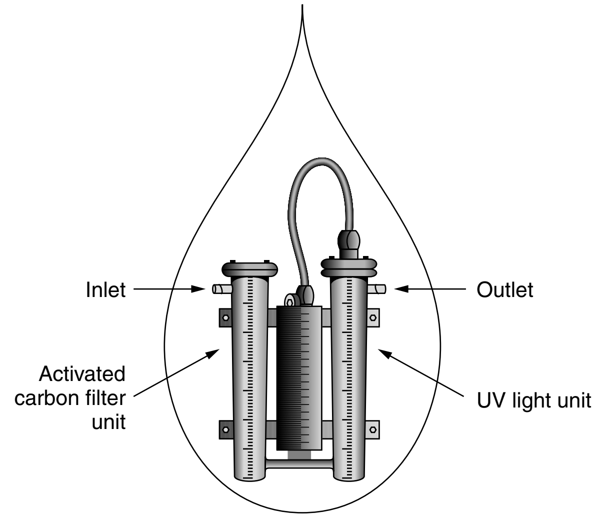

Figure 1. UV water purification system used in the case study.

Table 1. Typical Water Quality at the Water Sources Used in the Case Study

Water Quality Parameter	Location			
	Boat Yard	Ranmal	San Michele	University Well
Total coliform, per mL	167	127	510	97
Fecal coliform, per mL	134	116	481	92
Turbidity, NTU	4.1	3.7	5.7	1.1
DO, mg/L	5.7	6.3	7.7	5.9
Alkalinity, mg/L $CaCO_3$)	0.4	0.4	0.25	0.3
Ammonia, mg/L	1.04	0.56	0.28	1.20
Conductivity	218	1678	925	168.6
pH	7.14	7.43	7.45	7.5
Color, Hazen	40	20	15	5
Temperature, °C	29	29	28.5	29

A simple apparatus was set up using a ground level tank, a pump, a constant head overhead tank to feed the water through the UV purification system at a constant head, and devices to control and measure the flow rate and collect samples of raw and treated water for analysis at various times. Surface water from three locations of a large lake called Bolgoda Lake, several kilometers away from each other, and ground water from a dug well in the University premises were used as raw water sources. A map showing the locations of raw water sources for the case study is given in (Fig. 2). The typical water quality parameters at these water sources are given in Table 1.

Studies on the performance of the UV purification system in the laboratory showed that the total and fecal coliform numbers in the treated effluent from samples of raw water from all four sources, which initially contained high total coliform and fecal coliform, as shown in Table 1, were all zero after passing through the system. The tests were carried out at several flow rates, from as low as

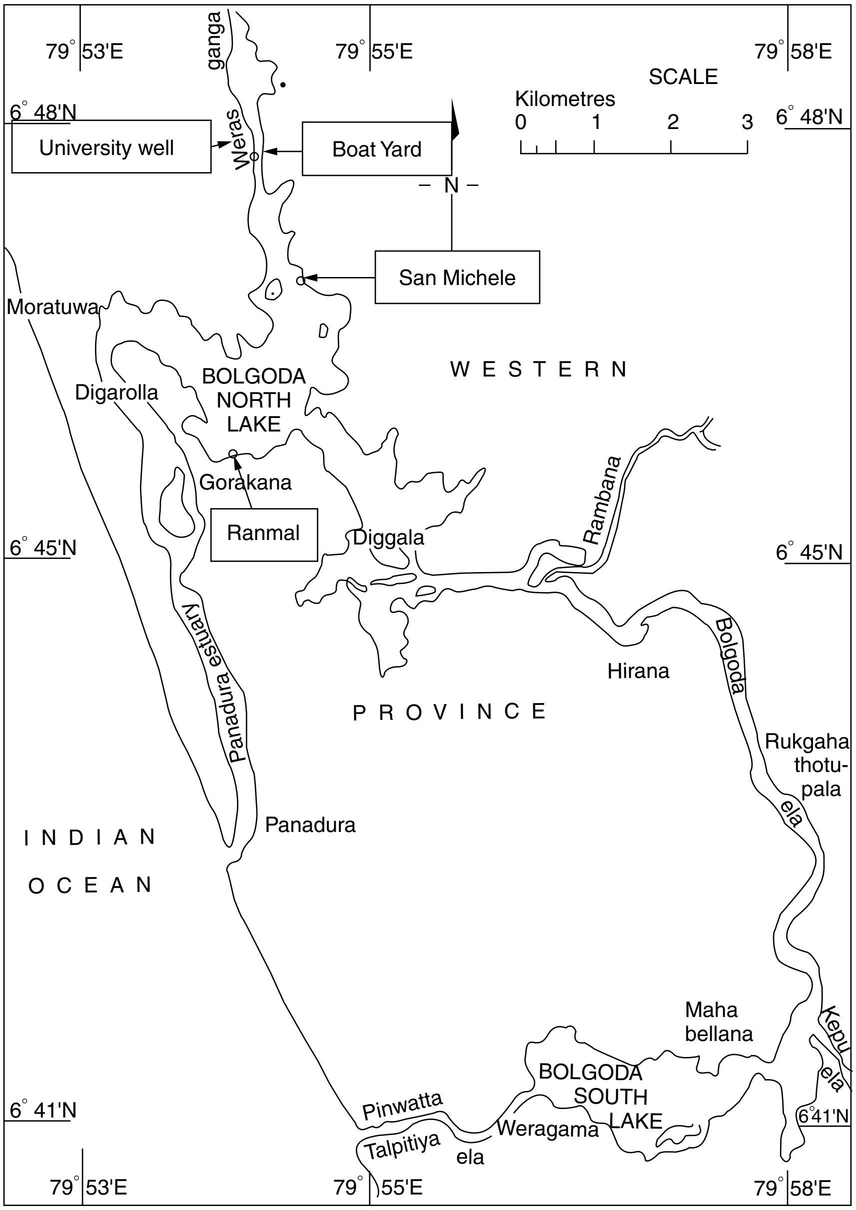

Figure 2. Map of locations of raw water sources taken for the case study.

10 liters per minute to 110 liters per minute, which is almost twice the maximum rate specified for the UV system. The resulting total and fecal coliform contents were always zero, when sampled after 5 minutes of UV lamp operation. The performance of the UV system when used for water from Bolgoda Lake at San Michele and Ranmal Hotel locations and the University well water are shown graphically in (Fig. 3). As can be observed, the total and fecal coliform counts were slightly reduced by the activated carbon filter, but the water cannot be considered safe because the coliforms were still present at all flow rates. However, when sampled 5 minutes after the UV lamp was switched on, none of the samples contained total or fecal bacteria. There is a slight improvement in the turbidity of the water after the activated carbon filter, but no significant change can be seen from the UV radiation.

During the first 5 minutes of system operation, initial studies showed that the coliform numbers do not immediately drop to zero but showed a gradual drop to zero, as shown in (Fig. 4). It was apparent that up to 4 minutes were needed to obtain water that was free from total and fecal coliform bacteria. However, further studies carried out using the same UV system showed that the total and fecal coliform numbers fall to zero when effluent samples are collected as early as after 30 seconds using the same water sources. Tests carried out using the larger UV system which is installed in the University canteen and has a design flow rate of 300 L/h, showed that the total and fecal coliform numbers in the water after passing through the system were zero when tested after 1, 2, 3, 4 and 5 minutes at the design flow rate, as well as double the design flow rate (600 L/h); this is the maximum flow rate possible from the tap. Thus, no

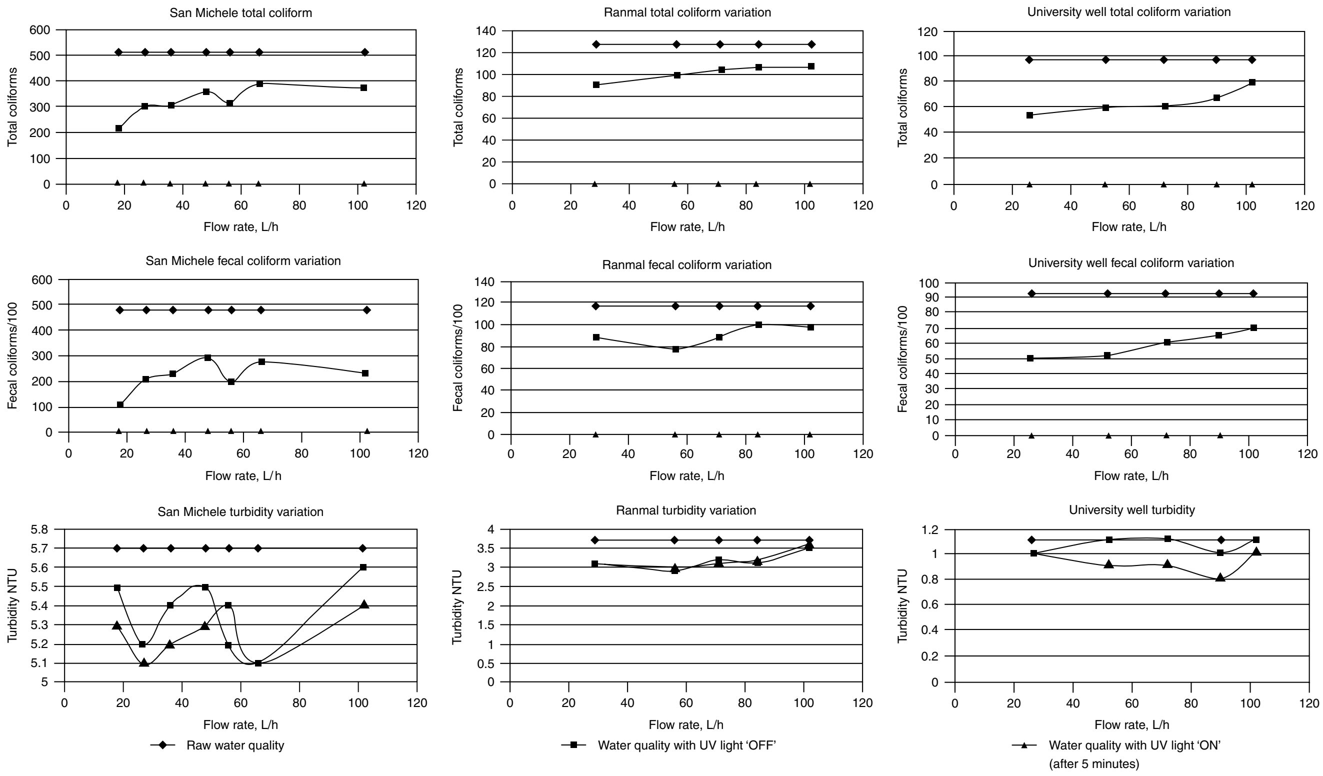

Figure 3. Performance of the UV system on water from different sources at different flow rates.

conclusive evidence could be gathered to determine the required lag time for the water. The time lag experienced earlier could be partly due to the time taken to replace the untreated water in the pipeline, particularly, at low flow rates.

When the system was used with the UV light in the "off" position, there was some improvement in the water quality, including a 20–40% reduction in the coliform content. However, this was not reduced to acceptable levels in any of the trials, and the reduction in total and fecal coliforms was lower at higher flow rates. Thus, it could be inferred that the 100% coliform removal obtained was due to UV radiation. Conversely, it is important to establish the fact that the water would not be safe from pathogenic organisms if the UV light were not turned "on." The function of the activated carbon filter is to remove turbidity, color, and organic matter in the raw water, so that the UV disinfection can be more effective.

A comparative study was done on the same source of raw water, to find the chlorine requirement of the water. It was attempted to find the "break point" for chlorination, but no clear break point was identified, presumably due to the low concentrations of ammonia that were present in the water sources. In raw water from the San Michele location of Bolgoda Lake, a slight depression of residual chlorine could be noticed (see Fig. 5) at an added chlorine concentration of 2.5 to 3.0 mg/L. However, at the Boat Yard location, there was no noticeable break point, but a high initial chlorine demand was noticed, probably due to the more polluted and stagnant status of the water at this location.

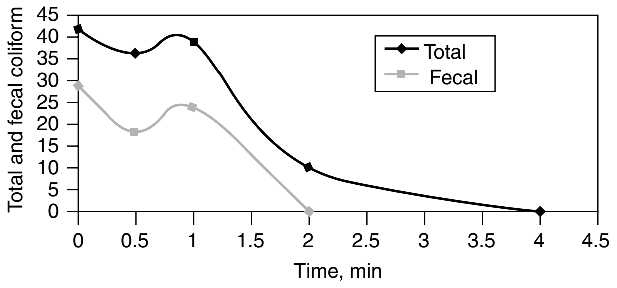

Figure 4. Variation of total and fecal coliform content in treated effluent within the first 5 minutes (initial trials).

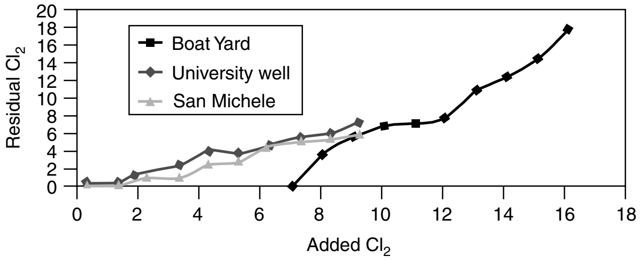

Figure 5. Chlorine demand of the source water.

The following example was used in comparing the cost of the UV treatment and chlorination using bleaching powder:

Water Source: Bolgoda Lake at San Michele location
Selected point of consumption: University canteen
Number of hours used per day: 16
Average flow rate: 300 L/h

— Annual consumption of water: 1,752,000 liters

UV treatment		
Cost of installation of the UV system	SLR 32344.00	US$ 359.38
Assumed lifetime of the unit	5 years	—
Price of a UV bulb	SLR 1200.00	US$ 13.33
Power of a UV bulb for 300 L/h unit	8 watts	—
Lifetime of a UV bulb	8000 hours operation	—
⇒ Number of bulbs required for 5 years	4	—
Cost of bulb per year	SLR 960.00 per annum	US$ 10.60
Annual cost of power for UV unit	SLR 280.00 per annum	US$ 3.11
At a discount rate of 15% per annum,	—	—
capital cost of UV unit as an annuity	SLR 9660.00 per annum	US$ 107.33
⇒ Total Annual Cost of UV unit	**SLR 10,900.00 per annum**	**US$ 121.11 per annum**
Chlorination using bleaching powder		
Required chlorine dosage at break point	2.5 mg/L	—
Required amount of chlorine per year	4.38 kg per annum	—
Percentage of chlorine in bleaching powder used	Approx. 8.9%	—
Required amount of bleaching powder,	44.88 kg per annum	—
	—	—
Total amount of bleaching powder required allowing 10% wastage	50 kg per annum	—
Cost of bleaching powder @Rs. 80 per kg	SLR 4000.00 per annum	US$ 44.44
Assumed annual O&M cost	SLR 12,000.00 per annum	US$ 133.33
⇒ Total Annual Cost of Chlorination	**SLR 16,000.00 per annum**	US$ 177.77

This example showed that the UV radiation equipment is more economical for this application than using bleaching powder for disinfection, considering the local conditions.

CONCLUSION

The case study showed that ultraviolet radiation preceded by activated carbon filtration was a very effective method of disinfecting water from surface as well as groundwater. Thus, it is a very useful method of disinfection when dual water supplies are used, and the treated public water supplies are supplemented by well or surface water, which does not require conventional treatment. The activated carbon unit reduces the color, turbidity, and dissolved organic substances that interfere with disinfection by ultraviolet radiation. However, if the water contains high amounts of the above substances, the filter unit will need frequent replacement. Therefore, it is most suitable for use with well water, rather than surface water. UV disinfection itself is more suitable for well water disinfection, as the disinfection effect of UV radiation on spore forming bacteria such as *Giardia lamblia* is not yet well established.

In comparing the cost of disinfection using UV radiation and chlorination, even though the UV equipment needs a rather high capital investment, the total annual cost of the UV system is lower than that of disinfection using bleaching powder. Considering the ease of operation and ability to have the unit on-line, it is a very attractive option for applications where a large number of users are expected to use water during short periods of time, such as in canteens and hostels.

However, care must be taken to maintain the system and to make sure that the UV lamp is turned on when the water is being used for consumption.

Acknowledgments

The case study was carried out by Messers H.G.C.K. Haththotuwa, S. P. Ranasinghe, K. Amarasinghe, and N.P. Liyanage, undergraduates of the Department of Civil Engineering, University of Moratuwa, under the supervision of the author during 1999, 2000, and 2001.

Next Page

BIBLIOGRAPHY

1. Black & Veach. (1995). EPRI's Community Environment Center December 1995, *UV Disinfection for Water and Waste Water Treatment*. Report CR—105252 www.energy.ca.gov/water/publication/public.
2. Cairns, W.L. *Comparing Disinfection by Ultraviolet Light and Chlorination—The Implications of Mechanism For Practice*. Trojan Technologies, London, Ontario, Canada.
3. Fahey, R.J. (1990). The UV effect on wastewater. *Water Engineering and Management* **137**(12): 1–4.
4. Grombach, P. (1983). *Disinfection in Water Treatment Plants*. Asian National Development. 32.
5. Harris, G.D. et al. (1987). The influence of photoreativation and water quality on UV disinfection of secondary municipal wastewater. *J. Water Pollution Control Federation* **59**: 781.
6. Metcalf & Eddy. (1990). *Waste Water Engineering Treatment, Disposal, Reuse*. New York, pp. 351–352.
7. Pontius Fedrick, W. (1990). *Water Quality and Treatment*. Washington, DC, p. 924.
8. Teresa, G. (1998). Canadian water and wastewater association. *A Technical & Socio-Economic Comparison of Options*, Chapter 5.
9. Thampi, M.V. (1990). Basic guidelines for specifying the design of UV disinfection systems. *J. Pollution Engineering*: 65.
10. US EPA. (1986). *Municipal Wastewater Disinfection Design Manual*. EPA/625/1861021, Washington, DC.
11. Water Pollution Control Federation. (1986). *Wastewater Disinfection*. Manual of Practice NO. FD-10, Washington, DC.
12. White, G.C. (1986). *Handbook of Chlorination*, 2nd Edn. Von Nostrand Reinhold Company, New York, p. 186.
13. World Health Organization. (1993–1998). Guidelines for drinking-water quality, 2nd Edn. Vol. 1: Recommendations (1993), pp. 122–130 & 114–121; Vol. 2: *Health criteria and other supporting information*, (1996) pp. 940–949 & 908–915; Addendum to Vol. 2 (1998) pp. 281–283.
14. *Survey of Effluent Disinfection Practices*. (1996). www.fuseinc.com/1996survey.htm - 95 k.
15. *Water disinfection methods*. www.pureflow.com/Artl.htm - 18 k.
16. *Comparing Disinfection by UV Light and Chlorination. www.trojanuv.com/papers/COMP_UY_CHL.htm - 54 k.*

DRINKING WATER QUALITY STANDARDS (DWQS)-UNITED STATES

Babs A. Makinde-Odusola
Public Utilities
Riverside, California

In 1974, the U.S. Congress, through the Safe Drinking Water Act (SDWA), created the first-ever mandatory national program to protect consumers from harmful contaminants in drinking water. There have been several amendments to the Act—the most recent in 1996. The SDWA requires the U.S. Environmental Protection Agency (EPA) to set national drinking water quality standards (DWQS). The EPA has set primary and secondary DWQS for chemical, physical, radiological, and microbiological contaminants listed in Tables 1 and 2. More Americans are now served by water systems without health risk violations since the EPA (2) began tracking violations around 1980.

INTRODUCTION

Public Water Systems (PWS) are required to comply with the SDWA. A PWS, by definition, provides water to at least 15 connections or 25 persons for at least 60 days out of the year Public Water Supply—United States. PWS provide domestic water to about 275 million Americans (3); about 42 million people, or 16% (4) depend on private sources, such as house wells. EPA does not regulate the quality of water in private wells, but many state and local governments have regulations to protect users of private wells.

Media coverage of waterborne disease outbreaks, such as the 1993 outbreak of cryptosporidiosis in Milwaukee, and popular culture movies, such as "A Civil Action" and "Erin Brockovich" increased public interest in the health risks of contaminated water. At least 50 people died and more than 400,000 people suffered from gastrointestinal illness during the 1993 incident in Milwaukee (7). Many consumers resorted to drinking bottled water. Bottled water is regulated by the U.S. Food and Drug Administration, and is required to meet DWQS. Some consumers rely on "Point of Use" or "Point of Entry" devices to improve the quality of domestic water.

EPA considers costs and health benefits when setting DWQS (see Fig. 1). The annual costs of operating U.S. PWS exceed $3.5 billion (5). The EPA (6) estimated that it would cost about $31 billion (Table 3) over a 20-year period to comply with existing and proposed SDWA regulations. The SDWA requires PWS to have the technical, financial, and managerial capability to provide safe drinking water (2). There are many programs to assist PWS in complying with DWQS, such as the drinking water State Revolving Funds that has a budget authority of $9.6 billion for the fiscal period 1994–2003 [http://www.epa.gov/OW/regs/intro.html].

The SDWA includes a multiple-barrier protection framework against pollution—source water protection, treatment, distribution system integrity, and public information. PWS conduct Source Water Assessments to reduce the potential for contamination (Well Head Protection). The Underground Injection Control (UIC) program of the SDWA and other Federal environmental laws, such as the Clean Water Act (CWA), are important for source water protection (1). U.S. EPA is using source water protection [WELLHEAD PROTECTION], and water safety and security to leverage the linkages between the CWA and SDWA. The Public Health Security and Bioterrorism Preparedness and Response Act of 2002 (SDWA Section 1433(a)) requires certain CWS to conduct Vulnerability Assessments (VAs), and prepare/revise Emergency Response Plans (ERPs) to protect the security ["quality"] of drinking water. Under the CWA, U.S.

BIBLIOGRAPHY

1. Black & Veach. (1995). EPRI's Community Environment Center December 1995, *UV Disinfection for Water and Waste Water Treatment*. Report CR—105252 www.energy.ca.gov/water/publication/public.
2. Cairns, W.L. *Comparing Disinfection by Ultraviolet Light and Chlorination—The Implications of Mechanism For Practice*. Trojan Technologies, London, Ontario, Canada.
3. Fahey, R.J. (1990). The UV effect on wastewater. *Water Engineering and Management* **137**(12): 1–4.
4. Grombach, P. (1983). *Disinfection in Water Treatment Plants*. Asian National Development. 32.
5. Harris, G.D. et al. (1987). The influence of photoreativation and water quality on UV disinfection of secondary municipal wastewater. *J. Water Pollution Control Federation* **59**: 781.
6. Metcalf & Eddy. (1990). *Waste Water Engineering Treatment, Disposal, Reuse*. New York, pp. 351–352.
7. Pontius Fedrick, W. (1990). *Water Quality and Treatment*. Washington, DC, p. 924.
8. Teresa, G. (1998). Canadian water and wastewater association. *A Technical & Socio-Economic Comparison of Options*, Chapter 5.
9. Thampi, M.V. (1990). Basic guidelines for specifying the design of UV disinfection systems. *J. Pollution Engineering*: 65.
10. US EPA. (1986). *Municipal Wastewater Disinfection Design Manual*. EPA/625/1861021, Washington, DC.
11. Water Pollution Control Federation. (1986). *Wastewater Disinfection*. Manual of Practice NO. FD-10, Washington, DC.
12. White, G.C. (1986). *Handbook of Chlorination*, 2nd Edn. Von Nostrand Reinhold Company, New York, p. 186.
13. World Health Organization. (1993–1998). Guidelines for drinking-water quality, 2nd Edn. Vol. 1: Recommendations (1993), pp. 122–130 & 114–121; Vol. 2: *Health criteria and other supporting information*, (1996) pp. 940–949 & 908–915; Addendum to Vol. 2 (1998) pp. 281–283.
14. *Survey of Effluent Disinfection Practices*. (1996). www.fuseinc.com/1996survey.htm - 95 k.
15. *Water disinfection methods*. www.pureflow.com/Artl.htm - 18 k.
16. *Comparing Disinfection by UV Light and Chlorination. www.trojanuv.com/papers/COMP_UY_CHL.htm - 54 k.*

DRINKING WATER QUALITY STANDARDS (DWQS)-UNITED STATES

Babs A. Makinde-Odusola
Public Utilities
Riverside, California

In 1974, the U.S. Congress, through the Safe Drinking Water Act (SDWA), created the first-ever mandatory national program to protect consumers from harmful contaminants in drinking water. There have been several amendments to the Act—the most recent in 1996. The SDWA requires the U.S. Environmental Protection Agency (EPA) to set national drinking water quality standards (DWQS). The EPA has set primary and secondary DWQS for chemical, physical, radiological, and microbiological contaminants listed in Tables 1 and 2. More Americans are now served by water systems without health risk violations since the EPA (2) began tracking violations around 1980.

INTRODUCTION

Public Water Systems (PWS) are required to comply with the SDWA. A PWS, by definition, provides water to at least 15 connections or 25 persons for at least 60 days out of the year Public Water Supply—United States. PWS provide domestic water to about 275 million Americans (3); about 42 million people, or 16% (4) depend on private sources, such as house wells. EPA does not regulate the quality of water in private wells, but many state and local governments have regulations to protect users of private wells.

Media coverage of waterborne disease outbreaks, such as the 1993 outbreak of cryptosporidiosis in Milwaukee, and popular culture movies, such as “A Civil Action” and “Erin Brockovich” increased public interest in the health risks of contaminated water. At least 50 people died and more than 400,000 people suffered from gastrointestinal illness during the 1993 incident in Milwaukee (7). Many consumers resorted to drinking bottled water. Bottled water is regulated by the U.S. Food and Drug Administration, and is required to meet DWQS. Some consumers rely on “Point of Use” or “Point of Entry” devices to improve the quality of domestic water.

EPA considers costs and health benefits when setting DWQS (see Fig. 1). The annual costs of operating U.S. PWS exceed $3.5 billion (5). The EPA (6) estimated that it would cost about $31 billion (Table 3) over a 20-year period to comply with existing and proposed SDWA regulations. The SDWA requires PWS to have the technical, financial, and managerial capability to provide safe drinking water (2). There are many programs to assist PWS in complying with DWQS, such as the drinking water State Revolving Funds that has a budget authority of $9.6 billion for the fiscal period 1994–2003 [http://www.epa.gov/OW/regs/intro.html].

The SDWA includes a multiple-barrier protection framework against pollution—source water protection, treatment, distribution system integrity, and public information. PWS conduct Source Water Assessments to reduce the potential for contamination (Well Head Protection). The Underground Injection Control (UIC) program of the SDWA and other Federal environmental laws, such as the Clean Water Act (CWA), are important for source water protection (1). U.S. EPA is using source water protection [WELLHEAD PROTECTION], and water safety and security to leverage the linkages between the CWA and SDWA. The Public Health Security and Bioterrorism Preparedness and Response Act of 2002 (SDWA Section 1433(a)) requires certain CWS to conduct Vulnerability Assessments (VAs), and prepare/revise Emergency Response Plans (ERPs) to protect the security [“quality”] of drinking water. Under the CWA, U.S.

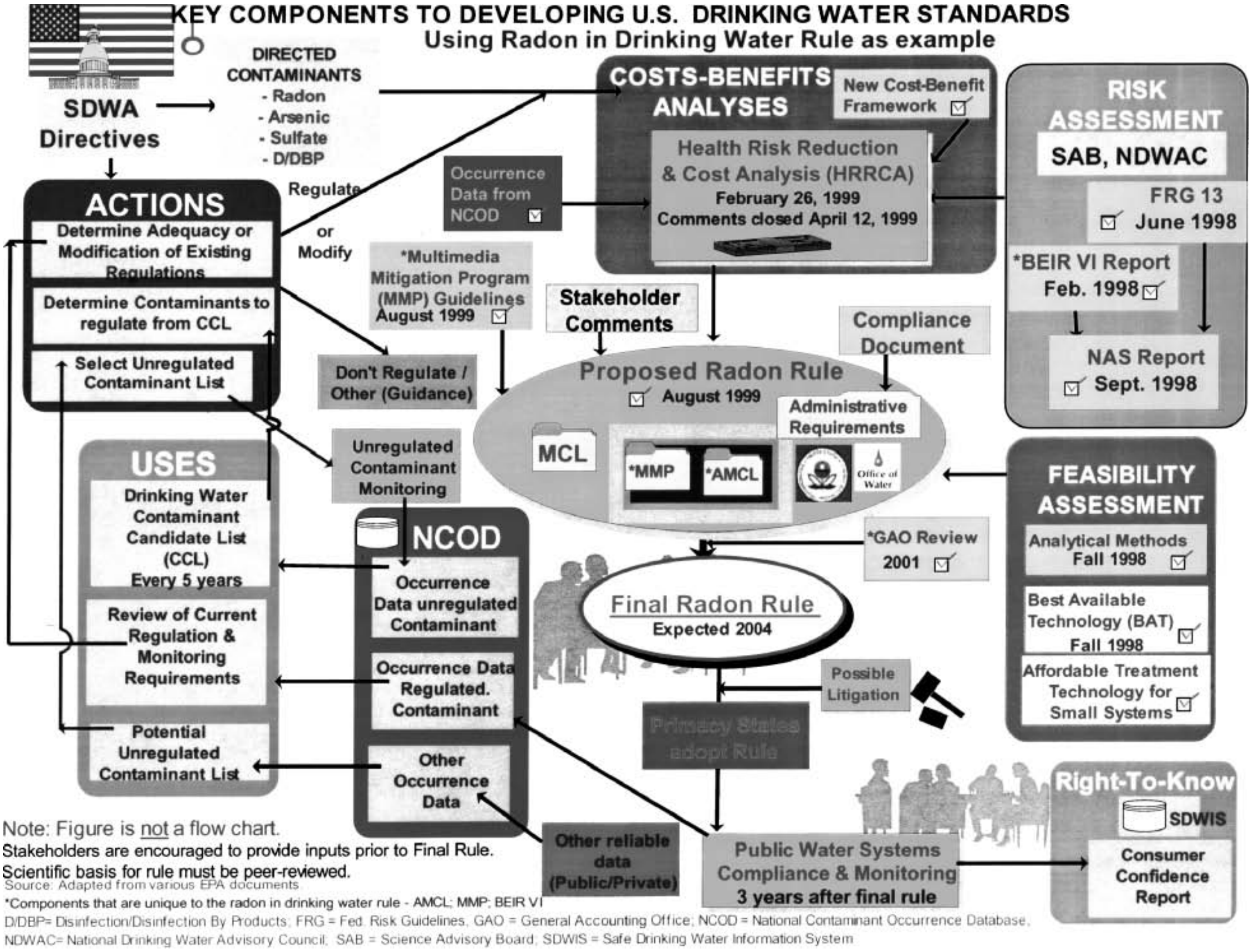

Note: Figure is not a flow chart.
Stakeholders are encouraged to provide inputs prior to Final Rule.
Scientific basis for rule must be peer-reviewed.

Source: Adapted from various EPA documents

*Components that are unique to the radon in drinking water rule - AMCL; MMP; BEIR VI

D/DBP= Disinfection/Disinfection By Products; FRG = Fed. Risk Guidelines, GAO = General Accounting Office; NCOD = National Contaminant Occurrence Database, NDWAC= National Drinking Water Advisory Council; SAB = Science Advisory Board; SDWIS = Safe Drinking Water Information System

Figure 1. Key components to developing and monitoring drinking water standards.

EPA established the Human Health Water quality criteria [http://epa.gov.waterscience/humanhealth/15table-fs.htm], which though not DWQS are numeric values that protect human health from the harmful effects of pollutants in ambient water based solely on science without considering socio-economic impacts. DWQS standards are often regarded as "Applicable or Relevant and Appropriate Requirements" (ARARs) in cleanup decisions affecting polluted water. Water conservation programs may reduce the need to use poorer quality sources for drinking water.

DEFINITIONS

Many terms are used to describe the numerical and/or narrative qualities of drinking water. Such terms include regulations, standards, goals, levels, limits, advisories, etc. (8). Terms used by the EPA (9,10) in published drinking water regulations include

Maximum Contaminant Level Goal (MCLG): The level of a contaminant in drinking water below which there is no known or expected risk to health. MCLGs allow a margin of safety and are nonenforceable public health goals. Table 1 is a list of the MCLGs.

Maximum Contaminant Level (MCL): The highest level of a contaminant that is allowed in drinking water. EPA sets the MCL as close to the MCLG as feasible, using the best available treatment technology (BAT) and considering cost. MCLs are enforceable standards and are listed in Table 1.

Maximum Residual Disinfectant Level Goal (MRDLG): The level of a drinking water disinfectant below which there is no known or expected risk to health (Table 1). MRDLGs do not reflect the benefits of using disinfectants to control microbial contaminants.

Maximum Residual Disinfectant Level (MRDL): The highest level of a disinfectant allowed in drinking water (Table 1).

Secondary MCLs or "SMCLs" are established for nuisance chemicals only as guidelines for aesthetic considerations, such as taste, color, and odor and are not considered to present a risk to human health at the SMCL. SMCLs (Table 2) are unenforceable at the federal level. Some states enforce SMCLs at the state level (1).

Treatment Technique (TT): A required process to reduce the level of a contaminant in drinking water. EPA

Table 1. United States Environmental Protection Agency National Primary Drinking Water Standards[1]

Microorganisms	MCLG, mg/L	MCL or TT, mg/L	Organic Chemicals	MCLG, mg/L	MCL or TT, mg/L
Cryptosporidium	zero	TT	Acrylamide	zero	TT
Giardia lamblia	zero	TT	Alachlor	zero	0.002
Heterotrophic plate count	n/a	TT	Atrazine	0.003	0.003
Legionella	zero	TT	Benzene	zero	0.005
Total coliforms[2]	zero	5.00%	Benzo(a)pyrene (PAHs)	zero	0.0002
Turbidity	n/a	TT	Carbofuran	0.04	0.04
Viruses (enteric)	zero	TT	Carbon tetrachloride	zero	0.005
	–	–	Chlordane	zero	0.002
Inorganic chemicals	–	–	Chlorobenzene	0.1	0.1
Antimony	0.006	0.006	2,4-D	0.07	0.07
Arsenic (01/01/06)	0	0.01	Dalapon	0.2	0.2
Asbestos (fiber $>$ 10 micrometers)	7 MFL	7 MFL	1,2-Dibromo-3-chloropropane (DBCP)	zero	0.0002
Barium	2	2	*o*-Dichlorobenzene	0.6	0.6
Beryllium	0.004	0.004	*p*-Dichlorobenzene	0.075	0.075
Cadmium	0.005	0.005	1,2-Dichloroethane	zero	0.005
Chromium (total)	0.1	0.1	1,1-Dichloroethylene	0.007	0.007
Copper (action level = 1.3)	1.3	TT	*cis*-1,2-Dichloroethylene	0.07	0.07
Cyanide (as free cyanide)	0.2	0.2	*trans*-1,2-Dichloroethylene	0.1	0.1
Fluoride	4	4	Dichloromethane	zero	0.005
Lead action level = 0.015	zero	TT	1,2-Dichloropropane	zero	0.005
Mercury (inorganic)	0.002	0.002	Di(2-ethylhexyl) adipate	0.4	0.4
Nitrate (measured as nitrogen)	10	10	Di(2-ethylhexyl) phthalate	zero	0.006
Nitrite (measured as nitrogen)	1	1	Dinoseb	0.007	0.007
Selenium	0.05	0.05	Dioxin (2,3,7,8-TCDD)	zero	3E-08
Thallium	0.0005	0.002	Diquat	0.02	0.02
	–	–	Endothall	0.1	0.1
Radionuclides[3]	–	–	Endrin	0.002	0.002
Alpha particles	zero	15 pCi/L	Epichlorohydrin	zero	TT
Beta particles and photon emitters	zero	4 mrem/yr	Ethylbenzene	0.7	0.7
Radium-226 and Radium-228 (combined) as of 12/08/03	zero	5 pCi/L	Ethylene dibromide	zero	0.00005
Uranium (12/08/03)	zero	30 µg/L	Glyphosate	0.7	0.7
	–	–	Heptachlor	zero	0.0004
Disinfectants & DBP	MRDLG	MRDL	Heptachlor epoxide	zero	0.0002
Bromate	zero	0.01	Hexachlorobenzene	zero	0.001
Chloramines (as Cl_2)	4	4	Hexachlorocyclopentadiene	0.05	0.05
Chlorine (as Cl_2)	4	4	Lindane	0.0002	0.0002
Chlorine dioxide (as ClO_2)	0.8	0.8	Methoxychlor	0.04	0.04
Chlorite	0.8	1	Oxamyl (Vydate)	0.2	0.2
Haloacetic acids (HAA5)	n/a	0.06	PCBs	zero	0.0005
Total trihalomethanes (TTHMs)	n/a	0.08	Pentachlorophenol	zero	0.001
	–	–	Picloram	0.5	0.5
	–	–	Simazine	0.004	0.004
	–	–	Styrene	0.1	0.1
	–	–	Tetrachloroethylene	zero	0.005
	–	–	Toluene	1	1
	–	–	Toxaphene	zero	0.003
	–	–	2,4,5-TP (Silvex)	0.05	0.05
	–	–	1,2,4-Trichlorobenzene	0.07	0.07
	–	–	1,1,1-Trichloroethane	0.2	0.2
	–	–	1,1,2-Trichloroethane	0.003	0.005
	–	–	Trichloroethylene	zero	0.005
	–	–	Vinyl chloride	zero	0.002
	–	–	Xylenes (total)	10	10

Notes: TT: Treatment Technique; pCi/L: picocuries per Liter MFL:million fibers per liter; PCBs: Polychlorinated biphenyls; mrem/yr: millirems per year; DBP: Disinfection by-products; MCL: Maximum Contaminant Level; MCLG: Maximum Contaminant Level Goal; MRDL: Maximum Residual Disinfectant Level; MRDLG: Maximum Residual Disinfectant Level Goal.

[1] There are many footnotes, not reproduced here.

[2] Including fecal coliform and *E. coli*

[3] Proposed radon rule: MCLG = 0, MCL = 300 pCi/L; Alternate MCL = 4,000 pCi/L + an approved Multimedia Mitigation Program.

Log on to URL of source for full details.

Primacy states may have more stringent standards.

Adapted from: http://www.epa.gov/safewater/mcl.html specifies a TT only if it is not "economically or technologically feasible" to ascertain the concentration of a contaminant (2,8).

Action Level: "The concentration of a contaminant which, if exceeded, triggers treatment or other requirements which a water system must follow. For lead or copper, it is the level which, if exceeded in more than 10% of the homes tested, triggers treatment (9).

Alternative Maximum Contaminant Level (AMCL): Unique to radon in drinking water, it is the level of radon in drinking water that would not increase the level of radon in indoor air beyond the natural level of radon in outdoor air RADON IN WATER.

Health Risk Reduction and Cost Analysis (HRRCA): The goal of the HRRCA is to provide a neutral and factual analysis of the incremental costs, benefits, and other impacts of controlling contaminant levels in drinking water based on system size.

Best available technology (BAT): The field-proven feasible treatment technique for removing contaminants from drinking water below the MCL. Historically, the EPA has based feasibility on technologies appropriate for large PWS, but the 1996 SDWA requires EPA to specify compliance technologies that can be feasibly implemented by small water systems.

Health Advisories [http://www.epa.gov/ost/drinking/standards] describe the levels of a contaminant at which adverse health effects would not be anticipated for different durations of exposure (10).

Some states have additional standards. The Office of Environmental Health Hazard Assessment (OEHHA) within the California EPA sets a Public Health Goal (PHG) based exclusively on health risk assessment. A PHG [http://www.oehha.org/water/phg/allphgs.html] is set at the level at which no known or anticipated adverse effect on health will occur with an adequate margin of safety; considering possible synergistic effects resulting from exposure to two or more contaminants; and the existence of groups in the population that are more susceptible to adverse effects of the contaminant. PHG is nonregulatory, but another state agency, the California Department of Health Services (DHS), is required to use the PHG as the basis for setting MCL.

Action Level (AL): In California, an AL is a health-based advisory level for a contaminant for which an MCL has not been adopted. DHS sets an AL when one of the following occurs: (1) a chemical is found in an actual or proposed drinking water source; or (2) a chemical is in proximity to a drinking water source, and guidance is needed, should it reach the source. DHS set the AL at 4 μg/L for perchlorate, the main active explosive ingredient in rocket fuel [http://www.dhs.ca.gov/ps/ddwem/chemicals/AL/actionlevels.htm].

Table 2. Secondary Maximum Contaminant Levels

Parameter	SMCL	Noticeable Effects Above the Secondary MCL
Aluminum	0.05 to 0.2 mg/L*	Colored water
Chloride	250 mg/L	Salty taste
Color	15 color units	Visible tint
Copper	1.0 mg/L	Metallic taste; blue-green staining
Corrosivity	Non-corrosive	Metallic taste; corroded pipes/fixtures staining
Fluoride	2.0 mg/L	Tooth discoloration
Foaming agents (MBAS)	0.5 mg/L	Frothy, cloudy; bitter taste; odor
Iron	0.3 mg/L	Rusty color; sediment; metallic taste; reddish or orange staining
Manganese	0.05 mg/L	Black to brown color; black staining; bitter metallic taste
Odor	3 TON	"Rotten-egg," musty or chemical smell
pH	6.5–8.5	Low pH: bitter metallic taste; corrosion high pH: slippery feel; soda taste; deposits
Silver	0.1 mg/L	Skin discoloration; graying of the white part of the eye
Sulfate	250 mg/L	Salty taste
TDS	500 mg/L	Hardness; deposits; colored water; staining; salty taste
Zinc	5 mg/L	Metallic taste
Bolero[1,2]	0.001 mg/L	Odor is detected at a lower concentration than its taste
MtBE[1]	0.005 mg/L	–

Notes: MtBE: Methyl tertiary-butyl ether
TON: Threshold Odor Number
TDS: Total Dissolved Solids
[1] California SMCL
[2] Also known as Thiobencarb
Adapted from: http://www.epa.gov/safewater/consumer/2ndstandards.html
http://www.dhs.ca.gov/ps/ddwem/chemicals/MCL/secondarymcls.htm
*MCL for Aluminum in California is 1 mg/L

HISTORY OF U.S. DWQS AND SDWA

Clean drinking water used to mean clear, palatable, nonodorous water. It was not until the early 1900s that DWQS other than for general clarity existed (2), even though the first U.S. community water system (CWS) began in 1799 (8). By 1900, many of the more than 3000 CWSs then contributed to major disease outbreaks because they "provided an efficient vehicle for the delivery of pathogenic bacteria" (8).

Federal authority to regulate drinking water began with the Interstate Quarantine Act of 1893 (1). In 1914, the U.S. Public Health Services (USPHS) set bacteriological DWQS for water systems that provided drinking water to interstate carriers such as ships and trains, as directed by the Act. Those standards were commonly referred to as "Treasury Standards" (1). In 1915, the federal government committed to reviewing DWQS on a regular basis (8). USPHS revised and expanded the standards in 1925, 1946, and 1962 (1). All 50 states adopted the USPHS

Table 3. 20-Year Regulatory Need (in Millions of January 1999 Dollars)

Regulations	Current Need	Future Need	Total Need
Existing SDWA regulations	–	–	–
Surface Water Treatment Rule[1]	$14,492.1	$4,873.3	$19,365.4
Total Coliform Rule[1]	$358.1	$112.8	$470.9
Nitrate/Nitrite Standard[1]	$197.1	$31.9	$229.0
Lead and Copper Rule	$1,039.6	$186.5	$1,226.2
Total Trihalomethanes Standard	$39.1	$60.6	$99.7
Other Regulations[2]	$430.8	$85.4	$516.2
Subtotal National Need	$16,556.9	$5,350.4	$21,907.4
Costs Associated with Proposed and Recently Promulgated Regulations (Taken From EPA Economic Analyses)[3]	–	$9,324.3	$9,324.3
Total National Needs	$16,556.9	$14,674.8	$31,231.7

Note: Numbers may not total due to rounding.

[1] Regulations for contaminants that cause acute health effects.

[2] Includes regulated VOCs, SOCs, IOCs, and Radionuclides

[3] Includes regulations for contaminants that cause acute and/or chronic health effects. In the Economic Analyses, the compliance costs for some regulations are given as a range. In calculating the $9.3 billion need, the survey used EPA's lead option, unless one was not available in which case the survey used the more conservative estimate.

Source: U.S. EPA. Drinking Water Infrastructure Needs Survey: 2nd Report to Congress. Office of Water, Washington, D.C. EPA 816-R-01-004, February 2001. http://www.epa.gov/safewater/dwsrf.html

standards with minor modifications either as regulations or as guidelines even though the federal government did not mandate them (11).

In the late 1960s, several studies found that many man-made chemicals from sources, such as intensive agriculture and manufacturing plants, were contaminating water supplies. One of the environmental laws passed to protect public health is the SDWA in 1974. The 1974 SDWA required EPA to establish legally enforceable DWQS for all PWS in the United States and a surveillance system for PWS. The 1974 law also included a schedule and procedures for developing new drinking water standards, which are to be reviewed at least every 3 years. In 1975, the EPA promulgated interim regulations based on the 1962 USPHS standards, with some additional health-effects information (1). The SDWA was amended and/or reauthorized in 1977, 1979, 1980, 1986, 1988, and 1996 (2).

The SDWA amendments of 1986 declared that the interim standards promulgated in 1975 are final and required EPA to regulate 83 contaminants within 3 years after enactment (2). EPA was required to regulate an additional 25 contaminants every 3 years and to designate the best available treatment technology (BAT) for each regulated contaminant. EPA was directed to require all PWS to disinfect. Filtration was mandated for surface water systems.

The SDWA Amendments of 1996 [http://www.epa.gov/safewater/sdwa/sdwa.html] modified the regulatory schedule and included new sections such as criteria and procedures for variance, operator certification requirements, and provisions for capacity development (1). The EPA published guidelines for certifying and recertifying operators of PWS [http://www.epa.gov/safewater/opcert/plan.html].

STANDARDS SETTING

EPA promulgates two types of drinking water regulations—the National Primary Drinking Water Regulations (NPDWRs) and the National Secondary Drinking Water Regulations (NSDWRs). NPDWRs address health risks/effects that may be acute or chronic and include enforceable maximum limits for drinking water contaminants (MCL) and disinfectants. A NPDWR also includes BAT, monitoring, reporting, recordkeeping provisions, and Regulatory Impact Analyses (Fig. 1). NSDWRs address drinking water aesthetics such as taste and odor that render drinking water unpalatable or undesirable to use. NSDWRs (SMCL) are not enforceable at the federal level. Regulating aesthetic standards deters consumers from seeking more pleasant but less safe sources of water (12). In California, the SMCL for methyl tertiary-butyl ether (MtBE) (5 μg/L) is lower than its MCL (13 μg/L).

The EPA (13) usually follows these steps when setting DWQS (Fig. 1): Determine whether to regulate a contaminant based on peer-reviewed science; set an MCLG DRINKING WATER QUALITY—STANDARDS AND CRITERIA, DRINKING WATER QUALITY—STANDARDS, REGULATIONS AND GOALS; propose an MCL or treatment technique (TT); finalize by, setting an enforceable MCL or TT; and authorize states to grant variances for PWS that serve up to 3300 people based on affordability criteria. Stakeholders can participate in the process.

The EPA selects from the Contaminant Candidate List (CCL) that contains the names of unregulated contaminants. The EPA is to revise the CCL [http://www.epa.gov/safewater/ccl/ccl_fr.pdf], every 5 years. The EPA has developed and maintains a publicly available database known as the National Contaminant Occurrence Database (NCOD), which contains information on regulated and unregulated contaminants that can be found in drinking water [http://www.epa.gov/ncod/].

COMPLIANCE AND SUPERVISION

The EPA administers the SDWA through primacy states. The SDWA allows the EPA Administrator to treat

Tribes, District of Columbia, Guam, Puerto Rico, the Northern Mariana Islands, the Virgin Islands, American Samoa, and the Trust Territory of the Pacific Islands as states for purposes of primacy. Primacy states usually adopt EPA DWQS, but some states set standards that are more stringent than those of the EPA. Nothing precludes a state from establishing DWQS for any water constituent not regulated by the EPA. The EPA, the states, the PWS, and the public cooperatively implement the SDWA.

The EPA provides grants to primacy states to assist states in developing and implementing drinking water programs and provides oversight in enforcing the standards. The EPA (2) has granted primacy to all states but Wyoming. None of the tribal governments has yet been granted primacy. The EPA has primary enforcement authority in states without primacy. States must adopt newly promulgated EPA regulations to maintain primacy status.

PWS are required to collect and analyze water samples at designated intervals and locations. Water samples must be analyzed in approved laboratories and the results reported to the state, which determines compliance. The three main types of violations (2) are

- MCL violation: occurs when contaminant level in treated water exceeds the MCL.
- TT violation: occurs when a PWS fails to treat water as prescribed by the EPA.
- Monitoring and reporting violation: occurs when a PWS either fails to test its water for certain contaminants or fails to report test results in a timely fashion.

An MCL or TT violation indicates a potential health risk. The EPA (2) estimated that in 1994, 89% of the population was served by CWS with neither MCL nor TT violations. The EPA plans to achieve a 95% level by 2005 (2). PWS in violation must notify the public. The EPA maintains a Federal database of violations known as Safe Drinking Water Information (SDWIS) [http://www.epa.gov/enviro/html/water.html#SDWIS].

The EPA, some water-related national associations, and more than 200 surface water utilities throughout the United States established a voluntary cooperative effort known as the Partnership for Safe Water [http://www.awwa.org/partner/partner2.htm?]. Utility members of the Partnership implement preventive programs to increase safety where legislation or regulation does not exist. The preventive measures are based around optimizing treatment plant performance and thus increasing protection against microbial contamination in the drinking water supply. PWS must provide their customers with "Consumer Confidence Reports" (CCR) CONSUMER CONFIDENCE REPORTS.

FUTURE TREND

In the 1980s and 1990s, the EPA set DWQS using the "Risk Assessment/Risk Management" (RA/RM) paradigm based on targeted mandates that prioritize and manage environmental risks, pollutant by pollutant and medium by medium. The proposed radon rule allowed the EPA to address health risk reduction within a multimedia framework. The EPA Science Advisory Board (14) proposed a conceptual framework for integrated environmental decision-making to "guide the Agency in the continuing evolution of environmental decision-making."

BIBLIOGRAPHY

1. American Water Works Association (AWWA). (2001). *Safe Drinking Water Advisor: A Compliance Assistance Resource*. AWWA, May 2001 CD-ROM Edition.
2. EPA. (1999). *25 Years of the Safe Drinking Water Act: History and Trends*. EPA 816-R-99-007.
3. EPA. (2000). Providing Safe Drinking Water in America. Office of Enforcement and Compliance Assurance, Washington, DC. EPA 305-R-00-002. http://www.es.epa.gov/ncepihom/index.html.
4. Solley, W.B., Pierce, R.R., and Perlman, H.A. (1998). Estimated Use of Water in the United States in 1995. U.S. Geological Survey Circular 1200. Denver, CO. http://water.usgs.gov/watuse/pdf1995/html/.
5. EPA. (1999). *Water Facts*. EPA 810-F-99-020.
6. EPA. (2001). Drinking Water Infrastructure Needs Survey: 2nd Report to Congress. Office of Water, Washington, DC. EPA 816-R-01-004. http://www.epa.gov/safewater/dwsrf.html.
7. Davis, J.P. et al. (1996). *Cryptosporidiosis-Associated Mortality Following a Massive Waterborne Outbreak in Milwaukee, Wisconsin*. Bureau of Public Health, Madison, WI.
8. Cotruvo, J.A. and Vogt, C.D. (1990). Chapter 1, Rationale for Water Quality Standards and Goals. *Water Quality and Treatment: A Handbook of Community Water Supplies*. 4th Edn. F.W. Pontius (Ed.). American Water Works Association, McGraw Hill, New York, pp. 1–62.
9. EPA. (2000). *Drinking Water Standards and Health Advisories*. Office of Water, Washington, DC. EPA 822-B-00-001.
10. EPA. (2001). *National Primary Drinking Water Standards*. Office of Water, Washington, DC. EPA 816-F-01-007.
11. Pontius, F.W. and Clark, S.W. (1999). Chapter 1, Drinking Water Quality Standards, Regulations and Goals. Water Quality and Treatment. 5th Edn. *American Water Works Association*. R.D. Letterman (Ed.). McGraw Hill, New York.
12. Tate, C.H. and Arnold, K.F. (1990). Chapter 2, Health and aesthetic aspects of water quality. *Water Quality and Treatment: A Handbook of Community Water Supplies*. 4th Edn. F.W. Pontius (Ed.). American Water Works Association, McGraw Hill, New York, pp. 63–156.
13. EPA. (1999). *Drinking Water Standards & Health Effects*. EPA 810-F-99-017.
14. EPA Science Advisory Board. (2000). *Toward Integrated Environmental Decision-Making*. EPA-SAB-EC-00-011.

See also CONSUMER CONFIDENCE REPORTS; DRINKING WATER QUALITY—STANDARDS AND CRITERIA; DRINKING WATER QUALITY—STANDARDS, REGULATIONS AND GOALS; PUBLIC WATER SUPPLY—UNITED STATES; RADON IN WATER; and WELL HEAD PROTECTION.

VALVES

Zacharia Michael Lahlou
Technical Assistance
Consultant

Valves direct, start, stop, mix, or regulate the flow, pressure, or temperature of a fluid. Valves range from simple water faucets to control valves equipped with microprocessors. Many different valve types exist; however, the most common types include gate, plug, ball, butterfly, check, pressure relief, and globe valves.

WHAT ARE THE VALVE FUNCTIONS?

Valve functions vary based on the position of the closure element in the valve. The closure position can be adjusted manually or automatically. Valves usually fall into one of three classes:

1. shut-off valves block the flow or allow it to pass;
2. anti-reversal valves allow flow to travel in one direction; and
3. throttling valves regulate flow at a point between fully open to fully closed.

However, specific valve-body designs may fit into one, two, or all three classifications.

WHAT ARE THE BASIC VALVES?

Manual Valves (See Figure 1)

Manual valves require manual operation, such as a hand wheel or lever, which are primarily used to stop and start flow (shut-off valves), although some designs can be used for basic throttling. A manual valve operator is any device that requires the presence of a human being to operate the valve, as well as to determine the proper action (open, closed, or a throttling position). Manual valves are also used to divert or combine flow through a three- or four-way design configuration. Four types of manual valves exist:

1. rotating valves, such as plug, ball, and butterfly valves;
2. stopper valves, such as globe and piston valves;
3. sliding valves, such as gate and piston valves; and
4. flexible valves, such as pinch and diaphragm valves.

Check Valves (See Figure 2)

Check valves are automatic valves that open with forward flow and close against reverse flow. Check valves, also known as non-return valves, prevent return or reverse flow and maintain pressure. Check valves do not require an outside power supply or a signal to operate. In fact, a check valve's operation depends upon the direction in which the water is flowing. A pump or a pressure drop may determine water flow. If the flow stops or if pressure conditions change so that flow begins to move backward, the check valve's closure element moves with the reverse

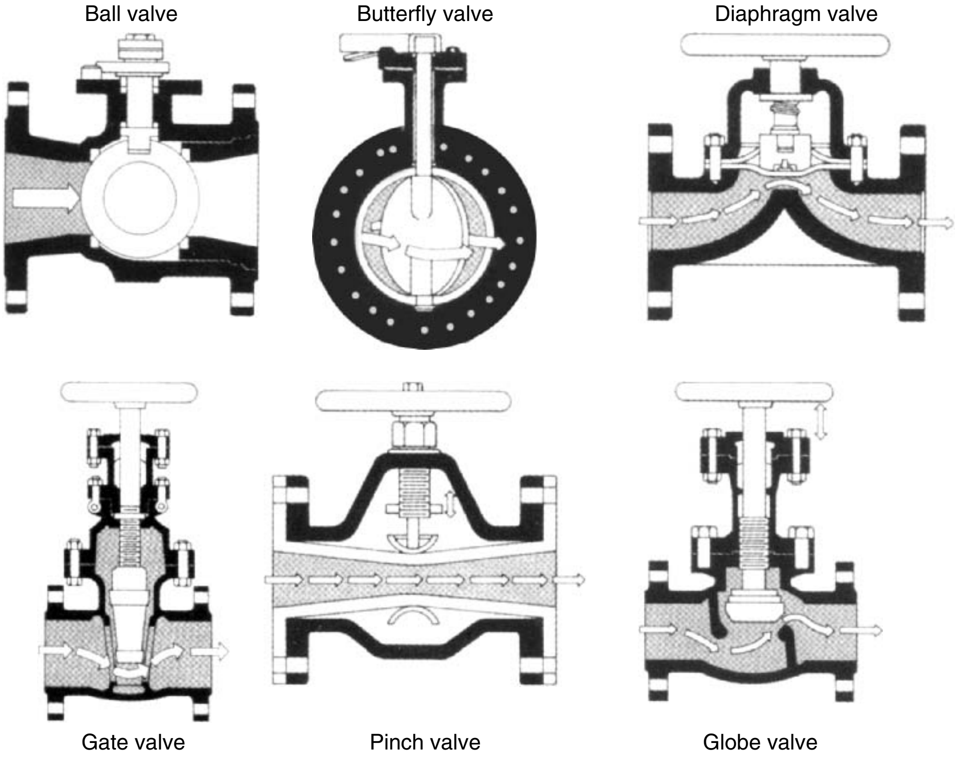
Ball valve
Butterfly valve
Diaphragm valve
Gate valve
Pinch valve
Globe valve

Figure 1. Manual valves.

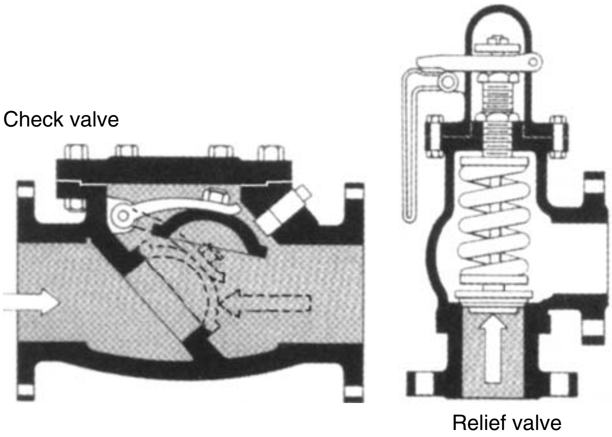

Figure 2. Check valves.

flow until it is seated, preventing any backward flow. There are different types of check valves, but they all have the same operating principle. Check valves include lift-check valves, swing-check valves, tilting-disk valves, split-disk valves, and diaphragm-check valves.

Pressure Relief Valves

A pressure-relief valve is used to protect against overpressurization of the system. When excess line pressure is detected, the pressure-relief valve automatically opens and relieves the excess pressure. A pressure-relief valve is installed where excess pressure constitutes safety concerns, such as pipes or equipment bursting. Following the depressurization of the water line to safe or normal limits, the pressure-relief valve automatically closes again to allow for normal system operation.

Control Valves (See Figure 3)

Control valves, also known as automatic control valves, are used to regulate flow anywhere from fully open to fully closed. Control valves are a fast growing segment of the valve industry because of the quickening pace of water-industry automation. Control valves are almost always equipped with some sort of actuator or actuation system (See definition under Manual Operators and Actuators below.) Manually operated valves and pressure regulators can stand alone in a throttling application, while a control valve cannot. Hence the difference: a control valve is a throttling valve, but not all throttling valves are control valves.

MANUAL OPERATORS AND ACTUATORS

With most valves, some mechanical device or external system must be devised to open or close the valve or to change the position of the valve. Manual operators, actuators, and actuation systems are those mechanisms that are installed on valves to allow this action to take place. Automatic valve control requires an actuator, which is defined as any device mounted on a valve that, in a response to a signal, automatically moves the valve to the required position using an outside power source.

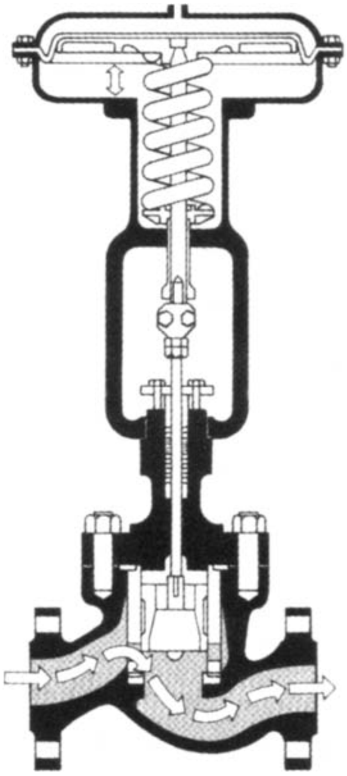

Figure 3. Control valves.

WHAT ARE COMMON VALVE PROBLEMS?

Pressure drop or pressure differential, which is the difference between the upstream and downstream pressures, makes water flow move through a valve. If the piping size is identical both upstream and downstream from the valve and the velocity is consistent, the valve will use frictional losses to reduce the fluid pressure and create flow. Because the pressure drop that a valve generates absorbs energy through frictional losses, an ideal pressure drop allows the full flow to pass through the valve's body without excessive velocity, absorbing less energy.

However, some systems may need to take a larger pressure drop through the valve. A high-pressure drop through a valve creates a number of problems, such as cavitation, flashing, choked flow, high noise levels, and vibration. Such problems present a number of immediate consequences: erosion or cavitation damage to the body and trim, malfunction or poor performance of the valve itself, attached instruments will not remain calibrated, piping fatigue, or hearing damage to nearby workers. In these instances, valves in high-pressure-drop applications require expensive trims, more frequent maintenance, large spare-part inventories, and piping supports. Such measures drive up maintenance and engineering costs.

Although users typically concentrate on the immediate consequences of high-pressure drops, the greatest threat a high-pressure drop presents is lost system efficiency. Usually, a pump adds pressure and, thus, energy to the system. As the system absorbs more energy, including the energy that valves with high-pressure drops lose, it must use larger pumps. Consequently, if the system is designed with few valves with high-pressure drops, the system is more efficient and able to use smaller pumps.

Cavitation

Cavitation happens when low-pressure bubbles suddenly form and then collapse within a small area of the valve within microseconds. Minor cavitation damage may be considered normal for some applications, which can be dealt with during routine maintenance. If unnoticed or unattended, severe cavitation damage can limit the life expectancy of the valve. It can also create excessive leakage, distort flow characteristics, or cause the eventual failure of the valve body and piping. In some severe high-pressure drop applications, cavitation can destroy valve parts within minutes.

One of three basic actions can control or eliminate cavitation. Operators can:

- modify the system;
- make certain internal body parts out of hard or hardened materials; or
- install special devices in the valve that are designed to keep cavitation away from valve surfaces or prevent the formation of the cavitation itself.

Flashing

When the downstream pressure is equal to or less than the vapor pressure, the vapor bubbles generated stay intact and do not collapse. This phenomenon is known as flashing. When flashing occurs, the fluid downstream is a mixture of vapor and liquid moving at a very high velocity, which results in erosion in the valve and in the downstream piping.

Unfortunately, eliminating flashing completely involves modifying the system itself, in particular the downstream pressure or the vapor pressure. However, not all systems are easily modified and this may not be an option. When flashing occurs, no solution can be designed into the valve, except possibly using hardened trim materials.

Choked Flow

The presence of vapor bubbles that cavitation or flashing cause significantly increase the specific volume of the fluid. This increase rises at a faster rate than the increase that the pressure differential generates. If upstream pressure remains constant, decreasing the downstream pressure will not increase the flow rate. Choked flow must be considered when sizing a valve.

HIGH VELOCITIES

Large pressure differentials create high velocities through a valve and in downstream piping. This in return creates turbulence and vibration if the velocities are not lowered. Lower velocities will reduce problems associated with flashing and erosion.

WATER HAMMER EFFECTS

A valve that is opened too quickly or slammed shut when the closure element is suddenly sucked into the seat (“bathtub stopper effect”) as the valve nears shutoff may cause water-hammer effect. Although water hammer generates considerable noise, the real damage occurs through mechanical failure. Water hammer can burst or damage piping supports and connections. In valves, water hammer can create severe shock through trim, gasket, or packing failure.

With valves, the best defense against water hammer is to prevent any sudden pressure changes to the system. For example, one solution may involve slowing the closure of the valve itself. Adding some type of surge protection to the piping system can also reduce water hammer.

WHAT ARE THE OPERATION AND MAINTENANCE REQUIREMENTS?

To avoid mechanical failure, design a practical valve maintenance program, which can result in cost savings for the water system. A job that appears to be mammoth becomes less daunting if the operator implements a systematic maintenance schedule. Operation and maintenance procedures for various types of valves are included in the manufacturer’s operation manuals and in the appropriate product standards. Valve records are essential for planning, operating, and verifying the system’s integrity. The valve record should contain information about valve condition, testing, and maintenance required.

WHAT ABOUT SAFETY/TERRORISM?

Recent events have understandably heightened concern that water supplies may be vulnerable to terrorist attacks—biological, chemical, and structural. Relief from this concern can come via knowledge, planning, and preparedness. It is crucial now to regularly inspect the location, accessibility, and operation of all the valves in the distribution system. This inspection will reveal the condition of the valve box and chamber.

It also is imperative to update the location of the valves on the map. E. H. Wachs Companies for example, offer “Valvecard” software that allows water utilities to manage their valve distribution system from a desktop computer. The software can be used to create valve inventory databases that require a secure user-friendly system to input, store, and analyze valve data or to collect valve information directly from each valve in a system. The software also can be used to operate and exercise valves in the system and record this information or monitor the system and keep it operational at all times. It can perform queries regarding valve properties, location, activities, and conditions as well as interface with Geographic Information Systems and map files.

After having complete and accurate information about all the valves in the system, the utility manager can run through a large number of “what if” scenarios to determine where a contaminant is likely to move and how valves or other utility operations can control its movement. To do this, the utility needs a calibrated, extendedperiod simulation model. There are a few types of software on the market for water distribution modeling and engineering that allow system managers to view scenarios graphically. Some of this software, such as WaterCad by Haestad Methods, can give a clear picture of what is occurring or will occur in the system in response to any operational measures that management proposes.

U.S. Environmental Protection Agency and the American Water Works Association in collaboration with Sandia National Laboratories continue to develop a risk- and consequence-based vulnerability assessment technique to improve the safety and security of water supply and distribution systems against emerging physical, chemical, and biological threats.

WHERE CAN I FIND MORE INFORMATION?

American Water Works Association. 1996. *Distribution Valves: Selection, Installation, Field Testing, and Maintenance (M44)*. AWWA: Denver, CO.

Haestad Methods. 2002. Proceedings from Water Security Summit. Haestad: Waterbury, CT.

National Environmental Services Center. 2002. *Small Drinking Water System Security*. NESC: Morgantown, WV.

Skousen, P. L. 1998. *Valve Handbook*. McGraw-Hill: New York, NY.

Zappe, R.W. 1999. *Valve Selection Handbook*. Fourth Edition. Gulf Publishing Company: Houston, TX.

Walski, T., D. Chase, and D. Savic. 2001. *Water Distribution Modeling*. Haestad Methods: Waterbury, CT.

TO ORDER OUR TECH BRIEFS

Call the NDWC at (800) 624-8301 or (304) 293-4191. You also may order online at *ndwc_orders@mail.estd.wvu.edu*. Additional copies of *Tech Briefs* are free. You also may download fact sheets from our Web site at www.ndwc.wvu.edu.

For further information, to comment on this fact sheet, or to suggest topics, contact the On Tap *editors via e-mail at* Kathy.Jesperson@mail.wvu.edu or Mark.Kemp-Rye@mail.wvu.edu.

Illustrations Source: McKetta, John J., Weismantel, Guy E. 1997. *Encyclopedia of Chemical Processing and Design*. Marcel Dekker, INC: New York, NY.

Z. Michael Lahlou holds a doctorate in environmental and natural resource economics, a masters in civil and environmental engineering, and an MBA. Formerly the technical assistance coordinator for the National Drinking Water Clearinghouse, Lahlou now resides in Huntington Beach, CA.

HAVE YOU READ ALL OUR TECH BRIEFS?

Tech Briefs, drinking water treatment and supply fact sheets, have been a regular feature in the National Drinking Water Clearinghouse (NDWC) publication *On Tap* for more than six years.

A package of *Tech Briefs* is now available as a product. A three-ring binder holds all the current *Tech Briefs* in print. New selections can easily be added to the package as they become available. To order this product, call the NDWC at the numbers listed at the bottom of this ad and ask for item #DWPKPE71. **The item is free.**

Additional copies of the *Tech Brief* fact sheets are also free; however, postal charges maybe added.

REMOVAL OF PATHOGENIC BACTERIA, VIRUSES, AND PROTOZOA

Kwok-Keung Au
Greeley and Hansen LLC
Chicago, Illinois

INTRODUCTION

Pathogenic bacteria, viruses, and protozoa can be removed from potable water by various processes. These removal processes can be categorized into three groups: (1) *initial*, (2) *principal*, and (3) *alternative/supplemental*. *Initial removal processes* are used to remove microbial pathogens before they enter or as they enter a treatment plant. The main objective of these processes is to reduce the particulate loadings, including microbial loadings, to the subsequent treatment processes of a treatment plant. Examples of initial removal processes are roughing filters, microstrainers, off-stream storage, bank filtration, and presedimentation with coagulation. *Principal removal processes* are those most frequently used to remove the majority of the microbes in the water being treated. These processes are sedimentation, flotation, and high-rate granular media filtration. They are often used in conjunction with coagulation and flocculation. *Alternative/supplemental removal processes* are used either as (1) alternatives to principal removal processes or (2) supplemental processes to enhance drinking water quality. These processes include slow sand filtration, precoat filtration, membrane filtration, bag filtration, and cartridge filtration.

The performance of removal processes is often quantified in terms of the log removal of pathogens, which is defined as log C_0/C, where C_0 and C are the pathogen concentrations in the water entering and leaving a process, respectively. A 1-log removal is equivalent to 90% removal, and a 2-log removal is equivalent to 99% removal. Drinking water regulations in the United States typically target log removals of pathogens. However, the performance of a removal process can also be quantified by the turbidity level in the outlet flow of the process or the removal of turbidity by the process. Turbidity is a measure of water clarity. A low turbidity level in the water leaving a treatment process is considered an indication that the process effectively removes pathogens.

INITIAL REMOVAL PROCESSES

Roughing Filters

Roughing filters are coarse granular media filters used to remove large particles from high turbidity waters prior to downstream processes. Although the term filter is used, roughing filters act more as a series of sedimentation boxes filled with media of decreasing size in the direction of flow. The media effectively reduce the settling distance of particles to the order of a few millimeters. The direction of flow through a roughing filter can be either horizontal or vertical. Typical roughing filter media are gravel, rock, crushed coconut, or other locally available materials.

According to a pilot study (1), a roughing filter can achieve average log removals of 0.28 for total bacteria and 0.20 for algal cells. It was also found that the removal of clay particles is improved when the filter has been ripened with algae. Another pilot study indicates a log removal of 0.64 to 0.96 for fecal coliforms, depending on filtration rate (2).

Roughing filters are often used in developing countries to serve as an initial process for slow sand filtration. Important design parameters for roughing filters are hydraulic filtration rate, filter depth (for vertical filters) or length (for horizontal filters), and media size. A review of design variables for roughing filters is available (2).

Microstrainers

Microstrainers are filtration units that use a thin medium, usually fabrics woven of stainless steel or polyester wires. The mesh sizes of microstrainer fabrics are less than 100 μm and typically from 15 to 45 μm. Microstrainers can be used to remove fine suspended particles and algal cells from surface waters. They are very effective for certain types of green algae, diatoms, and cyanophyta. A major application of microstrainers in potable water treatment is as an initial removal process for downstream membrane processes. They reduce particle loading to the membrane processes and also protect the membranes from debris.

The removal efficiencies of algae by microstrainers can vary from 0.22 to 0.52 logs, depending on the type of algae present (3). Large protozoa such as *Balantidium coli* cysts (whose diameter is about 60 μm) can also be removed by microstrainers. However, due to their small size, the removal of bacteria and viruses by microstrainers is usually insignificant.

Off-Stream Storage

Off-stream storage refers to reservoirs or basins whose locations in the treatment flow path are between the raw water source and treatment processes of a water treatment plant. Off-stream storage can improve microbial water quality through a combination of mechanisms. The major mechanisms are settling and die-off of pathogens and equalization of water quality spikes. The critical factors for the performance of off-stream storage are the storage's hydraulic residence time, flow patterns within the storage, and water temperature. Control of algal growth within the storage is important because algae can cause taste and odor problems in water. It is also very critical to avoid fecal contamination from animals and runoff from surrounding areas. Otherwise, the storage may actually deteriorate the quality of the water leaving it.

Settling is the primary mechanism by which pathogens are removed by off-stream storage. The longer the hydraulic residence time of off-site storage, the greater the amount of settling that takes place in it. The removal of coliform bacteria in off-stream reservoirs with hydraulic residence times greater than 40 days is from 0.7 to 2 logs (4). A 1.5 log reduction of enteroviruses was reported for an off-stream reservoir with a hydraulic residence time of about 100 days (5). In two studies, reductions of 1.4 to 2.0 logs for *Cryptosporidium*, 2.3 to 2.6 logs for *Giardia*, 2.2 logs for *Escherichia coli*, and 1.7 logs for fecal streptococci were found for storage reservoirs with residence times of 24 weeks or greater (6),(7). These results indicate that, for settling to be an effective method of removing microbes in off-stream storage, a long residence time and, therefore, a large land area are required.

Although not a removal process, the die-off of pathogens within off-site storage can improve microbial water quality. The die-off of pathogens is a kinetic process and may be considered a first-order reaction (8). Based on bench scale studies, the die-off rate constants found for *Cryptosporidium* are 0.01/day at 5 °C and 0.024/day at 15 °C (9). These are equivalent to a 0.5 log reduction in 50 days at 5 °C and in 21 days at 15 °C, indicating the significance of temperature in the die-off process.

Off-stream storage can also reduce the peak values of pathogen concentration spikes through the management of the source water pumping schedule. The peak levels of microbial contaminants are often associated with the turbidity spikes that follow rainfalls. By not using source water from the watercourse during these peaks and using the stored water instead, pathogen spikes of microbial contaminants can be avoided.

Bank Filtration

Bank filtration is a process by which surface water from a watercourse infiltrates into a groundwater aquifer and is later brought to ground level by well pumps. As the water passes through the ground, the aquifer acts as a natural filter and reduces the concentration of viable pathogens by two major mechanisms: deposition of the pathogens on porous aquifer materials and die-off of pathogens due to the long residence time within the aquifer. This process is also termed riverbank infiltration. It has been widely used as a water treatment process in European countries and, in the past few years, has received increased interest in the United States.

In a full-scale bank filtration study, no *Giardia* or *Cryptosporidium* were detected in wells, even though they were frequently detected in river water, and total coliform bacteria were never detected in wells (10). In a study of three different infiltration sites, the removals of bacteriophage were 3.1 logs and 4.0 logs when water passed through 2 m and 4 m of very fine dune sand, respectively (11).

Bank filtration can be highly effective in removing microbial pathogens. Aquifers suitable for bank filtration are composed of unconsolidated, granular, fine-grained materials and have open, interconnected pores that allow groundwater to flow (12). Other important factors affecting the treatment efficiency of bank filtration are the time of travel from the watercourse to the well (thus, the residence time of the water in the aquifer), the quality of the source surface water (e.g., turbidity), and the water temperature. USEPA proposes a treatment credit for *Cryptosporidium* of either 0.5 or 1.0 logs for bank filtration systems that meet some specific criteria. These criteria include aquifer compositions, distance from the pumping wells to surface water, and turbidity level in the well effluent (12).

Presedimentation with Coagulation

Presedimentation is used to remove gravel, sand, and other relatively large suspended materials from source water prior to principal treatment processes. It is often used for source waters that have highly variable raw water quality to reduce occasionally high solid loadings on the principal treatment processes. Microbial pathogens can be removed by this process if they are attached to or trapped inside larger particles. Adding coagulants can enhance the removal of pathogens by presedimentation.

According to a full-scale study of four presedimentation basins using surface loading rates of from 0.3 to 1.6 gpm/ft^2 and polymer and ferric sulfate as coagulants, the removal of *Bacillus subtilis* averaged 1.1 logs (12). The results of a full-scale study of inclined-plate presedimentation basins using a surface loading rate of 0.52 gpm/ft^2 and alum and polymer as coagulants indicated that the median removal of total aerobic spores and *Giardia* were 0.5 logs and 1.5 logs, respectively (12).

PRINCIPAL REMOVAL PROCESSES

Sedimentation

Sedimentation is a solid–liquid separation process that removes particles and pathogens from water by providing conditions that will cause them to settle under the force of gravity. The hydraulic loading per unit of surface area, hydraulic detention time, inlet/outlet conditions, and internal flow patterns are considered important design parameters for sedimentation basins, but the most important factor in removing microbial pathogens from sedimentation basins is the use of coagulation and flocculation processes upstream of the basins.

Due to their small size and low mass density, individual microbes have a very low settling velocity. In the absence of proper coagulation and flocculation, sedimentation does not achieve significant removals of microbial pathogens. Coagulation involves adding coagulants that either have or can form positively electrically charged species to neutralize negatively charged microbial particles. Because microbial particles are of like electrical charges, they repel each other. Neutralizing these charges facilitates contact between particles. When two or more particles come in contact, they form larger particles that settle more readily. The coagulant can also form precipitates or floc particles that attach to the naturally occurring particles in the water and form larger, heavier agglomerates. The flocculation process establishes conditions that increase the number of opportunities for particles and agglomerates to contact one another and result in large particles that have sufficient settling velocities to be removed in the sedimentation basins that follow.

To achieve good removal of pathogens by sedimentation, the type and dose of coagulant and the coagulation chemistry must be optimized. Also important in coagulation are rapid, thorough mixing, and gentle, relatively slow mixing and sufficient contact time in the flocculation process.

Based on pilot studies, it was reported that sedimentation achieved an average removal of 0.2 logs for *Cryptosporidium* under suboptimal coagulation conditions and the removal increased to 1.3 logs under optimal coagulation conditions (13). This indicates the importance of coagulation, as mentioned above.

Based on full-scale observations at treatment plants in various countries, sedimentation preceded by coagulation and flocculation achieved average removals of 0.14 to 0.59 logs for viruses, 0.17 to 0.89 logs for bacteria (total coliforms or fecal streptococci), and 0 to 1.22 logs for algae (14).

Flotation

Unlike sedimentation which removes materials whose mass densities are greater than that of water, flotation removes materials whose mass densities are less than that of water. In potable water treatment, the flotation of materials is induced and is termed dissolved air flotation (DAF). In DAF, dissolved air is injected into the process stream under high pressure. Then, the process stream is exposed to atmosphere pressure, and the dissolved air forms air bubbles. The bubbles form around and attach to particles in the process stream. This causes particle/air bubble agglomerates to have a mass density that is less than that of water and, thus, to float to the surface of the liquid where they are skimmed off. DAF is most applicable to waters with heavy algal blooms or waters with low turbidity, low alkalinity, and high color. These waters are difficult to treat using sedimentation because the coagulant floc particles produced typically have a low settling velocity.

The performance of DAF for *Cryptosporidium* oocyst removal depended on the pH, coagulant dose, flocculation time, and the ratio of the air-saturated water stream to process stream. According to several full-scale studies, DAF removes 1.4 to 2.0 logs of the algae *Aphanizomenon* and *Microcystis* (3). Study also shows that DAF achieved between 2 and 2.6 logs removal of *Cryptosporidium* oocysts, whereas conventional sedimentation resulted only in 0.6 to 0.8 log removals (15). Based on bench-scale testing, it was reported that DAF averaged 0.5 logs higher removal of *Cryptosporidium* than sedimentation (16).

High Rate Granular Filtration

High-rate granular filtration is the most widely used potable water treatment process for removing pathogens. This process is described separately in GRANULAR BED AND PRECOAT FILTRATION.

ALTERNATIVE/SUPPLEMENTAL REMOVAL PROCESSES

Slow Sand Filtration

Slow sand filtration is described separately in GRANULAR BED AND PRECOAT FILTRATION.

Precoat Filtration

Precoat filtration is described separately in GRANULAR BED AND PRECOAT FILTRATION.

Membrane Filtration

Membrane filtration is a separation process in which a thin, semipermeable membrane is used to remove contaminants from water. A differential hydraulic pressure across the membrane causes the water to pass through it. This differential pressure can be created by applying a higher pressure on one side of the membrane than the other and pushing the water through the membrane or by applying a negative pressure on one side and pulling the water through.

The most commonly used membrane processes in drinking water treatment are microfiltration (MF), ultrafiltration (UF), nanofiltration (NF), and reverse osmosis (RO). They differ from each other by their pore sizes. MF membranes have the largest pore size (of the order of 0.1 μm), followed by UF membranes (of the order of 0.01 μm) and NF membranes (of the order of 0.001 μm). RO membranes have the smallest pore size of the order of 0.0001 μm.

The pore size of the membrane is the key factor in determining which contaminants are removed by the membrane. Contaminants large than the membrane's pore size are retained and removed by the membrane. The pore size of the membrane also affects the operating hydraulic of the membrane process. Membranes that have smaller pore sizes require higher differential hydraulic pressures. The operating pressure of MF and UF is similar and from 30 to 50 kPa. It increases to 500 to 1000 kPa for NF and to 1000 to 5000 kPa for RO.

The sizes of protozoa and most bacteria are of the order of a micron or larger and can therefore be removed very effectively by MF. Removal of bacteria may sometimes be poor, if bacteria grow in the membrane system. Viruses have sizes from 0.01 to 0.1 μm and are likely to pass through MF membranes, unless coagulation pretreatment is provided. According to a pilot study using three MF membranes with pore sizes from 0.08 to 0.22 μm, log removals of greater than 4.6 to greater than 7.0 for *Giardia* and greater than 4.2 to greater than 6.9 for *Cryptosporidium* were achieved (17). Removals of MS2 bacteriophage were, however, less than 1 log. This was expected due to the small size of MS2, which is 0.025 μm. In a pilot study using MF with coagulation pretreatment, virus removals were more than 7 logs, even though the tested virus has a size of 23 nm and the membrane pore size is 100 nm (0.1 μm) (18).

In addition to bacteria and protozoa, UF can remove some viruses. In a study using UF membranes with pore sizes of 0.01 to 0.05 μm, neither *Giardia* nor *Cryptosporidium* were detected in the filtered water (17). The log removals were greater than 4.7 to greater than 7.0 for *Giardia* and greater than 4.4 to greater than 7.0 for *Cryptosporidium*. The membrane having the smallest pore size achieved MS2 bacteriophage removals of 6 logs and greater.

Because of their high costs, NF and RO are seldom used to remove microbial pathogens alone. NF is primarily used for softening because it can remove divalent, hardness-causing cations such as Ca^{++} and Mg^{++} and for removals of disinfection by-product precursors (dissolved organic carbon compounds). RO is primarily used for desalination because it can remove monovalent, salinity-causing ions such as Na^{+} and Cl^{-}. However, the ability of NF and RO to remove pathogens enhances their cost-effectiveness when they are used for these other treatment objectives.

During a bench-scale study using five different RO membranes, virus reductions of 2.7 logs to greater than 6.5 logs were achieved (19). A pilot study was conducted to investigate the efficiency of sequential membrane systems, using two different NF membranes with two different MF membranes as pretreatment (20). The removal of *Bacillus subtilis* endospores, a surrogate for *Cryptosporidium* oocysts and *Giardia* systs, varied from 8.0 to 11.0 logs.

Detailed descriptions of their design and operation are available in the literature (21). Maintenance of membrane integrity was very critical to process efficiency. The loss of membrane integrity (e.g., fiber breakage) certainly permits passage of pathogens into finished water.

Bag Filters and Cartridge Filters

A differential pressure is used to push the water through both bag filters and cartridge filters. Bag filters use nonrigid, fabric filter media bags housed in a pressure vessel. Water flows from inside the bags to outside the bags. Cartridge filters are quite the opposite. They use a rigid or semirigid, fabric filter media housed in a pressure vessel, and water flows from outside the cartridges to the inside of the cartridges.

These types of filters are often used for small systems and for point-of-use filtration applications. They can also be applied as an additional filtration process following the principle filtration processes such as high-rate granular filtration to improve water quality. High turbidity and algae cells can easily clog bag and cartridge filters, so they are appropriate only for high quality water. Addition of disinfectant prior to the filters should also be considered to minimize biofilm growth within filter media.

The typical pore size range for bag and cartridge filters is from 0.2 to about 10 μm. This is usually small enough to remove protozoa such as *Cryptosporidium* oocysts and *Giardia* cysts. Submicron-sized microbes, such as viruses and most bacteria, can pass through these filters. As water passes through a bag or cartridge filter, particles accumulated within the filter material increase the hydraulic pressure drop through the filter. When this pressure drop increases to a level that it is impractical for further operation, the bag or cartridge is replaced with a clean one.

Microbial removal by bag and cartridge filters can vary significantly. A range of 0.5 to 3.6 logs *Cryptosporidium* removal has been reported for bag and cartridge filters (12). Interestingly, there was not a correlation between nominal pore size and removal because the nominal pore size of a bag or cartridge filter medium, as reported by the manufacturer, represents only an average size. A pore size distribution exists, which means that some pores are of sizes greater than the nominal size and the pore size distribution values vary from one filter medium to another.

BIBLIOGRAPHY

1. Collins, M.D., Westersund, C.M., Cole, J.O., and Roccaro, J.V. (1994). *Evaluation of Roughing Filtration Design Variables*. American Water Works Association Research Foundation, Denver, CO.
2. Wegelin, M. (1988). Roughing gravel filters for suspended solids removal. In: *Slow Sand Filtration: Recent Developments in Water Treatment Technology*. N.J.D. Graham (Ed.). Ellis Horwood Ltd., Chichester, UK, pp. 103–122.
3. Mouchet, P. and Bonnelye, V. (1998). Solving algae problems: French expertise and world-wide applications. *Aqua* **47**: 125–141.
4. Bernhardt, H. (1995). Reservoirs and drinking water supply—a global perspective. *Aqua* **44**: 2–17.
5. Kors, L.J. and Bosch, A.D. (1995). Catchment protection of a multi-functional reservoir. *Aqua* **44**: 80–84.
6. Ketelaars, H. et al. (1995). Occurrence of *Cryptosporidium* oocysts and *Giardia* cysts in the river meuse and removal in the biesbosch reservoirs. *Aqua* **44**: 108–111.
7. van Breemen, L.W.C.A. and Waals, J.M.J. (1998). Storage of surface water in the Netherlands: Challenges of the future. *Water Supply* **16**: 375–381.
8. Oskam, G. (1995). Main principles of water quality improvement in reservoirs. *Aqua* **44**: 23–29.
9. Medema, G., Bahar, M., and Schets, F. (1997). Survival of *Cryptosporidium parvum, Escherichia coli, faecal Enterococci,* and *Clostridium perfringens* in river water: Influence of temperature and autochthonous microorganisms. *Water Sci. Technol.* **35**: 249–252.
10. Gollnitz, W.D., Clancy, J.L., Whitteberry, B.L., and Vogt, J.A. (2003). RBF as a microbial treatment process. *J. Am. Water Works Assoc.* **95**(12): 56–66.
11. Schijven, J.F. and Rietveld, L.C. (1997). *How Do Field Observations Compare with Models of Microbial Removal? Under the Microscope—Examining Microbes in Groundwater*. American Water Works Association Research Foundation, Denver, CO.
12. US EPA. (2003). *Long Term 2 Enhanced Surface Water Treatment Rule: Proposed Rule*. United States Environmental Protection Agency, Washington, DC.
13. Dugan, N.R., Fox, K.R., Owens, J.H., and Miltner, R.J. (2001). Controlling *Cryptosporidium* oocysts using conventional treatment. *J. Am. Water Works Assoc.* **93**(12): 64–76.
14. Gimble, R. and Clasen, J. (1998). International report: Removal of micro-organisms by clarification and filtration processes. *Water Supply* **16**: 203–208.
15. Plummer, J.D., Edzwald, J.K., and Kelley, M.B. (1995). Removing *Cryptosporidium* by dissolved-air flotation. *J. Am. Water Works Assoc.* **87**: 85–95.
16. Harrington, G., Chen, H., Harris, A., Xagoraraki, I., Battigelli, D., and Standridge, J. (2001). *Removal of Emerging Waterborne Pathogens*. American Water Works Association Research Foundation, Denver, CO.
17. Jacangelo, J.G., Adham, S.S., and Laîné, J-M. (1995). Mechanism of *Cryptosporidium, Giardia*, and MS2 virus removal by MF and UF. *J. Am. Water Works Assoc.* **87**(9): 107–121.
18. Matsui, Y. et al. (2003). Virus removal by ceramic membrane microfiltration with coagulation pretreatment. *Water Sci. Technol.* **3**(5–6): 93–99.
19. Adham, S.S., Trussell, R.S., Gagliardo, P.F., and Trussell, R.R. (1998). Rejection of MS-2 virus by RO membranes. *J. Am. Water Works Assoc.* **90**(9): 130–135.
20. Owen, C., Taylor, J.S., Robert, C., and Reiss, C.R. (1999). Microbial Challenge of Integrated Membrane System Large Scale Pilot Plants Treating a Highly Organic Surface Water. *Proc. AWWA Water Qual. Technol. Conf.*, American Water Works Association, Denver, CO.
21. Taylor, J.S. and Wiesner, M. (1999). Membranes. In: *Water Quality & Treatment*. R.D. Letterman (Ed.). McGraw Hill, New York, pp. 11.1–11.71.

WATER METER

Mario O. Buenfil Rodriguez
National University of Mexico &
Mexican Institute of Water
Technology
Cuernavaca, Morelos, Mexico

HISTORY OF WATER METERS

A water meter is a device for measuring and registering the amount of water that passes through a pipe or other outlet, usually for billing purposes. Water meters have various synonyms, used in different cities, countries, or circumstances. Some of these equivalent terms are volume counter, flow-gauging device, water totalizator, flowmeter, and accumulator.

The physical device used to measure the water consumed in a house or industry, or the water extracted from a well or similar supply source, is relatively simple compared to other quotidian instruments. It gauges and adds up the water volumes as they flow in the intake pipe (flow rate is the volume during a certain time span).

In cities when consumers must pay their water bills according to the volumes used, meters must be massively installed, so the logistics to assure their timely reading, quality, precision, repair, replacement, and several other aspects require thorough technical assessment and planning. Many cities of the world depend on them nowadays, although they are relatively modern (no more than 150 years) and are having a rapid evolution. Let us review a little of this evolution.

Humans have always liked to count and measure things ("Man is the measure of all things," Protagoras). Our troglodyte father counted his trophies, wives, possessions, or days passed with his fingers. Later, some other measuring units and quantifying devices were invented, although counting a shapeless and moving liquid was not so easy for some centuries.

Some indirect and ancient ancestors of our modern water meters were the water clocks or "clepsydra" (from Greek "water thieves") used since about 325 B.C., which were stone vessels with sloping sides that allowed water to drip at a nearly constant rate from a small hole near the bottom. Afterward, more sophisticated and artistic water clocks were used (as the huge Su-Sung water clock from China). These in a sense had the inverse purpose of water meters nowadays: first, it was measuring time from a known volume of water; now we want to measure a variable flow of liquid during a known span of time.

Flow, or water volume, must always be referred to a time unit and we must express it as liters per minute, m^3 per month, cubic kilometers per year, etc.

Time is measured with a calendar and/or a wristwatch. Volume can be counted (or estimated by calculus and assumptions) in many ways. One method is using a bucket or container of known dimensions; another procedure is measuring its velocity of travel in a pipe of known cross section (flow = velocity × area).

Most modern water-meters act on the principle of counting the number of turns made by a small reaction turbine moved by the water as it flows through it. This is the basis of Sir William Siemens' invention, who patented his water meter ('fluid meter') in April 1852 (1). Until that time, water metering was not possible; several attempts had been made to devise a suitable device, but all had failed. W. Siemens also invented meters for other applications (bathometer for the depth of the sea, electric pyrometer, etc.).

Ironically, England, where meters were invented, remains until now one of the few countries where almost no houses have water meters installed. This gives an idea of the complexity and expense of installing and reading them, particularly when they are not really indispensable (the United Kingdom has a relatively good quantity of rain and water to waste, but it is not the same in other countries).

Nowadays, water meters, instead of mechanical transmission of movements and signals (as in old clock gearing), send magnetic, electric, or digital pulses which can be stored (remote dial counter or display, data-logger, etc.).

Besides volume, there are multiple water parameters requiring measurement or pondering, for instance, temperature, color, chlorine or salt content, velocity, pressure, and environmental value. Each feature is useful for a given purpose. In leak detection, or billing consumption, and general management of a water utility, water volumes (referred to specific time spans) are key topics.

WATER SUPPLY MANAGEMENT AND CHARGING PRACTICES

Urban utilities use and need water meters as tools for the proper management of a valuable resource, which is water. Numerous water meter types exist that differ in design, material, size, precision, operating principle, reading display, purpose, and site where installed (Fig. 1 shows a brief sample of them). Before entering in any technical detail about how meters work, it is important to review why meters are required. This will give insight into the convenient meter type for a given task.

Water meters can be used either to measure the volumes supplied by the utility through different sources and aqueducts (macrometering), or to know the volumes consumed by clients (micrometering). There usually are differences among the volumes supplied and those consumed. Logically in this comparison, supply should be greater than real consumption, and their difference represents water losses, which must be studied to keep them within reasonable limits.

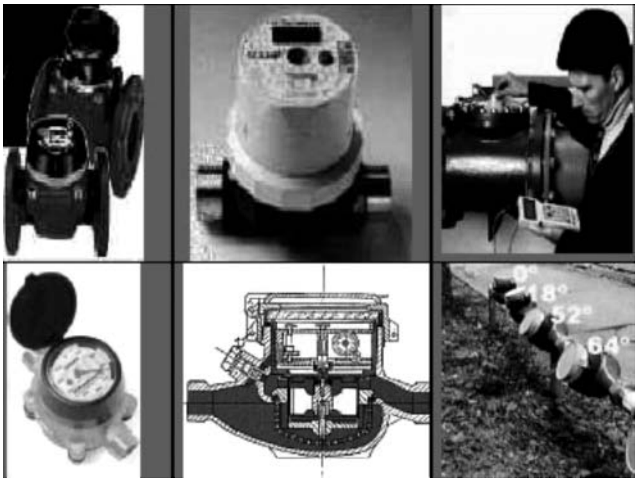

Figure 1. A variety of meters, installation and reading styles.

Metering is an important element to attain *water conservation* (*efficient water use*) in places where water availability is below potential demand. In fast growing cities with limited water resources, household water metering is primarily seen as an approach to control consumption (demand management). It avoids unrestricted use, so water supply and water demand can be balanced. Metering also leads the water undertaker into efficient use of other resources (money, energy, infrastructure, etc.) and good service. In some cases, "nonmetering" options can attain equivalent conservation results, but they must be evaluated carefully before any decision is made.

To be effective, water metering must be linked to appropriate tariffs and prices, together with ways to enforce them, for example, applying surcharges to consumers using more water than the average citizen or cancelling service to people who do not pay their water bills. This is particularly true when water conservation is a prime goal. The role of tariffs is to give a fair chance to everyone to use a reasonable quantity of water and to recover the operating and investment costs of the service. Tariff fixing is quite difficult and subjective when social fairness to consumers, historical privileges, and provisions for the future are involved.

Figure 2 illustrates different rate structure styles in a sample of 827 U.S. water utilities during 1996 (2). The most convenient style for conservation purposes is the "*increasing block rate*," where the price per each cubic meter is higher as total volume consumed grows.

REASONS AND PRIORITIES FOR HOUSEHOLD METERING

Water metering may have different approaches and purposes. Its most frequent bases are either slowing the growth of water demands via volumetric charges to consumers or improving the water utility's revenue through a fair pricing system. There may be other objectives, and any metering project obeys a combination of several of them.

Metering is not an exclusive way to obtain appropriate financial resources; other ways to get payments from consumers can be set up. Nevertheless, for *"consumptive" or "Western" mentalities*, billing water consumed through metering is a clear and fair charging system. The drawbacks are that meter installation, upkeep, and reading demand good organization and represent significant costs, particularly when massive radical changes are required in existing practices. It is relatively cheap when metering is done systematically in all new constructions and regular improvements in equipment and procedures take place.

The reasons and priorities to improve or adopt household metering may vary with place and time. For example in poor arid countries with fast growing urban populations, the priorities, in decreasing importance, may be:

1. Water conservation. Compel people to consume less water through volumetric charges. The water saved may allow extending the system, improving service standards, or protecting the environment.
2. Cost recovery and financial soundness of the water company. Appropriate revenue for all expenses (water supply, sewerage, pluvial drainage, or other services) and provision for future investments.
3. Individual household equity. Avoid discriminatory practices associated with fixed rate tariffs, and charge according to volume used.
4. Unaccounted for water reduction (*clandestine connections and leakages*) through better information on consumption.
5. Peak demand abatement. Cut down nonindispensable uses, or even provide the possibility of seasonal or hourly tariffs.
6. Social equity. Favor poor consumers, who use less water, with lower tariffs or through subsidies.
7. Better data about demand and variations to improve operation and planning of the water system.

Figure 3. Facets invovled in water meter planning.

Deciding the convenience of metering for a previously unmetered place, or upgrading it where it exists partially, requires careful evaluation of a broad spectrum of elements, ranging from local economy and politics to the family budget. It also involves family habits, health, institutional organization, and water service standards. Figure 3 represents the different elements that must be taken in account. Water metering planning requires a clear and careful bond of all technical, administrative, legal, and social aspects involved.

PRECISION TESTS FOR METERS

Various types of flow meters have mobile parts subject to wear, alteration, and clogging with debris or scaling (incrustation), so it is necessary to make periodic tests or replacements to assure their accuracy and proper performance. They could also have manufacturing defects or disarrangement during transport.

New and recently acquired meters usually are randomly tested at the utility's laboratory, on a testing bench, against the accepted standard (for example ISO 4064-1). Later, once installed, at intervals of some years, when

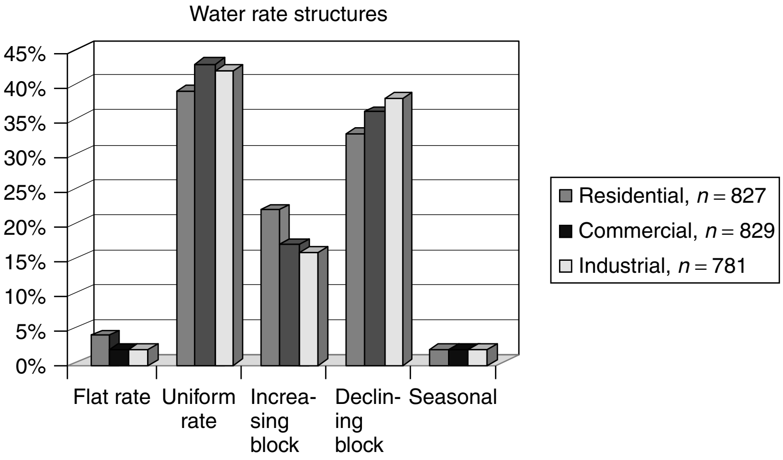

Figure 2. Water rate styles in various USA water utilities (year 1996).

a client complains about excessive charges, or there is a report of underregistration, the apparatus can be dismounted and taken to the lab to test its performance and to be repaired, although it is preferable to perform field tests (*in situ*, *without dismounting the house meter*) with a portable meter test kit. There is commercial equipment that provides everything needed to test domestic water meters accurately, contained in an easy-to-handle case. For more elaborate field assessment as in the case of intermittent service, where the main pipe could be empty, the utility may send a vehicle with all required elements for the test.

Water meters, like most gauging devices, even when new, are not always exact and reliable. Most meters do not perform well at very low flows, and at high flows, they also have errors. They have an appropriate working range where precision is high; out of it, errors increase exponentially. Meters of "*class D*" solve the underregistration presented by other types of meters when used in houses with frequently low inflow rates, for instance, those dwellings having a general storage tank with an admitting float ball-valve. These meters may be about 10% more expensive than "*class C*," but their impact on revenue (water sales) may be worth it. Even though they are more precise, they still cannot detect very low flows, for example, a leak through a dripping tap. Meters of "*class C*" are not precise for flows below 25 L/hour, and "*class D*" below 11 L/hour.

Meters by themselves, as any hydraulic fitting in a pipe, generate additional head losses and are obstacles that reduce the flow rate. It is important to know their potential effect because these losses may be reflected in additional pumping and energy costs for the undertaker in supplying appropriate pressure to some neighborhoods. Obviously meters that produce the lowest head losses are preferable. Figure 4 presents a typical precision graph and a head loss curve.

METER SELECTION AND SIZING

Many issues must be taken into account for the proper selection and sizing of a meter. The previous graph illustrates the need to select an apparatus yielding satisfactory precision without too much head loss. This evidently must be judged and compromised against purchasing and installation costs. Other factors to consider are house characteristics where meters will be installed; especial protection required against weather, vandalism, and tampering; or particular provisions for reading. When selecting a meter, it is necessary to assess the probable consumers demand (average flow, maximum flow), diameter and materials of the existing pipe, impact of water quality (corrosion, presence of sand or other suspended solids, scaling, dissolved air, temperature), effects of soil and environmental

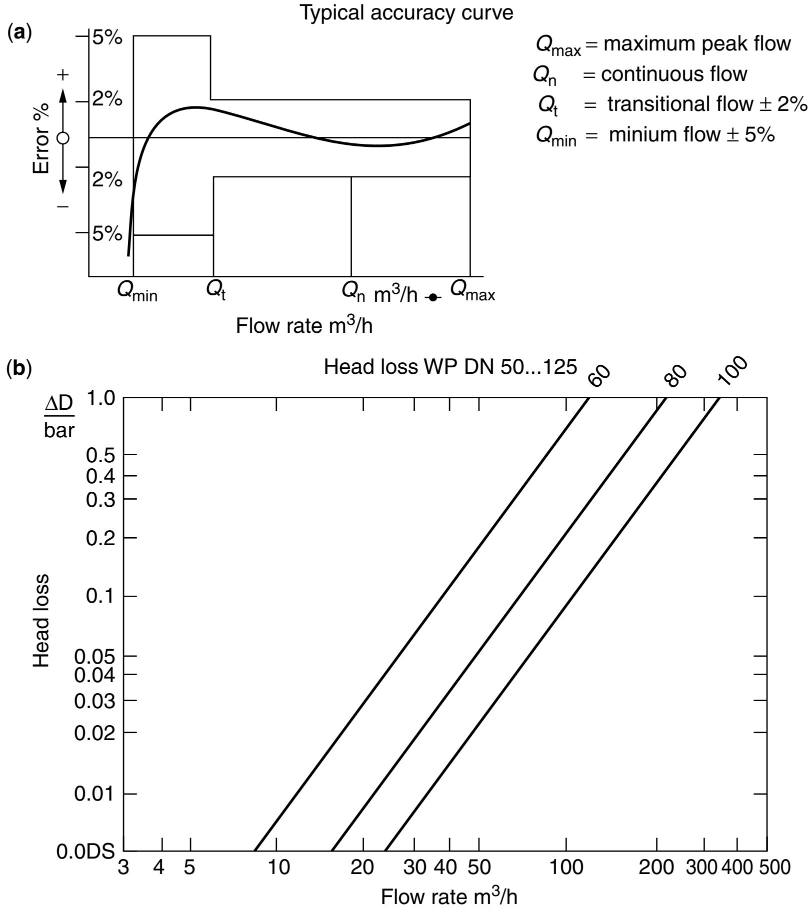

Figure 4. Examples of precision and pressure loss graphs for water meters.

aggression (temperatures, humidity, snow), tariff system and frequency for reading, inspection or replacement routines, risks of transport or from installation and bad practices (wrong position, inverted flow, tampering).

Still other issues affecting decisions are convenient stock for replacements (number of meters, models and their spare parts to be kept in reserve); available budget (acquisition, certification, storage, transport, installation, inspection, tests, visits, contracts and invoices to the consumer), financing sources, and legal and contractual adjustments (bylaws and codes to be modified or completed, inspections and enforcement required, renewal of contracts to customers).

Modern technology allows remote readings via radio, cable, or telephone, when the meter has the appropriate provisions for this, apart from the traditional direct display of its readings.

The installation of a macrometer (as for the study of a DMA, district-metering area) may require attaching some graphing or electronic recording device (data-logger) to it to monitor flow rates during different time spans.

In a consumer line with a rather big diameter, it is possible to save some money, sacrificing a little precision, by installing a “proportional meter.” This is placing a parallel detour (bypass) line with a smaller meter on it. Sometimes is possible to place “batteries of meters,” by means of parallel pipes with different meter diameters on each one.

Multirate meters and water charges through variable tariffs according to the time of day are quite feasible with modern electronics, although may be still expensive. Probably, in the next years, costs will drop and several cities will be using them (power companies in many parts of the world already use meters for electricity with such capability).

It is always important to select the appropriate meter in accord with consumer characteristics and their expected consumption patterns. The utility must have specific guidelines for different types of industries, commerce, service facilities, public offices housing, etc.

Meters can have a provision to restrict flow or pressure. Besides the billing effect of the meter in discouraging waste and high consumption, these adjustable restrictors associated with the counter, may lower peak demand during critical hours of the day. They can also reduce risks of leakage inside the house, or give equal chances to deliver fair and similar amounts of water to consumers in neighborhoods or streets too steep or long (with great differences in potential water pressure).

There can be approaches where nonpermanent installation of meters (random or temporary metering) or communal meters can be useful and economical. In old neighborhoods and in apartment buildings, the supply pipe is common to various dwellings, where installing individual meters, as well as periodic reading, may be extremely expensive and difficult.

TECHNICAL ASPECTS OF METERS

Flow meters work using basic physical and hydraulic principles of mass conservation or continuity, energy conservation (head losses due to flow and velocity energy converted to position energy), momentum, or inertia.

There is no space to describe the operating characteristics of the various measuring systems and instruments available and used nowadays. Figure 5 lists some of them (more details can be found in chapter 4 of Ref. 3).

Meters can be of various kinds, for example, volumetric, propeller, and remote sensing. All have different precision

Type of flow meter	Type of measurement
Venturi tube	Volumetric
Flow mazzle	Volumetric
Orifice plate	Volumetric
Magnetic	Velocity
Nutating disk	Volumetric
Oscillating piston	Volumetric
Bellows gas	Volumetric
Label impeller	Volumetric

Slide sense rotary Retracting vane rotary	Volumetric Volumetric
Helical gear	Volumetric
Turbine	Volumetric
Vortex shedding	Velocity
Vortex procession	Velocity

Type of flow meter	Type of measurement
Fluidic oscillating	Velocity
TOF ultrasonic	Velocity
Doppler ultrasonic Thermo-anemometer	Velocity Velocity (mass)
Calorimetric	Velocity (mass)
Corrolis mass	Mass flow
Relameter	Velocity

Figure 5. Physics operating principles of some meters.

and manufacturing standards, distinct installation and operating costs, and, of course, will render distinct benefits to the water company.

METERING DENSITIES, COSTS, AND IMPACTS

The following figures will give a rough idea of the costs, *per meter installed and working*, for the main concepts of a metering scheme: census, planning, and preliminary work \$2 US; device purchase \$30 US; installation \$15 US; operating costs (meter reading, billing, error correction, tests, etc.) as additional cost compared to fixed tariff system \$0.25 US/year, and meter reposition (in average every 15 years) \$38 US/15 years.

The most important impact of water counters is in reducing consumption, compared to a situation where meters do not exist. Several references report consumption reductions ranging from 5% to 54% when water price increases and there is a volumetric system. This is known as "*price elasticity of demand*". Figure 6 is an example, taken from a report by the World Bank, illustrating the consumption behavior (cubic meters per month per connection) in Brazilian utilities according to the percentage of meter coverage in each city (4). Reductions are on the water supply side and also on the volume of waste production and the required size of sewerage and treatment infrastructure.

"Universal metering" (all houses and industries in the city have a counter) is adequate only for rich cities, where the possibility of uncontrolled demand may occur randomly from any consumer. It is easy to finance when a meter is always installed in every new construction. Its cost can be recovered as part of the connection fee.

"Voluntary metering," "random metering," "selective metering," temporary metering, or neighborhood or hydrometric district control are ways to avoid expensive universal metering. These, if well planned, can lead to better budgetary results for the institution and to nearly the same water conservation results as universal metering.

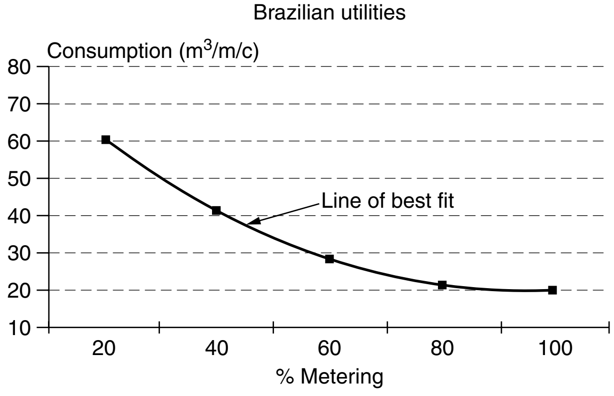

Source: Catalogo Brasileiro de Engenharia Sanitaria e Ambiental. CABES. 1990.

Figure 6. Impact of metering coverage on water consumption.

METER READING AND INFORMATION PROCESSING

Medium to big cities have thousands of meters to be read and accounts to be monthly billed and followed. This represents a basic regular task, requiring managerial ability from the water utility. Its success depends on good quality and reliability of the information handled and perfect synchronization among different departments and people.

Here, any gain to save time or money during reading and billing is desirable, and this is why new techniques are being tested and adopted continuously. Some utilities use contact rods to enter digital readings directly into a computer, instead of traditional visual reading and writing. Others use, reading through radio reception on a van circulating through the city streets. Reading through modems (telephone) is practiced as well.

In cities where traditional meters still exist, one important engineering task is optimizing the reading routes to save distance and time. These routes should be frequently reviewed and enhanced according to new customer contracts and technical improvements in metering and reading.

BIBLIOGRAPHY

1. Columbia Encyclopedia, 6th Edn. article on Siemens, Sir William, Infoplease learning network http://www.infoplease.com/ce6/people/A0845159.html.
2. AWWA. (1996). *Water Statistics*. American Water Works Association. http://www.awwa.org/h20stats/h20stats.htm http://www.awwa.org/h20stats/ratebar.htm.
3. EPA. (1998). *Compliance Assurance Monitoring (CAM). Technical Reference for Monitoring Equipment and Instruments. 4.4 Flow Rate Measurement Systems*. Environmental Protection Agency. http://www.epa.gov/ttn/emc/cam/sec4-5.pdf.
4. World Bank (Yepes, G. and Dianderas, A.). (1996). *Water and Waste Water Utilities Indicators*, 2nd Edn. Water and Sanitation Division, TWUWS, The World Bank. http://www.worldbank.org/html/fpd/water/pdf/indicators.pdf.

READING LIST

Aguas de Sevilla, Reglamento del suministro domiciliario, an example of specifications for water meter installation http://www.aguasdesevilla.com/Normativa/reglamento03.htm#art35.

AWWA. (1975). *Sizing Water Service Lines and Meters*. American Water Works Association.

Buenfil, R.M. (1992). *Water Metering Planning and Practices*. M.Sc. Thesis, Loughborough University.

Buenfil, R.M. (1993). *Needs and Illusions About Household Water Metering*. Management of Water Resources in North America, Arizona University & ASCE.

Buenfil, R.M. (1993). *Household Water Metering and Tariffs*. **Conserv-93**, AWWA-ASCE-AWRA Congr.

Buenfil, R.M. (2000). *Estrategia para la toma de lecturas en medidores*. CNA, IMTA.

Buenfil, R.M. (2002). *Clase 3, Medición, curso de uso eficiente del agua en ciudades*. DEPFI, UNAM (Universidad Nacional Autónoma de México). http://www.geocities.com/mario_buenfil/depfi/clase_3.htm.

Encyclopedia Britannica. article about Sir William Siemens, inventor of the water meter. http://34.1911encyclopedia.org/S/SI/SIEMENS_SIR_WILLIAM.htm.

Edwards, K. The fluid flow calculations website. LMNO Engineering, Research, and Software, Ltd. http://www.lmnoeng.com/index.shtml.

Hansen, R.P. (2001). *Elementos Secundarios de Medición de gasto Serie autodidáctica de medición*, IMTA.

ISO. (1993). *Measurement of water flow in closed conduits—Meters for cold potable water—Part 1: Specifications*, ISO 4064-1. International Standards Organization, http://www.dfmg.com.tw/dasp/standard/iso/d9772.html.

Lienhard, J. *Engines of our ingenuity. Episodes: 260 units of measure, 1363 man the measure, 1580 water clock*. University of Houston, College of Engineering http://search.uh.edu/engines_search.html http://www.uh.edu/engines/epi260.htm.

Mello, E.J. (2001). Conservação e Medição da Água various links and articles related to metering and water conservation. http://www.geocities.com/hidrometro/.

Nielse, M.J. Seleção e Dimensionamento de Medidores de Água. (principais critérios e fatores), http://www.geocities.com/hidrometro/dim_resum.htm.

NIST, International System of Units (SI), National Institute of Standards and Technology (USA), http://physics.nist.gov/cuu/Units/units.html.

Osaka Water Service. (2000). *How to Read Your Meter. How to Calculate Your Water Bill*. Water Service Handbook Osaka Water Service. http://www1k.mesh.ne.jp/osaka-water/english/benri/benri_5.html.

USBR. (1997). *Water Measurement Manual. Chapter 14—Measurements in Pressure Conduits*. United States Bureau of Reclamation. 3rd Edn. http://www.usbr.gov/wrrl/fmt/wmm/http://www.usbr.gov/wrrl/fmt/wmm/chap14_01.html.

MUNICIPAL WATERSHEDS

Timothy J. Downs
Clark University
Worcester, Massachusetts

INTRODUCTION

The Sustainable Development Summit in Rio de Janeiro in 1992 showed clearly that building and running more sustainable human settlements is one of the core challenges (1). Developed countries are characterized by having the majority of their population in urban settings, concentrations of relatively large numbers of people in a relatively small physical area. Transitional countries show a demographic shift from rural to urban settlement. In 2001, more than 54% of the world's 6 billion people lived in municipal centers (towns and cities); some countries have more than 90% (2). Developing and transitional countries often have megacities—Mexico City, Cairo, Sao Paulo, Jakarta, Bangkok—that completely dominate the watersheds they occupy. This form of municipal watershed lies at the least sustainable end of the scale because satisfying voracious demand is a priority, and water-producing zones of the watershed become either consumed by urban sprawl or cannot influence water policy.

In a humid environment, as a society moves from a preindustrial economy to an industrial economy, its watersheds experience three main stages of development (3): (1) few encroachments on water resources; more than enough water for everyone; water is free; (2) transitional infrastructure, including dams for energy and irrigation and interbasin transfers between water-rich and water-scarce watersheds; and (3) maximum streamflow regulation achieved, and aquifers are at maximum sustainable yield, so further development is both expensive and threatens to be unsustainable. Unfortunately, in the case of several megacity watersheds, a fourth stage of overexploitation by intensifying exploitive infrastructure and interbasin transfers is testing the functional viability of municipal watersheds.

TYPES OF MUNICIPAL WATERSHEDS

Megacities lie at one end of the municipal watershed scale (e.g., Mexico City); watersheds that have relatively equal-sized, similar economies and evenly distributed small towns form the other (e.g., Nashua River Watershed, Massachusetts, USA). In between, are watersheds that have a few large towns, a city or two, and several small towns; they have differing economies, needs, and priorities (Fig. 1).

As we move along the scale, the probability of achieving more integrated, sustainable watershed management declines primarily because of the rising conflicts of interest, increasingly significant cultural differences between municipalities, and growing pressure to overexploit natural resources that degrade the watershed ecosystem. In its most dramatic form, a megacity will also draw water resources from neighboring or external watersheds (e.g., Mexico City, Los Angeles), thereby becoming a regional 'sink' that has very significant ecological, sociopolitical, economic, and cultural impacts.

IMPACTS OF MUNICIPAL CENTERS ON WATERSHEDS

Water Quantity Impacts

Figure 2 shows a schematic of a typical municipal water cycle, which may also be considered to represent a composite of the water supply and wastewater sanitation infrastructure for all towns and cities of a municipal watershed. Humans perturb the stocks and flows of water in the watershed through withdrawals for water supply (4):

1. domestic (in-house and out-of-house);
2. industrial (e.g., factories, power stations), commercial (e.g., shops and hotels), and institutional (e.g., hospitals, schools);
3. agricultural (e.g., farms, crops, livestock);
4. public—for example, parks, fire fighting, sewer flushing;
5. Losses—distribution losses, consumer wastage, metering errors, and/or unrecorded consumption.

Domestic, industrial, and public demand tend to dominate municipal watersheds; agricultural demand dominates rural watersheds.

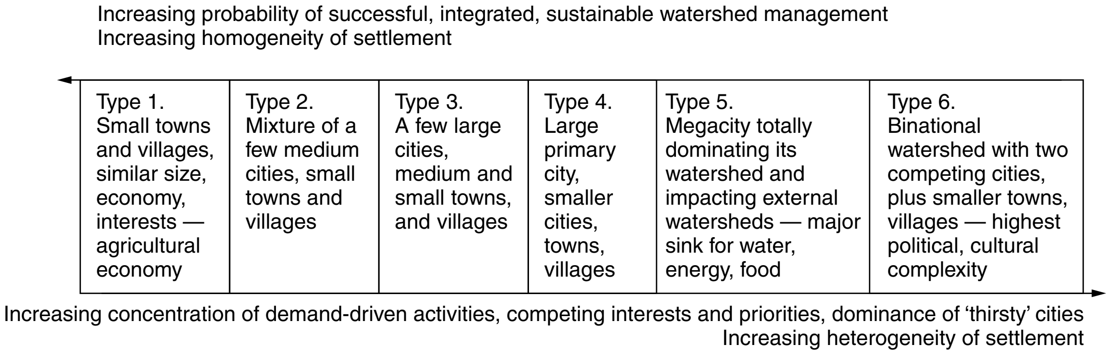

Type 1.	Type 2.	Type 3.	Type 4.	Type 5.	Type 6.
Small towns and villages, similar size, economy, interests — agricultural economy	Mixture of a few medium cities, small towns and villages	A few large cities, medium and small towns, and villages	Large primary city, smaller cities, towns, villages	Megacity totally dominating its watershed and impacting external watersheds — major sink for water, energy, food	Binational watershed with two competing cities, plus smaller towns, villages — highest political, cultural complexity

Figure 1. A simple typology of municipal watersheds.

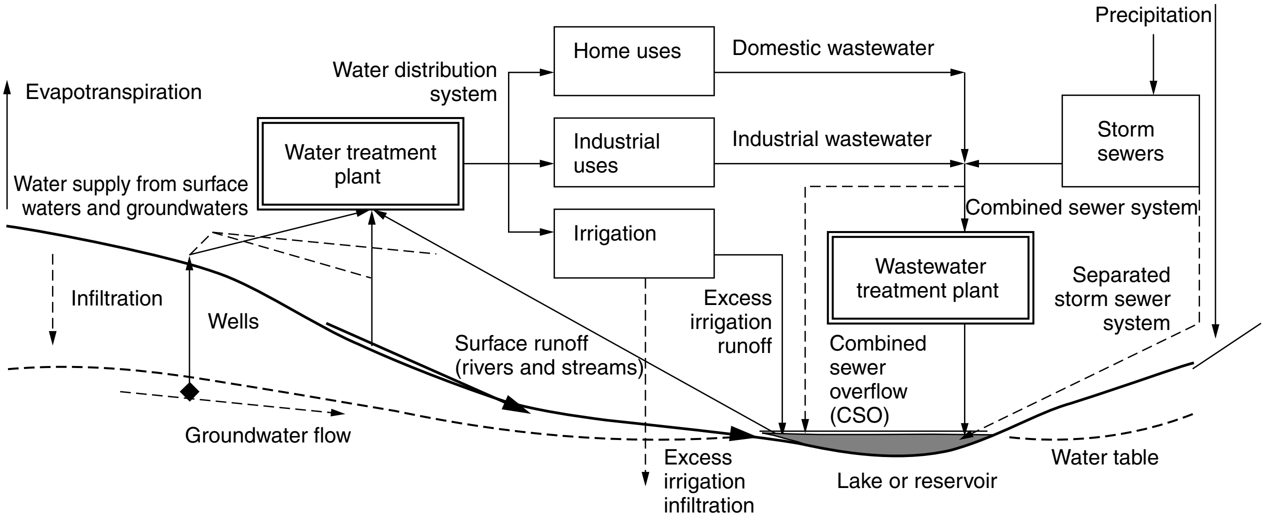

Figure 2. The water supply and wastewater sanitation cycle of a municipal center. The cycle also serves as a composite for the infrastructure serving a municipal watershed.

The estimation of water demand involves (4)

1. plotting the watershed population trend for the past 10–20 years and estimating the proportion likely to be due to immigration and that due to natural increase (births minus deaths);
2. dividing the supply area into different socioeconomic classes of domestic use and estimating the typical domestic consumption per capita in each class;
3. seeking values of future immigration and natural increase for the different classes of housing;
4. estimating distribution pipe loss rates, consumer wastage, and unsatisfied demand;
5. estimating the growth in industrial, commercial, and service demands (a function of population growth);
6. estimating growth in agricultural demand, hydroelectric demand, and public and ecosystem maintenance demand.

Water Quality Impacts

If the water supply–wastewater sanitation cycle is incomplete (usually the case in developing and transitional countries) and wastewater is untreated, it returns to the environment and pollutes receiving waters. The pollution by municipal wastewater also contaminates sediment and soil, air and biota; the distribution of contaminants is a function of their physicochemical properties and the biophysicochemical properties and processes of the environment. Even in a 'closed cycle' like (Fig. 2), municipal wastewater comprises stormwater runoff, combined sewer overflow (CSOs), and effluent from wastewater treatment plants (WWTPs), as well as industrial point source discharges. The location of the pollution sources, their magnitude, composition, frequency, and duration, plus the surface and groundwater hydrology of the receiving areas, will determine the impact. Specifically, it determines whether or not the assimilative capacity has been exceeded, causing pollutant accumulation. The temporal and spatial scales of impacts vary (after 2): (1) local scale for less than a day—acute toxic impacts, sediment resuspension, coarse material sedimentation; (2) subwatershed scale for days to weeks—BOD and DO changes, algal blooms; (3) watershed scale for months to decades—accumulation of pollutants in fish and sediments, cultural eutrophication, composition of aquatic species changes. As well as pollution potential, large municipal centers may affect microclimate by altering the energy regime, air quality, and circulation (2). Urban heat islands increase evaporation, whereas impervious surfaces reduce infiltration and increase runoff.

In developed countries, because of stricter laws, urban wastewater impacts mainly the recreational use of lakes and rivers, may contaminate shellfish, but rarely pollutes water supplies. Pathogens enter surface receiving waters from untreated CSOs, WWTP effluent, and stormwater runoff, posing risks to bathers, and leaking septic tanks can contaminate surface water and groundwater. In stark contrast, in developing countries, a polluted water supply, poor sanitation, and poor drainage cause a range of devastating impacts on public health (5,6): Vector-borne diseases like malaria and dengue fever (the former has an estimated annual global mortality of 1–2 million, morbidity 200 million, and more than 2 billion at risk); water-based diseases such as schistosomiasis (bilharziasis) and dracunculiasis (guinea worm); water-washed disease such as trachoma and scabia; and waterborne diseases such as cholera, bacillary dysentery, amebic dysentery, and typhoid (estimated annual mortality 4 million, morbidity more than 1.5 billion, 2 billion at risk).

In a municipal watershed, unregulated human activities produce a combination of ecological stressors: habitat destruction, flow modification, thermal pollution, invasive species, and toxic pollution from wastewater and solid waste. When these stressors act together, the possibility exists that degradation becomes synergistic, where one stressor magnifies the impacts of others. Clearly, from (Fig. 1), the municipal watershed ecosystems (including human populations) most at risk from such large-scale impacts are types 4, 5, and 6 because of the potential number, frequency, duration, and magnitude of stressors associated with large cities and megacities.

INTEGRATED MUNICIPAL WATERSHED MANAGEMENT (IMWM)—NEEDS

Municipal watershed management becomes *integrated* when actions on the local municipal scale are carried out in the context of connections with other settlements and the wider watershed. The integration can happen across dimensions described in terms of different kinds of capital. Scoones (7) identifies five categories of capital assets that sustain human livelihoods:

1. natural capital—land, water, wildlife, biodiversity—integration across types of natural resources because all are interrelated in the watershed ecosystem;
2. social capital—groups, networks, laws, regulations, institutions of people—integration across social groups (multistakeholder planning and management), and networking between communities, between government agencies, and between water users and government regulators;
3. human capital—skills, knowledge, healthy people—the integration of different disciplines, professions, interests, and abilities;
4. physical capital—basic infrastructure that supports shelter, transportation, energy, water supply and sanitation, health care, education, communications—the most obvious form of watershed integration, connecting water sources, supply and sanitation infrastructure serving different users;
5. financial capital—income, savings, credit—pooling of financial assets for investment in watershed management goals, including service charges, pollution taxes, and appropriate subsidies.

The last two are the most familiar; the first three require changes in thought, feeling, and action and are beginning to take place. It is notable that in the relatively few cases of 'successful' IMWM, integration has occurred, at some level or other, across each capital dimension. The goal of integrated water management has been described as: "the sustainable, coordinated management of water resources within a region, with the objectives of controlling and conserving water, minimizing adverse effects and achieving specified and agreed water management and social objectives" (8). The cycle of (Fig. 2) is made more sustainable as follows:

- optimizing the design of the engineered cycle to exploit sources responsibly and more sustainably and to minimize wastage, pipe leaks, unaccounted for losses, cross-contamination of supply and sanitation lines, and pollution of the receiving environment;
- reuse of wastewater for nondrinking uses (and drinking water if state-of-the-art tertiary treatment is economical, for example, in Windhoek, Namibia);
- water conservation measures by substituting more water efficient technologies (less water use), cleaner technologies (less wastewater production), and demand management through metering, pricing, and regulation.

Water Supply Enhancement

The World Health Organization (WHO) and the World Bank's basic water requirement for meeting survival needs is 20–40 liters/person/day (2). Many municipalities, especially in type 4, 5, 6 watersheds (Fig. 1), continue to develop new water supply projects and intra- and interbasin water transfer schemes to try to meet growing demand. New approaches emphasize conservation practices to regulate withdrawals (e.g., at or less than replacement levels for aquifers), enhance aquifer recharge, and exploit the economies of scale of conjunctive use of several water sources (2). Dual water supply systems—one for potable water, one for other uses—may be economical in areas of water scarcity, though retrofitting old water cycles is costly. Many new devices save water, like low-volume flushing toilets, shower heads, sprinklers, and washing machines. Considerable saving may be had from recycling rainwater or graywater (less polluted or partially treated wastewater). In a typical municipality, 20% of water use may be industrial, so improving process efficiency and reuse is an attractive option, especially for cooling waters that are polluted thermally. For dirty wastewaters, pretreatment options include ultrafiltration, activated carbon filtration, or reverse osmosis (2). Design of the treatment process depends on the effluent water quality, the required reuse quality, and the flow volume. Agricultural uses, though less important in a municipal watershed, may still account for 25%. Most water is used for crop irrigation and for cleaning dairy sheds (manure flushing), so reuse of graywater and rainwater is attractive. New hydroponic (soil-less) agriculture treats and recycles the water medium, using simple filtration (sand filters or biofilters) and ultraviolet disinfection (2). Xeriscaping—the planting of low water use plants for landscaping —is gaining popularity.

Wastewater Sanitation Enhancement

Development practitioners consider that 'environmental sanitation' encompasses wastewater, stormwater, and solid waste collection, handling, treatment and disposal, as well as personal, domestic, and communal hygiene practice. WHO considers that 'improved' sanitation consists of connection to sewers or septic tanks, pour-flush latrines, pit latrines and ventilated improved pit latrines; 'not-improved' devices include manually serviced bucket latrines and open latrines. The adoption of sanitation practices is heavily influenced by hygiene education and cultural norms. Black industrial wastewaters can undergo cheap pretreatment using unit processes like simple bioreactors (e.g., upflow anaerobic sludge bed technology) (2). The primary goal of wastewater treatment is to reduce total suspended solids (TSS), biochemical oxygen demand (BOD), chemical oxygen demand (COD), ammonia, and pathogens. Separation of wastewater streams at the source allows treating graywaters by simple primary filtration and settling processes (e.g., wetlands, sand filters) and then reuse; blackwaters are treated by primary (screening and settling) and secondary (aeration, bioreactors) unit processes (9). Urine may be collected, stored, and used as fertilizer; feces can be stored in a composting tank for later use as compost (10). In some places, biogas from biowaste decomposition is used for energy. Any agricultural reuse of wastewater and its sludge residue (biosolids) after treatment is carefully regulated so that food crops are not contaminated. Sludge processing can be costly, but the material can be used as a fertilizer.

Flood Protection and Drainage

By connecting land use/land cover planning with water flow management, the twin goals of flood defense and

Table 1. Key Aspects of Integrated Municipal Watershed Planning and Management

Information for Problem Definition[a]	Typical Priority Problems	Implementation Steps[b]
• Characterization of water resource capacities • Characterization of water quality • Inventory of aquatic life in the surface waters • Estimation of impacts of polluted water on humans and other species • Inventory of pollution sources (stormwater, CSOs, municipal, agricultural, and industrial point and nonpoint sources) • Inventory of existing water supply and sanitation infrastructure (treatment plant types, capacities) • Inventory of industrial and commercial wastewater discharges to municipal sewers • Land use/land cover inventory • Inventory of sensitive habitats • Soil contamination maps • Flood plain maps for various return periods • Groundwater hydrology and sustainable yield • Population data on numbers, growth, consumption	• Degraded habitats • Degraded water quality • Fisheries management • Supply deficits • Sanitation deficits • Leaky supply systems • User wastage	• Identify who the stakeholders are and their interests • Establish goals that are attainable, endorsed by stakeholders, and economically responsible • Develop a strategic plan on the appropriate (problem-determined) scale (local, subwatershed, watershed) • Develop the funding mechanism • Lobby politicians to gain support • Involve politicians in promoting the plan • Revise institutional roles and responsibilities to support the plan • Implement the plan • Monitor performance (postaudit)

[a] After (11).

[b] After (2).

habitat maintenance can be addressed. The former relies on educating people not to build in the floodplain, nonstructural measures such as laws and regulations, and structural measures such as reservoirs, levees, and revegetation. The considerable impervious surfaces in a municipal watershed reduce groundwater infiltration and increase rates of polluted runoff during precipitation. Stormwater management consists of three types (2):

1. policies and source controls—education, laws, regulations, including land-use zoning;
2. best management practices (BMPs) and community-level BMPs for the site—for example, rooftop detentions, ponding to increase infiltration, reducing slopes;
3. watershed-wide efforts—habitat conservation (especially wetlands), reducing soil erosion.

It has been shown that the total impervious area (TIA) of a watershed—a good development indicator—correlates with the biological integrity (BI—abundance and diversity of aquatic species) of streams (2): high BI—low TIA (<5–10%); moderate BI—moderate TIA (5–10% < TIA <35–45%); and low BI—high TIA (>35–45%). Stream restoration employs BMPs to provide sufficient, not excessive flows, and to mitigate pollution. In many cities of the developing world, periurban areas contribute highly polluted stormwater because of open sewers, open garbage dumps, and soil erosion from degraded slopes.

IMWM—Planning and Implementation

Not surprisingly, planning on a watershed scale requires a systems approach: (1) problem definition (needs and priorities); (2) description of alternative solutions; (3) evaluation of alternatives and selection of preferred option; (4) implementation of the preferred option; and (5) maintenance and monitoring of performance. Table 1 provides an overview of some key operational aspects.

CASE STUDY

The Nashua River Watershed (Fig. 3) and the Blackstone River Watershed are neighboring examples of municipal watersheds in New England, United States. Their management approaches illustrate many of the challenges for IMWM, and useful lessons can be drawn. Both rivers have similar industrial histories that caused severe pollution, equal areas (540 square miles), and both set out to clean up their rivers at the same time in 1969. Later, they set similar goals of meeting the 1972

Figure 3. Nashua River Watershed, a type 2 municipal watershed. Its integrated management has benefited from strong community leadership and a focus on water quality (12) (permission of Nashua River Watershed Association).

(and 1977 amendments) Clean Water Act standards of being ‘fishable and swimmable’ (Class B). Neither river has yet met this goal, originally set by the U.S. Environmental Protection Agency for *all* rivers in the country by 1983. But noticeable improvements in water quality have taken considerably longer in the Blackstone River than in the Nashua River. Why? Mailloux (13) found that Nashua’s watershed management has been more successful compared to Blackstone’s because of the strong community leadership that lobbied politicians and secured funding for paid positions in the Nashua River Watershed Association (NRWA). This community-based organization formed a cohesive, focused, highly motivated environmental movement devoted explicitly to improving water quality and able to take advantage of information resources and legislation. NRWA also formed strategic partnerships with industrial interests and government, significantly strengthening the social capital of their watershed. This process was aided by the relative homogeneity of the watershed settlements and constituent economies (type 2, (Fig. 1)). Blackstone did not enjoy the same homogeneity (type 3), complicated by bistate administration (Massachusetts and Rhode Island), and did not strengthen its social capital cohesively. Nor did it focus on the core issue of stemming water pollution; it was concerned primarily with recreational use and aesthetics.

CLOSING REMARKS

Municipal watersheds are complex socioecological systems that have many conflicting needs, interests, and priorities for the sustainable development of natural watershed resources: water, air, soil, energy, food, forests, and biodiversity. Without a watershed approach to stewardship, human needs sharply conflict with the needs of other species. With the rise of urbanization worldwide, municipal watersheds of different types are becoming geopolitically dominant in the landscape, placing strong demands on their water supply and wastewater sanitation systems and on their rural agricultural counterparts to produce food for them and often transfer water to them. In most developing countries, the problems are acute; few municipal watersheds have adequate systems; many have almost no sanitation. Unless major changes in policies and practices occur—those emphasizing integrated approaches, supply/sanitation optimization, cleaner production and efficiency—ever more settlements and habitats will be at risk from water stress, either from insufficient water quantity, degraded water quality, or both. Municipal watershed stewardship challenges our ingenuity and our political will, our core values, and the capacity of our institutions to respond. It is important, as always, to learn what works, and why, and draw on experience worldwide. Although the magnitudes of watershed problems differ from country to country, the underlying sociocultural obstacles to sustainable solutions are remarkably similar.

BIBLIOGRAPHY

1. UNCED (United Nations Conference on Environment and Development). (1992). *Earth Summit 1992—the United Nations Conference on Environment and Development*. The Regency Press, London.
2. Marsalek, J., Rochfort, Q., and Savic, D. (2001). Urban water as part of integrated catchment management. In: *Frontiers in Urban Water Management*. C. Maksimovic and J.A. Tejada-Guibert (Eds.). UNESCO and IWA, London, pp. 38–83.
3. Falkenmark, M. and Lindh, G. (1993). Water and economic development. In: *Water in Crisis, a Guide to the World’s Fresh Water Resources*. P. Gleik (Ed.). Oxford University Press, New York, p. 80.
4. Twort, A.C., Ratnayaka, D.D., and Brandt, M.J. (2000). *Water Supply*, 5th Edn. Arnold and IWA, London, pp. 1–35.
5. Masters, G. (1998). *Introduction to Environmental Engineering and Science*, 2nd Edn. Prentice-Hall, Upper Saddle River, NJ, p. 174.
6. UNEP (United Nations Environment Programme). (1993). *Environmental Data Report 1993–94*. Blackwell, Oxford.
7. Scoones, I. (1998). *Sustainable Rural Livelihoods: A Framework for Analysis*. Working Paper 72, Institute of Development Studies, Brighton, UK.
8. ASCE (American Society of Civil Engineers, Water resources planning and management division) and UNESCO (International hydrological programme IV). (1998). *Sustainability Criteria for Water Resource Systems*. Task committee on sustainability criteria, Project M-4.3, ASCE, Reston, VA.
9. Skjelhaugen, O.J. (1999). Closed system for local reuse of blackwater and food waste integrated with agriculture. *Water Sci. Technol.* **39**(5): 161–168.
10. Otterpohl, R. (2001). Options for Alternative Types of Sewerage and Treatment Systems Directed to Improve the overall performance. *Proc Interurba II, 2nd Int Conf on Interactions Between Sewers, Treatment Plants and Receiving Waters in Urban Areas*. Lisbon, Portugal, Feb. 19–22.
11. Geiger, W.F. and Hofius, K. (1995). Integrated water management in urban and surrounding areas—findings of the International Workshops in Essen 1992 and Gelsenkirchen 1994 by the German-Dutch IHP Committee to UNESCO, Project M3-3a. In: *Preprints of the Int. Symp. Integrated Water Manage in Urban Areas*. J. Niemczynowicz and K. Krahner (Eds.). Lund, Sweden, Sept. 26–30, pp. 127–151.
12. NRWA (Nashua River Watershed Association). (1995). *1995–2020 Vision for the Nashua River Watershed*. a guide for the next quarter century, NRWA Inc., Groton, MA.
13. Mailloux, J. (2003). *A Tale of Two Rivers, Environmental Science and Policy Honors Thesis*. Clark University, May.

PUBLIC WATER SUPPLY WORLD

Mario O. Buenfil-Rodriguez
National University of Mexico &
Mexican Institute of Water
Technology
Cuernavaca, Morelos, Mexico

VALUE OF A WATER SUPPLY

A good, dependable water supply promotes health, productivity, beauty, arts, safety, recreation, and most of the desirable goals of life and society. Probably, the first ancient human settlements simply occurred in places granted by nature with fruits, animals, and other food, and with enough and easily reachable water. People used to move or travel when conditions changed (*drought, flooding, climate change*) or simply because they wanted to explore new horizons. Later, the discovery of agriculture made sedentary groups larger, and towns grew. Water had a prime role in the possibility of producing food and having a stable place to live. Production and living depended on reliable, good quality water, requiring as little effort as possible in its conveyance, storage, and use. The best water sources were, and still are, regular rains, the stable flow of a river, or the constant level of a lake.

Growing Complexity

Meeting one or more of the previous conditions for accessible water is not always possible or easy. Evidently, when towns grow, new residents need to set up a bit farther away from the river or community storage than already existing houses. This creates, for the newer inhabitants, more toil to fetch the water, the need to have a donkey or a cart, expending time and effort carrying, or investing in pipes to convey the water to their place, and possibly to pay continuously for energy (*pumps*) to raise water to higher spots.

Moreover, the best sites for cities and water were occupied since ancient times, so new settlements, forced by the overwhelming human growth of the last two centuries, are occurring in places without the appeal and qualities of those first towns and cities. Unfit locations are often aggravated by deteriorated water quality due to wastewater and debris discharged by upstream housing, industries, and agriculture.

HEALTH, FAIRNESS, AND ECONOMIES OF SCALE

In cities of high housing densities, ***public water supply*** started as a measure to improve hygiene and health, as well as a need for fairness among citizens, to give them an equal chance to have water and to pay similar amounts via taxes or equivalent. It is also much cheaper and less chaotic to have collective solutions than individual ones.

A public water service can concentrate and pay specialized personnel, devoted full time to these highly technical matters of searching, treating, transporting, raising, distributing, maintaining, and charging for water service.

Water service in most cities of the world is provided by governmental institutions (*either national. regional. municipal*), but in many other locations is operated by private utilities. There also exist mixtures of private and governmental (*public*) responsibilities. Nowadays, there is a tendency, often promoted by lending institutions, to give up (*"privatize"*) governmental services that have too many conflicts and poor quality.

In this article, the term "***public water supply***" refers to the water service provided by some official concessionaire (*independently if is private or government*) that attends to the needs of various clients in a city. For its character and to benefit from economies of scale, it is regularly a "monopoly" within the same city or region. This means that usually no competitors, doing similar tasks, are allowed within a location.

SUPPLY DIFFERENCES AND STYLES IN THE WORLD

Developed countries generally have better natural water resources than poor "developing" (*alias "third world"*) nations. Certainly the ease and abundance of water resources explains many of the differences in material progress among nations. Besides, public water supply quality in "first-world" countries is better, reflected in the continuity, quantity, quality, pressure of water, and the readiness to attend to consumer complaints or requests.

Such reliable and quality service in rich countries doesn't mean that their service is superior or more "efficient" than that offered in poorer and drier countries. Often it is just the opposite because cities that lack water use more ingenuity, toil, and productivity to do much more with fewer resources (*smaller in water availability, energy usage, investments*), covering larger populations, and have radically different, or even nil, water fees.

Service in rich cities typically leaves little care and worry to the individual household, except to pay its taxes or water fees. But as governments are poorer, institutions are weaker, or water scarcity is severer, situations change. There can be various combinations and styles among tasks and investments left to the individual homeowner and among duties performed by the public water utility or through other intermediate providers. There usually are more responsibilities and "choices" posed on the tenant as situations are more stringent. Table 1 presents some examples of a water supply system's components (*infrastructure, equipment*) often left as obligations (*duties*) of each homeowner or group of them.

Sometimes the government or the public privatized service can be in charge, or help, in some individual solutions such as cistern cleaning, rainwater harvesting, toilet retrofitting, and leak repair inside houses, to save water and be able to meet the water demand.

COMPONENTS AND EVOLUTION OF A WATER SYSTEM

All humans need or desire water to cover the following purposes in decreasing priority level or intensity: drinking,

Figure 1. Communal rainwater sotrage.

Table 1. Components of a Water Supply System

Source
Rain catchment and storage at community level (Fig. 1)
Rain catchment at domestic level (*rain harvesting*) (Figs. 2 and 3)
Hand pump well
Wells (*sakia, noria, shallow wells*)
Distribution and pressure
Individual transport (Fig. 4)
Open channel or aqueduct (Fig. 5)
Public standpost
Home connections (*built by the homeowner, built by the public facility, built by a concessionaire*), with valves or meter or without them (Fig. 6)
Electric or manual pumps to raise water to places where needed (*2nd floor, etc.*)
tankers (*buy complementary water from private distributors*). (Fig. 7)
Storage
Public regulation and storage tank
On the roof container (*tinaco*)
In-ground individual cisterns
In-floor collective container (*for a group of houses*)
Water quality
Chlorination
Filters
Boiling (*gas or wood consumption*)
Buying bottled water for drinking
Water saving
Meters and fees
Low consumption toilets, water saving showerheads, etc.
Rota cuts, intermittent service (*regular rationing*)
Very low water pressure (*just enough to reach the water intake, the house connection to the main at street level*)
Laws and regulations to install water saving devices
Reuse of gray water (*sullage*)

bathing, clothes washing, cooking, gardening, pet care, housekeeping, recreation, manufacturing a product, etc. Priorities and quantities vary with cultural, natural resources, and technological differences among regions.

Around the world, cultural preferences and technology are varying rapidly in these decades, and sometimes their fast evolution is confusing many people in thinking that natural resources also evolve at our will (*a mistake which is creating serious risks and damages to other species*). Additionally, there is a strong tendency toward similar (*globalization*) aspirations for comfort and life standards in most communities, although these are not always possible to meet due to economic constraints. The world would be much better if the environmental aspects of each region are considered carefully before embarking on seeking such water system standardization.

Not very far off in time, even nowadays in rich countries, nobody thought about (*"needed"*) water faucets or showers inside their dwellings. The components of a water supply system could then consist of

Figure 2. Rural domestic rain collection.

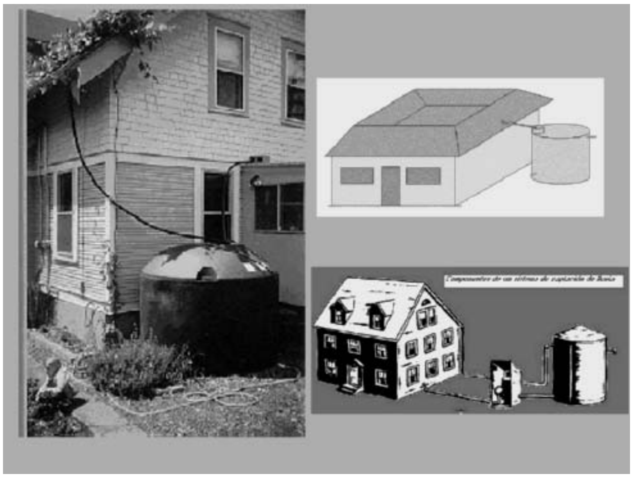

Figure 3. Urban domestic rain catchment.

Figure 4. Water transport.

Figure 5. Ancient acueduct.

Figure 6. House connections in dry tank.

Figure 7. Tanker supplying water.

- source (*storage for rain water, river intake, or shallow well*)
- gravity conduction or simple transport (*channel, simple aqueduct, buckets, animal cart, or pipes*)
- delivery at limited sites (*public fountains, stand posts*)

When water quality is not totally satisfactory or the quantity available from the source is insufficient during some seasons of the year or during some period of the day, a more "advanced" (*complex*) system is required (see Fig. 8) that has additional components:

- dam or long-term storage
- filter or chlorination
- short-term storage or regulatory tanks

Many public water supply systems in the world nowadays operate on a simple scheme, similar to those just mentioned. But as "progress" has advanced, pollution has increased, and people's expectations and habits have been transformed; the complexity of the systems has also enlarged. In such a case, some additional components can be (Fig. 9):

- regional or international agreements to use and share water
- deep wells
- long conduits
- pumping stations,
- metering and gauging stations
- leak detection equipment
- valves of different kinds
- dual distribution systems (*different water qualities on each*)
- house connections, consumption meters
- reading and billing mechanisms.

Much damage was done earlier by considering the water supply network isolated from the drainage or sewerage solution. Both must be understood as a unique, integrated, system. In this case, additional elements are

- sewerage network
- wastewater treatment plants
- reuse networks
- efficient water use and conservation programs
- legislation to prevent certain water uses and pollution of sources

It should also be realized that present ***technological progress*** has made different kinds of resources more intermingled and **interdependent**. Then, "public water supply" doesn't depend exclusively on water availability, but also on energy supply (*for pumping, treatment plants*), human resources (*requirements of trained personnel*), manufacturers and materials available in the region (*pipes, machinery*), and the political and legal framework (*standards, laws, institutions*).

PRESENT CONFLICTS AND NEEDS

Water supply and the possibility of having water and sanitation in our own houses, or at least very close to them, more than any law, invention, or medical advance is what

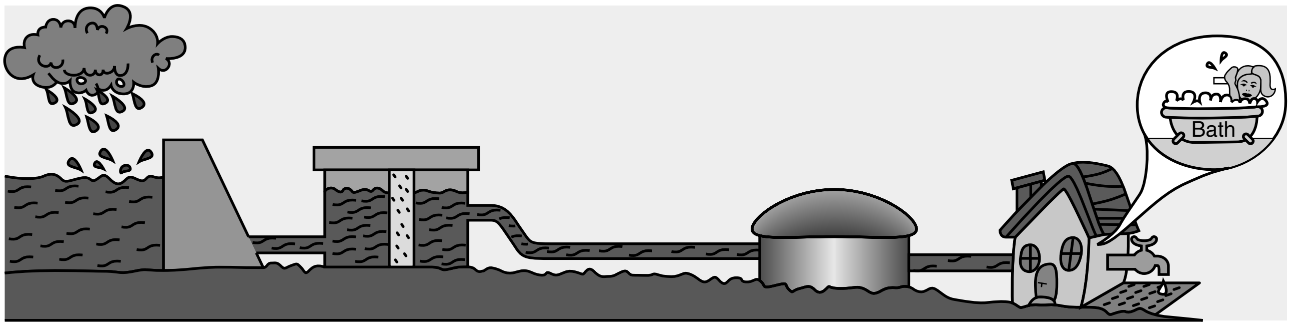

Figure 8. Elementary water supply and disposal system.

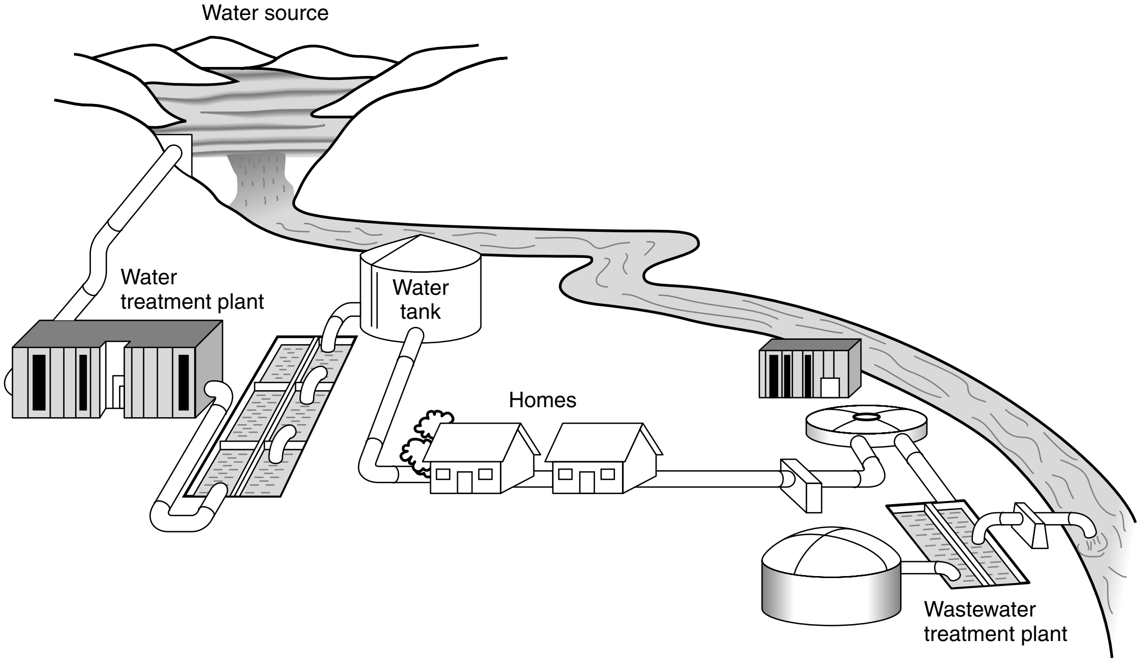

Figure 9. Increasing complexity for supply and disposal system.

has allowed extensive increments in comfort, health, and production and has accelerated the growth of cities. On the other hand, it is shocking that this main force, which made cities swell and consolidate in the past, is now causing the decline of security and comfort in many cities and regions. Past growth, mainly in Europe and well-gifted countries, was slower than that now occurring in poor regions. A slower pace gave the chance for trial and error, without so much risk as today, and also allowed time to mature and integrate useful capital (*material goods, infrastructure, knowledge, and educated people*).

The enormous global human population increase during the last decades, which will still continue during the next 40 years, has occurred almost entirely in the third world (*developing*) countries, ironically, where water scarcity is more evident, as shown by various studies of prestigious international organizations. These poorer countries, where 80% of humanity lives now, and where almost 90% will be in five decades (1–5), are facing “modernity” and high demands for welfare from their people, combined with mighty stresses on natural resources and without much time to think and try the best ways.

Risk or deterioration of those benefits provided by a water supply (*see value of a water supply*), comes from continuous and unrestricted demographic internal urban growth, aggravated by high immigration toward cities due to lack of employment, land, or water for irrigation in rural communities. Expectancies and water demands from these growing populations are difficult to meet without causing damage to other societies (*further in space or time*) or to the natural environment, particularly when the technologies and style of solutions proposed are still those used in the past (*no treatment of wastewater, intensive energy use for pumping, etc*).

The appeal of economies of scale, mentioned earlier, work in the opposite direction when cities are too big and complex (*smaller is better*). One example is Mexico City, where the *benefit/cost ratio* of augmenting the water supply is lower than in other cities of that country.

Some frequent and worrying signs of unsustainable water supply systems are

1. Water scarcity, caused when demands are greater than accessible volumes. Availability must consider the seasonal and regional variations of usable water.
2. Unrecoverable yearly drawdowns in aquifers.
3. Polluted rivers and lakes.
4. Lakes with decreasing water level, going dry or nearly extinguished.
5. Rivers not discharging any more into the ocean, due to deviations and use of their flows.
6. Deforestation threatening water recharge areas, creating erosion and flooding (*flow mode difficult to control*), and causing microclimate changes (*hotter spots*) that increase people's water demands.
7. Rubbish and contamination that menace aquifers and recharge areas.
8. Urban settlements and paving on zones of aquifer recharge or of high risk.
9. Untreated supplied water and free discharge of wastewater.
10. Supply services on rotation (*irregular*) and uncertainty of amounts provided.
11. Huge amounts of water lost through leaks, due to bad quality or status of water networks (*piping quality or poor operation and maintenance*).
12. Incorrect valuation of water. Fictitious or highly subsidized tariffs give the wrong message to the public (*hidden expenditure rob other services, regions, and/or the future*).
13. Institutional weaknesses as procrastination against nonpayment and clandestine connections.

WATER SERVICE BASIC DATA AND INDICATORS

To decide the size of the pipes and other items (*pumping stations, storage tanks, valves, etc*) of a public water network, it is essential to know first some characteristics of consumers' demands and of the supply sources. Studying the demand and ways to manage it, in case it surpasses the supply capabilities, is the prime step for any design. It is also important to know where and how is the existing infrastructure.

There are some typical key indicators used by most water utilities and recommended by international aid organizations (1,2,6). Indicators, parameters, or data describing the characteristics and status of a given water supply system can be grouped in various ways; one example is Table 2.

Tables 3, 4, and 5 list some indicators for three of the previous groups mentioned in Table 2.

This article cannot extend to a detailed discussion of these water service indicators or mention their representative values or ranks in some countries (*they can vary substantially. See reading list for more details*). Only two of them, coverage and demand, will be commented on briefly:

Table 2. Groups of Indicators of a Water Supply System

Water resources indicators
Personnel indicators
Physical indicators
Operational indicators
Quality of service indicators
Community organization and contributions (*e.g., tariff rates and structures*)
Financial indicators

Coverage

Coverage is the ratio of people served to the whole population in a location. It varies widely among countries. Developed nations have values of 100% or near it; in very poor countries, it may be as low as 10 or 20% (*Cambodia, Haiti, Mozambique, etc.*). Moreover, broad differences exist between urban and rural areas in many countries. Certain cities that have apparently satisfactory infrastructure coverage present serious deficiencies due to droughts or useless networks that imply continuous rationing of the supply.

Table 3. Operational Indicators

A. Water consumption
  A.1 Average water demand, requirement, or allotment
  A.2 Unit Consumption
  A.3 Water Consumption & Metering
  A.4 Distribution of Water Consumption
  A.5 Consumption by Main Users Category
  A.6 Ratio of Peak Day to Average Day
  A.7 Water Price & Income Elasticities
B. Water distribution system
  B.1 Coverage or population served
  B.2 Length of Water Piped Systems
  B.3 Storage Volume
  B.4 Pipe Breaks
  B.5 Pipe Breaks as a Function of Pipe Material
C. Unaccounted for water
  C.1 Water Losses
  C.2 Composition of UFW
  C.3 UFW Effective Reduction Programs
  C.4 Sustainability of UFW Reduction Programs
D. Wastewater
  D.1 Length of Sewer Systems
  D.2 Infiltration Flows in Sewer Systems
E. Wastewater treatment
  E.1 Typical Composition of Untreated Municipal Wastewater
  E.2 Typical Constituent Removal Efficiencies
  E.3 Removal of Microorganisms
F. Personnel
  F.1 Number of Staff
  F.2 Staff Composition
  F.3 Training Effort
G. Miscellaneous indicators
  G.1 Vehicles/1000 Water Connections
  G.2 Meter Reading
  G.3 Meter Maintenance and Replacement Practices

Table 4. Financial Indicators

A. Efficiency indicators
  A.1 Working Ratio
  A.2 Operating Ratio
  A.3 Accounts Receivable/Collection Period
  A.4 Percentage Contribution to Investment
B. Leverage indicators
  B.1 Debt Service Coverage Ratio
  B.2 Debt Equity Ratio
C. Liquidity indicator
  C.1 Current Ratio
D. Profitability indicators
  D.1 Return of Net Fixed Assets
  D.2 Return of Equity
E. Operating ratios
  E.1 Personnel
  E.1.1 Personnel Costs
  E.1.2 Staff Productivity Index
  E.2 Composition of Operational Costs
  E.3 Unit Operational Costs

Table 5. Community Organization Contributions

A. Tariff rates and structures
  A.1 Tariff Structure
  A.2 Domestic Tariff
  A.3 Average Charge & Incremental Cost
  A.4 Rate Discrimination by Consumer Group
  A.5 Water Billings, Consumption, & Users

Demand (average water requirement)

Demand represents the volume that needs to be withdrawn from the water sources to satisfy all uses and inhabitants in a city. Usually, it is expressed as a per capita daily volume (*liters/person-day*). It has some similarity to "unit consumption" but is not necessarily the same, especially when the sources are insufficient to meet people's desire for water. The "demand" is a broader concept which includes other water uses, besides domestic. Some "uses" that never should be forgotten when reviewing or designing a supply network are the leaks and other UFW (*unaccounted for water*) that inevitably have to be fed to the system, to be lost so a certain water quantity really reaches individual houses. Figure 10 compares some values in developed countries for unit consumption at household level. Figure 11 shows some typical values in demand variation among seasons, and (Fig. 12) represents distribution of water volumes in a typical Canadian dwelling.

The design, improvement, or revision of a water supply network is relatively easy using modern hydraulic simulation programs. The real labor and complication is in defining, finding, measuring, and gathering the appropriate and correct data to make those calculations and in defining reasonable specifications and standards for

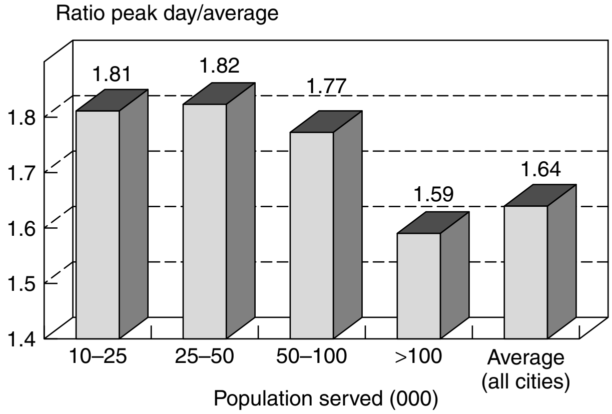

Source: AWWA. Water industry database; utility profiles, 1992.

Figure 11. Demand variations with town size.

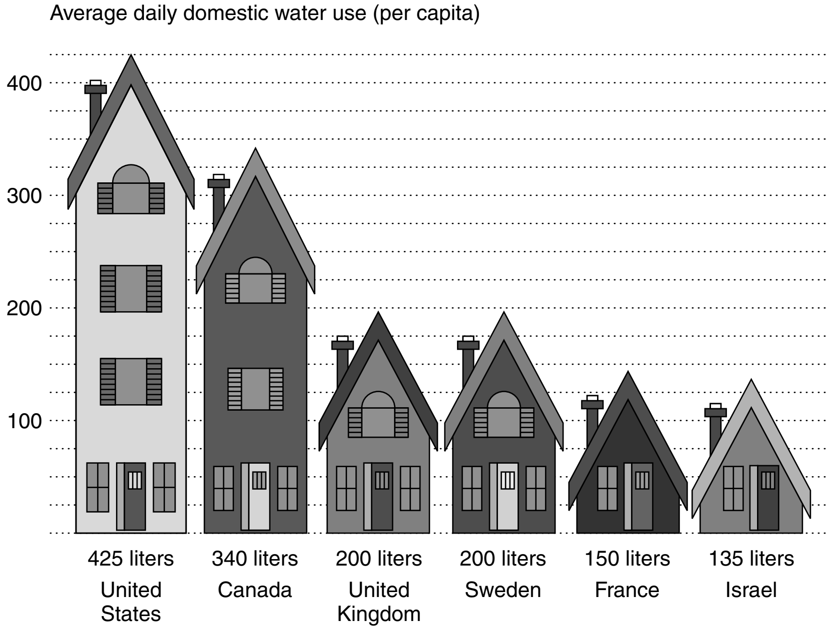

Figure 10. Comparison of household consumption in developed countries.

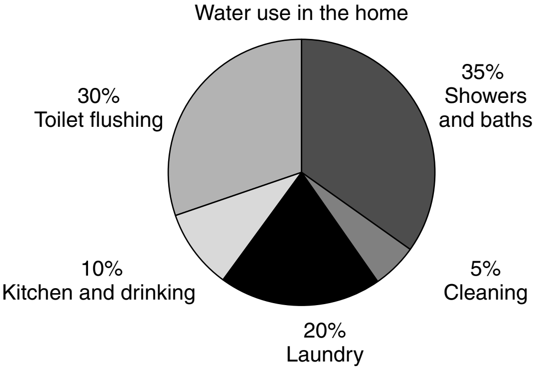

Figure 12. Water uses in a canadian house.

the system. Many textbooks and computer software forget to mention these difficulties. Basic data and indicators are not merely technical aspects, but are rather social, environmental, political and economic, and should be carefully tracked and understood by any person involved in a water supply project.

BIBLIOGRAPHY

1. WRI World Resources Institute. (1999). *Freshwater Data Tables*. WRI. http://www.wri.org/wri/facts/data-tables-freshwater.html.
2. WRI World Resources Institute, (Carmen, Revenga et al.). (1998). *Watersheds of the World. Ecological Value and Vulnerability*. Joint publication of World Resources Institute (WRI) and World Watch Institute. http://www.wri.org.
3. World Bank (Yepes, G. and Dianderas, A.). (1996). *Water and Waste Water Utilities Indicators*. 2nd Edn. Water and Sanitation Division, TWUWS, The World Bank, http://www.worldbank.org/html/fpd/water/pdf/indicators.pdf http://www.worldbank.org/html/fpd/water/topics/utility_o&m.html http://www.open.gov.uk/ofwat/leakage.htm.
4. World Bank. (2000). *World Development Report 2000/2001. Attacking Poverty*. The World Bank Group. http://www.worldbank.org/poverty/wdrpoverty/report/index.htm (*particularly see Table 9. Water use, deforestation, and protected areas*) http://www.worldbank.org/poverty/wdrpoverty/report/tab9.pdf.
5. National Geographic. (October, 1998). *Population. Millennium Supplement*. National Geographic Society. http://www.nationalgeographic.com/ngm/9810/index.html.
6. Alegre, H., Baptista, J.M., Hirner, W., and Cabrera, E. (2000). *The IWA System of Performance Indicators for Water Supply Services*. International Water Association IWA, manual of best practices (performance indicators). http://www.iwapublishing.com/template.cfm?name=isbn 1900222272.

READING LIST

CIWEM. (1984). *Water Distribution Systems. Water practice manual 4*. The Chartered Institution of Water and Environmental Management, CIWEM (*formerly Institution of Water Engineers and Scientists, IWES*). http://www.ciwem.org.uk/publications/manuals/index.asp.

CIWEM. (1983). *Water Supply and Sanitation in Developing Countries*. The Chartered Institution of Water and Environmental Management, CIWEM (*formerly Institution of Water Engineers and Scientists, IWES*). http://www.ciwem.org.uk/publications/manuals/index.asp.

CIWEM. (1986). *Water Services Planning. Water Practice Manual 6*. CIWEM (IWES). http://www.ciwem.org.uk/publications/manuals/index.asp.

Cullivan, D. and Edwards, D.B. (1988). *Guidelines for Institutional Assessment Water and Wastewater Institutions*. Water and sanitation for health project. WASH technical report No. 37, (reprint by EHP Environmental Health Project). http://www.dec.org/pdf_docs/pnaaz336.pdf.

Environment Canada. (1998). *The Management of Water*. National Water Issues Branch, Environment Canada. http://www.ec.gc.ca/water/en/manage/e_manag.htm.

US Environmental Protection Agency. (1977). *World of Fresh Water A Resource for Studying Issues of Freshwater Research*. EPA Table 1– 2. Water use by country. http://www.epa.gov/ORD/WebPubs/fresh/.

Gleick, P.H. (2000). *The Changing Water Paradigm. A Look at Twenty-first Century Water Resources Development*. International Water Resources Association IWRA. http://www.iwra.siu.edu/win/win2000/win03-00/gleick.pdf.

Hinrichsen, D. et al. (1999). *Solutions for a Water-Short World*. John Hopkins University. http://www.jhuccp.org/pr/m14edsum.stm http://www.jhuccp.org/prs/sm14/sm14chap1.stm.

James, W. (1999). *Urban Water Systems*. Guelph University, Canada. http://www.eos.uoguelph.ca/webfiles/james/homepage/Teaching/661/wj661Modules.html.

Lauria, D.T. (1994). *Planning Sustainable Water Systems in Developing Countries*. JAWWA Aqualink.

OIEAU, Office International de l' Eau (International Office for Water). (1998). *Paris Declaration*. International Conference on Water and Sustainable Development. http://www.oieau.fr/ciedd/ang/frames/final/declarfin.htm
http://www.oieau.fr/ciedd/esp/frames/final/declarfin.htm
http://www.waterlink.net/esp/decfinsp.htm.

Postel, S. (1984). *Water: Rethinking Management in an Age of Scarcity*. Worldwatch Institute paper 62, Worldwatch Institute, Washington, DC. http://isbn.nu/0916468623.

Saade, H.L. (1997). *Toward More Efficient Urban Water Management in Mexico*. Water International, IWRA. http://www.iwra.siu.edu/win/toc_22_3.html/mexsaade.htm.

Sumila, G. (2001). The Demand-Side Approach to Planning Water Supply. 1999, In: *The Challenge of Urban Government: Policies and Practices*. Freire, Mila, and Richard Stren. Herndon, VA, The World Bank Group. http://www.aemnet.hu/urbancourse2000/Reading/sumila%20toronto.doc http://www.worldbank.org/wbi/publicfinance/urban/publications.html#1 http://publications.worldbank.org/ecommerce/catalog/product?item_id=213497.

Van der Hoek, W., Konradsen, F., Ensink, J., Mudasser, M., and Jensen, P. (2001). *Irrigation Water as a Source of Drinking Water: Is Safe Use Possible? A Case Study from Pakistan*. Tropical Medicine and International Health, CIAGAR, International Water Management Institute, IWMI. http://www.cgiar.org/iwmi/home/drinking.htm.

Van Wijk-Sijbesma, C. (2001). *The Best of Two Worlds? Methodology for Participatory Assessment of Community Water Services*. Water and Sanitation Program (WSP), International Water and Sanitation Centre (IRC). http://www.irc.nl/products/publications/descr/tp38e.html.

TEN KEY TRENDS THAT WILL SHAPE THE FUTURE OF THE WORLD WATER INDUSTRY

Steve Maxwell
TechKNOWLEDGEy Strategic Group
Boulder, Colorado

As we move into the new millennium, economic observers, social critics, and stock watchers are turning their attention toward the challenges and issues that, surround the country's continuing supply of clean drinking water. Water issues are now routinely addressed on the front pages of the national press—*USA Today* recently featured a full pull-out section dealing with the issues of clean drinking water; water scarcity issues were featured in a recent cover article in the *Harper's Magazine*. Water is increasingly described in such terms as "the key commodity of the twenty-first century." Like oil in the twentieth century, the substance will determine the course of human development and economic progress, over which conflicts or even wars may arise.

Some observers have perhaps waxed a bit melodramatic; however, it *is* clear that these sorts of challenges and problems are not imaginary; nor are they somewhere in the far distant future. Water *quantity* issues have long impacted politics and economic development in the western United States, and now water *quality* issues are increasingly critical in the East. Debates over water rights are the source of intensifying conflict between countries or between regions within countries. The human health challenges and vast economic issue of poor drinking water in many less developed countries represent one of the world's most pressing problems; a recent United Nations report estimated that as many as 10 million people a year die from unsanitary drinking water and associated waterborne diseases.

RISING WATER PRICES

Water is still very cheap; the price of water is not yet even really a significant driver of change in the industry. Prices will continue to rise, particularly in countries like the United States where past government policies—and the failure to account fully for the true *economic* cost of water usage—have kept prices artificially low. As prices increase, decisions relating to water usage will begin to take on greater significance in the overall economy. Focus will begin to intensify on more efficient water usage and demand management techniques, and new treatment technologies and supply sources such as desalination and water recycling will begin to gain market share.

STRONGER ENFORCEMENT OF REGULATIONS

Regulatory attention and water quality enforcement levels will remain strong in the United States, both in terms of drinking water quality and wastewater discharges. Although public concern and congressional attention to broader environmental issues have waned in recent years, the topic of drinking water quality continues to move to the forefront of the public agenda. Regulation of drinking water standards will also be strengthened in most other regions of the world. The water industry will clearly continue to be driven and buffeted by regulatory forces; however, products and services that offer cost savings, revenue enhancements, improved productivity, or reduction of capital expenditures, will enjoy stronger market growth than those that offer just regulatory compliance services.

BROADER REGULATORY FOCUS

Regulation and enforcement has been shifting slightly toward nonpoint sources of runoff and water pollution. Early enforcement attention was focused on the more obvious and more concentrated "point" sources of pollution; admirable progress has been made in correcting these—typically industrial—problems, and recent focus has shifted more to nonpoint source pollution. This (mostly agriculture-related) problem has seen substantial improvement in recent years as well. Biological wetlands treatment systems for animal waste management and pesticide and herbicide runoff are one example of new technologies being perfected to treat agricultural waste problems. Both point and nonpoint sources, as well as storm-water and general "wet weather" management, are likely to receive continuing regulatory attention and scrutiny in the coming decades.

INCREASING PRIVATIZATION

Industrial outsourcing and municipal privatization will remain key trends in the business. The contract operation business has been growing as much as 25% per year, and many private firms continue to diversify into this business despite the fact that, to date, profitability has been elusive. The potential for large-scale water privatization is typically cited as the main driver behind the recent onslaught of foreign acquirers—a belief that the United States is the last large market where water supply is still largely under public ownership and operation.

To date, privatization has been driven by broad political predilections to minimize the scope of government or by the widely held assumption that focused and private for-profit companies can operate more efficiently than public organizations. Many of the institutional and tax barriers to more complete privatization have melted away. To operators and employees of publicly held water and wastewater plants, privatization pressures over the last few years have sent a clear message—"shape up or ship out." However, public agencies and municipalities *have* responded to both the threat and to the root causes of this incipient pressure to privatize. These forces to economize are leading to municipal "optimization" as an alternative to privatization. Optimization may involve cost-cutting and improvements in efficiency from initiatives such as benchmarking, automation of operations, staff reductions, and improved information and analysis to facilitate better decision-making. From a broader social point of view, if the simple threat of privatization ultimately improves the efficiency of municipal facilities, then the same economic and productivity goals will have been achieved. Privatization will undoubtedly continue to grow, but it seems unlikely that it will reach the levels of 80% or 90% of total municipal systems, as has been predicted in the past.

GREATER WATER REUSE

There is an inexorable trend toward broader water reuse and recycling systems as a key means of addressing water shortages. Reuse is not a new idea, and using technologies readily available today, wastewater can easily be cleaned to levels where it can be recycled back into primary usage. This general idea, however, has not yet gained widespread social acceptance; put simply, most people don't want to think about using—and particularly drinking—their own wastewater. We have all been conditioned to think of water as something that we take out of the environment, use, and then discharge back into a vaguely understood natural system. Most observers believe that the *direct* reuse of wastewater for human *drinking* is not likely to be a widespread reality in the foreseeable future. There is still the (clearly disproven but emotionally more appealing) assumption that wastewater is somehow cleaner if it momentarily flows down the Potomac River (or any other less than pristine natural waterway) a few miles, or if it is briefly discharged into and then pumped out of a "natural" underground aquifer. This type of *indirect* use of wastewater for drinking is clearly acceptable and widespread, but *direct* reuse of our own wastewater is still a bit of a stretch for most Americans.

However, the boundaries between "water" and "wastewater" are already beginning to fade. For example, on some major rivers in the United states, water is used and reused up to 20 times as it travels to the sea—the discharge water from one wastewater treatment plant comprises the raw water intake for a primary drinking water plant a few miles downstream. As a result of 30 years under the requirements of the Clean Water Act, discharged streams from wastewater treatment plants are often cleaner than the rivers and streams they flow into; in other words, treated wastewater is actually helping to clean up our rivers.

What many observers have failed to clearly recognize—and what suggests that there may be real business opportunities here in the near future—is the obvious fact that very little of our water is used for drinking. Compared to the roughly 1200 gallons of water per person per day that it currently takes to power the U.S. economy, each person drinks only about a gallon a day. This leaves 1199 gallons per person per day, a substantial portion of which could be recovered through reuse, without anyone ever having to drink "recycled" wastewater. The bottom line is that even if only small incremental gains could be made in terms of nonpotable water reuse, water availability concerns could be substantially impacted. Over the longer term, society will move beyond the "linear" thinking of today and develop a more "circular" philosophy of water usage.

MORE AND BETTER INFORMATION

As the Internet revolution races on, the effects of better and more rapid information are being felt in the water industry. For example, the introduction of the Consumer Confidence Report will undoubtedly spur broader awareness of water quality issues across the American public, in turn driving demand for better monitoring and treatment technologies to improve the quality of our drinking water. Real-time watershed monitoring technologies and telemetry systems are becoming more commonplace and will help improve the quality and lower treatment costs in public water systems. Automatic meter reading, and more accurate remote systems for monitoring water usage will in turn make us more aware and attentive to the issues of our often sloppy usage.

CONTINUING INDUSTRY CONSOLIDATION

The era of consolidation is not over in this industry. Despite all of the attention paid to Veolia, Suez, RWE and other consolidators during the past few years, the water business is still rather fragmented compared to many industries, and business leaders continue to promote loudly the opportunities for "another U.S. Filter." The large transactions have hit the front pages, but dozens and dozens of smaller and less visible deals are also constantly ongoing. This massive rearrangement—and a distinct shift toward foreign ownership of the water industry—will determine the water industry's future competitive landscape. The very occent announcements of major divestitures by both veolia and suez demonstrate that changing ownership will continue to be a major factor.

NEW TECHNOLOGY

The advance of modern technology is often held out by optimists as the answer to all of our environmental and natural resource challenges. Whether or not this will be true is open to debate, but there is no doubt that new technology will play a key role in the improved management and extension of our water resources. Key examples include the broader usage of reverse osmosis/membrane systems to desalinize seawater for human usage, the trend toward replacing chlorine and other water treatment chemicals by advanced oxidation and other emerging physical and mechanical technologies, water-saving drip irrigation systems for more efficient agriculture, a wide spectrum of recycling systems and technologies, and new packaging and transportation systems for moving freshwater from regions of abundance to regions of scarcity.

CONSERVATION AND MORE EFFICIENT USAGE

Almost all of the trends mentioned before underscore the questionable efficiency and economics of today's water system and infrastructure. Some simple steps in more efficient usage can go a long way toward solving water shortages. For example, even the most modern water transmission infrastructure systems often incur leakage of as much as 10 to 20%—losses which can be fixed relatively easily. Water conservation ideas are only just beginning to take advantage of the “low-hanging fruit”; a recent rebate plan in New York City to promote the use of low-flow toilets purportedly resulted in saving 90 million gallons a day (although many users find “low-flow” toilets to be a bit of a misnomer, as they simply have to be flushed more times)! As water prices rise, there will be increasing financial incentives for industrial companies to retool their manufacturing systems to use less water, or to reuse water more efficiently—and dollars get everyone’s attention. And different types of applications can use different kinds of water—urban planners have pointed to the fallacy of spending billions of dollars annually to clean *all* of our water supply to drinking standards when, as mentioned, only a tiny fraction is actually used for drinking. More sophisticated systems and infrastructure in the future (such as dual potable and nonpotable piping in the home) may allow us to be more efficient in our management and use of limited water resources.

MORE INNOVATIVE POLICIES AND REVOLUTIONARY APPROACHES TO WATER USE

From a broader social and economic perspective, we will see many changes in infrastructural approaches and broader water management policies. For example, water “carriage”—analogous to the wheeling of electricity in the power utility business, where one utility may use another’s infrastructure for delivering water—is being experimented with in England. One historical trend that is likely to begin to reverse in the future is the construction of large-scale dams for power and irrigation purposes; the broader social and economic impact of large-scale dams is being gradually reevaluated, and it is conceivable that in the future we could see some of the big dams coming down. Water and hydropower utilities are beginning to come together in certain regions, realizing that water management issues are the driving forces behind each. In short, the entire way we think about and use water will be markedly different just a couple of generations from now.

CONCLUSION

As the saying goes, “it is very difficult to make predictions, especially about the future.” The water industry has confounded many observers in the past and continues to perplex many today; for example, if the demands are so great and the opportunities so huge, why aren’t more water companies currently performing better in the stock market? No doubt, the business will continue to be a tough one to define, segment, and understand, and it will probably take some more unexpected turns in the future. One thing, however, is clear, and distinguishes this industry from almost all others—water is an essential prerequisite of life. A continuing and dependable source of clean water is a critical necessity for our standard of living and our modern industrial economy. *And we are not going to find a substitute for water*. For well-managed organizations—both private and public—and for the thoughtful and prudent investor, the fundamentals of the water industry suggest the potential for strong growth and consistent profitability far into the new millennium.

ZEBRA MUSSEL CONTROL WITHOUT CHEMICALS

Charles H. Sanderson
T. C. Molden
Magnatech Corporation
Fort Wayne, Indiana

Since the introduction and spread of the zebra mussels *(Dreissena polymorpha)* in the mid-1980s, industry and municipal water users of North America have been plagued by the mussel shells, which have clogged water intakes, costing vast amounts of downtime and loss of production (1).

The U.S. Coast Guard has since named the zebra mussel the number one threat to American waterways as well as a threat to fish and other aquatic species. The possible effects on numerous communities is also of great concern (2). In 1994, the Great Lakes Environmental Research Laboratory stated that the barnacle-like zebra mussel poses a multibillion dollar threat to industrial and public drinking water supplies and may become a costly nuisance to shippers, boaters, commercial fishermen, anglers, and beach-goers as well—far more costly in human terms than all previous Great Lakes invaders combined (3).

The life cycle of the *Dreissena polymorpha* begins at the egg stage, advances to the veliger, postveliger, juvenile-settling stage, and then to a full sized adult. In the settling stage, mussels colonize very rapidly in areas where calcium and algae are plentiful. A water velocity of 6 feet per second, with high concentration of algae and calcium, is an ideal location for the zebra mussel to settle and develop its shell consisting of 93% calcium (4). The mussel ideally settles in an intake pipe where the food source is constantly supplied by the moving water.

The adult life span is 3–5 years, and it reproduces year round, averaging 35,000 eggs per year (5). Zebra mussels look like small clams, and they have yellowish and/or brownish “D”–shaped shells, usually of alternating dark and light bands of color (thus the name “zebra”) (Fig. 1).

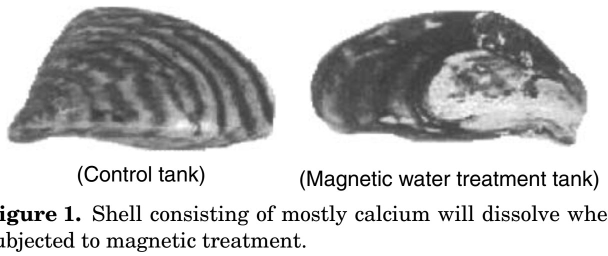

Figure 1. Shell consisting of mostly calcium will dissolve when subjected to magnetic treatment.

Zebra mussels can grow as large as 2 inches, but most are under one inch long, usually grow in clusters containing numerous individuals, and are generally found in shallow (6 to 30 feet deep) water. The zebra mussel is the ONLY freshwater mollusk that firmly attaches itself to solid objects, including rocks, boat hulls, and internal walls of pipes.

Methods to control the zebra mussel became top priority in many municipalities and industries that use water from the Great Lakes. In desperation, chemicals (mostly chlorine) became the quick fix; however, the toxic discharge back into the fresh water lakes increased the fish kill and killed other wildlife species. The U.S. Army Corps of Engineers sought other (nonchemical) means to control the spread of the mussel threat and set up workshops in which academic and industry professionals were invited participate.

The National Research Council listed several technologies more friendly to the environment than chlorine and other chemical biocides. Among those identified were *filtration systems, thermal techniques, electric pulse and pulse plasma techniques, ultraviolet treatment, acoustic systems, magnetic fields, deoxygenation*, and *antifouling coatings*.

Under the heading of magnetic fields, the Council stated, "Water to be treated is passed through a magnetic field of specified flux that is generated by ferromagnetic or electromagnetic devices. The biological and chemical effects of magnetic systems are not well understood, but it is thought that the organic and inorganic constituents of living organisms in the water are altered by the magnetic field" (6).

Magnetic conditioners up to capacities of 12,000 G.P.M. and pipe 60" in diameter (Fig. 2) treatment has been used successfully to control hard water scale ($CaCO_3$) in boilers, cooling towers, and other heat transfer systems for many years. The need for chemicals to control hardness and scale can be eliminated or greatly reduced and can provide energy savings (7). The mussel shell has a high calcium content (formed from the hard water intake), so a 2-year study began in 1993 to determine what effects, if any, magnetic fields would have on the zebra mussel. The results were very encouraging, as shown in (Fig. 3).

Dr. Graham O. Davies, president of the Water Board of Long Beach, Indiana, graciously offered the use of his facilities for the study and obtained the approval of the EPA for the installation of the magnetic unit and test equipment (Fig. 4). The Bardygula-Nonn study (8) used a flow-through magnetic water treatment device. Based on the results, it was suggested that magnetic water treatment might be a means of controlling zebra mussel populations to prevent macrofouling (Fig. 3). Besides the shell deterioration, the magnetic water treatment appeared to have other negative impacts on zebra mussels, including ciliary damage, byssal thread destruction, valve growth reduction, and increased mortality of juveniles.

A 4 year study started in 1995 at Purdue University (9) had three objectives: (1) to determine if the results of

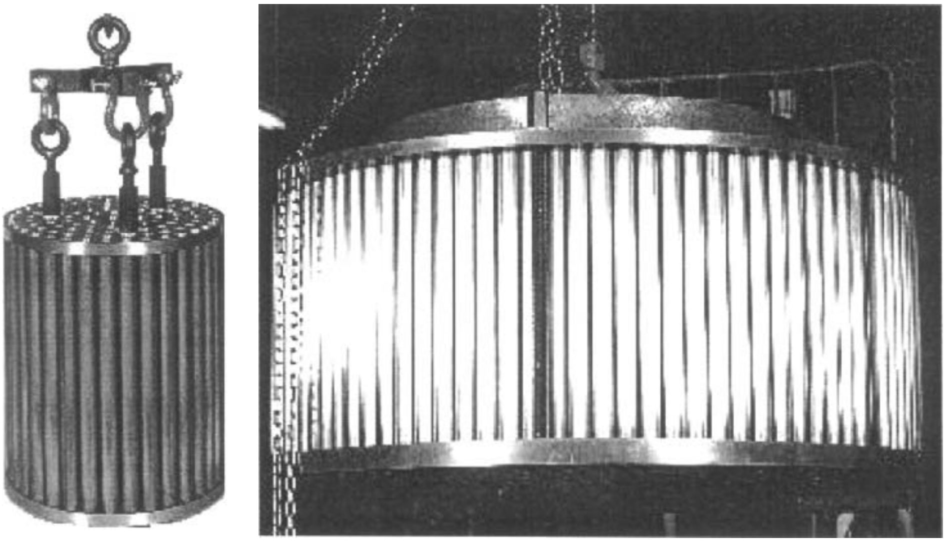
(20" dia. – 2,000 g.p.m.) ***Magnetic Perma-core™*** (60" dia. – 12,000 g.p.m.)

Figure 2. Magnetic assembly for inline treatment.

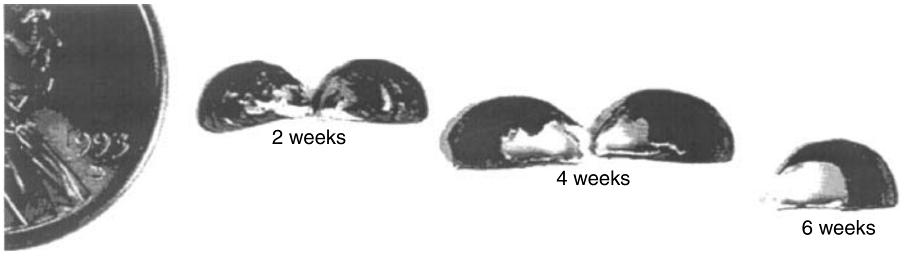

Figure 3. Zebra mussels exposed to magnetic water treatment- no chemicals were used.

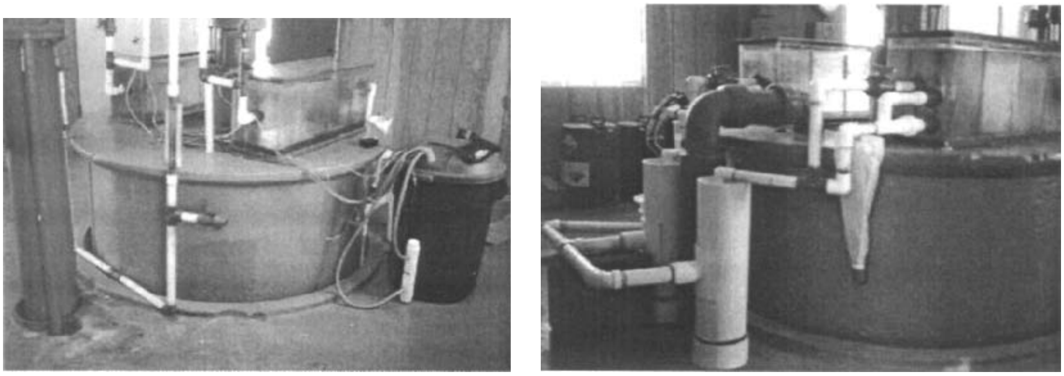

(Long beach facilities showing control tank (left) and magnetically treated water tank (right) with nutrient feeding tank)

(back view of the tanks with plankton nets filtering discharge water)

Figure 4. Research arrangement for side by side evaluation for mussels placed in the control tank vs. magnetic treated water.

the previous study could be reproduced in a controlled laboratory set-up, (2) to examine the effects of magnetic water treatment on adult zebra mussels in an industrial application, and (3) to examine the influence of magnetic water treatment on settling of veligers in an industrial application.

In the laboratory at Purdue, 65 of the 123 mussels from the control tank (those not used for dissection or elemental analyses) were alive at the end of the experiment, leaving 58 dead (Fig. 5). In the experimental tank, 33 of the 123 mussels were alive at the end of the experiment and 90 died during the course of the study (Table 1). The differential mortality levels observed from each tank were significantly different ($X_2 = 17.4 \text{ m df} = \text{in}, p < .005$). Therefore, significantly more zebra mussels died in the tank of magnetically treated water than in the tank of untreated water. This suggests that the magnetic treatment of water does have adverse effects on zebra mussels, under laboratory conditions, and increases the mortality rate. This coincides with the previous Bardygula-Nonn study.

INDUSTRIAL FIELD STUDIES ON ADULT MUSSELS

Mussels treated for 30 days had significantly lower wet-tissue weights than control mussels ($F = 8.10$, $p < .005$). Glycogen levels of the treated mussels were 17.6% lower compared with control mussels, a difference that was significant ($t = 2.23, \text{df} = 118, p = .028$). Treated mussels had a greater increase in shell length during the 30-day trial than control mussels. The difference was significant at ($t = 3.16, \text{df} = 118, p = .002$). No differences were observed between the control and treatment circuits for water quality or environmental parameters during the 30-day trial.

There was no significant increase in shell length of treated mussels during the 60-day trial compared with control mussels (as seen in the 30-day trial). Glycogen analyses were not done for these samples. The results of the Purdue study suggest that magnetic water treatment can reduce settlement of zebra mussels in an industrial setting and that the mortality of zebra mussels is increased when they are exposed to magnetically treated water. A combination of reduced chemicals such as chlorine and a magnetic unit may sometimes be required. Interestingly, the head chemist at an industrial field site reduced their first year's chemical use by 50%, then by 75%, and presently no chemicals at all are being used in the magnetically treated supply line (Fig. 6). The parallel control line remains at the 100% chemical-feed rate (9).

Table 1. Survival of Zebra Mussels Over the Course of the Laboratory Study

	Control Tank	Experimental Tank
Alive at end	65	33[a]
Died during	58	90
TOTALS	123	123

[a] $X_2 = 17.4$, $p < .005$; therefore, there is a significant difference between tanks in the survival of the mussels.

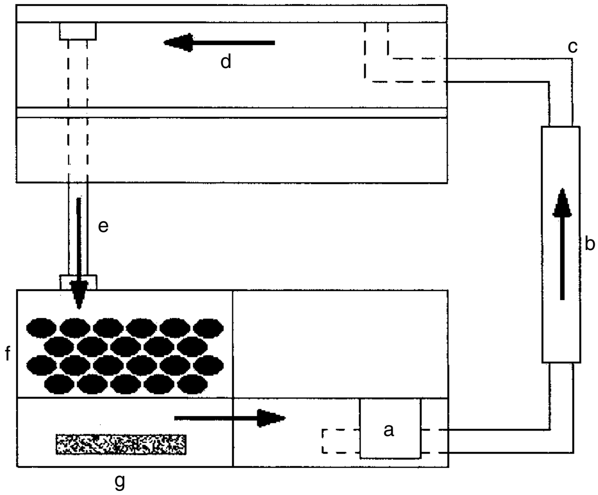

Figure 5. Diagram depicting setup of aquaria (control & treatment): (**a**) submersible pump, (**b**) magnetic treatment (treatment) or "dummy" unit without magnet (control), (**c**) water return line, (**d**) mussel bed, (**e**) overflow feed line, (**f**) biological filter media, and (**g**) aerator unit (air stone).

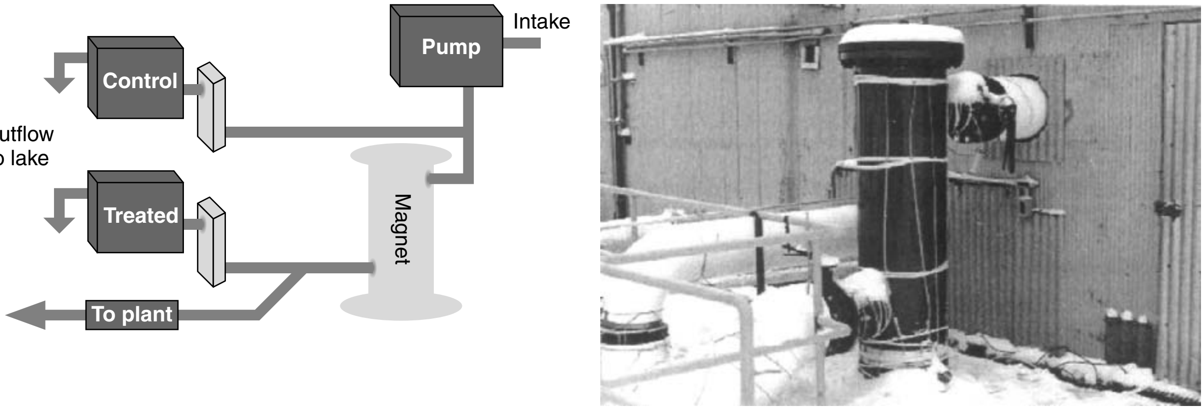

Figure 6. Diagram of industrial setup. Industrial magnetic water treatment system installed at a manufacturing plant on a parallel feed line with a flow rate of 1600 gallons per minute.

INDUSTRIAL FIELD STUDIES ON VELIGER SETTLING

A 53% reduction in settlement was observed in the treatment sampler compared with the control sampler. Chi-square analysis of settlement data revealed that the reduction was significant ($p > .001$). No significant differences were observed in recorded water quality parameters between the two sampling boxes. The results of this study suggest that magnetic water treatment can reduce settlement of veligers in an industrial water intake.

RESEARCH CONCLUSIONS

At this time, it is not clear how magnetically treated water could cause the greater mortality observed in the experimental tank compared with the control tank. Perhaps the changes, which occur when water is magnetically treated, serve as a stressor or influence cell membrane functions. It is known that magnetic water treatment causes the electrons of calcium ions to move, temporarily, into different orbitals (10). Perhaps this change in the ions can influence the zebra mussel's calcium uptake or metabolism. Calcium is especially important in shell formation, cell membrane function, muscular contraction, and ciliary contraction in zebra mussels. Therefore, if magnetism were affecting free calcium ions in the water, it would have potentially wide ranging effects on the zebra mussel's metabolic functions, which could generally explain an increase in mortality.

Ongoing research has shown that, by inserting probes at each of the alternating reversing magnetic pole sites and using active filters to eliminate any common noise (0.40 Hz) at the outputs, not only is a voltage induced in the conductor (water) but also a frequency. The frequency is a complex waveform whose fundamental frequency is <30 Hz. This falls into the extremely low frequency (ELF) spectrum of electromagnetic energy. More research is being done today on this aspect of the technology (9).

The biological effects of ELF radiation are well documented (11). ELF radiation lowers the energy barrier of cellular ion transport between the high-dielectric aqueous phase and low-dielectric lipid-containing barriers. This will influence the ion influx and efflux for the biological systems. The bonding interactions of ions with cellular membranes and tissue, which will consequently influence normal cellular functions, are also affected by ELF radiation. ELF, it has been shown, affects ion binding to cellular-membrane macromolecules, influences transmembrane ion transport, and alters membrane-signaling events. Ryan has suggested that ELF might affect the mussel's ability to assimilate calcium from water and also remove calcium from the shells and bodies of the zebra mussels. Proper containment of zebra mussel is necessary to prevent further spread of this nuisance non-indigenous species throughout the United States fresh water rivers and lakes (Fig. 7). Besides the distribution through connecting water ways, the mussel can survive a 3,000 mile trip by boat and trailer if the boat is not adequately cleaned after being pulled from a contaminated lake.

BIBLIOGRAPHY

1. Herbert, P.D.N., Muncaster, B.W., and Mackie, G.L. (1989). Ecological and genetic studies on dreissena polymorpha (pallas): A new mollusk in the great lakes. *Can. J. Fish. Aquat.* SC:**46**: 1587–1591.
2. Litherland, (1994). LCDR Phil, *Executive Officer, Naval Engineering Support Unit, U.S. Coast Guard*, Denver, CO. Zebra Mussel Workshop.
3. Beeton, A.M. (1994). *Director, Great Lakes Environmental Research Laboratory—Great Lakes Sea Grant Institute*. U.S. Army Corps of Engineering, Zebra Mussel Research program.
4. Burky, A.J. and Way, C.M. (1994). *Dayton University, U.S. Army Engineer Waterways Experimental Station, U.S. Army Corps of Engineers*, Zebra Mussel Workshop.
5. Liston, C.R. and McNabb, C.D. (1994). *Bureau of Reclamation. Research activities at the Bureau of Reclamation related to potential invasions of the zebra mussel*.
6. National Research Council—Stemming Tide. (1996). *Controlling Introduction of Non-indigenous Species By Ships' Ballast Water*. National Academy Press, Washington, DC.
7. Department of Energy Federal Technology Alert. (1998). "*Non-Chemical Technologies for Scale and Hardness Control*." DOE/EE-0162.
8. Bardygula-Nonn, L.G. (1994) *Acute and Sub-acute effects of magnetically treated water on adult zebra mussels (Dreissena polymorpha) ZMT Industries, Inc. Fort Wayne, IN*.

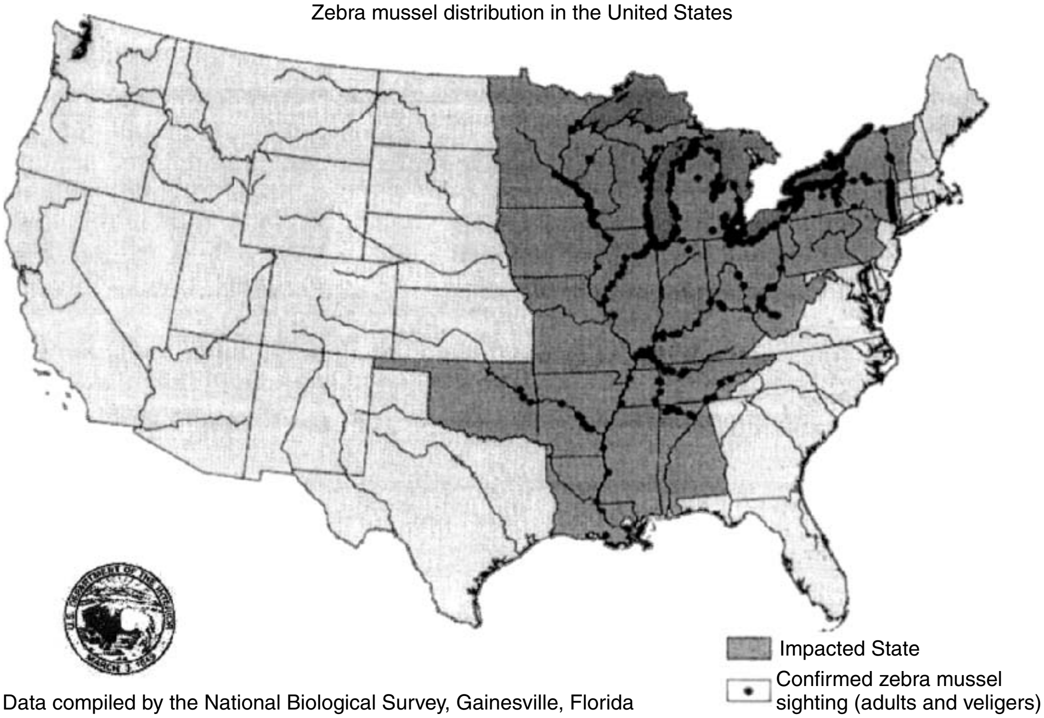

Figure 7. The zebra mussel was first discovered in Lake St. Clair in June of 1988 and has traveled throughout the Midwest region and other parts of the country.

9. Barnes, D.K., Camp, J.W. and Blaney, L.M. (1991) *Effects of Magnetically Treated Water on Zebra Mussels*. Purdue University, Westville, IN.
10. Molden, T.C. *Molden Associates—Research Report on frequency generated when flowing water passes through an alternating reversing polarity magnetic field. Water Services Technology Inc., Laporte, IN.*
11. Lin, I.J. (1981). *Conditions of Industrial Waters by Magnetic and Mechanical Means*. Department of Chemical Engineering, Technion—Israel Institute of Technology, Haifa, Israel.
12. Ryan, M.F. *Extremely Low Frequency Electromagnetism: An effective non-chemical method to control zebra mussel infestation*. Department of Chemistry and Physics, Purdue University Calumet, Hammond, IN.

PACKAGE PLANTS

National Drinking Water
Clearinghouse

Small water treatment systems often find it difficult to comply with the U.S. Environmental Protection Agency (EPA) regulations. Small communities often face financial problems in purchasing and maintaining conventional treatment systems. Their problem is further complicated if they do not have the services of a full-time, trained operator. The Surface Water Treatment Rule (SWTR) requirements have greatly increased interest in the possible use of package plants in many areas of the country. Package plants can also be applied to treat contaminants such as iron and manganese in groundwater via oxidation and filtration.

PACKAGE PLANTS: ALTERNATIVE TO CONVENTIONAL TREATMENT

What is a Package Plant?

Package technology offers an alternative to in-ground conventional treatment technology. They are not altogether different from other treatment processes although several package plant models contain innovative treatment elements, such as adsorptive clarifiers. The primary distinction, however, between package plants and custom-designed plants is that package plants are treatment units assembled in a factory, skid mounted, and transported to the site.

These units are most widely used to treat surface water supplies for removal of turbidity, color, and coliform organisms with filtration processes. However, many other treatment technologies are available to small systems as package plants. These technologies or a combination of them can be incorporated into a package plant to provide comprehensive water treatment, including:

- disinfection (chlorination, ozonation, ultraviolet radiation),
- filtration (bag and cartridge filters, membrane filtration including reverse osmosis or ultra filtration, slow sand filtration, pressure filtration, diatomaceous earth filtration),
- aeration,

- ion exchange,
- adsorption (using powdered activated carbon or granular activated carbon), and
- softening.

How to Select a Package Plant

Package plant systems are most appropriate for plant sizes that treat from 25,000 to 6,000,000 gallons per day (GPD) (94.6 to 22,710 cubic meters per day). Influent water quality is the most important consideration in determining the suitability of a package plant application. Complete influent water quality records need to be examined to establish turbidity levels, seasonal temperature fluctuations, and color level expectations. Both high turbidity and color may require coagulant dosages beyond many package plants design specifications. Pilot tests (tests that evaluate treatment processes and operations on a small scale to obtain performance criteria) may be necessary to select a package plant for more innovative designs using high flow rates and shorter detention time unit processes. Package treatment equipment manufacturers often perform these tests.

System Description and Design Considerations

Package plants can differ widely with regard to design criteria, and operating and maintenance considerations.

Package Plant Advantages

Package plants arrive on site virtually ready to operate and built to minimize the day-to-day attention required to operate the equipment.

Other major advantages of package plants are compact size, cost effectiveness, relative ease of operation, and design for unattended operation.

The main advantages of a packaged factory-finished system are savings in engineering, design and installation costs, and operation and maintenance. These features make package plants attractive to communities that must operate on a tight budget.

Package plants can effectively remove turbidity and bacteria from surface water of fairly consistent quality, provided that they are run by competent operators and are properly maintained. Package plants also can be designed to remove dissolved substances from the raw water, including color-causing substances and trihalomethane precursors (which are organic materials in source water that can react with chlorine to form disinfection by-products).

Package Plant Limitations

Highly variable influent water quality requires a high level of operational skill and attention, and that tends to negate the package plant advantages of low cost and automation.

Despite the relatively large number of package plants in use, many states are reluctant to endorse them completely.

The requirements of the Safe Drinking Water Act and its amendments might challenge package capability. Challenges include the possible inability of package plants to treat multiple types of contaminants.

Many communities are currently using package plants to treat water supplies, but little data has been collected to demonstrate long-term performance and operations of these systems. State agencies responsible for reviewing plans for the installation of package systems must review each potential plan on a case-by-case basis, with only their own experience to judge the potential for success or failure. Presently, there is no national verification process for package plants.

Types of Package Plant Filtration Systems

Conventional Package Plant. Conventional package plants are manufactured by several firms to a variety of specifications. As their name indicates, they contain the conventional processes of coagulation, flocculation, sedimentation, and filtration. Typical design standards for these units are:

- 20 to 30 minute flocculation detention time,
- 2 hour sedimentation detention time, and
- rapid sand filters rated at 2 gallons per minute (gpm) per square foot (1.34 liter/second/square meter).

Tube-Type Clarifier Package Plant. Tube-type clarifier package plants use tube settlers to reduce settling detention time (the average length of time water remains in the tank or chamber).

A flow diagram of a tube-type clarifier package plant is illustrated in (Fig. 1). This type of plant has two versions with different capacity ranges; one version can treat from 10 to 100 gpm (0.63 to 6.3 liters/second), and the other, equipped with dual units, can treat from 200 to 1,400 gpm (12.6 to 88.3 liters/second).

In these package systems, the disinfectant, primary coagulant, and polymer coagulant aid are added before the influent enters the flash mixer; then the water enters the flocculation chamber where mechanical mixers gently agitate the water for 10 to 20 minutes depending on the flow.

The flocculated water then enters the tube settlers, which consist of many 1 inch (2.5 centimeters [cm]) deep, 39 inches (99 cm) long split hexagonal shaped passageways. The large surface area of the many 1 inch (2.5 cm) deep tube settlers achieves an effective clarification overflow rate of less than 150 GPD/square foot (6.1 cubic meters/day/square meters). Adequate clarification is attained with less than 15 minute detention times.

The clarifier water then enters a gravity flow mixed-media filter (a filter with a coarse-to-fine gradation of filter media or several types of filter media). A constant filtration rate is maintained by a low-head filter effluent pump discharging through a float-operated, level control valve. After a preset headloss is reached, backwashing of the filter is initiated automatically. (Headloss is pressure or energy loss as a result of turbulence caused by the velocity of water flowing and the roughness of the channel walls.) A manual backwash cycle can be initiated any time (if desired). Settled sludge from the tube settlers is flushed during the backwashing cycle. Combining backwashing and tube settler flushing simplifies operations and reduces operator skill requirements.

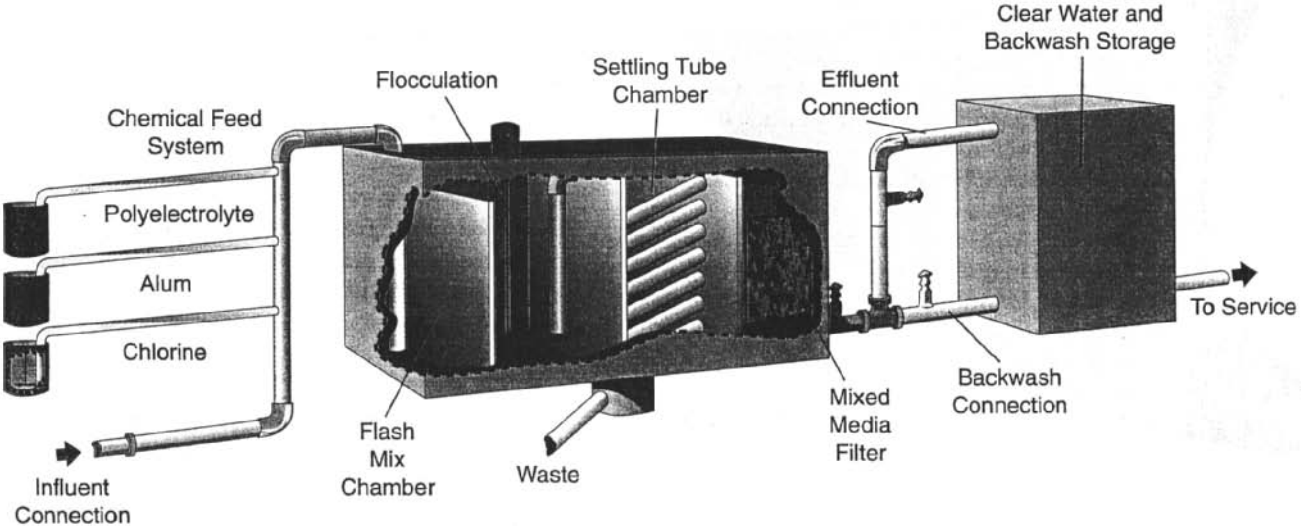

Figure 1. Tube-type clarifier package plant. *Source*: U.S. Environmental Protection Agency, 1989.

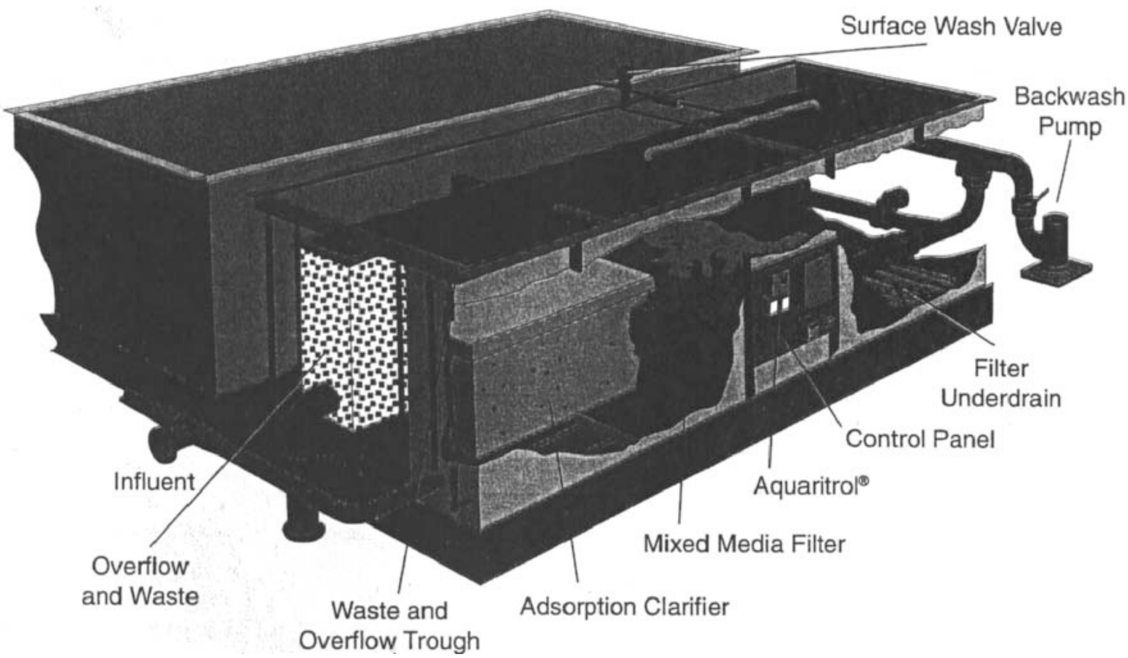

Figure 2. Adsorption clarifier package plant. *Source*: U.S. Environmental Protection Agency, 1989.

Adsorption Clarifier Package Plant. Adsorption clarifier package plants use an upflow filter with low density plastic bead media (called the adsorption clarifier) to replace the flocculation and sedimentation basin, thereby combining these two steps into one. A mixed media filter completes the treatment. Figure 2 shows a typical example.

While passing through the adsorption media, the coagulant and water are mixed, contact flocculated, and clarified. The mixing intensity, as measured by the mean velocity or gradient, ranges from 150 to 300 feet per second. Flocculation is accomplished through turbulence as water passes through the adsorption media. In addition, flocculation is enhanced by contact between the flocculated materials and the floccoated media. Turbidity is reduced through adsorption of the coagulated and flocculated solids onto the adsorption media and the previously adsorbed materials. The adsorption clarifier can achieve 95 percent or greater removal at 10 gpm/square foot (6.8 liters/second/square meter). This highly efficient clarification method results in extremely compact designs.

Adsorption clarifiers are cleaned by a combination of air scouring followed by water flushing. The air scouring starts the cleaning process for the plastic media used in the clarifier. Adsorption clarifier cleaning is initiated more frequently than filter backwashing because more solids are removed by the clarifier. The clarifier cleaning process is automatically initiated either by a timer or a pressure switch that continuously monitors headloss across the adsorption media.

The air introduced under the adsorption media causes a vigorous scrubbing action. The scrubbing action dislodges solids, which are washed away by the flow of the incoming water. Flushing is generally timed to occur between every fourth and eighth hour. Complete cleaning of the adsorption media is not desired because performance is enhanced by some residual solids.

Operation and Maintenance

Package plant operation is simplified by automated features such as effluent turbidimeters connected to chemical feed controls and other operating parameters, such as backwashing. Chemical feed controls are especially important for plants without full-time operators or with variable influent characteristics. Maintenance requirements are well documented in manuals. However, the operator needs to be well acquainted with water treatment principles and the plant manual, and should have attended a comprehensive training session.

The effluent turbidimeters and fail-safe controls are built into many plants to ensure that the finished water does not exceed set turbidity levels. Automated chemical feed systems are especially appropriate for plants without full-time operators or with highly variable influent characteristics.

Typical plant operator and maintenance manuals contain operating principles, methods of establishing proper chemical dosages, operating instructions, and trouble shooting guides.

Periodic visits by the manufacturer to make adjustments to the plant and inspect the equipment operation and performance are recommended. The first visit should be no more than six months after initial operation; the next should follow in another six months. Subsequently, annual visits should be sufficient.

Operators are the critical factor in overall success of any package plant, particularly in situations where raw water quality varies. When the automation fails, the operator needs to turn off the automatic controls/instrumentation and operate the plant manually.

Where Can I Find More Information?

Information in this fact sheet was obtained primarily from two sources:

(1) *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*, EPA/625/4-89/023; and
(2) “Package Plants for Small Systems: A Field Study” by Susan Campbell, Benjamin W. Lykins Jr., James Goodrich, Dallas Post, and Trudie Lay. *AWWA Journal*, Nov. 1995 pages 39–47.

Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities can be ordered free from the U.S. Environmental Protection Agency Office of Research and Development at (513) 569-7562.

The National Drinking Water Clearinghouse (NDWC) offers this document also, but at a cost to help recover photocopying expenses. *Technologies for Upgrading Existing or Designing New Drinking Water Treatment Facilities*, a 209-page book, costs $30.05. To order, call the NDWC at (800) 624-8301 or (304) 293-4191 and request item #DWBKDM04. NDWC products also may be ordered via e-mail at ndwc_orders@ndwc.wvu.edu. Postal charges are added to orders.

Also, the NDWC's Registry of Equipment Suppliers of Treatment Technologies for Small Systems (RESULTS) is a public reference database that contains information about technologies in use at small water systems around the country. For further information, call the NDWC at one of the above numbers.

Additional free copies of Tech Brief fact sheets are available at the above numbers or you may download Tech Briefs from our Web site at http://www.ndwc.wvu.edu.

ANAEROBIC SEWAGE TREATMENT

Lucas Seghezzo
Grietje Zeeman
Gatze Lettinga
Wageningen University and
Research Center
Wageningen, The Netherlands

WHAT IS SEWAGE?

Sewage is the wastewater generated by a community, namely: a) domestic wastewater, from bathrooms, toilets, kitchens, etc., b) raw or treated industrial wastewater discharged in the sewerage system, and sometimes c) rainwater and urban runoff (1). Domestic wastewater is the main component of sewage, and it is often taken as a synonym. Sand and coarse material (paper, bottles, etc.) are not considered part of sewage. They are transported by sewage but handled as solid waste when they arrive at a treatment facility. The sewage flow rate and composition vary considerably from place to place, depending on economic aspects, social behavior, type and number of industries in the area, climatic conditions, water consumption, type of sewers, etc. The main pollutants in sewage are suspended solids, soluble organic compounds, and fecal pathogenic micro-organisms. In addition, a variety of chemicals, like heavy metals, trace elements, detergents, solvents, pesticides, and other unusual compounds like pharmaceuticals, antibiotics, and hormones can also be detected in sewage. With urban runoff come potentially toxic compounds like oil from cars and pesticides that may reach the treatment plant and, eventually, a water body.

WHY SHOULD WE TREAT SEWAGE?

Direct discharge of raw or poorly treated sewage into the environment is one of the main sources of pollution on a global scale (2). Improved sanitation would have a significant impact on people's lives in terms of safety, privacy, convenience, and dignity (3). Sanitation is also a good starting point for addressing long-term poverty issues and reducing child mortality, because children are more susceptible to suffer from inadequate water supply and sanitation services. The lack of water, sanitation, and hygiene for all was dubbed as "... one of the biggest scandals of the last 50 years" (4). Simple, affordable, and efficient sewage treatment systems are urgently needed, especially in developing countries, where most of the conventional technologies currently in use in industrialized nations are too expensive and complex (5). Sustainable sewage treatment technologies will help to preserve water ecosystems and their biodiversity, indispensable for the provision of clean water, flood control, and other vital services.

Today, several technological options are available for sewage treatment, ranging from traditional waste stabilization ponds (WSP) to conventional aerobic systems (like trickling filters or activated sludge), from anaerobic reactors to integrated systems in which a variety of biological processes can be applied. Anaerobic processes are attracting the attention of sanitary engineers and decision-makers more and more, especially the upflow anaerobic sludge bed (or blanket) (UASB) reactor developed in the early 1970s by Lettinga and co-workers (6). Anaerobic sewage treatment in UASB reactors is an absolute success in tropical countries like India and Brazil, but it is also finding its way in other regions, even in subtropical and more temperate countries. Recent studies showed that it can be successfully applied at temperatures as low as 15 °C for a variety of different types of sewage (7–9). The lack of sewage treatment is not at all a technological or economic problem but a political one. In fact, governments, even from developing countries, have been spending billions of dollars per year on water and sanitation, usually to build expensive, centralized sewers and sewage treatment facilities for a few privileged citizens instead of low-cost appropriate sanitation systems for the needy majority.

THE PROCESS OF ANAEROBIC DIGESTION

Anaerobic bacteria, belonging to the domain Archaea, are the most ancient forms of life on earth. They evolved about 3.7 billion years ago in an oxygen-free atmosphere and fed from the nutrients in the aqueous environment around them. When, because of the action of cyanobacteria and algae, the atmosphere became dangerously full of oxygen, anaerobic bacteria were relegated to niches where anaerobic conditions could still be found. Anaerobic bacteria thrive today in the bottom of marshes, swamps, and wetlands, and in the stomach of ruminants, insects, and humans, where they degrade organic matter and produce biogas, a mixture of methane and carbon dioxide with a small percentage of hydrogen sulfide and traces of other gases.

The anaerobic (bio)degradation of complex, particulate organic material has been described as a multistep process of series and parallel reactions catalyzed by several groups of bacteria (10,11). Seven subprocesses are now recognized during anaerobic digestion: (1) hydrolysis of complex, particulate organic materials (proteins, carbohydrates, and lipids); (2) fermentation of amino acids and sugars; (3) anaerobic oxidation of long-chain fatty acids and alcohols; (4) anaerobic oxidation of intermediary products such as short-chain fatty acids (except acetate); (5) acetate production from carbon dioxide and hydrogen (homoacetogenesis); (6) conversion of acetate to methane (aceticlastic methanogenesis); and (7) methane production by reduction of carbon dioxide by hydrogen.

Hydrolysis was found to be the rate-limiting step in the degradation of wastewaters like sewage, high in volatile suspended solids (12). Hydrolysis rate in anaerobic systems is normally described as a first-order process with respect to the concentration of degradable particulate organic matter (13). Preliminary conversion mechanisms such as cell death and lysis are the first steps in the process of transforming viable micro-organisms to available substrate and can also be included in a kinetic model of the anaerobic digestion of aerobic sludges. However, these processes have been found not to be rate-limiting (14). Even if cell death and lysis were an important factor in the overall kinetics of anaerobic degradation, these processes are not applicable in the case of treatment of raw domestic sewage, and therefore hydrolysis may be considered as the only possible rate-limiting step in this case (15).

Some factors affecting the rate and extent of the anaerobic degradation of organic matter are temperature, pH, particle deposition, internal mixing in the reactor, and presence of toxic compounds. Further details on the effect of these factors is beyond the scope of this work and can be found in Heertjes and van der Meer (16), Rittmann and Baskin (17), Bouwer (18), Bohle (19), and van Haandel and Lettinga (1), among others.

ANAEROBIC SEWAGE TREATMENT

Anaerobic processes have been used to treat domestic and industrial wastewater for more than a century (20,21). The septic tank is the oldest and most widely used example. The striking success of the UASB reactor (Fig. 1) made anaerobic systems come back to the forefront, and they are now used for the treatment of several types of wastewater.

The UASB concept is based on the establishment of a dense sludge bed in the bottom of the reactor, in which all biological processes take place. This sludge bed is formed by accumulation of incoming suspended solids and by bacterial growth. In upflow anaerobic systems, and under certain conditions, it was also observed that bacteria could naturally aggregate in flocs and granules (22). These dense aggregates have good settling properties and are not susceptible to washout from the system under normal reactor conditions. Retention of granular or flocculent active sludge within the UASB reactor enables good treatment performance at high organic loading rates. Natural turbulence caused by the influent flow and the biogas production provides good wastewater biomass contact in UASB systems. Several configurations can be imagined for a wastewater treatment plant including a UASB reactor. In any case, a sand trap, screens for coarse material, and drying beds for the sludge must

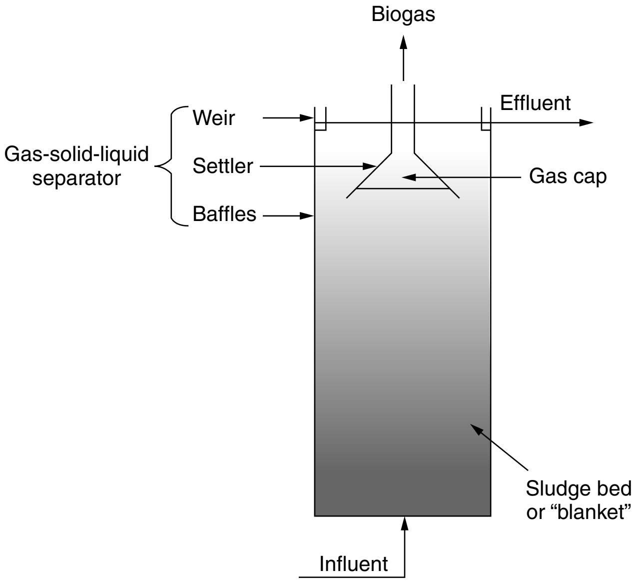

Figure 1. Basic scheme of a UASB reactor for sewage treatment.

exist. The UASB reactor may replace the primary settler, the anaerobic sludge digester, the aerobic step, and the secondary settler of conventional aerobic treatment plants like activated sludge or trickling filters.

UASB-based anaerobic treatment systems are recognized to be highly efficient, simple to construct and operate, and flexible to be applied on either a very large or a very small scale. UASB reactors have low space requirements and consume negligible amounts of energy, if any. Moreover, energy is produced during the process in the form of methane. The production of sludge is very low in comparison with aerobic methods, because of the slow growth rates of anaerobic bacteria. Besides, the sludge is well stabilized for final disposal, and it can be preserved for considerable periods of time without a significant reduction of its activity, allowing its use for the startup of new reactors. The process requires no additional nutrients, and an adequate and stable pH can be maintained without the addition of any chemicals. Anaerobic systems are well suitable for campaign industries because the adapted anaerobic sludge can be stored unfed within the reactor for entire seasons. The use of anaerobic systems can lead to a high degree of self-sufficiency.

Some disadvantages of anaerobic sewage treatment have been put forward, like low pathogen and nutrient removal. In fact, pathogens are only partially removed, except helminth eggs, which are effectively captured in the sludge bed. The removal of nutrients is also not complete and a post-treatment is generally required if the nutrients cannot be reused for agriculture or aquaculture within a reasonable distance from the treatment plant. The startup of anaerobic reactors usually takes longer compared with aerobic processes, unless good inoculum is available to speed up this stage. Hydrogen sulfide is produced during the anaerobic process, especially when high concentrations of sulfate exist in the influent. A proper handling of the biogas is then required to avoid bad smell. A significant proportion of the total amount of methane produced by the reactor is actually dissolved in the effluent. Its recovery may be required to minimize smell nuisances and methane emissions to the atmosphere. Substantial improvements have been made in tackling most of these alleged disadvantages (23).

Anaerobic treatment was considered economically more attractive and conceptually more holistic than aerobic treatment (24). The costs of aeration and sludge handling, the two largest costs associated with aerobic sewage treatment, would be reduced dramatically because no oxygen is needed in the process and the production of sludge is much smaller than in aerobic treatment (17). Moreover, the sludge (biomass) produced in aerobic processes has to be stabilized before it can be safely disposed of, which adds to the complexity of aerobic treatment plants. Anaerobic treatment can make a substantial contribution to the protection of the environment and the preservation of natural resources as part of a sustainable and appropriate wastewater treatment system for developing countries.

EXAMPLES OF SEWAGE TREATMENT IN ANAEROBIC REACTORS

UASB Reactors

The full-scale application of UASB reactors to domestic wastewater has been a success in tropical areas with sewage temperature above 25 °C (25). Studies in this field started many years ago in Cali, Colombia, where UASB reactors were found to be cheaper than facultative ponds and oxidation ditches, especially when capital costs were included (26). The efficiency of UASB reactors for sewage treatment was found to be similar in tropical and subtropical regions (sewage temperatures around 18–23 °C) using the same design parameters (9). The application of UASB reactors to sewage treatment under even lower temperature conditions has been studied in the Netherlands since 1976 (27,28). Promising results were obtained in Jordan, with UASB reactors treating strong raw sewage at a temperature of 18 °C in winter and 25 °C in summer (8). Encouraging results have also been obtained in the Mediterranean area (29). Although substantial experience on the design and operation of UASB reactors for sewage treatment has been gathered lately, most of the results have not been published (30). Under specific conditions, the use of two-stage anaerobic reactors (8,31–33), additional heated sludge digesters (7), or chemically enhanced primary sedimentation (34) have been proposed to further improve the performance of anaerobic systems.

Other Technologies

Technologies other than the UASB, like the anaerobic filter (35), traditional anaerobic digesters currently in use in China, the anaerobic attached film expanded bed (AAFEB) system (36), plug-flow reactors, and modified anaerobic baffled reactors (37), among others, have also been used for the anaerobic treatment of sewage.

Post-Treatment

Anaerobic sewage treatment systems generally fail to comply with COD discharge standards, as established by Council Directive 91/271/EEC on Urban Waste Water Treatment, dictated by the European Union Council of Ministers (38) (125 mgCOD/L), or the guideline proposed by the World Health Organization (39) for unrestricted irrigation (less than 1000 fecal coliform per 100 mL and less than 1 helminth egg per L). Therefore, a post-treatment step is mandatory in most cases to remove remnant COD, fecal coliform (as an indicator of pathogenic micro-organisms), helminth eggs, and even nitrogen and phosphorus when direct reuse is not feasible. WSP are among the most efficient and cost-effective post-treatment methods available (40). Other post-treatment methods are the biorotor system, or rotating biological contactor (41), integrated duckweed and stabilization pond system (42), trickling filters (43), and activated sludge (44), among others. See the review presented by Tawfik (41).

On-Site Systems

Application of modified UASB reactors for single households in isolated locations, like farms and recreational facilities not connected to the centralized sewerage system, was studied under different temperature conditions in the Netherlands (45) and Indonesia (46). These systems were called UASB septic tanks, because they shared features of both methods. Sludge gradually accumulates in the reactor, as in septic tanks, but they are operated in upflow mode, as UASB reactors. The design is almost as simple as that of conventional septic tanks, but the treatment efficiency is much higher (47). Suggestions for improving the treatment efficiency in these systems include the use of two- or three-stage UASB reactors and the adoption of post-treatment methods, like small aerobic lagoons (46). The UASB process was also applied to treat sewage from small-size communities in Brazil (48). The potential of anaerobic technology in so-called DESAR (Decentralized Sanitation And Reuse) schemes is huge, and more efforts should be directed to test these systems in the field (23).

WHAT IS THE FUTURE OF ANAEROBIC SEWAGE TREATMENT?

Anaerobic reactors should be more widely applied in tropical countries as the main pretreatment step for sewage. Conditions for anaerobic digestion are so good under these conditions that the application of any other system to treat the organic pollution in sewage is highly questionable on technical, environmental, economic, and social grounds. The application of UASB reactors for sewage treatment could be safely expanded to subtropical regions, and clear indications exist that even in temperate regions its use should not be discarded altogether. Anaerobic systems should also be used in DESAR schemes where the preservation of water, nutrients, and energy is indispensable. Anaerobic systems can significantly contribute to the achievement of more sustainable sanitation in most developing countries, and their potential should be recognized and explored by universities, private companies, and governments.

BIBLIOGRAPHY

1. van Haandel, A.C. and Lettinga, G. (1994). *Anaerobic Sewage Treatment. A Practical Guide for Regions with a Hot Climate*. John Wiley, Chichester, UK.
2. Gijzen, H.J. (2002). Anaerobic digestion for sustainable development: a natural approach. *Water Sci. Technol.* **45**(10): 321–328.
3. United Nations. (2003). *Water for People-Water for Life*. The United Nations World Water Development Report, UNESCO Publishing, Berghahn Books, Paris, France.
4. WSSCC (Water Supply and Sanitation Collaborative Council). (2003). *A guide to investigating one of the biggest scandals of the last 50 years*, Third World Water Forum, Kyoto, Japan.
5. Grau, P. (1996). Integrated water and waste management. *Water Sci. Technol.* **33**(8): 39–46.
6. Lettinga, G., van Nelsen, A.F.M., Hobma, S.W., de Zeeuw, W., and Klapwijk, A. (1980). Use of the upflow sludge blanket (USB) reactor concept for biological wastewater treatment, especially for anaerobic treatment. *Biotechnol. Bioengineer.* **22**: 699–734.
7. Mahmoud, N.J.A. (2002). *Anaerobic pre-treatment of sewage under low temperature (15 °C) conditions in an integrated UASB-digester system*, Ph.D. thesis, Wageningen University, Wageningen, the Netherlands.
8. Halalsheh, M.M. (2002). *Anaerobic pre-treatment of strong sewage*, Ph.D. thesis, Wageningen University, Wageningen, the Netherlands.
9. Seghezzo, L. (2004). *Anaerobic treatment of domestic wastewater in subtropical regions*, Ph.D. thesis, Wageningen University, Wageningen, the Netherlands.
10. Gujer, W. and Zehnder, A.J.B. (1983). Conversion processes in anaerobic digestion. *Water Sci. Technol.* **15**: 127–167.
11. Pavlostathis, S.G. and Giraldo-Gomez, E. (1991). Kinetics of anaerobic treatment: a critical review. *Crit. Rev. Environ. Control*. **21**(5,6): 411–490.
12. de Baere, L.A., Rozzi, A., and Verstraete, W. (1984). Solubilization of particulate organic matter as the rate-limiting step in anaerobic digestion loading. *Trib. Cebedeau.* **37**(484): 75–81.
13. Batstone, D.J., Keller, J., Angelidaki, I., Kalyuzhyni, S.V., Pavlostathis, S.G., Rozzi, A., Sanders, W.T.M., Siegrist, H., and Vavilin, V.A. (2002). Anaerobic digestion model No.1 (ADM1), Scientific and Technical Report No.13, IWA Task Group for Mathematical Modelling of Anaerobic Digestion Processes, IWA Publishing, London, UK.
14. Pavlostathis, S.G. and Gossett, J.M. (1985). A kinetic model for anaerobic digestion of biological sludge. *Biotechnol. Bioengineer.* **28**: 1519–1530.
15. Sanders, W.T.M. (2001). *Anaerobic hydrolysis during digestion of complex substrates*, Ph.D. thesis, Wageningen University, Wageningen, the Netherlands.
16. Heertjes, P.M. and van der Meer, R.R. (1978). Dynamics of liquid flow in an up-flow reactor used for anaerobic treatment of wastewater. *Biotechnol. Bioengineer.* **20**: 1577–1594.
17. Rittmann, B.E. and Baskin, D.E. (1985). Theoretical and modelling aspects of anaerobic treatment of sewage, In: *Proceedings of the Seminar/Workshop on Anaerobic Treatment of Sewage*. M.S. Switzenbaum (Ed.). Amherst, MA, pp. 55–94.
18. Bouwer, E.J. (1987). Theoretical investigation of particle deposition in biofilm systems. *Water Research*. **21**(12): 1489–1498.
19. Bohle, A.G. (1993). Performance of static flocculators. *Water Sci. Technol.* **27**(10): 181–194.

20. McCarty, P.L. (1981). One hundred years of anaerobic treatment. In: *Anaerobic Digestion*. D.E. Hughes et al. (Eds.). Elsevier, Amsterdam, p. 3.
21. McCarty, P.L. (2001). The development of anaerobic treatment and its future. *Water Sci. Technol.* **44**(8): 149–156.
22. Hulshoff Pol, L.W., de Zeeuw, W.J., Velzeboer, C.T.M., and Lettinga, G. (1983). Granulation in UASB-reactors. *Water Sci. Technol.* **15**: 291–304.
23. Lettinga, G. (2001). Digestion and degradation, air for life. *Water Sci. Technol.* **44**(8): 157–176.
24. Mergaert, K., Vanderhaegen, B., and Verstraete, W. (1992). Applicability and trends of anaerobic pre-treatment of municipal wastewater. *Water Research*. **26**(8): 1025–1033.
25. Alaerts, G.J., Veenstra, S., Bentvelsen, M., and van Duijl, L.A. (1993). Feasibility of anaerobic sewage treatment in sanitation strategies in developing countries. *Water Sci. Technol.* **27**(1): 179–186.
26. Schellinkhout, A., Lettinga, G., van Velsen, L., and Louwe Kooijmans, J. (1985). The application of the UASB-reactor for the direct treatment of domestic wastewater under tropical conditions, In: *Proceedings of the Seminar/Workshop on Anaerobic Treatment of Sewage*. M.S. Switzenbaum (Ed.). Amherst, MA, pp. 259–276.
27. Lettinga, G., Roersma, R., Grin, P., de Zeeuw, W., Hulshof Pol, L., van Velsen, L., Hobma, S., and Zeeman, G. (1981). Anaerobic treatment of sewage and low strength waste wasters, In: *Proceedings of the 2nd International Symposium on Anaerobic Digestion*. D.E. Hughes et al. (Eds.). Travemünde, Germany, pp. 271–291.
28. Grin, P., Roersma, R., and Lettinga, G. (1985). Anaerobic treatment of raw domestic sewage in UASB reactors at temperatures from 9–20 °C, In: *Proceedings of the Seminar/Workshop on Anaerobic Treatment of Sewage*. M.S. Switzenbaum (Ed.). Amherst, MA, pp. 109–124.
29. Collivignarelli, C., Urbini, G., Farneti, A., Bassetti, A., and Barbaresi, U. (1991). Economic removal of organic and nutrient substances from municipal wastewaters with full-scale UASB fluidized- and fixed-bed reactors. *Water Sci. Technol.* **24**(7): 90–95.
30. Wiegant, W.M. (2001). Experiences and potential of anaerobic wastewater treatment in tropical regions. *Water Sci. Technol.* **44**(8): 107–113.
31. Wang, K. (1994). *Integrated anaerobic and aerobic treatment of sewage*, Ph.D. thesis, Wageningen Agricultural University, Wageningen, the Netherlands.
32. Elmitwalli, T.A. (2000). *Anaerobic treatment of domestic sewage at low temperature*, Ph.D. thesis, Wageningen University, the Netherlands.
33. Mgana, S.M. (2003). *Towards sustainable and robust on-site domestic wastewater treatment for all citizens*, Ph.D. thesis, Wageningen University, Wageningen, the Netherlands.
34. Kalogo, Y. and Verstraete, W. (2000). Technical feasibility of the treatment of domestic wastewater by a CEPS-UASB system. *Environ. Technol.* **21**: 55–65.
35. Young, J.C. and McCarty, P.L. (1969). The anaerobic filter for waste treatment. *J. Water Pollut. Control Fed.* **41**: R160.
36. Jewell, W.J., Switzenbaum, M.S., and Morris, J.W. (1981). Municipal wastewater treatment with the anaerobic attached microbial film expanded bed process. *J. Water Pollut. Control Fed.* **53**(4): 482–490.
37. Yu, H. and Anderson, G.K. (1996). Performance of a combined anaerobic reactor for municipal wastewater treatment at ambient temperature. *Resources Conservation Recycling*. **17**: 259–271.
38. European Union Council of Ministers. (1991). *Official J. Eur. Commission*. **L135**: 40–52.
39. WHO. (1989). *Health Guidelines for the Use of Wastewater in Agriculture and Aquaculture*. Technical Report Series No 778, WHO, Geneva, Switzerland.
40. Cavalcanti, P.F.F. (2003). *Integrated application of the UASB reactor and ponds for domestic sewage treatment in tropical regions*, Ph.D. thesis, Wageningen University, Wageningen, the Netherlands.
41. Tawfik, A. (2002). *The biorotor system for post-treatment of anaerobically treated domestic sewage*, Ph.D. thesis, Wageningen University, Wageningen, the Netherlands.
42. van der Steen, P., Brenner, A., van Buuren, J., and Oron, G. (1999). Post-treatment of UASB reactor effluent in an integrated duckweed and stabilization pond system. *Water Research*. **33**(3): 615–620.
43. Chernicharo, C.A.de L. and Nascimento, M.C.P. (2001). Feasibility of a pilot-scale UASB/trickling filter system for domestic sewage treatment. *Water Sci. Technol.* **44**(4): 221–228.
44. von Sperling, M. (1996). Comparison among the most frequently used systems for wastewater treatment in developing countries. *Water Sci. Technol.* **33**(3): 59–72.
45. Bogte, J.J., Breure, A.M., van Andel, J.G., and Lettinga, G. (1993). Anaerobic treatment of domestic wastewater in small scale UASB reactors. *Water Sci. Technol.* **27**(9): 75–82.
46. Lettinga, G., de Man, A., van der Last, A.R.M., Wiegant, W., van Knippenberg, K., Frijns, J., and van Buuren, J.C.L. (1993). Anaerobic treatment of domestic sewage and wastewater. *Water Sci. Technol.* **27**(9): 67–73.
47. Zeeman, G. (1997). *Decentralized Sanitation Systems*. The role of anaerobic treatment in UASB-septic-tank systems, Paper presented to the Third Japan-Netherlands Workshop on Integrated Water Quality Management, the Netherlands, May 26–30, p. 6.
48. Vieira, S.M.M., Carvalho, J.L., Barijan, F.P.O., and Rech, C.M. (1994). Application of the UASB technology for sewage treatment in a small community at Sumare, São Paulo State. *Water Sci. Technol.* **30**(12): 203–210.

PERSISTENCE OF PATHOGENS IN WATER

ABID M. NASSER
Water Quality Research
Laboratory, Ministry of Health
Tel-Aviv, Israel

WATERBORNE PATHOGENS SIGNIFICANT TO HEALTH

Natural water sources become contaminated with pathogenic micro-organisms by the discharge of inadequately treated wastewater. Four major pathogen groups can be found in wastewater: over 100 types of enteric viruses, 7 types of pathogenic bacteria, 10 types of protozoan parasites, and 6 types of helminthes of health significance to humans and animals. The concentration of pathogenic micro-organisms in wastewater is affected mainly by the amount of precipitation, the season of the year, and the socioeconomic status of the community. For most pathogenic micro-organisms, the methods of detection are expensive, time-consuming, and difficult to perform. Therefore, data on the prevalence of pathogenic micro-organisms in wastewater effluents and natural water sources is limited to types for which a detection technique was developed. The recently developed molecular biology-based techniques enhanced our capability in detecting human pathogenic micro-organisms in water sources.

The viruses that are detected most often in polluted water are enteroviruses, because of the availability of tissue culture for virus cultivation and enumeration. In wastewater, researchers reported on the presence of 1000 pfu/L enteroviruses (1). Discharge of inadequately treated wastewater may result in the viral contamination of natural water sources. Enteric viruses have been detected in rivers, lakes, and even in groundwater. Using reverse transcription-PCR (RT-PCR), it was reported that 24 out of 48 well water samples (50%) were found positive for enteric viruses, including enteroviruses, rotaviruses, hepatitis A virus, adenoviruses, and noroviruses, and 10 out of 12 river water samples (83%) were found positive for viruses by RT-PCR (2).

Giardia cysts and *Cryptosporidium* Oocysts were detected in 100% raw wastewater samples at a concentration of 10^3 cysts/L and 5×10^2 oocysts/L, respectively (3). Most conventional wastewater treatment processes result in the reduction of parasites in the treated effluent; however, the remaining low numbers may pose a health risk. *Cryptosporidium* was detected in 50% of effluent samples of oxidation ponds. Mayer and Plamer (4) reported 99.9% removal for *Giardia* and 99% for *Cryptosporidium* by biological treatment of wastewater. *Cryptosporidium* and *Giardia* were removed by 90% and 99%, respectively, by activated sludge treatment. Despite difficulties in detecting procedures, the number of Cryptosporidium oocysts detected ranges from 0.005 to 252.7 oocysts/L of surface waters in different countries of the world. In a study conducted in Japan, it was found that oocysts of *Cryptosporidium* increased in numbers from late summer to early autumn, and the maximum number of oocysts detected in three rivers was 3.3 to 5 oocysts/L (5). The reported frequencies of occurrence of contamination of surface water with *Giardia* and *Cryptosporidium* are from 60–96% in the United States and from 20–64% in Canada. Their levels in surface waters were found to be very low, ranging from 0.5 to 5,000 oocysts in 100 liters of water. Analysis of 139 surface water samples in Finland revealed the presence of *Giardia spp.* in 19 (13.7%), *Cryptosporidium* oocysts in 14 (10.1%), and Noroviruses in 13 (9.4%) of the samples (6). The pathogens were isolated less frequently during the winter than during the summer months. Based on risk analysis, it has been proposed that the action levels to prevent outbreaks with these protozoa should occur when concentration in 100 liters water sampled are >5 *Giardia* cysts and 10 to 30 oocysts of *Cryptosporidium* (7).

PERSISTENCE OF PATHOGENS IN WATER

Most of the persistence data for micro-organisms was generated in laboratory experiments. Micro-organisms are seeded into water of certain quality and incubated at test temperature. The concentration of micro-organisms in the suspension is measured as a function of time. To investigate the effect of environmental factors under field conditions on the survival of micro-organisms, investigators have used environmental chambers. The chambers are of small volume that contain the micro-organisms in the matrix, but allow limited exchange with the surrounding environment by means of permeable membrane. Using sentinel chambers, Keswick and Gerba (8) reported the survival of micro-organisms ranked in increasing order were *E.coli*, fecal streptococcus, f2 bacteriophages, SA-11 rotavirus, echovirus 1, coxsackievirus B3, and poliovirus 1. Enteric viruses and bacteria were reported to survive longer in groundwater than in surface water, which was attributed to a number of factors, including lower temperatures, protection from sunlight, and limited microbial antagonisms.

Temperature is the most important factor that influences pathogenic micro-organisms persistence in the natural water and in the environmental setting in general (9). Inactivation rates increase with temperature, and temperature can therefore be a useful predictor of pathogen die-off in the environment (10). Moisture content and loss of moisture are also related to pathogen survival (10). In general, the survival of many micro-organisms is prolonged at lower temperatures, whereas at elevated temperatures, inactivation or die-off is fairly rapid. It is important to examine more than one stressor because, under natural environmental settings, multiple factors may influence the persistence of pathogens. A study was conducted in our laboratory to examine the impact of temperature, salinity, and organic material load on the persistence of *E.coli*, Coxsackievirus A9, and *Cryptosporidium* parvum. At 30 °C, the fastest die-off was observed for Cox A9 virus, where die-off of 5 orders of magnitude was recorded regardless of the water quality. At 15 °C, Cox A9 persistence was similar to that of *Cryptosporidium*, and no change was observed in the concentration of both organisms after 30 days incubation. Whereas a die-off of 5 orders of magnitude was recorded for *E.coli*. The die-off of *E.coli* was faster than that of the other tested micro-organisms regardless of the water quality In a previous study, we compared the persistence of *E.coli*, F+coliphages, poliovirus 1, and hepatitis A virus in phosphate buffered saline (PBS), groundwater, and raw wastewater. In this study, it was determined that micro-organism type and temperature are the most important factors influencing the persistence of micro-organisms in natural waters. Indigenous F+ bacteriophages persisted longest followed by Poliovirus 1, HAV, and the bacterial indicator *E.coli*. The enhanced inactivation of viruses at higher temperatures may be attributed to greater microbial activity in the natural waters. The microbial activity of *P.aeroginosa* has been shown to accelerate the die-off of enteric viruses and other pathogenic micro-organisms. At lower temperatures, where the microbial activity is low, *Cryptosporidium* oocysts' concentration was reduced by only one order of magnitude after 160 days of incubation. Microbial activity has been shown to enhance the die-off of HAV in wastewater, and the time needed for the inactivation of 90% of the virus at 25 °C was 8.3 days as compared with 15.1 days in autoclaved waste (11).

Short-wave ultraviolet (UV) components of sunlight are lethal to enteric pathogens. UV is used for disinfection in water and wastewater treatment processes. The site of the UV damage is the nucleic acid, where the damage inhibits the multiplication capability of the micro-organisms. UV damage may be reduced by the DNA repair mechanisms present in some cells. In marine water, the inactivation rate of *Cryptosporidium* oocysts, poliovirus 1, *Giardia* cysts, and salmonella increased when exposed to natural sunlight, with oocysts exhibiting the highest persistence Sattar et al. (12) demonstrated a detrimental effect on oocyst survival of exposure to sunlight and UV light.

Most of the data generated for the prevalence of the protozoan parasites *Cryptosporidium* and *Giardia* in natural water sources is based on microscopic determination of their presence and viability. As the concentration of the parasites in natural water is low, there were no attempts reported to correlate the viability and the infectious state of naturally occurring parasites in natural waters. Although laboratory experiments have shown that the protozoan parasites may persist for longer periods of time in natural waters, their infectivity potential should be determined in order to more accurately assess the health risk posed by these parasites in natural waters.

PREDICTING THE PERSISTENCE OF PATHOGENS IN WATER

Data on the persistence of pathogens in water generated in laboratory and field experiments can be used to establish models to predict the time needed to inactivate pathogens under various natural sittings. The die-off of sensitive micro-organisms can be described by linear regression equation

$$Y = b + aX$$

where Y represents the response, b represents the y-axis intercept, X represents the independent variable (time), and a is the slope. The value of the slope a is influenced by factors affecting the persistence of pathogens in water. For example, at 15 °C, the slopes recorded for *E.coli*, Coxsackievirus A9, and *Cryptosporidium* in stream water were -0.14, -0.06, and -0.01, respectively. At 30 °C, the die-off slopes for *E.coli* and *Cryptosporidium* did not change; however, that of coxsackie A9 increased to -0.14. The effect of part of the environmental variables on the survival of micro-organisms was found to be significant. The effect of temperature on the persistence of pathogenic micro-organisms was found significant in numerous studies (9). Therefore, the effect of temperature can be predicted by the linear regression equation. On the other hand, water quality effect was not found to be significant. Furthermore, the interaction of water quality with temperature was not found to be significant. The effect of micro-organism type on the persistence in water is significant, and also the interaction of micro-organism type and temperature was found to be significant (9). It is possible to classify micro-organisms into three groups according to their capacity to persist in waters and, consequently, use a multiple linear regression equation in which a slope value for each independent variable will be assigned (10).

CONCLUSIONS

1. Pathogenic micro-organisms may be introduced to natural water sources and may be present at concentrations that may pose health risks.
2. At low temperatures, viruses and protozoa parasites may persist for prolonged periods of time in natural water sources, which will enable their transmission to people consuming the contaminated waters.
3. Most studies in which the persistence of fecal coliforms was compared with those of pathogenic viruses and protozoan parasites revealed that the bacterial indicator is more sensitive than the pathogens and, therefore, it dose not accurately represent the health risks posed by these pathogens in natural waters.

BIBLIOGRAPHY

1. Ferguson, C., Husman, A.M., Altvilla, N., Deere, D., and Ashbolt, N. (2003). *Crit. Rev. Environ. Sci. Technol.* **33**: 299–361.
2. Borchardt, M.A., Haas, N.L., and Hunt, R.J. (2004). *Appl. Environ. Microbiol.* **70**: 5937–5946.
3. Nasser, A.M., Molgen, S., and Rubin, K. (1998). *Prevalence of Cryptosporidium and Giardia in Waste and Suface Water in Israel*. OECD Workshop Interlaken 98 on Molecular Technologies for safe drinking water, Interlaken, Switzerland.
4. Mayer and Plamer. (1996).
5. Tsushima, X. et al. (2002). *J. Vet. Med.* **65**: 121–123.
6. Horman, A. et al. (2004). *Appl. Environ. Microbiol.* **70**: 87–95.
7. Hass, C.N. and Rose, J.B. (1995). *J. Am. Water Works Assoc.* **87**: 81–84.
8. Keswick, B.H. and Gerba, C.P. (1980). *Environ. Sci. Technol.* **14**: 1290.
9. Nasser, A.M. and Oman, S. (1999). *Water Res.* **33**: 1748–1752.
10. Hurst, J.C. (1991). *Modeling the Environmental Fate of Microorganisms*. C. Hurst (Ed.). ASM, Washington, DC, pp. 137–159.
11. Deng, Y.M. and Cliver, D.O. (1995). *Appl. Environ. Microbiol.* **61**: 87–91.
12. Sattar, S.A., Chauret, C., Springthorpe, V.S., Battigelli, D.A., Abbaszadegan, M., and LeChevallier, M. (1999). *AwwARF and Ontario Ministry of Environment and Energy*. Denver, CO and Toronto, ON.

READING LIST

Rose, J.B. (1988). *J. Am. Water Works Assoc.* **80**: 53–58.

Stezenbach, L.D., Arrowood, J.M., Mershell, M.M., and Sterling, C.R. (1988). *Wat. Sci. Technol.* **20**: 193–198.

LeChevallier, M.W., Norton, W.D., and Lee, R.G. (1991). *Appl. Environ. Mocrobiol.* **57**: 2610–2616.

Ongerth, J.E., Hunter, G.D., and DeWalle, F.B. (1995). *Water Res.* **29**: 1295–1299.

Nasser, A.M. (1994). *CRC Crit. Rev. Environ. Sci. Technol.* **24**(4): 281–323.

WELL HEAD PROTECTION

Babalola Makinde-Odusola
Public Utilities
Riverside, California

Groundwater, especially karst aquifers (1) or groundwater under the direct influence (2) of surface water, is vulnerable to contamination from possible contaminating activities (PCAs) (3) related to land use. A wellhead protection (WHP) plan (WHPP) essentially includes the integrated proactive management practice(s) that protect the quality of groundwater that is a source of drinking water. WHP is part of an integrated (from source to consumer's tap) multibarrier system to reduce risks to public health from drinking water. WHP fits into the World Health Organization's (WHO's) Water Safety Plan that recommends the use of hazard analysis and critical control points (HACCPs) as a risk management tool in drinking water operations (4). In the United States, many public water systems (PWSs) are required to conduct source water assessment (SWA) (3,5–7). For PWSs that source from groundwater, a WHPP can be part of their SWA program (SWAP). A PWS must develop its WHPP in accordance with guidelines established by regulatory agencies (1–3).

An important step in WHP is the delineation of one or more WHP zones (WHPZs) to form a WHP area (WHPA) that includes surface and subsurface zone(s) that reasonably contribute water to the well(s) that are being protected from potential source(s) of contamination. Methods and criteria for delineating protection zones range from simple geometric shapes to complex numerical groundwater models based on the specified time of travel (TOT) of contaminant (1,8–11). The most appropriate criterion and method for delineating the WHP zones are selected after careful assessment of many factors, including regulatory acceptance, susceptibility of the aquifer(s) to contamination, available hydrogeological data, computational and technical resources, and the hydrogeologic setting of the aquifer. "Derived/estimated" protection zones must be adjusted to meet the required minimal distances mandated by regulations. Delineation of a WHPA by itself does not protect groundwater quality. Delineation must be followed with the implementation of appropriate management strategies to prevent groundwater contamination and periodic evaluation of their effectiveness. Using a computer-based geographical information system (GIS) to map the protection zones is recommended (12). Cooperation among appropriate jurisdictional agencies and stakeholders is beneficial and may increase the effectiveness of the WHPP.

INTRODUCTION

Groundwater is vulnerable to contamination from many PCAs such as leaching effluents from septic and underground storage tanks (USTs) or leaching pesticides and/or fertilizers. Emerging major concerns are xenobiotics (13) and organic wastewater compounds (14) in effluents from wastewater treatment plants and/or confined animal feeding operations (CAFOs) that may impair the beneficial use and/or environmental value of the groundwater. Xenobiotics include pharmaceuticals, hormones, and other compounds found in medicines and other personal care products. The U.S. Environmental Protection Agency (EPA) reported that over 200 different chemicals have been detected in groundwater in 38 states (15).

The goal of WHP is to manage PCAs to reduce, or preferably prevent, potential impairment of groundwater to assure public health. The U.S. EPA (16) estimated that, between 1971 and 1996, contaminated source waters (surface and groundwater) were the cause of 86% of waterborne disease outbreaks within the United States. The protection of surface water may often be linked to the protection of groundwater and vice versa. The failure of a pumping station from a lightning strike led to an outbreak of cryptosporidiosis in Brushy Creek, Texas, following the accidental discharge of raw sewage into a creek 2 mi upstream of the 120- and 200-ft deep municipal wells (17).

The U.S. EPA (16) observed that deep wells are not immune from biological contamination. Lawson et al. (18) reported an outbreak of gastroenteritis caused by Norwalk virus in water produced from a 600-ft deep well. In the aftermath of the Walkerton incident, Health Canada (19) identified WHP as one of the protective measures that should be implemented and maintained "if source water is subject to human fecal contamination or if human enteric viruses have been responsible for past waterborne outbreaks." The U.S. EPA (16) proposed the Groundwater Rule (GWR) to address risks of consuming waterborne pathogens in groundwater.

Many states require source water assessment before permitting new drinking water sources (such as wells) for PWSs (3). For proposed PWS wells, WHP begins with the proper siting and construction of well(s) in accordance with good sanitary practices and recommended well design standards (20). A successful WHPP protects public health and avoids the expense of treating polluted water or drilling replacement wells (7). There are additional benefits from implementing WHP (21–23). The U.S. EPA identified over 90 measures as possibilities for measuring success of source water contamination prevention at the national level (24).

The United Kingdom Margate Act (11) of 1902 and Brighton Act of 1924 are examples of earlier efforts directed toward the protection of groundwater produced from wells used for drinking. The Margate Act empowered a water authority to control drains, cesspools, and so on over an area of 1500 yards from any well (11). In 1986, the U.S. Congress introduced a Wellhead Protection Program (WHPP) as part of the Amendments to the Safe Drinking Water Act (SDWA) (25). A wellhead protection area (WHPA) as defined by that Act as "the surface and subsurface area surrounding a water well or well field, supplying a public water system through which contaminants are reasonably likely to move toward and reach such well or well field." The U.S. EPA (16) sometimes refers to a WHPA as "groundwater protection area." The California Department of Pesticide Regulation (DPR) defines a groundwater protection area (GWPA) as "an area of land that is vulnerable to the movement of pesticides to ground water according to either leaching or runoff processes" (26).

Amendments to the SDWA in 1996 required states to establish source water assessment programs (SWAPs) (6). Sourcewater protection is the generic term for describing management activities to protect the quality of water used for drinking water irrespective of the nature of the source: surface water, groundwater, or groundwater under the influence of surface water. Source water protection is the leading first barrier of the multibarrier approach to protecting the quality of drinking water delivered to consumers. Other elements of the multibarrier protection framework against pollution are source water treatment (including disinfection), distribution system integrity (including cross-connection control programs), and public information. The Minnesota Department of Health (7) differentiates between "source water assessment" plan and a "wellhead protection plan" as two different documents, with separate and distinct purposes. A PWS that sources from groundwater can use the completed SWA plan in developing a WHPP (7).

DEVELOPING A WHPP FOR A PWS

Many countries have developed regulations governing the practice of WHP (2,9–11). In the United States, the SDWA mandated states to develop WHP programs that must include some specified elements (9,22). The U.S. EPA must approve state WHP programs before implementation. As of 2004, the U.S. EPA has approved WHP programs for all states except Virginia (27). WHP programs developed by states provide regulatory guidance to PWSs on how to develop appropriate WHP plans (3,7).

Typical steps for developing a WHP plan include the following: (1) defining the statement and purpose of the WHP; (2) defining the members, roles, and duties of the team; (3) delineating the WHP zones; (4) inventorying the PCAs; (5) assessing the vulnerability; (6) developing management programs; (7) developing contingency plans for alternative water supplies; and (8) developing an optional source water protection program (9,22,28). Those steps are mainly sequential, but some steps can also be iterative.

Many PWSs have developed WHP plans, some of which included stakeholder participation. Some examples are available on the Internet (29,30). The U.S. EPA (31,32) maintains a directory of useful tools for developing a WHPP. Witten et al. (28) identified some financing strategies for wellhead protection.

DELINEATION CRITERIA AND METHODS

An important step is the delineation of one or more protection zones to which WHP management measures would be tailored. Five major criteria that may be used singly or in combination to delineate a protection zone are distance, drawdown, time of travel (TOT), assimilative capacity, and flow boundaries (8,9). Factors that determine the preferred delineation criterion/criteria include local regulatory requirements, number of wells involved, complexity of aquifer(s), spatial extent, and available resources and professional skills. The delineation methods can be deterministic or stochastic. Kunstmann and Kinzelbach (33) describe stochastic methods for delineating a WHPA.

The distance criterion is the simplest and least expensive. It is based on a fixed perimeter radius around the well professionally "judged" adequate to provide reasonable protection. The drawdown criterion is based on the zone of influence (ZOI) where groundwater level declines due to pumping from the well. The TOT criterion requires the delineation of isochrones (i.e., contours of equal time) representing the duration it takes for water to reach a well or a contaminant to reach a well in concentration exceeding a target level. The assimilative capacity criterion permits reduction in the spatial extent of the zone of contribution (ZOC) of the WHPA when contaminants can be attenuated to concentrations lower than the target level before reaching the well through biogeochemical processes such as dilution, dispersion, sorption, or biodegradation. The flow boundary criterion uses physical and hydrogeologic features that control groundwater flow to define the spatial extent, that is, the ZOC of the WHPA.

The U.S. EPA (9) categorized the methods for delineating into four major groups of increasing complexity: geometric, simple analytic (SA), hydrogeologic mapping (HM), and computer modeling (CM). Geometric methods involve the use of a predetermined fixed radius or simplified shapes that have been precalculated for a range of pumping and aquifer conditions. Examples are the arbitrary fixed radius (AFR), calculated fixed radius (CFR), and modified CFR.

AFR involves drawing a circle of defined distance around the well on a base map. The defined distance (radius) is based on professional judgment and regulatory acceptance. CFR is similar to the AFR but the radius is calculated based on the volume of cylinder with pore volume equal to the volume of water pumped during the specified period.

SA methods use appropriate analytical solutions to groundwater flow, such as the uniform flow equation, and known aquifer characteristics to derive protection zones. HM involves the use of a combination of hydrological, geological, and geochemical data and geologic mapping techniques to delineate the area contributing water to the well(s). HM methods are appropriate in settings where geologic features greatly influence groundwater flow such as in karst aquifers (15,34). CM methods numerically simulate groundwater flow and contaminant transport to calculate rate and direction of flow, travel times, and so on. The model developed may be used to compare the system's response to proposed management options (15).

The U.S. EPA (9) lists the advantages and disadvantages of each delineation method. The number of person-hours required to delineate a WHPA per well range from 1–5 for AFR to 10–200+ for numerical modeling (8). Figures illustrating the delineation methods are in given in Reference 9. Some states prefer specific methods for delineating WHPAs for special wells. California (13) prefers detailed hydrogeologic analyses when delineating WHPAs for nonvertical wells (e.g., horizontal wells). California (13) prefers using the delineation method that utilizes the most available detailed information. A combination of methods can be used to delineate one or more zones. For example, the innermost zone to the well may be based on a fairly simple geometry while the outermost protection zone may be delineated using complex numerical models.

WHP ZONES

A WHPA consists of one or more WHPZs, to which appropriate management measures can be tailored. The innermost WHPZ is usually delineated to protect against biological contaminants and to protect the integrity of the well. Some outermost zones regarded as buffer zones may be optional. In California (3), the suggested approach is to define four zones and an optional buffer zone. The required zones in California are (Fig. 1):

Well Site Control Zone (WSCZ). The closest zone to the wellhead is the area immediately surrounding the well. The WSCZ is managed to prevent seepage of contaminants, vandalism, or tampering. The recommended minimum radius is 50 ft.

Zone A2 or the Microbial/Direct Chemical Contamination Zone. This is the surface area above the aquifer that contributes water to well(s) within a 2-yr TOT. This zone was established to comply with the requirement of the proposed Groundwater Rule (8) based on research that bacteria and viruses are not likely to survive beyond 2 yr in soil and groundwater. Recommended minimum radii

$$R_t = \sqrt{(70,267Qt/\eta\pi H)}$$

where:

R_t = radius in feet of protection zone for time t
Q = Average pumping rate in gallons per minute (gpm)
t = Time of travel (TOT) to well in years
π = 3.1416
η = effective porosity of aquifer
H = open interval or length of well screen (feet)

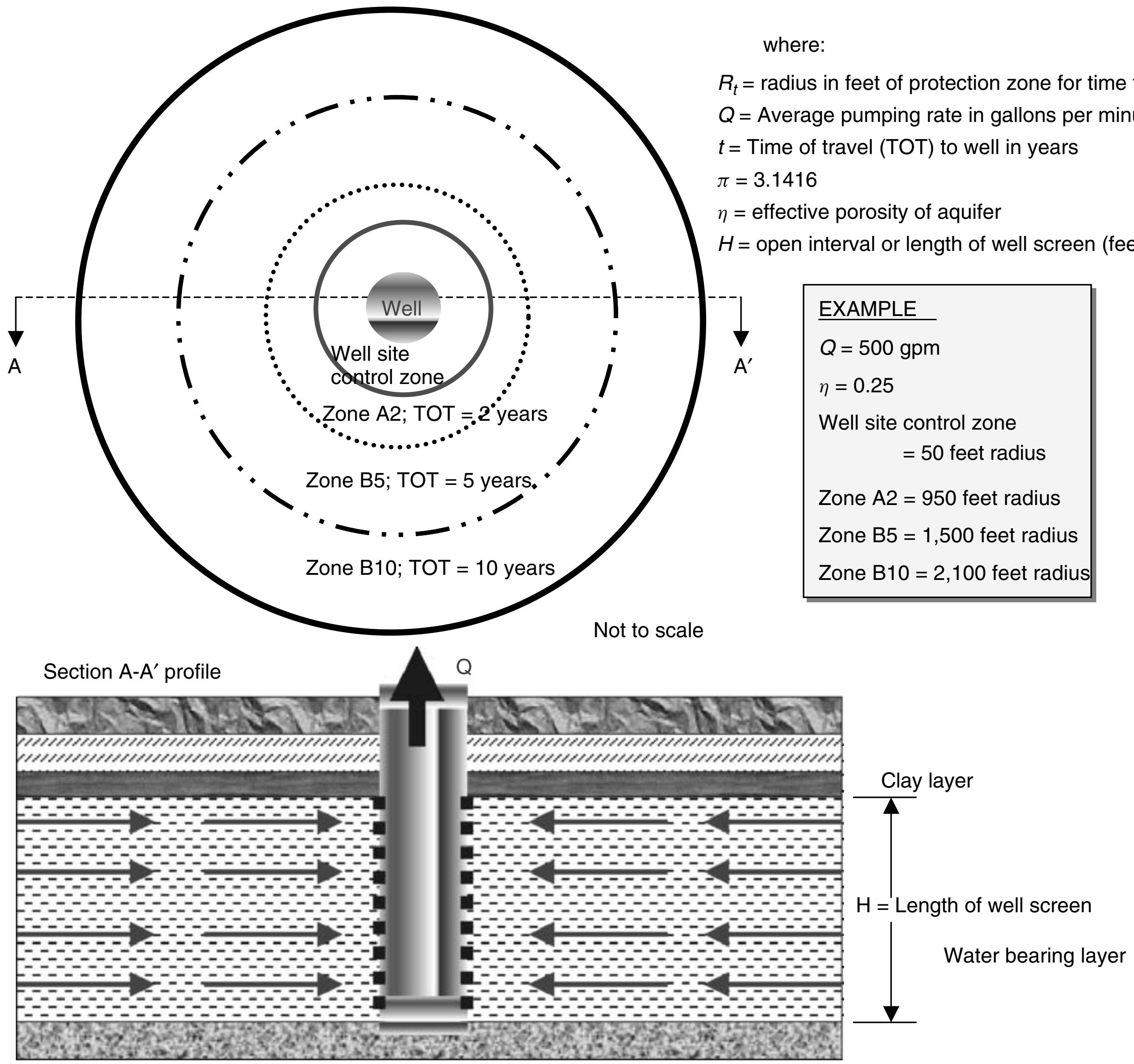

EXAMPLE
Q = 500 gpm
η = 0.25
Well site control zone
= 50 feet radius
Zone A2 = 950 feet radius
Zone B5 = 1,500 feet radius
Zone B10 = 2,100 feet radius

Figure 1. Wellhead protection zones delineated by calculated fixed radius (CFR) method.

are 600 ft and 900 ft for porous aquifer (PA) and fractured rock aquifer (FRA), respectively.

Zone B5. This chemical contamination zone is that surface area above the aquifer between the 2- and 5-yr TOT. This zone provides additional response time to manage a chemical spill. Recommended minimum radii are 1000 ft and 1500 ft for PA and FRA, respectively.

Zone B10. This chemical contamination zone is that surface area above the aquifer between the 5- and 10-yr TOT. This zone allows for some natural attenuation of contamination and development of mitigation measures. Recommended minimum radii are 1500 ft and 2250 ft for PA and FRA, respectively.

Buffer Zone (Optional). This zone is generally beyond Zone B10. The zone offers a higher level of protection and may be extended to include the entire recharge area, especially where there are potential sources of significant contamination such as landfills or other hazardous materials.

The delineated zones can be refined in shape and/or size based on professional judgment and/or local knowledge of some site-specific hydrogeology. In California, the final assessment map must be based on a USGS quadrangle 7.5-minute series topographic map (3).

INVENTORY OF PCAS

The purpose is to identify all possible contaminating activities (PCAs) by location within the WHPA. Typical resources used in developing the inventory of PCAs include land use maps, business license records, hazardous waste databases, and the Internet (35,36). A field reconnaissance can identify other potential sources of contamination.

The comprehensive inventorying of PCAs can be in phases with initial focus on sources that could pose the greatest risk. The inventory should include past and present sources of contamination and might require geophysical investigations for inactive sources such as abandoned wells and waste disposal sites. The information gathered should include the type of source, potential contaminants from each source, and location of source. The information gathered must be updated periodically. Many PWSs use global positioning systems (GPSs) to locate PCAs and GIS system to map the locations. The Groundwater Foundation (37) has developed a primer on using such technologies to develop an inventory of PCAs.

VULNERABILITY ASSESSMENT (VA) OR SUSCEPTIBILITY ASSESSMENT

The U.S. EPA (38) defines groundwater "vulnerability" as "the relative ease with which a contaminant introduced into the environment can migrate to an aquifer under a given set of management practices, contaminant properties, and aquifer hydrogeologic characteristics." Groundwater vulnerability can be specific (specified contaminant) or intrinsic (any contamination in general) and must consider long-term effects (39). The purpose of VA in WHP planning is to evaluate and rank the risks of the identified PCAs that pose the most significant threats to groundwater quality within the WHPA. VA under WHP should not be confused with the VA required under the U.S. Public Health Security and Bioterrorism Preparedness and Response Act of 2002 (PL 107–188). The Bioterrorism Act addresses the vulnerability of the entire water system (not just source waters) to intentional acts (e.g., terrorism) intended to substantially disrupt the ability of the system to provide a safe and reliable supply of drinking water. The U.S. EPA (16) proposed a groundwater rule that requires hydrogeologic assessments to identify wells sensitive to fecal contamination.

There are many methods for evaluating and mapping groundwater vulnerability to contamination ranging from a simple process to detailed numerical modeling of fate and transport of chemicals of concern (9,26,39–42). VA must factor the type and proximity of the PCA and site-specific hydrogeological information such as the presence of karst formations and any physical barrier that may affect the fate and transport of the PCA (3). DRASTIC [Depth to water table, net Recharge, Aquifer media, Soil media, Topography (slope), Impact to vadose zone, and hydraulic Conductivity of the aquifer] is a widely used method for evaluating relative vulnerability to groundwater contamination (9,41). Many VA models have technical limitations. Focazio et al. (40) illustrated the use of scientifically reliable tools for groundwater VA.

MANAGEMENT MEASURES (MMS)

Among the factors that influence the choice of MMs are availability of alternate sources of water, types of contaminants, hydrogeologic setting, and community and regulatory acceptance. MMs must be tailored to be most appropriate for each WHPZ. Land uses in the zone closest to the well are usually managed to avoid all possible risks, including those from bacteria, viruses, and chemical spills (2).

WHP MMs may be regulatory or nonregulatory. Nonregulatory MMs include public education to increase community awareness of the WHPP, economic incentives, land acquisition (43), and conservation easements. Regulatory MMs include local ordinances and zoning/land use planning. Additional examples of MMs are provided by the U.S. EPA (8,9,15). It is preferable that a WHP plan, that includes land use controls, use scientifically and legally defensible criteria to delineate the protection zones (40). Where feasible, installation of groundwater quality monitoring well(s) between the significant potential sources of contamination and the wellhead and periodic sampling could provide an early warning detection system.

CONTINGENCY PLANNING

A contingency plan (CP) should be developed to deal with emergency threats to groundwater quality and natural disasters such as floods. It is advisable to coordinate with the local emergency planning committee. The following elements are typically included in the CP: basic water system information, WHPA boundaries, a list of potential contaminant sources and locations, short-term and long-term water supply options, water rationing plans, and emergency response plans (ERPs). The development of a CP can be coordinated with other efforts such as that of the ERP required under the Bioterrorism Act. The U.S. EPA (44) maintains a list of ERP guidance documents and planning tools.

CONCLUSION

Wellhead protection planning is now one of the major programs within the SDWA related to protecting the quality of groundwater. The other programs are groundwater rule, sole source aquifer, source water assessment, underground injection control (UIC), source water petition, and comprehensive groundwater protection grants. The U.S. EPA is strengthening the linkages between the Clean Water Act (CWA) and the SDWA. The U.S. EPA is coordinating the source water protection program in each state with federal and state underground storage tank programs using GIS to map underground storage tank sites within source protection zones of drinking water sources.

Wellhead protection zones and plans should be periodically reviewed and updated. Regular sanitary surveys and vulnerability assessments could reveal operational deficiencies especially after special events such as floods. Curriero et al. (45) found that over half of the waterborne disease outbreaks in the United States in the past 50 years were preceded by heavy rainfall. Outbreaks due to groundwater contamination, which accounted for approximately 36% of all outbreaks, were more associated with extreme precipitation occurring within a 3-month lag preceding the outbreaks (45).

BIBLIOGRAPHY

1. U.S. EPA. (2002). *Delineation of Source-Water Protection Areas in Karst Aquifers of the Ridge and Valley and Appalachian Plateaus Physiographic Provinces: Rules of Thumb for Estimating the Capture Zones of Springs and Wells*. EPA 816-R-02-015. U.S. EPA Office of Water, Washington, DC.
2. Ontario Ministry of the Environment. (2001). Protocol for delineation of wellhead protection areas for municipal groundwater supply wells under direct influence of surface water, Ontario, Canada. Available at http://www.ene.gov.on.ca/envision/techdocs/index.htm. Accessed November 2004.
3. California Department of Health Services (DHS), Division of Drinking Water and Environmental Management. (2000). *Drinking Water Source Assessment and Protection (DWSAP) Program*. California Department of Health Services, Division of Drinking Water and Environmental Management, Sacramento, CA. Available at http://www.dhs.ca.gov/ps/ddwem/dwsap/DWSAP_document.pdf. Accessed November 2004.
4. http://www.who.int/water_sanitation_health/dwq/gdwq3/en/. Accessed November 2004.
5. U.S. EPA. (1997). *State Source Water Assessment and Protection Programs Guidance*. Final Guidance EPA-816-R-97-009. U.S. EPA Office of Water, Washington, DC.
6. http://www.epa.gov/safewater/protect/swap.html. Accessed November 2004.
7. http://www.health.state.mn.us/divs/eh/water/swp/swa/index.htm. Accessed November 2004.
8. U.S. EPA. (1987). *Guidelines for Delineation of Wellhead Protection Areas*. EPA/440/6-87-010. Office of Ground-Water Protection, Washington, DC.
9. U.S. EPA. (1994). *Handbook Ground Water and Wellhead Protection*. EPA/625/R-94-001. Office of Research and Development, Washington, DC.
10. Australian and New Zealand Environment and Conservation Council. (1995). *Guidelines for Groundwater Protection in Australia*. Canberra, ACT 2600.
11. National Groundwater and Contaminated Land Centre. *Groundwater Source Protection Zones*. Environment Agency, Olton, Solihull, UK. Available at http://www.environment-agency.gov.uk/commondata/105385/gspz_on_web.pdf. Accessed November 2004.
12. Bice, L.A., Van Remortel, R.D., Mata, N.J., and Ahmed, R.H. (2000). *Source Water Assessment Using Geographic Information Systems*. National Risk Management Research Laboratory, Office of Research and Development, U.S. EPA, Cincinnati, OH. Available at http://www.epa.gov/ORD/NRMRL/wswrd/SWAPJ18.PDF. Accessed November 2004.
13. Ongerth, J.E. and Khan, S. (2004). Drug residuals: how xenobiotics can affect water supply sources. *J. AWWA* **94**(5): 94–101.
14. Lee, K.E. et al. (2004). *Presence and Distribution of Organic Wastewater Compounds in Wastewater, Surface, Ground, and Drinking Waters, Minnesota, 2000–02*. U.S. Geological Survey Scientific Investigation Report 2004–5138.
15. U.S. EPA. (1993). *Wellhead Protection: A Guide for Small Communities*. EPA/625/R-93/002. Office of Research and Development, Washington, DC.
16. U.S. EPA. (2000). National primary drinking water regulations: groundwater rule; Proposed Rule 40 CFR Parts 141 and 142. *Fed. Reg.* **65**(91): 30194–30274.
17. http://www.epa.gov/safewater/ndwac/ndwac051701day1.html. Accessed November 2004.
18. Lawson, H.W. et al. (1991). Waterborne outbreak of Norwalk virus gastroenteritis at a southwest U.S. resort: role of geological formations in contamination of well water. *Lancet* **337**: 1200–1204.
19. Health Canada. (2003). *Virological Quality of Drinking Water*. Document for Public Comment—prepared by the Federal-Provincial-Territorial Committee on Drinking Water. Available at http://www.hc-sc.gc.ca/hecs-sesc/water/publications/virological_quality_dw/toc.htm.
20. California Department of Water Resources. (1990). *California Well Standards*. Bulletin 74–90. Department of Water Resources, Sacramento, CA.
21. Fenske, B.A. and Williams, M.B. (2004). *Demonstrating Benefits of Wellhead Protection Programs*. AWWA Research Foundation, Denver, CO.
22. Makinde-Odusola, B. (2003). Wellhead protection. In: *Encyclopedia of Water Science*. X. Stewart and X. Howell (Eds.). Marcel Dekker, New York, pp. 1024–1028. Available at http://www.dekker.com. Accessed November 2004.
23. http://www.health.state.mn.us/divs/eh/water/swp/whp/benefits.html. Accessed November 2004.
24. http://www.epa.gov/safewater/protect/strategyupdate.html. Accessed November 2004.

25. American Water Works Association (AWWA). (2001). *Safe Drinking Water Advisor: A Compliance Assistance Resource*. AWWA, CD-ROM Edition.
26. http://www.cdpr.ca.gov/docs/gwp/gwpamaps.htm. Accessed November 2004.
27. http://www.epa.gov/safewater/whpnp.html. Accessed November 2004.
28. Witten, X. et al. (1995). *A Guide to Wellhead Protection*. Planning Advisory Service Report No. 457/458. American Planning Association, Chicago, IL.
29. http://www.ci.redmond.wa.us/insidecityhall/publicworks/environment/grhistory.asp.
30. http://www.gem.msu.edu/casestd/casestd.html. Accessed November 2004.
31. http://www.epa.gov/safewater/protect/sources.html. Accessed November 2004.
32. http://www.epa.gov/safewater/protect/feddata/delineation.html. Accessed November 2004.
33. Kunstmann, H. and Kinzelbach, W. (2000). Computation of stochastic wellhead protection zones by combining the first-order second-moment method and Kolmogorov backward equation analysis. *J. Hydrol.* **237**: 127–146.
34. Bachman, S. et al. (1997). *California Groundwater Management*. Groundwater Resources Association of California, Sacramento, CA.
35. http://www.epa.gov/safewater/protect/feddata/inventory.html.
36. Stevens, K.B. et al. (2003). *Chemical Occurrence Data Sets for Source Water Assessment*. AWWA Research Foundation, Denver, CO.
37. The Groundwater Foundation. (2003). *Using Technology to Conduct a Contaminant Source Inventory—A Primer for Small Communities*. Lincoln, NE. Available at http://www.groundwater.org/ProgEvent/ACTT_Primer1202.pdf. Accessed November 2004.
38. U.S. EPA. (2004). *Handbook of Groundwater Protection and Cleanup Policies for RCRA Corrective Action*. EPA530-R-04-030. Office of Solid Waste and Emergency Response, Washington, DC.
39. National Research Council. (1993). *Groundwater Vulnerability Assessment, Contamination Potential Under Conditions of Uncertainty*. National Academy Press, Washington, DC. Available at http://books.nap.edu/books/0309047994/html. Accessed November 2004.
40. Focazio, M.J. et al. (2002). *Assessing Ground-Water Vulnerability to Contamination: Providing Scientifically Defensible Information for Decision Makers*. USGS Circular 1224.
41. Rupert, M.G. (2001). Calibration of the DRASTIC groundwater vulnerability mapping method. *Groundwater* **39**(4): 625–630.
42. U.S. EPA. (1993). *A Review of Methods for Assessing Aquifer Sensitivity and Groundwater Vulnerability to Pesticide Contamination*. EPA-813-R-93002.
43. Stone, D. and Schindel, G.M. (2002). The application of GIS in support of land acquisition for the protection of sensitive groundwater recharge properties in the Edwards aquifer of south-central Texas. *J. Cave Karst Studies* **64**(1): 38–44.
44. http://cfpub.epa.gov/safewater/watersecurity/home.cfm?program_id=8. Accessed November 2004.
45. Curriero, F.C. et al. (2001). The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948–1994. *Am. J. Public Health* **91**(8): 1194–1199.

CHEMICAL DRINKING WATER STANDARDS, PAST, PRESENT, AND FUTURE

Patrick Sullivan
Franklin Agardy
Forensic Management Association
San Mateo, California

Water resources in the United States have already been shown to contain the chemical byproducts of water treatment and disinfection, industrial chemicals (both organic an inorganic), human drugs (prescription and nonprescription), veterinary and human antibiotics, pharmaceuticals (e.g., from the clinical treatment of cancer using chemotherapy drugs), sex and steroidal hormones, antioxidants, detergents and detergent metabolites, fire retardants, plasticizers, polycyclic aromatic hydrocarbons, petroleum products, and a wide range of pesticides that are either regulated, anticipated to be regulated, or unregulated. As any given water resource will contain a mixture of individual chemical compounds, scientists, engineers, and environmental policy professionals need to ask the question, can existing drinking water standards, that are based on individual chemical-specific water quality criteria, be shown to protect the public health? In attempting to answer this question, we hope to put into context the validity of chemical drinking water standards in a world where the boundaries between water and wastewater are becoming increasingly blurred.

CHEMICAL POLLUTANTS OF THE PAST

Prior to the 1940s and the mass production of synthetic organic chemicals, the vast majority of industries produced chemicals that were dominated by inorganic compounds or biodegradable organic compounds, for example, canneries, dairies and feedlots, distilleries, fertilizers, foundries (iron and steel), metal plating, mining and smelting, pulp and paper, paint products, meat packing, tanneries, textiles, and timber products. In addition, a growing number of petroleum-based compounds were also being produced and used in coal-derived chemicals manufacturing, manufactured gas production, petroleum refining, and wood treating. As a result, common water pollutants included inorganic acids (chloride and sulfate), arsenic, copper, cadmium, chromium, cyanide, lead, nickel, nitrate, and zinc as well as phenols and a wide spectrum of petroleum hydrocarbons.

Some of these common pollutants were contained in the 1925 standards for drinking water (1). The 1925 standards state that, with respect to chemical and physical characteristics, for water to be suitable for drinking and culinary purposes, it should be free from toxic salts and should not contain an excessive amount of soluble mineral substances or any chemicals employed in treatment. In addition, when any of the following compounds exceed their standards, the water should be rejected for use:

Constituent	Concentration (mg/L)
Copper	0.2
Iron	0.3
Lead	0.1
Magnesium	100
Zinc	5.0
Chloride	250
Sulphate	250
Total Solids	1,000

By 1946 (2), grounds for rejection of the water supply was expanded and based on the following standards:

Constituent	Concentration (mg/L)
Arsenic	0.05
Chromium (6+)	0.05
Lead	0.1
Fluoride	1.5
Selenium	0.5

whereas the following chemical substances should preferably not occur in excess of the following standards:

Constituent	Concentration (mg/L)
Copper	3.0
Iron	0.3
Magnesium	125
Zinc	15
Chloride	250
Sulphate	250
Phenolic	0.001

Phenolic compounds (commonly associated with wastes produced by manufactured gas production, wood treating, and synthetic chemical manufacturing) were the first organic chemicals added to the drinking water standards. However, by the 1950s, the high-volume use of the synthetic organic surfactant, alkyl benzene sulfonate, in detergents and the widespread distribution of highly toxic halogenated pesticides, shown to be responsible for fish kills around the United States, resulted in the addition of more organic compounds to the 1962 water quality standards.

Based on the 1962 (3) standards, grounds for rejection of the water supply should occur when the following standards are exceeded:

Constituent	Concentration (mg/L)
Arsenic	0.01
Barium	1.0
Cadmium	0.01
Chromium (6+)	0.05
Cyanide	0.2
Lead	0.05
Selenium	0.01
Silver	0.05
Fluoride	0.6–1.7 (based on air temperature)
Copper	1.0

whereas the following chemical substances should not occur in excess of the following standards:

Constituent	Concentration (mg/L)
Alkyl Benzene Sulfonate	0.5
Carbon Chloroform Extract	0.2[1]
Phenols	0.001
Iron	0.3
Manganese	0.05
Zinc	5.0
Cyanide	0.01
Nitrate	45
Chloride	250
Sulphate	250
Total Dissolved Solids	500

As a result of the continued and extensive use of halogenated and nonhalogenated hydrocarbons since the 1940s and the recognition of halogenated disinfection byproducts in treated drinking water, the current drinking water standards now include 27 inorganic criteria and 63 organic criteria. The current National Primary Drinking Water Standards for chemicals are:

Contaminant	Maximum Contaminant Level(MCL) (Micrograms per liter)[2]
Inorganic Compounds	
Antimony	6.0
Arsenic	50.0
Barium	2000.0
Beryllium	4.0
Bromate	10.0
Cadmium	5.0
Chlorine gas (as Cl_2)	4000.0
Chlorine dioxide (as ClO_2)	800.0
Chlorite	1000.0
Chromium (total)	100.0
Copper	1300.0
Cyanide	200.0
Fluoride	4000.0
Lead	15.0
Mercury	2.0
Nitrate	10000.0
Nitrite	1000.0
Selenium	50.0
Thallium	2.0
Organic Compounds	
Acrylamide	1000.0
Alachlor	2.0
Atrazine	3.0
Benzene	5.0
Benzo(a)pyrene	0.2
Carbofuran	40.0
Carbon tetrachloride	5.0
Chloramines (as Cl_2)	4000.0
Chlordane	2.0
Chlorobenzene	100.0

2,4-D	70.0
Dalapon	200.0
DBCP	0.2
O-Dichlorobenzene	600.0
p-Dichlorobenzene	75.0
1,2-Dichloroethane	5.0
1,1-Dichloroethylene	7.0
cis-1,2-Dichloroethylene	70.0
trans-1,2-Dichloroethylene	100.0
Dichloromethane	5.0
1,2-Dichloropropane	5.0
Di(2-ethylhexyl)adipate	400.0
Di(2-ethylhexyl)phthalate	6.0
Dinoseb	7.0
Dioxin	0.00003
Diaquat	20.0
Endothall	100.0
Endrin	2.0
Epichlorohydrin	20000.0
Ethylbenzene	700.0
Ethylene dibromide	0.05
Glyphosate	700.0
Haloacetic acids (HAA5)[3]	60.0
Heptachlor	0.4
Heptachlor epoxide	0.2
Hexachlorobenzene	1.0
Hexachlorocyclopentadiene	50.0
Lindane	0.2
Methoxychlor	40.0
Oxamyl(Vydate)	200.0
Polychlorinated biphenyls	0.5
Pentachlorophenol	1.0
Picloram	500.0
Simazine	4.0
Styrene	100.0
Tetrachloroethylene	5.0
Toluene	1000.0
Total Trihalomethanes (TTHM)[4]	100.0
Toxaphene	0.003
2,4,5-TP (Silvex)	50.0
1,2,4-Trichlorobenzene	70.0
1,1,1-Trichloroethane	200.0
1,1,2-Trichloroethane	5.0
Trichloroethylene	5.0
Vinyl chloride	2.0
Xylenes (total)	10000.0

[1] Measures extractable organic residues (i.e., pesticides).

[2] The maximum level of a contaminant in drinking water at which no known or anticipated adverse effect on the health effect of persons would occur, and which allows for an adequate margin of safety.

[3] The HAA5 compounds are dibromoacetic acid, dichloroacetic acid, monobromacetic acid, monochloroacetic acid, and trichloroacetic acid.

[4] The TTHM compounds are bromodichloromethane, bromoform, chloroform, and dibromochloromethane.

The National Secondary Drinking Water standards are nonenforceable guidelines regulating contaminants that may cause cosmetic effects or aesthetic effects in drinking water. The United States Environmental Protection Agency (USEPA) recommends secondary standards for water systems but does not require systems to comply. However, individual states may choose to adopt them as enforceable standards.

Contaminant	Secondary Standard (Micrograms per liter)
Aluminum	0.05
Chloride	250
Iron	0.3
Manganese	0.05
Silver	0.1
Sulfate	250
Total Dissolved Solids	500
Zinc	5.0

All of these pollutants are associated with industrial and chemical use practices evolving over the last 100 years.

CHEMICAL POLLUTANTS OF THE PRESENT AND FUTURE

The National Research Council estimates that there are approximately 72,000 organic chemicals in commerce within the United States, with nearly 2000 new chemicals being added each year (4). Given this number of chemicals, it would seem impossible that the USEPA can identify every chemical that poses a hazard to the public health. In fact, it is impossible. For example, under the Toxic Substances Control Act that was established in 1979, the USEPA has conducted an assessment program to determine which new chemicals present an unreasonable risk to human health or the environment. Since 1979, the USEPA has only reviewed some 32,000 new chemical substances. With so many chemicals left to evaluate, the USEPA chose to exclude all chemicals that are produced in amounts less than 10,000 pounds per year as well as all polymers from further consideration. The remaining 15,000 chemical subset has been identified as being the broad focus of the USEPAs existing chemical testing and assessment program with the primary focus being on the 3000 high-production-volume chemicals that are produced/imported at levels above 1 million pounds per year (4).

Once toxicity data are collected from animal studies, the USEPA must still identify those chemicals that should be regulated in drinking water. As of 1998, USEPA had completed their evaluation of approximately 400 compounds for potential regulation in drinking water. Of these top 400, a list of 50 chemicals were selected for consideration as future drinking water standards. The 50 chemicals that make up the 1998 Drinking Water Contaminant Candidate List are:

Acetochlor
Alachlor ESA (and acetanilide degradation products)
Aldrin
Aluminum
Boron

Bromobenzene
DCPA nonacid degradate
DCPA diacid degradate
DDE
Diazinon
1,1-dichloroethane
1,1-dichloropropene
1,2-diphenylhydrazine
1,3-dichloropropane
1,3-dichloropropene
2,2-dichloropropane
2,4-dichlorophenol
Dieldrin
2,4-dinitrophenol
2,4-dinitrotoluene
2,6-dinitrotoluene
Disulfoton
Diuron
DPTC (s-ethyl-dipropylthiocarbanate)
Fonofos
Hexachlorobtadiene
p-isopropyltoluene (p-cymene)
Linuron
Manganese
Methyl bromide
Methyl-t-butyl ether (MTBE)
2-Methyl-phenol (o-cresol)
Metolachlor
Metrobuzin
Molinate
Naphthalene
Nitrobenzene
Organotins
Perchlorate
Prometon
RDX
Sodium
Sulfate
1,1,2,2-tetrachloroethane
Terbacil
Terbufos
Triazines (and degradation products)
2,4,6-trichlorophenol
1,2,4-trimethylbenzene
Vanadium

A minimum of 5 compounds must be selected for regulation under the primary drinking water standards. The 2003 revised Drinking Water Contaminant Candidate List is now composed of the following 42 chemicals:

1,1,2,2-tetrachloroethane
1,2,4-trimethylbenzene
1,1-dichloroethane
1,1-dichloropropene
1,2-diphenylhydrazine
1,3-dichloropropane
1,3-dichloropropene
2,4,6-trichlorophenol
2,2-dichloropropane
2,4-dichlorophenol
2,4-dinitrophenol
2,4-dinitrotoluene
2,6-dinitrotoluene
2-methyl-Phenol (o-cresol)
Acetochlor
Alachlor ESA and other acetanilide pesticide degradation products
Aluminum
Boron
Bromobenzene
DCPA monoacid degradate
DCPA diacid degradate
DDE
Diazinon
Disulfoton
Diuron
EPTC (s-ethyl-dipropylthiocarbamate)
Fonofos
p-Isopropyltoluene (p-cymene)
Linuron
Methyl bromide
Methyl-t-butyl ether (MTBE)
Metolachlor
Molinate
Nitrobenzene
Organotins
Perchlorate
Prometon
RDX
Terbacil
Terbufos
Triazines and degradation products
Vanadium

As this list becomes more focused, a major problem still remains. Approximately seven years after the first Drinking Water Contaminant Candidate List was proposed, no new chemicals have been added, which is a disturbing result as well as a disturbing trend. From a scientific and practical point of view, the USEPA never has had the time or resources to comprehensively evaluate and regulate the entire list of 72,000 chemicals. However, to limit selection to a minimum of five chemicals is just absurd. It is extremely difficult to accept the premise that of all the chemicals used in this country there are only five additional chemicals need be added to the Candidate

List every few years that are a hazard to human health. Consequently, many hazardous chemicals that should be regulated on the basis of their potential toxicity to humans will go unregulated, and when these unregulated chemicals occur in drinking water, there will be no requirement to warn the consumer[5]. Furthermore, based on the USEPAs performance to-date and their pace of evaluating and regulating chemicals in drinking water, we might all be dead by the time the USEPA ponders the fate of its initial list of 50 chemicals. Such delays are inherent in current federal policy. For example, the Toxic Substances Control Act requires the USEPA demonstrate that a chemical is dangerous before it can take any action against that chemical (i.e., regulate that chemical). It is because of this type of policy that the presence of unregulated chemicals in our drinking water poses one of the greatest threats to America's public health. The current Drinking Water Contaminant Candidate List does not even consider a wide range of chemicals that are present in home care products (e.g., phthalates and nonylphenols) or drugs and pharmaceuticals (e.g., Ibuprofen, Clofibrate [cholesterol treatment], Estrogen, Bleomycin [cancer treatment]) that have already been identified in our water resources.

The fact that there are 1) whole new classes of chemicals that are not even evaluated by the USEPA as a potential threat to drinking water, 2) chemical mixtures in drinking water that are not regulated, and 3) approximately 2000 new chemicals produced and/or used in the United States each year, drinking water standards cannot be demonstrated to actually protect the public health.

Given this condition, it is our recommendation that water provided by either public- or privately-operated water utilities not solely rely on federal or state drinking water standards to protect their consumers. Water utilities should treat their raw water resource with the appropriate water treatment technologies (e.g., granulated activated carbon, ion exchange, reverse osmosis, advanced oxidation) to remove all potentially harmful inorganic and organic pollutants from their final product.

Based on our interviews with both public- and privately-operated water utilities, the cost to apply advanced water treatment technologies to remove the maximum amount of either inorganic or organic compounds would only require a 20% increase in the water utility customer's water bill. Given this level of available treatment, the only technical issue that remains is how to monitor for pollutant removal. Treatment efficiency can be determined by producing a chemical fingerprint of the raw water and the treated product, which can be accomplished by using standard liquid and gas chromatography methods, Raman spectroscopy methods, or the use of new ion diffusion spectroscopy.

Protection of our drinking water resources will become increasing difficult because (5) the boundaries between water and wastewater are already beginning to fade. Therefore, the only real future defense against the consumption of polluted drinking water is to abandon our reliance on water quality standards and implement appropriate water treatment technologies to minimize pollutants in our drinking water.

BIBLIOGRAPHY

1. Anonymous. (1925). Report of Advisory Committee on Official Water Standards. *Public Health Reports*. **40**(15).
2. Anonymous. (1946). Public Health Service Drinking Water Standards. *Public Health Reports*. **61**(11).
3. USPHS. (1962). Public Health Service Drinking Water Standards. *Public Health Service Publication*. **956**.
4. National Research Council. (1999). *Identifying Future Drinking Water Contaminants*. National Academy Press, Washington, DC.
5. Maxwell, S. (2001). Ten key trends and developments in the water industry. *J. Amer. Water Works Assoc.* **93**(4).

---

[5] Community water systems are only required to notify consumers if regulated pollutants exceed their established criteria.

INDUSTRIAL WATER

MAGNETIC WATER CONDITIONING

Ingrid Ritchie
Indiana University-Purdue University
Indianapolis, Indiana

INTRODUCTION

Hard water contains minerals such as calcium, magnesium, and other hard ions that can form scale on a variety of surfaces that come into contact with the water, especially when it is heated. Magnetic water conditioning is a nonchemical technology that can prevent the formation of scale in hard water.

There are two basic categories of magnetic water conditioning devices: nonpermanent magnetic units (electromagnetic and electrostatic units) (Fig. 1) and permanent magnet units (Fig. 2). Electromagnetic units generate a magnetic field through the use of electromagnets, in which an electric current passes through a wire. Electrostatic units also use electricity to impose an electric field on the water flow, which attracts or repels ions in the water and generates a magnetic field.

Permanent magnetic units, the focus here, utilize permanent magnets rather than electricity to generate the magnetic field. Most permanent magnetic units utilize standard two-pole magnets, but one manufacturer utilizes a unique multiple, reversing pole magnet (Fig. 2) (1). The magnetic field exerts a force on ions as they pass through the field, which changes the crystallization behavior of the ions and promotes bulk solution precipitation rather than the formation of an adherent scale. Unwanted particulates can be removed by filtration, water/solids separation, blowdown, bleed-off, or draining.

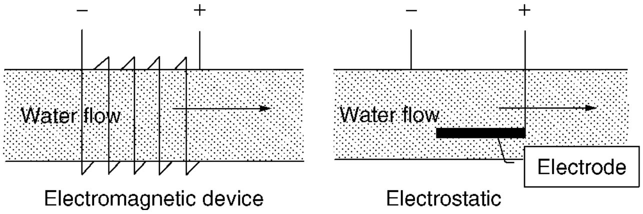

Figure 1. Diagram of classes of nonpermanent magnetic devices. (From Reference 1.)

Permanent magnetic units and electromagnetic units can be located so they are invasive (the magnet is located in line with the water flow) or noninvasive (the magnet is located so it is external to the water flow) as shown in Fig. 3. In-line systems are shielded magnets that are installed inside a section of pipe (see also Fig. 2). External units clamp onto the pipe through which the water flows.

THE PROBLEM OF HARD WATER AND SCALE

Hardness in water is due primarily to calcium and magnesium carbonates and bicarbonates (carbonate hardness that can be temporarily removed by heating) and calcium sulfate, calcium chloride, magnesium sulfate, and magnesium chloride (noncarbonated or permanent hardness that cannot be removed by heating). The sum of carbonate and noncarbonated hardness is total hardness, expressed as calcium carbonate. The World Health Organization (3) classifies hardness (in mg/L) as $CaCO_3$ as: 0–60, soft; 60–120, moderately hard; 121–180, hard; more than 180, very hard. Hard water, depending on the alkalinity, pH, temperature, and other factors, is responsible for the buildup of scale. Excessive hardness has a number of adverse effects in home, commercial, and industrial applications.

Hard water poses a significant problem in industrial boilers, cooling towers, and other heat transfer equipment. When hard water is heated, bicarbonates precipitate as carbonates and adhere to the pipe or vessel. In boiler and hot water tanks, the scale resulting from hardness reduces the thermal efficiency of heat transfer devices and eventually causes a restriction of the flow of water in piping systems. As a consequence, there is an increase in the system pressure that is required to move the water through a scale-restricted piping system. Both of these

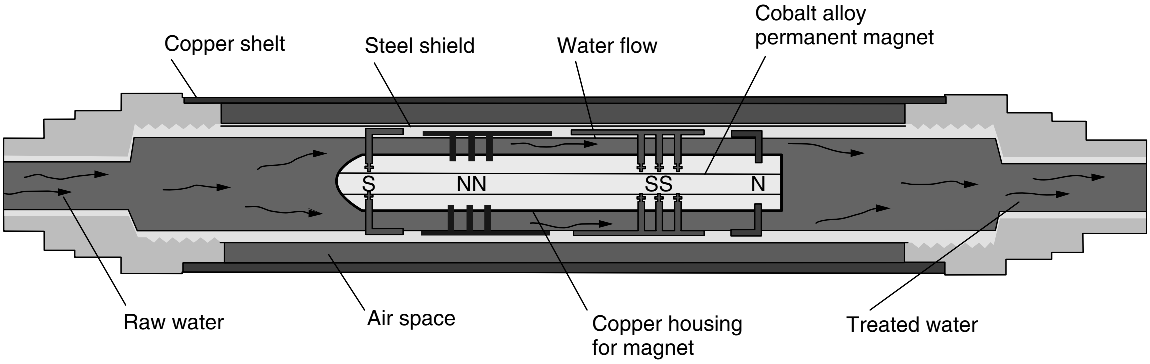

Figure 2. Diagram of a multipole permanent magnetic device. (From Reference 2. Available at: http://www.superiorwatercond.com/Products/our_superior_technology.html. Used with permission.)

problems increase the energy costs and maintenance costs of using and maintaining affected equipment. Table 1 provides example increases in energy consumption as a function of scale thickness.

The total cost of scaling in heat exchanger applications is estimated to be over $32 billion for the United States, Japan, German, the United Kingdom, Australia, and New Zealand combined, and over $14 billion in the United States alone (4).

In addition to increasing energy costs, hard water impairs the performance of detergents, which leads to dingy clothing and deterioration of fabrics and unsatisfactory cleaning of dishes and utensils in home, commercial, and industrial applications. Hard water results in spotting in car washes. Coffee makers, steamers, dishwashers, ice-makers, and other similar restaurant equipment are easily clogged by hard water and scale reduces the energy efficiency of these devices. In agricultural applications, hardness can reduce the efficiency of misting nozzles and irrigation systems.

APPLICATIONS OF PERMANENT MAGNETIC WATER CONDITIONING UNITS

According to the U.S. Department of Energy Technology Alert on nonchemical technologies for scale and hardness control, magnetic water conditioning "can be used as a replacement for most water-softening equipment," including lime or lime-soda softeners, ion exchange, and reverse osmosis (1). Magnetic water conditioning technology can also be used as a preconditioner to allow conventional devices to work more efficiently and last longer. The American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) identifies magnetic technology as a nonchemical treatment method that "has been used for scale control in boiler water, cooling water, and other process applications" (5).

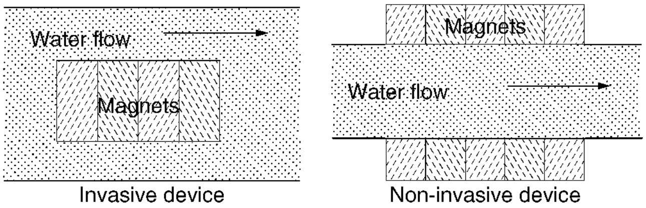

Figure 3. Diagram of invasive and noninvasive magnetic devices. (From Reference 1.)

Table 1. Example Increases in Energy Consumption as a Function of Scale Thickness

Scale Thickness, in.	Increased Energy Consumption, %
$\frac{1}{32}$	8.5
$\frac{1}{16}$	12.4
$\frac{1}{8}$	25.0
$\frac{1}{4}$	40.0

Source: Reference 1.

There are many benefits of this "green" technology over traditional technologies. The elimination or reduction of chemicals translates into dollar savings (which can reach thousands of dollars each year in some applications) when the costs of chemicals, labor, and equipment needed for chemical scale control are considered. Potential adverse health and environmental effects from the use of these chemicals are also eliminated or significantly reduced. Another advantage over chemical softening and ion exchange is that permanent magnetic water conditioning does not increase sodium concentrations in water and, thus, is a better option in those instances where sodium ions pose a health or other concern.

Periodic descaling of heat exchange equipment can be nearly eliminated, which, in turn, translates into savings because there is less downtime (as a consequence of scale formation), and the chemicals and labor needed for descaling are eliminated or significantly reduced. A related benefit is that heat exchanger tubes and related equipment failures caused by scale buildup are reduced, thus extending equipment life. In each of the applications in the section above on the problem of hard water and scale, permanent magnetic water conditioners can prevent the formation of scale and reduce existing scale. Because the technology keeps solids in solution, solids separation is needed in those applications where the process cannot tolerate particulate matter.

Even though applications are worldwide, magnetic water conditioning can be considered to be an emerging technology, whose future applications are likely to have significant impacts in colloidal chemistry and other fields. Any environmental, health, or industrial problem that is caused by hardness deposits could, in theory, be aided by magnetic treatment. A few examples to illustrate the diverse applications follow.

- In a dental study, patients who irrigated their teeth with an irrigator fitted with a magnetic device had 44% greater reduction ($p < 0.0005$) in calculus volume and a 42% greater reduction in calculus area ($p < 0.0001$) when compared to a control group that irrigated with conventional irrigation (without a magnetic device) (6).
- The Indianapolis 500 Brickyard Crossing Golf Course has realized a 71% reduction in the use of wetting agents and fertilizers and a related cost savings of 62% after the installation of the magnetic unit. The payback time for the magnetic unit at this facility, which does not pay for its water, is about 4.5 years based only on the wetting agent and fertilizer usage. The facility reports other benefits, including improved distribution and penetration of water, which substantially eliminates both wet and dry areas (7).
- In a case study, the DOE compared the costs and benefits of traditional lime softening and alternative permanent magnetic treatment for a recirculating boiler water system with a flow of 1000 gal/min at a hypothetical facility using extremely hard water (350 mg/L as CaCO3) (1). The magnetic technology had a simple payback of less than one year and

> produced an annual cost savings of over $200,000 per year over traditional lime softening. The DOE estimated a life-cycle savings of $2.77 million for the magnetic technology. Benefits assigned to the magnetic technology included energy savings and pollution reduction.

Permanent magnetic water conditioners are available in many sizes, ranging from small residential units (1 gal/min) to very large commercial and industrial units. The cost of magnetic water conditioning varies, depending on the amount of water to be treated. The costs are lower or in the same range as conventional water treatment systems; however, since the process is a physical rather than chemical process, there are no chemicals to buy and no on-going maintenance costs, which produces energy, labor, equipment, and chemical cost savings.

Magnetic technology, like other technologies, must be installed and used correctly. It is not appropriate in all applications. Additional guidelines developed by various manufacturers (1) include:

- *Installation*. Magnetic technology must be sized and installed correctly.
- *Water Quality*. Magnetic technology is not appropriate for soft water and some manufacturers report upper limits for hardness. Iron, which can cause fouling of the magnet, and silica, which can form scale with or without calcium, must be within specified limits.
- *Applications in Boilers and Heat Exchange Equipment*. Additional precautions must be taken with respect to alkalinity (to prevent corrosion), hydrogen sulfide (to prevent corrosion), and cycles of concentration.
- *Electrical Interference*. Magnetic units must be installed at an appropriate distance (at least 48 in.) from three-phase systems to avoid electrical interference, which causes magnetic lines of force to stray.

SCIENTIFIC SUPPORT FOR THE TECHNOLOGY

Historically, references to magnetic water conditioning date to the late 1800s, when lodestones (a naturally occurring permanent magnet) and naturally occurring magnetic mineral formations were used to decrease the formation of scale in cooking and laundry applications. From this modest beginning, magnetic water conditioning technologies have been and are being used throughout the world in a variety of applications to control scale. The acceptance of magnetic water conditioning, however, has not been universal, and there has been resistance to the technology in the United States, particularly by the water softening industry. Historically, the scientific literature has produced studies both positive and negative to magnetic water conditioning, and skeptics have centered their arguments against the technology largely on the lack of a consensus regarding the mechanism of action and lack of reproducibility in scientific studies. However, the more recent literature is increasingly supportive of the technology. Troup and Richardson (8) summarize earlier studies (pre-1978) for preventing scale formation and Baker and Judd (9) provide an informative review up to the mid-1990s. A few additional studies are noteworthy.

Barrett and Parsons (10) replicated the work of Higashitani and co-workers and found that magnetic treatment affects calcium carbonate formation under quiescent conditions, and the effect is reproducible. Both authors also determined that magnetic treatment produces a memory effect when the water is stored. Barrett and Parsons hypothesize that the formation of calcium carbonate is affected by the suppression of nucleation and acceleration in crystal growth.

Bogatin and others (11) propose that magnetic treatment of irrigation water increases the number of crystallization centers and the change of the free gas content, and that irrigation with magnetically treated water is the most effective for soils with high soda content. Other important parameters for effective treatment include flow rate through the apparatus, hardness, and pH.

Cho and co-workers (4) provide definitive evidence that physical water treatment (using a permanent magnet, solenoid coil, and high voltage electrode devices) is effective for the mitigation of fouling in cooling tower applications. (Fouling is contamination of a surface on which a liquid is boiling, that is, scale formation.) In this pivotal study, the authors demonstrated that multipole permanent magnets (with alternating fields) produced a significant reduction of 84% in fouling resistance compared to no treatment, and that a permanent magnet without alternating fields was not effective in reducing fouling. The authors also determined that previous studies showing no effect may have been limited because the treatment devices were not configured in an optimum way, giving support to manufacturers' instructions to install and operate the devices according to directions. The study proposed bulk precipitation as the mechanism of action. The magnetic field (or electric field, depending on the device) precipitates mineral ions in the water solution and forms clusters (i.e., colloidal particles of submicron size) in solution. As the temperature of the solution inside the heat transfer equipment increases, the solubility of the mineral ions decreases and the clusters grow in size and produce soft sludge coatings on the heat transfer surface rather than hard scale.

CONCLUSION

Permanent magnetic water conditioning technology is a physical process for reducing scale formation. It is a "green" technology that does not use chemicals or electricity to produce its effect. Because the magnetic unit does not have moving parts, it lasts for many years with very little, if any, maintenance.

These advantages enable permanent magnetic water conditioning to save energy, reduce the use of chemicals that treat and soften water, reduce the disposal of wastes from water treatment units such as water softeners, and conserve water. All of these benefits improve the environment and the health of our communities and reduce the costs of preventing and removing scale.

BIBLIOGRAPHY

1. U.S. Department of Energy. (1998). *Federal Technology Alert. Non-Chemical Technologies for Scale and Hardness Control*, DOE/EE-0162. U.S. Department of Energy, Pacific Northwest National Laboratory, Richland, WA.
2. Superior Manufacturing Division of Magnatech Corporation, Ft. Wayne, Indiana produces these units. The magnetic core is a single bar, multifield, multipole cobalt alloy permanent magnet.
3. *Guidelines for Drinking-Water Quality*. Vol. 2, p. 264. As cited in Salvato, J.A. (1992). *Environmental Engineering and Sanitation*, 4th Edn. Chap. 3, Water Supply. Wiley, Hoboken, NJ, p. 270.
4. Cho, Y.I., Lee, S., and Kim, W. (2003). Physical water treatment for the mitigation of mineral fouling in cooling-tower water applications, CH-03-3-3 (RP-1155). *ASHRAE Trans.* **109**: Pt. 1.
5. ASHRAE, Inc. (1999). Water Treatment, Chap. 47. In: *1999 ASHRAE Applications Handbook/SI Edition. HVAC Applications*, p. 47.5.
6. Watt, D.C., Rosenfelder, C.J., and Sutton, C.D. (1993). The effect of oral irrigation with a magnetic water treatment device on plaque and calculus. *J. Clin. Periodontal.* **20**: 314–317.
7. Ritchie, I.M. and Lehnen, R.G. (2001). Magnetic conditioning of fluids: an emerging green technology. In: *5th Annual Green Chemistry and Engineering Conference*, June 26–28. National Academy of Sciences, Washington, DC.
8. Troup, D.A. and Richardson, J.A. (1978). Scale nucleation on a heat transfer surface and its prevention. *Chem. Eng. Commun.* **2**: 167–180.
9. Baker, J.S. and Judd, S.J. (1996). Review paper, magnetic amelioration of scale formation. *Water Res.* **30**(2): 247–260.
10. Barrett, R.A. and Parsons, S.A. (1998). The influence of magnetic fields on calcium carbonate precipitation. *Water Res.* **32**(3): 609–612.
11. Bogatin, J. et al. (1999). Magnetic treatment of irrigation water: experimental results and application conditions. *Environ. Sci. Technol.* **33**: 1280–1285.

WATER IMPACTS FROM CONSTRUCTION SITES

JAMES A. JACOBS
Environmental
Bio-Systems, Inc.
Mill Valley, California

Surface water and, to a lesser degree, groundwater are impacted by runoff from construction sites, parking lots, roadways, and other nonpoint sources. Soil particles and dissolved constituents in runoff reduce the beneficial uses of water and have become a major regulatory challenge. Erosion rates at construction sites may be 10 to 20 times greater than for undisturbed lands (1). Construction sites, including environmental remediation activities and redevelopment, might contain a variety of natural and manmade contaminants that might exceed regulatory levels. Sediment runoff from nonpoint sources is primarily the result of wind and water erosion. Construction sites are locations where special erosion control measures may be needed to prevent storm water runoff and sediment buildup in nearby waterways and groundwater recharge areas (1,2).

To put this challenge into perspective, approximately 1.7 to 1.9 billion m^3 (1.4 to 1.5 million acre-feet) of the storage areas of reservoirs and lakes are permanently filled each year with sediment. The U.S. government annually spends over \$500 million each year to remove the sediment from harbors, rivers, and other waterways, primarily through the oversight of the U.S. Army Corps of Engineers. It has been estimated that erosion-related pollutants cost the United States about \$3.2 billion to \$13 billion each year (3). Since 1990 when the report was written, the challenge from nonpoint source pollution is better understood and studied. Over the years, the methods have been designed for effective sediment and erosion control to significantly reduce pollution and sediment volume from construction sites and other nonpoint sources. A good summary of the erosion control at construction sites and various engineered solutions are found in Fifield (1,2).

The objective of sediment and erosion control is to reduce the water-quality impacts from nonpoint sources, and many of the regulations and statutory requirements that attempt to ensure that construction projects are performed in a way that minimizes potential impact to the environment. The U.S. Congress has passed several laws to protect the environment from nonpoint sources, including construction sites. These laws include the National Environmental Policy Act, the Clean Water Act, the Coastal Zone Management Act, the Safe Drinking Water Act, the National Wild and Scenic Rivers Act, the Endangered Species Act, the Resource Conservation and Recovery Act, and the Federal Insecticide, Fungicide and Rodenticide Act. In addition to these federal acts, various state and local laws or guidelines have also been enacted.

Erosion occurs when raindrops, moving water, or wind transports soil particles. Both wind and water are transport agents for soil particles. Wind erosion is normally a problem in semiarid and arid regions where ground vegetation may be limited. Controlling wind erosion at construction sites is common and is frequently managed by spraying water to keep the dust and soil particles from becoming airborne. Silt fences consist of vertical plastic sheeting or other materials that catch and retain soil particles from leaving construction sites.

Water erosion at construction sites is common in humid regions where rainfall events are frequent. According to Fifield (2), six types of water erosion exist: Splash erosion is the dislodging of soil particles by raindrop impacts; sheet-flow erosion is the uniform removal of the saturated soil particles conveyed in storm runoff; rill erosion is a long, narrow depression or soil incision caused by increased topographic relief and higher runoff velocities; gully erosion is the deep and wide depression or incision caused by concentrated storm runoff flows; stream bank erosion is the removal of soil by a natural drainage pattern such as toe cutting and bank sloughing; and shoreline erosion is the removal of soil by high-energy wave action, which results in sloughing and mass wasting of shoreline cliff faces. Water-related erosion on construction sites is usually the result of raindrop impacts and sheet flows, although specific conditions at construction sites can contribute to water erosion near rivers or beaches.

One way to minimize water impacts at construction sites is to use sediment containment systems. Sediment containment systems create conditions for deposition of soil particles that are in suspension. Two types of sediment containment units exist: retention and detention systems. Retention systems tend to be large and, for individual construction sites, may be difficult for lack of space. Properly designed sediment retention systems provide the necessary volume for containment of all the incoming storm water runoff, including both the water and sediment portions. Within the sediment storage area, a properly designed retention system will have the capacity to allow for the sedimentation of the suspended soil particles. Finally, the sediment retention system will be designed to discharge the storm water at a controlled rate. Evaporation and seepage must be calculated to allow for the retention system to drain before the next incoming storm water surge. Retaining all storm waters from a construction site is generally not feasible because of space and time constraints.

Instead of trying to retain all storm water, a well-designed sediment detention system captures only a specific size of particles while being able to drain the large volume of storm water passing through. To be effective at reducing the suspended soil particles, detention containment systems must hold the storm water only long enough for the deposition of the suspended soil particles to occur (2).

Surface water runoff from construction sites, parking lots, roadways, and other nonpoint sources can enter groundwater through recharge areas. Turbidity, naturally occurring heavy metals as well as pesticides, hydrocarbons, and other organic contaminants, can be major sources of nonpoint pollution into surface waters during construction, which entails disturbing existing ground cover, trees, rocks, and soil. The total suspended solids (TSS) are the total amount of sediment in the water. The dissolved portion of solids in water is called total dissolved solids (TDS). In addition to high TSS and TDS, a large volume of sediments in surface water includes changes in water chemistry, such as too much salt, iron, lead, copper, and zinc or other metals. Along with the sediment, an influx of chemicals poisonous to fish and other aquatic life can occur. Other effects include the destruction of fish breeding grounds, the flattening of streambed channels, and the increased potential of surface flooding as river and lake storage capacity is reduced because of sediment buildup. Other wastes from construction include wash water from concrete mixers, paints and wash water from painting equipment, and wastes from cleaning of vehicles and equipment.

The solution to erosion and sediment control issues is site-specific erosion plans and sediment and waste containment systems. Site-specific sediment control systems in the past have used commonly available materials (plantings, rocks, sand bags, and straw-bales). More recently, engineered systems with synthetic materials have been tested and proven in the field. Surface and groundwater impacts from nonpoint source pollution, especially at construction sites, can be minimized and managed effectively.

BIBLIOGRAPHY

1. Fifield, J.S. (2004). *Designing for Effective Sediment and Erosion Control on Construction Sites*. Forester Communications Inc., Santa Barbara, CA. Available: www.foresterpress.com.
2. Fifield, J.S. (2004). *Field Manual on Sediment and Erosion Control Best Management Practices for Contractors and Inspectors*. Forester Communications Inc., Santa Barbara, CA. Available: www.foresterpress.com.
3. Forest, C. (1990). *Erosion Control in the United States Today: An Overview. Proceedings of Conference XXI of the International Erosion Control Association*, February 14–17.

INDUSTRIAL COOLING WATER—BIOFOULING

T.R. Bott
University of Birmingham
Birmingham, United Kingdom

The accumulation of living matter in cooling water systems may be the result of the activity of microorganisms such as bacteria, algae, and fungi, or the growth of macroorganisms that include mussels, barnacles, hydroids, and serpulid worms and plant material such as water weeds. The extent of the deposition of both micro- and macroorganisms, depends very much on the origin of the water employed in the cooling water circuit. In general, the water is taken from a natural source, which contains the organisms and nutrients to sustain life and growth. Some sources, such as borehole water, are relatively free of organic material, but in “open” cooling water circuits, it rapidly becomes contaminated from contact with the atmosphere. Water from other sources, including freshwater rivers, canals or lakes, estuarine or seawater, is already heavily contaminated before it is taken into an industrial system.

The presence of living matter on surfaces, particularly heat transfer surfaces, may promote fouling by mechanisms, including particle deposition, scale formation, and corrosion. In general, this is due to the metabolic activity of the living matter that creates localized changes in the properties of the water, particularly pH.

To maintain efficient operation of a cooling water system, it is necessary to combat the detrimental effects of the accumulation of unwanted deposits. It is particularly important with respect to heat transfer surfaces because the deposit represents a resistance to heat flow across the surface, thereby seriously reducing the capability of the system to operate according to the design requirements. Generally, the principal difficulties in cooling water circuits arise from the presence of microorganisms, because it is possible to reduce the gross effects of macroorganisms generally, by filtration at the cooling water intake.

SURFACES AND BIOLOGICAL FOULING

For both micro- and macroorganisms, although there are some exceptions, settlement on a surface is an essential element in biological activity. The advantages of residing on a surface include a continuous supply of nutrients, removal of waste products by the movement of the water, and some measure of protection from predators. Furthermore, it has been suggested (1) that for bacteria, the surface provides an enhanced protection from changes in the surrounding environment, compared to those microorganisms that are “free swimming” (planktonic) in the water. The changes that organisms might experience include variations in water pH, changes in nutrient concentration, and in some systems, the inorganic salt concentration. Changes in water velocity, in turn, affect the shear forces acting on the biofilm and hence the tendency for removal. It has been suggested that the colonization of a surface by bacteria depends on an adsorbed conditioning layer made up of macromolecules that are in plentiful supply in water from natural sources due to the breakdown of living matter. It is considered that the bacteria form a chemical bond with this layer of macromolecules. The movement of water in relation to the atmosphere in cooling towers and spray ponds usually means that the water becomes saturated with oxygen, conducive to the growth of organisms that depend on oxygen for survival, for example, aerobic bacteria.

Bacteria

Mozes and Rouchet (2) suggested that three zones are present in a biofilm consisting of bacteria:

1. the layer of cells in direct contact with the surface
2. the bulk biofilm
3. the layer of cells residing at the interface between the bulk biofilm and the water phase

Conditions are likely to be different for each of these groups of cells. The activity of the microorganisms will vary, depending on the mass transfer of nutrients, including oxygen for aerobic species, to and through the biofilm, and similarly for the removal of waste products of metabolism. If the biofilm is relatively thick, it is more than likely that the nutrient concentration will become zero either at the solid surface or somewhere in the bulk biofilm. Under these conditions, it is possible for the cells to die with potential consequences for the integrity of the biofilm.

Mass transfer depends on a concentration driving force, the extent of the turbulence at the water/biofilm interface, and the openness of the biofilm structure that either restricts or facilitates nutrient transport. The turbulence at the biofilm surface is very dependent on the water velocity across the surface.

Heat exchanger surfaces are particularly suitable for bacterial growth, because of these desirable benefits and also because the surface temperature is usually conducive to growth. In natural conditions, bacteria colonize stones and rocks, but in cooling water circuits, they use the metal surfaces in contact with water. Metal surfaces give rise to metal ions that could interfere with adhesion of the biofilm to the surface (3). Some metallic ions, for example, copper ions, may act as a biocide. As the colonies of bacteria grow on a surface, extracellular products (usually polysaccharides) are formed and accumulate around the clusters of cells to give a complex matrix of cells and these extracellular polymers. Apart from keeping the structure together, it is understood that the function of extracellular material is first and foremost to protect the cells. Protection may be afforded against predators or noxious substances in the water that might adversely affect the cells or kill them. It is also believed that the extensive growth of polysaccharides, when nutrients are in ample supply, is a method of storing nutrient material against a time when the supply is reduced or eliminated, so that survival is ensured. The spaces within the matrix are filled with water that represents 90–95% of the mass of the biofilm. Work by Stoodley et al. (4) demonstrated that water flows in the duct-like voids in a biofilm and facilitates the availability of nutrients for the cells in the matrix and the removal of waste products. The structure of the biofilm, therefore, has implications for the mass transfer of nutrients from the bulk water to the individual organisms in the biofilm, particularly those remote from the bulk water flow. The structure of the biofilm is very much a function of the conditions under which it was formed, particularly the magnitude of the water velocity across the cells that colonized the surface. Higher velocities tend to produce more compact structures. The use of nutrients by the cells reduces the local nutrient concentration in the vicinity of the biofilm, thereby enhancing the concentration driving force for mass transfer between the bulk water and the biofilm. It is possible that the cells close to the solid surface may be starved of nutrients that may cause changes in metabolism or even cell death.

OPERATING CONDITIONS AND BACTERIAL BIOFILM FORMATION

Because of the versatility of bacteria and the favorable conditions for growth in heat exchangers, bacterial biofilms on heat transfer surfaces represent a serious problem in cooling water circuits. The principal difficulties are the reduction of heat transfer efficiency and the increased energy required to pump the water through the system for a given bulk water flow rate necessary to achieve the desired heat removal. Substantial increases in operating costs may be the result.

The development of a biofilm generally follows the accepted pattern for most fouling phenonema, as shown by the idealized curve in Fig. 1. For a short interval of the order of a few days, there is little evidence of any biofilm accumulation. This is when the adsorption of macro organic molecules takes place and the surface becomes "conditioned," allowing microorganisms to colonize the surface. Once the surface has been colonized and providing sufficient nutrients are available, there is relatively rapid growth of the biofilm. After a period of accelerated growth, the rate of growth reduces to a point where the biofilm thickness remains reasonably constant, although variations in thickness about a mean are likely due to sloughing of lumps of the biofilm from the surface. It is generally understood that sloughing is due to the effects of the shear forces at the water/biofilm interface induced by the water velocity. As the biofilm develops and becomes thicker, its structure becomes less robust and more susceptible to the effect of fluid shear. As a result, there is a balance between growth and removal of the biofilm.

Figure 1 provides an indication of the general shape of the biofilm growth curve, but the shape of individual curves varies depending on the prevailing conditions. For instance, higher velocities are likely to produce thinner but robust biofilms, compared with biofilms formed under low velocities that have a more open structure. Biofilms formed at a temperature near the optimum for maximum growth for the prevailing species are thicker than those formed at a lower or higher temperature. Higher availability of nutrients encourages growth. As for all living material, the presence of trace elements is vital for healthy growth. All these influences change with the seasons, including the cooling water velocity which is likely to be lower in the winter months when the lower temperature of the water reduces the demand for cooling water for a given cooling requirement.

The topography of the surface on which the biofilm develops is also a factor. In general terms, rough surfaces favor biofilm accumulation, whereas smooth surfaces are less hospitable to biofilm growth. Bacteria carry a surface charge, so that its value in relation to the electric charge on the surface also influences the rate of colonization.

Suspended particulate solids in the water flow can also affect the development of biofilms. Reduced biofilm accumulation can be anticipated due to the scouring action of particles, although the nature and size of the particles themselves also exert an influence. Sand particles can remove biofilms very effectively, but at the same time, erosion of the metal surface could become a serious problem. On the other hand, softer clay particles may become part of the deposit structure and, under those circumstances, may reduce the robustness of the biofilm matrix.

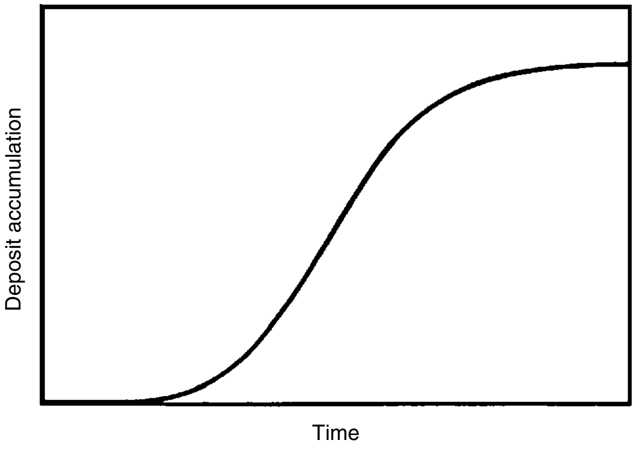

Figure 1. Idealized fouling curve.

Most bacteria flourish in an environment where the water pH is near 7. Variations affect the growth of biofilms, although most microorganisms can tolerate small changes in pH.

Algae

Algae require sunlight, carbon dioxide, and inorganic chemicals as their primary source of nutrients and energy for metabolism. Their natural habitat is attachment to rocks and stones in aqueous environments, or they may be "free floating" in lakes or reservoirs. Algae are larger than bacteria and exist as single entities or as filamentous colonies in an aqueous environment or attached to solid surfaces. In cooling water systems, they may be found where areas resemble those in nature that have access to sunlight, such as cooling tower basins, spray ponds, and serpentine coolers where the water, exposed to the atmosphere, flows over banks of horizontally mounted tubes. The formation of algal biofilms is initiated by the adsorption of water-soluble organic material, such as glycoproteins, and colonization by bacteria. Diatoms, unicells that have a symmetrical silica outer surface, are often early algal colonizers.

Algae themselves are not generally a problem in the enclosed parts of a cooling water circuit such as heat exchangers and pipelines, because of the lack of sunlight. The problem of algae growth in other parts of the system, however, is the production of organic debris that may give rise to blockage or be a source of nutrients for other microorganisms, such as bacteria that accumulate on heat transfer surfaces and in piping.

Fungi

Fungi require a fixed organic source of carbon. Their rigid cell wall limits them to being saprophytic on organic substrates or parasites living on higher forms of life, for example, so-called "athletes' foot" in humans. They lack chlorophyll, so that they cannot use photosynthesis in their metabolism.

In relation to cooling water systems, the nuisance value of fungi is not unlike that of algae. In sofar as they require a consumable substrate, they are to be found, for instance, on the wooden structures of older cooling towers. The attack of fungi on wooden structures involves enzymes (produced by the fungi cells) that open up the cellulose structure and make the structure liable to attack by other forms of microbial life. The presence of large quantities of living and dead material on the weakened wooden structure is likely to make it collapse. Modern cooling towers generally contain polymer packing that is resistant to fungal attack. The principal problem that fungi pose in cooling water circuits is the breakdown products and debris that arise from fungal growth. Operating problems may develop due to deposition on heat transfer surfaces and in pipelines, and by providing nutrients for microbial activity, particularly bacteria, as discussed earlier.

MACROFOULING OF SURFACES

Macrofouling is largely associated with sea or estuarine water involving the growth of macroinvertebrates. The fouling is usually confined to water intakes and culverts. Whitehouse et al. (5) summarized the conditions that support macrofouling of surfaces, including the continuous flow of seawater to provide oxygen and nutrients and reduced competition from algae and predators. Colonization of surfaces may be facilitated by bacterial biofilms (slime layers) already on the surface. The spectrum of species in the water varies with location and season.

Mussels

The growth of mussels is very much dependent on the season. Higher temperatures of summer facilitate growth, but at the lower winter temperatures, growth is restricted. Rapid growth, however, can occur at temperatures above about 10 °C. Following fertilization and a free swimming larval stage, metamorphosis occurs to produce a “spat.” Settlement onto a surface then occurs to provide a permanent adult site, where rapid growth can take place under suitable conditions. Settlement can be strongly affected by the water velocity across the surface; low velocities favor deposition and attachment. It is not practical to filter the water prior to the settlement process to remove the “spats,” because the small mesh size required would give rise to rapid “blinding” of the filter.

Barnacles

Barnacles are more of a problem in tropical regions than in more temperate zones. Planktonic larvae progress toward metamorphosis into small barnacles that can settle on surfaces under a higher water velocity than that for mussels. Blockage of filter screens is a common problem. Larvae can be carried forward into heat exchangers, where, if the velocity is relatively low, they may settle and develop to cause problems of blockage, pressure loss, and reduced heat transfer efficiency.

Hydroids

Hydroids are surface colonizers that facilitate the attachment of mussels. They are primitive creatures living in colonies of numerous microscopic units. There is a free swimming stage in development. The outer walls of hydroids are not calcified like barnacles and mussels, so their fouling potential is not great. Large accumulations, however, can severely restrict water flow.

Serpulid Worms

Calcareous tubes built by serpulid worms can cause fouling problems. Encrustations in pumps and valves can cause malfunction.

Other Organisms

There are many living organisms contained in estuarine and seawater that have not been listed but can cause severe operating problems in cooling water systems. The examples cited illustrate the general problems of macrobiofouling.

CONTROL OF BIOFILMS

It is essential to control biofilm growth in cooling water circuits particularly in heat exchangers, to maintain cooling efficiency and thus optimize costs. The two basic methods are chemical and physical.

Chemical Control

Chemicals, particularly chemical biocides that kill bacteria, have been used for many years. The action of the additive has essentially three phases:

1. adsorption onto the cell wall
2. penetration into the cell wall
3. chemical combination of the biocidal chemical and essential cell structures or components that directly affects cell metabolism or the integrity of the cell structure itself.

Chlorine has been the preferred biocide for many years on account of its availability and relatively low cost. Free chlorine reacts with water in the following way:

$$Cl_2 + H_2O \longrightarrow HOCl + HCl$$

and

$$HOCl \longrightarrow H^+ + OCl^-$$

Because it is a strong oxidizer, chlorine reacts with organic material in water. Many of these compounds are cancer forming, and chlorine can enter the food chain. It has, therefore an environmental risk, so its use is declining and in some places, it is not possible to discharge water that contains chlorine or its derivatives back into the environment from which it came without treatment. Other oxidizing biocides include hydrogen peroxide, ozone, and chlorine dioxide. The advantage of hydrogen peroxide and ozone is that they break down to water and oxygen.

There are many other chemical biocides available for the control of biofilm formation. In recent years, so-called “environmentally friendly” biocides have been developed that have a short life in aqueous solution and break down to innocuous by-products or nutrients for microorganisms that effectively disposes of these additives. Other chemicals that are used to control biofilm growth are biostats. These compounds reduce the activity of microorganisms, so that their influence is limited. Biostats may be included with appropriate biocides in proprietary mixtures for microbial control.

Physical Control

The use of physical techniques to control biofouling is not as widespread as the use of chemical additives. The principal reason for this is that chemicals are carried to all parts of the system, whereas this is not so for physical methods of control. Furthermore, the geometry of the cooling water system, particularly the heat exchangers, does not lend itself easily to the use of physical techniques.

One technique that has been successful in controlling biofouling in power station condensers is the so-called Tapproge system that involves circulating sponge rubber balls with the cooling water through the tubes of the condensers. The technology is suitable because the large number of tubes in a condenser are uniform in diameter. The balls are slightly larger in diameter than the internal diameter of the tubes, so that as the balls are forced through the tubes by the water flow, they are "squeezed" and wipe the inside surface of the tubes. The technique involves collecting balls, rejecting worn balls, and reinjecting acceptable balls. The effectiveness of the system depends on the availability of sufficient balls and the probability that balls pass through each tube reasonably frequently. Under ideal conditions, it has been estimated that each tube should receive 12 balls per hour (6). A potential drawback of the system is that balls may get stuck in the tubes or at a tube entrance on the tube plate, thereby causing a reduction in effective heat transfer area and an increase in pressure drop.

Potential future physical methods for biofouling control have recently been discussed (7). The use of ultrasound, the inclusion of wire-wound or spring-like inserts in tubes, and the circulation of robust polymer fibers are under development. The two former techniques depend, to some extent at least, on heat exchanger geometry and on increasing the removal forces acting on the biofilm generated by increased turbulence in the flowing water. The inclusion of polymer fibers in the flowing water, on the other hand, is more akin to the use of chemicals because the fibers are transported to all parts of the system by the water. The removal action is assumed to be from contact between the fibers and the deposit.

Ultraviolet light has been used to control biofilm growth, but the technique has severe drawbacks in cooling water systems. To be effective, the UV light source has to "see" the biofilm or the planktonic organisms in the circulating water. It is impossible, therefore, to control biofilm growth on heat transfer surfaces. It is possible, however, to pass the cooling water through a UV light unit that destroys organisms in the water. The problem that arises is that unless the kill is 100%, the remaining microorganisms can become colonizers of heat exchanger surfaces that are untouched and therefore are free to develop unwanted biofilms.

Treatment to modify the quality of surfaces likely to become fouled is a method of biofouling control. Electropolishing of stainless steel reduces fouling (8). The use of coatings to change the hydrophilicity of the surface has also been successfully used. Ion implantation has been effective in fouling control (9). The additional capital costs. imposed by these techniques and maintenance of the integrity of the treated surface are reasons why they are not used extensively in industry.

A more extensive discussion of cooling water and biofouling is to be found in (10).

BIBLIOGRAPHY

1. Fletcher, M. (1992). Bacterial metabolism in biofilms. In: *Biofilms—Science and Technology*. L.F. Melo, T.R. Bott, M. Fletcher, and B. Capdeville (Eds.). Kluwer Academic, Dordrecht, pp. 113–124.
2. Mozes, N. and Rouxhet, P.G. (1992). Influences of surfaces on microbial activity. In: *Biofilms-Science and Technology*. L.F. Melo, T.R. Bott, M. Fletcher, and B. Capdeville (Eds.). Kluwer Academic, Dordrecht, pp. 126–136.
3. Vieira, M.J., Oliveira, R., Melo, L.F., Pinheiro, M.M., and Martins, V. (1993). Effect of metallic ions on the adhesion of biofilms formed by *Pseudomonas fluorescens*. *Colloids Surf. B: Biointerfaces* **1**: 119–124.
4. Stoodley, P., de Beer, D., and Lewandowski, Z. (1994). Liquid flow in biofilms. *Appl. Environ. Microbiol.* **60**: 2711–2716.
5. Whitehouse, J.W., Khalanski, M., Saroglia, M.G., and Jenner, H.A. (1985). *The Control of Biofouling in Marine and Estuarine Power Stations*. CEGB Northern Region, UK.
6. Eimer, K. (1985). *Recommendations for the Optimum Cleaning Frequency of the Taprogge Tube Cleaning System*. Technical Report 85-26, Taprogge Gesellschaft mbH, Wetter.
7. Bott, T.R. (2001). Potential physical methods for the control of biofouling in water systems. *Chem. Eng. Res. Design* **79**: 484–490.
8. Mott, I.E.C. and Bott, T.R. (1991). The adhesion of biofilms to selected materials of construction for heat exchangers. *Proc. 9th Int. Heat Transfer Conf.* Jerusalem **5**: 21–26.
9. Zhao, Q. and Muller-Steinhagen, H. (1999). Influence of surface properties on heat exchanger fouling. In: *Proc. Int. Conf. Mitigation Heat Exchanger Fouling Its Econ. Environ. Implications*. T.R. Bott, A.P. Watkinson, and C.B. Panchal (Eds.). Banff, Canada, July 1999. UEF and Begell House, New York, pp. 217–228.
10. Bott, T.R. (1995). *Fouling of Heat Exchangers*. Elsevier Science, Amsterdam.

INDUSTRIAL COOLING WATER—CORROSION

T.R. Bott
University of Birmingham
Birmingham, United Kingdom

Corrosion is often a problem when metals are in contact with an aqueous environment. It may be defined as the deterioration and loss of metal due to some form of chemical attack. In general, it results from impurities in the water, which need only to be in trace quantities to initiate corrosion. Trace impurities can include gases dissolved from the atmosphere, particularly oxygen and carbon dioxide, that play important roles in the corrosion process. Small amounts of mineral salts in water can also become involved in corrosion reactions. Where the metal makes up an essential component of a structure, such as a bridge or equipment forming part of an industrial process, for example, a cooling water circuit, constant vigilance is required to maintain the integrity of the structure and the equipment involved. Corrosion is a complex phenomenon dependent on chemical and physical conditions within the particular system.

In general, corrosion is associated with iron and steel, although under certain circumstances, other metals and alloys may attacked. However, to illustrate the mechanisms of corrosion, reactions involving iron will be used.

The reddish-brown deposit that is generally evident when iron corrodes is the iron compound ferric hydroxide, usually mixed with ferrous oxide. It is produced by a sequence of electrochemical reactions in which a difference in electric potential develops between two different metals or between different areas on the same metal. The difference in potential enables electrons to flow from the positive to the negative regions which, is an electric current. Thus, there is a movement of electrons which affects the chemical nature of the surface, where they originate or the area to which the migrate, by localized electrochemical reactions. The mechanism is often referred to as "galvanic action." These centers of electrochemical activity, usually termed the "galvanic cell," have an anode and a cathode. Metal dissolution takes place at the anode. The overall corrosion process is facilitated by dissolved salts, usually inorganic, in the water, because they affect the electrical conductivity of the water. Water that has a high salt concentration is more conductive than water of a relatively low salt content. The basic migrations are shown in Fig. 1.

At the anode, electrons are released from the metal enabling it to dissolve in the water in ionic form.

$$\mathrm{Fe} \longrightarrow \mathrm{Fe}^{2+} + 2\mathrm{e} \tag{1}$$

where e represents an electron.

Electron flow through the metal constitutes an electric current. At the cathode, it is possible for electrons to take part in electrochemical reduction reactions.
Examples are

$$2\mathrm{H}^+ + 2\mathrm{e} \longrightarrow \mathrm{H}_2 \uparrow \tag{2}$$

The H ion arises from the disassociation of water molecules which occurs to a small extent in water.

$$\mathrm{H_2O} \longrightarrow \mathrm{H}^+ + \mathrm{OH}^- \tag{3}$$

$$2\mathrm{H_2O} + 2\mathrm{e} \longrightarrow \mathrm{H}_2 \uparrow + 2\mathrm{OH}^- \tag{4}$$

$$\mathrm{O}_2 + 2\mathrm{H_2O} + 4\mathrm{e} \longrightarrow 4\mathrm{OH}^- \tag{5}$$

As the combination of hydrogen ions takes place to form hydrogen gas, the depletion of hydrogen ions causes Equation 3 to move to the right, thereby releasing further OH^- ions. The result is a significant increase in alkalinity, represented by the increased concentration of OH^- ions. The presence of Fe and OH^- ions can give rise to chemical reactions and hence relatively insoluble deposits such as ferrous hydroxide:

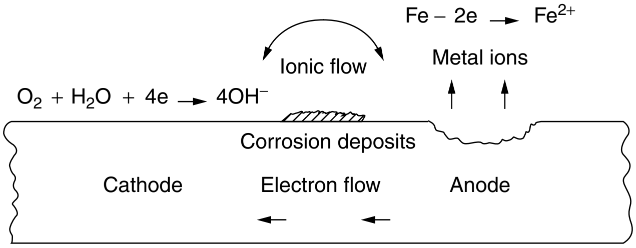

Figure 1. Electrochemical corrosion.

$$\mathrm{Fe}^{2+} + 2\mathrm{OH}^- \longrightarrow \mathrm{Fe(OH)_2} \downarrow \tag{6}$$

Ferrous hydroxide itself can further react to give ferric hydroxide:

$$4\mathrm{Fe(OH)_2} + \mathrm{O_2} + 2\mathrm{H_2O} \longrightarrow 4\mathrm{Fe(OH)_3} \downarrow \tag{7}$$

Ferric hydroxide can react further:

$$2\mathrm{Fe(OH)_3} \longrightarrow \mathrm{Fe_2O_3} + \mathrm{H_2O} \tag{8}$$

Thus a mixture of corrosion products can be formed, and within the deposits, there may be inclusion of solids from the water, such as suspended particles comprised of clay, sand, and other mineral matter, along with microbial cells.

The character of the deposit can affect the continued rate of corrosion. For instance, if the deposit is porous, corrosion can continue. On the other hand, a dense impervious structure prevents ionic migration and thereby protects the underlying metal from further corrosion. The "oxide layer," as it is called, is often relied upon to provide protection and may be deliberately laid down for corrosion control.

Equations 2 to 5 demonstrate that hydrogen gas or hydroxyl ions are produced at the cathode, as a result of corrosion. If hydrogen and hydroxyl ions remain in the vicinity of the cathode, they represent a barrier that interferes with the diffusion of oxygen gas or hydrogen ions to the cathode that would enable further chemical reactions to take place. The barrier therefore acts as a natural corrosion inhibitor. The overall process is termed "polarization."

This barrier, however, cannot be relied on as an effective method of corrosion control because it is easily disturbed. In most situations, the water is flowing and the "flushing" that is created removes the hydroxyl ions and the hydrogen bubbles from the surface. Furthermore, because the water is likely to contain dissolved carbonates and carbon dioxide in ionic form that can produce hydrogen ions, the pH of the water is reduced. The chemical reactions involved are

$$\mathrm{H_2O} + \mathrm{CO_2} + \mathrm{CaCO_3} \longrightarrow \mathrm{Ca(HCO_3)_2} \tag{9}$$

$$\mathrm{Ca(HCO3)_2} \longrightarrow \mathrm{Ca}^{2+} + 2\mathrm{HCO}^{3-} \tag{10}$$

and

$$\mathrm{H_2O} + \mathrm{CO_2} \longrightarrow \mathrm{H_2CO_3} \tag{11}$$

$$\mathrm{H_2CO_3} \longrightarrow \mathrm{H}^+ + \mathrm{HCO_3}^- \tag{12}$$

The increased concentration of H^+ ions allows reactions with the OH^- in the barrier to form water, thereby in effect removing the barrier.

The discussion so far, has centered on the corrosion of iron, but other metals immersed in water can corrode and produce a measurable potential. The alkaline earth and alkali metals hold their outer shell of electrons rather loosely and have a greater potential difference than iron. On the other hand, so called "noble" metals such as gold produce very few electrons and therefore have relatively low potential. Those metals displaying lower potential, it could be assumed, corrode less extensively and less easily than those of higher potential.

The discussion has emphasized the importance of H^+ ions in the corrosion process. The pH of an aqueous solution is the negative $\log_{10}$ of the molar concentration of H^+ ions. The range of pH values is 0–14. Low pH (<7) signifies an "acidic" condition, whereas high pH (>7) represents alkaline conditions. The effect of pH on metals is dictated by the behavior of the metal oxide. If, for instance, a metal oxide is soluble under acidic conditions, the metal will exhibit a high corrosion rate under these conditions. Alternatively, if the metal oxide is soluble under alkaline conditions, $\text{pH} > 7$, the metal will corrode in contact with water. Some metals have oxides that are soluble in both acid and alkaline conditions. Zinc, tin, and aluminum fall into this group and are usually referred to as amphoteric metals. Metals that are insoluble at any pH are called "noble," as already indicated.

Iron corrosion displays what might be considered inconsistent behavior. At low pH up to about 4, acidic conditions, the behavior is similar to that of an acid soluble metal. Between a pH of around 4 and up to about 11, the pH of the water has relatively little effect on the corrosion rate because oxygen depolarization is considered the principal influence on the corrosion process. Further increase in pH values reduces corrosion even further, the minimum is at a pH of around 12. At higher pH values, iron displays amphoteric behavior and corrosion again increases.

In industrial operations, notably cooling water applications, there are generally influences from the processing environment, in which the water is being used, for instance, the use of a biocide to control biofouling. Chlorine has been the preferred biocide to control microbial growth in cooling water systems for many years on account of its availability and effectiveness, but it can influence the corrosion of the cooling water system. Chlorine is soluble in water and reacts with it to give hypochlorous acid.

The chemical equation is

$$Cl_2 + H_2O \longrightarrow \underset{\text{Hypochlorous acid}}{HOCl} + \underset{\text{Hydrochloric acid}}{HCl} \tag{13}$$

The acids ionize to give H^+ ions that reduce the pH and hence increase the potential for corrosion. The associated equations are

$$HOCl \longrightarrow H^+ + OCl^- \tag{14}$$

and

$$HCl \longrightarrow H^+ + Cl^- \tag{15}$$

Unwanted chemicals in the system may cause corrosion. For instance, ammonia may be introduced into cooling water from leaks of process fluids in the coolers of a chemical works. Ammonia can be highly corrosive to copper. The chemical reaction involved is

$$2Cu + 2H_2O + O_2 + 8NH_3 \longrightarrow 2Cu(NH_3)_4{}^{2+} + 4OH^- \tag{16}$$

The cuprammonium ion $Cu(NH_3)_4{}^{2+}$ is very soluble, and it can cause rapid deterioration of copper components in a system.

A common problem in cooling water systems is the accumulation of micro- and macroorganisms on surfaces, particularly on heat transfer surfaces. The metabolism of these organisms can give rise to modified local conditions that are often acidic, and thereby produce local corrosive conditions under the microbial deposit. If the attack on the metal is sustained, this can give rise to a corrosive condition known as "pitting."

Some microbial growth can produce unique corrosion conditions. A microbial species known as *Desulfovibrio desulfuricans,* a sulfate reducing bacterium, gives rise to the production of hydrogen sulfide from the sulfates commonly found in water used for cooling purposes. Hydrogen sulfide dissolved in water lowers the pH and hence initiates corrosion. Hydrogen sulfide in solution in contact with iron or steel produces iron sulfide. The iron sulfide that is cathodic with respect to iron promotes galvanic corrosion.

Other Combinations of Conditions That Can Lead to Corrosion

When two dissimilar metals are in contact and are immersed in a salt solution, as described earlier, a potential difference can develop. This potential difference provides the opportunity for corrosion to occur, and this corrosion can be severe. The basic conditions for the interaction of two dissimilar metals are illustrated in Fig. 2.

Metal X has a lower electrode potential than metal Y. Ions migrate into the conducting water, and electrons flow across the junction of the two metals. Metal X is corroded at Z.

Immersion of the bimetallic couple in conducting water indicates only that corrosion is possible, but the rate of corrosion is governed by other external influences. The rate of corrosion, for instance, increases in proportion to the ratio of cathode to anode area. Generally, corrosion is also affected by other influences. Corrosion is the result of chemical reactions, so increased temperature accelerates corrosion. The velocity of the water across the corroding metal can exert an effect on the corrosion rate because it affects the resistance to the mass transfer of chemical components across the laminar sublayer.

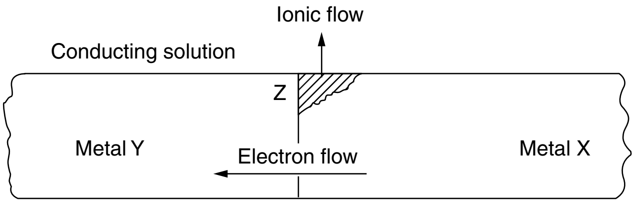

Figure 2. Corrosion at the junction of two dissimilar metals.

The metallurgy of the metal or alloy undergoing corrosion can also affect the extent of corrosion. Surface flaws such as scratches can become anodic to the rest of the metal. Anodic sites may also be established where there is stress within the metal at grain boundaries. Such a condition is often referred to as "stress cracking." So-called "caustic embrittlement" is considered a version of "stress cracking," and may result from the production of high concentrations of sodium hydroxide in the water in contact with the metal. Inclusion of metal particles that are not homogeneous with the principal metal in the system can also create a corrosion cell. For instance, brass is prone to this form of corrosion. Brass is an alloy of zinc and copper, if the zinc is removed from the surface of the metal, a porous copper layer remains that can establish corrosion cells. The process is usually referred to "dezincification." A combination of a corrosive environment and repeated working may result in early failure called "corrosion fatigue cracking." The fatigue is brought about by the repeated application of stress, such as in the tubes of a shell and tube heat exchanger that vibrate in response to operating conditions. If the protective oxide layer on the surface of the metal, discussed earlier, is subject to impingement of solid particles, or even bubbles traveling at high velocity in the water flow, the protective layer may be removed or not even formed, thereby allowing corrosion to occur.

CORROSION CONTROL

There are two principal ways in which corrosion of metal components can be eliminated or, at least, reduced by additives in circulating water. They include corrosion inhibitors and "controlled scaling." In general terms, the protection is achieved by positioning an electrical barrier on the metal that is under threat of corrosion, so that the electrical circuit is broken and the electron flow restricted. As indicated earlier, some systems have "natural" resistance to corrosion through the development of an impervious oxide layer on the metal surface that acts as an effective electrical barrier.

Corrosion, as described earlier, involves anodic and cathodic reactions. Chemicals that arrest or restrict these reactions halt or reduce corrosion. The terms anodic and cathodic inhibitors are used to classify the chemicals that are used for this purpose. Anodic inhibitors prevent or restrict electron flow (see Fig. 1), whereas cathodic inhibitors form a barrier at the cathode, restricting hydrogen and oxygen transport to the cathode surface. These are some chemical additives and their disadvantages:

Anodic Inhibitors	
Chromates	Generally not used now due to their toxicity
Nitrites	Nutrient for aquatic plant life
Silicates	Slow acting
Tannins	Nutrient for biological activity
Benzoates	Unreliable except perhaps for aluminum
Orthophosphates	Nutrient for aquatic plant life

Cathodic Inhibitors	
Zinc salts	Unlikely to be environmentally acceptable
Polyphosphates	Sludge forming and nutrient for aquatic plant life
Polyphosphonates/ molybdates	Require O_2 for effective protection; expensive.

So called "filming amines" that cover the entire surface of a metal liable to corrosion are used, particularly in the presence of steam condensate. An alternative is to apply an electrical current to the metal that is susceptible to corrosion to reverse the reactions described by Equations 2, 4, and 5, applying Le Chatelier's principle.

A more comprehensive discussion of corrosion in industrial cooling water can be found in Bott (1).

BIBLIOGRAPHY

1. Bott, T.R. (1995). *Fouling of Heat Exchangers*. Elsevier Science, Amsterdam.

INDUSTRIAL COOLING WATER—SCALE FORMATION

T.R. Bott
University of Birmingham
Birmingham, United Kingdom

Any source of naturally occurring water used of industrial cooling has been in contact with the earth, soil, and rock formations. As a result, soluble inorganic salts have been taken into solution that can give rise to "scaling" in industrial equipment. The phenomenon is apparent in domestic kettles that have been used to boil so-called "hard water." A whitish-brown deposit forms on the areas within the kettle that are subject to the highest temperatures, for example, the heating element in electric kettles. In cooling water circuits that usually employ water from natural sources such as rivers or lakes, deposition of inorganic salts, or "scaling" as it is often called, can represent a serious problem. The presence of these salts usually results in operating problems. The accumulation of deposits on heat transfer surfaces can lead to severe loss in cooling efficiency if not effectively controlled.

In general, there are two categories of salt present in water that is natural in origin, "normal" and "inverse solubility" salts. As the temperature of the water is increased, the solubility of the normal salt also increases, for example, sodium chloride. As the temperature of an inverse solubility salt solution is increased, the solubility of the salt decreases, and, if the solubility limit is exceeded, the salt begins to precipitate. Inverse solubility salts, for example, calcium carbonate, are usually sparingly soluble, whereas normal solubility salts have relatively high solubility. It will be seen from Table 1 that many common mineral salts display the properties of inverse solubility salts. They are often referred to as "hardness salts."

Table 1. Common Inverse Solubility Salts

Calcium	carbonate hydroxide phosphate silicate sulfate
Lithium	carbonate sulfate
Magnesium	hydroxide silicate

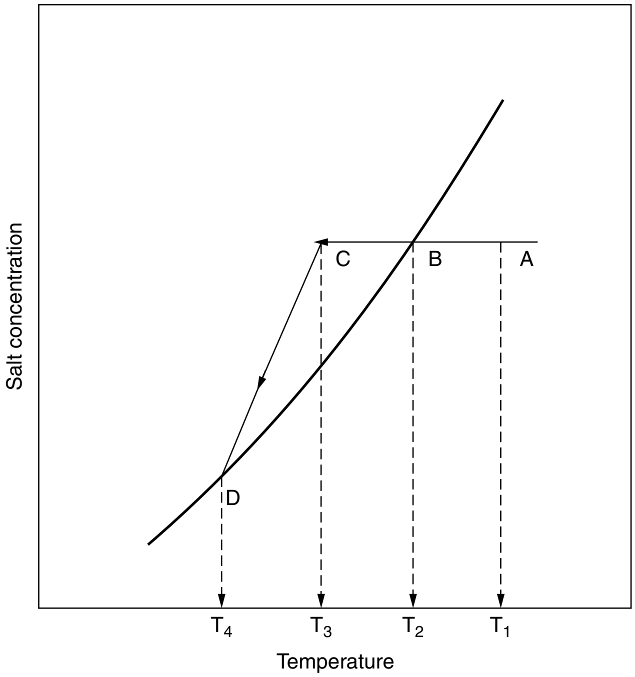

Figure 1. Cooling a normal solubility salt solution.

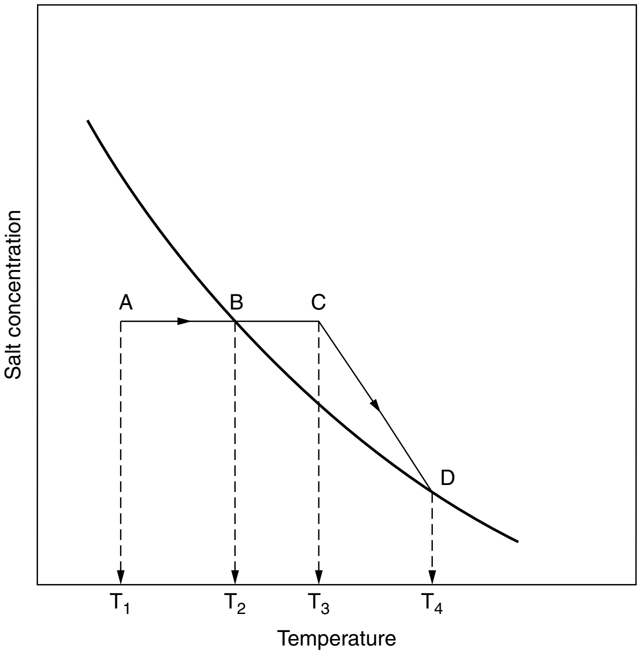

Figure 2. Heating an inverse solubility salt solution.

Whenever the temperature distribution changes in a water system containing dissolved salts, such as in an industrial cooling water system, the temperature sensitive solubility can affect successful operation. Deposition of inverse solubility salts is often one of the causes of operating problems in cooling water circuits, particularly on heat exchanger surfaces. The potential for scaling in industrial equipment is usually very high.

A solution that is in equilibrium with a soluble solid phase is said to be "saturated." The concentration in the water is the maximum that is possible at the given temperature. It is possible, however, under certain conditions, that as a saturated solution of an inverse solubility salt is raised slightly in temperature, precipitation is retarded, and the solution is said to be "supersaturated." Supersaturation can also occur with normal solubility salts. If the temperature of a saturated solution of a normal solubility salt is lowered slightly, a supersaturated solution may be produced.

Supersaturation is a necessary prerequisite for crystallization from a salt solution to occur. Its importance is illustrated in Figs. 1 and 2.

The curves in both figures represent the variation of saturated concentration with temperature. For a normal solubility salt, if the unsaturated solution represented by 'A' at temperature T_1 in Fig. 1. is cooled, that existing concentration will represent saturated conditions at the lower temperature T_2 at 'B'. Further cooling down to T_3 produces supersaturated conditions at 'C'. Supersaturation is represented by the temperature difference T_2-T_3, sometimes referred to as the "maximum allowable undercooling." 'C' represents a metastable condition; any disturbance such as the introduction of particles or crystals results in precipitation. If cooling (and precipitation) is continued, eventually a new equilibrium is reached at 'D' on the solubility curve, where saturation occurs at temperature T_4. Crystallization can occur either in the bulk or at the cold surface, through which heat is being withdrawn from the solution. The latter gives rise directly to deposit formation (scaling), and the crystallites produced in the solution may move toward the surface, under the influence of the hydrodynamics of the system, to be incorporated into the deposit on the surface.

The sequence of events, when an inverse solubility salt solution is heated is illustrated in Fig. 2. As the solution represented by 'A' is heated, it moves toward the saturation point 'B' on the solubility curve. As the temperature is further increased, point 'C' is reached where there is a degree of supersaturation and metastable conditions exist. The supersaturation is T_3-T_4. As the temperature is further increased, accompanied by a disturbance of the metastable condition, precipitation occurs, and the concentration of the salt in solution decreases till point 'D' is reached, representing equilibrium saturation conditions at temperature T_4. In a cooling water system, this could represent an operational problem in the coolers. In cooling operations, the highest temperature in contact with the coolant water is at the heat exchanger surface/water interface. Under these conditions, deposit formation, or "scaling," on the heat transfer surface can be expected. Furthermore, disturbance of the metastable supersaturation in the bulk water gives rise to crystallites that are likely, under the prevailing hydrodynamic conditions, to migrate toward the heat transfer surface with the further possibility of incorporation into the developing deposit.

In summary, the complex sequence of steps leading to scale formation is as follows: temperature changes that lead to supersaturation; disturbance of the metastable conditions, usually caused by solid particles, for instance, crystallites formed elsewhere in the system or particles of corrosion material that initiate the formation of crystallites, and ultimately crystals, either in the bulk water or at the heat transfer surface. Local hydrodynamic and temperature conditions affect the location of crystal and deposit formation.

Under constant heat flux, the presence of scale, a resistance to heat transfer, changes the temperature distribution across the heat transfer surface. These changes may affect the deposit structure and hence its strength, by reorienting the crystal lattice, with implications for the control of deposit accumulation. The degree of supersaturation may also be affected.

The previous discussion emphasizes that the key to scaling is the degree of supersaturation that is present, and it is likely that this will be different in different parts of the system. The following is a brief resume of the conditions that could lead to supersaturation; more detail is to be found in Bott (1).

Evaporation of water vapor from a solution, increases the salt concentration that could lead to supersaturation. The principal points in a system where this can occur is a cooling tower or a spray pond. It is here that the cooling water temperature is reduced for reuse by removing heat due to the latent heat required to vaporize some of the water. The "buildup" of dissolved salts is controlled by "blowdown," the removal of water containing relatively high concentrations of dissolved solids, to be replaced by "makeup" water that has the lower concentrations of the source water.

Mixing of different waters containing a common ion may lead to precipitation, the so-called "common ion effect" and attendant supersaturation. For instance, mixing a $Ca\ Cl_2$ solution with a solution of $Na_2\ CO_3$ may cause $Ca\ CO_3$ to precipitate if supersaturation occurs as a result.

Mixing of saturated or nearly saturated solutions may result in supersaturation and consequent precipitation.

Carbon dioxide, from the atmosphere dissolves in water in a cooling tower or spray pond. The chemical equilibria are

$$H_2O + CO_2 \longrightarrow H_2CO_3$$
$$H_2CO_3 \longrightarrow H^+ + HCO_3{}^-$$

The bicarbonate ion can further ionize:

$$HCO_3{}^- \longrightarrow H^+ + CO_3{}^{2-}$$

If the water is already saturated with dissolved $CaCO_3$ the additional $CO_3{}^{2-}$ ions can produce supersaturation of $CaCO_3$, and crystallization of $CaCO_3$ is possible. Hydrogen ions are involved in the reactions, so the solution pH will also affect the extent of supersaturation (2).

Crystallization

Before crystallization can occur that might lead to scale formation, it is necessary to "seed" the solution to initiate the process; usually this is accomplished by minute particles. The process is generally referred to as "heterogeneous crystallization." The particulate solids may be foreign bodies, such as pieces of scale broken off the deposit upstream, corrosion products, or crystallites that have been spontaneously formed in solution. Suitable nucleation sites may also be present on surfaces such as heat exchanger surfaces within a cooling water system. It is possible that this process is responsible for the initial deposition that provides a structure onto which further crystallization can occur or into which crystallites and larger crystals can be accommodated. Spontaneous crystallization occurs without the need for seeding. The mechanism is not well understood, but it is considered that it is related to random clustering of molecules in suitable juxtaposition to allow a tiny crystal to form. Spontaneous crystallization, however, is unlikely to occur in industrial cooling water systems because of the almost universal presence of particulate matter that will initiate scaling before spontaneous crystallization can occur.

Depending on the prevailing conditions, it is possible for a robust crystal structure to be formed on solid surfaces, particularly on heat transfer surfaces, where the necessary supersaturation is present. The structure may be weakened, making removal easier, by the presence of impurities, such as particles, in the structure.

The factors and conditions that affect scaling may be summarized as follows (1):

System Variables

Total dissolved solids concentration

Concentration of particular species

Alkalinity/acidity defined as free alkali (e.g., $CaCO_3$) and free acid (e.g., H_2SO_4) expressed as mg/L.

pH

Temperature

Plant Factors

Residence time

Evaporation, that is, concentration effects

Changes in heat load that affect the temperature distribution across the heat transfer surface, thereby changing the surface/water interface temperature

Water velocity

The general design of the equipment may affect the deposition process.

SCALING CONTROL

The conditions and factors listed that affect the scaling process are generally inherent in any cooling water system and its associated processes. In most scaling problems, therefore, the operating conditions cannot be modified. Even in a medium-sized plant, due to the large volumes of water in the cooling water system, the amount of $CaCO_3$ present, even in small concentrations, will be large and may amount to several tonnes; this has to be kept in solution by suitable water treatment.

Due to the complex chemistry of cooling water containing dissolved salts, it is difficult to assess whether or not the water will be scale forming or corrosive. As a result "indexes" were proposed to provide guidance. If the pH of water saturated with $CaCO_3$ by contact with the solid salt is denoted pH_s and pH is the measured pH of the water in the system, Langelier (3) proposed that $(pH - pH_s)$ could be used as an index to indicate whether the water would produce scale or be corrosive. If the Langelier index $(pH-pH_s)$ is negative $(pH_s > pH)$, it is assumed that the water would be corrosive. If on the other hand, the index is positive $(pH > pH_s)$ the water is saturated in $CaCO_3$, and scale formation is a distinct possibility.

Ryznar (4) suggested an improved index, $(2pH_s-pH)$. If the index is above 7, it is likely that the water would be corrosive. An index of 6 or lower suggests the possibility of scale formation.

Although these indexes have been used for many years in connection with water treatment in cooling circuits, there is some evidence that they are not altogether reliable because they are based on $CaCO_3$. Such a procedure ignores the complex water chemistry of associated cooling water systems and plant operating conditions.

Acid Treatment

The indexes discussed above suggest that a method of scale control could be the addition of acid, for example, sulphuric acid, to reduce the pH. There is a distinct risk, however, that severe corrosion may occur as a result of the acid treatment, so that in general, more sophisticated techniques are employed.

Water Softening (Removal of Hardness Salts)

Lime softening, ion exchange, and the use of membranes are potential methods for the control of scaling in industrial cooling water systems, but because of the large volumes of water to be treated, only lime treatment is used to any extent from this group of techniques. The more sophisticated techniques of ion exchange are far too expensive, except in closed cooling water circuits. Pretreatment with lime involves precipitation of carbonate hardness by hydrated lime, as calcium carbonate and magnesium hydroxide. An alternative is to pretreat with soda ash (Na_2CO_3) that precipitates other calcium salts in solution in addition to the carbonate. Soluble by-products in the form of sodium salts remain in solution.

Threshold Treatment (Sequestration)

When some chemicals are added to cooling water, even in relatively small quantities, they prevent the precipitation of alkaline salts and other scaling compounds, particularly calcium ions and manganese salts. Sometimes the result of the treatment is the production of sludge rather than scale that adheres to the heat transfer surface. The sludge can be removed from the system by suitable flushing. Phosphorous compounds are often used as sequestering agents. Table 2 provides some information on phosphorous compounds and their advantages and disadvantages for use with cooling water.

Crystal Modification

In the discussion of crystallization, it was stated that impurities included in the scale can reduce the cohesiveness of the crystal structure. It is possible, therefore, to effect scale control by adding so-called "scale modifiers" to cooling water. The action of these chemical additives does not prevent precipitation, as does threshold treatment, but simply reduces the opportunity for strong crystalline structures to form on heat transfer surfaces. The weaker structure of the scale makes it more susceptible to removal forces produced by the water velocity at the water/solid interface. Naturally occurring organic compounds, such as lignin or tannin, have been used for this purpose, but synthetic chemicals such as polycarboxylic acids are now more widely used as crystal modifiers. Polymaleic acid, for instance, produces a soft deposit with precipitating calcium carbonate that is easily removed.

Dispersants

It is possible to add a dispersing agent to the cooling water to keep the individual crystallites apart as they form in the water, and so reduce the opportunity of forming coherent crystal structure. Surface active agents on solid surfaces, such as heat transfer surfaces, may also hold the crystallites away from the surfaces, and so restrict scale formation. Dispersants that have been used for

Table 2. Organophosphorous Compounds Used as Sequestering Agents[a]

Compound Type	Advantages	Disadvantages
Polyphosphates	Hydrolysis to orthophosphate which reacts with calcium to form a sludge	Nutrient for microorganisms
Phosphonates	Effective against both alkaline and nonalkaline scaling conditions	If supersaturation levels are allowed to rise, uncontrolled precipitation may occur
Phosphate esters	Wide range available; biodegradable	Less effective than phosphonates
Phosphoncarboxylic	Similar to phosphonates	Some rise in supersaturation may be tolerated without attendant precipitation

[a] Reference 1.

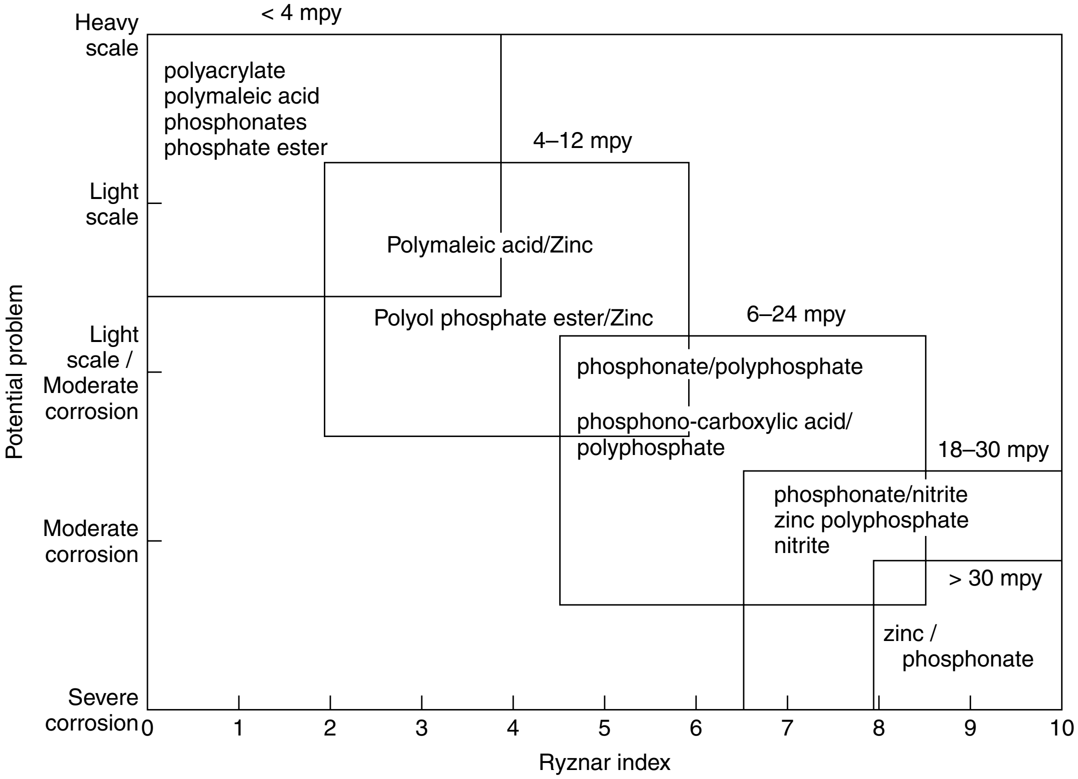

Figure 3. Scale and corrosion control based on the Ryznar index.

this purpose include organophosphorous compounds and polyelectrolytes.

General

Treatment programs for scale prevention are likely to include a range of chemical additives so that a number of techniques are working in parallel to control scale formation and also for corrosion control and biofouling mitigation.

Figure 3 is a diagram published some years ago that demonstrates how chemical additives may be used to control scaling and corrosion for different Ryznar indexes (5).

BIBLIOGRAPHY

1. Bott, T.R. (1995). *Fouling of Heat Exchangers*. Elsevier Science, Amsterdam.
2. Kemmer, F.N. (1988). *NALCO Water Handbook*. McGraw-Hill, New York.
3. Langelier, W.F. (1936). The analytical control of anticorrosion water treatment. *J. Am. Water Works Assoc.* **28**: 1500–1521.
4. Ryznar, J.W. (1944). A new index for determining the amount of calcium carbonate formed by water. *J. Am. Water Works Assoc.* **36**: 472–486.
5. Atkinson, A., Gerrard, A.F., and Harris, A. (1979). Development of additives to control inorganic fouling in industrial water systems. In: *Fouling Science or Art? Inst. Corr. Sci. and Inst. Chem. Engrs.* University of Surrey.

ECONOMICS OF INDUSTRIAL WATER DEMANDS

Steven J. Renzetti
Brock University
St. Catharines, Ontario,
Canada

This article examines what is known regarding the economic characteristics of industrial water use. Understanding the economic dimension of industrial water use is important for several reasons. First, in some countries, industrial water use is already an important part of total water use. Furthermore, in a number of countries, the proportion of total water use accounted for by industry is rising rapidly. Second, much of industrial water use results in significant alterations in water quality. Third, industrial water use differs from water use in other sectors because many firms recycle water and because many firms may choose their source of water supply.

INTRODUCTION

Manufacturing firms, mines and power generating plants typically withdraw water from the natural environment for three main reasons. First, water is applied in a production process. This includes cleaning and moving intermediate inputs as well as inclusion in final output (the production of beer). Second, water is used to cool

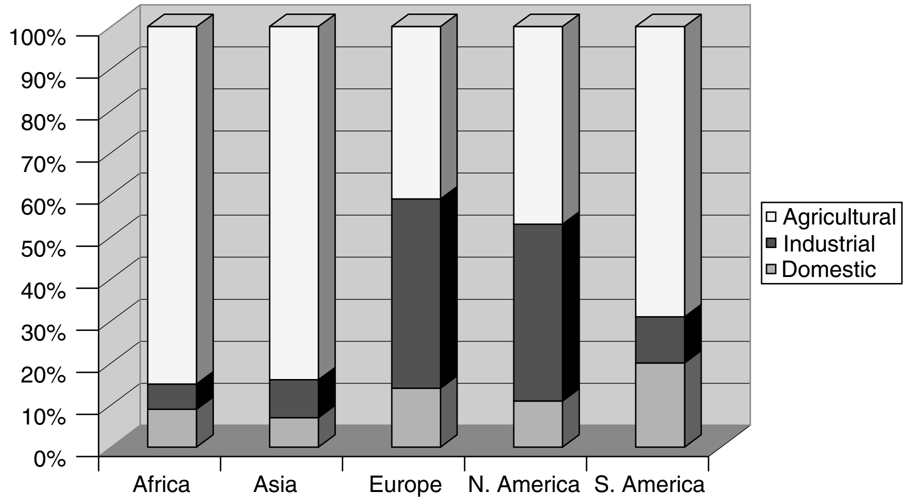

Figure 1. Sectoral shares in water use (1990) (1).

intermediate inputs (the production of petroleum-based fuels) or to produce steam (thermal electric generating plants). Third, water is used for a variety of miscellaneous purposes such as plant cleaning and personal sanitation.

The share of national water use accounted for by industrial water use (including water use in electrical power production) ranges from less than 5% in a number of low-income countries to more than 70% in a number of northern, high-income countries (1) (Fig. 1). Furthermore, in a number of countries experiencing rapid economic development and industrialization such as China, Indonesia and India, industrial water use's share is rising quite rapidly (2).

The next section of this article outlines briefly the economic theory that supports modeling industrial water use. The third section reviews the body of empirical industrial water demand models and the results that flow from this set of work. The last section concludes and provides suggestions for future research. Any brief survey must be selective in its coverage. A number of topics that are related to industrial water use are not discussed here. These include water-based commercial transportation, industrial wastewater discharges, and the valuation of industrial water use. These topics are covered in other articles found in this *Encyclopedia*. In addition, readers interested in a more detailed treatment of the issues surrounding the economic features of industrial water use may refer to Young and Gray (3), Stone and Whittington (4), Gibbons (5), and Renzetti (6,7).

The Economic Theory of Industrial Input Demands

Until recently, it was commonplace to assume that the amount of water required to produce a given quantity of output was fixed (8). In addition, aside from minor variations from differences in technology and the age of capital stocks, it was assumed that the ratio of water use to output was not sensitive to changes in the price of water or the prices of other inputs. In contrast, economists who believe that profit-maximizing firms' use of any input is sensitive to its own price, the prices of inputs, and the price of output have criticized this perspective. To see this, suppose that a firm's technology is represented by a production function $y = f(x_1, \ldots x_N)$ that shows that maximal amount of output (y) that can be produced from input quantities ($x_1, \ldots x_N$). Assuming that the firm faces a price of output (p) and prices of inputs ($w_1, \ldots w_N$) and that the firm seeks to choose output and input quantities to maximize its profits, then the firm's problem may be cast as the following:

$$\max_{\{y,x_i\}} p \cdot y - \sum_i w_i \cdot x_i, \quad i = 1 \ldots N$$

Carrying out this optimization yields the firm's profit-maximizing output supply and input demand equations (Ref. 7, chap. 4; 9):

$$\begin{aligned} y &= g(p, w_1, \ldots w_N), \\ x_i &= h(p, w_1, \ldots w_N), \quad i = 1 \ldots N \end{aligned}$$

It is clear from this presentation that the demand for any input (including intake water) is a function of its own price, all other input prices, the price of output, and the firm's technology. Only in exceptional circumstances should it be expected that the ratio of water intake to output be invariant to prices. Economic theory predicts that both increases in the price of output and decreases in the price of intake water will increase the demand for water. The impact of changes in other input prices on the demand for water will depend on whether those inputs are substitutes or complements to intake water in the firm's production process. Economists typically report these types of relationships as an "elasticity." For example, the own-price elasticity of intake water demand is the change in the quantity of water intake used by the firm in response to a small change in the price of intake water (both changes are expressed in percentages). The strengths of all of these relationships (that is, the magnitude of the elasticities) is a matter for empirical investigation.

EMPIRICAL MODELS AND RESULTS

The main purposes for estimating or calibrating empirical models of industrial water use are to understand water's role in production processes; to examine the relationship between intake water, internal water recirculation, and other inputs; and to measure the sensitivity of water demands to prices. These models may be used to forecast water demands as well as predict the sensitivity of industrial water uses to policy measures aimed at promoting water conservation. There are two main obstacles standing in the way of this type of industrial water demand research. The first is the fact that much of industrial water intake is self-supplied rather than being supplied by a municipal or regional water utility. The result of this is that many industrial water users face either no price for their intake water or only an administrative price that bears little resemblance to the value of water. The second obstacle stems from the lack of data on industrial water intake, recirculation and discharge. Few countries regularly collect and report these data (9).

Despite these challenges, a number of researchers have considered the structure of industrial water demands. The modern literature concerned with industrial water demands begins with a major research effort undertaken by analysts at Resources for the Future (10–12). These studies are noteworthy for their careful application of economic principles and measurement techniques and combining of econometric and engineering methods to represent firms' water use patterns. Since that time, researchers have employed econometric and programming models to examine industrial water demands. Both of these approaches are discussed briefly here.

Initial econometric efforts use ordinary least squares to fit linear equations to plant-level data (13,14). The data included observations on the price of water, the quantity of intake water, and the level and type of output. Both authors report elasticities whose values vary by industrial sector and use of water. De Rooy finds that price and output elasticity estimates vary according to the use of the intake water: −0.894 and 1.212 (cooling); −0.354 and 1.359 (process); −0.590 and 1.243 (steam generation).

An important issue in these and subsequent studies concerns the specification of the price of intake water. For publicly supplied firms, industrial water prices are typically a combination of a connection fee and a complex rate schedule (9). When price is a function of quantity, this creates the possibility of introducing a simultaneity bias into the estimated demand equation. Ziegler and Bell (15) and Renzetti (16) address this issue by estimating demand equations using an instrumental variable technique.

An even more serious challenge arises from industrial water users who are self-supplied. In this case, there may be no external price for intake water. Self-supplied firms incur several types of cost for their water use. First, the firm incurs internal costs resulting from pumping, treatment prior to use (chlorination, descaling, etc.), and on-site storage. Second, the firm may face external costs arising from fees and charges set by government. Third, the firm may incur internal costs for water disposal such as treatment, removal of effluents, pumping, and on-site storage. It is very likely that information regarding the first and third sources of cost will not be publicly available and firms may not even collect this information themselves. This means that it is difficult to specify and measure the cost of an incremental increase in water intake. Grebenstein and Field (17) address this challenge by relying on water utility prices to proxy self-supplied firms' water costs, whereas Renzetti (16) relies on plant-level observations on water utility prices, government water charges, and internal water use costs to construct a proxy for the price of water.

A number of authors (16–19) have used a system of equations approach to estimate the industrial demand for water. This increases the efficiency of estimation and allows the authors to examine the relationship between water and other inputs. Each of the authors assumes that the technology of the manufacturing sector may be represented by a cost function that includes the prices of capital, labor, energy, materials, and water as well as the level of output, as explanatory variables. A system of input demands is then derived from the cost function and estimated. Babin, Willis, and Allen find that the own-price elasticity of intake water ranges from 0.0 to −0.66 and an estimate of −0.66 for all sectors combined. Water's relationship to other inputs also varies across sectors. Water is a substitute for capital in some sectors and is a complement in others. Renzetti (16) finds that water intake's price elasticities range from −0.153 to −0.589, depending on the sector. Water recirculation is seen as a substitute for intake for all industries. Industries that are heavy water users (paper and wood products, chemicals, petroleum, food and beverage) also are the most sensitive to price changes.

More recently, Dupont and Renzetti (20) take a different approach in examining industrial water demands. The authors incorporate water recirculation and technological change in their estimation model and examine the impact of water use regulations on water demands. Because self-supplied firms in Canada must obtain a permit before withdrawing water, the authors consider whether firms are free to choose the optimal quantity of intake water (as is assumed in deriving input demand equations). Two alternative models are estimated: one assuming that water intake is a variable input and the other assuming that water intake is fixed. Specification tests support the choice of modeling water as a variable input. Water intake is found as a substitute for energy, labor, capital, and water recirculation. The relationship between water intake and recirculation is stronger when water use is process-related rather than related to cooling and steam production. Technological change has been biased in the direction of increased water intake and decreased water recirculation.

An alternative to using econometric methods to model water demands is to use linear programming techniques. Linear programming is a technique that is employed to find optimal solutions to problems when the objective function is linear and constraints are expressed as linear inequalities. For example, the objective function may measure a firm's revenues, and the constraints may reflect production processes that relate combinations of input quantities to levels of output. The solutions to linear programming problems have several interesting features that differentiate them from the estimated econometric demand models discussed before. For example, input demand equations derived from linear programming models are typically piecewise, continuous and linear. Discontinuities in the demand equation correspond to the adoption of new production processes (such as the installation of a cooling tower that permits recirculation and reuse of cooling water). Simulations conducted by Calloway, Schwartz, and Thompson (21) demonstrate the significance of this feature. Their results indicate that for intake water prices anywhere below 3.2¢ (1971 US$) per 1000 gallons, a plant producing ammonia employs a water intensive 'once-through' cooling system. At 3.2¢ per 1000 gallons, however, the plant switches to recycling its cooling water, this reduces water intake by 95%, and further price increases do little to encourage further water conservation.

Thompson and Young (22), Singleton, Calloway, and Thompson (23), and Kindler and Russell (24) are representative of this approach. Thompson and Young (22), for example, use engineering data and linear programming techniques to examine water use in thermal power generating stations. They find that there are a number of options for conserving water intake and these options become financially feasible at differing water prices. In addition, the authors find that the power plants can be encouraged to reduce water intake through policy measures such as heat discharge taxes.

Firms that use water in their production processes must find some way to dispose of that water. Water quality regulations typically impose limits on the quantity and/or concentration of effluents that can be discharged. Some analysts argue that water quality regulations have been a driving force in compelling firms to reduce their emissions (25,26). Empirical evidence also suggests that economic forces are relevant. A number of authors (12,16,27) have examined the sensitivity of firms' water discharge decisions to economic factors such as sewage prices, discharge fees, and fines. For example, Renzetti (16) finds that, for the Canadian manufacturing sector as a whole, the price elasticity of water discharge is −0.9752.

The discussion thus far has concentrated on water use in the manufacturing sector. A limited number of studies have considered commercial and institutional water use. The limited amount of work done in this area suggests that this sector's water use is sensitive to water prices and the firm's level of output. Schneider and Whitlatch (28) estimate a series of demand equations to calculate price and income elasticities for several user groups, including commercial, industrial, government, and education. The account-specific price elasticities (following the order of user groups set out before) are −0.918, −0.438, −0.781 and −0.956. Given the diversity of water users in these sectors, generalizations regarding the nature of their water use is quite difficult.

Water use in the mining and electricity power production sectors also has received relatively little attention, maybe because mining water use in most countries is a small share of total estimated withdrawals (it is approximately 1% in the United States and Canada; (26). Despite mining's small share in total withdrawals, water use by mines can be important regionally if they are located in areas where water is scarce such as the American Southwest or the western provinces of China. Mining water use can also be important because of the water contamination that frequently results from mining operations. More attention has been paid to water use in electricity production. The primary concern here, however, has been in establishing the value of water in this use, as opposed to estimating water's role in this technology (29). Young and Gray (3), Gibbons (5), and Renzetti (7) survey the methods used to estimate the value of water use in electricity production.

CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

Industrial water use is important for several reasons: in a number of countries, it makes up a significant part of total water use; firms derive significant benefit from its use, it has the potential for causing substantial alterations in water quality, and it often involves decisions by the firm regarding the degree of internal recirculation to undertake. For these reasons, a number of researchers have sought to understand the economic character of industrial water use. The available empirical evidence indicates quite clearly that decision-making with respect to water does not differ significantly from that for other inputs. Thus, industrial water use is sensitive to its own price, the price of output, and other input prices. Future directions for research will include efforts to combine econometric and engineering process models of industrial water use. These approaches are quite complementary: process models are very detailed and satisfy materials balance constraints, and econometric models are based on a range of technologies and market prices. Another area that remains understudied is the set of factors that influence firms' decisions regarding whether (or how much) to recirculate water.

BIBLIOGRAPHY

1. World Resources Institute. (2000). *World Resources 2000–2001*. Oxford University Press, New York.
2. Biswas, A.K. (1997). Water development and environment. In: *Water Resources: Environmental Planning, Management and Development*. A.K. Biswas (Ed.). McGraw-Hill, New York.
3. Young, R.A., Gray, S.L., Held, R.B., and Mack, R.S. (1972). *Economic Value of Water: Concepts and Empirical Estimates*. Report to the National Water Commission, NTIS no. PB210356 (National Technical Information Service, Springfield, IL).
4. Stone, J. and Whittington, D. (1984). Industrial water demands. In: *Modeling Water Demands*. J. Kindler and C. Russell (Eds.). Academic Press, London.
5. Gibbons, D. (1986). *The Economic Value of Water*. Resources for the Future, Washington, DC.
6. Renzetti, S. (2001). Introduction. In: *Economics of Industrial Water Use*. S. Renzetti (Ed.). Edward Elgar, London, pp. 1–20.
7. Renzetti, S. (2002). *The Economics of Water Demands*. Kluwer Academic Press, Boston.
8. Giuliano, G. and Spaziani, F. (1985). Water use statistics in industry. Experiences from regional surveys and planning

studies in Italy. *Statistical Journal of the United Nations* **3**: 229–235.

9. Dinar, A. and Subramanian, A. (Eds.). (1997). *Water Pricing Experiences: An International Perspective*. World Bank Technical Paper No. 386, The World Bank, Washington, DC.
10. Bower, B. (1966). The Economics of industrial water utilization. In: *Water Research*. A. Kneese and S. Smith (Eds.). Resources for the Future, Washington, DC.
11. Löf, G. and Kneese, A. (1968). *The Economics of Water Utilization in the Beet Sugar Industry*. Resources for the Future, Washington, DC.
12. Russell, C.S. (1973). *Residuals Management in Industry: A Case Study of Petroleum Refining*. Resources for the Future, Washington, DC.
13. Rees, J. (1969). *Industrial Water Demand: A Study of South East England*. London School of Economics Research Monograph 3, London.
14. DeRooy, Y. (1974). Price responsiveness of the industrial demand for water. *Water Resources Research* **10**(3): 403–6.
15. Ziegler, J. and Bell, S. (1984). Estimating demand by self-supplying firms. *Water Resources Research* **20**(1): 4–8.
16. Renzetti, S. (1992). Estimating the structure of industrial water demands: The case of Canadian manufacturing. *Land Economics* **68**(4): 396–404.
17. Grebenstein, C. and Field, B. (1979). Substituting for water inputs in U.S. manufacturing. *Water Resources Research* **15**(2): 228–232.
18. Babin, F., Willis, C., and Allen, P. (1982). Estimation of substitution possibilities between water and other production inputs. *American Journal of Agricultural Economics* **64**(1): 148–151.
19. Dupont, D. and Renzetti, S. (1998). Water use in the Canadian food and beverage industry. *Canadian Journal of Agricultural Economics* **46**(2): 1–11.
20. Dupont, D. and Renzetti, S. (2001). Water's role in manufacturing. *Environmental and Resource Economics* **18**(4): 411–432.
21. Calloway, J., Schwartz, A., and Thompson, R. (1974). Industrial Economic Model of Water use and Waste Treatment for Ammonia. *Water Resources Research* **10**(4): 650–8.
22. Thompson, R. and Young, H.P. (1973). Forecasting water use for policy making. *Water Resources Research* **9**(4): 792–799.
23. Singleton, F., Calloway, J., and Thompson, R. (1975). An integrated power process model of water use and waste water treatment in chlor-alkali production. *Water Resources Research* **11**(4): 515–25.
24. Kindler, J. and Russell, C. (Eds.). *Modeling Water Demands*. Academic Press, London.
25. Brown, G. and Johnson, R. (1984). Pollution control by effluent charges: It works in the federal republic of Germany, Why not in the U.S. *Natural Resources Journal* **24**: 929–966.
26. Solley, W., Pierce, R., and Perlman, H. (1999). *Estimated Use of Water in the United States in 1995*. United States Geological Survey Circular 1200. Available: http://water.usgs.gov/watuse/pdf1995/html/.
27. Sims, W. (1979). The response of firms to pollution charges. *Canadian Journal of Economics* **12**(1): 57–74.
28. Schneider, M. and Whitlatch, E. (1991). User-specific water demand elasticities. *Journal of Water Resources Planning and Management* **117**(1): 52–73.
29. Gillen, D. and Wen, J.F. (2000). Taxing hydroelectricity in Ontario. *Canadian Public Policy* **26**(1): 35–49.

ELECTRIC GENERATING PLANTS—EFFECTS OF CONTAMINANTS

Robin J. Reash
American Electric Power, Water & Ecological Resource Services
Columbus, Ohio

INTRODUCTION

Environmental pollution incidents concerning electric generating facilities have often been widely publicized. Some examples of these include choking air emissions, catastrophic coal ash spills, temporary releases of radioactivity, and contamination of groundwater. In some cases, a direct threat to human health was the cause for concern. In the movie "Erin Brokovich," an activist (played by actress Julia Roberts) draws public attention to the contamination of groundwater by potentially toxic trace metals that have leached through a power plant ash disposal facility. Most of these spectacular pollution events, however, occurred before full, widespread implementation of major environmental statutes (e.g., Clean Air Act, Safe Drinking Water Act, Resource Conservation and Recovery Act, Clean Water Act), enacted by the U.S. Congress during the period 1970–1980. These statutes provide minimum (baseline) environmental protection requirements for the release of pollutants to air, water, and land. Although most states have implemented the requirements of these statutes for several years, concerns regarding pollution released from electric generating facilities still remain in some locations.

This article concerns contaminants released or discharged to the water environment by electric generating facilities. The focus of this article will be on contaminants originating from the combustion and use of fossil fuels. Releases from nuclear-powered facilities will not be considered, largely because the only significant contaminant released by these stations is heated water. The pathway by which power plant contaminants affect aquatic ecosystems may be either direct or indirect. Some pollutants emitted to the air by coal combustion facilities, for example, have the potential to degrade surface waters or affect human health. A general historical review of power plant environmental impacts and environmental control will be discussed, followed by an assessment of sources and effects of pollutants in the water environment.

Historical Review of Power Plant Generation and Contamination

In general, the historical pattern of environmental effects caused by electric generating facilities mirrors the extent and magnitude of fossil fuel (largely coal) combustion. More precisely, the efficiency and extent of pollution control technology in the mining, transport, and combustion of coal and coal-derived materials has been a significant determinant concerning the release of contaminants. The extraction and combustion of coal has a legacy of environmental degradation to air, land, and water. Ecological and human health effects from fossil fuel combustion, which occurred through the first half of the twentieth century, played a significant role in the public's demand for a cleaner environment. Most of these perturbations occurred in the coal-bearing Appalachian region, which is not to say that significant environmental degradation during the first 50–70 years of the twentieth century was limited to fossil fuel combustion; numerous other environmental problems (e.g., discharge of raw sewage, indiscriminant use of persistent pesticides, urban air pollution) affected many industrialized countries during this period. In short, the congressional adoption of the major environmental statutes during the 1970s was driven, at least in part, by the desire to mitigate the environmentally harmful effects of coal mining, transport, combustion, and waste disposal.

As a result of the oil embargo instituted by OPEC countries against the United States during the early to mid-1970s, a national policy of reducing United States dependency on foreign oil was instituted, which resulted in the unprecedented construction of coal-fired (and nuclear-powered) electric generating stations during the mid-to-late 1970s. An increased demand for coal began as the construction of more proposed coal-burning facilities was approved by government regulators. As greater and greater amounts of coal were burned, the magnitude of pollutants emitted (and waste disposed of) began to reach peak levels. During this period, several environmental concerns became widely publicized: acid rain, thermal pollution, particulate matter, acid mine drainage, groundwater contamination, and others. Ultimately, new federal laws and regulations were successful at reducing or mitigating the harmful effects of power plant contaminants.

During the past 30 years, the quality of ambient air and water has improved considerably. This success has resulted largely because of two factors: the adoption of more stringent ambient standards and the requirement to meet best available technology controls for each industry sector. The investment of pollution control technologies and process modifications made by the power industry has not been insignificant. New coal combustion facilities must meet "new source standards" for environmental controls. A new era of "clean coal" technologies is now being implemented. In 2002, combustion of coal (and coal-derived fuels) represented approximately 53% of the U.S. electricity demand (1). This high percentage is likely to remain for some time, as the domestic coal reserve is anticipated to last for about 200 years, considerably longer than reserves for oil and natural gas (2).

It is important to note that the process of coal combustion does not "create" new pollutants. Rather, the extraction and combustion of coal changes the forms, mobility, and environmental compartmentalization of natural constituents in coal. Coal mining mobilizes constituents that normally would be secured in geological strata. When coal is combusted, the natural constituents in this fuel (trace metal, radionuclide, and organic compounds) are transformed into volatile forms (air emissions) or concentrated in ash (solid waste) because of chemical reactions caused by high combustion temperatures.

Air Emission Contaminants

In terms of total mass of pollutants released to the environment, power plant air emissions far outweigh the amount of pollutants released to land and water. Table 1 indicates the types of contaminants released to the air during coal combustion and potential adverse effects on human health and terrestrial/aquatic receptors. Depending on site- and region-specific fate and transport processes, many contaminants emitted by fossil fuel combustion can eventually transfer to the aqueous (water) medium; contaminants such as mercury and selenium do have the potential to be bioaccumulated in surface water

Table 1. Typical Contaminants Released to the Air During Fossil-Fuel Combustion, and Potential Human Health and Environmental Effects

Contaminant Class	Representative Contaminants	Potential Human Health and Environmental Effects
Gases	CO, SO_2, SO_3, NO_2, O_3	Respiratory irritants; atmospheric warming; acidic wet deposition; eutrophication of surface waters; increased ozone and haze
Trace elements	As, Ba, Cd, Cr, Pb, Hg, Ni, Se, Tl	Respiratory irritants; bioaccumulation in crops; bioaccumulation in terrestrial and aquatic food webs
Radionuclides	Isotopes of U, Th, Ra	Radiation exposure via inhalation, dermal, or ingestion exposure routes
Carbonyl compounds	Formaldehyde, acetaldehyde	Respiratory irritants
Aromatic hydrocarbons	Dioxins, phenol	Potential carcinogenic effects; respiratory irritants; food chain bioaccumulation
Acids	Hydrofluoric acid, hydrochloric acid, sulfuric acid	Respiratory irritants
Fine particulates	Sulfur particulates	Respiratory irritants; eye and throat irritants

systems. Other power plant contaminants emitted by coal-fired power plants return to the soil or are sequestered in biomass (trees, grasses, crops). The potential of these sequestered contaminants to elicit adverse effects in the water environment is typically small or negligible.

In 1998, U.S. EPA issued a report to Congress on the levels, and associated health risks, of hazardous air pollutants (HAPs) emitted by utilities using fossil fuel combustion (3). A total of 67 pollutants were assessed by the agency. The report indicated that, for most toxic contaminants emitted to the air, the inhalation (respiratory) route of exposure was most important. Emissions of mercury from coal-fired utilities was the HAP of greatest potential health concern, because mercury is highly persistent and can bioaccumulate to potential harmful levels in foodstuffs (mostly fish and shellfish). Three HAPs were identified that pose potential health concerns, although considerable uncertainties remain: dioxins, arsenic, and nickel. Dioxin, like mercury, has the potential to accumulate in crops or food web compartments. The mass of dioxin emitted from power plants, however, is very small compared with other HAPs. U.S. EPA concluded that, for the majority of generating facilities studied, the cancer risk from inhalation exposure is estimated to be less than 1 in one million (1×10^{-6}).

Damage to aquatic ecosystems from the atmospheric drift and fallout (deposition) from industrial point sources has been well documented in certain geographic areas. Pervasive ecological damage (caused by deposition of several trace metals and resulting bioaccumulation) was documented in several lakes downwind of the Sudbury, Ontario smelter (4). Fossil-fueled power plants emit approximately 45 tons of mercury each year in the continental United States (5). The transport and fate of mercury emitted from coal-fired facilities is complex, however, and is driven largely by chemical speciation of mercury and local meterological conditions. Some studies have shown a close association between mercury emissions from a power plant and levels observed in proximate environmental compartments [e.g., (6)], whereas others have not [e.g., (7)].

Solid Waste and Water Contaminants

Although some contaminants emitted to air during fossil fuel combustion have the potential to deposit on surface water watersheds to levels that could cause adverse ecological effects, contaminants released from dry ash and aqueous (wet) ash disposal facilities have a greater potential to enter aquatic ecosystems because of the much greater proximity to surface waters and groundwater. Table 2 indicates the contaminants that are typically released to groundwater and/or surface waters as a result of solid and aqueous disposal of coal combustion byproducts.

Coal combustion results in voluminous amounts of ash, which must be disposed. Two types of ash predominate: fly ash and bottom ash. Fly ash (which represents about 80% of the total coal combustion waste) is composed of fine-grained particles that have a variable morphology and consist primarily of an amorphous glassy material. The elemental composition of fly ash is highly variable and directly related to compositional variations in the parent coals and to the operational characteristics of individual power plants. Some elements are concentrated (enriched) on fly ash particle surfaces. Some trace metals are particularly enriched on fly ash particles. Suloway et al. (8) reported that arsenic, selenium, antimony, and lead typically have concentrations in ash that are 2–10 times the concentration in parent coal samples. Depending on factors such as pH, some trace metals may desorb from the ash particles and become soluble in ash ponds (treatment facilities used to precipitate metals, salts, and other constituents prior to discharge). Fly ash leachates may be toxic to aquatic receptors, depending on the concentration and bioavailable fraction.

Table 2. Typical Coal Combustion Byproduct Contaminants Released to Groundwater and/or Surface Waters from Solid and Aqueous Disposal Facilities

Coal Combustion Byproduct Disposal Method	Typical Contaminants	Potential Groundwater/Surface Water Effects
Solid waste- dry fly ash disposal	Trace metals	Leachate may contaminate aquifer if impermeable liner has not been installed. Collected leachate may be discharged to surface waters
Solid waste—FGD Sludge Disposal	Salts (sulfate, chloride, calcium); some trace metals	Leachate may contaminate aquifer if impermeable liner has not been installed. Collected leachate may be discharged to surface waters
Fly ash and bottom ash—aqueous disposal	Trace metals; particulates; extreme pH	Direct toxicity to aquatic life if toxic (bioavailable) form predominates. Suspended particulates may precipitate on stream substrate or lake bottom. Proximal sediments may be toxic to aquatic life. Bioaccumulative metals may cause indirect food web effects (effects on top predators or aquatic-dependent wildlife)

Bottom ash is composed of heavier, agglomerated ash particles that are too dense to be carried in the flue gas. Bottom ash is typically gray to black in color, is quite angular, and has a porous surface structure. Normally, bottom ash is less enriched in trace metals and salts compared with fly ash.

Another type of power plant combustion byproduct that may affect surface waters or groundwater is flue gas desulfurization (FGD) material. FGD systems are installed at power plants to provide enhanced removal of sulfur compounds from flue gas, a process that involves the capture of sulfur gases with sorbents. Limestone (calcium carbonate), lime (calcium oxide), and ammonia are the most widely used sorbents. FGD wastes are typically enriched with various calcium and sulfur-based salts. If released from disposal facilities, these constituents can alter surface water ionic strength and, subsequently, cause ionic imbalances in aquatic life.

Contamination of groundwater is a concern if the affected aquifer is a source of drinking water. If ash disposal facilities are not placed on impermeable geological formations (e.g., clay) or not adequately enclosed with a synthetic liner, leachate materials (salts and/or trace metals) can infiltrate into nearby aquifers [e.g., (9, 10)].

EFFECTS OF CONTAMINANTS—WASTEWATER DISCHARGES

Cooling Water Contaminants

For power plants that use once-through cooling (ambient water withdrawn for heat transfer in the condenser pipes), a few contaminants may be added to the heated water at concentrations that may affect nearby aquatic life. On a periodic basis, condenser pipes are flushed with a strong oxidant (typically chlorine or bromine) to reduce biofouling buildup on the pipes. Biocides may also be used to directly kill small mollusks or other larger fouling organisms. Adverse effects to aquatic life have been demonstrated when chlorine is used for biofouling control [e.g., (11)]. In addition, synergistic effects have been observed when aquatic life is exposed to both chlorinated water and heated water, simultaneously.

Condenser pipes experience chemical corrosion over time. Thus, metals may slough off these pipes and, depending on the available dilution, can be present in potentially toxic amounts. Wright and Zamuda (12) measured elevated levels of particulate and dissolved copper downstream of a once-through cooled power plant in Maryland. These levels were clearly above ambient concentrations.

Chemicals may be deliberately injected into the condenser system to help prevent alloy metal corrosion. These chemicals can be toxic to aquatic life if the dosing regime is not adequately regulated.

Coal Ash Byproduct Contaminants

Some useful reviews of the chemical composition, toxicity, and environmental hazards of power plant ash have been published (13–15). The extensive review by Rowe et al. (15), in particular, provides excellent coverage of the chemical characteristics of coal combustion wastes and the range of documented toxicological and ecological effects caused by contaminants in these wastes. As the chemical characteristics of coal ash differ considerably because of the blend of coals burned, generalizations regarding the environmental hazards of coal ash wastes are only speculative. Site-specific assessments of final coal ash discharges (concentration, dissolved and particulate phases, mitigating characteristics) are needed to make definitive predictions of potential toxicity. Often, basic characteristics of the receiving waters (dilution capacity, buffering capacity, suspended solids concentrations) dictate whether *in situ* adverse effects actually occur.

Numerous studies have been conducted documenting the adverse effects of wet ash discharges. Table 3 summarizes the range of observed adverse effects caused by coal ash wastewater discharges, grouped by type of ecological aberration. See Rowe et al. (15) for a more comprehensive review of adverse effects from the discharge of coal ash combustion wastes to surface waters.

Table 3. Types of Adverse Ecological Effects Caused by Coal Ash Wastewater Discharges, with Selected Literature Examples

Ecological Effect	Selected Examples and Reference Citations
Direct toxicity of seepage/wastewater	Acidic, iron-enriched ash pond seepage caused direct mortality to fish (16)
Simplification/modification of invertebrate or fish community	Structural and functional modifications to stream invertebrates (17, 18)
	Tolerant invertebrates and fish species downstream of fly ash pond discharge (19, 20)
Sediment toxicity/metal accumulation	Elevated levels of metals in sediments influenced by fly ash pond discharge (21, 22)
Trace metal bioaccumulation	All food web compartments with elevated concentrations of some trace metals (21, 22)
	Elevated levels in specific fish tissues (23)
Physiological/biochemical/hematological aberrations	Several assays in affected fish significantly different compared to reference fish (24, 25)
Reproductive toxicity	Top predator fish in power plant cooling lakes (receiving coal ash inputs) showed reproductive toxicity due to selenium (26, 27)

REGULATORY CONTROLS—WASTEWATER DISCHARGES

Power plant wastewaters are currently regulated under technology-based national categorical effluent standards. These standards were promulgated in the Effluent Guidelines and Standards for the Steam Electric Power Generating Point Source Category in 40 CFR 423.12 in 1982 and 1983. These limitations include monthly and daily maximum limitations for pH, total suspended solids, oil and grease, and free chlorine (all regulated outfalls); maximum limitations for copper and chromium for boiler chemical cleaning wastes; and maximum chromium and zinc limitations for cooling tower blowdown. For all outfalls, no detectable amounts of PCBs are allowed. Most power generating facilities are able to comply with these categorical standards, as adequate treatment facilities can be designed well in advance of plant operation.

More problematic for power plant wastewater discharges, in terms of compliance, is the requirement to meet water quality-based effluent limitations (WQBELs). These limitations are based on numeric, not-to-exceed water quality criteria that often differ among states. Water quality criteria are estimates of pollutant-specific concentrations that would not cause adverse effects to aquatic life or humans (based on contact recreation or consumption of fish). In contrast to the fixed, categorical effluent limitations, WQBELs can change with time because water quality criteria change (often becoming lower in concentration, or more stringent). In many instances, power plant facilities have installed costly treatment systems to remove constituents (typically trace metals) in order to meet certain WQBELs.

Like other source categories, power plant wastewaters are subject to complying with whole effluent toxicity criteria and/or limitations. Whole effluent toxicity criteria are meant to protect against unacceptable toxicity in the aggregate wastewater discharge. Typically, discharges are assayed using standardized laboratory testing procedures. If effluent toxicity is high, a facility may be required to identify the causative toxicant and remove (or mitigate) the source of the pollutant.

SUMMARY

As an energy source, the combustion of fossil fuels has been a principal factor in the rapid industrialization of many countries during the twentieth century. The legacy of environmental degradation caused by coal combustion is well documented, from coal extraction activities (mining) to release of contaminants into the atmosphere. The adoption and promulgation of several major environmental statutes, however, has lessened the overall environmental impacts of coal combustion. New coal-fired power plants use coal that has low levels of sulfur, mercury, and other environmentally important constituents. The electric utility industry has made prudent investments in cost-effective pollution control technologies. The pressures to implement new, less polluting technologies will drive the industry to balance use of this abundant energy source with cost-effective environmental controls.

Acknowledgments
Tim Lohner and John Van Hassel provided helpful comments on a draft version of the manuscript.

BIBLIOGRAPHY

1. U.S. Department of Energy, Energy Information Administration. (2002). Electricity InfoCard. Accessed at http://www.eia.doe.gov/neic/brochure/elecinfocard.html.
2. Siegel, J.S. and Temchin, J.R. (1991). Role of clean coal technology in electric power generation in the 21st century. In: *Energy and the Environment in the 21st Century*. J.W. Tester, D.O. Wood, and N.A. Ferrari (Eds.). MIT Press, Cambridge, MA, p. 623.
3. U.S. EPA. (1998). Study of Hazardous Air Pollutant Emissions from Electric Utility Steam Generating Units—Final Report to Congress. EPA-453/R-98-004a. U.S. EPA Office of Air Quality, Research Triangle Park, NC.
4. Bradley, R. and Morris, J. (1986). Heavy metals in fish from a series of metal-contaminated lakes near Sudbury, Ontario. *Water, Air, and Soil Pollution.* **27**: 341–354.
5. EPRI. (2000). An Assessment of Mercury Emissions from U.S. Coal-Fired Power Plants. Technical Report 100608. EPRI, Palo Alto, CA, p. 6–1.
6. Smith, K. and Anderson, W. (1981). Distribution and accumulation of trace metals at a coal-fired power plant and adjacent cooling lake. *Illinois Natural History Survey Bulletin*. **32**: 691–730.
7. Pinkney, A., Logan, D., and Wilson, H. (1997). Mercury concentrations in pond fish in relation to a coal-fired power plant. *Archives of Environmental Contamination and Toxicology*. **33**: 222–229.
8. Suloway, J., Roy, W., Skelly, T., Dickerson, D., Schuller, R., and Griffin, R. (1983). Chemical and Toxicological Properties of Coal Fly Ash. Environmental Geology Notes 105, Illinois State Geological Survey, Champaign, IL, p. 22.
9. Cherkauer, D. (1980). The effect of flyash disposal on a shallow ground-water system. *Ground Water*. **18**: 544–550.
10. Simsian, G., Chesters, G., and Andren, A. (1987). Effect of ash disposal ponds on groundwater quality at a coal-fired power plant. *Water Res.* **21**: 417–426.
11. Giattina, J., Cherry, D., Cairns, Jr., J., and Larrick, S. (1981). Comparison of laboratory and field avoidance behavior of fish in heated chlorinated water. *Trans. Amer. Fisheries Soc.* **110**: 526–535.
12. Wright, D. and Zamuda, C. (1991). Copper contamination in the Patuxent River, Maryland. *Hydrobiologia*. **215**: 31–41.
13. Carlson, C. and Adriano, D. (1992). Environmental impacts of coal combustion residues. *J. Environ. Qual.* **22**: 227–247.
14. Theis, T. and Gardner, K. (1990). Environmental assessment of ash disposal. *Crit. Rev. Environ. Contr.* **20**: 21–42.
15. Rowe, C., Hopkins, W., and Congdon, J. (2002). Ecotoxicological implications of aquatic disposal of coal combustion residues in the United States: A review. *Environ. Monitor. Assess.* **80**: 207–276.
16. Coutant, C. et al. (1978). Chemistry and biological hazard of a coal ash seepage stream. *J. Water Pollution Control Fed.* **4**: 747–753.
17. Van Hassel, J. and Wood, K. (1984). Factors affecting aquatic macroinvertebrates below a fly ash pond discharge. *Freshwater Ecol.* **2**: 571–585.
18. Specht, W., Cherry, D., Lechleitner, R., and Cairns, Jr., J. (1984). Structural, functional, and recovery responses of

stream invertebrates to fly ash effluent. *Canadian J. Fisheries Aquat. Sci.* **41**: 884–896.

19. Cherry, D., Guthrie, R., Sherberger, F., and Larrick, S. (1979). The influence of coal ash and thermal discharges upon the distribution and bioaccumulation of aquatic invertebrates. *Hydrobiologia.* **62**: 257–267.
20. Reash, R., Van Hassel, J., and Wood, K. (1988). Ecology of a Southern Ohio stream receiving fly ash pond discharge: changes from acid mine drainage conditions. *Arch. Environ. Contam. Toxicol.* **17**: 543–554.
21. Cherry, D. and Guthrie, R. (1977). Toxic metals in surface waters from coal ash. *Water Resources Bull.* **13**: 1227–1236.
22. Reash, R., Lohner, T., Wood, K., and Willet, V. (1999). Exotoxicological assessment of bluegill sunfish inhabiting a selenium-enriched fly ash stream. In: *Environmental Toxicology and Risk Assessment: Standardization of Biomarkers for Endocrine Disruption and Environmental Assessment.* D. Henshel, M. Black, and M. Harras (Eds.). 8th Vol. ASTM STP 1364. American Society for Testing and Materials, West Conoshohocken, PA, pp. 423–445.
23. Lohner, T., Reash, R., Willet, V., and Rose, L. (2001). Assessment of tolerant sunfish populations (*Lepomis* sp.) inhabiting a selenium-laden coal ash effluent. 1. Hematological and population level assessment. *Ecotoxicol. Environ. Safety.* **50**: 203–216.
24. Lohner, T., Reash, R., and Williams, M. (2001). Assessment of tolerant sunfish populations (*Lepomis* sp.) inhabiting a selenium-laden coal ash effluent. 2. Tissue biochemistry and histochemical evaluation. *Ecotoxicol. Environ. Safety.* **50**: 217–224.
25. Lohner, T., Reash, R., Willet, V., and Fletcher, J. (2001). Assessment of tolerant sunfish populations (*Lepomis* sp.) inhabiting a selenium-laden coal ash effluent. 3. Serum chemistry and fish health indicators. *Ecotoxicol. Environ. Safety.* **50**: 225–232.
26. Olmsted, L., Degan, D., Carter, J., and Cumbie, P. (1991). Ash basin effluents as a concern of fisheries managers: a case history and perspective. In: *Reservoir Fisheries Management: Strategies for the 80's*. G. Hall and M. Van Den Avyle (Eds.). American Fisheries Society, Bethesda, MD, pp. 261–269.
27. Baumann, P. and Gillespie, R. (1986). Selenium bioaccumulation in gonads of largemouth bass and bluegill from three power plant cooling reservoirs. *Environ. Toxicol. Chem.* **5**: 695–701.

ENERGY DISSIPATION

T. R. Bott
University of Birmingham
Birmingham, United Kingdom

Many process industries, such as oil refining or chemical manufacture, need to cool products or intermediate streams that may be either liquids or gases, before storage or further processing. To conserve heat and hence costs, it is usual to transfer as much heat energy as possible from an outgoing "hot" stream to an incoming cold stream or intermediate. However as the temperature of the "hot" stream is reduced, there comes a point where it is no longer economical or feasible to reduce the temperature of the stream to be cooled by further interchange of heat. As the temperature of the "hot" stream approaches the temperature of the incoming stream, the temperature driving force for heat transfer is reduced so that a relatively large heat transfer area is required to complete the cooling process. In general, naturally occurring water has a relatively low temperature and because there is so much of it in the world—in seas, lakes, rivers, canals, and boreholes—water is used as a universal cooling medium. The increasing world shortage of water, however, requires effective water management to conserve supplies and to protect the environment.

Apart from process industries, cooling is essential for efficient use of fuel for power generation. In the power industry, steam is usually made by combusting fossil fuel, biomass, or flammable waste, such as domestic refuse. The steam produced is used to power turbines, which drive electricity generators. It is possible to arrange for the exhaust steam from the turbines to be discharged at atmospheric pressure, but if the exhaust steam is condensed, the pressure drop across the turbine is increased so that additional energy from the steam can be converted into electrical energy. Heat removal from the steam that reduces the steam pressure is usually achieved in water cooled condensers, adding significantly to the overall efficiency of the power station in converting the chemical energy stored in the fuel to electrical energy.

The use of water for cooling purposes has several advantages. It is generally plentiful, although in some arid parts of the world it is not available in adequate quantities. Under these circumstances, air cooling is likely to be preferred. Water has a relatively high heat capacity and a relatively wide operating temperature range. It is easily transported around a cooling circuit by pumping. It is reusable, provided that its temperature, after use as a coolant, can be reduced and it is suitably treated to prevent operational problems that would otherwise occur.

COOLING WATER SYSTEMS

Three cooling water systems are used in industry: enclosed recirculating, open recirculating, and "once through" systems.

Enclosed Recirculating Systems

In the enclosed recirculating system, the water is totally enclosed and isolated from the atmosphere. In operation, the water takes heat from the process stream or condensing steam in a suitable heat exchanger and is then pumped through air blown coolers. Usually the tubes in these coolers are finned on the air side to improve the heat transfer from the water to the air stream. The air passes through the cooler by natural draft or is forced through by a suitable fan. The air may be drawn through the heat exchanger by a fan mounted on top of the exchanger, in which case it is known as an "induced draft" cooler. Alternatively, a fan mounted below the cooler is called a "forced draft" system. Provided that the water is suitably treated to avoid deposition and corrosion problems and the water quality is monitored frequently, the enclosed system can operate for long periods of time with a minimum of attention. The system is not common, however, at least in temperate zones, because of the additional capital cost involved, largely due to the fact that the heat is removed from the water as sensible heat. It is more usual to remove the heat gained by the cooling water by evaporating some of the water to the atmosphere in an open recirculating system. The removal of latent heat necessary to vaporize the water lowers the water temperature.

Open Recirculating Systems

It is possible to effect the necessary evaporation in a "spray pond." Here, water is sprayed into the atmosphere out of nozzles mounted on horizontal pipes. The pipes themselves are mounted in a concrete "pond." The evaporating droplets of water fall back into the pond for reuse in the recirculating system. The desired cooling can be achieved by adjusting the spray conditions.

In an open recirculating system, however, it is more usual to employ a cooling tower to evaporate some of the recirculating water, as illustrated in Fig. 1. In this technique, the water falls over suitable packing contained in the cooling tower. For many years, the packing in cooling towers consisted of a large number of wooden slats over which the water trickled as a film. The purpose of the packing was to maximize the contact area between the airflow and the water film. The use of wood was far from ideal because the damp wood and the favorable temperature conditions led to the growth of fungi on the packing and this, as well as the increased weight, tended to weaken the integrity of the wooden structure. Failure of the wooden structure was not uncommon in the past. In recent times, plastic materials that resist fungal growth have been used to fabricate packing. The use of these materials, molded to shape, has allowed more sophisticated designs of packing to be introduced, which maximize evaporation of water in the tower under given operating conditions.

It is usual to maintain the airflow by natural convection, as illustrated in Fig. 1. It is possible, however, to have special designs where the airflow is driven by a motorized fan, as described for the enclosed cooling system to create induced or forced circulation.

In the cooling tower, water is lost from the system by evaporation and what is known as "windage," as illustrated in Fig. 2. The windage is made up of droplets of water that leave the top of the tower and are often seen as a plume downwind from the tower. Water loss also occurs from any leakage from the system and from "drift loss" in the region of the cooling tower basin. The water used in open recirculating systems is usually taken from natural sources such as lakes, rivers, or canals and therefore contains dissolved solids in addition to biologically active material, such as microorganisms.

As water is removed from the system by evaporation, the concentration of the dissolved solids increases. Therefore, it is necessary to discard some of the recirculating water, usually back to the source, and replace it with "makeup" water from the source, which has a lower concentration of dissolved solids. The deliberate discharge of water is usually referred to as "blowdown." It may be continuous or intermittent, depending on the rate of buildup of dissolved solids. The "makeup" water (M) has to replace that lost from the system by all mechanisms and from Fig. 2,

$$M = E + W + L \tag{1}$$

A concentration factor n may be defined by

$$n = \frac{\text{Dissolved solids concentration in the blowdown water}}{\text{Dissolved solids concentration in the makeup water}} \tag{2}$$

In general n is in the range of 3–5, but it depends on the operating conditions in the cooling tower.

Blowdown B may be calculated ignoring windage and leakage water losses which are usually small compared with the evaporation, so that

$$B = \frac{E}{(n-1)} \tag{3}$$

Once Through System

In the once through system, water is taken from the source, passed through the system, and discharged back to the source without cooling. The technique reduces the capital cost of the installation, but there may be environmental problems associated with the temperature of the discharge water in relation to the temperature of the water source, a problem which is usually referred to as "thermal pollution." Local regulations may restrict the allowable thermal pollution of the natural source

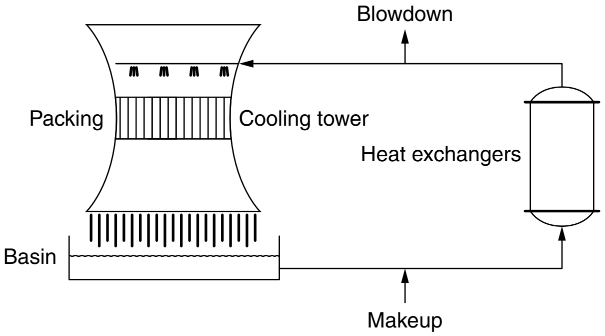

Figure 1. Recirculating cooling water system with cooling tower.

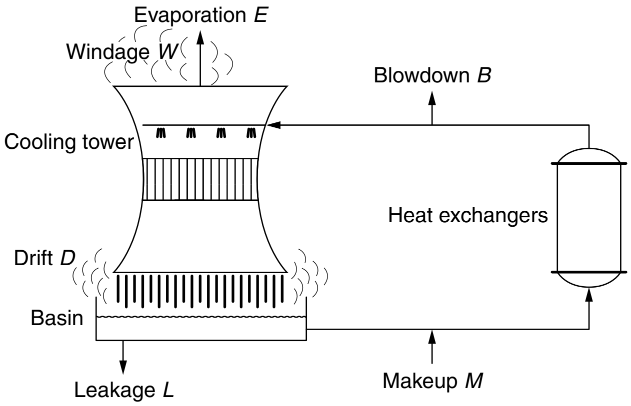

Figure 2. Water loss from a cooling tower circuit.

on account of its possible effect on the ecology of the water source.

Water Treatment

The origin of the water from a natural source in cooling water systems means that it can give rise to operating problems, generally associated with the heat transfer surfaces. Fouling may result from particulate deposition, microbial growth, or corrosion. These phenomena, the factors which influence them, and the methods available for control, are described elsewhere in this encyclopedia. Fouling is often controlled by chemical additives. Of general concern in open recirculating systems employing cooling towers or spray ponds is the presence of *Legionella* bacterium in the circulating water. Unless adequate precautions are taken, the bacterium may be present in the plume of water droplets, which leaves the top of the cooling tower or in the droplets above a spray pond. As a result, personnel who happen to be downwind of the tower or spray pond are susceptible to infection from the contaminated water. It is imperative for health and safety, quite apart from other considerations, to monitor the quality of the water within the system frequently and regularly.

CONCLUDING REMARKS

It is likely that the use of water for energy dissipation will continue in the future, but with greater emphasis on environmental protection and health and safety. It will increasingly be subject to regulation. To comply with the demands of legislators, the costs of operating the system are likely to be substantial. Effective management is necessary to keep the costs to a minimum.

WATER USE IN ENERGY PRODUCTION

Herbert Inhaber
Risk Concepts
Las Vegas, Nevada

The United States uses far more water to produce energy and electricity than any other sector. A slight reduction in water use in the electricity sector would far outweigh other modes of saving—toilets using less water, low-flow shower heads,—in the household sector.

Most attention here will be devoted to thermoelectric and hydroelectric plants. Thermoelectric plants can be divided into fossil fuel—coal, oil, and natural gas—and nuclear. These constitute almost all the electricity producers in this country.

However, this may not always be the case in the future. Some government reports and projections have said that renewables—solar, wind, geothermal, ocean thermal—will be used more in coming years. The so-called Kyoto Protocol (unratified at this time) anticipates a decrease of thermoelectric plants (due to their emissions of carbon dioxide, a greenhouse gas) and increased use of renewables. For example, in Denmark now about 10% of its electricity is supplied by windpower.

Water use is usually defined and measured in terms of withdrawal or consumption. The former is what is taken, and the latter is that which is used up. Water withdrawal in energy production is much greater than consumption. In hydroelectric plants, far more water is used than in any other energy source. However, the amount of water consumed—that is, not available for further use—is a small fraction of the water used and is attributable to evaporation from lakes behind hydro dams. Energy consumption for thermoelectric plants is of the order of 2% of water used.

Some of the water used to produce energy becomes contaminated with pollutants. This aspect of energy production is not discussed in this article; it is treated in detail in other articles in this publication.

Units

Units in the discussion of energy water use are both English and metric. The graphs here use metric units. Where the original form is English, metric units follow in parentheses, using the conversion factor $1\ \mathrm{m}^3 = 275$ gallons.

OVERALL WATER USE IN ENERGY PRODUCTION

More water is used for energy production than in any other sector in the United States, according to the Geological Survey (1). Of the 402,000 million gallons ($1460 \times 10^6\ \mathrm{m}^3$) per day of withdrawals, 189,000 million (690×10^6), or 47%, was used in thermoelectric plants. Irrigation was second at 134,000 million gallons ($490 \times 10^6\ \mathrm{m}^3$) per day.

During 2000, total U.S. net generation of electricity was $3{,}800 \times 10^9$ kWh (2). Of this, coal generated 1,967 (in the same units), or about 52%. This was followed by 754 from nuclear (20%), 612 from gas (15%), 273 from hydro (7%), 109 from petroleum (3%), and 84 from renewables (2%). Total fossil fuel electricity production was 2,688 billion kWh.

Of the renewables in the first six months of 2001, 0.37% of total production was from geothermal, 1.7% from biomass, 0.19% from wind, and 0.018% from solar (2).

CONVENTIONAL ENERGY SYSTEMS

Thermoelectric Water Use and Consumption

Most water withdrawn for thermoelectric power production is used for condenser and reactor cooling. Consumption of water was much less, at about 3,700 million gallons ($13.5 \times 10^6\ \mathrm{m}^3$) per day, about 2% of the total amount of water withdrawn.

What is the rate of water withdrawal for the four types of thermoelectric plants noted before? Official data are divided into fossil fuel and nuclear (3). It is not clear how geothermal data are handled, but this forms only a small proportion of total electricity production. Fossil fuels have a total withdrawal of 135,000 million gallons ($490 \times 10^6\ \mathrm{m}^3$) per day and consumption of about

2,800 (10×10^6 m^3) in the same units. For nuclear, the corresponding values are 54,500 (198×10^6 m^3) and 900 (3.3×10^6 m^3).

It would be of interest to determine (1) water use per unit of energy generated for each of the three fossil fuels—coal, gas, petroleum—compared to nuclear, (2) how the values varied within each type of energy, and (3) the relationship of water use to thermal efficiency of each type, but the data are not available in the official document.

However, the data above show that fossil fuels have a withdrawal of 67 m^3 per megawatt-hour (MWh) and consumption of 1.4 in the same units. For nuclear, the corresponding values are 97 and 1.6, respectively. The results are shown in Fig. 1. Hydroelectric values, to be discussed later, are also included.

Figure 1 uses a logarithmic scale because of the large amount of water use per unit of energy output for hydroelectricity. Values for water use in nuclear power are higher than for fossil fuels, primarily because of the lower thermal efficiency of the former. However, water consumption for both nuclear and fossil fuels are about the same. The ratio of consumption to water use for hydro, while not specified in official data, is much less than the approximately 2% for thermoelectric plants.

Two major documents dealing with water use in energy production in recent years, other than the Geological Survey document, are Gleick (4) and a German official publication (5). The values shown in the German document for thermoelectric plants using noncirculating water use (6), 160–220 m^3 per MWh, are somewhat higher than those in Fig. 1. Cooling water losses, or consumption, usually stay below 2% of water use, again in accordance with Fig. 1. The document also refers to small water requirements, in addition, of the order of 0.1 m^3 per MWh, for topping up the steam cycle, cooling the ashes, and operating certain types of flue-gas purification equipment (spray absorption, wet processes).

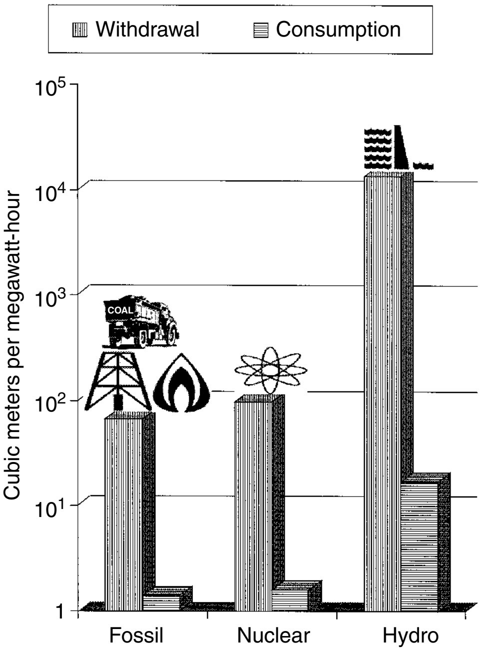

Figure 1. Water withdrawal and consumption rates for fossil fuels, nuclear and hydroelectricity, in cubic meters per MWh of electricity generated.

Reference 7 shows “consumptive water use” (taken to mean consumption as opposed to withdrawal) for the three main fossil fuels plus nuclear in m^3 per 10^{12} joules (thermal). Reference 7 considers many aspects of the fuel cycle, in addition to the generation phase, for each of the four major electricity sectors, that is, an attempt is made to estimate water requirements for the entire fuel cycle. This is in contrast to the Geological Survey and German documents, which concentrate on the generation phase. However, no indication is given in Ref. 7 about which of the components of water use for each technology is in widespread use.

For example, under nuclear, nuclear fuel reprocessing is shown as requiring about 50 m^3 of consumption per 10^{12} joules (thermal), the largest, by a factor of 2.5, of any of the nuclear components. But this aspect of nuclear technology is not in use in the United States.

Similarly, for the oil cycle, six possibilities are shown for enhanced oil recovery. The highest, micellar polymer, has the largest water consumption of any of the components for the four major technologies listed in the table. But it is not stated how common this technology is.

A later table (8) in Ref. 4 shows the average water consumption per megawatt-year for just the generating phase of the four major electricity sources. Results are then on the same basis as that of the official data. The values for fossil fuels with once-through cooling are comparable to those of Fig. 1, so they will not be repeated here. The results for coal and oil-fired generation are shown for both once-through cooling and cooling towers, although most fossil-fuel generation in the United States uses the former cooling method. Cooling towers consume about twice the amount of water of once-through cooling. The results in Ref. 8 for coal, oil, and gas are similar, so averaging over the three fossil fuels in Fig. 1 is reasonable.

Nuclear Power Water Use

Reference 3 states that nuclear plants use and consume somewhat higher amounts of water per MWh produced than fossil fuel plants. However, there are some proposed models of nuclear plants that will use and consume less. For example, the so-called MHTGR (modular high temperature gas reactor), due to its greater thermal efficiency, should require both less water use and consumption than other types of nuclear plants (9). At present, the MHTGR is not deployed in the United States.

In Ref. 8, only cooling towers are shown for nuclear generation, although most nuclear plants, like fossil-fuel plants, use once-through cooling. This produces a consumption value of about twice the official data, and tends to make nuclear look more water-consuming than is the case.

Hydroelectricity Water Use

Hydroelectricity produced about 6% of all U.S. electricity in the first 8 months of 2001. The proportion varies slightly from year to year, depending on rainfall, runoff, and other factors. In some countries, such as Canada and Norway, the proportion is much higher. Hydro is clearly a renewable system, but it is included in this section because it is conventional and widely used.

Water use for hydroelectric purposes is by far the largest of all U. S. energy sources, at 3,160,000 million gallons ($11{,}500 \times 10^6$ m^3) per day. This is about 2.6 times the total of the total annual runoff of the lower 48 states (10). This refers only to the withdrawal of water, the amount that passes through lakes and hydro dams.

Water is consumed, or evaporates, from the lakes behind dams. One set of estimates is from Ref. 11. The range of values for California dams is enormous, a factor of about 5,000 between the most water-conserving dams (measured in water consumption per MWh) and the most profligate. The variation will depend on such factors as the size of the lake, dam height and type, contours of the land, and outdoor temperatures.

For the United States, the average value is about three times that of the California average (12). The results are shown in Fig. 1. Average water consumption, per MWh, for hydro plants is about one-quarter the total withdrawal of fossil-fuel plants. If the most water-consuming dam had been used as a maximum for hydro, its value would have been greater than the entire water use for fossil and nuclear.

In 1995, 310×10^9 kWh of hydroelectricity were produced. This required 3,160,000 million gallons ($11{,}500 \times 10^6$ m^3) per day of withdrawals. So about 10 gallons (0.036 m^3) day, or 3700 gallons (13 m^3) per year, is required per kWh. The national average cost per kWh (enough to light a 100-watt light bulb for 10 hours) is about 8¢; so this implies that about 450 gallons (about 1.6 m^3) of water are required to produce about 1¢ of hydroelectricity value.

NONCONVENTIONAL ENERGY WATER USE

The key point in estimating water use and consumption by renewable energy systems was stated by the California Energy Commission (13): "In determining whether a particular fuel is 'clean' or not, consideration must be given to environmental impacts across the entire fuel cycle of the type of generation proposed (e.g., including fuel production, transportation, refinement, and generation)."

Renewable energy made up about 2% of all electricity production in the United States in the first 8 months of 2001. This fraction has changed only slightly in recent years. It has been widely claimed that this fraction will increase substantially in coming decades. Whether these claims are true can be verified only by the passage of time. Their present total water use is extremely small. All values presented here are only approximations. As more data are gathered, it is likely that many, if not all, of the values in this section will change. The German study (14) discusses the water implications of renewable energy sources, but nonquantitatively. It does make clear, however, that the entire fuel and energy cycle should be considered, not merely one aspect. This is also the implication of Reference 7.

Material Requirements for Energy Systems

To compare renewables with conventional energy systems in a consistent manner, the ratio of water use, both indirect and direct, per unit of energy or electricity output is required. The place to start is materials requirements for energy systems. The results of one calculation are shown in Fig. 2 (15).

Materials requirements for renewables, per unit of energy output, tend to be much higher than those for conventional energy systems. The main exception is hydro, but most of the requirements there are rock and earth, which require little water use. Solar and windpower generally are very dilute per unit volume compared to fossil and nuclear fuels. Thus considerably more collecting apparatus, again per unit energy output, is required. In turn, the collecting apparatus requires considerable material (and associated labor).

Water Use for Materials in Energy Industries

The next step is to estimate the water use for each of the materials shown in Fig. 2. For steel, the German study noted before estimates 5 m^3 per metric ton (16). This is in accordance with other estimates. On the other hand, "mini-mills generate up to 80 m^3 (of water use) per metric ton of steel product" (17). The value of 5, used here, may be different in non-Western countries. It is estimated that China requires 23–56 m^3 per ton (18).

For aluminum, one estimate is 1.5 tons of water (19) per ton of product. Another detailed look at the aluminum

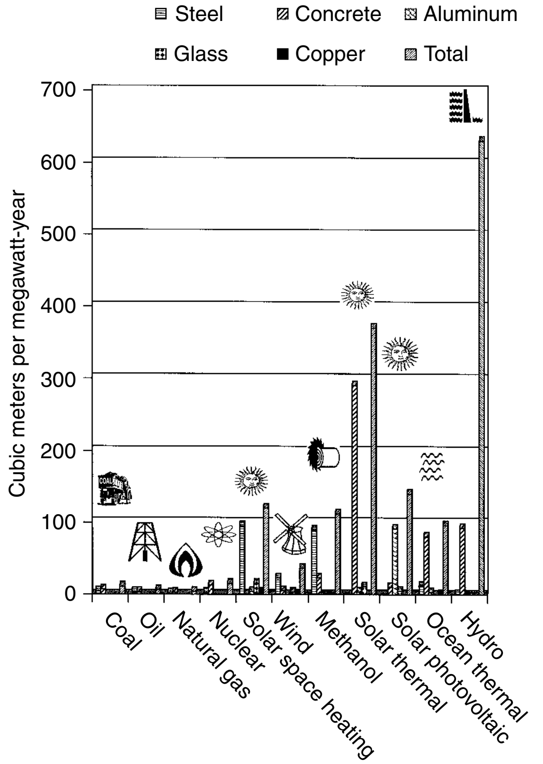

Figure 2. Material requirements for energy systems, per megawatt-year net output. Fuel—coal, oil, gas, and nuclear fuel—is not included.

industry suggests 5 m^3 of process water per metric ton of alumina and 1.88 m^3 of alumina per metric ton of aluminum. This then implies 2.6 m^3 of water per ton of aluminum, and this value is used here (20).

For copper, although water is undoubtedly used in its production, it proved difficult to obtain reliable data. For glass, "The average water consumption in a glass works should be less than 1 m^3/t of glass produced" (21). For cement, about 1–1.6 m^3 of water per metric ton of cement is required, depending on whether the dry or wet process is used (22). Concrete is one part cement to 6–8 parts other materials, so water use is adjusted accordingly.

Noted that values for water in materials production are likely use, not consumption. The data sources do not always make the distinction. The results for water use are as follows in m^3 per metric ton: steel, 5; aluminum, 2.6; copper, not available; glass, 1; cement, 1–1.6; and concrete, 0.1–0.25.

Indirect Water Use for Energy Systems

The indirect water use attributable to the materials used in all energy systems may be estimated by multiplying the values for specific materials in the preceding paragraph by the materials used per unit of energy output in Fig. 2. No values could be found for copper, so it is not included in the calculations. Its inclusion would change results only slightly. Minima and maxima of the results are shown as the bottom two components in the columns of Fig. 3.

Because of the wide differences between both components and the total of each column, the results are shown on a logarithmic scale. In each column, the lowest entry is for the minimum of indirect water use. The next highest entry is for the maximum of this component. Hydro has the highest amount of material use in Fig. 2, but almost all is rock and earth, assumed to have negligible water requirements. For coal-, oil-, and gas-fired, nuclear and hydro plants, values for direct water use are taken from Fig. 1. These are the next highest components of each column. Some technologies, such as solar photovoltaic, have negligible direct water use, so this component for their columns is blank. For solar thermal plants, an estimate of direct water consumption, as opposed to use, is available (23). It is assumed here that the ratio of water use to consumption for thermoelectric plants, approximately 50, also applies to this technology. Solar space heating, although included in Fig. 2, is excluded here because its output is thermal, not electrical.

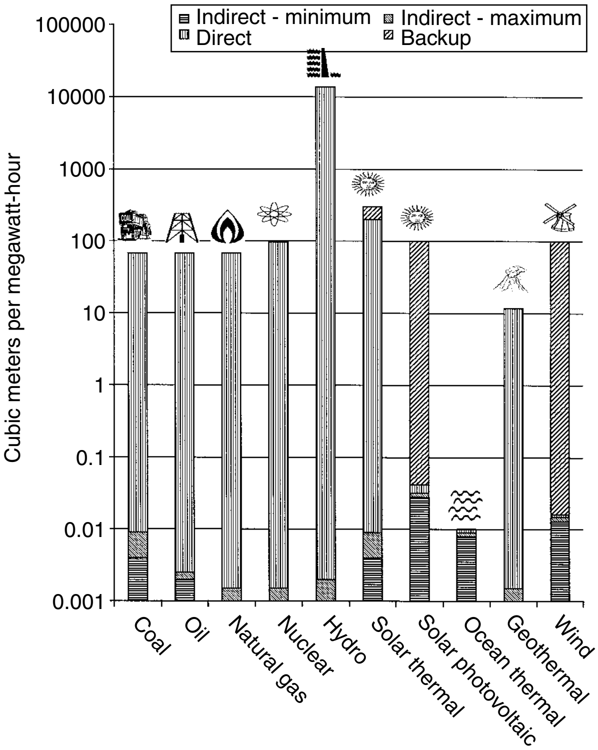

Figure 3. Water use per kWh output for conventional and renewable systems.

For geothermal, only values of consumption, not use, are available (23). Three values, for plants in the United States and New Zealand, have data, and the average was taken. It is likely that water use is higher than consumption for geothermal. One source states (24), "A geothermal plant produces much more wastewater than a coal-fired or atomic power plant," although no data on the proportion seem available. Backup water use, to be discussed in the next section, is shown for three energy systems.

Backup Energy and Electricity

Many articles dealing with renewable energy systems (other than geothermal and biomass) acknowledge the need for backup when the sun does not shine and the wind does not blow (25–29). The assumption is that consumers want reliable power, at least as reliable as present conventional systems. Given the fact that utilities are deluged with complaints after a power outage due to weather or other reasons, this is a reasonable assumption.

As a specific example of backup needed, consider Solar Two, one of the world's largest solar thermal plants. "During one stretch in the summer of 1998, the plant operated for 32 of 39 days", according to a U.S. Department of Energy Website dubbed "success stories" (30). Although this was only a demonstration project, in real life, consumers would be concerned about no power for 7 out of 39 days. Its average energy production was about 20% of its rating; fossil-fuel and nuclear plants routinely operate at 80% or more.

What proportion of backup is necessary? Estimates vary, depending on a host of factors: the variation of sunlight and wind, how much reliability is expected of renewable sources, whether the source is to be for peak power demands, and so on. Sørensen (31) notes that the "fuel-based supply system [for renewables] ... is a tenth of average load," that is, about 10% of conventional—fossil fuel, nuclear, and hydro in national proportions—electricity sources will be assumed as backup. Systems like geothermal, ocean thermal, and biomass require no backup because they produce energy continuously.

This proportion will change from one part of the country to another. For example, the proportion of hydro in the Pacific Northwest is higher than that in other parts. This would tend to make water use and consumption for backup higher there than elsewhere, on the basis of Fig. 1.

Backup would not apply to such renewable sources as biomass and geothermal, which produce power about as reliably as conventional energy systems. Using the values of Fig. 1 and the proportions of electricity production from fossil fuels, nuclear and hydro from earlier, an average national kWh of electricity would require 1.0 m^3 of water withdrawal and 0.0025 of consumption. Even though hydro is only about 7% of national electricity production, it generates about 95% of electricity water withdrawal and about half of water consumption. If an average backup of 10% is assumed, this implies 0.1 m^3 of withdrawal and 0.00025 of consumption per kWh of underlying energy attributable to backup.

SUMMARY

Most attention in water use and consumption has been focused on fossil-fuel and nuclear plants, though renewable energy systems, both present and in the future, can use considerable water when the entire energy cycle is considered. Hydro is classified as a renewable energy source by the Energy Information Administration and many other authors. It uses about 100 times the water and about ten times the consumption of thermoelectric plants, per MWh generated. Thermoelectric and hydro plants produce about 98% of all U.S. electricity. But hydro requires about 95% of all electricity water use and about half of electricity water consumption.

A number of conclusions can be drawn from Fig. 3. First, it clearly is important to take all aspects of the fuel or energy cycle into account when estimating water use and consumption. Second, hydro clearly has the highest water use (and consumption) of any major present or future energy electricity source. Third, although the material requirements estimated in Fig. 1 are useful, they affect overall results only slightly. Fourth, although Ref. 4 suggests that water use by most renewable energy sources will be negligible, Fig. 3 shows that this is not the case. Hydro and solar thermal rank the highest. Fifth, ranges of data are not shown in this figure except for water use associated with materials requirements, but remember that values in most cases are approximate and vary considerably. Sixth, ocean thermal (OT) apparently has the lowest water use of any of the technologies considered. OT uses considerable water in operation, but all of it is saline. Seventh, backup requirements make a considerable difference in water use for some renewable energy systems. Eighth, geothermal ranked second lowest of the systems, but its water use is probably underestimated in the graph, for the reasons explained in the caption.

In summary, to make rational decisions about which systems use more or less water, the entire fuel or energy cycle must be taken into account. Figure 3 shows the results of such calculations.

BIBLIOGRAPHY

1. Solley, W.B., Pierce, R.R., and Perlman, H.A. (1998). *Estimated Use of Water in the United States in 1995*. U.S. Geological Survey Circular 1200, US Government Printing Office, Offstream Use, Washington, DC.
2. Energy Information Administration (2001). *Electric Power Monthly*. November 2001. EIA, Washington, DC.
3. Ref. 1, op cit., *Thermoelectric Power*.
4. Gleick, P.H. (1994). Water and energy. *Annual Review of Energy and the Environment* **19**: 267–299.
5. German Federal Ministry for Economic Cooperation and Development (GFM). (1995). *Environmental Handbook*. Documentation on monitoring and evaluating environmental impacts, Vol. 2. Agriculture, Mining/Energy, Trade/Industry, GTZ GmbH, Eschborn, Germany.
6. Ref. 5, GFM, op. cit., Section 41.2.2.
7. Ref. 4, op. cit., Table 4, *Consumptive Water Use for Energy Production*.
8. Ref. 4, op. cit., Table 5, *Consumptive Water Use for Electricity Production*.
9. Neylan, A.J. Dilling, D.A., and Cardito, J.M. (1991). Environmental aspects of MHTGR operation. *Energy* **16**(1–2): 441–448.
10. Solley et al. op. cit., *Hydro Power*.
11. Ref. 4, op. cit., Table 7. *Average Annual Water Losses from California Hydroelectric Facilities*, in m^3 per 10^3 kWh per year.
12. Ref. 4, op. cit., Table 5.
13. California Energy Commission. (2001). *Energy Innovations Small Grant Program*.
14. Ref. 5, GFM, op. cit., Section 43, Renewable sources of energy.
15. Inhaber, H. (1982). *Energy Risk Assessment*. Gordon & Breach, New York, p. 36. Earlier versions of this study had some errors, which were removed by the time the book was published. About two-thirds of the book is devoted to responses to critics. One of the major critics of the original study later came to similar conclusions about the relative risk ranking of the technologies considered. This was noted by Watson (Watson, A.P. et al. (1983). *Health Implications of Unconventional Energy Production and Use, in Alternative Energy Sources*, Vol. 9, Policy/Environment, Hemisphere Publishing, Washington, DC, p. 249, who stated that “the conclusions of the two analyses [Inhaber’s and that of a major critic] are the same” [in terms of where the greatest health risks are]. A similar point is made by Sims (Gordon, H.E.S. (1990). *The Anti-Nuclear Game*. University of Ottawa Press, Ottawa, Canada) and others, too numerous to list here.
16. Ref. 5, op. cit., Section 49.2.
17. World Bank Group. (1998). *Pollution Prevention and Abatement Handbook*.
18. (1998). *Population Reports* **26**(1): 20.
19. Gleick, P.H. (2001). *Scientific American* **284**(2): 40.
20. (1997). *Energy and Environmental Profile of the US Aluminum Industry*. Energetics, Columbia, MD, pp. 29–30.
21. Ref. 5, GFM, op. cit., Section 48.2.
22. Ref. 5, GFM, op. cit., Section 46.2.
23. Ref. 4, op. cit., Table 5.
24. Found at http://www.geol.binghamton.edu/faculty/naslund/205.lect20c.html (Geology 205, Earthquakes and Volcanoes, Lecture 17, Geothermal Energy). The site was developed by H. Richard Naslund, Professor of Geology, State University of New York at Binghamton.
25. Cao, J. and Christensen, R.N. (2000). Modeling an integral dual solar/gas-fired generator for a water–lithium bromide absorption chiller. *Journal of Energy Resources Technology* **122**(4): 217–223. “Natural gas serves as the backup heat when the solar energy is unavailable or insufficient.”

26. Rhodes, R. and Beller, D. (2000). The need for nuclear power. *Foreign Affairs* **79**(1): 30–45. "A global solar-energy system without fossil or nuclear backup would also be dangerously vulnerable to drops in solar radiation from volcanic events such as the 1883 eruption of Krakatoa, which caused widespread crop failure during the "year without a summer" that followed."
27. Williams, R.H. (1990). Low-cost strategies for coping with CO_2 emission limits. *Energy Journal* **11**(4): 35. "Luz (solar thermal) units are hybrid systems with natural gas firing capacity added as backup"... "PV systems (and wind systems as well) can be used in hybrid configurations with natural gas-based backup power, as in the case of solar thermal units."
28. Carta, J.A. and Gonzalez, J. (2001). Self-sufficient energy supply for isolated communities: wind-diesel systems in the Canary Islands. *Energy Journal* **22**(3): 115. "A wind-driven generator to convert wind energy to electrical energy, back-up diesel generators (for when there is no wind)."
29. Hansen, U. (1998). Technological options for power generation. *Energy Journal* **19**(2): 63. "The principal disadvantage of solar is the dependence on variable sunshine hours and the resulting short capacity utilisation as well as the need for energy storage and back-up systems."
30. Available: http://www.eren.doe.gov/success_stories/opt_solartwo.html.
31. Sørensen, B. (2000). *Renewable Energy: Its Physics, Engineering, Use, Environmental Impacts, Economy and Planning Aspects*. Academic, San Diego, CA, p. 676.

EVALUATION OF TOXIC PROPERTIES OF INDUSTRIAL EFFLUENTS BY ON-LINE RESPIROMETRY

Petros Samaras
Technological Educational
Institute of West Macedonia
Kozani, Greece

Industrial effluents often contain chemical compounds that have potential toxic properties which upon co-treatment with municipal wastewaters may affect the operation of a wastewater treatment plant. Respiration measurements, corresponding to the estimation of oxygen uptake rate (OUR) of living microorganisms, may be used to evaluate the toxic properties of industrial wastewaters, estimate their biological treatability, and predict the impact on the performance of a wastewatewater treatment plant. Several commercial on-line respirometric biosensors have been developed for monitoring various parameters such as carbon/nitrogen biological removal, anaerobic digestion, and activated sludge settling; these systems have common characteristics such as a simple measurement principle, low amount of chemicals required during their use, and the absence of a sample pretreatment step. On-line biosensors may be used as early warning tools in treatment plants, to protect plant operation and to increase efficiency, contributing to the development of toxicity management plans, especially in plants treating industrial wastes.

INTRODUCTION

The activated sludge process is considered today the most effective method for the biological treatment of wastewaters. The activated sludge process is based on the development of a heterogeneous microbial community in an aeration basin, which allows the system to be flexible with regard to considerable fluctuation in the influent wastewater composition. The activity and the concentration of the microorganism populations are crucial for the effective operation of the system, and the presence of certain toxic substances in increased concentrations in the influents may result in the reduction of microorganism activity. The effect of toxic substances on the more sensitive nitrifying bacteria is especially important, as these bacteria have a very low reproduction and growth rate (1,2). Toxic throughput is one of the major causes of the failure of biological treatment plants and results in noncompliance with discharge permit limits.

The operation of wastewater treatments plants has become an essential issue in European Union countries. According to Directive 91/271/EEC, all wastewaters in the territories of the European Union have to be collected and treated by appropriate methods prior to their discharge into a water receiver. In addition, municipalities with more than 15,000 inhabitants should have constructed and operated a wastewater treatment plant by the year 2000, whereas municipalities of more than 2000 inhabitants should have a secondary wastewater treatment plant by the end of 2005 to be able to discharge their wastewaters to a freshwater receiver. Furthermore, industrial wastewaters should be subjected to the same provisions as municipal wastewaters. In practice, industrial wastewaters are often discharged to a domestic collecting system and are fed to a municipal wastewater treatment plant for co-treatment. Thus, several municipal wastewater plants are receiving industrial wastewaters, which are often partially treated or even untreated; as a result, failure of the operation of municipal wastewater treatment plants is often attributed to the presence of certain compounds of industrial origin whose properties are potentially toxic to activated sludge microorganisms. Therefore, continuous monitoring of the potential toxic properties of influents is very important, prior to their introduction into the biodegradation process, to improve effluent quality, reduce operating cost, and increase reliability.

Conventional chemical analysis methods alone have been found inadequate for evaluating the toxic properties of wastewaters and for ensuring that the influents will not have significant effects on the receiving sewage works. Industrial wastewaters are characterized as complex mixtures of varying concentration of pollutants that may affect system performance. However, inhibitory metals and specific organic compounds may be detected in an industrial wastewater, but these data alone do not prove that the growth of bacteria and protozoa comprising the activated sludge biomass of the aeration basin will be suppressed. In addition, analytical chemical scans are very expensive and time-consuming and provide circumstantial evidence of the inhibitory nature of the waste. Furthermore, the toxicity of multicomponent mixtures to activated sludge microorganisms may differ from an additive response because synergistic or antagonistic interactions between mixture components may make the mixture more toxic than predicted by summing the effects of each individual toxicant (3). Consequently, an alternative method is necessary to evaluate the impact of industrial wastewaters on the performance of a domestic activated sludge treatment plant, based on assessing the toxicity of influents in a wastewater treatment plant.

TOXICITY ASSESSMENT OF INDUSTRIAL WASTEWATERS

Several biotests have been developed for evaluating wastewater toxicity to determine the toxic level of compounds to aquatic organisms. Test organisms that have been incorporated in bioassays include plants, invertebrates, fish, and microorganisms (4). Species that are widely used for determining the ecotoxic quality of freshwaters include (5,6): brook trout (*Salvelinus fontinalis*), Coho salmon (*Oncorhyncus kisutch*), bluegill (*Lepomis macrochirus*), channel catfish (*Ictalurus punctatus*), and invertebrates such as crayfish (*Pacifastacus leniusculus*) and Cladocera (*Daphnia magna, Daphnia pulex, Ceriodaphnia* spp.). In addition, marine species have been used for examining the ecological condition of seawater, including red algae (*Champia parvula*) and mysid shrimp (*Mysidopsis bahia*).

However, most of these bioassays are time-consuming and require scientifically skilled personnel and specific equipment. Moreover, the use of higher organisms, such as a fish, as a test species may be ethically undesirable. As a result, the development of an early warning system for monitoring the toxicity of influents is necessary. The last years of research has been directed to examinating of appropriate toxicity tests that are rapid, easy to operate, and inexpensive. Bacterial screening tests using microorganisms have been studied for their ability to assess the toxicity of wastewaters based on various principles such as microbial transformations, determination of microbial enzymes, and measurement of luminescent activity or respiration rate (5,7).

Most of the bacterial screening tests, such as the Microtox®, ToxAlert®, and the Mutatox® tests, which are used to detect cytotoxic and genotoxic agents, respectively, measure bioluminescence (8–10). The basic technology of these tests is using luminescent bacteria, specifically the strain *Vibrio fischeri*, to measure the toxicity of environmental samples. When properly grown, luminescent bacteria produce light as a byproduct of their cellular respiration. Cell respiration is fundamental to cellular metabolism and all life processes. Bacterial bioluminescence is tied directly to cell respiration, and any inhibition of cellular activity (toxicity) results in a decreased rate of respiration and a corresponding decrease in the rate of luminescence. The more toxic the sample, the greater the percentage of light lost from the test suspension of luminescent bacteria. Bacterial bioluminescence has proven a convenient measure of cellular metabolism and consequently, comprises a reliable sensor for measuring the presence of toxic chemicals in aquatic samples and for determining the toxicity of single compounds, mixtures of compounds, and industrial wastes (11). The particular strain was originally chosen for acute and chronic tests because it displayed high sensitivity to a broad range of chemicals.

The Microtox® acute test has been successfully used for a variety of environmental applications:

- wastewater treatment plant influent and effluent testing,
- toxicity reduction evaluations (TREs) and toxicity identification evaluations (TIEs), and
- assessment of soil contaminated with metals and energetic compounds.

However, the relevance of the information obtained by these bioassays is not considered essential for operating a municipal wastewater treatment plant, and continuous monitoring of the toxicity of industrial wastewaters is not achieved by such methods. Toxicity detected by *Vibrio fischeri* cannot always be associated with a biochemical cause that will exert the same activity reduction on the biomass culture in an activated sludge process (12). In general, toxicity assessment by Mictotox® may result in overestimating the acute toxicity effect on the biomass operating in a plant (13).

The key determinant in the response of a wastewater treatment plant is microbiological behavior. Thus, a proper estimate of the effect of toxic influents on an activated sludge unit should incorporate a system that closely mimics the conditions in the target plant because the response of a bacterial culture to a change in environmental conditions is highly dependent on the culture's physiological state, that is, it is determined by its growth history. Therefore, alternative techniques for on-line detection of toxicity relying on respirometric techniques, using activated sludge as the biological material, are considered the most suitable. Devices based on these techniques are easily operated, need minimum maintenance, and give fast and relevant responses (14–16). Respirometry was recognized many years ago as a key parameter for process control because it is a direct measure of the viability and activity of biological organisms. Therefore, respirometry plays an important role in detecting toxicity and in process control of wastewater treatment plants.

TOXICITY ASSESSMENT OF INDUSTRIAL EFFLUENTS BY RESPIROMETRIC TECHNIQUES

Respirometry is the measurement of oxygen uptake rates (OUR) of living biomass, such as activated sludge microorganisms, and can be used as an indicator of microbial activity in an aerobic biomass. On-line respirometry thus provides dynamic measurements of biological activity as well as a measure of the organic strength of a wastewater, which in turn gives an indication of the effective capacity of the system when sampled throughout the system.

Activated sludge is a flocculent mixture of microorganisms, mainly bacteria, which degrade multiple substrates to grow. The active organisms in the biomass of an activated sludge facility are contained in the solid phase and depend for growth on the carbohydrates, proteins, nucleic acids, and lipids contained in the influent to the plant. The organisms require dissolved oxygen to oxidize the organic food in wastewater and to provide organic carbon to synthesize the compounds necessary for reproduction. This oxygen consumption is referred to as respiration. Respiration is intimately linked to the growth of the activated sludge and to biodegradation.

In activated sludge heterotrophic bacteria which consume organic carbon, the carbon taken up is divided between that which is broken down in respiration and that which is built up into macromolecules for growth. The relationship between respiration, growth, and total carbon uptake (which equals biodegradation) is shown in Fig. 1.

By measuring the respiration rate (A), it is possible to calculate the amount of organic carbon (B) that is being used to provide energy. (A) can therefore be used to predict (B). Respiration rate (A) can also be used to predict the energy used in growth. It can be assumed that there is a proportional relationship between this energy and the amount of carbon being used for growth (C). (A) can therefore be used to predict (C). (B) + (C) equals the total carbon uptake, which equals the rate of biodegradation. Respiration rate (A) predicts (B) and (C), so it can be used to predict the rate of biodegradation. The situation in nitrifiers is simpler because the oxygen uptake rate provides a direct measure of ammonia biodegradation. These two mechanisms show the central role of the respiration rate (or more strictly, oxygen uptake rate, OUR) in the aerobic breakdown process.

In heterotrophic bacteria, toxic compounds in wastewater can inhibit either growth or respiration. Because of the tight linkages among respiration, growth, and biodegradation, it is obvious that if growth is inhibited, the requirement for energy decreases, and the respiration rate is reduced. Less carbon is required for respiration, so the rate of carbon uptake or biodegradation decreases. At the same time, the amount of carbon that is incorporated in growth is reduced. This also leads to reduced biodegradation. Conversely, if the toxic compound inhibits respiration, the amount of carbon for respiration is reduced, leading to reduced biodegradation. A decreased supply of energy from respiration results in a reduced growth rate and hence a lowered rate of carbon uptake or biodegradation. From this it can be seen that wherever the toxicity acts, there is inhibition of both the respiration rate and the rate of biodegradation. Nitrifiers are particularly prone to inhibition by toxic substances. Inhibition of either growth or the energy yielding oxidation reactions results in a decrease in both oxygen uptake rate and rate of uptake biodegradation of ammonia. The oxidative processes of respiration and nitrification that underlie the breakdown and removal of organic carbon and ammonia may be measured by respirometry, using specific biosensors.

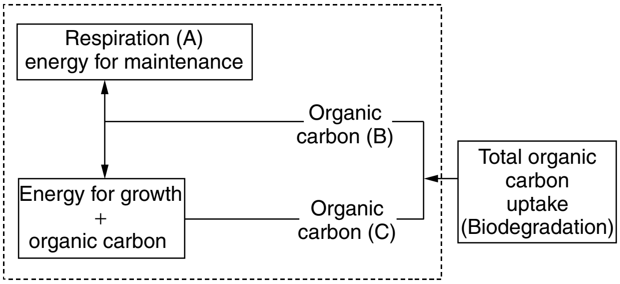

Figure 1. Relationship between respiration, growth, and biodegradation in heterotrophic bacteria.

Eight different types of respirometers have been developed, and in general they may be distinguished according to two criteria (17): (1) the phase where oxygen concentration is measured (liquid or gas) and (2) the method of liquid/gas introduction (flowing or static). Most respirometric methods are based on measuring the oxygen content in the liquid phase by using a dissolved oxygen (DO) probe. Static gas respirometers may be operated with a static or a flowing liquid phase. The static gas–static liquid respirometer is operated by withdrawing a sample of activated sludge from the aeration tank of a treatment plant, transferring it into a small reactor vessel, and then monitoring the decline of DO concentration by the time following a short aerated phase. The use of static gas–liquid respirometers is restricted due to potential problems from oxygen limitation. This type of respirometer, however, has the benefits of simple construction and measurement principle (18–20).

Flowing gas–static liquid respirometers, on the contrary, are continuously aerated and have the advantage that higher sludge concentrations can be used because there is a continuous input of oxygen, though oxygen limitation is unlikely (21). In this case, the oxygen transfer coefficient and the saturation DO concentration have to be known to calculate the respiration rate. Static gas–flowing liquid respirometers measure the DO concentration at both the inlet and the outlet of a closed respiration chamber (22). Aerated sludge is pumped continuously through the respiration chamber. The oxygen uptake rate is calculated by making an oxygen mass balance over the respiration chamber using the input and output DO concentration and the residence time in the vessel.

Several commercial biosensors have been presented for determining the toxicity of industrial effluents and assessing activated sludge plant performance, such as the Rodtox (Fig. 2a), the Stiptox (Fig. 2b), and the Manotherm. The Rodtox consists of a temperature controlled reactor vessel, which is filled with about 10 liters of activated sludge continuously aerated and stirred, while the dissolved oxygen concentration is monitored (21). Injection of a biodegradable calibration liquid in the reactor leads to a peak in the dissolved oxygen concentration. The total surface and the maximum height of the peak are proportional to the amount of BOD added, and the maximum slope of the respirogram is a measure of the activity of the activated sludge. The BOD of an unknown sample (short-term BOD, $\mathrm{BOD_{st}}$) is obtained by comparing the respirogram of the sample with that of the calibration liquid (23). Toxicity can be detected by comparing the peak of the calibration liquid before and after the addition of the sample. Stiptox is a toxicity analyzer with immobilized turbulent-bed biology where the microbes grow on the

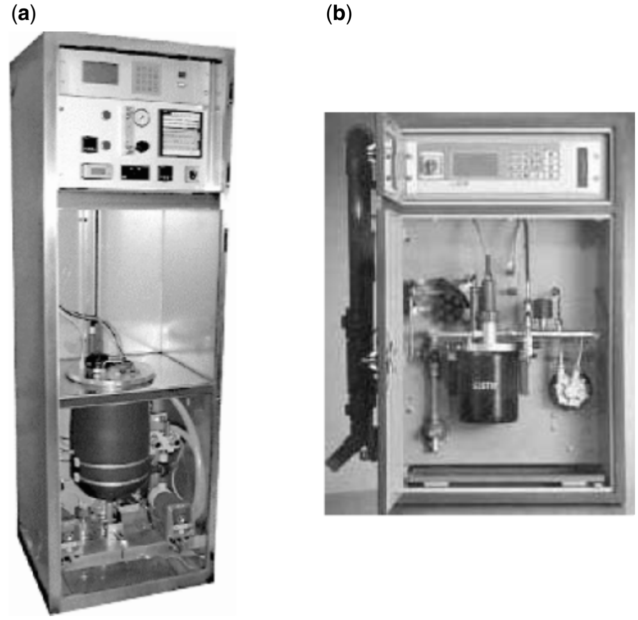

Figure 2. Commercial on-line toxicity biosensors: (**a**) Rodtox; (**b**) Stiptox.

inner surface of small hollow cylinders (24). Like the activated sludge of a treatment plant, the microbial population in the bioreactor is adapted to the conditions of the wastewater. As long as the wastewater is not toxic to the adapted biology, the organisms in the bioreactor take up dissolved oxygen. A toxic impact inhibits the respiration of the organisms, causing an increase in the dissolved oxygen level.

An alternative biosensor incorporating bacterial cells that can be used for rapid ecotoxicity analysis has been developed by Evans et al. (25). The amperometric biosensor device was called Cellsense and used ferricyanide, a soluble electron mediator, to divert electrons from the respiratory system of the immobilized bacteria to a suitable poised carbon electrode. The resulting current was, thus, a measure of bacterial respiratory activity, and the perturbation by pollutants could be detected within minutes as a change in the magnitude of the current.

In addition to carbonaceous removal, respirometric measurements are useful for monitoring nitrifying sludge activity and biological nitrogen removal. The interest in the nitrification capacity of activated sludge results from the higher sensitivity of nitrifiers to toxic compounds compared with heterotrophs (1). However, the main problem in measuring nitrifying capacity is separation of the nitrification oxygen uptake from the oxygen uptake for carbon substrate oxidation and endogenous metabolism. In general, batch experiments using activated sludge in the endogenous state may be used to determine nitrifying activity. In static gas–static liquid respirometers, the increase in OUR due to nitrification may be observed just after the addition of a certain amount of NH_4^+ (26). Respirometric methods consisting of pulse substrate additions to activated sludge cannot be used for estimating nitrification activity, as fresh substrate for both carbon oxidation and nitrification is continuously provided by the fresh wastewater entering the plant. Thus, selective nitrification inhibitors should be added; in this case, the experimental procedures consist of two subsequent phases. Initially, the total OUR of the sludge is measured, a nitrification inhibitor is added to the sludge, and OUR is measured again. The actual nitrification rate is then estimated by the difference between the two OUR values (20). Furthermore, respirometric experiments have been used to identify and quantify nitrogen nutrient deficiency in activated sludge processes (27). These methods may be applied to evaluate potential nitrogen nutrient deficiency and to estimate the amount of nitrogen required to remedy a nutrient deficiency in the sludge.

Respirometers have been used for detecting influent toxicity in wastewater treatment plants, using nitrifying bacteria as the indicating organisms, by the addition of specific nitrifying inhibitors (19,28). Nevertheless, specific techniques have been reported based on the implementation of bacteria from nitrifying enriched cultures containing no or little heterotrophic bacteria, where the addition of nitrification inhibitors is not necessary (19,28).

CASE STUDY: EXAMINATION OF INDUSTRIAL EFFLUENTS BY A RESPIROGRAPHIC BIOSENSOR

To evaluate the effect of industrial discharges on the performance of a municipal wastewater treatment plant, a biosensor (Rodtox, Kelma, NV) was installed in the industrial area of Thessaloniki, Greece. Several production plants are included in the Greater Industrial Area such as metal processing units, food industries, chemical-agrochemical, and pharmaceutical companies. Effluents from each process are partially treated on site and are discharged into a common pipe. Currently, industrial wastewaters are further treated in an industrial wastewater treatment plant; in the near future, industrial effluents will be fed to the Thessaloniki municipal wastewater treatment plant. The inhibition of the activity of activated sludge microorganisms due to the addition of industrial discharges to the respirometric biosensor, was estimated from the corresponding respirometric curves (21).

During the measuring period, activated sludge from the Thessaloniki municipal wastewater treatment plant was used as the biological material in the respirometric sensor. Activated sludge properties were considered as remaining almost similar during the monitoring time, as the wastewater treatment plant was fed by domestic wastewaters and its operation was maintained at constant conditions. In addition, the influence of the properties of activated sludge on the inhibitory results were periodically examined using calibration substances of known toxicity (solution of $ZnCl_2$), and slight differences were observed in the respirometric response. Similar results regarding the effect of the activated sludge origin have been already reported, underlying the insignificant influence of activated sludge on inhibition data (11).

Characteristic values measured in industrial wastewater samples are presented in Table 1. As shown in this

Table 1. Characteristic Values of Wastewaters from the Thessaloniki Industrial Area, mg/L

Parameter	Min	Max	Average
pH	7.1	8.3	7.6
BOD_5	71	2640	553
COD	155	4027	975
Ammonium nitrogen	2	317	62
Kjeldahl nitrogen	9	526	90
Total phosphorous	2	8	5
Suspended solids	148	376	249

table, industrial wastewaters in general were more polluted than municipal wastewaters; organic compounds reached up to 1000 mg/L COD. As a result, continuous monitoring of the toxicity of industrial discharges was very important to protect the co-treatment plant. Fluctuations in characteristic values could be attributed to seasonal variations, nontreated effluents, or effluents from processing units for which on-site wastewater treatment systems presented operating problems.

The short-term BOD of industrial wastewaters (BOD_{st}) and the corresponding inhibition of the respiration of activated sludge microorganisms during a 6-month period are shown in Figs. 3 and 4, respectively. High variation in both inhibition values and BOD_{st} was observed as a result of the complex nature of the samples. Short-term BOD values reached up to 3300 mg/L, and the corresponding inhibition values reached up to 98%, indicating that these wastewaters were toxic to the activated sludge microorganisms and could adversely affect the operation of the municipal treatment plant.

From the analysis of the relation between short-term BOD and inhibition of microorganism respiration during the particular period, it was concluded that the organic content of the industrial samples did not correlate directly with the toxicity of the samples. Samples of high toxicity, exceeding 35% inhibition, had BOD_{st} values ranging from low values up to 3000 mg/L. On the other hand, samples of low toxicity had a variation of organic loading from low levels up to levels exceeding 3000 mg/L. According to this observation, factors other than organic substances might affect the toxicity of the samples, such as inorganic compounds and especially metals and their speciation (29).

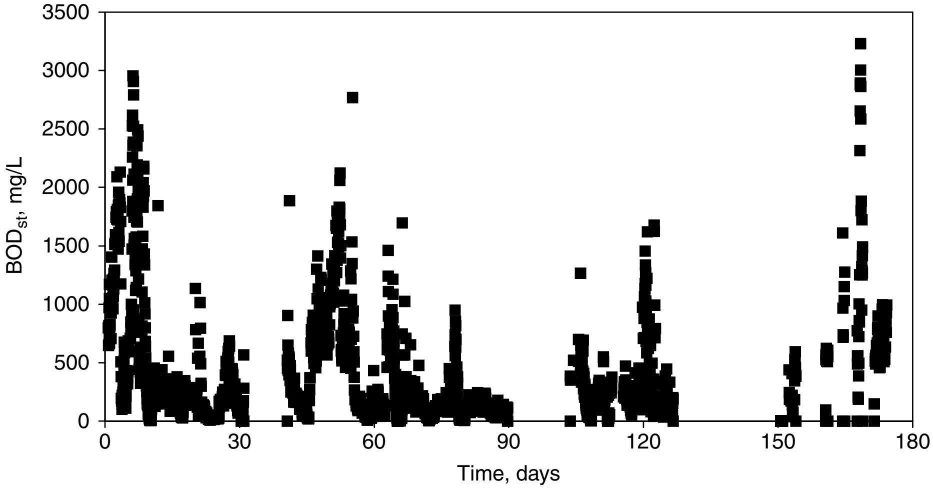

Figure 3. Short-term BOD values of industrial wastewaters, measured by an on-line toxicity meter.

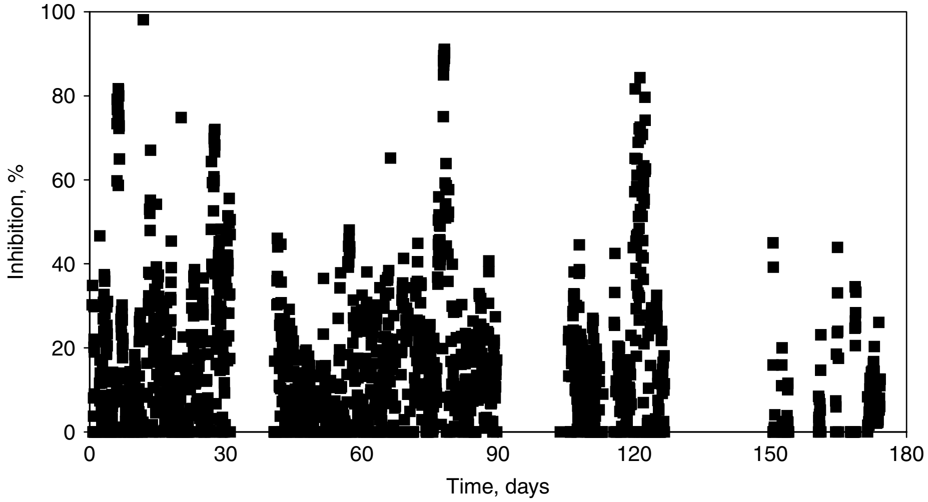

Figure 4. Inhibition of the respiration of active sludge microorganisms by the addition of industrial wastewaters.

Statistical analysis of the collected inhibition data was performed to obtain the daily toxicity distribution of industrial effluents; the characteristic values are shown in Table 2. In addition, the weekly frequency of toxicity values exceeding 35% inhibition is included in this table. As shown, the highest average toxicity values were observed on Saturday and Sunday, reaching about 18.5 and 20%, respectively. However, the maximum toxicity was observed on Tuesday evening. In addition, industrial effluents of high toxicity were supplied to the system on Thursdays and during weekends, corresponding to about 70% of the total number of values exceeding 35% inhibition. Furthermore, from the examination of the daily variation of inhibition values exceeding 35%, it was found that 63% of the peak inhibition values were observed during the night from 23:00 to 07:00. These results could be attributed to the reduced control of existing wastewater treatment systems in industrial plants, resulting in discharge of wastewater of low quality.

In conclusion, respirometry is an indispensable tool for assessing the biological treatability of industrial wastes separately and in combination with municipal wastewaters and for providing significant information about the impact of toxic effluents on the performance of a wastewater treatment plant. The benefit of on-line respirometric measurement is that it is a microcosm of the treatment facility and of what really affects process success. It shows the relative strength of the wastewater, how much treatment is needed, and the quality of the effluent produced.

Common characteristics of respirometric biosensors are the simple measurement principle, the low amount of chemicals required during their use, and the absence of a sample pretreatment step. Several biosensors have been developed to monitor various parameters such as biological removal of carbon and nitrogen, anaerobic digestion, and activated sludge settling characteristics that provide detailed information about important activated sludge processes. On-line respirometric biosensors may be used as early warning systems in influent toxic load detection, protecting the wastewater treatment plant from toxic shocks (30). Thus, their use in wastewater treatment plants supports the development of appropriate control strategies (31) to protect the efficient operation of a plant, to optimize system performance by better control of the treatment plant, and to improve effluent quality.

Table 2. Average Values and Daily Toxicity Variation of Industrial Discharges, %

Day	Min	Max	Average	Values Exceeding 35% Inhibition
Monday	0	79.7	13.7	10.7
Tuesday	0	98.1	7.7	0.9
Wednesday	0	64.4	8.4	2.8
Thursday	0	81.7	17.6	21.5
Friday	0	88.4	13.5	15.3
Saturday	0	91.1	18.5	23.9
Sunday	0	84.3	19.9	24.8

BIBLIOGRAPHY

1. Blum, D.J.W. and Speece, R.E. (1991). A database of chemical toxicity to environmental bacteria and its use in interspecies comparisons and correlations. *Res. J. Water Pollut. Control Fed.* **63**: 198–207.
2. Krois, P., Schweighofer, P., Frey, W., and Matsche, N. (1992). Nitrification inhibition—a source identification method for combined municipal and/or industrial wastewater treatment plants. *Water Sci. Technol.* **26**(5–6): 1135–1146.
3. Kungolos, A., Samaras, P., Kipopoulou, A.M., Zoumboulis, A., and Sakellaropoulos, G.P. (1999). Interactive toxic effects of agrochemicals on aquatic organisms. *Water Sci. Technol.* **40**(1): 357–364.
4. U.S. Environmental Protection Agency. (1991). *Technical Support Documents for Water Quality-Based Toxics Control*. EPA 502/2-90-001. Office of Water, U.S. EPA, Washington, DC.
5. Farre, M., Pasini, O., Carmen-Alonso, M., Casillo, M., and Barcelo, D. (2001). Toxicity assessment of organic pollution in wastewaters using a bacterial biosensor. *Anal. Chim. Acta* **426**: 155–165.
6. Farre, M. and Barcelo, D. (2003). Toxicity testing of wastewater and sewage sludge by biosensors, bioassays and chemical analysis. *Trends Anal. Chem.* **22**(5): 299–310.
7. Anderson, K., Koopman, B., and Bitton, G. (1988). Evaluation of INT-dehydrogenase assay for heavy metal inhibition of activated sludge. *Water Res.* **22**(3): 349–353.
8. Chen, C.Y., Huang, J.B., and Chen, S.D. (1997). Assessment of the microbial toxicity test and its application for industrial wastewaters. *Water Sci. Technol.* **36**(12): 375–382.
9. Chen, C.Y., Chen, J.N., and Chen, S.D. (1999). Toxicity assessment of industrial wastewater by microbial testing method. *Water Sci. Technol.* **39**(10–11): 139–143.
10. Dos Santos, L.F., Defrenne, L., and Krebs-Brown, A. (2002). Comparison of three microbial assay procedures for measuring toxicity of chemical compounds: ToxAlert®10, CellSense and Biolog MT2 microplates. *Anal. Chim. Acta* **456**: 41–54.
11. Dalzell, D.J.B., Alte, S., Aspichueta, E., de la Sota, A., Etxebarria, J., Gutierrez, M., Hoffmann, C.C., Sales, D., Obst, U., and Christofi, N. (2002). A comparison of five rapid direct toxicity of pollutants to activated sludge. *Chemosphere* **47**: 535–545.
12. OECD Method 209. (1987). Activated sludge, respiration inhibition test. In: *OECD Guidelines for Testing of Chemicals*. OECD, Paris, France.
13. Ricco, G., Tomei, M.C., Ramadori, R., and Laera, G. (2004). Toxicity assessment of common xenobiotic compounds on municipal activated sludge: Comparison between respirometry and Microtox®. *Water Res.* **38**: 2103–2110.
14. Vanrolleghem, P.A., Kong, Z., and Coen, F. (1996). Full-scale on-line assessment of toxic wastewaters causing change in biodegradation model structure and parameters. *Water Sci. Technol.* **33**(2): 163–175.
15. Ubay, E., Sozen, S., Orhon, D., and Henze, M. (1998). Respirometry analysis of activated sludge behaviour-I. Assessment of the readily biodegradable substrate. *Water Res.* **32**(2): 461–475.
16. Gutierez, M., Etxebarria, J., and de las Fuentes, L. (2002). Evaluation of wastewater toxicity: Comparative study between Microtox® and activated sludge oxygen uptake inhibition. *Water Res.* **36**: 919–924.
17. Spanjers, H., Vanrolleghem, P., Olsson, G., and Dold, P. (1996). Respirometry in control of the activated sludge process. *Water Sci. Technol.* **34**(3–4): 117–126.

18. Drtil, M., Nemeth, P., and Bodik, I. (1993). Kinetic constants of nitrification. *Water Res.* **27**: 35–39.
19. Gernaey, K., Bogaert, H., Massone, A., Vanrolleghem, P., and Verstraete, W. (1997). On-line nitrification monitoring in activated sludge with a titrimetric sensor. *Environ. Sci. Technol.* **31**: 2350–2355.
20. Gernaey, K., Vanderhasselt, A., Bogaert, H., Vanrolleghem, P., and Verstraete, W. (1998). Sensors to monitor biological nitrogen removal and activated sludge settling. *J. Microbiological Methods* **32**: 193–204.
21. Vanrolleghem, P.A., Kong, Z., Rombouts, G., and Verstraete, W. (1994). An on-line respirographic biosensor for the characterization of load and toxicity of load and toxicity of wastewaters. *J. Chem. Technol. Biotechnol.* **59**: 321–333.
22. Vanrolleghem, P. and Spanjers, H. (1994). Comparison of two respirometric principles for the determination of short-term biochemical oxygen demand. *Proc. 49th Ind. Waste Conf.*, May 9–11, Purdue University, West Lafayette, IN, pp. 177–188.
23. Kong, Z., Vaerewijck, M., and Verstraete, W. (1996). On-line stBOD measurement and toxicity control of wastewaters with a respirographic biosensor. *Environ. Tech.* **17**: 399–406.
24. Teutscher, M. and Grosser, J. (1994). Experiences with online toxicity analysers in municipal and industrial waste water treatment plants. *Eur. Symp. Biol. Resort Continuous Monitoring Water Qual.*, October 13–14, Nancy, France.
25. Evans, M.R., Jordinson, G.M., Rawson, D.M., and Rogerson J.G. (1998). Biosensors for the measurement of toxicity of wastewaters to activated sludge. *Pest. Sci.* **54**: 447–452.
26. Nowak, O. and Svardal, K. (1993). Observations on the kinetics of nitrification under inhibiting conditions caused by industrial wastewater compounds. *Water Sci. Technol.* **28**(2): 115–123.
27. Ning, Z., Patry G.G., and Spanjers, H. (2000). Identification and quantification of nitrogen nutrient deficiency in the activated sludge process using respirometry. *Water Res.* **34**(13): 3345–3354.
28. Aivasidis, A., Hochscherf, H., Rottman, G., Hagen, T., Mertens, M.T., Reiners, G., and Wandrey, C. (1992). Neuere Konzepte zur Prozessuberwachung und—regelung bei der biologischen Stickstoffelimination. *Abwassertechnik* **42**(5): 48–55.
29. Akio, I. and Earnest, G.F. (1996). Speciation of chromium (III) in activated sludge. *Water Sci. Technol.* **68**(3): 301–310.
30. Temmink, H., Vanrolleghem, R., Klapwijk, A., and Verstraete, W. (1993). Biological early warning systems for toxicity based on activated sludge respirometry. *Water Sci. Technol.* **28**(11–12): 415–425.
31. Copp, J.B. and Spanjers, H. (2004). Simulation of respirometry-based detection and mitigation of activated sludge toxicity. *Control Eng. Pract.* **12**: 305–313.

POLYCYCLIC AROMATIC HYDROCARBONS

Samuel C. Ashworth
Idaho National Engineering and
Environmental Laboratory
Idaho Falls, Idaho

Polycyclic aromatic hydrocarbons (PAHs) are a large group of organic compounds that have two or more fused aromatic rings. They have relatively low solubility in water and are highly lipophilic. The term polycyclic organic matter (POM) defines a broad class of compounds that generally includes all organic structures that have two or more fused aromatic rings. Polycyclic organic matter has been identified with as many as seven fused rings and, theoretically, millions of POM compounds could exist; however, only about 100 species have been identified and studied, and typically only a small fraction of them are regularly tested for as part of emissions measurement programs (1). Any effort to quantify emissions of POM relies on the group of compounds or analytes targeted by the test method employed. The EPA has defined these eight major categories of compounds in the class known as POM (2).

- polycyclic aromatic hydrocarbons (PAHs);
- aromatic hydrocarbons that contain nitrogen in a heterocyclic ring;
- aromatic hydrocarbons that contain a carbon–nitrogen double bond (C=NH);
- aromatic hydrocarbons that contain a one-ring carbonyl divalent group (C=O), also known as quinones; contain two-ring carbonyl divalent groups;
- carbonyl arenes that contain hydroxy groups and possibly alkoxy or acyloxy groups;
- oxa arenes are aromatic hydrocarbons that contain an oxygen atom in a heterocyclic ring; thia arenes are aromatic hydrocarbons that contain a sulfur atom in a heterocyclic ring; and
- some polyhalo compounds, such as polychlorinated dibenzo-*p*-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs), may be considered POM although they do not have two or more fused aromatic rings.

PAHs, the focus of this article, include naphthalene, phenanthrene, anthracene, fluoranthene, acenaphthalene, chrysene, benz[*a*]anthracene, cyclopenta[*cd*]pyrene, benzopyrenes, indeno(1,2,3-cd)pyrene, benzo(ghi)perylene, coronene, and some alkyl derivatives of these compounds. The structural representations of some of these PAHs are presented in this section (Fig. 1).

There are other naming divisions used also. If two or more benzene rings share pairs of carbon atoms resulting in fused aromatic rings, they have been known as polycyclic aromatic hydrocarbons (PAHs) and polynuclear aromatic compounds (PNAs). The PNA classification is more general and includes heterocyclic aromatic compounds (as discussed above in the EPA definitions). The rings constituting PAHs contain only carbon (3). PAH nomenclature is based on 35 IUPAC-prescribed compounds and naming sequences described in detail elsewhere (3).

Most POM compounds are solids that have high melting and boiling points and are extremely insoluble in water. The PAHs are primarily planar, nonpolar compounds whose melting points are considerably higher than 212 °F (100 °C). Phenanthrene, whose melting point is 214 °F (101 °C) and benzo(c)phenanthrene, whose melting point is 154 °F (68 °C) are two exceptions. The molecular weights,

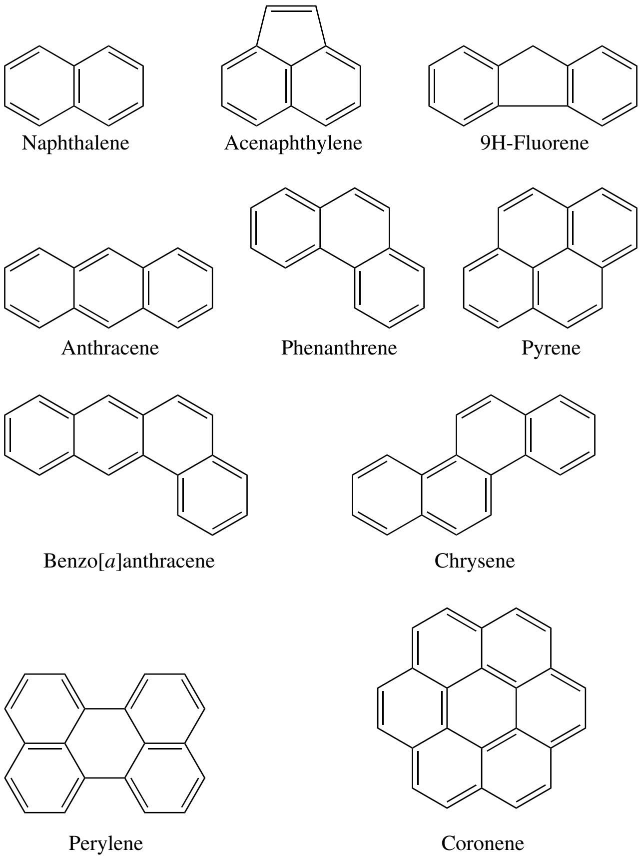

Figure 1. PAH structural representations.

melting points, and boiling points of selected PAH species are listed in Table 1.

The vapor pressures of POM compounds vary depending on the ring size and the molecular weight of each species. The standard vapor pressures of pure PAH compounds vary from 6.8×10^{-4} mmHg for phenanthrene to 1.5×10^{-12} mmHg for coronene (4). Table 1 includes vapor pressures at 25 °C for selected PAHs. Ultraviolet absorption spectra are available for many POM compounds. Most polycyclic aromatic hydrocarbons absorb light at wavelengths found in sunlight (>300 nm) and are believed to be photochemically reactive by direct excitation.

The chemistry of POMs is quite complex and differs from one compound to another. Most of the information available in the literature concerns polycyclic aromatic hydrocarbons. Generally, PAHs are more reactive than benzene, and reactivities toward methyl radicals tend to increase with greater conjugation. Conjugated rings are structures that have double bonds alternating with single bonds. Conjugated compounds are generally more stable, but they are more reactive toward free radical addition (6). For example, in comparison to benzene, naphthalene and benz(a)anthracene, which have greater conjugation, react with methyl radicals 22 and 468 times faster, respectively.

PAHs undergo electrophilic substitution reactions quite readily. An electrophilic reagent attaches to the ring to form an intermediate carbonium ion; to restore the stable aromatic system, the carbonium ion then gives up a proton. Oxidation and reduction reactions occur to the stage where a substituted benzene ring is formed. Rates of electrophilic, nucleophilic, and free radical substitution reactions are typically greater for PAHs than for benzene.

Environmental factors also influence the reactivity of PAHs. Temperature, light, oxygen, ozone, other chemical agents, catalysts, and the surface areas of particulates onto which the PAHs are adsorbed may play a key role in the chemical reactivity of PAHs.

The principal formation mechanism for POM occurs as part of the combustion process in many different types of sources. A secondary formation mechanism, primarily represented by naphthalene production and use categories, is volatilization of lightweight POM compounds. However, the combustion mechanism is much more significant in overall POM formation, and it also much more complex.

Most PAHs that have low vapor pressure in the air are adsorbed on particles. Dissolved in water or adsorbed on particulate matter, PAHs can undergo photodecomposition when exposed to ultraviolet light from solar radiation. In the atmosphere, PAHs can react with pollutants such as ozone, nitrogen oxides, and sulfur dioxide, yielding diones, nitro- and dinitro-PAHs, and sulfonic acids, respectively.

Table 1. Physical Constants of PAHs[a]

PAH	Mol. wt.	Solubility (μg/L)	Vap. Pressure at 25 °C (mmHg)	Log K_{ow} (Log K_{oc})	Henry's Law Constant,[b] atm-L/mole	MP °C	BP °C
Naphthalene	128.2	12,500 to 34,000	1.8×10^{-2}	3.37	4.5	80.5	218
Acenaphthylene	152.2	3420	10^{-3}–10^{-4}	4.07 (3.40)	0.24	96.2	279
Fluorene	166.2	800	1.64×10^{-5}	4.18 (3.86)	0.074	116	295
Anthracene	178.2	59	2.4×10^{-4}	4.5 (4.15)	1.8×10^{-3}	217	340
Phenanthrene	178.2	435	6.8×10^{-4}	4.46 (4.15)	—	100	340
Pyrene	202.1	133	6.9×10^{-7}	4.88 (4.58)	0.013	153	360
Benz[*a*]anthracene	228.3	11.0	1.1×10^{-7}	5.63 (5.30)	1.2×10^{-3}	160	435
Chrysene	228.3	1.9	5.3×10^{-9}[b]	5.63 (5.30)	6.7×10^{-4}	252	448
Perylene	252.3	2.4	—	6.21	—	274	500
Coronene	300.3	0.14	1.5×10^{-11}	7.36	1×10^{-6}	438	525

[a] Reference 5.

[b] Polycyclic aromatic hydrocarbons (PAHs) database in alphabetical order, http://chrom.tutms.tut.ac.jp/JINNO/DATABASE/00alphabet.html.

PAHs may also be degraded by some microorganisms in the soil (7,8).

PAHs are formed mainly as a result of pyrolytic processes, especially incomplete combustion of organic materials during industrial and other human activities, such as coal and crude oil processing, natural gas combustion, including heating, combustion of refuse, vehicle traffic, cooking and tobacco smoking, as well as in natural processes such as carbonization. There are several hundred PAHs; the best known is benzo[*a*]pyrene (BaP). In addition, a number of heterocyclic aromatic compounds (e.g., carbazole and acridine), as well as nitro-PAHs, can be generated by incomplete combustion (7).

Examination of a number of drinking-water supplies for six PAHs (fluoranthene, benzo[*b*]fluoranthene, benzo[*k*]fluoranthene, BaP, benzo[*ghi*]perylene, and indeno[1,2,3-*cd*]pyrene) indicated that the collective concentrations generally did not exceed 0.1 μg/L. The concentrations of these six PAHs were between 0.001 and 0.01 μg/L in 90% of the samples and higher than 0.11 μg/L in 1%. Concentrations of BaP in drinking water range to 0.024 μg/L (9).

PAH compound treatment depends upon the end user requirements. Some of the criteria are provided in Table 2. From a treatment perspective, it is only practical to target the objective at the criterion or detection limit (DL), whichever is smaller. In some instances, the practical quantitation limit (PQL) would be the target. Some of the standards shown in Table 2 require values below the PQL (e.g., for benzo [*a*] anthracene $0.013 < 0.044 < 0.13$).

Napthalene and other PAHs by extension are classified as aromatic compounds because they exhibit some of the properties of benzene. Napthalene has the required structure of an aromatic compound (6). It contains six-membered rings, and the atomic orbitals show that the structure can provide π clouds that contain six electrons, as shown in Fig. 2. As illustrated, ten carbons lie at the corners of two fused hexagons. Each carbon is attached to three other atoms by σ bonds. The bonds result from the overlap of trigonal sp^2 orbitals, so all carbon and hydrogen atoms lie in a single plane that has π electron clouds above and below forming a shape like fused toroids by the overlap of p orbitals (6).

Table 2. PAH Detection Limits and Criteria

PAH	Detection Limits[a] DL, μg/L	Detection Limits[a] PQL,[b] μg/L	Drinking Water, μg/L	Surface Water, μg/L
Naphthalene	1.8	18		
Acenapthylene	2.3	23	1200[c]	2700[c]
Acenapthene	1.8	18		
Fluorene	0.21	21	400[d]	
Phenanthene	0.64	6.4		
Anthracene	0.66	6.6	3000,[d] 9600[c]	110000[c]
Fluoranthene	0.21	2.1	400[d]	
Pyrene	0.27	2.7	960[c]	11000[c]
Benzo [*a*] anthracene	0.013	0.13	0.044[c]	0.49[c]
Chrysene	0.15	1.5		
Benzo [*b*] fluoranthene	0.018	0.18	0.2,[d] 0.044[c]	0.49[c]
Benzo [*k*] fluoranthene	0.017	0.17	0.044[c]	0.49[c]
Benzo [*a*] pyrene	0.023	0.23	0.2,[d] 0.7,[e] 0.044[c]	0.49[c]
Dibenz [*a*, *h*]-anthracene	0.03	0.3	0.044[c]	0.49[c]
Benzo [*g*, *h*, *i*] perylene	0.076	0.76		
Indeno [1,2,3-*cd*]-pyrene	0.0043	0.43	0.044[c]	0.49[c]

[a] U.S. EPA Method 8310.
[b] Practical quantitation limit = 10 × detection limit (DL).
[c] State of Virginia, ftp://ftp.deq.state.va.us/pub/watqual/tox/appc.pdf.
[d] State of Wisconsin, http://www.dhfs.state.wi.us/eh/ChemFS/pdf/pah.pdf.
[e] World Health Organization, http://www.who.int/water_sanitation_health/GDWQ/Summary_tables/Tab2b.htm.

PAH compounds, like other aromatic compounds, are resonance hybrids or resonance structures. The following Kekulé diagram illustrates the resonance equilibria (Fig. 3). Unlike benzene, all carbon–carbon bonds are not the same length. However, from a general approach, naphthalene and other PAH compounds are expected to have many of the typical aromatic properties, especially those relating to the π bond and resonance.

Two of the reactions discussed in Morrison and Boyd, (6) including oxidation (Fig. 4) and reduction (Fig. 5), reduce the bond order in PAHs. These can be helpful for researchers and designers in determining some of the potential reactions and treatment in water and possibly solid media. It is also speculated that PAH compounds can be treated via high-energy chemistry (e.g., UV/H_2O_2) to reduce toxicity. This may be assisted by using surfactants because, as previously discussed, PAHs are not very soluble in water and prefer a solid phase. This is discussed after the classical naphthalene reactions.

The first reaction reduces the bond order. The product may or may not be toxic, but the use of chromium is excluded, because it is a listed toxic metal. The second reaction has been commercialized in a different form known as the aqueous electron. However, due to the energetic sodium metal, this is not used for aqueous liquids (10).

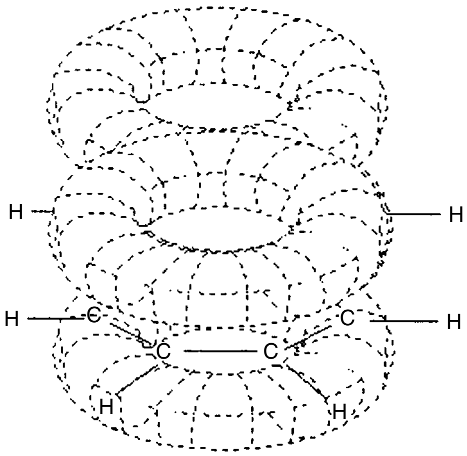

Figure 2. Napthalene π clouds.

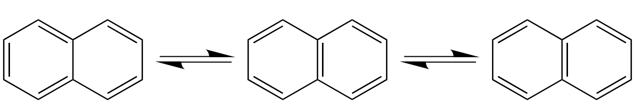

Figure 3. Napthalene Kekulé diagram.

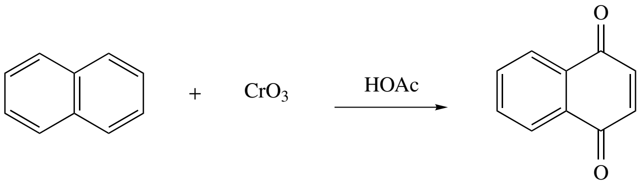

Figure 4. Oxidation using chromate.

A possible reaction path for naphthalene is shown in Fig. 6. This shows that naphthalene produces an aldehyde-ring compound that would be expected to undergo further additions thereby destroying the ring structures and likely forming complex aldehydes.

However, PAHs are sparingly soluble, so only the small amount in the aqueous phase would be removed, and the equilibrium would shift from the solid, if present, to maintain the solubility until there is no appreciable PAH left. It may be possible to use a surfactant to promote solubility and enhance this reaction. This requires evaluation because there are UV interferences, including phosphates, hydroxyl scavengers, and precipitates, all common in waters, that can impact the feasibility of this technology. The section on organic treatment discusses some of these technologies (11).

Because of low volatility, PAH compounds are not generally good candidates for removal by air stripping. The stripping factor (11) shows that quite large quantities of air are required:

$$\frac{G}{L} = \frac{R}{H} \tag{1}$$

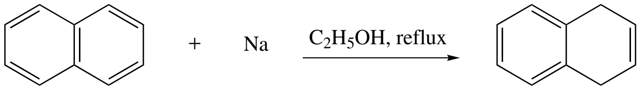

Figure 5. Reduction by sodium in alcohol.

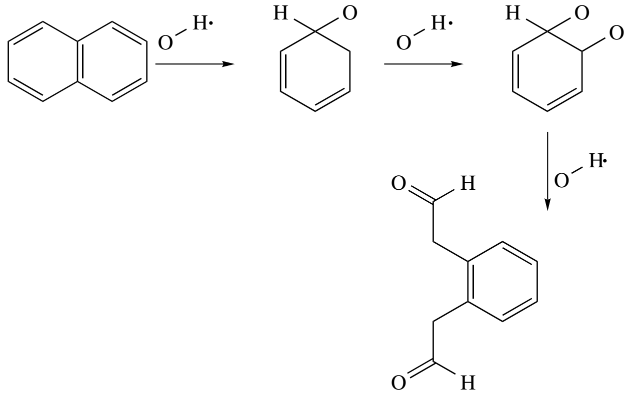

[2-(2-Oxo-ethyl)-pheny]-acetaldehyde

Figure 6. Possible reaction mechanism using UV/H_2O_2.

Using the highest H for Napthalene in the above and a break-even $R = 1$,

$$\frac{G}{L} = \frac{1}{4.5\ \text{atm} - \text{L/mole}}{}^{*}0.0823\ \text{L} - \text{atm/mole/K}^{*}298\ \text{K} = 5.45$$

Assuming a flow of 10 kg/min, the air rate would need to be 54.5 kg/min or 42 standard m^3/ min. This is feasible, but it represents excessive air requirements and becomes unrealistic for the other PAH compounds shown in Table 1.

Frequently, estimates of the air, water, and solid phase concentrations are desired. Assuming that the components of the phases are in equilibrium, the fugacities are equal (12).

$$f_{i,\text{gas}} = f_{i,\text{liquid}} = f_{i,\text{solid}} \tag{2}$$

Therefore, the gas and solid fugacities can be equated with the liquid fugacity. The result for a closed system (e.g., a closed vessel at atmospheric pressure) is

$$x_L = \frac{x}{K_D S + HV + L} \tag{3}$$

$$x_S = \frac{xK_D}{K_D S + HV + L} \tag{4}$$

$$x_V = \frac{xH}{K_D S + HV + L} \tag{5}$$

To illustrate the hydrophobic tendencies and also explain the meaning of K_{ow} in Table 1, the above values are calculated for fluorine. K_D is found from K_{oc} and f_{oc} (13):

$$K_D = K_{\text{oc}}{}^{*}f_{\text{oc}} \tag{6}$$

The octanol–water partition coefficient is a measure of the distribution of an organic compound between water and octanol. It is a general measure of a compound's hydrophobic tendencies. High values indicate hydrophobicity; low values indicate hydrophilic tendencies and hence higher solubility, for example, in alcohol.

K_{oc} can be found by a relation of the sort:

$$\ln K_{\text{oc}} = a \ln K_{\text{ow}} + b \tag{7}$$

However, the value listed for fluorine is log (K_{oc}) = 3.86 so that $K_{\text{oc}} = 7,240$ L/g. Assuming that the organic fraction of the solid is 1% ($f_{\text{oc}} = 0.01$), the solid phase has 10 g, the liquid phase and gas phases are 1 L each, and adding $x = 10$ mg,

$$K_D = K_{\text{oc}} f_{\text{oc}} = 72.4$$

$$x_L = \frac{10\ \text{mg}}{72.4\ \text{L/g}^{*}10\ \text{g} + 0.074\ \text{L} - \text{atm/mole/}\,(0.0823\ \text{L} - \text{atm/mole K}^{*}298\ \text{K})^{*}1\ \text{L}^{*}0.082 + 1\ \text{L}}$$

$$x_L = 0.0138\ \text{mg/L}$$

$$x_S = \frac{10\ \text{mg} \times 72.4\ \text{L/g}}{725} = 1\ \text{mg/g}$$

$$x_V = \frac{10\ \text{mg} \times 0.00302}{725\ \text{L}} = 4.17 \times 10^{-5}\ \text{mg/L}$$

This estimate demonstrates conclusively that PAH compounds are distributed preferentially on solids containing organics. In addition, the solids do not necessarily need to contain organics. However, the presence of the organic fraction ensures high distribution on the solid phase.

If there is no solid, then the fluorine will split into the water phase and a separate PAH phase consisting of nearly pure fluorine,

$$10\text{ mg} = 0.8(\text{mg/L}) \times L(L) + x\text{mg}$$

or $x = 9.2$ mg in a separate phase with a distinct vapor pressure of 1.64×10^{-5} mmHg. Therefore, the presence of a solid has a very large effect on the distribution between the phases and must be considered in treatment strategies.

Granular activated carbon (GAC) is an excellent technology for removing PAH compounds from water due to its demonstrated isotherms. Freundlich adsorption isotherms available for PAH compounds show that K's vary from 115 for acenapthalene to 376 for anthracene (14). The article in this Encyclopedia on organic treatment (11) provides examples of removing benzene that has a poor Freundlich isotherm, $K = 1.0$, $1/n = 1.6$. Removal of PAH compounds using GAC would require far less carbon than that for benzene and be much more efficient.

High-energy chemistry processes for bond reduction and detoxification of PAHs are potentially feasible processes. High-energy chemistry is defined as

$$E > \kappa T \tag{8}$$

This energy is not supplied in the form of heat but through the impact either of an electron, ion, atom, molecule, or photon or by application of electric, magnetic or other fields (15). PAHs are expected to be less refractory (i.e., more reactive toward free radicals) than benzene. As previously discussed, the main problem for these processes is low solubility. However, some of them (e.g., Fenton's reagent) do not depend on light transmittance and can be readily implemented (11).

NOMENCLATURE

a	Constant
b	Constant
E	Energy ML^2/t^2
f	Fugacity Pa
f_{oc}	Fractional organic carbon in solid
G	Gas rate in stripping M/t/L^2
H	Henry's law constant
K	Freudlich isotherm coefficient (capacity constant) mg/g*(L/mg)$^{1/n}$
K_D	Distribution coefficient, L/M
K_{oc}	Organic-carbon distribution coefficient, L/M
K_{ow}	Octanol–water distribution coefficient, L/M
L	Liquid phase or liquid rate in stripping, M/t/L^2
$1/n$	Freundlich bonding constant
R	Stripping factor
S	Solid phase
x	Concentration, M/L^3
V	Vapor phase
κ	Boltzmann's constant J/K

BIBLIOGRAPHY

1. U.S. Environmental Protection Agency. (1980). *POM Source and Ambient Concentration Data: Review and Analysis*. EPA Report No. 600/7-80-044, Washington, DC.
2. U.S. Environmental Protection Agency. (1975). *Scientific and Technical Assessment Report on Particulate Polycyclic Organic Matter (PPOM)*. EPA Report No. 600/6-75-001, Washington, DC.
3. Watts, R.J. (1998). *Hazardous Wastes: Sources, Pathways, Receptors*. John Wiley & Sons, New York.
4. U.S. Environmental Protection Agency. (1978). *Health Assessment Document for Polycyclic Organic Matter*. EPA Report No. 2/102, External Review Draft. Research Triangle Park, NC, pp. 3–1 to 3–47.
5. Neff, J.M. (1979). *Polycyclic Aromatic Hydrocarbons in the Aquatic Environment: Sources, Fates and Biological Effects*. Applied Science, Essex, UK.
6. Morrison, R.T. and Boyd, R.N. (1978). *Organic Chemistry*, 3rd Edn. Compounds. Allyn and Bacon, Chap. 30.
7. WHO. (1987). Polynuclear Aromatic Hydrocarbons (PAH). *Air Quality Guidelines for Europe*. Copenhagen, World Health Organization Regional Office for Europe, pp. 105–117.
8. Agency for Toxic Substances and Disease Registry. (1994). *Toxicological Profile for Polycyclic Promatic Hydrocarbons (PAHs): Update*. US Department of Health and Human Services, Public Health Services, Atlanta, GA.
9. IARC. (1983). *Polynuclear Aromatic Compounds. Part 1. Chemical, Environmental and Experimental Data*. Lyon, International Agency for Research on Cancer. (IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Vol. 32).
10. Ashworth, S.C. (1998). *K Basin Sludge Polychlorinated Biphenyl Removal Technology Assessment*. HNF—3095 DOE Information Bridge. Available: http://www.osti.gov/bridge/.
11. Ashworth, S.C. (2004). *Hydrocarbon Treatment Techniques*. IW71.
12. Thibodeaux, L.J. (1979). *Chemodynamics*. Wiley-Interscience.
13. Hemond, H.F. and Fechner, E.J. (1994). *Chemical Fate and Transport in the Environment*. Academic Press.
14. Pontious, F.W. (1990). *Water Quality and Treatment*, 4th Edn. McGraw-Hill, New York.
15. Bugaenko, L.T., Kuzmin, M.G., and Polak, L.S. (1993). *High-Energy Chemistry*. Ellis Horwood PTR Prentice-Hall.

HYDROCARBON TREATMENT TECHNIQUES

SAMUEL C. ASHWORTH
Idaho National Engineering &
Environmental Laboratory
Idaho Falls, Idaho

There are various criteria concerning hydrocarbons for water needs. Some water users have little concern for dissolved hydrocarbons (e.g., agriculture), and some have strict standards (e.g., drinking water). In general the presence of organic compounds, including hydrocarbons, in water can lead to a buildup of biofilms. Biofilms can be problematic in many situations including drinking water, cooling water, recirculating water, and water used for various other industrial processes. During biofilm formation, the first substances on the surface are not bacteria but trace organics (1). Almost immediately after the clean pipe surface comes into contact with water, an organic layer deposits on the water/solid interface (2). These organics form a conditioning layer that neutralizes excessive surface charge and surface free energy that may prevent a bacteria cell from approaching near enough to initiate attachment. In addition, the adsorbed organic molecules often serve as a nutrient source for bacteria.

Though there are many criteria, the thrust of this article is on the treatment, including the removal and destruction of hydrocarbons for whatever purpose. It is assumed in this article that the hydrocarbons are dissolved in water, so that technologies such as incineration or decantation will not apply. The focus will be on water-phase separation and destruction. No economic evaluation is considered. In reality, the trade-offs would need to be considered, for example, waste disposal costs, electrical cost, and materials.

BACKGROUND

In general, water from natural sources will lead to poor cooling performance unless adequate steps are taken to remove hydrocarbon contaminants (3). The extent of the problem, however, will depend on the quality of the raw makeup water. The deposits that occur on heat exchanger surfaces are complex and are likely to include particulate matter, crystalline salts, corrosion products, and biofilms. It is necessary to counteract this problem of deposition onto surfaces to maintain heat exchanger effectiveness.

The shape and structural arrangement of a biofilm growing in a flowing fluid will influence the mass transfer characteristics of the biofilm system as well as the drag force exerted on individual biofilm structures (4). If the biofilm is a highly compliant material, the shape will vary through the growth cycle of the biofilm and also due to variations in fluid shear stress (5). Fluctuations in biofilm shape will also affect the hydrodynamic drag that in turn will influence the detachment rate and pressure losses in a flowing system. In addition, it is thought that biofilm viscoelasticity may explain the large pressure drops observed in biofilm fouled pipes (6).

There are also regulatory standards for certain waters (e.g., drinking water and water discharged to aquifers and bodies of water). Some of these that might apply include the Safe Drinking Water Act (42 U.S.C. 300f et seq.), Clean Water Act (33 U.S.C. 1251 et seq.), Resource Conservation and Recovery Act (42 U.S.C. 6901 et seq.), and a host of others. Many of them provide standards for organic compounds including hydrocarbons in effluents and source waters.

The above are a few reasons for removing organic compounds but there are others, including the pharmaceutical and semiconductor industries that have specific needs and criteria in this regard.

TREATMENT

Separation

Separation is the process of removing a compound from one phase to another (e.g., activated carbon adsorption). One of the simplest methods for separating volatile hydrocarbons and other organic compounds from water is air stripping. An example is benzene, a ring molecule shown in Fig. 1.

A typical stripper is a packed system using Raschig rings or other packing that provides high surface area, as shown in Fig. 2.

Descending water containing dissolved benzene enters the top of the column and is distributed to flow through the packing. The air enters at the bottom contacts the descending water, and the benzene transfers from the water to the air, an interphase mass transfer operation. The clean water is used for the purposes required, and the air is ejected to the atmosphere, regulations permitting, or treated accordingly (e.g., vapor-phase activated carbon).

The driving force for mass transfer is the gradient from the water to the air. At the interface, it is normally assumed that the interfacial water concentration and the air partial pressure are in equilibrium. If the amount of dissolved organic in the water is small enough (water mole fraction not appreciably different from 1.0), equilibrium is expressed using Henry's law (7). Henry's law is a special case of the following relation for gas–liquid

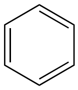

Figure 1. Benzene molecule.

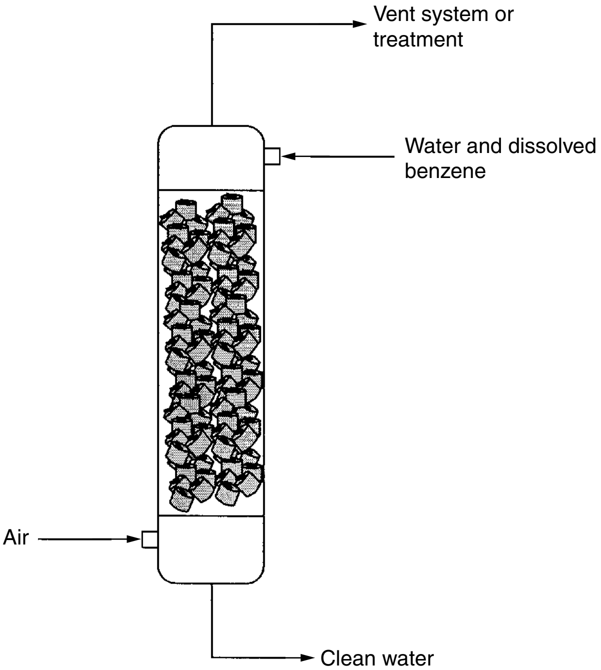

Figure 2. Packed-bed air stripper.

thermodynamic equilibrium[a]:

$$y_i \phi_i P_T = x_i P^{vap} \gamma_i$$

$$p_i = \frac{P^{vap} \gamma_i}{\phi_i} x_i = H x_i$$

Or the partial pressure of i equals Henry's constant times the liquid mole fraction of i where H (Henry's constant) is constant for small mole fractions in the liquid phase. The rate of mass transfer is based on using Henry's constant. For a dissolved gas, there are two resistances, gas-phase and liquid-phase. The mass transfer (F) flux is

$$F = K_L(C_B - C^*) = K_G(p^* - p_B)$$

The C^* and p^* above are virtual properties; C^* is the concentration in equilibrium with the bulk pressure if there were a liquid present. Similarly, p^* is the partial pressure that would be in equilibrium with the bulk liquid. Figure 3 illustrates these relations. Because C^* is not known, p_B/H is substituted for it. Similarly, C_BH is substituted for p^*. This is the easiest form as all that needs to be known are the bulk liquid concentration and bulk partial pressure.

The mass transfer based on individual coefficients is (see Fig. 4)

$$F = k_G(p_i - p_B) = k_L(C_B - C_i)$$

By combining these and using Henry's Law, the following is determined for the overall mass transfer coefficients either one can be used:

$$K_L = \frac{1}{\dfrac{1}{Hk_G} + \dfrac{1}{k_L}}$$

$$K_G = \frac{1}{\dfrac{1}{k_G} + \dfrac{H}{k_L}}$$

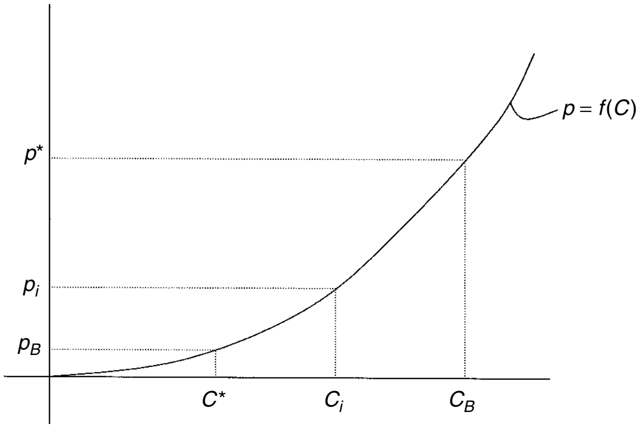

Figure 3. Stripper concentration and partial pressure relations.

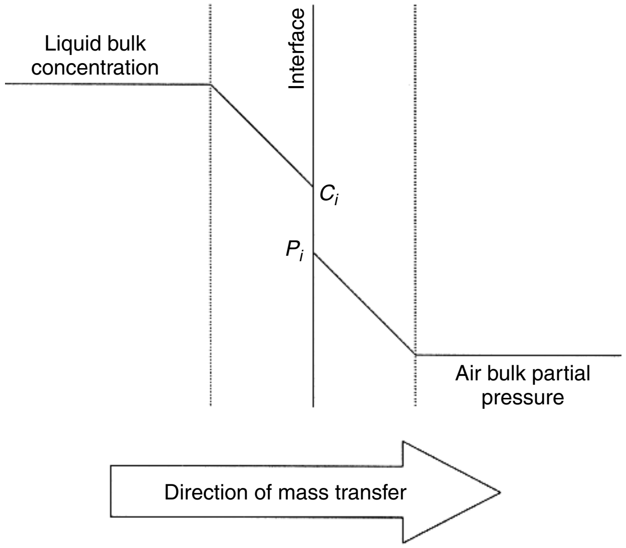

Figure 4. Interphase mass transfer.

There are many references that cover this thoroughly (8,9,11) but if Henry's Law applies, the height of a stripping column is

$$h = HTU * NTU$$

$$HTU = \frac{L}{K_L a}$$

$$NTU = \frac{R}{R-1} * \ln \frac{C_i/C_o * (R-1) + 1}{R}$$

$$R = \frac{HG}{L}$$

Using benzene as an example with an MCL of 5 μg/L, the height will be determined to achieve this starting with 50 μg/L and the data in Table 1. This would take a stripper approximately 5 m high (15–20 ft) to meet the MCL for benzene.

The R parameter is called the stripping factor and must be greater than one ($R > 1$) for stripping to be feasible. As observed by its definition, a small H can be somewhat overcome by a large G/L. The trend is such that volatile materials are easy to strip, whereas the semivolatiles and large organic compounds that have low vapor pressures are difficult to strip.

Adsorption

The use of activated carbon for organic removal from a water source or effluent is another very common and useful separation process. There are two types in use, granular and powdered. Only the granular type is discussed here; the powdered type is covered in other sources (10). Normally, the system is designed so that water flows into the top of the column and exits the bottom, as shown in Fig. 5.

Some methods for multicomponent adsorption are available. One is the multicomponent Langmuir that doesn't always give good predictions (11). There has also been some work on multicomponent systems (2). Some

[a] a Actually, $H_i \equiv \lim_{x_i \to 0} \dfrac{f_i^L}{x_i}$

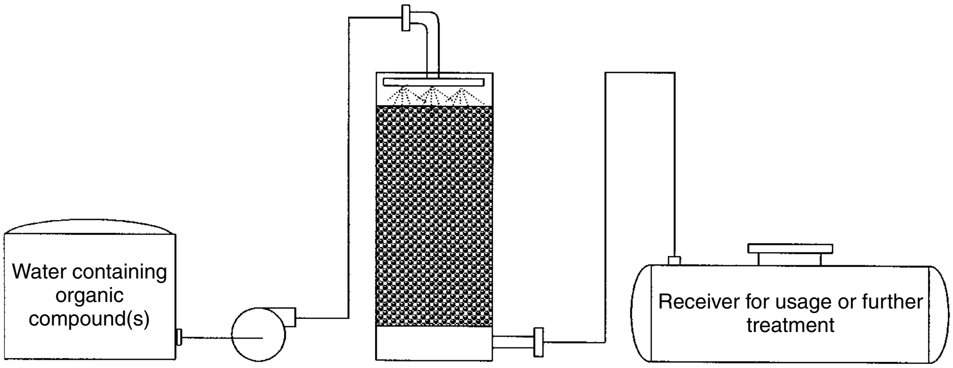

Figure 5. GAC system.

Table 1. Data for Air Stripping Benzene

Given Data	Values
C_i, μg/L	50
C_o, μg/L	5
G, $m^3/m^2/h$	1000
H, atm L/mg	5.52E−05
H, Dimensionless	1.80E−01
K_{L_a}, hr^{-1}	10^a
L, $m^3/m^2/h$	20
Flow, gpm	10
Calculated	
R	9.00
HTU, m	2
NTU	2.47
h, m	4.94
A, m^2	0.11
D, m	0.38

[a] This was an assumed value. Note that K_La is a combined, overall liquid mass transfer coefficient where a is the surface area to volume ratio of the packing and is usually measured or estimated as a single quantity. It is found similar to the K_G and K_L and there are correlations for it in many references, e.g., 8–10.

computer software is available for sizing but is limited to 8–10 components. Vendors also have programs and rule of thumb methods available. Freundlich isotherms for many of the compounds are readily available, and hence, this method is used for single organic compounds. The Freundlich isotherms are sometimes used for scoping multicomponent systems. However, the Freundlich method is not amenable to multicomponent systems. Therefore, it is assumed that they are additive, each compound adsorbs independently of the others without competition. Column dynamic testing of the actual liquid is preferred for multicomponent systems.

The Freundlich isotherm is explained by

$$q_e = KC_e^{1/n}$$

The additive type of method assumes that the organic compounds adsorb in layers; the strongest bonding molecules profile toward the top of the equilibrium zone, and the less strongly bonded toward the bottom (K is related to capacity, and $1/n$ is related to bonding strength). This is shown ideally in Fig. 6. It shows that the organic compounds that have the highest affinity for GAC profile toward the top of the column and those of low affinity profile toward the bottom. Hence, breakthrough of benzene would occur first if in a multicomponent system.

The procedure is to determine the amount of GAC for a component individually via the method of Snoeyink (Pontius 1990) (11) and determine the amount of GAC required. Using benzene again at 50 μg/L (it does not have a very favorable isotherm),

$$q_{e,Benzene} = 1 * 0.05^{1.6} = 0.0083 \text{ mg/g}$$

The amount of water per GAC, according to Snoeyink is,

$$Y = \frac{q_e}{C_i - C_o} * \rho_{\text{GAC}} = \frac{0.0083 \text{ mg/g}}{0.05\text{–}0.005 \text{ mg/L}} * 500 \text{ g/L}$$
$$= 92 \text{ gal/galGAC}$$

The amount of GAC is then the volume to be treated/Y. For 1000 gallons this is

$$\text{GAC} = \frac{1000}{92} = 11 \text{ gal}$$

or 11 gallons of GAC per 1000 gallons treated.

To avoid immediate breakthrough, the bed length must be greater than the mass transfer zone (MTZ). The MTZ is calculated from column dynamic tests as

$$\text{MTZ} = L\frac{t_s - t_b}{t_s}$$

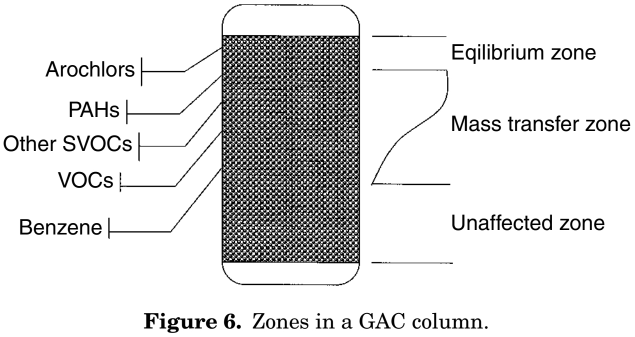

Figure 6. Zones in a GAC column.

The most important GAC adsorber design parameter is contact time, most commonly described by its empty bed contact time (EBCT) (11), V/Q. The empty bed contact time (EBCT) is an important design parameter in GAC column design. The range for 47 plants was 3–34 minutes, and the median was 10 minutes (Pontius 1990)(11). Longer contact times are normally used for highly concentrated solutions, so this design uses the value of 5 minutes. Based on this, the minimum column volume is

$$V_{\text{column}} = EBCT * Q = 5 \text{ min} * 10 \text{ gpm} = 50 \text{ gal}$$

A typical hydraulic loading is 3–4 gpm/ft^2. Using 10 gpm, the area is 2.5 ft^2. The height is

$$h = \frac{50 \text{ gal} * \text{ft}^3/7.48 \text{ gal}}{2.5 \text{ ft}^2} = 2.67 \text{ ft}$$

Destruction

The final technologies discussed for removing organic compounds from water are destruction/reaction techniques based on high-energy chemistry. High-energy chemistry is defined as

$$\text{E} > \kappa\text{T}$$

So-called advanced oxidation technology is a subset of high-energy chemistry that uses free radicals to oxidize the target compound. All of these processes work by generating free radicals followed by attack on the target's bonds by the free radicals. All of these processes are commercially available. However, some of them have certain restrictions on their use (e.g., nitrates, liquid only,). The following are some of the processes comprising advanced oxidation.

The UV/H_2O_2 process consists of a chemical reactor (batch or flow-through) that uses UV light to produce free-radical hydroxyls (Fig. 7)

$$\text{H}_2\text{O}_2 + h\nu \rightarrow \text{OH}\cdot$$

The free radicals extract electrons from the target and create a free radical target product. The ensuing mechanism is a chain reaction eventually degrading the target to CO_2 and H_2O and other simple, less regulated compounds. The overall reaction can usually be modeled by first-order kinetics. The design relation used for this type of reactor is the EE/o, the electrical energy per volume per order of magnitude (in kwh/1000 gal/order). This relationship provides all of the information required to determine efficiency (i.e., once known, any efficiency can be obtained by adjusting the power).

There are several similar systems that produce free-radical hydroxyls, including

- UV/ozone
- ozone/peroxide
- Fentons' reagent
- UV-vis/Peroxide

It has been speculated that UV is required for some compounds like polychlorinated biphenyls (PCBs), to activate the bonds, and OH^- is required to react with the activated bonds (i.e., synergism) (13).

For some compounds (e.g., NDMA), direct irradiation of organic pollutants by high-intensity UV light provides a significant destructive pathway. The target compounds degrade after absorbing UV light. The target must strongly absorb UV.

The hydrated electron is a very strong reducing agent that can react with halogenated alkanes and alkenes. The hydrated electron interacts with the chlorine–carbon bonding electrons and provide energy to break the bond and demineralize the target. There are several commercially available processes involving reduction via aqueous electrons. The aqueous electron is produced by several mechanisms including nuclear and high-energy processes, photochemistry, and chemical. Some processes require transferring high-energy electrons through thin films of water because the free path or linear energy transfer (LET) is small for electrons (high voltage process).

The baseline process is a UV-catalyzed process. The chemical added that is shown is a proprietary catalyst. The catalyst interacts with UV light shown by the energy arrow that produces the aqueous electron (e^-_{aq}). An example of this production reaction is

$$\text{Fe(CN)}_6^{-4} + h\nu \rightarrow \text{Fe(CN)}_6^{-3} + e^-_{\text{aq}}$$

The aqueous electron then interacts with the chlorine–carbon bond producing chloride ion and a free radical chlorinated aromatic:

$$e^-_{\text{aq}} + \text{C}_{12}\text{H}_5\text{Cl}_5 \rightarrow \text{C}_{12}\text{H}_5\text{Cl}_4^- + \text{Cl}^-$$

The free radical goes on to extract electrons from another molecule, and a chain reaction usually occurs. The rate constants for reductions are normally quite large. This process is particularly well suited to compounds not amenable to advanced oxidations, for example, CCl_4. The UV-induced process is commercially available (e.g., Calgon Advanced Oxidation Technologies).

Supercritical water oxidation, sometimes known as hydrothermal waste processing, uses the solvating traits of water in its supercritical state to destroy liquid organic wastes. As water is heated beyond its critical temperature (374.1 °C) and critical pressure (250 Mpa-s, about 3219 psi), the density of the water drops dramatically (typical operating densities are 0.15–0.2 g/cm^3). These changes in density and hydrogen bonding make organics highly soluble, and inorganic substances become nearly insoluble. The organic material is dissolved in an oxygen-rich environment where conversion occurs rapidly due to the high temperature of the process. Under such high pressure and temperature, organic materials are rapidly decomposed by oxidation at removal efficiencies of 99.9999% or greater.

Ultrasound can induce unusual high-energy chemistry through the process of acoustic cavitation, the formation, growth, and implosive collapse of bubbles in a liquid. Cavitation can occur in both clouds of collapsing bubbles (multi-bubble cavitation) or with high symmetry for isolated bubble (single-bubble cavitation, SBC). Multi-bubble cavitational collapse produces localized, transient hot spots with intense local heating (approx > 5000 K), high pressures (>2000 atm), and short lifetimes (sub-microsecond) in an otherwise cold liquid. From hydrodynamic modeling of this cavitational collapse, it has been estimated that both the heating and cooling rates are in excess of 1010 K/s. Acoustic cavitation is a unique means of creating high-energy chemistry, easily and inexpensively. Aqueous sonochemistry produces, supercritical water conditions on a microscopic scale. In this regime of temperature and pressure, the sonochemistry of water is an extreme limiting case of supercritical phenomena and is closely related to hydrothermal oxidation. For example, the ultrasonic irradiation of water produces a variety of extraordinarily reactive species (including OH^-, H^+, and HO_2^-) that can decompose many organic compounds.

The last design example is treating benzene-contaminated water to the same degree as the previous two examples using UV/H_2O_2. Figure 7 illustrates a schematic batch test reactor. Benzene has an EE/o = 2–5 kWh/1000 US gal/order (14). Therefore, the dosage of UV required in a flow-through system at 10 gal/min is (using 5 for the worse case) and the industry design relation (14):

$$\begin{aligned}&\text{Power(kWatt)}\\&= \frac{5\ \text{kWh} * 10\ \text{gal/min/order} * \log_{10}(50/5) * 60\ \text{min/hr}}{1000}\\&= 3\ \text{kW}\end{aligned}$$

So a UV lamp of 3 kW is sufficient to meet the requirement. Of course, if the flow rate were 100 gpm, it would require 30 kW, and so on. There will be economic trade-offs, electrical power costs, and chemical costs.

It is highly useful at this point to consider reaction kinetics and reactor design as there are some unusual features for this type of reactor. For a batch system, as shown in Fig. 8, the material balance and first-order reaction are

$$\frac{dC}{dt} = -kC$$

$$\ln\left(\frac{C_i}{C_o}\right) = kt$$

The first-order reaction coefficient is inversely proportional to the EE/o:

$$k = \frac{\ln 2 * P(kW)}{V * \text{EE/o} * \log 2}$$

The power for a batch reactor is then the same used for the industry design equation (14):

$$P(kW) = \frac{\ln(C_i/C_o) * V * \text{EE/o} * \log 2}{\ln 2 * t}$$

In a plug flow reactor (see Fig. 8), a material balance on a differential volume element yields:

$$QC - Q(C + dC) - kCdV = V\frac{dC}{dt}$$

At steady-state,

$$\ln\frac{C_i}{C_o} = k\frac{V}{Q}$$

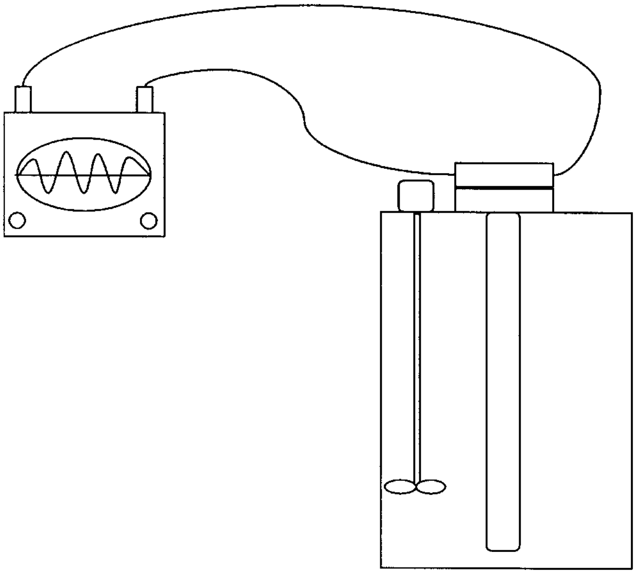

Figure 7. Batch reactor with lamp and agitator.

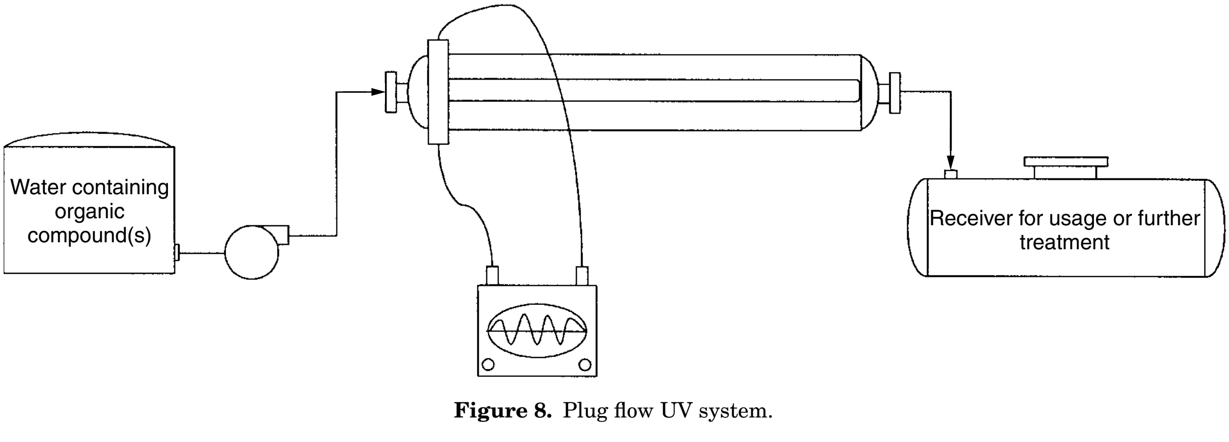

Figure 8. Plug flow UV system.

The power, based on this k, is then the industry design equation (14).

$$\mathrm{P(kW)} = \frac{\mathrm{EE/o} * \log 2 * Q * \ln(C_i/C_o)}{\ln 2}$$

For a continuously stirred reactor (CSTR), the steady-state material balance yields

$$\frac{C_i}{C_o} = \frac{kV}{Q} + 1$$

This does not yield the industry design equation. The answer to this seemingly paradoxical result is that the plug flow design applies, regardless of the reactor type for flow-through reactors. This is so because the reaction must be in the vicinity of the lamp surface; a solute in a CSTR must, at some time in its residence history, approach the lamp sufficiently close. This is one reason that flow-through reactors require appropriate design to obtain turbulence.

Nomenclature

A	Cross sectional area
C_B	Bulk liquid concentration
C_e	Equilibrium concentration
C_i	Inlet concentration
C_o	Outlet concentration
$C*$	Fictitious concentration in vapor phase
D	Diameter
EE/o	Electrical energy/1000 gallons/order
F	Mass transfer flux
G	Gas mass velocity
H	Henry's Law constant
$h\nu$	Photon energy
HTU	Height of a transfer unit
k	1st order reaction coefficient
K	Freundlich constant (related to capacity)
K_G	Overall mass transfer coefficient based on gas
K_L	Overall mass transfer coefficient based on liquid
k_G	Gas phase mass transfer coefficient
k_L	Liquid phase mass transfer coefficient
K_{L_a}	Overall liquid phase mass transfer coefficient
L	Liquid mass velocity, length of column
NTU	Number of transfer units
$1/n$	Freundlich exponent (related to bonding strength)
P	Power
P_T	Total pressure
p_B	Bulk gas partial pressure
p_i	Partial pressure at interface
P^{vap}	Pure component vapor pressure
$p*$	Fictitious pressure in liquid phase
q_e	Quantity adsorbate/adsorbent
Q	Liquid flow rate
R	Stripping factor
t_b	Breakthrough time
t_s	Saturation time (influent = effluent)
V	Volume
x_i	Liquid phase mole fraction
y_i	Gas phase mole fraction
γ_i	Activity coefficient
κ	Boltzmann's constant
ϕ_i	Fugacity coefficient

BIBLIOGRAPHY

1. Edstrom. (2002). Available: http://www.edstrom.com/Resources.cfm?doc_id=23.
2. Mittelman, M.W. (1985). Biological fouling of purified-water systems: Part 1. *Bacterial Growth and Replication, Microcontamination* **3**(10): 51–55, 70.
3. Bott, T.R. (1998). Techniques for reducing the amount of biocide necessary to counteract the effects of biofilm growth in cooling water systems. *Applied Thermal Engineering* **18**(11).
4. Stoodley, P., Lewandowski, Z., Boyle, J.D., and Lappin-Scott, H.M. (1999). Structural deformation of bacterial biofilms caused by short-term fluctuations in fluid shear: An *in situ* investigation of biofilm rheology. *Biotechnology and Bioengineering* **65**(1).
5. Boyle, J.D., Dodds, I., Stoodley, P., and Lappin-Scott, H.M. (1997). Stress management in biofilms. In: *Biofilms: Community Interactions and Control*. J.W.T. Wimpenny, P.S. Handley, P. Gilbert, H.M. Lappin-Scott, and M. Jones (Eds.). BioLine, Cardiff, UK, pp. 15–22.
6. Picologlou, B.F., Zelver, N., and Characklis, W.G. (1980). Biofilm growth and hydraulic performance. *J. Hydraul. Div. Am. Soc. Civ. Eng.* **106**(HY5): 733–746.
7. Prausnitz, J.M., Lichtenthaler, R.N., and de Azevedo, E.G. (1986). *Molecular Thermodynamics of Fluid-Phase Equilibria*. Prentice-Hall, Englewood Cliffs, NJ
8. Perry, R.H. and Green, D.W. (1984). *Perry's Chemical Engineers' Handbook*, 6th Edn. McGraw-Hill, New York.
9. Treybal, R.E. (1987). *Mass-Transfer Operations*. McGraw-Hill, New York.
10. Perrich, J.R. (1981). *Activated Carbon Adsorption for Wastewater Treatment*. CRC Press, Boca Raton, FL.
11. Pontius, F.W. (1990). Adsorption of organic compounds. *Water Quality and Treatment*. 4th Edn. McGraw-Hill, New York.
12. Khan, A.R., Al-Bahri, T.A., and Al-Haddad, A. (1997). Adsorption of phenol based organic pollutants on activated carbon from multi-component dilute aqueous solutions. *Wat. Res.* **31**(8): 2102–2112.
13. Ashworth, S.C. (1998). *K Basin Sludge Polychlorinated Biphenyl Removal Technology Assessment*. HNF–3095 DOE Information Bridge. Available: http://www.osti.gov/bridge/.
14. Calgon Carbon Corporation. (1996). *Advanced Oxidation Technologies Handbook* **1**(1): MI-AOT-10/96.

USE OF ANAEROBIC-AEROBIC TREATMENT SYSTEMS FOR MAIZE PROCESSING PLANTS

María-del-carmen Durán-de-bazúa
UNAM, National Autonomous
University of Mexico
University City, México

After the discovery of the American continent by the Spaniards in the fifteenth century, corn (*Zea mays*), an indigenous plant from Mexico, spread all over the world and became a staple food for many human groups. However, the traditional ways of eating the many products derived from this cereal in Mexico, inherited from the Mayan, Aztec, and other Mexican indigenous groups, were not disseminated until the last 20 years. Due to globalization, one of these maize products, "tortillas," is now an extremely popular food that can be found in most countries of the world. "Tortillas" are a sort of unleavened bread in circular form that, due to a mild flavor, can be combined with vegetables, meats, pulses, etc. They are the bread equivalent to "chapattis" for the people from India and "falafel" for the people of the Middle East (Fig. 1).

The traditional production of "tortillas" involves a Precolumbian technique, known as "*nixtamalization*," derived from the Aztec words *nextli* = lime ashes and *tamalli* = cooked corn dough. This ancient process, almost as old as corn domestication and cultivation, is a time- water- and energy-consuming technique (Fig. 2). Modernization of the traditional process to produce cornmeal instead of dough to lengthen its shelf life as well as some other changes for mass production have been introduced in the last 50 years. However, these changes affect the sensory characteristics of "tortillas," mainly "rollability," "sturdibility," and softness, so it has been a common practice among large-scale producers to introduce some chemical additives to make the new "supermarket tortillas" desirable from the sensory point of view. Unfortunately, from the nutritional point of view, it has not yet been proved that these additives do not affect the health of consumers. Some innovative processes that maintain these desirable nutritional and sensory characteristics, and at the same time, reduce energy and water consumption, and most importantly, processing time, have also been developed and started to be used (2–5). However, the traditional plants (more than 100,000 small-scale and 25 large-scale in Mexico) are still in operation, so methods to treat and stabilize the wastewater generated in the traditional process (so-called "nixtamal" mills that process around 0.5 to 4 metric tons maize per day) or in its industrial modifications (so-called precooked cornmeal or masa-harina factories that process from 100–1,000 metric tons maize per day) and contain appreciable amounts of soluble and insoluble organic matter (Table 1) have also been developed. These methods are based on anaerobic and aerobic systems, and the interesting issue is that by-products from the treatment can be further used, giving added value to these treatment processes.

The biological processes developed were screened out after testing at laboratory level. These processes are commonly used in biological wastewater treatment, such as activated sludge, aerated lagoons, facultative lagoons, low and high rate anaerobic reactors, etc. (7). Figure 3 shows the method developed for treating the wastewater from maize alkaline processing to produce "tortillas," particularly for those factories that process considerable amounts of corn. After primary settling to recover broken maize pieces and peelings, the anaerobic process is carried out to transform most of the biodegradable dissolved matter in the wastewater into methane-rich biogas. This is followed by the aerobic process that polishes the treatment, removes the traces of sulfide and other undesirable compounds in treated water, and converts the remaining biodegradable matter into protein-rich biomass. A tertiary treatment using activated carbon or any organic adsorbent will render recyclable water for the process. The water absorbed by corn is added as makeup freshwater to the overall process. The recycling of the by-products generated is shown in Fig. 3. The amount of methane from biogas obtained from the anaerobic treatment of the wastewaters is 9.6–16.8 m^3 per metric ton of maize processed (considering 5 m^3 of wastewater produced per metric ton maize at a conversion of 80% of the dissolved biodegradable matter into biogas). If this gas contains roughly 80% methane, the yield of this energy-rich gas is 7.7–13.48 m^3 methane per metric ton of maize processed. This methane-rich biogas may be washed and compressed to be used as a secondary source of energy during the lime-cooking of corn grains (nixtamalization). The amount of biomass and suspended solids obtained may be processed by extrusion to form pellets or flakes for fish feedstuff, considering the mass balances from the aerobic and anaerobic processes is of 23 kg SS and 10.6 kg biomass per metric ton of maize cooked. These biomass and suspended solids pellets or flakes proved comparable

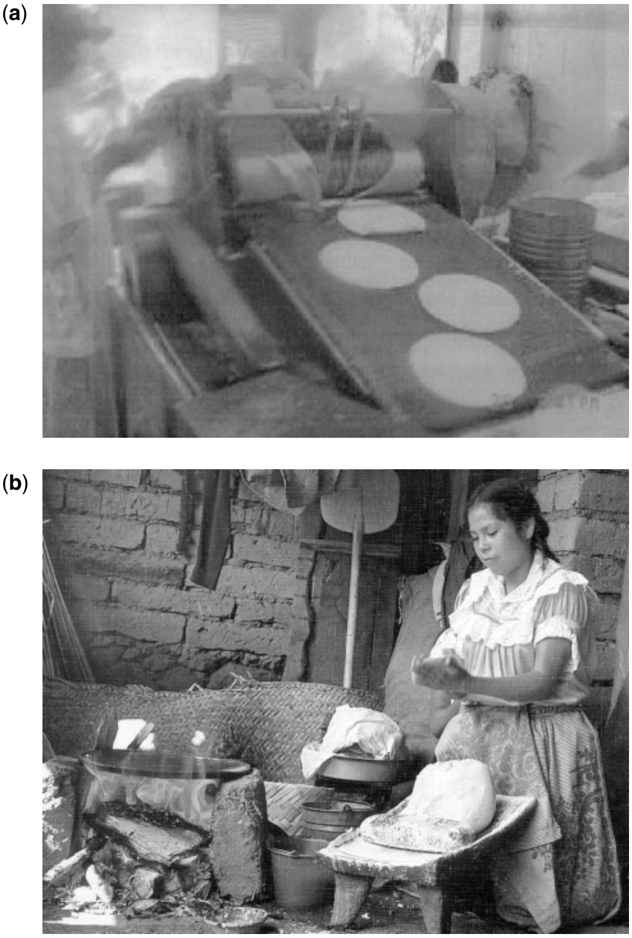

Figure 1. (**a**) Modern "tortilla" making machine. (**b**) Traditional "tortilla" hand-making (1).

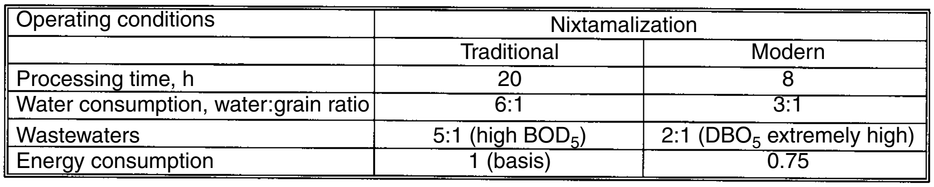

Operating conditions	Nixtamalization	
	Traditional	Modern
Processing time, h	20	8
Water consumption, water:grain ratio	6:1	3:1
Wastewaters	5:1 (high BOD_5)	2:1 (DBO_5 extremely high)
Energy consumption	1 (basis)	0.75

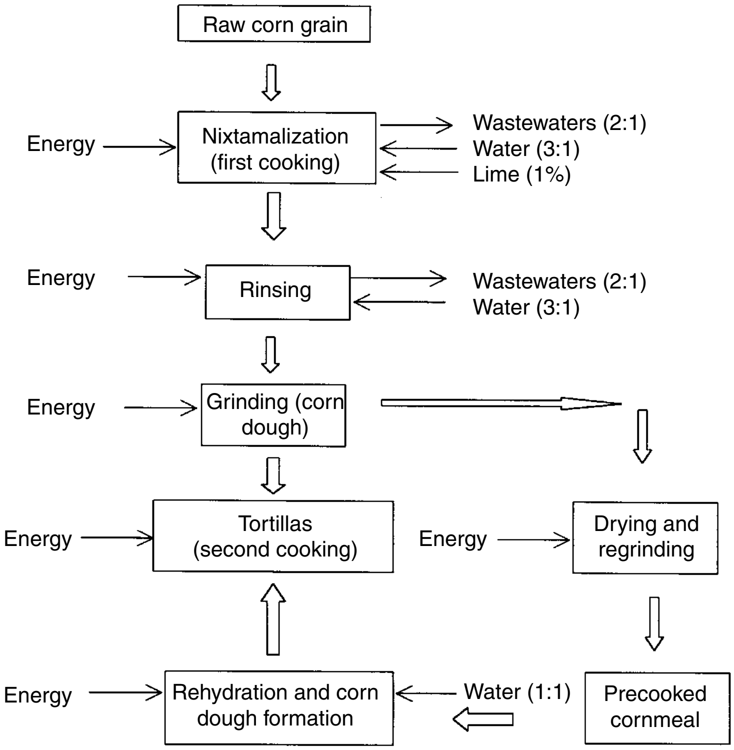

Figure 2. Diagram of the traditional nixtamalizatiŏn process (2).

Table 1. Maize Lime-Cooking (Nixtamalization) Wastewater Average Composition[a]

Characteristics	Average Values	UE Normativity[b] for Surface Water
Suspended solids, kg/m^3	2.4–4.6	0.025 (25 mg/L)
Dissolved organic carbon, kg/m^3	3.0–5.0	—
Biochemical oxygen demand (5 days, 20 °C), kg/m^3	1.5–3.0	<0.007 (7 mg/L)
Chemical oxygen demand (dissolved), kg/m^3	7.5–11.0	0.030 (30 mg/L)
Nitrogen (Kjeldahl), mg N/L	80–270	0.003
Phosphates, mg $PO_4{}^{-3}$/L	7–18	0.0007
pH Value	10–14	5.5–9.0
Color	Dark yellow	Colorless

[a] Reference 6.

[b] European Community: 16/6/75 (after physical and chemical treatment including desinfection).

to that for commercial feedlots for carp and other edible fish (8).

The reactors most suitable for this type of carbonaceous wastewater are those that develop active biofilms, particularly because the biomass formed tends to show more cohesion and separates more easily from the liquid phase by sedimentation. For the anaerobic step, packed bed anaerobic reactors or upflow anaerobic sludge blanket reactors are those that render the highest conversion into methane-rich biogas at the lowest hydraulic residence times between 24 and 48 hours. For the aerobic step, rotating biological reactors are excellent for transforming the soluble remaining biodegradable dissolved and colloidal organics into an easily settleable aerobic biomass. The energy consumption for the rotational speed needed is very low, and the area occupied by the reactors is also very small. Figure 4 shows the system to be used in cornmeal factories. Both systems, the anaerobic and the aerobic, are compact and easy to operate. The saving in water and energy consumption, wastewater disposal costs, and the sale of biomass for fish protein-rich feedlots make the treatment and recovery process economically feasible. The

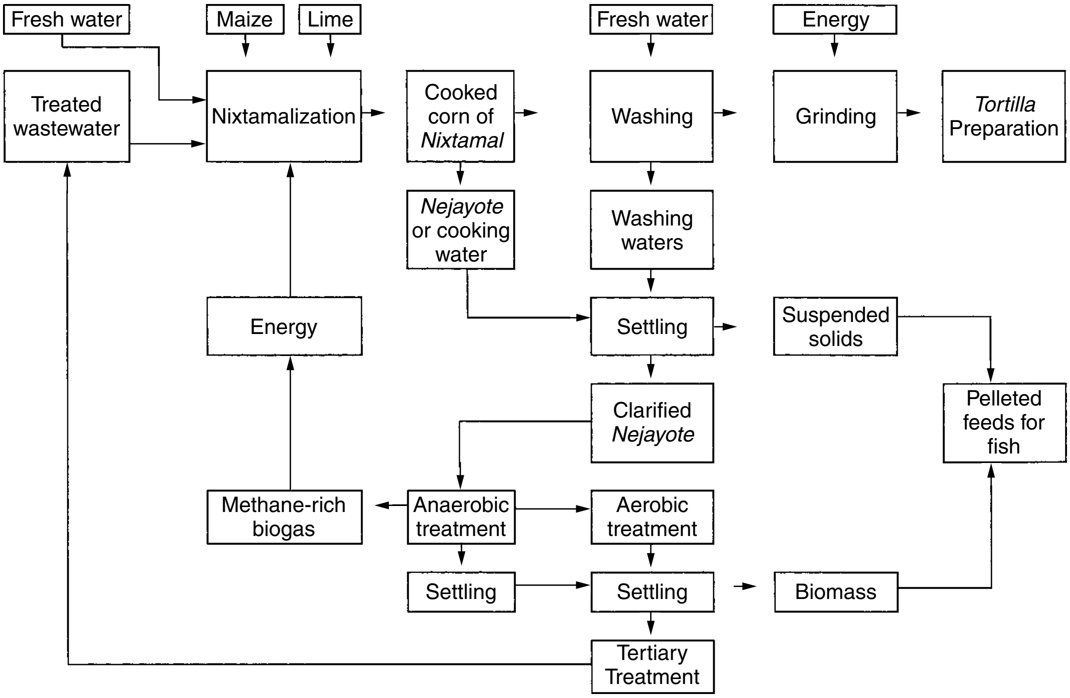

Figure 3. Wastewater treatment for maize alkaline cooking process for "tortilla" preparation ("*nixtamal*", "*nejayote*", Aztec derived words from ***nextli*** = lime ashes, ***tamalli*** = cooked maize dough, ***ayotl*** = broth, watery solution).

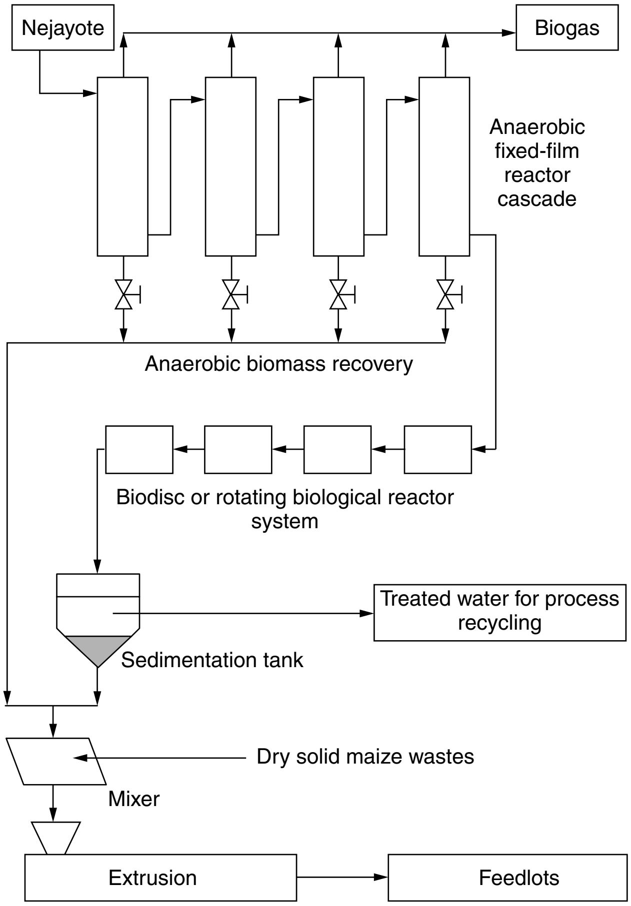

Figure 4. Wastewater treatment in a precooked cornmeal factory that uses a combined anaerobic-aerobic treatment system (7).

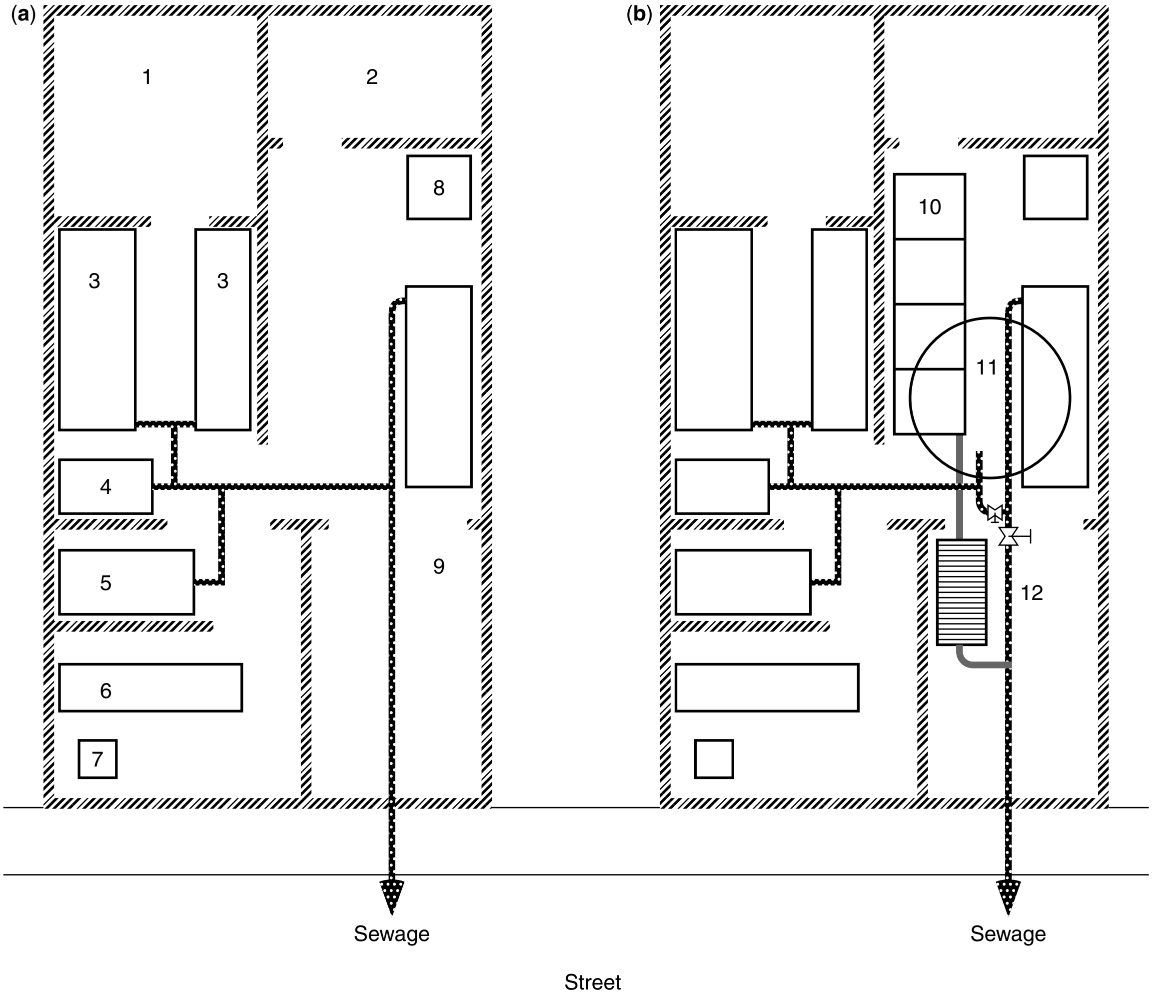

Figure 5. Wastewater treatment in a corn traditional mill (7). (**a**) Unimproved mill: (1) grain storage, (2) boiler, (3) nixtamalization tanks, (4) wash water tank, (5)stone grain mill, (6) selling shelf, (7) scales, (8) water reservoir, and (9) lime and corn solid waste storage room. (**b**) Changes for introducing the aerobic wastewater treatment system that includes an (11) underground settling tank with a sludge pump attached to the bottom exit pipe that sends the sludge to a (12) filter press, and the filtered wastewater to a (10) rotating biological reactor.

turnover rate is 1.5 years, that is, the money invested in the wastewater treatment plant together with the accessories to recover biogas and biomass and to reuse them is recovered in 18 months.

Due to the size of nixtamal mills, it is not so economical to install an anaerobic-aerobic process for treating their wastewaters (between 2.5 to 20 m^3). An aerobic system is then used for these small establishments (Fig. 5). The solid material collected from the filter press can be fed directly to small farm animals and, although the protein cannot be directly assimilated by the animals, it does not damage them. In this case, partially treated wastewater is sent to the municipal sewer system. Finally, it should be mentioned that this combined anaerobic-aerobic process may be applicable to wastewater from other food industry sectors, particularly to those that contain carbonaceous substances. When proteinaceous wastewater is to be degraded, the wastewater may be anaerobically treated, but the gases must be washed to remove the sulfur compounds separated from protein matrices before using the biogas as a combustion or fuel source. Hydrogen sulfide is extremely corrosive and toxic, but once dissolved in water, may be used by aerobic organisms to be reconverted into biomass protein.

BIBLIOGRAPHY

1. Bonfil, G. (Ed.). (1984). *SEP Cultura: El Maíz, Fundamento de la Cultura Popular Mexicana*. Museo Nacional de las Culturas Populares. G.V. Editores, S.A. México D.F. México.
2. Durán-de-Bazúa, C. (1972). *Report on Novel Technologies: Application of the Extrusion Process for the Production of Mexican Traditional Corn Foods*. University of California, Berkeley, College of Nutritional Sciences, Berkeley, CA.
3. Durán-Domínguez, M.C. (1978). *Procedure for Cooking Corn by Extrusion/Procedimiento Para Cocer Maíz or Extrusion*. Patent 172126. April 7. Dir. Gral. Invenciones y Marcas. Dept. Patentes. Secretaría de Patrimonio y Fomento Industrial, México D.F. México.

Next Page

4. Durán-Domínguez-de-Bazúa, M.C. (1992). *Procedure for Cooking Cereals and Other Grains (Legumes and Oilseeds) by Alkaline Extrusion/Procedimiento para Cocer Gramíneas (Cereals, Leguminosas y Oleaginosas) por Extrusion Alcalina*. Patent requested on May 13. Dir. Gral. Invenciones y Marcas. Dept. Patentes. Secretaría de Patrimonio y Fomento Industrial, México D.F. Mexico.
5. Durán-Domínguez-de-Bazúa, M.C., Sánchez-Tovar, S.A., and Solís-Fuentes, J.A. (2002). *Procedure for Cooking Grains (Cereals, Legumes, and Oilseeds) and Agroindustrial by-Products by Alkaline Extrusion/Procedimiento Para Cocer Gramíneas (Cereals, Leguminosas y Oleaginosas) y Subproductos Agroindustriales por Extrusion Alcalina*. Patent requested on November 16. Instituto Mexicano de la Protección Industrial, México D.F. Mexico.
6. Durán-de-Bazúa, C. and Hartmann, L. (1979). *Produktion Von Biomasse aus Prozessabwaessern Der Lebensmittelindustrie*. Internationales Seminar, Universität Karlsruhe, 7 IIBB-UNIKA, Karlsruhe, German Federal Republic.
7. Durán-Domínguez-de-Bazúa, M.C. (1987). Effluents of the food industry in México: Environmental impacts on soil and water resources and possible solution using the biotechnological approach. Case problem: The corn industry. In: *Global Bioconversions*. Vol. II, Chap. 4. CRC Press, Boca Raton, FL, pp. 75–119.
8. Durán-Domínguez, M.C., Pedroza-Islas, R., Rosas-Vázquez, C., Luna-Pabello, V.M., Sánchez-Zamora, A., Capilla-Rivera, A., Paredes-Gómez, L., Valderrama-Herrera, S.B., and Vázquez-Cedeño, I. (1991). Producción de alimentos para peces: Utilización de subproductos del tratamiento de aguas residuales. In: *Premio Nacional Serfín El Medio Ambiente*. J.J. De Olloqui (Ed.). Futura Eds., Villa Nicolás Romero, Méx., Mexico, pp. 79–106.

BONDING OF TOXIC METAL IONS

E.N. Peleka
Kostas A. Matis
Aristotle University
Thessaloniki, Greece

T. Erwe
V. Mavrov
Saarland University
Saarbrücken, Germany

Worldwide environmental problems are becoming more and more acute. It is known that many specific industrial wastewater streams with large flows contain toxic metals in concentrations up to 500 mg/L, which have to be removed prior to water recycling, indirect discharge into a sewage system, or direct discharge into surface waters. Some examples are (1) rinse water in metal working enterprises, containing Cu, Ni, Zn, Sn, Cd, Pd, Ag, Al, Au, Cr, Mo, etc.; (2) rinse water in the semiconductor industry, containing Cu, Sn, Pb, Sb; (3) wastewater from the copper industry, containing As, Se; (4) other types of water in which the presence of toxic metals causes problems; (5) drinking water from ground water contaminated with As; and (6) mine water contaminated with different toxic metals (e.g., U, Ra, As).

When discharged directly into rivers, polluted wastewater poses a great risk to the aquatic ecosystem, and discharge into a sewage system negatively affects biosludge activity and leads to contamination of the sludge to be disposed of. As a result of the standards specified in the Water Resources Act, which was amended in 1986, industry takes precautions against these risks by treating dangerous components in a partial stream before mixing with other types of wastewater. So, research was focused, among others, on the development of highly selective bonding agents (often denoted as BA) with fast reaction kinetics for removing heavy metal ions (Fig. 1a). The latter may lead to better use of the capacity of bonding agents, resulting in smaller units and low residual concentrations of toxic metals in treated water streams, to comply with individual standards for water reuse or discharge—sorption kinetics constitutes another article in this *Encyclopedia of Water*.

In addition to zeolites (1,2), other typical materials extensively used are apatites (3,4) and various iron sorbents (5,6). These BA materials are broadly classified in three classes: carbon sorbents (active carbons, activated carbon fibers, molecular carbon sieves, fullerenes, heterofullerenes, nanomaterials), mineral sorbents (silica gel, activated alumina, oxides of metals, hydroxides of metals, zeolites, clay minerals, inorganic nanomaterials) and others (synthetic polymers, composite sorbents, mixed sorbents, etc.).

A review paper on bonding agents recently presented to a conference (7) also mentioned new types such as mesoporous materials and chelating resins and their potential applications to metal bonding from aqueous solutions. It also presented the equilibrium sorption capacity for metal ions of certain sorbents and stated that the functionalized silica materials have shown higher uptake capacity than that of other materials found in the literature. An inherent disadvantage of these materials is their low loading capacities and relatively small metal ion binding constants.

To circumvent these limitations, promising heavy metal sorbents have been prepared by coupling chelating ligands (e.g., thiol, amine, or crown ether functions) to support matrices consisting of inorganic oxides (e.g., silica, alumina, or clay) or organic polymers (e.g., polystyrene, cellulose, or polymethyl-methacrylate). Such functionalized materials have relatively high metal ion loading capacities and strong binding affinities for selected metal ions. This performance can be attributed to the presence of the surface-bound ligands, which can be specifically tuned to accommodate the selective sorption of targeted metal ions. Although superior in performance to conventional ion exchangers, functionalized matrices remain relatively inefficient because only a fraction of the immobilized ligands is accessible for metal complexation. The discovery of mesoporous metal oxide molecular sieves (by Mobil researchers) and organoceramic sorbents has given rise to new prospects for sorbent and catalyst design (7).

The knowledge of the governing chemical state of the metal ion to be removed is a prerequisite, so that any new development is based on a solid scientific

4. Durán-Domínguez-de-Bazúa, M.C. (1992). *Procedure for Cooking Cereals and Other Grains (Legumes and Oilseeds) by Alkaline Extrusion/Procedimiento para Cocer Gramíneas (Cereals, Leguminosas y Oleaginosas) por Extrusion Alcalina*. Patent requested on May 13. Dir. Gral. Invenciones y Marcas. Dept. Patentes. Secretaría de Patrimonio y Fomento Industrial, México D.F. Mexico.
5. Durán-Domínguez-de-Bazúa, M.C., Sánchez-Tovar, S.A., and Solís-Fuentes, J.A. (2002). *Procedure for Cooking Grains (Cereals, Legumes, and Oilseeds) and Agroindustrial by-Products by Alkaline Extrusion/Procedimiento Para Cocer Gramíneas (Cereals, Leguminosas y Oleaginosas) y Subproductos Agroindustriales por Extrusion Alcalina*. Patent requested on November 16. Instituto Mexicano de la Protección Industrial, México D.F. Mexico.
6. Durán-de-Bazúa, C. and Hartmann, L. (1979). *Produktion Von Biomasse aus Prozessabwaessern Der Lebensmittelindustrie*. Internationales Seminar, Universität Karlsruhe, 7 IIBB-UNIKA, Karlsruhe, German Federal Republic.
7. Durán-Domínguez-de-Bazúa, M.C. (1987). Effluents of the food industry in México: Environmental impacts on soil and water resources and possible solution using the biotechnological approach. Case problem: The corn industry. In: *Global Bioconversions*. Vol. II, Chap. 4. CRC Press, Boca Raton, FL, pp. 75–119.
8. Durán-Domínguez, M.C., Pedroza-Islas, R., Rosas-Vázquez, C., Luna-Pabello, V.M., Sánchez-Zamora, A., Capilla-Rivera, A., Paredes-Gómez, L., Valderrama-Herrera, S.B., and Vázquez-Cedeño, I. (1991). Producción de alimentos para peces: Utilización de subproductos del tratamiento de aguas residuales. In: *Premio Nacional Serfín El Medio Ambiente*. J.J. De Olloqui (Ed.). Futura Eds., Villa Nicolás Romero, Méx., Mexico, pp. 79–106.

BONDING OF TOXIC METAL IONS

E.N. Peleka
Kostas A. Matis
Aristotle University
Thessaloniki, Greece

T. Erwe
V. Mavrov
Saarland University
Saarbrücken, Germany

Worldwide environmental problems are becoming more and more acute. It is known that many specific industrial wastewater streams with large flows contain toxic metals in concentrations up to 500 mg/L, which have to be removed prior to water recycling, indirect discharge into a sewage system, or direct discharge into surface waters. Some examples are (1) rinse water in metal working enterprises, containing Cu, Ni, Zn, Sn, Cd, Pd, Ag, Al, Au, Cr, Mo, etc.; (2) rinse water in the semiconductor industry, containing Cu, Sn, Pb, Sb; (3) wastewater from the copper industry, containing As, Se; (4) other types of water in which the presence of toxic metals causes problems; (5) drinking water from ground water contaminated with As; and (6) mine water contaminated with different toxic metals (e.g., U, Ra, As).

When discharged directly into rivers, polluted wastewater poses a great risk to the aquatic ecosystem, and discharge into a sewage system negatively affects biosludge activity and leads to contamination of the sludge to be disposed of. As a result of the standards specified in the Water Resources Act, which was amended in 1986, industry takes precautions against these risks by treating dangerous components in a partial stream before mixing with other types of wastewater. So, research was focused, among others, on the development of highly selective bonding agents (often denoted as BA) with fast reaction kinetics for removing heavy metal ions (Fig. 1a). The latter may lead to better use of the capacity of bonding agents, resulting in smaller units and low residual concentrations of toxic metals in treated water streams, to comply with individual standards for water reuse or discharge—sorption kinetics constitutes another article in this *Encyclopedia of Water*.

In addition to zeolites (1,2), other typical materials extensively used are apatites (3,4) and various iron sorbents (5,6). These BA materials are broadly classified in three classes: carbon sorbents (active carbons, activated carbon fibers, molecular carbon sieves, fullerenes, heterofullerenes, nanomaterials), mineral sorbents (silica gel, activated alumina, oxides of metals, hydroxides of metals, zeolites, clay minerals, inorganic nanomaterials) and others (synthetic polymers, composite sorbents, mixed sorbents, etc.).

A review paper on bonding agents recently presented to a conference (7) also mentioned new types such as mesoporous materials and chelating resins and their potential applications to metal bonding from aqueous solutions. It also presented the equilibrium sorption capacity for metal ions of certain sorbents and stated that the functionalized silica materials have shown higher uptake capacity than that of other materials found in the literature. An inherent disadvantage of these materials is their low loading capacities and relatively small metal ion binding constants.

To circumvent these limitations, promising heavy metal sorbents have been prepared by coupling chelating ligands (e.g., thiol, amine, or crown ether functions) to support matrices consisting of inorganic oxides (e.g., silica, alumina, or clay) or organic polymers (e.g., polystyrene, cellulose, or polymethyl-methacrylate). Such functionalized materials have relatively high metal ion loading capacities and strong binding affinities for selected metal ions. This performance can be attributed to the presence of the surface-bound ligands, which can be specifically tuned to accommodate the selective sorption of targeted metal ions. Although superior in performance to conventional ion exchangers, functionalized matrices remain relatively inefficient because only a fraction of the immobilized ligands is accessible for metal complexation. The discovery of mesoporous metal oxide molecular sieves (by Mobil researchers) and organoceramic sorbents has given rise to new prospects for sorbent and catalyst design (7).

The knowledge of the governing chemical state of the metal ion to be removed is a prerequisite, so that any new development is based on a solid scientific

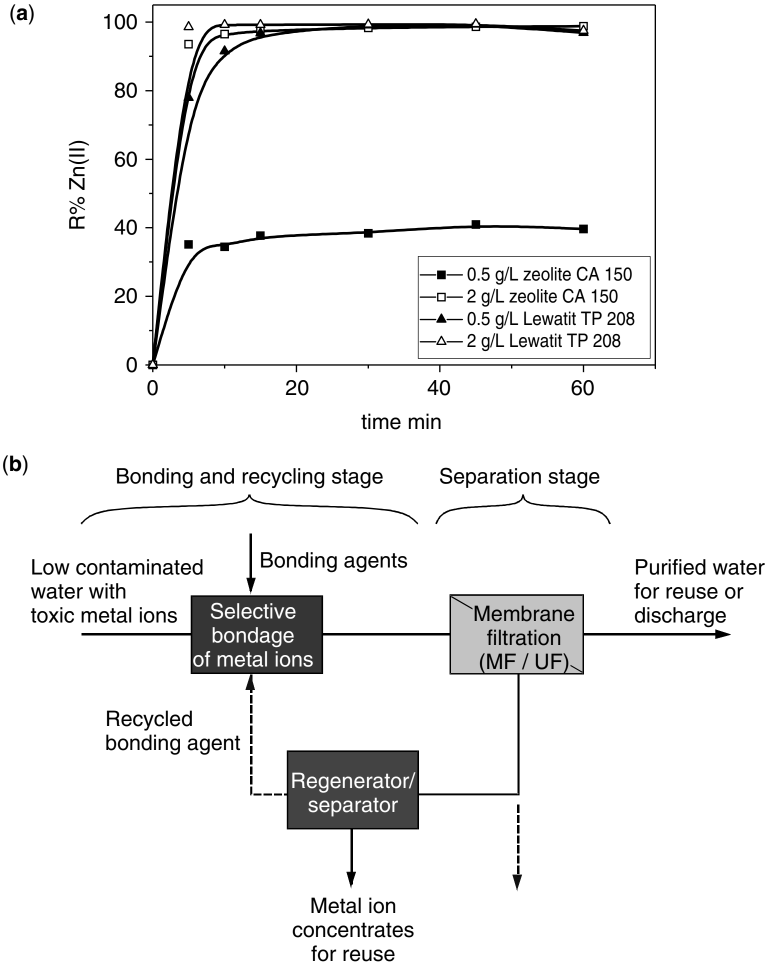

Figure 1. (**a**) Removal of zinc (50 mg/L, pH 6) by an ion exchange resin and zeolites for comparison (unpublished data). (**b**) Membrane separation of bonding agent/metal complexes for wastewater (METASEP project). Reprinted with gratitude from Reference 10.

foundation (8). The impact of chemical speciation is important in any process, including adsorption, ion exchange, or precipitation. These often act together, and the dominance of one specific is often hard to distinguish without careful chemical measurements and advanced analytical techniques. Identifying the bonding mechanism of metals on the agent/sorbent used is significant because, for example, the mechanism could dictate the mobilities of the species in their existing environment.

A new treatment process aimed at water and metal reuse was developed for the selective separation of heavy metals from wastewater to reduce costs and mitigate the environmental impact of wastewater pollutants (9). One of the variants, divided into three stages (shown as Fig. 1b), was the separation of BA-metal complexes by cross-flow microfiltration/ultrafiltration from wastewater that has metal concentrations less than 50 mg/L. This project was a contribution to the key action of European Union "Waste water treatment and re-use," in the 5th Framework programme.

The use of iron-based bonding materials, such as goethite and akaganėite, was successful because they are inexpensive, easily synthesized, suitable for both cation and anion sorption, and present low risks of adding a further pollutant to the system. Typical results are presented in Fig. 2; the sorbents had high surface area and defined pore size distribution. The synthesis of these materials involved hydrolyzing aqueous solutions of ferric salts followed by membrane purification and freeze drying (6). Sorption, in general, depends on the solution pH and its ionic strength. These macroscopic observations were used to differentiate between the kind of adsorption, physical or chemisorption (13). Other additional indications were obtained by examining the influence of temperature during equilibrium experiments and/or calculating (from the van't Hoff equation) the change of the respective enthalpy.

Table 1. Information about Industrial Effluents[a]

Wastewater from	Metal Targeted	Cationic	Anionic	Cleaning	Recovery
Open pit copper mine	Cu	X		X	X
Copper foil production	Sb, Cu	X	X	X	X
Groundwater	As (as AsF_6^-)		X	X	
Copper production	Se, As		X	X	

[a] Reference 10.

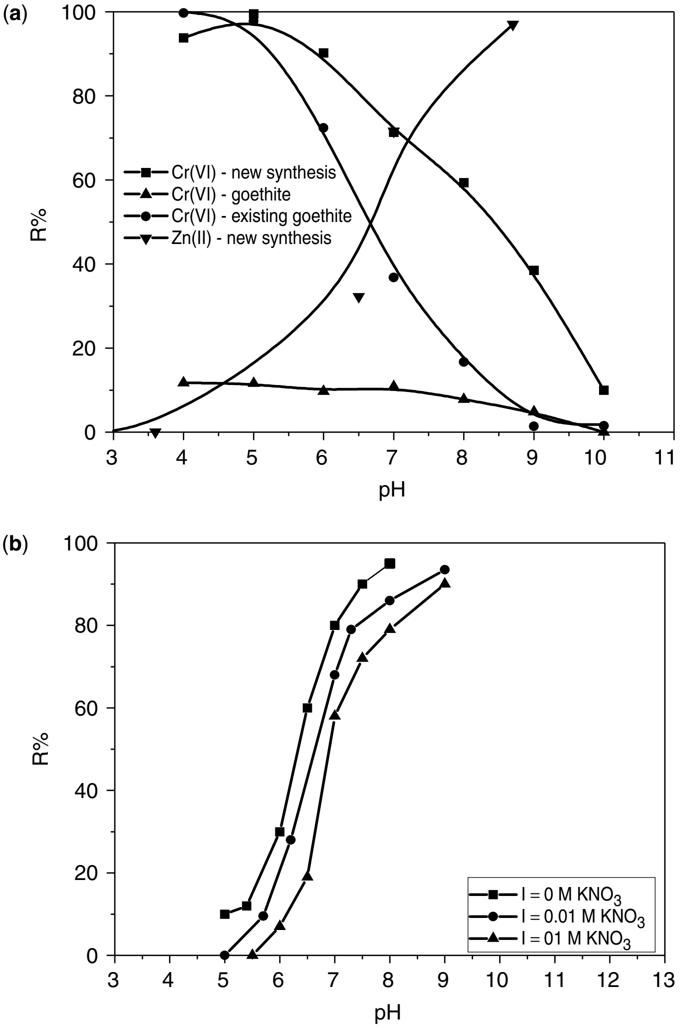

Figure 2. (**a**) Sorption of chromates (10 mg/L) and zinc ions (50 mg/L) by various goethite adsorbents, as a function of solution pH (1 g/L goethite, 2 h contact time). Reprinted with permission from Reference 11; copyright (2004) Scientific Journals. (**b**) Effect of pH on the adsorption of cadmium onto akaganéite for different values of ionic strength. Reprinted with permission from Reference 12; copyright (2004) Technical Faculty and Copper Institute, Bor Serbia.

The stability of a suitable bonding agent should be evaluated in a real process, for chemical and physical stability as well as losses in the proposed flowsheet. The selected industrial effluents in the aforementioned research project (e.g., METASEP) are presented in Table 1.

An effective mode of a sorption process is based on a fixed-bed contactor configuration (14). The sorption bed has to be porous to allow the liquid to flow through with minimum resistance but allow the maximum mass transfer into the particles—see also the article on Sorptive filtration in this *Encyclopedia*. Another possibility for this solid/liquid separation stage of loaded ultrafine material, as adsorbents and ion exchangers are, constitutes flocculation, as examined for goethite (Fig. 3a). The dynamics of flocculation were studied by a technique based on measuring the fluctuations of transmitted light in flowing dispersions.

An alternative for metals removal is *in situ* production, for instance, of ferric hydroxide (see Fig. 3b), which could be also conducted by electrocoagulation. Specifically, adsorbing colloid flotation is the process that involves removing a solute from an aqueous solution by adsorption on, coprecipitation with, or even occlusion in a carrier floc, often produced from ferric or aluminum salts by hydrolysis and solution pH modification. The aggregate, including the metals, is usually then floated after the addition of a suitable surfactant or collector (11).

The combination of flotation following sorption downstream (termed sorptive flotation) for metal cation removal by ion exchange materials, as the synthetic zeolites (of MAP type), gave promise for an improved metal separation process (16). A disadvantage of this is the application of surfactants, as flotation collectors; this represents an enormous demand, as flotation also applied in mineral processing, is currently met by numerous chemicals, mainly petroleum-based, hence, substances that are usually toxic to the environment and nonbiodegradable. For this reason, the successful application of biosurfactants was investigated (17).

Regeneration of bonding agents is, of course, a prerequisite often overlooked. The following two modes for different BA regeneration were envisaged: (1) The conventional, by changing the pH value; after separating it from the treated effluent, the loaded sorbent was mixed with a solution containing a strong acid (or alkal in the case of oxyanions) to shift the fixation equilibrium, to recover the metal in the acidic solution, and to produce a regenerated agent, where the protons have replaced the metal cations and could be recycled. (2) Electrodialysis with bipolar membranes for uncharged water-soluble bonding agents and modified electrodialysis for powdered agents (see Fig. 4). The loaded sorbent was fed into an electrodialysis reactor where, due to the electric field, the metals were separated from the material. The metal cations migrated in the electric field and crossed a cation-exchange membrane to enter into an other compartment, where they reacted with hydroxyl ions generated by a second bipolar membrane or by a cathode. The regenerated materials were recovered on one side, and a hydroxide sludge was recovered on the other side (18).

Elsewhere (19), another multistage process was proposed for removing aqueous cadmium consisting of, among others, elution at high concentrations of the sorbed metal ion from the biomass (the BA used here), recycling sorbent and eluant, wherever possible, and recovering cadmium electrochemically by a rotating cathode cell, in the form of metal powder as the end product. The flow diagram is shown with three cycles/routes for the combined process (investigated under an EU-funded research programme, too); the second-middle cycle represented flotation separation of the metal-loaded biomass. Biosorption, the uptake of heavy metals by dead biomass, has gained credibility during recent years, as it offers a technically feasible and economical approach. Several biological materials were investigated for heavy metals removal, including bacteria, yeasts, algae, and fungi.

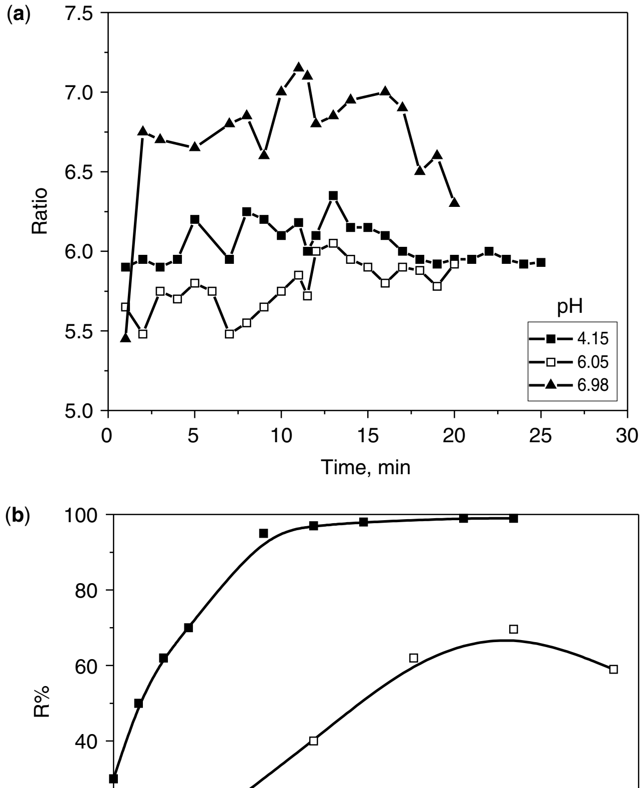

Figure 3. (**a**) Flocculation of synthetic goethite (0.5 g/L) fines by polyelectrolyte in the presence of As (III) (10 mg/L). Reprinted with permission from Reference 15; copyright (1999) Kluwer. (**b**) Removal of chromates and zinc as a function of Fe(III) concentration by in situ coprecipitation. Reprinted with permission from Reference 11; copyright (2004) Scientific Journals.

Concluding, the "*polluter pays*" principle ensures nowadays that companies have to pay the costs of environmental damage, or at least in part. Therefore, from their point of view, an economical and effective solution is required by developing techniques for improved treatment. On the other hand, many of the metal ions (cations or oxyanions) that are target compounds in effluents could be further concentrated for recycling; incineration of toxic sludge should be avoided, as this entails transferring pollution from one state to another. The strategy is to reuse water and metals, so "end-of-the-pipe" solutions will no longer be necessary. Attempts have been made to develop a more specific water treatment technique to provide industrial process water as close as possible to the source, applying quality assurance and efficiency control techniques for operational management. An innovative and cost-effective approach minimizes environmental impact from wastewater and prevents potential damage to health.

Acknowledgment

The financial help of the Energy, Environment and Sustainable Development programme of the European Union is gratefully acknowledged (acronym METASEP, contract no EVK1-CT-2000-00083).

BIBLIOGRAPHY

1. Colella, C. (1999). Environmental applications of natural zeolitic materials based on their ion exchange properties. In: *Natural Microporous Materials in Environmental Technology*. P. Misaelides, F. Macašek, T.J. Pinnavaia, and C. Colella (Eds.). Kluwer Academic, Dordrecht, The Netherlands, pp. 207–224.
2. Rees, L.V.C. (1997). Ion exchange in zeolites: Detergency and catalytic systems. In: *Progress in Ion Exchange—Advances and Applications*. A. Dyer, M.J. Hudson, and P.A. Williams (Eds.). RSC, pp. 393–402.
3. Ma, Q.Y., Traina, S.J., Logan, T.J., and Ryan, J.A. (1994). Effects of aqueous Al, Cd, Cu, Fe(II), Ni, and Zn on Pb immobilization by hydroxyapatite. *Environ. Sci. Technol.* **28**: 1219–1228.
4. Mandjiny, S., Zouboulis, A.I., and Matis, K.A. (1995). Removal of cadmium from dilute solutions by hydroxyapatite. Part I. Sorption studies. *Sep. Sci. Tech.* **30**: 2963–2978.
5. Lo, S., Jeng, H., and Lai, C. (1997). Characteristics and adsorption properties of iron-coated sand. *Water Sci. Technol.* **35**(7): 63–70.
6. Bakoyannakis, D.N. et al. (2003). Akaganeite and goethite-type nanocrystals: Synthesis and characterization. *Microporous Mesoporous Mater.* **59**: 35–42.
7. Asouhidou, D.D., Lazaridis, N.K., and Matis, K.A. (2004). Sorbent materials for metal ions removal from aqueous solutions: A review. *Int. Conf. Protection Restoration Environ.* VII, NTUA & Stevens Inst. Tech., Myconos, Greece.

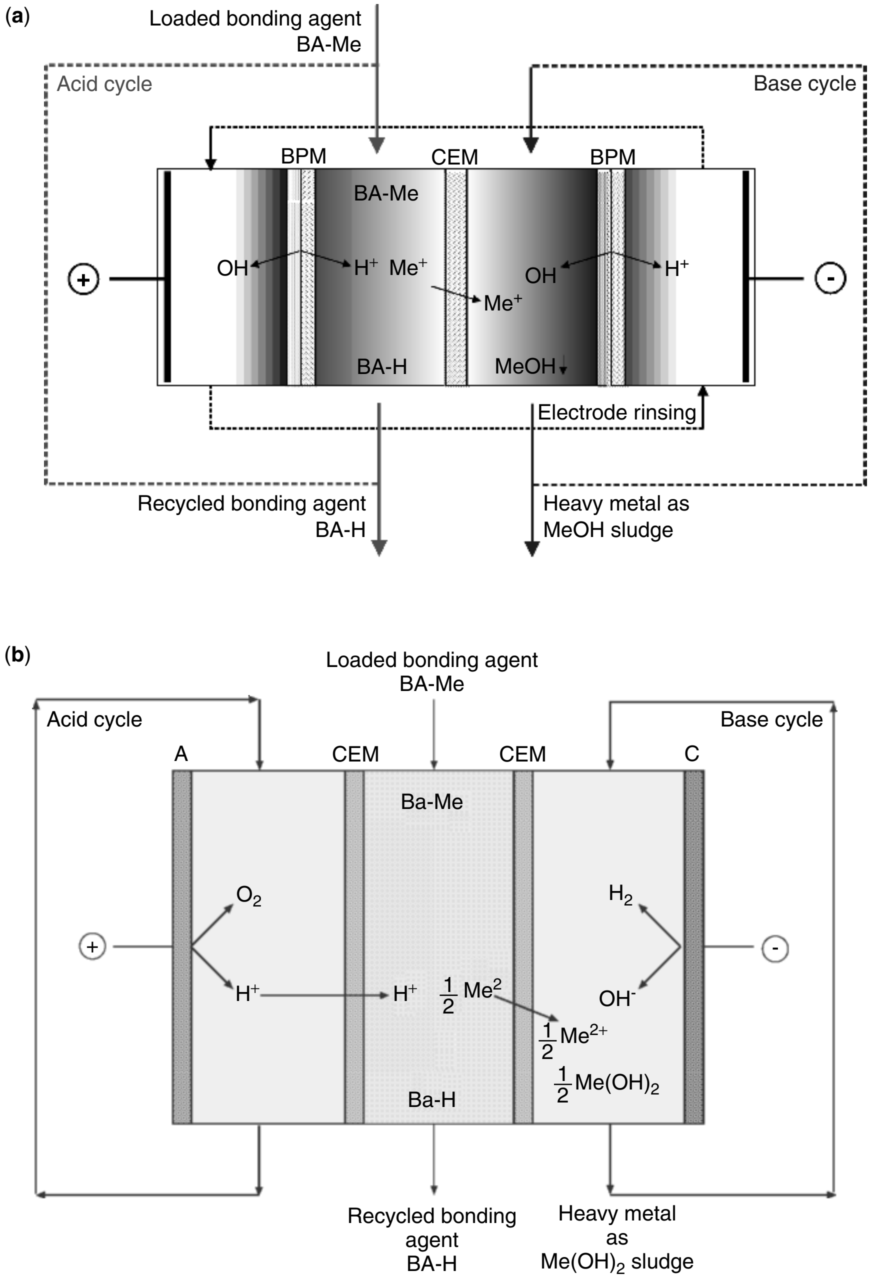

Figure 4. (**a**) Regeneration of water-soluble bonding agent by electrodialysis with bipolar membranes. (**b**) Regeneration of powdered bonding agent by modified electrodialysis. Reprinted with gratitude from Reference 10.

8. Evanko, C. and Dzombak, D. (1997). *Remediation of Metals-Contaminated Soils and Groundwater*. Ground-Water Remediation Technologies Analysis Center, Pittsburgh, PA.
9. Mavrov, V., Erwe, T., Blöcher, C., and Chmiel, H. (2003). Study of integrated processes combining adsorption, membrane separation and flotation for heavy metal removal from wastewater. *Desalination* **157**: 97–104.
10. EU FP 5 Project METASEP. (2004). Public Report. Brussels.
11. Deliyanni, E.A., Lazaridis, N.K., Peleka, E.N., and Matis, K.A. (2004). Metals removal from aqueous solution by iron-based bonding agents. *Environ. Sci. Pollut. Res.* **11**: 18–21.
12. Deliyanni, E.A., Bakoyannakis, D.N., Zouboulis, A.I., and Matis, K.A. (2004). Development and study of iron-based nanoadsorbents. *36th Int. Oct. Conf. Min. Metall.* Serbia-Montenegro.
13. Hayes, K.F., Papelis, Ch., and Leckie, J.O. (1988). Modelling ionic strength effects on anion adsorption at hydrous oxide/solution interfaces. *J. Coll. Interface Sci.* **125**: 717–726.
14. Lehmann, M., Zouboulis, A.I., and Matis, K.A. (2001). Modelling the sorption of metals from aqueous solutions on goethite fixed-beds. *Environ. Pollut.* **113**: 121–128.
15. Matis, K.A., Zamboulis, D., Zouboulis, A.I., and Lazaridis, N.K. (1999). Goethite mineral as a sorbent for heavy metals. In: *Natural Microporous Materials in Environmental Technology*. P. Misaelides, F. Macašek, T.J. Pinnavaia, and C. Colella (Eds.). Kluwer Academic, The Netherlands, pp. 425–433.
16. Matis, K.A. et al. (2004). Application of flotation for the separation of metal-loaded zeolites. *Chemosphere* **55**: 65–72.
17. Zouboulis, A.I., Matis, K.A., Lazaridis, N.K., and Golyshin, P.N. (2003). The use of biosurfactants in flotation: Application for the removal of metal ions. *Miner. Eng.* **16**: 1231–1236.

18. Schlichter, B., Mavrov, V., Erwe, T., and Chmiel, H. (2004). Regeneration of bonding agents loaded with heavy metals by electrodialysis with bipolar membranes. *J. Membrane Sci.* **232**: 99–105.
19. Butter, T.J. et al. (1998). The removal and recovery of cadmium from dilute aqueous solutions by biosorption and electrolysis at laboratory scale. *Water Res.* **32**: 400–406.

APPLICATION OF MICROFILTRATION TO INDUSTRIAL WASTEWATERS

T. Erwe
C. Blöcher
V. Mavrov
Saarland University
Saarbrücken, Germany

Kostas A. Matis
Aristotle University
Thessaloniki, Greece

The examination of membrane processes started mostly as a search for effective product separation that is crucial to economical operation in process industries. Certain types of materials are inherently difficult to separate, and their processing has become of increased importance in recent years, especially due to the growth of the biotechnological industry (1) This new technology using synthetic membranes for process separations has developed rapidly. Industrial membrane processes are classified according to the size range of materials, which they are to separate, and the driving force used in separation. The most common pressure-driven process is microfiltration (often denoted as MF) (2). This technique separate a size range of 10–0.1 μm; typical examples of materials separated are small particles, large colloids, and microbial cells.

Conventional processes for treating wastewater, precipitation, precipitation/reduction, ion exchange/sorption bed filtration, etc., often have disadvantages such as high use of treatment chemicals; large quantities of sludge produced that has to be treated and disposed of; inadequate selectivity; slow kinetics of the metals and their oxyanions and therefore, large water treatment units and high investment costs are required; low capacity of ion exchangers and sorbents; the residual metal concentration in the treated water streams can be high and water reuse or discharge is problematic; fouling and scaling problems, very low membrane permeability, and low water yield (up to 75%) are problems often encountered in conventional membrane processes; treatment processes are periodic and not continuous, followed by a regeneration step in which regeneration chemicals and rinsing water are employed, ultimately generating a new quantity of wastewater; treatment of wastewater streams with large volume flow and high concentrations is mostly not cost-effective. So, combinations of treatment processes are usually applied in practice.

Such a combination could, for instance, have as the first step the use of a suitable bonding agent followed by another separation technique downstream as a membrane process (3); see also the article on Bonding of toxic metal ions in this *Encyclopedia of Water*. Metal ion recovery by membrane technologies from aqueous waste streams was reviewed by several vendors in the field (4). At a lower particle size range than that for microfiltration, as aforementioned, other membrane processes are applied. Some examples of these applications involve, among others, ultrafiltration (UF) of fruit juices using tubular configuration polymeric and ceramic membranes (5) and reverse osmosis in treating sugar beet press water (6). Different polymeric membranes from commercial companies were tested for wine clarification (7), and the ceramic membranes were compared with polymeric ones for removing of natural organic matter from drinking water sources (8).

Ceramics are good materials for producing membrane filtration tubes (9). The advantages of ceramic membranes compared to polymer membranes include resistance to high temperatures up to 280 °C (in specially developed modules and systems up to approximately 700 °C); good corrosion resistance: resistance to organic solvents and a wide pH range; suitability for cleaning and steam sterilization; high mechanical strength: the possibility to back pulse, resulting in efficient removal of the dirt layer and the possibility of treating fluids with high viscosity; chemical inertness: a wide range of application possible in the chemical industry; long operational life; and high membrane flux from a composite structure. The disadvantages of ceramic membranes compared to polymeric membranes are brittleness and so they must be handled with care; the surface area/volume ratio is low so systems have larger dimensions; and the investment in ceramic membranes is high.

Microfiltration constitutes a physical barrier to solid particles (as bonding agents), but has limitations because of increased fouling at high concentrations. Conventionally, cross-flow processes are applied, where high shear rates on the membrane surface are created to prevent fouling of membranes. The corresponding high energy consumption led to the development of so-called low-energy systems (10), which are characterized by immersed membranes operated by vacuum or hydrostatic pressure. Fouling is often controlled by air blowing from below the membranes, creating shear forces near the membrane surface. The energy consumption of these systems is considerably lower than that of cross-flow systems, yet still significant (11). Furthermore, flux is also low, especially at higher concentrations of suspended solids in the reactor. Immersed membranes have been used for a variety of applications, mainly biological wastewater treatment and drinking water production (12).

New MF/UF membrane modules were developed (13) with emphasis on crossflow or submerged operation, optimized hydrodynamics (in contrast to commercial hollow fiber or spiral-wound modules) to work at higher bonding agent (BA) concentration in water, high fluxes in real process water, low fouling tendency, easy cleaning suited for backwashing, and low production costs. Both

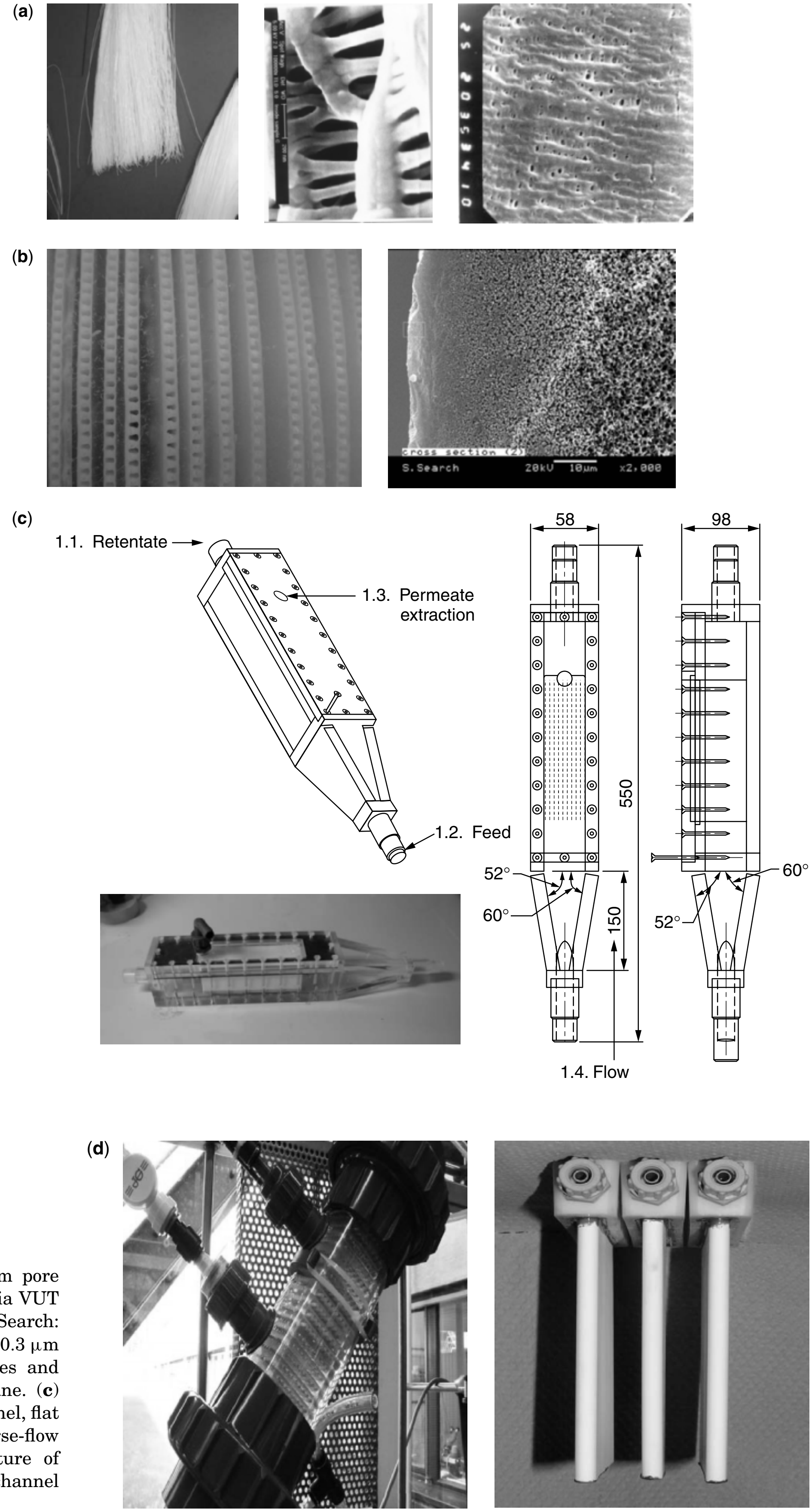

Figure 1. (**a**) PP membranes of 0.1 µm pore size produced by Eidos (and delivered via VUT Brno). (**b**) Membranes developed by S. Search: multichannel, flat sheet membranes (0.3 µm pore size), used for submerged modules and capillary outside skinned MF membrane. (**c**) Cross-flow module with PES, multichannel, flat sheet membranes. (**d**) Lab-scale transverse-flow membrane module by TNO, and picture of submerged module with ceramic multichannel flat sheet membranes.

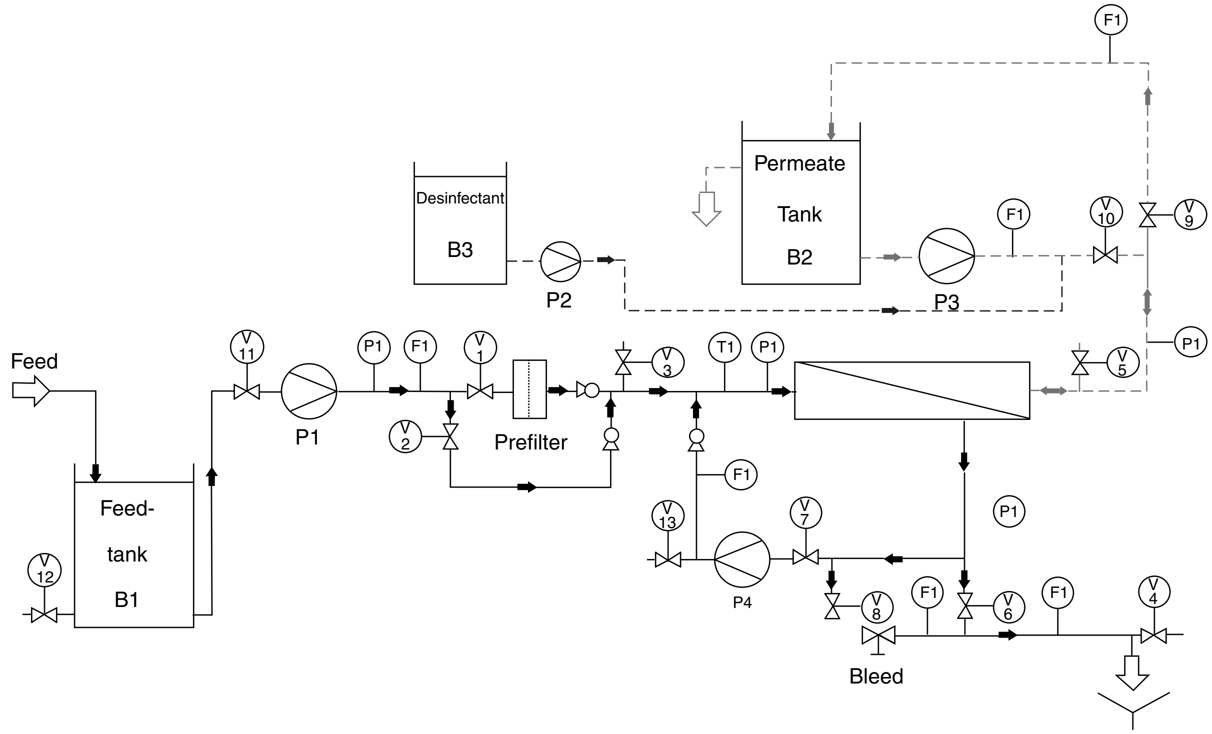

Figure 2. Process scheme of a bench-scale MF plant.

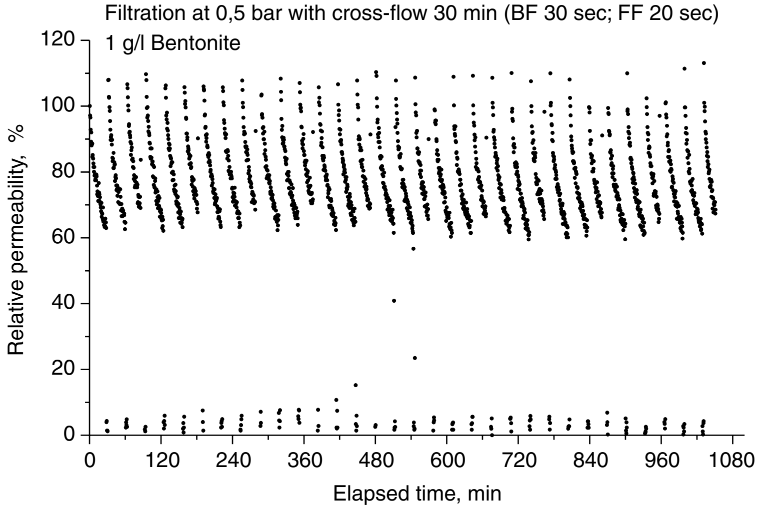

Figure 3. Filtration of bentonite suspension with PP membrane module in back-flush mode.

polypropylene (PP) as well as polyethersulphone (PES) based membranes were developed (see Fig. 1). PP fibers have a symmetrical wall, low flux and are inexpensive. PES capillaries have an asymmetrical wall, and higher flux, but are more expensive. Because of this trade-off, both membranes are of much interest in research. These PP fibers were used in common longitudinal membrane modules, as shown in Fig. 1a–d with bench-scale modules. Figure 1c shows a cross-flow module with PES multichannel, flat sheet membranes.

The PES capillary membranes were used for the transversal flow membrane module (TFM). Previous studies showed that within TFM modules, membranes tend to break because of the high stress at the potting-membrane interface. Much improved capillary membranes were developed. They possess three times the tensile strength at break (13). This was the result of adding a component to the extrusion mixture that resulted in a tighter sponge in the support layer. Figure 1d, left, shows the TFM module consisting of 10 active layers of 8 membranes next to each other in a grid.

The process scheme of the bench-scale plant used is presented in Fig. 2, based on feed-and-bleed processing; automatic back-flush was incorporated. Experiments with bentonite and the PP membrane modules in longitudinal outside mode showed that relative permeability decreased,

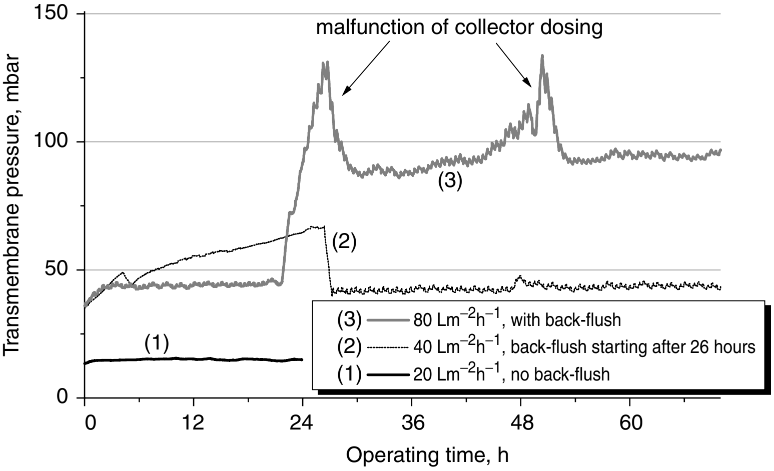

Figure 4. Transmembrane pressure as a function of time for the hybrid process at various operating modes, under the following conditions: $c(\text{zeolites}) = 8$ g/L, $c(\text{CTABr}) = 20$ mg/g, $u_g = 0.33$ cm/s, back-flushing (if applied) every 30 min for 10 s.

when the transmembrane pressure was increased. Somewhat surprising, but very promising was that the back-flush mode has been very effective. At least during 22 hours, permeability was constant (Fig. 3). When selecting membranes for efficiency and rejection of substances, permeability and pore size or molecular weight cutoff were examined as important criteria (14).

Today, there is a tendency for combined and more compact processes as, for example, membrane bioreactors (15). Membrane fouling was also said to be the main limitation to the faster development of the latter. A newly developed type that has the further advantage of energy saving is the submerged membrane bioreactor.

A large number of techniques have been used to limit membrane fouling, such as cross-flow liquid velocity, the use of baffles, back-flushing, transmembrane pressure (TMP) pulsing, and air sparging (16). Depending on the membrane process, the function of air can be different: a gas back-flush detaches and carries away deposited particles, the formation of a particle or concentration polarization may be prevented or limited, and even a compound may be transferred by gas to a liquid phase.

Air bubbling, under appropriate conditions, could also be used as the transport for flotation, which is another efficient separation method applied in water treatment (17). The aim of this innovative idea was integrating both processes in the same tank, taking advantage of dispersed-air flotation combined with microfiltration by submerged membranes. For this, ceramic flat-sheet membrane modules with multichannel geometry were tested; the pore size was 0.3 μm, pure water flux 3000 $\text{L m}^{-2}\ \text{h}^{-1}\ \text{bar}^{-1}$, the material of permeate collector polystyrene, and sealing material epoxy. The membrane distance was 20 mm, geometry 100 × 100 and area 0.054 m^2 (Fig. 1d, right). The hybrid cell was tested on-site at the Assarel-Medet open pit copper mine effluent, near Panagyurishte, Bulgaria (18). The flotation separation techniques employed and investigated were based on three different mechanisms: (1) ion flotation using xanthates, (2) precipitate flotation generating copper hydroxide, and (3) sorptive flotation using zeolites as sorbent material—see also the article on Flotation in this *Encyclopedia*.

A parallel economic study showed that the capital investment for the hybrid unit would be of the order of 0.92 M€. The operating cost of the plant is about 1.42 M€ per year. Comparison with separate systems (MF following flotation) showed that the hybrid system has a lower capital cost and lower operating costs. The former also requires larger membranes.

The advantageous influence of flotation may be apparent from Fig. 4. Membranes were operated at constant flux in all experiments. For an initial zeolite concentration of 8 g/L the membranes could be operated at 20 $\text{L/m}^2\text{h}$ without back-flushing. Resulting transmembrane pressure was as low as 15 mbar. However, to achieve stable membrane performance at higher fluxes, back-flushing was necessary. A cycle of 10 s back-flushing every 30 min was sufficient, resulting in TMP of 45 and 90 mbar (the mean of two cycles between backwashing) for imposed fluxes of 40 and 80 $\text{L/m}^2\text{h}$, respectively. The relatively long time between back-flushing resulted in a permeate loss of less than 5% due to backwashing (19).

In Fig. 4, it can also be seen that the increase in transmembrane pressure when operating the membranes at 40 $\text{L/m}^2\text{h}$ without back-flushing was fully reversible. Switching the operating mode of a membrane, that had been operated in continuous permeate withdrawal for 24 hours, to back-flushing resulted after less than 1 hour in the same transmembrane pressure as a membrane, which had been operated with back-flushing from the beginning. This proves that the loss in membrane performance was attributed to an easily removable cake layer, which was also visually observed. The resistance-in-series model (20) may be used in this case, which describes filtration resistance as the fraction of TMP to flux; the resistance of the cake layer can be evaluated according to

$$J = \frac{\text{TMP}}{\eta(R_\text{M} + R_\text{C})} 3.6 \times 10^8, \tag{1}$$

where J is permeate flux, R_M is membrane resistance, R_C is the resistance of the cake layer or, in general, fouling resistance; and η the viscosity of water to be filtered. Cake resistance is of the same magnitude as membrane resistance and thus, quite low, which can be attributed to a highly permeable cake layer that zeolites form. These results could be also considered in terms of the concept of "critical flux" (the flux up to which the relation between flux and TMP is linear), which is currently often applied to microfiltration processes (21). Although this concept does not imply back-flushes, from a process point of view, it can also be concluded that the membranes in our process were operated in the subcritical flux region. In all investigations, the air supply was determined by that required for flotation. No additional aeration of the membranes was carried out. This has to be considered when relating membrane performance to conventional immersed membrane systems because the superficial air velocity used in the hybrid cell was about 10 times lower than that usually applied in these systems.

Concluding, one of the main objectives of environmental technology today is to design and implement processes to clean effluents from industrial production; recycling valuable materials and microfiltration has certainly a definite role having presented very promising results. Product recovery and waste recycling have been addressed by providing novel integrated processes for reusing water and the possibility of metal recycling, aiming to meet regulatory requirements with environmental benefits, not merely by reducing risks.

Acknowledgments
The financial help of EU is gratefully acknowledged (contract no. EVK1-CT-2000-00083, acronym METASEP); all figures in this article are from the Public Report (13) of this project. Many thanks are due to our partners for their help and collaboration.

BIBLIOGRAPHY

1. Coulson, J.M., Richardson, J.F., Backhurst, J.R., and Harker, J.H. (1991). *Chemical Engineering*. 4th Edn. Vol. 2. Pergamon, Oxford, UK.
2. Porter, M.C. (1986). Microfiltration. In: *Synthetic Membranes: Science, Engineering and Applications*. P.M. Bungay, H.K. Lonsdale, and M.N. de Pinho (Eds.). D. Reidel, Dordrecht, pp. 225–247.
3. Mavrov, V., Erwe, T., Blöcher, C., and Chmiel, H. (2003). Study of new integrated processes combining adsorption, membrane separation and flotation for heavy metals removal from wastewater. *Desalination* **157**: 97–104.
4. Radecki, P.P., Crittenden, J.C., Shonnard, D.R., and Bulloch, J.L. (Eds.). (1999). *Emerging Separation and Separative Reaction Technologies for Process Waste Reduction—Adsorption and Membrane Systems*. AIChE, Center for Waste Reduction Technologies, New York, pp. 199–219.
5. Capannelli, G., et al. (1994). The use of membrane processes in the clarification of orange and lemon juices. *J. Food Eng.* **21**: 473–483.
6. Hinkova, A., Bubnik, Z., Kadlec, P., and Pridal, J. (2002). Potentials of separation membranes in the sugar industry. *Sep. Purific. Tech.* **26**: 101–110.
7. Urkiaga, A., Fuentes, de las L., Acilu, M., and Uriarte, J. (2002). Membrane comparison for wine clarification by microfiltration. *Desalination* **148**: 115–120.
8. Lee, S. and Cho, J. (2004) Comparison of ceramic and polymeric membranes for natural organic matter (NOM) removal. *Desalination* **160**: 223–232.
9. Scott, K. (1997). *Handbook of Industrial Membranes*. Elsevier, Oxford, UK, p. 413.
10. Côté, P.L., Smith, B.M., Deutschmann, A.A., Rodrigues, C.F.F., and Pedersen, S.C. (1994). Frameless array of hollow fibre membranes and method of maintaining clean fiber surfaces while filtering a substrate to withdraw a permeate. World Patent WO 94/11094, Zenon Environmental Inc.
11. Guibert, D., Ben Aim, R., Rabie, H., and Côté, P.L. (2002). Aeration performance of immersed hollow-fibre membranes in a bentonite suspension. *Desalination* **148**: 395–400.
12. Côté, P.L., Mourato, D., Güngerich, C., Russell, J., and Houghton, E. (1998). Immersed membrane filtration for the production of drinking water: Case studies. *Desalination* **117**: 181–188.
13. EU FP 5 Project METASEP. (2004). Public Report, Brussels.
14. Schlichter, B., Mavrov, V., and Chmiel, H. (2000). Comparative characterization of different commercial UF membranes for drinking water production. *J. Water Supply: Res. Tech.-AQUA* **49**: 321–328.
15. Blöcher, C. et al. (2002). Recycling of spent process water in the food industry by an integrated process of biological treatment and membrane separation. *Desalination* **144**: 143–150.
16. Matis, K.A. et al. (2004). Air sparging during the solid/liquid separation by microfiltration: Application of flotation. *Sep. Puri. Tech.*. **40**: 1–7.
17. Lazaridis, N.K., Blöcher, C., Dorda, J., and Matis, K.A. (2004). A hybrid MF process based on flotation. *J. Membrane Sci.* **228**: 83–88.
18. Lazaridis, N.K., Peleka, E.N., Karapantsios, Th.D., and Matis, K.A. (2004). Copper recovery from effluents by various separation techniques. *Hydrometallurgy* **74**: 149–156.
19. Blöcher, C. et al. (2003). Hybrid flotation-membrane filtration process for the removal of heavy metal ions from wastewater. *Water Res.* **37**: 4018–4026.
20. Cheryan, M. (1998). *Ultrafiltration and Microfiltration Handbook*. Technomic, Lancaster, PA.
21. Cho, B.D. and Fane, A.G. (2002). Fouling transients in nominally subcritical flux operation of a membrane bioreactor. *J. Membrane Sci.* **209**: 391–403.

WATER TREATMENT IN SPENT NUCLEAR FUEL STORAGE

Samuel C. Ashworth
Idaho National Engineering and
Environmental Laboratory
Idaho Falls, Idaho

Many facilities are changing their storage to dry mode, but in 1996, there were 100,000 tons of Zircaloy-clad spent nuclear fuel (SNF) in wet storage (1). Much of this fuel is stored in extended mode, beyond the period of time originally planned. Thus, this extended storage can create problems.

A typical spent nuclear fuel pool storage system is shown in Fig. 1. The main functions of water storage pools are to contain the fuel and any radionuclides that leak and to remove heat generated by radionuclide decay. Water is withdrawn from pools via overflow sumps and pumped through a filter to remove solids including microorganisms. The water is treated by ion exchange

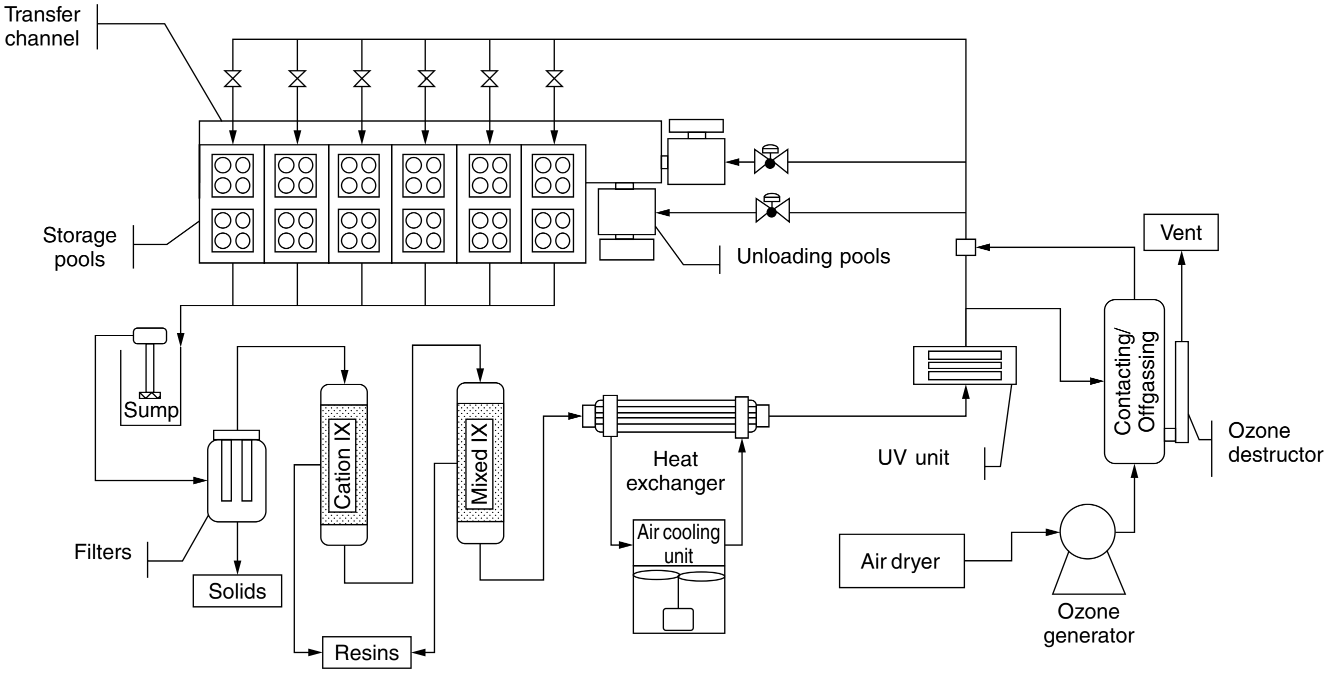

Figure 1. Typical spent fuel storage schematic.

to remove radionuclides and other dissolved substances. Heat exchangers remove the decay heat that the spent fuel transfers to the water. Some systems have ultraviolet (UV) light systems that kill microorganisms. An ozonator is also shown in Fig. 1, even though many systems do not have this. The supply makeup water normally has very low conductivity ($\kappa \leq 1$ μS/cm), so there is very little ion exchange capacity used up unless the fuel is in very bad shape and is leaking. There is also a need to provide basin chemistry/corrosion control to minimize corrosion of SNF, parts, and equipment. The following are typical limits imposed to minimize corrosion: $5.0 < \text{pH} < 8.0$, exclusion of heavy metals or their salts including Cu, Pb, Sn, Hg, Ni, and Co, keeping $\text{Cl}^- < 10$ ppm, and limiting conductivity <10 μmho/cm (μS/cm). There may be a potential for microbiologically induced corrosion (MIC) in storage pools, so it is important to control microorganisms. This article focuses on microbiological control that presents the biggest challenge facing pool storage systems. The corrosion of fuel element cladding must be minimized to ensure containment of radionuclides and worker protection. The general outline is a discussion of mechanisms, rationale for treatment, and treatment methods for storage pools.

MICROBIOLOGICAL MECHANISMS

Some of the variables that affect the growth of microorganisms include food/nutrient availability, temperature, pH, biocide concentrations, radiation, makeup water microorganism concentration, and predation. There are practical limitations to the types of activities that can be used to control microorganisms, for example, restricting the use of biocides or pH out of the range of water chemistry requirements.

The first substances on a surface during biofilm formation are trace organics (2). Almost immediately after the clean pipe surface comes into contact with water, an organic layer deposits on the water/solid interface (3). These organics form a conditioning layer that neutralizes excessive surface charge and surface free energy that may prevent a bacterial cell from approaching near enough to initiate attachment. In addition, the absorbed organic molecules often serve as a nutrient source for bacteria.

Some planktonic (free-floating) bacteria approach the pipe or equipment wall and become entrained within the boundary layer. Some of these cells strike and adsorb to the surface for some finite time and then desorb. This initial attachment is based on electrostatic attraction and physical forces. Some of the reversibly adsorbed cells begin forming structures that may permanently bond the cell to the surface. These cells become irreversibly adsorbed. Figure 2 provides a schematic of this stage of biofilm development.

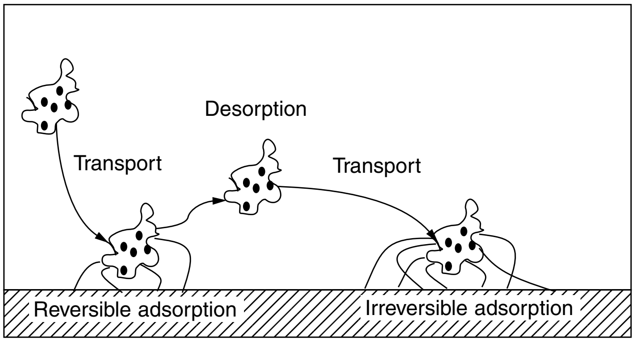

Figure 2. Biofilm development.

Biofilm bacteria excrete extracellular, polymeric substances that contain the biofilm and attach it to a wall. These polymer strands trap scarce nutrients and protect bacteria from biocides. Attachment is mediated by extracellular polymers that extend outward from the bacterial cell wall (3). The polymeric material, or glycocalyx, consists of charged and neutral polysaccharide groups that facilitate attachment and also act as an ion-exchange system for trapping and concentrating trace nutrients from the overlying water. As nutrients accumulate, the pioneer cells proceed to reproduce. The daughter cells then produce their own glycocalyx, greatly increasing the volume of ion-exchange surface. A viable colony of bacteria is thereby established (4).

A biofilm-covered surface is gelatinous and slippery because the glycocalyx matrix holds water (75–95%) (5). The glycocalyx net also absorbs other types of microbial cells through physical restraint and electrostatic interaction. These secondary colonizers metabolize wastes from the primary colonizers and produce their own waste that other cells then use in turn. Bacteria and other microorganisms develop cooperative colonies within the biofilm. An anaerobic biofilm may develop underneath the aerobic layer. The biofilm thickness will reach equilibrium as flowing water detaches cells extending out into turbulent flow.

Different species live in the films, helping each other to exploit food supplies and to resist biocides through synergistic interactions. Some species use toxic wastes produced by another species as food. By pooling their biochemical resources to build a colony, several species of bacteria, each with different enzymes, can break down food supplies that no single species could digest alone. The biofilms are permeated at all levels by a network of channels through which water, bacterial wastes, nutrients, enzymes, metabolites, and oxygen travel. Gradients of chemicals and ions between microzones provide the power to shunt substances around the biofilm (6).

A biofilm can spread at its own rate by ordinary cell division, and it also periodically releases new pioneer cells to colonize downstream sections of piping and equipment. As the film grows to a thickness that allows it to extend through the boundary layer into zones of greater velocity, some cells are sloughed off. These later pioneer cells more easily adsorb than their upstream predecessors because the parent film releases wastes into the stream that may serve as either the initial organic coating for uncolonized pipe sections downstream or as nutrient substances for other cell types (4).

The development of a mature biofilm may take several hours to several weeks, depending on the system (3). *Pseudomonas aeruginosa* is a common pioneering bacteria and is used frequently in biofilm research. In one experiment, researchers found that *Pseudomonas* cells adhere to stainless steel, even to electropolished surfaces, within 30 seconds of exposure (7). The association of bacteria with a surface and the development of a biofilm can be viewed as a survival mechanism. Bacteria benefit by acquiring nutrients and protection from biocides.

Potable water systems, especially high-purity water systems, are nutrient-limited environments, but even nutrient concentrations too low to measure are sufficient to permit microbial growth and reproduction. Bacteria and other organisms capable of growth in nutrient-limited environments are called oligotrophs. Bacteria have evolved the means to find and attach to surfaces to increase the chances of encountering nutrients.

In a study on the attachment of *Pseudomonas* to stainless steel surfaces (8), researchers put cells in a blender to remove the flagella. They found that the rate of cell attachment decreased at least 90% when flagella were removed.

Many organisms faced with the starvation conditions encountered in purified water systems respond by altering their cell wall structure to increase their affinity for surfaces. By altering the protein and lipid composition of the outer membrane, the charge and hydrophobicity can be changed. The cell wall becomes hydrophobic. Once in the boundary layer, they are attracted to a pipe surface (4).

Recent research has dispelled some earlier assumptions about bacteria and biofilms. In the past, microbiologists assumed that biofilms contained disorderly clumps of bacteria located in no particular structure or pattern. In most cases, the base of the biofilm is a bed of dense, opaque slime 5 to 10 μm thick. It is a sticky mix of polysaccharides, other polymeric substances, and water, all produced by the bacteria. Water currents that flow in channels between the colonies carry nutrients and waste (9).

Past researchers assumed that biofilm bacteria behaved much like solitary, free-floating microorganisms. Biofilm bacteria have exactly the same genetic makeup as their free-roving relatives, but their biochemistry is very different because they activate a different set of genes. The instant the bacteria dock to glass, they switch on certain genes involved in the synthesis of alginate (an unusually sticky form of slime), switching them off again once the bacteria are engulfed in alginate. It is now estimated that as many as 30 to 40% of the proteins in bacterial cell walls differ between sessile and planktonic bacteria. Some of the targets for biocides are not there any more, so the bacteria become difficult to kill.

Researchers are studying the chemicals (called sigma factors) which signal bacteria to change their biochemistry to life in a biofilm (9). If they can discover a reverse sigma factor which would change biofilm bacteria into planktonic free-floaters, it might be possible to dissolve biofilms by sending the equivalent of an evacuation signal (6).

The surface material has little or no effect on biofilm development. Stainless steel is just as susceptible as plastic pipe. Piping material to which microorganisms cannot adhere has yet to be discovered. Studies have shown that microbes adhere to stainless steel, Teflon, PVC, and PVDF (Kynar) nearly equally (4). Incorporating an antimicrobial additive into plastic to delay or reduce the attachment of biofilm is possible (10). However, the chemicals used are not normally safe for drinking water supplies. Some ion exchange resins have been silver-coated to prevent microbial growth. However, silver-tolerant bacterial populations can develop (11). There are no practical examples of toxic surface coatings for drinking water piping.

One major factor influencing biofilm development in purified-water systems is surface area. Industrial water systems, unlike most natural environments (lakes and rivers), offer a tremendous amount of surface area for attachment. RO membranes, ion exchange resins, storage tanks, cartridge filters, and piping systems all provide surfaces suitable for bacterial attachment and growth (3).

Although smoother surfaces delay the initial buildup of attached bacteria, smoothness does not significantly affect the total amount of biofilm that attaches to a surface. No surfaces have been found that are exempt from biofouling (12). Surface structure influences the rate of fouling, but only initially during the first few hours of exposure. In general, smooth surfaces foul at a slower initial rate than rough ones, but biofilm formation after a period of days is inevitable.

High water flow rates may alter biofilm growth but do not prevent the attachment of bacteria to pipe surfaces. Work by Mittelman (3), Patterson (13), and Meltzer (12) supports this conclusion. High flow does not prevent bacterial attachment or remove an existing biofilm for the following reasons:

- Low flow in the boundary layer. The velocity is zero at the wall of a pipe or other surface. Near the wall, a boundary layer exists that is controlled by viscous effects, whereas outside the boundary layer, the flow is not determined by viscous effects. Within this boundary layer, for both laminar and turbulent flow, particles and microorganisms can transport and attach to surfaces. Regardless of the water velocity, the velocity profiles are such that it flows most slowly in the layers adjacent to pipe surfaces. The laminar sublayer thickness was calculated (14) for various flow velocities and for five pipe sizes. The shear forces within the laminar sublayer are much less than the forces produced in biofilm attachment.
- Strong adhesion by exopolymers. In water systems where high-velocity flow is continuous, the bacteria that accumulate in a biofilm tend to be filamentous varieties (such as *Pseudomonas*) especially suited for attachment by filaments.
- Although high flow velocity does not prevent the attachment of bacteria to pipe surfaces, it does have the following effects on the biofilm structure:
  - Denser biofilm. At higher flow rates, a denser, somewhat more tenacious biofilm is formed (3).
  - Limited biofilm thickness. The maximum thickness of a biofilm can be considered the thickness of the laminar sublayer (see Table 1).
  - Limited nutrients. Like other living creatures, bacteria require certain nutrients for growth and reproduction. Limiting these nutrients limits, bacterial growth, but nutrient levels in high-purity systems are unequivocally sufficient to permit microbial growth and reproduction to a troublesome extent (15). Table 1 lists some sources of nutrients in purified water systems.

Bacteria can get some of their nutrients from the pipe and fittings in a water system (11). Most plastics are not biodegradable, but pipe cements and plasticizers that leach from epoxy resins, PVC pipe, and polyamide pipe can be organic carbon sources for bacteria. Cellulose-based RO membranes can also be a nutrient source. That is why RO feedwater needs chlorination. Bacteria can also obtain trace metal nutrients from stainless steel and other metal components.

Table 1. Nutrients for Bacterial Growth Found in Pure Water Systems[a]

Nutrient	Sources
Organic carbon	Humic and fulvic acids (source water), pipe plasticizers and solvents, fiberglass-reinforced plastics (FRPs), pump and gauge lubricants, microbial by-products, personnel, airborne dust
Nitrogen	Humic and fulvic acids (source water), nitrates and nitrites (source water), microbial by-products, airborne dust
Phosphorus	Phosphates (source water), microbial by-products, airborne dust
Sulfur	Sulfates (source waters), sulfuric acid (RO pretreatment), membrane surfactants, airborne dust
Trace metals and salts	Source waters, process piping, fiberglass-reinforced plastics (FRPs), stainless steel system components, RO pretreatment, chemicals, personnel, airborne dust

[a] Reference 3.

Currently available technology cannot reduce nutrient levels completely; so total control of bacteria is not achievable by simply controlling nutrients. Similarly, very small quantities of oxygen will adequately support bacterial growth, even if the bacteria do not revert to anaerobic respiration, which most bacteria can do. Therefore, a thriving bacterial population can exist, even in high-purity water systems (14). Nutrient-limiting environments can actually promote the attachment of bacteria to surfaces because that is where the trace organics accumulate and extracellular polymers in a biofilm capture trace nutrients. Although bacteria cannot be completely starved out, nutrient-poor reverse osmosis water will support less biofilm than regular tap water supplies.

Water samples collect only planktonic or free-floating bacteria. Free-floating bacteria in animal drinking water are either sloughed off the biofilm or pass through from the incoming water supply. If a plate count test is low, one should not assume that bacteria are not present in the watering system. More than 99% of the bacteria in water systems are in biofilms attached to pipe and equipment surfaces. A mature biofilm may not slough off many cells into the drinking water if the integrity has not been disrupted (e.g., by recent flushing or sanitization). However, it is still there.

TREATMENT RATIONALE

Some of the main reasons to control microorganisms include: water clarity/turbidity, heat transfer surfaces and other equipment problems, and microbiologically induced corrosion.

Water Clarity/Turbidity

Very clear water is required to see the bottom of a pool (≈10 m). Therefore, both algal and sedimentary, particles must be removed so that required clarity can be maintained. Turbidity refers to how clear the water is. The greater the amount of total suspended solids (TSS) in the water, the murkier it appears, and the higher the measured turbidity. Algal turbidity varies seasonally and with depth in a complex manner in response to physical, chemical, and biological changes. Figure 3 shows the effect of algae concentration on turbidity. A comparison with sediments is shown in Fig. 4. Turbidity is reported in nephelometric units (NTUs) referring to the type of instrument (turbidimeter or nephelometer) used for estimating light scattering from suspended particulate material. The nephelometer, also called a turbidimeter, has a photocell set at 90° to the direction of the light beam to estimate scattered rather than absorbed light. This measurement generally provides a very good correlation with the concentration of particles in the water that affect clarity.

Another inexpensive device is called a turbidity tube. It involves looking down a tube at a black and white disk (Secchi disk) and recording the depth of water that makes the disk disappear. Secchi disk depth, then, is a measure of water clarity. There have not been many efforts to correlate distance clarity and turbidity. However, a rough correlation is provided in Fig. 5. Turbidity units (NTU and JTU) are interchangeable units. They differ only in that their name reflects the device used to measure turbidity. One JTU is equivalent to 1 ppm of silica.

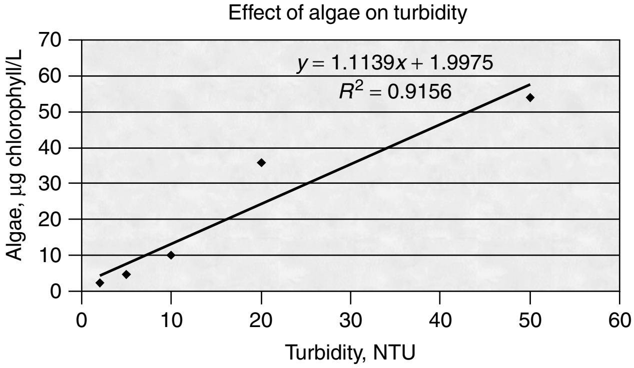

Figure 3. Turbidity vs. algae.

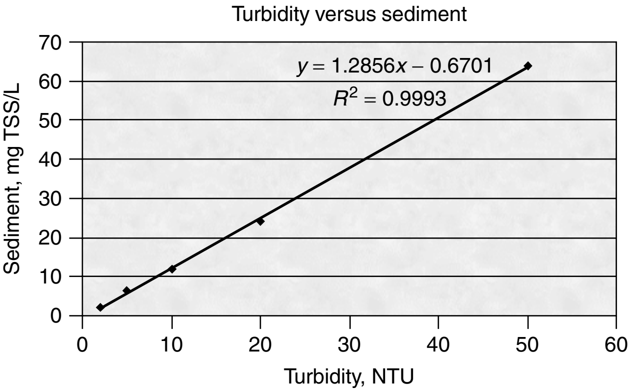

Figure 4. Turbidity vs. sediment.

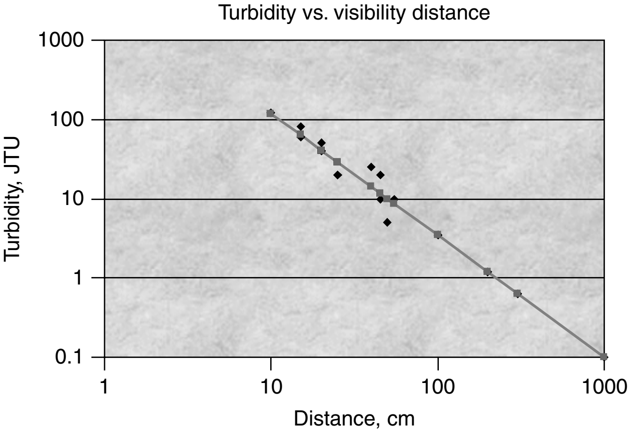

Figure 5. Turbidity vs. distance.

Heat Transfer/Equipment Problems

The waterside heat transfer systems of a treatment loop heat exchanger are susceptible to a biofilm that can reduce the overall heat transfer coefficient and hence efficiency. Microorganisms can cause biofilms to coat piping, valves, pumps, and other equipment. Further, biofilms can occur on ion exchange resins, filters, in-basin structures, and spent nuclear fuel. Minimizing microorganisms and biofilms to help prevent this is highly desirable. In general, water from natural sources leads to poor cooling performance unless adequate steps are taken (16). The extent of the problem, however, depends on the quality of the raw makeup water. Deposits on heat exchanger surfaces are complex, and are likely to include particulate matter, crystalline salts, corrosion products, and biofilms. It is necessary to counteract this problem of deposition onto surfaces to maintain heat exchanger effectiveness.

When using natural waters, the presence of microorganisms is probably the predominant cause of poor heat transfer efficiency. The conditions on the waterside of a cooler are usually conducive to the growth of microorganisms, notably bacteria, because of suitable temperature conditions and the availability of nutrients. The colonization of a heat transfer surface and the subsequent growth of a biofilm may, under favorable conditions, lead to a deposit 1 mm thick or more, unless preventive action is taken. This layer represents resistance to heat transfer: it acts as an insulating layer. The biofilm may also act as a trap for particulate matter and corrosion products. In addition, the biofilm may encourage corrosion of the underlying metal due to biological activity, thereby adding to operating and maintenance costs.

The shape and structural arrangement of a biofilm growing in a flowing fluid influences the mass transfer characteristics of the biofilm system as well as the drag force exerted on individual biofilm structures (17). If the biofilm is highly compliant, the shape varies through the growth cycle of the biofilm and also due to variations in fluid shear stress (18). Changes in biofilm shape affect its porosity and density, therefore affecting the transfer of solutes into and through the biofilm. Fluctuations in biofilm shape also affect the hydrodynamic drag that in turn influences the detachment rate and pressure losses in a flowing system. In addition, it is thought that biofilm viscoelasticity may explain the large pressure drops observed in biofilm-fouled pipes (19). However, very few studies have been conducted on the material properties of intact biofilms, in part, due to the technical difficulty of such testing. The small dimensions and pliability of biofilms makes sample handling extremely difficult, and removal of the biofilm from the substratum radically changes the integrity of the sample.

Boundary layers have a large role in biofilm formation, so some discussion concerning them is instructive for the purposes of this section. In fluid mechanics, Prandtl (20) called the region where the viscous forces cannot be ignored the boundary layer. Prandtl arbitrarily suggested that it be considered the region where the x component of the velocity is less that 99% of the free-stream velocity. This is an arbitrary division that does not correspond to any physically measurable boundary but corresponds to an arbitrary mathematical definition. However, it is a very useful concept, clarified numerous unexplained phenomena, and provides a much better intellectual basis for discussing complicated flows (21).

An analogous concept exists for heat and mass transfer. The momentum boundary layer represents the flow regions dominated by viscous forces, and the thermal boundary layer represents a region dominated by conduction. Analogously, the mass boundary layer represents a region dominated by diffusion. The dimensionless numbers corresponding to them are the Reynolds (*Re*), Prandtl (*Pr*), and the Schmidt (*Sc*) numbers, respectively. Figure 6 shows laminar boundary layers and profiles for momentum, heat, and mass. In general, these will not be the same. The flat plate, boundary layer thickness for momentum, heat, and mass are given elsewhere as are the complicated profiles near the wall in turbulent flow (22).

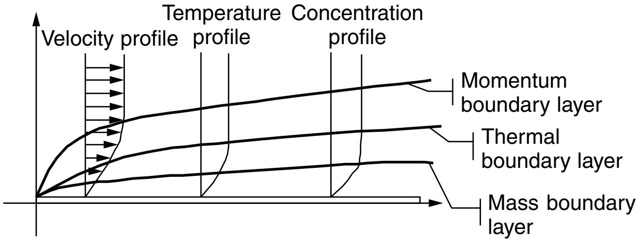

Figure 6. Laminar boundary layers.

The following provides an alternate, more practical correlation for determining the laminar, flat plate momentum boundary thickness (21):

$$\delta_{\mathrm{m}} \approx 5\sqrt{\frac{\nu\, x}{\nu_{\infty}}}. \tag{1}$$

Turbulent boundary layers differ in that there are several zones, including a laminar sublayer, a buffer zone, and a turbulent core. As the distance along a flat plate increases, a transition region develops followed by turbulence, as shown in Fig. 7. Estimating the turbulent boundary layer thickness can be done from (21)

$$\delta_{\mathrm{m}} = 0.37x\left[\frac{\nu}{\nu_{\infty}x}\right]^{1/5}. \tag{2}$$

The laminar sublayer is quite thin but represents the largest resistance to transport in the flow. However, the resistance is much less for turbulent than for laminar flow because the laminar sublayer is so thin and gradients are much steeper.

Microbiologically Induced Corrosion

Microbial cells on a metal surface can cause microbiologically influenced corrosion (MIC) or biocorrosion. The forms of corrosion caused by bacteria are not unique. Biocorrosion results in pitting, crevice corrosion, selective dealloying, stress corrosion cracking, and underdeposit corrosion. The basic corrosion equations for the corrosion cell shown in Fig. 8 are

$$\mathrm{M} \longrightarrow \mathrm{M}^{2+} + 2e^{-} \tag{3}$$

$$\mathrm{H_2O} + \tfrac{1}{2}\mathrm{O_2} + 2e^{-} \longrightarrow 2\mathrm{OH}^{-} \tag{4}$$

Nonuniform (patchy) colonies of biofilm result in the formation of differential aeration cells where areas under respiring colonies are depleted of oxygen relative to surrounding noncolonized areas. Having different oxygen concentrations at two locations on a metal creates a difference in electrical potential and consequently corrosion currents. Under aerobic conditions, the areas under the respiring colonies become anodic, and the surrounding areas become cathodic.

Oxygen depletion at the surface of stainless steel can destroy the protective passive film. Stainless steels rely

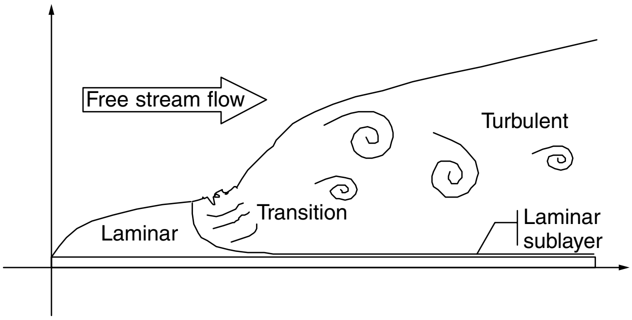

Figure 7. Turbulent boundary layer.

on a stable oxide film to provide corrosion resistance. Corrosion occurs when the oxide film is damaged or oxygen is kept from the metal surface by microorganisms in a biofilm. Oxygen depletion at the surface also provides a condition for anaerobic organisms such as sulfate-reducing bacteria (SRB) to grow. This group of bacteria is one of the most frequent causes for biocorrosion. They reduce sulfate to hydrogen sulfide that reacts with metals to produce metal sulfides as corrosion products. Aerobic bacteria near the outer surface of the biofilm consume oxygen and create a suitable habitat for the sulfate-reducing bacteria at the metal surface. SRBs can grow in water trapped in stagnant areas, such as dead legs of piping. Symptoms of SRB-influenced corrosion are hydrogen sulfide odor, blackening of waters, and black deposits. The black deposit is primarily iron sulfide (5). One method of limiting SRB activity is to reduce the concentration of their essential nutrients: phosphorus, nitrogen, and sulfate. Thus, purified (RO or DI) waters would have fewer problems with SRBs. Also, Any practices that minimize biofilm thickness (flushing, sanitizing, eliminating dead-end crevices) minimize the anaerobic areas in a biofilm that SRB need (5).

Bacteria can produce aggressive metabolites, such as organic or inorganic acids. For example, *Thiobacillus thiooxidans* produces sulfuric acid and *Clostridium aceticum* produces acetic acid. Acids produced by bacteria accelerate corrosion by dissolving oxides (the passive film) from the metal surface and accelerating the cathodic reaction rate. Many microorganisms produce hydrogen gas as a product of carbohydrate fermentation. Hydrogen gas can diffuse into metals and cause hydrogen embrittlement. Iron-oxidizing bacteria, such as *Gallionella, Sphaerotilus, Leptothrix,* and *Crenothrix,* are aerobic and filamentous bacteria that oxidize iron from a soluble ferrous (Fe^{+2}) form to an insoluble ferric (Fe^{+3}) form. The dissolved ferrous iron could be from either the incoming water supply or the metal surface. The ferric iron these bacteria produce can attract chloride ions and produce ferric chloride deposits that can attack austenitic stainless steel. The deposits of iron bacteria on austenitic stainless steel are typically brown or red-brown mounds.

TREATMENT METHODS

Once the microorganisms have attached, they can withstand normal disinfection processes. Biofilm bacteria display high resistance to biocides (23). Research demonstrates that biofilm associated bacteria may be 150–3000 times more resistant to free chlorine and 2–100 times more resistant to monochloramine than free-floating bacteria. Research (24) suggests that *Pseudomonas* has a clever way of eluding its attackers. It secretes a sticky slime that builds up on the pipe interior. A germicide flushed through the water distribution system kills free-floating microbes, but it is less effective against bacteria embedded in the slimy biofilm.

When bacteria are in a film, they are very resistant to biocides. They often produce more exopolymers after biocide treatment to protect themselves. To destroy the cell responsible for forming a biofilm, the disinfectant must first react with the surrounding polysaccharide network. The cells themselves are not actually more resistant; rather they have surrounded themselves with a protective

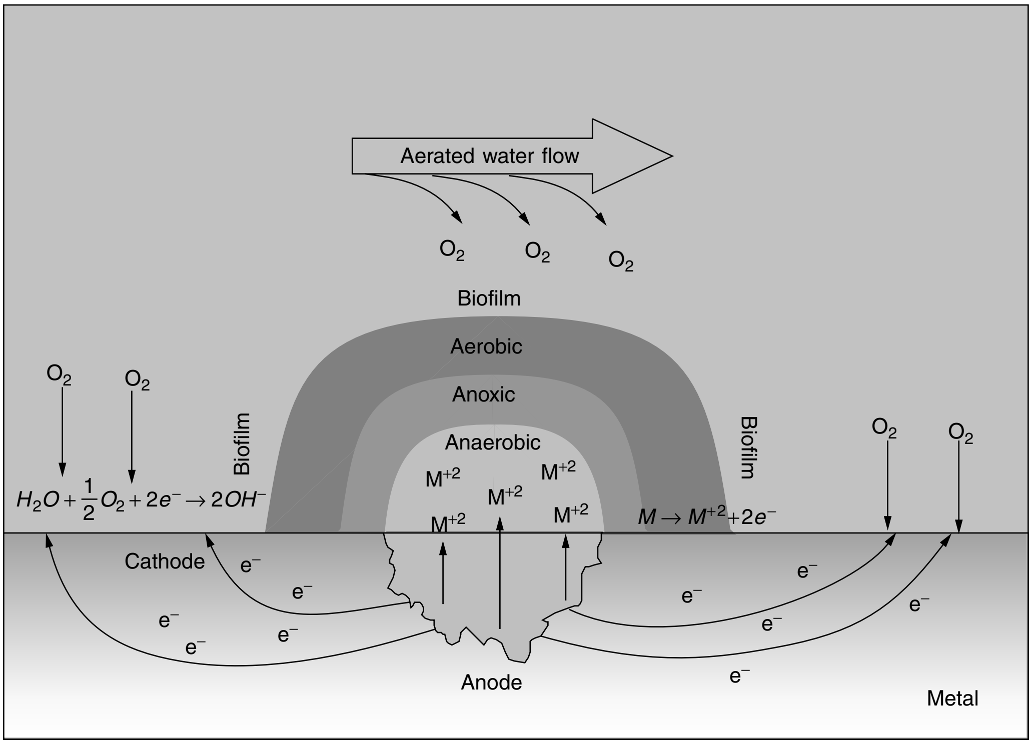

Figure 8. Biofilm corrosion mechanism.

film. The disinfectant's oxidizing power is used up before it can reach the cell.

When cells are attached to a pipe wall, delivery of the disinfectant is limited by the rate of diffusion of the compound across the boundary layer and through the film. It takes a higher concentration in a longer contact time for the disinfectant to reach the bacteria cells in a biofilm compared to free-floating organisms. Trace organics concentrate on surfaces. Extracellular polymers further concentrate trace nutrients from bulk water. Secondary colonizers use waste products from their neighbors. By pooling their biochemical resources, several species of bacteria, each that has different enzymes, can break down food supplies that no single species could digest alone.

Traditional disinfectant testing has been done using single-species free-floating laboratory cultures. The CT constant (concentration × time) for a disinfectant required to kill a particular bacteria should not be extrapolated to bacteria in biofilms. This explains how bacterial counts can be measured, even when water contains low levels of chlorine. Typical chlorine levels in tap water are from 0.5–2.0 ppm. This amount of chlorine kills free-floating bacteria but may not be enough to kill biofilm bacteria. Chunks of sloughed off biofilm can contain viable bacteria which show up in plate counts. This is a particular problem with *Pseudomonas* that is a great slime producer, and so is more chlorine resistant.

The development of biocide resistance is not understood, but recent studies have used a variety of model systems to determine how and why biofilms are so resistant to antimicrobial agents (25). As the importance of biofilms in water treatment has increased, much energy has been directed toward the study of the effects of antimicrobial agents on these surface-attached communities. Biofilm resistance mechanisms include physical or chemical diffusion barriers to antimicrobial penetration into the biofilm, slow growth of the biofilm owing to nutrient limitation, activation of the general stress response, and the emergence of a biofilm-specific phenotype.

Biofilms can be removed and/or destroyed by chemical and physical treatments. Chemical biocides can be divided into two major groups: oxidizing and nonoxidizing. Any chemical that may result in residuals, attack ion exchange resins, or cause corrosion must not be used to control microorganisms. Physical treatments include mechanical scrubbing and hot water. An article by Mittelman (26) has the most comprehensive information on treatment of biofouling in purified water systems. Table 2 provides typical dosage levels and contact times for various biocides.

Table 2. Typical Biocide Dosage Levels[a]

Biocide	Dosage Level, mg/L	Contact Time, h
Chlorine	50–100	1–2
Ozone	10–50[b]	<1
Chlorine dioxide	50–100	1–2
Hydrogen peroxide	10% (v/v)	2–3
Iodine	100–200	1–2
Quaternary ammonium compounds	300–1,000	2–3
Formaldehyde	1–2% (v/v)	2–3
Anionic and nonionic surfactants	300–500	3–4

[a] Reference 26.

[b] Ozone dosage is 10–50 mg/L, but the residual levels in water were 1–2 mg/L.

The effectiveness of oxidizing biocides in purified water systems based on an equal milligram-per-liter dosage decreases in the following order:

$$\text{Ozone } (O_3) > \text{chlorine dioxide } (ClO_2) > \text{chlorine } (Cl_2) > \text{iodine } (I_2) > \text{hydrogen peroxide } (H_2O_2).$$

Chlorine is probably the most effective and least expensive of all oxidizing and nonoxidizing biocides (26). The level of activity of chlorine against attached biofilms is particularly high; planktonic and biofilm bacteria are killed, and chlorine also reacts with and destroys the polysaccharide web and its attachments to the surface. By destroying extracellular polymers, chlorine breaks up the physical integrity of a biofilm. Characklis (27) recommends improving a chlorine treatment program by taking specific additional measures.

In addition to their biocidal activity, quaternary ammonium compounds are effective surfactants/detergents, which may be an important factor in their use for biofilm inactivation and removal from surfaces. Rinseability can be a problem because removal from a purified water system often requires exhaustive rinsing. Formaldehyde has been applied to pharmaceutical-grade systems. It is relatively noncorrosive to stainless steel. Its effectiveness against biofilms is questionable, and it is a toxic carcinogen. Anionic and nonionic surface-active agents (surfactant or detergent compounds) have limited biocidal activity against the bacteria in purified water systems. Applications may be found for these detergents in conjunction with other biocides to improve biofilms and other particulate removal.

Pharmaceutical water-for-injection systems use recirculating hot water loops (hotter than 80 °C) to kill bacteria. When these systems are used continuously, planktonic bacteria are killed, and biofilm development is reduced (26). Biofilms are even found in hot water (80 °C). Periodic hot water sanitization can also be used to destroy bacteria in a biofilm, but this requires a temperature of 95 °C for a period in excess of 100 minutes (28).

Heavy biofilms cannot be removed from storage tank walls by chemicals alone; mechanical scrubbing or scraping, high-pressure spraying, or a combination is also required. Mechanical removal of biofilm from distribution systems is impractical (26).

Unlike antibiotics used to fight bacteria associated with human, animal, and plant diseases, bacteria do not develop the same type of resistance to industrial biocides. The difference between antibiotics and industrial biocides is that an antibiotic may have a small number of target sites on or in a bacterial cell, but all oxidizing biocides have a multitude of potential target sites. It is thought that chlorine, for example, has more than a hundred potential target sites on or in microorganisms. It is virtually impossible for microorganisms to develop general resistance to such compounds (26). However, bacteria in a biofilm can resist biocides because they are shielded in slime.

Bacteria associated with biofilms are much more difficult to kill and remove from surfaces than planktonic organisms. Numerous investigators and plant operators have observed a rapid resumption of biofouling immediately following chlorine treatment (27). Incomplete removal of the biofilm allows it to return quickly to its equilibrium state, causing a rebound in total plate counts following sanitization. Initially, the bulk water bacteria count in one study dropped to zero after sanitization, but this was followed by a gradual increase in numbers to levels at or below pretreatment levels. In this example, regrowth started after 2 days and was back up to equilibrium levels after 20 days. Biofilms can also recover based on various mechanisms (27).

As an oxidizer, ozone is approximately twice as powerful as chlorine at the same concentrations. Like chlorine dioxide, ozone must be generated on-site because of its high reactivity and relative instability. Systems must be designed of appropriate ozone-resistant materials. Ozone is usually dosed continuously at 1–2 mg/L. Success in employing higher dosages noncontinuously has been limited, possibly because of the limited solubility of ozone in purified water; it is difficult to produce high concentrations of ozone in solution (26). Although chlorine is not as powerful as ozone when you compare 1–2 mg/L of each, chlorine can be used in higher sanitizing concentrations with equal disinfecting strength.

Ozone is soluble in water only up to the partial pressure above it; therefore, the concentration of ozone in the generator gas stream becomes very important. According to Henry's law, the efficiency of absorbing a specific gas into an aqueous phase is based on the partial pressure of that particular gas in the total gas flow. Therefore, the higher the partial pressure of ozone in the generator gas stream, the more complete is the concentration of ozone in the water. Table 3 shows the theoretical residuals of ozone that can be achieved in water, based on an ozone gas partial pressure of 1.0 atmosphere and Henry's law constants given in Perry (29). As would be expected, ozone solubility decreases as water temperature increases. Also as a consequence of Henry's law, the solubility of ozone increases as the partial pressure of ozone in the gas stream increases. In practical applications, ozone does not have sufficient contact time to achieve equilibrium conditions, resulting in solubility levels lower than those shown in Table 3. Figure 9 shows a multiple chamber diffusion contact system. The ozonated water stream is led along a pipe located at the bottom of each contact tank. The pipe is perforated by small holes to allow fine bubbles to escape and diffuse into solution.

Table 3. Maximum Ozone Solubility at p_{O_3} = 1.0 Atmosphere

T, °F	H, Atmosphere/x	Solubility, mg/L (ppm)
32	1940	1374.57
41	2180	1223.24
50	2480	1075.27
59	2880	925.93
68	3760	709.22
77	4570	583.52
86	5980	445.93

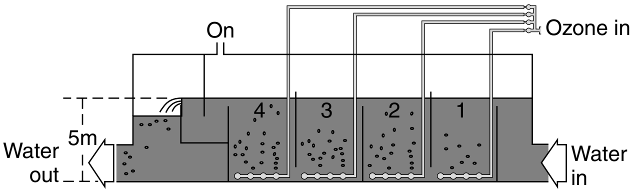

Figure 9. Diffusion contact chambers.

The latest ozone engineering principle involves using a self-contained ozone contacting/off-gassing system. This system means that the large contact tanks found in older system designs are no longer required. This results in significant capital savings on a new pool or retrofit (typically 50 to 75%). This system replaces the traditional contact tank design with a compact self-contained contacting system that is 98% efficient. Instead of a 20,000–30,000 gallon contact tank for a 100,000-gallon pool, the new system uses a powered contacting/mixing/off-gas unit.

Ozone is an oxidative and reactive gas that is harmful to humans above certain concentrations. Ozone gas is heavier than air; therefore, any undissolved ozone must be removed from the system and disposed of correctly. Several methods are available for ozone destruction, the simplest is by ozone adsorption and reaction with granulated activated carbon (GAC). Activated carbon adsorption is extensively used for small applications where air is the ozonator feed gas. This is an adsorption process whereby the reaction consumes the carbon medium. The material used is elemental carbon which has been steam activated to provide a large internal surface area. Carbon is a strong reducing agent; therefore, upon contact with ozone gas, the carbon is oxidized to carbon monoxide and carbon dioxide, resulting in destruction of the ozone molecule. This reaction degrades or powderizes the GAC; therefore, it has a finite life. To this end, the ozone consumes the carbon by slow-rate combustion. Due to the consumption of carbon, the medium must be replaced regularly. When the carbon is saturated with water and washed with sprays, the reaction becomes partially catalytic and requires a larger volume of GAC. For ozone installations where oxygen is used as the feed gas, carbon adsorption destructors must not be used, due to the dangers of combustion. The other method is heat treatment that degrades the ozone rapidly to oxygen.

Hydrogen peroxide is frequently used as a biocide in microelectronic-grade purified water systems because it produces no by-products and rapidly degrades to water and oxygen. A 10% by volume solution in purified water is effective in killing planktonic bacteria, but more studies are needed on the effectiveness against attached biofilms (26). The use of hydrogen peroxide, though not expected to be as effective as ozone, would be very simple to implement.

Peracetic acid, a derivative of hydrogen peroxide, displays a very wide bandwidth of attack against microbes. Research has shown that peracetic acid will be used to an ever-increasing degree in the field of human medicine due to its bacterial, fungicidal, sporicidal, and antiviral action. Because of the wide spectrum of attack, peracetic acid is very suitable as a wide bandwidth disinfectant for ion exchange resins (results of work done by Degussa Technical Applications Department in conjunction with Chemiewerk Homburg). Using a 0.2% peracetic acid solution (in water, with a reaction time of 1 hour), a slime concentration of 104–105/ mL, including mold, was reduced to almost zero. The short rinsing time after using peracetic acid is important (typically about 45 minutes or 10–15BV). In addition to the excellent disinfection action, peracetic acid (according to experiments) has a minimal effect on the ion exchange properties of cation or anion resins.

Peracetic acid (peroxyacetic acid), $C_2H_4O_3$, the peroxide of acetic acid, is a disinfectant that has the desirable properties of hydrogen peroxide: broad-spectrum activity against microorganisms, lack of harmful decomposition products, and infinite water solubility.

Peracetic acid forms in an equilibrium reaction where hydrogen peroxide and acetic acid react:

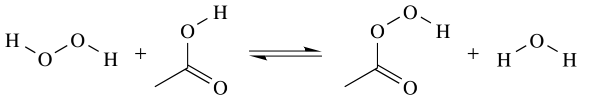
(5)

Peracetic acid also has greater lipid solubility and is free from deactivation by catalase and peroxidase enzymes. However, it is corrosive, and degradation products may have to be rinsed from the surface of disinfected materials. Peracetic acid is a more powerful antimicrobial agent than hydrogen peroxide and most other disinfectants (30–32). It has advantages for disinfection and sterilization not found in any other agent. Against spores of *Bacillus thermoacidurans,* it was reportedly the most active of 23 germicides (33); against a range of bacteria it is lethal at 6–250 ppm, toward yeasts at 25–83 ppm, fungi at 50–500 ppm, bacterial spores at 100–500 ppm, and viruses at 15–2,000 ppm (34). The values obtained are determined by the medium employed and the time necessary for inactivation.

Table 4 gives a comparison of peracetic acid and two other disinfectants against food-poisoning bacteria (35). Peracetic acid retains its activity better than many disinfectants at refrigeration temperatures and is more effective at lower pH values (31). Aqueous solutions are comprised of the acid in combination with hydrogen peroxide, acetic acid, sulfuric acid, water, and a stabilizing agent. All of these ingredients are necessary to keep it stable in storage; the concentration of the hydrogen peroxide in some formulations may considerably exceed that of the peracetic acid. Peracetic acid vapor, like that of hydrogen peroxide, is active against bacterial spores and has been found most effective at 80% relative humidity, it has little activity at 20% relative humidity.

Table 4. Comparison of Disinfectants Against Food-Poisoning Bacteria[a,b]

Organism	Peracetic Acid	Active Chlorine	Benzalkonium Chloride
	At 20 °C		
Listeria monocytogenes	45	100	200
Staphylococcus aureus ATCC 6538	90	860	500
Enterococcus faecium DSM 2918	45	300	250
	At 5 °C		
Listeria monocytogenes	90	860	500
Staphylococcus aureus ATCC 6538	90	1,100	750
Enterococcus faecium DSM 2918	90	450	500

[a] Lethality in 5 min.
[b] Concentrations of disinfectants given in ppm.

Peracetic acid has been accepted worldwide in the food processing and beverage industries as an ideal for clean-in-place systems (36); it does not require rinsing where the breakdown product, acetic acid, is not objectionable in high dilution. Peracetic acid is more toxic than hydrogen peroxide and is a weak carcinogen but can be used with safety when diluted. Like all peroxides, it is a powerful oxidizer and should be handled with proper safety precautions. It is more corrosive to metals and plastics than hydrogen peroxide (36).

UV energy is found in the electromagnetic spectrum between visible light and X rays and can best be described as invisible radiation. The energy employed for UV water treatment is further categorized into two primary levels measured as wavelengths—254 nm and 185 nm, where nm = 1/1000 of a micron. For disinfection and ozone destruction, the 254-nm wavelength is used. All living organisms contain DNA (deoxyribonucleic acid). DNA provides the mechanism for all functions needed to sustain life.

UV light disinfection systems emit UV light that penetrates the outer cell membrane of microorganisms, passes through the cell body, reaches the DNA, and alters the genetic material. The microorganism is thereby destroyed nonchemically and cannot reproduce.

Ultraviolet processing involves using radiation from the ultraviolet region of the electromagnetic spectrum for disinfection (37). Typically, the wavelength for UV processing ranges from 100 to 400 nm, as shown in Fig. 10. This range may be further subdivided into UVA (315 to 400 nm) normally responsible for changes in human skin that lead to tanning; UVB (280 to 315 nm) that can cause skin burning and eventually lead to skin cancer; UVC (200 to 280 nm), called the germicidal range, because it effectively inactivates bacteria and viruses, and the vacuum UV range (100 to 200 nm) that can be absorbed by almost all substances and thus can be transmitted only in a vacuum.

The germicidal properties of UV irradiation are mainly due to DNA mutations induced through absorption of UV light by DNA molecules. This mechanism of inactivation

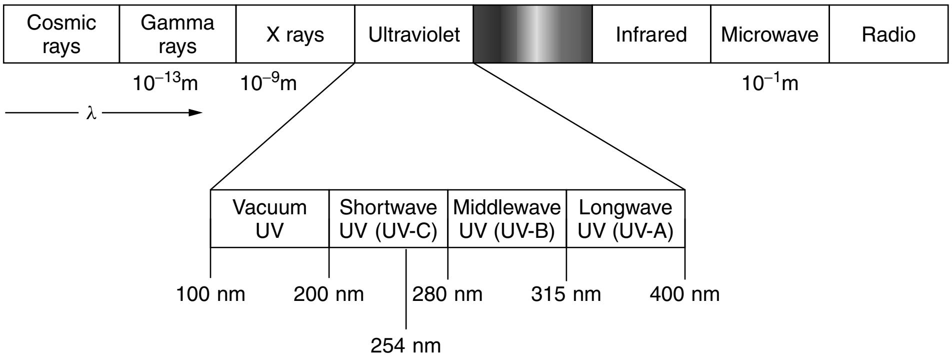

Figure 10. UV light spectrum.

results in a sigmoidal curve of microbial population reduction. To achieve microbial inactivation, the UV radiant exposure must be at least 400 J/m^2 in all parts of the product. Critical factors include the transmissivity of the product; the geometric configuration of the reactor; the power, wavelength and physical arrangement of the UV source(s); the product flow profile; and the radiation path length. UV may be used in combination with other alternative processing technologies, including powerful oxidizing agents such as ozone and hydrogen peroxide. Applications include disinfection of water supplies and food contact surfaces. Recently, interest has increased in using UV to reduce microbial counts in juices.

The shape of the curve for microbial inactivation by UV light is sigmoidal. Figure 11 is an idealized depiction of a sigmoidal curve. The initial plateau is due to an injury phase of the microorganism in response to UV exposure. After the initial plateau, the maximum amount of injury has been surpassed; thus, minimal additional UV exposure would be lethal for microorganisms, and survivor numbers rapidly decline. The end of the curve has a tailing phase due to UV resistance of the microorganisms and to experimental components, such as suspended solids, that may block the UV irradiation.

Much of the prior literature has focused on UV disinfection of water supplies. The literature is insufficient to develop comprehensive microbial inactivation reaction kinetics data or models. There are, however, studies relative to the UV radiant exposure required to obtain a four-log reduction of various microorganisms. These data suggest that the log reduction is related to the UV radiant exposure. The curve has a linear section with a shoulder and tailing effects (38). All organisms tested were reduced by four-log cycles when UV radiant exposure was less than 400 J/m^{237}. In some bacterial cells, photoreactivation, a repair mechanism that is enhanced by visible light in the blue spectral range, may occur. Photoreactivated cells show greater resistance to UV radiation than nonreactivated cells.

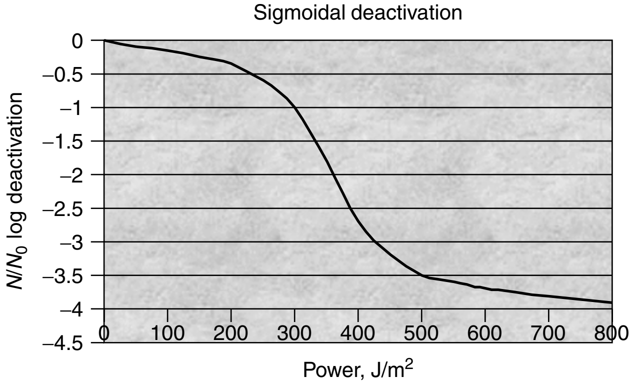

Figure 11. Idealized microorganism deactivation.

The germicidal properties of ultraviolet irradiation are due to DNA absorption of the UV light, causing cross-linking between neighboring pyrimidine nucleoside bases (thymine and cytosine) in the same DNA strand (39). Due to the mutated base, formation of hydrogen bonds to purine bases on the opposite strand is impaired. DNA transcription and replication is thereby blocked, compromising cellular functions and eventually leading to cell death. The amount of cross-linking is proportional to the amount of UV exposure. The level of mutations that can be reversed depends on the UV repair system in the target microorganism. Once the threshold of cross-linking is exceeded, the number of cross-links is beyond repair, and cell death occurs (39). This phenomenon is reflected in the shape of the inactivation curve described before. The cross-linking threshold corresponds to the point of rapid decline after the initial plateau phase on the sigmoidal survival plot of UV exposure.

To disinfect water, it is essential that all parts (each volume element) of the product receive UV radiant exposure of at least 400 J/m^2 (at 254 nm) to reduce human pathogens and viruses by at least four-log cycles (40). Thus, the homogeneity of the flow pattern and the radiation field may have critical effects on disinfection. An additional critical factor is the transmissivity of the material being disinfected. If the material is highly transparent to UV light, disinfection may be more effective; turbid materials attenuate and scatter UV radiation, resulting in less microbial inactivation. The thickness of the radiation path through a material is also important because attenuation increases with the length of passage; thus, the geometric configuration of process systems is critical. Another critical factor to consider is the UV wavelength used because it affects microbial inactivation.

The growth of organisms increases as temperature increases. The effect of temperature on the reaction rate of a biological process is usually expressed in the following form (41):

$$r_T = r_{20}\theta^{(T-20)} \tag{6}$$

This is similar to the Arrhenius equation used in chemistry; an approximate doubling of reaction rate for each 10 °C temperature rise. The range is limited because there is little biological activity in ice or boiling water. Figure 12 illustrates this and shows that a decrease from 80 to 60 °F has a significant impact on the growth rate. The growth rate can be reduced, but the temperature changes that are practical in pools do not kill the organisms.

Increasing temperature also increases the kill rate (decreases kill time). The effect of temperature on the kill rate of microorganisms is well represented by a form of the Van't Hoff–Arrhenius relationship (41). In terms of the kill time (t in the equation) required for a given percentage kill, the relationship is

$$\ln\frac{t_1}{t_2} = \frac{E(T_2 - T_1)}{RT_1T_2} \tag{7}$$

Using $T_2 = 20\,^\circ\mathrm{C}$ and chlorine at pH = 7.0, the effects of temperature on kill rate are illustrated in Fig. 13. Figure 13 shows that kill time can be significantly reduced by modest increases in temperature.

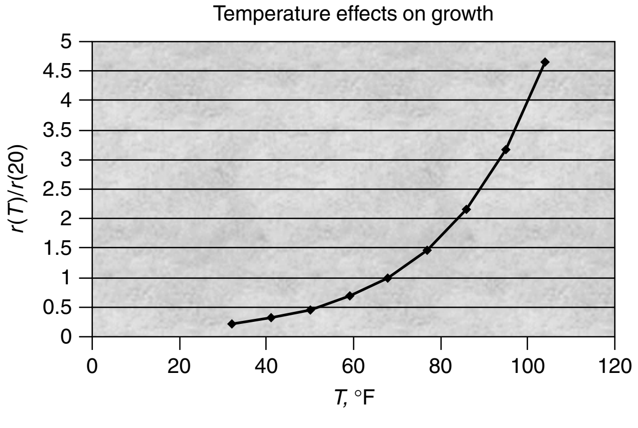

Figure 12. Temperature effects on biological growth.

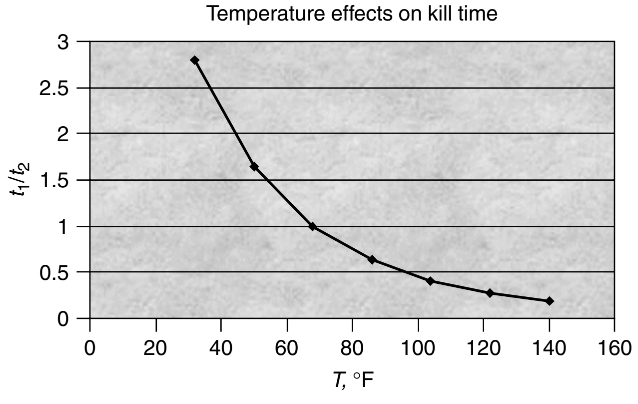

Figure 13. Temperature effects on kill time.

The reduction of light may reduce the growth rate of some microorganisms. Microbes may be classified according to their energy and carbon requirements (42). Growth from light is expected for some organisms as a result of photosynthesis. For example, algae use carbon dioxide in photosynthetic activity to produce new cells, oxygen, and water.

The criteria for chemical addition and other microbiological control for a typical nuclear fuel storage pool include

- minimal corrosion
- no damage occurs to other parts and/or systems
- no residual remains that can adsorb onto ion exchange resins
- personnel safety.

This severely limits the treatment chemicals that can be used, for example, chlorine cannot be used. Although there may be beneficial affects from controlling and/or manipulating light and temperature, the focus is on treatment to remove microorganisms and keep them minimized. Also, as previously discussed, the chemicals focus on those that will not leave residuals or cause corrosion. Therefore, the focus is on ozone, peracetic acid, and hydrogen peroxide. The International Atomic Energy Agency has studied this in some detail (1). They provided a list of potential methods for removing microorganisms from spent fuel pools that should be considered in a final study or evaluation. Ion exchange columns must be bypassed when ozone is at a high concentration in the water because ozone attacks and degrades the resins (43). It is believed that ozone is effective in long residence times and for sustained residuals. This expectation is partly based on experience from vendors using ozone in biofilm-fouled swimming pools (44).

The use of chemical agents must be demonstrated for personnel safety. The maximum possible dissolved ozone in the pools that transfers to the atmosphere would be related to Henry's law. At 86 °F, $H = 5980$ atmospheres/mole fraction, and the residual is 1 mg/L. The mole fraction (x) is then,

$$x = \frac{1\ \mathrm{mg}}{L} \times \frac{L}{1,000\ \mathrm{g}} \times \frac{18\ \mathrm{g}}{\mathrm{mole}} \times \frac{\mathrm{mole}}{48\ \mathrm{g}} \times \frac{\mathrm{g}}{10^3\ \mathrm{mg}} = 4\times10^{-7}$$

The partial pressure (maximum) in equilibrium with the pool surface is based on Henry's law constant (H) in Table 3:

$$\begin{aligned} p_i &= Hx = 5980 \times 4 \times 10^{-7} \\ &= 2 \times 10^{-3}\ \text{atmospheres (2,000 ppm)} \end{aligned}$$

The STEL/Ceiling/TWA is 0.1 ppm, so this is significantly higher. A material balance is used to determine the effective vapor phase concentration. Figure 14, material balance shows the ozone transferring from the water surface into the ventilation air. The material balance is thus:

$$Q_{\mathrm{ozone}} = (Q_{\mathrm{vent}} + Q_{\mathrm{ozone}})\, y_{\mathrm{ozone}} \tag{8}$$

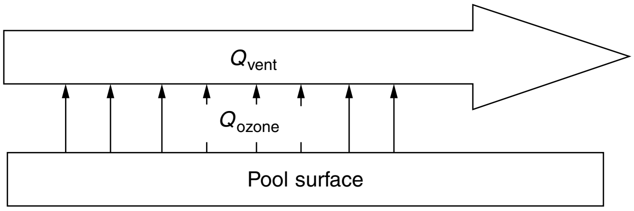

Figure 14. Material balance [SCA1][nlh2]diagram.

A dissolved gas has two resistances, gas-phase and liquid-phase. The mass transfer (F) flux is

$$F = K_L(C_B - C^*) = K_G(p^* - p_B) \tag{9}$$

The flux is also equal to the individual driving forces as

$$F = k_G(p_I - p_B) = k_L(C_B - C_I) \tag{10}$$

By combining these and using Henry's law, the following are determined for the overall mass transfer coefficients, and either one can be used:

$$K_L = \frac{1}{\dfrac{1}{Hk_G} + \dfrac{1}{k_L}} \tag{11}$$

$$K_G = \frac{1}{\dfrac{1}{k_G} + \dfrac{H}{k_L}} \tag{12}$$

According to Thibodeaux (45), the gas-phase and liquid-phase coefficients for water and CO_2 are 3000 cm/h and 20 cm/h, respectively. The ozone value can be directly used as 20 because the molecular weight falls in the range for direct substitution. However, the value of the gas-phase coefficient is determined from the ratio of the cube roots of molecular weights (46):

$$k_{G-O_3} = k_{G-H_2O}\left(\frac{18}{48}\right)^{1/3} = 2163$$

The overall coefficient is then

$$K_L = \frac{1}{\dfrac{1}{5,980 \times 2,163} + \dfrac{1}{20}} = 20$$

Thus it is seen that the transfer is totally controlled by the liquid phase. The material balance then becomes

$$\frac{p_B}{P_T}\frac{M_{O_3}}{M_{\text{air}}} = \frac{K_L A\left(C_B - \left(\frac{p_B}{H}\right)\right)}{K_L A\left(C_B - \left(\frac{p_B}{H}\right)\right) + Q_{\text{vent}}} \tag{13}$$

For small p_B, this can be simply solved as

$$p_B = \frac{K_L A C_B}{K_L A C_B + Q_{\text{vent}}}\frac{M_{\text{air}}}{M_{O_3}}P_T \tag{14}$$

The estimated partial pressure of ozone was based on the values in Table 3. The NIOSH time weighted average (TWA) is 0.1 ppm. The estimated concentration is 1.27 ppm that exceeds the limit. Therefore, the liquid concentration would need to be controlled at less than 0.1 mg/L to ensure personnel protection. However, an ozone system that intermittently supplies high-concentration ozone was effective in preventing biofilm formation using a 1-mg/L ozone dose for 10 minutes (47). Asea Brown Boveri (ABB) carried out tests with ozone as a biocide on the water of a small industrial cooling tower in Switzerland (48). The test results showed that ozone is an excellent biocide. In 1988, an ozonation plant was installed at the side stream cooling water system of a heating power station in Germany (49). The operation showed that ozone works well at 0.1 mg/L for 2–3 hours contacting time per day. Ozone destroys and reacts with pollutants in the air. If this is considered, the ozone concentration in air in pool areas will be much lower than the worse case calculation (Table 5).

The concentration above the pools using the 10-minute treatment is then

$$C_{10\,\text{min}} = 1.27 \times \frac{10\,\text{min}}{24\text{ h} \times 60\text{ min/h}} = 0.00882$$

This result is well below the TWA of 0.1 ppm.

NOMENCLATURE

A	Area
C	Concentration
C_I	Concentration at interface
C_B	Concentration in bulk solution
C^*	Fictitious concentration in bulk gas
CT	Concentration × time
E	Activation energy, J/g mole
F	Mass transfer flux
H	Henry's law constant
k_G	Local gas mass transfer coefficient
k_L	Local liquid mass transfer coefficient
K_G	Overall mass transfer coefficient, based on gas
K_L	Overall mass transfer coefficient, based on liquid
M	Molecular weight
p	Partial pressure, atmospheres
p_I	Partial pressure at interface, atmospheres

Table 5. Calculation of Ozone Above Pools

Parameter	Value
M_{O_3}, g/mole	48.00
M_{air}, g/mole	29.00
P_T, atm	0.87
Q_{vent}, scfm	28,100
Q_{vent}, g/h	6.2E+07
C_B, mg/L	1.00
K_L, cm/h	20.00
K_L, m/h	0.20
A, m^2	648.14
$F \times A$, g/h	129.63
Calculated p, atm	1.10E−06
Calculated O_3, ppm	1.27

p_{B}	Partial pressure in bulk, atmospheres
p^*	Fictitious pressure in bulk liquid
P_{T}	Total pressure, atmospheres
Q	Gas mass flow rate
R	Gas constant
Re	Reynolds number
r_{T}	Reaction rate at $T\,^{\circ}$C
r_{20}	Reaction rate at $20\,^{\circ}$C
T	Temperature
t	Time
v	Velocity
x	Mole fraction
y	Vapor mass fraction
δ	Boundary layer thickness
κ	Conductivity
ν	Kinematic viscosity
θ	Temperature-activity coefficient

BIBLIOGRAPHY

1. International Atomic Energy Agency. (1988). *Durability of Spent Nuclear Fuels and Facility Components in Wet Storage*. IAEA-TECDOC-1012, Vienna, Austria.
2. Edstrom. (2002). Available: http://www.edstrom.com/Resources.cfm?doc_id=23.
3. Mittelman, M.W. (1985). Biological fouling of purified-water systems: Part 1, Bacterial growth and replication. *Microcontamination* **3**(10): 51–55, 70.
4. Mayette, D.C. (1992). The existence and significance of biofilms in water. *Water Rev.* Water Quality Research Council, Lisle, IL, pp. 1–3.
5. Geesey, G.G., Lewandowski, Z., and Flemming, H.-C. (Eds.). (1994). *Biofouling and Biocorrosion in Industrial Water Systems*. Lewis, Ann Arbor, MI.
6. Coghlan, A. (1996). Slime city. *New Sci.* **15**(2045): 32–36.
7. VanHaecke, E. et al. (1990). Kinetics of *Pseudomonas aeruginosa* adhesion to 304 and 316-L stainless steel: Role of cell surface hydrophobicity. *Appl. Environ. Microbiol.* **56**(3): 788–795.
8. Stanley, P.M. (1983). Factors affecting the irreversible attachment of *Pseudomonas aeruginosa* to stainless steel. *Can. J. Microbiol.* **29**(11): 1493–1499.
9. Costerton, J.W. et al. (1995). Microbial biofilms. *Annu. Rev. Microbiol.* **49**: 711–745.
10. Hamilton, N.F. (1988). Antimicrobial controls effects of bioslime. *Modern Plast.* 166–168.
11. Flemming, H.C. and Geesey, G.G. (Eds.). (1991). *Biofouling and Biocorrosion in Industrial Water Systems*. Springer-Verlag, New York.
12. Meltzer, T.H. (1993). *High-Purity Water Preparation for the Semiconductor, Pharmaceutical, and Power Industries*. Tall Oaks, Littleton, CO.
13. Patterson, M.K., Husted, G.R., Rutkowski, A., and Mayette, D.C. (1991). Isolation, identification, and microscopic properties of biofilms in high-purity water distribution systems. *Ultrapure Water* **8**(4): 18–24.
14. Pittner, G.A. and Bertler, G. (1988). Point-of-use contamination control of high purity water through continuous ozonation. *Ultrapure Water* **5**(4): 16–22.
15. Husted, G.R., Rutkowski, A.A., and Couture, A. (1994). Response of oligotrophic biofilm bacteria in high-purity water systems to stepwise nutrient supplementation. *Ultrapure Water* **11**(6): 43–50.
16. Bott, T.R. (1998). Techniques for reducing the amount of biocide necessary to counteract the effects of biofilm growth in cooling water systems. *Appl. Thermal Eng.* **18**(11): 1059–1066.
17. Stoodley, P., Lewandowski, Z., Boyle, J.D., and Lappin-Scott, H.M. (1999). Structural deformation of bacterial biofilms caused by short-term fluctuations in fluid shear: An *in situ* investigation of biofilm rheology. *Biotechnol. Bioeng.* **65**(1): 5.
18. Boyle, J.D., Dodds, I., Stoodley, P., and Lappin-Scott, H.M. (1997). Stress management in biofilms. In: *Biofilms: Community Interactions and Control*. J.W.T. Wimpenny, P.S. Handley, P. Gilbert, H.M. Lappin-Scott, and M. Jones (Eds.). BioLine, Cardiff, UK, pp. 15–22.
19. Picologlou, B.F., Zelver, N., and Characklis, W.G. (1980). Biofilm growth and hydraulic performance. *J. Hydraul. Div. Am. Soc. Civ. Eng.* **106**(HY5): 733–746.
20. Prandtl, L. (1904). *Ueber Fluessigkeitbewegung mit Kleiner Reibung (Concerning Fluid Movements with Small Friction)*, *Verhandl. III Int. Math.*-Kong., Heidelberg; also reprinted in *Vier Abhandlungen zur Hydrodynamik*. Goettingen, 1927.
21. De Nevers, N. (1970). *Fluid Mechanics*. Addison-Wesley, Reading, MA.
22. Bird, R.B., Stewart, W.E., and Lightfoot, E.N. (1960). *Transport Phenomena*. John Wiley & Sons, New York.
23. LeChevallier, M.W., Cawthon, C.D., and Lee, R.G. (1988). Inactivation of biofilm bacteria. *Appl. Environ. Microbiol.* **54**(10): 2492–2499.
24. Anderson, R.L. et al. (1990). Effect of disinfectants on *Pseudomonas* colonized on the interior surface of PVC pipes. *AJPH*. 17–21.
25. Mah, T.C. and O'Toole, G.A. (2001). Mechanisms of biofilm resistance to antimicrobial agents. *Trends Microbiol.* **9**(1): 34–39.
26. Mittelman, M.W. (1986). Biological fouling of purified-water systems: Part 3, Treatment. *Microcontamination* **4**(1): 30–40, 70.
27. Characklis, W.G. and Marshall, K.C. (Eds.). (1990). *Biofilms*. John Wiley & Sons, New York.
28. Collentro, W.C. (1995). Microbial control in purified water systems—case histories. *Ultrapure Water* **12**(3): 30–38.
29. Perry, R.H. and Green, D.W. (1979). *Perry's Chemical Engineers' Handbook*, 6th Edn. McGraw-Hill, New York.
30. Eggensberger, H. (1979). *Zentralbl. Bakteriol. Mikrobiol. Hyg. [B]* **168**: 517–524.
31. Baldry, M.G.C. (1983). *J. Appl. Bacteriol.* **54**: 417–423.
32. Krzywicka, H. et al. (1975). *Resistance of Microorganisms to Disinfectants*. W.B. Kedzia (Ed.). Polish Acad. of Sciences, Warsaw.
33. Hutchins et al. (1949).
34. Block, S.S. (1991). *Disinfection, Sterilization, and Preservation*, 4th Edn. Lea and Febiger, Philadelphia, PA, pp. 167–181.
35. Orth, R. and Mrozeck, H. (1989). *Fleischwirtsch* **69**: 1575–1576.
36. Dychdala, G.R. (1988). *Proc. 4th Conf. Prog. Chem. Disinfection*. Binghamton, NY, pp. 315–342.
37. FDA. (2000). Available: http://vm.cfsan.fda.gov/~comm/ift-uv.html.

38. Hoyer, O. (1998). Testing performance and monitoring of UV systems for drinking water disinfection. *Water Supply* **16**(1/2): 419–442.
39. Miller, R., Jeffrey, W., Mitchell, D., and Elasri, M. (1999). Bacterial responses to ultraviolet light. *Am. Soc. Microbiol.* **65**(8): 535–541.
40. Bernhardt, H. (1994). Desinfektion Aufbereiteter Oberflachenwasser mit UV-Strahlen. *Wasser-Abwasser*. **135**(12): 677–689.
41. Tchobanoglous, G. and Burton, F.L. (1991). *Wastewater Engineering; Treatment, Disposal, and Reuse*, 3rd Edn. Metcalf & Eddy, McGraw-Hill, New York.
42. Frobisher, M., Hinsdill, R.D., Crabtree, K.T., and Goodheart, C.R. (1994). *Fundamentals of Microbiology*. Saunders, Philadelphia, PA.
43. Purolite. (2002). Personal conversation, Sam Ashworth with Francis Boodoo, The Purolite Company, Bala Cynwyd, PA.
44. Finney, A. (2002). Personal communication, Sam Ashworth with Allan Finney (email), Envron Inc., Regina, Sk., Canada.
45. Thibodeaux, L.J. (1979). *Chemodynamics: Environmental Movement of Chemicals in Air, Water, and Soil*. Wiley-Interscience, New York.
46. Crowl, D.A. and Louvar, J.F. (1990). *Chemical Process Safety Fundamentals with Applications*. Prentice-Hall, Englewood Cliffs, NJ.
47. Nakayama, S. et al. (1985). Anti-biofouling ozone system for cooling water circuits. II. An application of seawater. *Ozone: Sci. Eng.* **7**(1): 31–46.
48. Wellauer, R. (1990). Cooling water treatment with ozone. *Ozone: Sci. Eng.* **12**(3): 243–253.
49. Liechti, P.A. et al. (1992). One year full-scale study of ozone cooling water treatment at a German electric power station. *Ozone: Sci. Eng.* **14**(6): 531–544.

INDUSTRIAL MINE USE: MINE WASTE

Constantin Von Der Heyden
School of Geography and the
Environment
Oxford, United Kingdom

INTRODUCTION

Intensive industrial development and particularly the production of effluents from mining operations has led to worldwide concern for environmental and health impacts following decades of contaminated effluent discharge into natural waterways (1). Developing economies often depend heavily on extractive or primary industry, and many third-world governments face the difficult task of reconciling economic development with environmental and social protection. Concern arising from the discharge of polluted effluent from the mining industry was recorded as early as 1556 (2), and the legacy of unsustainable development is clearly chronicled in sediments worldwide (3–6). as the discharge of mining and industrial effluents has led to a substantial increase in riverine, estuarine, and marine sediment metal concentrations (7–9).

Several types of acid and toxic effluents are produced by mining and metallurgical processes during operations and following site decommission (Table 1). These effluents include tailings excess decant (rainwater or process water running off tailings impoundments), process acid streams, and mine drainage waters. Drainage from tailings impoundments is often circumneutral in pH and contains low concentrations (<20 mg/L) of total dissolved metals (10). Process acid stream effluents originate from the use or generation of acids during metallurgical processes and can contain very high levels of dissolved solids and toxic metals. Mine drainage waters are commonly of three main types: (1) saline formation waters; (2) acidic, heavy-metal- and sulfate-containing waters formed during biologically mediated and abiotic oxidation of sulfides (so-called acid mine drainage or AMD); and (3) alkaline, hydrogen-sulfide-containing, heavy-metal-poor waters resulting from buffering reactions (usually carbonate) and sulfate reduction within AMD (11,12). AMD surfaces as discharge from adits and abandoned workings of sulfidic mines or as drainage from sulfide-rich overburden material or spoil heaps (13).

ACID MINE DRAINAGE GENERATION

Acidity generation due to sulfide oxidation occurs by two processes: proton acidity and metal acidity. The former occurs during the oxidation of sulfides, and the latter is from the hydrolysis of metal species in solution (15). Each is discussed here in more detail.

Iron pyrite, the most common mineral leading to AMD, reacts with oxygen and water according to the following reaction (16):

$$2FeS_2 + 7O_2 + 2H_2O \leftrightarrow 2Fe^{2+} + 4SO_4{}^{2-} + 4H^+ \qquad (1)$$

This reaction often occurs in underground workings or deep within spoil heaps and overburden, resulting in the removal of oxygen and the discharge of an anaerobic effluent. Accordingly, oxygen is a limiting substrate within the deeper workings or lower levels of the spoil heap where anaerobic conditions prevail throughout the year. Younger (17) differentiates the production of AMD from flooded workings into "juvenile" and "vestigial" acidity: the former is the dissolution of weathering salts formed on sulfides when the water table within the workings rises following cessation of dewatering activities; the latter is the continuing oxidation of sulfides in the seasonally flooded zone of the workings as the water table rises and drops within the annual hydrologic cycle.

Table 1. Typical Mine Drainage Composition[a]

	Coal Mines	Cu–Pb–Zn Mixed Sulfide Mines	EPA Statutory Limits
pH	2.6–6.3	2.0–7.9 mg/L	6–9
Al	1–58		
Cu		0.005–76	0.05
Fe	1–473	8.5–3200	3.5
Mn	1–130	0.4	2
Pb		0.02	0.2
Zn		0.04–1600	0.2–0.5

[a] Reference 14.

On contact with the atmosphere, which often commences only at the site of effluent discharge from flooded workings or spoil heaps, the oxidation of ferrous to ferric iron and the precipitation of ferric oxyhydroxides generates further acidity:

$$Fe^{2+} + O_2 + 4H^+ \leftrightarrow Fe^{3+} + 2H_2O \qquad (2)$$

$$Fe^{3+} + 3H_2O \leftrightarrow Fe(OH)_3 + 3H^+ \qquad (3)$$

As noted, Equations 1 and 2 are dependent on an aerobic environment. However, once ferric iron has been formed, Fe^{3+} acts as the oxidizing agent and the oxidation of pyrite can proceed in an anaerobic environment:

$$FeS_2 + 14Fe^{3+} + 8H_2O \leftrightarrow 15Fe^{2+} + 2SO_4{}^{2-} + 16H^+ \qquad (4)$$

The reaction kinetics in Equations 1 and 3 are significantly higher than those of Equation 2, and accordingly the oxidation of ferrous iron is the rate-limiting step in this series (11). However, numerous iron oxidizing bacteria can markedly increase the oxidation of ferrous to ferric iron. Of these, *Thiobacillus ferrooxidans* and *Leptospirilium ferrooxidans* are likely to be the most important organisms, increasing the rate of reaction by six orders of magnitude (10^6) (11). These organisms function optimally below pH 3, thus requiring a degree of pyrite oxidation and ferric hydrolysis to lower the solution pH to optimal levels (18). Johnson and Hallberg (19) and Hallberg and Johnson (20) show that a host of other organisms known as moderate acidophils, which include sulfur-oxidizing bacteria and archaea, heterotrophic microorganisms, and some lower eukaryotic life-forms, function optimally at pH 3–5.5 and may be involved in the initial sulfide oxidation and proton acidity generation.

Other members of the pyrite family, such as arsenopyrite (FeAsS) and chalcopyrite ($CuFeS_2$), also generate proton acidity upon oxidation, releasing As or Cu into solution in addition to Fe. Monosulfide minerals, such as chalcocite (Cu_2S), sphalerite (ZnS), and galena (PbS), are also often encountered in sulfidic deposits. Although the oxidation of these minerals does not generate proton acidity (see Equation 5), subsequent hydrolysis of the metal ions generates metal acidity as in Equation 3:

$$ZnS + 2O_2 \leftrightarrow Zn^{2+} + SO_4{}^{2-} \qquad (5)$$

TYPES OF MINE EFFLUENT REMEDIATION FACILITIES

Mine effluent discharge is almost always a point source, so mine-water problems may be addressed by isolating the contaminant source and either treating the discharge *in situ* (i.e., within the working or the spoil heap by suppressing the reactions releasing contaminants) or by active or passive water treatment following discharge. Innovative *in situ* treatment techniques such as galvanic suppression, application of bactericides, and introduction of neutralizing or reducing agents are receiving increasing attention and represent real potential. However, such techniques have not yet achieved widespread application and require further applied and fundamental research on both field and laboratory scales.

Active or conventional treatment processes for effluent remediation usually involve expensive technology and infrastructure and high process costs. Kuyucak (10) lists some of the common active, physicochemical processes:

1. active neutralization systems
2. ion exchange materials, such as resins and zeolites (21)
3. membrane processes, such as reverse osmosis, ultrafiltration/microfiltration
4. solvent extraction
5. electrochemical treatment (22)
6. bioremediation through controlled sulfate reduction (23,24)

Passive treatment technology, which as the name implies has significantly lower capital and maintenance costs, commonly includes (13)

1. aerobic, surface-flow wetlands (reed beds)
2. subsurface flow systems, such as reducing and alkalinity producing systems—RAPS (25,26)
3. combination wetlands, such as compost wetlands (27)
4. anoxic and oxic limestone drains (28,29)
5. biosorbents (30–32)

Active neutralization facilities are still most widely used to treat mine waters but vary greatly in the degree of sophistication, depending on the geography and local specifications of the site, the nature of the effluent, and various other environmental, technical, and financial constraints. As such, the facilities can range from the simple addition of neutralizing agents to the effluent stream, to complex chemical plants consisting of reactors, clarifiers, and sludge dewatering technology. Lime as CaO or $Ca(OH)_2$ is the most common neutralizing agent used for treating low pH effluent, owing to the widespread availability, low cost, and high reactivity of the material. However, lime neutralization results in poor quality effluent and the need to dispose of large volumes of sludge; accordingly, certain situations require using other neutralizing reagents, such as $Mg(OH)_2$, Na_2S, NH_3, Na(OH) and $CaCO_3$ (10).

The current state-of-the-art active lime neutralization process is called the high density sludge process (HDS), where more than one reactor is used to perform the neutralization, and a mixture of sludge, recycled from the clarifier, and lime is used as the neutralizing agent in the first reactor (Fig. 1). Polymers are added to improve flocculation, and a clarifier is used to enhance solid/liquid separation, decreasing the volume and increasing the stability of the disposal sludge.

PROCESS IN MINE EFFLUENT REMEDIATION

The activity of hydrogen ions in solution (the pH of the solution) is the key determinant in the composition of mine waters and of their toxicity (16). Metal mineral solubility

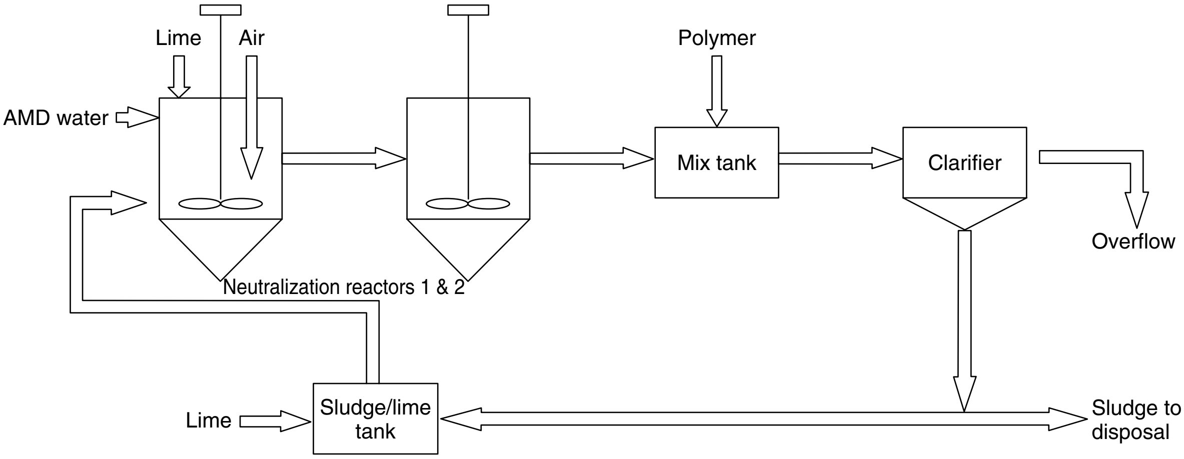

Figure 1. The high density sludge process (10).

increases markedly at low pH, and the precipitation of the dissolved metal cation as an insoluble species is low. Moreover, adsorption, which is often a primary process of metal removal from anthopogenically polluted systems (7,9), is also often strongly pH dependent (33). Adsorption approaches zero as the pH decreases below the adsorption edge of the metal, which for common metals such as Cu, Co, Ni, Pb, and Zn is around pH 4–5 (34). Anions, for example SO_4, demonstrate the converse; a pH decrease leads to greater anion adsorption (34,35). The free aquo ion is the most bioavailable, so the pH of the solution also affects its toxicity, directly through pH-mediated effects on organisms and indirectly as a result of its effect on the concentration of anions (e.g., oxyanions of Cr and As) and toxic metals such as Cu, Pb, and Zn (36–38). As such, correcting the pH is the most important and first consideration in mine effluent remediation (13).

Assuming that effluent pH has been regulated within an acceptable range (pH 5–9), Hammer (39) and Dunbabin and Bowmer (40) list the most common chemical and biological processes by which mine effluent quality is improved:

1. oxidation and hydrolysis
2. reduction reactions
3. adsorption of pollutant complexes to organic and inorganic surfaces
4. uptake of pollutants by plants
5. physical filtering of suspended solids by plants and substrate
6. Neutralization and precipitation by the formation of ammonium (NH_4^+) and carbon dioxide (CO_2)

Oxidation and Hydrolysis

The change in valence state following oxidation of certain metals can dramatically influence their solubility: oxidation of Fe^{2+} and Mn^{2+}, both highly soluble in the reduced state, to Fe^{3+} and Mn^{4+} results in rapid metal precipitation as oxide and (oxy)hydroxides.

The complexation of metal cations with hydroxide ions results in the formation of insoluble metal hydroxide (Equation 6), oxide, and oxyhydroxide precipitates (33,41). Hydrolysis and oxidative hydrolysis are the most common processes of metal precipitate formation in aerobic surface waters (13,42). However, hydroxide precipitation decreases the pH of the water due to the production of hydrogen ions: the metal acidity discussed above (Equation 6). As such, oxidation of effluent with high metal concentrations [e.g., Wheal Jane, see (43)] can rapidly lower the pH of the water and hence affect the mobility and solubility of other metals (15,44).

$$M^{2+(or)3+} + H_2O \leftrightarrow M(OH)_{2(or)3} \downarrow +2(or)3H^+ \quad (6)$$

Reduction Reactions

Reduction processes are used in active bioreactors and in RAPS, and numerous researchers identify bacterially mediated sulfate reduction as the most effective way to decrease sulfate and heavy metal concentrations in contaminated mine drainage (23,45,46). The reduction of sulfate and the production of hydrogen sulfide lead to an increase in alkalinity and the precipitation of insoluble metal sulfides (Equations 7 and 8):

$$SO_4 + C_xH_yO_z \leftrightarrow H_2S + wHCO_3 \quad (7)$$

$$M^{2+} + H_2S \leftrightarrow MS \downarrow +2H^+ \quad (8)$$

Adsorption by Organic and Inorganic Surfaces

Adsorption and surface complexation commonly occurs on sediments, suspended or colloidal matter, and plant surfaces (47). Within contaminated drainage, transport of pollutants is often facilitated by adsorption on a variety of suspended and colloidal solid surfaces. Colloidal aggregate formation and settling and the settling of suspended solids are common processes whereby such contaminants are incorporated into sediments.

Adsorption is highly pH-dependent (33,41), for example, SO_4 adsorption onto $Fe(OH)_3$ approaches 100% below pH 3, whereas metal cations dominate the adsorption surface above pH 4.5 (34). The most important and widely studied adsorption surfaces in mine effluent impacted systems are (oxy)hydroxide precipitates of iron and manganese (35,48–51). The order of metal affinity for MnO_2 binding sites is $Co > Ni \geq Zn > Cu$ (52), and that for FeOOH sites is $Pb > Zn \geq Co \geq Ni > Cu$ (34).

Metals are also complexed with a range of organic material such as algae, bacterial cells, detritus, and organic coatings on mineral surfaces (40,47,53,54). The organic adsorption affinity of metals decreases in the order $Cu = Fe \gg Zn \gg Mn$ (55,56). Adsorption and surface complexation are rapid processes producing pollutant attenuation from the water column but are transitory due to marked pH dependency (33,41). Ions immobilized as surface complexes are rapidly released during changes in effluent pH conditions.

The Effect of Aquatic Vegetation

The uptake and accumulation of metals in plant biomass has been extensively studied under laboratory conditions (57–59). Dunbabin and Bowmer (40) investigated the partitioning of metals within emergent hydrophytes and found that the majority of metals were stored in the roots and rhizomes. Floating wetland plants have been shown to hyperaccumulate Cu and Fe up to 78 times their concentration in wastewater (57). However, direct uptake and accumulation of metals within plant biomass usually constitutes a minor component of the overall removal processes (14,60–62).

In addition to metal uptake, wetland plants attenuate metal contaminants through a number of further mechanisms:

- Plants release oxygen through their roots, creating a zone of aerobic conditions in the substrate (14,40). The oxidizing conditions within the substrate increase metal oxidation with insoluble precipitate formation [commonly Fe (oxy)hydroxide plaques]. These precipitates provide a further surface for trace metal adsorption (63).
- Bacterial biofilms forming on root surfaces are strong adsorbents of trace metals (63).
- The turnover of aquatic vegetation is the dominant source of simple carbohydrates, which are the main substrate for sulfate-reducing bacteria and other fermentative organisms important in the production of sulfides, ammonium, carbon dioxide, and alkalinity (14,40,64,65).
- The humic substances of organic decay play an important role in aquatic contaminant chemistry through the adsorption of metals and subsequent transport as colloidal complexes or incorporation into the sediment as settled solids.

The Role of Microbial Processes

Microbial activity is integral to many of the processes described above. The importance of the microbiology of mine effluent remediation has recently been recognized, as evidenced by numerous studies in this discipline (66–69). Nordstrom (70) provides a good review of the state of knowledge and the gaps in the present understanding of the microbiology of acid mine waters. The processes that use or depend on, microbial action have been described above. Hence, listed here are the most common examples of microbial involvement in processes of mine effluent remediation:

- oxidation of Mn (71,72)
- oxidation and hydrolysis of Fe (73,74)
- oxidation of sulfides (11)
- sulfate reduction
- products of organic decay (CO_2, NH_4^+, and organic acids)
- surfaces for adsorption

BIBLIOGRAPHY

1. Vernet, J.-P. (1991). *Heavy Metals in the Environment*. Elsevier, Amsterdam.
2. Agricola, G. (1556). *De Re Metallica*. H.C. Hoover and L.H. Hoover (Eds.). (1950). Dover, New York.
3. Chen, M.H. and Hung, T.W. (1993). Copper, cadmium and lead in sediment from Kaohsiung River and its harbour area, Taiwan. *Mar. Pollut. Bull.* **30**(12): 879–884.
4. Everaarts, J.M. (1989). Heavy metals (Cu, Zn, Cd, Pb) in sediment of the Java Sea, estuarine and coastal areas of East Java and some deep sea areas. *Neth. J. Sea Res.* **23**: 403–413.
5. Sabri, A.W., Rasheed, K.A., and Kassim, T.I. (1993). Heavy metals in the water, suspended solids and sediments of the River Tigris Impoundment at Samarra. *Water Res.* **27**(6): 1099–1103.
6. Sin, S.N., Chua, H., Lo, W., and Ng, L.M. (2001). Assessment of heavy metal cations in sediments of the Shing Mun River, Hong Kong. *Environ. Int.* **26**: 296–301.
7. Förstner, U. (1983). Assessment of metal pollution in rivers and estuaries. In: *Applied Environmental Geochemistry*. I. Thornton (Ed.). Academic Press, London, pp. 395–423.
8. Förstner, U. and Wittmann, G.T.W. (1979). *Metal Pollution in the Aquatic Environment.* Springer-Verlag, Berlin.
9. Tessier, A. and Campbell, P. (1988). Partitioning of trace metals in sediments. In: *Metal Speciation: Theory, Analysis and Application*. H.E. Allen (Ed.). Lewis, Ann Arbor, MI, pp. 183–199.
10. Kuyucak, N. (1998). Mining, the environment and the treatment of mine effluents. *Int. J. Environ. Pollut.* **10** (2): 315–325.
11. Singer, P.C. and Stumm, W. (1970). Acid mine drainage: The rate limiting step. *Science* **167**: 1121–1123.
12. Banks, D., Younger, P.L., Arnesen, R.T., Iversen, E.R., and Banks, S.B. (1997). Mine-water chemistry: The good, the bad and the ugly. *Environ. Geol.* **32**(3): 157–174.
13. Younger, P.L., Banwart, S.A., and Hedin, R.S. (2002). *Mine Water: Hydrology, Pollution, Remediation*. Kluwer Academic, Dordrecht, the Netherlands.
14. Gazea, B., Adams, K., and Kantapoulos, A. (1996). A review of passive treatment systems for the treatment of acid mine drainage. *Miner. Eng.* **9**(1): 23–42.
15. Younger, P.L. (1998). Coalfield abandonment: Geochemical processes and hydrochemical products. In: *Energy and the Environment: Geochemistry of Fossil, Nuclear and Renewable Resources*. K. Nicholson (Ed.). Society for Environmental Geochemistry and Health, McGregor Science, Aberdeen, Scotland, pp. 1–29.

16. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry*. John Wiley & Sons, New York.
17. Younger, P.L. (1997). The longevity of minewater pollution: A basis for decision-making. *Sci. Total Environ.* **194**: 457–466.
18. Bigham, J.M. (1996). Influence of pH on mineral speciation in a bioreactor simulating acid mine drainage. *Appl. Geochem.* **11**(6): 845–849.
19. Johnson, D.B. and Hallberg, K.B. (2003). The microbiology of acidic mine waters. *Res. Microbiol.* **154**(7): 466–473.
20. Hallberg, K.B. and Johnson, D.B. (2003). Novel acidophiles isolated from moderately acidic mine drainage waters. *Hydrometallurgy* **71** (1–2): 139–148.
21. Riveros, P.A. (2004). The extraction of Fe(III) using cation-exchange carboxylic resins. *Hydrometallurgy* **72**(3–4): 279–290.
22. Chartrand, M.M.G. and Bunce, N.J. (2003). Electrochemical remediation of acid mine drainage. *J. Appl. Electrochem.* **33**(3–4): 259–264.
23. Buisman, C.J. and Schultz, C.E. (2000). From microbiology to industrial application of sulphate reducing bacteria. *WISA-Minewater Drainage Technol. Conf.* Rhodes University, Grahamstown, January 23–28, 2000.
24. Kaksonen, A.H., Riekkola-Vanhanen, M.L., and Puhakka, J.A. (2003). Optimization of metal sulphide precipitation in fluidized-bed treatment of acidic wastewater. *Water Res.* **37**(2): 255–266.
25. Sikora, F.J., Behrends, L.L., Brodie, G.A., and Taylor, H.N. (2000). Design criteria and required chemistry for removing manganese in acid mine drainage using subsurface flow wetlands. *Water Environ. Res.* **72**(5): 536–544.
26. Stephenson, B., Harris, T.M., and Rosche, W.A. (2001). Treatment of mine drainage with subsurface flow wetlands. *Abs. Papers Am. Chem. Soc.* **221**: 506.
27. Wieder, R.K. (1993). Ion input/output budgets for 5 wetlands constructed for acid coal-mine drainage treatment. *Water Air Soil Pollut.* **71**(3–4): 231–270.
28. Cravotta, C.A. and Trahan, M.K. (1999). Limestone drains to increase pH and remove dissolved metals from acidic mine drainage. *Appl. Geochem.* **14**(5): 581–606.
29. Cravotta, C.A. (2004). Size and performance of anoxic limestone drains to neutralize acidic mine drainage. *J. Environ. Qual.* **33**(3): 1164–1164.
30. Xie, J.Z., Chang, H.L., and Kilbane, J.J. (1996). Removal and recovery of metal ions from wastewater using biosorbents and chemically modified biosorbents. *Bioresource Technol.* **57**(2): 127–136.
31. Kratochvil, D. and Volesky, B. (1998). Biosorption of Cu from ferruginous wastewater by algal biomass. *Water Res.* **32**(9): 2760–2768.
32. Utgikar, V., Chen, B.Y., Tabak, H.H., Bishop, D.F., and Govind, R. (2000). Treatment of acid mine drainage: I. Equilibrium biosorption of zinc and copper on nonviable activated sludge. *Int. Biodeterioration Biodegradation* **46**(1): 19–28.
33. Langmuir, D. (1997). *Aqueous Environmental Geochemistry*. Prentice-Hall, Englewood Cliffs, NJ.
34. Dzombak, D.A. and Morel, F.M. (1990). *Surface Complexation Modelling: Hydrous Ferric Oxide*. John Wiley & Sons, New York.
35. Rose, S. and Elliott, W.C. (2000). The effects of pH upon the release of sulphate from ferric precipitates formed in acid mine drainage. *Appl. Geochem.* **15**(1): 27–34.
36. Campbell, P.G.C. and Stokes, P.M. (1985). Acidification and toxicity of metals to aquatic biota. *Can. J. Fish. Aquatic Sci.* **42**(12): 2034–2049.
37. Gerhardt, A. (1993). Review of the impact of heavy metals on stream invertebrates with special emphasis on acid conditions. *Water Air Soil Pollut.* **66**: 289–314.
38. Mason, C.F. (1996). *Biology of Freshwater Pollution*. Addison Wesley Longman, Edinburgh.
39. Hammer, D.A. (1990). *Constructed Wetlands for Wastewater Treatment*. Lewis, Ann Arbor, MI.
40. Dunbabin, J.S. and Bowmer, K.H. (1990). Potential use of constructed wetlands for treatment of industrial wastewater containing metals. *Sci. Total Environ.* **111**: 151–168.
41. Pankow, J.F. (1991). *Aquatic Chemistry Concepts*. Lewis, Ann Arbor, MI.
42. Eger, P. (1994). Wetland treatment for trace metal removal from mine drainage: The importance of aerobic and anaerobic processes. *Water Sci. Technol.* **29**(4): 249–256.
43. Hamilton, Q.U.I., Lamb, H.M., Hallett, C., and Proctor, J.A. (1999). Passive treatment systems for the remediation of acid mine drainage at Wheal Jane, Cornwall. *J. Chartered Inst. Water Environ. Manage.* **13**(2): 93–103.
44. Younger, P.L. (2000). Mine water pollution in the long-abandoned Cleveland Ironstone field, north-east England. *7th Natl. Hydrol. Symp.* British Hydrological Society, Newcastle Upon Tyne, September 6–8, 2002.
45. Cloete, E. (2000). Biological sulphate reduction in artificial acid mine drainage using different carbon sources. *WISA-Minewater Drainage Technol. Conf.* Rhodes University, Grahamstown, January 23–28, 2000.
46. Hedin, R.S., Hammack, R., and Hyman, D. (1990). Potential importance of sulphate reduction processes in wetlands constructed to treat mine drainage. In: *Constructed Wetlands for Wastewater Treatment*. D.A. Hammer (Ed.). Lewis, Ann Arbor, MI, pp. 508–514.
47. Drever, J.I. (1997). *The Geochemistry of Natural Waters*. Prentice-Hall, Englewood Cliffs, NJ.
48. Kay, J.T., Conklin, M.H., Fuller, C.C., and O'Day, P.A. (2001). Processes of nickel and cobalt uptake by a manganese oxide forming sediment in Pinal Creek, Globe Mining District, Arizona. *Environ. Sci. Technol.* **35**(24): 4719–4725.
49. Kim, J.-Y. and Chon, H.-T. (2001). Pollution of a watercourse impacted by acid mine drainage in Imgok Creek of the Gangreung coal field, Korea. *Appl. Geochem.* **16**: 1387–1396.
50. Schemel, L.E., Kimball, B.A., and Bencala, K.E. (2000). Colloid formation and metal transport through two mixing zones affected by acid mine drainage near Silverton, Colorado. *Appl. Geochem.* **15**: 1003–1018.
51. Yu, J.-Y. and Heo, B. (2001). Dilution and removal of dissolved metals from acid mine drainage along Imgok Creek, Korea. *Appl. Geochem.* **16**: 1041–1053. .
52. Suarez, D.L. and Langmuir, D. (1976). Heavy metal relationships in Pennsylvania soil. *Geochim. Cosmochim. Acta* **40**: 589–598.
53. Noller, B.N., Woods, P.H., and Ross, B.J. (1994). Case studies of wetland filtration of mine waste water in constructed and naturally occurring systems in Northern Australia. *Water Sci. Technol.* **29**(4): 257–265.
54. Stevens, S.E., Dionis, K., and Stark, L.R. (1990). Manganese and iron encrustation on green algae living in acid mine drainage. In: *Constructed Wetlands for Wastewater Treatment*. D.A. Hammer (Ed.). Lewis, Ann Arbor, MI, pp. 765–773.

55. Cohen, R.H. (2000). Conjunction of geochemical modelling and pilot scale experiments for testing bioreactor removal of arsenic and chromium. *WISA-Minewater Drainage Technol. Conf.* Rhodes University, Grahamstown, January 23–28, 2000.
56. Tipping, E. and Hurley, M.A. (1992). A unifying model of cation binding by humic substances. *Geochim. Cosmochim. Acta* **56**: 3627–3641.
57. Jain, S.K., Vasudevan, P., and Jha, N.K. (1989). Removal of some heavy metals from polluted water by aquatic plants: Studies on duckweed and water velvet. *Biological Wastes* **28**: 115–126.
58. Zayed, A., Gowthaman, S., and Terry, N. (1998). Phytoaccumulation of trace metals by wetland plants: Duckweed. *J. Environ. Qual.* **27**: 715–721.
59. Zhu, Y.L., Zayed, A.M., Qian, J.-H., de Sousa, M., and Terry, N. (1999). Phytoaccumulation of trace metals by wetland plants: Water hyacinth. *J. Environ. Qual.* **28**: 339–344.
60. Cohen, R.H. and Staub, M.W. (1992). *Technical Manual for the Design and Operation of a Passive Mine Drainage Treatment System*. Report to the U.S. Bureau of Land Reclamation. Colorado School of Mines, Golden, CO.
61. Hedin, R.S., Nairn, R.W., and Kleinmann, R.L.P. (1994). *Passive Treatment of Polluted Coal Mine Drainage, Bureau of Mines Information Circular 9389*. United States Department of the Interior, Washington, DC.
62. Shutes, R.B., Ellis, J.B., Revitt, D.M., and Zhang, T.T. (1993). The use of *Typha latifolia* for heavy metal pollution control in urban wetlands. In: *Constructed Wetlands for Water Quality Improvement*. G.A. Moshiri (Ed.). Lewis, Ann Arbor, MI.
63. Hansel, C.M. and Fendorf, S. (2001). Characterisation of Fe plaque and associated metals on the roots of mine-waste impacted aquatic plants. *Environ. Sci. Technol.* **35**(19): 3863–3868.
64. Batchelor, A. (2000). Some perspectives on the use of wetlands for the amelioration of mine drainage. *WISA-Minewater Drainage Technol. Conf.* Rhodes University, Grahamstown, January 23–28, 2000.
65. Pulles, W. (2000). Development of passive mine water treatment technology. *WISA-Minewater Drainage Technol. Conf.* Rhodes University, Grahamstown, January 23–28, 2000.
66. Costley, S.C. and Wallis, F.M. (2001). Bioremediation of heavy metals in a synthetic wastewater using a rotating biological contactor. *Water Res.* **35**(15): 3715–3723.
67. Glombitza, F. (2001). Treatment of acid lignite mine flooding water by means of microbial sulphate reduction. *Waste Manage.* **21**(2): 197–203.
68. Pagnanelli, F., Toro, L., and Veglio, F. (2002). Olive mill solid residues as heavy metal sorbent material: A preliminary study. *Waste Manage.* **22**(8): 901–907.
69. Podda, F., Zuddas, P., Minacci, A., Pepi, M., and Baldi, F. (2000). Heavy metal coprecipitation with hydrozincite $[Zn\text{-}_5(CO_3)_2(OH)_6]$ from mine waters caused by photosynthetic micro-organisms. *Appl. Environ. Microbiol.* **66**(11): 5092–5098.
70. Nordstrom, D.K. (2000). Advances in the hydrogeochemistry and microbiology of acid mine waters. *Int. Geol. Rev.* **42**(6): 499–515.
71. Chapnick, S.D., Moore, W.S., and Nealson, K.H. (1982). Microbially mediated manganese oxidation in a fresh-water lake. *Limnol. Oceanogr.* **27**(6): 1004–1014.
72. Thornton, F.C. (1995). Manganese removal from water using limestone filled tanks. *Ecological Eng.* **4**: 11–18.
73. Burke, S.P. and Banwart, S.A. (2002). A geochemical model for removal of iron (II) (aq) from mine water discharges. *Appl. Geochem.* **17**(4): 431–443.
74. Zhang, C.L. et al. (1997). Physiochemical, mineralogical, and isotopic characterization of magnetite-rich iron oxides formed by thermophilic iron-reducing bacteria. *Geochim. Cosmochim. Acta* **61**(21): 4621–4632.

SUGARCANE INDUSTRY WASTEWATERS TREATMENT

María-del-Carmen
Durán-de-Bazúa
UNAM, National Autonomous
University of Mexico
México D.F., México

Sugar (sucrose) is a sweet, crystalline, white or colorless substance, derived from the juice of several plants. World sugar production amounts approximately to 120 million tonnes, of which two-thirds come from sugarcane and one third from sugarbeet. Sugarcane is a perennial grass that grows between the tropics (30° latitude north and 30° latitude south) in more than 100 countries. It is the most efficient earth grass plant for storing solar energy as biomass and reaches field yields up to 150 ton/ha. A variety of products are obtained from this versatile plant (Fig. 1). Many studies on sugarcane are published in journals and books. Sugarcane processing to produce sugar represents one of the oldest "industries" in modern times. In America, the first sugarcane processing plants were located in the New Spain territories as well as in the Caribbean islands since the sixteenth century. In Mexico, for example, the first sugarcane mill was established in Veracruz on the Gulf of Mexico in 1525 (1,2).

The power for sugarcane mills originally came from water sources (rivers) and from cane bagasse burning. Considering that up to the first half of the twentieth century, industry was a synonym for smoke, wastewater, and wastes in general, the sugarcane industry was no exception. Therefore, considering modern sustainability concepts, the more than 2000 sugar processing plants still operate around the world with relatively poor technologies from an environmentally friendly point of view. Figure 2 shows a schematic diagram of a sugarcane factory, including the wastes generated.

Roughly, one metric tonne of sugarcane renders the following products:

350	kg wet bagasse	35%
100	kg sugar	10%
60	kg straw and leaves	6%
40	kg final syrup, molasses	4%
40	kg cachasses	4%
100	kg cane heads	10%
310	kg evaporated water	31%
1000	kg sugarcane	100%

After sugar, the by-product with more added value is the final syrup, known as molasses. Molasses is a

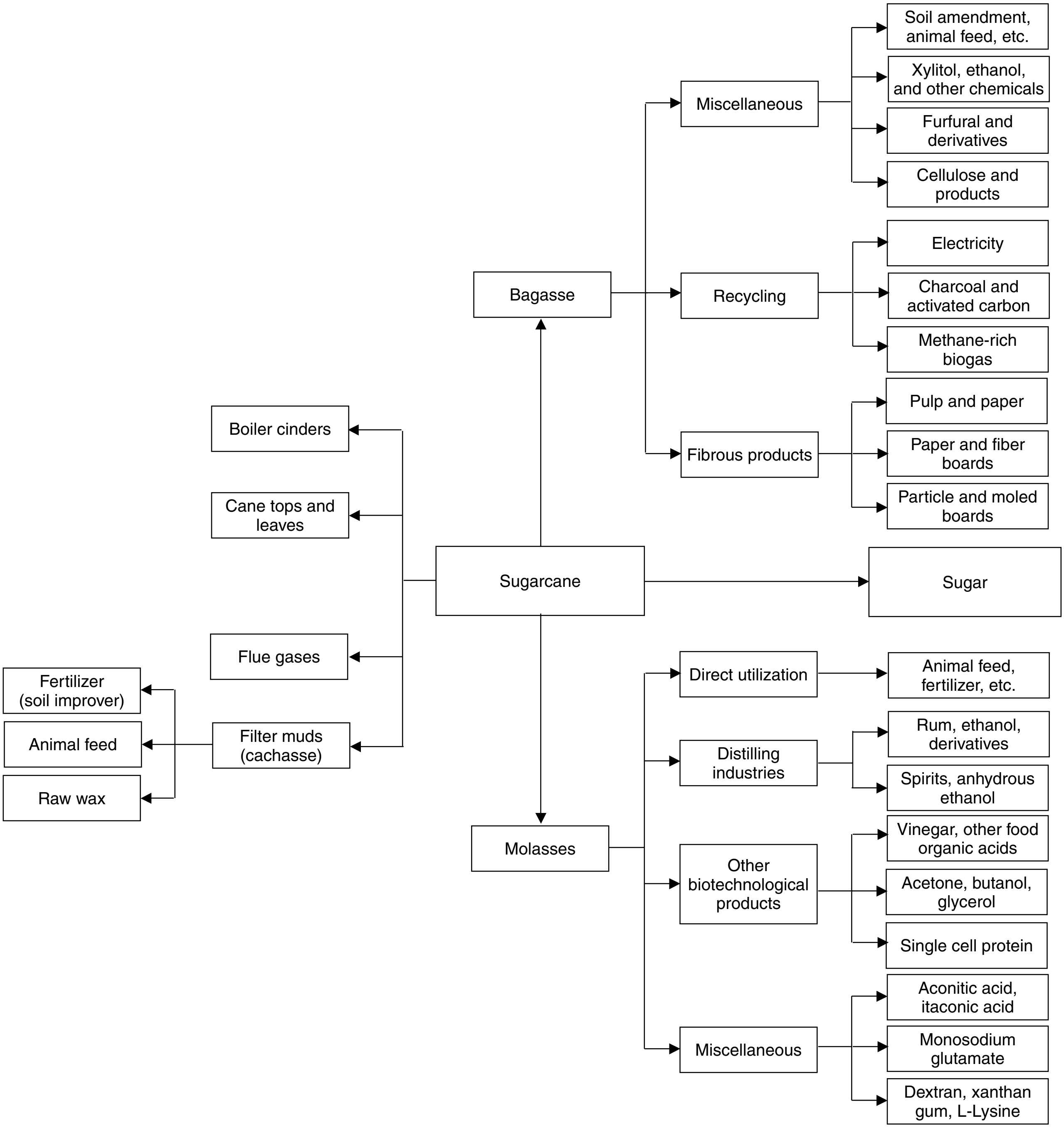

Figure 1. Versatile sugarcane use.

very inexpensive carbon source for many biotechnological products, including vaccines and antibiotics (Table 1). Of course, fermented and distilled beverages, such as rum, are more widely known biotechnological products.

In Figs. 3 and 4, diagrams for sugar and ethyl alcohol production are presented. In Fig. 3, the example is the production of raw sugar. The two other most popular sugar commercial products are the so-called plantation white sugar or mill white sugar (in Mexico, it is known as "standard" sugar) and refined sugar. Literature presents block diagrams for the production of these different types of sugar (1). Wastewaters are the most conspicuous wastes, and many processes have been devised to reduce their polluting impact and to gain some added value from treating them (5). Tables 2 and 3 show the average composition of some of these wastewaters, and Table 4 presents the regulatory limits that Mexico established in the 1990s for this agro-based industry.

It is clear from Table 3 that liquid effluents from a distillery in the sugarcane mill have a very important impact on the overall composition of its wastewaters. Also, to comply with the regulations for both sugar production and distilled products, namely ethanol (ethyl alcohol), a removal efficiency of more than 95 and

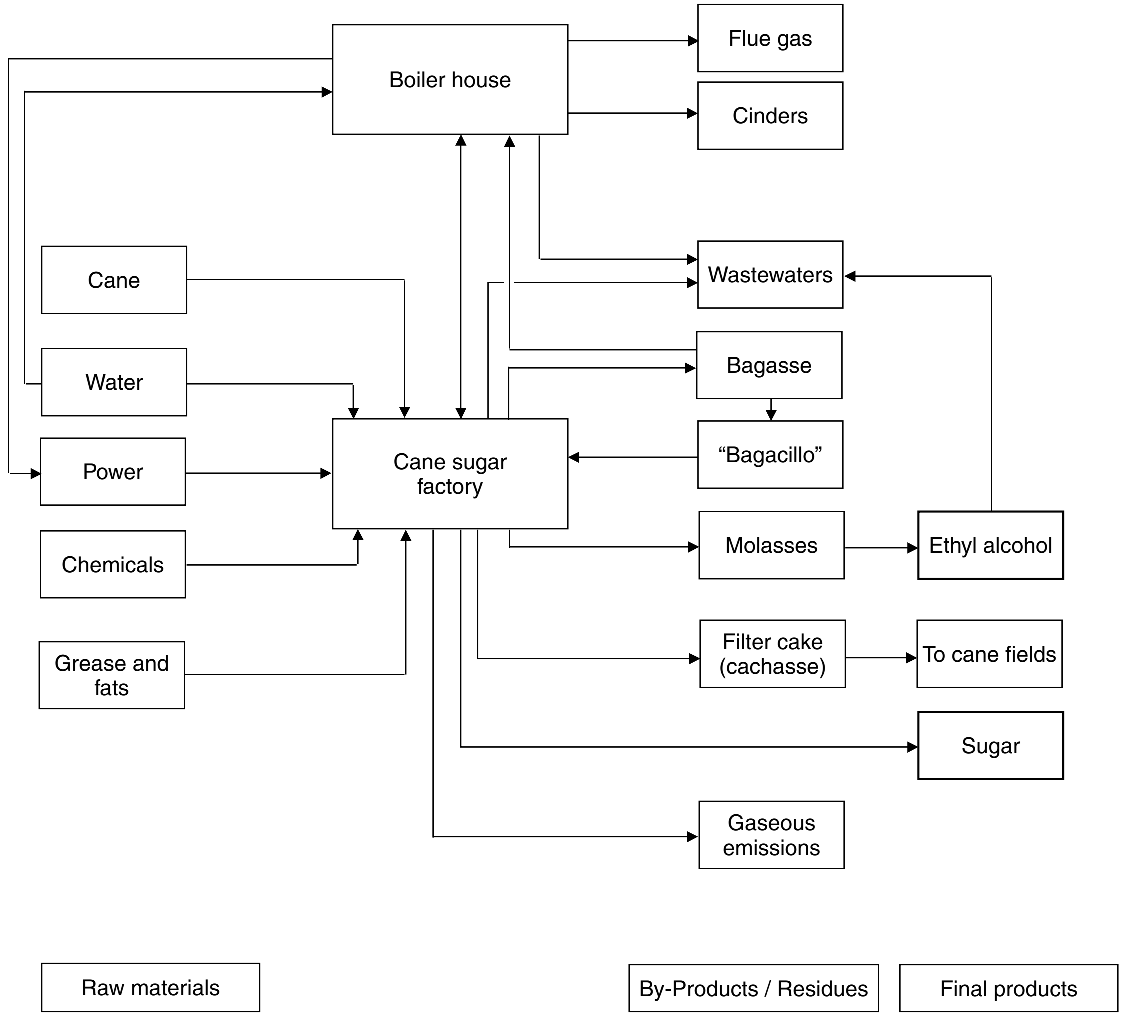

Figure 2. Raw materials, products, and by-products of the sugarcane industry.

99.8% must be reached for general wastewaters and vinasses, respectively. Most of these vinasses, to date, are discharged to receiving bodies (soil, surface water sources) without any treatment or with a partial depuration because of the notion that these wastewaters improve soil. However, this is true only when soils are very poor. For good agricultural soils, they have a negative effect (9–11).

To reduce costs, biological systems are preferred to physicochemical processes. For that reason, in general, anaerobic systems are chosen for treating those process streams that contain biodegradable compounds because methane-rich biogas can be a plus from the bioconversion. To improve the biotreatment and reuse the water in the process, aerobic polishing and sometimes a physicochemical treatment are added. From the aerobic treatment, the biomass produced can be used in feedlots, particularly for fish (12). Leftover treated water can be used for irrigating cane fields. In tropical countries, it is very common to "cultivate" river shrimp (*Cammarus montezumae*), known in Mexico as "*acociles*" (from the Nahuatl or Aztec language, atl = water, cuitzilli = bent, that bends in the water). Then, an interesting biological cycle is rendered. Aerobic biomass together with carotenoproteins from cephalothorax and exoskeleton from river shrimp (the unedible portion), are pelleted for fish feedlots, and the unedible residues from fish are ground and used for shrimp cultivation. Figure 3 shows a complete nature-like cycle process instead of the typical man-made linear polluting processes (Fig. 2).

When wastewaters contain mainly high concentrations of inorganic compound and they are highly soluble, as

Table 1. Molasses Composition[a]

Parameter	% weight	Parameter	% weight
Water	20	K_2O	3.5
Sucrose	32	CaO	1.5
Glucose	14	P_2O_5	0.2
Fructose	14	Carbonates, CO_3^{2-}	1.6
Nitrogen compounds and other	10	Sulfate ions	0.4
Organic products		Other inorganics	0.8
		Density, g m/L	1.42

[a] Reference 3.

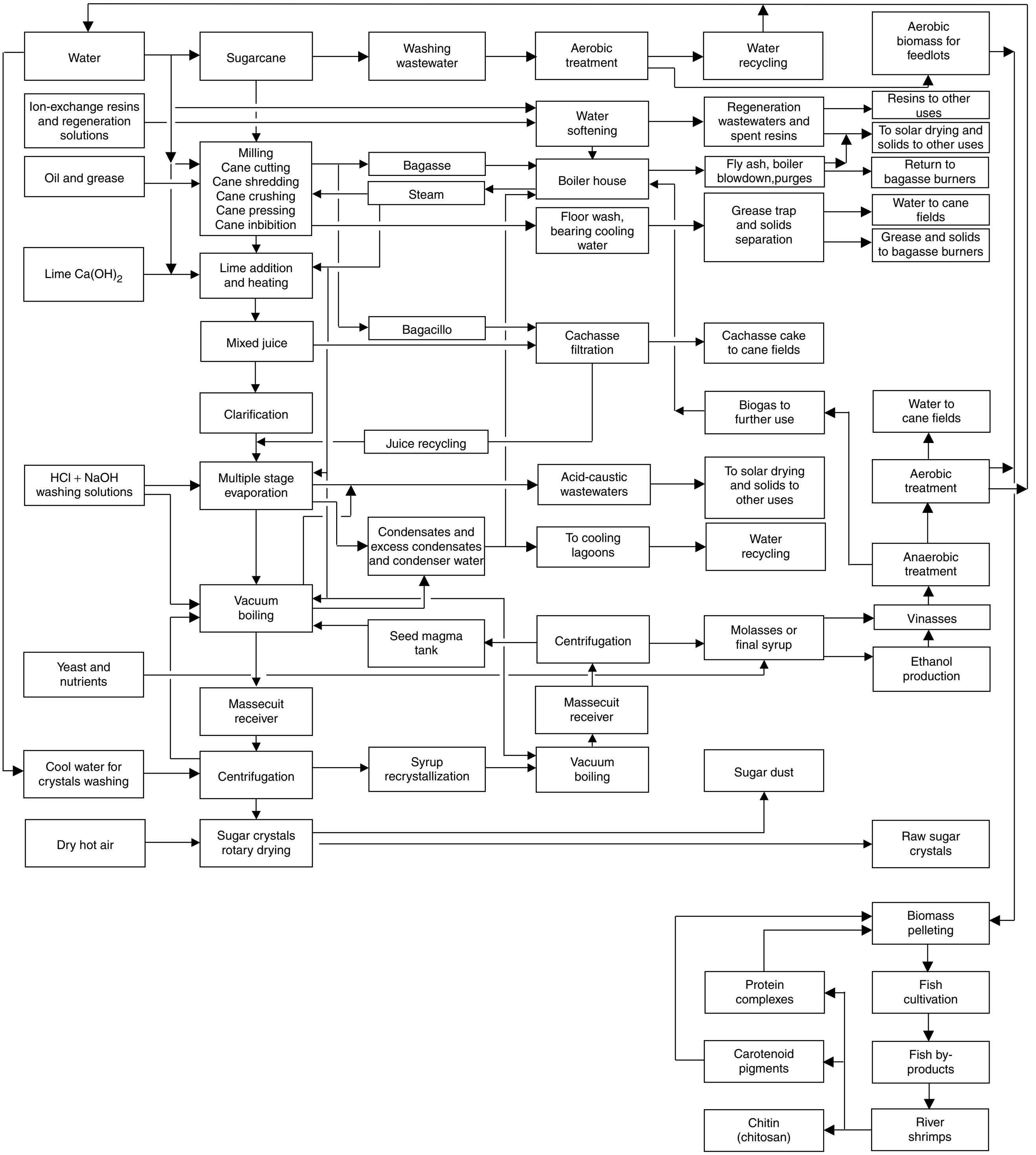

Figure 3. Example of water use, wastewater generation, and treatment proposals for raw sugar.

happens with boilers purges and acid-caustic solutions used to clean heat transfer surfaces in evaporators and vacuum boiler pans, the most suitable process is to eliminate water by evaporation (preferably solar evaporation, taking advantage of climatic conditions in tropical areas). The resulting impure salts, collected as dry solids, may be recycled by the companies that sell the caustic and acid products used as raw materials in the sugarcane mill, reducing the environmental impact of both enterprises (2).

Figures 5 and 6 present schemes of an aerobic plant for treating wastewaters generated in a plantation sugar production plant (14). This plant wastewater treatment system had a different approach. The wastewater

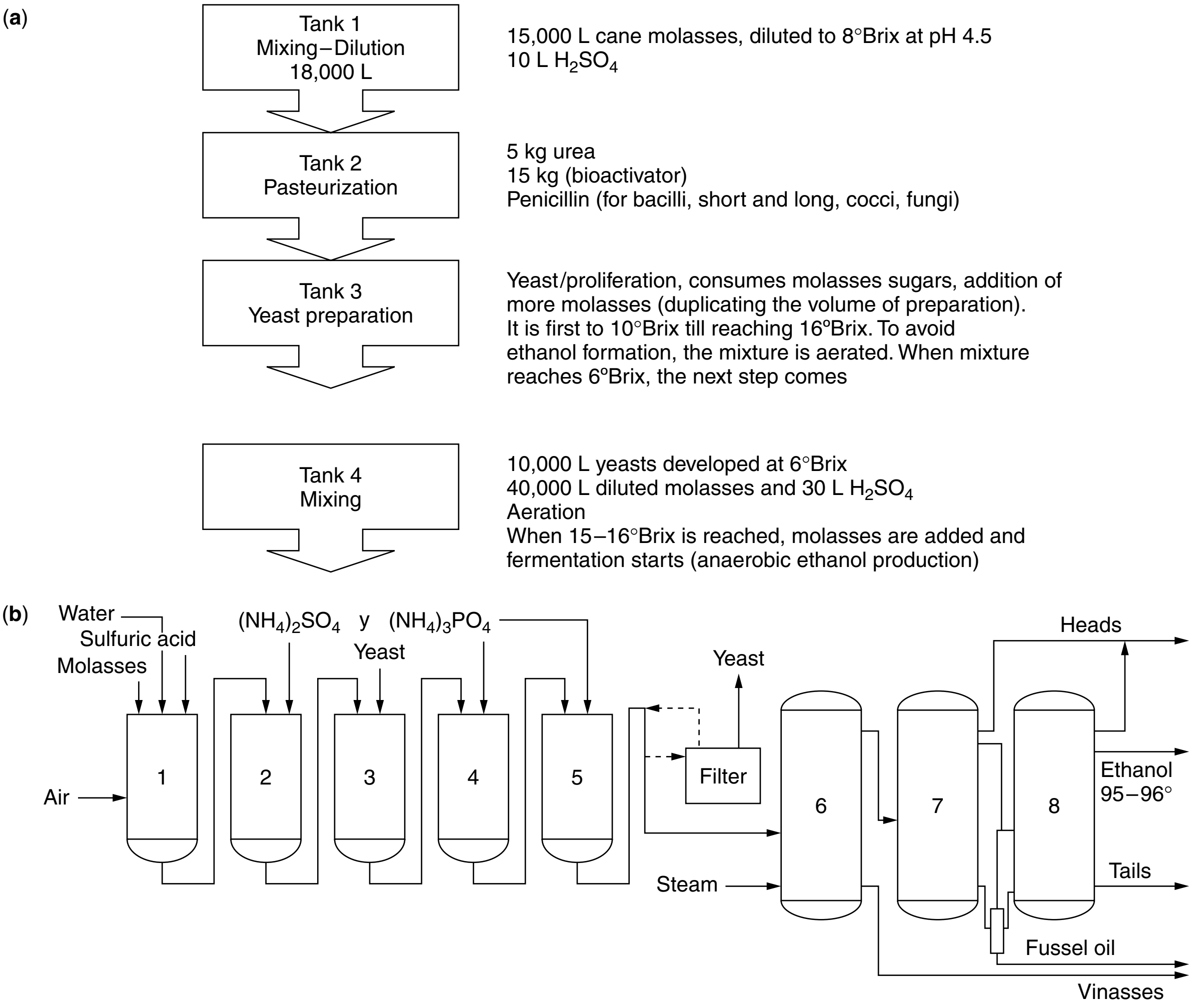

Figure 4. (**a**) Sequential steps for ethanol production in a typical Mexican sugarcane mill using the final syrup or molasses as a carbon source (4). (**b**) Ethanol production flow in a typical Mexican sugarcane mill using the final syrup or molasses as a carbon source (4).

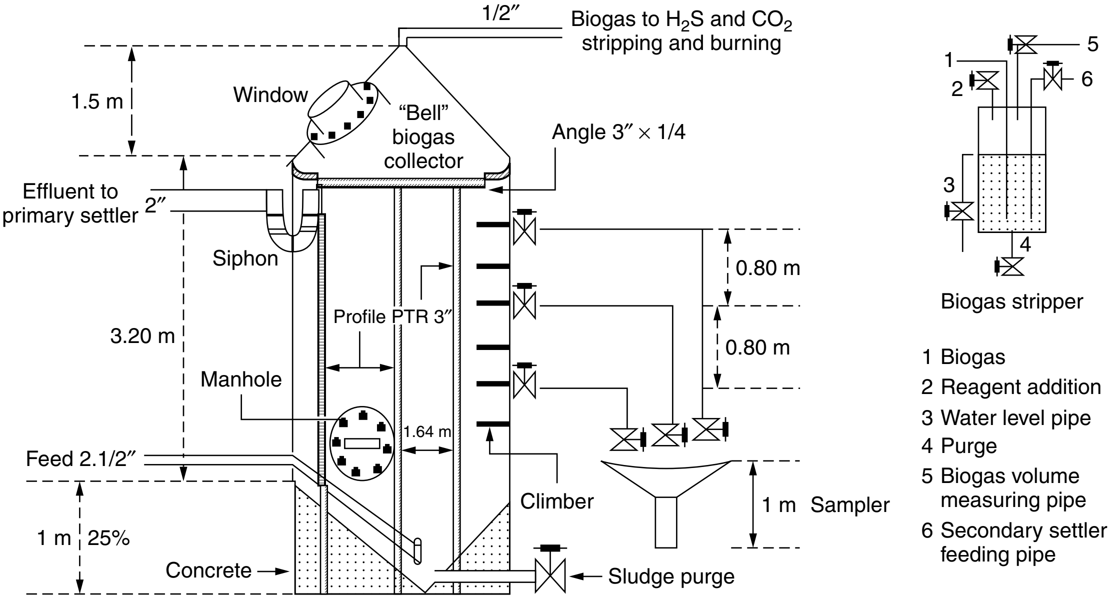

Figure 5. Upflow anaerobic sludge bed reactor (UASB) and biogas absorber tower (13).

Table 2. Liquid Effluent Characteristics of Some Cane Sugar Processing Plants in the World

Parameter	Puerto Rico[a]	Hawaii[a]	Philippines[a]	Louisiana[a]	India[a]	Mexico[b]
pH	5.3–8.8	n.r.[d]	5.3–7.9	n.r.	6.8–8.4	6–10
Biochemical oxygen demand (BOD_5), mg/L	112–225	115–699	130–1220	81–562	67–660	20–36,700[c]
Chemical oxygen demand (COD), mg/L	385–978	942–2340	50–1880	720–1430	890–2236	47–176,635[c]
Total suspended solids, mg/L	500–1400	3040–4500	n.r.	409	792–2043	20–46,190
Total nitrogen, mg/L	n.r.	n.r.	n.r.	n.r.	n.r.	0.2–1260
Total phosphorus, mg/L	n.r.	n.r.	n.r.	n.r.	n.r.	0.2–2000
Grease and oil, mg/L	n.r.	n.r.	n.r.	n.r.	n.r.	0–570

[a] Reference 6.
[b] Reference 7.
[c] When vinasses are considered.
[d] n.r.: not reported.

composition is similar to those of other countries that do not contain vinasses (Table 2), so once the suspended matter is eliminated in a primary settler, an aerobic reactor is adequate for removing most of the dissolved biodegradable pollutants. Here, the feasibility of using an anaerobic reactor to "digest" the aerobic biomass when it is not used for feedlots is shown (13). This reactor gives the added value of methane-rich biogas production, once the gas is stripped in a column using the treated water to dissolve H_2S leaving the methane-rich gas free of this corrosive compound and "enriching" the treated wastewater in sulfur compounds before sending it to cane fields as irrigation water.

Table 3. Average Vinasse Compositions[a]

Parameter		Parameter	
pH	4.8–4.9	Total volatile solids, g/L	76–83
Biochemical oxygen demand, mg/L	90,000	Total fixed solids, g/L	21–23
Chemical oxygen demand, mg/L	106,000	Dissolved volatile solids, g/L	69–76
Organic nitrogen, mg/L	446	Dissolved fixed solids, g/L	18–22
Total nitrogen, mg/L	730		
Ammonium ion, mg/L	310	Phosphorus, mg/L	150
Grease and oil, mg/L	2	Calcium, mg/L	2,960
Settleable solids, mL/L	27	Magnesium, mg/L	1,370
Total solids, g/L	97–106	Sodium, mg/L	310
Total suspended solids, g/L	8–10	Potassium, mg/L	2,550
Total dissolved solids, g/L	87–97	Sulfate ions, mg/L	10,500

[a] Reference 3.

BIBLIOGRAPHY

1. Chaux, D., Durán-de-Bazúa, C., and Cordovés, M. (1991). *Towards a Cleaner and More Profitable Sugar Industry*. UNIDO, Vienna, Austria.
2. Durán-de-Bazúa, C., Cordovés, M., and Zedillo, L.E. (1994). *Demonstration of Cleaner Production Techniques for the Sugar Cane Agroindustry. Final Draft, Consultancy Report 1994.* United Nations Industrial Development Office. Project US/INT/91/217/15-01-2. UNIDO/UNDP, Mexico City, Mexico.
3. Jiménez, R., Martínez, M., Espinoza, A., Noyola, A., and Durán-de-Bazúa, C. (1995). *La Caña de Azúcar, su Entorno Ambiental*. Parte II. Tratamiento de vinazas en una planta piloto en México en un reactor anaerobio de lecho de lodos. Informe técnico de proyecto VIN-01-95. GEPLACEA, CNIIAA, PIQAYQA-FQ-UNAM. Facultad de Química, UNAM, México D.F., México.
4. Castro-González, A., Bernal-González, M., and Durán-de-Bazúa, C. (2004). Tratamiento de vinazas de plantas

Table 4. Regulatory Limits Established in Mexico for Cane Sugar and Ethanol Production Plant Liquid Effluents[a]

	Cane Sugar		Distilling Industry	
Parameter	Daily Average Value	Instantaneous Value	Daily Average Value	Instantaneous Value
pH	6–9	6–9	6–9	6–9
BOD_5, mg/L	60	72	200	240
COD, mg/L	n.r.[b]	n.r.	260	360
Total suspended solids, mg/L	n.r.	n.r.	200	240
Settleable solids, mg/L	1.0	1.2	1.0	2.0
Total nitrogen, mg/L	n.r.	n.r.	10	12
Total phosphorus, mg/L	n.r.	n.r.	5	6
Grease and oils, mg/L	15	20	10	20
Phenols, mg/L	0.5	0.75	n.r.	n.r.

[a] Reference 8.
[b] n.r. not reported

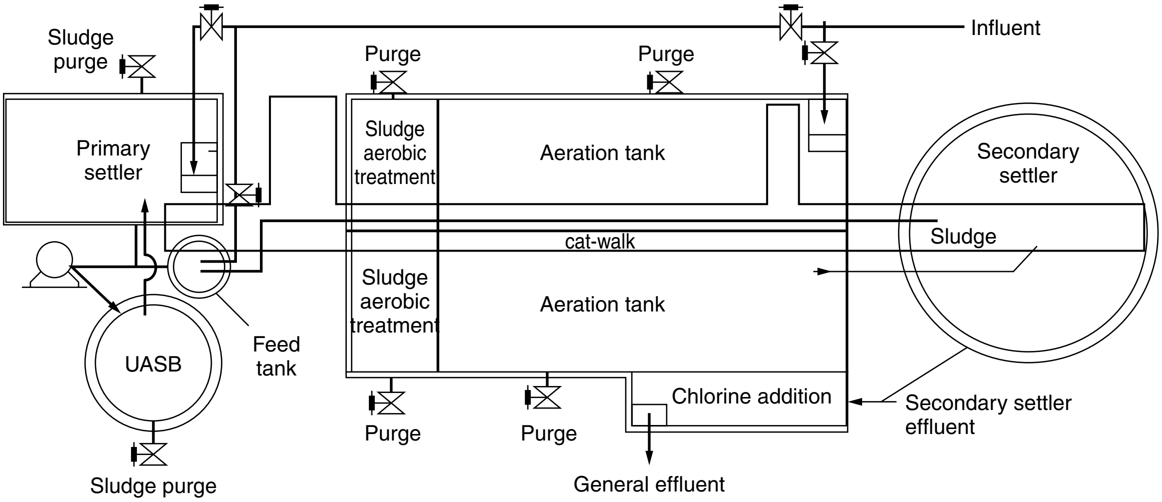

Figure 6. Complete sugarcane mill wastewater treatment plant, including a primary settling tank, an anaerobic feed tank, a UASB reactor, and a biogas absorption tank (biogas stripping) (14).

destiladoras de alcohol usando consorcios microbianos anaerobios. *Bebidas Mexicanas* **13**(3): 12–14, 16–20, 22–25.

5. Durán-de-Bazúa, C. (2004). *Tratamiento Biológico de Aguas Residuales Industriales*. PIQAyQA-Facultad de Química, UNAM, 5th Edn. (1994). México, D.F. Mexico. 6th Edn., in press.
6. Bao-Guo-Yu and Chen-Shi-Zhi. (1990). *Treatment and Utilization of Pollution Effluents of Cane Sugar Factories in China*. UNIDO. Workshop on pollution control and low waste technologies in agro-based industries. Shanghai, China.
7. Anonymous. (1976). *Uso del Agua y Manejo Del Agua Residual en la Industria*. Vol. 8. Azúcar. Secretaría de Recursos Hidráulicos. México, D.F. Mexico.
8. DOF. (1993, 1995). *Normas Oficiales Mexicanas en Materia de Protección Ambiental*. NOM-002-ECOL-1993, Proy. NOM-064- ECOL-1995. Diario Oficial de la Federación, Secretaría de Desarrollo Social (Sedesol), Mexico City, Mexico.
9. Bautista-Zúñiga, F., Reina-Trujillo, T. de J., Villers-Ruiz, L., and Durán-de-Bazúa, C. (2000). *Mejoramiento de Suelos Agrícolas Usando Aguas Residuales Agroindustriales*. Caso: Vinazas crudas y tratadas. Serie: Química Ambiental del Suelo. Vol. 1. PIQAyQA-FQ, UNAM. México D.F., México.
10. Bautista-Zúñiga, F., Durán-de-Bazúa, C., Reyna-Trujillo, T., and Villers-Ruiz, L. (2000). Agroindustrial organic residues: Handling options in cane sugar processing plants. Part I and Part II. *Sugar y Azúcar* **95**(9): 32–45; **95**(10): 23–37.
11. Villatoro-Reséndiz, J. (1998). *Estudio de la Transformación de la Materia Orgánica Biodegradable de la Vinaza Cruda y Tratada en Los Suelos Acrisol y Vertisol, Del Municipio Miguel Alemán en el Estado de Veracruz, México*. Tesis profesional (Biología). UNAM, Facultad de Ciencias, México D.F., México.
12. Durán-Domínguez, M.C., Pedroza-Islas, R., Rosas-Vázquez, C., Luna-Pabello, V.M., Sánchez-Zamora, A., Capilla-Rivera, A., Paredes-Gómez, L., Valderrama-Herrera, S.B., and Vázquez-Cedeño, I. (1991). Producción de alimentos para peces: Utilización de subproductos del tratamiento de aguas residuales. In: *Premio Nacional Serfín El Medio Ambiente*. J.J. De Olloqui (Ed.). Futura Eds., Méx., México, pp. 79–106.
13. Castro-González, A., Enríquez-Poy, M., and Durán-de-Bazúa, C. (2001). Design, construction, and starting-up of an anaerobic reactor for the stabilisation, handling, and disposal of excess biological sludge generated in a wastewater treatment plant. *Anaerobe (Biotechnology)* **7**: 143–149.
14. Castro-González, A., Durán-de-Bazúa, M.C., Enríquez-Poy, M., and Pliego-Bravo, Y. (1998). Consideraciones en un ingenio azucarero para minimizar las descargas líquidas y mejorar las condiciones de operación de su planta de tratamiento de aguas residuales. *Ingenio* (México) **3**(34): 2–7.

ESTIMATED USE OF WATER IN THE UNITED STATES IN 1990 INDUSTRIAL WATER USE

U.S. Geological Survey

Industrial water use includes water for such purposes as processing, washing, and cooling in facilities that manufacture products. Major water-using industries include, but are not limited to, steel, chemical and allied products, paper and allied products, and petroleum refining.

Many States have developed permit programs that require reporting of industrial withdrawals and return flows. Estimates for 1990 are improved over those of previous years because of the availability of more comprehensive inventories of industrial facilities and more complete water-use records. Information on deliveries from public suppliers to industrial users were estimated from a variety of methods if not available directly from public suppliers. Consumptive-use estimates generally were based on coefficients ranging from 3 to 80 percent (depending on the type of industry) of withdrawals and deliveries.

Industrial water use (freshwater withdrawals, public-supply deliveries, saline water withdrawals) during 1990 was an estimated 19,300 Mgal/d of self-supplied

---

Table 1. Industrial Water Use By Water-Resources Regions [Figures May Not Add to Totals Because of Independent Rounding. All values in Million Gallons Per Day]

	Self-Supplied Withdrawals							
	By Source and Type							
	Ground water		Surface water		Total			Reclaimed Waste Water
Region	Fresh	Saline	Fresh	Saline	Fresh	Saline	Total	
New England	96	0.0	382	68	479	68	547	
Mid Atlantic	361	.2	1,370	1,470	1,730	1,470	3,200	6
South Atlantic-Gulf	896	0	1,920	94	2,810	94	2,910	
Great Lakes	235	3.7	3,950	0	4,190	3.7	4,190	
Ohio	532	0	1,840	0	2,370	0	2,370	
Tennessee	23	0	1,170	0	1,190	0	1,190	
Upper Mississippi	349	0	618	0	967	0	967	
Lower Mississippi	501	.6	2,120	67	2,620	67	2,690	
Souris-Red-Rainy	1.3	0	47	0	49	0	49	
Missouri Basin	114	0	57	0	171	0	171	
Arkansas-White-Red	67	0	301	0	368	0	368	
Texas-Gulf	141	1.1	600	1,460	741	1,460	2,200	2
Rio Grande	11	0	1.0	0	12	0	12	
Upper Colorado	2.9	0	2.5	0	5.4	0	5.4	
Lower Colorado	49	0	124	0	174	0	174	
Great Basin	77	2.3	29	0	106	2.3	108	
Pacific Northwest	336	0	691	36	1,030	36	1,060	
California	126	0	4.8	25	130	25	156	
Alaska	5.2	0	106	0	111	0	111	
Hawaii	20	.6	23	0	43	.6	44	
Caribbean	11	1.2	0	50	11	51	62	
Total	3,950	9.7	15,400	3,260	19,300	3,270	22,600	9

Table 2. Industrial Water Use By States [Figures May Not Add to Totals Because of Independent Rounding. All values in Million Gallons Per Day]

	Self-Supplied Withdrawals							
	By Source and Type							
	Ground water		Surface water		Total			Reclaimed Waste Water
State	Fresh	Saline	Fresh	Saline	Fresh	Saline	Total	
Alabama	31	0.0	753	0.0	784	0.0	784	0.0
Alaska	5.2	0	106	0	111	0	111	0
Arizona	39	0	124	0	163	0	163	2.3
Arkansas	99	0	78	0	177	0	177	0
California	125	0	3.4	25	129	25	154	.8
Colorado	33	0	85	0	118	0	118	0
Connecticut	19	0	61	68	80	68	148	0
Delaware	18	0	47	6.0	65	6.0	71	0
D.C.	.5	0	0	0	.5	0	.5	0
Florida	282	0	121	56	403	56	459	0
Georgia	346	0	311	33	657	33	689	.5
Hawaii	20	.6	23	0	43	.6	44	0
Idaho	170	0	26	0	196	0	196	0
Illinois	155	0	309	0	464	0	464	0
Indiana	129	0	2,350	0	2,480	0	2,480	0
Iowa	71	0	148	0	219	0	219	0
Kansas	50	0	3.8	0	53	0	53	.5
Kentucky	93	0	220	0	313	0	313	0
Louisiana	289	.6	2,070	67	2,360	67	2,430	0
Maine	9.8	0	244	0	254	0	254	0

Table 2. (*continued*)

	Self-Supplied Withdrawals							
	By Source and Type							
	Ground water		Surface water		Total			Reclaimed Waste Water
State	Fresh	Saline	Fresh	Saline	Fresh	Saline	Total	
Maryland	21	0	49	379	70	379	450	63
Massachusetts	65	0	22	0	87	0	87	0
Michigan	175	3.7	1,510	0	1,680	3.7	1,690	0
Minnesota	65	0	89	0	154	0	154	0
Mississippi	144	0	126	0	269	0	269	0
Missouri	53	0	32	0	85	0	85	0
Montana	30	0	27	0	57	0	57	0
Nebraska	39	0	2.4	0	41	0	41	0
Nevada	9.4	0	.8	0	10	0	10	0
New Hampshire	.3	0	37	0	37	0	37	0
New Jersey	53	.2	273	1,020	326	1,020	1,340	0
New Mexico	4.6	0	1.7	0	6.3	0	6.3	0
New York	85	0	189	0	274	0	274	0
North Carolina	63	0	328	5.5	390	5.5	396	0
North Dakota	2.	0	6.6	0	8.8	0	8.8	0
Ohio	123	0	230	0	354	0	354	0
Oklahoma	3.3	0	32	0	35	0	35	0
Oregon	31	0	254	0	284	0	284	1.6
Pennsylvania	180	0	1,690	0	1,870	0	1,870	0
Rhode Island	2.5	0	9.1	0	12	0	12	0
South Carolina	47	0	585	0	632	0	632	0
South Dakota	5.0	0	10	0	15	0	15	0
Tennessee	69	0	813	0	882	0	882	0
Texas	143	1.1	741	1,460	884	1,460	2,340	22
Utah	77	2.3	29	0	106	2.3	108	0
Vermont	1.0	0	43	0	44	0	44	0
Virginia	195	0	300	66	495	66	561	0
Washington	104	0	397	36	501	36	536	0
West Virginia	106	0	26	0	132	0	132	0
Wisconsin	58	0	409	0	468	0	468	0
Wyoming	6.0	0	9.9	0	16	0	16	0
Puerto Rico	11	0	0	0	11	0	11	0
Virgin Islands	.1	1.2	0	50	.1	51	51	0
Total	3,950	9.7	15,400	3,260	19,300	3,270	22,600	90

freshwater, 5,190 Mgal/d of public-supplied freshwater, and an additional 3,270 Mgal/d of saline water. (See Table 1: water-resources regions and Table 2: States.) Industrial freshwater use during 1990 was 13 percent less than during 1985 and represents 7 percent of total freshwater use for all offstream categories. Surface water was the source for about 82 percent of self-supplied industrial withdrawals; ground water, 18 percent; and reclaimed wastewater, only a fraction of 1 percent. Public-supplied deliveries to industries accounted for 13 percent of public-supply withdrawals.

The source and disposition of water for industrial purposes are shown in the pie charts below (or as a GIF file or PostScript file (94 Kb)). The consumptive use of freshwater for industrial purposes during 1990 was 3,330 Mgal/d, or 14 percent of freshwater withdrawals and deliveries; saline consumptive use was 913 Mgal/d, or about 28 percent of saline water withdrawals.

In 1990, the Great Lakes and Mid Atlantic water-resources regions had the largest withdrawals for industrial purposes as shown in figure 20 (GIF file), or (PostScript file (620 Kb)). Indiana, Louisiana, Texas, Pennsylvania, and Michigan reported the largest state withdrawals for industries as shown in figure 21 (GIF file) or (PostScript file (508 Kb)). Indiana, Louisiana, Pennsylvania, and Michigan reported the largest freshwater use (figure 22 (GIF file)), or (PostScript file (508 Kb)), and Maryland and Texas reported the largest quantities of reclaimed wastewater used by industries.

WASTE WATER TREATMENT

AERATION

Arnim R.H. Hertle
GHD Pty Ltd
Wembley, Washington
Australia

OVERVIEW

The aeration of wastewater and byproducts is a key process in the operation of most modern wastewater treatment plants (WWTP), which is reflected in the energy consumption of a WWTP's aeration system, which can be up to 70% of the WWTP's total energy consumption.

As most wastewater treatment plants and many sludge treatment plants include aerobic biological processes, the transfer of oxygen into the wastewater (or sludge) is a key operation.

As a result of the large quantities of CO_2 produced in the course of the aerobic degradation of organic matter, the pH in such a reactor can be significantly lower than that of its feed. These CO_2-related effects can be amplified when nitrification occurs or when the liquor's alkalinity is low. This drop in pH can lead to decreased performance of the plant as well as to impacts such as concrete and metal corrosion if not managed properly. The second key function of an aeration system is therefore to remove CO_2 from the reactor (1).

One of the important functions that an aeration system has to provide is mixing to prevent settling of the biomass in the reactors and optimize the contact of wastewater with the biomass.

TERMS AND DEFINITIONS

The following terms and definitions will be used herein:

1. α (alpha) factor is the ratio of the apparent volumetric mass transfer coefficient kL in wastewater to that in clean water. α is normally <1.
2. Oxygen transfer rate (OTR) is the mass of oxygen dissolved by a given aeration system in a given volume of water under given water temperature, barometric pressure, and operating conditions [initial dissolved oxygen (DO) concentration, gas flow rate, power input] per unit time.
3. Standard oxygen transfer rate (SOTR) is the OTR in clean water at 20°C, whose initial DO concentration is zero in all parts of the waterbody, and is at 101.3 kPa barometric pressure.
4. Oxygen transfer efficiency (OTE) is the fraction of oxygen in an injected gas stream that is transferred into water at given initial DO, gas flow rate, water temperature, and barometric pressure.
5. Standard oxygen transfer efficiency (SOTE) is the OTE in clean water at 20°C, whose initial DO concentration is zero in all parts of the waterbody, and is at 101.3 kPa barometric pressure. The oxygen transfer efficiency per meter submergence for a given aeration system under given ambient and operating conditions is relatively constant for installation depths in the technically most important range around 4 to 5 meters. It is, therefore, often quoted in percent per meter submergence to allow the user of this information to calculate the OTE or SOTE case-by-case for the respective aerator installation depths rather than providing a long list of values for different submergences.
6. Aeration efficiency (AE) is the OTR divided by the total power input for the aeration system, measured as either brake power or wire power.
7. Standard aeration efficiency (SAE) is the AE based on the SOTR.

Further explanations and definitions of further parameters and terms used to characterize aeration systems and their design are provided in Refs. 2 and 3.

AERATION SYSTEMS

Numerous aeration mechanisms are used in wastewater treatment plants; each has its particular advantages. The oxygen transfer from air (or pure oxygen where it is applied) into liquor can take place only through a common surface of the two media; so the common goal of all aeration mechanisms is to make as much active surface area available as possible. A classification of the technical systems used for aeration purposes can be based on the mechanical equipment that is used. Alternatively, the systems can be classified on the basis of where the oxygen transfer mainly takes place, which indicates how the design calculations are carried out. Aeration systems are therefore differentiated into

- — surface aeration and
- — submerged aeration.

SURFACE AERATION

Natural Aeration

The specific characteristic of natural aeration is that no energy is directly applied and/or no mechanical equipment is directly involved in the oxygen transfer. Examples are

- — the oxygen transfer through the water surface of a facultative wastewater treatment pond;
- — the aeration of wastewater in a trickling filter where natural draft is relied on to move the air through the filter when it is brought into contact with the wastewater; or
- — postaeration of secondary effluent in a cascading system before it is discharged into the receiving waters.

The first two types of systems are typically designed on the basis of specific loads of biochemical or chemical oxygen demand (BOD or COD) and/or nitrogen (e.g., in kg BOD/(ha · d)) without explicitly designing the aeration process per se. Depending on the system, the contribution of algae to the oxygen balance in a pond may also be considered in the design calculation. These systems are not discussed here.

The secondary effluent can be postaerated by natural aeration in a cascade, although other systems such as fine bubble diffused air aeration can also be applied. The key design parameter of a cascade postaeration system is its height and the shape of the cascade, for example, a weir or steps. The required height is calculated based on the oxygen deficit ratio, the ratio of the difference between the saturation concentration and the actual oxygen content to the difference of the saturation concentration and the target effluent concentration. As the saturation concentration decreases with increasing water temperature, these systems typically have to be designed for the highest temperature when the deficit ratio and hence the required cascade height is greatest.

Mechanical Surface Aeration

An aeration system is referred to as mechanical surface aeration system when air is entrained in the liquor at the surface by moving mechanical equipment. These systems are usually operated with uncompressed ambient air. Low- and high-speed aeration turbines, rotating brush aerators, and aspirating mixer aerators are probably the most commonly used types of mechanical surface aerators.

Aeration turbines and brush aerators work on the principle of lifting some liquid up and accelerating and dispersing it in the air above the liquid surface, which achieves oxygen transfer mainly through the large total surface area of the small liquor droplets in the air and, to a lesser extent, through the turbulent surface of the liquor in the tank and air entrainment when the accelerated liquid falls back down and hits the liquor surface.

Aspirating mixers are high-speed mixers with a hollow inclined or vertical shaft. They use the vacuum created by the rapid motion of the liquid at the mixer tip to suck ambient air through the hollow mixer shaft into the zone of rapid liquid movement at the mixer tip and in the mixer's jet stream. The high shear forces in this area result in mixing the two streams.

Application of Mechanical Aeration Systems

Mechanical surface aeration systems are typically preferred if one or more of the following criteria are met:

- Shallow reactor. Diffused air aeration systems are inefficient here; their efficiency is a function of depth, whereas the efficiency of mechanical aeration systems is almost unaffected by depth. Note that, although this criterion of shallowness is often met in lagoons and mechanical aerators are, therefore, often used in lagoons, it would be wrong to deduce a genuine requirement to aerate lagoons with mechanical aerators because diffused air aeration systems are available that are specifically designed for lagoons.
- Very deep reactor. Fine bubble diffused air aeration is not preferred in such installations because of their reduced CO_2-removal efficiency. Instead, systems that have a lower oxygen transfer efficiency such as coarse bubble diffused air aeration or mechanical aeration are preferred for deep tanks (see below).
- Reactor and process configuration lend itself for it. A typical example is oxidation ditches. These reactor/process layouts typically benefit from the combined aerating and mixing action of mechanical, rotating-brush surface aerators, whereas diffused air aeration systems would require additional horizontal mixers to provide sufficient velocity of mixed liquor circulation.
- Simplicity is imperative. The installation of mechanical aeration systems comprises less components and is significantly simpler than that of diffused air aeration systems that have associated benefits regarding maintenance, for example, that all equipment can be easily accessed from the surface.
- Medium unsuitable for fine bubble diffused air aeration. A typical example would be the aeration of sludge, where fine bubble diffused air aeration loses its efficiency advantage due to effects of the medium, most of all bubble coalescence. Systems that produce coarser bubbles, such as coarse bubble diffused air aeration or mechanical aeration, lose their respective disadvantage, and other benefits, such as simplicity, become the driver for equipment selection. Other examples would be wastewaters that have a high tendency to scale up or foul the diffusers or one that is aggressive to the rubber or other compounds used in the diffusers, all of which can render their use impracticable and be a driver toward mechanical aeration.

The biggest drawback of mechanical aeration compared with fine bubble diffused air aeration is its lower AE in medium deep tanks of around 4 m depth and deeper. In particular, in medium and large municipal installations, this difference in AE often outweighs the advantages of mechanical aeration in capital and maintenance costs and leads to lower life cycle costs for diffused air aeration. Typical AE for mechanical aeration systems is in the range of 1.3–2 kg O_2/kWh. Note that fine bubble diffused air aeration systems may end up in the 2 kg O_2/kWh range as well after a few years of operation, which can lead to process problems. Mechanical aeration systems, in contrast, will stay on their initial AE level, and the oxygen input is at a point source; thus a tapered DO profile can be achieved that can enhance nitrogen removal.

SUBMERGED AERATION SYSTEMS

Submerged aeration systems can be differentiated into

- mechanical submerged aeration systems and
- diffused air aeration systems.

Mechanical Submerged Aeration Systems

A mechanical submerged aeration system entrains ambient air or compressed air in the liquor by moving mechanical equipment at a point close to the tank floor. The concept is to increase the contact time between air and liquor, and hence the oxygen transfer, by letting the air bubbles travel through the liquor and to attempt to produce fine bubbles via agitation. Mechanical submerged aeration systems comprise floor-mounted aspirating or sparger (when operated with compressed air) mixers and ejector pumps.

Aspirating and sparging mixers are high-speed, submerged mixers that introduce air at or below the rotor where it is dispersed and mixed with the liquid, as shown in Fig. 1. In the aspirating mode, the working principle is as described before for aspirating mixers for surface aeration. However, in some cases, the oxygen demand requires a supply of compressed air (sparger mixer). Figure 2 shows an installation of a submerged aspirating mixer.

Ejector pumps use a venturi nozzle on the pump's discharge to entrain ambient air or compressed air into a jet of pumped liquid. They work with the vacuum created by the acceleration of the liquid in the nozzle, which sucks air into this highly turbulent zone where the two streams are mixed.

Diffused Air Aeration

Diffused air aeration is a system where air (or pure oxygen) is compressed and then diffused into small bubbles by a passive component.

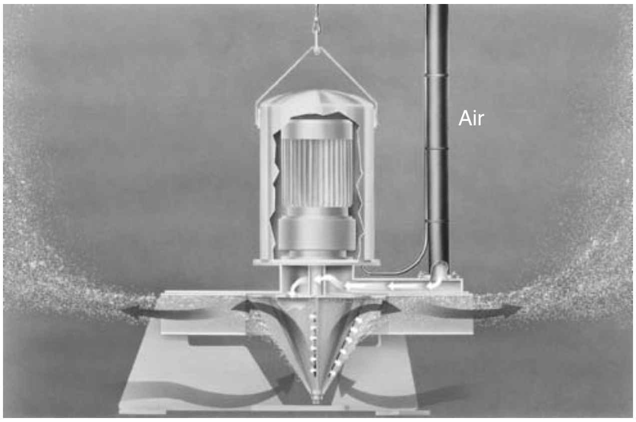

Figure 1. Working principle of a submerged aerating mixer.

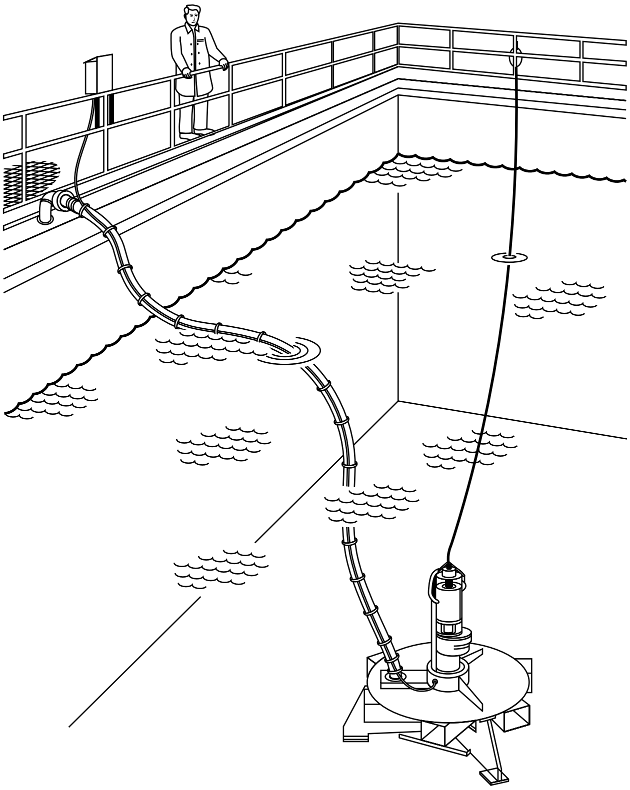

Figure 2. Installation of a submerged aspirating mixer.

Medium or coarse bubble diffusers are being used for applications requiring mixing, in deep tanks, or for the aeration of sludge. They have relatively large openings for the air to pass through. Rubber membranes and rigid porous materials are rarely used for these diffuser types, which include static mixer type arrangements, compressed air jets, perforated pipes, and orifices. Some examples of coarse bubble diffusers are shown in Fig. 3.

Fine bubble systems are used for wastewater aeration applications. Compressed air is passed through floor-mounted, finely slotted rubber membranes or rigid porous diffuser bodies to form small bubbles that rise through the mixed liquor to the surface. An example of a rubber

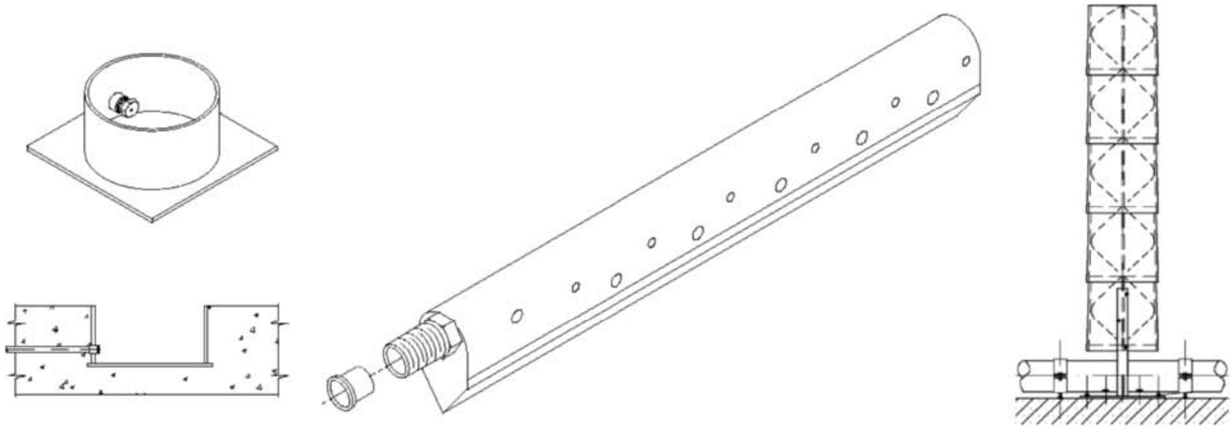

Figure 3. Coarse bubble diffuser types.

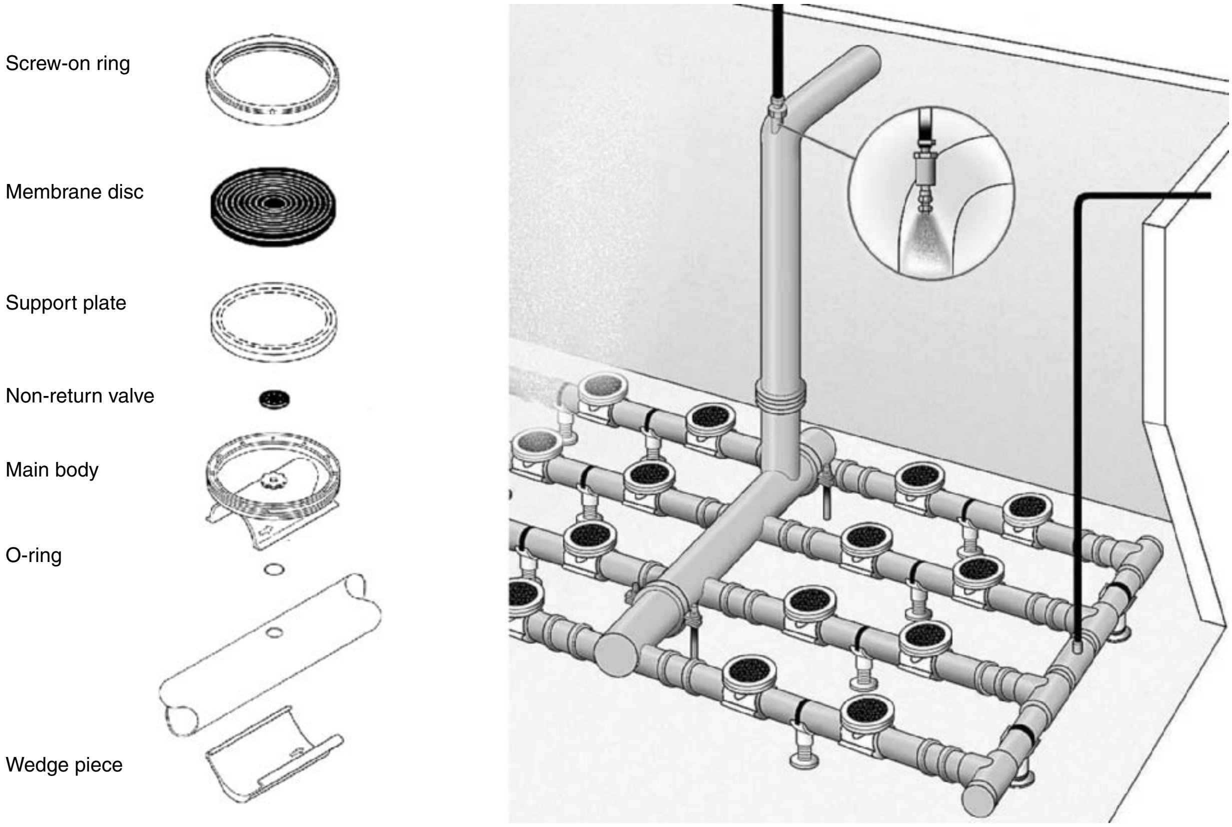

Figure 4. Typical membrane fine bubble diffuser details and installation.

membrane fine bubble diffuser for a fine bubble diffuser installation is shown in Fig. 4.

The mechanical and process design of fine bubble diffused air aeration systems can be quite involved. A dedicated section, therefore, expands on these systems.

BIBLIOGRAPHY

1. Hartley, K.J. (1985). *Operating the Activated Sludge Process*. GHD, Brisbane, pp. 75–77.
2. ASCE Standard No. ANSI/ASCE 2–91. (1993). *Measurement of Oxygen Transfer in Clean Water*, 2nd Edn. American Society of Civil Engineers, New York.
3. Metcalf & Eddy. (2003). *Wastewater Engineering: Treatment and Reuse*, 4th Edn. revised by G. Tchobanoglous, F.L. Burton, and H.D. Stensel, McGraw-Hill, New York, pp. 425–456.

FINE BUBBLE DIFFUSED AIR AERATION SYSTEMS

Arnim R.H. Hertle
GHD Pty Ltd
Wembley, Washington,
Australia

DIFFUSER TYPES

Fine bubble diffusers are typically used for aerating activated sludge tanks in municipal wastewater treatment plants because medium or coarse bubble diffusers are not economical in medium depth tanks, which are prevalent in municipal applications.

Two types of fine bubble diffusers are available on the market:

- — rigid porous diffusers; and
- — elastic/rubber membrane diffusers.

A general comparison of rigid porous and rubber membrane diffuser systems is given in Table 1.

Rigid Porous Fine Bubble Diffusers

The first rigid porous diffusers were made of sintered ceramics such as sintered alumina; many diffusers are still made of this material. Its biggest advantage lies in its chemical inertness. Ceramic diffusers can be refired, for example, in a brick kiln, or acid cleaned to remove scaling and fouling and bring their permeability/head loss back to nearly new condition (1). Practical experience indicates that they need to be refired every 5 to 10 years and that this can be done only two to three times. However, refiring may not often be practical once tank downtime or total cost, including replacement of lost (broken) diffusers, is accounted for.

Alternative materials used to manufacture rigid porous diffusers include sintered HDPE granulate and round quartz sand in a phenol–resin matrix. These diffusers cannot be cleaned thermally because the temperatures

Table 1. Comparison of Rubber Membrane and Rigid Porous Diffusers

	Rubber Membrane Diffusers	Rigid Porous Diffusers
Continuous operation	Well suited	Well suited
Intermittent operation	Reasonably suited	Not suited
Turndown ratio	Large	Moderate
High water hardness	Scaling may lead to short membrane replacement intervals	Carbonate scaling can be controlled with in-line acid dosing
High water temperature	Not suited	Depending on diffuser material
High oil and grease concentrations	EPDM: not suited Silicone: well suited	Well suited
In-line cleaning	Bumping	Acid gas injection
Typical head loss	Higher	Lower
Typical oxygen transfer efficiency	Higher	Lower
Typical asset life	Replace membrane after 3 to 5 years Replacement of membrane brings diffuser back to new condition	Refire or acid clean after 5–10 years Refiring or acid cleaning can bring diffuser back to nearly new condition

in a kiln would destroy them. However, they can be acid cleaned and, due to the more regular and smoother geometry of the particles from which they are made, high-pressure hosing may have a greater effect than on diffusers made from genuine ceramic. The use of a resin matrix also makes it possible to include fouling inhibitors (bactericides) in the matrix and thus reduce bacterial growth on the diffuser and in its pores.

Rigid porous diffusers are manufactured in many shapes and types such as domes, cylinders, circular disks, and plates (noncircular disk). The method of mounting the actual porous diffuser element to the pipework is important in selecting a diffuser to provide flexibility for replacement, which is critical for converting a rigid porous diffuser system to a membrane diffuser system and vice versa, if required. Further, some suppliers offer features such as a nonreturn valve or a pressure sample port that can be integrated into the diffuser carrier. A nonreturn valve prevents intrusion of larger quantities of mixed liquor into the diffuser and the pipework, if the airflow should drop below the required minimum or should cease. A pressure sample port in the diffuser carrier in combination with a pressure probe in the header pipe and a static pressure measurement allows a determination of the dynamic wet pressure of the diffuser to judge whether a cleaning cycle is required.

The lifetime of rigid porous diffusers is determined mainly by

- the lifetime of the bond between the diffuser body particles (disintegration);
- reduction of the pore diameter by scaling (e.g., $CaCO_3$ precipitation);
- reduction of the pore diameter by fouling (intrusion/growth of bacteria); and
- blockage of pores due to dust and other particles in the air.

Due to blockage of its pores, the head loss of a rigid porous diffuser can only increase over its lifetime. At the same time, its oxygen transfer efficiency (OTE) may increase as the pores, and hence the air bubbles, produced become smaller, leading to reduced required air throughput. If they occur, these two effects may offset each other to some degree and may thus expand the useful life of the diffusers. There will, however, be a point where the aeration blowers cannot deliver sufficient quantities of air at the required, increased pressure any longer, and the wastewater treatment process becomes underaerated (oxygen deficient).

Rubber Membrane Fine Bubble Diffusers

Rubber membrane diffusers are made of a finely slotted rubber membrane mounted on a carrier, typically made of plastic. Most rubber membranes are made of EPDM because of its favorable physical properties and acceptable chemical resistance at a low price. However, the drawback of EPDM in wastewater applications is that its plasticizers are washed out over time by oils, grease, and solvents, which are always present in small quantities in wastewater and in compressed air, which in turn causes the membrane to become stiff over time and lose its ability to contract fully and close its pores when the air is cut off. The speed at which this process progresses is a function of various parameters; one of them is the oil and grease content of the wastewater. EPDM membranes are, therefore, unsuitable for wastewaters that have a high concentration of these constituents.

Alternative membrane materials that overcome some of the shortfalls and diffuser membranes are also available now in materials such as silicone rubber and polyurethane rubber. These materials are resistant to oil and grease in wastewater because they do not contain plasticizers.

Note that the resistance to oil and grease is not the only material characteristic where the various rubber types differ. Table 2 shows a number of physical parameters that demonstrate how different these materials are and that, for example, EPDM and silicone are not uniquely specified substances but can be produced to match different physical and chemical requirements. It can also be seen that EPDM varieties have higher tear resistance than the alternative materials; however, once a tear has formed, it will grow longer more easily than in the other materials.

An important aspect that is not captured in Table 2 is how the membranes perform over time ("aging"). EPDM membranes become stiffer, whereas silicone membranes tend to become softer. If not replaced before these effects become noticeable and with other effects,

Table 2. Examples of Diffuser Membrane Materials[a]

Material	EPDM (1)	EPDM (2)	Silicone (1)	Silicone (2)	Polyurethane
Plasticizer content, %	12	18	0	0	0
Tear resistance, MPa[a]	13	14.4	10	10.5	11
Ultimate elongation, %[a]	700	580	650	680	700
Tear strength, N/mm[b]	11.4	10.5	40	44.5	23

[a] Reference 2.

[a] DIN-53504.

[b] Various testing standards.

e.g., scaling, left aside, this different aging process may lead to slightly finer bubbles and an associated increase in OTE for EPDM membranes and slightly larger bubbles and an associated decrease in OTE for silicone membranes.

Due to the different properties and aging patterns of rubber membrane diffusers, compared with rigid porous diffusers, maintenance and cleaning are also different. A wash down with water is a typical means at water pressures often quite low to avoid damaging the membranes. Acid injection is rarely used to clean membranes. Rubber membrane diffusers can be "bumped" to remove scale, which is done by sharply increasing the airflow and then dropping back to a low (or nil) flow. By doing so, deposits on the diffuser surface break and loosen when the rubber membrane expands and are washed away.

The most common shapes of rubber membrane diffusers are circular discs, tubes of circular and noncircular cross section (e.g., cylindrical diffuser elements of a round cross section), and plates (noncircular disk).

Most diffuser types comprise a carrier and a separate membrane mounted on the carrier. Tube types have the diffuser membrane all around the carrier; the flat disk and plate types have a membrane sheet that is attached to the upper side of the carrier. Apart from these three main types of diffusers less frequently used designs also exist, such as diffuser tubes without an internal carrier where the membrane tubes are fixed to a support grid to hold them in place.

The most important difference between rigid porous diffusers and membrane diffusers is that membrane diffusers have built-in backflow prevention properties, which results from the elastic rubber membrane totally contracting when the air is cut off and thereby closing its pores.

Note that the membrane lifts off the carrier when pressurized and returns to its initial position and lies tightly on the carrier when the air is cut off. Further, all diffuser assemblies are designed so that the membrane closes the air entry ports in the carrier when it is resting, which is facilitated by the static head of the liquor above the diffuser. In addition to these nonreturn mechanisms, some disk diffusers include a nonreturn flap in the air entry port of the carrier. These properties make it possible to cut off the air supply to submerged diffusers without flooding the diffusers or the piping, which is a characteristic that is vital for diffused air aeration systems in sequencing batch reactors or in plants that have large diurnal load fluctuations, where diffusers have to be cut off at certain times.

The lifetime of rubber membrane diffusers is determined mainly by

- — operating temperature (higher temperatures, e.g., in tropical regions or in industrial applications, lead to a shorter lifetime);
- — the lifetime/residence time of the membrane plasticizer (leading to decreased membrane elasticity and increased head loss);
- — other aging effects on the rubber, such as UV radiation or oxidation;
- — chemically or biochemically induced hydrolysis of the rubber;
- — loss of pores self-closing properties from scaling (e.g., $CaCO_3$ precipitation) due to reduced elasticity and/or deposits in the pores;
- — reduction of the pore diameter by scaling (e.g., $CaCO_3$ precipitation);
- — reduction of the pore diameter by fouling (intrusion/growth of bacteria); and
- — blockage of pores due to dust and other particles in the air.

Over its lifetime, the head loss of an EPDM rubber membrane diffuser typically first drops when the membrane elongates slightly as a result of the expansion–contraction cycles experienced during operation. Once the head loss has reached its minimum, it can only increase from there on as a result of the above impacts. The OTE of rigid porous diffusers, may increase and lead to reduced air throughput requirements. The increased head loss can reach a point where the aeration blowers cannot deliver sufficient quantities of air at the required pressure. The increased operating pressure of the (then stiffer) diffuser also includes an increased risk of membrane failure.

Mixed Diffuser Systems

It is possible to combine rigid porous diffusers and rubber membrane diffusers in one installation. By doing so, some of the wastewater treatment plant's diffuser installation can benefit from the long lifetime of rigid porous diffusers, and some can benefit from the large turndown ratio of rubber membrane diffusers, for example,

- — rigid porous diffusers in the main nitrification zone and rubber membrane diffusers in a bivalent zone ("swing zone") that can be operated anoxic in summer to maximize nitrogen removal and aerobic in winter to compensate for the lower nitrification rate at lower temperatures by providing more aerobic tank volume and hence longer aerobic sludge age.

- a modified Ludzack–Ettinger process arrangement that has rigid porous diffusers in the main nitrification zone and rubber membrane diffusers in the last aerated zone to minimize aeration there and thus oxygen recycling back into the anoxic zone.
- a plant that has strong diurnal or seasonal load fluctuations where rigid porous diffusers cover the base load and additional rubber membrane diffusers can be switched on to cover the additional oxygen demand during peak hours or peak season.

Such systems have been built and are being operated successfully, but many factors have to be considered. The diffuser types need to be suitable for the particular wastewater, and the grid and control system layout also needs to be adapted. For example, because the head loss for the two diffuser types varies differently when the diffusers are aging, one must be aware that unless the airflow to each subsystem is automatically controlled separately (e.g., air supply valve modulating based on a DO reading) in each tank, the adjustment of the flow split between the two subsystems is a recurrent additional task for plant operators. Other factors are increased piping costs and increased spares storage.

DESIGN CONSIDERATIONS FOR DIFFUSED AIR AERATION SYSTEMS

Oxygen Transfer Efficiency

Aeration systems, in general, and diffused air aeration systems, in particular, are often compared on the basis of their SOTE, which is measured in a laboratory and refers to standardized conditions. Both the testing procedures and the reference conditions are set out in standards, such as the ASCE standard (3), which is probably, internationally, the most frequently used standard for this application.

The SOTE of diffused air aeration systems is a function of the average bubble diameter produced by the diffuser as well as of other parameters that are discussed further below. The dependence on bubble size is demonstrated by the following typical values for the SOTE, quoted in % oxygen transferred per meter of diffuser submergence:

- for coarse bubble systems: 2–3%/m;
- for medium bubble systems: 4–5%/m; and
- for fine bubble systems: 5.5–6.5%/m.

The SOTE can vary among suppliers and materials, and, even for a given diffuser system, it is still widely variable, depending on factors such as

- Specific air throughput per meter of diffuser length or square meter diffuser area: The SOTE decreases with increasing specific air throughput because the average diameter of the bubbles produced increases, which leads to a decreasing surface area to volume ratio for the gas bubbles.
- Floor coverage, the ratio of installed diffuser surface area to tank floor area: The SOTE increases with higher floor coverage and more even diffuser distribution, which is due to the decreased formation of “rolling” motion patterns in the liquor, which are typically observed in line aeration systems, where the liquor is accelerated upward by the rising bubbles from the aeration line (airlift pump principle) and flows back downward in the unaerated areas. These patterns lead to shortened detention time of the bubbles in the water and hence a decreased SOTE.
- strong currents in the basin due to mixers that may coalesce fine bubbles but may also add a horizontal flow component that increases the bubble detention time. The effect on the SOTE can, therefore, be detrimental as well as positive.
- tank depth: The effect of deep tanks on the design of aeration systems is quite complex and is discussed separately later.

Head Loss

Diffused air aeration systems require low head loss to be efficient. However, because the head loss is also a function of bubble size, a balance between head loss and oxygen transfer efficiency must be maintained to provide optimal aeration efficiency.

The increase in head loss from coarse to medium and fine bubble diffusers is usually compensated for by the associated increase in oxygen transfer efficiencies. The overall power requirements and, thus, operating costs in typical municipal wastewater applications are, therefore, typically lower for fine bubble diffusers than for coarser systems. When fine bubble diffusers start to clog, bubble size decreases, and oxygen transfer efficiency and head loss increase. In practice, these two effects may compensate for each other initially; in particular, for rigid porous diffusers, this effect is well documented (4). However, there is a point where the increased head loss cannot be compensated for any longer and the overall efficiency drops or where the blowers cannot any longer deliver the required quantities of air at the increased head.

Other impacts on the head loss of a diffuser are the material from which it is made, its condition (new/old, fouled, scaled, etc.), and the specific air throughput.

Deep Tanks

A tank is generally considered “deep” when its depth is 6 meters or more. The aeration of wastewater in deep tanks requires the designer to account for some physicochemical effects that come into play only in deep tanks (5).

The alpha factor (α) is basically not affected, but the SOTE is affected by

- the higher average static pressure and related higher solubility of oxygen, as can be deduced from Henry’s law (it increases the SOTE).
- the lower average oxygen content (partial pressure) in the bubble because its detention time in the liquor

is longer and hence more oxygen can be transferred from the bubble into the liquor, leading to a higher degree of oxygen depletion in the bubble. This effect decreases the SOTE.

- A potentially lower average bubble size due to higher total mass transfer from the bubble into the liquor because the solubility of air in water increases with pressure and the quantity of oxygen transferred per m^3 of air injected is also increased. This effect can be positive as well as detrimental. If the diffuser would produce relatively large bubbles, it helps to increase the bubble detention time and hence the SOTE. However, if the diffuser already produces small bubbles, the bubble surface becomes "stiffer," and mass transfer may be slower due to a slower surface renewal rate.

Overall, the effect is typically a lower SOTE per meter of diffuser submergence (%/m or $g/(Nm^3 \cdot m)$) compared with shallower tanks at the same specific air throughput and floor coverage. This effect can be compensated for to a certain extent by increasing the number of diffusers to increase floor coverage and decrease specific air throughput; however, this leads to increased capital and maintenance costs.

Published SOTE data typically refer to tank depth of around 3.5 to 5 meters, which is higher than the average SOTE for a deep tank. Care has to be taken when extrapolating from such data; preferably it should be avoided, and the supplier be contacted.

Normal aeration tanks show a pH drop of around 0.3 units depending on initial alkalinity/buffer capacity and the degree of nitrification and denitrification, which will be aggravated in deep tanks due to the increased solubility of CO_2 and reduced aeration air volumes, which in turn reduces the CO_2 stripping capacity of the system. A highly efficient diffuser system is not recommended for a deep tank because the more efficient the system is, the less air is required, and the more CO_2 builds up in the aerated liquor. Deep tanks are, therefore, often aerated by medium or coarse bubble diffusers or mechanical aeration systems because of their lower SOTE and associated higher air throughputs.

Fouling, Scaling, and Clogging of Diffusers

Fouling, scaling, and clogging can affect the performance of diffusers quite significantly. Fouling is mainly a function of how well a biofilm can be established on the diffuser surface. Some systems claim potential advantages of their diffuser system (ability to clean or inhibition of biofilm growth). Fouling and scaling should always be taken into account in the performance of rubber membrane and rigid porous diffused air aeration systems.

Fouling of rigid porous diffusers can be removed in-line by acid cleaning; an acidic gas such as HCl is dosed into the aeration air. However, in particular, in low alkalinity wastewaters, care has to be taken to avoid negative effects on the process due to impacts on the pH. For rubber membrane diffusers, diffusers are normally bumped to remove biofilms that may have formed. High-pressure water cleaning is an effective means of off-line biofilm removal from all diffuser types; ceramic diffusers offer the additional option of refiring.

The two main reasons for diffuser scaling are water hardness and chemical precipitation. The risk and severity of scale formation is easier to assess than fouling. Scaling from water hardness can be removed via acid cleaning and bumping; high-pressure water jetting, however, may not be effective. Scaling from phosphorus precipitation or other precipitation reactions may not lend itself as easily to acid cleaning or refiring but should, in most cases, still be removable to a sufficient extent by diffuser bumping.

Clogging is caused mainly by particles in the compressed air. These particles typically do not dissolve in acids nor can they be quantitatively removed by refiring. Clogging is, therefore, an irreversible process for rigid porous diffusers. Although diffuser bumping would probably push some of these particles through a rubber membrane diffuser, a good practice is to minimize dust and oil in the compressed air by installing appropriate filters.

The need for a cleaning cycle can be determined by measuring the head loss over a diffuser element, which can be done in-line. Measurement of the delivery pressure in the supply air header can also give an indication if temperature effects are considered and the delivery pressures are compared at similar air throughputs.

Note that the "bumping capability" of rubber membrane diffusers decreases when the membrane loses elasticity. Therefore, the effectiveness of bumping to remove fouling, scaling, or clogging decreases over time. Also, the more bumping treatment occurs, the more likely the diffuser is to fail.

Life Cycle Costs

Aeration systems, in general, are a prime example for an asset whose ongoing operating and maintenance costs exceed investment costs. A thorough life cycle cost analysis is critical to decide on the type of aeration system and supplier.

The major component of the operating costs of all aeration systems is the cost of electrical power. For flexible membrane diffuser systems, membrane replacement is a key contribution to life cycle costs. The factors affecting cost are

- the SOTE at the given floor coverage and typical air throughput;
- total replacement costs for membranes, including equipment and labor for installation;
- bankable guarantees on membrane lifetime;
- for existing systems: Effect of the head loss at typical air throughput on blower power consumption and hence on AE.

Other Design Aspects

In applications that have low process oxygen demands (for example, the last aerated section in a MLE process), the air quantities required to supply oxygen may become so small that satisfactory mixing cannot be achieved by the aeration system alone. In these instances, checking the calculated air demands with the air throughputs required for mixing is recommended. Aeration rates of

0.6 to 0.9 m^3/h per m^3 tank volume are suggested (6) for a diffused air aeration system where the diffusers are evenly distributed across the tank floor. For line aeration systems that have a spiral roll liquid motion, the suggested aeration rates are higher at 1.2 to 1.8 m^3/h per m^3 tank volume.

Where diffusers have to be selected for an existing wastewater treatment plant, consideration should be given to conducting on-site trials to determine the alpha factor. If a rubber membrane system is being considered, tests should be conducted to determine the performance of various membrane materials over time.

CONCLUSIONS

Fine bubble diffused air aeration is successfully used on many wastewater treatment plants, regardless of their size. It can provide significant advantages over other aeration systems in energy efficiency. Two main types of diffused air aeration systems are known, rigid porous systems and membrane systems. Of these, the latter offer, by far, better turndown properties. However, rigid porous systems are used more often again due to the longer lifetime of the diffuser element and the lower life cycle cost often associated with that.

Notwithstanding the common use of fine bubble diffused air aeration systems, their successful use depends on choosing the right application and designing the aeration system properly. In some applications, for example, deep tanks or shallow tanks, alternative aeration systems are often to be preferred, as discussed earlier.

BIBLIOGRAPHY

1. Hartley, K.J. and Gobbie, M.R. (1991). Fouling and Cleaning of Fine Bubble Ceramic Dome Diffusers. AWWA Australian Water & Wastewater Association, *Proc. 14th Federal Convention*. Perth, Vol. 1, pp. 109–116.
2. St. Schmidt-Holthausen. (2002). Supratec GmbH, correspondence.
3. ASCE Standard No. ANSI/ASCE 2–91. (1993). *Measurement of Oxygen Transfer in Clean Water*, 2nd Edn. American Society of Civil Engineers, New York.
4. Libra, J.A. et al. (2002). Comparison of the efficiency of large-scale ceramic and membrane aeration systems with the dynamic off-gas method. *Water Sci. Technol.* **46**(4–5): 317–324.
5. ATV. (1996). *Hinweise zu tiefen Belebungsbecken—Arbeitsbericht des ATV-Fachausschusses 2.6*. Korrespondenz Abwasser 6/96, pp. 1083–1086
6. WPCF. (1988). *Aeration—Manual of practice FD 13*. Water Pollution Control Federation, Alexandria, VA.

AIR STRIPPING

Jeff Kuo
Cerritos, California

Air stripping is a physical process that enhances volatilization of compounds from water by passing clean air through it. It is commonly used in treating groundwater contaminated with volatile organic compounds (VOCs) and in removing ammonia in wastewater.

An air stripping system creates air and water interfaces to enhance mass transfer of target compound(s) between the air and liquid phases. Several system configurations, including tray columns, spray systems, diffused aeration, and packed towers, are commercially available. Packed towers (also known as air strippers) are the most popular alternative for groundwater treatment.

A typical air stripper includes a liquid distribution system (e.g., a spray nozzle) at the top of the tower. It sprays influent water over the packing in the column. As the water trickles down, air enters the lower portion of the tower and is forced up through the column by a blower. Packing materials within the tower increase the surface area of the contaminated water that is exposed to air, thus maximizing the migration of VOCs from the liquid stream to the air stream. A sump at the bottom of the tower collects the decontaminated water. The exhaust from the top of the tower, which is commonly referred to as "off-gas," is either discharged to the atmosphere or directed to an off-gas treatment system.

The design of an air stripping system includes determining the diameter and height of the column, type and size of packing materials, air-to-water ratio, blower rating, and treatment of off-gas from the tower. The information needed for the design includes flow rates and quality of the influent water (contaminant concentrations, mineral content, pH, and temperature), quality of the influent air (temperature and contaminant concentrations), height restrictions on the tower; and restrictions on contaminant concentrations in the effluent water and air discharge from the tower.

Although air stripping is widely used to treat VOC-contaminated water, there are several limitations and concerns:

- Air stripping transfers only contaminants from water to air, and the contaminants are not destroyed. Consequently, treatment of the off-gas, which is laden with contaminants, is often necessary.
- Biological and inorganic fouling of packing may require pretreatment or periodic column cleaning.
- Air stripping is effective only for water contaminated by compounds of high vapor pressure and low solubility. Modifications of a conventional air stripping operation may be needed. The temperature of the influent liquid may need to be raised to increase the volatility of the target compound(s). For ammonia stripping, the pH of the water should be sufficiently high, in the range of 10.5 to 11.5, so that most of the ammonia is converted to a gas.
- The visual impacts associated with air strippers and the noise level of the air blower should also be considered. They are often viewed as nuisances when located in residential areas.

Removal efficiencies greater than 95% are typical for towers that have 4.6 to 6 meters (15 to 20 feet) of conventional packing and are removing compounds amenable to stripping, such as B/T/E/X, chloroethane, TCE, DCE, and PCE. When the height of the stripper may be a problem, alternatives such as installation of multiple columns in series, aeration tanks, or low-profile air strippers should be considered. Aeration tanks strip volatile compounds by bubbling air into a tank (diffused aeration) through which contaminated water flows, and their profiles are much lower. The low-profile stripper has a number of trays that are set almost horizontally. Water is cascaded over the trays to maximize air–water contact while minimizing vertical space. Equipment that can be added to the basic air stripper includes an air heater to improve removal efficiencies and automated control systems with sump level switches and safety features, such as differential pressure monitors, high sump level switches, and explosion-proof components.

LAND APPLICATIONS OF WASTEWATER IN ARID LANDS: THEORY AND CASE STUDIES

J.G. Mexal
D.S. Rodriguez
T. Sammis
G. Picchioni
W. Zachritz II
C. Erickson
Z. Samani
New Mexico State University
Las Cruces, New Mexico

Land application of wastewater is recommended by the Environmental Protection Agency as a method of recycling nutrients and organic matter while protecting freshwater resources (1). Wastewater has been applied to agronomic crops, rangelands, forests, and recreation areas, including parks and golf courses, and to disturbed lands such as mine spoils sites (2–5). Benefits from recycling wastewater include creating a supplemental irrigation source, thereby conserving drinking water, and reducing fertilizer costs because wastewater contains nitrogen and phosphorus (6). As wastewater moves through the soil profile, the soil and plants act as filters that trap and treat, through various mechanisms, waste particles and contaminants allowing the remaining water to drain through the soil profile (7,8). Wastewater provides an effective source of nutrients that vegetation roots can absorb and allow the plant to assimilate. The benefit of this system is both the effective remediation of wastes and the recycling of water, nutrients, and carbon via biomass production (4).

Previous studies have demonstrated the effectiveness of saline wastewater application to horticultural systems (9) and to agronomic crops such as cotton (*Gossypium hirsutum* L.), alfalfa (*Medicago sativa* L.), and cantaloupe (*Cucumis melo* L.) (10). Both studies have shown that plants take up and store salt ions in vegetative organs.

Conventional land application systems are accepted technology for many communities, but many communities lack the resources and infrastructure to support them. Any technology considered for implementation should have low capital input and low operating cost. The wastewater treatment system should also stimulate the local economy and be easy to maintain (11). Unfortunately, there is limited information to guide land managers in highly fragile arid and semiarid environments where wastewater may be the only source of supplemental irrigation.

A land application system design, as depicted in (Fig. 1), requires balancing the input of water required by plants against the amount of nitrogen that can be applied without adverse environmental impact (12,13). The wastewater application rate is governed by the total amount of nitrogen allowed to enter the groundwater. Typically, this is expressed as the amount of NO_3-N in the wastewater leached to the groundwater that cannot exceed 10 mg NO_3-N/L. Yearly water (Eq. 1) and nitrogen mass balance (Eq. 2) equations were used to develop the design criteria.

$$L_{w(p)} = \mathrm{PET} - P + W_p \tag{1}$$

where $L_{w(p)}$ = wastewater hydraulic loading rate (cm/mo);
PET = potential evapotranspiration rate (cm/mo);
P = precipitation rate (cm/mo);
W_p = design percolation rate (cm/mo);

$$L_{w(n)} = \frac{[C_p{}^*(\mathrm{P} - \mathrm{ET}) + (U^*4.4)]}{[(1 - \mathrm{f})^* C_n - C_p]} \tag{2}$$

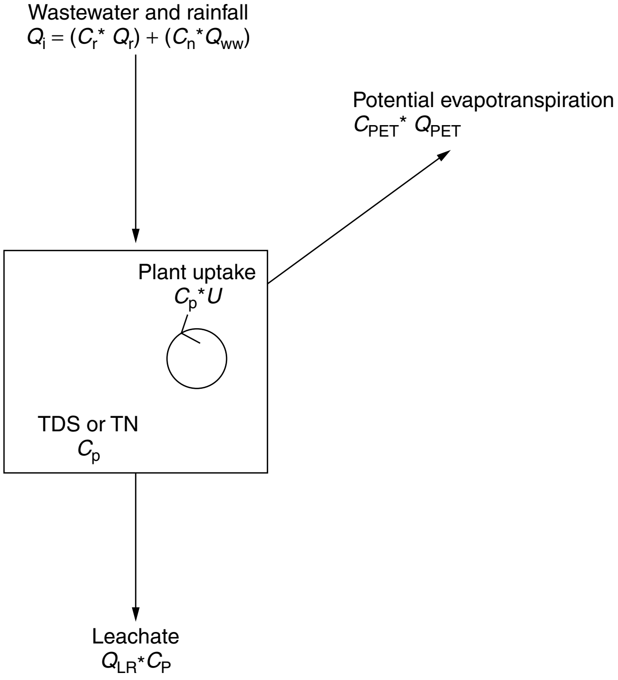

Figure 1. Mass flow diagram of the land application system with inputs, outputs, and process transformations where Q = flow; I = input; r = rainfall; ww = wastewater; LR = leachate; PET = potential evapotranspiration; C = Nitrogen concentration; r = rainfall; n = wastewater; p = percolating water; LR = leacheate; PET = potential evapotranspiration; U = Crop nitrogen uptake; TDS = Total dissolved solids; TN = Total nitrogen (12).

where $L_{w(n)}$ = allowable hydraulic loading rate (cm/yr);
ET = design PET rate (cm/yr);
P = design precipitation rate (cm/yr);
C_p = total nitrogen in percolating water (mg/L);
C_n = total nitrogen in applied wastewater (mg/L);
U = crop nitrogen uptake rate (kg/ha/yr); and
f = fraction of applied total nitrogen removed by denitrification and volatilization

Salt loading, soil salinity and its impact on plant survival and growth, must also be considered in plant selection. Salt loading is not incorporated into the design equations, but guidelines are provided in several sources (13,14). Most plants are intolerant to salt in the seedling stages than in more mature stages, so alternative sources of water may be required during the establishment period. When excess soluble salts accumulate in the rhizosphere, growth is reduced. Salt buildup can be controlled by applying additional irrigation water to leach salts below the root zone (15).

To use the limited amount and potentially saline wastewater, land managers must balance wastewater quality, crop water usage, nitrogen (fertilizer) requirements, and potential soil salinization. Crop coefficients can be calculated from local climate data to estimate evapotranspiration for particular crops from estimates of potential evapotranspiration (PET). Irrigation scheduling models have been developed for crop production in agronomic crops (16).

The following case studies explore the use of saline wastewater for pulpwood, nursery production, and wastewater treatment and disposal in arid regions (Table 1).

Ojinaga, Chihuahua, Mexico

Ojinaga, Chihuahua, Mexico is situated at the confluence of the Rio Grande and Rio Conchos on the U.S.–Mexican border (29.6°N 104.4°W). The climate of this part of the Chihuahuan Desert is arid; the annual rainfall is less than 250 mm, and approximately 75% of the rainfall occurs between June and October; annual potential evapotranspiration is 2450 mm. The maximum temperature is 50 °C, and the minimum temperature is −10 °C.

Municipal sewage is piped directly into an anaerobic lagoon, which acts as a settling pond for separating the solids from the waste stream and providing some reduction in waste strength. A 2-hectare (60,000 m^3) single-cell-anaerobic lagoon constructed in 1995 receives both industrial and municipal wastewater. Currently, one-half of the Ojinaga population is connected to the municipal wastewater system. According to Junta Municipal de Agua y Saneamiento (JMAS), the wastewater from Ojinaga is almost exclusively domestic. Industries in the area include a slaughterhouse and several cottage industries that contribute little in the way of metals or toxic organics. The contribution of the slaughterhouse wastewater is intermittent but high in organic waste. The influent is treated by a manual bar screen to remove large solids. Grit and materials that pass the bar screen are collected in the anaerobic cell by simple sedimentation with no active system for removing these materials. The single-stage anaerobic lagoon is designed to reduce biological oxygen demand (BOD), total suspended solids (TSS), and fecal coliform (FC) to some degree. Wastewater from this system flows to the land application unit for final treatment and disposal.

A pilot land application site (~1.1 ha total area) was established in April 1997 with black locust [*Robinia pseudoacacia* (L.)], clones of eucalyptus [*Eucalyptus camaldulensis* (Dehnh.)] and hybrid poplar [*Populus* (L.)]. Tree species selection was based on the criteria of drought tolerance, salt tolerance, and a short rotation cycle for wood fiber. Monitoring included wastewater effluent and groundwater quality characteristics, soils, and plant growth.

After 6 years, there was minimal overall impact on groundwater, no fecal coliform contamination or increase in NO_3-N, and about a 10% increase in salinity caused by high transpiration of the crop trees. The best tree species was *Eucalyptus* clone 4019. (Fig. 2). The growth rate exceeded the expected results and may allow harvesting on a 6-year rotation, which is 1 to 3 years faster than other studies with similar species. The advantages of commercial forestry include (1) providing a source of income and employment opportunities to the local economy, (2) improving abandoned saline agricultural lands, (3) maintaining irrigation district infrastructure, (4) diversifying the local economy, and (5) establishing an organization to produce wood fiber commercially. This endeavor could produce community-based jobs and have a positive cash flow for operation and maintenance. Ojinaga has been involved in master plan development for the past 3 years to develop a new treatment plant.

Table 1. Site Characteristics of Three Wastewater Use Studies

Parameter/Site	Ismailia, Egypt	Ojinaga, Mexico	Las Cruces, US
Soil type	Sand	Clay loam	Sand
Soil depth	10 m (to groundwater)	3 m (to groundwater)	3 m (to impervious layer)
Species	*Cupressus, Pinus*	*Populus, Eucalyptus*	*Larrea tridentata, Prosopis glandulosa*
Rotation	15–20 yr	6–7 yrs	n/a
Irrigation	Drip	Flood	Sprinkler
TDS	550	1950–2220	2500–3500
BOD	46	29–43	28
TKN	45	14–37	8–21

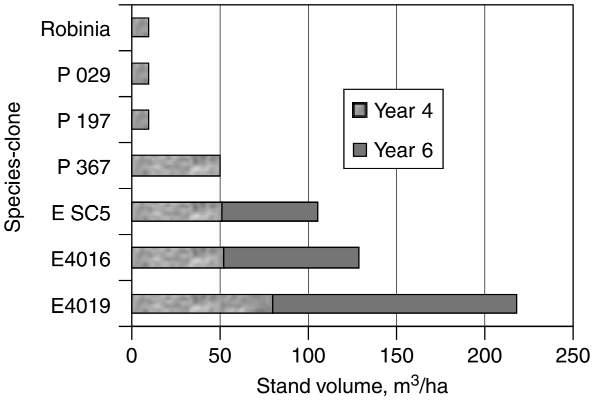

Figure 2. Ojinaga, Mexico, pulpwood stand volume 4 and 6 years after establishment.

Ismailia, Egypt

Ismailia, Egypt (30.60°N 32.29°E) built the Serrabium Wastewater Treatment Plant with Egypt and US AID support in 1995; it is operated by the Suez Canal Authority. The average mean temperature is 21.4 °C; monthly temperatures vary from 13.2 °C in January to 28.5 °C in July and August. Annual potential evapotranspiration rates are 2340 mm and annual precipitation is 25 mm. The treatment plant's designed capacity is 90,000 m^3/day, and the current flow is 80–85 thousand m^3/day. The amount of nitrogen in the wastewater is 45 mg/L, and groundwater depth is 10 m.

Pretreatment for the facility consists of three self-cleaning bar screens and two grit chambers. Flow is divided into two parallel treatment branches with three lagoons (aerated, facultative, and polishing lagoons) for each branch. The aerated lagoon has 20 aerators; 10 operate at any one time; the facultative lagoon has 10 aerators that are used only as needed to increase dissolved oxygen. The polishing lagoon is 3.5 m deep to facilitate the removal of fecal coliform bacteria and *Ascaris* eggs. The total detention time of the system is 11 days, designed to reduce BOD, TSS, *Ascaris* eggs, and fecal coliform bacteria.

The land application facility uses an area of 200 ha. The plants are drip irrigated with tertiary or secondary treated wastewater (90,000 m^3/day). The nursery at Serrabium Forest has a capacity to produce 100,000 seedlings/yr in polybags (12 × 15 cm) with a sand/clay/peat moss medium. Tree species produced are *Pinus pinea* (L.), *P. halepensis* (Mill.), *P. brutia* var. *eldarica* (Medv.), *Khaya senegalensis* (Desr.), *Cupressus arizonica* (Green), *C. sempervirens* (L.), *C. macrocarpa* (Hart.), *Morus alba* (L.), *Morus japonica* (Bail.), and *Cassia* sp. (L.). The wastewater is supplied to the site via an underground pipeline and is filtered through an in-line screen and then through several sand filters before it is delivered to the irrigation system.

There was no apparent soil layering or caliche development; however, root development appeared stunted in the upper 40 cm of the soil surface. Soils were saturated below the surface at the midpoint between trees indicating overapplication of wastewater and a potential risk of groundwater contamination. The stunted roots may have also resulted from the overirrigation of wastewater which may have contained some untreated dissolved organics which resulted in oxygen deficiency in the soil column.

Las Cruces, New Mexico, US

The study site located in Las Cruces, New Mexico, USA (32°18′ N, 106°55′ W) is semiarid the average rainfall is 220 mm, and approximately 50% of the annual rainfall occurs between July and September. The annual mean temperature is 15.8 °C, monthly temperatures vary from 5.5 °C in January to 26.6 °C in July. The potential evapotranspiration of the site is 1800 mm. The City of Las Cruces disposes of treated wastewater onto 36 ha of native Chihuahuan Desert vegetation including creosote bush *Larrea tridentata* (DC.) Cov. mesquite *Prosopis glandulosa* (Torr.) and annual and perennial forbs and grasses. The treatment plant designed capacity is 1500 m^3/day; current flows are 200 m^3/day applied to the 36-ha site.

The wastewater originates from tenants in an industrial park involved in dairy processing and metal wire fabrication. The total salinity (TDS) has reached as much as 4000 mg/L and is comprised mostly of sodium and chloride. (Fig. 3). The total nitrogen applied to the research application site in 2002 was 43.9 kg/ha. The wastewater

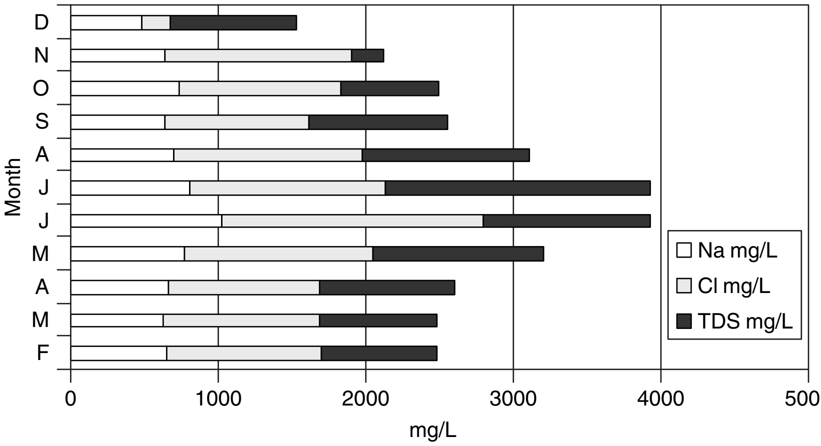

Figure 3. Las Cruces, NM, Effluent characteristics during first year of wastewater application.

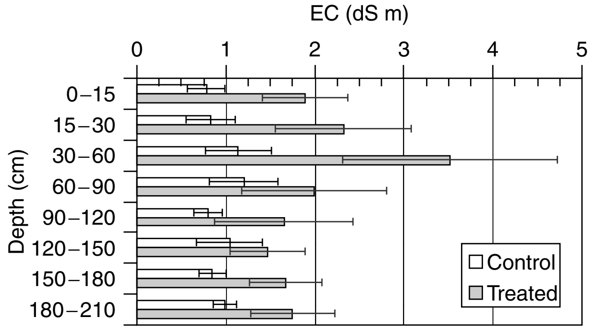

Figure 4. Electrical conductivity of soil vs. depth after 64 cm of wastewater application.

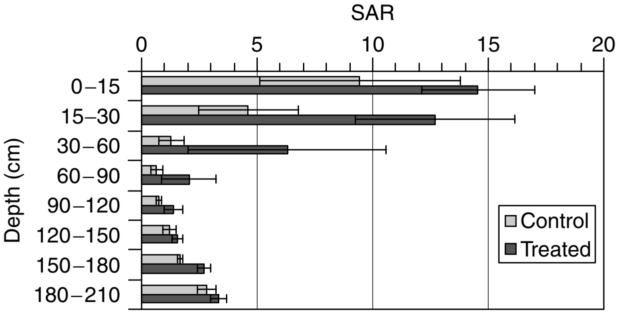

Figure 5. Sodicity measurements on soil after 64 cm of wastewater application.

is sent to a treatment train and then into aerated lagoons with a complete mix basin that has a fixed floor fine bubble diffuser system and a settling basin. The wastewater is then sent to a final holding pond before being land applied through a fixed head sprinkler system. After 64 cm of wastewater application during a growing season, there was an increase in soil salinity. Electrical conductivity (EC) peaked at 3.25 dS/m in the 30–60 cm sampling depth (Fig. 4). The sodium adsorption ratio (SAR) increased to a potential sodium hazard in the top 15 cm (Fig. 5). After 1 year, wastewater effectively deposited the salt ions in the top 60 cm; little or no salt accumulated in the lower depths. Further application will further deposit the salts beyond the root zone and limit growth reduction of the vegetation.

CONCLUSION

These wastewater land application case studies are located in arid regions that have minimal precipitation and high evapotranspiration. The applied water must be balanced against the needs of the plant and salt and nitrogen loading to ensure the development of a sustainable system. Therefore, the applied wastewater must meet the plant's water use needs. When seasonal plant demands are lower in the winter months, the wastewater must be applied at depths to prevent soil salinity increases in the rhizosphere. These interactions must be resolved through proper plant selection, site characteristic considerations, and water quality. The goals of the system in terms of the final product need to be incorporated into the process. These systems can be privately or publicly owned, but infrastructure and community support must be developed in advance and are essential to a project's success.

BIBLIOGRAPHY

1. U.S. Environmental Protection Agency. (1992). *Wastewater Treatment/Disposal for Small Communities Manual*. EPA.625-R-92-005. EPA, Washington, DC.
2. Sopper, W.E. and Kardos, L.T. (1973). Vegetation responses to irrigation with treated municipal wastewater. In: *Recycling Treated Municipal Wastewater and Sludge Through Forest and Cropland*. W.E. Sopper and L.T. Kardos (Eds.). The Pennsylvania State University Press, University Park, PA.
3. Sopper, W.E., Seaker, E.M., and Bastian, R.K. (Eds.). (1982). *Land Reclamation and Biomass Production with Municipal Wastewater and Sludge*. The Pennsylvania State University Press, University Park, PA.
4. Bastian, R.K. and Ryan, J.A. (1986). *Design and Management of Successful Land Application Systems*. *Proc. Utilization, Treatment, Disposal Waste Land*. Soil Science Society of America, Madison, WI, pp. 217–234.
5. Luecke, D.F. and De la Parra, C. (1994). From pollution to park. Experiment in Tijuana: A low-tech approach to wastewater management. *Calif. Coast Ocean* **10**(1): 7–19.
6. Al-Jamal, M.S., Sammis, T.W., Mexal, J.G., Picchioni, G.A., and Zachritz, W.H. (2002). A growth-irrigation scheduling model for wastewater use in forest production. *Agric. Water Manage.* **56**: 57–79.
7. Thomas, R.E. (1973). The soil as a physical filter. In: *Recycling Treated Municipal Wastewater and Sludge Through Forest and Cropland*. W.E. Sopper and L.T. Kardos (Eds.). The Pennsylvania State University Press, University Park, PA.
8. Watanabe, M.E. (1997). Phytoremediation on the brink of commercialization. *Environ. Sci. Technol.* **31**(4): 182–186.
9. Gori, R., Ferrini, F., Niece, F.P., and Lubello, C. (2000). Effect of reclaimed wastewater on the growth and nutrient content of three landscape shrubs. *J. Environ. Hortic.* **18**(2): 108–114.
10. Hoffman, W.C., Karpiscak, M.M., and Bartels, P.G. (1987). Response of cotton, alfalfa, and cantaloupe to foliar-deposited salt in an arid environment. *J. Environ. Qual.* **16**: 267–272.
11. Coate, C. (Ed.). (1998). *FORUM of the Association for Arid Lands Studies. Volume XV Number 1*. International Center for Arid and Semiarid Land Studies, Lubbock, TX.
12. Metcalf & Eddy, Inc. (1991). *Wastewater Engineering: Treatment, Disposal, and Reuse*. McGraw-Hill, New York.
13. Water Pollution Control Federation (WPCF). (1990). *Natural Systems for Wastewater Treatment*. Water Pollution Control Federation, Alexandria, VA.
14. Food and Agriculture Organization of the United Nations (FAO). (1992). CROPWAT—A computer program for irrigation planning and management. Irrigation and Drainage Paper No. 46. Rome, Italy, p. 126.
15. Sammis, T.W. and Herrera, E. (2000). *Leaching Requirements of Pecan and Fruit Trees, Guide H-644*. New Mexico State Agricultural Experiment Station, Las Cruces, NM.
16. Sammis, T.W., Mapel, C.L., Lugg, D.G., Lansford, R.R., and McGuckin, J.T. (1985). Evapotranspiration crop coefficients predicated using growing-degree-days. *Trans. ASAE* **28**(3): 773–780.

TECHNOLOGIES FOR ARSENIC REMOVAL FROM CONTAMINATED WATER SOURCES

I. Katsoyiannis
A. Zouboulis
Aristotle University of Thessaloniki
Thessaloniki, Greece

The majority of established arsenic removal technologies can be divided into three main categories, depending on the specific mechanism that is responsible for removing arsenic. These categories are usually referred as coagulation–precipitation, membrane, and sorption processes.

COAGULATION–PRECIPITATION PROCESSES

Three processes in this group have been developed for arsenic removal; aluminum or iron coagulation, lime softening, as well as a combination of iron and manganese removal with arsenic.

Coagulation Using Aluminum or Iron Salts

The coagulation/filtration process is considered a common water treatment technology, used to remove suspended, colloidal, and dissolved constituents from water sources or from wastewater. The coagulation process promotes the aggregation of suspended solids to form larger flocs, which can be subsequently separated/removed by sedimentation (settling) and/or filtration. Coagulation consists of three steps; coagulant addition and hydrolysis, particle destabilization, and interparticle collisions (floc growth). The first two steps, coagulant formation and particle destabilization, occur during initial rapid mixing, whereas the third step takes place during the following slow mixing, which enhances the flocculation of solids.

Aluminum (III) and iron (III) salts, such as aluminum sulfate (usually referred to as alum) or ferric chloride, are the most common coagulant reagents, used in water treatment (drinking or wastewater). Coagulation by aluminum or iron (III) salts, or a combination of them, such as the mixed sulfate salt, can also be used to remove dissolved inorganic contaminants such as arsenic from drinking water. The two primary mechanisms for removing dissolved inorganic contaminants are adsorption and occlusion. During adsorption, the dissolved contaminant attaches to the surface of a particle or precipitate. Occlusion occurs when the dissolved contaminant is adsorbed to a particle and then entrapped, as the particles continue to agglomerate. Several factors affect the efficiency of the coagulation process: coagulant dosage, pH, turbidity, natural organic matter, the presence of other anions or cations in the solution to be treated, the zeta potential of the system, and the temperature.

Most of the studies of arsenic removal have focused on As(V) treatment because As(V) can be removed more effectively than As(III) due to its speciation in aqueous solutions (i.e., the formation of dissociated/charged species). Furthermore, As(III) can be rather easily oxidized to As(V) by using a strong oxidant, such as chlorine. Ferric and alum salts are almost equally effective in removing arsenic on a molar basis, but when compared on a weight basis, ferric salts proved more effective than aluminum salts. These results occur because the iron hydroxides produced by the hydrolysis of ferric salts are less soluble than the respective aluminum hydroxides over a wide pH range. As(V) removal proved pH-independent in the pH range from 5.5–8.5, when coagulation took place from the addition of iron salts. When aluminum salts were used, arsenic removal decreased above pH 7. As(V) removal was independent of the initial arsenic concentration in the range of interest for drinking water treatment, and it correlated directly with coagulant dosage (1–3).

Lime Softening

Lime softening is commonly used for removing hardness from water. Hardness is due primarily to dissolved calcium and magnesium species. The addition of lime provides the necessary hydroxide ions, which increase the pH of water, resulting in calcium and magnesium removal by the formation of insoluble $CaCO_3$ and $Mg(OH)_2$ precipitates. Lime softening has also been applied successfully to remove heavy metals, radionuclides, dissolved organics, and viruses by adsorption and occlusion within the calcium carbonate and magnesium hydroxide produced. The typical lime softening treatment process train includes initial rapid mixing of lime, followed by flocculation of solids and separation by sedimentation (settling). These three steps are usually combined in one single unit, called a solids-contact softener. Filtration is usually used for subsequent treatment of the supernatant. Lime or a lime–soda mixture is used for softening, whereas the selection between them depends primarily on the specific type of hardness.

Relatively few studies have been conducted to evaluate the removal of arsenic during lime softening, indicating that arsenic removal is strongly pH-dependent. Removal of both As(III) and As(V) species is relatively small at pH values lower than 10. However, at a pH over 10.5, As(V) removal can reach almost 100%, whereas the removal of As(III) is about 75%. The presence of orthophosphates, as well as of carbonates, could limit the removal of arsenates by $Mg(OH)_2$ precipitates. If manganese is also present in the source water, some additional arsenic removal could be achieved through sorption onto the $Mn(OH)_2$ precipitate, although the formation of $Mn_3(AsO_4)_2$ solid was not considered a major mechanism. The initial arsenate concentration does not seem to affect the efficiency of the method in the range from 5–75 μg/L (2,3).

Combined Removal of Arsenic by Iron and Manganese

A number of processes, which are used to remove soluble iron and manganese cations from contaminated groundwater, involve oxidizing these ions, transforming them to the respective insoluble hydroxides, which can be further separated by sedimentation or filtration. Either iron or manganese ions can be oxidized by using a strong chemical reagent (oxidant), such as chlorine (4). Another way to oxidize iron and manganese cations is by biological

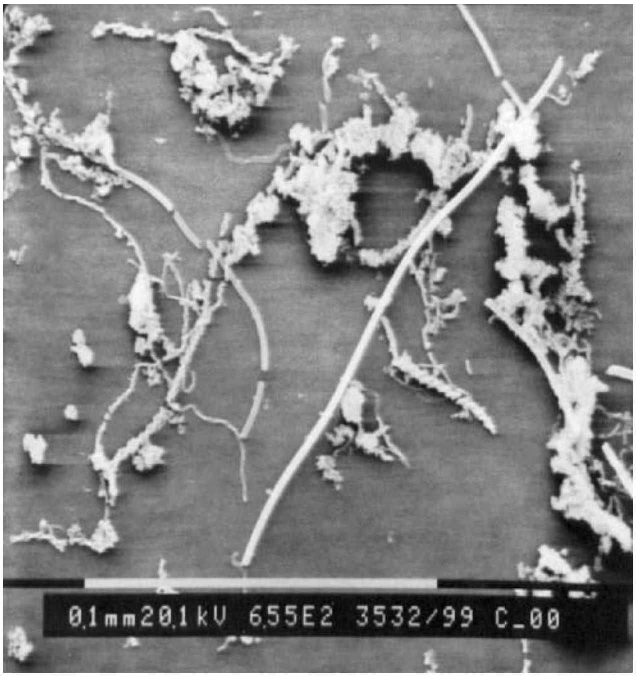

Figure 1. Common iron oxidizing bacteria of the *Gallionella* and *Leptothrix* genera.

oxidation, using so-called “iron oxidizing bacteria”; these facultative bacteria are indigenous to most groundwater (Fig. 1). Biological oxidation involves only preliminary aeration of groundwater to allow bacterial growth and then, the bacteria effectively oxidize iron and manganese to their insoluble hydroxides (5).

In both oxidation cases (chemical and biological), arsenic in the water can be subsequently sorbed onto the iron and manganese hydroxides formed (6). The following major steps have been suggested during iron and manganese oxidation, when used for simultaneous removal of arsenic:

(1) Indigenous “iron oxidizing bacteria” are colonized and immobilized on appropriate inert surfaces (filter media), such as granular sand or polymer beads, forming biofilm deposits and provoking the oxidation of soluble iron and manganese cations that form the respective insoluble precipitates.
(2) The precipitates are subsequently separated from the water by filtration.
(3) Arsenic attaches mainly to the iron hydroxides removed and to manganese oxides by adsorption and/or coprecipitation.

In this method, the initial concentration of iron seems to be critical because the efficiency of arsenic removal is increased by increasing the iron concentration (7). The initial concentration of arsenic does not affect the effectiveness of the method. During the application of biological oxidation for removing iron and manganese, the linear velocity does not affect arsenic removal up to 20–22 m/h (8). A typical biological iron removal (pilot-scale) plant is presented in Fig. 2.

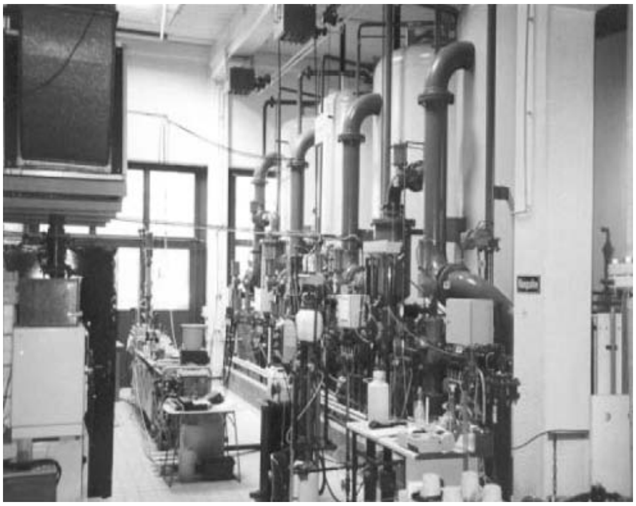

Figure 2. Pilot plant unit, which was used for the biological oxidation and removal of iron and manganese (Umweltbundesamt, Institute for Hygiene of Water, Soil and Air, Berlin, Germany).

MEMBRANE PROCESSES

Two types of membrane processes remove arsenic efficiently from water sources: reverse osmosis and electrodialysis.

Reverse Osmosis

Two types of reverse osmosis units are used for arsenic removal: nanofiltration and hyperfiltration (Fig. 3).

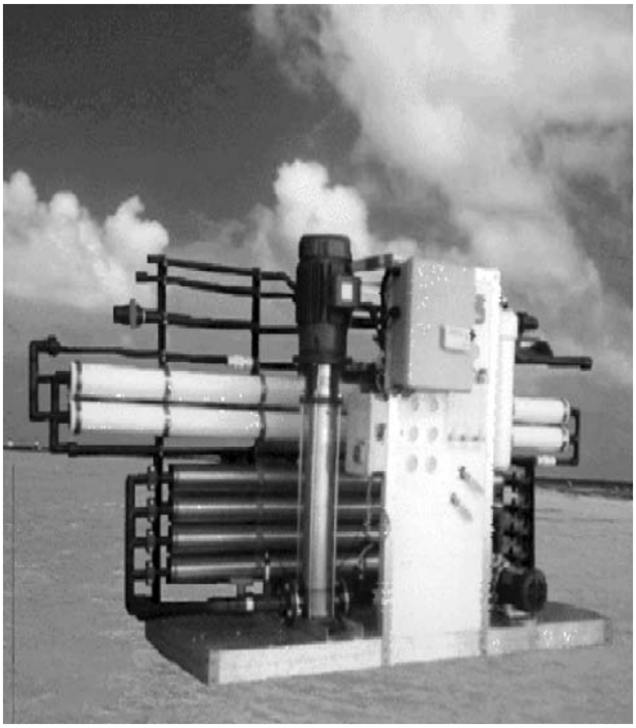

Figure 3. Typical commercial reverse osmosis unit used for drinking water treatment.

Nanofiltration is a relatively lower pressure reverse osmosis process that removes primarily the larger dissolved solids, compared to hyperfiltration. Nanofiltration is also known as a membranous softener because it removes divalent calcium and magnesium cations from water, which cause water hardness. Nanofiltration can also remove sulfates, as well as monovalent dissolved solids, such as chlorides and sodium.

Conversely, hyperfiltration operates at higher pressures, resulting in greater removal of all dissolved solids. A typical example of hyperfiltration is seawater desalination.

Nanofiltration and hyperfiltration are both effective for removing As(V), and the respective efficiencies are up to 95%, but when As(III) is present, a preoxidation step is usually required because As(III) removal efficiencies reportedly reach only 80%. The chemical reagents used to oxidize As(III), such as chlorine, are rather harmful to membranes, which is a major disadvantage of this oxidation method. Furthermore, reverse osmosis rejects/removes simultaneously large amounts of water, the desired product, which is another consideration for drinking water treatment applications (3).

Electrodialysis

Electrodialysis is another type of membrane process, using the passage of direct electric current through the aqueous solution to be treated and appropriate electrodes. Although the main applications of this method are to treat brackish waters, it can be also used for drinking water treatment. Dissolved solids exist as cations (positively charged, such as calcium and magnesium) and anions (negatively charged, such as sulfates and arsenic), so the cations can be moved toward the negatively charged electrode, and the anions to the positively charged electrode. The respective cell is divided into chambers, using specific ion exchange membranes that allow the selective passage of only cations or anions.

The problems in reverse osmosis also exist for electrodialysis. Strong oxidizing agents are usually harmful to the membranes. Large amounts of water are also rejected, eliminating the efficiency of the method. However, there has been little research on arsenic removal by electrodialysis (3).

ADSORPTION PROCESSES

Adsorption processes are another category of arsenic removal technologies; they are based on the adsorption of arsenic onto suitable adsorbing media. Two different technologies have been applied to remove arsenic, activated alumina and ion exchange.

Activated Alumina

Sorption onto activated alumina is considered mainly a physical (and to a lesser extent chemical) process, by which ions are removed from aqueous solutions by available adsorption sites on the surface of aluminum oxide. Activated alumina, usually prepared by dehydrating $Al(OH)_3$ at high temperatures, consists of amorphous and γ-aluminum oxide (Al_2O_3), usually placed in columns. Activated alumina removes arsenic in a number of cases. Although the process has not been yet fully characterized, it is believed that arsenic is adsorbed onto the surface of activated alumina mainly by electrostatic forces. A number of parameters affect the removal efficiency of this treatment method, such as arsenic oxidation state (i.e., tri- or pentavalent), the pH, the presence of other competitive anions, as well as the empty bed contact time (EBCT), which is the main operational/modeling parameter for column operation.

Similarly to the other previously presented treatment methods, As(V) can be removed more effectively than As(III), and the breakthrough point of 0.05 mg/L (the previous concentration limit for arsenic in drinking water) is reached after the treating 300 bed volumes (BV) for As(III), whereas for As(V) after 23,400 BV. The pH also affects the removal of arsenic significantly. Anions are better adsorbed below a pH of 8.2, which is the typical point-of zero charge for activated alumina. Below this value, the surface of activated alumina holds a positive charge, which can be balanced by adsorbing anions, such as arsenates. Several studies showed that As(V) removal is independent of pH in the range between 5.5 and 6. As(III) removal in that pH range is not equally efficient because As(III) is present as nonionic (molecular) arsenious acid (H_3AsO_3).

The presence of other competitive anions, such as phosphates and carbonates, can also affect the method's efficiency for arsenic removal because activated alumina has specific preferences for certain ions. Finally, the correlation between EBCT and arsenic removal is linear, and therefore, as EBCT is increased, the efficiency of arsenic removal is also increased (9,10).

Ion Exchange

Ion exchange is a physical/chemical process, by which ions held electrostatically on the surface of an appropriate solid phase can be exchanged with ions of similar charge from an aqueous solution. The solid phase is typically a synthetic ion exchange resin placed in columns and used preferably to remove particular contaminants of concern; certain natural products, such as zeolite, have been also used occasionally. Ion exchange is commonly used in drinking water treatment to remove hardness from water sources and also to remove nitrate, arsenate, chromate, and selenate anions from municipal water. Anion exchange resins are available in two main types, strong-base resin (SBA) and weak-base resin (WBA). Typically, SBA resins are used for arsenic removal because they tend to be more effective over a wider pH range than WBA resins.

Ion exchange does not remove As(III) because it occurs predominantly as uncharged species (H_3AsO_3) at the usual pH of most water sources. On the other hand, As(V) is negatively charged, and thus, it can be removed by appropriate ion exchange resins. When all the available sites on the resin have been exhausted, the bed can be regenerated by the passage of brine solution (chloride exchange).

The efficiency of the ion exchange process for arsenic removal is strongly affected by the presence of other competitive ions, such as total dissolved solids and sulfates. Other factors affecting the operating efficiency of the ion exchange process are the empty bed contact time, and the disposal of spent regenerant solutions (10).

Adsorptive Filtration Using Iron Oxide Coated Sand

This technology, relatively new in water treatment, is based on the adsorption of cationic metals and/or oxyanions onto the surface of iron oxide, and the particulate matter is simultaneously filtered from the water. Iron oxides form coatings on the surfaces of appropriate granular filter media placed in tanks or in columns, such as sand; subsequently, the modified material can act together as a filter and as an adsorber.

To remove oxyanions from contaminated water, the pH has to be adjusted below 7.2, which is the typical point-of-zero charge of iron oxides. Below this value, iron oxides are positively charged and can remove arsenic effectively. In this technology, As(V) is also treated more efficiently than As(III), because as discussed previously, As(III) is present as an uncharged species, up to pH 8. Above pH 8, As(III) starts to form anionic species, but iron oxides are also negatively charged at these high pH values; thus they cannot interact. Another crucial parameter for column operation is empty bed contact time. By increasing the respective contact time, arsenic removal is also increased. The initial arsenic concentration does not affect the efficiency of the method (11).

OXIDATION OF AS(III)

Arsenic can be removed from water by several technologies, but most of these techniques cannot remove As(III) effectively due to its speciation in water. Therefore, an oxidation step is usually required to transform As(III) into As(V). Oxygen is the preferred oxidant because using it avoids problems that other chemicals present, but the oxidation of arsenic by dissolved oxygen is very slow and therefore, cannot be effectively used. The most feasible oxidants today for arsenic removal are potassium permanganate and Fenton's reagent (a mixture of H_2O_2/Fe^{2+}). Chlorine or hypochlorites have also been used as oxidants of trivalent arsenic. Other techniques are UV irradiation and the use of ozone (9).

BIBLIOGRAPHY

1. Edwards, M. (1994). Chemistry of arsenic removal during coagulation and Fe-Mn oxidation. *J. Am. Water Works Assoc.* **86**(9): 64–78.
2. Fields, K., Chen, A., and Wang, L. (2000). *Report Arsenic Removal from Drinking Water by Coagulation / Filtration and Lime Softening Plants*. EPA/600/R-00/063. Cincinnati, OH.
3. Kartinen, E.O. and Martin, C.J. (1995). An overview of arsenic removal processes. *Desalination* **103**: 79–88.
4. Fields, K., Chen, A., and Wang, I. (2000). *Report Arsenic Removal from Drinking Water by Iron Removal Plants*. EPA/600/R-00/086. Cincinnati, OH.
5. Mouchet, P. (1992). From conventional to biological removal of iron and manganese in France. *J. Am. Water Works Assoc.* **84**(4): 158–166.
6. Lehimas, G.F., Chapman, J.I., and Bourgine, F.P. (1999). Use of biological processes for arsenic removal: A cost effective alternative to chemical treatment for As(III) in groundwater. Available at: www.saur.co.uk/poster/html.
7. U.S. E.P.A. (2000). Office of Groundwater and Drinking Water, Report Arsenic in drinking water-treatment technologies: Removal. Available at: www.epa.gov/OGWD/ars/ars4.html, Cincinnati, OH.
8. Seith, R. and Jekel, M. (1997). Bioxidation of AS(III) in fixed bed reactors. *Vom Wasser* **89**: 283–296.
9. Jekel, M.R. (1994). Removal of arsenic in drinking water treatment. In: *Arsenic in the Environment: Part I: Cycling and Characterization*. J.O. Nriangu (Ed.). John Wiley & Sons, New York, pp. 119–130.
10. Wang, L., Chen, A., and Fiels, K. (2000). *Report Arsenic Removal from Drinking Water by Ion-Exchange and Activated Alumina Plants*. EPA/600/R-00/088. Cincinnati, OH.
11. Benjamin, M.M., Sletten, R.S., Bailey, R.P., and Bennet, T. (1996). Sorption and filtration of metals using iron-oxide coated sand. *Water Res.* **30**: 2609–2620.

BIOCHEMICAL OXYGEN DEMAND

Michael A. Butkus
Joe D. Manous, Jr.
United States Military Academy
West Point, New York

BIOCHEMICAL OXYGEN DEMAND IN NATURAL AND ENGINEERED SYSTEMS

A vast number of organisms, primarily bacteria, survives and even thrives in an environment containing no oxygen. However, most living creatures with which we commonly interact are aerobic; they need oxygen to survive. Aerobic creatures include all of the higher life forms, such as fish, birds, mammals, and reptiles, as well as simpler species such as some bacteria and protozoa. The 4.7 billion year development of the earth and the evolution of species in response to this development have created many types of ecosystems; most support a balance between the demands of aerobic organisms for oxygen and the availability of oxygen. These "balanced ecosystems" are found in aquatic, terrestrial, and atmospheric biomes.

In a natural aquatic system, a large number of aerobic organisms exists. Interestingly, the higher life forms, fish, mammals, and shellfish, do not normally place a large oxygen demand on an aquatic system. Bacteria, and in particular those bacteria that decompose organic material, such as dead aquatic vegetation, for example, algae, can create a high oxygen demand due to the sheer number of organisms that reproduce in a short period of time. Aquatic bacteria that decompose organic material are endemic in the environment. Their populations are typically balanced by the natural system, and the available oxygen is generally sufficient; but human influences can upset this balance. The primary anthropogenic (human) sources of organic material in aquatic systems are sewage and other waste stream discharges. This increase in "food" can stimulate the growth of bacteria and can cause oxygen consumption at a rate greater than can be resupplied naturally. In addition to organic material, waste stream discharges along with overland runoff from agricultural fertilizers can increase the concentration of nutrients such as nitrogen and phosphorus in an aquatic system. This increased nutrient availability can stimulate the growth of aquatic vegetation, especially algae. When this vegetation dies, it serves as a source of organic material and stimulates aerobic bacterial growth and a corresponding oxygen demand. This process is called cultural eutrophication.

Certain chemical compounds (organic and inorganic) can also exert a demand for oxygen if the resulting combination of the chemical compound and oxygen forms a thermodynamically favorable spontaneous reaction. The *combined* biological and chemical requirement for oxygen in a waterbody is termed the biochemical oxygen demand (BOD). Material that makes up BOD can be either in dissolved or particulate (solid) form.

Whether BOD is considered a problem depends on two factors: the magnitude of the oxygen demand and the rate at which oxygen is resupplied to the water. The primary oxygen source in all natural waterbodies is the diffusion of molecular oxygen, O_2, from the atmosphere into water. This transfer of oxygen is limited by the solubility of oxygen in water (see Henry's law) and physical movement of oxygen molecules across the air–water boundary at the water surface.

MICROBIAL GROWTH

The life cycle for bacteria may only last a few days to weeks, but the growth rate can be rapid in response to the availability of some limiting factor. This rate of growth can be described, based on the growth factor that is in shortest supply, by using the empirical Monod model as follows:

$$\frac{dX}{dt} = \frac{\mu_m S X}{K_s + S} - k_d X \tag{1}$$

where X = concentration of biomass (mass/volume)
t = time
K_s = half saturation (limiting factor) constant (mass/volume)
k_d = endogenous decay (death) rate constant (time^{-1})
S = concentration of limiting factor (mass/volume)
μ_m = maximum growth rate of the organisms (time^{-1})

Limiting factors include an energy source (organic or inorganic), a carbon source to build new cell material, and nutrients.

MICROBIAL ENERGETICS

Microorganisms require a source of energy, a source of carbon, and nutrients to grow and reproduce. Microorganisms commonly obtain energy for metabolism via reduction—oxidation (redox) reactions. Redox reactions involve the transfer of an electron from a reduced compound to an oxidized compound. All complete redox reactions must contain a compound that becomes oxidized (loses an electron) and a compound that becomes reduced (gains an electron.) To gain energy, microorganisms oxidize a compound, which is called the "electron donor" because it loses an electron. The electron donor is also referred to as "food." Because electron transfer requires a complete electrical circuit, an electron acceptor is required to complete the transfer. The molecule that receives an electron is called the terminal electron acceptor (TEA).

Electron donors often consist of organic material found in sewage or natural systems, for example, dead aquatic vegetation. These highly reduced compounds contain more electrons and have a higher standard free energy than oxidized compounds that serve as TEAs (1). Inorganic compounds, such as ammonia and ferrous iron, can also serve as electron donors. The concentration of electron donor material in waste is sometimes referred to as the "strength" of the waste.

When oxygen is present in a particular system, it is commonly the preferred TEA over other oxidized compounds, such as nitrate and sulfate, because it is a stronger oxidizer. The chemical reduction half-reaction for molecular oxygen in aquatic systems can be summarized as follows:

$$2e^- + 2H^+ + \tfrac{1}{2}O_2 \longrightarrow H_2O \tag{2}$$

This "consumption" of 1/2 mole of O_2 for every two electrons transferred from an electron donor is the source of the oxygen demand exerted on a system and often the cause of oxygen depletion in waterbodies that contain sewage or other waste discharges. Organisms that use oxygen as a TEA are called aerobic.

MICROBIAL COMPOSITION

In addition to using oxygen as a TEA, microbes also use oxygen as a nutrient. In this case, oxygen is incorporated directly into the cell structure. The microbial cell structure is commonly modeled as follows: $C_5H_7O_2N$ (1). Thus, two moles of oxygen are consumed for every mole of cells produced.

Reaction Kinetics

In a natural or engineered system, there are typically numerous chemical and biological reactions that consume oxygen. Consequently, it is difficult to estimate oxygen demand based on stoichiometry or a biological assay. One method for estimating the quantity of oxygen-demanding material in a waste is to make a gross measurement of the quantity of oxygen consumed by the waste via chemical and biological reactions. For domestic sewage and many natural systems, the gross rate of oxygen consumption over time (sum of chemical and biological reactions) is commonly modeled as the following first-order reaction:

$$\frac{d(\text{DO})}{dt} = -k\text{DO} \tag{3}$$

where DO = dissolved oxygen concentration (mass/volume, e.g., mg/L)
t = time
k = reaction rate (oxygen consumption rate) constant (time^{-1})

For a closed system, Eq. 3 can be integrated from time 0 to t, which yields the following:

$$\mathrm{DO}(t) = \mathrm{DO}_0 e^{-kt} \quad (4)$$

where $\mathrm{DO}(t)$ = the dissolved oxygen concentration at a particular time (mass/volume)
DO_0 = the initial dissolved oxygen concentration (mass/volume)

This relationship is illustrated in Fig. 1, which indicates that there is an exponential decline in oxygen over time.

Similarly, the biochemical oxygen demand, BOD(t), at any time t can be defined as the quantity of oxygen *consumed* up to a given time (t) via chemical and biological reactions:

$$\mathrm{BOD}(t) = \mathrm{BOD}_{\mathrm{ult}}(1 - e^{-kt}) \quad (5)$$

where $\mathrm{BOD}_{\mathrm{ult}}$ = maximum or ultimate demand for oxygen that can be exerted under given conditions (mass/volume)

As illustrated in Fig. 2, BOD(t) increases exponentially over time until $\mathrm{BOD}_{\mathrm{ult}}$ is reached.

Many compounding factors influence BOD, but a significant environmental condition is the temperature of the system. Standard methods have been established for sampling and measuring BOD for comparison and analysis, but a temperature correction is typically required to correlate these laboratory results with field conditions. The correction is determined by using the relationship presented as Eq. 6 (see Ref. 2) and is applied to k (reaction rate constant). This correction is empirical and accounts for changes in both chemical and biological components of the reaction rate constant:

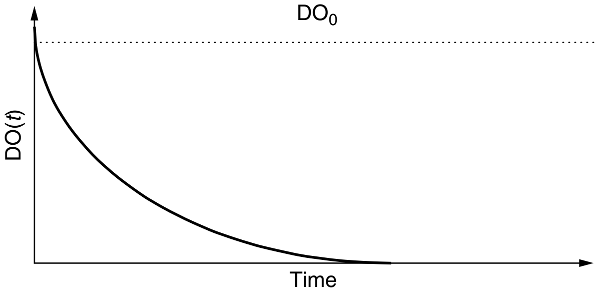

Figure 1. Change in dissolved oxygen with time.

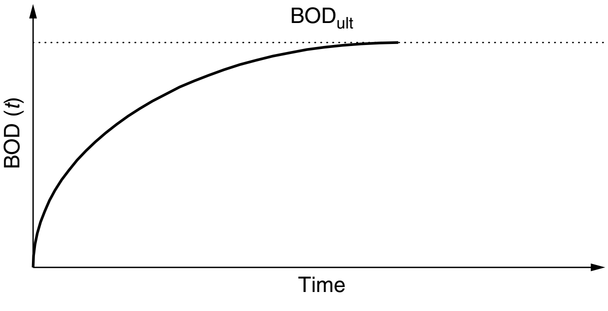

Figure 2. Exertion of BOD in a system.

$$k_{\mathrm{temp}} = k_{20}\theta^{(T-20)} \quad (6)$$

where k_{20} = reaction rate constant at 20 °C
T = temperature (°C)
$\theta = 1.135$ when $4\,^\circ\mathrm{C} \leq T \leq 20\,^\circ\mathrm{C}$ (3)

MEASUREMENT OF BOD

The BOD test is one method for quantifying the concentration of oxygen-demanding material in a waste. In this test, the quantity of oxygen consumed by the waste over a period of 5 days is measured. This measurement includes all viable oxygen-consuming chemical and biological reactions. The standard procedures for BOD measurement are detailed in *Standard Methods for the Examination of Water and Wastewater* (4). The basis of the Standard Methods procedure is measurement of the change in dissolved oxygen concentration in a standard size container after 5 days of incubation under a set of standard environmental conditions (i.e., sealed from the atmosphere, 20 °C, dark conditions). The resulting change in DO is related to the total biochemical oxygen demand of the waste. If the "strength" of the waste is high (high BOD), the standard sample may consume all available oxygen in the sealed sample before the end of the incubation period and yield inconclusive results. In such cases, the sample water is diluted with aerated water to ensure that a nonzero value of DO is measured at the end of the 5-day incubation. The BOD experiment is illustrated in Fig. 3. The water is aerated to saturate it with oxygen (see Henry's law for a discussion of saturation). A seed (freeze-dried microorganisms) can be added to ensure that sufficient microorganisms are present to consume the waste (electron donor). If used, a seed will place an additional oxygen demand on the sample, which must also be considered.

The samples are placed in an incubator at 20 °C for 5 days and are shielded from sunlight to minimize the influence of photosynthesis on the dissolved oxygen concentration. The following relationship is used to

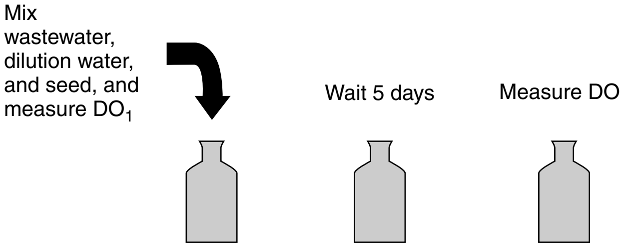

Figure 3. BOD experiment.

calculate the BOD (5) of a sample in which no seed is added:

$$\text{BOD} = \frac{(\text{DO}_1 - \text{DO}_2)}{P} \tag{7}$$

where

$$P = \frac{V_{\text{waste}}}{V_{\text{waste}} + V_{\text{dilution}}}$$

DO_1 = dissolved oxygen concentration of waste/dilution mixture immediately after preparation (mass/volume)
DO_2 = dissolved oxygen concentration of waste/dilution mixture immediately after a 5-day incubation (mass/volume)
V = volume

The BOD (4) is typically less than BOD_{ult} for most wastewaters. However, the BOD (4) is typically used for water quality standards and engineering designs. The total demand for oxygen (BOD_{ult}) by the water body can be determined from the BOD (5) measurement using Eq. 5 and t equal to 5 days.

CONTROL STRATEGIES

For centuries, rivers and lakes have been used as receptacles for human waste. These wastes are major sources of organic material, which can stimulate the growth of aerobic bacteria. In addition, waste streams can contain reduced chemical compounds that react with dissolved oxygen. Contrary to popular belief, the primary purpose of sewage treatment is not to rid the waste of harmful pathogens (though an important side benefit), but to ensure that the carbon and sometimes the nutrient contribution of the sewage to the receiving waterbody does not create a situation where lack of oxygen becomes limiting to aquatic life.

The most common Environmental Protection Agency (EPA) approved control strategy is removal of oxygen-demanding material from waste streams before the discharge enters a natural watercourse. Solid organic material is removed by settling processes (primary treatment), and dissolved organic material is removed through accelerated biological activity (secondary treatment). In secondary treatment, bacteria that are commonly present in domestic waste are provided with sufficient oxygen (a process called aeration) to consume dissolved organic material. In addition, as the organisms reproduce, additional dissolved organic material in the waste is converted into biomass, which is then separated (typically by gravity settling) from the water.

Following secondary treatment, various biological and chemical techniques can be employed (tertiary treatment) to remove additional "contaminants" such as nitrogen, phosphorous, and heavy metals. Primary and secondary treatment of municipal wastewater is required by law in the United States. The requirement for tertiary treatment is determined, case-by-case, based on a combination of total contaminant being discharged in the waste stream and the receiving waterbody's ability to assimilate the contaminant.

BIBLIOGRAPHY

1. Grady, C.P.L., Daigger, G.T., and Lim, H.C. (1999). *Biological Wastewater Treatment*, 2nd Edn. Marcel Dekker, New York.
2. Droste, R.L. (1997). *Theory and Practice of Water and Wastewater Treatment*. John Wiley & Sons, New York.
3. G. Tchobanoglous and F.L. Burton (Eds.). (1991). *Wastewater Engineering—Treatment, Disposal and Reuse*, 3rd Edn. Metcalf and Eddy, Inc., McGraw-Hill, New York.
4. AWWA. (1995). *Standard Methods for the Examination of Water and Wastewater*, 19th Edn. A.D. Eaton, L.S. Clesceri, and A.E. Greenberg (Eds.). American Water Works Association, American Public Health Association, and Water Environment Federation, Washington, DC.

READING LIST

Bitton, G. (1999). *Wastewater Microbiology*, 2nd Edn. Wiley-Liss, New York.

Davis, M.L. and Cornwell, D.A. (1998). *Introduction to Environmental Engineering*, 3rd Edn. McGraw-Hill, New York.

MOLECULAR BIOLOGY TOOLS FOR MONITORING BIODIVERSITY IN WASTEWATER TREATMENT PLANTS

LENA CIRIC
ANDY WHITELEY
Centre for Ecology and Hydrology
Oxford, United Kingdom

JIM PHILP
Napier University
Edinburgh, Scotland, United Kingdom

INTRODUCTION

The group of organisms most directly associated with wastewater treatment is bacteria. They dominate all other groups, both in numbers and biomass, and lead to mineralization and elimination of organic and inorganic compounds (1). Therefore, it seems a natural progression to study the biodiversity and ecology of these organisms in relation to wastewater treatment.

In recent years, molecular methods have been used to study microbial community structure and function due to the revelation that the majority of microbes in the environment are unculturable on standard laboratory agars and conditions (2). It is thought that the percentage of culturable organisms is in the range of 1–10% of the total community; the figure for the number of bacteria described so far is thought to be approximately 1% (3). Taking this into consideration, the use of molecular methods has provided a far more objective and unbiased picture of the total community in wastewater systems.

The methods available can be used in various ways to provide separate pieces of information. Some are used to give a profile of the total community at the site, whereas others identify the organisms performing a specific function. Another set of methods focuses on fluorescent probing of specific whole cells. More recently, new technologies have emerged, which are substrate based, labeling organisms according to their catabolic potential.

COMMUNITY PROFILING METHODS

Community profiling methods produce rapid surveys, which provide us with a phylogenetic profile of the microbial population at a particular site. The speed of application and specificity of these techniques can be used to assess the community composition across space, through time, down pollution gradients, and under various treatments.

Methodologies that provide "molecular fingerprints" are most commonly based on polymerase chain reaction (PCR) amplification of 16S rRNA genes. The 16S rRNA gene is essential as it encodes for a subunit of the prokaryotic ribosome, and therefore is present in all prokaryotic life forms. The equivalent gene in all eukaryotes is the 18S rRNA gene. The gene contains variable regions, which allow identifying organisms down to various taxonomic levels. Fingerprinting studies can also be carried out by amplifying particular functional genes and subsequently analyzing the evolution and phylogeny of enzymes of particular interest.

Denaturing Gradient Gel Elecrophoresis

Denaturing gradient gel electrophoresis (DGGE) has been used successfully in many investigations of community structure. It is now one of the most widespread and well-established methods used in obtaining culture independent microbial profiles.

DGGE is based on analytical separation of DNA fragments of identical or near-identical length based on their sequence composition (4). Separation is based on the changing electrophoretic mobilities of DNA fragments migrating in a gel containing a linearly increasing gradient of DNA denaturants. Changes in fragment mobility are associated with partial melting of the dsDNA in discrete regions, the so-called melting domains. Each band shown on the gel represents a taxonomic unit in the environment, and the band intensity can be associated with the species' abundance within it. Once the gel has been visualized, it is also possible to cut out bands directly for sequencing.

The methodology derives from one used in the medical sciences (4) and was subsequently modified for microbial community analyses (5), where the procedure is performed on the total community nucleic acid. An example of a DGGE gel is shown in Fig. 1 (6). The communities shown were extracted from various points of a diesel-contaminated groundwater remediation system.

DGGE is an efficient and inexpensive method of analyzing a community structure and diversity. It also allows higher phylogenetic resolution than many other community analysis methods. Furthermore, one has an advantage in being able to excise bands for sequencing. The only drawback of the method is that only sample sets of up to 50 can be analyzed on one gel, as gel-to-gel comparisons can be difficult.

Length Heterogeneity PCR

More recently, other methods have become available, such as length heterogeneity PCR (LH-PCR); (7). LH-PCR works by exploiting the natural variation in length of the 16S rDNA gene due to group-specific "variable regions" of microbial community members (8). The PCR is carried out using one fluorescently labeled primer, and the amplicons lengths are then separated on a sequencer, where the resultant peaks, comparable with DGGE gel bands, represent various phylogenetic groups. An example of an LH-PCR output can be seen in Fig. 1, where peaks shown represent various phylogenetic groups in a community in a diesel-contaminated groundwater sample.

LH-PCR is reliable for determining the identity of organisms down to class/subclass level and allows processing large sample sets where samples are directly comparable due to an internal size standard run on the sequencer. One drawback of LH-PCR is that organisms can be effectively identified down only to class/subclass level and single peaks may represent multiple bacterial genera within the same taxonomic subclass (e.g., γ-*Proteobacteria*). LH-PCR is especially suitable for tracking known community members through time or during different treatments. In time, this method will become more reliable when the databases available for various environmental samples are improved.

Terminal Restriction Fragment Length Polymorphism (tRFLP)

tRFLP has been used as a reliable community profiling technique by sizing variable length restriction fragment digest patterns of amplified 16S rDNA (9). The method is similar to LH-PCR but involves an endonuclease restriction step and gives a higher phylogenetic resolution based on the specific sequence variance within the 16S rRNA gene.

tRFLP is reliable for determining the identity of organisms down to group level and, in parallel with LH-PCR, it allows processing large sample sets where samples are directly comparable; an example is shown in Fig. 1. The problem of identifying organisms below group level is somewhat mitigated by tRFLP. It must be emphasized that the choice of enzyme for digests to resolve bacterial taxonomic groups during tRFLP analyses is critical and requires some a priori knowledge of the bacterial groups in the samples for effective phylogenetic targeting.

THE USE OF MOLECULAR PROBES IN COMMUNITY ANALYSIS

Rather than providing a community fingerprint, these methodologies are used to probe for certain known community members. These methods require some a priori knowledge of the population or can be used to determine whether an organism is present. A molecular probe is applied, which hybridizes with the DNA or RNA of the specific organism. The probe can be applied *in situ* or *ex situ*. These methods are powerful tools, which avoid the possible biases of cloning and PCR amplification techniques, and can yield a more direct measure of the target groups of interest.

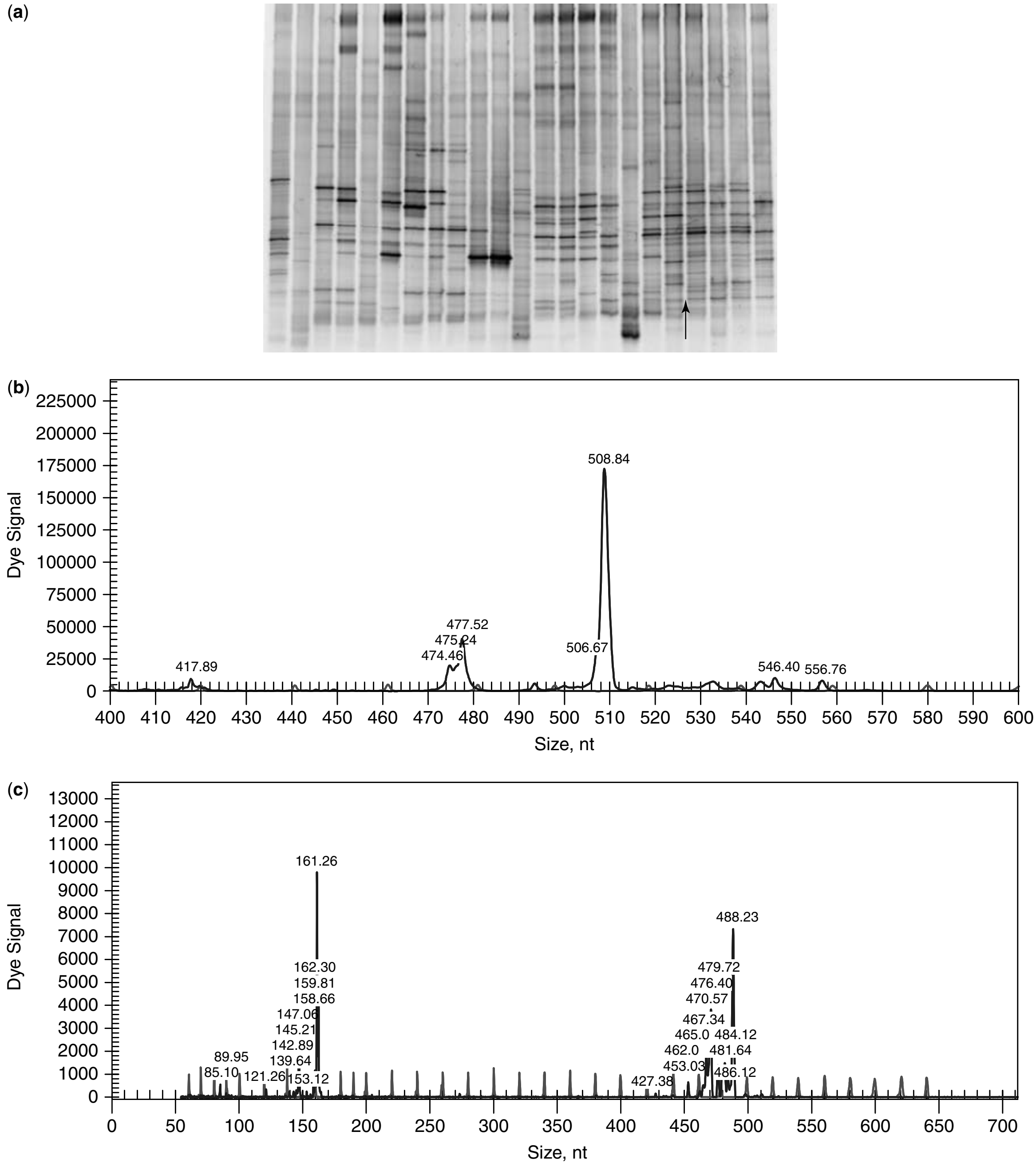

Figure 1. (**a**) DGGE profile showing samples taken from a groundwater diesel remediation system. (**b**) An example of an LH-PCR output. (**c**) An example of a tRFLP output. Both outputs correspond to lane 20 on the DGGE gel.

DNA Microarrays

A DNA microarray (a DNA microchip or DNA chip) is an orderly, high-density matrix of hundreds (or thousands) of individual longer cDNA probes or short oligonucleotides bound directly or indirectly to a solid surface (10). Unlike membrane hybridization, the chip is a high-density format that allows simultaneous hybridization of a labeled DNA or RNA target to a large set of probes, thus providing high throughput. Most applications have been in cell biology such as applied to drug discovery or to monitor gene expression patterns in pure culture, but more recently, the technology has been used successfully in environmental studies (11). Microarrays are a useful tool for high turnover screening of large numbers of samples. They are, however, rather expensive to construct.

Reverse Sample Genome Probing

Reverse sample genome probing (RSGP) uses the entire genome of a microorganism as a specific probe that allows detecting it in the environment. Whole genome probes have been used to detect *Mycobacterium, Mycoplasma, Chlamidia, Bacteroides, Pseudomonas, Sphingomonas*,

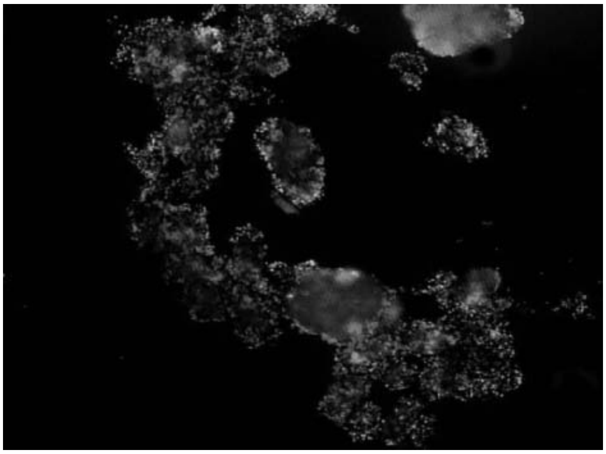

Figure 2. FISH probing of *Thauera* species in a wastewater treatment system. The *Thauera* specific probe is labeled in red. The species Eub358 is shown in green.

and *Campylobacter* species. A good example of the use of this method was looking for sulfate-reducing organisms in contaminated oil fields (12).

Fluorescent *In Situ* Hybridization of Whole Cells

Whole cell *in situ* hybridization with fluorescently labeled oligonucleotides in community analysis was first developed in the late 1980s (13). The procedure involves fixing the environmental sample to permeabilize the cells while maintaining their morphological integrity. The cells are then immersed in hybridization solution containing fluorescently labeled oligonucleotide. After washing to remove unbound probes, the sample is viewed by epifluorescence microscopy. Cells that show specific hybridization with the fluorochrome-labeled probe can be identified and enumerated, as shown in Fig. 2.

THE USE OF LABELED SUBSTRATES TO PROBE FOR ORGANISMS

In addition to community profiling methods and the use of specific oligonucleotide probes, substrate-based procedures have been developed that are culture independent and can be employed *in situ*. These methods, in particular, employ stable isotopes, such as ^{13}C, to determine exactly which organisms are involved in the breakdown of specific contaminants. These methods allow identification of the organisms involved in processing the contaminating compound and are some of the most sophisticated methods used in the field.

Polar Lipid-Derived Fatty Acid-Based SIP

^{13}C labeled substrate is pulsed into the microbial community in the environment and results in labeling polar lipid-derived fatty acids (PLFAs) from assimilating organisms. PLFAs are then extracted, separated, and analyzed for ^{13}C enrichment by isotope ratio mass spectrometry (IRMS). Because specific phylogenetic groups produce signature PLFA profiles, the stable isotope enrichment of certain PLFAs reveal which organisms were dominating the metabolism of the labeled substrate.

The first PLFA-SIP investigation was carried out by Boschker et al. (14), who identified the microorganisms responsible for oxidizing the greenhouse gas methane in a freshwater sediment.

DNA Stable Isotope Probing

More recently, nucleic acids have been used in SIP studies. These methods are more user-friendly because sequence information is so widely available, and nucleic acid extraction from environmental samples is easy.

Stable isotope labeled DNA can be isolated from mixed microbial communities based on the increase in buoyant density from isotopic enrichment. Density centrifugation in cesium chloride gradients was used to separate "heavy" for natural DNA, and 16S rDNA clone libraries constructed from heavy DNA were sequenced to obtain the identity of organisms assimilating the ^{13}C labeled substrate used in the study. This technique has been used to attribute methanol to use particular *Proteobacteria* in an oak forest soil (15).

RNA Stable Isotope Probing

Due to the fact that DNA synthesis is associated only with cell replication, the amounts of DNA synthesized in the duration of a ^{13}C pulse may be low in a natural environment with low cell replication. The use of RNA in SIP, however, offers the same sequence-based resolution but avoids the limitations of labeling due to its high turnover rate. The procedure is performed as DNA-SIP above, and reverse transcription of the heavy RNA to identify organisms is added. RNA-SIP has been applied to an operative industrial phenol-degrading wastewater treatment system to identify organisms responsible for metabolizing phenol (16). Figure 3 shows the fractionation of natural and heavy ^{13}C labeled RNA. The fractionation,

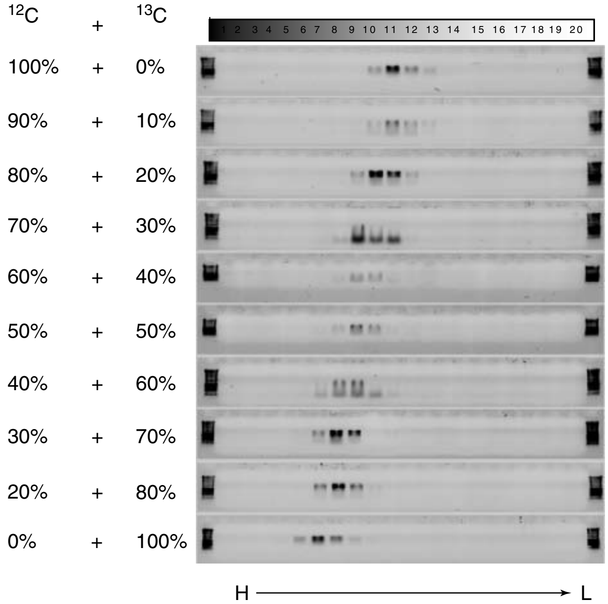

Figure 3. The appearance of natural (^{12}C) and heavy (^{13}C labeled) RNA shown in RNA-SIP. The top gradient fractions contain the natural weight RNA, and the heavy RNA becomes apparent as we move down the fractions (17).

extraction, and subsequent sequencing of the heavy RNA resulted in identifying the dominant phenol degrader in a wastewater system as a *Thauera* species.

Fluorescence *In Situ* Hybridization and Secondary Ion Mass Spectrometry

Recently, investigations in which individual cells or mixed aggregates of cells in methane-consuming communities were identified by fluorescence *in situ* hybridization and, subsequently, analyses for ^{13}C content by secondary ion mass spectrometry (18). The natural abundance of ^{13}C in methane is low, thus enabling the association of cells harboring depleted ^{13}C signatures with methane consumption.

All SIP techniques outlined before are sophisticated methods of associating an organism with a specific function, their catabolic potential. The one problem of SIP is that the labeled substrates are hard to obtain and may require custom synthesis, which is a rather costly process.

BIBLIOGRAPHY

1. Amann, R., Lemmer, H., and Wagner, M. (1998). Monitoring the community structure of wastewater treatment plants: a comparison of old and new techniques. *FEMS Microbiol. Ecol.* **25**: 205–215.
2. Amann, R.I., Ludwig, W., and Schleifer, K.H. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *FEMS Microbiol. Rev.* **59**: 143–169.
3. Cox, C.B. and Moore, P.D. (2000). Biodiversity. *Biogeography: An Ecological and Evolutionary Approach*, 6th Edn. Blackwell Science, New York.
4. Myers, R.M., Maniatis, T., and Lerman, L.S. (1987). Detection of localisation of single base changes by denaturing gradient gel electrophoresis. *Methods Enzymol.* **155**: 501–527.
5. Muyzer, G., De Wall, E.C., and Uitterlinden, A.G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* **59**: 695–700.
6. Ciric, L., Whiteley, A., Manefield, M., Griffiths, R., Philp, J., and Bailey, M. (2003). Approaches for analyzing the diversity of diesel degrading microbial communities. *Proc. 2nd Eur. Bioremediation Conf.*, pp. 449–452.
7. Suzuki, M., Rappé, M.S., and Giovannoni, S.J. (1998). Kinetic bias in estimates of coastal picoplankton community structure obtained by measurements of small subunit rRNA gene PCR amplicon length heterogeneity. *Appl. Environ. Microbiol.* **64**: 4522–4529.
8. Whiteley, A.S., Griffiths, R.I., and Bailey, M.J. (2003). Analysis of the microbial functional diversity within water stressed soil communities by flow cytometric analysis and CTC+ cell sorting. *J. Microbiol. Methods* **54**: 257–267.
9. Liu, W.T., Marsh, T.L., Cheng, H., and Forney, L.J. (1997). Characterisation of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl. Environ. Microbiol.* **63**: 4516–4522.
10. Service, R.F. (1998). Microchip arrays put DNA on the spot. *Science* **282**: 396–399.
11. Guschin, D.Y., Mobarry, B.K., Proudnikov, D., Stahl, D.A., Rittmann, B.E., and Mirzabekov, A.D. (1997). Oligonucleotide microchips as genosensors for determinative and environmental studies in microbiology. *Appl. Environ. Microbiol.* **63**: 2397–2402.
12. Voordouw, G., Voordouw, K.J., Jack, T.R., Foght, J., Fedorak, P.M., and Westlake, D.W.S. (1992). Identification of distinct communities of sulfate-reducing bacteria in the oil fields by reverse sample genome probing. *Appl. Environ. Microbiol.* **58**: 3542–3552.
13. DeLong, E.F., Wickham, G.S., and Pace, N.R. (1989). Phylogenetic strains: ribosomal RNA-based probes for the identification of single microbial cells. *Science* **243**: 1360–1363.
14. Boschker, H.T.S., Nold, S.C., Wellsbury, P., Bos, D., de Graaf, W., Pel, R., Parkes, R.J., and Cappenberg, T.E. (1998). Direct linking of microbial populations to specific biogeochemical processes by ^{13}C-labelling of biomarkers. *Nature* **392**: 801–805.
15. Radajevski, S., Ineson, P., Parekh, N.R., and Murrell, J.C. (2000). Stable-isotope probing as a tool in microbial ecology. *Nature* **403**: 646–649.
16. Manefield, M., Whiteley, A.S., Griffiths, R.I., and Bailey, M.J. (2002). RNA stable isotope probing: A novel means of linking microbial community function to phylogeny. *Appl. Env. Microbiol.* **68**: 5367–5373.
17. Manefield, M., Whiteley, A.S., Ostle, N., Ineson, P., and Bailey, M.J. (2002). Technical considerations for RNA-based stable isotope probing: An approach in associating microbial diversity with microbial function. *Rapid Commun. Mass Spectrom.* **16**: 2179–2183.
18. Orphan, V.J., House, C.H., Hinrichs, K.-U., McKeegan, K.D., and DeLong, E.F. (2001). Methane-consuming Archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science* **293**: 484–486.

BIOSOLIDS

IZRAIL S. TUROVISKIY
Jacksonville, Florida

Wastewater treatment processes remove organics, inorganic materials, and contaminants in sludge. Biosolids are treated municipal wastewater sludge that can be beneficially used, especially as a soil amendment, in accordance with standards.

The Water Environment Federation (WEF) has adopted a policy of encouraging the use of the word ‘biosolids’ in place of sludge to promote public acceptance of reused materials. Land application of biosolids has been and continues to be successfully implemented by many countries. Many countries have rules and regulations associated with biosolids reuse and disposal. Almost all rules are based on reducing pathogens and pollutants. However, rules and regulations vary in different countries. Criteria and methods of obtaining results to determine the use of biosolids also vary.

The United States EPA Standards (US EPA, 1993,1995) for using sludge includes pathogen requirements, vector attraction reduction requirements, and pollutant ceiling concentrations. In accordance with them, two classes of requirements—Class “A” and Class “B” were accepted as rules. To achieve Class “A” (pathogen) requirements, either the density of fecal coliform bacteria in the sludge shall be less than 1000 most probable number (MPN) per gram of total solids, or the density of *Salmonella* sp. bacteria shall be less than 3 MPN per 4 grams of total solids. In addition, the sludge must contain less than one unit of plaque-forming virus and less than one helminth ova per 4 grams of total solids.

Class “B” sludge product must meet one of three alternative requirements. In the first of these, the geometric mean of the density of fecal coliform (expressed either by MPN or colony-forming units per gram of total solids) for seven samples must be less than 2,000,000. The geometric mean is the antilogarithm of the arithmetic average of the logarithms of seven samples. Alternatively, the sludge may be classified as Class “B” by being treated by a process to significantly reduce pathogens (PSRP) or by a process determined to be equivalent to PRSP by the permitting authority.

Class “A” sludge is prone to regrowth of pathogenic bacteria, such as *Salmonella* sp., after treatment because it lacks an adequate density of competitive organisms. Concerns for regrowth led to several additional requirements. First, pathogen reduction must be achieved prior to or at the same time that certain vector attraction reduction requirements are met. Second, pathogen density requirements must be met when the biosolids are used or disposed of, not when sludge is treated. Third, Class “A” biosolids applied to the land for vector attraction reduction must be applied within 8 hours after discharge from the pathogen treatment process.

Class “B” biosolids cannot be sold or given away in a bag or other container for land application. Biosolids, which meet Class “B,” can be applied to agricultural land, forests, public contract sites, or reclamation sites. Land application of Class “B” has restrictions for the harvesting of crops and turf, grazing of animals, and public access.

These restrictions are designed to provide time for the natural environment to reduce the pathogenic organisms in the biosolids. Food crops that touch the biosolids/soil mixture and are aboveground cannot be harvested for 14 months after application of biosolids to land. Root crops cannot be harvested for 20 to 38 months after biosolids application. In this case, the lowest time period applies if biosolids remains on the land surface for 4 months prior to incorporation into the soil, and the higher value applies when it does not. Class “A” and Class “B” biosolids have some vector attraction reduction requirements for land application.

The disposal alternatives are based on the treatment level provided. The most common methods for achieving Classes “A” or “B” biosolids include digestion, composting, thermal drying, and alkaline stabilization.

DIGESTION

Several sludge digestion technologies that consist of anaerobic and aerobic processes are widely used.

Anaerobic Digestion

Anaerobic digestion is a biological process of reducing volatile solids by microorganisms in the absence of oxygen; it reduces odor and pathogen content. Digesters are cylindrical reservoirs with conical bottoms; the upper section of the reservoir has a sealed cover with a device for collecting gas. One-stage and two-stage digesters are in use. In two-stage digesters, the first stage has covered heated reservoirs, and the second stage has open, unheated reservoirs.

Two types of anaerobic digestion processes are in use—mesophilic and thermophilic. Mesophilic processes function in the temperature range of 32–35 °C. The thermophilic process is operated at a higher temperature (50–55 °C) to reduce organic solids and pathogen content further. The process can use conventional standard rate and high-rate digesters with solid retention time from 30–60 days to 10–20 days. To produce Class “B” biosolids, digested sludge has to meet one of the following: reduce volatile solids by a minimum of 38%, or demonstrate that further anaerobic digesting of a portion of the sludge in a bench-scale lab unit for 40 additional days at a temperature between 30 and 37 °C will reduce the volatile solids by less than 17%.

The end products of anaerobic digestion are biosolids, gas, and water. The following gas composition can be expected: methane —60 to 75%; carbon dioxide —16 to 40%; nitrogen, hydrogen, and oxygen —0.4 to 6%. The average heat of combustion of this gas is 21 million joule/m^3. Depending on the chemical composition of the sludge, the liberation of gas varies from 5 to 20 m^3 per m^3 of sludge or 1 m^3 for every 1 kg of a disintegrated sludge’s organic content.

Anaerobic digestion requires energy in the form of heat and electricity. Supplying this energy can be expensive unless part of the energy required is received directly from using the gas. When less than 10–12 m^3 of gas is liberated per 1 m^3 of sludge, the resulting heat balance may be negative. If the gas obtained is in excess of 12 m^3 per 1 m^3 of sludge, it may be used in boilers, heating systems, and so on.

Methane is characterized by explosion hazard and may poison people, if the gas leaks. The volume of gasholders is determined in accordance with gas production and the demand schedule. The volume of gasholders is approximately equal to 2–4 hours of gas production.

Aerobic Digestion

Aerobic digestion is a process of oxidation of the organic part of sludge by microorganisms in special tanks in the presence of oxygen (air aeration of sludge). Aerobic digestion has been widely used in wastewater treatment plants (WWTP) for many years. The purpose of aerobic sludge digestion is to stabilize raw sludge and produce biosolids for further treatment and disposal. Equipment for aerobic digestion is simpler, with in design and operation than that for anaerobic digestion. The more useful sludge aerobic digestion processes are conducted in open structures of the aeration tank type. The duration of the volatile part of solids oxidation depends on the food/microorganism ratio, temperature, intensity and quantity of aeration, and also on wastewater composition and technological demand. This process is more useful for digesting and stabilizing thickened activated sludge.

The detention time to reduce volatile solids and stabilize a mixture of waste activated sludge and raw primary sludge may be 10–15 days at an operating temperature of 20 °C (68 °F). Traditionally, digesters have been designed for a detention time of 20–30 days. However, to meet Class "B" sludge regulations by aerobic digestion, Federal Regulations (40 CFR 257) require detention times of 40 and 60 days at temperatures of 20 °C and 15 °C (68 °F and 59 °F), respectively, and reduction of volatile solids by a minimum of 38%. Long aerobic digestion times require large tank volumes. To reduce tank volumes, the solids concentration can be increased by gravity, belt, or drum thickeners.

Aerobically digested biosolids should be settled for 1.5–5 hours in a sedimentation zone inside the aeration tank or settling in sedimentors. The supernatant BOD is about 100 mg/L, and the COD varies from 350 to 700 mg/L. The moisture content of the biosolids is 95–98% after sedimentation.

Oxidation of the organic part of sludge releases about 3.6 kcal of heat per gram of volatile suspended solids (VSS) in aerobic autoheating mesophilic digestion or autothermal thermophilic aerobic digestion. Temperature control is important in both processes. These processes use heat, created by biochemical oxidation of the organic substances by air introduced into the reactors. The processes are realized in hermetically sealed reactors. The heat balance has to take into account heat loss to surroundings, effluent gas, effluent sludge, and so on.

Autothermal thermophilic aerobic digestion (ATAD) systems for sludge stabilization and disinfection include two hermetically sealed reactors which can be operated with an existing anaerobic digester (prestage ATAD) or operated as conventional ATAD. By using the ATAD process for heat balance, activated sludge has to be thickened to 5% or more of dry solids using belt or rotary drum thickeners. Technological schemes include waste-activated sludge holding and buffer tanks, thickeners, pumps, heat exchangers, reactors, treated sludge holding tanks, and communications. Depending on the temperature and process duration, ATAD can be used to produce Class "B" or "A" biosolids. Class "A" biosolids should be heated to 55–60 °C and aerated for about 10 days. Class "B" aerobic and anaerobic digested biosolids may be converted to Class "A" by maintaining a temperature of at least 70 °C for a minimum of 30 minutes. This can be done by inserting hot air or gases directly into the digested biosolids or by using heat exchangers.

COMPOSTING

Composting of domestic waste and manure has been used for thousands of years. However, the production of compost from wastewater sludge of large-scale municipal wastewater treatment plants began in the 1960–1970s.

Composting can be performed as aerobic or anaerobic biothermal processes which release heat during the decomposition of organics. Aerobic biothermal processes release almost 30 times more heat than anaerobic processes. That is why aerobic processes are used more often. The composting process reduces the organic material in the sludge by approximately 25–30%. During composting, the heat generated by the decomposition of the organic portion of the sludge, reduces the moisture content of the sludge, stabilizes the sludge, and renders the residual harmless by transforming it into usable biosolids.

The heat generated by decomposition of 1.0 kg of organic material averages 20–21 million joules (Mjoule). This generated heat drives the biothermal process and moves the composting process from the mesophilic phase (25 to 40 °C) to the thermophilic phase (55 to 65 °C). Most pathogens are destroyed during the thermophilic phase.

The generated heat also facilitates the evaporation of moisture. Approximately 4.0 Mjoule of heat will evaporate 1.0 kg of moisture (taking into account heat losses and heating of the compost material). Thus, decomposition of 1.0 kg of organic material facilitates the removal of approximately 5.0 kg of moisture from the sludge (21 Mjoule/4 Mjoule/kg of water).

Before composting, it is necessary to dewater sludge. Dewatering reduces the volume of sludge, and it also decreases the amount of moisture to be evaporated by the composting process.

The aerobic process can be performed by using different systems—windrow, aerated static piles, or in-vessel. Sludge is composted together with bulking materials, such as saw dust, wood chips, peat, paper product waste, and ready compost, which provide the mixture with the necessary moisture, porosity, and carbon content.

For vector attraction reduction requirements and to achieve Class "B" biosolids during composting, the temperature has to be above 40 °C (104 °F) for a minimum of 14 days and average more than 45 °C for the same period. To obtain Class "A" compost, the US EPA requires that the temperature in the piles have to be at least 55 °C (131 °F) for 3 days.

Windrow piles or trenches are effective, but mixing and turning does not provide enough air, that is why anaerobic conditions (sometimes) take place in these devices, which generates odors. More often aerated static pile composting is used. Piles have been made by mixing dewatered sludge with wood chips or saw dust. Arrangements of dewatered sludge and amendments prior to mixing follow recycled wood chips (or saw dust) on the floor, dewatered sludge on the top, or when fresh, wood chips (or saw dust) or recycled compost. To prevent heat loss and protect against insects and rodents, piles are covered by ready compost.

In aerated static piles, perforated pipes mounted underneath the composting pile provide aeration. Pipes are covered with layers of sand, gravel, or bulking material and are connected to blowers.

Aerated static pile composting is an effective process. It is easy to operate and does not require high personnel qualification. Compost is a dry, disinfected, and stabilized product. However, the speed of the process depends on several factors, such as bulking materials; proportion of bulking materials to dewatered sludge; efficacy of mixing bulking materials with dewatered sludge; temperature, moisture, and porosity of the mixture; mass of the mixture being composted; quantity of air and duration of aeration; correction between C and N; quantity of recycling compost and bulking material, and toxicity to biological decomposition.

The moisture content of mixed sludge-bulking material has to be 60–67%. The quantity of air is 10–25 m^3/hour per ton of volatile solids. Increased air allows an increase in process speed, but excess air decreases the process speed and temperature. Aeration begins 5–10 days after starting the process and lasts 25–30 days for 8 hours every three days. The index of aeration is the concentration of CO_2. When the concentration of CO_2 increases to more than 7% of gas volume, inhibition of microorganisms and a decrease in temperature occur. At the beginning of the process, the carbon to nitrogen (C:N) ratio is 30:1–34:1 in a pile with saw dust and 25:1–27:1 in a pile with ground bark. Compost is cured for 1–3 months. In the United States, more than 200 composting facilities are in operation now, mostly aerated static piles. But windrow and in-vessel composting also takes place.

Windrow piles are turned with mobile equipment every 3–5 days until the temperature is maintained at 55–65 °C and composted for a period of 3 to 5 weeks, when cured and stored in piles for 1–2 months. The process is based on constructing a windrow from a mixture of dewatered sludge, bulking agent, and recycled compost that provides composting material of proper moisture content, porosity, and carbon source for microorganisms.

Composting processes such as windrow, static piles, aerated static piles, and different types of in-vessel (vertical towers, vessels, reactors, containers, and so on) are widely used in Europe, Japan, and other countries. In-vessel composting processes are accomplished inside enclosed containers that provide a completely controlled environment, temperature, aeration, and odor control. However, in-vessel systems are more complicated and expensive than windrow or aerated static piles and require qualified personnel.

THERMAL DRYING

Thermal drying is designed to disinfect and decrease the weight and volume of wastewater sludge. Thermal drying permits economical sludge transportation from treatment plants and subsequent use.

Thermal drying of sludge is conducted in drying systems consisting of a drying device (dryer) and auxiliary equipment, which includes furnaces with a fuel supply system, feeders, cyclones, scrubbers, blowing equipment, conveyors and bins, as well as monitoring and measuring instruments and automatic control equipment.

To reduce fuel required, sludge should be dewatered before thermal drying by using belt process, centrifuges, or plate presses. For thermal drying of sludge, use fluidized bed dryers, drum dryers, opposite jet dryers, dryers with peripheral discharge of the dried sludge, or Pelletech drying systems. In dryers, the drying temperature should reach 500–800 °C, and drying time is approximately 10–15 minutes.

The Pelletech dryer is a vertically oriented multistage unit that uses steam or thermal transfer fluid in a closed loop to achieve 90% or greater dry solids content in the product. This process minimizes formation of dust and oversized chunks. Pelletizing and drying involves heat and mass transfer, mixing, combustion, conveying, classification, and evaporation.

Thermal dryer sludge is a Class "A" biosolid when sludge achieves 75% dry solids concentration, if no unstabilized primary is present, and 90% of dry solids content, if unstabilized primary sludge is present, by direct or indirect contact with hot gases. Either the temperature of the sludge particles exceeds 80 °C (176 °F) or the wet bulk temperature of the gas in contact with the sludge as the sewage sludge leaves the dryer exceeds 80 °C (176 °F).

The thermal drying process exceeds Class "A" biosolids regulations and can be used for soil improvement without additional restrictions; the solids can be bagged or sold in bulk as fertilizer to landscape contractors or the general public.

ALKALINE STABILIZATION

To produce biosolids that meet pathogen and vector attraction reduction requirements, wastewater treatment plants often use alkaline stabilization. For sludge stabilization, which meets Class "A" or "B," quicklime or hydrated lime are used. Lime is added to sludge before or after dewatering. If quicklime is added to sludge, it initially reacts with water to form hydrated lime. This reaction is exothermic and releases approximately 15,300 cal/g mole (27,500 Btu/lb/mole). The reaction between quicklime and carbon dioxide is also exothermic, releasing approximately 43,300 cal/g mole (78,000 Btu/lb/mole). To achieve Class "A" pathogen reduction, the quantity of added lime is determined so that the pH of the sludge and lime mixture is raised to 12 or greater for 72 hours; the temperature of the sludge during this period must be higher than 52 °C for 12 hours or longer. In addition, air-dried resultant biosolids have to achieve greater than 50% solids content. This process is often called sludge pasteurization. The lime pasteurization process provides a higher level of pathogen reduction than lime stabilization and the result is biosolids that meets Class "A" requirements.

For vector attraction reduction that meets Class "B," the mixture of sludge and lime has to be raised to pH 12 for at least 2 hours and remain over pH 11.5 for an additional 22 hours, all without further alkali addition.

Disinfection of sludge with quicklime or hydrated lime is widely used in Europe to reduce odor, pathogens, and putrefaction of the sludge. Lime stabilization can be part of a sludge conditioning process prior to dewatering (prelime stabilization) or following a dewatering step (postlime stabilization). As a rule, lime is stored in silos and mixed with sludge by various types of mixers.

BIOSOLIDS USE

Biosolids are often used as an organic-mineral nitrogen-phosphorus-potassium fertilizer, soil amendment, or soil conditioning material. Biosolids may be used as a fertilizer on agricultural lands, forests, public contact sites, reclamation sites, rangeland, pastures, lawns or home gardens, or as Class "A" biosolids that are given away in a bag or other container for application to land. Biosolids contain necessary macro- and micronutrients for plant growth. Biosolids may be considered a lime fertilizer when the lime is used for sludge stabilization or pasteurization. The mineral part of biosolids is usually represented by combinations of calcium, silica, aluminum, and iron. Several types of micronutrients are present in biosolids, which increase the speed of many biochemical reactions during plant growth. Most micronutrients can be considered microfertilizers. For example, copper increases the harvest of wheat on marshy and sandy soils. Manganese contributes to increasing the harvest of beets, corn, and some other agricultural crops. A deficiency in iron and zinc in plant growth causes serious breaches in vital activities of plants, specifically related to fruit plants, grape, and cotton plants. Boron is very important for crops such as flax, sugar beets, cotton, feed legumes, peas, clover, alfalfa and several other fruits, berries, and vegetable crops. Micronutrients also contribute to assimilation of organics from the sludge by plants. However, the increased concentration of microelements exercises a negative influence on the growth and quality of plants.

According to the standards for biosolids use, it is necessary to reduce pathogens as mentioned above, but it is not enough. It is very important to determine the pollutants concentration in biosolids, which have to be less than limits and whether biosolids meet cumulative pollutant loading rates (kg/ha) and annual polluting loading rates to soil (kg/ha-year).

Several microelements are essential and some not essential for plants and animals, depending on their doses. Some of them are heavy metals that require some restrictions in use. Among them are arsenic, cadmium, chromium, copper, lead, mercury, molybdenum, nickel, selenium and zinc. Some chemicals, such as barium, beryllium, carbon disulfide, dioxins, and phenol, should be considered hazardous enough to be a potential threat to the health of people and wildlife. The United States currently regulates nine chemical pollutants, but the EPA has already considered regulating six more pollutants. There are special rules and regulations in several economically developed countries that restrict the content of heavy metals in biosolids used as fertilizers and also restrict doses of embedding biosolids in soils.

The requirements for heavy metal content in biosolids become less restrictive when using the soils for shrubs, flowers, for fast-growing trees, like willows, for development of low productive soils and their recuperation, for reinforcement of ravines and hillsides, and for planting trees and shrubs on former industrial waste sites.

If biosolids contain more pollutant concentrations than the requirements of standards, these biosolids should be buried in landfills or incinerated. To reduce expenditures for incineration, sludge organics should be used as fuel.

CONCLUSION

Biosolids are treated wastewater sludge that can be beneficially used as a soil amendment. When considering the use of biosolids, how and where biosolids may be used and how they should be prepared for use corresponding to the requirements of Standards should be determined.

Several processes for sludge treatment such as different methods of digestion, thermal, biothermal, or chemical treatment provide reduction of pathogens and vector attractions to achieve Class "A" or "B" biosolids. Aerobic and anaerobic digestion of sludge are widely used processes of sludge stabilization and produce biosolids that meet Class "B" sludge regulations. Autothermal thermophilic digestion systems operated in hermetically sealed reactors allow production of Class "A" biosolids.

Thermal drying, lime stabilization, and lime pasteurization are used to destroy pathogens, eliminate odor, reduce most of the water content and the volume of sludge, and lower transportation costs of biosolids. These processes are in use for Class "A" or "B" biosolids.

Composting is an effective way of stabilizing and reducing pathogens in sludge. Many utilities use wastewater sludge composting systems, such as windrow, aerated static piles, and in-vessel.

Composting delivers high quality product, which has macro-and micronutrients and can be used as a soil conditioning material, amendment, or fertilizer—either as Class "A" or "B."

Biosolids, after appropriate treatment, may be used as a fertilizer on agricultural lands, forests, public contact sites, reclamation sites, rangeland, pastures, lawns, or home gardens. It should be determined if Class "A" or Class "B" pathogen requirements can be reached and also if the sludge meets one of the vector attraction reduction criteria. After that, pollutant concentration and doses of embedding biosolids in soils should be determined.

Extensive use of dewatered and disinfected biosolids received from municipal wastewater treatment plants as fertilizers will contribute to the solution of the problem of their effective, economical, and ecologically acceptable removal and disposal.

READING LIST

Outwater, A.B. (1994). *Reuse of Sludge and Minor Wastewater Residuals*. CRC Press, Boca Raton, FL.

Epstein, E. (1997). *The Science of Composting*. Technomic, Lancaster, PA.

Shimp, G. and Childress, B. (2002). Improving the process. *Civ. Eng.* **72**(9): 74–77.

Haug, R.T. (1993). *The Practical Handbook of Compost Engineering*. Lewis, Boca Raton, FL.

Lue-Hing, C., Zenz, D.R., and Kuchenither, R. (1992). *Municipal Sewage Sludge Management*. Vol. 4, Technomic, Lancaster, PA.

Manual of Practice #8. (1992). *Design of Municipal Wastewater Treatment Plans*. Vol. 2. WEF, ASCE Manual and Report of Engineering Practice #76.

Porteous, J. (1998). Controlled aerobic digestion of thickened sludge. *Water Eng. Manage.* August: 26–28.

Pelletier, R.A., Slooh, D.S., and Lothrop, T.L. (2001). To lime or not to lime... that is the question. *Fl. Water Resour. J.* February: 37–38; 40.

Spellman, F.R. (1996). *Wastewater Biosolids to Compost*. Technomic, Lancaster, PA.

Turovskiy, I.S. (1999). Beneficial use of wastewater sludge biosolids in Russia. *Fl. Water Resour. J.* May: 23–26.

Turovskiy, I.S. (2001). Technological improvements for the aerobic digestion of sludge. *Water Eng. Manage.* August: 33–36.

Turovskiy, I.S. and Westbrook, J.D. (2002). Wastewater sludge composting. *Water Eng. Manage.* October: 29–32.

U.S. EPA. (1999). *Biosolids Generation, Use, and Disposal in the United States*. EPA530-R-99-009.

U.S. EPA. (1989). *Summary of In-vessel Composting of Municipal Wastewater Sludge*. EPA625/8-89-016.

U.S. EPA. (1993). *The Standards for the Use or Disposal of Sewage Sludge. Title 40 of the Code of Federal Regulation 503*. Federal Register 58 FR 9248 TO 9404. U.S. EPA, Washington, DC.

Wave, H.H. (2003). *U.S. Filter's Mag.* **2**(2): 40.

INTEGRATED CAPACITY BUILDING NEEDS FOR WATER SUPPLY AND WASTEWATER SANITATION

Timothy J. Downs
Clark University
Worcester, Massachusetts

INTRODUCTION

Sustainable development is "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (1). This was the unifying theme of the 1992 Earth Summit in Rio de Janeiro that produced Agenda 21, Blueprint for Sustainable Development. To do most of what the new sustainable development paradigm challenges society to do requires us to strengthen our collective capacity to respond to those challenges and opportunities. "Capacity building is the sum of efforts needed to develop, enhance and utilize the skills of people and institutions to follow a path of sustainable development" (2). A UNDP program, Capacity 21 (2), seeks to build capacity to implement Agenda 21 (3). At the core of integrated watershed management (IWM) lies the engineered human water supply and wastewater sanitation cycle, operating within the hydrologic constraints of the surface watershed and aquifer system (Fig. 1). We must remember, however, that although water resources are the prima facie focus, water cycling depends on a functioning ecosystem comprising interactions of air, soil, climate, forests, and biodiversity.

It is evident from the simplified schematic diagram of the water cycle that this is a complex system that has many types of capacities needed to sustain the sociopolitical, economic, and technological networks it entails. In a typical settlement, there are diverse users of water/producers of wastewater. In a typical watershed, there are many such settlements of varying size and productive activity, comprising people living in water-producing zones (aquifer recharge zones, upland forest zones) and those who consume most water and produce most wastewater (cities and major farming/irrigation zones).

One common focus for societal capacity building is education and training, centered on strengthening human resources. Another common clientele for capacity building is government institutions. Howe (4) has identified key institutional water management requirements for a watershed scale: (1) coordinated management of surface water and groundwater resources, (2) coordinated management of both water quantity and water quality, (3) provision of incentives for greater economic and technical efficiencies in water use, and (4) protection of public values associated with water service (e.g., reliable, safe, clean, affordable supply). To achieve this, the water institutions must develop the following characteristics (after 1):

- the capacity to coordinate water plans with other agencies (e.g., urban planning, agriculture, public health, environment, industrial);
- the capacity to solve water problems creatively using a variety of options and approaches (e.g., laws, pricing, taxes, tradable supply and/or pollution permits, subsidies);
- the foresight to separate roles and responsibilities for water resource planning and management activities

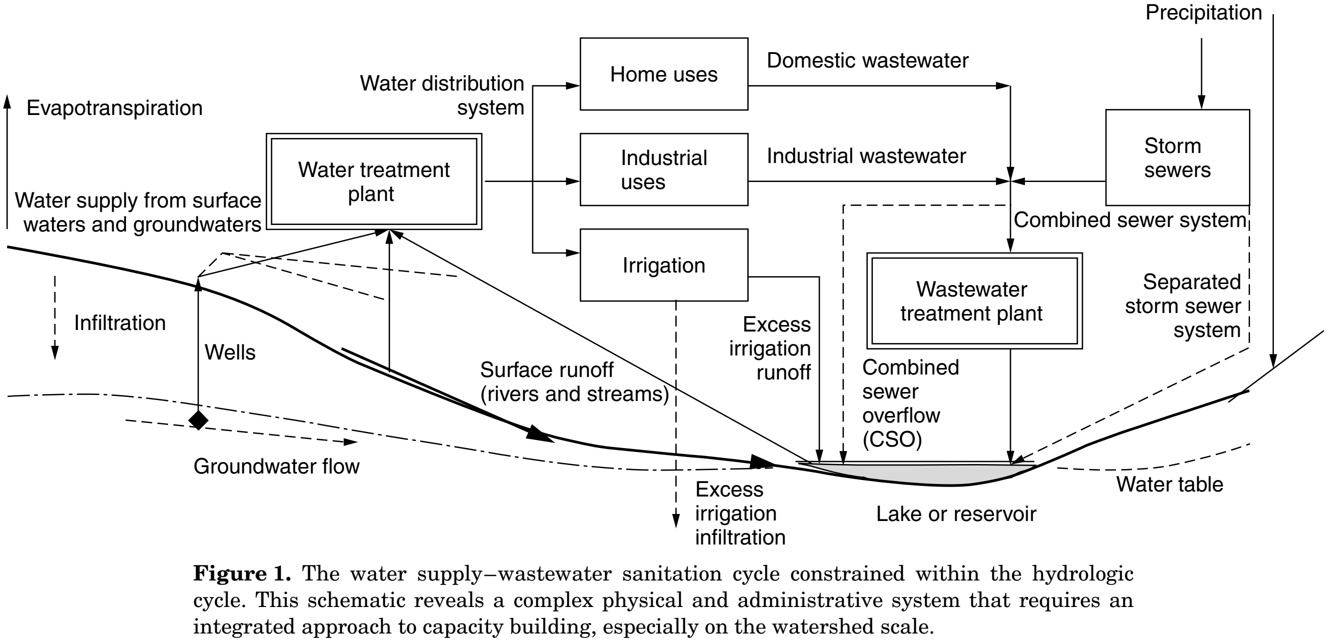

Figure 1. The water supply–wastewater sanitation cycle constrained within the hydrologic cycle. This schematic reveals a complex physical and administrative system that requires an integrated approach to capacity building, especially on the watershed scale.

from construction activities (i.e., avoid conflicts of interest);

- the multidisciplinary capacity to undertake multicriteria/multiobjective planning and evaluation of alternatives;
- devolve decision-making power to the lowest practical level—national, state/regional (provincial), local or municipal—consistent with the scale of the water issue;
- capacity and willingness to use appropriate participatory methods involving different stakeholders at different stages of a project (preplanning/conceptual, planning and design, implementation, maintenance, monitoring and evaluation);
- ability to reward innovation and adapt to changing conditions and priorities.

Wherever capacities are missing or weak, a process of integrated multiobjective, multidisciplinary capacity building is needed, a very common requirement worldwide.

Laws and regulations express societal norms and values and are crucial instruments that help us respond to the challenge of water sustainability. Designing and implementing such responsive legal instruments also reveals the need for an integrated capacity building approach. These legal instruments should (after 5)

- Encourage administration at the appropriate hydrologic scale: watershed, multiwatershed and/or aquifer system.
- Foster internalization of the values and ethics of sustainable resource development.
- Encourage integrated approaches to water supply and sanitation, ecology, and public health.
- Prevent water allocation/usage policy-making from fragmenting among agencies.
- Promote integrated appraisals, notably environmental, economic, and sociopolitical impacts assessment of alternative actions.
- Enforce reward and penalty incentives that encourage sustainability.

And to achieve this, de facto, they should also

- encourage integrated capacity strengthening of governmental institutions, NGOs, community associations, and businesses to transition into more sustainable policies and practices, and work collaboratively.

Current practices of natural resource management and current development policies do not exploit the codependency between coupled water sustainability topics such as (1) water quantity management (e.g., conservation, higher efficiency, and demand management through equitable pricing), (2) water quality management (e.g., standards demanded by diverse users and uses), and (3) aquifer recharge protection and upland forest conservation. This occurs mainly because there is no systematic way to do this that accommodates diverse stakeholder interests and concerns and bridges institutional barriers while strengthening the required social and technological assets (6). Existing methods—with very few notable exceptions—are largely ineffective and not adaptive to changing sociopolitical and geophysical conditions. A recent body of work called participatory integrated capacity building (PICB) addresses this challenge.

PARTICIPATORY INTEGRATED CAPACITY BUILDING

What is participatory integrated capacity building (PICB)? From 1998–2000, local working groups in collaboration with the Mexican National Water Commission (CNA), coordinated by Downs (7), developed a PICB approach to urban water planning and management. Following an analysis that compared relatively sustainable development projects worldwide during the past 10–15 years (those yielding a steady stream of benefits *after* external support was removed) with a much larger number of *un*sustainable projects, six broad synergistic categories of capacity building emerged as critical components for success: (1) strengthening political and financial commitment; (2) strengthening human resources, including education, training, and awareness-raising; (3) strengthening information resources for policy-making (e.g., monitoring and GIS tools for data integration); (4) strengthening policies, regulations, enforcement, and verification; (5) applying appropriate technology and basic infrastructure (e.g., for water and wastewater treatment); and (6) stimulating local enterprise development (i.e., support products and services that provide socioeconomic sustainability. Each one builds on those before it with positive feedback.

This six-story framework is made up of elements that are operational, that is, for which we can design an action plan (Table 1). Table 1 was synthesized from fieldwork, workshops, and literature, notably Alaerts et al. (8), UNDP (9), UNCED (10,11). Although Table 1 considers the needs of an urban area, the same six levels can be extended to determine the needs of periurban and rural areas of the watershed. This integrated approach can be applied on different scales and contexts of settlement, so it can, in theory, also be applied to networks of settlements that comprise a watershed as a whole.

Using Table 1 (or a rural, periurban, or watershed-scale version of it) as our diagnostic template, local working groups undertaking strategic planning can assess existing capacity, then prioritize elements using several criteria, including cultural acceptability, potential cost-effectiveness, and strategic importance to solving the priority problems that have been identified as objectives. In the case study (7), strategic planning was carried out to identify the capacity-building needs of three pilot cities: Mérida, Yucatán; Ciudad Juárez, Chihuahua; and Atizapán de Zaragoza in the State of México. The project was successful in developing strategic work plans that were used to leverage support for the ongoing implementation phase. The main lesson learned in the field was that horizontal multistakeholder processes worked well, even where the norm is nonproductive competition among social groups. The author also saw the need to strengthen sociopolitical support and seed

Table 1. PICB Elements for Sustainable Water Supply and Sanitation in an Urban Context[a]

#	Element Name	Description
		Level 1. Strengthen sociopolitical support and seed financial support
1.1	Sociopolitical support	Build local, state, and federal political support, and commitments of in-kind support (labor and materials).
1.2	Seed finance	Obtain seed funding then leverage other funds for capital projects
		Level 2. Strengthen education and awareness-raising
2.1	Water and sanitation education at the community and school levels	Awareness-raising of water and sanitation issues in the community and the community's rights, obligations, and roles. Education of children at kindergarten, primary, secondary, and preparatory school levels.
2.2	Promotion of a sustainable water and sanitation culture	Use of media instruments and campaigns (TV, radio, print, Web) to promote water savings, efficiency and sanitation goals, and fair pricing. Encourage recognition of water as an economic good and provider of ecological services.
2.3	Train the local trainer	Train local trainers for community, scholastic, and professional education to remove dependency on external experts and make training self-sufficient (links to 2.1 and 2.5).
2.4	Education at university level	Local/regional university education at undergraduate and postgraduate levels to form skilled young professionals and researchers who support water management.
2.5	Professional training	Training in specialist technical and administrative topics for public and private sector participants in water management.
		Level 3. Strengthen information and monitoring to inform policy
3.1	Integrated data collection, processing, and interpretation for informed decision making	Design integrated collection, processing, and interpretation of key water, public health and ecology data, implement it, and integrate information in a GIS format for easy access and decision support.
3.2	Analysis by certified laboratories	Build capacity of laboratories to provide quality-assured analysis of key indicators from 3.1 (links to 2.4, 2.5, 3.1). Instigate locally accessible certification via partnerships with U.S. or Canadian institutions.
3.3	Periodic sustainability diagnosis	Check progress of the sustainability/capacity-building process at intervals for decision-making. Links to 3.1, 4.1.
3.4	Make applied research more cost-effective using a regional applied research network	Facilitate and instigate improved communication and collaboration among researchers to eliminate duplication of effort. Identify, quantify, and prioritize risks. Develop interventions and communication programs with water/sanitation stakeholders.
3.5	Regional institutional network	Facilitate and instigate improved communication and collaboration among public and private institutions (e.g., public health, water, ecology, urban planning).
3.6	Multistakeholder forum for decision making	Facilitate and instigate improved communication and collaboration among representatives from the community, public sector, private sector, and academia to take joint decisions.
		Level 4. Strengthen policy making, regulations, and compliance
4.1	Pollution prevention and reduction policies	Statutory endorsement of sustainable pollution prevention/reduction goals by political leaders at regional, national, state, and local levels.
4.2	Responsive regulatory framework	Review and improve the water supply and sanitation regulatory framework at the municipal, state, and federal levels to make them responsive to municipal, state, and federal needs.
4.3	Regulatory compliance program	Design, develop, and implement a compliance program, including verification and incentives with emphasis on voluntary compliance.
4.4	Tariff and rights schemes	Review and improve the legal and judicial schemes used to assign pollution tariffs and rights. Implement schemes with user buy-in. Equitable supply pricing and allocation for basic needs.
		Level 5. Strengthen basic sanitation infrastructure
5.1	Sustainable water supply	Improved coverage and quality of safe drinking water supply, increasing user's willingness to pay and financial viability of service. Combine sustainable use of groundwater and surface water with demand management and wastewater reuse.
5.2	Wastewater handling/disposal	Best appropriate available technologies (BAATs) for domestic, industrial, and agricultural wastewaters. Includes a spectrum from small-scale rural 'ecological' sanitation to treatment plants with reuse of effluent and viable biosolids.
5.3	Pollution prevention and waste minimization	Design, develop, and implement pollution-prevention programs including pretreatment of industrial effluents, source reduction, and alternative process technologies.
5.4	Solid waste handling/disposal	BAATS for solid waste, including landfill and safe incineration with energy cogeneration.
5.5	Stormwater drainage	Effective evacuation of rainwater and flood prevention. Includes avoidance of stagnant water bodies that harbor disease vectors.

(*continued overleaf*)

Table 1. (*Continued*)

#	Element Name	Description
	Level 6. Strengthen local enterprise development for support products and services	
6.1	Utility company performance	Improve utility company's effectiveness and efficiency at the technical (e.g., leak detection, metering) and administrative levels (e.g., billing)
6.2	Water supply tariffs and rights scheme	Review and improve the legal and judicial schemes used to assign water supply and sanitation tariffs by capacity to pay and rights to water by use priority (direct links to 4.1, 5.3)
6.3	Water supply and sanitation service quality	Improve water supply and sanitation service quality (e.g., continuity, pressure and water quality), in turn boosting users' willingness to pay and economic sustainability
6.4	Develop local water supply and sanitation market	Stimulate local and external provision of products and services to support the water and sanitation sector. Emphasize local providers' participation.
6.5	Stimulate sector investment	Stimulate enterprise development. Provide regulatory and fiscal incentives.

[a] Reference 7.

financial support (Level 1 of Table 1) from the outset and not wait for work plans to be designed. A related issue was to remove any dependency on external, nonlocal support and so make success contingent *primarily* on community resources complemented by some federal support for the first phase of implementation that addressed priority actions.

It is worth noting that the sixth PICB level—enterprise development and the 'market'—recognizes water as an economic good that should be priced according to users' ability to pay. For subsistence communities, a full public subsidy is appropriate to cover basic needs. Overprivatization risks occur when water services are run purely for profit, and marginalized communities unable to pay incur crippling water debts (12). This has occurred in South Africa and touches a nerve in many other places where water is viewed as both a public good and a human right that should not be controlled by commercial interests. We also recognize that the issue of user willingness to pay for water services is cultural (habits of free water) and also one of poor service quality: Why should a person pay for water that is not clean and only comes intermittently? Experience in one pilot site for the PICB project—Ciudad Juárez, a water-scarce, rapidly growing city across the border from El Paso, Texas—demonstrates that even low-income users *are* willing to pay for a reliable, good quality water and sanitation service in a way that allows public utilities to break even on costs.

Significantly, the Ciudad Juárez case was one in which the community of water users was involved in planning and designing from the outset such that a strong sense of ownership of the water supply and sanitation infrastructure was developed. When special interests attempted to take control of the project, the community mobilized to prevent this, one of the few instances of successful public mitigation of corruption in Mexico. This lends credence to our hypothesis that PICB can weather political instabilities and corrupt power plays and also change the public's sense of power over its destiny, restoring faith in the public process and partnerships with responsible institutions. In this way, the three types of sustainability—economic, social, and environmental—are satisfied through a societal capacity building process.

The other caveat to the sixth (and most controversial) level is that PICB places emphasis on *local* providers of products and services such that local entrepreneurs have incentives and opportunities to compete fairly in the water market with external suppliers. External suppliers too often supply inappropriate technological solutions to developing countries where the capacity to operate and maintain the foreign plant is weak or nonexistent. Unfortunately, there are far too many cases, for example, of wastewater treatment plants and analytical laboratories that have been abandoned because of a lack of personnel trained to operate and maintain them (6,7).

OPERATIONAL CONSIDERATIONS

Operational sustainability is a function of integrated capacity building (7). Experience of this capacity building project (7) revealed the need to be clear and concise about the objectives and rationale behind what appears to be a very ambitious and complex process. Integrated watershed management—any environmental management problem—must pass through three key stages *before* any capacity building occurs (although when projects are done in a participatory way, one can argue that societal capacity is already being built). These preplanning stages are (1) a baseline survey of existing conditions and priority needs; (2) problem formulation based on the assessment—identification of priority watershed problems; and (3) project objectives, often policy questions that seek to design and implement 'sustainable solutions' to priority problems. Once preferred solutions/management options are chosen for each priority problem using a multicriteria method (e.g., environmental and social impacts assessment, ESIA), the integrated capacity required *to sustain* each solution is determined. The way PICB fits into typical watershed project stages is shown in Fig. 2.

A participatory, integrated approach to capacity building has one major strategic advantage. Solutions of priority environment–development problems share significant amounts of capacity building requirements. For example, marginalized communities often face coupled sets of problems: (1) unsafe water supply; (2) inadequate wastewater and solid waste sanitation; (3) inadequate water-related disease detection, prevention, and control; (4) grossly inefficient irrigation; and (5) no conservation/mitigation of soil erosion caused by surface runoff. By addressing them

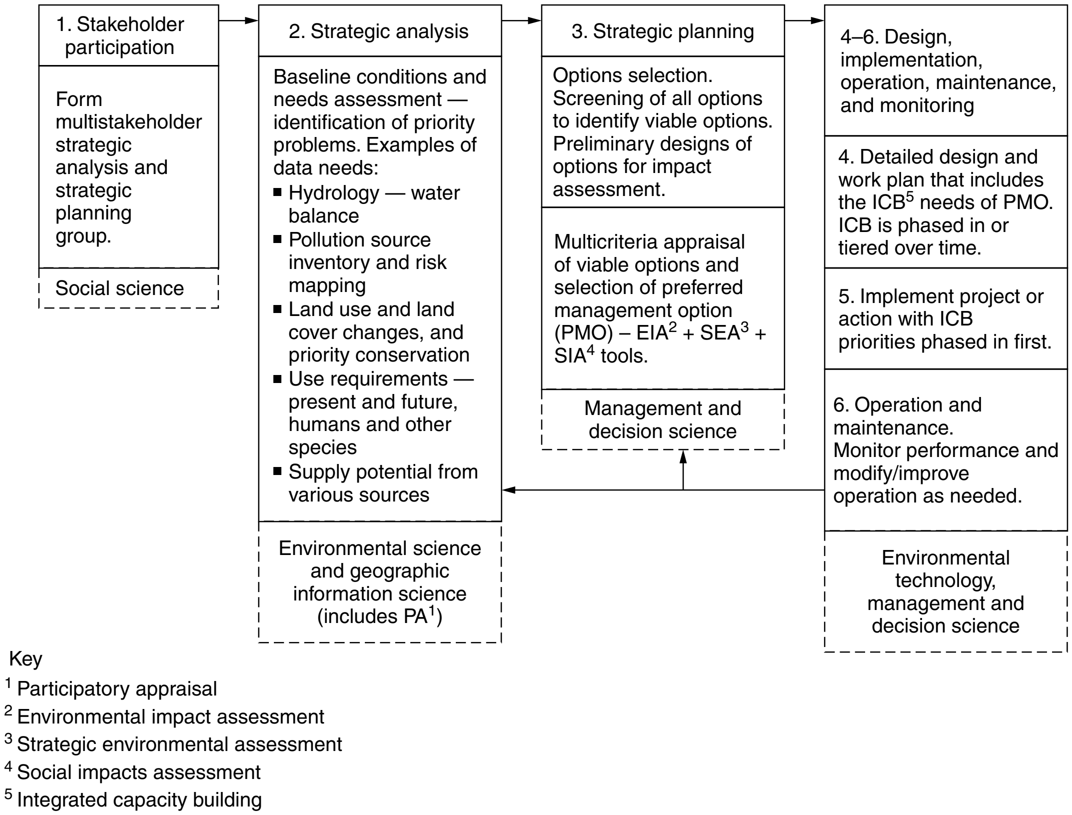

Figure 2. The key role of integrated capacity building (ICB) in sustainable watershed planning and management (6).

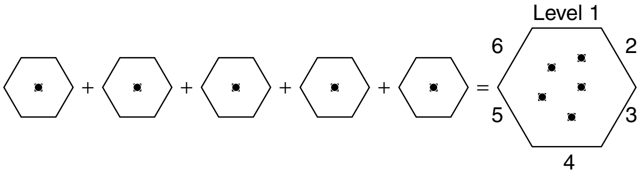

Figure 3. Strategic economies of scale emerge. Six levels of participatory integrated capacity building (PICB) support preferred solutions to five hypothetical priority problems. The integrated six levels are shown as hexagons, solutions as centered dots—the hexagons sustaining the dots. The final plan is shown as a large hexagon of integrated capacities sustaining the five coupled water-centered priority problems. The 'hexagonal' process of integrated capacity building is iterative, cyclical, continuous, just as the process of Fig. 2 is adaptive to new needs and priorities.

collectively, significant economies of scale emerge in the required societal capacity, as shown in Fig. 3.

MULTICAPITAL CONTEXT

Capacity building that supports sustainable development may also be viewed in the theoretical context of investments in, and returns on, strengthening different kinds of capital. Scoones (13) identifies five categories of capital assets that sustain human livelihoods:

1. *Natural Capital*: land, water, wildlife, biodiversity;
2. *Social Capital*: groups, networks, institutions of people;
3. *Human Capital*: skills, knowledge, healthy people;
4. *Physical Capital*: basic infrastructure that supports shelter, transportation, energy, water supply and sanitation, health care, education, and communications;
5. *Financial Capital*: income, savings, credit.

This means that to sustain water resources and watersheds requires the support of all five types of capital and further reinforces the concept that capacity building to strengthen this capital must be multifarious and integrated. Although the concept of capital is human-centered, it can be viewed in the context of ecosystem management, where humans are integral, supermodifying parts of the system. Watersheds are logical ecosystem delineations based on surface hydrology and topography.

Social capital is worthy of special note. Social capital is considered in institutional capacity building, but commonly, formal government agencies are addressed by donors. The central government in developing countries is often given funding from development agencies to build its capacity. Local nongovernmental organizations (NGOs) and community-based organizations (CBOs) typically have to rely on volunteer personnel and donations to function and build capacity. Universities still tend to build their own capacities through academic research grants, though some do partner with NGOs, CBOs, and local government to carry out projects that have clear practical goals. Empirical evidence points strongly to community-based natural resource management and participatory processes as keys to social sustainability. But how can this be done on the watershed scale required for ecological sustainability? The answer lies in strengthening the social capital of the watershed by forming collaborative networks of communities, principally CBOs, local NGOs, local (municipal) government, and local universities. When this is done, human capital and financial capital are also strengthened and become aligned with the mutual gains and common imperatives of watershed stewardship. For example, often in developing countries, marginalized rural agroforesters occupy the upper reaches of the watershed and become, de facto, ‘producers’ of water resources for downstream users because their actions affect zones of upland precipitation, groundwater recharge, and runoff. Both ‘producers’ and users share the common interest of watershed sustainability, and both must be involved in strategic planning. By networking and through capacity building coordinated by NGOs, marginalized communities can gain the social capital/political power they need to negotiate for mutual gains with influential user groups and government agencies. In Mexico, for example, the country has been divided into 26 watershed councils that administer water services to 100 million Mexicans. Unfortunately, the sustainability of this approach is called into question by the lack of participation by ‘producers’ and the clear domination of demand-driven policy-making by influential industrial and agricultural users. This is a useful example because it is a common equity issue in many countries that undermines watershed/ecosystem sustainability in three ways:

- conceptually, because equity issues are at the core of sustainable development;
- thermodynamically, because water-producing zones drive the water cycle; and
- practically, because without effective stakeholder participation, neither can diverse interests be reconciled, nor required integrated societal capacity be built.

CLOSING REMARKS

The concepts and practices of sustainable development, integrated watershed/ecosystem planning and management, equitable stakeholder participation, capital assets, and integrated capacity building are closely connected. The process may appear complex because of its multifaceted nature, but a strategic approach allows us to identify the most important issues, negotiate mutual gains for diverse stakeholder interests, and exploit the considerable economies of scale that emerge. Another advantage of integrated capacity building is that as well as mutual gains, the process delineates responsibilities, roles, and contributions as a collaborative socioecological enterprise.

BIBLIOGRAPHY

1. WCED (World Commission on Environment and Development). (1987). *Our Common Future*. Oxford University Press, New York.
2. UNDP (United Nations Development Programme). (2001). About Capacity 21. Available at: www.sdnp.undp.org/c21.
3. UNCED. (1992). *Earth Summit 1992—the United Nations Conference on Environment and Development*. The Regency Press, London, UK.
4. Howe, C. (1995). *Guidelines for the Design of Effective Water Management Institutions Utilizing Economic Instruments*. Report presented at Workshop on the use of economic principles for the integrated management of freshwater resources, United Nations Environment Programme (UNEP), Nairobi, Kenya.
5. Smith, D. and Rast, W. (1998). Environmentally sustainable management and use of internationally shared freshwater resources. In: *Watershed Management—Practice, Policies and Coordination*. R.J. Reimold (Ed.). McGraw-Hill, New York, p. 294.
6. Downs, T.J. (2002). A participatory integrated capacity building approach to the theory and practice of sustainability—Mexico and New England Watershed Case Studies. In: *International Experiences on Sustainability*. W.L. Filho (Ed.). Peter Lang, Frankfurt am Main, pp. 179–205.
7. Downs, T.J. (2001). Making sustainable development operational: Capacity building for the water supply and sanitation sector in Mexico. *J. Environ. Plann. Manage.* **44**(4): 525–544.
8. Alaerts, G.J., Blair, T.L., and Hartvelt, F.J.A. (Eds.). (1991). *A Strategy for Water Sector Capacity Building*, *IHE Report Series* 24. International Institute for Hydraulic and Environmental Engineering (IHE), Delft, The Netherlands and UNDP, New York.
9. UNDP (United Nations Development Programme). (1998). *Capacity Building for Sustainable Management of Water Resources and Aquatic Environment: A Strategic Framework for UNDP*. UNDP Water Strategy, UNDP, New York.
10. UNCED (United Nations Conference on Environment and Development). (1992). Transfer of environmentally sound technology, cooperation and capacity building, Agenda 21, Chapter 34. Available at: www.un.org/esa/sustdev/agenda21.
11. UNCED. (1992). National mechanisms and international cooperation for capacity building in developing countries, Agenda 21, Chapter 37. Available at: www.un.org/esa/sustdev/agenda21.
12. Gleik, P., Wolff, G., Chalecki, E.L., and Reyes, R. (2002). *The New Economy of Water: The Risks and Benefits of Globalization and Privatization of Fresh Water*. Pacific Institute, Oakland, CA.
13. Scoones, I. (1998). *Sustainable Rural Livelihoods: A Framework for Analysis*. Working Paper 72, Institute of Development Studies, Brighton, UK.

WASTEWATER CHARACTERIZATION

Absar Ahmad Kazmi
Nishihara Environment
Technology, Inc.
Tokyo, Japan

An understanding of the nature of wastewater is fundamental to the design and operation of wastewater collection, treatment, and reuse facilities. Wastewater is characterized in terms of its physical, chemical, and biological composition. Greater emphasis is being placed on wastewater characterization because of its changing characteristics and the imposition of stricter limits on wastewater discharges that are used beneficially. Thorough characterization of wastewater in the design and optimization of biological treatment processes is increasingly important due to the advent of process modeling. Process modeling for activated sludge, as it is currently conceived, requires experimental assessment of kinetic and stoichiometric constants.

The principal physical properties and the chemical and biological constituents of wastewater and their significance are as follows.

PHYSICAL CHARACTERISTICS

Physical characteristics are those characteristics that respond to the senses of sight, touch, taste, and smell. The most important physical characteristic of wastewater is its total solids content, which is composed of floating matter, settleable matter, colloidal matter, and matter in solution. Other important physical characteristics include particle size distribution; turbidity; color; transmittance; temperature; conductivity; and density, specific gravity, and specific weight. Table 1 illustrates common physical characteristics, their analyses, and importance in wastewater treatment.

CHEMICAL CHARACTERISTICS

The chemical constituents of wastewater are typically classified as inorganic and organic. Inorganic chemical constituents of concern include nutrients, nonmetallic constituents, metals, and gases. Organic chemical constituents of concern are BOD, COD, and specific organic compounds.

Inorganic Chemical Characteristics

The sources of inorganic nonmetallic and metallic constituents in wastewater derive from background levels in the water supply, from the additions resulting from domestic use, from the addition of highly mineralized water from private wells and groundwater, and from industrial use. Because the concentration of various inorganic constituents can greatly affect the beneficial uses for the waters, the constituents in each wastewater must be considered separately. Table 2 illustrates the common inorganic chemical characteristics, their analyses, and importance in wastewater treatment.

Organic Chemical Characteristics

Organic compounds or organics are normally composed of combinations of carbon, hydrogen, and oxygen, together with nitrogen in some cases. Many organics are soluble in water. Most natural organics consist of decay products of organic solids, synthetic organics are usually the result of wastewater discharges or agricultural practices. The organic matter in wastewater typically consists of proteins (40–60%), carbohydrates (25–50%), and oils and fats (8–12%). Along with the proteins, carbohydrates, fats and oils, and urea, wastewater typically contains small quantities of a very large number of different synthetic molecules, whose structures range from simple to complex.

Organic constituents of interest in wastewater are classified as aggregate and individual. Aggregate organic constituents are comprised of a number of individual compounds that cannot be distinguished separately. Laboratory methods commonly used today to measure aggregate

Table 1. Common Analyses Used to Assess the Physical Constituents in Wastewater

Test[a]	Abbreviation/Definition	Use or Significance of Test Results
Total solids	TS	To assess the reuse potential of wastewater and to determine the most suitable type of operations and processes for treating it
Total volatile solids	TVS	
Total fixed solids	TFS	
Volatile suspended solids	VSS	
Fixed suspended solids	FSS	
Total dissolved solids	TDS (TS-TSS)	
Volatile dissolved solids	VDS	
Total fixed dissolved solids	FDS	
Settleable solids		To determine those solids that will settle by gravity in a specified time period
Particle size distribution	PSD	To assess the performance of treatment processes
Turbidity	NTU[b]	Used to assess the quality of treated wastewater
Color	Light brown, gray, black	To assess the condition of wastewater (fresh or septic)
Transmittance	% T	Used to assess the suitability of treated effluent for effluent disinfection
Odor	TON[c]	To determine if odors will be a problem
Temperature	°C or °F	Important in the design and operation of biological processes in treatment facilities
Density	ρ	
Conductivity	EC	Used to assess the suitability of treated effluent for agricultural applications

[a] Details of various tests may be found in Reference 1.

[b] NTU = nephelometric turbidity unit.

[c] TON = threshold odor number.

Table 2. Common Analyses Used to Assess the Inorganic Chemical Constituents in Wastewater

Test[a]	Abbreviation/Definition	Use or Significance of Test Results
Free ammonia	NH_3	Used as a measure of nutrients present and the degree of decomposition in the wastewater; the oxidized form can be taken as a measure
Organic nitrogen	Org N	
Total Kjeldahl nitrogen	TKN (Org N + NH_4^+)	
Nitrites	NO_2^-	
Nitrates	NO_3^-	
Total nitrogen	TN	
Inorganic phosphorus	Inorg P	
Total phosphorus	TP	
Organic phosphorus	Org P	
pH	$pH = -\log[H+]$	A measure of the acidity or basicity of an aqueous solution
Alkalinity	$HCO_3^- + CO_3^{2-} + OH^- - H^+$	A measure of the buffering capacity of wastewater
Chloride	Cl^-	To assess the suitability of wastewater for agricultural reuse
Sulfate	SO_4^{2-}	To assess the potential for the formation of odors and may impact the treatability of waste sludge
Metals	As, Cd, Ca, Cr, Co, Cu, Pb, Mg, Hg, Mo, Ni, Se, Na, Zn	To assess the suitability of wastewater for reuse and for toxicity effects in treatment. Trace amounts of metals are important in biological treatment
Specific inorganic elements and compounds		To assess the presence or absence of specific constituents
Various gases	O_2, CO_2, NH_3, H_2S, CH_4	To assess the presence or absence of specific gases

[a] For test details, see Reference 1.

Table 3. Common Analyses Used to Assess the Organic Chemical Constituents in Wastewater

Test[a]	Abbreviation/Definition	Use or Significance of Test Results
Five-day carbonaceous biochemical oxygen demand	$CBOD_5$	A measure of the amount of oxygen required to stabilize a waste biologically
Ultimate carbonaceous biochemical oxygen demand	UBOD (also BOD_U, BOD_L)	A measure of the amount of oxygen required to stabilize a waste biologically
Nitrogenous oxygen demand	NOD	A measure of the amount of oxygen required to oxidize biologically the nitrogen in the wastewater to nitrate
Chemical oxygen demand	COD	Often used as a substitute for the BOD test
Total organic carbon	TOC	Often used as a substitute for the BOD test
Specific organic compounds and classes of compounds	MBAS[b], CTAS[c]	To determine the presence of specific organic compounds and to assess whether special design measures are needed for removal

[a] For test details, see Reference 1.
[b] MBAS = Methylene blue active substances
[c] CTAS = cobalt thiocyanate active substances

organic matter (typically greater than 1 mg/L) in wastewater include BOD, COD, and TOC. Both aggregate and individual organic constituents are of great significance in the treatment, disposal, and treatment of wastewater. Table 3 illustrates common organic chemical constituents, their analyses, and importance in wastewater treatment.

BIOLOGICAL CHARACTERISTICS

The biological characteristics of wastewater are of fundamental importance in controlling diseases caused by pathogenic organisms of human origin and because of the extensive and fundamental role played by bacteria

Table 4. Common Analyses Used to Assess the Biological Constituents in Wastewater

Test[a]	Abbreviation/Definition	Use or Significance of Test Results
Coliform organisms	MPN (most probable number)	To assess the presence of pathogenic bacteria and the effectiveness of the disinfection process
Specific microorganisms	Bacteria, protozoa, helminths, viruses	To assess the presence of specific organisms in connection with plant operation and for reuse.
Toxicity	TU_a and TU_c,	Toxic unit acute, toxic unit chronic

[a] For test details, see Reference 1.

and other microorganisms in decomposing and stabilizing organic matter in nature and wastewater treatment plants. Most species of pathogens such as bacteria, viruses, and protozoa can survive in wastewater and maintain their infectious capabilities for a significant period of time. Table 4 illustrates common biological characteristics, their analyses, and importance in wastewater treatment.

BIBLIOGRAPHY

1. APHA. (1998). *Standard Methods for the Examination of Water and Wastewater*, 20th Edn. American Public Health Association, Washington, DC.

READING LIST

Metcalf & Eddy. (2003). *Wastewater Engineering: Treatment, Disposal, Reuse*, 4th Edn. McGraw-Hill, New York, pp. 10, 30, 31.

CHEMICALLY ENHANCED PRIMARY TREATMENT OF WASTEWATER

Frédéric J.F. Chagnon
Donald R.F. Harleman
Massachusetts Institute of Technology
Cambridge, Massachusetts

Chemically enhanced primary treatment (CEPT) is a wastewater treatment process in which small doses of chemical compounds are added to wastewater in order to increase the operating efficiency of sedimentation basins. The chemicals cause the suspended particles to clump together via the processes of coagulation and flocculation. The particle aggregates, or flocs, settle faster and thus enhance the treatment efficiency, which is measured as the removal of solids, organic matter, and nutrients from the wastewater. The chemicals utilized in CEPT are the same ones commonly added in potable water treatment (e.g., metal salts and/or organic polymers).

CEPT allows the sedimentation basins to operate at twice the overflow rate (defined as the flow per unit area in the sedimentation basin) of conventional primary treatment, while maintaining higher removal rates of total suspended solids (TSS) and biochemical oxygen demand (BOD). (The increase in BOD removal by CEPT is usually larger than that of suspended solids because of the removal through precipitation of colloidal BOD.) The treatment infrastructure is thus smaller, which reduces capital costs. The capital costs of a CEPT plant are approximately 25% of the capital costs of a conventional secondary treatment plant of equal capacity, while the operational costs are one-half those of a conventional secondary treatment plant (1). Additionally, CEPT provides the opportunity for either reducing the size of subsequent biological treatment units, or increasing the capacity of existing conventional treatment plants.

Table 1. Comparison of Removal Efficiencies (1)

Treatment Method	TSS, %	BOD, %
Conventional primary treatment	55	35
Conventional primary + biological secondary treatment	90	85
Chemically enhanced primary treatment	85	55

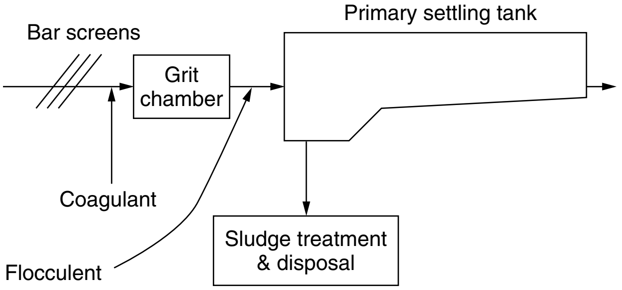

Figure 1. Schematic of conventional primary treatment and CEPT. (The addition of a flocculent in the form of organic polymers is optional.)

CEPT is almost as efficient as secondary treatment with respect to removal efficiencies (Table 1). CEPT may be implemented using a dedicated “CEPT tank” (i.e., a settling tank specially designed for CEPT) or by retrofitting a conventional primary treatment facility or stabilization ponds (2–5). The latter two incarnations of CEPT are relevant when upgrading overloaded existing systems (6).

A conventional primary treatment process consists of bar screens, a grit chamber, and a settling tank (Fig. 1). To upgrade a conventional primary treatment facility to a CEPT facility, all that is needed is the addition of a chemical coagulant (and optionally a flocculent). With CEPT’s high surface overflow rate, the sedimentation basins will not need to be large when compared to conventional primary sedimentation basins (7).

CEPT has been used for over one hundred years, yet it is not as commonly found as would be expected upon analysis of its performance. The misconception is that CEPT dramatically increased sludge production. However, CEPT is used today with a minimal coagulant dosage (10–50 mg/L), and the chemicals themselves make only a slight contribution to the total sludge production. The greatest portion of the increase of sludge production is due to the increased solids removal in the settling tank.

CEPT does not preclude secondary or tertiary treatment. It makes any subsequent treatment smaller and less costly due to the increased efficiency. CEPT is a relatively simple technology providing a low-cost and effective treatment, which is easily implemented over existing infrastructure (6,8,9). CEPT is also one of the least expensive wastewater treatment processes in which the effluent can be effectively disinfected.

BIBLIOGRAPHY

1. National Research Council. (1992). *Wastewater Management in Urban Coastal Areas*. National Academy Press, Washington, DC.

2. Ødegaard, H., Balmer, P., and Hanaeus, J. (1987). Chemical precipitation in highly loaded stabilization ponds in cold climates: Scandinavian experiences. *Water Sci. Technol.* **19**(12): 71–77.
3. Hanaeus, J. (1991). Chemical precipitation in ponds for wastewater treatment. *Vatten* **47**(2): 108–116.
4. Hanaeus, J. (1991). Sludge accumulation in ponds for wastewater treatment using alum precipitation. *Vatten* **47**(3): 181–188.
5. Hanaeus, J. (1991). *Wastewater Treatment By Chemical Precipitation in Ponds*. [Ph.D. thesis]. Lulea University of Technology, Lulea, Sweden.
6. Harleman, D.R.F. and Murcott, S.E. (1992). *Upgrading and Multi-Stage Development of Municipal Wastewater Treatment Plants: Applicability of Chemically Enhanced Primary Treatment*. Tech. rep., World Bank.
7. Harleman, D.R.F., Harremöes, P., and Yi, Q. (1997). Hong Kong Harbor cleanup: international panel reviews plans for treatment upgrade. *Water Environ. Technol.*, 47–50.
8. Harleman, D.R.F. and Murcott, S. (2001). CEPT: challenging the status quo. *Water* **21**: 57–59.
9. Harleman, D.R.F. and Murcott, S. (2001). An innovative approach to urban wastewater treatment in the developing world. *Water* **21**: 44–48.

GETTING OUR CLEAN WATER ACT TOGETHER

DAN WOLZ
City of Wyoming, Michigan
Wyoming, Michigan

Sometimes, having ideas or understanding ideas takes a long time. Maybe it is like fine wine. This idea was born in the current business climate of mission statements, defining the product, public perception, reality and principles and practices of Total Quality Management (TQM—which has reached the stage of maturity where an individual can now be a Certified Master TQMer). The idea: our industry's main product is clean water. No duh! (Not a TQM phrase.) This may not be a great revelation to anyone except maybe me. So what is my point? It is simply a question, "Why do we call them wastewater plants when the main product is clean water?" This could be rephrased to say "What cookie factory calls itself the flour, lard, baking soda, raisin, cinnamon, chocolate chip factory'?" This is the topic of this article.

As I began to think about this idea, I kept asking myself why we did not call the end product "clean water" right from the start. I can understand the use of the term "wastewater treatment plant" but not the shortened term "wastewater plant" now in common use. The subtle implication is that somehow "wastewater" is the product. It was starting to bother me. The message is inconsistent. There is something subliminally wrong with giving a negative connotation to something that does not deserve it. The well-worn axiom may fit here: If it walks like a duck and swims like a duck and quacks like a duck, it's a duck. I came to the conclusion it was time to call a duck "clean water." We spend a great deal of time and money trying to convince the dozens of students and the visiting public how well we clean it—just to call it "wastewater."

Before proceeding, I need to thank the author or authors of the word "wastewater" (you know, the word that is never recognized by your spell checker until you add it). Incidentally, if Virgil Langworthy, a Michigan resident and long-time water professional, is not the originator of the word, he can certainly be credited with being one of the early purveyors of it in the early 1960s. "The reason for the change" he said, "was the bad perception of the term 'sewage.'" And this was before TQM.

The word "wastewater" is far better than "sewerage" or "sewage" and was a great addition to our descriptive word quiver for our evolving industry language. These words have served us well over recent years, but the time has come to add and use the term "clean water" as the focus or our business. I must hasten to add that these traditional words should not be abandoned but further defined and applied when these new definitions warrant their use. "Wastewater" should now properly be redefined "as the used water leaving a household or industry and entering a sewer system." It could be also used as another way to describe plant influent but NOT what leaves a plant.

I read once that traditional Eskimo culture has more than 30 terms in its language to describe snow. The words were needed to describe the varying conditions of the harsh climate. We add more words to language to communicate concept, especially where we spend great amounts of time and study. One additional example would be the in computer industry. How many of us knew what "W.Y.S.I.W.Y.G." was in the early 1980s? (For those unfamiliar with that term, it was an acronym for "What you see is what you get," but it lost its periods and made its way into the dictionary as a genuine word.)

I decided to put action to my thoughts and began to use "clean water plant" in conversation with my co-workers and neighboring plants. Surprising resistance to the idea came from some of my brothers in drinking water plants who said things like "You are the 'dirty water people' and we are the 'clean water people'." The other comment was "if you use the term 'clean water' at a wastewater plant, the public will be confused." My response is simply to point out how frequently after a tour of the wastewater plant many people comment "I always wondered where my drinking water came from" when they leave. At that point, I feel like a total failure as a public educator. I do not think the public cares about our heterohydrophobic (fear of different kinds of water) distinctions, as long as they are receiving good service.

If there is confusion about the use of the term, it is because we in the clean water industry have failed to appropriate it and use it. The United States Congress gave us the "clean water" title 25 years ago with the passage of the Clean Water Act (CWA) which set discharge standards for WASTEWATER PLANTS since 1972. They then passed the Safe Drinking Water Act (SDWA) a few years later for DRINKING WATER PLANTS. What are we waiting for? Is not more than 25 years long enough to wait and see if the 800-pound gorilla (CWA) sitting in our living room is here to stay permanently? NEWS FLASH: He's *here to stay and while we were waiting, his whole family moved in and one of them, some say, is bigger and uGLIer (GLI is*

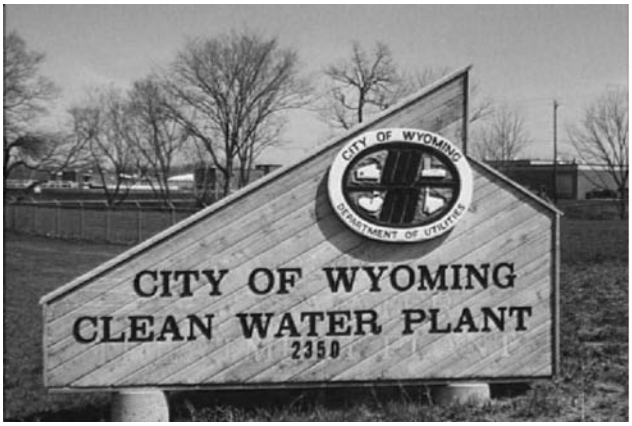

Figure 1. Sign at Wyoming waste water treatment plant.

an intentional reference to the Great Lakes Initiative). And they all want bananas!

Our society wants clean water, and it takes lots and lots of bananas to make it happen and keep it happening. We are giving society clean water now and we need to continue this good work, BUT make sure we acknowledge it is discharged from a Clean Water Plant. We all know bananas do not grow on trees (or is that money), and we do know how demanding gorillas can be. A word to the wise: feed the gorillas and suggest family planning.

In summary, I propose a global renaming of our "wastewater treatment plants" to "clean water plants." The drinking water plants are left with titles such as "safe drinking water plants," "drinking water plants," and/or "water filtration plants." I know of no drinking water plants that use the term "clean water" in their title. So, it is time to drive this idea out of the giant parking lot of ideas whose time has not yet come onto the freeway of ideas in everyday usage. I know the City of Wyoming, Michigan's Clean Water Plant is already on the freeway entry ramp hoping to see you at our clean water destination—the banana plantation (Fig. 1).

INADEQUATE TREATMENT OF WASTEWATER: A SOURCE OF COLIFORM BACTERIA IN RECEIVING SURFACE WATER BODIES IN DEVELOPING COUNTRIES—CASE STUDY: EASTERN CAPE PROVINCE OF SOUTH AFRICA

M.N.B. Momba
C. Mfenyana
University of Fort Hare
Alice, South Africa

INTRODUCTION

Sanitation, in its broadest meaning, refers to the formulation and application of measures designed to protect public health. It also covers the purification of wastewater and sewage before it is returned to the water cycle. Usually, wastewater treatment involves collecting the wastewater in a central, segregated location (the wastewater treatment plant) and subjecting the wastewater to various treatment processes (biological treatment and disinfection of the final effluent). The characteristics of these wastes, which are significant in pollution, are the suspended solids, oxygen demand of the organic matter (BOD), the coliform bacteria, and other pathogenic micro-organisms (1).

The presence of coliform bacteria in drinking water is the most common reason for the violation of the water standards. Coliform bacteria have a long history in water quality assessment, mainly because of their association with fecal pollution and relatively easy and rapid detection. Some members of the group are almost conclusively of fecal origion, whereas others may also multiply in suitable water environments (2,3). The South African General and Special Standards stipulate that treated sewage should comply with a standard of nil fecal coliforms/100 ml (Act 96 of May 18, 1984 No. 9225, Regulation 991). The efficiency of a wastewater treatment plant in removing harmful micro-organisms from the influent is therefore of utmost importance. Failure to remove such micro-organisms will result in harmful pathogens returning into the receiving surface water body. Being a major source of pathogens that are carried in water, monitoring sewage for pathogens has been demonstrated to be an excellent epidemiological tool for determining what diseases may be prevalent in the community at any one moment (4).

In the developing world, much of the population consumes untreated and nonpiped drinking water. Typically, people collect water from any available source and store it in a vessel or a container in the dwelling for domestic and potable use, often without treatment or protection from further contamination. It is therefore not surprising that an estimated 15–30% (5) of community gastrointestinal diseases in developing countries are attributed to unsafe drinking water, with studies indicating similar percentages (15–20%) of waterborne diarrheal diseases (6). These problems could be attributed to the unavailability of proper sanitation methods or when the existing methods are not functioning adequately (7). Severe problems with the microbial quality of water could be expected, especially in rural and informal settlements. In these areas, widespread scarcity, gradual destruction, and increased pollution of water sources exists. The situation is aggravated by the existence of many insufficient, poorly operated or maintained sanitation services (8). Fecal contamination of water supplies by untreated and/or inadequately treated sewage effluents entering rivers and dams that serve as the source of municipal water supplies create conditions for the rapid spread of pathogens, a problem encountered in both the developing and the developed parts of the world (9).

Although the Eastern Cape province of South Africa, where this study was conducted, is predominantly rural in composition, both rural and urban sectors can be found within the province. The Eastern Cape Province is one of South Africa's poorest provinces, where the official poverty rate in 1998 was at least 70%, and only a quarter of all households possessed an inside water tape (10). Although great strides have been made in effort to provide access to clean water, many small villages still lack safe water supplies (11), whereas large communities have water treatment plants, which fail to provide potable water of high quality to their consumers (12). The areas without water supplies use water directly from available and often contaminated sources without any treatment. In both cases, communities are therefore exposed to water-related diseases. This situation also pertains in most of the communities in the poor areas of developing countries.

The aim of this study was to evaluate the efficiency of the urban, semiurban, and the rural wastewater treatment plants for the removal of coliform bacteria in order to establish the relationship between the microbiological quality of the final effluent and that of the receiving water body, which may further influence infection and disease in the community. Our intention was to provide information that could assist water authorities in developing countries to address problems in the management of wastewater treatment plants in terms of microbiological standards of the effluent as set down by official guidelines, such as those in South Africa (13,14). Two of the four wastewater treatment plants used during the present survey, i.e., in Alice and Fort Beaufort, serve predominantly rural areas, whereas the other two wastewater treatment plants, i.e., East London and Dimbaza, serve urban and semiurban communities, respectively.

MATERIALS AND METHODS

Study Sites

Four wastewater treatment plants that serve the Buffalo City and Nkonkobe Municipal areas in the Eastern Cape Province of South Africa were used in the present study. The wastewater treatment plants are located in urban (East London–East Bank Reclamation Works), semiurban (Dimbaza), and rural areas (Alice and Fort Beaufort). The activated sludge system is the biological wastewater treatment used in all plants, followed by chlorination of the final effluent.

The final effluent from the East Bank Reclamation Works is discharged into the Indian Ocean between Nahoon and Eastern Beach at Bats Cave and into a pond for the irrigation of a nearby golf course. Supernatant liquor from the sedimentation tanks is channeled into a fish pond located within the plant premises. The Dimbaza wastewater treatment plant discharges its final effluent into a stream that empties into the Tembisa sewerage dams. The Alice wastewater treatment plant is situated on the banks of the Tyume River, which is also used as the receiving water body for the final effluent from the plant. The final effluent from Fort Beaufort Sewage Works is discharged into the Kat River. All these final effluents or receiving water bodies are often used by the communities for one or various purposes, which include domestic, agricultural, or recreational purposes.

Sample Collection

Wastewater samples were collected weekly from different stages (the raw influent, the aerobic zone, the clarifier, the final effluent, and the receiving water body) of the four plants between August and October 2003. For the microbiological quality analyses, sampling was done aseptically into sterile glass bottles. For the effluent samples, sodium thiosulphate (*ca* 17.5 mg/L) was added to bottles before autoclaving. The samples were then placed in ice bags and transported to the base laboratory at the University of Fort Hare for analyses within 2–4 h after collection.

Analyses of Wastewater and Receiving Water Body Samples

Free Chlorine Residual. The concentrations of free chlorine residual in the treated effluents were determined using a multiparameter ion-specific meter (Hanna BDH-laboratory).

Coliform Counts. The membrane filtration technique was considered, but the method proved to be too sensitive for the analyses of the samples. Consequently, the standard spread plate procedure was used during the study period. Three different selective media, namely Chromocult coliform agar (Merck), Fluorocult *E. coli* 0157: H7 agar (Merck), and Sorbitol–MacConkey agar (Mast Diagnostics) were used for the isolation of coliforms and other *Enterobacteriaceae* (Table 1). The agars were prepared according to the manufacturer's instruction. Water samples were analyzed for the above microorganisms using internationally accepted techniques (15). The enumeration of coliforms was based on the color of colonies corresponding to presumptive coliforms as indicated in Table 1.

Identification of Coliform Isolates. Bacterial colonies from the influent, final effluent, and the receiving water

Table 1. Summary of the Various Bacterial Colonies and Presumptive Coliform Strains Isolated from Different Culture Media

Culture Medium	Presumptive Bacterial Strain	Color of the Colony
Chromocult coliform agar	*Citrobacter freundii*	Salmon to red
	Escherichia coli	Blue to violet
	Salmonella enteritidis	Colorless
Fluorocult *E. coli* 0157:H7 agar	*Proteus Mirabilis*	Brown
	Enterobacter aerogenes	Yellow
	Escherichia coli 0157:H7	Colorless
	Salmonella typhimurium	Yellow with black center
Sorbitol–MacConkey agar	*Escherichia coli 0157:H7*	Colorless

body samples differing in size, shape, and color were randomly selected from different plates and transferred onto the same medium by streak plate technique and incubated at 37 °C for 24 h. These were further isolated on MacConkey agar (Biolab) by the same method and incubated at 37 °C for 24 h, thereafter they were further purified by the same method at least three times using nutrient agar biolab before gram staining was done. Oxidase tests were then conducted on those colonies that were gram negative. The 20E API kit was used for the oxidase-negative colonies and the strips were incubated at 37 °C for 24 h. The strips were then read and the final identification was secured using API LAB PLUS computer software (BioMérieux, Marcy l'Etoile, France).

RESULTS

Concentration of Free Available Chlorine (Residual) in the Final Effluents

The concentrations of the free chlorine residuals fluctuated in all the plants with the exception of the East London wastewater treatment plant (Table 2). Chlorine overdosing occurred during the first two months of the study period (August and September) in the Fort Beaufort plant and in September in the Dimbaza plant. A constant concentration of free chlorine residual was noted during the whole study period in the East London plant, whereas low concentrations of free chlorine residuals were noted in the final effluent of the Alice wastewater treatment plant (Table 2). Although the 1996 South African Guidelines do not specify any standard for the concentration of free chlorine residual in the treated effluent, this study considered those for domestic water supplies, which recommend ranges of 0.3–0.6 mg/l as ideal free chlorine residual concentration and 0.6–0.8 mg/l as good free chlorine residual concentration with insignificant risk of health effects (14). With the exception of the final East London treated effluents, none of the wastewater plant treated effluents complied with the above-mentioned recommended limits for domestic water supplies.

Table 2. Concentrations of Free Chlorine Residual in the Final Effluents from during the Study Period (Ranges) (No. of the Samples for Each Plant; 12)

	Chlorine Residual (mg/l)	
Wastewater Treatment Plant	Ranges	Means
Dimbaza	0.93–2.5	1.85
East London	0.59–0.66	0.58
Alice	0.03–0.66	0.24
Fort Beaufort	0.16–4.33	2.07

Performances of Wastewater Treatment Plants for the Removal of Presumptive Coliforms

The four wastewater treatment plants were investigated for their efficiency for the removal of coliforms. The enumeration of coliforms in each zone of different wastewater treatment plant was done in order to assess whether there was in fact a decrease in the number of presumptive coliforms as the wastewater flowed through the treatment processes, a practice that also served to identify possible malfunctioning in any of the wastewater treatment plant zones.

In general, a gradual removal of presumptive coliforms was observed through the different zones of the wastewater treatment plants. The data, however, reflected variations with regard to both the patterns and the efficiency of each plant for the removal of coliform micro-organisms (Figs. 1 and 2).

In the influent samples, both the Dimbaza and East London wastewater treatment plants had higher mean coliform counts (ranging between 4.6 log cfu/ml and 6.6 log cfu/ml for Dimbaza and between 5.3 log cfu/ml and 7 log cfu/ml for East London) than those found in Alice (5–6 log cfu/ml) and Fort Beaufort (4.8–6.2 log cfu/ml). This observation results from the fact that both Dimbaza and East London are urban and semiurban areas, respectively, and they are also both industrialized and, hence, much more densely populated than Alice and Fort Beaufort. Results in Figs. 1 and 2 also show that, although in Dimbaza, East London, and Alice wastewater treatment plants, a noticeable decrease in the mean coliform counts was observed in the clarifier zone, high mean coliform counts were still noted in the Fort Beaufort wastewater treatment plant, which was because of the absence of a functional clarifier.

The mean ranges for coliforms (0–0.2 cfu/ml) in the Dimbaza wastewater treatment's final effluent complied with the South African General and Special Standards, which stipulate that treated sewage effluents must have a standard of nil fecal coliforms (Act 96 of May 18, 1984 No. 9225, Regulation 991). The Dimbaza's effluent was also within the limits set for agricultural purposes (irrigation), which are ≤10,000 count/100 ml for fecal coliform and ≤1 counts/100 ml for *E. coli* (16). In East London, Alice, and Fort Beaufort wastewater treatment plants, the mean coliform ranges in the treated effluent samples were 1.5–2.9 log cfu/ml, 0.8–2.6 log cfu/ml, and 0.6–2.9 log cfu/ml, respectively. These results suggest that the general microbiological qualities of these effluents did not comply with the limits set by the South African Authorities, especially in terms of fecal coliforms.

Generally, an increase occurred in presumptive coliform counts from the final effluent samples to the receiving water body samples of the Dimbaza, Alice (presumptive *E. coli* being the exception), and Fort Beaufort wastewater treatment plants, whereas a decrease occurred in the mean counts of all presumptive coliforms from the East London water receiving body samples (Fig. 2). The means for coliforms in the receiving water bodies ranged between 0.2 and 1.5 log cfu/100 ml for Dimbaza, between 0 and 2.5 log cfu/100 ml for East London plant, between 0.25 and 3 log cfu/100 ml for the Alice plant, and between 1.5 and 3.25 log cfu/100 ml for Fort Beaufort. The results suggest that the receiving water bodies of all the plants did not comply with the limits set for domestic (0–10 counts/100 ml for total coliform, 0 counts/100 ml for fecal coliforms) and recreational (0–130 counts/100 ml for fecal coliform, *E coli* 0–130 counts/100 ml) use (13,14,17).

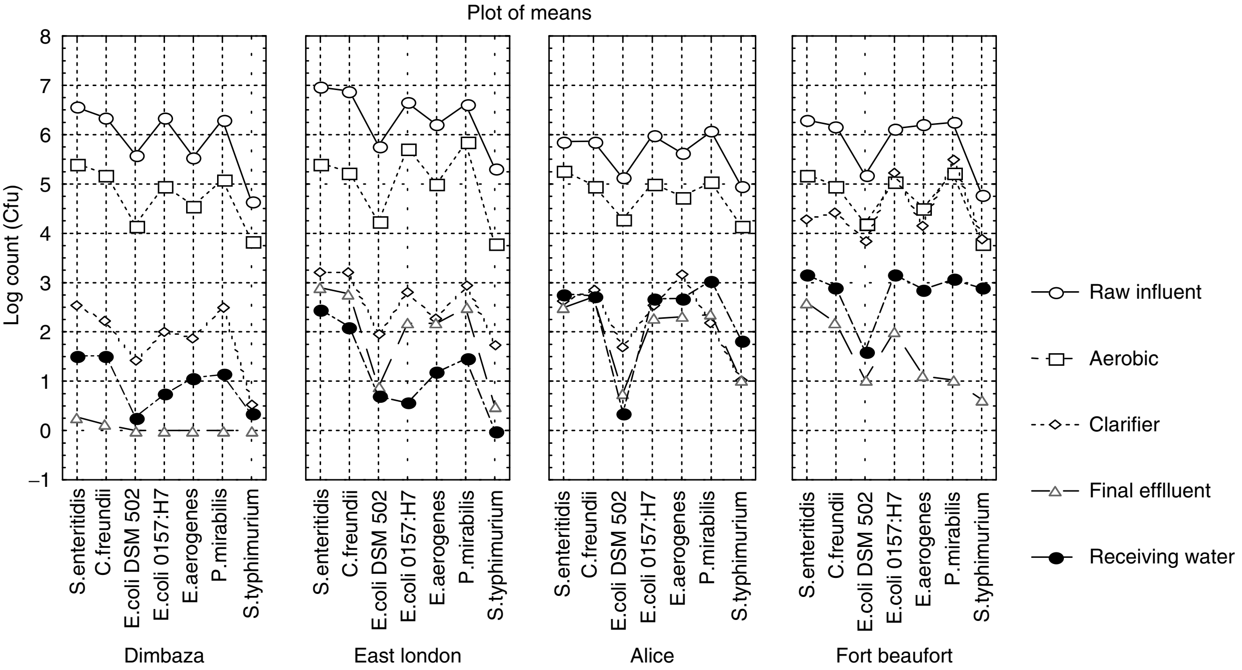

Figure 1. Plot of mean counts for presumptive coliforms obtained from the different zones of the Dimbaza, East London, Alice, and Fort Beaufort wastewater treatment plants and their water receiving bodies.

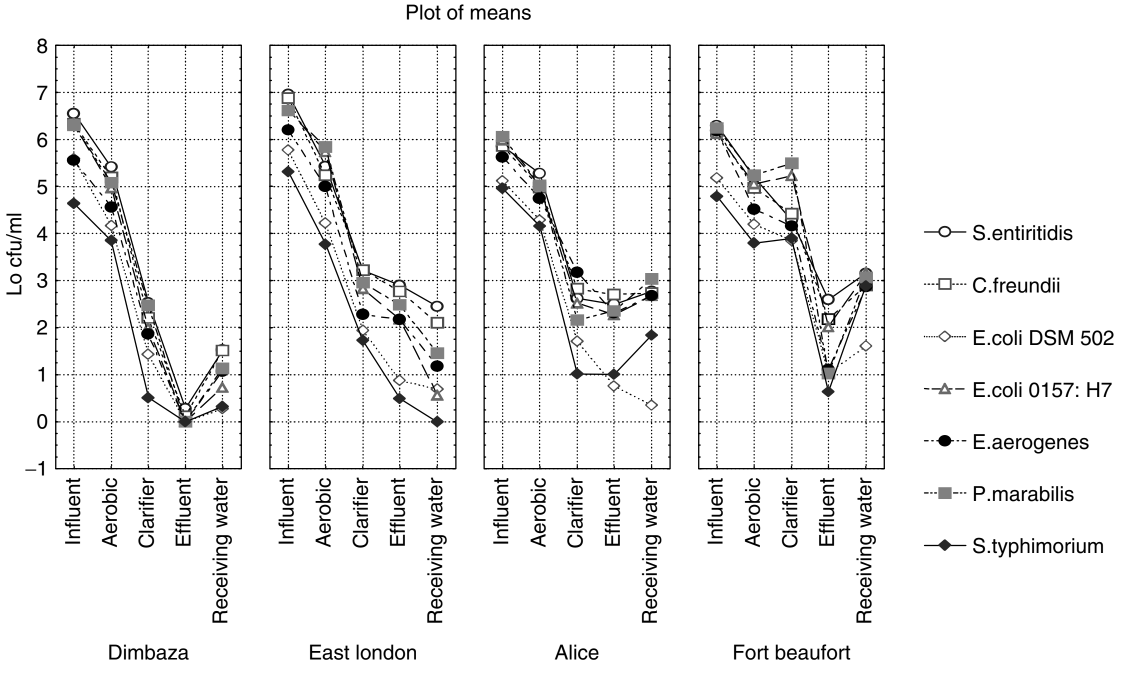

Figure 2. Performance of the different wastewater treatment plants for the removal of the presumptive coliform bacteria.

Identification and Bacterial Isolates

According to the API 20E system, the oxidase test and the gram staining test, species of *Proteus* were not among the identified organisms that belonged to the presumptive coliform species (Table 3). However, species of *Escherichia* (Dimbaza: influent; East London: receiving water body; Alice: influent and final effluent; Fort Beaufort: influent, final effluent, and receiving water body), *Citrobacter* (Alice: influent, final effluent, and receiving water body; East London: influent and receiving water body; Alice: influent; Fort Beaufort: influent and final effluent), and *Salmonella*

Table 3. Bacterial Isolates Identified from the Influent, Effluents, and Receiving Water Body Samples Analyzed (Number of Samples: 12 for each Stage of the Plant)

Wastewater Plants	Water Sources	Bacterial Isolate Identified
Dimbaza	Influent	*Aeromonas hydrophila, Citrobacter freundii, Escherichia coli, Klebsiella ornithinolytica, Vibrio cholerae, Vibrio parahaemolytica, Serratia liquefaciens.*
	Effluent	*Citrobacter freundii*
	Receiving body	*Citrobacter freundii, Enterobacter aerogenosa, Salmonella arizonae, Serratia odorifera*
East London	Influent	*Aeromonas hydophila, Chromo violaceum, Citrobacter freundii, Enterobacter aeogenosa, Klebsiella ornithinolytical, Pseudomonas putida, Vibrio fluvialis, Serratia ficaria, Serratia odorifera, Chromo violaceum.*
	Effluent	*Aeromonas hydophila, Chromo violaceum, Klebsiella ornithinolytica, Pseudomonas putida, Serratia odorifera, Serratia ficaria, Vibrio fluvialis.*
	Receiving body	*Aeromonas hydrophila, Citrobacter freundii, Escherichia coli, Klebsiella oxytoca, Klebsiella ornithinolytica, Morganella morgani, Vibrio fluvialis*
Alice	Influent	*Aeromonas hydophila, Citrobacter freundii, Enterobacter cloacae, Escherichia coli, Klebsiella ornithinolytica, Klebsiella oxytoca, Klebsiella ozaenae, Kluyvera* spp., *Pseudomonas putida, Vibrio cholerae*
	Effluent	*Aeromonas hydophila, Escherichia coli, Klebsiella ornithinolytica, Enterobacter amnigenus, Klebsiella ozaenae, Klebsiella oxytoca, Kluyvera* spp., *Pseudomonas putida*
	Receiving body	*Aeromonas hydophila, Enterobacter amnigenus, Klebsiella ornithinolytica, Klebsiella oxytoca, Pseudomonas aerogenosa, Pseudomonas putida*
Fort Beaufort	Influent	*Aeromonas hydrophila, Citrobacter freundii, Escherichia coli, Klebsiella ornithinolytica, Klebsiella oxytoca, Pasteurella pneumoniae, Salmonella arizonae, Serratia odorifera, Vibrio fluvialis.*
	Effluent	*Aeromonas hydrophila, Citrobacter freundii, Escherichia coli, Klebsiella oxytoca, Vibrio fluvialis*
	Receiving body	*Aeromonas hydrophila, Escherichia coli, Enterobacter cloacae, Pseudomonas putida, Shewan putrefaciens, Vibrio fluvialis*

(Dimbaza: receiving water body; Fort Beaufort: influent) were identified in some zones of the plants. Other organisms belonging to 11 other different species were also identified from different zones of the plants. These species included *Aeromonas* (Dimbaza: influent; Alice, East London, and Fort Beaufort: influent, final effluent, and receiving water body), *Klebsiella* (Dimbaza: influent; East London and Fort Beaufort: influent, final effluent, and receiving water body; Fort Beaufort: influent and final effluent), *Serratia* (Dimbaza: influent and receiving water body; East London: influent and effluent; Fort Beaufort: influent), *Enterobacter* (Dimbaza: receiving water body; East London: influent; Alice: influent, effluent, and receiving water body; Fort Beaufort: receiving water body), *Pseudomonas* (East London: influent and effluent; Alice: influent, effluent, and receiving water body; Fort Beaufort: receiving water body), *Vibrio* (Dimbaza: East London and Fort Beaufort: influent; Alice: influent, effluent, and receiving water body), *Chromo* (East London: influent and effluent), *Morganella* (East London: receiving water body), *Kluyvera* (Alice: influent and effluent), *Erwinia* (Alice: receiving water body), and *Shewanella* (Fort Beaufort: receiving water body) (Table 3). The efficiency of the Dimbaza wastewater treatment plant was also confirmed with the identification tests when compared with the other three plants. Among the 7 species identified from the Dimbaza's influent, only *Citrobacer freundii* was still prevalent in the treated final effluent and also in the receiving water body. It is obvious from the results in Table 3 that the final effluent was not the source of organisms such *Enterobacter aerogenosa, Salmonella arizonae*, and *Serratia odorifera* identified from the receiving water body. *Aeromonas hydrophila*, which was found to be the dominant strain in all influents, was still prevalent in the East London, Alice, and Fort Beaufort's treated effluents and received water bodies.

DISCUSSION

Access to a clean, pathogen-free water supply is a major priority for any community if it is to remain disease-free. In today's highly urbanized society, the best way to achieve this objective is to recycle water by treating used or wastewater through a treatment plant system. The efficiency of wastewater treatment plants in removing pathogenic micro-organisms from their final effluents can only be achieved by disinfection. According to White (18), the most prevalent practice of disinfection is free chlorine, which is also the practice in many developing countries. Disinfection with chlorine is always influenced by indicator organism concentration, disinfection concentration, contact time, temperature, and pH. The availability of the free chlorine residuals in the final effluents gives an indication of the efficacy of the disinfection process and thus a rapid indicator of the probable microbiological safety of the treated effluent.

In terms of the concentration of free residual in the effluents, results indicated the availability of free chlorine residual in all the wastewater plants, although in some plants (Fort Beaufort and Dimbaza, Table 2) these concentrations exceeded the limit recommended by water authorities, which are in the range of 0.3–0.6 mg/l as ideal free chlorine residual concentration and 0.6–0.8 mg/l as good free chlorine residual concentration for insignificant risk of health effects (14). In contrast to the other plants, the free chlorine concentrations complied with (East London) or were lower than the standards (Alice) (Table 2).

Although free chlorine was available in the various final effluents, with the exception of the Dimbaza plant, which showed a removal of coliform organisms at 85.71%, the results suggest that the microbiological quality of the final effluent in East London, Alice, and Fort Beaufort plants exceeded the standards for domestic and recreational use. The results also indicated that the effluents from all the plants were also one of the sources of coliforms in the receiving water bodies. In the Dimbaza plant, the *Citrobacter freundii* found in the final effluent also appeared in the receiving water body. In the Alice, Fort Beaufort, and East London wastewater treatment plants, 62.5%, 60%, and 43% of the organisms found in the final effluent were also prevalent in the receiving water body, respectively (Table 3). The results suggest that the microbiological quality of the effluents examined poses a serious health risk to the community, which is consistent with the findings of previous investigators, such as Muyima and Ngcakani (9), who pointed out that fecal contamination of drinking water supplied by untreated or inadequately treated sewage effluents entering rivers and dams that serve as the source of municipal water supplies creates conditions for the rapid spread of pathogens. It also became clear from the findings that the main reason behind the inefficiencies shown by the wastewater treatment plants for the removal of the coliforms stem from inadequate disinfection practices and/or inadequate maintenance of the facilities, as observed with the Fort Beaufort plant. Although the East London wastewater treatment plant was technically well equipped with five aerators and six clarifiers, this equipment was under repair during the study period and not all the clarifiers were in a working condition. This fact has been reiterated by Pearson and Idema (19), who, commenting on the disinfection of effluents in many cases in the developing countries, stated that a high level of reliability of water supply schemes, particularly the treatment process, is the exception rather than the rule. The authors further pointed out that various factors such as cost, operator training, and problems with maintenance of the infrastructure could be contributory factors to these problems.

A number of different species of potentially pathogenic micro-organisms were isolated from both the final effluents and the receiving water bodies of the different wastewater treatment plants. Although differences existed with the types of species isolated from the various effluents, potentially pathogenic organisms were isolated from all the effluent and receiving water body samples in various degrees of preponderance (Table 3). *Aeromonas hydrophila*, for example, was found as the dominant micro-organism identified from final effluents and receiving water bodies in East London, Alice, and Fort Beaufort wastewater plants. *Aeromonas* can be infectious, producing focal or systemic infections of varied severity. The most common clinical infection that has been associated with this organism is diarrhea (20–23). The preponderance of *A. hydrophila* in the final effluents and receiving water bodies is cause for much concern, because it is common for communities in the Eastern Cape to use the effluent receiving water body such as a river (in Alice and Fort Beaufort), a dam (Dimbaza), or ocean beach (East London) as a water source for drinking, bathing, washing (i.e., clothes, dishes, etc.), or recreational purposes (i.e., swimming) and the final effluent as a water source for irrigation purposes. The microbiological quality of these water sources could influence infection and disease in the communities.

CONCLUSIONS

A relationship existed between the microbiological quality of the final effluent and that of the receiving water body, and the relationship was such that the better the quality of the final effluent, the better the quality of the receiving water body. From the results, it could also be deduced that the appearance of coliform bacteria in the receiving water body, although no evidence exists of such in the final effluent, means that the source of pollution was from a source outside of the wastewater treatment process.

The disinfection practices and guidelines in terms of chlorine residual concentrations were found not to be sufficient for the removal of coliforms from the effluent. Although the chlorine residual concentration may fall within the recommended limits for no risk, the occurrence of coliforms could still be detected in the final effluent. This study, therefore, suggests that other factors, such as the chlorine demand of the water, the contact time, the initial indicator organisms, or the use of powerful disinfectants, be considered by the operators in wastewater treatment plants for the treatment of the effluents before they are discharged into the receiving water bodies.

Acknowledgments

The authors wish to express gratitude to the National Research Foundation who sponsored this investigation and to the late Mr. R.A. Agbandzi from the University of Fort Hare's Programme Unit of Statistics for his help during the study period; may his soul forever rest in peace.

BIBLIOGRAPHY

1. Eckenfelder, Jr., W.W. and O'Connor, D.J. (1961). *Biological Waste Treatment*. Pergamon Press, New York, pp. 1, 7.
2. Grabow, W.O.K. (1996). Waterborne diseases: Update on water quality assessment and control. *Wat. SA.* **22**(2): 193–197.
3. Mnqumevu, B.V. (1998). Detection of faecal coliform bacteria and heterotrophic plate count bacteria attached to household containers during storage of drinking ground water in rural communities. B.Sc. (Honors) Thesis, Department of Biochemistry and Microbiology, University of Fort Hare, Alice, South Africa.
4. Moore, B. (1948). Monthly bulletin. *Ministry of Health* **7**: 241.
5. Payment, P., Richardson, L., Siemiatycki, J., Dewar, M., and Franco. (1991). A randomized trial to evaluate the risks of

gastrointestinal diseases due to consumption of drinking water meeting current microbiological standards. *Am. J. Publ. Hlth.* **81**: 703–708.

6. Quick, R. et al. (1999). Diarrhea prevention in Bolivia through point-of-use disinfection and safe storage: a promising new strategy. *Epi. Infect.* **122**: 83–90.
7. Fatoki, O.S., Gogwana, P., and Ogunfowokan, A.O. (2003). Pollution assessment in the keiskamma river and in the impoundment downstream. *Water S.A.* **1**(2): 183.
8. Venter, S.N. (2001). Microbial water quality in the 21st Century. *S.A. Water Bull.* January/February.
9. Muyima, N.Y.O. and Ngcakani, F. (1998). Indicator Bacteria And Regrowth Potential Of The Drinking Waterin Alice, Eastern Cape, Department of Biochemistry and Microbiology, University of Fort Hare. *Water SA.* **24**: 1.
10. Mey, J. (1998). *Poverty and Inequality in South Africa*. Report prepared for the Executive Deputy President and Inter-Ministerial Committee for poverty and inequality, Durban.
11. Momba, M.N.B., Malakate, V.K., and Theron, J. (2004). Abundance of pathogenic Escherichia.coli, Salmonella typhimurium and Vibrio cholerae in Nkonkobe drinking water sources. *J. Water Health*. Submitted for peer review.
12. Momba, M.N.B., Tyafa, Z., and Makala, N. (2004). Rural water treatment plants fail to provide potable water to their consumers: the Alice water treatment plant in the Eastern Cape Province of South Africa. *S. African J. Sci.* **100**: 307–310.
13. Department of Water Affairs and Forestry (DWAF). (1996). *South African Water Quality Guidelines*, 2nd Edn. Vol. 1: Domestic Use. Pretoria, South Africa.
14. Department of Water Affairs and Forestry (DWAF), Department of Health and Water Research Commission. (1998). *Quality of Domestic Water Supplies*. Vol. 1: Assessment Guide. *Water research Commission* No: TT101/98 ISBN No: 1 86845 416 9.
15. American Public Health Association. (1998). *Standard Methods for the Examination of Water and Wastewater*, 20th Edn. AWWA. Water Environmental Federation, Washington, DC.
16. Department of Water Affairs and Forestry (DWAF). (1996). *South African Water Quality Guidelines*, 2nd Edn. Vol. 4: Agricultural Use and Irrigation. Pretoria, South Africa.
17. Department of Water Affairs and Forestry (DWAF). (1996). *South African Water Quality Guidelines*, 2nd Edn. Vol. 2: Recreational Use. Pretoria, South Africa.
18. White, G.C. (1992). *Handbook of Chlorination and Alternative Disinfectants*. Van Nostrand, New York.
19. Pearson, I. and Idema, G. (1998). An assessment of common problems associated with drinking water disinfection. *WRC* No. 649/1/98.
20. Davis, W.A., Kane, J.G., and Garagusi, V.F. (1978). Human *Aeromonas* infections: a review of the literature a case report of endocarditis. *J. Med.* **57**: 267–277.
21. Agbonlahor, D.E. (1983). The role of *Aeromonas* in acute diarrhoeal diseases in Nigeria. *Cent. Afric. J. Med.* **29**: 142–145.
22. Bhat, P., Shantnakumari, S., and Rajan, D. (1984). The characterization significance of *pleisionomasshigelloides* and *aeronomas hydrophilla* isolated from an epidemic of diarrhea. *Indian J. Med. Res.* **52**: 1051–1060.
23. Watson, I.M., Robinson, J.O., Burke, V., and Grace, M. (1985). Invasiveness of *Aeromonas* spp in relation to biotype. *J. Clinic. Microbiol.* **22**: 48–51.

DENITRIFICATION IN THE ACTIVATED SLUDGE PROCESS

Michael H. Gerardi
Linden, Pennsylvania

Denitrification in the activated sludge process is the use of nitrate ions (NO_3^-) by facultative anaerobic (denitrifying) bacteria to degrade soluble carbonaceous BOD (cBOD). Although denitrification most often is observed in the secondary clarifier, denitrification occurs whenever an anoxic condition exists. Denitrification in the secondary clarifier often is referred to as "clumping," "rising sludge," and "dark sludge rising." Clumping is highly undesired, because it results in poor compaction of solids, loss of solids, increased operational costs, and possible permit violations.

Besides the anoxic condition known as clumping, an undesired anoxic condition can occur in the sewer system and other treatment units and a desired anoxic condition can be produced in a denitrification tank or designated anoxic tank. A denitrification tank is used to satisfy a total nitrogen discharge limit, while a designated anoxic tank is used to improve process performance. Desired process performance includes destruction of undesired filamentous organism growth and strengthening of floc particles.

Unless an industry discharges nitrate ions to the activated sludge process, activated sludge processes that denitrify usually nitrify. Industrial discharges that may contain nitrate ions include pretreated leachate, pretreated meat processing wastewater, meat processing wastewater containing flavoring compounds, and steel mill wastewater. The presence of nitrate ions in the sewer system is highly undesired. Here, nitrate ions permit the rapid degradation of soluble cBOD. The degradation of soluble cBOD results in a decrease in the quantity of cBOD in the influent to a wastewater treatment plant. This decrease in cBOD makes it difficult for the operator of an activated sludge process to achieve an 85% removal efficiency for cBOD.

There are four factors that must be satisfied in order for an anoxic condition to occur. First, an abundant and active population of denitrifying bacteria must be present. Second, nitrate ions must be present. Third, free molecular oxygen (O_2) must be absent or an oxygen gradient must present. Fourth, a source of soluble cBOD must be present.

Denitrifying bacteria enter an activated sludge process through fecal waste and inflow and infiltration (I/I) as soil and water organisms. Approximately 80% of the bacteria in the activated sludge process are facultative anaerobes (Table 1). These organisms are present in millions per milliliter of bulk solution and billions per gram of floc particle, and most denitrifying bacteria reproduce every 15–30 min. The most significant genera of bacteria containing denitrifying species are *Alcaligens*, *Bacillus*, and *Pseudomonas*.

Denitrifying bacteria can use either free molecular oxygen or nitrate ions to degrade soluble cBOD. However, the bacteria prefer free molecular oxygen to nitrate ions to degrade soluble cBOD, because the use of free molecular

Table 1. Genera of Activated Sludge Bacteria that Contain Denitrifying Species

Achromobacter
Acinetobacter
Agrobacterium
Alcaligenes
Bacillus
Chromobacterium
Corynebacterium
Denitrobacillus
Enterobacter
Escherichia
Flavobacterium
Glucononbacer
Halobacterium
Hyphomicrobium
Kingella
Methanonas
Moraxella
Neisseria
Paracoccus
Propionibacterium
Pseudomonas
Rhizobium
Rhodopseudomonas
Spirillum
Thiobacillus
Xanthomonas

oxygen to degrade soluble cBOD results in the production of more offspring (bacterial cells or sludge) and more useable cellular energy than the use of nitrate ions. For example, the use of free molecular oxygen to degrade 1 lb of sugar results in the production of approximately 0.6 lbs of bacterial cells or sludge. The use of nitrate ions to degrade 1 lb of sugar in the activated sludge process results in the production of approximately 0.4 lbs of bacterial cells or sludge.

Nitrate ions may enter an activated sludge process from an industrial discharge. Typically, the nitrate ions used by denitrifying bacteria are produced in the activated sludge process through nitrification. Nitrate ions produced in the activated sludge process may be discharged to a secondary clarifier, denitrification tank, thickener, and anaerobic digester. Nitrate ions also may be recycled in the wastewater treatment plant to the headworks, primary clarifier, or influent to the activated sludge process. Depending on operation conditions in each treatment unit, the nitrate ions received in these units may undergo denitrification and may contribute to operational problems.

Nitrate ions are used by denitrifying bacteria in the absence of free molecular oxygen or in the presence of an oxygen gradient. An oxygen gradient is established when the dissolved oxygen concentration outside the solids is <1 mg/L and the size of the solids is $>100\ \mu$m. The oxygen gradient does not permit oxygen to reach the core of the solids. Therefore, bacteria on the perimeter of the floc particle degrade soluble cBOD using free molecular oxygen, while bacteria in the core of the solids use nitrate ions to degrade soluble cBOD. In the presence of an oxygen gradient, measurable oxygen is detected yet denitrification occurs.

The presence of soluble cBOD is the most important factor affecting the occurrence of an anoxic condition. The more soluble cBOD available to the facultative anaerobic bacteria, the more quickly the bacteria exhaust the available oxygen and begin to use nitrate ions. There are several sources of soluble cBOD that can be used by denitrifying bacteria. Chemical compounds that often are used in denitrification tanks include acetate, ethanol, glucose, and methanol. Methanol is the most commonly used source of soluble cBOD for denitrification tanks. Methanol is simplistic in structure, highly soluble, quickly absorbed by bacterial cells, and easily degraded. However, domestic wastewater often is used in an anoxic tank to achieve denitrification and destroy filamentous organisms and strengthen floc particles.

Denitrification proceeds in a step-by-step series of events through the reduction of nitrate ions. Nitrite ions are reduced to nitrite ions (NO_2^-), nitric oxide (NO), nitrous oxide (N_2O), and finally molecular nitrogen (N_2). When denitrification occurs there are three gases that are produced and released by the denitrifying bacteria. These gases in order of quantity produced are molecular nitrogen, carbon dioxide (CO_2), and nitrous oxide. Many bubbles of these gases escape from the wastewater, while some bubbles become entrapped in solids. The entrapped bubbles render the solids buoyant and result in a loss of settleability and solids from the secondary clarifier.

Denitrification also results in the production of alkalinity (Eq. 1).

$$\begin{aligned} 6\ NO_3^- + 5\ CH_3OH \longrightarrow\ & 3\ N_2 + 5\ CO_2 \\ & + 7\ H_2\ O + 6\ OH^- \end{aligned} \tag{1}$$

Alkalinity is returned in the form of the hydroxyl ion (OH^-) and the production of bicarbonate alkalinity (HCO_3^-) from carbon dioxide. Approximately 50% of the alkalinity lost during nitrification is returned to the activated sludge process through denitrification.

Clumping in the secondary clarifier can be prevented or controlled through several operational measures. Appropriate measures should be selected depending on whether a nitrification requirement exists for the activated sludge process.

If an activated sludge process is not required to nitrify, then nitrification may be terminated. This can be done by reducing solids inventory in the aeration tank, reducing the mean cell residence time (MCRT), reducing the dissolved oxygen level in the aeration tank, and taking an aeration tank off-line. If an aeration tank is taken off-line, then the remaining on-line aeration tanks have reduced hydraulic retention time (HRT) and increased organic loading. Reduced hydraulic retention time and increased organic loading disfavor nitrification.

If an activated sludge process is required to nitrify, then nitrification must be maintained. Clumping in the secondary clarifier can be prevented or controlled by increasing the return activated sludge (RAS) rate or treating secondary clarifier solids with a polymer or coagulant and increasing the RAS rate. Clumping also can be prevented or controlled by using plug flow mode of operation that incorporates an anoxic tank in the first tank (Fig. 1), or periodically terminating aeration for 1–2 h in an aeration tank while maintaining mixing action. Terminating aeration provides for the use of nitrate ions to degrade soluble cBOD in the aeration tank before they can be discharged to the secondary clarifier. However, once aeration is terminated to a tank, the tank can no longer nitrify.

Indicators of denitrification in the secondary clarifier are numerous. Rising bubbles of molecular nitrogen, carbon dioxide, and nitrous oxide are present in the clarifier. The bubbles may be found rising freely in

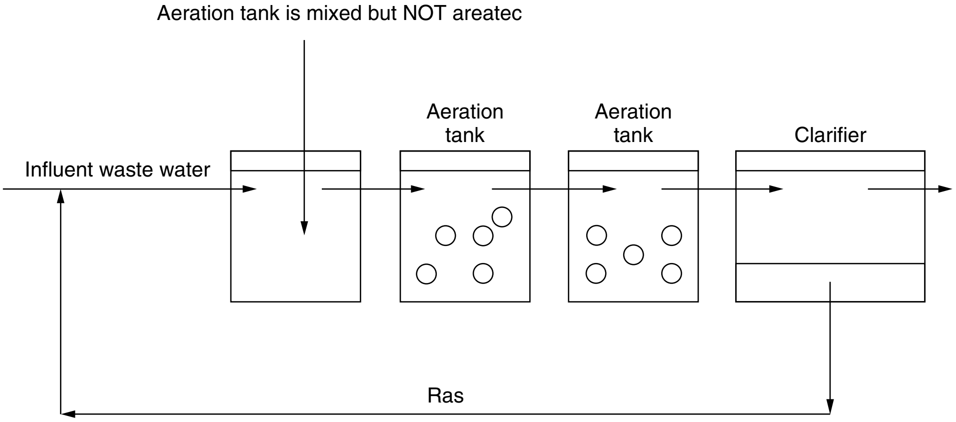

Figure 1. Plug flow mode of operation using an anoxic tank.

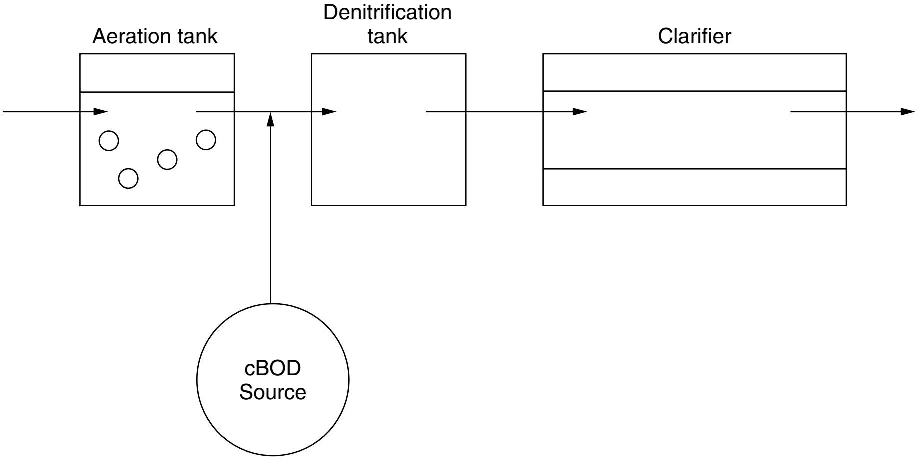

Figure 2. Denitrification tank.

the bulk solution, on the surface of rising solids, or released from the solids when the solids burst upon reaching the surface of the clarifier. The solids rising to the surface usually are dark due to the high MCRT required to grow a large population of nitrifying bacteria. Because nitrate ions are used during denitrification, the influent concentration of nitrate ions to the clarifier is greater than the effluent concentration of nitrate ions. Because denitrification returns alkalinity to the wastewater, the influent concentration of alkalinity to the clarifier is less than the effluent concentration of alkalinity from the clarifier. If sufficient alkalinity is returned to the secondary clarifier, an increase in pH across the clarifier occurs.

A denitrification tank is used to reduce the quantity of nitrogen in the effluent of an activated sludge process as required by permit (Fig. 2). A denitrification tank is located downstream of the activated sludge process. The denitrification tank receives the nitrate ions produced through nitrification in the aeration tank as well as the aeration tank's facultative anaerobic bacteria, residual dissolved oxygen, and residual soluble cBOD. In order to denitrify in the tank, slow subsurface mixing action is used to suspended large numbers of facultative anaerobic bacteria and place them in contact with the residual dissolved oxygen. Because the quantity of soluble cBOD entering the tank from the aeration tank is relatively small, soluble cBOD must be added to the denitrification tank. Approximately 3 mg/L of soluble cBOD are consumed in the denitrification tank for each mg/L of nitrate ion present. The soluble cBOD compound most often used in the denitrification tank is methanol. Retention times in the denitrification tank vary greatly, but most retention times are usually 30–60 min in length.

DETERGENTS

Kauser Jahan
Rowan University
Glassboro, New Jersey

Detergents are formulations that have cleaning and solubilization properties. Their main fields of application are industry (cleaning products, food, industrial processing), household (laundry, dishwashing), and personal care (soaps, shampoos, cosmetics). Detergents consist of surface-active agents known as surfactants and other subsidiary components that include boosters, builders, and fillers. In 1996, the production of surfactants including soaps was reportedly around 30×10^9 kg per year worldwide (1). Detergents are thus environmentally important because they are used in huge quantities. Historically, potential detergent contamination of the environment followed when the use of soap-based detergents changed to synthetic detergents. Detergents are discharged to the environment typically through wastewater treatment plant effluent and the use of sewage sludge on land. Industrial discharges to surface waters, oil spills, and surfactant-enhanced remediation of contaminated soil can also be other sources. Biological treatment of wastewater typically removes 95% of the detergents and their metabolites (2). However, recent research has generated scientific and regulatory concern because certain detergents and their metabolites appear toxic to aquatic organisms (3–7). These studies also indicate that detergent metabolites are also more toxic than their parent compounds. The complexity of detergent products, it has also been shown, leads to synergistic effects (8).

DETERGENT STRUCTURE AND USE

Detergents consist of surface-active agents, known as surfactants, and other subsidiary components that include boosters, builders, and fillers (9). Surfactants, as defined by Rosen (9), are substances that adsorb onto surfaces or interfaces and alter the surface or interfacial free energies of those surfaces when present at low concentrations in a system. The term interface means a boundary between any two immiscible phases: the term surface implies an interface where one phase is a gas. Surfactants have a characteristic molecular structure consisting of a structural group that is solvent hating (*lyophobic*) and a group that is solvent loving (*lyophilic*). This type of structure is called *amphipathic*. A typical surfactant molecule is shown in Fig. 1.

This nature of the surfactant causes concentration of the surfactant at a surface and the reduction of surface tension and also the orientation of the molecule at the surface: its *lyophilic* group is in the solvent phase, and its *lyophobic* group is oriented away from it. In an aqueous solution, the *lyophobic* term is *hydrophobic*, and the lyophilic term is *hydrophilic*. The hydrophobic group is typically a long-chain hydrocarbon, whereas the hydrophilic group is an ionic or highly polar group. Depending on the nature of the charge of the hydrophilic group, surfactants are classified as anionic (negative charge), cationic (positive charge), nonionic (no charge), and zwitterionic (both charges).

Anionic surfactants are used extensively in toilet soap bars, toothpaste, hair preparations and shampoos, wax strippers, textile lubricants, and degumming silk. Various types of anionic surfactants are listed in Table 1.

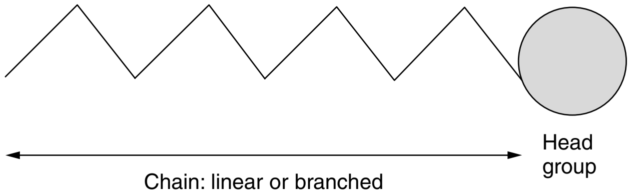

Figure 1. Structure of a surfactant molecule.

Table 1. Types of Anionic Surfactants[a]

Type	Example
1. Carboxylic acid salts	Soap
2. Sulfonic acid salts	LAS—linear alkyl benzenesulfonate in industrial detergents; petroleum sulfonates for oil recovery
3. Sulfuric acid ester salts	Na lauryl sulfate—toothpaste, fabric detergent
4. Phosphoric and polyphosphoric acid esters	Emulsifying agents in pesticides and herbicides
5. Perflourinated anionics	Control of oil and gasoline fires

[a] Reference 9.

Cationic surfactants are compatible with nonionic and zwitterionic surfactants. They carry a positive charge, so they strongly adsorb onto most solid surfaces. These surfactants exhibit poor detergency and are more expensive than other surfactants. Typical types of cationic surfactants and their uses are shown in Table 2.

Nonionic surfactants are widely used because they are compatible with other types of surfactants. They are generally available as 100% active material, free of electrolyte. They have poor foaming properties and are excellent carbon dispersing agents. Typical types of nonionic surfactants and their use are shown in Table 3.

Zwitterionic surfactants are compatible with all other kinds of surfactants. They are used as bactericides, corrosion inhibitors, pigment dispersion aids, cosmetics, fabric softeners, and soap detergent formulations.

Of all the types of surfactants mentioned, LAS (anionic) and APE (nonionic) are the major forms of surfactants found in detergent wastewater. APEs are further classified as OPEs (octylphenolethoxylates) and

Table 2. Types of Cationic Surfactants[a]

Type	Example
1. Long chain amines and their salts	Corrosion inhibitors; anticaking agents for fertilizers; flotation agents for metal ores
2. Di- and polyamines and their salts	Pigment coating; ore flotation
3. Quaternary ammonium salts	Textile softeners; germicides; disinfectants; sanitizers; health products
4. Polyoxyethylenated long-chain amines	Emulsifying agent for herbicides, insecticides, polishes, and wax emulsions
5. Quaternized polyoxyethlenated long-chain amines	Textile antistatic agent; dyeing leveler; corrosion inhibitor;
6. Amine oxides	Foam stabilizer for detergents, shampoos, and dishwashing liquids

[a] Reference 9.

Table 3. Types of Nonionic Surfactants[a]

Type	Example
1. Polyoxyethylenated alkylphenols; alkylphenol ethoxylates (APE)	Foam control; paint emulsifier; cosmetic emulsions; liquid detergents; dyeing retarders
2. Alcohol ethoxylates (AE)	Heavy-duty liquid detergents
3. Polyoxyethylenated polyoxypropylene glycols	Dispersants for pigments in latex paints; scale removal in boilers; petroleum deemulsifiers; foam control in laundry detergents
4. Polyoxyethylenated mercaptans	Cleaning and scouring of wool; metal cleaning; shampoos
5. Long-chain carboxylic acid esters	Food and pharmaceutical industries; cosmetics; textile antistats;
6. Alkanolamides	Foam stabilizers; thickeners for shampoos and liquid detergents
7. Tertiary acetylenic glycols	Wetting agents for use in powdered solids and emulsion paints
8. Polyoxyethylenated silicones	Wetting agents for polyester and polyethylene

[a] Reference 9.

NPEs (nonylphenolethoxylates), depending on the number of carbons. The general chemical structures of LAS and APE are presented in Fig. 2.

Analytical Methods

The qualitative and quantitative analysis of surfactants using titrimetric, spectrometric, and chromatographic techniques has been reported (3,4,10). Dissolved surfactant can be separated from dilute aqueous solution using the Wickbold method (11). This method is independent of the type of surfactant and isolates the surfactant as a residue suitable for analysis. The method, however, is lengthy and requires large sample volumes. Therefore, some studies have investigated and reported the use of SPE (solid-phase extraction) columns (12–13). The MBAS method (10) is useful for estimating the anionic surfactant content in water and wastewater samples. The CTAS (10) method is used to estimate the nonionic surfactant content in water if interference from other surfactants such as cationic and anionic can be eliminated. For more accurate and individual oligomers, normal-phase, high-performance liquid chromatography (HPLC) techniques have been reported for both LAS and APE surfactants. Ahel et al. (3,4,7) outlined analytical methods for APEs and their metabolites using normal-phase and reverse-phase HPLC for water, wastewater, and sludge samples. Ethoxylated nonionic surfactants are not volatile; therefore, their analysis is not well suited to gas chromatography. Typically, only lower molecular mass surfactant components can be analyzed by gas chromatography but only at high column temperatures (14). LAS analysis by HPLC has been described in detail by Matthijs and De Henau (15). Major metabolites for APE surfactants such as nonylphenol (NP) and octylphenol (OP) can also be analyzed by HPLC techniques (16–18).

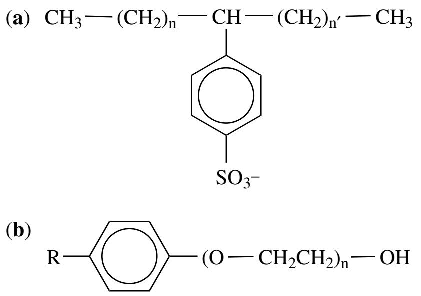

where R: C_9H_{19} n-Nonyl or C_8H_{17} O Octyl and n ~ 1 to 50

Figure 2. General chemical structure of (**a**) LAS and (**b**) APE.

OCCURRENCE OF DETERGENTS IN WASTEWATER AND SEWAGE SLUDGE

PT benzene (propylene tetramer benzene sulphonate) was used as the major surfactant in detergents until 1960. It was found that PT benzene is recalcitrant and causes foaming in rivers. This led to the use of more biodegradable straight-chain alkyl surfactants such as LAS. LASs currently represent 40% of all surfactants used (19). The other commonly used class of surfactants is APEs; the worldwide production is reportedly around 500×10^3 tons pa (19). Therefore, literature has focused more on the environmental impacts of LASs and APEs. Measured concentration of surfactants in municipal sewage effluents reportedly range from 0.008 to 6.2 mg/L and in river waters from 0.0042 to 2.6 mg/L (20–21). Literature indicates that a significant portion of the surfactants in raw sewage adsorbs to particulate matter. Primary sludge removed from primary settling tanks are reportedly rich in LAS, and concentrations vary from 5,000–15,000 mg/L (19). Most surfactants also do not biodegrade under anaerobic conditions. Therefore, sludge treated via anaerobic processes is rich in surfactants. Anaerobically digested sewage sludge reportedly contains 0.3–1.3% LAS (19).

Dentel (22) reported the load of LAS and APE in a typical WWTP and their subsequent fate. His data are presented in Fig. 3. The figure indicates that significant

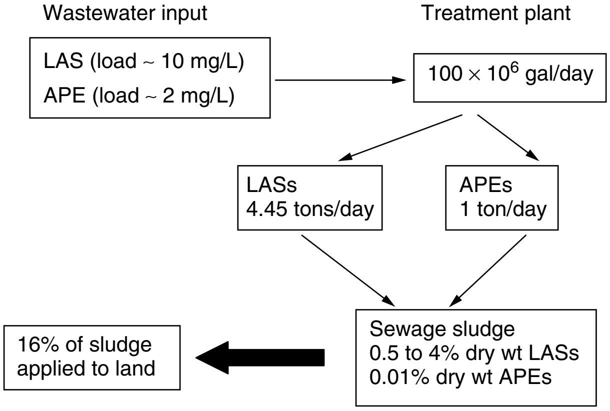

Figure 3. Surfactant transport through a typical WWTP (22).

amounts of surfactants are transported to the environment through WWTP effluent discharge and sludge disposal.

BIODEGRADATION OF DETERGENTS

Biodegradation of surfactants can be classified as primary or ultimate. Primary biodegradation is the loss of detergency, whereas ultimate indicates complete conversion to CO_2, CH_4, water, salts, and biomass. The biodegradation of LASs and APEs is well documented. LASs are highly biodegradable—rates are as high as 97–99% under aerobic conditions. Typically, the process involves breakdown of the straight-chain alkyl chain, then the sulphonate, and finally the benzene ring. There is no evidence for comparable LAS degradation under anaerobic conditions. Biodegradation of LAS, it has been shown, is dependent on dissolved oxygen concentrations, the presence of other organic contaminants and other surfactants, and pH.

In contrast to LASs, APEs are less biodegradable: values of 0–20% have been reported (19). APEs undergo complete primary degradation in aerobic environments, and metabolic products accumulate. APEs consist of NPES and OPEs. Nonylphenol (NP) and octylphenol (OP) have been identified as the major metabolites of these two classes of surfactants. Both are toxic and xenoestrogenic; octyplenol is more toxic than nonylphenol (23). NP is approximately 10 times more toxic than its ethoxylate precursor. Biodegradation is accomplished by stepwise shortening of the ethoxylate chain. This produces a complex mixture of compounds that can be divided into three main groups: short-chain ethoxylates, alkylphenoxy carboxylic acids, and alkylphenols such as NP and OP. As the chain gets shorter, the molecule becomes less soluble. The alkylphenoxy carboxylic acids and longer chain APEs are soluble in water; the shorter chain APEs are insoluble in water, particularly NP and OP. Both compounds have low water solubility and tend to adsorb onto suspended solids or sediments.

The biodegradation of surfactants in sewage sludge poses some problems. Typically a common initial step in WWTP is primary settling of particulates. Sludge from this tank can be rich in surfactants because surfactants tend to adsorb to particulates. Anaerobic sludge treatment at high temperature is typically the norm for WWTPs. Most common surfactants are not biodegradable under anaerobic conditions. Therefore, anaerobically digested sewage sludge is a potential source of surfactants for the soil.

LAS are not biodegradable either by mesophilic or thermophilic anaerobic digestion. Therefore, anaerobically digested sludge can contain 0.3–1.2% LAS and thus be a potential source. Jensen (24) reported that LAS concentrations in sludge are highly dependent on the sludge treatment method employed. LAS in sludge-amended soil, however, was rapidly biodegraded in the aerobic soil. Concentrations of APE are much higher in anaerobically digested sludge (900–1,100 mg/kg) than in aerobically digested sludge (0.3 mg/kg) (19). Marcomini (25) reported that sewage-amended soil in aerobic environments exhibited rapid NP degradation.

Cationic surfactants are considered readily biodegradable under aerobic conditions. These surfactants carry a positive charge, so they have a strong affinity for negatively charged particulate surfaces in sewage sludge. The adsorption of cationic surfactants to particulate matter makes the sludge treatment method employed important. However, there is no evidence of anaerobic biodegradation for this class of surfactants.

Data for soap indicates that it is readily biodegradable in both aerobic and anaerobic environments (26). Fatty acid esters (FES) are also readily biodegradable under aerobic conditions (27) but poorly degradable under anaerobic conditions (28). Steber and Werich (29) found rapid degradation of FES in aerobic soils amended with anaerobically treated sludge.

Fatty alcohol sulphates (FES) are readily biodegradable under both aerobic and anaerobic conditions (30). Research also indicates that alcohol ether sulfates (AES) are readily biodegradable under aerobic and anaerobic conditions. Fatty alcohol ethoxylates (AE) are known to be readily biodegradable and thus maybe a suitable replacement for APEs.

TOXICITY OF DETERGENTS

Toxicity of detergents has received a lot of attention in recent years. Detergents and their metabolites are common constituents of municipal effluents, so they also contribute to the toxicity. Toxicity data for surfactants to various marine species of freshwater and marine life have been documented as early as 1950 (31). In most cases, these data were generated from standard toxicity tests in the laboratory under controlled conditions. Detergent products and their ingredients are relatively toxic to aquatic life. Lewis (35) has provided an extensive review on the toxicities of surfactants to freshwater and marine life. Household washing detergents and softeners, it has been shown, are toxic from 0.07 to 35.4 mg/L (32). Both LAS and NPE exhibit relatively high acute toxicities; 48-h LC_{50} values are typically in the range of 5 to 10 and 0.1 to 2.0 mg/L. There is a large and growing body of literature on the acute and chronic toxicity of NP and NPEs toward aquatic organisms (33–36). Exposure of fathead minnows to 22.7 mg/L of NP in water for 20 days resulted in a bioconcentration factor of 344 (34). Selected values for NP toxicity determined by Naylor (36) are presented in Table 4.

REGULATORY STANDARDS

There are no regulatory restrictions on the use of APEs in the United States. Massive monitoring studies conducted in the United States (36) indicate that NP does not pose a significant risk to aquatic organisms. The USEPA is clarifying the regulatory status of NPEs by issuing water quality guidance (37–38). The draft guidance indicates the following levels for nonylphenol:

Freshwater 6.6 ppb water (4-day average) and 25 ppb (1-hour average)

Saltwater 1.6 ppb (4-day average) and 6.2 ppb (1-hour average)

The APE Research Council supports the draft EPA water quality guidelines for NP.

Table 4. Nonyphenol Aquatic Toxicity[a]

Species	Type	LC_{50} (96 h), ppb	No Observable Effect, ppb	Lowest Observable Effect, ppb	Maximum Acceptable Toxicant, ppb
Mysidopsis bahia	Acute	43			
Marine shrimp	Chronic	Length at 28 days	3.9	6.7	5.1
		Reproduction	6.7	9.1	7.8
		Survival	6.7	9.1	7.8
Pimephales promelas	Acute	300			
		135			
Fathead minnow, freshwater	Chronic	Hatch rate	23	>23	>23
		length, 28 days	23	>23	>23
		Survival, 33 days	7.4	14	10.2
Cyprinodon variegates	Acute	310			
Sheepshead minnow, marine					
Salmo gairdneri	Acute	230			
Trout freshwater		480 (24 h)			
Daphnia magna	Acute	190 (48 h)	77	160	111
Flea freshwater		440 (48 h)			
	Chronic	Reproduction rate	24	39	31
Selanastrum capricornutum	Acute	EC_{50} 410	92	190	132
Freshwater green alga					
Skelatonema costatum	Acute	EC_{50} 27	10	20	14.1
Marine algae					
Rana Catesbiana	Acute 10 day	260 mg/kg			
Tadpole, freshwater	NP in sediment				
Chironomus tentans	Subacute, 14 day NP in sediment				26.1 mg/kg
Midge larva, freshwater					
Mytilus edulis	Acute	2600			
Mussel marine	Chronic	Growth and strength, 32 days		56	

[a] Reference 36.

The Soaps and Detergents Association found that the present use of NPES in detergents and cleaning products poses little or no risk to the environment in the United States. Environmental monitoring studies (39–41) demonstrate that APEs and their metabolite concentrations in rivers and lakes in the United States are below toxic levels. These studies also demonstrate that 92 to 99% of NPEs are effectively removed in wastewater treatment plants in the United States. Based on research conducted in the USA, APEs do not appear to pose a significant ecological risk.

The Canadian Government is currently reviewing the environmental and health characteristics of NP and NPEs under its PSL2 program. European countries, however, have started phasing out the use of APEs in their products. In 1986, Germany instituted voluntary restrictions and Switzerland banned the use of surfactants in laundry detergents (39). A voluntary ban on APE use in household products began in 1995 throughout northern Europe (42). Denmark has introduced an environmental quality standard for NP and NPEs of 1 μg/L. Studies conducted in Japanese wastewater treatment plants (43) indicated widespread pollution by NPEs and their metabolites.

CONCLUSIONS

Literature indicates that the presence of detergents in wastewaters may pose problems, depending on the treatment methodology for the wastewater and sludge. Most detergents are biodegradable under aerobic conditions. However, concerns with degradation products of APEs, particularly nonylphenol, remain. There is also concern over the resistance of all major classes of detergents to biodegradation in anaerobic environments, especially because this is the predominant method of sludge treatment from primary tanks. The amphiphilic nature of surfactants makes them prone to adsorption to the sludge during primary tank settling. Therefore a significant portion of the surfactants in wastewaters may be passed untreated into the sludge. Studies indicate that application of sludge in aerobic soil environments can lead to further biodegradation. However, more studies are required to determine the fate of toxic intermediates in soil.

BIBLIOGRAPHY

1. Schroder, H.Fr. (2001). Tracing of surfactants in the biological wastewater treatment process and the identification of their metabolites. *J. Chromatogr.* **926**: 127–150.
2. Swisher, R.D. (1987). *Surfactant Biodegradation*. Marcel Dekker, New York.
3. Ahel, M., Giger, W., and Koch, M. (1994). Behaviour of alkylphenol polyethoxylate surfactants in the aquatic environment-I. Occurrence and transformation in sewage treatment. *Water Res.* **28**: 1131–1142.
4. Ahel, M., Giger, W., and Koch, M. (1994). Behaviour of alkylphenol polyethoxylate surfactants in the aquatic

environment-II. Occurrence and transformation in rivers. *Water Res.* **28**: 1143–1152.

5. Banat, F., Prechtl, S., and Biscof, F. (2000). Aerobic thermophilic treatment of sewage sludge contaminated with 4-nonylphenol. *Chemosphere* **41**: 297–302.
6. Bennett, E.R. and Metcalfe, C.D. (1999). Distribution of degradation products of alkylphenol ethoxylates near sewage treatment plants in the lower great lakes, North America. *Environ. Toxicol. Chem.* **19**(4): 784–792.
7. Brunner, P.H., Capri, S., Marcomini, A., and Giger, W. (1988). Occurrence and behaviour of linear alkylbenzenesulphonates, nonylphenol, nonylphenol mono- and nonylphenol diethoxylates in sewage and sewage sludge treatment. *Water Res.* **22**: 1465–1472.
8. Lewis, M.A. (1992). The effects of mixtures and other environmental modifying factors on the toxicities of surfactants to freshwater and marine life. *Water Res.* **28**(8): 1013–1023.
9. Rosen, M.J. (1989). *Surfactants and interfacial phenomena*, 2nd Edn. John Wiley & Sons, New York.
10. AWWA, APHA, and WEF. (1995). *Standard Methods for the Evaluation of Waste Water*. Washington, DC.
11. Wickbold, R. (1972). On the determination of nonionic surfactants in river and wastewater. *Ten. Surf. Deterg.* **9**: 173–177.
12. Riu, J. et al. (2000). Determination of LAS in WWTPs and coastal waters by automated solid phase extraction followed by capillary electrophoresis. *J. Chromatogr.* **889**: 221–229.
13. Stephanou, E. and Giger, W. (1982). Persistent organic chemicals in sewage effluents: quantitative determinations of NP and NPEs by glass capillary gas chromatography. *Environ. Sci. Technol.* **16**: 800–805.
14. McEvoy, J. and Giger, W. (1986). Determination of LAS in sewage sludge by high resolution gas chromatography/mass spectrometry. *Environ. Sci. Technol.* **20**: 376–383.
15. Matthijs, E. and De Hanau, H. (1987). Determination of LAS in aqueous samples, sediments, sludges and soils using HPLC. *Tens. Surf. Deterg.* **24**: 193–199.
16. Giger, W., Brunner, P.H., and Schaffner, C. (1984). 4-nonylphenol in sewage sludge: accumulation of toxic metabolites from nonionic surfactants. *Science* **225**: 623–625.
17. Field, J.A. and Reed, R.L. (1996). NPEC metabolites of nonionic surfactants in US paper mill effluents. *Environ. Sci. Technol.* **30**: 3544–3550.
18. Lee, H.-B. and Peart, T.E. (1995). Determination of 4-nonylphenol in effluent and sludge from sewage treatment plants. *Anal. Chem.* **67**: 1976–1980.
19. Scott, M.J. and Jones, M.N. (2000). The biodegradation of surfactants in the environment. *Biochim. Biophys Acta* **1508**: 235–251.
20. Dorn, P., Salanitro, J., Evans, H., and Kravetz, L. (1992) Assessing the aquatic hazard of some branched and linear non-ionic surfactants by biodegradation and toxicity. *Environ. Toxicol. Chem.* **12**: 1751–1762.
21. Hennes-Morgan, E.C. and Oude, N.T. (1993). Detergents In: *Handbook of Ecotoxicology*. P. Calow (Ed.). Blackwell Scientific, Oxford, UK, pp. 130–154.
22. Dentel, S.K., Allen, H.E., Srinivasarao, C., and Divincenzo, J. (2002). *Third Year Completion Report Project No. 6, Water Resources Center*. University of Delaware, Newark, DE. Available: http://bluehen.ags.udel.edu/dewre/surfact.htm.
23. Johnson, A.C., White, C., Bhardwaj, L., and Jurgens, M.D. (2000). Potential for octylphenol to biodegrade in some english rivers. *Environ. Toxicol. Chem.* **19**(10): 2486–2492.
24. Jensen, J. (1999). Fate and effects of LAS in the terrestrial environment. *Sci. Total Environ.* **226**: 93–111.
25. Marcomini, A., Capel, P.D., Goger, W., and Hani, H. (1988). Residues of detergent derived organic pollutants and PCBs in sludge amended soil. *Naturwissenschaften* **75**: 460–462.
26. Birch, R.R., Gledhill, W.E., Larson, R.J., and Nielsen, A.M. (1992). Role of Anaerobic Biodegradability in the Environmental Acceptability of Detergent Materials, *Proc. 3rd CESIO*. London, Int. Surf. Congr. Exhib., 26, pp. 26–33.
27. Gode, P., Guhl, W., and Steber, J. (1987). Environmental compatibility of fatty acid, alpha-sulfomethyl esters. *Fat. Sci. Technol.* **89**: 548–552.
28. Maurer, E.W., Weil, J.K., and Linfield, W.M. (1965). The biodegradation of esters of α-sulfo fatty acids. *J. Am. Oil Chem. Soc.* **54**: 582–584.
29. Steber, J. and Weirich, P. (1987). The anaerobic degradation of detergent range fatty alcohol ethoxylates-studies with ^{14}C labeled surfactants. *Water Res.* **21**: 661–667.
30. Thomas, O.R.T. and White, G.F. (1989). Metabolic pathway for the biodegradation of sodium dodecyl sulphate by *pseudomonas* sp-c12b. *Biotechnol. Appl. Biochem.* **11**: 318–327.
31. Holt, M.S., Mitchell, G.C., and Watkinsson, R.J. (1992). The environmental chemistry, fate and effects of nonionic surfactants. In: *Anthropogenic Compounds*. Vol. 3, Part F, Handbook of Experimental Chemistry, O. Hutzinger (Ed.). Springer, Berlin, pp. 89–144.
32. Ankley, G.T. and Burkhard, P. (1992). Identification of surfactants as toxicants in a primary effluent. *Environ. Toxicol. Chem.* **11**: 1235–1248.
33. Argese, E. et al. (1994). Submitochondrial particle response to LAS, nonylphenol polyethoxylates and their biodegradation derivative. *Environ. Toxicol. Chem.* **13**(5): 737–742.
34. Staples, C.A. (1998). An evaluation of the aquatic toxicity and bioaccumulation of C8- abd C9- alkylphenol ethoxylates. *Environ. Toxicol. Chem.* **17**: 2470–2480.
35. Lewis, M.A. (1992). The effects of mixtures and other environmental modifying factors on the toxicities of surfactants to freshwater and marine life. *Water Res.* **26**(8): 1013–1023.
36. Naylor, C.G. (1998). *Environmental Fate and Safety of Nonylphenol Ethoxylates*. The Alkylphenols and Alkylphenol Ethoxylates Review, APE Research Council, Washington, DC.
37. Renner, R. (1997). European bans on surfactants trigger transatlantic debate. *Environ. Sci. Technol.* **31**: 316A–320A.
38. VandenHenvel, C. (2000). A more certain future for nonylphenol ethoxylates. *Chem. Times Trends* **Spring**: 38–39.
39. Hale, R.C. et al. (2000). Nonylphenols in sediments and effluents associated with diverse wastewater outfalls. *Environ. Toxicol. Chem.* **19**(4): 946–952.
40. Bennett, E.R. and Metcalfe, C.D. (2000). Distribution of degradation products of APEs near sewage treatment plants in the lower great lakes, north america. *Environ. Toxicol. Chem.* **19**(4): 784–792.
41. Naylor, C.G., Williams, J.B., Varineau, P.T., and Webb, D.A. (1998). Nonylphenol ethoxylates in an industrial river. *Alkylphenols Alkylphenol Ethoxylates Rev.*: 44–53.
42. Warhurst, A.A. (1985). *Environmental and Human Safety of Major Surfactants*. Friends of the Earth Scotland, Edinburgh, Scotland.
43. Fujita, M. et al. (2000). Behavior of nonylphenol ethoxylates in sewage treatment plants in Japan—biotransformation and ecotoxicity. *Water Sci. Technol.* **42**: 23–30.

ECOLOGICAL WASTEWATER MANAGEMENT

GUENTER LANGERGRABER
BOKU—University of Natural
Resources and Applied Life
Sciences
Vienna, Austria

Conventional sanitation concepts are neither an ecological nor an economical solution for rural areas in both industrialized and developing countries. Ecological sanitation (EcoSan) represents a holistic approach toward ecologically and economically sound sanitation. The underlying aim is to close nutrient and water cycles with as small an expenditure on material and energy as possible to contribute to sustainable development. EcoSan is a systemic approach and an attitude. Single technologies are only a means to an end. Therefore, EcoSan-technologies may range from nearly natural wastewater treatment techniques to compost toilets, simple household installations, and to complex, mainly decentralized systems. These technologies are not ecological per se but only in relation to the observed environment. Promotion of EcoSan concepts, therefore, is the strategy for achieving the goal—closing the loop in wastewater management and sanitation.

DEFINITION OF ECOLOGICAL SANITATION

A sanitation system that provides ecological sanitation (EcoSan) is a cycle—a sustainable, closed-loop system (Fig. 1). The EcoSan approach is resource minded, not waste minded. Human excreta are treated as a resource and are usually processed on-site and then, if necessary, processed further off-site until they are completely free of disease organisms. The nutrients contained in excreta are then recycled by using them, for example, in agriculture.

The basic motivation behind the need to reshape the management of nutrients and streams of organic residuals in society may be found in the so-called "basic system conditions for sustainable development" for water and sanitation management, formulated in Agenda 21 (1):

1. The withdrawal of finite natural resources should be minimized.
2. The release of nonbiodegradable substances to the environment must be stopped.
3. Physical conditions for circular flows of matter should be maintained.
4. The withdrawal of renewable resources should not exceed the pace of regenerating them.

If a sanitation system shall contribute toward the goals of ecological sanitation, it has to meet or at least to be on the way toward meeting the following criteria, as given by Esrey et al. (2):

1. *Prevent disease*: A sanitation system must be capable of destroying or isolating pathogens.
2. *Affordable*: A sanitation system must be accessible to the world's poorest people.
3. *Protect the environment*: A sanitation system must prevent pollution, return nutrients to the soil, and conserve valuable water resources.
4. *Acceptable*: A sanitation system must be aesthetically inoffensive and consistent with cultural and social values.
5. *Simple*: A sanitation system must be robust enough to be easily maintained within the limitations of

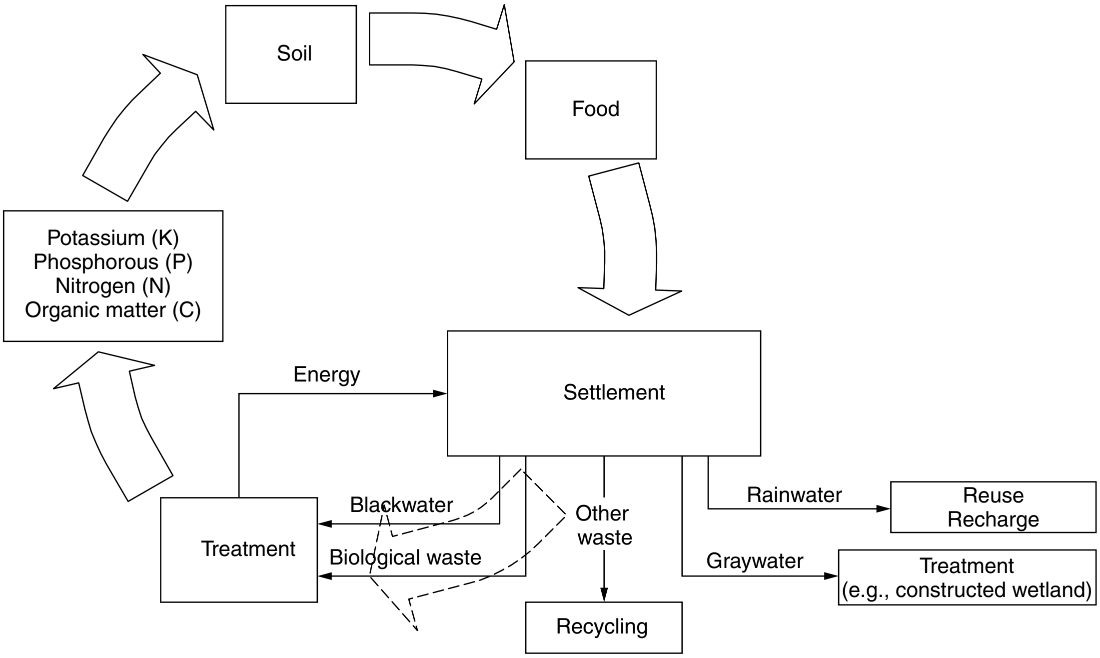

Figure 1. An ecological sanitation system.

the local technical capacity, institutional framework, and economic resources.

6. *Comfortable*: A sanitation system must be nearly as comfortable as a flush toilet. That means it has to be indoors or accessible under the same roof of the house.

Successful implementation of sanitation systems requires understanding all components of the system. The components have to be considered together when designing and making sanitation systems work. Following are the main components (Fig. 2):

- *Nature*: The most relevant natural variables are climate (humidity, temperature), water (amount available, groundwater level), and soil (stability, permeability).
- *Society*: The factors that describe the society include the settlement pattern (concentrated/dispersed, low/high rise), attitude (fecophobic/fecophilic), habits, beliefs and taboos related to human excreta, and the economic status of the community.
- *Process*: Physical, chemical, and biological processes turn human excreta into a nondangerous, inoffensive, useful product. Dehydration and decomposition are the principal processes.
- *Device*: The device is the on-site structure specifically built for defecation and urination.

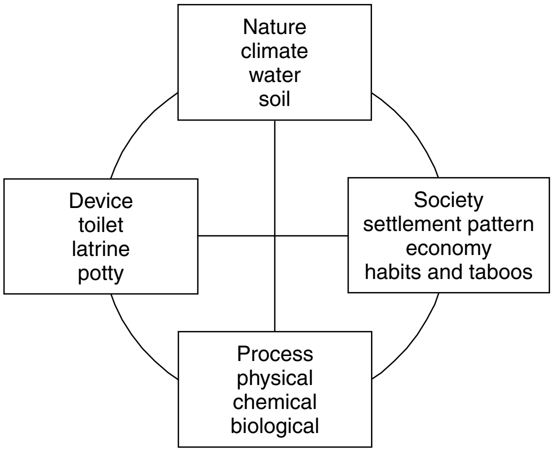

Figure 2. Main components of the “sanitation” system.

It is essential to sanitize human excreta before they are recovered and reused. Figure 3 shows the different approaches for handling human excreta:

- *Mix and drain*: In conventional sanitation systems, urine and feces are mixed and flushed away with water.
- *Mix and evaporate*: Excreta are mixed; however, they are not flushed away but treated on-site, for example, composted.
- *“Don’t mix” and dehydrate*: Urine and feces are collected and treated separately.

The principles underlying EcoSan are not novel. Sanitation systems based on ecological principles have been used in different cultures, for hundreds of years. EcoSan systems are still widely used in parts of East and Southeast Asia. In Western countries, this option was largely abandoned, as flush-and-discharge became the norm, but in recent years, there has been a revival of interest in EcoSan (2).

WASTEWATER IS A RESOURCE

Wastewater from households contains urine (yellowwater), feces (brownwater), and graywater. Graywater is the part of the wastewater which is not mixed with excreta (from kitchens, bathrooms, and laundries). If urine and feces are mixed, the resulting mixture is called blackwater.

For a long time, wastewater has been regarded as a problem because wastewater involves hygienic hazards and contains eutrophying substances in the form of organic matter, nitrogen, and phosphorus. These substances cause problems in seas, lakes, and streams, but, on the other hand, they would be valuable to farmers. Nitrogen (N), phosphorus (P), and potassium (K) in wastewater can be used instead of artificial fertilizer, and the organic material increases the humus content. Recirculating nutrients from wastewater as fertilizer reduces the need for industrially produced fertilizer and also reduces

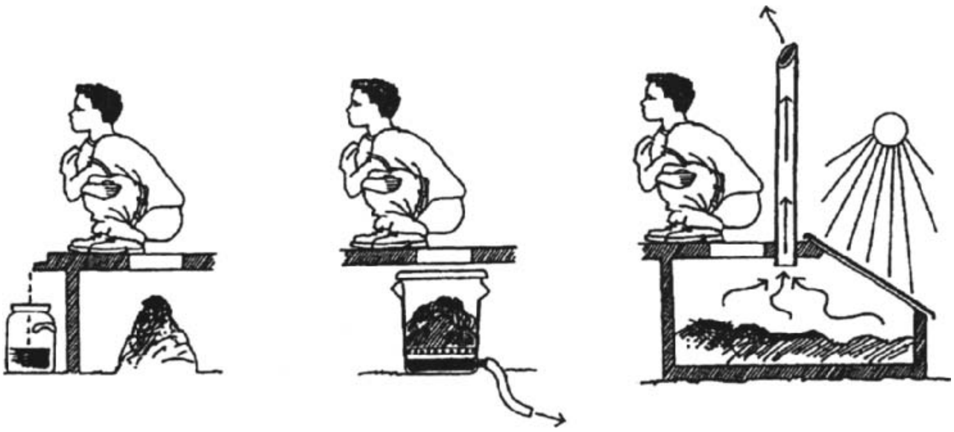

Figure 3. Different approaches for handling excreta: “Don’t mix” and dehydrate (left); mix and drain (middle); mix and evaporate (right).

discharges of nutrient-rich water from treatment plants into watercourses.

Wastewater Characteristics

One person produces about 500 liters of urine and 50 liters of feces per year. The same person, additionally, produces a range of 25,000 to more than 100,000 liters of graywater. Blackwater and graywater have very different characteristics (Table 1). Most of the nutrients essential in agriculture (N, P, K) occur in urine. Feces contain smaller amounts of nutrients, and the quantities in graywater are insignificant. If blackwater is collected separately at low dilution, it can be converted, for example, to a safe natural fertilizer that replaces synthetic products and prevents the spread of pathogens and other pollutants to receiving waters.

Sources of Hygienic Hazards

Watersheds may be affected directly by excreta that contain large amounts of pathogens from humans and animals, or indirectly through wastewater outlets, from large-scale wastewater treatment plants and from smaller units of wastewater or graywater on-site sanitation. Stormwater and runoff from agricultural lands may also carry large amounts of pathogens to watersheds emanating directly from excreta, or pathogens occurring in sludge, excreta, and manure applied to land. Stormwater and runoff water may also carry pathogens from domestic and wild animals and birds that may affect humans. Organic waste from human settlements and activities may also be a source of pathogenic organisms.

- *Human fecal excreta* may be harmless but can contain large amounts of pathogenic organisms. The risk depends on the frequency of infected persons and symptomless carriers in the population. Anyway, human fecal excreta are responsible for the major part of hygienic hazards.
- *Human urine* does not normally contain pathogenic organisms that will transmit enteric disease to other individuals. Fecal material is thus the main source of infectious organisms. Only in special cases, for example, a systemic infection with fever, will pathogenic organisms be present in urine.
- All microorganisms in *wastewater* originating from human excreta occur in amounts reflecting their occurrence in infected persons or carriers connected to the system. Their concentrations also depend on the dilution of the water. Untreated wastewater should always be regarded as a potential carrier of pathogenic organisms.
- *Graywater* normally contains small amounts of pathogenic organisms. However, due to a relatively high load of easily degradable organic substances, regrowth of indicator organisms of fecal pollution may occur.
- *Stormwater* may have high loads of fecal contamination. This is of special concern in areas of the world where open-air defecation is practiced, because high loads of pathogens, as in wastewater, may occur. Stormwater may also contain high loads of zoonotic pathogens originating from animal or bird feces.

Table 1. Typical Characteristics of the Main Components of Household Wastewater[a]

		Graywater	Urine	Feces
Volume (l.p^{-1}.yr^{-1})		25,000–100,000	∼500	∼5
Yearly loads (kg.p^{-1}.yr^{-1})				
N	∼4–5	∼3%	∼87%	∼10%
P	∼0.75	∼10% (P-free detergents)	∼50%	∼40%
K	∼1.8	∼34%	∼54%	∼12%
COD	∼30	∼41%	∼12%	∼47%

[a] Reference 3.

Resource Wastewater?

Summarizing the characteristics of wastewater and the sources of hygienic hazards, the following conclusions can be drawn:

- Most of the soluble nutrients are found in urine. If urine is segregated and converted to agricultural usage, the biggest step toward nutrient reuse and highly efficient water protection will have been taken.
- The hygienic hazards of wastewater originate mainly from fecal matter. Segregation opens the way to sanitation and finally to an excellent end product.
- Wastewater that is not mixed with feces and urine is a great resource for high quality reuse of water.
- Source control should include evaluating all products that end up in the water. High quality reuse will be far easier when household chemicals are not only degradable (original substance disappears, even if metabolites do not degrade) but can be mineralized by the available technology. Additionally, pipes for drinking water should not emit pollutants (e.g., copper or zinc).
- To reduce stormwater runoff, local infiltration and/or trenches to surface waters for relatively unpolluted rainwater can be used. Prevention of pollution includes avoiding copper or zinc gutters and roof materials that can cause heavy metal pollution of rainwater runoff.

TREATMENT SYSTEMS

Limitations of Conventional Sanitation

In conventional sanitation systems, human excreta are mixed with water and flushed away by conventional flush toilets. The wastewater is then collected and transported in sewers and treated in a centralized plant.

This results in high water demand, the spread of potentially dangerous pathogens and micropollutants (e.g., residues of pharmaceuticals) in a large volume of water and also the loss of the option to reuse graywater and to produce fertilizer. The initially small amount of feces could be hygienized easily by cheap methods.

For the strange mixture called "municipal wastewater," hygienization is an expensive further treatment step. Hormones and medical residues reveal another weakness of sanitation systems. These substances reach receiving waters easily especially because of their polarity and usually very low degradation rates in treatment plants.

Sewerage systems have a couple of severe disadvantages. In general, they are a very costly part of the infrastructure (if sufficient rehabilitation is done) and drain large amounts of water from the region. Combined systems emit raw wastewater into receiving waters with overflows. Storage tanks are very expensive if the number of overflows is low. Separate systems often have a large number of wrong connections (4).

Centralized Versus Decentralized Wastewater Treatment

Sanitation systems can be classified as follows:

- *Decentralized or 'on-plot' systems* in which safe disposal of excreta takes place on or near a single household or a small settlement.
- *Centralized or 'off-plot' systems* in which excreta are collected from individual houses and carried away from the plot to be treated off-site.

The selection of the most appropriate sanitation system is influenced by ecological, technical, social, cultural, financial, and institutional factors.

Proper decisions on where to connect houses to a sewerage system and where to build on-site facilities or small decentralized plants are the key issue for the economics of the whole wastewater infrastructure. Good regional planning can avoid the deadlock of very expensive sewerage systems that use all the money which could serve the environment in highly efficient decentralized treatment and collection systems.

There are cost calculation procedures that include long-term development to balance operating- and investment costs and products (reuse water, fertilizer, soil improver). The price of secondary products can be very relevant in economically weak and water-scarce countries where water and industrial fertilizers are no longer subsidized. Reuse-oriented sanitation can easily exceed the performance of the most advanced high-tech end-of-the-pipe plant often at much lower costs (4). Water and sanitation projects and programs will fail to be sustainable if they are not planned and designed to meet the needs of the end user (5).

Decentralized Wastewater Treatment Systems

The use of decentralized wastewater treatment systems offers these advantages (6). They save money, protect the homeowner's investment, promote better watershed management, offer an appropriate solution for low density communities, provide a suitable alternative for varying site conditions, and furnish effective solutions for ecologically sensitive areas.

Decentralized systems prevent unnecessary costs by focusing on preventive measures (assessment of community conditions/needs and maintenance of existing systems), instead of reacting to crises. They further maximize the potential for homeowners, who have existing septic systems, to continue to benefit from their original investment and avoid the potentially large transfers of water from one watershed to another that can occur in centralized systems. In small communities of low population density, decentralized systems are the most cost-effective option. These systems are variable, can be designed for various sites that have, for example, shallow water tables, shallow bedrock, low-permeability soils, and small property lot size. Additionally, decentralized systems can provide cost-effective solutions for areas that require advanced treatment, such as nutrient removal and/or disinfection, while recharging local aquifers and providing reuse opportunities close to points of wastewater generation.

Technology Options

Decentralized wastewater treatment alternatives for small communities can be broadly placed in the following categories that represent the basic approaches to wastewater conveyance, treatment, and/or disposal (6, revised):

- *Natural systems* use soil as a treatment and/or disposal medium, including land application, constructed wetlands, and subsurface infiltration. Some sludge and septage handling systems, such as sand drying beds, land spreading, and lagoons, are included.
- *Conventional treatment systems* use a combination of biological, physical, and chemical processes and employ tanks, pumps, blowers, rotating mechanisms, and/or mechanical components as part of the overall system. These include suspended growth, fixed growth, and combinations of the two. This category also includes some sludge and septage management alternatives, such as digestion, dewatering and composting systems, and appropriate disposal.
- *Alternative collecting systems* that use lightweight plastic pipe buried at shallow depths, have fewer pipe joints and less complex access structures than conventional gravity sewers. These include pressure, vacuum, and small-diameter gravity sewer systems.
- *Alternative treatment systems* use source control and separating systems. An example of an alternative treatment system for ecological sanitation is given later.

None of the described systems is a priori or not an EcoSan system. A number of criteria as given above have to be met for a sanitation system to be called an EcoSan system.

Low Water Consumption and Water-Free Toilets and Urinals. A major part of alternative treatment systems are devices suitable for reducing water consumption and/or for separating feces and urine. Therefore, the available systems of low water consumption and water-free toilets and urinals are described:

1. *Conventional water-flush toilet* (listed as the reference system): The conventional water-flush toilet is

standard in most industrialized countries. A precondition to the effective disposal of excrements in this manner is sufficient availability of flushing water (about 6–8 liters per flush) as well as a corresponding disposal system (sewer system and wastewater treatment plant or other disposal). At the high dilution of human excretions (approximately 100 to 250 g feces and 1 to 1.5 liters urine), the use of a conventional flush toilet requires about 40 liters of water per day, and the wastewater produced contains both nutrients and pathogens.

2. *Water-conserving toilet*: The water-conserving toilet reduces the required consumption of water to 1.0 to 4.5 liters per flush. Together with a reduction in water costs and an increase in the available capacity of the existing wastewater collection system (e.g., collection pits), water-conserving toilets produce a more concentrated wastewater that can be used further more easily (e.g., biogas extraction, solid waste separation). Due to the smaller volume of flushing water, blockages in the pipes must be prevented by a flushing device, and certain structural requirements must be observed (e.g., minimum gradient in the downstream pipe).
3. *Water-conserving toilet with waste segregation*: In addition to the reduction in water consumption, the water-conserving toilet with waste segregation allows the division of feces and urine. Therefore, urine can be used for, for example, agricultural purposes, and costs can be reduced by avoiding removal of nutirents in sewage treatment. The installation of a waste-segregation toilet is only possible in combination with a corresponding collection system (urine collection tank in the house, regular emptying of the tanks, agricultural application). Feces and urine are divided in a user-friendly manner by an adapted effluent vent in the toilet bowl.
4. *Vacuum toilets*: Vacuum toilet systems dispose of flushing water and excrement by using a pipeline network under vacuum (approximately 0.5 bar) connected to a collection tank. Water is required only to rinse the toilet bowl, not for facilitating transport (approximately 0.7 to 1.0 liter per flush). Until now, the vacuum system has been used mostly in ships, trains, and aircraft (limited flush and wastewater capacities). The wastewater is very concentrated, so it is suitable for energy generation in a biogas plant. Drawbacks to the system are the technical requirements (operation and maintenance of the vacuum unit) and the associated financial costs (with susceptibility to breakdown, for example, blockages).
5. *Toilet systems without water (dry toilets)*:
    - *Compost toilets*: The compost in this system is formed from an aerated mixture of excrements and composting earth held in a container. For proper function, a minimum air temperature of 10 °C is required as well as regular checks (approximately once/month) of the structural composition and equipment. Regular extraction of the humus is necessary.
    - *Dehydration toilets* are based on drying excrements, which are odorless and almost sterile in dehydrated form and can be used as a soil conditioner. It is therefore important to segregate urine and feces. Urine can then be collected for agricultural applications, drained away, or evaporated. The dehydration process can be accelerated by using moisture-absorbing material or heat, such as exposure to the sun. Dehydration toilets are ideal for dry and arid climates.
    - *Water-free urinals* are made of ceramics or plastics, that must be regularly impregnated with an antibacterial agent. Odor is contained by an insert with a special sealant that has a lower specific mass than urine and is buoyed upward or by using a float. As crystalline precipitation occurs only on contact of urine with water, such deposits cannot occur in the down pipe. A further advantage of the water-free urinal is the low installation and maintenance costs (no connection to a water supply and no mechanical or electrical flushing fixtures).

It has been clearly demonstrated that urine sorting toilets are feasible (e.g., Sweden has more than 3000 installations; Ref. 7). Drawbacks have been observed from too small diameters of urine pipes that clog from scaling. For waterless separating toilets, one major problem remains: Men are often reluctant to sit down to urinate. This would cause a loss of urine and mix urine with feces. A luxury solution for this problem would be a private waterless urinal.

Dry toilets will fail if those who plan/design do not understand the basics, wrong materials are used, or there is poor workmanship, and users are not involved and sufficiently instructed. “Fail” usually means that the content of the processing chambers turns wet, resulting in odors, fly breeding, and incomplete sanitation.

Example of an Alternative Treatment System

Figure 4 shows an alternative treatment system that is suitable for single houses and rural settlements and is based on sorting toilets (6). The example presented is more suitable for developed countries due to the number of technical systems used.

Urine from separating toilets and waterless urinals flows into a storage tank where it stays until it is used mainly for agricultural purposes. The storage period should be at least half a year because this is an appropriate time for collection and part of the eventual medical residues can be destroyed during this period.

Feces are flushed by an appropriate amount of water (e.g., 4 or 6 liters) and discharged into one chamber of a two-chamber composting tank (with filter floor or filter bag) where the solids are precomposted. After a 1-year collecting, dewatering, and composting period, the flow is directed to the second chamber; the first one is not fed for 1 year. This allows further dewatering and precomposting and makes removal from the tank safer (although the matter is not sanitized then).

The products removed from the composting tank are brought further to full composting. They could be

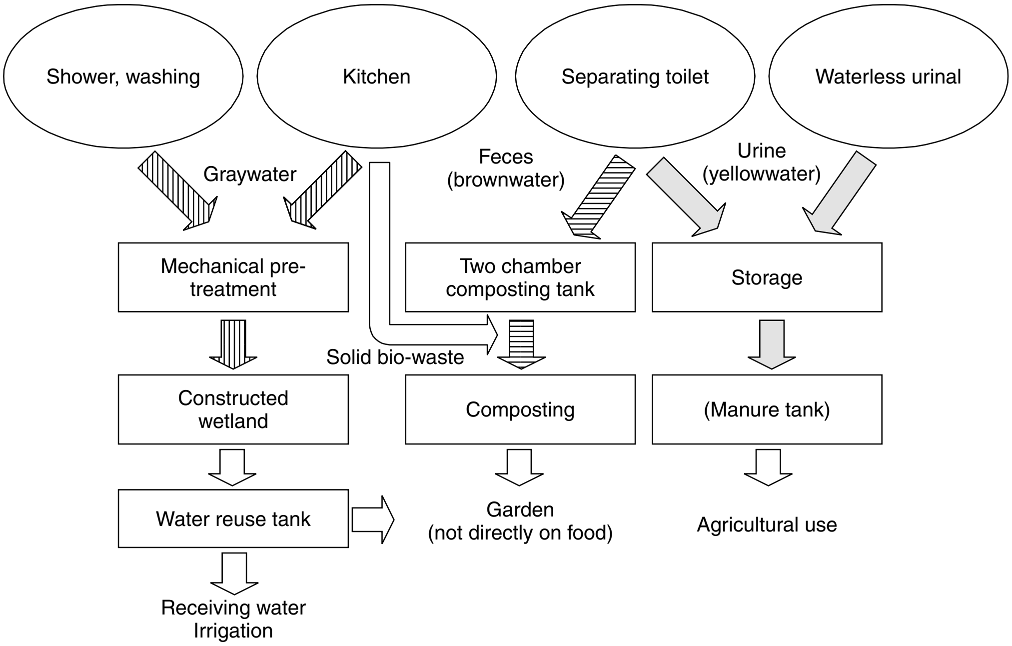

Figure 4. Concept for an alternative treatment system.

mixed with solid kitchen and garden waste to decompose completely and allow further sanitizing. The filtrate from the composting tank is low in nutrients due to the previous separation of urine. Therefore, the filtrate can be treated together with the graywater unless high-quality reuse is planned.

Graywater is pre-treated mechanically and afterward treated by a vertical-flow constructed wetland, intermittently loaded. The treated graywater is either stored and used in the garden and for irrigation or is discharged.

SUMMARY

The underlying aim of ecological sanitation (EcoSan) is to close nutrient and water cycles to contribute to sustainable development. Single technologies are only a means to an end to reach EcoSan goals. EcoSan-technologies, therefore, may range from nearly natural wastewater treatment techniques to separating toilets, simple household installations, and to complex, mainly decentralized systems. Technologies are not ecological per se but only in relation to the observed environment.

The main objectives of sanitation systems are that they have to minimize health risks and protect the environment. Ecological sanitation systems, additionally, have to return nutrients to the soil and conserve valuable water resources. They have to be affordable and therefore accessible to the world's poorest people, acceptable, aesthetically inoffensive, and consistent with cultural and social values, simple and robust in design and operation, and as comfortable as conventional systems.

A sanitation system consists of the following components: nature (climate, water, and soil as main factors), society (including settlement pattern, attitudes, beliefs and taboos related to human excreta, and the economic status of the community), the processes occurring (that convert human excreta into a nondangerous, inoffensive, useful product), and the device (the on-site structure specifically built for defecation and urination). These components have to be considered together when designing sanitation systems and making them work.

EcoSan systems have a number of advantages that can be summarized as follows:

- *Advantages to the environment*: If EcoSan systems could be adopted on a large scale, it would protect groundwater, streams, lakes, and the sea from fecal contamination. Less water would be consumed. Farmers would require smaller amounts of commercial fertilizers, much of which today washes out of the soil into water, thereby contributing to environmental degradation.
- *Advantages to households and neighborhoods*: Urine separating systems, if properly managed and maintained, do not smell or produce flies and other insects. This is a great advantage over ordinary pit toilets. Urine and feces do not come into contact and produce odor. Moisture levels are too low for fly breeding. Over half the population of the developing world has no sanitary system for excreta disposal. The market for appropriate sanitation devices is enormous, and there is a big demand. The majority of separation toilets do not require expensive or high-tech equipment. Jobs could be created for builders and for collectors of urine and sanitized feces. These products can be sold to farmers or used in the garden.
- *Advantages to municipalities*: Municipalities all over the world are experiencing greater and greater difficulty in supplying water to households and neighborhoods. EcoSan systems do not use these scarce water resources and may create therefore,

a more equitable allocation of water to rich and poor households.

A wastewater infrastructure is usually built for extremely long service. The lifetimes of existing house installations, sewerage systems, and treatment facilities have to be considered. A change to EcoSan systems is easier for newly constructed settlements or rehabilitation of complete houses. The lifetime of a house installation is far shorter than that of sewerage systems. Components of source control sanitation could be installed in each renovated flat and be connected to conventional systems first. This can be economical based on the water saving from the beginning; later, after conversion of a group of houses, separate treatment can be implemented.

Acknowledgments
Special thanks to Elke Müllegger, Martin Edthofer, and Phil Russell for their work in proofreading.

BIBLIOGRAPHY

1. Agenda 21. (1992). *The United Nations Program of Action from Rio*. United Nations, New York.
2. Esrey, S., Gough, J., Rapaport, D., Sawyer, R., Simpson-Hèbert, M., Vargas, J., and Winblad, U. (1998). *Ecological Sanitation*. SIDA, Stockholm, Sweden.
3. Lange, J. and Otterpohl, R. (2000). *Abwasser—Handbuch zu einer zukunftsfähigen Wasserwirtschaft (Wastewater—Manual for sustainable water management)*, 2nd Edn. MALLBETON-Verlag, Donaueschingen-Pfohren, Germany [*in German*].
4. Otterpohl, R. (2002). Resource efficient wastewater concepts—technical options and initial experience. *Proceedings IFAT 2002*. Munich, Germany.
5. Parry-Jones, S. (1999). *Optimising the Selection of Demand Assessment Techniques for Water Supply and Sanitation Projects*. Final Report Project/Task No: 207. Water and Environmental Health at London and Loughborough (WELL), UK.
6. Lens, P., Zeeman, G., and Lettinga, G. (Eds.). (2001). *Decentralized Sanitation and Reuse—Concepts, Systems and Implementation*. IWA Publishing, London, UK.
7. Johansson, M., Jönsson, H., Höglund, C., Richert, S. A., and Rodhe, L. (2001). *Urine Separation—Closing the Nutrient Cycle*. Report, Stockholm Water Company, Stockholm, Sweden.

WASTE TREATMENT IN FISH FARMS

Alexander Brinker
Fischereiforschungsstelle des
Landes Baden-Württemberg
Langenargen, Germany

Asbjorn Bergheim
RF-Rogaland Research
Stavanger, Norway

WASTE TREATMENT: GENERAL ASPECTS

With increasing aquacultural production, there is a greater need to reduce the amount of waste in fish farm effluents. These days, simple dilution is no longer considered an appropriate treatment (1). Generally, waste in fish farms is treated for two reasons: (1) in recirculating systems, it can negatively affect the fish and (2) in open systems, it can negatively affect the environment. Pollution control in general and in fish farming in particular should aim first at prevention rather than a cure, which can be achieved by using the best available technology (BAT) and best management practices (BMP). These have been reviewed and critically discussed in a recent comprehensive workshop (2). The use of highly digestible diets and stress-free husbandry have greatly reduced waste production per quantity of produced fish; for example, salmonid farms have achieved a reduction of about 80% in the last twenty years (3). However, despite BAT and BMP, the effluent often still needs to be treated and aquacultural waste management still relies heavily on end-of-pipe solutions. For treatment to be economically sustainable in terms of capital costs, running costs, and space requirements, some specific features of aquacultural waste loading must be taken into account.

The waste loads in aquacultural facilities have some properties that make treatment difficult. These include very low but strongly fluctuating concentrations of solids and nutrients and high flow. Table 1 compares some important effluent load parameters in salmonid farms with those in municipal waste waters. (Fig. 1). The average effluent load from a land-based flow-through fish farm is normally much lower than that in the treated water leaving a sewage plant. However, the flow from fish farms can be very high—greater than 500 $L\cdot s^{-1}$—and the total solid load can vary by two orders of magnitude in the course of a single day. Aquaculture treatment devices have to meet these diverse challenges but without the benefit of many of the technologies applied in sewage plants such as flocculation chemicals or biological treatments, which are often too expensive or otherwise impracticable under fish farm conditions.

With respect to effluent treatment, it is important to differentiate between four types of fish culture systems: (1) flow-through fish farms, (2) recirculating aquaculture systems (RAS), (3) open net cages, and (4) pond or integrated aquaculture.

In flow-through fish farms, the water quality for the fish is ensured by the steady discharge of wastes with the bulk flow. The relative loading is usually low, but due to the high flow rates, the total load in the recipient water body may become elevated, thus causing environmental problems. Waste treatment in flow-through systems is almost exclusively restricted to mechanical techniques. In RAS, where only compensatory water is added, wastes accumulate and remain within the system. The accumulated wastes must be removed to maintain physiologically adequate conditions in the culture water. In this case, waste control usually involves mechanical separation for the removal of particulate matter and the use of some kind of biofilter to remove dissolved compounds. In open net cages, waste control is very difficult as the waste can pass freely in nearly all directions. Particle traps provide the only means of collecting settled uneaten feed or fecal pellets. New

Table 1. Waste Loads of Flow-Through Fish Farms (Range of Values from Germany) Compared With Municipal Waste Loadings (11)

Source	Total Suspended Solids, mg·L^{-1}	Total Phosphorus, mg·L^{-1}	Total Nitrogen, mg·L^{-1}	BOD$_5$^{a}, mg·L^{-1}
Salmonid farms	0.5–10	<0.1	<1.4	0.5–5
Municipal wastes				
Low strength	350	4	20	110
Medium strength	720	7	40	190
High strength	1230	12	70	350

^{a}BOD$_5$, biological oxygen demand in 5 days.

advanced waste control measures use adjacent cultures of seaweeds, filtering organisms like mussels, or sediment feeders to transform the wastes near the net cages into valuable products. In pond or integrated aquaculture, the waste from one species is beneficial to other species and treatment is generally not desired. Some pond systems employed in Asian countries raise ducks or chickens on platforms from which manure falls directly into ponds stocked with filtering and omnivorous fish. The overflow may be used to irrigate paddy fields.

The principal wastes that can reasonably be expected to be removed from fish farm effluents comprise solids, biodegradable organics, and phosphorus and nitrogen compounds, including toxic ammonia and nitrite. Particles can be removed in a comparatively easy and effective manner; but the removal of the soluble components of effluents requires more sophisticated methods that must often be designed specifically for the culture system in use, taking account of the individual circumstances. Table 2 gives an overview of different waste handling systems and their capacities and restrictions. Irradiation and ozonation to control pathogens are special cases and will not be considered further here.

SOLID REMOVAL

Solids can be removed using settling basins, microscreen filters, granular media filters, and flotation (by air or foam fractionation as well as by ozonation). The two latter techniques are rarely applied in fish farming practice. The effectiveness of mechanical treatment depends on the particle size distribution and the ratio of bound particles to dissolved wastes. Results of empirical studies show considerable variations in the size distribution of suspended solids and in the efficiency of mechanical treatments (4).

Settling basins are the most traditional and widespread method for removing solids. They are simple to operate and have moderate costs; but as they are dependent on flow and sinking velocity of the particles, they require large

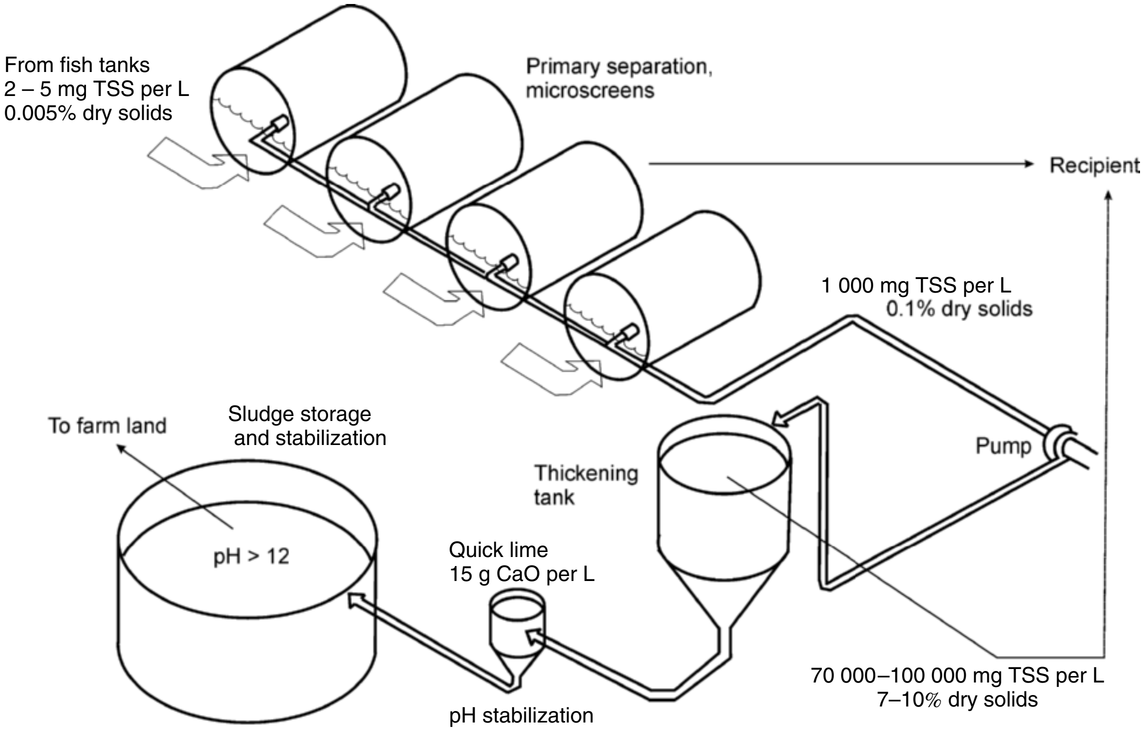

Figure 1. Diagram of an integrated effluent treatment and sludge processing system (2).

Table 2. Different Separation Processes for Aquacultural Facilities, Removal Potential, Advantages, and Disadvantages

Waste System	Treatment Types	Effective Removals	Advantages	Disadvantages
Sedimentation	Settling basin, swirl settler	Solids larger than 100 μm	Robust, easy to handle, cheap	High space requirements, re-release of wastes from settled solids, efficiency depends on TSS concentration
Microscreen	Drum filter, disk filter, Triangel filter	Solids larger than 30 μm	Low space requirements, easy to handle, tolerating high hydraulic loads, little TSS dependency	High capital and energy costs
Granular media filter	—	Solids larger than 15 μm, dissolved matter	Dual system removing particles and dissolved matter, low space requirements	Large amount of backwash water, suboptimal removal of solids and dissolved matter, high energy costs
Flotation	Dissolved air systems, dispersed air systems	Solids smaller than 30 μm	Capable of removing very small solids, low space requirements	Insufficient removal of large solids, high energy costs, dependent on bubble sizes and foam stability
Ozonation	—	Organic solids and dissolved matter, disinfections	Combines waste removal and control of pathogens	Toxic substance, system must be built with inert materials to avoid corrosion, high costs, underflow only
Biofilter	Submerged filter, trickling filter, rotating biological contactors, pressurized-bead filters, fluidized-bed biofilters	Solids and dissolved matter	Able to treat total wastes including toxic ammonia and nitrite	Costs sensitive to pH fluctuations, temperature, growth phase etc., low hydraulic loading capacity
Constructed wetlands	No standards available	Solids and dissolved matter	Very high removals (more than 90% of most compounds), can be integrated into the environment	High capital costs, each constructed wetland has to be developed for the individual situation, no long-term data available, dependent on climatic conditions

spaces. Solids settle as a result of their density. As the density of solids in aquaculture is typically close to that of water with values of 1.005–1.2, very small solids settle very slowly. Effective removal therefore requires very long hydraulic residence time. As a result, settling basins are often unable to effectively treat total suspended solid (TSS) loads lower than 10 mg·L^{-1} (dry weight) or solids smaller than 100 μm, and they do not attain solid effluent concentrations of less than 6 mg·L^{-1} (dry weight) (5,6).

Microscreen filters act as sieves that strain suspended particles larger than the mesh size. The collected particles are cleared from the screen with a dispersed hydraulic jet, thereby generating a backwash volume of 0.2–2% of the bulk flow. According to Cripps (7), mesh sizes below 60 μm give little advantage in solid removal but superproportionally decrease hydraulic capacity and increase backwash water. TSS removal in flow-through farms averages 68–80% compared with 22–70% in RAS. In practice, there are three main varieties of microscreen, namely, drum filters, Triangel filters, and disk filters.

Swirl settlers or hydroclones utilize the "tea-cup effect" generated by water injected at the outer radius of a conical tank. This effect improves settlement of the particles and forces them to the center of the tank, where they are removed from the flow. This process has a distinct advantage over classical sedimentation, and in fish farming, swirl settlers are used in small underflows comprising just 5–10% of total flow.

In recirculating aquaculture systems—in contrast to the other systems mentioned above—small solids (less than 50 μm) constitute most of the solid mass. Small solids are difficult to remove by the methods described above, but flotation is an efficient measure. In a flotation chamber, bubbles from the bottom float upward against the bulk flow, "collecting" surface-active particles, which become attached to the bubbles. The particles are carried to the surface where the bubbles break and form a foam layer. The small solids are released into the foam, which can be separated from the bulk flow by a weir. Another method sometimes used in recirculating systems employs granular media filters, which function as both mechanical and biological filters (8). Buoyant, inert beads up to 5 mm in diameter are carried through by the water stream and small particles are strained out or precipitated on the bead surface. The effects of microbial settlement on the beads are described in the next section. To clear the filter, the flow is stopped and the beads are washed forcibly backward. The backwash volume is 1–5% of the bulk flow. The dual function of this kind of filter brings considerable advantages but does not allow optimal removal of both solids and dissolved or colloid compounds.

Settling basins, flotation, and granular media filters provide opportunities for significant microbial degradation, leaching processes, and resuspension of settled waste, all of which lead to re-release of wastes back into the bulk flow. Only microscreen filters and swirl settlers ensure steady and permanent separation of solid-bound wastes from the flow that is treated.

DISSOLVED AND COLLOID COMPOUNDS REMOVAL

Ammonia and dissolved organic compounds are accessible for degradation by microorganisms attached to submerged or wetted surfaces. Ammonia is oxidized by autotrophic bacteria in a two-step process called nitrification. Biodegradable organic compounds are oxidized by heterotrophic bacteria, protozoa, and micrometazoa. When applied in aquacultural situations, this natural process is called biofiltration. The capacity of a biofilter depends on the surface area available for biological growth and on the oxygen supply. The efficiency of the biofilter depends on the accessibility of the media surface and the rate of mass transfer into the biofilm. It also varies with the growth phases of the filter (lag, log, stationary, or death phase). The voids ratio is critical for the hydraulic performance. A large variety of biofilters are available and the appropriate one must be chosen according to the individual fish farm. An overview is given by Summerfelt (6).

INTEGRATED TREATMENT SYSTEMS

Recently, there has been an increasing effort to use artificially or naturally constructed wetlands of different types for the removal of waste loads from fish farm effluents. This kind of treatment is very effective for both solids and dissolved or colloidal compounds. A current study by Michael (9) showed that an abatement pond coupled with a constructed wetland was able to remove more than 98% of solids, over 90% of biodegradable organics and phosphorus, and more than 84% of ammonia. However, further studies and monitoring are needed to assess the long-term effectiveness and economic viability of such treatments.

REUSE AND DISPOSAL OF SLUDGE

For each separation technique leading to an accumulation of sludge, it is crucial that the sludge is thickened to minimize its volume. Figure 1 gives a diagram of integrated effluent treatment combined with sludge processing. On average, the solid concentrations are increased from 2 to 5 $\mathrm{mg{\cdot}L^{-1}}$ in the effluent entering the microscreens to 7–10% dry solids in the thickened sludge, an increase of more than 20,000 times. Beneficial reuse of aquacultural sludge includes four possibilities: agricultural application, composting, vermiculture, and reed drying bed. Sludge from fish farms is rich in organic nitrogen (3–9% of dry matter) and phosphorus (1–4% of dry matter) but poor in potassium (less than 0.3% of dry matter) (10). Decomposition processes are necessary to make the nutrients accessible for plants. Heavy metal concentrations are usually below regulatory limits.

This brief review gives an insight into the tasks, challenges, and possibilities for treating aquacultural wastes. In reality, effective waste removal is often based on a combination of different waste treatment systems (11). Standard proceedings can seldom be recommended unmodified. Thus, the consultation of experts for the individual situation is often necessary.

BIBLIOGRAPHY

1. Davenport, J., Black, K., Burnell, G., Cross, T., Culloty, S., Ekaratne, S. et al. (2003). *Aquaculture: the Ecological Issues*. Blackwell Science Ltd., London.
2. Boyd, C.E. (2003). Guidelines for aquaculture effluent management at the farm-level. *Aquaculture* **226**(1–4): 101–112.
3. Bergheim, A. and Brinker, A. (2003). Effluent treatment for flow through systems and European environmental regulations. *Aquacultural Eng.* **27**: 61–77.
4. Kelly, L.A., Bergheim, A., and Stellwagen, J. (1997). Particle size distribution of wastes from freshwater fish farms. *Aquaculture Int.* **5**: 65–78.
5. Henderson, J.P. and Bromage, N.R. (1988). Optimising the removal of suspended solids from aquacultural effluents in settlement lakes. *Aquacultural Eng.* **7**: 167–181.
6. Summerfelt, S.T. (1999). Waste-handling systems. In: *CIGR Handbook of Agricultural Engineering*, Vol. II *Aquaculture Engineering*. F.W. Wheaton (Ed.). American Society of Agricultural Engineers, Baltimore, MD, pp. 309–350.
7. Cripps, S.J. and Bergheim, A. (2000). Solids management and removal for intensive land-based aquaculture productions systems. *Aquacultural Eng.* **22**: 33–56.
8. Chen, S., Timmons, M.B., and Malone, R.F. (1993). Production, characteristics, and modeling of aquaculture sludge from a recirculating aquaculture system using a granular media biofilter. In: *Techniques for Modern Aquaculture*. J-K. Wang (Ed.). American Society of Agricultural Engineers, pp. 16–25.
9. Michael, J.H.J. (2004). Nutrients in salmon hatchery wastewater and its removal through the use of a wetland constructed to treat off-line settling pond effluent. *Aquaculture* **226**: 213–225.
10. Bergheim, A., Cripps, S.J., and Liltved, H. (1998). A system for the treatment of sludge from land-based fish-farms. *Aquat. Living Resour.* **11**(4): 279–287.
11. Tchobanoglous, G., Franklin, B.L., and Stensel, D.H. (2003). *Wastewater Engineering—Treatment and Reuse*, 4th Edn. McGraw-Hill, New York.

FLOTATION AS A SEPARATION PROCESS

Kostas A. Matis
Aristotle University
Thessaloniki, Greece

The importance of the flotation process to the economy of the whole industrial world is considered enormous.

Without this process, many familiar metals and inorganic raw materials would be exceedingly scarce and costly because the high-grade ores that could be processed by simple physical and mechanical methods have long since been used up (1). Flotation initially originated from the field of mineral processing, usually termed froth flotation. For many years, various particulate solids besides minerals have been extracted from water by using this effective gravity separation method that is based on the idea of applying rising gas bubbles as the transport medium; the attachment of bubbles to particles transfers the solids from the body of water to the surface. As opposed to settling, flotation is a solid–liquid separation technique that is applied to particles whose density is lower or has been made lower than the liquid they are in. These flotation applications include mainly the treatment of water and wastewater (2,3); in the former paper, a useful discussion forum was placed on the Internet by the conference secretariat, Mr. Eero Teerikangas. Today, for example, applications of flotation exist in paper manufacturing for deinking and waste paper recycling (4), emulsified oil from various industrial wastewaters (5,6), and the separation of used plastics (7).

The typical classification of flotation processes is based on the method used to generate bubbles; so, two broad categories exist:

1. dispersed-air flotation (including electroflotation), and
2. dissolved-air flotation (DAF).

These are essentially two different techniques for bubble size, flow conditions, separation aim, and economics. Very few comparison studies, in similar experimental conditions, of these techniques are in the literature (8).

Dispersed-air flotation uses aeration, a well-known step in wastewater treatment as also is its design (guides are given by various companies, for example, Sanitaire). Fine bubble aeration is a subsurface form of diffusion, in which air is introduced in the form of very small bubbles; fine pore diffusers are usually made from ceramic, plastic, or flexible perforated membranes. Several papers have been published on bubble size and distribution, such as (9,10).

Dissolved-air flotation (see Fig. 1) is the dominant flotation technique in water and wastewater treatment. It is based on the theory of a recycle chemical reactor for its operation and on Henry's law for the dissolution of air in water as a function of pressure (1,11).

Computational fluid dynamics (CFD) codes have been also used to study the flow characteristics of this system. Multiphase flow equations for the conservation of mass, momentum, and turbulence quantities must be solved using an Eulerian–Eulerian approach in which the phases

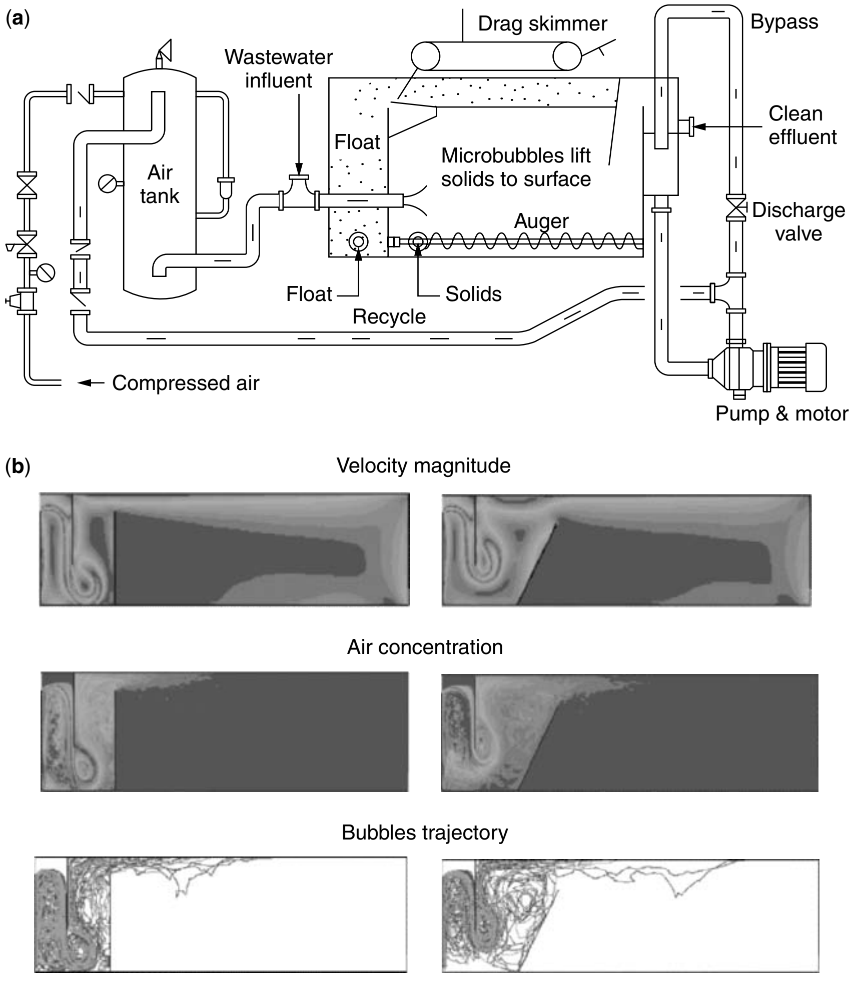

Figure 1. (**a**) The dissolved-air flotation technique (Courtesy Ecologix Environmental Systems) (top). (**b**) Impact of inclined internal baffle on the separation process obtained by CFD simulation—(unpublished results of EU METASEP project) (bottom).

are treated as interpenetrating continua. A four-phase model has been used for water, air bubbles, solid particles, and the aggregate of bubbles with particles (Fig. 1b).

Generally, low air volumes are available for dissolved-air flotation; this is not a parameter easily changed, and the reason is that the pressure in the saturation tank is raised up to 5–6 atm. For this reason, effective flocculation is a prerequisite mainly for dissolved-air flotation (12).

The trend in the development of DAF technology for potable water is to very thick microbubble beds with high flow rates, advertised as 3rd generation technology (2). A flow rate of more than 60 m/h was given, where flotation operation approaches turbulent flow. A compact (downstream filtration in the same basin) so-called re-flotation plant with hydraulic flocculation was built, for instance, for the Tampere, Finland, water and sewerage works in 1997 by the company OY Rictor AB. One of the advantages of flotation is its speed, hence its ability to operate inside a factory (Fig. 2).

Sorption combined with flotation has been investigated (8). The combination of processes is a necessity because of high existing standards and media pressure, not to mention the scarcity of water in many places. The process examined (often termed sorptive flotation) involves abstracting metal ions onto sorbents, in a fine or ultrafine size range, followed by flotation to separate metal-loaded particles. The sorbents used in the first stage were also industrial by-products (1,3). In this way, treated clear water is produced as an underflow and a foam concentrate, from which recovery of metals is possible, leading to clean technology.

A comparison between dispersed-air and dissolved-air flotation, applied to metal-loaded goethite, is presented in Fig. 3a. The recycle was the main parameter studied (at least 20% was necessary) for dissolved-air flotation (8). A flocculant was also required, particularly for DAF. Generally both flotation techniques gave good flotation recoveries; nevertheless, dispersed-air flotation looks preferable. Metal removals were near 100%.

Many industrial wastewater streams with large flow contain toxic metals, which have to be removed prior to water recycling, indirect discharge into a sewage system, or direct discharge into surface waters. Metal compounds such as copper, zinc, manganese, lead, and arsenic dominate the list of greatest releases. Mines and processing plants are often situated in water catchment areas (mountain slopes), and so environmental action plans are required to limit acid drainage and prevent pollution.

Copper recovery was investigated with a wastewater from the Assarel mine (in Bulgaria) composed of Cu(II) 240, Mn(II) 40, Fe(III) 8, Pb(II) 0.2, SO_4^{2-} 4200, Ca(II) 270, and Mg(II) 260 mg/L as the main constituents (13). Almost complete copper removal was obvious for all cases (Fig. 3b), even though the requisite xanthate dose was doubled for real effluents; the quantity needed was around 1 g/L for synthetic wastewater and around 2 g/L for real effluents. In all cases, the residual copper concentration was lower than 0.11 mg/L; the limit in Europe for indirect discharge of chemical industry effluents is 0.5 mg/L copper.

Xanthates are known to exhibit a high level of chemical reactivity for heavy metals. Precipitate flotation is the mechanism that involves first the immobilization of ions as precipitates, for example, raising the respective concentrations may lead to precipitation of the ions as a surfactant floatable product before air is injected (1). This means that it is not a solution any longer but rather a dispersion. In Reference 13, economic considerations of the various processes were also given, including microfiltration downstream.

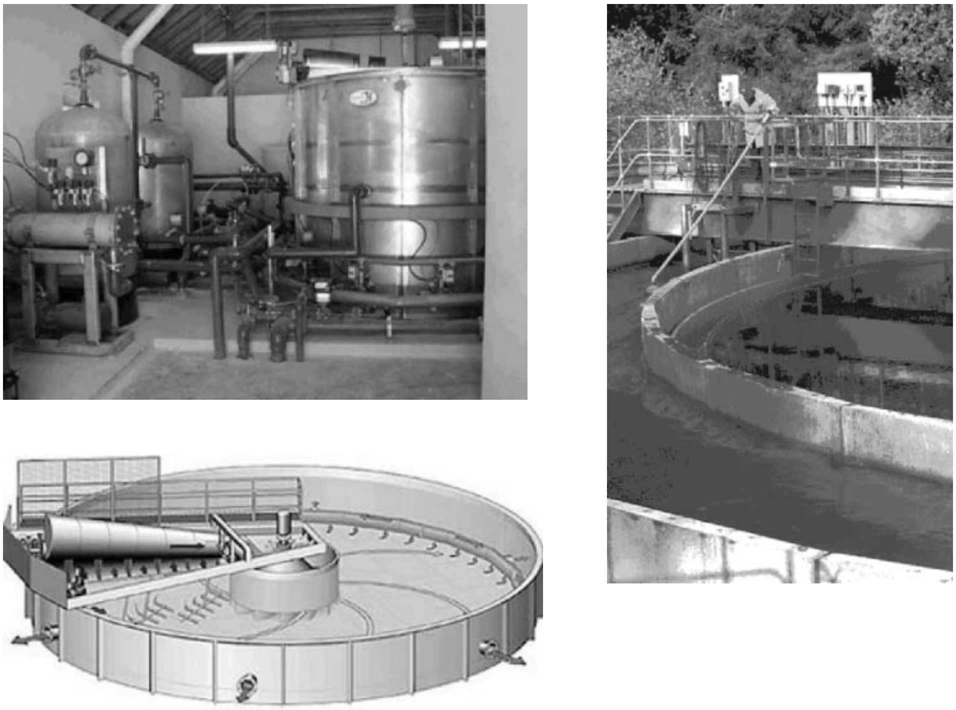

Figure 2. Possibility for the flotation process to operate in a closed factory (top) or open air. Gratitude is given to the companies Krofta (Sandfloat design), Meri Technologies (Deltapurge clarifier), and Purac Ltd. (bottom right).

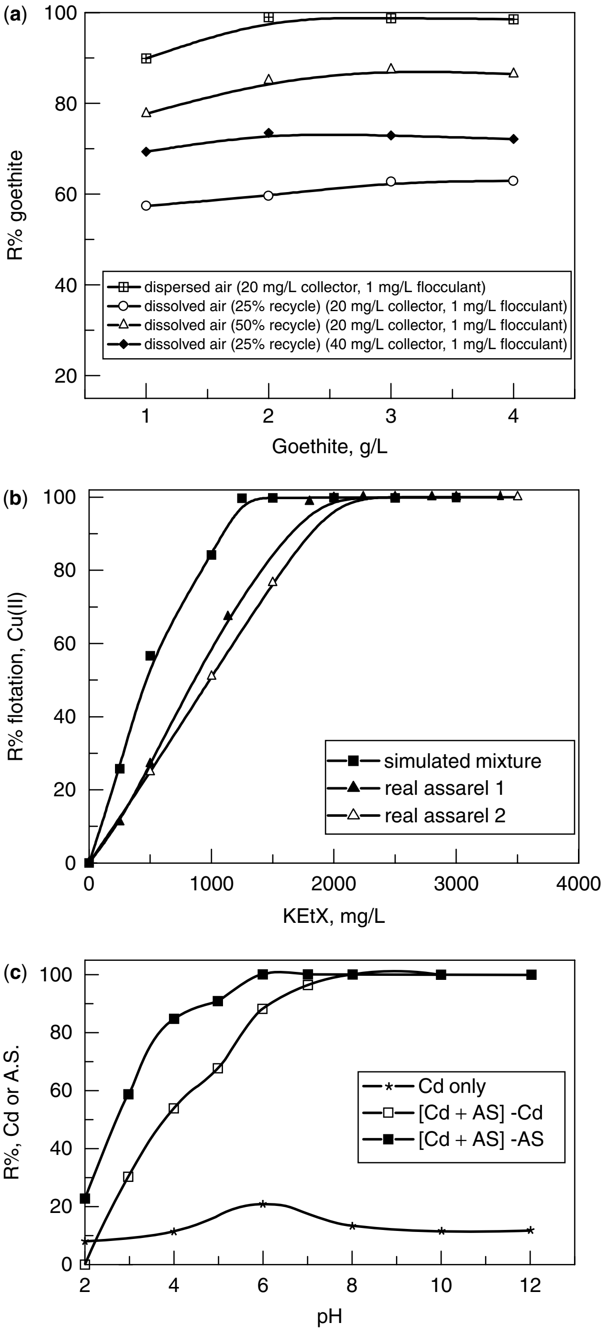

Figure 3. (**a**) Comparison of dispersed-air with dissolved-air flotation for Cr(VI) removal, applying dodecylamine as the flotation collector. Copyright Technical Faculty and Copper Institute, Bor Yugoslavia (top). (**b**) Copper ion flotation applied to a mine effluent: influence of xanthate collector concentration on copper removal. Reprinted from the periodical Hydrometallurgy, Ref. 13, copyright (2004) with permission from Elsevier. (**c**) Biosorptive electroflotation of cadmium on anaerobically digested activated-sludge biomass (denoted as AS): effect of solution pH. Reprinted from Ref. 2.

To the various abnormal reasons for serious contamination of water and soil by toxic heavy metals, one has to add ecological accidents. Like those at Aznalcóllar, Spain, in 1998, an accident in an area close to the Doñana Natural Park in South Andalusia released 2 million m^3 of tailings and 4 million m^3 of water contaminated by heavy metals into the Guadiamar River and that at Baia Mare, Romania, in 2000 where a tailings pond burst releasing approximately 100,000 m^3 of waste water containing up to 120 tonnes of cyanide and heavy metals into the Lapus River; this then traveled downstream into the Somes and Tisa Rivers into Hungary before entering the Danube.

Electrolytic flotation or electroflotation constitutes an unconventional technique in which the bubbles necessary for the flotation are produced by electrolysis of the medium. Electrodes are usually mounted horizontally and receive a low voltage current (6,14). A large amount of very small bubbles is formed with minimum turbulence. Clarification can be effected with effluents that previously would not have been considered suitable for treatment by flotation. The electric field gradient between the electrodes aids flocculation of suspended matter. Gas production, residence time, and the other operating conditions can be checked quickly and are easily controlled. No high-pressure pumps, pressurized vessels, or other complex machinery are needed, and moving parts are minimal.

When aluminum alloy electrodes were used (Fig. 3c), the results were improved; biomass recovery was possible from pH 7 approximately. Results were not highly affected by shorter process duration (10 min) and even lower current density (65 A/m^2). The two curves (of cadmium removal on biomass and of floated Cd-loaded biomass almost coincide with pH variation. Attention should be paid to metals aqueous speciation.

Since the discovery that the activated sludge treatment process can also remove metals from wastewaters, much progress in research has been made. Nowadays, it is known that biosorption, using appropriate usually nonliving biomass, is one of the innovative methods for removing toxic metals from effluent; the more conventional methods are precipitation, adsorption, and ion exchange. The process is based on several mechanisms; the most important are physical adsorption, ion exchange, surface complexation, and surface precipitation. The source of biosorbents can be an industrial waste material, as in the application of biomass from large-scale fermentation processes such as filamentous fungi and bacteria from pharmaceutical production (8).

The cost of an electroflotation device plus a sand filter compared favorably with a conventional treatment system using cyanide oxidation/alkaline precipitation/polymer-aided clarification (15). By the former system, a saving of approximately 43% was found when cyanide is present in the groundwater.

If required, the sorbents may be recycled after metal desorption, which is also no problem. A two-stage countercurrent scheme was selected based on extensive studies, with one parallel of elution; flotation was included following each respective biosorption stage (16). This scheme is similar to a conventional activated-carbon system for separating liquids. The partially loaded biomass was guided from the polishing to the leading biosorption stage, where the wastewater feed was introduced.

Concluding, flotation applications (apart from mineral processing) include mainly the treatment of water and wastewater, but also bacteria, coal, clays, corn, resins, protein, fats, rubber, dyes, glass, fruit juices, and cane sugar. Flotation is beneficial for a large variety of dissolved chemical species (e.g., ions and molecules), which can either be concentrated from the solution/dispersion or may be selectively separated from one another, leading to sustainability. Sustainable water and wastewater treatment plant is a combination of technology and practice that meets the multiple and changing requirements of the society in optimal economical and ecological ways during the life cycle of a treatment plant. The flotation techniques applied are often based on the differences in surface activity of the various substances. They have been called, according to the removal mechanism followed, ion flotation or sorptive flotation. Usually, the separation action is aided or induced by introducing appropriate chemical reagents.

Acknowledgments

Many thanks are due to my good colleagues in the Flotation group at AUTh, for the more than 20 years of fruitful collaboration and to the European Union for funding our environmental research projects.

BIBLIOGRAPHY

1. Matis, K.A. and Zouboulis, A.I. (1995). An overview of the process. In: *Flotation Science and Engineering*. K.A. Matis (Ed.). Marcel Dekker, New York, p. 1.
2. IWA. (2000). *Proc. 4th Intl. Conf. Flotation in Water and Waste Water Treatment*. Finnish Water and Wastewater Works Assoc., Helsinki, Finland.
3. Rubio, J., Souza, M.L., and Smith, R.W. (2002). Overview of flotation as a wastewater treatment technique. *Miner. Eng.* **15**: 139.
4. Finch, J.A. and Hardie, C.A. (1999). An example of innovation from the waste management industry: Deinking flotation cells. *Miner. Eng.* **12**: 467.
5. Zouboulis, A.I. and Avranas, A. (2000). Treatment of oil-in-water emulsions by coagulation and dissolved-air flotation. *Coll. Surf. A*. **172**: 153.
6. Backhurst, J.R. and Matis, K.A. (1981). Electrolytic flotation in effluent treatment. *J. Chem. Tech. Biotechnol.* **31**: 431.
7. Frenay, J. (1998). Recycling of plastics: Application of flotation. In: *Mineral Processing and the Environment*. G.P. Gallios and K.A. Matis (Eds.). Kluwer Academic, Dordrecht, The Netherlands, p. 295.
8. Matis, K.A. and Lazaridis, N.K. (2002). Flotation techniques in water technology for metals recovery: Dispersed-air vs. dissolved-air flotation. *J. Min. Metall. A* **38**: 1.
9. Burns, S.E., Yiakoumi, S., and Tsouris, C. (1997). Microbubble generation for environmental and industrial separations. *Sep. Purific. Tech.* **11**: 221.
10. Filippov, L.O., Joussemet, R., and Houot, R. (2000). Bubble spargers in column flotation: adaptation to precipitate flotation. *Miner. Eng.* **13**: 37.
11. Zabel, T.F. (1992). Flotation in water treatment. In: *Innovations in Flotation Technology*. P. Mavros and K.A. Matis (Eds.). Kluwer Academic, Dordrecht, The Netherlands, p. 431.
12. Valade, M.T., Edzwald, J.K., Tobiason, J.E., Dahlquist, J., Hedberg, T., and Amato, T. (1996). Pretreatment effects on particle removal by flotation and filtration. *JAWWA* **88**(12): 35.
13. Lazaridis, N.K., Peleka, E.N., Karapantsios, Th.D., and Matis, K.A. (2004). Copper recovery from effluents by various separation techniques. *Hydrometallurgy*. **74**: 149.
14. Matis, K.A. and Zouboulis, A.I. (2001). *Metals loading on sorbents and their separation*. *ESF Workshop In-Situ Reactive Barriers versus Pump-and-Treat Methods for Groundwater Remediation*. BAM, Berlin.
15. Poon, C.P.C. (1997). Electroflotation for groundwater decontamination. *J. Hazardous Mater.* **55**: 159.
16. Matis, K.A., Zouboulis, A.I., and Lazaridis, N.K. (2003). Heavy metals removal by biosorption and flotation. *Water Air Soil Pollut. Focus* **3**: 143.

DEGRADATION OF CHLORO-ORGANICS AND HYDROCARBONS

Frank J. Castaldi
Brown and Coldwell
Austin, Texas

Chloro-organics include chlorinated aromatic and aliphatic compounds. The chemical characteristics of chlorinated aromatic compounds cover a range of compounds that include chlorophenols, chlorobenzenes, chloroanilines, polychlorinated biphenyls, and many pesticides. Chlorinated aliphatic compounds make up a major portion of the Environmental Protection Agency (EPA) Hazardous Substances List of halogenated volatile constituents.

Typical chlorinated aliphatics include trichloroethylene, tetrachloroethylene, 1,1-dichloroethylene, 1,2-dichloroethylene, carbon tetrachloride, chloroform, 1,1, 1-trichloroethane, and chloroethylene. Hydrocarbons are petroleum-based compounds that contain carbon and hydrogen. Hydrocarbons can be separated into aliphatic, alicyclic, and aromatic hydrocarbon groups. Aliphatic hydrocarbons are straight or branched-chain hydrocarbons of differing length. These hydrocarbons are divided into alkane, alkene, alkyne, and unsaturated alkyl groups. Alicyclic hydrocarbons include the components of petroleum oils and complex substituted compounds. These hydrocarbons can be grouped into cycloalkanes and cycloalkenes. Aromatic hydrocarbons are comprised of benzene and its derivatives and polynuclear aromatic hydrocarbons. The potential and extent of substance degradation of each of these compound groups differs with the complexity of the compound and the nature of the aquatic environment. The principal focus of this discussion is that of microbial degradation in water and wastewater.

DEGRADATION OF CHLORINATED AROMATIC COMPOUNDS

Chlorinated aromatic compounds are typically produced for direct commercial use and as chemical intermediates during chemical synthesis. They are used in industrial operations to produce dyes, pharmaceuticals, pesticides, synthetic rubbers, solvents, cleaners, deodorants, pigments, and related formulations.

Microbial Transformations

Chlorinated aromatic compounds include chlorophenols, chlorobenzenes, chloroanilines, polychlorinated biphenyls, and many pesticides. Bacteria can remove chlorine from aromatic compounds by several different mechanisms. In aerobic systems, a hydroxyl group prior to ring cleavage can replace the chlorine atom. However, the more usual route involves removing the chlorine after ring cleavage through bacterial metabolism by a modified ortho pathway. Opening of the aromatic ring under aerobic conditions requires the presence of two hydroxyl groups. The hydroxyl groups derived from molecular oxygen can be added to the ring in reactions catalyzed by either monooxygenase or dioxygenase enzymes. Generally, the enzymes involved in the biodegradation of chlorinated aromatic compounds have broad substrate specificity that allows them to degrade or detoxify a number of other compounds (e.g., benzene, toluene, ethylbenzene, phenol) that also support growth (1).

Monochlorophenols are readily biodegradable compounds; degradation proceeds rapidly when they are exposed to acclimated cultures of activated sludges in conventional biological wastewater treatment processes. Dichlorophenols are also readily biodegradable, but degradation proceeds at a lower rate than that of monochlorophenols. Pentachlorophenol (PCP) is also biodegradable, but its degradation is very slow, and bacteria must be acclimated for prolonged time periods before effective biological treatment can be achieved. The biodegradation of chlorophenols is characterized by an increasing period of acclimation and a decreasing degradation rate in the following order (2):

Acclimation period for

Phenol < monochlorophenol < dichlorophenol
< trichlorophenol < tetrachlorophenol
< pentachlorophenol

Rate of biodegradation for

Phenol > monochlorophenol > dichlorophenol
> trichlorophenol > tetrachlorophenol
> pentachlorophenol

The rate of biodegradation is inversely proportional to the number of chlorine atoms in the molecule for most chlorinated aromatic compounds. This is demonstrated by chlorobenzoic acids. Monochlorobenzoic acids are readily biodegradable, but most dichloro derivatives and trichlorobenzoic acids are much more stable. For polychlorinated biphenyl (PCB), there is also a relationship between the rate of biodegradation and the number and position of chlorine atoms in the molecule. The resistance of PCB in the environment is directly proportional to the number of chlorine atoms in the molecule. Chloro-substituted diphenylmethanes behave similarly. The rate of chlorinated aromatic compound biodegradation in the activated sludge process usually decreases in the order of mono-, di-, tri-, tetra-, and pentachloro derivatives, depending on the position of the chlorine in the molecule (3).

Chloroanilines are biodegraded more slowly than chlorophenols because of the presence of both Cl and NH_2 groups that retard the biodegradability of the compound. As for chlorophenols, an increased number of chlorine atoms in the chloroaniline molecule results in increased resistance to biological treatment by the activated sludge process. There is also a retardant effect due to the number, type, and position of substituents on the aromatic ring. In addition to the halogens, the substituents that retard biodegradation are NH_2, NO_2, and SO_3H. These substituents retard the rate of biodegradation by decreasing the electron density on the aromatic ring (3). Alternatively, substituents that facilitate biodegradation by increasing electron density on the aromatic ring are OH, COOH, CHO, and CH_3. Consequently, the susceptibility of chloroaniline to biodegradation depends on whether the NH_2 group is readily transformed to OH by oxidative deamination. In certain compounds, a substituent (e.g., halogens) is not eliminated before the bacterium cleaves the aromatic ring. Chloroanilines are biodegraded most slowly when the $Cl–NH_2$ group is meta-substituted, whereas para-substituted derivatives are degraded rapidly (4).

The enzyme toluene dioxygenase catalyzes the hydroxylation of chloro-substituted benzenes (e.g., chlorobenzenes) as well as methyl- and chloro-substituted phenols. The dioxygenase system for initial attack, combined with the chlorocatechol degradation pathway, allows complete degradation of a range of industrial chemicals (1). Because of the nonspecific nature of the enzymes that transform benzoate to catechol, many aerobic bacteria can cometabolize chlorinated aromatic compounds. However, this biodegradation is not complete because chlorinated benzoates and catechols are the final end product of the oxidative biodegradation (5). The complete aerobic mineralization of chlorinated aromatic compounds is not typically seen, and the persistence of these compounds in the environment illustrates the ineffectiveness of bacteria for these degradations.

It is known that bacteria can mediate the anaerobic dehalogenation of chlorobenzoate compounds. Reductive dehalogenation has been confirmed for a number of chlorinated aromatic acids, chlorobenzenes, chlorophenols, chlorophenoxyacetate, herbicides, and PCBs (6). During reductive dechlorination, the chlorinated aromatic compound serves as the electron acceptor. Theoretical calculations of the Gibbs free energy available from dehalogenation indicates that bacteria can benefit from the use of chlorinated aromatic compounds as electron acceptors under anaerobic conditions (7). Nevertheless, most anaerobic dehalogenation is probably the result of cometabolism.

Methanogenic metabolism has been used successfully to dehalogenate a number of chlorinated aromatic compounds, including 2,4,5-trichlorophenoxyacetate, 3-chlorobenzoate, 2,4-dichlorophenol, 4-chlorophenol, 2,3,6-trichlorobenzoate, and 2,4-, 2,5-, 2,6-, 3,4-, and 3,5-dichlorobenzoates (6,8). Haggblom et al. (9) found that methanogenic bacteria preferentially remove ortho-substituted chlorine and that the meta- and para-substituted chlorine atoms are removed at much slower rates. Methanogenic metabolism does not result in the final mineralization of chlorinated aromatic compounds. To achieve mineralization, a balanced microbial consortium is required to perform the sequential biodegradations so that the transformation products are not more toxic than the parent compound.

Anaerobic dechlorination of aromatic compounds can be stimulated by adding electron donors. The addition of volatile fatty acids (i.e., acetate, butyrate, propionate) and ethanol increases the rate of dechlorination and the extent of biodegradation of many chlorinated aromatic compounds (8). Haggblom et al. (9) found that the addition of *p*-cresol and propionate enhances the methanogenic degradation of 2,4-dichlorophenol.

The presence of electron acceptors also influences the extent of anaerobic dechlorination in the environment. Certain electron acceptors may block the desired reduction reactions through competitive inhibition of methanogenic metabolism. However, Haggblom et al. (9) report that sulfate-reducing conditions actually enhance the apparent rates of biodegradation of 4-chlorophenol, 3-chlorophenol, 2-chlorophenol, and 2,4-dichlorophenol. Chlorophenols act as a source of carbon and energy and are degraded under sulfidogenic conditions by sulfate-reducing bacteria. In groundwater environments that contain sulfate-reducing bacteria, sulfate may be the preferred electron acceptor, and methanogens and sulfate reducers probably compete for suitable electron donors.

DEGRADATION OF CHLORINATED ALIPHATIC COMPOUNDS

Chlorinated aliphatic compounds are typically used as solvents for cleaning and reagents for chemical synthesis. The uses include paint and ink formulations, dry cleaning, synthetic rubber production, fumigants, paint and varnish removers, degreasers, pesticide solvents, adhesives, photographic supplies, pharmaceutical products, and household and office supplies.

Microbial Transformations

Microbial degradation of chlorinated aliphatic compounds can use one of several metabolic processes. These include oxidation of chlorinated alkanes for an energy source, cometabolism under aerobic conditions, and reductive dechlorination under anaerobic conditions. The response of different chlorinated aliphatic compounds to these metabolic processes differs, depending on the nature of the contamination, the redox condition, and the available electron acceptors.

When molecular oxygen is the electron acceptor, one- and three-atom substituted chlorinated aliphatic compounds are transformed by three types of enzymes: oxygenases, dehalogenases, and hydrolytic dehalogenases (10). The transformation products of oxygenases are alcohols, aldehydes, or epoxides. Dehalogenase transformation products are an aldehyde and glutathione. The glutathione is required as a cofactor for nucleophilic substitution by the dehalogenase enzyme. Hydrolytic dehalogenases hydrolyze the aliphatic compound and yield alcohols as a transformation product. Higher chlorinated compounds, particularly when all available valences on carbon are substituted (e.g., tetrachloride or tetrachloroethylene), have not been transformed under aerobic conditions. They must be transformed by reductive dechlorination.

Chlorinated aliphatic compounds may be either oxidized or reduced, depending on their structure and the redox potential of the aqueous environment where they are found. Reduction is possible because of their electron negative character. Consequently, polychlorinated aliphatic compounds often behave as an electron acceptor or the oxidant in a redox reaction. The more chlorinated a compound, the higher its oxidation state, and the more susceptible it is to reduction. This explains why the rate of dechlorination decreases under anaerobic metabolic conditions as, for example, tetrachloroethylene is dechlorinated to vinyl chloride. The reaction proceeds as follows:

$$\text{tetrachloroethylene} >>>> \text{trichloroethylene} >>> \text{1,2-dichloroethylene} >> \text{vinyl chloride} > \text{ethylene}$$

Vinyl chloride (monochloroethylene) is more reduced, so the thermodynamic equilibrium tends to stabilize vinyl chloride as the typical end product of trichloroethylene degradation in aquifers that have negative redox potentials (i.e., an anaerobic environment).

Aerobic Processes

Biotransformation of some chlorinated aliphatic compounds has been demonstrated under aerobic conditions (11). Under aerobic conditions, many soil microorganisms can oxidize vinyl chloride. Most chlorinated aliphatic compounds are eventually mineralized to carbon dioxide. The aerobic degradation capabilities of these microorganisms for chlorinated aliphatic compounds have provided successful treatment processes in seeded (i.e., cultured) activated sludge reactors.

Some chlorinated aliphatic compounds are degraded by cometabolism under conditions that support aerobic metabolism. These aerobic microbes generate oxygenase enzymes of broad-substrate specificity that oxidize chlorinated aliphatic compounds. These include microorganisms that belong to the genera *Alcaligenes, Mycobacterium, Pseudomonas, Nitrosomonas, Xanthobacter*, and *Ancylobacter* (12). For example, *Nitrosomonas europaea* catalyzes the aerobic transformation of vinyl chloride, *cis*- and *trans*-dichloroethylene, *cis*-dibromoethylene, and trichloroethylene (13).

During the cometabolism of chlorinated alkenes, other microorganisms derive their energy from organic compounds such as methane, propane, phenol, and toluene. Several toluene-using microorganisms can degrade trichloroethylene by cometabolism. Even under varied pH and temperature, significant rates of microbial degradation are measured (14). Trichloroethylene degradation decreased by 30% at 4 °C compared with that at 30 °C (14). Phenol-oxidizing bacteria have a much higher capacity to degrade trichloroethylene than methane-oxidizing microorganisms (i.e., methanotrophs). Trichloroethylene degradation by phenol-oxidizing bacteria reportedly removed greater than 90% of TCE after phenol injection (15). *In situ* studies of trichloroethylene degradation have demonstrated that phenol-using microorganisms can be readily stimulated in the environment. Phenol addition is a good primary substrate; it achieves degradation of *cis*- and *trans*-dichloroethylene and trichloroethylene *in situ*. The apparent rate of chlorinated alkene transformation increases as phenol concentrations increase.

Many investigators have confirmed the biotransformation of trichloroethylene by methanotrophic bacteria using methane as a primary substrate. These methanotrophs have the monooxygenase enzyme that will incorporate one oxygen atom from molecular oxygen into methane to produce methanol. The monooxygenase enzyme can hydroxylate many alkanes and aromatic compounds and form epoxides from chlorinated alkenes because it is not compound-specific (10). The products of these reactions are not further oxidized by methanotrophs, so a diverse community of microorganisms is needed to achieve complete mineralization of a given constituent. Trichloroethylene has been successfully degraded aerobically to carbon dioxide with methane in air, although the rate of transformation was less than that for dichloroethylenes (16). However, there are toxicity problems because some trichloroethylene oxidation products are toxic to many methanotrophic bacteria. It also appears that trichloroethylene concentrations greater than 50 mg/L inhibit methane use by methanotrophs (10).

Another serious limitation of methane-oxidizing bacteria is that they cannot transform tetrachloroethylene (i.e., perchloroethylene) or higher chlorinated aliphatic compounds. The less chlorinated the compound, the greater the rate and extent of the transformation under aerobic conditions, as expected from thermodynamics. The lower the oxidation state of the compound, the less difficult it is to oxidize. Trichloroethylene is more oxidized than vinyl chloride, so it is more difficult to oxidize this compound further. As noted previously, the opposite is true when reducing an oxidized compound. The higher the degree of oxidation, the easier it is to reduce that compound.

Anaerobic Processes

Many chlorinated aliphatic compounds are transformed under anaerobic conditions. These compounds are mineralized in the presence of a diverse community of microorganisms. One of the predominant mechanisms for transforming chlorinated aliphatic compounds is reductive dechlorination. Reductive reactions result in replacing the chlorine atom by dihaloelimination. The reductive process usually occurs through cometabolism.

Chlorinated aliphatic compounds are transformed by reductive dechlorination even at low concentrations of less than 200 parts per billion (ppb). During reductive dechlorination, the chlorinated organic compound serves as an electron acceptor. The rate of dechlorination under anaerobic conditions is linked to the rate of primary substrate oxidation. Electrons from the oxidation of a primary substrate carry out the dechlorination. The control of these reactions for bioremediation requires an understanding of the redox conditions and the influence and availability of specific electron acceptors and donors on the overall metabolic condition of the bacteria that perform the reduction.

The availability of electron acceptors in anaerobic systems affects reductive dechlorination by competing with the chlorinated compounds for reducing potential. For example, nitrate and sulfate can inhibit the dechlorination of some chlorinated alkenes. It has been reported that the addition of nitrate to natural soil microcosms completely blocked the dechlorination of tetrachloroethylene (17).

According to thermodynamic principles, microorganisms couple half reactions that yield the greatest free energy. As redox conditions become more reducing, more chlorinated compounds undergo transformation. Thus more compounds are transformed under methanogenic conditions than under other anaerobic respirations (e.g., sulfate or nitrate). Additionally, several biologically active donors (e.g., acetate and H_2) and the ferrous ion have lower reduction potentials than most chlorinated aliphatic compounds. As a result, they can be involved in chlorine removal by reduction.

Sulfate also influences these reactions. Using sulfate as the electron acceptor (e.g., sulfate reduction by *Desulfovibrio sp.*), the dechlorination of tetrachloroethylene proceeds at a slower rate than if carbon dioxide is the electron acceptor (e.g., methanogenesis). Sulfate may also block the dechlorination of trichloroethylene (18). The specific influence of sulfate and other electron acceptors cannot be generalized. Different microbial systems and different aquifer and groundwater chemical conditions shift the thermodynamic equilibrium. Therefore, treatability studies are always required to assess the particular situation.

Chlorinated aromatic hydrocarbons can serve as both the electron acceptor and the electron donor in a reductive dechlorination reaction (7). However, chlorinated alkenes need an additional electron donor to support anaerobic dechlorination. Typical electron donors are the following: methanol, ethanol, glucose, sucrose, benzoate, lactate, formate, acetate, and butyrate. Volatile fatty acids produced under methanogenic conditions are generally considered the most effective electron donors for enhancing dechlorination. Because some methanogens consume hydrogen as an electron donor, microbial fermentations involve the reoxidation of a reduced electron carrier, such as nicotinamide adenine dinucleotide (NAD)H_2, as part of the metabolic reaction. This means that $NADH_2$ is oxidized to NAD and H_2 in the presence of methanogens. Usually, nonmethanogenic bacteria provide the hydrogen for methanogenesis (19).

Abiotic Transformations

Chlorinated aliphatic compounds include trichloroethylene, tetrachloroethylene, 1,1-dichloroethylene, 1, 2-dichloroethylene, carbon tetrachloride, chloroform, 1,1,1-trichloroethane, and chloroethylene (vinyl chloride). In addition to biological transformations, these compounds undergo abiotic transformations in the environment. The important abiotic transformations include substitution, dehydrohalogenation, and reduction in water (20). A typical substitution is the addition of water resulting in hydrolysis. The nucleophiles of OH^- and H_2O are the principal species responsible for abiotic dehydrohalogenation

Table 1. Ions Capable of Abiotic Displacement of Halogen[a]

Oxygenated Waters	Anaerobic Waters
H_2O	$SO_3{}^{2-}$
OH^-	$S_2O_3{}^{2-}$
Cl^-	NH_3
Br^-	$NO_2{}^-$
$SO_4{}^{2-}$	$S_n{}^{2-}$
$HCO_3{}^-$	$R\text{–}C_6H_{13}S^-$
—	$C_6H_5S^-$

[a] Reference 20.

of chlorinated aliphatic compounds in water. However, a variety of other species can displace the chlorine. These are presented in Table 1. Under anaerobic conditions, the sulfur nucleophiles are generally the most powerful. Sulfides react with chlorinated aliphatic compounds via substitution to produce mercaptans.

These reactions proceed slowly in the absence of biological activity. The half-lives for the monochloroalkanes are approximately a month at 25 °C (20). Polychlorinated species can have half-lives as long as 40 years. Stronger nucleophiles such as HS^- can reduce the half-lives of the abiotic degradation of chlorinated aliphatic hydrocarbons. Microbial enzymes also catalyze these reactions and reduce half-lives significantly.

A variety of transition metals, including nickel, iron, chromium, and cobalt, can reduce chlorinated aliphatic compounds. As a result of this oxidation–reduction (redox) reaction, the metals are oxidized. Vogel et al. (11) defined the reduction products and metals that mediate such reactions. The transition metal reduces a chlorinated compound, removing the chlorine and creating an alkyl radical that readily picks up a hydrogen atom from water, resulting in the formation of an alkane. The reduction of polychlorinated alkanes can result in both alkanes and alkenes.

DEGRADATION OF PETROLEUM HYDROCARBONS

Petroleum hydrocarbons include the common components of petroleum oils and products. They are found in wastewaters from petroleum refineries, petrochemical facilities, fuel storage and transportation facilities, and industrial organic chemical production facilities. Benzene and other single-ring aromatic hydrocarbons are used as solvents in industrial processes. Substituted forms of single-ring aromatic compounds are used in many industrial processes for preparing dyes, resins, antioxidants, polyurethane foams, fungicides, stabilizers, coatings, insulation materials, fabrics, and plastics. Polynuclear aromatic hydrocarbons (PAHs) are associated with petroleum refining and coal tar distillation. PAHs are also associated with waste by-products from coal gasification and coke production.

Microbial Transformations

Petroleum hydrocarbons can be classified into aliphatic, alicyclic, and aromatic hydrocarbon groups. Aliphatic hydrocarbons are divided into alkane, alkene, alkyne, and unsaturated alkyl groups. Alicyclic hydrocarbons can be grouped into cycloalkanes and cycloalkenes. Aromatic hydrocarbons are comprised of benzene and its derivatives and polynuclear aromatic hydrocarbons.

The presence of molecular oxygen as a terminal electron acceptor is required for successful microbial degradation of petroleum hydrocarbons. Nitrate and sulfate also serve as alternative electron acceptors during anaerobic respiration of hydrocarbons. Generally, the biodegradation rate decreases with decreasing redox potential. Only negligible biodegradation is observed under strictly anaerobic (i.e., fermentative) conditions. Consequently, hydrocarbons remain for a relatively long period of time in bottom sediments and other anaerobic portions of the aquatic environment.

Petroleum hydrocarbons are generally hydrophobic compounds. Bacteria and fungi frequently attach to oil droplets because intimate contact between a microorganism and the surface of the petroleum hydrocarbon is necessary for biodegradation. Consequently, dispersing the petroleum oil in water makes it more susceptible to microbial attack. Bacteria often produce extracellular surfactants that aid in solubilizing petroleum hydrocarbons. These bacterial surfactants are complex mixtures of proteins, lipids (e.g., rhamnolipids, phospholipids), and carbohydrates (21).

Aliphatic and Alicyclic Hydrocarbons

The biodegradation potential of alkanes is a function of carbon chain length. Short carbon chains of less than 10 carbons are more difficult to biodegrade than longer chains. Because of their higher solubility, short-chain hydrocarbons also exhibit a high degree of toxicity in the aquatic environment (22). Longer chain aliphatic hydrocarbons are more easily biodegraded than the short-chain variety. A large number of facultative anaerobic bacteria are prevalent in the aquatic environment that can use aliphatic hydrocarbons as a source of carbon and energy. These bacteria include the genera *Acinetobacter, Alcaligenes, Arthrobacter, Flavobacterium, Methylococcus, Mycobacterium, Nocardia*, and *Pseudomonas* (21). Numerous fungi and yeast also biodegrade aliphatic hydrocarbons. Although fungi are more versatile than yeast in biodegrading short-chain hydrocarbons, both are effective in using long-chain alkanes.

Aerobic biodegradation of a long-chain aliphatic hydrocarbon requires incorporating molecular oxygen into the compound. Oxygenase enzymes (i.e., monooxygenases and dioxygenases) mediate this degradation reaction (23). The pathway of alkane biodegradation is oxidation at the terminal methyl group of an alcohol and then of the corresponding fatty acid. The terminal oxidation proceeds by successive removal of two carbon units, termed the beta-oxidation sequence.

Alkene biodegradation is more varied because bacteria attack at either the methyl group or the double bond. Unsaturated straight-chain hydrocarbons are usually biodegraded less easily than saturated compounds. Consequently, bacterial metabolism results in forming intermediates that consist of unsaturated alcohols and fatty acids, primary or secondary alcohols, methyl ketones, epoxides, and diols (21). The methyl group oxidation is the more likely biodegradation pathway.

Petroleum hydrocarbons that have branch chains are less susceptible to biodegradation. Quaternary carbon and β-alkyl-branched compounds are generally considered recalcitrant and accumulate in the environment. However, combining chemical oxidative processes with biodegradation is effective for treating recalcitrant hydrocarbons.

The biodegradation of cycloalkanes is usually by oxidation of the terminal methyl group and yields a primary alcohol. Hydroxylation must occur to initiate the biodegradation of cycloalkanes. The bacteria that can oxidize noncyclic alkanes can also hydroxylate cycloalkanes. Several alternate metabolic pathways exist for microbial attack on alicyclic hydrocarbons, and numerous intermediate compounds have been identified during their degradation. Substituted cycloalkyl compounds vary in their capacity for biodegradation, but those that contain carboxylic acid groups are readily biodegraded. Bacteria capable of degrading cycloalkyl carboxylic acids are numerous in the environment (24).

Aromatic Hydrocarbons

The nature and extent of the biodegradation of aromatic hydrocarbons depends on the number of rings in the structure, the number of substitutions, the type and position of the substituted groups, and the nature of the atoms in heterocyclic compounds. The solubility of the aromatic hydrocarbon greatly affects its potential for biodegradation under either aerobic or anaerobic conditions. Mixtures of aromatic compounds can also influence the rate of biodegradation.

Aerobic microbial attack on single-, double-, and triple-ring aromatic compounds involves the foundation of a dihydrodiol compound. Oxidative attack on the dihydrodiol compound results in forming an alkyl catechol, a common intermediate formed during the oxidation of many aromatic hydrocarbons. Additional metabolic oxidation results in ring fission, forming either an aldehyde or an acid. This step results in the destruction of the aromatic ring leaving an oxidized aliphatic hydrocarbon, which is easily biodegraded, releasing hydrogen (25). Bacteria of the genus *Nocardia* can oxidize substituted aromatic hydrocarbons, such as *p*- and *m*-xylene, and use these compounds as a sole source of carbon and energy (4).

A second metabolic pathway for degrading aromatic hydrocarbons involves oxidation of alkyl substitutes, which results in forming aromatic carboxylic acids that are then oxidized to dihydroxylated ring fission products (i.e., aldehydes and acids). The aldehydes and acids are then readily biodegraded by the beta-ketoadipic and meta fission pathways (26). In general, alkyl-substituted aromatic hydrocarbons are less biodegradable, the longer the chain length, or the more numerous the alkyl groups.

Single-ring aromatic hydrocarbons can be transformed anaerobically by denitrifying, sulfate-reducing, iron-reducing, and methanogenic bacteria (27). These are all anaerobic respirations where the nitrate, sulfate, ferric, and carbonate ions act as terminal electron acceptors (TEA) for energy metabolism. The use of these compounds as electron acceptors in microbial energy metabolism is called dissimilative metabolism. During dissimilative metabolism, a comparatively large amount of the TEA is reduced, and the reduced product is excreted into the aquatic environment. The possible end products of these reductions are HS^-, N_2, $NO_2{}^-$, N_2O, Fe^{2+}, and CH_4.

The presence or absence of oxygen in the structure of an aromatic compound impacts both the degradation mechanism and the rate of biodegradation. The initial step in the degradation of aromatic hydrocarbons is conversion of the compound to an oxygenated form. Under anaerobic metabolic conditions, oxygen is incorporated from water into the aromatic structure by hydroxylation (28). The microbial fermentation of benzene and toluene is characterized by end products that are both partly oxidized and partly reduced. The oxidation might include both methyl group oxidation and ring oxidation. The reduction generally results in forming saturated cyclic rings (3).

Grbic-Galic (29) observed that the biodegradation of benzene is initiated by ring oxidation, resulting in the formation of phenol. Three pathways are possible for toluene, starting with ring oxidation to *p*-cresol or *o*-cresol and methyl group oxidation to benzyl alcohol. Thereafter, the biodegradations proceed along pathways that are similar to the anaerobic transformations of oxygenated aromatic compounds (30).

The biodegradation of multiring aromatic hydrocarbons (PAHs) is a function of the complexity of the chemical structure of the compound. In general, PAHs that contain four or more aromatic rings are much less biodegradable than compounds that contain only two to three rings (3). Several of the higher ring number PAHs and the intermediate products of their biodegradation are either toxic or carcinogenic. They are also strongly hydrophobic, which predicts that their concentration in the aqueous phase is always relatively low. Therefore, significant portions of the PAHs are found adsorbed on particles and possibly trapped in the micropores of these particles. Consequently, the rate of biodegradation is controlled by the sorption–desorption kinetics of the strongly sorbed PAH compounds. Enhancing solubilization by introducing bacteria that produce extracellular surfactants improves the *in situ* biodegradation of PAHs in groundwater.

Biodegradation of the unsubstituted di- and tri-ring PAHs in marine and freshwater is well documented in the literature (25,31,32). The bacterial degradation rates for phenanthrene and anthracene appear to be related to the water solubilities of these compounds. As for benzene and its derivatives, the oxidation of di- and tri-ring PAHs involves the formation of dihydrodiol intermediates. Catechol is the principal intermediate product of these microbial degradations (25). Bacteria that can degrade anthracene and phenanthrene include the genera *Aeromonas, Beijerinckia, Flavobacterium, Nocardia*, and *Pseudomonas* (31).

Biodegradation of unsubstituted PAHs that contain four or more aromatic rings (e.g., fluorene, pyrene, benzo(a)pyrene, benzo(a)anthracene, dibenzo(a)anthracene) has been documented in the literature (33–35). The bacteria that can degrade the higher molecular weight PAHs were also identified in this literature; they include the genera *Alcaligenes, Beijerinckia, Mycobacterium*, and

Pseudomonas. These research studies have indicated that cometabolism coupled with analog substrate enrichment may be necessary to treat PAHs that contain four or more aromatic rings. Cometabolism appears to be the principal mechanism for biodegrading benzo(a)anthracene (35). The presence of biphenyl, *m*-xylene, and salicylate were necessary to induce oxidation of benzo(a)anthracene to carbon dioxide and a mixture of *o*-hydroxy polyaromatic compounds. Because the high molecular weight PAHs do not induce enzyme production in many bacteria, the addition of naphthalene as an analog substrate was necessary to biodegrade benzo(a)anthracene in this study.

BIBLIOGRAPHY

1. Spain, J.C., Haigler, B.E., and Nishino, S. (1989). Development of bacteria for biodegradation of chloroaromatic compounds. In: *Biotechnology Applications in Hazardous Waste Treatment*. G. Lewandowski, P. Armenante, and B. Baltzis (Eds.). Engineering Foundation, New York, pp. 129–147.
2. Compeau, G.C., Mahaffey, W.D., and Patras, L. (1991). Full-scale bioremediation of contaminated soil and water. In: *Environmental Biotechnology for Waste Treatment*. G.S. Sayler, R. Fox, and J.W. Blackburn (Eds.). Plenum Press, New York, pp. 91–109.
3. Pitter, P. and Chudoba, J. (1990). *Biodegradability of Organic Substances in the Aquatic Environment*. CRC Press, Boca Raton, FL.
4. Gibson, D.T. (1989). Recent advances in the microbial degradation of aromatic hydrocarbons. In: *Biotechnology Applications in Hazardous Waste Treatment*. G. Lewandowski, P. Armenante, and B. Baltzis (Eds.). Engineering Foundation, New York, pp. 149–165.
5. Reineke, W. (1984). Microbial degradation of halogenated aromatic compounds. In: *Microbial Degradation of Organic Compounds*. D.T. Gibson (Ed.). Marcel Dekker, New York, pp. 319–360.
6. Young, L.Y. (1984). Anaerobic degradation of aromatic compounds. In: *Microbial Degradation of Organic Compounds*. D.T. Gibson (Ed.). Marcel Dekker, New York, pp. 487–523.
7. Dolfing, J. (1992). Gibbs free energy of formation of halogenated aromatic compounds and their potential role as electron acceptors in anaerobic environments. *Environ. Sci. Technol.* **26**: 2213–2218.
8. Gibson, S.A. and Suflita, J.M. (1990). Anaerobic biodegradation of 2,4,5-Trichlorophenoxyacetic acid in samples from a methanogenic aquifer: stimulated by short chain organic acids and alcohols. *Appl. Environ. Microbiol.* **56**: 1825–1832.
9. Haggblom, M.M., Rivera, M.D., and Young, L.Y. (1991). Anaerobic degradation of chloroaromatic compounds under different reducing conditions. *EPA Symposium on Bioremediation of Hazardous Wastes*. Falls Church, VA.
10. Semprini, L., Grbic-Galic, D., McCarty, P.L., and Roberts, P.V. (1992). Methodologies for evaluating in-situ bioremediation of chlorinated solvents. U.S. Environmental Protection Agency, EPA/600/R-92/042.
11. Vogel, T.N., Criddle, G.S., and McCarty, P.L. (1987). Transformation of halogenated aliphatic compounds. *Environ. Sci. Technol.* **21**(8): 722–736.
12. Rainwater, K. and Scholze, R.J. (1991). In-situ biodegradation for treatment of contaminated soil and groundwater. In: *Biological Processes-Innovative Hazardous Waste Treatment Technology Series*. Vol. 3, H.M. Freeman and P.R. Sferra (Eds.). Technomic, Lancaster, PA, pp. 107–122.
13. Vanelli, T., Logan, M., Arciero, D.M., and Hooper, A.B. (1990). Degradation of halogenated aliphatic compounds by the ammonia-oxidizing bacterium *nitrosomonas europaea*. *Appl. Environ. Microbiol.* **56**(4): 1169–1171.
14. Shields, M. (1991). Treatment of TCE and degradation products using *Pseudomonas cepacia*. *Symposium on Bioremediation of Hazardous Wastes: EPA's Biosystems Technology Development Program*, Abstracts, Falls Church, VA., April 16–18.
15. Hopkins, G.D., Semprini, L., and McCarty, P.L. (1993). Microcosm and in-situ field studies of enhanced biotransformation of trichloroethylene by phenol-utilizing microorganisms. *Appl. Environ. Microbiol.* **59**: 2277–2285.
16. Tsien, H.C., Bousseau, A., Hanson, R.S., and Wackett, L.P. (1989). Biodegradation of trichloroethylene by *Methylosinus trichosporium*. *Appl. Environ. Microbiol.* **55**: 3155–3161.
17. Sewell, G.W., Gibson, S.A., and Russell, H.H. (1990). Anaerobic in-situ treatment of chlorinated ethenes. *Workshop on In-Situ Bioremediation of Groundwater and Contaminated Soils*. Water Pollution Control Federation Hazardous Wastes Committee, WPCF, Annual Conference, October 7–11, Washington, DC, pp. 68–79.
18. Bouwer, E.J. and Wright, J.P. (1988). Transformations of trace halogenated aliphatics in anoxic biofilm columns. *J. Contam. Hydrol.* **2**: 155–169.
19. Baek, N.H. and Jaffe, P.R. (1988). Anaerobic mineralization of trichloroethylene, presented at the *International Conference on Physicochemical and Biological Detoxification of Hazardous Wastes*. Vol. II. X.C. Wu (Eds.). Technomic, Lancaster, PA, pp. 772–782.
20. Reinhard, M., Barbash, J.E., and Kunzle, J.M. (1988). Abiotic dehalogenation reactions of haloaliphatic compounds in aqueous solution, presented at the *International Conference on Physicochemical and Biological Detoxification of Hazardous Wastes*, Vol. II, X.C. Wu (Ed.). Technomic, Lancaster, PA, pp. 722–741.
21. Britton, L.N. (1984). Microbial degradation of aliphatic hydrocarbons. In: *Microbial Degradation of Organic Compounds*. D.T. Gibson (Ed.). Marcel Dekker, New York, pp. 89–130.
22. Bitton, G. and Gerba, C.P. (1984). *Groundwater Pollution Microbiology*. John Wiley & Sons, New York, pp. 23–24.
23. Wackett, L.P., Brusseau, G.A., Householder, S.R., and Hansen, R.S. (1989). Survey of microbial oxygenases: trichloroethylene degradation by propane-oxidizing bacteria. *Appl. Environ. Microbiol.* **55**: 2960–2964.
24. Trudgill, P.W. (1984). Microbial degradation of the alicyclic ring: structural relationships and metabolic pathways. In: *Microbial Degradation of Organic Compounds*. D.T. Gibson (Ed.). Marcel Dekker, New York, pp. 131–180.
25. Gibson, D.T. and Subramanian, V. (1984). Microbial degradation of aromatic hydrocarbons. In: *Microbial Degradation of Organic Compounds*. D.T. Gibson (Ed.). Marcel Dekker, New York, pp. 181–252.
26. Arvin, E., Jensen, B., Godsy, E.M., and Grbic-Galic, D. (1988). Microbial degradation of oil and creosote related aromatic compounds under aerobic and anaerobic conditions, presented at the *International Conference on Physicochemical and Biological Detoxification of Hazardous Wastes*. Vol. II, X.C. Wu (Ed.). Technomic, Lancaster, PA, pp. 828–847.
27. Huling, S.G., Pivetz, B., and Stransky, R. (2002). Terminal electron acceptor mass balance: light nonaqueous phase liquids and natural attenuation. *J. Environ. Eng.* **128**: 246–252.

28. Vogel, T.N. and Grbic-Galic, D. (1986). Incorporation of oxygen from water into toluene and benzene during anaerobic fermentative transformation. *Appl. Environ. Microbiol.* **52**: 200–202.
29. Grbic-Galic, D. (1990). *Microbial Degradation of Homocyclic and Heterocyclic Aromatic Hydrocarbons under Anaerobic Conditions*. Department of Civil Engineering, Environmental Engineering and Science, Stanford University, Stanford, CA.
30. Grbic-Galic, D. and Vogel, T.M. (1987). Transformation of toluene and benzene by mixed methanogenic cultures. *Appl. Environ. Microbiol.* **53**: 254–260.
31. Hogan, J.A., Toffoli, G.R., Miller, F.C., Hunter, J.V., and Finstein, M.S. (1988). Composting physical model demonstration: mass balance of hydrocarbons and PCBs, presented at the *International Conference on Physicochemical and Biological Detoxification of Hazardous Wastes*. Vol. II, X.C. Wu (Ed.). Technomic, Lancaster, PA, pp. 742–758.
32. Mahaffey, W.R., Compeau, G., Nelson, M., and Kinsella, J. (1990). *Development of Strategies for Bioremediation of PAHs and TCE*. *In-Situ Bioremediation of Groundwater and Contaminated Soils*. WPCF Annual Conference, Washington, DC, pp. 3–47.
33. Environmental Protection Agency. (1993). *Pilot-Scale Demonstration of a Slurry-Phase Biological Reactor for Creosote-Contaminated Soil—Applications Analysis Report*. EPA/540/A5-91/009.
34. Grifoll, M., Casellas, M., Bayona, J.M., and Solanas, A.M. (1992). Isolation and characterization of a fluorene-degrading bacterium: identification of ring oxidation and ring fission products. *Appl. Environ. Microbiol.* **58**: 2910–2917.
35. Mahaffey, W.R., Gibson, D.T., and Cerniglia, C.E. (1988). Bacterial oxidation of chemical carcinogens: formation of polycyclic aromatic acids from benzo(a)anthracene. *Appl. Environ. Microbiol.* **54**: 2415–2423.

LANDFILL

Jim Philip
Napier University
Edinburgh, Scotland, United Kingdom

INTRODUCTION

All countries rely to a greater or lesser degree on landfilling to dispose of the huge quantities of municipal solid waste (MSW) generated. For example, the United Kingdom has traditionally relied very heavily on landfilling of MSW as it has a relatively poor recycling infrastructure compared with some other European countries. The United Kingdom produces roughly 28 million tons of MSW per annum, of which about 80% is landfilled. Only some 12% is recycled and about 8% incinerated with energy recovery. Moreover, the amount of waste being generated is growing some 3–4% per annum.

Globally, landfilling has had a checkered history, and poor practice in design and operation of landfills has led to serious environmental problems. With regard to the threats to water bodies, the production of landfill leachate is by far the most significant. Liquid leachate develops at a site when its water holding capacity is exceeded. If the site is unlined, the leachate makes its way off-site to groundwater or to a surface water body and can cause drastic water pollution.

The Landfill Directive of the EU (1) has the overall aim "to prevent or reduce as far as possible negative effects on the environment, in particular the *pollution of surface water, groundwater* (authors' italics), soil and air, and on the global environment, including the greenhouse effect, as well as any resulting risk to human health, from the landfilling of waste, during the whole life-cycle of the landfill." This succinctly crystallizes the objectives of modern landfill design, and the Directive has far-reaching consequences for the way we handle and dispose of MSW. There are various stringent requirements of the Directive, the most important of which for preventing water pollution are:

- Higher engineering and operating standards are to be followed.
- Biodegradable waste has to be progressively diverted away from landfill. By 2020, the amount going to landfill will be 35% of that of 1995.
- Banning of disposal of liquid wastes to landfill sites, along with certain hazardous and other wastes, will be implemented.

THE MICROBIOLOGY OF REFUSE DECOMPOSITION

One view of a landfill is as an enormous, solid-state fermenter in which naturally selecting microbial populations, usually bacteria, anaerobically decompose refuse components, ultimately to their mineral constituents. The process of anaerobic decomposition is microbiologically complicated, and a great many details remain to be elucidated. However, the overall process can be summarized as in Fig. 1. Although this image seems to separate the various processes, it should be kept in mind that all happen contemporaneously. As the landfill is microbiologically active over long periods, often decades, only the anaerobic processes (those in blue in Fig. 1) concern us here: hydrolysis and fermentation, acetogenesis, and methanogenesis. One of the reasons that these processes occur over decades is that

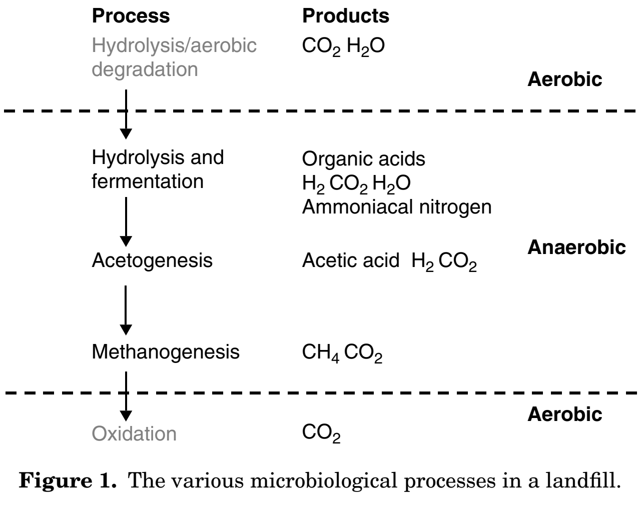

Figure 1. The various microbiological processes in a landfill.

the energy available from anaerobic processes is much lower than that from aerobic processes. For example, the aerobic mineralization of glucose to CO_2 and H_2O liberates more than seven times as much energy as the anaerobic mineralization of glucose to CO_2 and methane (CH_4).

Hydrolysis and Fermentation

Long-chain, insoluble carbohydrates, lipids, and proteins are not in a form in which they are readily metabolized by microorganisms. Hydrolytic reactions break down these long-chain molecules to smaller, more water-soluble molecules that can then be metabolized. Fermentation reactions produce soluble fermentation end products such as short-chain volatile fatty acids (VFA) and gaseous products of variable water solubility such as CO_2 and hydrogen.

Acetogenesis

The soluble acids from fermentation are converted to acetic acid, CO_2, and hydrogen by acetogenic bacteria. Other bacteria, acidogenic bacteria, convert carbohydrates, CO_2, and hydrogen to acetic acid. An important point is that the conversion of fermentation products to acetic acid occurs only at low concentrations of hydrogen. Hydrogen is produced at several stages, so it must be removed to prevent the inhibition of acidogenesis. If the partial pressure of hydrogen is too high, reduced organic acids, such as propionic, lactic, butyric, valeric, and caproic, start to accumulate, producing smells. Either from lowering the pH, mobilizing toxic metals, or inhibition due to the presence of high concentrations of these acids, there would be subsequent inhibition of methanogenesis.

Methanogenesis

Hydrogen concentration is kept low as a result of consumption by strictly anaerobic sulfate-reducing bacteria and methanogenic (methane-producing) bacteria. The methanogenic pathways of all species have the conversion of a methyl group to methane in common; however, the origin of the methyl group is variable. Although most isolated species can reduce CO_2, the majority of biological methanogenesis (about 70%) originates from conversion of the methyl group of acetate to methane, although in most cases acetate is not used as an energy source. Others acquire the methyl group directly from substrates such as methanol or methylamines.

Syntrophism

The low-energy yields involved in the anaerobic conversion of refuse to methane forces these different organisms into very efficient cooperation. Such cooperations are known as syntrophic relations. Syntrophism is a special case of symbiosis between two metabolically different types of bacteria, which depend on each other for degradation of a certain substrate, usually for reasons of energetics.

There is a classic example in landfill microbiology. Metabolism of low molecular weight fatty acids, such as propionate by *Syntrophobacter* and caproate and valerate by *Syntrophomonas*, produce acetate, hydrogen, and carbon dioxide, which are used by methanogens. The result is the overall conversion of fatty acids to methane. The conversion of fatty acids to acetate is energetically unfavorable, and the reaction depends on the removal of hydrogen by methanogens. Likewise, the methanogens need hydrogen for their metabolism.

Landfill Microbiology Is Mass Transfer Limited

Very often mass balance calculations of methane generation at landfills suggest that the overall process is not working nearly as efficiently as it might, even considering that several steps along the way have poor thermodynamics, which leads to very protracted timescales for return of the site to stability, and the consequent need is for long-term monitoring. If the bottleneck is not in thermodynamics, then where is it? The answer lies in a very common observation; that newspaper, although very rich in calories, can still be read after excavation from a landfill site decades after its disposal there.

Cellulose, the most abundant polymer on the planet, represents about 50% of the organic material going to landfills. The rate limitation in a landfill is much more likely to be associated with solid substrates, such as paper. Here the limitation is not thermodynamic, but mass transfer. The substrate has to be converted from the solid form into an aqueous form before other metabolic associations can continue the degradation. The enzymatic hydrolysis of polymers to monomers has long been known as the rate-limiting step in the conversion of cellulose to methane and in the digestion of refuse.

ENVIRONMENTAL IMPACTS OF LANDFILLS

Landfilling domestic refuse creates a whole host of negative environmental impacts, summarized in Fig. 2. These impacts have variable magnitudes and can be categorized as *localized* (odor, noise, litter, transport risks, public health risks created by birds and vermin, explosive gas migration), *diffuse* (groundwater, surface water, drinking water contamination), and *global* (greenhouse gas generation). In the present context, the effects on water pollution are the primary concern. And the water pollution problems are a direct result of leachate generation and its off-site migration.

Leachate Composition

Leachate is the result of water infiltration to the site that exceeds the water holding capacity of the waste and other site materials. The water balance of a landfill site can be summarized as

$$\mathrm{LC} = \mathrm{PR} + \mathrm{SRT} - \mathrm{SRO} - \mathrm{EP} - \mathrm{ST}$$

where LC = leachate
PR = precipitation
SRO = surface run-off
SRT = surface run-to (zero on a well designed site)
EP = evapotranspiration
ST = change in water storage

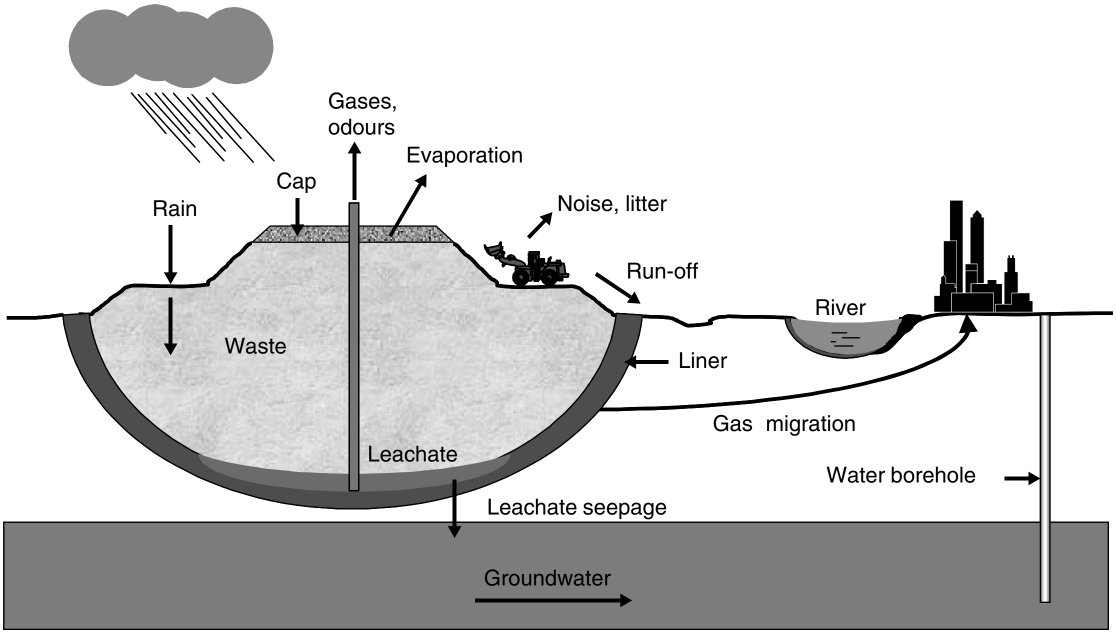

Figure 2. Environmental impacts of landfill practice.

Leachates vary in composition from site to site and also according to the age of the leachate. Leachates generated during the early stages of anaerobic decomposition are characterized by high concentrations of VFA, acidic pH, high BOD:COD ratio, and high levels of ammoniacal nitrogen and organic nitrogen. Ammonia is largely produced from the degradation of proteins. The low redox potential of such leachates facilitates the production of soluble, reduced-state metals, including chromium, iron, and manganese. These "young" leachates are much more environmentally damaging than mature leachates produced during the stable methanogenic phase.

By the time that methanogenesis is occurring at a high rate, many of the fatty acids have been converted ultimately to methane and CO_2. Methanogenic leachates are more likely to have a higher pH (resulting in lower heavy metal concentrations due to precipitation), lower levels of ammoniacal nitrogen, and a lower BOD:COD ratio (from the biodegradation of fatty acids).

LANDFILL DESIGN AND CONSTRUCTION IN RELATION TO ENVIRONMENTAL MITIGATION

It is not an objective to discuss detailed engineering design. Rather, in modern landfills, several design and construction measures can be taken to minimize the environmental effects of the site; and it is these that are summarized here. For example, a critical element of any landfill design is capacity, which is influenced by factors such as waste density, amount of daily cover, and the thickness of the final cap. However, for this discourse, capacity has little relevance to environmental effects.

Phasing and Cell Construction

A phase is a subsection of the total landfill to be filled; generally, it has an operational period of 12 to 18 months. Longer than 18 months results in excessive leachate production in wet areas. Cells are subsections of phases, which vary in size according to operational exigencies. Cell sizes are minimized according to the surface area required for maneuvering large machinery on the site. It may also be possible to size each cell in which the rate of vertical filling exceeds the rainfall plus water holding capacity to minimize leachate generation.

The practice of phasing (Fig. 3) has the objective of progressive excavation and filling of the site. As a result, at any one time, part of the site may be restored, part may be in the process of being capped, part is being prepared to receive waste, and only a relatively small part is being actively filled. When properly done, there will also be sufficient space for storage and protection of materials for subsequent restoration, and also coordination of haul roads and access routes. The environmental benefits of the phase/cell strategy are:

- reduction of leachate generation;
- progressive installation of leachate and gas control systems;
- segregation of clean surface water run-off within and outside the site;
- protection of local amenity.

Phases are generally filled from base to cap in a continuous operation, then capped and restored, leaving a temporary unrestored face sloping to the landfill base (Fig. 3). In deep landfills, such as those constructed in old quarries or opencast mining sites, the phases are vertically tiered so that, overall, the site is a three-dimensional honeycomb of cells.

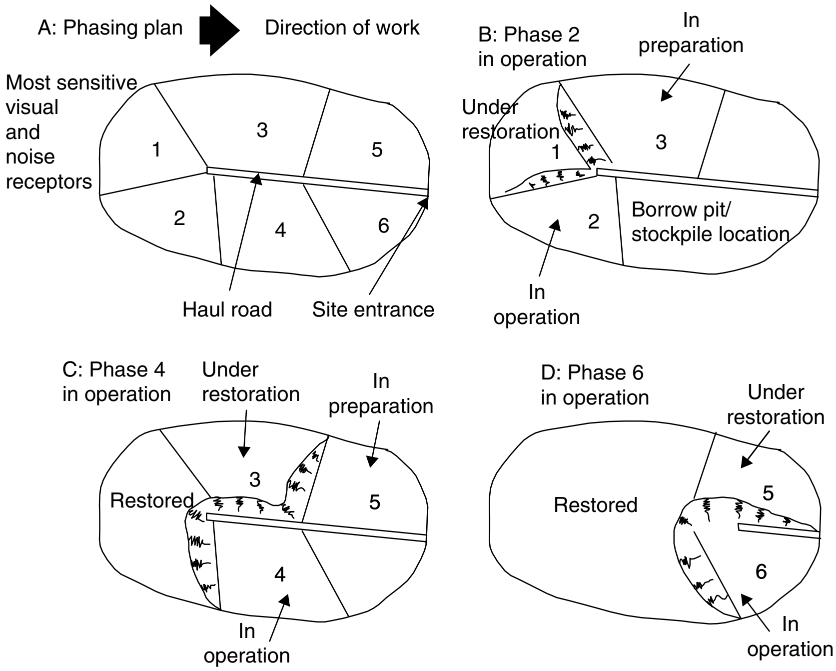

Figure 3. The phased approach to site management (adapted from Ref. 2).

Daily Cover

This practice involves covering refuse to a certain minimum depth daily with the following environmental benefits:

- prevention of wind-blown litter;
- suppression of odor;
- deterrence to vermin and birds;
- improvement of the site appearance.

From the point of view of water (and therefore leachate) and gas management, the material used for daily cover should be sufficiently permeable that it does not impede water or gas flow. Impermeable material creates perched conditions and makes it difficult to extract leachate. Ideally, the material of daily cover is soil excavated from within the boundaries of the landfill to prevent net consumption of void.

Liners

A key design parameter for MSW is to attain an impermeability of 10^{-9} m/s and so prevent leachate breakthrough to the unsaturated zone. The objective should be to select a new site on soil with a hydraulic conductivity lower than this. If this is not possible, then a variety of materials can be used to line the new site, either as single or multiple layers.

Clay liners varying in thickness from about 0.5 m (for imported clay) to 2 m (for *in situ* clay) are natural liners of high ion exchange capacity to retard the movement of toxic metals. Bentonite, the typical clay, is extremely absorbent. The hydraulic conductivity of dry, unconfined bentonite is 10^{-9} m/s. When saturated, however, it drops to less than 10^{-12} m/s^{-1}.

Synthetic liners, normally made of high-density (HDPE) or low-density (LDPE) polyethylene, are available in thicknesses of 0.5 to 2.0 mm. Although of very low permeability, their installation requires great care to prevent tearing.

Geosynthetic clay liners (GCL) (Fig. 4) are relatively new products gaining acceptance as barrier systems in municipal waste landfills. These offer some advantages over traditional bottom liners and covers and retain low hydraulic conductivity. Advantages include

- fast and easy to install;
- self healing of rips and tears due to the swelling property of bentonite;
- cost-effective in regions where clay is not readily available;
- their thin cross-section compared to a clay liner maximizes the capacity of a landfill and still protects groundwater.

Tests show that holes up to 75 mm diameter will self-heal when the clay hydrates and swells. Stitch-bonding or needle-punching creates small holes in the geotextile that heal due to swelling of the bentonite. The geotextile is often a blend of HDPE and very low-density polyethylene (VLDPE).

Drainage and Leachate Collection/Recirculation/Treatment

Leachate collected by the liner system must be removed to accelerate stabilization of the site and prevent liner damage, which is accomplished by drainage to collection sumps at low points via a granular layer containing perforated pipes of sufficient slope to allow gravity drainage. Leachate is then removed from the sumps, either by pumping in vertical wells or by gravity drains in a valley

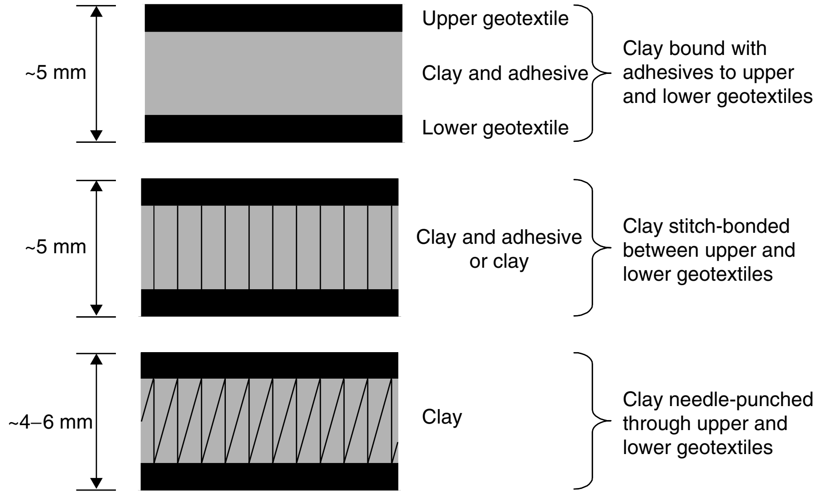

Figure 4. Geosynthetic landfill liner types.

site, where the leachate can either be treated on-site by a dedicated wastewater treatment plant or transported off-site for treatment.

Using a proper leachate collection system, it is possible to spray leachate back onto emplaced waste, which effectively uses the landfill site as a flushing anaerobic bioreactor and can improve landfill gas generation by uniformly wetting the waste.

Gas Abstraction and Use

This is mentioned here for completeness, although methane is highly insoluble in water and therefore contributes very little to the water-related environmental problems of landfills. During the stable methanogenic phase of a landfill, by far the longest phase, landfill gas, consists mainly of methane and carbon dioxide, both of which are greenhouse gases; the former is highly explosive; the latter is relatively water-soluble and corrosive. Most of the landfill gas management systems for landfills are designed with the characteristics of methane in mind. By appropriate siting of vertical or horizontal gas abstraction wells, it is possible to collect the gas, flare it, or, if economically viable, burn it for energy generation.

BIBLIOGRAPHY

1. Council Directive 1999/31/EC on the Landfill of Waste. (1999).
2. Department of the Environment. (1995). Landfill design, construction and operational practice. *Waste Management Paper 26B*. HMSO, London.

READING LIST

Kiely, G. (1996). *Environmental Engineering*. McGraw-Hill, London.

Kjeldsen, P., Barlaz, M.A., Rooker, A.P., Baun, A., Ledin, A., and Christensen, T.H. (2002). Present and long-term composition of MSW landfill leachate: A review. *Crit. Rev. Environ. Sci. Technol.* **32**: 297–236.

Qian, X., Koerner, R.M., and Gray, D.H. (2002). *Geotechnical Aspects of Landfill Design and Construction*. Pearson Education, Harlow.

Senior, E. (1995). *Microbiology of Landfill Sites*, 2nd Edn. Lewis, Boca Raton, FL.

LANDFILL LEACHATES, PART I: ORIGIN AND CHARACTERIZATION

A.I. Zouboulis
Aristotle University of Thessaloniki
Thessaloniki, Greece

Petros Samaras
Chemical Process Engineering Research Institute
Thermi-Thessaloniki, Greece

INTRODUCTION

Sanitary landfilling is the most widely used method for disposing of urban solid wastes around the world. The extensive use and the public awareness of this disposal method have raised concerns, over the negative environmental impacts and the pollution potential that this practice creates, as well as by the by-products of landfills (e.g., leachates, biogas, odors, etc.). Among them, leachates are considered the most important environmental burden. Depending on the composition and extent of decomposition of the disposal of refuse, as well as on the hydrological parameters existing in the landfill site, leachates may become highly contaminated wastewaters.

Landfill leachate, as defined in the U.S. Environmental Protection Agency Code of Federal Regulations (CFR) Title 40, Part 258.2, is the liquid that has passed through, or emerged from the disposal of solid wastes and contains soluble, suspended, or miscible materials from these wastes. Over time, the seepage of water through the landfill mainly from precipitation increases the mobility of pollutants and the potential for transferring them into the surrounding environment. As water passes through the layers of disposed solid wastes, it may "leach" pollutants from them, moving them deeper into the soil. The mobility of pollutants may present a potential hazard to public health, as well as to the environment, causing significant pollution problems in the groundwater aquifer and in adjacent surface waters. As a result, understanding and predicting leachate generation routes, as well as containing it and the subsequent appropriate treatment are required for environmentally proper handling of these heavily polluted wastewaters.

A simple measure to prevent the movement of toxic and hazardous waste constituents from a landfill is a liner operated in conjunction with a leachate collection system. Leachates are typically collected from a collection system placed at the bottom of the landfill. Leachates may also be collected by using slurry walls, trenches, or other containment systems. The leachate generated may vary from landfill site to site, based on a number of factors, which include the types of waste accepted for disposal, the operating practices (such as shredding, daily cover with soil, or capping), the depth of fill, the applied compaction of wastes, the annual precipitation at the landfill site, and the landfill operational age.

LEACHATE GENERATION

Leachates are the combined wastewater, containing organic and inorganic constituents, produced when water and/or other liquids seep through solid wastes, deposited in urban or hazardous solid waste landfills. The quantity of leachates is influenced by several interacting factors, such as annual precipitation, runoff, infiltration, evaporation, transpiration, mean ambient temperature, waste composition, waste density, initial moisture content, and underlying soil conditions (depth) (1). A number of techniques have been reported, using the water budget analysis through a landfill site, to estimate the amount of leachate generated (2). The various components of moisture used in the water budget are shown in Fig. 1. According to this analysis, the primary source of moisture is precipitation over the landfill site. A part of this moisture results in surface runoff, another part is returned to the atmosphere in the form of evapotranspiration from the soil and the surfaces of plants, and the remainder is added to soil moisture storage.

The maximum moisture that can be retained, without continuous downward percolation by gravity, is known as field capacity. Whenever the moisture content exceeds the field capacity of the soil, water percolates down into (through) the solid waste. The addition of moisture to solid waste over a period of time saturates the solid waste to its field capacity, resulting thereafter in leachate generation. The various moisture components, which constitute the processes taking place in a landfill that produce leachate, are affected by several parameters, such as (3):

1. Precipitation, which varies geographically and seasonally.
2. Surface runoff and infiltration, which depend on the intensity and duration of storms, surface slope, permeability of soil cover, and amount and type of vegetation.
3. Evapotranspiration at a landfill site, which is affected by the type of the soil and vegetation.
4. Soil moisture storage capacity, which is continually changing; it increases due to infiltration and decreases due to evapotranspiration.

Several methods used water balance calculations of these components to assess leachates generation rates. In general, following are the characteristics of the leachates produced (4):

1. A higher leachate generation rate is expected in humid than in dry areas.
2. Leachate generation follows a pattern similar to that of precipitation (rain), then remains at constant flow for a time period.
3. Production of leachates may be minimized by proper and efficient covering operations, careful drainage design, selection of vegetative cover, etc.
4. The quantity of leachate generated and its qualitative characteristics are significant for designing and constructing the most appropriate collection and treatment devices.

In general, leachates are generated over a long time period, unless percolation is prevented by the closure procedures and the final land use (5). In this case, leachate generation will cease shortly after the completion and closure of the landfill. However, for the proper design of

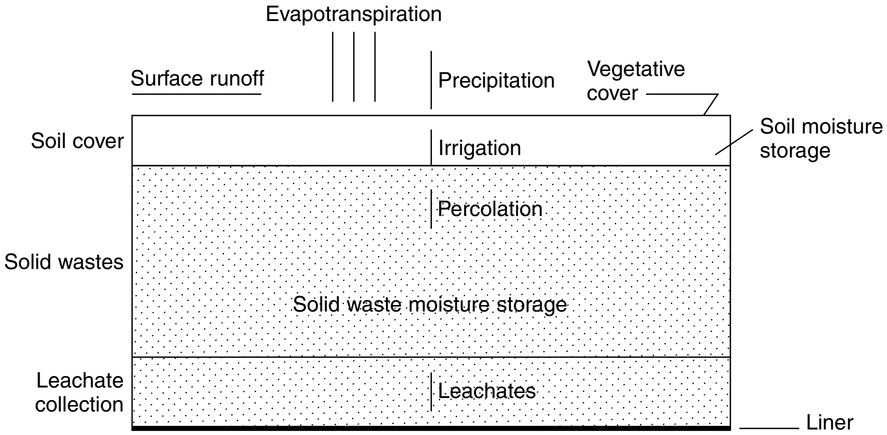

Figure 1. The various moisture components in a sanitary landfill.

a leachate treatment system, the characteristics of the leachate are necessary in addition to quantitative data.

LEACHATE CHARACTERISTICS

As leachates pass through or emerge from deposited solid wastes, they may contain soluble, suspended, or miscible materials from the wastes. Several factors may affect leachate quality, such as (6)

- specific types of solid waste accepted/deposited
- operating practices (shredding, cover, or capping)
- amount of infiltration
- depth of fill
- compaction
- age

The specific waste types received for disposal are the most representative characteristic of a landfill and, therefore, of the respective wastewater generated because the main contaminants in this wastewater are derived from the materials deposited in the site. The amount of infiltration and the age of a landfill are the primary factors that affect the concentration of contaminants in the leachate produced. The remaining factors influence mainly the rate of infiltration.

The highest concentrations of contaminants are typically present in leachates of new or very young landfills (7). However, the overall loads (i.e., the mass) of pollutants are generally not very large because new landfills typically generate low volumes of leachate. As the volume of waste approaches the capacity of the landfill and the production of leachate increases, both the pollutant loadings (i.e., flow × concentration) and the concentrations of certain contaminants, which are mainly organic pollutants, tend to increase. The increase of pollutant concentration is attributed to the onset of decomposition within the landfill and to the leachates that traverse the entire depth of the refuse. Therefore, large pollutant loadings from a typical landfill occur during the period of high leachate production and contain high levels of contaminants. The periods of varying leachate production cannot be quantified readily because they are site specific and depend on each of the aforementioned variables.

Over a period of time (as the landfill ages and leaching continues), the concentration of contaminants in the leachate decreases. The landfill may continue to generate substantial quantities of leachate; however, gradually the load of pollutants become lower because of the lower concentrations of soluble, suspended, or miscible contaminants that remain in the landfill. As the decomposition process within the landfill continues, the landfill attains a stabilized state of equilibrium, where further leaching produces leachates with a pollutant load lower, than during the period of peak leachate production. This stabilized state is presumably the result of decomposition of landfill waste by indigenous microorganisms that remove (biodegrade) many of the organic contaminants usually susceptible to further leaching.

Leachate characteristics change over time because there is a shift from the initial relatively short period of aerobic decomposition toward a longer period of anaerobic decomposition that has two distinct subphases (6, 8). The biological decomposition of landfilled municipal refuse is often based on the anaerobic breakdown of organic wastes. The biological activity occurs in a landfill shortly after the deposition of urban wastes, containing a large percentage of organic materials. Initially, the solid wastes, which contain high moisture content, can be decomposed rapidly under aerobic conditions, creating large amounts of heat. As oxygen is depleted, the intermediate anaerobic stage of decomposition begins. This change from aerobic to anaerobic conditions occurs unevenly through the landfill and depends upon the rate of oxygen diffusion into the fill layers.

In the first stage of anaerobic decomposition, the so-called "acidic phase," extra-cellular enzymes convert complex organic wastes, including carbohydrates, proteins, and fats, to more soluble organic molecules. Once the organic wastes are solubilized, their conversion to simpler organic molecules, such as acetic, propionic, butyric, isobutyric, valeric, isovaleric and hexanoic acids takes place; acetic acid is the main catabolic product of anaerobic fermentation. As a group, the low molecular weight, but highly polar, organic acids are termed volatile fatty acids (VFA). These soluble organic acids enter the leachate percolating through a landfill, resulting in a decreased pH of the leachate and increased oxygen demand. VFA impart to the leachate from this phase their characteristic "barnyard" odor and comprise the majority of its organic load. Anaerobic activity in the landfill can also lower the oxidation–reduction (redox) potential of the wastes, which under low pH conditions, can cause an increase in the concentration of dissolved inorganic contaminants.

Eventually, in the second or "methanogenic" phase of anaerobic decomposition, methane gas-forming bacteria within the landfill begin to convert the organic acids to methane and carbon dioxide, reducing the organic strength. The fraction of organic carbon, remaining after this degradation process, tends to be more oxidized, but has a higher molecular weight, higher than 500 amu. The absence of organic acids in the landfill increases the pH of the leachate, toward neutral or alkaline, which can subsequently decrease the solubility of inorganic contaminants and lower their concentrations in the resulting leachate.

The age or degree of decomposition of a landfill may be ascertained by observing the concentration of various leachate "gross" parameters, such as BOD_5, COD, TDS, or the organic nitrogen (N_{org}) concentration (9). The values of these leachate parameters can vary over the decomposition life of a landfill, depending on the specific phase. Typically, leachates from the early, acidic phase of anaerobic decomposition may be up to 35 times stronger than domestic wastewater and can have a COD content of more than 20,000 mg/L, BOD_5 greater than 12,000 mg/L, high volatile fatty acids concentration of about 6000 mg/L, and high content of inorganic compounds, such as 1300 mg/L chloride, but low phosphorous concentration (less than 1 mg/L) (10). Leachates from the older landfills have lower values of BOD_5 and COD, as well as of most organic pollutants, indicating the presence of smaller amounts of degradable compounds, derived from the aged stabilized waste.

The COD of leachates from the "methanogenic" phase tend to be lower, between 1500 and 4000 mg/L, and the significant decrease in the VFA concentration results in an increase in the pH to 7, or even higher. In addition, aged leachates can contain high levels of compounds existing in reduced form, such as ammonia (greater than 1000 mg/L), as well as a high concentration of chlorides because of the anaerobic environment of the landfill. Furthermore, certain metals such as iron, lead, and zinc tend to form stable complexes with the high molecular weight organic compounds (i.e., higher than 50,000), increasing their respective concentrations in the leachate (6).

However, using only these parameters, other refuse-filling variables, such as the processing of wastes prior to disposal or the fill depth, would not be taken into consideration. To compensate for these additional variables, several researchers have proposed examining certain ratios of leachate parameters over time (7). The most important (and widely used) such ratio is BOD_5:COD. Leachates from younger landfills typically exhibit BOD_5:COD ratios of approximately 0.8, whereas older landfills exhibit lower ratios (in many cases as low as 0.1). The decline in the BOD_5:COD ratio with age is due primarily to readily biodegradable material (e.g., phenols, alcohols, VFA) that degrade faster than the more recalcitrant compounds (such as the heavy molecular weight organic compounds, including humic and fulvic acids), which are much more difficult to treat biologically. As a result, as the landfill ages, the BOD_5 of the leachate decreases faster than the COD. Other ratios that decrease over time include the volatile solids to fixed (inorganic) solids (VS:FS), volatile fatty acids to total organic carbon (VFA:TOC), and sulfate to chloride (SO_4:Cl), which is inversely related to the redox potential (ORP).

As a result of the variation in leachate strength, leachates are commonly distinguished as young, acid-phase leachates and old, methanogenic ones, as well as medium- and low-strength leachates (11). For "young" leachates, a typical ratio of BOD_5/COD around 0.7 has been suggested; the corresponding values for mature leachates are 0.5, for aging 0.3, and for "old" 0.1. A typical time period for the transition from young to older leachate types is between 3 to 10 years from the landfill start-up, but may be as short as, 2 years in specific cases.

In summary, the following conclusions can be drawn regarding landfill leachate quality and their treatability (3):

- Leachate characteristics are highly variable.
- The quality of leachates changes with age, and therefore, the treatment facility should be flexible enough to handle/treat appropriately the changing leachate quality.
- A reliable estimate of the chemical quality of leachates should include analytical experimental data measured under the particular conditions prevailing in the landfill area.

The characteristics of leachates are usually very different from those of domestic wastewaters and similar to heavily loaded industrial wastewaters, indicating the need to use advanced treatment methods for the effective removal of pollutants before leachates are discharged.

BIBLIOGRAPHY

1. Leckie, J.O., Pacey, J.C., and Halvadakis, C. (1979). Landfill management with moisture control. *J. Environ. Eng. Div., ASCE* **110**(4): 780–796.
2. Tchobanoglous, G., Theisen, H., and Vigil, S. (1993). *Integrated Solid Waste Management: Engineering Principles Management Issues*. McGraw-Hill, New York.
3. Qasim, S.R. and Chiang, W. (1994). *Sanitary Landfill Leachate: Generation, Control and Treatment*. Technomics, Lancaster, PA.
4. Tatsi, A. and Zouboulis, A.I. (2002). A field investigation of the quantity and quality of leachate from a municipal solid waste landfill in a Mediterranean climate (Thessaloniki, Greece). *Adv. Environ. Res.* **6**(3): 207–219.
5. Barlaz, M.A. et al. (2002). A critical evaluation of factors required to terminate the post-closure monitoring period at solid waste landfills. *Environ. Sci. Technol.* **36**(16): 3457–3464.
6. Forgie, D.J.L. (1988). Selection of the most appropriate leachate treatment methods. Part 1. A review of potential biological leachate treatment methods. *Water Pollut. Res. J. Can.* **20**(2): 308–328.
7. Environmental Protection Agency. (2000). *Development Document for Final Effluent Limitations Guidelines and Standards for the Landfills Point Source Category*. EPA-821-R-99-019, Washington, DC.
8. Kjeldsen, P.K. et al. (2003). Present and long term composition of MSW landfill leachate—a review. *Crit. Rev. Environ. Sci. Technol.* **32**(4): 297–336.
9. Al-Yaqout, F. and Hamoda, M.F. (2003). Evaluation of landfill leachate in arid climate—a case study. *Environ. Int.* **29**(5): 593–600.
10. Harrington, D.W. and Maris, P.J. (1986). The treatment of leachate: A UK perspective. *Water Pollut. Control* **85**(1): 45–56.
11. Henry, J.G., Prasad, D., and Young, H. (1987). Removal of organics from leachates by anaerobic filter. *Water Res.* **21**(11): 1395–1399.

LANDFILL LEACHATES: PART 2: TREATMENT

PETROS SAMARAS
Chemical Process Engineering
Research Institute
Thermi-Thessaloniki, Greece

A.I. ZOUBOULIS
Aristotle University of Thessaloniki
Thessaloniki, Greece

INTRODUCTION

The characteristics (composition) of sanitary landfill leachates are very different from domestic wastewaters, and their quality varies from landfill to landfill, as well as with the particular landfill age. Hence, their treatment is based largely on industrial wastewater treatment processes. However, no single treatment method is considered efficient enough to achieve the high removal rates of pollutants, usually required; therefore, several treatment trains are currently used, including combinations of aerobic/anaerobic processes and several modes of physical–chemical treatment systems. The selection and design of leachate treatment facilities requires knowledge of several parameters, such as leachate quantity and quality, degree of necessary treatment, disposal methods, and effluents guidelines. These are the main problems that have to be considered:

- Specific treatment schemes applied in a particular landfill may not be transferable to other sites.
- Leachate quantity and quality vary seasonally, depending on climatic and hydrologic factors.
- The composition of disposed of solid wastes greatly affects the composition of leachates.

During plant design, the fluctuations in the leachate generation rate and its composition should be considered. Furthermore, the treatment system should also be flexible enough to treat the “young” leachates during the preliminary stages of landfill operation, as well as the “older” leachates produced during landfill aging.

During the early development of appropriate methods for leachates, the main efforts were focused on the application of treatment processes commonly used for municipal wastewaters; they were based mainly on biological processes. Physical and chemical systems used were later complementary, aimed at the development of more efficient overall techniques (1).

BIOLOGICAL PROCESSES FOR TREATING LEACHATES

Both aerobic and anaerobic biological units have been used to treat landfill leachates. The number of landfill facilities that use variations of biological treatment as part of landfill wastewater treatment systems in the United States has been reviewed by the EPA (2), and it is shown in Table 1. According to this table, most of the biological treatment systems use aerobic processes, including suspended growth processes (i.e., activated sludge, sequencing batch reactors, lagoons, etc.), as well as fixed-film processes (i.e., trickling filters or rotating biological contactors). However, the use of anaerobic systems is rather limited; these systems are most effective for treating high strength leachates (i.e., whose COD values are over 4000 mg/L) and for wastewaters containing refractory (not easily biodegradable) contaminants because of the effectiveness of methanotropic microorganisms in metabolizing these compounds. An important disadvantage of anaerobic treatment systems is the sensitivity of the applied methanotropic microorganisms to certain toxic substances, commonly found in many leachates.

The design of aerobic systems is based mainly on the requirements for removing organic loading in terms of BOD_5 and COD or on requirements for nitrogen removal. The selection of the design estimates depends on the effluents guidelines that have to be met and the problems anticipated from the existing high ammonia concentrations. A rule of thumb proposed for selecting the most appropriate design criteria is the ratio of BOD_5/N-content: when this ratio is less than 1, leachates are characterized by very high nitrogen (mainly ammonia) concentrations, and then nitrification criteria prevail; when this ratio is greater than 1, then the organic removal criteria dominate (3).

Table 1. Biological Treatment Facilities for Leachates in U.S. Landfills[a]

Type of Biological Treatment	% of Nonhazardous Facilities	% of Hazardous Facilities
Activated sludge	9	33
Aerobic lagoon systems	10	—
Facultative lagoons	7	—
Trickling filters	0	—
Anaerobic systems	2	—
Powdered activated carbon treatment (PACT)[b]	1	—
Nitrification systems	2	—
Rotating biological contactors (RBC)	0	—
Sequencing batch reactors (SBR)	1	33
Denitrification systems	1	—
Other biological treatment systems	13	—

[a] Reference 2.
[b] In combination with activated sludge.

Suspended growth biological treatment systems usually include mechanically based aerators to provide the required oxygen to the microbial population and for mixing the liquor components. The ranges of typical design and performance parameters of activated sludge systems are as follows (4):

- hydraulic retention time: 1–10 days.
- solids retention time between 1 and 10 days.
- food to microorganism (F/M) ratio from 0.02–0.4 kg BOD_5/kg MLVSS/day.
- average nutrient requirement ratio BOD_5: N:P = 100:3.2:0.5.
- Removal efficiencies in terms of BOD_5 and COD from 90–99%, depending on experimental conditions and the properties of raw leachates.

Several measurements of operational parameters have indicated that a large part of the organic compounds in raw leachates are usually not readily biodegradable and require prolonged reaction times and extensive biological activity to oxidize them. Efforts to determine the removal rates of various compounds in an aerobic reactor resulted in discriminating four distinct and successive steps of substrate use by microorganisms, in order of gradually increased degree of difficulty: carbohydrates, fatty acids, amino acids, and humic substances of high molecular weight (4). Residual organics consist mainly of fulvic-like materials with molecular weights in the range 500–10,000, which are not readily biodegradable.

Introducing an anoxic stage may enhance the removal of nitrogen from leachates, achieving more than 70% total nitrogen removal. However, the highest removal rates may be achieved by adding an organic carbon source, such as methanol. High metal removal rates (up to 99%) have also been observed in activated sludge treatment systems of leachates; this was attributed to the oxidation of metals, forming insoluble compounds, and the incorporation of the respective precipitates into bacterial flocs (2). However, when metal content is high (e.g., 80 mg/L iron and 10 mg/L manganese), a pretreatment step may be necessary for their efficient reduction to prevent the resulting low MLVSS/MLSS ratios and certain mixing problems.

Stabilization ponds and aerated lagoons have also been used as pretreatment steps for leachates, prior to disposal in municipal sewers or recycling into the landfill. Although extended aeration and lagoon-type systems are favored for treating leachates because of low manpower requirements and operational simplicity, however, this application has a primary drawback, the extended land required. Additional problems connected with suspended growth systems include the intense foaming of leachates, high power consumption, potential inhibition of biological activity by increased concentrations of metals, high sludge production rate, filamentous bulking of sludge, and decreased biological activity due to deficiency in certain nutrients (usually phosphorous) (5).

Fixed-film aerobic systems provide an appropriate substrate for attaching and growing aerobic and facultative (anaerobic) bacteria. However, their use is currently limited for treating leachates. Typical loading rates for such systems are detention time about 10 h and loading rates from 2 to 5 g of NH_3-N/m^2/day. The systems achieve high removal efficiencies and result in effluents whose ammonia content is lower than 1 mg/L and BOD_5 lower than 20 mg/L (2).

Anaerobic biological systems used for treating leachates are often based on fixed-film type reactors using inert media, as well as on suspended growth systems, such as the upflow anaerobic sludge blanket (UASB) reactor. Detention times reported for these systems range from 1–15 days for anaerobic filters and from 1–6 days for UASB units; for the latter systems, removal capacities greater than 80% have been observed, and for anaerobic filters, the corresponding values range from 70–99% COD removal efficiency. The determination of BOD_5/COD and COD/TOC ratios in effluents from anaerobic treatment units showed that the composition of these effluents compared with those of leachates created from landfills of intermediate age. It may be concluded that a substantial part of biodegradable organics can be removed by the landfill itself, acting as an anaerobic bioreactor, and thus, the subsequent application of biological treatment methods and in particular of anaerobic processes are only moderately effective for the removing the remaining organic matter from "older" leachates (1).

In general, aerobic biological treatment of leachates is possible for "young" leachates that have high BOD_5/COD ratios, higher than 0.4, as well as low ammonia content. Activated sludge systems are the most common treatment techniques, and fixed-film systems are best used for nitrification of "older" leachates rich in ammonia. In this case, phosphorous addition may be necessary to provide the proper nutrient balance for sufficient cell growth. Anaerobic treatment systems may also be successful for treating leachates because of the advantages of relatively simple design; low capital, operating, and maintenance costs; and the ability to treat leachates with high BOD_5/COD ratios. However, anaerobic systems also present certain drawbacks that limit their application: slow biomass establishment, requirement for higher (at least mesophilic) temperatures (i.e., difficult use in cold climates); and poor solids separation (2,5).

Combined treatment of leachates in an existing municipal wastewater treatment plant is considered convenient; it has been applied in several cases, and it is the preferred disposal method for leachates, when the principal following requirements are met: availability of a sewer system, wastewater treatment plant capacity high enough to accept the heavily loaded leachates, process compatibility with the specific (composition) characteristics of leachates, and a sludge treatment facility large enough to handle the increased sludge production rates (4). Several studies of the cotreatment of municipal wastewaters and landfill leachates concluded that the overall treatment process and the effluent quality are not seriously affected by the addition of leachates up to 10% by volume at the municipal sewage, although this depends mainly on the loading strength of the leachates (6,7).

However, several problems may arise during the cotreatment that are connected to the possible negative effect and accumulation of heavy metals, the conversion of ammonia, the variations in temperature, the (much higher) sludge production, the (usually intense) foaming, and (poor) solids settleability. As a result, the cotreatment of leachates and municipal wastewaters has to be studied case-by-case, considering the significant problems that may appear due to the presence of toxic compounds, as well as the specific portion of leachate that has to be cotreated. The introduction of a pretreatment facility, such as a simple aerated lagoon, may satisfy the requirements for a preliminary polishing step in these cases.

PHYSICAL–CHEMICAL TREATMENT OF LEACHATES

Physical–chemical techniques are becoming increasingly common for treating industrial wastewater and for reclaiming municipal wastewaters, especially when intended for reuse. These techniques include mainly processes such as equalization, neutralization/pH adjustment, chemical precipitation and coagulation, chemical oxidation, activated carbon adsorption, air-stripping, ion exchange, and membrane separation. The application of physicochemical treatment methods to leachates offers the advantages of short start-up periods, relative stability to temperature variations, and the potential for automation.

Equalization

The composition and generation rates of leachates at landfills may vary widely due to their direct relationship to rainfall, storm water run-on and runoff, groundwater entering the waste-containing zone, and the moisture content and absorptive capability of disposed of wastes. To allow equalization of pollutant loadings and flow rates, the leachates are often collected, prior to treatment, in tanks or ponds that have sufficient capacity to hold the peak flows generated at the landfill facility (2). A constant flow is delivered to the treatment system from these holding tanks to dampen the variation in hydraulic and pollutant loadings to the wastewater treatment system. This reduction in hydraulic and pollutant variability increases the performance and reliability of treatment systems applied downstream and can reduce the size of subsequent treatment tanks as well as the chemical or polymer feed rates of supplementary reagents by reducing the maximum flow rates and the concentrations of pollutants to be removed.

The equalization systems consist of steel or fiberglass holding tanks or lined ponds that can provide sufficient capacity to contain peak flow. Detention times determined by using a mass balance equation and depend on site-specific generation rates and treatment design criteria. Detention times can range from less than a day up to 90 days; the median value is about 2 days. Equalization systems usually contain either mechanical mixing or aeration systems; they enhance the equalization process by keeping the tank contents well mixed and prohibiting settling of solids.

pH Adjustment

The pH of wastewater generated by landfills may have a wide range of values, depending on the specific types of wastes deposited in the landfill. In many instances, the raw wastewater may require neutralization to eliminate either high or low pH that may upset the treatment system subsequently applied, such as an activated sludge biological treatment. The landfill facilities may use neutralization systems in conjunction with chemical treatment processes, such as chemical precipitation, to adjust the pH of the wastewater and to remove metals to optimize process control. Acids, such as sulfuric acid or hydrochloric acid, are added to reduce pH, whereas alkalis, such as sodium hydroxide or lime, are added to raise the pH. Neutralization may be performed in a holding tank, in a rapid mixing tank, or in an equalization tank. Typically, the neutralization systems applied at the end of a treatment system are designed to control the pH of the final discharge between 6 and 9.

Chemical Precipitation and Coagulation

Suspended particulates and colloidal matter contained in surface waters or wastewaters can be removed by coagulation using multivalent cations, such as Ca^{2+}, Fe^{3+} or Al^{3+}. As a result, several investigations of the treatment of leachates dealt with the use of chemical precipitation and coagulation (8).

In chemical precipitation, soluble metallic ions and certain anions, found in landfill wastewaters, are converted to insoluble forms that precipitate from the solution. Most metals are relatively insoluble as hydroxides, sulfides, or carbonates. Coagulation is used in conjunction with precipitation to facilitate their removal by agglomerating suspended and/or colloidal materials. The precipitated metals can be subsequently removed from the wastewater stream by filtration, settling clarification (sedimentation), or some other type of gravity-assisted separation. Other treatment processes such as equalization, chemical oxidation, or reduction (as in the case of hexavalent chromium) usually precede chemical precipitation. The performance of the chemical precipitation process is affected mainly by other chemical interactions, temperature, pH, the solubility of waste contaminants, and mixing effects (2).

Common precipitating reagents used at landfills usually include lime, sodium hydroxide, soda ash, sodium sulfide, or alum. Other chemicals also used in precipitation and coagulation, as well as for pH adjustment, include sulfuric and phosphoric acids, ferric chloride, and polyelectrolytes (synthetic organic polymers). Often, landfills use an appropriate combination of these chemicals. Precipitation by sodium hydroxide or lime is the most conventional method for removing metals from leachates. Hydroxide and coagulant precipitation has proven effective for removing several metals, such as trivalent chromium, pentavalent arsenic, copper, lead, nickel, and zinc. However, sulfide precipitation may also be used, instead of hydroxide precipitation, to remove certain metal ions, such as mercury, lead and silver more effectively.

Carbonate precipitation is another method of chemical precipitation; it is used primarily to remove antimony or lead. Use of alum as a precipitant/coagulant agent results in the formation of aluminum hydroxides in wastewaters, containing calcium or magnesium bicarbonate. Aluminum hydroxide is an insoluble gelatinous floc, which settles slowly and entraps suspended materials. It is considered particularly effective for removing certain metals, such as arsenic or cadmium.

Lime is less expensive than sodium hydroxide, so it is used more frequently at landfills employing hydroxide precipitation. However, lime is more difficult to handle and feed, as it must be slaked, prepared in a slurry, and mixed intensively, often plugging the feed system lines. Lime precipitation also produces a larger volume of sludge.

In addition to the type of chemical agent selected for treating leachates, another important design factor in the operation of chemical precipitation is the pH. Metal hydroxides are usually amphoteric, meaning they can react chemically both as acids or bases; as such, their solubilities increase at both lower (acidic) and higher (alkaline) pH levels. Therefore, there is an optimum pH value for the precipitation of each metal, which corresponds to its minimum solubility. Another key consideration in chemical precipitation is the necessary detention time during the sedimentation phase of the process. The optimal detention time depends on the wastewater being treated and on the desired effluent quality.

The first step in chemical precipitation is pH adjustment and the addition of coagulants. This process usually takes place in separate mixing and flocculation tanks. After mixing the wastewater with the appropriate chemical reagents, the resulting mixture agglomerates in the flocculation tank, and it is mixed slowly by mechanical means, such as mixers, or by recirculation pumping. The wastewater then undergoes further separation, by clarification (by settling) or filtration, where the precipitated metals are removed from the (cleared) solution. In a clarification system, an organic flocculant, such as a synthetic polymer (e.g., polyacrylamide), is sometimes added to help the settling. The resulting sludge from the clarifier or from the filter must be further treated, disposed of, or recycled.

Several studies have been reported on the examination of coagulation–flocculation for the treating landfill leachates, aimed to optimize performance, by selection of the most appropriate coagulant, determination of experimental conditions, assessment of the pH effect, and investigation of flocculant addition (9,10). Aluminum sulfate (alum), ferrous sulfate, ferric chloride, and ferric chlorosulfate are commonly used as coagulants (3). Iron salts proved more efficient than aluminum salts, resulting in COD reductions of up to 56%, whereas the corresponding values for alum or lime were 39 and 18%, respectively (11). Additionally, high COD removal capacities have been observed during the combined action of alum and lime on stabilized leachates (12). Furthermore, the addition of flocculants together with coagulants may substantially enhance the floc settling rate (9).

The coagulation–precipitation process has been investigated mainly by using stabilized or biologically pretreated landfill leachates, as a final polishing treatment stage. However, limited information exists on the efficiency of this physicochemical process, when used to remove pollutants from leachates, partially stabilized by recirculation or from recently produced ("fresh") leachate. This technique may be important for enhancing leachate biodegradability prior to biological treatment. High COD removal capacities (about 80%) have been obtained during the addition of ferric chloride to partially stabilized leachates, whereas low COD reductions (lower than 35%) have been measured during the addition of coagulants to raw samples (13).

In general, the coagulation and/or precipitation of raw leachates by the addition of lime resulted in the removal of multivalent cations, suspended solids, and color from raw leachates, but the effect on organic matter removal was rather negligible. In an early work of Slater et al. (14), it was found that only a small percentage (about 4–6%) of organic compounds with molecular weight of 10,000 or more was contained in an industrial raw leachate, whereas most of the organics had a molecular weight of 500 or less. However, after lime addition, the higher MW fraction disappeared, whereas the other fractions of MW 500 or less and between 500 and 10,000 remained almost untouched. As a result, lime addition to "young" leachates is not expected to be effective for removing organics because this type of leachate contains mainly high amounts of lower MW volatile fatty acids. Coagulation by lime may be an efficient method for treating "older" or biologically treated leachates that contain a large fraction of high MW substances, such as humic and fulvic acids (1). However, the use of ferric salts or of alum has proven more effective than lime, possibly due to different optimum pH process conditions that range between 4.5 and 5.5.

Chemical Oxidation

Chemical oxidation processes can generally be used in wastewater treatment to remove ammonia, to oxidize cyanide, to reduce the concentration of residual organics, and to reduce the bacterial and viral content. Chemical oxidation for treating leachates has been successful, based on several oxidants, including chlorine gas, calcium hypochlorite, potassium permanganate, hydrogen peroxide, and ozone (2). Both chlorine and ozone are two chemicals that are commonly used to destroy residual organics in biologically pretreated wastewater. When these chemicals are used for leachate treatment, the resulting disinfection of the wastewater is usually an added benefit.

Chemical oxidation is a potential treatment option for removing certain organic pollutants from leachates or groundwater. The amount of oxidant required in practice is generally greater than the theoretical mass calculated. The reasons for this are numerous and include incomplete oxidant consumption and oxidant demand caused by the simultaneous presence of other oxidizable species in solution. Oxidation reactions depend on the presence of appropriate catalysts, as well as on pH control, which is an important design variable. For many facilities using chemical oxidation, partial oxidation of organics, followed by additional treatment options, may be more efficient and cost-effective than using a complete oxidation treatment scheme alone.

The use of chlorine gas in leachates has been tested and resulted in high color and iron removal rates, but in limited reduction of organic matter, which is possibly due to (1) the presence of ammonia, which has to be initially destroyed by break-point chlorination before any organic oxidation and (2) the presence of relatively difficult to oxidize organics. On the other hand, the use of hydrogen peroxide presents several benefits: control of odor from stored leachates, removal of sulfides discharged to municipal sewers, and growth control of undesirable microorganisms near discharge of leachates. In addition, the study of the MW distribution of organics in leachates treated with hydrogen peroxide showed that it enhances the percentage of compounds of MW less than 1000 and therefore, increases the possibility of further biological treatment.

Ozone treatment of leachates provides several benefits, such as the removal of color, the degradation of particular organics (such as polyaromatic hydrocarbons), and the reduction of phenols and toxicity (15,16). Furthermore, ozone application may enhance the biodegradability of leachates by converting pollutants to end products or to intermediate products that are more readily biodegradable or can be more easily removed by adsorption (17). Ozone application proved very effective for removing color and iron, but was less efficient in removing COD. It was concluded that ozonation should not be used in leachates of high volatile fatty acid content, especially when acetic acid is present, due to their strong resistance to chemical oxidation (8).

Activated Carbon Adsorption

Activated carbon adsorption is a physical separation process, in which organic and inorganic materials are removed from wastewaters by sorption, attraction and accumulation of the contaminants on the surface of carbon granules. Most organic compounds and some metals typically found in landfill leachates can be effectively removed by using granular activated carbon (GAC). Although the primary removal mechanism is adsorption, biological degradation and filtration are additional pollutant removal mechanisms, also provided by an activated carbon filter. Adsorption capacities of 0.5 to 10% by weight are rather typical in many industrial applications. Spent carbon can be either regenerated on site by thermal processes, such as wet-air oxidation or steam stripping, or for smaller operations, it can be regenerated off site or sent directly for disposal in to hazardous waste landfills.

Several studies have been presented concerning the use of activated carbon adsorption for treating landfill leachates. In general, this process is very effective in removing residual organics that remain after prior biological treatment of leachates, and thus it could be used as a final polishing step for biologically pretreated leachates and/or for well-stabilized "old" leachates (8). Furthermore, the combination of powdered activated carbon with an activated sludge system results in enhanced removal capacities of organic matter up to 98% of BOD_5 (2).

In conclusion, the adsorption of higher MW organic compounds is enhanced by the properties of these compounds. As the MW of organics increases, their polarity, solubility, and branching properties decrease, resulting in an increase in carbon adsorption. As a result, lower MW volatile fatty acids representative of "younger" leachates are poorly adsorbed on activated carbon particles, whereas higher MW compounds, such as fulvic acids, found in "old" leachates, are adsorbed on activated carbon to a greater extent. Both powdered and granular activated carbon may be used for leachate treatment, but special consideration should be given during the design period, because of the high cost of this material.

Air Stripping

Stripping is an effective method for removing dissolved volatile organic compounds from wastewater. Removal is accomplished by passing air or steam through the agitated waste stream. Air stripping is used to the treat leachates, mainly to remove ammonia. In this case, the pH must be increased to between 10.8 to 11.5, usually by adding NaOH or $Ca(OH)_2$ solutions; ammonia stripping takes place by blowing large volumes of air upward through the leachate bulk volume. This process is carried out in towers, where leachates trickle down over some type of inert material or in a shallow aerated reaction vessel. Air stripping has been proved efficient for the extensive reduction of ammonia (up to 93%); the residual concentrations no longer inhibit nitrification (18). However, this method has two drawbacks; the cost of chemicals for pH adjustment and the problem of freezing as the air and the leachate temperature approach 0 °C (2).

Ion Exchange

Ion exchange is an adsorption process that uses appropriate (usually synthetic organic) resins as media to remove charged contaminants from wastewater. Ion exchange is commonly used to remove heavy metals from relatively low-concentration waste streams. A key advantage of the ion exchange process is that it allows recovery and reuse of the removed metals. Ion exchange can also be designed to be selective for certain metals and can effectively remove them from wastewater that contains high concentrations of "background" metals, such as iron, magnesium, and/or calcium. A specific disadvantage of this treatment method is that the resins used are subjected to fouling by oils or other natural (high MW) polymers. However, the use of ion exchange to treat leachates is limited, by the high operating costs of this method, including the cost of the exchange media, the necessary chemical reagents used as regenerants, and the regenerants disposal costs. As a result, this method is appropriate for the supplementary removal of metals, as posttreatment final polishing step, that results in low residual ion concentrations of less than 1 mg/L (2,8).

Membrane Filtration

Membrane filtration systems employ a semipermeable polymeric membrane and a pressure differential to separate constituents of different size (from microparticles down to soluble ions) from an aqueous phase. Nanofiltration, ultrafiltration, and reverse osmosis are the most commonly used membrane filtration processes.

Ultrafiltration uses a semipermeable microporous membrane, through which the wastewater is passed under pressure. Water and low molecular weight solutes, such as salts and surfactants, pass through the membrane and can be removed as permeate. Emulsified oils and suspended solids are rejected by the membrane and are removed with part of the wastewater as a more concentrated liquid. The concentrate is usually recirculated through the membrane unit, until the permeate flow drops substantially. Ultrafiltration is commonly used for removing substances whose molecular weights are greater than 500, including suspended solids, oil and grease, large organic molecules, and complexed heavy metals. Ultrafiltration is commonly used, when the solute molecules are greater than 10 times the size of the solvent molecules (usually water) and less than 0.5 μm.

Reverse osmosis is a separation process that uses selective semipermeable membranes to remove dissolved solids, such as metal salts, from water. The respective membranes are more permeable to water than to contaminants or impurities. The wastewater is forced through the membrane at a pressure that exceeds the osmotic pressure caused by the dissolved solids. Molecules of water pass through the membrane as permeate, and the contaminants are rejected along the surface of the membrane and exit as a concentrated stream. The concentrate (rejection) flow from a reverse osmosis system ranges from 10–50% of the feed flow; the concentrations of dissolved solids and of contaminants within this stream approach at least 10 times that of the feed (raw) wastewater. The percentage of permeate that passes through the membrane is a function of operating pressure, membrane type, and concentration of the contaminants in the feed.

Cellulose acetate, aromatic polyamide, and thin-film composites are commonly used membrane materials. Membrane pore sizes for a typical reverse osmosis system range from 0.0005–0.002 μm, and pressures of 20–30 bars are usually required. Therefore, reverse osmosis feed water must be very low in turbidity to avoid direct blocking of the membrane. As a result, pretreatment of landfill wastewater prior to reverse osmosis may be necessary, including chemical addition and clarification or cartridge filtration (using 5-μm filters), to remove suspended particulates from the influent and protect pumps and membranes. Carbon adsorption is also recommended as a pretreatment for membranes sensitive to chlorine. Biofouling can be prevented by chlorination and dechlorination of the feed water. To maintain sufficient solubility of metals, such as calcium, magnesium, and iron, and avoid the formation of precipitates, the pH should be appropriately adjusted with acid. Aside from pH adjustment, chemical reagents may also be used as bactericides, as well as for dechlorination (2).

Several reports on the use of ultrafiltration and especially on the application of reverse osmosis membranes for treating leachates showed that these methods are best used, following biological pretreatment, or for treating "older" leachates (19). In addition, during the treatment of leachates by membranes, compounds whose MW was lower than 200 were not rejected, in comparison with those having a MW higher than 200. The performance of reverse osmosis membranes may be further optimized, after careful adjustment of pH to values around 8 and by using polythylamine membranes that results in almost 94% reduction of total organic carbon (TOC) (8). However, severe membrane fouling was experienced in several cases, suggesting that reverse osmosis is most effective as a posttreatment step (following biological treatment) for removing residual COD and dissolved solids. Trebouet et al. (20) have shown that high pollutant removal can also be achieved by nanofiltration, especially for "older" leachates. Nanofiltration can be run at lower pressures than reverse osmosis, hence presenting lower operating costs and less membrane fouling.

Recirculation of Leachates into the Landfill

The recirculation of leachates is the redistribution of leachates that have been collected at the bottom of a landfill back to the top of it. The recirculation can usually be performed by spraying leachates onto the exposed surface of the landfill or by distribution through perforated pipes, located just beneath the surface of the landfill. As the recirculated leachate trickles downward through the fill, the disposed of solid waste materials in the landfill become an appropriate medium for developing anaerobic microorganisms, and as a result, an anaerobic treatment process is initiated.

Therefore, the landfill becomes an uncontrolled anaerobic digester and the biodegradable organics in the leachates are initially converted to volatile fatty acids and then to methane. Under these conditions, an initially low BOD_5 leachate, which is similar in composition to that of "older" leachates, may be produced during periods up to 18 months of recirculation. It has been suggested that the recirculation of leachates can more rapidly develop an active anaerobic bacterial population of methane-forming bacteria within the landfill. The rate of removal of organic compounds may be further enhanced by the addition of excess sewage sludge, acting as additional carbon source, which is produced in biological wastewater treatment plants, as well as by appropriate pH control (8,21).

The problems of recirculation usually include the development of odors, the high capital and maintenance cost of the recirculation system for leachates, and the precipitation of carbonates and iron oxides that may clog both the spraying equipment and the surface of the landfill and decrease the percolation rate. However, recirculation of leachates has a number of benefits, such as production of leachates with more uniform properties, the acceleration of overall landfill stabilization, the delay in the starting time for the application of other treatment systems, and reduction in the strength of the treated leachates. Nevertheless, the recirculation of leachates does not finally result in reduction of generated wastewater volumes and cannot provide a sufficient treatment process for leachates, because the treated leachates may have relatively high COD content (higher than 3000 mg/L), as well as high ammonia concentrations (2).

In general, physical–chemical treatment methods are considered an effective means for treating leachates, which contain organic compounds of MW less than 500 and a BOD/COD ratio lower than 0.1. This is particularly important for the use of chemical coagulation and activated carbon adsorption because these methods are very sensitive to the MW distribution of organics. Chemical oxidation may be used for removing dissolved metals (mainly iron), but has little effect on COD, when applied separately. Air stripping can be highly effective in removing ammonia, but at highly alkaline pH. Finally, membrane separation processes may have some potential for the treating leachates, but they are subject to membrane fouling.

In conclusion, neither physical–chemical treatment alone, nor biological methods may be able to treat leachates completely, when applied separately. As a result, the integration of several treatment processes is required to produce an effluent of acceptable quality; the selection of the most appropriate treatment train is very significant for an integrated leachate management system.

SELECTING THE APPROPRIATE COMBINATION OF TREATMENT PROCESSES

The characteristics of leachates may vary from place to place, as well as with time; thus, the construction of an appropriate treatment scheme is a difficult task, requiring the development/design of appropriate processes, consisting of the following subsequent steps: the first step involves assessment of leachate quality, an estimate of the (seasonal) quantity, and the type of treatment techniques available; the second step includes selection of the optimum biological treatment method for removing major pollutants from the leachate, followed by selection of the applicable polishing (final) stages as a third step (4,22).

Selecting Treatment Techniques Based on Leachate Quality

During this phase, the influent leachate quality is assessed, and the corresponding effluent characteristics are determined, based on the estimates of BOD_5, COD, nitrogen (as ammonia, i.e., NH_3-N), phosphorous, and metal content. Additional data that might be used in this phase include the concentration of volatile fatty acids, TOC, and total suspended solids.

The most important parameter for the preliminary screening of treatment techniques for leachates is the BOD_5/COD ratio. Leachates from a relatively new landfill have high COD values, usually higher than 10,000 mg/L; low NH_3-N content; and BOD_5/COD ratios ranging from 0.4–0.8. These leachates are representative of "young" leachates, containing high amounts of easily biodegradable organic substances, which are amenable to both aerobic and anaerobic biological treatment. However, physical–chemical treatment may not be an appropriate method for such leachates due to the presence of low molecular weight volatile fatty acids, prevailing over higher molecular weight compounds.

"Old" (mature) leachates are characterized by lower COD values, usually less than 3000 mg/L; higher NH_3-N concentrations, due to the anaerobic decomposition of organic nitrogen content within the landfill; and BOD_5/COD ratios lower than 0.4. In this case, aerobic biological treatment is required because it can provide extensive nitrogen removal through nitrification, in addition to removing organic matter (COD values). However, for BOD_5/COD ratios lower than 0.1, the remaining organics consist mainly of nonbiodegradable materials, and physical–chemical treatment becomes a more attractive option.

A second parameter, which can be used as an indicator for the preliminary selection of a treatment process, is the molecular weight partitioning of the organic content. In this case, biological treatment would be an effective technique for leachates that contain organic substances of MW lower than 500, whereas physical–chemical treatment is favored for removing organics, of molecular weights higher than 1000. However, this parameter is not easily measured in samples of leachates, and it is not used as a general index. Other less important parameters may include the BOD_5/NH_3-N ratio and the metal content, which are usually considered in the following steps of the design procedure.

Selecting the Appropriate Biological Treatment Method

To treat "young" leachates that have high BOD_5/COD ratios and high amounts of low MW (i.e., easily biodegradable) organics, biological treatment is the most efficient technique for reducing organic (carbonaceous) matter. The first option in this case is the examination of recirculation as a cost-effective method for preliminary reduction of leachate strength. However, when recirculation is not feasible, then aerobic or anaerobic biological treatment should be used.

Anaerobic biological treatment is the most appropriate method for handling leachates of high BOD_5/COD ratio, low NH_3-N content, high temperature, and high VFA content; the optimum process in this case is the application of anaerobic filters. The effluents from anaerobic processes would be similar to "old" leachates that have COD values between 1000 to 3000 mg/L and require an additional posttreatment step to reduce residual organics.

Aerobic biological treatment is a common method, which may be applied either for anaerobically pretreated or for raw leachates. In addition, biological denitrification may also be included in the aerobic process for effective removal of nitrates. Two significant problems are associated with the aerobic treatment of leachates; the first is the phosphorous deficiency in leachates that requires supplementary phosphorous, and the second is the high ammonia content that may inhibit the nitrification capacity of the system. In this case, the initial high ammonia concentration should be reduced by preliminary air stripping of the leachate. Additional problems include possible high metal content; long sludge retention time (up to 30 days), hence requiring the use of larger basins; foaming of leachates during aeration; and the potential for metal precipitate formation that affects the operation of aeration systems. The effluents from the aerobic processes usually have low BOD_5 values, lower than 100 mg/L, but rather high COD values, up to 1500 mg/L, that require subsequent physical–chemical treatment (22).

Selection of Physical–Chemical Treatment Method

The third step in the treatment design procedure for leachates is selecting a process for posttreating the effluent. "Old" and biologically pretreated leachates may contain nonbiodegradable high MW organic compounds, particulates, and metal ions. Residual organics may be removed by chemical oxidation, using hydrogen peroxide and/or ozone, but these techniques are expensive, due to the high dosages of the necessary chemicals. Coagulation is an alternative method, which is effective for removing higher MW organics that represent about 50% of the residual organic matter. Activated carbon adsorption is also a viable process for removing lower MW organics. Membrane separations may be used for removing organics, but these methods are best applicable as final polishing stages, due to high costs and problems of membrane fouling.

Chemical precipitation using lime or chemical oxidation followed by sedimentation and/or filtration has been suggested to decrease the metal content of leachates. Similar techniques can be also used to reduce particulates and metals content, such as coagulation followed by sedimentation/filtration, chemical oxidation, membrane separation, or ion exchange.

Several alternative units for treating leachates are operating currently worldwide, consisting of a combination of subsequently applicable treatment steps. The

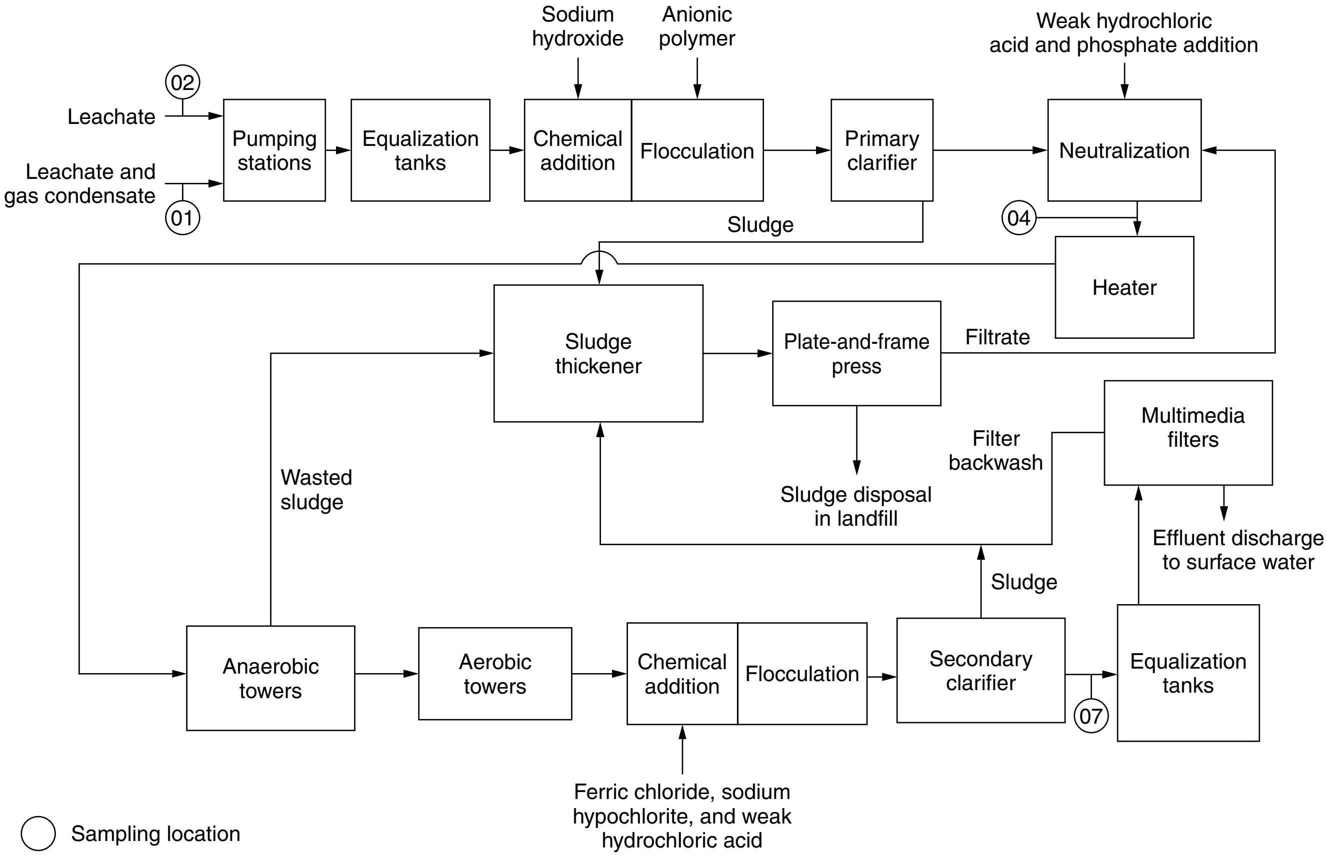

Figure 1. Flow sheet of a typical leachate treatment plant (2).

U.S. EPA (2) has evaluated the performance of a number of treatment systems and representative results are presented in the following. A typical treatment system for leachates from a sanitary landfill is shown in Fig. 1 (2). The system employs equalization tanks, coagulation/sedimentation, pH adjustment, biological treatment consisting of anaerobic towers followed by aerobic ones, coagulation/sedimentation, and multimedia filtration; the sludge treatment unit includes a sludge thickener and a plate-and-frame filter press. The results of the system operation are presented in Table 2.

As shown in Table 2, the biological treatment unit experienced good overall removal for TOC (93%), COD (90.9%), and NH_3-N (99.1%). The biological unit operation alone did not demonstrate high removals of BOD_5 (10.2%), TSS (9.3%), or for various metals (showing in general less

Table 2. Performance Data of a Typical Leachate Treatment Plant

	Biological Unit			Entire System		
Parameter	Influent, mg/L	Effluent, mg/L	Removal, %	Influent, mg/L	Effluent, mg/L	Removal, %
BOD_5	39.2	35.2	10.2	991	35.2	96.5
Total suspended solids	11.8	10.7	9.3	532.8	10.7	98.0
NH_3-N	135	1.1	99.1	193.3	1.1	99.4
COD	1742	159.4	90.9	4028	159.4	96.0
NO_3-N	1.5	130.5	0.0	0.693	130.5	0.0
TDS	5960	5181	13.1	5012	5181	0.0
TOC	758	52.8	93.0	1316	52.8	96.0
Total phenols	0.2	0.05	72.5	1.2	0.05	95.9
Barium	0.010	0.022	0.0	2.43	0.022	99.1
Boron	3	2.9	8.9	4.33	2.9	32.5
Chromium	0.012	ND[a]	—	0.036	ND[a]	70.3
Strontium	ND[a]	0.082	0.0	2.9	0.082	97.2
Zinc	ND[a]	0.012	0.0	0.144	0.012	91.6

[a] ND = not detected.

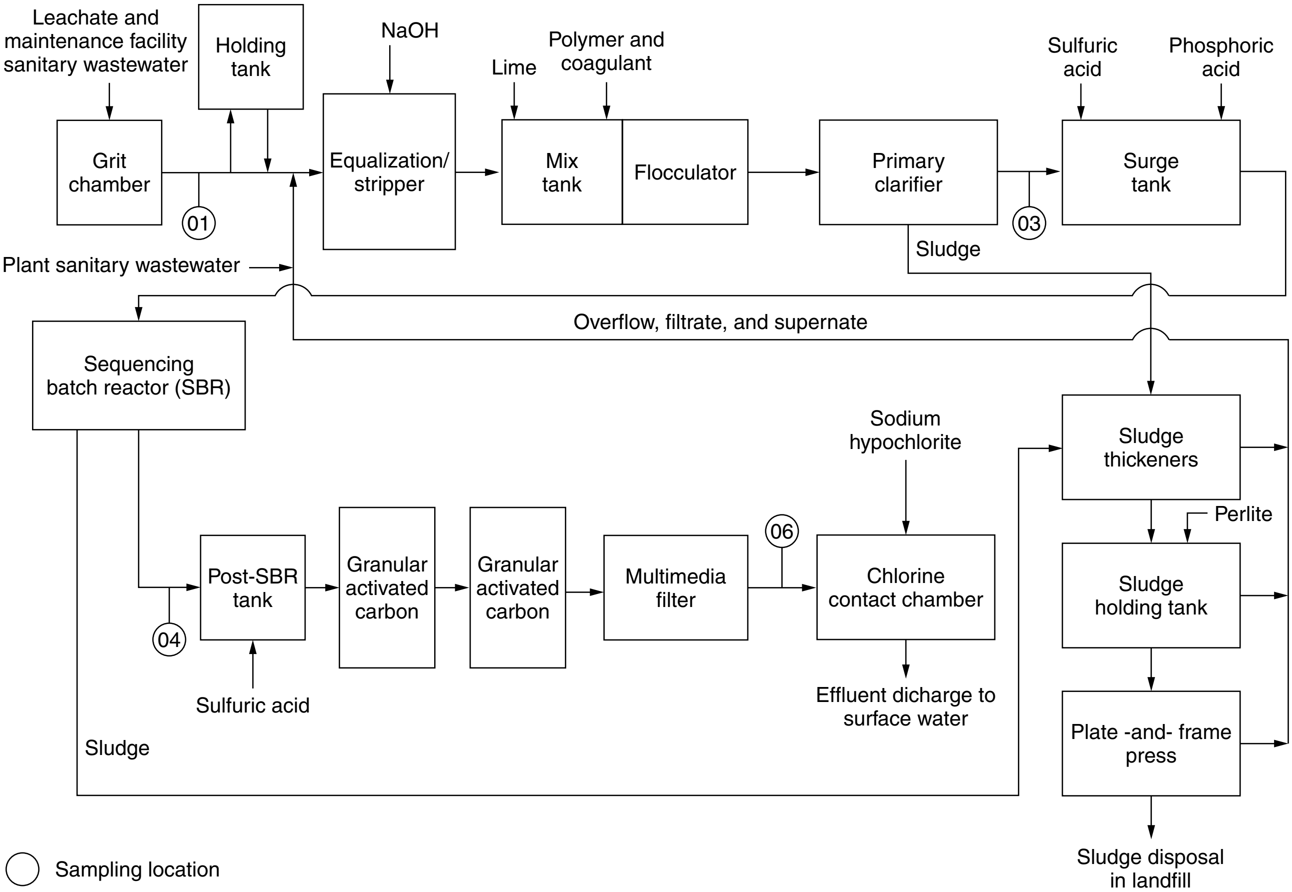

Figure 2. Flow diagram of an alternative leachate treatment plant using an SBR biological treatment unit.

than 10% removal) because the pollutants were generally not present in the biological treatment unit influent at treatable levels. The influent BOD_5 in the treatment unit was rather low (39.2 mg/L), TSS was 11.8 mg/L, and most metals were not at detectable levels, even though the raw wastewater at this facility exhibited an initial BOD_5 of 991 mg/L, TSS of 532.8 mg/L, and several metals at treatable levels.

The biological treatment unit influent had low concentrations of pollutants because this facility employed large aerated equalization tanks and a chemical precipitation system prior to biological treatment. The equalization tanks had a retention time of approximately 15 days and were followed by a chemical precipitation system using sodium hydroxide. Due to the long retention time and the aeration of wastewater, significant biological activity also occurred in these tanks. The resulting insoluble pollutants were removed in the primary clarifier prior to entering the biological towers. The entire treatment system showed good removals for BOD_5, TSS, NH_3-N, COD, TOC, and total phenols. Most metals had good percentage removals or were removed to nondetectable levels.

Another leachate treatment plant was evaluated by the EPA, including ammonia removal, hydroxide precipitation, biological treatment using a sequencing batch reactor, granular activated carbon adsorption, and multimedia filtration. A flow diagram of the landfill wastewater treatment system is presented in Fig. 2 (2). The wastewater treatment process used at this (nonhazardous) facility was primarily treated landfill generated wastewater and a small amount of sanitary wastewater flow from the on-site maintenance facility. A summary of percentage removal data collected for the biological treatment unit operation (SBR) and for the entire treatment system is presented in Table 3.

As shown in this table, the SBR treatment unit showed moderate overall removals for TOC (43.45%), COD (24.7%), and BOD_5 (48.7%). Improved removal efficiencies were observed for TSS (82.9%), total phenols (74.2%), and NH_3-N (80.7%). Metals, such as barium, chromium, and zinc had low removal efficiencies. However, these metals in the influent of the biological treatment system were measured at low concentrations, often close to the detection limit. Other metals also had poor removal efficiencies, including boron and silicon. The entire treatment system showed good removals for BOD_5, TSS, NH_3-N, COD, TOC, and total phenols. Each of the metal parameters also experienced good removal rates through the treatment system.

An alternative system for treating leachates from a nonhazardous facility was also constructed; it employed a two-stage reverse osmosis system and a multimedia filter. The flow diagram of this unit is shown in Fig. 3, and the corresponding performance data are given in Table 4.

Table 3. The Performance Data of an Alternative Leachate Treatment System that Employed an SBR Biological Treatment Unit

	Biological Unit			Entire System		
Parameter	Influent, mg/L	Effluent, mg/L	Removal, %	Influent, mg/L	Effluent, mg/L	Removal, %
BOD_5	232.6	119.3	48.7	1088	201	81.5
TSS	59.6	10.2	82.9	93.4	ND[a]	95.7
NH_3-N	134.8	26.04	80.7	295.9	12.06	95.9
COD	635.0	478.2	24.7	2932	251	91.4
NO_3-N	14.4	87.8	0.0	0.494	87	0.0
TDS	4024	3987	0.9	6232	3834	38.5
TOC	212.6	120.4	43.4	1098	82	92.5
Total Phenols	0.2	0.052	74.2	0.940	ND[a]	94.7
Barium	0.019	0.032	0.0	0.283	0.0426	85.0
Boron	2.842	2.483	12.6	6.7	2.334	65.2
Chromium	0.010	0.011	0.0	0.0906	ND[a]	87.7
Strontium	0.193	0.237	0.0	1.935	0.249	87.1
Zinc	0.025	0.058	0.0	0.494	0.027	94.5

[a] ND = not detected.

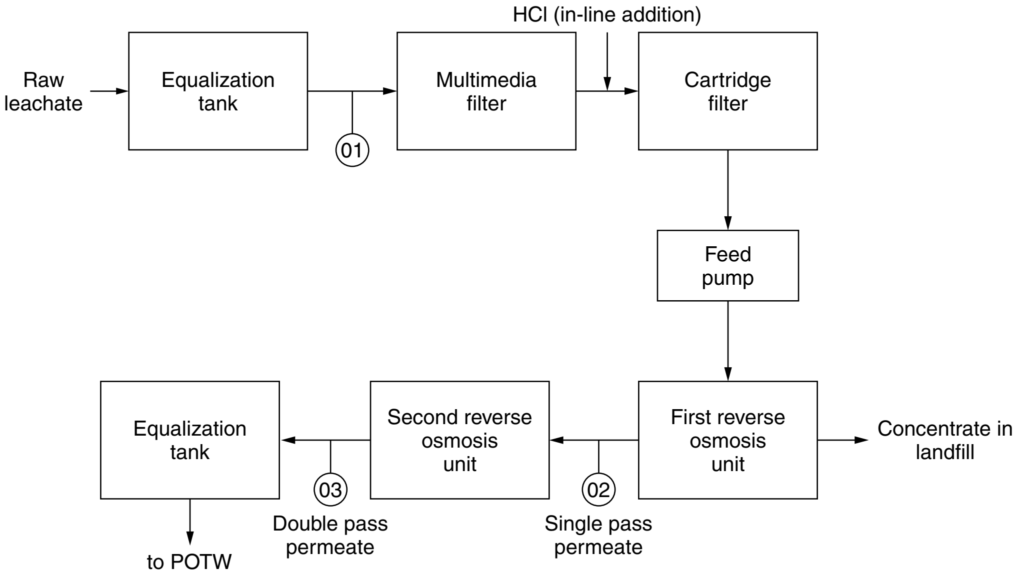

Figure 3. Flow diagram of a leachate treatment plant using a two-stage reverse osmosis unit.

As shown in Table 4, the single-pass reverse osmosis treatment system demonstrated good overall removals for a number of parameters, including TSS, TOC, BOD_5, TDS, and COD. Total phenol and NH_3-N% removals were observed at 75.1 and 76.7%, respectively. Metals with quantitative percentage removals included arsenic (87.4%), boron (54.1%), and strontium (92.9%). The additional polishing reverse osmosis unit caused the removal efficiency of most parameters to increase further. These parameters include BOD_5, NH_3-N, COD, TDS, TOC, and total phenols. The percentage removal for boron also increased from 54.1% in the single-pass reverse osmosis system up to 94.4% in the two-stage reverse osmosis treatment system.

In general, selecting the most appropriate treatment train should be based on leachate characteristics and cost estimates. For "young" leachates, which have BOD_5/COD ratios higher than 0.4, biological treatment will prevail in the overall treatment system. For the highest BOD_5/COD ratios (in the range of 0.6 to 0.8), recirculation of leachates should be included as a preliminary treatment stage, followed by anaerobic or aerobic treatment. For BOD_5/COD ratios lower than 0.4, aerobic biological treatment becomes the most important method, especially when nitrification is required; in all cases, physical–chemical treatment is necessary, as a polishing step for biologically pretreated effluents.

In some cases, physical–chemical treatment may precede aerobic treatment, aiming, for example, to reduce ammonia or metals. When the BOD_5/COD ratio is lower than 0.1, physical–chemical methods are the most appropriate, and aerobic biological treatment may be additionally used for nitrification/denitrification. During the design period, it is very important to consider the variation of

Table 4. Performance Data of a Leachate Treatment Unit, Including a Two-Stage Reverse Osmosis System

	Single-stage Reverse Osmosis[a] System			Entire Treatment System[a]		
Parameter	Influent, mg/L	Effluent, mg/L	Removal, %	Influent, mg/L	Effluent, mg/L	Removal, %
BOD_5	1182	54	95.4	1182	5.4	99.5
TSS	171.8	ND	97.7	171.8	ND	97.7
NH_3-N	58.48	13.6	76.7	58.48	0.608	99.0
COD	1526	72.2	95.3	1526	11.4	99.3
NO_3-N	1.3	0.666	48.8	1.3	0.502	61.4
TDS	2478	116.6	95.3	2478	ND	99.6
TOC	642.6	25	96.1	642.6	ND	98.4
Total phenols	1.26	0.316	75.0	1.26	0.063	95.0
Barium	0.28	0.006	98.0	0.28	0.001	99.5
Boron	1.808	0.83	54.1	1.808	0.10	94.4
Chromium	ND	ND	—	ND	ND	—
Strontium	1.406	ND	92.9	1.406	ND	92.9
Zinc	ND	ND	—	ND	ND	—

[a] ND = not detected.

leachates properties over time that requires particular attention to the treatment plant design to achieve operating flexibility to cope with varying influent characteristics.

BIBLIOGRAPHY

1. Lema, J.M., Mendez, R., and Blazquez, R. (1988). Characteristics of landfill leachates and alternatives for their treatment: A review. *Water Air Soil Pollut.* **40**: 223–250.
2. Environmental Protection Agency. (2000). Development document for final effluent limitations guidelines and standards for the landfills point source category. EPA-821-R-99-019, Washington, DC.
3. Ehrig, H.J. (1984). Treatment of sanitary landfill leachate: biological treatment. *Waste Manage. Res.* **2**: 131–152.
4. Qasim, S.R. and Chiang, W. (1994). *Sanitary Landfill Leachate: Generation, Control and Treatment*. Technomic, Lancaster, PA.
5. Forgie, D.J.L. (1988). Selection of the most appropriate leachate treatment methods. Part 1. A review of potential biological leachate treatment methods. *Water Pollut. Res. J. Can.* **20**(2): 308–328.
6. Kelly, H.G. (1987). Pilot testing for combined treatment of leachate from a domestic waste landfill site. *J. Water Pollut. Control Fed.* **59**(5): 254–261.
7. Diamadopoulos, E., Samaras, P., Dabou, X., and Sakellaropoulos, G.P. (1997). Combined treatment of landfill leachate and domestic sewage in a sequencing batch reactor. *Water Sci. Technol.* **36**(2–3): 61–68.
8. Forgie, D.J.L. (1988). Selection of the most appropriate leachate treatment methods. Part 2. A review of recirculation, irrigation and potential physical-chemical treatment methods. *Water Pollut. Res. J. Can.* **20**(2): 329–340.
9. Amokrane, A., Comel, C., and Veron, J. (1997). Landfill leachates pre-treatment by coagulation flocculation. *Water Res.* **31**(11): 2775–2782.
10. Sletten, R.S., Benjamin, M.M., Horng, J.J., and Ferguson, J.F. (1995). Physical–chemical treatment of landfill leachate for metals removal. *Water. Res.* **29**(10): 2376–2386.
11. Diamadopoulos, E. (1994). Characterization and treatment of recirculation-stabilized leachate. *Water Res.* **28**(12): 2439–2445.
12. Loizidou, M., Kapetanios, E.G., and Papadopoulos, A. (1992). Assessment of leachate characteristics and its treatability. *Fresenius Environ. Bull.* **1**: 748–753.
13. Tatsi, A., Zouboulis, A.I., Matis, K.A., and Samaras, P. (2003). Coagulation-filtration pretreatment of sanitary landfill leachates. *Chemosphere* **53**(7): 737–744.
14. Slater, C.S., Uchrin, C.G., and Ahlert, R.C. (1985). Ultrafiltration processes for the characterization and separation of landfill leachates. *J. Environ. Sci. Health A* **20**(1): 97–111.
15. Wenzel, A., Gahr, A., and Niessner, R. (1999). TOC removal and degradation of pollutants in leachate using a thin-film photoreactor. *Water Res.* **33**(4): 937–946.
16. Marttinen, S.K. et al. (2002). Screening of physical–chemical methods for the removal of organic material, nitrogen and toxicity from low strength landfill leachates. *Chemosphere* **46**: 851–858.
17. Beaman, M.S., Lambert, S.D., Graham, N.J.D., and Anderson, R. (1998). Role of ozone and recirculation in the stabilization of landfills and leachates. *Ozone Sci. Eng.* **20**: 121–132.
18. Cheung, K.C., Chu, L.M., and Wong, M.H. (1997). Ammonia stripping as a pretreatment for landfill leachate. *Water Air Soil Pollut.* **94**: 209–221.
19. Peters, T. (1999). Past and future of membrane filtration for the purification of landfill leachate. *Proc. 7th Int. Waste Manage. Landfill Symp.*, Cagliari, Italy, pp. 335–344.
20. Trebouet, D. et al. (1999). Effect of operating conditions on the nanofiltration of landfill leachates: Pilot scale studies. *Environ. Technol.* **20**: 587–596.
21. Morris, J.W.F., Vasuki, N.C., Baker, J.A., and Pendleton, C.H. (2003). Findings from long-term monitoring studies at MSW landfill facilities with leachate recirculation. *Waste Manage.* **23**(7): 653–666.
22. Forgie, D.J.L. (1988). Selection of the most appropriate leachate treatment methods. Part 3. A decision model for the treatment train selection. *Water Pollut. Res. J. Can.* **20**(2): 341–355.

MACROPHYTES AS BIOMONITORS OF POLYCHLORINATED BIPHENYLS

Karina S.B. Miglioranza
Mirta L. Menone
Universidad Nacional de Mar del Plata
Mar del Plata, Argentina
and
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET),
Buenos Aires, Argentina

Julia E. Aizpún de Moreno
Víctor J. Moreno
Universidad Nacional de Mar del Plata,
Mar del Plata, Argentina

INTRODUCTION

The freshwater, estuarine, and marine ecosystems around the world are now being increasingly subjected to greater stress from various human activities. As a result of several contaminants, significant changes in freshwater and marine plant communities have occurred (1).

Polychlorinated biphenyls (PCBs) are one of the more ubiquitous and toxic organic pollutants. Altogether, some 209 different individual PCB compounds exist, although only about 130 of these are found in commercial mixtures. PCBs have a number of physical and chemical characteristics that have contributed to their persistence in the environment, such as low aqueous solubility, resistance to oxidation and hydrolysis, and high volatility (2). The use of PCBs is now restricted to nondispersive systems (e.g., as dielectric fluids in condensers and transformers), but until few years ago, they were important constituents of a number of industrial chemical formulations, including paints, inks, plastics, pesticides, and copying papers.

In most industrialized countries, the use of PCBs has been banned since the late 1970s, resulting in a decline in PCB releases in the environment. However, in several aquatic ecosystems, the biota contamination remained relatively stable since the mid-1980s (3–5).

The concern about the presence and effects of PCBs on the biota and, ultimately, on human health, have mobilized worldwide research in order to find remediating PCB levels in the environment and therefore prevent possible future undesirable consequences. Toxic effects in animals include reproductive impairment, mutagenesis, carcinogenesis, and teratogenesis (6). Recently, several studies have demonstrated the ability of PCBs to induce oxidative stress (a cellular situation characterized by an elevation in the steady-state concentration of reactive oxygen species) (7) in fishes, birds, reptiles (8), and dinoflagellates (9); but in plants, a lack of studies about this subject exist. PCBs are very hydrophobic contaminants and preferentially adsorb onto sediment particles. Particle size distribution and organic carbon content seem to be important factors in determining the extent of partitioning of PCBs to natural sediment (10). In particular, the high octanol/water partition coefficient (Kow) indicates that PCBs have a high affinity for suspended solids, especially those rich in organic carbon. Mobilization of PCBs from sediments is one of the causes of the persistence of these contaminants (11,12), because sediments were, in the past, the most important sink of these pollutants.

THE ROLE OF MACROPHYTES IN THE AQUATIC ECOSYSTEM

Macrophytes include various emergent, submergent, and floating leaved species. Together with algae, periphyton, and phytoplankton, macrophytes form the base of most aquatic food chains. Macrophytes are important in nutrient cycling and respond rapidly to water quality changes. It is well known that aquatic vascular plants can enhance a freshwater ecosystem in many ways, such as by improving water clarity, contributing oxygen via photosynthesis, and providing spawning sites and protection for fish, mollusks, crustaceans, and other invertebrate species (13,14).

Plants interact with their environment through processes that include contaminant bioconcentration and excretion, shading, and organic matter production and decomposition. As a result of these interactions, aquatic plants may significantly affect water and sediment quality (15).

As plants are the base of most food chains, they will experience the effects of toxic compounds released into the aquatic media sooner than will organisms occupying higher trophic levels. In this respect, plants may be able to act as an “early warning signal” of impending contaminant impacts on other trophic levels of aquatic environments.

Contaminants present in aquatic macrophytes may be transferred to higher levels of the food chain, when consumed as live plants by herbivores or as detritus by detritivores (16,17).

Aquatic macrophytes in the littoral zones of lakes have two fundamental properties, which make them attractive as limnological indicators. First, they react slowly and progressively to changes in nutrient conditions, in contrast to bacteria and microalgae, i.e., over several years. Macrophytes, therefore, function as integrators of environmental conditions to which they are subjected, and thus can be used as long-term indicators with high spatial resolution. Second, the littoral zone may experience patterns of nutrients and pollutant concentrations (18,19), caused by natural or artificial inflows as well as by diffuse and nonpoint sources. The latter are difficult to localize and quantify, but even chemical analysis of point sources is often neglected in limnological routine work (20).

Vascular plants like bulrushes (*Schoenoplectus* spp.), duckweeds (*Lemna* spp.), and pickerel weed (*Pontederia* spp.) are the more common species used in North American wetlands treatment systems to improve water quality in a variety of ways, including binding soil and reducing the resuspension of muds (15). Bioaccumulation by plants can remove substantial quantities of potential toxicants (like organochlorines) and nutrients of water entering and passing through wetlands. Thus, Menone et al. (21) calculated that about 75 μg of Heptachlor-epoxide per square meter of cordgrass marsh exists, or over the 32 km^2 of cordgrass marsh, roots of *Spartina* densiflora concentrates 2 Kg of this contaminant. Total PCBs in the same estuary can be estimated as 22 μg per square meter or 0.7 Kg over the 32 km^2 of cordgrass marsh.

EXPOSURE AND UPTAKE OF MACROPHYTES TO CONTAMINANTS

The uptake of contaminants depends on the physicochemical characteristics of the compounds and on the life form of the macrophyte (floating-leaved, free-floating, well-rooted, or rootless species). Free-floating plants, such as *Lemna spp., Eichornia spp.*, and *Pistia spp.*, take up contaminants from the water by roots and/or leaves. Likewise, the rootless *Ceratophyllum spp.* takes up contaminants mainly through its finely divided leaves. The situation of species with a well-developed root-rhizome system and totally submersed foliage, such as *Myriophyllum spp., Potamogeton spp.*, and *Vallisneria spp.* species, is much more complex. Thus, these submersed aquatic macrophytes grow at the interface of two distinct environments, being their leaves exposed to the water column while their roots are in contact to the sediment. Both media (the water column and the sediment pore water) are potential sources of uptake of contamination (22,23). Indeed, the rooted aquatic macrophyte *Hydrilla verticillata*—common throughout the southern United States—has been shown to take up contaminants from both the sediment pore water and the overlying water column. It also has the ability to translocate sediment-incorporated contaminants from the roots into the vegetative portion of the plant (24).

Other species, such as *Schoenoplectus californicus, Spartina densiflora* (mainly from the Southern Hemisphere), and *Spartina alterniflora* (mainly from the Northern Hemisphere), have the root-rhizome system well developed, but have not totally submersed foliage. The first two macrophytes form the bulk of the emergent vegetation in many of the shallow lakes and estuarine areas of the Southern America regions. The pollutant sequestration in these macrophytes occur across many interfaces, sediment, water, and air (21,25,26).

Several environmental factors, including water and sediment pH, water current, sediment texture, organic carbon, and mineral composition, are known to influence contaminant adsorption by sediments and macrophytes growth, so they may contribute to overall site quality.

BIOMONITORING

During the past, direct chemical measurement of the concentrations of toxic and hazard substances have been carried out for the evaluation of water quality in contaminated natural environments. However, for environmental management, contaminant concentrations do not necessarily account for, or enable prediction of, the impairment of biota. If the objective is to monitor and improve environmental quality of the ecosystems, biotic measurements of contamination are more useful than only measures of water or sediment contaminant concentrations. Furthermore, the cost of evaluation and measurement of persistent and toxic organochlorines, such as PCBs, is substantial. Therefore, for ecologic and economic reasons, it is important to develop biological monitoring that is useful and valuable. The use of biomonitoring to assess and control discharges of toxic chemicals into the environment has been promoted as being a desirable alternative to more expensive, less realistic, and time-consuming chemical analysis (27). Thus, the presence of indicator organisms (biomonitors) provides a measure of cumulative exposure to contaminants over time and avoids the need for frequent sampling. The use of biomonitors *in situ* to identify and quantify toxicants in an environment is referred to as biomonitoring (28). This technique takes advantage of the ability of organisms to accumulate contaminants in their tissues through bioaccumulation and bioconcentration. The primary objective of biomonitoring investigations is to assess the quality of water in an area by relating observed responses of organisms that live within a suspected polluted site to the concentrations of contaminants detected within their tissues. An additional and very significant advantage of biomonitoring is that the bioaccumulated sublethal levels of contaminants within the tissues of organisms indicate the net amount of contaminants that have been integrated over a period of time (27).

The response of biota to pollution stress can be observed at the ecosystem, community, population, individual, and suborganismal levels of organization (29). Environmental assessments may be made by establishing quantitative relationships between (a) concentrations of pollutants that are accumulated within the tissues of organisms residing within a particular area and (b) manifest biological effects (30).

In general, three historical stages exist in the biomonitoring with plants, based on the use of different parameters:

1. various physiological, morphological, and community parameters (pigment content, photosynthetic activity, diversity indices, biomass, etc.)
2. environmental concentrations of pollutants in plant tissues
3. early warning systems or biomarkers for assessing contaminant exposure and effects (histological, biochemical, or genetic)

The presence of certain toxicants may induce a physiological response in an organism, often involving a heightened production of enzymes that are capable of metabolizing and/or degrading the toxicant in question. In this way, the quantity and activity of such xenobiotic-metabolizing enzymes may be used as indicators of the bioavailability of a specific contaminant in the environment. An increasing number of studies have used biochemical as well as physiological endpoints for assessing toxic effects on plants. These effects are often more sensitive, but their environmental relevance and their relationship to the impact of toxic chemicals on biomass are not known. However, they can be combined with other parameters of chemical exposure to predict the ecological consequences of chemical-specific contamination (31). For a biomarker to be applied in the field, it should be correlated with a significant effect, such as survival, growth, or reproduction. Although many biomarkers have been validated for use in evaluating animal health and exposure to toxic substances, both in the laboratory and in natural ecosystems, very few biomarkers have been validated and used to assess plant exposure and health effects under field conditions (31). Padinha et al. (32) has shown that variations in some indices of physiological stress, like thiolic protein concentrations in *Spartina maritima*, could be used as a tool to monitor contamination by heavy metals. On the other hand, Wall et al. (33) have found no adverse influences of PCBs in terms of peroxidase activity (POD), glutathione concentration (tGSH), photosynthesis, and transpiration on *S. alterniflora* from a Superfund site contaminated with $46.0 \pm 52.7\ \mu g\, g^{-1}$ dry weight in sediments. Therefore, the utility of POD as well as other detoxication enzymes as biomarkers of PCB stress deserves attention but needs further study.

In addition, laboratory research can complement and contribute to the best understanding of data from field studies (biomonitoring). Among laboratory experiments, we can mention:

— *Toxicity tests*: Toxicity tests involve exposing a well-defined test organism to a dilution series of a suspected toxicant under controlled laboratory conditions. The goal of toxicity tests is to correlate the level of toxicants to observed organismal responses. Of the multitude of organismal responses that could be observed as endpoints in toxicity test, the one most frequently used is survivorship. In ecotoxicology, survivorship is usually expressed as LD_{50}, which is the dosage of the suspected toxicant that is seen to cause mortality in half of the individuals tested within a specified time period.

— *Bioassays*: Bioassays are often used to assess the toxic effects of mixtures of compounds on biota by exposing test organisms to naturally contaminated water or sediment samples.

When a pollutant enters an aquatic environment, it is expected that its initial effect on an exposed organism will be a suborganismal one—either biochemical or genetic. In this manner, biochemical and genetic indicators may be able to detect the presence of minimum contaminant concentrations compared with the levels of toxicants required to elicit a response at the level of the entire organism (e.g., death) (28).

MACROPHYTES AS BIOMONITORS

Plant biomonitors provide an integrated description of pollution within an ecosystem (34). Freshwater species were used as sentinels of contaminant stress for many years, such as in the biomonitoring study of Wang and Williams (35), who used *Lemna minor* to examine the phytotoxicity of industrial effluents. Macrophytes have also proved to be useful for assessing organochlorine contamination in laboratory experiments (36,37) and in the field (27,28). The organochlorine bioaccumulation in submersed macrophytes may be very high; macrophytes may be three to four times more contaminated than sediment, and 6000 to 9000 times more contaminated than the water (38).

Under field conditions, *Schoenoplectus californicus* has already shown potential as a biomonitor of organochlorine pesticides. Moreover, this species has demonstrated the ability to function as phytoremediator—plant use for the remediation of contaminated environments—of these pesticides, because, in combination with other aquatic biota, about 40% of the more hydrophobic pesticides have been retained in the lake environment, leading to their lower release through the effluent creek (25,26,39). *Spartina densiflora* grows abundantly and contributes significantly to the primary productivity of the estuarine ecosystems on the Southern America coast. In Mar Chiquita coastal lagoon (Argentina), it can bioaccumulate both organochlorine pesticides and PCBs, and the total amount of these compounds in cordgrass biomass may represent a significant proportion of the total PCBs burden in the estuary system (21). The same species is the dominant plant in Humboldt Bay salt marshes. *S. densiflora*'s ability to rapidly expand in bare areas has implications for marsh mitigation and restoration activities (40).

Other known macrophytes, like *Vallisneria americana*, var. americana—one of the most abundant macrophyte in the Great Lakes (41,42)—*Potamogeton spp., Najas spp., Myriophyllum spp.*, and *Elodea spp.*, accumulate contaminants within its tissues and have also shown potential as biomonitors of organic contaminants in the field (27,28).

MACROPHYTES AND PCBS

Laboratory studies indicate that uptake of PCBs from contaminant water or sediments in aquatic macrophytes is expected (43). In addition, considering their limited mobility and their abundance in many aquatic systems, they could function as *in situ* biomonitors of water contaminants, like PCBs. However, the extent of sequestration in natural populations remains almost unknown. Thus, little impetus has developed for studying the role of submersed aquatic macrophytes in PCBs biogeochemical cycling in freshwater ecosystem, as evidenced by our literature survey, which revealed a lack of field data documenting the incidence of hydrophobic organic contaminants in feral aquatic macrophytes.

Exposure to sediment-borne PCBs is particularly important for rooted macrophytes. As a result of the hydrophobicity of these compounds, sediments frequently contain higher concentrations of contaminants than the surrounding water (44,45). Thus, plants grown in contaminated environments typically have higher concentrations of hydrophofobic pollutants in below-ground tissues than in foliage and, because their mobility within the plant tissues is very limited (23), they tend to accumulate in roots (28). Macrophytes collected from different sites may contain different concentrations of contaminants within their tissues, reflecting sediments, water, and air loads at each location (26,28). In a shallow lake from Argentina, it has been demonstrated that *S. californicus* accumulate PCBs in direct relationship to the sediment PCB concentrations. Moreover, the higher PCB levels were found in root tissues revealing the high ability of this macrophyte to function as PCB biomonitor (Miglioranza, personal communication).

The Bioaccumulation Factor (BF) is the ratio of the concentration in the biota and the concentration in the soil/sediment. It primarily depends on the properties of the soil/sediment and the biota, particularly the ratio of lipid in the biota and the organic carbon content of the soil/sediment (2). Root Bioaccumulation Factors (RBF) have been calculated for *Schoenoplectus californicus* and *Spartina densiflora* from Los Padres lake and Mar Chiquita coastal lagoon, respectively. The values in *S. californicus* ranged between 49 and 600 for different congeners of PCBs, being the highest RBF for the lower chlorinated congeners. In *S. densiflora*, the RBF ranged between 0.75 and 26.5. These results show the magnitude of bioaccumulation of PCBs in freshwater and estuarine macrophytes species and their importance in the ecology of these environments.

The investigations of Butler et al. (46) demonstrated for the first time that plant cells are capable of hydroxylating and glycosylating a chlorinated biphenyl in a manner similar to what has been reported for animals (47). The metabolism of PCBs varies between the plant species and is affected by the substitution pattern and the degree of chlorination (48,49). Wilken et al. (49) analyzed 12 different terrestrial plant species and showed that lower chlorination grade is associated with higher metabolism rates. Recent studies (50,51) have shown the important role of plant cytochrome P450 in metabolism of different toxicants, but have admitted involvement of peroxidases too (52,53).

Despite the scarce information about PCB concentrations in macrophytes under field conditions, we can conclude that macrophytes can play a crucial role in the biomonitoring and remotion of these toxic organic compounds from the environment, not only accumulating but also transforming them. For this reason, phytoremediation has been proposed as an alternative or complementary technique to treat sediment polluted by PCBs, but still needs much basic research.

BIBLIOGRAPHY

1. Orth, R.J. (1994). *Lake Reservoir Manage*. **10**: 49–52.
2. Connell, D.W. (1997). Lewis Publishers, New York, p. 505.
3. Borgmann, U. and Whittle, D.M. (1992). *J. Great Lakes Res.* **18**: 298–308.
4. Picer, M. and Picer, N. (1995). *Water Res.* **29**: 2707–2719.
5. Stow, C.A., Carpenter, S.R., Eby, L.A., Amrhein, J.F., and Hesselberg, R.J. (1995). *Ecol. Appl.* **5**: 248–260.
6. Rand, G.M. (1995). Ecological Services Inc., North Palm Beach, FL, p. 1125.
7. Scandalio, J.G. (1990). *Advances in Genetics* **28**: 1–12.
8. Cheung, C.C.C., Zheng, G.J., Lam, P.K.S., and Richardson, B.J. (2002). *Marine Pollution Bull.* **45**: 181–191.
9. Leitao, Ma da S., Cardozo, K.H.M., Pinto, E., and Colepicolo, P. (2003). *Arch. Environ. Contam. Toxicol.* **45**: 59–65.
10. Reid, B.J., Jones, K.C., and Cemple, K.T. (2000). *Environmental Pollution* **198**: 103.
11. Mac Donalds, C.R., Metcalfe, C.D., Balch, G.C., and Metcalfe, T.L. (1993). *Environ. Toxicol. Chem.* **12**: 195–224.
12. Larsson, P., Lennart, O., and Cronberg, G. (1998). *Can. J. Fish. Aquat. Sci.* **55**: 1926–1937.
13. Kilgore, K.J., Morgan, R.P., II, and Rybicki, N.B. (1989). *Nat. Amer. J. Fish. Manag.* **9**: 101–111.
14. Posey, M.H., Wigand, C., and Stevenson, J.C. (1993). *Estuarine Coast. Shelf Sci.* **37**: 539–555.
15. Lewis, M.A. and Wang, W. (1997). In: W. Wang, J.W. Gorsuch, and J.S. Hughes (Eds.). Lewis Publishers, New York, pp. 141–175.
16. Lodge, D.M. (1991). *Aquat. Bot.* **41**: 195–224.
17. Wetzel, R. (1995). *Freshwater Biol.* **33**: 83–89.
18. Dave, G. (1992). *Water, Air Soil Pollut.* **63**: 187–200.
19. Drake, J.C. and Heaney, S.I. (1987). *Freshwater Biol.* **17**: 513–523.
20. Melzer, A. (1999). *Hydrobiologia*. **395/396**: 181–190.
21. Menone, M.L. et al. (2000). *Estuaries*. **23**: 583–592.
22. Barko, J.W., Gunnison, D., and Carpenter, S.R. (1991). *Aquat. Bot.* **41**: 41–65.
23. Guilizzoni, P. (1991). *Aquat. Bot.* **41**: 87–109.
24. Byl, T.D., Sutton, H.D., and Klaine, S.J. (1994). *Environ. Toxicol. Chem.* **13**: 509–515.
25. Miglioranza, K.S.B. et al. (2002). *Environ. Sci. Pollut. Res.* **9**: 250–256.
26. Miglioranza, K.S.B., de Moreno, J.E.A., and Moreno, V.J. (2004). *Water Res.* **38**: 1765–1772.
27. Lovett-Doust, L., Lovett-Doust, J., and Biernacki, M. (1994). *J. Great Lakes Res.* **20**: 333–354.
28. Lovett-Doust, J., Schmidt, M., and Lovett-Doust, L. (1994). *Biologic. Rev.* **69**: 147–186.
29. Cairns, J. and Niederlehner, B.R. (1992). *J. Environ. Sci.* **4**: 1–9.
30. Sandhu, S.S. and Lower, W.L. (1992). In: S.S. Sandhu, D.M. DeMarini, M.J. Mass, M.M. More, and J.L. Munford (Eds.). Plenum Press, New York, pp. 145–160.
31. Lytle, J.S. and Lytle, T.F. (2001). *Environ. Toxicol. Chem.* **20**: 68–83.
32. Padinha, C., Santos, R., and Brown, M.T. (2000). *Marine Environ. Res.* **49**: 67–78.
33. Wall, V.D. et al. (2001). *Arch. Environ. Contam. Toxicol.* **40**: 10–17.
34. Dennison, W.C. et al. (1988). *Biosciences* **43**: 86–94.
35. Wang, W. and Williams, J.M. (1988). *Environ. Toxicol. Chem.* **7**: 645–652.
36. Biernacki, M., Lovett-Doust, J., and Lovett-Doust, L. (1995). *J. Appl. Ecol.* **32**: 761–777.
37. Biernacki, M., Lovett-Doust, J., and Lovett-Doust, L. (1995). In: *Ecovision Worls Monograph Series*. M. Munawar, T. Edsall, and J. Leach (Eds.). S.P.B. Academic Publishing, The Netherlands, pp. 413–426.

38. Painter, S. (1990). In: *Proc. Aquat. Ecosystem Health Sympos*. Waterloo, Ontario.
39. González Sagrario, M.A., Aizpún de Moreno, J.E., Moreno, V.J., and Escalante, A.H. (1998). *Environ. Sci.* **6**: 153–169.
40. Kittelson, P.M. and Boyd, M.J. (1997). *Estuaries* **20**: 770–778.
41. Schloesser, D.W., Edsall, T.A., and Manny, B.A. (1985). *Can. J. Bot.* **63**: 1061–1065.
42. Catling, P.M., Spicer, K.W., Biernacki, M., and Lovett-Doust, J. (1994). *Can. J. Plant Sci.* **74**: 883–897.
43. Gobas, F.A.P.C., McNeil, E.J., Lovett-Doust, L., and Haffner, G.D. (1991). *Environ. Sci. Technol.* **25**: 924–929.
44. Di Toro, D. et al. (1991). *Environ. Toxicol. Chem.* **10**: 1541–1583.
45. Burton, A. (1992). In: A. Burton (Ed.). Lewis Publishers, Boca Raton, FL, Ch. 8, pp. 167–181.
46. Butler, J.M., Groeger, A.W., and Fletcher, J.S. (1992). *Bull. Environ. Contam. Toxicol.* **49**: 821–826.
47. Sandermann, H. (1982). In: *Environmental Mutagenesis, Carcinogenesis, and Plant Biology*. E.J. Klekowski (Ed.). Vol. 1, Praeger, New York, pp. 2–32.
48. Lee, I. and Fletcher, J.S. (1992). *Plant Cell Reports* **11**: 97–100.
49. Wilken, A., Bock, C., Bokern, M., and Harms, H. (1995). *Environ. Toxicol. Chem.* **14**: 2017–2022.
50. Pflugmacher, S., Geissler, K., and Steinberg, C. (1999). *Ecotoxicol. Environ. Safety* **42**: 62–66.
51. Stiborová, M., Schmeiser, H.H., and Frei, E. (2000). *Phytochemistry* **54**: 353–362.
52. Koller, G., Moder, M., and Czihal, K. (2000). *Chemosphere* **41**: 1827–1834.
53. Chromá, L., Kucerová, P., Macek, T., and Macková, M. (2002). *Book of Abstracts, The Second PCB Workshop*. I. Holoubek, and I. Holoubkova (Eds.). Masaryk University Brno, p. 153.

WASTEWATER MANAGEMENT FOR DEVELOPING COUNTRIES

KONSTANTINOS P. TSAGARAKIS
University of Crete
Rethymno, Greece

INTRODUCTION

A country is characterized as a developing one according to specific economic indicators. The World Bank has linked gross domestic product (GDP) to pollution for developing nations (1) (Fig. 1). This link breaks when incentives to protect the environment are introduced, followed by the adoption of cleaner and more efficient technologies, which can be adjusted to the case of water pollution originating from the disposal of untreated wastewater. Once increasing pollution has had negative effects on the wellbeing and economy of an area, incentives for wastewater treatment are induced. Legislation is introduced requiring polluters to pay for the treatment of wastewater at a certain level. Construction of wastewater treatment plants (WTPs) in compliance with legislation reduces pollution from wastewater and has a positive effect on a quality index based on a combination of treated effluent qualitative parameters.

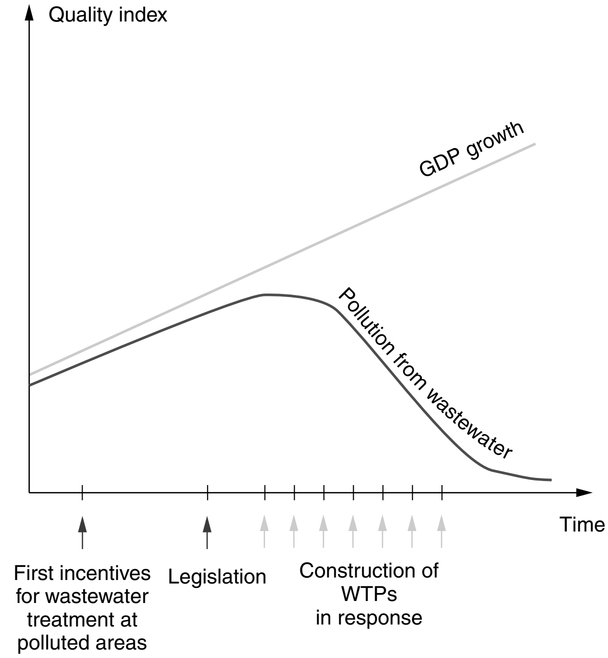

Figure 1. Breaking the link between economic growth and pollution from wastewater (1).

Wastewater management involves collection treatment and disposal/reuse. Potable water supply is inevitably related to them, hence some principles discussed below will include or directly apply to such issues. Ujang and Buckley (2) summarize sanitation problems for developing countries as lack of environmental awareness, insufficient expertise, inappropriate policies, insufficient funding, insufficient water resources, inappropriate management systems, and institutional support.

Following this introduction, the application of existing technologies to developing countries (DC) is discussed, focusing on technology and knowledge transfer and the role of international experts. Some differences over wastewater production and its quality are stressed with special reference to the high standards often set by decision makers in DC. Finally, issues on sustainable technology selection and key points referring to methodologies and indexes are discussed.

APPLICATION OF EXISTING TECHNOLOGIES TO DEVELOPING COUNTRIES

Wastewater management must be considered as an integral part of the development process and national plans should be formulated (3). Conditions applicable when planning wastewater treatment facilities for DC are not identical to those prevalent during planning in developed countries in the past. Although some similarities exist, many differences also exist. One reason is that water pollution issues are not the main concern in DC because of other more pressing issues such as national or racial security, food availability, and epidemic control (2).

Technology Transfer

Today, technology is available for the treatment of wastewaters of any origin and strength. DC cannot be expected to play a key role in the development of cutting edge technology in wastewater management or water research. However, research is needed in order for available technology to be applied to specific country-region conditions. This kind of research is an investment for those countries concerned. The existing transfer of technology should be done wisely because today, developed countries are regarded as an area of investment for water companies. For example, the design approach in rural Egypt was to select well-established technologies; little stress was placed on selecting more innovative technologies. However, less sophisticated treatment technologies, such as stabilization ponds, have proven capable of meeting the required effluent standards (4). A similar situation was reported in Greece. From the 147 small ($<$10,000 p.e.) WTPs, 80% were activated sludge systems, resulting in one out of three being out of operation. If some natural treatment technologies had been properly installed instead of activated sludge systems, failures would have been less frequent (5).

An approach by UNIDO (6), for the acquisition of foreign technology in DC, recommends the following steps:

a. It should be determined that the technology has been commercially proven yet is not obsolete;
b. Alternative technologies that may be available should be evaluated comparatively covering the following points: cost of obtaining such technologies, principal inputs required and their local availability, and estimation of manufacturing costs and profitability.
c. When a country cannot use a certain technology because of foreign exchange limitations or other constraints, a comparative evaluation should still be made, as in (b) above, for purposes of negotiation.

In most cases, when a project is to be evaluated for a DC, shadow pricing should be undertaken to adjust market prices to reflect opportunity costs. A conversion factor, i.e., shadow rate, should be calculated, which then needs to be multiplied with the market prices for unskilled labor wage, the foreign exchange rate, the opportunity cost of capital, and the shadow price of land, water, and other direct inputs (7).

In DC, institutional barriers often exist for hampering the effective application of projects. In many countries, emphasis has been given to increasing the coverage of service facilities to all parts of the population. However, the absence of institutional safeguards, at both the community and government levels, to ensure that these service facilities can be operated and maintained in a sustainable way has caused the facilities to deteriorate, creating an unacceptable financial burden on governments using traditional supply-driven projects (8).

Knowledge Transfer

Knowledge is less available in DC than in developed countries. At the same time, contemporary means of communication have made knowledge more accessible. The World Wide Web, international organizations, networks of excellence, and textbooks all play a positive role in its availability. Textbooks and other sources of knowledge are invaluable tools for design engineers but should always be used wisely. Not all information is applicable to all areas without screening. In fact, if necessary adjustments to local conditions are not made, projects may face serious problems (9).

International Experts and Advisors

Very commonly, advice is sought from international experts/advisors who are asked to give assistance on a specific project. Such assistance is useful when it applies to the technology itself, but to make any wastewater project or plan work, the adaptation should incorporate local conditions. Local experts should therefore always be involved. International experts or advisors can also be employed for solving specific operational problems. The selection of international experts/advisors should be done carefully. Reid (3) pinpoints nine cases of problems associated with advisors. He discusses how some may act like promoters of specific projects to a government resulting in low returns or even losses. In addition, too many advisors on the same project can cause confusion. When advisors are given a difficult task but not given all the information they need to solve a problem, no solution may be found. Although donor countries may offer financial and technical assistance for projects, this can be a misallocation of funds if they do not also take the recipient country's more immediate needs into account.

WASTEWATER QUALITY AND QUANTITY FOR DEVELOPING COUNTRIES

Wastewater Production and its Quality

Historical data on the flow rate is necessary for design purposes, therefore daily, seasonal, and annual variations should be acquired where possible. Reid (3) supports urging each country to establish their own resources agency with responsibility to collect pertinent data and to plan with a regional approach rather than a case-by-case approach.

Where this is not possible, accurate estimations should be taken according to local conditions and information. Crites and Tchobanoglous (10) proposed a domestic base flow rate of 245–300 L/inh.·d. These values apply mostly to developed countries. Uncritical adoption of the above values or other values may lead to overdesigning of the installation. The size of the community served, its socioeconomic conditions, and other parameters may influence the per inhabitant flow rate. Pujol and Lienard (11) proposed 150 (±50) L/inh.·d for rural areas of less than 5,000 population equivalent (p.e.). Campos and Von Sperling (12) developed a model for wastewater production in Brazil that showed significant variation according to the income of the people served (74–210 L/inh.·d). For Greek urban and rural communities, these values are on average 165 L/inh.·d and 100–120 L/inh.·d, respectively, and for Turkish cities 170 L/inh.·d (13,14). Wastewater production may also vary with time and conditions in the same area. Shaheen (15) reports that, in the West Bank, the water consumption was 58 l/inh.· d in 1991, but under a specific scenario, this will rise to 122 l/inh.·d by 2010.

Usually, wastewater in DC is characterized by having high concentrations of basic pollutants. A study in Jordan showed that the average values for 5 WTP in terms of BOD_5, COD, and TSS, were 616–1419 mg/L, 1159–3707 mg/L, and 458–1647 mg/L, respectively (16). In contrast, average BOD_5/COD/SS concentrations were 187/477/236 mg/L, 268/634/302 mg/L, 171/450/237 mg/L, and 267/581/426 mg/L for Belgium-Flanders, France, Netherlands, and Slovenia, respectively (17). One of the major problems in DC is the illegal discharge of industrial effluents into public sewers, unauthorized connections, and unregistered population (18).

Effluent Quality and Standards

Low levels for effluent requirements and standards are frequently introduced in DC, but these levels are, in reality, rarely met. No means exists of imposing any kind of effluent requirements or of adopting the most advanced technology of wastewater treatment, when no mechanism exists to safeguard the maximum possible efficiency.

Many of the high standards used in the developed world have only superficial environmental justification, and the costs of implementation are often not considered by the quality regulators. The regulation and monitoring of standards carries a cost to society that is often considered or appreciated by those setting standards (19).

A list of common problems associated with setting up and implementing standards in developing countries are discussed by von Sperling and Chernicharo (20). Of those, the most important are:

- Standards are often copied by developed countries.
- Protection measures that do not lead to immediate compliance with the standards do not obtain licensing or financing.
- Developing countries attempt to reach developed countries' status too quickly.
- No institutional development exists that could support and regulate the implementation of standards, thus standards are not actually enforced.

von Sperling and Chernicharo (20) propose a sustainable approach for several situations whereby wastewater treatment systems (and effluent standards) should first implement a less efficient process, or a process that removes fewer pollutants. This entails transferring to a second stage the improvement toward a system more efficient or more wide-reaching in terms of pollutants, and so on.

CENTRALIZED AND DECENTRALIZED TREATMENT

Cities in DC, if no wastewater facilities have been established, usually manage their wastewaters in septic tanks, absorption pits, or collect the wastewater and discharge untreated effluent into the environment. To establish a new treatment plant, they need the support of the population that is to be served.

In order to gain this support, people should be informed about the environmental benefits and be willing to pay for the construction and the O&M costs of the WTP. The acceptability of such projects may decline when direct costs are applied to the population to be served (21).

The task of wastewater treatment comes after the collection of the wastewater. Onsite technologies do not have some of the barriers on efficient wastewater management in DC as management is on the owners. The number and location of the WTP is determined by taking two main directions into consideration. The first is that the centralization of the wastewater produced in a large region will give an additional cost per capita, as more meters of sewer and pumping are needed and second, the larger a plant is, the less per capita it will cost because of economies of scale. Therefore, this will give an optimum degree of centralization as shown in the theoretical case of (Fig. 2).

Centralization indicates an increasing number of houses or clusters of houses connected to a main sewerage network instead of being treated onsite. Each cluster of houses, suburb, or neighboring village has an additional cost for connecting itself to a main sewerage network. Although an economy of scale exists in treatment, this is not the case with the cost of centralization of sewers when remote areas are added to the main network. In centralized treatment, extra costs are needed for pumping stations and piping. Furthermore, during the transport of wastewater over long distances, hydrogen sulphide is produced. Extra costs are then needed for odor prevention or reduction. In addition, extra costs are necessary for maintaining the system and pumping energy (22,23). In any case, the cost of centralization should be considered on top of the cost of treatment, with an economic analysis to investigate its feasibility. In a hypothetical case study, after a number of agglomerations have been centralized, it is not worth including others, because the additional cost of connecting them would offset the economy of scale achieved by the treatment of the wastewater.

A combination of onsite and centralized treatment cannot be overruled. In rural areas with a number of

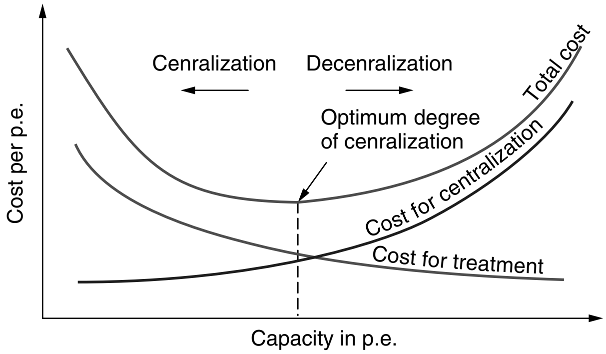

Figure 2. Effect of centralization to the total economic cost of sanitation projects.

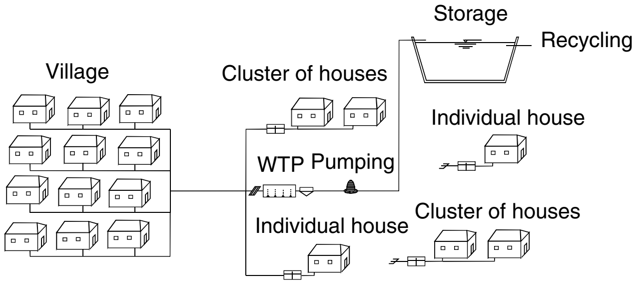

Figure 3. Combined centralized and decentralized wastewater management approaches.

villages, centralized treatment for the villages can exist together with onsite treatment for houses away from the WTP. For some other villages or cluster of houses, onsite septic tanks can be used adjacent to the villages, and then effluent can be carried to the WTP as shown in (Fig. 3).

In a study focusing on small communities of the MENA region, it is proposed that solids are removed in an interceptor tank, which is part of a household connection. Settled sewerage costs are quite low in comparison with conventional sewerage, mainly because of the shallow excavation depths, use of small diameter PVC paperwork, and simple inspection chambers (24).

TECHNOLOGY SELECTION

Sustainable Solutions

A great range of technical solutions are available today. From them are selected those that can produce such effluent quality as to protect receiving surfaces and underground water bodies, in all cases having no negative effects on the environment. These solutions must also protect public health and be acceptable to society. They should be affordable to the users and be operated effectively. The environment, society, and economy are the three aspects that must be given priority when selecting among different solutions (25,26) (Fig. 4).

Reid (3) defines as an appropriate technology for wastewater treatment the one that is acceptable to users and can be maintained by the community. No ideal solution applicable to all conditions exists, and each situation/project should be considered individually (27,28).

Society barriers and limitations may develop from social acceptance, amenity value, community participation, public awareness institutional requirements, local parameters, and stimulation of sustainable behavior (29).

The effluent destination is also a critical factor in the choice of technology for wastewater treatment. Reuse should be encouraged where possible. It is necessary to carry out careful investigations, which are based on local requirements and conditions, rather than uncritically adopting practices from abroad (30).

Methodologies/Indexes

It is possible to use indicators and methodologies to evaluate the different alternatives. These should be used as case study tools and should never lead to a universal classification of all different options. Balkema et al. (25) point out that some of the indicators are hard to quantify and ensure the integration and multidimensional character of the sustainability assessment.

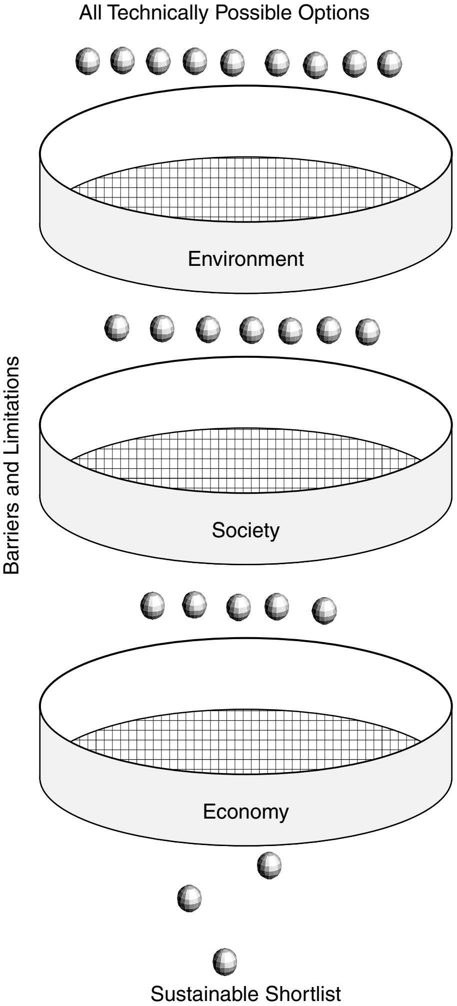

Figure 4. Sustainable short listed wastewater management solutions.

Phybus and Schoeman (31) present various monitoring and performance indicators for developing areas. They conclude that the adoption and use of performance indicators at the community level should be made a stimulating experience for the players involved. It should introduce an element of competition within the initial stages and can lead from there to comparison with neighboring communities.

Schutte (32) suggests that for an organization responsible for water supply and sanitation services in developing areas, five key success factors should be applied:

- Create an organization-wide culture of service to customers.
- Ensure reliability in the water supply.
- Ensure community involvement.
- Create a culture of awareness for focusing on generation and collection of income.
- Create a culture of cost consciousness and focus on minimization of losses.

Next Page

BIBLIOGRAPHY

1. World Bank. (1992). *World Development Report 1992: Development and the Environment*. Oxford University Press, Oxford, UK.
2. Ujang, Z. and Buckley, C. (2002). Water and wastewater in developing countries: present reality and strategy for the future. *Water Sci. Technol.* **46**(9): 1–9.
3. Reid, G.W. (1978). Introduction: interface for decision. In: *Appropriate Methods of Treating Water and Wastewater in Developing Countries*. G.W. Reid and K. Coffey (Eds.). Bureau of Water and Environmental Resources Research, Univ. of Oklahoma, Washington, DC.
4. Gaber, A., Antill, M., Kimball, W., and Wahab, R.A. (1993). Planning replaceable small flow wastewater treatment facilities in developing nations. *Water Sci. Technol.* **28**(10): 1–8.
5. Tsagarakis, K.P., Mara, D.D., Horan, N.J., and Angelakis, A.N. (2000). Small municipal wastewater treatment plants in Greece. *Water Sci. Technol.* **41**(1): 41–48.
6. UNIDO. (1973). *Guidelines for the Acquisition of Foreign Technology in Developing Countries*. United Nations Industrial Development Organization, Vienna, Austria, p. 43.
7. Mara, D. (1996). *Low Cost Urban Sanitation*. John Wiley, Chichester, UK, pp. 171–192.
8. Garn, H.A. (1997). *Lessons from Large-Scale Rural Water and Sanitation Projects: Transition and Innovation*. Water and Sanitation Program Working Paper. World Bank, Washington, DC.
9. Tsagarakis, K.P., Mara, D.D., Horan, N.J., and Angelakis, A.N. (2001). Institutional status and structure of wastewater quality management in Greece. *Water Policy* **3**(1): 81–99.
10. Crites, R. and Tchobanoglous, G. (1998). *Small and Decentralised Wastewater Management Systems*. WCB and McGraw-Hill, New York.
11. Pujol, R. and Lienard, A. (1990). Qualitative and quantitative characterisation of wastewater for small communities. *Water Sci. Technol.* **22**(3–4): 253–260.
12. Campos, H.M. and von Sperling, M. (1996). Estimation of domestic wastewater characteristics in a developing country based on socio-economic variables. *Water Sci. Technol.* **34**(3–4): 71–77.
13. Tsagarakis, K.P., Mara, D.D., and Angelakis, A.N. (2001). Wastewater treatment in Greece: experience and lessons for developing countries. *Water Sci. Technol.* **44**(6): 163–172.
14. Karagözoglu, B., Altin, A., and Degirmenci, M. (2003). Flowrate and pollution characteristics of domestic wastewater. *Int. J. Environ. Pollut.* **19**(3): 259–270.
15. Shaheen, H.Q. (2003). Wastewater reuse as means to optimize the use of water resources in the West Bank. *Water Int.* **28**(2): 201–208.
16. Jamrah, A.I. (1999). Assessment of characteristics and biological treatment technologies of Jordanian wastewater. *Bioprocess Engineering* **21**: 331–340.
17. Pons, M.N., Spanjers, H., Baetens, D., Nowak, O., Gillot, S., Nouwen, J., and Schuttinga, N. (2004). *Wastewater Characteristics in Europe - A Survey*. European Water Management Online, Official Publication of the European Water Association (EWA), Brussels, Belgium.
18. Karn, S.K. and Harada, H. (2001). Surface water pollution in three urban territories of Nepal, India, and Bangladesh. *Environ. Manag.* **28**(4): 483–496.
19. Johnstone, D.W.M. and Horan, N.J. (1994). Standards, costs and benefits: an international perspective. *J. IWEM.* **8**: 450–458.
20. von Sperling, M. and Chernicharo, C.A.L. (2002). Urban wastewater treatment technologies and the implementation of discharge standards in developing countries. *Urban Water*. **4**: 105–114.
21. Roomratanapum, W. (2001). Introducing centralised wastewater treatment in Bangkok: a study of factors determining its acceptability. *Habitat Int.* **25**: 359–371.
22. Townley, G. (1978). Cost effectiveness of the socio-cultural and health benefits of sewerless alternatives for domestic waste disposal in developing countries. In: *Appropriate Methods of Treating Water and Wastewater in Developing Countries*. G.W. Reid and K. Coffey (Eds.). Bureau of Water and Environmental Resources Research, Univ. of Oklahoma, Washington, DC, pp. 521–619.
23. Leeuw, E.J. and Van't, E. (1996). Oever, process selection, design and operation of the EDE WWTP. *Water Sci. Technol.* **33**(12): 57–63.
24. Bakir, H.A. (2001). Sustainable wastewater management for small communities in the Middle East and North Africa. *J. Environ. Manag.* **61**: 319–328.
25. Balkema, A.J., Preisig, H.A., Otterpohl, R., and Lambert, F.J.D. (2002). Indicators for the sustainability assessment of wastewater treatment systems. *Urban Water*. **4**: 153–161.
26. Bradley, B.R., Daigger, G.T., Rubin, R., and Tchobanoglous, G. (2002). Evaluation of onsite wastewater treatment technologies using sustainable development criteria. *Clean Techn. Environ. Policy* **4**: 87–99.
27. von Sperling, M. (1996). Comparison among the most frequently used systems for wastewater treatment in developing countries. *Water Sci. Technol.* **33**(1): 59–72.
28. Tsagarakis, K.P. (1999). *The Treatment of Municipal Wastewater in Greece*. PhD. Thesis, University of Leeds, School of Civil Engineering, Leeds, UK.
29. Burhard et al. (2000).
30. Tsagarakis, K.P., Tsoumanis, P., Charzoulakis, K., and Angelakis, A.N. (2001). Water resources status including wastewater treatment and reuse in Greece: related problems and prospectives. *Water Int.* **26**(2): 252–258.
31. Phybus, P. and Schoeman, G. (2001). Performance indicators in water and sanitation for developing areas. *Water Sci. Technol.* **44**(6): 127–134.
32. Schutte, C.F. (2001). Managing water supply and sanitation services to developing countries: key success factors. *Water Sci. Technol.* **44**(6): 155–134.

MERCURY REMOVAL FROM COMPLEX WASTE WATERS

K. Thomas Klasson
Oak Ridge National Laboratory,
Oak Ridge, Tennessee

The Remediation Technology Group in the Chemical Technology Division at Oak Ridge National Laboratory obtains information on mercury contamination throughout

---

Previous Page

BIBLIOGRAPHY

1. World Bank. (1992). *World Development Report 1992: Development and the Environment*. Oxford University Press, Oxford, UK.
2. Ujang, Z. and Buckley, C. (2002). Water and wastewater in developing countries: present reality and strategy for the future. *Water Sci. Technol.* **46**(9): 1–9.
3. Reid, G.W. (1978). Introduction: interface for decision. In: *Appropriate Methods of Treating Water and Wastewater in Developing Countries*. G.W. Reid and K. Coffey (Eds.). Bureau of Water and Environmental Resources Research, Univ. of Oklahoma, Washington, DC.
4. Gaber, A., Antill, M., Kimball, W., and Wahab, R.A. (1993). Planning replaceable small flow wastewater treatment facilities in developing nations. *Water Sci. Technol.* **28**(10): 1–8.
5. Tsagarakis, K.P., Mara, D.D., Horan, N.J., and Angelakis, A.N. (2000). Small municipal wastewater treatment plants in Greece. *Water Sci. Technol.* **41**(1): 41–48.
6. UNIDO. (1973). *Guidelines for the Acquisition of Foreign Technology in Developing Countries*. United Nations Industrial Development Organization, Vienna, Austria, p. 43.
7. Mara, D. (1996). *Low Cost Urban Sanitation*. John Wiley, Chichester, UK, pp. 171–192.
8. Garn, H.A. (1997). *Lessons from Large-Scale Rural Water and Sanitation Projects: Transition and Innovation*. Water and Sanitation Program Working Paper. World Bank, Washington, DC.
9. Tsagarakis, K.P., Mara, D.D., Horan, N.J., and Angelakis, A.N. (2001). Institutional status and structure of wastewater quality management in Greece. *Water Policy* **3**(1): 81–99.
10. Crites, R. and Tchobanoglous, G. (1998). *Small and Decentralised Wastewater Management Systems*. WCB and McGraw-Hill, New York.
11. Pujol, R. and Lienard, A. (1990). Qualitative and quantitative characterisation of wastewater for small communities. *Water Sci. Technol.* **22**(3–4): 253–260.
12. Campos, H.M. and von Sperling, M. (1996). Estimation of domestic wastewater characteristics in a developing country based on socio-economic variables. *Water Sci. Technol.* **34**(3–4): 71–77.
13. Tsagarakis, K.P., Mara, D.D., and Angelakis, A.N. (2001). Wastewater treatment in Greece: experience and lessons for developing countries. *Water Sci. Technol.* **44**(6): 163–172.
14. Karagözoglu, B., Altin, A., and Degirmenci, M. (2003). Flowrate and pollution characteristics of domestic wastewater. *Int. J. Environ. Pollut.* **19**(3): 259–270.
15. Shaheen, H.Q. (2003). Wastewater reuse as means to optimize the use of water resources in the West Bank. *Water Int.* **28**(2): 201–208.
16. Jamrah, A.I. (1999). Assessment of characteristics and biological treatment technologies of Jordanian wastewater. *Bioprocess Engineering* **21**: 331–340.
17. Pons, M.N., Spanjers, H., Baetens, D., Nowak, O., Gillot, S., Nouwen, J., and Schuttinga, N. (2004). *Wastewater Characteristics in Europe - A Survey*. European Water Management Online, Official Publication of the European Water Association (EWA), Brussels, Belgium.
18. Karn, S.K. and Harada, H. (2001). Surface water pollution in three urban territories of Nepal, India, and Bangladesh. *Environ. Manag.* **28**(4): 483–496.
19. Johnstone, D.W.M. and Horan, N.J. (1994). Standards, costs and benefits: an international perspective. *J. IWEM.* **8**: 450–458.
20. von Sperling, M. and Chernicharo, C.A.L. (2002). Urban wastewater treatment technologies and the implementation of discharge standards in developing countries. *Urban Water*. **4**: 105–114.
21. Roomratanapum, W. (2001). Introducing centralised wastewater treatment in Bangkok: a study of factors determining its acceptability. *Habitat Int.* **25**: 359–371.
22. Townley, G. (1978). Cost effectiveness of the socio-cultural and health benefits of sewerless alternatives for domestic waste disposal in developing countries. In: *Appropriate Methods of Treating Water and Wastewater in Developing Countries*. G.W. Reid and K. Coffey (Eds.). Bureau of Water and Environmental Resources Research, Univ. of Oklahoma, Washington, DC, pp. 521–619.
23. Leeuw, E.J. and Van't, E. (1996). Oever, process selection, design and operation of the EDE WWTP. *Water Sci. Technol.* **33**(12): 57–63.
24. Bakir, H.A. (2001). Sustainable wastewater management for small communities in the Middle East and North Africa. *J. Environ. Manag.* **61**: 319–328.
25. Balkema, A.J., Preisig, H.A., Otterpohl, R., and Lambert, F.J.D. (2002). Indicators for the sustainability assessment of wastewater treatment systems. *Urban Water*. **4**: 153–161.
26. Bradley, B.R., Daigger, G.T., Rubin, R., and Tchobanoglous, G. (2002). Evaluation of onsite wastewater treatment technologies using sustainable development criteria. *Clean Techn. Environ. Policy* **4**: 87–99.
27. von Sperling, M. (1996). Comparison among the most frequently used systems for wastewater treatment in developing countries. *Water Sci. Technol.* **33**(1): 59–72.
28. Tsagarakis, K.P. (1999). *The Treatment of Municipal Wastewater in Greece*. PhD. Thesis, University of Leeds, School of Civil Engineering, Leeds, UK.
29. Burhard et al. (2000).
30. Tsagarakis, K.P., Tsoumanis, P., Charzoulakis, K., and Angelakis, A.N. (2001). Water resources status including wastewater treatment and reuse in Greece: related problems and prospectives. *Water Int.* **26**(2): 252–258.
31. Phybus, P. and Schoeman, G. (2001). Performance indicators in water and sanitation for developing areas. *Water Sci. Technol.* **44**(6): 127–134.
32. Schutte, C.F. (2001). Managing water supply and sanitation services to developing countries: key success factors. *Water Sci. Technol.* **44**(6): 155–134.

MERCURY REMOVAL FROM COMPLEX WASTE WATERS

K. Thomas Klasson
Oak Ridge National Laboratory,
Oak Ridge, Tennessee

The Remediation Technology Group in the Chemical Technology Division at Oak Ridge National Laboratory obtains information on mercury contamination throughout

---

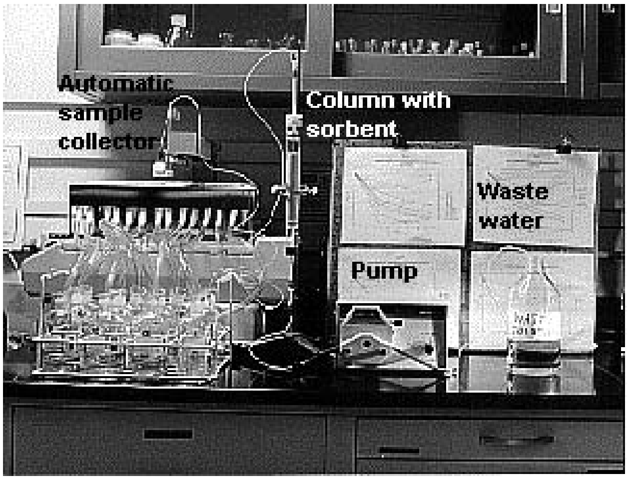

Figure 1. The typical experimental setup for long-term studies employs a continuous flow of waste water through a column packed with sorbent.

the Department of Energy (DOE) complex, and conducts research and development activities to serve end-users.

Mercury-contaminated soils, sludges and aqueous residues at DOE sites usually contain sufficient radioactivity to invoke special handling under the Atomic Energy Act and Resource Conservation and Recovery Act. Typical examples are high-salt acidic wastes stored at Idaho Chemical Processing Plant, alkaline waste stored at Savanah River Site (SRS), and mercury-bearing lithium hydroxide and low-salt neutral pH groundwater at the Oak Ridge Y-12 Plant. Difficulties in removing mercury from these wastes are due to its concentration, speciation, interference from other dissolved species, and pH of the bulk media.

Figures 2 and 3 shows data from a feasibility study conducted on a surrogate waste similar to the SRS waste water. Typically, a large number of sorbents are screened for applicability before an isotherm study is conducted. Based on the results from these experiments, sorbents are selected for continuous column experiments in which data are collected to determine the long-term performance of sorbents. The data collected in these type of experiments can be used to design full-scale systems.

Every waste is different, so there is always a need for these type of experiments. We prefer to work with the actual waste whenever possible; however, sometimes the cost is prohibitive.

In Fig. 2, are the results of isotherm studies with SRS waste. The three sorbents performed quite differently; the Mersorb (sulfur impregnated activated carbon pellets) appeared to be able to adsorb far more mercury than the other sorbents.

Column studies bring another dimension into the project. In Fig. 3, the concentration of the mercury in the effluents from columns packed with SR-3 or Mersorb sorbent are displayed. There is no mercury exiting the column until about 2200 bed volumes (BV) had passed through the system when SR-3 was used as sorbent. The breakthrough was sharp, indicating efficient use of the column and the mass transfer zone was 6.2 mm deep. The flow rate was about 11 BV per hour. In the case of Mersorb about 1000 BV could be processed before mercury appeared in the effluent. In this case the breakthrough was more gradual over time, which is less desirable.

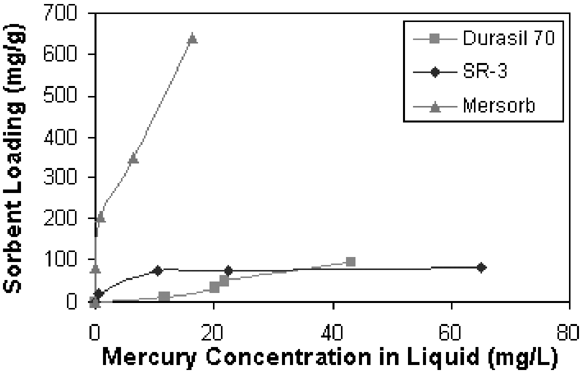

Figure 2. These data are from an experiment conducted with a waste water similar to tho one stored at the Savannah River Site. The original mercury concentration was 108 mg/L. Various amounts of sorbent were placed in contact with 50 mL of water for 24 h before the mercury was measured in the free liquid. The purpose of the experiment was to determine the isotherms.

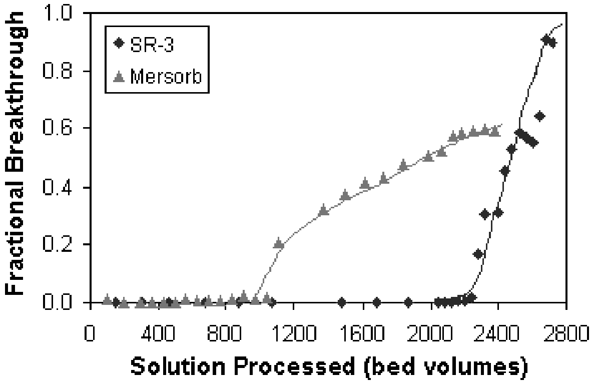

Figure 3. These data are from an experiment conducted with surrogate SRS waste passed through a column packed with a sorbent. As is noted, the effluent from the column is free of mercury until the sorbent becomes saturated.

METAL SPECIATION AND MOBILITY AS INFLUENCED BY LANDFILL DISPOSAL PRACTICES

Samuel C. Ashworth
Idaho National Engineering &
Environmental Laboratory
Idaho Falls, Idaho

Currently, landfills are constructed of a system of liners, leachate collection systems, and gas venting and collection systems. Hazardous wastes normally have two synthetic liners on the bottom and two leachate collection systems, one within the waste and one between the liners. These bottom liners typically rest on bentonite clay. Landfills have one bottom liner (New York requires two) resting on clay and one leachate collection system within the waste. A schematic of a typical landfill system is shown in Fig. 1. The landfill has a top cover system consisting of a synthetic liner with a topsoil cover designed to deflect precipitation. Not all of the liner components are shown, geotextiles and sands are used in various combinations for cushioning and inducement of preferential flow paths.

The following discusses and summarizes the mechanisms involved in metal speciation and mobility, modeling including references that provide more modeling information, and specific research in some of the areas.

Some precipitation inevitably enters the waste from liner permeation, holes, and cracks. However, most precipitation is deflected to stormwater collection, called runoff in Fig. 1. The dominant flow mechanisms into the waste are from the construction defects and not diffusion through the synthetic liners (1).

The precipitation that enters landfill wastes interacts with waste components and leaches metals and other substances. There are very complex interactions between the leachate and waste components that depend on pH, the metal and other component speciation, the presence of ligands, overall waste composition, redox potential, and others. Most of the leachate is collected in the leachate collection and treatment system. However, some of the leachate penetrates the bottom liner system from mechanisms similar to infiltration, and enters the soil column, and subsequently enters the groundwater.

Metal speciation and mobility are influenced by physical, chemical, and biological processes, and complex interactions between these processes may result in transporting specific contaminants at different rates (1). Attenuation of mobile metals depends on the amount of time that the contaminant is in contact with the material, the physical and chemical characteristics of the material, and the characteristics of the leachate.

Physical processes that control the transport of contaminants in the subsurface include advection (convection), the mixing effects of dispersion and diffusion, and physical sorption. Chemical processes that are important in controlling subsurface transport include precipitation/dissolution, chemical sorption, redox reactions, ion exchange, and complexation. These processes are generally reversible. The reversible processes tend to retard transport but do not permanently remove a contaminant from the system. Sorption and precipitation are generally the dominant mechanisms retarding contaminant transport (1). However, as discussed later, complexation can be dominant in landfills.

One model commonly used in chemical and environmental engineering that is often used to estimate leachate concentrations as a function of time and space, $C(t,x,y,z)$, is

$$\left(1+\frac{\rho_B}{\alpha}K_D\right)\frac{dC}{dt} = \nabla\bullet D\nabla C - \vec{v}\bullet\nabla C + R \qquad (1)$$

If the diffusion coefficient is constant and a single dimension is used, this becomes

$$\left(1+\frac{\rho_B}{\alpha}K_D\right)\frac{dC}{dt} = D\frac{\partial^2 C}{\partial x^2} - v_x\frac{\partial C}{\partial x} + R \qquad (2)$$

where C concentration, M/L^3
D diffusion (or dispersion) coefficient, L^2/t
K_D leachate/solid distribution coefficient, L^3/M
R reaction rate, M/L^3/t
x spatial variable, L
t time
v velocity, L/t
α solid porosity
ρ_B solid bulk density, M/L^3

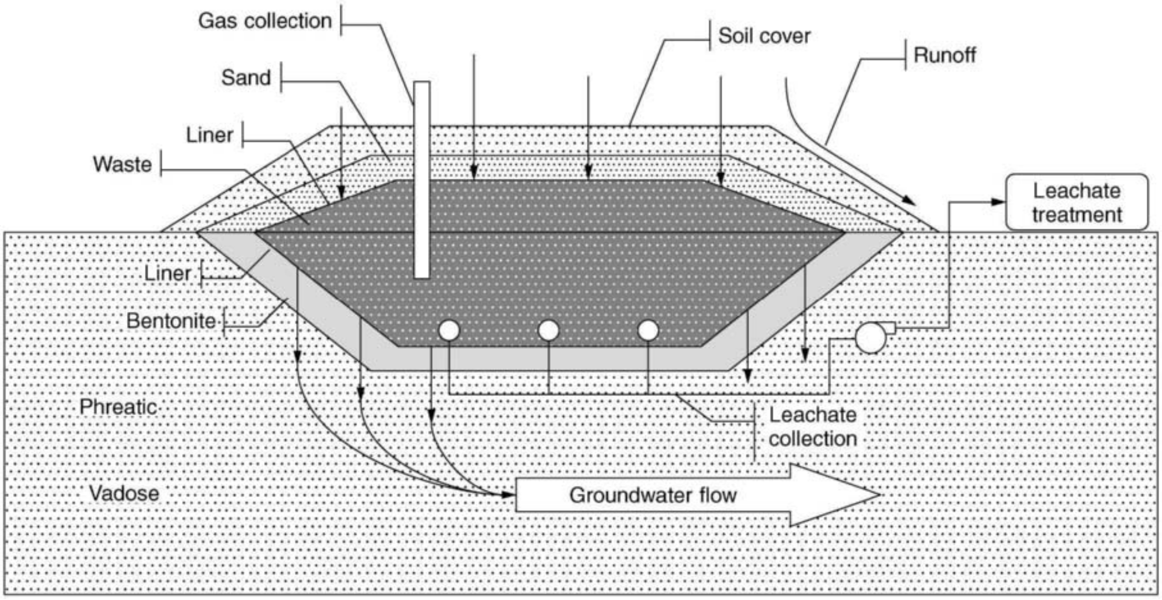

Figure 1. Schematic of landfill system.

The term on the left is the rate concentration change with time modified by the retardation based on sorption. The distribution coefficient for organic compounds can be estimated from the organic carbon content of the waste (2) but is likely to require measurement for metals. The first term on the right is the diffusion and/or dispersion. The second term on the right is the advection and the third term is the chemical reaction term. Equation 2 does not account for all phenomena. However, lumped sum parameters can be used, for example, the reaction rate R can include several types of generation and degradation terms lumped together. This equation or a variation thereof is normally solved numerically using appropriate boundary conditions.

Mathematical models to simulate the expected fate and transport of leachate metals are frequently used in groundwater and waste piles. Travers and Sharp–Hansen provide a thorough review of pertinent factors (3). Each facility is different, and no one generic model is appropriate in all situations. The conceptual model and data are used to select a mathematical model that accurately represents the conceptual model (4). A number of groundwater models, including saturated flow, solute transport, heat transport, fracture flow, and multiphase flow models, are summarized in van der Heijde et al. (5). A report by van der Heijde and Beljin provides detailed descriptions of 64 groundwater flow and solute transport models that were selected for use in determining wellhead protection areas (6). Reviews of groundwater flow and transport models for the unsaturated zone is presented in Oster (7). Groundwater flow and transport models are summarized by Bond and Hwang (8). Finally, Travers and Sharp-Hansen summarize models that may be applicable to problems of leachate generation and migration from landfill units (3). In an idealized situation (e.g., no complexation), metal mobility is controlled by the solubility of the species in the leachate. Precipitation is incipient when the metal exceeds its solubility, for the solubility product constant of a divalent metal,

$$K_{\mathrm{sp}} = [\mathrm{M}^{+2}][\mathrm{OH}^{-}]^2 \tag{3}$$

The precipitation of a dissolved substance may be initiated by changes in pressure, temperature, pH, concentration, or redox potential (9).

Chemical adsorption/desorption is a common mechanism that affects metal contaminant migration in soils. Solutes become attached to the solid phase by adsorption. Adsorption/desorption and precipitation/dissolution are reversible processes. However, adsorption/desorption generally occurs at a relatively rapid rate compared to precipitation reactions (1).

Oxidation and reduction (redox) reactions involve the transfer of electrons. Depending on the redox potential, metals may exist in various oxidation states that can impact their solubility. An example is iron that is very insoluble in the plus-three valence state. The valence state of iron depends on the redox potential and the pH, as shown in Eq. 4 (10):

$$4\mathrm{Fe}^{2+} + \mathrm{O}_2 + 4\mathrm{H}^+ = 4\mathrm{Fe}^{3+} + 2\mathrm{H}_2\mathrm{O} \tag{4}$$

Redox reactions and pH affect the solubility, complexing capacity, and sorptive behavior of constituents and thus control the presence and mobility of many substances in water. Microorganisms are responsible for a large proportion of redox reactions that occur in groundwater and, by inference, wastes in landfills.

Ion exchange occurs primarily at exchange sites on layered silicate clays and organic matter that have a permanent negative charge. The ability of soils and other solids to exchange cations is measured by the cation exchange capacity. The type and quantity of clay mineral present, the amount of organic matter present, and the pH of the soil or solids affect cation exchange capacity. Major cations in leachate (Ca, Mg, K, Na) usually dominate the exchange sites, resulting in little attenuation in soils of trace metals in the leachate, depending on the material selectivity and leachate metals concentrations. A smaller ion exchange effect for anions is associated with hydrous oxides. Soils typically have more negatively charged clay particles than positively charged hydrous oxides. Therefore, the transport of cations is attenuated more than the transport of anions.

Complexation reactions occur when metal ions react with inorganic anions or organic ligands to form coordination compounds. Complexation can either increase the concentration of a constituent in solution by forming soluble complex ions or decrease the concentration by forming a soluble ion complex with a solid. It is often difficult to distinguish among sorption, solid–liquid complexation, and ion exchange. Therefore, these processes are usually grouped together as one mechanism.

The biological degradation cycle of landfills normally occurs in four phases, aerobic, acetogenic, methanogenic that last on the order of 15 years, followed by a humic phase where the relatively refractory humic substances are stable. The degradation refers to organic compounds. The last phase is the phase in which no more organic degradation occurs and the remaining organic compounds consist of humic substances. The aerobic decomposition phase is short; it lasts approximately 1 month when all of the oxygen is used up (11). The second phase can last several years producing simple compounds such as fatty acids with high BOD_5 and an acidic leachate of pH approximately 5–6. The leachate contains high levels of mobile metals. Methanogenesis is slower than acetogenesis, so simple organic compounds are consumed during the third phase. Sulfates can be reduced to sulfides during the third phase that subsequently precipitate metals and hence change their speciation and reduce their mobility. The pH during the third phase is of the order of 7–8.

Recent work provides some specific examples of the way landfill practice influences speciation and mobility of metals. A plethora of information exists in the literature. Therefore, only some of the references found are cited. The general conclusion of the studies is that metal speciation and mobility are highly dependent on the landfill phase. The humic phase of the landfill is when the metals are least mobile and are speciated by chelation with humic substances. This phase may be short or long-lived, depending on the waste exposure to air. Humic oxidation can likely be minimized and/or eliminated by precluding air infiltration. The phase of a landfill can be accelerated by recycling water thus rendering metals less mobile sooner. It is also concluded that metals can exist in colloidal form and can complex with dissolved organic and/or inorganic compounds.

Several of the studies cited below correlate the speciation and mobility with chelation by humic compounds. Therefore, some discussion of these substances is needed. Humic compounds are derived from the amorphous part of dark brown to brown-colored substances in soil, lignite, peat, and brown coal produced during the humification of decaying organic materials. A conceptual model of a humic acid structure (12) is shown in Fig. 2. Figure 2 shows a possible sequestration mechanism of nickel, cadmium, and plutonium by humic acid. Humic acid/humic substance chemical structures have not yet been clearly defined. Approximate models describe humic acids as three-dimensional macromolecules whose molar mass is between 25,000 and 100,000. Their basic structure always has two distinctive features: (1) a central core with high aromaticity and (2) strong cross-linkage and peripheral functional groups linked by bridge binding (13). A substantial fraction of the mass of the humic acids is in carboxylic acid functional groups, which endow these molecules with the ability to chelate positively charged multivalent ions (Mg^{2+}, Ca^{2+}, Fe as well as hazardous metals such as Cd^{2+} and Pb^{2+}) (14). A major effect of humic acids in the speciation and mobility of metals is detoxification of heavy metals by chelation.

Organic matter degradation and hence the potential efflux of heavy metals in a long-term perspective was studied by defining some scenarios for waste deposits containing organic compounds, different longevity and functions of covers, and different water and air intrusion rates. The scenarios were based on various transport processes as well as different landfill constructions. The main conclusion is that the degradation of humic matter and hence the release of toxic metals can be substantially decreased if potential buildup of hydraulic gradients is avoided and if the landfill is located below the water surface (15). The amount of oxygen available for oxidation of the humic material and sulfides is reduced by water saturation. The assumption is that humic material and sulfides are readily oxidized and thus release the metals upon oxidation. The fifth phase of the landfill is the humic oxidation stage, if sufficient air can contact the waste and oxidation rates are fast. Additional research is needed in this area.

A conceptual model has been developed for waste deposits (16). The important processes occurring during the humic phase were quantified. The scenarios included the main mechanisms based on various transport processes as well as different landfill constructions, for example, binding capacities of sulfides and humic substances. They also included transport mechanisms by which the reactant oxygen can intrude into a deposit, sorption capacities of hydrous ferric oxides, and pH-buffering reactions. Scoping calculations have shown that the binding capacity of humic substances is sufficient to bind all toxic metals Cd, Cr, Pb, Zn, and Hg. In addition, the humics could also bind a smaller part of Ca, Fe, and Al, provided that much of the organic wastes remain as humic substances. Sulfides, on the other hand, can bind approximately twice the amount of all toxic metals. The binding capacity of hydrous ferric oxides, which can be formed by oxidation reactions during the humic phase, is estimated at three times the total content of metals that can sorb on hydrous ferric oxides. The alkalinity of the wastes is high enough to buffer the acidity produced by the oxidation of sulfides and by the degradation of organic matter, as well as that added by acid precipitation. Therefore, the main conclusion is that higher remobilization rates of heavy metals due

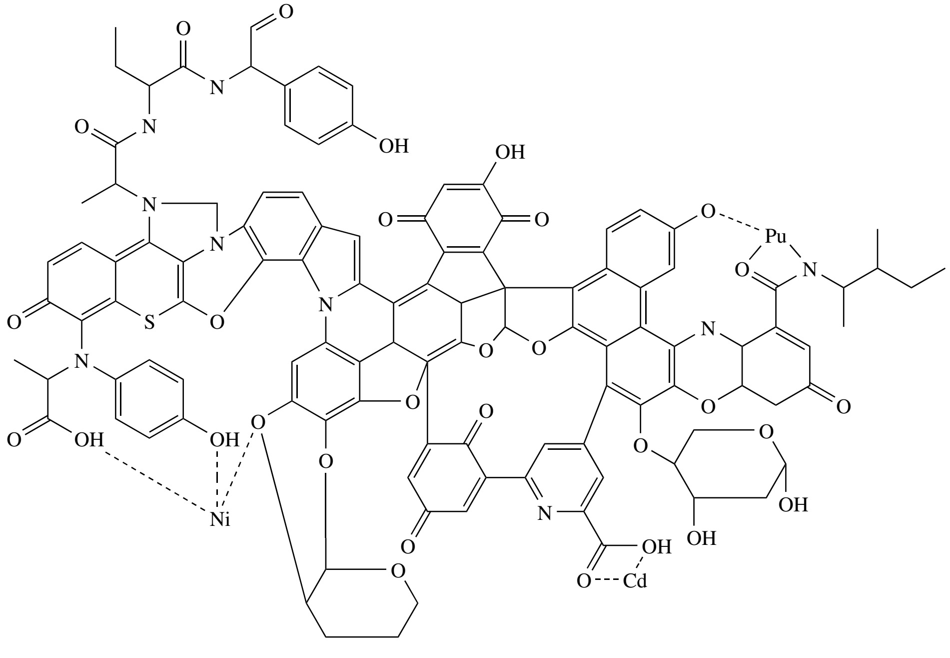

Figure 2. Model structure for humic acid.

to lowering of pH are not expected for many thousands of years.

Dissolved organic compounds can also complex with metals that increase mobility. Leachates after treatment from a landfill in Japan contain the same order as the influent (17). This was attributed to complexation with organic compounds, as treatment by elevated pH precipitation was effective when the organic compounds were removed. The ability of dissolved organic carbon (DOC) from landfill leachate-polluted groundwater to form complexes with the heavy metals cadmium (Cd), nickel (Ni), and zinc (Zn) was investigated (18). The DOC samples originated from the leachate pollution plume at Vejen Landfill, Denmark, and were studied in the original matrix with a minimum of manipulation. The results showed that DOC from landfill leachate-polluted groundwater can form complexes with Cd, Ni, and Zn and that the distribution coefficients were a factor of 2–6 lower in the presence of DOC. The migration velocity of the metals was increased by the presence of DOC but did not exceed 1.2% of the water migration velocity, indicating that the effect of DOC on the mobility of Cd, Ni, and Zn may have only minor environmental importance.

Rainfall and the recycling of leachate impact landfill leachate metal speciation and mobility. Leachate volume increases linearly with rainfall, and the leaching coefficient decreases, according to studies in Taiwan (19). Landfill bioreactor technology offers important advantages in managing and treating municipal solid waste, including accelerated waste stabilization rates, enhanced gas production, facilitated leachate management, volume reduction, and minimized long-term liability. These advantages have been documented in laboratory, pilot, and full-scale investigations. Although challenges remain in implementing the technology, bioreactor landfills are designed and operated at increasing frequency (20). Metal washout is initially one of the important mechanisms in conventional landfills, but sulfide and hydroxide precipitation followed by complexation with humic compounds are the primary mechanisms in recirculating landfills. However, metals can be remobilized, once humic oxidation starts. Recycling of leachate and/or addition of water through the waste accelerates stabilization. Water addition is a means of decreasing the degradation time span and thus accelerating metal stability. The water helps spread nutrients and bacteria to other parts of the landfill (21) and provides a medium conducive to biochemical reactions.

The physical forms of the species are also important in considering leachate metal speciation and mobility. Some of the metals can exist in colloidal form as, very small particles that are not easily filterable. This can have a large impact on characterization and leachate treatment processes. According to research on Danish landfills, metals exist in colloidal form. The dominant part of the colloidal material was in the smallest colloidal size fraction (0.001–0.40 mm); total organic carbon (TOC) was one of the major components of the colloidal mass. Si, Al, Ca, Fe, and S were also found in the colloidal fractions. This indicated that clay type particles, organic matter, and precipitates (most likely sulfides) were present in the leachate samples (22). The behavior of the colloidal mass can be markedly different from a dissolved metal speciation and needs to be accounted for in leachate treatment systems. In another similar study, anaerobic leachate-polluted groundwater samples were collected downgradient of a Danish landfill (23). The groundwater samples were spiked with heavy metals at concentrations within the range of metal concentrations found in landfill leachates and divided under anaerobic conditions into colloidal (screen-filtration and cross-flow ultrafiltration), organic (anion-exchange), and dissolved inorganic species of the heavy metals. The distribution of the heavy metals among the different size fractions showed that a substantial, but highly varying part of the heavy metals was associated with the colloidal fractions. These colloidally bound metals were, furthermore, associated primarily with the organic fraction, except for Zn, which was associated mainly with the inorganic fractions. Dissolved Cd, Cu, and Pb were mainly associated with dissolved organic carbon; Ni was also present as carbonate complexes and Zn as carbonate complexes and free divalent Zn. The results indicate that heavy metals in leachate-polluted groundwater are strongly associated with small-size colloidal matter and organic molecules.

BIBLIOGRAPHY

1. US EPA. (1987). *Background Document on Bottom Liner Performance in Double-Lined Landfills and Surface Impoundments*. EPA/530/SW-87/013. US EPA, Washington, DC. NTIS PB-87-182291.
2. Hemond, H.F. and Fechner, J.F. (1994). *Chemical Fate and Transport in the Environment*. Academic Press, San Dicego, CA.
3. Travers, C.L. and Sharp-Hansen, S. (1991). *Leachate Generation and Migration at Subtitle D Facilities: A Summary and Review of Processes and Mathematical Models*. Prepared for US EPA, Environmental Research Laboratory, Athens, GA.
4. Donigian, A.S. and Rao, P.S.C. (1990). Selection, application, and validation of environmental models. *Proc. Int. Symp. Water Quality Modeling Agric. Non-Point Sources*. Part 2.
5. van der Heijde, P.K., Bachmat, Y., Bredehoeft, J., Andrews, B., Holtz, D., and Sebastian, S. (1985). *Groundwater Management: The Use of Numerical Models*. American Geophysical Union, Washington, DC.
6. van der Heijde, P.K. and Beljin, M.S. (1988). *Model Assessment for Delineating Wellhead Protection Areas*. EPA-440/6-88-002. US EPA, Washington, DC.
7. Oster, P.A. (1982). *Review of Ground-Water Flow and Transport Models in the Unsaturated Zone, NUREG/CR-2917, PNL-4427*. Pacific Northwest Laboratory, Richland, WA.
8. Bond, F. and Hwang, S. (1988). *Selection Criteria for Mathematical Models Used in Exposure Assessments: Groundwater Models*. EPA/600/8-88/075. US EPA, Washington, DC.
9. Aller, L., Bennett, T., Lehr, J.H., Petty, R.J., and Hackett, G. (1987). *A Standardized System for Evaluation of Ground Water Pollution Potential Using Hydrogeologic Settings*. EPA 600/2-87-035. Kerr Environmental Research Lab, US EPA, Ada, OK.
10. Tchobanoglous, G. and Schroeder, E.D. (1987). *Water Quality*. Addison-Wesley, Reading, MA.

11. Irene, M. and Lo, C. (1996). Characteristics and treatment of leachates from domestic landfills. *Environ. Int.* **22**(4): 433–442.
12. Adaptation from a structure by Kickuth on a German Website, http://www.pharmawerk-weinboehla.de/english/humine/humin2.html.
13. http://www.pharmawerk-weinboehla.de/english/humine/humin2.html.
14. http://www.thekrib.com/Chemistry/humic.html.
15. Bozkurt, S., Moreno, L., and Neretnieks, I. (1999). Long-term fate of organics in waste deposits and its effect on metal release. *Sci. Total Environ.* **228**: 135–152.
16. Bozkurt, S., Moreno, L., and Neretnieks, I. (2000). Long-term processes in waste deposits. *Sci. Total Environ.* **250**: 101–121.
17. Urase, T. et al. (1997). Effect of high concentration of organic and inorganic matter in landfill leachate on the treatment of heavy metals in very low concentration level. *Water Sci. Technol.* **36**(12).
18. Christensen, J.B., Jensen, D.L., and Christensen, T.H. (1996). Effect of dissolved organic carbon on the mobility of cadmium, nickel and zinc in leachate polluted groundwater. *Water Res.* **30**(12): 3037–3049.
19. Paris, H.C. (1996). Assessment of leachates from sanitary landfills: impact of age, rainfall, and treatment. *Environ. Int.* **22**(2).
20. Reinhart, D.R. and Al-Yousfi, B.A. (1996). The impact of leachate recirculation on municipal solid waste landfill operating characteristics. *Waste Manage Res.* **14**(4): 1.
21. Mosher, F.A., McBean, E.A., Crutcher, A.J., and MacDonald, N. (1997). Leachate recirculation for rapid stabilization of landfills. *WQI*.
22. Jensen, D.L. and Christensen, T.H. (1999). Colloidal and dissolved metals in leachates from four danish landfills. *Water Res.* **33**(9): 2139–2147.
23. Jensen, D.L., Ledin, A., and Christensen, T.H. (1999). Speciation of heavy metals in landfills-leachate polluted groundwater. *Water Res.* **33**(11): 2642–2650.

MICROBIAL FOAMING IN THE ACTIVATED SLUDGE PROCESS

MICHAEL H. GERARDI
Linden, Pennsylvania

The production and accumulation of microbial (bacterial) foam in activated sludge processes is common and often contributes to increased operational costs and permit violations for total suspended solids (TSS), biochemical oxygen demand (BOD), and floating solids. There are several significant operational conditions that permit foam production in the aeration tank and its accumulation throughout numerous tanks in the treatment process. Each condition allows for the production of foam with its own characteristic texture and color (Table 1). Operational conditions responsible for microbial foam production include the undesired growth of foam-producing filamentous organisms, the presence of a nutrient deficiency, increasing sludge age, and undesired zoogloeal growth or viscous floc.

Table 1. Operational Conditions Associated with the Production of Microbial Foam

Operational Condition	Microbial Foam
Foam-producing filamentous organisms	Viscous chocolate-brown
Nutrient deficiency at a young sludge age	Billowy white
Nutrient deficiency at an old sludge age	Greasy gray
Sludge aging (young to old)	Billowy white to crisp white to crisp brown to viscous dark brown
Zoogloeal growth	Billowy white

Foam consists of entrapped air or gases beneath a thin layer of solids or biological secretions. Gases entrapped in foam consist of those released during the aerobic and anaerobic degradation of carbonaceous wastes. These gases include carbon dioxide (CO_2), molecular nitrogen (N_2), and nitrous oxide (N_2O).

Microbial foam is produced in the aeration tank and is discharged from the aeration tank to other treatment units such as the secondary clarifier, thickener, digester, dewatering facilities, and chlorine contact tank. In these units the accumulation of foam often results in increased operational costs. As foam is discharged from the aeration tank to other treatment units, the air and gases escape, and the foam collapses. The collapsed foam is sometimes referred to as scum.

FOAM-PRODUCING FILAMENTOUS ORGANISMS

In order of occurrence in activated sludge processes in North America, foam-producing filamentous organisms include nocardioforms, *Microthrix parvicella*, and type 1863. These organisms produce viscous, chocolate-brown foam.

Foam production by filamentous organisms is due to the presence of a large population of living organisms and a large population of dead organisms. The living organisms release insoluble lipids that coat floc particles and entrap air and gases. Dead organisms undergo autolysis and release surfactants (biosurfactants) that reduce the surface tension of the wastewater and permit microbial foaming. Surfactants released include ammonium ions (NH_4^+) and fatty acids. Because dead organisms release surfactants, control measures that destroy large numbers of foam-producing filamentous organisms over a relatively short time period often result in the rapid production of a copious quantity of foam.

The rapid and undesired growth of each foam-producing, filamentous organism can be associated with specific operational conditions (Table 2). By monitoring and preventing these conditions, the undesired growth of these organisms and their production of foam can be reduced.

Nocardioforms are a specialized group of bacteria that are known as actinomycetes. These organisms display some growth characteristics such as true branching that are found in the fungi. Nocardioforms and related

Table 2. Operational Conditions Associated with the Undesired Growth of Foam-Producing Filamentous Organisms

Operational Condition	Filamentous Organism		
	Nocardioforms	*Microthrix parvicella*	Type 1863
High MCRT (>10 d)		X	
Fats, oils, and grease	X	X	X
High pH (>8.0)		X	
Low DO and high MCRT		X	
Low F/M (<0.05)	X	X	
Low MCRT and high F/M			X
Low nitrogen or phosphorus	X		
Low pH (<6.5)	X		
Readily degradable substrates	X		
Slowly degradable substrates	X	X	
Winter proliferation		X	

genera that are most often associated with foam production in activated processes include *Actinomadura, Arthrobacter, Corynebacterium, Micromonospora*, and *Nocardia*. *Nocardia* is the most commonly observed foam-producing actinomycetes, and the most frequently reported species of *Nocardia* include *N. amarae, N. asteroides, N. caviae, N. pinesis*, and *N. rhodochrus*. These organisms are relatively short (<50 μm), highly branched, and gram-positive. Nocardioforms can be found in the foam and mixed liquor within floc particles, extending into the bulk solution from the perimeter of the floc particles, and free-floating between the floc particles. Usually, nocardioforms as well as *Microthrix parvicella* and type 1863 are found in higher concentration in the foam than in the mixed liquor.

Because foam-producing filamentous organisms are found in high concentrations in the foam, the foam represents a source of "reseeding" of the mixed liquor with filamentous organisms. Therefore, treatment of the foam should be addressed when attempting to control the growth of foam-producing filamentous organisms (Table 3).

Microthrix parvicella is a nonbranched, relatively short (<200 μm), gram-positive filamentous organism. Intracellular granules in the filamentous organism may appear as a string of "beads." *Microthrix parvicella* usually is found in tangled mats in the floc particle or bulk solution.

Table 3. Treatment Measures for the Control of Filamentous Organism Foam

- Collapsing foam with the application of cationic polymer
- Collapsing foam with the application of effluent through "bib" sprinklers
- Collapsing foam with nonpetroleum-based defoaming agent
- Digesting foam with the application of bacterial cultures that contain the lipid-digesting enzyme lipase
- Physically removing the foam
- Spraying foam with a 10–15% sodium hypochlorite solution and spraying foam with effluent after 2–3 h of foam exposure to the sodium hypochlorite solution

Blooms of *Microthrix parvicella* often are associated with cold wastewater temperatures (winter proliferation). The proliferation of *Microthrix parvicella* during cold wastewater temperature may be due to the congealing of fats, oils, and grease. These compounds are preferred substrates for this filamentous organism.

Type 1863 also is a nonbranched, relatively short (<50 μm), gram-negative filamentous organism. Type 1863 is found free floating in the bulk solution and extending into the bulk solution from the perimeter of the floc particles. The filamentous organism possesses "chains" of rod-shaped cells that are constricted at each end.

NUTRIENT DEFICIENCY

Nutrient deficiencies are commonly observed in activated sludge processes and usually are due to the presence of nutrient-deficient industrial wastewater (Table 4). Nutrients that are most often deficient are nitrogen and phosphorus.

During a nutrient deficiency, soluble substrate that is absorbed by bacterial cells in floc particles cannot be degraded. The nondegraded food is converted by bacterial cells to insoluble polysaccharides and stored outside the bacterial cells. The polysaccharides are less dense than water and hinder the settleability of floc particles. Often, polysaccharides are deposited in the floc channels that permit the moving of water, air, and gases through the floc particle. When these channels become heavily laden with polysaccharides, air and gases are entrapped in the channels and desired settleability continues to deteriorate and foam appears on the surface of the aeration tank.

Foam produced during a nutrient deficiency is billowy white at a young sludge age and greasy gray at an old sludge age. The difference in texture and color of nutrient-deficient foam is due to the accumulation of oils in the floc particles. Young bacterial cells produce relatively little oil that accumulates in floc particles as compared to old bacterial cells. The transfer of oil to the foam from the floc

Table 4. Nutrient-Deficient Industrial Wastewaters for Nitrogen or Phosphorus

Industrial Wastewater	Nutrient Deficient	
	Nitrogen	Phosphorus
Bakery	X	
Beverage—alcoholic and soda drink	X	X
Citrus	X	
Chemical		X
Coffee	X	
Corn	X	
Cotton kerning	X	
Dairy—milk		X
Dairy—cottage cheese	X	
Food processing	X	X
Fruit and vegetable	X	X
Leather tanning		X
Petroleum refining		X
Pharmaceutical		X
Pulp and paper	X	X
Textile	X	

particle results in the production of greasy gray foam as the sludge age increases.

A nutrient deficiency within an activated sludge process can be corrected by identifying nutrient-deficient industrial wastewaters and ensuring that adequate nutrients are added to these wastewaters. Also, nutrients can be added to the activated sludge process as needed. Chemical compounds that release ammonium ions (NH_4^+) or orthophosphate (HPO_4^{2-}) can be added to the mixed liquor influent or primary clarifier effluent. Appropriate recycle streams also can be discharged to the activated sludge process when needed, if these streams contain adequate quantities of ammonium ions or orthophosphate and relatively low quantities of soluble wastes. Recycle streams include decant from digesters, centrate, and filtrate.

INCREASING SLUDGE AGE

Several types of foam are produced through changes in microbial activity with increasing sludge age. Billowy white foam is produced at a young sludge age when the bacterial population or mixed liquor suspended solids (MLSS) value is relatively small, that is, <1000 mg/L. This small population of bacteria lacks sufficient enzymatic ability to adequately degrade the surfactants that are discharged to the activated sludge process. The nondegraded surfactants produce billowy white foam.

With increasing sludge age the bacterial population or MLSS value increases. This population has sufficient enzymatic ability to adequately degrade the surfactants. With degradation of the surfactants, the foam becomes crisp white.

As the bacterial population (or MLSS) continues to age, large quantities of oils secreted by the bacteria accumulate in floc particles. The color of the oils darkens crisp white foam to crisp brown foam.

Finally, as the bacterial population becomes very old, large numbers of slow-growing filamentous organisms such as nocardioforms and *Microthrix parvicella* may proliferate. The growth of these lipid-producing organisms contributes to the production of viscous dark brown foam.

Undesired quantities of foam as a result of young or old sludge ages can be corrected by decreasing or increasing the sludge age by regulating sludge wasting rates. Sludge wasting rates should be uniform over as long a period of time as possible in order to prevent the development of pockets of young bacterial growth and old bacterial growth. If these pockets of growth do develop, concentric circles of light foam and dark foam can be observed on the surface of the aeration tank when aerators are turned off.

ZOOGLOEAL GROWTH

Zoogloeal growth or viscous floc is the rapid and undesired proliferation of floc-forming bacteria such as *Zoogloea ramigera*. This growth is associated with the production of large quantities of insoluble gelatinous material that entrap air and gases. The gelatinous material and entrapped air and gases result in the production of billowy white foam.

The occurrence of zoogloeal growth in activated sludge processes is due to high MCRT, long HRT, nutrient deficiency, organic acids, and significant changes in F/M. Zoogloeal growth and its foam can be controlled by exposing the growth to anoxic periods.

The texture and color of foam may vary due to changes in operational conditions. Generally, foam becomes more billowy and lighter in color with decreasing sludge age and more viscous and darker in color with increasing sludge age. Also, several operational conditions for the production of foam can occur at the same time, resulting in a mixture of textures and colors. For example, nutrient-deficient foam and filamentous organism foam can occur simultaneously.

Additional foaming conditions that often occur in activated sludge processes include surfactant foam, alkalinity foam, and polymer foam. If surfactants are not quickly degraded in the aeration tank, surfactants change the surface tension of the wastewater and produce billowy white foam. Also, an increase in alkalinity in the wastewater changes the surface tension of the wastewater, resulting in the production of billowy white foam.

If cationic polymers containing amino groups ($-NH_2$) are applied in an excess quantity for sludge thickening or sludge dewatering, and the excess quantity of polymer enters the aeration tank, billowy white foam is produced. When the polymer is degraded in the aeration tank, ammonium ions are produced from the release of amino groups. The ammonium ions represent an increase in alkalinity and a change in the surface tension of the wastewater. The change in the surface tension of the wastewater results in the production of billowy white foam.

INTRODUCTION TO WASTEWATER MODELING AND TREATMENT PLANT DESIGN

David L. Russell
Global Environmental Operations, Inc.
Lilburn, Georgia

INTRODUCTION

There are many ways of approaching the design of a wastewater treatment plant. Of the many methods, the field of practice generally uses three or four: design codes, rules of thumb, simple Monod calculations; and rigorous mathematical modeling. Most states still subscribe to variants of the first two methods because they work and are written in code documents. Depending on the backgrounds of designers, they may or may not have been exposed to the Monod equation and/or mathematically rigorous approaches to development of a static or dynamic model.

We confine this article to discussing readily available activated sludge and related models. Anaerobic digestion or film type of biological treatment systems is not discussed because the modeling effort for film modeling is not well enough defined to permit accurate modeling, and anaerobic modeling is still relatively new and has not become widely available. Another reason is that many of the models would require a full book to summarize adequately, and their use and calibration are often the subject of Ph.D. theses.

The older design code approach is often characterized by reference to individual state and national design codes such as the Ten States Standards. This approach is acceptable but produces a design which is between 30 and 50% in excess of that required and leads to sloppy and wasteful operation.

The rationale model is a group of codified experience based principally on food to microorganism ratio, mixed liquor suspended solids (MLSS), and a number of rules of thumb about the amount of air to supply per pound of waste treated, the quantity of cells produced per pound of biochemical oxygen demand treated, and hydraulic retention time.

For example, one could find rationale in the literature and the codes for designing a treatment plant with an F/M ratio of 0.4 to 0.05, a retention time between 12 and 30 hours, an MLSS of between 1000 and 4000 mg/L, and an oxygen transfer rate of 2.2 lb oxygen supplied per horsepower/hour per pound of BOD removed. That plant would probably work well, but under the ranges of the assumptions given, one could wind up with tankage that varies by a factor of 10 times or more. Still, no discussion of the rules for designing a clarifier system to remove solids has occured. The clarifier design can vary over a smaller range of perhaps 4:1. There is a wide latitude in what works, but the principal objective is to get the best combination of size and efficiency for the cost.

To varying degrees, the individual and company design groups have computerized their designs on spreadsheets (principally Excel®) or in limited versions for numerical calculators. These programs are adequate as far as they go, but designing an expensive and energy-intensive facility should consider the dynamic behavior of the wastewater and the performance of the plant.

It can be argued, because I have done so, that repeated running of the static models at varying conditions is an adequate substitution for dynamic modeling. At one time I believed that. However, the static models available are just not up to the performance of the dynamic models, and they do not address the different types of conditions and alternatives that can be evaluated using dynamic models.

Modeling the operation of a wastewater treatment plant is neither easy nor simple. Rather, it requires a concerted effort to acquaint oneself with the process of model development, the development of a mindset, and a willingness to research the somewhat arcane area of mathematical modeling formulations and development.

Modeling a wastewater treatment plant requires mastery of the elements of modeling and a basic understanding of the fundamentals of model development and chemistry. It also requires understanding the limitations of the model(s) used and their interactions.

STATIC AND DYNAMIC MODELS AVAILABLE

A number of wastewater models are currently available. The principal static models are available on the World Wide Web or from individuals via e-mail:

The Steady© model was developed by Luis Aburto Garnica and Gerald E. Speitel, Jr. It is available from the University of Texas via its website. It solves the activated sludge equations by an iterative approach. The graphics are limited, and the model is also very limited, but it is free. It can be obtained from: http://www.ce.utexas.edu/prof/speitel/steady/steady.htm.

Dr. William McHarg developed a spreadsheet model that is also free. The documentation is available online from Dr. McHarg, and a copy of the spreadsheet is available from him. The documentation is very good, and it is available at: http://www.wideopenwest.com/~bmcharg/bioxmdl/bioxmdl.htm.

The Wisconsin Department of Natural Resources has developed a model for the activated sludge process, named ENBIR. The ENBIR model is available from the following web address: http://www.dnr.state.wi.us/org/water/wm/ww/biophos/enbir.htm. The model is DOS-based and may not run under all versions of Windows. The documentation and theory development are available in a Power Point® slide from the following address: http://www.dnr.state.wi.us/org/water/wm/ww/biophos/remproc.ppt. The model was developed by Dr. Jim Park in 1996 to address Wisconsin's concerns about cold weather effects on wastewater treatment plants and the decrease of nutrient removals during cold weather. The documentation is reasonably good but does not consider all the ramifications of design and has limited ability to handle industrial wastes and nonstandard conditions.

Dynamic models come in all shapes and sizes. As far as can be ascertained, the only two free dynamic models are from Clemson University, called simulation of single sludge Processes (SSSP), and the corresponding program from University of Uppsala, Sweden, called Java activated sludge process simulator (JASS).

The SSSP model was developed by Dr. Les Grady and one of his graduate students around 1986. It is a DOS-based model. It was one of the earliest works that incorporated dynamic modeling and is now available free from Clemson. As far as can be ascertained, it was developed about the same time as the ASM1 model (see below), and it can model a multitank activated sludge plant. Steven Birstrip was one of Dr. Grady's graduate students, and he developed the DOS code for the model. The model is relatively easy to use, has good documentation, and can be learned and used in a matter of hours. It is limited in that it does not model the clarifier portion of the wastewater treatment plant well, but other than that, it is functional. The model can be obtained from the following website at the Rich Environmental Research Center of Clemson University: http://www.ces.clemson.edu/ees/sssp/index.htm.

The JASS model can be found at http://www.syscon.uu.se/~psa/. It is free, but it runs on their network and cannot be downloaded to a private personal computer. The graphics are unsophisticated, and the modeling is based on the ASM1 model. In times past, when the site was under development, some of the comments on the bulletin board accompanying the site indicated that there were mistakes in the program that were being fixed. The site is not necessarily convenient, nor is it reliable, and it lacks the power and flexibility of some of the commercial software packages.

Based on my own research, I prefer the SSSP program to the JASS program. This does not say that others cannot and have not developed free spreadsheet and other dynamic models, but none are as well documented as the SSSP Model.

For those of you who have the Mathcad® program, there is an activated sludge simulator that can be freely downloaded, but can be used only with Mathcad®.

The next group of programs is available for operation on mainframe computers where the base operating program is Matlab. Matlab is designed for solving large groups of complex equations, and a number of models of the activated sludge programs are available on Matlab. Matlab is not for everyone and is primarily a research tool. A single site license for Matlab costs in excess of $2500 per year.

The last group of programs is commercially available. They are made by several different groups. The principal programs are BIOWIN, GPS-X, SIMBA, STOAT, and WEST. All models have good graphical interfaces. Some are slightly easier to use than others. Depending on the desires and personal tastes of the user, one particular program may suit a user need better than others. The pricing structure and the modeling capabilities are somewhat similar but differ in outcome and computation time, graphical interface, and ease of use. All products allow the user to develop plant configurations of any type or size and readily make changes in the configuration.

Some programs offer a variety of models and configurations; others offer only a single model with adapted differences for enhanced performance. The real advantage of dynamic modeling lies in its ability to model an entire plant and show the designer or operator how the system operates. In all programs, it is possible to come up with an almost infinite variety of configurations and plant arrangements. For example, there are at least six dynamic models of clarifiers. Some models are only for primary clarifiers, others are for secondary clarifiers. When combined with the many variants of activated sludge and other process models, there is almost an infinite variety of plant configurations one can develop for modeling an activated sludge process before one has to consider tank and reactor dynamics and before one considers the various control systems and parameters. The result can often be quite overwhelming because the models construct and then solve more than 50 differential equations for a simple plant configuration.

The differences in the software are qualitative: power, flexibility, user base, development, features, and pricing structure. There is also a trend toward using a simulator software as a supervisory control and data acquisition system (SCADA), but, at this time, only WEST offers that capability.

The author of this article is affiliated with HEMMIS, the producer of WEST software, and, accordingly, all descriptions of competitors could be construed as commercial criticism or endorsements, but are very general in nature. Nothing should be construed as either an endorsement or as a criticism of a particular software product.

BIOWIN® is produced by EnviroSim Associates in Canada. The web address is http://www.envirosim.com/products/. It is very popular and has a reasonable interface.

GPS-X has a very good graphical icon interface, and it is reasonable to work with. It is produced by Hydromantis in Toronto, Canada. The Web address is http://www.hydromantis.com.

SIMBA is a German product produced by Ifak System GmbH, Schleinufer 11, D-39104 Magdeburg, Germany. SIMBA 4.0 is shipped together with the Matlab®/Simulink™ software of MathWorks Inc.; it runs under the Windows 9x, 2000, and NT 4.0 operating systems.

STOAT is produced by WRc Plc Ltd. in England. It is a very intuitive program, and one can quickly set up a configuration, but the targets on the connecting graphical elements are a bit small. The completed program creates an output to an Excel® file, and all the calculations are done at one time. The program is mathematically rigorous and good to use. It was designed by engineers, so it is quite intuitive, and the documentation is very good. Because it does the computation at one time, it is relatively quick compared with other programs. The Internet link for information on the WRc program is http://www.wrcplc.co.uk/asp/wastewater_tmnt.asp.

WEST® is the product of HEMMIS in Kortrijk, Belgium. It is an extremely rigorous and powerful program and has good graphics. Several features of WEST® include the ability to adjust the modeling parameters (such as oxygen uptake rate) on the fly; to adjust flow, control, and other parameters during the run; and instantaneously see the output. The program has a wide variety of models and options, including all the IWA models, plus others that have been rigorously tested by the Biomath Department at the University of Ghent. On the HEMMIS website, there is a free download of a sample of WEST®. The website is http://www.hemmis.com.

A word about pricing. Prices vary significantly with user needs. Base prices start at somewhere around $3,000 and go up to about $20,000 or higher depending on the user needs and the package. The cost of an adequate system for modeling a variety of conditions is between $10,000 and $15,000. Do not be fooled by different marketing offers, and make sure that the program you buy has the features you need and want. Some sites provide a simple matrix of programs at a modest to low price, and then if you want more complexity and power, you have to buy advanced models at a much higher price. Plan on a $10,000–$22,000 price tag for a complete system.

IWA

In the software packages discussed, all models have been developed by, or are based on, work performed by the International Water Association (formerly the International Water Quality Association) in London. The oldest of the models is Activated Sludge Model Number 1 (ASM1). The model formulation and description are presented by the IWA, and anyone who wishes can take the documentation and prepare an individual version of the model. If one is planning on doing any modeling, it is imperative that one has the IWA model documentation.

The available models are:

Activated Sludge Model Number 1 (ASM1), the oldest model and the most widely accepted. It allows computation of nitrification and denitrification.

Activated Sludge Model Numbers 2 and 2d. These models are similar but not identical to ASM1. The models were developed to consider nitrification, denitrification, and phosphorous uptake and release by the cells. The difference between Model 2 and Model 2d is that ASM2d allows for chemical precipitation and removal as well as for uptake and removal by the biomass.

Activated Sludge Model Number 3 is a BOD-based model that differs in formulation from the other three models. It has been tested and produces reliable results, but many of the experienced modelers do not like the general formulation of the model, and to date, it has not gained wide acceptance. It is handy because it is a BOD-based model, whereas all other models are COD based. In the United States, where the standards are BOD based, it is a bit easier to use because the model outputs can be used directly and do not have to be converted from COD back into BOD.

Anaerobic Digestion Model Number 1 is the newest of the IWA models. It has been released relatively recently and has not yet gained wide acceptance.

OTHER DYNAMIC MODELS AVAILABLE

This discussion has been primarily about modeling an entire plant, so a brief digression is in order. Most of the IWA models deal with the aeration portion of the plant and do not deal with the clarifier or other portions of the plant. Of specific interest are the associated or "ancillary" programs that often make up a plant configuration. These include the selection of primary and secondary clarifiers, sludge thickeners, dissolved air flotation units, centrifuges, filters, filter presses, disinfection units, and various types of headworks such as degritters and equalization systems. A word about these models and their characterization is necessary.

First, degritters and some headworks devices are nice showy pieces for modeling but do not really affect the plant modeling effort very greatly because we, as engineers and scientists, often do not characterize the grit and sand loads in the wastewater stream for two reasons. First, grit and sand often fall to the bottom of a pipe and are not sampled or characterized. Second, they are not characterized in models because they represent inert solids and the driving models for wastewater treatment are mostly COD based, not TSS based. In many systems, they represent a distraction as they add some numbers but affect the outcome by less than 5%.

Sand does not have any measurable COD, and most degritters are designed to remove particles of sand and rocks. The removal is proportional to the difference in settling velocities by Stokes' law and is proportional to the difference between the fluid and the particle density. Organic particles often have a specific gravity of 1.01–1.15; sand and rock have SG of 2.65. So, a model designed to remove sand may look impressive, but it provides only a minimal removal of the organic loading for the dynamic modeling of the wastewater treatment plant.

Clarifier models are for primary or secondary treatment. To some extent, they may be the same model with slightly different applications or literature on which the model was built. For example, the Lessard–Beck primary clarifier model (1) is similar in many aspects to the Takacs secondary clarifier model (2) and the Mars–Libelli clarifier model. The similarities include the manner in which the mass balance in the clarifier is analyzed, the location of the inlet port, and the approach to settling versus upward bulk flow. The clarifier is analyzed in ten layers in all models, and the mass balance is taken across the layers. The solids settle against the rising bulk velocity of the fluid and, as they reach the lower levels of the clarifier, depending on the model used, they collect and compress and generally follow the research work of Vesilind on sludge accumulation and compression (3).

When the sludge reaches the bottom of the clarifier, it is drawn off by different methods. Lessard–Beck uses a scour coefficient for primary clarifiers that could just as easily be adapted to a secondary clarifier. Mars–Libelli uses a slightly different configuration, and Takacs still a different configuration for sludge withdrawal and thickening. The point is that the model assumptions do make a difference in behavior. Where thickening behavior is to be investigated, the Mars–Libelli model is often preferred. For general behavior, the Takacs model is preferred, and for primary treatment, use the Lessard–Beck Model.

Where sludge thickening and centrifugal processing are involved, it is well to look carefully at the equations. The modeling assumptions often come down to a complex set of equations and elaborate calculations with a very interesting simplifying assumption. For example, several of the dissolved air flotation system and centrifugal models look very complex, but finally resolve down to a "stickiness factor" or a collection or efficiency factor, often experimentally determined, which controls the degree of separation. After all the trouble in the calculations, one finally has a simple constant on which the removal hangs, and that can be simulated by a single line equation or a simple clarifier model without any difficulty.

COD, BOD, and Modeling

BOD is a common but useless parameter. Unfortunately, we use it for much of our historical data and much of the permit work in the United States and elsewhere. It is useless because it is so unreliable. A quick check of Standard Methods indicates that the BOD test has a variability of around 20% at a 200 mg/L level. At lower levels, it is so unreliable that the issue of using it for permits is laughable, which makes for bad control and even worse permitting. Single-digit BOD permit values may be unrealistic, but, unfortunately, they are finding their way into the permit structure in some watersheds.

Another reason that BOD is useless as a control parameter is that the test takes too long for the results. Given that most tests take 5 days or 20 days (ultimate BOD), we still have to adjust the BOD test to prevent oxidation of ammonia and make sure that the seed is acclimatized. Even under extreme conditions, the retention time of wastewater in a treatment plant, except for very large lagoons, is seldom over 48 hours and is often between 4 and 24 hours; a standard activated sludge plant takes 12 hours.

As a control or performance parameter, it is also useless. Hoover, Jacewizc, and Porges made this appropriate comment about the BOD test in 1953 (4):

> The BOD test is paradoxical. It is the basis of all regulatory actions and is used routinely in almost all control and research studies on sewage and industrial waste treatment. It has been the subject of a tremendous amount of research, yet, no one appears to consider it adequately understood or well adapted to his own work.

Question: How do you control a plant when you can only find out what happened 5 days later?

Answer: You cannot, which is one reason why many plants are not operated but are maintained and marginally operated. There is a fundamental lack of understanding and appreciation of the basics underlying the operation and the value and the role of the operator. When an operator only finds out what happened to his plant a week later, there is no positive feedback. Dissolved oxygen and ammonia and VSS/TSS are more important control parameters. But we cannot run a plant intelligently on those parameters alone.

There are many other reasons for a lack of operation as well, but fundamentally underlying all of them is the idea that we cannot control the process, and it is a series of operator experience and guesses that get the plant through rather than any intelligent control decisions. About the only reason that an operator will get out of the air-conditioned office on a hot Friday afternoon is if the air conditioner breaks, the plant is in crisis, or the office catches fire. And who can blame them. We have set up our systems with so much slop that a typical design has between 30% and 50% excess capacity before any allowance for expansion or population growth (5).

This discussion is about how to get better process control and how to set up and evaluate a plant for this control and evaluation.

Relating BOD and COD

Both BOD and COD are measures of the amount of carbon in wastewater. There are two types of COD in use, permanganate and dichromate. Permanganate is a weaker oxidizer and gives lower results than dichromate COD. The U.S. standard is dichromate oxidation. If you apply European or foreign COD data, you have to make sure that you know the basis for the test.

For many municipal wastewaters, the relationships in Table 1 apply (6).

Table 1. Various Indicator Strengths

	Concentration		
	Weak	Medium	Strong
TSS	350	720	1200
TDS	250	500	850
Fixed TDS	145	300	525
Volatile TDS	105	200	325
BOD-5	110	220	400
COD	250	500	1000
TOC	80	160	290
TKN	20	40	85
NH_3	12	25	50
P-Total	4	8	15
Organic P	1	3	5
Inorganic P	3	5	10

The BOD does not measure ammonia uptake because we inhibit the test to prevent ammonia demand. COD does not measure ammonia either.

Various sources have estimated the relationships between COD and BOD. The overall range is that the BOD is between 40% and 60% of the COD. In Table 1, the three ratios are approximately 40%. A rough figure of BOD = 50% COD is not a bad estimate and depends on whose text is used. A standard figure of about 50% has often been employed. Figure 1, from Eckenfelder (7), would seem to contradict that, but given the variability of the BOD test, the 50% figure holds reasonably well. If one is considering the effluent BOD after biological treatment, then it is about 25% of the COD.

For modeling purposes, however, BOD is generally not used because it does not represent what is happening in the reactor tanks, it is not a consistent parameter, and the

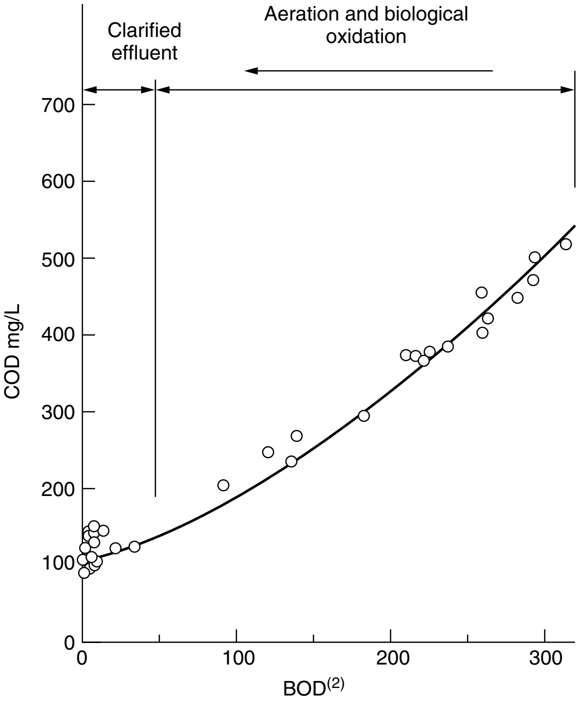

Figure 1. Relationship between COD and BOD (7).

Table 2. Various Indicator Strengths

Ratio	Low	Typical	High
COD/BOD	1.5–2.0	2.0–2.5	2.5–3.5
COT/TKN	6–8	8–12	12–16
COT/Tot. P	30–35	35–45	45–60
BOT/TKN	3–4	4–6	6–8
BOD/Tot. P	10–15	15–20	20–30
VSS/SS	0.4–0.6	0.6–0.8	0.8–0.9
COD/TOC	2–2.5	2.5–3	3–3.5

reproducibility of the test is uncertain. It does not make a good basis for modeling.

Additional Guidance on BOD/COD Relationships

Table 2 gives typical values for BOD/COD and other relationships in wastewaters. A high COD/BOD ratio indicates that organic matter is difficult to degrade. A high COD/TKN ratio indicates a waste that may require denitrification. A high VSS/SS ratio indicates a high organic matter relationship in the suspended materials (8).

MODELING THE BIOLOGICAL PROCESS

In the material that follows, we look briefly at the Monod equation and some of the mass balances around a typical activated sludge system. Then we continue briefly through the model development associated with the Peterson Matrix—which is used by the IWA in presenting and describing the mathematical development of the various models. Finally, we look at some of the balances around an activated sludge plant and consider the variables in the ASM1- Activated Sludge Model Number 1, which handles both nitrification and denitrification. The notation used and the pictures are taken from some of the modeling documents associated with HEMMIS' WEST® modeling platform and software and are used by permission.

Biological Growth Occurs According to the Monod Equation

The Monod equation is

$$\mu = (\lambda S)/(K_s + S)$$

where μ = specific growth rate coefficient
λ = maximum growth rate coefficient, which occurs at 0.5 μ_{max}
S = concentration of limiting nutrient: BOD, COD, TOC, etc.

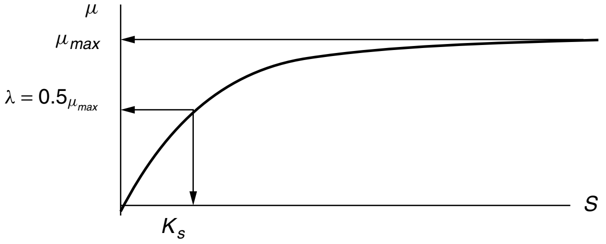

Figure 2. Graph of Monod equation.

K_s = Monod coefficient, also called the half-saturation coefficient because it corresponds to the concentration at which μ is one-half of its maximum, which can be seen from the Monod equation by setting S equal to K_s. K_s occurs at $\lambda = 0.5\ \mu_{max}$ (Fig. 2).

The curve is a plot of specific growth rate coefficient versus concentration of growth-limiting substrate when there is no inhibition.

$$\text{Organics} + \text{Bacterial} + \text{Nutrients} + \text{Oxygen} \longrightarrow \text{New Bacteria} + CO_2 + H_2O + \text{Residual organics} + \text{Inorganics}$$

$$\text{Rate of bacterial growth} = Rg = \mu X$$

where X is the microorganism concentration in mass/volume and μ = specific growth rate per unit of time.

$$\text{Cell Yield Coefficient} = Y_{obs} = \frac{-R_g}{R_{su}}$$

where Y_{obs} = observed yield coefficient, R_{su} = substrate utilization rate, and R_g = cell growth rate.

If we are looking only at the carbon variable, we get the following relationships:

$$Q = \text{Influent} = \text{Effluent} - \text{waste} = Q - Q_w + Q_w$$

= or if we define a new variable, Q_e, as the effluent, then

$$Q = Q_e + Q_w$$

For convenience, define the flow between the aeration tank and the clarifier as Q_2 with variable X_2. X_2 representing the combined flow of MLSS out of the tank and S_2 representing the substrate in the tank.

COD Relationships in Modeling

The models for activated sludge and biological processes use COD. The following are the relationships of the variables, and they will walk you through the formulation of the basics of the activated sludge models (ASMs).

In this example, we will first look at the carbon variable and then at the fractionation of the wastewater for treatment.

The basic activated sludge plant looks something like the schematic shown in Fig. 3.

Starting with a water balance, $Q_1 = Q_3 + Q_w$.

Running a mass balance around the system, we get

$$Q_1C_1 - r_{vs}V_2 = Q_3C_3 + Q_wC_r$$

where r_{vs} is the reaction rate for substrate utilization.

Note that in the above equation, no specific definitions are implied, so C can be NO_2, NH_3, COD, or anything else. However, one must define the unit of volume as well as the reaction rate, which means that the reaction rate can be r_{vs} or r_{xs} together with the unit of the volume V_2 and for r_{sx}, the activated sludge concentration X_2.

The units must be internally consistent. The activated sludge concentration X_2 can be measured in kg of SS/m^3,

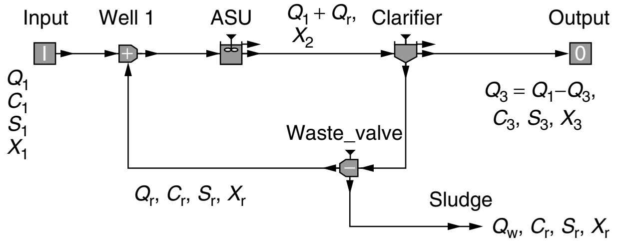

Figure 3. Activated sludge plant. Plant model was created using graphics from WEST® Software.

kg of VSS/m^3, or kg of COD/m^3, but the units must be consistent in the numerator and denominator.

Another way of looking at the mass balance is the following. The materials must be hydrolyzed before they can be accessed by the bacteria. So if you have some substrate such as BOD or COD, it cannot be used until it is solubilized, which means that the basic balance will look like the following when we consider the same mass balance as in Fig. 3, only now we have added the growth and hydrolysis terms to the equations:

$$\text{input} + \text{hydrolysis} + \text{growth} = \text{effluent}$$

$$\text{Input} + \text{sludge return} + \text{growth}$$

$$Q_1 S_{s1} + K_h * X_{S2} V_2 + [-(1/Y)]\mu[\mathrm{Ss2}/(S_{sw} + K_s)]$$
$$[S_{O_2}/(K_{sO_2} + S_{O_2})] X_{BH} V_2 = Q_3 S_3 + C_r$$

(note that the growth term removes the substrate).

X_{BH} is the heterotrophic biomass and S_{O_2} is the oxygen concentration, etc as shown below.

We can also add other terms to the equations to compensate for the oxygen and substrate terms, and those are most conveniently multipliers expressed in the form of the Monod equation: $S/(S+K)$. A few of them are shown below, and you will note that a few of them act as "switches" because when S goes below a specific value, the term tends toward zero, and the entire multiplier falls out of the balance equation.

Some of those terms are:

Now the Matrix:

Component	S_s	X_s	X_I	X_{BH}	S_{O_2}	Reaction Rate r_v
Process						
Aerobic heterotrophic growth	$-(1/Y)$			1	$(1-Y)/Y$	$\mu[S_s/(S_s+K_s)]*$ $[S_{O_2}/(K_{sO_2}+S_{O_2})]$ X_{BH}
Heterotrophic decay		$1-f_{XB}$	f_{XB}	-1		$b_H * X_{BH}$
Hydrolysis	1	-1				$K_h * X_s$
Units			kg COD/m^3			
	Easily degraded organic matter	Slowly degradable organic matter	Inert suspended organic matter	Heterotrophic biomass	Oxygen	

This is the start of formulating the equations for most of the activated sludge models developed by the IWA.

How to Use the Matrix and the Equations. Given the plant depicted in Fig. 3 and an effluent concentration of 10 mg/L COD (about 20 mg/L BOD). Arbitrarily, we use a flow rate of 5 million gallons per day which is a Q of 3.785 L/Gal*5 MGD= 18.925 M L/D = 18925 m^3/D at a strength of 500 mg/L (= 0.5 kg COD (S)/m^3. The plant has no recycle. Find the aeration tank volume.

Moreover, we have the following information provided: r_{xs} = 3 kg COD (S)/kg(B) COD 3 kg of substrate consumed per kg of bacteria, and the yield for substrate is 0.4 g COD(B)/kg COD (S) – 0.4 kg COD of bacteria per kg of substrate consumed.

Look at the equations above, and find the necessary volume for the tank:

$$\mathrm{V} = (Q_1 C_1 - Q_3 C_3)/(r_x X_b)$$

and

$$X_b = Y * (C_1 - C_3)$$
$$X = 0.4(500 - 10) = 0.4 \times 490 = 196.0$$

and

$$V = (Q_1 C_1 - Q_3 C_3)/(r_x X_b)$$
$$= (18925 \times 500 - 18925 \times 10)/(3 \times 196)$$
$$= 15770.83 \text{ cubic meters}$$

Now if the plant has recycle, $Q_w > 0$.

We can go on and look at the overall process doing mass balances in any of a number of ways and at various points. The critical element is to take the internal tank reactions and the recycle rate into consideration.

If we performed a mass balance around the aeration tank and use that to calculate the tank effluent concentration, we get the following:

$$X_1 C_1 + X_r C_r + (Q_1 + Q_r)(C_{\mathrm{ATIn}} - C_{\mathrm{ATOut}})$$
$$Y = (Q_1 + Q_4) * X_2$$

Obviously, that requires a bit more information.

The overall substrate removal rate is

$$r = (\mu/Y)(S_2/(S_2 + K_s)(S_{O_2}/(S_{O_2} + K_{sO_2}) * X_2$$

where S_2 is the concentration of organic matter in the aeration tank.

The mass balance for the entire plant is

$$\text{Input} + \text{hydrolyzed} - \text{removed} = \text{Output}$$
$$Q_1 * S_1 + r_x V_2 v_{xs} - r_{vs} V_2 = Q_3 S_3$$

where v_{xs} is the hydrolysis coefficient.

All the above is for a simple set of reactions, but it forms the basics for modeling.

The IWA models currently in use are ASM1, ASM2, ASM2d, and ASM3. ASM1 is the oldest, dating from about 1987, and ASM3 is the newest. The problem with ASM1 is that it is difficult to fractionate the influent in the manner required for the model without a lot of trouble. ASM1 and ASM3 can be used to model nitrate removal.

ASM2 and 2d are written with the phosphorous variable in mind. They are rewritten rather to consider the fact that activated sludge is comprised of cellular biomass that has the ability to store and use phosphorous. ASM2d is specifically for phosphorous removal with the addition of precipitates (metal hydroxides).

ASM3 is a BOD-based model, and it is somewhat easier to use but still requires a basic understanding of how the model works and the kinetics. ASM3 is not as widely used because it is newer and because of the model formulation with regard to cellular internal storage.

The ASM models are written in the same matrix notation used above in the table, and in the simplest model, ASM1, there are eight rate equations: aerobic growth of heterotrophs, anoxic growth of heterotrophs, aerobic growth of autotrophs, anoxic growth of autotrophs, decay of heterotrophs, decay of autotrophs, ammonification of soluble organic nitrogen, hydrolysis of entrapped organics, and hydrolysis of entrapped organic nitrogen. So, eight equations that have 13 variables all expressed in matrix form.

The use of the words heterotroph and autotroph refers to the wastewater bacteria that can use exterior or interior sources of carbon to oxidize the wastewater.

Heterotrophs are assumed to be the utility organisms, capable of doing a wide variety of things in a biomass system. They grow aerobically, anoxically, and may be active in anaerobic fermentation. They are responsible for hydrolysis of particulate substrates and can use all degradable organic substrates under all conditions.

Autotrophs are nitrifying organisms that are responsible for nitrification. They are obligate aerobes, classified as chemo-litho-autotrophs, and are responsible for ammonia oxidation to nitrate (*Nitrosomonas* and *Nitrobacter*).

The variables and their notation in WEST are as follows:

For ASM1

Note that the IWA model description uses almost identical notation, but uses subscripts: thus S_i would be expressed by the IWA rather than the S_I notation used here

Name	Description
H2O	Water
S_I	Inert soluble matter
S_S	Readily biodegradable matter
S_O	Dissolved oxygen
S_NO	Nitrate and nitrite
S_NH	Free and ionized ammonia
S_ND	Soluble biodegradable organic nitrogen
S_ALK	Alkalinity
X_I	Inert particulate matter
X_S	Slowly biodegradable matter
X_BH	Heterotrophic biomass
X_BA	Autotrophic biomass
X_P	Particulate products resulting from biomass decay
X_ND	Particulate biodegradable organic nitrogen

These components are also used to characterize the influent of the wastewater treatment plant.

Additionally, the following parameters are also required for specifying the state variables in the modeling process.

Variable	Description	Units
Y_H	Yield For heterotrophic biomass	g COD/g COD
i_XB	Mass of nitrogen per mass of COD in biomass	g N/g COD
Y_A	Yield for autotrophic biomass	g COD/g N
f_P	Fraction of biomass converted to inert matter	—
i_XP	Mass of nitrogen per mass of COD in products formed	g N/g COD
K_S	Half-saturation coefficient for heterotrophic biomass	g COD/m^3
K_OH	Oxygen half-saturation coefficient for heterotrophic biomass	g O_2/m^3
K_NO	Nitrate half-saturation coefficient for denitrifying heterotrophic biomass	g NO_3-N/m^3
b_H	Decay coefficient for heterotrophic biomass	1/day
mu_H	Maximum specific growth rate for heterotrophic biomass	1/day
n_g	Correction factor for anoxic growth of heterotrophs	—
K_OA	Oxygen half-saturation coefficient for autotrophic biomass	g O_2/m^3
K_NH	Ammonia half-saturation coefficient for autotrophic biomass	g NH_3-N/m^3
b_A	Decay coefficient for autotrophic biomass	1/day
mu_A	Maximum specific growth rate for autotrophic biomass	1/day
k_a	Maximum specific ammonification rate	m^3/(g COD.d)
K_X	Half saturation coefficient for hydrolysis of slowly biodegradable substrate	g COD/g COD
k_h	Maximum specific hydrolysis rate	g COD/(g COD.d)
n_h	Correction factor for anoxic hydrolysis	—
-Kla	Oxygen transfer coefficient	1/day
S_O_Sat	Oxygen saturation concentration	g/m^3

Y_H and Y_A must be larger then zero ($Y\_H > 0$ and $Y\_A > 0$).

Fortunately, many of these values are published in the IWA Model and are constant for many waste streams. The documentation considers published values plus ranges of these values in the models.

In the next article on plant modeling and hydraulics, we consider some of the challenges and opportunities associated with setting up a plant model.

BIBLIOGRAPHY

1. Lessard, P. and Beck, M.B. (1988). Dynamic simulation of primary sedimentation. *J. Environ. Eng.* **114**: 753.
2. Takacs, I., Patry, G.G., and Nolasco, D. (1991). A dynamic model of the thickening/clarification process. *Water Res.* **25**: 1263–1271.
3. Vesilind, P.A. (1979). *Treatment and Disposal of Wastewater Sludges*. Ann Arbor Science, Ann Arbor, MI.
4. Gaudy, A.F., Jr. (1972). Biochemical oxygen demand. In: *Water Pollution Microbiology*. R. Mitchell (Ed.). Vol. 1, John Wiley & Sons, New York, pp. 305–332.
5. Russell, D.L., Dudley, J., and Amerlinck, Y. (2002). Does your plant need an upgrade? *Water Environ. Technol.* October.
6. Liu and Liptak. (1994). *Environmental Engineer's Handbook*. Lewis, Boca Raton, FL, Table 7.18, p. 520.
7. Thackston, E. and Eckenfelder, W. (Eds.). (1972). *Process Design in Water Quality Engineering*. Jenkins, New York.
8. Henze, M., Harromotes, P., la Cour Jansen, J., and Arven, E. (2002). *Wastewater Treatment—Biological and Chemical Processes*. Springer-Verlag, New York, p. 31.

PRACTICAL APPLICATIONS OF WASTEWATER MODELING AND TREATMENT PLANT DESIGN

David L. Russell
Global Environmental
Operations, Inc.
Lilburn, Georgia

INTRODUCTION

In this article, we consider some of the specifics for setting up a model and plant hydraulics, and some of the practical considerations and limitations of modeling.

NOTES ON PLANT MODELING

The following notes on setting up a mathematical model of a wastewater treatment plant were prepared as a part of a training document for using WEST software. The proprietary parts have been removed, and the generic suggestions and ideas have been incorporated as a trial procedure based on personal experience. The procedure is designed for an ASM 1 model.

By following the steps outlined below, you will be able to set up and run a model and either match existing conditions or proposed conditions:

1. Information gathering
2. Model selection
3. Reactor hydraulics
4. Laboratory and other data organization
5. Flow sheet setup & model organization in WEST—the WEST C
6. Model compilation and setup—the WEST-S
7. Input and output file preparation
8. Initialization of model parameters and first runs
9. Parameter adjustments

Step 1: Information Gathering

Your objective is to make a model that represents a form of reality. That reality may represent either the dynamic behavior of a new facility or the dynamic behavior of an existing facility. The purposes are predictive and or historical. In the first case, you may want to find out how the facility will work and whether a specific design is adequate. In the second case, you may be looking to duplicate the performance of a real plant by preparing a model and matching the data from the model to the data from the real plant. In this latter instance, you can then use the data to predict future performance or investigate plant modifications.

Modeling is a very good but imperfect process. In the process of gathering information about a facility, you may never have enough information to satisfy your requirements. It will be easier to make assumptions about a new facility than to model the behavior of an existing one.

When you are dealing with an existing facility, you will almost never have enough information about the facility to satisfy your model and requirements. You will have to make some intelligent and reasonable guesses at certain things. You also may never know all that you want about a particular facility, but at the same time, you may know more about the facility than you ever dreamed of knowing because the data are often in the wrong place or of the wrong type.

This is not quite an "abandon all hope—ye who enter here" type warning, but rather a practical admission that you will be trying to reconstruct an event or make estimates of what may be happening with a model in an imperfect world.

You will need:

- An initial idea of which model you might want to use and the requirements of that model.
- Plant drawings and flow sheets.
- Aerator sizes and capacities.
- Physical dimensions of units.
- Maximum flow to each unit, which is particularly important in pumped flow where you have maximum capacities for pumps. Locations include clarifier underflow, and in nitrification plants, the return rates or recirculation rates for the system.
- Chemical data on the plant:
  - Influent data
  - Effluent data
  - Water quality data and effluent requirements
  - Internal plant operating records and logs
  - Chemicals added and operating set points

Most of the models are based on COD and most of the data you have are BOD. You may have to go back into the laboratory for more information to establish the necessary correlations to create the partition for the influent file.

Depending on how the model is established, you may or may not need additional data. If there is a system to enable you to transform your data into corresponding data for the model, it may be used after examination and careful consideration. One of the problems you may encounter is that the conversion routine makes some arbitrary decisions about the relationship between BOD and COD and the other parameters, and so you may have to consider that in the setup of the model. Also, the relationships (parameters and fractions) may change over time and with temperature and other influences. The program does not make those changes.

If you are working on a new design, you can make many choices about the type and kind of data you use.

Depending on the need for accuracy and performance, you may have to go into the laboratory and reproduce oxygen uptake rate (OUR) data, as well as yield and specific growth data, and also data for settling characteristics. If you are not going to be reproducing data from an existing plant, the effort in gathering detailed data may not be worth the trouble or expense.

A major caution should be raised here. If you are seeking to model any type of industrial waste, you probably should perform some minimum characteristics testing and should include some pilot plant testing as well as testing for settling characteristics because toxicity and hindered settling can often be associated with industrial waste treatment systems.

Step 2: Model Selection

Table 1 lists the input variables for each of the three principal activated sludge models, ASM1, ASM2d, and ASM3.

Model ASM1 is the oldest and has the most history behind it. It can be used for nitrogen removal in BNR plants. ASM2 and ASM2d were formulated later to consider the specific ability of the cell to store food materials. The ASM2d is formulated to allow the addition of metal hydroxides for phosphorous precipitation.

The ASM3 model is the newest, and many in the modeling community do not like it because of its formulation. They have indicated that the lumping of wastewater fractions and overall formulation of the model away from observed cellular behavior is objectionable. The difference in model formulation between ASM1 and ASM3 is substantial. ASM3 was formulated to address some of the deficiencies and difficulties with ASM1 resulting from assignment of the COD fractions of the influent. (See the discussion on COD and BOD and preparation of the influent file.)

Quoting and paraphrasing from the IWA description of ASM3 and ASM1 differences:

- ASM1 does not include kinetic expressions that can deal with the nitrogen and alkalinity limitations of heterotrophic organisms. The result is that the computer code cannot be based on the original form of ASM1, where a negative concentration of ammonia may occur under certain circumstances.
- ASM1 includes biodegradable soluble and particulate organic nitrogen as model compounds. These cannot be easily measured.
- ASM1 differentiates inert particulate organic material depending on its origin, influent or biomass decay, but it is impossible to differentiate these two fractions in reality.
- ASM1, ASM2, and AMS3 do not directly predict the MLSS or TSS concentration or the BOD concentration in a waste stream.
- In ASM1, hydrolysis has a dominating effect on predicting oxygen consumption and denitrification by heterotrophs. In reality, the process contains a number of coupled reactions.

ASM3 is a different model formulation all together.

An example of the differences among the models is shown in Table 1. These are the listed variables in the ASM1, ASM2d, and ASM3 models (not the parameters) expressed in the notation used in the WEST modeling platform.

Before we leave model selection, a couple of points are particularly important.

- The models do not define the conventional parameters as conventionally measured by the wastewater treatment industry. Everything is related to COD, and moving backward and forward through COD may introduce errors in modeling that are both frustrating and easily made.
- In most experiments, the researcher is often interested in modeling reductions in BOD, TSS, ammonia, phosphorous, and, perhaps, MLSS and COD. MLSS is often a control parameter for plant operations. The problem is that the values for many of these parameters are 'reconstructed' or measured indirectly from the COD. Ammonia and phosphorous are often directly modeled and measured, but MLSS, TSS, and BOD are calculated from COD data, and they can be difficult to obtain for control purposes.

In a recent modeling experiment, this point was again driven home by the frustration of trying to model conventional parameters. The experimenter was trying to model TSS accumulation in a membrane bioreactor system and was being frustrated by the model's internal composition.

The ASM2 model (and the other models as well) has a complex definition of TSS. The definition of TSS includes part of the soluble and insoluble fractions of the particulates; part of the heterotrophic and autotrophic bacterial mass; and part of the internal storage product of phosphorous accumulating organisms, metal phosphates, and polyphosphates. Each factor is multiplied by partitioning coefficients, and a change in heterotrophic concentration or other parameters may influence the measurement of TSS. In the other models, the definition of TSS is somewhat less complex and more easily measured.

Table 1. Variables in ASM Models

ASM1 Variables		ASM2d Variables		ASM3 Variables	
Name	Description	Name	Description	Name	Description
H2O	Water	H2O	Water	H2O	Water
S_I	Inert soluble matter	S_I	Inert soluble matter	S_I	Inert soluble organic matter
S_S	Readily biodegradable matter	S_O	Dissolved oxygen		
		S_N2	N_2	S_S	Readily biodegradable organic substrate
S_O	Dissolved oxygen	S_F	Fermentable, readily biodegradable organic matter		
S_NO	Nitrate and nitrite			S_O	Dissolved oxygen
S_NH	Ionized and un-ionized ammonia			S_NH	Ammonia
		S_A	Fermentation products considered to be acetate	S_N2	Dinitrogen
S_ND	Soluble biodegradable organic nitrogen			S_NO	Nitrate
				S_ALK	Alkalinity
S_ALK	Alkalinity	S_NO	NO_2^- and NO_3^- nitrogen	X_I	Inert particulate organic matter
X_I	Inert particulate matter				
X_S	Slowly biodegradable matter	S_PO	Inorganic soluble phosphorus, primarily orthophosphates	X_S	Slowly biodegradable substrates
X_BH	Heterotrophic biomass			X_H	Heterotrophic organisms
X_BA	Autotrophic biomass	S_NH	Ammonium nitrogen, NH_4-N	X_STO	Cell internal storage product of heterotrophic organisms
X_P	Particulate products resulting from biomass decay	S_ALK	Alkalinity of the wastewater	X_A	Nitrifying organisms
X_ND	Particulate biodegradable organic nitrogen	X_I	Inert particulate matter	X_TS	Total suspended solids
		X_S	Slowly biodegradable matter		
		X_H	Heterotrophic biomass		
		X_PAO	Phosphate accumulating organisms: PAO		
		X_PP	Polyphosphate		
		X_PHA	A cell internal storage product of phosphorus accumulating organisms		
		X_AUT	Autotrophic biomass		
		X_TSS	Total suspended solids TSS		
		X_MeOH	Metal hydroxides		
		X_MeP	Metal phosphate ($MePO_4$)		

This is also true of BOD. In some of the models, BOD is calculated through the COD where it factors into the ultimate BOD, and then to the 5-day BOD. Needless to say, the conversion factors play a large part in the conversion. Another example is that the COD to BOD conversion formula is different for the influent and effluent parameters.

- At least one company, WRc, Ltd., in England, has a model that is formulated strictly on BOD (STOAT).
- The ASM3 model is still COD based, but it is a lot easier to use than ASM1 and ASM2. The ASM3 model has its limits, discussed above, but it is generally more forgiving, if, arguably, less accurate than the ASM1 and ASM2 models.

Step 3: Reactor Hydraulics

Reactor hydraulics is an extremely important consideration in modeling. Reactor hydraulics are critical in determining both the type of model used and the relationships involved in the plant design. For complete mix plants, the concentration of solids in the aeration tank is the same at both influent and effluent, or $X_{\text{in}} = X_{\text{out}}$.

But for a plug flow plant, the concentration of X varies through the aeration tank. The characteristics of the concentration in the effluent depend on the reactions and the tank hydraulics.

The computer models of activated sludge systems use continually stirred tank reactors (CSTRs). The ideal flow pattern in a long narrow tank approaches plug flow. The reactor flow patterns need to be addressed in the model so that the hydraulics of the model match the hydraulics of the wastewater tankage.

The subjects of reactor hydraulics and dispersion are beyond the scope of this article but need to be considered to ensure adequate hydraulic modeling for the plant.

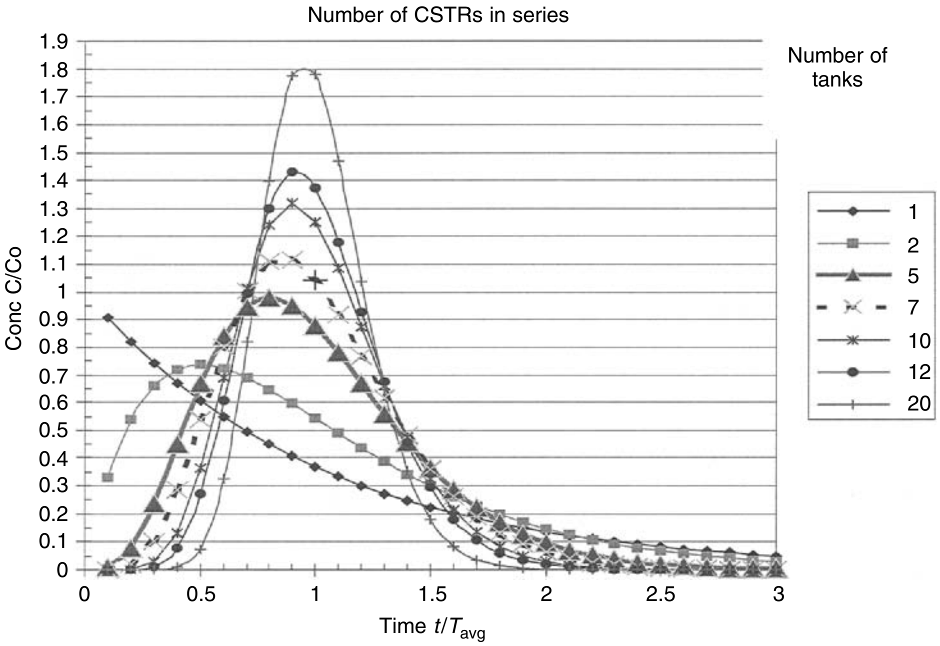

Figure 1. Changes in concentration with time.

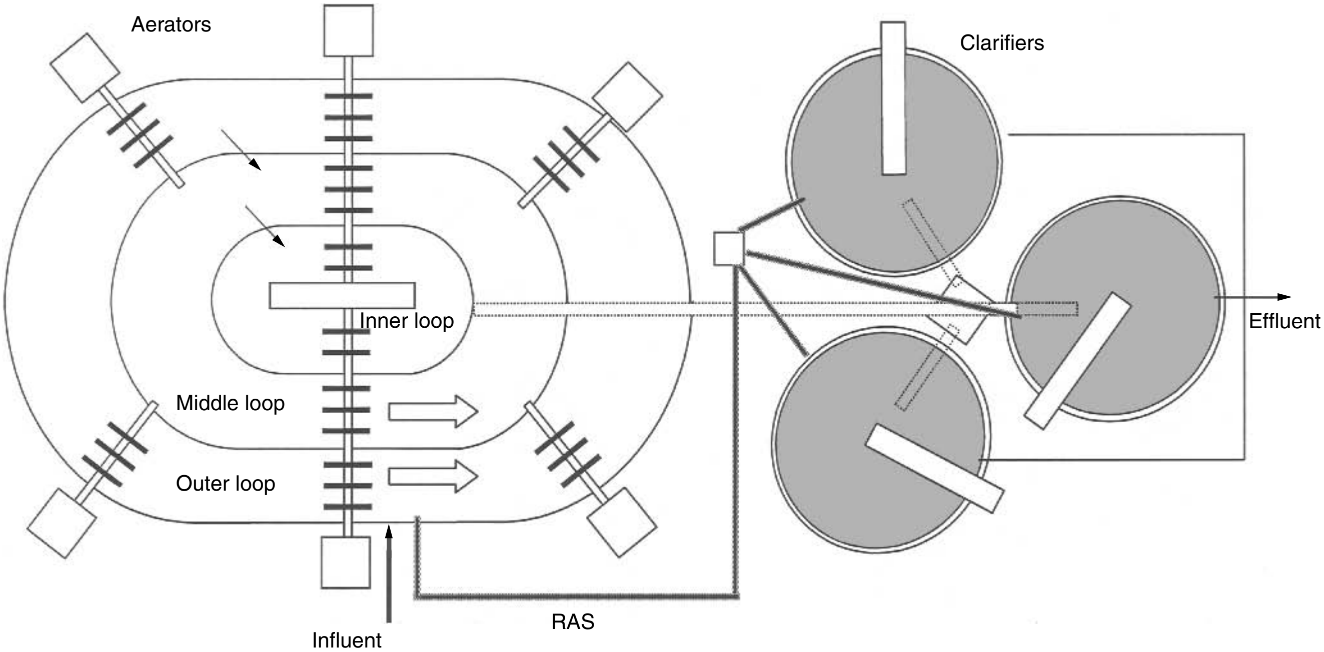

Figure 2. The general configuration of the ORBAL followed by configuration in WEST.

The relationship between the number of CSTRs in series and the concentration is shown in Fig. 1.

The difference between a true plug flow reactor and a true complete mix reactor is defined in terms of the axial turbulent diffusion coefficient, average flow velocity, length of the reactor, and the inverse of the Peclet number, or D/UL. The important thing to consider is that the higher the peak and the sharper the curve, the more the reactor acts like a plug flow reactor. (This also means that, for a greater number of tanks, one will have to do a lot more work in setting up the model.)

In Fig.1, the value for one tank is comparable with a complete mix reactor, whereas the value for 20 tanks approaches that of a plug flow reactor. There is a balance formula for this computation that is given by the following:

A correlation that relates the internal mixing characteristics (number of compartments or stages) to the physical properties (for diffused air systems) was used. The correlation, in SI units, is

$$N = 7.4\, Q_S\, (1 + R)\, L/WH$$

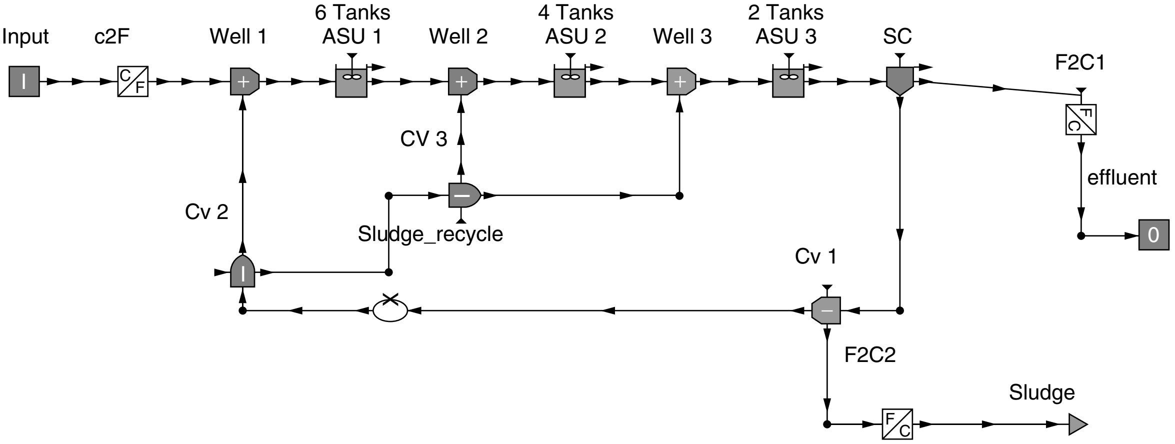

Figure 3. ORBAL plant configuration in WEST (*Source*: Dynamic calibration of the Orbal extended aeration plant by Peter A. VanRolleghem and H. Gucu Insel, Biomath Department, University of Ghent, October, 2002).

where N = number of compartments;
Q_S = flow, m^3/s;
L = length, m;
W = width, m; and
H = depth, m.
R = RAS ratio;

This formula was first published by Chambers (1). The units have to be in the overall range of the following:

$$28 < L < 500$$

$$2 < B < 20$$

$$2.4 < H < 6$$

$$0.7 < R < 1.5$$

$$1.3 \text{ hours} < t < 3 \text{ hours}$$

where $t = 3600\ Q\ (1 + R)$.

The following example may provide some guidance.

Recently, we had to model an ORBAL plant using WEST (Fig 2). The ORBAL has three concentric circles and a number of aerators with complex flow patterns because there is a submerged opening or, in some cases, a wall opening between the concentric tanks to allow the waste to pass from the outer tanks to the inner tanks.

The purpose of the figures is to introduce the reader to a practical application of the consideration of reactor hydraulics. In the ORBAL, there is a natural break point for each of the series of ORBAL channels, and those are the aerators (Fig. 3). The location of the aerators naturally divides the channels into logical units because of reaeration considerations. It can be argued that the channels could be divided into twice as many tanks for more accurate modeling, where the first tank would reflect the aeration input by the aerator and the second tank would reflect the lack of aeration input, which might be a better model, but it is definitely twice as complex and will require at least twice as long to compile and run. It may or may not be worth the effort.

Step 4: Laboratory and Other Data Organization

The objective is to gather enough data to be able to make an estimate of the magnitude and change in the flow sufficient to characterize the waste stream and its elements.

Depending on the model you are planning to use, you may have to go into the laboratory and obtain additional information. The number of samples and complexity of the sampling program is up to you and the budget you have available. We have observed everything from 12–15 samples run to establish some basic correlations, to a complete sampling campaign using automated equipment that will allow you to collect and analyze up to six sampling points and multiple analyses for each sampling location. This latter equipment program is described by Dr. Bruce Beck in his Environmental Process Control Laboratory or EPCL. The web address for the EPCL is http://hilbert.forestry.uga.edu/epcl/issues&info/flowchart.html.

It is difficult to provide specific guidance on the sample reliance for an unknown waste stream, but it would be nice if the laboratory data one obtained were reproducible to within about 10–20%. The entire discussion of precision and accuracy and the number of samples required therein is outside the scope of this paper. For a discussion of the issues involved in sampling and analysis, see Ref. 2.

Here are some suggestions for running the ASM1 testing and ASM2 and ASM2d testing from a proposal to a client. For additional guidance, you will want to see the descriptions of the models themselves.

> Any modeling or characterization effort requires a substantial amount of data. In this instance, we will probably need access to the plant analytical and operating records, chemical addition, and other records for at least 6 months, and perhaps a full year. The scope of the request is somewhat more extensive than the description might indicate, as it includes such routine operating data as the pump characteristics and rate, hours of operation, timing of the operations, aerator horsepower and hours of operation for each day, etc. The data should also include the quantity and type of chemicals used in various parts of the plant.

> Wastewater characteristics are divided into several elements. The principal elements include
>
> - Total and filterable COD
> - Total and filterable TKN
> - Total ammonia and nitrate
> - TSS & VSS
> - Total phosphate, filterable phosphate, and orthophosphate
> - Alkalinity
> - Volatile fatty acids (measured as acetate) (only for ASM2 and ASM2d)
>
> Analyses should be run in parallel with routine influent and effluent analyses for a period sufficient to determine reliable parameter variation and fractionation with routinely measured variables.
>
> A sampling campaign between 1 and 2 week long is suggested for the wastewater characterization study. The study should collect and analyze daily composite samples of influent and effluent. The study should be conducted in parallel with the routine sampling data being collected at the plant now.

Automatic sampling equipment that collects hourly grab samples of water and wastewater is often as good as can be obtained in a sampling campaign. Daily composite samples can be used but are often unreliable because the waste stream fluctuates with time and composite data will mask the diurnal and other variations.

Flowmeter readings are often unreliable, even when the flowmeter works. Calibration and inspection is suggested.

For internal calibration, you will need some data relating to TSS and VSS (filtered and unfiltered in various parts of the plant), along with DO measurements, and possibly some respirometry data, recycle rates, sludge levels in the clarifier, sludge wasting data, and just about everything you can imagine that an operator would need to run the plant properly. The trouble is that the data do not often exist or exist in the wrong form.

An example might be the notation in the plant operating log—"wasted sludge for two hours". That tells you that the operator wasted sludge at between 1.5 and 2.5 hours, but nothing else unless you have either the valve settings and the MLSS at the time of the wasting.

If you are fortunate, you will find that he took an MLSS sample immediately before or after the wasting. If not, you may have to guess from the pump flow rate and talking to the operator. The key is that the operators almost always have a favorite method of operation that can often be recovered. They may know that if they turn the valve wide open, something will happen that will cause them work down the way, and so they leave the valve partially shut so that they do not have to clean up the floor of the sludge treatment facility.

The operation of the aeration equipment is often manual as well. It is safe to assume that about the only instrumentation you may find is the flowmeter, and that may not even work.

Generating the Flows When You Do Not Have the Data. When you have only BOD data or BOD, COD, and TSS data, WEST and some other programs have both a BOD transformer and/or a BOD/COD transformer that can be used to fractionate the flow into the necessary components. Remember the cautions expressed on letting a program fractionate the data for you.

If you have little data and have to generate dynamic and/or diurnal flow patterns from your source, you may want to use a sinusoidal pattern to generate the data. All good books on environmental engineering have some form of generic flow pattern for domestic wastewaters. The flows are often highest between 9 A.M. and 4 P.M. and lowest in the early morning hours when most people are asleep. Often there is a secondary peak in flow during the dinner and early evening hours when food preparation is the greatest.

You may want to look at your time data when it was collected from the influent and try to fit a time curve with each flow and concentration element to see if there is any variability in the data and determine whether you need to have the biosolids and chemical composition run counter or with the hydraulic basis. The following example may help to explain this point:

> Assume that a process discharges a uniform concentration of material and the flow varies. Then the concentration of contaminants would vary directly with the flow.
>
> If, however, the process discharged at a constant rate and contaminant level and if we add periodic boiler blowdown and filter backwash to the waste stream and some kitchen wastewater (which peaks from about 1–2 hours either side of midshift), then the flows and the concentration may vary inversely with each other or may spike at midshift. It may depend on the quantity of blowdown and backwash.

Getting the Hydraulics and the Tankage Correct. As part of the overall process and before you set up the flow sheet, you will need to address the issue of hydraulic similitude for the system, which really compares how well the model you are using models the hydraulic performance and hydraulic mixing of the system you are modeling. The best way to determine how well your system is performing is to conduct dye or tracer studies on the tank system. The purpose of the dye testing is to find out whether you are operating in a plug flow or complete mix or some other regimen. The differences among the types of reactors were discussed earlier in this article.

If you cannot perform dye testing to determine the mixing requirements, you should examine the system you are modeling carefully. There may be some logical divisions in the system that suggest that it should be modeled using a specific number of tanks. For example, an extended aeration system, such as an oxidation ditch, may have a number of aerators, and the physical positioning of the aerators may suggest a logical number of tanks to use for modeling the flow.

Each tank or tank division will have its own aeration capacity (*Kla*) that can be adjusted. Multiple aerators may suggest a configuration that needs more than one tank and an interlaced flow diagram to suggest the right type of flow connections, which is a matter that is settled by the judgment of the engineer.

At this point, it may be a good idea to rough out the flow sheet for your model, which is also a first

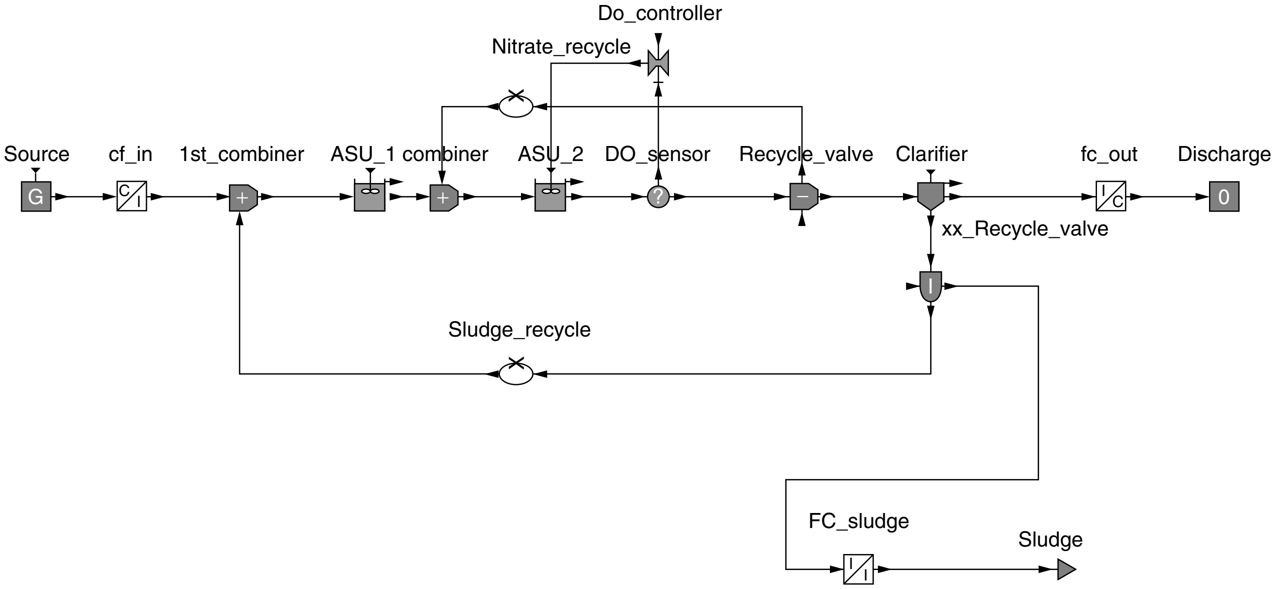

Figure 4. Labeled flow sheet in WEST.

draft of the process and instrumentation diagram (P & ID). It may not be detailed but rather simple. You can always fill in controls and details later when you set up the compositional model. Clarifier selection is also an important choice for both the model and the complexity of the models.

A simple flowsheet is shown in Fig. 4.

Start with an elementary drawing, and put the basic icons in place. Annotate it with the details you need to prepare the drawing of the plant, so that it reflects the P & ID as shown above. Then start your process simulator, and begin your work.

After you have completed the physical placement of the model elements and the selection of the model, you still have to several steps to go.

Step 5: Model Compilation and Setup

Setting up a model is a multipart process:

1. Initial definition of the model critical model parameters
2. Initialization of the variables
3. Setting up the integrator

Start by initializing the model. First, start with the parameters. Most of them are biological parameters that you will not have to adjust and probably should not adjust.

Concern yourself with the physical parameters first. You will want to look at the physical parameters and set things like tank volumes and initial guesses to determine the *Kla* (oxygen transfer and uptake rates), suspended solids levels in the tanks, and various physical parameters.

You will also have to go through and estimate the initial operating conditions for the reactors by going into the variables and filling them out.

One of the other types of errors that can be generated in ASM1, particularly, is *negative ammonia*. Negative ammonia is caused by an unbalanced waste in the model. The configuration of ASM1 is such that it does not have a “trap” for negative ammonia, and what the model is really saying is that the feed conditions are such that the ammonia is consumed at a faster rate than it is supplied by degradation of the organic material in the waste. Obviously it is an impossible condition, but the model does not know the difference. Sometimes, the autotrophs in the waste are too high, and a simple adjustment in the feed will correct that problem.

This is one reason why it is often more convenient to set up and run a steady-state model for the initial run. The steady-state system allows you to balance the model before you begin dynamic modeling.

After the steady-state model is balanced and running properly and after you have set the various values at the approximate levels you need for your initial conditions, then it is time to go back and recalibrate the model using your successful run data from the steady-state case. The purpose of this step is to bootstrap the model with the values you will be using rather than wait for it to reach equilibrium.

Integrator Settings and Data Collection. All equations in the IWA models are expressed as rate equations that need to be integrated. The simulator package you are using does this. In some packages, you have a choice of how fine you want your integration settings. On other simulators, you may not. Remember that if you use too large an increment, your program may be fast, but you sacrifice accuracy. Fine steps in the integrator provide nice smooth curves but take additional computation time and can produce very large effluent files.

Step 6: Input and Output File Preparation

The setup of the influent file is strongly dependent on which model is being used. Some models use comma delimited files, other tab delimited files, and some use an input and output from Excel files.

Step 7: Initialization of the Model Parameters and First Runs

The first run of a model may produce unexpected results. The unusual results are not unexpected, and to some extent, that is why one sets up a static model first. When the static model has been evaluated and balanced, then it is necessary to use the final values for variables and modify them to represent starting variables for the first runs.

What to Balance or Adjust. A limited number of things can be adjusted in most models:

- Primary clarifier underflow rate (removes solids).
- Oxygen transfer rates (*Kla* values) can be adjusted in each tank.
- Flow proportion in each tank (if flow sheet permits).
- Recycle rate on loops (can take almost infinite adjustment).
- Secondary clarifier underflow (increases sludge return and improves effluent solids).
- Sludge wasting rate (decreases solids in the system).
- *Adjustment of the rate coefficients and parameters is not recommended on initial runs!* (Adjusting these coefficients prematurely may mess up your results in later runs.)

What to Key in on During Your Modeling. Key in on the same things that the plant operator keys in on: suspended solids, oxygen levels, sludge wasting rates, hours of operation of certain processes, and flow rates. Look specifically at weekends because many plants are essentially unmanned on weekends.

These are the operator's only tools. The two or three hardest to measure are, frequently, the MLSS values in the tanks, the oxygen levels in the tanks when and where it is measured, and any other chemical parameters.

When you have the model balanced for steady-state flows, then start dynamic modeling. If you are fortunate enough to have good quality data on the plant, you may be able to get traces where peaks and valleys in the model results can be compared with the peaks and valleys in the performance data for the plant. It is almost like curve matching, but you are looking to see that any disturbances in the plant (and there will always be some) match disturbances in the model. The results may be very interesting; the peaks may not match in height, but they should be at about the same times and for the same duration as obtained from the plant data.

Step 8: Parameter Adjustment

Adjustment of the parameters is often as much art as science. It needs to be done only after a detailed analysis of the waste and the characteristics and after evaluating the modeling results compared with a case where there are sufficient data to enable comparison.

From the literature, the one case where parameter adjustment may be necessary is in phosphorous modeling. The sensitivity of the various coefficients and uptake rates can affect the storage of phosphate and the final results. The adjustment of parameters should be undertaken only after careful consideration and experience and probably substantial laboratory testing and/or a number of model runs to compare results with actual performance.

Unless you have specific chemical data or parametric data to indicate that a specific change in a model is desirable, do not plan on changing the given model parameters as it may cause problems later, which is especially true for dynamic modeling. If you do not have any good indication of what and where to change things, try leaving them alone and make adjustments only as a last resort after you have tried everything else and need to make those adjustments based on dynamic data or specific knowledge of the plant.

FINAL NOTES

Look carefully at your data and your modeling system. If you understand it and the way it is put together, fine! If you do not understand what is happening, and why, you are like the student who says, “Don't bother me with the theory, just show me where to plug in the numbers!”. That is uncertain territory and can cause difficulties for the modeler. Unexpected results do occur and one has to understand both the results and how they relate to the physical model.

Acknowledgments

I need to provide special thanks to Mr. Youri Amerlinck of Hemmis, NV, and Mr. Dirk Van der Stede, President of Hemmis, NV, and Dr. Peter Van Rolleghem at the Biomath Department of the University of Ghent, Belgium, for their invaluable technical assistance. Without their help and education about modeling and WEST in particular, this article could not have been written.

BIBLIOGRAPHY

1. Chambers, B. (1992). Design methodology for optimization of aeration efficiencyin activated sludge plants. In: *Proc. 6th Fourm Appl. Biotechnol.*, Med. Fac. Ladnbouww, Univ. Ghent, 17, 1631–1642.
2. Russell, D.L. (1980). Monitoring and Sampling liquid effluents. *Chem. Eng.* October.

NEW YORK CITY HARBOR SURVEY

David J. Tonjes
Cashin Associates, PC
Hauppauge, New York

The New York City Harbor Survey, begun in 1910 and conducted continuously since 1912, may have the longest operational history of any annual water quality monitoring program in the world. Initiated to document decaying water quality in New York Harbor and, therefore, cause changes in waste management practices, it now demonstrates the effectiveness of New York City's comprehensive wastewater treatment program. Data from the 53 stations of the Survey (some of which are monitored year-round using the dedicated Harbor Survey vessel) show general compliance throughout the Harbor with New York State water quality standards today, despite historical pollution problems.

INTRODUCTION

New York City is composed of islands and shorelines. The proximity of so much water to so many people has meant that much of the wastes produced by its residents during the past 300 years has been discarded into the waterways (1).

One impetus to understanding the impacts of water pollution has been to document damage to the environment (and associated threats to human health) from these releases. It is easier to provoke action when the magnitude of the problem is displayed. This concept was understood by the New York City Metropolitan Sewage Commission in 1909. To quantify conditions in the Harbor, parts of which were described as open sewers, where swimming was called a dangerous health hazard, and where a once unmatched oyster industry was soon to be abandoned, the Commission established the Harbor Survey (2).

The first New York Harbor Survey was undertaken in 1910; samples have been collected every year since 1912 (generating 91 years of data as of 2001). This appears to make it the longest continuously conducted water quality monitoring program in the world.

PHYSICAL SETTING

Water quality is determined by the amount of chemical contaminants present and by their fate and transport. New York Harbor (Fig. 1) has a complex physical oceanographic setting that must be understood to interpret sampling data.

Generally, the Harbor can be considered a drowned river valley (3), formed as the continental glaciation receded from Staten Island 20,000 years ago (4). The Hudson River is tidal to the Troy Dam, 150 miles north of Manhattan, and its flow is entirely freshwater only for the 90 miles north of Poughkeepsie. Flow over the Troy Dam (90% of the freshwater in the river) has an average annual peak of 40,000 cubic feet per second (cfs), usually in spring due to snowmelt and greater groundwater discharges. The other major rivers feeding the Harbor (the Passaic and Hackensack Rivers into Newark Bay and the Raritan River into Raritan Bay) contribute a quarter of the Hudson River flow (5,6).

Generally, the Harbor is a partially stratified estuary (the Hudson River, Upper Bay, Lower Bay, Raritan Bay, Newark Bay, the eastern portion of the Upper East River, and western Long Island Sound), where tidal currents mix salt water into riverine water. The earth's rotation causes concentration of less dense (fresher) water on the west side of each waterbody (7). Mixing of tidally advected salt water dilutes the freshwater, so that flows at The Battery often exceed 400,000 cfs (8). Strong tidal currents can break down the stratification caused by the Coriolis effect, especially in the Upper and Lower Bays (9).

However, this straight-forward depiction of water flows is complicated by the multiple ocean connections (most estuaries have only one), where the Upper Bay connects to the Atlantic Ocean through the Lower Bay and also through the East River into Long Island Sound and to the Ocean. In addition, minor, well-mixed tidal straits, including the Harlem River (connecting the East River to the Hudson River), the Kill van Kull (connecting Newark Bay to Upper Bay), and the Arthur Kill (connecting Newark Bay to Raritan Bay), are not partially stratified. Jamaica Bay, which tends to be well mixed, and although not contiguous with the other parts of the Harbor, is also considered part of the Harbor (8).

Flow patterns in the Lower East River are not simple. Flow is driven by the mismatch in tides between the Upper Bay and Long Island Sound. Velocities can reach 4.5 knots (still much less than 10 knot currents once measured before a large rock was removed) (10), and it tends to be well mixed. There are six water pollution control plants (WPCPs) along the East River, and they input 750 million gallons per day (MGD) of effluent [15% of the total flow of 6 billion gallons per day (BGD)] (11).

Jamaica Bay also contains an abundance of sewage effluent; its natural tributaries were captured by sewers during urbanization. Up to half of its water can be sewage effluents (12), partially because water residence time is more than a month (up from 11 days due to dredging, which increased the depth of the Bay from 3 feet to 16 feet) (13). The Bay is generally shallow enough that stratification caused by summer insolation is of short duration due to mixing by winds (12).

New York Harbor was once famed for its natural beauty and seafood resources (14). The physical changes in the Harbor have caused impacts which in many cases cannot be differentiated from those from pollution. The predredged Harbor was 18–20 feet deep (9). Ships sailed Long Island Sound to avoid the shoals of the Lower Bay (10). Channels criss-cross the harbor, at depths as great as 45 feet, and plans are to increase some to more than 50 feet (15). The Harbor used to hold gently sloping coastal marshes and mudflats but now has hardened, steep-sided, deepened channels. Losses include 61,000 acres of tidal marsh (17), and fills have created 7,000 acres in the City (18). The islands in the Harbor have been removed, expanded, connected to other islands, and otherwise altered (19).

SOCIOLOGICAL SETTING

New York was attractive to immigrants in colonial days because of its natural resources. Its location made it (and the Hudson Valley) the key site of the Revolutionary War. The construction of the Erie Canal made New York City the economic capital of the United States (20). New York Harbor and water power available in New Jersey made the region a center of the Industrial Revolution and, ultimately, the country's economic engine (21).

Thus, New York City and its environs grew with time. It had 25,000 residents at the turn of the nineteenth

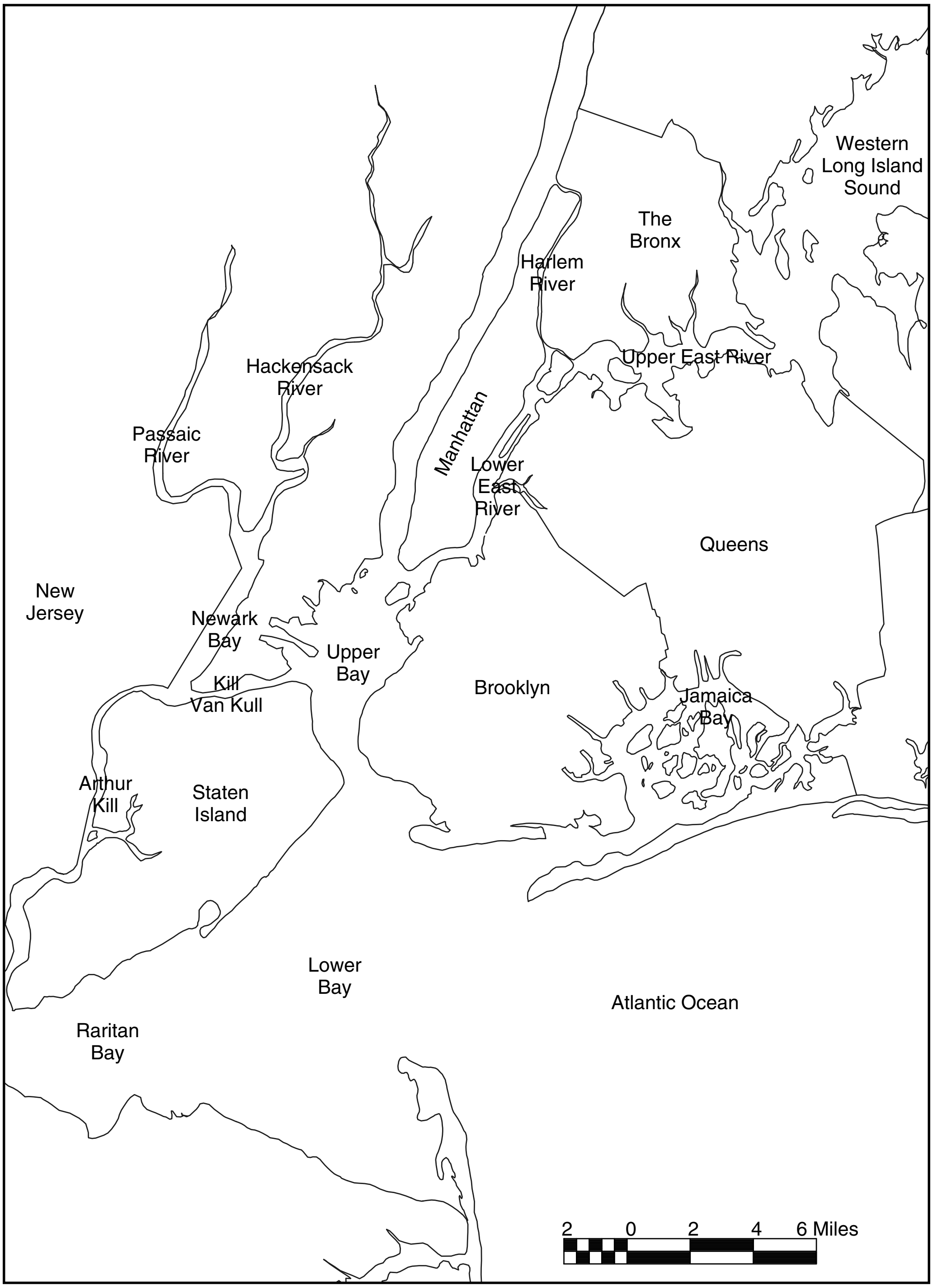

Figure 1. New York Harbor.

century (1). Population in the Harbor drainage area exceeded four million by the twentieth century (22) and more than 20 million at the turn of the twenty-first century (23). These multitudes produced solid and septic wastes, and the Harbor was the local disposal point since the 1600s. City solid waste was deposited into the Harbor or nearby waters until the early 1900s (24). Untreated sewage, likewise, was sent to the Harbor for dilution and transport away from land. As of 1905, 5.3 million people in New York and New Jersey sent sewage to the Harbor (25), augmented by industrial effluents (26).

In 1910, seven small WPCPs in New York City processed approximately 19 MGD, but the "process of purification has not materially reduced either the suspended matter or matters in solution." Typhoid outbreaks occurred in 1904 and 1918. In 1925, oyster fishing was banned. The result was construction of a sewage system for the City in the 1930s (25), although complete waste treatment was not achieved until 1989 (27).

CURRENT NEW YORK CITY SEWAGE TREATMENT INFRASTRUCTURE AND PROGRAMS

The 8.0 million residents of New York City are served by 14 WPCPs that generate 1.3 BGD of effluent (Table 1). All of these plants achieve secondary treatment (defined as 85% removal of BOD and TSS) (28).

New York City uses a combined stormwater–wastewater collection system of 1930s' design (1). Its engineers were not anticipating modern pollution controls for stormwater but intended to conduct as much water management as possible using the least amount of piping (which nonetheless totals 6344 miles) (28).

The combined system is stressed in the rain. To prevent volume overloads at the WPCPs, wastewater bypasses the plants. The combined sewer overflows (CSOs) that result remain a serious pollution concern. Water conservation measures led to City-wide decreases in consumption from 1.4 BGD in 1990 to 1.24 BGD in 1999, despite a population increase of 400,000. The reductions mean that dry weather flows to the WPCPs use only 70% of the system's capacity. Thus, the WPCPs can now manage more than 50% of the runoff (annually), and CSOs have decreased (29). The City has further committed to a $1.5 billion program to build CSO abatement facilities (28).

In the 1970s and 1980s, CSOs were the source of a local problem: floatables—washup of medical wastes on local beaches. Close examination of the wastes and transport calculations showed they were washed off City streets into CSOs (30). The problem has diminished due to the capture of materials in the Harbor and at catchbasins and an emphasis on street-cleaning (28).

The NYCDEP has decreased the toxicity of its sewage effluent through source control—700 sewer users have decreased metals loadings from 3000 pounds per day (1974) to 205 pounds per day. The Shoreline Survey Program monitors storm sewers for dry weather flows and then tracks down illegal discharges. The Sentinel Monitoring Program uses coliform to identify contaminated outfalls for investigation and remediation (28).

Table 1. NYC Water Pollution Control Plants

WPCP	Primary Treatment	Secondary Treatment	Last Upgrade	Current Capacity (MGD)
Coney Island	1936	1995	1995	100
Wards Island	1937	1937	1997[a]	275
Bowery Bay	1939	1942	1973	150
Tallman Island	1939	1939	1976	80
Jamaica	1943	1943	1971	100
26th Ward	1944	1951	1990	85
Hunts Point	1952	1952	1979	200
Owls Head	1952	1952	1995	120
Rockaway	1952	1962	1971	45
Port Richmond	1953	1978	1979	60
Oakwood Beach	1956	1956	1978	40
Newtown Creek	1967	1967[b]	1967[b]	310
North River	1986	1991	1991	170
Red Hook	1987	1989	1990	60

[a] Additional tanks installed in 1998 increased capacity by 25 MGD.
[b] Currently operated as modified aeration. Upgrade to full secondary operation planned for 2007.

THE HARBOR SURVEY

Six stations were sampled in 1910. Sampling was continued in 1912 and has been conducted each year since. A special section of the NYCDEP, the Marine Science Section, conducts the Survey. It has its own 55-foot, aluminum-hulled, twin-engine diesel research vessel, the HSV OSPREY. The OSPREY has an onboard laboratory and is crewed by six to eight. It sails 5 days per week from June to September, when all 53 stations (Fig. 2) are sampled biweekly (a subsection of stations is sampled monthly, year-round) (28).

The Survey aims to determine the impacts of sewage treatment on the Harbor. Therefore, the parameters keenly tracked are coliform, dissolved oxygen (DO), and nutrients, especially nitrogen compounds. Temperature and salinity data are collected and used to determine the physical structure of the Harbor waters, placing laboratory results in context. Ecological data (chlorophyll *a* concentrations and some phytoplankton identification) are also collected. Beginning in 1985, samples were collected from the water surface (within 1 meter of the surface) and from the bottom of the water column (1 meter from the bottom). Statistical analyses of the data showed strong correlations between top and bottom data; beginning in 2001, the Survey collected field parameters (temperature, pH, turbidity, salinity, and DO) from top and bottom waters and water quality samples of surface water only (28). Before 1985, nearly all lab data were generated from composite samples (27).

DOCUMENTED WATER QUALITY TRENDS

Harbor Survey data have been used to address two basic questions:

- How bad was water quality? and
- Has a resolution of sewage-related pollution been reached now?

Suszkowski (25) analyzed the first 60 years of Survey data, using basinwide averages of summer data (ie., all summer sample results from the Hudson River were summed and averaged to generate one "Hudson River" value) and Harbor-wide averages (the mean of all basins). The findings were that the mean DO, measured by saturation percent, decreased from approximately 80% of saturation in 1910 to 40% in the late 1930s. In the 1960s, the DO slowly increased to 70%. The lowest values were found in the inner areas of the Harbor. The DO trends were linked to biological oxygen demand loadings from raw sewage and WPCP effluents. The correlation indicated that organic loading from raw and treated sewage was the prime factor in decreased DO levels.

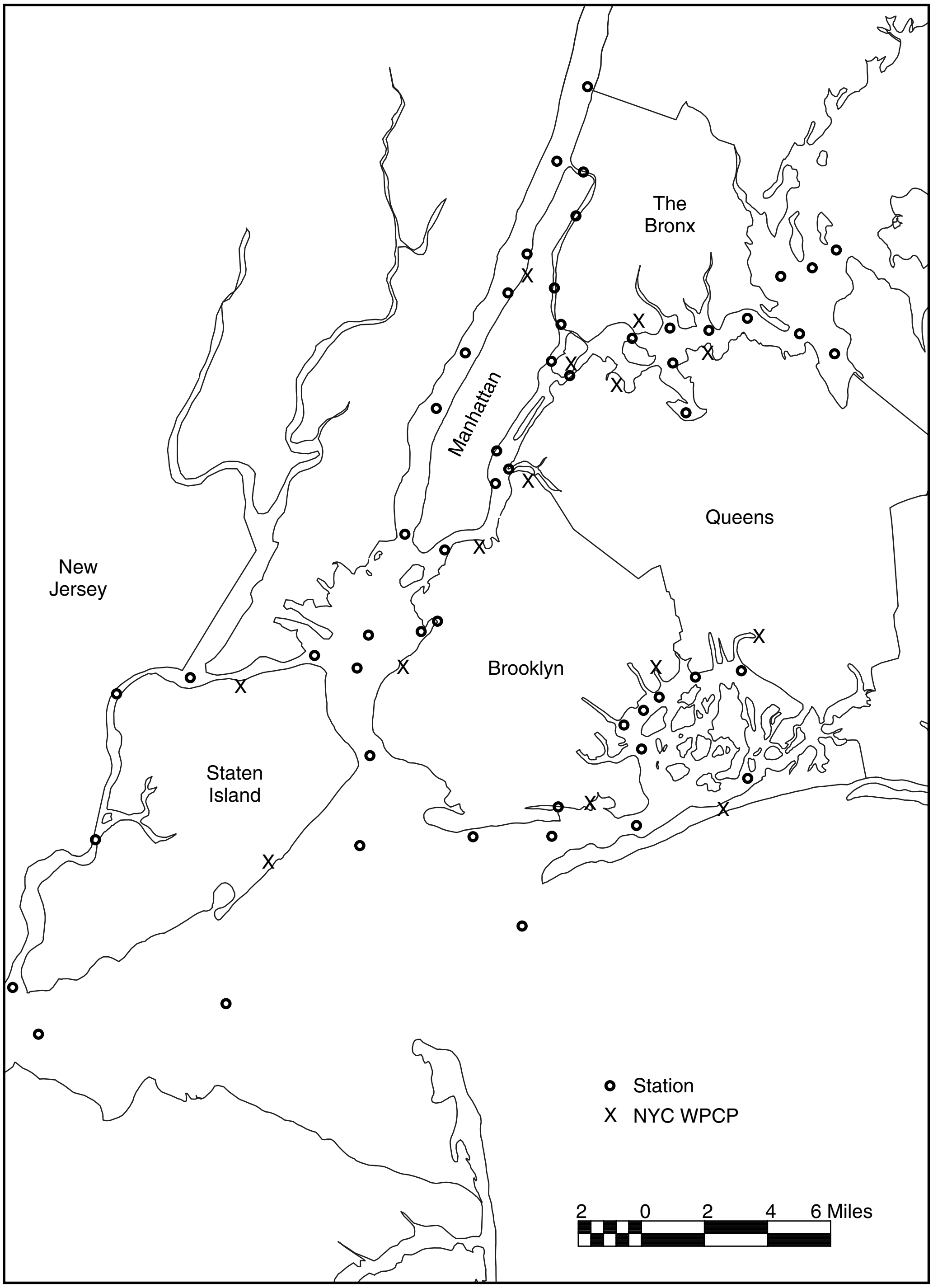

Figure 2. Harbor survey stations.

Beginning in 1989, NYCDEP began to self-report the results of the Survey; that year also marked the end of routine disposal of untreated sewage. NYCDEP had confidence that the Survey would show improvements in water quality, and so these reports could serve as public outreach to tout City environmental programs (27). Beginning in 1999, academic institutions prepared the longer of the two annual reports; the intent was an increase in rigor and a broadened perspective.

The NYCDEP summary reports show improving trends for all analytes nearly harborwide. The improvements have been so great that New York State water quality goals are being met routinely, as measured by DO and fecal coliform summer means, reported both as waterbody means and individual sample results (31).

SUNY at Stony Brook used a more sophisticated statistical analysis to identify trends. Its methodology also resulted in determinations of improved water quality since

1985. Order of magnitude reductions in fecal coliform and striking increases in the DO for nearly every station were found, and most of the linear trends were statistically significant at $p < .05$. The improvements were greatest in the late 1980s, probably from large reductions in raw sewage inputs as the entire City was connected to sewage treatment (28).

THE HARBOR SURVEY AND PUBLIC POLICY

One region where improvements in water quality are not apparent is western Long Island Sound. Long term trends for bottom DO do not show any improvements since the 1940s and may exhibit a declining trend (28). The Long Island Sound Study has identified hypoxia (low bottom DO) as the most serious problem in the Sound. A policy was set to reduce nitrogen inputs to the Sound by nearly 60%, intending to force higher bottom DO by reducing the impact of eutrophication (32).

The New York City East River WPCPs export most of their effluent (and associated nutrient loads) to the Sound. The Sound Study management plan requires multibillion dollar upgrades of the WPCPs to achieve tertiary treatment levels to reach agreed-upon nutrient reductions (33). Data from plant operations and the Harbor Survey show that 30% reductions in nitrogen loadings during the period 1992 to 2000 (caused by operational changes) concur with statistically significant declines in nitrogen concentrations in western Long Island Sound. However, the pattern of water column nitrogen concentrations does not correlate with changes in WPCP nitrogen releases, even using time lags of several months—in addition to the absence of increases in bottom DO in western Long Island Sound (28). The simple conceptual model, used by the Sound Study, that elevated wastewater nutrient concentrations in the East River cause increased phytoplankton concentrations in western Long Island Sound, leading to reductions in bottom DO due to microbial consumption of the deceased plankton (32), appears to be inadequate to explain the data. Cleorn (34), in an analysis of estuarine responses to eutrophication, showed that local physical conditions and biological community structures determine the particular reaction of waterbodies to elevated nutrient inputs. NYCDEP remains hopeful that its monitoring data may allow the City to find another way of addressing this complicated problem of low bottom DO [a situation that affects 25% of all estuarine waters in the northeast United States each summer (35)].

CONCLUSIONS

New York Harbor, a complex estuary, has been polluted by waste disposal for more than three centuries. For nearly 100 years, New York City has conducted a water quality monitoring program to measure impacts from sewage and WPCP effluent. This program, the Harbor Survey, documented the impacted water quality found in the Harbor before the City's wastewater treatment system was completed. Monitoring since 1985 has measured impressive improvements in parameters of concern. The program documents the effectiveness of the City's sewage treatment program, and it also serves as a public outreach and education tool and a means of helping to determine root causes of serious environmental problems in the Harbor.

BIBLIOGRAPHY

1. Melosi, M.V. (2000). *The Sanitary City*. The Johns Hopkins University Press, Baltimore, MD.
2. Metropolitan Sewerage Commission. (1912). *Present Sanitary Conditions of New York Harbor and the Degree of Cleanness which is Necessary and Sufficient for the Water*. Report of the Metropolitan Sewerage Commission, Wyncoop Halenbeck Crawford, Co., New York.
3. Pritchard, D.W. (1967). What is an estuary: physical viewpoints. *Estuaries*. AAAS Pub. 87, pp. 3–5.
4. Sanders, J.E. (1974). Geomorphology of the Hudson Estuary. In: *Hudson River Colloquium. Ann. NY Acad. Sci.* O.A. Roels (Ed.). V. 250, New York, pp. 5–38.
5. Moran, M.A. and Limburg, K.E. (1985). The Hudson River ecosystem. In: *The Hudson River Ecosystem*. K.E. Limburg, M.A. Moran, and W.H. McDowell (Eds.). Springer-Verlag, New York, pp. 7–39.
6. USGS Water Resources of New Jersey website. http//:nj.usgs.gov.
7. Oey, L.Y., Mellor, G.L., and Hires, R.I. (1985). Tidal modeling of the Hudson-Raritan estuary. *Estuarine Coastal Shelf Sci.* **20**: 511–527.
8. Abood, K.A. (1974). Circulation in the hudson river estuary. In: *Hudson River Colloquium. Ann. NY Acad Sci.* O.A. Roels (Ed.). V. 250, New York, pp. 39–111.
9. Duedall, I.W., O'Connors, H.B., Wilson, R.E., and Parker, J.H. (1979). *The Lower Bay Complex. MESA New York Bight Atlas Monograph 29*. New York Sea Grant Institute, Albany, NY.
10. Swanson, R.L, Parker, C.A., Meyer, M.C., and Champ, M.A. (1982). *Is the East River, New York, a River or Long Island an Island?* NOAA Technical Report, NOS 93, National Oceanic and Atmospheric Administration, Rockville, MD.
11. Blumberg, A.F. and Pritchard, D.W. (1997). Estimates of transport through the East River, New York. *J. Geophys Res.* **102**(C3): 5685–5703.
12. HydroQual, Inc. (1998). *A Water Quality Model for Jamaica Bay: Calibration of the Jamaica Bay Eutrophication Model (JEM)*. HydroQual, Inc., Mahwah, NJ.
13. Swanson, R.L., West-Valle, A.S., and Decker, C.J. (1992). Recreation vs. waste disposal: the use and management of Jamaica Bay. *Long Island Hist. J.* **5**(1): 21–41.
14. Juet, R. (1609). The third voyage of Master Henry Hudson. In: *Collections of the New York Historical Society in the Year 1809*. V. 1. 1811. I. Riley, New York, pp. 102–145.
15. US Army Corps of Engineers New York District website. http://www.nan.usace.army.mil/harbor.
16. Squires, D.F. (1990). A historical review of changes in near-shore habitats in the Sound-Harbor-Bight system. In: *Cleaning Up Our Coastal Waters: An Unfinished Agenda*. M.T. Southerland and K. Swetlow (Eds.). Dynamac Corp., Rockville, MD, pp. 403–428.
17. Waldman, J. (1999). *Heartbeats in the Muck*. The Lyons Press, New York.
18. Squires, D.F. and Bone, K. (1997). The beautiful lake: the promise of the natural systems. In: *The New York Waterfront:*

Evolution and Building Culture of the Port and Harbor. K. Bone (Ed.). Monacelli Press, New York, pp. 16–33.

19. Swanson, R.L. and Squires, D.F. (2002). Ellis Island, New York and New Jersey. *New York Hist.* **83**(3): 248–270.
20. Adams, A.G. (1996). *The Hudson Through the Years*, 3rd Edn. Fordham University Press, New York.
21. Cunningham, J.T. (1978). *This is New Jersey*, 3rd Edn. Rutgers University Press, New Brunswick, NJ.
22. Suszkowski, D.J. (1990). Conditions in the New York/New Jersey harbor estuary. In: *Cleaning Up Our Coastal Waters: An Unfinished Agenda*. M.T. Southerland and K. Swetlow (Eds.). Dynamac Corp., Rockville, MD, pp. 105–132.
23. US Census Bureau website. http//:www.census.gov.main/www/cen2000.
24. Miller, B. (2000). *Fat of the Land*. Four Walls Eight Windows, New York.
25. Suszkowski, D.J. (1973). *Sewage Pollution in New York Harbor: A Historical Perspective*. Masters Thesis, SUNY at Stony Brook, Stony Brook, NY.
26. Mueller, J.A. (1990). Toxic inputs and fate in the New York-New Jersey Harbor, Bight, and Long Island Sound. In: *Cleaning Up Our Coastal Waters: An Unfinished Agenda*. M.T. Southerland and K. Swetlow (Eds.). Dynamac Corp., Rockville, MD, pp. 317–354.
27. Brosnan, T.M. and O'Shea, M.L. (1996). Long-term improvements in water quality due to sewage abatement in the lower Hudson River. *Estuaries* **19**(4): 890–900.
28. Swanson, R.L. and Tonjes, D.J. (2001). *The New York City 2000 Regional Harbor Survey*. P. Heckler, B. Ranheim, N-J. Yao, and B.W. Stephens (Eds.). New York City Department of Environmental Protection, New York.
29. Swanson, R.L. and Tonjes, D.J. (2001). Water conservation cleans Long Island Sound. *Clearwaters* **31**(2): 8.
30. Bell, T.M., Schubel, J.R., and Swanson, R.L. (1989). *Floatable Wastes and the Region's Beaches: Answers to Some Common Questions*. Marine Sciences Research Center Special Report 85, SUNY at Stony Brook, Stony Brook, NY.
31. NYCDEP. (2001). *2000 New York Harbor Water Quality Regional Report*. New York City Department of Environmental Protection, New York.
32. LISS. (1994). *The Comprehensive Conservation and Management Plan*. The Long Island Sound Study, Stamford, CT.
33. LISS. (1997). *Proposal for Phase III Actions for Hypoxia Management*. EPA 840-R-97-001. The Long Island Sound Study, Stamford, CT.
34. Cleorn, J.E. (2001). Our evolving conceptual model of the coastal eutrophication problem. *Mar. Ecol. Prog. Ser.* **210**: 223–253.
35. USEPA. (2000). *Ambient Water Quality Criteria for Dissolved Oxygen (Saltwater): Cape Cod to Cape Hatteras*. United States Environmental Protection Agency, Washington, DC.

NITRIFICATION IN THE ACTIVATED SLUDGE PROCESS

Michael H. Gerardi
Linden, Pennsylvania

There are two reduced forms of nitrogen that are of concern to wastewater treatment plant operators and state and federal regulatory agencies. These forms contain hydrogen and consist of nonionized ammonia or ammonia (NH_3) and ionized ammonia or ammonium ion ($NH_4{}^+$). Although ammonia and ammonium ions are both present in wastewater, the amount of each is dependent on the pH of the wastewater (Fig. 1).

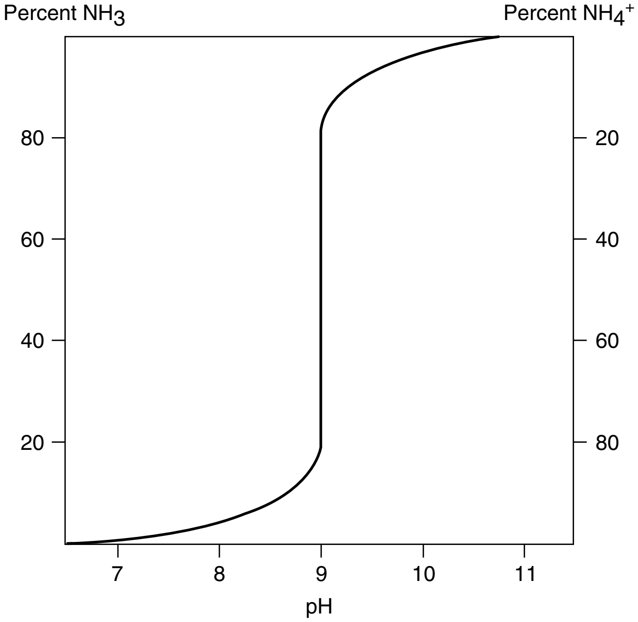

Figure 1. Relative distribution of ammonia and ammonium ions as determined by pH.

Wastewater nitrification is the biological oxidation of ammonium ions to nitrite ions ($NO_2{}^-$) (Eq. 1) or the biological oxidation of nitrite ions to nitrate ions ($NO_3{}^-$) (Eq. 2).

$$NH_4{}^+ + 1.5\ O_2 \xrightarrow{\textit{Nitrosomonas}} NO_2{}^- + 2\ H^+ + H_2O + \text{energy} \tag{1}$$

$$NO_2{}^- + 0.5\ O_2 \xrightarrow{\textit{Nitrobacter}} NO_3{}^- + \text{energy} \tag{2}$$

Nitrification occurs because nitrifying bacteria oxidize ammonium ions and nitrite ions in order to obtain cellular energy. There are two principal genera of nitrifying bacteria, *Nitrosomonas* and *Nitrobacter*. Although both genera of bacteria are strict aerobic organisms, *Nitrosomonas* can oxidize only ammonium ions, and *Nitrobacter* can oxidize only nitrite ions.

There are four nitrogen-containing molecules that are involved in wastewater nitrification. These molecules include ammonia, ammonium ions, nitrite ions, and nitrate ions. Because nitrite ions are highly unstable, they usually do not accumulate in wastewater and are considered to be a short-lived intermediate molecules that are quickly converted to nitrate ions. However, there are operation conditions that do permit the accumulation of nitrite ions (Table 1).

Ammonium ions enter activated sludge processes from nitrogenous compounds that are found in domestic and industrial wastewater. Domestic wastewater contains urea, amino acids, and proteins. Urea and simple amino acids degrade in the sewer system through hydrolysis

Table 1. Operational Conditions that Permit the Production and Accumulation of Nitrite Ions

Operational Condition	Description
Depressed temperature	$\leq 15\,^{\circ}C$
Limiting process conditions	Change in pH Inhibition Low dissolved oxygen level Slug discharge of soluble cBOD

and deamination, respectively. Once degraded, these compounds release ammonium ions. Due to the hydrolysis of urea and deamination of simple amino acids, most municipal activated sludge processes contain 25–30 mg/L of influent ammonium ions.

Stable amino acids and proteins degrade in the aeration tank and release ammonium ions. Urea, amino acids, and proteins are organo-nitrogen compounds. Additional organo-nitrogen compounds that are found and degraded in aeration tanks include surfactants and polymers. Ammonium ions and organo-nitrogen compounds that release ammonium ions can be found in many industrial wastewaters that are discharged to activated sludge processes (Table 2). Nitrite ions and nitrate ions that are found in activated sludge processes usually are produced in the aeration tank. However, there are some industrial dischargers of nitrite ions and nitrate ions (Table 2).

Activated sludge processes that are required to satisfy an ammonia discharge limit or total nitrogen discharge limit must nitrify. Although many activated sludge processes are not required to nitrify, operators of these processes may promote nitrification to ensure process control or use the nitrate ions produced through nitrification during controlled anoxic (denitrification) periods to obtain improved floc particle structure and decrease operational costs (Table 3). Finally, many activated sludge processes that are not required to nitrify may "slip" into nitrification. Regardless of permit requirements, activated sludge processes may enter an undesired form of nitrification that results in increased operational costs, process upsets, and permit violations.

Table 2. Industrial Discharges of Ammonium Ions, Nitrite Ions, and Nitrate Ions

	Nitrogenous Compound		
Industrial Discharge	NH_4^+	NO_2^-	NO_3^-
Automotive	X		
Chemical	X		
Coal	X		
Corrosion inhibitor		X	
Fertilizer	X		
Food	X		
Leachate	X		
Leachate (pretreated)		X	X
Livestock	X		
Meat	X		
Meat (flavoring)			X
Meat (preservative)		X	
Meat (pretreated)		X	X
Ordnance	X		
Petrochemical	X		
Pharmaceutical			X
Primary metal	X		
Refineries	X		
Steel	X	X	X
Tanneries	X		

Regardless of the presence or absence of a nitrification requirement, nitrogenous compounds are of concern to operators of activated sludge processes (Table 4) and regulatory agencies (Table 5) due to their undesired impacts on the activated sludge process and receiving body of water, respectively. The nitrogenous compounds of concern consist of inorganic compounds (NH_3, NH_4^+, NO_2^-, and NO_3^-) and organic-nitrogen compounds (total kjeldahl nitrogen or TKN).

There are several forms of nitrification that can occur in activated sludge processes (Table 6). These include one complete and four incomplete forms. The identification of the form of nitrification that occurs in the activated sludge process is of value to an operator to ensure acceptable nitrification and correct undesired incomplete nitrification.

If nitrification occurs as incomplete #3 or incomplete #4, nitrite ions are produced. The accumulation of these

Table 3. Benefits Obtained Through the Use of Controlled Anoxic Periods (Denitrification)

Decrease sludge production
Destroy undesired filamentous organism growth
Improve process control: ensure adequate cBOD removal
Improve process control: ensure the presence of a "healthy" biomass
Return alkalinity to the treatment process
Strengthen floc particles

Table 4. Nitrogenous Compounds of Concern to Operators of Activated Sludge Processes

Compound	Formula	Impact
Ammonia	NH_3	Toxicity
Ammonium ion	NH_4^+	Oxygen demand upon nitrification to NO_2^- Primary nitrogen nutrient for bacterial growth Toxicity upon conversion to NH_3
Nitrite ion	NO_2^-	Denitrification ("clumping") in the secondary clarifier Increased chlorine demand (chlorine "sponge") Oxygen demand upon nitrification to NO_3^- Toxicity
Nitrate ion	NO_3^-	Denitrification ("clumping") in the secondary clarifier Secondary nitrogen nutrient for bacterial growth Toxicity upon reduction to NO_2^- by *E. coli*
Organic-nitrogen(TKN)		Oxygen demand upon degradation Release of cBOD upon degradation Release of nBOD (NH_4^+) upon degradation

Table 5. Nitrogenous Compounds of Concern to Regulatory Agencies

Compound	Formula	Impact
Ammonia	NH_3	Toxicity
Ammonium ion	NH_4^+	Oxygen demand upon nitrification to NO_2^- Toxicity upon conversion to NH_3
Nitrite ion	NO_2^-	Oxygen demand upon nitrification to NO_3^- Toxicity
Nitrate ion	NO_3^-	Causative agent for methemoglobinemia Primary nitrogen nutrient for aquatic plants Undesired growth of aquatic plants, especially algae
Organic-nitrogen(TKN)		Oxygen demand upon degradation Release of cBOD upon degradation Release of nBOD (NH_4^+) upon degradation

Table 6. Forms of Nitrification

Form of Nitrification	Mixed Liquor Effluent Filtrate Concentration, mg/L		
	NH_4^+	NO_2^-	NO_3^-
Complete	<1	<1	As great as possible
Incomplete #1	<1	As great as possible	<1
Incomplete #2	>1	<1	>1
Incomplete #3	<1	>1	>1
Incomplete #4	>1	>1	>1

ions in the activated sludge process is known as the chlorine "sponge," nitrite "kick," and nitrite "lock," because nitrite ions react quickly with chlorine and consume large quantities of chlorine. Approximately 13 pounds of chlorine are consumed per mg/L NO_2^- produced and accumulated per million gallons of flow. Nitrite ions in the activated sludge process interfere with the destruction of filamentous organisms in the mixed liquor or return activated sludge (RAS) via chlorination. Nitrite ions also interfere with the destruction of indicator organisms and pathogens in the final effluent via chlorination. To correct for the chlorine sponge, appropriate operational measures should be used to compensate for depressed temperature or correct for the responsible limiting factor.

Forms of incomplete nitrification occur as a result of depressed wastewater temperature (<15 °C) and limiting process conditions. Limiting process conditions include (1) change in pH, (2) a slug discharge of soluble cBOD, (3) temporary low dissolved oxygen level, and (4) toxicity. Additional factors that influence nitrification include alkalinity deficiency, decreased hydraulic retention time (HRT), and phosphorus deficiency. By identifying and correcting the condition responsible for incomplete nitrification, the return of acceptable nitrification should occur within 2–3 d.

The adverse impact of depressed wastewater temperature upon nitrification can be overcome through the use of appropriate changes in operational conditions. These changes include increasing the HRT of the aeration tanks, removing more colloidal and particulate BOD in the primary clarifiers, increasing dissolved oxygen concentration in the aeration tanks, and using bioaugmentation products. These products consist of commercially prepared cultures of nitrifying bacteria and organotrophic (saprophytic) bacteria that remove carbonaceous BOD quickly.

There are several significant operational conditions that impact nitrification in the activated sludge process. These conditions include temperature, mean cell residence time (MCRT), dissolved oxygen concentration, inhibition, alkalinity, and pH. Of these conditions temperature and MCRT are the most critical.

Due to the relatively small quantity of energy obtained from the oxidation of nitrogenous compounds, nitrifying bacteria reproduce very slowly. The generation times for *Nitrosomonas* and *Nitrobacter* under optimal laboratory conditions are approximately 8 and 10 h, respectively. In the activated sludge process the generation time of these organisms is estimated to be approximately 2–3 days. Therefore, high MCRTs are required at activated sludge processes to permit the growth of large numbers of nitrifying bacteria. However, warm wastewater temperatures provide for increased bacterial activity. Therefore, the number of nitrifying bacteria and the MCRT can be reduced. However, cold wastewater temperatures provide for decreased bacterial activity, and the number of nitrifying bacteria and the MCRT must be increased (Table 7).

Because nitrifying bacteria are strict aerobes, nitrification in the activated sludge process is influenced by dissolved oxygen concentration in the aeration tank. Nitrification is considered to proceed efficiently within the dissolved oxygen range of 2–3 mg/L. At dissolved oxygen concentration above 3 mg/L, nitrification may become more efficient, if carbonaceous BOD is more quickly removed from the aeration tank. If carbonaceous BOD is more quickly removed, then more hydraulic retention time is provided for nitrification. Approximately 4.6 pounds of dissolved oxygen are consumed for each pound of ammonium ions oxidized to nitrate ions (Table 8).

Generally, whatever is inhibitory to the organotrophic bacteria (cBOD-removing bacteria) also is toxic to nitrifying bacteria and is toxic to nitrifying bacteria at lower values. The more sensitive nature of nitrifying bacteria to inhibition is due to the relatively small quantity of energy obtained by nitrifying bacteria from the oxidation of nitrogenous compounds as compared to the quantity of energy obtained by organotrophic bacteria from the oxidation of carbonaceous compounds. The relatively small quantity of energy available to nitrifying bacteria

Table 7. Temperature and MCRT Recommended for Nitrification

Temperature, °C	MCRT, days
10	30
15	20
20	15
25	10
30	7

Table 8. Approximate Quantity of Oxygen Consumed During Nitrification

Biochemical Reaction	O_2, lb	O_2, kg
1 kg NH_4^+ to 1 kg NO_2^-		1.6
1 lb NH_4^+ to 1 lb NO_2^-	3.4	
1 kg NO_2^- to 1 kg NO_2^-		0.5
1 lb NO_2^- to 1 lb NO_2^-	1.2	
1 kg NH_4^+ to 1 kg NO_3^-		2.1
1 lb NH_4^+ to 1 lb NO_3^-	4.6	

limits their ability to repair cellular damage caused by inhibitory conditions.

There are several forms of inhibition that can damage nitrifying bacteria and adversely affect nitrification in the activated sludge process (Table 9). Those that commonly occur include inorganic (Table 10), organic (Table 11), and substrate inhibition. A unique form of inhibition that affects nitrifying bacteria is "recognizable," soluble carbonaceous BOD inhibition.

Substrate inhibition occurs through the production of free ammonia (NH_3) at elevated pH values and nitrous acid (NNO_2) at depressed pH values. Free ammonia and free nitrous acid are toxic to nitrifying bacteria. In the presence of relatively high ammonium ion concentrations (>500 mg/L) in the aeration tank, either free ammonia is produced with increasing pH or free nitrous acid is produced with decreasing pH. Therefore, the discharge of high concentrations of ammonium ions to the aeration tank or the release of high concentrations of ammonium ions in the aeration tank should be prevented.

Table 9. Forms of Inhibition that Affect Nitrifying Bacteria and Nitrification

Form of Inhibition	Description or Example
Free chlorine residual	Chlorination of the mixed liquor
Inorganic	Cyanide or heavy metals
Organic	Phenols or recognizable, soluble carbonaceous BOD
pH	<5
Substrate	Free ammonia or free nitrous acid
Sunlight	Ultraviolet radiation
Temperature	<5 °C

Table 10. Inhibitory Threshold Concentrations of Some Inorganic Wastes that Affect Nitrifying Bacteria and Nitrification

Inorganic Waste	Concentration, mg/L
Chromium (hexavalent)	0.25
Chromium (trivalent)	0.05
Copper	0.35
Cyanide	0.5
Mercury	0.25
Nickel	0.25
Silver	0.25
Sulfate	500
Zinc	0.3

Table 11. Inhibitory Threshold Concentrations of Some Organic Wastes that Affect Nitrifying Bacteria and Nitrification

Organic Waste	Concentration, mg/L
Allyl alcohol	20
Aniline	8
Chloroform	18
Mecaptobenzothiazole	3
Phenol	6
Skatol	7
Thioacetamide	0.5
Thiourea	0.1

Recognizable, soluble carbonaceous BOD inhibition occurs in the presence of simplistic organic compounds (Table 12). These compounds are capable of inhibiting enzymatic activity within the nitrifying bacterial cell. This inhibitory effect is not resolved until the organic compounds are either reduced in concentration or removed completely from the waste stream. Many of these inhibitory organic compounds are produced through fermentative reactions that occur in the biofilm and sediment within sewer systems and help to prevent nitrification in the sewer system. Additional factors that prevent nitrification in the sewer system include lack of an adequate population of nitrifying bacteria, short retention time, and absence of dissolved oxygen or a relatively small quantity of dissolved oxygen.

Alkalinity is removed and destroyed during nitrification. Alkalinity is removed from the wastewater by nitrifying bacteria as a source of carbon for growth, repair, and, most importantly, reproduction. Alkalinity is destroyed in the wastewater through the production of nitrous acid by *Nitrosomonas* during the first biochemical reaction of nitrification. As alkalinity is removed and destroyed, the pH of the aeration tank decreases.

Approximately 7.14 mg of alkalinity are consumed for each mg of ammonium ion oxidized to nitrate ion. To ensure adequate alkalinity for proper nitrification, the mixed liquor effluent of the aeration tank should contain at least 50 mg/L of alkalinity after complete nitrification. If an alkalinity deficiency exits for proper nitrification, alkalinity can be added to the mixed liquor with the use of an appropriate chemical (Table 13).

The pH of the aeration time affects enzymatic activity of the nitrifying bacteria. Activated sludge processes

Table 12. Examples of Recognizable, Soluble Carbonaceous BOD

Organic Compound	Formula
Aminoethanol	$CH_3NH_2CH_2OH$
Ethanol	CH_3CH_2OH
Ethyl acetate	$CH_3COC_2H_5$
i-Propanol	$(CH_3)_2CHOH$
Methanol	CH_3OH
Methylamine	CH_3NH_2
n-Butanol	$CH_3(CH_2)_2CHOH$
n-Propanol	$CH_3CH_2CH_2OH$
t-Propanol	$(CH_3)_3COH$

Table 13. Chemicals Suitable for Alkalinity Addition

Chemical Name	Formula	Common Name
Sodium bicarbonate	$NaHCO_3$	Baking soda
Calcium carbonate	$CaCO_3$	Calcite, limestone, whiting, chalk
Sodium carbonate	Na_2CO_3	Soda ash
Calcium hydroxide	$Ca(OH)_2$	Lime
Magnesium hydroxide	$Mg(OH)_2$	Magnesia
Sodium hydroxide	NaOH	Caustic soda

that nitrify efficiently do so at near neutral pH values (6.8–7.2). Nitrification at pH values greater than 7.2 are used, but nitrifying bacteria adjust slowly to significant swings in pH.

Nitrification can be achieved in one-stage systems or two-stage systems. One-stage systems consist of one aeration tank or a series of aeration tanks that remove carbonaceous BOD and also nitrify. Two-stage systems consist of a series of tanks. The first tanks in the series remove carbonaceous BOD only. The last tanks in the series nitrify only. Two-stage systems provide better process control than one-stage systems, and with increasing regulatory requirements to nitrify more and more activated sludge processes, especially in the temperate regions of North America, may need to nitrify using two-stage systems. Regardless of the nitrification requirements placed upon activated sludge processes, a knowledge of the biology of nitrification and its application to activated sludge processes are of value in achieving and maintaining cost-effective nitrification in compliance with permitted discharge limits.

EFFLUENT LIMITATIONS AND THE NPDES PERMIT

L. Donald Duke
University of South Florida
Tampa, Florida

THE CLEAN WATER ACT AND INDUSTRIAL DISCHARGES

The Federal Water Pollution Control Act of 1972 was enacted by Congress with the stated objective "to restore and maintain the chemical, physical, and biological integrity of the Nation's waters" (1). The act acquired its common name, the Clean Water Act (CWA), in its first reauthorization by Congress in 1977. CWA in its essence can be considered in three conceptual parts: funding for local agencies to construct wastewater treatment plants; requirement for the states to develop water quality standards, using a set of criteria developed under CWA as goals for the chemical constituent conditions of waters of the United States; and creation of a series of regulations for permitting wastewater discharges from industrial facilities and from treatment plants. The CWA enacted a permit requirement for wastewater discharges, prohibiting discharges in the absence of a permit, and developed a permit system with roles for federal, state, and local agencies. CWA requirements have teeth: substantial financial and criminal penalties are authorized for discharging without a permit as well as for violation of the permits, with enforcing agencies directed to apply a sliding scale related to the severity of a violation, up to $25,000 per day during which the discharge is in violation, and higher fines plus jail time for violations found to be negligent, intentional, and/or to cause injury to humans (2).

The three parts are designed to work together to drive toward protection and improvement of water quality nationwide. The first, funding (and associated regulation) for locally-operated wastewater treatment plants, created an extensive physical and institutional infrastructure over time that implements controls on household, municipal, and industrial wastewater discharges to waters of the United States. The Publicly Owned Treatment Works (POTW) became an instrument of the Clean Water Act, accountable to federal requirements but operated by local agencies and exercising local control over a local network of dischargers.

The second conceptual part of CWA initiated a system of ambient standards for waters of the United States or conditions in which any given water body can be documented to attain those uses the public chooses for that water body. A standard is defined to consist of two parts: a designated beneficial use for the water body and the physical conditions under which that use is supported. The embodied concept of protecting waters for particular, specified uses continues to drive water quality goals in the United States. The conditions determined to meet those uses typically are derived from the federally-developed criteria—logically avoiding conducting the same research multiple times when each state chooses conditions for specified designated uses. As of 1999, water quality criteria had been developed for some 115 pollutants, including 65 named classes or categories of priority (toxic) pollutants (2). In this way, standard development incorporates the federal role of providing scientific information on satisfactory conditions for attaining a given use (e.g., chemical, temperature, and flow requirements to support particular species of sports fisheries) and states' preferences in selecting designated uses for each water body within their borders (although this is limited by a CWA stricture against allowing degradation, which effectively precludes failure to protect designated uses in place as of 1975, without extensive review and justification).

The third conceptual part of CWA is a system of permits for wastewater discharges, applicable to POTWs; industrial wastewater discharges; and other miscellaneous discharges. The controls take a form known as "uniform numeric limitations," and as such are designed to require facilities of a given type nationwide to attain a common standard, recognizing wastewater quality as an equal goal, and equal burden, for discharges in all parts of the United States. The uniform limit has a powerful theoretical basis in the scheme of regulatory systems, which is that by specifying a numeric target rather than a selected technology, dischargers are encouraged to identify the least-cost approach that would attain that target for their own industrial facility, which might be substantially different from the treatment method that serves best for other facilities. This structure provides an incentive to promote research and development of improved, lower cost treatment technologies that would attain the same target in a more effective manner. That incentive is imperfect, as long as the limit remains greater than zero, because research and development is encouraged only to the extent of identifying less costly means to attain the specified limit. In contrast, economic approaches, such as emission trading or charges-and-standards, also encourage research and development on the most cost-effective ways to reduce discharges of pollutants extending below the specified standard, in principal also encouraging research into low-emission or zero-discharge technologies that are not equally encouraged by the numeric limit approach. The uniform numeric limitation approach also foregoes the advantages of a competing system, one that could attain greater economic efficiency by allowing trading pollutants among different dischargers with different attainment costs, of the kind employed by certain Clean Air Act regulations. An extensive body of environmental economics literature addresses these concepts; see, for example, Tietenberg (3).

CWA engendered an interwoven system of regulations and policy goals spanning federal, state, and local agencies across the United States. The Act was envisioned as dynamic, to be revised and refined over time, taking advantage of research advances and the expected water quality improvements; and to some extent, that has occurred. Two major reauthorizations, in 1977 and 1987, encompassed substantial changes, and periodic rulemaking has introduced new directions, such as the emphasis on watershed management in the early 1990s. However, since that time, little of substance has changed (4). Annual updates on CWA and related issues can be found in the Congressional Research Service's Issue Brief series [e.g., (5)]. A legislative and conceptual overview of CWA, its history and intent, can be found in (4) and in a 25-year summary prepared by the Water Environment Federation (1). The entire Clean Water Act may be seen in the Code of Federal Regulations (CFR) under Title 40, Protection of the Environment; Chapter I, Environmental Protection Agency; sections 401–499, Clean Water Act. A searchable version of the CFR is maintained online by the National Archives and Records Administration, updated quarterly as regulations are modified or expanded. The website for CFR Title 40 is http://www.gpoaccess.gov/topics/environment.html.

NPDES: CWA'S NUMERIC EFFLUENT SYSTEM

The National Pollutant Discharge Elimination System (NPDES) is the mechanism developed in the CWA to specify numeric limitations for effluent from both industrial facilities and POTWs. In the 1972 CWA, Congress authorized the newly-organized Environmental Protection Agency (EPA) to develop, promulgate, and enforce NPDES requirements.

Uniform numeric limitations for industry are specified in two forms. The first is by substance: EPA is directed to develop a list of allowable limitations for both "conventional pollutants," which were widely recognized as of 1972, and "priority pollutants," consisting largely of toxic substances about which information was and is still emerging. The second is by industrial category: numeric limitations are listed for each of a range of industry groups (categories) selected by EPA according to facility types likely to have similar wastewater characteristics and problems, and similarity of the treatment technologies likely to be successful in removing pollutants from wastewater typical of that industry's characteristics. These categorical standards within a given industry recognize two types of discharges: direct discharges into waters of the United States and indirect discharges into municipal wastewater systems, known as "pretreatment" standards, because after onsite treatment, the discharges are intermingled and treated again by the receiving POTW. For some substances in some industry categories, the specified numeric limitations are somewhat different for direct and for indirect discharges.

In order to be uniform across the United States, logically it is not possible to select numeric effluent limitations based on water quality. No uniform standard could address the wide variation of water bodies' assimilative capacities, mixing properties, biological communities, and other considerations or the wide variation of discharges' volumes, timing, intermingled substances, and similar characteristics. The NPDES rules were expected to attain water quality improvement simply by specifying a uniform target for all discharges that would require industrial facilities, and POTWs, to provide some treatment to remove pollutants recognized as potentially harmful to the environment. Therefore, the NPDES limitations were of the type we can term "technology based": the numeric limitations were selected by EPA after a thorough review of the kinds of wastewater treatment technologies available, those known to have been constructed and operated by at least some facilities in the United States. The operating characteristics of the technologies provided knowledge of the attainable chemical characteristics of wastewater after treatment and some information about the cost to treat wastewater to achieve a specified standard. EPA was directed to use that information to specify a numeric limit, not a technology that all dischargers must adopt, but a number that all discharges must attain, based on the conclusion that available technologies could successfully attain the number.

Both the substance-specific and the category-specific limitations achieve the same conceptual advantages of any numeric limitation, i.e., the discharger (and the entity responsible for wastewater treatment) may select the most cost-effective technology for his own situation, encouraging innovative research and development to develop lower cost processes if possible. This concept today remains central to the goals of EPA and NPDES: "A facility's drive to identify cheaper, more effective ways to achieve compliance is consistent with EPA's mission of clarifying and simplifying environmental regulatory control" [Sector Notebook series, (6)]. The structure of NPDES regulations does not permit EPA to specify treatment technologies, instead requiring EPA to specify a numeric target (although EPA is allowed to justify its selection of a limitation by documenting that a cost-effective treatment exists that would attain the limitation).

The technology-based structure does, however, generate another conceptual weakness in the NPDES numeric limitation scheme: Limitations based on current technology have the effect of "freezing" the technology at that level. Research and development on lower cost ways to attain the specified limitation are strongly encouraged, but no incentive exists to invest in developing or implementing improved technology that would attain any greater pollutant removal than specified in the limitation. It may be inferred that Congress' intent under CWA was to routinely revise the limitations as more information becomes available through continuing research on the effects and treatability of toxic materials. (In fact, as the term "elimination" in the title suggests, Congress envisioned that industry and POTWs would attain effectively zero discharge of substances found to cause water quality problems in receiving waters.) The history of CWA reauthorizations does not show a robust program of routinely revisiting and revising the limitations for most substances once they have been initially selected.

EFFLUENT LIMITATIONS BY SUBSTANCE

EPA developed the numerical limitations for its NPDES requirements using information about technologies available during the 1970s, when the requirements were written, as well as the best information at the time about technological capabilities expected in the future. The initial set of limitations focused on the "conventional" pollutants and on the kinds of treatment technologies then well known and proven to be economically feasible for typical industrial facilities; these were known as "best practicable control technologies" (BPT). CWA specified a timeframe by which dischargers needed to meet the BPT requirements, which were, of course, specified in the form of a numeric limit, not a specified technology. Conventional pollutants subject to BPT were:

—Suspended solids

—Biochemical oxygen demanding substances (BOD)

—Fecal coliform

—Acidity (pH)

A second set of limits was promulgated that were to be attained on a later timeframe and were termed "best available technology" found to be economically achievable (BAT). These BAT requirements also included more stringent regulations for conventional pollutants. That feature places NPDES into the subcategory of technology-based rules known as "technology forcing," because the increasingly stringent requirements as a rule tended to drive dischargers to develop increasingly effective treatment technologies. The BAT requirements focused on limitations for nonconventional pollutants (heavy metals, pesticides, and other organic chemicals), including a list of 125 "priority" or toxic pollutants. Priority toxic pollutants and their effluent limitations appear in the Clean Water Act section 307(a).

EFFLUENT LIMITATIONS BY INDUSTRIAL CATEGORY

Further BPT and BAT point-source limitations are written for certain substances in discharges from certain industries, specified separately for each of a number of industrial categories. These are uniform within the industry, but may differ from the broader uniform limitations in any of a number of ways: some are more stringent requirements for the same substance; some impose mass-based limitations in addition to the uniform concentration-based limitations; and others specify requirements for different substances, common to the industry category but not included in the broader uniform requirements. These requirements, in general, encompass limitations for both direct and indirect point-source discharges.

Table 1 lists the categories of industrial wastewater discharges subject to NPDES categorical effluent limitations, along with the CFR section where those limitations can be found (for instance at the website given above). Using that information, the searchable CWA referenced above is a good way to locate current categorical standards. The industry categories of NPDES are not grouped according to the familiar Standard Industrial Classification (SIC) system, but instead use the groups developed by EPA in its CWA analyses.

For many of these industry categories, a number of useful descriptive and advisory documents have been prepared by EPA, other government agencies, and a range of private sector and industry-association groups. Two useful sets of documents from EPA are the Sector Notebooks and the Compliance Assurance documents, developed by EPA in the mid- to late-1990s. The last column in Table 1 includes the Federal report number of those documents that address particular NPDES categories, where such documents have been developed. Those guides do not contain the full text of the regulations and are not to be understood as the complete list of actions required by the regulated community, so the reader should refer to complete regulations to ensure compliance. Instead, they summarize the totality of regulatory requirements of which an operator should be aware. These are perhaps at their most useful in summarizing the range of environmental regulations that apply to a particular sector, including placing CWA requirements in the context of other regulations and listing the kinds of measuring, reporting, and documenting that can fulfill one or several of these requirements singly or jointly. Counter to the NPDES approach, many of those useful documents do follow the traditional SIC categorization, so the facility operator needs to find his or her own activities using a broad and flexible definition of his own processes.

Information about other industries, not named as CWA point-source categories, includes:

Dry cleaning	EPA/310-R-95-001
Wood furniture and fixtures	EPA/310-R-95-003
Motor vehicle assembly	EPA/310-R-95-009
Printing	EPA/310-R-95-014
Air transportation	EPA/310-R-97-001
Ground transportation	EPA/310-R-97-002
Water transportation	EPA/310-R-97-003
Shipbuilding and repair	EPA/310-R-97-008
Agricultural crop production	EPA/310-R-00-001
Agricultural livestock production	EPA/310-R-00-002
Aerospace	EPA/310-R-98-001

Table 1. Industry Sectors Subject to Category Numeric Limitations for Wastewater Discharges; Sources of Information

Industry Sector (Point Source Category)	CFR Section (Effluent Limitations Listed)	Additional Information*
Dairy products processing	405	
Grain mills	406	
Canned and preserved fruits and vegetables processing	407	
Canned and preserved seafood processing	408	
Sugar processing	409	
Textile mills	410	EPA/310-R-97-009
Cement manufacturing	411	EPA/310-R-95-017 (stone, clay, glass, and concrete)
Concentrated animal feeding operations (CAFO)	412	EPA/310-R-00-002017 (agricultural livestock production)
Electroplating	413	
Organic chemicals, plastics, and synthetic fibers	414	EPA/310-R-95-012 (organic chemicals); EPA/310-R-97-006 (plastic resin, manmade fibers); EPA/310-R-95-016 (rubber and plastic)
Inorganic chemicals manufacturing	415	EPA/310-R-95-006 (lumber and wood products)
Soap and detergent manufacturing	417	
Fertilizer manufacturing	418	EPA/310-R-00-003 (agricultural chemical, pesticide, and fertilizer industry)
Petroleum refining	419	EPA/310-R-95-013
Iron and steel manufacturing	420	EPA/310-R-95-005
Nonferrous metals manufacturing	421	EPA/310-R-95-010
Phosphate manufacturing	422	
Steam electric power generating	423	EPA/310-R-97-007
Ferroalloy manufacturing	424	
Leather tanning and finishing	425	
Glass manufacturing	426	EPA/310-R-95-017 (stone, clay, glass, and concrete)
Asbestos manufacturing	427	EPA/310-R-95-017 (stone, clay, glass, and concrete)
Rubber manufacturing	428	EPA/310-R-95-016 (rubber and plastic)
Timber products processing	429	EPA/310-R-95-006 (lumber and wood products)
Pulp, paper, and paperboard	430	EPA/310-R-95-015
Meat products	432	
Metal finishing	433	
Coal mining	434	
Oil and gas extraction	435	EPA/310-R-99-006
Mineral mining and processing	436	EPA/310-R-95-011 (non-fuel, non-metal mining)
Centralized waste treatment	437	
Metal products and machinery	438	EPA/310-R-95-007 (fabricated metal products)
Pharmaceutical manufacturing	439	EPA/310-R-97-005
Ore mining and dressing	440	EPA/310-R-95-008 (metal mining)
Transportation equipment cleaning	442	EPA/310-R-95-018
Paving and roofing materials	443	EPA/310-R-95-017 (stone, clay, glass, and concrete)
Waste combustors	444	
Landfills	445	
Paint formulating	446	EPA-305-S-97-005 (paint, . . ., and allied products manufacturing)
Ink formulating	447	
Gum and wood chemicals	454	
Pesticide chemicals	455	EPA/310-R-00-003 (agricultural chemical, pesticide, and fertilizer industry)
Explosives manufacturing	457	
Carbon black manufacturing	458	
Photographic	459	
Hospitals	460	
Battery manufacturing	461	
Plastics molding and forming	463	
Metal molding and casting	464	EPA/310-R-97-004
Coil coating	465	
Porcelain enameling	466	
Aluminum forming	467	
Copper forming	468	
Electrical and electronic components	469	EPA/310-R-95-002
Nonferrous metals forming and metal powders	471	

*Sector Notebook Project, developed by EPA's Office of Compliance, maintained online at http://www.epa.gov/compliance/resources/publications/assistance/sectors/notebooks/.

NPDES AND WATER QUALITY-BASED LIMITATIONS

Since its inception, CWA has included the stated intent that water quality considerations would overlay the uniform effluent limitations and, where necessary to attain ambient standards, would justify more stringent requirements for particular discharges, which was to be done by the states' routinely assessing the attainment of water quality standards in receiving waters. In cases where uniform limitations had been fully attained, and the water quality continued to fall short of state-specified standards, the states were to develop plans that would lead to attaining the specified ambient standards in all waters. NPDES is envisioned as a tool to attain the standards: states and/or POTWs may issue more restrictive NPDES limitations than the uniform effluent limits. This condition is identified as water quality-based effluent limitations.

For a variety of institutional reasons, this approach was applied only to a limited extent over roughly the first 20 years of CWA implementation. Since the mid-1990s, largely driven by court findings that the states should be held to water quality-based protections, EPA and the states have been busily developing programs to attain ambient standards, in large part within the Total Maximum Daily Load (TMDL) program of watershed-based planning and water quality protection. That program is addressed more fully in other parts of this encyclopedia. From the standpoint of effluent limitations on industrial wastewater discharges, the key aspect of this program is that specified limitations under NPDES permits may be made more stringent than the uniform limitations specified under CWA regulations. Selection of those limitations is guided not by economic considerations, but wholly by the requirement to attain water quality standards of the effluent's receiving waters and of systems connected to those waters.

STORM WATER EFFLUENT LIMITATIONS

The discharge of storm water runoff appears as a separate category of effluent limitation under CWA, except for a small number of industry categories (petroleum refining, for example) where storm water is included in wastewater to be collected, treated, and discharged under the wastewater NPDES permits. For most industry categories, storm water runoff requirements are a separate set of rules promulgated pursuant to the 1987 reauthorization of CWA, effective in 1992. Storm water runoff was identified in the original CWA as a nonpoint source, separate from the NPDES regulations designed for point sources, and set aside for regulation under future rulemaking. After a court finding in the 1980s that storm water can be considered a point source—in many cases, it does, after all, enter waters of the United States from a discrete channel or pipe leaving a municipal system or an industrial facility—EPA was required to include discharges of storm water runoff under NPDES.

The approach EPA has adopted for storm water runoff is substantially different from that of other industrial point sources in at least two respects. One is that EPA does not specify or require numeric effluent limitations on storm water discharges; instead, storm water discharges both by industrial facilities and by local agencies (through municipal separate storm sewer systems, MS4s) are subject to narrative standards, such as that the discharge not cause or contribute to water quality problems in waters of the United States. The second key difference is the use by EPA and the states of the general permit approach. NPDES permits for storm water discharges associated with industrial activities have been promulgated as permits encompassing rules for the entire regulated community, rather than written specifically for each discharger. That is, the permits identify by type the facilities, and the discharges, that are subject to the permit; specify a series of actions required of facilities subject to the permit; and require dischargers that fit that type to identify themselves and complete the specified requirements (7). Failure to comply carries the same penalties as other discharges without an NPDES permit, and EPA and the states are expected to implement compliance and enforcement actions to promote and require compliance; but facilities do not hold site-specific permits, and in almost no cases are they held to numeric effluent limitations. As a result of these major differences from the overall structure of numeric effluent limitations in other NPDES requirements, industrial discharge of storm water runoff is not treated within this section.

BIBLIOGRAPHY

1. Water Environment Federation. (1997). The Clean Water Act 25th Anniversary Edition. Water Environment Federation, Alexandria, VA.
2. Copeland, C. (1999).
3. Tietenberg, T. (2001). *Environmental Economics and Policy*. Addison Wesley Longman Inc., New York.
4. Copeland, C. (2003). *CRS Issues Brief for Congress: Water Quality: Implementing the Clean Water Act*. The Library of Congress: Congressional Research Service, Resources Science and Industry Division, Order Code IB89102, Washington, DC.
5. Copeland, C. (2004). *CRS Issues Brief for Congress: Clean Water Act Issues in the 108th Congress*. The Library of Congress: Congressional Research Service, Resources Science and Industry Division, Order Code IB10108, Washington, DC.
6. U.S. Environmental Protection Agency. (2002). Laws & Regulations: Clean Water Act History; A Brief Introduction to the Clean Water Act; and Finding Clean Water Act Updates. Available at http://www.epa.gov/region5/water/cwa.htm.
7. U.S. Environmental Protection Agency. (1995). Final National Pollutant Discharge Elimination System Storm Water Multi-Sector General Permit for Industrial Activities. *Federal Register* **60**(189): 50803–51319.

READING LIST

U.S. Government Printing Office National Archives and Records Administration. (2004). *Code of Federal Regulations*. Title 40, Protection of the Environment; Chapter I, Environmental Protection Agency; Parts 401–499, Clean Water Act.

U.S. Environmental Protection Agency. (2000). National Pollution Discharge Elimination System Permitting Program. Available at http://www.epa.gov/owm/npdes.htm.

U.S. Environmental Protection Agency. (1997). Office of Compliance, Office of Enforcement Compliance and Assurance. EPA Sector Notebook Project, 1995 through 2002. Available at www.epa.gov/compliance/resources/publications/assistance/sectors/.

ODOR ABATEMENT IN WASTEWATER TREATMENT PLANTS

JIM PHILP
Napier University
Edinburgh, Scotland, United Kingdom

CHRISTOPHER POLLEY
SUE MCLEOD
William Forrest and Sons,
Omoa Works
Newarthill, Motherwell, United Kingdom

INTRODUCTION

Odorous emissions from sewer systems and wastewater treatment plants are produced generally by hydrogen sulfide (H_2S) and volatile or semivolatile organic compounds at very low concentrations; they are present in the liquid phase and are emitted into ambient air at the liquid–gaseous interface (1). These odors cause serious complaints in the vicinity of wastewater treatment plants. As a result, odor control has become a key issue facing wastewater treatment plant management (2). Developments are under way to improve techniques for assessing odor and producing new technologies for destroying odor.

Volatile organic compounds (VOCs) are often amenable to biological treatment through a variety of reactor formats. However, several other technologies exist that could be used in treating them to abate odor. The eventual choice of a technology largely depends on the efficiency of treatment and capital and operating costs. The advent of the “triple bottom line” now considers sustainability of the technique also, although this concept is in its infancy. Biological treatment undoubtedly represents a sustainable green option, but it must also be proven economically viable and efficient.

HOW IS ODOR MEASURED?

Olfactometry is the most widely used method for measuring odor concentrations (3). Odor is “measured” by diluting the odor-bearing gas in a dynamic system where it is presented to groups of volunteers at various dilutions. The threshold odor number (TON) is the concentration at which half of the panel of volunteers can detect the smell (4) or fail to distinguish it from odor-free air. Although far from perfect at present (5), a quantitative approach to the development of standards in olfactometry is being developed successfully in Europe (6).

Electronic sensing systems—the so-called “electronic nose”—are being researched as a replacement for, or complement to, olfactory measurements. Arrays of nonspecific sensors can respond to many thousands of chemical species due to the broad selectivity of the sensor surfaces (7). This technique can be used to produce a unique odor profile or fingerprint, which can be further interrogated by pattern recognition techniques or neural networks. Although this is far from a complete technology, results have shown that a nonspecific sensor array can measure odor concentration from a specific wastewater treatment works (2).

WHAT CAUSES ODOR AT WASTEWATER TREATMENT PLANTS?

Odor-producing substances created during wastewater treatment are small, relatively volatile molecules whose molecular weight is about 30–150. The odor is composed of a large number of different chemicals. Many of these substances develop from the anaerobic decomposition of organic matter containing sulfur and nitrogen. Inorganic gases produced during wastewater treatment include hydrogen sulfide (H_2S), ammonia (NH_3), carbon dioxide (CO_2), and methane (CH_4), of which only hydrogen sulfide and ammonia create odor problems.

Many of the causal molecules contain sulfur and are produced in trace concentrations during anaerobic decomposition. For example, mercaptans (Table 1) are highly offensive at very low threshold concentrations and are common in wastewater treatment. They are reduced compounds, analogous to alcohols with a substitution of sulfur for oxygen in the OH radical (8). Hydrogen sulfide is the predominant odorant from sewage, and it results from bacterial sulfate reduction, a dissimilatory, anaerobic process using sulfate as a terminal electron acceptor.

Anaerobic decomposition of proteins, amino acids, and urine results in a variety of odorous, nitrogen-containing compounds (Table 1). The resulting amines

Table 1. Odorous Chemicals from Wastewater Treatment[a]

Compound	Formula	Odor Character
Hydrogen sulfide	H_2S	Rotten eggs
Diallyl sulfide	$(CH_2CHCH_2)_2S$	Garlic
Diethyl sulfide	$(C_2H_5)_2S$	Nauseating
Dimethyl disulfide	$(CH_3)_2S_2$	Putrification
Allyl mercaptan	CH_2CHCH_2SH	Garlic
Butyl mercaptan	C_4H_9SH	Unpleasant
Crotyl mercaptan	$CH_3CHCHCH_2SH$	Skunk, rancid
Methyl mercaptan	CH_3SH	Decayed cabbage, garlic
Thiocresol	$CH_3C_6H_4SH$	Skunk, rancid
Thiophenol	C_6H_5SH	Putrid, nauseating, decay
Ammonia	NH_3	Sharp, pungent
Methylamine	$CH_3\ NH_2$	Fishy
Trimethylamine	$(CH_3)_3N$	Fishy, ammoniacal
Cadaverine	$NH_2(CH_2)_5\ NH_2$	Decomposing meat
Indole	$C_8H_6\ NH$	Fecal, nauseating
Skatole	$C_9H_8\ NH$	Fecal, nauseating
Acetic acid	CH_3COOH	Vinegar
Butyric acid	C_3H_7COOH	Rancid, sweaty
Valeric acid	C_4H_9COOH	Sweaty
Formaldehyde	$HCHO$	Acrid, suffocating
Butyraldehyde	C_3H_7CHO	Rancid, sweaty
Acetone	CH_3COCH_3	Fruit, sweet
Butanone	$C_2H_5COCH_3$	Green apple

[a] (Adapted from References 8 and 9)

have an offensive, fishy smell, and indole and skatole smell strongly of feces. Anaerobic fermentation end products include a range of short-chain volatile fatty acids that have a rancid, disagreeable odor. Aldehydes, alcohols, and ketones are byproducts of carbohydrate fermentation. Although many of these compounds have an odor described as sweet or fruity, their contribution with the other components leads to odors characteristic of wastewater treatment plants. Thus, it is hardly surprising that over 40% of respondents in a study of 100 German wastewater treatment plants (10) identified sludge treatment processes as sources of odor.

VOLATILITY

VOCs can be defined as those compounds whose vapor pressure is 0.01 kPa or more. The volatility of a compound is a function of its vapor pressure. Vapor pressure may be thought of as the pressure exerted by a chemical on the atmosphere. Compounds that have high vapor pressure exert higher pressure on the atmosphere and therefore have an increasing driving force to volatilize. High water solubilities and low vapor pressures tend to decrease the potential for volatilization of dilute species from water. However, compounds that have a low vapor pressure may still have a high tendency to escape if their water solubility is low. For example, methanol as a pure compound is highly volatile but has a low tendency to evaporate from aqueous solution as it has infinite water solubility. So vapor pressure alone is not the best guide to volatility.

Henry's law constant is the best indicator of the tendency of a compound to volatilize from water. It can be expressed in two forms; one has units and one has a dimensionless form. Henry's law in the dimensionless form (H_D) can be expressed as

$$H_D = C_A/C_W \tag{1}$$

where C_A = the concentration in air (mol/m^3)
C_W = the concentration in water (mol/m^3).

Henry's law is the ratio of the partial pressure of a compound in the atmosphere to its concentration in water. Henry's law constant H_D has units of atm-m^3/mole and can be converted to its dimensionless form by

$$H = H_D(RT)^{-1}$$

where R = the universal gas constant [8.2×10^{-5} m^3-atm (mole-K)$^{-1}$]
T = the temperature on the Kelvin scale (K)

Thus, high values for Henry's constant indicate that the compound is volatile and low values indicate low volatility. Figure 1 is a working guide to the volatility of compounds based on Henry's law constants.

BIODEGRADABILITY

The ability of microorganisms to degrade environmental pollutants depends on the chemical functional groups present; some are more easily attacked than others. Although there is a well-accepted order of biodegradability of functional groups, the predictability of biodegradation is not yet an exact science:

$$\begin{aligned}&\text{Ester} \rightarrow \text{anhydride} \rightarrow \text{carbonyl} \rightarrow \text{sites of unsaturation}\\ &\rightarrow \text{methyl} \rightarrow \text{amide} \rightarrow \text{hydroxyl}\\ &\rightarrow \text{epoxide} \rightarrow \text{benzene} \rightarrow \text{methylene}\end{aligned}$$

$$\xrightarrow{\hspace{10cm}}$$

Decreasing biodegradability

As the knowledge of biodegradability is woefully incomplete (less than 0.01% of known organic compounds have been tested) at present, there has been an initiative started to gather the information available and organize it systematically (12).

AQUEOUS PHASE PARTITION

When dealing with VOCs, a rate-determining step is likely to be aqueous solubility. Convincing evidence now exists that microbial biodegradation happens only in the aqueous phase, and therefore the overall rate of biodegradation of hydrophobic pollutants may be mass transfer limited. Of relevance to VOCs is the concept of threshold (13). If a chemical is so volatile that only a very low concentration exists in aqueous solution, then this may limit microbial growth. At some lower value, all the energy available from the carbon entering the cell may be used to keep the cell alive, and none is used for growth. The substrate (the VOC in this instance) is being metabolized, but the population size and biomass are not increasing. This concentration represents the threshold.

Hydrophobicity is the final key determinant of bioavailability and thus biodegradability. As a general rule, polar,

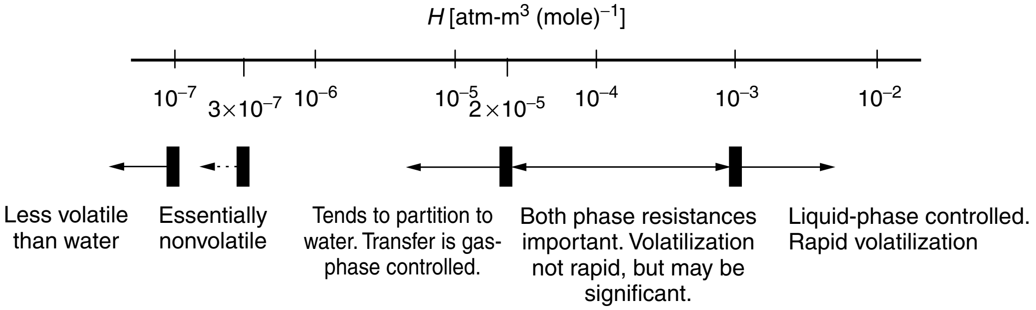

Figure 1. Ranges and relative values of Henry's law constants (adapted from Reference 11).

water-soluble, hydrophilic chemicals are more readily available to organisms than nonpolar, water-insoluble, hydrophobic ones. The polarity of a chemical has a strong negative correlation with its log octanol–water partition coefficient (log K_{ow}). K_{ow} is the ratio of a chemical's concentration in octanol to its concentration in water at equilibrium. K_{ow} was developed by the pharmaceutical industry as an index to predict the behavior of a drug in the body because partitioning between octanol and water roughly mimics partitioning between body fat and water. It is used as a convenient measure of hydrophobicity.

The above highlights the fact that the biodegradability of individual VOCs depends on the interplay of several factors. But it is also used to stress that VOCs have to be made available to an aqueous phase before biodegradation is feasible.

COMPETING TECHNOLOGIES

There are several noteworthy technologies for treating VOCs to abate odor.

Recovery/Removal

Recovery/removal treatments extract the VOC from the air stream for recycling. Whereas recovery has obvious benefits in industries using solvents, there is no benefit in the animal rendering business.

Adsorption. This is the physical adhesion of molecules to the surfaces of an adsorbent without chemical reaction. The adsorbent is usually present as a finely powdered material to maximize the surface area for contact. For activated carbon, the surface area available is very high, of the order of 1000 m^2/g. Where both the VOC concentration and the flow rate are low, activated carbon is the normal absorbent of choice. In such instances, the carbon is removed periodically and sent for regeneration or disposal. When the VOC concentrations or the flow rates are high, an *in situ* regeneration system is usually used.

Scrubbing. Water is the most common scrubbing liquid. Alkaline solutions are used for acidic components and acid solutions for alkaline compounds such as ammonia (14).

Condensation. At high VOC concentrations, an air stream may be cooled sufficiently to allow condensing the VOC as a liquid. Cryogenic condensation may be required to achieve sufficiently low concentrations. Such techniques are relatively energy-intensive, which contributes to high running costs (See Fig. 7).

Destruction

Destructive technologies aim to break down VOCs to carbon dioxide and water.

Oxidation. Oxidation can destroy most organic compounds. It is currently the most widely used technology for VOC treatment, but both thermal and catalytic oxidation systems have very high capital and running costs. Thermal oxidation requires raising the exhaust gas temperature to over 800 °C; heat recovery is employed to improve cost-effectiveness. Catalytic incineration operates at around 350 °C.

A summary of the available technologies for treating VOCs is shown in Fig. 2.

BIOLOGICAL TREATMENT TECHNOLOGIES

The biotreatment plants available for VOC treatment are adaptations of traditional, highly successful fixed film

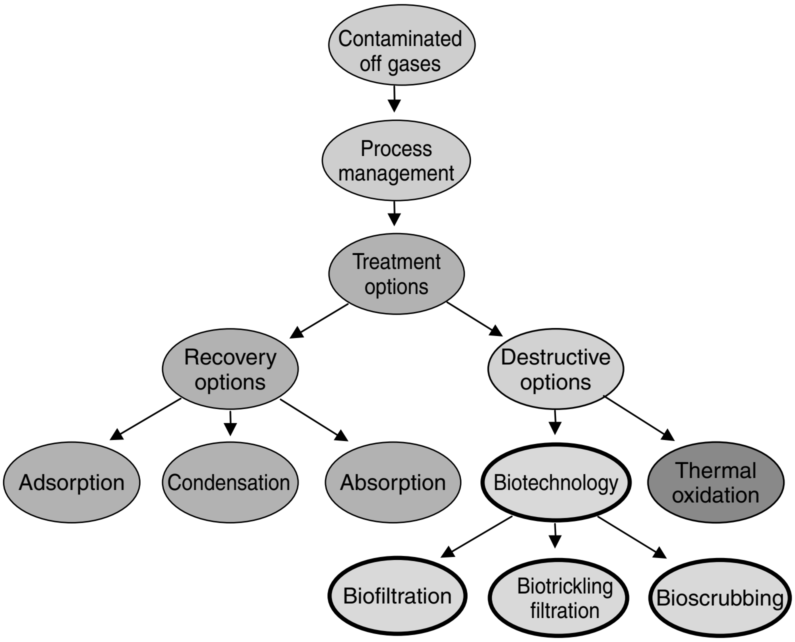

Figure 2. Available options for treating VOCs.

technologies used for wastewater treatment. The most crucial modification is the means of dissolving the VOC in the aqueous phase. They are simple mechanically, which means that maintenance and operation are simple and inexpensive. Expected life spans are of the order of 20 years. The most frequent maintenance required is periodic cleaning of the filter support material and sludge disposal.

Biofilters

VOC-containing stream(s) are humidified and discharged to the top of a reactor filled with an inert support medium to which the appropriate microorganisms attach and form a treating biofilm. Treatment is in the down-flow mode (Fig. 3); thus treatment is only effected by biofilm organisms and not by freely suspended biomass. Natural supports (e.g., bark, peat, compost, heather) improve the sustainability of the operation. Plastic supports can be used, but those available for wastewater trickling filters are probably too large; the void volume is so large that short-circuiting may be possible.

Biotrickling filter. The VOC-laden airstream enters at the bottom of the reactor and passes upward through the biofilm immobilized on a high surface area synthetic support (Fig. 4). Water downflow ensures the delivery of inorganic nutrients and keeps the biomass moist. The VOCs dissolved in the aqueous phase are mineralized in the process. Note that filter is a misnomer—filtration plays only a small part in the operation; the major contribution comes from biological oxidation.

Bioscrubber. VOCs are first absorbed in a liquid phase (usually water) in a tower packed with an inert biomass support (A, Fig. 5). The solution is then pumped to an aerated tank containing a suspended flocculating biomass in an activated sludge process (B) where biodegradation takes place. The treated water is settled to return biomass to the process, and the wastewater discharged.

CHOICE OF OPTIONS

A wide range of factors other than costs govern the technology choice. The principal factors are concentration and flow rate. Figure 6 gives a technology selection based on these two factors. However, bear in mind that plants are application-specific. For example, biofilters appear to be the best option (of the biological and other treatments) for dealing with animal carcass disposal VOCs (14).

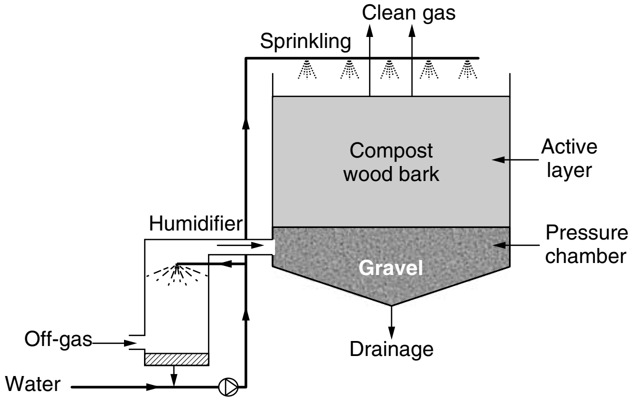

Figure 3. Biofilter characteristics.

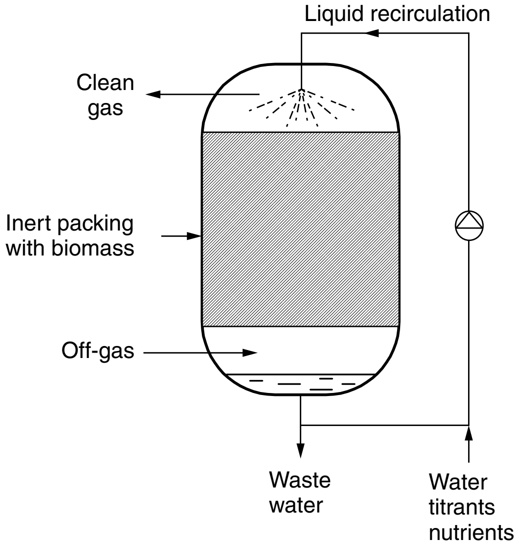

Figure 4. Biotrickling filter characteristics.

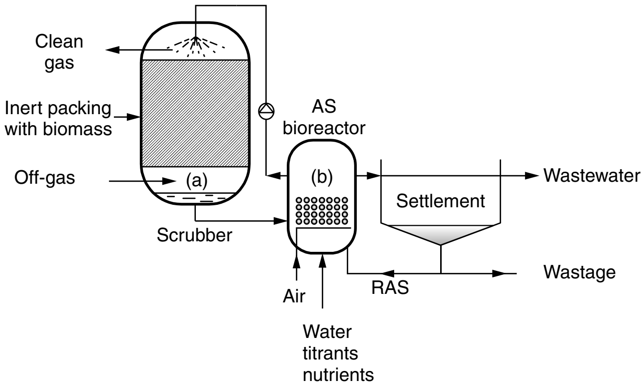

Figure 5. Bioscrubber characteristics.

THE BOTTOM LINE

Biological treatments for VOC degradation must compete with the other available technologies in treatment efficiency and also cost. The ambient conditions for treatment, low maintenance, and reasonable capital costs mean that biological methods compare very favorably with other competing technologies (Fig. 7). The size and the complexity of the plant are the determining factors that affect treatment cost. Plant size is a function of the volume of off-air to be treated daily and how well the VOCs dissolve in water.

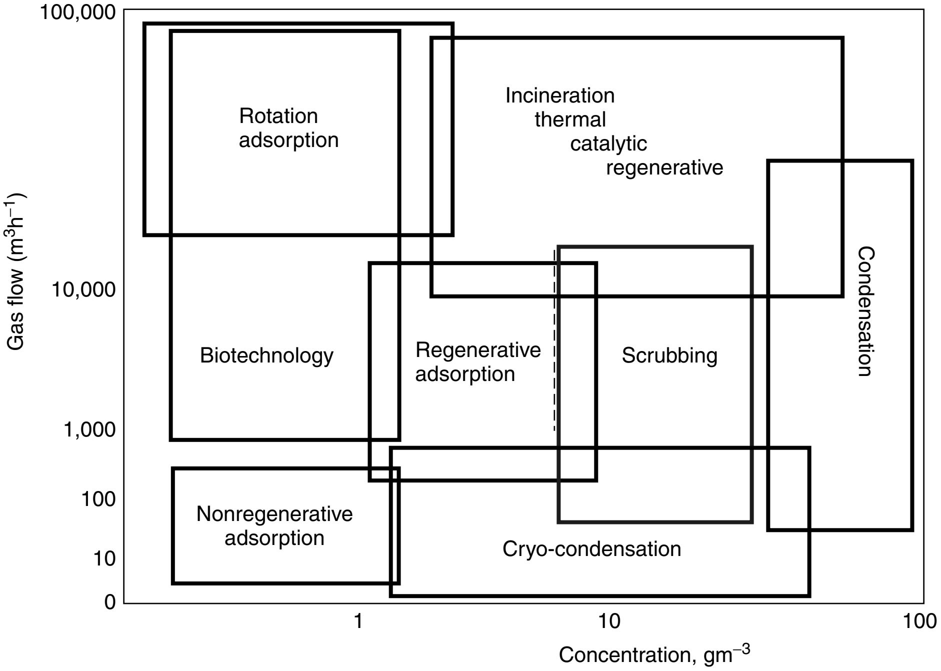

Figure 6. Effect of concentration and flow rate on choice of abatement option.

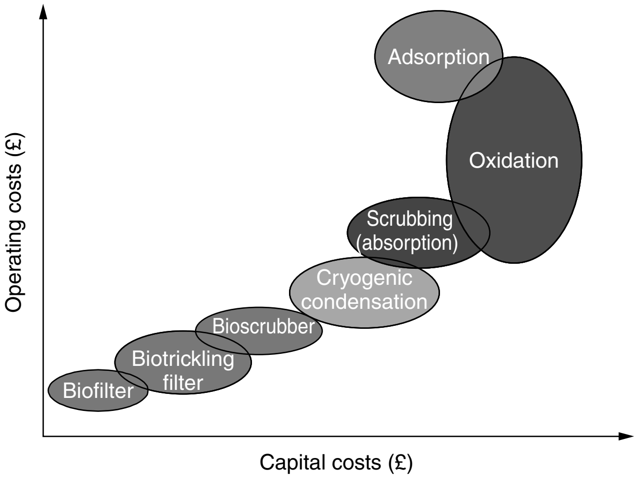

Figure 7. Relative costs of biological treatment versus other systems for VOCs.

BIBLIOGRAPHY

1. Frechen, F-B. and Köster, W. (1998). Odour emission capacity of wastewaters—standardization of measurement method and application. *Water Sci. Technol.* **38**: 61–69.
2. Stuetz, R.M., Fenner, R.A., and Engin, G. (1999). Assessment of odours from sewage treatment works by an electronic nose, H_2S analysis and olfactometry. *Water Res.* **33**: 453–461.
3. Defoer, N., De Bo, I., Van Langenhove, H., Dewulf, J., and Van Elst, T. (2002). Gas chromatography-mass spectrometry as a tool for estimating odour concentrations of biofilter effluents at aerobic composting and rendering plants. *J. Chromatogr. A* **970**: 259–273.
4. Porteous, A. (2000). *Dictionary of Environmental Science and Technology*, 3rd Edn. John Wiley, Chichester, UK.
5. Bliss, P.J., Schulz, T.J., Senger, T., and Kaye, R.B. (1996). Odour measurement—factors affecting olfactometry performance. *Water Sci. Technol.* **34**: 549–556.
6. Schulz, T.J. and van Harreveld, A.P. (1996). International moves towards standardisation of odour measurement using olfactometry. *Water Sci. Technol.* **34**: 541–547.
7. Persaud, K.C., Khaffaf, S.M., Hobbs, P.J., and Sneath, R.W. (1996). Assessment of conducting polymer odour sensors for agricultural malodours measurement. *Chem. Sci.* **21**: 495–505.
8. EPA/625/1-85/018. (1985). *Odor and Corrosion Control in Sanitary Sewerage Systems and Treatment Plants*. United States Environmental Protection Agency, Office of Research and Development, Cincinnati, OH.
9. Gostelow, P., Parsons, S.A., and Stuetz, R.M. (2001). Odour measurements for sewage treatment works. *Water Res.* **35**: 579–597.
10. Frechen, F-B. (1988). Odour emissions and odour control at wastewater treatment plants in West Germany. *Water Sci. Technol.* **20**: 261–266.
11. Watts, R.J. (1997). *Hazardous Wastes: Sources, Pathways, Receptors*. John Wiley & Sons, New York.
12. Wackett, L.P. and Ellis, L.B.M. (1999). Predicting biodegradation. *Environ. Microbiol.* **1**: 119–124.
13. Alexander, M. (1999). *Biodegradation and Bioremediation*. Academic Press, San Diego, CA.
14. Bio-Wise. (2000). *VOC and Odor Abatement. A Review of Biological Abatement Technology*. Lewis, Boca Raton, FL.

READING LIST

Gostelow, P., Longhurst, P.J., Parsons, S.A., and Stuetz, R.M. (2003). *Sampling for Measurement of Odours*. IWA, Colchester, UK.

Stuetz, R.M. and Frechen, F-B. (2001). *Odours in Wastewater Treatment*. IWA, Colchester, UK.

AQUEOUS REACTIONS OF SPECIFIC ORGANIC COMPOUNDS WITH OZONE

James B. Duncan
Kennewick, Washington

Ozone is an allotropic form of oxygen and exists as a pale-blue gas (O_3) with a pungent odor. It is very reactive chemically and decomposes without difficulty ($2O_3 \rightarrow 3O_2$). Within the arena of water treatment, it is used for disinfection and for taste and odor control. Ozone is 13 times more soluble in water than in oxygen, with a half-life of approximately 40 minutes at 14.6 °C at pH = 7.6. The density of ozone ranges from 1.09 kg/m at 0 °C to 0.14 kg/m^3 at 60 °C (1). From the table presented by Ullmann, the data fit a polynomial regression of Eq. 1.

$$y = -1.66\text{E-}06x^3 + 3.9\text{E-}04x^2 - 3.3\text{E-}02x + 1.086 \qquad R = 0.9998 \qquad (1)$$

where $y = O_3$, kg/m^3 and x = temperature, °C.

At higher temperatures and pH, the half-life decreases significantly. Although an excellent disinfectant, ozone oxidizes ammonia in a negligible amount in water treatment operations and ozone has little to no residue within a water distribution system. In addition to the above-mentioned aspects of ozone use, the other uses of ozone are:

- Control of excessive color
- Oxidation of iron and manganese
- Oxidation of organics
- Flocculation aid
- Does not form trihalomethanes (THMs); ozone has the ability to remove precursors to THM formation (2)

To analytically determine the amount of ozone residual in water, the 1995 edition of *Standard Methods, Method 4500-O_3 B.*, calls out the Indigo Colorimetric Method as the approved method (3).

Ozone-induced oxidations in aqueous environments can generally be described in the sequence of reactions as depicted in Fig. 1 (4,5). An amount of the ozone dissolved in an aqueous environment reacts directly with solutes (M). Usually these reactions are highly selective and demonstrate slow kinetics, on the order of minutes. Part of the ozone will decompose before reaction with solutes leading to free radicals. Among the free radicals, the hydroxyl radical is one of the most reactive oxidants. The hydroxyl radical can easily oxidize all types of organic contaminants along with several inorganic solutes. The hydroxyl radicals demonstrate rapid kinetics and are consumed in microseconds while demonstrating little substrate selectivity. In measured oxidations in model solutions, the indication is that up to 0.5 mole of hydroxyl radical ($OH^\bullet$) is formed per mole of ozone decomposed. As the pH increases, so does the decomposition of ozone, which is catalyzed by the hydroxyl ion (OH^-). Ozone decomposition may be accelerated by an autocatalyzed sequence of reactions in which radicals formed from decomposed ozone act as chain carriers. Some solutes will react with the hydroxyl radicals and from secondary radicals ($R^\bullet$), which may still act as chain carriers. Other species, such as bicarbonates, will inhibit chain reactions, which transform the primary radicals to terminal species (Φ). The rate of ozone decomposition is a function of the ozone concentration and pH as well as the solutes present (6,7).

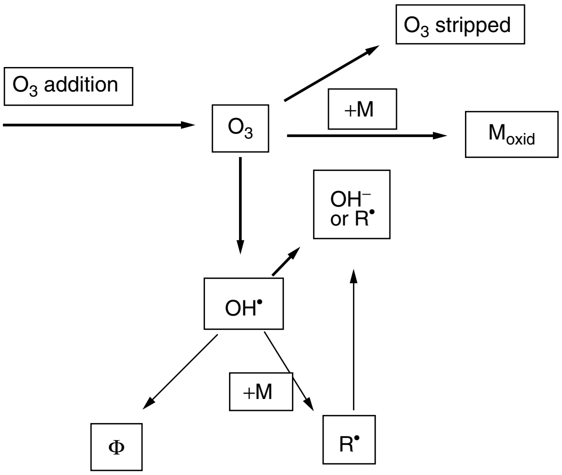

Figure 1. Ozone-induced oxidation.

When ozone attacks an organic carbon–carbon double bond (alkene), it first forms a molozonide, which quickly rearranges itself to an ozonide before the release of product (8) (see Fig. 2). Likewise, ozone will attack an alkyne bond in the following manner (8) (see Fig. 3). In an aqueous environment, ozone has different reactions because of its electronic configuration and may directly oxidize organic compounds. The reactivity of ozone is strongly influenced by electron density at sites of attack. Reactions of ozone with aromatics substituted with electron donor groups (OH, CH3, and NH_2) are exacerbated. However, aromatics substituted with electron withdrawing groups (COOH and NO_2) are retarded. The direct reactions can be divided into two categories: (1) the cyclo addition (Criegee mechanism) and an electrophilic reaction. The Criegee mechanism is selective to carbon–carbon double bonds, which then forms carbonyl compounds. An excellent review of the above-mentioned reactions may be found in Refs. 4 and 9. Rice reported on the following organic groups open to reactions with ozone (10) (see Fig. 4).

Discussions of ozone reactions with aliphatic compounds are found in Refs. 6 and 11–14. References 15–19

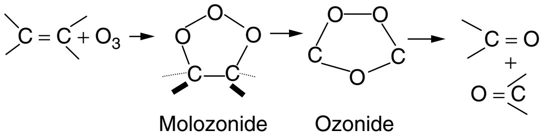

Figure 2. Molecular alternation.

Internal alkyne $R{-}C{\equiv}C{-}R' + O_3 \rightarrow RCOOH + R'COOH$

Terminal alkyne $R{-}C{\equiv}C{-}H + O_3 \rightarrow RCOOH + CO_2$

Figure 3. Ozone/alkyne relationship.

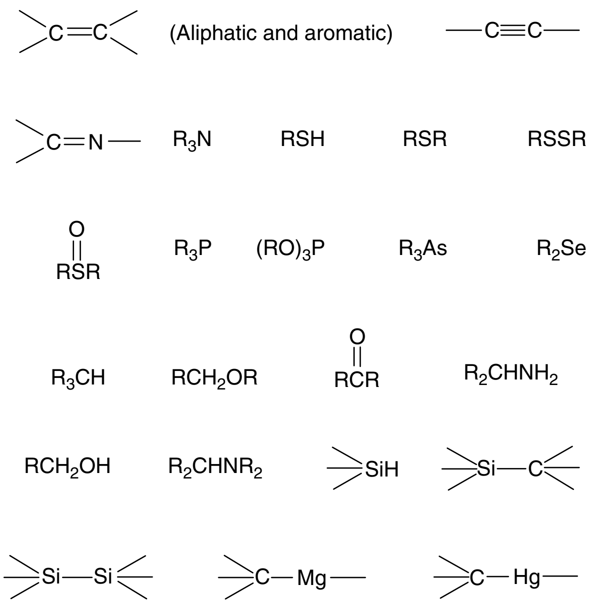

Figure 4. Organic groups reacting with ozone.

discuss reactions with aromatics with nonelectron withdrawing groups. For reactions with aromatics such as phenol, see Refs. 20–23.

BIBLIOGRAPHY

1. (1991). *Ullmann's Encyclopedia of Industrial Chemistry*, 5th Edn. VCH Verlagsgesellschaft, Germany, pp. 415–419.
2. De Zuane, J. (1997). *Handbook of Drinking Water Quality*, 2nd Edn. John Wiley & Sons, New York, pp. 138, 463.
3. APHA, AWWA, WEF. (1996). *Standard Methods, Method 4500-O_3 B., Indigo Colorimetric Method*, 19th Edn. American Public Health Association, Washington, DC, pp. 4–104–4-106.
4. Strumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry*, 3rd Edn. John Wiley & Sons, New York, p. 692.
5. Hoigne, J. and Bader, H. (1978). *Prog. Wat. Technol.* **10**: 657.
6. Langlais, B., Reckhow, D.A., and Brink, D.R. (1991). *Ozone in Water Treatment*. Applicatiion and Engineering. Lewis Publishers, Chelsea, MI.
7. Staehelin, J. and Hoigne, J. (1985). *Environ. Sci. Technol.* **19**(12): 1206–1213.
8. McMurray, J. (1992). *Organic Chemistry*, 3rd Edn. Brooks/Cole Publishing Company, Pacific Grove, CA, p. 239.
9. Beltran, F.J. (2004). *Ozone Reaction Kinetics for Water and Wastewater Systems*. Lewis Publishers, CRC Press Company, New York, pp. 7–13.
10. Rice, R.G., (1997). *Ozone Sci. & Eng.* **18**(6): 477–515.
11. Decoret, C., Royer, J., Legube, B. and Dore, M. (1984). *Environ. Technol. Lett.* **5**: 207–208.
12. Hoigne, J. and Bader, H. (1983) *Water Res.* **17**: 185–194.
13. Hoigne, J. and Bader, H. (1983). *Water Res.* **17**: 173–183.
14. Fronk, C.A. (1987). *Ozone-Science Engineering* **9**: 265–287.
15. Dore, M., Langlais, B., and Legube, B. (1980). *Ozone-Science Engineering* **2**: 39–54.
16. Legube, B., Langlais, B., Sohm, B., and Dore, M. (1981). *Ozone-Science Engineering* **3**: 33–48.
17. Legube, B., Langlais, B., and Dore, M. (1981). *Water Sci. Technol.* **13**: 553–570.
18. Legube, B., Guyon, S., Sugimitsu, H., and Dore, M. (1983). *Ozone-Science Engineering* **5**: 151–170.
19. Schalekamp, M., Dietlicher, K., Valenta, J., and Gould, J.P. (1982). *Water Sci. Technol.* **14**: 1042–1043.
20. Gurol, M.D. and Singer, P.C. (1983). *Water Res.* **17**: 1173–1181.
21. Gurol, M.D. and Vatistas, R. (1987). *Water Sci. Technol.* **19**: 1177–1181.
22. Singer, P.C. and Gurol, M.D. (1983). *Water Res.* **17**: 1163–1171.
23. Gould, J.P. (1987). *Ozone-Science Engineering* **9**: 207–216.

THE FATE OF PERSISTENT ORGANIC POLLUTANTS (POPs) IN THE WASTEWATER TREATMENT PROCESS

C. Samara
A. Katsoyiannis
Aristotle University of Thessaloniki
Thessaloniki, Greece

INTRODUCTION

Municipal wastewaters may contain a wide range of organic compounds deriving as components of domestic sewage, the urban runoff, including wet and dry deposition from the atmosphere, and industrial discharges (1,2). Among them are chlorinated biphenyls, dioxins and furans, polycyclic aromatic hydrocarbons, pesticides, pharmaceutical and personal care products, detergents, dyes, and solvents (3–7). Modern wastewater treatment plants (WWTPs) can effectively accomplish carbon and nitrogen removal, as well as microbial pollution control. However, conventional treatment technologies have not been specifically designed for the different organic contaminants of wastewater (natural or synthetic). The removal efficiencies of these compounds are influenced, apart from their physicochemical properties, by microbial activity and environmental conditions (2,8). Several studies have shown that the elimination of organic contaminants is often incomplete (2,7,9), rendering WWTPs important sources of toxic chemicals to the receiving environment (10–15).

Persistent organic pollutants (POPs) are a class of chemicals with pronounced persistence to biological/chemical degradation and, as such, are expected to be less efficiently eliminated in the conventional biological treatment of wastewater. Included in this class are the intended industrial products organochlorine pesticides (OPs), among them Dieldrin, Aldrin, Endrin, DDT, Toxaphene, Mirex, Chlordane, hexachlorobenzene, hexachlorocyclohexanes (HCHs), chlorinated biphenyls, and PCBs and the unintended polychlorinated dibenzo-*p*-dioxins and furans (PCDD/Fs). In addition to their

Table 1. Selected Physicochemical Properties[a] of POPs

	$\log K_{ow}$	H_c (atm·m^3· mol^{-1})	WS
PCDD/Fs	4.8–8.2 (Ref. 24)	$10^{-4.8}$–10^{-7} (Ref. 26)	10^{-4}–10^{-9} mol/L (Ref. 28)
PCBs	4.3–8.2	3–5 × 10^{-4} (Ref. 27)	0.1–160 ppb (Ref. 29)
OPs	3.6–6.2 (Ref. 25)	4.6 × 10^{-7}–5 × 10^{-4} Ref. 27	0.7–8000 ppb

[a] K_{ow}, *n*-octanol–water partition coefficient; H_c, Henry's law constant; WS, water solubility.

persistence, POPs are hydrophobic compounds with a strong tendency for bioaccumulation. They have toxic effects on animal reproduction, development, and immunological function. POPs span a large range of volatility and condensation temperatures, and some may present high mobility through the global distillation process (16). Several actions have been taken by environmental organizations and governments to restrict the inputs of POPs to the environment. The emissions of PCDD/Fs from their major sources—waste incinerators—have been strictly regulated. The production of PCBs and OPs has been banned in the United States and in Europe since the mid-1970s, although some of them were still being used until lately in developing countries, and in 2001 the Stockholm Convention for POPs banned the manufacture and use of these chemicals worldwide.

Several studies worldwide have indicated the presence of POPs in wastewaters treated in WWTPs (2,6,12,15, 17–23). Their influent concentration levels depend on the origin of the wastewater, and mainly on the contribution of industrial discharges. The fate of POPs in WWTPs is influenced by several factors: the physicochemical properties (Table 1) of the compounds, their biodegradability potency, and the composition of treated wastewater and the operational characteristics of the plant.

THE ACTIVATED SLUDGE TREATMENT PROCESS

Conventional WWTPs include two reactors—the primary reactor (primary sedimentation tank) and the biological reactor (aeration and secondary sedimentation tanks)—and, finally, a disinfection stage. The sludge that is produced during the primary and secondary sedimentation also undergoes treatment comprising digestion (usually anaerobic), thickening, homogenization, and dewatering (Fig. 1).

Behavior of POPs in the Primary Reactor

POPs entering the primary treatment stage of a WWTP are likely to be removed through sorption onto suspended solids and subsequent sedimentation, advection out of the tank to the biological reactor (either in the dissolved or the adsorbed phase), volatilization to the atmosphere through diffusional exchange at the air–water interface, and biotransformation (30). Given the hydrophobic character of most POPs and their resistance to biodegradation, the principal removal mechanisms from this stage are sorption, advection, and possibly volatilization, whereas biotransformation is considered to have only a minor influence. Sorption to the sludge has been shown to remove considerable fractions of POPs, in general between 18% and 90% for individual compounds (2,31–33). The physicochemical parameter that best describes the tendency of POPs to be sorbed onto suspended particles is the *n*-octanol–water partition coefficient (K_{ow}). The tendency of POPs to be sorbed could be generalized as follows (34):

$\log K_{ow} < 2.5$	Low sorption potential
$\log K_{ow} > 2.5$ and < 4.0	Medium sorption potential
$\log K_{ow} > 4.0$	High sorption potential

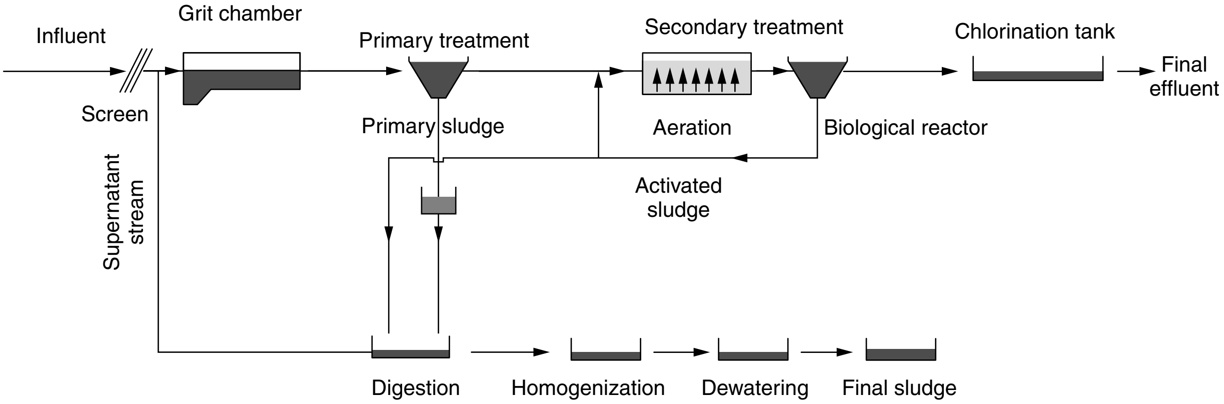

Figure 1. Flowchart of a conventional wastewater treatment plant.

The log K_{ow} values of individual POPs vary between 3.8 for HCHs and 8.2 for highly chlorinated dioxins and furans (Table 1), suggesting different sorption rates for different compounds within this class of chemicals.

Katsoyiannis and Samara (35), investigating the distribution of 20 POPs (13 OPs and 7 PCBs) between the sorbed and the aqueous phase of untreated and treated wastewater and sludge in a conventional WWTP, found that phase partitioning, expressed by the distribution coefficient K_d is influenced by two parameters: the concentration of solids, either suspended or settling (SS), and the concentration of dissolved organic carbon (DOC). For both parameters, an inverse relationship with log K_d was observed, suggesting that DOC and nonsettling microparticles can be an important carrier of POPs in the dissolved phase of treated effluents. Morris and Lester (32) also suggested that the existence of miscible organic solvents and nonsettling fine particles ($<100\ \mu$m) in the wastewater can increase the percentage of hydrophobic chemicals that remain in the effluent of the primary sedimentation tank.

The removal efficiency of POPs in the primary stage due to sorption can also be affected by the operational characteristics of the plant. According to Byrns (30), the removal of POPs due to sorption in settled sludge from the primary sedimentation stage might be influenced (decreased) following a reduction in sedimentation efficiency resulting, for instance, from an increased influent flow rate.

Volatilization and biodegradation/biotransformation are not expected to have significant contribution to the removal of POPs in the primary sedimentation tank. However, remarkable losses have been observed in the mass balance of DDT around the primary reactor, which were attributed to biodegradation under aerobic conditions (35,36).

Behavior of POPs in the Biological Reactor

In the activated sludge stage, a similar array of mechanisms controls the fate and intermedia transport of POPs. In this stage, air stripping due to the forced injection of air into the mixed liquor is considered as opposed to the diffusional exchange under the quiescent conditions characterizing the primary stage.

Sorption on activated sludge has been shown to remove up to 65% of Heptachlor, 18% of Lindane, and 60% of Aroclor 1254 entering the biological reactor (22,37). Byrns (30) suggested that almost 96% of DDT and 2,3,7,8-TCDD but only 34% of Dieldrin are likely to be adsorbed to surplus sludge. The secondary sorption rates are usually lower than the primary ones.

Data for the biodegradation/biotransformation of POPs in biological treatment plants are sparse and relatively few studies have been undertaken to determine rate coefficients (rate coefficients as low as 0.0001 have been reported in Ref. 38 for DDT, Dieldrin, and 2,3,7,8-TCDD). Biotransformation of a chemical in the dissolved phase is governed by the hydraulic retention time of the mixed liquor in the biological reactor, whereas for chemicals strongly sorbed to the biomass solids, the average reaction time for biodecay will be governed by the operating solids retention time as a result of the biosorption phenomenon (30). Biodegradation/biotransformation in the secondary treatment has been suggested for certain POPs, such as for Lindane (33,39), for Aldrin to Dieldrin, for Isodrin to Endrin, for Heptachlor to Heptachlorepoxide (37), and for DDT to DDE and DDD (21,35). Highly chlorinated compounds such as PCBs and PCDD/Fs are considered as resistant to degradation under aerobic conditions and would only degrade under anaerobic conditions (34).

Volatilization of POPs during the secondary treatment is facilitated by the forced injection of air in the aeration tank, which is important for some POPs. Volatilization occurs from the dissolved fraction since the fraction sorbed to the solid phase is not directly available, under equilibrium conditions, for mass transfer across the water–air interface. The significance of volatilization losses of specific organic compounds during sewage treatment can be estimated using the following empirically defined categories based on Henry's law constant (H_c, in atm·m^3·mol^{-1}) (34):

$H_c > 1 \times 10^{-4}$	High volatilization potential
$H_c < 1 \times 10^{-4}$	Low volatilization potential

The H_c values of POPs vary between 3×10^{-4} for PCBs to 5×10^{-7} for Endrin (Table 1), suggesting that removal due to volatilization is also a compound-specific mechanism. Bamford et al. (40), studying the air–water exchange of PCBs, suggested that approximately 10% of the dissolved PCB content in a water column exchanges with the atmosphere each day. In their study, Petrasek et al. (37) showed that Heptachlor and Aroclor 1254 might be removed up to 50% by volatilization, whereas Lindane would hardly be volatilized. Finally, Byrns (30), modeling the fate of organic xenobiotics in conventional WWTPs, predicted very low removal percentages (<0.1%) for DDT, Dieldrin, and 2,3,7,8-TCDDs due to volatilization in the activated sludge stage.

Behavior of POPs in the Disinfection Stage

Disinfection is the final treatment step in a WWTP before the discharge of the effluent wastewater. Disinfection is usually done by use of free chlorine, which is followed by formation of chlorinated disinfection by-products, especially when the wastewater contains high concentrations of natural organic matter. These chlorinated by-products are basically trihalogenated methanes (THMs), but it has been suggested that formation of some POPs might also occur (41). No formation of POPs has been reported when other disinfection reagents are used (e.g., ClO_2).

The Fate of POPs in the Sludge Treatment Stream

The amounts of POPs that are removed primarily and secondarily by sorption/sedimentation are transported to the sludge treatment stream. It has been estimated that compounds with log $K_{ow} > 3.5$ will concentrate in sewage sludge in concentrations more than 200 times higher than the ones found in the raw wastewater, while for compounds with log $K_{ow} > 5$, the concentration factor may be even 1000 (30).

The persistence of POPs during the digestion of sludge under anaerobic conditions has not been completely clarified and many controversial studies have been published on this issue. Hill and McCarty (42) investigated the effect of anaerobic conditions during sludge digestion on the degradation of DDT, Aldrin, Dieldrin, Endrin, Heptachlor, and γ-HCH. They noticed enhanced degradation in comparison to the aerobic conditions for all the compounds except Dieldrin. The resistance to degradation followed the order Dieldrin > Aldrin > DDD > DDT > Endrin > Heptachlor > γ-HCH. Zitomer and Speece (43) reported that reductive dehalogenation takes place under anaerobic conditions, reducing the level of chlorination of organochlorine pesticides, thus making them more amenable to further degradation and in general rendering residues less toxic. Dechlorination of OPs is facilitated by the presence of suitable microorganisms in sewage sludge; for instance, the anaerobic bacterium *Clostridium rectum* is capable of degrading γ-HCH (44). In the case of PCDD/Fs, the dechlorination pathway is from the higher chlorinated 2,3,7,8-subtituted PCDD/Fs to the lower ones, which, in antithesis to PCBs and pesticides, have a higher toxic equivalence factor (TEF) rating, something that increases the overall toxicity of sewage sludge (45,46). Fu et al. (47) reported that abiotic dechlorination of PCDD/Fs in sludge leads to DiCDD/Fs as end products. This suggestion was also prompted by Stevens et al. (48), who found that DiCDFs were the predominant congeners in U.K. sewage sludges. However, this suggestion was not confirmed by Klimm et al. (49) and by Disse et al. (50), who under strictly anaerobic conditions did not observe any formation or destruction of PCDD/Fs. On the contrary, Klimm et al. (49) observed formation of hepta- and octa-CDDs under semianaerobic conditions, something that led to a twofold increase in the concentration of these congeners. Under the same conditions, there was no formation of other PCDD/Fs or PCBs congeners.

From the above-mentioned, it is difficult to decide if the POPs degrade or do not degrade during the anaerobic digestion of sewage sludge. The fact that POPs can be detected in almost all sewage sludges after treatment (46) suggests that they are rather resistant to this process.

The concentration of POPs, especially of PCBs and of PCDD/Fs, in sewage sludges is of great importance, since sewage sludge is often used in agriculture for soil amendment. The European Union (51), in order to improve the long-term protection of soils, is working on a new directive, which includes maximum permissible concentrations for use of sludge in agriculture. According to this upcoming directive (51), the sum concentration of 7 PCB congeners (IUPAC −28, −52, −101, −118, −138, −153, −180) should not exceed 800 μg/kg (dry matter, dm), the concentration of PCDD/Fs should not exceed 100 ng TEQ/kg (dm), and the sum of organochlorine compounds should not exceed 500 mg/kg (dm). In studies dealing with the occurrence of POPs in sewage sludges, ΣPCB levels have been found to vary dramatically from 22.7 to 8000 μg/kg (dm) (1,2,52–62). The reported PCDD/Fs concentrations are also greatly variable ranging between 0.7 and 4100 ng TEQ/kg (dm) (52,63–67).

The Fate of POPs Throughout the Overall Treatment Process

Studies have shown that POPs are recalcitrant to the conditions prevailing in typical biological WWTPs and persist in one or more phases within the treatment plant. Reported removal efficiencies throughout the overall treatment process range within 18–100% for PCBs and 75–90% for several OPs (2,6,31). There are no available data concerning the removal of PCDD/Fs in WWTPs. The distribution of POPs within the WWTP is dependent on the physicochemical properties of the chemicals and the operating conditions within the plant. The principal removal mechanism for the most hydrophobic POPs is through sorption to sludge particles and transfer to the sludge processing system. Advective transport into the final effluent, in association with suspended solids or in the dissolved phase, is also important for less hydrophobic compounds. It has been reported that from the 50 kg of all PCBs annually entering a WWTP, 60% is removed through the wasted sludge, 26% is discharged into the recipient with the final effluent, while 14% is lost due to volatilization or biotransformation (35). Therefore, the long-term ecotoxicological effects on both terrestrial and aquatic organisms need to be assessed for safe disposal of the products of the wastewater treatment process.

BIBLIOGRAPHY

1. Blanchard, M. et al. (2001). Origin and distribution of polyaromatic hydrocarbons and polychlorobiphenyls in the urban effluents to wastewater treatment plants of the Paris area (France). *Water Res.* **35**(15): 3679–3687.
2. Katsoyiannis, A. and Samara, C. (2004). Persistent organic pollutants (POPs) in the conventional activated sludge treatment process: occurrence and removal. *Water Res.* **38**(11): 2685–2698.
3. Bedding, N.D., McIntyre, A.E., Perry, R., and Lester, J.N. (1982). Organic contaminants in the aquatic environment. I sources and occurrence. *Sci. Total Environ.* **25**: 143–167.
4. Valls, M., Fernandez, P., and Albaiges, J. (1989). Broad spectrum analysis of organic contaminants in urban wastewaters and coastal receiving systems. In: *Organic Contaminants in Wastewater, Sludge and Sediment: Occurrence, Fate and Disposal*. D. Quagherbeur, I. Temmerman, and G. Angeletti (Eds.). Elsevier Applied Science, New York, pp. 19–34.
5. Paxeus, N., Robinson, P., and Balmer, P. (1992). Study of organic pollutants in municipal wastewater in Goeterborg, Sweden. *Water. Sci. Technol.* **25**: 249–256.
6. Pham, T.T. and Proulx, S. (1997). PCBs and PAHs in the Montreal urban community (Quebec, Canada) wastewater treatment plant and in the effluent plume in the St. Lawrence River. *Water Res.* **31**(8): 1887–1896.
7. Carballa, M. et al. (2004). Behavior of pharmaceuticals, cosmetics and hormones in a sewage treatment plant. *Water Res.* **38**: 2918–2926.
8. Johnson, A.C. and Sumpter, J.P. (2002). Removal of endocrine disrupting chemicals in activated sludge treatment works. *Environ. Sci. Technol.* **256**: 163–173.
9. Manoli, E. and Samara, C. (1999). Occurrence and mass balance of polycyclic aromatic hydrocarbons in the Thessaloniki sewage treatment plant. *J. Environ. Qual.* **28**: 176–187.
10. Bergh, A. and Peoples, R. (1977). Distribution of polychlorinated biphenyls in a municipal wastewater treatment plant and environs. *Sci. Total Environ.* **8**(3): 197–204.

11. Shannon, E., Ludwig, F., and Waldmanis, I. (1977). Polychlorinated biphenyls (PCBs) in municipal waste waters. An assessment of the problem in the Canadian lower Great Lakes. *Environmental Canada Research Report*, No 49.
12. McIntyre, A., Perry, R., and Lester, J. (1981). The behaviour of polychlorinated biphenyls and organochlorine insecticides in primary mechanical waste water treatment. *Environ. Poll. Ser. B* **2**: 223–233.
13. Barrick, R. (1982). Flux of aliphatic and polycyclic aromatic hydrocarbons to central Puget Sound from Seattle (Westpoint) primary sewage treatment effluent. *Environ. Sci. Technol.* **16**: 682–692.
14. Eganhouse, R. and Kaplan, I. (1982). Extractable organic matter in municipal wastewater. 1. Petroleum hydrocarbons: temporal variations and mass emission rates to ocean. *Environ. Sci. Technol.* **16**: 180–186.
15. Granier, L., Chesterikoff, A., Chevreuil, M., and Letolle, R. (1988). Essai de bilan des polychlorobiphenyles (PCB) a la station d'epuration de l'agglomeration parisienne (Acheres). *C.R. Acad. Sci. Paris* **306**(Ser. II): 1175–1178.
16. Wania, F. and Mackay, D. (1996). Tracking the distribution of persistent organic pollutants. *Environ. Sci. Technol. News* **30**(9): 390–396.
17. Pujadas, E. et al. (2001). Application of the new C18 speedisks™ to the analysis of polychlorinated dibenzo-*p*-dioxins and dibenzofurans in water and effluent samples. *Chemosphere* **43**: 449–454.
18. Peters, T., Nestrick, T., and Lamparski, L. (1984). The determination of 2,3,7,8-tetrachlorodibenzo-*para*-dioxin in treated wastewater. *Water Res.* **18**(8): 1021–1024.
19. De Alencastro, L.F. and Taradellas, J. (1983). Study of the concentration of PCBs in used water from a wastewater treatment plant. *Gaz Eaux Eaux Usees* **3**: 113–122.
20. Van Luin, A.B. and Van Starkenburg, W. (1985). Hazardous substances in wastewater. *Water Sci. Technol.* **17**: 843–853.
21. Saleh, F.Y., Lee, G.F., and Wolf, H.W. (1980). Selected organic pesticides, occurrence, transformation, and removal from domestic wastewater. *JWPCF* **52**: 19–28.
22. Hannah, S.A., Austern, B.M., Eralp, A.E., and Wise, R.H. (1986). Comparative removal of toxic pollutants by six wastewater treatment processes. *JWPCF* **58**: 27–34.
23. Nicoud, S., Humbert, B., De Alencastro, L.F., and Taradellas, J. (1988). Organic micropollutants at the effluents of the wastewater treatment plant, at the water of Rhone and the water of Leman-Campagne 1987. *Rapport sur les études et recherches entreprises dans le basin Lemanique*, pp. 225–234.
24. Lorber, M., Barton, R., Winters, D., Bauer, K., Davis, M., and Palausky, J. (2002). Investigation of the potential release of polychlorinated dioxins and furans from PCP-treated utility poles. *Sci. Total Environ.* **290**(1–3): 15–39.
25. Huckins, J.N., Petty, J.D., Prest, H.F., Clark, R.C., Alvarez, D.A., Orazio, C.E., Lebo, J.A., Cranor, W.L., and Johnson, B.T. (2000). *A Guide for the Use of Semipermeable Membrane Devices (SPMDs) as Samples of Waterborne Hydrophobic Contaminants*. Columbia Environmental Research Center, USGS, USDI and California Analytical Division, Agilent Technologies, Inc., Palo Alto, CA.
26. Eitzer, B. and Hitte, R. (1989). Atmospheric transport and deposition of polychlorinated dibenzo-*p*-dioxins and dibenzofurans. *Environ. Sci. Technol.* **23**: 1396–1401.
27. U.S. EPA. (1982). *Fate of Priority Pollutants in Publicly Owned Treatment Works*. Final Report, Vol. I. EPA 440/1-82/303. Effluent guidelines division WH-552. U.S. EPA, Washington, DC.
28. Wang, Y. and Wong, P. (2002). Mathematical relationships between vapor pressure, water solubility, Henry's law constant, *n*-octanol–water partition coefficient and gas chromatographic retention index of polychlorinated-dibenzo-dioxins. *Water Res.* **36**: 350–355.
29. Anonymous (www.es.lancs.ac.uk/ecerg/kcjgroup/11.pdf).
30. Byrns, G. (2001). The fate of xenobiotic organic compounds in wastewater treatment plants. *Water Res.* **35**(10): 2523–2533.
31. Chevreuil, M., Granier, L., Chesterikoff, A., and Letolle, R. (1990). Polychlorinated biphenyls partitioning in waters from river, filtration and wastewater plant: the case for Paris (France). *Water Res.* **24**(11): 1325–1333.
32. Morris, S. and Lester, N. (1994). Behaviour and fate of polychlorinated biphenyls in a pilot wastewater treatment plant. *Water Res.* **28**(7): 1553–1561.
33. Kipopoulou, A.M., Zouboulis, A., Samara, C., and Kouimtzis, Th. (2004). The fate of Lindane in the conventional activated sludge treatment process. *Chemosphere* **55**: 81–91.
34. Rogers, H.R. (1996). Sources, behaviour and fate of organic contaminants during sewage treatment and in sewage sludges. *Sci. Total Environ.* **185**: 3–26.
35. Katsoyiannis, A. and Samara, C. (2004). Persistent organic pollutants (POPs) in the conventional activated sludge treatment process: fate and mass balance. *Environ. Res.* in press.
36. Pereira, W.E., Hostettler, F.D., and Rapp, J.B. (1996). Distributions and fate of chlorinated pesticides, biomarkers and polycyclic aromatic hydrocarbons in sediments along a contamination gradient from a point-source in San Francisco Bay, California. *Mar. Pollut. Bull.* **41**(3): 299–314.
37. Petrasek, A.C. et al. (1983). Fate of toxic organic compounds in wastewater treatment plants. *JWPCF* **55**(10): 1286–1296.
38. Howard, P., Boethling, R., Jarvis, W., Meylan, W., and Michalenko, E. (1991). *Handbook of Environmental Degradation Rates*. Lewis Publishers, Chelsea, MI.
39. Dorussen, H.L. and Wassenberg, W.B.A. (1997). Feasibility of treatment of low polluted waste water in municipal waste water treatment plants. *Water Sci. Technol.* **35**(10): 73–78.
40. Bamford, H.A., Ko, F.C., and Baker, J.E. (2002). Seasonal and annual air–water exchange of polychlorinated biphenyls across Baltimore harbor and the northern Chesapeake Bay. *Environ. Sci. Technol.* **36**: 4245–4252.
41. Ahmed, F.E. (2003). Analysis of polychlorinated biphenyls in food products. *TRAC–Trends Anal Chem.* **22**: 170–185.
42. Hill, D. and McCarty, P. (1967). Anaerobic degradation of selected chlorinated hydrocarbon pesticides. *J. Water Pollut. Cont. Fed.* **39**(8): 1259–1277.
43. Zitomer, D. and Speece, R. (1993). Sequential environments for enhanced biotransformation of aqueous contaminants. *Environ. Sci. Technol.* **27**(2): 227–244.
44. Mogilevich, N. (1982). Microbial destruction of organohalogen compounds. *Sov. J. Water Chem. Technol.* **43**(3): 98–109.
45. Barkowskii, A.L. and Adriaens, P. (1996). Microbial dechlorination of historically present and freshly spiked chlorinated dioxins and diversity of dioxin-dechlorinating populations. *Appl. Environ. Microbiol.* **62**: 675–681.
46. Stevens, J., Green, N., and Jones, K.C. (2003). Fate of 1,2,3,4,6,7,8-heptachlorodibenzofuran and pentachlorophenol during laboratory-scale anaerobic mesophilic sewage sludge digestion. *Chemosphere* **50**: 1227–1233.

47. Fu, Q.S., Barkowskii, A.L., and Adriaens, P. (1999). Reductive transformation of dioxins: an assessment of the contribution of dissolved organic matter to dechlorination reactions. *Environ. Sci. Technol.* **33**: 3837–3842.
48. Stevens, J., Green, N.J.L., and Jones, K.C. (2001). Survey of PCDD/Fs and non-ortho PCBs in UK sewage sludges. *Chemosphere* **44**: 1455–1462.
49. Klimm, C., Schramm, K.W., Henkelmann, B., Martens, D., and Kettrup, A. (1998). Formation of octa- and heptachlorodibenzo-*p*-dioxins during semi anaerobic digestion of sewage sludge. *Chemosphere* **37**: 2003–2011.
50. Disse, G., Weber, H., Hamann, R., and Haupt, H.J. (1995). Comparison of PCDD and PCDF concentrations after aerobic and anaerobic digestion of sewage sludge. *Chemosphere* **31**: 3617–3625.
51. EU. (2000). *Working Document on Sludge*. 3rd draft—unpublished.
52. Eljarrat, E., Caixach, J., and Rivera, J. (2003). A comparison of TEQ contributions from PCDDs, PCDFs and dioxin-like PCBs in sewage sludges from Catalonia, Spain. *Chemosphere* **51**: 595–601.
53. Stevens, J., Northcott, G., Stern, G., Tomy, G., and Jones, K.C. (2003). PAHs, PCBs, PCNs, organochlorine pesticides, synthetic muscs, and polychlorinated *n*-alkanes in UK sewage sludge: survey results and implications. *Environ. Sci. Technol.* **37**: 462–467.
54. Ottaviani, M., Crebelli, R., Fusselli, S., La Rocca, C., and Baldassari, L.T. (1993). Chemical and mutagenic evaluation of sludge from a large wastewater treatment plant. *Ecotox. Environ. Safety* **26**: 18–32.
55. Sulkowski, W. and Rosinka, A. (1999). Comparison of the efficiency of extraction methods for polychlorinated biphenyls from environmental wastes. *J. Chromatogr. A* **845**: 349–355.
56. Paulsrud, B., Wien, A., and Nedland, K.T. (2000). *A Survey of Toxic Organics in Norwegian sewage Sludge, Compost and Manure*. Aquateam, Norwegian Water Technology Centre, Oslo, Norway.
57. Alcock, R. and Jones, K.C. (1993). Polychlorinated biphenyls in digested UK sewage sludge. *Chemosphere* **25**(12): 2199–2207.
58. Kedding, M., Langenohl, T., and Witte, H. (1989). PCB, Dioxine und Furane im Klarschlamm und deren Auswirkungen bei der landwirtschaftlichen Klarschlammverwertung. *Korrespondenz Abwasser.* **35**: 19–26.
59. Lazzari, L., Sperni, L., Bertin, P., and Pavoni, B. (2000). Correlation between inorganic (heavy metals) and organic (PCBs and PAHs) micropollutant concentrations during sewage sludge composting processes. *Chemosphere* **41**: 427–435.
60. Taradellas, J., Muntau, H., and Beck, H. (1985). Abundance and analysis of PCBs in sewage sludge. In: *COST 681 Symposium on Processing and Use of Organic Sludge and Liquid Agricultural Wastes*. October 8–11, Rome, Italy.
61. Blanchard, M., Teil, M.J., Ollivon, D., Legenti, L., and Chevreuil, M. (2003). Polycyclic aromatic hydrocarbons and polychlorobiphenyls in wastewaters and sewage sludges from the Paris area (France). *Environ. Res.* **95**(2): 184–197.
62. McGrath, D., Postma, L., McCormack, R.J., and Dowdall, C. (2000). Analysis of Irish sewage sludges: suitability of sludge for use in agriculture. *Irish J. Agric. Food Res.* **39**(1): 73–78.
63. Fabrellas, B., Sanz, P., Abad, E., Rivera, J., and Larrazabal, D. (2004). Analysis of dioxins and furans in environmental samples by GC-ion-trap MS/MS. *Chemosphere* **55**(11): 1469–1475.
64. Eljarrat, E., Caixach, J., and Rivera, J. (1999). Decline in PCDD and PCDF levels in sewage sludges from Catalonia (Spain). *Environ. Sci. Technol.* **33**(15): 2493–2498.
65. Langenkamp, H., Part, P., Erhardt, W., and Pruess, A. (2001). *Organic Contaminants in Sewage Sludge for Agricultural Use*. A publication of the European Union. Available at http://europa.eu.int/comm/environment/waste/sludge/organics_in_sludge.pdf.
66. Rappe, C., Andersson, R., Studer, C., and Karlaganis, G. (1997). Decrease in the concentrations of PCDDs and PCDFs in sewage sludge from Switzerland. *Organohal. Comp.* **33**.
67. Rappe, C., Andersson, R., Karlaganis, G., and Bonjour, R. (1994). PCDDs and PCDFs in samples from various areas in Switzerland. *Organohal. Comp.* **20**: 79–84.

THE ROLE OF ORGANOCLAY IN WATER CLEANUP

George R. Alther
Biomin, Inc.
Ferndale, Michigan

Organically modified clays, also called organoclays, have been used to clean up water since 1985. Their prime function is as a prepolisher for activated carbon, ion exchange resins, and membranes. They are also used in a stand-alone mode after dissolved air flotation (DAF) and oil/water separation units. Their main use is to remove oils, greases, and other large hydrocarbons of low solubility from water. They are very adept at removing chlorinated hydrocarbons. In this application, organoclays are usually blended with anthracite in a ratio of 30% organoclay and 70% anthracite. The reason is that the organoclay, in its pure form, would collect so much oil in its interstitial pore spaces, and due to swelling of the clay, that it would last no longer than activated carbon, which removes 8–10% of oil based on its weight, before its pores are blinded. The organoclay blend, on the other hand, removes 50–70% of oil based on its weight, some seven times as much as activated carbon. The economic benefit for the end user, the one who pays for the cleanup, is a savings of 50% or more of operating costs.

Organoclays can be called prepolishers to carbon, but it can also be said that carbon is a postpolisher to organoclay. The reason is that the organoclay also removes other compounds, such as PNAHs, BTEX, PCBs, and other hydrocarbons of low solubility, with extreme efficiency. This has been shown in many publications through the last 15 years (1–5).

Organoclays are bentonites modified with quaternary amines. Bentonite is a volcanic, chemically altered rock that consists primarily of the clay mineral montmorillonite. Montmorillonite contains inorganic exchange ions, particularly sodium, calcium, and magnesium that hydrate in the presence of water and produce a hydrophilic environment on the surface of the clay. Mixing a cationic quaternary alkyl ammonium chloride or bromide compound with the bentonite makes the clay organophilic. Bentonite, which is hydrophilic in its natural state, becomes hydrophobic and organophilic when modified with

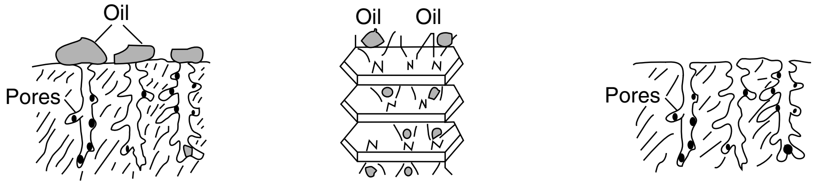

Activated carbon granule pore spaces of activated carbon, blinded by emulsified oil.

Clay platelets, modified with quaternary amine, removal of emulsified oil on the clay surface.

Activated carbon downstream of organoclay, ready to remove more soluble compounds.

Figure 1. How activated carbon and organoclay remove oil from water.

quaternary amines (6,7). The positively charged end of the amine chain, which consists of a carboxylic head that includes a nitrogen ion, exchanges for a sodium or calcium ion on the surface of the bentonite. The cationic amine now becomes neutral, and the thus formed organoclay has turned into a nonionic surfactant that has a solid phase. The swelling capacity of sodium bentonite in water is up to 15 times its volume, but nearly zero in hydrocarbon fuels and solutions (8). After the organic modification, the organoclay swells some 15 times in a fuel such as gasoline and by some 20% when placed into a column and exposed to fuels such as gasoline (9–11). Its swelling capacity in water is very low. If alcohol is added to a fuel, the organoclay will swell even more and turn the system into a gel, or even grease. For this reason, organoclays have been used as thickeners for paints, drilling muds, greases, printing inks, and many other systems since the 1950s.

Figure 1 shows how the two media differ in removing oil from water. The organoclay removes it by partition, whereas the activated carbon removes oil and other hydrocarbons by adsorption. Partition takes place outside the clay particle; adsorption takes place inside the pores, which is the reason that they become blinded by organic compounds equal to or exceeding the diameter of the pores.

During the late 1950s, it was discovered that organoclays remove organic compounds from water, including benzene (12). It was found that organoclays can remove chlorinated organic hydrocarbons of low solubility efficiently (13,14). Excellent descriptions of the mechanisms of organoclay interactions with organic compounds are also given by Lagaly (15,16).

The removal of these compounds from water by organoclays is the result of a partition mechanism similar to the process when immiscible organic compounds such as octanol are added to water contaminated with organic compounds (16,17–19). The interlayer phase of the organoclay acts as a partition medium for oils, greases, and other hydrocarbons (1,20–24). A portion of these organic compounds moves out of the water and into the organic compound where it is more soluble; like dissolves into like. The relative solubility of the contaminant determines the amount retained in each phase. The terminology can be extended to contaminants partitioning from the water phase into a solid phase such as organic cations sorbed to a clay surface. The higher the solution concentration of a compound and the lower its solubility, the larger the quantity removed by the organoclay by partitioning. The organic compounds are held closely by the quaternary amine by coulombic forces (25), and the contaminant is not easily leached off. If an amount of amine is exchanged into the bentonite that exceeds its stoichiometric capacity, the amine chains will still attach to the organoclay in a tail-to-tail interaction. The result is that this organoclay now has a positive charge and will remove organic and inorganic anions from water, such as humic acids and hexavalent chromium (Fig. 2) (26).

Only part of the clay surfaces is covered with the quaternary amines, so a portion remains free, available for cation exchange with heavy metals such as lead, zinc, nickel, cadmium, and iron. The cation exchange capacity of bentonite ranges from 70–95 meq/100 gram. Column studies were conducted with pure organoclay and an organoclay/anthracite blend, testing for a number of metals simultaneously, and then testing for the removal capacity for single metals (27).

The main application of organoclays has been in groundwater remediation. Removal of oil from water is the organoclay's main function. Figure 3 shows a column test with organoclay/anthracite to determine the sorptive capacity of this medium for a vegetable oil. A 30-inch long (76.2 cm) by 3-inch diameter (7.62 cm) column was constructed from polyvinyl chloride (PVC) and filled with about 6 pounds of sorbent material to be studied. A peristaltic pump forced an aqueous solution containing 680 mg/L of vegetable oil through the column, after the column was backwashed with water to displace any air pockets. Samples were collected periodically at the outflow of the column and analyzed for their organic content using chemical oxygen demand (COD) analysis. The results are shown in Table 1 and Fig. 3.

The percent removal capacity for a mineral oil such as Bunker C would be much higher because it is less soluble than vegetable oil. Figure 3 is a graphic description of these data.

Figure 4 shows the results of a column study comparing the removal capacity for oil between organoclay/anthracite and bituminous activated carbon. To gain some background data, the ability of powdered, nonionic organoclay to remove a variety of oils from water was tested in jar tests (26,28). The results displayed in Figs. 5–8 show that organoclay removes all oils from water exceptionally well, as long as they are refined. It is far superior to activated carbon, which is why it is used as a prepolisher. If oil is not refined, which means its composition includes polar compounds, either a cationic organoclay has to be used, or the oil, which is now partially chemically emulsified, must be deemulsified (1,21,23,28–30). Figure 9 shows

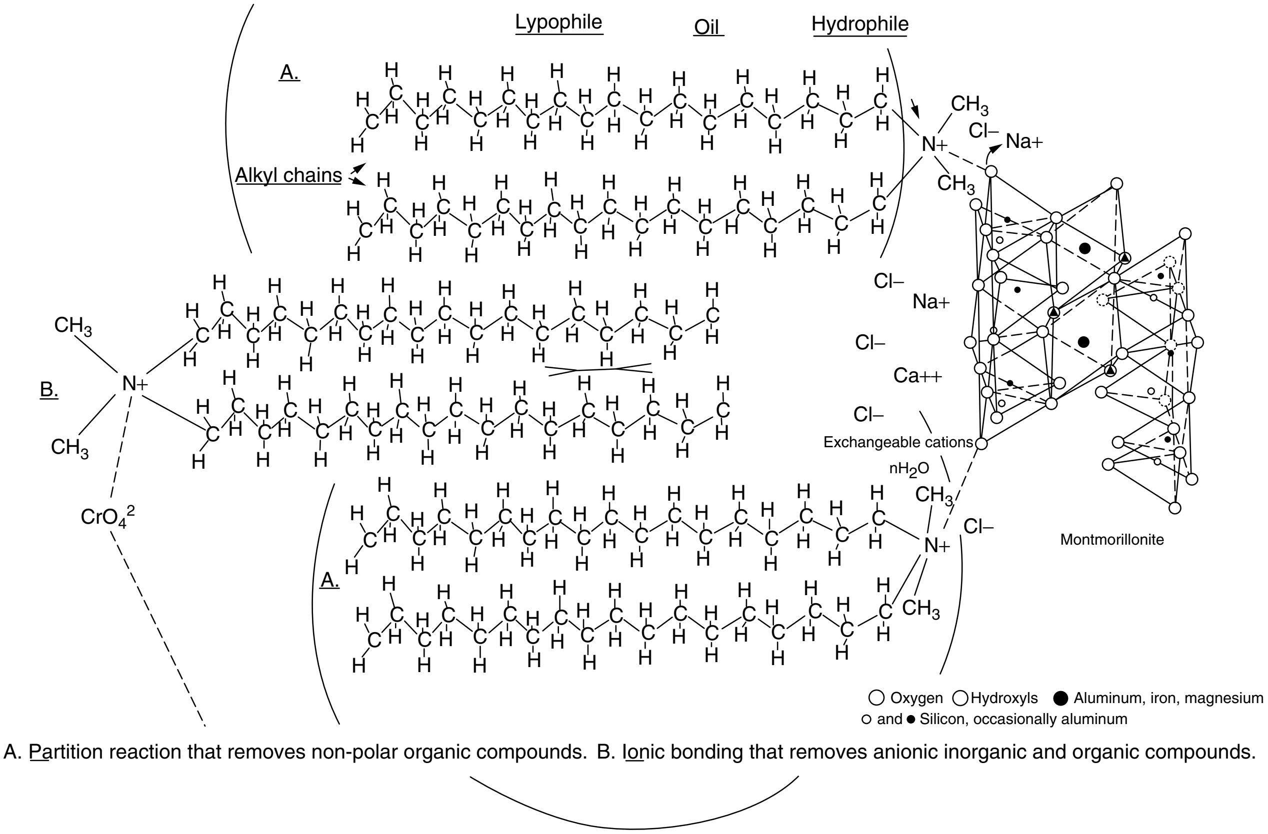

Figure 2. Model of an organoclay indicating partition and ion exchange adsorption mechanisms.

the effect of surfactants, which act as emulsifiers for oil, on the performance of organoclay. The effects are the same as on activated carbon; as solubility increases due to emulsification, the sorptive capacity decreases. Only the nonionic surfactant shows little effect on organoclay performance; probably it does not polarize the oils to any significant extent. Tests have shown that organoclay is just as effective in the removing surfactants from water as activated carbon (30). However, organoclay always prefers oil to any other compound; thus once the emulsion is split, the organoclay is used to remove the oil, followed by carbon which removes the surfactants.

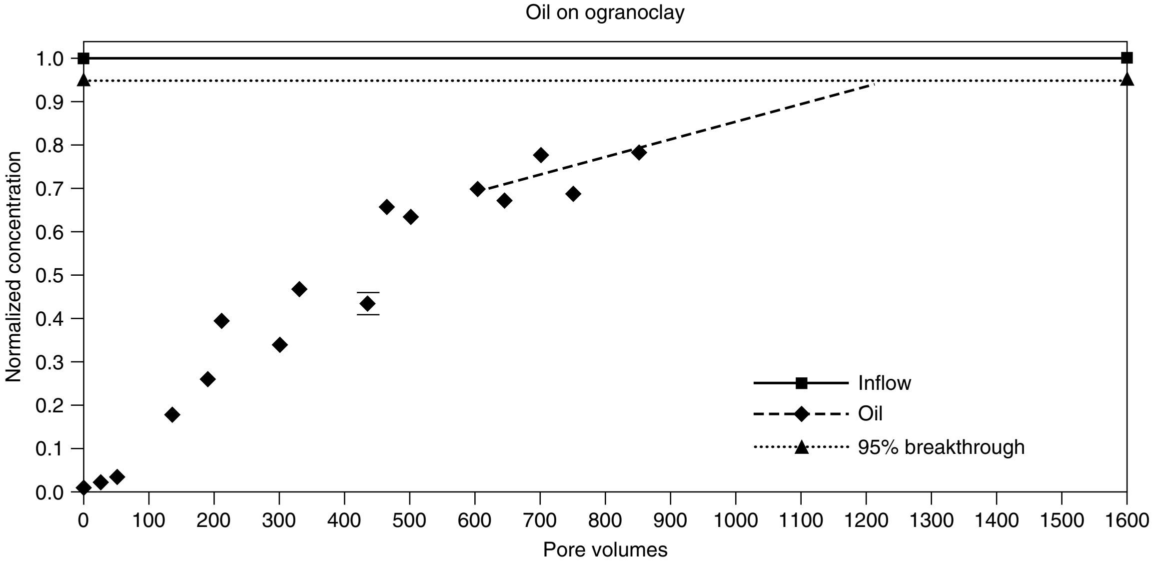

Figure 3. Breakthrough curve of oil on organoclay.

Table 1a. Sorbent Mass, Porosity, Flow Rate, and Residence Time

Mass Sorbent,		Porosity	Flow Rate,		Residence Time
kg	lb	%	mL/min	gal/h	min
0.141	0.31	0.3	15.45	0.23	8

Table 1b. 95% Breakthrough for Organoclay/Anthracite Given in Pore Volumes and Minutes Along with Estimated Mass of Oil Sorbed Per Mass of Sorbent in mg/kg, lb/lb, and On a Percent Basis

Breakthrough		Mass Sorbed		Mass Sorbed per Mass Sorbent		
PV	min	g/lb		g/kg	lb/lb	% by sorbent
1150	9,200	65.8	0.14	475	0.475	47.5

A minicolumn test was used to determine the ability of organoclay to remove such compounds as benzene, toluene, xylene, naphthalene, and PCB, and to be able to compare the data with those of bituminous activated carbon. The minicolumn method consists of spiking water with the compound to be evaluated and pumping that water through 1 gram of sorbent powder, which is tightly packed into a very small column. Pumping is performed until the influent concentration equals that of the effluent concentration (26). The sorbents are of 200 U.S. mesh size. This method is more comparable to large-scale, real situations than isotherms.

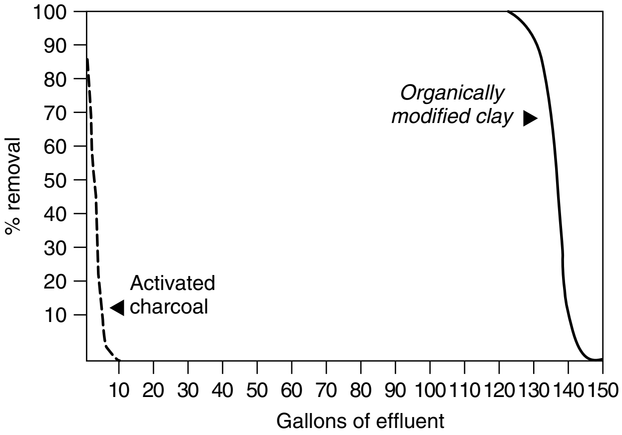

Figure 4. Removal of oil from water.

RESULTS

Figure 10 is a graphical illustration of minicolumn tests. When testing the removal capacity of the sorbents for benzene, toluene, *o*-xylene, and naphthalene, the organoclay performs similarly to carbon and performance improves as the solubility of the compounds decreases. Nonionic organoclay outperforms activated carbon with PCB 1260, as well as with motor oil, which is nearly insoluble in water. Surprisingly, organoclay removes methylene chloride much more effectively than activated carbon. The reasons are unclear, except that organoclay has an affinity for chlorinated compounds. Earlier results for vinyl chloride proved similar. It is theorized

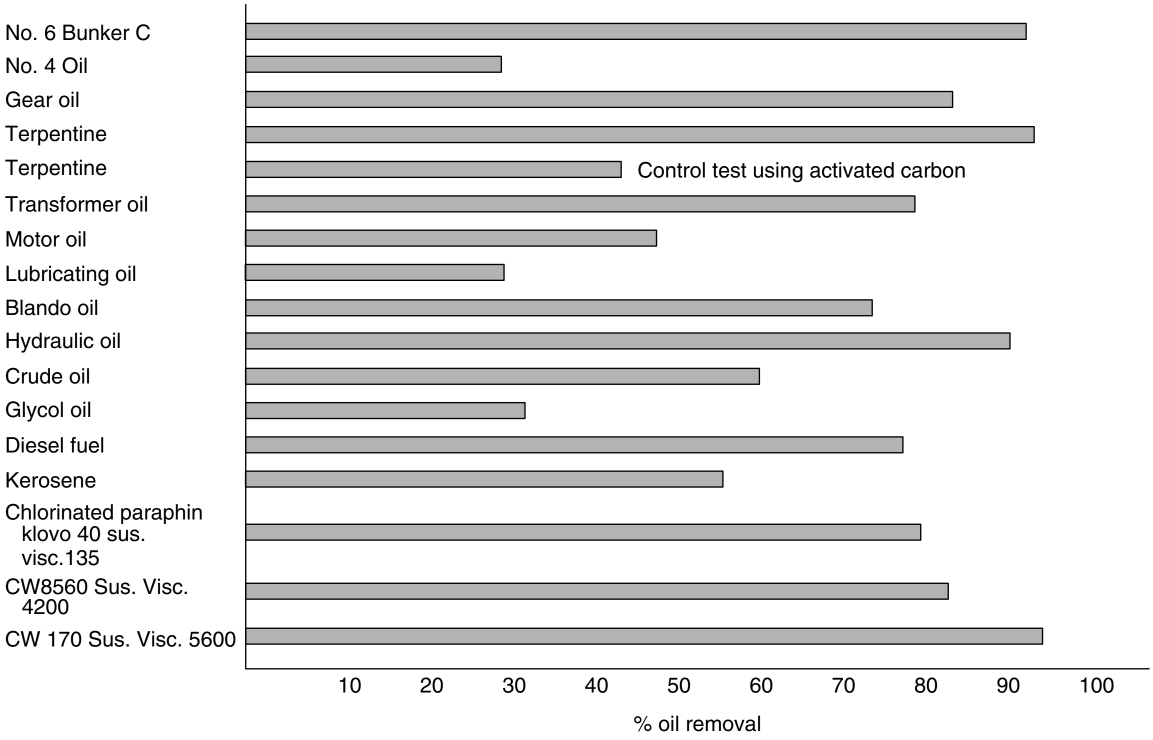

Figure 5. Removal of mineral oils from water by organoclay.

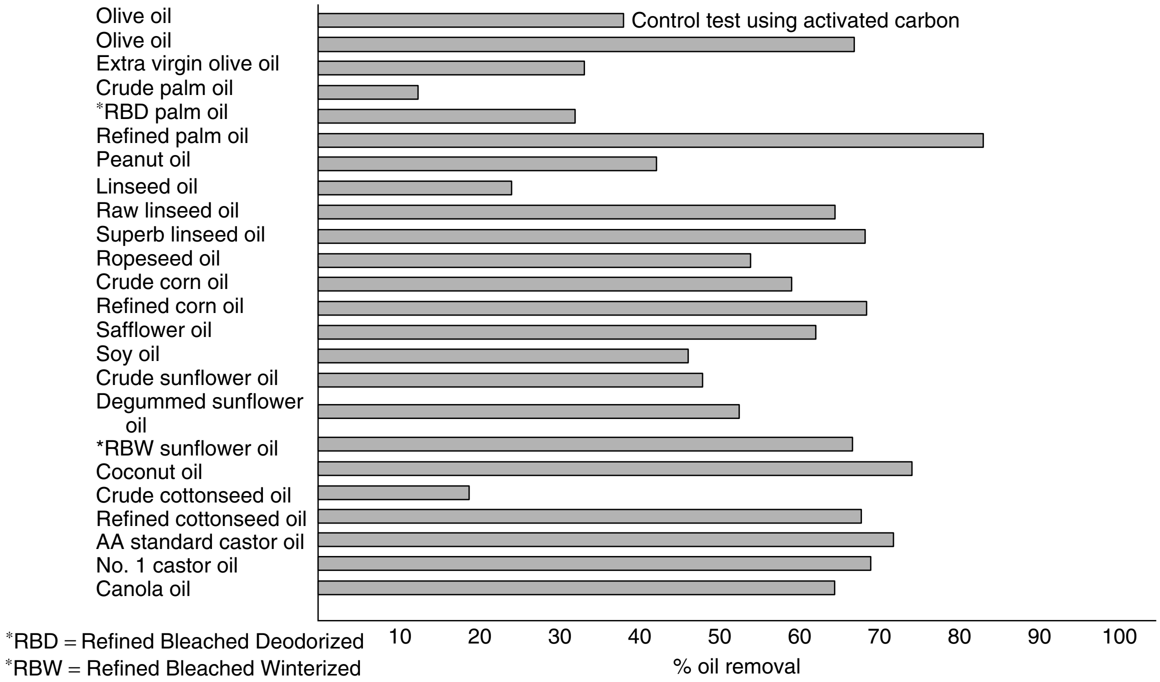

Figure 6. Removal of plant oils from water by organoclay.

that compounds such as methylene chloride have high electronegativity due to the presence of large amounts of halogens such as chlorides. The organoclay possesses positive charges on the surface due to the presence of inverted quaternary amine chains, so methylene chloride could chemically bond to these charges via their electronegativity. Therefore, two removal mechanisms, partitioning and ionic bonding, account for organoclay removal capacity for these compounds.

These tests were followed by a set of tests using a ternary effluent, which is a wastewater containing three different organic hydrocarbons, naphthalene, benzene, and toluene. This allowed the observation of competition among these compounds for adsorption sites. Of each compound, 900 mL/gram were added to water. It was possible to add that much naphthalene because benzene and toluene helped dissolve it. Usually, its solubility is only 10 mg/L. This concentrate was pumped separately through a column of organoclay, powdered activated carbon, and organoclay/carbon combined. In that case, the bottom of the minicolumn contained 0.5 grams of activated carbon, and the upper half contained 0.5 grams of organoclay. This was done to determine if the organoclay/carbon combination is more effective than each sorbent alone.

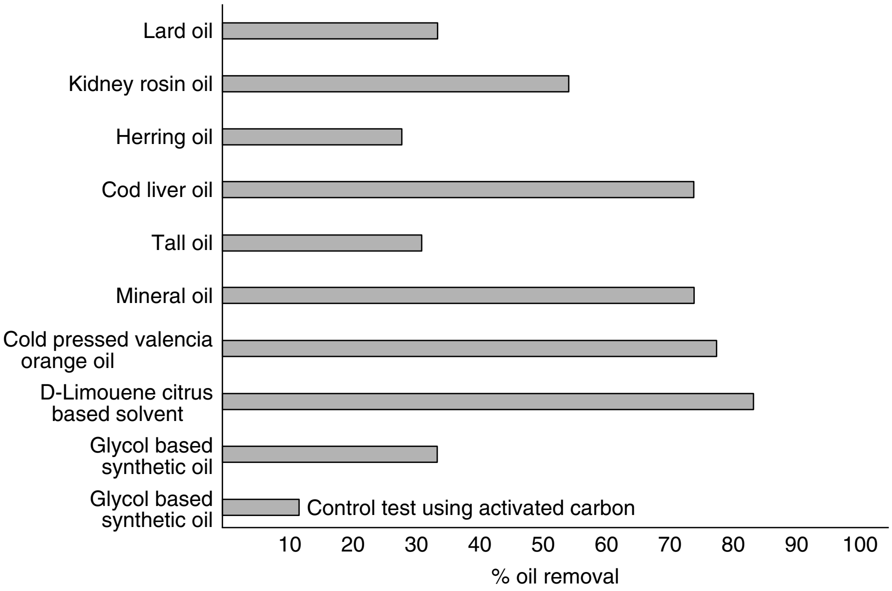

Figure 7. Removal of miscellaneous oils from water.

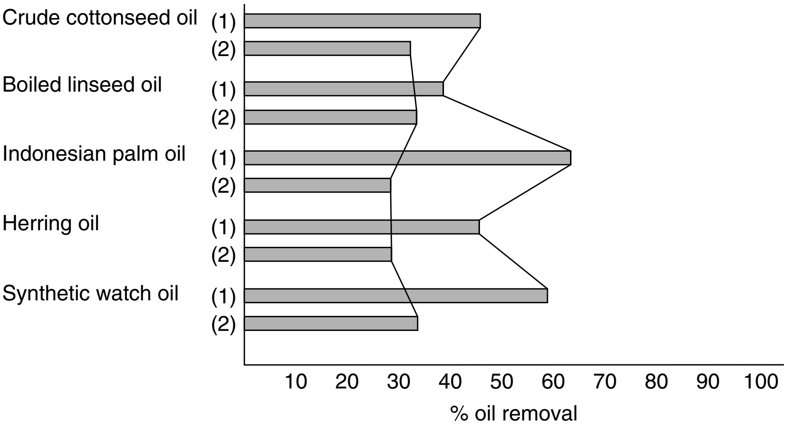

Figure 8. Removal of miscellaneous oils from water by different organoclays.

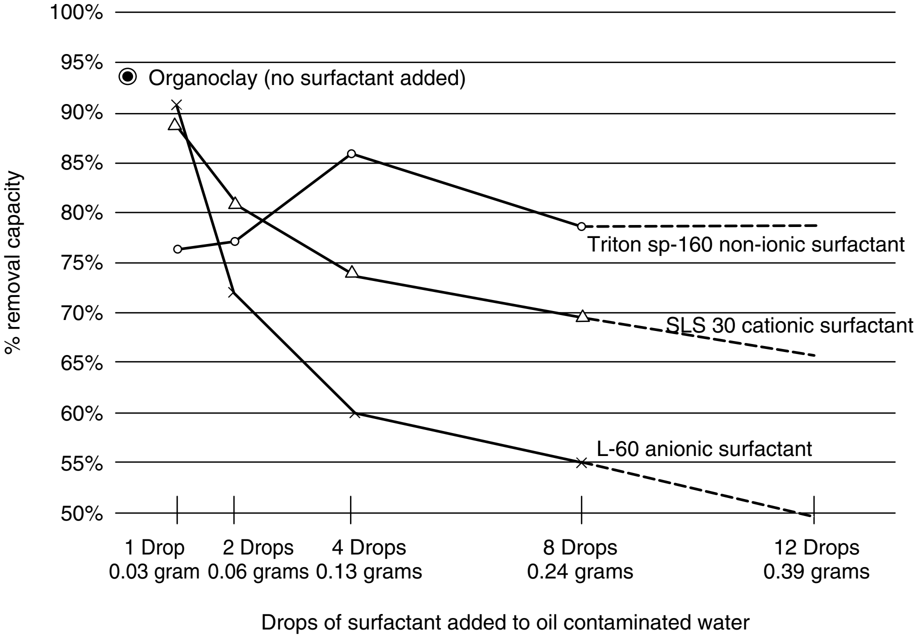

Figure 9. Percent removal capacity of organoclay for oil after adding a surfactant.

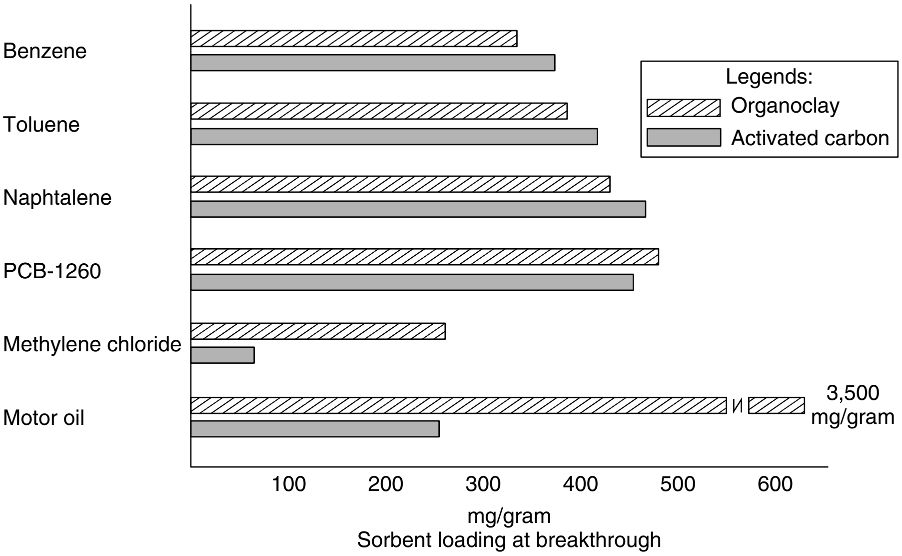

Figure 10. Minicolumn tests.

The results in Fig. 11 illustrate that benzene breaks through first, followed by toluene, and lastly naphthalene. This was expected based on their solubilities in water. The competition, however, is not 100% proof. The total adsorbed amounts are higher than the individually adsorbed amounts of the three solvents at breakthrough, probably because the geometrically arranged packing of the solvents of different sizes, either within the carbon pores or around the amine chains, favors a higher packing density.

The most important result is shown in the "total combined" graph. By placing the organoclay in front of the carbon, the removal capacity is doubled compared with the removal capacity of carbon and organoclay individually. This is also shown in a standard permeation column experiment with gasoline (Fig. 12). Again, the combination of organoclay/anthracite, followed by activated carbon, is much more effective in removing gasoline from water than either sorbent by itself, even though the amount of sorbent in each column is twice that in the combined column. Table 2 shows results from an actual groundwater cleanup project. The organoclay removes the oil completely and also a significant amount of other solvents and increases the effectiveness of activated carbon in removing volatile organic compounds. Because of the roll-off phenomenon, where less soluble compounds such as toluene and xylene, knock benzene off sorption sites within activated carbon, thus recontaminating the effluent, it is important to prepolish the water with organoclay, followed by activated carbon, which these results clearly illustrate.

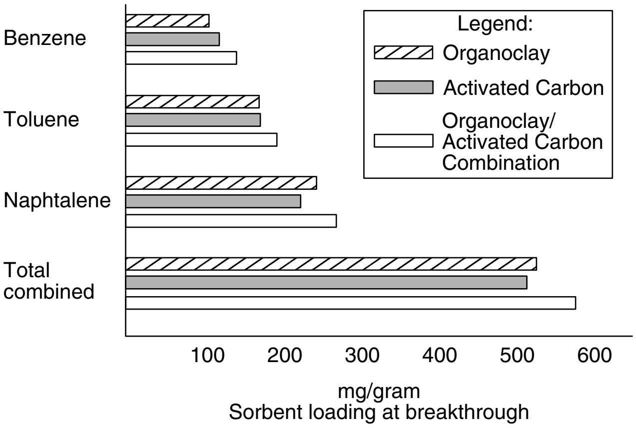

Figure 11. Minicolumn tests.

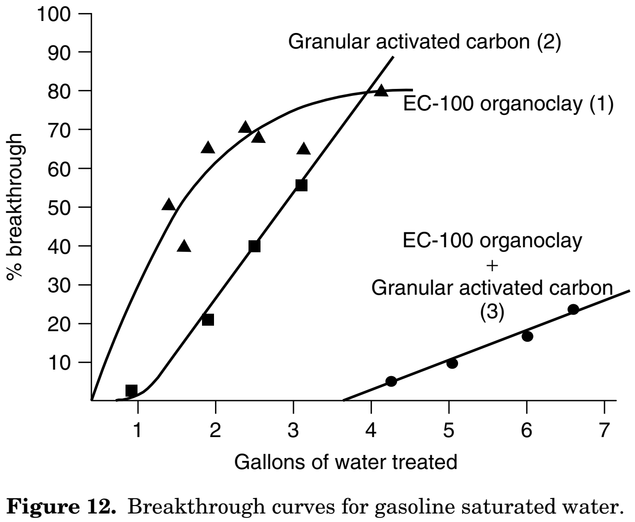

Figure 12. Breakthrough curves for gasoline saturated water.

The improved performance of an organoclay as the chlorination of the organic compound increases, along with a corresponding decrease in solubility in water, is shown in Figs. 13 and 14 (8). Organoclay is effective in removing phenol from water, but much more so, and far superior to activated carbon, for pentachlorophenol (PCP).

Freundlich isotherms confirm the results previously discussed, illustrated in Figs. 15–18. Figure 19, which shows the effectiveness of organoclay in removing methylene chloride, is impressive.

Figure 20 shows a time study, comparing how effectively nonionic organoclay and activated carbon remove turpentine from water. This test was conducted by pumping water spiked with turpentine through minicolumns and removing and analyzing a water sample every 50 minutes. Once the curve flattens out, saturation of the sample has occurred. These test results confirmed that the retention time of 6–10 minutes for water in an adsorber

Table 2. Organoclay/Carbon Sequence for Treating Contaminated Groundwater at an Abandoned Manufacturing Site

Organic Compound	Solubility, mg/L at 20–25 °C	Influent, μg/L	Effluent After O. Clay, μg/L	Effluent After Carbon, μg/L
Oil, mg/L	0.02	5.0	ND	ND
1,1,1-Trichloroethane	480–4000	42,622	26,044	ND
Trichloroethane	110	688	271	ND
1,1-Dichloroethene	335	285	ND	ND
Toluene	535	967	242	ND
pH		8.64	8.01	9.2

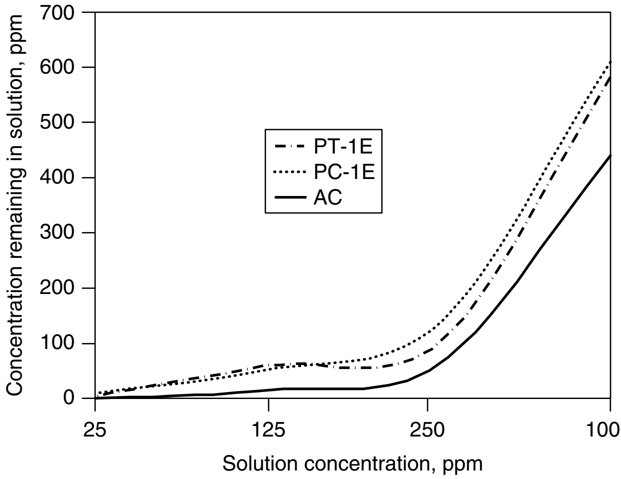

Figure 13. Adsorption of phenol by organoclay and activated carbon.

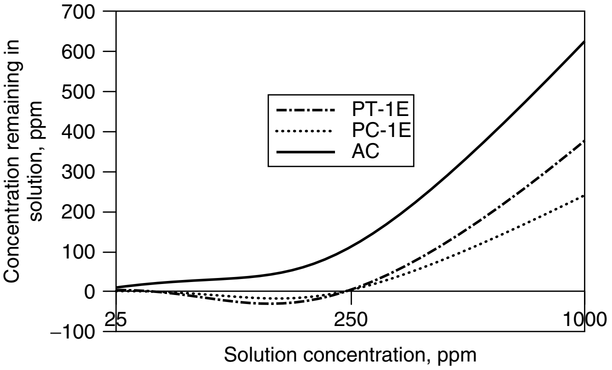

Figure 14. Adsorption of pentachlorophenol (PCP) by organoclay and activated carbon.

filled with either medium as is standard operating procedure, is applicable to both media. Another test showed that the organoclay may need a slightly longer retention time (8–12 minutes?) for more soluble benzene.

Bentonite is a natural ion exchange resin. Thus is an organoclay, even though a portion of its surface is covered with the chains of the quaternary amines. Some capacity for removing metals is still available in both media.

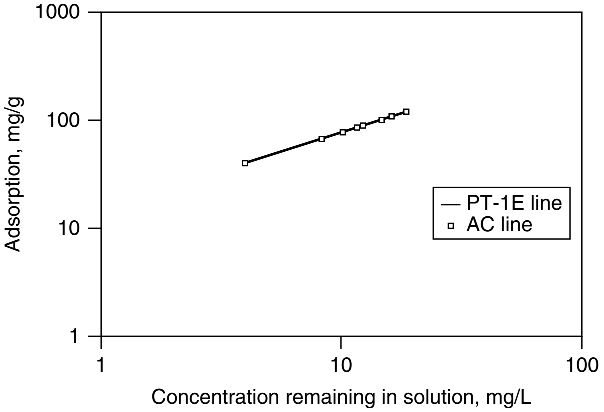

Figure 15. Isotherm showing a comparison of organoclay and activated carbon for *o*-xylene adsorption.

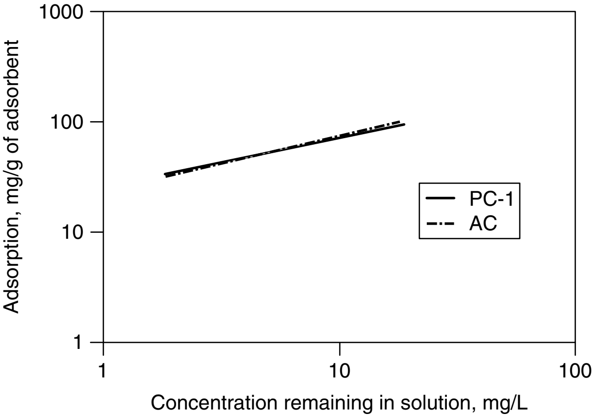

Figure 16. Isotherm showing a comparison of the removal capacity of organoclay and activated carbon for toluene.

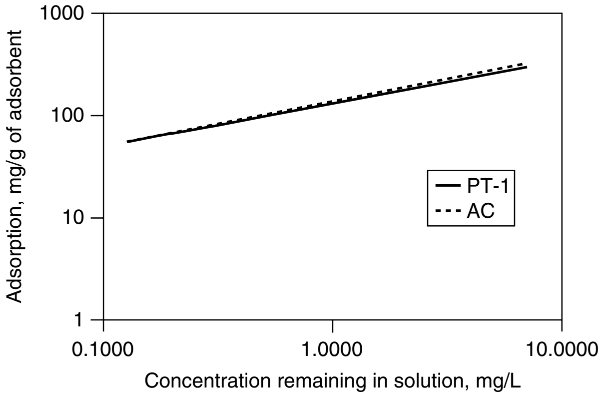

Figure 17. Isotherm showing a comparison of organoclay and activated carbon for the adsorption of naphthalene.

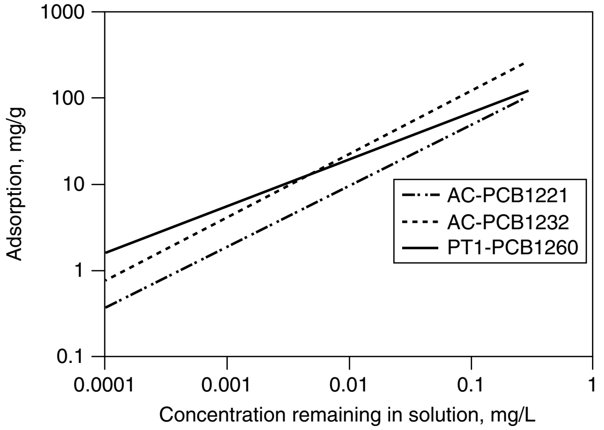

Figure 18. Isotherms showing a comparison of organoclay and activated carbon for the adsorption of PCB.

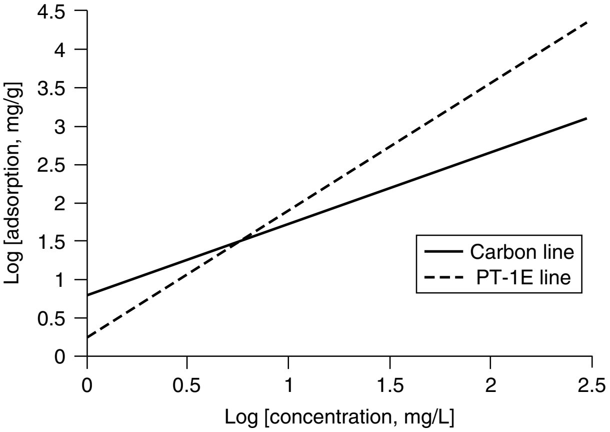

Figure 19. Isotherms showing a comparison of organoclay and activated carbon for the adsorption of turpentine.

Cation Exchange Capacity (CEC)

Organoclay/anthracite; 0.04 meq/gram; Organoclay: 0.04 meq/gram.

Surface Area (m^2/gram)

Organoclay/anthracite: 0.97 ± 0.05; Organoclay: 1.23 ± 0.69. A series of column tests were conducted in the

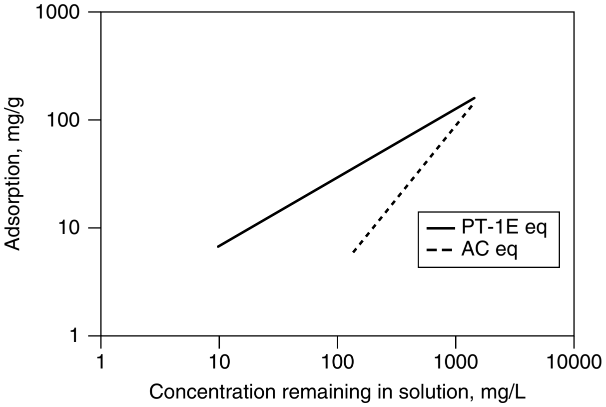

Figure 20. Isotherm showing a comparison of organoclay and activated carbon for the adsorption of methylene chloride.

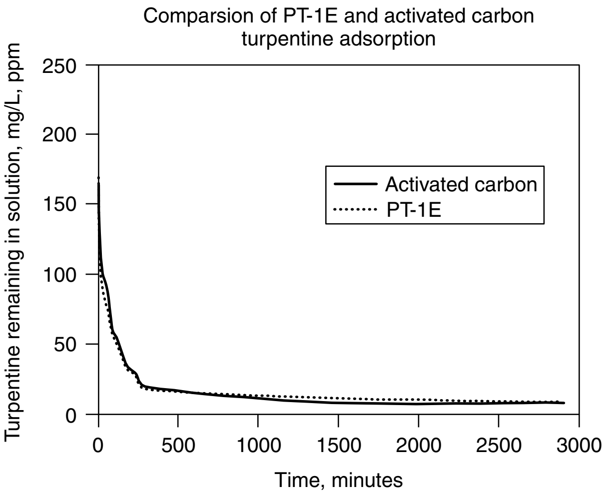

Figure 21. Time study showing a comparison of organoclay and activated carbon for the adsorption of turpentine from water.

same manner as that described for oil removal (Fig. 3), and the capacities of organoclay/anthracite and straight organoclay for removing various metals were determined. Bar charts were then constructed to illustrate the results, shown in Figs. 21 and 22 (24). The fact that the straight organoclay is not much better than organoclay/anthracite suggests that the diluting action of the anthracite results in improved access to sites on the organoclay. The U.S. Standard mesh size of the medium is 8 × 30 mesh. This capacity to remove small amounts of metals is not of great importance, but it can be a factor in calculating whether an ion exchange resin needs to be added to the treatment train if metal removal is required.

Figure 23 shows the ability of a cationic organoclay to remove chlorine from water. This isotherm proves that a cationic organoclay is an excellent medium for removing humic and fulvic acids (natural organic matter) and is far superior to bituminous activated carbon (Fig. 24). A minicolumnn test compares the removal capacity of cationic organoclay, bituminous activated carbon, and an "organotrap" ion exchange resin, again showing superior results for the cationic organoclay (Fig. 24). These results are thoroughly discussed in Alther (26).

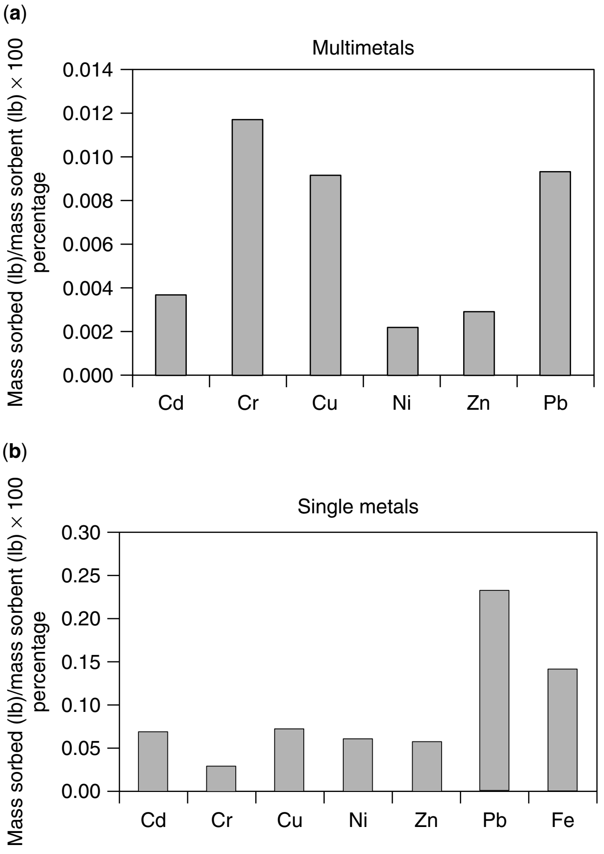

Figure 22. Adsorption of heavy metals from water by an organoclay/anthracite blend: Column study.

Table 3 shows the results of minicolumn tests to determine the removal capacity of cationic organoclay for negatively charged metals, including hexavalent chromium, selenite, arsenate, and fluoride. These results are in line with the capacity of this clay for chlorine removal.

An anionic organoclay was developed and tested in a minicolumn to show its capacity to remove cations, including ethylenediaminetriacetate (EDTA) (Fig. 25). The results are that coal based activated carbon removes 21.8% EDTA from the spiked solution, nonionic organoclay removes 47.3%, and anionic organoclay removes 70% of the EDTA.

Another set of minicolumn tests (Fig. 26) on water spiked with monoethanol amine revealed the following results: coal-based activated carbon removed 41.4%, nonionic organoclay removed 45.2%, anionic organoclay removed 29%, and cationic organoclay removed 79% of the monoethanolamine from the spiked water.

These results reveal the extreme versatility of organoclays in removing a variety of contaminants.

As a final test, the iodine numbers were determined for the organoclays, using the ASTM D-4607-94 testing method, to see if this test could be used to compare the two media. The results are as follows:

activated carbon (coal based): 700–900
nonionic organoclay: 275
cationic organoclay: 190
anionic organoclay: 410

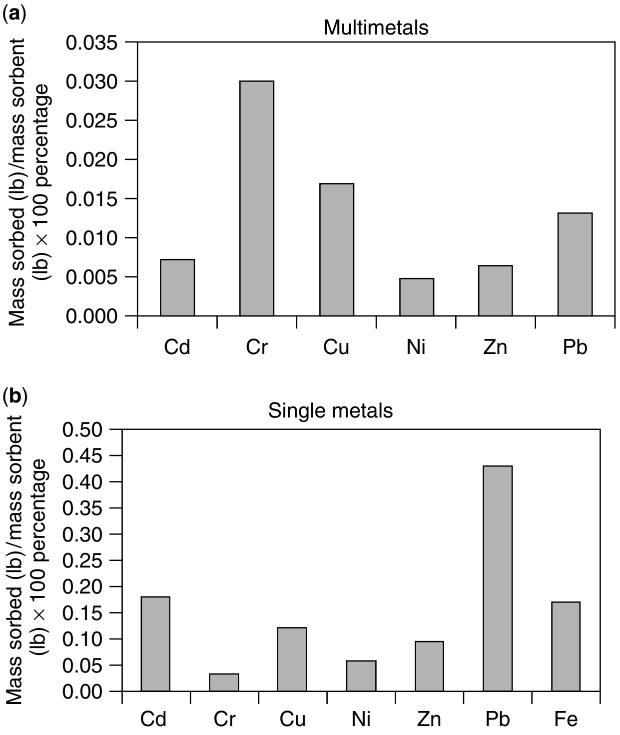

Figure 23. Adsorption of heavy metals from water by straight organoclay: Column study.

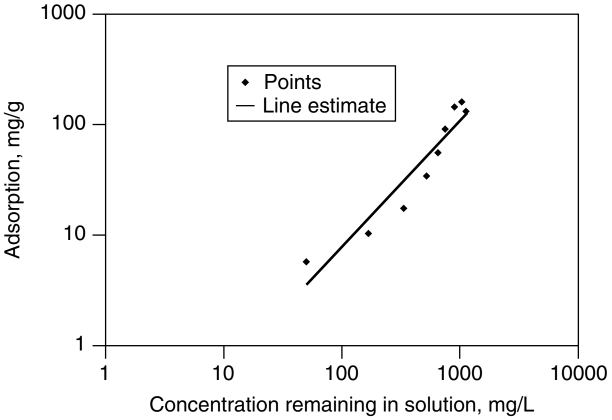

Figure 24. Isotherm showing the adsorption capacity of cationic organoclay for chlorine from water.

Table 3. Laboratory Column Study with Cationic Organoclay

Arsenate

Amount of cationic organoclay used: 16.5 grams (16 × 30 mesh)

Amount of spiked water passed through column: 8.31 liters

Removal capacity of cationic organoclay for arsenate by weight: 0.3%, or 3 grams per 1000 grams cationic organoclay

Fluoride

Amount of cationic organoclay used for fluoride removal: 16.5 grams (16 × 30 mesh)

Amount of spiked water passed through column: 0.54 liters.

Removal capacity of cationic organoclay by weight: 0.1% or 1 gram fluoride per 1000 gram of cationic organoclay

Chromate

Batch test: 100 mL water spiked with 5 mg/L hexavalent chromium, added 2 grams powdered cationic organoclay:

Removed 96% of 5 mg/L.

Conclusion

Aside from the extreme efficiency with which cationic organoclay removes humic acids from water, it also removes negatively charged metals, such as those shown above, and selenite.

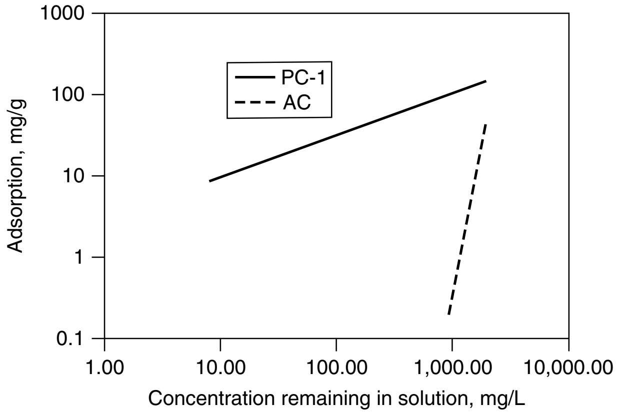

Figure 25. Isotherm showing the adsorption capacity of cationic organoclay with that of activated carbon.

Performance of activated carbon in this test is superior because iodine molecules fit inside the pores and do not cause blinding and all adsorption on the organoclay takes place outside the clay platelets. Therefore, it is concluded that this test is of no use for organoclay, although the results are interesting.

Below are several case histories, which give a practical appreciation of organoclays to the engineer who designs remediation systems.

Case Histories

1. A creosote superfund site on the East Coast installed a pump and treat system consisting of two filter vessels, each containing 20,000 lb activated carbon. The flow rate was 170 gpm. The COD consisted of 40–60 ppm, including benzene, VOCs, and phenols. The activated carbon lasted about 2 weeks with a breakthrough of 7–12 ppm COD; then it had to be replaced. After another vessel containing 19,000 lb of organoclay was installed, the effluent after the activated carbon was not detectable. Furthermore, there was a TSS content of 32–35 ppm (discharge limit is 40 ppm), primarily due to the presence of ferric iron. Once the organoclay was installed, the TSS content in the effluent was 3 ppm because the organoclay, a bentonite, also removes heavy metals by ion exchange.

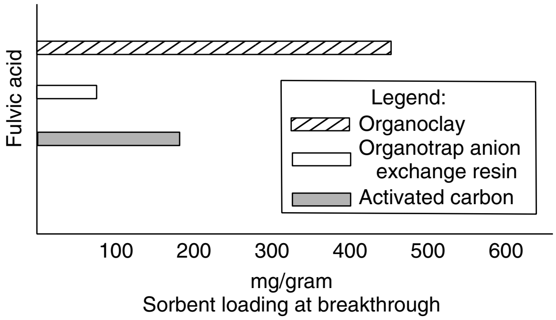

Figure 26. Microcolumn study data comparing the adsorption capacity of cationic organoclay and activated carbon for fulvic acids.

2. An old wood-treating site in Colorado is situated above an aquifer, which had a concentration of 30 ppm of an oil and 25 ppm of PCP. The discharge limit for PCP is 50 ppb. When an activated carbon system was installed, replacement was required within 1 month. After 20,000 lb of organoclay was installed prior to the activated carbon, discharge limits were met, and replacement was required only after 12 to 15 months.
3. An old railroad site in southeastern United States, where railroad ties were once treated with creosote, required excavating the soil and thermally treating it to destroy the creosote. A condensate built up that contained PCP. Rather than accepting the high cost of incinerating the condensate water, it was passed through an organoclay/carbon system and discharged locally.

This brief description of the use of organoclays for water treatment should convince the reader of their usefulness. Anyone who is interested in more detail should consult the references.

BIBLIOGRAPHY

1. Alther, G.R. (2002). Removing oils from water with organoclays. *J. Am. Water Works Assoc.* **94**(7): 115–121.
2. Alther, G.R. (1997). Working hand in hand with carbon for better results. *Soil & Groundwater Cleanup* **May**: 17–19.
3. Sheng, G., Xu, S., and Boyd, S.A. (1996). Co-sorption of organic contaminants from water by hexadecyltrimethylammonium exchanged clays. *Water Res.* **30**(6): 1483–1489.
4. Smith, J.A. and Jaffe, P.R. (1994). Benzene transport through landfill liners containing organophilic bentonite. *J. Environ. Eng.* **120**(6): 1559–1577.
5. Wolfe, T.A., Demiral, T., and Bowman, E.R. (1985). Interaction of aliphatic amines with montmorillonite to enhance adsorption of organic pollutants. *Clays Clay Miner.* **33**(4): 301–311.
6. Jaynes, W.F. and Boyd, S. (1991). Clay mineral type and organic compound sorption by hexadecyltrimethylammonium-exchanged clays. *Soil Sci. Soc. Am. J.* **55**(1): 43–48.
7. Jordan, J.W. (1949). Organophilic bentonites. I. Swelling in organic liquids. *J. Phys. Colloid Chem.* **53**(2): 294–306.
8. Alther, G.R. (2004). A winning combination. *Water Wastewater Prod.* **July/August**: 22–27. Available at www.wwp-online.com.
9. Alther, G.R., Evans, J.C., and Tarites, R. (1991). The use of organoclays for stabilization of hazardous wastes. *Proc. 4th Annu. Hazardous Waste Manage. Conf./Central*, pp. 547–552.
10. Alther, G.R., Evans, J.C., and Pancoski, S. (1989). *A Composite Liner System to Retain Inorganic and Organic Contaminants*. Hazardous Materials Research Institute, Superfund 89, Silver Spring, MD.
11. Alther, G.R., Evans, J.C., and Pancoski, E.S. (1988). Organically modified clays for stabilization of organic hazardous waste. Hazardous Materials Control Research Institute *9th Natl. Conf., Superfund 88*, Silver Spring, MD, pp. 440–445.
12. Cowan, C.T. and White, D. (1962). Adsorption by organoclay complexes, I. *Proc. 9th Natl. Clay Conf.*, pp. 4509–4517.
13. Slabough, W.H. and Hanson, D.B. (1969). Solvent selectivity by an organoclay complex. *J. Colloid Interface Sci.* **29**(3): 460–463.
14. Street, G.B. and White, D. (1963). Adsorption by organoclay derivatives. *J. Appl. Chem.* 288–291.
15. Lagaly, G. (1981). Characterization of clays by organic compounds. *Clay Miner.* **16**: 1–21.
16. Lagaly, G. (1979). The "layer charge" of regular interstratified 2:1 clay minerals. *Clays Clay Miner.* **27**: 1–10.
17. Fashan, A., Tittlebaum, M., and Cartledge, F. (1993). Nonionic organic partitioning onto organoclays. *Hazardous Waste Hazardous Mater.* **10**(3): 313–322.
18. Mortland, M.M. (1970). Clay–organic complexes and interactions. *Adv. Agron.* **22**: 75–117.
19. Theng, B.K.G. (1974). *The Chemistry of Clay-Organic Reactions*. Halstead Press, John Wiley & Sons, New York.
20. Alther, G.R. (1999). Organoclays remove oil, grease, solvents and surfactants from water. *CleanTech 99 Proc.*, Witter, Flemington, NJ, pp. 72–79.
21. Alther, G.R. (1998). Put the breaks on oil and grease. *Chem. Eng.* **March**: 82–88.
22. Alther, G.R. (1996). Organically modified clay removes oil from water. *Waste Manage.* **15**(8): 623–628.
23. Alther, G.R. (1995). Organically modified clay removes oil from water. *Waste Manage.* **15**(8): 641–650.
24. Jaynes, W.F. and Vance, G.F. (1996). BTEX sorption by organo-clays: Cosorptive enhancement and equivalence of interlayer complexes. *Soil Sci. Soc. Am. J.* **60**: 1742–1749.
25. Mortland, M.M., Shaobai, S., and Boyd, S.A. (1986). Clay–organic complexes as adsorbents for phenols and chlorophenols. *Clays Clay Miner.* **34**(5): 581–585.
26. Alther, G.R. (2001). Organoclays remove humic substances from water. In: *Humic Substances*. E.A. Ghabbour and G. Davis (Eds.). Royal Society of Chemistry, Cambridge, UK, pp. 277–288.
27. Tillman, F.D., Bartelt-Hunt, S.L., Craver, V.A., Smith, J.A., and Alther, G.R. (2004). Relative metal ion sorption on natural and engineered sorbents: Batch and column studies. *Environ. Eng. Sci.* (submitted).
28. Alther, G.R. (2002). *Using Organoclays to Enhance Carbon Filtration. Waste Management* 22. Pergamon Press, New York, pp. 507–513.
29. Alther, G.R. (2000). Maximize water cleanup performance. *Environ. Prot. Mag.* **Feb.**: 37–39. Available at www.wp-online.com.
30. Alther, G.R. (1999). Organoclay filtration technology for oil removal. *Fluid/Part. Sep. J.* **12**(2): 96–102.

COMBINED SEWER OVERFLOW TREATMENT

C.A. Prochaska
A.I. Zouboulis
Aristotle University of Thessaloniki
Thessaloniki, Greece

INTRODUCTION

Combined sewer overflows (denoted hereafter as CSOs) occur when flows exceed the hydraulic capacity of either the wastewater treatment plant (denoted as WWTP) or the collection system that transports the combined flow of storm water and sanitary sewage to the WWTP. The principal components of a combined sewer system include (1) the contributing drainage area (catchment) and wastewater sources, (2) the combined sewer pipe network and interceptor(s), (3) the regulator and diversion structures, and (4) the CSO outlets (Fig. 1).

When an overflow occurs, the excess flows tend to be discharged into the neighboring receiving body of surface water. CSOs typically discharge a variable mixture of raw sewage, industrial/commercial wastewater, polluted runoff, and scoured materials that build up in the collection system during dry weather. These discharges contain a variety of pollutants that may adversely impact the receiving waterbody, including pathogenic microorganisms, viruses, cysts, and chemical and floatable materials. Health risks associated with bacteria-laden water may result from dermal contact with the discharge, from ingestion of contaminated water, as well as from consumption of fish or shellfish.

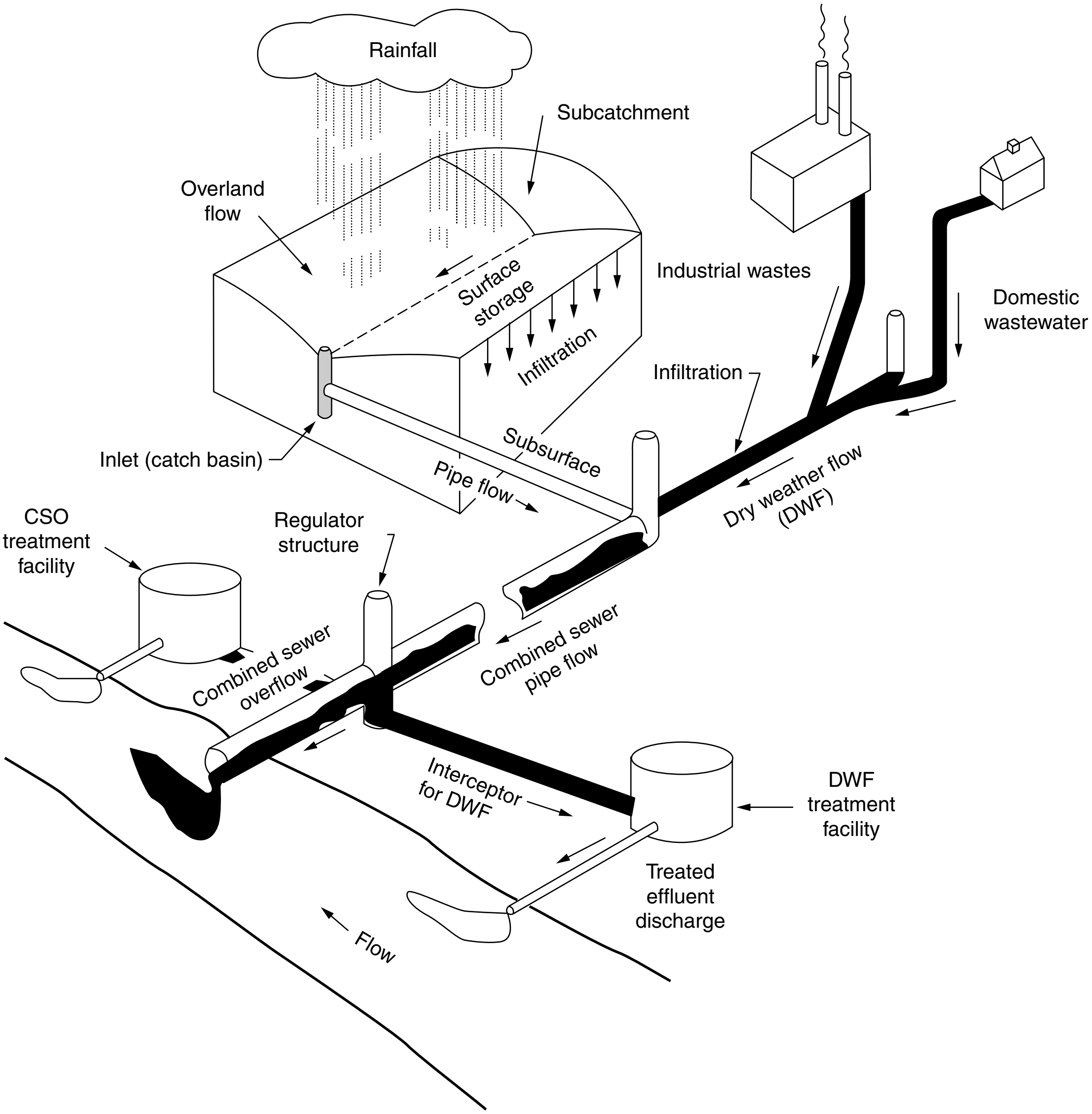

Figure 1. Schematic diagram of a combined sewer system (2).

The methods used to treat CSOs can be classified as physical, chemical, and biological and methods that include a combination of some or all the above, such as treatment by constructed reed beds (1).

PHYSICAL TREATMENT

Physical treatment alternatives include sewer separation, retention basins, swirl/vortex technologies, screening, netting systems for floatable control, dissolved air flotation, and filtration. Most of these physical unit operations have been in use for many years and are considered reliable. Physical treatment operations are usually flexible enough to be readily automated and can operate over a wide range of flows. They can also stand idle for long periods of time without affecting treatment efficiency (1).

Sewer Separation

Separation is conversion of a combined sewer system into separate storm water and sanitary sewage collection systems. This alternative, historically considered the ultimate answer to CSO pollution control, has been reconsidered in recent years because of increased cost and major disruptions to traffic and other daily community activities from separated collection systems. Several potential benefits of sewer separation might warrant its consideration in specific cases:

1. Eliminating CSOs and preventing untreated sanitary sewage from entering the receiving waters during wet weather. Sanitary sewage is a more objectionable source of certain pollutants, such as TSS, sanitary floatables, and bacteria.
2. Reduced volume of flow to be treated at the publicly owned treatment works (POTWs), thus reducing operating and maintenance (O & M) costs by eliminating surface runoff inflows during wet weather.
3. Reduced infiltration and excess flow to a POTW for new sanitary sewer construction, replacing old combined sewers.
4. Reducing upstream flooding, as well as overflows, when the existing combined sewers are undersized and back up frequently during storms.
5. Being more effective and economical than treatment facilities for remote segments of a combined sewer system, serving relatively small areas.

Retention Basins

CSO retention basins (RBs) capture and store some of the excess combined sewer flow that would otherwise be bypassed to receiving waters. Stored flows are subsequently returned to the sewer system during dry weather, when the in-line flows are reduced and more capacity is available at the treatment facility. RBs can be designed to control both flow rate and water quality.

Figure 2 shows an example of a multistage CSO RB that has some treatment capabilities. This facility handles peak flows by routing them through a mechanical bar screen and then pumping them into the first compartment. The main function of the first compartment is to allow primary settling and grit removal. If the flows continue to rise, the

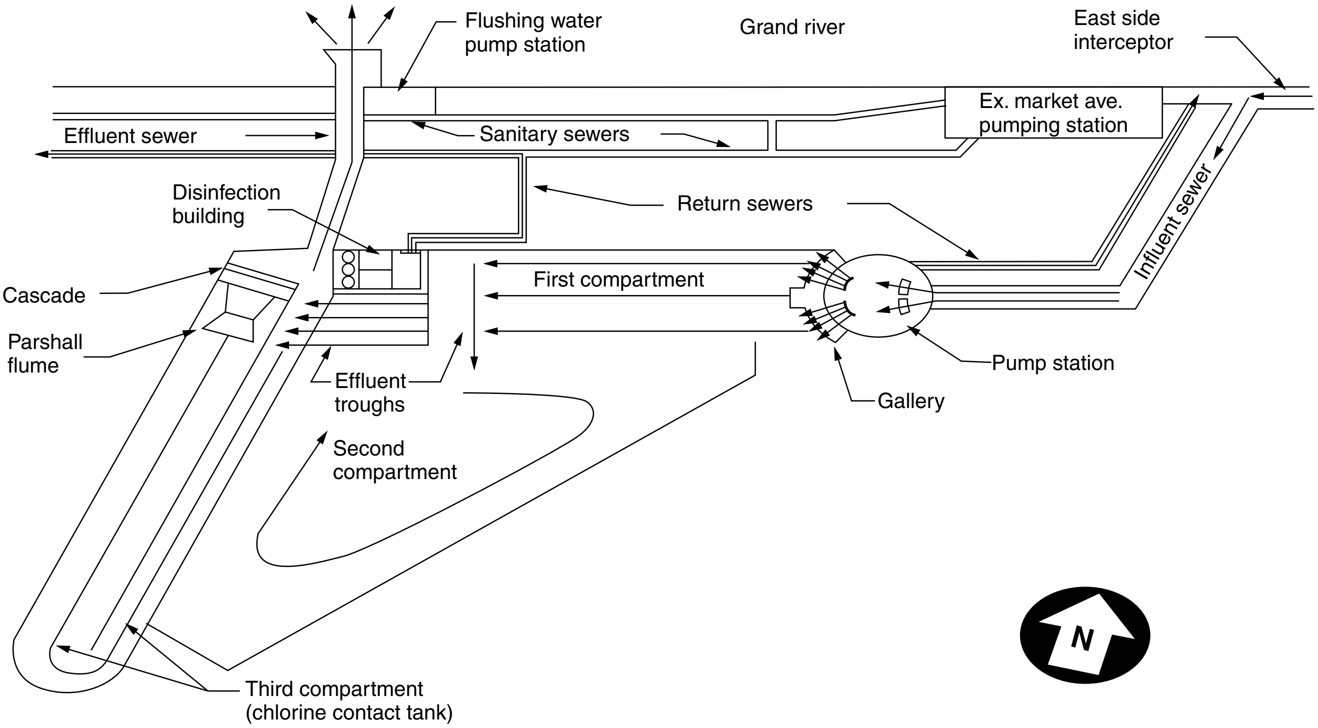

Figure 2. Multistage CSOs RBs (3).

first compartment fills and then spills over into the second compartment. This compartment is designed specifically to store most of the overflow from the first compartment. The second compartment is also equipped with a floor wash system that flushes all settled sediments into a collection trough. If the flows continue to rise, the water spills over into a series of troughs, where sodium hypochlorite is applied for disinfection. The flow is then routed to a contact tank (third compartment), which eventually returns the water to the nearby surface waters (2).

These are the primary concerns in the operation of RBs:

1. Managing flows to and from the retention basin.
2. Preventing the combined sewage from becoming septic or handling the wastewater appropriately after it has become septic.
3. Removing accumulated solids and floatables.
4. Disinfecting basin overflows to receiving waters.

Swirl/Vortex Technologies

Solids separation devices, such as swirl concentrators and vortex separators, have been used in Europe and (to a lesser extent) in the United States. These devices are relatively small, compact solids separation units with no moving parts. A typical vortex-type CSO solids separation unit is illustrated in Fig. 3. During wet weather, the outflow from the unit is throttled, causing the unit to fill and to self-include a swirling vortex-like flow regime. Secondary flow currents rapidly separate settleable grit, as well as floatable matter. The concentrated foul matter is intercepted for treatment, whereas the cleaner, treated flow can be discharged to receiving surface waters. These devices are usually intended to operate under very high hydraulic flow regimes.

Screening

Generally, there are two types of bar screens, coarse and fine. Both are used at CSOs control facilities; with each type provides a different level of solids removal efficiency. Although there is no industrial standard for classifying screens based on aperture size, coarse bar screens generally have a 0.04 to 0.08 m clear spacing between bars, whereas fine screens generally have rounded or slotted openings of 0.3 to 1.3 cm clear space.

Coarse Screens. Coarse screens are constructed of parallel vertical bars and are often referred to as *bar racks* or *bar screens*. In CSO control and treatment facilities, coarse screens are usually the first unit of equipment in the system. These screens are usually set at 0 to 30° from the vertical and are cleaned by an electrically or hydraulically driven rake mechanism that removes the collected material from the screen continuously or

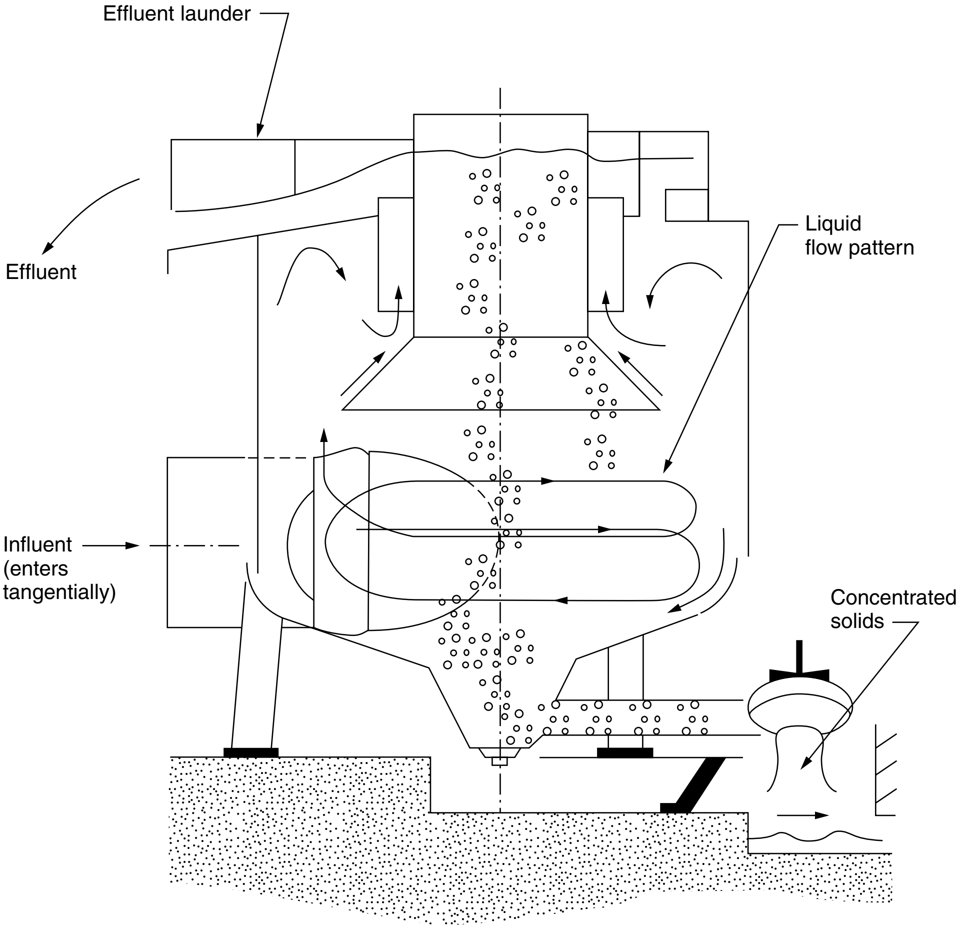

Figure 3. Cross section through a typical vortex-type solids separation device (1).

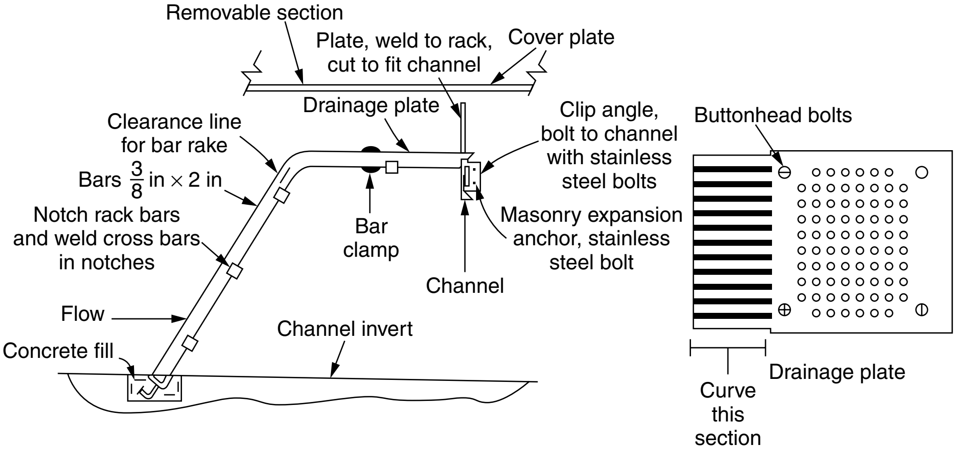

Figure 4. Diagram of trash rack used for treating CSOs (2).

periodically. The most common type of bar screen used at CSO control facilities is a trash rack. Trash racks typically have 0.04 to 0.08 m clear spacing between bars. Figure 4 is a diagram of a typical trash rack.

Fine Screens. Fine screens at CSO facilities typically follow coarse bar screening equipment and provide the next level of physical treatment in removing the smaller solid particles from the waste stream. Both fixed (static) and rotary screens have been used in CSO treatment facilities.

Fixed fine screens are typically provided with horizontal or rounded slotted openings of 0.02 to 1.27 cm. The screens are usually constructed of stainless steel in a concave configuration, at a slope of approximately 30° from the vertical. Flow is discharged across the top of the screen. The flow then passes through the slotted openings, and solids are retained on the screen surface. Solids are discharged from the screen surface by gravity and by washing onto a conveyor belt or other collecting system.

Rotary fine screens include externally and internally fed screens. Externally fed screens allow the wastewater to flow over the top of the drum mechanism and through the screens surfaces, while collecting solids onto the screen surface. As the screen rotates, a system of cleaning brushes or sprayed water removes debris from the drum. Internally fed systems discharge wastewater in the center of the drum, allowing the water to pass through the screen into a discharge channel, while solids are removed from the screen surface by cleaning brushes or a water spray.

In response to the need for solids and floatables control during storms, proprietary screen products, such as the ROMAG^{TM} screen, have been designed for wet weather applications (Fig. 5). The ROMAG^{TM} screen partitions the flow, sending screened flow to the CSO discharge point, while keeping solids and floatables in the flow directed toward the sanitary sewer. This screen works as follows: excess flow enters the screening chamber, flows over a spill weir, and proceeds through the screen into a channel, which discharges flow to a neighboring receiving waterbody. Floatables trapped by the screen move laterally along the face of the screen via combs/separators to the transverse end section of the pipe, where they can be directed to the sanitary sewer line for ultimate removal at the wastewater treatment plant. Screen blinding is prevented by a hydraulically driven rake assembly.

Netting Systems for Floatables Control

Floatables control technologies are designed to reduce or eliminate the visible solid waste that is often present in CSO discharges. The Netting Trash-Trap^{TM} system is a modular floatables collection system, located at the CSO outfall. It uses the passive energy of the effluent stream to drive the floatable materials into disposable mesh bags. These bags are suspended horizontally in the CSO flow stream within a support structure. The construction methodology and method of installation at the outfall are determined site-by-site. Ever since, several other end-of-pipe, but also in-line configurations have been developed and implemented.

The standard nets used in the system are designed to hold up to 0.7 m^3 of floatables and a weight of 227 kg each. For the floating units, the effluent stream and the collected floatables are directed into the bags by two floating booms and curtains, which run from the front corners of the pontoon to either side of the outfall, where they attach to a vertical piling that has a roller

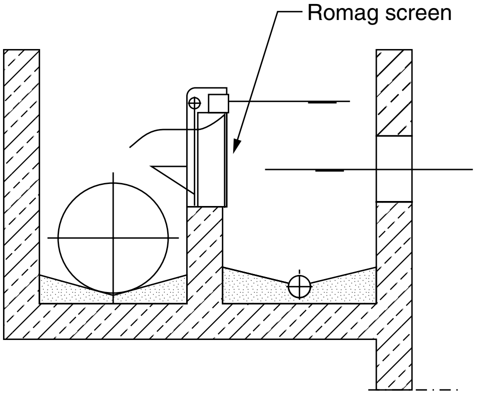

Figure 5. ROMAG^{TM} "combing" mechanical screen (vertical) for CSO floatables control (4).

mechanism or a shoreline support. This design allows the boom to float and accommodate changes in the water level. The extended curtains are weighted to conform to the water bottom. The maximum high water level, expected at the site, determines the depth of the curtain. Certain modifications of the outfall design may include adding structural support, attaching structural struts and strut anchor support, and adding foundations.

Dissolved Air Flotation

Dissolved air flotation (DAF) removes solids by introducing fine air bubbles into the CSO stream. The air bubbles can attach to solid particles suspended in the liquid, causing the solids to float toward the surface, from where they can be skimmed off. This technology has been tested in several CSO applications. A major advantage of DAF is its relatively high overflow rate and short detention time, which results in reduced facility size, compared to conventional sedimentation. Oil and grease are also more readily removed by DAF. Operating costs for DAF are high, due to larger energy demand, and skilled operators are usually required for efficient operation.

Filtration

Dual media high rate filtration has been piloted for treating of CSO flows. A two-layer bed, consisting of coarse anthracite particles on top of less coarse sand, was used. After backwash, the less dense anthracite remains on top of the sand. Filtration rates of 16 gal/ft^2/min or more were used, resulting in substantially smaller space requirements, compared with sedimentation. Demonstration test systems included pretreatment by fine-mesh screens. The addition of chemical coagulant agents improved the performance considerably. Filtration is more appropriately applied after pretreatment by fine screening. Operation may be automated but tends to be rather O & M intensive. Intermittently operated sand filters can also be used for CSO treatment and show considerable promise for larger scale operations (3).

CHEMICAL TREATMENT (MAINLY DISINFECTION)

Chlorine

Chlorine has long been the disinfectant of choice for most disinfection systems. It offers reliable reduction of pathogenic microorganisms at reasonable operating costs. Disinfection by chlorine is the most common method used to kill pathogenic microorganisms at WWTPs, but this methodology may not be feasible at all CSOs for several reasons:

1. CSOs occur intermittently, and their flow rate is highly variable, thus making it difficult to regulate the addition of disinfectant.
2. CSOs have high concentrations of suspended solids.
3. CSOs can vary widely in temperature and bacterial composition.
4. Disinfectant residuals, following the use of chlorine, may be prohibited from entering receiving waters.
5. CSO outfalls are often located in remote areas and thus, may require automated controls for the disinfection systems.

In addition to these problems, increased health and safety concerns for using chlorine to disinfect CSOs have prompted the development of alternative disinfectant agents/methods, which often present fewer problems and health hazards. Alternative methods to chlorine addition have been developed and evaluated for the continuously disinfecting wastewater discharges to small streams or sensitive waterbodies and are now also being considered for treating CSOs and other episodic discharges. These include the use of chlorine dioxide, the application of ozonation or of ultraviolet radiation, the addition of peracetic acid, and electron beam irradiation (e-beam) (5).

Ozone

Ozone is a strong oxidizer and is applied to wastewater as a gas mixture with air. Its use in CSO treatment facilities for wastewater disinfection is relatively new; few facilities are currently using ozone for disinfection. This can be attributed to the higher initial capital costs of ozone generating equipment. Ozone is equal or superior to chlorine in “killing” power for pathogenic microorganisms, but it does not cause the formation of harmful by-products (halogentated organics), as does chlorination (5).

Ultraviolet Radiation

UV radiation is electromagnetic radiation used for disinfection. UV disinfection incorporates the spectrum of light between 40 and 400 nm. Germicidal properties range between 200 and 300 nm; 260 nm is the most lethal. The primary method for using UV disinfection is to expose wastewater to a UV lamp. UV radiation is not a chemical disinfection method; it avoids the addition of chemical reagents, and it disinfects without altering the physical or chemical properties of water. However, UV efficiency is affected by the presence of suspended solids in the CSOs, which scatter and absorb light and lower the method’s efficiency. Thus, UV disinfection is not very effective for CSOs that containing high TSS (5).

Peracetic Acid

Peracetic acid (CH_3COOOH, denoted PAA), also known as ethaneperoxoic acid, peroxyacetic acid, or acetyl hydroxide, is a very strong oxidant. Based on limited demonstration data for disinfecting secondary treatment plant effluents, peracetic acid appears to be an effective disinfectant and should be further evaluated for treating CSOs. The equilibrium mixture of hydrogen peroxide and acetic acid that produces PAA is too unstable and explosive to transport; therefore, PAA must be produced on site. The decomposition of PAA results in the formation of acetic acid, hydrogen peroxide, and oxygen (5).

Electron Beam Irradiation

Electron beam irradiation (e-beam) uses a stream of high energy electrons that is directed into a thin film of water or sludge. The electrons break water molecules apart and produce a large number of highly reactive chemical species (mainly radicals), including oxidizing hydroxyl radicals, reducing aqueous electrons, and hydrogen atoms. These are the main disadvantages of this method:

1. Increased safety considerations due to the use of high-voltage technology and the generation of X-ray radiation.
2. There is no full-scale application experience for CSOs.
3. High capital costs.
4. High O & M costs.
5. Thin process flow stream.
6. Sufficient pretreatment straining of influent is also required to remove most of suspended solids for efficient application of this system (5).

BIOLOGICAL TREATMENT AND COMBINED SYSTEMS

The use of biological treatment, combined with certain of the aforementioned physical–chemical treatment processes, for treating CSO presents certain serious limitations:

1. The biomass used to assimilate the nutrients in the CSOs must also be kept alive during dry weather, which can be rather difficult, except at an existing treatment plant.
2. Biological processes are subject to upset under to erratic loading conditions.
3. The land requirements for this type of treatment plant can be excessive near an urban area.
4. Operation and maintenance can be costly, and the facilities require highly skilled operators.

Some biological treatment technologies are used in CSO control as elementary parts of a WWTP. Pump-back or bleed-back flows from CSO storage facilities commonly receive secondary (biological) treatment in the treatment plant, once wet weather flows have subsided. In a WWTP, which has maximized the wet weather flows that are accepted, the hydraulic flows are sometimes split; only a portion of the primary treated flows is subjected to secondary treatment to avoid process upset. The split flows are blended again before the exit and disinfected appropriately for final discharge (1).

Constructed Wetlands

Constructed wetlands are artificial wastewater treatment systems, consisting of shallow (usually 1 m deep) beds, that have been planted with aquatic plants, and which rely upon natural microbial, biological, physical, and chemical processes to treat CSOs. They typically have impervious clay or synthetic liners and engineered structures to control the flow direction, liquid detention time, and water level. Depending on the specific type of system, they may contain inert porous media, such as rock, gravel, or sand. Constructed wetlands have been classified in the literature into two types.

Free water surface (FWS) wetlands (also known as surface flow wetlands) closely resemble natural wetlands in appearance; they contain aquatic plants, and water flows through the leaves and stems of the plants. Vegetated submerged bed (VSB) systems (also known as subsurface flow wetlands) do not resemble natural wetlands because they have no standing water. Wastewater (i.e., CSOs) stays beneath the surface of the medium, flows in contact with the roots and rhizomes of the plants, and is not visible or available to wildlife. Finally, the term vertical flow wetland is used to describe a typical vertical-flow sand or gravel filter, which has been planted with aquatic plants. Successful operation of this type of system depends mainly on its operation as a filter (i.e., frequently applying dosing and draining cycles) (7).

Figure 6 shows a typical cross section of a horizontal subsurface flow wetland (also known as a reed bed). These systems are used to treat the excess combined sewer flow that would otherwise be bypassed to receiving waters, with good treatment efficiency that meets the tight permitting conditions. During exceptionally dry weather, secondary or tertiary treated effluent can be diverted to this system to conserve the plantings and the microorganism's population in the system (6).

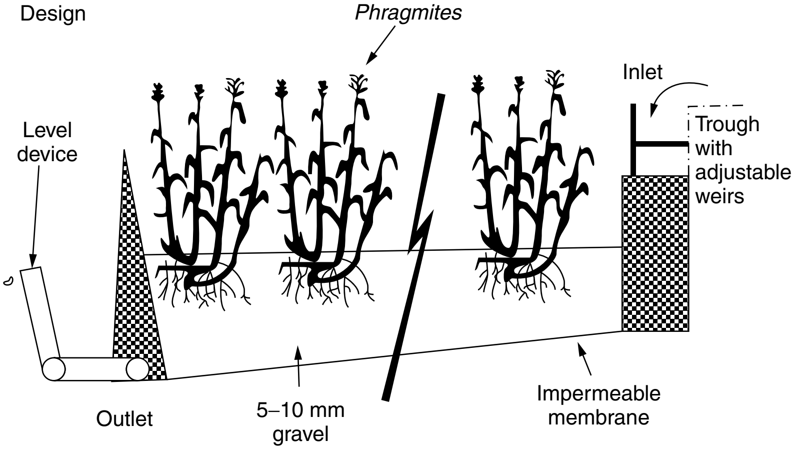

Figure 6. Diagrammatic longitudinal section of a horizontal subsurface flow wetland (6).

BIBLIOGRAPHY

1. U.S. Environmental Protection Agency. (1993). *Manual: Combined Sewer Overflow*. Report EPA/625/R-93/007. Cincinnati, OH.
2. Metcalf & Eddy, Inc. (1991). *Wastewater Engineering: Treatment, Disposal and Reuse*. McGraw-Hill, Singapore.
3. U.S. Environmental Protection Agency. (1999). *Combined Sewer Overflow Technology Fact Sheet Retention Basins*. Report EPA/832/F-99/042. Washington, DC.
4. Pisano, W.C. (1995). Comparative assessment: vortex separators, rotary sieves, and "Combing" screens for CSO floatable control. Presented at the *Water, Environ. Fed. Annu. Conf.*, Miami, FL.
5. U.S. Environmental Protection Agency. (1999). *Combined Sewer Overflow Technology Fact Sheet Alternative Disinfection Methods*. Report EPA/832/F-99/033, Washington, DC.
6. Green, M.B. and Upton, J. (1995). Constructed reed beds: Appropriate technology for small communities. *Water Sci. Technol.* **332**: 339–348.
7. U.S. Environmental Protection Agency. (2000). *Manual: Constructed Wetlands Treatment of Municipal Wastewaters*. Report EPA/625/R-99/010, Cincinnati, OH.
8. Prochaska, C.A. and Zouboulis, A.I. (2003). Performance of intermittently operated sand filters: A comparable study, treating wastewaters of different origins. *Water, Air, Soil Pollut.* **147**: 367–388.

BIOLOGICAL PHOSPHORUS REMOVAL IN THE ACTIVATED SLUDGE PROCESS

Michael H. Gerardi
Linden, Pennsylvania

Phosphorus often is the limiting nutrient that promotes the excessive growth of aquatic plants, especially algae, and excess quantities of phosphorus often are present in the effluent of activated sludge processes in quantities greater than those required for the growth of aquatic plants. Excess phosphorus promotes not only the undesired growth of aquatic plants but also the undesired impacts of the death of the aquatic plants upon the receiving water. Therefore, state and federal regulatory agencies limit the quantity of phosphorus in the effluent. The requirement limiting the quantity of phosphorus in the effluent of activated sludge processes is becoming more and more stringent for municipal wastewater treatment plants. For example, discharge limits for total phosphorus of 2 mg/L or lower have been applied broadly to many plants in the lower Susquehanna River Basin.

Several environmental concerns related to the excessive growth of aquatic plants include clogging of receiving waters and the production of color, odor, taste, and turbidity problems when the receiving waters are used for potable water supplies. The die-off of large numbers of aquatic plants results in oxygen depletion when the plants decompose. Those portions of the aquatic plants that do not decompose accumulate in the receiving waters and contribute to eutrophication or rapid aging of the receiving waters. Additionally, some algae release toxic compounds.

Municipal wastewater contains 10–20 mg/L of total phosphorus. The total phosphorus consists of inorganic and organic phosphorus. Inorganic phosphorus consists of phosphorus-containing compounds that do not contain carbon and hydrogen. Organic phosphorus-containing compounds do contain carbon and hydrogen. Significant and common inorganic phosphorus-containing compounds are orthophosphate (PO_4-P) and polyphosphates. Orthophosphate makes up approximately 50–70% of the total phosphorus in influent and approximately 90% of the phosphorus in the effluent of municipal wastewater treatment plants.

Orthophosphate is the preferred phosphorus nutrient for aquatic plants. The forms of orthophosphate found in the influent and effluent of wastewater treatment plants are pH dependent. At pH values <7 the $H_2PO_4{}^-$ form is dominant, whereas the $HPO_4{}^{2-}$ form is dominant at pH values >7. The forms of orthophosphate in the influent and effluent are produced through dissociation (Eq. 1).

$$H_2PO_4{}^- \rightleftharpoons HPO_4{}^{2-} + H^+ \tag{1}$$

Significant and common organic phosphorus-containing compounds include phytin, nucleic acids, and phospholipids. Phytin is an organic acid found in vegetables such as corn and soybean. Phytin is difficult to digest and is found in domestic wastewater. Nucleic acids are large complex molecules that contain genetic material. Phospholipids also are large and complex molecules that are used in the production of structural materials. Phytin, nucleic acids, and phospholipids degrade slowly in the activated sludge process. Their degradation results in the release of orthophosphate. It is the orthophosphate that is used as the phosphorus nutrient by bacteria and incorporated into cellular material or MLSS (mixed liquor suspended solids) in the activated sludge process.

The degradation of organic compounds in activated sludge processes and the incorporation of phosphorus into new cellular material are achieved by a large and diverse population of bacteria. However, there are some bacteria that are capable of removing and storing phosphorus in quantities larger then their cellular needs. Bacteria that are capable of removing phosphorus in excess quantities are known as "phosphorus-accumulating organisms" (PAO) or "poly-P bacteria" (Table 1). Of all poly-P bacteria, *Acinobacter* is the most commonly known and studied. Phosphorus is removed from the wastewater in the orthophosphate form and stored by the poly-P bacteria as polyphosphate granules. Removal of phosphorus by poly-P bacteria often is termed "luxury uptake of phosphorus."

There are several operational measures that can be used in activated sludge processes to remove phosphorus

Table 1. Genera of Wastewater Bacteria that Contain Poly-P Species

Acinobacter	*Escherichia*
Aerobacter	*Klebsiella*
Aeromonas	*Moraxella*
Arthrobacter	*Mycobacterium*
Beggiatoa	*Pasteurella*
Enterobacter	*Pseudomonas*

Table 2. Biological and Chemical Measures Available for Phosphorus Removal

Operational Measure	Description
Chemical precipitation	Use of alum, ferric chloride, ferrous sulfate, or lime at a high pH; results in increased operational costs for chemical addition and disposal of chemical sludge
Assimilation	Phosphorus incorporated into cellular material (MLSS) as 1–3% dry weight and wasted from the activated sludge process
Biological phosphorus removal	Phosphorus incorporated into cellular material (MLSS), especially poly-P bacteria, as 6–7% dry weight and wasted from the activated sludge process
Biological/chemical techniques	Precipitation of phosphorus released by poly-P bacteria under anaerobic condition and chemically precipitated with alum, ferric chloride, ferrous sulfate, or lime

from the effluent. These measures include biological and chemical techniques (Table 2). Phosphorus as orthophosphate can be removed from wastewater through its chemical precipitation as a metal salt at a high pH. Chemical compounds commonly used to precipitate orthophosphate include alum, ferric chloride, ferrous sulfate, and lime. Orthophosphate also can be removed through its incorporation as cellular material. By increasing the MLSS, more orthophosphate is incorporated into bacterial cells. However, the amount of phosphorus in the bacterial cells is only 1–3% by dry weight. Orthophosphate also can be removed by combined biological and chemical techniques and through biological phosphorus removal.

Of the biological and chemical measures available for phosphorus removal, biological phosphorus removal offers several advantages when compared to other measures. For example, biological phosphorus removal is relatively inexpensive due to the reduction in chemical costs and sludge disposal costs associated with chemical addition to precipitate orthophosphate. Biological phosphorus removal is also capable of removing phosphorus to low effluent concentrations.

Biological phosphorus removal or luxury uptake of phosphorus occurs when orthophosphate uptake by poly-P bacteria exceeds cellular requirements. If luxury uptake of phosphorus does not occur in an activated sludge process, the phosphorus content of the activated sludge is approximately 1–3% on a dry weight basis. If luxury uptake of phosphorus does occur, the phosphorus content of activated sludge is approximately 6–7%.

Bacteria that are capable of luxury uptake of phosphorus enter activated sludge processes in fecal waste and through inflow and infiltration as soil and water bacteria. These bacteria are unique and remove phosphorus in excess of cellular needs in the presence of rapidly degradable organic compounds when transferred from an anaerobic (fermentative) tank to an aerobic tank.

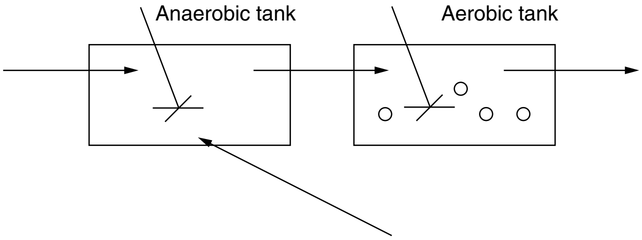

Figure 1. Microbial activity in the anaerobic tank. In the anaerobic tank, soluble cBOD is fermented in the absence of free molecular oxygen and nitrate ions. The fermentation process produces a variety of fatty acids. The acids are rapidly absorbed by the poly-P bacteria and stored as an insoluble starch (PHB). In order to absorb the fatty acids and store them as starch, energy in the form of orthophosphate is released by the poly-P bacteria to the bulk solution.

Table 3. Soluble Fatty Acids Produced in the Anaerobic Tank

Fatty Acid	Formula
Formic acid	HCOOH
Acetic acid	CH_3COOH
Propionic acid	CH_3CH_2COOH
Butyric acid	$CH_3CH_2CH_2COOH$
Valeric acid	$CH_3CH_2CH_2CH_2COOH$
Caproic acid	$CH_3CH_2CH_2CH_2CH_2COOH$

In the anaerobic tank (Fig. 1) that has a retention time of 1–2 h, fatty acids are produced in large quantities through fermentation (Table 3). Fermentation is the microbial degradation of organic compounds without the use of free molecular oxygen (O_2) or nitrate ions ($NO_3{}^-$). Poly-P bacteria quickly absorb the fatty acids produced through fermentation. Although the fatty acids are absorbed, the acids are not degraded. Instead, the soluble fatty acids are stored in the form of insoluble starch granules (poly-β-hydroxybutyrate or PHB).

The conversion of fatty acids to insoluble starch granules and the storage of the granules requires the expenditure of energy by the poly-P bacteria. The expenditure of energy results in a release of orthophosphate from the poly-P bacteria into the fermentative tank. With the production of PHB in the poly-P bacteria, the fermentative tank contains two "pools" of phosphorus—influent phosphorus and poly-P bacteria released orthophosphate.

In the aerobic tank (Fig. 2) that has a retention time of 1–2 h, PHB granules are solubilized and degraded with the use of free molecular oxygen. The degradation of PHB granules results in the release of a large quantity of energy that is captured and stored by the poly-P bacteria. The energy is stored in the bacteria in the form of insoluble phosphate granules or voluntin. Phosphorus is removed from the activated sludge process when MLSS (bacteria) is wasted.

There are several processes available for the biological removal of phosphorus. Often biological phosphorus removal is combined with nitrification and denitrification. Nitrification is the oxidation of ammonium ions ($NH_4{}^+$) to nitrate ions ($NO_3{}^-$), while denitrification is the reduction

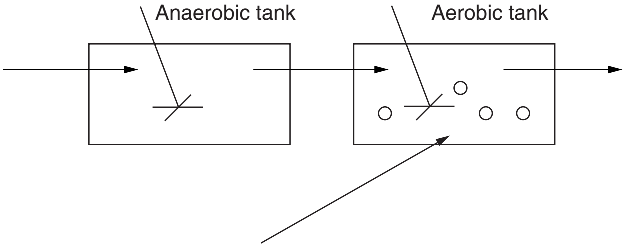

Figure 2. Microbial activity in the aerobic tank. In the aerobic tank, PHB is solubilized and degraded in the presence of free molecular oxygen. The degradation of PHB results in the production of carbon dioxide, water, and new bacterial cells. Phosphorus released in the anaerobic tank as well as phosphorus present in the waste stream are absorbed by poly-P bacteria and stored in phosphorus granules.

of nitrate ions to molecular nitrogen (N_2) and nitrous oxide (N_2O). Nitrification and denitrification are responsible for the biological removal of nitrogen from wastewater. When nitrification and denitrification are combined with biological phosphorus removal, these three biological processes are known as biological nutrient removal (BNR). There are five significant processes used for biological phosphorus removal or BNR. Several of these processes are proprietary. These processes are the A/O, Phostrip, A2O, Bardenpho, and UCT.

Biological nutrient removal processes are either mainstream or sidestream (Table 4). A mainstream process contains an anaerobic tank within the major wastewater flow from influent to effluent. A sidestream process contains an anaerobic tank outside the major wastewater flow.

Of the nutrient removal processes, two are designed to remove phosphorus only. These processes are the A/O and the Phostrip. The A/O (anaerobic/oxic) process is a mainstream process (Fig. 3). The A/O process is patented by Air Products and Chemicals, Incorporated and is similar to a conventional activated sludge process.

The Phostrip process is a sidestream process and includes biological and chemical phosphorus removal (Fig. 4). A stripper tank is included in the Phostrip process. This tank has an anaerobic condition where phosphorus is released by poly-P bacteria from the return activated sludge (RAS). The released phosphorus is removed from the stripped tank by elutriation water. Lime is added to the elutriation water as it leaves the stripper tank. The addition of lime results in the precipitation of phosphorus as calcium phosphate.

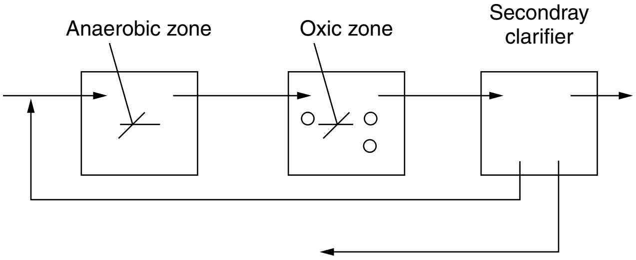

Figure 3. The A/O process.

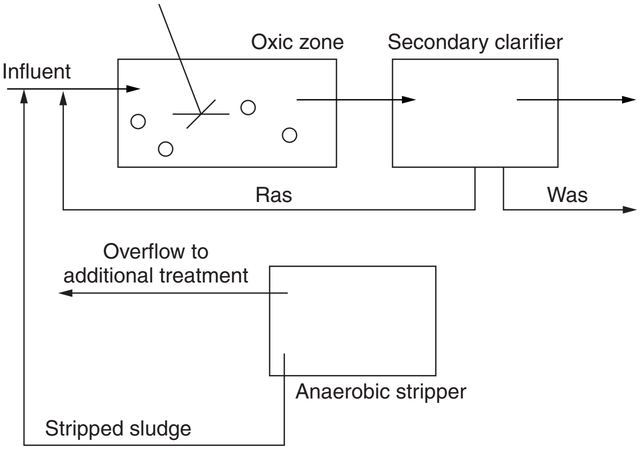

Figure 4. The Phostrip process.

The A2/O process consists of an anaerobic zone, an anoxic zone, and an oxic zone. As wastewater and bacteria move through these three zones, phosphorus is removed biologically and nitrogen is removed through nitrification and denitrification. Fermentation occurs in the anaerobic zone and phosphorus is released to the bulk solution by poly-P bacteria. In the anoxic zone nitrate ions are used (denitrified) by facultative anaerobic bacteria (denitrifying bacteria) to degrade soluble carbonaceous BOD. In the oxic zone ammonium ions (NH_4^+) in the wastewater and ammonium ions released from nitrogen-containing compounds are oxidized (nitrified) to nitrate ions (NO_3^-).

The Bardenpho process is licensed and marketed in the United States by Eimco Process Equipment Company. The Bardenpho process includes five zones (anaerobic zone, anoxic zone, oxic zone, anoxic zone, and oxic zone). The University of Capetown or UCT process also contains anaerobic, anoxic, and oxic zones. However, the UCT

Table 4. Nutrient Removal Processes

Process Name	Nutrient Removed		Process		Chemical Precipitation
	Nitrogen	Phosphorus	Mainstream	Sidestream	
A/O		X	X		
Phostrip		X		X	X
A2/O	X	X	X		
Bardenpho	X	X	X		
UCT	X	X	X		

process is designed to reduce the quantity of nitrate ions returned to the anaerobic zone in order to optimize the phosphorus removal.

As phosphorus and nitrogen discharge limits become more stringent for activated sludge processes in the United States, these nutrient removal systems will become more popular. The choice of the biological nutrient removal system is based on cost, wastewater composition, and nutrient removal requirements.

PHOTOCATALYTIC MEMBRANE REACTORS IN WATER PURIFICATION

Raffaele Molinari
Università della Calabria
Rende, Italy

L. Palmisano
Università di Palermo
Palermo, Italy

INTRODUCTION

General

Potable water, industrial water, and wastewater are often polluted by toxic organic species. Some classical methods (1,2) to cleanup waters, before sending them to rivers or to municipal drinking water supplies, transfer pollutants from one phase to another, hence creating further waste streams. New methods, such as those involving heterogeneous photocatalytic reactions, allow in many cases a complete degradation of organic pollutants to small and non-noxious species, without using chemicals, thus avoiding sludge production and its disposal.

Membrane separation processes, thanks to the selective property of membranes, have already proved to be competitive with other separation processes (3–7). Photocatalytic membrane reactors (PMRs) have some advantages compared to conventional photoreactors. Indeed, confining the photocatalyst to the reaction environment as a result of the presence of the membrane enables operation with high amounts of catalyst, control of the residence time of the molecules in the reactor, and realization of a continuous process with simultaneous product(s) separation from the reaction environment. Some PMR configurations using membranes ranging from microfiltration (MF) to reverse osmosis (RO) have been investigated (8–13).

The influence of various operating parameters on the photodegradation rate of pollutants present in aqueous effluents by means of discontinuous and continuous photocatalytic processes in the presence of NF membranes has been reported (14). Moreover, the possible use of solar radiation (15–18) in PMRs is of particular interest as the energy cost is one of the main drawbacks for industrial applications. Although many papers on photocatalysis have been published, the cases where membrane and photocatalyst are coupled are very few.

Historical Background

The use of photocatalysis in waste treatment is usually based on the electronic excitation of a polycrystalline semiconductor caused by light absorption that drastically alters its ability to lose or gain electrons, promoting decomposition of pollutants to harmless by-products. Photocatalytic processes in liquid phase have been applied to the degradation of several organic compounds (19–24). Only a few compounds such as chlorofluorocarbons, trifluoroacetic acid, and 2,4,6-trihydroxy-1,3,5-triazabenzene (cyanuric acid) cannot be completely degraded by photocatalytic methods (24). The formation of transient by-products, more toxic than the starting substrates, could be observed in some cases. The research on types of reactors that can be used in photocatalysis is very active, and slurry photocatalytic reactors suitable for the requirement of continuous operation are described in the literature (25,26). Nevertheless, as the chemical industry is characterized by an almost exclusive use of continuous processes, a photocatalytic powder for potential application purposes should possess a suitable size and mechanical characteristics in addition to good catalytic properties. To date, only rare examples of pilot plant photoreactors have been reported, as the difficulty of making a clear assessment of the costs of the photocatalytic processes (which are typically efficient only in dilute systems) have prevented a wider development of them from an application perspective. Some authors have immobilized the semiconductor on Pyrex glass sheets (27,28), on aerogels (29), or on particles having high surface area (e.g., alumina or silica gel) (30,31). In all these methods, a drawback is the mass transfer resistance of the reacting species, which could control the reaction rate. Despite the potential advantages of using hybrid membrane photoreactors, the research on coupling photocatalysis and membranes is not yet sufficiently developed. Some papers reported in the literature concern the use of cellulose microporous membranes (32); porphyrin containing membranes (33); porphyrins immobilized as photosensitizers on sulfocationic membranes (34); filtration coupled to catalysis (35,36); and TiO_2 immobilized inside the membrane (37), physically deposited on the membrane surface (38), or confined in suspension by means of the membrane (8,9,10,39).

FUNDAMENTALS OF HETEROGENEOUS PHOTOCATALYSIS

A simple definition of heterogeneous photocatalysis implies only the acceleration of a photoreaction by the action of a solid catalyst, which may interact with the species to be degraded and/or with the intermediates, depending on the reaction mechanism. The catalytic nature of the process should be demonstrated by checking that the turnover number (TON) is greater than 1. TON can be defined as the ratio of the number of photoinduced transformations for a given period of time to the number of photocatalytic sites. The total surface area of the solid photocatalyst can be considered when the number of sites is unknown, but figures of TON obtained in this way are lower limits. The most widely used semiconductor

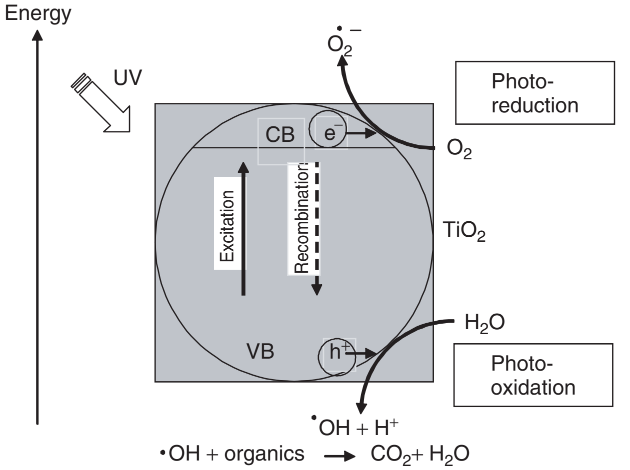

Figure 1. Basic photocatalytic mechanism occurring when a semiconductor particle (e.g., TiO_2) is irradiated.

photocatalyst is polycrystalline TiO_2, in the allotropic phases of both anatase and rutile (the anatase phase appears generally more photoactive), although the use of many other solids (ZnO, WO_3, CdS, MoS_2, CdSe, Fe_2O_3, etc.) has also been reported (40–42). When a semiconductor is illuminated with light of suitable energy (i.e., greater than its bandgap), electrons are promoted from the valence band (vb) to the conduction band (cb), acquiring the reducing power of the cb energy; positive holes are created in the vb, acquiring the oxidation power of the vb energy. This photoproduced pair can give rise to recombination with emission of thermal energy and/or luminescence or to redox (reduction–oxidation) reactions involving electron acceptor or donor species adsorbed on the surface of the catalyst particles (Fig. 1). Consequently, the recombination rate of the photoproduced electron–hole pairs should be as low as possible in order to favor their availability on the surface of the catalyst particle.

A photocatalyst should possess some essential characteristics: (1) light absorption should occur in the near UV and possibly in the visible wavelength ranges; (2) the stability should be such that its re-utilization is possible; and (3) some thermodynamic and kinetic constraints should be fulfilled (43). The photoreactivity depends not only on the intrinsic electronic characteristics of the photocatalysts, but also on their structural, textural, and surface physicochemical features. Among them we can cite the vb and cb energies, the bandgap value, the lifetime of photogenerated electron–hole pairs, the crystallinity, the allotropic phase, the particle sizes, the presence of defects and dopants, the specific surface area, the porosity, the surface hydroxylation, and the surface acidity and basicity. The relative weight of their importance for the studied reaction determines the final level of the photoreactivity. For this reason, it is often difficult to explain and to discuss exhaustively the observed photoreactivity trends by considering only a few properties of the photocatalysts. The steps shown in Table 1 occur when polycrystalline TiO_2 is used in aqueous medium in the presence of O_2 and a generic substrate.

The oxidant radicals along with the holes can give rise to oxidant attacks on a wide variety of substrates. The presence of O_2 is essential in order to trap photoproduced electrons, improving the charge separation and consequently the availability of the holes.

The application of heterogeneous photocatalysis to the purification of aqueous effluents containing dissolved organic and/or inorganic species (e.g., CN^-) has been widely studied. Nevertheless, it is worth noting that analytical studies to check if noxious intermediates are produced under irradiation are essential when the scale-up of photoreactors is proposed for application purposes.

FUNDAMENTALS OF MEMBRANE PROCESSES

Membrane processes are separation methods at the molecular level that have received more interest in recent years for their possibility of being used in many industrial applications. The main goals of such processes are concentration of a solute by removing the solvent, purification of a solution by removing nondesirable components, and fractionation of liquid or gaseous mixtures. Membrane separation processes offer interesting opportunities in pollution control, in the production of drinking water (44), and in the treatment of industrial wastewater (45). A membrane can be defined as a selective barrier between two phases (46,47); it can be thin or thick, natural or synthetic, neutral or charged, the structure can be homogeneous or heterogeneous; and the mass transport can be active or passive. In the last case the driving force can be due to a difference of pressure, concentration, or temperature. The driving force has the capacity to transport a component more rapidly

Table 1. Some Essential Steps When Polycrystalline TiO_2 Is Used in Photodegradation Processes of Noxious Organic and Inorganic Substrates

Role of TiO_2	*Role of* O_2
$TiO_2 + h\nu \rightarrow TiO_2\ (e^-_{(cb)} + h^+_{(vb)})$	$O_2 + e^-_{(cb)} \rightarrow {}^{\bullet}O_2{}^-$
$OH^- + h^+_{(vb)} \rightarrow {}^{\bullet}OH$	${}^{\bullet}O_2{}^- + H^+ \rightarrow {}^{\bullet}HO_2$
	$2{}^{\bullet}HO_2 \rightarrow O_2 + H_2O_2$
	$H_2O_2 + {}^{\bullet}O_2{}^- \rightarrow OH^- + {}^{\bullet}OH + O_2$

Substrate Degradation

Substrate $+ {}^{\bullet}HO_2 \rightarrow$ products

Substrate $+ {}^{\bullet}OH \rightarrow$ products

Substrate $+ h^+_{(vb)} \rightarrow$ products than others owing to different physical and/or chemical properties between the membrane and the components. The membrane is assembled in a module whose geometry is generally plane or cylindrical and in which the feed is separated into two streams called retentate (the treated feed) and permeate.

Each membrane process is characterized by the employement of a particular type of membrane. Some processes are microfiltration, ultrafiltration, nanofiltration, and reverse osmosis (48) in decreasing order of particles size and increasing pressures (0.1–0.2 to 50–100 bar) used as driving forces. Other membrane processes concern the separation of ionic species by applying an electrical potential (electrodialysis), the separation of mixtures of volatile liquids (pervaporation), and the separation of water from nonvolatile solutes (membrane distillation) by means of a temperature difference (49).

SOME CASE STUDIES OF PHOTOCATALYTIC MEMBRANE REACTORS USED TO PHOTODEGRADE POLLUTANTS

Preliminary Remarks

The pioneering studies for coupling photocatalysis and membrane separations focused on the optimization of the system configuration (14). A continuous operation system (feed feeding and permeate withdrawing) was reported by Molinari et al. (39) involving the use of TiO_2 particles in suspension (Fig. 2). This configuration seems appropriate for industrial applications, so experimental results deriving from this system will be mostly reported in the following. Membrane rejection of the target species with and without photodegradation was measured to obtain information on the performance of the system. Degussa P25 TiO_2 (specific surface area ≅ 50 m^2/g; crystallographic phase, ~80% anatase and 20% rutile) was used as the photocatalyst. Some of the molecules chosen as model pollutants were 4-nitrophenol (4-NP) and two dyes—patent blue ($C_{27}H_{31}N_2NaO_6S_2$) and congo red ($C_{32}H_{22}N_6Na_2O_6S_2$). The extent of the degradation of the contaminant was determined by UV-visible measurements and total carbon (TC), total organic carbon (TOC), and total inorganic carbon (TIC) determinations.

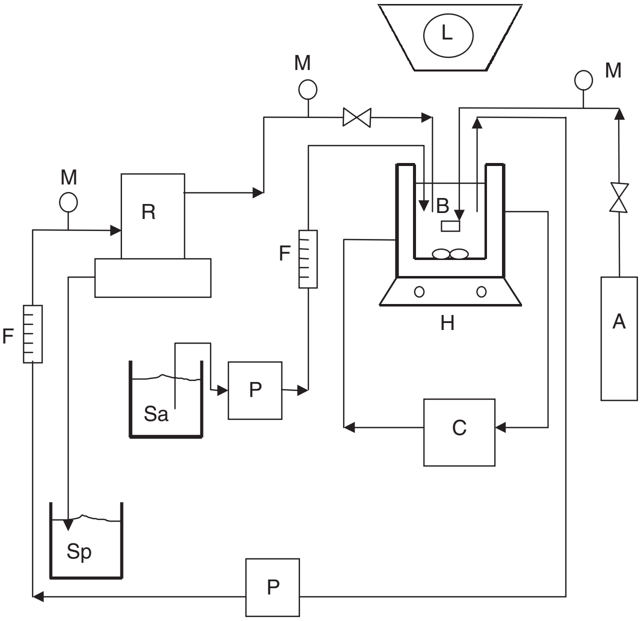

Figure 2. Scheme of a continuous membrane photoreactor system with suspended catalyst. A, oxygen cylinder; B, recirculation reservoir (reactor); C, thermostatting water; L, UV lamp; M, pressure gauge; F, flowmeter; R, membrane cell; H, magnetic stirrer; P, peristaltic pump; Sa, feed reservoir; Sp, permeate reservoir (39).

4-Nitrophenol Degradation in Batch and Continuous Membrane Photoreactors

When NF-PES-010 or N30F nanofiltration membranes were tested (the first one was the most permeable), the permeate fluxes through the membranes in the presence of suspended catalyst were a little lower than those found in the absence of catalyst. Deposition of the catalyst on the membrane was mimimized by using a cell geometry that guaranteed turbulence and presence of vortexes in the bulk of the solution above the membrane.

Photodegradation of 4-NP in batch (recycle) and continuous system configurations showed, for both types of membranes, bell-shaped curves of permeate concentration as a function of time. The concentration of 4-NP in the permeate was attributed to three factors: rejection, photocatalytic degradation, and adsorption. The increase of the initial concentration of permeate in the bell-shaped curves was lowered by the photocatalytic degradation that was responsible for the decrease of the concentration both in the retentate and in the permeate. Although 4-NP concentration in the retentate for the continuous configuration decreased less quickly than for the discontinuous one, the continuous system seems more promising for industrial application (39). In this system the optimum choice of the ratio between the irradiated volume and the total volume, V_i/V_t, was important. When the total suspension volume was increased from 400 to 700 mL, for instance, V_i/V_t increased owing to a constant recycle volume and, consequently, 4-NP abatement was higher due to an increased percentage of irradiated with respect to recycled suspension. The UV radiation mode was also important. The immersed lamp was found to be three times more efficient than the external lamp. Indeed, 99% w/w 4-nitrophenol degradation was achieved after about 1 h in the first case, whereas about 3 h were needed in the second case (39).

Photodegradation of Other Pollutants

NF-PES-010 and NTR-7410 nanofiltration membranes were tested in degradation runs after determination of their permeability and rejection for a variety of pollutants (50). It was found that membranes hold both catalyst and pollutants, but the NTR-7410 membrane tested at 8 bar gave a higher water permeate flux (105 $L/h \cdot m^2$) than the NF-PES-010 one (30 $L/h \cdot m^2$). The NTR-7410 membrane was also able to retain small molecules carrying negative charges (like the membrane) as a result of Donnan exclusion and, at 6 bar, yielded fluxes ranging between 20 and 40 $L/h \cdot m^2$. The results of rejection tests and photodegradation studies of humic acids, patent blue dye, and 4-NP in a system utilizing this membrane are described in the following.

The NTR-7410 membrane, not surprisingly, provided 100% retention of humic acids as they are made up of oligomers with molecular sizes greater than the membrane cutoff (600–800 g/mol). For patent blue (molar weight = 567 g/mol and initial concentration 10 mg/L), membrane rejection was about 78.6%. With respect to 4-NP, 0% rejection occurred at pH 6.75, while at pH 11 a rejection of about 77% and a negligible adsorption were observed. The pH dependence of rejection of 4-NP is to be expected by taking into account its acid–base properties and the resultant electrostatic interaction between 4-NP and the membrane (51,52):

$$C_6H_4NO_2OH + H_2O \rightleftharpoons C_6H_4NO_2O^- + H_3O^+$$

$$R\text{—}SO_3{}^-H^+ + H_2O \rightleftharpoons RSO_3^- + H_3O^+$$

Humic acids are found in many natural waters (53) and often induce fouling problems when membranes are used. The interaction between humic acid and the particulate TiO_2 photocatalyst has been explored by Lee et al. (54) who found that (1) humic solution or TiO_2 suspension alone produced essentially a constant membrane flux of 160–170 $L/h \cdot m^2$ and (2) the mixture gave initial flux of 145 $L/h \cdot m^2$ that decreased to 101 $L/h \cdot m^2$ (a decrease of approximately 30%) in the first 30 min of operation. When the photoreactor was illuminated, a constant flux was measured during 6 h, consistent with destruction of humic acid. These authors proposed that the photocatalyzed conversion of humic acid into smaller and/or less absorptive species was occurring and concluded that "photocatalytic reactions appear to be attractive for the control of fouling materials such as natural organic matter."

Continuous Membrane Photoreactor at High Pollutant Concentrations

The detrimental effect caused by a high pollutant concentration could be minimized by taking advantage of the ability of the membrane to retain both the catalyst and the pollutant. In order to investigate this, the photocatalytic system was tested in a continuous process with an initial concentration of pollutant in the photoreactor equal to zero, that is, in the presence of distilled water. A concentrated solution was continuously fed with a flow rate equal to that of the removed permeate. This solution was immediately diluted in the reactor, so photodegradation was effective at low pollutant concentration, although it was continuously fed at high concentration. Control of the residence time of the pollutant in the reactor was possible with the result that very low concentrations in the permeate could be obtained.

This approach was tested by performing humic acid photodegradation studies (50) with an initial concentration in the photoreactor equal to zero and feeding a 200-mg/L solution. It was possible to maintain steady-state pollutant concentrations lower than 5 mg/L and 2 mg/L in the retentate and in the permeate, respectively. It was observed that the humic acid rejection was not 100% and this was mainly due to the lower size of the humic acid oligomers produced during the photodegradation process.

The continuous process was also tested for high concentration (500 mg/L) feeds of patent blue dye and 4-NP. The photodegradation rate of patent blue was found to be lower than that obtained for other pollutants, possibly because of adsorption of the acid dye on the amphoteric catalytic surface, preventing UV light absorption. Indeed, at the end of the run, the catalyst was a dark blue color.

Degradation of Dyes in the Continuous Membrane Photoreactor

In order to achieve a better control of the residence time of pollutants, such as the dyes during the photodegradation process, a hybrid photoreactor was used in which the nanofiltration membrane was able to confine selectively dyes (congo red and patent blue) and catalyst in the reaction ambient while the permeate was withdrawn (55). The experimental results of runs carried out with patent blue in the membrane photoreactor with suspended catalyst showed that the photodegradation reaction followed pseudo-first-order kinetics (observed rate constant equal to $3.76 \times 10^{-3}\ min^{-1}$). A similar run carried out in the absence of membrane showed an observed rate constant of $1.02 \times 10^{-2}\ min^{-1}$. The lower reaction rates for both dyes obtained by using the membrane with respect to that obtained in its absence were due to the smaller volume of irradiated suspension (320 against 500 mL) because a part (180 mL) circulated in the pipes of the plant and in the membrane cell.

The possibility of successfully treating highly concentrated solutions of both dyes was examined, allowing the setup shown in Fig. 2 to work as a continuous system. In particular, the transient condition in the membrane photoreactor was studied by separating the effects of accumulation, adsorption, and photodegradation. Results of three runs with congo red for which the initial concentration of pollutant inside the photoreactor was zero are reported in Fig. 3.

The first run, carried out in the absence of UV light and photocatalyst, indicated that the initial rate of dye accumulation in the photoreactor was 0.151 mg/min. The second run, carried out in the absence of UV light but in the presence of TiO_2, indicated that no increase of dye concentration occurred in the photoreactor during the first 45 min of continuous working of the plant because the congo red feed was adsorbed onto the catalyst surface. After the active sites of the catalyst were saturated, the concentration of the dye in the retentate increased linearly with an accumulation rate of 0.136 mg/min, very close to that determined for the first run. In the third run the continuous degradation of congo red in the presence of UV light and catalyst was performed and concentrations in the retentate and in the permeate versus irradiation time are reported. It can be noticed that, due to the concurrent

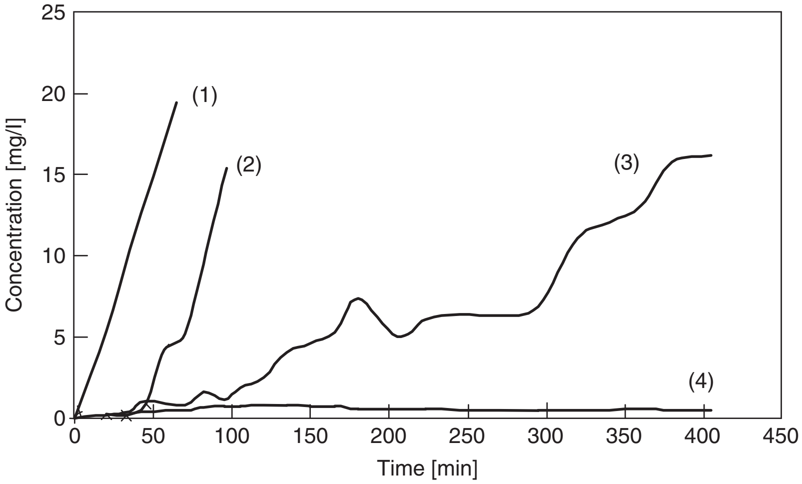

Figure 3. Concentration of congo red in the retentate and in the permeate versus time for three different continuous runs. (1) Absence of UV light and TiO_2; (2) absence of UV light and presence of TiO_2; (3) presence of UV light and TiO_2 for retentate; and (4) presence of UV light and TiO_2 for permeate. $V = 500$ mL; $T = 303$ K; $C_0 = 0$ mg/L; $C(O_2) = 22$ ppm; TiO_2 amount = 1 g/L; $C_{\text{feed}} = 500$ mg/L; initial permeate flux, $J_{p_{\text{in}}} = 74.2$ L/(m$^2 \cdot$ h); final permeate flux, $J_{p_{\text{fin}}} = 29.8$ L/(m$^2 \cdot$ h); lamp, 125-W medium pressure Hg immersed lamp; initial pH = 6.42; membrane, NTR-7410; $\Delta P = 3.5$ bar.

effect of dilution, adsorption, and photodegradation, the accumulation rate in the retentate was lower than that observed during the other two runs. The concentration of congo red in the permeate was virtually zero because the membrane maintained the substrate in the reacting ambient. The permeate flux (J_p) throughout all the runs decreased from the initial value of 74.2 L/m$^2 \cdot$ h to the value of 29.8 L/m$^2 \cdot$ h and, consequently, the dye feeding rate also decreased. It is worth noting that the average photodegradation rate (0.274 mg/min) calculated for the overall run was lower than the average feeding rate (0.416 mg/min), calculated for the first 180 min, while it was 0.049 mg/min higher for longer time (average feeding rate 0.225 mg/min).

Degradation of patent blue was also tested in the continuous system under the same experimental conditions used for congo red and the results are reported in Fig. 4. During the transient state (250–300 min in this specific case) the dye accumulated in the photoreactor because the photodegradation rate (0.570 mg/min) was lower than the feeding rate (0.863 mg/min). Subsequently, steady-state conditions were achieved, owing to the lower permeate flowrate, and no difference was observed between photodegradation and feeding rates. The rejection of NTR-7410 membrane at steady-state conditions with respect to patent blue was 44.6%.

Use of the membrane was beneficial because in addition to its role as a barrier for the catalyst, the product [i.e., cleaned up water (the permeate)] contained a very low concentration of dye with respect to the feed. It was approximately 1% in the case of congo red and approximately 11% in the case of patent blue with respect to 500 mg/L of the feed. It is worth noting that the concentration of the product corresponded to that of the retentate (approximately 3% for congo red and approximately 22% for patent blue) if the membrane was not used.

CONCLUSIONS

The continuous membrane photoreactor that combines both the advantages of classical photoreactors (catalyst in suspension) and membrane processes (separation at molecular level) appears very promising. Photocatalytic degradation can be carried out in reasonable times due to the high irradiated surface area of the suspended particles in the batch.

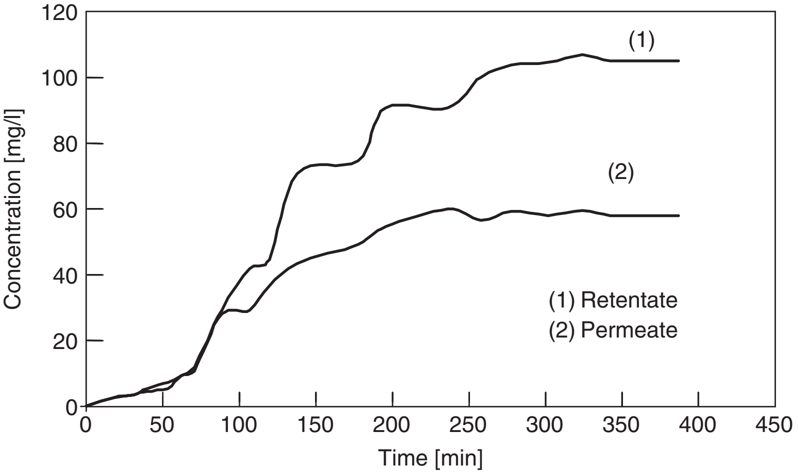

Figure 4. Concentration of patent blue in the retentate and in the permeate versus time for a continuous photodegradation run carried out in the setup showed in Fig. 2. $V = 500$ mL; $T = 303$ K; $C_0 = 0$ mg/L; $C(O_2) = 22$ ppm; TiO_2 amount = 1 g/L; $C_{\text{feed}} = 500$ mg/L; regime permeate flux, $J_{p_{\text{regime}}} = 78.4$ L/(m$^2 \cdot$ h); lamp, 125-W medium pressure Hg immersed lamp; initial pH = 5.61; membrane, NTR-7410; $\Delta P = 7.0$ bar.

Properly selected membranes should have both the capability to retain the catalyst and to partially reject organic species, enabling control of the residence time in the reacting system.

In studies of 4-nitrophenol degradation, three factors—rejection, photocatalytic degradation, and adsorption—were found to contribute to maintain the steady-state 4-NP concentration in the permeate at very low values. The adsorption phenomenon is particularly important when oscillating concentrations of pollutant are fed to the membrane photoreactor, resulting in a negligible variation of concentration in the permeate.

Photoreactors with an immersed lamp are generally more efficient than systems with external lamp. In addition, the pH of the polluted water, the molecular weight of the pollutant, and the type of pollutant and membrane are variables influencing pollutant rejection as a result of charge repulsion (Donnan exclusion) effects.

In order to select a suitable membrane, rejection should be determined during operation of the photoreactor. The pressure in the membrane cell, the pH of the polluted water, the molecular size of the pollutants, and the photogenerated by-products and intermediate species can influence the permeate flux of the membrane and consequently its choice.

High initial concentrations of the pollutants (e.g., 4-NP, patent blue, congo red) lower the photodegradation rate; however, this problem can be solved by diluting the feed in the reactor and by controlling the residence time of the pollutant by means of the membrane.

The experimental results available in the literature indicate that the choice of a suitable membrane is essential for applying the photocatalytic membrane processes to the treatment of real effluents. Furthermore, use of photocatalysis combined with RO processes for drinking water production can eliminate the membrane fouling problem and the need for plant sanitizing.

The hybrid continuous photoreactor where a nanofiltration membrane is used can give advantages over other approaches: simplification of clean-up or purification of various types of waters (for industrial, municipal/domestic, and agricultural uses), no sludge production, and saving in chemicals usage. It is expected that these hybrid processes will be considered particularly when plant upgrade is planned and, especially, if sunlight energy can be used for irradiation.

Acknowledgment
The authors thank MIUR (Italy) for financial support for this work.

BIBLIOGRAPHY

1. Gorenflo, A., Hesse, S., and Frimmel, H. (1998). A concept for advanced biodegradation. In: *Proceedings of the European Research Conference on Natural Waters and Water Technology: Catalytic Science and Technology for Water*. Acquafredda di Maratea, (I), October 3–8, p. 32.
2. Martin, R.J., Iwugo, G., and Pawlowski, L. (Eds.). (1989). *Physicochemical Methods for Water and Wastewater Treatment*. Elsevier, New York, p. 265.
3. Drioli, E. (Ed.). (1994). *Proceedings of the International Symposium on Membrane and Membrane Processes*. Hangzou, China, April 5–10.
4. Howell, A.J. and Noworyta, A. (Eds.). (1995). *Towards Hybrid Membrane and Biotechnology Solutions for Polish Environmental Problems*. Wroclaw Technical University Press, Wroclaw.
5. Caetano, A., De Pinho, M.N., Drioli, E., and Muntau, H. (Eds.). (1995). *Membrane Technology: Applications to Industrial Wastewater Treatment*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
6. Pramauro, E. et al. (1993). Preconcentration of aniline derivatives from aqueous solutions using micellar-enhanced ultrafiltration. *Analyst* **118**: 23.
7. Pramauro, E. and Bianco Prevot, E. (1998). Application of micellar ultrafiltration in environmental chemistry. *Chim. Ind. Milan* **80**: 733.
8. Sopajaree, K., Qasim, S.A., Basak, S., and Rajeshwar, K. (1999). An integrated flow reactor–membrane filtration system for heterogeneous photocatalysis. Part I: Experimental and modelling of a batch-recirculated photoreactor. *J. Appl. Electrochem.* **29**: 533.
9. Sopajaree, K., Qasim, S.A., Basak, S., and Rajeshwar, K. (1999). An integrated flow reactor–membrane filtration system for heterogeneous photocatalysis. Part II: Experiments on the ultrafiltration unit and combined operation. *J. Appl. Electrochem.* **29**: 1111.
10. Molinari, R., Palmisano, L., Drioli, E., and Schiavello, M. (2002). Studies on various reactor configurations for coupling photocatalysis and membrane processes in water purification. *J. Membr. Sci.* **206**: 399.
11. Ollis, D.F. (2002). Integrating photocatalysis and membrane technologies for water treatment. In: *Proceedings of 2nd European Meeting on Solar-Chemistry and Photocatalysis: Environmental Applications*. Saint Avold, France, May 29–31, Plenary Lecture 1.
12. Karakulski, K., Morawski, W.A., and Grzechulska, J. (1998). Purification of bilge water by hybrid ultrafiltration and photocatalytic processes. *Sep. Purif. Technol.* **14**: 163.
13. Xi, W. and Geissen, S-U. (2001). Separation of titanium dioxide from photocatalytically treated water by crossflow microfiltration. *Water Res.* **35**: 1256.
14. Molinari, R., Giorno, L., Drioli, E., Palmisano, L., and Schiavello, M. (2004). Photocatalytic nanofiltration reactors, Chap. 18. In: *Nanofiltration: Principles and Applications*. A.I. Schaefer, D. Waite, and A. Fane (Eds.). Elsevier Science, London.
15. Wyness, P., Klausner, J.F., Goswami, D.Y., and Schanze, K.S. (1994). Performance of nonconcentrating solar photocatalytic oxidation reactors, part I: flat-plate configuration. *J. Solar Energy Eng.* **116**: 3.
16. Zhang, Y., Crittenden, J.C., Hand, D.W., and Perram, D.L. (1994). Fixed-bed photocatalysts for solar decontamination of water. *Environ. Sci. Technol.* **28**: 435.
17. Liu, G. and Zhao, J. (2000). Photocatalytic degradation of dye sulforhodamine B: a comparative study of photocatalysis with photosensitization. *New J. Chem.* **24**: 411.
18. Augugliaro, V. et al. (2004). Degradation of lincomycin in aqueous medium: coupling of solar photocatalysis and membrane separation. In: *Serie Ponencias, The Improving Human Potential Programme*. Access to research infrastructures activity. Research results at Plataforma Solar de Almeria within the year 2003 access campaign. CIEMAT, Madrid, pp. 43–52.

Next Page

19. Barbeni, M. et al. (1984). Photodegradation of 4-nitrophenol catalyzed by TiO_2 particles. *Nouv. J. Chim.* **8**: 547.
20. Okamoto, K. et al. (1985). Heterogeneous photocatalytic decomposition of phenol over TiO_2 powder. *Bull. Chem. Soc. Jpn.* **58**: 2015.
21. Augugliaro, V., Palmisano, L., Sclafani, A., Minero, C., and Pelizzetti, E. (1988). Photocatalytic degradation of phenol in aqueous titanium dioxide dispersion. *Toxicol. Environ. Chem.* **16**: 89.
22. Augugliaro, V. et al. (1991). Photocatalytic degradation of nitrophenols in aqueous titanium dioxide dispersion. *Appl. Catal.* **69**: 323.
23. Pramauro, E., Vincenti, M., Augugliaro, V., and Palmisano, L. (1993). Photocatalytic degradation of monuron in TiO_2 aqueous dispersions. *Environ. Sci. Technol.* **27**: 1790.
24. Pichat, P. and Enriquez, R.E. (2001). Interactions of humic acids, quinoline, and TiO_2 in water in relation to quinoline photocatalytic removal. *Langmuir* **17**: 6132.
25. Brandi, R.J., Alfano, O.M., and Cassano, A.E. (2000). Evaluation of radiation absorption in slurry photocatalytic reactors. 1. Assessment of methods in use and new proposal. *Environ. Sci. Technol.* **34**: 2623.
26. Brandi, R.J., Alfano, O.M., and Cassano, A.E. (2000). Evaluation of radiation absorption in slurry photocatalytic reactors. 2. Experimental verification of the proposed method. *Environ. Sci. Technol.* **34**: 2631.
27. Aguado, M.A., Anderson, M.A., and Hill, C.J., Jr. (1994). Influence of light intensity and membrane properties on the photocatalytic degradation of formic acid over TiO_2 ceramic membranes. *J. Mol. Catal.* **89**: 165.
28. Cheng, S., Tsa, S., and Lee, Y. (1995). Photocatalytic decomposition of phenol over titanium oxide of various structures. *Catal. Today* **26**: 87.
29. Walker, S.A., Christensen, P.A., Shaw, K.E., and Walker, G.M. (1995). Photoelectrochemical oxidation of aqueous phenol using titanium dioxide aerogel. *J. Electroanal. Chem.* **393**: 137.
30. Loddo, V., Marcì, G., Palmisano, L., and Sclafani, A. (1998). Preparation and characterisation of Al_2O_3 supported TiO_2 catalysts employed for 4-nitrophenol photodegradation in aqueous medium. *Mater. Chem. Phys.* **53**: 217.
31. Lepore, G.P., Persaud, L., and Langford, C.H. (1996). Supporting titanium dioxide photocatalysts on silica gel and hydro-physically unmodified silica gel. *J. Photochem. Photobiol. A Chem.* **98**: 103.
32. Bellobono, I.R. (1995). Photosynthetic membranes in industrial waste minimization and recovery of valuable products. In: *Membrane Technology Applications to Industrial Wastewater Treatment*. A. Caetano et al. (Eds.). Kluwer Academic Publishers, Dordrecht, The Netherlands, p. 17.
33. Solovieva, A.B. et al. (2000). Porphyrin containing membrane photocatalytic systems for steroid olefine oxidation. *Euromembrane*: 345.
34. Kotova, S.L., Rumjantseva, T.N., Solovieva, A.B., Zav'jalov, S.A., and Glagolev, N.N. (2000). Porphyrins immobilized on sulfocationic membranes as photosensitizers in singlet oxygen generation. *Euromembrane*, 355.
35. De Smet, K., Vankelecom, I.F.J., and Jacobs, P.A. (2000). Filtration coupled to catalysis: a way to perform homogeneous reactions in a continuous mode. *Euromembrane*: 356.
36. Puhlfürß, P., Voigt, A., Weber, R., and Morbé, M. (2000). Microporous TiO_2 membranes with a cut off <500 Da. *J. Membr. Sci.* **174**: 123.
37. Artale, M.A. et al. (2001). Preparation and characterization of membranes with entrapped TiO_2 and preliminary photocatalytic tests. *Ann. Chim. (Rome)* **91**: 127.
38. Molinari, R. et al. (2000). Study on a photocatalytic membrane reactor for water purification. *Catal. Today* **55**: 71.
39. Molinari, R., Grande, C., Drioli, E., Palmisano, L., and Schiavello, M. (2001). Photocatalytic membrane reactors for degradation of organic pollutants in water. *Catal. Today* **67**: 273.
40. Schiavello, M. (Ed.). (1985). *Photoelectrochemistry, Photocatalysis, Photoreactors. Fundamentals and Developments*. D. Reidel Publishing, Dordrecht, The Netherlands.
41. Pelizzetti, E. and Serpone, N. (Eds.). (1989). *Photocatalysis. Fundamentals and Applications*. John Wiley & sons, Hoboken, NJ.
42. Di Paola, A., Marcì, G., Palmisano, L., and Schiavello, M. (2000). Trattamenti innovativi delle acque mediante l'utilizzo di radiazione UV e catalizzatori: la fotocatalisi eterogenea. In: *Processi e Metodologie per il Trattamento delle Acque, a cura di L. Palmisano*. Edizioni Spiegel, Milano, Chap. 13, p. 237.
43. Palmisano, L. and Sclafani, A. (1997). Thermodynamics and kinetics for heterogeneous photocatalytic processes. In: *Heterogeneous Photocatalysis*, Vol. 3, *Wiley Series in Photoscience and Photoengineering*. M. Schiavello (Ed.). John Wiley, Chichester, Chap. 4, p. 110, and references therein.
44. Aptel, P. and Vial, D. (1992). In: *Membranes in Water Treatment and Potabilization*. C. Haber and E. Drioli (Eds.). Proc. Membrane Separation Processes, Rio de Janeiro, Brazil, May 3–8, p. 291.
45. Scott, K. (1995). *Handbook of Industrial Membranes*. Elsevier Advanced Technology, Oxford, UK.
46. Mulder, M. (1991). *Basic Principles of Membrane Technology*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
47. Hwang, S.T. and Kammermeyer, K. (1975). *Membranes in Separations*. Wiley Interscience, Hoboken, NJ.
48. Toyomoto, K. and Higuchi, A. (1992). Microfiltration and ultrafiltration. In: *Membrane Science and Technology*. Y. Osada and T. Nakagawa (Eds.). Marcel Dekker, New York, Chap. 8, p. 289.
49. Baker, R.W. (2004). *Membrane Technology and Applications*, 2nd Edn. Wiley, London.
50. Molinari, R., Borgese, M., Drioli, E., Palmisano, L., and Schiavello, M. (2002). Hybrid processes coupling photocatalysis and membranes for degradation of organic pollutants in water. *Catal. Today* **75**: 77.
51. van Der Bruggen, B., Schaep, J., Wilms, D., and Vandecasteele, C. (1999). Influence of molecular size, polarity and charge on the retention of organic molecules by nanofiltration. *J. Membr. Sci.* **156**: 29.
52. Meares, P. (1986). *Synthetic Membranes: Science, Engineering and Applications*. D. Reidel Publishing, Dordrecht, The Netherlands.
53. Schäfer, A.I. (2001). *Natural Organic Matter Removal Using Membranes: Principles, Performance and Cost*. CRC Press, Boca Raton, FL.
54. Lee, S.A. et al. (2001). Use of ultrafiltration membranes for the separation of TiO_2 photocatalysts in drinking water treatment. *Ind. Eng. Chem. Res.* **40**: 1712.
55. Molinari, R., Pirillo, F., Falco, M., Loddo, V., and Palmisano, L. (2004). Photocatalytic degradation of dyes by using a membrane reactor. *Chem. Eng. Proc.* **43**: 1103.

Previous Page

EPA'S NATIONAL PRETREATMENT PROGRAM, 1973–2003: THIRTY YEARS OF PROTECTING THE ENVIRONMENT

Office of Water—United States
Environmental
Protection Agency

EPA's National Pretreatment Program has led the way to dramatically reduce or eliminate discharges of pollutants to sanitary sewer systems and to the nation's water bodies. The Program controls a complex array of industrial wastestreams in order to prevent interference or pass—through of municipal treatment system processes. Without these controls, a number of harmful pollutants could make their way into the nations' waters. Federal, state, and local partnerships are central to the successful implementation of the Program. Renewed commitment and support to the Pretreatment Program will conserve the environmental gains of the last 30 years, strengthen strategic partnerships, and prepare communities to meet the pollution challenges of the 21st Century.

THE PRETREATMENT PROGRAM

Pollutants in industrial wastewater may compromise municipal treatment plants' processes or contaminate the nation's waters. To protect municipal treatment plants and the environment, the Pretreatment Program requires industrial dischargers to use treatment techniques and management practices to reduce or eliminate the discharge of harmful pollutants to sanitary sewers. The Pretreatment Program is a core part of the Clean Water Act's National Pollutant Discharge Elimination System (NPDES), and it has helped communities:

- Maintain and restore watershed quality, at a much lower cost than upgrading treatment;
- Encourage pollution prevention;
- Increase beneficial uses of sewage sludge;
- Prevent formation of poisonous gases in the sanitary sewer system;
- Meet wastewater discharge standards; and
- Institute emergency—prevention measures.

Goleta, California's pretreatment program makes it possible for marine life—fish, sea stars, and coral—to thrive at the city's Sanitary District outfall

The Pretreatment Program's strategic partnerships go beyond ensuring the success of Publicly Owned Treatment Works (POTWs). The partnerships—involving approximately 1,500 communities and 27,000 industrial facilities nationwide—promote:

- Protection of drinking water supplies, by reducing contaminants released into source waters by POTWs;
- Prevention of overflows that include raw sewage from sewers, through controls on oil and grease;
- Extension of the life of the nation's wastewater infrastructure, which has an estimated funding gap of over $6 billion per year, through controls on corrosion;
- Worker safety, by protecting workers from harmful fumes through limits on the discharge of dangerous gases and gas—forming substances; and
- Homeland security, by ensuring proper disposal of wastes from decontamination showers.

PARTNERSHIPS THAT WORK

Through the Pretreatment Program, POTWs have worked intensively in a federal, state, and local partnership that is a model of intergovernmental cooperation.

National data affirm the Pretreatment Program's success

- Industries release fewer toxic contaminants to POTWs. Data from EPA's Toxics Release Inventory show that industries have reduced the amount of pollutants they discharge to sewage treatment plants during the past 10 years.

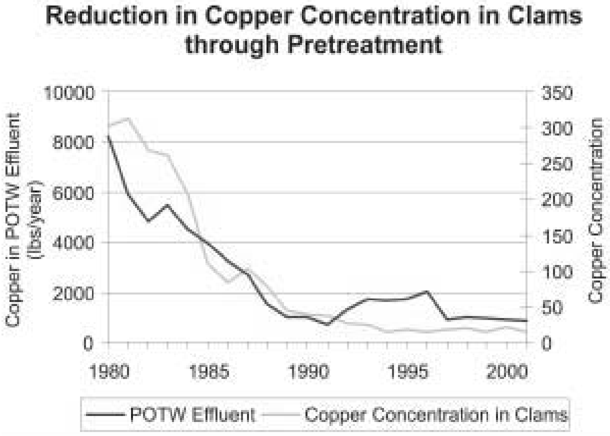

Through its pretreatment program, Palo Alto, California has reduced pollutants discharged to San Francisco Bay. As the quantity of copper in the POTW's effluent has improved, copper levels in clams near the city's outfall have decreased

---

- POTWs now generate sewage sludge that poses fewer threats to public health and the environment. An Association of Metropolitan Sewerage agencies (AMSA) survey of biosolids concluded that the Pretreatment Program was directly responsible for reductions in metals found in sewage sludge. Surveyed POTWs experienced a modest 13 percent decline in metals not regulated by the Pretreatment Program. However, concentrations of metals regulated under the Pretreatment Program decreased considerably, dropping 59 percent from 1986 to 1997.
- Because of the successful implementation of industrial discharge limitations, more POTWs are able to apply sewage sludge to land. For example, six Midwestern states saw a significant increase in the percentage of POTWs whose sludges qualified for the term 'biosolids.' More than 90 percent met metals limitations that allow POTWs to land apply their biosolids. As a result of this national progress, many POTWs, like the Unified Sewerage Agency of Washington County, Hillsboro, Oregon, have been able to use 100 percent of the biosolids they produce. Hillsboro applies 11,000 dry metric tons annually to amend soils and improve crop production.

SUPPORTING KEY ENVIRONMENTAL AND PUBLIC HEALTH PROGRAMS

Drinking Water Protection. EPA's Pretreatment Program has been helping communities protect their sources of drinking water for the last 30 years by:

- Protecting POTWs' treatment processes from upsets, which could compromise the treatment facility's ability to treat harmful substances, such as pathogens. Elevated pathogen levels could have substantial impacts on public health if the pathogens exit POTWs and enter surface waters. Downstream drinking water treatment facilities may be challenged by significantly higher—than—normal pathogen levels.

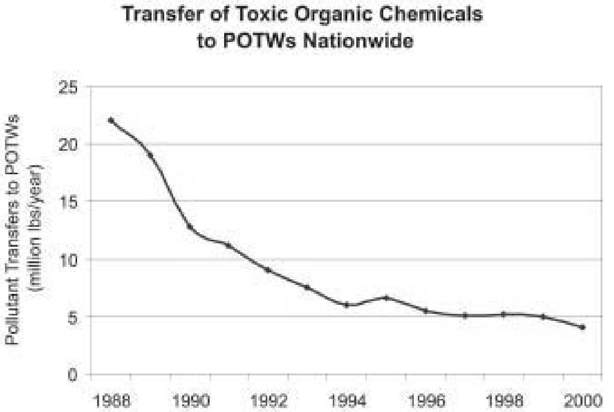

The Toxics Release Inventory is a publicly available EPA database that contains information on toxic chemical releases and other waste management activities reported annually by covered industry groups and the federal government

- Controlling industrial releases of carcinogenic contaminants that might otherwise be discharged to municipal sewage treatment plants. Without controls on industrial dischargers, carcinogens might pass through the sewage treatment plant and enter reservoirs tapped for drinking water supplies. Therefore, the Pretreatment Program eliminates the need to install additional, expensive water treatment equipment to protect the long—term health of consumers.

CSO, SSO, and storm water management—Combined and Sanitary Sewer Overflows (CSOs and SSOs) contain raw sewage, and may also carry industrial waste and debris. EPA's wet—weather programs recognize the importance of the Pretreatment Program in helping communities:

- Mitigate overflows. For example, many cities halt discharges of waste flows from industrial facilities to the combined sewer system during wet weather

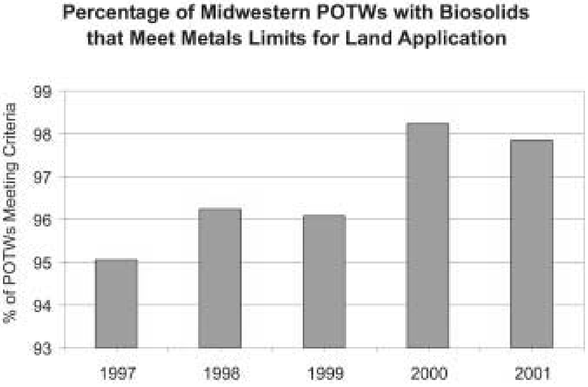

Data obtained from EPA's Permit Compliance System, representing information reported to EPA by POTWs in Illinois, Michigan, Minnesota, Ohio, and Wisconsin

Overflowing sewer due to a sewer blockage events. In addition, industries may retain storm water from their properties and release it to the sewer system after the storm has ended and sewer system capacity is adequate.

- Control grease buildup in sewers—a major cause of SSOs. As part of their pretreatment programs, many communities require food service establishments to recycle all fats, oils, and greases or to use oil and grease removal equipment. Annapolis, Maryland's pretreatment program practically eliminated SSOs associated with the restaurant community.
- Inventory and control non—domestic sources of pollutants, the types of pollutants discharged, and the volume of industrial flow. These inventory efforts not only protect wastewater treatment plants, but also help identify industrial contributions to municipal separate storm sewer systems.

Infrastructure Protection—Pretreatment programs help to protect underground infrastructure from costly damage and the need for premature replacement. The gap between the average annual funding needed and the funding available for wastewater treatment and collection systems is estimated to be $6.1 billion per year from 2000 to 2019. EPA's Pretreatment Program helps to extend the life of infrastructure and postpones costly replacement. For example, limits on pH prevent corrosion of collection systems and treatment plants from acidic discharges. Discharge limits on sulfides and sulfate—bearing wastewaters also protect infrastructure from corrosion caused by bacteria.

Worker and Public Safety—Ensuring safety is an important goal of EPA's Pretreatment Program. The Program accomplishes this goal by:

- Protecting the physical integrity of the sewer system. Volatile organic compounds discharged to sewers may accumulate in the head space of sewer lines, increasing the potential for explosions that may cause significant damage. Discharge limitations and management practices required by the Pretreatment Program reduce the likelihood of such catastrophes.

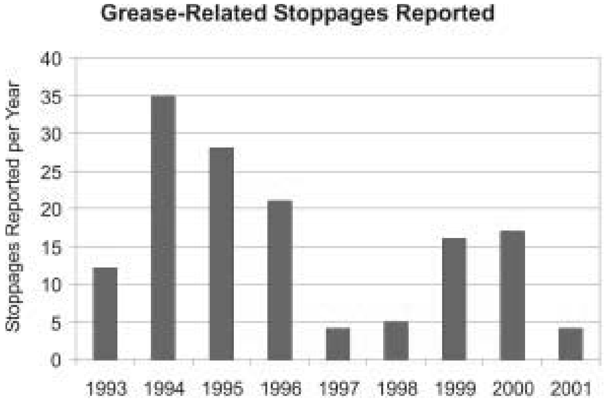

San Antonio, Texas achieved significant reductions in grease—related stoppages through its pretreatment program

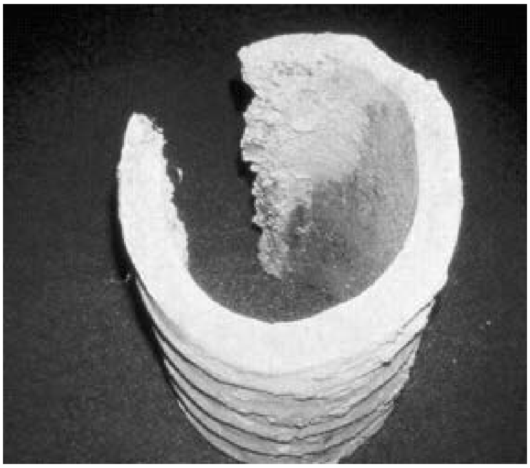

Example of a sewer pipe damaged by hydrogen sulfide corrosion

- Preventing the buildup of poisonous gases. Discharges of toxic organics can generate poisonous gases, through various kinds of mixing and chemical reactions. Appropriate pretreatment discharge limits prevent this gas build up.

Homeland Security Initiatives—As they increase their preparedness for possible terrorist attacks, communities across the country will rely on their local pretreatment programs. The Pretreatment Program helps communities build contingency plans for the control and disposal of decontaminated wastes, such as wastes from decontamination showers. Recently, pretreatment authorities in Boca Raton, Florida managed the disposal of anthrax wastes as part of clean-up activities in that city.

FUTURE CHALLENGES

The Pretreatment Program must evolve to face tough new challenges in protecting public health and the environment. Future challenges include:

- *Addressing emerging industries and pollutants*. Every year, new industries and new pollutants of concern challenge POTWs. For example:

A safety issue: Discharges of hexane from an industry into Louisville, Kentucky's sewer system collapsed sewer lines and destroyed streets throughout the city in 1981

  - The Pretreatment Program must keep pace with the constant shifts in industrial processes and the development of new industries and chemicals. For example, in the 1980s and 1990s, the rapid growth of the semi—conductor industry required new pretreatment program strategies. Mounting evidence now suggests that persistent, bioaccumulative, and toxic (PBT) chemicals, surfactants, and hormone—disrupting substances can pass through wastewater treatment plants and endanger human health. Through their pretreatment programs, communities continue to address changes in industrial operations.
  - New effluent guidelines for transportation equipment cleaning and centralized waste treatment facilities address highly complex industries with a history of disrupting POTW treatment. Communities will rely on the Pretreatment Program as the vehicle to implement these guidelines and control these highly variable and highly toxic wastestreams.
- *Water conservation and reuse*. Industries and municipalities have a growing understanding of the economic benefits of using water more efficiently. Effluent waters from POTWs are increasingly used for irrigation and for cooling in power generation and industrial processes. The high quality waters needed for these uses emphasizes how important pretreatment is to the growing area of water reuse and conservation.
- *Improving watershed quality through Total Maximum Daily Loads (TMDLs)*. TMDLs are established to ensure that rivers and streams meet their intended uses, such as recreational areas and drinking water supplies. State and local pretreatment personnel are increasingly called upon to provide expertise in developing appropriate TMDLs and community—based strategies. These individuals are aware of the contributions from industrial pollutants within the local watershed and sewer—shed, and can share their knowledge of how to implement pollutant limits while pursuing complex watershed—based solutions.

Communities will rely on the leadership of the Pretreatment Program to meet these and other unanticipated challenges. While the Program has had many successes in the last 30 years, a firm commitment to the federal, state, and local partnerships established under the Program is critical to protecting public health and the environment in the future.

PROBLEMS ENCOUNTERED DURING PIPE REPAIR AND RENEWAL

Sanjiv Gokhale
Vanderbilt University
Nashville, Tennessee

INSTALLATION PROBLEMS

Installation Problems—Sliplining

The major problem encountered during sliplining of an existing sewer is improper preparation of the interior of the sewer prior to the inserting the sliplining pipe.

1. Obstructions such as roots, large joint offsets, extended laterals, rocks, or other debris that would prevent the passage of or the damage to the liner pipe sections must be removed or repaired prior to the installation of the new liner pipe.
2. Prior to the insertion of the new liner pipe, a test mandrel or pipe section should be pulled through the existing line to ensure that the liner may be inserted without damage or restrictions.
   (a) The mandrel should be of a length equal to the longest pipe that will be inserted.
   (b) The mandrel test may indicate that short lengths of pipe should be used if the angular deflection negates using longer pipe.

The accurate location and installation of the laterals presents a difficult problem for the Contractor.

1. Existing lateral and sewer connection locations must be carefully determined.
2. The installation of the laterals to provide a leak-tight system is critical.

The diameter of the slipliner pipe is smaller than the host or existing sewer, so inverts and connections to existing manholes must be watched closely.

Grouting the annular space between the existing sewer and the slipliner pipe must be done with great care.

1. Flotation during grouting is a potential problem. The liner pipe should be flooded to a depth of 2 to 3 feet prior to beginning the placement of the grout.
2. The grout should be placed to ensure that the lift heights, grout density, and the sewage flow depth are coordinated to control the line pipe deformation within allowable limits. Point loads causing excessive dents or bulges should be avoided. The maximum pipe diameter deflection for sliplined pipe is 5%. Close monitoring of grout pressure is essential during grouting to prevent buckling of the liner pipe.

Installation Problems—Cured-in-Place Pipe (CIPP)

A difficult problem with CIPP is the location of laterals. The Contractor relies on the “dimples” in the new liner and accurate measurements.

A difficult problem is wrinkles or folds in the flow line in the finished product.

1. This problem could result from improper sizing of the liner bag size.
2. Wrinkles also occur due to head loss in the filling operations.

Inadequate curing of the liner, resulting from the wrong temperature or time duration, may result in soft spots. The Contractor inserts a section of material called dry liner at the soft spot in the liner, which corrects this problem when cured.

Care should be taken to determine if laterals that are to be abandoned are indeed dead. A good rule of thumb to follow is that if you cannot see a plug or cap from inside the main line, then the lateral must be reinstated.

Although not a Contractor problem, a homeowner who has a bad trap on the house sewer may experience a highly offensive odor from the chemical reaction at cure-out. It is offensive but not harmful. This problem can be reduced if the Contractor instructs the homeowners to fill the traps of floor drains with water.

Installation Problems—Gunite

Care should be taken to remove the “rebound.” Some Contractors may try to trowel the rebound and leave it at the bottom of the pipe.

One of the keys for a successful gunite operation is the skill of the nozzleman. Most specifications list qualifications for the nozzleman. Apprentices should be allowed to work only with proper supervision.

Make certain that all leaks in the substrate are stopped. Running water will prevent the material from achieving the proper bond to the substrate.

RADIOACTIVE WASTE

James Crocker
Richland, Washington

Radioactive waste is a concern with regard to water quality when the waste is released into the environment or has the potential to be released into the environment. Air and water containing radioactive contamination may be released directly into the environment if the activity of the contamination is within defined levels. The World Health Organization (WHO) has established release guidelines as explained below. In the United States, the Nuclear Regulatory Commission (NRC) regulates radioactive discharges from commercial nuclear facilities. The NRC regulations are found in Codified Federal Regulations (CFR) 10 CFR 20.106. Limits for radioactive air and water are given in 10 CFR 20, Appendix II, Table 2 for releases to the environment, and in 10 CFR 20, Appendix II, Table 3 for releases of radioactive water into sewers. The United States Department of Energy (DOE) has also established requirements for release of radioactive water and air from DOE facilities. These requirements are provided in DOE Order 4500. The United States Environmental Protection Agency (EPA) provides standards for releases into water sources that may potentially be used as drinking water (see Drinking Water Limits). Of the relatively few instances of radionuclides exceeding levels of concern, most are from natural sources (1,2).

RADIOACTIVITY AND RADIATION

If the energy released during radioactive decay is high enough to strip electrons from other atoms, or split the nuclei of other atoms, the radiation is called ionizing. It is this ionizing radiation that poses the health risk associated with radionuclides. Four types of ionizing radiation exist: alpha, beta, gamma, and neutron. Alpha, beta, and neutron radiation are particles that are released when the isotope decays. An alpha particle is the nucleus of a helium-4 atom, thus consisting of two neutrons and two protons, and no electrons, therefore having a +2 charge. Alpha radiation is easily stopped by only moderate shielding, for example, two or three pieces of paper, or human skin. Neutrons have no charge and therefore do not interact as easily with other atoms when released as radiation. Beta particles are energetic electrons and have a −1 charge. Gamma rays are not particles, but rather electromagnetic radiation, similar to x-rays. Gamma rays and beta particles have both the ability to penetrate farther than alpha particles and interact more than neutrons, and they are therefore typically the more hazardous form of ionizing radiation. However, if alpha particles are ingested, they may present a significant health hazard as an internal contamination. The time for the atoms of a radioactive compound to decay to one half of their starting concentration is called the half-life of the isotope. Two sources of radionuclides exist in drinking water, naturally occurring and anthropogenic.

NATURALLY OCCURRING SOURCES

Naturally occurring radioactive isotopes originate in Earth’s crust and the upper atmosphere. Cosmic rays and other extraterrestrial ionizing radiation form the radioactive isotope tritium in Earth’s upper atmosphere. This isotope is incorporated into natural rainfall and therefore eventually reaches Earth’s surface. Radioactive isotopes, including uranium-238, radium-226, radium-228, and strontium-90, were formed in Earth's crust during the birth of the Solar System about 5 billion years ago. Radon is a daughter product formed when uranium decays. Groundwater may therefore contain levels of radionuclides that are somewhat elevated.

Of the two kinds of naturally occurring isotopes, stable and unstable, only unstable isotopes pose a health concern (3). Therefore, the natural isotopes of primary interest with regard to water quality are uranium-238, radium-226, radium-228, uranium-238, radon, and tritium.

Anthropogenic (Manmade) Sources

The primary sources of manmade radioactive contamination in water are tritiated cooling water effluent from nuclear reactors, tritiated water from nuclear fuel-reprocessing facilities, medical sources, and nuclear fallout from nuclear weapons testing. Effluent from nuclear facilities must meet strict regulatory requirements. In comparison with the large number of safe operating hours accumulated by nuclear facilities over the years, instances of accidental release of radioactive contamination into the environment are rare.

DRINKING WATER LIMITS

The World Health Organization (3) provided guidelines for upper limits of radioactive contamination in drinking water consumed by large populations over a lifetime as:

Gross beta contaimination	1000 pCi/L
Radium-226	10 pCi/L
Strontium-90	30 pCi/L

In the United States, if water is discharged into a water source that may be used for drinking water, such discharge is regulated by the EPA (4). The EPA has updated its standards for radionuclides in drinking water, and it has set a new standard for uranium, as required by the 1986 amendments to the Safe Drinking Water Act. The standards are as follows:

Combined radium-226/228	5 pCi/L
Beta emitters	4 mrems
Gross alpha	15 pCi/L
Uranium	30 μg/L

TREATMENT TECHNOLOGIES

Unlike other chemical contaminants, radionuclides cannot be altered by chemical reactions to form nonradioactive products through normal chemical means. The isotopes that emit radiation will continue to do so no matter what chemical compounds they form. Only through the passage of time will the radioactive compounds decay naturally into their associated daughter products, eventually becoming radioactively stable isotopes. Although the radionuclides will continue to emit radiation, the chemical compounds emitting this radiation can be removed by conventional water treatment methods. Cox (5) stated that 70–90% of the radioactive materials appearing in surface waters are removed in conventional water treatment facilities. According to DeZuane (1), more than 99% of radioactive contaminants that percolate into soils via rainwater are removed naturally. If further removal is required beyond these levels, several standard water treatment methods may effectively remove the contamination (2).

Radionuclides of primary interest that may require removal from water sources are uranium-238, radium-226, radium-228, radon, and strontium-90. Uranium may be removed by lime softening, ion exchange (anion and/or cation), reverse osmosis, electro-dialysis, activated alumina, or a combination of coagulation with sedimentation and filtering. Radium-226 and radium 228 may be removed by lime softening, cationic ion exchange, reverse osmosis, or electro-dialysis. Strontium-90 may be removed by lime softening, cationic ion exchange, reverse osmosis, ultrafiltration (UF), or electro-dialysis. A combination of air stripping and granulated activated carbon is an effective technique for removing radon from water (2).

MEMBRANE FILTRATION PROCESSES

In membrane filtration, particles in the range of 0.0001 μm to 1.0 μm are removed from water. This process involves removal of dissolved constituents, as opposed to other filtration processes that remove colloids and larger particles. In UF, the particles are separated from the water by a sieving mechanism. In reverse osmosis (RO), small particles are rejected by the water layer adsorbed on the surface of the memberane (6). An electrodialysis system passes an electric current through the solution, which causes the cations to move to the anode and the anions to move to the cathode. This movement results in alternating cells of concentrated and dilute salts, which are flushed out from the units and in this way removed from the waste water. Ref. 6, pp. 1104–1137, for additional information.

AIR STRIPPING

Air stripping is an effective method for removing radon from radon-contaminated water (Fig. 1a). In air-stripping equipment, the contaminants in the water phase are transferred into the air phase. The contact area between the clean air phase and the water phase controls the removal rate and therefore the process efficiency. Higher contact areas provide increased removal efficiencies. Water is usually introduced in the top of a cylindrical tower and air is blown in from the bottom, although other arrangements are used. Air stripping may be carried out in several types of equipment, including packed towers, sparged or agitated vessels, or tray towers. In packed towers, high-surface area packing maximizes interphase surface area. Sparged vessels, also called bubble columns, introduce the air into the water to form small bubbles that rise upward through the water. Tray towers contain multiple vertically stacked trays that are made to allow the air to move upward through the water via small openings. The water flows across the trays and down through baffled weirs to the bottom of the tower, which produces foaming and the gas–liquid contact required to produce the mass transfer (6,7).

Several gas stripping systems are being used in the United States to remove radon from drinking water. A good summary is provided in Table 1 (8).

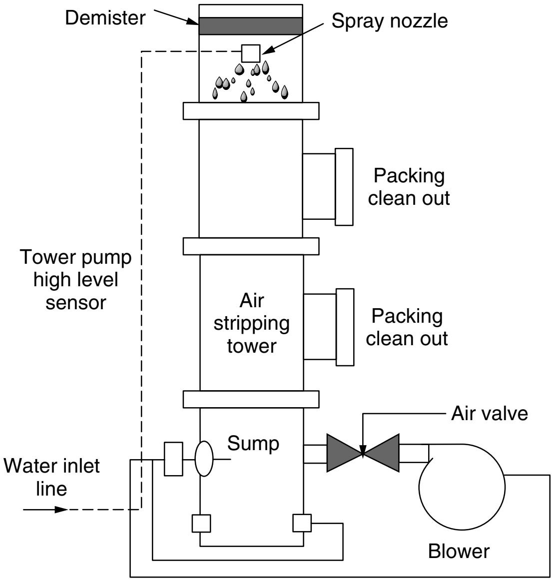

Figure 1. Air stripping town.

CHEMICAL PRECIPITATION

Heavy metals such as uranium and strontium may be removed from contaminated water by converting their soluble salts to insoluble salts that will precipitate. Typically, the pH of the waste stream is first adjusted, and an appropriate chemical is added to convert the soluble metal salts into carbonates, sulfides, or hydroxides of the metal (see LIME SOFTENING). A flocculating agent is then added to promote discrete particle agglomeration, usually performed in a vessel with slow mixing. The flocculated particles are then removed from the stream by either filtration or settling (9). The chemical precipitants and flocculants required depend on the chemistry of the specific waste water system being considered for treatment (6).

LIME SOFTENING

Lime softening is used in a variety of ways to precipitate Group II A ionic species from water (for example, ions of magnesium, calcium, strontium, and radium) and may be used to remove uranium contamination. Lime (calcium hydroxide) is added in either lump, powder, or slurry form to carry out the following simple reactions:

1. $\mathrm{H2CO3} + \mathrm{Ca(OH)2} \rightleftharpoons \mathrm{CaCO3(s)} + \mathrm{2H2O}$
2. $\mathrm{Ca(HCO3)2} + \mathrm{Ca(OH)2} \rightleftharpoons \mathrm{2CaCO3(s)} + \mathrm{2H2O}$

Reaction 1 shows the reaction of the lime with the carbonic acid component typically in water. This reaction does not precipitate any of the desired ionic contaminants, but it must be considered because it creates a lime demand. Reaction 2 shows the removal of two calcium ions. Radioactive ionic contaminants may be removed in a similar manner. Ferrous sulfate is often used in conjunction with lime softening. Many additional chemical reactions may also occur during lime softening. For more information on these reactions, and dosing calculations, refer to Ref. 10, p. 270, and Ref. 6, p. 495 (10–12).

ION EXCHANGE

Ion exchangers operate on the principle of exchanging ions from an insoluble resin with the ions that are present in water. Ion exchange is a common method for removing heavy metals from waste water streams. Resins that typically remove metals include zeolites, anionic and cationic resins, and chelating resins. Chelating and ionic resins that are selective for uranium, strontium, and radium have been developed (6,9).

ACTIVATED CARBON TREATMENT

Activated carbon has been used in conjunction with other waste removal systems such as air stripping to remove radon (8). The radon is adsorbed onto the activated carbon, and it may be regenerated by heating the carbon to sufficient temperatures. See AIR STRIPPING.

BIBLIOGRAPHY

1. De Zuane, J. (1997). *Handbook of Drinking Water Quality*, 2nd Edn. John Wiley & Sons, New York, pp. 341–352.

   CRC. (1995). *Handbook of Chemistry and Physics*, 75th Edn. CRC Press, Boca Raton, FL.

2. Droste, R.L. (1997). *Theory and Practice of Water and Wastewater Treatment*. John Wiley & Sons, New York, pp. 225–226.

   Davis, R.E., Gailey, K.D., and Whitten, K.W. (1984). *Principles of Chemistry*. Saunders College Publishing, New York.

3. World Health Organization. (1963).

   Alberty, R.A. (1987). *Physical Chemistry*, 7th Edn. John Wiley & Sons, New York, p. 218.

Table 1. Efficiencies, Flow, and Construction Costs for Mitigation Systems Being Used in the United States to Remove Radon from Drinking Water

Treatment Method	Removal Efficiency, %	Flow Range, $m^3\ d^{-1}$	Unit Construction Cost, $\$m^{-3}\ d^{-1}$	No. of Systems Evaluated
I. Aeration Methods				
1. Packed tower (PTA)	79 to >99%	49 to 102,740	18 to 481	11
2. Diffused bubble				
a. Single-stage	93	431	312	1
b. Multistage	71 to >99	65 to 6540	11 to 433	8
3. Spray Aeration	∼88	1,025	5.3	1
4. Slat tray	86 to 94	1989 to 2453	5.3 to 124	6
5. Cascade aeration	∼88	5,450	7.9	1
6. Surface aeration	83 to 92	54,504	42	1
II. Granular Activated Carbon	20 to >99	11 to 981	77 to 365	5

A.S.C.E—A.W.W.A. (1990). *Water Treatment Plant Design*, 2nd Edn. McGraw-Hill, New York.

4. U.S. Environmental Protection Agency. *Safe Drinking Water Act and Associated Final Rulings*. Government Printing Office, Washington, DC.
5. Cox, C.R. (1964). *Operation and Control of Water Treatment Processes*. World Health Organization, Brussels, Belgium.
6. Tchobanoglous, G., Burton, F.L., and Stensel, D.H. (2003). *Wastewater Engineering—Treatment, Disposal and Reuse*, 4th Edn. Metcalf & Eddy, Inc., McGraw-Hill, New York, pp. 1104–1137, 1162–1196.

Dean, J.A. (1985). *Lange's Handbook of Chemistry*, 15th Edn. McGraw-Hill, New York.

7. Treybal, R.E. (1988). *Mass Transfer Operations*, 3rd Edn. McGraw-Hill Book Companies, New York, pp. 140–144, 158–211, 641–646.
8. Commission on Life Sciences. (1999). *Risk Assessment of Radon in Drinking Water*. Committee on Risk Assessment of Exposure to Radon in Drinking Water, Board on Radiation Effects Research, National Research Council, National Academy Press, Washington, DC.
9. Stocchi, E. (1990). *Industrial Chemistry*. Ellis Horwood, London, pp. 103–167.
10. Benefield, L.D., Judkins, J.F., and Weand, B.L. (1982). *Process Chemistry for Water and Wastewater Treatment*. Prentice-Hall, Englewood Cliffs, NJ, pp. 270–271, 293–294.

Galperin, A.L. (1992). *Nuclear Energy, Nuclear Waste*. Chelsea House Publishers, New York.

11. Hafele, W. (1990). Energy from nuclear power. *Scientif. Am.* September.
12. U.S. Environmental Protection Agency. (1985). *Guidance Manual For Compliance With the Filtration and Disinfection Requirements For Public Water Systems Using Surface Water Sources*. Science and Technology Branch-Criteria and Standards Division-Office of Water, USEPA Publications, Washington, D.C., or NTIS, Springfield, VA.

READING LIST

Montgomery, C.W. (1992). *Environmental Geology*, 3rd Edn. Wm. C. Brown Publishers, Chicago, IL.

Hazardous Waste Cleanup Information, United States Environmental Protection Agency. Available: http://www.clu-in.com.

Federal Remediation Technologies Roundtable. Available: http://www.frtr.gov.

RECLAIMED WATER

Absar Ahmad Kazmi
Nishihara Environment
Technology, Inc.
Tokyo, Japan

INTRODUCTION

Reclaimed water is water from a wastewater treatment plant (WWTP) that has been treated and can be used for nonpotable uses such as landscape irrigation, cooling towers, industrial processes, toilet flushing, and fire protection. The inclusion of planned water reclamation, recycling, and reuse in water resource systems reflects the increasing scarcity of water resources to meet societal demands, technological advancements, increased public acceptance, and improved understanding of public health risks.

Per capita water use in the United States has quadrupled since the beginning of the twentieth century. Americans typically consume between 60 and 200 gallons (230 to 760 liters) per capita each day. The use of reclaimed water for nonpotable purposes can greatly reduce the demand on potable water sources—this use is encouraged by diverse organizations such as FEMP, EPA, and the American Water Works Association (AWWA). Municipal wastewater reuse amounts to about 4.8 billion gallons (18 million m^3) per day (about 1% of all freshwater withdrawals). Industrial wastewater is far greater—about 865 billion gallons (3.2 billion m^3) per day.

Reclaimed water contains valuable nitrogen, phosphorus, and other nutrients, which promote plant growth. At the same time, the water meets stringent disinfection standards. Experience has shown that contact with reclaimed water does not promote waterborne disease transmission. In fact, reclaimed water quality standards are more stringent than those for surface streams, rivers, and irrigation channels. Reclaimed water is delivered through pipelines that are completely separate from the potable water system.

In some areas of the United States, reclaimed water may be referred to as irrigation quality ("IQ") water, but potential uses can extend well beyond irrigation. Using higher levels of treatment, such as reverse osmosis, reclaimed water as a potable source is technically and economically feasible. New technological breakthroughs in membrane filtration and combined biological and filtration treatment offer unprecedented opportunities for water recycling, especially in isolated locations and regions where the water supply is severely limited.

BENEFITS OF USING RECLAIMED WATER

- It saves millions of gallons of drinking water each day.
- Its use for nonpotable (nondrinking) purposes is less expensive for the vast majority of of customers.
- It delays the need for developing costly new water sources and building very expensive treatment plants.
- There is no odor or staining from its use.
- It allows a city to comply with permits relating to its water supply and wastewater treatment.
- It minimizes negative effects around underground water sources, preserving the quality of life for plants and wildlife.
- It reduces fertilizing costs because reclaimed water is rich in nitrogen and phosphorus.

POTENTIAL USES OF RECLAIMED WATER

Urban public water supplies are treated to satisfy requirements for potable use. However, potable use

Table 1. Summary of EPA Suggested Guidelines for Water Reuse[a]

Levels of Treatment	Types of Reuse	Reclaimed Water Quality	Reclaimed Water Monitoring	Setback Distances
1. Disinfected tertiary[b]	Urban reuse[c]	pH = 6–9	pH = weekly	15 m (50 ft) to potable water supply wells[d]
	Food crop irrigation	$BOD_5 \leq 10$ mg/L Turb. ≤ 2 NTU	BOD = weekly Turb. = cont.	
	Recreational impoundments	*E. coli* = none Res. $Cl_2 \geq 1$ mg/L	*E. coli* = daily Res. Cl_2 = cont.	
2. Disinfected secondary	Restricted-access-area irrigation	pH = 6–9	pH = weekly	30 m (100 ft) to areas accessible to the public (if spray irrigation)
		BOD_5 = 30 mg/L	BOD = weekly	
	Food crop irrigation (commercially processed)	TSS = 30 mg/L	TSS = cont.	
		E. coli = 200/100 mL	*E. coli* = daily	
			Res. Cl_2 = cont.	90 m (300 ft) to potable water supply well
	Nonfood crop irrigation Landscape impoundments (restricted access) Construction Wetlands habitat	Res. $Cl_2 \geq 1$ mg/L		

[a] From Reference 1.

[b] Filtration of secondary effluent.

[c] Uses include landscape irrigation, vehicle washing, toilet flushing, fire protection, and commercial air conditioners.

[d] Setback increases to 150 m (500 ft) if impoundment is not sealed.

(drinking, cooking, bathing, laundry, and dishwashing) represents only a fraction of the total daily residential use of treated potable water. The remainder may not require water of potable quality. In many cases, water used for nonpotable purposes, such as irrigation, may be drawn from the same ground or surface source as municipal supplies, creating an indirect demand on potable supplies. There are opportunities for substituting reclaimed water for potable water or potable supplies for uses where potable water quality is not required. Specific water use where reuse opportunities exist include

- urban
- industrial
- agricultural
- recreational
- habitat restoration/enhancement, and
- groundwater recharge

FEDERAL GUIDELINES FOR RECLAIMED WATER REUSE

The U.S. Environmental Protection Agency (1) has suggested reclaimed water quality guidelines for the following reuse categories:

- urban reuse
- restricted-access-area irrigation
- agricultural reuse—food crops
- agricultural reuse—nonfood crops
- recreational impoundments
- construction uses
- industrial reuse
- groundwater recharge
- indirect potable reuse

Levels of treatment, minimum reclaimed water quality, reclaimed water monitoring, and setback distances are suggested for each reuse category (1). The guidelines are summarized in Table 1 for the two principal levels of treatment—disinfected tertiary (filtered secondary effluent) and disinfected secondary effluents.

CONSTITUENTS OF RECLAIMED WATER

The constituents of municipal wastewater subject to treatment may be classified as conventional, nonconventional, and emerging. Typical constituents included under each category are described in Table 2. The term conventional is used to define those constituents measured in mg/L that are the basis for designing most conventional wastewater treatment plants. Nonconventional applies to those constituents that may have to be removed or reduced using advanced wastewater treatment processes before the tank can be used beneficially. The term emerging is applied to those classes of compounds measured in the micro- or nanogram/L range that may pose long-term health concerns and environmental problems as more is known about the compounds. In some cases, these compounds cannot be removed effectively, even by advanced treatment processes.

WATER RECLAMATION TECHNOLOGIES

As noted in the previous section, the constituents of wastewater subject to treatment may be classified as conventional, nonconventional, and emerging. Conventional

Table 2. Classification of Typical Constituents in Wastewater

Classification	Constituent
Conventional	Total suspended solids
	Colloidal solids
	Biochemical oxygen demand
	Chemical oxygen demand
	Total organic carbon
	Ammonia
	Nitrate
	Nitrite
	Total nitrogen
	Phosphorus
	Bacteria
	Protozoan cysts and oocysts[a]
	Viruses[b]
Nonconventional	Refractory organics
	Volatile organic compounds
	Surfactants
	Metals
	Total dissolved solids
Emerging	Prescription and nonprescription drugs[c]
	Home care products
	Veterinary and human antibiotics
	Industrial and household products
	Sex and steroidal hormones
	Other endocrine disrupters

[a] Value per 100 mL.
[b] Plaque-forming units/100 mL.
[c] Pharmaceutically active substances.

constituents are removed by conventional treatment technologies. Advanced treatment technologies are used most commonly for removing nonconventional constituents. The removal of emerging constituents occurs in both conventional and advanced treatment processes, but the levels to which individual constituents are removed are not well defined. Typical performance data for selected treatment combinations are presented in Table 3.

PLANNING FOR WASTEWATER RECLAMATION AND REUSE

In effective planning for wastewater reclamation and reuse, the objectives and basis for conducting the planning should be defined clearly. The optimum water reclamation and reuse project is best achieved by integrating both wastewater treatment and water supply needs into one plan. This integrated approach is somewhat different from planning for conventional wastewater treatment facilities where planning is done only for conveyance, treatment, and disposal of municipal wastewater. Effective water reclamation and reuse facilities should include the following elements: (1) assessment of wastewater treatment and disposal needs, (2) assessment of water supply and demand, (3) assessment of water supply benefits based on water reuse potential, (4) analysis of reclaimed water market, (5) engineering and economic analyses of alternatives, (6) implementation plan

Table 3. Treatment Levels Achievable from Various Combinations of Unit Operations and Processes Used for Water Reclamation

	Typical Effluent Quality, mg/L, Except Turbidity, NTU						
	TSS	BOD$_5$	COD	Total N	NH$_3$-N	PO$_4$-P	Turbidity
Activated sludge + granular medium filtration	4–6	<5–10	30–70	15–35	15–25	4–10	0.3–5
Activated sludge + granular medium filtration + carbon adsorption	<5	<5	5–20	15–30	15–25	4–10	0.3–3
Activated sludge/nitrification single stage	10–25	5–15	20–45	20–30	1–5	6–10	5–15
Activated sludge/nitrification denitrification separate stages	10–25	5–15	20–35	5–10	1–2	6–10	5–15
Metal salt addition to activated sludge +nitrification/ denitrification separate stages	≤5–10	≤5–10	20–30	3–5	1–2	≤1	0.3–2
Biological phosphorus removal[a]	10–20	5–15	20–35	15–25	5–10	≤2	5–10
Biological nitrogen and phosphorus removal + filtration	≤10	<5	20–30	≤5	≤2	≤2	0.3–2
Activated sludge + granular medium filtration + carbon adsorption +reverse osmosis	≤1	≤1	5–10	<2	<2	≤1	0.01–1
Activated sludge/nitrification-denitrification +granular medium filtration + carbon adsorption + reverse osmosis	≤1	≤1	2–8	≤1	≤0.1	≤0.5	0.01–1
Activated sludge/nitrification-denitrification and phosphorus removal + microfiltration +reverse osmosis	≤1	≤1	2–8	≤0.1	≤0.1	≤0.5	0.01–1

[a] Removal process occurs in the main flowstream as opposed to sidestream process.

with financial analysis, and (7) a public information program.

TECHNICAL CONSIDERATIONS

For a successful reclaimed water project, one or more of the following ingredients are required: (1) high cost water or a need to extend the drinking water supply, (2) local public policy encouraging or mandating water conservation, (3) availability of high quality effluent from a WWTP, and (4) recognition of environmental or other nontangible benefits of water reuse.

Technologies vary with end uses. In general, tertiary or advanced secondary treatment is required; either usually includes a combination of coagulation, flocculation, sedimentation, and filtration. Viruses are inactivated by granular carbon adsorption plus chlorination or by reverse osmosis, ozonation, or UV exposure. Dual water systems are beginning to appear in some parts of the country where the water supply is limited, such as Southern California. Office buildings may have two water lines coming in—one for "freshwater" and the other for reclaimed water. The former is for all potable uses, the latter for nonpotable uses.

Piping and valves used in reclaimed water systems should be color-coded with purple tags or tape. This minimizes piping identification problems and cross-connection problems when installing systems. Liberal use of warning signs at all meters, valves, and fixtures is also recommended. Note that potable water mains are usually color-coded blue, and sanitary sewers are green. Reclaimed water should be maintained at 10 psi (70 kPa) lower pressure than potable water mains to prevent backflow and siphonage in the event of accidental cross-connection. Although it is feasible to use backflow prevention devices for safety, it is imperative never to connect reclaimed and water piping directly. One additional precaution is to run reclaimed water mains at least 12 in. (30 cm) lower (in elevation) than potable water mains and to separate them from potable or sewer mains by a minimum of 10 ft (3 m) horizontally. Reclamation can be complex when the water supplier and the wastewater utility are not the same. In addition, issues of water ownership are withdrawn from one use to accommodate another.

BIBLIOGRAPHY

1. U.S. EPA. (1992). *Manual—Guidelines for Water Reuse*. U.S Environmental Protection Agency and US Agency for International Development, Washington, DC.

READING LIST

Metcalf & Eddy, Inc. (2003). *Wastewater Engineering: Treatment, Disposal, Reuse*, 4th Edn. McGraw-Hill, New York, pp. 1345–1438.

Asano, T. (Ed.). (1998). *Water Quality Management*. Vol. 10, Technomic, Lancaster, PA.

WPCF. (1989). *Water Reuse: Manual of Practice*. 2nd Edn. Water Pollution Control Federation, Alexandria, VA.

WASTEWATER TREATMENT AND RECYCLING TECHNOLOGIES

Imram Ali
C.K. Jain
National Institute of Hydrology
Roorkee, India

Due to the continuous population growth, uneven distribution of water resources, periodic droughts, and water pollution, it is very important to develop water treatment and recycling technologies for economic, effective, and fast water treatment and reclamation. Different types of water treatment and recycling techniques and their basic principles, applications, costs, maintenance, and suitability are discussed. In addition, a systematic approach to water treatment and recycling, involving understanding, evaluating, and selecting parameters, is presented. A comparison of the technologies is discussed from the view points of performance, sludge production, life cycle, operation, and maintenance. Brief guidance for selecting appropriate technologies for specific applications is also evaluated.

INTRODUCTION

Water is a very important constituent of the ecosystem on the earth and an essential component of life. The quality of our water resources is deteriorating day by day due to the continuous addition of undesirable constituents to them. The main sources of water contamination are industrial, domestic, and agricultural activities and other environmental and global changes. The groundwater and surface water at many places in the world are not suitable for drinking because of the presence of toxic pollutants. If this continues, the world will be in great trouble in a few decades. Therefore, the importance of water quality preservation and improvement is essential. Continuous population growth, uneven distribution of water resources, and periodic droughts have forced scientists to search for new and effective water treatment and recycling technologies (1). There is a great need for developing suitable, inexpensive, and rapid wastewater treatment and reuse or conservation methods.

In this article, different types of wastewater treatment and recycling techniques and a systematic approach to water treatment and recycling methods are discussed. Technologies are compared in terms of their performance, sludge production, life cycle, operation, and maintenance costs. The main purpose of this article is to provide guidance for selecting appropriate technologies or combinations for specific applications.

WASTEWATER TREATMENT AND RECYCLING TECHNOLOGIES

Water treatment technologies are used for three purposes: water source reduction, wastewater treatment, and recycling. At present, unit operations and processes are combined to provide what are called primary, secondary, and tertiary treatment. Primary treatment includes a very preliminary physical purification process (filtration by bar screen, grit chamber, etc.). Secondary treatment deals with chemical and biological processes for treating wastewater. In the tertiary treatment process, wastewater (treated by primary and secondary processes) is converted into good quality water which can be used for different purposes such as drinking, industrial, and medicinal. In the tertiary process, up to 90 to 99% of the pollutants are removed, and water is safe for the specific use. In a complete water treatment plant, all three processes are combined to produce good and safe quality water.

Despite various advanced technologies for water treatment and reclamation, economic, effective, and rapid water treatment and reclamation on a commercial level is still a challenging problem. Prior to water treatment and reclamation, one should be aware of the qualitative and quantitative nature of water pollutants. Managing the removed pollutants (sludge) should also be kept in mind. A systematic approach to water treatment and recycling technologies involves understanding the technology which includes construction and operating cost, maintenance, and management of removed pollutants. A comparison of these wastewater treatment and reclamation technologies is presented in Table 1. A detailed literature survey of water treatment and recycling technologies has been carried out through analytical, chemical, water abstracts, and other journals, and a brief discussion of them is presented here. Water treatment and recycling technologies classified on the basis of their working principles are given below:

1. physical technologies
2. chemical technologies
3. electrical technologies
4. thermal technologies
5. biological technologies

Physical Technologies

Treatment and recycling technologies involving physical forces are known as physical technologies. These include screening, filtration and centrifugal separation, micro- and ultrafiltration, reverse osmosis, crystallization, sedimentation and gravity separation, flotation, and adsorption.

Screening, Filtration, and Centrifugal Separation. Pieces of cloth, paper, wood, cork, hair, fiber, kitchen refuse, and fecal solids in wastewater are removed by screening. The main idea of screening is to remove solid wastes from wastewater. Generally, screening is used as the very first step in a wastewater treatment plant. Screens of various sizes are used for this purpose, the size of the screen selected is based on the size of the solids in the wastewater.

Table 1. A Comparison of Wastewater Treatment and Recycling Technologies

Wastewater Technologies	Applicability[a]	Suitability[b]	Cost (US$ per Million Liters of Treated Water)
	A. Physical technologies		
Screening, filtration, and centrifugal separation	Ss & Sl IOB	RSrT	20–450
Micro- and ultra-filtration	Sl IOB	RSrT	10–400
Reverse osmosis	Sl IOB	RSrT	10–450
Crystallization[c]	Sl IO	RSrT	50–150
Sedimentation and gravity separation	Ss IOB	RSrT	2–10
Flotation	Ss IOB	RT	5–25
Adsorption	Ss & Sl IOB	RSrT	50–150
	B. Chemical technologies		
Precipitation[c]	Sl IO	RT	15–500
Coagulation	Ss & Sl I	RT	20–500
Oxidation	Sl IO	RSrT	100–2000
Ion exchange	Sl IO	RSrT	50–200
Solvent extraction	Sl OV	RSrT	250–2500
	C. Electrical technologies		
Electrodialysis	Sl IO	RSrT	10–400
Electrolysis	Sl IO	RSrT	—
	D. Thermal technologies		
Evaporation[c]	Sl & Ss IOB	RSrT	10–200
Distillation	Sl IOB	RT	10–2000
	E. Biological technologies		
Aerobic	Sl & Ss O	RT	10–200
Anaerobic	Sl & Ss O	RT	10–200

[a] Sl: soluble; Ss: suspended; I: inorganics; O: organics; V: volatiles; B: biologicals.
[b] R: reclamation; T: treatment; and Sr: source reduction.
[c] Rarely Used.

Filtration is a very simple physical process in which insoluble contaminants are removed by passing the water through a setup of pores of different sizes, depending on the presence of solid contaminants. It is used to remove suspended solids, greases, oils, and bacteria. Various types of filters such as membranes and cartridges, made of sand, gravel and other granular materials are used. The filtration technique is applicable below 100 mg/L suspended solids and 25 mg/L oil and grease. These constituents can be reduced up to 99%. Filtration is used for both water treatment and recycling. Water produced by filtration is used in adsorption, ion exchange, or membrane separation processes. Potable water is also produced by filtration. The cost of filtration varies from 20 to 450 US$ per million liters of treated water (1,2).

In centrifugal separation, suspended noncolloidal solids (up to 1 μ) are separated from water by centrifugal forces. Wastewater is placed in centrifugal devices and rotated at different speeds, and the solids (sludges) are separated and discharged. The extent of separation of suspended solids depends on their densities and the speed of the centrifuge. The applications include source reduction and separation of oils and greases. The different types of centrifuges available and in use are solid-bowl, basket type, directflow, and countercurrent flow. The cost of wastewater treatment ranges from 20 to 450 US$ per million liters of treated water (1,2).

Micro- and Ultrafiltration. Particles and other microbes from 0.04 to 1 μ are removed by microfiltration provided that the total suspended solids do not exceed 100 mg/L. The filters used are in the form of cartridges. Commercially available cartridges are made of cotton, wool, rayon, cellulose, fiberglass, polypropylene, acrylics, nylon, asbestos, and fluorinated hydrocarbon polymers. These are arranged in as tubular, disc, plates, spiral, and hollow fiber forms. The life of cartridges varies from 5 to 8 years depending upon the concentration of dissolved solids. Preremoval of suspended solids is an important factor in the life of cartridges. The operating pressure in this process ranges from 1 to 3 bar. Applications include removal of solids and microbes. Water purified by this technique is used for the food and drink industries, soft drinks, pharmaceuticals, photofilm processing, swimming pools, and drinking (2,3). It has also been used as a wastewater source reduction technique.

Ultrafiltration is a low-pressure membrane separation process that removes high molecular weight materials, colloids, pyroxenes, microorganisms, and suspended solids from wastewater. Ultrafiltration membranes are manufactured from a wide variety of polymers, and minerals in the range of 0.005 to 0.10 μ. The membranes are made of polysulphonates, polyacrylonitriles, polyamides, PVDF, and zirconium oxides. To achieve the required filtration, membranes are arranged in tubular, disc, plates, spiral, and hollow fiber forms. The life of membranes varies from 5 to 8 years and may be increased as discussed above (2). The cost of treated water varies from 10 to 400 US$ per million liters.

Reverse Osmosis. Reverse osmosis (RO), also known as hyperfiltration, is a classical method of purification that came into existence since the advancement of semipermeable membranes. It has received great attention nowadays as the best water recycling technique. The separation and concentration of a dissolved species is achieved due to the hydraulic gradient across the semipermeable membrane. Pressure greater than osmotic pressure is applied for the process. The most commonly used membranes are made of cellulose, nylon, polyether, polyethyl urea, polyphenyl oxides, phenylenes, and polyamide. To achieve the required filtration, membranes are arranged in tubular, disc, plate, spiral, and hollow fiber forms. The partition coefficients of solutes between water and the membrane play an important role in removing water pollutants. The free energy of interaction between water and membrane sites is also responsible for the RO process. The pH, pressure, size, and molecular weight of the solute and time of operation are considerable factors in RO.

RO has been used as a separation and concentration technique at macro- and microlevels for removing large, nonpolar, ionic, and toxic substances. Up to 85–99% total dissolved solids (TDS), organic dissolved matter (ODM), and bacteria can be removed by this method. It has been used for treating wastewater from sanitary wastes, municipal leachates, petrochemicals, electroplating, textiles, coal, gasification, pulp and paper, steel, and electronic industries (3,4). It rejects 100% of bacteria, viruses, and other microbes, and, therefore, it is used to prepare ultrapure water for pharmaceuticals, medicines, and electronics. In addition, it has been used for source reduction. RO is today's most economical process for potable water production from saline water.

The life of RO membranes is 2–5 years, depending on the nature of the wastewater treated. The flux and the quality of the permeate may decrease over a long period of time due to membrane fouling from humic acids, bacterial slimes, or scales that may accumulate on the RO membranes. Phenols also clog the membranes. To increase the efficiency and life of RO systems, pretreatment is necessary to minimize the concentration of colloidal and dispersed solids. Physicochemical coagulation with lime has been used to minimize colloids, turbidity, dispersed oil phases, metal ions, and suspended matters. Sodium hydroxide solution (pH 9–11) has been used to clean RO membranes in case of silica and sulfate fouling (5,6). Silica can also be removed from membranes by ion retarding resins that have high affinity for strong acids, together with conversion of the weak acid [$Si(OH)_4$] into much stronger acid (H_2SiF_6). Bacterial inhibitor solutions are circulated (to check the bacterial growth) into the RO tubules or discs before stopping the process for a long period (5). Phenolics may be removed from RO membranes by circulating hydrogen peroxide solution. The cost of the process varies from 10 to 450 US$ per million liters of treated water.

Crystallization. In this process, soluble constituents are removed by raising their concentrations to the point where they start to crystallize. This is done either by evaporation, by lowering the temperature of the water, or by adding other solvents. It is useful for treating wastewater that has high concentrations of TDS, including soluble organics and inorganics. During the process, other constituents such as bicarbonate, ammonia, and sulfite may break down and may be converted into various gases and, therefore, crystallization sometimes may be used for pH control. The treated water from this process is of high quality. Crystallization is generally used for wastewater released from cooling towers, coal and gas fired boilers, paper, and dyeing plants. It is also used for source reduction. The commonly used devices for crystallization include forced circulation, draft tube baffle, surface cooled crystallizers, and fluidized suspension. The cost of the technique ranges from 50 to 150 US$ per million liters of treated water (3,7).

Sedimentation and Gravity Separation. In this process, suspended solids, grit, and silt are removed by allowing water to remain undisturbed/semidisturbed for different time periods. The suspended solids settle by gravity (1,2). The time period depends on the size and density of the solids. Various types of tanks are designed for this purpose. Some chemicals such as alum are used to adjust pH and augment the process. Gravity separation can reduce oil concentrations and suspended solids up to 99% and 60%, respectively. Generally, sedimentation is carried out prior to a conventional treatment process. It is a very useful method for treating effluents from the paper and refinery industries. Water treated in this process is used for industrial water supply, water for ion exchange, and membrane processing. The technique is also used for source reduction. The cost of treated water varies from 2 to 10 US$ per million liters.

Flotation. This technique removes suspended solids, oils, greases, and biological solids by adhering them, to air or gas bubbles (1,2,8). The solids thus adhered to gas or air bubbles form agglomerates, which in turn accumulate at the water surface and are removed. Some chemicals such as alum and activated silica help in the flotation process. Compressed air is allowed to pass through water, which helps in the flotation process. Some workers have also used electroflotation as an effective process for water treatment and recycling. Up to 75% and 99% of suspended solids and oil/grease are removed, respectively, by this process. Flotation requires water tanks of different sizes. Flotation is a common and essential component of a conventional water treatment plant. It is a very effective technique for treating wastewater from the paper and refinery industries. The cost of the operation varies from 5 to 25 US$ per million liters of treated water.

Adsorption. Adsorption (1,2,9) is a surface phenomenon defined as the increase in concentration of a particular component at the surface or interface between two phases. Adsorption efficiency depends on a number of parameters such as pH, temperature, concentration of pollutants, contact time, particle size of the adsorbent, and nature of adsorbents and pollutants. Suspended particles, oils, and greases reduce the efficiency of the process and, therefore, pre-filtration is required. It is considered a universal water treatment and reclamation process because it can be applied to remove soluble and insoluble organics, inorganics, and biological solids. Different types of adsorbents are used in the adsorption process. The most commonly used adsorbents are activated carbon, fly ash, metal oxides, zeolites, moss, biomass, and geothites. At the industrial level, pollutants are removed from wastewater by using columns and contactors filled with the required adsorbents. The extent of removal varies from 90 to 99%. Adsorption is used for source reduction, wastewater treatment, and reclamation for potable, industrial, and other purposes. The basic problems of adsorption are regeneration of columns and column life. The cost of the technique ranges from 50 to 150 US$ per million liters of treated water.

Chemical Technologies

Water treatment methods involving the use of chemicals are chemical technologies. Precipitation, coagulation, oxidation, ion exchange, and solvent extraction are the main chemical methods for wastewater treatment and reclamation.

Precipitation. Dissolved contaminants may be converted into solid precipitates by adding chemicals (2,10) that react with the soluble pollutants and form precipitates. The most commonly used chemicals for this purpose are different types of alum, sodium bicarbonate, ferric chloride, ferric sulfate, ferrous sulfate, and lime. pH and temperature are the controlling factors in the precipitation process. Precipitation is carried out in sedimentation tanks; 40 to 60% removal of pollutants by precipitation has been reported. The presence of oil and grease may cause a problem in precipitation. The applications of precipitation include wastewater treatment (from nickel and chromium plating) and water recycling. Specific applications include water softening and removal of heavy metals and phosphate from wastewater. The major problem in precipitation is managing the large volume of sludge produced. The cost varies from 15 to 500 US$ per million liters of treated water.

Coagulation. The suspended nonsettleable solids in wastewater are allowed to settle by the addition of certain chemicals in a process called coagulation (1,11). The commercially available chemicals are alum, starch, iron compounds, activated silica, and aluminum salts. In addition, synthetic cationic, anionic, and nonionic polymers are very effective coagulants but are usually more costly. pH, temperature, and contact time are the most important controlling factors in the coagulation process. In a biological treatment plant, microbes and other organics floated on the surface are removed by the addition of certain coagulants. It is the main component of a wastewater treatment plant and its application includes wastewater treatment, recycling, and removal of heavy metal ions and fluoride. The cost of the treated water varies from 20 to 500 US$ per million liters.

Oxidation. In chemical oxidation, organic compounds are converted into water and carbon dioxide or some other products such as alcohols, aldehydes, ketones and carboxylic acids which are biodegradable (1,12). Chemical oxidation is carried out by potassium permanganate, chlorine, ozone, peroxides, air, and chlorine dioxides. The rate of chemical oxidation depends on the nature of the oxidants and pollutants. pH, and temperature also play a crucial role in the rate of chemical oxidation. Ammonia, cyanide, sulfides, phenols, hydrocarbons, and some pathogens may be removed by chemical oxidation. Chemical oxidation is used for wastewater treatment and recycling for industry and irrigation. It is also a useful and effective method for source reduction. The cost of the technique ranges from 100 to 2000 US$ per million liters of treated water.

Ion Exchange. Ion exchange is a process in which ions in wastewater are exchanged with solid materials called ion exchangers (1,13). It is a reversible process and requires low energy. The ion exchangers are of two types, cation and anion exchangers, that can exchange cations and anions, respectively. Ion exchangers are resins of natural or synthetic origins that have active sites on their surfaces and, generally, are in the form of beads. The most commonly used ion exchangers are sodium silicates, zeolites, polystyrene sulfonic acid, acrylic and methacrylic resins. Ion exchange is used to remove low concentrations of inorganics and organics (up 250 mg/L). The concentration of organics and inorganics can be reduced up to 95%. Applications include the production of potable water, water for industries, pharmacy, research, and softening for boiler feed, fossil fuels, nuclear power stations, paper, and, electronic industries. It has also been used for source reduction. Pretreatment of water is required in the presence of oil, grease, and high concentrations of organics and inorganics. One million liters of wastewater are treated for 50 to 200 US$.

Solvent Extraction. Organic solvents that are immiscible with water and can dissolve water pollutants are added to wastewater to remove pollutants. The technique is called solvent extraction (14). A maximum concentration of TDS of 2000 mg/L can be reduced up to 90% by solvent extraction. The most commonly used solvents are benzene, hexane, acetone, and other hydrocarbons. The technique is effective to remove only the dissolved organics, oils, and greases in wastewater. However, certain metal ions and actinide chemicals may be removed by the method. It is also used for water treatment and recycling in chemical process plants, phenol, gasoline, and acid industries. It has also been used for water source reduction. The presence of suspended solids may cause a problem in solvent extraction and, hence, requires pretreatment. The cost varies from 250 to 2500 US$ per million liters of treated water.

Electrical Technologies

Water pollutants are removed under the influence of electric current in electrical water treatment and recycling technologies. Electrical water treatment technologies are summarized below.

Electrodialysis. In this technique, water soluble ions are allowed to pass through ion selective semipermeable membranes under the influence of an electric current (1–3,15). The ion selective membranes are made of ion exchange material. They may be cation and anion exchangers, which permit outflow of cations and anions, respectively. The process, operated either in a continuous or batch mode, has two electrodes on which an emf. is applied. To obtain the desired degree of demineralization, membranes are arranged either in parallel or series. The dissolved solids removal depends on pH, temperature, the amount of current applied, the nature of the pollutants, selectivities of the membranes, the wastewater flow rate, fouling and scaling by wastewater, and the number and configuration of stages. Applications include production of potable water from brackish water. This technique has also been used for water source reduction. A maximum concentration of 200 mg/L of TDS can be reduced by 90% by electrodialysis. Membrane fouling occurs as in reverse osmosis. Cleanup and other precautions should be taken as discussed in the reverse osmosis section. The cost of treated water varies from 10 to 400 US$ per million liters of water.

Electrolysis. The technique in which the soluble inorganics and organics are either deposited or decomposed on the surface of electrodes by an electrochemical redox reaction is called electrolysis (16). Metal ions are deposited on the electrode surface, and organics are decomposed into carbon dioxide and water or some other products. It has been used to remove turbidity and color from wastewater. This method is effective for the removal of TDS below 200 mg/L and, therefore, requires pretreatment of wastewater. The technique comprises a water tank or tanks in series with two or a series of electrodes of the required metal. The electrodes are specific with respect to the dissolved metal ions in wastewater. The most important controlling factors for this process are pH, temperature, amount of current applied, and contact time. Electrolysis as a technique for wastewater treatment is not yet developed completely and is still at the research and development stage. It has been rarely used commercially for wastewater treatment. However, its applications include treating some industrial effluents especially enriched with metal ions and some organics. It may be used as a water source reduction technique. The advantages of this technique comprise the further use of deposited metal ions without any waste management problem.

Thermal Technologies

Techniques involving the use of heat energy for water treatment and recycling are thermal technologies. The most commonly used techniques for wastewater treatment and recycling are evaporation and distillation.

Evaporation. Evaporation is a natural process and is generally used to reduce waste liquid volume. In modern development, it has been used as a water treatment method (2,3,17). The water surface molecules escape from the surface under natural conditions, and these escaped molecules are collected as pure liquid water. Mechanical evaporators have been used for water recycling. Sometimes, vacuum evaporation has been used for wastewater recycling, and these are operated by steam or electric power. Evaporation is effective for removing inorganic and organic (except volatile organics) contaminants, and it works even at very high concentrations (about 10%) of pollutants. Foaming, scaling, and fouling along with suspended solids and carbonates are the major problems in evaporation because they create maintenance problems. Evaporation applications include treating wastewater from the fertilizer, petroleum, pharmaceutical, and food processing industries. It is also used for water supply to ion exchangers and membrane processes. Water from evaporation has been used for cooling in towers and boilers. It can be used as a technique for water source reduction. The cost of water production varies from 10 to 200 US$ per million liters.

Distillation. In distillation, water is purified by heating it up to 100 °C at which liquid water vaporizes and leaves the pollutants behind (1,18). The vapors generated are cooled to liquid water. The wastewater should be free of volatile impurities. Water produced by this technique is about 99% free from impurities. Various types of boilers with multistage and double distillation are used in this process. The size of the boilers depends on the quantity of water required. Applications of distillation in water treatment and reclamation include water supplies in laboratories, pharmacy, and medicinal preparations. Distillation is effective for preparing potable water from sea and brackish water. The cost of water production varies from 10 to 2000 US$ per million liters.

Biological Technologies

Biological treatment has a reputable place in various water treatment and recycling methods (1–3,19,20). Soluble and insoluble organic pollutants are oxidized by microbes in this process. Water is circulated in a reactor that maintains a high concentration of microbes, and the microbes convert organic matter into water, carbon dioxide, and ammonia. Sometimes, the organic matter is converted into other products such as alcohol, glucose, and nitrate. Wastewater should be free of toxic organics and inorganic pollutants. The maximum concentrations of TDS, heavy metals, cyanides, phenols, and oil should not exceed 16,000, 2.0, 60.0, 140, and 50 mg/L, respectively. Biological treatment includes aerobic and anaerobic digestion of wastewater. The cost of Biological treatment varies from 10 to 200 US$ per million liters of treated water.

Aerobic Process. When air or oxygen in dissolved form is available freely to wastewater, then the biodegradable organic matter undergoes aerobic decomposition, caused by aerobic and facultative bacteria. The extent of the process depends on oxygen availability, retention time, temperature, and the biological activity of the bacteria. The rate of biological oxidation of organic pollutants may be increased by adding chemicals required for bacterial growth. The technique is effective for removing dissolved and suspended volatile and nonvolatile organics. The concentration of biodegradable organics can be reduced up to 90%. Applications include treating industrial wastewater to reduce BOD, COD, nitrogen, and phosphorous. The disadvantage of this method is the production of a large quantity of biosolids, which require further costly management.

Anaerobic Process. If free or dissolved oxygen is not available to wastewater, then anaerobic decomposition called putrefaction occurs. Anaerobic and facultative bacteria convert complex organic matter into simpler organic compounds of nitrogen, carbon, and sulfur. The important gases evolved in this process are nitrogen, ammonia, hydrogen sulfide, and methane. The applications of the anaerobic process to organic pollutant digestion are as discussed in the aerobic process.

CONCLUSION

There is a scarcity of safe and good quality water at present, and it will become more problematic in the near future. Wastewater treatment technologies differ from each other in terms of their principles, scope of application, speed, and economy. All techniques of primary and secondary treatment processes have their own importance and are essential. Therefore, no comparison can be done, whereas wastewater recycling technologies (tertiary treatment) can be compared. The feasibility of any water recycling technique at a commercial level depends on the costs of construction, maintenance, and operation. Sludge management is also an important factor in the selection of technology. Adsorption is considered the best and a universal technique, and it can be used to remove a wide variety of pollutants. It is also a rapid process with a very low cost of construction, maintenance, and operation using low cost adsorbents. Reverse osmosis is also used widely because the water quality is good but the costs of construction and maintenance are comparatively high, and it is rarely used in developing countries. Beside these two, other techniques such as micro- and ultrafiltration, ion exchange, electrodialysis, solvent extraction, and distillation are used for specific purposes, but their use is restricted at potable and industrial levels. Electrolysis as a water recycling technique is in its development stage, and hopefully it will be the best technique on a commercial scale in terms of its cost, maintenance, and speed. The advantages of this technique are a wide range of applications.

BIBLIOGRAPHY

1. Droste, R.L. (1997). *Theory and Practice of Water and Wastewater Treatment*. John Wiley & Sons, New York.
2. Nemerow, N. and Dasgupta, A. (1991). *Industrial and Hazardous Waste Treatment*. Van Nostrand Reinhold, New York.
3. Zinkus, G.A., Byers, W.D., and Doerr, W.W. (1998). Identify appropriate water reclamation technologies. *Chem. Eng. Prog.* **94**(5): 19–31.
4. Williams, A.R. (1991). The use of reverse osmosis for the purification of coal gasification liquors. In: *Effective Industrial Membrane Processes: Benefits and Opportunities*. M.K. Turner (Ed.). Elsevier Science, Barking Essex.
5. Sun, L., Perdue, E.M., and McCarthy, J.F. (1995). Using reverse osmosis to obtain organic matter from surface and ground water. *Water Res.* **29**: 1471–1477.

6. Culkin, B., Plotkin, A., and Monroe, M. (1998). Solve membrane fouling problems with high shear filtration. *Chem. Eng. Prog.* **94**(1): 29–33.
7. vander Ham, F., Witkamp, G.J., deGrauw, J., and van Rosmalen, G.M. (1998). Eutectic freeze crystallization: Application to process streams and wastewater purification. *Chem. Eng. Process* **37**(2): 207–213.
8. Sinev, I.O., Sinev, O.P., and Linevich, S.N. (1997). Apparatus of flotation treatment of natural waters and wastewater. *Izobreteniya* **26**: 369–370.
9. Gupta, V.K. and Ali, I. (2002). Adsorbents for water treatment: Low cost alternatives to carbon. *Encyclopedia of Surface and Colloid Science*. Vol. 1. A. Hubbard (Ed.). Marcel Dekker, New York, pp. 136–166.
10. Nagasaki, Y. (1998). Treatment of Organic Wastewaters by coagulation and precipitation. Jpn.Kokai Tokyo Jp. 1057967 [9857967] (Cl. C02F1/52).
11. Latifossglu, A.A., Surucu, G., and Evirgen, M. (1997). Improvement to the dewaterability of ferric sludge produced from chemical treatment of wastewaters. *Water Pollut. IV: Modelling Meas. Predicting, 4th Int. Conf.* pp. 733–742.
12. Bigda, R.J. (1995). Consider Fenton's chemistry for wastewater treatment. *Chem. Eng. Prog.* **89**: 62–66.
13. McNulty, J.T. (1984). Anion exchange resin kinetics in mixed bed condensate polishing. In: *Ion Exchange Technology*. D. Naden and M. Streat (Eds.). Ellis Norwood, Chichester.
14. Ahn, J.W. and Ahn, J.G. (1997). Solvent extraction for the treatment of industrial wastes. *Chawn Risaikring* **6**(4): 48–54.
15. Gottberg, V., Antonia, J.M., and Siwak, L.R. (1998). Electrodialysis reversal process. *Int. Desalination Water Reuse Q.* **7**(4): 33–37.
16. Coin, R.J., Niksa, M.J., and Elyanow, D.I. (1996). Wastewater treatment by electrochemistry. *Environ. Prog.* **15**: 122–127.
17. Matoba, T. and Furuishi, H. (1998). Sea Water Desalination System. Jpn. Kokai Tokyo Jp. 1052683 [9852683] (Cl. C02F1/04).
18. Bom, P.R. (1998). Method for desalination containing water, single effect or multiple - Effect distillation apparatus and modular element suitable for a single effect or multiple effect distillation apparatus. PCT Int. Appl. WO 9825679 (Cl. B01D1/26).
19. Clark, T. and Stephenson, T. (1998). The effect of chemical adsorption on aerobic biological treatment of municipal wastewater. *Environ. Technol.* **19**(6): 579–590.
20. Kato, M.T., Field, J.A., and Lettinga, G. (1997). The anaerobic treatment of low strength wastewater in UASB and EGSB reactors. *Water Sci. Technol.* **36**: 375–382.

WASTEWATER TREATMENT PROCESSES AND WATER REUSE

Kevin S. McLeary
Pennsylvania Dept. of
Environmental Protection
Harrisburg, Pennsylvania

TREATMENT PROCESSES

The selection of wastewater treatment processes depends on the influent characteristics, the quality of effluent required, and the cost. Treatment can be divided into these stages: preliminary, primary, secondary, and advanced treatment.

Preliminary Treatment Processes

The objective of preliminary treatment is to prepare wastewater for further treatment by removing large objects and grit from the wastewater. These materials may otherwise impede the efficiency or increase the maintenance of downstream processes. Typical preliminary treatment may include screening, grit removal, comminution, and flow equalization.

Screening. Screening removes large objects and trash that could interfere with, clog, or damage downstream equipment, such as pumps, valves, actuators, mechanical aerators, and biological filters. Bar screens typically consist of inclined steel bars spaced at equal intervals in a channel through which the wastewater flows. Large solids are removed from the wastewater by the bars, which must be cleaned either manually or mechanically. Due to their putrescible nature, screenings are usually landfilled. When designing bar screens, bar size, bar spacing, angle of inclination, and wastewater approach velocity must be considered (1).

Manually cleaned screens usually have bar spacings ranging from 1 to 2 inches, and the bars are inclined from 30° to 45° to the horizontal (1). The screens are cleaned by raking the screenings onto a sieve plate to allow for drainage.

Mechanically cleaned screens typically have bar spacings ranging from 0.5 to 1.5 inches and are inclined from 0° to 30° from the vertical (1). The screen automatically starts the cleaning process after a preprogrammed operating time but also has a high-level override.

Grit Removal. Grit consists of sand, gravel, and other non- or minimally putrescible material that may wear mechanical equipment or accumulate in process tanks. Grit has a high settling velocity and settles quickly (1). Grit chambers are frequently aerated, which keeps the light organic materials in suspension and allows the heavy grit particles to settle to the bottom of the tank to be collected (1).

Comminution. Comminutors shred materials without removing them from the wastewater. Comminutors have a bypass channel and a screening device that may be used if the comminutor must be taken out of service (1). The basic parts of a comminutor include a screen and cutting teeth on a revolving drum. The drum has slots cut into it. The teeth shred the material as it is trapped against the screen. The shredded solids then pass through the drum slots and out of the bottom opening through an inverted siphon (2). Shredded materials cause problems in downstream process units, so comminutors are not frequently installed in new wastewater treatment plants (1).

Equalization. Equalization is used in plants that have significant variations in hydraulic or organic loading. It provides relatively uniform loading throughout the day. Uniform loading increases the efficiency and reliability of downstream process units. In addition, the downstream treatment units no longer have to treat highly variable influent, so the sizes of the units may be decreased (1).

Equalization tanks may be designed either as in-line or side-line units. The sizing should be based on an influent flow hydrograph during wet weather. To prevent detained wastewater from becoming septic, a dissolved oxygen concentration of 1 mg/L must be maintained in the tank (3). Equalization tanks should be mixed to prevent solids from settling.

Primary Treatment Processes

Sedimentation. Primary sedimentation is the oldest and most widely used operation in treating wastewater (1). It removes solids to produce an effluent that is suitable for downstream biological treatment.

Solids removal is governed by retention time and surface settling rate. The surface settling rate is the volumetric flow rate over the surface area of the clarifier. Particles whose settling velocity is greater than the surface settling rate are removed from the wastewater stream. However, if the detention time is too long, the wastewater turns septic.

Clarifier design is based on the average design flow and the peak hourly flow. Flows are generally limited to 1000 gallons per day/ft^2 for design flow and 1500 gpd/ft^2 at peak hourly flow (3). The calculated settling area is based on both flows, and the larger clarifier is selected.

Secondary Treatment Processes

Activated Sludge Processes. Activated sludge is a biological process that uses microorganisms to treat wastewater in an aerobic environment. Wastewater is introduced into the reactor tank where it is mixed with microorganisms that are returned from the clarifier. This mixture of raw wastewater and return activated sludge is called mixed liquor. The microorganisms are then allowed to flocculate and settle under quiescent conditions in a clarifier; treated wastewater flows over weirs for further treatment or discharge. There are several variations on the conventional activated sludge process, including plug flow, step feed, tapered aeration, complete mix, contact stabilization, sequencing batch reactor, and extended aeration.

Plug Flow. A plug flow reactor is configured so that the wastewater flows through a long, narrow channel during treatment. It approximates flow through a pipe. In an ideal plug flow reactor, there is no longitudinal mixing of wastewater.

Step Feed. Step feed is a variation of the plug flow reactor. In this variation, the wastewater influent is introduced into the reactor at several points, rather than adding the entire influent stream at the beginning of the reactor. It provides equalization of the load over the reactor volume.

Tapered Aeration. Tapered aeration is another variation on the plug flow reactor. The reactor influent is introduced at the front end of the reactor, but the aeration pattern is different. Approximately 50% of the aeration capacity is in the first third of the reactor, where the oxygen demand is the highest, and less is supplied further down the reactor, where the demand is less.

Complete Mix. A complete mix reactor is the opposite of a plug flow reactor. All of the wastewater is completely mixed throughout the reactor. To facilitate mixing, complete mix reactors are shorter and wider than plug flow reactors. Due to the rapid mixing of the contents, complete mix reactors can tolerate shock loads better than plug flow reactors.

Contact Stabilization. Contact stabilization is a process where treatment takes place in two tanks. The wastewater is introduced into the contact tank with microorganisms. The organics are not degraded in the contact tank but are absorbed into the microbial cells. After clarification, the settled sludge is returned to a reaeration tank, where the microorganisms then degrade the organics that were absorbed during the contact process.

Sequencing Batch Reactor. A sequencing batch reactor is a variation of the complete mix reactor. Both wastewater stabilization and settling take place in the same tank in sequenced phases. Wastewater is introduced into the tank (fill phase), where it is mixed and aerated. Following a "react phase," both mixing and aeration are turned off, and the wastewater is allowed to settle quiescently (settle phase). The treated wastewater is then decanted from the top of the reactor (decant phase), and the sequence begins again.

Extended Aeration. The extended aeration process is a suspended growth system that operates in the endogenous respiration phase of microbial growth. This process requires low organic loading rates, long aeration times, and high mixed liquor concentrations. This process is commonly used in prefabricated package treatment plants for small communities.

Aeration Requirements. The dissolved oxygen concentration in the aeration tanks should be greater than 2 mg/L at all times. To supply sufficient air for treatment, all of the treatment processes except extended aeration should be designed to provide 1.1 lb of oxygen per pound of biochemical oxygen demand (BOD) oxidized. Because of the additional aeration needs of endogenous respiration, extended aeration systems require 1.5 lb of oxygen per pound of BOD oxidized. This does not include the air required for nitrification (3).

Fixed Film Processes. Fixed film processes are biological treatment processes using microorganisms attached to a support medium. As the wastewater flows over the medium, the microorganisms absorb the organics. The biofilm builds up, as wastewater is treated, and eventually sloughs off. Like activated sludge, the sloughed microorganisms are then allowed to settle under quiescent conditions in a clarifier.

In a trickling filter, the wastewater flows through a rotating arm, which distributes it over the top of the medium. To maintain proper liquid flow for efficient operation of the filter, a portion of treated wastewater must be recirculated back to the filter.

In a rotating biological contactor, microorganisms are attached to a plastic disk, which is rotated through the wastewater. When the microorganisms are submerged, they absorb organics. During the time the microorganisms are exposed to the air, they receive the oxygen that they require for treatment. Unlike trickling filters, however, no recirculation is required.

Stabilization Ponds. Stabilization ponds are typically large, lined basins, which may be aerobic, anaerobic, or facultative. Ponds are designed around a detention time measured in days, rather than hours, and are relatively shallow compared with other biological treatment processes. These design criteria mean that a large land area is required for ponds, and they are usually used only in small communities. Their advantages include low construction and operating costs.

Advanced Treatment Processes

Nitrification. Ammonia nitrogen is converted to nitrate in a two-step process. Ammonia is first converted to nitrites, and the nitrites are then converted to nitrates. This conversion is oxygen intensive. For each milligram of ammonia converted to nitrate, 4.6 mg of oxygen is required. In addition, each milligram of ammonia converted consumes 7.14 mg of alkalinity (4).

Nitrification may occur in the same tank as carbon oxidation in a single sludge process, or it may take place in a separate nitrification tank. Because nitrifying organisms have a slower growth rate than the organisms for carbon oxidation, the process requires longer detention time and longer mean cell retention time.

Biological Phosphorus Removal. Phosphorus removal can be enhanced in a biological system by first creating an anaerobic zone followed by an aerobic zone. In biological phosphorus removal, from 2.5 to 4 times more phosphorus can be removed than in a secondary treatment process (4).

To generate energy for cell growth in the anaerobic stage, phosphorus is released from the internal polyphosphates of the cell, resulting in an increase in the liquid phosphorus concentration (3). In the aerobic zone, there is a rapid uptake of the soluble phosphorus for the resynthesis of intracellular polyphosphates. More phosphorus is absorbed by the cells than was released in the anaerobic zone.

There are three major biological phosphorus removal methodologies: the Anaerobic/Oxic (A/O) process, the PhoStrip process, and the sequencing batch reactor. The A/O process is proprietary, and phosphorus removal depends on the influent ratio of BOD to P. The PhoStrip process is also proprietary. Phosphorus removal does not depend on the BOD:P ratio, but chemicals must be used to precipitate the phosphorus (4). The SBR can be designed to provide anaerobic conditions during the treatment cycle, which release phosphorus. When the reactor is then aerated, the phosphorus is absorbed from the wastewater and is incorporated into the biomass.

Denitrification. Denitrification is the removal of nitrogen from wastewater. In an anoxic environment, several species of bacteria can use nitrates, rather than oxygen, as their energy source. Denitrification converts the nitrates into nitrogen gas and additional biomass (4). The process requires a carbon source for completion. In wastewater treatment, it is common to use the wastewater itself for the carbon supply. The raw wastewater flows into an anoxic zone with return sludge and a large mixed liquor recycle. The recycle ratio is determined by the ammonia concentration and the required effluent nitrate concentration. The anoxic zone then denitrifies by using the nitrates created in the mixed liquor. Following the anoxic zone, the wastewater flows to an aerobic zone to strip nitrogen gas. The process may be repeated for additional nitrogen removal. Denitrification is normally done in a plug flow type system, an oxidation ditch, or a sequencing batch reactor (4).

Biological Dual-Nutrient Removal. Biological dual-nutrient removal is the reduction of both nitrogen and phosphorus in wastewater by biological methods. Biological dual-nutrient removal is achieved through several proprietary treatment processes, including the A^2O process, the Bardenpho process, the University of Capetown (UCT) process, and the Virginia Initiative (VIP) process (4). These processes use the aerobic process for carbon oxidation, the anoxic process for denitrification, and the anaerobic process for biological phosphorus removal, although arrangement of the processes varies. The UCT process and the VIP process are further complicated by the use of internal recycle streams.

Air Stripping. Air stripping is a method of removing volatile compounds from a solution. Air is introduced at the bottom of a packed tower. Wastewater flows down the tower from the top and contacts the air countercurrently. The driving force in air stripping is the concentration difference between the air and the wastewater. The tower medium may become fouled, resulting in high operating and maintenance costs.

Coagulation/Sedimentation. Coagulation/sedimentation uses chemicals to enhance the sedimentation of solids, precipitate pollutants, or remove phosphorus. The chemicals most commonly used in the coagulation/sedimentation process are lime, alum, iron salts, and polymers (2). Coagulation involves destabilizing colloidal particles through any of several processes, including double layer compression, charge neutralization, enmeshment, or interparticle bridging (5). The particles then aggregate and settle out.

Alum is typically used in the chemical removal of phosphorus, although iron salts may also be used. Phosphorus removal occurs by the formation of an insoluble precipitate of aluminum or iron phosphate. Alum and iron also react with hydroxyl radicals in the water, forming hydroxides in addition to phosphates.

Filtration. Filtration is the removal of wastewater solids by passing the wastewater through granular media. Some of the media that have been used include sand, anthracite coal, diatomaceous earth, perlite, and granular activated carbon. Sand filters are the most commonly used filters in wastewater treatment, although filters can also consist of multiple types of media, such as coal over sand or coal over silica sand over garnet sand (5).

Particles may be removed by interstitial straining. However, smaller particles must be transported to the surfaces of the media, where an attachment mechanism retains the particles. Transport mechanisms may include gravitation, diffusion, and interception. These processes depend on the physical characteristics of the media. The attachment mechanism may include electrostatic attraction, chemical bridging, or adsorption. These processes are functions of the coagulant and the chemical characteristics of the wastewater and media (5).

Filters are classified as slow filters, rapid filters, or pressure filters. Slow filters require a buildup of solids on the top surface of the filter through which the wastewater must pass, which requires a low application rate. This buildup strains particles from the wastewater. Rapid filters and pressure filters use the entire depth of the media and may be operated at higher loadings than slow filters.

Activated Carbon Adsorption. Adsorption is a process where molecules of a compound adhere to a solid surface. The most commonly used adsorbent in wastewater treatment is activated carbon. Activated carbon comes in two forms, powdered and granular. Powdered activated carbon (PAC) is added to the mixed liquor in the aeration tanks and is removed from the wastewater by settling. Granular activated carbon (GAC) is used in a packed bed (2).

The adsorptive capacity of the carbon is a function of the material and method used to create the activated carbon as well as the chemical properties of the compound to be adsorbed. In general, organics are completely removed until the adsorptive capacity is exhausted. At this point, the effluent concentrations increase (2). Spent activated carbon may be regenerated by heating.

Membrane Systems. Membrane processes use a semipermeable barrier that allows the water to flow through but retains the contaminants. There are several types of membrane systems in wastewater treatment, including reverse osmosis, nanofiltration, microfiltration, and ultrafiltration. All of these processes use pressure to force water through the membrane.

Ultrafiltration may be used to remove molecules that have a molecular weight of 500 or greater and have a low osmotic pressure at moderate concentrations. This includes bacteria, viruses, proteins, and clays (5). Reverse osmosis is used to separate small molecules whose osmotic pressure is high. Microfiltration and nanofiltration are membrane systems that lie between ultrafiltration and reverse osmosis.

Membrane processes are subject to fouling of membranes. These processes should be pilot tested to determine which process and membrane work best for any given application.

Disinfection Processes

Chlorination/Dechlorination. Chlorine has been used as a disinfectant for many reasons, including inactivation of a wide range of pathogens, maintenance of a residual, and cost. As chlorine dissolves in water, it forms hypochlorous acid, which dissociates into hypochlorite ions and hydrogen ions and decrease the pH. Lower pH values cause less dissociation, which is preferable, because hypochlorous acid is a much more effective disinfectant than hypochlorite. Sodium and calcium hypochlorites also form hypochlorous acid when dissolved, but they also liberate hydroxyl ions, and thus increase the pH of the wastewater. Chlorine is toxic, so dechlorination may be required, which is usually done by using sulfur dioxide to reduce the chlorine to chlorides. Sodium metabisulfite or sodium bisulfite may be used instead of sulfur dioxide in small facilities. The reactions are nearly instantaneous, and detention times are less than 2 minutes.

Ozonation. Ozone is a powerful oxidant that can disinfect wastewater using less contact time and lower dosages than other chemical methods. It has high germicidal efficiency against a wide range of organisms, and it does not leave a residual (6). Because of its instability, ozone must be generated on-site. Ozone is applied to wastewater in closed contactors. The off-gas from the contactors contains high concentrations of ozone, which must be destroyed before it is discharged to the atmosphere.

Ultraviolet Light Disinfection. Ultraviolet light is a form of electromagnetic radiation at wavelengths of 100–400 nm. Electromagnetic radiation at wavelengths from 240–280 nm inactivates microorganisms by damaging their nucleic acid (6). Ultraviolet lamps operate in much the same way as fluorescent lamps. The UV radiation is generated by passing a current through mercury vapor. The mercury lamps may be low-pressure or medium-pressure lamps. Low-pressure lamps emit most of their energy at a wavelength of 253.7 nm, which is in the optimal range. Medium-pressure lamps generate a lesser portion of their energy in the optimal range, but the intensity of the radiation is much greater than that of the low-pressure lamps, and fewer lamps are required (6).

REUSE

Types of Reuse Applications

Urban Reuse. Reclaimed wastewater may be used in an urban setting to irrigate public parks, recreation areas, and residential landscaping (7). Some water codes even consider the use of potable water for landscape irrigation as “a waste or an unreasonable use of such water” when reclaimed water is available (2). Other urban uses include fire protection, dust control for construction activities, and concrete production. The water for urban reuse activities is obtained from a dual distribution system (one for potable water and one for reclaimed water) (7). Reclaimed wastewater may also be used inside buildings for toilet flushing (2).

Industrial Reuse. Reclaimed water can be used in industrial activities, for cooling water (once-through or recycled), industrial process water (tanning, textiles, pulp and paper manufacturing), and boiler feed water. Recycled cooling water must be of higher quality than once-through cooling water because evaporation concentrates pollutants. Industrial process water quality depends on the quality of the product. Higher quality products require higher quality process water. High-pressure boilers require nearly pure water, whereas lower pressure boilers may be able to operate with less pure water (7). Potential problems in the industrial reuse of wastewater include mineral scaling and biological growth (4).

Agricultural Irrigation. Reclaimed water includes nutrients, such as nitrogen and phosphorus, which are used by plants. Reclaimed water also includes valuable trace elements that are vital to plant growth at low concentrations. Salinity is the most important factor in agricultural irrigation (7). As soil salinity increases, plants expend more energy to adjust the salt concentration within their cells, and less energy is available for plant growth. Soil salinity increases as a function of evapotranspiration (4). The ability of plants to tolerate salinity varies widely, from sensitive plants such as citrus fruits and berries to tolerant crops, such as barley and cotton. Salt is especially detrimental to plants at the germinating and seedling stages. Salinity may be decreased by leaching or the overapplication of wastewater to carry away excess accumulated salt (7).

Habitat Restoration/Enhancement and Recreational Reuse. These uses typically include wetlands, recreational impoundments, stream augmentation, and snowmaking. Wetlands serve valuable functions, such as flood attenuation, wildlife and waterfowl habitat, aquifer recharge, and water quality enhancement. Recreational impoundments can serve noncontact uses, such as water hazards on golf courses; human contact activities, such as fishing, boating, and swimming; and the manufacture of snow for skiing. The quality of reclaimed water for contact activities must be higher than that for noncontact uses (7).

Groundwater Recharge. Groundwater can be recharged by surface irrigation or subsurface injection. Subsurface injection may require a higher degree of treatment because water does not receive the benefit of additional treatment by surface soils. Groundwater recharge applications are often used to establish barriers to saltwater intrusion in coastal areas, to provide additional treatment for later reuse, to augment aquifer levels, to provide water storage, and to prevent or control the subsidence of soils (7). Other types of reuse, such as agricultural or landscape irrigation may also provide additional benefits by recharging groundwater (4).

Table 1. Suggested Treatment

Type of Use	Recommended Degree of Treatment
Urban	
Landscape irrigation, fire protection	Secondary, filtration, disinfection
Restricted access irrigation	Secondary, disinfection
Construction	Secondary, disinfection
Industrial Reuse	
Industrial cooling, once-through	Secondary
Industrial cooling, recirculated	Secondary, disinfection
Agricultural Irrigation	
Food products, not commercially processed	Secondary, filtration, disinfection
Food products, commercially processed	Secondary, disinfection
Nonfood products	Secondary, disinfection
Habitat Restoration/Recreational	
Recreational impoundments	Secondary, filtration, disinfection
Landscape impoundments	Secondary, disinfection
Environmental (wetlands, stream augmentation)	Secondary, disinfection
Groundwater Recharge	
Surface irrigation	Primary
Subsurface injection	Secondary
Augmentation of Potable Supplies	
Indirect reuse, aquifer spreading	Secondary, disinfection
Indirect reuse, aquifer injection	Secondary, filtration, disinfection, advanced
Indirect reuse, surface augmentation	Secondary, filtration, disinfection, advanced

Augmentation of Potable Supplies. Potable supplies may be augmented by either direct or indirect methods. Direct potable reuse is objectionable and is used only in one installation in the world, in Windhoek, Namibia, and there it is used only intermittently. Indirect reuse may involve the discharge of treated effluent into a body of water upstream of a water intake. The effluent may then be further treated by natural processes and by dilution. Indirect use may also involve applying effluent to groundwater by surface irrigation or subsurface injection, where the wastewater is further treated by the soil and aquifer before being withdrawn for potable uses (7).

Treated Effluent Characteristics. The quality of the effluent required depends on the use of the wastewater. For stream discharges or applications that are not being used for drinking water or in public contact areas, secondary treatment may be all that is required. However, for more sensitive uses, such as water that will eventually be used for potable water supplies, for high-quality industrial applications, or for uses where there will be significant human contact, advanced treatment may be required.

Note that there are no national treatment standards for reuse applications. Different jurisdictions may have different treatment requirements, depending on use. EPA has recommended degrees of treatment for various applications, as shown in Table 1 (7).

BIBLIOGRAPHY

1. Water Environment Federation and American Society of Civil Engineers. (1992). *Design of Municipal Wastewater Treatment Plants*. WEF Manual of Practice No. 8/ASCE Manual and Report on Engineering Practice No. 76, 3rd Edn.
2. Water Environment Federation. (1989). *Water Reuse*. Manual of Practice SM-3, 2nd Edn. Alexandria, VA.
3. Great Lakes-Upper Mississippi River Board of State and Provincial Public Health and Environmental Managers. (1997). *Recommended Standards for Wastewater Facilities*. Health Education Service, Inc., Albany, NY.
4. Metcalf & Eddy, Inc. (1991). *Wastewater Engineering: Treatment, Disposal, Reuse*, 3rd Edn. McGraw-Hill, Inc., New York.
5. Weber, W. J. (1971). *Physicochemical Processes for Water Quality Control*. John Wiley & Sons, New York.
6. US EPA. (1999). *Alternative Disinfectants and Oxidants Guidance Manual*. EPA 815-R-99-014.
7. US EPA. (1992). *Guidelines for Water Reuse*. EPA/625/R-92/004.

WASTEWATER RECLAMATION AND REUSE RESEARCH

Jorg E. Drewes
Colorado School of Mines
Golden, Colorado

Water reclamation is the treatment of wastewater to make it reusable. Water reuse is the beneficial use of the treated water and can serve either nonpotable or potable applications (1). Nonpotable water reuse includes all water applications other than augmenting drinking water supplies. Nonpotable reuse is occurring as agricultural and landscape irrigation, recreation, wildlife habitat maintenance (stream-flow augmentation), in-building applications (such as toilet flushing), industrial cooling, and groundwater recharge. Potable water reuse refers to the use of highly treated reclaimed water to augment drinking water sources. Direct potable reuse, where reclaimed water after advanced treatment is piped directly into the potable water system, is limited to rare cases. For indirect potable reuse, reclaimed water can either be discharged into a surface water or infiltrated into the subsurface to augment, in part, a drinking water supply source. Infiltration into the subsurface can be accomplished by either direct injection into the aquifer or by surface spreading which provides additional soil–aquifer treatment (SAT) when water percolates through the subsurface. The majority of research activities in wastewater reclamation and reuse reflects the main concerns about potable and nonpotable water reuse, such as (1) protection of public health, (2) reliable treatment of wastewater to meet strict water quality requirements for the intended reuse, and (3) gaining public acceptance (1,2).

WATER REUSE RESEARCH TO PROTECT PUBLIC HEALTH

The majority of research activities in water reuse has been conducted on health effects which can be classified as direct effects, indirect effects, and issues related to the aesthetic quality of reclaimed water (3,4). Figure 1 illustrates health effects related to water reuse.

Direct and indirect health effects that might occur from consumption of pathogens (including enteric viruses, parasites, and enteric bacteria) and organic and inorganic chemicals in reclaimed water could be manifested via short-term exposure and acute effects or through chronic exposure and latent effects. These chronic effects of organic and inorganic constituents or their metabolites are of particular concern to the extent that they relate to incidences of cancer. There is also increasing concern about endocrine-disrupting compounds (EDC) and pollutants originating from pharmaceuticals and active ingredients in personal care products (PPCPs), which are present in municipal wastewater and continually introduced as complex mixtures to the aquatic environment (5,6). EDCs are exogenous agents that interfere with the production, release, transport, metabolism, binding, action, or elimination of the natural hormones in the body responsible for maintaining homeostasis and regulating developmental processes. Among compounds that have endocrine activity are both synthetic chemicals produced industrially (such as surface active agents, pesticides, plasticizers, food additives, birth control pills, herbal supplements, and cosmetics) and naturally occurring compounds (such as sex steroids, plant-produced estrogens, and heavy metals) (7). The occurrence of PPCPs in domestic effluents is reported in studies since the late 1980s (8,9). The major concern for the presence of pharmaceuticals in

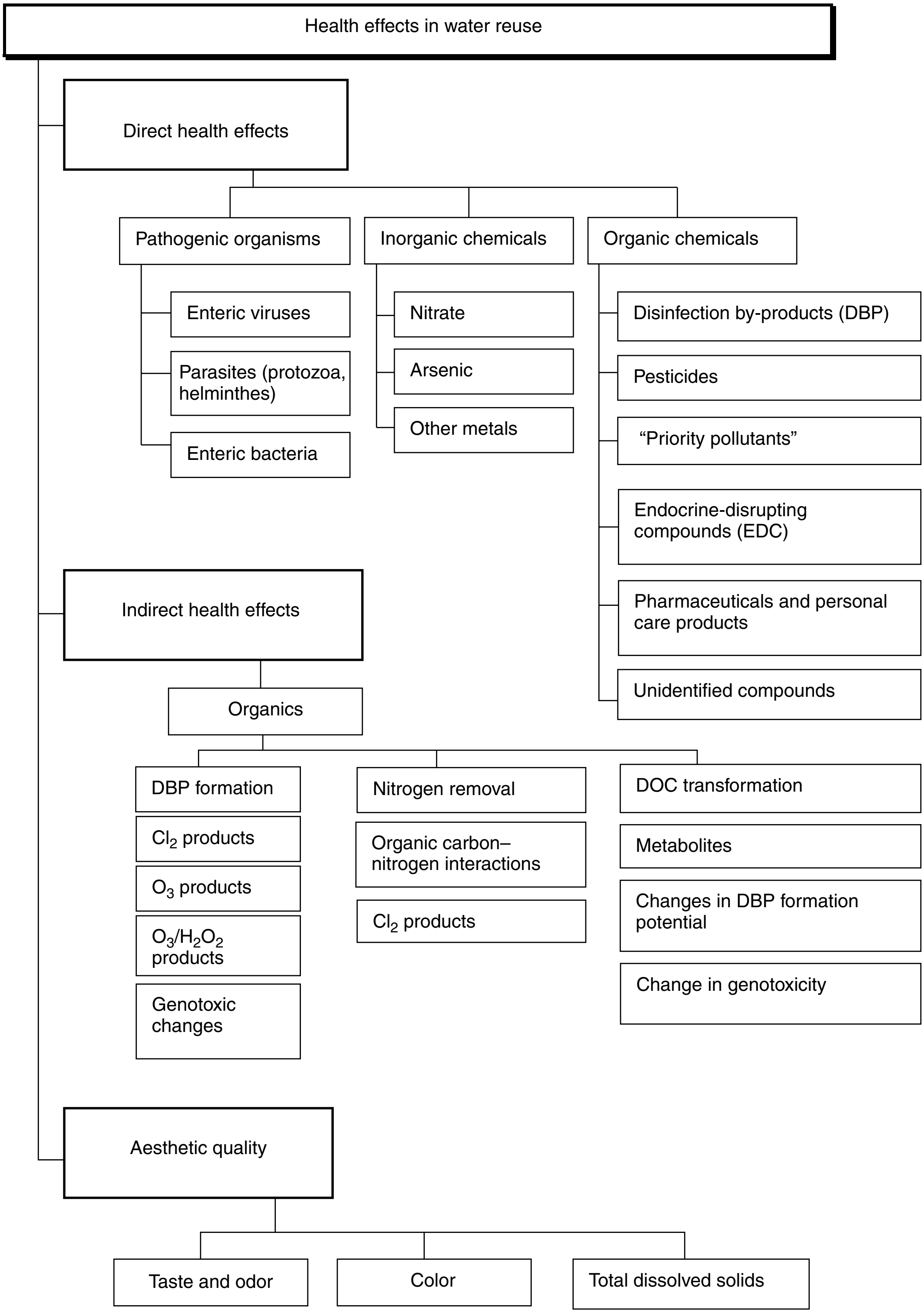

Figure 1. Research activities related to health effects in water reuse.

reclaimed water ultimately used for human consumption is not necessarily acute effects on human health, but rather the manifestation of imperceptible effects that can accumulate over time ultimately to cause truly profound changes (5). Effects might occur over long periods of time in certain populations, so changes would not be distinguishable from natural events or ecological succession. A continuous research effort is directed to unknown and unidentified organic contaminants. The toxicological risk, or safety, of these mixtures may never be known precisely. This presents an ever-present element of uncertainty for all reclaimed water projects involving potable reuse and justifies continuous monitoring and research. The aesthetic quality of reclaimed water is determined by taste and odor, color, and total dissolved solid (TDS) concentration.

Nonpotable Reuse

The main research issues in nonpotable reuse are the fate and transport of pathogens, nitrogen, and TDS during water reclamation and reuse. Today, these issues do not present a significant barrier to implementing nonpotable reuse projects, and direct nonpotable reuse today is the dominant planned reuse venue for supplementing public water supplies worldwide. Table 1 summarizes key research studies on pilot and full scale in nonpotable reuse and lists their main research objectives, locations, and references. This selection of important research studies is not complete nor does it indicate any type of ranking regarding their scientific and technical merits, but it directs the reader to important research in the field.

Enteric microorganisms are the contaminants of greatest concern in reusing reclaimed wastewater. Pathogen removal has been addressed in several research studies that focus on the efficiency of different pretreatment technologies for virus inactivation during subsequent disinfection, the survival of enteric microorganisms on food crops that have been irrigated with reclaimed water, as well as disease outbreaks from consuming contaminated foods, and the fate and transport of microorganisms in reclaimed water during artificial groundwater recharge. Recent research advances also resulted in better methods of detecting pathogens. Findings of these research studies suggest that tertiary treatment (such as coagulation/flocculation, lime clarification, dual-media filtration) and disinfection are necessary for many reuse applications to ensure pathogen reduction. Current research activities on pathogen removal address treatment plant reliability, removal of new and emerging enteric pathogens of concern, and the ability of new technologies to effect pathogen reduction (20).

For water reuse, the nonmetals ammonia, nitrite, and nitrate are of particular health significance because nitrate concentrations significantly higher than 10 mg N/L can cause methemoglobinemia in infants. Research is directed to aboveground treatment technologies and the efficiency of soil–aquifer treatment systems in removing nitrogen from reclaimed water. In the sustainable operation of local water reuse systems, the total solids concentration represents a significant challenge to the water reuse research community to develop strategies for minimizing salt buildup in local water cycles and cost-efficient treatment and disposal techniques.

Potable Reuse

Potable reuse involves a broad spectrum of potential health concerns. The major concerns are that adverse health effects could result from introducing pathogens, toxic organic chemicals, or inorganic constituents into groundwater that is consumed by the public (21). Treatment efficiency and reliability in removing pathogens are very important in potable reuse where even short-term exposure to pathogens could result in significant risk to the exposed population (22). Enteric viruses were not found in reclaimed water at field sites employing tertiary treatment but were observed in reclaimed water that had secondary treatment followed by chlorination (16). Table 2 summarizes key research studies at pilot- and full-scale potable water reuse facilities.

An assessment of health risks from potable reuse cannot be considered definitive because of limited chemical and toxicological data and inherent limitations in available toxicological and epidemiological methods. The mix of contaminants in the watershed of a reclamation facility might vary from site to site, so health effects studies

Table 1. Key Research Studies on Pilot and Full Scale in Nonpotable Water Reuse

Research Objective	Type of Water Reuse	Location	Research Study	Reference
Behavior of nitrogen and organic contaminants	Groundwater recharge	Phoenix, Arizona	Flushing Meadows project	10, 11
Virus inactivation	Irrigation Groundwater recharge Stream-flow augmentation	County Sanitation Districts of Los Angeles County, California	Pomona Virus Study evaluated different pretreatment processes for subsequent inactivation of viruses	12, 13
Virus inactivation	Storage and recovery—irrigation	Tucson, Arizona	Virus inactivation study during SAT	14–16
Efficiency of soil–aquifer treatment (SAT) in removing of pathogens and nutrients	Irrigation	Dan Region, Israel	Efficiency of SAT	17
Irrigation with reclaimed water	Irrigation	St. Petersburg, Florida	Treatment, storage, and distribution of reclaimed water for irrigation	18
Implementing nonpotable reuse	Urban and agricultural irrigation Industrial reuse Toilet flushing	Irvine Ranch Water District, California	Water quality and water distribution studies	19

Table 2. Key Research Studies on Pilot and Full Scale in Potable Water Reuse

Research Objective	Type of Water Reuse	Location	Research Study	Reference
Health effects of direct reuse	Direct potable reuse	Denver, Colorado	Direct potable reuse demonstration project	23
Health effects of chemicals and microorganisms	Groundwater recharge by direct injection into a potable aquifer	Palo Alto, California	Mutagenic activity and chemical characterization during groundwater injection	24
		Orange County, California	Water Factory 21	25, 26
Health effects of chemicals and microorganisms	Groundwater recharge by surface spreading into a potable aquifer	Montebello Forebay County Sanitation Districts of Los Angeles County, California	Health effects study Epidemiological assessment Birth outcomes study	27–29
Sustainability of soil–aquifer treatment (SAT)			Sustainability of organic carbon, nitrogen, and virus removal during SAT	16
Health effects of chemicals and microorganisms	Augmentation of surface water supplies	San Diego, California	Aqua III demonstration project	30
		Tampa, Florida	Tampa recovery project	31
Fate of trace organic compounds during advanced water treatment	Groundwater recharge by direct injection into a potable aquifer	West and Central Basin Municipal Water District, California	Removal of unidentified trace organics	32

are applicable only to a specific location (4). In findings from research studies conducted at full-scale indirect potable reuse facilities, there is no indication that health risks from using reclaimed water are greater than those from using existing water supplies or that concentrations of regulated chemicals and microorganisms in product water exceeded established drinking water standards set by the U.S. EPA (23,27–29,31). The limited data and extrapolation methodologies used in toxicological assessment may present limitations and uncertainties in overall risk characterization. Similarly, epidemiological studies may suffer from the fact that many cancers have latency periods of 15 years or more. In addition, these types of studies require large populations to uncover the generally low risks from low concentrations of potential carcinogens.

The scope of inorganic chemical health issues in potable water has not changed significantly in several decades. In general, the health hazards from inorganics are well established. However, the threshold for risk protection has changed because of new information for some inorganics such as lead and arsenic. The fact that most forms of nitrogen can be readily converted to nitrate ions either in the treatment process or in soils makes managing nitrogen a significant issue in potable reuse projects.

WATER REUSE RESEARCH IN WASTEWATER TREATMENT TECHNOLOGIES

Research in water reuse has contributed to significant progress in developing sound technical approaches to produce a high quality and reliable water source from reclaimed wastewater (2,33). The evolution of water reclamation technology is illustrated in Fig. 2 for water treatment processes that lead to indirect potable reuse.

Findings of research studies and the advent of new water treatment technologies have resulted in an evolution of water reclamation technologies since the first water reclamation facilities for indirect potable reuse were established in the early 1960s (Fig. 2a). Primary treatment and secondary treatment are still the basic treatment in current water reclamation facilities for removing particulate matter and degradable organic material, respectively. In the early 1970s, treatment trains were developed to improve the removal of pathogens and trace organics by adding lime clarification, granular activated carbon, and reverse osmosis (Fig. 2b,d). Since the late 1990s, the use of microporous membranes in ultrafiltration (UF) and microfiltration (MF) as a pretreatment in reverse osmosis has become the industry standard for indirect potable reuse applications (Fig. 2c,e). Disinfection, conducted over decades using chlorine, is critically reviewed by many utilities and more and more substituted by disinfection using ultraviolet light. With the proper dosage and installation, UV radiation is a proven and effective bactericide and viricide for wastewater and does not contribute to the formation of known toxic by-products (Fig. 2e). This dynamic in incorporating new unit operations in water reclamation is proceeding with the advent of membrane-coupled bioreactors that can substitute for conventional primary and secondary treatment, including tertiary filtration. Membrane bioreactors (MBR) consist of a biological reactor where suspended biomass and solids are separated by microfiltration membranes whose nominal pore sizes range from 0.1 to 0.4 μm. MBRs can operate at much higher mixed liquor suspended solids (MLSS) concentrations (15,000 to 25,000 mg/L) than conventional activated sludge processes. These (**a**) Montebello Forebay surface spreading grounds, County Sanitation Districts of Los Angeles County, California; continuous, started in 1962

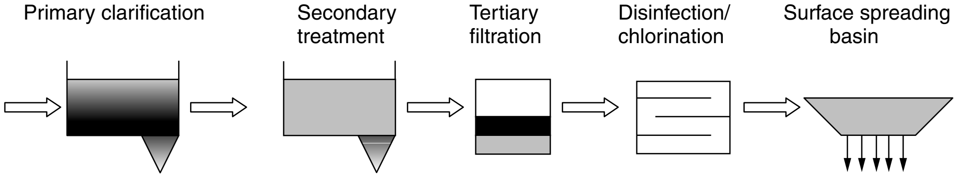

(**b**) Denver potable water demonstration project, City of Denver, Colorado; initiated in 1974, completed in 1990

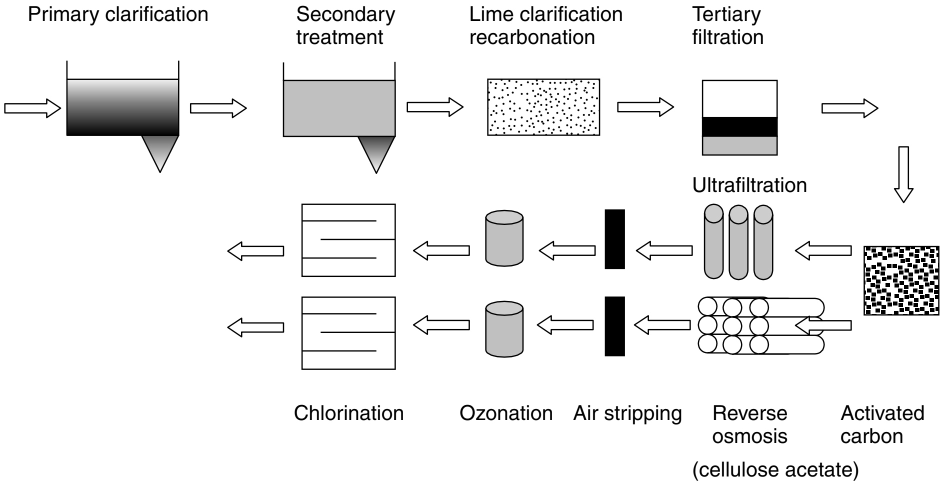

(**c**) Scottsdale Water Campus, City of Scottsdale, Arizona; continuous, started in 1999

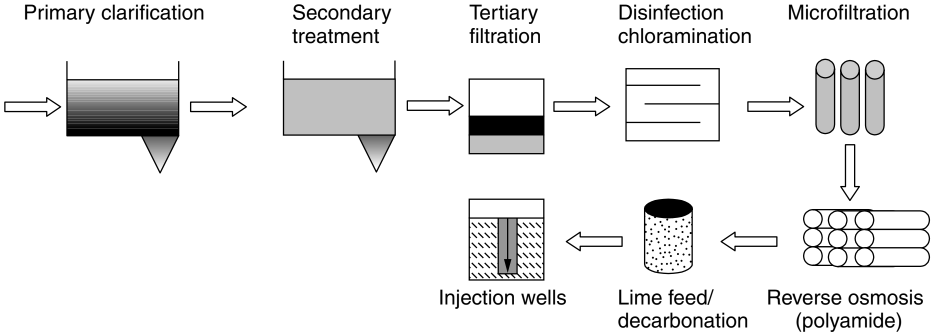

Figure 2. Evolution of water reclamation process trains leading to indirect potable reuse.

advantages are enhanced by a small footprint and high effluent quality with respect to turbidity, total suspended solids (TSS), biochemical oxygen demand (BOD), and pathogens (2).

WATER REUSE RESEARCH RELATED TO PUBLIC PERCEPTION

The ultimate success of any water reuse program is determined by its level of public acceptance. Several indirect potable reuse projects have failed to emphasize public support and resulted in project termination, although the technical and scientific evaluation was well done. Gaining public acceptance has and will initiate research programs related to human health and public involvement and education (34).

RESEARCH NEEDS IN THE FUTURE

Water quality requirements for nonpotable reuse are quite tractable, and treatment requirements are not likely to change significantly in the future, but drinking water quality standards are likely to become more rigorous. The number of contaminants to be monitored is increasing, and, for many contaminants, the maximum contaminant levels and action levels are decreasing. Thus, it is appropriate to anticipate more restrictive treatment requirements and water quality limits and (**d**) Water Factory 21, Orange County Water District; 1977

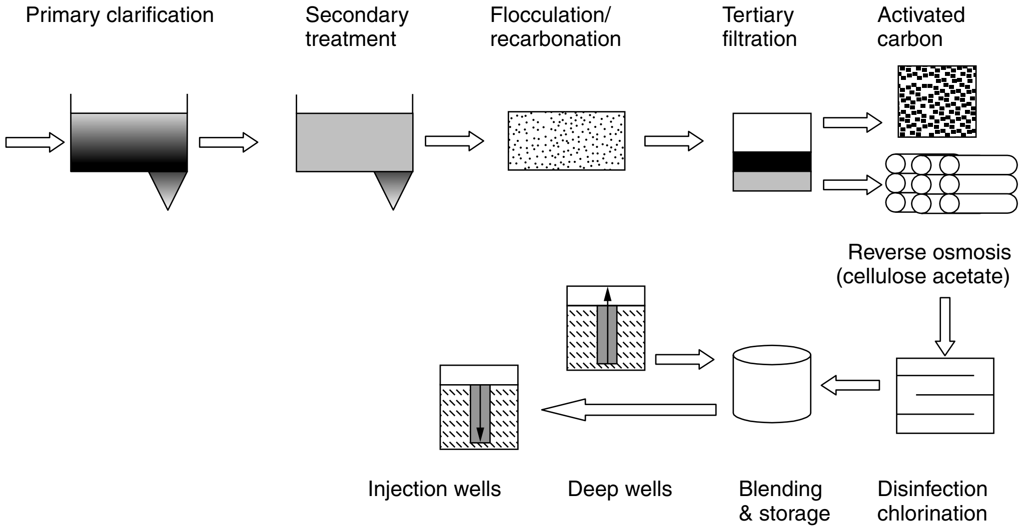

(**e**) Water Factory 21, Orange County Water District; 2002

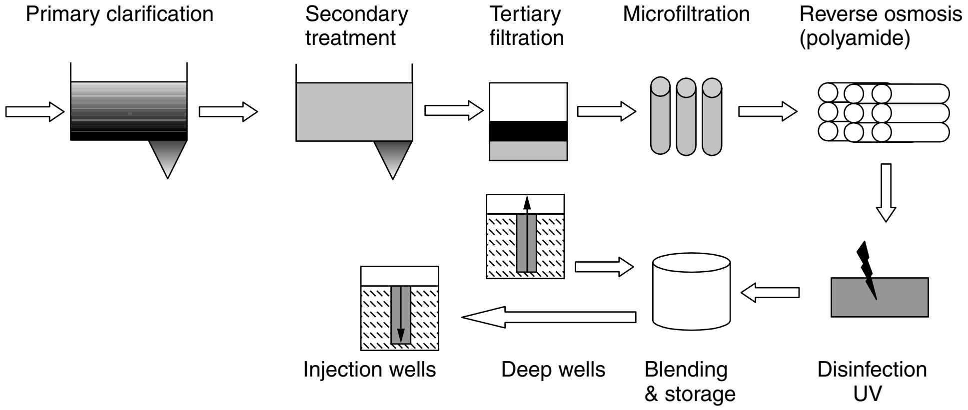

Figure 2. (*continued*).

to design facilities accordingly. Continued and future research will contribute to further progress in designing and operating water reclamation and reuse facilities. Some key topics will include assessment of health risks from trace pollutants in reclaimed water; improvement of monitoring techniques to evaluate microbiological quality; optimization of treatment trains and integration of membrane processes in producing reclaimed water; brine disposal strategies for membrane treatment processes; evaluation of the fate of microbiological, chemical, and organic contaminants; and development of surrogate parameters in combination with *in vitro* and *in vivo* assays to detect unidentified organic compounds in reclaimed water (1,4).

BIBLIOGRAPHY

1. Asano, T. (Ed.). (1998). *Wastewater Reclamation and Reuse*. Water Quality Management Library. Vol. 10. Technomic, Lancaster, PA, p. 1528.
2. Metcalf & Eddy, Inc. (2003). *Wastewater Engineering: Treatment and Reuse*, 4th Edn. Revised by G. Tchobanoglous, F.L. Burton, and H.D. Stensel, McGraw-Hill, New York.
3. Drewes, J.E., Fox, P., and Nellor, M.H. (2000). In: *Efficiency and Sustainability of Soil-Aquifer Treatment for Indirect Potable Reuse of Reclaimed Water*. I. Chorus, U. Ringelband, G. Schlag, and O. Schmoll (Eds.). Water Sanitation & Health, IWA, London, pp. 227–232.
4. National Research Council. (1998). *Issues in Potable Reuse—The Viability of Augmenting Drinking Water Supplies with Reclaimed Water*. National Academy Press, Washington, DC, p. 263.
5. Daughton, C. and Ternes, T. (1999). Pharmaceuticals and personal care products in the environment: agents of subtle change? *Environ. Health Perspect* **107**(suppl. 6): 907–938.
6. Drewes, J.E. and Shore, L. (2001). Concerns about pharmaceuticals in water reuse, groundwater recharge, and animal waste. In: *Pharmaceuticals and Personal Care Products in the Environment: Scientific and Regulatory Issues. Symposium Series 791*. C. Daughton and T. Jones-Lepp (Eds.). American Chemical Society, Washington, DC, pp. 206–228.

7. National Research Council. (1999). *Hormonally Active Agents in the Environment*. National Academy Press, Washington, DC.
8. Richardson, M.L. and Bowron, J.M. (1985). The fate of pharmaceutical chemicals in the aquatic environment. *J. Pharm. Pharmacol.* **37**: 1–12.
9. Halling-Sørensen, B. et al. (1998). Occurrence, fate and effects of pharmaceutical substances in the environment—a review. *Chemosphere* **36**: 357–394.
10. Bouwer, H. et al. (1980). Rapid-infiltration research—the flushing meadows project, arizona. *J. Water Pollut. Control Fed 52* **2**: 457.
11. Bouwer, E.J. et al. (1984). Organic contaminant behavior during rapid infiltration of secondary wastewater at the Phoenix 23rd Avenue Project. *Water Res.* **18**: 463.
12. County Sanitation Districts of Los Angeles County. (1977). *Pomona Virus Study—Final Report*. Whittier, CA.
13. Chen, C., Kuo, J., and Stahl, J.F. (1998). The role of filtration for wastewater reuse. In: *Wastewater Reclamation and Reuse, Water Quality Management Library*. Vol. 10. T. Asano (Ed.). Technomic, Lancaster, PA, pp. 219–262.
14. Gerba, C.P. et al. (1991). Fate of viruses in treated sewage effluent during soil aquifer treatment designed for wastewater reclamation and reuse. *Water Sci. Technol.* **24**: 95–102.
15. Powelson, D.K., Gerba, C.P., and Yahya, M.T. (1993). Virus transport and removal in wastewater during aquifer recharge. *Water Res.* **27**: 583–590.
16. Fox, P., Houston, S., Westerhoff, P., Drewes, J.E., Nellor, M., Yanko, W., Baird, R., Rincon, M., Arnold, R., Lansey, K., Bassett, R., Gerba, C., Kapiscak, M., Amy, G., and Reinhard, M. (2001). *Soil-Aquifer Treatment for Sustainable Water Reuse*. Final report. American Water Works Association Research Foundation (AWWARF), Denver, CO.
17. Kanarek, A. and Michail, M. (1996). Groundwater Recharge with Municipal Effluents, Dan Region Reclamation Project, Israel, *Proc. 18th Bienn. IAWQ Int. Conf. Water Qual.* June 23–26, Singapore, pp. 228–234.
18. Johnson, W.D. and Parnell, J.R. (1998). Wastewater reclamation and reuse in the City of St. Petersburg, Florida. In: *Wastewater Reclamation and Reuse*. Vol. 10. T. Asano (Ed.). Water Quality Management Library, Technomic, Lancaster, PA, pp. 1037–1104.
19. Young, R.E., Thompson, K.A., McVicker, R.R., Diamond, R.A., Gingras, M.B., Ferguson, D., Johannessen, J., Herr, G.K., Parsons, J.J., Seyde, V., Akiyoshi, E., Hyde, J., Kinner, C., and Oldewage, L. (1998). Irvine Ranch Water District's reuse today meets tomorrow's conservation needs. In: *Wastewater Reclamation and Reuse*. Vol. 10. T. Asano (Ed.). Water Quality Management Library, Technomic, Lancaster, PA, pp. 941–1036.
20. Yates, M. and Gerba, C. (1998). Microbial Considerations in Wastewater Reclamation and Reuse. In: *Wastewater Reclamation and Reuse*. Vol. 10. T. Asano (Ed.). Water Quality Management Library, Technomic, Lancaster, PA, pp. 627–704.
21. Crook, J. (1998). Water Reclamation and Reuse Criteria. In: *Wastewater Reclamation and Reuse*. Vol. 10. T. Asano (Ed.). Water Quality Management Library, Technomic, Lancaster, PA, pp. 627–704.
22. Tanaka, H., Asano, T., Schroeder, E.D., and Tchobanoglous, G. (1998). Estimating the safety of wastewater reclamation and reuse using enteric virus monitoring data. *Water Environ. Res.* **70**(1): 39–51.
23. Lauer, W.C. and Rogers, S.E. (1998). The demonstration of direct potable water reuse: Denver's landmark project. In: *Wastewater Reclamation and Reuse*. Vol. 10, T. Asano (Ed.). Water Quality Management Library, Technomic, Lancaster, PA, pp. 1269–1334.
24. McCarty, P.L. et al. (1980). *Mutagenic Activity and Chemical Characterization for the Palo Alto Wastewater Reclamation and Groundwater Injection Facility*. EPA-6001-81-029. U.S. Environmental Protection Agency, Washington, DC.
25. Fujita, Y., Ding, W., and Reinhard, M. (1996). Identification of wastewater dissolved organic carbon characteristics in reclaimed wastewater and recharged groundwater. *Water Environ Res* **68**(5): 867–876.
26. Mills, W.R., Bradford, S.M., Rigby, M., and Wehner, M.P. (1998). Groundwater recharge at the Orange County Water District. In: *Wastewater Reclamation and Reuse*. Vol. 10. T. Asano (Ed.). Water Quality Management Library, Technomic, Lancaster, PA, pp. 1105–1142.
27. Nellor, M.H. et al. (1984). *Health Effects Study*. Final Report, County Sanitation Districts of Los Angeles County, Whitter, CA.
28. Sloss, E.M., Geschwind, S.A., McCaffrey, D.F., and Ritz, B.R. (1996). Groundwater recharge with reclaimed water. *An Epidemiologic Assessment in Los Angeles County*. 1987–1991. Prepared for the Water Replenishment District of Southern California, Rand, Santa Monica, CA, p. 124.
29. Sloss, E.M., McCaffrey, D.F., Fricker, R.D., Geschwind, S.A., and Ritz, B.R. (1999). *Groundwater Recharge with Reclaimed Water. Birth Outcomes in Los Angeles County, 1982–1993*. Prepared for the Water Replenishment District of Southern California, Rand, Santa Monica, CA, p. 148.
30. Western Consortium for Public Health. (1997). *Total Resource Recovery Project Aqua III San Pasqual Health Effects Study Final Summary Report*. Berkeley, CA.
31. Hemmer, J. et al. (1994). Tampa Water Resource Recovery Project. *Proc. Am. Water Works Assoc/Water Environ. Fed. Water Reuse Symp*. Dallas, TX. American Water Works Association, Denver, CO, p. 557.
32. Levine, B. et al. (2000). Water quality assessment for indirect potable reuse: a new methodology for controlling trace organic compounds at the West Basin water recycling plant (California, USA). *Water Sci. Technol.* **43**(10): 249–527.
33. Joint Task Force of the Water Environment Federation and the American Water Works Association. (1998). *Using Reclaimed Water to Augment Potable Water Resources, Water Environment Federation/American Water Works Association*. Alexandria, VA/Denver, CO, p. 357.
34. Wegner-Gwidt, J. (1998). Public support and education for water reuse. In: *Wastewater Reclamation and Reuse*. Vol. 10. T. Asano (Ed.). Water Quality Management Library, Technomic, Lancaster, PA, pp. 1417–1462.

WASTEWATER RECLAMATION AND REUSE

RASHEED AHMAD
Khafra Engineering Consultants
Lilburn, Georgia

Because of growing population and increasing urbanization, many communities throughout the world are approaching or reaching the limits of their available water supplies. Wastewater reclamation and reuse has become an attractive option for conserving and extending available water supplies. Wastewater reclamation and reuse may also present communities an opportunity for pollution abatement when it replaces effluent discharge to sensitive surface waters.

The terminology currently used in wastewater reuse engineering is derived from sanitary and environmental engineering practice. The water potentially available for reuse includes municipal and industrial wastewater, agricultural return flows, and storm water. Of these, return flows from agriculture irrigation and storm water are usually collected and reused without further treatment. *Wastewater reclamation* involves treating or processing wastewater to make it reusable, and *wastewater reuse* or *water reuse* is the beneficial use of the treated water. Reclamation and reuse of water frequently require water conveyance facilities for delivering the reclaimed water and may require intermittent storage of the reclaimed water prior to reuse. In contrast to reuse, *wastewater recycling* or *water recycling* normally involves only one use or user, and the effluent from the user is captured and redirected back into that use scheme. In this context, the term *wastewater recycling* is applied predominantly to industrial applications such as in the steam—electric, manufacturing, and mineral industries.

The use of reclaimed wastewater where there is a direct link from the treatment system to the reuse application is termed *direct reuse*. Direct reuse provides water for agricultural and landscape irrigation, industrial applications, urban applications, and dual water systems. *Indirect reuse* includes mixing, dilution, and dispersion of reclaimed wastewater by discharge into an impoundment, receiving water, or groundwater aquifer prior to reuse, such as in groundwater recharge.

Indirect wastewater reuse, through effluent disposal to streams and groundwater basins, has been an accepted practice throughout the world for centuries. Communities that are situated at the end of major waterways have a long history of producing potable water from river water sources that have circulated through multiple cycles of withdrawal, treatment, and discharge. Similarly, riverbeds or percolation ponds may recharge underlying groundwater aquifers with treated wastewater, which in turn, is withdrawn by downgradient communities for domestic water supplies. These kinds of uses are considered unplanned indirect reuse. However, indirect reuse can also be planned, for example, the artificial groundwater recharge program in Los Angeles County, California, where reclaimed water has provided a source of water since the 1960s.

It is also important to differentiate between *potable* and *nonpotable* reuse applications. *Potable water reuse* refers to the use of highly treated reclaimed water to augment drinking water supplies. Although direct potable reuse is limited to extreme cases, it consists of incorporating reclaimed water into a potable water supply system without relinquishing control over the resource. *Nonpotable water reuse* includes all water applications other than drinking water supplies. Currently, on an international scale, direct nonpotable water reuse comprises the dominant mode of wastewater reuse for supplementing public water supplies for uses such as landscape and agricultural irrigation.

The potential health risks of wastewater reclamation and reuse are related to the extent of direct exposure to the reclaimed water and the adequacy, effectiveness, and reliability of the treatment system. The goal of each water reuse project is to protect public health without unnecessarily discouraging wastewater reclamation and reuse. Regulatory approaches stipulate water quality standards in conjunction with requirements for treatment, sampling, and monitoring. In 1992, the U.S. Environmental Protection Agency (USEPA) issued “Guidelines for Water Reuse” (1); however, specific criteria for wastewater reclamation in the United States are developed by individual states, often in conjunction with regulations on land treatment and disposal of wastewater. The World Health Organization (2) has published guidelines for reuse in agricultural irrigation. These guidelines stipulate stringent microbial water quality requirements. In recent years, many state agencies in the United States have developed their own reclamation and reuse standards. Besides microbial water quality criteria, these standards include standards for total suspended solids (TSS), biochemical oxygen demand (BOD), pH, and, in some cases, residual chlorine.

Wastewater reclamation and reuse are becoming more common and are gaining acceptance among communities and regulatory agencies. They have proven effective and successful in creating a new and reliable water supply, while not compromising public health. Nonpotable reuse is a widely accepted practice that will continue to grow. The primary drawback is cost, as these systems often require traditional water treatment facilities and a separate distribution system. The current advances in wastewater technology, especially in membrane processes, are expected to reduce the cost of wastewater reclamation and reuse in the future.

BIBLIOGRAPHY

1. USEPA. (1992). *Guidelines for Reuse Manual*. USEPA, Washington, DC.
2. WHO. (1989). *Health Guidelines for the Use of Wastewater in Agriculture and Aquaculture*. World Health Organization Scientific Group, Geneva, Switzerland.

READING LIST

Asano, T. (Ed.). (1998). *Wastewater Reclamation and Reuse*. Technomic, Lancaster, PA.

Rowe, D.R. and Abdel-Magid, I.M. (1995). *Handbook of Wastewater Reclamation and Reuse*. Lewis, Boca Raton, FL.

USEPA. (1991). *Municipal Wastewater Reuse: Selected Readings on Water Reuse*. USEPA, Washington, DC.

WASTEWATER RECLAMATION AND REUSE TREATMENT TECHNOLOGY

Rasheed Ahmad
Khafra Engineering Consultants
Lilburn, Georgia

The effective treatment of wastewater to meet water quality objectives for water reuse applications and to protect public health is a critical element of water reuse systems. Municipal wastewater treatment consists of a combination of physical, chemical, and biological processes and operations to remove solids, organic matter, pathogens, metals, and sometimes nutrients from wastewater. General terms used to describe different degrees of treatment, in order of increasing treatment level, are preliminary, primary, secondary, tertiary, and advanced treatment. A disinfection step for control of pathogenic organisms is often the final treatment step prior to distribution or storage of reclaimed wastewater. Wastewater reclamation, recycling, and reuse treatment systems are derived from applying technologies used for conventional wastewater treatment and drinking water treatment. The goal in designing a wastewater reclamation and reuse system is to develop an integrated cost-effective treatment scheme that can reliably meet water quality objectives.

The degree of treatment required in individual water treatment and wastewater reclamation facilities varies according to the specific reuse application and water quality requirements. The simplest treatment systems involve solid/liquid separation processes and disinfection, whereas more complex systems involve combinations of physical, chemical, and biological processes employing multiple-barrier treatment approaches for contaminant removal.

Primary treatment is the removal of particulate matter, typically by settling, with or without coagulants. In conventional wastewater treatment facilities, primary treatment includes screening and comminution for removal of large solids, grit, and sedimentation. Conventional primary treatment is effective in removing particulate matter 50 microns or larger. This process, in general, removes nearly 50% of the suspended solids and 25–50% of the biochemical oxygen demand (BOD_5) from the untreated wastewater. The removal efficiency of primary treatment processes can be increased by adding coagulants before gravity sedimentation. For most wastewater reuse applications, primary treatment alone is not adequate to meet water quality objectives.

Secondary treatment systems consist of biological treatment processes coupled with solid/liquid separation. They are intended to remove the soluble and colloidal organic matter that remains after primary treatment. Biological treatment consists of applying a controlled natural process, in which microorganisms remove soluble and colloidal organic material from the waste and are, in turn, removed themselves. To carry out this natural process in a reasonable time, the biological treatment systems are designed to maintain a large active mass of bacteria within the system. The basic principles remain the same in all biological processes, but the techniques used in their application vary widely. A useful classification divides these systems into *attached (film) growth or suspended growth processes*. Attached growth processes use a solid medium on which bacterial solids are accumulated to maintain a high population. Attached growth processes include trickling filters, rotating biological filters, fluidized beds, intermittent sand filters, and a variety of similar systems.

Suspended growth processes maintain an adequate biological mass in suspension within the reactor by employing either natural or mechanical mixing. In most processes, the required volume is reduced by returning bacteria from a secondary clarifier to maintain a high solids concentration.

Suspended growth processes include activated sludge and its various modifications, oxidation ponds, and sludge digestion systems. The effluent from conventional secondary treatment processes contains levels of suspended solids and BOD_5 ranging from about 10 to 30 mg/L. Depending on process operation, from 10% to 50% of the organic nitrogen is removed during conventional secondary treatment, and phosphorus is converted to phosphate. The sludge produced during secondary processes is treated by aerobic or anaerobic digestion, composting, or other solids processing technologies. For many wastewater reclamation and reuse systems, secondary treatment can adequately remove organic matter from wastewater. Frequently, secondary treatment is supplemented by filtration for additional removal of particles and by disinfection for microbial inactivation.

Tertiary treatment is the additional removal of colloidal and suspended solids by chemical coagulation and granular medium filtration. The advanced treatment is directed toward reduction in ammonia, organic nitrogen, total nitrogen, phosphorus, refractory organics, and dissolved solids. Tertiary and advanced treatment processes are normally applied downstream of biological treatment. Advanced treatment processes, in addition to chemical coagulation and granular medium filtration, refer to ion exchange, air stripping, chemical oxidation, adsorption, membrane treatment, and disinfection. Coagulation processes involve the addition of chemicals to wastewater to promote aggregation of particles for enhanced solid/liquid separation by sedimentation and filtration. Common coagulants used are alum (aluminum sulfate), ferric sulfate, and ferric chloride. Various types of polyelectrolytes or polymers are used in conjunction with chemical coagulants to improve process effectiveness. Chemical coagulation processes have a role in removing phosphorus, ammonia, and particulate matter.

Granular medium filtration uses a column consisting of a filter medium such as sand or anthracite. Wastewater is passed through the granular medium and particles are removed by impaction, interception, and physical straining. The filter is cleaned (this process is known as backwashing) after it reaches head-loss or turbidity breakthrough. Filters are generally used as a polishing unit and also, in some cases, as a pretreatment step for disinfection, activated carbon adsorption, ion exchange, or reverse osmosis. Filters are very effective in reducing pathogen concentrations.

Ion exchange is a chemical phenomenon in which materials in solution are removed by interchange with other ions immobilized within a solid matrix through which the flow passes. This process is used for removing nitrogen, metals, and dissolved solids.

Membrane processes include microfiltration, ultrafiltration, nanofiltration, and reverse osmosis. A membrane is a thin layer of natural or synthetic material capable of separating substances when a driving force is applied across the membrane. Reverse osmosis is used primarily to remove dissolved salts.

Nanofiltration is used to soften waters and remove color. Ultrafiltration and microfiltration are used to remove turbidity, pathogens, and other particulate matter. Membranes are normally classified by solute exclusion size, which is sometimes referred to as *pore size*. Membrane filtration works by passing water at high pressure through a thin membrane in the form of hollow fiber or spirally wound composite sheets. The contaminants are retained on the high-pressure side and frequently must be cleaned by reversing the flow and flushing the waste. Periodic chemical cleaning may be required to remove a persistent contaminant. Membrane assemblies are contained in pressure vessels or cartridges. Low-pressure membranes, in the form of either ultrafiltration (UF) or microfiltration (MF), have become more economical in both capital and operating costs and have received increased application in wastewater reuse and reclamation. Fouling of the membrane is a major problem challenging widespread application of this technology.

In recent years, the use of membrane bioreactors is proving to be an alternative to conventional biological processes. Membrane bioreactor technology combines a biological treatment process with a membrane system to provide organic and suspended solids removal. Installation within a biological reactor system typically replaces sedimentation and filtration as a means of separating mixed liquor suspended solids (MLSS) from treated effluent. Flow passes through the membranes, and solids remain in the biological treatment process. The benefit of immersed membrane technology is the amount of membrane surface area that is in contact with the process fluid. This large surface area allows operating membranes at much lower pressures.

Disinfection is an essential treatment component for almost all wastewater reclamation and reuse applications. The most common type of disinfection system in wastewater reclamation is chlorine disinfection. Ultraviolet (UV) disinfection has earned a reputation as a viable alternative to chemical disinfection processes in wastewater reclamation and reuse applications. The other technologies mentioned such as air stripping, chemical oxidation, and carbon adsorption are used in specific situations.

Water reuse for some applications, such as irrigation, is seasonal. Storage represents an important step between wastewater treatment and water reuse. Storage acts to equalize flow variations and to balance the production of reclaimed water with the use of water. Another benefit of storage is the additional residence time and treatment afforded.

Wastewater reclamation and reuse treatment technologies are drawn from currently practiced water and wastewater treatment technologies. However, some of the technologies such as membrane and ultraviolet disinfection, which are not widely used in conventional wastewater treatment plants, are gaining more viability in wastewater reuse and reclamation applications. This may be attributed to lower cost of operation and better process understanding and control of membrane and ultraviolet technologies. It may also stem from stricter regulatory standards and emphasis on multiple-barrier treatment techniques. The selection of treatment technologies for wastewater reclamation and reuse application depends on the type of reuse. For *urban reuse*, wastewater may require filtration and disinfection in addition to secondary treatment, whereas for *agricultural reuse*, secondary treatment with disinfection can meet reclaimed water quality standards. However, *indirect potable reuse* generally has to meet the strictest water quality standards. Hence, advanced wastewater treatment may be required to meet *indirect potable reuse* standards.

READING LIST

Asano, T. (Ed.). (1998). *Wastewater Reclamation and Reuse*. Technomic, Lancaster, PA.

Rowe, D.R. and Abdel-Magid, I.M. (1995). *Handbook of Wastewater Reclamation and Reuse*. Lewis, Boca Raton, FL.

USEPA. (1991). *Municipal Wastewater Reuse: Selected Readings on Water Reuse*. USEPA, Washington, DC.

USEPA. (1992). *Guidelines for Reuse Manual*. USEPA, Washington, DC.

WHO. (1989). *Health Guidelines for the Use of Wastewater in Agriculture and Aquaculture*. World Health Organization Scientific Group, Geneva, Switzerland.

SEWAGE

Absar Ahmad Kazmi
Nishihara Environment
Technology, Inc.
Tokyo, Japan

When human feces and urine are diluted with flushing water or other gray water (such as from washing, bathing, and cleansing activities), it becomes sewage, domestic wastewater, or sanitary wastewater. In other words, from the standpoint of sources of generation, sewage or domestic wastewater may be defined as a combination of the liquid- or water-carried wastes from residences, institutions, and commercial and industrial establishments, together with such groundwater, surface water, and stormwater as may be present. Sewage can be classified into two types.

- *Domestic sewage or domestic wastewater*: human excrement, waterborne human excretion, or water-carried wastes from liquid or nonliquid culinary purposes, washing, cleansing, laundering, food processing, or ice production;
- *Municipal sewage or municipal wastewater*: municipal liquid waste originating primarily from residences, but may include contributions from

1. holding tanks in recreational vehicles, boats, and houseboats;
2. commercial, institutional, and industrial sources; and
3. inflow and infiltration.

SEWAGE COMPOSITION AND CONTAMINANTS

Body wastes, food waste, paper, rags, and biological cells form the bulk of suspended solids in sewage. Even inert materials such as soil particles become fouled by adsorbing organics to their surfaces. Although suspended solids are biodegradable by hydrolysis, biodegradable material in sewage is usually considered soluble organics. Soluble organics in sewage are composed chiefly of proteins (40–60%), carbohydrate (25–50%), and lipids (approximately 10%). Proteins are chiefly amino acids; carbohydrates are compounds such as sugars, starches, and cellulose. Lipids include fats, oil, and grease. All of these materials contain carbon that can be converted to carbon dioxide biologically, thus exerting oxygen demand. Proteins also contain nitrogen, and thus a nitrogenous oxygen demand is also exerted. The biochemical oxygen demand (BOD) test is therefore used to quantify biodegradable organics. All forms of waterborne pathogens may be found in sewage wastewater. These include bacteria, viruses, protozoa, and helminthes. These organisms are discharged by persons who are infected with disease. A list of contaminants commonly found in sewage, along with their sources and environmental consequences, is given in Table 1.

The quantity and composition of sewage vary widely from location to location depending on, for example, food diet, socioeconomic factors, weather, and water availability. Quantitatively, constituents of sewage may vary significantly, depending on the other kinds of wastewater and the amount of dilution from the infiltration/inflow into the collection system. The results of analyzing a typical municipal wastewater or sewage from a municipal collection system are given in Table 2. The composition of wastewater from a given collection system may change slightly on a seasonal basis, reflecting different water uses. Additionally, daily fluctuations in quality are also observable and correlate well with flow conditions. Generally, smaller systems with more homogenous uses produce greater fluctuations in wastewater composition.

Table 2. Typical Analysis Of Municipal Wastewater

Constituent, mg/L	Concentration Strong	Medium	Weak
Solids, total	1200	720	350
Dissolved, total	850	500	250
Fixed	525	300	145
Volatile	325	200	105
Suspended, total	350	220	100
Fixed	75	55	20
Volatile	275	165	80
Settleable solids, mL/L	20	10	5
Biochemical oxygen demand, 5-day, 20 °C (BOD$_5$)	400	220	110
Total organic carbon (TOC)	290	160	80
Chemical oxygen demand (COD)	1000	500	250
Nitrogen, total as N	85	40	20
Organic	35	15	8
Free ammonia	50	25	12
Nitrites	0	0	0
Nitrates	0	0	0
Phosphorus, total as P	15	8	4
Organic	5	3	1
Inorganic	10	5	3
Chlorides	100	50	30
Alkalinity, as CaCO3	200	100	50
Grease	150	100	50

SEWAGE TREATMENT

If untreated sewage is allowed to accumulate, decomposition of the organic material it contains can lead to the production of large quantities of malodorous gases. In addition, untreated sewage usually contains numerous pathogenic or disease-causing microorganisms that dwell in the human intestinal tract. Sewage also contains nutrients, which can stimulate the growth of aquatic

Table 1. Important Wastewater Contaminants

Contaminant	Source	Environmental Significance
Suspended solids	Domestic use, industrial wastes, erosion by infiltration/inflow	Cause sludge deposits and anaerobic conditions in aquatic environment
Biodegradable organics	Domestic waste	Cause biological degradation, which may use up oxygen in receiving water and result in undesirable conditions
Pathogens	Domestic waste	Transmit communicable diseases
Nutrients	Domestic and industrial waste	May cause eutrophication
Refractory organics	Industrial waste	May cause taste and odor problems, may be caustic or carcinogenic
Heavy metals	Industrial waste, mining etc.,	Are toxic; may interfere with effluent reuse
Dissolved inorganic solids	Increase above level in water supply by domestic and/or industrial use	May interfere with effluent reuse

plants, and it may contain toxic compounds. For these reasons, the immediate and nuisance-free removal of sewage from its sources of generation, followed by treatment and disposal, is not only desirable but also necessary in an industrialized society. In the United States, it is now mandated by numerous federal and state laws. Sewage treatment is generally classified into on-site or off-site treatment systems according to nonsewered or sewered facilities.

ON-SITE SEWAGE TREATMENT SYSTEMS

They generally refer to individual septic tanks followed by soakpits. A septic tank is used to receive the wastewater discharged from individual residences and other nonsewered facilities. A septic tank followed by a soil absorption system constitutes what is known as conventional on-site sewage management system. Sludge has settled to the bottom of the septic tank over a period of years, and the liquid and surface scum layer is called septage. Septage and sludge generally require some level of treatment prior to final disposal or reuse. In many cases, septage is discharged to a municipal wastewater treatment plant and treated as a wastewater source. In separate treatment situations, the options of treatment include conventional wastewater treatment processes, conventional sludge treatment processes, land treatment, or a combination of aquatic treatment and constructed wetlands. Sludge can be treated in digesters or sludge lagoons as a liquid, or it can be dewatered and treated by land composting or land application.

OFF-SITE SEWAGE TREATMENT SYSTEMS

Off-site sewage treatment refers to the sewerage system that conveys wastewater to a chosen off-site disposal system. It is composed of a combination of unit operations and processes to reduce certain constituents of wastewater to an acceptable level. Many different combinations are possible. Municipal wastewater treatment systems are often divided into primary, secondary, and tertiary subsystems. The purpose of primary treatment is to remove solid materials from the incoming wastewater. Secondary treatment usually consists of biological conversion of dissolved and colloidal organics into biomass that can be subsequently removed by sedimentation. Contact between microorganisms and organics is optimized by suspending the biomass in wastewater or passing the wastewater over a film of biomass attached to solid surfaces. Secondary systems produce excess biomass that is biodegradable through endogenous catabolism by other microorganisms. Secondary sludges are usually combined with primary sludge for further treatment by anaerobic biological processes.

READING LIST

Crites, R. and Tchbanoglous, G. (1998). *Small and Decentralized Wastewater Management Systems*. McGraw-Hill, New York, pp. 943–950.

Peavy, H.S, Rowe, D.R., and Tchbanoglous, G. (1985). *Environmental Engineering*, International ed. McGraw-Hill, New York, pp. 209–214.

DOMESTIC SEWAGE

Kevin S. McLeary
Pennsylvania Dept. of
Environmental Protection
Harrisburg, Pennsylvania

CHARACTERISTICS OF DOMESTIC SEWAGE

Physical Characteristics

Fresh sewage typically has a light brown color. As it ages, oxygen is depleted, and anaerobic microorganisms begin to use other electron acceptors, including sulfates, nitrates, iron, and manganese. As sulfates are reduced, hydrogen sulfide is produced, which can combine with metals in the sewage. The color gradually changes from light brown to gray and eventually to black as a result of the metal sulfides.

The odor of fresh sewage is usually not objectionable, but as it ages, anaerobic microorganisms produce odors that are much more offensive than the odor of fresh sewage.

Odor. Some of the disagreeable compounds formed by the anaerobic decomposition of sewage include hydrogen sulfide, indole, skatole, cadaverine, and mercaptans (1).

Solids. Total solids are made up of soluble solids and suspended solids and are determined by evaporating a known volume of sewage at a temperature of 103–105 °C and weighing the residue. Total suspended solids (TSS) are determined by filtering a sample, followed by drying and weighing (in mg) the residue, and dividing by the volume of the sample (in L). Dissolved solids are the difference between the total and suspended solids. The suspended solids are further separated into settleable and nonsettleable solids. The settleable solids are those that will settle to the bottom of a cone-shaped container in a 60-minute period. The solids (total, suspended, or dissolved) can further be divided into volatile and fixed fractions. Samples are dried and then heated to 550 °C. The volatile materials are burned off, and the fixed solids are weighed. The volatile solids are presumed to be organic material, and the nonvolatile solids are presumed to be inorganic.

Temperature. The temperature of sewage is generally warmer than that of the water supply because of heat added by homes, businesses, and industrial activities. Sewage temperatures vary with the seasons and geography. In cooler climates, temperatures may vary from 7–18 °C (45–65 °F), and in warmer climates, the temperatures may vary from 13–24 °C (55–75 °F) (1).

Chemical Characteristics

Organic. The chief constituents of the organic compounds found in sewage are proteins (40–60%), carbohydrates (25–50%), and fats and oils (10%). Other compounds that may be present in small quantities include surfactants, pesticides, organic priority pollutants, and volatile organic compounds (2).

The organic strength of sewage is measured by a test called biochemical oxygen demand (BOD). The test does not directly measure the concentration of the organics, but rather measures the amount of oxygen that is used as the organic compounds are degraded by microorganisms.

Inorganic. The major inorganic constituents of sewage include alkalinity, sulfates, and metals as well as compounds of nitrogen and phosphorus. Alkalinity is a measure of buffering capacity against changes in pH and is a result of the presence of anions such as hydroxides, carbonates, and bicarbonates. Sulfur is an important element in the formation of proteins and is released when they are degraded. Metals occur naturally at low concentrations in water. Table 1 shows metal concentrations that may be found in sewage.

Small quantities of metals are vital to the biological treatment process as trace nutrients. However, at higher concentrations, metals are toxic to the treatment process.

Nitrogen. Inorganic nitrogen is in the form of ammonia, nitrites, and nitrates. Ammonia exists as both dissolved ammonia gas (NH_3) and the ammonium ion (NH_4^+). Nitrite comprises only a small portion of the inorganic nitrogen and is formed by the oxidation of ammonia. Nitrate is the most oxidized form of inorganic nitrogen in sewage treatment. Nitrate concentrations are a concern in water reuse applications because of methemoglobinemia. The drinking water standard for nitrate has been established at 10 mg/L as nitrogen.

Phosphorus. Inorganic phosphorus may be present as several orthophosphate ions (PO_4^{3-}, HPO_4^{2-}, $H_2PO_4^-$, and H_3PO_4). Phosphorus is an important nutrient for storing energy in cells. It is important for microorganisms, and phosphorus is frequently the limiting nutrient for phytoplanktonic (e.g., algae) growth in the aquatic environment. As the limiting nutrient, the growth of algae can be related to the phosphorus concentration. Excessive growth of algae can cause eutrophication of lakes and other slow moving bodies of water.

Gases. As sewage flows through the collection and conveyance system, atmospheric gases such as nitrogen, oxygen, and carbon dioxide may dissolve in the water. In addition to atmospheric gases, the organic compounds in the sewage begin to decompose, and the decomposition products, such as ammonia, hydrogen sulfide, and methane, also dissolve in the sewage (2).

Table 1. Metals Concentrations in Sewage Treatment Plant Influent[a]

	Concentration, mg/L	
Metal	Range	Average
Arsenic	0.0001–5	0.187
Chromium	0.0006–9	0.255
Copper	0.004–20	0.637
Lead	0.001–5	0.138
Mercury	0.0001–0	0.034
Nickel	0.002–7	0.294
Silver	0.0002–5	0.173
Zinc	0.024–2,027	39.4

[a] Reference 3.

Biological Characteristics

Untreated sewage is an ideal medium for microbiological organisms. It provides inorganic and inorganic nutrients that are needed for cell growth (3). Almost every type of microorganism can be found in sewage. The most abundant types include bacteria, viruses, protozoa, and helminths.

Bacteria. Bacteria are the principal microorganisms requiring the treatment of sewage. Bacteria convert carbonaceous materials into various gases and additional cell mass (2). One of the most common pathogens found in sewage is *Salmonella*. *Shigella*, the leading cause of recreational waterborne outbreaks in lakes and rivers, is also found. Some of the other bacteria that have been identified in sewage include *Vibrio, Mycobacterium, Clostridium, Leptospira, Yersinia, Escherichia coli*, and *Campylobacter* (4).

Viruses. Viruses are parasitic particles that consist of a strand of DNA or RNA. They invade living cells and redirect the cell to reproduce the virus, which accumulates within the cell. When the host cell dies, the viruses are released (2). Viruses of concern in sewage include enteroviruses, rotaviruses, reoviruses, parvoviruses, adenoviruses, and hepatitis A virus.

Protozoa. Protozoa are organisms that feed on bacteria and other microorganisms and are vital to the operation of biological treatment systems. Important protozoa in sewage treatment include amoebas, flagellates, and ciliates. There are also pathogenic protozoa in sewage. *Giardia lambia* and *Cryptosporidium* are included in this category (2).

Helminths. Sewage may contain a number of worms (helminths) that are of great concern. There are two major phyla of helminth that may be present, platyhelminthes and aschelminthes. Platyhelminthes are flatworms, including flukes and tapeworms. Nematodes are the major aschelminthes (roundworms) of concern in sewage. Of special concern are *Trichinella, Necator, Ascaris*, and *Filaria* (2).

Pathogenic Organisms. Pathogens are organisms that are highly infectious. Pathogens usually found in sewage are discharged by humans who are infected with or are carriers of a disease. These pathogens typically cause typhoid fever, diarrhea, cholera, and other diseases of the gastrointestinal tract.

Indicator Organisms. There is a large number and variety of pathogenic organisms in sewage. The time, cost, and difficulty of analyzing for all of these organisms would be excessive. To provide a level of confidence in pathogen concentrations, a substitute organism should be found for analysis. The ideal organism would be present when pathogenic organisms are present, be absent when pathogenic organisms are absent, be easy and inexpensive to analyze, and not be a pathogen itself. An organism that meets these criteria is called an indicator organism (5). The most widely used indicator organism in sewage treatment is fecal coliform bacteria.

TREATMENT TECHNOLOGY

Sewage must receive a minimum of a secondary treatment. A secondary treatment is defined as a 30-day average concentration of biochemical oxygen demand and total suspended solids not to exceed 30 mg/L, where the 7-day average does not exceed 45 mg/L. The definition also includes an 85% removal requirement for BOD and TSS. Sewage effluent that may contact humans or be indirectly used for water supply may require additional treatment. Selection of treatment technology is dependent in large part on the influent characteristics, the quality of the effluent required, and cost.

Treatment is normally classified as either physical/chemical or biological and can be broken down into several stages: preliminary treatment, primary treatment, secondary treatment, and tertiary treatment.

Preliminary Treatment Processes

Preliminary treatment is a physical process intended to remove large objects and grit from sewage (5). The removal of these materials is necessary because they could reduce the efficiency or increase the maintenance of downstream processes (6). Preliminary treatment may include the following processes: screening, grit removal, comminution, and flow equalization.

Screening. Screening removes large objects that could clog or damage downstream equipment. Screens typically consist of inclined steel bars spaced at equal intervals in a sewage channel. Common practice is to use a mechanically cleaned bar screen that has an emergency bypass channel containing a manually cleaned screen (7). Design parameters for bar screens include bar size, bar spacing, angle of inclination, channel width, and sewage approach velocity (6).

Grit Removal. Grit consists of sand, gravel, and other high specific gravity material that may abrade and wear mechanical equipment or may accumulate in treatment tanks. A common method for grit removal involves using aerated grit chambers, in which diffused air is introduced to the sewage along the bottom of one side of a rectangular chamber. This creates a rolling motion that keeps the lighter organic materials in suspension but allows the heavier grit particles to settle to the bottom of the tank, where they are removed (6).

Comminution. Comminutors are devices that shred materials without removing them from the sewage. They are typically designed with a bypass channel containing a screening device that may be used if the comminutor is removed from service. The basic parts of a comminutor include a screen and cutting teeth on a revolving drum that has slots cut into it. The cutting teeth shred the material as it is trapped against the screen. The shredded solids then pass through the drum slots and out of the bottom (7). These shredded materials can cause problems, such as clogging, in downstream process units. As a result, comminutors have fallen out of favor in the design of sewage treatment plants (6).

Equalization. Equalization is used in plants to dampen variations in hydraulic or organic loadings. Sewage flows into the unit at a variable rate but is removed at a relatively constant rate. Equalization is used to provide relatively uniform loading throughout the day and to increase the performance of downstream process units (6). Equalization tanks may be designed either as in-line or side-line units.

Equalization tanks must be aerated to prevent the contents from becoming septic and should be mixed to prevent solids from settling. Aeration equipment should be capable of maintaining a dissolved oxygen concentration of 1 mg/L (8). Mixing may be provided by diffused air systems, mechanical mixers, or both.

Primary Treatment Processes

Sedimentation. Primary sedimentation is the oldest and most widely used process in treating sewage (6). It is a physical process whose goal is to achieve solids separation.

Solids removal by sedimentation is a function of retention time and surface settling rate. The surface settling rate is defined as the volumetric flow rate over the surface area of the clarifier in units of velocity. Particles whose settling velocity is greater than the surface settling rate are removed from the sewage stream. However, if the detention time is too long, the sewage turns septic, and gas bubbles formed in the sewage reduce the efficiency of the process. A typical minimum side water depth for primary clarifiers is 10 feet. To allow for adequate settling, a minimum distance of 10 feet should separate the inlet and the outlet (8). Clarifier design is typically based on two flows, the average design flow and the peak hourly flow. The calculated size of the clarifier is based on both flows, and the larger clarifier is selected.

Secondary Treatment Processes

Activated Sludge Processes. Activating sludge is a biological treatment process using a suspension of microorganisms to treat sewage in an aerobic environment. The microorganisms are allowed to flocculate and settle under quiescent conditions, and treated sewage then flows over weirs for further treatment or discharge. Solids from the bottom of the clarifier are recycled to the reactor to provide an adequate concentration of microorganisms for treatment. The contents of the reactor, called mixed liquor, must be aerated and mixed by using either mechanical aerators or diffused air.

There are several variations of the conventional activated sludge process. These include plug flow reactors, including step feed, tapered aeration, extended aeration, and complete mix reactors, including sequencing batch reactors.

A plug flow reactor has a configuration in which the sewage flows through a long, narrow channel for treatment. It approximates flow through a pipe. In an ideal plug flow reactor, there is no longitudinal mixing of the sewage. A step feed reactor is a variation of the plug flow reactor in which the sewage is introduced into the reactor at several places. This allows more equal distribution of the organic load. Tapered aeration is another variation of the plug flow reactor. In tapered aeration, the majority of the aeration capacity is provided at the head of the reactor, where the organic load is the highest, and less aeration is provided where the organic load is lower. Extended aeration is a treatment process requiring long detention times (typically greater than 24 hours) and low organic loadings. Extended aeration is commonly available in package-type treatment plants and is economical for small treatment plants.

A complete mix reactor is the opposite of a plug flow reactor. All of the sewage is completely mixed in a short, wide reactor. Due to rapid and complete mixing of the reactor contents, complete mix reactors can tolerate shock loads better than plug flow reactors. A sequencing batch reactor (SBR) is a variation of the complete mix reactor; stabilization, settling, and equalization take place in the same tank, eliminating the need for a clarifier.

Aeration Requirements. The dissolved oxygen concentration in aeration tanks should be greater than 2 mg/L at all times. Treatment processes (except extended aeration) should be designed to provide 1.1 lb of oxygen per pound of BOD treated. Extended aeration systems generally require 1.5 lb of oxygen per pound of BOD treated. These aeration requirements do not include the aeration capacity needed for nitrification. If nitrification is required, an additional 4.6 lb of oxygen per pound of ammonia nitrogen is required (8).

Fixed Film Processes

Trickling Filters. Trickling filters are a fixed film process where the microorganisms are attached to a stone or plastic medium. The sewage flows through a rotating arm, which distributes it over the medium. As the sewage flows over the medium, the microorganisms absorb organics from the sewage. When the sewage is not being applied to that specific section of the medium, air flows through the filter, providing the oxygen that the microorganisms need for respiration.

Sewage is recirculated back to the filter to maintain a proper application rate for efficient operation of the filter, to equalize the organic loading to the filter, and to prevent the microorganisms from drying out.

Rotating Biological Contactors. Rotating biological contactors (RBCs) are another version of the fixed film process. The microorganisms are attached to a plastic disk, which is partially submerged and rotated through the sewage. When the microorganisms are submerged, they absorb organics. During the time the microorganisms are exposed to the air, they receive the oxygen that is required for treatment. Treatment efficiency is a function of the surface area of the disks, more surface area provides greater treatment. Unlike trickling filters, no recirculation is required for rotating biological contactors.

Stabilization Ponds. Stabilization ponds are large, lined basins, that may be aerobic, facultative, or anaerobic. Ponds use detention time measured in days, rather than hours, and are typically relatively shallow compared with other biological treatment processes. Thus, a large land area is required for ponds, and they are usually used only in small communities. Their advantages include low construction and operating costs.

Aerobic ponds may be aerated mechanically or naturally. Natural aeration is by atmospheric diffusion and production of oxygen by algae. Facultative ponds have several stratified layers—an upper, aerated section; a lower, anaerobic section; and an intermediate section consisting of both aerobic and anaerobic processes. Anaerobic ponds may be up to 30 feet deep and are used for treating high strength (typically industrial) waste. Deep ponds maximize anaerobic conditions.

Tertiary Treatment Processes

Nitrification. Ammonia nitrogen is converted to nitrate in a two-step process, in which ammonia is first converted to nitrites and the nitrites are then converted to nitrates. The rate-limiting step is the conversion of ammonia to nitrite. Nitrification can co-occur with carbon oxidation, or it may take place in a separate nitrification tank. The reaction rate is slower and, therefore, requires a longer detention time than carbon oxidation. Nitrifying organisms have a slower growth rate than the organisms for carbon oxidation, and the process requires a longer mean cell residence time (sludge age).

Biological Phosphorus Removal. Biological phosphorus removal can be enhanced in a two-step process. The first step takes place anaerobically. The microorganisms release phosphorus to generate energy for the uptake of organics. The second step is aerobic. In this step, the microorganisms absorb large amounts of phosphorus to replace the phosphorus that was lost in the anaerobic step, as well as to store additional energy for the next "feast or famine" feeding cycle.

There are three major biological phosphorus removal methodologies—the Anaerobic/Oxic (A/O) process, the PhoStrip process, and the sequencing batch reactor (2). Both the A/O process and the PhoStrip process are proprietary. Phosphorus removal in the A/O process is dependent on the BOD:P ratio; the PhoStrip process is independent of this ratio. The PhoStrip process does require additional chemicals for phosphorus removal, however. The sequencing batch reactor may be cycled to achieve biological phosphorus removal but usually is used for smaller flows and with more limited design data.

Denitrification. Denitrification is the removal of the inorganic nitrogen from sewage. Several species of bacteria can use nitrates, rather than oxygen, as their energy source. These bacteria convert the nitrates into nitrogen gas (2).

In the denitrification process, raw sewage flows into an anoxic zone with return sludge and return mixed liquor from an aerobic zone. The anoxic zone denitrifies by using the nitrates in the mixed liquor. Following the anoxic zone, the sewage flows to an aerobic zone where nitrates are created. The nitrates are then recycled to the anoxic zone for removal. Denitrification is normally done in a plug flow type system or in an oxidation ditch, although a sequencing batch reactor may be programmed for denitrification (2).

Biological Dual Nutrient Removal. Biological dual nutrient removal is the reduction of both nitrogen and phosphorus from sewage by microorganisms. Several proprietary treatment processes, including the A^2/O process, the Bardenpho process, the University of Capetown (UCT) process, and the Virginia Initiative Plant (VIP) process, have been developed for dual nutrient removal (2).

These processes are a combination of the denitrification process and the biological phosphorus removal process. The proprietary systems may use from three to five stages to achieve the desired nutrient removal, but all have the use of an anaerobic zone in common, followed by an anoxic zone, followed by an aerobic zone. Some of the processes may use two anoxic zones and/or two anaerobic zones with different recycle streams to achieve greater nutrient removal, but the treatment principles are the same.

Coagulation/Sedimentation. Coagulation/sedimentation requires chemical addition to enhance the sedimentation of solids, precipitate pollutants, or remove phosphorus. The chemicals most commonly used are lime, aluminum salts, ferric salts, and polymers (7).

Chemical phosphorus removal occurs by the addition of chemicals to the sewage, which create an insoluble phosphate precipitate. Alum is frequently used in the chemical precipitation of phosphorus, although iron salts may also be used. Alum also reacts with hydroxyl radicals in the water, forming aluminum hydroxide, in addition to aluminum phosphate. Iron (III) reacts in the same manner.

Filtration. Filtration is the removal of solids by passing the sewage through a bed of granular media. Although the most commonly used filters are composed of sand, filters may also consist of multiple types of media, such as coal over sand or coal over silica sand over garnet sand (9). Filters may be classified as slow filters, rapid filters, or pressure filters. Slow filters require a buildup of a biological mat on the upper surface of the filter, which provides greater treatment, but requires a low application rate, and therefore requires a larger area. Rapid filters and pressure filters depend on the entire depth of the media for filtration and may be operated at higher loading rates than slow filters, although backwashing of the media is required.

Activated Carbon Adsorption. Adsorption is a process by which a compound adheres to a solid surface. In sewage treatment, activated carbon is the most commonly used adsorbent. Activated carbon comes in two forms, powdered and granular. Powdered activated carbon (PAC) is applied in slurry form at the head of the aeration tanks and is removed in the final clarifiers. Granular activated carbon (GAC) is used in a filter bed (7). Carbon adsorption is used only where highly treated effluent is required.

Membrane Systems. Membrane processes involve the use of a semipermeable barrier. The membrane allows the water to flow through and retains the contaminants. There are several types of membrane systems in sewage treatment, including reverse osmosis, nanofiltration, microfiltration, and ultrafiltration. All of these processes require pressure to force water through the membrane. Ultrafiltration requires the least pressure, whereas reverse osmosis requires the greatest pressure. Membrane processes are subject to fouling of the membranes. These processes should be pilot tested to determine which process and membrane will work best. Like carbon adsorption, membrane processes are used when only high-quality effluent is required.

Disinfection Processes

Chlorination/Dechlorination. Chlorine has been used as a disinfectant for sewage for several reasons, including inactivation of a wide range of pathogens, maintenance of a residual, and economy. There are several forms of chlorine that may be used: gaseous chlorine, sodium hypochlorite, and calcium hypochlorite. Chlorine is toxic to aquatic life, so the recent trend has been to dechlorinate the sewage before discharge to the receiving stream, which is usually done by using sulfur dioxide to reduce the chlorine to chlorides. Sodium metabisulfite or bisulfite may be used as a substitute for sulfur dioxide in small facilities. Reaction times are nearly instantaneous, and detention times are usually less than 2 minutes.

Ozonation. Ozone is a very powerful oxidant. It can inactivate sewage pathogens with less contact time and a lower dosage than other disinfection methods. It is effective against a wide range of organisms, and it does not leave a toxic residual (10). Ozone must be generated on-site because it is unstable. Ozone is generated by corona discharge, which consists of passing clean, dry air or oxygen through electrodes, which are separated by a dielectric and a gap.

Ultraviolet Light Disinfection. Ultraviolet radiation whose wavelengths are in the range of 240–280 nm inactivate microorganisms by causing damage to their DNA (10). Ultraviolet lamps operate in the same way as fluorescent lamps—the radiation is generated by passing an electrical current through ionized mercury vapor. The mercury lamps may operate at low or medium pressures. Low-pressure lamps emit the majority of their energy at 253.7 nm, which is in the optimal range for inactivation. Medium-pressure lamps generate a smaller portion of their energy in the 240–280 nm range, but the intensity of their light is much greater. Therefore, fewer medium-intensity lamps are required for the same amount of disinfection (10).

BIBLIOGRAPHY

1. Linsley, R.K. and Franzini, J.B. (1979). *Water Resources Engineering*, 3rd Edn. McGraw-Hill, New York.
2. Metcalf & Eddy, Inc. (1991). *Wastewater Engineering: Treatment, Disposal, Reuse*, 3rd Edn. McGraw-Hill, New York.
3. Water Environment Federation. (1994). *Wastewater Biology: The Life Processes*. Alexandria, VA.
4. US EPA. (1992). *Guidelines for Water Reuse*. EPA/625/R-92/004.
5. Sincero, A.P. and Sincero, G.A. (1996). *Environmental Engineering: A Design Approach*. Prentice-Hall, Upper Saddle River, NJ.
6. Water Environment Federation and American Society of Civil Engineers. (1992). *Design of Municipal Wastewater Treatment Plants*. WEF Manual of Practice No. 8/ASCE Manual and Report on Engineering Practice No. 76, 3rd Edn.
7. Water Environment Federation. (1989). *Water Reuse, Manual of Practice SM-3*, 2nd Edn. Alexandria, VA.
8. Great Lakes-Upper Mississippi River Board of State and Provincial Public Health and Environmental Managers. (1997). *Recommended Standards for Wastewater Facilities*. Health Education Service, Inc., Albany, NY.
9. Weber, W.J. (1971). *Physicochemical Processes for Water Quality Control*. John Wiley & Sons, New York.
10. US EPA. (1999). *Alternative Disinfectants and Oxidants Guidance Manual*. EPA 815-R-99-014.

SOLIDIFICATION/STABILIZATION OF HAZARDOUS SOLID WASTES

Th. A. Ioannidis
A.I. Zouboulis
Aristotle University of Thessaloniki
Thessaloniki, Greece

INTRODUCTION

Solidification/stabilization is an established technology that has been used for almost 20 years to treat a variety of toxic or hazardous wastes. Historically, this technology belongs to the top five source control treatment technologies, used at Superfund remedial sites, as presented in Fig. 1 (1).

Stabilization is a process by which sufficient quantities of specific additives (reagents) are added to hazardous materials to reduce the toxic nature (properties) of a waste by converting the toxic constituents into an appropriate solidified form. As a result, (a) the rate of contaminant migration into the surrounding environmental media is minimized, and (b) the level of toxicity is substantially reduced, as determined by the application of the appropriate leaching tests, such as TCLP. The procedure of solidification/stabilization has shown considerable promise, and it is commonly used for the fixing different waste types. The wide use of stabilization for treating (mainly) inorganic-laden wastes derives primarily from

1. the lack of better alternatives; for example, toxic metals do not biodegrade and they do not change in atomic structure/properties, when incinerated.
2. the well-defined physicochemical mechanisms taking place, such as precipitation and adsorption (2).

SOLIDIFICATION/STABILIZATION TECHNOLOGY

Solidification and stabilization (denoted hereafter as S/S) waste treatment processes involve mixing specialized additives or reagents with hazardous waste materials to reduce, by physical or chemical means, the solubility or mobility of contaminants in the surrounding environmental matrix. Solidification and stabilization are closely related; both use several chemical, physical, and thermal processes or an appropriate combination of them, to detoxify hazardous wastes.

The term "solidification/stabilization" refers to a general category of processes, that are used to treat a wide variety of hazardous (or toxic) wastes, including mostly solids, but also liquids. Nevertheless, solidification and stabilization are considered distinct technologies.

Solidification refers to processes that encapsulate a waste to form a solid material and to restrict the migration of contaminant by decreasing the available surface area, which is exposed to leaching, when contacting liquids and also by coating the waste with low-permeability materials. Solidification can be accomplished by a chemical reaction between the hazardous waste and the binding (solidifying) reagents or by the application of mechanical processes (e.g., compaction). The solidification of fine solid waste particles (i.e., those whose diameters are below 100 μm) is referred to as *microencapsulation*, whereas

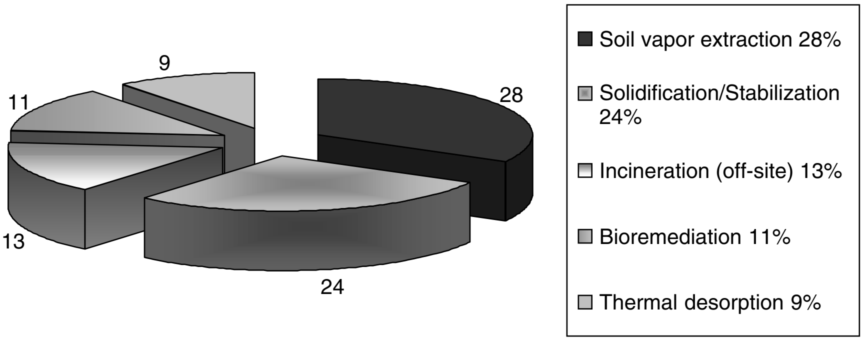

Figure 1. The top source control treatment technologies used at Superfund remedial sites.

the solidification of a large block or container of waste is referred to as *macroencapsulation*.

Stabilization refers to processes that involve chemical reactions, which reduce the leachability of the product (stabilized waste). Stabilization immobilizes the hazardous materials or reduces their solubility through appropriate chemical reactions. The physical nature of the waste may or may not be changed by this process.

Solidification refers to processes that encapsulate the waste in a monolithic solid, which has sufficient structural integrity. The encapsulation may be that of compacted fine waste particles, or of a large block, or container of wastes. Solidification does not necessarily involve a chemical interaction between the waste and the solidifying reagents, but the binding of waste in the product (monolith) may be mechanical. Migration of contaminants is restricted by the large decrease in surface area, which is exposed to eventual leaching, and by isolating the waste within an impervious capsule.

Stabilization refers to processes, which reduce the risk posed by a waste, through the conversion of contaminants into a less soluble, immobile, or less toxic form, whereas the physical nature of the waste may not be necessarily changed. Phosphates, sulfides, carbonates, and several other materials have been used as appropriate treatment reagents.

In many instances, stabilization takes place exclusive of solidification. A stabilized product should have low leaching characteristics. Many of the reagents used for the S/S method can also be used in other chemical treatment (oxidation) processes, such as in dechlorination of liquid wastes. S/S systems can be used to treat contaminated soil or wastes in place, or they can be employed to treat excavated wastes externally, preparing them for subsequent disposal (3).

Solidification involves the formation of a solidified matrix that physically binds the contaminated material. Stabilization, also referred to as *fixation*, usually uses a chemical reaction to convert the hazardous constituents of the waste to a less mobile form. The general approach for S/S treatment processes involves mixing or injecting treatment agents into the contaminated soils. Inorganic binders, such as cement, fly ash, or blast furnace slag, as well as organic binders, such as bitumen, have been used to form a crystalline, glassy, or polymeric framework around the waste.

The dominant mechanism by which metals are immobilized in the absence of anions such as phosphates, carbonates, halides, or sulphates, is the precipitation of respective hydroxides within the solid matrix. S/S technologies are not useful for certain forms of metal contamination, such as species that exist as anions [e.g., Cr(VI) or arsenic], or for metals whose hydroxides are very soluble (e.g., mercury).

The S/S technology may not be applicable at polluted sites, containing wastes that include organic compounds, especially when volatile organic contaminants are present. The application of mixing and heating processes, associated with the hydration of binders, may release organic vapors. Pretreatment, such as air stripping or incineration, may be used to remove the volatile organics during preliminary treatment and to prepare the waste for the S/S of residues. The chemical composition of the contaminated matrix, the amount of water present, and the ambient temperature also affect the application of S/S technologies. Some factors can interfere with the effectiveness of the S/S process by inhibiting bonding of the waste to the binding material, retarding the setting of the mixtures, decreasing the stability of the matrix, or reducing the strength of the solidified products.

Cement-based binders and stabilizers are commonly used materials, when implementing S/S technologies (4). Portland cement, a mixture of Ca silicates, aluminates, alumino ferrites and sulfates, is an important cement-based material. Pozzolanic materials, which consist of small spherical particles, formed by coal combustion in lime and cement kilns, such as fly ash, are also commonly used for S/S. Pozzolans exhibit cement-like properties, especially when their silica content is relatively high. Portland cement and pozzolans can be used alone or together to obtain optimal stabilization properties for a particular waste or polluted site (5).

Organic binders may also be used to treat metals by polymer microencapsulation. This process uses organic materials, such as bitumen, polyethylene, paraffins, waxes, and other polyolefins, as thermoplastic or thermosetting resins. For polymer encapsulation, the organic materials are heated and mixed with the contaminated matrix at elevated temperatures (120–200 °C). The organic materials polymerize, agglomerate the waste, and as a result, the waste matrix is encapsulated (5). The organics are volatilized and collected, and the treated material is extruded for disposal or possible reuse (e.g., as paving material).

The contaminated material may require a certain extent of pretreatment to separate rocks and other debris and to dry the feed material. Polymer encapsulation requires more energy and more complex equipment than cement-based S/S operations. Bitumen (asphalt) is the cheapest and most common thermoplastic binder (5). S/S is achieved by mixing the contaminated material with appropriate quantities of binder/stabilizer and with water. The mixture sets and cures to form a solidified matrix, which contains the hazardous constituents. The cure time and handling characteristics of the mixture and the final properties of the hardened cement depend on the specific composition of the binder/stabilizer used (i.e., amount of cement, pozzolan, water, etc.).

Vitrification is the process of converting solid materials into a glass-like substance. This process is increasingly being considered for treating various toxic or hazardous wastes. The mobility of metal contaminants can be highly decreased by the application of high temperatures to the contaminated waste, which results in the formation of vitreous material, usually as oxide solids. During this process, the increased temperature may also cause the volatilization and/or destruction of organic contaminants or of volatile metal species (such as Hg), which must be carefully collected for further treatment or disposal. Most contaminated soils can be treated by vitrification, and a wide variety of inorganic and organic contaminants can be targeted. As a stabilization process, vitrification may immobilize inorganics by incorporating them into the glass structure or by encapsulating them within the product (glass). Vitrification may be performed *ex situ* or *in situ*, although the *in situ* processes are preferred, due to lower energy requirements and cost. The performance and characteristics of vitrification make it the focal point of treatment systems applied to high-level radioactive waste (HLW) around the world (6).

Ex situ S/S processes can be easily applied to excavated soils because the necessary mechanical methods are available to provide vigorous mixing, which is needed to combine the binder/stabilizer with the contaminated material. Pretreatment of a solid waste may be necessary to screen and crush large rocks and debris. Mixing can be performed via *in-drum, in-plant*, or whole area mixing processes. In-drum mixing may be preferred for treating small volumes of wastes or for toxic wastes. In-plant processes use rotary drum mixers for batch processes or pug mill mixers for continuous treatment. Larger volumes of wastes may be excavated and moved to a contained area for area mixing. This process involves layering contaminated material with the stabilizer/binder and subsequent mixing with a backhoe or similar equipment. Mobile, as well as fixed treatment plants are available for *ex situ* S/S treatment. Smaller pilot-scale plants can treat up to 100 tons of contaminated waste or soil/day; larger portable plants typically process 500 to over 1000 tons/day.

S/S techniques are also available to mix binder/stabilizer with the waste or contaminated soil *in situ*. *In situ* S/S technology is less labor- and energy-intensive than *ex situ* processes, which require excavation, transport, and disposal of the treated material. The *in situ* S/S method is also preferred, when volatile or semi-volatile organics are present in a waste because otherwise the excavation would expose these contaminants to the air. However, the presence of bedrock, large boulders, cohesive soils, oily sands, and clays may preclude the application of the *in situ* S/S method at certain sites. It is also more difficult to provide uniform and complete mixing, when applying *in situ* processes. The binder may be mixed with the contaminated matrix using in-place mixing, vertical auger mixing, or injection grouting. In-place mixing is similar to *ex situ* area mixing, except that the waste or soil is not excavated prior to treatment.

The *in situ* process is useful for treating surface or shallow contamination and involves spreading and mixing binders with the waste, using conventional excavation equipment, such as draglines, backhoes, or clamshell buckets. Vertical auger mixing uses a system of augers to inject and mix the binding reagents with the waste. Larger (2–4 m diameter) augers are used for shallow drilling (depths of 3–13 m) and can treat 400–800 m^3/day. Deeper applications of S/S methods (up to 50 m deep) can be achieved by using ganged augers (up to 1 m in diameter each), which can treat 100–300 m^3/d. Finally, injection grouting may be performed to inject the binder, containing suspended or dissolved reagents, into the treatment area under pressure. The binder permeates the surrounding soil and cures in place (7).

The main generic elements of typical S/S are presented in Fig. 2.

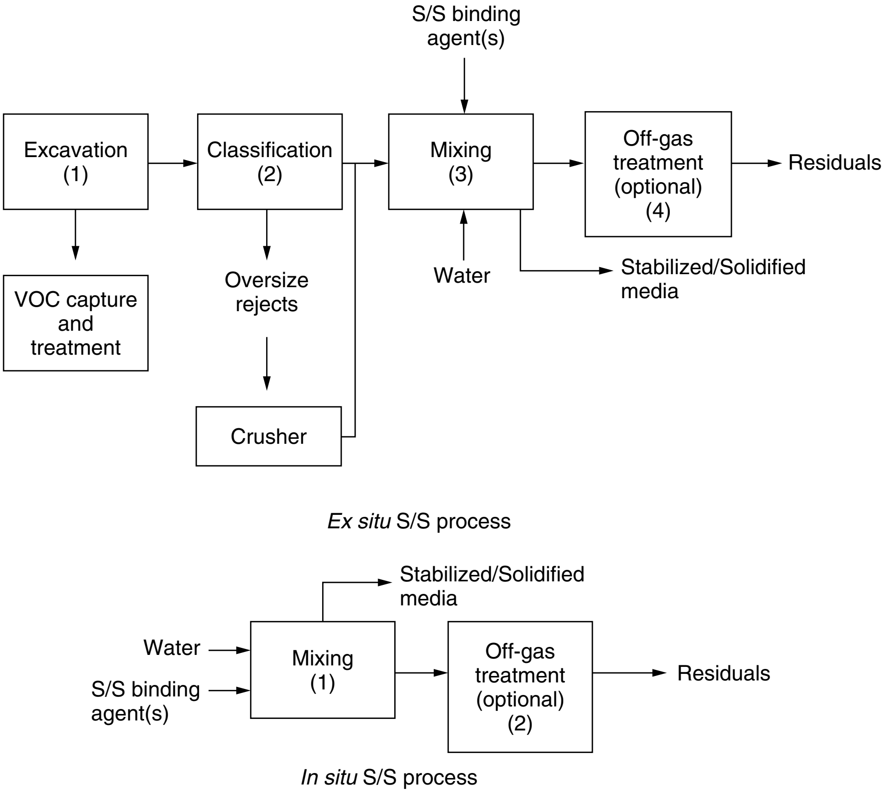

Figure 2. Generic elements of typical S/S processes.

KEY FEATURES OF S/S PROCESSES FOR BEST MANAGEMENT PRACTICES (BMPs)

The key features of S/S processes for BMPs are briefly presented in the following:

- This process requires the mixing of reagents, either on- or off-site.
- As a result of S/S application, contaminants are immobilized.
- Similarly to other immobilization technologies, S/S does not destroy inorganic wastes but may alter or change organic wastes.
- Stabilization can be combined with encapsulation or other immobilization technology.
- The application of S/S may increase the total volume of materials, which must be handled as waste for disposal.
- The wastes treated by the S/S method may be amenable to further reuse, following this treatment.
- Field application may involve the installation of any or all of the following:
  - a) auger type drilling and mixing equipment for *in situ* applications.
  - b) dust collection systems.
  - c) volatile emission control systems.
  - d) bulk storage tanks.

CROSS-MEDIA TRANSFER POTENTIAL OF POLLUTANTS AFTER S/S

There is a certain potential for the cross-media transfer of pollutants, after S/S:

1. There is always a risk of inaccurate site characterization, which has to be considered during the application of technology for treating soils. The solid material encountered at the remedial site may not be similar to the soils studied during laboratory treatability or pilot-scale tests. Additional contaminants may be encountered, whereas the percentage of the fine-grained fraction may be significantly different from that expected. These factors may lead to long-term storage or generation of a high volume of residuals that increases the potential for cross-media transfer of pollutants.
2. Also, during remediation, implementation, including staging and site preparation (e.g., clearing, grubbing), drilling, well installation and trenching operations, mobilization and demobilization of equipment, excavation, transport of materials across the site, and certain treatment activities, there is high potential for fugitive dust emissions due to the movement of equipment at the site. In addition, these activities can enhance the volatilization of VOCs (volatile organic compounds), SVOCs (semivolatile organic compounds), and of other potentially hazardous materials into the atmosphere.
3. During pretreatment operations, such as excavation, storage, sizing, crushing, dewatering, neutralization, blending, and feeding, there is the potential for dust and VOC emissions from the contaminated media (solid or liquid wastes).
4. The migration of contaminants to uncontaminated areas may occur during mobilization or demobilization.
5. VOC and SVOC emissions tend to increase during periods of hot and dry weather.
6. Leaching of contaminants to neighboring surface waters can occur from uncovered stockpiles and from excavated pits.
7. Improper handling and disposal of residues, such as sediment/sludge residuals or postwashing wastewaters, may allow contaminants to migrate into and pollute uncontaminated areas.
8. The posttreatment discharge of wastewater produced, when improperly managed, can also cause migration of contaminants.

Additional concerns for the application of S/S technology include

1. The leaching of contaminants or excess reagents to groundwaters from the treated wastes, which are disposed of on-site.
2. The long-term degradation of the stabilized mass, creating a potential for release of solidified hazardous constituents, reagents, VOCs, etc. from the treated waste.

WASTE CHARACTERISTICS THAT MAY INCREASE THE LIKELIHOOD OF CROSS-MEDIA CONTAMINATION FROM THE APPLICATION OF S/S TECHNOLOGIES

The effectiveness of S/S applications can be compromised and provoke cross-media contamination under certain conditions, which are identified in the following (8). However, some of these limitations might be overcome with various technology-specific modifications and variations.

1. The main physical mechanisms that can interfere the effective application of S/S process include
   - a. Incomplete mixing due to the presence of high moisture or organic chemical content, resulting in only partial wetting or coating of the waste particles with the stabilizing and binding agents and aggregation of untreated waste into lumps.
   - b. Disruption of the gel structure of the curing cement or pozzolanic mixture by hydrophilic organics in the waste soil to be treated.
   - c. Undermixing of dry or pasty wastes.
2. The main chemical mechanisms that can interfere with the S/S process include
   - a. chemical adsorption
   - b. precipitation
   - c. nucleation

Additional factors, which can interfere with the S/S/process include

a. the precise tailoring of waste composition to the S/S process used.
b. The treatment of waste containing oil and grease in moderate to high concentrations.

APPLICATION AND PERFORMANCE OF S/S

Types of Applications

The majority of S/S projects at Superfund remedial sites are *ex situ* applications, where inorganic binders and additives were used to treat metal-containing waste. Organic binders were used for treating specialized wastes, such as radioactive wastes, as well as those containing specific hazardous organic compounds. The S/S method was used to treat wastes containing mainly organic contaminants in a small number (6%) of the total number of projects.

Performance

Most performance testing for S/S waste products has been conducted after curing was completed, whereas only limited data are available on long-term performance of S/S at Superfund remedial sites. The available performance data for the behavior of metals in these projects showed that S/S can meet established performance goals. Also, only limited data were available on the behavior of organics; however, the S/S process met the established performance goals for several projects.

SPECIFIC TYPES OF S/S APPLICATIONS

Figure 3 shows a breakdown of S/S projects by specific application type, which include *ex situ* S/S, *in situ* S/S, *ex situ* stabilization only, and *in situ* stabilization only. A specific type of S/S application was identified for 88% of the completed projects. S/S (*in situ* and *ex situ*) represents 63% of the S/S projects, compared to 15% for stabilization only (*in situ* and *ex situ*).

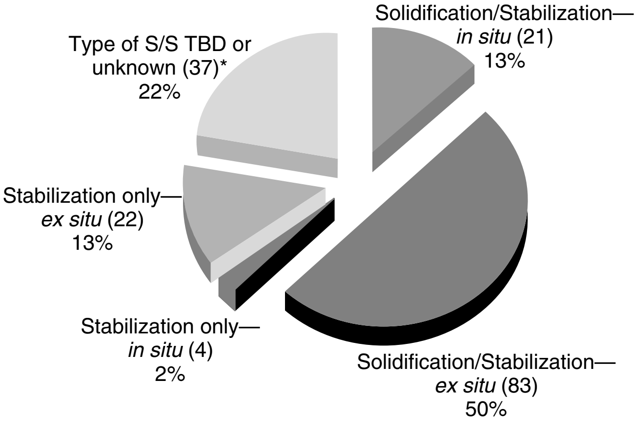

* Includes 3 *in situ* projects.

Note: The number of projects by type is shown in parentheses. TBD= (still) to be determined.

Figure 3. S/S projects by type. Total number of projects: 167.

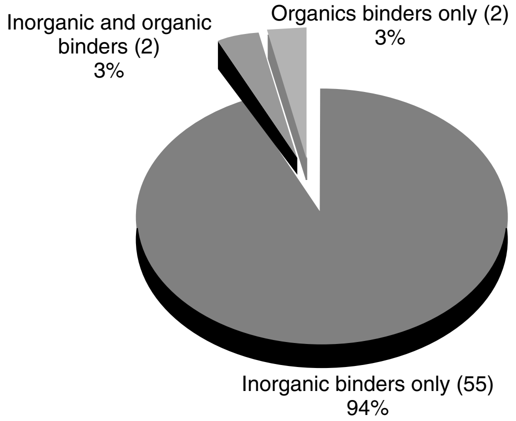

Note: Inorganic binders consist of cement, fly ash, lime, soluble silicates, or sulfur. Organic binders consist of asphalt, organophilic clay, or activated carbon.

Figure 4. Binder materials used for S/S projects. Total number of projects: 59.

Figure 4 shows the types of binders used for S/S projects at Superfund remedial sites, including inorganic binders, organic binders, and appropriate combinations of organic and inorganic binders. Many of the binders used included one or more proprietary additives. Specific examples of inorganic binders included cement, fly ash, lime, soluble silicates, and sulfur-based binders; organic binders included asphalt, epoxide, polyesters, and polyethylene. More than 90% of the S/S completed projects used inorganic binders. In general, inorganic binders are less expensive and easier to use than organic binders, which are generally used to solidify radioactive wastes or other specific hazardous organic compounds.

COST DATA

Information about the cost of using S/S to treat wastes at Superfund remedial sites was available for 29 completed projects. The total costs for S/S projects ranged from \$75,000 to \$16 million, including the cost of excavation, treatment, and disposal (for *ex situ* applications). The cost per m^3 of treated waste ranged from \$10 to approximately \$1500/m^3. The average cost/m^3 for these projects was \$340, including two projects with relatively high costs (approximately \$1500/m^3). Excluding those two projects, the average cost was \$250/m^3 (1).

Additional information (3,4) indicates that the cost of *in situ* S/S can range from as little as \$25 to \$50/m^3 to as much as \$120 to \$250/m^3, depending on the volume to be treated, the structure of the soil (porosity), the treatment depth, the specific type of main contaminant to be stabilized/solidified, and the desired posttreatment objectives (leachability, permeability, or bearing ratio). The low end of the cost range would apply to solidifying, whereas the high end would apply to treating high concentrations of contaminants at great depths. Typical costs for application at a hazardous waste site, consisting of sands to silts, at a depth up to 9 m, has been estimated at around \$90 to \$120/m^3; around 20% would be the cost of reagent/additive.

BIBLIOGRAPHY

1. US EPA. (2000). *Solidification / Stabilization Use at Superfund Sites*. EPA Report 542-R-00-010. Cincinnati, OH.
2. LaGrega, M.D., Buckingham, P.L., and Evans, J.C. (1994). *Hazardous Waste Management*. McGraw-Hill, New York.
3. US EPA. (1997). *Recent Developments for in situ Treatment of Metal Contaminated Soils*. Office of Solid Waste and Emergency Response, Technology Innovation Office, Washington, DC.
4. US EPA. (1997). *Technology Alternatives for the Remediation of Soils Contaminated with As, Cd, Cr, Hg, and Pb*. Engineering Bulletin, EPA Report 540/S-97/500. Cincinnati, OH.
5. US EPA. (1989). *Stabilization / Solidification of CERCLA and RCRA Wastes*. EPA Report. Cincinnati, OH.
6. US EPA. (1992). *Vitrification Technologies for Treatment of Hazardous and Radioactive Waste*. Handbook. Cincinnati, OH.
7. Evanko, C.R. and Dzombak, D.A. (1997). *Remediation of Metals-Contaminated Soils and Groundwater*. Technology Evaluation Report, GWRTAC.
8. US EPA. (1997). *Best Management Practices (BMPs) for Soil Treatment Technologies, Suggested Operational Guidelines to Prevent Cross-Media Transfer of Contaminants During Cleanup Activities*. EPA Report 530-R-97-007. Cincinnati, OH.

WASTEWATER TREATMENT—SMALL SCALE

Guenter Langergraber
BOKU—University of Natural Resources and Applied Life Sciences
Vienna, Austria

Small wastewater treatment plants are defined as those to which fewer than 2’000 people are connected or those whose flow is less than 200 m^3/d. These are treatment plants for small communities, villages, institutions, hotels, and camping resorts.

There are a tremendous number of small wastewater treatment plants worldwide. Small treatment works are needed mainly in rural areas to treat municipal wastewater. However, technologies to treat wastewater from industry are also important. The separate treatment of wastewater from certain processes provides opportunities to reuse water and to reduce the overall consumption of freshwater.

In rural areas and also in large cities that are not yet provided with sewers, sanitation requirements can be met at a reasonable price and on a reasonable timescale by constructing small but highly compact and efficient plants to serve clusters of dwellings.

Small wastewater treatment plants are required to satisfy very different needs which can be integrated into infrastructures at very different stages of development and sophistication. Engineers are called upon to provide technologies that are reliable, easy to operate and control, and tolerant of variations in the influent. The task is difficult and needs an interdisciplinary approach that takes into account local conditions.

REQUIREMENTS FOR SMALL WASTEWATER TREATMENT PLANTS

Rural areas have some general characteristics:

- small communities, groups of houses, and also single houses sometimes where large distances are between them,
- low population density; and
- primarily agricultural use and only little industry.

These properties result in

- highly fluctuating wastewater flows, and
- high concentrations of wastewater constituents with high fluctuations.

Additionally only few trained personnel are available to operate wastewater treatment plants. As a result, some general requirements for small wastewater treatment plants in rural areas can be formulated:

- simplicity of technology,
- simple operation and maintenance,
- high robustness,
- large volume to buffer the high fluctuations of flow and concentrations,
- high stability, and
- low surplus sludge production.

TREATMENT METHODS

Wastewater treatment alternatives (for small communities) can be broadly subdivided into the following categories that represent the basic approaches to wastewater conveyance, treatment, and/or disposal (1):

- *Natural systems* use soil as a treatment and/or disposal medium, including land application, constructed wetlands, and subsurface infiltration. Some sludge and septage handling systems, such as sand drying beds, land spreading, and lagoons, are included.
- *Conventional treatment systems* use a combination of biological, physical, and chemical processes and employ tanks, pumps, blowers, rotating mechanisms, and/or mechanical components as part of the overall system. These include suspended growth, fixed growth, and combinations of the two. This category also includes some sludge and septage management alternatives, such as digestion, dewatering, and composting systems and appropriate disposal.
- *Alternative collecting systems* that use lightweight plastic pipe buried at shallow depths that have fewer pipe joints and less complex access structures than conventional gravity sewers. These include pressure, vacuum, and small-diameter gravity sewer systems.
- *Alternative treatment systems* use source control and separating systems. Wastewater can be separated

into blackwater—waste from toilets—and graywater—the part of the wastewater which is not mixed with excreta (from kitchens, bathrooms, and laundries). Blackwater can further be separated into urine (yellowwater) and feces (brownwater).

Depending on the local situation, different requirements arise for the treatment of organic matter, nutrients, and pathogens. In principle, all treatment methods that are used for large-scale wastewater treatment can be applied in small wastewater treatment plants. However, treatments that meet the special requirements of small systems should be favored.

Table 1 summarizes treatment alternatives to fulfill different effluent requirements. In many countries, systems for partial treatment combined with chemical precipitation are the only treatment and may serve as pre-treatment for additional future treatment steps. It is well understood that, in most cases, the relatively low removal rates in primary clarifiers and septic tanks do not comply with treatment goals and that septic tanks, especially, are perfectly suited for pre-treatment in infiltration systems and constructed wetlands. Although construction and operation of anaerobic pre-treatment is simple, only appropriate sludge control, it has been shown, guarantees settling of solids and partial hydrolysis by methanogenic fermentation. In all other treatment systems, biochemical oxidation is the most important process and determines, to a great extent, the size and costs of the treatment system.

Pathogen removal in nitrifying systems is generally higher than that in nonnitrifying systems. If very low microbial contamination is required, disinfection methods have to be applied.

Table 2 classifies treatment systems as low and high rate systems. In low rate systems, only lower hydraulic and contaminant loading rates can be applied, oxygen transfer is limited to natural diffusion and convection processes, the biomass concentration is lower, and therefore the specific volume and area requirements are higher. However, they are easier to operate and maintain and require, therefore, less skilled personnel.

The advantages of using natural systems are their “low-tech/no-tech” nature, which means that these systems

Table 1. Treatment Alternatives to Fulfill Different Effluent Requirements[a]

Treatment Requirement	Treatment Alternatives
Partial treatment (pre-treatment)	• Primary clarifier • Septic tank • Imhoff tank • Anaerobic pond • Surface-flow constructed wetlands
Physicochemical treatment	• Chemical precipitation combined with partial treatment one- or two-stage (+flocculation)
Complete organic matter removal	• Nonaerated and aerated pond • Rotating biological contactor • Trickling filter • Biofilter • Moving bed reactor • Subsurface horizontal-flow constructed wetlands • (Activated sludge)
including phosphorus removal	• Chemical precipitation in combination with most other technologies
including nitrification	• Extended aeration • Sequencing batch reactors (SBR) • Rotating biological contactor • Trickling filter • Biofilter • Moving bed reactor • Subsurface vertical-flow constructed wetlands • Sand infiltration
including denitrification	• Activated sludge • Sequencing batch reactors (SBR) • Rotating biological contactor • Trickling filter • Biofilter • Moving bed reactor • Subsurface-flow constructed wetlands
Pathogen removal	• Activated sludge, sequencing batch reactors (SBR) • Soil infiltration • Subsurface vertical-flow constructed wetlands • Disinfection (UV, chlorine, ozone)

[a] References 2 and 3.

Table 2. Low and High Rate Treatment Systems[a]

Low Rate Systems	High Rate Systems
• Sand infiltration • Soil infiltration • Constructed wetlands • Ponds	• Conventional activated sludge systems • Sequencing batch reactors (SBR) • Rotating biological contactor • Trickling filter • Biofilter • Moving bed reactor • Extended aeration

low	←	loading rates	→	high
low	←	oxygenation rates	→	high
low	←	oxidation rates	→	high
low	←	biomass concentration	→	high
high	←	volume and area requirements	→	low
low	←	skilled control	→	high

[a] Reference 2.

are relatively easy to construct and operate; and their low costs (especially operating costs), which makes them attractive to communities with a limited budget. However, their simplicity and low cost may be deceptive in that the systems require frequent inspections and constant maintenance to ensure smooth operation (low maintenance does not mean no maintenance).

Mainly because of the land requirements for natural systems, many communities prefer technical systems, which tend to require less land and permit better control of the operation. However, these systems generally are high cost and require more skilled personnel to operate them. The disadvantages generally relate to the cost of construction and ease of operation. Mechanical systems can be costly to build and operate because they require specialized personnel. Generally, the complexity and cost of wastewater treatment technologies increase with the quality of the effluent produced.

Table 3 compares the specific reactor area and volume of systems used in small wastewater treatment plants. Additionally, the capability of the systems to nitrify is shown.

All systems described can be used for decentralized or 'on-plot' treatment and centralized or 'off-plot' treatment. For decentralized solutions, safe disposal of excreta takes place on or near a single household or a small settlement. Centralized solutions collect excreta from individual houses and carry them away from the plot to be treated off-site. The selection of the most appropriate sanitation system is influenced by ecological, technical, social, cultural, financial, and institutional factors.

COSTS

The costs of a wastewater treatment plant can be subdivided into investment and operating costs. To include the payback of the investment, the investment costs are transformed into yearly costs (using an economical interest rate; 3.5% is used in the examples given below). In the examples, the assumed lifetime of the treatment system (SBR and constructed wetland) is 20 years; for the sewer system, a lifetime of 40 years is assumed.

Table 4 compares specific investment, operating, and yearly costs of a technical treatment system (SBR) and a constructed wetland (subsurface vertical-flow, intermittently loaded). Additionally, two sizes of the treatment plant are compared, 100 and 500 PE (people equivalent). Both treatment systems have to fulfill the same effluent requirements (nitrification); the same sewer line configuration is used in both cases. The investment and operating costs for the sewer are estimated at about $1500\ \mathrm{EUR} \cdot \mathrm{PE}^{-1}$ and $5\ \mathrm{EUR} \cdot \mathrm{PE}^{-1} \cdot \mathrm{yr}^{-1}$, respectively. The investment costs for the two treatment systems are similar; specific costs increase as plant size decreases. However, the constructed wetland system shows lower operating and therefore lower yearly costs.

Seven different treatment scenarios for individual solutions for wastewater treatment of a single house (5 PE) are compared:

1. *Technical treatment system*: A conventional technical treatment system is used, and the effluent is discharged.

Table 3. Specific Reactor Area and Volume of Systems Used in Small Wastewater Treatment Plants[a]

System	Area/PE, m^2	Volume/PE, m^3	Contact Time	Nitrification
Nonaerated pond	10	10–15	>20 d	—
Soil infiltration	4–20	7–12	—	+
Surface-flow constructed wetland	5–10	5–15	>10 d	—
Aerated pond	3	4–7	>3–6 d	±
Subsurface horizontal-flow constructed wetland	4–8	3–5	1–4 d	—
Buried sand filter	4–6	4–6	1 h–7 d	+
Subsurface vertical-flow constructed wetland	2–5	1–3	1–3 d	+
Nitrifying trickling filter	0.17–0.3	0.45–0.6	6–10 min.	+
Extended aeration	0.12–0.25	0.35–0.6	1–3 d	+
Sequencing batch reactors (SBR)	0.1–0.2	0.3–0.5	1–3 d	+
Nitrifying rotating biological contactor	0.1–0.18	0.17–0.25	10–20 h	+
Nonnitrifying trickling filter	0.05–0.08	0.13–0.18	3–6 min	—
Nonnitrifying rotating biological contactor	0.04–0.07	0.07–0.13	8–15 h	—
Nitrifying biofilter	0.005–0.01	0.02–0.03	30–50 min.	+
Nonnitrifying biofilter	0.005–0.01	0.013–0.03	20–40 min.	—

[a] Reference 2, revised.

Table 4. Comparison of Specific Investment, Operating, and Yearly Costs for a Technical Treatment System (SBR) and a Constructed Wetland[a]

System		SBR		Constructed Wetland	
Size of the Treatment Plant	PE	500	100	500	100
Investment costs	$EUR \cdot PE^{-1}$	430	770	420	690
Operating costs	$EUR \cdot PE^{-1} \cdot yr^{-1}$	22	44	15	36
Yearly costs	$EUR \cdot PE^{-1} \cdot yr^{-1}$	128	174	120	160

[a] Reference 4, modified.

2. *Constructed wetland*: A constructed wetland system is used instead of the technical solution. The effluent is discharged.
3. *Cesspit and disposal* of cesspit waste: All wastewater is collected in a cesspit, and the cesspit waste is disposed of periodically at a wastewater treatment plant.
4. *Cesspit and agricultural use* of cesspit waste: The same as No. 3, but the cesspit waste is used in agriculture.
5. *Cesspit for blackwater (disposal* of cesspit waste) and constructed wetland for *graywater treatment*: Blackwater and graywater are separated. Only blackwater is collected in the cesspit and disposed of periodically. Graywater is treated using a constructed wetland.
6. *Cesspit for blackwater* (*agricultural use* of cesspit waste) and constructed wetland for *graywater treatment*: The same as No. 6, but the cesspit waste is used in agriculture.
7. *Urine separation* (agricultural use of urine and feces) and constructed wetland for graywater treatment: Separation of urine and feces; both are used in agriculture. Graywater is treated using a constructed wetland.

Table 5 compares the investment, operating, and yearly costs for different treatment scenarios.

The constructed wetland system has yearly costs lower than to the conventional technical treatment system. When all wastewater is collected in a cesspit, the yearly costs of the scenario for agricultural use of the cesspit waste are only about 75% of the yearly costs when disposing of the waste at a wastewater treatment plant. However, all scenarios using source separation have the lowest operating and yearly costs. Separating toilet water from graywater leads to a tremendous reduction of the volume that has to be collected, and, therefore, the operating and also the yearly costs drop drastically. Urine separation has the lowest costs and, additionally, closes water and nutrient cycles on a local scale. Therefore, it is a promising system toward more ecologically sound sanitation.

SUMMARY

Wastewater treatment plants for small communities, groups of houses, and also single houses have to handle special wastewater characteristics such as highly fluctuating wastewater flows and high concentrations of wastewater constituents (and additionally with high fluctuations). Therefore, treatment systems for small-scale applications should have high process robustness and stability and a large buffer volume. Additionally, they should have simple technology, they should be easy to operate and maintain, and should have low surplus sludge production.

Table 5. Comparison of Investment, Operating, and Yearly Costs for the Treatment Alternatives for a Single Household with 5 PE[a]

Scenario		1	2	3	4	5	6	7
System[b]		Technical System	CW	CP	CP	CP (BW)	CP (BW)	US
Disposal of cesspit waste				WWTP	AU	WWTP	AU	AU
Separation of black-/graywater		No	No	No	No	Yes	Yes	Yes + US
Graywater treatment						CW	CW	CW
				Investment costs				
Treatment unit	$EUR.PE^{-1}$	1450	1450	1780	1780	1120	1120	1160
Sewer	$EUR.PE^{-1}$	350	350	230	230	410	410	290
				Operating costs				
Treatment unit	$EUR.PE^{-1}.yr^{-1}$	240	170	370	230	160	130	90
Sewer	$EUR.PE^{-1}.yr^{-1}$	5	5	5	5	5	5	5
Yearly costs	$EUR.PE^{-1}.yr^{-1}$	**362**	**292**	**468**	**336**	**246**	**208**	**192**

[a] Reference 4, modified.

[b] CW: constructed wetland; CP: cesspit; CP(BW): cesspit only for blackwater; WWTP: wastewater treatment plant; AU: agricultural use; US: urine separation.

Different requirements for treatment efficiency arise, depending on the local situation. Several treatment methods can be used to meet these requirements. The selection of the most appropriate sanitation system is influenced by ecological, technical, social, cultural, financial, and institutional factors.

For a small wastewater treatment plant, especially, low operating costs are essential. It was shown that for single households, solutions with source control separating at least blackwater and graywater yield significantly lower costs compared to solutions where the different types of wastewater are mixed and, therefore, a large volume has to be treated.

BIBLIOGRAPHY

1. Lens, P., Zeeman, G., and Lettinga, G. (Eds.). (2001). *Decentralized Sanitation and Reuse-Concepts, Systems and Implementation*. IWA Publishing, London, UK.
2. Boller, M. (1997). Small wastewater treatment plants—a challenge to wastewater engineers. *Water Sci. Technol.* **35**(6): 1–12.
3. US EPA. (2002). *Onsite Wastewater Treatment Systems Manual*. Report No. EPA/625/R-00/008. US EPA, Cincinnati, OH.
4. BMLFUW. (2002). Kosten-Nutzen-Überlegungen zur Gewässerschutzpolitik in Österreich mit besonderer Berücksichtigung des ländlichen Raumes (Cost-benefit analysis of water pollution control in Austria with a special focus on rural areas). *Schriftenreihe des BMLFUW*. Vienna, Austia [*in German*].

READING LIST

Lange, J. and Otterpohl, R. (2000). *Abwasser—Handbuch zu einer zukunftsfähigen Wasserwirtschaft* (*Wastewater—Manual for sustainable water management*), 2nd Edn. MALLBETON-Verlag, Donaueschingen-Pfohren, Germany [*in German*].

Watanabe, Y. (Ed.). (1997). Small wastewater treatment plants III. *Water Sci. Technol.* **35**(6).

Wilderer, P.A. and Deininger, A. (Eds.). (2000). Small wastewater treatment plants IV. *Water Sci. Technol.* **41**(1).

MICROBIAL FOAMING AND BULKING IN ACTIVATED SLUDGE PLANTS

Fiona M. Stainsby
James C. Philp
Sandra Dunbar
Napier University
Edinburgh, Scotland, United Kingdom

Irena B. Ivshina
Maria S. Kuyukina
Russian Academy of Sciences
Perm, Russia

BACKGROUND

Microbial foaming and bulking are among the most frequent and widespread problems in activated sludge (AS) wastewater treatment plants (WWTPs). Bulking has plagued AS plants almost since their inception nearly a century ago, whereas microbial foaming was not documented until around 30 years ago. Both problems are associated with the excessive growth of various filamentous bacteria, although the mechanisms by which they form are not fully understood. Initial attempts to identify these bacteria shifted attention from engineering aspects to the microbiological activities fundamental to the activated sludge process. Subsequently, filament identification keys and abundance scales became routine tools for monitoring filamentous populations in activated sludge. More recently, molecular biological techniques have enabled more detailed and precise study of the diversity and ecology of bacterial communities associated with foaming and bulking.

CURRENT REGULATION OF WASTEWATER TREATMENT

Within the European union, treated effluents must comply with the standards set out in the EU Urban Wastewater Treatment Directive (1). The Directive states the minimal standards for the treatment and disposal of wastewater and waste sludge via WWTPs. These standards include the maximum biological oxygen demand (BOD; 20 mg/L) and level of suspended solids (SS; 30 mg/L), known as the 20:30 standard, allowed in treated effluents (2). In AS plants, the quality of the final effluent is governed by factors that affect both biological oxidation in the aeration basins and sludge separation in the final clarification tanks.

FOAMING AND BULKING ARE SLUDGE SEPARATION PROBLEMS

Central Role of Activated Sludge Floc

Activated sludge is comprised of three-dimensional aggregate structures, known as flocs, that contain microorganisms and abiotic matter. The structure of these flocs, including firmness, shape, and size determines the settling and compacting characteristics of the sludge solids. In ‘healthy’ flocs, filamentous bacterial growth creates a framework onto which other floc-forming bacteria can attach, and provides the strength required for the floc to remain intact in the turbulent environment of the aeration basin (3). However, excessive growth of various types of filaments can alter floc structure and interfere with the separation and recycling of sludge solids; this is observed in both microbial foaming and bulking.

Mechanism of Foam Formation

Microbial foam manifests on the surfaces of activated sludge aeration tanks as a dense, frothy, grey-brown, viscous scum, often described as having a ‘chocolate mousse’ appearance (Fig. 1). The majority of microbial foams are extremely stable and persistent and can accumulate to depths of more than a meter (4).

Foams consist of aggregates of air bubbles in liquid, flocs, and large numbers of filamentous bacteria, which selectively accumulate on the surface of mixed liquor

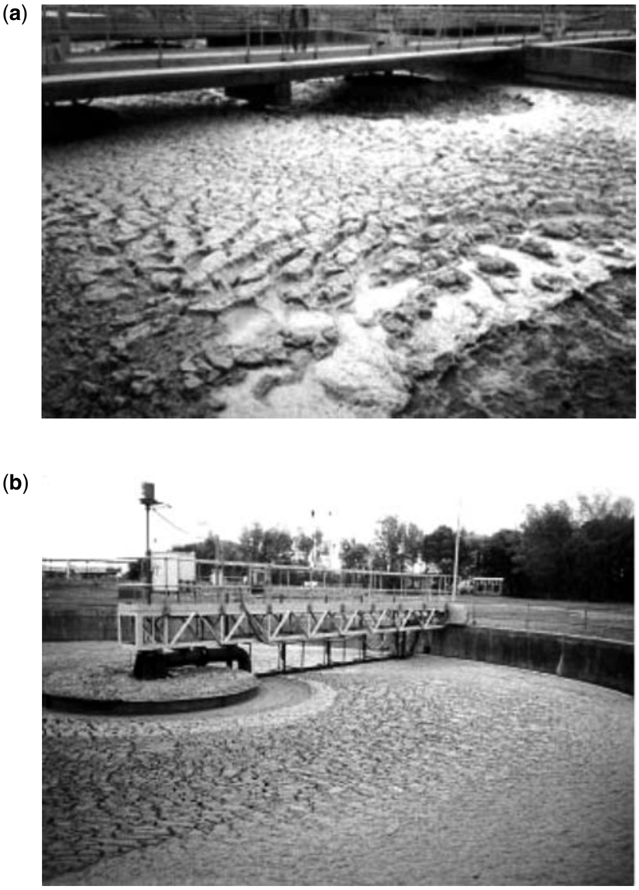

Figure 1. Microbial foam on the surface of (**a**) a mechanically agitated aeration basin and (**b**) a final clarifier, where it has partially dried out placing the tank out of operation.

by flotation. It is thought that excessive growth of hydrophobic bacteria in the mixed liquor renders flocs hydrophobic and hence amenable to attachment to air bubbles that are introduced to the mixed liquor by aeration. Large open flocs, it is thought, require only a few hydrophobic sites to adhere to air bubbles (5). The air bubble–floc aggregates have a lower density than water and rapidly rise to the surface of aeration basins where they accumulate. Hydrophobic solids stabilize the foam by bridging the bubble lamellae, preventing liquid drainage and subsequent breakdown.

Microbial Bulking

Bulking is caused by extensive and uncontrolled growth of various filamentous bacteria, both within and extending from the confines of flocs into the bulk liquor (Fig. 2). This results in a diffuse, open floc structure and interfloc bridging that can interfere with the settling and compaction of sludge solids in the final clarifiers. The supernatant is usually clear because bulking flocs filter out suspended particles (6). Although a blanket of scum may appear on the mixed liquor surface, bulking sludge is generally less easy to detect visually than foam. Several empirical methods have been devised and are widely used in the wastewater industry to monitor sludge characteristics and diagnose bulking. These include the diluted sludge volume index (DSVI; 7) total extended filament length (TEFL), and subjective filament abundance scales (4,8).

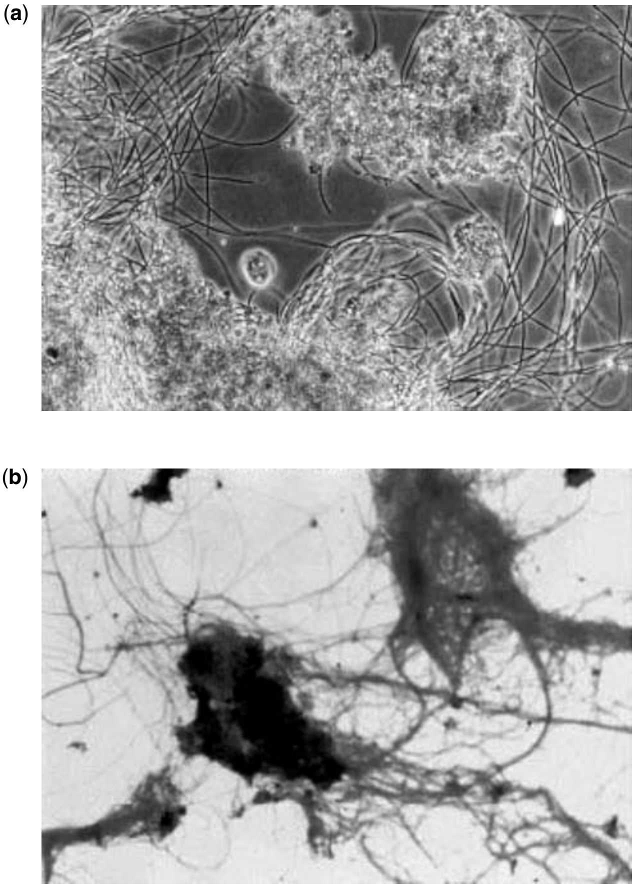

Figure 2. Microscopic examination of bulking activated sludge: (**a**) wet mount preparation viewed under phase contrast (×330 magnification) showing excessive growth of Eikelboom type 021N extending into the bulk phase and bridging adjacent flocs and (**b**) Gram-stained preparation viewed under bright-field (× 330) showing 021N filaments stained gram-negative forming large entangled knots (upper right) and surrounding the flocs (lower left).

Problems Caused by Foaming and Bulking

Foaming and bulking lead to a number of problems that may compromise operational efficiency and effluent quality and have inevitable financial and regulatory repercussions:

- Overflow of foam or blanket scum to the final clarifiers may lead to increased BOD and SS levels in the final effluent.
- Loss of solids means increased substrate-loading rates per unit of sludge in the aeration basins.
- Poor sludge settling and compaction caused by bulking affects the dewatering characteristics of waste activated sludge.

- Thick foams inhibit mixing and oxygen transfer at the surfaces of mechanically aerated basins, thereby restricting biological oxidation.
- Spillage of foam onto plant walkways can create hazardous working conditions, limit access to sections of the treatment unit, and create cleaning and odor problems.
- Release of poorly clarified effluent to receiving waters may disperse enterobacteria and viruses that are pathogenic to humans or animals.
- Pathogens may be carried in aerosols arising from windblown scum, which may be inhaled or ingested by plant workers or local inhabitants (9).

PREVENTION AND CONTROL OF FOAMING AND BULKING

There are no definitive procedures for preventing and controlling either foaming or bulking. Most proposed measures are derived from experiences within one plant or laboratory scale system or from questionnaire-based surveys and hence not universally applicable (4,10–12). Chemical treatment methods, such as chlorine or hydrogen peroxide dosing, are only temporary measures; these oxidants are nonselective; therefore, organisms that are beneficial to the AS process may also be adversely affected. Antifoaming agents, bioadditives containing enzymes, microorganisms and nutrients, coagulants and iron salts, used to treat foams, vary in success but are generally not cost-effective. Filaments implicated in foaming and bulking are considered slow growers and may be washed out of the AS system by reducing the sludge age or mean cell residence time (MCRT). This approach is usually not suitable in plants where nitrification is essential because the MCRT necessary for nitrifying bacteria exceeds that required to washout filaments. Selector technology is an engineering approach, whereby modified or additional bioreactor tanks impose environmental conditions that place undesirable filaments at a competitive disadvantage to other floc-formers (13,14). However, selector tanks are not effective against all foaming or bulking filament types as these can vary considerably in their metabolic activities.

BACTERIA INVOLVED IN FOAMING AND BULKING

Filament Identification Schemes

Eikelboom (15) published the first key recommended for identifying filaments in activated sludge. In this scheme, filaments were identified by taxonomic name or given a morpho-type number and assigned to groups based on morphology, staining properties, example, Gram's stain, and the occurrence of cell inclusions, example, sulfur granules, determined by microscopy. This artificial classification, used in conjunction with filament abundance scales, has enabled comparative studies of filament composition and floc structure in AS plants in different regions of the world, under different plant operating conditions, and during bulking and foaming episodes. Microscopic examination is relatively simple and inexpensive, and regular monitoring can provide an early indication of foaming and bulking (16,17). Several additional schemes are available for filament identification (4,8,18), including color posters (19) and computer software formats (20).

The Usual Suspects

Numerous foaming and bulking plants have been surveyed worldwide to establish the relative abundance of different filament types. The same filaments are often reported in different geographical regions, though their relative abundance may vary, and several filaments occur in both foaming and bulking sludges. Foaming was originally attributed to gram-positive branched filaments identified as members of the genus *Nocardia* (21). Subsequent studies have reported many other species related to *Nocardia* in foams, including nonfilamentous types; these organisms are phylogenetically related and are termed collectively mycolic acid-containing actinomycetes (22,23). The gram-positive, unbranched filament *Microthrix parvicella* (24) has also been reported in foaming AS plants in many parts of the world (Fig. 3).

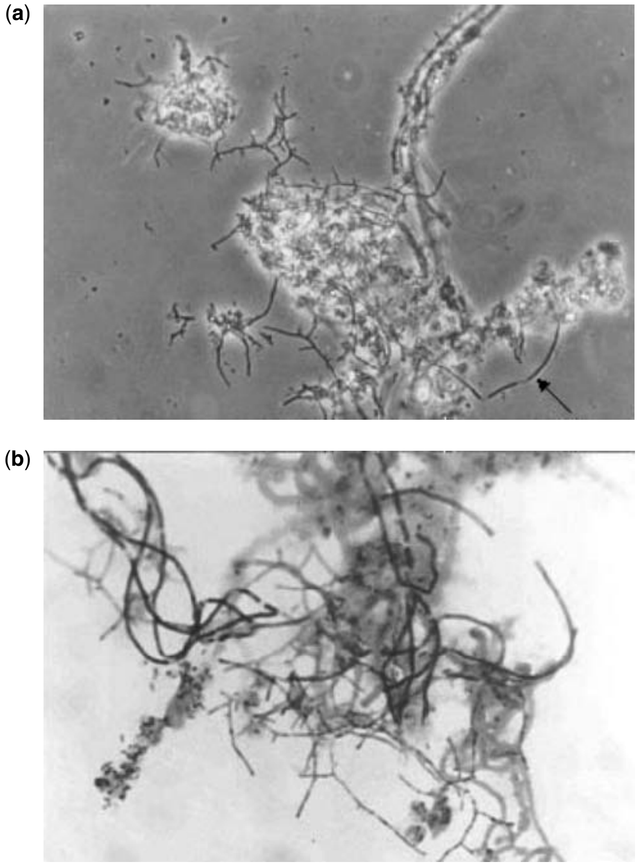

Figure 3. Microscopic examination of foaming activated sludge: (**a**) Wet mount preparation viewed under phase contrast (× 1000 magnification) and (**b**) Gram-stained preparation viewed under bright-field (× 1250). The dominant filaments are the spaghetti-like *Microthrix parvicella* with empty sections in trichome (arrow) and right-angled branched filaments with irregular staining patterns identified as mycolic acid-containing actinomycetes.

Other dominant filaments include *Halisocomenobacter hydrossis*, *Nostocoida limicola*, *Sphaerotilus* spp., and Eikelboom types 1863, 0092, 0041, and 0675 (13).

Filaments regularly associated with bulking in various parts of the world include *H. hydrossis*, *M. parvicella*, *Nocardia* spp., *N. limicola*, *Sphaerotilus natans, Thiothrix* spp., and types 021N, 1701, 0041, 0092, and 0675 (4,16). Geographical variations in filament populations may reflect differences in climate, wastewater composition, plant design, and operating conditions.

Isolation and Characterization of Filaments

Filament identification schemes provide a simplified view of the bacterial communities in activated sludge. Few bacteria can be accurately identified using only microscopy. Selective isolation studies, whereby dilutions of activated sludge or individual filaments extracted by micromanipulation are plated onto various nutrient media (23,24), enable the cultivation of bacteria in the laboratory for further characterization and identification. Studies that examine the genetic, phenotypic, and physiological traits of such organisms reveal that filaments with identical morphologies may be only distantly related in evolutionary terms and have quite different metabolic requirements (25,26). This, in turn, has profound implications for the development of engineering strategies to the control foaming and bulking organisms.

FACTORS AFFECTING THE GROWTH OF FOAMING AND BULKING ORGANISMS

Plant operating parameters, including mode of aeration, feed to mass ratio, and sludge retention time or MCRT, reportedly influence the proliferation of particular filament types (4). The composition of influent wastewater, substrate concentrations, and abiotic factors such as dissolved oxygen concentration, ambient temperature, mixed liquor pH, and redox potential are also important. The growth rates of bacteria and hence their ability to remain within the AS system also depend on their nutrient requirements and metabolic versatility. Mycolic acid-containing actinomycetes provide a good example because they can metabolize a diverse range of substrates. Many members of this group also produce biosurfactants (27,28), which aid foam stabilization and provide selective access to oils and fats prevalent in wastewater (29), which gives them a competitive advantage over many other bacteria.

MOLECULAR BIOLOGICAL TOOLS

Molecular biological techniques developed in the last decade are providing new insights into the complexity, dynamics, and ecophysiology of bacterial populations implicated in foaming and bulking. Fluorescent *in situ* hybridization (FISH) enables detecting and enumerating of particular phylogenetic groups of bacteria in activated sludge, irrespective of their morphology (30). Fluorescent-labeled nucleic acid probes can be designed to target cells of the whole bacterial community or at different phylogenetic levels down to individual species or even subspecies (31–33). Quantitative microautoradiography (QMAR) used together with FISH, allows *in situ* measurement of the uptake of specific radiolabeled substrates by individual bacterial cells that belong to particular phylogenetic groups (34). Information on the substrate utilization kinetics of different filamentous bacteria will help us to understand why they predominate in bulking or foaming AS plants and the operating conditions under which this occurs.

EXAMPLE: MONITORING FOAM-FORMING BACTERIA USING FISH

The micrograph image in Fig. 4a shows an example of FISH using two fluorescent-labeled probes. The technique involves fixation of an activated sludge sample to preserve the bacterial populations present, followed by permeabilization of the cells to allow the probes to enter and hybridize with target nucleic acid sequences. Treated samples are mounted on microscope slides and viewed by epifluorescence microscopy using a confocal laser scanning microscope (CLSM). Images of randomly selected fields of view can be captured, and the cells in a known volume

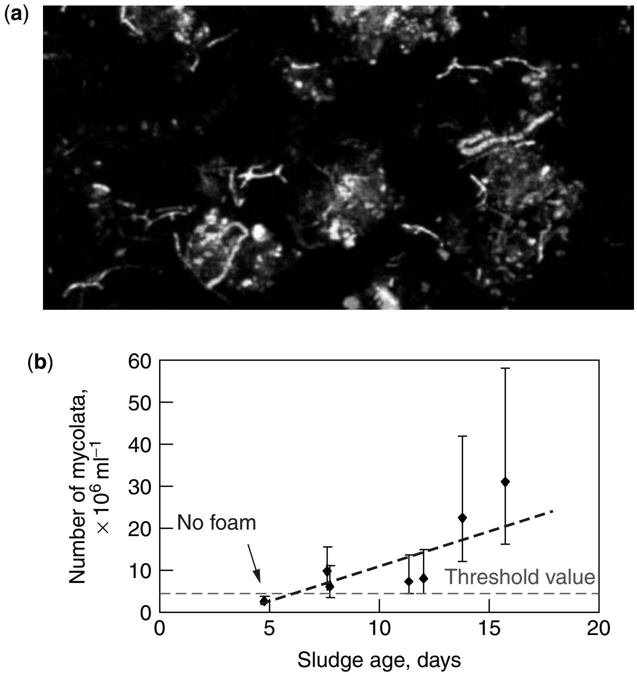

Figure 4. Fluorescent *in situ* hybridization: (**a**) Confocal laser scanning micrograph of activated sludge taken from a foaming WWTP. The sample is hybridized with two nucleic acid probes, a eubacterial-specific probe (Bact338) labeled with fluorescein isothiocyanate (FITC), which fluoresces green, and a mycolata-specific probe (Myc657) labeled with tetrarhodamine isothiocyanate (TRITC), which fluoresces red. Only bacterial cells that hybridize with both probes and appear yellow-orange in color are mycolata (mycolic acid-containing actinomycetes). (**b**) Plot showing the relationships between sludge age, numbers of mycolic acid-containing actinomycetes determined using quantitative FISH and foaming in an AS reactor. The error bars indicate 95% confidence intervals (courtesy of R. J. Davenport, unpublished data; 2003).

of activated sludge calculated using computer software. FISH has been used to monitor changes in the abundance of mycolic acid-containing actinomycetes in an AS plant operated at different sludge ages. The graph in Fig. 4b shows that a significant rise in the numbers of these organisms occurs with increasing sludge age and that there is a threshold in numbers that triggers foaming (32).

BIBLIOGRAPHY

1. European Commission Council Directive. Official Journal L135. (1991). Urban Waste Water Treatment Directive 91/271/EEC. EEC, Brussels, Belgium.
2. Gray, N.F. (1989). *Biology of Wastewater Treatment*. Oxford University Press, Oxford, UK.
3. Li, D-H. and Ganczarczyk, J.J. (1990). *Biotechnol. and Bioeng.* **35**: 57–65.
4. Jenkins, D., Richard, M.G., and Daigger, G.T. (1993). *Manual on the Causes and Control of Activated Sludge Bulking and Foaming*, 2nd Edn. Lewis, Boca Raton, FL.
5. Lemmer, H., Lind, G., Müller, E., Schade, M., and Ziegelmayer, B. (2000). *Acta Hydrochima et Hydrobiologica* **28**: 34–40.
6. Blackall, L.L. (1999). Bulking. In: *The Microbiology of Activated Sludge*. R.J. Seviour and L.L. Blackall (Eds.). Chapman & Hall, London, pp. 147–160.
7. Wanner, J. (1994). *Water Sci. Technol.* **29**: 193–203.
8. Eikelboom, D.H. and van Buijsen, H.J.J. (1983). *Microscopic Sludge Investigation Manual*, 2nd Edn. TNO Research Institute for Environmental Hygiene, Delft.
9. Carducci, A. et al. (2000). *Water Res.* **34**: 1173–1178.
10. Blackall, L.L., Harbers, A.E., Greenfield, P.F., and Hayward, A.C. (1991). *Water Res.* **25**: 313–317.
11. Foot, R.J., Robinson, M.S., and Forster, C.F. (1994). *Water Sci. Technol.* **29**: 213–220.
12. Pitman, A.R. (1996). *Water Sci. Technol.* **34**: 291–298.
13. Soddell, J.A. (1999). Foaming. In: *The Microbiology of Activated Sludge*. R.J. Seviour and L.L. Blackall (Eds.). Kluwer, Boston, pp. 161–202.
14. Grady, C.P.L. and Filipe, C.D.M. (2000). *Water Air Soil Pollut.* **123**: 117–132.
15. Eikelboom, D.H. (1975). *Water Res.* **9**: 365–388.
16. Eikelboom, D.H., Andreadakis, A., and Andreasen, K. (1998). *Water Sci. Technol.* **37**: 281–289.
17. Madoni, P., Davoli, D., and Gibin, G. (2000). *Water Res.* **34**: 1767–1772.
18. Wanner, J. and Grau, P. (1989). *Water Res.* **23**: 883–891.
19. Seviour, E.M., Lindrea, K.C., Seviour, R.J., and Soddell, J.A. (1997). Colour Posters to Assist in the Identification of Filamentous Bacteria and Protozoa in Activated Sludge Biomass. *Proc. 2nd Int. Conf. Microorganisms Activated Sludge Biofilm Processes*. July 21–23, Berkeley, CA, pp. 633–635.
20. de Arcaya Alvarez, J.M., Blanco, J.M., Urbano, C., and Palmer, A. (1997). A Computer Program for the Identification of Activated Sludge Microorganisms. *Proc. 2nd Int. Conf. Microorganisms Activated Sludge Biofilm Processes*. July 21–23, Berkeley, CA, pp. 67–670.
21. Lechevalier, M.P. and Lechevalier, H.A. (1974). *Int. J. Syst. Bacteriol.* **24**: 278–288.
22. Soddell, J.A. and Seviour, R.J. (1998). *J. Appl. Microbiol.* **84**: 1–13.
23. Stainsby, F.M., Soddell, J., Seviour, R., Upton, J., and Goodfellow, M. (2002). *Water Sci. Technol.* **46**: 81–90.
24. Blackall, L.L. et al. (1996). *Int. J. Syst. Bacteriol.* **46**: 344–346.
25. Seviour, E.M. et al. (1997). *J. Appl. Microbiol.* **82**: 411–421.
26. Kanagawa, T. et al. (2000). *Appl. Environ. Microbiol.* **66**: 5043–5052.
27. Lang, S. and Philip, J.C. (1998). *Antonie Leeuwenhoek* **74**: 59–70.
28. Philp, J.C. et al. (2002). *Appl. Microbiol. Biotechnol.* **59**: 318–324.
29. Soddell, J.A. and Seviour, R.J. (1996). *Water Sci. Technol.* **34**: 113–118.
30. Davenport, R.J. and Curtis, T.P. (2002). *Water Sci Technol.* **46**: 529–533.
31. Wagner, M., Amann, R., Lemmer, H., Manz, W., and Schleifer, K-H. (1994). *Water Sci. Technol.* **29**: 15–23.
32. Davenport, R.J., Curtis, T.P., Goodfellow, M., Stainsby, F.M., and Bingley, M. (2000). *Appl. Environ. Microbiol.* **66**: 1158–1166.
33. de los Reyes, F.L. and Raskin, L. (2002). *Water Res.* **36**: 445–459.
34. Nielsen, J.L., Christensen, D., Kloppenborg, M., and Nielsen, P.H. (2003). *Environ. Microbiol.* **5**: 202–211.

AQUEOUS BEHAVIOR OF ELEMENTS IN A FLUE GAS DESULFURIZATION SLUDGE DISPOSAL SITE

Frank J. Castaldi
Brown and Caldwell
Austin, Texas

FLUE GAS DESULFURIZATION PROCESS DESCRIPTION

Most conventional combustion processes emit sulfur oxides, nitrogen oxides, carbon oxides, particulate matter, and other potentially harmful pollutants into the air. Solid wastes from combustion processes or from associated control technologies present disposal and environmental health problems. Adverse water-related health and ecological effects might result when chemical compounds and metals are leached from solid residues.

Flue gas desulfurization (FGD) systems are end of process pollution abatement techniques used primarily for removing sulfur dioxide from coal-fired utility and industrial boiler combustion gases. The technology has also been successfully applied to Claus sulfur recovery unit tail gas and sulfuric acid plant tail gas streams. Sulfur dioxide (SO_2) and particulate emissions are controlled by a flue gas scrubber system. The scrubber is often designed as an integral part of the steam generation plant. It consists of a two-stage venturi-absorber scrubber module designed to treat boiler flue gas. The flue gas from the boiler enters the air quality control system through a common plenum. The gas stream is sent to the individual scrubber modules from this plenum.

The sulfur-laden flue gas first passes through an electrostatic precipitator or venturi scrubber for particulate removal and then into a multistage absorber where it contacts the absorption slurry. Scrubbed gas passes through a demister and is reheated prior to discharge to the atmosphere to provide buoyancy and prevent condensation. Sulfur dioxide removal is normally 85–90%. Up to 99% of the particulate matter is removed from the gas stream at this point (1). The particles are entrained in the liquid that drops into the bottom sump. The gases then pass through the SO_2 absorber section. The SO_2 is removed by adsorption as the gas stream is drawn through stainless steel sieve trays that are sprayed with the absorption slurry. This slurry with the absorbed SO_2 also drops into the bottom sump. The gas stream then passes through a demister section in which excess moisture and mist are removed. Then, the gases are reheated to increase the gas temperature for improved gas buoyancy and to reduce the probability of deposits on the induced air fans. Upon exiting the scrubber, the gases are sent to the stack.

Wet desulfurization processes produce large quantities of nonregenerable waste. Wet scrubbing processes may use calcium carbonate ($CaCO_3$), calcium oxide (CaO), alkaline fly ash, and sodium carbonate (Na_2CO_3) as scrubbing materials (2,3). Limestone-based FGD systems are nonregenerable processes where the reactive component in the absorbent slurry combines chemically with SO_2 to form a sludge that consists of fly ash, water, calcium sulfate, and calcium sulfite. The calcium sulfate and calcium sulfite sludges are waste by-products. Sodium-based scrubbing systems produce sludge that typically contains unspent scrubber material, sodium sulfite and sulfate, and entrained fly ash. Unlike calcium-based systems, the waters that characterize sodium-based FGD processes and their associated sludge have high concentrations of dissolved solids dominated by sodium and sulfate ions. Sodium-based FGD processes and their associated sludge pond systems may also generate hydrogen sulfide as a by-product, which could be emitted into the atmosphere.

LIMESTONE/LIME-BASED FGD SYSTEMS

In limestone-based FGD systems, ground limestone mixed with water is used for making up the scrubber slurry. The slurry water is recycled from settling ponds where solids accumulate. The SO_2 and sulfur trioxide (SO_3) are removed from the flue gas stream by reaction with the aqueous solution of limestone. The major component of limestone is $CaCO_3$. The products from the SO_2 reaction are carbonic acid (H_2CO_3) and calcium sulfite hemihydrate ($CaSO_3{\bullet}\frac{1}{2}H_2O$). The products from the SO_3 reaction are carbonic acid and calcium sulfate dihydrate ($CaSO_4{\bullet}2H_2O$). The reactions are presented here:

$$SO_2 + 1.5\ H_2O + CaCO_3 \longrightarrow H_2CO_3 + CaSO_3{\bullet}\tfrac{1}{2}\ H_2O$$

$$SO_3 + 3\ H_2O + CaCO_3 \longrightarrow H_2CO_3 + CaSO_4{\bullet}2\ H_2O$$

The pH of the slurry is maintained at 5.5 to 6.0. The optimum pH for the chemical reactions is 5.6 to 5.8. If the pH exceeds 5.8, the amount of limestone required increases, and soft scale begins to accumulate rapidly. If the pH drops below 5.6, a hard gypsum ($CaSO_4{\bullet}2H_2O$) scale builds up on scrubber surfaces. Adjusting the rate at which limestone is added to the slurry controls the pH of the slurry.

In a lime-based FGD process, lime and sulfur dioxide react to form calcium sulfite or sulfate as follows:

$$CaO_{(s)} + SO_2 + \tfrac{1}{2}H_2O \longrightarrow CaSO_3{\bullet}\tfrac{1}{2}\ H_2O$$

$$CaO_{(s)} + SO_2 + 2\ H_2O + \tfrac{1}{2}\ O_2 \longrightarrow CaSO_4{\bullet}2\ H_2O$$

In this process, first lime hydrates and dissolves in the aqueous phase. Next, sulfur dioxide dissolves in the aqueous phase where it participates in numerous ionic reactions. Finally, calcium sulfite and sulfate form precipitates that are removed from the aqueous phase.

Oxidation is important in this system. If the system oxidation is very low, all of the calcium sulfate produced will leave in the liquid phase. If the oxidation is below a critical value, generally considered to be approximately 20%, the sulfur will leave as a solid solution of calcium sulfite/sulfate. If the oxidation rate is high, the sulfate builds to supersaturation, and gypsum precipitates. If the extent of supersaturation is above 1.3 to 1.4, scaling can occur (4–6).

Under ideal conditions, a forced-oxidized limestone-based FGD system will produce wallboard grade gypsum. An important factor in the quality of the gypsum produced by such a FGD system is the amount of chlorine in the coal. The chlorine content of the fuel and the chloride concentration in the makeup water control the volume of liquid bleed needed to manage the chloride level in the FGD system. Because few chloride ions can leave the process in the washed gypsum product, almost all of the chloride that enters with the coal must be purged from the system as a blowdown. This blowdown is the FGD scrubber purge water, which will require treatment before it can be discharged to most receiving waters.

The solid wastes generated by combustion are removed from the system as slurries. The electric power industry generally uses ponds and landfills to dispose of sludge. When an oxidation step is included, the limestone-based FGD system yields a sludge that is mostly calcium sulfate (gypsum). Inhibited or nonforced oxidation systems produce sludges that are mostly calcium sulfite. The solids that settle to the bottom of the settling ponds are periodically dredged. The majority of the solid waste is disposed of in a company-owned landfill.

FGD SCRUBBER PURGE WATER CHARACTERISTICS

Table 1 presents the chemical characteristics of purge waters from forced and inhibited oxidation limestone-based FGD systems (7). These data indicate that purge water from FGD systems that are designed for forced oxidation are lower in reduced sulfur species than nonforced oxidation processes but may still contain high concentrations of thionates and sulfur–nitrogen compounds. Trithionate is the dominant thionate in nonforced oxidation systems, whereas dithionate is most prevalent in forced oxidation purge water. Thiosulfate is the dominant reduced sulfur species in nonforced oxidation FGD system purge water.

These wastewaters contain numerous alkali metals and alkaline earth metals, as indicated in Table 2; calcium,

Table 1. FGD Scrubber Purge Water Characterization[a]

Parameter	Nonforced Oxidation, mg/L	Forced Oxidation, mg/L
pH units	6.3	6.9
Total alkalinity as $CaCO_3$	130	450
Total dissolved solids	58,200	11,700
Chemical oxygen demand	1,340	1,260
Chloride	28,900	3,460
Fluoride	30	23
Ammonia as nitrogen	<5	<5
Total Kjeldahl nitrogen	9.5	7.8
Nitrate as nitrogen	<2	<2
Nitrite as nitrogen	<0.1	6.4
Sulfide	<5	<5
Sulfite	41	<6
Sulfate	815	4,730
Thiosulfate	1,010	<12
Total hydrolyzable sulfur (as sulfate)	2,570	3,600
Dithionate	54	162
Trithionate	427	<6
Tetrathionate	102	<6
Hydroxylamine disulfonate (HADS)	386	127
Selenium	0.536	0.204
Nickel	0.929	0.972
Iron	8.98	0.019
Calcium	6,770	1,190
Magnesium	4,420	1,450

[a] Reference 7.

magnesium, sodium, potassium, and barium are present at the highest concentrations (8). The high chloride content of the combusted coal is the principal reason for the elevated concentrations of calcium and magnesium in the inhibited (nonforced) oxidation limestone-based FGD system purge water. Aluminum, cadmium, chromium, cobalt, copper, lead, manganese, molybdenum, silver, nickel, and iron also are present at detectable concentrations, but only nickel exists at levels requiring treatment.

Table 2. FGD Scrubber Purge Water Metals Analysis[a]

Parameter	Nonforced Oxidation, mg/L	Forced Oxidation, mg/L
Aluminum	0.707	0.0733
Antimony	<0.00532	0.0022
Arsenic	<0.00345	<0.00345
Barium	1.02	0.245
Beryllium	<0.00198	<0.00198
Cadmium	0.00562	0.00954
Calcium	6,770	1190
Chromium	<0.00122	0.0061
Cobalt	0.0531	0.0538
Copper	<0.00166	0.00436
Lead	<0.00273	<0.00273
Magnesium	4420	1450
Manganese	1.44	40
Molybdenum	0.1	0.0663
Potassium	174	18.4
Silver	0.00164	0.00268
Sodium	1580	123

[a] Reference 8.

These purge waters also contain selenium at concentrations that require treatment to achieve an acceptable discharge quality. It is known that selenium forms complexes with both thionates and thiosulfate. Typical selenium–thionate complexes include selenopentathionate [$Se(S_2O_3)_2{}^{2-}$] and selenotrithionate; selenothiosulfate ($SeS_2O_3{}^{2-}$) is a typical complex of selenium and thiosulfate (7). These complex ions of selenium can increase the concentration of dissolved selenium in FGD sludge disposal pond waters.

ASH POND AND FGD SLUDGE POND WATER QUALITY

The ash pond and the FGD sludge pond are the major wastewater treatment facilities at electric utilities. The pond waters from FGD sludge disposal sites are usually slightly alkaline and oxidized; pH values are between 7 and 9, and oxidation–reduction potential (ORP) values are from +490 to +290 mV, respectively. Below the oxidized surface layer, the sludge and porewaters are generally highly reducing; ORP values average approximately −300 mV. In this environment, rainfall can impact the dissolution of sulfate, sulfite, and sulfide solids which in turn controls the leachate concentrations of metals (e.g., barium, calcium, cadmium, copper, iron, nickel, lead, strontium, and zinc), metalloids (e.g., antimony, boron, arsenic), and nonmetals (e.g., fluorine and selenium). Field observations have indicated that anaerobic, heterotrophic sulfate-reducing and denitrifying microbes are present in significant populations in FGD sludge and impact the geochemistry of these materials (2,3). Consequently, the low ORP measurements typically observed in FGD sludge porewater are indicative of active dissimilative metabolism that results in microbial reduction of sulfate to hydrogen sulfide.

FGD sludge porewater chemical characteristics for both calcium-based and sodium-based SO_2 scrubber systems are presented in Table 3 (2,3). These data show the range of constituent concentrations that may be found in FGD sludge porewater for a variety of coals and sulfur removal processes. Table 4 presents a comparison of constituent chemistries for FGD sludge liquors and elutriates from a variety of calcium-based and sodium-based SO_2 scrubber systems for power plants that burn either western or eastern US coals (9,10). The large variation in the water chemistry of desulfurization sludges is indicative of the complex nature of FGD sludge–water systems as well as the potential environmental impact that may result from improper disposal of these materials. Elevated concentrations of arsenic, boron, chloride, fluoride, and selenium in FGD sludge liquors and elutriates present unique problems for environmental management and control.

A high potential for H_2S odors from FGD sludges is suggested by the porewater sulfur speciation data. The substantial quantity of reduced sulfur species observed in porewaters results partly from dissimilatory sulfate reduction by *Desulfovibrio sp*. and related microorganisms (2). It is surmised that the variation in the concentrations of reduced sulfur species may result from microbial use of thiosulfate and sulfite as well as sulfate and the degree of oxygen penetration into the underlying sludge porewaters.

Table 3. FGD Sludge Porewater Chemical Characteristics

Constituent	Calcium-Based SO_2 Scrubber Systems (3), mg/L	Sodium-Based SO_2 Scrubber Systems (2), mg/L
Sulfide	<0.4–305	0.7–2,046
Thiosulfate	<0.25–345	BDL–5,200
Sulfite	<0.1–98.7	BDL–66.4
Sulfate	1,480–12,000	15,420–158,910
Aluminum	<0.05–0.28	BDL[a]
Arsenic	0.01–7.9[b]	ND[d]
Arsenic (III)	0.008–6.7[c]	ND
Arsenic (V)	<0.012–1.19[c]	ND
Barium	0.014–0.09	BDL
Boron	15.4–111	713–1,220
Calcium	607–874	7.0–682
Cadmium	<0.002–0.009	BDL
Chloride	466–4,690	13,440–86,900
Chromium	<0.003–0.3	BDL
Copper	<0.015–0.25	BDL
Fluoride	ND	249–811
Iron	<0.02–0.12	BDL
Lead	<0.028–0.23	BDL
Magnesium	<8.4–720	122–4,210
Molybdenum	<0.1–3.47	BDL
Nickel	<0.03–0.35	BDL
Potassium	ND	691–3,200
Selenium	<0.01–1.03	ND
Selenium (IV)	<0.01–0.083[c]	ND
Selenium (VI)	0.01–0.28[c]	ND
Sodium	33.6–6,290	73,200–79,000
Strontium	7.0–16.1	0.52–25
Vanadium	0.024–0.04	BDL
Zinc	<0.02–0.21	BDL

[a] BDL = Below detection limits.
[b] Mean of two samples.
[c] Mean of nine samples.
[d] ND = Not determined.

Table 4. Comparison of Constituents in FGD Sludge Liquors and Elutriates[a]

Constituent	FGD Sludge Liquors Eastern Coals, ppm	FGD Sludge Liquors Western Coals, ppm
Antimony	0.46–1.6	0.09–0.22
Arsenic	<0.004–1.8	<0.004–0.2
Beryllium	<0.0005–0.05	0.0006–0.14
Boron	41	8.0
Cadmium	0.004–0.1	0.011–0.044
Calcium	470–2,600	240–45,000[d]
Chromium	0.001–0.5	0.024–0.4
Cobalt	<0.002–0.1	0.1–0.17
Copper	0.002–0.4	0.002–0.6
Iron	0.02–0.1	0.42–8.1
Lead	0.002–0.55	0.0014–0.37
Manganese	<0.01–9.0	0.007–2.5
Mercury	0.0009–0.07	<0.01–0.07
Molybdenum	5.3	0.91
Nickel	0.03–0.91	0.005–1.5
Selenium	<0.005–2.7	<0.001–2.2
Sodium	36–20,000[b]	1,650–9,000[b]
Zinc	0.01–27	0.028–0.88
Chloride	470–5,000	1,700–43,000[c]
Fluoride	1.4–70	0.7–3.0
Sulfate	720–30,000[c]	2,100–18,500[b]

[a] References 9 and 10.
[b] The high-end concentration of sodium is for dual alkali FGD sludge and reflects a single measurement on an unwashed filter cake.
[c] The high-end concentrations of chloride and sulfate are for limestone scrubbing of FGD sludge and reflect a closed-loop process using cooling tower blowdown for process makeup water.
[d] Chloride forms weak complexes with many metals, including calcium, and these may be scavenged from solution by fly ash.

Ash pond water at a typical bituminous coal-fired power plant may contain a number of pollutants that can impact the local soil and groundwater. These include the metals aluminum, copper, mercury, lead, nickel, and thallium; the metalloids arsenic and antimony; and the nonmetal selenium. Generally, fly ash handling wastes are the largest source of pollutants to an ash pond (11,12). Other smaller sources include bottom ash handling, plant sump wastes, and cooling tower discharges. Coal pile runoff may also be an intermittent source of pollution along with fly ash sluice water, which is generally at a low pH.

Under typical conditions, ash ponds provide sufficient residence time to reduce suspended solids concentrations below discharge limits (13,14). The solid particles that remain are those that are near or below the density of water and those that are so small that they exhibit properties that create a repelling force and prevent agglomeration and settling in the residence time available to them. These particles are generally clays and fine silt soils from naturally occurring portions of the fly ash waste stream or from commingled stormwater runoff. Heavy metal contaminants such as aluminum and iron in the discharge of ash ponds are directly related to the concentration of suspended effluent solids.

Arsenic and selenium are predominately present in their dissolved state in ash pond water and are poorly removed by this treatment. The forms most commonly found in ash pond water are arsenate (Arsenic V) and selenite (Selenium IV). The speciation of these constituents affects wastewater management because each element has a different treatment requirement that results from differences in their aquatic toxicity. For example, arsenate is more toxic to aquatic life than arsenite (Arsenic III). Fortunately, arsenate is effectively removed by iron coprecipitation, which is also an effective treatment for removing selenite from ash pond water (15). However, neither arsenite nor selenate (Selenium VI) are effectively removed by ferric oxyhydroxides. The fact that soluble arsenic and selenium are not effectively removed by ash ponds suggests that coprecipitation does not occur under normal operating conditions. This indicates a lack of either a source of ferric or ferrous ions within these ponds.

The wet sluicing of fly ash to a pond will increase the apparent concentration of copper and selenium in the pond for certain coal types (8). Wet sluicing will also affect the water quality throughout the pond. The elevated selenium concentrations in ash pond water are due primarily to the increased loading of the soluble form of this constituent that results from the physical act of sluicing. Copper concentrations are also high because of the lower pH conditions that are a consequence of the wet sluicing operation. The relationship between the copper concentration and the pH is supported by solubility theory in that the apparent concentration of copper is an order of magnitude higher at pH 6 than at pH 7 in slightly acidic ash ponds. The percent sulfur in the coal and the calcium content in the dry fly ash influence the pH of the ash sluice water. The solubilities of other metals are not as impacted as is copper by pH changes in the range 5 to 7.

Similarly, the character of the ash influences the concentrations of all heavy metals in an ash pond. For example, acidic fly ash contains high levels of carbonates and sulfates that affect the solubility of certain metals (e.g., copper, nickel, zinc) in the pond water. The amount of iron hydroxide present in the water influences the removal of trace metals by precipitation. Although the range of pH changes vary in different ash pond systems, heavy metal concentrations generally increase with decreasing water pH (8).

Both arsenic and selenium have been removed from wastewater using ferric oxyhydroxide treatment at low pH (7,15). Arsenate is removed effectively by iron coprecipitation at pH values up to 9, and selenite can be treated up to pH 7. However, both constituents are removed more effectively at lower pH. Ferric oxyhydroxide treatment is also effective for removing chromium (III), cadmium, lead, nickel, zinc, and vanadium from FGD sludge pond water. However, it is less effective for removing chromate (chromium VI) and completely ineffective for removing selenate (selenium VI) and boron from FGD sludge pond water.

The metals concentrations in ash and FGD sludge ponds are different because the capture mechanisms of the electrostatic precipitator (ESP), which removes fly ash from the flue gas, and the FGD scrubber, which removes SO_2 from the flue gas, are different. The FGD system captures less ash but more of the volatile metals than the ESP system (13). This results in a higher concentration of dissolved metals in FGD sludge pond water. The two treatment systems are also different because of differences in pond pH. The pH in the FGD sludge pond is usually lower than the pH in an ash pond system because of the nature of the limestone absorption system used for acid gas control in FGD scrubbers (16).

Metals that volatilize and stay in the gaseous form are captured preferentially in the FGD scrubber. Mercury in flue gas exists either as oxidized or elemental mercury. Generally, oxidized mercury is easily captured by the FGD system (17). It is also common to detect nickel at higher concentrations in FGD sludge ponds than in ash ponds because of the effect of chloride on nickel solubility. The presence of chloride at concentrations greater than 10 mM/L in FGD process liquor results in higher nickel solubility by forming soluble nickel–chloride complexes (8). Higher nickel concentrations in FGD sludge ponds also result from the fact that this metal is added with the limestone slurry. The loading of nickel from limestone slurry accounts for approximately 75% of all nickel found in the water of FGD sludge ponds.

The water-phase concentrations of most metals (i.e., copper, mercury, nickel, and zinc) are higher in FGD sludge ponds than in ash ponds. Copper and zinc concentrations are highest at a pH below 6 but decrease as pH values increase. The concentrations of aluminum and arsenic may be higher in ash ponds than in FGD sludge ponds because the concentrations of these constituents tend to increase at the higher pH of the ash pond (8).

Selenium is generally found at higher concentrations in water from FGD sludge ponds than in typical ash ponds (7). Selenium lies just below sulfur in the periodic table, and its chemistry is similar in many respects. Selenium exhibits variable valence from −II to +VI; the most common forms are Se^{2-}, Se^{4+}, and Se^{6+}. Several compounds containing selenium in the various oxidation states are shown in Table 5 with their sulfur analogs. Selenite (selenium IV) is the predominate form of selenium found in waters from ash ponds, and selenate (selenium VI) is the predominate form of selenium found in the waters from FGD sludge ponds (8).

Selenide is the dissociated form of hydrogen selenide (H_2Se). Like its sulfur analog (H_2S), hydrogen selenide is volatile in its associated form, although its vapor pressure is slightly lower than that of hydrogen sulfide. However, hydrogen selenide has much less tendency to volatilize than hydrogen sulfide because its pK_a (3.9) is much lower than that of $H_2S (pK_a = 7.0)$. The principal forms of selenium in FGD sludge pond water are selenite and selenate, which are stable in water and act independently of each other. Generally, selenite is present at lower concentrations because it occurs mainly in acidic and moderately oxidizing waters whereas selenate is predominant in waters that are alkaline and oxidizing. The occurrence of selenium as an oxyanion in water gives it acid characteristics. Selenous acid (H_2SeO_3) is a weak acid ($pK_a = 2.8$) that dissociates in water to $HSeO_3^-$ and SeO_3^{2-}. The concentrations of SeO_3^{2-} and $HSeO_3^-$ are pH dependent; SeO_3^{2-} occurs predominantly at a neutral pH. The hydroxides of iron readily adsorb selenite, thus fixing it in soils as a ferric–selenite complex. However, selenic acid (H_2SeO_4) is a strong acid ($pK_a = -3.0$) that dissociates to give SeO_4^{2-}. Selenate salts are more soluble than the corresponding selenite salts and are not affected by iron hydroxides (7). Selenate behaves like the sulfate ion in natural waters. Consequently, selenium can readily enter the food chain as selenate through soils, into plants, and finally to animals. Elemental selenium exists in several crystalline and amorphous forms; the crystalline metallic form is the most stable and relatively insoluble in water.

The remaining selenium in FGD sludge pond water is probably present as selenotrithionate, a complex ion (7). It is surmised that the addition of the thiosulfate ion displaces sulfite from selenotrithionate to give

Table 5. Forms of Selenium and Sulfur in FGD Sludge Pond Water (7)

Valence	Selenium	Sulfur
−II	Selenide (Se^{2-})	Sulfide (S^{2-})
0	Elemental (Se^0)	Elemental (S^0)
+IV	Selenite (SeO_3^{2-})	Sulfite (SO_3^{2-})
+VI	Selenate (SeO_4^{2-})	Sulfate (SO_4^{2-})

Table 6. Results of TCLP Analysis of Sludges from Treatment of Oxidized FGD Scrubber Water[a]

Metal	Sludge Concentration,[b] mg/kg	TCLP Leachate Concentration,[c] mg/L	Allowable Leachate Regulatory Limit,[d] mg/L
Arsenic	0.0418	<0.0233	5.0
Barium	1.65	0.414	100
Cadmium	0.308	0.0146	1.0
Chromium	12.0	0.0047	5.0
Lead	13.8	<0.0107	5.0
Mercury	0.206	<0.0003	0.2
Nickel[e]	12.4	0.116	NA[f]
Selenium	17.3	0.341	1.0
Silver	0.0179	0.0049	5.0

[a] Reference 8.

[b] The solids resulted from selenium treatment applied to oxidized FGD scrubber water.

[c] After the leaching period, the solids were recombined with the original filtrate before the TCLP was performed. The total solution was then filtered and collected for analysis.

[d] Allowable TCLP regulatory limits in 40CFR261.24.

[e] Nickel is not a TCLP metal.

[f] NA = Not applicable.

selenopentathionate, as follows:

$$Se(SO_3)_2{}^{2-} + 2S_2O_3{}^{2-} \longrightarrow Se(S_2O_3)_2{}^{2-} + 2SO_3{}^{2-}$$

Solutions of selenopentathionate undergo slow decomposition in a FGD sludge pond to produce elemental selenium and the tetrathionate ion. This process probably accounts for the loss of selenium from nonforced-oxidized FGD sludge pond waters when these waters are retained in surface impoundments at ambient temperature for several weeks or more. Usually, the presence of selenopentathionate and selenotrithionate in FGD sludge pond water increases the concentration of dissolved selenium in pond waters. Therefore, it is surmised that inhibition of the removal of selenium from FGD sludge pond water is due to the formation of selenium–thionate complexes that are formed by the presence of thiosulfate in the pond (7).

The results of the Toxicity Characteristic Leaching Procedure (TCLP) performed on sludges from the treatment of oxidized FGD scrubber water are presented in Table 6 (8). The treatment involves contacting the FGD scrubber water with the ferrous ion and hydrated lime to remove the oxyanions of selenium and several divalent and trivalent metal cations such as chromium, lead, and nickel. These constituents appear at elevated concentrations in the resultant wastewater treatment sludge. Nevertheless, the results of the TCLP analysis on this sludge indicate that all constituents are within the allowable leachate regulatory limit.

BIBLIOGRAPHY

1. Battelle, C. (1994). *A Study of Toxic Emissions from a Coal-Fired Power Plant Utilizing an ESP/WET FGD System*. US Department of Energy, Pittsburgh Energy Technology Center, DOE/PC/93251-T2 (Volume 1).
2. Battelle, C. (1994). Pacific Northwest Laboratories. *Sodium-Based Flue Gas Desulfurization Sludge Disposal Ponds*. Electric Power Research Institute, Report TR-103915.
3. Battelle, C. (1994). Pacific Northwest Laboratories. *Calcium-Based Flue Gas Desulfurization Sludge Disposal Ponds*. Electric Power Research Institute, Report TR-103914.
4. Black & Veatch Consulting Engineers. (1982). *Full-Scale Scrubber Characterization of Conesville Unit 5*. Electric Power Research Institute, Report CS-2525.
5. Radian Corporation. (1980). *Evaluation of Chiyoda Thoroughbred 121 FGD Process and Gypsum Stacking*. Vol. I, Electric Power Research Institute, Report CS-1579.
6. Radian Corporation. (1984). *Field Investigation of FGD System Chemistry*. Electric Power Research Institute, Report CS-3796.
7. Castaldi, F.J., Behrens, G.P., and Hargrove, O.W. (2001). Removal of Selenium from FGD Scrubber Purge Water. *Proc. Water Environ. Fed. 74th Ann. Conf.* Atlanta, GA.
8. Castaldi, F.J. (2002). *Personal Data Files on Forced-Oxidation Process Limestone-Based FGD Scrubber Sludges*. Austin, TX.
9. Radian Corporation. (1979). *Review and Assessment of the Existing Data Base Regarding Flue Gas Cleaning Wastes*. Electric Power Research Institute, Report FP-671.
10. Radian Corporation. (1982). *Chemical and Physical Characterization of Western Low-Rank Coal Waste Materials*. US Department of Energy, Grand Forks Energy Technology Center, Report DOE/FC/10200-T4.
11. Radian Corporation. (1975). *Coal Fired Power Plant Trace Element Study*. For US Environmental Protection Agency, Region VIII, EPA Contract 68-01-2663.
12. Radian Corporation. (1975). *Environmental Effects of Trace Elements from Ponded Ash and Scrubber Sludge*. Electric Power Research Institute, Report 202.
13. Radian Corporation. (1994). *A Study of Toxic Emissions from a Coal-Fired Power Plant Utilizing an ESP While Demonstrating the ICCT CT-121 FGD Project*. US Department of Energy, Pittsburgh Energy Technology Center, DCN 93-643-004-03.
14. Southern Research Institute. (1994). *Characterizing Toxic Emissions from a Coal-Fired Power Plant Demonstrating the AFGD ICCT Project and a Plant Utilizing a Dry Scrubber/Baghouse System*. US Department of Energy, Pittsburgh Energy Technology Center, DOE/PC/93254-T1.
15. Benjamin, M.M., Hayes, K.F., and Leckie, J.O. (1982). Removal of toxic metals from power generation waste streams by adsorption and coprecipitation. *J. Water Pollut. Control Fed.* **54**(11): 11.
16. US EPA. (1977). *Disposal of By-Products from Nonregenerable Flue Gas Desulfurization Systems*. EPA-600/7-77-052.
17. US EPA. (1980). *Environmental Assessment of a Coal-Fired Controlled Utility Boiler*. EPA-600/7-80-086.

SLUDGE TREATMENT AND DISPOSAL

Ludovico Spinosa
National Research Council
BARI, Italy

The management of sewage sludge is a major problem in wastewater treatment. Sludge amounts to about 2% by volume of processed sewage, but handling it accounts for up to 50% of total operating costs. Furthermore, during the last few years, there has been a worldwide movement toward a common strategy for any kind of waste; the priorities are reusing waste materials and taking advantage of their energy content.

Treatment of sludge is aimed at rendering it suitable for use or disposal at minimum cost. The low solids concentration (2–3%), and the presence of pathogens, heavy metals, organic pollutants and potentially putrescible organic matter are the main problems to be faced and solved. Other factors influencing sludge processing technology are the need to extract as much energy and material as possible from the sludge itself and the introduction of stricter regulations that require higher levels of treatment before disposal.

PRODUCTION AND CHARACTERIZATION

Three principal categories of sewage sludge can be identified: primary, derived from mechanical and preliminary treatments; secondary or waste activated, from biological treatments; and tertiary, from final physicochemical treatments.

Typical sludge quantities and concentrations are reported in Table 1. They range from 0.2 to 5.0 L/cap/d and 0.7% to 10.0%, respectively; 2 L/cap/d at 4% solids concentration is the typical production of primary plus activated sludge from municipal plants.

The selection of the best processing sequence preliminarily requires knowledge of the sludge characteristics. Total, suspended, and volatile solids content, density of dry solids (1.3–2.1 g/cm^3), pH, sludge volume index (SVI), specific resistance to filtration (values of the order of 10^{12} m/kg at 49 kPa generally indicate good filterability), capillary suction time (CST, optimal values around 10 s using a 10-mm reservoir), compressibility coefficient, floc strength, caloric value (ranging from 21 to 28×10^3 kJ/kg volatile solids), and nutrient content are the most common parameters.

Rheological properties are also of great importance because they influence almost all treatment and utilization/disposal operations. This is particularly true for land application and storage and transportation steps because the selection of the most suitable equipment and procedure is strongly dependent on physical consistency and flow characteristics.

TREATMENT

Treatment options are numerous, but (1) thickening and dewatering to reduce sludge volume and subsequent handling costs and (2) stabilization to reduce putrescibility due to organic matter are of prevailing interest. Disinfection also plays an important role.

Thickening

Thickening, normally carried out before stabilization reduces sludge volume by two to three times. Generally, solids concentration remains below 10%, the sludge behaves as a liquid, and is still pumpable by conventional equipment.

Gravity thickening in tanks is often the most cost-effective method. A thickener may be provided with vertical pickets to facilitate the release of water and gases. Loading rates range from 1.6 to 3.3 kg/m^2/h for mixed sludges, and hydraulic retention times are lower than 24 h to avoid odor problems. As an alternative, sludges can also be thickened by various mechanical systems, such as drainage belts, rotary-drum thickeners, and centrifuges, which sometimes require adding a polyelectrolyte for sludge conditioning. Sludges can be also thickened by dissolved air flotation units.

Dewatering

By dewatering, volume reduction greater than by thickening is obtained, and the sludge becomes paste-like. To improve performance, dewatering is generally preceded by conditioning (chemical or physical).

In chemical conditioning, the particle charge is neutralized, and flocculation is promoted by organic reagents (polymeric macromolecular compounds characterized by monomer type, molecular weight, ionic charge, and degree of hydrolysis) or inorganic reagents (such as iron and aluminum salts, lime, or a combination).

The conditioner type and dosage can be assessed by general laboratory tests, for example, a jar test, specific resistance to filtration and CST, and specific tests for each type of dewatering technology. Physical methods include thermal conditioning, freezing, and the use of inorganic admixtures. Thermal conditioning involves heating sludge at 180–220 °C for 30–90 minutes; sludge sterilization also occurs. Conditioning by freezing seems to be justified only if accomplished by natural means. The use of inorganic substances (e.g., ash, diatomaceous earth) can produce a mixture that has improved filtering characteristics and is usually less compressible than the sludge alone. Dewatering can be accomplished naturally (by drying beds) or mechanically (by filtration or centrifugation).

Drying Beds

When land is available, dewatering by nature can be attractive. On the beds, sludge is placed 15–30 cm deep,

Table 1. Typical Sludge Quantities and Characteristics

Type	Quantity, L/cap/d	Solids Concentration, %	Nitrogen, %-DM	Phosphorus, %-DM	Potassium, %-DM
Raw primary	0.9–2.2	2.0–8-0	1.5–7.0	0.3–2.8	<1.0
Raw activated	1.4–7.3	0.5–1.5	3.0–10.0	1.0–7.0	0.1–0.9
Raw pr. + act.	1.8–2.8	3.0–6.0	4.0–6.0	1.0–1.2	—
Dig. pr. + act.	0.5–1.0	2.0–12.0	1.0–6.8	0.2–5.7	<4.0
Tertiary	0.2–8.0	5.0–10.0	—	—	—

Note: DM = dry matter.

and the time required to reach a liftable state ranges from a few days to several months, depending on climate and sludge characteristics. The area required ranges from 0.05 to 1.40 m^2/cap. Lagoons are similar, but sludge is placed at depths 3 to 4 times greater than in a bed. Lagoons are generally used for long-term storage and, eventually, for cold digestion.

Mechanical Equipment

Mechanical dewatering can take place by filtration, with a filter press or belt press, or by centrifugation. Schemes of these machines are shown in Fig. 1. Filtration under pressure is carried out for 1.5–6.0 h at 0.5–1.4 MPa and is the only operation that allows a cake concentration >30%. Conventional plate filter press operation requires a great deal of labor to open and clean equipment, and yield is low. The automation of plate movement and cloth washing make labor reduction possible, and yield can be increased by using membrane type machines.

In belt pressing, dewatering takes place through an initial drainage step followed by compression under rollers, where sludge also undergoes shearing action due to the relative movement of the two belts. Specific flow rates of 2–3 m^3/h/m belt width are common for municipal sewage sludge, and cake concentrations of 25–30% can be expected. Special care must be taken in belt washing which requires a rinsing water flow rate of 50–200% of that of input sludge.

The type of centrifuge used mostly consists of a cylindrical–conical shell (bowl), that has an internal Archimedean screw (scroll/conveyor) that revolves at a speed slightly lower than that of the bowl. The solid–liquid separation takes place like sedimentation but at g values up to 3000. Pool depth, beach length, and bowl/conveyor differential speed are the most important variables that affect performance. The main advantages consist of indoor solid–liquid separation and limited equipment size, but cake concentrations higher than 25% are normally difficult to obtain. The main general features of filter presses, belt presses, and centrifuges are summarized in Table 2.

Innovations include the development of screw presses, continuous filter presses, high-pressure belt presses, and new generation centrifuges, such as the centripress.

Stabilization

Stabilization is employed mainly for odor control, although pathogen reduction is also achieved. Available options are biological stabilization (anaerobic digestion, aerobic stabilization, and composting), normally permanent in effect, and chemical stabilization (by lime addition), normally temporary. The addition of various oxidizing

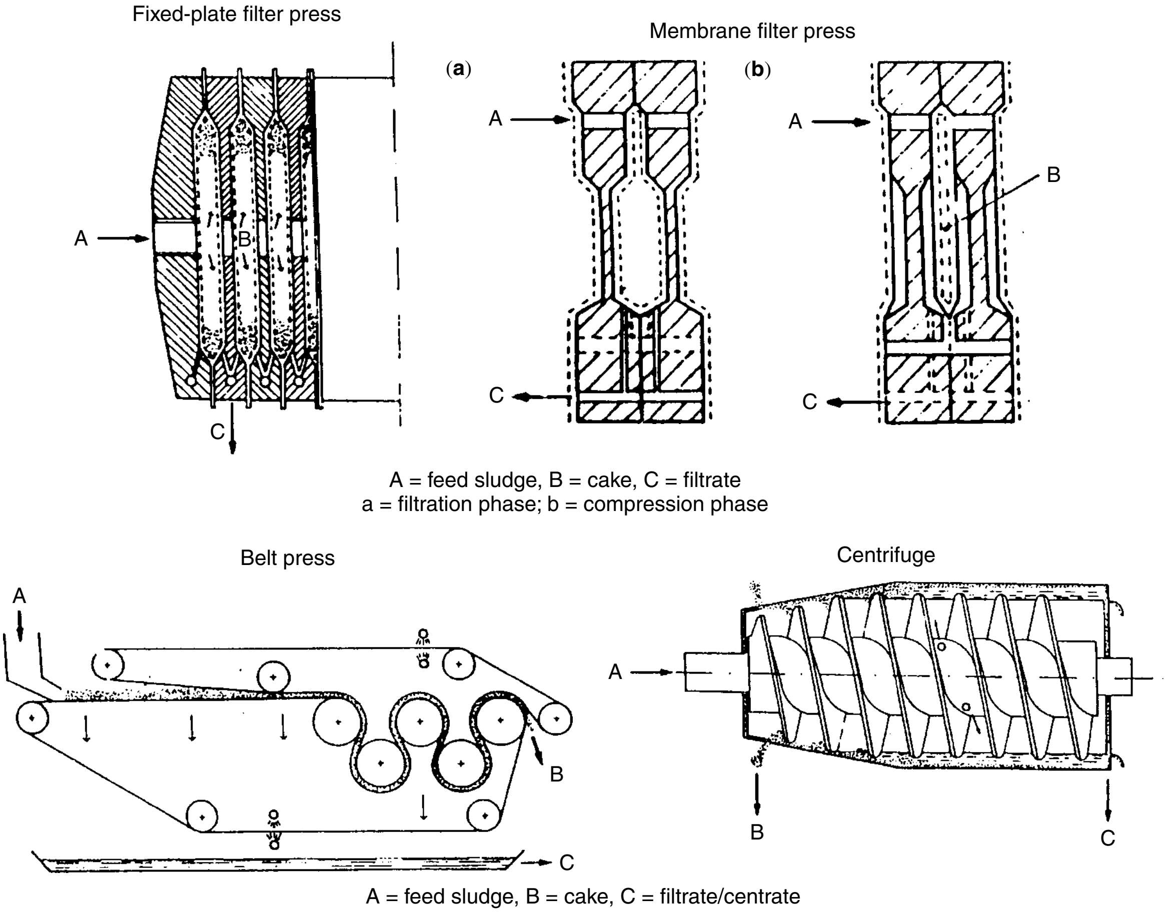

Figure 1. Schemes of dewatering machines.

Table 2. Dewatering Equipment Features

Type	Cake Concentration, %	Separation Efficiency, %	Area Requirement	Energy Demand
Filter press	>30	>95	high	low
Belt press	25–30	around 95	medium–high	medium
Centrifuge	<25	<95	low	medium–high

chemicals (ozone, hydrogen peroxide, etc.) is another possibility, but less used.

Anaerobic Digestion

Anaerobic digestion consists of biological degradation of organic substances in the absence of free oxygen. The organic matter is first liquified and converted by acid-forming bacteria to short-chain acids, then converted by strictly anaerobic methanogenic bacteria to methane, carbon dioxide, and other trace gases.

The simplest process is to store the sludge in open basins for several months, but now this has limited application, especially in cold climate, because of operational and safety problems. A heating process at 35 °C in mesophilic conditions is the norm. The advantages include methane production, no need for oxygen supply, lower net sludge production, and easier solubilization of complex organics. Among the disadvantages, operating control, maintenance problems, heat requirements, supernatant quality, and high capital costs must be mentioned.

Factors that affect the process are feeding modalities, mixing (external pumping, internal mechanical mixing, and internal gas mixing), temperature, pH, bicarbonate alkalinity, volatile acids, and toxics. Typical systems are (1) the low-rate (unheated, unmixed, and intermittently fed); (2) the one-stage high-rate (heated, mixed, and continuously fed); and (3) the two-stage high-rate, a combination of above systems (Fig. 2).

Biogas production of 0.75–1.10 m^3 per kg of volatile solids destroyed, consisting of 65–70% CH_4, 30–35% CO_2, and traces of H_2, H_2S, and volatile solids destruction as high as 60% can be expected.

The most important criterion for digester design is the volumetric loading rate. Reported values range from 0.4 to 8.0 kg volatile solids/m^3/d, depending on digester type. However, as a result of better understanding of digestion kinetics, the solids retention time parameter is now recognized as more important. Hydraulic retention times range from 7 to 30 days.

The development of low-cost spark ignition engines suitable for operation on biogas gives the possibility of

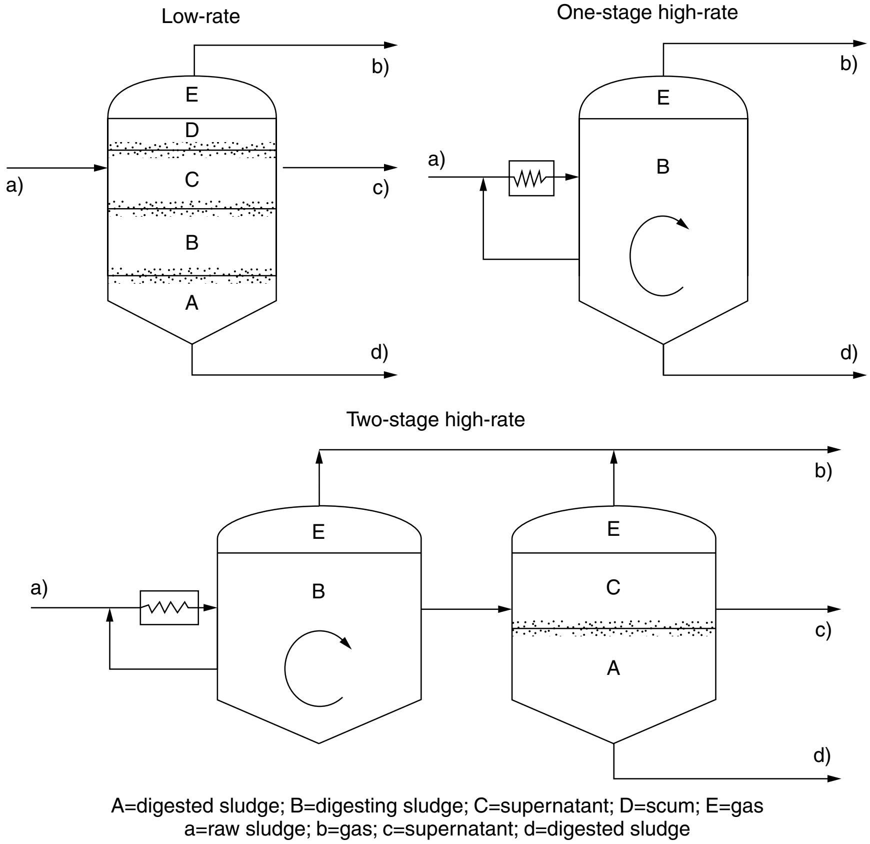

Figure 2. Anaerobic digestion systems.

generating power at small works, even rural, thus making the operation of digesters more attractive economically.

Aerobic Stabilization

Sludge can be stabilized by simply aerating it in open basins. Alternatives are cold stabilization at ambient temperature and autoheated stabilization in thermophilic conditions at high temperature. Cold stabilization involves relatively large energy use for aeration and mixing. Retention times of 10–20 days are normally necessary, but 50 days or more are not uncommon in cold climates. This process is justified at small plants, where the realistic maximum reduction of volatile solids is 40%. In an autoheated process, the heat produced by oxidation of organic matter is conserved to produce a temperature rise to 55 °C or more, thus involving stabilization in only 3–5 days and also sanitizing. The essential prerequisite of this process is effective heat insulation.

As an alternative, oxygenation can be provided by pure oxygen. The gas throughput and consequent heat losses can be minimized and the process operated in more favorable conditions. Another system is the dual-digestion process in which sludge is treated at 50–70 °C by thermophilic aerobic digestion followed by anaerobic digestion without the need for further heating.

Composting

Composting is an aerobic process in the solid phase that converts biodegradable organic substances to a stable, hygienic, and humus-like material under the combined activity of a mixed population of bacteria, fungi, actinomycetes, and protozoa. Composting is considered a treatment option to sludge agricultural refuse.

The main operating variables that affect performance are (1) moisture content (optimal 50–60%) to support microbial activity; (2) carbon to nitrogen ratio (optimal 25–30) to avoid slow processing at high ratios or ammonia volatilization at low ratios; (3) aeration to permit metabolism and respiration of microorganisms and oxidation of organic molecules and avoid development of malodors; (4) temperature (optimal 55–60 °C) to obtain pathogen reduction without inhibiting microorganism growth; and (5) pH, even if quite a wide range is acceptable (5.5–8.0).

Several techniques can be adopted; they differ mainly in aeration and mixing modalities, but all of them consist of basic steps including (1) addition of bulking agents to provide a sufficiently porous and soft mixture suitable for processing; (2) aeration and attainment of a temperature of about 60 °C to destroy pathogens and reduce moisture; (3) further storage (curing) to complete stabilization; and (4) refining treatments, generally including separation and recycling of bulking agents, screening, granulation, and packaging. The general categories of composting plants are open systems, confined ones, and closed or in-vessel systems (Fig. 3).

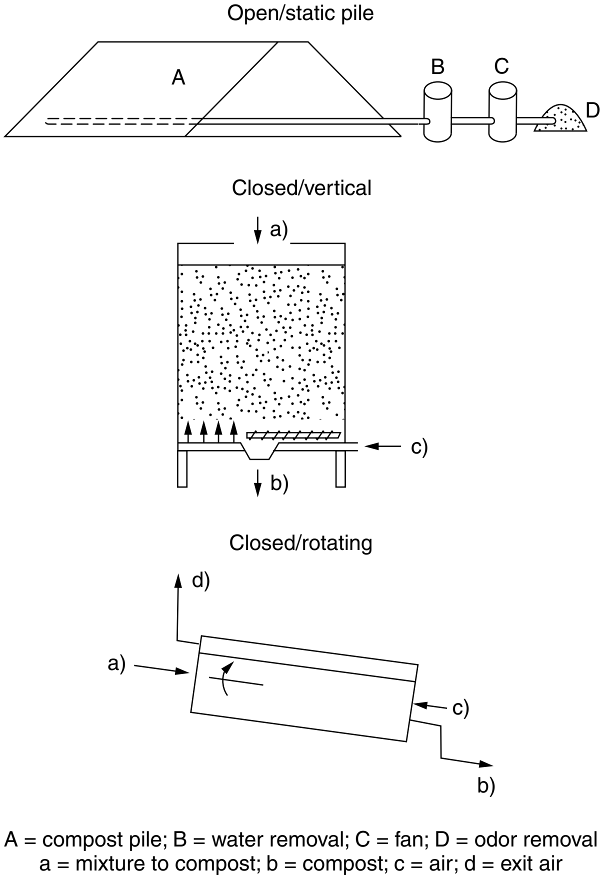

Figure 3. Schemes of composting systems.

In open systems, the material is arranged in windrows and aerated either by simply turning the mass (turned pile) or by forced ventilation (static pile). Stabilization occurs in 15–40 days, followed by a suitable period of maturation. A long, high-temperature period is to be avoided because most microorganisms do not survive temperatures above 60 °C for long time, even if this has a positive effect in reducing pathogens. Therefore, several processes include an initial phase of suction, which permits temperature to rise for a few days, followed by blowing with temperature control. Confined and in-vessel systems, both horizontal and vertical, comprise mechanical equipment that provides better control of turning, ventilation, and other operating variables. The advantages of closed systems over open ones include more effective odor control, indoor operation, insensitivity to climate, and low labor requirements.

Chemical

The process consists of adding lime to sludge to raise the pH above 12 and maintain it for 2 hours, at least. Capital costs are generally low, but the overall economics depend very much on the local price of lime. Calcium hydroxide (slaked lime) or calcium oxide (quick lime) can be used. Quicklime also has the ability to dewater sludge. Lime tends to eliminate odors and reduce pathogenic content, but lime-treated sludges are not chemically stable.

Disinfection

Disinfection reduces the number of pathogenic microorganisms in a raw sludge to a level where they are no longer a risk to human, animal, or plant health. Disinfection can be attained by several methods that act by different mechanisms. Only pasteurization and irradiation have disinfection as the main purpose; other processes disinfect as a secondary effect.

Pasteurization is a heat treatment process carried out at a temperature of 70 °C for 30 minutes. The process is often combined with a stabilization process; it is normally carried out before anaerobic mesophilic digestion.

Ionizing radiation for disinfection uses directly ionizing particles (electrons), produced by accelerators and indirectly ionizing electromagnetic radiation (X), obtained from radionuclides sources, such as ^{60}Co and ^{137}Cs. In general, the doses required for sludge disinfection are 500 krad for liquid sludge and 1000 krad for dewatered sludge.

In composting, the heat generated by biological oxidation is the most important lethal factor, but microbial competition also plays an important röle. A temperature of 55–60 °C for a few days seems to be the lower limit for disinfection.

When lime is added to a liquid sludge, the lethal factor for pathogens is the high pH; when calcium oxide is added to dewatered sludge, an additional factor is the temperature rise to 80 °C. In aerobic thermophilic digestion, biological oxidation results in a temperature rise through which disinfecting action may be obtained; temperatures of 60–70 °C can be reached when aerating with pure oxygen. This process should be carried out preferably as a batch process to prevent microorganisms in raw sludge from reaching the end product before they have been killed. Limited hygyenic effects can also be obtained through anaerobic and aerobic digestion. Sludge processed by thermal conditioning is sterilized and also easily dewaterable; in addition, the sterilizing effect is particularly persistent.

UTILIZATION/DISPOSAL

Land application, thermal processing, and landfilling are available and well-known practices for utilization/disposal of sludge.

Land application

Sludge can be used in agriculture by spreading it on land directly or after composting. Direct use leads to several advantages through recycling of nutrients (i.e., inorganic and organic nitrogen and phosphorus) and organic matter which positively affect soil structure by increasing porosity, stability of aggregates, water retention, and also pH and cation exchange capacity.

However, this practice raises several problems, mainly from the presence of toxic and phytotoxic pollutants, both inorganic and organic, and pathogens. Although numerous processes that can ensure disinfection have been developed, no such reliable methods exist to neutralize or eliminate heavy metals, so regulations set limits for the maximum concentrations of these metals in sludges and the soil, as well as for the maximum disposable quantities for a certain number of years.

Annual nutrient requirements depend on the crop. In most cases, sludge amounts that supply adequate nitrogen also supply phosphorus, but in excess of crop needs, at possible risk of polluting ground- and surface waters. It is also necessary to consider the actual amounts of nutrients available in the first year and those that will become available in subsequent years. Additional sources of potassium are generally required. Annual nitrogen, phosphorus, and potassium use by selected crops are listed in Table 3.

Land application is recommended for sludges from small and medium wastewater works that are close to disposal fields and serve comparatively non-industrialized areas. Direct application to land is, however, subject to great variability over time, depending on crop type and weather conditions, but sludge production is continuous. For these reasons, the agricultural use of compost is often preferred because this material can be more easily stored, transported, and used at times and on sites different from those of production. Composting also involves the production of a safer and more hygienic product.

Thermal processing

Thermal processing includes incineration, pyrolysis, and drying.

Incineration. Incineration involves complete oxidation of volatile matter and the production of an inert residue (ash). If enough water has been removed, the organic material, will sustain combustion; therefore, sludge incineration must be preceded by effective dewatering or drying. The types of sludge incinerators most commonly used are the multiple-hearth furnace (MHF) and the fluidized bed furnace (FBF). The rotary kiln furnace (RKF) is typically used for hazardous waste incineration, but it is also suitable for sewage sludge because of its great adaptability (Fig. 4).

An MHF consists of a vertical, cylindrical, refractory-lined reactor containing a number of horizontal hearths. Rabble arms, supported by a single central shaft, rake

Table 3. Annual N, P, and K Use by Selected Crops

		Nutrient Requirements, kg/ha		
Crop	Yield, t/ha	N	P_2O_5	K_2O
Field				
Corn (grain)	12.3	267	111	256
Rice	7.8	122	67	189
Sugar beets	73.6	306	94	611
Wheat	7.3	194	89	157
Vegetable				
Asparagus	3.3	106	55	133
Lettuce	49.1	106	33	222
Potatoes	25.4	278	128	395
Tomatoes	73.6	278	89	534
Fruit				
Apples	36.8	111	50	200
Oranges	73.6	133	44	194
Forage				
Alfalfa	19.6	500	89	534
Timothy	9.8	167	28	278
Turf				
Bentgrass	6.1	250	89	178

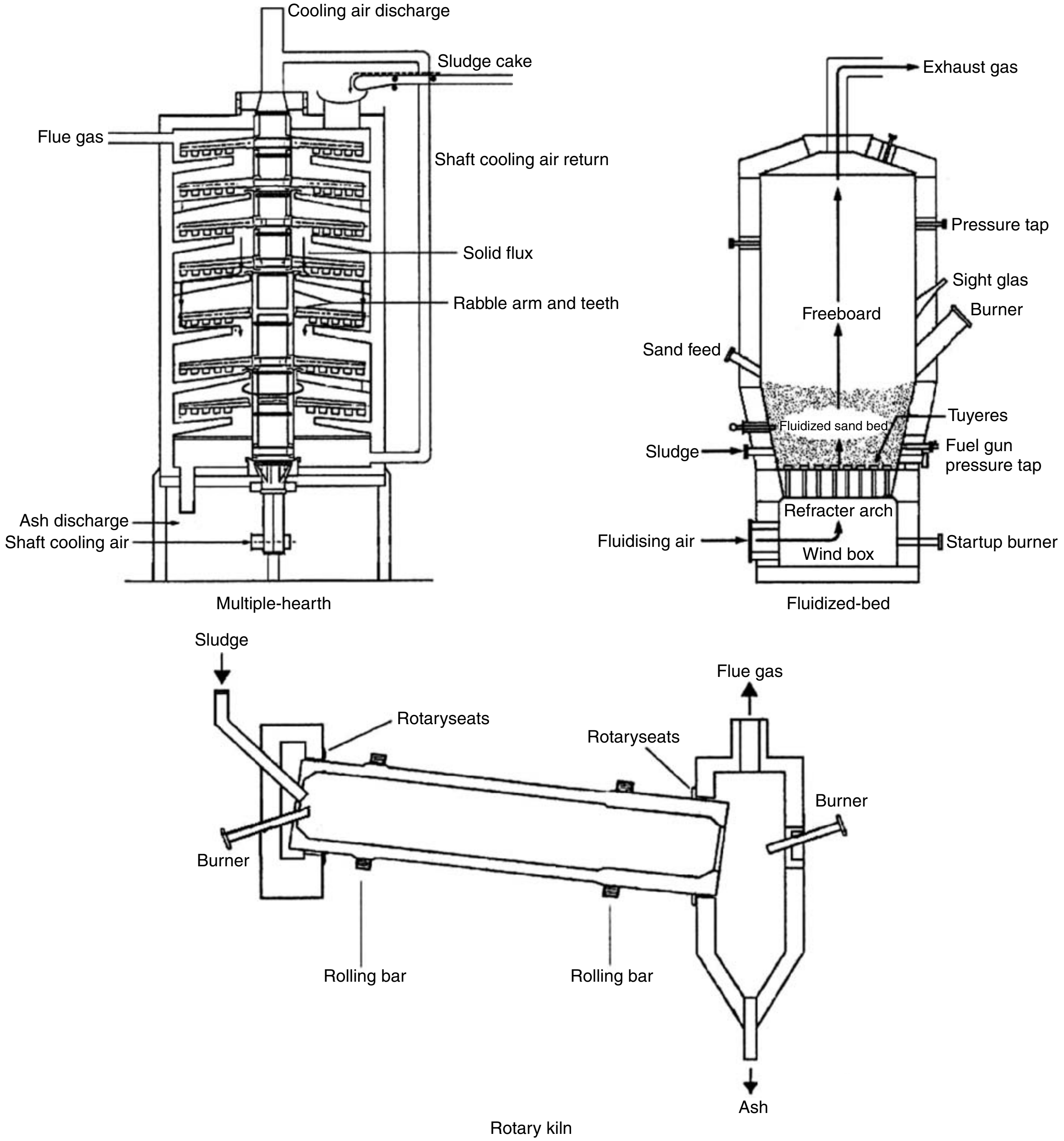

Figure 4. Typical cross sections of furnaces.

the sludge radially across the hearths from the top to the bottom, countercurrently with air and hot gases. Three zones can be distinguished in the furnace: the drying zone in the upper part (gas temperature up to 400 °C), the burning zone in the central–low part (gas and solid phase temperatures of 850–900 °C), and the ash cooling zone in the lowest part (temperatures of ash and air generally <200 °C). The typical design values of a MHF are 2–8 m diameter, 4–14 hearths, a hearth loading rate of 30–60 kg wet sludge/m^2/h, and 100–125% excess air.

The advantages are flexibility with respect to feed quality and loading rates and low fuel consumption due to effective heat recovery inside the equipment. The disadvantages include possible odor problems and emissions of volatile substances (due to the low temperature of the exhaust gas), high need for excess air (due to low turbulence), and high maintenance costs because of many moving parts. Moreover, high fuel consumption is needed, if afterburning of exhaust gases is required to destroy volatile compounds and products of incomplete combustion.

An FBF consists of a cylindrical, refractory-lined shell containing a sand bed that is fluidized, during operation, by air through a distributor plate below the bed. The temperature of the bed is controlled at about 750 °C. FBFs fall into two categories: bubbling and circulating; the latter allows a higher fluidization velocity and very intensive mixing. Particles are carried out of the vertical combustion chamber by the flue gas and are removed in a cyclone to be returned to the FBF through a loop seal. Typical design parameters of a bubbling FBF, which is much more common than the circulating type, are 3–8 m bed diameter, a solid loading rate of 150–300 $kg/m^2/h$, 40–60% excess air, two to three times of fixed bed height (bed expansion), and 1.5–10% of sand losses in 100 h.

The advantages of FBF are low excess air (due to the high turbulence), low NO_x production (due to effective control of combustion temperature), reliability (no moving parts), flexibility for shock load, adaptability to sludges of different moisture content (dewatered, partially dried, fully dried), heat storage capacity by sand bed, and possible abatement of acidic compounds within the bed using additives such as limestone and dolomite. The disadvantages include ash and sand carryover and possible formation of a block of vitrified sand when salts of low melting points are present. This problem can be attenuated by an adding chemicals to bind the alkaline salts.

An RKF consists of a refractory-lined cylindrical shell mounted at a slight incline from the horizontal plane (2–3%) that slowly rotates (0.25–1.50 rpm). Varying the rotational speed allows control of the solid residence time and ensures adequate mixing. The excess air requirement ranges from 100–200%. The advantages include possible melting of ash (but blockage must be avoided), no need for pre-treatments, and adaptation to many feed mechanism designs.

Good performance of thermal process plants also depends on providing proper auxiliary equipment and devices, which include receiving and storage systems, pre-treatment equipment, a feeding system, flue gas cleaning, heat recovery, ash handling, wastewater disposal, and process monitoring.

The main problem in incineration is the potential toxicity of gaseous emissions, but several devices are available for emission abatement at high efficiency. The particulate is generally small, and the legal standards can be easily met; a few metals, particularly Pb and Cd, are poorly retained, but the environmental hazard is slight primarily because the amounts in sludge are small. The problem of toxic substances, such as pesticides and PCBs, can be overcome by afterburning at high temperature (>1000 °C). On the other hand, flue gases represent a potential source of energy, usable to preheat the incoming furnace air, in sludge conditioning, or for external uses.

The volume reduction by incineration is more than 90% compared to the volume of dewatered sludge. The ash is free of pesticides, viruses, and pathogens, and the metals are in the less soluble oxide form.

Pyrolysis

Pyrolysis is a process in which organic material is decomposed at high temperature in an oxygen-deficient environment. This action causes irreversible chemical changes and produces gas, oil, and char (solid residue). The residence time, temperature, and pressure in the reactor are controlled to produce various combinations and compositions of these products.

True pyrolysis involves a total absence of air and applying all required heat externally to the reactor. Partial combustion, also known as starved-air combustion, involves adding a small amount of oxygen to the reactor: the oxygen sustains combustion of a portion of the reactor contents which, in turn, produces the heat required to pyrolyze the remainder of the contents. Pyrolysis has the potential advantage of reducing air pollution and producing useful by-products.

Drying

Drying is the simplest thermal process for producing a solid product. If metals and organic contaminants are low, dried sludge is an acceptable fertilizer. The main drawback is the cost to evaporate the water from the sludge cake; moreover, the end product must be sold to make the operation economical.

Landfilling

Landfilling is a convenient solution where enough space is locally available at reasonable fees. In any case, it is a necessary support to all other systems to dispose of materials that cannot be reused and for maintenance and/or emergency during shutdown periods.

Only well-dewatered sludges are suitable for landfilling. Solids concentrations of at least 20–25% are generally required, but values up to 30–35% are often necessary because the corresponding physical consistency could be too low to support the cover material. A good level of stability is also necessary to avoid possible emissions of bad odors. Co-disposal, with solid wastes or soil, is often practiced.

The basic landfill types are the (1) trench (narrow and wide), (2) area, (3) in-fill mound, (4) fill layer, and (5) dike containment modes. Sludge is spread in layers within a confined area and then daily covered with a thin, continuous layer of inert material.

The main problems are leachate and biogas control. Leachate may be controlled through natural conditions, imported soils or amendments used as liners and/or cover, membrane liners, and collection and treatment. Soil's natural permeability can be reduced by adding imported clays or synthetic materials such as high-density polyethylene (HDPE).

Biogas production typically starts a few months after deposition, reaches a maximum after 5–7 years, and continues for many years at a reducing rate. Gas control techniques can be classified as permeable and impermeable methods. Permeable methods usually entail installing a gravel-filled trench outside or wells inside the filled area to intercept and vent gas into the atmosphere or to an energy recovery system; a forced vacuum extraction system is often appropriate. Gas migration is also minimized by placing a low permeability barrier around the landfill.

Next Page

PRODUCTION OF USABLE MATERIALS

Many other methods to recover usable materials from sludge are available. The recovery of nitrogen (separated by stripping or as struvite) and phosphorus (generally by chemical/physical processes) is becoming of great interest, together with that of the organic fraction as a raw material for activated carbon. Other significant alternatives include the production of (1) slag, char, and Portland cement from dewatered sludge cake and (2) slag, brick, tile, and artificial lightweight aggregate (ALWA) from incinerated ash. Use as animal feed can also be mentioned.

CODISPOSAL WITH MUNICIPAL SOLID WASTES

The combined management of sewage sludge and municipal solid wastes allows, in most cases, overcoming specific technical problems arising from handling them separately, and obtaining significant economic advantages and environmental benefits.

In composting, the different characteristics of solid wastes and sewage sludge can be integrated usefully to obtain a final product of better quality because the relatively higher solids content and carbon to nitrogen ratio of solid wastes can counterbalance the lower solids concentration and carbon to nitrogen ratio of sludge.

In co-incineration, sewage sludge drying can take place by using the excess heat recovered from solid waste combustion, but greater attention in designing and operating furnaces and exhaust gas abatement systems is required.

Co-landfilling provides faster waste stabilization, better leachate quality, and higher biogas production, but the operating procedures must be carefully planned.

WASTEWATER SLUDGE

Izrail S. Turoviskiy
Jacksonville, Florida

TYPES AND CHARACTERISTICS OF SLUDGE

Types of Sludge

The Environmental Protection Agency (EPA) has defined sewage sludge as any solid, semisolid, or liquid residue removed during the treatment of municipal wastewater or domestic sewage, including solids removed during primary, secondary, or advanced wastewater treatment, scum, septage, and sewage sludge products. Instead of using the term sludge, some specialists have used the terms residual, sediments, solids, slime-solid, or biosolids, depending on its condition or its usefulness and whether it meets the applicable criteria for that particular term.

Sludge is a suspension including solid and semisolid materials separated and generated from the liquid wastewater stream of a treatment plant during purification of municipal wastewater, industrial wastewater, or natural water in mechanical, biological, or physicochemical facilities.

Depending on the type of wastewater plants and facilities, the following sludge can be generated:

- —primary raw sludge: settleable solids separated from the wastewater stream during sedimentation in clarifiers (primary settling tanks);
- —activated sludge: a complex of microorganisms of the colloidal type with adsorbed and partially oxidized admixtures, precipitating in the secondary clarifiers in the biological treatment of wastewater;
- —sludge generated in industrial wastewater purification. This sludge has a different chemical composition, quantity, and moisture content, depending on the type of industry and wastewater treatment processes, some levels of radionuclides may be present;
- —sludge from the treatment of natural water (groundwater, surface water) that is generated during the production of potable water: the composition of this sludge depends on the composition of the natural water and the types of reagents used to purify the water.

The wastewater sludge (biosolids) products can also be classified by the type of treatment process as the following:

- —aerobically digested activated sludge or a mixture of it with primary sludge;
- —anaerobically mesophilic or thermophilic digested primary sludge or its mixture with thickened activated sludge;
- —dewatered sludge from mechanical dewatering devices;
- —dried sludge from sludge beds;
- —thermally treated or dried;
- —biothermally treated (compost).

The main and more useful process of municipal and industrial wastewater purification is the biological activated sludge process. The by-products of this process are two types of sludge, such as primary and waste activated sludge, containing up to 99% of contaminants removed from wastewater.

Primary Raw Sludge

Definition. Primary raw sludge is the sediment from clarifiers whereby wastewater and solids are separated to produce clarified effluent and sludge. The main raw sludge components are proteins, nitrogenous compounds, cellulose, sugar, carbohydrates, grease, fats, macronutrients (nitrogen, phosphorous, potassium), micronutrients, bacteria, and viruses. It has high moisture and poor dewatering ability. Fresh sludge is a gray or light brown colored suspension; its particles are of different sizes and composition, and it has a less intense sour odor than septic sludge. Because of the high content of organic material, it decays rapidly, the sludge condition becomes septic, its color changes to dark gray or black, and it generates an objectionable sour odor. Sludge is a hazardous

Previous Page

PRODUCTION OF USABLE MATERIALS

Many other methods to recover usable materials from sludge are available. The recovery of nitrogen (separated by stripping or as struvite) and phosphorus (generally by chemical/physical processes) is becoming of great interest, together with that of the organic fraction as a raw material for activated carbon. Other significant alternatives include the production of (1) slag, char, and Portland cement from dewatered sludge cake and (2) slag, brick, tile, and artificial lightweight aggregate (ALWA) from incinerated ash. Use as animal feed can also be mentioned.

CODISPOSAL WITH MUNICIPAL SOLID WASTES

The combined management of sewage sludge and municipal solid wastes allows, in most cases, overcoming specific technical problems arising from handling them separately, and obtaining significant economic advantages and environmental benefits.

In composting, the different characteristics of solid wastes and sewage sludge can be integrated usefully to obtain a final product of better quality because the relatively higher solids content and carbon to nitrogen ratio of solid wastes can counterbalance the lower solids concentration and carbon to nitrogen ratio of sludge.

In co-incineration, sewage sludge drying can take place by using the excess heat recovered from solid waste combustion, but greater attention in designing and operating furnaces and exhaust gas abatement systems is required.

Co-landfilling provides faster waste stabilization, better leachate quality, and higher biogas production, but the operating procedures must be carefully planned.

WASTEWATER SLUDGE

Izrail S. Turoviskiy
Jacksonville, Florida

TYPES AND CHARACTERISTICS OF SLUDGE

Types of Sludge

The Environmental Protection Agency (EPA) has defined sewage sludge as any solid, semisolid, or liquid residue removed during the treatment of municipal wastewater or domestic sewage, including solids removed during primary, secondary, or advanced wastewater treatment, scum, septage, and sewage sludge products. Instead of using the term sludge, some specialists have used the terms residual, sediments, solids, slime-solid, or biosolids, depending on its condition or its usefulness and whether it meets the applicable criteria for that particular term.

Sludge is a suspension including solid and semisolid materials separated and generated from the liquid wastewater stream of a treatment plant during purification of municipal wastewater, industrial wastewater, or natural water in mechanical, biological, or physicochemical facilities.

Depending on the type of wastewater plants and facilities, the following sludge can be generated:

- —primary raw sludge: settleable solids separated from the wastewater stream during sedimentation in clarifiers (primary settling tanks);
- —activated sludge: a complex of microorganisms of the colloidal type with adsorbed and partially oxidized admixtures, precipitating in the secondary clarifiers in the biological treatment of wastewater;
- —sludge generated in industrial wastewater purification. This sludge has a different chemical composition, quantity, and moisture content, depending on the type of industry and wastewater treatment processes, some levels of radionuclides may be present;
- —sludge from the treatment of natural water (groundwater, surface water) that is generated during the production of potable water: the composition of this sludge depends on the composition of the natural water and the types of reagents used to purify the water.

The wastewater sludge (biosolids) products can also be classified by the type of treatment process as the following:

- —aerobically digested activated sludge or a mixture of it with primary sludge;
- —anaerobically mesophilic or thermophilic digested primary sludge or its mixture with thickened activated sludge;
- —dewatered sludge from mechanical dewatering devices;
- —dried sludge from sludge beds;
- —thermally treated or dried;
- —biothermally treated (compost).

The main and more useful process of municipal and industrial wastewater purification is the biological activated sludge process. The by-products of this process are two types of sludge, such as primary and waste activated sludge, containing up to 99% of contaminants removed from wastewater.

Primary Raw Sludge

Definition. Primary raw sludge is the sediment from clarifiers whereby wastewater and solids are separated to produce clarified effluent and sludge. The main raw sludge components are proteins, nitrogenous compounds, cellulose, sugar, carbohydrates, grease, fats, macronutrients (nitrogen, phosphorous, potassium), micronutrients, bacteria, and viruses. It has high moisture and poor dewatering ability. Fresh sludge is a gray or light brown colored suspension; its particles are of different sizes and composition, and it has a less intense sour odor than septic sludge. Because of the high content of organic material, it decays rapidly, the sludge condition becomes septic, its color changes to dark gray or black, and it generates an objectionable sour odor. Sludge is a hazardous waste which must be stored, disposed of, or reused, in accordance with hazardous waste regulations. Sludge regulations limit sludge disposal alternatives based on the treatment level provided, pathogen removal, and the heavy metals content. Treatment processes include conditioning, dewatering, stabilization (control of odor, pathogens, biodegradable toxins, and vectors), and disinfecting resulting in a valuable land application and soil conditioning product that has many useful properties.

Quantity and Moisture. The mean moisture content of the sludge discharged from primary clarifiers is 95% for gravity-flow removal and 93.8% for removal by plunger pumps. The quantity of raw primary sludge is approximately 0.4–0.5% of the volume of wastewater treated. The quantity of municipal wastewater primary sludge with 95% moisture can be estimated by assuming that 39.0 cu ft/1000 persons is produced daily.

Chemical Composition. The composition of sludge from primary settling tanks (clarifiers) depends on the nature of the wastewater being treated and particularly the type and quantity of industrial wastewater treated together with municipal wastewater. The dry solids of sludge consist of organic matter (60–75%) constituted of protein, fats, and carbohydrates. The ultimate composition of the dry solids of sludge includes carbon, hydrogen, sulfur, nitrogen, and oxygen. The typical composition of raw primary sludge is as follows (% of dry solids): grease and fats 6–30, protein 20–30, nitrogen 1.5–8.0, phosphorus 0.8–2.8, potassium 0.1–1.0, and cellulose 8–15. The mineral constituents include SiO_2, Al_2O_3, Fe_2O_3, CaO, MgO, K_2O, Na_2O, and other minerals. Wastewater sludge may also contain heavy metals such as cadmium, chromium, cobalt, copper, lead, mercury, nickel, and zinc. Alkalinity and pH are the most important of the easily measured chemical parameters affecting sludge conditioning. Raw primary sludge has a pH range of 5.0–8.0, alkalinity (mg/L as $CaCO_3$) of 500–1500, and organic acid (mg/L as HAc) 200–2000.

Waste Activated Sludge

Definition. Activated sludge gets its name from the interaction between wastewater and microorganisms (bacteria, protozoa, and rotifers) in the presence of dissolved oxygen in a biological process of wastewater purification. Association of microorganisms that use organic material as food, remove suspended solids, and adsorb and mineralize wastewater contaminants are given the name activated sludge. The average sizes of activated sludge flocs are 1–4 mm; they have a light gray, yellow gray, or dark brown color. As the biological process of wastewater purification takes place, activated sludge increases. After the biological process of wastewater purification, activated sludge settles in secondary clarifiers. From the secondary clarifiers, the main part of the activated sludge is moved into aeration tanks to participate in the biological process; the other part receives treatment as waste activated sludge.

Activated sludge contains mostly bacterial cells that are viscous and difficult to dewater. The parameters of activated sludge sedimentation are sludge volume index (SVI), sludge density index (SDI = SVI/100), and sludge age determined by a settleometer test. An SVI less than 100 mL/g indicate an older, denser, fast settling sludge that has a thick, scummy, dark tan foam in aeration tanks. An SVI of more than 100 mL/g indicates a young, slow settling, light density sludge. An optimal SVI is 70–130 mL/g. An SDI of more than 1 g/cm^3 indicates old sludge; an SDI of less than 1 g/cm^3 indicates a young sludge. Sludge age refers to the number of days for which the suspended solid particles remain under aeration (5–15 days). The following microorganisms are Indicators of activated sludge age in order from young to old: amebas, flagellates, ciliates, and rotifers.

Quantity and Moisture. The moisture content of the activated sludge discharged from secondary clarifiers after aeration tanks is 99.2–99.7%. Gravity or belt gravity thickening of waste activated sludge is a very important process because of the high volume and moisture content in waste activated sludge. After gravity thickening, waste activated sludge has a moisture content of 97–98%, and volume is decreased 5–10 times. After belt gravity thickening, waste activated sludge has a moisture content of about 95%. In the process of belt gravity thickening, added polymers can reduce the moisture content. Thickened activated sludge is often treated together with primary sludge. For approximate computations, the quantity of the mixture of primary sludge and gravity thickened activated sludge at an average moisture content of 96.2% can be assumed as 0.6–1% of the volume of wastewater treated.

Chemical Composition. Activated sludge consists of microorganisms and adsorbed particles. The activated sludge dry solids consist of 70–75% organic matter. The dry solids contain carbon, hydrogen, nitrogen, sulfur, and oxygen. Activated sludge contains (% of dry solids) 6–7.5 fats, 2.4–7.5 nitrogen, 2.8–11.0 phosphorous, potassium up to 0.4%, and approximately 2.5–3 times less carbohydrates and two times more protein than primary raw sludge, the major minerals present in dry solids. Activated sludge can contain Cr, Cu, Ni, Pb, Cd, and other elements. The alkalinity of activated sludge is 580–1100 (mg/L as $CaCO_3$), organic acids 1100–1700 (mg/L as HAc), and pH 6.5–7.5.

Physical Properties

Granulometric Composition. Primary raw sludge contains 5–20% of particles larger than 7–10 mm, 9–33% 1–7 mm, and 50–88% below 1 mm in which about 45% have sizes less than 0.2 mm of the total weight of dry solids. In activated sludge, the quantity of particles less than 0.2 mm is 90%, below one mm 98%, particles of 1–3 mm 1.6%, and over 3 mm 0.4% of the weight of dry solids. The organic part of sludge decays rapidly, and an increase in the quantity of finely dispersed and colloidal particles and bound water results in a decrease in the water separation from the sludge and poor dewatering ability.

Density and Fluidity. The average density of activated sludge is 0.7–1.3 g/cm^3. The density of primary sludge is about 1 g/cm^3, and the density of the sludge dry solids is 1.2–1.4 g/cm^3. Primary raw sludge whose moisture is higher than 90% is a fluid; when moisture is 86–90%, it looks like sour cream; when moisture is 82–86%, it looks like a slush; and when moisture is less than 82%, it looks like a light thin soil. Waste activated sludge whose moisture is 88–91% looks like sour cream, and when moisture is 85–87% like thin soil. At a concentration of solids above 5% of primary sludge and above 3% of activated sludge, they are non-Newtonian which means that head losses are not proportional to velocity and viscosity. They are also thixotropic which means that they become less viscous when stirred.

Dewatering Characteristics. Dewatering is a process of natural or mechanical removal of water from sludge. The dewatering characteristics of sludge can be obtained by measuring the volume of filtrate collected from sludge and the time it takes to filter using varying doses of conditioning reagents. The most commonly used tests are the Specific Resistance Test, the Buchner Funnel Filtration Test, and the Capillary Suction Time Test (CST).

The water in sludge may be present as free water and water bound physicomechanically, physicochemically, or chemically. The more bound water present in sludge, the more energy or reagents must be used to condition sludge to remove bound water by dewatering. The separation of water from sludge depends on the size of the solid particles; the smaller the particles, the poorer the water separation from the sludge. Any process that reduces the size of the suspended solids particles has a negative effect on conditioning and dewatering. The chemical composition of sludge exerts a significant influence on its treatment. Compounds of iron, aluminum, chromium, and copper, as well as acids and alkalis, improve the processes of precipitation, thickening, and dewatering and reduce the consumption of chemical reagents for conditioning of sludge before dewatering. Oils, fats, and nitrogen compounds intensify anaerobic sludge digestion but interfere with thickening and conditioning processes.

Thermophysical Characteristics. The specific heat of a mixture of primary and thickened activated sludge is (3.5–4.7) × 10,000 Joule/(kg K). The heat of combustion of sludge dry solids equals 16.7–18.4 MJ/kg, the heat for incinerating fuel matter is 23.4–26.9 MJ/kg of sludge organic. The heat value is higher for raw sludge and lower for activated sludge. Sludge burns at a temperature of 430–500°C (800–1000°F); to eliminate odors, the temperature needs to be raised to 800–850°C (1500°F). For an increase in the moisture and ash of the sludge, there is a decrease in the heat of combustion and liberation of volatiles. In the process of thickened activated sludge aerobic digestion, 3.6 kcal are released/g volatile (organic) suspended solids oxidized (15 MJ/kg). The reduction of 1 kg of organic sludge during composting of dewatered sludge creates an average of 21 MJ/kg of heat. Raw primary sludge has a thermal content 6,800–10,000 Btu/lb.

Bacteriological Content

The activated sludge process is the most efficient biological process for removing coliforms, pathogenic bacteria, and virus particles from wastewater; they are transported to primary and activated sludge. The primary sedimentation of sewage allows reducing 30–70% of microorganisms and bacteria. After activated sludge treatment, the reduction of microorganisms and bacteria reaches 90–99%. The average level of indicator bacteria and pathogens, such as coliforms, *Streptococcus, Salmonella*, enteric viruses, and parasite ova/cysts reaches millions n/g dry weight of sludge. The diversity of microbial flora makes it difficult to enumerate the total population.

Biosolids

The Water Environment Federation (WEF) has adopted a policy of encouraging the use of the word biosolids in place of sludge to promote public acceptance of reused water projects. The term biosolids is used to connote the primary organic solid product of treatment that meets US EPA or other applicable criteria for beneficial use. The term biosolids has been used by WEF in the last few years instead of sludge. The type and level of wastewater and sludge treatment has an effect on the type, quantity, and quality of the biosolids generated. Biosolids are the solid organic matter produced as by-products of municipal wastewater treatment processes (also known as sewage sludge) that can be beneficially used, especially as a soil amendment, in accordance with standards and requirements. Sludge regulations limit biosolids disposal alternatives based on the treatment level provided.

CONCLUSION

Primary and waste activated sludge are two main types of sludge (biosolids) formed in wastewater purification. The quantity of a mixture of primary and thickened waste activated sludge whose average moisture is 96.2% reaches 1.0% of the treated wastewater. High moisture with bound water and small solid particles make sludge difficult to dewater. The dry solids of sludge consist of organic matter and minerals. The microbiological population of sludge includes millions of coliforms, pathogenic bacteria, and viruses. Sludge (biosolids) contains macronutrients, such as nitrogen, phosphorus, potassium, and micronutrients. After treatment (thickening, conditioning, stabilization, dewatering, disinfecting), biosolids meet regulatory requirements for pathogens, vector attraction reduction, and heavy metal content and become beneficial valuable products, which can be applied to land for soil conditioning, preventing soil erosion, and as fertilizer.

READING LIST

Bitton, G. (1944). *Wastewater Microbiology*. Wiley-Liss, New York.

Bruce, A.M. and Fisher, W.J. (1994). *Sewage Sludge Stabilization and Disinfecting*. Ellis Hardwood Ltd., Chichester, UK.

WEF. (1992). *Design of Municipal Wastewater Treatment Plants*. MOP 8, Vol. 2. WEF. #76.

Foess, G.M. and Siger, R.B. (1993). Pathogen/vector attraction redaction requirement of the sludge rules. *Water Eng. Manage.* **Jun.**: 25.

US EPA. (1997). *Land Application of Biosolids Process Design Manual*. US EPA.

Lue-Hing, C., Zenk, D.R., and Kuchenither, R. (1992). *Municipal Sewage Sludge Management*. Technomic, Lancaster, PA.

Siger, R.B. (1993). Practical guide to the new sludge standards. *Water Eng. Manage.* **Nov.**: 26.

Spellman, F.R. (1997). *Dewatering Biosolids*. Technomic, Lancaster, PA.

US EPA. (1990). *State Sludge Management Program Guidance Manual*. US EPA.

Tehobanoglous, G.B., Franklink, L., and Metcalf and Eddy, Inc. (1992). *Wastewater Engineering: Treatment, Disposal, and Reuse*. McGraw-Hill, New York.

Turovskiy, I.S. (2000). Biosolids or sludge? *Water Eng. Manage.* **April**: 19–21.

Turovskiy, I.S. (2001). Reduction of energy consumption in wastewater sludge treatment. *Fl. Water Resour. J.* **March**: 34–36.

US EPA. (1995). *A Guide to the Biosolids Risk Assessments for EPA Part 503 Rule*.

US EPA. (1993). *Office of Wastewater Management*. EPA 832-B-93–005.

US EPA. (1999). *Biosolids Generation, Use, and Disposal in the United States*. 530-R-99. September 10.

Vesilind, P.A. (2003). *Wastewater Treatment Plant Design*. WEF and IWA Publishing.

Ward, R.L., McFeters, G.A., and Yeger, J.G. (1984). *Pathogens in Sludge*. Sandia Report. 83–0557. TTC-0428. Sandia Natural Laboratory, Albuquerque, NM.

PROCESSING OF SLUDGE

Izrail S. Turoviskiy
Jacksonville, Florida

Wastewater sludge is a high moisture suspension in which water is bound by small solid particles. The main parts of the sludge solids are organic and contain different types of microorganisms, pathogenic bacteria, viruses, and parasites. On the other hand, sludge contains valuable organic macronutrients, such as nitrogen, phosphorus, potassium, and micronutrients, such as elements and metals. Wastewater sludge is hazardous waste which must be stored, disposed of, or be reused in accordance with standards and regulations based on the treatment level provided.

The main purpose of treatment is to prepare sludge for removal from wastewater treatment plants and to produce biosolids that can be used as a soil amendment and conditioner. To achieve this goal, treatment processes of sludge should include thickening and dewatering to decrease sludge volume and moisture and disinfection and stabilization to control odor, pathogens, and vectors.

SLUDGE TREATMENT

Thickening

Thickening is an economically effective process to increase sludge concentration and decrease sludge volume by removing some free water to the extent that the sludge remains in the fluid state. The objectives of thickening are to produce a relatively solids-free supernatant and to produce a sludge that can be pumped without difficulty. Sludge can be thickened by using gravity thickeners, gravity belt thickeners, rotary drums, separators, centrifuges, and flotators. Rotary drums and gravity belt thickeners are mechanical devices that remove free water from wastewater sludge using gravity. Centrifugation is a process in which centrifugal force (usually about 500 to 3000 times the force of gravity) is applied to a sludge slurry to accelerate the separation of the solid and liquid fractions. Flotation thickening is a solid–liquid separation caused by introducing fine air bubbles into the liquid phase. Adding polymers in waste activated sludge thickening allows removing more water.

Gravity thickening is a common method of solid–liquid separation to reduce the sludge volume handled in the dewatered/sludge disposal of a wastewater treatment facility. During the process of waste activated sludge gravity thickening, the concentration of dry solids increases, on average, from 0.2 to 2.0% and the volume of sludge is reduced 10 times. However, when the thickening process lasts 8–10 hours or more, organic putrefaction occurs, the microorganisms of the activated sludge perish without air, the amount of colloids increases, and part of the free water is transformed into a bound state of water with dry solids. Unthickened activated sludge usually has better dewatering abilities than thickened sludge. On the other hand, dewatering or digesting nonthickened activated sludge does not make sense due to the large volume and low initial concentration of dry solids. The kinetics of the process of activated sludge thickening determines the rational concentration of dry solids.

Conditioning

Conditioning is a chemical or physical process that improves the dewaterability of a sludge during its preparation for dewatering. Sludge conditioning consists of such methods as inorganic chemical conditioning, organic chemical conditioning, thermal conditioning, elutriating, and freeze–thawing. Chemical conditioning is a commonly used method. By this process, chemicals such as ferric chloride, lime, organic polymers, and others are added to sludge to coagulate or flocculate the fine particles and decrease bound water. It is a unique process vital to the successful operation of sludge thickening and dewatering systems. Thermal conditioning uses elevated temperatures and pressures to promote the separation of solids and liquid through the release of cell-bound water. By using thermal conditioning, sludge can often be mechanically dewatered without using chemicals. Freeze–thaw conditioning substantially decreases consumption of reagents required for conditioning sludge that was mechanically dewatered.

Dewatering

Dewatering is the process of natural or mechanical removal of water from sludge during which the sludge loses its fluidity, becomes damp solids, and can be transported in bulk. The dewatering processes currently in use include natural methods such as air drying on drying beds, lagoons, and mechanical methods such as belt press filtration, centrifugation, vacuum filtration, and pressure filtration. Moisture in sludge may be bound physicomechanically, physicochemically or bonded chemically and can also be present as "free water." Conditioning of sludge before dewatering allows increasing part of the free water, and more water can be removed by mechanical dewatering. Primary raw sludge has better dewatering ability than mixtures of primary raw sludge and thickened activated sludge and better than digested sludge. Dewatering often is followed by sludge aerobic or anaerobic digestion. During the dewatering process, the water is not completely removed from the sludge. The moisture remains within 70–80% limits. At that moisture level, sludge loses its fluidity and may be moved by conveyors. Sludge, generally, should be dewatered before it is thermally dried, composted, alkali stabilized, or incinerated.

Stabilization

Stabilization is one of the more frequently used methods of sludge treatment (especially if the sludge will be used as a fertilizer). All sludge requires some form of treatment, whether stabilization, thickening, or dewatering, possibly followed by drying, composting, and incineration, or a combination of one or more of these processes, before being discharged into the natural environment. Stabilization is a combination of processes of sludge treatment to meet U.S. EPA or other applicable criteria for beneficial use. The process of stabilization is necessary to eliminate the potential of putrefaction of sludge's organic part, to prevent offensive odor dissemination, and to reduce volatile and pathogen content. Two criteria typically used to measure biosolids stability include the volatile solids content and pathogen indicator organism reduction. The following methods of sludge stabilization are in use: aerobic digestion, anaerobic digestion, alkaline stabilization, composting, thermal drying, heat/wet air oxidation, acid (oxidation) disinfecting, and some others.

Aerobic Digestion. Aerobic digestion has been widely used in wastewater treatment plants (WWTP) for many years. Aerobic digestion is a process of oxidizing the organic part of the sludge by microorganisms in special tanks in the presence of oxygen (air aeration of sludge). Aerobic sludge digestion stabilizes raw sludge and produces biosolids for further treatment and disposal. The process of aerobic sludge digestion may be conducted using several technological schemas. The duration of the volatile part of solids oxidation depends on the food/microorganism ratio, temperature, intensity and quantity of the air aeration, and also on wastewater composition and technological demand. This process is more useful for digesting and stabilizing thickened activated sludge.

Anaerobic Digestion. Anaerobic digestion is a biological process that reduces volatile solids by using microorganisms in the absence of oxygen and reduces odor and pathogen content. Digesters are cylindrical reservoirs with conical bottoms; the upper section of the reservoir has a sealed cover with a device for collecting the gas. Two types of anaerobic digestion processes are in use: mesophilic and thermophilic. Mesophilic processes occur in the temperature range of 32–35 °C. The thermophilic process operates at higher temperature (50 to 55 °C) to reduce organic solids and pathogen content further. The quantity of gas obtained during digestion is approximately 1.0 m^3 for every 1.0 kg of disintegrated organic sludge. The following composition of the gas can be expected: methane—60 to 70%; carbon dioxide—16 to 34%; nitrogen, hydrogen, and oxygen—0.4 to 6%. The heat of combustion of this gas averages around 21 million Joule/m^3(MJ/m^3). Anaerobic digestion is a widely used stabilization process for primary or a mixture of primary and thickened waste activated sludge.

Alkaline Stabilization. Alkaline stabilization of sludge produces biosolids that are reduced in pathogen and vector attraction, and meet Class "A" requirements. Of the chemicals used for sludge stabilization, the most common is quicklime or hydrated lime, which is added to sludge before or after dewatering. The quantity of lime added is determined so that the pH of the sludge and lime mixture is raised to 12.0 or above for a period of 2 hours.

Composting. Various composting processes are used, such as windrow composting, aerated static piles, and in-vessel composting. This technology includes mixing dewatered sludge with an added bulking agent (sawdust, peat, wood chips, bark, etc.) and aerating the mixture. In sludge composting, a biothermal process takes place in which microorganisms reduce the sludge's organic in the presence of oxygen. This aerobic process is accompanied by a rise in temperature to about 55–65 °C and a decrease in moisture content. The quantity of organic sludge reduced during composting averages 25%. A reduction of 1.0 kg of sludge organic creates an average 21 MJ/kg of heat. Taking in to account heat losses and heating of compost material, it is necessary to spend approximately 4 MJ of heat to evaporate 1.0 kg of water. Thus, the reduction of 1.0 kg of organic sludge allows removing 5.0 kg of water from the sludge. Besides that, part of the moisture is removed from the sludge by natural evaporation. The total quantity of moisture removed from the sludge depends on climate factors, the season of the year, the dimensions of the piles, duration of composting, and intervals between shoveling over. Removal of moisture from the sludge produces compost at a moisture content of approximately 50–55%. During composting, the heat generated by the decomposition of the organic portion of the sludge stabilizes and renders the sludge harmless transforming it into usable soil conditioning biosolids.

Thermal Drying. Thermal drying is used to destroy pathogens, to eliminate odor, and to reduce most of the water content and the volume of sludge and transportation costs of the biosolids. Thermal drying of sludge is conducted in drying systems consisting of a drying device and the auxiliary equipment, which includes furnaces with a fuel supply system, feeders, cyclones, scrubbers, blowing equipment, conveyors and bins, monitoring and measurement instruments, and automatic control equipment. Depending on the consistency of the sludge, its end uses and quantity, spray dryers, drum dryers, opposed jet dryers, dryers with suspended and fluidized beds, pneumatic pipe dryers, and other different types of heat dryers can be used. Thermal drying can produce dry biosolids whose moisture content ranges from 10–40%, but it does require fuel for processing.

Incineration. Sludge is incinerated if it is impossible or economically infeasible to use, if storage is limited or unavailable, and also when it is required for sanitary and hygienic considerations. The most commonly used types of incinerators are multiple-hearth and fluidized-bed furnaces. Incineration dramatically reduces the volume of sludge and completely disinfects it. The preparation of sludge for incineration requires dewatering and/or thermal drying. When considering incineration, it is important to choose a method that yields safe exhaust gases and techniques to reduce the amount of energy required and/or provide some, if not all, of the required energy from another sludge process.

CONCLUSION

The purpose of sludge treatment is to reduce the moisture content and the volume of sludge; it renders the sludge harmless, prepares it for biosolids use, and moves biosolids from the WWTP. The most common types of sludge treatment are thickening, dewatering, and stabilization, which can be used one after the other consecutively and in various combinations. Stabilization refers to a number of processes, which reduce volatile solids, pathogen levels, vector attraction, and odor. Sludge must be stabilized before use and disposal. The variety of wastewater sludge treatment alternatives allows comparing and choosing cost-effective technology for each wastewater treatment plant.

READING LIST

Operation and Maintenance of Sludge Dewatering Systems. (1987). *Manual of Practice*. #OM-8. WPCF.

MOP. (1988). *Sludge Conditioning*. MOP ED-14.

WEF. (1996). *Standards for the Use and Disposal of Sewage Sludge—Final Rule and Phased—In Submission of Sewage Sludge Permit Application*. WEF.

WEF. (1996). *Wastewater and Sludge Treatment Processes*. WEF.

(1997). High-temperature sludge digesters and duckweed ponds reduce the cost of sludge disposal. *Chem Eng*. **February**: 81–83.

Bruce, A.M. and Fisher, W.J. (1994). *Sewage Sludge Stabilization and Disinfection*. Ellis Hardwood, Chichester, UK.

WEF. (1992). *Design of Municipal Wastewater Treatment Plants*. MOP 8, Vol. 2. WEF.

Epstein, E. (1997). *The Science of Composting*. Technomic, Lancaster, PA.

Foess, G.M. and Siger, R.B. (1993). Pathogen/vector attraction reduction requirement of the sludge rules. *Water Eng. Manage*. **June**: 25.

Ghosh, S. (1987). Improved sludge classification by two-phase anaerobic digestion. *Environ. Eng*. **June**: 1265–1284.

Lawler, D.F. and Chung, V.J. (1986). Anaerobic digestion. Effect on particle size and dewaterability. *W.P.C.F.* **December**: 1107.

Lue-Hing, C., Zenk, D.R., and Kuchenither, R. (1992). *Municipal Sewage Sludge Management*. Technomic, Lancaster, PA.

McDonald, G.J. (1995). Applying sludge to agricultural land-within the rules. *Water Eng. Manage.* **February**: 38.

Metcalf and Eddy, Inc. (1992). *Wastewater Engineering, Treatment, Disposal, and Reuse*. McGraw-Hill, New York, p. 1334.

Parkin, G.F. (1986). Fundamentals of anaerobic digestion of waste-water sludges. *Environ. Eng*. **May**: 867–920.

Spellman, F.R. (1997). *Dewatering Biosolids*. Technomic, Lancaster, PA.

Spellman, F.R. (1997). *Incinerating Biosolids*. Technomic, Lancaster, PA.

Turovskiy, I.S. (2001). Technological improvement for the aerobic digestion of sludge. *Water Eng. Manage.* **August**: 33–36.

US EPA. (1999). *Biosolids Generation, Use, and Disposal in the United States*. 530-R-99.

MUNICIPAL STORM WATER MANAGEMENT

C.A. Prochaska
A.I. Zouboulis
Aristotle University of Thessaloniki
Thessaloniki, Greece

INTRODUCTION

Enlargement of urbanization and industrial activities around the countryside have significantly altered the natural landscape of watersheds. The hydrological changes resulting from urbanization are shown in Fig. 1 (1).

This, in turn, has adversely affected both the quantity and the quality of storm water runoff and has contributed to the chemical, physical, and biological impairment of receiving waters. Several studies have shown that heavy metals, synthetic organics, pesticides, fuels, waste oils, and pathogens commonly contaminate storm water that originates from urban and industrial areas. To help improve the quality of storm water discharges, the U.S. Congress amended the Clean Water Act in 1987, which directs the Environmental Protection Agency (EPA) to develop the National Pollutant Discharge Elimination System (hereafter termed NPDES). Under NPDES regulations, the municipalities must develop storm water management plans that include specific controls to reduce the discharge of pollutants to the maximum extent practicable. The management plans must also address the legal, administrative and financial aspects of the municipality's storm water control program.

The municipal storm water management programs all involve similar elements. These include mainly public information/participation, elimination of illegal

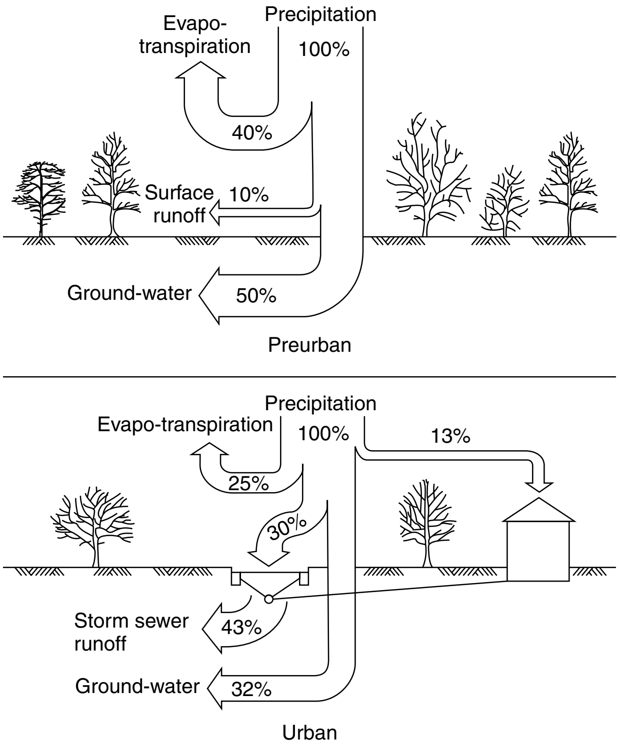

Figure 1. Comparison of water distribution before and after urbanization (1).

discharges, public agency activities, control of industrial/commercial storm water discharges, new development management, storm water treatment, program evaluation, and monitoring. The activities associated with each of these essential program components are presented briefly in the following (2):

PUBLIC INFORMATION/PARTICIPATION

This element is considered the most important early action and is the cornerstone of effective pollution prevention. Its objectives are to inform the public, commercial entries, and industries about the proper use and disposal of materials and waste and to correct the practices of storm water runoff pollution control. The public information activities include the development of general and focused information materials, as well as public service announcements. The participation activities include citizen monitoring programs, stenciling of storm drain inlets with “no dumping” signs and organized creek cleanups.

ELIMINATION OF ILLEGAL DISCHARGES

The elimination of illicit connections to the storm drain system and the prevention of illegal dumping are other essential early action elements. The objective is to ensure that only storm water or otherwise authorized discharges can enter the storm drains. The relevant activities include inspection of storm drain outfalls, surveillance of storm drain systems, and enforcement actions.

PUBLIC AGENCY ACTIVITIES

Many public agency activities may affect storm water pollution. Some activities prevent or remove storm water pollution; other activities are actually sources of pollution. The objective of this element is to ensure that routine municipal operations and maintenance activities are initiated or improved, to reduce the likelihood that pollutants are discharged to the storm drain system. The relevant activities include street sweeping; maintenance of storm drain inlets, lines, and channels and catch basins; corporation yard management; and the application of specific recycling programs. Coordination of road maintenance and flood control activities with storm water management program is also included.

CONTROL OF INDUSTRIAL/COMMERCIAL STORM WATER DISCHARGES

Industrial and commercial sources may contribute substantial pollutant loading to a municipal storm drain system. The objective of this element is to identify and effectively control the industrial and commercial sources of concern. The relevant activities include compilation of a list of industrial and commercial sources, identification of appropriate pollution prevention and control measures, and inspection of respective facilities. The focus is not only on facilities associated with industrial activity, as defined in the storm water regulations, but on any facility that conducts industrial activities, as well as commercial facilities, such as automotive operations and restaurants. This effort is expected to complement federal and state industrial storm water permitting efforts.

NEW DEVELOPMENT MANAGEMENT

New development (and redevelopment) areas offer the greatest potential for implementing the most effective pollution prevention and control measures. The objective of this element is to reduce the likelihood of pollutants entering the storm drain system from areas of new development or significant redevelopment during and after the construction period. The relevant activities include mainly the review of existing local permitting procedures and the modification of the procedures to identify and assign appropriate site design, erosion control, and permanent storm water control measures.

STORM WATER TREATMENT

The initial focus of storm water management programs is on pollution prevention and source control. The subsequent treatment of collected storm water is expected to be a rather costly alternative. There may be opportunities, however, for installation or retrofitting of structural control. The objectives of this element are to study the various treatment alternatives available, to test the feasibility of conducting the activities, and to determine the effectiveness of the treatment through pilot-scale projects. The main available treatment system alternatives are the following:

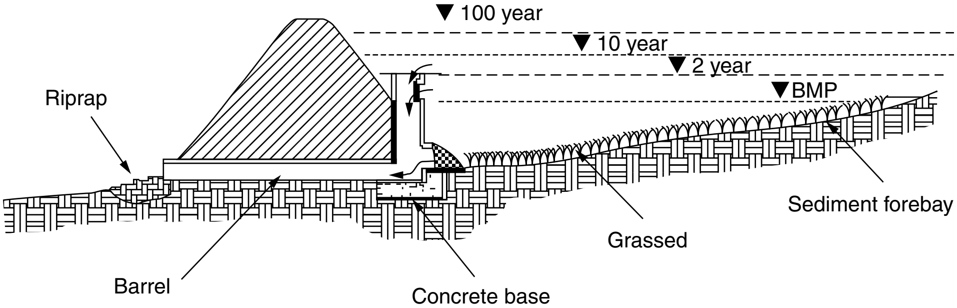

Figure 2. Detention basin (2).

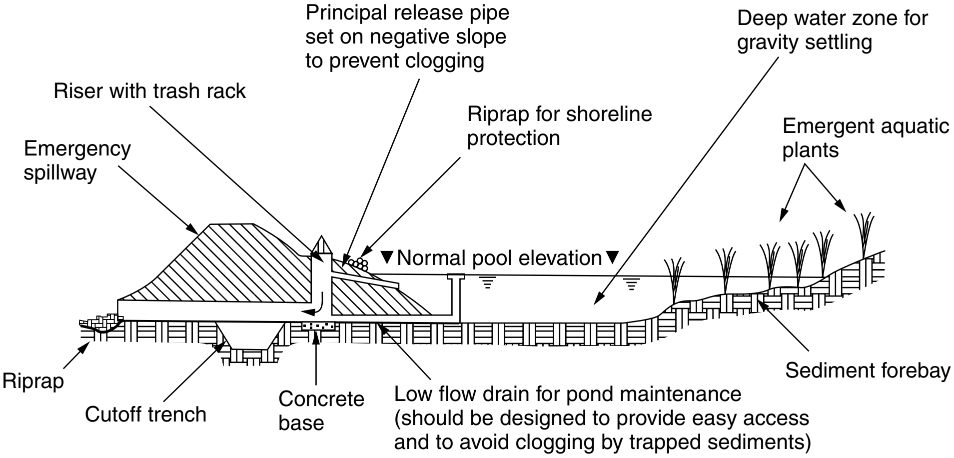

Figure 3. Retention pond (2).

1. *Infiltration systems*, which capture a volume of runoff and infiltrate it into the ground. Infiltration facilities may include infiltration basins, infiltration trenches, or porous pavement systems.
2. *Detention systems*, which capture a volume of runoff and temporarily retain that volume for subsequent release. Detention systems do not retain a significant permanent pool of water between runoff events; a common type is shown in Fig. 2.
3. *Retention systems*, which capture a volume of runoff and retain that volume until it is displaced in part or in total by the next runoff. Retention systems, therefore, maintain a significant permanent pool volume of water between runoff events. The details of a retention pond are shown in Fig. 3.
4. *Constructed wetland systems* are similar to retention and detention systems, except that a major portion of the water surface area (in pond systems) or bottom (in meadow-type systems) contains wetland vegetation. This group also includes wetland channels. A typical wetland system design is shown in Fig. 4 (3).
5. *Filtration systems* use some combination of granular filtration media, such as sand, soil, organic material, carbon or a membrane, to remove constituents found in runoff.
6. *Vegetated systems (biofilters or bioretention systems)*, such as swales and filter strips, are designed to convey and treat either shallow flow (swales) or sheet flow (filter strips) runoff. A diagram of a typical bioretention area is shown in Fig. 5.
7. The method of *minimizing directly connected impervious surfaces* describes a variety of practices that can be used to reduce the amount of surface area directly connected to the storm drainage system by minimizing or eliminating the traditional curb and gutter. This is considered to a nonstructural practice, but it has been included here because of the need to design and construct alternative conveyance and treatment options.
8. *Miscellaneous and vendor-supplied systems*, which include a variety of proprietary and miscellaneous systems that do not fit under any of the above categories. These may include catch basin inserts, hydrodynamic devices, and filtration devices.

PROGRAM EVALUATION

Storm water management programs are expected to change as they mature. Consequently, they should have built-in flexibility to allow for changes in priorities, needs, or levels of awareness. The objective of this element is to provide a comprehensive annual evaluation and report of the program's effectiveness. The measures of effectiveness include mainly quantitative monitoring to assess the effectiveness, the specific control measures, and the detailed accounting of program accomplishments and funds and staff hours expended. The annual report provides an overall evaluation of the program and sets forth plans and schedules for the coming year. The annual report is considered a program's self-audit and

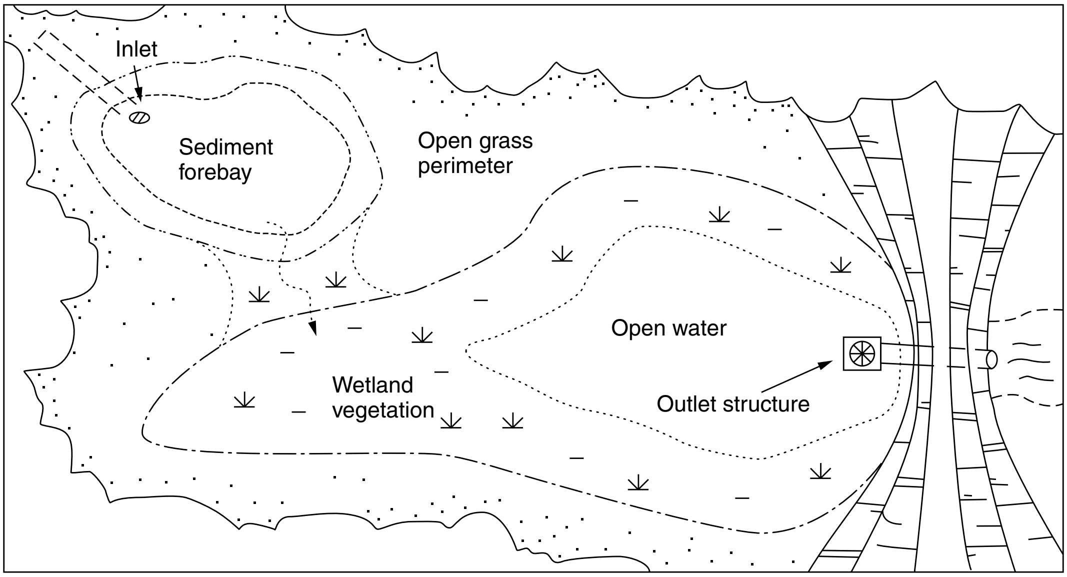

Figure 4. Constructed wetland system (3).

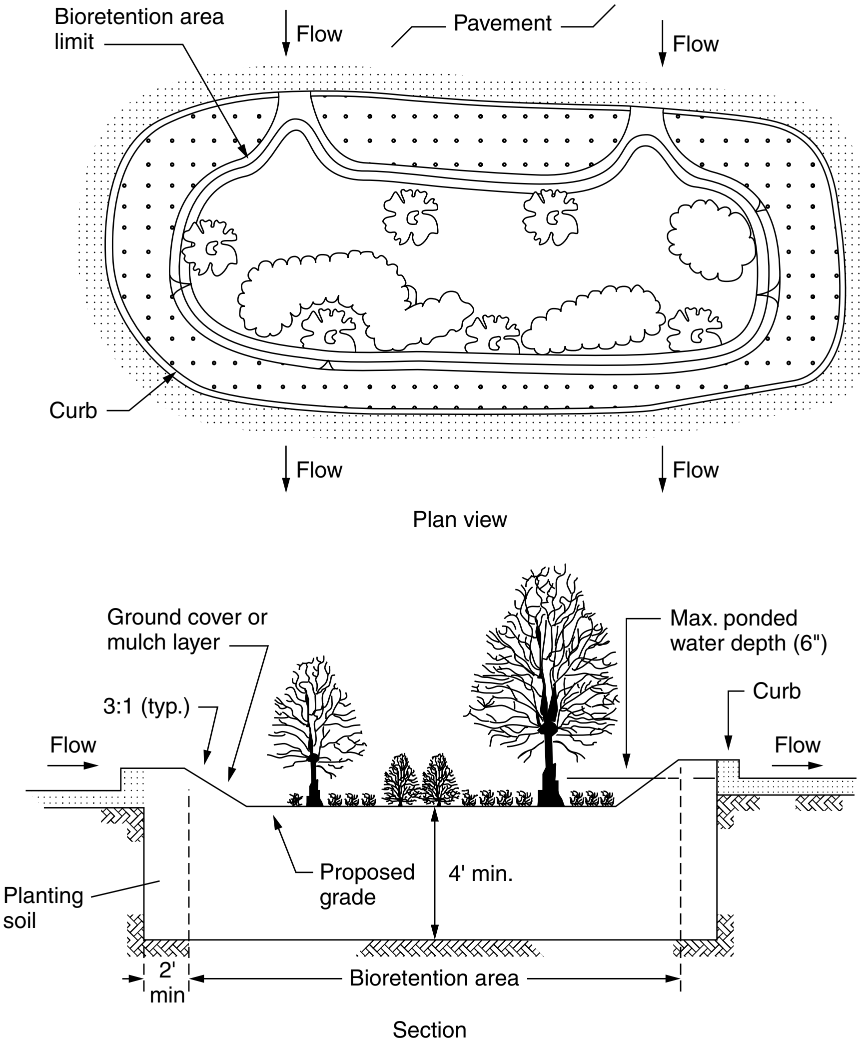

Figure 5. Bioretention system (3).

provides a mechanism to propose modifications to the storm water management plan in response to program accomplishments or failures. The annual report also serves as the key regulatory tool for providing accountability and public review in accordance with the respective NPDES permit.

MONITORING

Monitoring is an essential component of any pollution control program. The objectives are to obtain quantitative information to measure program progress and effectiveness, to identify the sources of pollutants, and to document the reduction of pollutant loads (if any). The success of the monitoring program can be measured by the ability to make more informed decisions on a program's direction and effectiveness. The monitoring activities include primarily the baseline monitoring of storm drain discharges and receiving waters and is focused on special studies to identify sources of pollutants and to evaluate the effectiveness of specific control measures. The types of monitoring may include water column measurements, sediment measurements, and nonsampling and analysis measurements, such as the number of outfalls inspected or the amount of material removed by regular maintenance actions. Toxicity identification and evaluations are also integral components of monitoring programs.

BIBLIOGRAPHY

1. Quasim, S.R. (1999). *Wastewater Treatment Plants: Planning, Design and Operation*. Technomic, Lancaster, PA.
2. U.S. EPA. (1998). *N.P.D.E.S. Storm Water Phase II Fact Sheets*, Washington, DC.
3. U.S. EPA. (1995). *National Conference on Urban Runoff Management: Enhancing Urban Watershed Management at the Local, County and State Levels*. Report EPA/625/R-95/003. Chicago, IL.

WHAT WASTEWATER UTILITIES CAN DO NOW TO GUARD AGAINST TERRORIST AND SECURITY THREATS

U.S. Environmental Protection
Agency—Office of
Wastewater Management

One consequence of the events of September 11th is a heightened concern among citizens in the United States over the security of their critical wastewater infrastructure. The nation's wastewater infrastructure consisting of approximately 16,000 publicly owned wastewater treatment plants, 100,000 major pumping stations, 600,000 miles of sanitary sewers and another 200,000 miles of storm sewers, is one of America's most valuable resources, with treatment and collection systems valued at more than $2 trillion. Taken together, the sanitary and storm sewers form an extensive network that runs near or beneath key buildings and roads, and is contiguous to many communication and transportation networks. Significant damage to the nation's wastewater facilities or collection systems would result in: lose of life, catastrophic environmental damage to rivers, lakes and wetlands, contamination of drinking water supplies, long term public health impacts, destruction of fish and shellfish production, disruption to commerce, the economy and our normal way of life. Although many wastewater utilities have already taken steps to increase security, the following recommendations provide many straightforward, commonsense actions to increasing security and reducing threats from terrorism. Many of these actions are recommended by the Association of Metropolitan Sewer Agencies, the Water Environment Federation, and other leading professional organizations. The recommendations include:

GUARDING AGAINST UNPLANNED PHYSICAL INTRUSION

- Lock all doors and set alarms at your office, pumping stations, treatment plants, and vaults, and make it a rule that doors are locked and alarms are set;
- Limit access to facilities and control access to pumping stations, chemical and fuel storage areas, giving close scrutiny to visitors and contractors;
- Post guards at treatment plants, and post "Employee Only" signs in restricted areas;
- Control access to storm sewers;
- Secure hatches, metering vaults, manholes and other access points to the sanitary collection system;
- Increase lighting in parking lots, treatment bays, and other areas with limited staffing;
- Control access to computer networks and control systems, and change the passwords frequently;
- Do not leave keys in equipment or vehicles at any time.

MAKING SECURITY A PRIORITY FOR EMPLOYEES

- Conduct background security checks on employees at hiring and periodically thereafter;
- Develop a security program with written plans and train employees frequently;
- Ensure all employees are aware of communications protocols with relevant law enforcement, public health, environmental protection, and emergency response organizations;
- Ensure that employees are fully aware of the importance of vigilance and the seriousness of breaches in security, and make note of unaccompanied strangers on the site and immediately notify designated security officers or local law enforcement agencies;
- Consider varying the timing of operational procedures if possible so if someone is watching the pattern changes.

- Upon the dismissal of an employee, change passcodes and make sure keys and access cards are returned;
- Provide Customer Service staff with training and checklists of how to handle a threat if it is called in.

COORDINATING ACTIONS FOR EFFECTIVE EMERGENCY RESPONSE

- Review existing emergency response plans, and ensure they are current and relevant;
- Make sure employees have necessary training in emergency operating procedures;
- Develop clear protocols and chains-of-command for reporting and responding to threats along with relevant emergency management, law enforcement, environmental, public health officials, consumers and the media. Practice the emergency protocols regularly;
- Ensure key utility personnel (both on and off duty) have access to crucial telephone numbers and contact information at all times. Keep the call list up to date;
- Develop close relationships with local law enforcement agencies, and make sure they know where critical assets are located. Request they add your facilities to their routine rounds;
- Work with local industries to ensure that their pretreatment facilities are secure;
- Report to county or State health officials any illness among the employees that might be associated with wastewater contamination;
- Report criminal threats, suspicious behavior, or attacks on wastewater utilities immediately to law enforcement officials and the relevant field office of the Federal Bureau of Investigation.

INVESTING IN SECURITY AND INFRASTRUCTURE IMPROVEMENTS

- Assess the vulnerability of collection system, major pumping stations, wastewater treatment plants, chemical and fuel storage areas, outfall pipes, and other key infrastructure elements;
- Assess the vulnerability of the storm water collection system. Determine where large pipes run near or beneath government buildings, banks, commercial districts, industrial facilities, or are contiguous with major communication and transportation networks;
- Move as quickly as possible with the most obvious and cost-effective physical improvements, such as perimeter fences, security lighting, tamper-proofing manhole covers and valve boxes, etc.;
- Improve computer system and remote operational security;
- Use local citizen watches;
- Seek financing for more expensive and comprehensive system improvements.

While wastewater utilities are the key to improving security of our wastewater treatment plants and collection systems, EPA, other Federal agencies, and both industry and managerial trade associations also provide help and support. EPA is working with AMSA and other groups to develop training courses and technical materials for wastewater utilities and State personnel on assessing vulnerabilities and improving security. EPA is working collaboratively with the Association of Metropolitan Water Agencies and other groups to develop an Information Sharing and Analysis Center to bolster coordinated notification and response to threats and vulnerabilities at both water and wastewater facilities. A number of technical projects are underway to help increase security of the nation's critical wastewater infrastructure.

FOR MORE INFORMATION

For more information please visit the following web sites:

EPA Counter-terrorism: http://www.epa.gov/ebtpages/ecounterterrorism.html
EPA Alert on Chemical Accident Prevention and Site Security: http://www.epa.gov/ceppo/pubs/secale.pdf
Association of Metropolitan Sewer Agencies: http://www.amsa-cleanwater.org
Association of Metropolitan Water Agencies: http://www.amwa.net/isac/amwacip.html
Water Environment: http://www.wef.org
National League of Cities: http://www.nlc.org/nlc_org/site/newsroom/terrorism_response

WASTEWATER TREATMENT TECHNIQUES—ADVANCED

Miguel A. Valenzuela
Instituto Politecnico
Nacional—ESIQIE. MEXICO

ADVANCED OXIDATION PROCESSES (AOP)

All advanced oxidation processes are characterized by a common chemical feature: production of $OH^\bullet$ radicals.

Table 1 and Fig. 1 show a list of AOP and their applicability.

These radicals are suitable for achieving complete abatement and mineralization of pollutants. AOP usually operate at or close to ambient temperature and pressure. The potentialities offered by AOP can be exploited to integrate biological treatments by oxidative degradation of toxic substances, entering or leaving the biological stage (1–2). The usual two AOP are the Fenton process and photocatalysis:

Fenton Process

Production of $OH^\bullet$ radicals by Fenton's reagent occurs when addition of H_2O_2 is added to Fe^{2+} salts (3): It has been demonstrated that Fenton's reagent can destroy toxic

Table 1. Some Advanced Oxidation Technologies

Reactions
Fenton-Type reactions
$Fe^{2+} + H_2O_2 \rightarrow {}^{\bullet}OH + Fe^{3+} + OH^-$
Ozone—peroxide—UV systems
$O_3 + {}^{-}OH \rightarrow O_2{}^{-} \rightarrow {}^{\bullet}OH$
$3O_3 + UV(< 400\ nm) \rightarrow 2{}^{\bullet}OH$
$H_2O_2 + UV(< 400\ nm) \rightarrow 2{}^{\bullet}OH$
$H_2O_2 + O_3 \rightarrow 2{}^{\bullet}OH$
$H_2O_2 + O_3 + UV \rightarrow {}^{\bullet}OH$
Semiconductor oxides—UV systems
$TiO_2 + hv \rightarrow TiO_2(h^+ + e^-)$
$H^+ + OH^- \rightarrow {}^{\bullet}OH$
Radiolysis (high-energy beams)
$H_2O \rightarrow e^-{}_{aq} + H^{\bullet} + {}^{\bullet}OH + (H_2, H_2O_2, H_3O^+)$
Wet oxidation (WO) systems
$RH + O_2 \rightarrow R^{\bullet} + HO_2{}^{\bullet}$
$RH + HO_2{}^{\bullet} \rightarrow R^{\bullet} + H_2O_2$
$H_2O_2 + M \rightarrow 2\ OH^{\bullet}$
$RH + OH^{\bullet} \rightarrow R^{\bullet} + H_2O$
$R^{\bullet} + O_2 \rightarrow ROO^{\bullet}$
$ROO^{\bullet} + RH \rightarrow ROOH + R^{\bullet}$
Sonolysis (ultrasound)
$H_2O \rightarrow H^{\bullet} + {}^{\bullet}OH$

Source: Reference 26.

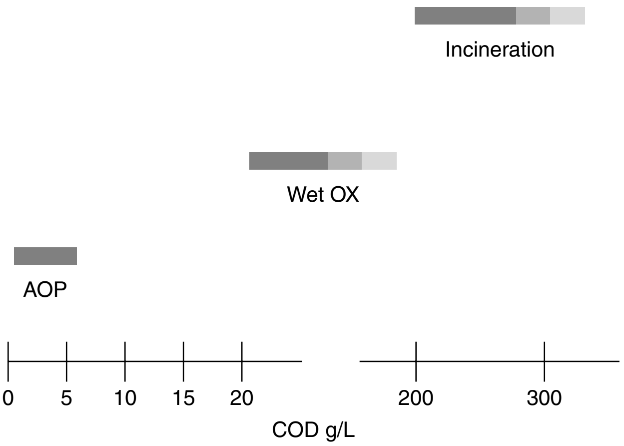

Figure 1. Suitability of water treatment technologies according to COD contents. (*Source*: Reference 1).

compounds such as phenols and herbicides in wastewaters. Irradiation by UV-vis light strongly accelerates the degradation rate of organic pollutants (4). The application of the Fenton process requires strict pH control; sludges can be formed which create disposal problems.

Photocatalysis

In this process, hydroxyl and other radicals are generated at the surface of an UV-absorbing powder (called a photocatalyst). The most widely used photocatalyst is the wide band-gap (3.2 eV) semiconductor TiO_2 in its anatase crystalline form (5–6). TiO_2 absorbs UV light at wavelengths below ~ 380 nm creating an excess of electrons in the conduction band ($e^-{}_{cb}$) and holes in the valence band ($h^+{}_{vb}$):

$$TiO_2 \longrightarrow (e^-{}_{cb} + h^+{}_{vb})$$

The carriers can diffuse to the surface where they react as follows:

$$h^+{}_{vb} + OH^- \longrightarrow OH^{\bullet}{}_{ads}$$

and

$$e^-{}_{cb} + O_2 \longrightarrow O_2{}^{\bullet -}{}_{ads}$$

where ads = adsorbed to the surface of TiO_2 particles.

Organic pollutants may adsorb on the surface of TiO_2 particles, and there they are attacked by the adsorbed $OH^{\bullet}$ radicals and holes. The $O_2{}^{\bullet -}$ radicals can further disproportionate as follows:

$$2O_2{}^{\bullet -}{}_{ads} + 2H^+ \longrightarrow H_2O_2 + O_2$$

Although the quantum yield of TiO_2 photocatalyzed reactions is rather low, the system does have the advantage that it can use UV photons in the near UV (blacklight UV fluorescent lamps or the UV portion of solar radiation) (Fig. 2). Compilations of substances which can be mineralized using photocatalysis are now available (7).

COMPLEXATION/FLOCCULATION

It has been shown that dissolved humic substances (DHS), bind (complex) organic solutes via hydrophobic interactions, forming humic–contaminant complexes in the aqueous phase. The use of DHS in flushing solutions to enhance desorption of hydrophobic contaminants from sediments was suggested by several researchers (8–10). It was further proposed to remove the humic–contaminant complexes by flocculation using alum or ferric salts, followed by press filtration and incineration or disposal of the resulting precipitate.

The treatment process follows two stages: (1) binding of Dissolved Humic Acid (DHA) by the dissolved contaminants to form complexes (complexation stage) and (2) precipitation of DHA and the associated contaminant by using a flocculant (alum or ferric chloride, flocculation stage). This process can be applied to remove various classes of hydrophobic organic pollutants such as PAHs, PCBs and chloro-organo pesticides from industrial wastewater. Additionally, this technology has the advantage that it may be coupled to the general water treatment process (Fig. 3) (11)

CONDUCTING POLYMERS

Conducting polymers have ion exchange properties induced by charging and discharging processes (12). For instance, polypyrrole (PPy) can function as an anion exchanger, whereas PPy modified with polystyrenesulfonate anions (PPy/PSS$^{\bullet}$) works as a cation exchanger (Fig. 4) (13,14). Such a modified polymer can be used as an electrochemically switchable ion exchanger for water softening (15). This ion exchanger can be regenerated without chemical additives or aqueous electrolysis.

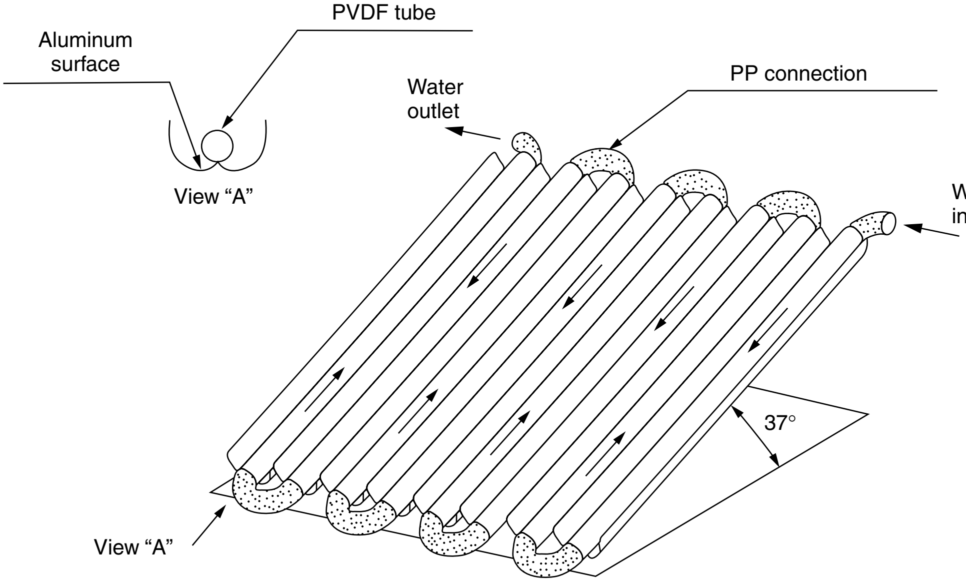

Figure 2. Scheme of one CPC module used for solar detoxification of water. (*Source*: J.M. Hermann, et al. (1998). *Appl. Catal. B: Environmental* **17**: 15).

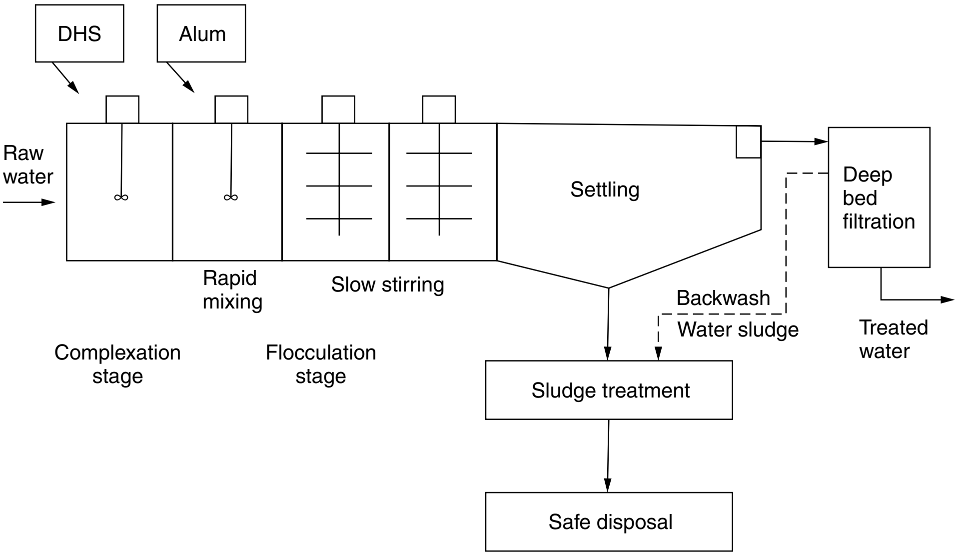

Figure 3. Flow diagram of the complexation—flocculation process coupled with a common water treatment operation. (*Source*: Reference 11).

IONIZING IRRADIATION

High energy irradiation (γ rays, X rays, and electron beams) interacts with water to generate a variety of free radicals, principally $OH^{\bullet}$, $H^{\bullet}$ and hydrated electrons. If H_2O_2 or O_3 is present in the water, the $H^{\bullet}$ and the hydrated electrons are converted efficiently to $OH^{\bullet}$ radicals. This process is based on electron accelerators. An attempt has been made to use combined electron beam and ozone for treating municipal wastewater in aerosol flow (Fig. 5) (16,17).

MEMBRANE/SONICATION/WET OXIDATION

Hybrid systems are becoming popular for treating waste streams that are otherwise difficult to handle. For instance, the powder-activated carbon–activated sludge system (PACT system by Zimpro Environmental, Inc.) is a classic example of such systems. OXYMEM is another hybrid process, where wet oxidation and nanofiltration were used together to treat bioresistant industrial wastewater containing polyethylene glycol. It has also been demonstrated that sonication followed by wet oxidation (SONIWO) is a useful hybrid process for treating refractory waste. Conventional bioprocesses may not be amenable to biodestruction of the effluent from reactive bath dye. "Membrane-sonication-wet oxidation" (MEMSONIWO) is a hybrid process applied to water conservation via recycling. The membrane unit allows concentrating the waste, and then the permeate (mostly water) can be recycled. The concentrate from the membrane unit can, then, be treated by sonication to

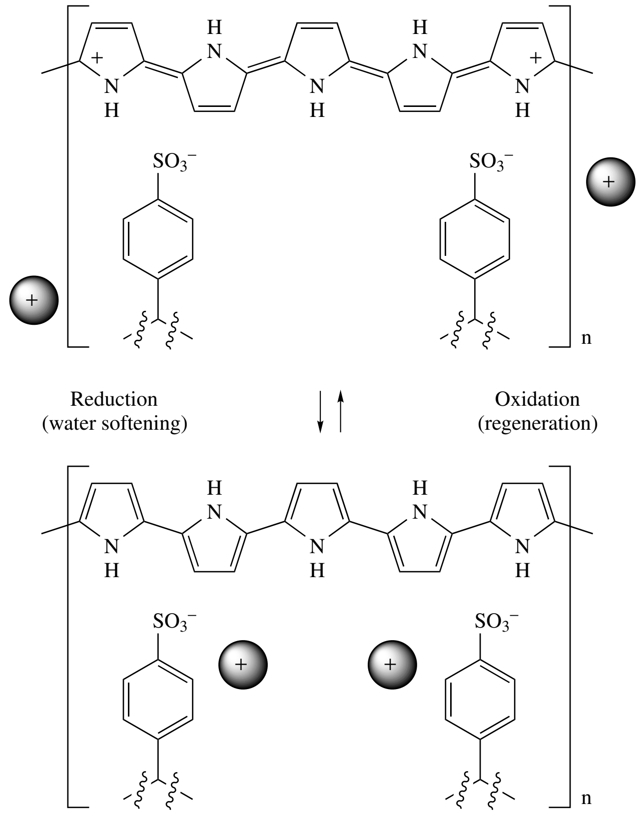

Figure 4. Polypyrrole with incorporated PSS anions working as a cation exchanger. (*Source*: Reference 15).

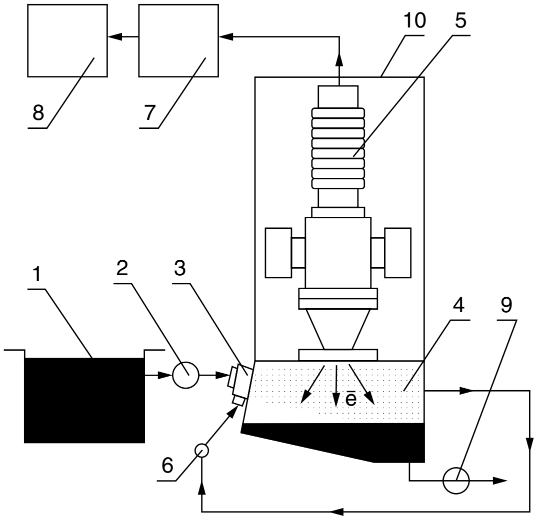

Figure 5. Scheme of pilot plant for combined electron-beam and ozone treatment of municipal wastewater in aerosol flow. (1) Reservoir of wastewater intake (2) Electric pump unit for wastewater (3) Sprayer unit (4) Irradiation chamber (5) Electron accelerator (6) Turboblower (7) Power supply (8) Control desk (9) Electric pump unit for purified water removal (10) Biological shielding (*Source*: Reference 17).

make it suitable for wet oxidation. After wet oxidation, the water can be discharged or recycled (18).

SORPTION BY ZEOLITES

It is well established that the sorption characteristics of zeolite-type materials are defined by pore size and charge properties (19). Most naturally occurring zeolites bear a relatively high framework charge arising from Al^{3+} substitution for Si^{4+} in the crystal lattice; this results in a structure of high cation-exchange capacity. Such zeolites have been used as ion exchangers to treat water and are incorporated into systems for treating radioactive waste (removal of $^{137}Cs^+$ and $^{90}Sr^{2+}$) and for removing NH_4^+ from wastewater. Zeolites that have high SiO_2/Al_2O_3 ratios have a low capacity to retain cations but are more hydrophobic and can, therefore, sorb uncharged molecules. In laboratory studies using batch sorption equilibria, high Si large-pore mordenite (MOR) and ZSM-5, it was found, have sorption properties for methyl tert-butyl ether (MTBE) and trichloroethylene (TCE) that are superior to those of activated carbon (Table 2) (20).

SUPERCRITICAL WATER OXIDATION

Supercritical water oxidation (SCWO) is considered a promising technology for treating several wastes (21–24). SCWO is a process where oxidation takes place in water above its critical point (647 K, 22.1 MPa). SCWO is an environmentally acceptable technology that produces a disposable clean liquid (pure water), clean solid (metal oxides, salts), and clean gas (CO_2, N_2). Recently, there has been increasing interest in using heterogeneous catalysts in SCWO (Fig. 6). Catalysts can increase the oxidation rates, reduce the residence times and temperatures

Table 2. Solution Concentrations and Percent Removal after Equilibration of 100 μg/L Solutions with 5 mg of Solid Phase[a]

	MTBE		$CHCl_3$		TCE	
Sorbent	μg/L	% Removal	μg/L	% Removal	μg/L	% Removal
MOR	4.0 ± 0.3a	96	62.2 ± 5.2a	38	23.2 ± 4.4a	77
ZSM-5	36.6 ± 9.4b	63	21.9 ± 1.2b	78	< 3.0b[b]	> 97
Y	94.9 ± 6.1c	5	99.6 ± 3.9c	< 1	90.9 ± 1.7c	9
AC	48.2 ± 3.9b	52	43.8 ± 2.2d	56	32.2 ± 1.6a	68

[a] Mean values within columns followed by the same letter are not significantly different from each other at $p \leq .05$.

[b] 3 μg/L was the detection limit for the method used.

Source: Reference 20.

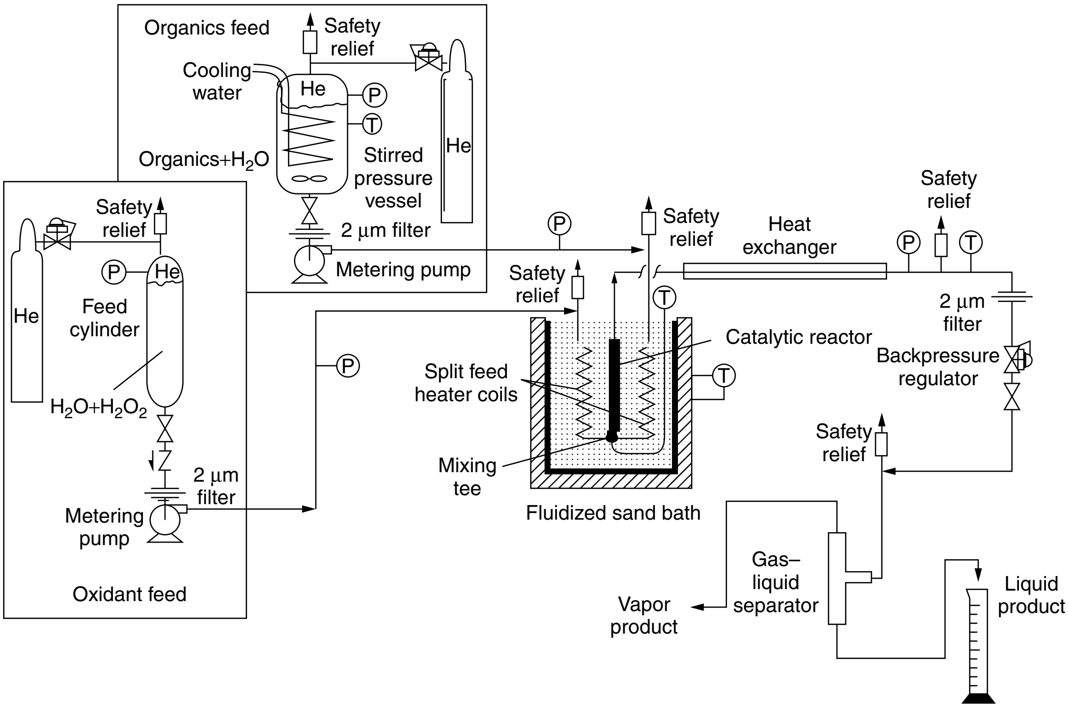

Figure 6. Experimental apparatus for catalytic SCWO. (*Source*: Reference 25).

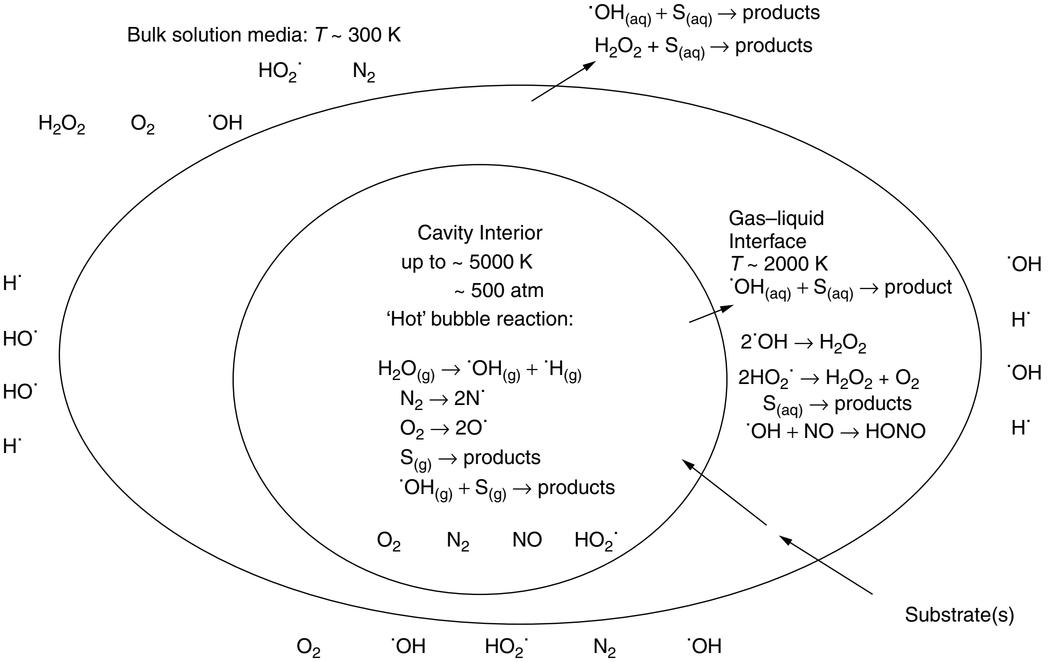

Figure 7. Three reaction zones in the cavitation process. (*Source*: Reference 26).

required for treatment, and possibly control the selectivity of the reaction pathways (25).

ULTRASONIC IRRADIATION

Sonochemical effects are due to the phenomenon of "cavitation," the nucleation and the behavior of bubbles in a liquid (26,27). In wastewater treatment, a bubble of cavitation may function as a microreactor which destroys volatile organic compounds inside (28–30). The cavity may also be thought as a $H^{\bullet}$, $OH^{\bullet}$, $OOH^{\bullet}$ radical source that react with pollutants in the bulk of the solution (Fig. 7). Several potential applications of ultrasonic irradiation have been reported recently.

BIBLIOGRAPHY

1. Andreozzi, R., Caprio, V., Insola, A., and Marotta, R. (1999). *Catal. Today* **53**: 51.
2. Ollis, D.F. (1993). In: *Comparative Aspects of Advanced Oxidation Processes, Emerging Technologies in Hazardous Waste Management III*. D.W. Tedder and F.G. Pohland (Eds.). Chap. 2, ACS Symposium Series 518, Washington, DC.
3. Haber, F. and Weiss, J. (1934). *Proc. R. Soc. Ser. A* **147**: 332.
4. Bauer, R. et al. (1999). *Catal. Today* **53**: 131.
5. Mills, A. and Hunte, S.L. (1997). *J. Photochem. Photobiol. A: Chem.* **108**: 1.
6. Bahnemann, D. (1999). *The Handbook of Environmental Chemistry*. Vol. 2, Part I, Environmental Photochemistry, P. Boule (Ed.). Springer-Verlag, Berlin.
7. Blake, D.M. (1999). *Bibliography of Work on the Photocatalytic Removal of Hazardous Compounds from Water and Air*. Report NREL/TP-570-26797, National Renewable Energy Laboratory, Golden, CO.
8. Johnson, W.P. and Amy, G.L. (1995). *Environ. Sci. Technol.* **29**: 807.
9. Rebhun, M., Smedt, F.D., and Rwetabula, J. (1996). *Water Res.* **30**: 2027.
10. Liu, H. and Amy, G.L. (1993). *Environ. Sci. Technol.* **27**: 1553.
11. Rebhun, M., Meir, S., and Laor, Y. (1998). *Environ. Sci. Technol.* **32**: 981.
12. Shimidzu, T., Ohtani, A., and Honda, K. (1988). *J. Electroanal. Chem.* **88**: 323.
13. Zhong, C., Doblhofer, K., and Weinberg, G. (1989). *Faraday Discuss. Chem. Soc.* **88**: 307.
14. Ren, X. and Pickup, P.G. (1993). *J. Phys. Chem.* **97**: 5356.
15. Weidlich, C., Mangold, K-M., and Juttner, K. (2001). *Synth. Met.* **119**: 263.
16. Pikaev, A.K. et al. (1996). *Radiat. Phys. Chem.* **48**: 75.
17. Pikaev, A.K., Podzorova, E.A., and Bakhtin, O. (1997). *Radiat. Phys. Chem.* **49**: 155.
18. Dhale, A.D. and Mahajani, V.V. (1999). *Ind. Eng. Chem. Res.* **38**: 2058 and references therein.
19. Newsam, J.M. (1986). *Science* **231**: 1093.
20. Anderson, M.A. (2000). *Environ. Sci. Technol.* **34**: 725.
21. Mishra, V.S., Mahajani, V.V., and Joshi, J.B. (1995). *Ind. Eng. Chem. Res.* **34**: 2.
22. Matatov-Meytal, Y.I. and Sheintuch, M. (1998). *Ind. Eng. Chem. Res.* **37**: 309.
23. Gloyna, E.F. and Li, L. (1995). *Environ. Prog.* **14**: 182.
24. Levec, J. (1997). *Chem. Biochem. Eng. Q* **11**: 47.
25. Yu, J. and Savage, P.E. (1999). *Ind. Eng. Chem. Res.* **38**: 3793 and references therein.
26. Adewuyi, Y.G. (2001). *Ind. Eng. Chem. Res.* **40**: 4681 and references therein.
27. Thompson, L.H. and Doraiswamy, L.K. (1999). *Ind. Eng. Chem. Res.* **38**: 1215.
28. Pétrier, C. and Francony, A. (1997). *Ultrasonic Sonochemical* **4**: 295.
29. Phull, S.S. et al. (1997). *Ultrasonic Sonochemical* **4**: 157.
30. Pandit, A.B., Gogate, P.R., and Mujumdar, S. (2001). *Ultrasonic Sonochemical* **8**: 227.

TRENCHLESS REPAIR AND REHABILITATION TECHNIQUES

Sanjiv Gokhale
Vanderbilt University
Nashville, Tennessee

Trenchless technology (TT) consists of a wide range of methods, materials, and equipment for installing new or rehabilitating existing underground pipelines and utility systems with minimal excavation of the ground. According to the North America Society of Trenchless Technology (NASTT), trenchless construction is “a family of methods, materials, and equipment capable of being used for the installation of new or replacement or rehabilitation of existing underground infrastructure with minimal disruption to surface traffic, business, and other activities.”

Open-cut trench construction has proven expensive, especially in congested urban areas, because it requires disruption of surface activities. Some problems caused by the open-cut method include traffic disruption, economic impact on local businesses, damage to existing utilities, and concerns for worker safety. The development and use of trenchless techniques has expanded rapidly over the past 10 years. The reason for this exceptional growth is the desire to install or rehabilitate underground pipeline systems with minimum impact on society and the environment. The benefits of trenchless technology are quite apparent compared with the conventional open-cut process. However, it is necessary to evaluate the suitability and appropriateness of any rehabilitation system, trenchless or otherwise, on a project-by-project basis with due consideration of all project criteria, such as size and material of pipe, level of deterioration, hydraulic capacity, etc.

According to a 2003 survey (http://www.oildompublishing.com/UC/uchome.html) of trenchless construction methods used by municipalities, the percentage of all municipal projects using trenchless construction methods has grown by 180% (new installation) and 270% (rehabilitation), respectively, over the past 5 years.

BASICS OF PIPELINE RENOVATION USING TRENCHLESS TECHNIQUES

Trenchless pipeline renewal methods offer several advantages over conventional dig-up and repair/replace methods:

- minimize disturbance to existing site, underground utilities, and environment;
- are more suitable for difficult underground conditions, such as high water table and unstable soils;
- require less exposed work area, therefore involving lesser risk to the workers and public; and
- minimize the need for spoil removal and disposal.

The primary trenchless pipeline-system renewal methods can be divided into five categories (Table 1):

Table 1. Trenchless Pipeline Renewal Methods

Method	Diameter Range, in.	Maximum Installation Lengths, ft.	Liner Material[a]	Application
		Sliplining:		
Segmental	12–150	5000	PE, PP, PVC, GRP	Gravity & pressure
Continuous	4–60	1000	PE, PP, PVC, GRP	Gravity & pressure
Spiral wound	4–100	1000	PE, PP, PVC, PVDF	Gravity
		CIPP:		
Inverted in place	4–108	3000	Thermoset resin	Gravity & pressure
Winched in place	4–54	500	Thermoset resin	Gravity & pressure
		Close-Fit:		
Swaged	4–15	700	HDPE, PVC	Gravity & pressure
Folded	3–24	1000	HDPE, MDPE	Gravity & pressure
Expanded spiral	4–36	1000	HDPE, MDPE	Gravity & pressure
		Spray-on Lining:		
Cement mortar	4–36	NA	Cement mortar	Gravity & pressure
Shotcrete	>42	NA	Cement mortar	Gravity & pressure
Epoxy	4–24	NA	Epoxy	Gravity & pressure
		Point Source Repairs:		
Robotic	8–30	NA	Epoxy & cement mortar	Gravity
Grouting	NA	NA	Chemical grouting	Any
Mechanical sleeve	4–24	NA	Mechanical sleeves	Any
Point CIPP	4–24	50	Fiberglass/polyester resin	Gravity

[a] **Definitions of Acronyms**: PE: Polyethylene; PP: Polypropylene; PVC: Polyvinyl chloride; PVDF: Polyvinylidene chloride; GRP: Glassfiber reinforced polyester; HDPE: High density polyethylene; MDPE: Medium density polyethylene.

- sliplining
- cured-in-place pipe (CIPP)
- close-fit pipe
- spray lining
- point source repair

Sliplining

Sliplining is one of the earliest forms of trenchless pipeline rehabilitation. There are three main types of sliplining: continuous, segmental, and spiral wound. A new pipe of smaller diameter is inserted by pulling, pushing, or spiral winding it into the host pipe, and the annulus between the existing pipe and the new pipe is grouted. Small liners may be pulled in manually but most require a winch (Fig. 1). The winch applies a steady, progressive pull to place the liner inside the host pipe. The liner pipe is generally butt-fused to its design length. Numerous designs of pipe pushing machines, both manual and hydraulic, are available. In most instances, an insertion pit is required for the pushing machine. The pushing machine grips the liner pipe and pushes it forward into the host pipe. The gripping mechanism is then released and returned to the starting position, and the process is repeated. Segmental liners are generally used to reduce the size of the insertion pit. Pipe joints are generally of the mechanical type with either a snap-fit or a screw-on mechanism. For spirally wound liners, strips of PVC material are passed through a winding machine and a helically wound liner pipe is manufactured *in situ*. To increase the stiffness of the liner pipe, ribbed strips are used, with "T-beams" forming on the outer surface. The winding machine is normally located down a manhole or a small excavation. The tube travels down the host pipe as more strips are fed into the machine. The whole tube is rotating during the installation, so the limiting factor is usually the friction and weight of the liner that the winding machine is capable of turning. An alternate spirally wound technique overcomes this drawback by using a winding machine that travels through the host pipe, thereby removing the need to rotate the liner itself.

This method has the merit of simplicity and is relatively inexpensive. One of the chief drawbacks of sliplining is the resulting decreased cross-sectional area. In some instances, however, despite the reduced cross section, the hydraulic capacity of the pipeline may actually increase due to the superior flow characteristics of the new pipe. The reconnection of laterals and service lines in conjunction with sliplining of gravity pipelines usually necessitates excavation. Excavation must take place and the lateral must be disconnected before grouting. Electrofusion is commonly used to connect laterals to PE liners in the same way as new installations. For glass-fiber reinforced polyester resin (GRP) pipe, laterals may be typically reconnected to the new liner pipe using "tees."

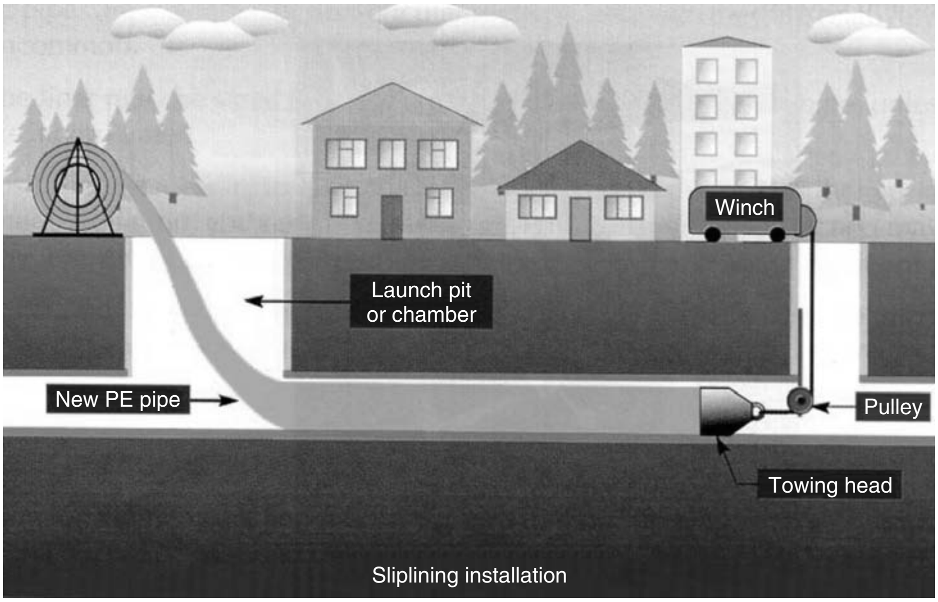

Figure 1. Continuous sliplining.

Cured-In-Place Lining (CIPP)

The main alternative to sliplining and its variants in the non-man-entry pipeline renovation market is the cured-in-place lining (CIPP), sometimes referred to as "*in situ* lining." Although several patented systems are currently available, the common feature of CIPP is the use of a polyester or epoxy resin impregnated fabric tube. The tube is inserted into the existing host pipe and inflated against the wall of the host using a hydrostatic head or air pressure (Fig. 2). The inflated liner is cured by recirculating hot water or steam. The CIPP process creates a "close-fit" pipe that has quantifiable structural strength and can be designed for specific loading conditions.

The chief advantages of CIPP is that it minimizes the reduction in cross section and the liner pipe can conform to noncircular cross sections. The laterals can be reopened remotely after lining by using a remote controlled robotic cutter. The chief disadvantage of this method is the need to take the host pipe out of service during installation and curing. Diversion or bypass pumping adds to the installation cost. CIPP is not cost-effective for large diameters.

As in all trenchless renovation systems, thorough cleaning and preparation are essential. In non-man-entry pipelines, CCTV inspection should be carried out prior to and after installation. All silt and debris should be removed completely by jetting or other means prior to installation. Intruding connections, encrustation, and other deposits should be removed by mechanical means or high-pressure jetting, followed by cleaning to remove debris. There may be short-term environmental implications with CIPP systems when using polyester resins. The solvents in the systems can give off strong odors and in high concentrations, the vapor can pose a health risk. Such levels are not typically found in CIPP installations; nonetheless, adequate ventilation of a CIPP work site is essential. This problem applies only until the resin is cured.

Close-Fit Pipe

The use of liners that are deliberately deformed prior to insertion, and then revert to their original shape once

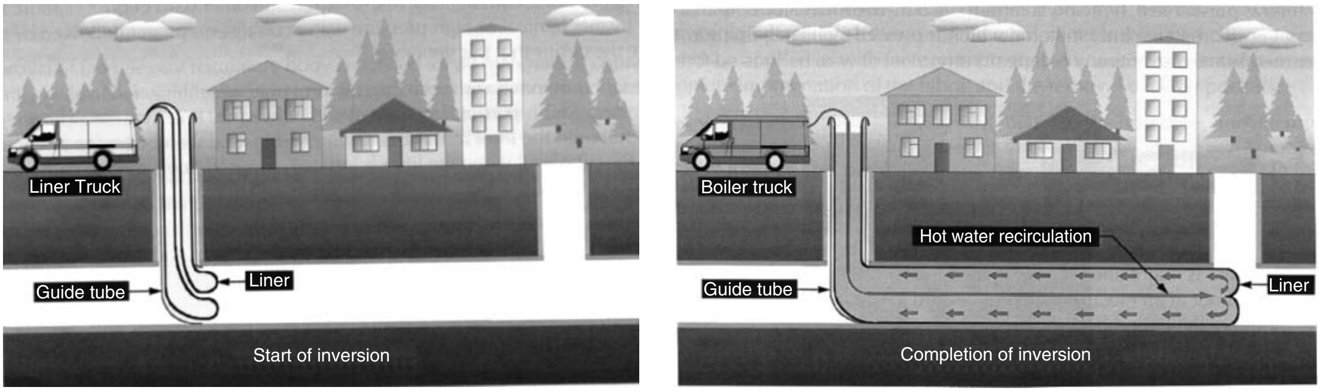

Figure 2. Cured-in-place pipe (CIPP) lining.

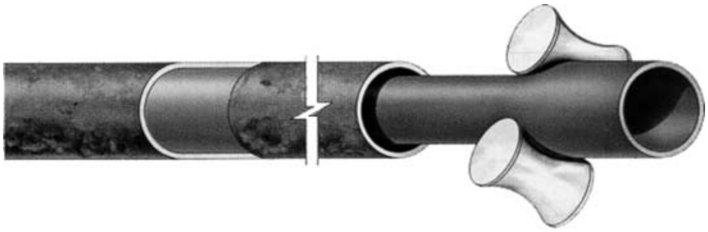

The pipe is reduced in diameter by passing it through rollers or dies, and is reverted to its original size after insertion

Figure 3. Swaged liner. Diagram courtesy of Subterra.

in position, so that they fit closely inside the host pipe, is called "close-fit lining." Such techniques are a logical development of basic sliplining described previously. The close-fit pipe method takes two principal forms, both of which take advantage of the built-in memory of some polymeric pipe materials. In one method, the diameter of the liner is temporarily reduced by drawing it through a set of rollers, a process referred to as "swaging," so that it can be inserted into the existing pipe (Fig. 3). Once installed in place, the liner is restored to its original size by pressurizing the line. Due to the limitations on size reduction, this technique is better suited to pressure pipes than to gravity sewers that have dimensional irregularities. The second method involves folding the liner into a "U" or "C" shape prior to insertion (Fig. 4), and then using heat and/or pressure to restore its original shape. This technique is often described as "fold-and-form." The liner diameter and wall thickness are the main limitations of this process, but it can be used for both pressure pipes and gravity sewers. Some versions of the spirally wound lining technique described previously offer the facility to expand the liner to provide a close fit with the host pipe. During installation, the joint between the adjacent turns of the helix is prevented from slipping by a locking wire. Once the liner is in position, the locking wire is pulled back progressively to allow the joint to slip and the helix to increase in diameter (Fig. 5).

Close-fit lining minimizes the need for cross-sectional reduction. Theoretically speaking, in this type of method, there is no annulus and therefore no need for annulus grouting. The curing time is reduced compared with a cured-in-place (CIPP) liner. As in CIPP, the lateral connections can be reopened remotely after lining by using a remote controlled robotic cutter. The chief disadvantage is the size limitation (for non-man-entry pipe sizes).

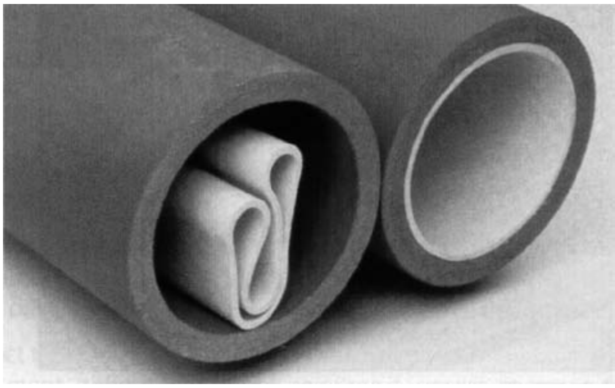

Folded PVC liner for sewer renovation, showing close fit after reversion

Figure 4. Fold-and-form pipe.

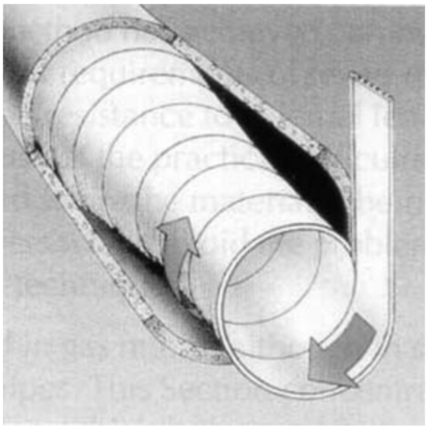

The spirally wound liner is expanded after insertion by allowing slippage between adjacent turns of the helix

Figure 5. Spirally wound liner being installed in a host pipe.

Spray-On Linings

Spray-On linings have a long history in potable water pipelines, where the principal objective is to remove the scale and corrosion of old pipelines, generally made of cast iron, and then apply a coating that inhibits further deterioration and seals minor leaks and cracks. The most common materials used for this purpose are *cement mortar linings* or *epoxy resin linings*. For small-to medium sized pipelines this can be achieved by a robotic spraying machine that is winched through the pipeline at a constant, predetermined rate. Development continues in achieving a cost-effective application of spray-on linings in non-man entry sewer pipeline renovation, but so far no method has commercial prominence. This may be partly due to the differing requirements for sewer renovation, where the aim is to increase the structural strength of the pipe, rather than prevent corrosion, and partly due to the practical difficulty of ensuring that the inflow into the system is completely stopped while the material is being applied and cured. For large, man-entry pipelines, the application of a mortar lining can be done manually or mechanically.

Spray-on linings are usually intended as protective coatings and rely on the bond to the host pipe for their function. Hence the preparation of the host pipe is critically important. Old water mains, particularly those made of cast iron, often have heavy internal deposits of corrosion and scale. Cleaning techniques include high-pressure water jetting, scraping, pigging, etc. There is often a balance drawn between removing all traces of corrosion and avoiding damage to the pipe wall; some of the more aggressive cleaning techniques should be used with caution.

Pipe scrapers are designed to remove hard deposits. Spring-loaded steel blades are mounted on the central shaft of a scraper. A towing eye is fitted to each end allowing the scraper to be pulled through the pipe in either direction. Wire brush pigs consist of circular wire brushes mounted on a central shaft and are used to remove the debris loosened by the scrapers or by themselves to remove loose debris from the pipe. Cleaning pigs are available in a wide range of designs and are normally driven through the pipe under water pressure. In a heavily encrusted pipe, pigging may be carried out in stages, using pig sizes of increasing diameters. Pull-through pigs, also known as "squeezes," remove fine material and fluids from pipes. They consist of rubber disks fitted to the central shaft that are towed through the pipe. Pull-throughs are often used in the final stage of pipe preparation to produce a clean, dry surface, to which the spray-on lining is applied.

The application of a *cement mortar lining* is a common and a relatively inexpensive method for water main renovation. The cement mortar serves two important functions: the alkalinity of the cement inhibits corrosion of an iron pipe, and the relative smoothness improves flow characteristics. Note that cement mortar linings are applied to many new cast iron and ductile iron pipes to inhibit corrosion. Application is generally by a spraying machine that is either fed by a hose from the surface or, for large pipes, may have its own hopper containing premixed mortar. The forward speed of the sprayer is controlled to achieve a consistent mortar thickness. Spray application may be followed by troweling. This is often carried out by rotating spatulas fitted to the spraying machine.

For renovation of large diameter brick and concrete pipelines, *shotcrete* is often used as a cost-effective means to enhance the structural integrity of a pipeline. Shotcrete is a process in which mortar or concrete is conveyed through a hose and projected at high velocity onto the surface of a pipe. Shotcrete may be applied dry or wet. The method of applying dry-mix shotcrete through a hose using compressed air is also given the trade name, *Gunite*. There are advantages and disadvantages to both methods. For instance, a wet-mix shotcrete application results in less rebound, less dust, and yields a higher compressive strength. The drawbacks of the wet-mix are limited conveying distance, increased demand on aggregate quality, and increased clean-up costs. As in other lining methods, surface preparation is critical. All deteriorated material from the pipe walls must be removed along with dirt, oil, and other bond inhibiting material. The pipe surface must be dampened in the wet method so as to be saturated surface dry (SSD) prior to application. Prepackaged mixtures of ready-to-use cementitious mortars containing various proportions of silica fume and fiber reinforcement, and admixtures for high-strength, low-shrinkage shotcrete are commercially available. For additional strength, reinforcing bars or welded wire fabrics may be used. Both wet and dry-shotcrete are shot perpendicularly to the pipeline surface to minimize rebound and to create a smooth pattern. After applying the shotcrete, it is allowed to stiffen for about 10 minutes before "bumpy" areas are troweled to give a uniform surface. Shotcrete is applied in layers (approximately 1-inch thick) until the required design thickness is reached. Before applying the next layer, the previous layer must be allowed to reach an initial set. This can take from 45 minutes to several hours depending on mix consistency and ambient temperature.

Epoxy linings are an alternative to cement mortar linings (Fig. 6). The resin bonds to the surface of the pipe, forming a coating, which inhibits water penetration and corrosion. Epoxy coatings are much thinner than cement mortar linings and therefore do not significantly reduce the pipe cross section. However, any defect in the coating may allow corrosion to start and unlike cement mortars, they do not have alkalinity to inhibit chemical deterioration. Epoxy resins are also relatively expensive compared to cementitious materials. Epoxy resins should not be used for lining water lines unless specifically approved by NSF. The resin is applied by a spraying machine with a rotating nozzle. The thickness of the coating is controlled by the flow rate and forward speed of the sprayer. In most systems, the resin base and the hardening agent are fed through separate hoses and are combined by a static mixer just behind the spray nozzle. Ideally, the cure time should be as short as possible to minimize the time during which the pipe is kept out of service. However, too rapid a cure carries the risk of causing blockage in the static mixer or the nozzle. Unlike cement mortar linings, the resin is not smoothed or troweled after spraying, and the

Figure 6. Application of epoxy resin lining.

surface quality depends on the application technique and the properties of the material.

Epoxy linings do not fulfill any structural function and are not recommended for pipes that leak significantly or for pipes with significant structural defects.

Point Source Repairs

Various localized repair systems have been developed, most aimed at sewer renovation but include some that are designed to seal joints in pressure pipes. Many techniques are adaptations of the full-length lining systems seen previously.

The economics of "spot repair" versus renovation of the entire pipe length is assessed by the Design Engineer, but as a general rule spot repair methods may prove economical where defects in the pipe are restricted to less than 25% of the length. Preinspection of the pipeline and cleaning are just as essential to localized repairs as to full-length renovation. There are four broad categories of localized repair systems: robotic, grouting, mechanical seals, and point CIPP. Sometimes point repair systems are classified into two categories, structural and nonstructural. Robotic methods, mechanical seals, and point CIPP produce increased structural strength and are hence regarded as "structural" renovation techniques; chemical grouting is often termed "nonstructural."

Point Source Repairs. Robotic techniques, used mainly in gravity pipelines, consist of a grinding robot and a sealing robot (Fig. 7). The grinding robot can be fitted with various shapes of diamond carbide cutters suitable for clayware, concrete, polymeric pipe, and even steel. The wheels are driven by an electric motor and cooled by a water jet spray. The robotic operation is monitored by a CCTV camera attached to the head. Typically, the grinding robot mills cracks to a width and depth of up to 1 inch, after which the area around the crack is cleaned. A grinding robot can also remove intruding laterals, grout deposits, and hard encrustation. Like the grinding robot, the sealing robot is self-propelled and equipped with an onboard CCTV camera. Epoxy is applied by the sealing robot to completely fill the slot milled by the grinding robot. In addition to filling slots, the sealing robot can apply epoxy around poorly made connections or leaking joints to seal connection and joints.

All robotic functions are controlled from a central console housed in a vehicle, which contains the power pack, hoist for lifting and lowering robots into manholes, and other ancillary equipment. Robotic repairs are versatile, but to be cost-effective, a consistent program of work is needed.

Grouting. Resin injection systems fall into two categories: those whose principal function is to seal the pipeline against infiltration and exfiltration and those that restore structurally damaged pipe.

A common method of sealing leaking joints in gravity pipelines is by using a special packer that combines the functions of leakage testing and grout injection. A packer with inflatable elements is positioned across a pipe joint and pressurized to isolate the joint. The joint is tested for pressure loss and if necessary a sealing gel is injected through the packer to seal the joint. The grout has little intrinsic strength but turns the ground around the leaky joint into an impermeable mass thereby enhancing the structural stability of the pipe.

In the past, the use of acrylamide grouts was commonplace. However, these grouts have been banned in the United States because the unreacted components pose a serious health risk. Today the use of polyurethane (PU) grouts is common. PU grouts react with the free water in the soil to form a viscous gel that hardens over time. Many PU grouts contain acetone to reduce viscosity. The flammability of acetone should be considered during storage and handling.

For leak sealing the entire system—main sewer, laterals, and manholes—a different approach is taken, often referred to as "fill and drain." The section of the sewer to be sealed is first isolated and then filled from manhole to manhole with an environmentally safe chemical solution (usually *sodium silicate*). After a predetermined interval to allow the chemical to permeate through leaking joints and cracks, the solution is pumped out quickly, and the section filled once again with a second chemical solution

Photo courtesy of KA-TE system AG

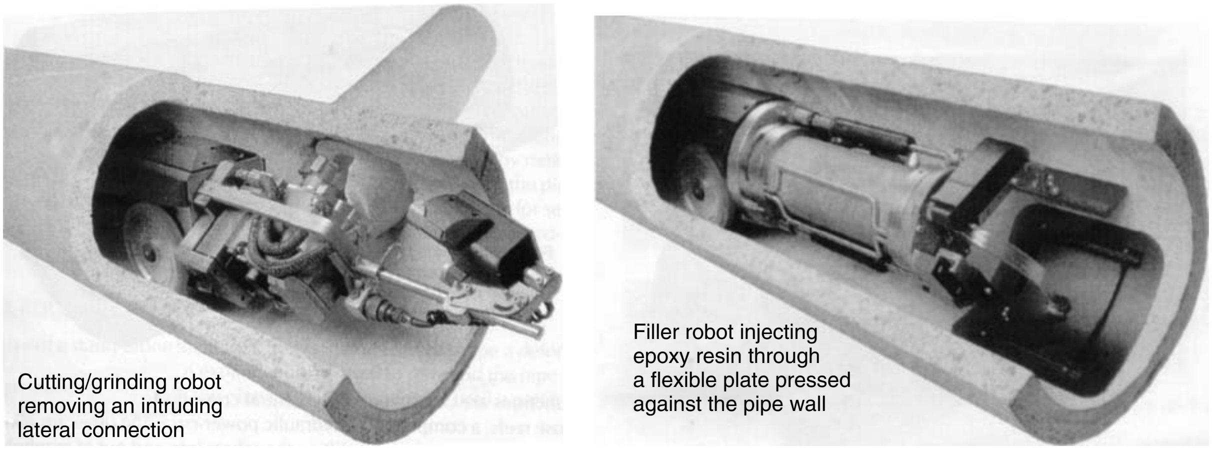

Figure 7. Robotic repairs using grinding and filling remotely controlled robots.

(usually *proprietary*), which reacts with the residue of the first chemical to form a waterproof gel. The second chemical is then pumped out and the pipe cleaned before returning it to service.

An epoxy resin system is used to address structural defects in pipes. An inflatable packer isolates the defect and injects a rapid-setting epoxy resin into the crack, fracture, or hole in the pipe wall. The packer is left in position until the resin has cured, thus forming an internal collar inside the pipe.

Mechanical Seals. Mechanical seals are used primarily for sealing leaking joints. This method involves installing a metal band or clip (usually stainless steel) faced with an elastomeric material across the joint. The elastomeric material forms a seal with the inner surface of the pipe. Repair modules are installed by an inflatable packer which expands the clip and presses the elastomeric material against the pipe wall. The packer is then deflated and withdrawn. Systems of this type are available for gravity and pressure pipe. Mechanical sealing systems have the advantage of being quick and easy to install; however, the material cost is higher than in other methods. Tapered versions are available to seal between pipes of varying diameter and to seal the annulus at the ends of sliplined pipes.

Point CIPP. Most point or "patch" repairs with cured-in-place (CIPP) liner sections entail impregnating a fabric with suitable resin, pulling the patch into place within a sewer by using an inflatable packer or mandrel, and then filling the packer with water, steam, or air under pressure to press the patch against the existing sewer wall while the resin cures. After curing, the packer is deflated and removed. The repair is then inspected by CCTV, and any lateral connections present are reopened using the same techniques as for full-length CIPP liners.

Both thermal and ambient cure systems are currently available. In many respects, point repairs are a short version of cured-in-place liners, although often the fabrics and resins are specialized to achieve greater structural strengths in the patch. The fabric is commonly polyester needle-felt (unwoven), either on its own or in combination with glass fiber. Some systems use a multilayer sandwich; the glass fiber provides strength, and the felt acts as a resin carrier. Although polyester resin is used in full-length liners for economy, epoxy resins are a common alternative for local repair systems. Epoxies, unlike polyester resins, are not affected by the presence of water. This is especially relevant in techniques designed for installation without diverting the flow in the pipeline. Impregnation of the fabric is usually, but not always, carried out on site. In this case, care is needed to avoid health risks and the spillage of chemicals, some of which are toxic in the unreacted stage. During mixing and impregnation of the resin, it is important to remove all entrapped air because air can weaken the material and in severe instances results in porosity. To overcome this, some systems use vacuum impregnation.

In both ambient and thermally cured systems, it is essential to limit the rise in the temperature of the material until the patch is inflated within the host pipe. Exothermic cure begins as soon as the resin components are mixed, and the rate of temperature rise depends on the volume of resin mixed. One of the most common causes of point system failure is premature cure, where the patch has started to harden before it is in position.

BASICS OF MANHOLE RENOVATION USING TRENCHLESS TECHNIQUES

Trenchless manhole renewal methods offer several advantages over conventional dig and repair/replace methods. These advantages are similar to those in trenchless pipeline renovation, explained previously.

There are four basic methods of manhole rehabilitation:

- spot repairs (chemical grouting, plugging, and patching);
- cementitious coatings (spray and troweled);
- protective coatings (epoxy and polyurea); and
- structural repair systems (monolithic *in situ* liners)

Spot Repairs. Most injection processes use materials similar to those employed in pipe joint sealing, polyurethane or acrylic grouts. The primary purpose of pressure grouting is for I/I control and stabilization of voids around manholes.

The process generally involves drilling injection holes in a predetermined pattern, through which the grout is injected to create a flexible external membrane. Polyurethene grouts are water-reactive and can be injected without a catalyst if there is sufficient free water in the surrounding ground.

Rapid setting hydraulic cements may be used for plugging holes and cracks and are generally used prior to grout injection to prevent washout of the grout. Some hydraulic cements have a setting time of just a few seconds.

Careful inspection of the contractor's work during the actual grouting operation is necessary to ensure adequate coating of the exterior of the manhole structure. A follow-up dye test is recommended for each manhole.

Cementitious Coatings. Cementitious coatings are typically used to inhibit I/I and are not structural. Cementitious coatings are typically made of portland cement, finely graded mineral fillers, and proprietary chemical additives. Cementitious coatings are generally brushed on in two applications, one gray and one white, to provide a visual indication of the coverage of each layer, to a total thickness of about 1/2 inch. Before applying the cementitious coatings, it is essential to prepare the interior by washing with high pressure jets and plugging and patching as required.

Protective Coatings. Protective chemical coatings, such as epoxies and polyureas, are used as barriers to bacterial corrosion on new or recently installed precast manholes. They are not used in rehabilitating badly deteriorated or structurally unsound manholes. The recommended coating thickness differs greatly among products and manufacturers, although the typical range is from 60 to 200 mils. A combination of semistructural cementitious liners with epoxy coatings has recently been introduced to provide surface restoration and protection.

Structural Repair Systems. Cementitious liners have been widely used over the past 5 years for structural enhancement of brick manholes which have lost some of their mortar joints but remain basically sound. The liners are made of high-strength, quick-setting mortars of portland or high alumina cements. Thorough preparation of the internal surface is essential, and a bonding or stabilizing agent may be necessary depending on the condition of the manhole. Structural rehabilitation generally requires a cementitious liner thickness of between 1 to 4 inches (Fig. 8).

In situ structural replacement systems fall into two categories: inserts and *in situ* lining systems. A variety of inserts made of fiberglass, polyurethane, PVC, etc., can be installed inside existing manholes and then back-grouted to create a new structure inside the host manhole. The concept is similar to sliplining in pipeline renovation. This method requires excavating the ring and the cone to allow placing the insert inside. If the manhole is tapered, it may be necessary to excavate several feet before the insert can be positioned. Careful attention is required around pipe inlets because it is difficult to make a seal when the material composition of the liner is different from that of the inlet pipe.

In situ lining systems are reconstruction techniques that require no excavation and are not disruptive. These systems use a variety of spiral-wound, cured-in-place, etc., materials to create a liners inside existing deteriorated manholes. Cured-in-place liners are generally recommended for structural rehabilitation of manholes where future corrosion is expected. A felt or a glass-reinforced “bag” is custom-made to the size of the manhole. Resin is impregnated into the bag at the job site. The bag is then lowered into the manhole, steam pressure injected for 1–2 hours to achieve curing, and holes are cut at the invert and pipe inlet locations before returning the manhole to service. Surface preparation is crucial to the success of this method. All voids and recesses need to be filled prior to undertaking cure-in-place lining. The helically wound lining technique described previously for pipeline renewal can also be used for renovating manholes. PVC strips are fed into the manhole opening and spirally wound by a winding machine into the existing manhole. Preparation involves removing ladders, step-irons, and internal backdrops before cleaning the surface by high pressure water jetting. Annulus grouting is carried out to bond the liner to the walls of the manhole.

Figure 8. Application of one-coat cementitious structural liner.

BASICS OF UNDERGROUND WATER AND SEWER PIPELINE ASSESSMENT, REPAIR, AND REHABILITATION

Sanjiv Gokhale
Vanderbilt University
Nashville, Tennessee

In March 2001, ASCE released a Report Card for America’s Infrastructure. The wastewater and drinking water infrastructure categories received a discouraging D+ overall and an estimated need for an investment of $1.3 trillion to bring conditions to acceptable levels. In September 2003, ASCE released a Progress Report (http://www.asce.org/reportcard/) that examines the current trends for addressing the nation’s deteriorating infrastructure. ASCE determined that the conditions and performance had not changed significantly in 2 years.

America’s industries rely on clean water to carry out activities that contribute over $300 billion to our economy each year. However, the challenge to continue providing clean water remains because our existing national wastewater infrastructure is aging, deteriorating, and in need of repair, replacement, and upgrading. In fact, EPA has reported that without improvements to the nation’s wastewater treatment infrastructure, we face the very real risk of losing the environmental gains we have achieved during the last three decades since the passage of the Clean Water Act of 1972. Drinking water quality remains good, but the infrastructure of the nation’s 54,000 drinking water systems is aging rapidly. Federal funding remains flat, and the infrastructure needs continue to increase. There is an annual shortfall of $11 billion needed to replace or rehabilitate facilities that are nearing the end of their useful lives and to comply with federal water regulations.

The nation’s 16,000 wastewater systems face enormous needs. Some sewer systems are 100 years old and many treatment facilities are past their recommended life expectancies. For the fiscal year 2001, Congress allocated $1.35 billion for wastewater infrastructure, but currently there remains a $12 billion annual shortfall in funding for infrastructure needs. Sewer overflows and discharge of untreated or undertreated flows into streams and rivers has made more than one-third of U.S. surface waters fail to meet water quality standards.

The forecast for our nation’s drinking water systems indicates a downward slope. Drinking water received a D on the 2001 Report Card, yet the situation continues to worsen as aging systems—some developed more than a century ago—continue to service our ever-growing population.

Traditionally, municipalities have constantly been a "step behind" in managing the municipal, and this is especially true of sewer systems. Municipalities have addressed the design, construction, maintenance, and rehabilitation of sewer systems by a "crisis-based" approach. Sewers are not built and/or rehabilitated until a major failure (i.e., overflow, collapse) occurs. This reactive management practice has led to inefficient use of funds and causes more frequent sewer failures, which end in difficult and costly repairs. For municipalities to be a "step ahead" in managing sewer systems, they need to have a better understanding of the existing condition of the network, which in turn will help them plan and manage maintenance to prevent major failures.

PIPELINE ASSESSMENT

A variety of ways and means exist by which existing underground pipelines may be evaluated:

- ground penetrating radar (GPR)
- sonar
- televising
- smoke testing
- dye testing
- flow monitoring
- walk-through
- pressure leak tests
- electronic leak testing

Ground Penetrating Radar

Ground penetrating radar (GPR) works by launching pulses of electromagnetic energy into the soil under investigation. These pulses propagate through the ground and are reflected by subsurface targets or at interfaces between different media. The radar measures the time taken for a pulse to travel to and from the target, which indicates its depth and location. The transit time of the pulse depends on the dielectric properties of the material. In addition, the depth and penetration obtained may vary considerably with soil conditions; for example, loam is an easy medium to penetrate, whereas clay is more difficult.

The difficulty with GPR lies in selecting the correct frequency; different types of soil respond differently to radar frequency. A large study undertaken in Canada to detect leaks in buried plastic water distribution pipes was inconclusive. However, more recent advances in GPR technology and in particular the use of multisensory systems seem to hold promise. Field trials indicate that multisensory systems are more accurate, faster, and capable of working in a variety of ground conditions. At present, however, the effectiveness of this technology in pinpointing inactive leaks in underground pipes is undocumented.

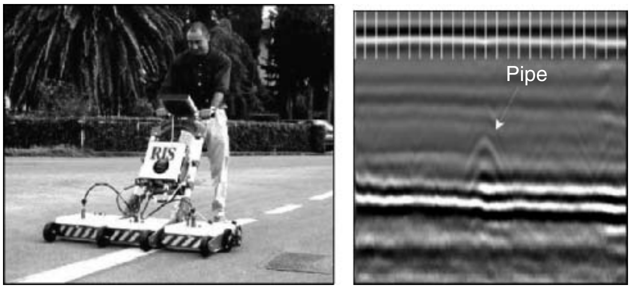

Figure 1. Multisensory ground penetration radar (GPR) used in locating underground pipe.

GPR is used today primarily to locate underground pipelines. GPR equipment is dragged across the surface in a precise fashion, and the output (radar signal) is stored and later graphed after filtering out the "noise" (Fig. 1). Computer enhancements have allowed GPR technology to become very effective in situations where an underground discontinuity (e.g., a pipeline) is known, but the plan location of a buried pipeline is unknown. However, GPR may be able to indicate that a discontinuity exists underground, but there is often an uncertainty about what the discontinuity actually is and also what is the exact depth at which this discontinuity exists.

GPR technology is an evolving field and it is expected that in the future this may very well become the most dominant method of pipeline location and assessment.

Sonar

Sonar can give a relatively accurate picture of the profile of a pipe wall and surrounding soil. The results, however, are very much open to interpretation, and a highly skilled operator is vital. Sonar techniques can, however, show flow regimes under water and are thus useful in detecting infiltration into surcharged pipes.

Televising

Pipeline televising is used to identify problems primarily in sanitary and storm sewer pipe. It is usually referred to as closed-circuit television inspection (CCTV) and consists of a remote television camera passing through a pipe that sends the live picture to a screen aboveground. The remote television camera unit is composed of a camera mounted on a skid or equipped with wheels (Fig. 2). The camera can be directed through the pipe using manholes as access points to insert cables that connect to the remote TV unit. Other cables connected to the remote TV unit carry the signal back to the screen, which is usually mounted in a truck. The screen mount in the truck incorporates a videocassette recorder slot so that data can be saved and viewed by others.

The sizes and types of pipe that can be televised are virtually unlimited, especially with newer technologies that allow smaller TV units to be pushed through small diameter pipe such as house laterals. The line televising crew usually consists of an operator stationed in the truck and one or two operators stationed near the manhole(s).

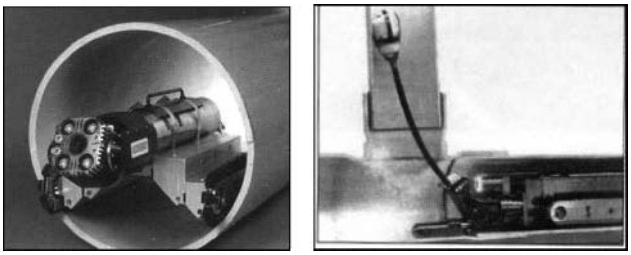

Figure 2. CCTV Systems: Mainline inspection and lateral inspection (left to right).

In most cases, the line televising crew is employed by the local utility, although private contractors perform this service as well.

Line televising is performed as part of routine operation and maintenance by most wastewater utilities. It provides a good indication of the structural integrity of the pipe and can pinpoint problems in a way that was not possible before its invention (Fig. 3).

A utility may also require postconstruction televising at the completion of a new sewer or storm pipe to determine adequacy of joint connections, service connections, or other indications of proper pipe installation. A further application might be the review of lined sewer pipe to ensure that services were properly reinstated. The applications of line televising are numerous. The information generated, in conjunction with proper consultation with the Engineer, will guide the utility toward better decisions regarding the maintenance and operation of the network.

Smoke Testing

Smoke testing is a simple test used to identify unauthorized connections to a sanitary sewer system. In systems where the sanitary and storm sewers are separate, unauthorized connections include roof leaders, downspouts, and yard drains. Smoke testing can also identify direct connections between the sanitary and storm sewer system and, to a lesser extent, locations of broken sewers or “lost” manholes. In these latter two cases, line televising is more effective.

To smoke test, a blower unit is inserted over a manhole to allow smoke to be blown into the sewer (Fig. 4). Pipe plugs are inserted into adjacent manholes to isolate a section of sewer before “smoke bombs” are inserted into the manhole where the blower is located. The smoke blown into the sewer then, in theory, has no place to escape except through unauthorized connections, direct connections from the storm sewer, plumbing vents, or defects in the sewer system. However, there is a potential that smoke can enter homes or businesses if the plumbing that connects the building to the sewer is defective or was installed improperly or if the traps are dry. For example, the absence of a vent pipe in a home may cause smoke to enter the house. For these reasons, advance warning to the residents potentially impacted is not only a good idea, but also a necessity. Further, coordination with the local Police and Fire Departments is necessary to avoid panic or alarm with the potential of smoke detected in resident’s homes. Smoke testing can be a valuable tool in determining who is connected to the sanitary, or storm, sewer system.

Dye Testing

Dye testing, like smoke testing, can be used to determine connections to a sanitary, or storm, sewer system or interconnections between the two. It can also be used to determine leaking sewers that may be discharging to creeks or other bodies of water. Improper connections, such as yard drains discharging to a sanitary sewer, can be effectively identified by a dye test. It can also be used to identify positively whether a house is actually connected

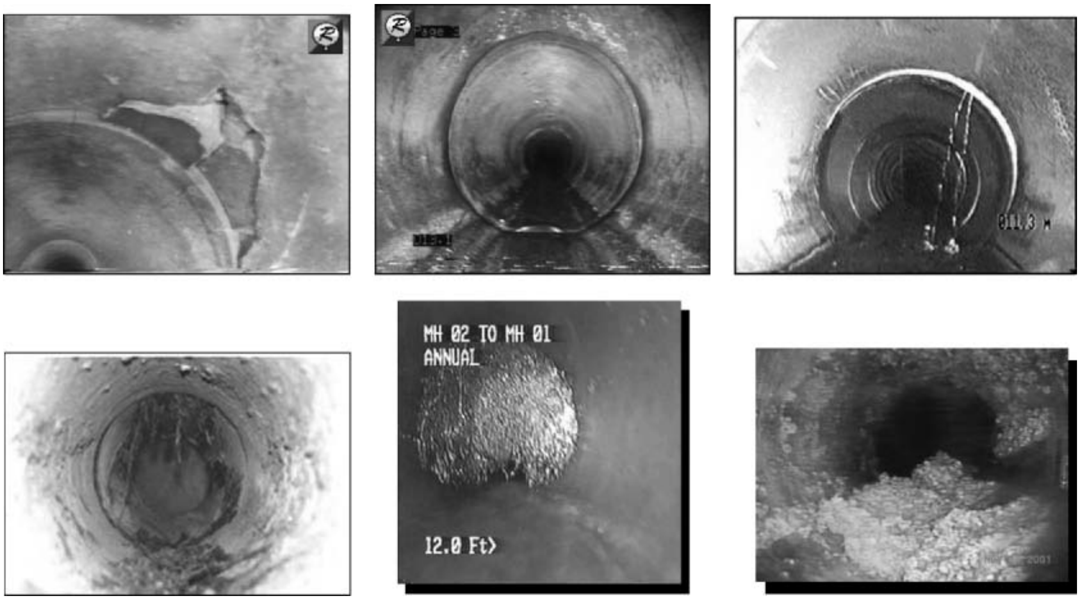

Figure 3. Pipeline Defects: Cracked pipe, offset joint, infiltration root intrusion, root blockage, and grease blockage (clockwise from top left).

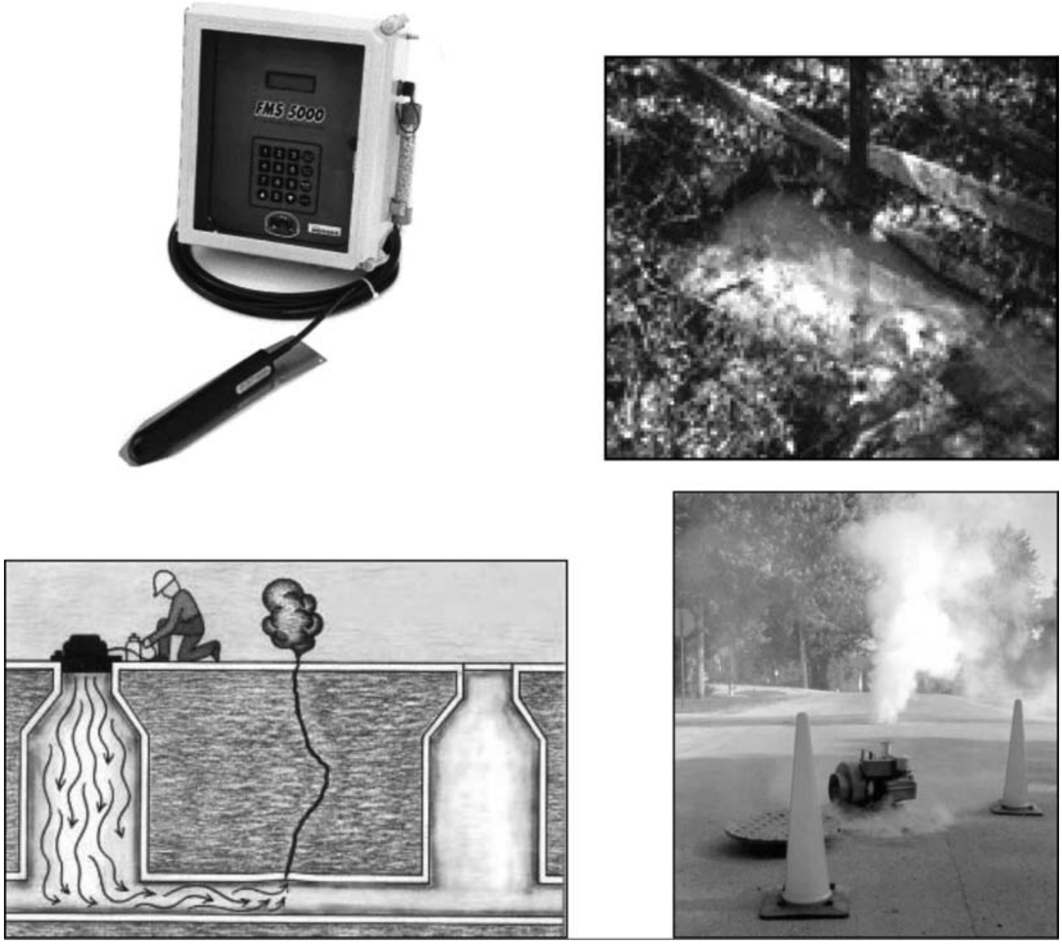

Figure 4. Flow monitoring device, dye testing, and smoke testing in sewers (clockwise from top left).

to the sanitary sewer. Rare instances where house laterals were either inadvertently not connected or wrongly connected to the storm sewer system, unfortunately, sometimes occur. A dye test is a good tool for these cases.

The dye can be in powder, tablet, or liquid form and is simply inserted into the sewer (Fig. 4). Although safe and harmless, the wastewater utility should be contacted to ensure that no regulations against the dye are in place and as a “heads-up” if appreciable quantities of the dye reach the treatment plant. Similarly, the local environmental agency should be contacted if dye testing is proposed in storm sewers. Some dyes can be mistaken for things such as “antifreeze” spills and can cause undue alarm and attention. Like smoke testing, public education is a necessity in most applications. Where access to households is necessary to test for positive connections of laterals or drains, it is probably best to work through local city or town officials.

Flow Monitoring

Flow monitors (or meters) are useful as part of a statistical modeling exercise for I&I (Fig. 4). Data on water use, and hence discharge to sewers, can be obtained through a few strategically placed flowmeters and when combined with network modeling, can present a very accurate overall picture of the system operation.

Walk-Through

The walk-through method of pipe assessment is obviously used only for large diameter pipes or large storm sewer box structures. This method can yield important information on the structural integrity of the pipe and structure, especially if trained personnel perform the inspection. The Construction Inspector must be fully aware of the safety implications of entering these pipes or structures. The decision on whether to enter a pipe or structure should be made with full regard for Occupational Safety and Health Administration (OSHA) confined space entry regulations http://www.osha.gov/.

Pressure Leak Testing

For pressure pipe, the methods of assessment for pipe in service are limited. One method used is leak testing using sound waves to pinpoint the location of the leak. The test is primarily on water mains and is usually performed by specialized firms. However, the cost of leak detection equipment has decreased to the point where most medium to large water utilities have invested in this methodology.

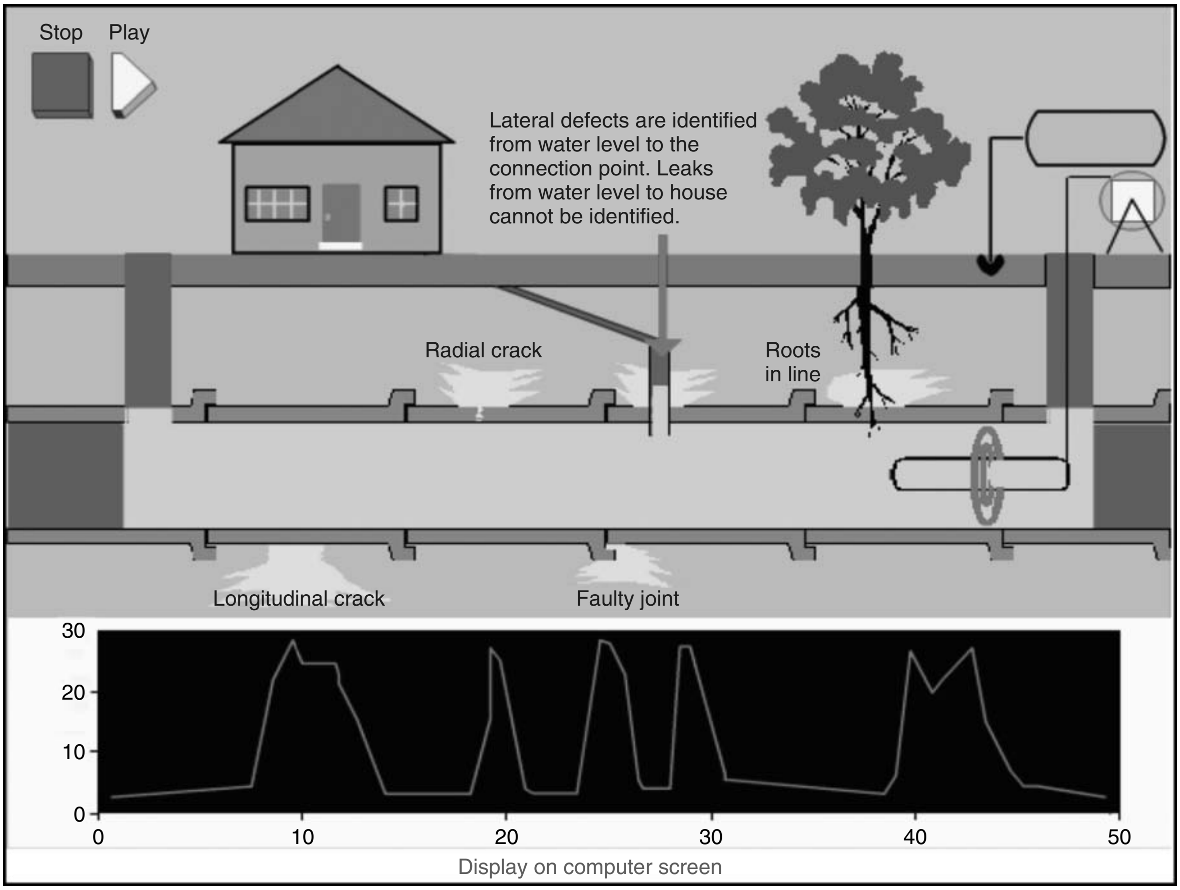

Figure 5. The FELL system.

Leak testing uses sound waves to identify the point in the pipe where a leak has occurred. The technology has evolved to a point where digital readouts can specify the exact location of the leak relative to the wire connections on water main appurtenances, usually valves. This technology could be useful when recently installed or repaired pipe cannot pass specified pressure tests.

Electronic Leak Testing

The focused electrode leak location (FELL) system measures electrical current flow between a probe that travels in a pipe and a surface electrode. Pipe defects that allow liquids to flow into or out of the pipe cause a spike in the electrical signal, thereby locating the sources of infiltration or exfiltration. The intensity of the measured current can be correlated with the magnitude of the leaks.

The FELL test system uses a specially constructed electrode called a “sonde” that generates an electric field. The sonde is 30 in. long and 3 in. in diameter. The electric field is focused into a narrow disc, 1 in. wide set normal to the longitudinal axis of the sonde. A surface electrode (usually a metal stake) is put into the ground at the surface. When the sonde is placed in a *nonconducting* pipe that contains sewage (and/or water), the electric current flow between the sonde and the surface electrode is very small. Defects in the pipe that would allow flow of fluid either *into* or *out of* the pipe provide an electrical pathway from the sonde, through the wall of the nonconducting pipe, and through the ground, to the surface electrode. When the sonde is close to such a defect, the current between the sonde and the surface electrode increases. Figure 5 illustrates the operation of the FELL system.

PIPELINE REPAIR AND RENEWAL—EXCAVATION

Sanitary and storm sewer systems and water distribution systems develop defects over time. Defects occur from a variety of causes including improper initial installation, freeze/thaw damage, and cutting by another contractor. Pressure lines typically exhibit failures rapidly; it may take years for a defect in a gravity line to be obvious. A failing pipe system can be repaired or renewed by replacing the entire pipe or manhole or by replacing only the defective portion. Construction requirements for repair and renewal are similar to original construction requirements. The following discussion highlights the different methods of repairing or renewing pipes and manholes.

PIPELINE REPAIR AND RENEWAL—WHOLE PIPE

Gravity Sewers

When the Engineer has determined that a pipe and/or manhole has deteriorated beyond trenchless rehabilitation, replacement is typically the prudent alternative. To replace a whole segment (structure to structure), the Contractor must excavate to the pipe. Paved surfaces must be saw cut prior to excavation to prevent damage beyond the construction limits. The original backfill is removed; if it is acceptable material by today's standards, it can be reused; otherwise, it must be removed. When the pipe is exposed, it is removed along with the bedding material. Lateral connections should be cut before the pipe is removed to prevent fracturing the lateral. Once enough of the pipe is removed, new segments of pipe may be installed following procedures for new pipe installation. Bedding, haunching, and backfilling are all important components of successful pipe replacement. Although it is often more difficult to prepare the bedding below the pipe because water may be flowing in the trench, proper bedding is essential for proper pipe installation. In addition, if the original pipe settled, the bedding remaining under the existing pipe will not be on grade for the installation of the new pipe. It is important to check the grade of the bedding prior to laying the new pipe.

Force Mains

If the line to be replaced is a force main, it may be designed in a new location so the existing force main can remain in service until the latest possible time. In this case, the force main installation is the same as new construction until it is time to make the final connection to the pump end of the force main. If the new force main connects to the existing force main, a pressure coupling must be used. If the new force main is connected at the discharge of the pumps, properly sized spool pieces must be on hand to complete the final connections as rapidly as possible.

Couplings to join the new force main to the old force main must be designed for the pressure the line will undergo. Pressure couplings typically have a rubber sealing sleeve surrounded by a steel or stainless steel sleeve. The sleeve is fastened to the pipe by tightening bolts provided on the clamp.

All testing of the force main must be completed prior to putting it into operation. The same tests required for new force mains must be completed on replacement force mains.

The Contractor may have to provide external pumping capacity or temporary piping to allow pumping the wastewater or storm water downstream. This can also be accomplished by a vacuum truck, with appropriate permissions to discharge into a downstream manhole.

Manhole Connections

The rubber gasket or boot used to create a watertight seal between the manhole and the pipe may need to be replaced. The importance of the gasket, whatever type is used, is to prevent groundwater from entering the pipe–manhole joint. The gasket must seal tightly to the manhole and to the pipe.

Lateral Connections

The Engineer will specify if the laterals are to be replaced to the right-of-way or easement limits or only within the trench. In either case, the lateral wye or tee must be installed in the main line approximately at the same location as the previous one. The Contractor must excavate along the lateral until sound (not broken or cracked) pipe is found and replace the lateral from that point to the main. The connection between the new pipe and old pipe is made with a special coupling. These couplings are manufactured to fit various pipe sizes and materials. They are manufactured from rubber compounds and have some type of clamp to seal the coupling to the pipe. Some of the couplings have a stainless steel sleeve to prevent misalignment of the pipe.

Testing

Testing requirements are specified in the Contract Documents. Rehabilitated pipe may be tested similarly to a new pipe installation with a low-pressure air test and a deflection test. If the main has lateral connections, the lateral connection coupling must be removed and the lateral plugged prior to the air test. To help reduce the cost of testing, the pipe can be air tested (if permitted by Contract Documents) once the pipe has been bedded and backfilled up to the elevation of the lateral couplings. Deflection testing should be completed after the required waiting time. In special circumstances, the Engineer may specify joint testing in lieu of air testing of the whole pipe. The Engineer may also specify deflection and/or television inspection of the main after the repair has been completed. Television inspection after installation is similar to the inspection described in assessment of pipes earlier in this article.

Sewage Control

The Contractor may choose (or the Engineer may require) to plug the upstream sewer line and pump the flow around the segment being replaced. This creates easier working conditions for preparing the bedding for the new pipe. As work progresses on the line segment, the Contractor may have to cope with flow from laterals. If the flow is light, the Contractor may choose to "just work in it" and let the sewage flow through the trench. This is acceptable as long as the crew is able to compact the bedding and backfill properly. If the Contractor does not bypass the wastewater, workers need to be aware of the potential for disease transmission. Rubber boots, gloves, and eye protection should be provided. Facilities must be provided for washing hands.

PIPELINE REPAIR AND RENEWAL—POINT REPAIRS FOR GRAVITY SEWERS

Gravity Sewers

If a sewer or force main has a small section that is deteriorated, it may be replaced while leaving the remaining line in service. The Contractor saw cuts pavement as necessary and excavates down to the pipe. The bad section of pipe is exposed. The pipe must be cut with a saw past the limits of the damage. It is important that the pipe is cut back to sound pipe for the repair coupling to seal against. A new section of pipe is cut to the same length as the piece removed. Repair couplings are placed on the replacement piece and the new piece is inserted into place.

Bedding

The repair piece must be bedded properly to prevent settlement or heaving of the new piece of pipe. If flexible couplings are used, this is all the more critical because the flexible couplings will deflect rapidly creating an offset in the pipe or a tear in the coupling if the offset is severe enough. If couplings with a steel sleeve are used, the offset may cause the pipe to crack if it is not bedded properly. Lean concrete or flowable fill may be used to bed the pipe.

Repair Clamps

- *Flexible Repair Coupling*: these clamps are manufactured to fit snugly on various sizes and types of pipe. Clamps can be purchased to fit VCP on one end and PVC on the other because they have different outside diameters. The proper clamp must be used to ensure a watertight connection. The rubber sleeve is tightened onto the pipe by tightening the clamps supplied with the repair coupling
- *Repair Couplings with Sleeves*: these clamps are similar to the flexible clamps described above, but they have a stainless steel (usually) sleeve that wraps around the rubber sleeve to prevent the pipes sections from being misaligned or offset.
- *PVC Repair Clamps*: these clamps can be used to repair PVC pipe with PVC pipe. The clamp is a double bell.

Testing

The most common tests for point repairs are

- *Segmental Low-Pressure Air Test*: a testing ring is inserted into the line and positioned over the repair. The positioning is observed with a television camera. The testing ring is inflated to the required test pressure.
- *Mandrel Deflection Test*: Completed in the same manner as new construction.
- *Television Inspection*: Completed in the same manner as assessment.

PIPELINE REPAIR AND RENEWAL—POINT REPAIRS FOR FORCE MAINS

Force Mains

Force mains may be point repaired by methods similar to those for gravity lines. The main difference is in the type of repair clamp and testing. If the defect in the force main is a single crack, a repair clamp of the proper pressure rating and length may be used without replacing any pipe. If the damaged section cannot be repaired by one clamp, a section of the force main may be removed by saw cutting. A new section of pipe is then secured into the line with one or two clamps (one if it is possible to use a push-on or bolt-on joint). The bedding must be moved away from the pipe to install the clamp. The pipe must be cleaned to allow the clamp to seat properly.

Bedding

The replaced section of pipe or the clamped area must be rebedded to support the pipe properly. Lean concrete or flowable fill may be used for support under the repaired area.

Repair Clamps

The surface of the existing pipe and the replacement piece must be relatively clean to allow the repair clamp to seal completely on the pipe. The clamp will leak if there is debris on the pipe.

- *Bolt-On Flanges*: for most pressure pipe, flanges can be bolted onto the existing pipe and a new flanged piece (or bolt-on flanges on the repair piece) bolted into place.
- *Pressure Couplings*: these couplings are typically made from a sheet of rubber to create the sealing surface and a stainless steel band with multiple clamps to hold and tighten the clamp onto the pipe. They are manufactured for various pressure ratings.

Testing

Typically, the repair is tested after the pipe is bedded but before backfilling is completed. The pressure is returned to the line; if no leaks are observed, backfilling is completed.

MANHOLE REPAIR AND RENEWAL

Manholes

To replace a manhole or storm structure, the Contractor must saw cut the existing pavement and excavate around the manhole. When the excavation reaches the elevation of the sewers, the excavation must follow along the pipes to allow the pipes to be saw cut. Once all lines have been freed from the manhole or structure, the manhole can be removed. The Contractor will have to provide a means of controlling wastewater flow into the excavation, as described earlier in this section.

Once the manhole has been removed, the bedding can be removed and replaced. The bedding must be compacted well, leveled, and brought up to the correct elevation. The new manhole is set similarly to new construction. Once the manhole is set, the pipes must be reconnected. New sections of pipe are inserted into the manhole, and the other end is connected to the existing pipe with a coupling as, described in the Point Repair section.

Testing

Replaced manholes are typically vacuum tested and the line connections are tested, as described in the Point Repair section.

WATER MAIN REPAIR AND RENEWAL—POINT REPAIRS FOR FORCE MAINS

Point Repairs of Water Mains

Point repairs on water mains are similar to repairs on force mains, except that water mains may be under constant pressure. Normally, a sufficient number of valves in the water distribution system can be closed to reduce or eliminate any pressure in the water main. Point repairs can be completed using a repair clamp or sleeve with the proper pressure rating and size for the pipe to be repaired. Water main pipe materials vary from ductile iron, cast iron, old sand cast iron with variable diameters and rough surfaces to asbestos cement and concrete to smooth wall PVC and polyethylene plastic. Repair clamps are normally stainless steel with rubber gaskets and seating surfaces. Repair sleeves are usually ductile iron with rubber gaskets.

If the portion of the water main to be repaired is large enough, sections of pipe may need to be removed, replaced, and connected with one or two repair clamps/sleeves. If a new bell joint can be used at one end of the repair, only one clamp/sleeve would be needed.

Pipe bedding must be removed below the clamp/sleeve for proper cleaning and preparation of the pipe to receive the repair clamp/sleeve. Once repairs have been made, the repaired section may be pressurized to check for leaks at the joints and repair clamps/sleeves.

After satisfactory testing or observation of the repair under pressure, pipe joints and repair clamps/sleeves need to be properly bedded. After bedding of the entire repair section, the trench is backfilled and compacted, as it would be for new construction. Backfill materials may be excavated soil, granular materials, and flowable fill.

Line Stops

At times, there may not be an adequate number of valves in the existing system to isolate the water main to be repaired. In these instances, it may be necessary to install line stops to complete the repair. A line stop consists of a tee installed on the water main to provide access to install a flexible rubber plug in the water main that acts as a closed valve. Line stops can be installed in a water main while the water main is still pressurized.

The tee is bolted to the water main and a specially constructed valve is installed on the branch of the tee. A cutting tool is inserted through the valve and a circular hole is cut into the existing water main. The cutting tool is retracted, the valve is closed, and the cutting tool is removed. The rubber stopper or plug is inserted through the valve and pushed into the water main, filling the entire pipe with a rubber plug. The line stop acts as a temporary closed valve. Repairs are completed to the water main using normal pipe repair methods, and once the repairs are complete, the plug and valve are removed from the line stop tee. A threaded cap is installed in the tee to close the tee. The tee is backfilled as part of the pipe but can be used again in the future if a temporary closed valve is necessary.

Insertion Valves

At times, there may not be an adequate number of valves in the existing system to isolate the water main to be repaired. In these instances, it may be desirable to insert a permanent valve to allow making the repair and to provide a way to isolate water main sections in the future. The procedures to insert a valve are similar to those for installing a line stop.

An insertion valve body consisting of at least two sections is bolted to the water main and a temporary valve is installed on top of the valve body. All work to cut the wall of the water main and to install the valve gate is done through the temporary valve. A cutting tool is inserted through the valve and a circular hole is cut in the existing water main. The cutting tool is removed and the valve gate and operator are installed in the valve body and water main. The valve body and operator are fastened together and the temporary valve is removed.

The insertion valve and joints are tested for leaks and once accepted, can be bedded and backfilled as new construction of a water main valve.

Freezing and Clamping of Service Pipes

Repair of service pipes can sometimes be completed by freezing or by clamping a service line. Wrapping dry ice around the entire circumference of the pipe usually completes freezing. For freezing to work, there must be no flow through the pipe, so one end of the service line must be plugged or crimped to stop all flow. Once the water in the service pipe is frozen, the plugged/crimped end can be cut off and repairs made.

Clamping a service line is limited to use on polyethylene plastic service pipe. Polyethylene can be completely clamped tight and once the clamp is removed, it rebounds to match its original shape closely.

Testing

As discussed before, testing usually consists of physically observing the repaired section of pipe and any valves and fittings for leaks after pressure is restored to the water main. Any leaks must be repaired to the satisfaction of the water utility or municipal personnel. It is advantageous to have the circumference of all joints exposed so any dripping on the bottom of the pipe can be observed.

Disinfection

Under normal repair conditions, the repaired water main cannot be disinfected by methods used for new construction. Repairs and water in the water main are not usually disinfected after the repair is complete because the water mains usually serve existing customers. Instead, the repair pipe, valves, fittings, line stops, etc., are cleaned and disinfected during repair procedures. Cleaning and disinfecting is done by swabbing the materials with a strong solution of chlorine prior to installation.

After the repair is complete, the water main is usually flushed with a high flow rate of water by opening one or several fire hydrants near the repair section. Many water utilities and municipalities also issue a “boil order” to have customers boil their water for 1 or 2 days following a water main repair.

Public Notification

Customers affected by reduced pressure or a water main shutdown should be notified of the repairs being made and precautions to take once water service is restored. "Boil orders" are often used to protect public health and welfare. Boil orders are issued by personal contact, door-to-door notice, daily newspaper articles, and radio/television notice. The water utility or municipality normally takes responsibility to notify its customers of repairs and precautions.

WATER HAMMER: QUANTITATIVE CAUSES AND EFFECTS

Rathnavel Ponnuswami
Care2
Redwood City, California

Water hammer is a hydraulic phenomenon that has to be considered for the safe design of water-carrying pipes. It is basically a pressure fluctuation that would result in disturbing noises and, more importantly, rupture of pipes and fittings (1,2).

DEFINITION

Water hammer can be defined as a transitory increase in pressure in a water system when there is a sudden change of direction or velocity of the water. Water hammer may also be known by names like shockwave effect or hydraulic shock.

If the velocity of water flowing in a pipe is suddenly checked, the energy given up by the water will be utilized in compressing the water itself. This pressure energy is transferred to the pipe wall and this leads to the setup of shock waves in the system. These shock waves travel backward (with a speed similar to the speed of sound), until encountering a solid obstacle (the valve), and crash into the obstacle back and forth repeatedly. This repeated crashing produces an undesirable sound and may lead to rupture of valves and pipes.

The concept of water hammer is normally explained by drawing an analogy to the movement of a train. Assume a real train, instead of slowing to a stop, hits into a mountainside. The back of the train continues forward even though the front cannot go anywhere. Similarly, the hammer occurs because an entire train of water is being stopped so fast that the end of the train hits up against the front end and sends shock waves through the pipe.

EFFECTS OF WATER HAMMER

Water hammer cannot be ignored as another hydraulic effort in water conveyance, as it poses undesirable effects during water conveyance. The major effects of water hammer are the following:

1. Water hammer will cause disturbing noises that tend to be highly problematic in houses and industrial areas.
2. Water hammer may account for the malfunctioning of pipe appurtenances like pressure relief valves, mud valves, backflow preventers, or stop valves.
3. Water hammer may also lead to rupture of the pipes and connected fittings.
4. Water hammer may at times lead to minor flow losses.

CAUSES OF WATER HAMMER

The problem of water hammer can be encountered in varied places like houses, irrigation systems, domestic water supply, and industries. The causes for water hammer (i.e., causes for sudden change in flow velocity/direction) are varied, yet they can be grouped under five major cases:

- Rapid opening or closing of control valves.
- Starting and stopping of pumps.
- Recombination of water after water column separation.
- Rapid exhaustion of all air from the system.
- Power failure in water pumping system.

COMPUTATION OF WATER HAMMER

The pressure rise caused by the water hammer should be quantified to understand its impact and to sort out remedial measures. The rise in pressure due to water hammer can be attributed to pipe velocity, pipe diameter, and pressure wave travel in the pipe. The maximum pressure caused by water hammer can be obtained by the expression

$$H_{\max} = a \times V_n/g$$

where $H_{\max}$ is the maximum pressure rise due to water hammer, V_n refers to the velocity of flow in the pipe, g refers to the acceleration due to gravity, and a is the velocity of pressure wave travel, which in turn can be obtained by the expression

$$a = 1425/(1 + kd/EC)^{1/2}$$

where k is the bulk modulus of water, d is pipe diameter, C is wall thickness of the pipe in meters, and E is the modulus of elasticity of the pipe material.

The pressure exerted due to water hammer is taken into consideration for the structural design of pipelines for public water supplies. In case of unavailability of data and as a rule of thumb, Table 1 presents the extra pressures taken depending on pipe size.

Table 1. Water Hammer Pressure for Different Pipe Diameters

Diameter of pipe, mm	<250	300	400	500	600	750	>750
Water hammer pressure, kg/cm^2	8.50	7.75	7.00	6.30	5.75	5.60	4.90

Today there are a lot of computational facilities also available for the quantification of water hammer pressure.

CONTROL OF WATER HAMMER

As a hydraulic effect, water hammer is undesirable but inevitable. However, there are a lot of techniques to minimize the pressure rise due to the water hammer, which need to be exercised for proper flow profile.

1. The best and simplest technique would be to close or open the valves slowly. Even the first 80% of the valve travel may be executed conveniently, but the last 20% should be done at a snail's pace to avoid problems. The usage of flywheels for this purpose is advisable, and automatically controlled valves with slow closing are desirable.
2. The provision of check valves, surge relief valves, and similar instruments may help in reducing the water hammer. Today, specialized water hammer arrestors are commercially available for this purpose. Apart from these surge-absorbing water hammer arrestors, pressure regulation devices can also be used. An example is a pressure snubber, a device for slowing the rate of change of system flow that will protect from water hammer damage if installed (using a proper size) near the input of a pressure sensor.
3. The pumps are also a matter of concern. Pump startup problems can usually be avoided by increasing the flow slowly to collapse or flush out the voids gently. Also, an uninterrupted power supply for pumps would prevent water hammer. The above arrangements may also contribute to lower horsepower and maximum operating efficiency.
4. Great care should be taken to incorporate the pressure rise due to water hammer during the structural design of pipelines.

There has not been much study about the phenomenon of water hammer and its significance in the design and performance aspects of a water system, but the current scientific forum is coming up with better descriptions.

BIBLIOGRAPHY

1. Singh, G. (1994). *A Text Book of Water Supply Engineering*. Standard Publishers & Distributors, New Delhi, pp. 198–201.
2. Lahlou, M. (2003). Tech brief—water hammer. *On Tap Mag.*, Winter.

CONSTRUCTED WETLANDS

Guenter Langergraber
BOKU—University of Natural
Resources and Applied Life
Sciences
Vienna, Austria

The term constructed wetlands describes artificial wetlands designed to improve water quality. A constructed wetland is a complex mixture of water, substrate, plants, litter (fallen plant material), and a variety of microorganisms (most importantly, bacteria). The mechanisms for improving water quality include settling of suspended particulate matter; filtration and chemical precipitation; chemical transformation; adsorption and ion exchange on surfaces of plants, substrate, and litter; breakdown, transformation, and uptake of pollutants and nutrients by microorganisms and plants; and predation and natural die-off of pathogens.

Constructed wetlands provide the optimal treatment conditions found in natural wetlands but have the flexibility of being constructed. For optimal support of the treatment mechanisms, various types of constructed wetlands optimized for different applications are used. Free-water-surface constructed wetlands and subsurface-flow constructed wetlands are the main types of constructed wetlands. The applications include treatment of domestic, industrial, and agricultural wastewaters, storm water, and landfill leachate.

Wetland treatment systems are effective in treating organic matter, nitrogen, phosphorus, and additionally for decreasing the concentrations of heavy metals, organic chemicals, and pathogens. A good pre-treatment for suspended solids is essential for long-term operation.

MECHANISMS FOR WATER QUALITY IMPROVEMENT

Organic Matter

Treatment wetlands receive large external supplies of carbon. Degradable carbon compounds are utilized in wetland processes. At the same time, degradation processes produce biodegradable carbon compounds. Microorganisms use degradable organic matter as a substrate for respiration and growth.

In general, wetlands are efficient in reducing organic matter. Organic matter exists in particulate and dissolved form at the influent of constructed wetlands. The concentration of organic compounds is often reduced to the background level at the effluent.

In aerobic zones where dissolved oxygen is available, organic matter is decomposed by respiration of heterotrophic organisms that produce carbon dioxide (CO_2) and water. Under anoxic conditions (no dissolved molecular oxygen is available), nitrate is used as an electron acceptor and is reduced to dinitrogen (denitrification). In anaerobic zones (neither molecular nor bound oxygen is available), degradation of organic matter is a multistep process in which CO_2 and methane (CH_4) are produced. Under anaerobic conditions, organic matter can also be reduced by sulfate and iron. However, these pathways play a minor role in subsurface-flow constructed wetlands.

Particulate matter in subsurface-flow wetlands has a large impact on the porous substrate. Particles settle into the pores, and the settled matter is degraded as described before. If the decomposition rate of the particulates is less than the settling rate, the pore volume is reduced, and this causes clogging. The potential for clogging is higher near the inlet of subsurface-flow systems. To prevent clogging, only pre-treated wastewater can be used in combination with subsurface-flow constructed wetlands.

Nitrogen

Nitrogen compounds are among the most important constituents of wastewater due to their role in eutrophication, their effect on the oxygen content of receiving waters, and their toxicity to aquatic organisms and animals. The most important inorganic forms of nitrogen in wetlands are ammonia (NH_4^+), nitrite (NO_2^-), nitrate (NO_3^-), and dissolved elemental nitrogen or dinitrogen gas (N_2). Nitrogen is also present in many organic forms, including urea, amino acids, amines, purines, and pyrimidines. Organic nitrogen compounds are also a fraction of the dry weight of plants, microorganisms, detritus, and soils.

The various forms of nitrogen are continually involved in biochemical transformations. The major processes involved in the nitrogen cycle are ammonification, nitrification, denitrification, nitrogen fixation, nitrogen assimilation, and ammonia volatilization (Fig. 1).

Ammonification. Ammonification is the transformation of organic nitrogen to ammonia. This process occurs during the breakdown of organic matter. The rate of ammonification is directly related to the rate of degradation of organic matter.

Nitrification. Nitrification is the process of oxidizing ammonia to nitrate. The oxidation is done by autotrophic bacteria, which are entirely dependent on generating energy from nitrification. Nitrification is strictly an aerobic process (the presence of oxygen is indispensable) and takes place in two steps. *Nitrosomonas sp.* are responsible for the first step of nitrification: ammonia is oxidized, and nitrite is produced. The nitrite produced is further oxidized to nitrate by *Nitrobacter sp*. Due to the high oxygen demand of the nitrification process (4.57 g oxygen are required for 1 g of ammonia nitrogen), the elimination of ammonia nitrogen is one of the major requirements of wastewater treatment to guarantee a sufficient oxygen concentration in receivers.

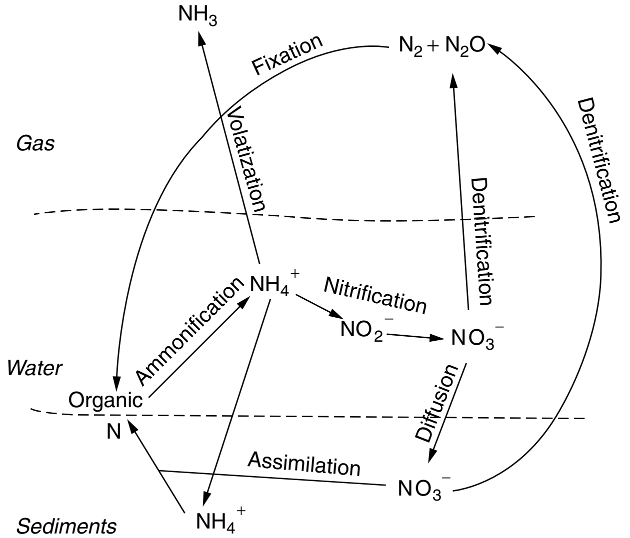

Figure 1. Simplified wetland nitrogen cycle.

Denitrification. Under anoxic conditions and if easily biodegradable organic matter is present, heterotrophic bacteria reduce nitrate to molecular dinitrogen by denitrification. Besides dinitrogen, small amounts of nitrous oxide (N_2O) and nitric oxide (NO) are produced. These dissolved gases are released to the atmosphere, and this reduces the total nitrogen content of treated water. Denitrification and nitrification occur simultaneously in soils where both aerobic and anoxic zones exist nearby. The lack of carbon in the anoxic zones where nitrate is available is the limiting factor for denitrification in wetlands.

Nitrogen Assimilation. Nitrogen assimilation refers to a variety of biological processes that convert inorganic nitrogen into organic compounds. Ammonia uptake is more favored by wetland plants and bacteria than nitrate uptake. In general, the amount of nitrogen taken up by plants and incorporated into the bacterial biomass is small compared to the nitrogen influent load of constructed wetlands for wastewater teatment.

Other Processes. Bacteria, blue-green algae, and plants can fix nitrogen. Nitrogen gas in the atmosphere diffuses into solution and is reduced to ammonia. However, this is generally not observed in constructed wetlands because fixation is more effective in nitrogen-poor environments.

Un-ionized ammonia (NH_3) is relatively volatile, so it can be removed from solution to the atmosphere via diffusion. Usually less than 1% of the total ammonia is present in un-ionized form.

Ammonium can be adsorbed onto active sites of a soil matrix. Only intermittent loading of a system will show rapid removal of ammonia by adsorption due to depletion of ammonia on the adsorption sites during resting periods.

Phosphorus

Phosphorus is typically present in wastewater as orthophosphate and organic phosphate. Phosphorus is released during the degradation of organic matter. Orthophosphate is incorporated into the biomass of organisms during growth. Phosphorus removal in wetland treatment systems occurs by adsorption, plant uptake, complexation, and precipitation.

Plant uptake is not a suitable measure of the net removal rate in a wetland because most of the stored phosphorus is returned to the water by decomposition processes. Due to the limited sorption sites in the soil matrix, phosphorus removal via adsorption is time dependent. After a phosphorus removal rate of almost 100% at the start-up of the operation of the constructed wetland, the removal rate becomes very low, once all sorption sites are filled. During times of low phosphorus influent concentration, also a release of phosphorus can be observed.

Other Substances

Besides the main constituents of wastewater, organic matter, nitrogen, and phosphorus, other pollutants such as heavy metals, specific organic compounds, and pathogens can be removed from water using constructed wetlands.

Heavy Metals. Wetlands interact strongly with heavy metals in a number of ways and thus are capable of significant metal removal. The three major mechanisms are

- binding to soil, sediments, particulates, and soluble organic matter;
- precipitation as insoluble salts, principally sulfides and oxyhydroxides; and
- uptake by plants and bacteria.

A criterion for heavy metal removal is the contact time between dissolved metals and the soil matrix. The longer the contact time, the more exchange processes occur. A particularly significant effect can be observed on the adsorbed metal concentration in the rhizosphere by using different plant species. The plant organisms can, however, become stressed at high heavy metal concentrations, so that the treatment is more appropriate to low metal influent levels. Harvesting the plants used for metal removal then provides a means to prevent metals cycling in the wetland ecosystem and a sustainable means of disposal, if the harvested plants are then appropriately dealt with. Experimental evidence has shown that heavy metal concentrations are highest in the plant roots, followed by the stems, and then the leaves, so that the whole plants must be harvested, not just the foliage.

Specific Organic Compounds. Many constructed wetlands deal with domestic wastewater where BOD and DOC are used as a sum parameter for organic matter. However, other special organic compounds can be removed by using constructed wetlands. Experience exists with waters containing surfactants, solvents, mineral oils, and pesticides. Some specific compounds that were successfully treated by constructed wetlands are, for example, MTBE, trichloroethane, BTEX, cyanide, and explosives (e.g., TNT). Major removal mechanisms are volatilization, photochemical oxidation, sedimentation, sorption, and microbial degradation by fermentation or aerobic and anaerobic respiration. Bioaugmentation of the sediment and sorption by macrophytes is particularly important. In general, there is a lack of knowledge on the detailed removal pathways for organic compounds.

Pathogens and Parasites. Human pathogens are typically present in untreated wastewater. Their populations are highly variable and depend on the health of the human population contributing to the waste. Typical pathogen populations in wastewater include viruses, bacteria, fungi, and protozoa. Parasites such as helminth worms are also common.

Natural treatment technologies can reduce populations of human pathogens by filtration, natural die-off rates, and hostile environmental conditions. Constructed wetlands reduce pathogen populations at varying but significant degrees of effectiveness. The elimination efficiency varies with macrophyte type and increases with the retention time of the wetland.

The Role of Plants

Macrophytes growing in constructed wetlands have several properties in relation to the treatment process. This makes plants an essential part of the design of constructed wetlands.

The most important effects of the macrophytes ragarding the treatment process are physical effects. The roots provide surface area for attached microorganisms, and root growth maintains the hydraulic properties of the substrate. The vegetation cover protects the surface from erosion, and shading prevents algae growth. Litter provides an insulating layer on the wetland surface (especially for operation during winter).

As far as is known, the effect of plant uptake plays a minor role in common wastewater parameters compared to the degradation by microorganisms. For other pollutants such as heavy metals and special organic compounds, the selection of different plant species can, however, play a major role in enhancing treatment efficiency. Both plant productivity and pollutant removal efficiency are relevant in finding an appropriate plant for a given application.

If the wetland is not harvested, the substances incorporated in the plant will be returned to the system during decomposition of the plants. When dead plants are degraded, the organic constituents can act as an additional carbon source for denitrification. Some plants also release organic compounds from their roots, which can also be used for denitrification.

Oxygen release from roots into the rhizosphere is well documented but *in situ* release rates, still are a matter of controversy. Compared to the amount of oxygen brought into the system from the atmosphere via convection and diffusion, the release of oxygen from plant roots can be neglected.

Plants also have other functions not directly related to the treatment process. They provide a habitat for wildlife (including birds and reptiles in large systems) and give the treatment system a more aesthetic appearance.

TYPES OF CONSTRUCTED WETLANDS

Surface-Flow Constructed Wetlands

Surface-flow (SF) or free-water-surface (FWS) wetland technology is strongly related to natural wetlands. Wetlands have been used for wastewater discharge for as long as sewage has been collected. When monitoring was initiated at some of the discharges, an awareness of the water purification potential of wetlands began to emerge. The ‘technology’ started in the 1970s in North America with ecological engineering of natural wetlands for wastewater treatment.

In surface-flow wetlands (Fig. 2), water flows over the soil surface from an inlet to an outlet point. Inflow water containing particulate and dissolved pollutants slows down and spreads through the area of shallow water. The main removal mechanism for particulates is settling. Particulates contain organic matter and enter the biogeochemical cycle in the water column and wetland soil surface. Dissolved pollutants are sorbed by plants, soil, and active microbial populations throughout the wetland and then also enter the overall mineral cycles of the wetland ecosystem. The amount of sorption processes is related to the area where these processes can take place. Due to

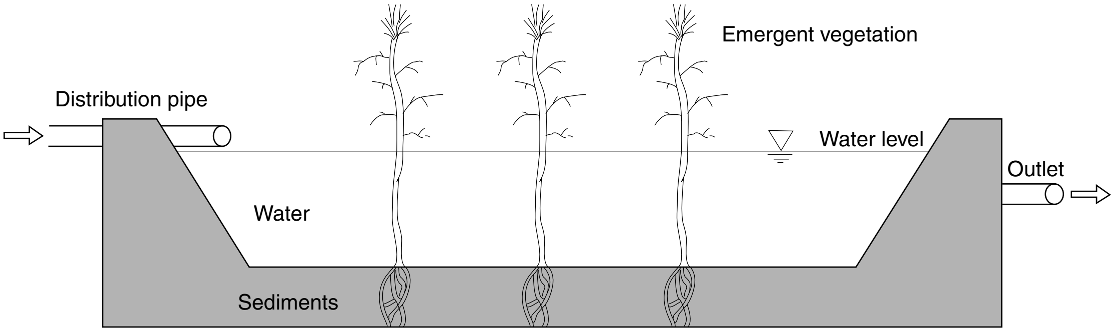

Figure 2. Surface-flow wetland containing emergent macrophytes.

the small surface area of the wetland soil, surface-flow constructed wetlands require a large area to reach good elimination rates of dissolved pollutants.

Surface-flow constructed wetlands can be classified by the plants that are used. Most treatment wetlands use emergent macrophytes (e.g., common reed—*Phragmites australis*), plants whose parts extend above the wetland waters (Fig. 2). Other types use free-floating macrophytes (e.g., water hyacinth—*Eichornia crassipes*), floating-leaved, bottom-rooted macrophytes (rooted in the soil, but their leaves float on the water surface, e.g., water lilies—*Nymphea* spp.), submersed macrophytes (plants that are buoyant and suspended in the water column), or floating mats (some emergent macrophytes can form floating mats, e.g., cattails—*Typha* spp. and common reed—*Phragmites australis*).

Summarizing, one can say that the main disadvantages of free-water-flow constructed wetlands are that they require a large area, they may freeze in temperate climates, and they provide a breeding ground for mosquitoes.

Subsurface-Flow Constructed Wetlands

Subsurface-flow (SSF) constructed wetlands can be subdivided into horizontal flow (HF) and vertical flow (VF) systems depending on the direction of water flow through the porous medium (soil or gravel). The use of subsurface-flow constructed wetlands is limited to pre-treated water, which contains only a low particulate content. Compared to surface-flow systems, the contact area of water with bacteria and substrate is much larger. This enhances the process rates of the system and therefore decreases the area requirement of the treatment system. Pathogen reductions for SSF compared with SWF for similarly sized and managed systems are higher. Subsurface-flow constructed wetlands are frequently planted with common reed (*Phragmites australis*). Therefore a common term is a reed bed treatment system. Other emergent plants like cattails (*Typha* spp.) and bulrushes (*Scirpus* spp.) are also used.

Horizontal-Flow Constructed Wetlands. In a horizontal-flow constructed wetland (Fig. 3), water is fed in at the inlet and flows slowly under the surface through the porous media until it reaches the outlet zone, where it is collected and discharged. Despite regular topography and homogenous plant growth, obtaining an ideal flow is not guaranteed, so that tracer studies are advisable to ensure good hydraulic design. During the passage through the system, water comes in contact with a network of aerobic, anoxic, and anaerobic zones in the substrate. Aerobic zones occur in the upper parts of the water table, anaerobic conditions at the bottom of the wetland.

Oxygen required for aerobic processes is supplied mainly from the atmosphere via diffusion. The oxygen transport from the roots into zones under the water table is too weak to facilitate aerobic processes. Therefore, anoxic and anaerobic processes play an important role in horizontal-flow constructed wetlands. Organic matter is decomposed aerobically and anaerobically and is, therefore, removed well. Due to the poor oxygen supply, only incomplete nitrification can take place.

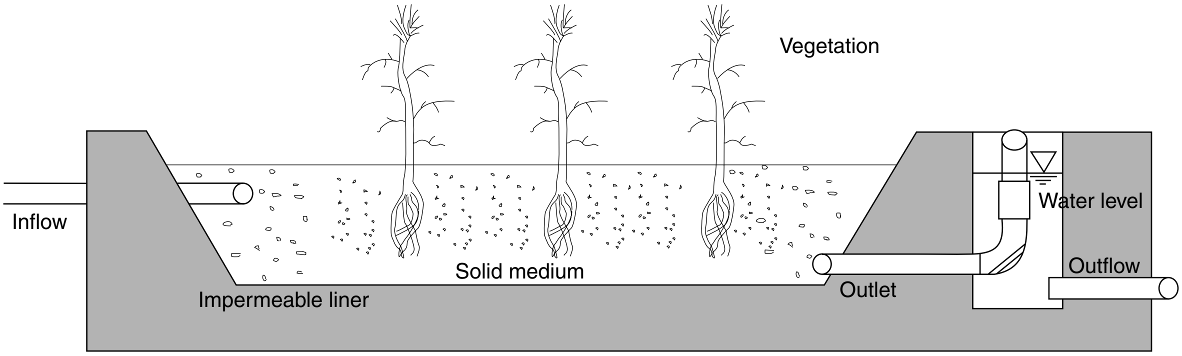

Figure 3. Longitudinal cross section of a horizontal subsurface-flow constructed wetland.

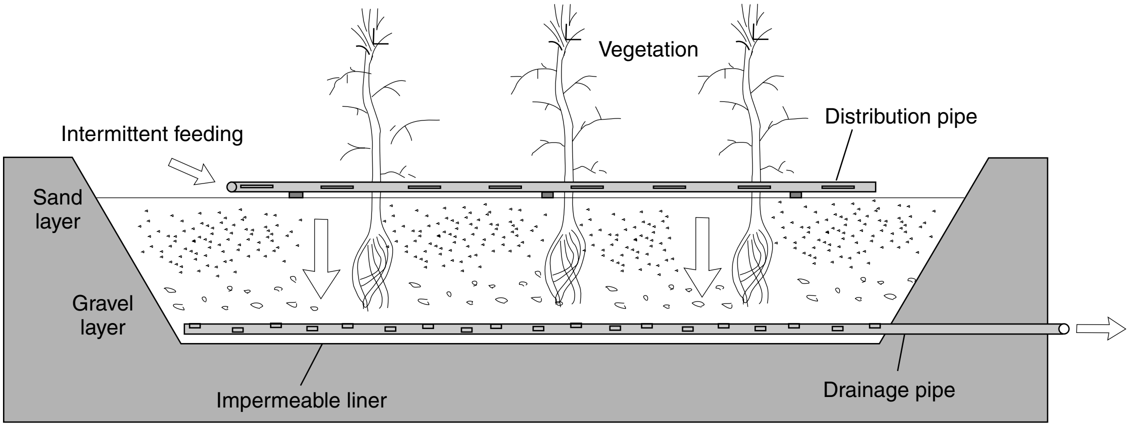

Figure 4. Typical construction of a vertical-flow constructed wetland.

Vertical-Flow Constructed Wetlands. Vertical-flow constructed wetlands look like the system shown in Fig. 4. Water is fed to the system intermittently. The large amount of water from a single feeding causes flooding of the surface. The water infiltrates into the substrate, then gradually drains down vertically, and is collected by a drainage network at the base. Until the next loading, oxygen re-enters the system, and good oxygen transfer into the system is possible. Vertical-flow constructed wetlands are, therefore, suitable when nitrification and other strictly aerobic processes are important factors in the treatment process. When high ammonia nitrogen elimination is required, only vertical-flow constructed wetlands intermittent loading can be used.

APPLICATION OF CONSTRUCTED WETLANDS

The applications of constructed wetlands for water treatment are widespread and include treatment of domestic wastewater, stormwater management in urban areas, treatment of surface water and industrial wastewater, including special wastewater (e.g., from hospitals), and sludge consolidation. Some of these of are discussed more fully here.

Domestic Wastewater. Constructed wetlands for domestic wastewater treatment are generally applied as a main or tertiary treatment stage. For use as a main treatment stage, only subsurface-flow constructed wetlands are used. For tertiary treatment, both surface and subsurface-flow constructed wetland are widely used.

As a main treatment stage, either horizontal-flow or vertical-flow constructed wetlands can be used, depending on the effluent quality desired. Good pre-treatment is necessary to reduce the loading of suspended solids. If low ammonia effluent concentrations are required, only vertical-flow constructed wetlands with intermittent loading fed can guarantee good nitrification. To improve denitrification rates, combined systems with horizontal and vertical types of SSF constructed wetlands can be used.

Sufficient oxygen supply is the main factor for good performance of vertical-flow constructed wetlands. Design recommendations can be based on calculations of oxygen demand and oxygen supply. Considerations of long-term operation of vertical-flow constructed wetlands should, in particular, consider pore size reduction from influent inorganic suspended solids.

Agricultural Wastewater and Food Wastes. Constructed wetlands were used to treat agricultural wastewater from farms with animal production. Crop runoff and pesticides such as atrazine have also been effectively treated Food processing wastes usually have high organic loads that are easily biodegradable. Constructed wetlands have been used to treat potato processing water, olive oil mill water, and dairy wastewater.

Industrial Wastewater. Constructed wetlands can be used to treat several kinds of industrial wastewater: coal and metal mining water, oil extraction, refinery effluents, oil-sand processing water, and wastewater produced by the pulp and paper industries. Many organic and heavy metal wastes have been effectively retained or removed. Constructed wetland technology was also successfully applied to treat hospital wastewater.

Landfill Leachate. Landfills generate highly pollutant leachates as they decompose anaerobically over many years. Constructed wetlands are used for leachate treatment due to their long-term sustainable treatment and low operating and maintenance costs. The retention capacity of wetlands can also be sufficient to prevent contamination of downstream water sources in a heavily contaminated area.

Stormwater and Runoff Management. Uncontrolled urban stormwater has been identified as a major contributor to the nonpoint source pollution of surface waters. Small amounts of rainwater often carry large amounts of pollutants. Three approaches are used to control urban stormwater using constructed wetlands: dry and wet detention ponds (where the ponds remain dry or wet between floods) and stormwater wetlands. Detention ponds collect stormwater and release it during dry weather. The main removal mechanism is settling.

Surface Water. Heavily polluted surface water can also be treated using constructed wetland technology. The treated water can be used, for example, for groundwater recharge and for restoring contaminated surface waters. However, it is not possible to produce an effluent quality that allows direct use as drinking water.

SUMMARY

Constructed wetlands are artificial wetlands designed to improve water quality. They provide the optimal treatment conditions found in natural wetlands but have the flexibility of being constructed. Wetland treatment systems are effective in treating organic matter, nitrogen, phosphorus, and additionally for decreasing the concentrations of heavy metals, organic chemicals, and pathogens. Various types of constructed wetlands optimized for different applications are used for optimal support of the treatment mechanisms. The main types are free-water-surface constructed wetlands and subsurface-flow constructed wetlands with horizontal or vertical flow.

In general, the use of constructed wetlands provides a relatively simple, inexpensive, and robust solution for treating water. Compared to other treatment options, constructed wetlands usually need less operation and maintenance. Additional benefits include their tolerance to fluctuations in flow, the facility of water reuse and recycling, the provision of habitat for many wetland organisms, and the more aesthetic appearance of a natural system compared with technical treatment options.

Acknowledgments
Special thanks to Mr. Phil Russell for his work in proofreading and editing.

BIBLIOGRAPHY

1. Cooper, P.F., Green, M.B., and Shutes, R.B.E. (1996). Reed beds and constructed wetlands for wastewater treatment. WRc Swindon, Swindon, UK.
2. Haberl, R., Grego, S., Langergraber, G., Kadlec, R.H., Cicalini, A.R., Martins Dias, S., Novais, J.M., Aubert, S., Gerth, A., Hartmut, A., Hartmut, T., and Hebner, A. (2003). Constructed wetlands for the treatment of organic pollutants. *JSS—J Soils & Sediments* **3**(2): 109–124.
3. Kadlec, R.H., Knight, R.L., Vymazal, J., Brix, H., Cooper, P., Haberl, R. (Eds.). (2000). *Constructed Wetlands for Pollution Control—Processes, Performance, Design and Operation*. IWA Scientific and Technical Report No. 8, IWA Publishing, London, UK.
4. Kadlec, R.H. and Knight, R.L. (1996). *Treatment Wetlands*. CRC Press, Boca Raton, FL.
5. Langergraber, G. and Haberl, R. (2001). Constructed wetlands for water treatment. *Minerva Biotechnologica* **13**(2): 123–134.
6. US EPA. (2000). *Constructed Wetlands Treatment of Municipal Wastewaters*. Report EPA/625/R-99/010. US EPA, Cincinnatti, OH.
7. Vymazal, J., Brix, H., Cooper, P.F., Green, M.B., and Haberl, R. (Eds.). (1998). *Constructed Wetlands for Wastewater Treatment in Europe*. Backhuys Publishers, Leiden, The Netherlands.
8. Wissing, F. (1995). Abwasserreinigung mit Pflanzen. Ulmer Verlag, Stuttgart, Germany [*In German*].

USING ECOSYSTEM PROCESSES IN A CONSTRUCTED WETLAND TO TREAT MINE WASTEWATER IN IRELAND

Aisling D. O'Sullivan
Marinus L. Otte
University College Dublin
Belfield, Ireland

WASTEWATER GENERATION

The development and expansion of industrial and agricultural activities in the Western world has led to a corresponding increase in the release of potentially toxic chemicals into the environment. Although growth has had many socioeconomic benefits, it has also led to negative implications for wetland ecosystems. For instance, the intensive farming practices developed in Ireland in the 1970s led to eutrophic (overenriched) rivers and lakes as a result of inadvertent inputs of excess nitrogen and phosphorus. This resulted in algal blooms, subsequent fish mortality, and reduced biodiversity of aquatic systems. The breaching of tailings dam walls in mine storage ponds has also resulted in catastrophic impacts on neighboring watersheds, which in many circumstances has led to irreversible ecosystem damage. Therefore, a balance is sought between reducing undesirable impacts while facilitating positive economic progress. By understanding the complex processes that affect the biogeochemistry and cycling of contaminants in aquatic ecosystems, innovative treatment technologies can be applied to mitigate and treat these wastes.

During metal mining, the oxidation of ores frequently results in increased reactivity and bioavailability and possibly toxicity of heavy metals (1). Many contaminants such as heavy metals and radionuclides cannot be chemically degraded, but microbial deactivation (mainly through immobilization) is most suitable for the treatment of these types of wastes (2). Other pollutants that prove difficult to dispose of are polychlorinated biphenyls (PCBs) which were used liberally as additives in the manufacture of paints and plastics until their toxicity was realized in the mid-1960s. PCBs have very high chemical, thermal, and biological stability; these properties resulted in their bioaccumulation in sediments and biota. Excess plant nutrients such as nitrogen and phosphorus are typically easier to treat in waters because they can be easily assimilated by plants and algae.

CURRENT TREATMENT PRACTICES AND ALTERNATIVE OPTIONS

Conventional practices for treating wastewater in Ireland adopt predominantly chemical applications that require high energy demands. These traditional technologies typically include the chemical manipulation of contaminants and subsequent physical precipitation of suspended solids in the water. Other means of treating wastewater contaminated with biodegradable organic wastes have proven very successful in the past few decades as wastes are biologically converted into gaseous and solid forms. Treatment practices relying on biological processes include continuous activated sludge reactors, which operate either aerobically or anaerobically. In suspended reactor vessels, microorganisms are kept in suspension in the wastewater, whereas in attached systems, microbial masses adhere typically to an inert medium. Various modifications of the aerobic treatment processes have been developed, including stabilization ponds, aerated waste reservoirs, and rotation and filtration techniques (3). These treatment techniques rely principally on activated processes, that is, they require substantial energy inputs and risk failure without continuous energy inputs.

Nonactivated biological processes employed to degrade contaminants are a more recent practice in wastewater treatment. Worldwide, particularly in North America and Australia, emerging practices for treating wastewater include constructed wetlands. These systems operate by using natural processes and usually do not require substantial energy inputs. The biological processes are typically solar-driven and use carbon and nutrients in the substrate to drive the microbial and plant processes. Therefore, constructed wetlands can be economical and ecologically acceptable. Typically, constructed wetlands are designed specifically for the type of wastewater being treated. Nutrients can be removed in aerobic systems, and sulfates and metals can be removed from wastewater under chemically reducing conditions. Wastewater contaminated by metals has usually been treated primarily to reduce acidity, but some treatment options for this type of waste have shown that substantial metals can be removed concurrently (4).

THE TARA MINES CASE STUDY

Tara Mines Ireland, a subsidiary of the Finnish company Outokumpu-Zinc, is the largest producer of zinc in Europe and the fifth largest lead-zinc mine worldwide. They have been operating since 1977 and supply approximately 20% of the European demand for zinc (5). The local bedrock from which the ore is extracted is classified as Lower Carboniferous calcite ($CaCO_3$) and dolomite ($CaMgCO_3$), which buffers the wastewater used in the processing operations to a pH of approximately 7.8 (5). Spent water is pumped to large storage ponds for treatment at a site about 5 km from the ore extraction site. Some of this wastewater is fed to the experimental treatment wetlands described here.

Treatment wetlands were engineered on-site at Tara Mines in 1997. These wetlands, built specifically to treat sulfate and metal-enriched wastewater emanating from the mine, are the only treatment wetlands of this kind in Europe at this time. Most applications employing microbes to reduce levels of contaminants in wastewater have focused on metal removal (6); the research at Tara Mines was concerned primarily with removing sulfate from the alkaline mine water.

Construction Design at Tara Mines Wetlands

Two similar experimental treatment systems were built adjacent to large tailings ponds. Each system measured 12 m (length) × 3 m (width) × 2 m (depth) or 72 m^3 with a 4:1 (length to width) aspect ratio (7) and comprised three 12 m^2 in-series surface-flow cells: inflow, wetland, and outflow (see Fig. 1 for details). Both systems were lined with a 2000 gauge (light-insensitive) impermeable PVC sheet, and each wetland cell was planted with cattails (*Typha latifolia*), reed (*Phragmites australis*), and floating sweetgrass (*Glyceria fluitans*) in the ratio of 4:9:7 per m^2. Three wooden baffles coated with an industrial-grade waterproof varnish were also placed in each wetland cell to increase the length of the flow path of the water. Mesocosm experiments indicated that an optimal combination of plant growth and substrate permeability was achieved by using a medium containing spent mushroom substrate and fine grit in a 1:3 ratio by volume of. Each cell was filled with a similar mix of this substrate to a depth of about 50 cm (Fig. 1).

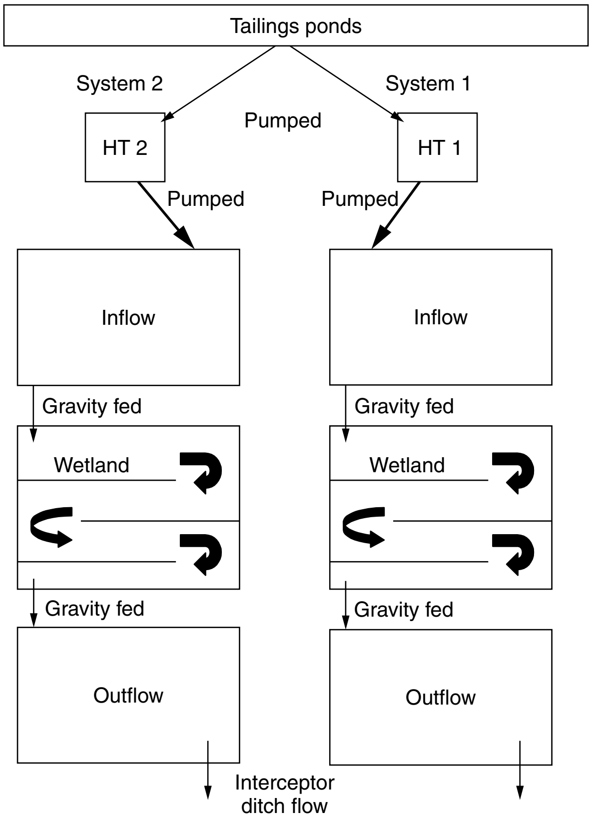

Figure 1. Schematic representation of the experimental treatment wetlands at Tara Mines, Ireland. Arrows indicate the flow path of the water, which was pumped from the storage ponds to the header tanks (HT 1 and HT 2) and subsequently to the inflow cells of each system. From there on, water traveling between cells was gravity fed (by head differences between cells) to the wetland cells and then finally to the outflow cells.

BIOGEOCHEMISTRY—THE KEY TO REMEDIATION

Wetland plants can remove pollutants from contaminated soils and water through uptake, translocation, and compartmentalization in storage tissues (2,8). However, the more significant influence of plants on metal removal from wastewater is indirect by accumulation of metals in the sediments surrounding the roots (8,9). Sufficient organic matter, provided by seasonal plant dieback, is also important in treatment systems that operate on natural biological processes. The accumulation of these contaminants is governed by microbial reactions mediated in suitable substrates and under appropriate conditions. For instance, the substrates used in the treatment wetlands at Tara Mines contained indigenous populations of sulfate-reducing bacteria, identified as the genera *Desulfovibrio, Desulfobulbus, Desulfotomaculum*, and *Desulfoccus*. The systems were permanently flooded and this provided net anaerobic substrate conditions conducive to the chemical reduction of sulfate (SO_4^{2-}) to sulfide (S^{2-}). This reaction occurred as the microorganisms assimilated sulfate in the absence of oxygen, thus reducing it to sulfide through the transfer of electrons produced by the simultaneous oxidation of the organic substrate. Sulfide anions resulting from the reaction are very unstable and readily react with free or sorbed metal cations to form metal sulfides such as zinc sulfide (ZnS), lead sulfide (PbS), and iron sulfide (FeS) (10). However, sulfide can also react with hydrogen to form hydrogen sulfide. Once the overlying water became saturated with hydrogen sulfide, it evolved to the atmosphere as a gas. Although evolution of this gas has not yet been quantified, it appears that substantial amounts were produced in the Tara Mines treatment wetlands, as evident from the frequent detection (olfactory). This may be a concern in the application of a larger scale treatment wetland and has not yet been fully addressed. Nonetheless, in the Tara Mines case study, the capacity to treat metal and sulfate contaminated wastewater using natural ecosystem processes was successfully demonstrated.

SUCCESS STORY—IN MANY WAYS

Treatment Capacity

The experimental treatment wetlands were modeled on the surface-flow design previously established in North America. Effectively, treatment occurred through bioimmobilization of the contaminants in the water as substrate, vegetation, and microbial assemblages interacted at the sediment–water interface. Although sulfate concentrations exiting the treatment systems did not comply with acceptable discharge levels (of 200 mg L^{-1}) stipulated in national legislation, up to 69% of the influent concentration (900 mg L^{-1}) was removed, and this equaled a removal rate of 29 g m^{-2} day^{-1}. Similarly, zinc and lead were removed by up to 99% and 64%, respectively, of the original concentrations supplied of 1.8 mg Zn L^{-1} and 0.2 mg Pb L^{-1}. Concentrations of these metals and of sulfides in the sediments were significantly greater than those in the original substrates, and concentrations in the soil water were also elevated compared to background water levels. Plants contained less than 1% of sulfur in their tissues, in accordance with other studies. Vegetation (including algae that voluntarily colonized the systems) sequestered metals to some extent, but plant roots had almost double the concentration of metals, compared to plant shoots. This may be explained by selective uptake and translocation mechanisms exhibited by some plants that can prevent contaminants from traveling to their shoots. The effect may also be explained by metal hydroxides that can form on plant roots (11). These hydroxides form when localized oxygenation by wetland plant roots and rhizomes induce precipitation of metals. Once metals are precipitated from solution, they become relatively immobile and thus generally less available to living organisms. The vegetation was not harvested from the treatment wetlands, so metals in these tissues were not actually removed from the systems because plant uptake and decay result in cycling of metals within such ecosystems. The most important attributes for ensuring long-term success and sustainability of constructed wetlands that treat these kinds of wastes appear to be wetland size and chemical loading rates. Therefore, these considerations must be clearly evaluated before constructing treatment wetlands.

Ecological and Societal Benefits

Wetlands are recognized for their hydrologic, economic, ecological, and aesthetic values. Natural, restored, and constructed wetlands have become the focus of many scientists in recent years for their capacity to treat wastewater in an ecologically and economically cost-effective way (12). The ecological benefits provided by the Tara Mines treatment wetlands were diverse and innumerable. Soon after they were constructed, macroinvertebrates (35 taxa), plants (6 species), algae, and microorganisms voluntarily colonized the systems. The vegetation afforded refuge to birds (e.g., moorhen) and other wildlife (foxes, rabbits, shrews, etc.), and also provided niches for microbial and invertebrate communities.

It is obvious that surface-flow wetlands offer many ancillary benefits by creating new habitats for wildlife. Initiatives to develop treatment wetlands have frequently involved local communities and educational establishments, thereby integrating educational principles into their goals. Aspects of the study at Tara Mines have been delivered to a wide spectrum of society from high-school teenagers in the local community to undergraduate and postgraduate students in universities in Ireland, Europe, and the United States. Additionally, the work has been well received by industries and local interest groups. By involving communities in the construction of wetlands, awareness of the value of such ecosystems and their capacity to recycle wastes by biological and ecologically acceptable techniques is generated.

TREATMENT SUCCESS COMPARED WITH OTHER SYSTEMS

The performance of the Tara Mines treatment wetlands for sulfate and metal removal was compared with that of other studies. Many of the other investigations quantified removal as a percentage of the influent concentration, but mass removal rates represent a more realistic estimate of the actual amount removed (4). In this instance, contaminant removal is calculated from a loading perspective; considers flow rates, concentrations, and wetland size; and expresses removal per unit time. Sulfate removal at Tara Mines was comparable with that published by Eger (13) of up to 27 g m^{-2} day^{-1}. Metal removal (from a percentage perspective) at Tara Mines was usually greater than that of other studies (e.g., Reference 14; removal of 33% Zn and 26% Pb). However, metal levels in the wastewater treated in the other studies were higher than those in the Tara Mines wastewater, and so higher mass removal rates probably occurred in these other studies.

STIMULUS FOR REVISING WASTE MANAGEMENT PRACTICES

In the past 25 years, most treatment wetland studies were concerned with removing plant nutrients from water. In the last decade, this focus has shifted toward remediating mine wastewater, principally to reduce the impact of acid mine drainage (4). However, constructed wetlands built for treating mine waste were primarily developed in North America, and the scope of similar systems in Europe has yet to be demonstrated. However, the question of long-term reliability of treatment wetlands is still not conclusively answered because most systems of this type have operated less than 15 years. Many scientists will advocate that, once they are appropriately designed and managed, they can remain self-renewing for several decades. The development of constructed wetland technology has demonstrated attractive success rates, and incurred limited financial operational costs. They can also provide ancillary ecological benefits. Increasing pressure in preparing for mine closure following ore exhaustion, coupled with societal concerns regarding environmental quality, is pressuring industries to reevaluate conventional treatment practices. Recent Irish and European Union legislation has contributed in some ways toward increasing the awareness of alternative cost-efficient treatment practices allied with sustainable development. By exploiting this newly found awareness, the possibility of implementing some innovative, ecologically sound waste technologies looks promising.

Acknowledgments
The authors acknowledge the financial support of Tara Mines Ireland and Enterprise Ireland for the research project. Dr. Declan Murray is also acknowledged for his involvement in the research project and our colleagues in the Department of Botany (UCD, Ireland) for their assistance in the field while constructing the treatment wetlands.

BIBLIOGRAPHY

1. Schnoor, J. and Zehnder, A. (1996). *Environmental Modeling: Fate and Transport of Pollutants in Water, Air and Soil*. John Wiley & Sons, New York, pp. 381–451.
2. Salt, D. et al. (1995). Phytoremediation: a novel strategy for the removal of toxic metals from the environment using plants. *Biotechnology* **13**: 468–473.
3. Gray, N. (1999). *Water Technology—an Introduction for Environmental Scientists and Engineers*. John Wiley & Sons, New York, pp. 302–320.
4. Hedin, R., Nairn, R., and Kleinmann, R. (1994). *Passive Treatment of Coal Mine Drainage*. Information Circular #9389, Bureau of Mines, US Department of the Interior.
5. O'Leary, W. (1996). Wastewater recycling and environmental constraints at a base metal mine and process facilities. *Water Sci. Technol.* **33**(10–11): 371–379.
6. Song, Y., Piak, B., Shin, H., and La, S. (1998). Influence of electron donor and toxic materials on the activity of sulfate reducing bacteria for the treatment of electroplating wastewater. *Water Sci. Technol.* **38**(4–5): 187–194.
7. Crites, R. (1994). Design criteria and practice constructed wetlands. *Water Sci. Technol.* **29**(4): 1–6.
8. Otte, M., Kearns, C., and Doyle, M. (1995). Accumulation of arsenic and zinc in the rhizosphere of wetland plants. *Bull. Environ. Contam. Toxicol.* **55**: 154–161.
9. Ye, Z., Baker, A., Wong, M., and Willis, A. (1998). Zinc, lead and cadmium accumulation and tolerance in *Typha latifolia* as affected by iron plaque on the root surface. *Aquatic. Bot.* **61**: 55–67.
10. Morse, J., Millero, F., Cornwell, J., and Richard, D. (1987). The chemistry of the hydrogen sulfide and iron sulfide systems in natural waters. *Earth Sci. Rev.* **24**: 1–42.
11. Doyle, M. and Otte, M. (1997). Organism-induced accumulation of iron, zinc, and arsenic in wetland soils. *Environ. Pollut.* **96**: 1–11.
12. Mitsch, W. and Gosselink, J. (2000). *Wetlands*, 3rd Edn. Van Nostrand Reinhold, New York, pp. 688–697.
13. Eger, P. (1994). Wetland treatment for trace metal removal from mine drainage: the importance of aerobic and anaerobic processes. *Water Sci. Technol.* **29**(4): 249–256.
14. Mays, P. and Edwards, G. (2001). Comparison of heavy metal accumulation in a natural wetland and constructed wetlands receiving acid mine drainage. *Ecol. Eng.* **16**: 487–500.
15. Hedin, R. (1996). Environmental engineering forum: long-term effects of wetland treatment of mine drainage. *J. Environ. Eng.* **122**: 83–85.

READING LIST

Catallo, W. (1993). Ecotoxicology and wetland ecosystems—current understanding and future needs. *Environ. Toxicol. Chem.* **12**(12): 2209–2224.

Dunbabin, J. and Bowmer, K. (1992). Potential use of constructed wetlands for treatment of industrial wastewaters containing metals. *Sci. Total Environ.* **111**: 151–168.

Foras Forbatha. (1990). *Water Resources Division in Eutrophication (1986–1990)*. Publication of the (Irish) National institute for planning and construction research.

Mora, A. (2001). *Solutions Applied to Aznalcollar Disaster*. Abstracts of the COST Action 837 (WG2) Workshop; Phytoremediation of trace elements in contaminated soils and water. C. Camara (Ed.). University of Madrid, Spain, April 5–7, p. 29.

Schuurmann, G. and Markert, B. (1998). *Ecotoxicology*. John Wiley & Sons, London.

WATER AND WASTEWATER PROPERTIES AND CHARACTERISTICS

Absar A. Kazmi
Department of Civil
Engineering
Roorkee, Uttaranchal, India

Water in nature is most nearly pure in its evaporative state. Because the very act of condensation usually requires a surface, or nuclei, water may acquire impurities at the very moment of condensation. Additional impurities are added as the liquid travels through the remainder of the hydrologic cycle and comes into contact with materials in the air and on or beneath the surface of the earth. Human activities contribute further impurities in the form of industrial and domestic wastes, agricultural chemicals, and other less obvious contaminants. Ultimately, these impure waters complete the hydrologic cycle and return to the atmosphere as relatively pure water molecules. However, the water quality in the intermediate stage is of greatest concern because the quality at this stage affects human use of water.

The impurities accumulated by water throughout the hydrologic cycle and from human activities may be in both suspended and dissolved form. Suspended material consists of particles larger than molecular size that are supported by buoyant and viscous forces within the water. Dissolved material consists of molecules and ions that are held by the molecular structure of water. Colloids are very small particles that technically are suspended but often exhibit many of the characteristics of dissolved substances.

Water pollution is the presence of impurities in water in such quantity and of such nature as to impair the use of the water for a stated purpose. Many parameters and characteristics have evolved that qualitatively reflect the impact that various impurities have on selected water uses. Knowledge of properties/characteristics/parameters of water and wastewater treatment processes is essential for environmental scientists and engineers.

PHYSICAL WATER QUALITY PARAMETERS

Physical parameters define those characteristics of water that respond to the senses of sight, touch, taste, or smell. Suspended solids, turbidity, color, taste and odor, and temperature fall into this category.

Suspended Solids

Solids can be dispersed in water in both suspended and dissolved forms. Solids suspended in water may consist of inorganic or organic particles or of immiscible liquids. Inorganic solids such as clay, silt, and other soil constituents are common in surface water. Organic materials such as plant fibers and biological solids (algal cells, bacteria, etc.) are also common constituents of surface waters. These materials are often natural contaminants resulting from the erosive action of water flowing over surfaces. Other suspended material may result from human use of water. Domestic wastewater usually contains large quantities of suspended solids that are mostly organic. Industrial use of water may result in a wide variety of organic or inorganic suspended impurities. Immiscible liquids such as oils and greases are often constituents of wastewater.

Suspended material in water may be objectionable for several reasons. It is aesthetically displeasing and provides adsorption sites for chemical and biological agents. Suspended organic solids may be degraded biologically, resulting in objectionable by-products. Biologically active (live) suspended solids may include disease-causing organisms as well as organisms such as toxin-producing strains of algae.

Suspended solids are likely to be organic and/or biological and are an important parameter of wastewater. The suspended solids parameter is used to measure the quality of wastewater influent, to monitor several treatment processes, and to measure the quality of the effluent. The U.S. Environmental Protection Agency (EPA) has set maximum suspended solids at 30 mg/L for most treated wastewater discharges.

Turbidity

Direct measurement of suspended solids is not usually performed on samples from natural bodies of water or on potable (drinkable) water supplies. The nature of the solids in these waters and the secondary effects they produce are more important than the actual quantity. For such waters, a test for turbidity is commonly used. Turbidity is a measure of the extent to which light is either absorbed or scattered by suspended material in water.

Most turbidity in surface waters results from the erosion of colloidal material such as clay, silt, rock fragments, and metal oxides from the soil. Vegetable fibers and microorganisms may also contribute to turbidity. Household and industrial wastewaters may contain a wide variety of turbidity producing material. Soaps, detergents, and emulsifying agents produce stable colloids that result in turbidity. Although turbidity measurements are not commonly run on wastewater, discharges of wastewaters may increase the turbidity of natural waterbodies.

When turbid water in a small, transparent container, such as a drinking glass, is held up to the light, an aesthetically displeasing opaqueness or “milky” coloration is apparent. The colloidal material associated with turbidity provides adsorption sites for chemicals that may be harmful or cause undesirable tastes and odors and for biological organisms that may be harmful. Disinfection of turbid waters is difficult because of the adsorptive characteristics of some colloids and because the solids may partially shield organisms from the disinfectant.

In natural waterbodies, turbidity may impart a brown or other color to water, depending on the light-absorbing properties of the solids, and may interfere with light penetration and photosynthetic reactions in streams and lakes. Accumulation of turbidity-causing particles in porous streambeds results in sediments that can adversely affect the flora and fauna of the stream.

Color

Pure water is colorless, but water in nature is often colored by foreign substances. Water whose color is partly due to suspended matter is said to have apparent color. Color contributed by dissolved solids that remain after removal of suspended solids is known as true color.

After contact with organic debris such as leaves, conifer needles, weeds, or wood, water picks up tannins, humic acid, and humates and takes on yellowish-brown hues. Iron oxides make water reddish, and manganese oxides make water brown or brackish. Industrial wastes from textile and dyeing operations, pulp and paper production, food processing, chemical production, and mining, refining, and slaughterhouse operations may add substantial coloration to water in receiving streams.

Colored water is not aesthetically acceptable to the general public. In fact, given a choice, consumers tend to choose clear, uncolored water of otherwise poorer quality over treated potable water supplies with an objectionable color. Highly colored water is unsuitable for laundering, dyeing, papermaking, beverage manufacturing, dairy production and other food processing, and textile and plastic production. Thus the color of water affects its marketability for both domestic and industrial use.

Color is not a parameter usually included in wastewater analysis. In potable water analysis, the common practice is to measure only the true color produced by organic acid resulting from decaying vegetation in the water. The resulting value can be taken as an indirect measurement of humic substances in the water.

Taste and Odor

The terms taste and odor define this parameter. Because the sensations of taste and smell are closely related and often confused, a wide variety of taste and odors may be attributed to water by consumers. Substances that produce an odor in water almost invariably impart a taste as well. The converse is not true; many mineral substances produce tastes but no odor. Many substances with which water comes into contact in nature or during human use may impart perceptible taste and odor. These include minerals, metals, and salts from the soil, end products from biological reactions, and constituents of wastewater. Inorganic substances are more likely to produce tastes unaccompanied by odor. Alkaline material imparts a bitter taste to water, and metallic salts may give a salty or bitter taste. Organic materials, on the other hand, are likely to produce both taste and odor in water; petroleum based products are prime offenders. The biological decomposition of organics may also result in taste-and odor-producing liquids and gases in water. Consumers find taste and odor aesthetically displeasing for obvious reasons. Because water is thought of as tasteless and odorless, the consumer associates taste and odor with contamination and may prefer to use a tasteless, odorless water that might actually pose more than a problem of simple aesthetics because some of those substances may be carcinogenic.

Temperature

Temperature is not used directly to evaluate either potable water or wastewater. It is, however, one of the most important parameters in natural surface water systems. The temperature of surface waters governs to a large extent the biological species present and their rates of activity. Temperature has an effect on most chemical reactions that occur in natural water systems. Temperature also has a pronounced effect on the solubilities of gases in water. The temperature of natural water systems responds to many factors, ambient temperatures are the most universal.

CHEMICAL WATER QUALITY PARAMETERS

Water has been called the universal solvent, and chemical parameters are related to the solvent capabilities of water. Total dissolved solids, alkalinity, hardness, fluorides, metals, organics, and nutrients are chemical parameters of concern in water quality management.

Total Dissolved Solids (TDS)

The material remaining in water after filtration for a suspended solids analysis is considered dissolved. Dissolved material results from the solvent action of water on solids, liquids, and gases. Like suspended material, dissolved substances may be organic or inorganic. Inorganic substances that may be dissolved in water include minerals, metals, and gases. Water may come in contact with these substances in the atmosphere, on surfaces, and within the soil. Materials from the decay products of vegetation, from organic chemicals, and from organic gases are common dissolved organic constituents of water. Many dissolved substances are undesirable in water. Some dissolved minerals, gases, and organic constituents are carcinogenic.

Because no distinction is made among the constituents, the TDS parameter is included in the analysis of water and wastewater only as a gross measurement of dissolved material. This is often sufficient for wastewaters, but it is frequently desirable to know more about the composition of the solids in water for use in potable supplies, agriculture, and some industrial processes. When this is the case, tests for several of the ionic constituents of TDS are made.

Akalinity

Alkalinity is the quantity of ions in water that will react to neutralize hydrogen ions. Alkalinity is thus a measure of the ability of water to neutralize acids. The constituents of alkalinity in natural water systems include CO_3^{2-}, HCO_3^-, and OH^-. These ions result from the dissolution of mineral substances in the soil and atmosphere. Phosphates may also originate from detergents in wastewater discharges and from fertilizers and insecticides from agricultural land. Hydrogen sulfide and ammonia may be products of the microbial decomposition of organic material. In large quantities, alkalinity imparts a bitter taste to water. The principal objection to alkaline water, however, is the reactions that can occur between alkalinity and certain cations in the water. The resultant precipitates can foul pipes and other water system appurtenances.

Hardness

Hardness is the concentration of multivalent cations in solution. At supersaturation, hardness cations react with anions in the water to form solid precipitates. The multivalent metallic ions most abundant in natural waters are calcium and magnesium, and for all practical purposes, hardness may be represented by the sum of calcium and magnesium ions.

Soap consumption by hard water represents an economic loss to the water user. Sodium soaps react with multivalent cations to form a precipitate, thereby losing their surfactant properties. Lathering does not occur until all the hardness ions are precipitated, at which point the water has been softened by the soap. Boiler scale, the result of the carbonate hardness precipitate, may cause considerable economic loss through fouling of water heaters and hot water pipes. Changes in pH in water distribution systems may also result in deposits of precipitates. Bicarbonate begins to convert to the less soluble carbonates at a pH above 9.0.

Magnesium hardness, particularly associated with the sulfate ion, has a laxative effect on persons accustomed to it. Magnesium concentrations of less than 50 mg/L are desirable in potable waters, although many public water supplies exceed this amount. Calcium hardness presents no public health problem.

Fluoride

Generally associated in nature with few types of sedimentary or igneous rocks, fluoride is seldom found in appreciable quantities in surface waters and appears in groundwater in only a few geographical regions. Fluoride is toxic to humans and other animals in large quantities, but small concentrations can be beneficial.

Metals

All metals are soluble to some extent in water. Excessive amounts of any metal may present health hazards, but only those metals that are harmful in relatively small amounts are commonly labeled toxic; other metals fall into the nontoxic group. Sources of metals in natural waters include dissolution from natural deposits and discharges of domestic, industrial, or agricultural wastewaters. Metals in water are usually measured by atomic absorption spectrometry.

Organics

Many organic materials are soluble in water. Organics in natural water systems may come from natural sources or may result from human activities. Most natural organics consist of the decay products of organic solids; synthetic organics are usually the result of wastewater discharges or agriculture practices. Dissolved organics in water are usually divided into two broad categories: biodegradable and nonbiodegradable (refractory).

Biodegradable Organics. Biodegradable material consists of organics that can be used for food by naturally occurring microorganisms within a reasonable length of time. In dissolved form, these materials usually consist of starches, fats, proteins, alcohols, acids, aldehydes, and esters. They may be the end products of the initial microbial decomposition of plant or animal tissue, or they may result from domestic or industrial wastewater discharges. Although some of these materials may cause color, taste, and odor problems, the principal problems from biodegradable organics are the secondary effects from the action of microorganisms on these substances.

Nonbiodegradable Organics. Some organic materials are resistant to biological degradation. Tannic and lignic acids, cellulose, and phenols are often found in natural water systems. These constituents of woody plants biodegrade so slowly that they are usually considered refractory. Molecules that have exceptionally strong bonds (some of the polysaccharides) and ringed structures (benzene) are nonbiodegradable.

Nutrients

Nutrients are elements essential to the growth and reproduction of plants and animals. Aquatic species depend on the surrounding water to provide their nutrients. Although a wide variety of minerals and trace elements can be classified as nutrients, those required in most abundance by aquatic species are carbon, nitrogen, and phosphorus. Carbon is readily available from many sources. Carbon dioxide from the atmosphere, alkalinity, and decay products of organic matter all supply carbon to the aquatic system. In most cases, nitrogen and phosphorus are the nutrients that limit aquatic plant growth.

BIOLOGICAL WATER-QUALITY PARAMETERS

Water may serve as a medium in which literally thousands of biological species spend part, if not all, of their life cycles. Aquatic organisms range in size and complexity from the smallest single-cell microorganism to the largest fish. All members of the biological community are to some extent, water quality parameters, because their presence or absence may indicate in general terms the characteristics of a given body of water. As an example, the general quality of water in a trout stream would be expected to exceed that of a stream in which the predominant species of fish is carp. Similarly, abundant algal populations are associated with a water rich in nutrients.

Biologists often use a species-diversity index (related to the number of species and the relative abundance of organisms in each species) as a qualitative parameter for streams and lakes. A body of water that hosts a large number of species with well-balanced numbers of individuals is considered a healthy system. Based on their known tolerance for a given pollutant, certain organisms can be used as indicators of the presence of pollutants.

Pathogens

From the perspective of human use and consumption, the most important biological organisms in water are pathogens, organisms that infect or transmit diseases to humans. These organisms are not native to aquatic systems and usually require an animal host for growth and reproduction. They can, however, be transported by natural water systems, thus becoming temporary members of an aquatic community. Many species of pathogens can survive in water and maintain their infectious capabilities for significant periods of time. These waterborne pathogens include species of bacteria, viruses, protozoa, and helminths (parasitic worms).

Bacteria. Bacteria are single-cell microorganisms, usually colorless, and are the lowest form of life that can synthesize protoplasm from the surrounding environment. Gastrointestinal disorders are common symptoms of most diseases transmitted by waterborne bacteria. Among the most violent waterborne bacterial diseases, cholera causes vomiting and diarrhea that, without treatment, result in dehydration and death.

Viruses. Viruses are the smallest biological structures known to contain all the genetic information necessary for their own reproduction. So small that they can only be "seen" with the aid of an electron microscope, viruses are obligate parasites that require a host in which to live. Symptoms of waterborne viral infections usually involve disorders of the nervous systems rather than of the gastrointestinal tract. Waterborne viral pathogens cause poliomyelitis and infectious hepatitis.

Protozoa. The lowest form of animal life, protozoa are unicellular organisms more complex in their functional activity than bacteria or viruses. They are complete, self-contained organisms that can be free-living or parasitic, pathogenic or nonpathogenic, microscopic or macroscopic. Highly adaptable, protozoa are widely distributed in natural waters, although only a few aquatic protozoa are pathogenic. Protozoal infections are usually characterized by gastrointestinal disorders of a milder order than those from bacterial infections.

Helminths. The life cycles of helminths, or parasitic worms, often involve two or more animal hosts, one of which can be human or animal waste that contains helminths. Contamination may also be via aquatic species of other hosts, such as snails or insects. Aquatic systems can be the vehicle for transmitting helminthal pathogens, but modern water treatment methods are very effective in destroying these organisms. Thus, helminths pose hazards primarily to those persons who came into direct contact with untreated water.

READING LIST

Metcalf & Eddy, Inc. (2003). *Wastewater Engineering: Treatment, Disposal, Reuse*, 4th Edn. McGraw-Hill, New York, pp. 27–139.

Peavy, H.S, Rowe, D.R., and Tchbanoglous, G. (1985). *Environmental Engineering*, International Edn. McGraw-Hill, New York, pp. 11–57.

ANAEROBIC WASTEWATER TREATMENT

Look Hulshoff Pol
Lettinga Associates Foundation
Wageningen, The Netherlands

Anaerobic digestion is a natural process in which different microorganisms of the biological kingdoms of Bacteria and Archaea work together to convert organic compounds through a variety of intermediates into biogas, a mixture of methane and carbon dioxide and small amounts of hydrogen sulfide and hydrogen. This ancient process, brought about by living species long before the presence of oxygen in the atmosphere, is presently gaining an increased interest because of its potential in the treatment of solid organic waste, sludge, and wastewater.

HISTORY OF THE APPLICATION OF ANAEROBIC PROCESSES

Volta is considered to be the first to realize that there was a relationship between decaying vegetation and the occurrence of inflammable gas. In 1776 he showed that "combustible air" was formed from sediments in lakes, ponds, and streams (1). The initial use of anaerobic fermentation for pollution prevention has been for the treatment of domestic wastewater using anaerobic filters and hybrid systems, the latter consisting of a combination of an anaerobic tank and an anaerobic filter. The first recorded anaerobic treatment process was an air-tight chamber called the Mouras Automatic Scavenger, which was developed in the 1860s in France (2). Based on this concept Cameron developed a kind of septic tank in England to treat the wastewater of the city of Exeter in 1895, and Talbot developed a similar tank with baffles in Illinois in the United States in 1894.

In the first half of the twentieth century, anaerobic processes were especially applied for the digestion of sewage sludge. During the 1920s and 1930s, interest in the utilization of methane generated in sludge digesters grew, especially in Germany. Gas was used for heating of digesters, and in 1927 the Ruhrverband in Germany started to use sludge gas to generate power for a biological treatment plant. The generated (waste) heat was used for heating the digester. This is now common practice throughout the world.

The development of anaerobic treatment of industrial wastewater started in the second half of the twentieth century, also thanks to extensive studies in the first half of the century by Buswell. To reduce the size of the treatment systems, in the 1950s in South Africa, Stander (3) developed the clarigester, a modified Imhoff tank with an internal settler on top. In the same period, Schroepfer (4) used a reactor with recirculation of sludge from the settling tank, similar to aerobic treatment. This process was called the anaerobic contact process.

The development of anaerobic industrial wastewater treatment got a big boost after the energy crisis in the early 1970s, as anaerobic treatment is more energy efficient than aerobic treatment. A major factor in the development was also the recognition of the importance of sludge retention. New reactor concepts resulted in a further reduction of reactor volumes that could accommodate much higher loading rates than conventional aerobic treatment processes. The most prominent of these new systems is the UASB-process developed by Lettinga in The Netherlands (5).

In the last two decades a renewed interest has emerged in the application of anaerobic pretreatment for domestic wastewater in countries with a warm climate. The anaerobic treatment step is usually followed by an aerobic post-treatment.

THE ANAEROBIC PROCESS

The most important polluting compounds in wastewater are usually organic polymers such as fats, proteins, and carbohydrates. For the degradation of these polymers four main steps can be distinguished in anaerobic fermentation: hydrolysis, acidogenesis, acetogenesis, and methanogenesis (see Fig. 1).

In the hydrolysis step, extracellular complex organic polymers (proteins, carbohydrates, and lipids) are hydrolyzed by exoenzymes into smaller molecules (mono- and oligomers: amino acids, sugars, and long-chain fatty acids), which can pass through bacterial cell membranes for further decomposition. For complex wastewaters, containing a large fraction of polymers, hydrolysis is the rate-limiting factor. During acidogenesis, the mono- and oligomers are acidified, which means that they are converted into simple organic acids, hydrogen, and carbon dioxide. In the third step, called acetogenesis, acetate is formed out of the simple (volatile) acids other than acetate. Under standard conditions, acetogenic conversions are endothermic and cannot take place. Only by immediate consumption of the products of this conversion step (hydrogen and acetate) by the methanogenic Archaea, can this conversion become thermodynamically favorable (exothermic). Finally, in the last step (methanogenesis) biogas is formed, either out of hydrogen and carbon dioxide or out of acetate. Roughly 70% of the biogas is produced via the acetate route. Due to the low growth rates of acetotrophic methanogens (the growth rate $\mu_{\max}$ of *Methanosaeta soehngenii*, which is usually the most predominant methanogen in the anaerobic sludge, is around 0.1 d^{-1}), acetotrophic methanogenesis is a crucial conversion in the total anaerobic digestion process.

If sulfate is present in the wastewater, part of the organic compounds will be degraded by sulfate-reducing bacteria. Sulfate mediates the degradation of organic compounds and will thereby be reduced to hydrogen sulfide, which will partly end up in the biogas.

BIOGAS PRODUCTION

For more concentrated types of wastewater, biogas production forms a welcome source of renewable energy. Theoretically, about 0.5 m^3 biogas can be produced per kilogram of converted COD. The basic principle of anaerobic fermentation is that part of the carbon is completely oxidized to CO_2, whereas the other part is completely reduced to CH_4 with the result that the average oxidation state of the carbon stays the same. N and O will stay completely reduced. Buswell derived an equation to determine the theoretical biogas production for any given organic compound. Assuming that of an organic molecule with the general formula of $C_nH_aO_bN_d$ a fraction x of C goes to CH_4 (oxidation state $C = -4$) and a fraction $(1 - x)$ of C goes to CO_2 (oxidation state $C = +4$) and taking into account that that does not change, it can be derived that the average oxidation state of the organic molecule $(2b + 3d - a)/n$ equals $-4x + 4(1 - x)$. From this the Buswell equation is obtained:

$$\begin{aligned} C_nH_aO_bN_d &+ (n - a/4 - b/2 + 3d/4)H_2O \\ &\rightarrow (n/2 + a/8 - b/4 - 3d/8)CH_4 \\ &\quad + (n/2 - a/8 + b/4 + 3d/8)CO_2 \\ &\quad + dNH_3 \end{aligned}$$

The real amount depends on various factors such as temperature, atmospheric pressure, pH, heat production, biodegradability of the pollutants in the wastewater, amount of COD used for cell maintenance and growth, and sulfate concentration in the wastewater. The CO_2 content is strongly related to the pH: the higher the pH, the lower the partial pressure of CO_2 in the biogas. Sulfate will reduce the CH_4 content of the biogas. Theoretically, sulfate-reducing bacteria will reduce 2 g of COD per gram of sulfate, with the effect that the generated electrons are used for formation of H_2S rather than CH_4. This results in a biogas with lower methane content.

Biogas can be used as a replacement for natural gas in different ways. It can be used directly in burners, in boilers, or in gas engines or fuel cells for electricity generation.

ANAEROBIC TREATMENT TECHNOLOGY

Anaerobic treatment has some clear advantages over conventional aerobic treatment. It has to be noted though that anaerobic treatment still needs further aerobic treatment of the remaining BOD. Most importantly, the aerobic activated sludge process needs energy for aeration (about 100 kWh per 100 kg COD), whereas in anaerobic

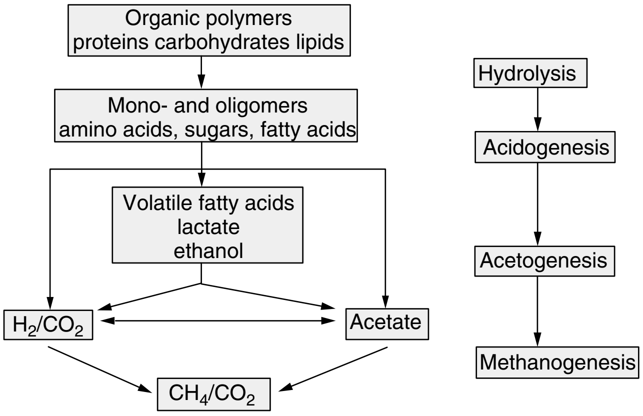

Figure 1. Schematic representation of the anaerobic fermentation of complex organic polymers into biogas.

Table 1. Comparison of Various Characteristics of Aerobic and Anaerobic Degradation of Glucose

Characteristic	Aerobic	Anaerobic
Reaction	$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$	$C_6H_{12}O_6 \rightarrow 3CO_2 + 3CH_4$
Energy release	$\Delta G^0 = -2840$ kJ/mol	$\Delta G^0 = -393$ kJ/mol
Carbon balance	50% → CO_2 50% → biomass	95% → $CH_4 + CO_2$ (biogas) 5% → biomass
Energy balance	60% → biomass 40% → heat	90% retained in CH_4 5% → biomass 5% → heat
Biomass production	Yield > 0.5 kg VSS/kg COD	Yield = 0.05—0.15 kg VSS/kg COD
Energy input for aeration	1 kWh/kg COD	None

treatment energy is produced in the form of biogas. Over 285 kWh of electrical power can be produced per 100 kg COD.

In aerobic treatment much more sludge is produced than in anaerobic treatment. This is caused by the big difference in sludge yield factors (= grams of biomass formed per gram of degraded COD). For aerobic sludge this factor usually is around 0.5, whereas for anaerobic sludge it is generally below 0.15. As sludge management is a major cost factor in wastewater treatment, a strongly reduced sludge production is a big advantage. Contrary, comparing the effluent qualities of aerobic and anaerobic treatment shows that aerobic treatment is superior to anaerobic treatment in COD removal efficiency. Usually anaerobic treatment needs further post-treatment to meet common effluent discharge standards. However, an anaerobic system with appropriate post-treatment, in which the bulk of the organic pollutants is treated in the anaerobic step, is more cost effective than conventional aerobic treatment. The above mentioned difference between aerobic and anaerobic degradation is further illustrated for glucose as a model pollutant in Table 1.

ANAEROBIC REACTOR SYSTEMS

Key to the worldwide interest in anaerobic treatment is that it allows for an extreme uncoupling of the solid retention time from the hydraulic retention time. This uncoupling can be achieved by various means of sludge retention, such as sedimentation, immobilization on a fixed matrix or moving carrier material, and granulation. Granulation is a form of autoimmobilization typical for anaerobic biomass. Under appropriate conditions, dispersed sludge gradually transforms into a granular type of sludge. Granules are usually 0.5–3 mm in diameter and have high specific methanogenic activities and high settleabilities (up to 60 m/h). They are far more resistant to external shear forces than flocs. In practice, uncoupling of the hydraulic retention time from the solids retention time means that small reactor systems can be applied at high volumetric loading rates.

Tables 2, 3, and 4 give an overview of the most common low rate, high rate, and super high rate processes that are presently applied on full scale.

The anaerobic filter (AF) was used in early applications for domestic sewage treatment. It was further developed in the United States (6,7) for the treatment of industrial wastewater treatment. The reactor is filled with a packing material for bacterial attachment. Sludge retention is also achieved by accumulation of anaerobic biomass in the crevices of the packing. This reactor type can be operated both in an upflow and downflow mode.

In the Upflow Anaerobic Sludge Bed (UASB) reactor (5,8) sludge retention is achieved by sedimentation of granular or thick flocculent sludge in an internal settler located in the upper part of the reactor. This internal settler is also called a gas–liquid–solids (GLS) or three-phase separator. In this separator biogas is captured in a gas hood, creating a quiet zone above this hood, allowing solids to settle and return to the lower reactor compartment where the anaerobic conversion takes place. The water leaves the reactor on top of the reactor as effluent. About 60% of all full-scale anaerobic reactors that have been built worldwide are of the UASB type. Alternative sludge bed processes are the Anaerobic Baffled Reactor (ABR) (9,10) and the Anaerobic Migrating Blanket Reactor (AMBR) (11). UASB and AF reactors are used for the treatment of both industrial and domestic wastewater.

Common super high rate processes are the Expanded Granular Sludge Bed (EGSB) reactor (12) and the Internal Circulation (IC) reactor (13) presented in Table 4. These reactor systems can treat at volumetric loading rates up to 30 kg COD/m^3· d. The EGSB is operated with anaerobic granular sludge. EGSB reactors are more widely applied than fluidized bed (FB) reactors. The upflow velocity in EGSB reactors is up to 10 m/h. Like the UASB reactor, the EGSB system has an internal settler on top of the reactor. IC reactors are also operated with anaerobic granular sludge. In this reactor an internal liquid recycling is created by a so-called gas lift of the water by the generated biogas. This internal recycling results in optimal mixing and an intense contact between wastewater and sludge. Contrary to external recirculation, the internal recirculation does not lead to higher average upflow velocities in the system. EGSB and IC reactors are tall and slim and therefore have a small footprint.

Table 2. Most Common Low Rate Anaerobic Treatment Systems (Volumetric Loading Rate <5 kg COD/m^3·d)

Reactor Type	Characteristics
Completely Stirred Tank Reactor (CSTR)	
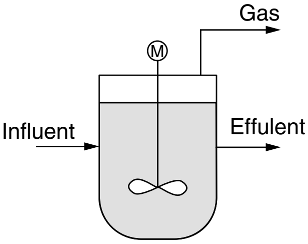	• Suitable for sludge and manure digestion • No sludge retention • Long hydraulic retention times (15–30 days)
Anaerobic Contact Process	
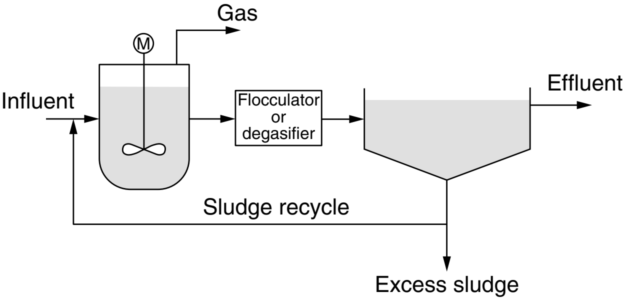	• Suitable for wastewaters rich in suspended solids • Sludge retention by external sedimentation and sludge recycling
Covered Anaerobic Lagoon	
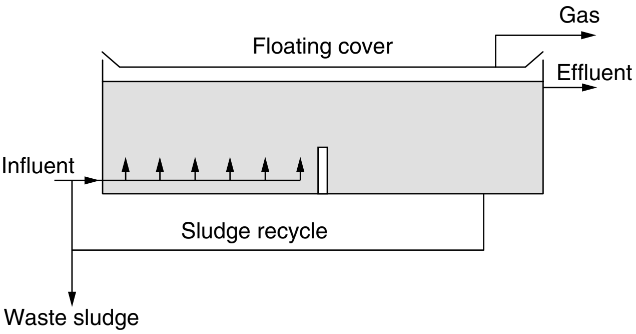	• Suitable for wastewaters rich in suspended solids • Sludge retention by sedimentation • Large reactor with large footprint • Long hydraulic retention times
Anaerobic Sequencing Batch Reactor (ASBR)	
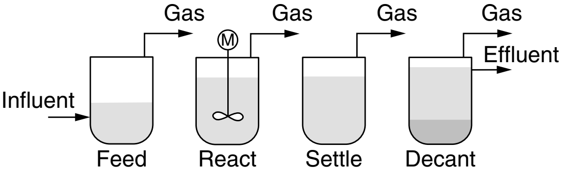	• Suitable for wastewaters rich in suspended solids • Sludge retention by sedimentation

BASIC DESIGN PRINCIPLES

There are three important factors that determine the design of anaerobic reactors:

1. The biological conversion capacity.
2. The maximum allowable hydraulic loading rate.
3. The maximum allowable gas loading rate.

For the design of a system treating concentrated wastewaters the biological conversion capacity is the determining factor. The conversion capacity is determined by the amount of biomass that can be retained in the reactor and the specific methanogenic activity of this biomass. This activity is influenced by various factors such as pH, temperature, intensity of the contact between wastewater and the biomass, COD composition, presence of inhibiting or toxic compounds, and presence of sufficient macro- and micronutrients.

For diluted wastewaters the limiting factor is the hydraulic capacity of the reactor. If the hydraulic load

Table 3. Most Common High Rate Anaerobic Treatment Systems (Volumetric Loading Rate 3–20 kg COD/m^3·d)

Reactor Type	Characteristics
	Anaerobic Filter (AF)
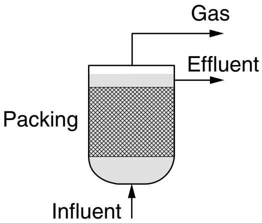	• Sludge retention by attachment on filter material and entrapment in voids in the filter • Can be operated upward and downward
	Upflow Anaerobic Sludge Bed (UASB) Reactor
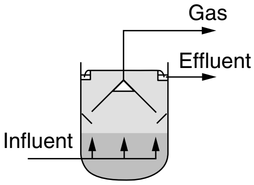	• Sludge retention by granulation and settling • Internal three-phase separator to separate gas, sludge, and water • Optimal contact of water with biomass by gas mixing

Table 4. Most Common Super High Rate Anaerobic Treatment Systems (Volumetric Loading Rates 10–30 kg COD/m^3·d)

Reactor Type	Characteristics
	Expanded Granular Sludge Bed (EGSB) Reactor
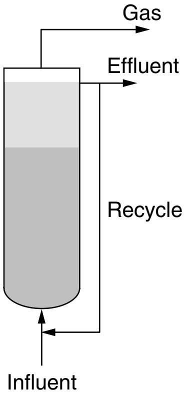	• Suitable for medium to low strength wastewater and low temperatures (<25 °C) • Comparable to fluidized bed reactor • External recirculation • Upflow flow rate in the reactor is up to 10 m/h
	Internal Circulation (IC) Reactor
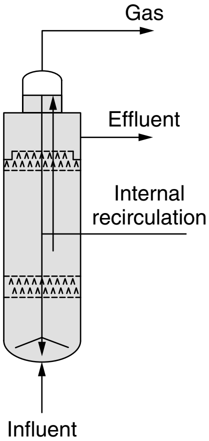	• Requires anaerobic granular sludge • Two internal three-phase separators • Internal circulation of water by gas lift principle

exceeds the hydraulic capacity of the system, the sludge washout from the reactor can become higher than the sludge yield by bacterial growth. This will result in a gradual drop in the quantity of active biomass in the reactor and ultimately in reactor failure.

PHYSICAL SEPARATION OF ACIDOGENESIS FROM METHANOGENESIS

Fairly common in anaerobic treatment is the physical separation of hydrolysis and acidification in an acidifying reactor and acetogenesis and methanogenesis in a second reactor. This phase separation was supported by a school of researchers (14–17), who claimed that such a separation would lead to better control of the overall digestion process. Under optimal mesophilic conditions a slight preacidification of the wastewater is certainly beneficial, but generally this is already accomplished during the transport of the wastewater in the sewer system. For soluble and not or partially acidified carbohydrate-containing wastewaters, phase separation is therefore not required. For the formation of granular sludge it is even better to have no phase separation. However, for treatment of hardly acidified carbohydrates under psychrophilic conditions, the use of a preacidification step is required: the in-growing voluminous acidifying organisms do not decay sufficiently fast, and therefore they will accumulate in the retained methanogenic sludge, creating a bulking type of sludge. As a result, the sludge characteristics will deteriorate (18).

Staging of anaerobic treatment systems can be considered beneficial for the treatment of various types of complex wastewaters, such as domestic sewage or wastewaters containing slowly biodegradable or inhibitory compounds (19). In such staged reactors in principle all phases of the anaerobic degradation process are allowed to proceed to some extent simultaneously in each reactor module. A staged reactor system will provide a higher treatment efficiency, because more difficult compounds like intermediates such as propionate, or possibly even xenobiotic compounds (when present in the wastewater), will find a more optimal environment for degradation due to the development of appropriate microbial communities in each stage. The process stability of a staged system is also substantially higher than in the present commonly practiced one-step systems.

APPLICATION OF ANAEROBIC WASTEWATER TREATMENT

Anaerobic treatment has been applied primarily for medium to high strength wastewaters of (agro-) industrial origin with COD concentrations over 1000 mg/L. Common industries generating wastewater that can be well treated by anaerobic processes are starch factories, breweries, pulp and paper mills, and distilleries. More recently, other industries with wastewater containing organic pollutants have started using anaerobic treatment, such as chemical and pharmaceutical industries.

Since 1985 interest has grown in anaerobic treatment of domestic sewage in countries with tropical or subtropical climates (20). The full-scale plants mostly use UASB reactors for the anaerobic step and they are constructed all over South America and Asia, especially in Brazil and India. As indicated in Fig. 2 the advantage of anaerobic sewage treatment over aerobic treatment is that four components of the latter (primary clarifier, activated

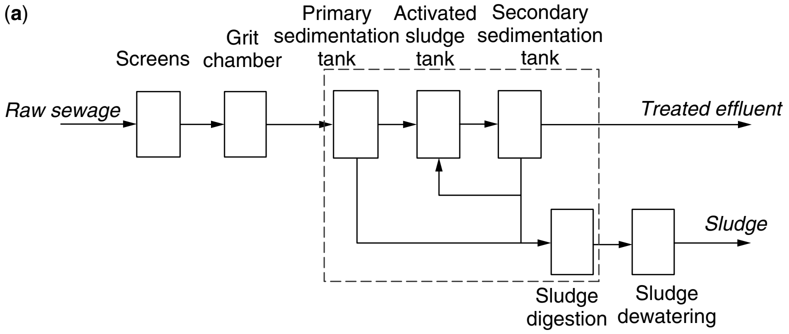

Aerobic sewage treatment

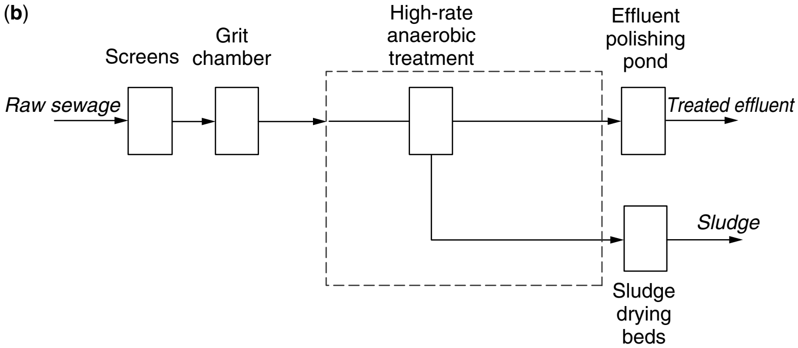

Anaerobic sewage treatment

Figure 2. Flowsheets of aerobic sewage treatment (**a**) using the activated sludge process and anaerobic sewage treatment (**b**) with post-treatment in a polishing pond.

sludge tank, secondary clarifier, and sludge digester) can be combined in one single reactor. In this reactor removal of COD and sludge stabilization takes place simultaneously.

Anaerobic treatment has traditionally been applied under optimal mesophilic temperature conditions (around 30 °C). Anaerobic digestion however, is, feasible under a wide range of conditions, such as temperatures ranging from 4 to 70 °C, high salinity, the presence of recalcitrant or toxic compounds, and for wastewater with very low COD/sulfate ratios.

CONCLUSION

Anaerobic digestion is implemented for the treatment of industrial and domestic wastewater, sludge, and solid waste for the removal of biodegradable compounds from the waste. Anaerobic processes require only a small energy input and a renewable energy source is generated in the form of biogas. Anaerobic treatment is a standard technology for wastewater of agroindustrial origin in the industrialized world. In developing countries, interest in and application of anaerobic treatment of domestic sewage are emerging.

The applicability of anaerobic processes is further extended to hot and cold wastewater as well as wastewater from the (petro-) chemical, pharmaceutical, textile, and mining industries.

BIBLIOGRAPHY

1. Barker, H.A. (1956). Biological formation of methane. In: *Bacterial Fermentations*. Wiley, Hoboken, NJ, p. 1.
2. McCarty, P.L. (2001). The development of anaerobic treatment and its future. In: *Anaerobic Digestion for Sustainable Development, Proceedings of the farewell seminar of Prof. Gatze Lettinga*. March 29–30, Wageningen, The Netherlands.
3. Stander, G.J. (1966). Water pollution research—a key to wastewater management. *J. Water Pollut. Control Fed.* **38**: 774.
4. Schroepfer, G.J., Fullen, W.J., Johnson, A.S., Ziemke, N.R., and Anderson, J.J. (1955). The anaerobic contact process as applied to packinghouse wastes. *Sewage Ind. Wastes.* **2**: 117–142.
5. Lettinga, G. et al. (1980). Use of the Upflow Sludge Blanket (USB) reactor concept for biological wastewater treatment, especially for anaerobic treatment. *Biotechnol. Bioeng.* **22**: 699–734.
6. Young, J.C. and McCarty, P.L. (1969). The anaerobic filter for waste treatment. *J. Water Pollut. Control Fed.* **41**: R160.
7. Young, J.C. and Dahab, M.F. (1983). Effect of media design on the performance of fixed bed anaerobic reactors. *Water Sci. Technol.* **15**: 369–383.
8. Lettinga, G. and Hulshoff Pol, L.W. (1991). UASB process design for various types of wastewaters. *Water Sci. Technol.* **24**: 87–107.
9. Bachmann, A., Beard, V.L., and McCarty, P.L. (1985). Performance characteristics of the Anaerobic Baffled Reactor. *Water Res.* **19**(1): 99–106.
10. Barber, W.P. and Stuckey, D.C. (1999). The use of the Anaerobic Baffled Reactor (ABR) for wastewater treatment: a review. *Water Res.* **33**: 1559–1578.
11. Angenent, L.T., Banik, G.C., and Sung, S. (2000). Psychrophilic anaerobic pretreatment of low-strength wastewater using the Anaerobic Migrating Blanket Reactor. In: *Proceedings of the 73rd Annual Water Environment Conference*, New Orleans, LA.
12. Zoutberg, X. and de Been, X. (1997).
13. Yspeert, P., Vereijken, T., Vellinga, S., and de Vegt, A. (1993). The IC reactor for anaerobic treatment of industrial wastewater. In: *Proceedings of the 1993 Food Industry Environmental Conference*, Atalanta, GA, pp. 487–497.
14. Breure, X. (1986).
15. Breure, X. et al. (1985).
16. Cohen, X. et al. (1985)
17. Pohland, X. and Ghosh, X. (1971).
18. Rebac, X. (1999).
19. van Lier, J.B. et al. (2001). Advances in high-rate treatment: staging of reactor systems. *Water Sci. Technol.* **44**(8): 15–25.
20. van Haandel, A.C. and Lettinga, G. (1994). *Anaerobic Sewage Treatment. A Practical Guide for Regions with a Hot Climate*. John Wiley, Chichester, UK.

SEWERAGE ODORS—HOW TO CONTROL

Bradley A. Striebig
Gonzaga University
Spokane, Washington

WHAT ARE ODORS?

Odors result from the perception of chemicals by the olfactory system. The chemicals are detected in the mouth and nasal passages. Sensations are conveyed to the brain where they may be negatively perceived. The olfactory nerve (first cranial nerve) conveys the perception of chemical odorants to the brain. The trigeminal nerve (fifth cranial nerve) relates the pungency or irritability due to chemical exposure to the brain. Many of the chemicals that cause a negative reaction, bad odors, are related to bacterial emissions that may indicate the presence of pathogenic organisms. The human olfactory system can detect many of these pathogenic indicator chemicals at concentrations of only a few parts per billion (ppb). The chemicals that cause the olfactory response are called odorants. The human olfactory system is capable of detecting a wide variety of odorants.

The human response to odorants present in the air depends on the odor concentration, intensity, persistence, and character. Odor concentration can be measured as a dilution ratio and results may be reported as a detection threshold, recognition threshold, or dilution-to-threshold ratio. Odorants can be measured analytically in the field or laboratory and their concentrations are typically reported in parts per million (ppm) or parts per billion (ppb) by volume. Odor intensity is often related to the equivalent intensity produced by exposing an odor panel to the sample of interest and correlating the odor panel response to a butanol-based standard scale. Odor persistence is interpreted as the duration of exposure and is reported as a dose–response function. The odor character is a descriptive scale, which describes what the odor smells like based on categorical descriptive terms. These terms provide scientists and engineers with the measurement and descriptive techniques relating odorants to odors and hold the key to understanding odor control strategies. Prior to describing control techniques, it is helpful to understand what causes odors and how they are generated in the sewers.

WHAT CAUSES ODORS?

Wastewater or sewage water is a complex mixture of organic and inorganic wastes. The organic wastes consist of a mixture of human wastes, food wastes, and industrial wastes. Simple organic compounds such as sugars and carbohydrates are broken down aerobically into carbon dioxide and water, and anaerobically into methane. More complex organic molecules, such as proteins and amino acids, are also broken down into carbon dioxide, water, and methane; however, the remaining sulfur and nitrogen present in these compounds may result in the production of hydrogen sulfide, ammonia, and other strong odorants as illustrated in Fig. 1.

Aerobic wastewaters commonly contain many odorants, including indole, skatole, organic acids, esters, alcohols, and aldehydes. However, microbial activity in the sewers depletes the oxygen, creating anaerobic conditions. When anaerobic conditions develop, the types and concentrations of odorants in the sewer rapidly increase. Most sewage odor problems are related to the odorants formed under anaerobic conditions, including hydrogen sulfide, mercaptans, ammonia, amines, and volatile fatty acids.

Hydrogen sulfide is typically the odorant of greatest concern in sewers due to its low detection threshold, high concentration, and acidic nature. Hydrogen sulfide is generated from the anaerobic biological reduction of sulfate ($SO_4{}^{2-}$) or thiosulfate. One common mechanism is described by

$$SO_4{}^{2-} + 2C + 2H_2O \xrightarrow{\text{anaerobes}} 2HCO_3{}^{-} + H_2S$$

Hydrogen sulfide is generated in the slime layers and sludge deposits in sewer collection systems as illustrated in Fig. 2. Key parameters that control sulfide generation include the concentration of organic materials and nutrients in the sewage, sulfate concentrations, temperature, dissolved oxygen, pH, and residence time of the sewage and sludge deposits in the sewer system.

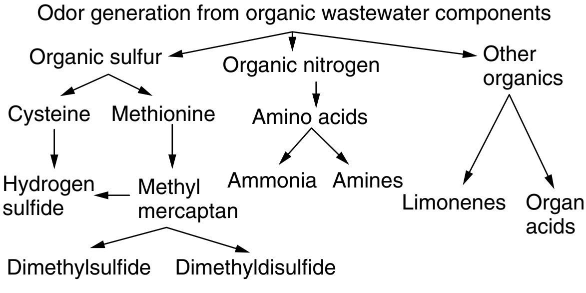

Figure 1. Odorants generation mechanisms from organic wastes.

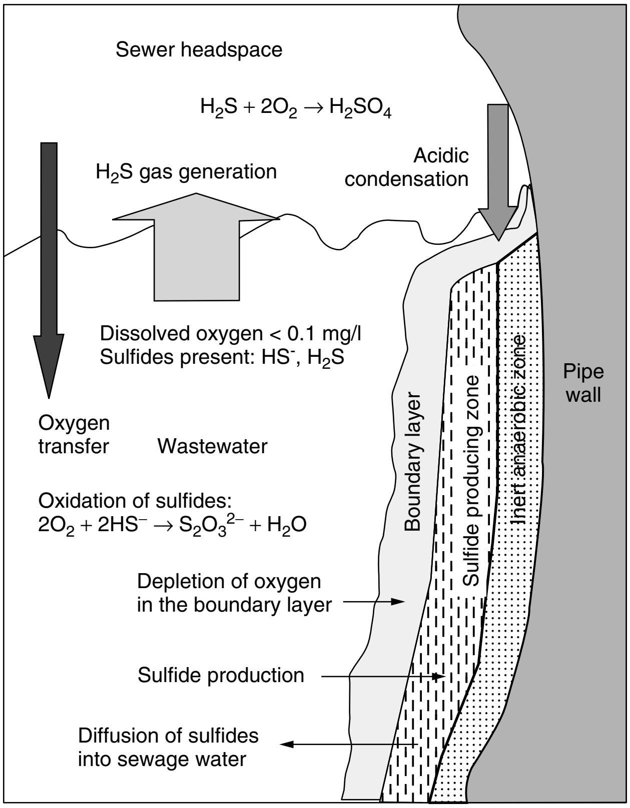

Figure 2. Sulfide generation in sewers (adapted from Ref. 1).

Pomeroy and Parkhurst (1) developed a predictive model for estimating sulfide generation in wastewater collections systems:

For pipes flowing full:

$$S_2 = S_1 + (M)(t)[\text{EBOD}(4/d + 1.57)]$$

For less than full pipes:

$$S_2 = S_{\text{lim}} - \frac{(S_{\text{lim}} - S_1)}{\log^{-1}\left[\dfrac{m(su)^{3/8}t}{2.31d_m}\right]}$$

where S_2 = predicted sulfide concentration at time t_2 (mg/L)
S_1 = sulfide concentration at time t_1 (mg/L)
S_{lim} = limiting sulfide concentration, $S_{\text{lim}} = (M'/m)\ \text{EBOD}(su)^{-3/8}(\text{P/b})$ (mg/L)
M = sulfide flux coefficient, typically 0.5×10^{-3} to 1.0×10^{-3} (m/h)
M' = effective sulfide flux coefficient in gravity sewers, typically 0.32×10^{-3} (m/h)
m = empirical coefficient for sulfide loss, typically 0.96
t = detention time in the sewer reach with constant diameter and flow (h)

EBOD = effective biochemical oxygen demand, EBOD = BOD × $1.07^{(T-20)}$ (mg/L)
d = pipe diameter (m)
T = wastewater temperature (°C)
s = slope (m/m)
u = stream velocity (m/s)
d_m = mean hydraulic depth, equal to area of flow divided by surface width (m)
P = wetted perimeter (m)
b = width of wastewater stream at the surface (m)

Odor emission of other compounds present in the sewage, like hydrogen sulfide, increase with increased turbulence of flow. The solubility of the odorant, concentration, temperature, pH, and mass transfer coefficient are all factors that affect the emission rate of odorant in the sewer. Weirs, junction chambers, and holding tanks may represent significant sources of odor.

Odor emissions are dependent on the chemical nature of the odorant of concern and the composition of the wastewater. Emissions are also dependent on the mechanical operating parameters of the sewer. Odor emissions are of greatest concern when odorants are in close proximity to receptor populations. Nuisance odors are often reported because pressurization of the sewer atmosphere results in odorant emissions from manholes, house vents, and other structures. Pressurized conditions can result from inverted siphons, drop structures, discharges of forces mains, reductions in pipe diameter, and sags in the sewer that restrict air movement.

Odor control measures are designed to prevent the formation of hydrogen sulfide or remove hydrogen sulfide from sewer exhaust. Prevention of hydrogen sulfide formation is usually accomplished by either improving oxygen transfer or the addition of chemical or microbial additives to the sewer.

ODOR PREVENTION

Improving Oxygen Transfer

Ventilation of the sewer system can be improved through proper maintenance or the addition of appurtenances to increase the rate of oxygen transfer. Sewer system ventilation is a dynamic parameter that varies with change in barometric pressure along the sewer, wind velocities past vents, sewage drag, sewage flow rates, and the relative density of sewer and outside air. Partial blockages of flow and buildup of the slime layer within the sewers significantly affect ventilation and increase hydrogen sulfide emissions. These conditions can be mitigated through a regular program of inspections and cleaning. Regular cleaning has been shown to temporarily reduce the rate of sulfide buildup, particularly when sewage deposit buildups are problematic (2).

Oxygen can be transferred mechanically into the sewer to increase the dissolved oxygen concentration in the wastewater and reduce hydrogen sulfide generation. Air or oxygen may be directly injected into the sewer. The rate of air flow required for odor reduction with direct air injection depends on the oxygen uptake rate, detention time in the downstream sewer, temperature and pressure, and degree of odor control required. Direct air injection rates utilized in practice in the United States are highly variable, but typical ranges are 0.75–2.25 m^3 air/m^3 wastewater or 0.7–1.3 m^3/h/cm-pipe-diameter. Other appurtenances used to increase oxygen transfer into the wastewater include venture aspirators, air lift pumps, U-tube injection lines, and pressurized air tanks, which are described in more detail in the referenced literature (1–8). The costs associated with air injection are illustrated in Table 1.

Additives

A variety of chemical additives have been employed to reduce odors from sewers and wastewater treatment facilities. The additives reduce hydrogen sulfide emissions by chemical oxidation, microbial inhibition of sulfate reduction, precipitation, and pH control. The most effective chemical additives for control of sewage odors have been chlorine agents, peroxides, and metal salts.

Chlorine has been used effectively to control hydrogen sulfide emissions from wastewater. Chlorine may be added to wastewater as hypochlorite or as chlorine gas. Hypochlorite solutions may be used for small chlorine dosages, but chlorine gas is more cost effective when 2.3 kg/d of chlorine or more is required. Chlorination systems are designed based on the level of sulfide control required, the characteristics of the wastewater, and the degree of process control required. Sulfide is oxidized to

Table 1. Typical Costs for Increasing Oxygen Transfer in 2003 Dollars[a]

Condition	Small Air Injection System	Large Air Injection System	Small Oxygen Injection System	Large Oxygen Injection System
Flow, m^3/d	3,785	37,850	3,785	37,850
Pipe diameter, cm	25.4	61.0	36	76
Pipe length, m	1,600	1,600	1,600	1,600
Air pressure, kPa	158	158	NA	NA
Air flow, m^3/min	2.5	15.3	NA	NA
Oxygen required, kg/d	NA	NA	93	310
Capital cost, $	37,000	96,000	35,000	89,000
Oxygen cost, $/yr	NA	NA	16,000	35,000

[a] Update by the Consumer Price Index from 1984 dollars (1).

sulfate in wastewater when excess chlorine is present:

$$HS^- + 4Cl_2 + 4H_2O \longrightarrow SO_4{}^- + 9H^+ + 8Cl^-$$

At least 8.87 parts by weight of chlorine are required to oxidize sulfide according to the above equation. However, higher dosages are required in full-scale applications due in part to limitation on the amount of mixing. Field application rates typically required 10–15 kg Cl_2/kg H_2S. Full-scale studies should be performed for a period of several weeks to ensure adequate odor control is achieved based on the control system and chlorine dose. Chlorine feed rates may be optimized through monitoring residual chlorine, sulfide, and wastewater flow rate. Table 2 illustrates the typical costs associated with chlorine injection systems.

Hydrogen peroxide has been utilized to oxidize hydrogen sulfide and reduce odorous emissions from sewage. The hydrogen peroxide oxidation occurs rapidly and excess peroxide results in a higher dissolved oxygen concentration. Hydrogen peroxide is also an attractive reagent because it can be used for gravity and pressurized sewers, feed systems are relatively simple, and peroxide provides effective sulfide control for up to 4 h.

The chemical mechanism of sulfide oxidation is dependent on the pH of the wastewater:

$$\text{For pH} < 8.5: H_2O_2 + H_2S \longrightarrow S + 2H_2O$$

$$\text{For pH} > 8.5: 4H_2O_2 + S^- \longrightarrow SO_4{}^- + 2H_2O$$

The stoichiometric peroxide dose required is 1 g H_2O_2/g H_2S. Dosage varies with the BOD, pH, and temperature of the wastewater and the hydraulic characteristics of the sewer. Typical peroxide dosages in practice range from 0.9 to over 3.0 g H_2O_2/g H_2S. The peroxide dose can be optimized by monitoring wastewater flow rates, sulfide concentrations, and chemical feed rates. The costs associated with peroxide addition systems are shown in Table 2.

Many metal ions react with dissolved sulfides to form insoluble salts that precipitate from the solution. Iron and zinc salts have been added for control of sulfur emission. The proposed mechanism of sulfur removal with iron salts is represented by

$$Fe^{2+} + 2Fe^{3+} + 4HS^- \longrightarrow Fe_3S_4 + 4H^+$$

The overall removal mechanism can be expressed as

$$FeSO_4 + H_2S \longrightarrow FeS + H_2SO_4$$

Stoichiometric removal of sulfur requires 1.6 g Fe/g H_2S removed. Actual dosages require field testing to optimize the iron dose required for the desired odor reduction. The costs associated with iron injection for hydrogen sulfide control are shown in Table 2.

Nitrates, alkalis, ozone, and potassium permanganate have been utilized for short-term odor control measures. These additives have been shown to help control short-term odor problems and spikes in odor at specific locations that require only small doses of the additives. However, these chemicals have not been cost effective for long-term sulfide control in large sewers.

Biological additives have been marketed to help control odor emissions from sewage. These products may alter the metabolism of existing microorganisms or be "designer" microorganisms that out-compete odor-producing microorganisms and thus reduce emissions. Metabolic modification with anthraquinone products has been effective in treating small diameter pipes and force mains where there is a large slime layer or septic sediments are problematic. The product is cost effective when used as a preventative measure for small lines with high sulfide concentrations and slow moving gravity lines. Recommended dosages are based on the surface area of the slime layer and are approximately 0.24 kg of the product per 1000 m^2 of surface area. Nitrate and metabolic modifiers may be a cost-effective combination for minimizing odors.

ODOR REMOVAL

If odor prevention is not cost effective or sufficient to control nuisance odors, various odor control technologies are capable of removing odorous compounds exhausted from confined sources. The most common technologies for

Table 2. Typical Costs for Chemical Additives in 2003 Dollars[a]

Condition	Chlorine Injection System	Hydrogen Peroxide Injection System	Iron ($FeSO_4$) Injection System
Flow, m^3/d	3,785	3,785	3,785
[H_2S], mg/L	5	5	5
Dose, mg/L	30	10	23
Capital cost, $	32,000	44,000	18,000
Chemical cost, $/yr	25,000	37,000	23,000
Flow, m^3/d	37,850	37,850	37,850
[H_2S], mg/L	5	5	5
Dose, mg/L	30	10	23
Capital cost, $	71,000	89,000	23,000
Chemical cost, $/yr	243,000	335,000	230,000

[a] Update by the Consumer Price Index from 1984 dollars (1).

treating sewer-related odors are carbon adsorption, biofiltration, absorption, ozonation, and thermal oxidation.

Four major factors need to be considered to select and design a cost-effective system to remove odorants from exhaust air. The concentration and composition of the odorants in the air stream must be determined from sampling and laboratory analysis. The flow rate of the exhaust air stream must be minimized. The cost associated with the control technologies, both capital and operational costs, must be considered. Selection and design of control technologies should include pilot scale testing to ensure odors are adequately removed. Experience and judgment are inherent and necessary parts of designing an effective odor control system.

Adsorption Systems

Carbon adsorption systems utilize the attractive van der Waals forces on porous granulated activated carbon (GAC) surfaces to capture and contain organic pollutants. GAC systems are commonly applied to a wide variety of odor control situations due the simplicity of design, low capital costs, and minimal maintenance requirements. Carbon adsorption systems are ideal for small applications dictated by space and low capital investment.

The activated carbon acts as a capture and control device. The odorants adhere to the surface of the GAC and eventually consume all available surface sites. The mass-transfer zone (MTZ), where odorants are being removed, gradually moves from the inlet side of the carbon bed to the outlet side. As time passes and more odorants are adsorbed, the amount of spent carbon increases. Breakthrough occurs when the zone of mass transfer reaches the exit of the carbon bed and there is no longer an excess of active carbon sites available. When the carbon adsorption bed efficiency decreases, as the bed nears saturation, the carbon must be replaced or regenerated. The vendor of the material typically regenerates the GAC off-site.

Monitoring of the carbon adsorption system is critical to prevent and detect breakthrough of the odorants as the bed nears the end of its useful life span. Regular monitoring of the system allows accurate measurement of cycle times and service contracts should be created to minimize episodes of poor efficiency. The capital costs, illustrated for two different air flow rates in Table 3, and the operational costs may be considerable.

Biological Control Systems

Biofilters have been used for many years in the United States for odor control and have a reputable history for odor control in many European countries. Odorants are degraded by microorganisms in the biofilter. Odorant removal occurs in a thin liquid film or biofilm. The end products of the degradation are carbon dioxide, water, biomass, and (when removing high concentrations of hydrogen sulfide) sulfuric acid. Because biotrickling filters have the ability to "store" the contaminants as a food source, they are often applied to situations that have variable influent loading rates.

Biofilters represent a low cost alternative for treating low concentration, high volume exhaust streams as shown in Table 3. The capital and operational costs for biofiltration systems are relatively low. In addition, biofilters produce no secondary waste if properly operated, except for periodic media replacement. Historically, biofilters were limited in their application due to variable removal efficiencies caused in part by their intolerance to fluctuation in air flow rates, concentrations, and temperatures. In addition, biofilters typically require a very large area footprint compared to other types of organic control technologies.

Absorption Systems

Absorption systems or scrubbers involve the selective transfer of the odorants from the gas phase to a contacting liquid. The odorants must have preferential solubility in the liquid. The soluble odorants diffuse from the gas through a gas–liquid interface and the odorants are dispersed in the liquid.

Absorption systems are ideal for controlling water-soluble odorants, such as hydrogen sulfide, ammonia, and organic acids. The soluble pollutants must be continually destroyed or treated to maintain operational efficiencies. For both ammonia and hydrogen sulfide removal, the pH of the system must be continuously monitored and controlled. Absorption systems are sometimes combined with chemical additives to increase aqueous phase reaction rates. Absorption systems are often used to remove large quantities of hydrogen sulfide prior to a second stage scrubber or carbon bed used to remove less concentrated odorants such as mercaptans and volatile acids, which without treatment would result in nuisance odors.

Ozone Contactors

Ozone is utilized for odor control in the exhaust air as well as preventing odors from forming in the wastewater. Sufficient time and ozone concentrations are required in the contacting chamber to ensure adequate ozone removal. Detailed ozone reaction mechanisms are complex and

Table 3. Typical Costs for Odor Removal from Confined Space Exhaust Air in 2003 Dollars[a]

Air Flow Rate, m^3/min	Adsorption		Biofiltration		Absorption		Ozone Contactor	Thermal Oxidizer
	Capital	Annual	Capital	Annual	Capital	Annual	Capital	Capital
28	53,000	11,000	17,000	500	69,000	4,000	54,000	56,000
280	227,000	85,000	158,000	5,000	136,000	35,000	120,000	91,000

[a] Update by the Consumer Price Index from 1984 dollars (1).

highly dependent on the chemical composition of the treated wastewater or air stream. Examples of overall reactions of ozone with hydrogen sulfide, amines, and mercaptans are illustrated by

$$H_2S + O_3 \xrightarrow{\text{major pathway}} S + H_2O + O_2$$

$$H_2S + O_3 \xrightarrow{\text{minor pathway}} SO_2 + H_2O$$

$$R_3N + O_3 \longrightarrow \underset{\text{(amine oxide)}}{R_3NO} + O_2$$

$$\underset{\text{Methyl mercaptan}}{CH_3SH} + O_3 \longrightarrow CH_3S - SCH_3$$

$$+ O_3 \longrightarrow \underset{\text{methyl sulfonic acid}}{CH_3SO_3H} + O_2$$

Ozone is an unstable gas that must be produced on-site. Ozone is typically generated by corona discharge, which requires significant electrical consumption and excess heat. Three to four ppm of ozone are required in the exhaust air to sufficiently control odors. Reaction times, typically in the range of 10–60 s, vary widely depending on the degree of control required, odorant concentration, humidity, and ozone concentration. The ozone dosage must be controlled to minimize the discharge of unreacted ozone or excess ozone can be discharged into the wastewater. Occupational and environmental health and safety aspects must be considered due to the potential exposure to ozone and electrical currents. Typical costs for ozone contactors are provided in Table 3.

Thermal Oxidation Systems

Thermal oxidation systems reliably maintain nearly complete destruction of odor-causing compounds. Thermal oxidizers are designed based on operating temperature, residence time, and turbulence or mixing in the reactor. Temperature requirements for destruction range from typical design temperatures of 480 to 870 °C. Thermal oxidation systems can maintain high destruction efficiencies even with wide fluctuations in concentration. However, thermal oxidation systems do not tolerate wide flow rate fluctuations well. Thermal oxidation systems also consume large quantities of fossil fuels and as a result they are sources of nitrogen oxides, carbon dioxide, and possibly acid gases for sulfur-containing waste streams. Furthermore, thermal oxidation systems are typically run 24 h/d due to the long start-up and shut-down times required for operation.

The volumetric flow rates of the fuel and air streams along with the size of the reaction chamber are important to ensure adequate retention time and destruction efficiency. For low concentration air streams, additional make-up air is not necessary if there is sufficient oxygen in the polluted air stream to maintain combustion. A mass balance and enthalpy balance should be performed to estimate the required fuel flow rate to maintain the desired operational temperature.

Fuel costs make up the majority of operational costs and must be considered in order to select the most cost-effective treatment systems. Adequate design and operation of thermal oxidation systems is dependent on the temperature, the residence time of the gas, and turbulence or mixing within the reaction chamber. These variables are dependent on one another. The kinetic rate constants increase exponentially with temperature. Reaction times on the order of 0.1–0.5 s are usually sufficient to allow the reactants to reach the desired degree of chemical destruction. Turbulence within the reaction chamber ensures sufficient mixing. Therefore, a higher reaction temperature results in a shorter residence time, a smaller combustion chamber, and lower capital costs. A higher residence time lowers the operating temperature, results in less fuel usage, and in higher capital costs. The operational costs versus capital investment should be considered during selection and design of thermal oxidation systems. Factors that should be considered when determining the capital costs illustrated in Table 3 include the materials of construction, instrumentation, costs of heat exchangers, engineering fees, and construction fees.

CONCLUSION

Odor control requires a comprehensive maintenance program, preventive measures to reduce odor control costs, and possible implementation of control systems where sewage odor is a nuisance. Research, experience, and comprehensive testing programs are important for determining the most cost-effective odor control methods. Regardless of the technologies implemented, there are ongoing costs for preventing and controlling odors that must be included in the operational budget of the wastewater authority.

BIBLIOGRAPHY

1. US EPA. (1985). *Design Manual, Odor and Corrosion Control in Sanitary Sewerage Systems and Treatment Plants*. U.S. Government Printing Office, Washington, DC.
2. Thistlewayte, D.K.B. (1972). *The Control of Sulphides in Sewerage Systems*. Ann Arbor Sciences and Publishers, Ann Arbor, MI.
3. WEF. (2004). *Control of Odors and Emissions from Wastewater Treatment Plants*. WEF Manual of Practice No. 25, Water Environment Federation, Alexandria, VA, pp. 5–7, 133–146, 239–293.
4. Tchobanoglous, G., Burton, F.L., and Stensel, H.D. (2003). *Wastewater Engineering Treatment and Reuse*, 4th Edn. Metcalf and Eddy/McGraw Hill, Boston, pp. 1650–1693.
5. Pomeroy, R.D. and Bowlus, F.D. (1946). Progress report on sulfide control research. *Sewerage Works J.* **18**(4): 597–640.
6. WPCF. (1976). *Chlorination of Wastewater*. Manual of Practice No. 4, Water Pollution Control Federation, Washington, DC.
7. WPCF. (1969). *Design and Construction of Sanitary and Storm Sewers*. Manual of Practice No. 9, Water Pollution Control Federation, Washington, DC.
8. Duggan, S.W. (1980). *A Perspective on Hydrogen Sulfide in Sewers*. Presented at New Jersey Water Pollution Control Association Technology Transfer Seminar, November 19, 1980.

ULTRAFILTRATION—COMPLEXATION IN WASTEWATER TREATMENT

Raffaele Molinari
Pietro Argurio
Universitá della Calabria
Rende, Italy

GENERAL INTRODUCTION

General

Wastewater treatment represents one of the most important fields of study today in the wide subject of pollution problems solving. A "rational hydrologic resource management" is necessary because of the world's increased demand of water, particularly in these last years, owing to a lack of this resource (1). As a result, the main challenge is to create new resources and to fully reuse the existing ones (2). In particular, water frequently contains numerous ionic solutes, many of which are not desirable, and it is used either for residential or industrial purposes.

In the last fifty years, pressure-driven membrane processes have become a routine technique for the removal of environmentally relevant and hazardous substances from aqueous systems (3–10). A membrane can be defined as a selective barrier between two phases (5,11). Transport through the membrane takes place when a driving force is applied. The main goals of such processes are: a) concentration of a solute by removing the solvent, b) purification of a solution by removing nondesirable components, and c) fractionation of liquid or gaseous mixtures.

The separation of solutes with ionic dimensions can be accomplished by using the reverse osmosis operation, but this will result in high operative costs, low permeate flow rate, and low ions selectivity. In order to overcome these problems, the hybrid ultrafiltration-complexation process was introduced (12–17). It was named polymer-assisted ultrafiltration (PAUF) or polymer-enhanced ultrafiltration (PEUF). This process can be applied for various purposes such as the treatment of waste effluents, groundwater, and seawater (13–15,18,19). The advantages of this method are the low energy requirements involved in UF, the conceptual simplicity, the modularity typical of membrane processes, the high permeate flow rate, the high removal efficiency because of effective binding while reducing the initial waste volume significantly, the selectivity achieved when an appropriate complexing agent is considered, and the optimal quality of treated water (6,16,20–22).

The separation process will be successful if the polymer meets the following requirements: good solubility, high selectivity, regeneration possibility, chemical and mechanical stability, low toxicity, high molecular weight with low viscosity, and low cost (12).

The complexation-ultrafiltration technique is mainly applied today in the separation of metal ions from aqueous solutions, covering processes ranging from production of potable water to leaching and recovery of metals from washing water of contaminated soil or from ores to detoxification of process water and wastewater, also for water recycling and reuse (23–27). Indeed, metal contamination is a dangerous cause of water pollution and it constitutes a big health hazard (28–34). Some metal ions play an essential role in many biological processes, and their deficiency, unusual accumulation or imbalance, may lead to biological troubles, e.g., Cu^{2+} ion is an essential nutrient, but when people are exposed to copper levels of above 1.3 mg/l for short periods of time, stomach and intestinal problems occur. Long-term exposure to Cu^{2+} leads to kidney and liver damage (35,36), producing DNA mutation, evidence of its cancerogeneous character.

In the following, the application of PAUF technique is reported with particular focus on metal removal from water.

Historical Background

In 1980, Nguyen et al. (37) considered the application of ultrafiltration to the concentration and separation of solutes of low molecular weight in water. Their results showed a high rejection of these species by complexing them with a suitable soluble macromolecule. Furthermore, they demonstrated that it is possible to separate a specific cation from a mixture by using a polymer that shows a marked selectivity for one cation.

Another similar approach was published in 1982 by Renault et al. (38), which studied the recovery of chromium from effluents by using ultrafiltration.

In 1984, Buffle and Staub (39) gave a fundamental contribution to increase research interest on PAUF method with a work in which they applied the ultrafiltration for measurement of complexation equilibrium constants of metal ions in water in natural conditions. Thus, the coupling of the terms complexation and ultrafiltration was used.

In successive years, several research efforts were made to study mainly the technical and economical feasibility of PAUF to meet the limits fixed by pollution laws for metal removal from water of various origins.

Tabatabai et al. (17) studied the feasibility in the removal of Ca^{2+} and Mg^{2+} ions from hard water by using the sodium polystyrene sulfonate (PSS) polymer (water softening). They demonstrated (with some economical considerations) that the PSS needs to be recovered from the retentate and regenerated appropriately to be reused.

Juang and Shiau (20) studied the metal removal from aqueous solutions using chitosan-enhanced membrane filtration, and in two other works (2,40), Juang and Chiou considered the problem of technical feasibility on the use of PAUF for brackish water softening and wastewater treatment by using three weakly basic, water-soluble polymers like chitosan, polyethylenimine (PEI), and poly(diallyl dimethylammonium chloride) to remove ions like Ca^{2+}, Mg^{2+}, Na^+, K^+, Cu^{2+}, and Zn^{2+}.

An et al. (28) evaluated the ability of crab shell (practically chitosan) to remove heavy metals from aqueous solutions by comparison with several sorbents. They found, for crab shell heavy metals, removal capacity higher than cation exchange resins, zeolite, powered activated carbon, and granular activated carbon. Besides, this process is selective, removing Pb and Cr in preference to Cd and Cu.

Barron-Zambrano et al. (23) investigated the separation of mercury from aqueous solutions using PEI as polymeric complexing agent. They considered the development of a two-stage process: The first one enables mercury concentration and the production of a purified stream, whereas the second one was required to separate the mercury from the polymer and to recycle the polymer (chemical regeneration).

Steenkamp et al. (35) considered the copper(II) removal from polluted water with alumina/chitosan composite membrane, giving attention prevalently to the problems related to the synthesis of the composite support and to the factors that influence metal removal efficiency, like pore radii variation with temperature and powder mixtures used and chitosan coating thickness.

Vieira et al. (41) studied the metal removal from wastewater of the pulp and paper industry. Zakrzewska-Trznadel et al. (42) tested the application of PAUF for radioactive waste processing purposes. Their results showed that this process could be an alternative to reverse osmosis.

Canizares et al. (18) studied a semicontinuous laboratory-scale application of polymer-enhanced ultrafiltration for the recovery of lead and cadmium from aqueous effluents. They emphasized that this process includes two stages: a) metal retention, where a permeate stream free of heavy metals is obtained and b) polymer regeneration, where the polymer is regenerated in order to be reused. The pH for metal retention and polymer regeneration and the binding capacity to know the metal amount that can be treated are also important parameters.

FUNDAMENTALS OF POLYMER-ASSISTED ULTRAFILTRATION

The idea of the PAUF process is that ultrafiltration can be used for removal of ions from aqueous streams, provided that they are preliminarily bound to water-soluble polymers (2,12,43,44). The unbound ions pass through the membrane, whereas the polymers and their complexes are retained (13,14,42,45).

Low-molecular-weight species such as metal ions can be bound to macromolecules by intermolecular forces, mainly ionic interaction and complex binding, or the combination of both.

Formation of complexes is significantly more selective than ionic interactions. An example of this binding mechanism is the complexation reaction among the polymeric agent (PEI), the proton (H^+), and the metal cation (Cu^{2+}), which is represented by the equilibrium equations:

$$\mathrm{PEI} + \mathrm{n\ H_2O} \Longleftrightarrow \mathrm{PEIH_n^{n+}} + \mathrm{n\ OH^-} \tag{1}$$

$$\mathrm{PEI} + \mathrm{a\ Cu^{2+}} \Longleftrightarrow \mathrm{PEICu_a^{2a+}} \tag{2}$$

where $0 \leq \mathrm{n} \leq \overline{\mathrm{n}}$ and $0 \leq \mathrm{a} \leq \overline{\mathrm{a}}$ with $\overline{\mathrm{n}}$ equal to the number of monomers contained in a single polymeric chain and $\overline{\mathrm{a}}$ representing the maximum complexation ratio of the polymers with copper ions ($\overline{\mathrm{a}} = \frac{\overline{\mathrm{n}}}{4}$ for PEI-Cu complex because of Cu^{2+} tetra-coordination with four nitrogen of PEI). In particular, considering that commercial PEI, widely used in literature, has a polymeric chain of MW 60 kDa and considering that a monomeric unit $-CH_2-CH_2-NH-$has MW 43.062 Da, $\overline{n} = 1393$ is obtained (30).

An ionic interaction mechanism can be described, for example, in the removal of dihydrogen arseniate ion with a polymer R^+Cl^-, by the following reactions:

$$\mathrm{R^+Cl^-} \Longleftrightarrow \mathrm{R^+} + \mathrm{Cl^-} \tag{3}$$

$$\mathrm{R^+} + \mathrm{H_2AsO_4{}^-} \Longleftrightarrow \mathrm{RH_2AsO_4} \tag{4}$$

This is an ion-exchange reaction, similar to that which takes place in anion-exchange resin. In another case, the water-soluble polymers polyacrylic acid (PAA) and polyacrylic acid sodium salt (PAASS) interact with copper cation by ion-exchange reactions, described by the following equations:

$$\mathrm{PAA} + \mathrm{n\ Cu^{2+}} \Longleftrightarrow \mathrm{PAACu_n} + \mathrm{2n\ H^+} \tag{5}$$

$$\mathrm{PAASS} + \mathrm{m\ Cu^{2+}} \Longleftrightarrow \mathrm{PAACu_m} + \mathrm{2m\ Na^+} \tag{6}$$

Ionic exchange interaction mechanism has low selectivity and the disadvantage to release another ion (H^+ or Na^+ in this case) in the feed solution, so that to remove an ion from an aqueous solution, another one must be released. In contrast, reactions such as in Eq. 2 do not present this disadvantage.

It should be taken into account that, in general, solid resins and water-soluble polymers have similar functional groups, which would result in similar chemical properties, e.g., the ability to bind certain ions. These analogies in the properties can be used in order to predict the behavior of an unknown hydrophilic polymer if the properties of the functional group of the resin are known.

The PAUF process can be economically more feasible if the polymer could be regenerated, releasing the metal to separate, and be reused. The general scheme of the overall process is represented in Fig. 1.

Polymer regeneration could be carried out by three major methods (12):

1. *Chemical regeneration* means the change of pH of the retentate in order to cleave the polymer–metal bond (21,30,46);
2. *Electrochemical regeneration* means the electrolysis of the retentate resulting in a deposition of the metal on an electrode, whereas the polymeric agent remains in the solution (44);

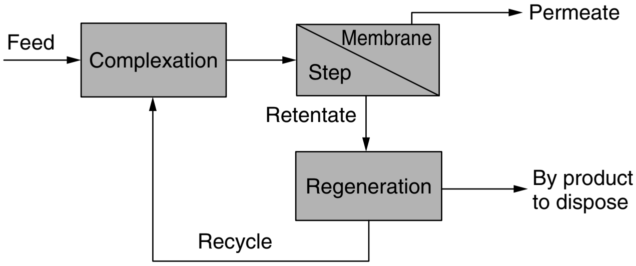

Figure 1. Flowsheet of PAUF separation process.

3. *Thermal regeneration* could be also possible if the polymer-metal bond can be cleaved by heating the retentate, but it has found no practical application so far.

SOME STUDIES OF PAUF IN THE REMOVAL OF METAL IONS FROM AQUEOUS SYSTEMS

In the following, attention will be focused on the application of the complexation-ultrafiltration process in the removal of the model ion copper(II) from aqueous systems, using water-soluble polymers as chelating agents.

Copper(II) Removal from Water by Using Polyethylenimine (30)

The mechanism of the water-soluble polymer PEI—copper interaction can be described by the previously reported equilibrium reactions (1) and (2). In those reactions, a competition exists between Cu^{2+} and H^+ for the polymer because, depending on pH, the PEI is able to complex copper ions by means of Eq. (2), but, at low pH, it stays in aqueous solution in the $PEIH_n^{n+}$ form incapable to interact with copper.

Thus, the first step in the application of the complexation ultrafiltration technique consists in the determination of optimal chemical conditions (pH) for copper complexation (bound) and de-complexation (release) at isothermal conditions (e.g., temperature of 25 °C). To quantify the copper-polyethylenimine (Cu-PEI) complex formation, the spectrophotometric technique was used by reading at 620 nm wavelength.

The complexation-decomplexation process was quantified by plotting vs. the pH the complexation percentage $C\% = (ABS/ABS_{max}) \times 100$, where ABS_{max} is the maximum value of the absorbance that corresponds the maximum amount of complex (100%). The results, reported in Fig. 2, show that PEI is able to complex copper ion at pH 6 or higher, whereas the decomplexation happens at pH < 3.

Similar results were obtained by working at different polymer concentrations observing that maximum binding pH does not depend on polymer concentration. This behavior agrees with the chemical mechanism of polymer-copper interactions (Eqs. 1–2). Indeed, at high pH, the complexation reaction (2) takes place.

To determine the binding capacity of PEI (maximum copper amount (mg) that can be complexed by a fixed amount (1 g) of polymer), some complexation tests were carried out with a polymer concentration of 150 mg/l and changing copper concentration at a fixed pH (~6). The obtained value of the binding capacity was 0.333 mg Cu^{2+}/mg PEI. It must be observed that a high-binding capacity means a lower cost of the PAUF process.

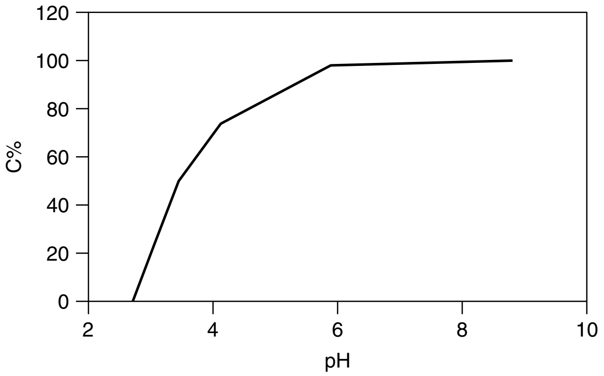

Figure 2. Cu-PEI complex formation $C\%(= (ABS/ABS_{max}) \times 100)$ vs pH in complexation tests of PEI (150 mg/l) with copper (50 mg/l).

Ultrafiltration tests were carried out by using five different membranes (Table 1), two operative transmembrane pressures (2 and 4 bar), pH ca. 6, and five different weight concentrations of PEI and Cu^{2+} with the same ratio (150/50, 270/90, 375/125, 480/160, 600/200).

Working at increasing PEI and Cu^{2+} concentrations and maintaining the same ratio, it is permitted to simulate the increase of retentate concentration in a hypothetical industrial plant where the permeate, free of metals, is withdrawn using the PAUF technique.

It was obtained that simultaneously increasing copper and polymer concentrations (ratio $PEI/Cu^{2+} = 3$ fixed) in the retentate, the separation efficiency (R%) decreased, resulting in an increase of copper and polymer concentrations in the permeate and a little decrease of permeate flux. As a result of increased concentration in the retentate, rejection first decreased, but then increased, because of the formation of a selective dynamic layer (by concentration polarization), which caused a little decrease of permeate flux too, because of mass transfer resistance increase.

It should be observed that an optimal PAUF process should produce a high permeate flux (J_P) and a low copper concentration (C_p). So, in order to compare membrane performances, an appropriate parameter Jp/Cp was introduced. This parameter has no dimensional significance, but it answers the previous requirements to optimize PAUF processes by choosing the membrane that gives the highest ratio.

To evaluate the possibility of polymer regeneration, some UF tests were carried out (operative conditions:

Table 1. Some Characteristics of the UF Membranes Tested in the PAUF Process

Membrane Type	Material	Cut-off (kDa)	Producer	Water Flux [l/h × m^2] (2–4 bar)
Iris 10	Poly ether sulphone (PES)	10	Tech-Sep	33.85–55.00
FS 40 PP	Fluoride-polypropylene	40	Dow	220.0–397.7
GR 40 PP	Polysulphone-polypropylene	40	Dow	220.0–444.3
Iris 30	Poly ether sulphone (PES)	30	Tech-Sep	114.2–207.3
PAN 40	Polyacrylonitrile	40	Tech-Sep	291.1–528.9

PEI = 150 ppm; Cu = 50 ppm; pH = 3) withdrawing the permeate at established time and analyzing copper and TOC (Total Organic Carbon) concentrations. Obtained data showed that all the copper passed through the membrane, whereas the polymer remained in the retentate (rejection of 95% with PAN 40 kDa membrane); that means a good possibility of polymer regeneration, recovery, and reuse.

Comparison of Copper(II) Removal from Waters by Using Various Polymers (21)

Some water-soluble polymers, such as polyetilenimine (PEI), polyacrylic acid (PAA), polyacrilic acid sodium salt (PAASS), and poly(dimethylamine–co–epichlorohydrin–co–ethylenediamine) (PDEHED) as chelating agents (Table 2), have been tested by using the Cu^{2+} as model ion.

Optimal Chemical Conditions. For PAA and PAASS, an ionic interactions mechanism, such as the equilibrium reactions (5) and (6), occurs. It is influenced by the pH: at low pH, the protonation of carboxylic group of the polymer is favored, whereas at higher pH complexation, the reaction is shifted right, meaning macromolecular complex formation with the metal ion. The interaction between the copolymer PDEHED and copper ion is more complicated: both the lone pair of the nitrogen in dimethylamine and ethylenediamine monomers and the oxygen of epichlorohydrin could bind copper with both complex bond and ionic interactions.

From Fig. 3 it can be observed that copper ion is complexed by PEI, PAA or PAASS, and PDEHED at pHs higher than 6, 4.6, and 8, respectively. The decomplexation reactions took place at pH < 3.

Binding capacity was 0.333 mg Cu^{2+}/mg polymer for PEI and PAA. At pH 8.5, working with a polymer concentration of 50 mg/l and changing copper concentration, a binding capacity of 2 mg Cu^{2+}/mg PDEHED was obtained; at higher ratios, a cloud solution was observed, probably because of limited complex solubility.

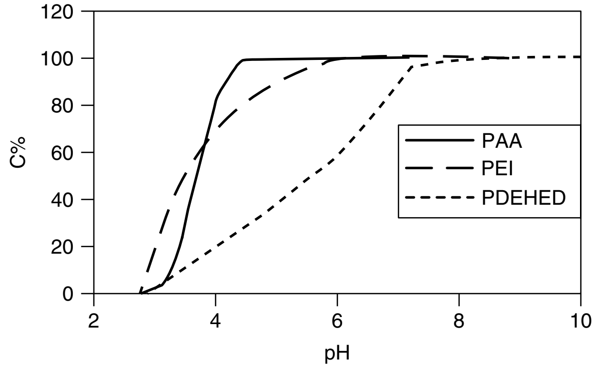

Figure 3. Comparison of Cu-polymer complex formation C% (= $(ABS/ABS_{max}) \times 100$) vs pH in complexation tests of PEI (150 mg/l), PAA (150 mg/l), and PDEHED (150 mg/l) with copper (50 mg/l).

Ultrafiltration Tests. Ultrafiltration tests were realized at two transmembrane pressures (2 and 4 bar), by using five different membranes, and determining flux and rejection. Copper concentration was fixed at 50 mg/l to simulate the treatment of the same pollution load.

The results, summarized in Table 3, show that the fluxes (J_P) obtained with the PDEHED are lower than that registered using PEI and PAA. Regarding the separation efficiency, measured by R%, this is in the order PDEHED > PAA > PEI. This behavior could be caused by the higher membrane fouling and/or polarization concentration caused by the copolymer.

Inspection of membranes at the end of the experimental runs showed a thin layer on the filtering surface: the cake was cerulean with the color of polymer-copper complexes, and it appeared like an incrustation in the case of PDEHED, whereas for PAA and PEI it was a simpler deposit easily removable.

The polymer PDEHED is useful if the objective of wastewater treatment is to obtain a complete copper removal. In opposition, the polymer PAA with the

Table 2. Some Polymeric Binding Agents

Polymer	General Formula	Average Molecular Weight (kDa)
Polyacrylic acid (PAA)	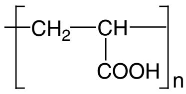 $\left[CH_2-\underset{COOH}{CH} \right]_n$	100
Polyethylenimine (PEI)	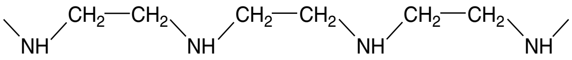	60
Polyacrylic acid, sodium salt (PAASS)	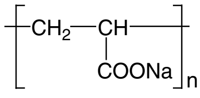 $\left[CH_2-\underset{COONa}{CH} \right]_n$	30
Poly(dimethylamine-co-epichlorohydrin-co-ethylenediamine) (PDEHED)	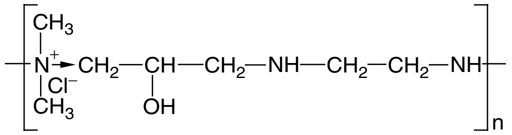 $\left[\overset{CH_3}{\underset{CH_3}{N^+}}\,Cl^- \rightarrow CH_2-\underset{OH}{CH}-CH_2-NH-CH_2-CH_2-NH \right]_n$	75

Table 3. Results of Ultrafiltration Tests Using: PEI (150 mg/l) and Cu^{++} (50 mg/l) at pH = 6.2; PAA (150 mg/l) and Cu^{++} (50 mg/l) at pH 4.6; PDEHED (50 mg/l) and Cu^{++} (50 mg/l) at pH = 8.5

	Membrane	Iris 10 kDA		Iris 30 kDa		PANGKSS HV3/T		PANGKSS HV2/T		UTC 60 ROPUR	
Polymer	Pressure [bar]	J_p[l/hm^2]	R%	J_p[l/hm^2]	R%	J_p[l/hm^2]	R%	J_p[l/hm^2]	R%	J_p[l/hm^2]	R%
PDEHED	2	93.08	100.0	110.0	99.9	122.7	99.9	131.2	99.9	50.8	100.0
	4	55.0	99.9	67.7	99.9	71.9	99.8	71.9	99.8	50.8	100.0
PEI	2	120.58	99.0	156.5	98.7	241.2	99.1	253.9	99.1	29.6	99.9
	4	173.47	98.9	232.7	98.5	262.3	98.7	262.3	98.6	63.5	99.9
PAA	2	105.78	99.1	156.5	98.3	258.1	99.5	275.0	99.6	25.4	100.0
	4	143.85	99.5	190.4	99.3	220.0	99.5	215.8	99.5	50.8	99.9

membrane PAN GKSS HV2/T can be used when very low metal concentration is not required because of the higher flux (rejection 99.6% and permeate flux 275 l/h·m^2 at 2 bar). Indeed, it is better operating at transmembrane pressure of 2 bar rather than at 4 bar because the little increase of the permeate flux obtained at steady state for PEI and PAA polymers does not justify the higher costs (e.g., electrical energy and cooling).

Data of the optimization parameter J_p/C_p for the transmembrane pressure of 2 bar, for the PEI and PAA polymers, showed that PAN GKSS membranes gave the best combination of the two parameters. Furthermore, a higher J_p/C_p for PAA was registered, meaning more interesting performances in copper removal from waters. The optimization parameter for PDEHED has no practical significance, in this case, because of copper concentration next to the zero.

Membrane Washing and Reuse. The possibility of membrane reuse in the complexation-ultrafiltration process was evaluated by carrying out three UF runs in series by using the polymer PDEHED, which gave the highest fouling. Each run was composed by four steps in sequence:

1. membranes characterization;
2. UF test carried out until reaching steady-state conditions;
3. washing of membranes and system with tap water for 2 hours without recycle (open loop);
4. washing of membranes and system with 20 l of demineralized water without recycle; and
5. return to (1).

Steps 3 and 4 were carried out at maximum cross flow and minimum transmembrane pressure to avoid further cake compaction during membrane washing.

The results showed that, after a flux decrease was observed in the second run, membrane performance remained the same in the third run, which is interesting for a long-time use of the same membrane.

CONCLUSION

The experimental work available in the literature on the PAUF process show that satisfactory results were obtained by applying it in the separation and concentration of metallic cations from wastewaters. This technique combines both the advantages of classical adsorption (i.e., ion exchange and complexation interactions) method for metal removal from aqueous systems and of membrane processes.

The chemical fundamentals of the process have to be preliminarly studied in order to find the optimal chemical conditions of: 1) pH for the metal retention stage (complexation) and for the polymer regeneration stage (decomplexation) and 2) polymer-binding capacity (loading ratio = g metal/g polymer).

These results have to be transferred in the realization of the two stages of metal separation and recovery process: (1) metal retention, where a permeate stream free of heavy metals can be obtained, and (2) polymer regeneration, where the polymer is regenerated in order to be recycled.

Several factors influence the separation of the target substance, such as membrane type, composition of water to treat, pH, binding capacity of the polymer, polymer adsorption on the membrane (fouling), and hydrodynamics.

The complexation-ultrafiltration technique could be competitive in the near future provided a significant knowledge on the main process parameters are realized: (1) design and preparation of polymeric binding agents with the desired properties (good solubility, high selectivity, regeneration possibility, chemical and mechanical stability, low toxicity, high molecular weight with low viscosity, and low cost); (2) proper membrane choice; (3) accurate approach to the fluid dynamics and to the chemistry of the process; (4) appropriate study of polymer regeneration; and (5) appropriate study on membrane washing and reuse.

The complexation-ultrafiltration process is a relatively new separation technique, but the results reported in the literature clearly indicate its potential in wastewater treatment with some reasonable technological improvements.

BIBLIOGRAPHY

1. Asano, T. and Levine, D. (1996). Wastewater reclamation, recycling and reuse: past, present and future. *Wat. Sci. Technol.* **33**: 1–14.

2. Juang, R.S. and Chiou, C.H. (2001). Feasibility of the use of polymer-assisted membrane filtration for brackish water softening. *J. Membr. Sci.* **187**: 119–127.
3. Scott, K. (1995). *Handbook of Industrial Membranes*. Elsevier Advanced Technology, Oxford, UK.
4. Aptel, P. and Vial, D. (1992). In: *Membranes in Water Treatment and Potabilization*. C. Haber and E. Drioli (Eds.). Proc. Membrane Separation Processes, Rio de Janeiro, Brazil, p. 291.
5. Mulder, M. (1996). *Basic Principles of Membrane Technology*, 2nd Edn. Kluwer Academy Publishers, Dordrecht, the Netherlands.
6. Uludag, Y., Ozbelge, H.O., and Yilmaz, L. (1997). Removal of mercury from aqueous solutions via polymer enhanced ultrafiltration. *J. Membr. Sci.* **129**: 93–99.
7. Molinari, R., Argurio, P., and Romeo, L. (2001). Studies on interactions between membranes (RO and NF) and pollutants (SiO_2, $NO_3{}^-$, Mn^{++} and Humic Acid) in water. *Desalination* **138**: 271–281.
8. Oussedik, S. et al. (2000). Enhanced ultrafiltration of bovine serum albumin with pulsed electric field and fluidized activated alumina. *Desalination* **127**: 59–68.
9. Huotari, H.M., Huisman, I.H., and Trägårdh, G. (1999). Electrically enhanced crossflow membrane filtration of oily waste water using the membrane as cathode. *J. Membr. Sci.* **156**: 49–60.
10. Toyomoto, K. and Higuchi, A. (1992). Microfiltration and ultrafiltration, In: *Membrane Science and Technology*. Y. Osada and T. Nakagawa (Eds.). Marcel Dekker, New York, Chap. 8, p. 289.
11. Hwang, S.T. and Kammermeyer, K. (1975). *Membranes in Separations*. Wiley Interscience, New York.
12. Geckler, K.E. and Volchek, K. (1996). Removal of hazardous substances from water using ultrafiltration in conjunction with soluble polymers. *Env. Sci. Technol.* **30**: 725–734.
13. Volchek, K., Krantsel, E., Zhilin, Yu., Shtereva, G., and Dynersky, Yu. (1993). Polymer binding/ultrafiltration as a method for concentration and separation of metals. *J. Membr. Sci.* **79**: 253–272.
14. Rumeau, M., Persin, F., Sciers, V., Persin, M., and Sarrazin, J. (1992). Separation by coupling ultrafiltration and complexation of metallic species with industrial water soluble polymers. Application for removal or concentration of metallic cations. *J. Membr. Sci.* **73**: 313–322.
15. Volchek, K., Keller, L., Velicogna, D., and Whittaker, H. (1993). Selective removal of metal ions from groundwater by polymeric binding and microfiltration. *J. Membr. Sci.* **89**: 247–262.
16. Tabatai, A., Scamehorn, J.F., and Christian, S.D. (1995). Water softening using polyelectrolyte-enhanced ultrafiltration. *Sep. Sci. Technol.* **30**: 211–224.
17. Tabatabai, A., Scamehorn, J.F., and Christian, S.D. (1995). Economic feasibility study of polyelectrolyte-enhanced ultrafiltration (PEUF) for water softening. *J. Membr. Sci.* **100**: 193–207.
18. Canizares, P., Perez, A., Camarillo, R., and Linares, J.J. (2004). A semi-continuous laboratory-scale polymer enhanced ultrafiltration process for the recovery of cadmium and lead from aqueous effluents. *J. Membr. Sci.* **240**: 197–209.
19. Bodzek, M., Korus, I., and Loska, K. (1999). Application of the hybrid complexation-ultrafiltration process for the removal of metal ions from galvanic wastewater. *Desalination* **121**: 177–121.
20. Juang, R.S. and Shiau, R.C. (2000). Metal removal from aqueous solutions using chitosan-enhanced membrane filtration. *J. Membr. Sci.* **165**: 159–167.
21. Molinari, R., Argurio, P., and Poerio, T. (2004). Comparison of polyethylenimine, polyacrylic acid and poly(dimethylamine–co–epichlorohydrin–co–ethylenediamine) in Cu^{2+} removal from wastewaters by polymer–assisted ultrafiltration. *Desalination* **162**: 217–228.
22. Bisset, W., Jacobs, H., Koshti, N., Stark, P., and Gopalan, A. (2003). Synthesis and metal ion complexation properties of a novel polyethyleneimine N-methylhydroxamic acid water soluble polymer. *Reactive Functional Polymers* **55**: 109–119.
23. Barron-Zambrano, J., Laborie, S., Viers, Ph., Rakib, M., and Durand, G. (2002). Mercury removal from aqueous solution by complexation-ultrafiltration. *Desalination* **144**: 201–206.
24. Thompson, J.A. and Gordon, J. (1999). Using water-soluble polymers to remove dissolved metal ions. *Filtration Sep.*: 28–29.
25. Kruithof, J.C. and Kopper, H.M.M. (1989). Experiences with groundwater treatment and disposal of the eliminated substances in the Netherlands. *Aqua.* **38**: 207–216.
26. Bohdziewicz, J. (2000). Removal of chromium ions (VI) from underground water in the hybrid complexation-ultrafiltration process. *Desalination* **129**: 227–235.
27. Alpatova, A., Verbych, S., Bryk, M., Nigmatullin, R., and Hilal, N. (2004). Ultrafiltration of water containing natural organic matter: heavy metal removing in the hybrid complexation-ultrafiltration process. *Sep. Purif. Technol.* **40**: 155–162.
28. An, H.K., Park, B.Y., and Kim, D.S. (2001). Crab shell for the removal of heavy metals from aqueous solution. *Water Research* **35**: 3551–3556.
29. Mulligan, C.N., Yong, R.N., and Gibbs, B.F. (2001). Heavy metal removal from sediments by biosurfactants. *J. Hazardous Materials* **85**: 111–125.
30. Molinari, R., Gallo, S., and Argurio, P. (2004). Metal ions removal from wastewater or washing water from contaminated soil by ultrafiltration–complexation. *Water Research* **38**: 593–600.
31. Mulligan, C.N., Yong, R.N., and Gibbs, B.F. (2001). Heavy metal removal from sediments by biosurfactant. *J. Hazardous Materials* **85**: 111–125.
32. Yurlova, L., Kryvoruchko, A., and Kornilovich, B. (2002). Removal of Ni(II) ions from wastewater by micellar enhanced ultrafiltration. *Desalination* **144**: 255–260.
33. Korus, I., Bodzek, M., and Loska, K. (1999). Removal of zinc and nickel ions from aqueous solutions by means of the hybrid complexation ultrafiltration process. *Sep. Purif. Technol.* **17**: 111–116.
34. Aliane, A., Bounatiro, N., Cherif, A.T., and Akretche, D.E. (2001). Removal of chromium from aqueous solution by complexation-ultrafiltration using a water soluble macroligand. *Water Research* **35**: 2320–2326.
35. Steenkamp, G.C., Keizer, K., Neomagus, H.W.J.P., and Krieg, H.M. (2002). Copper(II) removal from polluted water with alumina/chitosan composite membranes. *J. Membr. Sci.* **197**: 147–156.
36. Petrov, S. and Nenov, V. (2004). Removal and recovery of copper from wastewater by a complexation-ultrafiltration process. *Desalination* **162**: 201–209.

37. Nguyen, Q.T., Aptel, P., and Neel, J. (1980). Application of ultrafiltration to the concentration and separation of solutes of low molecular weight. *J. Membr. Sci.* **6**: 71–82.

38. Renault, M., Aulas, F., and Rumeau, M. (1982). Recovery of chromium from effluents using ultrafiltration. *Chem. Eng. J.* **23**: 137.

39. Buflle, J. and Staub, C. (1984). Measurements of complexation properties of metal ions in natural condition by ultrafiltration: measurement of equilibrium constants for complexation of zinc by synthetic and natural ligands. *Anal. Chem.* **56**: 2837–2842.

40. Juang, R.S. and Chiou, C.H. (2000). Ultrafiltration rejection of dissolved ions using various weakly basic water-soluble polymers. *J. Membr. Sci.* **177**: 207–214.

41. Vieira, M., Tavares, C.R., Bergamasco, R., and Petrus, J.C.C. (2001). Application of ultrafiltration-complexation process for metal removal from pulp and paper industry wastewater. *J. Membr. Sci.* **194**: 273–276.

42. Zakrzewska-Trznadel, G. and Harasimowicz, M. (2002). Removal of radionuclides by membrane permeation combined with complexation. *Desalination* **144**: 207–212.

43. Juang, R.S. and Shiau, R.C. (2000). Ultrafiltration rejection of dissolved ions using various weakly basic water-soluble polymers. *J. Membr. Sci.* **177**: 207–214.

44. Villoslada, I.M. and Rivas, B.L. (2003). Retention of metal ions in ultrafiltration of mixtures of divalent metal ions and water-soluble polymers at constant ionic strength based on Freundlich and Langmuir isotherms. *J. Membr. Sci.* **215**: 195–202.

45. Blocher, C. et al. (2003). Hybrid flotation-membrane filtration process for the removal of heavy metal ions from wastewater. *Water Research* **37**: 4018–4026.

46. Molinari, R., Poerio, T., and Argurio, P. (2004). Ultrafiltration-Complexation for copper-citric acid chelates removal from washing solutions of contaminated soil. *J. Appl. Electrochem.* in press.

ACID MINE DRAINAGE—EXTENT AND CHARACTER

ARTHUR W. ROSE
Pennsylvania State University
University Park, Pennsylvania

Acid mine drainage (AMD), also known as acid rock drainage (ARD), is an extensive environmental problem in areas of coal and metal mining. For example, the Appalachian Regional Commission (1) estimated that 5700 miles of streams in eight Appalachian states were seriously polluted by AMD. AMD is also serious near major metal mining districts such as Iron Mountain, CA and Summitville, CO (2,3). In streams affected by AMD, fish and stream biota are severely impacted and the waters are not usable for drinking or for many industrial purposes (4). In addition to deleterious effects of dissolved constituents (H^+, Fe, Al) on stream life, Fe and Al precipitates can cover the stream bed and inhibit stream life, and suspended precipitates can make the water unusable. In metal mining areas, heavy metals can add toxicity. General references on chemistry of AMD are Rose and Cravotta (5) and Nordstrom and Alpers (6).

CHEMISTRY OF FORMATION

AMD is formed by weathering of pyrite (FeS_2, iron sulfide) and other sulfide minerals, including marcasite (another form of FeS_2), pyrrhotite ($Fe_{1-x}S$), chalcopyrite ($CuFeS_2$), and arsenopyrite (FeAsS). The following reactions, involving oxygen as the oxidant, occur when pyrite is exposed to air and water:

$$FeS_2 + 3.5O_2 + H_2O = Fe^{2+} + 2SO_4{}^{2-} + 2H^+ \quad (1)$$

$$Fe^{2+} + 0.25O_2 + H^+ = Fe^{3+} + 0.5H_2O \quad (2)$$

$$Fe^{3+} + 3H_2O = Fe(OH)_3 + 3H^+ \quad (3)$$

The sum of these, representing complete oxidation and Fe precipitation, is

$$FeS_2 + 3.75O_2 + 3.5H_2O = Fe(OH)_3(s) + 2SO_4{}^{2-} + 4H^+ \quad (4)$$

In addition, the Fe^{3+} formed in Eq. 2 is a very effective oxidant of pyrite:

$$FeS_2 + 14Fe^{3+} + 8H_2O = 15Fe^{2+} + 2SO_4{}^{2-} + 16H^+ \quad (5)$$

These reactions also generate considerable heat, which tends to increase temperature and reaction rate.

In the above equations, the Fe precipitate is shown as $Fe(OH)_3$, but other ferric Fe phases can precipitate, depending on the conditions. Goethite (FeOOH), hematite (Fe_2O_3), ferrihydrite ($Fe_5HO_8{\cdot}4H_2O$), schwertmannite ($Fe_8O_8(OH)_6(SO_4)$), and jarosite ($KFe_3(SO_4)_2(OH)_6$) are among the products. The latter two products represent "stored acidity" that can react further to release additional acidity. Under evaporative conditions, $FeSO_4$ and other Fe sulfates can precipitate to form stored acidity.

Acid generation is dependent on a large number of factors, including the pH of the environment, temperature, the surface area of the pyrite or other source, the atomic structure of the pyrite, bacterial activities, and oxygen availability.

Oxidation of Fe^{2+} (Eq. 2) is relatively slow at pH below about 5. However, certain bacteria, such as Thiobacillus ferrooxidans, can catalyze the oxidation reaction under acid conditions. Bacterial action increases the reaction rate by a factor of about 10^6 (7). In addition, Fe^{3+}, the most effective oxidant via Eq. 4, has negligible solubility above about pH 3.5. As a result of these effects, severe AMD only develops in conditions where the water in contact with pyrite is highly acid and Fe-oxidizing bacteria are present (8). At higher pH, acid generation is relatively slow.

During natural weathering of pyrite-bearing rocks, the oxidation reactions happen slowly. In contrast, mining and other rock disturbances, such as road building, can result in greatly increased exposure of pyrite to oxidizing conditions, with resulting rapid acid generation. The water flowing from many underground mines is deficient in oxygen, and the above sequence proceeds only as far as reaction (1) [or perhaps reactions (1), (2) and (4)]. As a result, outflowing water contains elevated Fe^{2+} that oxidizes after it reaches the surface and generates additional acid owing to Fe precipitation after exposure to air. In such cases, pH can decrease downstream.

CHEMISTRY OF ACID MINE DRAINAGE

The H^+ generated by pyrite oxidation attacks various rock minerals, such as carbonates, silicates, and oxides, consuming some H^+ and releasing cations. For this reason, AMD commonly contains moderate to high levels of Ca, Mg, K, Al, Mn, and other cations balancing SO_4, the dominant anion. These reactions consume H^+ and increase the pH. If reaction with rock minerals is extensive, the resulting water may have a pH of 6 or even higher, and if oxidizing conditions exist, Fe may be relatively low. If carbonates are present in the affected rocks, the "AMD" may contain significant alkalinity as HCO_3. AMD is characterized by $SO_4{}^{2-}$ as the dominant anion but can have a wide range of pH, Fe, and other cations (Table 1).

The pH of AMD typically ranges from about 2.5 to 7, but the frequency distribution of pH is bimodal, with most common values in the range 2.5 to 4 and 5.5 to 6.5 (5). Relatively fewer values are in the range 4 to 5.5. An extreme value of negative 3.6 is reported (2). Common ranges of other constituents are up to 100 mg/L Fe, up to 50 mg/L Al, up to 140 mg/L Mn, and up to 4000 mg/L SO_4.

A key variable characterizing AMD is "acidity." Acidity is commonly expressed as the quantity of $CaCO_3$ required to neutralize the sample to a pH of 8.3 by reaction (8). The acidity includes the generation of H^+ by reactions (1) and (3), as well as the effects of other cations that generate

Table 1. Analyses of Typical "Acid Mine Drainage" (5)

Constituent	Units	WMB A1	WMS 2K	LMS S2-15
pH		2.6	4.2	6.9
Acidity	mg/L $CaCO_3$	688	270	0
Alkalinity	mg/L $CaCO_3$	0	0	730
Fe	mg/L	174	34	5.7
Mn	mg/L	25.5	67	7.5
Al	mg/L	68.9	26	<0.14
Ca	mg/L	83	270	650
Mg	mg/L	74	280	230
SO_4	mg/L	913	1600	1900
Spec. Cond.	μS/cm	2120	1860	3890

acidity at $pH < 8.3$, such as

$$Al^{3+} + 3H_2O = Al(OH)_3 + 3H^+ \quad (6)$$

In common AMD, acidity includes contributions from H^+, Fe^{2+} (Eqs. 2 and 3), Fe^{3+} (Eq. 3), Al^{3+}, and Mn^{2+}. To accurately measure the contribution of these solutes, the acidity determination should include a step in which Fe and Mn are oxidized, commonly by addition of H_2O_2 and heating (9,10). In waters from metal mining areas, other heavy metals, such as Cu, may contribute to acidity. Acidity of AMD from coal mining areas is commonly up to 1000 mg/L as $CaCO_3$.

BIBLIOGRAPHY

1. Appalachian Regional Commission. (1969). *Acid Mine Drainage in Appalachia*. Appalachian Regional Commission, Washington, DC, p. 126.
2. Nordstrom, D.K., Alpers, C.N., Ptacek, C.J., and Blowes, D.W. (2000). Negative pH and extremely acidic mine waters from Iron Mountain, California. *Environ. Sci. Technol.* **34**: 254–258.
3. Plumley, G.S., Gray, J.E., Roeber, M.M., Coolbaugh, M., Flohr, M., and Whitney, G. (1995). The importance of geology in understanding and remediating environmental problems at Summitville. *Colorado Geological Survey* Special Publication **38**: 13–19.
4. Earle, J. and Callaghan, T. (1998). Impacts of mine drainage on aquatic life, water uses and man-made structures. In: *Coal Mine Drainage Prediction and Pollution Prevention in Pennsylvania*. PA Department of Environmental Protection, pp. 1-1–1-22. Available at: http://www.dep.state.pa.us/deputate/minres/districts/CMDP/main.htm.
5. Rose, A.W. and Cravotta, C.A. (1998). Geochemistry of coal mine drainage. In: *Coal Mine Drainage Prediction and Pollution Prevention in Pennsylvania*. PA Department of Environmental Protection, pp. 1-1–1-22. Available at: http://www.dep.state.pa.us/deputate/minres/districts/CMDP/main.htm.
6. Nordstrom, D.K. and Alpers, C.N. (1999). Geochemistry of acid mine waters. In: *Environmental Geochemistry of Mineral Deposits, Reviews in Economic Geology*. Vol. 6A. G.S. Plumley and M.J. Logsdon (Eds.). Society of Economic Geologists, Littleton, CO, pp. 133–160.
7. Singer, P.C. and Stumm, W. (1970). Acidic mine drainage—The rate-determining step. *Science* **167**: 1121–1123.
8. Kleinmann, R.L.P., Crerar, D.A., and Pacelli, R.R. (1981). Biogeochemistry of acid mine drainage and a method to control acid formation. *Mining Eng.* **33**: 300–313.
9. American Public Health Association. (1998). Acidity (2310)/titration method. In: L.S. Clesceri et al. (Eds.). *Standard Methods for the Examination of Water and Wastewater*, 20th Edn. American Public Health Association, Washington, DC, pp. 2.24–2.26.
10. U.S. Environmental Protection Agency. (1979). Method 305.1, Acidity (Titrimetric). In: *Methods for Chemical Analysis of Water and Wastes*. U.S. Environmental Protection Agency Report EPA/600/4-79-020. Available at: http://www.nemi.gov.

THE CONTROL OF ALGAL POPULATIONS IN EUTROPHIC WATER BODIES

Joanna Davies
Syngenta
Bracknell, Berkshire, United Kingdom

Peter M. Scarlett
Winfrith Technology Centre
Dorchester, Dorset, United Kingdom

INTRODUCTION

Eutrophication is a natural aging process occurring in lakes and reservoirs, which is characterized by increasing nutrient levels in the water column and increasing rates of sedimentation. For water bodies in urban and agricultural landscapes, this process is accelerated by increased nutrient inputs from agricultural fertilizers, sewage effluents, and industrial discharges. Eutrophication is accompanied by increased macrophyte and algal populations, which, without appropriate management, can develop to nuisance proportions.

Algae are microscopic plants that can reproduce rapidly in favorable conditions. Some species can form scum or mats near or at the water surface. The presence of excessive algae can disrupt the use of a water body by restricting navigational and recreational activities and disrupting domestic and industrial water supplies. In particular, algal blooms can have a severe impact on water quality, causing noxious odors, tastes, discoloration, and turbidity. Blue-green algal blooms are particularly undesirable due to their potential toxicity to humans, farm livestock, and wild animals. In addition, algae may block sluices and filters in water treatment plants and reduce water flow rates, which may in turn encourage mosquitoes and increase the risk of waterborne diseases such as malaria and bilharzia (schistosomiasis). Ultimately, the presence of excessive algal populations will limit light penetration through the water column, thus inhibiting macrophyte growth and leading to reduced biodiversity.

Methods for controlling algal populations can be divided into categories as follows:

1. Environmental methods involve limiting those factors, such as nutrients, that are essential for algal growth.
2. Chemical methods involve the application of herbicides, or other products such as barley straw, that have a direct toxic effect on algae.
3. Biomanipulation involves the control of zooplanktivorous fish to favor algal grazing by invertebrates and also describes the use of microbial products.

ENVIRONMENTAL METHODS

As algae are dependent on the availability of limiting nutrients required for growth, namely, phosphorus and occasionally nitrogen, the long-term control of algal populations requires measures to reduce the entry of nutrients from external sources and to reduce the internal release of nutrients from sediments.

Nutrients may enter water bodies from point sources, such as sewage and industrial effluents, and diffuse sources, such as runoff from agricultural land following the application of animal slurries or fertilizers. Methods for reducing external nutrient loading from point sources include diversion, effluent treatment, or installation of artificial wetlands. Methods for reducing loading from diffuse sources include buffer strips and adoption of good agricultural practice.

For many shallow water bodies, attempts to reduce external loading have been less successful due to the internal release of phosphate from the sediment, which delays the decline in the total phosphorus load. Under these circumstances, the successful management of algae requires simultaneous measures to reduce internal phosphate cycling. Techniques available for this purpose include dilution and flushing, hypolimnetic withdrawal, hypolimnetic aeration, artificial circulation, phosphorus inactivation, sediment oxidation, sediment sealing, and sediment removal.

A brief description of each method for reducing external and internal nutrient loading is provided below.

Diversion

Nutrient inputs can be reduced by the diversion of effluents away from vulnerable water bodies to water courses that have greater assimilative capacity. This option is only viable where there is an alternative sink within the vicinity of the affected water body as construction of pipe work to transport water over long distances is prohibitively expensive. Diversion also reduces the volume of water flushing through the water body and, therefore, may not be viable if such a reduction is predicted to significantly affect water body hydrology. Examples where diversion has successfully reduced external phosphate loading, leading to reductions in algal biomass, include Lake Washington in the United States (1). In this case, total phosphorus concentrations were reduced from 64 μg/L prior to diversion in 1967, to 25 μg/L in 1969. This reduction was accompanied by a fivefold decrease in the concentration of chlorophyll *a* over the same period.

Effluent Treatment

Reductions in external nutrient loading can be achieved by removing phosphate and/or nitrates from effluents, prior to their discharge. Phosphorus can be removed from raw sewage or, more commonly, final effluents, by the process of stripping, which involves precipitation by treatment with aluminum sulfate, calcium carbonate, or ferric chloride. The resulting sludge is spread onto land or transferred to a waste-tip. In contrast, removal of nitrates from effluents is more complex requiring the use of ion exchange resins or microbial denitrification. The process of phosphate stripping is a requirement of the Urban Waste Water Treatment Directive in some European countries including Switzerland and The Netherlands. Phosphate stripping has successfully reduced algal biomass in Lake Windermere in the United Kingdom. In this case, total internal phosphate concentrations were reduced from 30 μg/L in 1991 to 14 μg/L in 1997, while biomass of the filamentous algae, *Cladophora*, was reduced by 15-fold between 1993 and 1997 (2).

Constructed Wetlands

External nutrient loads may also be reduced by passing effluents through detention basins or constructed wetlands, which are areas of shallow water, planted with macrophytes, that are designed to retain and reduce nutrient concentrations by natural processes. Wetlands are particularly effective for reducing nitrate concentrations by denitrification and retaining phosphorus that is bound to particles. However, they will not permanently retain soluble phosphorus and may release phosphorus from sediments at certain times of year. The development of artificial wetland systems can also be prohibitively expensive as they require large areas of land and regular maintenance, including sediment dredging and macrophyte harvest, to remain efficient. An example where constructed wetlands have been developed as part of a lake restoration scheme is provided by Annadotter et al. (3).

Good Agricultural Practice and Buffer Strips

External nutrient loading from diffuse agricultural sources can be reduced by the adoption of good agricultural practices designed to minimize fertilizer use and reduce runoff into adjacent water courses. Strategies for reducing nutrient inputs include minimizing fertilizer applications, where possible, and using slow-release formulations. Measures to minimize opportunities for runoff include avoiding applications in wet weather, incorporating fertilizer into soil by ploughing after application, maintaining ground cover for as long as possible to minimize exposure of bare ground to rainfall, and, finally, ploughing along contours.

Nutrient loading from diffuse sources can also be reduced by the creation of buffer strips between cultivated land and vulnerable water courses, in which fertilizer applications are prohibited. The development of semi-natural vegetation within these strips will also serve to intercept and assimilate the nutrients in runoff, as described for constructed wetlands (4). Under current legislation in the European Union, buffer strips are only mandatory in situations where they are essential for protecting drinking water supplies from nitrate inputs and, in particular, in areas where drinking water abstractions have concentrations exceeding 50 mg nitrate per liter. As nitrate concentrations that contribute to eutrophication are much lower than 50 mg/L, this legislation is unlikely to assist in the control of algae. Various schemes exist at member state level to promote adoption of buffer strips as a normal farming practice although experience in the United Kingdom indicates that such schemes have not been widely adopted.

Dilution and Flushing

Dilution and flushing involve the influx of large volumes of low nutrient water into the affected water body, thus diluting and therefore reducing nutrient concentrations, and washing algal cells out of the water body. This approach is dependent on the availability of large volumes of low nutrient water and systems for its transport to the affected water body. Examples where dilution has been successfully used to control algae include Green Lake in Washington State (USA). In this case, phosphorus and chlorophyll *a* concentrations were reduced by 70% and 90%, respectively, within six years of initiation of dilution (5).

Phosphorus Precipitation from the Water Column

Surface water concentrations of phosphorus can be reduced by the application of aluminum salts to the water column. At water pH values between 6 and 8, these salts dissociate and undergo hydrolysis to form aluminum hydroxide, which is capable of binding inorganic phosphorus. The resulting floc settles to the bottom of the water body, where the precipitated phosphorus is retained. The precipitate also effectively seals the sediment, thus retarding the further release of phosphorus from the sediment. At pH values below 6 or above 8, the aluminum exists as soluble ions that do not bind phosphorus. As well as being ineffective for phosphorus precipitation, these forms of aluminum present a toxic hazard to fish and aquatic invertebrates. Therefore, the successful and safe use of aluminum salts requires careful calculation of the necessary dose based on water pH and may require the use of a buffer solution such as sodium aluminate. This technique is widely used in eutrophic water bodies and Welch and Cooke (6) report several case studies. Effects are typically rapid and have been reported to reduce phosphorus release from sediments for between ten and fifteen years, although the effectiveness and longevity of the treatment may be compromised by natural sedimentation and benthic invertebrate activity.

Alternatively, phosphorus can be precipitated from the water column by the application of iron or calcium salts. These salts do not pose such a toxicity hazard as aluminum but their successful use often requires additional management techniques such as aeration or artificial circulation in order to maintain the necessary water pH and redox conditions. Consequently, the use of iron and calcium salts is less widely reported (7,8).

Hypolimnetic Withdrawal, Aeration, and Artificial Circulation

The release of phosphorus from sediment can be inhibited by increasing the concentration of dissolved oxygen in hypolimnetic waters at the water–sediment interface. The hypolimnion is the layer of water directly above the sediment, which, in thermally stratified water bodies, is usually too deep to support photosynthesis. Continued respiration leads to depletion of dissolved oxygen and the concomitant release of phosphorus from iron complexes. In stratified lakes with low resistance to mixing, wind action and the resulting turbulence may cause temporary destratification and movement of phosphorus from the hypolimnion, through the metalimnion, and into the upper epilimnion layers. In susceptible water bodies, this natural process can be circumvented by implementation of hypolimnetic withdrawal, aeration, or artificial circulation processes. By default, these approaches have the advantage of extending the habitat available for colonization by fish and zooplankton.

Hypolimnetic withdrawal involves the removal of water directly from the hypolimnion through a pipe installed at the bottom of the water body. This process requires low capital investment and reduces the detention times of water in the hypolimnion, thus reducing the opportunity for the development of anaerobic conditions. The successful implementation of these systems depends on the availability of a suitable sink for discharge waters and measures to avoid thermal destratification, caused by epilimnetic waters being drawn downward, which would otherwise encourage the transport of hypolimnetic nutrients to surface waters. This risk can be reduced by careful control of the rate of withdrawal and redirection of inlet water to the metalimnion or hypolimnion. Examples where hypolimnetic withdrawal systems have been installed in lakes are reported by Nurnberg (9).

Aeration of the hypolimnion can be achieved by mechanical agitation, whereby hypolimnetic waters are pumped onshore where the water is aerated by agitation before being returned to the hypolimnion. More usually, the hypolimnion is aerated using airlift or injection systems, which use compressed air to force hypolimnetic waters to the surface where they are aerated on exposure to the atmosphere. Water is then returned to the hypolimnion with minimal increase in temperature. In contrast, artificial circulation involves mechanical mixing of anaerobic hypolimnetic water with the upper water body using pumps, jets, and bubbled air. By default, complete circulation causes destratification of the water body, with potential adverse consequences for cold water fish species, but may also serve to reduce the concentration of algal cells in the upper water body by increasing the mixing depth and relocating algal biomass to deeper water with reduced light availability (10). Detailed examples of these processes are provided by Cooke et al. (5).

Sediment Oxidation

The release of phosphorus from sediment can also be inhibited by oxidation of sediment by the “Riplox” process. This process involves the direct injection of calcium nitrate solutions into the sediment in order to stimulate microbial denitrification and thus restore an oxidized state, conducive to the binding of interstitial phosphate with ferric hydroxide. In some cases, where the inherent iron content of the sediment is inadequate, iron chloride is initially added to generate ferric iron. Similarly, calcium hydroxide may also be added in order to raise the sediment pH to levels required for optimum microbial activity (11). This process has been documented to reduce phosphate release by up to 90% under laboratory conditions and by 50–80% following the treatment of Lake Noon in the United States (5). However, sediment oxidation is only suitable for water bodies where phosphate binding is modulated by iron redox reactions. In shallow water bodies where phosphate release is predominantly influenced by fluctuating pH and temperature at the water–sediment interface, sediment oxidation may not significantly reduce phosphorus release.

Sediment Sealing

Sediment sealing involves the physical isolation of the sediment from the water column using plastic membranes or a layer of pulverized fly ash, a solid waste product from coal-burning power stations. These techniques may prove effective at reducing algal biomass in small enclosures used for recreational purposes or industrial water supplies but are unsuitable for conservation or restoration purposes, as fly-ash may contain high levels of undesirable heavy metals and open sediment is required to support aquatic plant growth (12).

Sediment Removal

The release of nutrients from the sediment can be averted by sediment removal either using conventional excavation equipment after drawdown or using a suction-dredger and pump. In both cases, the resulting wet sediment is then transferred to a suitable disposal site. Direct dredging is more commonly practiced but has the disadvantage that phosphates or other toxins adsorbed to sediment may be released into the water column when sediment is disturbed. Furthermore, benthic organisms will inevitably be disturbed and removed during dredging, although recolonization will mitigate any long-term effects. Examples where sediment removal projects have been implemented include Lake Finjasjon in Sweden and the Norfolk Broads in the United Kingdom (3,13).

CHEMICAL METHODS

Herbicides

The availability of aquatic herbicides has been limited by their small market value and the stringent toxicological and technical challenges presented by the aquatic environment. Aquatic herbicides, particularly those used to control algae, must be absorbed rapidly from dilute and often flowing aqueous solution. More recently, the number of products available for use in water has been reduced by the prohibitive cost of meeting the increasingly stringent requirements of pesticide registration. In 2004, the few products that remain on the market for use on a large scale are based on the active ingredients of copper, diquat, endothall, and terbutryn. Additional products based on other active ingredients are available for amateur use in enclosed garden ponds but are not registered for use in larger water bodies.

Use of herbicides to control algal blooms is often limited where water is required for domestic drinking water supply or for the irrigation of crops or livestock. Their use under these circumstances requires strict adherence to label recommendations and observation of recommended irrigation intervals. Despite these restrictions, chemical control may be preferable where immediate control is required or alternative, long-term measures are prohibited due to excessive cost.

Herbicide applications to water are generally made using hand-operated knapsack sprayers operated from bank or boat, or spray booms mounted to boats, tractors, helicopters, or planes. Much of this equipment is modified from conventional agricultural sprayers and nozzles, although the injection of herbicides into deep water or onto channel beds may require the use of weighted, trailing hoses fitted to boat-mounted spray booms.

Disadvantages associated with the use of herbicides include potential adverse effects on nontarget organisms including aquatic invertebrates and fish, development of herbicide-resistant algal strains, and excessive copper accumulation in treatment plant sludges, which may lead to disposal problems. The control of algae with herbicides can also create large quantities of decaying tissue, which cause deoxygenation of the water due to a high bacterial oxygen demand. This may lead to the death of fish and other aquatic organisms, particularly during summer months when deoxygenation is more rapid due to lower dissolved oxygen levels and increased rates of decomposition caused by higher water temperatures. Deoxygenation can largely be avoided by restricting applications to the early growing season. Where later treatments are essential, applications should be restricted to discrete localized areas of a water body, or slow-release formulations should be used to avoid a sudden buildup of decaying tissue.

Barley Straw

Research in the United Kingdom has led to the use of barley straw for the control and prevention of algal blooms in a range of water bodies including lakes, reservoirs, rivers, streams, and ponds (14). During decomposition, barley straw is known to release compounds, including active oxygen and hydrogen peroxide, which are toxic to algae (15). Although similar effects have been demonstrated with wheat, linseed, oil-seed rape, and lavender straw, barley straw provides the most effective and long-lasting control. For effective control, chopped straw is typically added to the water in loosely packed bales enclosed in mesh sacks, netting, or wire cages that are anchored to the bank or to the bottom of the water body. Depending on the type of water body and the severity of the algal bloom, typical application rates vary between 100 and 500 kg/ha. In order to promote the decomposition process and the release and distribution of the algicidal components, bales should be positioned at, or near, the water surface where water temperatures, movement, and aeration are the greatest. Depending on water temperature, the decomposition process may continue for 2–8 weeks after application before the straw releases sufficient quantities of the active components to cause an effect. These chemicals continue to be released until decomposition is complete, which may take up to six months. Therefore, for maximum efficacy, straw is typically applied in spring before algal blooms reach their peak and again in the autumn.

The use of straw has few, if any, adverse effects on nontarget organisms such as macrophytes, invertebrates, or fish. The main disadvantage associated with its use is the risk of deoxygenation caused by the high bacterial oxygen demand of those microorganisms responsible for straw decomposition. Deoxygenation may lead to the death of fish and other aquatic organisms, particularly during summer months when rates of decomposition are increased due to higher water temperatures. This risk can be reduced by early application during the spring and avoiding applications during prolonged periods of hot weather.

BIOMANIPULATION

Removal of Zooplanktivorous Fish

Algal populations may be controlled through the biomanipulation of foodwebs to favor algal grazing by invertebrates, by removing fish (16). Methods for the direct removal of fish include application of the piscicide, rotenone, to anesthetize fish prior to removal or to kill them outright. However, use of rotenone for this purpose is discouraged in some countries, including the United Kingdom, and requires special consent from the appropriate government authority. Alternative methods for fish removal include systematic electrofishing, seine netting, and fish traps. In smaller water bodies, drawdown may be used to concentrate fish into increasingly smaller water volumes to simplify their capture. Even where complete drawdown is possible, elimination of all fish is unlikely and additional measures may be necessary to prevent successful spawning of remaining fish. Spawning can be prevented by nets, placed in traditional spawning areas, to capture eggs, which are then removed.

The long-term success of fish removal schemes also requires measures to prevent recolonization from tributaries and floodwaters. In cases where water can be diverted and tributaries do not carry boat traffic, isolation of water bodies may be possible by the construction of dams. Where boat traffic requires access, the installation of electronic netting gates, which automatically lower to allow access, or engineered locks, which dose the water with rotenone on opening, may be required (12). Recolonization can also be avoided by creating fish-free enclosures, using fishproof barriers, within affected water bodies. Once restoration is complete within an enclosure, more enclosures may be built and eventually joined together until a large proportion of the water body is enclosed. This approach was used during the restoration of Hoveton Broad in the United Kingdom (12).

Alternative methods to control fish populations involve the introduction of piscivorous fish, such as pike or pikeperch in the United Kingdom or American large-mouthed bass in the United States. The addition of piscivores can have immediate and dramatic impacts on invertebrate populations and has been practiced as part of restoration schemes in Lake Lyne in Denmark and Lakes Zwemlust and Breukeleveen in The Netherlands (12). However, the success of this approach is unlikely to be sustained without regular restocking as predation of the fish population will invariably lead to a decline in the piscivore population.

Microbial Products

The use of bacteria to control algal populations is a recent innovation adapted from the wastewater industry. As bacteria have a high surface area to volume ratio and a high uptake rate for nutrients relative to unicellular algae, they can out-compete algae for limiting nutrients, such as nitrogen and phosphorus, and have been demonstrated to suppress the growth of algal cultures under laboratory conditions (17). This observation has led to the commercial development of microbial products, containing bacteria and enzymes, that are designed to supplement natural microbial populations to the levels required to have a significant impact on algae. The number of available products has increased as the use of chemical algicides has become more restricted. However, few researchers report a significant reduction in algal growth following the use of microbial products under experimental conditions and their use on a large scale has yet to be widely documented (18).

CONCLUSIONS

While chemical control methods, such as the application of herbicides or barley straw, can provide rapid, short-term reductions in algal populations, the long-term and sustainable management of algae requires consideration of the cause and source of eutrophication and the implementation of techniques to reduce nutrient loading and to restore natural foodweb interactions. Evaluation of the suitability of the techniques discussed here, for use in individual cases, requires detailed assessments to determine the trophic status of the affected water body and, in particular, the relative contribution of point source, diffuse, and internal nutrient sources to total nutrient concentrations. Only when the causes of eutrophication are clearly identified can the symptom of excessive algal growth be efficiently managed.

BIBLIOGRAPHY

1. Edmondson, W.T. and Lehman, J.R. (1981). The effect of changes in the nutrient income on the condition of Lake Washington. *Limnol. Oceanography* **1**: 47–53.
2. Parker, J.E. and Maberley, S.C. (2000). Biological response to lake remediation by phosphate stripping: control of *Cladophora*. *Freshwater Biol.* **44**: 303–309.
3. Annadotter, H. et al. (1999). Multiple techniques for lake restoration. *Hydrobiologia* **395/396**: 77–85.

4. Abu-Zreig, M. et al. (2003). Phosphorous removal in vegetated filter strips. *J. Environ. Qual.* **32**(2): 613–619.
5. Cooke, G.D., Welch, E.B., Peterson, S.A., and Newroth, P.R. (1993). *Restoration and Management of Lakes and Reservoirs*, 2nd Edn. Lewis Publishers, Boca Raton, FL.
6. Welch, E.B. and Cooke, G.D. (1999). Effectiveness and longevity of phosphorous inactivation with alum. *J. Lake Reservoir Manage.* **15**: 5–27.
7. Randall, S., Harper, D., and Brierley, B. (1999). Ecological and ecophysiological impacts of ferric dosing in reservoirs. *Hydrobiologia* **395/396**: 355–364.
8. Prepas, E.E. et al. (2001). Long-term effects of successive $Ca(OH)_2$ and $CaCO_3$ treatments on the water quality of two eutrophic hardwater lakes. *Freshwater Biol.* **46**: 1089–1103.
9. Nurnberg, G.K. (1987). Hypolimnetic withdrawal as lake restoration technique. *J. Environ. Eng.* **113**: 1006–1016.
10. Brierly, B. and Harper, D. (1999). Ecological principles for management techniques in deeper reservoirs. *Hydrobiologia* **395/396**: 335–353.
11. Ripl, W. (1976). Biochemical oxidation of polluted lake sediment with nitrate—a new restoration method. *Ambio* **5**: 132.
12. Moss, B., Madgwick, J., and Phillips, G. (1997). *A Guide to the Restoration of Nutrient-Enriched Shallow Lakes*. Environment Agency, UK.
13. Phillips, G. et al. (1999). Practical application of 25 years research into the management of shallow lakes. *Hydrobiologia* **396**: 61–76.
14. Barrett, P.R.F and Newman, J.R. (1992). Algal growth inhibition by rotting barley straw. *Br. Phycol. J.* **27**: 83–84.
15. Everall, N.C. and Lees, D.R. (1997). The identification and significance of chemical released from decomposing barley straw during reservoir algal control. *Water Res.* **30**: 269–276.
16. Moss, B. (1992). The scope for biomanipulation in improving water quality. In: *Eutrophication: Research and Application to Water Supply*. D.W. Sutcliffe and J.G. Jones (Eds.). Freshwater Biological Association, Cumbria, UK, pp. 73–81.
17. Brett, M.T. et al. (1999). Nutrient control of bacterioplankton and phytoplankton dynamics. *Aquat. Ecol.* **33**(2): 135–145.
18. Duvall, R.J., Anderson, W.J., and Goldman, C.R. (2001). Pond enclosure evaluations of microbial products and chemical algicides used in lake management. *J. Aquat. Plant Manage.* **39**: 99–106.

ARSENIC COMPOUNDS IN WATER

CORINNA WATT
X. CHRIS LE
University of Alberta
Edmonton, Alberta, Canada

Arsenic is the twentieth most abundant element in the earth's crust; it occurs naturally in the environment in both inorganic and organic forms. Arsenic is also released to the environment by anthropogenic activities such as pesticide use, wood preservation, mining, and smelting. Human exposure to arsenic by the general population occurs primarily from drinking water and food. In areas of endemic arsenic poisoning such as Bangladesh, India, Inner Mongolia, and Taiwan, the main exposure is through drinking water where inorganic arsenic levels can reach concentrations in the hundreds or thousands of micrograms per liter. The arsenic is released from natural mineral deposits into the groundwater in endemic areas. Groundwater is the primary drinking water source in these areas.

Arsenic is present in a variety of inorganic and organic chemical forms in water. This is a result of chemical and biological transformations in the aquatic environment. The specific arsenic compound present determines its toxicity, biogeochemical behavior, and environmental fate.

In natural waters, arsenic is typically found in the +5 and +3 oxidation states (1–6). The most common arsenic compounds detected in water are arsenite (As^{III}) and arsenate (As^{V}). Monomethylarsonic acid (MMA^{V}), monomethylarsonous acid (MMA^{III}), dimethylarsinic acid (DMA^{V}), dimethylarsinous acid (DMA^{III}), and trimethylarsine oxide (TMAO) have also been detected in water (Table 1). Several reviews provide important additional information regarding the cycling and speciation of arsenic in water and the environment (1,3,4,7–12).

Arsenic and arsenic compounds are classified as Group 1 carcinogens in humans by the International Agency for Research on Cancer (IARC). "The agent (mixture) is carcinogenic to humans" and "the exposure circumstance entails exposures that are carcinogenic to humans" (13). Chronic exposure to high levels of arsenic in drinking water has been linked to skin cancer, bladder cancer, and lung cancer, as well as to several noncancerous effects (8,11,14). Noncancerous effects from arsenic exposure include skin lesions, peripheral vascular disease (blackfoot disease), hypertension, diabetes, ischemic heart disease, anemia, and various neurological and respiratory effects (8,11,14). The reproductive and developmental effects of arsenic exposure have also been reported (14). The National Research Council and the World Health Organization have reviewed the most significant studies on arsenic exposure, toxicity, and metabolism (8,11,14). The association of arsenic with internal cancers has led to increased pressure for stricter guidelines for arsenic in drinking water.

Table 1. Arsenic Species Present in Water

Arsenic Species	Chemical Abbreviation	Chemical Formula
Arsenite, arsenous acid	As^{III}	H_3AsO_3
Arsenate, arsenic acid	As^{V}	H_3AsO_4
Monomethylarsonic acid	MMA^{V}	$CH_3AsO(OH)_2$
Monomethylarsonous acid	MMA^{III}	$CH_3As(OH)_2[CH_3AsO]$
Dimethylarsinic acid	DMA^{V}	$(CH_3)_2AsO(OH)$
Dimethylarsinous acid	DMA^{III}	$(CH_3)_2AsOH$ $[((CH_3)_2As)_2O]$
Trimethylarsenic compounds	TMA	$(CH_3)_3As$ and precursors
Trimethylarsine oxide	TMAO	$(CH_3)_3AsO$
Oxythioarsenic acid		H_3AsO_3S

The World Health Organization (WHO) guideline for arsenic in drinking water is 10 μg/L. The Canadian guideline is 25 μg/L and is currently under review. The United States has recently lowered its maximum contaminant level (MCL) for arsenic in drinking water from 50 μg/L to 10 μg/L. The previous MCL was established as a Public Health Service standard in 1942 and was adopted as the interim standard by the U.S. Environmental Protection Agency (EPA) in 1975. The Safe Drinking Water Act (SDWA) Amendments of 1986 required the EPA to finalize its enforceable MCL by 1989. The EPA was not able to meet this request partly because of the scientific uncertainties and controversies associated with the chronic toxicity of arsenic. The SDWA Amendments of 1996 require that the EPA propose a standard for arsenic by January 2000 and promulgate a final standard by January 2001. In 1996, the EPA requested that the National Research Council review the available arsenic toxicity data and evaluate the EPA's 1988 risk assessment for arsenic in drinking water. The National Research Council Subcommittee on Arsenic in Drinking Water advised in its 1999 report that, based on available evidence, the MCL should be lowered from 50 μg/L (8). After much debate on what would be an appropriate MCL (3, 5, 10, or 20 μg/L), the EPA published its final rule of 10 μg/L in January 2001 under the Clinton administration. This was initially rescinded by the Bush administration. The reasons cited for this rejection were the high cost of compliance and incomplete scientific studies. In 2001, the National Research Council organized another subcommittee to review new research findings on arsenic health effects and to update its 1999 report on arsenic in drinking water (14). This second subcommittee concluded that "recent studies and analyses enhance the confidence in risk estimates that suggest chronic arsenic exposure is associated with an increased incidence of bladder and lung cancer at arsenic concentrations in drinking water that are below the MCL of 50 μg/L" (14). The MCL of 10 μg/L was approved later by the EPA and became effective in February 2002. The date for compliance has been set at January 23, 2006.

Most controversies over an appropriate MCL arise from a lack of clear understanding of the effects on health from exposure to low levels of arsenic. There are limited studies available that have examined the increased risk of cancer at low levels of arsenic exposure. Therefore, the National Research Council and the EPA had to base their assessments on epidemiological studies where the arsenic exposure is very high (8,11,14). Extrapolation of health effects from very high exposure data to the much lower exposure scenarios involves large uncertainties. Experimental animal studies were also consulted as well as the available human susceptibility information (8,14). Animal studies are of limited value when examining arsenic effects in humans because of differences in sensitivity and metabolism. Experimental animals are often exposed to very large doses of arsenic that are not representative of typical human exposure. In addition, the studies used to establish the new MCL do not take into account arsenic exposure through food intake and other beverages. Other confounding factors may include nutrition, metabolism, and predetermined genetic susceptibility (8,14).

The exact mechanism of arsenic's tumorigenicity is not clear. Therefore it is difficult to ascertain the risk of cancers from very low exposures to arsenic. Due to the limited information on the cancer risks posed by low-level arsenic exposure and to the unknown shape of the dose–response curve at low doses, the EPA used a default linear dose–response model to calculate the cancer risk. The linear extrapolation to zero assumes that there is no safe threshold of exposure at which health effects will not occur, whereas others argue that a safe threshold level or sublinear dose–response relationship may exist.

The cost of compliance with a lower MCL and the monitoring and treatment technology are other important considerations in setting the new MCL. According to the EPA, water systems that serve 13 million people and representing 5% of all systems in the United States would have to take corrective action at an MCL of 10 μg/L (15). However, the EPA believes that the new MCL "maximizes health risk reduction at a cost justified by the benefits" (15).

INORGANIC ARSENIC

As^{III} and As^{V} are the dominant arsenic species detected in most natural waters. As^{III} and As^{V} have been detected in all forms of natural water including groundwater, freshwater, and seawater (Table 2). In oxygen-rich waters of high redox potential, the As^{V} species H_3AsO_4, $H_2AsO_4^{-}$, $HAsO_4^{2-}$, and AsO_4^{3-} are stable (1,4). As^{III} species, which exist in reduced waters, may include H_3AsO_3, $H_2AsO_3^{-}$, and $HAsO_3^{2-}$ (1,4). However, in most natural waters, As^{III} is present as H_3AsO_3 (as its pK_a values are 9.23, 12.13, and 13.4) (8,16). The pK_a values of As^{V} are 2.22, 6.98, and 11.53 (8). At natural pH values, arsenate is usually present as $H_2AsO_4^{-}$ and $HAsO_4^{2-}$ (1,4,16).

The distribution of As^{V} and As^{III} throughout the water column varies with the season due to changes in variables such as temperature, biotic composition, pH, and redox potential (4,17–28). Biological activity also contributes to changes in speciation (1,3,4,6,20,22–24,29–39). The uptake of As^{V} by phytoplankton and marine animals results in the reduction of As^{V} to As^{III} and the formation of methylated arsenic compounds (4,19,23,24,29,32,34,37–42).

As^{V} is typically dominant in oxygen-rich conditions and positive redox potentials (4,6,33,43,44). As^{V} has been detected as the predominant species in most natural waters (Table 2). As^{III} is expected to dominate in anaerobic environments (4,7,10,25,33,45). As^{III} has been detected as the predominant species in groundwater (46–48) and is significant in the photic zone of seawater (43,45,49). As^{III} has also been associated with anoxic conditions in estuaries (31,50), seawater (33,43), and marine interstitial water (6,30). In lake interstitial water, 55% of the dissolved arsenic was present as arsenite and a few percent as DMA^{V} (51).

The As^{III}/As^{V} ratio typically does not reach thermodynamic equilibrium (1,4,6,20,24,33,50,52,53). This reflects biological mediation. The kinetics of the As^{V}–As^{III}

Table 2. Inorganic Arsenic Detected in Water[a]

Source	As^{III}	As^{V}	Reference
	Groundwater		
Lead–zinc mine, Coeur d'Alene, Idaho, USA	0.1–42.3	N.A.	81
	0.1–1336.1[b]		81
Southwest United States	<5	21–120	44
Bowen Island, British Columbia, Canada	N.D.–220	N.D.–477.7	82
Six districts in West Bengal, India	10–2600	16–1100	83
Bangladesh (tubewell for drinking)	30–1200[b]		84
Bangladesh (deep tubewell for drinking)	50–80[b]		84
Bangladesh (deep tubewell for irrigation)	30–200[b]		84
Purbasthali (Burdwan), India	Present in 1 sample	N.D.–133.6	85
Taiwan (Pu-Tai) endemic area	285–683 (462 ± 129)	33–362 (177 ± 109)	48,86
Taiwan (I-Lan) comparison area	537–637 (572 ± 42)	24–67 (38 ± 18)	86
Taiwan (Hsin-Chiu) control area	<3	<12	86
Taiwan (Pu-Tai)	~20–720[b]		87
Taiwan (Fushing)	70.2 ± 2.6	870 ± 26	88
Taiwan (Chiuying)	51.6 ± 1.8	601 ± 22	88
Region Lagunera, northern Mexico	Trace–217	4–604	89
Fukuoka Prefecture, Japan	15–70	11–220	90
Kelheim, Lower Bavaria, Germany[c]	N.D.–176.3	N.D.–147.7	91
	Rivers		
Hillsborough River, Tampa, Florida, USA	N.D.	0.25 ± 0.01	92
Withlacoochee River, Tampa, Florida, USA	N.D.	0.16 ± 0.01	92
Tamiami Canal, Florida, USA	0.47	0.07	49
Colorado River, Parker, Arizona, USA	0.11	1.95	49
Nine rivers in California, USA	0.02–1.3	0.07–42.5	49
Four rivers in California, USA	0.7–7.4[b]		20
Colorado River Slough, near Topock, California	0.08	2.25	93
North Saskatchewan River, Canada	0.21 ± 0.08	0.32 ± 0.01	94
Haya-kawa River, Japan	28[b]		95
River in Taiwan (Erhjin)	2.95 ± 3.52	4.59 ± 1.31	88
River in Taiwan (Tsengwen)	0.12 ± 0.06	2.57 ± 1.17	88
	Lakes, Ponds, Reservoirs		
Mono Lake, California, USA (highly alkaline lake)	17250[b]		20
Pyramid Lake, California, USA	97.5[b]		20
Six lakes in California, USA	0.3–11.2[b]		20
Davis Creek Reservoir (filtered samples), California, USA	Trace–1.2	0.01–1.5	20
Davis Creek Reservoir (unfiltered samples), California, USA	0.8–2.6[b]		20
Elkhorn Slough, California, USA	1.1[b]		20
Donner Lake N. Shore, California, USA	0.06	0.07	49
Saddleback Lake S. End, California, USA	0.05	0.02	49
Squaw Lake (surface), California, USA	0.87	2.76	49
Squaw Lake (2-m depth), California, USA	0.53	2.96	49
Senator Wash Reservoir, California, USA	0.58	2.47	49
Lake Echols, Tampa, Florida, USA	2.74 ± 0.01	0.41 ± 0.01	92
Lake Magdalene, Tampa, Florida, USA	0.89 ± 0.01	0.49 ± 0.01	92
Remote pond, Withlacoochee Forest, Tampa, Florida, USA	N.D.	0.32 ± 0.01	92
University research pond, Tampa, Florida, USA	0.79 ± 0.01	0.96 ± 0.01	92
Coot Bay Pond, Everglades, Florida, USA	0.81	1.27	49
Paurotis Pond, Everglades, Florida, USA	0.04	0.02	49
Lake Washington, Seattle, USA	40% of TDA	60% of TDA	51
Lake Washington interstitial water, Seattle, USA (one sample)	55% of TDA	N.D.	51
Subarctic lakes[c] (pH 1), Yellowknife, NWT, Canada	N.A.	0.7–520	74
Subarctic lakes[c] (pH 6), Yellowknife, NWT, Canada	0–22	N.A.	74
Lake Biwa, Japan	<0.2	<1.9	64
Lake in Taiwan (Nanjen Hu)	0.05 ± 0.00	1.67 ± 0.07	88
Lake Pavin, France (measurements at different depths)	0–8.25	0.28–4.4	53

Table 2. (*Continued*)

Source	$\mathrm{As}^{\mathrm{III}}$	As^{V}	Reference
Seawater			
Pacific Ocean (<100 m)	1.1–1.6[b]		39,96
Pacific Ocean, SE Taiwan	0.05 ± 0.02	1.28 ± 0.24	88
Western Atlantic Ocean (Stn.10)—surface	∼0.3[b]		45
Western Atlantic Ocean (Stn.10)—deep water (2200–3300 m)	1.5 ± 0.1[b]		45
Western Atlantic Ocean (Stn.10)—bottom water (3950–4460 m)	1.5 ± 0.03[b]		45
Western Atlantic Ocean (Stn.10)—80-m deep mixed layer (maximum)	0.03	N.D.	45
Western Atlantic Ocean (Stn.10)—700 m to bottom (maximum)	0.007	N.D.	45
Antarctic Ocean	0.003	1.0	54
East Indian Ocean	0.2	0.4	54
North Indian Ocean	0.2	0.5	54
Baltic Sea	N.D.–0.92	N.D.–1.01	43
Southern North Sea (over several months)	0.1–1.5[b]		65
China Sea	0.2	0.3	54
Indonesian Archipelago	0.2	0.4	54
Estuaries and Coastal Waters			
Scripps Pier, La Jolla, California, USA	0.01–0.03	1.70–1.75	93
San Diego Trough, California, USA (surface to 100 m below surface)	0.01–0.06	1.32–1.67	93
Suisun Bay, California, USA	1.5[b]		20
Southern California Bight (varying depths and locations)	N.D.–0.87	0.16–1.45	49
Tidal flat, Tampa, Florida, USA	0.6 ± 0.01	1.3 ± 0.01	92
Bay, Causeway, Tampa, Florida, USA	0.1 ± 0.01	1.4 ± 0.01	92
McKay Bay, Tampa, Florida, USA	0.06 ± 0.01	0.3 ± 0.01	92
Chesapeake Bay[c], Maryland, USA	N.D.–0.24	0.13–1.02	31
Patuxent River Estuary, Maryland, USA (over 2 years)	0.1–0.2	0.1–1.1	27
Saanich Inlet, Vancouver, British Columbia, Canada (oxic stations)	0.07–0.19	0.77–2.51	33
Saanich Inlet, Vancouver, British Columbia, Canada (anoxic stations)	0.07–1.91	N.D.–2.20	33
Hastings Arm[c] porewaters, British Columbia, Canada	0.80–11.5	2.0–12.1	30
Alice Arm[c] porewaters, British Columbia, Canada	2.9–28.8	2.3–28.8	30
Quatsino Sound[c] porewaters (one station), British Columbia, Canada	N.A.	<3–10.5	30
Holberg Inlet[c] porewaters (one station), British Columbia, Canada	N.A.	1.5–<13.5	30
Rupert Inlet[c] porewaters (two stations), British Columbia, Canada	N.A.	<1.5–9.0	30
Surface waters of the Beaulieu Estuary, Hampshire, UK	<0.02–0.40	N.A.	17
	0.17–1.20[b]		17
Deep waters of the Beaulieu Estuary, Hampshire, UK	0.43–1.07[b]		17
Tamar Estuary[c], southwest England (primarily As^{V})	∼2.7[b]		63
Tamar Estuary[c] (porewaters), southwest England	∼29.6[b]		63
Tamar Estuary[c] (porewaters), southwest England	5–60[b]		97
Southampton water, UK	0.03–0.1	N.A.	26
Humbar Estuary[c], UK (over four seasons)	0.7–2.3[b]		18
Thames Estuary and plume[c], UK (Feb.1989)	3.3 ± 0.9	N.A.	23
Thames Estuary and plume[c], UK (July 1990)	2.1 ± 1.0[b]		23
Seine Estuary, France (over several months)	N.D.–1.2	0.4–2.1	22
Continental shelf off the Gironde Estuary, France	0.01–0.3	0.8–1.5	29
Tagus Estuary[c], Atlantic coast of Europe	2.8–14.7[b]		67
Schelde Estuary[c], central Europe	0.06–0.4	1.8–4.8	50
Schelde Watershed[c], central Europe	0.003–8.2	0.9–27.7	50
Coastal water in Taiwan (Erhjin)	0.06 ± 0.04	0.88 ± 0.25	88
Coastal waters in Taiwan (Tsengwen)	0.03 ± 0.01	0.62 ± 0.05	88
Other			
Rainwater, Washington, USA	35% of TDA	65% of TDA	51
Rainwater, La Jolla, California, USA	<0.002	0.09–0.18	93
Hot Creek Gorge, Eastern Sierra Nevada, USA	0.61 ± 0.20		98
North drainage channel—Twin Butte Vista Hot Spring, Yellowstone National park, USA	610–1900	600–1900	99
Bottled water (carbonated)	1.6 ± 0.1	7.5 ± 0.3	100

N.A.: Not available (Not reported).
N.D.: not detected.
TDA: Total dissolved arsenic.
[a] All concentrations in μg/L.
[b] The sum of $\mathrm{As}^{\mathrm{III}} + \mathrm{As}^{\mathrm{V}}$.
[c] Anthropogenic contamination.

transformation in natural waters is also known to be chemically slow (5,33,37,52,53). As a result, As^{III} has been observed in oxic waters, and As^{V} has been found in highly sulfidic water (25), contradictory to thermodynamic predictions. As^{III} was present in higher than expected amounts in the oxic epilimnion of the Davis Creek Reservoir (20). As^{III} has been detected in the oxic surface waters of lakes (24,53) and seawater (45,54). As^{V} has also been detected in anoxic zones (24,33,43,53).

The instability of arsenic species in water samples and the procedures used for sample handling and analysis may be the reason that As^{V} is commonly reported as the predominant species in water. Oxidation of As^{III} to As^{V} can occur during sample handling and storage. As^{III} may be present in higher concentrations in groundwater than previously reported (7). Many methods for preservation have been tested to prevent oxidation of As^{III} to As^{V} (55–60). On-site methods of analysis have been developed to avoid the need for preservation (61,62). As^{III} was determined as the predominant species in most groundwater samples measured from tanks and wells in Hanford, Michigan (47). In thirteen of sixteen wells and three of four tanks, As^{III} was present as $86 \pm 6\%$ of the total arsenic (47).

METHYLATED ARSENIC COMPOUNDS

Biological mediation is primarily responsible for the production and distribution of methylated arsenic species (1,3,4,6,19,22,29,31,34,35,37,42,43,45,49,63). MMA^{V}, MMA^{III}, DMA^{V}, DMA^{III}, and TMAO are the methylarsenicals present in natural waters (Table 3). It is estimated that methylarsenicals account for approximately 10% of the arsenic in the ocean (43). Methylarsenicals accounted for up to 59% of the total arsenic in lakes and estuaries in California (20). The pK_a values of MMA^{V} are 4.1 and 8.7, and the pK_a of DMA^{V} is 6.2 (8). At neutral pH, MMA^{V} occurs as $CH_3AsOOHO^-$ and DMA^{V} occurs as $(CH_3)_2AsOO^-$ (4). Trimethylarsenic (TMA) species have also been detected in some natural waters (Table 3). In the contaminated Tagus Estuary, trimethylarsenic was present at 0.010–0.042 μg/L (67). TMAO has been detected in marine interstitial waters (30) but not in surface waters (34). Methylated arsenic species are at a maximum in the euphotic zone of seawater (36,39,43). Similar to inorganic arsenic, the distribution of methylarsenicals is affected by seasonal changes in natural waters (19,20,22–24,27,32,64–66). Methylarsenicals usually occur predominantly in the pentavalent (+5) oxidation state (4).

Pentavalent arsenic species (As^{V}, MMA^{V}, DMA^{V}, and TMAO) are chemically or biologically reduced to trivalent arsenic species (As^{III}, MMA^{III}, and DMA^{III}), and biological methylation results in methylarsenicals. MMA^{III} and DMA^{III} are intermediates in the two-step methylation process. MMA^{III} and DMA^{III} appeared in minor fractions in Lake Biwa (eutrophic zone), Japan, and DMA^{V} was dominant in the summer (28,64).

PARTICULATE ARSENIC

In oxidizing conditions, As^{V} becomes associated with particulate material and may be released in reducing conditions. The amount of arsenic adsorbed to particulate can be substantial (62,68,69). The failure to account for arsenic in particulate results in an underestimation of total arsenic and inefficient treatment and removal of arsenic in water. The particulate matter in natural waters may occur as undissolved mineral (1,4,10) and organic species (70,71). As^{V} readily adsorbs and/or coprecipitates onto Fe^{III} oxyhydroxide particles (1,4,8,10,16,71). As^{V} and As^{III} can also react with sulfide ions to form insoluble arsenic sulfide precipitates (1,8,10,16,72,73). Under highly reducing conditions, the organic/sulfide fraction predominates (16,73).

In groundwater samples in the United States, particulate arsenic accounted for more than 50% of the total arsenic in 30% of the samples collected (69). In the Nile Delta Lakes, the arsenic budget consisted of 1.2–18.2 μg/L dissolved arsenic and between 1.2 and 8.7 μg/g of particulate arsenic (73). Significant amounts of particulate arsenic have also been detected in estuaries and coastal waters (17,23,70). In the drainage waters of an area impacted by mine wastes, particulate matter was greater than 220 times more concentrated than dissolved arsenic. Suspended matter in the deep water of Lake Washington contained up to 300 mg/g arsenic in the summer (51). In the Gironde Estuary, arsenic-containing suspended particulate matter varied with depth (5.1–26.8 μg/g) (29). The level of arsenic in phytoplankton was estimated at 6 μg/g compared to 20–30 μg/g in iron-rich and aluminum-rich terrigenous particles (29).

UNCHARACTERIZED ARSENIC SPECIES

Substantial amounts of arsenic species remain to be characterized. There are many reports of unidentified arsenic species in water (8,21,29,34,74–78). After UV irradiation of surface water from Uranouchi Inlet, Japan, the inorganic arsenic and dimethylarsenic concentrations detected by hydride generation increased rapidly (75). The UV-labile arsenic fractions represented 15–45% and 4–26% of the total dissolved arsenic in Uranouchi Inlet and Lake Biwa, respectively. In sediment porewater of Yellowknife, Canada, there was an increase of 18–420% in total dissolved arsenic concentration observed after irradiation (74). The difference between the sum of known arsenic species and total arsenic in the euphotic layer of an estuary was 13% (29).

In coastal waters, the concentration of total dissolved arsenic increased by approximately 25% following UV irradiation of the samples (78). In a National Research Council of Canada river water standard reference material, approximately 22% of the arsenic was unidentified (76). In estuarine waters, uncharacterized arsenic compounds corresponded to approximately 20% and 19% of the total arsenic content in summer and winter samples, respectively (77). Identification of these compounds is necessary to complete our understanding of the biogeochemical cycling of arsenic in the environment.

ARSENOTHIOLS

There is limited evidence that the precursors to methylarsine species detected by hydride generation may

Table 3. Organic Arsenic Species Detected in Water[a]

Source	MMA	DMA	TMA	Reference
Groundwater				
Taiwan (Pu-Tai)	∼0.5–4.2	∼2.0–6.9	∼3.3–5.1	87
Region Lagunera, northern Mexico	<3	Trace–20		89
Rivers				
Withlacoochee River, Tampa, Florida, USA	0.06 ± 0.01	0.3 ± 0.01		92
Tamiami Canal, Florida, USA	<0.005	0.05		49
Colorado River, Parker, Arizona, USA	0.06	0.05		49
Nine rivers in California, USA	<0.007–0.56	<0.007–0.31		49
Colorado River Slough, near Topock, California	0.13	0.31		93
Haya-kawa River, Japan	N.D.	N.D.	2	95
Lakes, Ponds, Reservoirs				
Pyramid Lake, California, USA	N.D.	1.1		20
Six lakes in California, USA	N.D.–0.6	0.03–2.4		20
Davis Creek Reservoir (filtered samples), California, USA	0.18–0.3	0.03–0.9		20
Davis Creek Reservoir (unfiltered samples), California, USA	0.22–0.3	N.D.–0.8		20
Elkhorn Slough, California, USA	0.07	0.2		20
Donner Lake N. Shore, California, USA	<0.008	0.003		49
Saddleback Lake S. End, California, USA	<0.002	0.006		49
Squaw Lake (surface), California, USA	0.022	0.052		49
Squaw Lake (2-m depth), California, USA	0.020	0.11		49
Senator Wash Reservoir, California, USA	0.014	0.006		49
Lake Echols, Tampa, Florida, USA	0.11 ± 0.01	0.32 ± 0.01		92
Lake Magdalene, Tampa, Florida, USA	0.22 ± 0.01	0.15 ± 0.01		92
Remote Pond, Withlacoochee Forest, Tampa, Florida, USA	0.12 ± 0.01	0.62 ± 0.01		92
University research pond, Tampa, Florida, USA	0.05 ± 0.01	0.15 ± 0.01		92
Coot Bay Pond, Everglades, Florida, USA	<0.01	0.32		49
Paurotis Pond, Everglades, Florida, USA	<0.005	0.03		49
Lake Washington, Seattle, USA	N.D.	0.05		51
Lake Washington interstitial water, Seattle, USA (one sample)	N.D.	1		51
Subarctic lakes[c] (pH 1), Yellowknife, NWT, Canada	N.D.–0.5	N.D.–0.7	N.D.–0.2	74
Subarctic lakes[c] (pH 6), Yellowknife, NWT, Canada	N.D.–0.04	N.D.	N.D.	74
Lake Biwa, Japan	<0.05 <0.01 (MMA^{III})	<0.76 <0.01(DMA^{III})		64
Lake Pavin, France (measurements at different depths)	Not quantified	Not quantified		53
Seawater				
Pacific Ocean (<100 m)	0.009–0.02	0.02–0.2		39,96
Western Atlantic Ocean—upper 200 m	0.007	0.05		45
Antarctic Ocean	0.007	0.02		54
East Indian Ocean	0.03	0.05		54
North Indian Ocean	0.03	0.2		54
Baltic Sea	N.D.–0.03	N.D.–0.52		43
Southern North Sea (over several months)	0–0.1	0.05–0.3		65
China Sea	0.02	0.08		54
Indonesian Archipelago	0.03	0.09		54
Estuaries and Coastal Waters				
Scripps Pier, La Jolla, California, USA	0.01–0.02	0.12		93
San Diego Trough, California, USA (surface to 100 m below surface)	0.003–0.005	0.002–0.21		93
Suisun Bay, California, USA	0.1	0.07		20
Southern California Bight (varying depths and locations)	<0.002–0.031	0.01–0.26		49
Tidal flat, Tampa, Florida, USA	0.08	0.3		92
Bay, Causeway, Tampa, Florida, USA	N.D.	0.2		92
McKay Bay, Tampa, Florida, USA	0.07	1.0		92
Chesapeake Bay[c], Maryland, USA	N.D.–0.42	N.D.–0.34		31
Patuxent River Estuary, Maryland, USA (over 2 years)(max. values)	<0.3	0.2–0.6		27
Hastings Arm[c] porewaters, British Columbia, Canada	<0.03	0.09–0.23	<0.22 (TMAO)	30
Alice Arm[c] porewaters, British Columbia, Canada	<0.47	<0.19	<0.41 (TMAO)	30

Table 3. (*Continued*)

Source	MMA	DMA	TMA	Reference
Quatsino Sound^c porewaters (one station), British Columbia, Canada	<0.18	<0.22	<0.70 (TMAO)	30
Holberg Inlet^c porewaters (one station), British Columbia, Canada	<0.04	<0.04	<0.27 (TMAO)	30
Rupert Inlet^c porewaters (two stations), British Columbia, Canada	<0.33	<0.24	<0.25 (TMAO)	30
Surface waters of the Beaulieu Estuary, Hampshire, UK	<0.02–0.32	<0.02–0.44		17
Deep waters of the Beaulieu Estuary, Hampshire, UK	0.03–0.26	0.19–0.40		17
Tamar Estuary^c, southwest England	0.2–0.5	0.02–1.3		63
Tamar Estuary^c (porewaters), southwest England	0.4	0.5		63
Tamar Estuary^c (porewaters), southwest England	0.04–0.7	0.1–0.5		97
Southampton Water, UK	0.03–0.08	0.05–0.5		26
Humbar Estuary^c, UK (over four seasons)	N.D.–0.02	N.D.–0.2		18
Thames Estuary and plume^c, UK (Feb.1989)	N.D.	N.D.		23
Thames Estuary and plume^c, UK (July 1990)	0.036 ± 0.016	0.138 ± 0.127		23
Seine Estuary, France (over several months)	N.D.–0.05	N.D.–0.3		22
Continental shelf off the Gironde Estuary, France	N.D.	0–0.1		29
Tagus Estuary^c, Atlantic coast of Europe	0.01–0.06	0.07–0.2	0.01–0.04	67
Other				
Rainwater, La Jolla, California, USA	<0.002	<0.002–0.024		93

N.D.: not detected.
TDA: Total dissolved arsenic.
^{a}All concentrations in μg/L.
^{b}The sum of $As^{III} + As^V$.
^{c}Anthropogenic contamination.

be arsenic/sulfur compounds (40,74,79). Arsenic/sulfur compounds could dominate in reducing environments (4,34). The presence of arsenothiols in lake sediment porewater has been suggested (74). The presence of oxythioarsenate, $[H_3As^VO_3S]$, in water from an arsenic-rich, reducing environment has been demonstrated (80).

CONCLUDING REMARKS

Ingestion of arsenic from drinking water is a risk factor for several cancers and noncancerous health effects. Arsenic occurs naturally in the environment. Both inorganic and methylated arsenic species have been detected in aquatic systems. Inorganic As^V and As^{III} are predominant in most natural waters; methylated arsenic species occur at lower concentrations. Arsenothiols and uncharacterized arsenic species are also present in some aqueous environments. The relative abundance of various arsenic species in natural waters depends on both biological activity and chemical parameters, such as redox potential and pH.

BIBLIOGRAPHY

1. Ferguson, J.F. and Gavis, J. (1972). *Water Res.* **6**: 1259–1274.
2. Welch, A.H., Lico, M.S., and Hughes, J.L. (1988). *Ground Water* **26**: 333–347.
3. Tamaki, S. and Frankenberger, W.T. (1992). *Rev. Environ. Contam. Toxicol.* **124**: 79–110.
4. Cullen, W.R. and Reimer, K.J. (1989). *Chem. Rev.* **89**: 713–764.
5. Cherry, J.A., Shaikh, A.U., Tallman, D.E., and Nicholson, R.V. (1979). *J. Hydrol.* **43**: 373–392.
6. Andreae, M.O. (1979). *Limnol. Oceanogr.* **24**: 440–452.
7. Korte, N.E. and Fernando, L. (1991). *Crit. Rev. Environ. Control* **18**: 137–141.
8. National Research Council. (1999). *Arsenic in Drinking Water*. National Academy Press, Washington, DC.
9. Nriagu, J.O. (Ed.). (1994). *Arsenic in the Environment, Part I: Cycling and Characterization*. John Wiley & Sons, Hoboken, NJ.
10. Welch, A.H. (2000). *Ground Water* **38**: 589–604.
11. World Health Organization. (2001). *Environmental Health Criteria 224: Arsenic and Arsenic Compounds*, 2nd Edn. World Health Organization, Geneva, pp. 1–108.
12. Le, X.C. (2001). *Environmental Chemistry of Arsenic*. W.T. Frankenberger Jr. (Ed.). Marcel Dekker, New York, pp. 95–116.
13. IARC. (1987). International Agency for Research on Cancer, Suppl., Vol. 23. http://193.51.164.11/htdocs/Monographs/Suppl7/Arsenic.html.
14. National Research Council. (2001). *Arsenic in Drinking Water: 2001 Update*. National Academy Press, Washington, DC.
15. USEPA (U.S. Environmental Protection Agency) Federal Register. (2001). *National Primary Drinking Water Regulations. Arsenic and Clarifications to Compliance and New Source Contaminants Monitoring. Final Rule*, 40 CFR Parts 9, 141 and 142.
16. Mok, W.M. and Wai, C.M. (1994). *Arsenic in the Environment, Part I: Cycling and Characterization*. J.O. Nriagu (Ed.). John Wiley & Sons, Hoboken, NJ, pp. 99–117.
17. Howard, A.G., Arbab-Zavar, M.H., and Apte, S.C. (1984). *Estuar. Coast. Shelf Sci.* **19**: 493–504.
18. Kitts, H.J., Millward, G.E., Morris, A.W., and Ebdon, L. (1994). *Estuar. Coast. Shelf Sci.* **39**: 157–172.
19. Sanders, J.G., Riedel, G.F., and Osman, R.W. (1994). *Arsenic in the Environment, Part. I: Cycling and Characterization*. J.O. Nriagu (Ed.). John Wiley & Sons, Hoboken, NJ, pp. 289–308.
20. Anderson, L.C.D. and Bruland, K.W. (1991). *Environ. Sci. Technol.* **25**: 420–427.

21. Kuhn, A. and Sigg, L. (1993). *Limnol. Oceanogr.* **38**: 1052–1059.
22. Michel, P., Chiffoleau, J.F., Averty, B., Auger, D., and Chartier, E. (1999). *Cont. Shelf Res.* **19**: 2041–2061.
23. Millward, G.E., Kitts, H.J., Ebdon, L., Allen, J.I., and Morris, A.W. (1997). *Mar. Environ. Res.* **44**: 51–67.
24. Aurillo, A.C., Mason, R.P., and Hemond, H.F. (1994). *Environ. Sci. Technol.* **28**: 577–585.
25. Spliethoff, H.M., Mason, R.P., and Hemond, H.F. (1995). *Environ. Sci. Technol.* **29**: 2157–2161.
26. Howard, A.G. and Apte, S.C. (1989). *Appl. Organomet. Chem.* **3**: 499–507.
27. Riedel, G.F. (1993). *Estuaries* **16**: 533–540.
28. Sohrin, Y., Matsui, M., Kawashima, M., Hojo, M., and Hasegawa, H. (1997). *Environ. Sci. Technol.* **31**: 2712–2720.
29. Michel, P., Boutier, B., Herbland, A., Averty, B., Artigas, L.F., Auger, D., and Chartier, E. (1997). *Oceanol. Acta* **21**: 325–333.
30. Reimer, K.J. and Thompson, J.A.J. (1988). *Biogeochemistry* **6**: 211–237.
31. Sanders, J.G. (1985). *Mar. Chem.* **17**: 329–340.
32. Howard, A.G., Arbab-Zavar, M.H., and Apte, S.C. (1982). *Mar. Chem.* **11**: 493–498.
33. Peterson, M.L. and Carpenter, R. (1983). *Mar. Chem.* **12**: 295–321.
34. Francesconi, K.A. and Edmonds, J.S. (1997). *Adv. Inorg. Chem.* **44**: 147–189.
35. Maeda, S. (1994). *Arsenic in the Environment, Part I: Cycling and Characterization*. J.O. Nriagu (Ed.). John Wiley & Sons, Hoboken, NJ, pp. 155–187.
36. Francesconi, K.A. and Edmonds, J.S. (1994). *Arsenic in the Environment, Part I: Cycling and Characterization*. J.O. Nriagu (Ed.). John Wiley & Sons, Hoboken, NJ, pp. 221–260.
37. Sanders, J.G. (1980). *Mar. Environ. Res.* **3**: 257–266.
38. Johnson, D.L. and Burke, R.M. (1978). *Chemosphere* **8**: 645–648.
39. Santosa, S.J., Mokudai, H., Takahashi, M., and Tanaka, S. (1996). *Appl. Organomet. Chem.* **10**: 697–705.
40. Reimer, K.J. (1989). *Appl. Organomet. Chem.* **3**: 475–490.
41. Sanders, J.G. and Riedel, G.F. (1993). *Estuaries* **16**: 521–532.
42. Andreae, M.O. and Klumpp, D. (1979). *Environ. Sci. Technol.* **13**: 738–741.
43. Andreae, M.O. and Froelich, P.N. (1984). *Tellus* **36B**: 101–117.
44. Robertson, F.N. (1989). *Environ. Geochem. Health* **11**: 171–176.
45. Cutter, G.A., Cutter, L.S., Featherstone, A.M., and Lohrenz, S.E. (2001). *Deep Sea Res. II* **28**: 2895–2915.
46. Korte, N.E. (1991). *Environ. Geol. Water. Sci.* **18**: 137–141.
47. Hering, J.G. and Chiu, V.Q. (2000). *J. Environ. Eng.* 471–474.
48. Chen, S.L., Dzeng, S.R., Yang, M.H., Chiu, K.H., Shieh, G.M., and Wai, C.M. (1994). *Environ. Sci. Technol.* **28**: 877–881.
49. Andreae, M.O. (1978). *Deep Sea Res.* **25**: 391–402.
50. Andreae, M.O. and Andreae, T.W. (1989). *Estuar. Coast. Shelf Sci.* **29**: 421–433.
51. Crecelius, E.A. (1975). *Limnol. Oceanogr.* **20**: 441–451.
52. Cutter, G.A. (1992). *Mar. Chem.* **40**: 65–80.
53. Seyler, P. and Martin, J.M. (1989). *Environ. Sci. Technol.* **23**: 1258–1263.
54. Santosa, S.J., Wada, S., and Tanaka, S. (1994). *Appl. Organomet. Chem.* **8**: 273–283.
55. Raessler, M., Michalke, B., Schramel, P., Schulte-Hostede, S., and Kettrup, A. (1998). *Int. J. Environ. Anal. Chem.* **72**: 195–203.
56. Edwards, L., Patel, M., McNeill, L., Frey, M., Eaton, A.D., Antweller, R.C., and Taylor, H.E. (1998). *J. Am. Water Works Assoc.* **90**: 103–113.
57. Borho, M. and Wilderer, P. (1997). *SRT Aqua* **46**: 138–143.
58. Gallagher, P., Schwegel, C.A., Wei, X., and Creed, J.T. (2001). *J. Environ. Monitoring* **3**: 371–376.
59. Hall, G.E.M. and Pelchat, J.C. (1999). *J. Anal. At. Spectrom.* **14**: 205–213.
60. Lindemann, T., Prange, A., Dannecker, W., and Neidhart, B. (2000). *Fresenius J. Anal. Chem.* **368**: 214–220.
61. Kim, M.J. (2001). *Bull. Environ. Contam. Toxicol.* **67**: 46–51.
62. Le, X.C., Yalcin, S., and Ma, M. (2000). *Environ. Sci. Technol.* **24**: 2342–2347.
63. Howard, A.G., Apte, S.C., Comber, S.D.W., and Morris, R.J. (1988). *Estuar. Coast. Shelf Sci.* **27**: 427–433.
64. Hasegawa, H. (1997). *Appl. Organometal. Chem.* **11**: 305–311.
65. Millward, G.E., Kitts, H.J., Comber, S.D.W., Ebdon, L., and Howard, A.G. (1996). *Estuar. Coast. Shelf Sci.* **43**: 1–18.
66. Hasegawa, H. (1996). *Appl. Organometal. Chem.* **10**: 733–740.
67. de Bettencourt, A.M. (1988). *Neth. J. Sea Res.* **22**: 205–212.
68. Roussel, C., Bril, H., and Fernandez, A. (2000). *J. Environ. Qual.* **29**: 182–188.
69. Chen, H.W., Frey, M.M., Clifford, D., McNeill, L.S., and Edwards, M. (1999). *J. Am. Water Works Assoc.* **91**: 74–85.
70. Waslenchuk, D.G. and Windom, H.L. (1978). *Estuar. Coast. Shelf Sci.* **7**: 455–464.
71. Mucci, A., Richard, L.F., Lucotte, M., and Guignard, C. (2000). *Aqua. Geochem.* **6**: 293–324.
72. Kim, M.J., Nriagu, J.O., and Haack, S. (2000). *Environ. Sci. Technol.* **34**: 3094–3100.
73. Abdel-Moati, A.R. (1990). *Water Air Soil Pollut.* **51**: 117–132.
74. Bright, D.A., Dodd, M., and Reimer, K.J. (1996). *Sci. Total Environ.* **180**: 165–182.
75. Hasegawa, H., Matsui, M., Okamura, S., Hojo, M., Iwasaki, N., and Sohrin, Y. (1999). *Appl. Organomet. Chem.* **13**: 113–119.
76. Sturgeon, R.E., Siu, K.W.M., Willie, S.N., and Berman, S.S. (1989). *Analyst* **114**: 1393–1396.
77. de Bettencourt, A.M. and Andreae, M.O. (1991). *Appl. Organomet. Chem.* **4**: 111–116.
78. Howard, A.G. and Comber, S.D.W. (1989). *Appl. Organomet. Chem.* **3**: 509–514.
79. Bright, D.A., Brock, S., Cullen, W.R., Hewitt, G.M., Jafaar, J., and Reimer, K.J. (1994). *Appl. Organomet. Chem.* **8**: 415–422.
80. Schwedt, G. and Rieckhoff, M. (1996). *J. Chromatogr. A* **736**: 341–350.
81. Mok, W.M., Riley, J.A., and Wai, C.M. (1988). *Water Res.* **22**: 769–774.
82. Boyle, D.R., Turner, R.J.W., and Hall, G.E.M. (1998). *Environ. Geochem. Health* **20**: 199–212.

83. Chatterjee, A., Das, D., Mandal, B.K., Chowdhury, T.R., Samanta, G., and Chakraborti, D. (1995). *Analyst* **120**: 643–650.
84. Tanabe, K., Yokota, H., Hironaka, H., Tsushima, S., and Kubota, Y. (2001). *Appl. Organomet. Chem.* **15**: 241–251.
85. Nag, J.K., Balaram, V., Rubio, R., Alberti, J., and Das, A.K. (1996). *J. Trace Element Med. Biol.* **10**: 20–24.
86. Chen, S.L., Yeh, S.J., Yang, M.H., and Lin, T.H. (1995). *Biol. Trace Element Res.* **48**: 263–274.
87. Lin, T.H., Huang, Y.L., and Wang, M.Y. (1998). *J. Toxicol. Environ. Health* **53**: 85–93.
88. Hung, T.C. and Liao, S.M. (1996). *Toxicol. Environ. Chem.* **56**: 63–73.
89. Del Razo, L.M., Arellano, M.A., and Cebrian, M.E. (1990). *Environ. Pollut.* **64**: 143–153.
90. Kondo, H., Ishiguro, Y., Ohno, K., Nagase, M., Toba, M., and Takagi, M. (1999). *Water Res.* **33**: 1967–1972.
91. Raessler, M., Michalke, B., Schulte-Hostede, S., and Kettrup, A. (2000). *Sci. Total Environ.* **258**: 171–181.
92. Braman, R.S. and Foreback, C.C. (1973). *Science* **182**: 1247–1249.
93. Andreae, M.O. (1977). *Anal. Chem.* **49**: 820–823.
94. Yalcin, S. and Le, X.C. (2001). *J. Environ. Monitoring* **3**: 81–85.
95. Kaise, T.M. et al. (1997). *Appl. Organomet. Chem.* **11**: 297–304.
96. Santosa, S.J., Wada, S., Mokudai, H., and Tanaka, S. (1997). *Appl. Organomet. Chem.* **11**: 403–414.
97. Ebdon, L., Walton, A.P., Millward, G.E., and Whitfield, M. (1987). *Appl. Organomet. Chem.* **1**: 427–433.
98. Wilkie, J.A. and Hering, J.G. (1998). *Environ. Sci. Technol.* **32**: 657–662.
99. Gihring, T.M., Druschel, G.K., McCleskey, R.B., Hamers, R.J., and Banfield, J.F. (2001). *Environ. Sci. Technol.* 3857–3862.
100. Yalcin, S. and Le, X.C. (1998). *Talanta* **47**: 787–796.

ARSENIC HEALTH EFFECTS

I. Katsoyiannis
A.I Zouboulis
Aristotle University of
Thessaloniki
Thessaloniki, Greece

INTRODUCTION

The toxicity of inorganic arsenic species depends mainly on their valence state (usually −3, +3, or +5) and also on the specific physical and chemical properties of the compounds in which it occurs. Trivalent (arsenite, As^{III}) compounds are generally more toxic than pentavalent (arsenate, As^{V}) compounds, whereas arsenic compounds that are more soluble in water usually are more toxic and more likely to present systemic effects, in comparison with less soluble compounds, which are more likely to cause chronic pulmonary effects, if inhaled. Among the most toxic inorganic arsenic compounds is arsine gas (AsH_3). Additionally, note that laboratory animals (commonly used for toxicity evaluation) are generally less sensitive than humans to the toxic effects of inorganic arsenic compounds.

The toxicity "scale" of arsenic compounds is as follows (in the order of decreased toxicity):

$$\text{Arsine} > \text{Arsenites} > \text{Arsenates} > \text{Organoarsenic compounds}$$

Humans are usually exposed to arsenic compounds primarily through (contaminated) air, food, and/or water sources. The concentration of arsenic in air is usually in the range of a few ng As/m^3, although the relevant exposure may be higher in highly polluted areas. The consumption of contaminated water for drinking is also an important source of arsenic exposure. The concentration of arsenic is generally higher in groundwater than in surface water, especially when and where specific geochemical conditions favor dissolution of arsenic minerals.

Due to the widespread distribution of this element in major environmental compartments, it should not be surprising that most people experience a (more or less) measurable arsenic intake each day. It has been reported that exposures of the general population to inorganic arsenic are between 5 and 30 μg/day, which come from air, food, and water. In particular, it has been estimated that humans can get around 0.05 μg As/day from the air, 1–10 μg As/day from (drinking) water, and 5–20 μg As/day from food. Another 1–20 μg As/day can be absorbed, when someone is also a smoker. The U.S. Food and Drug Administration (U.S. FDA) has calculated the mean daily intake of inorganic arsenic at about 0.5 μg/kg body weight of an adult; this corresponds to a range of 30–38 μg/day for adults weighing 60–75 kg. The Food and Agriculture Organization (FAO) and the World Health Organization (WHO) have estimated the safe daily doses or tolerable daily intakes of inorganic arsenic. Provisional values for adults are 2.1 μg As/kg body weight, which equals about 126–160 μg As/day for adults weighing 60–75 kg.

From the previous information, it becomes obvious that the recent (up to 1998) maximum concentration limit (MCL) of 0.05 mg/L of arsenic in drinking water is not protective enough. Considering that humans consume about 2–3 liters of water daily, when the water contains about 50 μg/L of arsenic, this results in a daily arsenic intake of 100–150 μg/day, in addition to another 30–38 μg/day, which has been calculated by the U.S. FDA; therefore, the total arsenic intake of humans, who drink water containing 50 μg/L of arsenic, is up to 130–188 μg/day. This might be higher than the safe daily dose for arsenic, as proposed by the WHO. As a result, the concentration limit of arsenic in drinking water has recently been lowered. In Europe, the respective limit, according to the E.C. Directive 98/83, is now 10 μg/L. In the United States, after several discussions of this issue, the arsenic limit in drinking water remained 50 μg/L up to January 22, 2001. Then, the U.S. Environmental Protection Agency (1) published a relevant final rule, which includes the revised standard for arsenic in drinking water, of 10 μg/L.

METABOLISM AND DISPOSITION

Once arsenic compounds are ingested, the soluble forms of arsenic are readily absorbed from the gastrointestinal (GI) tract. Arsenate, whether in inorganic or in organic form, is better absorbed than arsenite because it is less reactive with the membranes of the GI tract.

The absorption of water-soluble inorganic arsenic compounds through the GI tract is very high. In humans, absorption rates of 96.5% for trivalent sodium arsenite and 94% for soluble sodium arsenate have been reported. In contrast, the GI absorption of less soluble trisulfide arsenate was reportedly only 20–30% in hamsters. The absorption of arsenic in the lungs depends on the size of the particles onto which arsenic compounds have been deposited, as well as on their respective solubility; respirable particles (i.e., 0.1–1 μm in diameter) can be carried deeper into the lungs, and therefore they are more likely to be absorbed.

Once absorbed, the blood transports arsenic to different body organs, such as the liver, kidney, lung, spleen, aorta, and skin. It is worth noting that, except for skin, clearance from these organs is relatively rapid. Arsenic also deposits extensively in the hair and nails (2). Arsenic compounds are then subject to metabolic transformation. Pentavalent arsenic compounds are reduced to trivalent forms, and then they are methylated in the liver to the less toxic methylarsinic acids. Typical levels in the blood of people who are not exposed to significant sources of arsenic pollution are in the range of 1–5 μg As/L (3). Finally, arsenic can be removed (cleared) from the body relatively rapidly, primarily through urine. Urinary excretion rates of 80% within 61 h, following oral doses, and 30–80% within 4–5 days following parenteral doses, have been measured in humans (4). Arsenic can also be lost from the body through the hair and nails, because they represent a nonbiologically available arsenic pool. In Fig. 1, the main routes of the fate, distribution, and

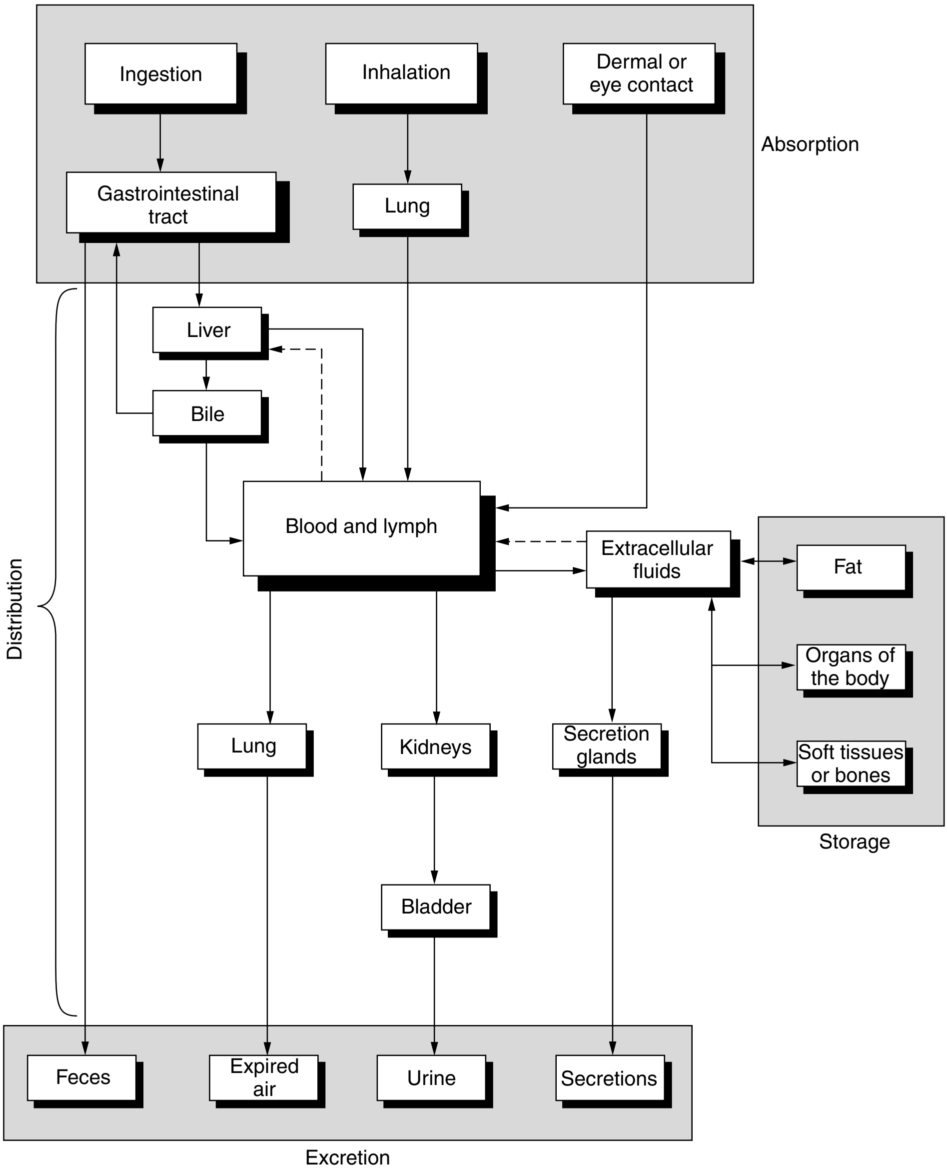

Figure 1. Fate, distribution, and excretion of toxic chemicals in the human body.

excretion of toxic chemicals, such as arsenic, for humans are presented.

MECHANISMS OF TOXICITY

The main inorganic arsenic species (i.e., As^{V} and As^{III}) have different mechanisms of action on which their toxicity depends. Arsenates behave similarly to phosphates. Consequently, they can substitute for phosphates in common cell reactions, whereas arsenites have high affinity for the thiol groups of proteins, causing inactivation of several enzymes.

In particular, the structural similarity of arsenates to phosphates allows them to substitute for phosphates in energy-producing reactions within the cell. First, arsenate can replace phosphate during glycolytic phosphorylation; if this occurs, glycolysis can continue, but ATP-producing reactions do not take place and, therefore, the cell produces less ATP. Second, arsenates can uncouple oxidative phosphorylation by substituting for phosphates in the ATP synthetase enzyme. As electrons are transferred to oxygen, ADP-arsenate (rather than ATP) is formed, which rapidly hydrolyzes. Hence, energy is wasted from the electron transport chain, because it cannot be stored appropriately. In addition, arsenate may also exert its toxic effects indirectly via reductive metabolism to arsenite.

The key to the toxicity of arsenite is its electrophilic nature; arsenite binds to electron-rich sulfydryl groups on proteins. Although such binding exerts adverse effects on structural proteins, such as the microfilaments and microtubules of the cytoskeleton in cultured cells, most of arsenite's toxicity is likely due to the inhibition of enzymes by binding to a thiol-containing active site. In particular, arsenite is known to inhibit enzymes of the mitochondrial citric acid cycle, and cellular ATP production decreases. Arsenite also uncouples oxidative phosphorylation. Arsenic's multipronged attack on the cell's energy production system can adversely affect cellular health and survival (5,6).

In contrast to inorganic arsenic, organoarsenic compounds, such as MMA or DMA, can bind strongly to biological molecules of humans, resulting in less toxic effects.

ARSENIC HEALTH EFFECTS

Acute arsenic exposure (i.e., high concentrations ingested during a short period of time) can cause a variety of adverse effects. The severity of effects depends strongly on the level of exposure. Acute high-dose oral exposure to arsenic typically leads to gastrointestinal irritation, accompanied by difficulty in swallowing, thirst, and abnormally low blood pressure. Death may also occur from cardiovascular collapse. The lethal dose to humans is estimated at 1–4 mg As/kg for an adult (3,7,8). Short-term exposure of humans to doses higher than 500 μg/kg/day can cause serious blood, nervous system, and gastrointestinal ill effects and also may lead to death.

Chronic arsenic uptake may have noncarcinogenic, as well as carcinogenic, effects on humans. Chronic exposure to low arsenic concentrations is of primary interest, when the health significance of arsenic in drinking water is evaluated. The most common signs of long-term, low-level arsenic exposure are dermal changes. These include variations in skin pigments, hyperkeratosis, and ulcerations. Vascular effects have also been associated with chronic arsenic exposure.

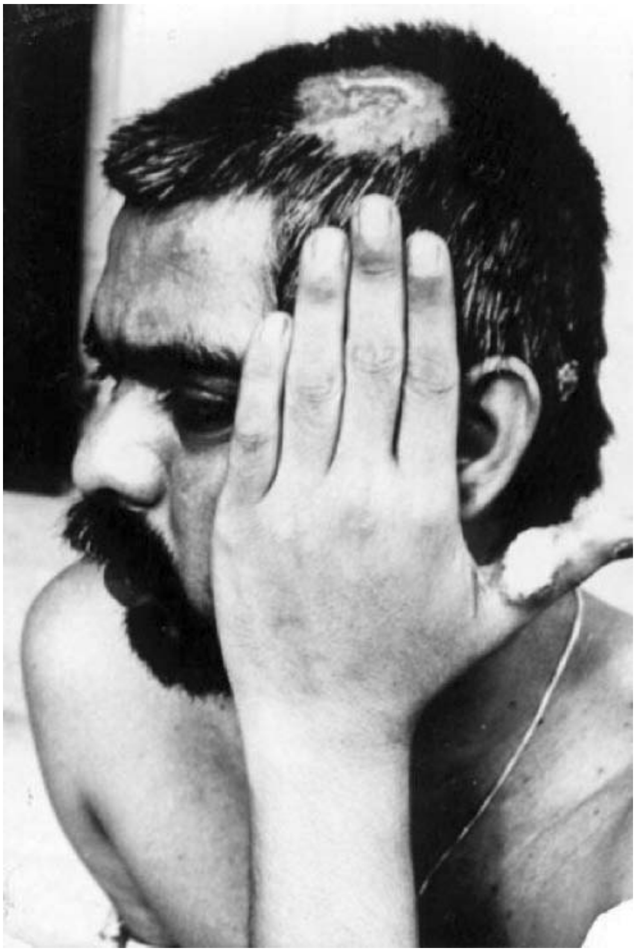

Figure 2. Skin cancer on the head caused by chronic arsenic intake.

Chronic arsenic exposure can also lead to carcinogenesis in humans. Arsenic is classified as a human carcinogen, according to the U.S. EPA. This classification was based mainly on human data because the respective animal data were inadequate. Several epidemiological studies have documented an association between chronic exposure to arsenic in drinking water and skin cancer. A correlation between chronic arsenic exposure and cancer in the liver, bladder, kidney, lung, and prostate has also been documented. Arsenic contamination is of major concern in several countries, such as Bangladesh, Chile, and Taiwan. Figures 2–5, taken of people suffering from arsenicosis in these countries, show the significance of the arsenic problem, mainly due to the consumption of contaminated drinking water, and indicate clearly the urgent need for effective treatment of groundwater to remove arsenic (3,7,8).

As can be noticed in these photos, chronic exposure to arsenic can cause skin cancer in several parts of the human body: on the head, on the hand, and on the foot. In India,

Figure 3. Skin cancer on the foot known as "black-foot" disease, caused by chronic arsenic intake.

Figure 4. Skin cancer on the hand.

in the area of West Bengal, more than 40 million people suffer from different kinds of arsenicosis, and 37% of the water samples, which have been analyzed, contain arsenic concentrations higher than 50 μg/L. In Bangladesh, more than 70 million people suffer from several kinds of cancer due to chronic exposure to arsenic. (9). This problem also exists in several areas of the United States, as well as in Europe. In the United States, more than 3 million people drink water whose arsenic concentrations are higher than 50 μg/L, whereas in Europe, arsenic concentrations over the 10 μg/L concentration limit have often been reported in Germany, Hungary, Finland, and Greece (7).

BIBLIOGRAPHY

1. U.S. EPA. Arsenic in drinking water: health effects research. Available at: www.epa.gov/OGWDW/ars/ars10.html.
2. U.S. Environmental Protection Agency. (1984). *Health Assessment Document for Arsenic*. Office of Health and Environmental Assessment—Environmental Criteria and Assessment Office, EPA Report 600/8-32-021F.
3. Pontius, F.W., Brown, K.G., and Chen, C.-J. (1994). Health implications of arsenic in drinking water. *J. Am. Water Works Assoc.* **86**(9): 52–63.

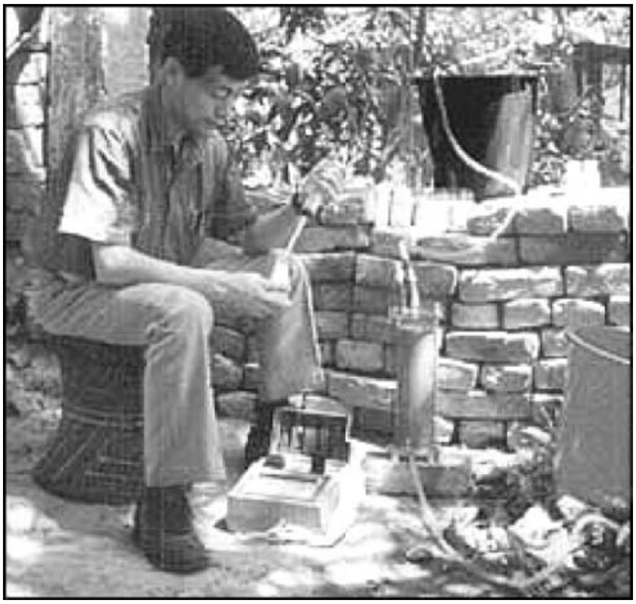

Figure 5. Scientist measuring arsenic in contaminated groundwater wells of Bangladesh. (Courtesy Stevens Institute of Technology.)

4. Crecelius, E.A. (1977). Changes in the chemical speciation of arsenic following ingestion by man. *Environ. Health Perspect.* **19**: 147–150.
5. Desesso, J.M. et al. (1998). An assessment of the developmental toxicity of inorganic arsenic. *Reprod. Toxicol.* **12**(4): 385–433.
6. Abernathy, C.O., Lin, Y.P., Longfellow, D., Aposhian, H.V., Beck, B., Fowler, B., Goyer, R., Menler, R., Rossman, T., Thompson, C., and Waalkes, M. (1997). Proceedings of the Meeting on Arsenic: Health Effects, Mechanisms of Action and Research Issues, Hunt Valley, MD, Sept. 22–24. *Environ. Health Perspect*. **107**(7): 593–597.
7. World Health Organization. Toxic effects of arsenic in humans. Available at: www.who.int/peh-super/Oth-lec/Arsenic/Series4/002.htm
8. Saha, J.C., Dikshit, A.K., Barley, M.K., and Saha, K.C. (1999). A review of arsenic poisoning and its effects on human health. *Crit. Rev. Environ. Sci. Technol.* **29**(3): 281–313.
9. Karim, M.M. (1999). Arsenic in groundwater and health problems in Bangladesh. *Water Res.* **34**: 304–310.

BACKGROUND CONCENTRATION OF POLLUTANTS

Mário Abel Gonçalves
Faculdade de Ciências da
Universidade de Lisoba
Lisoba, Portugal

The background concentration of an element or compound in a system may be considered as the averaged most common concentration defined from a set of representative samples. Traditionally, statistical techniques, such as moving averages and probability plots, are used to define these values quantitatively, which means that a statistically representative number of samples are required, and that samples represent the same general set of physical and chemical characteristics of the system. These approaches are most common to geochemical exploration surveys, through soil and rock analyses. Characterization of natural waters is based on the same principles, but the difficulties in achieving these principles are markedly different.

Defining background concentrations of pollutants in waters is of fundamental importance to establish water quality standards. Several pollutants occur in surface and groundwaters in a range of concentrations exclusively related to natural phenomena. However, anthropogenic activities induce large perturbations in most water systems. One major difficulty facing chemical characterization of surface and groundwater systems is their dynamic nature, which means that concentration patterns change with time. Once introduced in the water system, the fate and spatial distribution of a pollutant are determined by factors such as advection, diffusion, chemical reactivity, and biodegradation.

The large majority of the U.S. Environmental Protection Agency (EPA) list of priority pollutants refers to organic compounds that are solely due to human activities, and most of the inorganic pollutants occur in the natural waters at trace element concentrations (<1 mg/L). Thus, to define the background concentration of pollutants, one must distinguish those that occur naturally, with or without anthropogenic-induced disturbances, from those exclusively due to anthropogenic activities. Inorganic pollutants are clearly in the former group whereas organic compounds fall into the latter.

FACTORS DETERMINING WATER CHEMISTRY

Factors that influence the natural concentrations of major and minor elements in surface waters include the lithology, relief, climate, atmospheric dry and wet deposition, and human activities. Global river chemistry shows an enormous range in concentrations, ionic ratios, and proportion of ions in cation and anion sums spanning several orders of magnitude (1,2). It is unrealistic to define any kind of average of global water chemistry composition based on the large river basins. Besides, Maybeck (1) notes that river chemistry is very sensitive to alteration by human activities, and that, in the northern hemisphere, it is now difficult to find a medium-sized basin not significantly impacted by human activities. It must also be stressed that, between different chemical elements and species, different degrees of chemical reactivity exist, such as precipitation and incorporation into mineral phases, bonding to organic functional groups, and adsorption onto inorganic solid surfaces. These factors and the amount of colloids and suspended solid material in natural waters determine much of the element dispersion and concentration decay from source. Nriagu and Pacyna (3) estimated that human-induced mobilization of several trace metals far exceeds the natural fluxes, and enrichment factors in the range of 3–24 are reported for such elements as As, Cd, Pb, Se, and Hg.

Groundwaters are less sensitive to some of the external influences as the ones previously mentioned. However, understanding groundwater geochemistry is difficult because of the chemical heterogeneity of most aquifer systems. The residence time of groundwater is a factor that influences its final chemistry and depends on permeability and transmissivity and the amount of recharge of the aquifer. Groundwaters with long residence time in the reservoir become closer to chemical equilibrium with the minerals of the host rocks. In turn, low residence times of waters associated with the slow kinetics of silicate dissolution put severe constraints on groundwater chemistry. Nevertheless, cases where rapid ongoing dissolution and precipitation reactions are taking place, such as in some aquifers, result in relatively unaffected groundwater chemistry, irrespective of flow rate and direction. These are just some examples that illustrate the difficulty in defining acceptable background concentrations of chemical elements in several water systems.

BACKGROUND CONCENTRATION OF INORGANIC POLLUTANTS

Nitrate, ammonia, and trace metals are the most important inorganic pollutants present in waters. Several surface and groundwaters contain natural (background) concentrations of one or more of these and other chemical elements and species exceeding, for example, the U.S. EPA drinking water standards for reasons totally unrelated to human activities. Thus, identifying such natural sources is fundamental for regulatory decisions to avoid the assignment of unrealistic cleanup goals below such natural background concentrations. One of the most striking examples comes from groundwater arsenic contamination in Bangladesh (4), whose mechanisms, although attributable to iron oxide reductive dissolution, are still subject of intense debate. Of the millions of tube wells, about one-third have arsenic concentrations above the local drinking water standard (50 μg/L), and half of them do not meet the 10 μg/L guideline value of the World Health Organization (WHO) (4).

The approach to determine background concentrations of inorganic pollutants in water systems requires the sample of stream waters and/or groundwaters from boreholes in nearby areas in the same geological setting, where water chemistry has presumably not been affected by human activity. Regretfully, this approach is not always possible in several areas. In such cases, it might be reasonable to assume natural background concentrations equal to the measured concentrations in streams and groundwaters in the same general area and similar geologic environments, not forgetting that such factors as climate, relief, and aquifer recharge and hydraulic properties must also be evaluated. Sources of data on the "typical" chemical composition of groundwaters from different rock types also exist [see references in (5)]. However, the factors that influence water chemistry are varied and may differ between similar geological settings, and so the previous approaches are far from being reliable.

Use of cumulative probability plots of all the data for the area of interest can be a way to identify background concentrations, allowing the classification of samples into uncontaminated and contaminated groups (6). Some geostatistical techniques also aim to provide a separation between anomalous and background values in samples, especially factorial kriging (7). However, this does not avoid the need to have truly uncontaminated samples. Another shortcoming is that not all elements spread equally through a surface or groundwater system. Nitrates, for example, are highly mobile because they are not limited by solubility constraints. Adsorption onto solid phases is a mechanism that may rapidly remove trace metals from solution, such as Pb and Cd. Thus, what may be an uncontaminated water sample with respect to a given element may not hold in relation to other elements.

BACKGROUND CONCENTRATION OF ORGANIC POLLUTANTS

Organic pollutants are compounds that were mostly introduced in the environment by anthropogenic activities. Contrary to metals (the most important class of inorganic pollutants), several of these compounds are biodegradable, which, coupled to other mechanisms such as abiotic decomposition (including hydrolysis, oxidation reduction, and elimination), adsorption, dispersion, and dilution, contributes to what is called “natural attenuation.” Their source is varied and includes spilling and leakage from underground storage tanks and landfills, and several commercial and industrial activities, to name but a few. Background concentration of these pollutants is solely the result of human impact on the water system. In this case, organic pollutants should not overcome the maximum concentration level for drinking water standard in agreement with the various national regulations. For water quality assessment, it is important to have a good knowledge of organic pollutant dispersion and persistence in water systems, and the same basic principles presented also apply to sampling in this case. As a result of the previously mentioned mechanisms, these compounds have a complex behavior in water systems. Chlorinated solvents, for example, sometimes behave as conservative solutes that are rapidly transported, but they also undergo several microbial degradation processes causing them to rapidly disappear and be replaced by the lightly chlorinated ethenes. Some case studies describe the development of a dynamic steady state in plumes of hydrocarbons that stopped spreading because the rate of input of soluble hydrocarbons was balanced by biodegradation mechanisms that consumed hydrocarbons in the plume. Thus, knowledge of these processes is fundamental to evaluate and prevent organic pollutant concentrations to build up above levels considered potentially harmful to human and ecological health.

BIBLIOGRAPHY

1. Meybeck, M. (2004). Global occurrence of major elements in rivers. In: *Treatise on Geochemistry—Surface and Ground Water, Weathering, and Soils*. J.I. Drever (Ed.). Vol. 5, pp. 207–223, Elsevier, B.V.
2. Gaillardet, J., Viers, J., and Dupré, B. (2004). Trace elements in river waters. In: *Treatise on Geochemistry—Surface and Ground Water, Weathering, and Soils*. J.I. Drever (Ed.). Vol. 5, pp. 225–272, Elsevier, Canada.
3. Nriagu, J.O. and Pacyna, J.M. (1988). Quantitative assessment of worldwide contamination of air, water and soils by trace-metals. *Nature* **333**: 134–139.
4. BGS and DPHE. (2001). *Arsenic Contamination of Groundwater in Bangladesh*, D.G. Kinniburgh and P.L. Smedley (Eds.). British Geological Survey, Keyworth, England.
5. Langmuir, D. (1996). *Aqueous Environmental Geochemistry*. Prentice-Hall, Upper Saddle River, NJ.
6. Hann (1977). *Statistical Methods in Hydrology*. Iowa University Press, Iowa City, IA.
7. Wackernagel, M. (1995). *Multivariate Geostatistics*. Springer-Verlag, Berlin.

WATERBORNE BACTERIA

Barry H. Rosen
US Fish & Wildlife Service
Vero Beach, Florida

CHARACTERISTICS OF BACTERIA

Bacteria are so widespread that only the most general statements can be made about their life history and ecology. They are everywhere on the earth, even in the most hostile environments, including extremes of temperature, such as freshwater hot springs. Many live in water; however, they are also an important part of soils and live in and on plants and animals; others are parasites of humans, animals, and plants. Only a handful of waterborne bacteria cause diseases, usually by producing toxins. Most bacteria have a beneficial role in decomposing dead material and releasing nutrients back into the environment.

Bacteria are a group of microorganisms that lack membrane-bound organelles and hence are considered simpler than plant and animal cells. They are separated into gram-positive and gram-negative forms based on the staining properties of the cell wall. Most bacteria are unicellular and are found in various shapes: spherical (coccus), rod-shaped (bacillus), comma-shaped (vibrio), spiral (spirillum), or corkscrew-shaped (spirochete). Generally, they range from 0.5 to 5.0 micrometers in size. Motile species (those that can move on their own) have one or more fine flagella arising from their surface. Many possess an outer, slimy capsule, and some have the ability to produce an encysted or resting form (endospore). Bacteria reproduce asexually by simple division of cells and rarely by conjugation.

Bacteria have a wide range of environmental and nutritional requirements. One way to classify them is based on their need for oxygen. *Aerobic bacteria* thrive in the presence of oxygen and require it for continued growth and existence. *Anaerobic bacteria* thrive in oxygen-free environments, as often found in lake or wetland sediments. *Facultative anaerobes* can survive in either environment, although they prefer oxygen. The way bacteria obtain energy provides another means of understanding and classifying organisms. *Chemosynthetic bacteria* are autotrophic and obtain energy from the oxidation of inorganic compounds such as ammonia, nitrite (to nitrate), or sulfur (to sulfate). *Photosynthetic bacteria* convert sunlight energy into carbohydrate energy. Bacteria such as the purple sulfur bacteria, purple nonsulfur bacteria, green sulfur bacteria, and green bacteria have a bacterial form of chlorophyll. Cyanobacteria, commonly called blue-green algae, are a separate group that contains chlorophyll *a*, a pigment that is common to eukaryotic algae and higher plants. *Heterotrophic* bacteria form a diverse group that obtains energy from other organisms, either while they are alive (parasites) or dead (saprophytes).

BACTERIAL IDENTIFICATION

Bacterial size and a lack of visual cues prevent the identification of organisms by traditional microscopic techniques. Typical bacterial identification involves isolating organisms by culturing them on a variety of media that reveal the organism's physiological or biochemical pathways. For example, purple nonsulfur photosynthetic bacteria can be isolated from lake sediments by inoculating a mineral salt–succinate broth and incubating in light at 30 °C in a bottle sealed to ensure anaerobic conditions.

INDICATORS OF BACTERIAL/FECAL CONTAMINATION

One emerging water quality issue is contamination of surface and groundwater by bacteria and other microorganisms that are defined as pathogens—organisms that cause disease in animals or plants. The difficulty in direct detection of bacteria in water has led to the use of fecal bacteria as an indicator of the presence of pathogens and the risk of disease. Rapid, inexpensive techniques have been standardized for determining if fecal material has contaminated water, although the sources of that contamination are numerous. Wildlife, pets and companion animals, agricultural animals, and humans are all possible sources (Fig. 1).

Although indicator bacteria are not pathogenic themselves, high numbers may indicate fecal contamination from leaky septic tanks, animal manure, or faulty wastewater treatment facilities. Some species also live in soil and on plants and are harmless. Total coliform is the broadest category (Table 1) of indicator bacteria, and it was originally believed that it indicates the presence of fecal pollution. Numerous nonfecal sources make this indicator too generic. Fecal coliform, a subgroup of total coliform, originates from the intestinal tract of warm-blooded animals. This subgroup is the most commonly used indicator of bacterial pollution in watersheds. *Escherichia coli* is a member of the fecal coliform subgroup. This subgroup is used because it correlates well with illness from swimming and can cause gastrointestinal problems. Fecal streptococci, also called fecal strep, are another grouping of bacteria, similar to the coliforms that are associated with feces from warm-blooded animals. Enterococci are a subgroup of fecal strep bacteria. This subgroup is used because it correlates well with human illness from recreational waterbodies. Enterococci and *E. coli* are considered to have a higher degree of association with outbreaks of gastrointestinal illness than fecal coliforms, as indicated by the U.S. Environmental Protection Agency. State and local government agencies commonly monitor for total or fecal coliform, and some monitor *E. coli* and enterococci. States and local health agencies may have more stringent standards than national guidelines (Table 1).

Escherichia coli O157:H7 is a potentially deadly fecal bacteria that can cause bloody diarrhea and dehydration in humans. The combination of letters and numbers in the name of this bacterium refers to the specific molecular markers found on its cell surface; they distinguish it from other types of *E. coli,* most of which are part of the normal bacterial flora in warm-blooded animal intestinal tracts. This organism does pose a threat to bathers or others from bodily contact in contaminated waters, although the majority of outbreaks are from contaminated food.

BACTERIAL SOURCE TRACKING

DNA fingerprinting, one tool for bacterial source tracking (BST), consists of a family of techniques that are under development to identify the sources of fecal contamination in various waterbodies. The goal is to identify the source, whether human, livestock, pets and companion animals, or wildlife. In theory, this allows targeting management activities for bacterial reduction on the appropriate sources.

The EPA periodically reports data submitted by states on impairments to rivers, lakes, and estuaries. Bacteria and pathogens, based on fecal indicator bacteria data collected by states, frequently exceed water quality criteria. An individual waterbody or segment of a river that exceeds these criteria is listed by the state [EPA's 303(d) list] and may trigger additional efforts to improve the waterbody. One mechanism for improving water quality that is currently being used by the EPA is the total maximum daily load (TMDL). Microbial source tracking techniques are currently the best technology for tracking sources of fecal contamination and can play an important role in TMDL development. For example, if fecal contamination is a major issue in a particular watershed, BST might show that the contamination originates from humans, not livestock. The TMDL should then reflect an appropriate proportion of reduction in bacterial loading from agriculture and from urban areas (leaky septic systems or ineffective sewage treatment plants).

Understanding the concept of how BST works, including its limitations, can help determine if these techniques can be used in a particular situation. Although DNA fingerprinting has received the greatest amount of attention recently, other BST methods are in use and can play an important part in fecal source tracking.

For most of these techniques, bacteria from known sources (humans, swine, raccoons, deer, cows, etc.) are collected directly from the animal, then isolated and grown

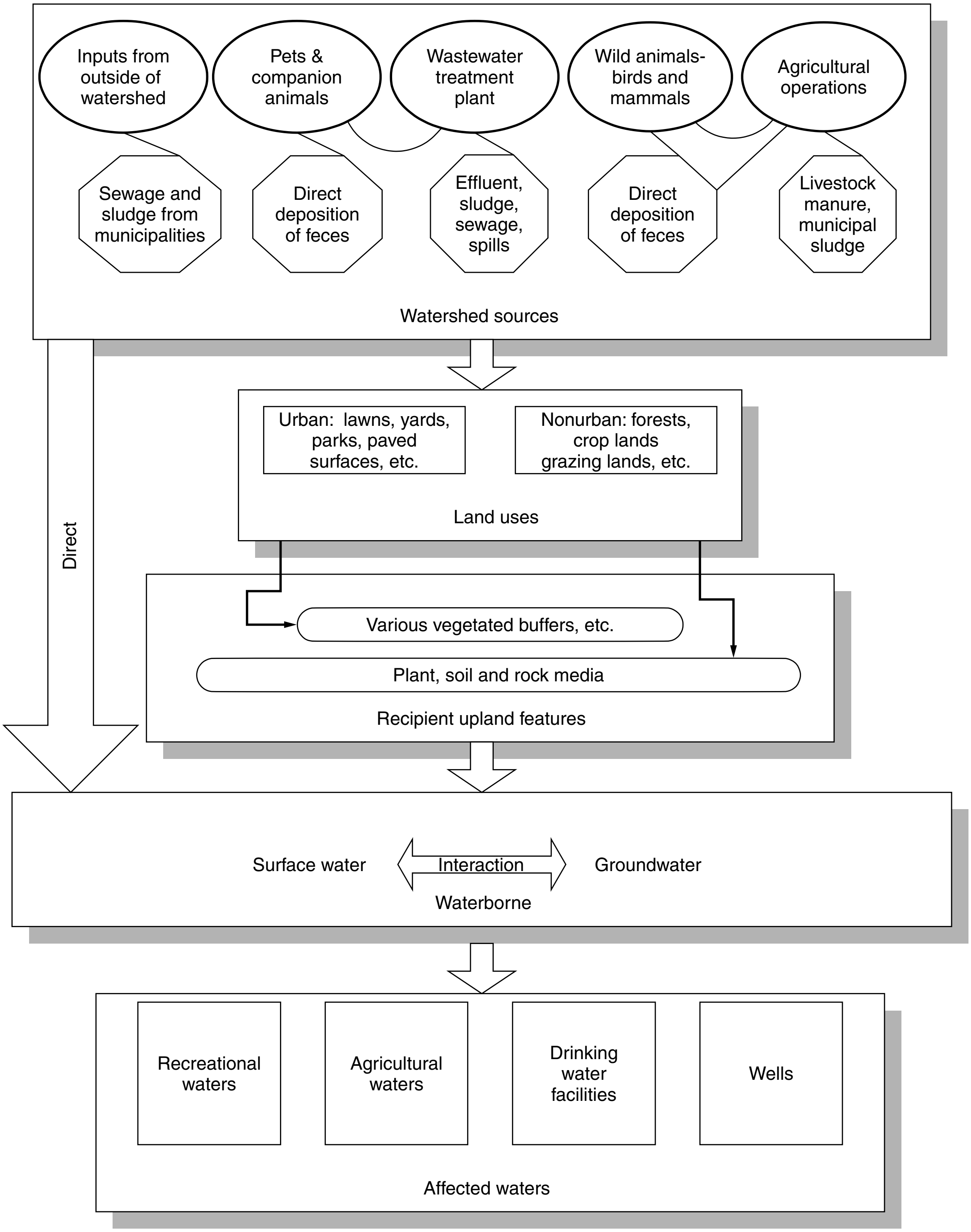

Figure 1. Potential pathways for pathogen movement into water.

in the laboratory. One species of bacteria, *E. coli*, has been extensively used, but others are more likely to be found in particular situations. For example, fecal streps (e.g., *Streptococcus feacalis*) are very numerous compared to *E. coli* in composted poultry litter or in high quality biosolids. These isolates grown in the laboratory are the basis of the library or database for subsequent comparison to unknown samples.

The current state of research normally involves isolating and culturing as many organisms as practical. Financial constraints usually limit this approach; several hundred may be needed from each major source to identify sufficiently the source in a water sample. In addition, regional differences in the genetics of isolates are just coming to light, potentially restricting the broader use of genetic libraries.

Table 1. Comparison of Fecal Bacteria Water Quality Indicators Commonly Used

Microbial Indicator	Properties	Federal Standard[a,b]
Total coliforms (TC)	• originally believed to indicate the presence of fecal pollution • widely distributed in nature: soils, water, flora, fauna • contains members of *Escherichia, Citrobacter, Klebsiella*, and *Enterobacter* identified by incubation at 35 °C	1000 CFU/100 mL
Fecal coliforms (FC)	• subgroup of TC • coliforms that originate specifically from intestinal tracts of warm-blooded animals • cultured by increasing the incubation temperature to 44.5 °C • remains the predominant indicator used to assess bacterial pollution in watersheds	200 CFU/100 mL
Escherichia coli	• member of the FC group • presence correlates with illness from swimming in both fresh and marine waters • has been shown epidemiologically to cause gastrointestinal symptoms • O157:H7 is a toxin-producing strain of this common bacterium	126 CFU/100 mL
Enterococci	• subgroup of the FC, including *Streptococcus faecalis, S. faecium,* and *S. avium* • commonly found in intestinal tracts of humans and other warm-blooded animals • presence correlates well with illness from both fresh and marine waters	33 CFU/100 mL

[a] Individual states may have higher standards but not lower; CFU = colony forming units.

[b] Primary contact water includes recreational use such as swimming and fishing.

CURRENT TECHNIQUES UNDER DEVELOPMENT

Molecular Methods (Genotype)

The overall intestinal environment in each type of animal is different enough to allow selective pressures on the microbial flora, resulting in populations of bacteria that have slightly different genetics from each species. The purpose of these molecular techniques is to find that difference among the genetics of the same type of bacteria isolated from these different animals. Commonly used methods, such as restriction endonuclease and polymerase chain reaction (PCR) used to amplify DNA, are sophisticated molecular techniques that rely on DNA extracted from isolates that have been cultured and purified before analysis.

Ribosomal RNA is the target of several molecular techniques because this portion of the bacterial genome is considered stable. The techniques associated with target ribosomal RNA are called “ribotyping.” Other molecular methods target parts of the bacterial DNA and may use pulse field gel electrophoresis and randomly amplified polymorphic DNA—a technique that amplifies selected portions of the DNA by PCR. Highly qualified personnel trained on rather elaborate equipment must perform all of these molecular techniques, currently limiting the availability and cost of these procedures.

Biochemical Methods (Phenotype)

Biochemical methods have some advantages over molecular techniques in cost and efficiency. Molecular techniques are more precise, but more time-consuming, and are not suitable for analyzing large numbers of samples at this time. A combination of the two approaches may be best, the simpler, quicker, less costly biochemical approach followed by detailed molecular analysis of selected bacterial isolates to confirm the results.

One technique that falls under biochemical methods is antibiotic resistance analysis. The same principle that applied to the discussion of the intestinal tract selecting certain genetics in bacterial flora applies to microbes that acquire resistance to antibiotics or other biochemical traits. Each source is expected to have its own pattern of resistance; the general concept is that human fecal bacteria have developed the greatest diversity of resistance. Recent findings, however, have shown that domestic animals that receive antibiotics are hosts of antibiotic resistant bacteria and these bacteria can be found in waterbodies adjacent to farms.

Other biochemical methods that are being used include the F-Specific (F+ or FRNA) coliphage (although this is targeted at viral sources), sterols or fatty acid analysis from the cell walls and membranes of *E. coli*, and nutritional patterns.

STATUS OF THESE TECHNIQUES

The current set of techniques can distinguish between humans and animal sources. Separation of wildlife from domestic animals is also successful, although less accurate than human–animal separation. Distinguishing different types of domestic animals or different types of wildlife from one another is still under development but likely in the near future.

All of these techniques should be considered under development. Studies are in progress with more than one technique applied to the same set of conditions to determine the extent of similarity, effectiveness, cost, strengths, and weaknesses. Many site-specific studies have proved effective and valuable; however, regional variation needs to be addressed to move beyond these site-specific findings. Most likely, a combination of methods will be useful, based on the particular situation. Other nonbacterial techniques may also indicate the presence of wastewater, such as optical brighteners from laundry detergents and caffeine. Collectively, BST techniques will prove valuable for tracing fecal pollution from a variety of sources. State and local public health officials are increasingly acquiring this technology and may be good resources for information applicable on areawide or watershed scales.

WATER ASSESSMENT AND CRITERIA

Parley V. Winger
University of Georgia
Atlanta, Georgia

INTRODUCTION

Human activities have historically altered aquatic systems to some extent; however, with the Industrial Revolution and the shift from hunting and agricultural use of the land to amassing in population centers, pollution of adjacent waterways reached intolerable levels (1,2). Untreated municipal and industrial wastes were released directly into river, stream, and lake systems, and even with the development of sewage treatment in the mid-1800s, sewage outfalls continued to reduce the quality of downstream waters (1). Federal intervention into activities influencing water quality began in the late nineteenth century (e.g., River and Harbors Act of 1899) and continues today, primarily under the auspices of the Clean Water Act and its amendments (3,4). With the advent of legislative regulation for water quality, procedures for measuring and describing the status and quality of aquatic systems were needed for management and decision making. Developing from this need has been a progressive advancement of practices and procedures used to describe and categorize the quality of aquatic environments, culminating in the present-day regional approach to biological assessment and the development of biological criteria that can be used to evaluate ecosystem health and classify the biological integrity of aquatic systems. Biological assessment is defined as "an evaluation of the condition of a waterbody using biological surveys and other direct measurements of the resident biota in surface waters," and biological criteria are "numerical values or narrative expressions that describe the reference biological condition of aquatic communities inhabiting waters of a given designated aquatic life use and are used as benchmarks for water resources evaluation and management decision making" (5). A mandate of the Clean Water Act was to restore and maintain biological integrity of aquatic systems. Biological integrity encompasses all factors affecting an ecosystem and is defined as the "capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of the natural habitat of the region" (6).

EARLY ASSESSMENTS

An extensive history and subsequent evolution of concepts, procedures, and understanding of processes associated with assessments of aquatic environments in relation to anthropogenic impacts exist (2). Basic to the development of appropriate and effective bioassessment procedures was the recognition that aquatic systems are more than just water and that the biological components represented by the myriad of microbes, plants, and animals are integral to these systems in carrying out the biological processes that ensure water quality. The diverse biological assemblages that reside in aquatic habitats, although generally unique for each system, are characteristic of the indigenous water quality. Early studies and measurements of water quality relied primarily on the chemical characteristics (pH, dissolved oxygen, etc.) of the water, but it was eventually recognized that these measurements provided little insight into the biological health of these bodies of water. Maintenance of clean, healthy, high-quality water is dependent on diverse and fully functional aquatic communities. Conversely, the biological components present in an aquatic system reflect the quality of the water and habitat.

DEVELOPMENT OF BIOASSESSMENT PROCEDURES

A number of factors and activities exist that can adversely impact aquatic systems. Aside from the general features that provide the underlying control of the structure and function of rivers and streams (climate, geology, soil types, etc.), anthropogenic activities can have marked influences on aquatic environments (7). Environmental factors (e.g., water quality, habitat structure, flow regime, energy source, and biotic interactions) influence the biological integrity of aquatic systems and anthropogenic activities generally influence one or more of these factors (5). Physical alterations of the habitat (e.g., agriculture, logging, channelization, flow alterations) can affect erosion, sedimentation, and hydrological characteristics of streams and rivers, and these, in turn, can drastically alter the biological communities. Inputs of contaminants and other pollutants from agricultural runoff and municipal and industrial effluents can also impact the biological components, but generally in a different manner from that elicited by physical disturbance. Similarly, different types of contaminants can adversely impact aquatic communities in unique ways (e.g., impacts from organic pollution are different from that shown by acid-mine drainage or an industrial chemical). Consequently, bioasssements need to be robust enough to be able to detect impairment of aquatic communities, regardless of the cause.

Although early assessments associated with organic pollution relied heavily on chemical analyses (basic water chemistry such as dissolved oxygen), the adverse effects of pollution on aquatic organisms were understood by the mid-1800s (4). However, most early documentation of the biological impacts from pollution in aquatic systems was qualitative. Results from biological measurements were generally lengthy lists of “indicator” organisms associated with zones of pollution or purification. For example, Kolkwitz and Marrson (8,9) identified the zones of purification/decomposition (“saprobia”) downstream from sewage outfalls as poly-, meso-, and oligosaprobia, with each zone characterized by a unique assemblage of organisms.

The use of indicator organisms in categorizing water quality provided a framework for focusing on the biological components of aquatic systems and for the development of more robust bioassessment procedures (10). As “indicator organisms” were more descriptive of organic pollution (sewage) and the tolerances of organisms to the many types of pollution (sewage, metals, industrial contaminants, pesticides, etc.) were not well known, there was an impetus for development of bioassessment programs that incorporated community-based approaches (11,12). Patrick (13) pioneered the concept of using aquatic communities in stream assessments and compared histograms depicting the numbers of species of various taxonomic groups (blue-green algae, oligochaetes and snails, protozoans, diatoms and green algae, rotifers and clams, insects and crustaceans, and fish) comprising the communities in test streams with those from a clean or reference site. Patrick recognized that healthy streams had a great number of species representing the various taxonomic groups and no species were represented by a great number of individuals, whereas streams impacted by pollution had reduced numbers of species and those remaining were in great abundance.

A major drawback, however, of including most of the various biological components (algae, protozoa, benthic macroinvertbrates, and fishes) in these types of bioassessments was the amount of time required and the expertise needed to collect and identify all of the organisms comprising the aquatic communities at sites of concern. As a result of this criticism, aquatic scientists began to specialize or use specific taxonomic groups in bioassessments as representative components or surrogates of the biotic communities (14). Of course, each specialist considered their particular taxonomic group as the most appropriate for use in bioassessments, and for the most part, although each group had its advantages and disadvantages, they each provided meaningful information and generally reflected some level of impairment (5), albeit in much less time than that required for assessment of the full biological complement of taxa. A number of taxonomic groups have been touted as useful in assessing the quality of aquatic systems: diatoms (15), algae (16), protozoa (17,18), benthic macroinvertebrates (19–22), and fishes (23,24). The advantages of using periphyton (algae), benthic macroinvertbrates, and fish are outlined in Barbour et al. (25). The structure and function of specific biotic assemblages or taxonomic groups, such as benthic invertebrates or fish, integrate information about the past and present water quality conditions. Consequently, their inclusion in aquatic bioassessments strengthens the ability to categorize the prevailing water quality.

Refinements to aquatic bioassessment procedures resulted in more widespread use and acceptance, as well as a movement toward standardization of field and laboratory procedures (3,26,27). The inclusion of biotic indices (28–31) and diversity indices (32–34) in the analysis of biological data provided additional measures of community structure and function that aided in the interpretation and identification of impairment from pollution. These indices were also useful in reducing and synthesizing large amounts of data (data that reflected biotic responses to complex interactions and exposures to pollution) into comprehensible numeric values descriptive of water quality. However, bioassessment procedures were still deficient in their ability to fully and accurately describe the overall health of aquatic systems (4).

CURRENT STATE-OF-THE-ART BIOASSESSMENT PROCEDURES

The most useful measures of biotic integrity of aquatic systems reflect the biological responses of organisms and their populations to the prevailing (past and present) environmental conditions. The ability to classify the quality of streams and rivers (with indigenous biological assemblages) provides information that is integral for making management decisions and identifying anthropogenic impacts (35). Development of an “Index of Biotic Integrity” (IBI) using a combination of 12 community parameters (metrics) for fish assemblages provided direction into a new and innovative way of obtaining this type of information (24). The combination of these community attributes (multimetrics) into an IBI provided a classification of the quality of fish communities as a measure of water/habitat quality. The metrics used in the construction of the IBI encompassed several levels of fish community structure and function—species composition/richness and information on ecological factors (trophic structure and health). Successful application of multimetric indices depends on selecting the appropriate metrics used to describe and delineate the biological assemblages (36). The use of multimetrics provides a way of measuring the response of biological assemblages at several levels of organization (e.g., composition, condition, and function) to anthropogenic influences, and these measurable biological attributes can be tested, verified, and calibrated for use in the respective geographical areas of study (37–39). The multimetric concept was expanded on and modified to include benthic macroinvertebrates as well as fish in a set of procedures developed for the U.S Environmental Protection Agency called the Rapid Bioassessment Protocols (RBPs) for use in streams and rivers (40). The RBPs provided a cost-effective framework for bioassessment that could be used by federal, state, and local agencies for management purposes of screening, site ranking, and trend monitoring of the biotic integrity of streams. Modifications, improvements, and additions to these procedures were included in a more recent revision of the RBPs (25). The integrated assessments recommended in the RBPs include information (field collected) on the physical habitat and water quality as well as a number of measures (metrics) describing the biological assemblages of periphyton, benthic macroinvertebrate, and fish.

REFERENCE SITES AND ESTABLISHING BIOCRITERIA

Knowing the organisms that comprise or could comprise the aquatic assemblages in unimpacted streams provides the basis for application of current biological assessment procedures, such as the RBPs, and the development of biological criteria. Consequently, reference streams/sites provide the basis or benchmark for establishing suitable criteria that can be used for comparison with other streams/sites to detect impairment (5,41). Unfortunately, few streams and rivers are pristine anymore because of the widespread influence of anthropogenic activities. Consequently, streams/sites that are the least impacted (minimally impaired) are selected to establish the "attainable" condition or level that is used for comparison with the streams under study.

Reference sites must also be representative of the aquatic systems that are being evaluated. Two types of reference conditions, site specific and regional, are currently used for field bioassessments (25). However, multiple reference sites within a geographical region provide a realistic representation of the community assemblages and compensate for and include the inherent variability associated with individual sites (41). The selection of reference sites on an ecoregional or subregional basis provides an unbiased estimate of the least impacted (attainable) biological assemblages for that particular region of study (41,42). Using a suite (population) of reference streams/sites also allows the use of statistical analyses to establish the natural variability and ranges for each of the metrics and indices to be used for comparison with the study streams. The number of reference sites needed to describe the expected or attainable conditions may vary with the geographical area, but generally 3 to 20 sites are acceptable (41). Once the biological assessments are complete for the reference sites, the suitability and strengths of the metrics included in the analysis and the overall rating index used in detecting impairment can be evaluated using box-and-whisker plots (5) or other statistical analyses (43,44). The box-and-whisker-type plots for the metrics/indices determined for reference streams/sites depict the median, interquartile range (25th and 75th percentiles) and the range (minimum and maximum values). Metric and index values from the study sites that are above (low values for reference conditions) or below (high values for reference conditions) the interquartile range shown by the reference sites demonstrate impairment. The interquartile ranges for the various metrics and combined overall indices (combined scores for the individual metrics) establish the numeric biological criteria for that region.

The reference streams/sites must be representative of the resources at risk. They provide the benchmarks for the expectation or attainable level of biological integrity/water quality that the study areas will be compared with for classification of their level of impairment, and once defined, these biocriteria will describe the best attainable condition (5,37). The acceptable level of difference between the established criteria representing the reference sites and those of the study sites will vary depending on the designated aquatic life use (e.g., cold water fisheries, warm water fisheries, and endangered species) of the study sites (5). The narrative description based on the quantitative database or the numeric criteria (indices) calculated from the database from the reference sites can also be used as the biological criteria for identifying water quality and level of impairment.

An alternative approach to using multimetrics (45,46) in establishing reference conditions relies on multivariate analyses and predictive modeling (47–49). Models are developed that will predict the expected species composition of a "pristine" site given the physical and chemical characteristics (50). The species composition (and associated metrics) from this expected (predicted) community assemblage can be compared with those from the field study sites included in the bioassessments; assessment sites are compared with model-predicted reference conditions. These predictive models require extensive information on the physical and chemical characteristics of pristine streams and the biological organisms that would be associated with them (5). In the absence of field data (or as alternatives for defining reference conditions in the event that suitable sites no longer exist), historical data, simulation modeling, and expert consensus can be used to supplement the data needed for these models (5). Under the multivariate approach, the predicted species compositions (and associated metrics/indices) would establish the biological criteria for streams in that region.

BIOASSESSMENTS CRITICAL IN PRESERVING BIOTIC INTEGRITY

Current practices in biological assessment and the establishment and use of biological criteria based on community structure and function at reference sites (or predicted by models) are effective tools in describing biological integrity and classifying water quality (5,51). As these procedures are used, refinements will most likely be made (particularly at the regional level) as more information is generated and the selection/calibration of metrics is improved. Biological monitoring that includes bioassessment and the establishment of biocriteria on an ecoregional basis (46) is integral to meeting the demands for ecosystem health (52) and measurement of ecological integrity (53) of our nation's aquatic resources.

BIBLIOGRAPHY

1. Hynes, H.B.N. (1960). *The Biology of Polluted Water*. Liverpool University Press, Liverpool, UK, p. 202.
2. Warren, C.E. and Dourdoroff, P. (1971). *Biology and Water Pollution Control*. W.B. Saunders, Philadelphia, p. 434.
3. Mackenthun, K.M. and Ingram, W.M. (1967). *Biological Associated Problems in Freshwater Environments: Their Identification, Investigation and Control*. U.S. Department of

the Interior, Federal Water Pollution Control Administration, Washington, DC, p. 287.

4. Davis, W.S. (1995). Biological assessment and criteria: building on the past. In: *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*. W.S. Davis and T.P. Simon (Eds.). Lewis Publishers, Boca Raton, FL, pp. 1–29.
5. Gibson, G.R. Jr., Barbour, M.T., Stribling, J.B., Gerritesen, J., and Karr, J.R. (1996). *Biological Criteria, Technical Guidance for Streams and Small Rivers* (Rev. Edn.). U.S. Environmental Protection Agency, Office of Water, Washington, DC, EPA/822-B-96-001.
6. Karr, J.R. and Dudley, D.R. (1981). Ecological perspective on water quality goals. *Environ. Manage.* **5**: 55–68.
7. Hynes, H.B.N. (1970). *The Ecology of Running Water*. Liverpool University Press, Liverpool, UK, p. 555.
8. Kolkwitz, R. and Marrson, M. (1908). Ecology of plant saprobia. *Rep. German. Botan. Soc.* **26a**: 505–519.
9. Kolkwitz, R. and Marrson, M. (1909). Ecology of animal saprobia. *Int. Rev. Hydrobiol. Hydrogeogr.* **1**: 126–152.
10. Wilhm, J.L. (1975). Biological indicators of pollution. In: *River Ecology*. B.A. Whitton (Ed.). University of California Press, Berkeley, pp. 375–402.
11. Hynes, H.B.N. (1965). The significance of macroinvertebrates in the study of mild river pollution. In: *Biological Problems in Water Pollution, Third Seminar-1962*. C.M. Tarzwell (Ed.). U.S. Public Health Service, Robert A. Taft Sanitary Engineering Center, Cincinnati, Ohio, pp. 235–240.
12. Cairns J. Jr. (1974). Indicator species vs. the concept of community structure as an index of pollution. *Water Res. Bull.* **10**: 338–347.
13. Patrick, R. (1949). A proposed biological measure of stream conditions, based on a survey of the Conestoga Basin, Lancaster County, Pennsylvania. *Proc. Acad. Natural Sci. Phila.* **101**: 277–341.
14. Hart, C.W. Jr. and Fuller, S.L.H. (1974). *Pollution Ecology of Freshwater Invertebrates*. Academic Press, New York, p. 389.
15. Patrick, R. (1957). Diatoms as indicators of changes in environmental conditions. In: *Biological Problems in Water Pollution. 1956 Symposium*. C.M. Tarzwell (Ed.). U.S. Department of Health, Education and Welfare, Robert A. Taft Sanitary Engineering Center, Cincinnati, Ohio, pp. 71–83.
16. Palmer, C.M. (1957). Algae as biological indicators of pollution. In: *Biological Problems in Water Pollution*. C.M. Tarzwell (Ed.). U.S. Department of Health, Education and Welfare, Robert A. Taft Sanitary Engineering Center, Cincinnati, Ohio, pp. 60–69.
17. Cairns, J. Jr. (1974). Protozoans (protozoa). In: *Pollution Ecology of Freshwater Invertebrates*. C.W. Hart Jr. and S.L.H. Fuller (Eds.). Academic Press, New York, pp. 1–28.
18. Cairns, J. Jr. (1978). Zooperiphyton (especially protozoa) as indicators of water quality. *Trans. Am. Microscop. Soc.* **97**: 44–49.
19. Gaufin, A.R. and Tarzwell, C.M. (1956). Aquatic macroinvertebrate communities as indicators of organic pollution in Lytle Creek. *Sewage Industr. Wastes*. **28**: 906–924.
20. Goodnight, C.J. (1973). The use of aquatic macroinvertebrates as indicators of stream pollution. *Trans. Am. Microsop. Soc.* **92**: 1–13.
21. Roback, S.S. (1974). Insects (arthropoda: insecta). In: *Pollution Ecology of Freshwater Invertebrates*. C.W. Hart Jr. and S.L.H. Fuller (Eds.). Academic Press, New York, pp. 314–376.
22. Klemm, D.J., Lewis, P.A., Fulk, F., and Lazorchak, J.M. (1990). *Macroinvertebrate Field and Laboratory Methods for Evaluating the Biological Integrity of Surface Waters*. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Cincinnati, Ohio, EPA/600/4-90/030.
23. Doudoroff, P. and Warren, C.E. (1957). Biological indices of water pollution with special reference to fish populations. In: *Biological Problems in Water Pollution*. C.M. Tarzwell (Ed.). U.S. Public Health Service, Robert A. Taft Sanitary Engineering Center, Cincinnati, Ohio, pp. 144–163.
24. Karr, J.R. (1981). Assessment of biotic integrity using fish communities. *Fisheries* **6**: 21–27.
25. Barbour, M.T., Gerritsen, J., Snyder, B.D., and Stribling, J.G. (1999). *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish*. U.S. Environmental Protection Agency, Office of Water, Washington, DC. EPA 841-B-99-002.
26. Slack, K.V., Averett, R.C., Greeson, P.E., and Lipscomb, R.G. (1973). *Methods for Collection and Analysis of Aquatic Biological and Microbiological Samples. Techniques of Water-Resource Investigations of the United States Geological Survey*. Book 5, Ch. 4, United States Department of the Interior, U.S. Government Printing Office, Washington, DC, p. 165.
27. C.I. Weber (Ed.). (1973). *Biological Field and Laboratory Methods for Measuring the Quality of Surface Waters and Effluents*. U.S. Environmental Protection Agency, Cincinnati, Ohio, EPA-670/4-73-001.
28. Beck, W.M. Jr. (1955). Suggested method for reporting biotic data. *Sewage Indust. Wastes* **27**: 1193–1197.
29. Chutter, F.M. (1972). An empirical biotic index of the quality of water in South African streams and rivers. *Water Res.* **6**: 19–30.
30. Hilsenhoff, W.L. (1982). *Using a Biotic Index to Evaluate Water Quality in Streams*. Technical Bulletin No. 132. Department of Natural Resources, Madison, WI, p. 22.
31. Hilsenhoff, W.L. (1987). An improved biotic index of organic stream pollution. *The Great Lakes Entomol.* **20**: 31–39.
32. Pielou, E.C. (1966). The measurement of diversity in different types of biological collections. *J. Theoretic Biol.* **13**: 131–144.
33. Wilhm, J.L. and Dorris, T.C. (1968). Biological parameters for water quality criteria. *BioScience* **18**: 477–481.
34. Washington, H.G. (1984). Diversity, biotic and similarity indices: a review with special relevance to aquatic ecosystems. *Water Res.* **18**: 653–694.
35. Hawkes, H.A. (1975). River zonation and classification. In: *River Ecology. Studies in Ecology*. Vol. 2. B.A. Whitton (Ed.). University of California Press, Berkeley, pp. 313–374.
36. Karr, J.R. (1999). Defining and measuring river health. *Freshwater Biol.* **41**: 221–234.
37. Barbour, M.T., Stribling, J.B., and Karr, J.R. (1994). Multimetric approach for establishing biocriteria and measuring biological condition. In: *Biological Assessment and Criteria, Tools for Water Resource Planning and Decision Making*. W.S. Davis and T.P. Simon (Eds.). Lewis Publishers, Boca Raton, FL, pp. 63–77.
38. Barbour, M.T. et al. (1996). A framework for biological criteria for Florida streams using benthic macroinvertebrates. *J. North Am. Benthologic Soc.* **15**: 185–211.
39. Fore, L.S., Karr, J.R., and Wisseman, R.W. (1996). Assessing invertebrate responses to human activities: evaluating alternative approaches. *J. North Am. Benthologic Soc.* **15**: 212–231.

40. Plafkin, J.L., Barbour, M.T., Porter, K.D., Gross, S.K., and Hughes, R.M. (1989). *Rapid Bioassessment Protocols for Use in Streams and Rivers. Benthic Macroinvertebrates and Fish*. U.S. Environmental Protection Agency, Office of Water Regulations and Standards, Washington, DC, EPA 440-4-89-001.
41. Hughes, R.M. (1995). Defining acceptable biological status by comparing with reference conditions. In: *Biological Assessment and Criteria, Tools for Water Resource Planning and Decision Making*. W.S. Davis and T.P. Simon (Eds.). Lewis Publishers, Boca Raton, FL, pp. 31–47.
42. Hughes, R.M., Larsen, D.P., and Omernik, J.M. (1986). Regional reference sites: a method for assessing stream potential. *Environ. Manage.* **10**: 629–635.
43. Hughes, R.M. et al. (1998). A process for developing and evaluating indices of fish assemblage integrity. *Can. J. Fish Aquat. Sci.* **55**: 1618–1631.
44. Kilgour, B.W., Somers, K.M., and Matthews, D.E. (1998). Using the normal range as a criterion for ecological significance in environmental monitoring and assessment. *Ecoscience* **5**: 542–550.
45. Karr, J.R. and Chu, E.W. (1999). *Restoring Life in Running Waters: Better Biological Monitoring*. Island Press, Washington, DC, p. 206.
46. Karr, J.R. and Chu, E.W. (2000). Sustaining living rivers. *Hydrobiologia* **422/423**: 1–14.
47. Moss, D., Furse, M.T., Wright, J.F., and Armitage, P.D. (1987). The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. *Freshwater Biol.* **17**: 41–52.
48. Hawkins, C.P., Norris, R.H., Hogue, J.N., and Feminella, J.W. (2000). Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecol. Applic.* **10**: 1456–1477.
49. Norris, R.H. and Hawkins, C.P. (2000). Monitoring river health. *Hydrobiologia* **435**: 5–17.
50. Clarke, R.T., Furse, M.T., Wright, J.F., and Moss, D. (1996). Derivation of a biological quality index for river sites: comparison of the observed with the expected fauna. *J. Appl. Stat.* **23**: 311–332.
51. U.S. Environmental Protection Agency. (2002). *Biological Assessment and Criteria: Crucial Components of Water Quality Programs*. Office of Water, Washington, DC, EPA 822-F-02-006.
52. Rapport, D.J. et al. (1999). Ecosystem health: the concept, the ISEH, and the important tasks ahead. *Ecosystem Health* **5**: 82–90.
53. Parrish, J.D., Braun, D.P., and Unnasch, R.S. (2003). Are we conserving what we say we are? Measuring ecological integrity within protected areas. *BioScience* **53**: 851–860.

PHYSIOLOGICAL BIOMARKERS AND THE TRONDHEIM BIOMONITORING SYSTEM

KARL ERIK ZACHARIASSEN
Norwegian University of
Science and Technology
Trondheim, Norway

Poor water quality is frequently the result of the presence of potentially toxic substances. Detection of such substances by chemical analysis may be complicated, time consuming, and expensive, and it does not reveal much about possible toxic effects.

WHAT IS A BIOMARKER?

Biomonitoring of water quality takes advantage of the fact that organisms in the water may be moderately affected by the substances. A biomarker is a cellular, biochemical, or physiological parameter, which may be affected when bioactive chemicals are present in the water. One biomarker can respond to the presence of a wide variety of substances and respond before manifest toxic effects occur. Biomarkers are, therefore, particularly useful in early warning monitoring.

Physiological Biomarkers

However, biomarkers should be used by scientists with care, and an appropriate interpretation of their response requires insight into the integrated functional system they are part of, that is, by considering the biomarkers in a physiological context. In a healthy organism, optimal chemical and physical conditions are maintained by a variety of homeostatic regulatory mechanisms. Well-known examples of regulated parameters are body temperature, body fluid pH, and body fluid concentrations of ions, organic solutes, and hormones. Whenever such parameters, for some reason, are displaced beyond optimal range, regulatory mechanisms are activated to compensate for the disturbing stress.

The impact of a chemical substance on an organism can range from a trivial disturbance, which the organism can compensate for, to a serious and eventually lethal disturbance of a regulated parameter. When an organism is injured or killed by a toxic agent, it is because the agent changes a regulated parameter beyond the tolerated range. Regulated parameters as well as the activity of regulatory mechanisms can be physiological biomarkers to provide information about the severity of the impact of a toxic agent (1). As different toxic agents interfere with organisms in different ways, the combined effect pattern of several biomarkers may be more or less specific to the toxic agents, and the effect pattern may identify the agent in cases when the cause of a toxic impact in the environment is unknown or disputed.

THE TRONDHEIM BIOMONITORING SYSTEM

The Trondheim Biomonitoring System (TBS) was developed to extract the appropriate information out of the changes that occur in an organism affected by a toxic chemical. It is a so-called active biomonitoring system; that is, it uses organisms such as mussels placed individually in chambers where they can be exposed to the water that is to be monitored. It operates with three categories of biomarkers: biomarkers for exposure, biomarkers for evaluation of toxic effects, and fingerprint biomarkers to identify the toxic agent (Fig. 1).

Phase 1, Exposure Biomarkers

Exposure biomarkers are the first changes that occur in an affected organism, and they reflect the mechanisms

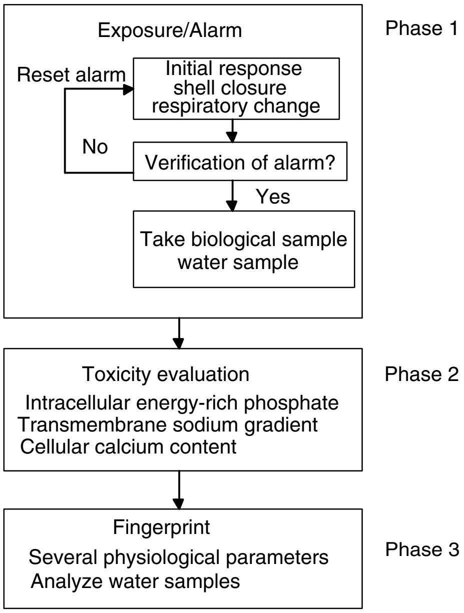

Figure 1. Trondheim Biomonitoring System.

that handle the impact. These changes may be increased activity of homeostatic regulatory mechanisms or the activation of detoxifying mechanisms such as cytochrome P_{450} for certain organic chemicals and metallothioneins (MTs) for certain toxic metals. It should be emphasized that the activation of detoxifying mechanisms such as P_{450} and MT production does not imply that an organism is seriously affected, only that it is compensating for a chemical impact. It should also be emphasized that these systems are activated even by substances of natural origin, and it may be difficult to distinguish natural from non-natural origins.

An ideal exposure biomarker should be sensitive to a wide variety of substances occurring at low concentrations, and it should be possible to monitor the biomarker continuously, for example, online on a computer. As moderate toxic impacts and the activation of many compensatory mechanisms are associated with changes in metabolic rate, the TBS has been using automatic measurement of oxygen consumption as an exposure biomarker. Water from chambers with and without animals is led to a separate chamber with an oxygen electrode connected to a computer, and the oxygen consumption is continuously determined from the difference in oxygen tension of water from chambers with and without animals.

Phase 2, Toxicity Biomarkers

As toxicity biomarkers, the TBS has so far been using tissue levels of energy-rich phosphates and cellular levels of sodium and calcium. The rationale behind the use of these parameters is that many toxic agents inhibit energy metabolism and ionic pumps or affect membrane permeability. Such impacts will affect cellular levels of ATP and active extrusion of sodium and calcium from cells, which will in turn increase the cellular levels of these ions. A substantially reduced transmembrane energy gradient for sodium or high cellular levels of calcium are inevitably lethal, and moderate changes in these parameters indicate that the cells are under serious stress and may become injured if the stress persists (2).

Phase 3, Fingerprint Biomarkers

To obtain pollutant-specific fingerprint biomarkers, a combination of exposure and toxicity stress biomarkers has been used. In addition, intracellular and extracellular concentrations of free amino acids have been included. The specificity of the response of the latter is enhanced because they include more than 20 different substances and they are influenced by a variety of stress factors such as a drop in the transmembrane sodium gradient, anaerobiosis, and cell volume changes. In principle, any measurable parameter may be used, but the number of parameters should be kept at the minimum required to obtain a pollutant-specific effect pattern. The development of an operative system of fingerprint biomarkers is complicated because different pollutants sometimes act together and the changes associated with a single pollutant may depend on its concentration (hormesis).

Toxic agents may also affect a variety of other processes, such as nervous transmission, hormonal control systems and immune responses, osmoregulation, respiration, and heart activity, and effects may be acute or caused by chronic, long-term exposure. Depending on the mechanisms of action of the various toxic chemicals, appropriate biomarkers may be selected and included in the biomarker categories of the TBS.

BIBLIOGRAPHY

1. Aunaas, T. and Zachariassen, K.E. (1993). Physiological biomarkers and the Trondheim biomonitoring system. In: *Biomonitoring of Coastal Waters and Estuaries*. K.J.M. Kramer (Ed.). CRC Press, Boca Raton, FL, pp. 107–130.
2. Børseth, J.F. et al. (1995). Transmembrane sodium gradient and calcium content in the adductor muscle of *Mytlius edulis* L. in relation to the toxicity of oil and organic chemicals. *Aquatic Toxicol.* **31**: 263–276.

BIOMARKERS, BIOINDICATORS, AND THE TRONDHEIM BIOMONITORING SYSTEM

Rita Triebskorn
Steinbeis-Transfer Center for
Ecotoxicology and
Ecophysiology
Rottenburg, Germany

Environmental effects resulting from chemical exposure can be traced at different levels of biological organization. In this context, the term “biomarker” is used mainly for functional measures expressed at the organism to suborganism level of biological organization, while the term “bioindicator” refers to effect endpoints at higher biological levels including population and community-level attributes (1).

During the last two decades, interest in using biomarkers or bioindicators as monitoring tools to assess environmental pollution has steadily increased. This is mainly due to the fact that chemical analyses, which are well established in ecotoxicological monitoring programs, can exclusively give information on chemicals present in the environment or accumulated in biota. However, they cannot provide information on exposure-related effects, which possibly occur in exposed organisms. In addition, chemical analyses can often only provide information on a defined spectrum of chemicals (being searched for) available and are mainly "snapshots" of a defined exposure scenario (depending on the time of sampling). Furthermore, low dose effects have received increasing attention during the last few years, especially in the context of endocrine disruption. Therefore, the detection of effects that may be attributed to chemicals present in the environment in concentrations at or beyond the detection limits has become more and more important. On the other hand, it has also become evident that the presence of a chemical in an organism or an environment does not necessarily imply that the organism or the environment is affected or threatened by that chemical.

Biomarkers are suitable not only for delivering information on the health status of exposed organisms (biomarkers of effect), but also on the quality and/or quantity of the exposure situation (biomarkers of exposure). Thus, they can be used both as measures of toxic effects and as fingerprints for chemical exposure (2–4). Depending on the character of biomarkers selected for risk assessment, the evaluator will receive more data on either adverse effects in organisms or on the exposure situation itself (Fig. 1).

Whereas established ecological surveys (e.g., the water quality assessment system) are often based on the disappearance of species under stressed conditions (thus using mortality as the endpoint of toxic effect), the biomarker approach uses responses of organisms (still) living under unfavorable conditions as early warnings before populations or ecosystems are severely damaged (5–7). Since biomarkers respond quickly to environmental stressors before ecological effects become apparent, they serve to detect effects at an early stage. They are highly sensitive and, in addition, integrate effects of animals' exposure to contaminants and confounding factors over time (8–10).

To date, most biomarker research has been conducted in aquatic ecosystems (summarized in Reference 1). In the last decade, however, the field of biomarker application in terrestrial environments has also expanded and several techniques have been adapted for soil organisms (11,12).

A concept to integrate physiological biomarkers into a monitoring system was elaborated by a group of scientists from Trondheim, Norway, ten years ago (2). This "Trondheim Biomonitoring System" (TBS) is a practically operative integrated system for biomonitoring of marine environments using the mussel *Mytilus edulis* as the monitoring organism. The technical setup of this system consists of ten exposure chambers each containing one mussel. The exposure chambers, through which seawater flows, are mounted into a submersible exposure unit and are connected to a respirometer. Thus, data for oxygen consumption can be recorded continuously for each mussel and displayed on a computer screen. In addition, samples (e.g., tissue, hemolymph) for other purposes can be taken from the exposed organisms. The aims of this biomonitoring system are as follows:

1. To detect the presence of environmental pollution at the earliest possible stage in order to enact possible countermeasures before ecological damage becomes obvious.
2. To evaluate the stress status of organisms or ecosystems exposed to environmental pollutants.
3. To identify the pollutants responsible for the detected stress status.

To achieve these aims, the TBS uses physiological biomarkers in three phases. In the first phase, "the

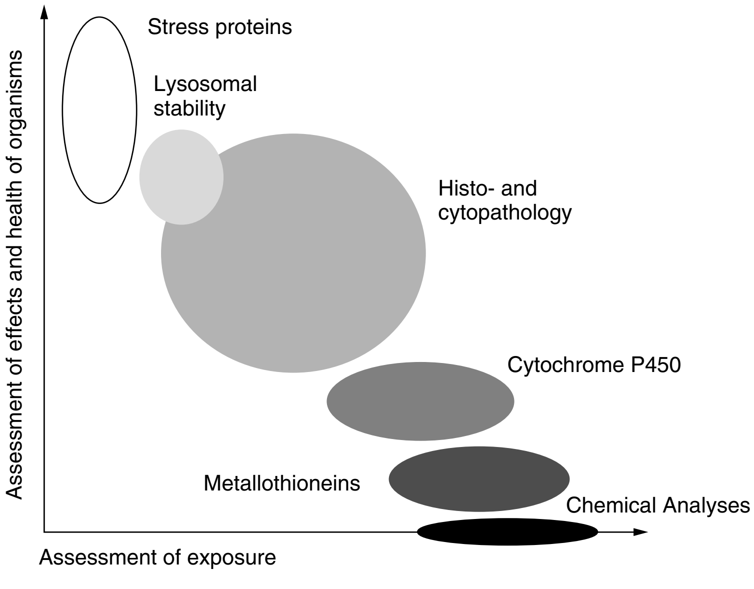

Figure 1. Suitability of different biomarkers to assess effects and health of organisms on the one hand and the exposure situation on the other. Chemical analyses only provide information on the exposure situation.

alarm phase," very sensitive biological parameters (e.g., shell closure or oxygen consumption of mussels) are continuously recorded as "alarm parameters." In the event the alarm is verified as not due to changes in natural environmental parameters, mussel samples are taken for phase 2.

In the second phase, "the toxicity evaluation phase," biomarkers indicating toxicity stress or strain (e.g., intracellular energy-rich phosphates, strombine formation, transmembrane sodium gradient, cellular calcium content) are used to assess the health status of the exposed organisms.

In the third phase, "the fingerprint phase," a set of chemical-specific biomarkers (e.g., transmembrane distribution of free amino acids) are applied in parallel with chemical analyses of samples automatically taken in phase 1. These investigations provide information on the nature of the chemicals responsible for the effects observed in phases 1 and 2.

For freshwater monitoring of pollution, the Multispecies Freshwater Biomonitor (MFB) has been established by LIMCO Int. (13). Similar to the TBS, the MFB uses flow-through exposure chambers for the exposure of limnic organisms including gammarids, chironomids, or fish. As biomarkers/bioindicators, breathing activity and locomotion of animals can be recorded continuously in this system. In case of an incidence, samples can be taken for different analytical purposes.

In order to evaluate biomarker responses with respect to their toxicological and ecological relevance (i.e., their chemical specificity and importance for effects at higher biological levels including populations and ecosystems), an extended monitoring project called VALIMAR (*vali*dation of bio*mar*kers for the assessment of small stream pollution) was conducted at two small streams in southern Germany between 1995 and 2000 using fish as monitoring organisms (10). The project was based on a "middle-up and middle-down" approach using flow-through aquaria in the field as intermediates between controlled lab and uncontrolled field situations.

According to the results of VALIMAR, for stream systems mainly influenced by wastewater release and agricultural activities, a combination of cellular (e.g., histopathology of kidney or liver, ultrastructure of gills or liver) with biochemical biomarkers (e.g., EROD activity, stress protein expression, catalase or esterase activity) can be recommended as useful tools for assessing risk in fish populations. Biomarkers proved to be sensitive early warning sentinels, which indicate the general stress status and the health of fish in streams influenced by different sources of pollution, but also in streams influenced by temporary or locally restricted spills (6,10). In addition, biomarkers are very useful in assessing the recovery of aquatic ecosystems (14). Although in the VALIMAR project the ultrastructure of the gills could clearly be identified as the best discriminator between the two streams in both brown trout and stone loach (15), the use of single biomarkers to assess the holistic effects of pollution, either structural or biochemical, is not recommended. Whereas biochemical biomarkers are sensitive and respond rapidly to stressors, their disadvantage is that they are also highly variable and are quick to recover following removal of stress. Structural responses, on the other hand, are generally less variable, but also less sensitive in response to short-term events. Since structural parameters require more time to recover than biochemical responses, structural responses are useful in understanding past water quality conditions for some days or weeks after a stress event. For the experimental situation of VALIMAR, a combination of biomarkers to discriminate between the two test streams in trout was the gill ultrastructure, esterase, and cytochrome P450 enzymes. A suitable combination of biomarkers in stone loach was gill ultrastructure, acid phosphatase, and alanine aminotransferase (15).

As cellular biomarkers, histological and ultrastructural changes in monitoring organs (e.g., liver or gills of fish, or hepatopancreas or kidney of invertebrates) were shown to be suitable for pollution assessment in terrestrial and aquatic ecosystems (16–18). The *in situ* evaluation of cellular injury allows a direct diagnosis of adverse effects qualitatively but also quantitatively by means of morphometry or different semiquantitative assessment methods (19). Since histology and cytology not only reflect the impact of xenobiotics but also the general metabolic status of an organism, it is essential for an investigator of cellular biomarkers to be aware of the diversity and plasticity of cellular reactions occurring in an organ under unpolluted conditions. Many studies using light and electron microscope techniques as tools of prospective as well as retrospective risk assessment have been conducted. The majority of this work has been done with fish, but experience is also available for molluscs, annelids, mites, diplopods, isopods, and collembolans in response to heavy metals and organic pollutants (reviewed in References 11 and 20).

At the subcellular/organelle level, the lysosomal system has been identified as a very sensitive target for xenobiotics (21). Lysosomes are membrane-bound vesicles containing acid hydrolases. In intact lysosomes, these hydrolases are prevented from reacting with cytoplasmic compounds by an intact membrane. In response to stress, the lysosomal membrane stability decreases and its permeability increases. As a result, hydrolases are released into the cytoplasm. One of the easiest techniques to determine lysosome membrane integrity is the neutral-red retention assay, which is based on the uptake of the dye neutral-red by transmembrane diffusion. The capacity of healthy lysosomes to take up and maintain the dye depends on the efficiency of membrane-bound proton pumps (22). In stressed cells, neutral-red gradually leaks from the lysosomes into the surrounding cytoplasm. The time of dye retention in the lysosomes is the measure for the intensity of stress. In the past, most experience with the neutral-red retention assay exists in the marine field (21); however, a few studies have also been conducted with soil invertebrates exposed to environmental pollutants (23,24).

At the physiological, biochemical, and molecular levels, several biomarkers are well established, which differ with respect to their capability to indicate exposure and/or effects. The most prominent biomarkers are biotransformation enzymes (e.g., cytochrome P450 and glutathione-*S*-transferase), metallothioneins, and stress proteins. In the TBS, respiratory rates, intracellular energy-rich phosphates, strombine formation, the transmembrane sodium gradient, the cellular calcium content, the transmembrane distribution of free amino acids, and other physiological parameters are recommended for evaluation of chemical stress in marine ecosystems. Physiological, biochemical, and molecular biomarkers are very sensitive and respond quickly to environmental stressors. Therefore, they are also useful tools to detect influences on organisms in minor contaminated environments. In contrast to cellular biomarkers, however, which show saturation response kinetics with increasing intensity of the stressor, response kinetics of biochemical and molecular biomarkers often follow an optimum curve (Fig. 2). Thus, low values for molecular biomarker responses are difficult to interpret since they may occur in healthy as well as in moribund organisms, in the latter having already surpassed the optimal biomarker response threshold. Whether these low values result from exposure to very low or very high stressor intensities can only be decided when biochemical markers are combined with markers that are characterized by optimum response kinetics, for example, histopathological diagnostics (25). In addition, the dependence of physiological, biochemical, and molecular marker responses on exogenous and/or endogenous stress factors other than xenobiotics (e.g., temperature or reproduction status) has to be taken into account by the investigator.

The question of whether biomarker responses are relevant for effects at higher biological levels (i.e., populations or ecosystems) can be answered either experimentally (as in Reference 26) or by using a weight of evidence approach (27). Köhler and colleagues (26) studied the relevance of stress protein responses in the slug *Deroceras reticulatum* exposed to metals for different population parameters. They showed that the hsp70 levels in slugs after 2–3 weeks of exposure to the respective metals reflected effects in fecundity, reproduction, and mortality of lifetime exposure. Thus, in this case, biomarker measurements were proved to be early warning tools to predict long range consequences of xenobiotic exposure at the level of life cycle parameters.

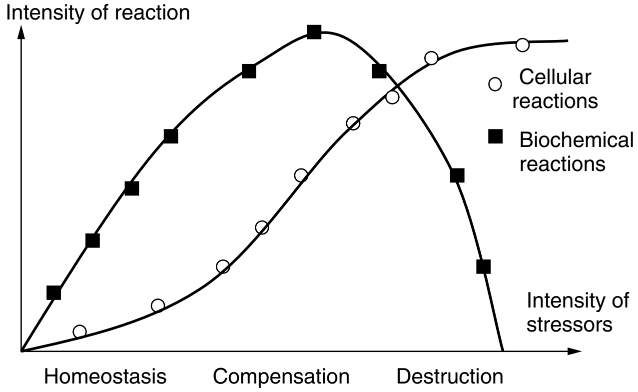

Figure 2. Response kinetics of cellular and biochemical biomarkers.

Since long generation times of many organisms often do not permit such direct evidence of causal relationship between effects at low and high biological levels, a weight of evidence approach commonly used in human medicine is suitable to relate causes and effects on the basis of epidemiological data. By means of Hill's causality criteria (28), for example, the criteria of biological plausibility, strength of association, consistency of association, or experimental evidence, causality can be established even when direct experimental proof of causality is lacking. Based on this, in the VALIMAR project (10), a simultaneous collection of data for responses at different levels of biological organization in combination with a detailed characterization of the exposure situation made it possible to establish causality between different levels of biological organization as well as between exposure and effects (27).

BIBLIOGRAPHY

1. Adams, S.M. (2002). Biological indicators of aquatic ecosystem stress: introduction and overview. In: *Biological Indicators of Aquatic Ecosystem Stress*. S.M. Adams (Ed.). American Fisheries Society, Bethesda, MD, pp. 1–11.
2. Aunaas, T. and Zachariassen, K.E. (1994). Physiological biomarkers and the Trondheim biomonitoring system. In: *Biomonitoring of Coastal Waters and Estuaries*. K.J.M. Kramer (Ed.). CRC Press, Boca Raton, FL, pp. 107–130.
3. Depledge, M.H. and Fossi, M.C. (1994). The role of biomarkers in environmental assessment (2). Invertebrates. *Ecotoxicology* **3**: 161–172.
4. Peakall, D.B. and Walker, C.H. (1994). The role of biomarkers in environmental assessment (3). Vertebrates. *Ecotoxicology* **3**: 173–179.
5. Ham, K.D., Adams, S.M., and Peterson, M.J. (1997). Application of multiple bioindicators to differentiate spatial and temporal variability from the effects of contaminant exposure on fish. *Ecotoxicol. Environ. Safety* **37**: 53–61.
6. Adams, S.M., Bevelhimer, M.S., Greeley, M.S., Levine, D.A., and Teh, S.J. (1999). Ecological risk assessment in a large river-reservoir: 6: bioindicators of fish population health. *Environ. Toxicol. Chem.* **18**(4): 628–640.
7. Adams, S.M. (2000). Assessing sources of stress to aquatic ecosystems using integrated biomarkers. *Biological Resource Management Connecting Science and Policy*, pp. 17–29.
8. Van Gestel, C.A.M. and van Brummelen, T.C. (1996). Incorporation of the biomarker concept in ecotoxicology calls for a redefinition of terms. *Ecotoxicology* **5**: 217–225.
9. Triebskorn, R. et al. (1997). Induction of heat shock proteins, changes in liver ultrastructure, and alterations of fish behavior: are these biomarkers related and are they useful to reflect the state of pollution in the field? *J. Aquat. Ecosys. Stress Recov.* **6**: 57–73.
10. Triebskorn, R. et al. (2001). The project VALIMAR (VALIdation of bioMARkers for the assessment of small stream pollution): objectives, experimental design, summary of results, and recommendations for the application of biomarkers in risk assessment. *J. Aquat. Ecosyst. Stress Recov.* **8** (3/4): 161–178.

11. Kammenga, J.E. et al. (2000). Biomarkers in terrestrial invertebrates for ecotoxicological soil risk assessment. *Rev. Environ. Contam. Toxicol.* **164**: 93–147.
12. Köhler, H.R. and Triebskorn, R. (2004). Stress im Boden: Früherkennung ökotoxikologischer Effekte durch Biomarker. *BIUZ* **4**: 240–248.
13. Gerhardt, A. (2001). A new Multispecies Freshwater Biomonitor for ecological relevant control of surface waters. In: *Biomonitors and Biomarkers as Indicators of Environmental Change*, Vol. II, F. Butterworth et al. (Eds.). Kluwer Plenum Press, New York, pp. 301–317.
14. Munkittrick, K.R., Servos, M.R., Carey, J.H., and van der Kraak, G.J. (1997). Environmental impacts of the pulp and paper wastewater: evidence for a reduction in environmental effects at North American pulp mills since 1992. *Water Sci. Technol.* **35**(2/3): 329–338.
15. Dietze, U. et al. (2001). Chemometric discrimination between streams based on chemical, limnological and biological data taken from freshwater fishes and their interrelationships. *J. Aquat. Ecosyst. Stress Recov.* **8**(3/4): 319–336.
16. Triebskorn, R. and Köhler, H.-R. (1996). The impact of heavy metals on the grey garden slug *Deroceras reticulatum* (Müller): metal storage, cellular effects and semi-quantitative evaluation of metal toxicity. *Environ. Pollut.* **93**: 327–343.
17. Gernhöfer, M., Pawert, M., Schramm, M., Müller, E., and Triebskorn, R. (2001). Ultrastructural biomarkers as tools to characterize the health status of fish in contaminated streams. *J. Aquat. Ecosyst. Stress Recov.* **8**(3/4): 241–260.
18. Triebskorn, R. et al. (2004). Toxic effects of the non-steroidal anti-inflammatory drug diclofenac. Part II: cytological effects in liver, kidney, gills and intestine of rainbow trout (*Oncorhynchus mykiss*). *Aquat. Toxicol.* **68**(2): 151–166.
19. Schramm, M., Müller, E., and Triebskorn, R. (1998). Brown trout *(Salmo trutta)* f. *fario* liver ultrastructure as biomarkers of small stream pollution. *Biomarkers* **3**(2): 93–108.
20. Triebskorn, R. et al. (1991). Invertebrate cells as targets for hazardous substances. *Z. Angew. Zool.* **78**(3): 277–287.
21. Moore, M.N. (1990). Lysosomal cytochemistry in marine environmental monitoring. *Histochemistry* **22**: 187–191.
22. Seglen, P.P. (1983). Inhibitors of lysosomal function. *Methods Enzymol.* **96**: 737–765.
23. Weeks, J.M. and Svendsen, C. (1996). Neutral red retention by lysosomes from earthworm (*Lumbricus rubellus*) coelomocytes: a simple biomarker of exposure to soil copper. *Environ. Toxicol. Chem.* **15**(10): 1801–1805.
24. Harreus, D., Köhler, H.-R., and Weeks, J. (1997). Combined non-invasive cell isolation and neutral-red retention assay for measuring the effects of copper on the lumbricid *Aporrectodea rosea* (Savigny). *Bull. Environ. Contam. Toxicol.* **59**: 44–49.
25. Triebskorn, R. and Köhler, H.-R. (2003). Cellular and molecular stress indicators as tools to assess effects and side-effects of chemicals in slugs. *BCPC Symp. Proc.* **80**: 69–76.
26. Köhler, H.-R. et al. (1998). Validation of hsp70 stress gene expression as a marker of metal effects in *Deroceras reticulatum* (pulmonata): correlation with demographic parameters. *Environ. Toxicol. Chem.* **17**: 2246–2253.
27. Triebskorn, R. et al. (2003). Establishing causality between pollution and effects at different levels of biological organization: the VALIMAR project. *Hum. Ecol. Risk Assessment* **9**(1): 171–194.
28. Bradford-Hill, A. (1965). The environment and disease. Association or causation? *Proc. R. Soc. Med.* **9**: 295–300.

ACTIVE BIOMONITORING (ABM) BY TRANSLOCATION OF BIVALVE MOLLUSCS

Roel Smolders
Judith Voets
Lieven Bervoets
Ronny Blust
University of Antwerp
Antwerp, Belgium

Victor Wepener
Rand Afrikaans University
Auckland Park, South Africa

DEFINITION

To start with, it is necessary to clearly define what is understood by active biomonitoring (ABM) using bivalve molluscs in this context. Here, the definition that will be used is: "the translocation of bivalve molluscs from one place to another for the purpose of environmental quality monitoring" (1–3).

The use of bivalves obviously indicates that the focus will be on the aquatic environment. However, ABM has also been performed in the terrestrial environment, using, for example, snails (4,5), mosses (6,7), lichens (8), or grasses (9).

Thus, an ABM approach employs organisms that are collected from an (generally) unstressed, unpolluted population and that, afterwards, are translocated to potentially polluted sites. The chemical and biological consequences of this translocation, which usually involves the caging of the organisms, can then be followed in space and time to estimate the effects of exposure on selected endpoints (3,10).

A second approach in which organisms are used to reflect environmental pollution is passive biomonitoring (PBM), which comprises "the collection of organisms from their habitats at sites where a natural population exists" (1–3). Although PBM has also frequently been used, ABM has a number of properties that advocate its use in environmental quality monitoring.

(DIS) ADVANTAGES OF ABM

The advantages of ABM for biomonitoring purposes are:

- Experiments can be performed for a **known exposure period**.
- All organisms have a **similar life history** at the start of the exposure.
- It is easy to **compare different sites**, even if the organisms are not normally present at the exposure locations.
- A **comparison between transplanted and indigenous organisms** can indicate to what extent the indigenous organisms are adapted to the location or can give information about uptake kinetics of pollutants. Indigenous species can be "genetically

protected," and thus the use of transplanted organisms can circumvent this adaptation process, making the comparison among different sites more precise and the use of nonadapted species more sensitive.

- Compared with laboratory-bred organisms, transplanted organisms are **better acclimated** to changing environmental conditions, and thus the impact of field exposure through, for example, climatological shock, will be less pronounced.

Although these advantages advocate the use of ABM under field conditions, a number of drawbacks also exist that have to be kept in mind when evaluating and interpreting the results of ABM studies:

- **Food availability** may skew the results by overestimating or underestimating the instream toxicity. As mussels depend mainly on algae as a food source, exposure in nutrient-poor but clean water may indicate stressful situations not related to pollutants but to a decreased quality of food resources (11,12). On the other hand, high amounts of food in eutrophic streams may mitigate the impact of pollutant exposure on aquatic organisms. Several authors have demonstrated the positive impact of increased food availability on growth and reproduction of pollutant-exposed organisms (13–15).
- The **loss of cages** containing the transplanted animals can make interpretation of the results much more difficult or even impossible.
- Compared with indigenous organisms, ABM might give **an overestimation of effects** because prolonged exposure to low levels of contamination eventually could lead to genetic adaptation of indigenous organisms. ABM does not take this kind of evolutionary processes into account.

WHY USE BIVALVE MOLLUSCS?

ABM studies have been performed with a wide variety of organisms, but bivalve molluscs have a number of important advantages compared with other aquatic organisms, which advocate their usefulness in biomonitoring.

- Although subsequent handling of organisms had a demonstrably negative effect on bivalve growth (16,17), **handling stress** and crowding are less important in bivalves compared with other organisms.
- Many bivalve molluscs are **relatively resistant to pollution**, although this feature is obviously species-dependent. The most commonly used test species for both marine and freshwater environments have a relatively high pollution tolerance, although this does not necessarily mean that they are insensitive to it.
- As filter feeders, mussels are **exposed to water, food, and suspended matter** as potential sources of contamination, which provides an integrated measure of pollution exposure, which is often not the case with other test species.
- Mussels have a **very high bioaccumulation** and a low biotransformation potential for both organic and inorganic contaminants (18,19), which means that there will be no or low transformation of, for example, organic micropollutants, and mussels can thus be used as long-term bioaccumulators.
- Mussels are **sedentary organisms** and are often easily collected. Mostly, they are attached to rocky or woody substrates in high numbers, which makes it easy to collect a large number of organisms from one location.
- Bivalves often have a **commercial value**, which makes them target organisms for environmental protection and legislation efforts.

However, a number of negative features connected to the use of bivalve molluscs as a biomonitoring tool also exist.

- **Reproduction**, an ecotoxicological endpoint with very high ecological relevance, is relatively difficult to measure and is highly dependent on seasonality (20).
- **No clear reference populations** exist, whereas other aquatic organisms (e.g., daphnids or zebrafish) have clearly defined responses under "no stress" or laboratory situations, which might complicate the comparison between different studies when organisms for transplantation are taken from geographically different collection sites.
- Mussels have no or a relatively **low biotransformation potential** (also a benefit, see above), which limits the use of biotransformation enzymes as biomarkers of exposure (18).

THE USE OF BIVALVE MOLLUSCS FOR BIOMONITORING

Biomonitoring is regarded as the regular systematic use of organisms to evaluate changes in the environment or water quality (21). It may take on many forms, from the measurement of chemical residues in tissues of living organisms (i.e., chemical monitoring) to the quantification of biological endpoints (i.e., biological effect monitoring or biomarkers) including changes to various biochemical, physiological, morphological, and behavioral factors (22).

The most popular bioindicator species used in the marine bioaccumulation and biomarker ABM are species from the family Mytilidae (e.g., *Mytilus edulis* and *M. galloprovincialis*) and *Perna* spp. from the family Pernidae. With the subsequent expansion of ABM to freshwater environments, the Dreissenidae (e.g., *Dreissena polymorpha*) and species from the family Unionidae (e.g., *Anodonta cygnea, Elliptio complanata*) are the most used bioindicator organisms. These four families make up over 70% of the bioindicator organisms used in ABM studies worldwide. An evaluation of three databases revealed over 350 references related to the implementation of transplanted bivalves in biomonitoring programs. These studies that were carried out between the early 1980s and 2004 showed two very definite trends in the application of bivalves in ABM.

Chemical Monitoring with Bivalve Molluscs

Physical and chemical measurements of dissolved and particulate bound compounds reveal valuable information concerning the presence and the dispersion of pollutants in aquatic environments. However, important shortcomings exist in predicting the impact of pollution in the environment. Due to changes in water composition and environmental conditions, representative samples cannot easily be obtained (23,24). Furthermore, measurements of pollutants in the water and/or sediments do not necessarily represent their bioavailability. Therefore, measurements of accumulated micropollutants in the biota can help to determine the presence, bioavailability, and impact of pollution in the environment.

During the last three decades, bivalves have widely been employed as biomonitors to evaluate the bioavailability of micropollutants. Marine mussels, such as *M. edulis, M. galloprovincialis*, and *P. viridis*, have been used since 1975 in monitoring programs to evaluate the bioavailability and effects of pollutants present in estuarine and coastal ecosystems (25–29). In several areas of the world, this so-called “Mussel Watch Program” has been implemented (30). Following these marine biomonitoring programs, freshwater mussels have been used since the late 1980s to monitor the quality of freshwater ecosystems. Freshwater bivalves that have been used are, among others, *E. complanata*, the Asiatic clam *Corbicula fluminea*, and *Anodonta* sp. (31,32). In Europe, the most frequently used bivalve species is the zebra mussel, *D. polymorpha* (33–36). One matter that is of special interest when molluscs are used for chemical monitoring involves the question whether concentrations of micropollutants in resident mussels might vary from transplanted mussels. Therefore, comparison of the accumulation of contaminants between transplanted and indigenous mussels is essential to evaluate and validate the applicability of caged organisms as a chemical monitoring tool. Among others, an extensive study by Bervoets et al. (37) showed that transplanted zebra mussels, exposed for a period of 6 weeks in summer, accumulated a broad range of micropollutants up to levels comparable with levels measured in resident mussels.

Table 1 lists a number of selected studies that have used ABM with bivalve molluscs to monitor the presence of pollutants in freshwater, estuarine, and marine environments.

Biological Monitoring with Bivalve Molluscs

The realization, however, that it is virtually impossible to monitor all contaminants present in the environment (both of anthropogenic and natural origin), together with the increased attention that the relationship between chemical body burdens and toxic effects was receiving, led to the application of biomarkers in ABM (59,60). Biomarkers involve the assessment of the overall quality of the aquatic environment by examining biochemical, physiological, behavioral, or population responses that reflect the potential of contaminants to impair biological processes in exposed organisms (61). From the early 1990s, biomarkers became frequently employed in ABM studies as indicators of biological responses, which were related to exposure and toxic effects of environmental chemicals. Various biochemical parameters in bivalves have been tested for their responses to toxic substances and their potential use as biomarkers of exposure or effect (22). The biomarkers that have been investigated the most extensively are related to those biochemical parameters that provide short-term indication of long-term effects, that is, responses at the lower level of biological organization (e.g., antioxidant enzymes, metallothioneins, acetylcholine esterase). However, the technical demands and lack of ecological relevance have been identified as a

Table 1. A Brief Overview of Chemical Monitoring Studies with Bivalve Molluscs

Area	Species	Common Name	Pollutant	Reference
Marine waters	*Mytilus edulis*	Common mussel	Heavy metals	39
	Mytilus edulis	Common mussel	PAHs	40
	Mytilus edulis	Common mussel	PCBs and hydrocarbons	41
	Mytilus galloprovincialis	Bay mussel	Heavy metals	42
	Mytilus galloprovincialis	Bay mussel	PAHs	43
	Mytilus californianus	California mussel	Heavy metals	44
	Macoma balthica	Baltic tellin	Heavy metals	39, 45
	Perna viridis	Green mussel	Heavy metals	46, 47
Estuarine	*Perna perna*	Brown mussel	Hg, MeHg	48
	Ragnia cuneata	Wedge clam	Heavy metals	49
	Ragnia cuneata	Wedge clam	Dioxins and furans	50
	Potamocorbula amurensis	Amur River clam	Hydrocarbons	51
Freshwater	*Dreissena polymorpha*	Zebra mussel	Organotins	52, 53
	Dreissena polymorpha	Zebra mussel	Heavy metals	37, 38, 54
	Dreissena polymorpha	Zebra mussel	PCBs, PBDEs, DDE, HCB	37
	Corbicula fluminea	Asian clam	Heavy metals	31
	Pyganodon grandis	Giant floater	MeHg	55
	Elliptio complanata	Eastern elliptio	PAH and PCB,	56, 57
	Hyridella depressa		Heavy metals	58

concern in the application of these biomarkers in ABM. In this regard, energy budgets at different levels of biological organization can potentially link these different levels and provide a holistic and integrative indication of the effects observed within exposed organisms. This concept has been applied from as early as the beginning of the 1980s to demonstrate the scope for growth changes in bivalves along a pollution gradient. In recent years, energetic responses are increasingly used as biomarkers in ABM and have been statistically correlated to negative effects on growth and condition in bivalves transplanted along a pollution gradient and other invertebrate community-related responses (2,3,62,63).

CONCLUSIONS

Bivalve molluscs have a number of properties that advocate their use in both chemical and biological monitoring studies. Active biomonitoring, being the translocation of organisms from one (usually unpolluted) location to another one (usually polluted), can provide valuable information about the state of the environment, the presence and biological availability of pollutants, and the effects of these pollutants on organisms and ecosystems.

BIBLIOGRAPHY

1. De Kock, W.C. and Kramer, K.J.M. (1994). In: *Biomonitoring of Coastal Waters and Estuaries*. K.J.M. Kramer (Ed.). CRC Press, Boca Raton, FL.
2. Bervoets, L., Smolders, R., Voets, J., and Blust, R. (2003). In: *Biological Evaluation and Monitoring of the Quality of Surface Waters*. J.J. Symoens and K. Wouters (Eds.). Cercle Hydrobiologique de Bruxelles, p. 129.
3. Smolders, R., Bervoets, L., Wepener, V., and Blust, R. (2003). *Hum. Ecol. Risk Assess.* **9**: 741–760.
4. Pihan, F. and de Vaufleury, A. (2000). *Ecotoxicol. Environ. Safety* **46**: 137–147.
5. Scheifler, R., Ben Brahim, M., Gomot-de Vaufleury, A., Carnus, J.-M., and Badot, P.-M. (2003). *Environ. Pollut.* **122**: 343–350.
6. Couto, J.A., Aboal, J.R., Fernández, J.A., and Carballeira, A. (2004). *Chemosphere* **57**: 303–308.
7. Fernández, J.A., Aboal, J.R., and Carballeira, A. (2004). *Ecotoxicol. Environ. Safety* **59**: 76–83.
8. Jeran, Z., Byrne, A.R., and Batic, F. (1995). *The Lichenologist* **27**: 375–385.
9. Dietl, C., Reifenhäuser, W., and Peichl, L. (1997). *Sci. Total Environ.* **205**: 235–244.
10. Salazar, M.H. and Salazar, S.M. (1999). *Standard Guide for Conducting in-situ Field Bioassays with Caged Bivalves*. ASTM International Guideline E2122-02.
11. Borcherding, J. (1995). *Malacologia* **36**: 15–27.
12. Madon, S.P., Schneider, D.W., Stoeckel, J.A., and Sparks, R.E. (1998). *Can. J. Fish. Aquat. Sci.* **55**: 401–413.
13. Bridges, T.S., Farrar, J.D., and Duke, B.M. (1997). *Environ. Toxicol. Chem.* **16**: 1659–1665.
14. Emery, V.L., Moore, D.W., Gray, B.R., Duke, B.M., Gibson, A.B., Wright, R.B., and Farrar, J.D. (1997). *Environ. Toxicol. Chem.* **16**: 1912–1920.
15. Stuijfzand, S.C., Helms, M., Kraak, M.H.S., and Admiraal, W. (2000). *Ecotoxicol. Environ. Safety* **46**: 351–356.
16. Salazar, M.H. and Salazar, S.M. (1996). In: M.A. Champ and P.F. Seligman (Eds.). *Organotin*. Chapman & Hall, London.
17. Monroe, E.M. and Newton, T.J. (2001). *J. Shellfish Res.* **20**: 1167–1171.
18. Daubenschmidt, C., Dietrich, D.R., and Schlatter, C. (1997). *Aquat. Toxicol.* **37**: 295–305.
19. Nasci, C. et al. (1999). *Environ. Pollut.* **39**: 255–260.
20. Seed, R. and Suchanek, T.H. (1992). In: *The Mussel Mytilus: Ecology, Physiology, Genetics and Culture*. E.M. Gosling (Ed.). Elsevier, Amsterdam.
21. De Zwart, D. (1995). *Monitoring Water Quality in the Future*. Vol. 3. *Biomonitoring*. National Institute of Public Health and Environmental Protection (RIVM), Bilthoven, The Netherlands.
22. Van der Oost, R., Beyer, J., and Vermeulen, N.P.E. (2003). *Environ. Toxicol. Pharmacol.* **13**: 57–149.
23. Luoma, S.N. (1983). *Sci. Total Environ.* **28**: 1–22.
24. van Griethuysen, C., Meijboom, E.W., and Koelmans, A.A. (2003). *Environ. Toxicol. Chem.* **22**: 457–465.
25. Goldberg, E.D. et al. (1978). *Environ. Conserv.* **5**: 101–125.
26. Farrington, J.W. et al. (1983). *Environ. Sci. Technol.* **14**: 490–496.
27. Walsh, A.R. and O'Halloran, J. (1998). *Environ. Toxicol. Chem.* **17**: 1429–1438.
28. Chase, M.E. et al. (2001). *Mar. Pollut. Bull.* **42**: 491–505.
29. de Astudillo, L.R. et al. (2002). *Arch. Environ. Contamin. Toxicol.* **42**: 410–415.
30. Blackmore, G. and Wang, W.X. (2003). *Environ. Toxicol. Chem.* **22**: 388–395.
31. Andres, S. et al. (1999). *Environ. Toxicol. Chem.* **18**: 2462–2471.
32. Gewurtz, S.B., Lazar, R., and Haffner, G.D. (2000). *Environ. Toxicol. Chem.* **19**: 2943–2950.
33. Regoli, L., Chan, H.M., de Lafontaine, Y., and Mikaelian, I. (1994). *Aquat. Toxicol.* **53**: 115–126.
34. Mersch, J., Wagner, P., and Pihan, J.C. (1996). *Environ. Toxicol. Chem.* **15**: 886–893.
35. Hendriks, A.J., Pieters, H., and de Boer, J. (1998). *Environ. Toxicol. Chem.* **17**: 1885–1898.
36. Zimmermann, S. et al. (2002). *Environ. Toxicol. Chem.* **21**: 2713–2718.
37. Bervoets, L. et al. (2004). *Environ. Toxicol. Chem.* **23**: 1973–1983.
38. Kraak, M.H.S., Scholten, M.C.T., Peeters, W.M.H., and Dekock, W.C. (1991). *Environ. Pollut.* **74**: 101–114.
39. Rainbow, P.S. (1995). *Mar. Pollut. Bull.* **31**: 183–192.
40. Baumard, P. et al. (1999). *Mar. Environ. Res.* **47**: 17–47.
41. Porte, C. and Albaiges, J. (1994). *Arch. Environ. Contamin. Toxicol.* **26**: 273–281.
42. Conti, M.E. and Cecchetti, G. (2003). *Environ. Res.* **93**: 99–112.
43. Baumard, P. et al. (1998). *Estuarine Coastal Shelf Sci.* **47**: 77–90.
44. Gutierrez-Galindo, E.A. and Munoz-Barbosa, A. (2003). *Cien. Mars.* **29**: 21–34.
45. Regoli, F. et al. (1998). *Arch. Environ. Contamin. Toxicol.* **35**: 594–601.
46. Yap, C.K., Ismail, A., Tan, S.G., and Omar, H. (2002). *Environ. Int.* **28**: 117–126.

47. Avelar, W.E.P. et al. (2000). *Water Air Soil Pollut.* **118**: 65–72.
48. Kehrig, H.D.A., Costa, M., Moreira, I., and Malm, O. (2001). *Environ. Sci. Pollut. Res.* **8**: 275–279.
49. Mcconnell, M.A. and Harrel, R.C. (1995). *Am. Malacologic Bull.* **11**: 191–201.
50. Harrell, R.C. and McConnell, M.A. (1995). *Estuaries* **18**: 264–270.
51. Prereira, W.E., Hostettler, F.D., and Rapp, J.B. (1992). *Mar. Pollut. Bull.* **24**: 103–109.
52. Stab, J.A. et al. (1995). *Environ. Toxicol. Chem.* **14**: 2023–2032.
53. Regoli, L., Chan, H.M., de Lafontaine, Y., and Mikaelian, I. (2001). *Aquat. Toxicol.* **53**: 115–126.
54. Camusso, M., Balestrini, R., and Binelli, A. (2001). *Chemosphere* **44**: 263–270.
55. Malley, D.F., Stewart, A.R., and Hall, B.D. (1996). *Environ. Toxicol. Chem.* **15**: 928–936.
56. Gewurtz, S.B., Drouillard, K.G., Lazar, R., and Haffner, G.D. (2002). *Arch. Environ. Contamin. Toxicol.* **43**: 497–504.
57. Gewurtz, S.B., Lazar, R., and Haffner, G.D. (2003). *J Great Lakes Res.* **29**: 242–255.
58. Vesk, P.A. and Byrne, M. (1999). *Sci. Total Environ.* **225**: 219–229.
59. De Maagd, P.G.J. (2000). *Environ. Toxicol. Chem.* **19**: 25–35.
60. Chapman, P.M. (1997). *Mar. Pollut. Bull.* **34**: 282–283.
61. McCarty, J.F. and Shugart, L.R. (1990). In: J.F. McCarty and R.L. Shugart (Eds.). *Biomarkers of Environmental Contamination*. Lewis Publishers, Boca Raton, FL, pp. 3–16.
62. Widdows, J., Phelps, D.K., and Galloway, W. (1981). *Mar. Environ. Res.* **4**: 181–194.
63. Smolders, R., De Coen, W., and Blust, R. (2004). *Environ. Pollut.* **132**: 245–263.

BIOCHEMICAL OXYGEN DEMAND AND OTHER ORGANIC POLLUTION MEASURES

Neil F. Pasco
Joanne M. Hay
Lincoln Ventures, Ltd.
Lincoln, New Zealand

Biochemical (also called biological) oxygen demand (BOD) is an empirical test indicator of biological activity and follows a specified laboratory procedure to determine the oxygen requirements of wastewaters, effluents, and polluted waters (1,2). The BOD test measures the amount of oxygen consumed in breaking down organic matter using both aerobic biological and chemical degradation processes, giving an indication of the organic strength of a wastewater. The test has its widest application in measuring waste loadings to treatment plants and in evaluating the BOD-removal efficiency of such treatment systems.

BOD is determined by incubating a water sample in a sealed container for a specified time and measuring the loss of oxygen that occurs between the beginning and end of the test. Samples must often be diluted prior to incubation.

WHAT DOES IT MEASURE?

Biodegradable organic matter is a composite substrate and is difficult to quantify directly. The BOD test estimates the amount of dissolved oxygen used during the degradation of organic waste as a surrogate for measuring the waste itself.

The amount of molecular oxygen measured during the incubation period can accrue from three sources:

1. The oxygen used for the biochemical degradation of organic material (carbonaceous biochemical oxygen demand, CBOD).
2. The oxygen used to oxidize inorganic materials such as sulfides and ferrous iron.
3. The oxygen used to oxidize reduced forms of nitrogen (nitrogenous demand), although a nitrifying inhibitor is often used to suppress the nitrogenous demand.

The earth's surface is mildly oxidizing and a tendency exists for reduced compounds, such as carbohydrates, to react with oxygen to form the more stable endproducts, carbon dioxide and water. BOD is the manifestation of this tendency in an aqueous environment, and two of the three sources listed above, (1) and (3), require microorganisms as the catalyst, whereas in (2), oxygen is capable of directly oxidizing some classes of inorganic materials.

Primarily, BOD values indicate the amount of carbonaceous organic material. The oxygen demand made by the oxidation of ammonia and organic nitrogen can be inhibited using 2-chloro-6-(trichloromethyl) pyridine or allyl thiourea in order to remove the nitrogenous demand and isolate the CBOD value. In the remainder of this article, BOD refers to CBOD only.

HISTORICAL CONTEXT

The complete biochemical oxidation of organic matter in a water sample can take up to 90 days, an impractical duration for an analytical test. The current international regulatory standard method for determining biodegradable organic compounds in wastewaters has been set as a five-day BOD assay (BOD_5). BOD_5 quantifies the amount of dissolved oxygen required for the microbial oxidation of carbonaceous organic material in five days under specified conditions (1). The five-day duration has no theoretical grounding but is based on an historical convention reported in 1912, as follows (3):

> In a report prepared by the Royal Commission on Sewage Disposal in the United Kingdom at the beginning of last century, it was recommended that a five-day, 18.3 °C, BOD value be used as a reference in Great Britain. These values were selected because British rivers do not have a flow time to the open sea greater than five days and average long-term summer temperatures do not exceed 18.3 °C. The temperature has been rounded upward to 20 °C, but the five-day time period has become the universal scientific and legal reference.

The standardized test was first published in the *American Public Health Association (APHA) Standard Methods* in 1917.

UNITS OF MEASURE

BOD is expressed either as milligrams of oxygen per liter (mg O_2/L) or parts per million (ppm).

WHAT IS IT USED FOR?

Despite its dubious ancestry, the BOD_5 test is still the most important and widely used environmental index for monitoring organic pollutants in wastewaters, mainly as proof of compliance with relevant legislation. Most developed countries strictly regulate the permissible BOD loadings of water as it is discharged from a treatment plant into receiving waters (e.g., sea or river), which is typically required to be less than 20 mg O_2/L. The BOD_5 value for potable water is close to zero; domestic wastewater effluent entering a treatment plant typically has a BOD_5 of around 300 mg O_2/L; and industrial effluents can have BOD_5 values greater than 20,000 mg O_2/L.

HOW DOES IT WORK?

Heterotrophic microorganisms require a metabolizable source of organic carbon, such as sugars, amino acids, nucleotides, organic acids, and fats, in order to grow. Depending on their physiology and the prevailing environmental conditions, microorganisms can harness energy from these food sources using either respiratory or fermentative biochemical pathways. Respiration is defined as the oxidation of an energy source (such as organic carbon) in which electrons are removed from the carbon source and donated to an inorganic terminal electron acceptor within the cell. Aerobic respiration uses oxygen as the terminal electron acceptor. Hence, through a complexity of genetics, physiology, and biochemistry, the breakdown of organic carbon is coupled to oxygen reduction and is facilitated by microorganisms.

Natural organic detritus plus organic waste from wastewater treatment plants, agricultural, and urban runoff provides food sources for waterborne microorganisms. The lab-based BOD_5 test is unique in that it simulates the microbial decomposition of organic wastes in the environment, thereby approximating the response of natural ecosystems (4), which helps explain the enduring relevance of the BOD test over so many years.

MEASUREMENT PROTOCOL

The BOD_5 test consists of a sample of wastewater held in a full, airtight bottle (usually 300 mL) (Fig. 1) for five days at 20 °C in the absence of light. The amount of dissolved oxygen present in the sample is measured at the beginning of the test period and at the end of five days using a dissolved oxygen probe (Fig. 2). The difference in dissolved oxygen between the day 0 and day 5 measurements represents the biochemical oxygen demand. The test requires the presence of three key ingredients: dissolved oxygen, organic substrate, and microorganisms, as prescribed by the standard test protocol. For complete specifications on how to set up and calculate a BOD_5 analysis, refer to Reference 1.

Figure 1. Standard 300 mL BOD bottle.

Figure 2. Dissolved oxygen probe.

COMPONENTS OF THE BOD_5 TEST

Dilution Water

The dilution water (saturated with dissolved oxygen) must be free of organic matter and bioinhibitory substances and has a twofold function:

- It ensures that aerobic conditions prevail throughout the test. Water samples with high BOD loadings contain insufficient dissolved oxygen for the complete oxidation of the compounds present in the sample. Without dilution, the sample will become anaerobic at some point during the five-day incubation period, causing the microorganisms to switch from a respiratory to a fermentative mode. Diluting these samples ensures that sufficient oxygen is maintained to oxidize all the organic matter present, with some residual oxygen left at the end of the test.
- It supplies the inorganic nutrients necessary for microbial growth (magnesium sulfate, calcium chloride, ferric chloride) and a buffering capacity in order to maintain a pH suitable for bacterial growth.

Seed (Microorganisms)

The seed refers to the microorganisms responsible for the biodegradation of organic matter in the sample. Although microorganisms occur naturally in the environment, a diluted water sample may require seeding with additional microbes to ensure a sufficient number of microbes are present to degrade the organic matter in the sample. The preferred seed is effluent from a wastewater treatment plant. The seed material has a BOD of its own, which must be measured separately in a control (seed in dilution water with no added substrate) and subtracted from the BOD of the samples. In a 300 mL BOD bottle, the volume of seed is typically 2 mL.

Wastewater Sample

The wastewater is the source of the organic matter to be measured. Given that biodegradable organic matter is a composite substrate, the BOD measurement is an aggregate of the wide range of substances present. It is also dependent on the microbial species present and the relative biodegradability of the substrates present. Depending on the expected BOD strength of the wastewater sample, a known volume is added to the BOD bottle and made up to 300 mL with dilution water.

STANDARDS

Glucose–Glutamic Acid (GGA)

The primary standard for calibrating BOD_5 tests is a glucose and glutamic acid solution (150 mg glucose/L, 150 mg glutamic acid/L). The GGA solution is intended to be a reference point for evaluating the dilution water quality, seed effectiveness, and analytical technique.

Organization for Economic Cooperation and Development (OECD)

The GGA solution has increasingly been questioned as a valid standard, particularly for the developing rapid BOD tests (see later). Many microorganisms preferentially use simple sugars like glucose and will actively repress the metabolism of other carbon sources when glucose is present, which has led to an underestimation of BOD values of samples containing compounds that are less readily biodegradable. A more appropriate standard solution with similar biodegradative properties to the sample being analyzed may be warranted, especially for rapid BOD tests. The OECD-defined wastewater standard solution (5) has been proposed as an alternative and more appropriate choice of standard for analysis of treated wastewater (6,7). The OECD solution, however, is not currently recognized as a regulatory standard.

QUALITY CONTROLS

No absolute BOD value of a sample exists, and BOD results are test defined, meaning BOD values are based on the parameters of the test method, not on any “true” BOD value. Some specific requirements exist that must be met for a BOD_5 analysis to be valid and which ensure that the methodology used is correct. Controls to be included each time a BOD_5 test is performed are a dilution water blank, a seed (the microorganisms) sample, and a GGA standard. If any of these quality control checks fall outside the recommended ranges (specified below), the results of the BOD_5 test should be rejected.

- *BOD blank* (bottle containing only dilution water and no seed) must not show a dissolved oxygen depletion of more than 0.2 mg O_2/L after five days.
- *Seed blank* (microorganisms added to water sample) should contribute between 0.6 and 1.0 mg O_2/L uptake per BOD bottle.
- *The glucose–glutamic acid standard* should report a BOD of 198 ± 31 mg O_2/L after a five-day incubation (i.e., the oxidation of the glucose–glutamic acid solution consumes 198 mg O_2/L over five days).

Sample dilutions should report a 50% decrease in dissolved oxygen over the five-day incubation period. At minimum, there should be an oxygen depletion of at least 2 mg O_2/L between the initial and final reading, plus a residual dissolved oxygen reading of at least 1 mg O_2/L.

HOW MUCH IS MEASURED?

Wastewater contains both biodegradable and non-biodegradable material, the latter not being represented in the BOD measurement. However, in general, the BOD_5 test does not represent the total biodegradable fraction of organic material present. Depending on the type of wastewater, the rates at which the biodegradable material in a given sample can be oxidized over five days can be between 50% and 90%.

Where the composition of the organic substrate is chemically defined, as is the case for the primary GGA standard, the amount of oxygen required for its complete oxidation can be calculated from a stoichiometric equation.

For example, the stoichiometric equation for the oxidation of glucose ($C_6H_{12}O_6$) is

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_20$$

For the complete oxidation of 1 mole of glucose, 6 moles of oxygen is required. A 150 mg glucose/L is reported to have a BOD_5 of 99 mg/L (1). On a molar basis, 150 mg glucose/L corresponds to 0.83 mmol/L, and the stoichiometric oxygen requirement is 5 mmol/L. The BOD_5 glucose requirement of 99 mg/L converts to 3.094 mmol O_2/L, and therefore the biological oxidation occurring after five days equates to 61.9% conversion. A similar exercise with glutamic acid, excluding the oxidation of nitrogen, equates to 58.4% conversion; thus the mean molar conversion of the standard glucose and glutamic acid mix in a one-hour BOD_5 test is 60.5% (8).

PROBLEMS

Although the BOD_5 test sounds simple, practical difficulties associated with it exist, and it has significant limitations. Although the test has been refined over the years, the basic approach of using a dilution technique has remained unchanged. Unfortunately, these dilutions reduce the concentration of substrates and microorganisms in the samples, thereby decreasing the overall kinetics. This decrease, along with the arbitrary time period of five days, means the actual BOD of the sample may not be reflected. The presence of toxic substances, such as copper, mercury, or chromium, even at low concentrations, in the wastewater can affect, or kill, the bacteria. A decrease in BOD_5 values may be because of a decrease in organic load or may be because of the presence of a toxin in the sample. The test is also insensitive and imprecise at low concentrations; the method requires experience and has an accepted margin of error of 15–20%, possibly because of the heterogeneity of the microbial populations in wastewater treatment systems and their widely differing responses to substrates.

However, the major drawback of the traditional BOD test continues to be the five-day measurement time. Such a delay is unacceptable for any form of active intervention for environmental monitoring and/or process control. It is suggested that the low solubility of oxygen in water is rate-limiting in the BOD_5 test; an O_2 saturated solution at 20 °C contains 9.07 mg O_2/L. The limited oxygen storage capacity of water is responsible for aqueous environments being so susceptible to nutrient deoxygenation and is why wastewater samples are often required to be substantially diluted to keep the oxygen depletion within the 1–9 mg O_2/L working range of the BOD_5 test.

ALTERNATIVE MEASUREMENT TECHNOLOGIES

In an attempt to overcome the difficulties in achieving rapid, reproducible results, a number of analytical wet chemical techniques have been developed to monitor organic pollution, as alternatives to BOD_5 measurements (9).

Chemical Oxygen Demand (COD)

The COD test measures the chemical oxidation of wastewater by using a strong oxidizing agent under acid conditions. As for BOD, it measures the amount of oxygen required as a surrogate for measuring the organic waste component directly. However, the COD test does not differentiate between biologically available and inert organic matter and measures the total quantity of oxygen required to oxidize all organic matter into carbon dioxide and water. COD values are always greater than BOD values, but the results can be obtained within two hours and the presence of toxic compounds in the sample do not affect COD measurements. In many cases, the BOD/COD ratio is used to provide an indication of the biodegradability of the wastewater. However, the chemicals used, such as acid, chromium, silver, and mercury, produce liquid hazardous waste that requires careful handling and disposal. Finally, the precision and accuracy of the COD test are questionable at low BOD values.

Total Organic Carbon (TOC)

This test measures carbon, where the carbon is oxidized by catalytic combustion (either at low temperatures by ultraviolet light and the addition of persulfate reagent, or by using a catalyst at high temperatures) to carbon dioxide, which is then measured. However, good correlations between BOD and TOC measurements have not been demonstrated, and again, the test is unable to differentiate between biodegradable and nonbiodegradable matter. The main advantage of the TOC test is its speed; determinations can be made within minutes, facilitating a greater number of measurements compared with either BOD or COD tests.

RAPID BOD TECHNOLOGIES

Oxygen-Based Sensors

In response to the restrictive five-day time lag inherent in the traditional BOD_5 test, a number of rapid BOD techniques have been devised. They have evolved down two development paths, both respiratory-based: microbial biosensors and respirometers.

Microbial BOD biosensors consist of a microbial biofilm immobilized onto a dissolved oxygen probe. Since the first BOD biosensor of this sort was developed (10), various sensor designs have been reported, and their response time ranges from 3 to 30 minutes. In all reported cases, the decrease in dissolved oxygen forms the basis of the analytical signal. One consequence of oxygen limitation for these sensors is that the amount of organic material biodegraded in a short-time assay is a very small fraction (~1%) of the total biodegradable organic content, and they can only respond to the easily assimilated compounds in the wastewater.

Commercially available respirometers, such as RODTOX®, BIOX®, BIOMONITOR®, and RACOD®, are designed to replicate the operation of a wastewater treatment plant, rather than the methodology of the

BOD test. They consist of one or more activated-sludge bioreactors maintained at steady-state conditions by continuous stirring. The inclusion of organic material increases microbial respiration, in turn increasing the rate of oxygen uptake, which is recorded by an oxygen sensor. Oxygen consumption is then correlated with BOD; these devices also report a short-term BOD estimation in 3–30 minutes.

Nonoxygen-Based Sensors

A second generation of rapid BOD sensors have been reported based on the measurement of charge transfer, either to an alternative electron acceptor, such as a synthetic redox-active mediator (8,11), or directly to an electrode (12), rather than the depletion of oxygen.

Mediated Sensors

The oxidation (loss of electrons and the concurrent transfer of charge) of organic substrates can be coupled to the reduction of a redox mediator, such as potassium hexacyanoferrate. These mediators readily accept or donate (i.e., shuttle) electrons in biochemical reactions. In a BOD application, the redox mediator substitutes for oxygen as the terminal electron acceptor during aerobic oxidation and transfers the electrons to an electrode, generating an analytical signal that can be converted to a BOD value. By replacing oxygen with a more soluble synthetic mediator, higher numbers of microorganisms can be incorporated into the assay, significantly reducing the incubation time to one or two hours (8).

Direct Electron Transfer

Direct electron transfer to an electrode is exhibited by some microorganisms, such as *Shewanella putrefaciens*, thereby eliminating the need for a redox mediator. The direct reducing action of the electrochemically active microorganisms can be exploited in a microbial fuel cell device, and the current produced has been shown to be dependent on the concentration of organics in the wastewater water sample (12).

CONCLUSION

For all its shortcomings and the availability of a number of related measurement techniques, the BOD_5 assay remains the preferred test for reporting the oxygen requirements of wastewaters, effluents, and polluted waters, which is, in part, because of its international regulatory status, but also because it uniquely characterizes the impact of oxygen depletion developing from biological activity in breaking down organic matter on receiving waters.

Concern over water quality continues to increase and has led to the implementation of stricter environmental pollution control legislation worldwide. Legislation, such as the European Urban Wastewater Treatment directive, is framed in terms of compliance with BOD_5 and will entrench the test even further. Although the need for a more rapid test is well recognized, the success of these alternative tests will remain limited until they can demonstrate robust comparability with the BOD_5 measurement.

BIBLIOGRAPHY

1. APHA (American Public Health Association). (1995). *Standard Methods for the Examination of Water and Wastewater*, 19th Edn. A.D. Eaton, L.S. Clesceri, and A.E. Greenberg (Eds.). APHA, Washington, DC, pp. 5.1–5.9.
2. JIS. (1993). *Japanese Industrial Standard Committee Testing Methods for Industrial Wastewater*. Tokyo, Japan.
3. Tchobanoglous, G. and Schroeder, E. (1987). *Water Quality*. Addison-Wesley, Reading, MA.
4. Clark, D.W. (1992). BOD: the modern alchemy. *Public Works* **50–51**: 86–89.
5. OECD. (1981). *OECD Guideline 209—Activated Sludge, Respiration Inhibition Test*. OECD Guideline for testing of chemicals. OECD, Paris, France.
6. Tanaka, H., Nakamura, E., Minamiyama, Y., and Toyoda, T. (1994). BOD biosensor for secondary effluent from wastewater treatment plants. *Water Sci. Technol.* **30**: 215–227.
7. Liu, J. and Mattiasson, B. (2002). Microbial BOD sensors for wastewater analysis. *Water Res.* **11**(15): 3786–3802.
8. Pasco, N.F., Baronian, K.H., Jeffries, C., and Hay, J. (2000). Biochemical mediator demand—a novel rapid alternative for measuring biochemical oxygen demand. *Appl. Microbiol. Biotechnol.* **53**(5): 613–618.
9. Bourgeois, W., Burgess, J.E., and Stuetz, R.M. (2001). On-line monitoring of wastewater quality: a review. *J. Chem. Technol. Biotechnol.* **76**(4): 337–348.
10. Karube, I., Matsunaga, T., Mitsuda, S., and Suzuki, S. (1977). Microbial electrode BOD sensors. *Biotechnol. Bioeng.* **19**: 1535–1547.
11. Yoshida, N., Yano, K., Morita, T., McNiven, S.J., Nakamura, H., and Karube, I. (2000). A mediator-type biosensor as a new approach to biochemical oxygen demand estimation. *Analyst* **125**(12): 2280–2284.
12. Kim, M. et al. (2003). Practical field application of a novel BOD monitoring system. *J. Environ. Monitor.* **5**(4): 640–643.

BIODEGRADATION

Christian D. Frazar
Silver Spring, Maryland

The beginnings of life on this planet arose during a time of extreme conditions. For life to arise in an inhospitable world, the first organisms needed to be simple but hearty. The life forms of today considered most similar to those early ancestors of life are bacteria and archaea—simple, primitive organisms. Despite their primitive nature, these organisms have a diverse and powerful metabolic potential.

As life and other processes transformed the earth's air, land, and water, early organisms were in a new and changing environment. A wider variety of ecological niches became available. For five-sixths of the earth's existence, it was inhabited exclusively by microorganisms (1). Given their short generation time, microorganisms have had unfathomable generations to adapt to earth's many chemicals. As a result it is quite possible that for every natural compound that exists, there is a microorganism that has evolved to metabolize it and use it for growth.

Microorganisms are at the base of the food chain and play a key role in both secondary production and in decomposition. The ability of microorganisms to decompose or break down compounds is known as biodegradation. This term is often used to refer to the degradation of environmental pollutants by bacteria or fungi. This discussion will focus on the degradation process itself and the factors that influence and limit biodegradation.

MICROBIAL DIVERSITY AND METABOLIC POTENTIAL

The use of biodegradation to remove a contaminant from the environment is known as bioremediation. Understanding microbial diversity and metabolic potential is essential to understanding how to use biodegradation for remediation projects. During the twentieth century, a vast number of synthetic and natural compounds were released into the environment. A complex community of microorganisms often has the mechanisms to use or detoxify many of the compounds that humans have released.

Some organisms have a very narrow range of metabolic potential, and others have a wider capacity. For example, it has been demonstrated that the organism *Pseudomonas putida,* strain F1, degrades approximately 105 compounds, and it has been suggested that it can degrade many more (2). Other compounds, many of which are synthetic, are resistant to microbial degradation. Resistant compounds are known as recalcitrant compounds; many factors contribute to a compound's recalcitrance, which will be discussed later.

Many synthetic compounds are considered xenobiotics. The term xenobiotic is often used to refer to compounds that are foreign to biological systems. The term originates in mammalian systems, not microbial systems. It does not always describe a compound accurately in microbial terms, because microbial diversity is so vast. What is foreign to one group of bacteria may be readily metabolized by a different group.

Microorganisms benefit from biodegradation by metabolizing a compound to yield carbon and energy. Energy is derived from the transfer of electrons and the conservation of that energy in the form of molecules such as adenosine triphosphate (ATP) and nicotinamide-adenine dinucleotide (NAD^+). ATP conserves this energy through high energy phosphate bonds and is used to drive energy-requiring reactions in the cell. NAD^+ is an electron carrier involved in oxidation–reduction reactions. If metabolism results in complete breakdown of a compound to inorganic products such as CO_2 and H_2O, this is referred to as mineralization. As the compound of concern disappears from the environment, the number of organisms that use this compound as a substrate for metabolism increases. This is known as growth-linked biodegradation, and mineralization is a common result.

ENRICHMENT CULTURES AND BIODEGRADATION RESEARCH

When research scientists attempt to identify an organism capable of degrading a particular compound, they collect microorganisms from the environment and inoculate them into a culture flask containing minimal medium, and the compound of interest provides the sole carbon source. The minimal medium provides other nutrients, such as nitrogen and phosphorus, that are necessary for microbial growth. Organisms unable to use this compound will not be able to grow, whereas those that can metabolize the compound will flourish. The organisms for this enrichment culture are often collected from a site where the compound of interest is found naturally or as a contaminant. For instance, if a researcher were attempting to isolate a xylene-degrading microorganism, a sample might be collected from an area of a gasoline spill because xylene is a component of gasoline.

Scientists need to be cautious of optimism because one organism may seem to degrade a chemical efficiently under laboratory conditions, but the actual environment where the degradation is to take place may have many inhibiting factors that cannot be controlled.

It is important for researchers to identify degradation products because some products are more toxic than the parent compound. For example, the hydrolysis of the herbicide 2,4-D is catalyzed by soil microorganisms, which yields a form of the compound toxic to plants (3).

FACTORS LIMITING BIODEGRADATION

The mere presence of a compound and a capable microorganism does not guarantee the compound's degradation. Microbial growth requires water, nitrogen, phosphorus, trace elements, carbon, energy, temperature, and suitable pH conditions. Biodegradation commonly supplies the carbon or energy or both. But growth may be inhibited by other toxins or limited by other essential elements or conditions. Under environmental conditions, predation can also be a factor that limits biodegradation. Protozoa consume bacteria; bacteriophages are viruses that can infect and lyse bacteria.

If no organisms can use the compound as the sole carbon or energy source, a compound may still be cometabolized. Cometabolism involves the degradation of a compound by an organism without providing carbon, energy, or any other factors required for growth. Unlike mineralization, cometabolism is not growth-linked. Even though the compound is not functioning as the carbon or energy source, it is possible for degradation to be quite rapid (4). Often other organic molecules in the environment provide the carbon or energy source, which leads to an increase in the microbial population that is responsible for the increased rate of degradation.

ENZYMATIC MECHANISMS

The proper enzyme must exist for metabolism or transformation of a compound. Over geologic time, microorganisms have adapted to compounds and structures to which they are exposed regularly. It is often believed that microbial enzymes have not evolved rapidly enough to cope with the onslaught of synthetic compounds that have been released into the environment in the last century (2).

Biodegradation may occur inside or outside the cell. Intracellular enzymes require the organism to have some means of transporting the compound inside the organism for degradation. Extracellular enzymes are released by the organism and degrade the compound before the products are transported into the cell, where the final degradation occurs. This is a common microbial tactic when the carbon or energy source is too large to internalize. Extracellular enzymes are manufactured in comparatively high concentrations compared to intracellular enzymes.

Degradative enzymes are often highly specific for a particular reaction. The enzyme must be able to fit the correct molecular structure into its active site. The active site of an enzyme is responsible for the cleavage or molecular rearrangement of a portion of the compound. The three-dimensional structure of the functional group fits neatly into the active site of the enzyme. Functional groups are often the targets of these enzymes (Table 1). Compounds can have a combination of functional groups, and the transformation of these compounds will depend on the spatial arrangement of the functional groups and the enzyme's ability to catalyze the degradation products. If it is known that an organism can degrade a structurally similar compound, then it may be possible to enrich for an organism capable of degrading the compound of concern.

ACCLIMATION PHASE

Initially, the organisms in the environment that can degrade a particular compound will often be insufficient to cause a noticeable change in the quantity of compound present. This initial period, referred to as the acclimation phase, can vary in length among compounds, environments, and organisms. Bioavailability and compound concentration can influence the length of the acclimation phase. If the compound is present in the environment at a concentration that is too high, then it can be toxic to organisms that would break it down (5).

Table 1. Common Functional Groups in Contaminants

Name	Structure	Example(s)
Halogen	R–X where X=Cl, Br, or F	Metolachlor
Aromatic ring	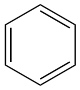	Benzene, toluene, ethylbenzene, xylenes (BTEX), polyaromatic hydrocarbons (PAHs)
Carboxcylic acid	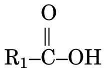 $R_1{-}\overset{\displaystyle O}{\overset{\|}{C}}{-}OH$	Picloram
Alkene	$R_1{-}C{=}C{-}R_2$	Aldrin, dieldrin, and heptachlor
Alcohol (hydroxyl group)	$HO{-}R_1$	2,4-D
Ester	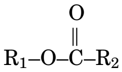 $R_1{-}O{-}\overset{\displaystyle O}{\overset{\|}{C}}{-}R_2$	Malathion
Ether	$R_1{-}O{-}R_2$	2,4-D, MTBE

Research has shown that organisms that have been exposed to a particular compound can respond more quickly to a second exposure to the same compound (6). The first exposure selects for a community of organisms capable of degrading the compound. Upon reintroduction of that same compound, the community is already capable of degrading the compound, and the acclimation phase is shortened. The shortened acclimation phase can be attributed either to a more metabolically active community or a more rapid increase in the number of organisms capable of degrading the compound (6). This response has been identified among organisms capable of degrading several different pesticides (6).

RECALCITRANCE

Recalcitrant molecules are either resistant to degradation or are degraded at a very slow rate. There are several reasons that this occurs. As suggested previously, it is possible that the necessary enzyme does not exist. It is also possible that the correct enzyme may exist, but it cannot reach the active site and catalyze the necessary reaction. Finally, the enzyme necessary for the degradation of the compound may never be turned on by the organism. In this case, the enzyme is encoded on the organism's DNA, but no signal induces enzyme production. Some compounds, such as 2,4-D, are readily degraded under aerobic conditions but not under anaerobic conditions (4). Other compounds, such as the organophosphate herbicide, fenitrothion, are preferentially degraded under anaerobic conditions (4). Thus, the environment can play a role in a compound's recalcitrance.

BIOAVAILABILITY

A compound must be bioavailable before degradation will occur. The organism and its enzymes must be capable of coming into contact with the compound. To understand fully the limitation that bioavailability presents, we need to examine the microscopic scale on which microorganisms exist. Most bacteria are no longer than 10 μm, and the distance between one soil particle and another may be a journey that a bacteria never makes. Most bacteria in soil or sediment are attached to particles, but some are motile. These motile organisms often simply pass through pore spaces and do not really contribute to the microenvironment. If examined in great detail, it would be possible to see that the surfaces of soil and sediment particles consist of depressions and valleys, which provide habitats for native microorganisms. Some microorganisms will never leave these tiny spaces. These organisms rely on the movement of compounds and nutrients for survival. Some pores are too small for microorganisms but are large enough to provide a protective area for the compound.

Over time, the bioavailability of a compound tends to decrease due to sequestration of the compound. When the compound is plentiful, microorganisms will use the molecules that are easiest to obtain. Available compound is used, whereas the excess becomes bound to soil and sediment particles. Compound bioavailability then limits biodegradation. Researchers have attempted to model the diffusion and sorption of a compound to soil or sediment (7). One model is the diffusion–sorption bioavailability model that attempts to describe the rate of biodegradation in the presence of soil or sediment aggregates (7). Many biodegradation studies are done in flasks and do not take into account environmental factors such as diffusion and sorption.

NONAQUEOUS PHASE LIQUIDS AND COMPOUNDS OF LOW WATER SOLUBILITY

Nonaqueous phase liquids (NAPLs) generally have low water solubility. Dense nonaqueous phase liquids can be found below an aquifer, whereas light nonaqueous phase liquids are found on the surface of the water. A very small fraction of a NAPL is soluble in water, and this quantity is available to microorganisms for biodegradation. The vast majority of NAPLs are not in the aqueous phase, so they are often considered unavailable to the organisms capable of degrading or metabolizing them, thus presenting a factor limiting biodegradation. NAPLs that have polluted the environment include hydrocarbons (mainly from oil spills) and industrial solvents.

Microbial growth on compounds of low water solubility is generally slow; however; in some cases, exponential growth has been observed (8). There are several ways by which a compound that has low water solubility can be degraded. The small fraction of the compound that enters the aqueous phase is readily accessible to the organism. The organism also can come in direct contact with the compound at the water–compound interface (9). Finally, some organisms can make the compound more accessible through a process known as pseudosolubilization (9). Pseudosolubilization occurs via compounds released by the organism, which act as carrier molecules (10). The compound becomes enclosed in a micelle, which may diffuse through the aqueous phase (10). Microorganisms in the aqueous phase can then use the carbon source within the micelle. Surface active compounds can also be released by the organism. These compounds are composed of various lipids and interact with low solubility compounds such as hydrocarbons to facilitate their uptake by the organism (11,12).

GROUNDWATER BIOREMEDIATION

Groundwater contamination is of high priority because so many people rely on groundwater as a source of drinking water. Porosity, the empty space between soil and sediment grains, determines how much water an aquifer can hold. Water moves slowly through underground aquifers and that movement is a function of the type of soil or rock that composes the aquifer. Darcy's law allows scientists to approximate the discharge of water from an aquifer and takes the material that makes up the aquifer and the hydraulic gradient into consideration. Once the plume is located and the concentration of compound has been determined, removing it can be difficult and costly.

For *in situ* biodegradation to occur, it is often necessary to supply oxygen and nutrients to the microorganisms. The absence or presence of oxygen can make a substantial difference to the active microbial community and its ability to degrade the contaminant. The contaminated groundwater can also be pumped from the ground into a reactor, where microorganisms degrade the contaminant (for a review, see Reference 13). These reactors function similarly to sewage treatment plants. Contaminated groundwater and nutrients are pumped into a reactor where microorganisms are grown so that they are attached to the reactor. As the contaminated groundwater passes the microorganisms, they degrade the compound. The water is then disposed of or returned to the aquifer. Bioreactors are often limited by their high costs. Pumping of NAPLs is often difficult due to their low solubility in water.

MARINE OIL SPILLS

Marine oil spills can be remediated by bioremediation. Dispersants are used to spread the nonaqueous phase contaminants across the water's surface to create more favorable biodegradation conditions. The dispersants themselves are biodegraded in the process. It is estimated that 50% of the oil from the *Exxon Valdez* spill was biodegraded on the water surface, on the shore, or in the water column (14). Fertilizers are often used to enhance the rate of oil degradation and to supply otherwise limiting nutrients.

ALGAL BIOREMEDIATION

Biodegradation is not limited only to bacteria. Fungi play an important role in soil systems; in many aquatic systems, algae can play an important role. Several algal species are capable of degrading some pesticides in freshwater systems (15). Algae can be used to clean industrial pulp mill effluent before the effluent is discharged. They are effective in removing both the dark color of the effluent and organic halides (16). Some algae can bioaccumulate some polycyclic aromatic hydrocarbons in marine systems (17). Bioaccumulation is the sequestration of a compound by an organism, where the compound remains unchanged and the organism is unharmed by its presence. Over time, significant quantities of compound can be accumulated by the organism.

The metabolic diversity of microorganisms is a powerful tool for degrading unwanted compounds. Numerous sources are available to provide additional information regarding biodegradation processes. The journal *Biodegradation*, published quarterly, contains scientific advancements in biodegradation. Biodegradation databases are available that provide information on biodegradable compounds and the enzymes and microbial species that degrade them. A useful database is the University of Minnesota Biocatalysis/Biodegradation Database, found on the web at: www.umbbd.ahc.umn.edu/index.html.

BIBLIOGRAPHY

1. Madigan, M., Martinko, J., and Parker, J. (2000). *Brock: Biology of Microorganisms*, 9th Edn. Prentice-Hall, Upper Saddle River, NJ.
2. Wackett, L. and Hershberger, C. (2001). *Biocatalysis and Biodegradation: Microbial Transformation of Organic Compounds*. ASM Press, Washington, DC.
3. Vlitos, A. and King, L. (1953). Fate of sodium 2,4-dichlorophenoxy-ethyl-sulphate in the soil. *Nature* **171**: 523.
4. Liu, D. et al. (2000). Factors affecting chemical biodegradation. *Environ. Toxicol.* **15**: 476–483.
5. Saxena, A., Zhang, R., and Bollag, J. (1987). Microorganisms capable of metabolizing the herbicide metolachlor. *Appl. Environ. Microbiol.* **53**: 390–396.
6. Robertson, B.K. and Alexander, M. (1994). Growth-linked and cometabolic biodegradation: possible reason for occurrence of absence of accelerated pesticide biodegradation. *Pest. Sci.* **41**: 311–318.
7. Scow, K. and Hutson, J. (1992). Effect of diffusion and sorption kinetics of biodegradation: theoretical considerations. *Soil Sci. Soc. Am. J.* **56**: 119–127.
8. Stucki, G. and Alexander, M. (1987). Role of dissolution rate and solubility in biodegradation of aromatic compounds. *Appl. Environ. Microbiol.* **53**: 292–297.
9. Goswami, P.C., Singh, H.D., Bhagat, S.D., and Baruah, J.N. (1983). Mode of uptake of insoluble solid substrates by microorganisms. I: sterol uptake by arthrobacter species. *Biotechnol. Bioeng.* **25**: 2929–2943.
10. Velanker, S.K., Barnett, S.M., Houston, C.W., and Thompson, A.R. (1975). Microbial growth on hydrocarbons—some experimental results. *Biotechnol. Bioeng.* **17**: 241–251.
11. Geurra-Santos, L., Käppeli, O., and Fiechter, A. (1984). *Pseudomonas aeruginosa* biosurfactant production in continuous culture with glucose as a carbon source. *Appl. Environ. Microbiol.* **48**: 301–305.
12. MacDonald, C.R., Cooper, D.G., and Zajic, J.E. (1981). Surface-active lipids from *Nocardia erthropolis* grown on hydrocarbons. *Appl. Environ. Microbiol.* **41**: 117–123.
13. Langwaldt, J.H. and Puhakka, J.A. (2000). On-site biological remediation of contaminated groundwater: a review. *Environ. Pollut.* **107**: 187–197.
14. Sugai, S., Lindstrom, J., and Braddock, J. (1997). Environmental influences on the microbial degradation of *Exxon Valdez* oil on the shorelines of Prince William Sound, Alaska. *Environ. Sci. Technol.* **31**: 1564–1572.
15. U.S. EPA. (1976). *Degradation of Pesticides by Algae*. Office of Research and Development, Environmental Research Laboratory, Athens, GA.
16. Dilek, F.B., Taplamacioglu, H.M., and Tarlan, E. (1999). Colour and AOX removal from pulping effluents by algae. *Appl. Microbiol. Biotechnol.* **52**: 585–591.
17. Kirso, U. and Irha, N. (1998). Role of algae in fate of carcinogenic polycyclic aromatic hydrocarbons in the aquatic environment. *Ecotoxicol. Environ. Safety* **41**: 83–89.

BIOLUMINESCENT BIOSENSORS FOR TOXICITY TESTING

Jim Philip
Jennifer Bell
Fiona Stainsby
Sandra Dunbar
Napier University
Edinburgh, United Kingdom

Siouxsie Wiles
Imperial College London
London, United Kingdom

Andrew Whiteley
Mark Bailey
Centre for Ecology and
Hydrology–Oxford
Oxford, United Kingdom

BACKGROUND

This article summarizes recent developments in whole cell bioluminescent biosensors for estimating toxicity at wastewater treatment plants (WWTPs). This started with the use of natural marine bioluminescent bacteria, and there is now at least a 20-year history of toxicity testing with these organisms. Molecular biology has given rise to many new bioluminescent strains for toxicity testing to overcome some specific limitations of *Photobacterium (Vibrio) fischeri*. Luminescence decay toxicity tests have also been developed in higher organisms, up to *Caenorhabditis elegans* (1), to try to improve the relevance of these tests. Custom-made biosensors for protecting wastewater treatment plants are already in development. These developments coincide with the pattern of environmental legislation in Europe.

LEGISLATIVE DRIVERS

The Urban Wastewater Treatment Directive and the Water Framework Directive

In the European Union, under the Urban Wastewater Treatment Directive (2), the functioning of wastewater treatment works is considered important, and reaching nontoxic levels is of primary concern (3). In common with most legislation in the area, the Directive is directed toward regulating discharges to the environment, rather than to the WWTP itself. The more recent EU Water Framework Directive (4) is concerned mostly with protecting receiving waters from pollution and toxic discharges.

IPPC

The Integrated Pollution Prevention and Control Directive, adopted in 1996 (5), is one of the cornerstones of the EU's environmental legislation. It is likely that fines based on toxicity levels, in addition to the current physicochemical parameters, will be assessed on companies that discharge industrial effluents at concentrations above guidelines (6).

ENVIRONMENTAL QUALITY STANDARDS

Industrial effluents tend to contain more toxic substances than domestic wastewater, and numerical limits are set for discharges to ensure compliance with Environmental Quality Standards (EQSs). Compliance is then monitored by chemical analysis. There are several problems in this approach:

- There are many substances for which there is no EQS (over 99%).
- There are no ecotoxicological data for thousands of chemicals, and most EQSs are based on limited data.
- Difficulties are experienced in predicting the interaction of chemicals with each other and the subsequent effect on the environment.
- There are analytical difficulties for many chemicals, and great cost in separating and identifying all constituents.

TOXICITY CONSENTS AND DIRECT TOXICITY ASSESSMENT

In 1993, the predecessor to the U.K. Environment Agency, the National Rivers Authority (NRA), began research into setting discharge consents based on toxicity. This became known as Direct Toxicity Assessment (DTA). DTA provides a direct measure of acute toxicity, and it is not necessary to identify the substances causing effluent toxicity to treat or reduce it (7); a property can be measured and treated directly. If toxicity is detected but considered acceptable, a toxicity consent can be derived and applied. If the toxicity of the effluent is unacceptably high, however, remedial action may be needed, and then a toxicity consent can be derived (Fig. 1). In the United States, a similar process is applied to discharges to receiving waters; the process is called Whole Effluent Toxicity (WET).

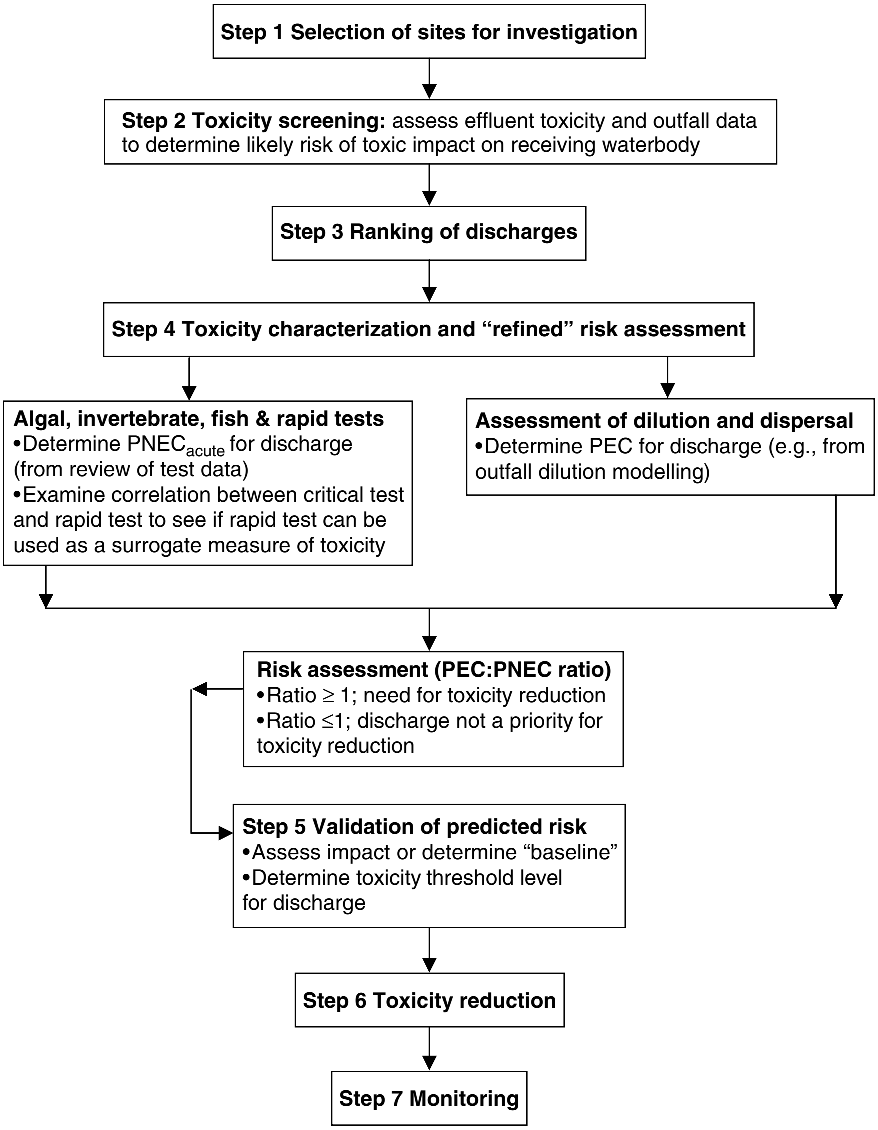

Figure 1. The Direct Toxicity Assessment process.

Two Types of "Legislative" Toxicity

From the above, it can be recognized that there is a need to measure toxicity in two different senses. On the one hand, there is the need to protect the biological WWTP, especially the activated sludge plant, from toxic influents that can kill the treating biomass, leading to loss of treatment and concomitant financial and regulatory implications. On the other hand, there is the need to measure the toxicity of the treated effluent from the WWTP for environmental protection.

Toxicity, An Imperfect Analyte

Toxicity is not a well-defined analyte for a biosensor. Acute toxicity refers normally to lethality as a result of exposure to a high concentration of a toxicant for a short period. Chronic toxicity may be lethal or sublethal but refers to exposure to a low dose for a long period. Although chronic toxicity does not necessarily cause death, it can result in long-term debilitation of some form.

Nevertheless, to make meaningful comparisons, it is necessary to put numbers to toxicity. These are some of the most commonly encountered numbers:

- LC_{50}: the concentration that produces 50% mortality in test organisms;
- LD_{50}: the dose that produces 50% mortality in test organisms;
- IC_{50}: the concentration at which 50% of growth or activity is inhibited;
- EC_{50}: concentration at which 50% of the predicted effect is observed;
- LOEC (the lowest observable effect concentration): the concentration at which the lowest effect is seen;
- NOEC (the no observable effect concentration): the maximum concentration at which no effect is observed.

In bioluminescence tests, the figures quoted are usually EC_{50} or IC_{50}. Their direct correlation with lethality (e.g., LC_{50}) in higher organisms is fraught with difficulty. It should be expected that the EC_{50} will be lower than the LC_{50} because impairment of function should precede total inhibition (8).

THE RELATIONSHIP BETWEEN BACTERIAL BIOLUMINESCENCE AND TOXICITY

The bacterial bioluminescence reaction involves oxidation of a long-chain aldehyde (RCHO) and reduced flavin mononucleotide ($FMNH_2$), resulting in the production of oxidized flavin (FMN) and a long-chain fatty acid (RCOOH), along with the emission of light (Fig. 2).

$FMNH_2$ production depends on functional electron transport, so only viable cells produce light. This relationship between cellular viability and light endows bioluminescence with the ability to report bacterial injury and recovery. It can be established readily that any substance that impairs the biochemistry will lead to a reduction in light output. Of crucial importance to the commercial development of the bacterial bioluminescence toxicity tests was the fact that the change in light output is proportional to the concentration of the toxicant. Thus, it was possible to develop quantitative tests and compare the results to toxicity tests with higher organisms (Fig. 3).

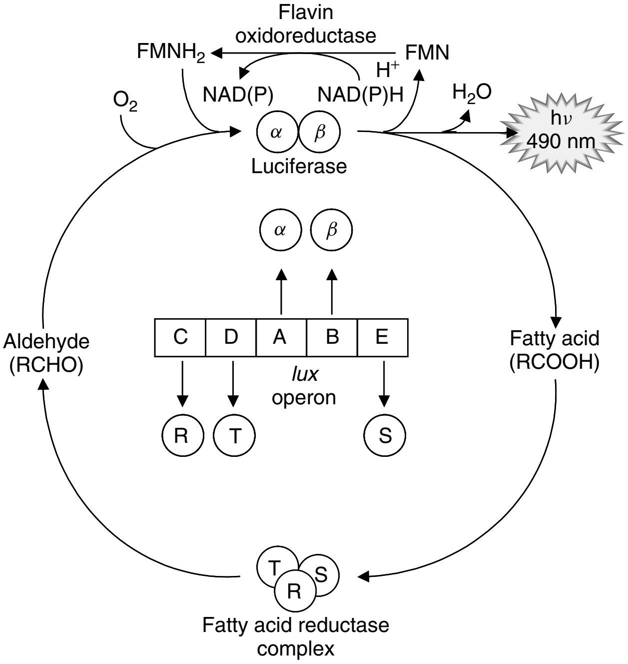

Figure 2. Genetics and biochemistry of bacterial bioluminescence. The rectangles represent genes and the circles represent enzymes or enzyme subunits.

BIOLUMINESCENT TOXICITY BIOSENSORS: POTENTIAL FOR USE IN DTA

What is required of these sensors is not necessarily greater sensitivity, resulting in lower EC_{50} values. The real value to DTA of treated effluents discharged to receiving water bodies would be relevance (good correlation) to the results of toxicity tests with higher organisms. The tests based on naturally occurring marine bioluminescent bacteria as used in the Microtox® test or the ToxAlert® test are in widespread use, and they correlate inhibition of luminescence with toxicity (9–12).

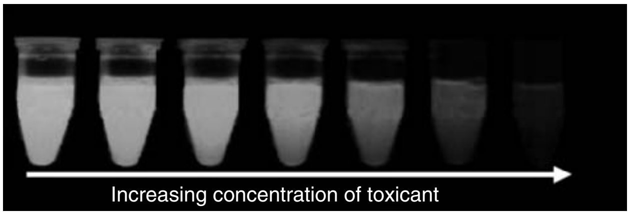

Figure 3. A series of tubes containing bioluminescent bacteria responding to increasing concentrations of toxicant by producing less light. (Courtesy of Remedios Ltd, http://www.remedios.uk.com/.)

The tests are rapid, and reliable toxicity testing has been demonstrated (13–19). Their use has resulted in extensive databases on the toxicity of pure chemicals (18). The potential for using the Microtox test as an alternative to animal testing has been demonstrated (19). It was found more sensitive and less prone to false positives than two common toxicity tests, the L-929 Minimal Eagle's Medium (MEM) elution cytotoxicity test and the Draize rabbit eye irritancy test. It offers the often quoted advantages of low cost and convenience, and it is far less controversial than these latter tests.

Despite these advantages, application to environmental samples has not always been successful (20,21), and correlation with other toxicity assays may require much more intrusive modification. In particular, the marine origin of these natural bacteria may act as a barrier to their use in DTA. For example, they are sensitive to pH and osmotic conditions. The revolution in genetic engineering, however, has made it routine to transfer genes for bioluminescence to other bacteria. This has opened up new possibilities for biosensor design.

BIOLUMINESCENT TOXICITY BIOSENSORS: MONITORING THE PERFORMANCE OF WWTPs

The limitations of marine bioluminescent bacteria become much more apparent when the biosensor is to be used for predicting toxicity in WWTPs. Then, it is not the treated, biodegraded effluent that is to be monitored, but the toxic influent before treatment; so the concentrations of pollutants are much higher, depending on the particular industrial wastewater. The issues for the plant operator are somewhat similar to those of DTA, regulatory and financial if a fine is levied for noncompliance. However, the situation for the operator is more serious if the influent to a biological wastewater treatment plant is toxic to the microbial community that performs the purification.

LOSS OF TREATMENT SPELLS REGULATORY AND FINANCIAL TROUBLE

Partial or complete loss of treatment through toxicity to mean liquor suspended solids (MLSS) will cause out-of-compliance and can also result in serious operational problems (such as sludge bulking and foaming) and the attendant need for untreated disposal or storage. In the event of total loss of viability, restoring the biological plant can take weeks to months, with very serious financial and regulatory consequences. Until recently, total loss of viability has been considered anecdotal, but real instances have been recorded. Much more frequent occurrences are instances of partial or chronic toxic events. For example, up to 50% inhibition of nitrification has been attributed to toxicity from industrial wastewater (22).

MARINE BIOLUMINESCENCE TOXICITY TESTS ARE NOT FIT FOR THE CHALLENGE

The irrelevance of marine bioluminescent biosensors can be demonstrated with phenol as an example. During coal coking, an aqueous effluent is produced containing several hundred mg/L of phenols. Biological treatment of these wastewaters is an accepted full-scale practice; high efficiencies of removal are achievable, but occasionally treatment fails through toxicity to the treating biomass. However, the 5-minute EC_{50} for phenol to the Microtox strain is 18 mg/L (23). Used as a toxicity biosensor upstream of such an activated sludge plant, this strain would cause continuous false alarms while the biodegrading community functioned normally.

ENTER GENETIC ENGINEERING

Thus, there is a need for biosensor strains that are more relevant to the wastewater treatment system under consideration. The best correlations between toxicity tests and the actual behavior of an activated sludge plant are obtained when microorganisms from the activated sludge plant are used in toxicity tests (24). In this manner, Kelly et al. (20) developed a bioluminescent toxicity biosensor by inserting the *luxCDABE* genes from *V. fischeri* into a *Pseudomonas* strain from an activated sludge plant. By similar means, a wide range of alternative luminous bacteria has been developed for use in toxicity testing of contaminated waters (25–30).

MEASURING LIGHT PRODUCTION AND DECAY

One reason for the popularity of bioluminescence biosensors is the ease of detecting and quantifying light. Light does not accumulate or diffuse and facilitates real-time and *in situ* measurements. It can be measured linearly over several magnitude orders, an extremely important consideration. As a result, different techniques are available for detection. The most widely used is luminometry.

Luminometry is rapid and sensitive, and its sensitivity is increasing with time due to technical innovations. Laboratory-based luminometers generally rely on high sensitivity photomultiplier tubes and have a range of advanced features. For field deployment of toxicity biosensors, a more portable instrument is required. The choice is normally silicon photodiode technology, which has advanced rapidly in the last 10 years, so that very inexpensive, compact, and sensitive photodiodes are now available, making them suited to field use. The technology has advanced to the stage where an instrument is available with 24 individual photodiodes so that a 24-well microtiter plate can be read in an instrument that is small enough to be handheld (Fig. 4).

Example Data

A generic example is presented here, where a handheld instrument was used to measure bioluminescent decay from a genetically modified biosensor designed to respond to phenol (30). The data in Fig. 5a show the response of the sensor to different concentrations of a toxicant (in this case trichlorophenol) with time. The instrument measured light every second, and each decay curve represents a concentration of toxicant. Note that an increasing concentration of toxicant depresses the end point of the

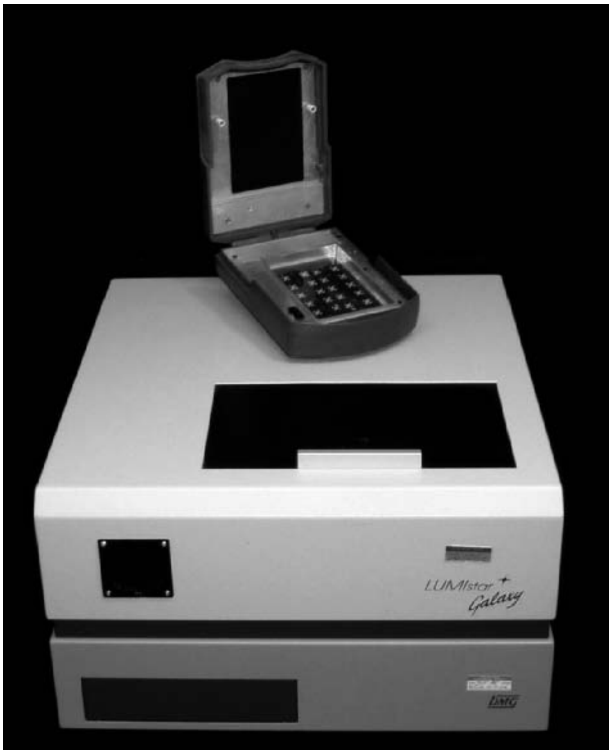

Figure 4. Two types of luminometer. On the bottom is a research laboratory instrument based on photomultiplier technology. On the top is a silicon photodiode instrument containing multiple photodiodes for microtitre plate reading (courtesy of Cybersense Biosystems, http://www.cybersensebiosystems.com/).

reaction with reference to a nontoxic control. These data were then manipulated to give a mean endpoint figure, which was then plotted against concentration (Fig. 5b).

BIBLIOGRAPHY

1. Lagido, C., Pettitt, J., Porter, A.J.R., Paton, G.I., and Glover, L.A. (2001). *FEBS Lett.* **493**: 36.
2. European Commission Council Directive. (1991). Urban Waste Water Treatment Directive 91/271/EEC. EEC, Brussels.
3. Farré, M. and Barceló, D. (2003). *Trends Anal. Chem.* **22**: 299.
4. European Commission Parliament and Council Directive. (2000). *Water Framework Directive* 2000/60/EC. EEC, Brussels.
5. European Commission Council Directive. (1996). *Integrated Pollution Prevention and Control Directive* 96/61/EC. EC, Brussels.
6. dos Santos, L.F., Defrenne, L., and Krebs-Brown, A. (2002). *Anal. Chim. Acta* **456**: 41.
7. Scottish Environmental Protection Agency. (2003). *Direct Toxicity Assessment*. Guidance no. 03-DLM-COPA-EA3.
8. Connell, D.W., Lam, P., Richardson, B., and Wu, R. (1999). *Introduction to Ecotoxicology*. Blackwell Science, Oxford.
9. Bulich, A.A. and Isenberg, D.L. (1981). *ISA Trans.* **20**: 29.
10. Isenberg, D.L. (1993). The Microtox™ toxicity test a developer's commentary (I). In: *Ecotoxicology Monitoring*. M. Richardson (Ed.). VCH Publishing, New York, p. 3.
11. Merck. (1997). *ToxAlert™ 10 Operating Manual*.
12. Microbics Corp. (1992). *Microtox System Operating Manual*.
13. Chang, J.C., Taylor, P.B., and Leach, F.R. (1981). *Bull. Environ. Contam. Toxicol.* **26**: 150.
14. Dutka, B.J., Kwan, K.K., Rao, S.S., Jurkovic, A., and Lui, D. (1991). *Environ. Monitor. Assess.* **16**: 287.
15. Kaiser, K.L.E., Ribo, J.M., and Kwasniewska, K. (1988). *Water Pollut. Res. J. Can.* **23**: 356.
16. Steinberg, S.M., Poziomek, E.J., Engelmann, W.H., and Rogers, K.R. (1995). *Chemosphere* **30**: 2155.
17. Yates, I. and Porter, J. (1982). *Appl. Environ. Microbiol.* **44**: 107.

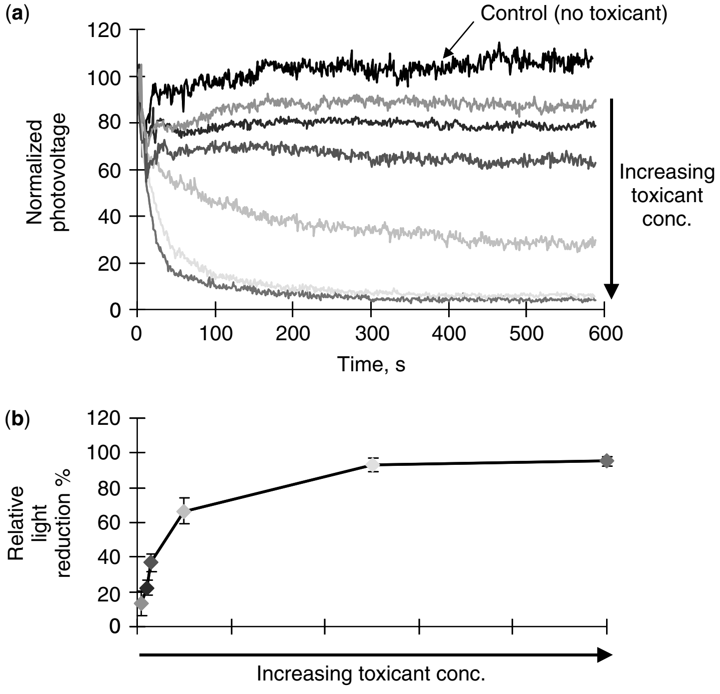

Figure 5. Sample data measuring the response of a genetically modified bioluminescent bacterium to the presence of trichlorophenol. (**a**) The output from an instrument recording light with time. (**b**) the percentage light reduction at a suitable endpoint has been measured against the nontoxic control [the black trace in (a)], and the relationship to concentration has been plotted. The shaded data points on (**b**) correspond to the shades of the decay curves on (**a**).

18. Kaiser, K.L.E. and Palabrica, V.S. (1991). *Water Pollut. Res. J. Can.* **26**: 361.
19. Bulich, A.A., Tung, K.K., and Scheibner, G. (1990). *J. Biolumin. Chemilumin.* **5**: 71.
20. Kelly, C.J., Lajoie, C.A., Layton, A.C., and Sayler, G.S. (1999). *Water Environ. Res.* **71**: 31.
21. Novaha, M., Vogel, W.R., and Gaugitsch, H. (1995). *Environ. Int.* **21**: 33.
22. Jönsson, K., Grunditz, C., Dalhammar, G., and Jansen, J.L.C. (2000). *Water Res.* **34**: 2455.
23. Blum, D.J.W. and Speece, R.E. (1991). *Res. J. WPCF* **63**: 198.
24. Gernaey, K., Petersen, B., Ottoy, J.-P., and Vanrolleghem, P. (1999). *Water Qual. Int.* June: 16.
25. Boyd, E.M., Killham, K., Wright, J., Rumford, S., Hetheridge, M., Cumming, R., and Meharg, A.A. (1997). *Chemosphere* **35**: 1967.
26. Brown, J.S., Rattray, E.A.S., Paton, G.I., Reid, G., Caffoor, I., and Killham, K. (1996). *Chemosphere* **32**: 1553.
27. Layton, A.C., Gregory, B., Schultz, T.W., and Sayler, G.S. (1999). *Ecotoxicol. Environ. Safety* **43**: 222.
28. Paton, G.I., Palmer, G., Burton, M., Rattray, E.A.S., McGrath, S.P., Glover, L.A., and Killham, K. (1997). *Lett. Appl. Microbiol.* **24**: 296.
29. Sinclair, G.M., Paton, G.I., Meharg, A.A., and Killham, K. (1999). *FEMS Microbiol. Lett.* **174**: 273.
30. Wiles, S., Whiteley, A.S., Philp, J.C., and Bailey, M.J. (2003). *J. Microbiol. Methods* **55**: 667.

BIOMANIPULATION

Masaki Sagehashi
University of Tokyo
Tokyo, Japan

Biomanipulation is a currently used lake purification method, which modifies a lake ecosystem to an ideal state. Prior to discussing biomanipulation, lake purification methods are reviewed, and the principles and current status of biomanipulation are described afterward.

CLASSIFICATION OF LAKE RESTORATION METHODS

Some measures have been designed and implemented to prevent the eutrophication of lakes. These methods can be divided into two types, the "temporary type" and the "sustainable type." The separation of artificially aggregated algae (1–4) and aeration (5–7) are temporary types of purification methods. These methods are always immediately and strongly effective, and special knowledge of the methods is not required because the mechanisms are relatively simple. These methods, however, do not sustain their effects, and whenever algal blooms appear, we should use these methods. Therefore, this type of method is not essentially a lake purification method. On the other hand, external nutrient loading reductions (8), flow control (9), dredging of sediments (10), and "biomanipulation" are the sustainable type. There are many unclear points as to the effects of these types of methods; however, they restore a lake ecosystem, and long-term stable effects can be expected in contrast to the temporary types of methods. External nutrient loading reductions are the most effective methods; however, it is difficult to prevent completely external nutrient loading due to the considerable amount of so-called nonpoint nutrient loading sources. Furthermore, internal nutrient elution from sediments is always abundant. In other words, it is difficult to prevent on algal bloom with only external nutrient loading or flow control. On the other hand, dredging of sediments is costly. That is why interest in "biomanipulation" has been growing.

PRINCIPLES OF BIOMANIPULATION

The theoretical possibility that ecosystems have more than one equilibrium has been discussed (11). Lakes make complicated ecosystems (Fig. 1) (12) and also have alternative equilibria, a clear one and a turbid one (13, 14). This idea is based on the observation that the restoration of turbid eutrophic lakes by nutrient reduction seems to be often prevented by ecological feedback mechanisms; moreover, the clear state also possesses a number of stabilizing feedback mechanisms (15). The negative effect of the vegetation of aquatic macrophytes on turbidity has been one of the reasons to expect alternative stable states in freshwater ecosystems. Scheffer (15) showed a schematic representation of the effect of vegetation on the nutrient–turbidity relationship and consequences for system equilibria if the vegetation completely disappears from a lake at a critical turbidity (Fig. 2). The equilibrium curve is also affected by the depth of the lake (Fig. 3). Scheffer summarized the relationship using his model and proposed the "marble-in-a-cup" theory (Fig. 4). The minima of the resulting curves represent stable equilibria, whereas the hilltops are unstable breakpoints. The hysteresis behavior of lake ecosystems (13) is thus derived from this relationship. At a low nutrient level, the system has one globally stable equilibrium, a clear water state. Any increase in the nutrient level gradually changes the shape of the stability landscape and gives rise to an alternative turbid equilibrium. However, no major disturbances occur, and

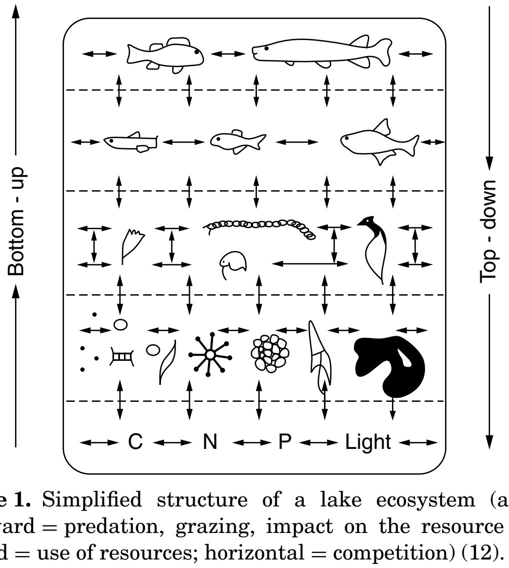

Figure 1. Simplified structure of a lake ecosystem (arrows: downward = predation, grazing, impact on the resource basis; upward = use of resources; horizontal = competition) (12).

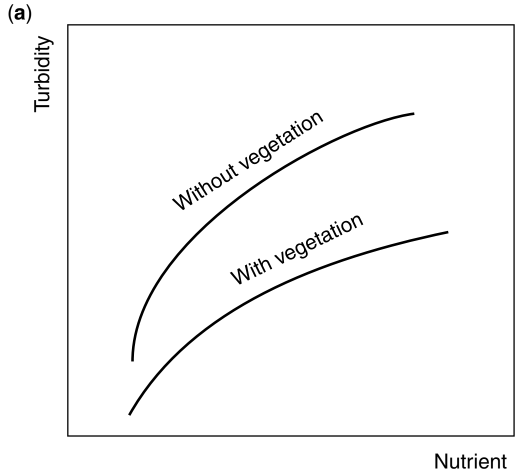

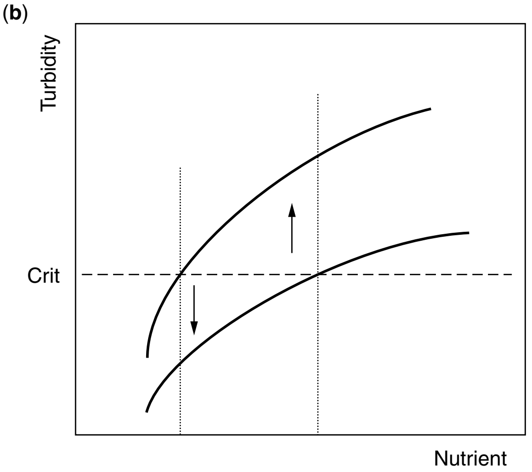

Figure 2. Schematic representation of (**a**) the effect of vegetation on the nutrient–turbidity relationship and (**b**) consequences for systems equilibria if vegetation completely disappears from a lake at a critical turbidity (crit) (15).

the system will stay in its current state, responding only slightly to enrichment. If the nutrient level is further raised, the stability of the clear state decreases, and slight perturbations are enough to cause a switch in the turbid equilibrium. At still higher nutrient levels, the clear water equilibrium disappears, inevitably causing an irreversible jump to a turbid state. Efforts to restore the system by decreasing the nutrient level change the stability landscape again, but even if the nutrient level at which the system was formerly clear is realized, there will hardly be a response to the measures. An alternative clear equilibrium may be present, but the locally stable turbid state is sustained. Only a severe reduction in nutrient level results in a switch to the clear state.

The reduction of the nutrient level as a measure to restore a lake ecosystem, which has the alternative equilibria mentioned before, gives poor results, but "biomanipulation" (16) as an additional measure can have significant effects, provided that the nutrient level has been sufficiently reduced to allow the existence of a stable alternative clear water equilibrium. A schematic diagram of the restoration of water quality by biomanipulation is shown in Fig. 5. Biomanipulation may be defined as the restructuring of a biological community to achieve a favorable response, usually a reduction in the algal biomass, attainment of clear water, and promotion of a diverse biological community (17). The theory of biomanipulation is based on the prediction that increased piscivore abundance results in decreased planktivore abundance and that increased zooplankton grazing pressure leads to reductions in phytoplankton abundance and improved water clarity (18).

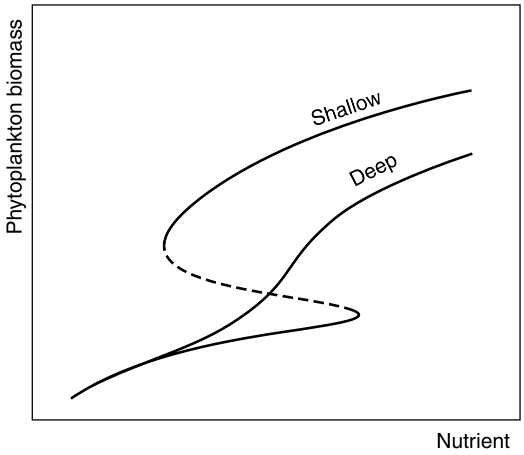

Figure 3. Equilibrium density of phytoplankton as a function of nutrient level computed from Scheffer's model for a shallow lake of homogeneous depth ("shallow") and for a deep lake with gradually declining slopes ("deep"). The dashed part represents unstable "breakpoint" equilibria (15).

REVIEW OF TRIALS OF BIOMANIPULATION AND ITS PERSPECTIVE

A number of fish manipulations in lakes and ponds have been done in various water ecosystems around the world. Some researchers have reviewed them (17,18,20). Based on these reviews, the biomanipulation trials so far attempted are summarized in Table 1. The manipulations, lake size, and outcomes are mentioned in this table. Table 1 shows that there has been a considerable spectrum of fish manipulation occurring throughout the world. Note that not every manipulation was aimed at algal growth suppression. Some of them were an unexpected change in fish biomass. In early stages, such as Hrbácek et al. described (21), little attention was given to reducing the algal biomass by manipulating the fish biomass.

These trials can be divided into several groups based on their results. The following categories are used for classifying the trials:

A: biomass of phytoplankton decreased ("successful")
B: dominated by larger zooplankton species or an increase in the biomass of zooplankton/obvious change in phytoplankton community was not observed ("insufficient")
C: no obvious changes in both the phytoplankton and zooplankton communities ("failure")
D: harmful algal dominance was observed ("risky")
E: others ("others")

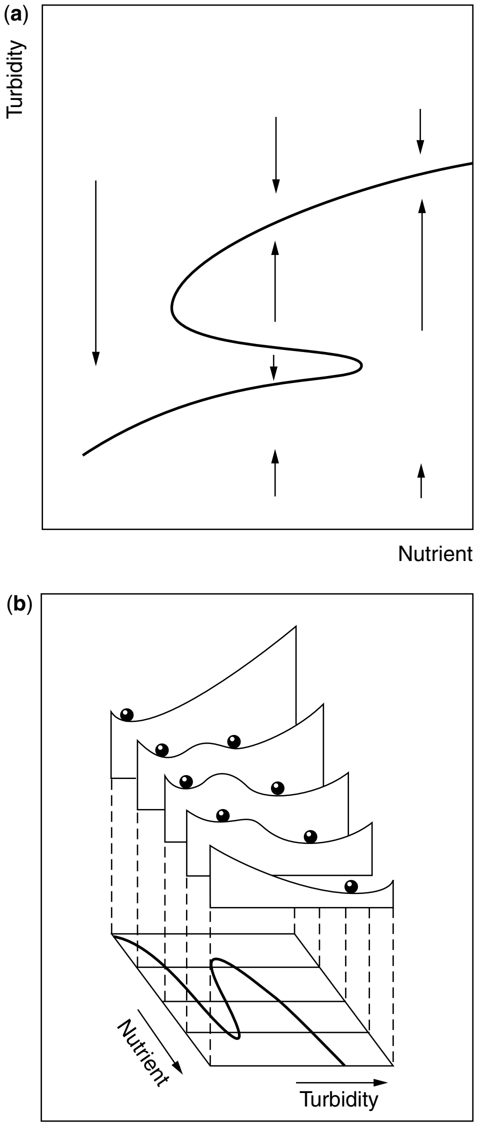

Figure 4. (**a**) Equilibrium turbidity as a function of nutrient level; direction and speed of change are indicated by arrows (example computed from vegetation–phytoplankton model). (**b**) "Marble-in-a-cup" representation of stability at five different nutrient levels. The minima correspond to stable equilibria; the tops to unstable breakpoints (15).

There are 35 trials listed in Table 1. Twenty cases of "success" (57.1%), five cases of "insufficient" (14.3%), three cases of "failure" (8.6%), four cases of "risky" (11.4%), and three cases of "others" (8.6%) are categorized. Categories A and B can be considered "positive cases"; C and D are the "negative cases." Based on this consideration, there are 71.4% positive and 20.0% negative results.

There are various manipulations in the "positive" cases. A-1, A-2, and B-3 are the cases that are naturally caused. Case A-2 was also affected by summer stratification. A-6–8, and B-4 are enclosure experiments, which examined the effect of planktivorous fish (A-6–8: bluegill, bream, roach, and crucian carp) or piscivores (B-4: walleye). Fish were killed by fish toxin (mainly rotenone) in A-4, 9, 10, 13, and 14, and B-1 and B-2. Selective catchment, which is more theoretical than poisoning, was found in cases A-12 and A-16–20. The stocking of piscivores was tried in cases A-3, 5, 11, 13, 15, 16, and 19 and B-1, 2, and 5. The piscivores stocked were pike, pikeperch, rudd, salmonine, trout, bass, and walleye. Case A-15 is one of the most perfect biomanipulation trials because the lake was emptied by pumping, and daphnids and macrophytes were stocked as well as piscivores.

Some reasons could be postulated for the "no effect" cases. In general, the removal of fish was not enough (53). This means that the calculation of sufficient biomass removal of fish to improve water quality is essential for successful biomanipulation.

The transition to an ecosystem dominated by harmful algae, which is the most undesirable result, was observed in cases D-1–4. In these cases, large size blue-green algae, which could not be predated by zooplankton, dominated the phytoplankton community. This is one of the contradictions in the effect of biomanipulation. Increased zooplankton predate small, edible phytoplankton. Large, inedible algae such as filamentous blue-green algae then occupy the phytoplankton's niche. Hosper and Meijer (48) pointed out that filamentous blue-green algae and the possible development of invertebrate predators (*Neomysis, Leptodora*) on *Daphnia* are uncertain factors for successful biomanipulation. Therefore, it should be determined whether or not inedible phytoplankton will increase in a waterbody prior to the biomanipulation.

The long-term stability of biomanipulation still needs to be investigated because these trials were performed over relatively short periods. Hosper and Meijer (48)

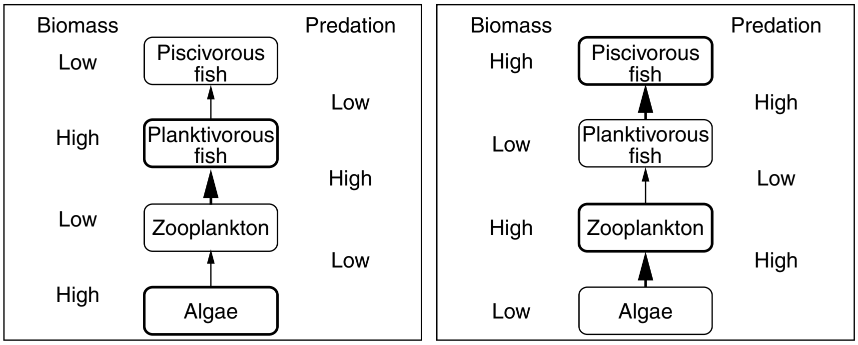

Algal-blooming state due to eutrophication

Preferable state created by biomanipulation

Figure 5. The paradigm of biomanipulation (19).

Table 1. Biomanipulation Trials Reported in the Literature

			Appearances	
Type #	Manipulation	Waterbody	Premanipulation[a]	Postmanipulation[a]
A-1	**Natural case: 1960s** decrease in *Neomysis mercedis* (planktivores) perhaps due to increase in long-fin smelt (22)	**Lake Washington**, Washington, USA	<1950s> **F:** large amount of *Neomysis* **P:** dominance of *Oscillatoria*	<1970–1980> **F:** abruptly scarce of *Neomysis* **Z:** daphnid dominance **P:** decrease in *Oscillatoria*
A-2	**Die-off (winterkill): in winter, 1964–65** DO depletion due to ice cover (23)	**Lake Severson**, Minnesota, USA area = 10.4 ha $z_{\max} = 5.3$ m; $z_{\text{ave}} = 2.2$ m	<1955–1965> **Z:** small species **P:** 300–600 mg chl.$a \cdot$m^{-3} water bloom of blue-greens **O:** summer stratification	<1965> **Z:** large species **P:** <30 mg chl.$a \cdot$m^{-3}; small greens **F:** disappearance of planktivores **M:** increased **O:** disappearance of stratification
A-3	**Stocking (piscivores): 1972** ca. 200 ind.$\cdot$ha^{-1} of trout (24)	**Hubenov Reservoir**, the Czech Republic area = 92.5 ha $z_{\max} = 31.5$ m $z_{\text{ave}} = 8.93$ m control; Klicana; Vrchlice res.	<1976> (control reservoirs) **F:** dense population: roach, perch, other small cyprinids **Z:** small cladocerans **P:** 10.5(annual);11.4 (summer) mg chl.$a \cdot$m^{-3} **S:** 2.1 (annual); 2.5 (summer) m **fit well to the Dillon–Rigler regression**	<1976> (Hubenov Reservoir) **Z:** large daphnids **P:** 2.4(annual);2.3 (summer) mg chl.$a \cdot$m^{-3} **S:** 4.9 (annual); 5.0 (summer) m **half or even less than half the value expected from the Dillon–Rigler regression**
A-4	**Poisoning (rotenone): Nov. 1973** (25,26)	**Lake Lilla Stocke-lidsvatten** Sweden area = 1 ha $z_{\max} = 8$ m;	<1973> **Z:** small cladocerans (*Bosmina*) **P:** 400 mm$^3 \cdot$m^{-2} (in August) **S:** 1.2 m (in August) **PP:** 18 mg C$\cdot$m$^{-3} \cdot$h^{-1} (average, surface)	<1975,76> **Z:** large copepods (*Eudiaptomus*) **P:** 25 mm$^3 \cdot$m^{-2} (in August) **O:** Secci: 5.8 m (in August) **PP:** 2 mg C$\cdot$m$^{-3} \cdot$h^{-1} (average, surface)
A-5	**Piscivore stocking: 1975–** An aggressive salmonine stocking program (27)	**Lake Michigan**, Wisconsin, USA area = 58,188 km^2 $z_{\max} = 285$ m $z_{\text{ave}} = 84$ m	<1975> **F:** Alewife = 110, 000 t total **Z:** calanoid copepods dominant **P:** 0.7–3.5 mg chl.$a \cdot$m^{-3} **S:** 3–8 m	<1983,84> **F:** alewife = 18, 000–20,000 t total **Z:** large daphnids dominant **P:** 0.2–2.5 mg chl.$a \cdot$m^{-3} **S:** 5–15 m
A-6	**Enclosure exp. (fish stocking): June 15–25, 1975** bluegill (0–4 ind.$\cdot$encl.$^{-1}$) (28,29)	**Enclosures in Pleasant Pond**, Minnesota, USA area = 0.25 ha $z_{\max} = 2.5$ m 1 m in diamter; 1.8 m in depth; closed bottom	[with fish] **Z:** small cladocerans **P:** increased with the biomass of fish (20—80 mg chl.$a \cdot$m^{-3})	[without fish] **Z:** large daphnids **P:** low (<5 mg chl.$a \cdot$m^{-3})
A-7	**Enclosure exp. (fish stocking): 1976** without fish and with fish (90 g m^{-2} of bream and roach) from June to October (30,31)	**Enclosures in Lake Trummen**, Sweden; enclosure: 3 m in diameter 2.2 m in depth open for sediments	[with fish] **Z:** rotifers, cyclopoid copepodites; **P:** 440 mg chl.$a \cdot$m^{-3} (in August) *Microcystis* dominant (in summer; 83 g fresh$\cdot$m^{-3}) [lake] soft-water, dense fish population (bream, roach, and pike)	[without fish] **Z:** large daphnids **P:** 20.4 mg chl.a m^{-3} (in August) a mixed population of cryptomonads, small blue-greens, and diatoms

(*continued overleaf*)

Table 1. (*Continued*)

			Appearances	
Type #	Manipulation	Waterbody	Premanipulation^{a}	Postmanipulation^{a}
A-8	**Enclosure exp. (fish stocking): 1976** without fish and with fish (220 g m^{-2} of crucian carp) June to October (30,31)	**Enclosures in Lake Bysjön**, Sweden Enclosure: 3 m in diameter 2.2 m in depth open for sediments	[with fish] **Z:** rotifers (*Keratella*) **P:** 286 mg chl.*a* m^{-3} *Microcystis, Aphanizomenon, Pediastrum, Scenedesmus, Synedra, Stephanodiscus, Melosira, Cryptomonas* [lake] hard water, low fish population (crucian carp and rainbow trout)	[without fish] **Z:** large daphnids **P:** <5 mg chl.*a* m^{-3} cryptomonads, *Aphanizomenon*
A-9	**Poisoning (rotenone): Sept. 22, 1980** 0.5 ppm; 165 kg·ha^{-1} of whitefish died in 2 days (32)	**Lake Haugatjern**, Norway area = 9.1 ha z_{ave} = 7.6 m; z_{max} = 15.5 m	<1980> **F:** whitefish (700 kg·ha^{-1}) **P:** diatoms (in June); *Anabaena* (in summer); *Staurastrum* (in autumn) summer biomass = 16 g ww/m^{3}	<1981,1982> **P:** reduced in biomass (1.5 g ww/m^{3}) growth rate: increase (five times) *Rhodomonas* (1981, 1982); *Cryptomonas* (1982)
A-10	**Poisoning (rotenone): late 1980** (33)	**Lake Haugatjern**, Norway area = 9.1 ha z_{max} = 15.5 m; z_{ave} = 7.6 m	**Z:** rotifers **P:** large-size *Anabaena flos-aquae* and *Staurastrum luetkemuelleri*	**Z:** daphnids; increased in size (1.3–1.8 mm) **P:** fast growing species, gelatinous greens, 78% decrease in biomass
A-11	**Stocking: May 1981** pikeperch of 20 specimens (2–4 kg ind.) (34,35)	**Lake Gjersjøen**, Norway area = 250 ha z_{max} = 63.5 m	<1980> **F:** juvenile roach = 12,000–15,000 fish·ha^{-1} (pelagic areas) roach dominance (>95%) (littoral areas) **P:** blue-greens (*Oscillatoria*) = 60%	<–1991> **F:** juvenile roach = 250 fish·ha^{-1} (pelagic areas) (1988–1991) perch dominance (>50%) (littoral areas) first spawning of pikeperch (1982) **P:** the biomass was reduced by half blue-greens (*Oscillatoria*) = 20%
A-12	**Removal: 1986–87** roach and bream: 160 kg·ha^{-1} (34,36–38)	**Lake Væng**, Denmark area = 15 ha z_{max} = 2 m; z_{ave} = 1.2 m	<1986> **Z:** rotifer dominance; 0.4 g DW·m^{-3} **P:** *Anabaena* and *Stephanodiscus* dominance; 5.9 g DW·m^{-3}	<1987,1988> **Z:** large cladoceran dominance; 1.3–2.7 g DW·m^{-3} **P:** cryptophytes and periodically larger species; 0.8–3.0 g DW·m^{-3}
A-13	**Poisoning (rotenone): Oct. 1987** 3 ppm rotenone **Stocking: 1988–1990** largemouth bass, walleye (39)	**Lake Christina**, Minnesota, USA area = 1619 ha z_{max} = 4.0 m; z_{ave} = 1.5 m	<1985–1987> **F:** planktivores and bentivores **Z:** small cladoceran (<100 ind L^{-1}) **P:** 40–60 mg-chl.*a*·m^{-3} (summer ave.); diatom, nonfilamentous blue-green dominance **S:** <0.3 m (midsummer)	<1988–1990> **Z:** large daphnids (>100 ind. L^{-1}) **P:** 1988,89;30–50 mg chl.*a*·m^{-3} (summer); filamentous blue-green dominance 1990; <20 mg chl.*a*·m^{-3} (summer); cryptophyceae dominance **S:** 0.3—1 m (midsummer)
A-14	**Poisoning (rotenone): Sept. 1987** 0.5 ppm all the fish (mainly whitefish: 4600 kg) were killed (40)	**Lake Mosvatn**, Norway area = 46 ha z_{max} = 3.2 m; z_{ave} = 2.1 m	<1978–1987> **Z:** rotifer dominance **P:** 23 mg chl.*a*·m^{-3} (in summer) **S:** 1.7 m (in summer)	<1988> **Z:** daphnid dominance; individually doubled in weight **P:** 7 mg chl.*a*·m^{-3} (in summer) **S:** >2.3 m (in summer)

Table 1. (*Continued*)

Type #	Manipulation	Waterbody	Appearances	
			Premanipulation^a	Postmanipulation^a
A-15	**Removal: Mar. 1987** all fish; emptied by pumping **Stocking: 1987** juvenile pike (1600); rud(140); *Daphnia*; macrophytes (37,41–44)	**Lake Zwemlust**, The Netherlands area = 1.5 ha $z_{\text{max}} = 2.5$ m; $z_{\text{ave}} = 1.5$ m	<–1986> **P:** *Microcystis* bloom; 250 mg chl.a·m^{-3} **M:** 5% of the lake bottom (6g DW m^{-2}) **F:** 800 kg·ha^{-1}; bream dominance (75%) **S:** 0.3 m	<1987–1990> **F:** rud = 395(1990), 300(1991) kg·ha^{-1} pike = 44 kg·ha^{-1} (1990–91) **P:** *Microcystis* bloom disappeared; 5–13 mg chl.a·m^{-3} **M:** 70% (87 g DW m^{-2}; 1988) and almost 100% (200 g DW m^{-2}; 1989) of the lake bottom **S:** often reached the lake bottom
A-16	**Removal: in Apr. 1987** 85% (2000 kg) of fish (breams of 1200 kg; carp of 550 kg) **Addition: May–July 1987** pike perch of 800 ind. (3.0 cm long) (37,45)	**Galgje site, Lake Bleiswijkse Zoom**, The Netherlands area = 3.1 ha $z_{\text{ave}} = 1.1$ m control: Zeeltje site area = 11.3 ha; $z_{\text{ave}} = 1.1$ m	(control area) in 1987 **F:** breams (430 kg·ha^{-1}); carp (240 kg·ha^{-1}) **Z:** 100 g·m^{-3}, large cladocera **P:** 130 mg chl.a·m^{-3}, greens and blue-greens	(treated area) in 1987 **F:** breams (45 kg·ha^{-1}), carp (59 kg·ha^{-1}) **Z:** 300 g·m^{-3}, large cladocera **P:** 10 mg·m^{-3}, greens and blue-greens
A-17	**Removal: 1988** bream: 100 kg·ha^{-1} (17% reduction) (36)	**Lake SøbygÅrd** Denmark area = 40 ha $z_{\text{max}} = 2$ m; $z_{\text{ave}} = 1.0$ m	<1984–1985> **Z:** 0.02 g DW·m^{-3}; rotifers dominance (in summer) **P:** 52 g DW·m^{-3}; *Scenedesmus* dominance	<1986–88> **Z:** 0.8–1.2 g DW·m^{-3}; cladoceran dominance (in summer) **P:** 14 g DW·m^{-3}; diverse community
A-18	**Removal: 1988–1990** major part of grass carp (46)	**Senec gravel-pit lake**, West Slovakia area = 116 ha uneven bottom $z_{\text{max}} = 6.5$–8.5 m	<1986–1988> **Z:** rotifers and Gastrortica **P:** 3.9–30.6 mg chl.a·m^{-3} (annual max.); *Microcystis* bloom (in 1987) **M:** considerably reduced	<1989–90> **Z:** rotifers (1989), copepods (1990) **P:** 10.7–16.1 mg chl.a·m^{-3} (annual max) **M:** gradually restored
A-19	**Removal: Nov. 1990–June 1991** 75% of fish (425,000 kg); mainly bream and roach **Stocking: May 1990** 600,000 young pikes (3–4 cm) (47,48)	**Lake Wolderwijd**, The Netherlands area = 2700 ha $z_{\text{max}} = 2.5$ m; $z_{\text{ave}} = 1.5$ m	<1990, 1991> **F:** 200 kg·ha^{-1}: bream(95), roach(40), ruffe(30) **Z:** *Daphnia* (in spring); hardly present (in summer) **P:** filamentous blue-green dominance 20–40 (in spring); 100–135 mg chl.a·m^{-3} **S:** 0.25–0.5 m **O:** mysid shrimp = 30 ind·m^{-2}	<1997> **F:** 45 (June); 115 (Sept.) kg·ha^{-1} mainly ruffe 0+ of cyprinids = 10–15 kg·ha^{-1} 92% of introduced pikes died from predatory fish = 5–10 kg·ha^{-1} **Z:** daphnid dominance (300 ind.·L^{-1}) (May); collapsed in July **P:** various species including filamentous blue-greens; <50 mg chl.a·m^{-3} **S:** 1.8 m max; 1.1 m average
A-20	**Removal: Nov. 1992–June 1993** fish (bream and roach) of 3582 kg (119.4 kg·ha^{-1}; 80%) (49)	**Lake Duinigermeer**, The Netherlands area = 28 ha $z_{\text{ave}} = 1$ m	<1992> **Z:** *Daphnia*; peak at the end of May (100 ind. L^{-1}) **P:** 60 mg chl.a·m^{-3} (in summer); cyanobacteria (*Oscillatoria*) **M:** disappeared **S:** 0.3–0.6 m (in summer)	<1993> **Z:** *Daphnia*; same biomass as 1992, appeared earlier, individuals were larger than in 1992 **P:** 20 mg chl.a·m^{-3} (in summer); cyanobacteria (*Oscillatoria*) **M:** covered 50% of the lake bottom **S:** 0.7 m–bottom

(*continued overleaf*)

Table 1. (*Continued*)

Type #	Manipulation	Waterbody	Appearances	
			Premanipulation[a]	Postmanipulation[a]
B-1	**Poisoning (rotenone): 1951** **Stocking (rainbow trout): every year** 15–18 cm; 370–740 ha^{-1} (50)	**Paul Lake**, Michigan, USA area = 1.21 ha $z_{\max} = 12.0$ m	<–1951> **F:** yellow perch, largemouth bass dominance **Z:** daphnids (70–90%); low biomass	<1951–> **F:** brown trout dominance **Z:** increased daphnid biomass
B-2	**Poisoning (Fishtox): 1955,1956** 68730 (907 kg) fish [346 (18.6 kg) pike] were killed **Stocking: 1957** largemouth bass (150 ind. ha^{-1}) (21)	**Poltruba backwater**, Czech Republic area = 0.18 ha $z_{\max} = 5.56$ m; $z_{\text{ave}} = 2.77$ m	<1955> **Z:** small species (*Bosmina* etc.) **P:** diatoms dominant	<1957> **Z:** large species (*Daphnia* sp.) **P:** cryptophycae dominant
B-3	**Dieoff (disease): spring, 1967** a massive dieoff of alewives (51)	**Lake Michigan**, Wisconsin, USA area = 58, 188 km^2 $z_{\max} = 285$ m; $z_{\text{ave}} = 84$ m southeastern part	<until 1966> **F:** alewives: uncommon (1954); dominant (1966) **Z:** large sp. (1954); small sp. (1966)	<1968> **Z:** large species
B-4	**Partitioning exp. (fish stocking): May–Oct. 1976** northern part: walleye fry (10,000 ind.) introduced southern part: control (28)	**Pleasant Pond** Minnesota, USA area = 0.25 ha $z_{\max} = 2.5$ m	[the southern part: control] **F:** minnow **Z:** rotifers **P:** *Chlamidomonas; Cryptomonas*; *Aphanizomenon* (two peaks); *Oscillatoria; Dactylococcopsis; Anabaenopsis; Kirchneriella*	[the northern part: fish stocked] **F:** 306 ind. (7.7 kg) of walleye (Oct.) **Z:** large cladocerans **P:** *Chlamidomonas; Cryptomonas*; *Aphanizomenon*; a variety of algal species
B-5	**Stocking: Mar. and Apr. 1981** predatory fish (perch and trout) 5150 kg (82 ind.) (52)	**A flooded quarry**, Dresden, Germany area = 0.044 ha $z_{\text{ave}} = 7.6$ m; $z_{\max} = 15.5$ m	<1979,1980> **F:** large population of *Leucaspius*, a few tehches and crucian carps (230 $\text{kg}\cdot\text{ha}^{-1}$ total) **Z:** rotifers, small cladocerans, cyclopoid copepods, Chaoborus (ca. 8–12 g $\text{ww}\cdot\text{m}^{-2}$ in summer) **P:** *Peridinium* and *Cryptomonas* (winter); various types (summer)	<1981> **F:** decreased rapidly **Z:** large daphnids (ca. 16–24 g $\text{ww}\cdot\text{m}^{-2}$ in summer) **P:** biomass was not changed drastically *Oocystis* and *Chlamydomonas* became dominant *Peridinium* vanished completely
C-1	**Removal: 1986** roach,bream, and crucian carp: 275 $\text{kg}\cdot\text{ha}^{-1}$ (78% reduction) **Stocking: 1986** perch:15 $\text{kg}\cdot\text{ha}^{-1}$ (15%) (36)	**Frederiksborg Castle Lake,** Denmark area = 21 ha $z_{\max} = 8$ m; $z_{\text{ave}} = 3.1$ m	<1986> **Z:** *Daphnia* and *Eudiaptomus* dominance; 0.8–1.4 g $\text{DW}\cdot\text{m}^{-3}$ **P:** 12–20 g $\text{DW}\cdot\text{m}^{-3}$; *Microcystis* dominance	<1987> **Z:** *Daphnia* and *Eudiaptomus* dominance; 1.2–1.3 g $\text{DW}\cdot\text{m}^{-3}$ **P:** 12–13 g $\text{DW}\cdot\text{m}^{-3}$; *Aphanizomenon* dominance **only minor changes**
C-2	**Removal: 1986–1989** 19.7 t (141 $\text{kg}\cdot\text{ha}^{-1}$) of bream and roach **Stocking: 1987–1989** asp (predatory cyprinid) (53)	**Lake Rusutjärvi**, Finland area = 140 ha $z_{\text{ave}} = 2.0$ m	<–1985> a eutrophic lake	<–1990> **Water quality has not improved**
C-3	**Removal: 1990–91** Vendace of 34 t (34)	**Lake Froylandsvatn**, Norway a shallow lake	<–1990> **F:** vendace and small trout **Z:** small biomass **P:** large biomass	<–1994> **A possible effect has not been monitored**

Table 1. (*Continued*)

			Appearances	
Type #	Manipulation	Waterbody	Premanipulation[a]	Postmanipulation[a]
D-1	**Stocking: 1977–** 20,000–80,000 of zander parrs were stocked every year pike fry and fingerlings (1985–) adult pike, eel, and catfish (1991–) **Restrictions catch: 1979–** piscivores (52,54)	**Bautzen Reservoir**, Dresden, Germany area = 533 ha $z_{\mathrm{ave}} = 7.4$ retention time = 193 d first impounded in 1974	<1974–1980> **F:** large pike and perch population (1974) (it was opened for fishing in May 1976); large stock of small perch and roach (1976) zooplankton **Z:** total 6.52 ± 2.85 (g ww·m^{-3}) daphnids: 2.30 ± 1.42 cyclopoid copepods: 1.07 ± 0.85 individual body weight: 0.011 ± 0.004 (mg ww) **P:** 7.75 ± 3.44 (g ww·m^{-3}) **S:** 1.71 ± 0.55 (m)	<1981-> **Z:** total 7.47 ± 1.57 (g ww·m^{-3}) daphnids: 4.71 ± 1.54 cyclopoid copepods: 0.63 ± 0.40 individual body weight: 0.024 ± 0.007 (mg ww) daphnid dominance **P:** 22.31 ± 17.44 (g ww·m^{-3}) lower (in 1981), higher (in 1982–) blue-green species increased **S:** 1.89 ± 0.53 (m)
D-2	**Poisoning (rotenone): Sept. 23, 1980** **Stocking (fish): Oct. 1980** largemouth bass; walleye; planktivorous bluegills channel catfish (55,56)	**Round Lake**, Minnesota, USA area = 12.6 ha $z_{\mathrm{max}} = 10.5$ m; $z_{\mathrm{ave}} = 2.9$ m	< –1980> **F:** largemouth bass (1960–1970) bluegills, and *Pomoxis nigromaculatus* (1976–1980) **Z:** small cladocerans **P:** *Cryptomonas*; various greens and filamentous blue-greens	<1981,1982> **Z:** large daphnids (1981) large daphnids -> small cladocerans (1982) **P:** *Cryptomonas*; filamentous blue-greens, filamentous blue-green algae were responsible for algal peak
D-3	**Removal: 1985–1992** selective catchment of planktivore **Stocking: 1988–1992** pikeperch (57)	**Lake Feldberger Haussee,** Germany area = 13 ha $z_{\mathrm{max}} = 12$ m; $z_{\mathrm{ave}} = 6.3$ m	< –1985> **Z:** rotifer dominance **P:** green dominance; rare filamentous blue-greens (2.2 g ww·m^{-3})	<1987– > **Z:** cladocerans **P:** filamentous blue-green dominance (21.5 g ww·m^{-3}) significant relationship between filamentous blue-green and *Eudiaptomus*
E-1	**Planktivore addition: 1955?** inadvertently introduced into an upstream of glut herring (planktivore) (58)	**Crystal Lake**, Connecticut, USA	<1942> **Z:** large species dominant	<1964> **Z:** small species dominant
E-2	**Poisoning(rotenone): June 4, 1960** 0.5 ppm rotenone was added (59)	**Fern Lake**, Washington, USA area = 21.2 acres $z_{\mathrm{max}} = 25$ feet	<until 1959> **Z:** appearance at all sights of the lake	<1960> **Z:** open water, shore edge:absent dense weed patches:resist
E-3	**Aeration: summer, 1975, 1976, 1978, 1979** to supply O_2 for low temperature hypolimnion (60)	**Tory Lake**, area = 1.23 ha $z_{\mathrm{max}} = 10$ m $z_{\mathrm{ave}} = 4.5$ m	<1973,1974> **Z:** localization at 0–2 m in depth (1974) **O:** anoxic situation was found at 0.3 m (May 1973) and 5 m (summer 1974)	<1975–1979> **F:** fat head minnow **Z:** uniform distribution at 0–8 m in depth (in 1975 and 1976) cladocerans: localized at surface; rotifers:also in hypolimnion where DO<1 ppm; small species dominance (1978 and 1979)

[a] F: fish; Z: zooplankton; P: phytoplankton; PP: primary productivity; M: macrophytes; S: Secci depth; O: others.

evaluated the long-term studies of biomanipulation in The Netherlands. They surveyed nine trials (Lakes Zwelmust, Bleiswijkse Zoom, Noorddiep, Breukeleveense plas, Klein Vogelenzang, IJzeren Man, Sondelerleijen, Wolderwijd, and Duinigermeer). Their survey indicated that six cases (Lakes Zwelmust, Bleiswijkse Zoom, Noorddiep, IJzeren Man, Wolderwijd, and Duinigermeer) showed positive results; however, only three lakes (Noorddiep, IJzeren Man, and Wolderwijd) showed long-term (>5 years) stability of clear water. Thus, long-term stability resulting from biomanipulation must be examined. In other words, the ecological sustainability of treated lakes determines their long-term stability. One of the essential methods for elucidating long-term stability is to quantify the material flow in a lake ecosystem as a means of numerical modeling. For this purpose, some numerical models that describe the lake ecosystem and predict the effect of biomanipulation are available in the literature (19,61).

BIBLIOGRAPHY

1. Bernhardt, H. et al. (1982). *Proc. Japanese-German Workshop Waste Water Sludge Treatment*, pp. 583–637.
2. Edward, J.K. and Wingler, B.J. (1990). *Aqua* **39**: 24–35.
3. Longhurst, S.J. and Graman, N.J.D. (1987). *Public Health Eng.* **14**(6): 71–76.
4. Zabel, T. (1985). *J. AWWA* **77**(5): 42–46.
5. Lindenschmidt, K.E. and Hamblin, P.F. (1997). *Water Res.* **31**(7): 1619–1628.
6. Little, J.C. (1995). *Water Res.* **29**(11): 2475–2482.
7. Soltero, R.A. et al. (1994). *Water Res.* **28**(11): 2297–2308.
8. Daley, R.D. and Pick, F.R. (1990). *Verh. Int. Verein. Limnol.* **24**: 314–318.
9. Johnson, M.G. et al. (1965). *Can. Fish. Cult.* **35**: 59–66.
10. Recknagel, F. et al. (1995). *Water Res.* **29**(7): 1767–1779.
11. May, R.M. (1977). *Nature* **269**: 471–477.
12. Benndorf, J. (1995). *Int. Rev. Gesamten Hydrobiol.* **80**(4): 519–534.
13. Hosper, S.H. (1989). *Hydrobiol. Bull.* **23**: 5–11.
14. Timms, R.M. and Moss, B. (1984). *Limnol. Oceanogr.* **29**: 472–486.
15. Scheffer, M. (1990). *Hydrobiologia* **200/201**: 475–486.
16. Shapiro, J. et al. (1975). *Proc. Symp. Water Qual. Manage. Biological Control.* University of Florida, pp. 85–96.
17. Perrow, M.R. et al. (1997). *Hydrobiologia* **342/343**: 355–365.
18. DeMelo, R. et al. (1992). *Limnol. Oceanogr.* **37**(1): 192–207.
19. Sagehashi, M. et al. (2000). *Water Res.* **34**(16): 4014–4028.
20. Shapiro, J. (1990). *Hydrobiologia* **200/201**: 13–27.
21. Hrbácek, J. et al. (1961). *Verh. Int. Verein. Limnol.* **14**: 192–195.
22. Edmondson, W.T. and Litt, A.H. (1982). *Limnol. Oceanogr.* **31**: 1022–1038.
23. Schindler, D.W. and Comita, G.W. (1972). *Arch. Hydrobiol.* **69**: 413–451.
24. Hrbácek, J. et al. (1978). *Verh. Int. Verein. Limnol.* **20**: 1624–1628.
25. Stenson, J.A.E. et al. (1978). *Verh. Int. Verein. Limnol.* **20**: 794–801.
26. Henrikson, L. et al. (1980). *Hydrobiologia* **68**: 257–263.
27. Scavia, D. et al. (1986). *Can. J. Fish. Aquat. Sci.* **43**: 435–443.
28. Lynch, M. (1979). *Limnol. Oceanogr.* **24**: 253–272.
29. Lynch, M. and Shapiro, J. (1982). *Limnol. Oceanogr.* **26**: 86–102.
30. Andersson, G. et al. (1978). *Hydrobiologia* **59**: 9–15.
31. Bjork, S. (1978). *Proc. MAB Project 5 Regional Workshop Land use Impacts on Lake and Reservoir Ecosystems*, Warsaw, May 26–June 2, 1978, pp. 196–219.
32. Reinertsen, H. and Olsen, Y. (1984). *Verh. Int. Verein. Limnol.* **22**: 649–657.
33. Langeland, A. (1990). *Hydrobiologia* **200/201**: 535–540.
34. Bratli, J.L. (1994). *Mar. Pollut. Bull.* **29**(6–12): 435–438.
35. Brabrand, A. et al. (1993). *Oecologia (Heidenberg)* **95**(1): 38–46.
36. Jeppesen, E. et al. (1990). *Hydrobiologia* **200/201**: 205–218.
37. Meijer, M.L. et al. (1994). *Hydrobiologia* **275–276**: 457–466.
38. Nielsen, S.N. (1994). *Ecol. Modelling* **73**: 13–30.
39. Hanson, M.A. and Butler, M.G. (1994). *Hydrobiologia* **279/280**: 457–466.
40. Sanni, S. et al. (1990). *Hydrobiologia* **200/201**: 263–274.
41. Janse, J.H. et al. (1995). *J. Aquat. Ecol.* **29**(1): 67–79.
42. Kornijow, R. et al. (1992). *Arch. Hydrobiol.* **123**(3): 337–347.
43. Kornijow, R. et al. (1992). *Arch. Hydrobiol.* **123**(3): 349–359.
44. Van Donk, E. et al. (1990b). *Hydrobiologia* **200/201**: 275–289.
45. Jayaweera, M. and Asaeda, T. (1996). *Ecol. Modelling* **85**: 113–127.
46. Horecka, M. et al. (1994). *Biologia, Bratislava* **49**: 141–146.
47. Meijer, M.L. et al. (1994). *Hydrobiologia* **275/276**: 31–42.
48. Hosper, S.H. and Meijer, M.L. (1999). *J. Jpn. Soc. Water Environ.* **22**(1): 18–24.
49. Van Berkum, J.A. et al. (1995). *Neth. J. Aquat. Ecol.* **29**(1): 81–90.
50. Kitchell, J.A. and Kitchell, J.F. (1980). *Limnol. Oceanogr.* **25**: 389–402.
51. Wells, L. (1970). *Limnol. Oceanogr.* **15**: 556–565.
52. Benndorf, J. et al. (1984). *Int. Rev. Gesamten Hydrobiol.* **69**: 407–428.
53. Pekkarinen, M. (1990). *Aqua Fennica* **20**(1): 13–25.
54. Benndorf, J. and Schweiz, Z. (1988). *Hydrobiologia* **49**(2): 237–248.
55. Shapiro, J. and Wright, D.I. (1984). *Freshwater Biol.* **14**: 371–383.
56. Wright, D.I. and Shapiro, J. (1984). *Verh. Int. Verein. Limnol.* **22**: 518–524.
57. Kasprzak P. et al. (1993). *Arch. Hydrobiol.* **128**(2): 149–168.
58. Brooks, J.L. and Dodson, S.I. (1965). *Science* **150**: 28–35.
59. Kiser, R.W. et al. (1963). *Trans. Am. Fish. Soc.* **92**: 17–24.
60. Taggart, C.T. (1984). *Can. J. Fish. Aquat. Sci.* **41**: 191–198.
61. Sagehashi, M. et al. (2001). *Water Res.* **35**(7): 1675–1686.

GENOMIC TECHNOLOGIES IN BIOMONITORING

Michael J. Carvan III
Matthew L. Rise
Rebecca D. Klaper
University of Wisconsin–Milwaukee
Milwaukee, Wisconsin

Biological monitoring, or biomonitoring, is defined as using measurements of the resident biota in surface waters to evaluate the condition of the aquatic environment (1). In part, biological monitoring in the United States is designed to satisfy the goals of the Clean Water Act, which aims to protect wildlife in and on the water and eliminate discharge in toxic amounts into navigable water. Biomonitoring has inherent advantages over simply measuring the presence or concentration of a particular stressor or toxin due to the fact that organisms represent an integrated response to factors in their environment over time. Biomonitoring in aquatic habitats traditionally involved measuring the diversity within the community of macroinvertebrates, fish, or periphyton using species diversity, composition with respect to functional groups, or genetic diversity (2). Community indices have inherent problems. For example, streams can include a large spatial variance that is not correlated with problems in habitat. In addition, sampling effort can affect diversity measurements, resulting in an increase in the number of species sampled with longer or more frequent sampling times (3). Identification of species is time consuming and can increase the costs associated with these types of measurements due to personnel-hours required (4).

As one alternative, measurement of the concentration of chemicals has been used as a surrogate for biological criteria. These chemical water quality criteria (TMDL or total maximum daily load) have been developed by determining concentrations of compounds that cause acute toxicity to standard test organisms. However, simply measuring the concentration of a chemical present in a system in one snapshot of time does not indicate the actual exposure of any particular organism to that chemical, neither to the organism as a whole nor to a target of action. Variation in exposure within a single species may be influenced by habitat utilization in the environment, feeding behaviors, and metabolism and physiology. To provide this information, chemical data can be used in conjunction with biomarkers, which are enzymes, proteins, or other products that can indicate an organism has been exposed to a toxin or that a toxic response has occurred. This involves sampling one or several species in a community to collect tissues, or whole organisms, for detection of the biomarker. A problem with this method in the past has been that biomarkers, for the most part, indicate a general response that can occur with exposure to several different compounds and are not specific enough to indicate a particular exposure or toxic response.

Genomic tools could potentially solve some of the problems with each of these techniques, while at the same time complementing traditional methods of monitoring. Examining the genes that are being expressed within an organism, or across several organisms, in the community provides hundreds to thousands of potential biomarkers for exposure and effects (5,6). As the number of biomarkers is several orders of magnitude greater, there is also the capability of differentiating exposures and effects to provide a more specific indicator, even from a complex mixture. Here, we provide an overview of the technologies of genomics (and other -omics) and provide a few examples of the development of their use for environmental monitoring.

HIGH THROUGHPUT TECHNOLOGIES AND THEIR APPLICATION TO BIOMONITORING

Genomics

There are a number of high throughput genomic technologies with application to environmental biomonitoring. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) is one of the most sensitive techniques for detection of low abundance mRNA. The reverse transcription process creates a complementary DNA molecule (cDNA) from an mRNA template. The PCR process amplifies the gene-specific cDNA target in a series of thermal cycles that include DNA denaturation, target-specific primer binding, and primer extension using a thermostable polymerase. Utilization of fluorescent labeling allows the quantitative detection of cDNA molecules following each round of amplification. Gene-specific fluorescent labeling of targets can be accomplished using TaqMan® (Applera Corporation) or molecular beacon (7) probes that are designed to anneal to an internal sequence of the PCR product. This approach allows multiple target sequences to be analyzed using multiple fluorescent labels in the same reaction (multiplex qRT-PCR). Intercalating agents like SYBR green are used as nonspecific probes that bind DNA and track the accumulation of PCR product. These assays are less expensive and very flexible because SYBR green will bind to any target. SYBR green can also be used to detect multiple targets if used in conjunction with melting point analysis of the PCR products (8).

Relative quantitation of the PCR products compares fluorescent intensity and the number of cycles required for the fluorescent signal to reach a fixed threshold (generally multiple standard deviations above background fluorescence) generally in the logarithmic portion of the amplification cycle. The parameter Ct (cycle threshold) is defined as the cycle number at which the fluorescence emission exceeds the fixed threshold. Ct is proportional to the number of target molecules in the sample and is used to derive mRNA abundance estimates (Fig. 1). The use of a normalizer gene, expressed at approximately the same level in the experimental and control samples used in microarray and qRT-PCR studies, corrects for technical inaccuracies (i.e., slight differences in amounts of RNA template due to pipetting error, or variability in efficiency of reverse transcriptase reactions). While there are no perfect normalizer genes, a number of transcripts are commonly reported in the literature, including glyceraldehyde-3-phosphate dehydrogenase, beta actin, major histocompatibility complex I, cyclophilin, various ribosomal proteins, and ribosomal RNA. Comparative quantitation uses Ct data to determine the relative abundance of specific transcripts in different samples. A new approach for the relative quantitation of targets in qRT-PCR (9) takes into account primer-dependent differences in the efficiency of amplification resulting in more accurate relative estimates of target abundance. Absolute quantitation of mRNA abundance in a sample requires comparison of Ct for the target gene to that of samples with known quantities of target. A standard curve can be generated using a dilution series of known targets and can then be used to estimate the number of target molecules in a particular sample. Some real-time PCR systems (e.g., Stratagene Mx3000P) include user-friendly software for performing data analyses such as conversion of Ct values to fold differences using comparative or absolute quantitation methods corrected with normalizer gene data.

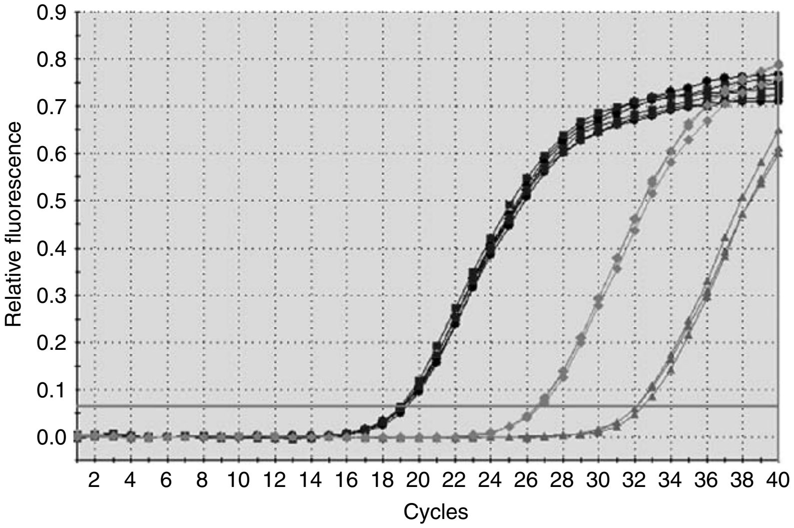

Figure 1. qRT-PCR amplification plots. Curves show amplification from template A (control, non-exposed tissue) and template B (exposed tissue). The horizontal line is the threshold. The Ct values for the amplification plots are virtually identical, indicating that the normalizer gene is expressed at approximately the same level in control and exposed tissues. Analysis of these triplicate data using Mx3000P software (Stratagene) and the relative quantification method (9) revealed that the gene of interest (GOI) was 44-fold downregulated in the exposed sample relative to the control sample.

DNA arrays (also known as Gene Chips®, DNA chips, or gene arrays) have revolutionized the way scientists think about biological problems. DNA arrays allow the simultaneous detection and quantitation of hundreds to thousands of DNA targets within a single complex mixture. The unlabeled probes for DNA arrays are printed (spotted) using high speed robots and immobilized on a solid substrate (generally nylon membranes, glass slides, or quartz wafers). Macroarrays are low density DNA arrays that are usually printed on nylon membranes or glass slides with spot diameters generally greater than 300 μm. Microarrays are high density DNA arrays that are usually printed on glass slides or quartz wafers with spot diameters generally less than 200 μm. The highest density microarrays utilize tens of thousands of oligonucleotide probes applied to, or synthesized directly on, quartz wafers. For DNA array analyses, the targets in a single sample are generally labeled with a fluorescent (or radioactive) tag and hybridized to the immobilized probes (spots on the DNA array). Hybridization of specific targets is quantified relative to standard genes on a typical one-color (or radioactive) array. Two- (or multi-) color arrays utilize multiple fluorescent labels, one for each sample to be compared, and allow the quantitation of the relative abundance of each of the specific targets in each sample. The most common experimental approach is to use the fluorescent dyes Cy3 and Cy5 to label cDNA from the control and experimental samples (or vice versa), which are then hybridized to the DNA array. This process is diagrammed in Fig. 2. Through competitive hybridization, the abundance and proportion of bound target in the two samples should reflect the abundance and proportion of corresponding transcripts in the original samples. Images of the hybridized microarray are generally acquired by laser confocal scanning using dye-specific lasers and/or filters. Quantitation of the data at each spot requires sophisticated image analysis and statistical software and the final output depicts the ratio of cDNA abundance between the two samples for each probe by scatter plot, dendrogram, heat map, or Venn diagram. The differential gene expression data from microarray experiments are usually confirmed by qRT-PCR.

Microarray analysis is an extremely sensitive technique for the analysis of mRNA abundance, and small variations in the experimental conditions may lead to dramatically different results and conclusions. As a result, the Microarray Gene Expression Data (MGED) Society was formed as an international organization of biologists, computer scientists, and data analysts that aims to facilitate the sharing of microarray data generated by functional genomics and proteomics experiments. The MGED has established a set of guidelines called the Minimum Information About a Microarray Experiment (MIAME) with the goal that such information will allow one to "interpret unambiguously and potentially reproduce and verify an array-based gene expression monitoring experiment" (www.mged.org) (10). Adherence to the MIAME guidelines and submission of all microarray images and data to a public data repository such as Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) are required by most of the top scientific journals in the United States and Europe.

Serial analysis of gene expression (SAGE) is a powerful tool that allows the analysis of overall gene expression patterns (11). SAGE can be used to identify and quantitate genes from model organisms with a sequenced genome (12) and from those with few known genes (13). The SAGE technique is based on the concept that short sequence tags (10–14 bp) can contain sufficient information to uniquely identify a transcript. Sequence tags are linked together in concatamers, which are then sequenced. Quantitation of the number of times a particular tag is observed provides the expression level of the corresponding transcript. SAGE has contributed to our understanding of numerous biological processes including differentiation, development, carcinogenesis, and responses to environmental chemicals and stressors (14).

Massively parallel signature sequencing (MPSS) technology generates sequence information from millions of DNA fragments without the need for individual sequencing reactions and the physical separation of DNA

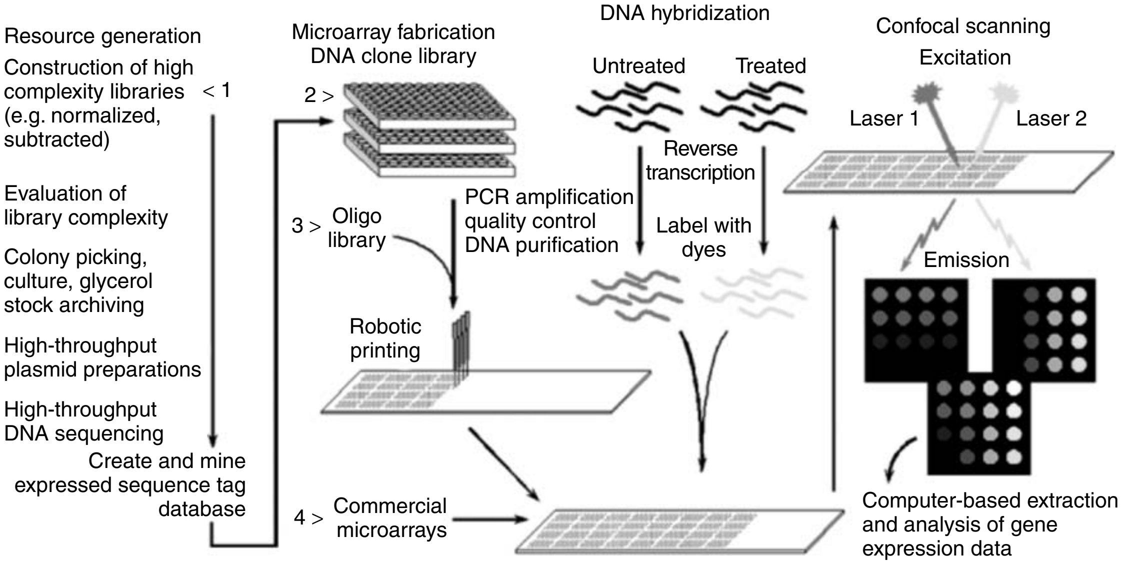

Figure 2. Microarray fabrication and utilization for transcriptional profiling. There are a number of entry points into the microarray fabrication process that are indicated by the arrows (>). The utilization of transcriptional profiling and other functional genomic technologies will require the generation of genomic resources for the vast majority of ecologically important species. This process begins, at step <1, with construction of high complexity libraries and expressed sequence tag (EST) databases. Investigations that utilize organisms for which DNA clone or oligonucleotide (oligo) libraries are available (entry points 2> and 3>, respectively) can derive fragments from these libraries for spotting on prepared glass slides. Investigations using human, mouse, *Drosophila*, yeast, or sequenced microorganisms can purchase commercial slides and enter at step 4>.

fragments (15,16). Amplified DNA molecules are bound to Megaclone™ microbeads, which are arrayed in a flow cell, and the reagents required for sequencing flow over the microbeads. The proprietary protocol elicits sequence-dependent fluorescent responses from the microbeads, which are recorded by a CCD camera after each cycle, to construct a short 17–20 bp signature sequence. The sequences are then grouped to determine abundance and compared to public and private databases to identify the gene represented by the individual signature sequences and to determine putative function of the gene product (17,18).

Proteomics

High throughput protein expression analysis is primarily performed using three techniques: two-dimensional (2D) gel electrophoresis, mass spectrometry, and protein arrays. Two-dimensional gel electrophoresis separates labeled proteins from a biological sample by their isoelectric point and electrophoretic mobility (size), producing a 2D fingerprint of protein content. Changes in protein expression can be determined by analyzing 2D gels from different samples and isolation of the elements that vary with experimental conditions. N-terminal protein sequencing can be used to identify the protein represented by a single element of a 2D gel (19). Alternatively, the isolated element can be analyzed by mass spectrometry to determine peptide mass and/or sequence, leading to its identification. The primary types of mass spectrometry used for protein expression analysis are matrix-assisted laser desorption ionization time of flight (MALDI-TOF) and Electrospray ionization MS quadrupole (ESI-MS). MALDI-TOF allows peptide fingerprinting of a protein sample following tryptic digestion. The peptide fingerprint is then compared to a database with data from known proteins or computationally derived data for theoretical peptides (20). ESI-MS provides peptide sequence information in addition to any post-translational modifications, for example, acylation, phosphorylation, or glycosylation (21).

Protein arrays are assay systems using immobilized probes on a solid support, which include glass, membranes, microtiter wells, and mass spectrometer plates. Protein arrays are used to assay a biological sample for a number of interactions including ligand binding, protein–protein, protein–DNA, protein–drug, receptor–ligand, and enzyme–substrate (22–24). Depending on the design of the array, proteins of interest can be either immobilized probes or labeled targets for binding to immobilized high affinity probes (antibodies, engineered scaffolds, single protein domains, peptides, or nucleic acids). Since this field is very new and rapidly evolving, a variety of methods are under development for utilization of protein from a number of sources, application to different formats and surfaces, immobilization of probes, fabrication of arrays, and detection of signal. Many of the techniques available today rely on adaptation of DNA array technologies and the bioinformatics tools developed for data analysis. We now know that the number of proteins produced by a given organism is, on the average, an order of magnitude greater than the number of genes in the genome. Many of the protein array technologies are focused on specific pathways and do not try to analyze the entire proteome.

Metabolomics

The foundation of metabolomics is the ability to detect changes in metabolic processes that are initiated in response to cellular stressors and enzyme system inhibitors. A number of analytical techniques can be used to generate metabolomic databases, including liquid chromatography in conjunction with mass spectrometry, high resolution nuclear magnetic resonance spectroscopy, high pressure liquid chromatography, and optical spectroscopic techniques (25–27). These techniques produce a comprehensive profile of small molecule signals. When samples derived from stressed tissues (diseased, toxicant exposed, etc.) are compared with appropriate control tissues, relatively small stresses can result in dramatic changes in metabolite profiles (28).

DEVELOPING TOOLS FOR ECOTOXICOGENOMIC RESEARCH AND BIOMONITORING

Genomics is defined as the comprehensive study of an organism's genes and their functions. Toxicogenomic research is aimed at understanding how changes in gene expression, elicited by chemical exposures, contribute to toxicological effects. DNA microarrays, important tools in toxicogenomic research, allow the relative expression of thousands of genes to be evaluated simultaneously. A primary goal of ecotoxicological research is to shed light on the mechanisms by which toxins exert their effects on various wildlife species, and ecotoxicogenomics involves the use of genomic resources and methods in ecotoxicological research.

Expressed sequence tag (EST) databases and DNA microarrays are important tools in ecotoxicogenomic research. EST databases, generally arising from high complexity (e.g., subtracted, normalized) cDNA libraries, allow researchers to identify the genes expressed in a given tissue under select conditions and at a particular point in time. For example, a suppression subtractive hybridization (SSH) cDNA library construction kit (e.g., from Clontech) could be used with mRNA from two tissues (e.g., toxin-exposed liver and control nonexposed liver) to generate reciprocal SSH libraries. One of these libraries would be enriched for genes upregulated in toxin-exposed liver relative to nonexposed liver, and the other library would be enriched for genes suppressed by the exposure. Characterization of these libraries, involving colony (clone) picking and overnight culture, plasmid DNA preparation, and automated sequencing of clone inserts (generally from one end, hence the term "tag" in EST) is generally performed in 96- or 384-well format. EST databases are designed by bioinformaticians to perform quality control such as trimming of poly A tails, low quality sequence, and vector. In addition, EST databases are generally designed to perform assemblies to identify ESTs contributing to contiguous sequences (putative transcripts). Assembled ESTs and their putative translations are aligned (BLASTN and BLASTX, respectively) with public nucleic acid or amino acid sequence databases (e.g., NCBI nt or nr, respectively) for identification. The top (most negative E-value) BLAST hits (gene names, regions of sequence overlap, percent identity/similarity, and significance of hit) and functional annotations associated with these genes (gleaned from public databases such as Gene Ontology Consortium or Swiss-Prot) may be linked to sequences within an EST database (4,29). With steadily decreasing costs associated with DNA sequencing, EST projects for ecologically important (e.g., indicator) species are now relatively affordable.

An EST project usually precedes design and construction of DNA microarrays (see Fig. 2). For model organisms (e.g., *D. rerio, X. laevis, C. elegans, Drosophila*, rat, and mouse) and some organisms of agricultural importance (e.g., rice, barley, cow, Atlantic salmon), genomic resources (e.g., EST databases and DNA microarrays) have been developed and are available. However, ecotoxicogenomic researchers generally wish to study wildlife species for which little or no genomic data previously exist. If a researcher has access to a fully equipped molecular biology laboratory, it is possible to develop genomic reagents for an ecologically important species for under \$50,000, not including personnel costs. Reciprocal SSH libraries can be constructed for about \$5000; 10,000 sequences from these high complexity libraries (likely representing approximately 5000 different transcripts) can be obtained for under \$20,000; 1000–2000 selected transcripts (e.g., relevant to xenobiotic metabolism) can be PCR amplified, cleaned, and spotted onto glass slides for under \$20,000; reagents for conducting preliminary microarray experiments (50 chips) would cost approximately \$5000. An example of such a project might be to study the effects of atrazine exposure on gene expression in a larval amphibian. Reciprocal SSH libraries could be constructed from atrazine-exposed versus nonexposed larval tissues. Transcripts of interest could be selected from the resulting EST database for inclusion on a microarray. It is now affordable for a researcher to partially characterize a nonmodel organism's transcriptome, generate genomic tools (EST database and DNA microarray), and acquire meaningful data on how the sentinel organism responds transcriptomically to a particular insult. The cost of constructing custom DNA microarrays will decrease as the number of array facilities increases. Methods for construction depend on whether the research desires a cDNA or oligonucleotide microarray.

A typical, simple toxicogenomic experimental objective might be to evaluate the impact of a single dose of toxin (e.g., atrazine) on global gene expression in a specific organ (e.g., liver) at various time points (Fig. 3). An alternative experimental approach would use microarrays to evaluate gene expression changes in a tissue associated with various toxin doses at a single time point (e.g., 24-h exposure). For all genes represented on the DNA microarrays used, expression can be compared in the exposed versus unexposed tissues. Microarray experiments are generally run on pooled samples (e.g., pooled $n = 5$ exposed liver versus

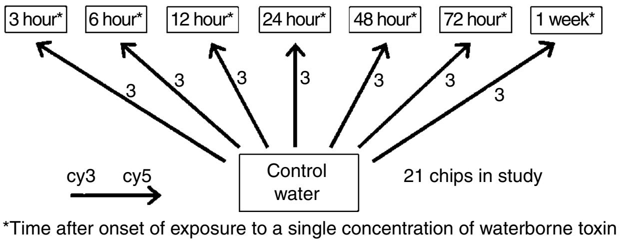

Figure 3. Microarray experimental design for determining expression of candidate molecular biomarkers of toxin exposure over time. All RNA samples are from a single toxin (e.g., atrazine) challenge at a single concentration (e.g., 10 ppb).

pooled $n = 5$ control liver), followed by real-time qRT-PCR for select transcripts of interest on individual fish samples to validate microarray results and reveal biological variability in expression of informative genes. Validated toxin-responsive genes can be functionally annotated using public sequence databases such as Swiss-Prot (http://us.expasy.org/sprot/). The example microarray experiment (Fig. 3) would be relatively easy to run, and statistical analyses of resulting data would be straightforward (e.g., K mean clustering of background-corrected, Lowess-normalized Cy5/Cy3 ratio values using software such as GeneSpring by Silicon Genetics, Fig. 4). It is important to note, however, that organisms in nature generally do not encounter a single toxin in a single dose for an extended period of time. Actual toxin exposures in the wild (e.g., exposure of larval amphibians to agricultural runoff) are continuums with fluctuating toxin concentrations and dynamic ambient conditions (e.g., temperature, pH). Furthermore, ecologically important organisms often encounter mixtures of toxins, such as effluents or automobile exhaust. Realistic modeling of this type of complexity will require extremely well-designed ecotoxicogenomic experiments supported by collaborations with mathematicians, statisticians, and computer scientists.

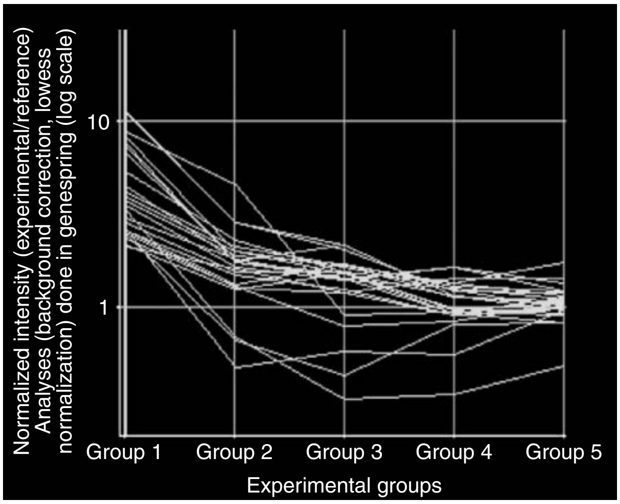

Figure 4. K means clustering identifies suites of genes having similar expression profiles.

RECENT APPLICATION OF GENOMIC TOOLS IN BIOMONITORING

Only very recently have genomic tools been used to understand environmental processes or to determine environmental quality, and we will present a few examples. Random shotgun sequencing of DNA from an environmental sample (e.g., a natural biofilm) can be used to determine species composition and provide evidence of survival strategies in these harsh environments (30). The relatively small number of dominant bacterial species in many harsh environments allows for extensive analyses without culturing organisms in the laboratory or having a previously sequenced genome to aid analyses (30).

Stressor-specific signatures in gene expression profiles can be used to diagnose which stressors are impacting populations in the field (31). Consensus PCR primers can be designed using sequence information from several animal sequences. It may then be possible to amplify some of these well-characterized stress-induced genes from organisms of interest in ecotoxicology. cDNAs representing stress-induced genes from a number of species can be arrayed on a single glass slide for rapidly screening environmental samples (31).

A few cDNA microarrays are available for fish species important in the aquaculture industry. These arrays are valuable for expression profiling following exposure to specific toxicants (32) or bacterial infection (28). While these studies are laboratory based, they demonstrate the potential for cDNA microarrays to provide insight into organism health as part of a larger biomonitoring effort. Currently available microarray resources can also be used to analyze samples from many closely related species (33,34). Thus, biomonitoring projects may be able to take advantage of genomic resources developed for model organisms in other fields and may not need to generate a species-specific microarray for every organism of interest. Targeted low density macroarrays from a few environmentally relevant species can detect alterations in gene expression following exposure to "environmental estrogens" (35,36). This technology is useful for identifying endocrine disruption; however, it may have limited utility in the analysis of the multiple stressors found in complex environments.

It is important to note that genomic information must be tied to traditional measures of toxicity to be useful. Gene expression profiling provides a picture in a snapshot of time where toxic endpoints are an accumulation of exposure and effect (5,37). To include microarray and other types of genomic data in risk assessments will involve measurements of gene expression over time and over several tissues and organisms.

BIBLIOGRAPHY

1. Gibson, G.A., Barbour, M.T., Stribling, J.B., Gerritsen, J., and Karr, J.R. (1996). *Biological Criteria: Technical Guidance for Streams and Rivers*. EPA Office of Science and Technology, Washington, DC.
2. Karr, J.R. and Chu, E.W. (1999). *Restoring. Life in Running Waters: Better Biological Monitoring*. Inland Press, Washington, DC.

3. Li, J. et al. (2001). *Freshwater Biol.* **46**(1): 87–97.
4. Resh, V.H. (1993). Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. In: *Freshwater Biomonitoring and Benthic Macroinvertebrates*. D.M. Rosenburg and V.H. Resh (Eds.). Chapman and Hall, New York, pp. 195–233.
5. Klaper, R. and Thomas, M.A. (2003). *BioScience* **54**: 403–412.
6. Thomas, M.A. and Klaper, R. (2004). *Trends Ecol. Evol.* **19**: 440–445.
7. Tyagi, S. and Kramer, F.R. (1996). *Nat. Biotechnol.* **14**(3): 303–308.
8. Siraj, A.K. et al. (2002). *Clin. Cancer Res.* **8**(12): 3832–3840.
9. Pfaffl, M.W. (2001). *Nucleic Acids Res.* **29**(9): e45.
10. Brazma, A. et al. (2001). *Nat. Genet.* **29**(4): 365–371.
11. Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. (1995). *Science* **270**(5235): 484–487.
12. Fujii, S. and Amrein, H. (2002). *EMBO J.* **21**(20): 5353–5363.
13. Lorenz, W.W. and Dean, J.F. (2002). *Tree Physiol.* **22**(5): 301–310.
14. Tuteja, R. and Tuteja, N. (2004). *Bioessays* **26**(8): 916–922.
15. Brenner, S. et al. (2000). *Nat. Biotechnol.* **18**(6): 630–634.
16. Brenner, S. et al. (2000). *Proc. Natl. Acad. Sci. U.S.A.* **97**(4): 1665–1670.
17. Brandenberger, R. et al. (2004). *BMC. Dev. Biol.* **4**(1): 10.
18. Reinartz, J. et al. (2002). *Brief. Funct. Genomic Proteomic* **1**(1): 95–104.
19. Hu, Y., Huang, X., Chen, G.Y., and Yao, S.Q. (2004). *Mol. Biotechnol.* **28**(1): 63–76.
20. Sheffield, L.G. and Gavinski, J.J. (2003). *J. Anim. Sci.* **81**(Suppl 3): 48–57.
21. Eymann, C. et al. (2004). *Proteomics* **4**(10): 2849–2876.
22. Espina, V. et al. (2004). *J. Immunol. Methods* **290**(1–2): 121–133.
23. Gerstein, M. and Echols, N. (2004). *Curr. Opin. Chem. Biol.* **8**(1): 14–19.
24. Predki, P.F. (2004). *Curr. Opin. Chem. Biol.* **8**(1): 8–13.
25. Frederich, M. et al. (2004). *Phytochemistry* **65**(13): 1993–2001.
26. Griffin, J.L. (2004). *Philos. Trans. R. Soc. London B* **359**(1446): 857–871.
27. Purohit, P.V., Rocke, D.M., Viant, M.R., and Woodruff, D.L. (2004). *OMICS.* **8**(2): 118–130.
28. Kell, D.B. (2004). *Curr. Opin. Microbiol.* **7**(3): 296–307.
29. Rise, M.L. et al. (2004). *Physiol. Genomics* **20**: 21–35.
30. Tyson, G.W. et al. (2004). *Nature* **428**(6978): 37–43.
31. Snell, T.W., Brogdon, S.E., and Morgan, M.B. (2003). *Ecotoxicology* **12**(6): 475–483.
32. Koskinen, H. et al. (2004). *Biochem. Biophys. Res. Commun.* **320**(3): 745–753.
33. Renn, S.C., Aubin-Horth, N., and Hofmann, H.A. (2004). *BMC. Genomics* **5**(1): 42.
34. Rise, M.L. et al. (2004). *Genome Res.* **14**(3): 478–490.
35. Larkin, P., Sabo-Attwood, T., Kelso, J., and Denslow, N.D. (2003). *Ecotoxicology* **12**(6): 463–468.
36. Larkin, P. et al. (2003). *EHP Toxicogenomics* **111**(1T): 29–36.
37. Neumann, N.F. and Galvez, F. (2002). *Biotechnol. Adv.* **20**(5–6): 391–419.

MACROPHYTES AS BIOMONITORS OF TRACE METALS

Shimshon Balanson
Tarun K. Mal
Cleveland State University
Cleveland, Ohio

INTRODUCTION

Aquatic macrophytes are macroscopic aquatic plants that play an important role in the health of aquatic ecosystems. Not only do macrophytes provide vital structural support in aquatic ecosystems (1), but they also contribute toward food production, nutrient regeneration, and habitat for a wide range of organisms. Because of their roles in aquatic ecosystem health, aquatic macrophytes can be utilized as indicators of stream health (2). In fact, aquatic macrophytes may be indicative of a variety of ecosystem stressors including nutrient runoff, changes in hydrologic regime, and exotic species invasion (3,4). Therefore, species richness of aquatic macrophytes, their composition, and density have all been incorporated into different macrophyte indicators (5,6). As such, bioindicators estimate the pollutants affecting the biota in an ecosystem and are purely qualitative in nature.

BIOINDICATORS VERSUS BIOMONITORS

Of course, the use of aquatic macrophytes as *qualitative* bioindicators of stream health does not reveal the amount of pollutants in the water column and sediments. Therefore, another technique must be utilized when attempting to *quantify* the direct impact of in-stream pollutants on aquatic macrophytes. The use of biota as monitors of ecosystem health relies on changes in the morphology, biochemical pathways, and life cycle of the organisms being tested as well as direct measurement of pollutant levels in the tissues (2,7,8). Biomonitors have been preferred to abiotic monitors in many instances. Abiotic monitors such as water and sediments do detail the amount of a pollutant in the ecosystem but do not reveal the quantities that are actually *available* to organisms living in an ecosystem (9).

In some ways, however, bioindicators and biomonitors share many of the same purposes. Like bioindicators, biomonitors can be utilized to identify changes in plants over the course of time, especially morphological and biochemical changes (8,10–12). Changes can be quantified at contaminated sites and patterns compared to those in plants unaffected by pollutants. Experiments can also be conducted to explore what pollutants and at what level are causing the changes in the organism. Furthermore, small discrete changes can be indicative of low contamination levels and can serve as early warning signs (12).

TRACE METAL POLLUTION IN THE AQUATIC ENVIRONMENT

Pollution is a global crisis and affects both terrestrial and aquatic environs. In aquatic ecosystems, pollution can enter through various vectors. Agricultural and industrial pollutants enter aquatic ecosystems through point sources and nonpoint sources. Atmospheric deposition, runoff, groundwater seepage, precipitation, and direct discharge are the pathways by which pollutants enter aquatic ecosystems. Trace metals are important pollutants that infiltrate aquatic ecosystems. Because many of them are essential nutrients for biota at low concentrations in certain forms and deadly at high concentrations in other forms (13), trace metals can affect aquatic ecosystems in a way akin to that of phosphorus and nitrogen.

Trace metals, in excessive quantities, have serious health ramifications for humans and other organisms. In many areas of the world, where industrial pollution is less regulated and sophisticated water treatment facilities are not in place, trace metal pollution can significantly impact drinking water and food supplies (14,15). The effects of certain trace metal species or high concentrations of essential trace metals on human health include and are not limited to: increased rates of cancer, allergic reactions, behavioral problems, lower cognitive function, lowered immunity, reduced fertility, and central nervous system dysfunction (13,16–20). Therefore, it is imperative that water resources utilized for recreation, irrigation, and drinking water supply be investigated for potential impacts of trace metal contamination on human health and ecosystem function.

AQUATIC MACROPHYTES AS BIOMONITORS

Aquatic macrophytes are prime candidates as biomonitors of trace metals in aquatic ecosystems. Their ability to accumulate trace metals many times the background levels in the environment is a significant instrument for assessing the availability of trace metals to biota (14,21–23). Aquatic macrophytes can amass trace metals by accumulation of metal ions in tissues through the roots, shoots, and leaves or through adsorption of metals to the surface of plants (21,22,24–27). The ability to bioaccumulate in roots, shoots, and leaves, termed hyperaccumulation, allows aquatic plants to be utilized as biomonitors in three distinct ways.

The tissues of aquatic macrophytes can be analyzed for trace metal concentration. Through a rather elaborate sample preparation process, with special attention to avoid introducing outside contaminants to the sample, tissue can then be tested for the concentrations of various metals generally using atomic absorption spectrophotometry (8,22,25,27–29). The direct testing of metal concentration in macrophyte tissues delivers a quantitative value regarding the accumulation potential of a given plant and can help identify the potential impact of trace metals to biota in the ecosystem (30,31). This is of primary importance when monitoring aquatic ecosystems utilized for agricultural and drinking water purposes (14). Direct testing also allows for quantification of accumulated trace metals in plant parts. This information is especially useful for plants that may be consumed. Active *in situ* experiments are preferred over passive ones by allowing temporal impacts and controlling background concentration (2).

However, that is not to say that direct tissue testing is not without its difficulties. The cost of directly testing plant tissues can be extraordinarily high and the results may not be indicative of current ecosystem conditions (32). Time-dependent accumulation and adsorption must also be taken into account (24,32). Background metal concentration also impacts uptake potential for aquatic plants. As such, *in situ* experiments must monitor sediment and water column metal concentrations and pay special attention to local environmental conditions. Finally, metal concentration is influenced by pH and other factors relating to sediment/soil composition (33).

Beyond direct testing for metal concentration in plant tissues, there exists the potential to indirectly monitor the concentration of metal through its impact on plants. Trace metals can be indirectly monitored by analyzing biochemical changes in aquatic macrophytes. Various trace metals, cadmium (Cd), lead (Pb), copper (Cu), and chromium (Cr), have been shown to reduce carbohydrate, protein, and chlorophyll production in aquatic macrophytes (see Fig. 1) (10,34). Reduction in essential nutrient uptake has also been observed in plants affected by trace metals (35), which may account for the decrease in production of other compounds in plants. Additionally, biomarkers of metal stress can indicate the presence of metal, and through research a stress-response pattern can be elucidated (36). Toxicity studies are quite similar in approach and may also help estimate levels of trace metal contamination that significantly reduce survivorship in aquatic macrophytes. Laboratory studies are far more reliable sources of information regarding indirect effects since *in situ* experiments may be confounded by synergistic effects of various compounds in the water column or other unidentifiable factors (32).

The third mechanism for testing the effects of trace metals on aquatic macrophytes is through studies of morphological changes. Changes in plant height, biomass, shoot length, and structural symmetry may all be indicative of metal stress (7,11,35,37,38). The goal in biomonitoring is to relate the degree of morphological changes in organisms to metal concentrations in the environment and available to biota. Thus, the long-term goal of utilizing morphological changes is linked to direct testing. Morphological monitoring is significantly limited,

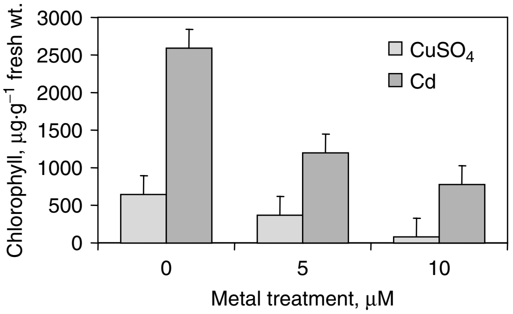

Figure 1. Changes in chlorophyll concentration as a function of trace metal treatment after eight days ($CuSO_4$, copper sulfate; Cd, cadmium). (Adapted from Refs. 10 and 34.)

much like indirect testing, to laboratory media where metal-specific responses can be controlled.

HYPERACCUMULATION POTENTIAL OF VARIOUS AQUATIC MACROPHYTES

Laboratory and *in situ* experiments on the hyperaccumulation capability of various aquatic macrophytes have yielded a wide range of results. Figure 2 exhibits the maximum observed accumulation potential of various metals by aquatic macrophytes. While aquatic macrophytes as a whole may accumulate significant levels of trace metals from the water column and sediments, no single species accumulates all metals well enough to be utilized as a single biomonitor for aquatic ecosystems (39). Figure 2 also indicates the maximum accumulation potential of various trace metals by different aquatic macrophytes. Please note that the greatest concentration of different metals has been observed in different aquatic macrophytes. However, the concentrations of metal in the tissue are not adjusted for background concentration.

While different aquatic macrophytes accumulate metals in different concentrations, it also has been found that different tissues vary in metal accumulation. In general, roots hyperaccumulate the greatest concentration of trace metals (28,30). Additional studies have shown metal concentrations in inflorescence and fruiting bodies of plants to be at startling levels especially considering their importance as food for both humans and other organisms (Fig. 3) (14,40).

Since roots are shown to be the greatest accumulators of heavy metals, it would seem obvious that rooted aquatic macrophytes would be preferred to floating macrophytes. However, this is not true in all cases. Floating macrophytes of the genus *Lemna* are utilized frequently in phytotoxicity studies (34,41). Ease of use and consistency with previous studies often influence the species of macrophytes to be used, especially for laboratory studies. *In situ* studies generally rely on the available local species, although submerged rooted species of the genus *Potamogeton* almost always have been investigated.

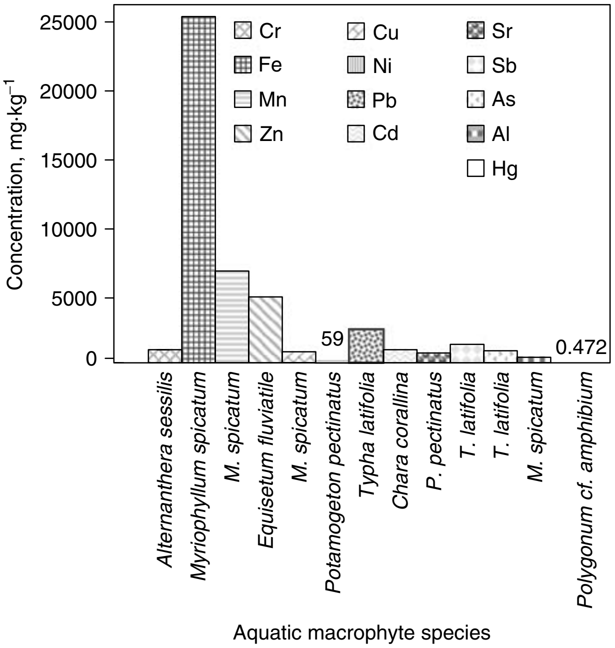

Figure 2. Maximum accumulation of trace metals by aquatic macrophytes (Cr, chromium; Fe, iron; Mn, manganese; Zn, zinc; Cu, copper; Ni, nickel; Pb, lead; Sr, strontium; Sb, antinomy; As, arsenic; Al, aluminum; Hg, mercury). (Adapted from Refs. 15, 21, 22, and 25.)

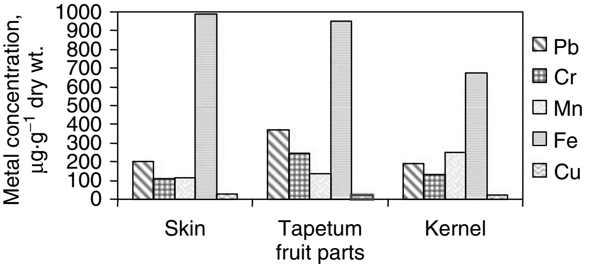

Figure 3. Accumulation of various metals in the edible plant *Trapa natans* (Pb, lead; Cr, chromium; Mn, manganese; Fe, iron; Cu, copper). (Adapted from Ref. 14.)

MACROPHYTES AS BIOMONITORS OF INDUSTRIAL POLLUTANTS

Aquatic macrophytes are already being utilized as keystone monitors of industrial pollutants in various bodies of water (15,21,22,42). For instance, aquatic macrophytes have been used to monitor pollutants in discharge of water from mines and ore processing facilities. Aquatic macrophytes, living in such highly polluted places, have adapted to extreme conditions and therefore are more likely to survive than any new species introduced as a monitor.

Aquatic macrophytes are also utilized as biomonitors of drinking water in parts of India (15). In instances where the indigenous population relies on streams and lakes for drinking water, industrial pollution can significantly harm the human population. In such cases, the concentration of metal in aquatic macrophytes can indicate to resource managers and decision-makers that something must be done to protect the residents. Direct testing of aquatic macrophytes can also indicate to farmers that irrigating crops from a specific body of water may have ramifications on human health down the line. Additionally, aquatic macrophytes may be a vital link in the foodweb dynamics of a region and significantly impact the health of animals as well as humans (22).

FUTURE RESEARCH OPPORTUNITIES

Before aquatic macrophytes can be fully employed as biomonitors of trace metals, various avenues of research must be undertaken. Direct correlations between biochemical changes and metal concentration in both plant tissues and background environments must be fully realized through intensive research. Synergistic effects of various metals should also be assessed. Studies of morphological changes due to heavy metal accumulation can be undertaken in both laboratory and *in situ* settings. Additionally, time-dependent mechanisms relating to hyperaccumulation of trace metals and their impacts on morphology and biochemical processes should be investigated. Those utilizing biomonitoring for practical purposes must know how long macrophytes retain metals in their tissues and if they are re-released into the water column in any significant way. Studies should also take into account the implication of introducing new aquatic macrophyte species as biomonitors and their potential impact on the aquatic ecosystem in general.

IMPORTANCE OF BIOMONITORING TO OTHER AREAS OF SCIENTIFIC RESEARCH

Biomonitoring of aquatic macrophytes can provide additional insight into other areas of science. The information on hyperaccumulation of heavy metals has a profound impact on the use of plants in bioremediation (8,15,42). Plants may be of vital use in extracting pollutants from a given site. Additional research on the re-release of trace metals back into the water column and sediment will have tremendous value for developing bioremediation programs as well.

Furthermore, understanding biomonitoring potential for aquatic macrophytes can also provide insight into the mechanisms and factors that influence the metal ion accumulation in aquatic plants (43). Metal accumulation by macrophytes is also a significant component of metal cycling in aquatic ecosystems (44). Thus, the scientific knowledge gained through research on aquatic macrophytes as biomonitors may have far-reaching impacts on other important realms of scientific exploration.

BIBLIOGRAPHY

1. Small, A.M. et al. (1996). A macrophyte-based rapid biosurvey of stream water quality: restoration at the watershed scale. *Restoration Ecol.* **4**: 124–145.
2. Lovett Doust, J., Schmidt, M., and Lovett Doust, L. (1994). Biological assessment of aquatic pollution: a review, with emphasis on plants as biomonitors. *Biol. Rev.* **69**: 147–186.
3. Haury, J. (1996). Assessing functional typology involving water quality, physical features and macrophytes in a Normandy river. *Hydrobiologia* **340**: 43–49.
4. King, S.A. and Buckney, R.T. (2000). Urbanization and exotic plants in northern Sydney streams. *Austral. Ecol.* **25**: 455–461.
5. Thiebaut, G., Guerold, F., and Muller, S. (2002). Are trophic and diversity indices based on macrophyte communities pertinent to monitor water quality? *Water Res.* **36**: 3602–3610.
6. Newbold, C. and Holmes, N.T.H. (1987). Nature conservation: water quality criteria and plants as water quality monitors. *Water Pollut. Control* **86**: 345–364.
7. Mal, T.K., Adorjan, P., and Corbett, A.L. (2002). Effect of copper on growth of an aquatic macrophyte, *Elodea Canadensis*. *Environ. Pollut.* **120**: 307–311.
8. Sinha, S., Saxena, R., and Singh, S. (2002). Comparative studies on accumulation of Cr from metal solution and tannery effluent under repeated metal exposure by aquatic plants: its toxic effects. *Environ. Monitor. Assess.* **80**: 17–31.
9. Phillips, D.J. (1978). Use of biological indicator organism to quantitate organochlorine pollutants in aquatic environments—a review. *Environ. Pollut.* **16**: 167–229.
10. Babu, T.S. et al. (2003). Similar stress responses are elicited by copper and ultraviolet radiation in the aquatic plant *Lemna gibba*: implication of reactive oxygen species as common signals. *Plant Cell Physiol.* **44**: 1320–1329.
11. Mal, T.K., Uveges, J.L., and Turk, K.W. (2002). Fluctuating asymmetry as an ecological indicator of heavy metal stress in *Lythrum salicaria*. *Ecol. Indicators* **1**: 189–195.
12. Lovett Doust, L., Lovett Doust, J., and Schmidt, M. (1993). Forum: in praise of plants as biomonitors—send in the clones. *Funct. Ecol.* **7**: 754–458.
13. NCEH (National Center for Environmental Health). (2002). *Toxicity of Heavy Metals and Radionuclides*. Centers for Disease Control, Atlanta. Available at http://www.cdc.gov/nceh/radiation/savannah/SRSHES_Toxicity_jan02.htm. Accessed on 10/04/2004.
14. Rai, U.N. and Sinha, S. (2001). Distribution of metals in aquatic edible plants: *Trapa natans* (Roxb.) Makino and *Ipomoea aquatica* Forsk. *Environ. Monitor. Assess.* **70**: 241–252.
15. Rai, U.N., Sinha, S., and Chandra, P. (1996). Metal biomonitoring in water resources of Eastern Ghats, Koraput (Orissa), India by aquatic plants. *Environ. Monitor. Assess.* **43**: 125–137.
16. Sta, F. et al. (1997). Changes in natural killer cell subpopulations in lead workers. *Int. Arch. Occup. Environ. Health* **69**: 306–310.
17. Tong, S. et al. (1996). Lifetime exposure to environmental lead and children's intelligence at 11–13 years: the PortPirie cohort study. *Br. Med. J.* **312**: 1569–1578.
18. Undeger, U., Basaran, N., Canpinar, H., and Kansu, E. (1996). Immune alterations in lead exposed workers. *Toxicology* **109**: 167–172.
19. Sokol, R.Z., Okuda, H., Nagler, H.M., and Berman, N. (1994). Lead exposure *in vivo* alters the fertility potential of sperm *in vitro*. *Toxicol. Appl. Pharmacol.* **124**: 310–316.
20. Assennato, G. et al. (1986). Sperm count suppression without endocrine dysfunction in lead exposed men. *Arch. Environ. Health* **41**: 387–390.
21. Samecka-Cymerman, A. and Kempers, A.J. (2004). Toxic metals in aquatic plants surviving in surface water polluted by copper mining industry. *Ecotoxicol. Environ. Safety* **59**: 64–69.
22. Hozhina, E.I., Khramov, A.A., Gerasimov, P.A., and Kumarkov, A.A. (2001). Uptake of heavy metals, arsenic and antimony by aquatic plants in the vicinity of ore mining processing industries. *J. Geochem. Explor.* **74**: 153–162.
23. Zurayk, R., Sukkariyah, B., and Baalbaki, R. (2001). Common hydrophytes as bioindicators of nickel, chromium and cadmium pollution. *Water Air Soil Pollut.* **127**: 373–388.
24. Keskinkan, O., Goksu, M.Z.L., Basibuyuk, M., and Forster, C.F. (2004). Heavy metal adsorption properties of a submerged aquatic plant (*Certaophyllum demersum*). *Bioresour. Technol.* **92**: 197–200.
25. Zakova, Z. and Kockova, E. (1999). Biomonitoring and assessment of heavy metal contamination of streams and reservoirs in the Dyje/Thaya River basin, Czech Republic. *Water Sci. Technol.* **39**: 225–232.

26. Siebert, A. et al. (1996). The use of the aquatic moss *Fontinalis anitpyretica* L. ex. Hedw. as a bioindicator for heavy metals. *Sci. Total Environ.* **177**: 137–144.
27. Lewander, M., Greger, M., Kaustky, L., and Szarek, E. (1996). Macrophytes as indicators of bioavailable Cd, Pb, and Zn flow in the river Przemsza, Katowice Region. *Appl. Geochem.* **11**: 169–173.
28. Cardwell, A.J., Hawker, D.W., and Greenway, M. (2002). Metal accumulation in aquatic macrophytes from southeast Queensland, Australia. *Chemosphere* **48**: 653–663.
29. Markett, B. (1995). Sample preparation (cleaning, drying, homogenization) for trace element analysis in plant matrices. *Sci. Total Environ.* **176**: 45–61.
30. Baldantoni, D. et al. (2004). Assessment of macro and microelement accumulation capability of two aquatic plants. *Environ. Pollut.* **130**: 149–156.
31. Stankovic, Z., Pajevic, S., Vuckovic, M., and Stojanovic, S. (2000). Concentrations of trace metals in dominant aquatic plants of the Lake Provala (Vojvodina, Yugoslavia). *Biol. Plantarium* **43**: 583–585.
32. Seaward, M.R.D. (1995). Use and abuse of heavy metal bioassays in environmental monitoring. *Sci. Total Environ.* **176**: 129–134.
33. Sparling, D.W. and Lowe, T.P. (1998). Metal concentrations in aquatic macrophytes as influenced by soil and acidification. *Water Air Soil Pollut.* **108**: 203–221.
34. Mohan, B.S. and Hosetti, B.B. (1997). Potential phytotoxicity of lead and cadmium to *Lemna minor* grown in sewage stabilization ponds. *Environ. Pollut.* **98**: 233–238.
35. Singh, R.P. et al. (1997). Response of higher plants to lead contaminated environment. *Chemosphere* **34**: 2467–2493.
36. Keltjens, W.G. and van Beusichem, M.L. (1998). Phytochelatins as biomarkers for heavy metal stress in maize (*Zea mays* L.) and wheat (*Triticum aestivum* L.): combined effects of copper and cadmium. *Plant Soil* **203**: 119–126.
37. Uveges, J.L., Corbett, A.L., and Mal, T.K. (2002). Effects of lead contamination on the growth of *Lythrum salicaria* (purple loosestrife). *Environ. Pollut.* **120**: 319–323.
38. Nedelkoska, T.V. and Doran, P.M. (2000). Characteristics of heavy metal uptake by plant species with potential for phytoremediation and phytomining. *Minerals Eng.* **13**: 549–561.
39. Schlacher-Hoenlinger, M.A. and Schlacher, T.A. (1998). Differential accumulation patterns of heavy metals among the dominant macrophytes of a Mediterranean seagrass meadow. *Chemosphere* **37**: 1511–1519.
40. Villar, C. et al. (1999). Trace metal concentrations in coastal marshes of the Lower Parana River and the Rio de la Plata Estuary. *Hydrobiologia* **397**: 187–195.
41. Blinova, I. (2004). *Use of Freshwater Algae and Duckweeds for Phytotoxicity Testing*. Wiley-Interscience, Hoboken, NJ. Available at http://www.interscience.wiley.com. Accessed 09/14/2004.
42. Karathanasis, A.D. and Johnson, C.M. (2003). Metal removal potential by three aquatic plants in an acid mine drainage wetland. *Mine Water Environ.* **22**: 22–30.
43. Nigam, K.D.P., Srivastav, R.K., Gupta, S.K., and Vasudevan, P. (1998). A mathematical model for metal ion uptake by aquatic plants for waste water treatment. *Environ. Model. Assess.* **3**: 249–258.
44. Jackson, L.J. (1998). Paradigms of metal accumulation in rooted aquatic vascular plants. *Sci. Total Environ.* **219**: 223–231.

BIOSORPTION OF TOXIC METALS

M.X. Loukidou
A.I. Zouboulis
Aristotle University of Thessaloniki
Thessaloniki, Greece

GENERAL ASPECTS

Metals may be disseminated in the environment "naturally" by geochemical and biological activities, but anthropogenic actions today produce a greater input. Several industrial activities, such as the mining, electroplating, and metal finishing industries, have intensified environmental pollution problems, as well as the discharge of heavy metals in the environment. It is well known that metals are not biodegradable and can be transformed from one chemical form to another. Growing attention is being given to the potential hazard shown by their increased concentrations in the environment because their accumulation in living tissues throughout the food chain may pose a serious human health problem. Therefore, metal-laden waste sources should be effectively treated (1).

Different factors, such as the degree of environmental risk assessment, the probable future increase in market price, and the depletion rate of reserves are usually applied to set them into different categories—high, medium, or low priority. Under public and media pressure, the European Union and the U.S. EPA have introduced and progressively enforced stricter regulations on metal discharges, particularly for effluents from industrial operations. Currently available techniques for treating metal-bearing effluents, including mainly precipitation, evaporation, adsorption, ion exchange, membrane processing, and solvent extraction, are not effective enough, or rather expensive and occasionally inadequate, considering the large wastewater quantities produced. The main performance characteristics of common heavy metal wastewater treatment technologies are summarized in Table 1 (2).

The method that consists of oxidation/reduction, chemical precipitation, and solid/liquid separation has been traditionally the most commonly applied. Depending on the volumes involved, this treatment can be operated either in batch or continuous mode. It is also essential to control carefully the addition of chemical agents, as well as the pH. One of the main problems of this method is the production of toxic sludge; it transforms an aquatic pollution problem into another of subsequent safe solid waste disposal.

As a consequence, research efforts are intensive to apply more efficient and cost-effective treatment technologies. Biosorption promises to fulfill these requirements because it is a competitive, effective, and rather cheap method.

Generally, biosorption is a treatment process that uses inexpensive biomass, usually by-product, to sequester toxic metals from contaminated effluents. Biosorbents can

Table 1. Performance Characteristics of Common Heavy Metal Removal Technologies to Treat Industrial Wastewater

Method	Disadvantages	Advantages
Chemical precipitation	• Difficult separation • Disposal of resulting toxic sludge • Not very effective	• Simple • Relatively cheap
Electrochemical treatment	• Applicable for high metal concentrations • Sensitive to specific conditions, such as the presence of certain interfering compounds	• Metal recovery
Reverse osmosis	• Application of high pressures • Membrane scaling/fouling • Expensive	• Pure effluent/permeate (available for recycling)
Ion exchange	• Sensitive to the presence of particles • Expensive resins	• Effective • Possible metal recovery
Adsorption	• Not very effective for certain metals	• Conventional sorbents

be prepared from natural abundant biomass sources, such as seaweeds, or from waste biomass of algae, fungi, or bacteria (Fig. 1) (3–6).

The unique capabilities of certain types of biomass to concentrate and immobilize particularly heavy metals can also offer some degree of selectivity. This property depends on the following parameters:

- the specific type of biomass,
- the mixing conditions of the suspension,
- the type of biomass preparation, and
- the specific physicochemical conditions.

Broad-range biosorbents can collect most of the heavy metals from an aqueous solution (wastewater), presenting rather a smaller degree of selectivity among them. A specific metal can be separated either during sorption uptake, by manipulating the properties of a biosorbent, or upon desorption during the regeneration cycle of the biosorbent.

Metal uptake by living or dead cells can consist of two different modes (5). The first, independent of metabolic activity, is referred to as *biosorption* or *passive uptake*. The second mode of metal uptake into the cell (and across the cell membrane) depends on cell metabolism and is referred to as *bioaccumulation, active uptake*, or *intracellular uptake*. Both living and dead cells can take up metals. The use of dead cells seems to be the preferred alternative for the majority of studies. The wider acceptability of dead cells is devoted to the absence of toxicity limitations, as well as of requirements for growth media and nutrient supplements in the solutions to be treated. The biosorbed metals can also be easily desorbed, creating a more concentrated solution, from which metals can be recovered for subsequent reuse by using convenient technologies, such as electrolysis; simultaneously, the regenerated biomass can also be

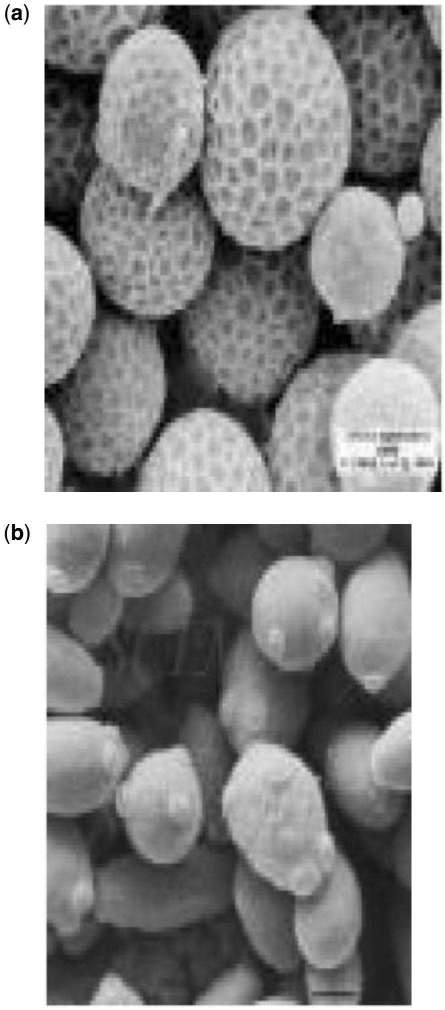

Figure 1. SEM microphotograph of **(a)** *Penicillium chrysogenum* fungi and **(b)** *Saccharomyces cerevisiae* yeast, commonly used as biosorbents.

Table 2. The Main Characteristics of Metal Biosorption and Bioaccumulation

Feature	Biosorption	Bioaccumulation
Metal affinity	High under favorable conditions	Toxicity will affect metal uptake
Rate of metal uptake	Rapid	Slower than biosorption
Selectivity	Poor	Better than biosorption
Temperature tolerance	Within a modest range	Inhibited by relatively low or high temperatures
Versatility	pH dependent, affected by anions or other molecules	Requires an energy source; dependent on plasma membrane

reused. The main characteristics of metal biosorption and bioaccumulation processes are presented in Table 2.

Biological processes to remove toxic metals can be more effectively used to treat metal solutions of low concentration (in the range of few mg/L), as a final polishing step to remove any residual metal. Therefore, the appropriate pretreatment of metal-laden wastewater containing elevated concentrations (in the range of several mg/L) is considered necessary and is usually done by common methods, such as chemical precipitation.

Although the use of living organisms is often successful for treating even highly toxic organic contaminants, living microorganisms in conventional biological treatment systems have not been found particularly useful for treating wastewater containing heavy metal ions. When the metal concentration becomes high enough, then the metabolism of the microorganism is disrupted, causing its death. The defense mechanisms of living microorganisms are responsible for the observed decrease of metal sorption. This disadvantage is not operative when nonliving microorganisms, or biological materials derived from microorganisms, are used to remove metal ions from dilute aqueous solutions.

DESORPTION

Whenever the biosorption process is used as an alternative treatment method for metal-laden wastewater, regeneration of loaded biosorbent is an important issue, to keep the process costs low and to establish the possibility of recovering the metal extracted from the liquid phase (4,7). The desorption process could result in a more concentrated metal solution, keeping the metal uptake high upon subsequent reuse of the biomass and when appropriately performed, without causing physical changes in or damaging the biosorbent. The use of dilute acids was found quite effective, and careful control of pH may also provide a simple means to increase the selectivity of metal recovery. Carbonates and/or bicarbonates were also suggested as desorptive agents; they appear to have the most commercial potential, though carbonates may sometimes affect the biomass structure.

The selection among possible metal recovery methods, subsequently applied to the resulting concentrated metal solutions, depends on the ease of metal removal from the biomass and on the commercial (or environmental) value of the metal or the biomass, as a reusable entity. When the waste biomass is relatively cheap or when it is used to accumulate valuable metals, destructive recovery by incineration or strong acid/alkali dissolution of the biomass may be considered economically feasible. The recovery method also depends on the mechanism of metal accumulation. Biosorption is often reversible, whereas metabolism-dependent intracellular accumulation is usually irreversible and, hence, necessitates the use of destructive treatment to recover metals.

MECHANISMS OF METAL BIOSORPTION

The complexity of a microorganism's structure implies that there are many different ways for the metal to be captured by a (living or dead) cell. Therefore, biosorption mechanisms can be varied, and in some cases, they are still not very well understood. They may be divided into several categories, according to the dependence on cell metabolism, metabolism-dependent and nonmetabolism-dependent, as well as on the location, where the metal will be removed from the solution, through extracellular accumulation/precipitation, cell surface sorption/precipitation, or intracellular accumulation, as shown in Fig. 2 (3).

Transport of the metal across the cell membrane yields intracellular accumulation, which depends on cell metabolism; this implies that this kind of biosorption may take place only with the viable cells. It is often associated with the active defense system of microorganisms, which react in a toxic environment (in this case, the metal).

In physicochemical interaction between a metal and the functional groups on a cell surface that takes place during physical adsorption, ion exchange, and/or surface complexation mechanisms, sorption onto a cell surface would not depend on metabolism. Cell walls of microbial biomass, composed mainly of polysaccharides, proteins, and lipids, can offer several metal-binding functional groups, such as carboxylate, hydroxyl, sulfate, phosphate, and amino groups. Biosorption of metals through physicochemical interaction occurs relatively rapidly and can be reversible. In this case, biomass has the major characteristics of an ion exchange resin, or of an activated carbon, implying certain advantages, such as lower cost, in the industrial application of biosorption.

Precipitation of metals may depend on cellular metabolism or may be independent of it. Metal removal from solution is often associated with the active defense system of microorganisms. Metal removal by precipitation, which does not depend on cellular metabolism, may be a consequence of the chemical or physical interaction

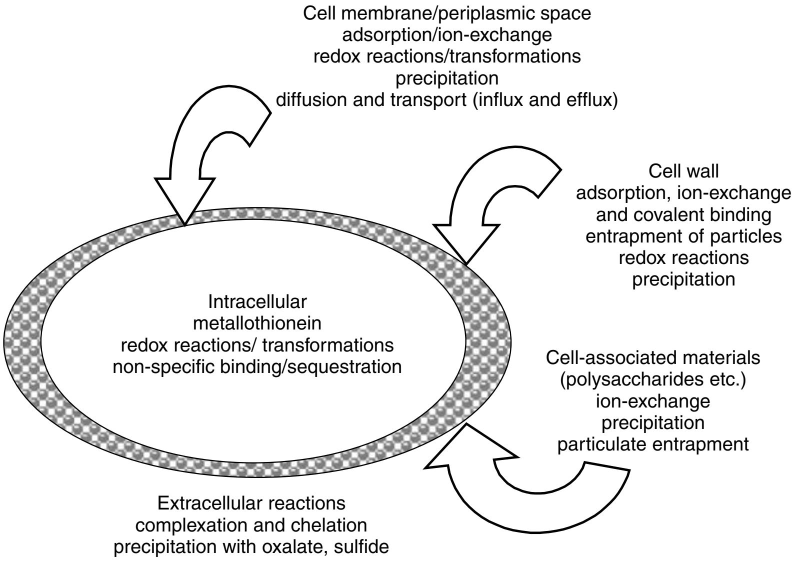

Figure 2. Several processes that contribute to microbial uptake and detoxification of aqueous solutions containing toxic metals.

between the metal and the cell surface. It becomes evident that the biosorption mechanisms are varied, and they can also take place simultaneously.

Generally, biosorption by living microorganisms consists of two basic steps. First, an independent metabolism binding step to cell walls and second, metabolism-dependent intracellular uptake, whereby metal ions are transported across a cell membrane into the cell (8). Cadmium, zinc, copper, and cobalt biosorption by the dead biomass of algae, fungi, and yeasts takes place mainly through electrostatic interactions between the metal ions in solution and cell walls. Cell walls of microorganisms contain polysaccharides as basic building blocks. The ion exchange properties of natural polysaccharides have been studied in detail, and it is well established that bivalent metal ions may exchange with the counterions of polysaccharides. Alginates of marine algae usually occur as natural salts of K^+, Na^+, Ca^{2+}, and/or Mg^{2+}. These metallic ions can be exchanged with counterions, such as Co^{2+}, Cu^{2+}, Cd^{2+}, and Zn^{2+}, resulting in biosorptive uptake of the metals. Knowledge of metal biosorption mechanisms could lead to the development of economically attractive sorbent materials.

MODELING

Mathematical modeling of biosorption offers an extremely powerful tool for a number of tasks on different levels; it is essential for process design and optimization. The biosorptive capacity and required contact time are two of the most important parameters. Sorptive equilibrium is established when the concentration of sorbate in the bulk solution is in dynamic balance with that on the interface. An equilibrium analysis is the most important information required to evaluate the affinity or capacity of a biosorbent. However, thermodynamic data can predict only the final state of a system from an initial nonequilibrium mode. It is important to determine how sorption rates depend on the concentrations of solution and how these rates are affected by sorptive capacity. From kinetic analysis, the solute uptake rate, which determines the residence time required to complete the sorption reaction, may be established (9). Equilibrium and kinetic analyses allow estimation of sorptive capacities and rates and also lead to suitable rate expressions that are characteristic of possible reaction mechanisms.

A wide range of equilibrium-based models has been used to describe the sorption of metals onto solid particles. They include the common isotherm equations, such as those developed by Langmuir and Freundlich. The Langmuir equation models monolayer sorption onto surfaces that contain a finite number of accessible sites:

$$q_{\mathrm{e}} = \frac{q_{\max} b C_{\mathrm{e}}}{1 + bC_{\mathrm{e}}} \tag{1}$$

where $q_{\max}$ is the maximum quantity of metal ions per unit weight of biomass to form a complete monolayer on the surface (mg/g) and b is a constant related to the affinity of binding sites for the metal ions (sorbate, L/mg). Note that $q_{\max}$ represents a practical limiting adsorptive capacity corresponding to the surface of a sorbent fully covered by metal ions. This quantity is particularly useful for assessing adsorption performance, especially when the sorbent does not reach its full saturation, because it enables an indirect comparison between different sorbents (10).

The empirical Freundlich equation accounts macroscopically for sorption onto heterogeneous surfaces:

$$q_e = K_{\mathrm{F}} C_{\mathrm{e}}^{1/n} \tag{2}$$

where K_{F} is an indicator of adsorptive capacity (L/g) and n is adsorptive intensity (dimensionless) (10).

Table 3 lists several types of biomass examined for their ability to remove Cd(II), Ni(II), Cu(II), and Cr(VI), which are metals considered of high environmental priority. The values of specific metal uptake q (mg of metal biosorbed per gram of biomass) are quite high in several cases

Table 3. Selected Literature Survey: Comparison of Different Biosorbents for Cd(II), Cu(II), Ni(II), and Cr(VI) Removal Capacity

Metal	Biosorbent	$q_{\max}$, mg/g	pH	T, °C	Biomass Concentration, g/L	Reference
Cd(II)	*B. licheniformis*	142.7	7	20	1	11
	B. laterosporus	159.5	1	25	1	11
	Sphaerotilus natans	115	6	25	0–1	3
	Sargassum fluitans	101.4	7	25	—	3
	Rhizopus arrhizus	26.8	7	26	3	11
	Saccharomyces cerevisiae	0.56	5	25	2	8
Cu(II)	*Chlorella vulgaris*	37.6	4	25	0.75	16
	Spirulina platensis	10	6	25	1	3
	Synechocystis sp.	23.4	4.5	25	1	3
Cr(VI)	*Chlorella vulgaris*	24	2	25	1	11
	Zooglera ramigera	3	2	25	—	3
	B. laterosporus	72.6	1	25	1	11
	Rhizopus arrhizus	62	2	25	1	13
	Rhizopus nigricans	123.45	2	25	1	13
Ni(II)	*Ascophyllum nodosum*	70	6	25	1	16
	Chlorella vulgaris	42.3	5	25	1	16
	Synechocystis sp.	15.8	5	25	1	16

and showed that selected biosorbents can be effective for removing heavy metals.

Sorption kinetics may be controlled by several independent processes that can act in series or in parallel, such as (i) bulk diffusion, (ii) external mass transfer (film diffusion), (iii) chemical reaction (chemisorption), and (iv) intraparticle diffusion (9). Any of these four steps, or any combination among these steps may be the rate-controlling factor. Transport in the solution is sometimes the rate-determining parameter in large-scale field processes. At sufficiently high agitation speed in the reaction vessel, the bulk diffusion step can be safely ignored because then sorption onto sorbent particles is decoupled from mass transfer in the bulk mixture.

IMMOBILIZATION

In the industrial application of biosorption, an immobilization procedure for biosorbent materials is often necessary to improve subsequent solid/liquid separation. The immobilization of biomass within appropriate solid structures creates a new sorptive material that has the appropriate size, mechanical strength, rigidity, and porosity necessary for use in conventional unit operations, such as fixed beds, adsorption columns, or mixing tanks. The principal techniques for the application of biosorption are based on entrapment in a strong, but rather permeable matrix, such as encapsulation within a membrane-like structure and bonding of smaller particles. It has been recognized that immobilizing biomass in a polymer matrix may improve the biomass performance and capacity, and, simultaneously, there is no need for the solid–liquid separation of a metal-laden biomass from the treated solution (11). Many polymeric matrices have been used for biomass immobilization, including polysulfone, polyurethane, alginate, polyacrylamide, and polyethyleneimine (PEI) (3). Several modifications of biomass immobilization have been used for metal removal, such as *Chlorella homosphaera* immobilized by sodium alginate; *Sargassum fluitans* and *Ascophyllum nodosum* cross-linked with formaldehyde, glutaraldehyde, or embedded in polyethylene imine; sphagnum peat moss immobilized in porous polysulfonate beads; and *Rhizopus arrhizus* immobilized in reticulated polyurethane foam (5). Modified polyacrylonitrile (denoted as PAN) was also examined as an efficient binder of fungal biomass (dead cells of *P. chrysogenum*, a waste by-product during antibiotic production) (12).

BIOSORPTION OF HEAVY METAL ANIONS

Several studies were carried out and highlighted that biosorptive removal of metal cations from aqueous solution depends mainly on the interactions between metal ions and specific groups, which are associated with the biosorbent surface. The uptake of anions by biosorption has become a growing concern, because they cannot be easily removed as insoluble hydroxides or salts by the common application of chemical precipitation.

Several chemically modified sorbents were examined for their enhanced ability to bind anions (3). The selective modification of *Rhizopus nigricans* using cetyl trimethylammonium bromide was also examined; polyethyleneimine improved the biosorptive efficiency to high levels (13). The modification of *P. chrysogenum* with common cationic surfactants (amines) or with a cationic polyelectrolyte overcame the low capacity of biomass for treating arsenic-contaminated wastewater (14). Extensive studies were carried out and reported on the biosorption of molybdate or (hexavalent) chromium anions (15). The removal of these anions can be related to the various functional groups of biomass surface, such as amino groups or proteins, which can provide positive surface charges, therefore, causing the electrostatic attraction of anions.

BIOSORPTION TREATMENT PLANTS

Following equilibrium and dynamic sorption studies, the quantitative basis for the sorption process can be established, including process performance models. Biosorption process feasibility is assessed for well-selected cases. It is necessary to realize that there are mainly two types of pilot plants, which eventually can be run side by side (Fig. 3) (4):

- biomass processing pilot plant
- biosorption pilot plant

The process of biosorption for removing metals can show performance comparable to its closest commercially used competitor, ion exchange treatment. Effluent qualities of the order of only a few ppb (μg/L) of residual metal can be achieved. Commercial ion exchange resins are rather costly, and the price tag of biosorbents can be an order of magnitude cheaper, only 10% of the ion exchange resin cost. The main attraction of biosorption seems to be the cost-effectiveness of this process. Biosorption is in its early developmental stages, and further improvements in both performance and costs can be expected in the near future.

EXISTING TECHNOLOGIES

Several biosorption processes have been developed and introduced for removing metal contaminants from surface water and groundwater. The BIOCLAIM® process employs microorganisms, principally bacteria of the genus *Bacillus*, which have been treated with strong caustic solution to enhance the accumulation of metal, washed with water to remove residual caustic, and immobilized in extruded beads using polyethyleneimine (PEI) and glutaraldehyde (3). The AlgaSORB™ process is a proprietary family of products that consists of several types of nonliving algae in different immobilization matrices. The BIO-FIX® process includes several types of biomass, including sphagnum peat moss, algae, yeast, bacteria, and/or aquatic flora, which are immobilized in polysulfone. Immobilized *Rhizopus arrhizus* has also been evaluated for recovering uranium from an ore bioleach solution. The application of flotation for the subsequent separation of metal-laden biomass, when applied in suspension, was also suggested (17).

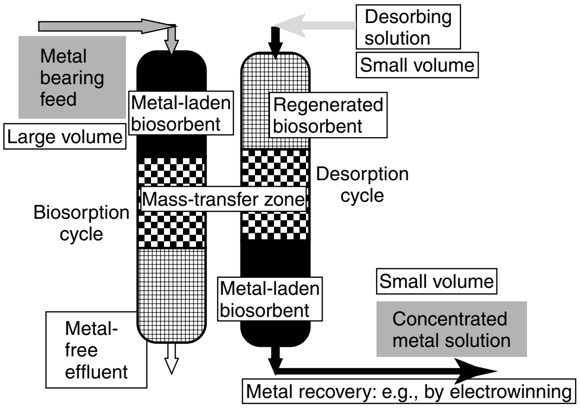

Figure 3. Schematic diagram of possible biosorption implementation, using packed bed columns for biosorption and desorption (4).

FUTURE ASPECTS

It should be evident from this brief survey that the possibilities for microbial metal removal and recovery are abundant. Biosorption is a multidisciplinary scientific area extending into several sciences, such as biology, chemistry, and engineering. The biosorption process could be used, even with a relatively low degree of understanding of its metal-binding mechanisms, but better understanding is necessary to increase the effectiveness and to optimize the relevant applications. Regarding technological development, scientific investigation should result in new families of products, as well as of biosorbents. The accompanying services should involve:

- assessing the type of problem to be remedied,
- assessing the applicability of biosorption,
- developing a customized biosorption process,
- designing and building the plant, and
- recovering the removed metals for subsequent reuse (4).

BIBLIOGRAPHY

1. Wase, J. and Forster, C. (1997). *Biosorbents for Metal Ions*. Taylor & Francis, London.
2. Eccles, H. (1999). Treatment of metal-contaminated wastes: Why select a biological process. *Trends Biotechnol.* **17**: 462–465.
3. Vegliò, F. and Beolchini, F. (1997). Removal of metals by biosorption: A review. *Hydrometallurgy* **44**: 301–316.
4. Volesky, B. (2001). Detoxification of metal-bearing effluents: Biosorption for the next century. *Hydrometallurgy* **59**: 203–216.
5. Volesky, B. (1990). *Biosorption of Heavy Metals*. CRC Press, Boca Raton, FL, pp. 84–301.
6. Kefala, M.I., Matis, K.A., and Zouboulis, A.I. (1999). Biosorption of cadmium ions by Actinomycetes and separation by flotation. *Environ. Pollut.* **104**: 283–293.
7. Gadd, G.M. (1990). Biosorption. *Chem. Ind.* **2**: 421–426.
8. Volesky, B. and Holan, Z.R. (1995). Biosorption of heavy metals—review. *Biotech. Prog.* **11**: 235–250.
9. Ho, Y.S., Ng, J.C.Y., and McKay, G. (2000). Kinetics of pollutant sorption by biosorbents: Review. *Sep. Purif. Methods* **29**: 189–232.
10. Yiacoumi, S. and Tien, C. (1995). *Kinetics of Metal Ion Adsorption from Aqueous Solutions*. Kluwer Academic, Boston.
11. Zouboulis, A.I., Loukidou, M.X., and Matis, K.A. (2004). Biosorption of toxic metals from aqueous solutions by bacteria strains isolated from metal-polluted soil. *Process Biochem.* **39**: 909–916.
12. Zouboulis, A.I., Matis, K.A., Loukidou, M., and Sebesta F. (2003). Metal biosorption by PAN-immobilized fungal biomass in simulated wastewaters. *Colloids Surf. A: Physicochem. Eng. Aspects* **212**: 185–195.

13. Bai, R.S. and Abraham, T.E. (2001). Biosorption of Cr(VI) from aqueous solution by *Rhizopus nigricans*. *Bioresource Technol.* **79**: 73–81.
14. Loukidou, M.X., Matis, K.A., Zouboulis, A.I., and Liakopoulou-Kyriakidou, M. (2003). Removal of As(V) from wastewaters by chemically modified fungal biomass. *Water Res.* **37**: 4544–4552.
15. Kratochvil, D. and Volesky, B. (1998). Advances in the biosorption of heavy metals. *Trends Biotechnol.* **16**: 291–300.
16. Donmez, G.C., Aksu, Z., Ozturk, A., and Kutsal, T. (1999). A comparative study on heavy metal biosorption characteristics of some algae. *Process Biochem.* **34**: 885–892.
17. Zouboulis, A.I., Lazaridis, N.K., and Matis, K.A. (2002). Removal of toxic metal ions from aqueous systems by biosorptive flotation. *J. Chem. Technol. Biotechnol.* **77**: 958–964.

BROMIDE INFLUENCE ON TRIHALOMETHANE AND HALOACETIC ACID FORMATION

Shankar Chellam
University of Houston
Houston, Texas

INTRODUCTION

General Background

Chemical disinfectants are added to drinking water primarily to kill or inactivate microorganisms, as discussed in other sections of this encyclopedia. However, an undesired and unintentional consequence of their use is that they react with natural organic matter (NOM) to form disinfection by-products (DBPs) (1,2). DBPs have been studied extensively since the first reports of trihalomethanes (THMs) and haloacetic acids (HAAs) in drinking water by Rook (3) and Christman et al. (4), respectively. THMs and HAAs constitute the two most important classes of DBPs measured on a weight basis in chlorinated drinking water (5). Because of length limitations, this section can only succinctly describe the important influence of the bromide ion on THM and HAA formation and speciation. The reader is directed to the references cited to obtain more detailed information. Additionally, only DBPs formed because of the application of free chlorine will be considered herein because most municipalities around the world use it for disinfection.

Simplistically, the formation of organic DBPs including THMs and HAAs can be written as follows:

$$\text{NOM} + \text{chlorine} + \text{bromide} \rightarrow \text{THMs} + \text{HAAs} + \text{other halogenated DBPs} \quad (1)$$

In the absence of the bromide ion, the only THM that can be formed is $CHCl_3$ (chloroform) whereas three HAAs, namely, chloroacetic acid (ClAA), dichloroacetic acid (Cl_2AA), and trichloroacetic acid (Cl_3AA), can be formed. However, bromide can be present in drinking water supplies due to natural geological occurrence, salt water intrusion, or human activities such as coal mine drainage or re-injection of oil field brine (6). When Br^- is present in water along with NOM, addition of chlorine can also result in formation of three additional THMs, namely, dichlorobromo methane, chlorodibromo methane ($CHClBr_2$), and bromoform ($CHBr_3$), and six additional HAAs, namely, bromo acetic acid (BrAA), dibromo acetic acid (Br_2AA), tribromo acetic acid (Br_3AA), bromochloro acetic acid (BrClAA), dichlorobromo acetic acid (Cl_2BrAA), and chlorodibromo acetic acid ($ClBr_2AA$). Hence, a total of four THMs and nine HAAs contain bromine and chlorine. As can be inferred from Equation 1, changes in the concentration of the reactants (including Br^-) will influence both the formation and relative speciation of the four THMs and nine HAAs.

Health Effects and Regulations

Epidemiological studies suggest a slightly increased risk of bladder, colon, and rectal cancers after long-term exposure to chlorinated water because of the presence of DBPs (7), warranting their designation as "primary" drinking water contaminants. However, partially because of limited data available for the occurrence, formation, and toxicology of the mixed chlorobromo acetic acids and Br_3AA, current federal regulations (8) simplistically address the sum of mass concentrations of THMs and only five HAAs (HAA5, viz, ClAA, Cl_2AA, Cl_3AA, BrAA, and Br_2AA) expressed in μg/L.

Literature reports indicate that increasing concentrations of the bromide ion cause a general shift from chlorinated toward brominated THMs (9–14), HAAs (15–18), as well as other DBPs classes (19). These studies have unequivocally established the formation of the unregulated mixed bromochloro HAA species (ClBrAA, Cl_2BrAA, and $ClBr_2AA$) and Br_3AA when chlorine is added to waters containing bromide. Importantly, recent evidence points to the increased cytotoxicity, mutagenicity, and genotoxicity of brominated acetic acids and nitromethanes compared with their chlorinated counterparts (20,21). Hence, an expressed need exists to delineate the role of bromide ion in controlling THM and HAA9 speciation, with the ultimate aim of reducing health risks associated with DBPs in drinking water.

CHEMISTRY OF BROMINATED THM AND HAA FORMATION

The relative proportion of individual THM and HAA species depends on several water chemistry factors including the concentrations of Br^- and NOM (quantified often as dissolved organic carbon, DOC), Cl_2 dosage, temperature, pH, and contact time (22). Additionally, NOM characteristics, such as polarity and molecular weight distribution, also influence THM and HAA speciation (23,24). Hence, seasonal variations in NOM characteristics, temperature, and bromide levels can be expected to induce changes in THM and HAA formation with season.

The oxidation of bromide by free chlorine (hypochlorous acid) to hypobromous acid is a very fast second-order reaction:

$$\text{HOCl} + \text{Br}^- \rightarrow \text{HOBr} + \text{Cl}^- \quad (2)$$

Note that this reaction also demonstrates that Br^- could exert a chlorine demand in natural waters. The simultaneous and parallel reactions of HOBr and HOCl with NOM results in the formation of halogenated DBPs. A higher fraction of HOBr ($pK_a = 8.8$) will remain undissociated at pH values commonly encountered in drinking water treatment than of HOCl ($pK_a = 7.5$), which allows increased formation of halogenated DBPs in bromide-containing water supplies. Starting from the early report of THM formation (3), bromine has been understood to be more reactive than chlorine in forming THMs and HAAs (13,14,16).

Trihalomethanes

Several investigations have shown that, for fixed NOM concentrations and chlorine doses, increasing bromide substantially increases total THM formation on a mass basis (9,12,13). This phenomenon is corroborated by the positive exponents for bromide ion in empirical THM models (25–29). Hence, Br^- is a THM precursor in addition to NOM. The increase in total THM levels is smaller when calculated on a molar basis because bromine has a higher atomic weight (79.9 g/mole) than does chlorine (35.5 g/mole).

Increasing Br^-/DOC ratio also induces changes in speciation under most circumstances, wherein $CHCl_3$ mole fraction monotonically decreases and $CHBr_3$ mole fraction monotonically increases. In contrast, $CHCl_2Br$ and $CHClBr_2$ mole fractions first increase and then decrease. The chlorine dose has to be increased to adjust for increasing chlorine demand for higher DOC waters in order to obtain a constant residual at the end of the predetermined incubation time when performing tests under simulated distribution system (30) or uniform formation conditions (31). These changes in Br^-/Cl_2 ratios also influence THM formation and speciation. Simultaneous changes in Br^-/DOC and Br^-/Cl_2 ratios may confound analysis of THM speciation and formation, even though the above-mentioned trends are generally observed (9,10,12–14). In general, high pH favors the formation of THMs (2,32) possibly because they are the end products of hydrolysis of several DBPs (33).

The degree of bromine substitution in THMs can be quantified for the bromine incorporation factor n (34):

$$n = \frac{\sum_{k=0}^{3} k \times [\mathrm{CHCl}_{3-k}\mathrm{Br}_k]}{\sum_{k=0}^{3} [\mathrm{CHCl}_{3-k}\mathrm{Br}_k]} \tag{3}$$

where each THM species is expressed in μM. If $CHCl_3$ is exclusively formed $n = 0$, whereas $n = 3$ corresponds to a situation when only $CHBr_3$ is formed. In most water treatment scenarios, $0 < n < 3$, which corresponds to the simultaneous formation of all or several THM species. As can be anticipated, increasing Br^-, decreasing DOC, and decreasing Cl_2 dose generally result in increasing THM bromine incorporation factor (9,10,12–14).

Haloacetic Acids

In contrast to the numerous THM studies available in the literature, investigations on the role of bromide ion on HAA9 formation and speciation had been limited until relatively recently by the nonavailability of commercial standards for the mixed chlorobromo and tribromo species. Early HAA9 studies employing instrumental standards for several mixed chlorobromo species synthesized in academic laboratories revealed a gradual shift from the chlorinated to the mixed bromochloro and then to the brominated species with increasing Br^- concentration (15,16). However, unlike the THMs, the molar yield of HAA9 does not appear to strongly depend on Br^- (15,16), which indicates that Br^- may not necessarily be a precursor for HAAs. Studies that only monitored HAA5 or HAA6 (HAA5 and BrClAA) have indicated a decrease in the regulated HAAs with increasing Br^- (10,35), which should be interpreted with caution because shifts toward the unmonitored and more brominated species would cause this phenomenon rather than a decrease in the total molar concentration of all nine HAAs. These patterns are consistent with the negative exponent for Br^- in empirical power law functions derived to model HAA6 formation (26) and purely chlorinated HAAs (27) and positive exponent for BrAA and Br_2AA (27).

Monohalogenated HAAs (ClAA and BrAA) are typically formed at very low levels after chlorination of drinking water supplies, whereas the di- (Cl_2AA, ClBrAA, Br_2AA) and trihalogenated (Cl_3AA, Br_3AA, Cl_2BrAA, and $ClBr_2AA$) HAAs are the dominant species. Their relative dominance may depend on specific ultraviolet absorbance at 254 nm and pH (16,36). As a general rule, HAAs are preferentially formed at low pH values partially because trihalogenated HAAs undergo base catalyzed hydrolysis at high pH values (33).

Equation 3 can be extended to define a HAA9 bromine incorporation factor n':

$$n' = \frac{\begin{array}{c}[\mathrm{BrAA}] + [\mathrm{BrClAA}] + [\mathrm{BrCl_2AA}] \\ +2 \times [\mathrm{Br_2ClAA}] + 2 \times [\mathrm{Br_2AA}] + 3 \times [\mathrm{Br_3AA}]\end{array}}{\begin{array}{c}[\mathrm{ClAA}] + [\mathrm{Cl_2AA}] + [\mathrm{Cl_3AA}] + [\mathrm{BrAA}] + [\mathrm{BrClAA}] \\ +[\mathrm{BrCl_2AA}] + [\mathrm{Br_2ClAA}] + [\mathrm{Br_2AA}] + [\mathrm{Br_3AA}]\end{array}} \tag{4}$$

Similar to the THM bromine incorporation factor, $0 \leq n' \leq 3$ and n' increases with the degree of bromine substitution.

EFFECTS OF TREATMENT PROCESSES ON SPECIATION

During drinking water purification, changes in the relative amounts of Br^- and DOC are most often brought about by decreases in NOM rather than by increases in Br^-. In other words, the Br^-/DOC ratio is increased by decreasing the denominator rather than by increasing the numerator, as in most previous investigations of bromide effects on THM and HAA formation and speciation (9,12,13,15,16,35). This result is because processes such as coagulation, precipitative softening, nanofiltration (NF), biological filtration, advanced oxidation, and granular activated carbon (GAC) treatment are capable of DBP precursor (or NOM) removal, whereas Br^- passes almost conservatively through them (37). These changes in Br^-/DOC inherent to drinking water treatment processes will induce changes in THM and HAA formation and speciation as discussed next. Because of length limitations, only NF and GAC will be considered herein. Additional information on other treatment processes can be found elsewhere (1,33,38).

Effects of Granular Activated Carbon Treatment

Granular activated carbon treatment has been designated as one of the best available technologies for DBP precursor removal by the U.S. Environmental Protection Agency and is widely employed for NOM removal. Several excellent reviews of GAC use in drinking water treatment are available including its use for DBP control (39,40). However, changes in DBP speciation have not been emphasized in these reviews. Hence, this section focuses primarily on speciation issues and the reader is directed elsewhere for a more holistic understanding of GAC use for DBP control.

Figure 1 depicts several aspects of THM and HAA precursor removal and changes in speciation after GAC treatment for a pretreated surface water. The breakthrough curves corresponding to a gradual increase in effluent DOC and UV_{254} values are shown in Fig. 1a. Consequently, the effluent Br^-/DOC ratio decreases (Fig. 1e) because GAC treatment has very little effect on bromide ion. Figure 1e also depicts decreasing Br^-/Cl_2 decrease with progressive breakthrough of NOM, which is caused by increasing disinfectant demand, which necessitates higher Cl_2 doses to maintain a ~1 mg/L residual at the end of the 24 h incubation time (SDS testing conditions). These changes result in an increase in the formation of brominated THMs and HAAs in the effluent with operating time (Fig. 1b and 1c). Note that after ~3200 bed volumes, $CHBr_3$ concentrations were greater in the effluent than in the influent. Similar changes in THM speciation have been discussed earlier (41,42). Importantly, it should be noted that even though each individual THM and HAA increased on mass basis in the effluent (Fig. 1b and 1c), only the mole fraction of chloroform increased substantially (0.10 to 0.27). This result was accompanied by a significant decrease in $CHBr_3$ mole fraction (0.33 to 0.13). Mole fractions of $CHCl_2Br$ remained approximately constant, whereas that of $CHClBr_2$ increased slightly from 0.18 to 0.23. Of the unregulated HAAs, two mixed chlorobromo species (ClBrAA and Cl_2BrAA) were detected in both the GAC influent and effluent, whereas Br_3AA and $ClBr_2AA$ were never formed (probably because of the relatively low influent levels of $Br^- = 72$ μg/L). In the feed water ($Br^-/DOC = 5.8$ μM/mM), the speciation of the detectable HAAs was in the order: $Cl_2AA > BrClAA > Cl_3AA \sim Br_2AA > Cl_2BrAA$. After 4800 bed volumes ($Br^-/DOC = 13.4$ μM/mM), the speciation changed to $Cl_2AA \sim Br_2AA > BrClAA > Cl_2BrAA \sim Cl_3AA$ because the DOC was reduced to 0.8 mg/L with Br^- remaining at 72 μg/L. This change in Br^-/DOC ratio increased the mole fraction of Br_2AA from 0.13 in the influent water to 0.21 in the effluent. It should be noted that different speciation patterns should be expected if Br^- levels in the influent were higher.

Decreasing Br^-/DOC with GAC operation (increasing bed volumes) reduced both THM and HAA bromine incorporation factors as shown in Fig. 1(d). Another method of analyzing halogen incorporation is to calculate DBP-X/DBP, which refers to the molar concentration of Cl or Br incorporated in THMs or HAAs divided by

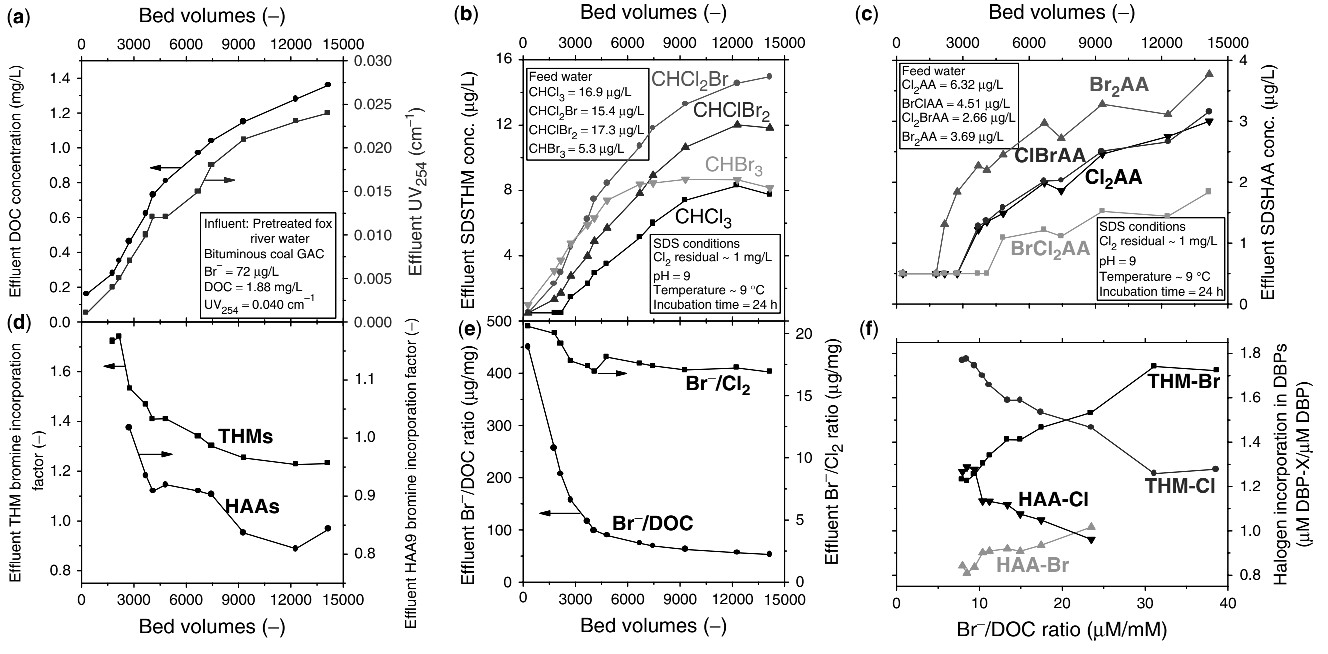

Figure 1. Several aspects of DBP control using GAC. (**a**) DOC and UV_{254} breakthrough curves; (**b**) and (**c**) formation of individual THMs and HAAs in the effluent, respectively; decreasing THM and HAA bromine incorporation factors (**d**) are caused by decreases in Br^-/DOC and Br^-/Cl_2 ratios (**e**). Increasing Br incorporation and decreasing Cl incorporation in GAC effluent for both THMs and HAAs with Br^-/DOC is shown in (**f**).

the total molar concentration of THM or HAA (Fig. 1f). Bromine incorporation into THMs and HAAs formed in GAC effluent was the highest at early times (high Br^-/DOC ratio) and decreased progressively with bed operation. As may be anticipated, the opposite trend was observed for chlorine incorporation. Bromine and chlorine incorporation in both THMs and HAAs were found to be equal at Br^-/DOC of ~20 μM/mM, corresponding to a Br^-/Cl_2 molar ratio of 15.5, which confirms that HOBr is more reactive than HOCl is in forming THMs and HAAs.

Effect of Nanofiltration

The early work by Taylor and co-workers simply documented that NF membranes were capable of very high NOM removal and capable of allowing municipalities to easily meet or exceed federal DBP regulations even if chlorine is employed for both primary and secondary disinfection (43). As demonstrated in more recent investigations (44–46), DBP precursor rejection decreases with increasing feed water recovery (Fig. 2a), which indicates that diffusion is the controlling mechanism of NOM and DBP precursor removal by NF membranes. Hence, increasing recovery reduces the Br^-/DOC ratio (Fig. 2f), which consequently increases the formation of the more chlorinated THMs ($CHCl_3$ and $CHCl_2Br$) while decreasing the more brominated ones ($CHClBr_2$ and $CHBr_3$) as depicted in Fig. 2c. Importantly, decreasing HAA9 and THM4 formation as recovery is lowered (increasing Br^-/DOC) indicates that NF waters are precursor limited.

Because NF achieved >90% removal of DOC (Fig. 2a) while allowing the almost complete passage of Br^-, a substantial increase in more brominated THM and HAA species was observed in the permeate compared with the feed water at a fixed operating recovery (70%) (Fig. 2b and 2e). Such speciation changes also decrease n and n' values with increasing recovery (Fig. 2d). Note that the mixed chlorobromo species were all detected in the permeate water. Similar to Fig. 1e, the Br^-/Cl_2 ratio also decreased with increasing DOC because greater Cl_2 doses were necessary to maintain a ~2 mg/L residual at the end of the 24 h holding time. Formation and speciation trends in Fig. 1 are qualitatively similar to those reported earlier (44,47).

Importantly, chlorination conditions (pH, temperature, incubation time, and dose) have been shown to influence THM and HAA speciation more than the membrane material and molecular weight cutoff or source water location (35,47).

Figure 3 depicts individual THM species formed in the permeate water normalized by the corresponding value in the feed water for three different NF membranes. Under the SDS conditions employed, NF preferentially reduced $CHCl_3$ and $CHCl_2Br$ concentrations, whereas NF increased concentrations of the more brominated species ($CHClBr_2$ and $CHBr_3$). Cumulatively, even though ~90% of THM precursors were removed by the same 400 Da molecular weight cutoff membrane shown in Fig. 1, it

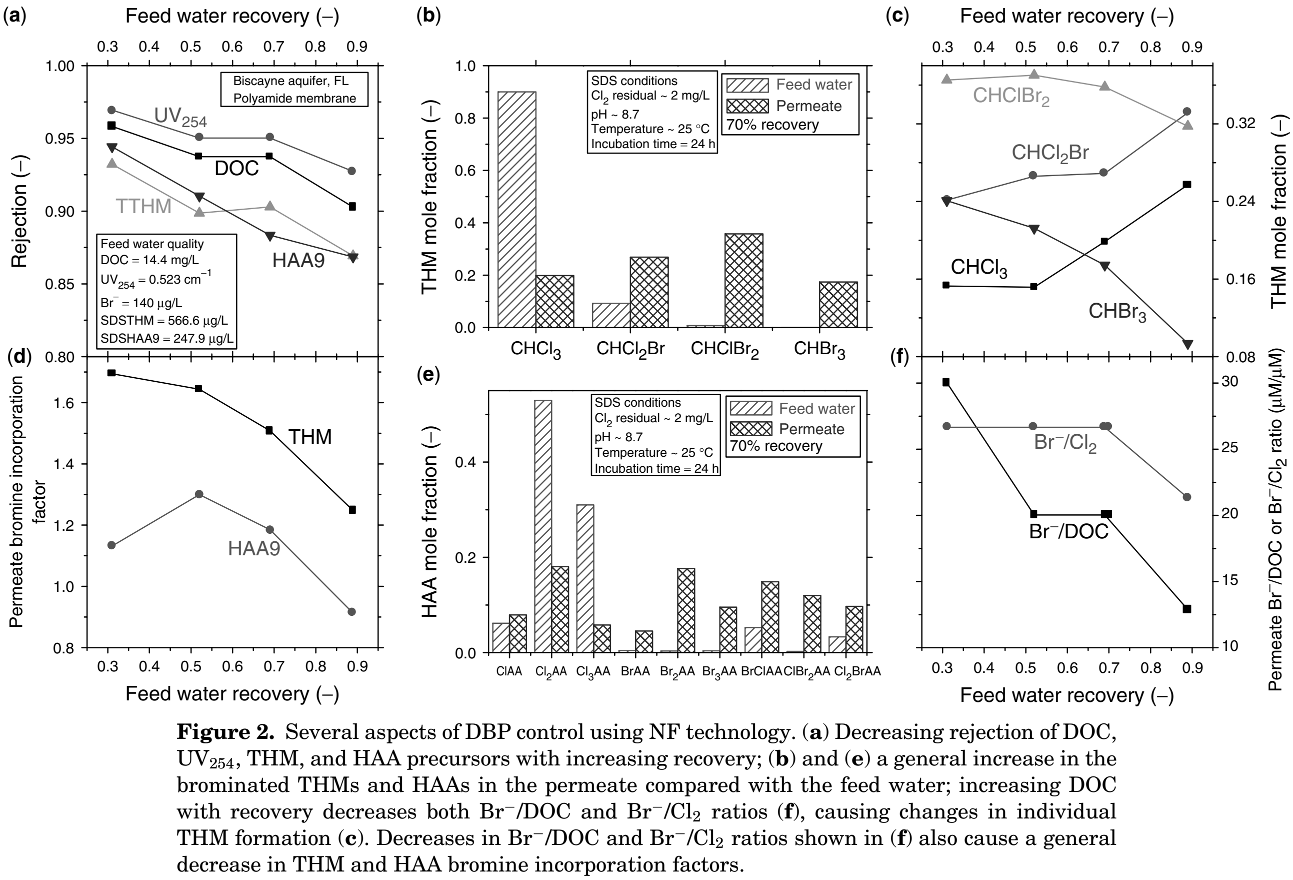

Figure 2. Several aspects of DBP control using NF technology. (**a**) Decreasing rejection of DOC, UV_{254}, THM, and HAA precursors with increasing recovery; (**b**) and (**e**) a general increase in the brominated THMs and HAAs in the permeate compared with the feed water; increasing DOC with recovery decreases both Br^-/DOC and Br^-/Cl_2 ratios (**f**), causing changes in individual THM formation (**c**). Decreases in Br^-/DOC and Br^-/Cl_2 ratios shown in (**f**) also cause a general decrease in THM and HAA bromine incorporation factors.

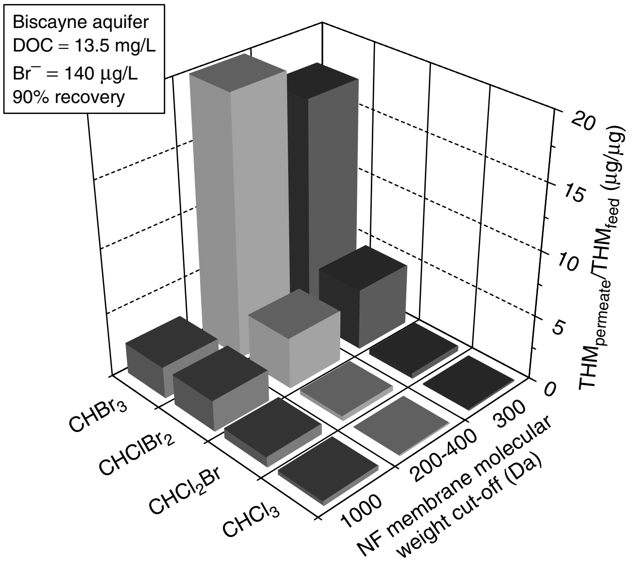

Figure 3. Effects of nanofiltration and membrane molecular weight cutoff on THM speciation for a naturally brominated groundwater measured under simulated distribution system conditions mentioned in Fig. 2.

also induced a fourfold increase in $CHClBr_2$ and a 20-fold increase in $CHBr_3$. As reported earlier (37), the shift toward the more brominated THMs was generally consistent with a decrease in MWCO, which increases Br^-/DOC ratio.

In summary, Figs. 1–3 as well as other investigations (44,47) unequivocally demonstrate the formation of the unregulated HAAs after advanced treatment. In fact, one study showed that the mole fraction of unregulated HAAs could be as high as ~50% following NF (47). Such substantial changes in speciation are not captured in existing DBP regulations, which only limit the sums of mass concentrations of four THMs and five HAAs. This important pitfall should be considered in formulating more equitable DBP regulations designed to protect human health.

Acknowledgments
This work has been funded by a grant from the National Science Foundation CAREER program (BES-0134301). The contents do not necessarily reflect the views and policies of the sponsors nor does the mention of trade names or commercial products constitute endorsement or recommendation for use. The assistance of Dr. Ying Wei from the City of Houston and Ramesh Sharma in measuring THMs and HAAs is appreciated.

BIBLIOGRAPHY

1. Singer, P.C. (Ed.). (1999). *Formation and Control of Disinfection By-Products in Drinking Water*. American Water Works Association, Denver, CO.
2. Singer, P.C. (1994). *J. Environ. Eng.* **120**: 727–744.
3. Rook, J.J. (1974). *Wat. Treat. Exam.* **23**: 234–243.
4. Christman, R.F. et al. (1983). *Environ. Sci. Technol.* **17**: 625–628.
5. Krasner, S.W. et al. (1989). *J. Am. Water Works Assoc.* **81**: 41–53.
6. Nawrocki, J. and Bilozor, S. (1997). *J. Water SRT-Aqua* **46**: 304–323.
7. Zavaleta, J.O. et al. *Formation and Control of Disinfection By-Products in Drinking Water*. P.C. Singer (Ed.). American Water Works Association, Denver, CO, pp. 95–117.
8. USEPA. (1998). *National Primary Drinking Water Regulations: Disinfectants and Disinfection By-Products*. Final Rule.
9. Minear, R.A. and Bird, J.C. (1980). *Water Chlorination: Environmental Impact and Health Effects*. R.L. Jolley, W.A. Brungs, and R.B. Cumming (Eds.). Vol. 3, Ann Arbor Science, Ann Arbor, MI, pp. 151–160.
10. Krasner, S.W. et al. (1996). *Disinfection By-Products in Water Treatment: The Chemistry of Their Formation and Control*. R.A. Minear and G.L. Amy (Eds.). CRC Press, Boca Raton, FL, pp. 59–90.
11. Krasner, S.W. et al. (1994). *J. Am. Water Works Assoc.* **86**: 34–47.
12. Oliver, B.G. (1980). *Water Chlorination: Environmental Impact and Health Effects*, Vol. 3, R.L. Jolley, W.A. Brungs, and R.B. Cumming (Eds.). Ann Arbor Science, Ann Arbor, MI, pp. 141–149.
13. Symons, J.M. et al. (1993). *J. Am. Water Works Assoc.* **85**: 51–62.
14. Symons, J.M. et al. (1996). *Disinfection By-Products in Water Treatment: The Chemistry of their Formation and Control*. R. Minear and G.L. Amy (Eds.). Lewis Publishers, Boca Raton, FL, pp. 91–130.
15. Pourmoghaddas, H. et al. (1993). *J. Am. Water Works Assoc.* **85**: 82–87.
16. Cowman, G.A. and Singer, P.C. (1996). *Environ. Sci. Technol.* **30**: 16–24.
17. Wu, W.W. and Chadik, P.A. (1998). *J. Environ. Eng.* **124**: 932–938.
18. Roberts, M.G. et al. (2002). *J. Am. Water Works Assoc.* **94**: 103–114.
19. Richardson, S.D. et al. (1999). *Environ. Sci. Technol.* **33**: 3378–3383.
20. Kargalioglu, Y. et al. (2002). *Teratogen. Carcinogen. Mutagen.* **22**: 113–128.
21. Plewa, M.J. et al. (2004). *Environ. Sci. Technol.* **38**: 62–68.
22. Singer, P.C. and Reckhow, D.A. (1999). *Water Quality and Treatment*, 5th Edn. R.D. Letterman (Ed.). McGraw Hill, New York, pp. 12.11–12.51.
23. Schnoor, J.L. et al. (1979). *Environ. Sci. Technol.* **13**: 1134–1138.
24. Liang, L. and Singer, P.C. (2003). *Environ. Sci. Technol.* **37**: 2920–2928.
25. Solarik, G. et al. (2000). *Natural Organic Matter and Disinfection By-Products: Characterization and Control in Drinking Water*. S.E. Barrett, S.W. Krasner, and G.L. Amy (Eds.). *ACS Symposium Series* Vol. 761. American Chemical Society, Washington, DC, pp. 47–66.
26. Sohn, J. et al. (2004). *Water Res.* **38**: 2461–2478.
27. Chowdhury, Z.K. and Amy, G.L. (1999). *Formation and Control of Disinfection By-Products in Drinking Water*. P.C. Singer (Ed.). American Water Works Association, Denver, CO, pp. 53–64.
28. Westerhoff, P. et al. (2000). *J. Am. Water Works Assoc.* **92**: 89–102.
29. Amy, G.L. et al. (1987). *J. Am. Water Works Assoc.* **79**: 89–97.
30. Koch, B. et al. (1991). *J. Am. Water Works Assoc.* **83**: 62–70.
31. Summers, R.S. et al. (1996). *J. Am. Water Works Assoc.* **88**: 80–93.

32. Reckhow, D.A. et al. (1990). *Environ. Sci. Technol.* **24**: 1655–1664.
33. Xie, Y.F. (2000). *Disinfection Byproducts in Drinking Water*. Lewis Publishers, Boca Raton, FL.
34. Gould, J.P. et al. (1983). *Water Chlorination: Environmental Impact and Health Effects*, Vol. 4, R.L. Jolley, W. Brungs, J. Cotruva, J. Mattice, and V. Jacobs (Eds.). Ann Arbor Science Publishers, Ann Arbor, MI, pp. 297–310.
35. Laine, J.-M. et al. (1993). *J. Am. Water Works Assoc.* **85**: 87–99.
36. Hwang, C.J. et al. (2000). *Natural Organic Matter and Disinfection By-Products: Characterization and Control in Drinking Water*. S.E. Barrett, S.W. Krasner, and G.L. Amy (Eds.). *Vol. ACS Symposium Series* 761, American Chemical Society, Washington, D.C., pp. 173–187.
37. Jacangelo, J.G. et al. (1995). *J. Am. Water Works Assoc.* **87**: 64–77.
38. Letterman, R.D. (Ed.) (1999). *Water Quality and Treatment*, 5th Edn. McGraw Hill, New York.
39. Snoeyink, V.L. et al. (1999). *Formation and Control of Disinfection By-Products in Drinking Water*. P.C. Singer (Ed.). American Water Works Association, Denver, CO, pp. 259–284.
40. Snoeyink, V.L. and Summers, R.S. (1999). *Water Quality and Treatment*, 5th Edn. R.D. Letterman (Ed.). McGraw Hill, New York, pp. 13.11–13.83.
41. Amy, G.L. et al. (1991). *Water Res.* **25**: 191–202.
42. Summers, R.S. et al. (1993). *J. Am. Water Works Assoc.* **85**: 88–95.
43. Taylor, J.S. and Jacobs, E.P. (1996). *Water Treatment Membrane Processes*. J. Mallevialle, P.E. Odendaal, and M.R. Wiesner (Eds.). McGraw Hill, New York, pp. 9.1–9.70.
44. Chellam, S. (2000). *Environ. Sci. Technol.* **34**: 1813–1820.
45. Allgeier, S.C. and Summers, R.S. (1995). *J. Am. Water Works Assoc.* **87**: 87–99.
46. Chellam, S. and Taylor, J.S. (2001). *Water Res.* **35**: 2460–2474.
47. Chellam, S. and Krasner, S.W. (2001). *Environ. Sci. Technol.* **35**: 3988–3999.

ACTIVATED CARBON: ION EXCHANGE AND ADSORPTION PROPERTIES

Basu Saha
Loughborough University
Loughborough, United Kingdom

INTRODUCTION

The importance of environmental pollution control has increased significantly in recent years. Environmentalists are primarily concerned with the presence of heavy metals, pesticides, herbicides, chlorinated hydrocarbons, and radionuclides in groundwater, surface water, drinking water, and aqueous effluents because of their high toxicity and impact on human and aquatic life.

Several techniques have been developed and used to remove and/or recover a wide range of micropollutants from water and a variety of industrial effluents. Adsorption using activated carbon is well established for the removal of organic molecules from aqueous solution, but to a much lesser extent for the removal of toxic heavy metals.

Here, the preparation, properties, and metal sorption performance of a range of as-received and oxidized samples of granular activated carbon that were either prepared or modified in the author's laboratory are discussed. Samples were evaluated for the removal of trace toxic metal ions from aqueous solutions. Batch and column experiments have been performed to elucidate the relationship between sorptive performance and the physical and chemical structure of these materials. A detailed discussion of adsorption and ion-exchange properties of engineered activated carbons and carbonaceous materials can be found in a recent publication by Streat et al. (1).

ACTIVATED CARBONS

Activated carbon is the generic name given to amorphous carbonaceous materials with an extensively developed internal pore structure. The characteristics of activated carbons depend on the carbonaceous precursor and the technique of activation used in their production. Activated carbon is obtained by thermal decomposition, combustion, and partial combustion of different carbonaceous substances. The carbonaceous substances are usually cheap vegetable material high in carbon content and low in ash content. The raw material can be wood, peat, fruit stones, nutshells, or sawdust. (2). Nonvegetable sources include blood, bones, leather waste, fish, petroleum products, and by-products (3). Currently, a bulk of the activated carbon comes from natural coal, which is because of its low cost and availability.

Activated carbons have proved versatile adsorbents mainly because of their large surface area and surface reactivity. They also possess favorable pore size, which makes the internal surface accessible, enhances the adsorption rate, and enhances mechanical strength (2). According to the International Union of Pure and Applied Chemistry (IUPAC) system, the pore structures are divided into three categories, namely, micropores (<2 nm in diameter), mesopores (between 2 and 50 nm in diameter), and macropores (>50 nm). The pores exist in different shapes, which can be V-shaped pores, tapered pores, capillary shaped with one end closed or open at both ends, contracted entrance pores, and so on.

Micropores constitute the bulk of the area, and their sizes are usually comparable with the adsorbate molecules. The mechanism of adsorption in micropores is micropore filling, and it comes as a result of the adsorption force field that encompasses the whole volume of micropores. This type of mechanism leads to higher heat of adsorption energy as compared with the adsorption energy on surface or mesopore regions. Mesopores have a dual function that involves adsorption and channels for the adsorbate to access the interior of the adsorbent. The mechanism of adsorption is via capillary condensation. Macropores serve as passages for the adsorbate to access the interior of the adsorbent, and they do not contribute significantly to the adsorption capacity.

Activated carbon is a microcrystalline, nongraphitic form of carbon and can be produced in the form of powders, granules, and shaped products satisfying any modern engineering requirements. Its exceptional adsorptive properties are attributed to the controlled size and distribution of pores within the carbon matrix. The strong market position of the activated carbons relates to their unique properties and lower cost compared with that of other possible competitive adsorbents.

Through the choice of precursor, the method of activation, and control of processing conditions, the adsorptive properties of activated carbons can be specially designed for many diverse applications such as the purification of potable water and adsorption of various emissions from motor vehicles. In drinking water treatment, the carbon is used to remove dissolved organics with molecular weights in excess of 45 g/mol (4).

HISTORICAL BACKGROUND

The use of activated carbon as an adsorbent dates back to the Egyptians in about 1500 B.C. (2). It was used in the form of carbonized wood mainly for medical purposes as well as a purifying agent. However, the basis of its industrial production and use was established in the late eighteenth century and its uses have broadened since then. The first industrially manufactured activated carbons were Eponit decolorizing carbons. They have been produced since 1909 by heating wood charcoal with steam and carbon dioxide in a furnace specially designed for this purpose. Activated carbons are now produced industrially by several different methods, as many new production technologies have appeared.

Granular activated carbon found uses for purification of gases and extraction of vapors from gas streams, recovery of liquid petroleum from natural gas, extracting benzene from manufactured coal gas, and recovering volatile solvents from various industrial applications. Activated carbon has also been applied to the removal of color, odor, taste, and many chemical undesirable species from aqueous solutions either by adsorption or ion exchange (1). It is now extensively used for potable, domestic, and industrial waste water purification, sugar refining, and metal recovery (2,3,5,6).

Over the last two decades, an interest has developed in activated carbon, and applications continue to evolve in response to growing demands for environmental protection and emerging technologies.

PRECURSORS

The materials used for the production of activated carbon are of high carbon content and contain a low concentration of inorganic substances. Precursor materials include wood, animal bones, nutshells, coconut shell, fruit stones, brown coal, bituminous coal, charcoal, peat, lignin, lignite, pitch, mineral oil products, and petroleum and lubricant wastes (3,7). However, only wood, bones, coal, lignite, coconut shell, and peat have been used on an industrial scale because they are currently the only economically viable raw materials (3,4).

As already mentioned, many potential raw materials exist for the production of activated carbon, but not all of them produce activated carbon of high quality. The reason behind this lies in the high amount of inorganics that exist in the structure of the raw materials as well as in their scarcity. Poor quality can also result from poor storage conditions and difficult workability of the raw materials.

The principal properties of manufactured active carbons depend on the type and properties of the raw material used. The properties of the final product will also depend on the nature of the raw material but mostly on the nature of the activating agent and the conditions of the activation process, which makes the carbonization and activation processes the most important factors in producing a good quality activated carbon. The next section gives a brief review on how activated carbon is produced and discusses how the activation of the carbon can result in a material suitable for specialized applications.

CARBONIZATION

The first stage of production of the activated carbons is the carbonization process, where the properties of the activated carbon can be modified. The carbonization process involves thermal decomposition and elimination of the noncarbon species and can introduce a simple pore structure that depends principally on the physical properties of the precursor and can be controlled by the rate of heating.

Generally, carbonization is conducted by heating the initial material up to 973 K in the absence of oxidizing gases. During this process, the raw material is dehydrated, that is, oxygen- and hydrogen-containing compounds are decomposed and eliminated from the precursor to produce a skeleton possessing a latent pore structure. If the temperature of carbonization is low, impurities such as tar may be left in the pores without decomposition. The presence of impurities, which block the pores, may result in a low surface area and limited surface activity. The mechanical and adsorptive properties of an activated carbon can also be introduced and controlled by various additives introduced during the stabilization and carbonization process.

ACTIVATION

The activation process is the second stage of production, where any disorganized carbon from the surface and the carbon from the aromatic sheets under the surface are burned off by contacting the carbonized material either with gases or with chemical reagents, which results in wider pores and new porosity with an oxidized surface.

Generally, two main activation processes are used by industry to manufacture activated carbon: physical (or thermal) and chemical activation. The main difference between the two is that in physical activation, the process is done at high temperature and the yield of activated carbon is low, whereas in chemical activation, the process is carried out at a lower temperature and results in a much higher yield.

Physical activation involves carbonization of the raw material and the activation of the carbonized product. Carbonization is generally performed at temperatures around 873–923 K in the absence of oxidizing gases (8). The resulting material is enhanced by activation via partial gasification with mild oxidizing agents such as steam, carbon dioxide, or a mixture of both at 1073–1273 K (9). In the case of activation with air or oxygen, excessive burning of the carbon exists and the reaction is difficult to control, which is the main reason that these oxidizing agents are not used for activation purposes. Generally, microporosity is introduced when activation is done using carbon dioxide, whereas meso- or macroporosity develops if steam is used for the activation.

Chemical activation, on the other hand, involves the reaction of a carbon precursor with a reagent. The contact of the raw material with the reagent takes place while the temperature increases from 623 K to 1173 K. The commonly used reagents for chemical activation in industry until 1970 were zinc chloride, sulfuric acid, and phosphoric acid. The preferred precursors exposed to this type of activation are lignocellulosic materials, wood being the most common. Today, only nonmineral precursors can be applied successfully, with wood and olive stones the most commercially used. Jankowska and Swiatkowski (10) reported that this activation process produces a carbon matrix that contains a considerable amount of heteroatoms (including residual of phosphates, sulfates, and zinc).

PHYSICAL AND CHEMICAL PROPERTIES OF ACTIVATED CARBONS

The structure of activated carbon is best described as a twisted network of defective carbon layer planes, cross-linked by aliphatic bridging groups (11). The lattice is composed of pores sizes whose size distribution can be controlled by the choice of carbon feedstock and mode of preparation. The surface of activated carbons is not uniform in either structural or in energetic respect. Atoms at the carbon surface are in a different electronic state than those in a pore phase, especially on the edges of the carbon layers; where defects, dislocations, and discontinuities of the layer planes are present. Such sites, called the “active” sites, are associated with high concentration of unpaired electron spin centers and are therefore expected to play a significant role in chemical and adsorptive interaction of the activated carbon with different compounds (12).

Today, it is well established that almost all types of activated carbons are covered with oxygen complexes, unless special care is taken to eliminate them. Therefore, the adsorptive properties of an activated carbon are not only determined by its porous structure but also by its chemical composition. These complexes are often the source of the properties, which make an activated carbon useful or effective in certain applications. Two main types of surface functional groups exist: acidic and basic surface groups, which are discussed in the following section.

ACID SURFACE FUNCTIONAL GROUPS

Several types of oxides having an acidic character have been identified to exist on the surface of activated carbons by different types of techniques: specific organic reactions, potentiometry, Fourier transform infrared (FTIR), and X-ray photoelectron spectroscopy (XPS). The structures presented in Fig. 1 are the oxygen containing groups that are found at the edges of the graphene layers. The most important ones are the carboxyl (a), carbonyl (f), and single hydroxyl (phenolic) (e) groups.

Other groups can then be generated from a combination of these three because the simultaneous presence of various types of organic groups may lead to more complex surface structures. Thus, carboxylic anhydrides (b) can be produced when carboxyl groups are close together, lactone groups (c) when hydroxyl and carboxyl groups are joint, and lactoles (d) when carboxyl and carbonyl groups come close enough (Fig. 2). Carbonyl groups can also be present on the surface of the carbon as conjugated structures called quinones (g). Lastly, ether type structures (h), which are stable even at high temperatures, can also be found on the activated carbon surface.

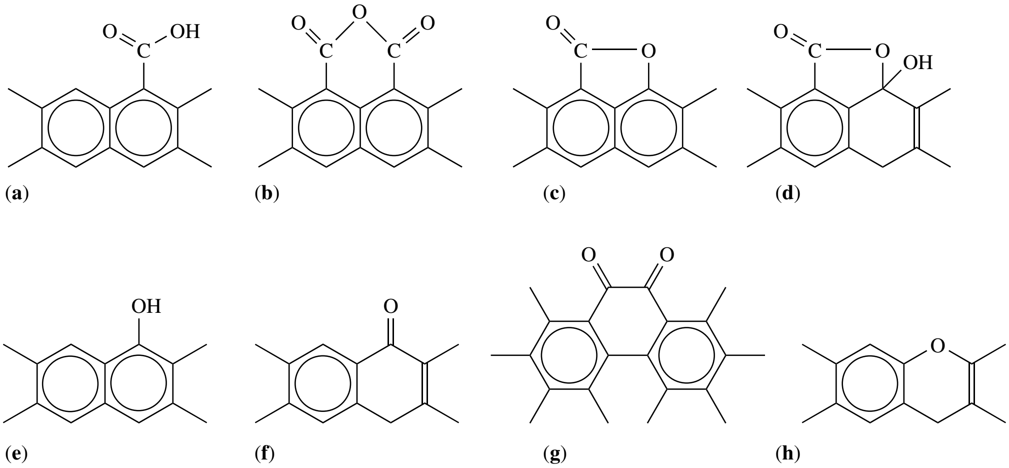

Figure 1. Possible structures of surface oxygen groups (redrawn from Ref. 13).

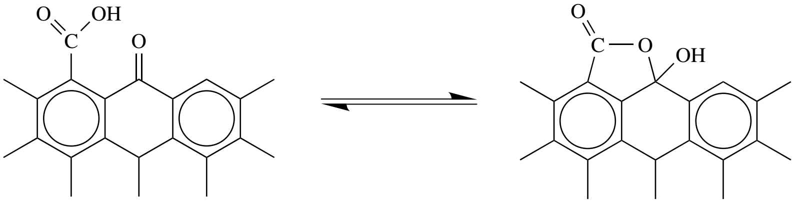

Figure 2. Possible interaction of oxygen-containing groups (redrawn from Ref. 13).

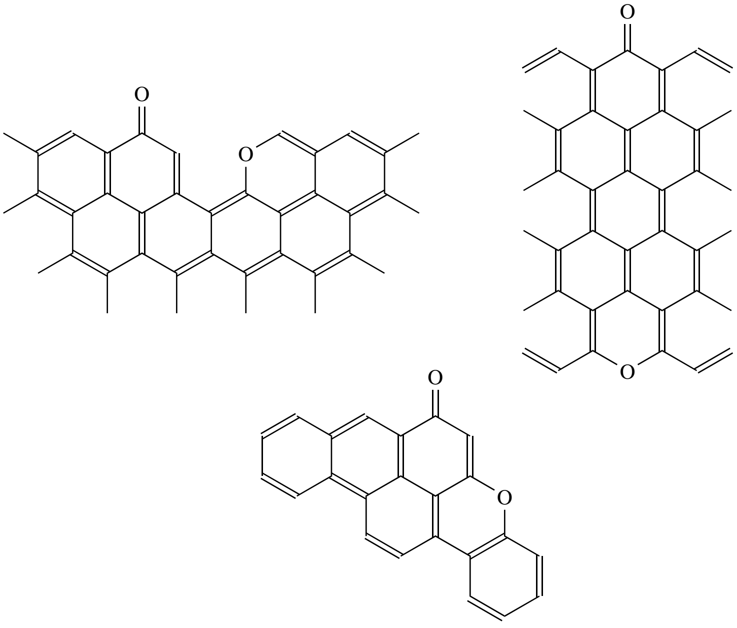

Figure 3. Possible structures of basic surface sites on a graphene layer (redrawn from Ref. 16).

BASIC SURFACE FUNCTIONAL GROUPS

Very scant information is available about basic surface functional groups. They are formed when a carbon is heat treated above 973 K, cooled in an inert atmosphere, and reexposed to oxygen at room temperature. In order to explain the uptake of acids, Garten et al. (14) and Garten and Weiss (15) proposed a chromene structure, but g-pyrone-like structures are more likely (16). These structures are schematically shown in Fig. 3, and they characteristically possess two oxygen groups situated on different rings of the graphitic layer.

Evidence also exists that protons can be adsorbed on basal planes either by reaction with pyrones or by donor–acceptor interactions. The end result is the formation of basic surface structures, which can be represented by pyrone-like structures with oxygen atoms, in general, located in two different rings of a graphite layer; the positive charge being stabilized by the resonance (Fig. 4).

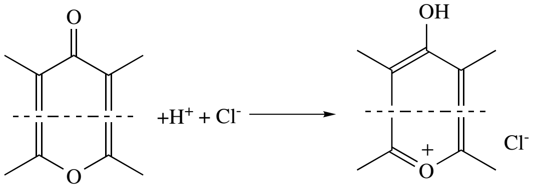

Figure 4. Reaction of proton with pyrone-like structure (redrawn from Ref. 16).

PREPARATION AND PROPERTIES OF ACTIVATED CARBON

The synthesis, structure, and adsorption properties of engineered activated carbons with reference to the body of work carried out in the author's research laboratory are reviewed. The main emphasis is granular activated carbon (GAC). The majority of engineered carbons discussed here have graphitic or disordered graphitic microstructures. Also, for most engineered carbon materials, the originating precursor is organic and the materials develop from heat treatment of the precursor in inert atmospheres (carbonization). A selection of technically important carbons developing from solid, liquid, and gaseous organic precursors are presented in Table 1.

The role of pore structure and surface area of engineered carbons and their impact on adsorbent performance are discussed elsewhere (1,17,18). The surface chemical structure affects interaction with polar and nonpolar molecules because of the presence of chemically reactive functional groups (see Fig. 1). The adsorption is a complex interplay between the chemical and porous surface structure of the carbon. Important factors including the nature and relative amounts of surface functional groups, the surface area, and pore size distribution, as well as the characteristics of the adsorbate molecule (e.g., size and

Table 1. Starting Precursors for Carbons (Adapted from Ref. 1)

Primary	Secondary	Example of Carbons
Hydrocarbon gases		Pyrocarbons, carbon blacks, vapor grown carbon fibers, matrix carbon[a]
Petroleum derived		Delayed coke, calcined coke
	Petroleum pitch	Needle coke, carbon fibers, binder and matrix carbon[a]
	Mesophase pitch	Mesocarbon microbeads, carbon fibers
Coals		Semicoke, calcined coke
	Coal chars	Activated carbons
	Coal tar pitch	Premium cokes, carbon fibers, binder and matrix carbons[a]
	Mesophase pitch	Mesocarbon microbeads, carbon fibers
Polymers	Polyaddition and polycondensation type resins	PAN-based carbon fibers
		Glassy carbons, binder and matrix carbons, carbon beads, graphite films and monoliths
Biomass	Coconut shells, apricot, olive and peach stones, pits	Activated carbons

[a] Precursor as a binder in granular carbons and graphites and in carbon–carbon composites.

nature of the cation/anion, polarity and chemical structure of the molecule, molecular/ionic dimensions) are discussed by Streat et al. (1). The importance of precise characterization of the surface and structure of engineered carbons in order to tailor adsorbents for specific applications is outlined by Saha et al. (17–19).

Granular oxidized carbon samples showed similar ion-exchange properties to active carbon fibers with respect to influence of solution pH on ion-exchange behavior. The pH dependence may be attributed to acidic surface groups that dissociate as a function of solution pH. The point-of-zero charge (PZC) plays an important role in the ion-exchange behavior of the carbon adsorbents in removing metal ions from solution (17).

SORPTION OF TRACE METALS ONTO ACTIVATED CARBON

Previous publications have discussed the adsorption of heavy metal ions from aqueous solutions by carbonaceous materials (20–24). It is widely accepted that surface acidic functional groups are responsible for metal ion binding (20). Chemical oxidation is commonly used to introduce these functional groups onto the surface of carbons.

The results of detailed minicolumn experiments to determine breakthrough characteristics and regeneration performance of selected oxidized activated carbons are outlined in this section. Samples of BGP (a wood-based activated carbon) were tested using solutions containing target metals to determine overall performance criteria, which consisted of testing samples of BGP against solutions containing copper, nickel, zinc, or cadmium. All breakthrough capacities were calculated based on the total amount of metal removed before 5% or 50% of the feed concentration of metals was detected in the outlet, which was to give an indication of uptake performance of each material for the four metals used. The results of the minicolumn experiments are summarized in Table 2.

Table 2. Uptake Capacity of Carbons from 1 mM Metal Solutions (pH 4.7)

			Breakthrough Capacity (mmol/g)	
Sample[a]	Metal	pH	5%	50%
BGP OxII	Cu	4.7	0.35	0.42
BGP AOx	Cu	4.7	0.04	0.05
BGP Unoxidized	Cu	4.7	—	0.004
BGP OxII	Zn	4.7	0.13	0.16
BGP AOx	Zn	4.7	—	0.13
BGP OxII	Ni	4.7	0.14	0.20
BGP AOx	Ni	4.7	0.03	0.06
BGP OxII	Cd	4.7	0.16	0.21
BGP AOx	Cd	4.7	—	0.02

[a] AOx, air oxidized; OxII, alkali washed, acid oxidized for 9 h.

An unoxidized sample of BGP was used to generate a breakthrough curve to compare copper sorption capacity with oxidized samples. The copper breakthrough curves can be seen in Fig. 5.

Results show that significant improvement occurred in copper sorption capacity on oxidation of the original sample. The copper sorption capacities increased by a factor of up to 100 by oxidizing the original material. This change can be attributed to oxygen surface groups introduced onto the surface of Ceca BGP. The 5% copper breakthrough capacities of BGP OxII and BGP AOx were 0.35 mmol/g and 0.04 mmol/g, respectively. Over 100 bed volumes of 1 mM solution were passed before any copper could be detected leaving the column using the acid oxidized sample, BGP OxII.

BGP AOx was also used to generate breakthrough curves for removal of nickel, zinc, and cadmium. Results for these experiments can be seen in Fig. 6. It shows that the metal sorption capacity of air-oxidized samples was much less than acid-oxidized samples for a feed concentration of 1 mM and pH of 4.7. Breakthrough capacities at 5% were 0.02 mmol/g for nickel, and for other samples, the 5% breakthrough was instantaneous. The low capacity of BGP

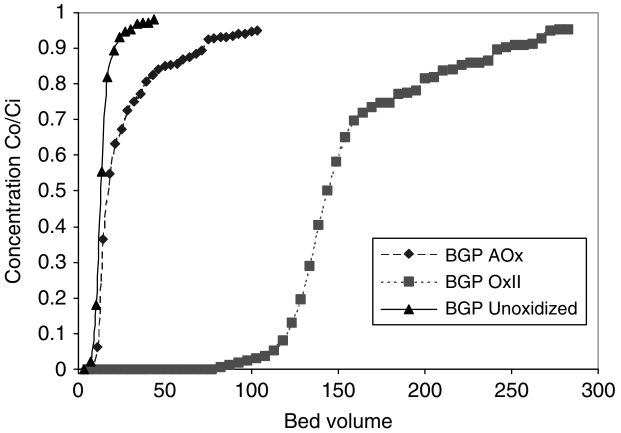

Feed concentration = 1 mM copper, feed pH = 4.7, flow rate = 5 BV/h

Figure 5. Copper breakthrough curves for samples of BGP.

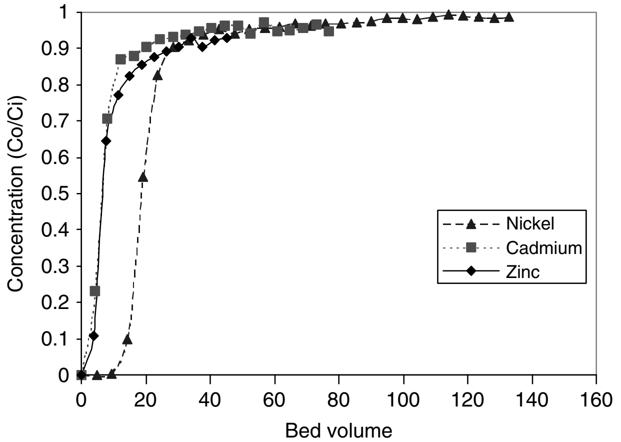

Feed concentration = 1 mM Ni, Zn or Cd, feed pH = 4.7, flow rate = 5 BV/h

Figure 6. Breakthrough curves for nickel, zinc, and cadmium using BGP AOx sample.

AOx sample can be partially attributed to the high reaction temperature during the oxidation rising above 723 K, which would have resulted in desorption of some oxygen surface groups. Very poor breakthrough performance was obtained for the BGP AOx sample with the exception of the nickel experiment when 10 bed volumes were treated before any metal was detected in the outlet.

A combined feed containing all four metals, copper, zinc, nickel, and cadmium, was passed through the column containing 1 g of BGP OxII (see Fig. 7). The feed concentration of all metals was 0.25 mM and the feed pH was 4.7, which indicate the following selectivity series: $Cu^{2+} > Ni^{2+}, Cd^{2+} > Zn^{2+}$. The increase in the concentration above the feed was because of chromatographic elution. A solution containing copper, zinc, nickel, and cadmium was used to obtain the selectivity of BGP AOx. The results shown in Fig. 8 indicate a selectivity order of $Cu^{2+} > Ni^{2+} > Zn^{2+}, Cd^{2+}$. These studies confirm that both materials are highly selective toward copper in the presence of other metal ions.

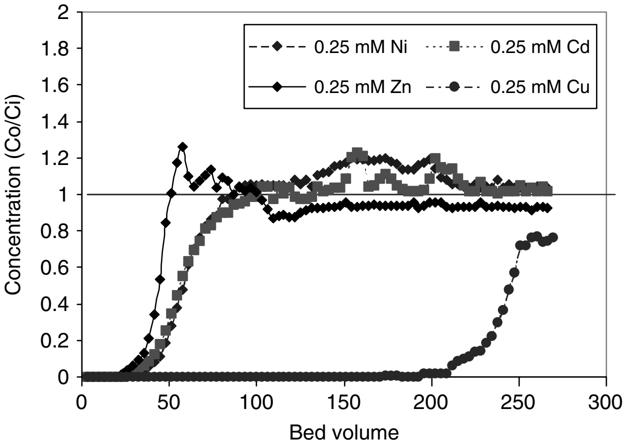

Feed concentration = 0.25 mM each of Cu, Ni, Zn and Cd, feed pH = 4.7, flow rate = 5 BV/h

Figure 7. Breakthrough curve using a feed solution containing copper, nickel, zinc, and cadmium with a sample of BGP OxII.

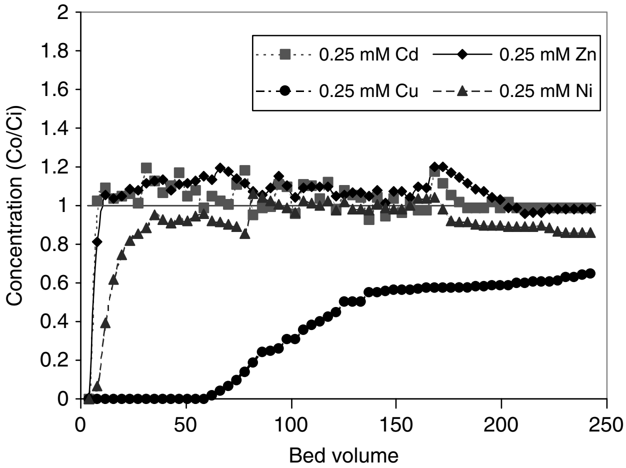

Feed concentration = 0.25 mM each of Cu, Ni, Zn and Cd, feed pH = 4.7, flow rate = 5 BV/h

Figure 8. Breakthrough curve using a feed solution containing copper, nickel, zinc, and cadmium with a sample of BGP AOx.

The BGP OxII elution results are shown in Fig. 9. Most of the metals were eluted with 0.1M HCl in the first 10 bed volumes except for cadmium. All the copper was recovered, but only 86% of nickel, 90% of zinc, and 85% of cadmium were eluted. Elution efficiency of BGP AOx was 100% for all metals. The results show that most of the metals were recovered before 10 bed volumes of eluant had been passed through the columns.

The sorptive capacity and selectivity of oxidized carbons vary for different metal ions, and higher valency metal ions are usually preferred to those of lower valency (20,21). It has also been observed that selectivity differs even within a series of metals with the same valency. Among divalent metals, Cu^{2+} is generally the most preferred ion. However, the reason for the higher affinity of oxidized carbons toward this particular metal ion has not been clearly identified. The formation of metal surface complexes on the oxidized carbon involving cooperative action seems quite likely (see Fig. 10).

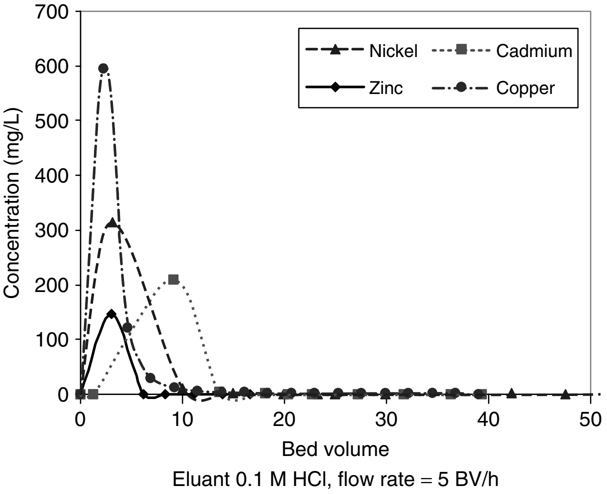

Figure 9. Elution curves of BGP AOx.

It can be concluded that air and acid oxidation significantly enhance the sorption capacity of Ceca BGP. BET and Langmuir surface areas were relatively unchanged by air oxidation, but some surface area was lost during acid oxidation. Acid-oxidized samples have a much higher copper uptake capacity than the as-received material. Air-oxidized samples show lower metal uptake capacity compared with acid-oxidized samples. Acid-oxidized samples are heat treated and alkali washed to remove organic by-products that are formed during the oxidation. At every stage, some material is lost, reducing the yield of material based on weight of as-received material compared with the final product. The sorption capacity of the materials was reduced by these treatments, indicating the metal binding capacity of these organic by-products formed during oxidation. Equilibrium capacity of modified samples decreased with a reduction in the pH because of the weakly acidic nature of the sorbents. Kinetic experiments indicate that about 80% of the metals are removed in the first twenty minutes for all samples. Minicolumn experiments show that modified carbons are efficient sorbents for the removal of trace metal ions from solution. After initial breakthrough, complete loading of carbons occurs slowly, which is most likely because of the presence of micropores and small mesopores within the structure of the materials. All samples are regenerated using 0.1 M HCl solution. Both modified samples demonstrated good regeneration efficiencies with 100% of the copper recovered during elution cycles compared with the amount of metal removed during the sorption experiments.

CONCLUSIONS

Engineered activated carbons can be manufactured in a range of physical forms and from a wide variety of starting precursor materials. Pyrolysis and activation combined with further surface pretreatment results in a range of adsorbent materials with widely varying physicochemical properties. It is argued that liquid-phase sorption of solutes on carbon surfaces involves a complex interplay between the delocalized π-electron system developing from the aromatic backbone of the graphite microstructure and the electron withdrawing surface functional groups (e.g., oxygen functional groups). This interplay not only lowers the point-of-zero charge of the surface but also reduces the dispersive adsorption potential by decreasing the π-electron density in the graphene layers (2). Optimization of the physical properties of carbonaceous adsorbents, that is, surface area and a well-developed pore structure, are essential for the application of the adsorbents in conventional water treatment technologies (e.g., using fixed-bed adsorbers). In addition, it is vital that the chemical properties of the adsorbent surface must be tailored to suit the individual application. Carbon surface functional groups influence the selectivity of the adsorbent for metal ions in solution. Oxidation treatment of commercially available carbons dramatically enhances their ion-exchange properties and results in reversible adsorption of metal ions from solution. Design of adsorption systems using engineered carbons requires careful optimization of process parameters with due attention to the speciation chemistry of metal ions and the solution pH that markedly influences the interaction of the solute species with the adsorbent surface. We are now in a position to reverse-engineer tailored structured carbonaceous adsorbents with optimized physicochemical properties suited for selective separation of solutes from aqueous streams.

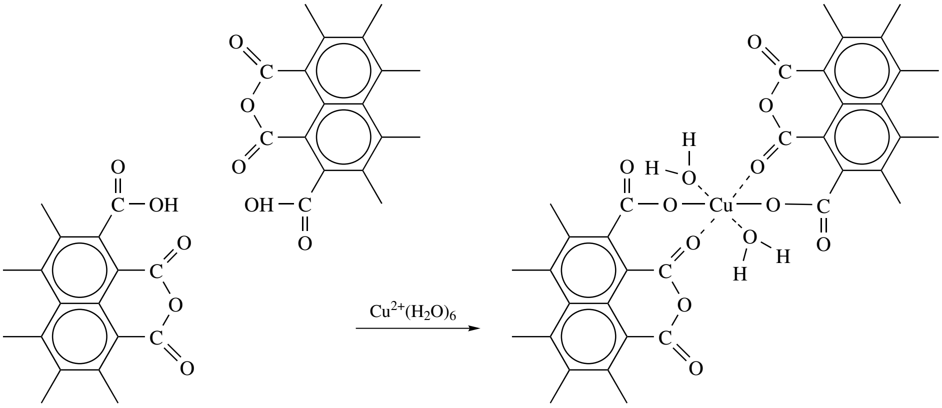

Figure 10. Postulated complexation reaction between copper (II) and oxidized carbon surface.

Acknowledgments

I would like to express my sincere gratitude to Professor Michael Streat for his support, continuous encouragement, and advice for conducting the research in the area of adsorption and ion exchange and development of tailored sorbent materials for environmental remediation. I am indebted to Dr. Hadi Tai, Purazen Chingombe, and Eleni Karounou for their painstaking experimental work. Funding from European Community, EPSRC, Severn Trent Water Ltd., and Commonwealth Scholarship Commission is gratefully acknowledged.

BIBLIOGRAPHY

1. Streat, M., Malik, D.J., and Saha, B. (2004). In: *Ion Exchange and Solvent Extraction*. Vol. 16. A.K. SenGupta and Y. Marcus (Eds.). Marcel Dekker, New York, pp. 1–84.
2. Bansal, C.R., Donnet, J.B., and Stoeckli, F. (1988). *Active Carbon.* Marcel Dekker, New York, pp. 1–17.
3. Hassler, J.W. (1967). Activated carbon. *Chemical and Process Engineering Series*. pp. 1–171.
4. Noll, K.E. (1991). *Adsorption Technology for Air and Water Pollution Control*, Lewis Publishers, Chelsea, MI, pp. 1–211.
5. Donnet, J.B. and Bansal, R.C. (1984). *Carbon Fibres*. Marcel Dekker, New York, pp. 1–40.
6. Roy, G.M. (1995). *Activated Carbon Applications in Food and Pharmaceutical Industries*. Technomic, pp. 1–15.
7. Mantell, C.L. (1968). *Carbon and Graphite Handbook*. John Wiley & Sons, Hoboken, NJ, pp. 6–34.
8. Smíšek, M. and Cerný, S. (1967). *Active Carbon: Manufacture Properties and Applications*. Elsevier, Amsterdam.
9. Freeman, J.J., Gimblett, F.G.R., Roberts, R.A., and Sing, K.S.W. (1987). *Carbon* **25**(4): 559–563.
10. Jankowska, H. and Swiatkowski, A. (1991). *Active Carbon*. Ellis Horwood Limited, pp. 9–74.
11. McEnaney, B. and Mays, T.J. (1989). *Introduction to Carbon Science*. Butterworths, London, pp. 153–156.
12. Puri, B.R. (1970). *Chemistry and Physics of Carbon*. Vol. 6. Marcel Dekker, New York, pp. 1–191.
13. Boehm, H.P. (1994). *Carbon* **32**(5): 759–769.
14. Garten, V.A., Weiss, D.E., and Willis, J.B. (1957). *Austral J. Chem.* **10**: 295–308.
15. Garten, V.A. and Weiss, D.E. (1957). *Rev. Pure. Appl. Chem.* **7**: 69–122.
16. Voll, M. and Boehm, H.P. (1971). *Carbon* **9**: 481–488.
17. Saha, B., Tai, M.H., and Streat, M. (2001). *Trans. IChemE.* **79**(Part B): 211–217.
18. Saha, B., Tai, M.H., and Streat, M. (2003). *Chem. Eng. Res. Des.* **81**: 1343–1353.
19. Saha, B., Tai, M.H., and Streat, M. (2001). *Trans. IChemE.* **79**(Part B): 345–351.
20. McEnaney, B. (1999). In: *Carbon Materials for Advanced Technologies*. T.D. Burchell (Ed.). Pergamon, New York, pp. 1–33.
21. Kadirvelu, K., Faur-Brasquet, C., and Le Cloirec, P. (2000). *Langmuir.* **16**: 8404–8409.
22. Shim, J.-W., Park, S.-J., and Ryu, S.-K. (2001). *Carbon* **39**: 1635–1642.
23. Corapcioglu, M.O. and Huang, C.P. (1987). *Water Res.* **21**: 1031–1044.
24. Budinova, T.K., Gergova, K.M., Petrov, N.V., and Minkova, V.N. (1994). *J. Chem. Technol. Biotechnol.* **60**: 177–182.

ACTIVATED CARBON—POWDERED

Absar A. Kazmi
Nishihara Environment
Technology
Tokyo, Japan

INTRODUCTION

The most commonly used adsorbent in water and wastewater treatment is activated carbon. Activated carbon is manufactured from carbonaceous material such as wood, coal, or petroleum residues. A char is made by burning the material in the absence of air. The char is oxidized at higher temperatures to create a very porous structure. The "activation" steps provide irregular channels and pores in the solid mass, resulting in a very large surface-area-per-mass ratio. Surface areas ranging from 500 to 1500 m^2/g have been reported, and all but a small surface area is in the pores. Once formed, activated carbon is pulverized to a very fine powder (Fig. 1). The size is predominantly less than 0.075 mm (200 sieve). Dissolved organic material adsorbs to both exterior and interior surfaces of the carbon. The characteristics of powdered activated carbon are summarized in Table 1.

APPLICATION IN WATER TREATMENT

Seasonal application of powdered activated carbon (PAC) at the raw water intake or rapid mix unit is used by some plants to correct short-term raw water quality problems such as algal blooms. PAC is used to correct taste and

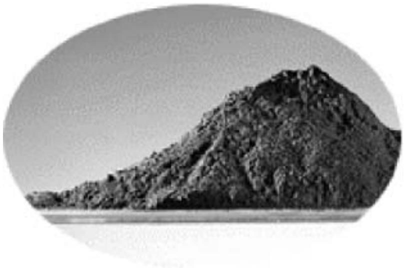

Figure 1. Powdered activated carbon.

Table 1. Characteristics of Powdered Activated Carbon

Parameter	Unit	Values
Total surface area	m^2/g	800–1800
Bulk density	kg/m^3	360–740
Particle density, wetted in water	kg/L	1.3–1.4
Particle size range	mm	5–50
Effective size	mm	—
Uniformity coefficient	UC	—
Mean pore radius	Å	20–40
Iodine number	—	800–1200
Abrasion number	Minimum	70–80
Ash	%	≤ 6
Moisture as packed	%	3–10

odor problems, which are primarily aesthetic qualities of water. Contact time is needed to allow adsorption to occur. Powdered activated carbon is contacted with water in open vessels where it is maintained in suspension for the necessary contact time and then removed by conventional solids removal processes (Fig. 2). The general feeling among water treatment plant operators is that, when everything else fails, it is time to use PAC. It is quite reliable, but its effectiveness depends on the type of carbon, the dosage, and the point of application. Usual carbon dosage ranges from 0.5 to 5.0 mg/L.

Normally, the PAC dose is based on operator experience, although laboratory tests may provide more precise dosages. In these tests, a series of carbon dosages is added to odorous water and the residual odor levels are determined. When a single odorous chemical is present, the initial and residual concentrations can be measured by using advanced laboratory identification techniques. Where such equipment is not available, the odor source is unknown, or numerous compounds are involved, the threshold odor number (TON) method may suffice. These data then can be plotted in one of several ways to clarify the removal response, to produce a predictive trend, and to determine the dosages to achieve acceptable water quality.

A simple plot of carbon dose versus TON (or the residual concentration) might be sufficient to highlight the relationship; however, the preferred method is to identify constants in an equation for removal and apply these to dosage calculations. The Freundlich and Langmuir isotherm equations are commonly used. The primary difference is that the Freundlich isotherm equation is strictly an empirical approach, accepted as standard; the Langmuir isotherm has a theoretical basis and has wider applicability.

The Freundlich isotherm is expressed by Eq. 1 and its corollary by Eq. 2

$$\frac{X}{M} = kC^{1/n} \tag{1}$$

$$\ln\left(\frac{X}{M}\right) = \ln(k) + \frac{1}{n}\ln C \tag{2}$$

X/M = mass of adsorbate per unit mass of adsorbent, mg adsorbate/g activated carbon
C = final or equilibrium concentration of the adsorbate, mg/L
k = empirical constant (y intercept)
n = slope inverse constant

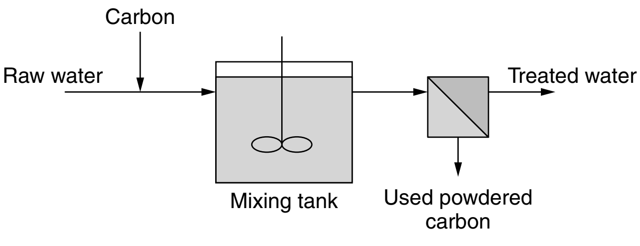

Figure 2. Powdered activated carbon for water treatment.

Similarly, the Langmuir isotherm is defined by Eqs. 3, 4, and 5:

$$\frac{X}{M} = \frac{abC}{1+bC} \tag{3}$$

$$\frac{C}{XM} = \frac{1}{ab} + \left(\frac{1}{a}\right)C \tag{4}$$

$$\left(\frac{1}{X/M}\right) = \left(\frac{1}{ab}\right)\left(\frac{1}{C}\right) + \frac{1}{a} \tag{5}$$

where a = maximum number of moles adsorbed per mass of adsorbent at monolayer saturation
b = empirical constant, L/mg

The Langmuir isotherm can be plotted in terms of $C/(X/M)$ versus C or $1/(X/M)$ versus $1/C$. The plot that produces a clear trend provides more accurate constants.

ANALYSIS AND DESIGN OF POWDERED ACTIVATED CARBON CONTACTOR

For a powdered activated carbon (PAC) application, isotherm adsorption data can be used in conjunction with a material mass balance to obtain an approximate estimate of the amount of carbon that must be added, as illustrated below. Here again, because of the many unknown factors involved, column and bench tests are recommended to develop the necessary design data. If the mass balance is written around the contactor (i.e., batch reactor) after equilibrium has been reached, the resulting expression is

$$\text{Amount absorbed} = \text{initial amount of absorbate present} - \text{final amount of absorbent present}$$

$$q_e m = VC_o - VC_e, \tag{6}$$

where q_e = adsorbate phase concentration after equilibrium, mg adsorbate/g adsorbent
m = mass of adsorbent, g
V = volume of liquid in the reactor, L
C_o = initial concentration of the adsorbate, mg/L
C_e = final equilibrium concentration of the adsorbate after absorption

If Eq. 6 is solved for q_e, the following expression is obtained:

$$q_e = \frac{V(C_o - C_e)}{m} \tag{7}$$

The equation can be rewritten as

$$\frac{V}{m} = \frac{q_e}{C_o - C_e} \tag{8}$$

The term V/m, defined as the specific volume, represents the volume of liquid that can be treated with a given amount of carbon. The reciprocal of the specific volume corresponds to the dose of adsorbent that must be used.

APPLICATION IN WASTEWATER TREATMENT

A proprietary process, “PACT,” combines the use of powdered activated carbon with the activated-sludge process. In this process, when the activated carbon is added directly to the aeration tank, biological oxidation and physical adsorption occur simultaneously. A feature of this process is that it can be integrated into an existing activated sludge system at nominal capital cost. The addition of powdered activated carbon has several advantages, including (1) system stability during shock loads, (2) reduction of refractory priority pollutants, (3) color and ammonia removal, and (4) improved sludge settleability. In some industrial waste applications, where nitrification is inhibited by toxic organics, the application of PAC may reduce or limit this inhibition.

The use of powdered activated carbon in secondary wastewater systems results in an inseparable mixture of biological solids and carbon. Thermal regeneration of the carbon also results in destruction of biomass, eliminating the need for other sludge processing and disposal techniques, but increases the size of the carbon regeneration system.

The dosage of powdered activated carbon and the mixed liquor–powdered activated carbon suspended solids concentration are related to solids retention time as follows:

$$X_\mathrm{p} = \frac{X_i \mathrm{SRT}}{\tau} \tag{9}$$

where X_p = equilibrium powdered activated carbon–MLSS content, mg/L
X_i = powdered activated carbon dosage, mg/L
SRT = solids retention time, days
τ = hydraulic retention time, days

Carbon dosages typically range from 20 to 200 mg/L. At higher SRT values, the organic removal per unit of carbon is enhanced, thereby improving process efficiency. Reasons cited for this phenomenon include (1) additional biodegradation due to decreased toxicity, (2) degradation of normally degradable substances due to increased exposure time to the biomass through adsorption on the carbon, and (3) replacement of low molecular weight compounds with high molecular weight compounds, resulting in improved adsorption efficiency and lower toxicity (Fig. 3).

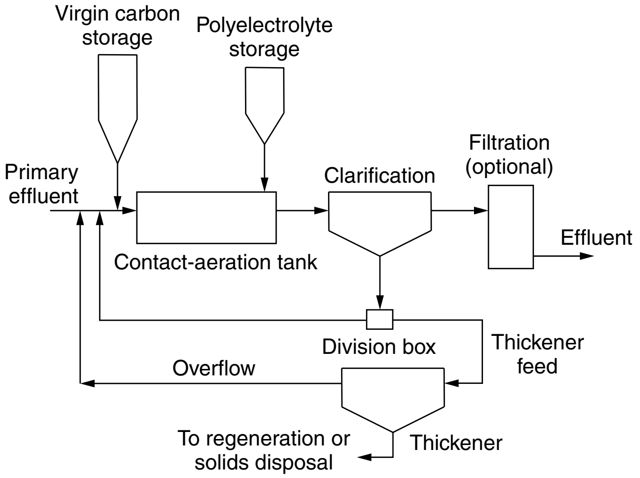

Figure 3. Schematic for the application of powdered activated carbon.

READING LIST

Metcalf & Eddy, Inc. (2003). *Wastewater Engineering: Treatment, Disposal, Reuse*, 4th Edn. McGraw-Hill, New York, pp. 1138–1162.

Peavy, H.S., Rowe, D.R., and Tchbanoglous, G. (1985). *Environmental Engineering*, International ed. McGraw-Hill, New York, pp. 195–197.

Qasim, S.R, Motley, E.M., and Zhu, G. (2000). *Water Works Engineering, Planning, Design & Operation*. Prentice-Hall, Upper Saddle River, NJ, pp. 437–463.

CHLORINATION

Absar A. Kazmi
Nishihara Environment
Technology
Tokyo, Japan

INTRODUCTION

Chlorine is added to water to kill disease-causing bacteria, parasites, and other organisms. At low concentrations, chlorine probably kills microorganisms by penetrating the cell and reacting with the enzymes and protoplasm. At higher concentrations, oxidation of the cell wall destroys the organism. Chlorination also removes soluble iron, manganese, and hydrogen sulfide from water.

CHLORINE COMPOUNDS

The principal chlorine compounds used in water treatment are chlorine (Cl_2), sodium hypochlorite (NaOCl), and calcium hypochlorite $Ca(OCl)_2$. Chlorine gas can be liquefied by compression and shipped to the site in compact containers. Because it can be regasified easily and has a solubility of approximately 700 mg/L in water at the pH and temperatures generally found in water treatment plants, this form of chlorine is usually preferred. Hypochlorites tend to raise the pH, thus driving the reaction more toward the less effective OCl^-. Commercially available calcium hypochlorite contains approximately 70–80% available chlorine, whereas NaOCl contains only 3–15% available chlorine. Some practical difficulty is involved in dissolving $Ca(OCl)_2$, and both hypochlorites are more expensive on an equivalence basis than liquefied Cl_2. There are other considerations, however, that sometimes dictate using hypochlorites.

CHLORINATION CHEMISTRY

Chlorine may be applied to water in gaseous form (Cl_2) or as an ionized product of solids [$Ca(OCl)_2$, NaOCl]. The reactions in water are as follows:

$$\mathrm{Cl_2 + H_2O \longrightarrow H^+ + HOCl} \tag{1}$$

$$\mathrm{Ca(OCl)_2 \longrightarrow Ca^{2+} + 2OCl^-} \tag{2}$$

$$\mathrm{NaOCl \longrightarrow Na + OCl^-} \tag{3}$$

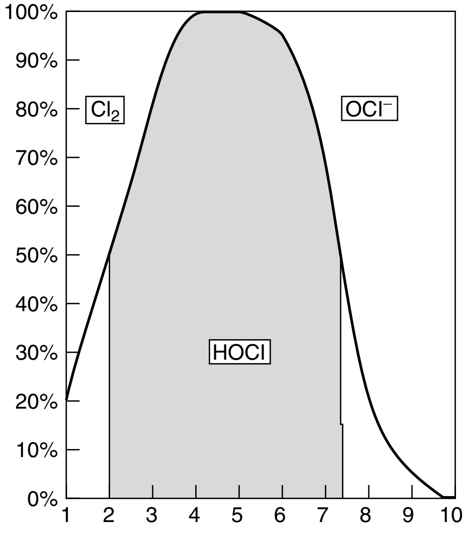

Figure 1.

The hypochlorite (HOCl) and the hypochlorite ion (OCl^-) in these equations are further related by

$$\mathrm{HOCl} \rightleftharpoons \mathrm{H^+ + OCl^-} \tag{4}$$

A relationship governed primarily by pH is shown in Fig. 1.

The sum of HOCl and OCl^- is called the free chlorine residual and is the primary disinfectant used. HOCl is a more effective disinfectant. As indicated in the previous equations, HOCl is produced on a one-to-one basis by the addition of Cl_2 gas, along with the reduction of pH, which limits the conversion of OCl^-. Cell membranes are rapidly penetrated by HOCl because lipid structures accept nonpolar materials. Passage of OCl^- is much slower, and this mitigates its killing. Although a low pH is best for disinfection by chlorine, the pH is often adjusted upward to reduce the chlorine-like odor and because swimmers suffer eye irritation when the concentration of HOCl is high.

AFFECTING FACTORS

Factors affecting the process are:

1. *Form of Chlorine.* Hypochlorous acid is more effective than the hypochlorite ion by approximately two orders of magnitude.
2. *pH.* Because the free-chlorine species are related to pH, one would expect a relationship between efficiency and pH. Empirically, it has been found that chlorine dosages must be increased to compensate for higher pH.
3. *Concentration, Contact Time.* The chlorine concentration and contact time relationship is often expressed by

$$C^n t_p = k, \tag{5}$$

C = concentration of chlorine, mg/L
t_p = time required for given percent kill, min
n, k = experimentally derived constants for a given system

4. *Temperature.* The effects of temperature variations can be modeled by the following equation derived from the van't Hoff–Arrhenius equation:

$$\ln \frac{t_1}{t_2} = \frac{E'(T_2 - T_1)}{R_{T_1,T_2}} \tag{6}$$

where t_1, t_2 = time required for given kills
T_1, T_2 = temperature corresponding to t_1 and t_2, K
R = gas constant, 1.0 cal/K·mol
E' = activation energy related to pH

CHLORINATION PROCESS

Chlorine is a strong oxidant and reacts with almost any material that is in a reduced state. In water, this usually consists of Fe^{2+}, Mn^{2+}, H_2S, and organics. Ammonia (NH_3) is sometimes present in small quantities or may be added for purposes to be presently discussed. These oxidizable materials consume chlorine before it has a chance to act as a disinfectant. The amount of chlorine required for this purpose must be determined experimentally because the exact nature and the quantity of oxidizable material in water are seldom known. A typical titration curve is shown in Fig. 2.

The products of organics oxidized by chlorine are undesirable. Organic acids (humic, fulvic) form chlorinated hydrocarbons or trihalomethanes (THMs) that are suspected of being carcinogenic. Because THMs are very seldom associated with groundwater, they are a primary concern where surface water is used. Lifetime consumption of water with THMs at a level greater than 0.10 mg/L is considered a potential cause of cancer by the Environmental Protection Agency. Minute quantities of phenolic compounds react with chlorine to form severe taste and odor problems. The original organics must be removed before chlorination, and undesirable compounds must be removed after chlorination, or the compounds must be prevented from forming. The compounds can be removed by adsorption onto activated carbon, or their formation can be prevented by the substitution of chloramines, which do not react with organics or phenols, for free chlorine. Chloramines can be formed by adding a small quantity of ammonia to the water. The reactions of chlorine with ammonia are as follows:

$$\mathrm{NH_3 + HOCl} \longrightarrow \mathrm{NH_2Cl}\ \text{(monochloramine)} + \mathrm{H_2O} \tag{7}$$

$$\mathrm{NH_2Cl + HOCl} \longrightarrow \mathrm{NHCl_2}\ \text{(dichloramine)} + \mathrm{H_2O} \tag{8}$$

$$\mathrm{NHCl_2 + HOCl} \longrightarrow \mathrm{NCl_3}\ \text{(nitrogen dichloride)} + \mathrm{H_2O} \tag{9}$$

These reactions are dependent on several factors; the most important are pH, temperature, and reactant quantities. At a pH greater than 6.5, monochloramine is the predominant species. Combined residuals

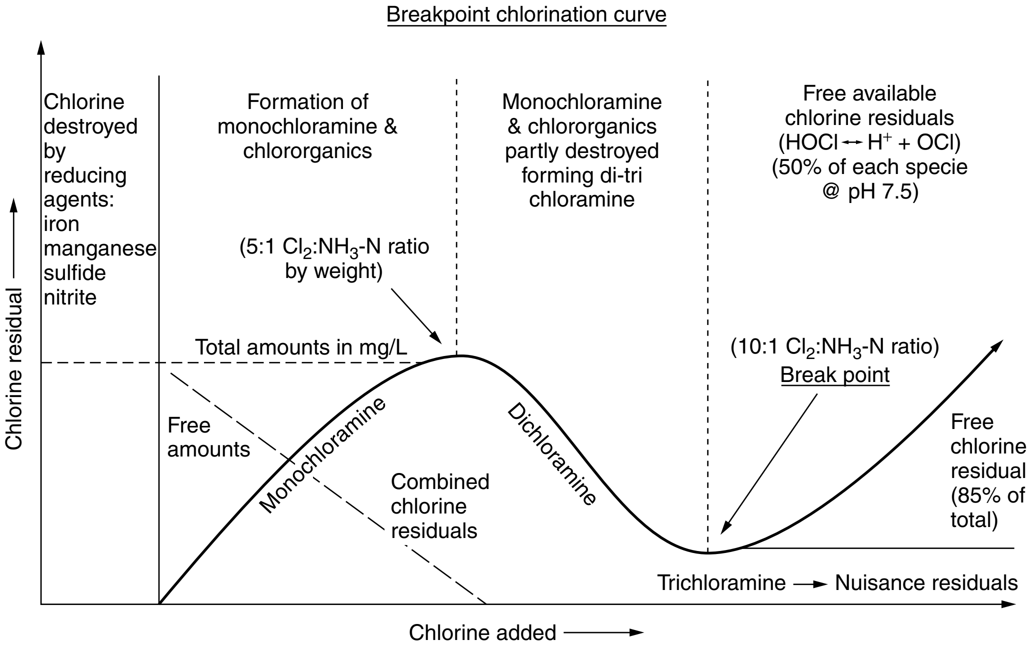

Figure 2.

are less effective as a disinfectant, so a concentration of 2–3 mg/L and contact time in excess of 30 min are often required. Chloramines are persistent and provide continued protection against regrowth in the distribution system.

APPLICATION AND DESIGN CONSIDERATIONS

Provision may be made for applying chlorine at several points within the water treatment processes. When treating raw water of good quality, no early application may be necessary, yet it is advisable to design a plant to allow for easy addition of early applications later, should future conditions require them.

Chlorine may be added to the incoming flow (prechlorination) to assist in oxidizing inorganics to arrest biological action that may produce undesirable gases in the sludge at the bottom of clarifiers. Chlorine is frequently added prior to filtration to keep algae from growing at the medium surface and to prevent large populations of bacteria from developing within the filter medium.

Mixing is one of the most important aspects of the chlorination process. The velocity gradient must be sufficient to ensure a uniform concentration of chlorine throughout the water and to break up any remaining flocculant material that might shield microorganisms from contact with chlorine. A contact chamber must be provided to ensure adequate kill time. In water treatment plant operations, mixing and contact operations may be accomplished by sectioning off part of the clear well.

CHLORINE GAS SAFETY CONSIDERATIONS

Chlorine gas is a very strong oxidant that is toxic to humans. It is heavier than air, so it spreads slowly at ground level. Therefore, extreme care must be exercised in its manufacture, shipping, and use. Accounts of evacuations of populated areas because of rail or barge accidents involving chlorine gas are common news items. The use of hypochlorite is often mandated when large quantities of chlorine are needed in treatment plants in highly populated areas.

Safe and effective application of chlorine requires specialized equipment and considerable care and skill on the part of the plant operator. Liquefied chlorine is delivered to water treatment plants in tanks containing from 75 to 1000 kg. Large plants may be designed to allow using chlorine directly from a tank car. In such cases, designers should be aware of the Interstate Commerce Commission (ICC) and Occupational Health and Safety Agency (OSHA) regulations for shipping and handling chlorine.

Safety considerations mandate storing chlorine tanks in a separate room. Storage and operating rooms should not be directly connected to other enclosed areas of the treatment plant. All doors to these facilities should be open to the outside, and windows should be provided for visual inspection from the outside. Safety equipment, including masks with air tanks, chlorine detection devices, and emergency repair equipment, should be provided in strategic locations.

READING LIST

Metcalf & Eddy, Inc. (2003). *Wastewater Engineering: Treatment, Disposal, Reuse*, 4th Edn. McGraw-Hill, New York, pp. 1218–1286.

Peavy, H.S, Rowe, D.R., and Tchbanoglous, G. (1985). *Environmental Engineering*, International ed. McGraw-Hill, New York, pp. 182–190.

CHLORINATION BY-PRODUCTS

J. Michael Wright
Harvard School of Public Health
Boston, Massachusetts

Raghuraman Venkatapathy
Oak Ridge Institute for Science and Education
Cincinnati, Ohio

June M. Weintraub
City and County of San Francisco Department of Public Health
San Francisco, California

BACKGROUND

Disinfection of drinking water is considered one of the most successful public health interventions, effectively eliminating many waterborne diseases (e.g., cholera) from many parts of the world. In 1908, chlorine was the first disinfectant to be used on a continuous basis to treat water in the United States (1). Chlorine can be used for primary or secondary disinfection via gaseous chlorine, sodium hypochlorite, or calcium hypochlorite. According to an American Water Works Association (AWWA) survey of medium (10,000–100,000 people served) and large (>100,000 people served) water systems in the United States, chlorine gas was used for disinfection by 84% of the surveyed utilities (2). Chlorination was also the predominant choice for small systems (<10,000 people served) surveyed by the AWWA, with all of the groundwater users relying on chlorine gas or sodium/calcium hypochlorite for disinfection (3). Chlorine is the primary disinfectant used for water treatment worldwide, but chlorine dioxide and ozone are also common in many European countries (see Alternative Disinfectants chapter).

Chlorination is a well-defined and fairly cost-effective treatment process. Chlorine provides adequate residual for treatment throughout the distribution system to combat bacteria, viruses, and some protozoa (e.g., *Giardia*) but may not efficiently inactivate resistant pathogens, such as *Cryptosporidium* (4). Gaseous chlorine is explosive and can present safety concerns, so many U.S. systems are switching to less hazardous types of application (5). An additional disadvantage of chlorine disinfection is the formation of potentially hazardous disinfection by-products (DBPs).

DBP FORMATION AND OCCURRENCE

In 1974, the trihalomethanes (THMs) were the first DBPs to be identified (6,7). DBPs are formed when organic or inorganic matter present in water combines with disinfectants such as chlorine. DBP formation can vary seasonally and is influenced by numerous water quality and treatment parameters, including contact time, pH, temperature, natural organic matter, chlorine dose and residual, and bromide levels (8). As shown in Table 1, different water treatment processes will influence the type and amount of DBPs formed. Recent analytical improvements have led to the identification of over 600 DBPs in laboratory scale studies and another 250 DBPs in distribution system samples (17). While numerous halogenated DBPs have been identified, they account for less than half of the total organic halide concentrations detected in chlorinated water (18).

The main chlorination DBPs include halogenated species such as the THMs, haloacetic acids (HAAs), haloacetonitriles (HANs), haloacetaldehydes (HAs), haloketones (HKs), and halonitromethanes (HNMs). The

Table 1. By-products Formed During Disinfection of Drinking Water

Disinfectant	Organohalogenated DBPs[a]	Inorganic DBPs	Nonhalogenated DBPs
Chlorine	THMs,[b] HAAs,[b] HANs, HAs, HNMs, HKs, NDMA, bromohydrins, chlorophenols, MX and brominated MX analogues	Chlorate, iodate	Aldehydes, cyanoalkanoic acids, alkanoic acids, carboxylic acids, benzene
Chloramines	THMs[b], HAAs,[b] HAs, NDMA, HKs, MX, cyanogen chloride, organic chloramines	Chlorate, nitrate, nitrite, hydrazine	Aldehydes, ketones
Chlorine dioxide	HAAs, HANs, HAs, HKs, MX	Chlorate, chlorite	Carboxylic acids
Ozone	Bromoform, monobromoacetic acid, dibromoacetic acid, dibromacetone, cyanogen bromide	Bromate, chlorate, iodate, ozonates, hydrogen peroxide, hypobromous acid, epoxides	Aldehydes, ketones, ketoacids, carboxylic acids, iodic acid, aldo acids

[a] Abbreviations: DBPs, disinfection by-products; THMs, trihalomethanes; HAAs, haloacetic acids; HAs, haloacetaldehydes; HANs, haloacetonitriles; HNMs, halonitromethanes; HKs, haloketones; NDMA, *N*-nitrosodimethylamine; MX, 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5*H*)-furanone.
[b] Includes the iodinated DBPs.
Source: References 9–16.

two most predominant groups of DBPs include the THMs and the HAAs (19). Chloroform (TCM) is typically the most prevalent THM found in chlorinated water, although the brominated THMs can predominate in bromine-rich water. Dichloroacetic acid (DCAA) and trichloroacetic acid (TCAA) are the most prevalent of the nine HAAs typically measured in drinking water.

THM and HAA concentrations found in U.S. finished water are usually less than 0.1 mg/L, while the HANs, HKs, HAs, and HNMs rarely exceed 0.01 mg/L (15,20,21). 3-Chloro-4-(dichloromethyl)-5-hydroxy-2(5*H*)-furanone (MX) is typically found at an order of magnitude less than most other halogenated DBPs (22,23), but recent findings in the United States and Russia have detected MX levels exceeding 300 ng/L (15,24). Other chlorinated DBPs such as *N*-nitrosodimethylamine (NDMA) are found in chlorinated and chloraminated water at concentrations usually below 10 ng/L (25,26).

DISINFECTION AND DBP REGULATIONS

As part of the Stage 1 Disinfection By-products Rule (DBPR), the U.S. EPA (27) established maximum residual disinfectant levels in water distribution systems for chlorine (4 mg/L as Cl_2), chloramines (4 mg/L as Cl_2), and chlorine dioxide (0.8 mg/L). Stage 1 also established maximum contaminant levels (MCLs) for bromate (0.010 mg/L), chlorite (1.0 mg/L), and HAA5 [sum of monochloroacetic acid (MCAA), DCAA, TCAA, monobromoacetic acid, and dibromoacetic acid] (0.060 mg/L) and reduced the THM4 [sum of chloroform (TCM), bromodichloromethane (BDCM), dibromochloromethane (DBCM), and bromoform (TBM)] MCL to 0.080 mg/L. The Stage 2 DBPR proposes to change the monitoring requirement for DBPs from a running annual average to a locational running annual average (28).

Members of the European Union are required to regulate total THMs at 0.1 mg/L and bromate at 0.25 mg/L, based on the Council of the European Union directive from 1998. As shown in Table 2, member countries have the discretion to establish other standards or set lower maximum acceptable concentrations (MACs) for individual DBPs. Canada has a similar MAC for THM4 and has recently proposed a MAC of 0.016 mg/L for BDCM (31). Australia has nonenforceable national guidelines for several DBPs, including 0.25 mg/L for THM4 (32). The World Health Organization (16) drinking water guidelines include values for TCM (0.20 mg/L), BDCM (0.06 mg/L), DBCM (0.10 mg/L), TBM (0.10 mg/L), dibromoacetonitrile (0.07 mg/L), MCAA (0.02 mg/L), TCAA (0.20 mg/L), cyanogen chloride (0.07 mg/L), formaldehyde (0.90 mg/L), and trichlorophenol-2,4,6 (0.20 mg/L). WHO provisional guideline values include DCAA (0.05 mg/L), bromate (0.01 mg/L), chloral hydrate (0.01 mg/L), dichloroacetonitrile (0.02 mg/L), chlorite (0.70 mg/L), and chlorate (0.70 mg/L).

Table 2. Standards for Disinfection By-products in Drinking Water

	Disinfection By-product Standards,[a] mg/L				
Country	THM4	BDCM	HAA5	Bromate	Chlorite
United States	0.08		0.06	0.01	1.00
Canada	0.10	0.016[b]			
United Kingdom	0.10			0.01	
Germany	0.01				0.20
Italy	0.03			0.25	0.80
Australia[c]	0.25				
WHO[c]		0.06		0.01	0.70[d]

[a] Abbreviations: THM4, sum of bromodichloromethane, dibromochloromethane, bromoform, chloroform; BDCM, bromodichloromethane; HAA5, sum of monochloroacetic acid, dichloroacetic acid, trichloroacetic acid, monobromoacetic acid, dibromoacetic acid; WHO, World Health Organization.

[b] Proposed guideline.

[c] Nonenforceable guidelines.

[d] Provisional guideline.

Source: References 16, 28–33.

HEALTH EFFECTS OF CHLORINE AND CHLORINATION BY-PRODUCTS

Existing toxicologic data on chlorine carcinogenicity is equivocal. No dose–response relationships have been observed, but positive findings for lymphoma and leukemia were detected among female rats (34,35). Chloroform is the most studied DBP and was the first to be recognized as a carcinogen in 1976. Liver and kidney cancer in mice and rats and central nervous system depression in mice have been reported following administration of high doses of chloroform (36,37). Similar effects have been noted in animals following exposure to BDCM (37–40). Other DBPs reported to be carcinogenic in animals include DBCM, TBM, MX, NDMA, the HANs, and the HAAs (41–45). Despite exposure assessment limitations, epidemiologic studies have also reported associations between DBP exposure and bladder cancer (46,47).

Much of the ongoing DBP research is focused on potential reproductive and developmental toxicity. Previous studies showed teratogenic effects in animals following administration of HAAs and HANs (48–53). Full litter resorptions were reported in rats following dosing with BDCM and bromoform, and sperm abnormalities have also been reported in male rats following ingestion of BDCM in water (54,55). Klinefelter and colleagues (56) and Tyl (57) have reviewed most of the existing reproductive and developmental toxicity data. Associations between DBPs and a variety of reproductive and developmental outcomes reported in the epidemiologic literature have also been reviewed in detail elsewhere (58–61). Overall, the weight of evidence for reproductive and developmental effects of DBPs is mixed for many endpoints. Toxicological and epidemiological studies examining various effects are ongoing and may further elucidate the relationship between DBP exposures and potential health effects.

ROUTES OF EXPOSURE

The main route of exposure for nonvolatile chlorinated DBPs (e.g., HAAs) is from ingestion of water or water-based beverages. Exposure to volatile DBPs, such as the THMs, can occur through inhalation, dermal absorption, or ingestion. Since inhalation and dermal exposures to volatile DBPs may be greater than from ingestion, activities such as bathing, showering, and swimming need to be considered when determining overall exposure levels (62). Point-of-use filtration devices can be used to minimize DBP ingestion exposures (24). Boiling water prior to consumption is also effective at removing many DBPs, especially the volatile compounds (63,64). In addition to point-of-use treatment, DBPs can be reduced from a variety of changes in water treatment processes or from source water protection efforts (see Alternative Disinfectants chapter). Changes in water treatment implemented to reduce DBPs need to be carefully balanced with protection against microbial pathogens, so that drinking water supplies are not compromised.

BIBLIOGRAPHY

1. Wigle, D.T. (1998). Safe drinking water: a public health challenge. *Chronic Dis. Can.* **19**(3): 103–107.
2. American Water Works Association (2000). Committee report: disinfection at large and medium-sized systems. *J. Am. Water Works Assoc.* **92**(5): 32–43.
3. American Water Works Association (2000). Committee report: disinfection at small systems. *J. Am. Water Works Assoc.* **92**(5): 24–31.
4. Korich, D.G. et al. (1990). Effects of ozone, chlorine dioxide, chlorine, and monochloramine on *Cryptosporidium parvum* oocyst viability. *Appl. Environ. Microbiol.* **56**(5): 1423–1428.
5. Weinberg, H.S., Unnam, V.P., and Delcomyn, C.D. (2003). Bromate in chlorinated drinking waters: occurrence and implications for future regulation. *Environ. Sci. Technol.* **37**(14): 3104–3110.
6. Rook, J.J. (1974). Formation of haloforms during chlorination of natural waters. *Water Treat. Exam.* **23**(2): 234–243.
7. Bellar, T.A., Lichtenberg, J.J., and Kroner, R.C. (1974). The occurrence of organohalides in chlorinated drinking water. *J. Am. Water Works Assoc.* **66**(22): 703–706.
8. Singer, P.C. (1994). Control of disinfection by-products in drinking water. *J. Environ. Eng.* **4**: 727–744.
9. Andrews, R.C. et al. (1990). Occurrence of the mutagenic compound MX in drinking water and its removal by activated carbon. *Environ. Technol.* **11**: 685–694.
10. Bichsel, Y. and von Gunten, U. (1999). Oxidation of iodide and hypoiodous acid in the disinfection of natural waters. *Environ. Sci. Technol.* **33**: 4040–4045.
11. Plewa, M.J. et al. (2004). Chemical and biological characterization of newly discovered iodoacid drinking water disinfection byproducts. *Environ. Sci. Technol.* **38**(18): 4713–4722.
12. Richardson, S.D. et al. (1994). Multispectral identification of chlorine dioxide disinfection byproducts in drinking water. *Environ. Sci. Technol.* **28**(4): 592–599.
13. Richardson, S.D. et al. (1999). Identification of new ozone disinfection by-products in drinking water. *Environ. Sci. Technol.* **33**(19): 3368–3377.
14. Richardson, S.D. et al. (2003). *Environ. Sci. Technol.* **37**(17): 3782–3793.
15. Weinberg, H.S., Krasner, S.W., Richardson, S.D., and Thruston, A.D. Jr. (2002). The occurrence of disinfection by-products (DBPs) of health concern in drinking water: results of a nationwide DBP occurrence study, EPA/600/R02/068, U.S. Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA, 2002. http://www.epa.gov/athens/publications/reports/EPA_600_R02_068.pdf.
16. WHO (2004). *Guidelines for Drinking-Water Quality*, 3rd Edn. World Health Organization, Geneva. http://www.who.int/water_sanitation_health/dwq/gdwq3/en/.
17. Sadiq, R. and Rodriguez, M.J. (2004). Fuzzy synthetic evaluation of disinfection by-products—a risk-based indexing system. *J. Environ. Manage.* **73**(1): 1–13.
18. Stevens, A.A., Moore, L.A., and Miltner, R.J. (1989). Formation and control of non-trihalomethane disinfection by-products. *J. Am. Water Works Assoc.* **81**(8): 54–60.
19. Singer, P.C. (2002). Occurrence of haloacetic acids in chlorinated drinking water. *Water Supply* **2**(5–6): 487–492.
20. Krasner, S.W. et al. (1989). The occurrence of disinfection byproducts in US drinking water. *J. Am. Water Works Assoc.* **81**: 41–53.
21. Roberts, M.G., Singer, P.C., and Obolensky, A. (2002). Comparing total HAA and total THM concentrations using ICR data. *J. Am. Water Works Assoc.* **94**: 103–114.
22. Wright, J.M. et al. (2002). 3-Chloro-4-(dichloromethyl)-5-hydroxy-2(5*H*)-furanone (MX) and mutagenic activity in Massachusetts drinking water. *Environ. Health Perspect.* **110**(2): 157–164.
23. Meier, J.R. et al. (1987). Studies on the potent bacterial mutagen 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5*H*)-furanone: aqueous stability, XAD recovery and analytical determination in drinking water and in chlorinated humic acid solutions. *Mutat. Res.* **189**: 363–373.
24. Egorov, A. et al. (2003). Exposures to drinking water chlorination by-products in a Russian City. *Int. J. Hyg. Environ. Health* **206**: 1–13.
25. Najm, I. and Trussell, R.R. (2001). NDMA formation in water and wastewater. *J. Am. Water Works Assoc.* **93**(2): 92–99.
26. California Department of Health Services (2002). Studies on the occurrence of NDMA in drinking water. http://www.dhs.ca.gov/ps/ddwem/chemicals/NDMA/studies.htm.
27. U.S. Environmental Protection Agency (2001). Office of Water (4606M), EPA 816-F-01-010, May. Stage 1 Disinfectants and Disinfection Byproducts Rule: A Quick Reference Guide. http://www.epa.gov/safewater/mdbp/qrg_st1.pdf.
28. U.S. Environmental Protection Agency (2003). Office of Water (4606M), EPA 816-D-03-002, November. The Stage 2 Disinfectants and Disinfection Byproducts Rule (Stage 2 DBPR) Implementation Guidance-Draft for Comment. http://www.epa.gov/safewater/stage2/pdfs/st2_dbpr_ig_11-11-03.pdf.
29. Hydes, O. (1999). European regulations on residual disinfection. *J. Am. Water Works Assoc.* **91**(1): 70–74.
30. Collivignarelli, C. and Sorlini, S. (2004). Trihalomethane, chlorite and bromate formation in drinking water oxidation of Italian surface waters. *J. Water SRT Aqua* **53**: 159–168.
31. Health Canada (2004). Federal–Provincial–Territorial Committee on Drinking Water. Trihalomethanes in Drinking Water. August. http://www.hc-sc.gc.ca/hecs-sesc/water/pdf/trihalomethanes_drinking_water.pdf.
32. Simpson, K.L. and Hayes, K.P. (1998). Drinking water disinfection by-products: an Australian perspective. *Water Res.* **32**(5): 1522–1528.
33. Drinking Water Inspectorate (2000). Water, England and Wales: The Water Supply (Water Quality) Regulations, No. 3184. http://www.dwi.gov.uk/regs/si3184/3184.htm#.
34. Soffritti, M., Belpoggi, F., Lenzi, A., and Maltoni, C. (1997). Results of long-term carcinogenicity studies of chlorine in rats. *Ann. N.Y. Acad. Sci.* **837**: 189–208.

35. National Toxicology Program (1992). NTP toxicology and carcinogenesis studies of chlorinated water (CAS Nos. 7782-50-5 and 7681-52-9) and chloraminated water (CAS No. 10599-90-3) (deionized and charcoal-filtered) in F344/N rats and B6C3F1 mice (drinking water studies). *Natl. Toxicol. Program Tech. Rep. Ser.* **392**: 1–466.
36. National Cancer Institute (1976). Report on the carcinogenesis bioassay of chloroform. NTIS PB-264-018. National Cancer Institute, Bethesda, MD.
37. Balster, R.L. and Borzelleca, J.F. (1982). Behavioral toxicity of trihalomethane contaminants of drinking water in mice. *Environ. Health Perspect.* **46**: 127–136.
38. Tumasonis, C.F., McMartin, D.N., and Bush, B. (1985). Lifetime toxicity of chloroform and bromodichloromethane when administered over a lifetime in rats. *Ecotoxicol. Environ. Safety* **9**: 233–240.
39. Dunnick, J.K., Eustis, S.L., and Lilja, H.S. (1987). Bromodichloromethane, a trihalomethane that produces neoplasms in rodents. *Cancer Res.* **47**: 5189–5193.
40. National Toxicology Program (1986). Toxicology and carcinogenesis of bromodichloromethane in F344/N rats and B6C3F1 mice. Technical Report Series No. 321. NIH Publication No. 88–2537, CAS No. 75-27-4. U.S. Department of Public Health and Human Services, Washington, DC.
41. Bull, R.J. et al. (1985). Mutagenic and carcinogenic properties of brominated and chlorinated acetonitriles: by-products of chlorination. *Fund. Appl. Toxicol.* **5**: 1065–1074.
42. Bull, R.J. et al. (1990). Liver tumor induction in $B6C3F_1$ mice by dichloroacetate and trichloroacetate. *Toxicology* **63**: 341–359.
43. Komulainen, H. et al. (1997). Carcinogenicity of the drinking water mutagen 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5*H*)-furanone in the rat. *J. Natl. Cancer Inst.* **89**(12): 848–856.
44. National Toxicology Program (1985). Toxicology and carcinogenesis studies of chlorodibromomethane in F344/N rats and $B6C3F_1$ mice. TR 282. DHHS Publ. No. (NIH) 85–2538. National Institutes of Health, Bethesda, MD.
45. National Toxicology Program (1989). Toxicology and carcinogenesis of tribromomethane (bromoform) in F344/N rats and B6C3F1 mice. Technical Report Series No. 350. NIH Publication No. 89–2805, CAS No. 75-25-2. U.S. Department of Public Health and Human Services, Washington, DC.
46. Villanueva, C.M. et al. (2004). Disinfection byproducts and bladder cancer: a pooled analysis. *Epidemiology* **15**: 357–367.
47. Morris, R.D. et al. (1992). Chlorination, chlorination by-products, and cancer: a meta-analysis. *Am. J. Public Health* **82**: 955–963.
48. Epstein, D.L. et al. (1992). Cardiopathic effects of dichloroacetic acid in Long-Evans rat. *Teratology* **46**: 225–235.
49. Hunter, E.S. III, Roger, E.H., and Schmid, R.A. (1996). Comparative effects of haloacetic acids in whole embryo culture. *Teratology* **54**: 57–64.
50. Smith, M.K. et al. (1988). Teratogenic effects of trichloroacetonitrile in the Long-Evans rat. *Teratology* **38**: 113–120.
51. Smith, M.K., Randall, J.L., Strober, J.A., and Read, E.J. (1989). Developmental toxicity of dichloroacetonitrile: a by-product of drinking water disinfection. *Toxicology* **12**: 765–772.
52. Smith, M.K., Randall, J.L., Read, E.J., and Strober, J.A. (1989). Teratogenic activity of trichloroacetic acid. *Teratology* **10**: 445–451.
53. Smith, M.K., Randall, J.L., Read, E.J., and Strober, J.A. (1992). Developmental toxicity of dichloroacetic acid in Long-Evans rat. *Teratology* **46**: 217–223.
54. Klinefelter, G.R., Suarez, J.D., Roberts, N.L., and DeAngelo, A.B. (1995). Preliminary screening for the potential of drinking water disinfection byproducts to alter male reproduction. *Reprod. Toxicol.* **9**: 571–578.
55. Bielmeier, S.R., Best, D.S., Guidici, D.L., and Narotsky, M.G. (2001). Pregnancy loss in the rat caused by bromodichloromethane. *Toxicol. Sci.* **59**: 309–315.
56. Klinefelter, G.R., Hunter, E.S., and Narotsky, M. (2001). Reproductive and developmental toxicity associated with disinfection by-products of drinking water. In *Microbial Pathogens and Disinfection By-Products in Drinking Water: Health Effects and Management of Risks*, G.F. Craun, F.S. Hauchman, and D.E. Robinson (Eds.). ILSI Press, Washington, DC, pp. 309–323.
57. Tyl, R.W. (2000). Review of animal studies for reproductive and developmental toxicity assessment of drinking water contaminants: disinfection by-products (DBPs). RTI Project No. 07639. Research Triangle Institute, Research Triangle Park, NC.
58. Nieuwenhuijsen, M.J. et al. (2000). Chlorination disinfection by-products in water and their association with adverse reproductive outcomes: a review. *Occup. Environ. Med.* **57**(2): 73–85.
59. Graves, C.G., Matanoski, G.M., and Tardiff, R.G. (2001). Weight of evidence for an association between adverse reproductive and developmental effects and exposure to disinfection by-products: a critical review. *Regul. Toxicol. Pharmacol.* **34**: 103–124.
60. Bove, F., Shim, Y., and Zeitz, P. (2002). Drinking water contaminants and adverse pregnancy outcomes: a review. *Environ. Health Perspect.* **110**(Suppl. 1): 61–74.
61. Hwang, B.F. and Jaakkola, J.J. (2003). Water chlorination and birth defects: a systematic review and meta-analysis. *Arch. Environ. Health* **58**(2): 83–91.
62. Backer, L.C. et al. (2000). Household exposures to drinking water disinfection by-products: whole blood trihalomethane levels. *J. Expo. Anal. Environ. Epidemiol.* **10**(4): 321–326.
63. Wu, W.W., Benjamin, M.M., and Korshin, G.V. (2001). Effects of thermal treatment on halogenated disinfection by-products in drinking water. *Water Res.* **35**(15): 3545–3550.
64. Krasner, S.W. and Wright, J.M. (2005). The effect of boiling water on disinfection by-product exposure. (In review).

CLASSIFICATION AND ENVIRONMENTAL QUALITY ASSESSMENT IN AQUATIC ENVIRONMENTS

Huan Feng
Victor Onwueme
Montclair State University
Montclair, New Jersey

Lewis Schneider
North Jersey District Water
Supply Commission
Wanaque, New Jersey

Water plays an important role in our daily life. Natural water in each compartment on Earth is balanced

through a hydrological cycle. With the population increase, industrialization, and rapid urbanization, water quality has been affected and an increased demand for water use has occurred. Anthropogenic substances are discharged into our aquatic environment. As a consequence, water contamination has been found in many rivers, streams, and lakes globally. Facing our current water environmental problems, water resources management and aquatic environmental protection and restoration are important issues to us. In the United States, water quality has been improved significantly since 1972 when the Clean Water Act was enacted by the U.S. Environmental Protection Agency. In this article, we introduce the classification of waters based on the reservoir they reside in and discuss the water quality, water pollution, and aquatic environmental legislations. This article gives a general overview of water classification, quality assessment, and legislations. Greater efforts are still needed to protect the aquatic environment and improve water quality.

INTRODUCTION

Water is essential to human life. A water molecule is made up of two positively charged atoms of hydrogen and one negatively charged atom of oxygen ($2H^+ + O^{2-} = H_2O$). Each hydrogen atom of the water molecule is linked to the oxygen atom by a covalent bond. It is a polar molecule. Because of the distribution of the charges, the molecules interact with each other by forming hydrogen bonds between the molecules. Therefore, the water molecule is stable. The properties of water are determined by the structure of the water molecule. It has a high specific heat (1 calorie $^\circ C^{-1}$ g^{-1}), a high latent heat of fusion (80 calories), and evaporation (540 calories). Light can only travel a maximum of a few hundred meters through water, sound can travel thousands of kilometers through water, and water is essential to life. Pure water freezes at 0 °C and boils at 100 °C at 1 atmosphere pressure. The maximum density of freshwater is at 4 °C (for seawater it is at its freezing point −1.9 °C) at 1 atmosphere pressure. The pH ($pH = -\log_{10}[H^+]$) value of pure water is pH = 7, rainwater pH = 5.0–5.6, river water pH = 6–7 (6.8 on average), and seawater pH = 7.5–8.5 (7.8 on average).

CLASSIFICATION OF WATER AND HYDROLOGIC CYCLE

Based on their physical and chemical properties and their storage reservoirs, the waters on Earth are classified as follows (1–3):

Rainwater—Precipitation from the atmosphere when water vapor in the air condenses and water droplets grow sufficiently large.

River (Stream) Water—Precipitation that does not evaporate or infiltrate into the ground runs off the surface and flows back toward the ocean. Best measure of water volume carried by a river is discharge that is defined as the amount of water that passes a fixed point in a given amount of time. Usually, it is expressed as cubic meter per second.

Lake and Pond—Lakes are inland depressions that hold standing freshwater year-round. Ponds are generally considered as small bodies of water.

Groundwater—The second largest reservoir of freshwater. The source of groundwater is caused by infiltration, a process of water percolating through the soil and into fractures and permeable rocks. Presently, groundwater is being removed faster than it can be replenished in many areas.

Seawater—Oceans contain more than 97% of all liquid water in the world. In open ocean, seawater has approximately 35 g of salts in 1000 g of seawater; thus, salinity = 35‰.

The hydrological cycle keeps the water balance on Earth. It describes the circulation of water when it evaporates from land, water, and organisms or transpires from plants, enters the atmosphere and condenses, and then precipitates back to the Earth's surface. Water can move underground by infiltration or overland by runoff into streams, rivers, lakes, and oceans. Renewable water supplies are made up of surface runoff and infiltration into accessible freshwater aquifers. It is estimated that 1.5 billion people lack access to an adequate supply of drinking water and nearly 3 billion lack acceptable sanitation (3). Regarding the water use in each sector, on the worldwide average, agriculture consumes about 69% of total water withdrawal and industry accounts for about 25% of all water use (3).

ENVIRONMENTAL QUALITY ASSESSMENT

Water Pollution

Since industrialization, many inorganic and synthetic organic chemicals have been used to make pesticides, plastics, pharmaceuticals, pigments, and so on. They are highly persistent and toxic and tend to bioaccumulate in food chains. These characteristics are also true of many heavy metals such as mercury, lead, cadmium, and nickel. The two most important sources of toxic inorganic and organic chemicals in water are improper disposal of industrial and household wastes and runoff of pesticides from high-use areas (such as fields, roadsides, and golf courses). It was reported that more than 25% of the mass of all materials dumped into the ocean is dredged materials from ports and waterways (2,4). Industrial waste is one major waste dumped at sea. Between 1890 and 1971, for example, 1.4×10^6 m^3 solid waste has been dumped into the water of the New York Bight (5,6). In 1972 the U.S. Marine Protection, Research, and Sanctuaries Act (the Ocean Dumping Act) prohibited dumping of all types of materials into ocean waters without a permit from the EPA. In fact, the United States has stopped its ocean dumping except for uncontaminated dredged materials with a permit from the U.S. Corps of Engineers.

Any physical, biological, or chemical change in water quality that adversely affects living organisms can be considered pollution. Pollutants can be from point sources (i.e., discharge pollutants from specific locations such as factories) and non-point sources (i.e., no specific location of discharge, such as agricultural fields and feedlots). At present, at least 2.5 billion people in less developed countries lack adequate sanitation, and about half lack access to clean drinking water (7). Sewage treatment plants and pollution-control devices can greatly reduce pathogens (8,9). To remove intestinal bacteria, drinking water is generally disinfected via chlorination.

Most urban sewer systems in the United States were built in the late 1800s and early 1900s. In 1972, the Federal Clean Water Act, in an effort to make U.S. water "fishable and swimmable," mandated the upgrade of sewage treatment to the secondary treatment level by 1977 (8,9). Although the discharge flow has been increasing over the years because of a population growth and an increasing discharge of waste by industry, total suspended solids, 5-day biochemical oxygen demand (BOD), and oil and grease have decreased in the waste stream. In most areas, heavy metal concentrations in the waste stream have been reduced as a result of enhanced sewage treatment and more strict federal and state policies and criteria that prevent these contaminants from entering the waste stream (10,11). The amount of dichlorodiphenyltrichloroethane (DDT) shows a dramatic decrease with time, although it is being gradually flushed from the land. Because of their toxic effect, polychlorinated biphenyls (PCBs) are also no longer used in a large quantity at present. Nutrient concentrations in coastal waters have been changed with land practice change, and the increasing supply of nutrients results in deoxygenation in some coastal waters as the surplus plant material decays (12–14).

Water Legislation

Water legislation in the United States was designed primarily to protect the nation's water from pollution, which included the Clean Water Act (CWA) (1972), a major revision of the prior Federal Water Pollution Control Act; and the Safe Drinking Water Act (1974) to protect populations consuming water and prevent groundwater contamination; as well as other laws such as the Resource Conservation and Recovery Act (RCRA) and the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) to prevent and mitigate contaminated groundwater. "The objective of this act (CWA) is to restore and maintain the chemical, physical and biological integrity of the Nation's waters." A major concern of the CWA was to prevent discharges to navigable water bodies of the United States. The CWA did this by making it illegal to discharge to a navigable waterway without a permit. The National Pollutant Discharge Elimination System (NPDES) defines this permitting process and implements the requirements for a point source discharger to monitor conventional, toxic, and nonconventional pollutants. Reduction (conventional pollutants) or limiting (toxicants) of these effluent discharges is achieved by the discharger through pollution control technologies. Pretreatment standards for indirect dischargers are also required in the CWA. Industries discharging indirectly to a publicly owned water treatment work (POTW) through a wastewater collection system must also comply with pollutant limits. Section 404 of the CWA includes permits oversighted by the U.S. Army Corp of Engineers to discharge dredged spoils to waters of the United States. Section 404 has been interpreted to protect adjacent and nonadjacent waterbodies that covers wetlands. Water quality standards are imposed to maintain the goal of "fishable/swimable waters," which are adopted by states with prior approval from the EPA. The water quality standards include designated use(s) of the waterbody, an antidegradation policy, and numerical and narrative criteria. Non-point source pollution is a key concern for the waterways of the United States and is addressed in the 1972 amendments to the CWA, which include state primacy in land use planning and regulations to control non-point source pollution. Storm water discharges may be considered point source discharges when conveyed through storm sewers. In 1987, the CWA amendments required discharges from municipal storm sewers to meet criteria under the act.

The primary goals of the SDWA were to ensure potable water at the consumer's tap and protect groundwater from contamination. The act applies to public water systems (PWSs), those servicing over 15 service connections or at least 25 persons. PWSs are in turn divided into community and noncommunity systems. The differentiation of community versus noncommunity is that community systems are used year-round, whereas use of noncommunity systems may vary from at least 60 days to more the 6 months. To protect the public health, the SDWA establishes primary and secondary standards. The primary standards are based on specific feasible maximum contaminant levels (MCLs) and maximum contaminant level goals (MCLGs) for the future. The primary standards include microbiological monitoring, inorganics (primarily metals and nitrate/nitrite), volatile organic compounds, synthetic organic compounds and pesticides, asbestos, and radionuclides. The SDWA establishes monitoring frequencies and analytical reporting, public notification, recordkeeping, and remediation requirements and procedures for these primary standards. Secondary contaminants are based on aesthetics, which include taste, odor, and appearance of the water, and adversely affect public welfare. Recommended ranges and limits are given for secondary standards that include physical, chemical, and biological characteristics. Similar to the primary standards, secondary standards establish monitoring requirements, public notification, and remediation. State underground injection wells are also regulated by the SDWA. Sole source aquifers are protected as well as wellhead areas and defined watersheds for surface water sources. Amendments to the SDWA, particularly in 1996, included grant and loan funding, a consumer confidence report designating water quality for annual customer distribution, and the use of risk assessment and cost–benefit analysis for adding new contaminant standards.

Water Quality Today

Significant progress in water protection and quality improvement has been made since the early 1970s. Most progress is from municipal sewage treatment facilities. In the United States, the Environmental Protection Agency has focused more on watershed-level monitoring and protection since 1998. States are required to identify waters not meeting water quality goals and develop total maximum daily loads for each pollutant and each listed waterbody. However, some problems remain. Greatest impediments to achieving national goals in water quality are sediment, nutrients, and pathogens, especially from non-point source discharges. Globally, according to Cunningham et al. (3), sewage treatment in wealthier countries of Europe generally equals or surpasses the United States. In Russia, only about half of the tap water supply is safe to drink. In urban areas of South America, Africa, and Asia, 95% of all sewage is discharged untreated into rivers. Two-thirds of India's surface waters are contaminated sufficiently to be considered dangerous to human health.

Pollution control has been advocated and practiced for many years, which includes source reduction and wastewater treatment. The public has recognized that the most effective way to reduce pollution is not producing it or not releasing it into our environment. To reduce the water pollution in our living environment, the wastewater treatment technologies have been developed, which include primary treatment (i.e., physical separation of large solids from the waste stream), secondary treatment (i.e., biological degradation of dissolved organic compounds), and tertiary treatment (i.e., removal of plant nutrients from secondary effluent) (8,9). The present water remediation technologies include oxidation, reduction, neutralization, or precipitation of polluted water. Bioremediation technologies mainly use living organisms to effectively break down contaminants in polluted waters (15).

CONCLUSION

Natural water is essential to our human life and living plants and organisms on Earth. The global hydrological cycle keeps the water balance in all compartments on Earth. Water resources management is important in regulating water use, discharge, and protecting the environment. With human population growth and increasing urbanization, demand on the water supplies is becoming increasingly high. Because of industrialization and economic development, water quality in some regions is becoming a serious environmental problem. Some rivers and streams are polluted by various anthropogenic contaminants. Since the 1970s, water pollution has been realized as a serious concern and water protection has received more and more attention. Since then, water quality has been improved to varying degrees in different countries. Overall, great efforts are still needed to conservatively use our water resource and protect the aquatic environment.

BIBLIOGRAPHY

1. Drever, J.I. (1997). *The Geochemistry of Natural Waters*, 3rd Edn. Prentice-Hall, Upper Saddle River, NJ.
2. Sverdrup, K., Duxbury, A.C., and Duxbury, A.B. (2003). *An Introduction to the World's Oceans*, 7th Edn. McGraw-Hill, New York.
3. Cunningham, W., Cunningham, M.A., and Saigo, B. (2005). *Environmental Science: A Global Concern*, 8th Edn. McGraw-Hill, New York.
4. Sverdrup, K., Duxbury, A.C., and Duxbury, A.B. (2003). *An Introduction to the World's Oceans*, 7th Edn. McGraw Hill, New York.
5. Freeland, G.L. and Swift, D.J.P. (1978). *Surficial Sediments: Marine EcoSystems Analysis (MESA) Program*, MESA New York Bight Atlas Monograph 10, New York Sea Grant Institute, Albany, NY.
6. Duedall, I.W., O'Connors, H.B., Wilson, R.E., and Parker, J.H. (1979). *The Lower Bay Complex*. MESA New York Bight Atlas Monograph, 29, New York Sea Grant Institute, Albany, NY.
7. WHO/UNICEF (World Health Organization/United Nations Children's Fund). (2000). *GlobalWater Supply and Sanitation Assessment 2000 Report*. New York.
8. LaGrega, M.D., Buckingham, P.L., and Evans, J.C. (2003). *Hazardous Waste Management*, 2nd Edn. McGraw-Hill, New York.
9. Metcalf and Eddy. (2003). *Wastewater Engineering*, 4th Edn. McGraw-Hill, New York.
10. Al-Sayed, M., Al-Mershed, A., and Al-Askar, I. (1999). Evaluation and classification of water use rate policy using the Points Credit Index. *Desalination* **123**(2–3): 233–239.
11. Guerbet, M. and Jouany, J.M. (2002). Value of the SIRIS method for the classification of a series of 90 chemicals according to risk for the aquatic environment. *Environ. Impact Assessment Rev.* **22**(4): 377–391.
12. Aller, R.C. et al. (1985). Early chemical disgenesis, sediment–water solute exchange, and storage of reactive organic matter near the mouth of the Changjian, East China Sea. *Continental Shelf Res.* **4**: 227–251.
13. Gu, D., Iricanin, N., and Trefry, J.H. (1987) The geochemistry of interstitial water for a sediment core from the Indian River Lagoon, Florida. *Florida Scientist* **50**: 99–110.
14. Ha., H. and Stenstrom, M.K. (2003). Identification of land use with water quality data in stormwater using a neural network. *Water Res.* **37**(17): 4222–4230.
15. Rittmann, B.E. and McCarty, P.L. (2001). *Environmental Biotechnology: Principles and Applications*. McGraw-Hill, New York.

READING LIST

NJDEP. (2002). *N.J.A.C 7:10A Licensing of Water Supply and Wastewater Treatment System Operators*.

NJDEP. (2004). *NJSDWA N.J.A.C. 7:10*.

USEPA. (1999). *US. 40 CFR 141 National Primary Drinking Water Regulations*.

Powell, F.M. (1998). *Law and the Environment*. West Educational Publications, Indiana.

Tarazona, J.V. et al. (2000). Assessing the potential hazard of chemical substances for the terrestrial environment. Development of hazard classification criteria and quantitative environmental indicators. *Sci. Total Environ.* **247**(2–3): 151–164

Federal Water Pollution Control Act. 33 U.S.C. 2002. USEPA. pp. 1251–1387.

Calamari, D. (2002). Assessment of persistent and bioaccumulating chemicals in the aquatic environment. *Toxicology* **181/182**: 183–186.

El-Manharawy, S. and Hafez, A. (2001). Water type and guidelines for RO system design. *Desalination* **139**(1–3): 97–113.

Mukherjee, A.B., Zevenhoven, R., Brodersen, J., Hylander, L.D., and Bhattacharya, P. (2004). Mercury in waste in the European Union: Sources, disposal methods and risks. *Resources, Conservation and Recycling* **42**(2): 155–182.

COAGULATION AND FLOCCULATION IN PRACTICE

Alper Ozkan
Selcuk University
Konya, Turkey

INTRODUCTION

All waters, particularly surface waters and process wastewaters, contain both dissolved and suspended particles. Very fine particles only a few microns in diameter settle extremely slowly by gravity alone. Therefore, several methods are used to increase the size of solid particles in solid/liquid separation. The processes of coagulation and flocculation are the major techniques used to aggregate particles. Solid particles can be flocculated into relatively large lumps, called flocs, that settle out more rapidly and thus particles are separated more easily from water (1,2). Although the terms coagulation and flocculation are often used interchangeably, or the single term "flocculation" is used to describe both, they are two different processes. Fine particles can be aggregated by neutralizing the electrical charge of the interacting particles, which is "coagulation," and bridging the particles with polymolecules is called "flocculation" (3,4).

COAGULATION

The major interactive forces between suspended particles are of two kinds, attractive and repulsive. The stability of particle suspensions depends on the properties of these forces. Attractive forces arise from van der Waals forces, which are effective only at very close range, whereas repulsive forces are due to the overlap of similarly charged electrical double layers of particles. Aggregation between particles can occur if the magnitude of the interparticle repulsive force does not exceed the corresponding attractive force. Electrostatic forces are responsible mostly for repulsion, so if suspension properties are adjusted to minimize such forces, particles can be coagulated (4,5).

Most solid particles suspended in water have a negative charge; since they have the same type of surface charge, they repel each other when they come close together. Repulsive forces prevent coagulation of particles and also retard their settlement. The specific surface charge of a particle and its magnitude, measured as zeta potential, are indicative of the strength of the repulsive forces due to charge. The charge on particle surfaces can be controlled by suspension pH. At some pH, called point-of-zero charge (PZC), the particles have no net charge and are unstable with respect to coagulation. pH control is a simple method for destabilizing suspensions. However, it may be impractical if the PZC occurs in an inconvenient pH range. Therefore, coagulants are often used to eliminate the negative charge on particle surfaces (4,6). Coagulants have a charge opposite to that of the particle surfaces; this causes charge neutralization that allows particles to come into contact and start to associate in bunches. Inorganic salts, such as $CaCl_2$, $MgCl_2$, $FeCl_3$, $AlCl_3$, $FeSO_4$, and $Al_2(SO_4)_3$, are used as coagulants. However, aluminum and iron salts are the most commonly used in wastewater treatment. The ability of most inorganic salt ions to compress the electrical double layer or the magnitude of the zeta potential is related to the valence of the ionic charge itself. In addition, hydroxy complexes of these are very active; thus they adsorb strongly on solid surfaces and can reverse the sign of the zeta potential. When hydrolyzable metal ions are used in the pH range and at the concentration level where the metal hydroxide is precipitated, sweep coagulation of suspension occurs by precipitating hydroxides. This effect leads to overall destabilization of a particle suspension (4–7).

FLOCCULATION

Polymeric flocculants are used effectively to destabilize fine particle suspensions. The advantage of polymeric flocculants is their ability to produce large, stronger flocs compared to those obtained by coagulation. The flocculating action of polymeric flocculants proceeds via either "charge patch attraction" or "polymer bridging." Charge patch attraction occurs when the particle surface is negatively charged and the polymer is positively charged. On the other hand, bridging is considered a consequence of the adsorption of segments of flocculant macromolecules onto the surfaces of more than one particle. Such bridging links the particles into loose flocs, and incomplete surface coverage ensures that there is sufficient unoccupied surface on each particle for adsorption during the collisions of chain segments attached to the particles. Most flocculants are synthetic polymers based on repeating units of acrylamide and its derivatives, which may contain either a cationic or anionic charge. Both anionic and nonionic polymers can be manufactured at very high molecular weights and thus can form large, rapid-settling, good-compacting flocs. They are widely used in industrial applications (2,4,8).

The kind of initial adsorption of polymers onto a particle varies, depending on the respective charges of both polymer and particle. It may be purely electrostatic if these charges are opposite in sign. If not, then other physicochemical reactions may take place. A cationic polymer, positively charged, can adsorb on the surface of negatively charged particles via electrostatic bonding. The most likely mechanism of adsorption for nonionic polymers is through hydrogen bonding. Anionic polymers and negatively charged suspensions may also adsorb via hydrogen bonding. In suspensions to which a cation has been added, polymer adsorption often also occurs through cation bridging; that is, the cations can form an electrostatic bridge between the negatively charged particle surface and the negatively charged polymer (4,5).

Theoretically, flocculants may be used either after destabilization of the suspension by coagulation, that is, predestabilization, or without destabilization. Polymeric flocculants are more effective in destabilized fine particle suspensions and so the first option that involves predestabilization by coagulation is always better. Destabilization prior to flocculant addition can usually be accomplished by eliminating the charge barrier by pH control or by adding an inorganic salt. This permits small flocs to form; the bridging action of the polymer links them into substantially larger units (2,6,9,10).

BIBLIOGRAPHY

1. Wills, B.A. (1985). *Mineral Processing Technology*, 3rd Edn. Pergamon Press, New York.
2. Laskowski, J.S. (2000). *Aggregation of fine particles in mineral processing circuits*. *Proc. 8th Int. Miner. Process. Symp.*, Antalya, Turkey, Oct. 16–18, 2000, Balkema, Rotterdam, pp. 139–147.
3. Laskowski, J.S. (1992). In: *Colloid Chemistry in Mineral Processing*. J.S. Laskowski and J. Ralston (Eds.). Elsevier, New York, Chap. 7, pp. 225–241.
4. Klimpel, R.R. (1997). *Introduction to Chemicals Used in Particle Systems*. ERC Particle Science & Technology, Gainesville, FL.
5. Somasundaran, P. (1980). In: *Fine Particle Processing*. P. Somasundaran (Ed.). AIME, New York, Vol. 2, pp. 947–975.
6. Hogg, R. (1999). Flocculation and dewatering. *Int. J. Miner. Process.* **58**: 223–236.
7. Letterman, R.D., Amirtharajah, A., and O'Melia, C.R. (1999). In: *Water Quality and Treatment*. R.D. Letterman (Ed.). McGraw-Hill, New York, Chap. 6, pp. 6.1–6.66.
8. Hogg, R. (1984). Collision efficiency factors for polymer flocculation. *J. Colloid Interface Sci.* **102**: 232–236.
9. Kitchener, J.A. (1972). Principles of action of polymeric flocculants. *Br. Polym. J.* **4**: 217–229.
10. Hogg, R. (1992). *Process design for flocculation, clarification and thickening of fine-particle suspensions*. *18th Int. Miner. Process. Congr.*, Sydney, Australia, May 23–28, The Australasian Institute of Mining and Metallurgy, Sydney, pp. 1315–1319.

COLLOIDS AND DISSOLVED ORGANICS: ROLE IN MEMBRANE AND DEPTH FILTRATION

Gary G. Leppard
National Water Research Institute
Burlington, Ontario, Canada

Jacques Buffle
University of Geneva
Geneva, Switzerland

INTRODUCTION

A colloid is operationally defined as any particulate with a least dimension in the range of 1.0–0.001 μm; the range includes macromolecules, but not conventional dissolved substances, which tend to be smaller (1,2). For environmental scientists, aquatic colloids are being redefined as any particulate that provides a molecular milieu into and onto which chemicals can escape from bulk water and whose movement is not significantly affected by gravitational settling (3). In practice, colloids and dissolved substances are often defined by a filter operation (4,5). Living microbes, when sufficiently small, can also be considered colloids, albeit colloids capable of very complex behavior (2).

Filtration for water scientists is the separation of particulates from a liquid, achieved by passing a suspension through a porous solid. Through the application of membrane filters and/or depth filters, filtration protocols are commonly used (1) to concentrate or remove particulate matter and (2) to separate suspended particulates into size classes for subsequent biological, chemical, or physical analyses. In water treatment, one discriminates between membrane filters, which are porous membranes that concentrate particulates at or close to the upper filter surface, and depth filters. The latter are relatively inexpensive filters with high collection efficiencies and flow rates that entrap particulates within the filter interior.

Filters for use in the analysis of aquatic samples may be separated into three categories:

1. filters made of a random assembly of fine fibers;
2. filters made of an impermeable material, which has discrete cylindrical holes traversing it and which acts as a true sieve; and
3. "depth filters," which have a spongy structure.

These three categories of filters and their relationships to applicable size ranges and to standard sieves and dialysis membranes are shown in Fig. 1. The various filters in use are complementary in pore size ranges, pore size distributions, and chemical nature, thus providing versatility in minimizing sample perturbation by filters.

PRACTICAL ADVANTAGE OF BETTER UNDERSTANDING

Understanding better the roles of colloids and dissolved organics in a filtration process is particularly useful when considering the following goals:

1. understanding the processes of fouling and biofouling as they relate to water treatment filters (6,7);
2. optimizing the design of filters for improved performance in specific situations (5,6), focused on filters designed to be cleaned *in situ*; and
3. improving the quality of size fractionations for limnologists and oceanographers focused on capturing and analyzing natural microbiological consortia with minimal artifact (7).

Some specific concepts and advances are noteworthy. Biofouling of filters is an unwanted progressive deposition of microorganisms and their consequent formation of biofilms, which clog filter pores and eventually establish an additional (but different) filter at the upper surface

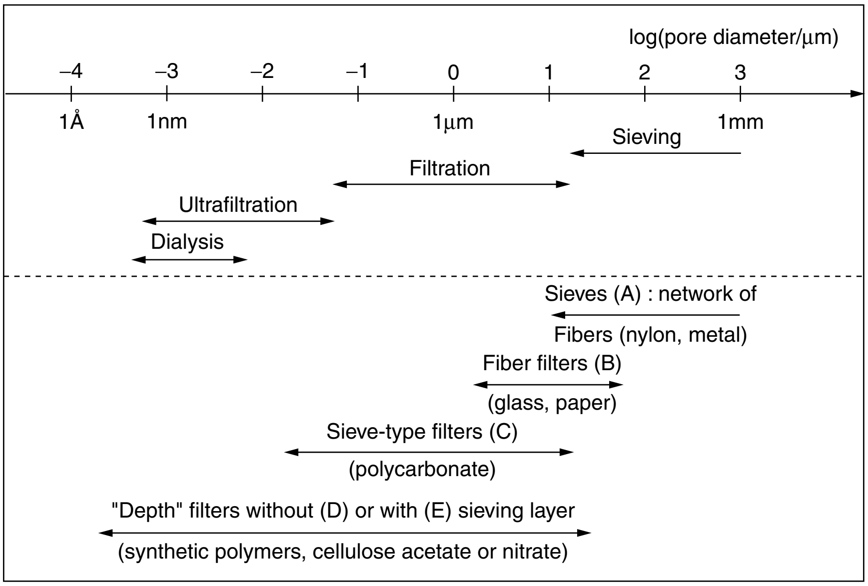

Figure 1. Size ranges for sieving and filtration, including depth filters, and for application of the most important filter types. Reproduced with permission from Buffle et al. (5).

of a membrane filter. The economic consequences of this process are considerable (6). New technology is leading to the identification of specific nanoscale organic entities that microorganisms produce to facilitate their attachment to filter surfaces (8). Such research may lead to better methods for cleaning filters *in situ*. In general, fouling may also result from purely physicochemical interactions between organic and inorganic colloids and a membrane surface, leading to colloid coagulation (5). In such cases, controls over the flow rate and physicochemical conditions are essential.

As realization increases of the many biogeochemical processes dominated by certain abundant living colloids in the oceans, the need to analyze isolated important biological species grows in importance (7). To isolate these microbes and separate them according to size (and then possibly according to dominant species) for subsequent characterization and assessment of their biogeochemical roles, there is a requirement for well-controlled filtration procedures. An additional and valuable consideration is filter selection according to filter compatibility with various techniques of analytical optical microscopy for studying fresh isolates and their minimally perturbed natural consortia while they sit on filters.

ABUNDANT COLLOIDS AND DISSOLVED ORGANICS

In natural waters, the most commonly encountered colloids are fractal aggregates of fulvic and humic acids, living cells of colloidal size, viruses, fibrils (linear aggregates of biopolymers rich in polysaccharide), refractory skeletal fragments of organisms, clay minerals, and oxyhydroxides of iron and manganese (9). Three major groups enable one to describe colloid coagulation properties: inorganic (compact) colloids, fibrillar (rather rigid) large biopolymers, and small (nanometers) fulvic acids (1). Fulvic and humic acids are at the borderline between dissolved and colloidal organics, whereas organic metabolites emanating from leaky cells, lysing cells, or excreta are usually dissolved. In engineered aquatic ecosystems, such as water treatment facilities, the same colloids and dissolved organics tend to be present but in different relative amounts. These colloids and dissolved organics can have an impact on the biological activities and clarity of both natural and engineered waters through their influences on flocculation (10,11). The impact on filtration processes of those most abundant can be considerable, leading to defective filter performance and misleading information (5).

KINDS OF FILTERS AND THEIR APPLICATION: GENERAL CONSIDERATIONS

An in-depth review of the various kinds of filters, used currently for water treatment and for the size fractionation of aquatic particles and colloids, is beyond the scope of this article. Water treatment plants and their apparatus are evolving toward greater cost-effectiveness through design changes, which are adaptations to the specific properties of the water being received. Thus, different sites are likely to have different kinds of membrane filters and use different flow rates, and changes are made continually. General considerations are found in Buffle et al. (5) and in Droppo (4). General research needs are outlined well in a 1998 Committee Report of the AWWA (12). For water treatment, there is a strong research effort to develop plastic filters with well-controlled properties, an ongoing work that is progressing well. Coupled to this are considerable interests in developing filters that can be cleaned *in situ* and improved filters for removing waterborne pathogens. Filtration can be used in conjunction with chemical aids, such as flocculants, to minimize fouling. This is a research topic that is likely to expand quickly.

BIOPHYSICOCHEMICAL INTERACTIONS BETWEEN FILTERS AND COLLOIDS/DISSOLVED ORGANICS

As stated earlier and revealed by Figs. 1 and 2, the various types of filters are complementary in properties. The nature of filtration problems, however, varies from one filter type to another, in particular because flow rates through the various membrane types differ.

Filters with a large pore size (>0.1 μm) are used for sieving and filtration (Figs. 1 and 2). A large solution flow rate can be achieved, even though a rather low pressure is applied to the filter. At a large flow rate, the concentration of colloids at the filter surface may become orders of magnitude larger than in the bulk solution (the so-called concentration polarization effect), leading to instantaneous coagulation (Fig. 3) and consequent clogging of the filter (5). Under such conditions, the nature of the membrane is not very relevant. The way to minimize clogging is to use minimum flow rates. Although very low flow rates are usually required to avoid surface coagulation, they cannot be so low as to permit bulk coagulation (Fig. 3).

Filtration membranes with pore sizes in the range 1–100 nm are used for ultrafiltration. High pressure (1–4 atm) must be applied to the membrane because of the small pore size, but flow rates are never large. When the effective hole size is of the same order of magnitude as the molecule, dialysis or reverse osmosis occurs, whereby the solvent does not move through the membrane. With such membranes and techniques, clogging due to surface coagulation is low or negligible, and it decreases with the membrane pore size. On the other hand, when molecules and pore sizes become more and more similar, interactions between the solute (the organics especially) and the membrane body become increasingly important (Fig. 3). In

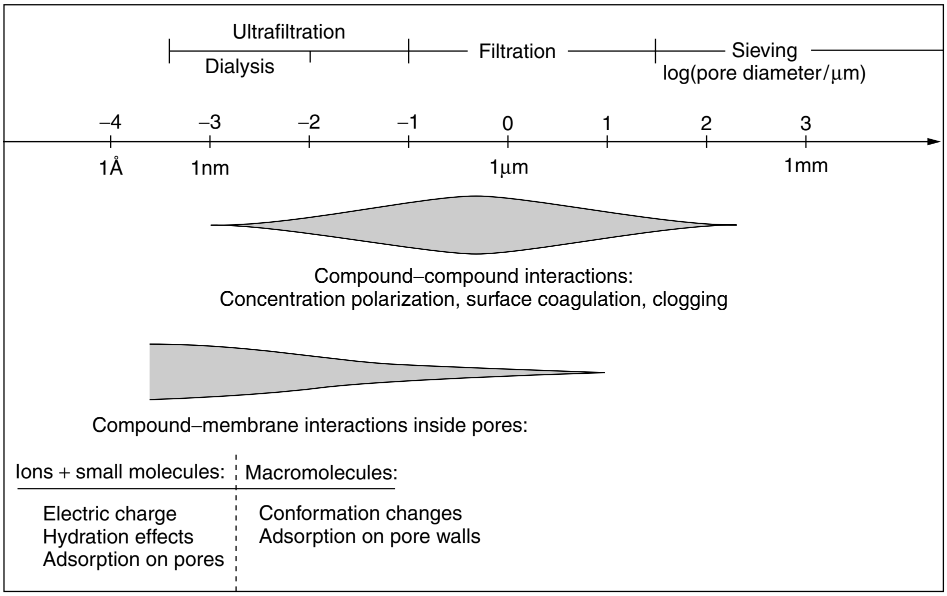

Figure 2. Schematic of size ranges where important secondary filtration effects are expected to play a significant role. The thickness of hatched zones reflects the relative importance of the corresponding factors. The word "compound" designates any component different from water, be it particulate or colloidal or dissolved. Reproduced with permission from Buffle et al. (5).

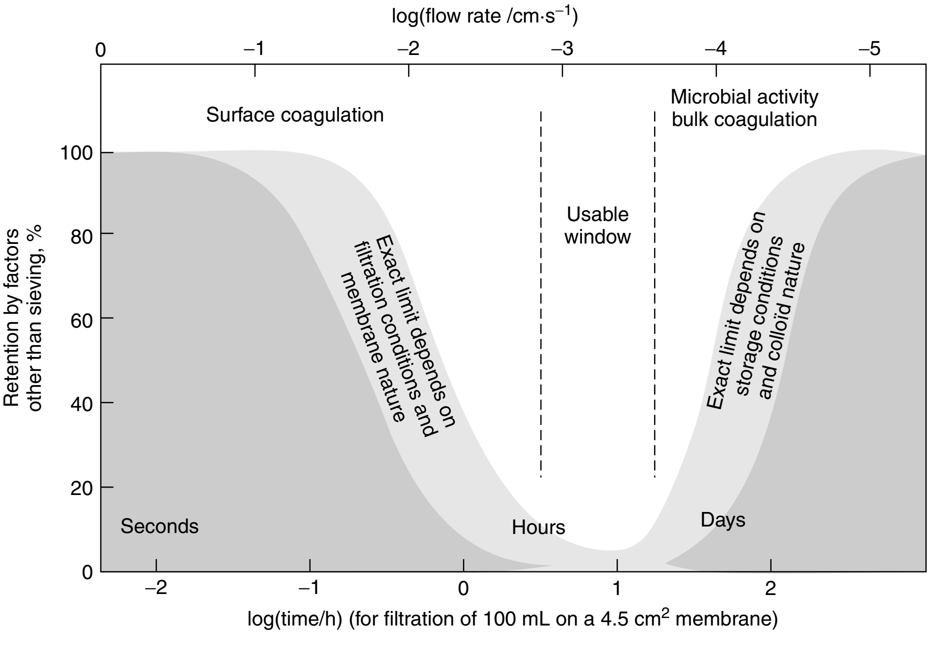

Figure 3. Semiquantitative representation of the change in retention with flow rate and with aggregation in the filtration cell, revealing the flow rate window of minimal artifact. Reproduced with permission from Leppard and Buffle (2).

such cases, the chemical nature of the membrane plays an important role. Thus it must be realized that the problems encountered in sieving and filtration are rather different from those in ultrafiltration, dialysis, and reverse osmosis. Detailed descriptions of the various types of interactions are given in Buffle et al. (5). The concept of the optimal flow rate window is illustrated in Fig. 3.

In addition to the aforementioned physicochemical interactions between membranes, suspended colloids, and dissolved organics, biological effects may play an important role. In particular, once microorganisms have been collected on the membrane filter, some of them secrete adhesive organic macromolecules (as cell attachment devices) directly onto membrane surfaces, leading to the development of a biofilm (an attached consortium of individual colloidal cells enmeshed in a porous matrix of their extracellular biopolymers). Attempts to identify specific important bioadhesives are receiving an impetus from transmission electron microscopy examinations of colloid/filter associations, using fouled membrane filters (8).

SIZE FRACTIONATION

Various sequential size fractionation procedures have been subjected to comparative analysis. In analytical applications, one can use either washing (also called diafiltration) or concentration techniques and, for each case, either sequential (also called cascade) or parallel filtration. In the concentration technique, the solution is pushed through the membrane by applying pressure. In the washing technique, a constant volume is maintained over the filtration membrane by compensating with synthetic solution. It has been found that the washing technique is more reliable because it avoids concentrating the particles in the retentate, thus minimizing coagulation problems. In a sequential filtration procedure, the same solution is filtered successfully through a series of membranes of decreasing pore size; the filtrate of one cell is filtered on the following membrane. The proportion of colloids in each size fraction is obtained by analyzing the various filtrates. In the parallel procedure, aliquots of the same sample are filtered through several membranes of different pore size. The proportions of colloids in the size fractions are obtained, based on the analysis of the filtrates and calculations. Because the reproducibility of filtration is not better than 5–10% error, the accumulation of errors becomes exceedingly large over five filtration steps in all cases, but sequential filtration is preferred because it minimizes aggregation problems. The rate of coagulation increases with particle concentration and heterogeneity.

OPTIMIZATION

The factors that affect the retention of colloids are summarized in Fig. 3. At large flow rates ($>3 \times 10^{-3}$ cm/s), large concentration polarization occurs and retention is due mostly to surface coagulation, which often leads to clogging. At very low flow rates, such surface coagulation is not important, but coagulation in the bulk solution becomes an important factor (for flow rates $<10^{-4}$ cm/s) because filtration times can be exceedingly long (days). For this reason, in chemical analysis, it is also preferable to avoid dialysis for which very long equilibration times are required. At intermediate flow rates ($10^{-4} - 3 \times 10^{-3}$ cm/s), the separation of colloids is effectively based on sieving and depends on membrane pore size.

The aforementioned limiting values of flow rates are valid for stirred conditions with a diffusion layer thickness (δ) of 10 μm. However, surface and bulk coagulation strongly depend on hydrodynamic conditions, which influence the value of δ (and thus surface coagulation), and on the bulk coagulation rate. Thus, both surface and bulk coagulation depend on the filtration mode (batch filtration with stirring, cross-flow filtration, pulsed filtration, filtration on rotary cylinders or disks, hollow fiber filtration). Decreasing δ below a few micrometers, to minimize surface coagulation, is very difficult; thus the schematic curves of Fig. 3 can be seen as the upper limits for both surface and bulk coagulation. It is interesting to note that the actual pore size of a membrane filter might not play a direct role in particle retention; the actual pore size might influence only the flow rate and thus the coagulation efficiency.

NEW VISTAS

In the last few years, novel chemical treatments have been applied to both porous and nonporous surfaces to create modified surfaces that kill bacteria on contact (13). The treatments are especially advantageous because they are not expected to exacerbate the problem of growing bacterial resistance to antibiotics. In practice, an antibacterial chemical (which disrupts drastically the integrity of the bacterial surface to kill bacteria before they can establish a biofilm) is bonded to the surface of interest so that it can neither be washed away nor modified upon interaction with oncoming bacteria. The treatments are currently applicable to glass surfaces, polymer surfaces, and carbohydrate-based porous materials. Can their use with membrane and depth filters be far behind? Considering the explanations given previously, this membrane treatment should be combined with a physicochemical means to minimize “cake” formation at the surface of filters by surface coagulation.

An interesting complementary idea is the creation of living filters in which a biofilm is purposely formed at the surface of the pores of a filter. Even though the concept needs considerable development, it can be envisioned that such a filter could be used (1) selectively to extract certain valuable substances from water for recovery or (2) selectively to degrade organic contaminants that become bound within the biofilm matrix (and thus are subject to enzymatic activities of nearby microorganisms).

Acknowledgments

The authors thank the Fonds National Suisse and the Natural Sciences and Engineering Research Council of Canada for generous long-term support.

BIBLIOGRAPHY

1. Buffle, J. et al. (1998). A generalized description of aquatic colloidal interactions: the three-colloidal component approach. *Environ. Sci. Technol.* **32**: 2887–2899.

2. Leppard, G.G. and Buffle, J. (1998). Aquatic colloids and macromolecules: effects on analysis. In: *The Encyclopedia of Environmental Analysis and Remediation*. R.A. Meyers (Ed.). John Wiley & Sons, Hoboken, NJ, Vol. 2, pp. 349–377.
3. Gustaffson, O. and Gschwend, P.M. (1997). Aquatic colloids: concepts, definitions, and current challenges. *Limnol. Oceanogr.* **42**: 519–528.
4. Droppo, I.G. (2000). Filtration in particle size analysis. In: *Encyclopedia of Analytical Chemistry*. R.A. Meyers (Ed.). John Wiley & Sons, Chichester, pp. 5397–5413.
5. Buffle, J., Perret, D., and Newman, M. (1992). The use of filtration and ultrafiltration for size fractionation of aquatic particles, colloids, and macromolecules. In: *Environmental Particles*. J. Buffle and H.P. van Leeuwen (Eds.). Lewis, Chelsea, MI, Vol. 1, pp. 171–230.
6. Flemming, H.-C., Griebe, T., and Schaule, G. (1996). Antifouling strategies in technical systems—a short review. *Water Sci. Technol.* **34**(5–6): 517–524.
7. Cole, J.J. (1999). Aquatic microbiology for ecosystem scientists: new and recycled paradigms in ecological microbiology. *Ecosystems* **2**: 215–225.
8. Liao, B.Q. et al. (2004). A review of biofouling and its control in membrane separation bioreactors. *Water Environ. Res.* (in press).
9. Leppard, G.G. (1992). Evaluation of electron microscope techniques for the description of aquatic colloids. In: *Environmental Particles*. J. Buffle and H.P. van Leeuwen (Eds.). Lewis, Chelsea, MI, Vol. 1, pp. 231–289.
10. Ferretti, R., Stoll, S., Zhang, J., and Buffle, J. (2003). Flocculation of hematite particles by a comparatively large rigid polysaccharide: schizophyllan. *J. Colloid Interface Sci.* **266**: 328–338.
11. Leppard, G.G. and Droppo, I.G. (2004). Overview of flocculation processes in freshwater ecosystems. In: *Flocculation in Natural and Engineered Environmental Systems*. I.G. Droppo, G.G. Leppard, T.G. Milligan, and S.N. Liss (Eds.). Lewis/CRC Press, Boca Raton, FL, pp. 25–46.
12. AWWA. (1998). American Water Works Association Membrane Technology Research Committee, Committee Report: Membrane Processes. *J. Am. Water Works. Assoc.* **90**(6): 91–105.
13. Borman, S. (2002). Surfaces designed to kill bacteria. *C&EN* **80**(23): 36–38.

COLUMN EXPERIMENTS IN SATURATED POROUS MEDIA STUDYING CONTAMINANT TRANSPORT

WOLFRAM SCHUESSLER
Institut für Nukleare Entsorgung
Karlsruhe, Germany

ROBERT ARTINGER
TÜV Industrie Service GmbH
München, Germany

Column experiments are one key technique to investigate transport phenomena of contaminants in water-bearing porous media. The main applications of column experiments are the determination of (1) sorption coefficients for contaminants in different porous media (e.g., aquifers), (2) transport phenomena of dissolved matter (e.g., filtration of aquatic colloids), and (3) reaction mechanisms of dissolved matter (e.g., dissociation kinetics of aquatic complexes).

The typical setup of a column experiment is given in Fig. 1. From a reservoir a fluid (e.g., groundwater or contaminated water) is pumped through a column and the effluent of the column is analyzed online or by collecting individual samples. To control the flow-through either a flowmeter or a balance is used. A second reservoir can be used to switch from one feeding solution to another [e.g., from a groundwater solution to a contaminated (spiked) groundwater]. In addition to the analysis of the effluent, the distribution of sorbing tracers or filtrated colloids in the column can be analyzed, either online or postmortem. Depending on the samples under investigation the whole experiment can be performed in a glove box, for example, for radiation protection reasons or to control the ambient air composition.

Most of the column experiments found in the literature fall into three categories: experiments with (1) intact natural porous media to the greatest possible extent, (2) artificially packed natural porous media, and (3) artificial porous media (e.g., ion exchange resins).

For the experiments in category 1 the most difficult task is to get an undisturbed (or least disturbed) column from the field site into the lab and to maintain the natural boundary conditions.

Category 2 experiments somehow represent a compromise between categories 1 and 3. The natural material can be well characterized before using it in the column experiment. During the process of column packing the natural steady state between the solid and the liquid phase is disturbed. Therefore, prior to a column experiment under “near natural” conditions the whole system, porous medium and feed solution, has to be equilibrated until a steady state is reached. This new steady state is not necessarily the one found in nature. To come as close as possible to natural conditions, parameters such as surrounding gas atmosphere or redox state have to be monitored and, if possible, controlled.

Categories 1 and 2 aim at the quantitative determination of parameters “directly” applicable to the natural systems under investigation. However, even being as close as possible to the natural conditions, one has to be aware of possible artifacts and the problems of upscaling, relevance, and transferability (1).

Category 3 experiments are used to investigate properties of dissolved matter like aquatic complexes of contaminants and aquatic colloids. Here the main goal is to identify and quantify processes under well defined conditions, which normally is not possible in a complex natural system. In these experiments well defined standard porous media (e.g., quartz sand) or porous media with special characteristics (e.g., ion exchange resins) are used.

COLUMN PREPARATION

The best preparation method strongly depends on the properties of the system under investigation. Therefore,

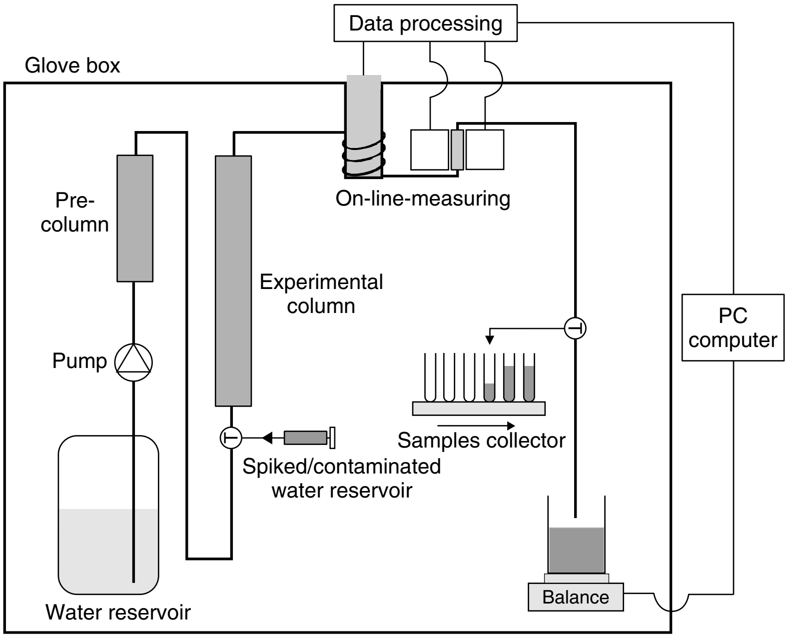

Figure 1. Experimental setup for column experiments.

only a few main hints, resulting from the long experience with column experiments of D. Klotz (2) and our own experiments with radionuclides, are given here.

The dimensions of the column are determined by the porous medium under investigation. It turned out that the column diameter should be at least 30 mm and larger than 25 mean grain diameter; the length should be larger than four times the column diameter. This setup enables evaluation of the experiments assuming an almost homogeneous distributed porous medium (not necessarily meaning homogeneous grain size) and minimizes the influence of boundary effects at the inlet and the outlet of the column.

Suitable materials for the columns are stainless steel or acrylic glass. The use of acrylic glass enables eye inspection of the readily prepared column. The disadvantage in case of experiments with natural samples is the growth of algae if the experiment is not run in the dark and at limited pressure. If the column is kept in the dark except for short periods of light (e.g., for sample taking), the growth of algae is prevented. At the inlet and the outlet of the column a frit is used in order to keep the porous medium inside the column and to distribute or sample the water over the whole cross section of the column.

As examples, two preparation methods (2) for columns are given here. In preparing a column with a cohesive medium, a slurry of one part porous medium and one part water is made. The slurry is filled into the vertical adjusted column up to a height of 1 cm. Then the water is extracted from the bottom of the column with a slow flow rate (1 cm^3/h or less). A second batch of slurry is put on top of the condensed porous medium in the column and the water is again extracted. This procedure is applied until the column is filled to the top. In the case of a sandy porous medium, the vertical adjusted column is filled with 1 cm of water and then 1 cm of the sandy medium is put into the column. The sandy medium is compacted by stamping with a bead-molding in the column and knocking with a rubber mallet from outside the column. Next, the supernatant is discarded and the process is continued unless the column is filled. This procedure guarantees a tightly packed reproducible bed inside the column.

EXPERIMENTAL PROCEDURE

In general, the flow through the column should be directed from the bottom to the top. Remaining gas bubbles originating from the packing procedure in the column are driven by the water flow and the buoyant force toward the outlet and are thus removed from the column during the equilibration process. If the column is operated from the top to the bottom, the buoyant force is always directed toward the column inlet and thus always some bubbles remain in the column, influencing the flow characteristics in an uncontrollable way.

In the case of sandy media, a pump is located at the inlet of the column and the effluent is recovered in an open sample reservoir at the outlet. For cohesive media, at the column outlet pressure is applied to extract the effluent from the column. At the outlet a sample collector can be used to allow solution analysis of individual fractions of the effluent. Flow-through can easily be monitored using a balance and a data acquisition program. In addition, online detection methods can be applied, for example, radiometric or spectroscopic methods or ion-sensitive electrodes.

Initially, the column is manipulated until a steady state in chemical composition of the effluent and hydraulic properties of the column is reached. The hydraulic properties are determined and monitored by using conservative tracers. When the steady state is reached, either a short pulse injection or a continuous step injection may be applied for a contaminated water. Which method is used depends on the system under investigation. Experiments with radioactive contaminants often use the pulse injection method to keep the activity introduced into the column as low as possible (3). In contrast, experiments with water contaminated by heavy metals frequently use the step injection method. However, the corresponding breakthrough curves of the tracers in both experiments contain the same information although they look very different at first glance (4).

A column may be used several times. However, one has to make sure that the subsequent experiments are not influenced by the preceding ones. This has to be kept in mind for the analysis of the breakthrough curves as well as for the postmortem analysis of the tracer distribution in the column.

EXAMPLES

Sorption Coefficient of Cadmium onto Sea Sand

The sorption coefficient of cadmium (Cd) onto sea sand (5) in a synthetic groundwater is determined using a plexiglas column (length 500 mm, diameter 50 mm) (Fig. 2). The column is equilibrated with the water for a period of 3 weeks. Bromide is applied as a conservative tracer for the determination of the hydraulic properties of the column. Cd breakthrough curves are determined from pulse injection (300–800 mL) experiments with initial Cd concentrations of 8.9×10^{-6} mol/L. The Cd concentration in the eluate is determined online as a function of time, that is, the Cd breakthrough curve, by means of a Cd-sensitive electrode and after collection of individual fractions by ICP-MS.

In the case where all chemical reactions involved in the experiment are in equilibrium, the sorption coefficient can be determined from the breakthrough curve according to

$$R_f = 1 + \gamma K_d V / V_p$$

where γ is the density of the packed column (here 1.7 $\text{g}\cdot\text{cm}^{-3}$), V is volume of the column (here 1000 cm^3), V_p is pore volume of the column (here 400 cm^3), and the retardation factor $R_f = v_\text{D}/v_t$ with v_D being the Darcy velocity and v_t the pore water flow velocity. Here $R_f = V_\text{max}/V_p = 8000/400 = 20$ and therefore $K_d = 4.5\ \text{g}\cdot\text{cm}^{-3}$. V_max is the solution volume eluted from the column corresponding to maximum tracer concentration in the effluent (maximum of the breakthrough curve).

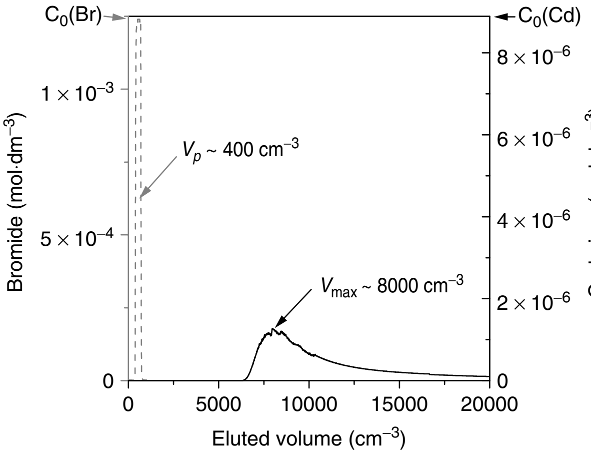

Figure 2. Column experiment to determine the sorption coefficient of Cd onto sea sand.

Transport Phenomena of Dissolved Matter, Size Exclusion of Humic Colloids

Column experiments (6) were carried out to investigate the influence of humic colloids on subsurface uranium migration (Fig. 3). The columns were packed with well characterized aeolian quartz sand and equilibrated with groundwater rich in humic colloids (DOC: 30 $\text{mg}\cdot\text{dm}^{-3}$). U migration was studied under an Ar/1% CO_2 gas atmosphere. Precolumns (250 mm long, and 50 mm in diameter) were used as filters to remove possibly generated larger particles and to pre-equilibrate the groundwater. The whole system was equilibrated with groundwater over a period of at least three months. Tritiated water (HTO) was used as a conservative tracer to determine the hydraulic properties of the column. The breakthrough curve of HTO was measured using a flow-through monitor for β-radiation; U was determined using single fraction analysis. The uranium-spiked groundwater was prepared by adding aliquots of an acidic U(VI) stock solution to the groundwater. The U concentration in the spiked groundwater was 10^{-6} $\text{mol}\cdot\text{dm}^{-3}$.

The breakthrough curves (Fig. 3) indicate that U was transported as a humic colloid-borne species with a velocity up to 5% faster than the mean groundwater flow (monitored by HTO). This could be attributed to size exclusion of the humic colloids from small water-bearing pores within the column.

Dissociation Kinetics of Am-Humic Colloids

The column experiments (3) were performed using the same groundwater/sand system as described for the U experiment (Fig. 4). The breakthrough curves of HTO and Am were measured using flow-through monitors for β and γ radiation and additionally by single fraction

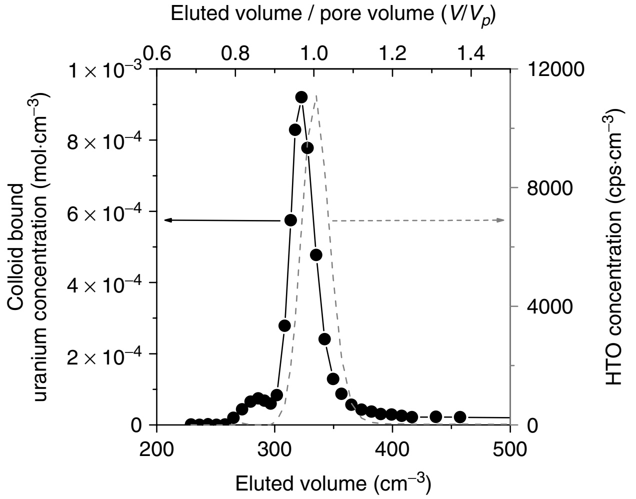

Figure 3. Column experiment investigating the size exclusion of humic colloids in a sandy aquifer.

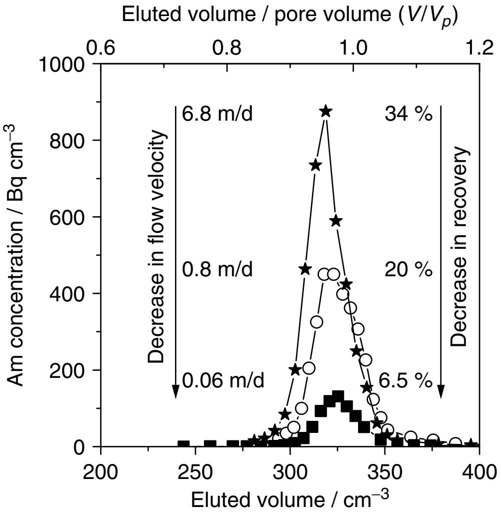

Figure 4. Column experiment investigating the dissociation kinetics of Am-humic colloids.

analysis. The ^{241}Am concentrations in eluate fractions and batch experiments were further determined by liquid scintillation counting. The Am-spiked groundwater was prepared in the same manner as described for U. The Am concentration in the spiked groundwater was 2×10^{-8} mol $\cdot$ dm^{-3}. The breakthrough curves again indicate that Am was transported as a humic colloid-borne species with a velocity up to 5% faster than the mean groundwater flow (monitored by HTO).

By varying the groundwater flow velocity about two orders of magnitude (0.06–7 m/d), an increase of the Am recovery (ΣAm injected into the column/ΣAm recovered from the column) is obtained ranging from 6% to 34% (Fig. 4). That means the longer the residence time of humic colloid-borne Am in a column, the higher is the amount of Am sorbed onto sand. The significant dependence of the Am recovery on the groundwater flow velocity is due to time-dependent turnover of Am from humic colloids onto the sand surface. These types of experiments were further used to investigate the actinide humic colloid interaction kinetics (7–10).

BIBLIOGRAPHY

1. Schüßler, W. et al. Humic colloid-borne contaminant migration: applicability of laboratory results to natural systems. *Nucl. Sci. Eng.* (in press).
2. Klotz, D. (1991). *Experiences with column experiments to determine contaminant migration (in German)*. GSF-Bericht 7/91.
3. Artinger, R., Kienzler, B., Schuessler, W., and Kim, J.I. (1998). Effects of humic substances on the ^{241}Am migration in a sandy aquifer: batch and column experiments with Gorleben groundwater/sediment systems. *J. Contam. Hydrol.* **35**: 261–275.
4. Bürgisser, C.S., Cernik, M., Borkovec, M., and Sticher, H. (1993). Determination of nonlinear adsorption isotherms from column experiments: an alternative to batch studies. *Environ. Sci. Technol.* **27**: 943–948.
5. Schüssler, W., Artinger, R., and Kienzler, B. (1998). Geochemical modelling of transport processes: surface complexation (in German). *Wiss. Ber. FZKA* **6051**: 57–76.
6. Artinger, R. et al. (2002). Humic colloid-borne migration of uranium in sand columns. *J. Contam. Hydrol.* **58**: 1–12.
7. Artinger, R. et al. (2000). Humic colloid-borne Np migration: influence of the oxidation state. *Radiochim. Acta* **88**(9–11): 609–612.
8. Artinger, R., Schüßler, W., Schäfer, T., and Kim, J.I. (2002). A kinetic study of Am(III)/humic colloid interactions. *Environ. Sci. Technol.* **36**: 4358–4363.
9. Artinger, R. et al. (2002). ^{241}Am migration in a sandy aquifer studied by long-term column experiments. *Environ. Sci. Technol.* **36**: 4818–4823.
10. Schäfer, T. et al. (2003). Colloid-borne americium migration in Gorleben groundwater: significance of iron secondary phase transformation. *Environ. Sci. Technol.* **37**: 1528–1534.

CYTOCHROME P450 MONOOXYGENASE AS AN INDICATOR OF PCB/DIOXIN-LIKE COMPOUNDS IN FISH

HANNAH AOYAGI
OLADELE A. OGUNSEITAN
University of California
Irvine, California

POLYCHLORINATED BIPHENYL AND DIOXIN-LIKE COMPOUNDS AS ENVIRONMENTAL TOXICANTS

Polychlorinated biphenyls, dioxins, and furans are ubiquitous and persistent organic pollutants (POPs). These three categories of chemicals are among the 12 priority POPs that have been scheduled for global phase-out according to the Stockholm Convention on Persistent Organic Pollutants (SCOPOP). SCOPOP was inaugurated in May 2001, and it has been ratified by more than 100 countries under the auspices of the United Nations Environment Program (1). Despite the stringent regulatory restrictions now placed on the manufacture and distribution of polychlorinated biphenyls and dioxin-like compounds, their recalcitrance in polluted aquatic environments necessitates the continuation of monitoring programs that have been implemented to reduce negative impacts on biotic and abiotic components of vulnerable ecosystems.

Polychlorinated biphenyls (PCBs) include 209 congeners consisting of two to ten chlorine atoms attached to a biphenyl structure. PCBs are highly soluble, clear, oily liquids at room temperature. They are chemically inert and thermally stable. These properties make them useful as insulators and coolants for capacitors and transformers in a variety of industries. Monsanto began production of PCBs in 1930, marketing congener mixtures under the brand name Aroclor. PCBs were also sold under the names Clophen (Germany), Kanechlor (Japan), Fenchlor (Italy), and Phenochlor (France). PCB production was banned in the United States in 1977 due to potential carcinogenicity (2,3).

PCBs tend to deposit in sediments and soils where they persist and bioaccumulate; their half-lives are in the range of months to decades. Due to their high affinity for lipids, PCBs have been detected in the adipose tissues of animals worldwide, including humans (3). Consumption of contaminated fish is the primary route of exposure for adults, whereas infants can be exposed to high concentrations of PCBs in breast milk (4).

The adverse health effects associated with PCB exposure include neurological impairment, developmental effects, carcinogenicity, hepatotoxicity, thyroid problems, dermal and ocular abnormalities, and changes in immune function (3). Occupational studies provided the first evidence of chloracne (skin eruptions) and neurological impairment, but detailed acute and chronic health effects of PCB exposure in the general population were provided by the Yusho and Yu-cheng poisoning incidents in Taiwan where 1843 adults, children, and pregnant women consumed PCB-contaminated rice oil (5,6).

The increasing recognition of human health effects due to PCB exposure has led to suspicions of their ecotoxicological impacts on various organisms inhabiting polluted ecosystems (7). Aquatic animals and their predators are at risk of cancer, neurotoxicity, reproductive toxicity, immune impairment, and endocrine effects. Due to the bioaccumulative nature of PCBs, health risks are highest for the top predators in the food chain (7). No threshold level of exposure has been identified for immune effects in animals. Moreover, it has been found that reproductive impairment can endure long after exposure in some organisms, indicating that PCBs may pose a serious threat to the fundamental survival of many species (2).

Chlorinated Dibenzofurans

CDFs are a class of chemicals composed of 135 congeners; each consists of a furan ring surrounded by two benzene rings, and one to eight chlorine atoms are attached to the benzene structures (Fig. 1). They are colorless solids, soluble in organic solvents, and do not interact with acids or alkalis. However, they are less stable than PCBs and degrade at a temperature of 700 °C (4). CDFs are not manufactured for commercial use, but they are by-products of PCB and chlorinated herbicide manufacturing, solid waste incineration, and bleaching processes. Currently there are no specific regulations pertaining to CDFs (8). They have been detected in the atmosphere and in surface waters but tend to adhere to sediments where they can bioaccumulate in fish and other aquatic organisms (4). Human exposure routes include eating contaminated fish and drinking milk from cows that have ingested feed contaminated by CDFs from atmospheric deposition. There are no data on the human health effects of CDFs alone because exposures have typically occurred simultaneously with PCBs and other similar compounds. Animal studies indicate that exposure to CDFs leads to adverse health outcomes similar to those of PCB exposure, including hepatotoxicity, thyroid alterations, and reproductive effects (8).

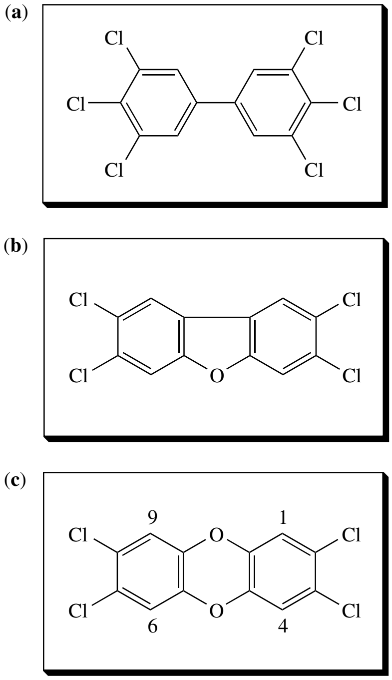

Figure 1. Schematic diagram of the structure of polychlorinated biphenyls (**a**), chlorinated dibenzofurans (**b**), and chlorinated dibenzo-*p*-dioxins (**c**) compounds.

Chlorinated Dibenzo-*p*-dioxins

CDDs comprise a group of 75 congeners; each consists of two benzene rings connected by two oxygen bridges, and one to eight chlorine atoms surround the structure (Fig. 1). CDDs are colorless solids or crystals of low solubility. They readily partition in air, water, and soil (4). CDDs were originally produced as herbicides, such as Agent Orange. In 1983, the Environmental Protection Agency (EPA) banned the use of chlorophenoxy herbicides on agricultural lands. CDDs are still emitted from incineration and chlorine bleaching processes in paper and pulp mills, as well as from the combustion of fossil fuels, wood, and cigarettes. Recent regulations have attempted to reduce paper and pulp mill effluents, as well as reducing dioxin emissions from incineration (9). CDDs are found worldwide in small concentrations, in the atmosphere (incinerator emissions), surface water (paper and pulp mill effluents), and in soils and sediments (herbicide spraying and atmospheric deposition). CDDs in soil and sediments can bioaccumulate through the food chain and reach humans through fish, meat, and milk. Occupational exposures can occur during the manufacture of pesticides and herbicides. Human and animal health effects are similar to PCBs, including cancer, immunosupression, reproductive and endocrine effects, chloracne, and neurotoxicity (9).

CYTOCHROME P450 ENZYMES

The cytochrome P450 superfamily (P450s) of microsomal hemoproteins is a set of oxygen activating terminal oxidases that catalyze several biosynthetic and metabolic processes. The monooxygenase activity of P450 enzymes is expressed through a series of electron transfer reactions involving the heme moiety (Fig. 2). The overall reaction can be summarized by the schematic equation

$$\mathrm{RH} + \mathrm{O_2} + 2\mathrm{H^+} + 2\mathrm{e^-} \longrightarrow \mathrm{ROH} + \mathrm{H_2O}$$

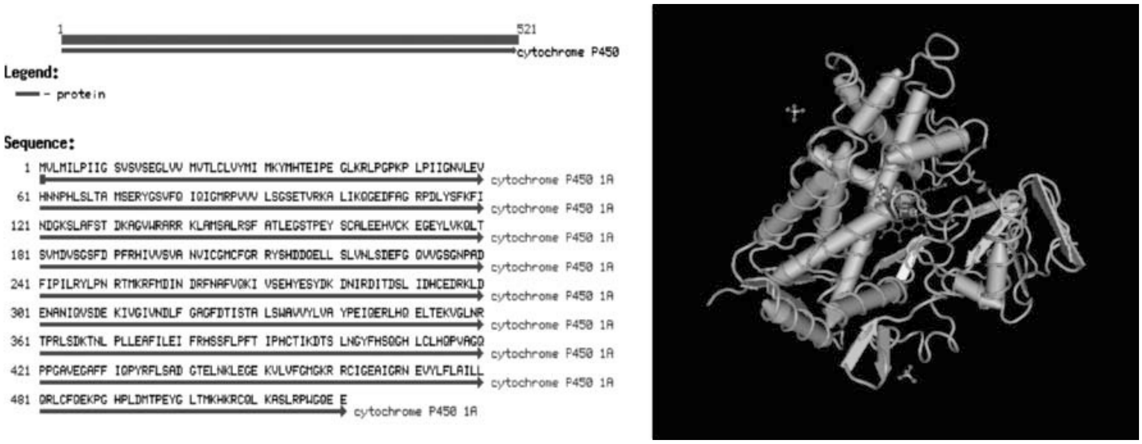

```
  1 MVLMILPIIG SVSVSEGLVV MVTLCLVYMI MKYMHTEIPE GLKRLPGPKP LPIIGNVLEV
                                                                    cytochrome P450 1A
 61 HNNPHLSLTA MSERYGSVFQ IQIGMRPVVV LSGSETVRKA LIKQGEDFAG RPDLYSFKFI
                                                                    cytochrome P450 1A
121 NDGKSLAFST DKAGVWRARR KLAMSALRSF ATLEGSTPEY SCALEEHVCK EGEYLVKQLT
                                                                    cytochrome P450 1A
181 SVMDVSGSFD PFRHIVVSVA NVICGMCFGR RYSHDDQELL SLVNLSDEFG QVVGSGNPAD
                                                                    cytochrome P450 1A
241 FIPILRYLPN RTMKRFMDIN DRFNAFVQKI VSEHYESYDK DNIRDITDSL IDHCEDRKLD
                                                                    cytochrome P450 1A
301 ENANIQVSDE KIVGIVNDLF GAGFDTISTA LSWAVVYLVA YPEIQERLHQ ELTEKVGLNR
                                                                    cytochrome P450 1A
361 TPRLSDKTNL PLLEAFILEI FRHSSFLPFT IPHCTIKDTS LNGYFHSQGH LCLHQPVAGQ
                                                                    cytochrome P450 1A
421 PPGAVEGAFF IQPYRFLSAD GTELNKLEGE KVLVFGMGKR RCIGEAIGRN EVYLFLAILL
                                                                    cytochrome P450 1A
481 QRLCFQEKPG HPLDMTPEYG LTMKHKRCQL KASLRPWGQE E
                                              cytochrome P450 1A
```

Figure 2. Amino acid sequence of salmon P4501A, and the three-dimensional structure of eukaryotic cytochrome P450 monooxygenase. The amino acid sequence was recovered from the National Center for Biotechnology Information database (http://www.ncbi.nlm.nih.gov/) (12). The image was generated from another P450 sequence imported into the Cn3D4.1 software program. The string and barrel structure shows the position of the essential iron atom.

The RH substrate represents a wide variety of compounds, including chlorinated aromatic pollutants. More than a hundred different P450 enzymes have been identified in animals, plants, fungi, and microorganisms, although they appear to play the most complex role in animals (10).

P450s perform a variety of functions, including the capacity to hydroxylate different structural types of organic molecules. They were first identified in liver tissue and later in tissues such as the adrenal gland and lymphocytes (10). In the liver, P450s catalyze the first step in metabolizing environmental toxicants as well as endogenous steroid hormones and drugs. These reactions have important implications for the use of cytochromes P450 as biomarkers of environmental toxicant exposures (12).

In the typical metabolic reaction, cytochrome P450 binds with the substrate, oxygenates it, and transforms it into a reactive state. In the second step of metabolism, the substrate undergoes a conjugation reaction (glucuronidation) and eventually elimination. In the reactive state, the substrate has the potential to form DNA adducts, thereby initiating carcinogenesis. However, ecotoxicologists are more interested in the mechanism of P450 induction and how changes in enzyme synthesis can be used as an indicator of exposure to specific environmental toxins. Cytochrome P450 assays are useful because genetic regulation of the enzyme can be induced by certain substrates (12).

There are two types of P450 induction; one requires phenobarbitone as a substrate, and the other requires carcinogenic polycyclic hydrocarbons such as PCBs, CDDs, and CDFs as a substrate. This second type of induction occurs with cytochrome P450 1A1 (CYP1A1), which is believed to have evolved at least 800 million years ago as a way for animals to metabolize plant chemicals (12). Cytochrome P450 1A1 can be induced in animals by planar molecules such as planar PCBs and CDDs; 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) is the most potent inducer known (12). These specific planar molecules bind readily to the cytosolic aryl hydrocarbon receptor (AhR) due to a similarity shared in the number of chlorine atoms present at the para and meta positions of the chemical structure. Although the main function of the AhR is unclear, it allows planar molecules to be translocated into the nucleus and induce CYP1A1 by suppressing normal genomic control of transcription (12).

Cytochrome P450 1A1 induction is a potent biomarker of ecosystem-wide exposure to PCBs and dioxin-like compounds due to its occurrence in most animal species. CYP1A1 induction is quantifiable, it indicates specific types of exposure, and its induction is manifested as a reversible subclinical effect (14). Rather than assessing the toxicological endpoints of PCB and dioxin exposure, CYP1A1 provides a measurement of alterations that affect the bioavailability and metabolism of carcinogens and other toxins (14). Sustained CYP1A1 induction can lead to an increase in the rate of metabolism and activation of carcinogens and mutagens, leading to increased formation of oxidative DNA adducts (12). Activated metabolites can also be transported to distant tissues where they cause oxidative damage and other adverse effects (15).

ECOLOGICAL IMPACTS OF PCBS AND DIOXIN-LIKE COMPOUNDS ON FISH

PCBs and dioxin-like compounds have impacted aquatic ecosystems through paper mill effluents, other industrial discharges, urban runoff, and atmospheric deposition (4). These contaminants persist in sediments and bioaccumulate, causing chronic exposure in many fish species (Table 1). Of all fish studied to date, salmonids are the most sensitive to chlorinated hydrocarbons, especially in regard to reproductive effects. They are thus useful as test organisms for the many possible physiological effects of exposure to PCBs and dioxin-like contaminants (16).

Chronic exposure to TCDD increases adult mortality in female rainbow trout (*Oncorhynchus mykiss*) in a dose-dependent fashion after a relatively long exposure of at

Table 1. Chemical Properties of PCBs, CDFs, and CDDs

Compounds	Commercial Uses and Sources into the Environment	Estimated Environmental Half-Life[a]
Polychlorinated biphenyl	Insulators, coolants in capacitors and transformers	Vapor phase: 12.9 days for monochlorobiphenyl to 1.31 years for heptachlorobiphenyl Freshwater: 6.8–9 years
Chlorinated dibenzofuran	By-product of heating PCBs, incineration, and chlorine bleaching	Surface water: 7 days–4 weeks Soil: 7 days–4 weeks Aerobic aquatic: 7 days–4 weeks Anaerobic aquatic: 28 days–16 weeks
Chlorinated dibenzo-*p*-dioxin	Herbicides, by-product of incineration and bleaching	Surface water: 1.15–1.62 years Soil: 1.15–1.62 years Aerobic aquatic: 1.15–1.62 years Anaerobic aquatic: 4.58–6.45 years

[a] References 4, 11.

least 26 weeks (16). TCDD exposure, it is also believed affects behavior in rainbow trout, making them less active and responsive (17). Reproductive outcomes are manifested in the decreased survival of young and a large number of deformities, although egg characteristics such as weight, lipid content, and total number of eggs produced do not differ from those of unexposed populations. TCDD-exposed rainbow trout young experience an unusually high incidence of yolk sac edema, a disease rarely seen in unexposed fish (17).

Chlorinated hydrocarbon exposure can also lead to AhR-mediated induction of CYP1A1 in rainbow trout and in many other species of fish, making it one of the most consistently observed outcomes in fish toxicology. CYP1A1 levels naturally fluctuate in female rainbow trout and decline as they approach their reproductive phase. It has been observed that the fluctuations are associated with the health of the fish and its environmental conditions. Although it is unclear exactly what role the enzyme plays in the life cycles of fish, it may be involved in mitigating yolk sac edema and a number of other adverse health effects, both directly and indirectly (16).

Several factors affect the intensity of cytochrome P450 induction and PCB/dioxin toxicity in fish. The potency of the compound, or mixture of compounds, determines the extent of induction and subsequent toxicity. In addition to being the most potent CYP1A1 inducer, 2,3,7,8-TCDD is considered the chlorinated polycyclic hydrocarbon most toxic to all life and thus has a toxic equivalency factor (TEF) of 1. In rainbow trout and carp, TCDD is followed in toxicity by CDFs, whose TEFs range from 0.359 to 0.028, and nonortho PCBs, whose TEFs range from 0.00016 to 0.000041 (17,18).

Lipid levels also play a role in CYP1A1 induction, toxicant bioavailability, and baseline health conditions of fish. Linoleic acid (18:2) controls hepatic cytochrome P450 concentrations, high levels of 18:2 decrease enzymatic activity (19). Chlorinated polycyclic hydrocarbons have high affinity for lipids and can be readily stored in the adipose tissues of fish. This type of storage is believed beneficial by reducing chronic exposure, although it can be harmful if the tissue is mobilized.

METHODS FOR ASSESSING CYP1A INDUCTION AND THE RATIONALE FOR INCLUDING IT IN ECOTOXICOLOGICAL ASSESSMENTS

CYP1A1 induction is an appropriate biomarker of PCB/dioxin exposure in fish because it is easily and rapidly quantified, specific, and has a temporary, reversible effect. There are various methods for assaying cytochrome P450 induction, using well-established criteria to characterize accurately the extent of induction in the liver and other tissues of interest. The level of induction may be used as a measure of potential adverse health effects in fish, without the manifestation of disease end points. Ideally, a biomarker should quickly signal adverse effects before they become serious and irreversible (14). CYP1A1 induction fits this criterion; it is not a disease end point itself, but an indicator of potential DNA damage through carcinogen activation, and may be reversed once exposure ends (15).

Induction can be examined in terms of mRNA production, protein synthesis, or enzymatic activity, and there are methods for assessing cytochrome P450 induction at each of these stages: mRNA quantification, immunologic protein determination, and assays of enzymatic activity (20). mRNA quantification is a relatively simple and rapid method compared to immunodetection and enzyme assays and may be a more useful tool for measuring aquatic pollution through cytochrome P450 induction. One process involves isolating CYP1A RNA from tissues of interest, synthesizing cDNA, designing primers to amplify specific regions of CYP1A cDNA, and quantifying using reverse transcription-polymerase chain reaction (RT-PCR) and gel electrophoresis techniques (20,21). Other studies have used slot blot analysis and RNA spectrophotometry to quantify CYP1A induction (22).

Immunodetection methods, such as Western blotting or immunohistochemistry, require identifying antibodies for specific cytochromes P450; monoclonal antibody 1-12-3 is specific to CYP1A1 and can detect CYP1A in all fishes studied thus far (23). The antibody and developer are applied to tissue samples and incubated until staining has developed. The resulting staining pattern indicates which cells have expressed CYP1A1 and to what degree. The intensity of CYP1A staining is measured on a scale of 0 to 15, where zero indicates no induction of the enzyme (23). The advantages of immunodetection methods are that they have the sensitivity to detect very small quantities of cytochrome P450 in all types of cells (15). The drawbacks are that these methods require a longer induction time than mRNA quantification, and it is uncertain whether appropriate antibodies can be found for all species of fish (20).

Enzyme assays measure the biotransformational activities of induced CYP1A on specific substrates. One type of assay examines the deethylation of 7-ethoxyresorufin (EROD) to resorufin, which can be compared to a resorufin reference (12). The dealkylation of 7-pentoxyresorufin (PROD) and UDP-glucuronosyltransferase activity (UDP-GT) are also used, and all three types of biotransformation may be quantified by using fluorescence spectrophotometry (24). Enzyme assays provide a good quantitative method for determining CYP1A induction, but the induced enzyme is subject to denaturation prior to EROD, potentially underrepresenting the extent of induction. Furthermore, certain pollutants may inhibit EROD, masking their effect on CYP1A1 induction (20).

CASE STUDIES USING P450 AS AN INDICATOR OF POLLUTANTS IN FISH

The Atlantic Tomcod in the Hudson River Estuary of the United States

The Hudson River contains aromatic hydrocarbon contaminants released from industry, municipal wastewater treatment plants, and sewage outflows. During a span of 30 years, hundreds of thousands of pounds of PCBs were released from two major sites impacting the Hudson River estuary (22). CDDs and CDFs were released during a 20-year period by a company manufacturing the herbicide 2,3,4-T, and from municipal waste incineration. Therefore, the estuarine sediments contain high levels of contaminants that have bioaccumulated in fish, including the Atlantic tomcod (22).

Yuan and co-workers (22) used Atlantic tomcod to compare hepatic CYP1A1 mRNA expression in relationship to sediment levels and hepatic burdens of chlorinated aromatic hydrocarbons and to evaluate CYP1A1 as a biomarker for mapping sediment contaminant concentrations in the Hudson River estuary. Previous studies had linked tomcod body burdens of PCBs, CDDs, and CDFs to regions of high sediment contamination and linked body burden of chlorinated aromatic hydrocarbons to CYP1A1 induction. The authors observed significant differences in CDD and CDF body burdens that varied across the estuary and sediment contamination levels. They also found that CYP1A1 induction varied significantly across sediment contamination levels. However, CYP1A1 induction did not correspond to the type of chlorinated aromatic hydrocarbon or to the TEF. The authors hypothesized that polyaromatic hydrocarbons (PAHs)—also present in the Hudson River estuary but not included in the study—were partially responsible for the observed CYP1A1 induction. Additionally, chronic exposure to PCBs, CDFs, and CDDs, it was shown, depletes AhR cells, thereby lowering the level of CYP1A1 inducibility and restricting the ability to predict exposure in fish based on enzyme levels.

Carp Fish in Hikiji River, Kanagawa Prefecture, Japan

Between 1992 and 2000, dioxin-contaminated raw sewage was released into the Hikiji River, Japan from an industrial waste incinerator (25). The toxic equivalent (TEQ) of 3.0 g of TCDD was released, and water concentrations ranged from 0.93–8.7 pg-TEQ/L in the year 2000. Agricultural runoff and municipal incinerators contributed a variety of other CDDs as well as CDFs. Sakamoto and co-workers examined the level of CYP1A induction in carp from the Hikiji River, both upstream and downstream from the effluent discharge, in conjunction with indicators of their reproductive health (25).

Muscle tissue analysis of carp from the contaminated downstream region showed higher concentrations of several different CDDs, CDFs, and coplanar PCBs. Liver microsomal CYP1A content was significantly higher in the exposed male and female carp, although EROD activity was increased only in females. Serum estrogen levels, an important factor in the initial stages of oocyte development, were low in the exposed fish and significantly correlated with CYP1A content. Reproductive health markers were also correlated to estrogen levels; exposed females produced fewer oocytes containing yolk protein, although the difference was not statistically significant. The investigators noted that estrogen poses a problem in interpreting CYP1A induction due to its interactions with the enzyme. It is known that, of all endogenous compounds, estrogen has the largest impact on CYP1A activity. High estrogen levels during fish reproductive phases are associated with low CYP1A levels. However, the study found a consistent pattern of increased CYP1A induction in exposed carp across all serum estrogen levels.

PROSPECTS FOR USING CYTOCHROME P450 AS A BIOINDICATOR

Though its mechanisms are not fully understood, CYP1A induction, it has been shown, is a very effective biomarker for fish exposure to PCBs and dioxin-like compounds. As Yuan et al. demonstrated, it is difficult to distinguish the effects of PCBs from those of CDDs, CDFs, and PAHs, and thus it is not yet possible to use CYP1A1 induction as a direct measure of one particular aquatic contaminant or class of chemicals. Nor are there adequate data on the level of response for all fish species residing in contaminated waterways. Future research must better quantify the dose–response relationship between exposure and induction across species and elucidate the various modifications of CYP1A induction by endogenous agents, fish health status, and other environmental conditions. Toward this end, the development and breeding of genetically engineered indicator fish with reporter genes such as bioluminescence or green fluorescent protein linked to the inducer locus of cytochrome P450 genes may ultimately prove indispensable.

Acknowledgments
This work was supported in part by the Program in Industrial Ecology at the University of California, Irvine, and by the UC Toxic Substances Research and Teaching Program.

BIBLIOGRAPHY

1. United Nations Environment Programme. (2003). *Stockholm Convention on Persistent Organic Pollutants*. ONLINE. Available: http://www.pops.int/ [10 July 2003].
2. Agency for Toxic Substances and Disease Registry. (2000). *Toxicological Profile for Polychlorinated Biphenyls*. ONLINE. Available: http://www.atsdr.cdc.gov/toxprofiles/tp17.html [18 June 2003].
3. U.S. Environmental Protection Agency. (2002). *Health Effects of PCBs*. ONLINE. Available: http://www.epa.gov/opptintr/pcb/effects.html [18 June 2003].
4. Food and Agricultural Organization. (2000). *Fact Sheets on Chemical Compounds*. ONLINE. Available: http://www.fao.org/DOCREP/003/X2570E/X2570E08.htm [10 July 2003].
5. Kashimoto, T. et al. (1985). PCBs, PCQs and PCDFs in blood of yusho and yu-cheng patients. *Environ. Health Perspect.* **59**(2): 73–78.
6. Hsu, S.T. et al. (1985). Discovery and epidemiology of PCB poisoning in Taiwan: a four-year follow-up. *Environ. Health Perspect.* **59**(2): 5–10.
7. Chiu, A. et al. (2000). Effects and mechanisms of PCB ecotoxicity in food chains: algae, fish, seal, polar bear. *Environ. Sci. Health* **18**: 127–152.
8. Agency for Toxic Substances and Disease Registry. (1994). *Toxicological Profile for Chlorodibenzofurans*. ONLINE. Available: http://www.atsdr.cdc.gov/toxprofiles/tp32.html [18 June 2003].
9. Agency for Toxic Substances and Disease Registry. (1998). *Toxicological Profile for Chlorinated Dibenzo-p-dioxins*. ONLINE. Available: http://www.atsdr.cdc.gov/toxprofiles/tp104.html [18 June 2003].
10. Omura, T., Ishimura, Y., and Fuji-Kuriyama, Y. (1993). *Cytochrome P-450*, 2nd Edn. Kodansha, Tokyo, pp. 6–7.
11. Howard, P.H. et al. (1991). *Handbook of Environmental Degradation Rates*. Lewis, Chelsea, MI, pp. 510, 657.
12. Ruckpaul, K. and Rein, H. (1990). *Frontiers in Biotransformation: Principles, Mechanisms and Biological Consequences of Induction*. Vol. 2, Taylor and Francis, Philadelphia, pp. 5, 6–10, 13–16, 39.
13. Arukwe, A. (2002). Complementary DNA cloning, sequence analysis and differential organ expression of beta-napthoflavone-inducible cytochrome P4501A in Atlantic salmon (*Salmo salar*). *Comp. Biochem. Physiol.* **133**: 613–624.
14. Gil, F. and Pla, A. (2001). Biomarkers as biological indicators of xenobiotic exposure. *J. Appl. Toxicol.* **21**: 245–255.
15. Pelkonen, O. and Raunio, H. (1997). Metabolic activation of toxins: tissue-specific expression and metabolism in target organs. *Environ. Health Perspect.* **105**(4): 767–774.
16. Giesy, J.P. et al. (2002). Effects of chronic dietary exposure to environmentally relevant concentrations of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin on survival, growth, reproduction and biochemical responses of female rainbow trout (*Oncorhynchus mykiss*). *Aquatic Toxicol.* **59**(1–2): 35–53.
17. Hornung, M.W., Zabel, E.W., and Peterson, R.E. (1996). Toxic equivalency factors of polybrominated dibenzo-*p*-dioxin, dibenzofuran, biphenyl, and polyhalogenated diphenyl ether congeners based on rainbow trout early life stage mortality. *Toxicol. Appl. Pharmacol.* **140**(2): 227–234.
18. Loganathan, B.G. et al. (1995). Isomer-specific determination and toxic evaluation of polychlorinated biphenyls (PCBs), polychlorinated/brominated dibenzo-*p*-dioxins and dibenzofurans, polybrominated biphenyl ethers, and extractable organic halogen in carp from the Buffalo River, New York. *Environ. Sci. Technol.* **29**(7): 1832–1838.
19. Landis, W.G. and Yu, M.H. (1999). *Introduction to Environmental Toxicology: Impacts of Chemicals upon Ecological Systems*, 2nd Edn. Lewis, Boca Raton, FL, pp. 144–145.
20. Cousinou, M. et al. (2000). New methods to use fish cytochrome P4501A to assess marine organic pollutants. *Sci. Total Environ.* **247**(2–3): 213–225.
21. Rees, C.B. et al. (2003). Quantitative PCR analysis of CYP1A induction in Atlantic salmon (*Salmo salar*). *Aquatic Toxicol.* **62**: 63–78.
22. Yuan, Z. et al. (2001). Is hepatic cytochrome P4501A1 expression predictive of hepatic burdens of dioxins, furans, and PCBs in Atlantic tomcod from the Hudson River estuary? *Aquatic Toxicol.* **54**(3–4): 217–230.
23. Moore, M.J. et al. (2003). Cytochrome P4501A expression, chemical contaminants and histopathology in roach, goby and sturgeon and chemical contaminants in sediments from the Caspian Sea, Lake Balkhash and the Ily River Delta, Kazakhstan. *Mar. Pollut. Bull.* **46**: 107–119.
24. Huuskonen, S. and Lindström-Seppä, P. (1995). Hepatic cytochrome P4501A and other biotransformation activities in perch (*Perca fluviatilis*): The effects of unbleached pulp mill effluents. *Aquatic Toxicol.* **31**(1): 27–41.
25. Sakamoto, K.Q. et al. (2003). Cytochrome P450 induction and gonadal status alteration in common carp (*Cyprinus carpio*) associated with the discharge of dioxin contaminated effluent to the Hikiji River, Kanagawa Prefecture, Japan. *Chemosphere* **51**(6): 491–500.

WATER-BASED DISEASE.

See Water Related Diseases

WATERBORNE DISEASE.

See Water Related Diseases

WATER RELATED DISEASES

Mark J. Walker
University of Nevada
Reno, Nevada

Water related diseases can be placed into four categories: (1) water-based disease, (2) waterborne disease, (3) water vectored disease, and (4) water washed disease. Infectious microbes that cause disease within these categories include bacteria, enteric viruses, parasitic protozoans, helminthes, and algae. The health effects of infection vary considerably, from symptomless to temporary mild illness to prolonged illness to death. Each category represents a distinct pathway and mechanism for infection by a microbial pathogen. Symptoms of infection, such as gastrointestinal upset, may be common across categories.

The categories represent differences in strategies that can be applied to control the prevalence of infection.

Water-based diseases are caused by microorganisms whose reproductive cycles involve passage through an aquatic animal. If the aquatic animal is not present, the microorganism cannot reproduce and, consequently, cannot infect humans. As an example, trematodes, *Shistosoma* spp. (blood flukes), cause the disease schistosomiasis. This disease occurs in China, Japan, the Philippines, Indonesia (infection with spp. *japonica*), throughout Africa (spp. *haematobium*), in Southeast Asia (spp. *mekongi*), and in western and central Africa (spp. *intercalatum*). Shistosomes must pass through a snail in the aquatic environment to develop to a form that infects people (the cercacia). The cercacia enters the bloodstream by passing through skin when people come in contact with contaminated water. The parasite is reintroduced into water as an egg contained in human feces. Control programs to prevent water-based diseases are often designed to eliminate the intermediate aquatic host, using molluscicides, snails that compete for the same habitat but do not act as intermediate hosts, or snail eating fish or birds. The most effective control programs include education about the life cycle, including proper disposal of human wastes.

Waterborne diseases are part of the fecal–oral cycle of infection. This type of disease may be transmitted by many routes, including water contaminated by feces from an infected host. Microorganisms, in a life cycle stage that may infect a new host, are excreted and enter water that is then consumed. Common waterborne microbial pathogens include protozoa (such as *Cryptosporidium* and *Giardia*), viruses (such as *hepatitis*), bacteria (such as *Salmonella* and *Shigella*), helminthes (such as *Ascaris*) and algae (such as *Schizothrix*). The likelihood of infection upon consumption depends upon the virulence of the strain present, survival in the aquatic environment, the number of organisms ingested, the estimated infectious dose, and the susceptibility of the host. As an example, cryptosporidiosis is a gastrointestinal upset caused by ingestion of the environmentally resistant stage of *Cryptosporidium parvum,* the oocyst. This parasitic protozoan has a wide range of hosts, including domesticated animals and wildlife. Oocysts of *Cryptosporidium* spp. are found in surface waters throughout the world. The disease is usually self-limiting, unless the infected subject is immunocompromised.

Immunocompetent hosts develop an immune response to the parasite which dramatically diminishes the likelihood of illness after initial exposure. Control measures for waterborne diseases include water treatment, using flocculants, settling and filtration processes, followed by application of disinfectants such as sodium hypochlorite. Education is an important component of strategies for reducing the prevalence of waterborne diseases, especially education about the fecal–oral route of infection and the role of proper human waste disposal and personal hygiene in preventing disease transmission.

Water vectored diseases are caused by microorganisms that have intermediate hosts in insects that reproduce in aquatic environments and transmit pathogenic microorganisms as part of feeding on human blood. If the aquatic environment needed for the insect to reproduce is not present, the prevalence of the disease is reduced. Malaria, yellow fever, and onchocerciasis are water vectored diseases associated with mosquitos and black flies, respectively. The parasitic microorganisms that initiate disease in humans undergo a portion of their life cycles in the insect host and are transmitted during the bite directly into the human bloodstream. In malaria, the protozoan species of genus *Plasmodium* must develop in an anopheline mosquito. When the female mosquito bites an infected person, it draws blood containing forms of the parasite that develop further within the mosquito, eventually entering the salivary glands. The infected mosquito then injects the developed form into the next human host in saliva secreted during the bite. The anopheline mosquito requires standing, fresh waters for reproduction. Common habitats for breeding include ponds and marshes. Control programs to prevent water vectored diseases focus on eliminating aquatic habitat for insect hosts and reducing the prevalence of disease in human hosts. Activities for this include draining perennially wet areas to remove standing water and application of pesticides or oils designed to kill developing mosquito larvae.

Water washed diseases are caused by contact with water contaminated with fecal material and microorganisms that colonize and reproduce in skin lesions or vulnerable tissues. Common microbial pathogens that cause this type of infection include *Pseudomonas* spp. and *Staphylococcus* spp. This type of water related illness is most associated with lack of adequate supplies, which leads to inadequate personal hygiene. Typical water washed diseases include skin irritations and ulcers, conjunctivitis, and trachoma. Increased supply, coupled with improved sanitation and education, often reduces the prevalence of water washed disease.

WATER VECTORED DISEASE. See Water Related Diseases

WATER WASHED DISEASE. See Water Related Diseases

DISHWASHING WATER QUALITY PROPERTIES

Eva Ståhl Wernersson
Erik Johansson
GS Development AB
Malmö, Sweden

Håkan Håkanson
University of Lund
Lund, Sweden

INTRODUCTION

Water is used as a rinse medium in dishwashing to remove soil from the dishes. An important requirement of the process is that the consumption of fresh water should be moderate, because the supply of fresh water in many countries is limited. Furthermore, the dishwater should reduce the microbiological activity (even when a dishwasher is not used) to diminish the risk of cross-contamination from the dishwater to the dishware.

In dishwashing, it is not only the dishwater that contributes to the cleaning of the dishware, but also the dishwashing process and the detergents. The dishware covers a range from glasses to pots and pans, which consist of a variety of materials such as glass, china, ceramics, thermostable plastics, stainless steel, and aluminum. The dishwashing processes use techniques from hand dishwashing or machine dishwashing to granule-assisted pot washing. For hand dishwashing, the dishware is soaked in water with hand dishwashing detergents (pH 7 to 8) before cleaning. The contact time between the dishware and the dishwater is in the range of 20 to 60 minutes, whereas in dishwashers, the contact time between dishwater and dishware is on the order of 1 to 10 minutes. The dishwashers use detergents with a pH from 12 to 14 that in the dishwater solution becomes pH 9 to 11. In all dishwashing processes, the importance of rinsing the dishware with fresh water is unquestionable (1). In the international standards (2,3), either hot fresh water at a temperature of 85 °C or cold fresh water with additives of disinfecting chemicals such as chlorine should be used to achieve hygienically clean dishware.

The microbiological status of the dishwater is of great importance, because this could cause cross-contamination and thereby the spread of infections (4,5). Therefore, this subject will be addressed more thoroughly.

CHARACTERIZING DISHWATER QUALITY

Dishwater consists of fresh water, detergent, and food residues. Depending on the hardness of the fresh water supply, detergents with more or less complex-acting properties are used. In hand dishwashing, the concentration of the washing-up detergent is around 0.05% (6), and in dishwashers, it is normally about 0.2% (7). The soil in the dishwater can be anything ranging from lipstick residues on glasses to carbohydrates like starch or vegetable fibers, proteins, animal or vegetable fats, and microorganisms from food residues.

There are several parameters to consider when characterizing dishwater, such as the conductivity, the pH, the temperature, the chemical oxygen demand (COD), the general microbiological level expressed as aerobic plate count (APC), and the characterization of the microbiological content. The amount of solid substances in dishwater is low; it varies from 5 mg/L (dishwasher for glasses and cutlery) to 5000 mg/L (pots and pans machine) according to Harpel (8) and our own measurements.

The conductivity is commonly used to monitor the concentration of the detergent used in dishwashers. However, the conductivity can be affected by the ions of the soil dispersed in the dishwater, and this could cause a malfunction of the dosage of the detergent.

The pH is a result of the amount and type of detergent used and the soil on the dishware. The organic soil in the dishwater is hydrolyzed by the high pH into smaller components. The influence of the pH on the microbiological growth in dishwater is shown in Fig. 1. From the figure, it can be seen that a change of the pH from 7 to 9 at 65 °C has a relatively small influence on the reduction of the concentration of the investigated bacteria, *Enterococcus feacium*. The samples were taken and analyzed 30 minutes after the addition of the bacteria, where the initial concentration was measured to be 50000 cfu/mL (^{10}log 4.7). The detergent used in the experiment was Gigant (Diskteknik, Sweden), an ordinary alkaline dishwasher detergent used in professional kitchens. The chosen test microorganism in this investigation is a rather thermostable bacteria, used in international standards (3) and in literature (8).

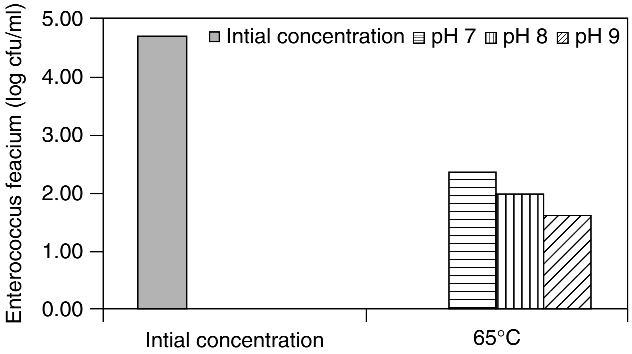

Figure 1. The reduction, 30 minutes after injection of *E. faecium* in dishwater, as a function of pH at 65 °C dishwashing temperature.

The temperature of the dishwater is determined by the process used. For hand dishwashing, the temperature is normally 40–45 °C (9), and for dishwashers, the temperature is 55–65 °C (7). The influence of the temperature is vital for the reduction of the concentration of bacteria in dishwater (see Fig. 2). For this microorganism, the reduction after 30 minutes is negligible for temperatures lower than 55 °C. However, at 65 °C there is a noticeable reduction of the concentration of *E. feacium*. For comparison, in Fig. 2, the reduction of the concentration of the bacteria for the conditions in

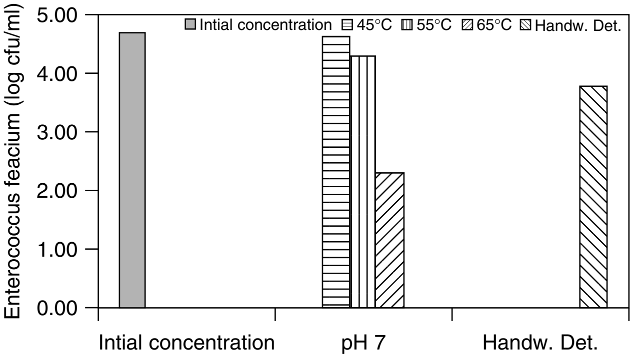

Figure 2. The reduction of *E. faecium,* 30 minutes after injection in dishwater, as a function of temperature of the dishwaters. The pH was kept to 7. For comparison, the reduction of *E. faecium* for the same time, using hand dishwashing detergent at 45 °C dishwashing temperature.

hand dishwashing is included. The decrease in microbiological activity depends solely on the antibacteriological properties of the detergent (Dizzy, Diskteknik, Sweden). Even with longer exposure times of microorganisms in dishwaters containing hand dishwashing detergents, the reduction rate will be low (6,10). Because of safety regulations in many countries, hand dishwashing cannot be performed at temperatures above 55 °C.

Another important factor of the temperature is the emulsion of fats into dishwater, which is considerably more effective at higher temperatures.

The COD depends mainly on the type and the amount of soil on the dishware. Also, the detergent adds to the value of the COD for dishwater. Values ranging from 500 to 5700 mg/L are reported in literature (8) and confirmed by our own measurements (see Fig. 3). Dishwater samples were taken in a professional kitchen from a pot and pan machine. The first sample was taken before the changing of dishwater. The second sample was taken after draining and filling the dishwater tank with fresh water and detergent, and the following ones after each dishwashing cycle. In this experiment, the increase of the COD is about 270 mg/L for each cycle. The manufacturers of one-tank machines recommend changing the dishwater after about 20 dishwashing cycles.

The microbiological content of the dishwater is more difficult to predict. There are several types of active bacteria in dishwater. In the studies of Mattick et al. (9) and Mospuye and van Holy (11), it is obvious that the distribution of the species found in the food residues are transferred to the dishwater. The most frequently reported microorganisms are *Bacillus cereus, Bacillus* spp., *Campylobacter* spp., *Clostridium prefringens, E.coli*, Enterobacteriaceae, *Salmonella* spp., and *Staphylococcus* spp. (9,11–13). Some of these are foodborne pathogens that pose a considerable risk of the spread of infection, if cross-contamination occurs (4).

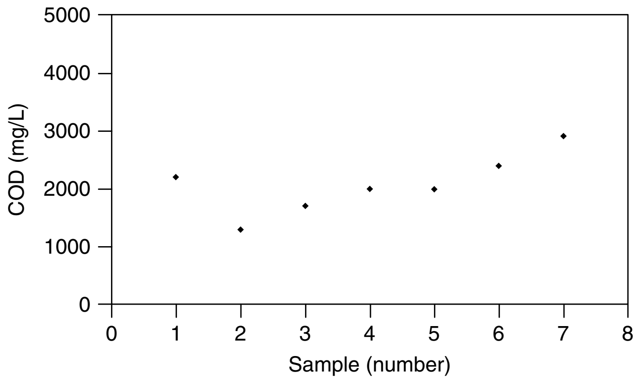

Figure 3. COD of dishwater samples as a function of sample numbers.

Reported values of the APC in dishwater show that hand dishwashing exhibits higher values of the APC than does dishwater from dishwashers (see Table 1) (9,12,13). The content of Enterobacteriaceae and *Bacillus cereus* in the dishwaters from the studies can also be seen in the same table. The level of the APC ranges from $^{10}\log 2.6 \pm 0.8$ to $^{10}\log 4.0 \pm 1.2$. The observations confirm that in hand dishwashing conditions, with a lower temperature and less aggressive detergents, the microbiological activity is higher than in dishwater from dishwashers.

ENVIRONMENTAL ASPECTS

An important factor from an environmental point of view is the consumption of water. The largest consumption of water is for hand dishwashing. In professional dishwashers, the dishwater is reused up to 20–30 cycles, and only diluted with fresh water from each rinse cycle with about 10% of the tank volume. The tank volume ranges from 50 to 250 L in dishwashers. In dishwashers, where the cleaning process is assisted by plastic granules, the granules are reused for about 2000 dishwashing cycles.

To reduce the microbiological activity in dishwater, several methods can be used, for instance, ozone, high temperature and high pH, and disinfecting chemicals such as chlorine (14). From an environmental point of view, high temperature is the most agreeable and efficient parameter to use (see Fig. 2). In many countries, detergents containing aggressive chemicals, like chlorine, are not recommended, which is consistent with an environmental approach. Even the use of ozone is questionable because its general reactivity is well documented.

Table 1. The APC, Entereobacteriaceae, and *Bacillus cereus* Found in Different Dishwaters (cfu/mL) from Professional and Domestic Kitchens

Type of Establishment	Domestic Homes (9)	Hospices (9)	Street-Vended Food (11)	Professional Kitchens (13)
Dishwashing process	Hand dishwashing	Hand dishwashing	Hand dishwashing	Dishwasher
Number of samples	52	10	18	16
APC				
Mean	$1.3 \cdot 10^5$	$2.3 \cdot 10^4$	$1.0 \cdot 10^4$	$2.7 \cdot 10^2$
Range	10^1 to $>3 \cdot 10^5$	$3 \cdot 10^3$ to $1 \cdot 10^5$	$3 \cdot 10^3$ to $3 \cdot 10^6$	<10 to 10^4
Enterobacteriaceae				
Mean	$7.1 \cdot 10^4$	$7.2 \cdot 10^2$	ND	1.2
Range	<1 to $>3 \cdot 10^5$	<1 to $>6 \cdot 10^3$	ND	<1 to 10^2
Bacillus cereus				
Mean	ND	ND	46% of 65 isolates were determined	$2.7 \cdot 10^2$
Range				<10 to 10^4

ND: not determined.

CONCLUSIONS

The temperature is the most important factor in maintaining a low level of microbiological activity in the dishwater. This parameter is also from an environmental point of view less harmful than ozone or detergents containing other oxidizing chemicals. According to the international standard, the heat load on dishware should exceed 3600 HUE (Heat Unit Equivalence) to achieve hygienically clean dishware.

To shorten the dishwashing cycle time, granule-assisted dishwashing could decrease the energy and detergent consumption in the cleaning process. The rinsing procedure with fresh water must follow the above-mentioned standard.

Hand dishwashing consumes more fresh water than do dishwashers. It also uses a lower temperature, which has a considerably lower effect on the reduction of the microbiological activity of the dishwater. It is important to keep an acceptable quality level of the reused dishwater, because otherwise the dishwashing process and the rinse process will not be able to obtain clean dishware. To improve the dishwater quality, some manufacturers have integrated cyclones or filters into their dishwashers.

BIBLIOGRAPHY

1. Cogan, T.A., Sladder, J., Bloomfield, S.F., and Humphrey, T.J., (2002). Achieving hygiene in the domestic kitchen: the effectiveness of commonly used cleaning procedures. *J. Appl. Microbiol.* **92**: 885–892.
2. American National Standard/NSF International Standard 3. (2001). Ann Arbor, MI.
3. Deutsche Norm DIN 10512. (2001). *Food Hygiene—Commercial Dishwashing with Onetank-Dishwashers—Hygiene Requirements*, type testing, Gewerbliches Geschirrspülen mit Eintank-Geschirrspülmaschinen Dezember, Beuth Verlag GmbH, Berlin, Germany.
4. Ståhl Wernersson, E., Johansson, E., and Håkanson, H. (2004). Cross-contamination in dishwashers. *J. Hospital Infection* **56**: 312–317.
5. Cowan, M.E., Allen, J.P., and Pilkington, F. (1995). Small dishwashers for hospital ward kitchens. *J. Hospital Infection* **29**: 227–231.
6. Charnock, C. (2003). The antibacterial efficacy of Norwegian hand dishwashing detergents. *Food Protection Trends* **23**: 10, 790–796.
7. Ståhl Wernersson, E., Johansson, E., and Håkanson, H. (2004). Granule-assisted dishwashing improves hygiene. *Food Service Technol.* **4**: 129–137.
8. Harpel, S., Tilkes, F., and Krüger, S. (1994). Forderungen an die maschinelle Dekontamination vor Geschirr unter hygienischen Gesichtspunkten sowie an eine geringe Belastung der Abwässer—ein Widerspruch? (Requirements to the decontamination in dish-washing machines under hygienic and environmental aspects—a contradiction in terms?) *Zbl. Hyg.* **195**: 377–383.
9. Mattick, K., Durham, K., Hendrix, M., Slader, J., Griffith, C., Sen, M., and Humphrey, T. (2003). The microbiological quality of washing-up water and the environment in domestic and commercial kitchens. *J. Appl. Microbiol.*, **94**: 842–848.
10. Kusumaningrum, H.D., van Putten, M.M., Rombouts, F.M., and Beumer, R.R. (2002). Effects of antibacterial dishwashing liquid on foodborne pathogens and competitive microorganisms in kitchen sponges. *J. Food Protection* **65**: 61–65.
11. Mosupye, F.M., and von Holy, A. (1999). Microbiological quality and safety of ready-to-eat street vended foods in Johannesburg, South Africa. *J. Food Protection* **62**: 1278–1284.
12. Mosupye, F.M. and von Holy, A. (2000). Microbiological hazard identification and exposure assessment of street food vending in Johannesburg. *Int. J. Food Microbiol.* **61**: 137–145.
13. Ståhl Wernersson, E., Håkanson, H., Lindvall, I., and Trägårdh, C. (2003). Hygiene in warewashers utilizing blasting granules that foodservice establishments use. *Food Protection Trends* **23**(10): 797–807.
14. von Krüger, S. and Zschaler, R. (1996). Überprüfung der Dekontaminationswirkung in Mehrtank-Spühlmaschinen. *Arch. Lebensmittelshygiene* **47**: 105–128.

DISINFECTION BY-PRODUCT PRECURSOR REMOVAL FROM NATURAL WATERS

Yeomin Yoon
Northwestern University
Evanston, Illinois

Shahnawaz Sinha
Malcolm Pirnie Inc.
Phoenix, Arizona

Jinsik Sohn
Kookmin University
Seoul, Korea

BACKGROUND

Water is one of Earth's most important natural resources, and its role in our everyday life cannot be ignored. Civilizations began and were centered within regions of abundant water supplies. However, water present in our natural environment is not always potable for human consumption. Many lives have been lost as recently as the early part of last century because of waterborne diseases such as cholera and typhoid fever. Tragically, this battle still continues elsewhere in developing countries where infant mortality is often the result of waterborne diseases (1). To be "potable," water needs to be disinfected with a method such as chlorination, which has been successfully practiced in the United States for over a century.

However, chlorinated disinfection by-products (DBPs) have been a concern for many years because of their adverse effects on human health. A great deal of research has been conducted on these chlorinated by-products. Several chlorinated DBPs have already been identified since the discovery of trihalomethanes (THMs) in 1974 (2). Other chlorinated disinfection by-products include haloacetic acids (HAAs), chloral hydrate, haloketone, and haloacetonitriles. One way to identify the total amount of chlorinated DBPs is to measure the total organic halide (TOX) concentration. However, if all identified chlorinated DBPs were considered, they would only account for about 50% of the TOX (3). The unidentified

DBPs may be too large (nonvolatile) or too small (polar) in molecular weight to be detected by currently accepted analytical methods such as gas chromatography.

In recent years, other alternative disinfectants have been discovered and used by scientists, including chloramines, chlorine dioxide, ozone, potassium permanganate, and ultraviolet irradiation. However, because of proven reliability, cost, and availability, chlorine has remained the chemical of choice. As a result, of all the municipal water supplies that are being chemically disinfected, at least 70% use chlorine (4). Unfortunately, in the process of attacking these unwanted microorganisms, the chlorine reacts with other impurities such as natural organic matter (NOM) and bromide, and it can lead to the formation of carcinogenic DBPs (5). Alternative disinfectants may help to minimize these DBPs. However, these oxidants/disinfectants may produce other DBPs, which may also be harmful (4).

Health-related studies on these carcinogenic by-products suggest that the presence of as low as 6 μg/L of chloroform (chlorinated DBP) can have a lifetime cancer risk of 10^{-6} (6). Another recent DBP-related risk study (7) showed that as the bromide concentration increases from 50 μg/L to 200 μg/L, it substantially increases overall DBP-related cancer risk. To reduce the public health risk from these toxic compounds, the U.S. Environmental Protection Agency (EPA) has regulated THMs since 1979, with a maximum contaminant level (MCL) of 100 μg/L. In view of recent studies, the EPA has proposed a more stringent disinfection/disinfection by-products (D/DBPs) rule, which will be implemented in two stages. Stage 1 (effective from 1998) of the D/DBP rule would reduce the current MCL for THMs from 100 μg/L to 80 μg/L, with a new MCL restriction of 60 μg/L for HAA_5 (sum of five HAA species). Stage 2 (effective from 2002) would further reduce both of these MCLs for THMs and HAAs (as HAA_6) by an additional 50% (8). However, the success of water utilities in meeting these stringent and challenging MCLs will require a better understanding of the DBP precursor material, NOM, and its reactivity with chlorine in forming DBPs.

The most effective strategy to decrease the formation of DBPs would be to reduce the amount of NOM, which is the main precursor of DBPs, before chlorination. The processes that are available to remove this precursor material include membranes, granular activated carbon adsorption, biofiltration processes, and chemical treatment by coagulation. Coagulation has remained the most widely accepted method by utilities to remove NOM from natural waters.

PRECURSOR REMOVAL BY COAGULATION

Coagulation

Coagulation is an age-old practice in water treatment and has not changed much over the years. Until recently, coagulation has primarily removed color (e.g., NOM) and turbidity from water. More attention has been given to NOM removal from water. In the early 1990s, the EPA proposed the enhanced coagulation (EC) rule to further reduce health-related risks from these DBPs and compelled utilities to practice EC to further remove NOM. In enhanced coagulation, NOM removal is based on raw water total organic carbon (TOC) and alkalinity content (Table 1). EC has already been evaluated by several researchers and has been found to be effective in both NOM removal and DBP minimization. Other benefits include reduction in chlorine demand in finished water (9). The main objective of EC is to achieve 15% to 50% removal of organic matter before disinfection. However, some of the disadvantages of EC also reported from these studies include increased chemical use, which translates into higher operational cost; higher sludge production; and additional chemical costs associated with pH adjustments. Thus, more research is still needed to overcome these disadvantages.

Table 1. Percent of Organic Removal for Enhanced Coagulation Requirements Proposed for the DBP Rule

	Alkalinity, mg/L as $CaCO_3$		
TOC, mg/L	<60	60–120	>120
2–4	35%	25%	15%
4–8	45%	35%	25%
>8	50%	40%	30%

Coagulants Used for NOM Removal

Various types of metal coagulants are available to remove NOM. However, two of the most common conventional metal coagulants in water treatment are aluminum sulfate ($Al_2(SO_4)_3 \cdot 14H_2O$, commonly referred to as alum) and ferric chloride ($FeCl_3 \cdot 6H_2O$, commonly referred to as iron salt). Alum is currently the most widely used. However, over the past years, the application of iron salt as a coagulant has increased dramatically (10). Conventional coagulants of other types include aluminum-based sodium aluminate, aluminum chloride, and ferric sulfate (11). However, the application of these coagulants in potable water has been limited. Recently, various other types of (alternative) coagulants have been introduced. These specialized coagulants are more common in Europe and Japan than in the United States (12). These coagulants (patented products) are claimed to be more effective than the conventional coagulants. Conventional coagulants are basically salts of a strong acid (e.g., HCl or H_2SO_4) and a weak base ($Al(OH)_3$ or $Fe(OH)_3$). Thus, it is a mixture of cations (from bases) and anions (from acids). However, alternative coagulants are specially prepared and may include additional anions as additives (i.e., anions such as sulfate).

Effectiveness of Coagulation in Precursor Minimization

Numerous studies have focused on removal of the "humic" fraction of NOM that is believed to be the precursor for THMs. Thus, most of the research has focused on the removal of this fraction by coagulation and other processes. Oliver and Lawrence (13) studied raw water samples collected from treatment plants in the Great Lakes Basin to investigate the effect of alum coagulation in haloform reduction and reported a dramatic drop in haloform formation by alum coagulation. Chadik and Amy

(14) studied natural waters in six different geographical locations to evaluate the effectiveness of metal coagulants (alum and ferric chloride) in removing THM precursors. THM formation potential (THMFP) reduction by alum was in a range between 36% and 66%, whereas the reduction by iron was in a range between 33% and 77%. Alum was found to be more effective than iron in reducing TOC and THMFP. Jodellah and Weber (15) investigated THM precursor removal by alum coagulation, lime softening, and activated carbon adsorption, and they found that precursor removal was variable and concluded that greater removal of organic matter does not necessarily indicate greater reduction of THMFP. No correlation was found between TOC removal by these treatments and overall THMFP reduction. It can be concluded that the reduction of THMFP is dependent on the qualitative nature of the organic matter, rather than on the quantitative amount of organic matter. Thus, a careful assessment of the physical and chemical composition remaining after different treatments of NOM mixtures is required before a mechanistic explanation can be provided.

Hubel and Edzwald (16) studied the THM precursor removal by alum, high-molecular-weight polymers, cationic polymers (usually used as coagulant aids), and various combinations of these coagulants. In this study, it was concluded that cationic polymers were less effective in precursor removal as compared with alum when used as a sole coagulant, where alum removed 50% to 60% more organic matter than did cationic polymers. This study also indicated that the type of polymers used (polyacrylamide versus polyquaternary amine) (i.e., charge density and molecular weight) had slight effects on precursor removal. Cationic polymers as sole coagulants were found not to be effective in THM precursor removal, but they were effective in turbidity removal by forming denser and better settling flocs than alum alone. Knocke et al. (17) studied the effect of seasonal temperature variation in precursor removal in natural water by alum and ferric sulfate. The results showed that lower temperature (coagulation at pH 5.0) significantly reduced the rate of THM formation (2 °C versus 22 °C) for equivalent organic content.

SUMMARY

This study provides a general overview on coagulation and its effectiveness in removing NOM from natural waters for DBP minimization with various metal coagulants. Although other technologies such as activated carbon adsorption and membrane filtration are available to reduce DBP precursors, these alternative technologies are comparatively expensive processes (18). For the time being, coagulation may remain the technology of choice for DBP minimization. In addition, identification of DBP precursors based on detailed NOM characterization and fractionation techniques can provide mechanisms for how these precursor materials could effectively be controlled in natural waters. Once identified, these potential precursor materials could then be effectively removed by in-depth knowledge of coagulants and coagulation techniques.

BIBLIOGRAPHY

1. Carol, H.T. and Fox, K. (1990). *Health and Aesthetic Aspect of Water Quality in Water Quality and Treatment—a Handbook of Community Water Supplies*, 4th Edn. American Water Works Association/McGraw Hill, New York.
2. Rook, J.J. (1974). Formation of haloforms during chlorination of natural water. *Water Treatment Exam* **23**: 234.
3. Krasner, S.W. et al. (1989). The occurrence of disinfection by products in US drinking water. *J. Am. Water Works Assoc.* **81**: 41–53.
4. Oxenford, J. (1996). *Disinfection By-products: Current Practices and Future Direction in Disinfection By-products in Water Treatment*. CRC/Lewis Publishers, Boca Raton, FL.
5. Rook, J.J. (1976). Haloforms in drinking water. *J. Am. Water Works Assoc.* **68**: 168–172.
6. Bull, R.J. and Kopfler, F.C. (1991). Health effect of disinfectants and disinfection by-products. Proceedings of American Water Works Association Annual Conference, Denver, CO.
7. Sohn, J., Sinha, S., Amy, G.L., and Zhu, H.W. (1997). The role of coagulation in DBP related risk reduction. Proceedings of American Water Works Association Water Quality & Technology Conference, Denver, CO.
8. Pontius, F.W. (1996). An update of federal regs. *J. Am. Water Works Assoc.* **88**: 36–46.
9. Sinha, S. (1996). *Disinfection by-products precursors removal by enhanced coagulation*, Masters Thesis, The University of Kansas-Lawrence.
10. Baltpurvins, K.A., Burns, R.C., and Lawrance, G. (1996). Effect of pH and anion type on the aging of freshly precipitated iron (III) hydroxide sludge. *Environ. Sci. Technol.* **30**: 939–944.
11. Van Benschoten, J.E. and Edzwald, J.K., (1990). Measuring aluminum during water treatments. *J. Am. Water Works Assoc.* **82**: 71–78.
12. Amirtharajah, A. and O'Melia, C.R. (1990). *Coagulation Process: Destabilization, Mixing and Flocculation, in Water Quality and Treatment*, 4th Edn. American Water Works Association/McGraw Hill, New York.
13. Oliver, B.G. and Lawrence, J. (1979). Haloforms in drinking water: a study of precursors and precursor removal. *J. Am. Water Works Assoc.* **71**: 161–163.
14. Chadik, P.A. and Amy, G.L. (1983). Removing trihalomethane precursors from various natural waters by metal coagulants. *J. Am. Water Works Assoc.* **75**: 532–536.
15. Jodellah, A.M. and Weber, W.J. (1985). Controlling trihalomethane formation potential by chemical treatment and adsorption. *J. Am. Water Works Assoc.* **77**: 95–100.
16. Hubel, R.E. and Edzwald, J.K. (1987). Removing THM precursor by coagulation, *J. Am. Water Works Assoc.* **79**: 98–106.
17. Knocke, W.R., West, S., and Hoehn, R.C. (1986). Effect of low temperature on the removal of trihalomethane precursors by coagulation. *J. Am. Water Works Assoc.* **78**: 189–195.
18. Singer, P.C. (1994). Control of disinfection by products in drinking water. *J. Environ. Eng.-ASCE* **120**: 727–744.

ALTERNATIVE DISINFECTION PRACTICES AND FUTURE DIRECTIONS FOR DISINFECTION BY-PRODUCT MINIMIZATION

June M. Weintraub
City and County of San Francisco Department of Public Health
San Francisco, California

J. Michael Wright
Harvard School of Public Health
Boston, Massachusetts

Raghuraman Venkatapathy
Oak Ridge Institute for Science and Education
Cincinnati, Ohio

INTRODUCTION

Disinfection of public water supplies is required in many countries to inactivate microbial pathogens and protect public health. Chlorine is the predominant disinfectant used worldwide for this purpose, and while chlorine provides significant benefit, it can create by-products that may cause adverse health effects. Disinfection by-products (DBPs) have been linked to bladder and rectal cancers and to adverse reproductive outcomes (1,2). The Safe Drinking Water Act amendments of 1996 required the United States Environmental Protection Agency (U.S. EPA) to develop regulations that balanced the conflicting goals of pathogen control with DBP minimization. The Stage 1 Disinfection By-products Rule established standards for trihalomethanes (THMs), haloacetic acids (HAAs), bromate, and chlorite in U.S. drinking water (3).

The use of alternative disinfectants in the United States has received increasing attention in light of DBP regulatory changes. The formation of DBPs depends both on the type of disinfectant used and the presence of precursors. Thus, there are essentially three approaches to reduce the occurrence of any particular DBP: (1) alternative disinfection, (2) other treatment options (e.g., that reduce either the presence of DBP precursors or the amount of disinfectant needed), and (3) approaches aimed at source water, such as watershed protection to limit the addition of DBP precursors to waterways, or the use of source water with naturally low precursor levels. The advantages and disadvantages of these options are briefly outlined below, including usage information, operational concerns, disinfectant properties, and potential impact on the occurrence of specific DBPs in drinking water supplies.

ALTERNATIVE DISINFECTANTS

A number of disinfectants are appropriate to treat drinking water on a large scale. Chlorine, chloramine, chlorine dioxide, ozone, and ultraviolet radiation are among the most widely used disinfectants. Ozone and chlorine dioxide are more common in some European countries, but most water systems use chlorine for either primary (predistribution) or secondary (residual) disinfection (4). Of all the choices, only chlorine and chloramine are suited for secondary disinfection to maintain residual protection in the distribution system.

Chloramine

Chloramines were first used to treat water in the United States and Canada in 1917 (5). Chloramination is increasing in the United States, with about 30% of medium and large utilities reporting its use in 1998, compared to 20% in 1989. Chloramination involves the addition of ammonia to chlorinated water; aqueous chlorine reacts with ammonia to form inorganic chloramines (6). A principal operational consideration is that the chlorine/ammonia dosage ratio must be monitored carefully to prevent the growth of nitrifying bacteria in distribution systems and storage reservoirs (7).

Chloramines are weaker disinfectants compared to chlorine, ozone, or chlorine dioxide. Chloramine *CT* (product of disinfectant concentration and contact time) values to control *Giardia*, *Cryptosporidium*, or viruses are up to four orders of magnitude higher than for the other disinfectants, making chloramines an impractical choice for primary disinfection (Table 1). Chloramines are therefore typically used only for secondary disinfection; the maximum residual disinfectant level (MRDL)

Table 1. *CT* (Concentration × Time) Values[a] for Alternative Disinfectants

		Virus			*Giardia*			*Cryptosporidium*
Disinfectant	Units	2 log	3 log	4 log[b]	1 log	2 log	3 log[b]	2 log[b]
Chlorine	mg · min/L	3	4	6	35	69	104	>4,000
Chloramine	mg · min/L	643	1067	1491	615	1230	1850	>25,000
Chlorine dioxide	mg · min/L	4	13	25	8	15	23	40[c]
Ozone	mg · min/L	<1	<1	1	<1	<1	1	3.9–10
UV radiation	mJ/cm^2	40	N/A[d]	N/A	N/A	N/A	N/A	12

[a] All values are at 10 °C, pH 6–9.

[b] Required removal for large public drinking water systems in the United States.

[c] Chauret et al. (8) report variable *CT* values for 2 log inactivation of *Cryptosporidium* at 21 °C, ranging from 75 to 1000 mg·min/L.

[d] N/A, not available.

Source: References 6, 8–10.

for chloramines is 4 mg/L. Compared to chlorine, chloramines are more stable secondary disinfectants and may reduce taste and odor complaints. Chloramination may also be more effective than chlorine at inhibiting biofilm and bacterial regrowth in distribution systems (11).

The use of chloramines reduces the formation of two of the major regulated DBP groups—trihalomethanes and haloacetic acids. However, some DBPs, including dihalogenated acids and NDMA (*N*-nitrosodimethylamine), may increase in chloraminated systems (12,13). Source water chemistry and ammonia application timing influence the formation of chloramination by-products. For example, in one chloraminated system, high levels of iodoacids were attributed to the simultaneous application of chlorine and ammonia to source water high in iodide. When ammonia application is applied after significant free chlorine contact time, iodoacid formation is minimized (14).

Chlorine Dioxide

During the 1940s, the Niagara Falls, New York water treatment plant was the first North American utility to use chlorine dioxide for water treatment (15). Chlorine dioxide was used by 8% of medium and large utilities in the United States in 1998, more than triple the proportion that utilized it in 1989 (7). Chlorine dioxide must be generated on site, which requires significant operator skill; thus, it is relatively expensive in comparison to chlorine. Chlorine dioxide is flammable and explosive, and its use therefore requires significant safety precautions.

Compared to chlorine, chlorine dioxide is more effective at inactivation of *Cryptosporidium* and *Giardia* but less effective for *E. coli* and rotaviruses (see Table 1). Although chlorine dioxide CT values of 40 mg · min/L at 10 °C are reported to accomplish 2 log removal, (i.e., 99%) of *Cryptosporidium* (6), highly variable CT values have been reported at other temperatures (8). An advantage of chlorine dioxide is that disinfection efficiency is not dependent on pH or the presence of ammonia (16). However, the half-life decreases at $\text{pH} > 9$ and with increasing chlorine dioxide concentrations (17). The MRDL for chlorine dioxide is 0.8 mg/L (6).

Chlorine dioxide is effective at controlling DBP formation, since it can destroy precursors to trihalomethanes and haloacetic acids and does not oxidize bromide ion to bromate. The predominant DBPs formed by chlorine dioxide are chlorate and chlorite, which may lead to adverse human health effects (see the chapter entitled Health Effects of Commonly Occurring Disinfection By-products in the Municipal Water Supply). Chlorine dioxide can produce other DBPs such as dihalogenated HAAs and haloacetaldehydes (18). Due to the inability of chlorine dioxide to remove hydroxyfuranone precursors, high levels of MX and the brominated MX analogues have been found in waters with elevated total organic carbon and bromide (19).

Ozone

Ozonation was first used in The Netherlands in 1893 and remains a popular disinfection method in Europe (6). While ozonation was used for disinfection by less than 10% of large water utilities in 1998, its use is increasing in the United States (7). In the proposed Stage 2 Disinfectants and Disinfection By-products Rule, the U.S. EPA reported that in 2000 there were 332 plants of various sizes using ozone and 58 plants that were planning to install ozonation (20). Ozone is one of the strongest oxidants used to disinfect water and is the predominant choice of the bottled water industry. Although ozonation is a fairly expensive and complex treatment process, it has several operational advantages. These advantages include short contact time, taste and odor control, and improved coagulation and filtration efficiency. Ozone is very effective at microbial inactivation and is effective against viruses, *Giardia*, and *Cryptosporidium* (see Table 1). Because ozone can also degrade more complex organic matter that may promote microbial regrowth in distribution systems (21), and because ozone does not provide a protective residual, it must be used in combination with chlorine or chloramine in systems where residual disinfection is necessary.

Use of ozone limits the formation of chlorinated DBPs, but it can combine with bromide in water to form bromate, brominated organics, and other brominated by-products (22,23). Ozonation can also produce organic DBPs such as the aldehydes, ketones, ketoaldehydes, carboxylic acids, aldo acids, keto acids, hydroxy acids, alcohols, esters, and alkanes (23). Furthermore, if ozonation is followed by chlorination, the formation of aldehydes and ketones may allow subsequent formation of their halogenated derivatives (e.g., chloropicrin and other halonitromethanes) (19,21,24).

Ultraviolet Radiation

Ultraviolet (UV) radiation is not widely used to treat drinking water in the United States, but has been used extensively to treat wastewater (20). UV radiation can inactivate viruses, bacteria, and protozoans (9). The use of UV radiation is somewhat limited by intensive operational and maintenance requirements, and it is most effective in low turbidity waters. Advantages of UV radiation include a very short detention time requirement, so dedicated contact chambers are not necessary (25). However, a slow flow rate is required, and this may pose operational constraints. UV disinfection does not provide residual protection against bacterial regrowth in the distribution system. UV radiation alone forms few DBPs but DBP formation can still occur when UV disinfection is combined with secondary disinfectants (26).

OTHER TREATMENT OPTIONS

In addition to alternative disinfection, many treatment options are available to water utilities to reduce DBP formation. These include pH adjustment, disinfectant timing or treatment process changes, and physical or oxidative removal of DBP precursors. Changes to these complex treatment processes can have different effects on the various types of DBPs.

The impact of pH on DBP concentrations is reported by a number of researchers (27–32). In general, the concentrations of THMs increase while those of HAAs, HANs, and haloketones decrease with increasing pH. Many of the halogenated DBPs are hydrolyzed above pH 8, reducing their formation (33).

Treatment plant operators may be able to limit DBP formation by reducing contact time or the initial disinfectant dose. Changing the order of disinfectant application within the treatment chain should also be considered. This may include moving the primary disinfectant step to follow filtration, or changing the application point for the secondary disinfectant application (34). Distribution system maintenance may decrease residual disinfectant demand, allowing operators to use smaller amounts of disinfectant.

The major precursors to DBPs of concern are organic substances (e.g., humic and fulvic acids), bromine, and iodine. If treatment processes or primary disinfection methods that reduce DBP precursors are used, then secondary disinfection with chlorine or chloramine will produce lower levels of their associated by-products. Precursor removal can be accomplished by coagulation, filtration, or oxidative processes.

Coagulation is an important step in converting soluble precursors to particulate, which can then be removed by filtration or settling (35). Enhanced coagulation and enhanced softening processes are those that improve removal of DBP precursors over conventional treatment. Coagulants specified by the U.S. EPA as appropriate for precursor removal to achieve compliance with the Stage 1 Disinfection By-Products Rule include aluminum or iron salts, polyaluminum chloride, and cationic polymers, as well as anionic and nonionic polymers (36).

Physical removal of precursors includes processes such as microfiltration and biofiltration, which are used primarily in larger plants. Biofiltration through the use of sand filters or granulated activated carbon can reduce formation of halogenated DBPs (37). Passage of ozonated and chlorinated water samples through a rapid sand filter reduced the concentration of aldehydes by 26–62%, due to biological activity in the sand filters (38,39). However, removal of organic contaminants through biofiltration does not affect bromide concentrations in the treated water and results in greater formation of brominated DBPs (39). Other physical precursor removal techniques like electroporation (40) show promise but are in the early stages of development for practical use.

Oxidative precursor removal is another effective means of reducing DBPs. Ozone is effective at destroying precursors to MX and trihalogenated HAAs (19). Potassium permanganate is very effective at removing DBP precursors, both by oxidation and by enhancing coagulation (6). Some advanced oxidation processes (AOPs) that combine oxidation techniques show promise for utilities (10). Typical AOPs include combinations of ozone, hydrogen peroxide, UV radiation, solid catalysts, or sodium hypochlorite. Titanium dioxide photocatalysis, a combination of titanium dioxide and UV radiation (26), may be applicable for smaller utilities.

WATER QUALITY PROTECTION MEASURES

In the third edition of its guidelines for drinking water quality published in 2004, the World Health Organization (WHO) emphasized source water quality protection as a critical element to providing safe drinking water (41). Source water quality is highly variable, with some regions having high quality source waters and naturally low levels of DBP precursors. For example, some waters contain low bromide levels and low levels of total organic carbon. Such high quality waters are less likely to form certain by-products upon disinfection and may require minimal treatment to remove DBP precursors.

Source water protection, such as covered reservoirs, may help reduce DBP levels by limiting the amount of runoff of organic matter into surface water supplies. Other watershed management practices can reduce the amount of dissolved organic carbon in a water source and concomitant production of DBPs (42). Highly protected source waters, such as naturally filtered groundwater, may require minimal disinfection to remove microbial pathogens.

Providing safe, palatable water necessitates the use of a multiple barrier approach in which watershed protection, water filtration and disinfection, and maintenance of the water distribution system are addressed (43). Future trends in drinking water treatment in developed countries may shift the focus from water treatment to pollution prevention (44) and concentrate resources on maintaining outdated distribution systems and protecting the quality of source water.

BIBLIOGRAPHY

1. Nieuwenhuijsen, M.J., Toledano, M.B., Eaton, N.E., Fawell, J., and Elliott, P. (2000). Chlorination DBPs in water and their association with adverse reproductive outcomes: a review. *Occup. Environ. Med.* **57**(2): 73–85.
2. Villanueva, C.M. et al. (2004). DBPs and bladder cancer: a pooled analysis. *Epidemiology* **15**(3): 357–367.
3. U.S. Environmental Protection Agency (2001). Office of Water (4606M), EPA 816-F-01-010, May. *Stage 1 Disinfectants and Disinfection Byproducts Rule: A Quick Reference Guide*. http://www.epa.gov/safewater/mdbp/qrg_st1.pdf.
4. Connell, G.F. (1998). European water disinfection practices parallel U.S. treatment methods. Chlorine Chemistry Council. *Drinking Water Health Q.* **4**(3).
5. Hall, E.L. and Dietrich, A.M. (2000). A brief history of drinking water. *Opflow* **26**(6).
6. U.S. Environmental Protection Agency (1999). Office of Water (4607), EPA 815-R-99-014, April. *Alternative Disinfectants and Oxidants Guidance Manual*. http://www.epa.gov/safewater/mdbp/pdf/alter/cover_al.pdf.
7. AWWA Water Quality Division Disinfection Systems Committee (2000). Committee report: disinfection at large and medium size systems. *J. Am. Water Works Assoc.* **92**(5): 32–43.
8. Chauret, C.P., Radziminski, C.Z., Lepuil, M., Creason, R., and Andrews, R.C. (2001). Chlorine dioxide inactivation of *Cryptosporidium parvum* oocysts and bacterial spore indicators. *Appl. Environ. Microbiol.* **67**(7): 2993–3001.

Next Page

9. Driedger, A.M., Rennecker, J.L., and Marinas, B.J. (2001). Inactivation of *Cryptosporidium parvum* oocysts with ozone and monochloramine at low temperature. *Water Res.* **35**(1): 41–48.
10. Mofidi, A.A. et al. (2002). *Advanced Oxidation Processes and UV Photolysis for Treatment of Drinking Water*. Metropolitan Water District of Southern California, La Verne, California.
11. Momba, M.N.B., Cloete, T.E., Venter, S.N., and Kfir, R. (1998). Evaluation of the impact of disinfection processes on the formation of biofilms in potable surface water distribution systems. *Water Sci. Technol.* **38**(8–9): 283–289.
12. Diehl, A.C. et al. (2000). DBP formation during chloramination. *J. Am. Water Works Assoc.* **92**(6): 76–90.
13. California Department of Health Services (2002). *Studies on the Occurrence of NDMA in Drinking Water*. http://www.dhs.ca.gov/ps/ddwem/chemicals/NDMA/studies.htm.
14. Plewa, M.J. et al. (2004). Chemical and biological characterization of newly discovered iodoacid drinking water DBPs. *Environ. Sci. Technol.* **38**(18): 4713–4722.
15. Aieta, E.M. and Berg, J.D. (1986). A review of chlorine dioxide in drinking water treatment. *J. Am. Water Works Assoc.* **78**(6): 62–72.
16. Hoehn, R.C. and Gates, D.J. (1999). Control of disinfection by-product formation using chlorine dioxide. In: *Formation and Control of Disinfection By-Products*. P.C. Singer (Ed.). American Water Works Association, Denver, pp. 205–221.
17. Gordon, G., Kieffer, R.G., and Rosenblatt, D.H. (1972). The chemistry of chlorine dioxide. *Prog. Inorg. Chem.* **15**: 201–286.
18. Krasner, S.W. et al. (2001). The occurrence of a new generation of DBPs (beyond the ICR). In: *Proceedings of the American Water Works Association Water Quality Technology Conference*. American Water Works Association, Denver.
19. Weinberg, H.S., Krasner, S.W., Richardson, S.D., and Thruston, A.D. Jr. (2002). *The occurrence of disinfection by-products (DBPs) of health concern in drinking water: results of a nationwide DBP occurrence study*, EPA/600/R02/068, U.S. Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA.
20. U.S. Environmental Protection Agency (2003). *40 CFR Parts 141,142 and 143*. National Primary Drinking Water Regulations: Stage 2 Disinfectants and Disinfection Byproducts Rule (Proposed Rules). 68 *Fed. Reg.* 49548 (August 18).
21. Reckhow, D.A. (1999). Control of disinfection by-product formation using ozone. In: *Formation and Control of Disinfection By-Products*. P.C. Singer (Ed.). American Water Works Association, Denver, pp. 179–204.
22. von Gunten, U. (2003). Ozonation of drinking water: Part II. Disinfection and by-product formation in presence of bromide, iodide or chlorine. *Water Res.* **37**(7): 1469–1487.
23. Richardson, S.D. et al. (1999). Identification of new ozone disinfection by-products in drinking water. *Environ. Sci. Technol.* **33**: 3368–3377.
24. Hoigné, J. and Bader, H. (1988). The formation of trichloronitromethane (chloropicrin) and chloroform in a combined ozonation/chlorination treatment of drinking water. *Water Res.* **22**(3): 313.
25. Malley, J.P. (1999). Control of disinfection by-product formation using ultraviolet light. In: *Formation and Control of Disinfection By-Products*. Singer, P.C. (Ed.). American Water Works Association, Denver, pp. 223–235.
26. Richardson, S.D. et al. (1996). Identification of TiO_2/UV disinfection byproducts in drinking water. *Environ. Sci. Technol.* **30**(11): 3327–3334.
27. Stevens, A.A., Slocum, C.J., Seeger, D.R., and Robeck, G.G. (1976). Chlorination of organics in drinking water. *J. Am. Water Works Assoc.* **68**(11): 615–620.
28. Lange, A.L. and Kawczynski, E. (1978). Controlling organics: the Contra Costa County Water District experience. *J. Am. Water Works Assoc.* **70**(11): 653–660.
29. Trussell, R.R. and Umphres, M.D. (1978). The formation of trihalomethanes. *J. Am. Water Works Assoc.* **70**(11): 604–612.
30. Miller, W.J. and Uden, P.C. (1983). Characterization of nonvolatile aqueous chlorination products of humic substances. *Environ. Sci. Technol.* **17**(3): 150–152.
31. Reckhow, D.A. and Singer, P.C. (1985). Mechanisms of organic halide formation during fulvic acid chlorination and implications with respect to preozonation. In: *Water Chlorination: Chemistry, Environmental Impact and Health Effects. Vol. 5*, R.L. Jolley, R.J. Bull, and W.P. Davis (Eds.). Lewis Publishers, Chelsea, MI, pp. 1229–1257.
32. Nieminski, E.C., Chaudhuri, S., and Lamoreaux, T. (1993). The occurrence of DBPs in Utah drinking waters. *J. Am. Water Works Assoc.* **85**(9): 98–105.
33. Singer, P.C. (1994). Control of disinfection by-products in drinking water. *J. Environ. Eng.* **4**: 727–744.
34. Carlson, M. and Hardy, D. (1998). Controlling DBPs with monochloramine. *J. Am. Water Works Assoc.* **90**(2): 95–106.
35. Huehmer, R.P. (2001). *Strategies for Disinfection Byproduct Mitigation in Low-Pressure Membrane Systems*. Presentation at Annual Conference of the Association of State Drinking Water Administrators.
36. U.S. Environmental Protection Agency (1999). Office of Water (4607), EPA 815-R-99-012, May. *Enhanced Coagulation and Enhanced Precipitative Softening Guidance Manual*. http://www.epa.gov/safewater/mdbp/coaguide.pdf.
37. Lykins, B.W. and Clark, R.M. (1988). *GAC for Removing THMs* (EPA 600/9-88/004). U.S. Environmental Protection Agency, Washington, DC.
38. Van Hoof, F., Janssens, J.G., and van Dijck, H. (1985). Formation of mutagenic activity during surface water preozonation and its removal in drinking water treatment. *Chemosphere* **14**(5): 501–509.
39. Sketchell, J., Peterson, H.G., and Christofi, N. (1995). Disinfection by-product formation after biologically assisted GAC treatment of water supplies with different bromide and DOC content. *Water Res.* **29**(12): 2635–2642.
40. Schlager, K.J. (2001). *Final Report: An Electroporation Drinking Water Disinfection System*. http://cfpub.epa.gov/ncer_abstracts/index.cfm/fuseaction/display.abstractDetail/abstract/1525/report/F.
41. World Health Organization (2004). *WHO Guidelines for Drinking-Water Quality*, 3rd Edn. WHO, Geneva. http://www.who.int/water_sanitation_health/dwq/en/gdwq3_contents.pdf.
42. Fleck, J.A., Bossio, D.A., and Fujii, R. (2004). Dissolved organic carbon and disinfection by-product precursor release from managed peat soils. *J. Environ. Qual.* **33**(2): 465–475.
43. Ford, T.E. (1999). Microbiological safety of drinking water: United States and global perspectives. *Environ. Health Perspect.* **107**(S1): 191–206.
44. U.S. Environmental Protection Agency (2000). Office of Water (4606M), EPA 816-F-00-002, *Drinking Water: Past, Present, and Future*. February. http://www.epa.gov/safewater/consumer/dwppf.pdf.

WATER QUALITY ASPECTS OF DREDGED SEDIMENT MANAGEMENT

Anne Jones-Lee
G. Fred Lee
G. Fred Lee & Associates
El Macero, California

INTRODUCTION

The United States has 25,000 miles of navigable waterways, which include coastal waterways, such as the Intercoastal Waterway along the Eastern and Gulf Coasts; the major inland rivers, such as the Mississippi and Ohio; and major ports and harbors, such as New York, Galveston, Houston, Los Angeles, San Diego Bay, San Francisco Bay, Puget Sound, and those of the United States–Canadian Great Lakes. The U.S. Army Corps of Engineers (COE) is charged by the U.S. Congress with maintaining the navigation depth of these waterways as they fill with sediment. Maintaining navigation depth is considered to be of national importance to minimize transportation costs of bulk materials. To accomplish this, the COE dredges, or issues permits for contractors to dredge, about 500 million yd^3/yr of sediment from major waterways as well as smaller harbors, marinas, ports, and channels.

In the 1960s, it was beginning to be recognized that the sediments of some U.S. waterways were highly contaminated with chemicals that have the potential to be pollutants—that is, to impair the beneficial uses of waterbodies as domestic water supply sources or for aquatic life/fisheries. As the conventional method of disposal of dredged sediment was to dump it in deeper open water, concern developed about the solubilization and desorption of chemical constituents that were associated with sediment particles when sediments were dredged and then introduced into a watercolumn at a disposal site. In response to water quality impact concerns, the U.S. Army Corps of Engineers developed a $30 million five-year Dredged Material Research Program (DMRP) to investigate the water quality significance of potential pollutants in dredged sediments. As part of that program, the authors conducted extensive laboratory and field studies to investigate the release of sediment-associated contaminants and their impacts associated with open water disposal of dredged sediment (1–3), results of which are summarized or referenced herein.

IMPACT OF TYPE OF DREDGING/DISPOSAL OPERATION

The development of an approach to evaluate and regulate potential water quality impacts of dredging and dredged sediment disposal requires an understanding of what transpires during such operations. In his *Handbook of Dredging Engineering*, Herbich (4) described the various methods of dredging and dredged sediment disposal. Dredging is typically accomplished by either mechanical or hydraulic means.

Mechanical Dredging/Barge Disposal

Mechanical dredging typically employs a clamshell or dragline bucket to remove sediments, which are then deposited onto a barge. Water introduced during dredging drains off the barge. After the barge is towed to the designated disposal site, the sediment is released and drops as a fairly cohesive mass to the bottom. As limited mixing occurs of the sediments being dredged with water, mechanical dredging results in limited opportunity for release of constituents from the sediments to the water, which limits the potential for sediment-associated contaminants to adversely impact water quality.

Hydraulic Dredging

Hydraulic dredging involves sucking up the sediment as a slurry (Fig. 1) and pumping it to an open water or confined on land disposal area or into the holding hoppers of a hopper dredge for transport to an open water disposal site. The sediment:water ratio of the slurry is characteristically 1:4 by volume. The slurrying results in the mixing of the sediments with water, which tends to promote the release of constituents associated with the sediment and interstitial water into the slurry water. If hopper dredge is used, the excess water is typically allowed to drain off at the dredging site or during transport to the disposal site to optimize efficiency of sediment removal.

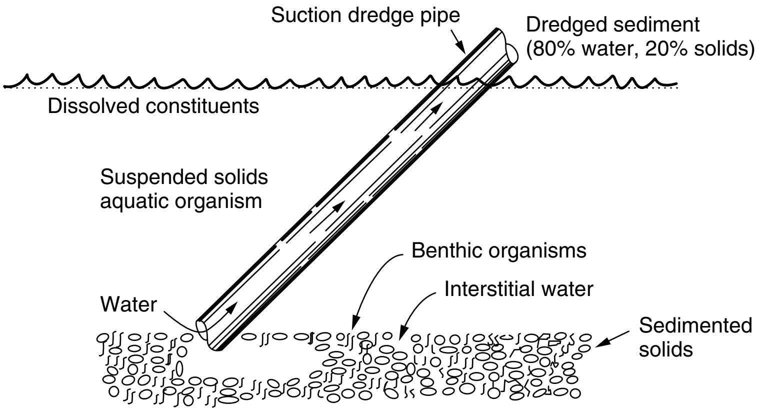

Figure 1. Diagrammatic representation of hydraulic dredging.

Hopper-Dredge Disposal

Figure 2 illustrates the watercolumn during disposal of dredged sediment from a hopper dredge. When the hopper dredge is at the designated disposal location, the hopper doors are opened and the sediment is released. Most of the sediment rapidly descends to the bottom as a fairly cohesive mass. As indicated in Fig. 2, the descent of the dredged sediment from the hopper to the bottom of the watercolumn forms a turbid cloud in the waters of the region.

In their evaluation of the nature and water quality impacts of dredging and dredged sediment disposal, Lee et al. (1) and Jones and Lee (2) monitored the watercolumn during more than 10 open water disposal operations by measuring more than 30 physical and chemical parameters at various depths in the watercolumn before, during, and after the passage of the turbid plume for each. Figure 3 shows the characteristic pattern of turbidity marking the passage of the turbid plume during open water disposal of dredged sediment. As indicated in Fig. 3, near the surface (2-meter depth) the turbidity persisted at a location a few tens of meters downcurrent from the dump for about 2 minutes. Near the bottom, at 14 meters, the turbid plume turbidity persisted for about 7 minutes. The plume typically dispersed to indistinguishable from ambient turbidity of the disposal area in about one hour.

Lee et al. (1) found that, in some instances, some increase occurred in concentration of some chemical constituents in this turbid cloud as well as some decrease in the dissolved oxygen (DO) concentration, as is shown in Fig. 4.

In the studies conducted by Lee et al. (1), the only constituent of potential concern that was released was ammonia. The potential impact of ammonia on aquatic organisms at an open water disposal site would have to be reviewed on a site-by-site basis, because it would be controlled largely by the sensitivity of the organisms there and the rate of dilution.

The turbid plume associated with open water dumping of mechanically dredged sediment is even less pronounced than with hopper-dredge disposal, because the

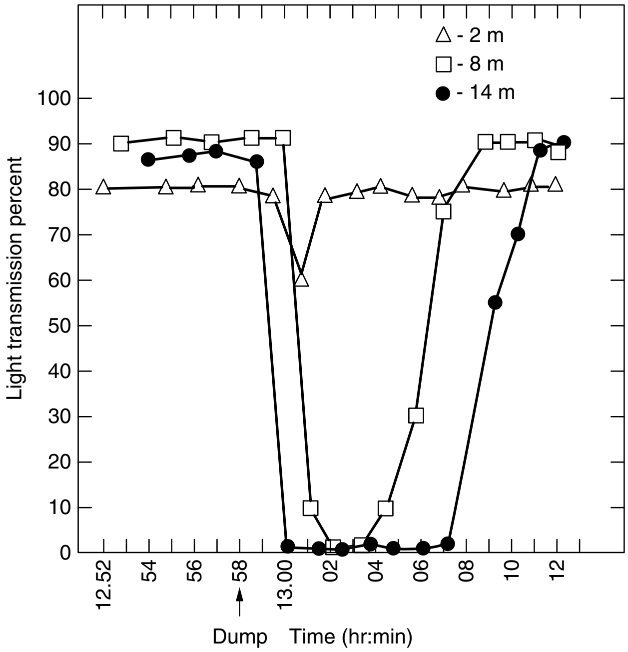

Figure 3. Percent light transmission with depth during passage of turbid plume from open water disposal of hopper-dredged sediment.

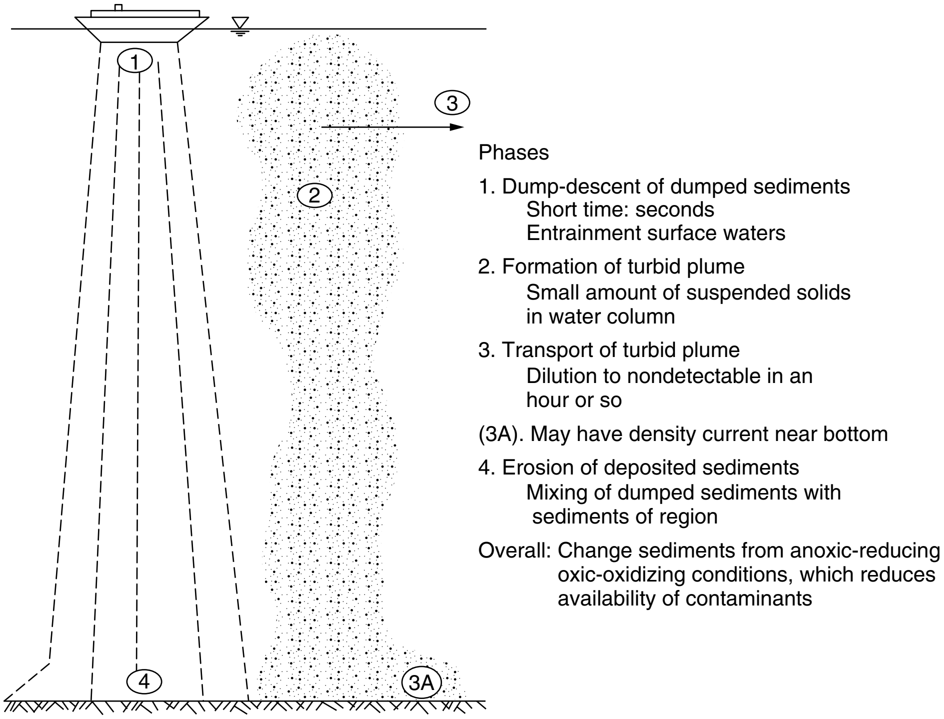

Figure 2. Hopper-dredge disposal of dredged sediment.

mechanically dredged sediment is a more cohesive mass. It was concluded by Lee et al. (1) that open water disposal of even contaminated sediments would not be expected to cause water quality problems in the disposal site water-column because of the limited releases and the short exposures that aquatic organisms could experience from such releases.

Pipeline Disposal

The open water pipeline disposal of hydraulically dredged sediments, illustrated in Fig. 5, presents a situation significantly different from that of dumping of mechanically or hopper-dredged sediments in open waters. In this process, the dredged sediment slurry is transported via pipeline to the discharge location where it is discharged. The slurry sinks to the bottom and moves downcurrent as an approximately one-meter-thick density current along the bottom. This density current has been found to persist for thousands of meters from the point of discharge during disposal operations. As this dredged sediment density current is typically characterized by low DO and/or the release of constituents such as ammonia, its presence at a location for extended periods of time (many hours to a day or so) could represent a significant adverse impact on aquatic organisms residing on the bottom in the path of the current. Ordinarily, however, such density currents do not persist in one location for extended periods because of the intermittent nature of pipeline disposal operations owing to frequent mechanical adjustments and the movement of the dredge. Further, the point of discharge is frequently moved because of the accumulations of sediments near the point of discharge.

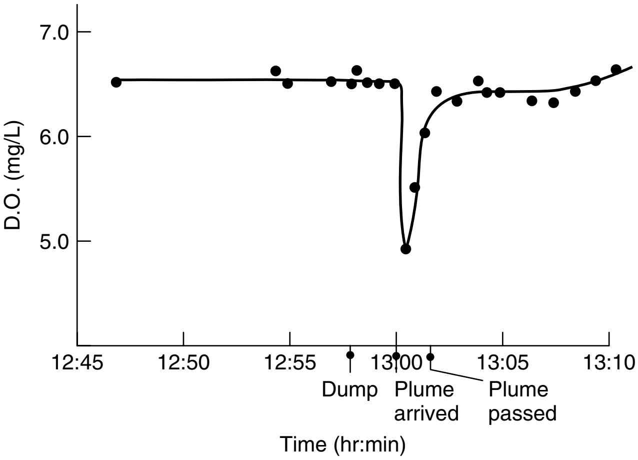

Figure 4. Dissolved oxygen concentration in near-surface waters during passage of turbid plume from hopper-dredge disposal of dredged sediment.

Upland-"Confined" Disposal of Dredged Sediments

It has been commonly assumed that "confined" disposal of dredged sediment on land or behind a dike to form an island in water is less environmentally damaging than open water disposal. This assumption misjudges the adverse impacts of open water disposal and misjudges the water quality protection afforded by "confined" disposal. It is now more widely recognized that upland disposal of contaminated dredged sediments has a greater potential for causing adverse environmental impact than open water disposal of dredged sediments.

As generally practiced, "confined" disposal does not truly confine deposited sediment or eliminate adverse impact. Rather, it provides a settling area where the larger, denser particles settle and the supernatant water and fine materials from the hydraulic dredging overflow the confinement and enter the nearshore areas of the nearby watercourse. Ironically, it is this overflow water that would represent a threat to water quality. It is the fine materials that do not readily settle in the disposal area that contain chemicals that could become available to adversely affect water quality. The area where the confined disposal overflow occurs (i.e., nearshore) is generally the

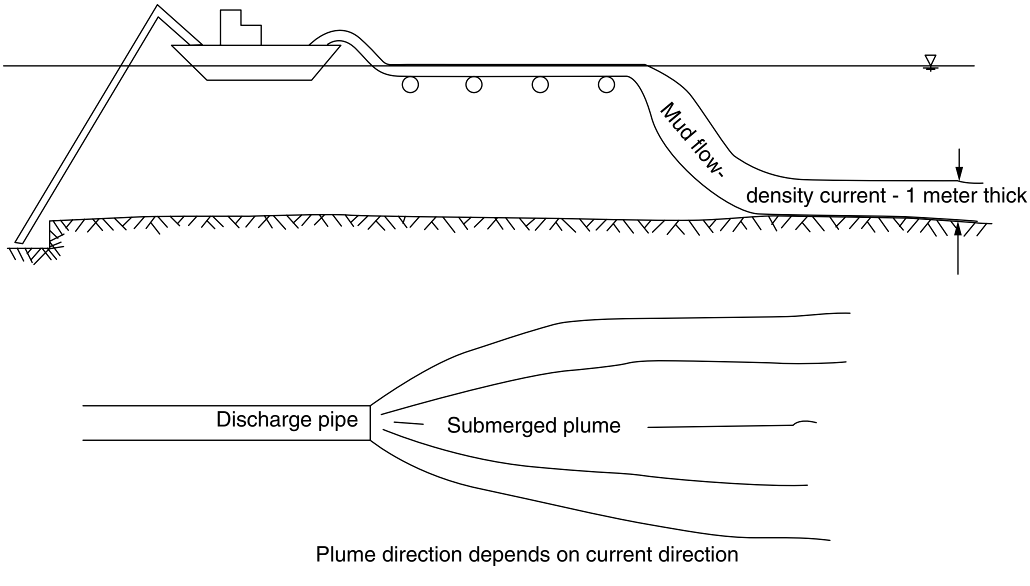

Figure 5. Hydraulic dredging with open water pipeline disposal.

most ecologically sensitive area of the waterbody. Thus, rather than being more protective of ecological/water quality, "confined" disposal operations pose a potentially greater risk than open water disposal.

Potential adverse impacts associated with the materials that remain in the confined area also exist. Studies conducted at the COE Waterways Experiment Station (5) have shown that when the sediments in a confined disposal area dry, they tend to release constituents such as heavy metals, which appears to be related to the oxidation of the amorphous sulfides present in the sediments, which keep the heavy metals in particulate form while the sediment is wet, and the development of acidic conditions. It may also be related to the aging of the ferric hydroxide precipitate (hydrous oxides). Although freshly precipitated ferric hydroxide has a substantial holding power for constituents, aged precipitates, especially those that dry out, lose some of this holding power. It is therefore not surprising to find that dredged sediments that dry in a confined disposal area release appreciable heavy metals when the area is subjected to precipitation or receives additional dredged sediment. The contaminants released then drain from the confinement area to the nearby nearshore waters. The potential for chemicals released from the confined dredged sediments to pollute area groundwater also exists. It is important, therefore, that those who advocate upland disposal of contaminated dredged sediments conduct a proper, critical review and evaluation of the potential adverse impacts of the constituents associated with the sediments that leave the confined disposal area during overflow during dredging operations or via drainage or seepage from the disposal site.

Contaminant Release from Redeposited Sediment

In addition to concern existing about the potential watercolumn impacts of contaminants in dredged sediment, a need exists to evaluate the potential for chemicals in the redeposited dredged material to be released to the watercolumn at the disposal site at sufficient rate to be adverse to water quality. In order to assess whether measured or unmeasured chemicals could have an adverse impact on watercolumn organisms or organisms that might colonize the redeposited dredged sediments shortly after deposition, Lee et al. (1) developed a dredged sediment elutriate screening toxicity test. The dredged sediment toxicity screening test involves introduction of standard test organisms into the settled elutriate from a standard oxic elutriate test procedure (oxic mixing of 1 volume sediment with 4 volumes water and oxic settling). Grass shrimp (*P. pugio*) were used for marine conditions and daphnids were used for freshwater. Lee et al. (1) and Jones and Lee (2) found that sediments collected near urban and industrial areas showed toxicity to aquatic life under the conditions of the laboratory test. Typically, in tests in which toxicity was found, from 10% to 50% of the test organisms were killed in the 96-hour test period. Toxicity was substantially less than may have been expected based on the elevated concentrations of the myriad measured and unknown contaminants in the sediments. Thus, although many of the sediments tested had high concentrations of heavy metals and chlorinated hydrocarbons of various types, those constituents were present in the sediments largely in unavailable, nontoxic forms. In the case of contaminated New York Harbor sediments, Jones and Lee (6) and Jones-Lee and Lee (7) found that the toxicity discovered in the toxicity testing was because of ammonia. Other investigators are also finding that ammonia is one of the principal causes of sediment toxicity to aquatic life. In general, however, it has been found that disposal sites are of sufficiently high energy and the rates of release of pollutants sufficiently slow that water quality problems in the watercolumn do not occur from redeposited sediments.

However, the bioaccumulation or buildup of chemical contaminants in higher trophic-level aquatic organisms that recolonize the disposal site is of concern. A few constituents exist, including methylmercury, PCBs, and PAHs, that have the potential for trophic magnification in aquatic environments. The bioaccumulation of constituents within organisms occurs from both direct uptake of constituents from the water and the consumption of particles that have constituents associated with them as well as the consumption of other organisms. The bioaccumulation and subsequent impact of chemical contaminants cannot be predicted based on their concentration in the sediment. Evaluation of this potential impact needs to be made on a site-specific basis using organism bioaccumulation tests of the type described by the U.S. EPA and COE (8,9).

REGULATING DREDGED SEDIMENT DISPOSAL

In the 1972 amendments to the Federal Water Pollution Control Act (PL 92–500), Congress specified in Section 404 that the disposal of dredged sediments in U.S. waters may take place, provided that an avoidance of "unacceptable effects" exists. It further stated that the disposal of dredged sediments should not result in violation of applicable water quality standards after consideration of dispersion and dilution, toxic effluent standards, or marine sanctuary requirements, and should not jeopardize the existence of endangered species. In order to implement these regulations, the U.S. EPA and Corps of Engineers developed two dredged sediment testing manuals, *Evaluation of Dredged Material Proposed for Ocean Disposal - Ocean Testing Manual* (8), and *Evaluation of Dredged Material Proposed for Discharge in Waters of the U.S. - Inland Testing Manual* (9). The Ocean Testing Manual was developed to implement Section 103 of Public Law 92–532 (the Marine Protection, Research, and Sanctuaries Act of 1972). These manuals prescribe testing procedures to assess biological effects such as toxicity and bioaccumulation of hazardous chemicals from dredged sediments.

The reliable evaluation of the potential adverse effects that may occur from dredging and dredged sediment disposal in a particular manner requires effort and understanding of the biological, chemical, and physical processes involved. It has been long recognized that potential adverse impacts cannot be determined based on the bulk chemical composition of the sediment; this understanding had, in fact, led to the DMRP studies of the 1970s. However, in an attempt to simplify the process and make determinations more administratively expedient, the U.S. EPA and other regulatory agencies have again attempted to regulate dredged sediment disposal based on chemical concentration-based "sediment quality criteria." One such approach that has been advanced is based on the "co-occurrence" in a sediment of a total concentration of a chemical and a biological response not even necessarily caused by the chemical of concern. Lee and Jones-Lee (10) reported on their reviews of co-occurrence-based so-called sediment quality guidelines and discussed why they are technically invalid and cannot provide a reliable evaluation for any purpose. Lee and Jones-Lee (11) recommended that a "best professional judgment" triad, weight-of-evidence approach be used to evaluate the water quality significance of chemicals in aquatic sediments. This approach is based on the integrated use of aquatic toxicity, bioaccumulation, organism assemblage information, and physical and chemical information focused on identifying the cause of aquatic life toxicity. Additional information on the water quality impacts of chemicals in sediments is available from Jones-Lee and Lee (7), DOER (12), and in the U.S. Army Corps of Engineers bulletin, "Dredging Research," available on request from www.wes.army.mil/el/dots/drieb.html.

DREDGING FOR SEDIMENT CLEANUP

In addition to dredging for maintaining navigation depth of waterways, dredging is being undertaken to remove (remediate) contaminant-laden sediments from Superfund and other sites. Evison of the U.S. EPA Office of Emergency and Remedial Response summarized the magnitude of the problem of contaminated sediments at Superfund sites in a presentation entitled, "Contaminated Sediment at Superfund Sites: What We Know So Far" at a U.S. EPA and U.S. Army Corps of Engineers national workshop (13). At the same workshop, Ellis (14), Sediments Team Leader with the U.S. EPA Office of Emergency and Remedial Response, presented a discussion entitled "Superfund Cleanup Issues at Contaminated Sediment Sites." The dredging of contaminated "Superfund" sediments and their disposal require attention to the same water quality/environmental quality, and contaminant availability, transport, and impact issues that are faced in dredging of contaminated sediment for channel depth maintenance.

An example of the use of dredging for sediment remediation is the current cleanup of PCB-contaminated sediments in the Hudson River in New York State. The Hudson River sediments from Hudson Falls to New York City (a distance of 200 miles) were contaminated by PCBs discharged for decades in wastewaters by General Electric Company. U.S. EPA (15) provides information on the past and proposed dredging of the PCB-contaminated sediments in that area. Another example of Superfund dredging is the dredging of PCB-contaminated sediments from the Wisconsin Fox River (16). There, an issue of considerable concern is the appropriateness of placing PCB-contaminated sediments in municipal landfills.

BIBLIOGRAPHY

1. Lee, G.F., Jones, R.A., Saleh, F.Y., Mariani, G. M., Homer, D.H., Butler, J. S., and Bandyopadhyay, P. (1978). Evaluation of the Elutriate Test as a Method of Predicting Contaminant Release during Open Water Disposal of Dredged Sediment and Environmental Impact of Open Water Dredged Materials Disposal, Vol. II: Data Report. Technical Report D-78-45, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
2. Jones, R.A. and Lee, G.F. (1978). Evaluation of the Elutriate Test as a Method of Predicting Contaminant Release during Open Water Disposal of Dredged Sediment and Environmental Impact of Open Water Dredged Material Disposal, Vol. I: Discussion. Technical Report D-78-45, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
3. Lee, G.F. and Jones-Lee, A. (2000). Water quality aspects of dredging and dredged sediment disposal. In: *Handbook of Dredging Engineering*, 2nd Edn. J.B. Herbich (Ed.). McGraw Hill, New York, pp. 14.1–14.42. Available at http://www.gfredlee.com/dredging.html.
4. Herbich, J.B. (Ed.). (2000). *Handbook of Dredging Engineering*, 2nd Edn. McGraw Hill, New York.
5. Palermo, M.R. (1986). Development of a Modified Elutriate Test for Estimating the Quality of Effluent from Confined Dredged Material Disposal Areas. Technical Report D-86-4, Department of the Army, U.S. Army Corps of Engineers, Washington, DC.
6. Jones, R.A. and Lee, G.F. (1988). Toxicity of U.S. waterway sediments with particular reference to the New York Harbor area. In: *Chemical and Biological Characterization of Sludges, Sediments, Dredge Spoils and Drilling Muds*. ASTM STP 976 American Society for Testing Materials, Philadelphia, PA, pp. 403–417.
7. Jones-Lee, A. and Lee, G.F. (1993). Potential significance of ammonia as a toxicant in aquatic sediments. In: *Proc. First International Specialized Conference on Contaminated Aquatic Sediments: Historical Records, Environmental Impact, and Remediation*. International Assoc. Water Quality, Milwaukee, WI, pp. 223–232, Available at http://www.gfredlee.com/psedqual2.htm.
8. U.S. EPA and COE (U.S. Army Corps of Engineers). (1991). *Evaluation of Dredged Material Proposed for Ocean Disposal—Ocean Testing Manual*. U.S. EPA Office of Water (WH-556F), EPA-503-8-91/001, Washington, DC.
9. U.S. EPA and COE (U.S. Army Corps of Engineers). (1998). *Evaluation of Dredged Material Proposed for Discharge in Waters of the U.S.—Inland Testing Manual*. Office of Water, EPA-823-B-98-004, Washington, DC.
10. Lee, G.F. and Jones-Lee, A. (2004). Unreliability of Co-Occurrence-Based Sediment Quality 'Guidelines' for Evaluating the Water Quality Impacts of Sediment-Associated Chemicals. Report of G. Fred Lee & Associates, El Macero, CA. Available at http://www.gfredlee.com/psedqual2.htm.
11. Lee, G.F. and Jones-Lee, A. (2004). Appropriate Incorporation of Chemical Information in a Best Professional Judgment 'Triad' Weight of Evidence Evaluation of Sediment Quality, Presented at the 2002 Fifth International Symposium on Sediment Quality Assessment (SQA5), In: *Aquatic Ecosystem Health and Management* M. Munawar (Ed.). **7**(3): 351–356. Available at http://www.gfredlee.com/BPJWOEpaper-pdf.
12. DOER (U.S. Army Corps of Engineers Dredging Operations and Environmental Research). (2002). Program Products of U.S. Army Corps of Engineers Operation and Maintenance

Navigation Program, Volume II (FY 1997–2002), CD available from U.S. Army Engineer Research and Development Center (ERDC) Waterways Experiment Station, Vicksburg, MS. Available at http://www.wes.army.mil/el/dots/doer.

13. Evison, L. (2003). Contaminated sediment at superfund sites: what we know so far. Presented at U.S. EPA and U.S. Army Corps of Engineers national workshop *Environmental Stability of Chemicals in Sediments*. San Diego, CA.
14. Ellis, S. (2003). Superfund cleanup issues at contaminated sediment sites. Presented at U.S. EPA and U.S. Army Corps of Engineers national workshop *Environmental Stability of Chemicals in Sediments*. San Diego, CA. Available at http://www.sediments.org/sedstab/agenda.pdf.
15. U.S. EPA. (2004). Hudson River PCBs. U.S. Environmental Protection Agency Region 2 Superfund, New York. Available at http://www.epa.gov/region02/superfund/npl/0202229c.htm.
16. U.S. EPA. (2004). Lower Fox River and Green Bay Area of Concern. U.S. Environmental Protection Agency Region 5, Chicago, IL. Available at http://www.epa.gov/glnpo/aoc/greenbay.html.

THE ECONOMICS OF WATER QUALITY

Keith O. Keplinger
Texas Institute for Applied
Environmental Research
Stephenville, Texas

The role of economics in water quality decision-making is a topic receiving increasing attention today. A number of factors have spurred this interest. On the policy front, the total maximum daily load program has emerged as a premier water quality program in the United States, engendering fresh opportunities for economic analysis, because the means of attaining ambient targets are not prescribed. The increased relevance of economic analysis in the policy arena has been accompanied by a blossoming of the discipline of environmental economics and greater public and governmental acceptance of pursuing cost-effective and market-based solutions to environmental problems.

The theory of environmental economics has been extensively developed, but the utility and form of policy instruments are highly influenced by the institutional landscape. Consequently, this report begins with a brief assessment of water quality policy in the United States, focusing on the landmark legislation that constitutes the lion's share of current water quality policy, the 1972 Clean Water Act (CWA). Next, the meaning of costs and benefits in a water quality context are discussed, and the merits of two commonly used economic instruments in water quality assessments are explored: cost–benefit analysis and cost-effectiveness analysis. Finally, equity issues and the economic implications of water quality policy alternatives are discussed, with particular emphasis on the historical development and potential for incentive-based mechanisms.

WATER QUALITY POLICY IN THE UNITED STATES

Federal water quality policy in the United States can be succinctly characterized as initially state-based and ambient-based, then switching to a federal system dominated by technology standards with passage of the 1972 CWA. More recently, federal water quality policy has seen a return to the ambient-based approach for impaired waters as the total maximum daily load (TMDL) program has emerged. An ambient-based water quality approach targets the attainment of a specified level of ambient (in-stream) water quality as its objective. Loads or concentrations of pollutants are typically used as water quality measures. Ambient-based water quality approaches contrast with technology-based approaches, which specify the application of controls (e.g., best available technology), based on industrial classification, without explicit regard for the ambient conditions of receiving waters.

A legal scholar has noted that the breathtakingly ambitious goals of the CWA suggest a virtually cost-blind determination to control water pollution (1). It might also be added that the CWA takes a benefits-blind approach as well. Although not completely absent, references to costs or benefits in the CWA are few. The major regulatory mechanism of the CWA, the National Pollutant Discharge Elimination System (NPDES), focuses on technical feasibility as the overriding factor in establishing pollution controls. Considering the political and historical context surrounding the passage of the CWA, this largely prescriptive approach to restoring water quality can be considered neither surprising nor irrational. After years of inaction under prior national clean water legislation in the face of deteriorating water quality and mounting political pressure, the Senate was keen to pass legislation that forced action (2).

Portney (3) opines that

> When the U.S. Congress drafted the environmental legislation of the early 1970's, these and other problems were thought to be so obvious and so serious—and, in fact, they were—that members simply could not envision directing EPA to address them only if it was affordable to do so. As a consequence, most of the early laws effectively directed the agency's administrator to disregard cost altogether in setting ambient air quality standards and some water quality standards.

Houck (2) attributes the CWA's shift in focus to technology standards to “a new ethical premise, that water should simply be clean.”

Unlike the former state administered ambient water quality approach, the prescriptive NPDES approach largely dispensed with the need for scientific endeavor and did not require the monitoring of water quality, estimation of the environmental impacts of control actions, or the estimation of economic costs and benefits.

During debate of the 1972 amendments, classic economic arguments were put forth in favor of continuing ambient-based federal water quality legislation. For example, in House testimony, the Chairman of the Council of Economic Advisors (CEA) stated:

> CEA is in agreement with the use of water quality targets appropriate to the conditions and expected uses of water in particular areas of the country. That is basic to the concept of relating the costs of programs to the benefits received from them. To abandon that concept for a nationally legislated standard which focuses on the level of pollutants removed and is unrelated to water quality uses and standards is economically unwise because it means a necessary misallocation of our inevitably scarce economic resources (4).

Countering this line of argumentation was a consensus, which ultimately prevailed, that the existing state-based water quality approach, which relied on ambient standards, had been given ample time to succeed, but had overwhelmingly failed. “The Senate found the standards weak, late, widely disparate, scientifically doubtful, largely unenforced, and probably unenforceable” (2). Thus, whatever the economic and scientific demerits of the prescriptive NPDES approach, it was at least viewed as an approach that could be effectively implemented and would actually produce cleaner water—the benefits of which were, presumably, greater than the administrative and social costs of the program.

Of the two approaches, it can be argued that the NPDES approach was preferable only because it actually produced quantifiable water quality improvements, whereas the former federal legislation was largely effete. Tietenburg (5) concludes, “the wrong inference was drawn from the early lack of legislative success [in achieving water quality],” suggesting that equal or greater water quality gains could have been made at a lesser cost if correctly structured and adequately enforced incentive-based programs had been implemented. It can also be argued, however, that prescriptive regulation may be preferable when the problem is dire, the benefits great, and when scientific data are lacking, conditions which appear to have largely existed when NPDES was drafted. The National Research Council (6), however, concludes, “the data and science have progressed sufficiently over the past 35 years to support the nation’s return to ambient-based water quality management.” The initial implementation of NPDES and subsequent return to an ambient water quality approach can also be seen within a historical context. Ackerman and Stewart (7) suggest that the prescriptive technology approach “made some sense as a crude first-generation strategy.” Portney (3) attributes the growing role of economics in environmental decision-making to often much higher costs and fewer benefits of additional pollution reduction, the impact of global integration, and to the loss (since the 1970s) of moral imperative status for pollution control.

It is also likely that Congress and the American public believed that effluent limitations based on best available technology applied to point sources would largely solve the nation’s water quality problems. Nonetheless, Section 303(d) of the CWA, the legislation establishing the TMDL program, was included as a backup measure to achieve state water quality criteria if technology measures failed. Reminiscent of and in fact modeled after the 1965 Clean Water Act, Section 303(d) enumerates a number of steps designed to achieve ambient water quality targets consistent with designated uses. The TMDL program, thus, exemplifies an ambient-based approach to water quality.

Largely ignored for the first two decades after passage of the 1972 CWA, a series of lawsuits starting in the 1980s forced both the states and the U.S. Environmental Protection Agency (EPA) to gear up and prepare for the onslaught of TMDL activity to come. The trickle of TMDLs seen in the 1980s has swelled to the point where TMDLs have become a prominent (and controversial) component of CWA compliance. Recent state lists indicate in excess of 21,000 impaired water bodies throughout the nation requiring in excess of 40,000 TMDLs (6). These figures suggest that the ambient-based TMDL program will continue as a major component of water quality control for decades to come.

Although Section 303(d) does not explicitly mention cost considerations, the nature of the legislation facilitates an important role for economic analysis. The goal of attaining quantitative water quality targets is paramount, but the means of attainment are unaddressed, thereby permitting a potentially broad array of attainment strategies; each has its own unique set of control measures and associated economic costs. Consequently, an analysis of viable alternatives, including their economic consequences, can play an important role within the TMDL paradigm.

AN ECONOMIC CHARACTERIZATION OF WATER QUALITY

Water pollution is a classic externality produced by various types of economic activity. An industrial facility, for example, may discharge toxic effluent into a river causing fish kills and environmental impairment further downstream. Because the fisherman and recreationists that suffer harm are not parties to the plant’s decision to discharge, that harm is external to the decision process. Another key feature of water pollution is that, in almost all cases, it harms many users and nonusers simultaneously; hence, in economic lingo, it is a *public* bad. From the reverse perspective, clean water is a public good that produces benefits to often-large groups of users but requires costs to attain it. The externality and the public good nature of water quality/water pollution often lead to market failure—a situation where market forces alone do not result in socially desirable outcomes. The major causes for market failure in water quality are undefined water quality rights and control obligations, and the practical obstacles that discourage numerous individuals who receive relatively small water quality benefits from coordinating their efforts and directly contracting with discharges. Societies have sought to correct market failures through various forms of social intervention. Although not exhaustive, alternative externality correction devices include market emergence, merger, economic incentives, regulation, prohibition, pseudomarkets, and moral suasion.

COMPONENTS OF ECONOMIC ANALYSIS

The two fundamental components of applied microeconomic analyses are costs and benefits. Water pollution involves sacrifices for some groups (costs) and economically benefits others. The benefits of attaining clean water are equivalent to the costs of water pollution, whereas the cost of attaining clean water is equivalent to the benefits that accrue to the production of water pollution. To avoid confusion, the following sections discuss costs and benefits from a water quality perspective.

Costs

Water quality is a public good requiring costs for its attainment. Due to the constraints imposed by nature and current technology, producing more of one good, including water quality, is typically accomplished only at the expense of producing less of other goods. Occasionally, analyses suggest that management or structural measures designed to improve water quality may also result in net benefits to dischargers. Such “win–win” solutions are the exception but might occur for one of two reasons: (1) current technology and resources are not being efficiently utilized, or (2) recent technology has opened up new opportunities for improving water quality while simultaneously producing benefits in other areas. Numerous water pollution control measures have been identified for municipalities, industry, and agriculture. Economic and engineering studies often estimate the costs of implementing specified measures and typically express those costs in monetary units, for example, dollars, which represent the foregone opportunity of producing and consuming other goods.

Benefits

The benefits of mitigating water pollution can be classified as environmental and economic. Environmental benefits are measured by a large number of environmental indexes, such as biological diversity and increased fish populations. Economic benefits are the monetary values society is willing to pay for the associated environmental benefits. Numerous studies have quantified the economic benefits of water quality (8,9). Following Freeman’s taxonomy, benefits include improvements in (1) recreation (swimming, fishing, boating, water fowl hunting), (2) nonuser benefits (amenity, aesthetic, and ecological benefits not directly associated with activities on or near a waterbody, but for which households may be willing to pay), (3) diversionary uses (reducing risk to human health and decreased costs for municipal water supply treatment), and (4) commercial fisheries. Many of the benefits listed above are nonmarket; that is, they are not purchased or sold in ordinary markets and therefore have no observable price. Lack of organized markets for water quality complicates the estimation of water quality benefits (10), though a number of indirect methods have been developed.

ECONOMIC INSTRUMENTS FOR INFORMING WATER QUALITY POLICY

Cost–Benefit Analysis

Applied to water quality, cost–benefit analyses generally estimate and compare both the economic costs and benefits of a water quality project. Such analyses are nominally prescriptive; benefits that exceed costs may represent a sufficient condition for proceeding with a project. A more exacting and universally accepted welfare economic criterion is the Pareto criterion, which states that everyone should be made at least as well off while nobody is made worse off. A policy producing net benefits (total benefits exceeding total costs) is considered a *potential* Pareto improvement because, in theory, the gains of the policy action can be redistributed such that everyone is made at least as well off as before. In practice, a strict Pareto improvement is difficult to achieve and seldom attempted. Thus, a potential Pareto improvement, that is, the generation of positive net benefits, is often used as a decision rule.

In theory, the optimal level of ambient water quality is such that the benefits produced by one additional unit of water quality just equals the costs to attain it. Thus, to fully assess the merits of a water quality project, detailed information on both its benefits and costs is required. Significant difficulties, limitations, and complications inherent in environmental and economic benefit estimation techniques, as described in succeeding sections, suggest one reason why cost–benefit analysis has not factored more prominently in water quality policy.

It is noteworthy that authors of environmental economics textbooks, for the reasons cited, typically exhort students to be cautious in their interpretation of cost–benefit analyses. For example, “Although cost-benefit analysis, with its reliance on numbers and economic theory, may seem as if it might be a precise science, this is not the case” (11). On somewhat of a counterpoint, Hartwick and Olewiler (12) argue, “Cost benefit analysis has been maligned by those who say that it is useless because the numbers are so bad. This criticism misses the point of the exercise. ... The estimate of the benefits of environmental control is crucial if we are to evaluate the tremendous costs of reducing pollution and make informed judgments about the social value of improving environmental quality.”

When benefits cannot be adequately quantified because of uncertainty or other technical hurdles, benefits can only be described. By contrast, estimating the costs of structural or managerial pollution control measures is relatively straightforward and generally involves much less uncertainty. Comparing quantified costs to descriptive, though potentially large, benefits can unduly discount benefits. Although they broadly support the use of cost–benefit analysis, a group of highly respected economists also highlighted some of its limitations in a draft set of principles (13). One of many caveats was that “care should be taken to assure that quantitative factors do not dominate important qualitative factors in decision-making” (14).

Another factor hindering the use of formal cost–benefit analyses as a decision-making tool is the perception that such analyses are prone to manipulation, especially when gaps in knowledge and sources of uncertainty are substantial. Recounting his experience as the senior economist for environmental and resource policy at the Council of

Economic Advisors, Jason Shogren (15) comments: “But now I understand that some people see our orthodoxy as not just simply confining but as downright prehistoric. . . .Because after all, cost–benefit analysis is just naked self-interest dressed up in banker’s pajamas, isn’t it?” Many studies, for instance, have indicated that the nonuse or existence value of environmental amenities may far outweigh their commercial and use value, while other studies omit the estimation of these values altogether.

A final factor discouraging the more widespread use of cost–benefit analysis for water quality is that the estimation of benefits has no practical relevance in either technology-based or ambient-based water quality policy. NPDES technology-based controls, for instance, are mandated on the basis of industrial classification and available technology , while TMDL ambient-based controls are designed to achieve ambient criteria, regardless of the economic benefits produced.

Estimating the Economic Benefits of Water Quality. Three linkages must be established and quantified to estimate water quality benefits: (1) the link between a pollution mitigation strategy and water quality indicators, (2) the link between water quality indicators and environmental benefits, and (3) the translation of environmental benefits into economic benefits. The first linkage is often estimated by simulating a pollution mitigation strategy by using a water quality model. Biophysical models may also be employed to estimate environmental benefits such as increased fish populations. Estimating economic benefits requires knowledge of both environmental benefits and how society values those benefits.

When water pollution causes direct damage to a market commodity resulting in either a decrease in production (e.g., in fisheries) or an increase in expense (e.g., increased treatment cost for municipal drinking water), calculating the value of water quality improvements is relatively straightforward. Because of the public good nature of many water quality benefits, however, many benefits are not directly reflected in markets. Economists have developed a number of indirect methods for estimating nonmarket values of water quality improvement. The hedonistic technique is often used to reveal a lower bound on amenity or aesthetic value. This technique infers the value of a nonmarket commodity by analyzing the value of a commodity whose value is influenced by the nonmarket commodity, for example, comparing waterfront property values across regions experiencing different levels of water pollution. Differences in human behavior, such as increased recreation, due to water quality improvements, can also be used in conjunction with travel cost models to estimate the value of recreational water quality benefits. The travel cost model assumes that the willingness to pay for a recreational trip is at least equal to the cost incurred in traveling to the recreational site. Hedonic and travel cost models can be classified as revealed preference methods because their estimation techniques depend on observable (or revealed) behavior.

If values for water quality are not intimately related to use, for example, the desire to provide clean water environments to current and future generations, they are not necessarily revealed through either changes in market prices or other aspects of human behavior, such as travel to a recreational site or home location. Observational methods, thus, cannot reveal nonuse valuations. Researchers have circumvented this hurdle by describing hypothetical markets to respondents who are asked to indicate (or state) their willingness to pay for a nonmarket commodity (16). This relatively recent benefits estimation methodology has come to be known as the contingent valuation (CV) method, because willingness to pay values are contingent upon the particular hypothetical market described. Because the situation is hypothetical and an actual transaction or observable behavioral pattern has not been effected, value estimates based on surveys are controversial. Participants may answer strategically, for example, may indicate too high a value because they think it will improve the chance that a water quality improvement will be made, or indicate too low a value if they think that their reported willingness to pay might result in increased taxes. If participants are convinced that their responses will have no effect on actual outcomes, the incentive to reveal an inaccurate valuation is reduced (8).

The public good nature of water quality improvements also complicates benefits estimation and adds to data collection expense (10). The value of water quality improvements is typically spread across a wide population of users, and researchers must determine the relevant population for a particular improvement. Another obstacle in estimating nonmarket benefits is that most people do not have well-formed values for the vast array of environmental resources and therefore may resort to heuristics and simple protocols to construct dollar value estimates. These protocols, however, vary widely among individuals, and many irrelevant issues, such as the appropriateness of the payment vehicle or who should pay, often influence responses (17). A few experiments suggest that stated willingness to pay for existence values may be two to ten times higher than actual contributions or payments (18).

Despite the theoretical and methodological difficulties, quantifying the benefits of achieving clean water has often changed the character of debate (18). It is typical that the costs required to achieve water quality are highly concentrated but that benefits are widely distributed. Because benefits to any one user may be quite small, only bearers of costs would have a great incentive to change the outcome of a debate. Given the public good nature of many water quality improvements, however, a CV analysis may show that benefits, though widely dispersed, far outweigh costs.

In sum, certain benefits of water quality are reflected in markets, but many are not. The nonmarket benefits of water quality are important and potentially very large. Measuring these benefits, however, presents unique challenges and methodological difficulties, which have been addressed by a number of ingenious nonmarket benefits estimation techniques; CV is among the most promising. The accuracy, completeness, and potential bias of such techniques, especially CV, however, remain controversial issues.

Cost-Effectiveness Analysis

Cost-effectiveness analysis circumvents the difficulty and controversy of estimating economic benefits by focusing on the costs of achieving a quantified noneconomic objective. Once an objective has been chosen, costs of alternative strategies that reach the objective, for example, a water quality indicator, can be compared. Because the estimation of water quality indicators is generally considered more reliable than estimating the economic benefits of different levels of those indicators, cost-effectiveness analyses generally engender less controversy than cost–benefit analyses.

Cost-effectiveness analyses, in a water quality context, estimate the costs and environmental effectiveness of defined control measures or combinations of measures. Effectiveness is often measured as a reduction in the ambient load or concentration of a pollutant. Cost-effectiveness ratios can be calculated by dividing the cost of a control measure by its effectiveness. This ratio quantifies the cost of achieving a one-unit improvement in the effectiveness measure. Cost-effectiveness ratios can be easily compared across control measures to determine the most cost-effective measures—those measures that have the lowest ratios. To achieve a water quality target when any one control measure does not single-handedly achieve the target, the most cost-effective solution is typically found by applying the most cost-effective measure first, followed by the second most cost-effective measure, and so on, until a specified water quality target is achieved.

Cost-effectiveness analysis is particularly adapted to TMDL deliberations for the following reasons. First, numerical water quality targets are an essential component of TMDL plans. Second, estimating the environmental effectiveness of control measures is also central to TMDL stakeholder deliberations. Descriptions of control measures and their effectiveness are required elements of TMDL implementation plans (19). Third, the means of attaining water quality targets are not prescribed, allowing the comparison of alternative strategies. Fourth, although not a requirement, there is no prohibition on considering the costs of alternative strategies and individual control measures in determining which set of measures to implement. The EPA has characterized the TMDL process as a "cost-effective framework" for achieving water quality and has endorsed cost-effectiveness assessments (20). Fifth, from a welfare economic perspective, reducing or minimizing the cost of achieving a given water quality target increases social welfare. Sixth, compared to estimation of economic benefits, the calculation of pollution mitigation costs is relatively straightforward and less controversial. TMDL assessments, thus, are natural candidates for the application of cost-effectiveness analyses. Despite the suitability of cost-effectiveness analysis within the TMDL framework, rigorous analyses have been pursued for relatively few TMDLs, and there remains a large potential for improving the economic performance of TMDL (and other ambient-based water quality program) allocations.

Cost Minimization. Cost-effectiveness ratio analysis, as outlined above, minimizes costs within the limitations imposed by ratio analysis, for example, linearity and additivity. Under more complex (and realistic) assumptions, cost-effectiveness analysis requires more completely defined relationships among costs, water quality indicators, and control measure adoption. When costs and control measure are not additive, combinations of control measures should be assessed to capture interactive effects. Control measure costs are not necessarily linear. The same control measure at different locations or implemented at different levels can involve different unit environmental responses and entail different unit costs. More sophisticated forms of analyses are needed to analyze partial and site-specific implementation of control measures. Nonlinear and nonadditive relationships among varying levels and combinations of control measures can be addressed by employing a generalized cost minimization framework capable of solving for cost minimizing levels of control measures.

EQUITY CONSIDERATIONS

The scope of equity considerations within technology-based programs, such as NPDES, is rather limited because the procedures for prescribing water quality remedies are highly delineated within the law and regulation. Equity considerations, however, can become highly important within the ambient-based TMDL approach because particular remedies are not prescribed. In many cases, numerous entities that often belong to two or more industries are responsible for exceeding ambient standards. Loads can be reduced to TMDL targets by any combination of control responsibilities as long as the load reductions sum to the total reduction needed to achieve the TMDL target. Cost minimization can be used as the sole criterion for assigning loads, but such assignments may be rejected because they are not perceived as equitable. Hence, the flexibility of the TMDL approach creates challenges to promulgating implementation plans that are deemed equitable. Solutions deemed equitable, however, may compromise economic efficiency, that is, implementing the most cost-effective solution.

Control costs are generally concentrated in particular industrial sectors or entities. The fairness of assigning control costs is almost always at issue in water quality deliberations. Subsidizing water quality improvements by transferring mitigation costs from a small group of dischargers to broader publics is sometimes viewed as an equitable means of distributing control costs. State and federal funds are often appropriated to defray additional expenses of new control obligations. Such transfers can promote the acceptance of new controls. When such funding is not available, the legality of mandating new control obligations, especially for nonpoint sources, is often challenged.

Equity criteria can be analyzed by rigorous analytical methods, such as the attainment of a Pareto optimal. More generally, however, equity issues are dealt with subjectively; nonetheless, concern for fairness is an important driver that influences many water quality decisions. This is particularly true within the TMDL program, where attaining ambient standards often necessitates additional efforts by industry and agriculture. Most states assemble advisory boards of watershed stakeholders, typically composed of public servants, environmental interests, and industry representatives, to recommend key TMDL decisions, including load allocations. Concern for fairness is an important factor influencing load assignments. Recommended load assignments are a product of the internal dynamics of advisory board deliberations, but appeals to equity or fairness are repeatedly raised in these deliberations and undoubtedly play a large role in determining load assignments. It may be deemed equitable for each identified source, or group of sources, to reduce its water quality impact by roughly the same percentage or absolute amount. There are many potential equity criteria, however, and often no concensus on which one is best. Imposing pollution reduction obligations on small businesses and family farms, especially when economic viability is jeopardized, may be deemed unfair (as well as politically infeasible). Even if concensus if reached rearding equitable control assignments, implementation strategies deemed equitable may significantly depart from cost-effective solutions, although a correctly structured incentive-based mechanism can potentially overcome this limitation.

WATER QUALITY POLICY INSTRUMENTS

Poor water quality is symptomatic of market failure and provides a rationale for social intervention. This intervention can take the form of direct regulation, prohibition, or one of a variety of incentive-based mechanisms. Complete prohibition of discharge is sometimes resorted to for extremely toxic substances, when even a very small amount might cause harm. Regulatory approaches, often referred to as “command and control” regulation, typically prescribe control remedies that allow little flexibility in the means for achieving goals. By contrast, incentive-based mechanisms encourage behavior that reduces pollution through market forces and signals and allow at least some degree of flexibility in the means for achieving goals.

The NPDES permit system for point sources, which bases acceptable discharges on best available technology, exemplifies the regulatory approach. The regulatory approach, and NPDES in particular, have been criticized in recent years for reasons that relate mostly to questions of economic efficiency: that it is overly prescriptive and lacks flexibility, that it seeks to regulate inputs rather than outputs (water quality), that it does not adequately consider costs, and even that it has reduced real incomes (7,21). Whether or not a particular market- or incentive-based policy is, more efficient in practice, than a particular “command and control” policy is, however, an empirical question. Alternative policy instruments possess unique biophysical and economic implications, and the efficacy of any given policy depends on a number of site-specific behavioral, environmental, and economic factors that may be difficult to establish a priori. Nonetheless, incentive-based mechanisms for achieving water quality have attracted considerable academic interest because of their theoretical efficiency characteristics, and a number of incentive-based programs have been initiated.

Incentive-Based Mechanisms for Improving Water Quality

Incentive-based mechanisms are advocated primarily as a means of reducing the financial burden of control costs while maintaining or making further progress on water quality. The defining hallmark of incentive-based systems is that they are designed to be incentive compatible, that is, by pursuing self-interests, such as profitability, environmental goals are simultaneously achieved. Unlike prescriptive regulation, incentive-based mechanisms also promote the development and implementation of new pollution control technology. Though recommended by environmental economists for the past several decades, environmental organizations have more recently promoted incentive-based mechanisms (particularly effluent trading) for achieving water quality (22–24).

While generally viewed as a cost saving measure, making pollution control more affordable can also enhance its political acceptability and, hence, promote additional pollution control. Tietenburg (25), for instance, maintains that “with the inclusion of a tradable permits program for sulfur in the [acid rain bill], the compliance cost was reduced sufficiently to make passage politically possible.” Incentive-based mechanisms for improving environmental quality include emission fees or taxes, tradable permits, deposit-refund systems, subsidies for pollution control, removal of subsidies with negative environmental impacts, reductions in market barriers, and performance standards (26–28). We focus here on the two most prominent incentive-based mechanisms, pollution taxes and tradable permits.

Pollution Taxes. Economists have long recognized the allocational inefficiencies of environmental externalities. In the early part of the twentieth century, the noted economist Arthur Cecil Pigou (29) recognized that “when the divergences between private and social costs are pervasive throughout the economy, market prices cannot be used as measurements of consumer satisfaction.” Pigou proposed an externality tax on pollution to internalize the damages caused by pollution. During his time, Pigou’s pollution tax was regarded as an academic exercise and did not gain practical significance. The environmental consciousness of the 1970s, however, spurred a revival of interest in pollution taxes as a solution to environmental externalities, and several European countries and Japan adopted pollution taxes. The consensus among economists is that these fees have typically not had noticeable effects on pollution because they have not been set at levels that affect behavior (27).

Simple models of pollution abatement costs and benefits suggest that pollution taxes yield results identical to those of tradable permits within a cap if the cap or tax is appropriately set. This conclusion can be drawn from a simple graph where pollution is represented on the horizontal axis, control costs are represented on the

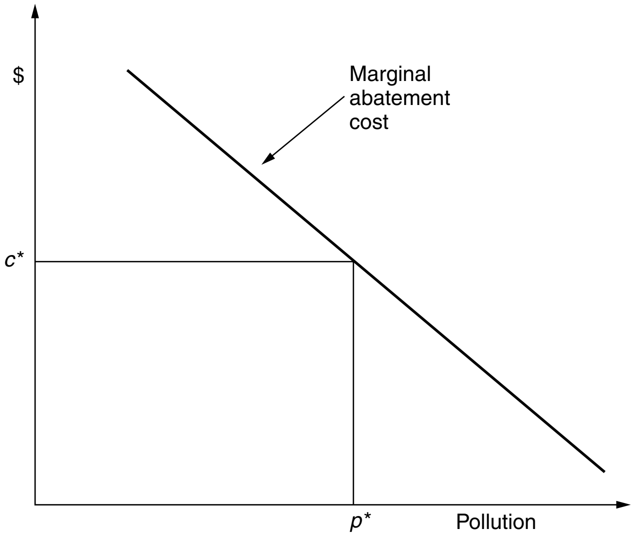

Figure 1. Theoretical equivalence of a pollution tax (c^*) and a pollution cap (p^*).

vertical axis, and declining marginal abatement costs are plotted within the graph (Fig. 1). Any point on the marginal abatement cost function can be mapped to a level of pollution (p^*) and a control cost (c^*). A pollution cap is represented as a vertical line and the equivalent pollution tax by a horizontal line, each intersecting the marginal abatement cost function at the same point. For a given pollution tax, the firm would have an incentive to reduce pollution up to the point where control costs exceed the tax, which could also be achieved by specifying a pollution cap at this level of pollution. Despite the theoretical equivalence of pollution taxes and "cap and trade" programs for the simple model portrayed in Fig. 1, a number of practical advantages are often cited in favor of tradable permits, especially when attaining ambient standards is the overriding consideration.

Tradable Permits. The theoretical underpinnings of a tradable permits system has been attributed to Coase (30), who, in his classic work, *The Problem of Social Cost*, forcefully argued the importance of clear liability rules and bargaining between parties to achieve socially desirable levels of harmful effects (31). Practical applications of Coase's basic premise were extended to air (32) and water pollution (33). Whereas Coase envisioned the bargaining of social harms between the producer and recipient of those harms (regardless of who was liable), Dales set out the parameters of the type of "cap and trade" program in widespread use today. The basic components of the program involve setting a cap on total waste, issuing emission permits, the aggregate of which equals the cap, and allowing the sale and purchase of permits among polluters (33). Marketable credits are formed when dischargers who have low mitigation costs reduce loads below permitted levels. These credits can in turn be sold to entities that have high mitigation costs so that both parties gain, thereby providing an incentive to trade.

Historical precedent for tradable permit systems for water quality is found in the allocation of natural resources, in water supply and fisheries (34). Tietenberg (25) traces the acceptance of tradable permits in the United States to the introduction of "pollution offsets" by the U.S. Congress in 1977, which opened the door to emissions trading in the air arena. In the years since, tradable emission permit systems have been successfully implemented at the federal level for the lead phaseout program, to reduce ozone-depleting chemicals under the Montreal Protocol, and for the sulfur allowance program (25). A number of programs have also been developed at the state level. Emissions trading in the air arena has generally been deemed quite successful. Burtraw and Mansur (35), for instance, estimate that trading of sulfur dioxide emissions resulted in a 13% reduction in compliance costs and produced health related benefits of nearly $570 million in 1995. The actual use of tradable permits for water quality control is of recent origin. In 1992, Willey wrote, "no water emissions markets exist yet," whereas several functioning effluent trading programs were operating by the late 1990s (26,36). Of the 37 trading and offset programs summarized by Environomics (36), 11 were considered well along in being implemented; trades were under way or completed; 5 had specific trading mechanisms approved and were very near implementation; 6 had completed the development and program approval process, but no specific trades had yet been identified; 12 were in various stages short of program approval, including study, discussion, planning, and/or development; 1 was exclusively a study; and 2 were inactive or discontinued.

The advantages of "cap and trade" programs relative to pollution taxes are, first and foremost, that a new pollution tax on industry may be politically unacceptable whereas the issuing of permits produces a potentially valuable property right. A revenue-neutral tax would reduce or eliminate this advantage, although such programs are rarely considered. Second, under conditions of uncertainty, a permit system would guarantee the attainment of the specified pollution level, whereas a large degree of guesswork would be involved in setting the appropriate tax level. Several adjustments of the tax might be required to find the level that would produce the sought for level of pollution control. Third, the amount of pollution reduction achieved by a given pollution tax is also unstable over time. Other things being equal, economic growth would foster increased emissions, as more pollution producing products and services were demanded and produced, because the pollution tax is a variable component of production costs. On the other hand, given expected advances in technology, levels of pollution might decrease if more cost-effective pollution technology were implemented over time. In summation, pollution taxes and tradable permits have relative advantages and disadvantages, but tradable permits systems are considered preferable when achieving quantitative pollution targets is the paramount consideration. Consequently, tradable effluent permits are conceptually well suited to achieving ambient water quality standards.

Because of numerous factors, TMDL and other ambient-based control allocations typically do not exhibit the greatest degree of cost-effectiveness. A tradable permit system holds the promise of reallocating any initial set of load assignments so that a least cost allocation may be attained. Whether or not a tradable permit system, in practice, is more efficient than its absence depends on the level of transactions costs relative to the gains from trade. In general, a trading system is deemed advisable only if the benefits attributable to trading exceed associated transactions costs. Transaction costs include the costs to industry of quantifying their mitigation costs, finding trading partners, initiating a trade, and bargaining, as well as the administrative costs (often borne by the public) of setting up and maintaining a market for pollution credits. If load allocations and implementation strategies are initially made in a relatively efficient manner, there is little potential for further reducing compliance costs that might be realized by a tradable permit system.

Trading has the most potential when there are large differences between marginal pollution control costs and potentially large amounts of pollution credits that can be traded. For these reasons, some researchers see the greatest potential in trades between point and nonpoint sources because the cost of nonpoint mitigation is often believed much less than that of point source mitigation (37,38). Point/nonpoint source trading, however, entails an additional set of challenges, including monitoring and measuring nonpoint pollution and uncertainty regarding the effectiveness of nonpoint controls (38,39).

Within a water quality context, effluent trading is most attractive for waterbodies where pollutants are well mixed and where there are at least several large dischargers and large differences in control costs. Spatial characteristics of water pollution can pose challenges to the development of viable markets. If pollutants are not well mixed, likely in a stream or river, the waterbody may need to be broken into several markets, or conditions to trading would need to be applied so that improvements in one area would not lead to unacceptable degradation in other areas. For this reason, it is also likely that sale of pollution credits downstream would be more likely than upstream sales for streams and rivers. Trading programs for water pollution control reflect the fact that pollution in waterbodies is generally more localized and spatially isolated than in air sheds. The potential for viable water pollution trading programs depends on various site-specific factors, many of them dictated by nature. Under the right conditions, however, a properly structured trading program can be a useful tool for achieving cost-effective water pollution control, either within or outside the TMDL process.

Having achieved some notable successes in the air arena, economists and policy makers have investigated the potential of applying the tradable permit concept to water pollution control and several programs have been developed (26,38). In 1996, the EPA issued a *Draft Framework for Watershed-Based Trading* designed to promote, encourage, and facilitate trading wherever possible, provided that equal or greater water pollution control could be attained at an equal or lower cost (40). Despite considerable effort by the EPA and the states to implement watershed trading, programs have been developed at relatively few locations, and trading within these programs has been limited, probably less than anticipated. It remains to be seen whether efforts by EPA to vigorously implement the TMDL program will greatly increase the use of watershed trading, as suggested by EPA (26), or if spatial and administrative challenges will continue to limit its adoption to a small percentage of impaired or potentially impaired waters.

SUMMARY

Like all environmental issues, the maintenance and attainment of water quality involves intrinsic economic dimensions. For a long time, however, the economic dimensions of water quality have been subordinated by prescriptive water quality policy. The emergence of an ambient-based water quality approach in the United States (in the form of the TMDL program), however, has opened the door to a more flexible approach to water quality attainment. While both benefits and costs are important elements of water quality, benefit estimation is fraught with inherent complexity and controversy. Neither technology-based nor ambient-based policy relies on cost–benefit criteria. Control costs and their estimation, however, are often important elements in TMDL (and other ambient-based program) deliberations. Hence, cost-effectiveness analyses are particularly well adapted to ambient-based water quality approaches.

Equity considerations can increase the costs of achieving water quality targets. A correctly structured incentive-based mechanism, however, can achieve both equity and economic efficiency. Pollution taxes and tradable permits are two incentive-based mechanisms that can, in principle, achieve water quality targets at lower cost. Because maintaining a quantified level of water quality (often measured by loads) is a paramount feature of ambient-based approaches, a tradable permit (cap and trade) system is better suited to TMDLs and other ambient-based programs than pollution taxes. Whether a "cap and trade" program is more efficient in practice than its absence is highly dependent on site-specific factors and the magnitude of transaction costs. The potential for viable tradable permit systems for water quality and the extent to which they may eventually be adopted remain to be seen. What is quite apparent, however, is the increasing use of economic analysis in the planning and implementation of individual water quality projects and the more prominent role that economic assessments have played in forming and evaluating water quality policy.

BIBLIOGRAPHY

1. Percival, R.V., Miller, A.S., Schroeder, C.H., and Leape, J.P. (1992). *Environmental Regulation: Law, Science, and Policy*. Little, Brown, Boston, Toronto, London.
2. Houck, O.A. (1997). TMDLs: the resurrection of water quality standards-based regulation under the clean water act. *Environ. Law Rep. Environmental Law Institute* **27**: 10329–10344.
3. Portney, P.R. (1998). Counting the cost: the growing role of economics in environmental decisionmaking. *Environment* March.

4. *Water Pollution Control Legislation—1972: Hearings on H.R. 11896, H.R. 11895 Before the Comm. On Public Works, 92d Cong. 483 (House Hearings)*. 1971. At 202–03 (statement of a panel composed of Russell E. Train, Council on Environmental Quality, and Paul V. McCracken, Council of Economic Advisors).
5. Tietenberg, T.H. (1992). *Environmental and Natural Resource Economics*, 3rd Edn. HarperCollins, New York.
6. National Research Council. (2001). *Assessing the TMDL Approach to Water Quality Management*. National Academy Press, Washington, DC.
7. Ackerman, B.A. and Stewart, R.B. (1988). Reforming environmental law: the democratic case for market incentives. *Columbia J. Environ. Law* **13**: 171–199.
8. Freeman, A.M. III (1982). *Air and Water Pollution Control: A Benefit-Cost Assessment*. John Wiley & Sons, Hoboken, NJ.
9. Ribaudo, M.O. (1989). *Water Quality Benefits from the Conservation Reserve Program*. AER-606. U.S. Dept. Agric., Econ. Res. Serv., Feb.
10. Ribaudo, M.O. and Hellerstein, D. (1992). *Estimating Water Quality Benefits: Theoretical and Methodological Issues*. TB-1808. U.S. Dept. of Agric., Econ. Res. Serv.
11. Kahn, J.R. (1998). *The Economic Approach to Environmental and Natural Resources*, 2nd Edn. Harcourt Brace College Publishers, Fort Worth.
12. Hartwick, J.M. and Olewiler, N.D. (1986). *The Economics of Natural Resource Use*. Harper & Row, New York.
13. *Inside EPA's Risk Policy Report*. (1996). *Leading Economist Warn There Are Limits to Benefit-Cost Analysis*. An Inside Washington Publication. Jan. 31.
14. Arrow, K.J., Cropper, M.L., Eads, G.C., Hahn, R.W., Lave, L.B., Noll, R.G., Portney, P.R., Russell, M., Schmalensee, R., Kerry Smith, V., and Stavins, R.N. (1996). *Benefit-Cost Analysis in Environmental, Health, and Safety Regulation: A Statement of Principles*. American Enterprise Institute, The Annapolis Center, and Resources for the Future. AEI Press.
15. Shogren, J.F. (1998). Do All the Resource Problems in the West Begin in the East? *J. Agric. Resour. Econ.* **23**(2): 309–318.
16. Mitchell, R.C. and Carson, R.T. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Resources for the Future, Washington, DC.
17. Schkade, D.A. (1995). *Issues in the Valuation of Environmental Resources: A Perspective from the Psychology of Decision Making*. In: *Review of Monetary and Nonmonetary Valuation of Environmental Investments*, T.D. Feather, C.S. Russell, K.W. Harrington, and D.T. Capan (Eds.). (Appendix B). U.S. Army Corps of Engineers. IWR Report 95-R-2.
18. Loomis, J.B. (1997). Use of non-market valuation studies in water resource management assessments. *Water Resour. Update* **109**: 5–9.
19. U.S. EPA 841-D-99-001. (1999). *Draft Guidance for Water Quality-based Decisions: The TMDL Process (Second Edition)*. Office of Water, Washington, DC.
20. Fox, C.J. (2000). *Letter to The Honorable Bud Shuster, U.S. House of Representatives*. April 5.
21. Meiners, R.E. and Yandle, B. (1993). Clean water legislation: reauthorize or repeal? In: *Taking the Environment Seriously*. R.E. Meiners and B. Yandle (Eds.). Rowman & Littlefield, Lanham, MD.
22. Oates, W.E. (1999). Forty years in an emerging field: economics and environmental policy in retrospect. *Resources* **137**: 8–11.
23. Faeth, P. (1999). *Market-Based Incentives and Water Quality. Drawn from a forthcoming report by WRI: Trading as an Option: Market Based Incentives and Water Quality by Paul Faeth. World Resources Institute*. <http://www.econet.apc.org/wri/wri/incentives/faeth.html>.
24. Environmental Defense. (2000). *From Obstacle to Opportunity: How Acid Rain Emissions Trading Is Delivering Cleaner Air*.
25. Tietenberg, T. (1999). Tradable permit approaches to pollution control: faustian bargain or paradise regained? In: *Property Rights, Economics, and the Environment*. M.D. Kaplowitz (Ed.). JAI Press, Stamford, CT.
26. Environmental Protection Agency (EPA). (2001). *The United States Experience with Economic Incentives for Protecting the Environment*. EPA-240-R-01-001. Office of the Administrator, Washington, DC.
27. Hahn, R.W. (2000). The impact of economics on environmental policy. *J. Environ. Econ. Manage.* **39**: 375–399.
28. Stavins, R.N. (1998). *Market Based Environmental Policies*. Discussion Paper 98–26. Resources for the Future, Washington, DC.
29. Pigou, A.C. (1920). *The Economics of Welfare*. Macmillan, London.
30. Tietenberg, T. *Editors Introduction to the Evolution of Emissions Trading: Theoretical Foundations and Design Considerations*. Introduction to the first volume of a two-volume collection of previously published articles. Environmental Economics Reference Series, Ashgate Press, in press.
31. Coase, R. (1960). The problem of social cost. *J. Law Econ.* October.
32. Crocker, T.D. (1966). The structuring of atmospheric pollution control systems. *The Economics of Air Pollution*. H. Wolozin (Ed.). W. W. Norton, New York, pp. 61–86.
33. Dales, J.H. (1968). Land, Water and Ownership. *Can. J. Econ.* **1**: 791–804.
34. Colby, B.G. (2000). Cap and trade challenges: a tale of three markets. *Land Econ.* **76**(4): 638–658.
35. Burtraw, D. and Mansur, E. (1999). *The Effects of Trading and Banking in the SO_2 Allowance Market*. Discussion Paper 99–25. Resources for the Future, Washington, DC.
36. Environomics. (1999). *A Summary of U.S. Effluent Trading and Offset Projects*. Prepared for: Dr. Mahesh Podar, USEPA, Office of Water. <http://www.epa.gov/owow/watershed/hotlink.htm>.
37. Letson, D., Crutchfield, S., and Malik, A. (1993). *Point-Nonpoint Source Trading for Managing Agricultural Pollutant Loadings: Prospects for Coastal Watershed*. USDA, ERS, AER-674.
38. Letson, D. (1992). Point/nonpoint source pollution reduction trading: an interpretive survey. *Nat. Resour. J.* **32**(spring): 219–232.
39. Stephenson, K. and Norris, P. (1998). Watershed-based effluent trading: the nonpoint source challenge. *Contemporary Econ. Policy* **16**(4): 412–421.
40. Environmental Protection Agency (1996). *Draft Framework for Watershed-Based Trading*. EPA 800-R-96-001. Office of Water, Washington, DC. <http://www.epa.gov/OW/watershed>

UNDERSTANDING *ESCHERICHIA COLI* O157:H7 AND THE NEED FOR RAPID DETECTION IN WATER

Zia Bukhari
Janice Weihe
American Water
Belleville, Illinois

Mark LeChevallier
American Water
Voorhees, New Jersey

INTRODUCTION

Escherichia coli is a facultative anaerobe that commonly inhabits the gastrointestinal tracts of humans within hours of birth. This gram-negative bacillus of the family Enterobactericeae is normally a harmless commensal that helps the host by suppressing pathogenic bacteria (due to competition) and by producing vitamin K and vitamin B complex (1). Individuals who suffer severe immunosuppression (both acquired and due to chemotherapy) or a breached gastrointestinal barrier can experience infection by nonpathogenic *E. coli*. More likely, however, are instances of infection by pathogenic *E. coli* that can include urinary tract infections, sepsis/meningitis, or diarrhea. Within the pathogenic *E. coli* responsible for diarrheagenic disease, differentiation of the strains has led to development of five classes (virotypes): enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), and enteroaggregative *E. coli* (EAEC) (2). Differentiation of pathogenic strains of *E. coli* is based on the organism's surface antigen profiles of O (somatic), H (flagellar), and K (capsular). This serological classification was originally proposed by Kauffman in 1944 and a modified form is still in use today. Although this serotyping of *E. coli* is a reproducible mechanism for identifying *E. coli* associated with particular clinical disease, the serotypes themselves are not responsible for virulence but are only identifiable chromosomal markers for virulence factors (2).

This article focuses on *E. coli* O157:H7, a member of the EHEC class of diarrheagenic *E. coli*, and specifically discusses transmission routes, particularly focused on the waterborne route and the methods that may be employed for detecting contamination.

HISTORY

Escherichia coli O157:H7 was first identified in 1975 by the Centers for Disease Control (CDC) but did not gain widespread attention until 1982 from the investigation of two outbreaks of gastrointestinal disease. Individuals were experiencing severe cramping, abdominal pains, watery diarrhea leading to bloody diarrhea, and low-grade fever. The illness, hemorrhagic colitis (HC), was associated with the consumption of undercooked hamburgers from a fast food restaurant, and examination of stool cultures from affected individuals revealed the presence of a rare *E. coli* serotype, O157:H7 (3). In addition, sporadic cases of hemolytic uremic syndrome (HUS) were reported by Karmali et al. (4), who found fecal cytotoxins and cytotoxin-producing *E. coli* in stool samples. Karmali et al. (5) also reported that verocytotoxin/Shiga-like toxins were the common virulence factor between HC and HUS. This led to the recognition of *E. coli* O157:H7 as an emerging enteric pathogen that was responsible for both intestinal and renal disease. To determine whether *E. coli* O157:H7 was really an emerging pathogen (rather than one that had not been detected previously), retrospective studies were conducted by the CDC and Public Health Laboratory Services (PHLS) in the United Kingdom, which indicated the presence of *E. coli* O157:H7 in only 1 of 3000 samples from 1973 to 1983 and in 1 of 15,000 samples from 1978 to 1982, respectively (2). It has been suggested that the reason for this pathogenic emergence was the acquisition of virulence factors such as the Shiga toxin gene by horizontal gene transfer and recombination (6). *Escherichia coli* O157:H7 produces Shiga-like toxins that allow categorizing it among other Shiga toxin-producing *E. coli* (STEC), alternatively called verocytotoxin-producing *E. coli* (VTEC). These terms are synonymous and are seen in the literature interchangeably. Various strains of *E. coli* O157:H7 can produce Shiga toxin 1 (stx1) and/or Shiga toxin 2 (stx2). Stx1 is virtually identical to the toxin produced by *Shigella dysenteriae* whereas stx2 is approximately 56% homologous to stx1 (6).

CLINICAL FEATURES AND TREATMENT

Infection by *E. coli* O157:H7 brings a wide range of clinical manifestations ranging from asymptomatic carriage to HC and HUS. The incubation period between exposure and disease onset ranges from 1 to 8 days. The illness typically presents with abdominal cramps and nonbloody diarrhea that progresses to bloody diarrhea after 1–2 days in 70% of cases. Blood content in stools can vary from low levels to almost entirely blood (7). Fever occurs in 30% of patients and vomiting in 30–60% of infected individuals (6). Most cases of HC resolve themselves within 1–2 weeks, and the treatment focuses on supportive care, emphasizing nutrition and hydration. It is not generally recommended to treat with antimotility agents, as this could delay O157 from being voided from the body, thus potentially increasing toxin exposure (8).

From 3% to 20% of the more severe cases of HC progress to HUS. Factors that may contribute to this progression include fever, age related susceptibility (i.e., children and the elderly are more susceptible); elevated serum leukocyte count; and treatment with antimotility agents (6). The role of antibiotic treatment in the progression to HUS remains controversial. Conflicting studies have shown both increased and decreased levels of HUS among patients who receive antibiotic therapy for HC. Death can occur in 3–5% of those infected with HUS, whereas 12–30% can experience various long-term complications, including permanent renal damage (2). There is no therapy for abatement of HUS, but dialysis and transfusions are supportive treatment, along with careful management of fluids, electrolytes, and blood pressure.

One highly touted treatment drug was Synsorb-Pk, which consists of a synthetic analog of the receptor for the Stx toxins (Gb3) linked to diatomaceous earth. This compound binds stx1 and stx2 from the intestines of affected patients, and by preventing additional toxin damage, it had the potential to limit the progression from HC to HUS. Initial clinical trials showed promise (2) but final phase III trials showed no effect of stx-binding Synsorb-Pk on the severity of HUS (9).

In some HUS patients, particularly the elderly, the disease also progresses to thrombotic thrombocytopenic purpura (TTP). TTP involves a greater incidence of fever and neurological deficits and involves organs other than the kidneys, specifically the heart and brain (7). There is also a distinct skin darkening that occurs with TTP in 90% of cases. Patients suffering from TTP typically have reduced platelet and red blood cell counts compared with HUS patients. Neurological findings in TTP range from confusion and headaches to coma and seizures. Cardiac involvement ranges from simple rhythmic disturbances to sudden cardiac arrest. Both cardiac and neurological problems are likely to be due to the occurrence of a classic TTP clinical finding: hyaline thrombi (clots composed of platelets and fibrin) that occlude the small arterioles and capillaries. Treatment of TTP consists largely of plasmapheresis (plasma volume exchange) or plasma infusions that continue until the platelet count stabilizes and hemolysis ceases. The increased use of two preventive measures has assisted in reducing mortality rates to 30%; however, some patients (11–28%) may experience a recurrent episode (10). Due to the lack of effective therapies for O157 infections, using vaccines to reduce colonization of the human intestine has been attempted by targeting proteins involved in attachment and effacement or by specifically targeting stx or the O157 lipopolysaccharide. Usually, children are most highly affected by O157 infections and implementation of vaccines in these populations can be difficult. Nonetheless, vaccination of cattle, which can be a major primary source of contamination in food and water, could reduce environmental burdens from this organism. The large numbers of cattle needing vaccination as well as the rapid turnover of these cattle may make such large-scale vaccination programs challenging (11).

PATHOGENESIS OF *E. COLI* O157:H7

Escherichia coli O157:H7 is highly acid resistant, survives the stomach environment, and passes into the gut (8). There, it adheres to the large intestine and disrupts the brush border, causing the characteristic attachment and effacing lesions of O157 infections. *Escherichia coli* O157:H7 strains can produce stx1 and/or stx2; the latter is more damaging to human renal cells (12) and leads to a greater number of HUS cases (13). The Shiga toxins are bacteriophage encoded and consist of five B subunits and one A subunit. The B subunits bind to globotriaosylceramide (Gb3), a glycolipid found in high levels in the human kidney (14), especially in infants and children (13). Upon binding to the cell membrane, the A subunit is internalized by receptor-mediated endocytosis. Within the cell, the A subunit is disrupted and produces a fragment that has RNA N-glycosidase activity, which then acts on the 28S rRNA of the 60S ribosomal subunit. This prevents the peptide chain elongation step of protein synthesis and results in cell death (8). It is believed that this process is the primary cause of renal cell damage in HUS (7).

Although the Shiga toxins produced by O157 are the most frequently discussed virulence factors, other factors can also be responsible for host pathogenesis. The presence of a 60-MDa plasmid that encodes for an enterohemolysin may cause lysis of erythrocytes as seen in the bloody diarrhea of HC. The heme and hemoglobin released from lyzed erythrocytes could be a source of iron for O157 (2,6). A second virulence factor important in O157 pathogenicity is the locus of enterocyte effacement (LEE). The LEE contains the *eaeA* and *tir* genes. The *eae* gene encodes for intimin, a bacterial cell surface protein, which along with a translocated intimin receptor, encoded by *tir*, allows O157 to adhere to intestinal mucosa. This produces the attachment and effacing lesions seen in *E. coli* O157 infections (7,15).

EPIDEMIOLOGY

The CDC estimated that the annual disease burden due to *E. coli* O157:H7 is around 20,000; up to 250 deaths occur at a cost of $250–350 million (16). However, the true incidence may be underestimated due to the failure of laboratories to isolate or detect this pathogen successfully. Some researchers estimate that the incidence is as high as 73,000 annually; 11,000 of those are from waterborne transmission (17). The disease is of greater importance in the developed countries of the Northern Hemisphere, where 80% of all STEC infections are caused by *E. coli* O157:H7 (13). Within the United States, more cases are reported in the northern states, specifically the Northeast, than in southern states. There is also an association with seasonality; most cases are reported in the summer months when peak shedding of O157 in animals has been noted (6).

Stx-producing *E. coli*, specifically O157:H7, are found in the flora of various animals, including cattle, sheep, goats, deer, pigs, cats, dogs, chickens, gulls, and even flies. Cattle are recognized as the important animal species responsible for harboring *E. coli* O157:H7; infection rates typically range from 2% to 16%, however, incidence rates up to 87% have been identified (18). Farm animals are likely reservoirs of human infection because the organism can survive for extended periods in animal feces (19), soil (20,21), and water (22,23).

In human infections, there is a wide variation in the shedding patterns of *E. coli* O157:H7; 66% of stool samples are negative for *E. coli* O157:H7 7 days after diarrhea, despite a lack of antibiotic therapy (24). A later study noted a median shedding period of 17 days (range 2–62 days) (25). Others found a median shedding period of 21 days (range 5–124 days, during which most of the patients were asymptomatic). In addition, changes in the *E. coli* profile of two patients were noticed over 2 weeks of shedding, specifically a loss in the Shiga toxin (Stx) gene, as noted by pulsed field gel electrophoresis (PFGE). This occurrence has important implications for epidemiological studies based on molecular techniques (26).

TRANSMISSION ROUTES

Enterohemorrhagic *E. coli* (EHEC) can be transmitted by any route where fecally contaminated material is ingested by a susceptible host. Epidemiological investigations during outbreaks have found that the infectious dose for EHEC is low (i.e., 10–200 organisms) (27). This low infectious dose combined with the possibility of asymptomatic carriage of *E. coli* O157:H7 allows direct person-to-person transmission. This is especially likely in institutional settings such as day-care facilities and hospitals (25). Incidences of infection from both recreational and occupational exposure to farm animals have been reported (28) as well as occupational exposure of healthcare and laboratory personnel (2,29). A documented outbreak of *E. coli* O157:H7 infection was linked to exposure from a contaminated building. Infected individuals had visited a county fair building, in which *E. coli* O157:H7 was isolated from the dust within the rafters. No other links were established, so this suggested airborne dispersion of the organism (30).

A common route of transmission is food, and a number of outbreaks have been reported from consumption of undercooked ground beef either in restaurants or within individual homes. This suggests important implications for modern food processing practices, where beef from various sources is ground together and distributed statewide or nationwide. In addition to meat, other foods were implicated, including mayonnaise, unpasteurized apple juice, fermented hard salami, lettuce, radish sprouts, and alfalfa sprouts. Indirect contamination of these foods may have occurred due to improper hygiene among food handlers or by contact between uncooked foods and raw meats contaminated with *E. coli* O157:H7. There is also the possibility of direct contamination of produce in the fields by manure from infected animals (6). Foodborne outbreaks have also occurred following consumption of yogurt and milk (31,32).

WATERBORNE TRANSMISSION

A number of *E. coli* O157:H7 outbreaks have been linked to recreational waters. Outside Portland, Oregon, 59 individuals were infected after swimming in a recreational lake in 1991 (33). Similarly, water from a children's paddling pool was responsible for transmitting infection to two other children in Scotland, UK, in 1992 (34). Epidemiological investigations did not reveal conclusive information; however, it was postulated that a child with diarrhea contaminated the pool with *E. coli* O157:H7, and as the pool water was not changed or disinfected, it served as the vehicle of infection for the other two children, who in turn infected others by person-to-person contact. A water park outbreak in Georgia in 1998 resulted in dozens of infections and the death of one child and was linked to low chlorine levels in the pool water (35).

The first municipal waterborne *E. coli* O157:H7 outbreak occurred in Missouri in 1989 (36); more than 240 people were infected and four individuals died. The outbreak, it was suspected, occurred from backflow during two water main breaks, leading to intrusion of the pathogen. In Wyoming, 157 people were infected by *E. coli* O157:H7 during the summer of 1998 as a result of a community water system supplied by a spring and two wells under the influence of surface water. The water in this system was not chlorinated, and the outbreak may have occurred from fecal contamination by wildlife near the spring (37). Another large waterborne outbreak occurred in New York in 1999 where an untreated well at a county fair was contaminated by an adjacent septic system. A total of 127 confirmed cases of *E. coli* O157:H7 infection occurred that led to 14 cases of HUS and 2 deaths (38).

The most publicized waterborne outbreak of *E. coli* O157:H7 occurred in May 2000 in Walkerton, Ontario. Approximately 1350 cases of gastroenteritis were reported in individuals exposed to Walkerton municipal water: overall, it was estimated that the number of cases associated with Walkerton exceeded 2300 individuals. Stool samples confirmed 167 cases of *E. coli* O157:H7 infection; 65 individuals were hospitalized, and 27 individuals developed hemolytic uremic syndrome. Four deaths were directly due to the *E. coli* outbreak in Walkerton, and *E. coli* O157:H7 was a contributing factor that led to three additional deaths. During the outbreak, a supply source well was contaminated with coliforms and *E. coli* O157:H7. This well was prone to surface contamination, especially following flooding conditions, which is what preceded the Walkerton incident. Furthermore, environmental testing of livestock farms, especially a farm adjacent to this well, indicated the presence of *E. coli* O157:H7 infections in the livestock. Had appropriate chlorine residuals been maintained during the flooding and well contamination, it would have prevented the widespread outbreak (39,40).

SUSCEPTIBILITY OF *E. COLI* O157:H7 TO CHLORINATION

The susceptibility of various strains of *E. coli* O157:H7 to chlorine disinfection has been examined in a number of studies, and, except for an occasional resistant strain (41), the organism can be rapidly inactivated by low levels of disinfectant. For example, in the study reported by Zhao et al. (41), free chlorine concentrations of 0.25, 0.5, 1.0, and 2.0 mg/L at 23 °C and sampling times of 0, 0.5, 1.0, and 2.0 min resulted in >7 log reduction. It was determined that a 0.25 mg/L free chlorine concentration required only 1 min to achieve these levels of inactivation, suggesting that this organism had no unusual tolerance to chlorine. In another study, Rice et al. (42) compared the chlorine susceptibility of *E. coli* O157:H7, isolated from two waterborne outbreaks, with four wild-type *E. coli* isolates and reported approximately 4 log inactivation using a chlorine residual of 1.1 mg/L and exposure times up to 120 s. The biocidal activity of chlorine is temperature dependent, and the temperatures used in the Rice et al. (42) study (5 °C) were intended to represent worst case conditions for both groundwater or winter surface water. This was probably a factor leading to lower inactivation levels in this study compared to those reported by Zhao et al. (41).

According to a Water Quality Disinfection Committee survey (1992), it was determined that United States water utilities maintain a median chlorine residual of 1.1 mg/L and a median exposure time of 45 minutes before the point of first use in the distribution system. This suggested that *E. coli* O157:H7 is unlikely to survive conventional water treatment in the United States at these levels of chlorination. In nondisinfected drinking water, *E. coli* O157:H7 has environmental survival or chlorine susceptibility similar to wild-type *E. coli* (43), which suggests that wild-type *E. coli* can be an adequate indicator organism of fecal contamination of water.

DIAGNOSTIC PROCEDURES

Biochemical

Unlike the majority of other *E. coli* isolates, most strains of *E. coli* O157:H7 do not ferment sorbitol within 24 h (44). This allows easy identification of O157 on sorbitol-containing media, such as sorbitol-MacConkey (sMAC) agar. Colonies of O157 appear colorless whereas sorbitol fermenters are pink or mauve. The use of sMAC agar is one of the primary screening methods used in the clinical laboratory to analyze patient specimens for the presence of serotype O157:H7 (45). It has been reported that sMAC agar has a detection sensitivity of 50–60% (14) and the use of sMAC agar is most successful early in the disease; the detection rate declines further (down to 33%) as the infection progresses. The use of cefixime-tellurite sorbitol-MacConkey agar (46) increases sensitivity and inhibits other nonsorbitol fermenting organisms. *Escherichia coli* O157:H7 also does not produce β-glucuronidase and therefore is negative by the methyl-umbelliferyl glucuronide (MUG) assay (47). A variety of selective media have been formulated to use these two biochemical characteristics by adding fluorogenic or chromogenic compounds (2). Although many of these selective media are useful, isolating and identifying *E. coli* O157:H7 by these methods has limitations. One of the greatest limitations involves the time-consuming procedures. Current USDA methods involve four enrichment and culturing steps followed by 10 biochemical confirmation tests that require a minimum of 3 days (17). In addition, there may be misidentification of *E. coli* O157:H7 because other intestinal bacteria (i.e., *E. hermanii* and *Hafnia* spp.) share similar phenotypes and resemble serotype O157:H7 on a sorbitol-containing medium. In addition, there are O157 that are not the H7 serotype or pathogenic and also cannot ferment sorbitol (48). There is also evidence to indicate that serotype O157:H7 in sorbitol-containing foods can mutate to become a sorbitol-fermenting phenotype (49). Isolation of sorbitol-fermenting *E. coli* O157 from HUS patients has also been reported (50) and these strains were also MUG assay positive. These findings were initially considered atypical, but an increasing prevalence of sorbitol-fermenting *E. coli* O157:H7 from HUS patients has been reported (51). There is also the issue of other flora within an environmental or clinical sample that mask low levels of O157:H7. In addition, it has been shown that *E. coli* O157:H7 can enter a viable but nonculturable state, especially when present for a long time in cold water. In such a situation, these culture techniques would not detect the organism isolated from water samples (22).

Immunologic

Currently, a number of tests are available commercially for detecting *E. coli* O157:H7; the majority of these assays are antibody based, using polyclonal, preadsorbed, affinity purified or monoclonal antibodies that target the surface-expressed antigens of O157 and H7, although some tests are specific for stx toxins. These antibody based tests have been used primarily in an enzyme-linked immunosorbent assay (ELISA); the most common assay is the "sandwich" type ELISA, where the same antibody is used both to capture and to detect the specific antigens. Such antibodies have also been effective for use in agglutination or latex agglutination tests and enable rapid screening or serological confirmation of isolates. Anti-O157 antibodies have also been coupled to paramagnetic beads (Dynal, Inc., Lake Success, NY) to isolate this pathogen selectively from the sample matrix by immunomagnetic separation (IMS) (52). IMS procedures have also been combined with ELISA, using colorimetric or fluorometric detection procedures. Simpler procedures have used lateral flow immunoassays, where the antibody–antigen complex is allowed to migrate along a membrane and then captured by a second antibody that is immobilized in a specific region of the membrane. The lateral flow format, which is similar to pregnancy test kits, produces a visual signal (with the aid of colloidal gold or dye encapsulating liposomes) when samples are positive and can be conducted without specialized skills or equipment. Obviously, using antibody based isolation and detection procedures can be prone to cross-reactivity with other organisms (i.e., *Citrobacter freundii*, *E. hermanii*, *Yersinia enterocolitica*) or other non-H7 serotypes; however, these can be reduced either by preadsorption procedures or by selecting specific target antigens (53).

Molecular

Molecular assays typically do not have the limitations of immunologic and biochemical methods. Molecular methods may include pulsed field gel electrophoresis (PFGE), ribotyping, and polymerase chain reaction (PCR). A number of PCR conditions have been developed for identifying *E. coli* O157:H7 based on virulence factors such as the stx, enterohemolysin, and intimin genes (2,14,15). Some multiplex PCRs have been developed that are more specific, namely, a PCR based assay that amplifies an allele in the *uidA* gene that is unique to serotype O157:H7, including phenotypic variants of serotype O157:H, that are sorbitol and methyl-umbelliferyl glucuronide positive (53). Coupled with primers specific for *stx* genes, this multiplex PCR assay simultaneously identifies isolates of serotype O157:H7 and the type of Shiga toxin they encode (54).

The advantages of molecular methods include specificity, sensitivity, and the ability to detect phenotypic variants of serotype O157:H7. However, the complexity and expense of these assays may limit their application in routine water microbiological analysis. Fortunately, some molecular detection procedures, such as the BAX® system by Dupont-Qualicon, do offer a user-friendly molecular assay. In this procedure, primers target sequences derived through random amplified polymorphic DNA (RAPD) analysis and all the PCR components are provided in a tablet format. Simple rehydration of the tablet with the lyzed test sample, followed by sample incubation in a thermal cycler (under preset thermocycling conditions) enables specific detection of *E. coli* O157. Although this assay provides a user-friendly format for specific detection of *E. coli* O157:H7, a drawback is that it uses proprietary equipment and software that necessitate an initial capital outlay of approximately $25,000.

CURRENT STATUS OF *E. COLI* O157:H7 MONITORING IN WATER

Currently, monitoring for microbial contamination of public water systems is regulated under a number of different United States Environmental Protection Agency (U.S. EPA) regulations such as the Safe Drinking Water Act (SDWA) of 1974, and amendments introduced in 1986 and 1996. A list published every 5 years contains contaminants that are known or anticipated in public water systems and could possibly require regulation. Microbial contaminants within this list are regulated by the Surface Water Treatment Rule and Total Coliform Rule of 1989. According to U.S. EPA regulations, a system that operates at least 60 days per year and serves 25 people or more or has 15 or more service connections is regulated as a public water system. Under the Total Coliform Rule, routine monitoring of tap water is required for all public water systems; the frequency and number of samples vary with the size of the system and the number of people that are served. The presence of total coliforms indicates that the system is contaminated or vulnerable to fecal contamination. For any size system, two consecutive total coliform positive samples from one site and one fecal coliform or *E. coli* positive sample result in an acute violation of the Maximum Contaminant Level (MCL) and lead to mandatory immediate notification of the public. Various procedures have been documented in *Standard Methods for the Examination of Water and Wastewater*, 20th ed. (55) to determine the presence of total coliforms, fecal coliform, and *E. coli*. However, *E. coli* O157:H7 has biochemical characteristics that do not allow detection by the techniques used for these broader groups.

The occurrence of waterborne or water-associated outbreaks has demonstrated the need for rapid user-friendly methods for isolating and detecting *E. coli* O157:H7 from source and finished water samples. The absence of such methods has hampered the accumulation of epidemiological information, the effectiveness of water treatment plants for removal or inactivation, and limited timely responsiveness during outbreak investigations, thus delaying employment of adequate control strategies to prevent further disease transmission. In a recent study, Bukhari et al. (56) evaluated four lateral diffusion (pregnancy type test kits) immunologic assays (Reveal®, Eclipse™, ImmunoCard STAT!, and Mionix) and one molecular assay (BAX). They determined that the Reveal test was the most promising and had a sensitivity level of 1×10^4 CFU. Using a molecular detection procedure (BAX® *E. coli* O157:H7 assay), a sensitivity level of 10 CFU was ascertained. Both the Reveal and BAX *E. coli* O157:H7 detection procedures, combined with an overall method, consisting of 200 mL water sample concentration (membrane filtration), enrichment in tryptic soy broth, and simultaneous capture by immunomagnetic separation, enabled rapid detection (less than 8 h) of 0.03–0.12 CFU/mL. The sensitivity and reproducibility of this overall procedure were determined by using spiked environmental water samples in multilaboratory round-robin trials, which indicated that the Reveal assay was more sensitive and reproducible when using spiked finished water samples; however, the BAX method had greater specificity, an important factor when evaluating raw water samples. It is anticipated that user-friendly procedures such as these will allow water utilities to employ these methods routinely to confirm the presence of *E. coli* O157:H7 rapidly in all total or fecal coliform positive samples. Alternatively, water providers could use these methods directly to gather epidemiological information, assess environmental survival, develop treatment strategies, or investigate risks during outbreak situations.

CONCLUSION

In conclusion, numerous waterborne outbreaks of human *E. coli* O157:H7 infection have occurred in the past; however, employing routine monitoring programs and coupling them with adequate disinfection would help reduce the risk of waterborne disease transmission. Historically, routine monitoring was not feasible for water utilities because the methods available were tedious and time-consuming; however, it has become viable with the methods recently described by Bukhari et al. (56), which would allow water utilities to rapidly arm themselves with the relevant information to use appropriate contingency measures to ensure continued safety from this significant human pathogen.

BIBLIOGRAPHY

1. Weir, E. (2000). *Escherichia coli* O157:H7. *CMAJ* **163**(2): 205.
2. Nataro, J.P. and Kaper, J.B. (1998). Diarrheagenic *Escherichia coli*. *Clin. Micobiol. Rev.* **11**: 142–201.
3. Riley, L.W., Remis, R.S., Helgerson, S.D., McGee, H.B., Wells, J.G., Davis, B.R., Hebert, R.J., Olcott, E.S., Johnson, L.M., Hargrett, N.T., Blake, P.A., and Cohen, M.L. (1983). Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N. Engl. J. Med.* **308**: 681–685.
4. Karmali, M.A., Petric, M., Lim, C., Fleming, P.C., and Steele, B.T. (1983). *Escherichia coli* cytotoxin, haemolytic-uraemic syndrome, and haemorrhagic colitis. *Lancet* **ii**: 1299–1300.

5. Karmali, M.A., Steele, B.T., Petric, M., and Lim, C. (1983). Sporadic cases of haemolytic uremic syndrome associated with faecal cytotoxin and cytotoxin producing *Escherichia coli* in stools. *Lancet* **i**: 619–620.
6. Mead, P. and Griffin, P. (1998). *Escherichia coli* O157:H7. *Lancet* **352**: 1207–1212.
7. Coia, J. (1998). Clinical, microbiological and epidemiological aspects of *Escherichia coli* O157 infection. *FEMS Immunol. Medical Microbiol.* **20**: 1–9.
8. Paton, J.C. and Paton, A.W. (1998). Pathogenesis and diagnosis of Shiga toxin-producing *Escherichia coli* infections. *Clin. Microbiol. Rev.* **11**: 450–479.
9. Trachtman, H., Cnaan, A., Christen, E., Gibbs, K., Zhao, S., Acheson, D.W., Weiss, R., Kaskel, F.J., Spitzer, A., and Hirshman, G.H. (2003). Effect of an oral Shiga toxin-binding agent on diarrhea-associated hemolytic uremic syndrome in children: a randomized control study. *JAMA* **290**(10): 1337–1344.
10. Rust, R.S. (2002). Thrombotic thrombocytopenic purpura and hemolytic uremic syndrome. Available: www.emedicine.com/neuro/topic499.html (accessed November 2003).
11. Horne, C., Vallance, B.A., Deng, W., and Finlay, B.B. (2002). Current progress in enteropathogenic and enterohemorrhagic *Escherichia coli* vaccines. *Expert Rev. Vaccines* **1**(2): 89–99.
12. Li, W. and Drake, M. (2001). Development of a quantitative competitive PCR assay for detection and quantification of *Escherichia coli* O157:H7 cells. *Appl. Environ. Microbiol.* **67**: 3291–3294.
13. Ochoa, T.J., and Cleary, T.G. (2003). Epidemiology and spectrum of disease of *Escherichia coli* O157. *Curr. Opin. Infect. Dis.* **16**: 259–263.
14. Kehl, S.C. (2002). Role of the laboratory in the diagnosis of enterohemorrhagic *Escherichia coli* infections. *J. Clin. Microbiol.* **40**: 2711–2715.
15. Ibekwe, M.A., Watt, P.M., Grieve, C.M., Sharma, V.K., and Lyons, S.R. (2002). Multiplex flurogenic real-time PCR for detection and quantification of *Escherichia coli* O157:H7 in dairy wastewater wetlands. *Appl. Environ. Microbiol.* **68**: 4853–4862.
16. Boyce, T.G., Swerdlow, D.L., and Griffin, P.M. (1995). Current concepts: *Escherichia coli* O157:H7 and the hemolytic-uremic syndrome. *N. Engl. J. Med.* **333**: 364–368.
17. Shelton, D.R. and Karns, J.S. (2001). Quantitative detection of *Escherichia coli* O157 in surface waters by using immunomagnetic electrochemiluminescence. *Appl. Environ. Microbiol.* **67**: 2908–2915.
18. Orr., R. 2000. The prevalence of *E. coli* O157:H7 in cattle. Available at: www.foodsafetynetwork.ca/animal/prevalence-O157-cattle.htm (accessed January 2004).
19. Wang, G., Zhao, T., and Doyle, M.P. (1996). Fate of enterohemorrhagic *Escherichia coli* O157:H7 in bovine faeces. *Appl. Environ. Microbiol.* **62**: 2657–2570.
20. Maule, A. (2000). Survival of verocytotoxigenic *Escherichia coli* O157 in soil, water and on surfaces. *Symp. Ser. Soc. Appl. Microbiol.* **29**: 71S–78S.
21. Ogden, I.D., Hepburn, N.F., MacRae, M., Strachan, N.J., Fenlon, D.R., Rusbridge, S.M., and Pennington, T.H. (2002). Long-term survival of *Escherichia coli* O157 on pasture following an outbreak associated with sheep at a scout camp. *Lett. Appl. Microbiol.* **34**: 100–104.
22. Wang, G. and Doyle, M.P. (1998). Survival of enterohemorrhagic *Escherichia coli* O157:H7 in water. *J. Food Prot.* **61**: 662–667.
23. Chalmers, R.M., Aird, H., and Bolton, F.J. (2000). Waterborne *Escherichia coli* O157. *J. Appl. Microbiol. Symp. Suppl.* **88**: 124S–132S.
24. Tarr, P.I., Neill, M.A., Clausen, C.R., Watkins, S.L., Christie, D.L., and Hickman, R.O. (1990). *Escherichia coli* O157:H7 and the hemolytic uremic syndrome: Importance of early cultures in establishing the etiology. *J. Infect. Dis.* **162**: 553–556.
25. Belongia, E.A., Osterholm, M.T., Soler, J.T., Ammend, D.A., Braun, J.E., and Macdonald, K.L. (1993). Transmission of *Escherichia coli* O157:H7 infection in Minnesota child daycare facilities. *JAMA* **269**: 883–888.
26. Karch, H., Russmann, J., Schmidt, H., Schwarzkopf, A., and Heesemann, J. (1995). Long-term shedding and clonal turnover of enterohemorrhagic *Escherichia coli* O157 in diarrheal disease. *J. Clin. Microbiol.* **33**: 1602–1605.
27. Willshaw, G.A., Thirwell, J., Jones, A.P., Parry, S., Salmon, R.L., and Hickey, M. (1994). Verocytoxin-producing *Escherichia coli* O157 in beefburgers linked to an outbreak of diarrhea, haemorrhagic colitis and haemolytic uraemic syndrome in Britain. *Lett. Appl. Microbiol.* **19**: 304–307.
28. Coia, J., Sharp, J.C.M., Campbell, D.M., Curnow, J., and Ramsay, C.N. (1998). Environmental risk factors for sporadic *Escherichia coli* O157 infection in Scotland: results of a descriptive epidemiology study. *J. Infect.* **36**(3): 317–321.
29. Griffin, P.M. (1995). *Escherichia coli* O157:H7 and other enterohemorrhagic *Escherichia coli*. In: M.J. Blaser, P.D. Smith, J.I. Ravdin, H.B. Greenberg, and R.L. Guerrent (Eds.), *Infections of the Gastrointestinal Tract*, Raven Press, New York.
30. Varma, J.K., Greene, K.D., Reller, M.E., DeLong, S.M., Trottier, J., Nowicki, S.F., DiOrio, M., Koch, E.M., Bannerman, T.L., York, S.T., Lambert-Fair, M., Wells, J.G., and Mead, P.S. (2003). An outbreak of *Escherichia coli* O157 infection following exposure to a contaminated building. *JAMA* **290**: 2709–2712.
31. Morgan, D., Newman, C.P., Hutchinson, D.N., Walker, A.M., Rowe, B., and Majid, F. (1993). Verotoxin producing *Escherichia coli* O157 infections associated with the consumption of yogurt. *Epidemiol. Infect.* **111**: 181–187.
32. Upton, P. and Coia, J.E. (1994). Outbreak of *Escherichia coli* infection associated with pasteurized milk. *Lancet* **344**(8928): 1015.
33. Keene, W.E., McAnulty, J.M., Hoesly, F.C., Williams, L.P., Hedberg, K., Oxman, G.L., Barrett, T.J., Pfaller, M.A., and Fleming, D.W. (1994). A swimming-associated outbreak of hemorrhagic colitis caused by *Escherichia coli* O157:H7 and *Shigella sonnei*. *N. Engl. J. Med.* **331**: 579–584.
34. Brewster, D.H., Browne, M.I., Robertson, D., Houghton, G.L., Bimson, J., and Sharp, J.C.M. (1994). An outbreak of *Escherichia coli* O157 associated with a children's paddling pool. *Epidemiol. Infect.* **112**: 441–447.
35. Gilbert, L. and Blake, P. (1998). Outbreak of *Escherichia coli* O157:H7 infections associated with a waterpark. *Georgia Epidemiol. Rep.* **14**: 1–2.
36. Swerdlow, D.L., Woodruff, B.A., Brady, R.C., Griffin, P.M., Tippen, S., Donnell, H.D., Geldreich, E., Payne, B.J., Meyer, A., and Wells, J.G. (1992). A waterborne outbreak in Missouri of *Escherichia coli* O157:H7 associated with bloody diarrhea and death. *Ann. Intern. Med.* **117**: 812–819.
37. Olsen, S.J., Miller, G., Breuer, T., Kennedy, M., Higgins, C., Walford, J., McKee, G., Fox, K., Bibb, W., and Mead, P. (2002). A waterborne outbreak of *Escherichia coli* O157:H7 infections and hemolytic uremic syndrome: Implications for rural water systems. *Emerg. Infect. Dis.* **8**: 370–375.

38. New York Department of Heath. (2000). Health commissioner releases *E. coli* outbreak report. Available: www.health.state.ny.us/nysdoh/commish/2000/ecoli.html (accessed November 2003).
39. Hrudey, S.E., Payment, P., Huck, P.M., Gillham, R.W., and Hrudey, E.J. (2003). A fatal waterborne disease epidemic in Walkerton, Ontario: Comparison with other waterborne outbreaks in the developed world. *Water Sci. Technol.* **47**: 7–14.
40. O'Conner, D.R. (2002). Report of Walkerton Inquiry: Part 1—The events of May 2000 and related issues. Available: http://web.utk.edu/~hydro/Geol685/Walkerton_Summary.pdf (accessed November 2003).
41. Zhao, T., Doyle, M.P., and Zhao, P. (2001). Chlorine inactivation of *Escherichia coli* O157:H7 in water. *J. Food Prot.* **64**: 1607–1609.
42. Rice, E.W., Clark, R.M., and Johnson, C.H. (1999). Chlorine inactivation of *Escherichia coli* O157:H7. *Emerg. Infect. Dis.* **5**: 461–463.
43. Rice, E.W., Johnson, C.H., Wild, D.K., and Reasoner, D.J. (1992). Survival of *Escherichia coli* O157:H7 in drinking water associated with a waterborne disease outbreak of hemorrhagic colitis. *Lett. Appl. Microbiol.* **15**: 38–40.
44. Farmer, J.J. III and Davis, B.R. (1985). H7 antiserum-sorbitol fermentation medium: A single-tube screening method for detecting *Escherichia coli* O157:H7 associated with hemorrhagic colitis. *J. Clin. Microbiol.* **22**: 620–625.
45. March, S.B., and Ratnam, S. (1986). Sorbitol-MacConkey medium for detection of *Escherichia coli* O157:H7 associated with hemorrhagic colitis. *J. Clin. Microbiol.* **23**: 869–872.
46. Zadik, P.M., Chapman, P.A., and Siddons, C.A. (1993). Use of tellurite for the selection of vero/cytotoxigenic *Escherichia coli* O157:H7. *J. Med. Microbiol.* **39**: 155–158.
47. Thompson, A., Hodge, D., and Borczyk, A. (1990). Rapid biochemical test to identify verocytotoxin-positive strains of *Escherichia coli* serotype O157. *J. Clin. Microbiol.* **28**: 2165–2168.
48. Willshaw, G., Smith, H., Roberts, D., Thirwell, J., Cheasty, T., and Rowe, B. (1993). Examination of raw beef products for the presence of verocytotoxin producing *Escherichia coli,* particularly those of serogroup O157. *J. Appl. Bacteriol.* **75**: 420–426.
49. Fratamico, P.M., Buchanan, R.L., and Cooke, P.H. (1993). Virulence of an *Escherichia coli* O157:H7 sorbitol-positive mutant. *Appl. Environ. Microbiol.* **59**: 4245–4252.
50. Gunzer, F., Bohm, H., Russman, H., Bitzan, M., Aleksic, S., and Karch, H. (1992). Molecular detection of sorbitol-fermenting *Escherichia coli* O157 in patients with hemolytic-uremic syndrome. *J. Clin. Microbiol.* **30**: 1807–1810.
51. Bitzen, M., Ludwig, K., Klemt, M., Konig, H., Buren, J., and Muller-Wiefel, D.E. (1993). The role of *Escherichia coli* O157 infections in the classical (enteropathic) haemolytic urameic syndrome: Results of a Central European, multicenter study. *Epidemiol. Infect.* **110**: 183–196.
52. Wright, D.J., Chapman, P.A., and Siddons, C.A. (1994). Immunomagnetic separation as a sensitive method for isolating *Escherichia coli* O157 from food samples. *Epidemiol. Infect.* **113**: 31–39.
53. Feng, P. (1993). Identification of *Escherichia coli* O157:H7 by DNA probe specific for an allele of uidA gene. *Mol. Cell Probes* **7**: 151–154.
54. Cebula, T.A., Payne, W.L., and Feng, P. (1995). Simultaneous identification of strains of *Escherichia coli* serotype O157:H7 and their Shiga-like toxin type by mismatch amplification mutation assay-multiplex PCR. *J. Clin. Microbiol.* **33**: 248–250.
55. *Standard Methods for the Examination of Water and Wastewater*. 20th ed. (1998). American Public Health Association, American Water Works Association, Water Environment Federation, joint publishers.
56. Bukhari, Z., Weihe, J., and LeChevallier, M. (2005). Detection methods for *Escherichia coli* O157:H7 American Water Works Association Research Foundation Report.

EUTROPHICATION AND ORGANIC LOADING

Scott A. Lowe
Manhattan College
Riverdale, New York

The connection between eutrophication and organic loading was first realized in the 1870s by engineers in London. They concluded that when sewage was discharged into local rivers, the level of dissolved oxygen (DO) dropped, and the amount of algae increased. In other words, the rivers were eutrophied.

Raw sewage contains organic matter (including pathogens that make it so harmful when mixed in drinking water). The typical composition by weight of raw sewage is about 42% carbon (C), 7% nitrogen (N), and 1% phosphorus (P). The remaining 50% is water. This is also our human stoichiometry.

When discharged into a receiving water such as a lake, river, or estuary, the sewage begins to break down into its components. The sewage carbon, for example, oxidizes to CO_2:

$$C \rightarrow CO_2 \tag{1}$$

Stoichiometrically, it can be seen that $2 \times 16\,\mathrm{g\,O}$ is required to oxidize 12 g of carbon, or expressed as a ratio, 2.67 g O/g C. Similarly, for sewage nitrogen, which is mainly bound up as ammonia, NH_3:

$$NH_3 \rightarrow NO_3{}^- + H_2O + H^+ \tag{2}$$

The stoichiometry is 4×16 g O for every 14 g N; as a ratio, 4.57 g O/g N.

Phosphorus in raw sewage is bound up mainly as orthophosphate, PO_4, which is already in an oxidized state and so remains unchanged.

The oxygen that is consumed in this degradation process is supplied by the receiving water DO—hence the drop in DO that was observed in those London rivers. The other observed consequence of high levels of organic loading is excessive algal activity. To understand this phenomenon, one must consider what algae, typically phytoplankton, require to grow.

Phytoplankton have an internal stoichiometry that is similar in proportion to that described above for sewage, that is, 42% C, 7% N, and 1% P. The rest is water. To grow, they need to consume C, N, and P in this proportion. Being plants, they obtain their carbon from dissolved CO_2 in water, in the well-known process of photosynthesis. This also means that phytoplankton require sufficient light to grow.

The N and P that phytoplankton require come from dissolved N and P in the water. This N and P results from the breakdown of organic matter, as described before. If all four items (light, C, N, and P) are supplied in excess, then extreme algal growth occurs. This creates obvious visual effects such as "pea soup" in a pond or lake or "red tides" in the ocean.

To control this phenomenon, C is not important because there is always excess dissolved CO_2 in the water compared to the other nutrients, N and P. Light is also considered uncontrollable in natural systems. Therefore, control efforts center around N and P. The ratio at which algae need N and P is called the nitrogen to phosphorus ratio, simply N/P. It is approximately 7; algae need approximately seven times as much N as P, per their internal composition (7%N and 1%P). Therefore, if N and P in the water are measured (N as dissolved NH_3 and NO_3 and P as dissolved PO_4), then the ratio of actual N/P can be compared to the algal N/P ratio of 7.

The implication of this comparison is that if the measured N/P ratio is greater than, say, 10, then there is excess nitrogen in the water and the lack of phosphorus is limiting algal growth. Conversely, a low N/P ratio would imply that a lack of nitrogen is limiting algal growth. This is known as the limiting nutrient concept. Once the limiting nutrient (either N or P) for a system is identified, then efforts at eutrophication control will focus on further restricting inputs of this nutrient to the system. If the system is relatively balanced in both N and P, then whichever of the two is easier and more cost-effective to control is focused on.

In general, it has been found that freshwater systems are P limited, whereas marine systems, including estuaries, are N limited. Some real world examples of attempts to control eutrophication include:

- the ban on phosphorus in detergents and fertilizers in the watersheds that drain to the Great Lakes, which are P-limited;
- the upgrading of wastewater treatment plants to remove N that discharge to Long Island Sound, which is N-limited;
- the installation of tertiary treatment to remove P in wastewater treatment plants on watersheds that drain into New York City's drinking water reservoirs, which are P-limited; and
- the upgrading of wastewater treatment plants to remove N that discharge to Chesapeake Bay, which is N-limited.

Two final points—first, organic loads can come from raw sewage, and also from a myriad of other sources such as

- agricultural runoff
- industrial discharges
- forests
- urban runoff
- stormwater
- wastewater treatment plants

Second, sediment often contains large amounts of organic matter as a result of organic loading. One consequence of this is that even when the external organic load is reduced, for example, by treating raw sewage in a wastewater treatment plant, there still remains an internal source of nutrients for algal growth. Additionally, sediments with high organic content also exert an oxygen demand on the water, known as SOD—sediment oxygen demand.

TRACE ELEMENT CONTAMINATION IN GROUNDWATER OF DISTRICT HARDWAR, UTTARANCHAL, INDIA

Chakresh K. Jain
National Institute of Hydrology
Roorkee, India

The trace element contamination in the groundwater of District Hardwar in the state of Uttaranchal (India) has been assessed to find the suitability of groundwater for drinking purposes. Forty-eight groundwater samples from shallow and deep aquifers were collected during both pre- and postmonsoon seasons in 2002 and analyzed for Fe, Mn, Cu, Ni, Cr, Pb, Cd, and Zn. The presence of trace elements in groundwater was recorded in all samples, but these were not significantly higher. Water quality standards were violated for iron and manganese at certain locations. The concentration of iron varies from 210 to 7004 μg/L during the premonsoon season and from 163 to 6555 μg/L during the postmonsoon season against the maximum permissible limit of 1000 μg/L; that of manganese varies from 7.5 to 839 μg/L during the premonsoon and from 9 to 859 μg/L during the postmonsoon season against the limit of 300 μg/L. The concentration of copper, chromium, lead, cadmium, and zinc were well within the permissible limits in all samples from District Hardwar.

INTRODUCTION

The contamination of groundwater by trace elements has received great attention in recent years due to their toxicity and accumulative behavior. These elements, in contrast to most pollutants, are not biodegradable and undergo a global ecobiological cycle in which natural waters are the main pathways. The major sources of trace elements in groundwater include weathering of rock minerals, discharge of sewage and other waste effluents on land, and runoff water. The water used for drinking should be free from any toxic elements, living and nonliving organisms, and excessive amounts of minerals that may be hazardous to health. Some trace elements are essential for humans, for example, cobalt and copper, but large quantities of them may cause physiological disorders. Cadmium, chromium, and lead are highly toxic to humans even in low concentrations.

The creation of the new state of Uttaranchal has posed many challenges for planners and policy makers. Problems such as drinking water, transportation, power sector, housing and construction, and safety from natural hazards are very serious and require immediate attention. For sustainable development of a society, it is essential that the natural resources are used judiciously for the benefit of the existing population and also to meet the needs and aspirations of future generations. Drinking water is one such precious commodity for which a planned strategy is needed for immediate demands and for sustainability for future needs also. A large part of the state of Uttaranchal lies in the hills, where distribution of the drinking water supply and its quality are major problems needing immediate attention. The physicochemical and bacteriological characteristics of the groundwater of District Hardwar in the state of Uttaranchal have been described in an earlier report (1). In this article, the groundwater quality is examined with reference to trace elements.

STUDY AREA

District Hardwar is part of the Indo-Gangetic plains that lie between latitude 29°30′ and 30°20′ N and between longitude 77°40′ and 78°25′ E in the state of Uttaranchal (Fig. 1). It is the largest district (populationwise) of Uttaranchal State and occupies an area of about 2360 km^2. Per the 2001 census, the population of the District Hardwar is 14,44,213 and the population density is 612 per km^2.

Physiographically, the area is generally flat except for the Siwalik Hills in the north and northeast. The area is devoid of relief features of any prominence except for deep gorges cut by ravines and rivers flowing through the area. The area is bounded by River Yamuna in the west and River Ganga in the east. The climate of the area is characterized by a moderate type of subtropical monsoon. The average annual rainfall in the region is about 1000 mm; the major part is received during the monsoon period. The major land use is agriculture, and there is no effective forest cover. The soils of the area are loam to silty loam and are free from carbonates. The most common groundwater use is by hand pumps and tube wells. Based on the lithologic logs and water table fluctuation data, two types of aquifers have been delineated in the area. The upper is a shallow unconfined aquifer that generally extends to depths around 25 m. The deeper is confined to semiconfined and located about 25 to 150 m below ground level separated by three to four aquifers at average depths of 25 to 55, 65 to 90, and 120 to 150 m. Water table contours in the area indicate the southward trend of groundwater flow in both unconfined and confined aquifers.

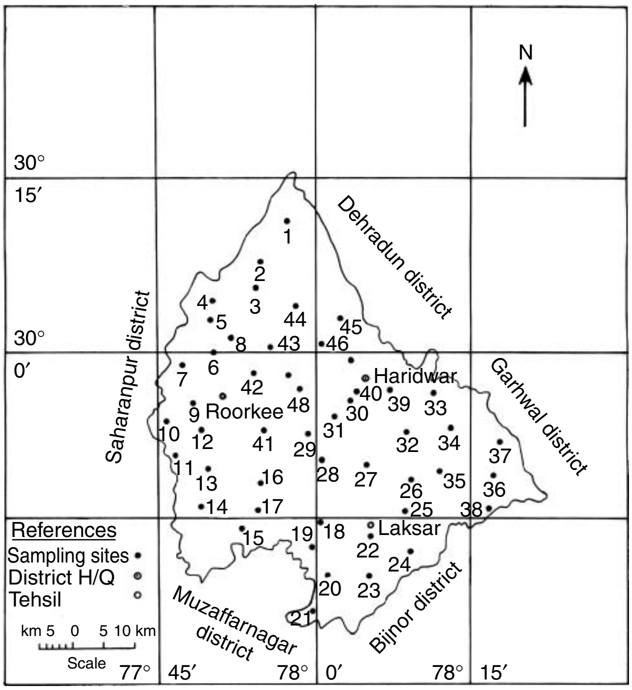

Figure 1. Study area showing location of sampling sites (District Hardwar).

EXPERIMENTAL METHODOLOGY

Forty-eight groundwater samples from District Hardwar were collected during both pre- (June 2002) and postmonsoon (October 2002) seasons from various abstraction sources at various depths covering extensively populated area, commercial, industrial, agricultural, and residential colonies, so as to obtain a good areal and vertical representation; they were preserved by adding ultrapure nitric acid to a pH of 2.0 (2,3). The hand pumps and tube wells were continuously pumped prior to sampling to ensure that groundwater sampled was representative of the groundwater aquifer. The details of sampling locations and source and depthwise distribution are given in Tables 1 and 2, respectively.

Metal ion concentrations were determined by atomic absorption spectrometry using a Perkin-Elmer Atomic Absorption Spectrometer (Model 3110) using an air–acetylene flame. Operational conditions were adjusted in accordance with the manufacturer's guidelines to yield optimal determination. Quantification of metals was based upon calibration curves of standard solutions of respective metals. These calibration curves were determined several times during the period of analysis. The detection limits for iron, manganese, copper, nickel, chromium, lead, cadmium and zinc are 0.003, 0.001, 0.001, 0.004, 0.002, 0.01, 0.0005, and 0.0008 mg/L, respectively.

RESULTS AND DISCUSSION

The Bureau of Indian Standards (BIS), earlier known as the Indian Standards Institution (ISI), laid down standard specifications for drinking water during 1983; they have been revised and updated from time to time. To enable the users to exercise their discretion toward water quality criteria, the maximum permissible limit has been prescribed especially where no alternate source is available. The national water quality standards describe essential and desirable characteristics that must be evaluated to assess the suitability of water for drinking (4). The trace element data for the two sets of samples collected during pre- and postmonsoon seasons are presented in Figs. 2–9. The toxic effects of these elements and the

Table 1. Description of Groundwater Sampling Locations in District Hardwar

Sample No.	Location	Source	Depth, m
1	Mohand	OW	10
2	Banjarewala	OW	10
3	Buggawala	HP	30
4	Kheri	HP	38
5	Dadapatti	HP	38
6	Bahbalpur	HP	15
7	Sikandarpur	HP	25
8	Bhagwanpur	HP	30
9	Chudiala	HP	40
10	Balswa Ganj	HP	25
11	Manakpur	HP	40
12	Iqbalpur	HP	30
13	Jharera	HP	30
14	Sherpur	HP	30
15	Narsen	HP	30
16	Manglour	HP	35
17	Libarheri	HP	35
18	Mahesari	HP	20
19	Sahipur	HP	35
20	Khanpur	HP	15
21	Chandpuri Kalan	HP	15
22	Laksar	HP	15
23	Kalsiya	HP	10
24	Niranjanpur	HP	15
25	Sultanpur	HP	15
26	Shahpur	HP	25
27	Pathri	HP	15
28	Subashgarh	HP	20
29	Marghubpur	HP	30
30	Bahadarabad	HP	20
31	Alipur	HP	20
32	Katarpur	HP	10
33	Kankhal	HP	25
34	Shyampur	HP	30
35	Rasiya Garh	HP	40
36	Gandikhatta	HP	10
37	Laldhang	HP	90
38	Kottawali	HP	40
39	Hardwar	HP	30
40	Jwalapur	HP	30
41	Roorkee	HP	10
42	Gumanwala	HP	30
43	Manubas	HP	10
44	Bandarjud	HP	15
45	Beriwala	HP	5
46	Hazara	HP	10
47	Aurangbad	HP	10
48	Daulatpur	HP	35

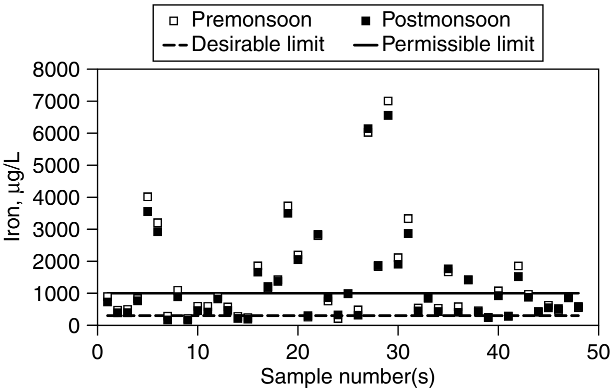

Figure 2. Distribution of iron at different sampling locations.

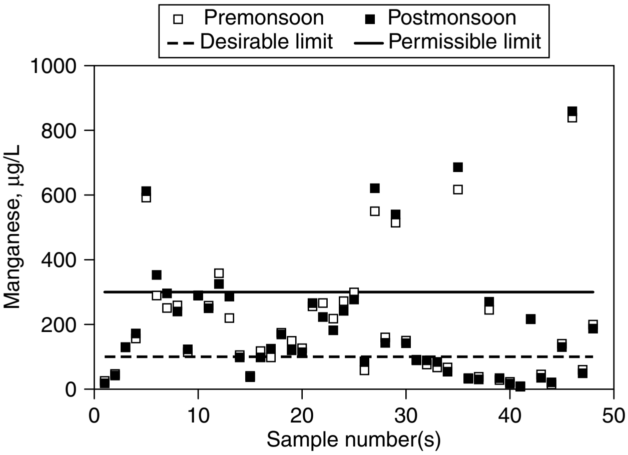

Figure 3. Distribution of manganese at different sampling locations.

extent of their contamination in the groundwater are discussed in the following sections.

Iron (Fe)

The concentration of iron in the groundwater of District Hardwar ranges from 210 to 7004 µg/L during the premonsoon season and from 163 to 6555 µg/L during the postmonsoon season; its distribution at different sites is shown in Fig. 2. The Bureau of Indian Standards has recommended 300 µg/L as the desirable limit and 1000 µg/L as the maximum permissible limit for drinking water (4). It is

Table 2. Source and Depthwise Distribution of Sampling Sites in District Hardwar

	Depth Range			
Source Structure	<0–20 m	20–40 m	>40 m	Total number
Hand pumps	6, 18, 20, 21, 22, 23, 24, 25, 27, 28, 30, 31, 32, 36, 41, 43, 44, 45, 46, 47	3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 26, 29, 33, 34, 35, 38, 39, 40, 42, 48	37	46
Tube wells	—	—	—	—
Open wells	1, 2	—	—	2
Total	22	25	1	48

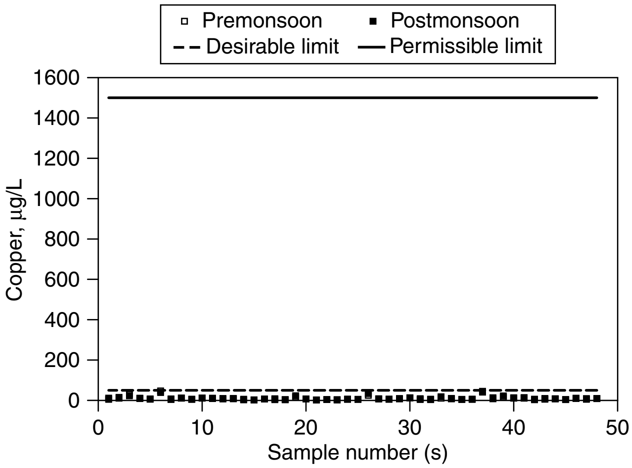

Figure 4. Distribution of copper at different sampling locations.

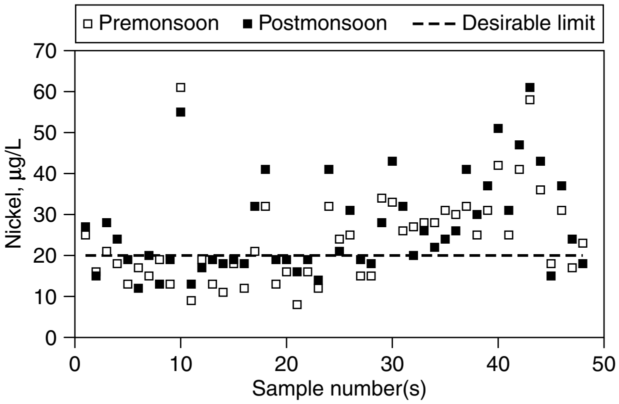

Figure 5. Distribution of nickel at different sampling locations.

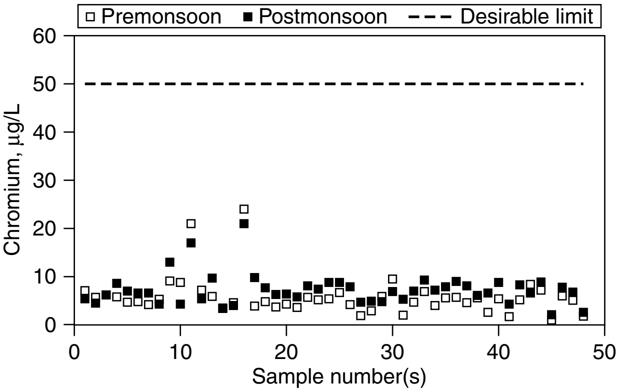

Figure 6. Distribution of chromium at different sampling locations.

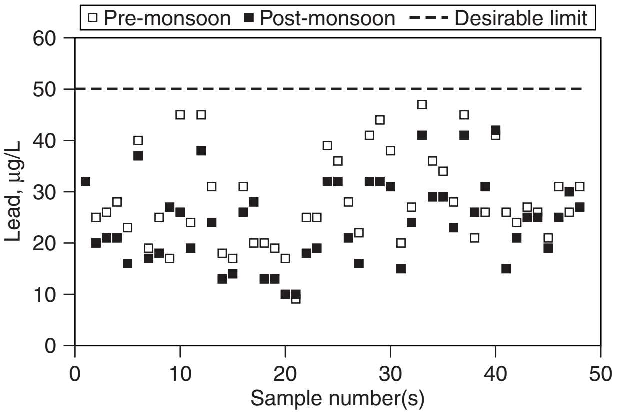

Figure 7. Distribution of lead at different sampling locations.

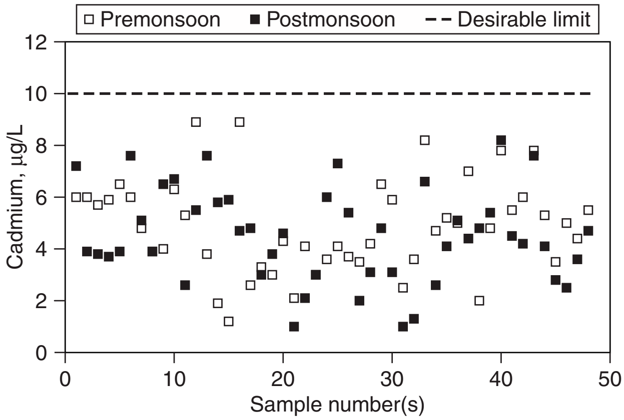

Figure 8. Distribution of cadmium at different sampling locations.

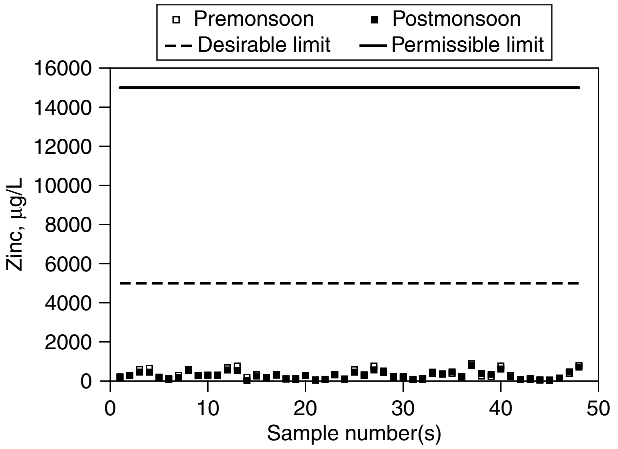

Figure 9. Distribution of zinc at different sampling locations.

evident from the results that only about 15% of the samples fall within the desirable limit of 300 µg/L and about 35% of the samples even cross the maximum permissible limit of 1000 µg/L. High concentrations of iron generally cause an inky flavor and a bitter, astringent taste. It can also discolor clothes and plumbing fixtures and cause scaling, which encrusts pipes. Excessive concentration may promote bacterial activities in pipe and service mains, causing objectionable odors and red-rot disease in water. Well water containing soluble iron remains clear while pumped out, but exposure to air causes precipitation of iron due to oxidation and a resulting rusty color and turbidity.

The increase in iron concentrations during the postmonsoon season may be due to mixing phenomena of recharge water that has more oxygen to react with iron ore, which may be available in clay lenses in the aquifer. The limits of iron in water supplies for potable use have not been laid down from health considerations but are due to the fact that iron in water supplies may cause discoloration of clothes, plumbing fixtures, and porcelain ware. The "red-rot" disease of water, caused by bacterial precipitation of hydrated oxides of ferric iron that give an unaesthetic appearance to water, clog pipes, pit pipes, and cause foul smells, is due to the presence of relatively high iron in water. The concentration of iron in natural water is controlled by both physicochemical and microbiological factors. The weathering of rock and discharge of waste effluents on land are generally considered the main sources of iron in groundwater. Iron migrates adsorbed to suspended matter, as insoluble hydrated iron compounds complexed to inorganic and organic ligands, and also as hydrated ions. Dissolved carbon dioxide, pH, and E_h of water affect the nature of aqueous iron species in the water. In groundwater, iron generally occurs in two oxidation states, ferrous (Fe^{2+}) and ferric (Fe^{3+}) forms.

Manganese (Mn)

The concentration of manganese was recorded at a maximum level of 839 μg/L during the premonsoon season and 859 μg/L during the postmonsoon season; its distribution at different sites is shown in Fig. 3. Manganese is an essential element, which does not occur naturally as a metal but is found in various salts and minerals frequently in association with iron compounds. In general, the concentration of manganese in groundwater is low due to geochemical control. A concentration of 100 μg/L has been recommended as a desirable limit and 300 μg/L as the permissible limit for drinking water (4). The WHO has prescribed 500 μg/L as a provisional guideline for drinking water (5). The presence of manganese above the permissible limit in drinking water often imparts an alien taste to water. It also has adverse effects on domestic uses and water supply structures.

It is evident from the results that about 35% of the samples from District Hardwar fall within the desirable limit of 100 μg/L, 50% of the samples cross the desirable limit but are within the permissible limits, and about 10–15% of the samples exceed the permissible limit of 300 μg/L. The high concentration of manganese at few locations may be attributed to the reducing conditions of the water and dissolution of manganese bearing minerals from the soil strata. Manganese may gain entry into the body by inhalation, consumption of food, and through drinking water.

Copper (Cu)

The concentration of copper was recorded at a maximum level of 47 μg/L during the premonsoon season and 41 μg/L during the postmonsoon season. The distribution of copper at different sites during the pre- and postmonsoon seasons is shown in Fig. 4. The Bureau of Indian Standards has recommended 50 μg/L as the desirable limit and 1500 μg/L as the permissible limit in the absence of alternate sources (4). Beyond 50 μg/L, the water has an astringent taste and causes discoloration and corrosion of pipes, fittings, and utensils. The World Health Organization has recommended 2000 μg/L as the provisional guideline for drinking water (5).

All samples from District Hardwar fall below the desirable limit of 50 μg/L. As such, the groundwater of District Hardwar can be safely used as a source of drinking water supplies. In general, the principal sources of copper in water supplies are corrosion of brass and copper pipe and addition of copper salts during water treatment for algae control. The toxicity of copper to aquatic life depends on the alkalinity of the water. At lower alkalinity, copper is generally more toxic to aquatic life. Copper, if present in excess amount in public water supplies, enhances corrosion of aluminum and zinc utensils and fittings. High intake of copper may result in damage to the liver. The industrial sources of copper that enhance the concentration in groundwater include industrial effluents from electroplating units, textiles, paints, and pesticides.

Nickel (Ni)

The concentration of nickel in the study area was recorded at a maximum level of 61 μg/L during the pre- and postmonsoon seasons; its distribution at different sites is shown in Fig. 5. The World Health Organization has recommended 20 μg/L as the guideline for drinking water (5).

More than 50% of the samples from District Hardwar exceed the WHO limit of 20 μg/L during pre- as well as postmonsoon seasons. The violation of the BIS limit could not be ascertained as no permissible limit of nickel has been prescribed in BIS drinking water specifications. Nickel at trace levels is essential to human nutrition, and no systemic poisoning from nickel is known in this range. The level of nickel usually found in food and water is not considered a serious health hazard. Some of the important nickel minerals include garnierite, nickeliferous limonite, and pentiandite. Certain nickel compounds have carcinogenic effects on animals; however, soluble compounds are not currently regarded as human or animal carcinogens.

Chromium (Cr)

The concentration of chromium in the study area was recorded at a maximum level of 24 μg/L during the premonsoon season and 21 μg/L during the postmonsoon season; in this range, it is not harmful to humans. The distribution of chromium at different sites during pre- and postmonsoon seasons is shown in Fig. 6. A concentration of 50 μg/L has been recommended as a desirable limit for drinking water (4). The WHO has also prescribed 50 μg/L as the guideline for drinking water (5). All samples from the study area fall well within the desirable limit for drinking water.

The two important oxidation states of chromium in natural waters are +3 and +6. In well-oxygenated waters, Cr(+6) is the thermodynamically stable species. However, Cr(+3) is kinetically stable and could persist bound to naturally occurring solids. Interconversions of Cr(+3) and Cr(+6) occur in conditions similar to those in natural waters. Municipal wastewater releases a considerable amount of chromium into the environment. Chromium is not acutely toxic to humans due to the high stability of natural chromium complexes in abiotic matrices. In addition, the hard acid nature of chromium imparts a strong affinity for oxygen donors rather than sulfur donors present in biomolecules. However, Cr(+6) is more toxic than Cr(+3) because of its high rate of adsorption through intestinal tracts. In the natural environment, Cr(+6) is likely to be reduced to Cr(+3), thereby reducing the toxic impact of chromium discharges.

Lead (Pb)

The concentration of lead in the study area was recorded at a maximum level of 47 μg/L during the premonsoon season and at 42 μg/L during the postmonsoon season. The distribution of lead at different sites during the pre- and postmonsoon seasons is shown in Fig. 7. The Bureau of Indian Standards has prescribed 50 μg/L lead as the desirable limit for drinking water (4). Beyond this limit, water becomes toxic. WHO has also prescribed the same guideline for drinking water (5).

All samples from District Hardwar fall within the permissible limit for drinking water during the pre- as well as the postmonsoon seasons, and therefore, the groundwater in the study area can be safely used as a source of drinking water supplies. The major source of lead contamination is the combustion of fossil fuel. Lead is removed from the atmosphere by rain, falls back on the earth's surface, and seeps into the ground. Lead passes from the soil to water and to plants and finally into the food chain. In drinking water, it occurs primarily due to corrosion of lead pipes and solder, especially in areas of soft water. Dissolution of lead requires extended contact time, so lead is most likely to be present in tap water after being in the service connection piping and plumbing overnight.

Cadmium (Cd)

Cadmium is a nonessential, nonbeneficial element known to have a high toxic potential. The cadmium content in the study area varies from 1.2 to 9.0 μg/L during the premonsoon season and from 1.0 to 8.2 μg/L during the postmonsoon season. The distribution of cadmium at different sites during the pre- and postmonsoon seasons is shown in Fig. 8. The Bureau of Indian Standards has prescribed 10 μg/L cadmium as the desirable limit for drinking water (4). Beyond this limit, water becomes toxic. WHO has prescribed 3 μg/L cadmium as the guideline for drinking water (5).

All samples from District Hardwar were within the desirable limit of 10 μg/L, as prescribed by BIS. It is obvious, therefore, that the groundwater of District Hardwar does not present any cadmium hazards to humans. The levels of cadmium in public water supplies are normally very low; generally, only small amounts exist in raw water, and many conventional water treatment processes remove much of the cadmium. Drinking water that contains more than 10 μg/L of cadmium can cause bronchitis, emphysema, anemia, and renal stone formation in animals. Cadmium can also enter the environment from a variety of industrial sources, including mining and smelting, electroplating, and pigment and plasticizer production. Drinking water is generally contaminated by galvanized iron pipe and plated plumbing fittings of the water distribution system. The U.S. EPA has classified cadmium as a probable human carcinogen based on positive carcinogenicity testing. However, no such health hazard is expected in the groundwater of District Hardwar.

Zinc

The concentration of zinc in the study area ranges from 44 to 878 μg/L during the premonsoon season and from 15 to 787 μg/L during the postmonsoon season. The distribution of zinc at different sites during the pre- and postmonsoon seasons is shown in Fig. 9. The Bureau of Indian Standards has prescribed 5000 μg/L zinc as the desirable limit and 15,000 μg/L as the permissible limit for drinking water (4). The WHO has prescribed 3000 μg/L as the guideline for drinking water (5). In the study area, all samples analyzed were within the desirable limits prescribed by the BIS (4) and the WHO (5).

CONCLUSIONS AND RECOMMENDATIONS

The groundwater quality in District Hardwar varies from place to place and with the depth of the water table. The water drawn for domestic applications should be tested and analyzed to ensure the suitability of groundwater for human consumption. The heavy metals except iron and manganese, which are present in appreciable concentration in groundwater, have been below the prescribed maximum permissible limits. The concentration of iron varies from 210 to 7004 μg/L during the premonsoon season and from 163 to 6555 μg/L during the postmonsoon season compared to the maximum permissible limit of 1000 μg/L; that of manganese varies from 7.5 to 839 μg/L during premonsoon and from 9 to 859 μg/L during the postmonsoon season compared to the 300 μg/L limit. The concentration of copper, chromium, lead, cadmium, and zinc were well within the permissible limits in all samples from the study area.

BIBLIOGRAPHY

1. Jain, C.K. (2003). *Ground water quality in District Hardwar, Uttaranchal*. Technical Report, National Institute of Hydrology, Roorkee, India.
2. Jain, C.K. and Bhatia, K.K.S. (1988). *Physico-Chemical Analysis of Water and Wastewater, User's Manual, UM-26*. National Institute of Hydrology, Roorkee, India.
3. APHA. (1992). *Standard Methods for the Examination of Water and Waste Waters*, 18th Edn. American Public Health Association, Washington, DC.
4. BIS. (1991). *Specifications for Drinking Water, IS:10500:1991*. Bureau of Indian Standards, New Delhi.
5. WHO. (1996). *Guidelines for Drinking Water, Recommendations*, Vol. 2. World Health Organization, Geneva.

IRON BACTERIA

Baikun Li
Pennsylvania State University
Harrisburg, Pennsylvania

GENERAL INTRODUCTION

Iron is one of the most abundant elements found in nature, accounting for at least 5% of Earth's crust. On the surface of Earth, iron exists naturally in two oxidation states, ferrous (Fe^{2+}) and ferric (Fe^{3+}). Dissolved or ferrous iron is usually found only in groundwater supplies that have not been exposed to oxygen. Water containing ferrous iron is clear and colorless. At contact with the air, the dissolved ferrous iron reacts with oxygen and forms insoluble ferric oxide, also known as iron oxide or "red rust." Water containing ferric iron has undesirable odors and tastes. Iron staining has caused problems in well systems and water distribution pipelines. With the presence of iron bacteria, iron may cause the corrosion of steel piping components.

Iron bacteria are found in all parts of the world, with the exception of the southern polar region. The "true" iron bacteria are microorganisms in which the oxidation of iron is an important source for their metabolic energy. This group is most often associated with filamentous and stalked forms that are encrusted with iron *(Thiobacillus*, *Leptothrix*, *Clonothrix*, *Gallionella*, *Sphaerotilus)*. During the oxidation, soluble ferrous iron (Fe^{2+}) is transferred to insoluble ferric iron (Fe^{3+}), which then precipitates out of solution. The overall reaction of spontaneous ferrous iron oxidation is

$$Fe^{2+} + 0.25\ O_2 + 2.5\ H_2O \rightarrow Fe(OH)_3 + 2H^+$$

The energy released from this reaction promotes the growth of thread-like slimes, which, with the ferric iron, form a voluminous mass that has caused clogging in the water supply system.

Thiobacullus ferrooxidans (Fig. 1) and *Leptothrix ferrooxidans* (Fig. 2) are the best-known iron-oxidizing bacteria. Both of them can grow autotrophically with ferrous iron (Fe^{2+}) as the electron donor. They are very common in acid-polluted environments such as coal-mining dumps. Also, acidophilic iron bacteria exist, such as *Ferrobacillus ferrooxidans* and *Ferroplasma acidophilum* (archaeon), which can oxidize iron in low pH environments (even $pH = 0$) at high temperature (50 °C), and are most commonly found associated with acid mine waste. The acidophilic bacteria are not commonly found in drinking water supplies.

Iron bacteria also cause the majority of problems in water pumps and pipes, blocking them with a brown, slimy, clay-like material (Fig. 3). Sometimes a bad odor comes from the water, like rotten eggs. Iron bacteria are not harmful to humans, but they are harmful to pipelines and pumps. If left untreated, they will cause the eventual blockage of the bore and failure of the pump.

Normally, iron bacteria refer to those in which the oxidation of iron is an important source for their metabolic energy, but another type of iron bacteria exists: iron-reducing bacteria, which get metabolic energy by reducing ferric iron to ferrous iron under anoxic or anaerobic condition with the presence of organic matters or other electron donors. Important ferric reducers include *Shewanella*, *Geobacter*, *Geospirillum*, and *Geovibrio*. Iron-reducing bacteria are important at natural cleanup and energy production.

Figure 1. *T. ferrooxidans* cell suspension viewed by an electron microscope (magnified 30,000 times) (1).

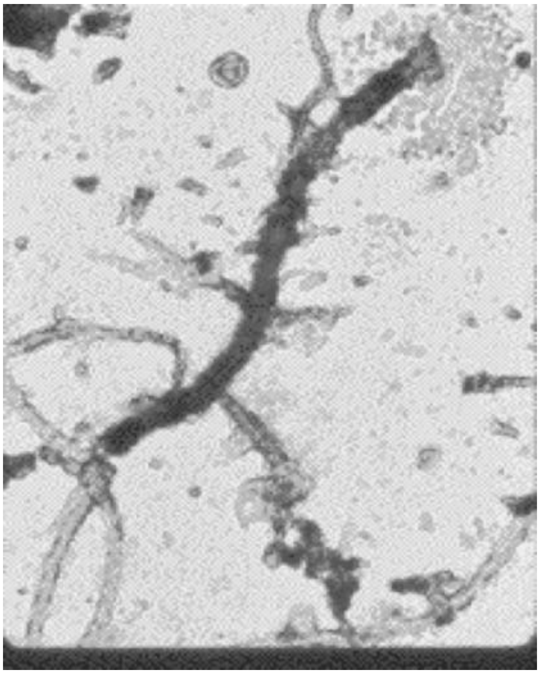

Figure 2. Iron and manganese oxide precipitants deposited on filaments of *Leptothrix ferrooxidans* (2).

IRON-OXIDIZING BACTERIA

Most iron-oxidizing bacteria grow and multiply in water and use dissolved iron as part of their metabolism. They oxidize ferrous (Fe^{2+}) iron into its insoluble ferric state (Fe^{3+}) and deposit it in the slimy gelatinous material that

Figure 3. The cut section of a cast-iron pipe shows a severe iron bacteria buildup on the pipe walls.

surrounds their cells. These filamentous bacteria grow in stringy clumps and are found in most iron-bearing surface waters. They have been known to proliferate in waters containing iron as low as 0.1 mg/L. To carry out the oxidation function, these aerobic bacteria need at least 0.3 ppm of dissolved oxygen in the water.

Iron-Oxidizing Bacteria in Groundwater and Surface Water System

Although iron bacteria do not cause health problems in people, they may have the following unpleasant and possibly expensive effects: cause unpleasant odors and tastes in water, corrode plumbing equipment, reduce well yields (clog screens and pipes), and increase chances of sulfur bacteria infestation for pipeline corrosion. Iron bacteria can build up in a stagnant section of the water supply system, such as the quick disconnect fittings of recoil hoses that are not regularly flushed. If the iron bacteria buildup becomes thick enough, anaerobic conditions may develop at the wall of the fitting or pipe that can cause corrosion of stainless steel or growth of sulfate-reducing bacteria.

The growth of iron bacteria also causes biofouling and blockage in groundwater bores. Although most groundwater contains some iron, iron bacteria do not affect all bores in the absence of oxygen. When iron bacteria infect a bore, the resulting growth may be suspended in water or deposited onto the bore hole equipment. Iron oxide (because of the oxidation by iron bacteria) suspended in the bore water is usually exhibited as biofouling in the irrigation pipes, which will deteriorate the pipeline surface and cause corrosion and shorten the pipeline lifetime. Bacterial growth deposits impact the bore hole equipment as well as the irrigation plumbing.

Normally, the presence of iron and oxygen in water triggers the growth of iron bacteira. Although iron is present in water, it is seldom found at concentrations greater than 10 mg/L or 10 parts per million (ppm). The U.S. Environmental Protection Agency (EPA) did not include iron in the National Primary (health-related) Drinking Water Regulations, but they did set a Secondary Regulation limit of 0.3 mg/L based on aesthetic and taste considerations. Surface water supplies usually have concentrations below the standard level (0.05–0.20 mg/L), although higher levels may be present because of corrosion of pipes. Higher iron levels (1.0–10 mg/L) are most common in groundwater supplies.

Iron-Oxidizing Bacteria in Acid Mining Drainage (AMD)

Mine drainage is metal-rich water formed from chemical reaction between water and rocks containing sulfur-bearing minerals. The runoff formed is usually acidic and frequently comes from areas where ore- or coal-mining activities have exposed rocks containing pyrite—a sulfur-bearing mineral with the major component of FeS. When pyrite is exposed to air and reacts with oxygen and water, sulfuric acid and dissolved iron (Fe^{2+}) are produced (Reaction 1). Ferrous iron is oxidized to ferric (Fe^{3+}) when reacting with oxygen (Reaction 2) and forms ferric hydroxide ($Fe(OH)_3$, Reaction 3). Ferric hydroxide precipitates to form the red, orange, or yellow sediments in the bottom of streams containing mine drainage. The acid runoff further dissolves heavy metals such as copper, lead, and mercury into ground or surface water. In the natural environment, the formation of acid mine drainage would take a long time to develop. But the rate and degree by which acid mine drainage proceeds can be dramatically sped up by the action of certain bacteria (e.g., *Thiobacillus ferrooxidans*).

$$4FeS_{2(s)} + 14O_{2(g)} + 4H_2O_{(l)} \rightarrow 4Fe^{2+}{}_{(aq)} + 8SO_4{}^{2-}{}_{(aq)} + 8H^+{}_{(aq)} \quad (1)$$

$$4Fe^{2+}{}_{(aq)} + O_{2(g)} + 4H^+{}_{(aq)} \rightarrow 4Fe^{3+}{}_{(aq)} + 2H_2O_{(l)} \quad (2)$$

$$4Fe^{3+}{}_{(aq)} + 12H_2O_{(l)} \rightarrow 4Fe(OH)_{3(s)} + 12H^+{}_{(aq)} \quad (3)$$

The overall reaction for pyrite oxidation and precipitation (Reaction 4) is as follows:

$$4FeS_{2(s)} + 15O_{2(g)} + 14H_2O_{(l)} \rightarrow 4Fe(OH)_{3(s)} + 8SO_4{}^{2-}{}_{(aq)} + 16H^+{}_{(aq)} \quad (4)$$

In acid creeks in mining regions like Contrary Creek, VA (Fig. 4), iron-rich precipitates often contain sulfur and are yellow. Many iron-oxidizing bacteria such as *T. thiooxidans* and *T. ferrooxidans* produce this "yellow boy," a yellowish-orange precipitate that turns the acidic runoff in the streams to an orange or red color and covers the stream bed with a slimy coating.

Although iron-oxidizing bacteria play a nuisance role in acid mine drainage, they are the central role for microbial mining, in which the leaching of copper from low-grade ore can recover copper from the drainage water of mines and improve the copper content and ore grade. Many ores that are sources of metals (copper) are sulfide (HS^-), which are highly insoluble minerals. If the concentration of metal in the ore is low, it may not be economically feasible to concentrate the mineral by conventional chemical means. On the other hand, the oxidized status of sulfide–sulfate ($SO_4{}^{2-}$) forms a water-soluble mineral, which can improve

Figure 4. Acid creeks in mining regions at Contrary Creek, VA (picture from USGS website) (3).

the copper content in ores. The natural oxidization from sulfide to sulfate is slow, but it could be catalyzed with the presence of iron-oxidizing bacteria (e.g., *T. ferrooxidans*), thus aiding in the solubilization of the metal (Reactions 5 and 6). Through these processes, the soluble copper sulfate can be recovered from ore (4).

$$Fe^{2+} + \tfrac{1}{4}O_2 + H^+ \rightarrow Fe^{3+} + \tfrac{1}{2}H_2O \quad \text{(by } \textit{T. ferrooxidans}\text{, or } \textit{L. ferrooxidans}\text{)} \qquad (5)$$

$$CuS(\text{insoluble copper}) + 8Fe^{3+} + 4H_2O \rightarrow Cu^{2+}(\text{soluble copper}) + 8Fe^{2+} + SO_4{}^{2-} + 8H^+ \qquad (6)$$

IRON-REDUCING BACTERIA IN BIOREMEDIATION AND ENERGY PRODUCTION

Besides aerobic iron-oxidizing bacteria (normally referred to as iron bacteria), another type of bacteria exists: anaerobic iron-reducing bacteria, which use ferric iron as an electron acceptor for energy metabolism. A wide variety of both chemoorganotrophic and chemolithotrophic bacteria can have this feature. Ferric reduction can be coupled to the oxidation of several organic and inorganic electron donors. Various organic compounds, including aromatic compounds, can be oxidized anaerobically by iron-reducers. Much research on the energetics of ferric iron reduction has been done in the gram-negative bacterium: *Shewanella putrefaciens*, in which ferric-dependent anaerobic growth occurs with various organic electron donors. Other important ferric reducers include *Geobacter, Geospirillum*, and *Geovibrio*.

Geobacter metallireducens has been a model for study of the physiology of Fe^{3+} reduction. This organism can oxidize acetate with Fe^{3+} as an acceptor as follows:

$$\text{Acetate} + 8Fe^{3+} + 4H_2O \rightarrow 2HCO_3{}^- + 8Fe^{2+} + 9H^+$$

Geobacter can also use H_2 or other organic electron donors including aromatic hydrocarbon toluene as the electron donor, which may be environmentally significant because toluene from accidental spills or leakage from hydrocarbon storage tanks often contaminates ferric-rich aquifers, and it has been suggested that organisms like *Geobacter* may be natural cleanup agents in such environments (4).

Another distinct feature of the *Geobacter* species is that it has the ability to directly transfer electrons onto the surface of electrodes, which has made it possible to design novel microbial fuel cells that can efficiently convert waste organic matter to electricity. It has been found that dissimilatory metal-reducing microorganisms, such as *Geobacter* and *Rhodoferax* species (5,6), have the novel ability to directly transfer electrons throughout the cell membrane onto the surface of electrodes, which has led to the construction of microbial fuel cells that are more efficient than are traditional microbial fuel cells in that previously described microbial fuel cells depended on the toxic electron shuttling mediator (e.g., thionine, benzylviologen, phenazines, and phenoxazines) (7,8) to transfer electrons from cell membrane to electrode. As *Geobacter* can transfer electrons onto the electrodes by themselves, no electron shuttling mediator is needed, which can reduce the toxicity to bacterium and make microbial fell cells more cost-effective (Fig. 5). In addition, because *Geobacter* can biodegrade many contaminants, it is possible to harvest electricity from many types of waste organic matter or renewable biomass (5).

SOLUTION FOR BIOFOULING AND WATER COLOR CAUSED BY IRON BACTERIA

Because of the expensive cost incurred by biofouling, corrosion, and red water in the water supply system, efficient approaches should be employed to prevent and control the growth of iron bacteria in water supply and groundwater systems. Currently, chemical and mechanical methods for treating iron bacteria problems are available.

Chemical Treatment

Because it is difficult to get rid of iron bacteria once they exist in well systems, prevention is the best safeguard against accompanying problems. For well drillers, prevention means disinfecting everything that goes into the ground with a strong chlorine solution (250 ppm). As iron bacteria are nourished with the presence of carbon and other organics, it is essential

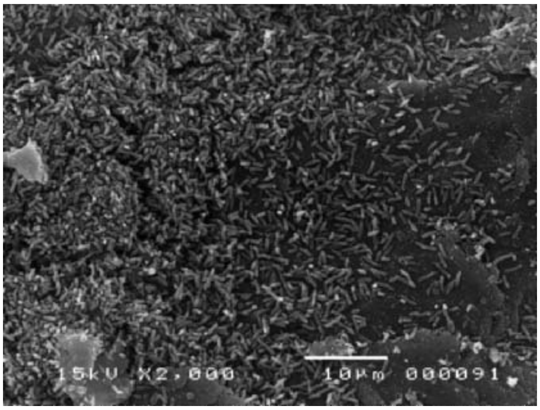

Figure 5. *Geobacter* colonizing a graphite electrode surface in a microbial fuel cell (9).

that these are not introduced into any part of the well system during the drilling process. Also, temperature is normally low in groundwater drilling systems, which leads to the much slower chemical reactions in wells. Therefore, bacterial cells need a long exposure to the chemical for the treatment to be effective.

Mechanical Treatment

In addition to the chemical treatment, mechanical approaches are available to control the iron bacteria in community water systems. Stagnant water conditions can be avoided by looping dead-end plumbing lines and periodically flushing low-flow lines to reduce bacteria. Forcing hot water or steam into a well to disperse the slime and kill the bacteria has also worked well. In addition, flushing large quantities of heated water into the aquifer has been found successful in field tests.

BIBLIOGRAPHY

1. Ehrlich, H.L. (1990). *Geomicrobiology*, 2nd Edn. Marcel Dekker, New York.
2. http://soils1.cses.vt.edu/ch/biol_4684/Microbes/Leptothrix.html.
3. http://pubs.usgs.gov/publications/text/Norrieviewing.html.
4. Madican, M.T., Martinko, J.M., and Parker, J. (2003). *Brock Biology of Microorganisms*. Person Education, Upper Saddle River, NJ, pp. 588 and 671.
5. Bond, D.R., Holmes, D.E., Tender, L.M., and Lovley, D.R. (2002). Electrode-reducing microorganisms that harvest energy from marine sediments, *Science* **295**: 483–485.
6. Chaudhuri, S.K. and Lovley, D.R. (2003). Electricity generation by direct oxidation of glucose in mediatorless microbial fuel cells. *Nature Biotechnol.* **21**(10): 1229–1232.
7. Roller, S.D. et al. (1984). *J. Chem. Technol. Biotechnol.* **34B**: 3.
8. Kim N., Choi, Y., Jung, S., and Kim, S. (2000). *Biotechnol. Bioeng.* **70**: 109.
9. Bond, D.R. and Lovley, D.R. (2003). Electricity production by *Geobacter sulfurreducens* attached to electrodes. *Appl. Environ. Microbiol.* **69**(3): 1548–1555.

CARTRIDGE FILTERS FOR IRON REMOVAL

D.V. Vayenas
G. Lyberatos
University of Ioannina
Agrinio, Greece

INTRODUCTION

The presence of iron in a water supply at concentrations that exceed the secondary drinking water standards of 0.2 mg/L (1) is undesirable for the following reasons (2): (1) iron precipitates give water a reddish color when exposed to air, (2) iron gives water an undesirable metallic taste, (3) deposition of iron in the distribution system can reduce the effective pipe diameter and eventually clog a pipe, and (4) iron is a substrate for bacterial growth in water mains. When iron bacteria die and slough off, bad odors and unpleasant tastes may result. Consequently, it is very important to remove iron from surface water before it may be used. Iron may be removed from water through a physical-chemical method, a biological method, or a combination of the two.

The oxidation state of iron may be either +2 (ferrous) or +3 (ferric). Surface water iron is generally found in its precipitated ferric form and does not require any further treatment to oxidize it. Groundwater iron is in its ferrous state in the deeper layers of some water reserves that lack oxygen, or in groundwater, in which case it is in a reduced dissolved form (Fe^{2+}) and often chelated. The oxidation state of iron in water depends on the pH and the redox potential. A dissolved form of iron (Fe^{2+} or $FeOH^-$) can be transformed to a precipitated form [$FeCO_3$, $Fe(OH)_2$, or $Fe(OH)_3$] by increasing either the oxidation potential, the pH, or both (3).

Aeration followed by solid–liquid separation is the most commonly used physical-chemical treatment method for iron removal. Aeration is generally recommended for oxidizing ferrous iron in water with high iron concentrations (>5 mg/L), so that costs for chemicals can be avoided (4). Sedimentation and/or filtration (sand or dual-media filtration) are used for the solid–liquid separation. If necessary, other treatments may be added, such as pH correction (lime softening), chemical oxidation by adding an oxidizing agent (such as chlorine, potassium permanganate, or ozone), ion oxidation, settling, and biological oxidation. Unfortunately, several problems are associated with physical-chemical treatments for iron removal that lead to relatively low iron removal rates. Substances such as humic acids, silicates, phosphates, and polyphosphates inhibit oxidation, precipitation, and filtration of ferric hydroxide. Thus, adding reagents is often necessary for pH correction and flocculation. The iron is not retained in a compact form, thus leading to low removal capacity, low filtration rates, large amounts of wash water, and difficult sludge treatment. The effective size of the filter medium can be between 0.5 and 1 mm and filtration rates between 5 and 15 m/h.

The rates of ferrous iron oxidation by air increase with pH; about 90% conversion may be achieved in 10–20 min at pH 7 (5). Iron oxidation is substantially slower at a pH below 6 and reduced forms may persist for some time in aerated waters (6). The rates of ferrous iron oxidation may be increased by the action of microorganisms.

Several genera of bacteria oxidize dissolved iron by different mechanisms. *Gallionella* sp., *Leptothrix ochracea*, and *Crenothrix polyspora* cause intracellular oxidation by enzymatic action; extracellular oxidation is caused by the catalytic action of polymers excreted by the filaments comprising the stalks produced by *Gallionella* sp.; and the sheaths from *Leptothrix* sp., *Crenothrix* sp., *Clonothrix* sp., and *Sphaerotilus* sp. Secondary extracellular oxidation is also caused by the catalytic action of polymers, excreted by the microscopic filaments of capillary appearance that are abandoned by the gliding trichomes of *Toxothrix thrichogenes* and extracellular polymers that are excreted by various Siderocapsaceae (7).

Biological iron removal systems have the following advantages compared with physicochemical iron removal systems: they require smaller installations because higher filtration rates may be used and/or because aeration and filtration can take place simultaneously in the same vessel; they allow longer lasting filtration runs because the retention time of iron in the filter increases due to the formation of denser precipitates and the use of coarser media; they generate denser backwash sludge that is easier to thicken and dewater; they lead to higher net production of clarified water because less water is required for backwashing, though it is possible to use raw water for backwash; they require no added chemicals; they do not deteriorate water quality over time; and they require lower capital and operating outlays (8).

The metabolic activities of iron bacteria are not fully understood, but it is believed that the same oxidation of iron is carried out by some variation in the physical-chemical reaction,

$$4Fe^{2+} + O_2 + 10H_2O \rightarrow 4Fe(OH)_3 + 8H^+ + \text{energy}$$

Whatever the metabolic pathway for the iron oxidation reactions, the biological process is catalytic and causes rapid oxidation. The red insoluble precipitates are all slightly hydrated iron oxides that, beneficially, are forms more compact than the precipitates formed when using physical-chemical processes. This feature partially explains the greater iron retention capacity between backwashes of biological filters compared to physical-chemical treatment filters.

It is important to recognize that iron bacteria catalyze oxidation reactions under conditions of pH and E_h that are intermediate between those of natural groundwater and those required for conventional treatment. The range of conditions allowing biological iron oxidation thus straddles the theoretical boundaries between the fields of Fe^{2+} and Fe^{3+} stability expected strictly by chemical thermodynamics. This can be visualized in a stability diagram, using pH and E_h as ordinates, as shown in Fig. 1.

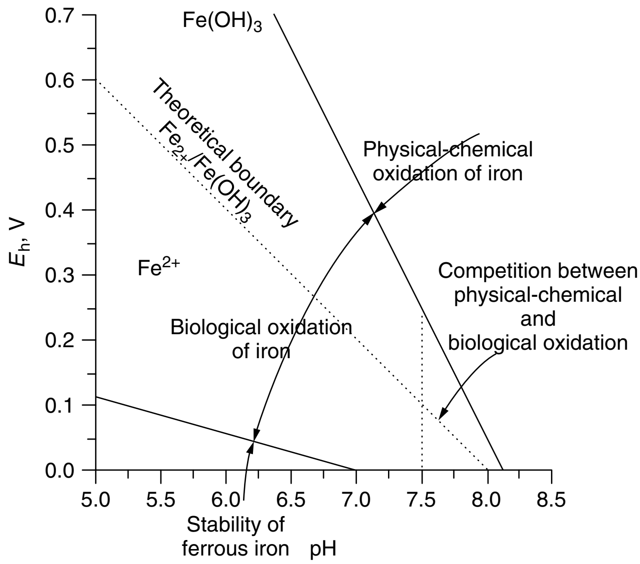

Figure 1. Stability diagram for biological iron removal.

By aerating the raw water and raising the dissolved oxygen concentration, the raw water's rH can be increased into the pH and E_h range where biological oxidation can take place. By controlling the quantity of supplied air, the rH can be adjusted to a level that is ideal for iron removal.

Once a biological iron removal plant is constructed, the system must be given time to "seed" with bacteria naturally present in the water source. This seeded biomass is naturally and continuously regenerated during the life of the plant and is periodically removed partially by backwashing. For iron removal plants, the seeding period is from 1 day to 1 week. At the end of the seeding period, the metal concentration in the effluent falls to near detection levels. The seeding period is affected by the temperature of the water source and is generally longer for cold water sources.

Iron bacteria are generally robust, and because of the variety of species involved, one type or another can thrive under most environmental conditions. Given the correct pH, from 6–8, and E_h, the bacteria can normally oxidize iron at temperatures ranging from 5 to 50 °C. Inhibition of the biological process can, however, be caused by H_2S, chlorine, NH_4^+, and some heavy metals normally not present in water sources.

AMMONIA AND MANGANESE INHIBITION OF BIOLOGICAL IRON REMOVAL

Water sources that contain iron often also contain ammonia and manganese. Ammonia, iron, and manganese may be removed chemically or biologically from a water supply. Biological removal of these pollutants is preferable because there is no need to add extra chemicals and the volume of sludge generated is appreciably smaller and hence easier to handle (9). Simultaneous biological removal of the three elements is very difficult mainly due to the different redox potential values needed to oxidize them. Mouchet (7) summarized the methods for iron and manganese removal from potable water and concluded that due to the different values required for dissolved oxygen concentration, redox potential, and pH, the basic design of plants for iron and manganese removal is substantially different. He also concluded that it is generally not possible to achieve simultaneous removal of Fe^{2+} and Mn^{2+} in a single filter, except in some cases of extremely low filtration rates. Gouzinis et al. (10) studied the simultaneous removal of ammonia, iron, and manganese from potable water using a pilot-scale trickling filter and the influences these pollutants have on filter performance and efficiency. They concluded that iron has a strong negative effect on ammonia removal and ammonia has a very low impact on iron removal. They also concluded that iron has a strong effect on Mn removal and manganese seems to have a much lower effect on iron removal. The main conclusion of that work was that iron should be removed before ammonia and manganese oxidation. For low iron concentration, oxidation by extended aeration could reduce iron at very low levels, and subsequent biological ammonia and manganese removal could be carried out without any problems. For higher iron concentration, an initial biological iron oxidation stage should be necessary.

FILTERS FOR IRON REMOVAL

A key feature of biological iron removal lies in the high rate of iron oxidation. Thus, the engineering systems that have been developed for biological iron removal can be classified by the type of water flow and by aeration into pressurized filters, gravity-flow units (9), and trickling filters.

Pressurized Filters

In a pressurized unit, the main components of the plant (Fig. 2) are the reactor, which is a filter specially designed to operate at high rates and contains sand somewhat coarser than normal for conventional filters; a calibrated raw water aeration system where oxygen is supplied either by an air injector fitted with a rotameter (to adjust the air flow rate, thereby ensuring that E_h and dissolved oxygen concentrations in the preaerated raw water are conducive to optimum growth of iron bacteria) or by recirculating a portion of the treated water, aerated near saturation (this configuration is selected when the dissolved oxygen content at the filtration inlet must be very low and tightly controlled); and an optional aeration system for filter effluent, followed by an optional nonchlorinated washwater storage tank and a storage tank for chlorinated water (7).

Filters can be backwashed by using raw water (in which case, the need for a washwater storage tank is eliminated) or treated water, always beginning with a simultaneous air scour. Aeration systems are simple baffle mixers. Filtration rates can in certain cases attain 40 or even 50 m/h.

Gravity-Flow Units

Gravity-flow units contain an atmospheric pressure aeration system followed by an open or closed filter (Fig. 3). Filtration can also be carried out by gravity flow in these units. The atmospheric pressure aeration system may be cascade aeration (designed to supply water that features optimum oxidation conditions for biological iron removal, as determined by the raw water pH) or spray aeration. These types of open aeration systems are particularly recommended when the raw water contains H_2S. Gravity filtration is usually designed for plants of high capacity.

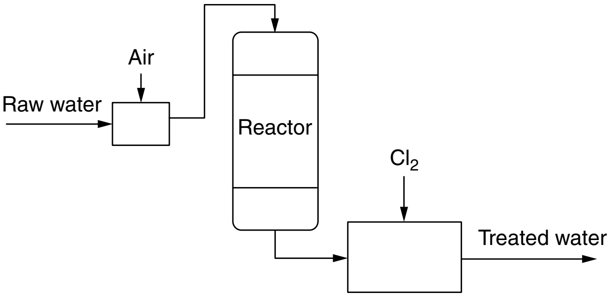

Figure 2. Process flow diagram for biological iron removal in a pressurized filter.

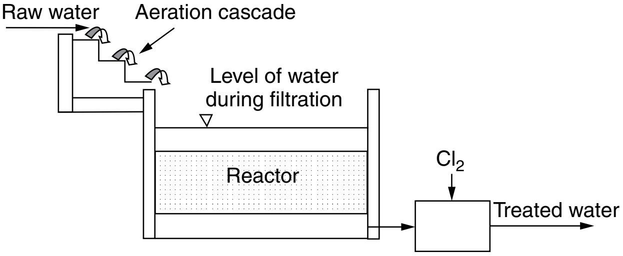

Figure 3. Process flow diagram for biological iron removal in a gravity plant.

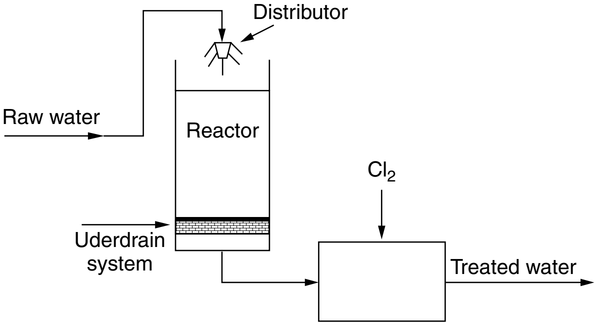

Figure 4. Process flow diagram for biological iron removal in a trickling filter.

Filtration rates can in certain cases attain 20–23 m/h. The filter medium is usually sand of an effective size between 1.0 and 1.5 mm (7).

Trickling Filters

The modern trickling filter consists of a bed of a highly permeable medium to which microorganisms are attached and through which water is percolated or trickled (Fig. 4). The filter media usually consist of either rock or a variety of plastic packing. Filters are constructed with an underdrain system for collecting the treated water and any biological solids that have become detached from the media. This underdrain system is important both as a collection unit and as a porous structure through which air can circulate. The collected liquid is passed to a settling tank where the solids are separated from the treated water (11).

An adequate flow of air is of fundamental importance to the successful operation of a trickling filter. The principal factors responsible for air flow in an open top filter are natural draft and wind forces. In natural draft, the driving force for airflow is the temperature difference between the ambient air and the air inside the pores. If the water is colder than the ambient air, the pore air will be cold and the direction of flow will be downward. If the ambient air is colder than the water, the flow will be upward. The latter is less desirable for mass transfer because the partial pressure of oxygen (and thus the oxygen-transfer rate) is lowest in the region of highest oxygen demand. Thus, a trickling filter has the advantage that it does not require an external air supply or an aeration system. If, in addition, the mean diameter of the filter media is sufficiently small (up to 5 mm), then complete aeration and very good filtration may be effected at the same time (12). Filtration rates are intermediate between those of pressurized filters and gravity-flow plants.

ADVANTAGES AND DISADVANTAGES OF FILTERS FOR IRON REMOVAL

Pressurized Filters

These filters are specially designed to operate at high rates and use particles of small mean diameter as support material, thus increasing the specific surface area available for ferrous iron oxidation. As a result, this type of filters exhibits high iron removal rates. On the other hand, the small particles lead to small pores, which may be clogged quite easily due to bacterial growth and iron precipitate filtration. Consequently, frequent backwashing is necessary. Finally, an external air supply is necessary, leading to increased operating cost.

Gravity Units

These filters are designed for plants where large quantities of drinking water have to be treated. The flow rates are quite low compared to the flow rates in pressurized systems. The support material is usually fine sand, which enables very good filtration but also needs frequent backwashing. The use of cascade aeration keeps the operating cost low and makes the system ideal for H_2S removal from drinking water. The capital cost of these systems is relatively high.

Trickling Filters

Trickling filters are quite simple constructions and do not require an external air supply or an aeration system, leading thus to low capital and operating costs. The support material is larger than that for pressurized systems to secure adequate air circulation. Filtration also takes place simultaneously with oxidation but a separate settling tank may be needed to ensure complete removal of iron precipitates. The rates in these filters lie between the high rates of pressurized systems and the low rates of gravity plants. Figure 5 presents typical iron concentration profiles along the depth of a pilot-scale trickling filter.

Technology offers various solutions for iron removal from drinking water. Cartridge filters combine physical-chemical and biological iron removal. The type of filter that is used depends on the physicochemical characteristics of the water, iron concentration, volumetric flow rate, and, of course, the costs. Thus, a feasibility study is necessary before deciding on the type of cartidge filter that will be used.

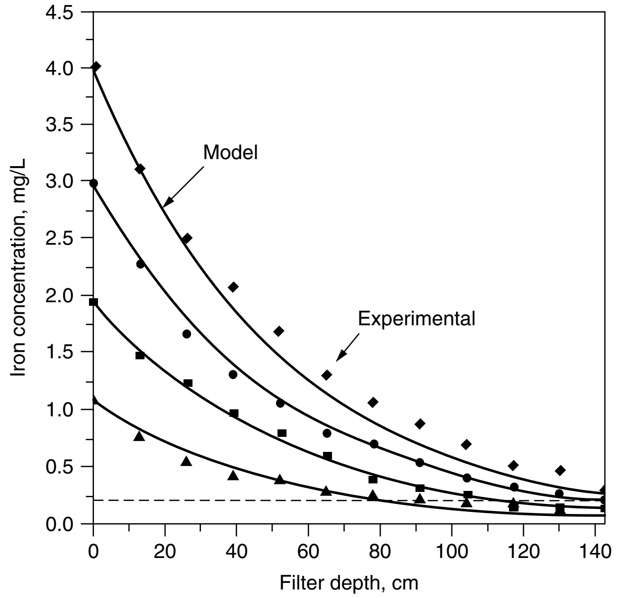

Figure 5. Typical iron concentrations profiles along the depth of a pilot-scale trickling filter (13).

BIBLIOGRAPHY

1. *EEC-Official Journal of the European Communities*, No 80/779, August 30, 1980.
2. Kontari, N. (1988). Groundwater, iron and manganese: an unwelcome trio. *Water/Eng. Manage.* 25–26.
3. Hem, J.D. (1961). Stability field diagrams as aids in iron chemistry studies. *J. Am. Water Works Assoc.* **53**(2): 211.
4. Wong, J.M. (1984). Chlorination-filtration for iron and manganese removal. *J. Am. Water Works Assoc.* **76**: 76–79.
5. Sundstrom and Klei (1979). *Wastewater Treatment*. Prentice-Hall, Englewood Cliffs, NJ.
6. Sawyer, C.N. and McCarty, P.L. (1978). *Chemistry for Environmental Engineering*, 3rd Edn. McGraw-Hill, Singapore.
7. Mouchet, P. (1992). From conventional to biological removal of iron and manganese in France. *J. Am. Water Works Assoc.* **84**(4): 158–166.
8. Gage, B., O'Dowd, D.H., and Williams, P. (2001). Biological iron and manganese removal, pilot and full scale applications. *3rd Water Works Assoc. Conf.*, Ontario.
9. Degremont, (1991). *Water Treatment Handbook*, 6th Edn. Lavoisier, Paris.
10. Gouzinis, A., Kosmidis, N., Vayenas, D.V., and Lyberatos, G. (1998). Removal of Mn and simultaneous removal of NH_3, Fe and Mn from potable water using a trickling filter. *Water Res.* **32**(8): 2442–2450.
11. Metcalf & Eddy. (1991). *Wastewater Engineering, Treatment, Disposal, Reuse*, 3rd Edn. McGraw-Hill, New York.
12. Dimitrakos Michalakos, G., Martinez Nieva, J., Vayenas, D.V., and Lyberatos, G. (1997). Removal of iron from potable water using a trickling filter. *Water Res.* **31**(5): 991–996.
13. Tekerlekopoulou, A.G., Vasiliadou, I., Vakianis, I., and Vayenas, D.V. (2004). Physicochemical and biological iron removal from potable water. *Water Res.* (submitted).

IRRIGATION WATER QUALITY IN AREAS ADJOINING RIVER YAMUNA AT DELHI, INDIA

Chakresh K. Jain
Mukesh K. Sharma
National Institute of Hydrology
Roorkee, India

The groundwater quality of areas adjoining the River Yamuna in Delhi (India) has been assessed for the suitability of groundwater for irrigation. Thirty-eight

groundwater samples from shallow and deep aquifers were collected each during pre- and postmonsoon seasons in the year 2000 and analyzed for various parameters. The suitability of groundwater for irrigation has been evaluated based on salinity, sodium adsorption ratio (SAR), residual sodium carbonate (RSC), and boron content. According to the U.S. Salinity Laboratory classification of irrigation water, more than 50% of the samples fall in water type C3-S1 (high salinity and low SAR). Such water cannot be used on soils that have restricted drainage. Even with adequate drainage, special management for salinity control may be required and plants with good tolerance should be selected. About 30% of the samples fall in water type C3-S2 (high salinity and medium SAR). Such water will induce an appreciable sodium hazard in finely textured soils that have good cation exchange capacity, especially under low leaching conditions.

INTRODUCTION

Groundwater plays an important role in agriculture for both watering crops and for irrigating dry season crops. It is estimated that about 45% of the irrigation water required is from groundwater sources. Unfortunately, the availability of groundwater is not unlimited, nor is the groundwater protected from deterioration. In most instances, the extraction of excessive quantities of groundwater has resulted in drying up of wells, damaged ecosystems, land subsidence, saltwater intrusion, and depletion of the resource. Groundwater quality is being increasingly threatened by agricultural, urban, and industrial wastes. It has been estimated that once pollution enters the subsurface environment, it may remain concealed for many years, disperses over wide areas of the groundwater aquifer, and renders groundwater unsuitable for consumption and other uses. The rate of depletion of groundwater levels and the deterioration of groundwater quality are of immediate concern in major cities and towns of the country.

The National Capital Territory (NCT) of Delhi is facing severe problems in managing groundwater quality and quantity (1). Surface waterbodies play a significant role in groundwater flow. The hydraulic gradient plays a significant role in the lateral and vertical migration of contaminants in groundwater aquifers. Therefore, the present study has been carried out to assess the suitability of groundwater for irrigation in areas adjoining the River Yamuna at Delhi.

STUDY AREA

Delhi generates about 1900 MLD of sewage against installed capacity of 1270 MLD of sewage treatment. The balance of untreated sewage along with a significant quantity of partially treated sewage is discharged into the River Yamuna every day. The river receives sewage and industrial wastes through sixteen drains, which join River Yamuna between Wazirabad and Okhla. Thus Delhi is the largest contributor of pollution to the River Yamuna, which receives almost 80% of its pollution load through these drains.

The climate of Delhi is influenced mainly by its inland position and the prevalence of the continental type air during the major part of the year. Extreme dryness with an intensely hot summer and cold winter are the characteristics of the climate. During the monsoon months, air of oceanic origin penetrates to this area and causes increased humidity, cloudiness, and precipitation. The normal annual rainfall in the National Capital Territory of Delhi is 611.8 mm. The rainfall increases from southwest to northeast; about 81% of the annual rainfall is received during the three monsoon months of July, August, and September; the balance of the annual rainfall is received as winter rains and as thunderstorm rain during pre- and postmonsoon months.

METHODOLOGY

Groundwater samples were collected in polyethylene bottles during the pre- and postmonsoon seasons during the year 2000 from areas adjoining the River Yamuna at Delhi and were preserved by adding an appropriate reagent (2,3). All samples were stored in sampling kits maintained at 4 °C and brought to the laboratory for detailed physicochemical analysis. The details of sampling locations are given in Table 1. The physicochemical analyses were performed following standard methods (2,3).

RESULTS AND DISCUSSION

The hydrochemical data for the two sets of samples collected from the areas adjoining the River Yamuna at Delhi (Fig. 1) during the pre- and postmonsoon seasons are given in Table 2. The pH of the groundwater of areas adjoining the River Yamuna are mostly within the range of 6.7 to 8.3 during the premonsoon season and 6.6 to 8.2 during the postmonsoon season; most of the samples point toward the alkaline range in both seasons. The total dissolved solids (TDSs) in the groundwater vary from 402 to 2266 mg/L during the premonsoon season and from 397 to 2080 mg/L during the postmonsoon season. Higher values of TDSs in areas nearby the River Yamuna indicate high mineralization of the groundwater.

Alkalinity varies from 116 to 380 mg/L during the premonsoon season and from 106 to 310 mg/L during the postmonsoon season. From the view point of hardness, only one sample exceeds the maximum permissible limit of 600 mg/L. The concentration of chloride was within the permissible limit. The sulfate content limit in groundwater is exceeded at two locations (Himmat Puri and Golf Course). The nitrate content in drinking water is considered important because of its adverse health effects. The nitrate content in the groundwater of areas adjoining the River Yamuna varies from 1 to 286 mg/L during the premonsoon season. The higher level of nitrate at certain locations may be attributed to surface disposal of domestic sewage and runoff from agricultural fields. It has also been observed that the groundwater samples collected from hand pumps at various depths have a high nitrate

Table 1. Description of Groundwater Sampling Locations

Sample No.	Label*	Location	Type of Well (Depth in ft)[a]
1	1	Bhagwanpur Khera	HP (25)
2	2	Loni Road	BW (20)
3	3	Kabul Nagar	OW (35)
4	4	Naveen Shahdara	HP (20)
5	5	Seelampur	HP (30)
6	6	Shastri Park	HP (25)
7	7	Lakshmi Nagar	HP (100)
8	8	Prit Vihar	TW (40)
9	9	Shankar Vihar	HP (200)
10	A	Pratap Nagar	TW (100)
11	B	Himmatpuri	HP (80)
12	C	Civil Lines	TW (250)
13	D	Rajpur road	OW (25)
14	E	Malka Gunj	HP (20)
15	F	Tripolia	BW (100)
16	G	Gulabi Bagh	BW (40)
17	H	Gulabi Bagh	BW (160)
18	I	Shastri Nagar	BW (100)
19	J	Shastri Nagar	BW (60)
20	K	Lekhu Nagar	HP (40)
21	L	Ram Pura	HP (40)
22	M	Punjabi Bagh West	HP (40)
23	N	Rajghat	BW (50)
24	P	JLN Marg	HP (60)
25	Q	GB Pant Hospital	BW (20)
26	R	Panchkuin Marg	TW (450)
27	S	Panchkuin Marg	HP (60)
28	T	Rajendra Nagar	BW (100)
29	U	Rajendra Nagar	BW (300)
30	V	Shankar Road	OW (20)
31	W	IARI, Pusa	TW (120)
32	X	Zoological Park	HP (60)
33	Y	Golf Course	BW (80)
34	Z	Rabindra Nagar	TW (150)
35	a	Teen Murti Chowk	HP (60)
36	b	Malcha Marg	BW (100)
37	c	Sardar Patel Road	TW (150)
38	d	Janpath	BW (250)

[a] HP: hand pump; OW: open well; BW: bore well; TW: tube well.

content, which may be attributed to wellhead pollution. The fluoride content was within the maximum permissible limit of 1.5 mg/L in all samples.

WATER QUALITY EVALUATION FOR IRRIGATION

Irrigation water quality refers to its suitability for agricultural use. The concentration and composition of dissolved constituents in water determine its quality for irrigation. The quality of water is an important consideration in any appraisal of salinity or alkali conditions in an irrigated area. Good quality water can result in maximum yield under good soil and water management practices. The most important characteristics of water that determine the suitability of groundwater for irrigation are as follows:

1. Salinity
2. Relative proportion of sodium to other cations
3. Residual sodium carbonate
4. Boron content

The safe limits of electrical conductivity for crops of different degrees of salt tolerance under varying soil textures and drainage conditions are given in Table 3. The quality of water is commonly expressed by classes of relative suitability for irrigation with reference to salinity levels. The recommended classification with respect to electrical conductivity, sodium content, sodium absorption ratio (SAR), and residual sodium carbonate (RSC) are given in Table 4. The values of SAR, % Na, and RSC for groundwater samples collected from areas adjoining the River Yamuna are given in Table 5.

Salinity

Salinity is broadly related to total dissolved solids (TDSs) and electrical conductivity (EC). A high concentration of TDSs and electrical conductivity in irrigation water may increase the soil salinity, which affect the salt intake of the plant. The salts in water affect plant growth directly and also affect the soil structure, permeability, and aeration, which indirectly affect plant growth. Soil water passes into the plant through the root zone by osmotic pressure. As the dissolved solid content of the soil water in the root zone increases, it is difficult for the plant to overcome the osmotic pressure and the plant root membranes are unable to assimilate water and nutrients. Thus, the dissolved solids content of the residual water in the root zone also has to be maintained within limits by proper leaching. These effects are visible in plants by stunted growth, low yield, discoloration, and even leaf burns at the margin or top.

Relative Proportion of Sodium to Other Cations

A high salt concentration in water leads to the formation of saline soil, and high sodium leads to the development of an alkaline soil. The sodium or alkali hazard in the use of water for irrigation is determined by the absolute and relative concentration of cations and is expressed in terms of the sodium adsorption ratio (SAR). If the proportion of sodium is high, the alkali hazard is high; conversely, if calcium and magnesium predominate, the hazard is lower. There is a significant relationship between SAR values of irrigation water and the extent to which sodium is absorbed by soil. If water used for irrigation is high in sodium and low in calcium, the cation-exchange complex may become saturated with sodium. This can destroy the soil structure owing to dispersion of clay particles. A simple method of evaluating the danger of high-sodium water is the sodium adsorption ratio (SAR) (4):

$$\mathrm{SAR} = \frac{\mathrm{Na}^+}{\sqrt{(\mathrm{Ca}^{2+} + \mathrm{Mg}^{2+})/2}} \tag{1}$$

The percentage sodium is calculated as

$$\%\mathrm{Na} = \frac{\mathrm{Na}^+ + \mathrm{K}^+}{\mathrm{Ca}^{2+} + \mathrm{Mg}^{2+} + \mathrm{Na}^+ + \mathrm{K}^+} \times 100 \tag{2}$$

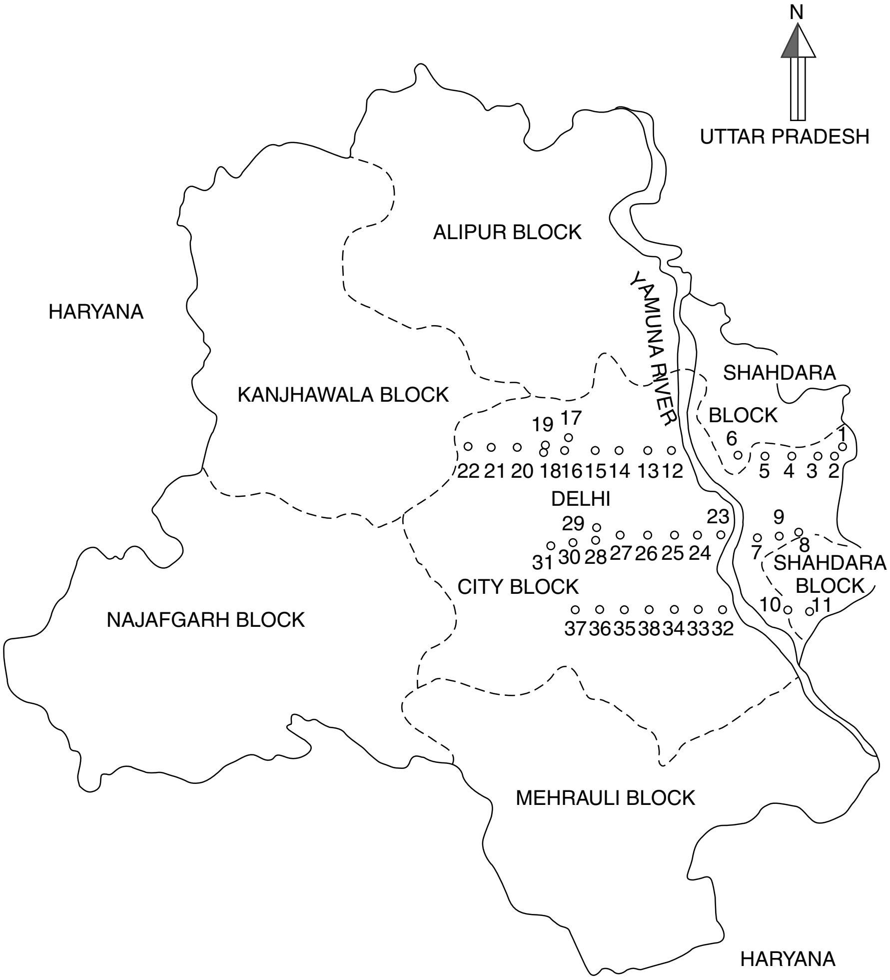

Figure 1. Study area showing sampling locations.

where all ionic concentrations are expressed in milliequivalents per liter.

The values of SAR and percentage sodium of the groundwater samples collected from the area adjoining the River Yamuna are given in Table 5. Calculation of SAR for a given water provides a useful index of the sodium hazard of that water for soils and crops. A low SAR (2 to 10) indicates little danger from sodium; medium hazards are between 7 and 18, high hazards between 11 and 26, and very high hazards above that. The lower the ionic strength of the solution, the greater the sodium hazards for a given SAR (4).

Table 2. Hydrochemical Data on Groundwater Samples from Delhi (Pre- and Postmonsoon, 2000)

Parameter	Minimum[a]		Maximum[a]		Mean[a]	
pH	6.7	(6.6)	8.3	(8.2)	7.1	(7.1)
Conductivity, μS/cm	628	(620)	3540	(3250)	1463	(1370)
TDSs, mg/L	402	(397)	2266	(2080)	936	(879)
Alkalinity, mg/L	116	(106)	380	(310)	201	(194)
Hardness, mg/L	116	(114)	841	(792)	235	(225)
Chloride	17	(17)	400	(390)	86	(86)
Sulfate, mg/L	43	(48)	690	(680)	180	(177)
Nitrate, mg/L	1.0	(ND)	286	(287)	78	(68)
Phosphate, mg/L	0.07	(0.03)	0.28	(0.21)	0.16	(0.08)
Fluoride, mg/L	0.33	(ND)	1.31	(0.76)	0.80	(0.41)
Sodium, mg/L	55	(51)	340	(322)	160	(157)
Potassium, mg/L	4.2	(2.2)	121	(103)	26	(20)
Calcium, mg/L	25	(21)	240	(227)	59	(59)
Magnesium, mg/L	9.0	(7.0)	64	(56)	21	(19)

[a] Values in parentheses represent postmonsoon data.

Table 3. Safe Limits of Electrical Conductivity for Irrigation Water

Sample No.	Nature of Soil	Crop Grown	Upper Permissible Safe Limit of EC in Water, μS/cm
1	Deep black soil and alluvial soils having clay content of more than 30% that are fairly to moderately well drained	Semitolerant	1500
		Tolerant	2000
2	Textured soils having clay content of 20–30% that are well drained internally and have good surface drainage	Semitolerant	2000
		Tolerant	4000
3	Medium textured soils having clay of 10–20%, internally very well drained, and having good surface drainage	Semitolerant	4000
		Tolerant	6000
4	Light textured soils having clay of less than 10% that have excellent internal and surface drainage	Semitolerant	6000
		Tolerant	8000

The SAR values in the groundwater of the study area ranged from 1.70 to 12.06 during the premonsoon season and from 1.65 to 11.24 during postmonsoon. As evident from the SAR values, the groundwater of the study area falls in the category of low sodium hazard except for two samples (Civil Lines and Sardar Patel Road), which reveals that the groundwater in the study area is free from any sodium hazard. The percentage sodium in the study area varied from 30.4 to 83.7 during the premonsoon season and from 34.5 to 82.9 during postmonsoon. More than 50% of the samples in the study area exceeded the recommended value of 60% during both pre- and postmonsoon seasons.

Residual Sodium Carbonate

In addition to total dissolved solids, the relative abundance of sodium with respect to alkaline earths and boron and the quantity of bicarbonate and carbonate in excess of alkaline earths also influence the suitability of water for irrigation. This excess is denoted by residual sodium carbonate (RSC) and is determined by the following formula:

$$\mathrm{RSC} = (\mathrm{HCO_3}^- + \mathrm{CO_3}^-) - (\mathrm{Ca}^{2+} + \mathrm{Mg}^{2+}) \qquad (3)$$

where all ionic concentrations are expressed in equivalents per mole (epm). Groundwater containing high concentrations of carbonate and bicarbonate ions tends to precipitate calcium and magnesium as carbonate. As a result, the relative proportion of sodium increases and is fixed in the soil, thereby decreasing soil permeability. If the RSC exceeds 2.5 epm, the water is generally unsuitable for irrigation. Excessive RSC causes the soil structure to deteriorate because it restricts water and air movement through soil. If the value is between 1.25 and 2.5, the water is of marginal quality; values less than 1.25 epm indicate that the water is safe for irrigation. During the present study, the RSC values were mostly negative. The RSC values clearly indicate that the groundwater in the study area does not have a residual sodium carbonate hazard.

Table 4. Guidelines for Evaluating Irrigation Water Quality

Water Class	Sodium (Na), %	Electrical Conductivity, μS/cm	SAR	RSC
Excellent	<20	<250	<10	<1.25
Good	20–40	250–750	10–18	1.25–2.0
Medium	40–60	750–2250	18–26	2.0–2.5
Bad	60–80	2250–4000	>26	2.5–3.0
Very bad	>80	>4000	>26	>3.0

Boron

Boron is essential to the normal growth of all plants, but the concentration required is very small and if exceeded may cause injury. Plant species vary both in boron requirement and in tolerance to excess boron, so that concentrations necessary for the growth of plants having high boron requirements may be toxic for plants sensitive to boron. Though boron is an essential nutrient for plant growth, generally it becomes toxic beyond 2 ppm in irrigation water for most field crops. It does not affect the physical and chemical properties of the soil, but at high concentrations, it affects the metabolic activities of plants.

U.S. Salinity Laboratory Classification

The sodium concentration is an important criterion in irrigation water classification because sodium reacts with soil to create sodium hazards by replacing other cations. The extent of this replacement is estimated by the sodium adsorption ratio (SAR). A diagram for studying the suitability of groundwater for irrigation is based on the SAR and the electrical conductivity of water expressed in μS/cm.

The chemical analysis data of groundwater samples collected from the area adjoining the River Yamuna in Delhi has been processed per the U.S. Salinity Laboratory classification (Fig. 2), and the results have been summarized in Table 6. It is evident from the results that during the premonsoon season, only three samples (7.8%) fall into water type C2-S1 (medium salinity and low SAR). Such water can be used if a moderate amount of leaching occurs, and plants with moderate salt tolerance can be grown in most cases without special practices for salinity control. More than 50% of the

Table 5. Values of Sodium Adsorption Ratio (SAR), Percentage Sodium (%Na), and Residual Sodium Carbonate (RSC)

Sample No.	Sample Location	Type (Depth)	Premonsoon 2000			Postmonsoon 2000		
			SAR	%Na	RSC	SAR	%Na	RSC
1	Bhagwanpur Khera	HP (25)	4.93	65.1	0.24	5.35	69.8	1.52
2	Loni Road	BW (20)	3.81	56.4	0.27	4.10	58.4	−0.31
3	Kabul Nagar	OW (35)	2.99	55.6	−0.04	2.35	54.1	0.18
4	Naveen Shahdara	HP (20)	2.74	51.9	−0.30	2.39	49.2	−0.43
5	Seelampur	HP (30)	1.70	41.7	0.11	1.65	41.0	0.17
6	Shastri Park	HP (25)	3.20	67.9	1.44	2.87	66.9	1.38
7	Lakshmi Nagar	HP (100)	2.79	55.1	0.58	2.89	54.8	0.41
8	Prit Vihar	TW (40)	4.84	64.1	1.76	4.86	65.7	1.75
9	Shankar Vihar	HP (200)	5.49	71.2	1.31	5.93	72.7	1.92
10	Pratap Nagar	TW (100)	5.88	71.7	1.05	6.31	74.1	0.98
11	Himmatpuri	HP (80)	7.40	68.3	−2.75	7.00	68.6	−2.22
12	Civil Lines	TW (250)	12.06	83.7	1.30	11.24	82.9	1.24
13	Rajpur road	OW (25)	2.26	48.7	0.39	2.30	50.7	0.57
14	Malka Gunj	HP (20)	3.01	59.6	0.22	4.83	69.0	0.31
15	Tripolia	BW (100)	6.15	67.4	−1.74	6.05	67.1	−1.86
16	Gulabi Bagh	BW (40)	3.31	44.9	−6.02	3.99	48.9	−6.01
17	Gulabi Bagh	BW (160)	1.94	30.4	−6.71	2.33	34.5	−5.46
18	Shastri Nagar	BW (100)	3.58	64.0	0.40	3.37	61.0	−0.10
19	Shastri Nagar	BW (60)	2.48	47.4	−2.31	2.64	47.7	−1.79
20	Lekhu Nagar	HP (40)	2.99	59.0	0.82	2.64	50.3	0.00
21	Ram Pura	HP (40)	2.39	52.2	0.60	2.48	52.4	0.53
22	Punjabi Bagh West	HP (40)	4.31	59.0	−0.69	4.94	60.4	−1.46
23	Rajghat	BW (50)	4.42	61.2	−1.07	4.25	60.8	−1.22
24	JLN Marg	HP (60)	8.61	82.0	2.04	8.15	81.8	2.12
25	GB Pant Hospital	BW (20)	4.28	65.9	−0.08	4.11	64.3	−0.23
26	Panchkuin Marg	TW (450)	8.00	74.3	0.75	6.74	68.4	−0.91
27	Panchkuin Marg	HP (60)	4.15	62.7	0.08	3.81	58.9	−0.43
28	Rajendra Nagar	BW (100)	2.22	54.8	0.01	2.05	49.1	−0.59
29	Rajendra Nagar	BW (300)	2.13	49.4	−0.32	2.47	48.9	−1.16
30	Shankar Road	OW (20)	2.70	49.9	−1.66	4.32	60.5	−0.05
31	IARI, Pusa	TW (120)	5.44	61.1	−1.89	5.28	59.7	−1.63
32	Zoological Park	HP (60)	3.49	54.5	−1.34	3.25	57.1	−0.62
33	Golf Course	BW (80)	4.33	44.3	−12.9	4.32	45.2	−12.06
34	Rabindra Nagar	TW (150)	5.84	65.1	−1.54	5.89	65.0	−1.27
35	Teen Murti Chowk	HP (60)	7.23	74.1	0.52	6.98	72.5	0.82
36	Malcha Marg	BW (100)	9.23	75.3	0.34	9.16	76.0	0.58
37	Sardar Patel Road	TW (150)	10.71	79.6	1.59	10.99	81.3	1.75
38	Janpath	BW (250)	5.63	67.6	−0.39	5.94	70.2	−0.16

samples fall into water type C3-S1 (high salinity and low SAR). Such water cannot be used on soils with restricted drainage. Even with adequate drainage, special management for salinity control may be required, and plants with good salt tolerance should be selected. About 30% of the samples fall into water type C3-S2 (high salinity and medium SAR). Such water cannot be used on soils with restricted drainage. Even with adequate drainage, special management for salinity control may be required, and plants with good salt tolerance should be selected. The water will also present an appreciable sodium hazard in finely textured soils that have good cation exchange capacity, especially under low leaching conditions. Three samples (7.8%) fall into water type C3-S3 (high salinity and high SAR). Such water cannot be used on soils with restricted drainage. Even with adequate drainage, special management for salinity control may be required, and plants with good salt tolerance should be selected. The water will also produce harmful levels of exchangeable sodium in most soils and will require special soil management, good drainage, high leaching, and organic matter additions. One sample falls into water type C4-S2 (very high salinity and medium SAR). Such water is not suitable for irrigation. High salinity is harmful to plant growth and changes soil structure, permeability, and aeration, which in turn affect plant growth and yield considerably. An almost similar trend was observed during the postmonsoon season.

CONCLUSIONS

The quality of groundwater varies from place to place with the depth of the water table. It also shows significant variation from one season to another. The U.S. Salinity Laboratory classification for irrigation water indicates that more than 50% of the samples fall into water type C3-S1 (high salinity and low SAR). Such water cannot be used on soils with restricted drainage. Even with adequate drainage, special management for salinity control may be

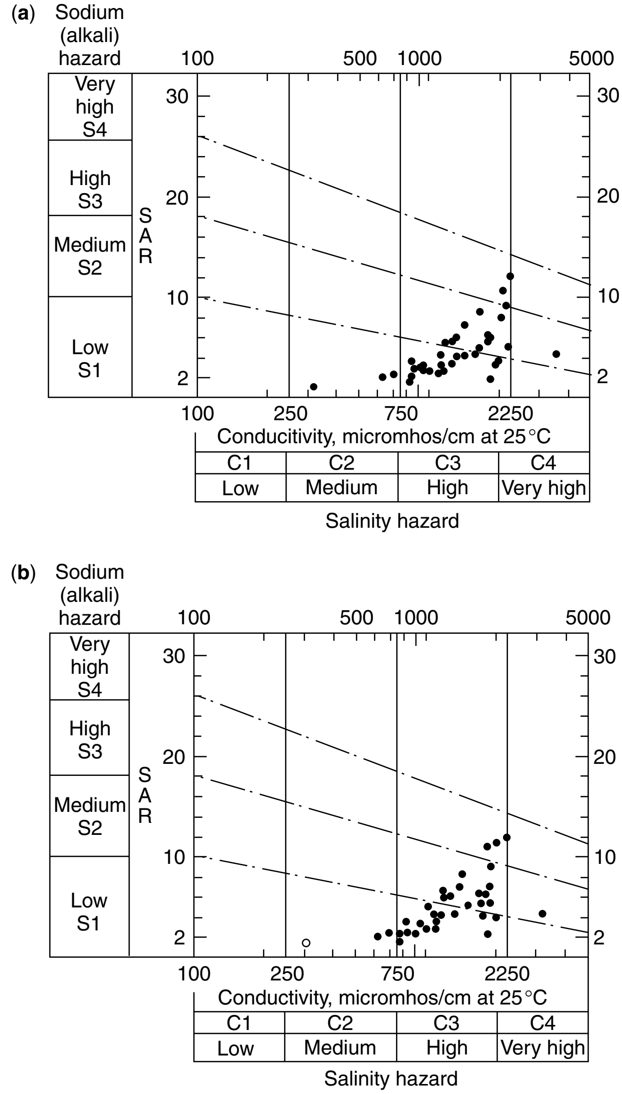

Figure 2. (**a**) U.S. Salinity Laboratory classification (premonsoon 2000). (**b**) U.S. Salinity Laboratory classification (postmonsoon 2000).

required, and plants with good salt tolerance should be selected. About 30% of the samples fall into water type C3-S2 (high salinity and medium SAR). Such water will also induce an appreciable sodium hazard in finely textured soils that have good cation exchange capacity, especially under low leaching conditions.

BIBLIOGRAPHY

1. CGWB and CPCB. (2000). *Status of Ground Water Quality and Pollution Aspects in NCT-Delhi*, January 2000.
2. APHA. (1992). *Standard Methods for the Examination of Water and Waste Waters*, 18th Edn. American Public Health Association, Washington, DC.
3. Jain, C.K. and Bhatia, K.K.S. (1988). *Physico-chemical Analysis of Water and Wastewater, User's Manual, UM-26*. National Institute of Hydrology, Roorkee.
4. Richards, L.A. (Ed.). (1954). *Diagnosis and Improvement of Saline and Alkali Soils, Agricultural Handbook 60*. U.S. Department of Agriculture, Washington, DC.

WATER SAMPLING AND LABORATORY SAFETY

Timothy J. Ryan
Ohio University
Athens, Ohio

Water sampling is done for a wide variety of reasons, including water quality, thermal effluent assessments, point and nonpoint industrial source impacts, microbial contamination from human wastes, and accidental releases. For example, samples can be taken from rivers, lakes, streams, and oceans to determine the level of a specific chemical component, such as nitrates or sulfates. Or tests may be done to investigate the biological oxygen demand (BOD) of the water, as an indicator of the overall health of the water body. Water is routinely sampled for the presence of harmful pathogens, which potentially pose a risk not only to the end user but to the sampling personnel as well. Because sampling may place the person in unsafe locations or even perilous environments, and because sample collection represents a significant allocation of time and resources, safeguards for personnel and samples alike are important.

The first decision in water sampling methodology is the categorical type of sample to be collected: general survey or site-specific. For general survey sampling, locations and analyses should be selected that are representative of the water source, treatment plant, storage facilities, distribution network, or points of use to be characterized. At a minimum, duplicate samples should be collected at each site. To accurately assess large or widely distributed bodies of water, consideration should be given to the use of a statistically sound sampling algorithm.

Table 6. Results of U.S. Salinity Laboratory Classification

	Sample Numbers	
Classification/Type	Premonsoon 2000	Postmonsoon 2000
C2-S1	21,28,29	3,21,28
C3-S1	2,3,4,5,6,7,13,14,15,16,17,18, 19,20,22,23,25, 27,30,32	2,4,5,6,7,8,13,14,16,17,18, 19,20,23,25,27,29,30,32
C3-S2	1,8,9,10,11,24,26,31,34, 35,38	1,9,10,11,15,22,24,26,31, 34,35,36,38
C3-S3	12,36,37	12,37
C4-S2	33	33

A site-specific sampling regimen is inherently simpler in that the sampling location and analysis parameters are known, easily anticipated, or even proscribed by a regulatory agency.

Actual procedures for sampling and analysis frequently depend on the type of analysis to be conducted. As a general truism, the following items are essential to conduct an effective sampling trip: labeled bottles appropriate for the analysis to be run, field equipment for commonly determined parameters (pH, water temperature, etc.), and sampling collection tools. Additional useful and typical equipment might include preservatives, ice packs and coolers, personal gear suitable for all weather conditions and intense insect exposure, a camera or video equipment (if needed to document the actual sampling operation), and laboratory chain-of-custody and analysis requisition forms. All sampling equipment and material should be checked before departing on the collection campaign, and real-time instruments, such as dissolved oxygen spectrophotometers should be calibrated, have their batteries fully charged, and be packed securely to avoid damage. Of all items taken on a sampling trip, the field log book is perhaps the most important. Although it will slow the sampling mechanics, all information pertinent to the sampling trip should be fastidiously recorded in it. Special bound books with water-resilient paper and bindings are available for such use, and they are well worth their added cost because of the increased information security they provide.

Immediately before sample collection, the sampling site should be adequately described in the log book. The use of global positioning satellite (GPS) transponders for precise location identification is always a sound practice and is in fact critical if repeated sampling in the same location on different dates is a possibility. Still-frame or video photographic documentation may be helpful in this regard in that readily identifiable landmarks can easily be recorded and communicated to third parties at a later date. The log book should also document the specific sampling collection information, such as time of day, prevailing wind direction, and weather conditions. Water-related aspects that may be important to the interpretation of specific sampling results include ambient temperature, water temperature, dissolved oxygen, pH, and salinity. Any unusual aspects about the site or water being sampled, such as oil slicks, excessive algae growth, or unusual color or odor, should be noted in the log, along with both typical or atypical human activities in the sampling vicinity. Such activities might include logging, commercial fishing with nets, dredging, industrial permitted point sources, construction work, aerial agricultural spraying, and utility company line clearing or defoliant operations.

The physical security and integrity of water samples is essential to ensuring the accuracy and legal defensibility of their analysis results. Toward this end, there are several things that one can to do assure and maintain sample quality in the field. First and foremost, one must take all necessary precautions to protect samples from contamination and deterioration. Maintaining the sterilized state of the bottle in which biological samples are to be collected is obviously important. Any specific sampling instructions from the analysis laboratory for handling sample containers must be closely followed. After samples are collected, they are normally kept refrigerated and care must be exercised that this does not result in the sample actually freezing. Unless specifically instructed otherwise, all samples should be place in a dark place away from other contaminant sources while awaiting transport or analysis. The person conducting the sampling should have clean hands and abstain from eating, gum chewing, or smoking while working with sampling media or other equipment that could be cross contaminated by such habits. To prevent sample or site confusion, bottles for particular sample types (e.g., volatile organics, chloride, total organic halides) can be color coded for more confident identification.

Sample collection can take place from a variety of locations, including from shore, a boat, access from a bridge, adjacent to the effluent stream, and from receiving waters. Different procedures and safety precautions should be taken with each situation. In general, it is best to take samples from mid-stream rather than near the shore to reduce the possibility of local shore effects contaminating the sample. It bears reiteration to note in the sampling log book the exact spot of sampling, so that if further sampling is needed (e.g., in response to a contested sample result), the identical location can be found for further rounds of testing. Water samples can also be taken in the home or business, such as from a water faucet, well head, or septic system. During these samplings, the same precautions need to be taken concerning the physical security and integrity of the sample. As for outdoor samples, notes should be kept in a log book about pertinent aspects of the water such as location in the building where the sample was collected, temperature of the water, time of day, and overall appearance of the water (e.g., turbid, milky, clear).

After field sample collection, precautions must also be taken by the individual handling water samples in the laboratory. As in the field, normal prohibitions against eating and smoking should be followed, both for the safety of personnel as well as for the integrity of the sample because contamination can result from such sources. Ensuring a hygienic laboratory environment when testing biological samples is a high priority, primarily as a safeguard against cross contamination but also to safeguard the health of the laboratory staff. Because water samples are most often collected in specially cleaned glass bottles (e.g., a standard 300 milliliter BOD bottle), containers should routinely be transported in shock-absorbent carriers to protect them from breakage, resultant sample loss, and possible injury. Note that even when using plastic bottles, precautions should be taken because such containers can break, leak, be punctured, or fail under pressure or heat.

In approximately 25 states and territories, laboratories handling water samples may be subject to federal Occupational Safety and Health Administration (OSHA) standards; laboratories in other states must comply with similar regulations promulgated under state authority. Regardless of their origin, laboratory safety regulations and hazard communication rules are to be followed by persons handling, processing, and otherwise analyzing water samples. Such work can involve exposure to highly corrosive agents (e.g., potassium chromate for chloride analyses), highly reactive catalysts (e.g., metallic sodium for total chlorine), or moderately toxic reagents (e.g., mercuric nitrate for total chlorine). Accordingly, basic minimal protective clothing should be mandatory for handling water samples. This personal protective equipment should include a laboratory coat or chemically resistant apron, gloves impermeable to the reagents handled, and an appropriately selected and worn respirator (where airborne hazards exist, and fumehoods are not available). Perhaps of most importance, protection of the eyes from chemical splashes is required, preferably in the form of chemical goggles or as safety glasses in conjunction with a face shield.

READING LIST

Title 29, Labor. *Code of Federal Regulations, Part 1910*. Standard 1910.1450, Occupational exposure to hazardous chemicals in laboratories.

Test Methods for Evaluating Solid Waste, Physical/Chemical Methods. EPA publication SW-846. Available: http://www.epa.gov/epaoswer/hazwaste/test/sw846.htm (accessed September 22, 2004).

WHO—Water Sampling and Analysis. Available: http://www.who.int/water_sanitation_health/dwq/en/2edvol3d.pdf (accessed September 11, 2004).

MUNICIPAL SOLID WASTE LANDFILLS—WATER QUALITY ISSUES

G. Fred Lee
Anne Jones-Lee
G. Fred Lee & Associates
El Macero, California

INTRODUCTION

Here we provide information on significant shortcomings inherent in the management of municipal solid waste (MSW) and industrial "nonhazardous" waste by the Subtitle D "dry tomb" landfilling practiced today. Focus is on why following the prescriptive measures outlined for such landfills cannot be relied on to provide protection of public health, groundwater resources, or the interests of nearby property owners.

Historically, the goal of municipal solid wastes and industrial "nonhazardous" waste management has been to get the wastes out of sight in the least costly manner. Solid wastes from urban areas were deposited on nearby low-value lands, frequently wetlands, to create waste dumps. This approach evolved into the deposition of wastes into areas excavated for greater disposal capacity. Wastes in the dumps were often burned to reduce volume. Beginning in some areas in the 1950s, wastes placed in dumps were covered daily with a layer of soil to reduce odors and access to the wastes by vermin, flies, birds, and so on, which was the beginning of the "sanitary landfill." Although the design of sanitary landfills diminished some environmental and public health concerns, it did not address issues of the potential for the wastes to cause groundwater pollution, for the gas generated in the landfill to cause explosions, or for the waste to cause other public health or environmental problems. The pollution of groundwater by landfill leachate from conventional sanitary landfills was recognized in the 1950s (2), but it was not until the 1990s that national regulations existed directed at control of groundwater pollution by landfills.

U.S. EPA SUBTITLE D DRY TOMB LANDFILLING

Recognizing the pollution of groundwaters by sanitary landfills, Congress passed the Resource Conservation and Recovery Act (RCRA) that directed the U.S. EPA to address the issue. As this was not accomplished in a timely manner, the U.S. EPA was sued; in response to that suit, the U.S. EPA developed the dry tomb approach to landfilling of MSW. Despite its acknowledgment of technical shortcomings in the approach, but under the pressure of the lawsuit, the U.S. EPA delineated the prescriptive standards for dry tomb landfills in Subtitle D of RCRA. These prescriptive standards have been adopted by some state regulatory agencies as minimum design standards for MSW landfills.

The concept of dry tomb landfilling was built on the premise that because water in contact with the wastes led to the formation of leachate that traveled to groundwater, isolation of the wastes in dry tombs would prevent groundwater pollution by landfill leachate. A dry tomb landfill, as implemented by the U.S. EPA, is illustrated in Fig. 1. It relies on a liner and cap to keep the wastes dry, leachate collection and removal systems to keep the wastes from polluting groundwater, and gas collection and removal systems. It also relies on groundwater monitoring to signal incipient pollution of groundwater by landfill leachate and specifies a "postclosure" period of 30 years, during which time the facility is to be maintained and monitored.

These provisions notwithstanding, the U.S. EPA realized in its development of these requirements that such a system would not, in itself, ensure the protection of groundwaters forever. As part of adopting the RCRA Subtitle D regulations, the U.S. EPA stated in the draft regulations (3),

> First, even the best liner and leachate collection system will ultimately fail due to natural deterioration, and recent improvements in MSWLF (municipal solid waste landfill) containment technologies suggest that releases may be delayed by many decades at some landfills.

Further, the U.S. EPA (4) Criteria for Municipal Solid Waste Landfills states,

> Once the unit is closed, the bottom layer of the landfill will deteriorate over time and, consequently, will not prevent leachate transport out of the unit.

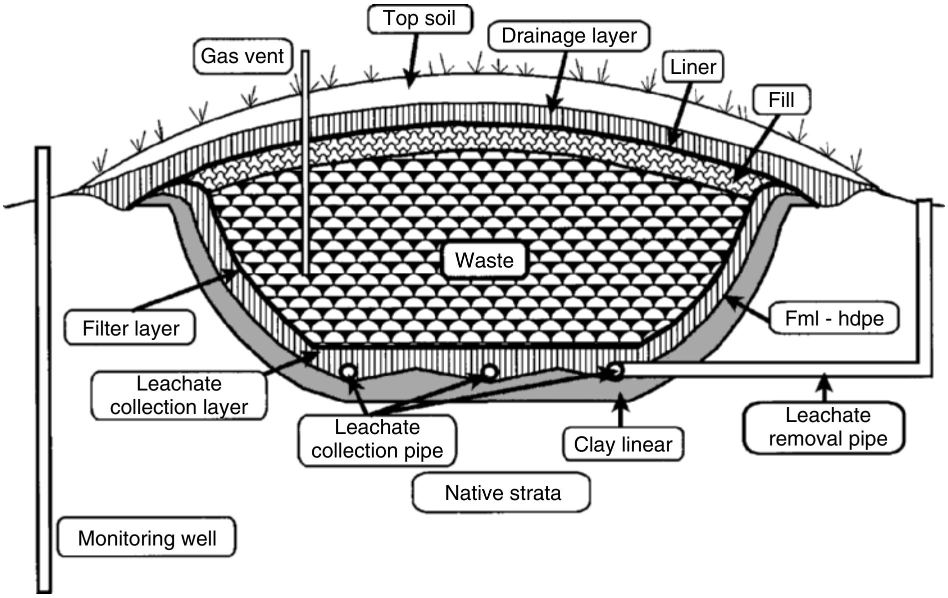

Figure 1. Single-composite liner landfill containment system.

Thus, even if the dry tomb system could be installed without flaw, it will eventually fail to isolate the wastes, which is significant because many of the contaminants in MSW landfills do not decompose to inert forms (e.g., heavy metals) and are hence a threat forever. Furthermore, although many organics can eventually be bacterially decomposed, such decomposition requires water, which is excluded from the landfill under ideal dry tomb landfill conditions. It is therefore important to consider the characteristics and shortcomings of the components of a dry tomb landfill to provide protection of groundwater quality for as long as the wastes in the landfill remain a threat, and how they can be improved to provide more reliable protection.

LANDFILL COVER

Allowed closure of Subtitle D landfills is begun by covering the waste with soil that is placed and shaped to serve as the base for a low-permeability plastic sheeting layer. The plastic sheeting is overlaid by a one- to two-foot deep drainage layer. The drainage layer is overlaid by a few inches to a foot or so of topsoil that serves as a vegetative layer designed to promote the growth of vegetation to reduce erosion of the landfill cover. In principle, water that enters the vegetative layer would either be taken up by the vegetation or penetrate through the root zone to the porous (drainage) layer. When the moisture reaches the low-permeability plastic sheeting layer of the cover, it is supposed to move laterally to the outside of the landfill (see Fig. 1).

However, in addition to deficiencies introduced during the installation of the plastic sheeting, landfill covers are subject to a variety of factors that can breach their integrity, including burrowing animals, differential settling of the wastes beneath the cover, and deterioration of the plastic sheeting. The typical approach to monitoring landfill covers for integrity in keeping moisture out of the landfill, that is advocated by landfill owners and operators and allowed by regulatory agencies, involves visual inspection of the surface of the vegetative soil layer of the landfill cover. If cracks or depressions in the surface are found, they are filled with soil. Such an approach, however, will not detect or remediate cracks in or deterioration of the plastic sheeting layer, which is the basis for the moisture removal system for the cover. As a result, moisture that enters the drainage layer and comes in contact with the plastic sheeting layer will penetrate into the wastes rather than be directed off the landfill. If this occurs during the postclosure care period, the increased leachate generation could be detected. However, it could also readily occur in year 31 after closure or thereafter, when there could well be no one monitoring leachate generation.

LEACHATE COLLECTION AND REMOVAL SYSTEM

The key to preventing groundwater pollution by a dry tomb landfill is collecting all leachate that is generated in the landfill, in the leachate collection and removal system. As shown in Fig. 1, a Subtitle D system prescribes a relatively thin (0.06 inch) plastic sheeting layer (high-density polyethylene—HDPE) and a 2-ft thick compacted soil/clay layer, which together, in intimate contact, form a “composite” liner beneath the wastes. Atop the liner is the leachate collection system consisting of gravel or some other porous medium, which is intended to allow leachate to flow rapidly to the top of the HDPE liner for removal. This porous layer is overlaid by a filter layer, which is supposed to keep the solid waste from migrating into the leachate collection system. Thus, in principle, leachate that is generated in the solid waste passes through the filter layer to the leachate collection layer beneath. Once it reaches the sloped HDPE liner, it is supposed to flow across the top of the liner to a collection pipe, where it would be transported to a sump from which it could be pumped from the landfill. According to regulations, the maximum elevation of leachate ("head") in the sump is to be no more than 1 ft.

However, leachate collection and removal systems, as currently designed, are subject to many problems. Biological growth, chemical precipitates, and "fines" derived from the wastes all tend to cause the leachate collection system to plug. This, in turn, increases the head of the leachate above the liner upstream of the area that is blocked, which further stresses the integrity of the system. Although the potential exists to back-flush some of these systems, this back-flushing will not eliminate the problem.

One of the most significant problems with leachate collection systems' functioning as designed is that the HDPE liner, which is the base of the leachate collection system, develops cracks, holes, rips, tears, punctures, and points of deterioration. Some of these are caused at the time of installation, and HDPE integrity deteriorates over time. When the leachate that is passing over the liner reaches one of these points, it starts to pass through the liner into the underlying clay layer. If the clay layer is in intimate contact with the HDPE liner, the rate of initial leakage through the clay is small. If, however, problems existed achieving or maintaining intimate contact between the clay and HDPE liner, such as the development of a fold in the liner, the leakage through the HDPE liner hole can be quite rapid. Under those conditions, the leachate can spread out over the clay layer and can leak at a substantial rate through the clay. As noted above, even in the establishment of the Subtitle D regulations, the U.S. EPA recognized that such a liner will deteriorate over time.

GROUNDWATER MONITORING

U.S. EPA Office of Solid Waste Emergency Response senior staff members have indicated that the fact that HDPE liners will fail to prevent leachate from passing into the underlying groundwaters does not mean that the Subtitle D regulations are fundamentally flawed because of the groundwater monitoring system requirement. They claim that when leachate-polluted groundwaters first reach the point of compliance for groundwater monitoring, they are detected by the groundwater monitoring system with sufficient reliability so that a remediation program can be initiated. The point of compliance for groundwater monitoring at Subtitle D landfills is specified as being no more than 150 meters from the downgradient edge of the waste deposition area and must be on the landfill owner's property. However, serious technical deficiencies exist in that position.

It was pointed out by Cherry (5) that initial leakage through HDPE-lined landfills will be through areas where there are holes, rips, tears, or points of deterioration of the HDPE liner, which, as illustrated in Fig. 2, can lead to relatively narrow plumes of polluted groundwaters at the point of compliance for groundwater monitoring. The typical groundwater pollution plume in a sand, gravel, or silt aquifer system will likely be on the order of 10 to 20 ft wide at the point of compliance. In order to detect incipient leakage from a Subtitle D landfill, therefore, these narrow plumes would have to be detected by the groundwater monitoring well array at the point of compliance. Typically, federal and state regulatory agencies allow monitoring wells to be placed 100 or more feet apart at the point of compliance; each monitoring well has a zone of capture of about 1 ft. Thus, if the wells were 200 ft apart, 198 ft would exist between wells where a 10- to 20-ft wide plume of leachate-polluted groundwater could pass undetected. It is virtually impossible to reliably monitor groundwater for leachate contamination in fractured rock or cavernous limestone areas. There, leachate can travel great distances and in unexpected directions via cracks and caverns.

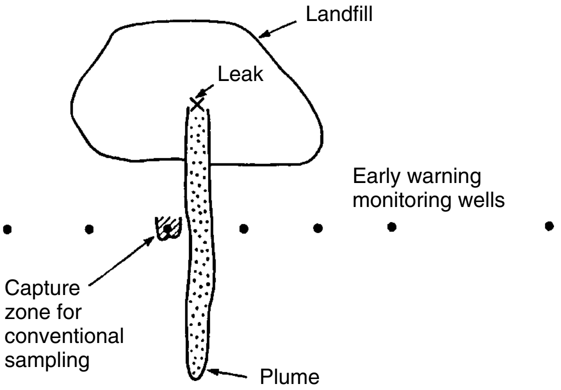

Figure 2. Leakage from HDPE-lined landfill (adapted from Ref. 5).

It is because of the unreliability of groundwater monitoring systems that are based on vertical monitoring wells at the point of compliance that some states (such as Michigan) require that a double-composite liner be used at municipal solid waste landfills. In that system, a leak detection system is in place between the composite liners to detect leakage through the upper composite before the lower liner is breached. Although this approach is not foolproof in its ability to detect when both liner systems fail, it has a much greater probability of detecting when the upper composite liner fails. It is for this reason that Lee and Jones-Lee (6) recommended that all Subtitle D landfills incorporate a double-composite liner with a leak detection system between the two liners.

LANDFILL GAS

Some organics in municipal solid wastes can serve as a source of food for bacteria that will, in a moist landfill environment, produce methane and CO_2 (landfill gas). MSW landfills also release a number of other volatile chemicals, including highly hazardous volatile organic compounds (VOCs) and odorous compounds, which are a threat to the health and welfare of those within the sphere of influence of the landfill. This sphere can extend for several miles, depending on the topography of the area and the tendency for atmospheric inversions to occur.

Although landfill advocates claim that the practice of daily cover of the wastes will reduce the gaseous (odorous) releases from landfills, even diligent covering does not prevent them. Further, when landfill owners/operators become sloppy in operations, greater-than-normal landfill gas emissions occur, which are typically detected through landfill odors. Although some attempt to minimize the significance of smelling landfill gas on adjacent properties as only being an aesthetic problem, in fact, as discussed by Shusterman (7), it is now known that noxious odors can cause illness in people. Odors should be controlled so that they do not trespass across the landfill adjacent property owner's property line.

The only way at present to reliably ensure the protection of adjacent and nearby property owners/users from landfill gas is to provide a sufficient landfill-owner-owned buffer land about the landfill. If an adjacent or nearby property owner/user can smell the landfill, then inadequate buffer land exists between the landfill and adjacent properties, which should make it necessary for the landfill owner/operator to either acquire adjacent buffer land or to use more than the minimum approach for controlling gaseous releases from the landfill. It is also important to control future land use within the landfill area so that releases from the landfill would not be adverse to the land use. For example, agriculture in these areas should be restricted, because releases from the landfill could contaminate some crops. As long as landfill owners are allowed to use adjacent properties for their waste disposal buffer zones, justified NIMBY ("not in my back yard") issues by adjacent property owners will exist.

A common misconception held by landfill applicants and some regulators is that landfill gas production will cease after a comparatively short time after closure of a dry tomb landfill. Estimates of landfill gas production rate and duration are typically based on a model assuming unbagged, homogeneous wastes that are allowed to interact with moisture, which is not the situation that is found in Subtitle D dry tomb MSW landfills. The key to landfill gas production is the presence of sufficient moisture to allow bacteria to metabolize certain of the organic materials in the landfill to methane and CO_2 (landfill gas). Thus, the rate of moisture penetration through the cover and the mixing of that moisture with the waste components control the rate and duration of landfill gas production. Once the landfill is closed, Subtitle D landfills aim to keep moisture out of contact with the buried waste; the dryer the waste is kept, the slower the rate of landfill gas production. Furthermore, because much of the municipal waste that is placed in Subtitle D landfills is contained within plastic bags, and because those plastic bags are only crushed and not shredded, the crushed bags will "hide" fermentable components of the waste. Thus, once moisture does breach the cover and is allowed to contact with the wastes, the period of landfill gas production will be prolonged further, until after the plastic bags decompose and the bagged wastes allowed to interact with the moisture, which can extend gas production many decades, to a hundred or more years.

POSTCLOSURE MONITORING AND MAINTENANCE

The Subtitle D relies on the 30-year postclosure period of monitoring and maintenance to continue the protection of groundwater. The mistaken idea existed that 30 years after closure, the waste in a dry tomb landfill would no longer be a threat. From the characteristics of wastes and their ability to form leachate, as well as the processes than can occur in a landfill, it is clear that 30 years is an infinitesimally small part of the time that waste components in a landfill, especially a dry tomb landfill, would be a threat to cause groundwater pollution through leachate formation. A critical review of the processes that can take place in a MSW landfill that can generate leachate shows that the containment system of a dry tomb landfill, for which there is at least an initial effort to reduce the moisture entering the wastes, will eventually fail to prevent groundwater pollution for as long as the wastes are a threat. The municipal solid wastes in a classic sanitary landfill where no attempt to prevent moisture from entering the wastes exists have been found to generate leachate for thousands of years. Dry tomb landfilling delays and then prolongs leachate generation.

Subtitle D regulations also require that a small amount of assured funding be available for 30 years of postclosure monitoring and maintenance. Some regulatory agencies will allow landfill companies to be self-insured or insured through an insurance company that is backed by a landfill company. Such approaches should not be allowed, because landfill companies are amassing large liabilities for the ultimate failure of the landfill liner systems and the groundwater pollution that will occur as a result of those failures. It is well understood that, ultimately, private landfill companies will not likely be able to comply with Subtitle D regulations for funding of remediation. The amount of postclosure monitoring and maintenance funding that is currently required is grossly inadequate compared with the funding levels that could be necessary during the 30-year mandatory postclosure period, much less the extended period over which the wastes in the landfill will be a threat. Although Congress, through Subtitle D, required that the regulations include provisions to potentially require additional funding at the expiration of the 30-year postclosure care period, the likelihood of obtaining this funding from private landfill companies, even if they still exist 30 years after a landfill has been closed, or from a public agency that developed or owns landfills, is remote.

Lee (8) discussed the importance of solid waste management regulatory agencies requiring that landfill owners, whether public or private, prepare for the inevitable failure of the landfill containment system and provide for funding to address this failure. Designation of a 30-year postclosure assured funding period in RCRA, which is implemented not as a minimum but rather as a definitive period, leaves the public to pay the significant balance of the cost of landfilling; the public is left to deal with the long-term impacts of MSW landfills on public health and the environment, and eventual remediation of the landfill. This significant deficiency in Subtitle D regulations is recognized not only in the technical community but also by various other groups and individuals who have reviewed this issue. For example, in the executive summary of its report *Funding of Postclosure Liabilities Remains Uncertain,* in a section labeled, “Funding Mechanisms Questionable,” the U.S. Congress General Accounting Office (GAO) (9) concluded,

> Owners/operators are liable for any postclosure costs that may occur. However, few funding assurances exist for postclosure liabilities. EPA only requires funding assurances for maintenance and monitoring costs for 30 years after closure and corrective action costs once a problem is identified. No financial assurances exist for potential but unknown corrective actions, off-site damages, or other liabilities that may occur after the established postclosure period.

Further, the U.S. EPA Inspector General (10), in a report entitled “RCRA Financial Assurance for Closure and Postclosure,” came to similar conclusions:

> There is insufficient assurance that funds will be available in all cases to cover the full period of landfill postclosure monitoring and maintenance. Regulations require postclosure activities and financial assurance for 30 years after landfill closure, and a state agency may require additional years of care if needed. We were told by several state officials that many landfills may need more than 30 years of postclosure care. However, most of the state agencies in our sample had not developed a policy and process to determine whether postclosure care should be extended beyond 30 years, and there is no EPA guidance on determining the appropriate length of postclosure care. Some facilities have submitted cost estimates that were too low, and state officials have expressed concerns that the cost estimates are difficult to review.

LANDFILL SITING ISSUES

In the development of Subtitle D landfill regulations, the U.S. EPA failed to address one of the most important issues that should be addressed in developing a minimum Subtitle D landfill, namely, the need to site landfills at geologically suitable sites for a landfill of this type. Although the EPA does require that minimum Subtitle D landfills not be sited too close to airports, where major bird problems could exist for aircraft, too near an earthquake fault, or within a flood plain, the EPA does not address the issue of siting minimum Subtitle D landfills where the underlying geological strata do not provide natural protection of the groundwaters from pollution by landfill leachate when the landfill liner systems eventually fail. In accordance with current regulations, minimum Subtitle D landfills can be sited over highly important aquifers that serve as domestic water supply sources. They can also be sited in fractured rock and cavernous limestone areas, where it is impossible to reliably monitor for the pollution of groundwaters by landfill leachate using vertical monitoring wells.

Subtitle D landfill regulations also fail to address one of the most important causes for people to object to landfills; the deposition of wastes is allowed very near the landfill property owner’s property line. Under these conditions, the landfill gases, blowing paper, birds, rodents, vermin, and so on associated with the landfill can impinge on adjacent properties. For example, it is well established that landfill gas can readily travel a mile or more from a landfill, and thereby be adverse to the adjacent property owners’ use of their properties. It is recommended that at least a mile, and preferably two miles, of landfill-owned buffer lands exist between the area where wastes are deposited and adjacent property owners’ property lines, which would provide for dissipation of releases from the landfill on the landfill owner’s property. Such an approach would greatly reduce the trespass of waste-derived materials from the landfill onto adjacent properties.

ROLE OF 3 Rs IN MSW LANDFILLING

Considerable efforts are being made in many parts of the country to increase reuse, recycling, and reduction (the “3Rs”) of MSW as part of conserving natural resources and landfill space. As discussed by Lee and Jones-Lee (11), the 3 Rs should be applied to the maximum extent practicable to reduce the number of new and expanded Subtitle D landfills that will eventually pollute groundwaters.

JUSTIFIED NIMBY

NIMBY is an acronym for “not in my backyard,” a commonly dismissed plea presumed to be made by those who, without justification, simply oppose having a landfill nearby. However, Lee and Jones-Lee (12) and Lee et al. (13) discussed many of the technical issues that, in fact, justify a NIMBY position for Subtitle D landfills. Table 1 presents typical, real, adverse impacts of landfills on nearby property owners/users.

IMPROVING LANDFILLING

The degree of protection of public health and the environment from adverse impacts of MSW disposed of in dry tomb landfills can be improved to address some of the deficiencies of the approaches and specifications currently commonly accepted.

Siting

Landfills should be sited so that they provide, to the maximum extent possible, natural protection of groundwaters when the liner systems fail. Siting landfills above geological strata that have groundwater whose flow

Table 1. Adverse Impacts of Dry Tomb Landfills on Adjacent/Nearby Property Owners/Users[a]

- Public health, economic and aesthetic impacts on groundwater and surface water quality
- Methane and VOC migration—public health hazards, explosions, and toxicity to plants
- Illegal roadside dumping and litter near landfill
- Truck traffic
- Noise
- Dust and wind-blown litter
- Odors
- Vectors, insects, rodents, birds
- Condemnation of adjacent property for future land uses
- Decrease in property values
- Impaired view

Source: Lee et al. (13).

paths are not readily amenable to monitoring for leachate-polluted groundwaters should be avoided. Of particular concern are fractured rock and cavernous limestone areas, as well as areas with sandy lenses.

Design

MSW landfills should incorporate double-composite liners with a leak detection system between the two liners.

Closure

MSW landfills should incorporate leak-detectable covers that will indicate when the low-permeability layer of the landfill cover first fails to prevent moisture from entering the landfill.

Monitoring

The primary monitoring of liner leakage should be associated with the leak detection system between the two composite liners. If vertical monitoring wells are used, the spacing between the vertical monitoring wells at the point of compliance should be such that a leak in the HDPE liner caused by a 2-ft-wide rip, tear, or point of deterioration at any location in the landfill would be detected based on the plume that is generated at the point of compliance with a 95% reliability.

Landfill Gas Collection

For those landfills that contain wastes that can produce landfill gas, a landfill gas collection system should be designed, installed, and maintained for as long as the wastes in the dry tomb landfill have the potential to generate landfill gas, giving proper consideration to how and for how long gas will be generated in the system. The landfill gas collection system should be designed to have at least a 95% probability of collecting all landfill gas generated at the landfill.

Maintenance

The maintenance of the landfill cover, monitoring system, gas collection system, and so on. should be conducted for as long as the wastes in the landfill will be a threat, with a high degree of certainty of detecting landfill containment system and monitoring system failure, which will extend well beyond the 30-year period typically established, especially with improved provisions for excluding moisture.

Funding

The funding for closure, postclosure monitoring, maintenance, and groundwater remediation should be established at the time the landfill is established, in a dedicated trust fund of sufficient magnitude to address plausible worst-case scenario failures for as long as the wastes in the landfill will be a threat. Unless appropriately demonstrated otherwise, it should be assumed that the period of time for which postclosure care funding will be needed will be infinite.

Buffer Land

At least several miles of landfill-owner-owned buffer lands should exist between where wastes are deposited and adjacent properties.

Adoption of these approaches (or as many of them as possible) will improve the protection of groundwater quality, public health, and environmental quality from adverse impacts of the dry tomb landfilling of MSW.

Further information on the topics discussed is provided by Jones-Lee and Lee (14), Lee and Jones-Lee (6,12,15,16,18), and Lee (17).

BIBLIOGRAPHY

1. U.S. EPA. (1991). Solid waste disposal facility criteria; final rule, part II. *Federal Register*. 40 CFR Parts 257 and 258, U.S. Environmental Protection Agency, Washington, DC.
2. ASCE. (1959). *Sanitary Landfill*. Report Committee on Sanitary Landfill Practice of the Sanitary Engineering Division of the American Society of Civil Engineers, New York.
3. U.S. EPA. (1988). Solid waste disposal facility criteria: proposed rule. *Federal Register* **53**(168): 33,314–33,422–. 40 CFR Parts 257 and 258, U.S. Environmental Protection Agency, Washington, DC.
4. U.S. EPA. (1988). *Criteria for Municipal Solid Waste Landfills*. U.S. Environmental Protection Agency, Washington, DC.
5. Cherry, J.A. (1990). Groundwater monitoring: some deficiencies and opportunities. In: *Hazardous Waste Site Investigations: Towards Better Decisions* B.A. Berven and R.B. Gammage (Eds.). Proceedings of the 10th ORNL Life Sciences Symposium, Gatlinburg, TN, Available from gfredlee@aol.com as "LF019."
6. Lee, G.F. and Jones-Lee, A. (1998). *Assessing the Potential of Minimum Subtitle D Lined Landfills to Pollute: Alternative Landfilling Approaches*. Proc. Air and Waste Management Assoc. 91st Annual Meeting, San Diego, CA. Available on CD ROM as paper 98-WA71.04(A46), p. 40 and http://www.gfredlee.com/alternative_lf.html.
7. Shusterman, D. (1992). Critical review: the health significance of environmental odor pollution. *Arch. Environ. Health.* **47**(1): 76–87.
8. Lee, G.F. (2003). *Workshop on Landfill Postclosure and Financial Assurance*. Comments submitted to Mike Paparian, California Integrated Waste Management Board, by G. Fred Lee & Associates, El Macero, CA. Available at http://www.gfredlee.com/paparian10-30-03T.pdf.
9. GAO. (1990). *Funding of Postclosure Liabilities Remains Uncertain*. GAO/RCED-90-64, U.S. General Accounting Office Report to the Congress, Washington, DC.
10. U.S. EPA. (2001). *RCRA Financial Assurance for Closure and Postclosure*. Audit Report, 2001-P-007, U.S. Environmental Protection Agency, Office of Inspector General, Washington, DC.
11. Lee, G.F. and Jones-Lee, A. (2000). *Three Rs Managed Garbage Protects Groundwater Quality*. Proc. AWMA 93rd annual conference, Salt Lake City, UT, paper 00-454 CD ROM, Air & Waste Management Association, Pittsburgh, PA. Available at http://www.gfredlee.com/3rpap_sli.pdf.
12. Lee, G.F. and Jones-Lee, A. (1994). Addressing Justifiable NIMBY: A Prescription for MSW Management.

Environmental Management Review, Government Institutes, Rockville, MD, No. 31, pp. 115–138. Available at http://www.gfredlee.com/funding.htm.

13. Lee, G.F., Jones-Lee, A., and Martin, F. (1994). *Landfill NIMBY and Systems Engineering: A Paradigm for Urban Planning*. In: Proc. National Council on Systems Engineering Fourth Annual International Symposium, Vol. 1, pp. 991–998. Available at http://www.members.aol.com/duklee2307/NIMBY-NCO2.pdf.
14. Jones-Lee, A. and Lee, G.F. (1993). Groundwater pollution by municipal landfills: leachate composition, detection and water quality significance. *Proc. Sardinia '93 IV International Landfill Symposium*. Sardinia, Italy, pp. 1093–1103. Available at http://www.gfredlee.com/lf-conta.htm.
15. Lee, G.F. and Jones-Lee, A. (1995). *Overview of Landfill Post Closure Issues*. Presented at American Society of Civil Engineers Convention session devoted to "Landfill Closures - Environmental Protection and Land Recovery." San Diego, CA. Available at http://www.gfredlee.com/plandfil2.htm#postclosure.
16. Lee, G.F. and Jones-Lee, A. (1998). *Deficiencies in Subtitle D Landfill Liner Failure and Groundwater Pollution Monitoring*. Presented at the NWQMC National Conference "Monitoring: Critical Foundations to Protect Our Waters. U.S. Environmental Protection Agency, Washington, DC. Available at http://www.gfredlee.com/nwqmcl.html.
17. Lee, G.F. (2002). Solid Waste Management: USA Lined Landfilling Reliability. An invited submission for publication in *Natural Resources Forum*, a United Nations Journal, New York. Available at http://www.gfredlee.com/UNpaper-landfills.pdf.
18. Lee, G.F. and Jones-Lee, A. (2004). *Flawed Technology of Subtitle D Landfilling of Municipal Solid Waste*. Report of G. Fred Lee & Associates, El Macero, CA. Available upon request from gfredlee@aol.com.

LAND USE EFFECTS ON WATER QUALITY

Ross A. Steenson
Geomatrix
Oakland, California

Factors that determine water quality at a site include climate, geology, biology, and land use (anthropogenic practices). Water quality varies naturally as the result of differing natural environments, but anthropogenic practices can override these natural variations. During the last several decades, research and regulatory efforts in the United States have largely focused on understanding and limiting water quality effects from specific or localized anthropogenic sources, which are termed point sources.

Point sources include farms; industrial facilities where chemicals are manufactured, used, or stored; landfills; sewage treatment plants; hazardous waste sites; mines; and construction sites.

There are numerous routes through which sources can affect water quality (e.g., downward migration of source chemicals, wet weather runoff, and direct discharge). The water quality effects observed from sources include increased concentrations of dissolved chemicals (e.g., fertilizers/nutrients, pesticides/herbicides, industrial chemicals, trace elements, and minerals), loss of dissolved gases (e.g., oxygen, carbon dioxide), and change in physical characteristics (e.g., sediment load, temperature). Waters potentially susceptible to these effects include nearby surface waterbodies (e.g., streams and rivers) and underlying shallow groundwater or groundwater that is connected hydraulically to affected surface waterbodies.

More recently, studies and regulatory efforts have shifted beyond water quality effects at specific sites to understanding the broader effects. These efforts at understanding the broader water quality effects of land use have included studies of larger geographic regions and studies of the effects of increased human populations. Regional studies have combined available water quality data with measures of local land use and/or census information and identified general correlations between areal water quality, land use, and population density (1). As a result of the broadened view of water quality effects, the concept of nonpoint sources has evolved. Nonpoint sources can include multiple point sources or more areal sources (or practices).

Nonpoint sources include agricultural activities and agricultural and urban runoff. These diffuse sources have a greater potential to affect water quality and are much harder to address than point sources. An example of a common contaminant that is largely the result of nonpoint sources is nitrate, which is considered the most widespread contaminant in groundwater, particularly in agricultural areas (2). Runoff from urban areas is a potential source of many chemicals (e.g., spilled oils), debris, and sediment that can be rapidly discharged to surface waterbodies following wet weather. Thus, urbanization of watersheds has been viewed increasingly as ever more significantly affecting water quality (3).

In the last decade, concern has increased regarding the indirect effects on water quality from human populations and the pharmaceuticals and personal care products used by individuals. The number of these products is rapidly increasing, and it is believed that many of these chemicals find their way primarily into surface waterbodies, but also into groundwater, through incomplete treatment at sewage treatment plants or disposal in landfills (4).

BIBLIOGRAPHY

1. Eckhardt, D.A.V. and Stackelberg, P.E. (1995). Relation of ground-water quality to land use on Long Island, New York. *Ground Water* **33**(6, Nov.–Dec.): 1019–1033.
2. Hallberg, G.R. and Keeney, D.R. (1993). Nitrate. In: *Regional Ground-Water Quality*, W.M. Alley (Ed.). Van Nostrand Reinhold, New York, pp. 297–322.
3. Horner, R.R., Skupien, J.J., Livingston, E.H., and Shaver, H.E. (1994). *Fundamentals of Urban Runoff Management: Technical and Institutional Issues*. Terrene Institute, Washington, DC, in cooperation with the U.S. Environmental Protection Agency.
4. Daughton, C.G. and Ternes, T.A. (1999). Pharmaceuticals and personal care products in the environment: agents of subtle change? *Environ. Health Perspect.* **107**(Suppl. 6, Dec.).

MONITORING LIPOPHILIC CONTAMINANTS IN THE AQUATIC ENVIRONMENT USING THE SPMD-TOX PARADIGM

B. Thomas Johnson
USGS—Columbia
Environmental Research Center
Columbia, Missouri

SPMD-TOX PARADIGM

SPMD-TOX (Fig. 1) is a monitoring tool that uses microscale techniques to assess the environmental impact of civilization on an aquatic resource. To determine a relevant profile, the sample must reflect the chemical contaminant's effect, not just its presence. To collect and concentrate bioavailable lipophilic contaminants, a semipermeable membrane device is used instead of live animals. To make risk assessments that are both effective and efficient, microscale toxicity tests (TOX) are employed. (Use of specific products by USGS and its laboratories does not constitute an endorsement.)

AQUATIC ENVIRONMENTAL SAMPLING

Water by its very physical and chemical nature is a natural repository and carrier of domestic, industrial, and agricultural products. In the last 50 years, analytical chemists have made significant strides in collecting, separating, and identifying these environmental waterborne contaminants (1). However, to determine the effect of these contaminants on an aquatic ecosystem, additional questions must be answered: Is the contaminant simply a chemical out of place, or is it a pollutant, a chemical toxin injurious to the health of the ecosystem? Where is the contaminant? Does the contaminant move? Is the contaminant a pollutant that is bioavailable and toxic to the ecosystem?

Bioavailability

Bioavailability is a critical concept in ecotoxicology because a harmful effect occurs only when biota are exposed to bioavailable pollutants (2). In an aquatic ecosystem, a chemical tends to exist in one of three forms: (1) dissolved in water, (2) passively sorbed directly into the sediment and on biotic or abiotic components suspended in the water column, or (3) actively integrated in some biota and bioaccumulated through respiratory or food ingestion pathways. It is evident that water-soluble contaminants are dissolved and bioavailable to all biota in the water column, but lipophilic contaminants are fat-soluble and, as a result, present a different challenge to identify their toxic effects on biota.

Samples

Samples from an aquatic environment come with some general knowledge. A site may be industrial, rural, urban, recreational, or commercial. The area of concern may be a result of stormwater, a chemical spill, a fish kill, a careless citizen, or a business error. Based on the Paracelsus adage, toxicologists know that all things are toxic and the dosage makes the "toxicant." Paracelsus (1493–1541) is often considered the father of modern toxicology. He brought empirical evidence into toxicology with his writings *"What is there that is not a poison? All things are poison and nothing without poison. Solely the dose determines that a thing is not a poison."* In this context, a toxicant must be defined both qualitatively (identified) and quantitatively (how much); toxicity is clearly dose-responsive. Simply stated, a waterborne chemical in the environment may be a contaminant at one concentration and a toxicant at another concentration; dosage makes the difference. Thus, the modern-day toxicity bioassay is predicated on the dose–response experimental design.

SPMDs

To collect and concentrate a lipophilic sample, a semipermeable membrane device (SPMD) that mimics the bioaccumulation of waterborne organic contaminants by fish is

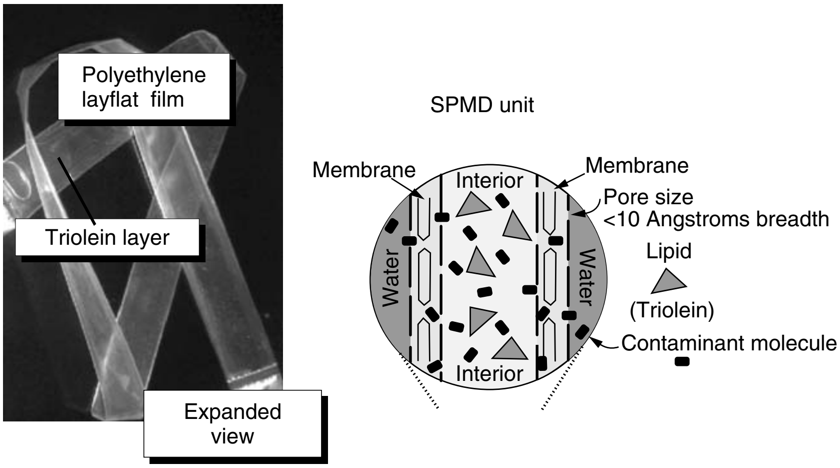

Dimensions: 2.5 cm × 15 cm

Figure 1. SPMD unit.

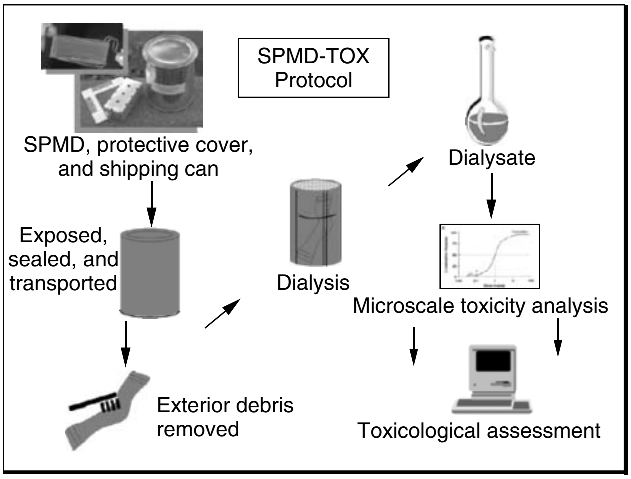

Figure 2. SPMD-TOX protocol.

used (SPMDs are sold exclusively by Environmental Sampling Technology in St. Joseph, MO); the SPMD permits estimates of tissue residues from the exposure site. The SPMD is a small low-density polyethylene lay-flat tube that contains neutral lipid triolein with a membrane surface area to lipid ratio of about $400\ \mathrm{cm}^2/\mathrm{mL}$ (Fig. 2). The SPMD is mounted in a protective stainless steel container and submerged generally 2 to 4 feet below the surface for about 30 days. Waterborne organic molecules diffuse through transitory openings (so called transport corridors) in the polyethylene membrane and sorb to the triolean, thus simulating the passive diffusion of a chemical across a fish gill's membrane. The SPMD monitors the lipophilic contaminant's bioavailability by a log K_{ow} value of less than one, a molecular weight of 600 daltons or less, and a size less than 10 Å. The entire unit is returned to the laboratory and the SPMD is chemically dialyzed using an organic solvent to recover sequestered chemicals.

The SPMDs as environmental contaminant-concentrating tools have many advantages. (1) They are abiotic, which means SPMDs do not metabolize sequestered products but provide a true reflection of bioavailable contaminants. (2) They are time integrative samplers, which means SPMDs will collect all lipophilic materials introduced into the water column during the exposure period, even if the introductions are episodic. (3) SPMDs can survive in heavily polluted, toxic environments where living organisms may not survive. (4) They are not temperature specific; SPMDs can be used in both cold and warm water. (5) SPMDs are easily transported to sites of interest for sampling and to laboratories for processing. (6) SPMD dialysates are analyzed by traditional GC/MS to identify the sequestered contaminants; algorithms and sampling rates are used to back-calculate concentrations in the water column. (7) SPMDs are cost-effective and, as a result, can be used in large monitoring programs.

TOXICITY ASSAYS

Microtox (3) determines the acute toxicity of SPMD dialysates by measuring the changes in light produced when samples are exposed to bioluminescent bacteria under standard conditions. Microtox bacteria *Vibrio fischeri* are clonal cultures, which diminish possible genetic differences and ensure quality control of the tester strain with greater assay sensitivity and precision. The freeze-dried bacteria are available on demand and all test media and glassware are prepackaged, standardized, and disposable. The test requires minimal laboratory space and limited dedicated equipment. The test is well defined, computer assisted, and user friendly.

Microtox is microscale. Dialysates are prepared in the basic dose–response design and placed in an SDI Model 500 Analyzer for incubation. [Columbia Environmental Research Center (CERC) uses Microtox materials and equipment sold by Strategic Diagnostics Inc. (SDI) in Newark, DE, to preserve the Microtox protocol. SDI provides comprehensive instructive guides, manuals and computer software to operate the Microtox test at their Web site (www.azurenv.com). The Microtox protocol described here is a standard USGS SOP.] The amount of light remaining in the sample is used to determine the sample's relative toxicity, which can then be compared to the standard reference's toxicity. As the toxicant's concentration increases, bacterial light emissions decrease in a dose-dependent manner. Supporting computer software is used to determine a 50% loss of light in the test bacteria, that is, the effective concentration (EC_{50}) value. Tests are completed and data are available in <30 min. The lower the EC_{50} value, the greater the acute toxicity of the sample.

Mutatox (4) determines the DNA-damaging substances in SPMD dialysates by measuring the increase in light produced when samples are exposed to dark mutant stains of *V. fischeri*. (Mutatox is prepared and sold exclusively by SDI. The Mutatox protocol described here is a standard USGS SOP.) Similar to Microtox, the test bacteria are stored freeze-dried under vacuum and require no preculturing; aseptic techniques are minimal; all test media and glassware are prepackaged, standardized, and disposable. The test requires minimal laboratory space and limited dedicated equipment. Because most bacterial cells such as these dark mutant strains fail to duplicate vertebrate biotransformation of progenotoxins into active genotoxins (DNA-damaging substances), a vertebrate activation system, typically rat or fish hepatic enzymes, is incorporated into the test.

Mutatox is microscale; all tests are conducted in microvolumes with microcuvettes. A single reaction cuvette contains the dark mutant bacteria, the diluent, the dialysate, a metabolic activation system, and a carrier solvent such as DMSO. Test samples are serially prepared over a 100-fold dose range, preincubated at 37 °C (metabolic activation) for 15 to 30 min, and then grown in the dark at 27 °C for 18 to 24 h. Benzopyrene, a commonly found PAH, is used as a positive control. A genotoxic response is defined as an increased light value of at least three times the light intensity of the solvent control blank. Mutatox is a qualitative test that provides a yes-no assessment of potentially DNA-damaging substances.

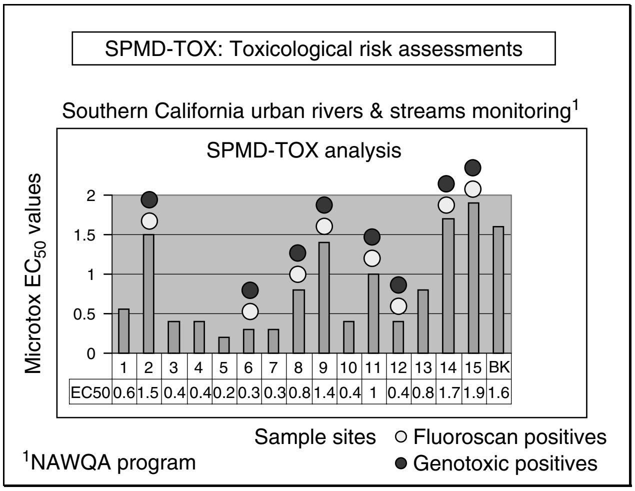

Figure 3. Risk assessment of Southern California urban streams with SPMD-TOX.

Case Study: A Simulated Oil Spill (5)

SPMD-TOX monitored a simulated oil spill during an 8-week period along the St. Lawrence River in Canada. An SPMD was placed weekly in each of the 20 oiled plots and removed 1 week later. The dialysates from the first 2 weeks fluoresced, indicating the presence of PAHs, and Mutatox analyses confirmed these findings with a yes response. Microtox analyses initially showed evidence of acute toxicity followed by a rapid decline and by week 4 no detectable acute toxicity. Samples from weeks 3 through 8 showed no evidence of the oil spill in the water column. SPMD-TOX analyses of these samples during the 8-week exposure period presented a clear picture of the occurrence and rapid recovery of the simulated oil spill site.

Case Study: Urban River Sites (6)

During a 30-day exposure period, SPMD-TOX was used to monitor 15 urban sites at rivers and streams in Southern California for the presence of bioavailable waterborne lipophilic ecotoxins. Fluoroscan, Microtox, and Mutatox data suggested the probable presence of PAHs and the initial identity of both acute toxic and genotoxic substances that were bioavailable in the water column (Fig. 3).

SUMMARY

SPMD-TOX is a sensitive, technically simple, and cost-effective assessment tool for monitoring waterways for bioavailable lipophilic chemical contaminants that may adversely affect aquatic communities and water quality. The SPMD-TOX paradigm can rapidly identify chemical incursions that may impact the quality of the water and provide the ecotoxologist and resource manager with valuable risk assessments of areas of concern.

BIBLIOGRAPHY

1. Manahan, S.E. (1989). *Toxicological Chemistry*. Lewis, Chelsea, MI.
2. Rand, G.M., Wells, P.G., and McCarty L.S. (1995). Introduction to aquatic toxicology. In: *Fundamentals of Aquatic Toxicology*, G.M. Rand, (Eds.). Taylor & Francis, Washington, DC. pp. 3–66.
3. Microbics Corporation. (1992). *Microtox Manual*. Microbics Corporation, Carlsbad, CA., Vols. I to V.
4. Johnson, B.T., Petty, J.D., and Huckins, J.N. (2000). Collection and detection of lipophilic chemical contaminants in water, sediment, soil and air—SPMD-TOX. *Environ. Toxicol.* **15**: 248–252.
5. Johnson, B.T. et al. (2004). Hazard assessment of a simulated oil spill on intertidal areas of the St. Lawrence River with SPMD-TOX. *Environ. Toxicol.* **19**: 329–335.
6. Johnson, B.T., Petty, J.D., Huckins, J.N., Goodbred, S., and Burton, C. (2002). Assessment of aquatic habitat quality in urban streams in the Western United States using SPMD-TOX. 23rd *Annu. Meet. Soc. Environ. Chem. Toxicol.*, November 15–20, 2002, Salt Lake City, Utah. Abstract.

Johnson, B.T. (1997). Microtox® Toxicity Test System—new developments and applications. In: *Microscale Testing in Aquatic Toxicology*, P.G. Wells, K. Lee, and C. Blaise (Eds.). CRC Press, Washington, DC, pp. 201–218.

Johnson, B.T. (2000). Revised: Activated Mutatox assay for detection of genotoxic substances. Technical Methods Section. *Environ. Toxicol.* **15**(3): 253–259.

USE OF LUMINESCENT BACTERIA AND THE *LUX* GENES FOR DETERMINATION OF WATER QUALITY

Nirit Ulitzur
Checklight Ltd.
Tivon, Israel

Shimon Ulitzur
Technion Institute of
Technology
Haifa, Israel

INTRODUCTION

Water quality monitoring tests have expanded over the years to include the use of bioassays for routine testing and screening of water quality in an attempt to overcome several drawbacks of using the physicochemical tests. The identification of chemicals by the sophisticated chemical and instrumental methods are conditioned by the presence of a reference sample. Such methods are unable to estimate possible biological synergistic or antagonistic interactions between chemicals, and their routine use is limited due to their high cost and the requirement for skilled personnel. The routine use of aquatic bioassays has been limited by the time required (usually days) to perform such tests as well as the cost and skill level. In order to overcome these limitations, numerous short-term bioassays have been developed in the past fifteen years, but the most commonly used and thoroughly validated rapid tests are the bioluminescence-based bioassays based on the use of luminescent bacteria. The following sections describe the various approaches and technologies that utilize luminescent bacteria for determination of water quality.

USE OF NATURAL LUMINESCENT BACTERIA FOR ON-SITE DETERMINATION OF WATER QUALITY

The generation of light by luminescent bacteria involves the activity of the enzyme luciferase that simultaneously catalyzes the oxidation of $FMNH_2$ and long chain aldehyde with the aid of molecular oxygen to yield FMN, long chain fatty acids, water, and light. Seven genes (*luxRICDABE*) are involved in the regulation and synthesis of the *Vibrio fischeri lux* proteins. The intricate molecular aspects of the bacterial *lux* system have been reviewed extensively (1–4).

The use of intact luminescent bacteria for analytical purposes has some clear advantages: the light of a single bacterium may reach about 30,000 quanta/s/cell; thus, one can use a simple luminometer to detect light generated by a few hundred cells/mL. Chemophysical and biological factors that affect cell respiration, the rate of protein or lipid synthesis, promptly alter the level of luminescence. In 1979, Bulich first showed that luminescent bacteria used for toxicity testing could be freeze-dried and provided to testing laboratories in a convenient reagent configuration (5). Freeze-dried (lyophilized) luminescent bacteria reagent can be stored for years at −18 °C. Upon hydration, freeze-dried cultures can yield 95–98% of their original viability, as well as their original level of *in vivo* luminescence. Thus, the use of these preparations for analytical purposes is not different, in practice, from other biochemical tests. Numerous studies described the use of natural luminescent bacteria or recombinant *lux*-carrying bacteria for rapid assessment of toxic components in aqueous environments. The next sections describe in more detail the bioluminescent tests for acute and chronic water toxicants.

MICROTOX TEST

The use of the bacterial luminescence system for the rapid assessment of aquatic toxicity was first suggested by Bulich (5) and later commercially marketed as Microtox. The bioassay is based on a preparation of lyophilized cultures of *Vibrio fischeri* (NRRL B11177). The lyophilized culture is hydrated with high purity water from which 10 μL aliquots each containing 10^6 cells are added to serial dilutions of the tested water supplemented with NaCl. Light measurements from the diluted water sample and clean water control are obtained after 5–15 min of incubation at 15 °C. The minimal concentration of tested water that results in 50% inhibition of light relative to the light level of the control is defined as IC_{50} (or EC_{50}) of the said water sample, or of the tested chemical spiked into it. Figure 1 shows typical results and determination of the IC_{50} value for a toxic agent spiked into a clean water sample. The Microtox test has been successfully applied worldwide for quality testing of industrial water, sewage, effluents, and contaminated sediments (6,7) and qualified as a standard for certain types of samples by various standards organizations in the United States and Europe. For many chemicals and toxic water samples, the sensitivity of the tested water in Microtox agreed well with the results obtained with daphnia or fish-based bioassays (8). In order to meet the need for the higher sensitivity required for drinking water analysis, we developed for AZUR Environmental a 22-h chronic test that involves bacterial growth and *de novo* synthesis of the *lux* system in *V. fischeri* (9,10). The test has been successfully used to screen drinking water samples for the presence of unexpected chemical contamination. The Microtox Test System is commercially available from Strategic Diagnostics, Inc.

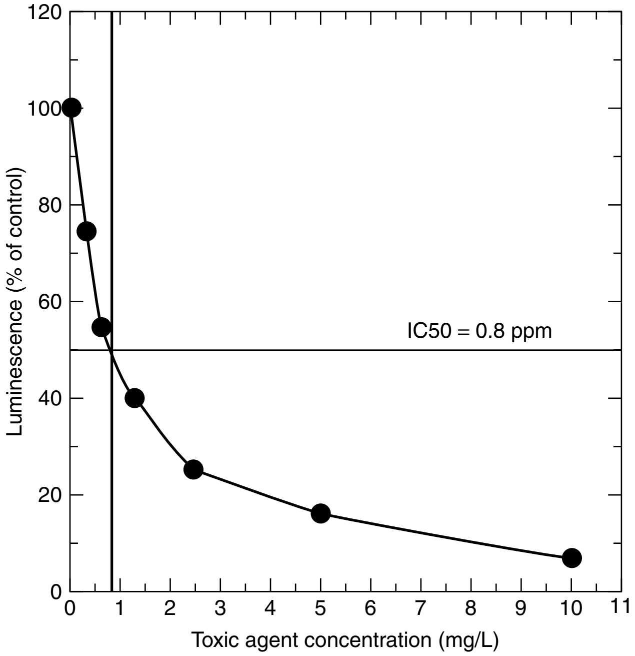

Figure 1. Typical response of luminescent bacteria to a toxic agent (see text for details).

TOXSCREEN TEST

A suitable test for drinking water toxicants should be rapid and sensitive enough to detect very low concentrations of diverse groups of toxic agents. Other requirements from such a test include simple operation, minimal false positive or negative responses, affordability, and the capability to run the test outside the laboratory. A novel bioassay (ToxScreen) that demonstrates those requirements has been developed by CheckLight (11). ToxScreen utilizes a marine luminescent bacterium (*Photobacterium leiognathi*) and applies unique assay buffers that markedly increase the sensitivity of the test and also enable the preliminary discrimination between organic and cationic metal toxicants. Cationic heavy metals can be detected at sub-milligram/liter levels; many of them were detected at 10–50 μg/L range. Similarly, pesticides, PAHs, and chlorinated hydrocarbons were detected in a separate test, all within 20–45 min. For most of the tested toxic agents, ToxScreen was found to be markedly (over 10–100-fold) more sensitive than Microtox (11). A practical feature of the new bioassay is the option of running the test at ambient temperatures (18–35 °C). In addition, the stability of the freeze-dried bacterial reagent preparation precludes the need for refrigeration or freezing during shipment. These features, together with its simple operation procedure, suggest that it has potential applications as a cost-effective prescreening tool to appraise chemical toxicity in various water sources. The higher sensitivity and simplicity of ToxScreen has recently been confirmed by the U.S. EPA's Environmental Technology Verification (ETV) program of rapid toxicity technologies (http://www.epa.gov/etv/verifications/vcenter1-12.html).

USE OF LUMINESCENT BACTERIA FOR ONLINE MONITORING OF WATER TOXICITY

The increasing concern for drinking water quality, together with the growing terrorism awareness, has emphasized the need for real-time monitoring of drinking water quality. Two kinds of bioluminescence-based online monitors for water quality have been developed so far. An online monitor that utilizes lyophilized cultures of luminescent bacteria was first developed by AZUR Environmental and Siemens in the mid-1990s. The device (Microtox-OS) was designed to automate the standard Microtox test and operate on-site, unattended for two weeks. The Microtox-OS Test System was equipped with a statistical process control software to continuously monitor data and was user-configurable to identify test results that were statistically different from preceding values. Due to its high operation cost, however, only a few units were sold.

The TOXcontrol-BioMonitor (now available through Microlan, The Netherlands) is based on freshly cultivated light-emitting *V. fischeri* cells that are continuously mixed with clean water and tested water. Light is detected using an online dual monitoring luminometer as it passes through twin flow cells. Any significant difference in light detected from the reference and monitoring streams results in a number of programmable events: an alarm can be programmed or a valve can be switched to collect samples for further analysis. No information has been published on the sensitivity of this automatic bioluminescent test. In general, luminescent bacteria growing in liquid medium in the presence of nutrients are not expected to show high sensitivity toward toxic agents. The system is mainly used for river biomonitoring and wastewater monitoring in Europe.

The integration of the ToxScreen test with an automatic online monitor was recently introduced by CheckLight Ltd. and Systea Srl and named MicroMac-ToxScreen. MicroMac-ToxScreen is an innovative automated online water quality monitoring system that utilizes the ToxScreen reagents to detect microgram/liter concentrations of toxic organic and inorganic chemical pollutants in surface or groundwater, as well as raw and treated drinking water. Freeze-dried luminescent bacteria are hydrated and kept in the device at 4 °C to maintain a stable luminescent culture. The instrument is resupplied with newly hydrated luminescent bacteria and a fresh inventory of liquid assay every two weeks. Automatic safeguards have been engineered into the system to assure reagent and data quality and appropriate instrument functioning. The instrument is also equipped with autocalibration features to assure reliable instrument performance; microprocessor-based system controls provide for data storage, data downloading, real-time communication with a remote PC, and user adjustable alarm levels.

Cho et al. (12) have recently adapted a 384-multiwell plate monitor for determination of water toxicity. Lyophilized *luxAB* carrying *Janthinobacterium lividum* cells are exposed to the water in question for 15 min followed by luminescence recording after addition of aldehyde. Similar to other terrestrial *lux*-carrying systems, this test is more sensitive to metals than the Microtox test.

LUX REPORTER GENES FOR DETERMINATION OF GENERAL STRESS-INDUCING AGENTS OR SPECIFIC TOXIC CHEMICALS

With the development of recombinant DNA technology, various recombinant bioluminescent bacteria have been engineered by fusing the *lux* gene with a stress-responding inducible promoter or with a specific promoter for a given toxicant (13). Exposing these recombinant cells to active toxic agents resulted in dramatic increase in bacterial luminescence. Such bacteria can be extremely useful for bioremediation studies; they can be utilized to determine the presence and concentrations of specific pollutants such as Hg(II) (14), or detection of linear alkanes (15) and naphthalene (15), or general responses to environmental stress (16,17).

A novel approach for identification of water toxicants by using *Escherichia coli* carrying *lux* genes fused with stress promoters was suggested by Ben Israel et al. (18). Twenty-three out of 25 tested chemicals were identified by a novel testing matrix using a 3-h test procedure. This biological recognition strategy also provides a tool for evaluating the degree of similarity between the modes of action of different toxic agents. Recently, Kim and Gu (19) described a high throughput toxicity monitoring and classification biosensor system using four immobilized bioluminescent *E. coli* strains that have plasmids bearing a fusion of a specific promoter to the *lux*CDABE operon. These strains are immobilized in a single 96-well plate using an LB-agar matrix. The bioluminescence increases when the water sample contains chemicals that cause either oxidative damage, membrane damage, DNA damage, or protein damage.

USE OF LUMINESCENT BACTERIA FOR DETERMINATION OF MUTAGENIC AND CARCINOGENIC AGENTS

Ulitzur et al. (20,21) and Weiser et al. (22) were the first to use luminescent bacteria for determination of mutagenic and genotoxic agents. The first test, later commercialized as Mutatox (23,24), was based on a spontaneous dark variant (M169) of the luminescent bacterium *Photobacterium leiognathi*. Light production was restored when cells were incubated in the presence of subacute concentrations of genotoxic agents, including base substitution, frameshift, DNA synthesis inhibitors, DNA damaging agents, and DNA intercalating agents (25). Mutatox test showed very good correlation with the Ames test (26). The primary genetic lesion of M169 was discovered later to be due to the presence of mutation in the *rpoS* gene (4).

Vitotox® is another example of a bioluminescent test for genotoxic agents (27,28). In this test, two recombinant *Salmonella typhimurium* TA104 strains are used to determine both genotoxicity and cytotoxicity of the tested sample. Under normal growth conditions, the luciferase reporter gene is repressed; in the presence of a DNA damaging agent, the SOS response is stimulated and leads to derepression of the strong promoter controlling the expression of the luciferase gene.

A new mutagenicity assay, in which a series of genetically modified strains of marine bacterium *Vibrio harveyi* was used for detection of mutagenic pollution of marine environments, has recently been described by Czyz et al. (29).

USE OF LUMINESCENT BACTERIA FOR DETERMINATION OF NUTRIENTS IN WATER

AOC (Assimilable Organic Carbon)

Regrowth potential of heterotrophic bacteria in potable water depends mainly on the presence of an assimilable organic carbon source. Many bacteria are capable of dividing in water containing as low as 5–10 μg/L of various carbon sources. Most of the standard or proposed AOC methodologies are hindered by their long duration and require between 5 and 14 days for completion. A bioassay that utilizes a nutrient-deprived culture of autoinducer-requiring mutant of *V. fischeri* has been developed and is commercially available (CheckLight Ltd). Luminescence in this strain is directly proportional to the concentration of utilizable organic material in the water sample. This 2-h test enables the detection of as little as 5 μg/L organic carbon and the test has shown high correlation with the standard 7-d long AOC test.

Haddix et al. (30) have recently described a test system in which *Pseudomonas fluorescens* P-17 and *Spirillum* species strain NOX bacteria were engineered with a *luxCDABE* operon fusion and inducible transposons that were selected on minimal medium. Independent mutants were screened for high luminescence activity and predicted AOC assay sensitivity. All mutants tested were able to grow in tap water under AOC assay conditions. Peak bioluminescence and plate count AOC tests were linearly related. Bioluminescence results that were obtained 2–3 d postinoculation were comparable with the 5-d ATP pool luminescence AOC assay and 8-d plate count assay.

BOD (Biochemical Oxygen Demand)

The biochemical oxygen demand (BOD) test (BOD5) is a crucial environmental index for monitoring organic pollutants in wastewater but is practically limited by the 5-d requirement for completing the test. A rapid BOD test based on an autoinducer-requiring mutant of *V. fischeri* is now commercially available (CheckLight Ltd.). The test procedure requires the water in question to be boiled for 30 min in the presence of 1 N HCl followed by addition of 1 N NaOH to neutralize the sample. This brief hydrolysis step is sufficient to break down proteins and complex carbohydrates into oligomers or monomers that can be assimilated by the bacteria. A good correlation was found between the 2-h bioluminescent test and the 5-d standard BOD test.

A BOD sensing system based on bacterial luminescence using recombinant *E. coli* containing *lux A-E* genes from *V. fischeri* has recently been described by Sakaguchi et al. (31). Frozen cells of luminescent recombinants of *E. coli* were applied to measure and detect organic pollution due to biodegradable substances in various wastewaters. The data from this study showed a similar correlation with that of the conventional method for BOD determination (BOD5).

Other Test Systems

Due to the intense use of synthetic nitrogen fertilizers and livestock manure in current agriculture practices, food (particularly vegetables) and drinking water may contain unacceptable concentrations of nitrate. A plasmid-borne transcriptional fusion between the *E. coli* nitrate reductase (*nar*G) promoter and the *Photorhabdus luminescens lux* operon provided *E. coli* with a highly bioluminescent phenotype when in the presence of nitrate concentrations as low as 5×10^{-3} mol/L (0.3 ppm) (32).

Recently, Gillor et al. (33) have described a reporter system for nitrogen availability. The promoter of the glutamine synthetase-encoding gene, P *gln*A, from *Synechococcus* sp. strain PCC7942 was fused to the *luxAB* luciferase-encoding genes of the bioluminescent bacterium *Vibrio harveyi*. The resulting construct was introduced into a neutral site on the *Synechococcus* chromosome to yield the reporter strain GSL. Light emission by this strain was dependent on ambient nitrogen concentrations. The linear response range of the emitted luminescence was 1 mM to 1 μM for the inorganic nitrogen species tested (ammonium, nitrate, and nitrite) and 10–50-fold lower for glutamine and urea. When water samples collected from a depth profile in Lake Kinneret (Israel) were exposed to the reporter strain, the bioluminescence of the reporter strain mirrored the total dissolved nitrogen concentrations chemically determined for the same samples. This test strain was shown to be a sensitive indicator of bioavailable nitrogen.

CONCLUSIONS

The bioluminescence tests for detecting water toxicants as well as for measuring elementary nutrients have been shown to be simple, rapid, and relatively inexpensive tools for monitoring water quality. Still, the relatively high cost ($3000–25,000) of commercially available luminometers restricts the routine applications of these technologies. The use of the *lux* genes in recombinant microorganisms has not been commercially applied yet mainly due to the restriction on using recombinant DNA in commercial products. Wider application of bioluminescent tests for screening water quality will ensure a better supply of adequate quality and will minimize the danger of water contamination by terrorist acts.

BIBLIOGRAPHY

1. Wilson, T. and Hastings, J.W. (1998). Bioluminescence. *Annu. Rev. Cell Biol.* **14**: 197–230.

2. Meighen, E.A. and Dunlap, P.V. (1993). Physiology, biochemical and genetic control of bacterial bioluminescence. *Adv. Microbol. Physiol.* **34**: 1–67.
3. Ulitzur, S. and Dunlap, P. (1995). The role of LuxR protein in regulation of the lux system of *Vibrio fischeri*. *Photochem. Photobiol.* **62**: 625–632.
4. Ulitzur, S., Matin, A., Fraley, C., and Meighen, E. (1997). H-NS protein represses transcription of the lux systems of *Vibrio fischeri* and other luminous bacteria cloned into *Escherichia coli*. *Curr. Microbiol.* **35**(6): 336–342.
5. Bulich, A.A. (1979). Use of luminescent bacteria for determining toxicity in aquatic environments. In: *Aquatic Toxicology, ASTM STP 667*. L.L. Markings and R.A. Kimerle (Eds.). American Society of Testing and Materials, Washington, DC.
6. Bulich, A.A. and Isenberg, D.L. (1981). Use of the luminescent bacterial system for the rapid assessment of aquatic toxicity. *ISA Trans.* **20**: 20–33.
7. Bulich, A.A. and Bailey, G. (1995). Environmental toxicity assessment using luminescent bacteria. In: *Environmental Toxicology Assessment*. M. Richardson (Ed.). Taylor and Francis, London, pp. 29–40.
8. Kaiser, K.L.E. (1993). Qualitative and quantitative relationships of Microtox data with toxicity data for other aquatic species. In: M. Richardson (Ed.). *Ecotoxicology Monitoring*, VCH, Weinheim.
9. Bulich, A.A. and Huynh, H. (1995). Measuring chronic toxicity using luminescent bacteria. *Can. Tech. Rep. Fish Aquat Sci.* **2050**: 23.
10. Backhaus, T., Froehner, K., Altenburger, R., and Grimme, L.H. (1997). Toxicity testing with *Vibrio fischeri:* a comparison between the long term (24 h) and the short term (30 min) bioassay. *Chemosphere* **35**: 2925–2938.
11. Ulitzur, S., Lahav, T., and Ulitzur, N. (2002). A novel and sensitive test for rapid determination of water toxicity. *Environ. Toxicol.* **17**: 291–296.
12. Cho, J.C. et al. (2004). A novel continuous toxicity test system using a modified luminescent freshwater bacterium. *Biosens. Bioelectron.* **20** (2): 338–344.
13. Ross, P. (1993). The use of bacterial luminescence systems in aquatic toxicity testing. In: M.L. Richardson (Ed.). *Ecotoxicology Monitoring*, VCH Verlagsgesellschaft, New York, p. 185–195.
14. Selifonova, O., Burlage, R., and Barkay, T. (1993). Bioluminescent sensors for detection of bioavailable Hg(II) in the environment. *Appl. Environ. Microbiol.* **59**(9): 3083–3090.
15. Werlen, C., Jaspers, M.C.M., and van der Meer, J.R. (2004). Measurement of biologically available naphthalene in gas and aqueous phases by use of a *Pseudomonas putida* biosensor. *Appl. Environ. Microbiol.* **70**: 43–51.
16. Belkin, S., Vollmer, A.C., van Dyk, T.K., Smulski, D.R., Reed, T.R., and LaRossa, R.A. (1994). Oxidative and DNA-damaging agents induce luminescence in *E. coli* harboring *lux* fusions to stress promoters. In: *Bioluminescence & Chemiluminescence: Fundamentals & Applied Aspects*. A.K. Campbell, L.J. Krica, and P.E. Stanely (Eds.). John Wiley, Chichester, UK, pp. 509–512.
17. DuBow, M.S. (1998). The detection and characterization of genetically programmed responses to environmental stress. *Ann. NYAS Online* **851**: 286–291
18. Ben-Israel, O., Ben-Israel, H., and Ulitzur, S. (1998). Identification and quantification of toxic chemicals by use of *Escherichia coli* carrying *lux* genes fused to stress promoters. *Appl. Environ. Microbiol.* **64**: 4346–4352.
19. Kim, B.C. and Gu, M.B. (2003). A bioluminescent sensor for high throughput toxicity classification. *Biosens. Bioelectron.* **18**(8): 1015–1021.
20. Ulitzur, S., Weiser I., and Yannai S. (1980). A new, sensitive, fast and simple bioluminescence assay for mutagenic compounds. *Mutat. Res.* **74**: 113–124.
21. Ulitzur, S. and Weiser, I. (1981). Acridine dyes and other DNA-intercalating agents induce the luminescence system of luminescent bacteria and their dark variants. *Proc. Nat. Acad. Sci. U.S.A.* **78**: 3338–3342.
22. Weiser, I., Ulitzur, S., and Yannai, S. (1981). DNA damaging agents and DNA synthesis inhibitors induce the luminescence in dark variants of luminescent bacteria. *Mutat. Res.* **91**: 443–450.
23. Kwan, K.K., Dutka, B.J., Rao, S.S., and Liu, D. (1990). Mutatox test: a new test for monitoring environmental genotoxic agents. *Environ. Pollut.* **65**(4): 323–332.
24. Bulich, A. (1992). Mutatox: a genotoxicity assay using luminescent bacteria. *Schriftenr Ver Wasser Boden Lufthyg* **89**: 763–770.
25. Ulitzur, S. (1986). Bioluminescence test for genotoxic agents. *Methods Enzymol.* **133**: 264–274.
26. Sun, T.S. and Stahr, H.M. (1993). Evaluation and application of a bioluminescent bacterial genotoxicity test. *J. AOAC Int.* **76**: 893–898.
27. van der Leile, D. et al. (1997). The VITOTOX test, and SOS bioluminescence *Salmonella typhimurium* test to measure genotoxicity kinetics. *Mutat. Res.* **389**: 279–290.
28. Verschaeve, L. et al. (1999). Vitotox® bacterial genotoxicity and toxicity test for the rapid screening of chemicals. *Environ. Mol. Mutagen.* **33**: 240–248.
29. Czyz, A. et al. (2002). Comparison of the Ames test and a newly developed assay for detection of mutagenic pollution of marine environments. *Mutat. Res.* **519** (1–2): 67–74.
30. Haddix, P.L., Shaw, N.J., and LeChevallier, M.W. (2004). Characterization of bioluminescent derivatives of assimilable organic carbon test bacteria. *Appl. Environ. Microbiol.* **70**(2): 850–854.
31. Sakaguchi, T. et al. (2003). A rapid BOD sensing system using luminescent recombinants of *Escherichia coli*. *Biosens. Bioelectron.* **19**(2): 115–121.
32. Prest, A.G., Winson, M.K., Hammond, J.R., and Stewart, G.S. (1997). The construction and application of a lux-based nitrate biosensor. *Lett. Appl. Microbiol.* **24**(5): 355–360.
33. Gillor, O. et al. (2003). *Synechococcus* PglnA::luxAB fusion for estimation of nitrogen bioavailability to freshwater cyanobacteria. *Appl. Environ. Microbiol.* **69**(3): 1465–1474.

WATER QUALITY MANAGEMENT

Chakresh K. Jain
K.D. Sharma
National Institute of Hydrology
Roorkee, India

INTRODUCTION

Water is one of the basic necessities for the survival of human beings and the prosperity of civilization. The ever increasing growth of population, industrialization, the steady rise in irrigation, urbanization, and the high level of living standards exert tremendous pressure on the available water which is highly uneven in its spatial and temporal distribution in quantity and quality. Water is put to different uses such as irrigation, industry, power generation, drinking, bathing, recreation, fisheries, wild life propagation, and pollution abatement. For each of the uses, water is required in appropriate quantity and quality. Nowadays, there is an increasing awareness of water quality, especially in urban areas. Due to the rapid increase in population and the growth of industrialization in the country, the pollution of natural water by municipal and industrial wastes has increased tremendously.

The term "water quality" is a widely used expression, which has an extremely broad spectrum of meanings. Each individual has vested interests in water for a particular use. The term water quality, therefore, must be considered relative to the proposed use of water. From the user's point of view, the term "water quality" is defined as the physical, chemical, and biological characteristics of water by which the user evaluates the acceptability of water. For example, for the sake of human health, we require a pure, wholesome, and potable water supply.

In recent years, rivers are being used indiscriminately for disposal of municipal, industrial, and agricultural wastes, thereby polluting the river water beyond permissible limits. Due to this, river waters are gradually becoming unfit even for irrigation in some places. Thus it has become essential to evaluate the environmental impacts of water resources to minimize the progressive deterioration in water quality. Therefore, a detailed study of water quality in all vulnerable rivers is imperative for better management and use of water.

WATER QUALITY MONITORING

Water quality problems stem from two factors, the natural hydrology of a river basin and the development and use of land and water resources by human beings. Depending on the interrelation of these two factors, a wide variety of quality problems can result. Each river basin, therefore, is unique, and it must be subjected to individual and intensive water quality assessment to provide a proper basis for judicious management of land and water.

The following points should be considered for proper planning of a water quality monitoring program:

What are the objectives of the program?
From where are samples to be taken?
Which determinants are of interest?
When and how often are samples to be taken?
What is to be done with the results?

These questions will provide a framework for defining measurement programs. The results obtained from such a program should be regularly reviewed to decide if any changes (e.g., determinants or sampling frequency) in the monitoring program are necessary.

Objectives of the Monitoring Program

It is obvious that the objectives of a program should be clearly and precisely formulated by the user. If it is not done, inappropriate analytical data may well be provided, and/or the user is likely to call for needlessly large or unduly small numbers of results. Furthermore, if the objectives are not precisely expressed, it will be difficult or impossible to decide the extent to which they are achieved.

It is suggested that analytical information should be requested regularly only when the user knows beforehand that the results will be used in a precisely known fashion to answer one or more defined questions on quality. Requests for analysis based on the thought that the results may ultimately prove useful should be avoided, particularly when the analytical sampling effort is limited. There is an almost infinite number of analyses that might be useful in most situations, but it is completely impracticable to attempt any such comprehensive coverage. Therefore, selection from all possible objectives and determinants is essential for proper planning of a monitoring program.

As a further means of optimizing measurement programs, users should formulate their information needs as quantitatively as possible. As an extreme and perhaps rather artificial example of a badly defined requirement, consider the statement, "to obtain information on the quality of the river." Such a statement is almost completely useless as a basis for designing a monitoring program for the following reasons:

1. The determinants are not specified, so that the analyses required are not known.
2. The particular river and the locations on the river are not defined, so that inappropriate sampling positions may be chosen.
3. No indication is given of the timescale or sampling frequency, so that too few or too many samples may be collected and analyzed.
4. As a result of all of the above deficiencies, there is no indication of the amount of data that needs to be processed and the nature of the data treatment, so that appropriate data handling techniques cannot be defined.

Therefore, users of analytical results must seek to avoid uncertainties such as those in the preceding example by careful and quantitative definition of every aspect of their requirements. Thus, an objective such as the above would be better expressed by a statement of the form, "to estimate each year the annual average concentration of ammonia ($NH_3 + NH_4^+$) at all river sites used for the production of potable water." Appropriate statements of this type for other determinants of interest then provide a set of quantitative targets essential in optimizing the choice of sampling, analytical, and data handling techniques.

Objectives of GEMS/WATER

The fundamental objectives of the water quality monitoring system within the GEMS/WATER program are to:

1. Assess the impacts of human activities upon the quality of the water and its suitability for required uses.
2. Determine the quality of water, in its natural state, which might be available to meet future needs.

3. Keep under observation the source and pathways of specified hazardous substances.
4. Determine the trend of water quality at representative stations.

The first objective is met by the establishment of impact stations, the second by baseline stations, the third by either impact or baseline stations depending upon whether the hazardous substance is of artificial or natural origin, and the fourth by trend stations.

Baseline stations are located where no direct diffuse or point sources of pollutants are likely to be found. They are used to establish the natural background level of variables, to check if no synthetic compounds are found in remote areas (e.g., DDT), and to assess the long-term trends of surface water quality resulting from global atmospheric pollution. Impact stations are situated in waterbodies where there is at least one major use of the water or which are greatly affected by human activities. Four type of impact stations can be identified according to different uses of water:

1. *Drinking water*: at the raw water intake before treatment for drinking water.
2. *Irrigation*: at the water intake before distribution for irrigation.
3. *Aquatic life*: river and lake stations representative of the general quality of the waterbody.
4. *Multiple impacts*: several water uses at the station and/or of the waterbody.

Trend stations are set up specially to assess the trends in water quality. They must be representative of a large area with various types of human activities. These stations should be more frequently sampled to increase the statistical significance of the average concentrations and to validate the trends.

Sampling Locations and Points

Sampling location and point mean the general position within a waterbody and the exact position at which samples are obtained. The objectives of a program sometimes immediately define the sampling locations. For example, when the concern is to measure the efficiency of a chemical plant for purifying water, sampling locations will be required before and after the plant. Similarly, when the effect of effluent discharge on the water quality of a receiving river is of interest, samples will be required from locations upstream and downstream of the discharge. For larger scale water bodies (e.g., a river basin, a large estuary, a large urban drinking water distribution system), however, the objectives may be defined in terms that do not indicate sampling locations. For example, objectives such as "to measure river quality within a river basin" or "to measure the quality of water in a distribution system" give no indication of which of the large number of possible sampling locations are of interest. Such broadly expressed objectives are completely inadequate for detailed planning of efficient sampling programs and should always be sharpened, so that they do indicate the positions of sampling locations. A commonly useful device for helping in this respect is to consider the intended use of the water because this aids in indicating these positions in a waterbody where quality is of key importance.

Figure 1 illustrates the hypothetical case of a river system along with criteria for choosing different sampling sites. Figures 2 and 3 illustrate the location of sampling sites for lakes and groundwater with corresponding criteria for choosing of different sampling sites.

Determinants of Interest

The particular determinants appropriate for a program depend critically on the type of water and the objectives of the study. Depending on the intended use of a receiving water, the parameters listed in Table 1 are significant for water quality characterization. They are guidelines for analysis of wastewater quality for treatment and control. The parameters listed in Table 2 are frequently used to identify various types of pollution from industrial wastewater.

Some of the most important and most frequently used tests for analyzing water are the nonspecific tests listed in Table 3. These tests often measure a property of a group of substances. For example, alkalinity indicates the capacity of water to neutralize hydrogen ions. Many of these tests are used to determine the suitability of natural water for industrial or municipal use and to determine the type and degree of treatment required. Table 4 lists some of the more frequently measured parameters in pollution studies.

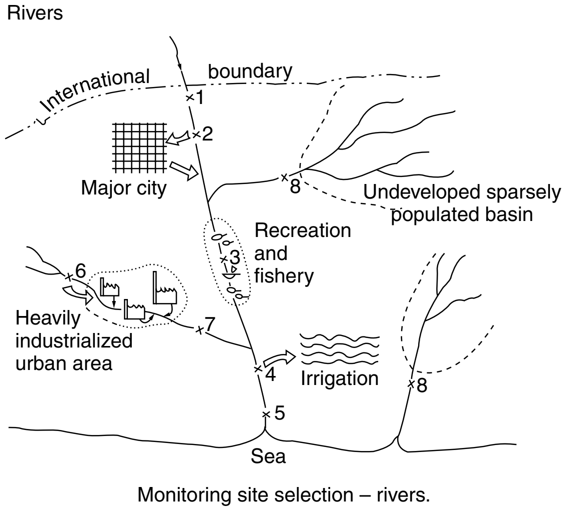

Station	Type	Criteria
1	Impact	Immediately downstream of an international (or baseline) boundary
2	Impact	Abstraction for public supply of large town
3	Impact	Important fishing, recreation, and amenity zone
4	Impact	Abstraction for large-scale agricultural area
5	Trend	Freshwater tidal limit of major river
6	Impact	Abstraction for large industrial activity
7	Impact	Downstream of industrial effluent discharge and important tributary influencing main river
8	Baseline	Station where water is in a natural state (no direct or indirect pollution, no water use)

Figure 1. Monitoring site selection—rivers.

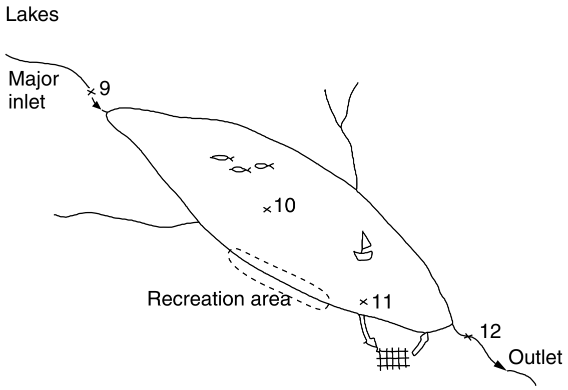

Monitoring site selection – lakes.

Station	Type	Criteria
9	Impact (or baseline)	Principal feeder tributary
10	Impact (or baseline)	General water quality of lake
11	Impact	Water supply of major city
12	Impact (or baseline)	Water leaving lake

Figure 2. Monitoring site selection—lakes.

Frequency and Time of Sampling

The quality of water in various waterbodies is rarely constant in time. There may be some relationship between the role of change in different variables, but others alter independently. In measuring the mean, maximum, and minimum values of variables over a period of time, the closeness of the monitored values to the true values depend upon the variability of the variables and the number of samples taken. The larger the number of samples from which the mean is derived, the narrower are the limits of the probable difference between the observed and true means. These confidence limits are not directly proportional to the number of samples but to the square of the number. Therefore, to double the reliability of a mean value, the number of samples must be increased fourfold.

Variations in water quality are caused by changes (increase or decrease) in the quantity of any of the inputs to a waterbody. Such changes may be natural or man-made and either cyclic or random. Water quality variation may therefore be similarly cyclic or random.

Variability differs among rivers, lakes, and underground waters. It is most pronounced in rivers, and the ranges are greater, the nearer the sampling point is to the source or sources of variability. As the distance from the source increases, longitudinal mixing smooths out irregularities and fewer samples are needed to meet given confidence limits. However, as the distance between the source of variability and the sampling point increases, there is a reduction in the range of variation, and there is also dilution. Some variables are reduced by self-purification,

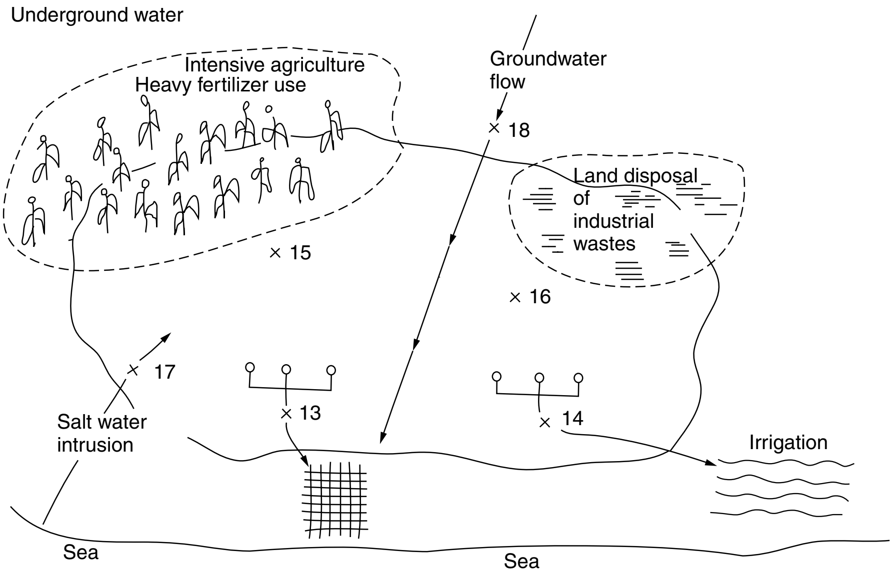

Monitoring site selection – underground water

Station	Type	Criteria
13	Impact	Water supply to large town threatened by fertilizer residues and saline intrusion
14	Impact	Water for large-scale irrigation threatened by leachate from waste tips
15 } 16 } 17 }	Impact	Water supply of major city
18	Baseline	No human activities in the groundwater recharge area

Figure 3. Monitoring site selection—underground water.

Table 1. Parameters for Water Quality Characterization

Water Use	Quality Parameters
Domestic water supply	Color, odor, taste
	Organic content: chlorine demand, COD, BOD, TOC, phenols
	Carcinogens and toxic compounds: insecticides, pesticides, detergents
	Turbidity, salinity
	Alkalinity, pH
	Total hardness, Ca, Mg, Fe, Si, etc.
	Pathogenic organisms, total bacterial count (37 °C), *E. coli*, plankton count
Fish, shellfish, wildlife, and recreation	Color, odor
	Toxic compounds
	Turbidity, floating matter, sludge deposits, salinity
	Temperature
	Dissolved oxygen, BOD
	Alkalinity, pH, pathogenic organisms, plankton count, nitrogen, phosphorus (inorganic nutrients that support algal blooms and other undesirable aquatic growth)
Agricultural irrigation	Salinity and Na-Ca content
	Alkalinity, pH
	Pesticides, growth regulators
	Persistent synthetic chemicals (e.g., polyethylene derivatives, asphalt sprays.)
	Pathogenic organisms
Watering livestock	Salinity
	Toxic compounds
	Pathogenic organisms
	Plankton count

Table 2. Significance of Parametric Measurements

Test(s)	Significance
Dissolved solids	Soluble salts may affect aquatic life or future use of water for domestic or agricultural purposes
Ammonia, nitrites	Degree of stabilization (oxidation) or nitrates and total organic nitrogenous matter, organic nitrogen
Metals	Toxic pollution
Cyanide	Toxic pollution
Phenols	Toxic pollution, odor, taste
Sulfides	Toxic pollution, odor
Sulfates	May affect corrosion of concrete, possible biochemical reduction to sulfides
Calcium and magnesium	Hardness
Synthetic detergents	Toxic pollution

deposition, and adsorption. These effects must be considered if a sampling station used for quality control is located some distance from the point of use.

In lakes, the mass of water and good lateral mixing provide inertia against any rapid changes from modifications in input and output. Many lakes exhibit marked seasonal variations due to thermal stratification, overturn, and biological activity. Depending on the type of lake, it may be sampled with a seasonal bias related to the natural cycles of the lake.

Underground water has a lower variability than that of either rivers or lakes. The rate of quality change depends on the depth of sampling, the volume of the aquifer, and hydraulic conductivity. The time elapsing between changes in land use and in surface recharge water and their effect upon the underground water depends on the time of percolation. Variations are often, but not invariably, seasonal with a time lag according to the rate of percolation. Direct injection into boreholes or saline intrusion from subterranean sources may take effect more rapidly.

The time of sampling is also of main concern when the quality of the water shows more or less regular variations, for example, diurnal variations in the concentration of dissolved oxygen in rivers and variations of quality on start-up and shutdown of industrial plants.

It is seldom that the quality of water at only one instant is of interest. Normally, information is required for a period during which quality may vary. The problem arises, therefore, of deciding the time at which to collect samples, so that they will adequately represent the quality during the period of interest.

The best technical solution to this problem is often to use an automatic, on-line instrument that provides continuous analysis of the water of interest. This approach can be of great value in that, in principle, a continuous record of quality is obtained and the problems of selecting particular times for sampling do not arise.

Table 3. Nonspecific Water Quality Parameters

Physical	Chemical	Physiological
Filterable residues	Hardness	Taste
Salinity	Alkalinity and acidity	Odor
Density	BOD	Color
Electrical conductance	COD	Suspended solids
	Total carbon	Turbidity
	Chlorine demand	

Table 4. Tests Used for Measuring the Pollution of Natural Waters

Nutrient Demand	Specific Nutrients	Nuisance	Toxic
DO	Nitrogen	Sulfide	Cyanide
BOD	Ammonia	Sulfite	Heavy metals
COD	Nitrate	Oil and grease	Pesticides
Total carbon	Nitrite	Detergents	
	Organic nitrogen	Phenols	
	Phosphorus		
	Orthophosphate		
	Polyphosphate		
	Organic phosphorus		

MAJOR WATER QUALITY ISSUES

Water Scarcity

Due to uneven distribution of rainfall in time and space and ever-increasing demand of water for agricultural, industrial, and domestic activities, water resources are overexploited. This results in shrinking or even drying up of many waterbodies for a considerable period in a year.

Maintaining targeted water quality in such waterbodies is very difficult. If the objectives of the Water Act of 1974 are to be fulfilled, maintaining a minimum level of water needs to be specified, for example, restrictions on water abstraction from different waterbodies (rivers, lakes, or ground water). The Central Ground Water Authority (CGWA) formed under the Environment (Protection) Act of 1986 has already taken such initiatives. The restrictions need to be extended also to rivers and lakes.

Targets for conservation include reducing demands by optimum use, minimization of waste, efforts to reduce percolation and evaporation losses, conservation efforts in domestic uses, groundwater recharging, rainwater harvesting, afforestation, recycling, and reuse.

There are a number of cases where salinity is increasing in both surface water and groundwater. The increase in groundwater salinity is mainly due to increased irrigation or seawater intrusion in coastal areas. The salinity in surface water is increasing mainly due to discharges of industrial wastewater or agricultural return water. Salinity impairs the fitness of water for drinking or irrigation. It may also affect the ecosystem in surface waters.

Oxygen Depletion

A large portion of wastewater mostly from domestic sources is discharged into watercourses without any treatment. Such wastewater contains high amounts of organic matter. Industries also discharge effluents containing high levels of organic matter, for example, agro industries. When oxidized in water by microbes, this organic matter consumes dissolved oxygen. Water has limited availability of oxygen, and when consumption exceeds the availability, oxygen depletion results. Thus, the survival of aquatic life becomes difficult.

In many waterbodies, a massive input of organic matter sets off a progressive series of chemical and biological events in downstream water. The stretch is characterized by high bacterial population, cloudy appearance, high BOD, and strong disagreeable odor—all indicating general depletion of oxygen. Masses of gaseous sludge rising from the bottom are often noticed floating near the surface of the water. During the monsoon, the sludge deposited in such stretches is flushed and stays in suspension, causing a rise in oxygen uptake downstream. Due to such sudden oxygen depletion, heavy fish mortality occurs every year during the first flushing after the onset of the monsoon.

Pollution Due to Urbanization

Urbanization has encouraged people to migrate from villages to urban areas. This has given rise to a number of environmental problems, such as water supply and wastewater generation, collection, treatment, and disposal. In urban areas, water is tapped for domestic and industrial use from rivers, streams, wells, and lakes. Approximately 80% of the water supplied for domestic use passes out as wastewater. In most cases, wastewater is let out untreated, and either it percolates into the ground and in turn contaminates the groundwater or is discharged into the natural drainage system causing pollution in downstream areas.

Nonpoint Source Pollution

Nonpoint sources of water pollution have been recognized as of greater importance than point sources. This is due in part to the continuing efforts to reduce pollution from point sources over the past few decades, as well as recognition that nonpoint sources, such as stormwater, may contain harmful contaminants. Pesticides and nutrients, particularly, are of major concern because of high toxicity and eutrophication. This has increased the need to identify and quantify major sources of nutrients and pesticides deposited in river systems.

For proper understanding of the nature and the magnitude of diffuse water pollution under different circumstances, it is necessary to consider some source related characteristics, briefly discussed below, for diffuse pollution from some more common sources.

In rural areas, most people use open fields for defecation, a few use pit latrines or septic tanks. Much of the bathing and washing (clothes, utensils, etc.) is in or near the waterbody, reducing abstraction and transport of water but causing *in situ* diffuse pollution.

Eutrophication

The discharge of domestic wastewater, agricultural return water or runoff water, and many industrial effluents contributes nutrients such as phosphates and nitrates. These nutrients promote excess growth of algae in waterbodies. This is not desirable for a balanced aquatic ecosystem.

Salinity

Salinity is increasing in many waterbodies, especially groundwater, due to leaching of salt buildup in agricultural areas under intense irrigation. A number of industrial activities discharge wastewater with high dissolved solids that increase the salinity of water.

Due to the discharge of toxic effluents from many industries and increased use of chemicals in agriculture and their subsequent contribution to waterbodies, many waterbodies in the country are polluted by toxic substances. Toxic substances impair water quality and make it unfit for human consumption, aquatic life, and irrigation.

Natural Contaminants

By far the greatest water quality problem in developing countries is the prevalence of waterborne diseases. Yet, in addition to human induced pollution problems, water supplies also have specific natural quality problems, which are mainly related to local geology. Some of the specific problems of natural origin include fluoride and arsenic.

Pathogenic Pollution

Waterborne diseases are the most important water quality issues in India. This is mainly due to inadequate arrangements for transport and treatment of wastewaters. A major portion of the wastewater generated from human settlements is not properly transported and treated before discharge into natural waters. This results in contamination of both surface water and groundwater. Moreover, the contribution of pathogens from diffuse sources is also quite significant. Thus, most of the surface waterbodies and many groundwater sources are contaminated.

A large population in the country still uses water directly for drinking or contact use without any treatment and thus are exposed to waterborne diseases. This is the single major cause of mortality due to water pollution.

Ecological Health

A large number of areas in our aquatic environment support rare species and are ecologically sensitive. They need special protection. The Water Act of 1974 provides for maintaining and restoring the wholesomeness of aquatic resources, which is directly related to the ecological health of waterbodies, so it is important that the ecological health of waterbodies is given first priority in water quality goals.

GROUNDWATER QUALITY

In recent years, groundwater pollution due to human activity has become an increasing threat. The adverse effects on groundwater quality are the result of human activity on the ground, unintentionally from agriculture, domestic and industrial effluents, and unexpectedly from subsurface or surface disposal of sewage and industrial wastes.

A major problem in urbanized areas is the collection and disposal of domestic wastewater. Because a large volume of sewage is generated in a small area, the waste cannot be adequately disposed of by conventional septic tanks and cesspools. Therefore, special disposal sites may be required to collect and dispose of such wastes in densely populated areas.

The quality of groundwater is of great importance in determining the suitability of a particular groundwater for a certain use (public water supply, irrigation, industrial applications, power generation, etc.). The quality of groundwater is the result of all the processes and reactions that have acted on the water from the moment it condensed in the atmosphere to the time it is discharged by a well. Therefore, the quality of groundwater varies from place to place, with the depth of water table, and from season to season, and is primarily governed by the extent and composition of dissolved solids in it.

The wide range of contamination sources is one of the many factors contributing to the complexity of groundwater assessment. It is important to know the geochemistry of chemical–soil–groundwater interactions to assess the fate and impact of pollutants discharged onto the ground. Pollutants move through several different hydrologic zones as they migrate through the soil to the water table. The serious implications of this problem necessitate an integrated approach to undertake groundwater pollution monitoring and abatement programs.

The supply of groundwater is not unlimited, nor is it always available in good quality. In many cases, the abstraction of excessive quantities of groundwater has resulted in drying up of wells, saltwater intrusion, and drying up of rivers that receive their flows in dry seasons from groundwater. Groundwater quality is increasingly threatened by agricultural, urban, and industrial wastes, which leach or are injected into underlying aquifers. Once pollution has entered the subsurface environment, it may remain concealed for many years, becoming dispersed over wide areas, and rendering groundwater supplies unsuitable for human use.

A vast majority of groundwater quality problems are caused by contamination, overexploitation, or combination of the two. Most groundwater quality problems are difficult to detect and hard to resolve. The solutions are usually very expensive, time-consuming, and not always effective. Groundwater quality is slowly but surely declining everywhere. Groundwater pollution is intrinsically difficult to detect; it may well be concealed below the surface, and monitoring is costly, time-consuming, and somewhat hit-or-miss by nature.

Groundwater in several areas, where sewage is discharged without proper treatment, has been adversely affected by contaminants from sewage. Due to indiscriminate use of fertilizers, very high concentrations of potassium and nitrate have been found in groundwater at several places in the states of Punjab, Haryana, and Uttar Pradesh. Excessive concentrations of nitrate in groundwater that may originate from fertilizers or sewage, or both, have been reported to the extent of several hundred, mg/L in various parts of the country. Groundwater pollution from some industrial sources has reached alarming levels. High levels of hexavelant chromium at Ludhiana and Faridabad, lead near Khetri in Rajasthan, nickel in Coimbatore, and cadmium in Kanpur and parts of Delhi are some of the manifestations of heavy metal pollution. Arsenic concentrations in groundwater have been found in West Bengal in excess of the permissible limit of 0.05 mg/L. The population in these areas is suffering from arsenic dermatosis from drinking arsenic-rich groundwater.

The problem of groundwater pollution in several parts of the country has become so acute that unless urgent steps for detailed identification and abatement are taken, extensive groundwater resources may be damaged. All activities carried out on land have the potential to contaminate groundwater, whether from urban, industrial, or agricultural activities. Large-scale, concentrated sources of pollution, such as industrial discharges, landfills, and subsurface injection of chemicals and hazardous wastes, are obvious sources of groundwater pollution. These concentrated sources can be easily detected and regulated, but the more difficult problem is from diffuse sources of pollution, such as leaching of agrochemicals and animal wastes, subsurface discharges from latrines and septic tanks, and infiltration of polluted urban runoff and sewage where sewerage does not exist or is defunct. Diffuse sources can affect entire aquifers, which are difficult to control and treat. The only solution to diffuse sources of pollution is to integrate land use with water management. Table 5 presents land-use activities and their potential threat to groundwater quality.

Common Groundwater Contaminants

- *Nitrates*: Dissolved nitrate is the most common contaminant in groundwater. High levels can cause blue baby disease (methemoglobinemia) in children, may form carcinogens, and can accelerate eutrophication in surface waters. Sources of nitrates include sewage, fertilizers, air pollution, landfills, and industries.
- *Pathogens*: Bacteria and viruses that cause waterborne diseases such as typhoid, cholera, dysentery, polio, and hepatitis. Sources include sewage, landfills, septic tanks, and livestock.
- *Trace metals*: Include lead, mercury, cadmium, copper, chromium, and nickel. These metals can be toxic and carcinogenic. Sources include industrial and mine discharges and fly ash from thermal power plants due to either fallout or disposal in ash ponds.
- *Organic compounds*: Include volatile and semivolatile organic compounds such as petroleum derivatives, PCBs, and pesticides. Sources include agricultural activities, street drainage, sewage landfills, industrial discharges, spills, and vehicular emission fallout.

Table 5. Land-Use Activities and Their Potential Threat to Groundwater Quality

Land Use	Activities Potential to Groundwater Pollution
Residential	Unsewered sanitation
	Land and stream discharge of sewage
	Sewage oxidation ponds
	Sewer leakage, solid waste disposal, landfill
	Road and urban runoff, aerial fallout
Industrial and commercial	Process water, effluent lagoon
	Land and stream discharge of effluent
	Tank and pipeline leakage and accidental spills
	Well disposal of effluent
	Aerial fallout
	Landfill disposal and solid wastes and hazardous wastes
	Poor housekeeping
	Spillage and leakage during handling of material
Mining	Mine drainage discharge
	Process water, sludge lagoons
	Solid mine tailings
	Oilfield spillage at group gathering stations
Rural	Cultivation with agrochemicals
	Irrigation with wastewater
	Soil salinization
	Livestock rearing
Coastal areas	Saltwater intrusion

LEGAL CONSIDERATIONS

The basic objective of the Water (Prevention and Control Pollution) Act of 1974 is to maintain and restore the wholesomeness of national aquatic resources. Due to the large variation in type, size, shape, quality, and quantity of water available from our aquatic resources, each one has a very specific waste load receiving capacity. This implies the need for prescribing different effluent standards based on the assimilative capacity of recent systems. Notwithstanding its merits, it is difficult to administer compliance with widely varying standards for specific types of effluents. To reduce the administrative difficulties of relating effluent quality to ambient water quality, the concept of Minimum National Standards (MINAS) was developed by the Central Pollution Control Board, whereby minimum effluent limits are prescribed for each category of discharge, regardless of receiving water requirements. Where water quality standards cannot be reached by imposing a standard level of treatment alone, then and only then will the conditions of the receiving waters dictate more stringent controls (State Boards can make the MINAS more stringent). The conventional methods of treatment cannot cope with these specific situations, and the polluters are required to introduce new methods for specific requirements. The number of such situations in our country is gradually rising due to water scarcity in many waterbodies and some times due to the nature of effluents.

APPROACH TO WATER QUALITY MANAGEMENT

The Water (Prevention and Control of Pollution) Act of 1974 reflects the national concern for water quality management. The basic objective of the Act is to maintain and restore the wholesomeness of water through prevention and control of pollution. The Act does not define the level of wholesomeness to be maintained or restored in different waterbodies of the country. To define these levels for different waterbodies, the Central Pollution Control Board (CPCB) had initially taken the use of water as a basis for identifying water quality objectives for different waterbodies. Hence, it has classified the national aquatic resources according to their uses. CPCB has also identified primary water quality criteria for different uses of water as a yardstick for preparing different pollution control programs. The use-based classification system is given in Table 6.

The "designated best use" yardstick, as explained above, is based on a limited number of parameters called primary criteria. While implementing such criteria, there are many practical difficulties, which need scientific support. The criteria should be such that they can support the designated uses and also the ecological sustainability of the waterbody (wholesomeness), which is the prime objective of the Water Act of 1974. Hence, the CPCB, with the help of an expert group, has revised the approach.

Table 6. Use-Based Classification of Surface Waters in India

Designated Best Use	Quality Class	Primary Quality Criteria
Drinking water source without conventional treatment, but with chlorination	A	6.5 to 8.5 (1); 6 or more (2); 2 or less (3); 50.5–200%, and 20–50% (4); NIL (5–8)
Outdoor bathing (organized)	B	6.5 to 8.5 (1); 5 or more (2); 3 or less (3); 500, 5%-2000, and 20%-500 (4); NIL (5–8)
Drinking water source with conventional treatment	C	6.5 to 8.5 (1); 4 or more (2); 3 or less (3); 5000, 5%-20000, and 20%-5000 (4); NIL (5–8)
Propagation of wildlife and fisheries	D	6.5 to 8.5 (1); 4 or more (2); NIL (3–4); 1.2 (5); NIL (6–8)
Irrigation, industrial cooling, and controlled waste disposal	E	60 to 8.5 (1); NIL (2–5); 2250 (6); 26 (7); 2 (8)

(1) pH.
(2) Dissolved oxygen, mg/L.
(3) BOD, (20 °C) mg/L.
(4) Total coliform (MPN/100 mL).
(5) Free ammonia, mg/L.
(6) Electrical conductivity, μmho/cm.
(7) Sodium adsorption ratio.
(8) Boron, mg/L.

The Revised Approach

Human interests cannot be ignored or downgraded in importance, but it is now widely accepted that the long-term interests of human beings themselves lie in maintaining the environment and ecosystems in an overall healthy condition. A system totally oriented to the objective of protecting direct beneficial uses by humans and of classifying water quality on that basis may not be adequate. In many cases (such as for small waterbodies or those with no significant current water use), the objective may have to be the protection of basic environmental qualities for sustaining ecosystems. Even where current or potential beneficial uses can be clearly defined and a "designated best-use," the first priority in water quality assessment and management would be maintaining and restoring the desirable level of general environmental quality or "wholesomeness." With this approach, protection of designated best use is not being abandoned, but the overall health of ecology is given its rightful place, ahead of any direct current beneficial uses by humans. Such ecological considerations shall ensure the biological integrity of a waterbody, including structural as well as functional integrity.

For effective management of water quality, it is necessary to involve local administrative bodies at the district, block, municipal, and village panchayat levels, NGOs, and the citizens themselves in conducting at least a crude, rapid, and overall assessment of water quality.

Water quality requirements for specific large-scale organized uses may continue to be prescribed and even strengthened, but the stakeholders could share responsibility for water quality monitoring and management in such cases. It would lead to a two-pronged approach; defining water-quality requirements and classification with the objective of (1) maintaining and restoring wholesomeness of water for the health of ecosystems and the environment in general (basic water quality requirement) and (2) protecting designated organized uses of water by humans and expecting stakeholders to share responsibility in monitoring and managing such quality by associating with them local administrations, NGOs, and citizens in monitoring and managing the desired quality to the extent possible.

All waterbodies whose existing water quality is below acceptable levels should be identified on a priority basis to take remedial measures to restore their quality to acceptable levels within a stipulated period.

For waterbodies whose water is of acceptable quality, but below desirable quality, the limits of the relevant parameters should be identified and a quality restoration program should be implemented within a reasonable period.

WATER QUALITY MANAGEMENT AND NONPOINT SOURCE CONTROL

Harry X. Zhang
Parsons Corporation
Fairfax, Virginia

INTRODUCTION OF NONPOINT SOURCE CONTROL

Nonpoint source (NPS) pollution, unlike pollution from industrial and sewage treatment plants, comes from many diffuse sources. NPS pollution is caused by rainfall or snowmelt moving over and through the ground. As the runoff moves, it picks up and carries away natural and human-made pollutants, finally depositing them into lakes, rivers, wetlands, coastal waters, and even our underground sources of drinking water.

The United States has made tremendous advances in the past 25 years to clean up the aquatic environment by controlling pollution from industries and sewage treatment plants. Despite those successes, approximately 39% of assessed rivers and streams and 39% of assessed lakes are not safe for fish consumption, and 37% of its surveyed estuarine area is not safe for basic uses such as swimming or fishing. The estimates for nonattainment of swimming in surveyed rivers and lakes were 32% and 30%. Today, nonpoint source pollution is the leading remaining cause of water quality problems. It is the main reason that approximately 40% of our surveyed rivers, lakes, and estuaries are not clean enough for basic uses such as fishing or swimming (1,2).

Nonpoint source pollutants include (1) excess fertilizers, herbicides, and insecticides from agricultural lands and residential areas; (2) oil, grease, and toxic chemicals from urban runoff and energy production; (3) sediment from improperly managed construction sites, crop and forest lands, and eroding stream banks; (4) salt from irrigation practices and acid drainage from abandoned mines; (5) bacteria and nutrients from livestock, pet wastes, and faulty septic systems; and (6) atmospheric deposition and hydromodification (3). NPS pollution is widespread because it can occur any time activities disturb the land or water. Agriculture, forestry, grazing, septic systems, recreational boating, urban runoff, construction, physical changes to stream channels, and habitat degradation are potential sources of NPS pollution. Careless or uninformed household management also contributes to NPS pollution problems. The most common NPS pollutants are sediment and nutrients. These wash into waterbodies from agricultural land, small and medium-sized animal feeding operations, construction sites, and other areas of disturbance.

EXISTING REGULATIONS RELATED TO NONPOINT SOURCE CONTROL

During the last 10 years, the United States has made significant progress in addressing NPS pollution. At the federal level, recent NPS control programs include the Nonpoint Source Management Program established under Section 319 of the 1987 Clean Water Act (CWA) Amendments, and the Coastal Nonpoint Pollution Program established by the 1990 Coastal Zone Act Reauthorization Amendments (CZARA).

The most prominent of these was the Nonpoint Source Management Program set up by the 1987 amendments (4). Under this program, states were to identify waterways that would not be expected to meet ambient standards because of nonpoint sources. For these areas, states were to develop NPS programs that identified best management practices for the categories of nonpoint sources involved. Because EPA had little or no authority to regulate many nonpoint sources, it relied upon state and local governments to implement controls. The program provides grants to states, territories, and tribes to implement NPS pollution controls described in approved NPS pollution management programs. In the 1987 amendments to the Clean Water Act, toxic waste and pollution from NPS received more attention than they had previously, and stormwater discharges from urban areas and industrial facilities were included in the NPDES permit program.

Under Section 303(d) of the 1972 Clean Water Act, states, territories, and authorized tribes are required to develop lists of impaired waters. The law requires that these jurisdictions establish priority rankings for waters on the lists and develop total maximum daily loads (TMDLs) for these waters. In 1991, EPA issued the Guidance for Water Quality-Based Decisions: The TMDL Process (5). In 1992, the EPA clarified the goals of the Nonpoint Source Management Program and moved toward a watershed protection approach, which tailors NPS pollution control strategies to fit conditions in particular watersheds and gives state and local governments flexibility. In addition to engineering controls on nonpoint source pollution, management strategies might include outreach and education programs as well as government action to change land use patterns and activities that contribute to NPS pollution.

The 1987 amendments also created the National Estuary Program (NEP), which provides a basis for managing NPS pollution. The NEP relied on existing statutory authority at the state and municipal levels to accomplish program goals. Each National Estuary Program is charged to create and implement a Comprehensive Conservation and Management Plan (CCMP) that addresses all aspects of environmental protection for the estuary, including issues such as water quality, habitat, living resources, and land use.

Another regulation on nonpoint source pollution is the Coastal Zone Act Reauthorization Amendments Section 6217, which addresses coastal NPS pollution problems in coastal waters. This program is administered jointly with the National Oceanic and Atmospheric Administration (NOAA). In its program, a state or territory describes how it will implement nonpoint source pollution controls, known as management measures, that conform with those described in Guidance Specifying Management Measures for Sources of Nonpoint Pollution in Coastal Waters (6). The plans must also specify abatement measures that are “economically achievable” and yet “reflect the greatest degree of pollutant reduction achievable by applying the best available nonpoint source control practices, technologies, processes, siting criteria, operating methods, or other alternatives.

RELATED FEDERAL PROGRAMS FOR NONPOINT SOURCE CONTROL

There are several federal programs available to address nonpoint source control.

U.S. Environmental Protection Agency (EPA)

The EPA administers Section 319 of the Clean Water Act, also known as the Nonpoint Source Management Program. Under Section 319, states, territories, and tribes apply for and receive grants from the EPA to implement NPS pollution controls. The EPA administers other sections of the CWA to help states, territories, and tribes plan for and implement water pollution control programs, which can include measures for NPS control. These include Section 104(b)(3), Water Quality Cooperative Agreements; Section 104(g), Small Community Outreach; Section 106, Grants for Pollution Control Programs; Section 314, Clean Lakes Program; Section 320, National Estuary Program; and Section 604(b), Water Quality Management Planning. Together with the NOAA, EPA helps administer Section 6217 of the 1990 Coastal Zone Act Reauthorization Amendments, a program that tackles nonpoint source pollution affecting coastal waters (7).

National Oceanic and Atmospheric Administration (NOAA)

NOAA administers Section 306 of the Coastal Zone Management Act that provides funds for water pollution control projects, including NPS management activities, in states that have coastal zones. Together with the EPA, NOAA also helps administer Section 6217 of the CZARA.

U.S. Department of Agriculture (USDA)

The USDA administers incentive-based conservation programs through the Consolidated Farm Services Agency, the Natural Resources Conservation Service (NRCS), and the U.S. Forest Service to help control NPS pollution from agriculture and forestry.

U.S. Department of Transportation (DOT)/Federal Highway Administration (FHWA)

Under the Intermodal Surface Transportation Efficiency Act of 1991, the Federal Highway Administration developed erosion control guidelines for federally funded construction projects on roads, highways, and bridges.

U.S. Department of the Interior (DOI)

Within the U.S. Department of the Interior, the Bureau of Reclamation (BLR), the Bureau of Land Management (BLM), and the Fish and Wildlife Service (FWS) administer several programs to help states manage NPS pollution by providing technical assistance and financial support.

NONPOINT SOURCE CONTROL BY SOURCE CATEGORIES

NPS Pollution from Agriculture

The latest National Water Quality Inventory indicates that agriculture is the leading contributor to water quality impairments; it degraded 60% of the impaired river miles and half of the impaired lake acreage surveyed by states, territories, and tribes (1). Agricultural activities that cause NPS pollution include confined animal facilities, grazing, plowing, pesticide spraying, irrigation, fertilizing, planting, and harvesting. The major agricultural NPS pollutants that result from these activities are sediment, nutrients, pathogens, pesticides, and salts. Agricultural activities can also damage habitat and stream channels. Agricultural impacts on surface water and groundwater can be minimized by properly managing activities that can cause NPS pollution.

Managing Sedimentation. Excessive sedimentation clouds water, which reduces the amount of sunlight reaching aquatic plants; covers fish-spawning areas and food supplies; and clogs the gills of fish. In addition, other pollutants such as phosphorus, pathogens, and heavy metals are often attached to soil particles and wind up in the waterbodies with the sediment. Farmers and ranchers can reduce erosion and sedimentation by 20–90% by applying management measures to control the volume and flow rate of runoff water, keep the soil in place, and reduce soil transport.

Managing Nutrients. When nutrients such as phosphorus, nitrogen, and potassium in the form of fertilizers or manure are applied in excess of plant needs, they can wash into aquatic ecosystems where they can cause excessive plant growth, which reduces swimming and boating opportunities, creates a foul taste and odor in drinking water, and kills fish. In drinking water, high concentrations of nitrate can cause methemoglobinemia, a potentially fatal disease in infants also known as blue baby syndrome. Farmers can implement nutrient management plans that help maintain high yields and save money on the use of fertilizers while reducing NPS pollution.

Managing Confined Animal Facilities and Livestock Grazing. Confined animal and livestock areas are major sources of animal waste. Runoff from poorly managed facilities can carry pathogens (bacteria and viruses), nutrients, and oxygen-demanding substances that contaminate shellfishing areas and cause other major water quality problems. Groundwater can also be contaminated by seepage. Discharges can be limited by storing and managing facility wastewater and runoff using an appropriate waste management system.

Overgrazing exposes soils, increases erosion, encourages invasion by undesirable plants, destroys fish habitat, and reduces the filtration of sediment necessary for building streambanks, wet meadows, and floodplains. To reduce the impacts of grazing on water quality, farmers and ranchers can adjust grazing intensity, keep livestock out of sensitive areas, provide alternative sources of water and shade, and revegetate rangeland and pastureland.

Managing Irrigation. Irrigation water is applied to supplement natural precipitation or to protect crops from freezing or wilting. Inefficient irrigation can cause water quality problems. In arid areas, for example, where rainwater does not carry residues deep into the soil, excessive irrigation can concentrate pesticides, nutrients, disease-carrying microorganisms, and salts—all of which impact water quality—in the top layer of soil. Farmers can reduce NPS pollution from irrigation by increasing water use efficiency.

Managing Pesticides. Pesticides, herbicides, and fungicides are used to kill pests and control the growth of weeds and fungus. These chemicals can enter and contaminate water through direct application, runoff, wind transport, and atmospheric deposition. To reduce NPS contamination from pesticides, people can apply integrated pest management (IPM) techniques based on the specific soils, climate, pest history, and crop for a particular field.

NPS Pollution from Urban Runoff

The most recent National Water Quality Inventory reports that runoff from urban areas is the leading source of impairments to surveyed estuaries and the third largest source of water quality impairments to surveyed lakes. Runoff from rapidly growing urban areas will continue to degrade coastal waters.

Increased Runoff. Nonporous urban landscapes such as roads, bridges, parking lots, and buildings do not let runoff slowly percolate into the ground. Receiving waters are often adversely affected by urban runoff due to alterations in hydraulic characteristics of streams receiving runoff such as higher peak flow rates, increased frequency and duration of high flows and possible downstream flooding, and reduced baseflow levels.

Increased Pollutant Loads. Urbanization also increases the variety and amount of pollutants transported to receiving waters. Sediment from development and new construction; oil, grease, and toxic chemicals from automobiles; nutrients and pesticides from turf management and gardening; viruses and bacteria from failing septic systems; road salts; and heavy metals are examples of pollutants generated in urban areas (8). Sediments and solids constitute the largest volume of pollutant loads to receiving waters in urban areas.

Measures to Manage Urban Runoff

To protect surface water and groundwater quality, urban development and household activities must be guided by plans that limit runoff and reduce pollutant loadings.

Managing Urban Runoff from New Development. New developments should attempt to maintain the volume of runoff at predevelopment levels by using structural BMP controls and pollution prevention strategies. Plans for managing runoff, sediment, toxics, and nutrients can establish guidelines to help achieve both goals. Management plans are designed to protect sensitive ecological areas, minimize land disturbances, and retain natural drainage and vegetation.

Managing Urban Runoff from Existing Development. Existing urban areas can target their urban runoff control projects to make them more economical. Runoff management plans for existing areas can first identify priority pollutant reduction opportunities, then protect natural areas that help control runoff, and finally begin ecological restoration and retrofit activities to clean up degraded waterbodies.

Managing Runoff for On-Site Disposal Systems. The control of nutrient and pathogen loadings to surface waters can begin with the proper design, installation, and operation of on-site disposal systems. These septic systems should be situated away from open waters and sensitive resources such as wetlands and floodplains and should also be inspected, pumped out, and repaired at regular time intervals.

NPS Pollution from Forestry

Sources of NPS pollution from forestry activities include removal of streamside vegetation, road construction and use, timber harvesting, and mechanical preparation for planting trees. Road construction and road use are the primary sources of NPS pollution on forested lands; they contribute up to 90% of the total sediment from forestry operations. Harvesting trees in the area beside a stream can affect water quality by reducing the streambank shading that regulates water temperature and by removing vegetation that stabilizes the streambanks. These changes can harm aquatic life by limiting sources of food, shade, and shelter (7).

Establishing Streamside Management Areas. Plans for buffer strips or riparian zones often restrict forestry activities in vegetated areas near streams, thereby establishing special Streamside Management Areas (SMAs). Vegetation in the SMA stabilizes streambanks, reduces runoff and nutrient levels in runoff, and traps sediment generated from upslope activities before it reaches surface waters. SMA vegetation moderates water temperature by shading surface water and provides habitat for aquatic life.

Managing Road Construction and Maintenance. Good road location and design can greatly reduce the transport of sediment to waterbodies. Roads should follow the natural contours of the land and be located away from steep gradients, landslide-prone areas, and areas of poor drainage. Proper road maintenance and closure of unneeded roads can help reduce NPS impacts from erosion over the long term.

Managing Timber Harvesting. Most detrimental effects of harvesting are related to access and movement of vehicles and machinery and dragging and loading trees or logs. These effects include soil disturbance, soil compaction, and direct disturbance of stream channels. Poor harvesting and transport techniques can increase sediment production ten to twenty times and disturb as much as 40% of the soil surface.

Other NPS sources include those induced by hydromodification/habitat alteration, marinas/boating, and roads/highways/bridges, which will not be discussed in detail here.

EMPLOYING BEST MANAGEMENT PRACTICES

Public and private groups have developed and used pollution prevention and pollution reduction initiatives and NPS pollution controls, known as management measures, to clean up our water efficiently. The use of best management practices (BMPs) in addressing runoff problems was frequently identified as a NPS control measure, although additional research and field studies on the performance of BMPs were also needed. A wide variety of BMPs, both structural and nonstructural, are available to address urban storm water runoff and discharges (9–12). In summary, many BMPs are used primarily for water quantity control (i.e., to prevent flooding), although they may provide ancillary water quality benefits. Some BMP types have been analyzed for performance in terms of site-specific pollutant removal, but not extensively enough to allow for generalizations. It is still difficult to develop comparisons of cost-effectiveness for various BMP types. Other BMP types, particularly nonstructural and those that do not have discrete inflow or outflow points, are difficult to monitor. In addition, only limited cost studies have been conducted for stormwater BMPs. Therefore, the pollutant removal performance of certain BMP types is essentially undocumented. The role of chemical pollutant monitoring versus receiving stream biological monitoring in evaluating BMP performance on a watershed scale is not well documented or needs further research.

One particular new design that is becoming popular and is widely used in recent years is low impact development (LID). LID is an innovative stormwater management approach using a basic principle that is modeled after nature; managing rainfall at the source using uniformly distributed decentralized microscale controls. Its goal is to mimic a site's predevelopment hydrology by using design techniques that infiltrate, filter, store, evaporate, and detain runoff close to its source (13). Another important concept in controlling urban NPS pollution is smart growth. Although it recognizes the many benefits of growth, smart growth practice invests time, attention, and resources in restoring community and vitality to existing cities and older suburbs. The purposes of smart growth and low impact development practices are to promote a number of practices that can lessen the environmental impacts of development, to preserve ecosystem functions as well as specific critical areas, and to design developments to maintain predevelopment conditions (14).

WATER QUALITY MANAGEMENT AND TMDL PROGRAMS

Water quality management has become increasingly more complicated. Problems such as toxic contaminants, sediments, nutrients, and habitat alteration result from a variety of point and nonpoint sources. The Water Quality Planning and Management Regulation (40 CFR 130) links a number of Clean Water Act sections, including Section 303(d), to form the water quality approach to protecting and cleaning up the nation's waters. Previous practices for implementing 303(d) have focused primarily on point sources and wasteload allocations (WLA). In recent years, nonpoint source contributions to water quality problems have become better understood, and it is now clear that EPA and state implementation of 303(d) must encompass nonpoint source pollution problems and seek to address problems occurring over large geographic areas (5).

The water quality approach emphasizes the overall quality of water within a waterbody and provides a mechanism through which the amount of pollution entering a waterbody is controlled by the intrinsic conditions of that body of water and the standards set to protect it. Section 303(d) of the Clean Water Act addresses those impaired waters that are not "fishable, swimmable" by requiring states to identify the waters and to develop a total maximum daily load (TMDL). EPA's regulations establish a TMDL as the sum of point source load allocation (WLA) plus nonpoint source allocation (LA) plus a margin of safety (MOS). TMDLs can be expressed in terms of either mass per time, toxicity, or other appropriate measures that relate to a state's water quality standard. If BMPs or other nonpoint source pollution control actions make more stringent load allocations practicable, then WLAs can be made less stringent. Thus, the TMDL process provides for nonpoint source control trade-offs.

The TMDL process is viewed as an effective approach to watershed protection and a water quality management tool for both point and nonpoint source pollution control on a watershed scale. A watershed protection approach is a strategy for effectively protecting and restoring aquatic ecosystems and protecting human health. This strategy has as its premise that many water quality and ecosystem problems are best solved at the watershed level rather than at the individual waterbody or discharger level (15). The current regulations that address water quality problems have tended to focus on particular sources, pollutants, or water uses and have not resulted in an integrated environmental management approach. Consequently, significant gaps exist in our efforts to protect watersheds from the cumulative impacts of a multitude of activities. Existing water pollution prevention and control programs are, however, excellent foundations on which to build a watershed approach. Increasingly, state and tribal water resource professionals are turning to watershed management to achieve greater results from their programs. Major features of a watershed protection approach include targeting priority problems, promoting a high level of stakeholder involvement, integrated solutions that use the expertise and authority of multiple agencies, and measuring success through monitoring and other data gathering.

TMDL IMPLEMENTATION AND NPS CONTROL

NPDES permits are usually sought for point sources, but nonpoint source controls may be established by implementing BMPs, so that surface water quality objectives are met. These controls should be based on load allocations developed by using the TMDL process. When establishing permits for point sources on the watershed scale, the record should show that in the case of any credit for future nonpoint source reductions, there is reasonable assurance that nonpoint source controls will be implemented and maintained or that nonpoint source reductions are demonstrated through an effective monitoring program.

However, it is difficult to ensure, *a priori*, that implementing nonpoint source controls will achieve expected load reductions. Nonpoint source control measures may fail to achieve projected pollution or chemical load reductions due to inadequate selection of BMPs, inadequate design or implementation, or lack of full participation by all contributing sources of nonpoint pollution. States should describe nonpoint source load reductions and establish a procedure for reviewing and revising BMPs in TMDL documentation. The TMDL implementation plan must contain a description of best management practices or other management measures. The implementation plan may deal with sources on a watershed basis, as long as the scale of the implementation plan is consistent with the geographic scale for which the TMDL pollutant load allocations are established. This makes water quality monitoring an important component during TMDL implementation.

The achievement of nonpoint source load reductions is a complex challenge. The variability in nonpoint source loadings due to hydrologic variability can often make it difficult to discern short-term trends. Therefore, achievement of water quality standards is tracked using selected long-term milestones and measures. Tracking the implementation of management actions over time will provide valuable information. The tracking of implementation will assist in determining the success of the load allocation, the adequacy of funding and resources, the potential for water quality improvement, and the need for corrective actions. For areas that have predominantly nonpoint source controls, the use of tracking information can support demonstration of progress in the absence of clear benefit through water quality monitoring (16).

SUMMARY AND CONCLUSION

Today, nonpoint source pollution remains the largest source of water quality problems in the United States. It is the main reason that approximately 40% of our surveyed rivers, lakes, and estuaries are not clean enough for basic uses such as fishing or swimming (1,17). This remaining problem impacting water quality is not easily remedied because NPS pollution comes from pipes and also from diffuse sources such as agricultural and forestry operations, urban areas, and construction sites.

Best management practices or other innovative management measures such as low impact development are frequently identified to address the NPS problem. However, additional research and field studies on the long-term performance of BMPs, especially on the watershed scale, are much needed.

The past decade has seen a shift toward an emphasis from "end-of-pipe" treatment for point source pollution to the watershed approach, which encourages a holistic take on identifying problems and implementing the integrated solutions that are needed to overcome multiple causes of water quality impairment, including NPS. In addition to the watershed approach, the water quality based TMDL program has evolved as an effective tool for addressing more complicated water quality problems (e.g., for waters that remain polluted after the application of technology-based standards). TMDL implementation and NPS control on a watershed basis are two interconnected components for successful water quality management.

BIBLIOGRAPHY

1. U.S. EPA. (1996). *Nonpoint Source Pollution: The Nation's Largest Water Quality Problem*. EPA841-F-96-004A.
2. U.S. EPA. (2002). *National Water Quality Inventory: 2000 Report*. EPA-841-R-02-001.
3. U.S. EPA. (1994). *What is Nonpoint Source (NPS) Pollution? Questions and Answers*. Report EPA-841-F-94-005, http://www.epa.gov/owow/nps/qa.html
4. Ortolano, L. (1997). *Environmental Regulation and Impact Assessment*. John Wiley & Sons, Hoboken, NJ.
5. U.S. EPA. (1991). *Guidance for Water Quality-Based Decisions: The TMDL Process*. EPA 440/4-91-001, Washington, DC.
6. U.S. EPA. (1993). *Guidance Specifying Management Measures for Sources of Nonpoint Pollution in Coastal Waters*. Report EPA-840-B-93-001c, Washington, DC.
7. U.S. EPA. (1996). *The Nonpoint Source Management Program*. EPA841-F-96-004D.
8. U.S. EPA. (1996). *Managing Urban Runoff*. EPA841-F-96-004G.
9. American Society of Civil Engineers (ASCE). (2001). National Stormwater Best Management Practices (BMPs) Database, http://www.bmpdatabase.org/.
10. American Society of Civil Engineers (ASCE) and Water Environment Federation (WEF). (1998). *Urban Runoff Quality Management*. ASCE, Alexandria, VA and WEF, Reston, VA.
11. U.S. EPA. (1999). *Preliminary Data Summary of Urban Stormwater Best Management Practices*. Report EPA-821-R-99-012.
12. U.S. EPA. (2002). *Urban Stormwater BMP Performance Monitoring*. Report EPA-821-B-02-001.
13. Low Impact Development Center, Inc. (LID). (2002). <http://www.lid-stormwater.net/>.
14. U.S. EPA. (2002c). Environmental Protection and Smart Growth, <http://www.epa.gov/smartgrowth/>.
15. U.S. EPA. (1996d). *Why Watershed Approach*. Report EPA800-F-96-001.
16. U.S. EPA. (1999). *Draft Guidance for Water Quality-based Decisions: The TMDL Process (Second Edition)*. Report EPA-841-D-99-001.
17. U.S. EPA. (2001a). *National Management Measures to Control Nonpoint Source Pollution from Agriculture*. Washington, DC.

READING LIST

Novotny, V. and Olem, H. (1994). *Water Quality: Prevention, Identification, and Management of Diffuse Pollution*. Van Nostrand Reinhold, New York.

U.S. EPA. (2001b). *National Management Measures to Control Nonpoint Source Pollution from Forestry (Draft)*. Washington, DC.

U.S. EPA. (2002d). Total Maximum Daily Loads, <http://www.epa.gov/owow/tmdl>.

U.S. EPA. (2002e). *National Management Measures to Control Nonpoint Source Pollution from Urban Areas*. Report EPA-842-B-02-003, Washington, DC.

Wanielista, M., Kersten, R., and Eaglin, R. (1997). *Hydrology: Water Quantity and Quality Control*. John Wiley & Sons, Hoboken, NJ.

WATER QUALITY MANAGEMENT IN AN URBAN LANDSCAPE

Russell N. Clayshulte
Aurora, Colorado

Urban landscape broadly defines a built environment influenced by different combinations of land uses. Residential and commercial development, industrial areas, parks, open space, and roadways are examples of urban land uses. The construction of wastewater, water, and stormwater pipelines, and power lines, generally called utility infrastructure, also change the natural environment and help shape the urban landscape. Land use patterns strongly influence surface water quality and to a lesser extent groundwater. Land use choices are interactive parts of water quality management, restoration, or enhancement programs. Urban land use decisions must consider water quality management strategies and goals. Conversely, water quality must be considered in zoning and plotting processes used by local governments. A goal in shaping urban areas is to produce a sustainable urban landscape that incorporates reasonable and effective water quality strategies to restore, protect, or enhance the urban environment.

Scale affects how water quality is managed and often who is responsible for this management. Water quality management targets overall health and the quality of water and environmental resources. Management also provides direction in fixing specific water quality problems in targeted waterbodies. Consequently, water quality management in urban settings occurs on two basic levels. Macroscale landscapes include all combined urban land uses for a single city or a combination of cities and towns. Water quality management on the macroscale is often linked to watershed restoration or protection efforts. Microscale landscapes can include smaller urban watersheds, drainages, or other limited geographic areas where site-specific water quality management is necessary to control targeted urban-caused pollution. Macroscale watersheds are often measured in square miles (square kilometers); microscale landscapes are identified in acres (hectares).

Water quality management programs in urban landscapes are divided into three general source categories: point sources, nonpoint sources, and stormwater runoff. Point sources are discrete discharges that go from a pipe source (e.g., wastewater treatment plant discharge) back into the environment. Point sources can be permitted to meet very specific discharge limits under local, state, and federal regulations. Nonpoint sources are a type of diffuse pollution without a single point of origin or not introduced into the environment from a single pipe. Common nonpoint sources are agriculture, forestry, urban, mining, construction, hydromodification (e.g., dams, channels, and ditches), land disposal, and saltwater intrusion. Stormwater runoff is a type of nonpoint source runoff usually associated with urban landscapes, which is subject to permitting requirements for medium and large cities. Stormwater and other nonpoint sources in urban landscapes are collectively called urban runoff.

Urban runoff can flow across residential and urban streets, roofs, lawns, open space, hard (impervious) surfaces, and other areas. Urban runoff occurs under both wet and dry weather conditions. Dry-weather runoff from irrigation and leaky pipes contributes various amounts of polluted runoff. Urban runoff carries many different types of chemicals, sediments, debris, and pathogens. An excessive amount of urban pollution impairs water quality in wetlands, streams, rivers, lakes, reservoirs, or groundwater (receiving waters). Construction or other development activities introduce large amounts of sediments and other pollutants into receiving waters. Sediment and solids contribute the largest volume of pollutant load to urban receiving waters. Increased loads can harm fish and wildlife populations, kill native vegetation, reduce or limit recreational uses, and contaminate drinking water.

The terms pollution and pollutants are different. This difference is very important from a regulatory perspective. The 1972 federal Clean Water Act (Public Law 92-500) defines the term "pollution" to mean the "man-made or man-induced alteration of the chemical, physical, biological, and radiological integrity of water." Generally, pollution is the presence of a substance in the environment that because of its chemical composition or quantity prevents the functioning of natural processes and produces undesirable environmental and/or health effects. The act defines the term "pollutant" to mean any substance introduced into the environment that adversely affects the usefulness of a resource or the health of humans, animals, or ecosystems. Common types of pollutants include dredged spoil, solid waste, incinerator residue, sewage, garbage, sewage sludge, munitions, chemical wastes, biological materials, radioactive materials, heat, wrecked or discarded equipment, rock, sand, cellar dirt, and industrial, municipal, and agricultural waste discharged into water. Specific pollutants can be targeted and regulated through a number of local, state, and federal processes. Harmful impacts from pollutants and pollution are often linked to risk studies with measurable impact levels.

Pollution management in urban landscapes is a relatively new requirement for cities and towns. Water quality management, as a national or even global problem, became a higher priority public concern in the late 1940s. The first legislation to control water pollution was the Water Pollution Control Act of 1948 (Public Law 845, 80th Congress). In 1972, the United States Congress enacted the first comprehensive national clean water legislation in response to growing public concern for serious and widespread water pollution. The federal Clean Water Act (Public Law 92-500; as amended in 1977, Public Law 95–217 and as amended in 1987, Public Law 100-4) is the primary federal law in the United States that protects national waters, including wetlands, lakes, rivers, aquifers, and coastal areas. The goal of the federal Clean Water Act is to make waters of the United States "fishable and swimmable, where applicable." Meeting the goals of the act requires an extensive and integrated approach to water pollution abatement.

Federal regulations require all states to maintain, improve, and in some cases restore the quality of their water resources. This federal law significantly altered water resource management in urban settings. Federal regulations provide a clear path to clean water and an effective national water program. In 1972, only a third of U.S. waters were safe for fishing and swimming. Wetlands losses were estimated in excess of 460,000 acres annually with an alarming loss of urban wetlands. The United States Environmental Protection Agency reported almost 100% wetland loss for many larger cities. Agricultural and construction runoff caused erosion of over 2.25 billion tons of soil per year. The deposition of large amounts of phosphorus and nitrogen, referred to as nutrients, into receiving waters was attributed, in a large part, to urban centers. This nutrient enrichment caused many ponds, lakes, and reservoirs to experience excessive algal growth, resulting lost fisheries, and water supply contamination. Sewage treatment plants or point sources served about 85 million people. A large amount of human wastes were flowing into receiving waters with little or no treatment.

Water quality in the United States was characterized in a 1998 report to the American Congress as part of a

Table 1. Leading Sources of Water Quality Impairment Related to Urban Landscapes for Rivers, Lakes, and Estuaries[a]

	Rivers and Streams	Lakes, Ponds, Rivers	Estuaries
Pollutants	Siltation (38%)[b]	Nutrients (44%)	Pathogens (47%)
	Pathogens (36%)	Metals (27%)	Organic enrichment (42%)
	Nutrients (28%)	Siltation (15%)	Metals (27%)
Sources[c]	Hydromodification (20%)	Hydromodification (20%)	Municipal point sources (28%)
	Urban runoff/storm sewers (12%)	Urban runoff/storm sewers (12%)	Atmospheric deposition (23%)
			Urban runoff/storm sewers (23%)

[a] Adapted from Reference 1.

[b] Values in parentheses represent the percentage of surveyed river miles, lake acres, or estuary square miles that are classified as impaired.

[c] Excluding unknown, natural, and "other" sources.

national water quality inventory (1). Table 1 lists leading sources of water quality impairment in urban areas. This report showed that about 40% of receiving waters still did not meet the fishing and swimming goals of the federal Clean Water Act. In fact more than 291,000 miles of assessed rivers and streams did not meet state adopted water quality standards. Sediment and siltation is one of the most common pollutants from urban development and construction. Other major pollutants are pathogens, nutrients, and metals. Runoff from agricultural lands and urban areas is a major source of these pollutants. Across all types of waterbodies, states, territories, tribes, and other jurisdictions have shown that poor water quality affects drinking water, aquatic life, fish consumption, swimming, and other recreational uses.

However, the national water quality picture has improved since the passage of the Clean Water Act. By 2001, two-thirds of the nation's waters were safe for fishing and swimming. The rate of annual wetlands losses is reduced to about 70,000–90,000 acres. The amount of soil lost due to agricultural runoff and urban construction activities has been cut by over 1 billion tons annually. Phosphorus and nitrogen levels in receiving waters are lower. Modern wastewater treatment facilities serve over 173 million people. Aggressive water quality management programs in urban areas have helped make these differences.

States adopt use classifications and numerical water quality standards for receiving waters in urban watersheds. Within urban watersheds, waters are divided into individual stream segments for classification and standard setting. Site-specific water quality classifications are used to protect any existing or intended use of the target water. Typical use classification categories include recreation, agriculture, aquatic life, domestic water supply, or wetlands. These use classification categories are referred to as "beneficial uses." Beneficial use classification types vary from state to state and, in some cases, even within a state. A water quality standard consists of four basic elements: (1) *designated uses* of the water body (e.g., recreation, water supply, aquatic life, agriculture), (2) *water quality criteria* to protect designated uses (numerical pollutant concentrations and narrative requirements), (3) an *antidegradation policy* to maintain and protect existing uses and high quality waters, and (4) *general policies* addressing implementation issues (e.g., low flows, variances, mixing zones).

Regulatory and enforcement programs provide the basic framework for water quality management in an urban environment. These programs generally reside at the state level. Although the federal Clean Water Act and other federal laws provide guidance for regulatory programs, states have the flexibility to establish programs to fit conditions within the state. Often standards and classifications of surface and groundwater drive management options. Urban watershed efforts ensure that waters within an urban landscape meet underlying standards and beneficial uses. Enforcement actions and lawsuits that target the amount of allowable pollutant load discharged into receiving waters (wasteload litigation) have triggered more water quality management programs in urban centers.

An urban watershed protection approach is an integrated, holistic strategy to protect or attain the desired beneficial uses of waters within an urban area. The approach is more effective than isolated efforts under existing programs that do not consider the watershed as a whole. Nonpoint source or stormwater runoff control on a watershed level can have a significant impact on the protection of beneficial uses. A watershed protection approach addresses point source discharges along with nonpoint source and stormwater pollutant loads. An urban watershed approach also considers other human activities that may affect the uses and quality of water resources.

Development patterns change the natural landscape and replace it with hard or impervious surfaces (i.e., concrete, asphalt, or hard-packed dirt), which in turn change the hydrologic or flow patterns from surfaces and into receiving waters. Urbanization causes increased stormwater runoff, which is more intense, has a higher volume, and has less residual runoff than natural runoff conditions. A typical city block characterized by pavement and rooftops generates over nine times more runoff than a woodland area of equal size. Less water gets into underlying groundwater and more water enters receiving waters as runoff flow. Aquifers under urban areas are called "starved" because water recharge is minimized. Natural runoff from an area without any kind of disturbance is called background or sometimes baseflow. Urbanization that alters runoff patterns can decrease baseflow and increase runoff flow. In most urban landscapes, lawn irrigation or other forms of dry-weather discharge can greatly increase runoff flow.

Urban hydrologic changes or major features like roads and parking lots cause many types of water quality problems such as nutrient enrichment, chemical pollutants, turbidity or sedimentation, biological pathogen introduction, toxic conditions, or increased temperature. Transportation systems have the highest pollutant concentrations compared with other urban land uses. Common types of transportation-related pollutants include gasoline, exhaust, motor oil and grease, antifreeze, undercoating, brake linings, rubber, asphalt, concrete, and diesel fuel. Landscaping in urban areas can contribute large amounts of nutrients from fertilizer application and pesticide or herbicide residues. Grass clippings and leaves in stormwater systems are oxygen-demanding substances, which can cause low oxygen conditions in receiving waters and stress aquatic life.

Widespread water quality data from urban areas across the country show that water quality problems are directly caused by increased watershed imperviousness. Ecological stress becomes apparent in an urban watershed when the amount of impervious cover reaches 10–20%. Typically, residential areas have impervious cover ranging from 25% to 60%; commercial areas can reach 100%. Increased pollutant loading is directly linked to the amount of impervious surface. Urban landscapes with greater than 50% impervious surface have highly stressed ecosystems. Low impact development is a new urban management concept designed to reduce the amount of impervious surface and promote more natural runoff patterns and processes.

Changes in flow regimes also cause increased frequency of flooding and peak flow volumes in receiving waters. These floods cause loss of aquatic or riparian habitat and changes in stream physical characteristics (channel width and depth). Stormwater management requires urban areas to develop stormwater drainage systems to avoid flood damage. These systems accumulate runoff from storms and send these flows away from the urban area. Flood control structures and modifications to natural drainage systems decrease natural stormwater purification functions. Runoff quality has not been well addressed in the design criteria of flood control structures. In some cases, stormwater systems are used as a convenient method of waste disposal.

Water quality monitoring for urban runoff facilitates the critical element of relating the physical, chemical, and biological characteristics of receiving waters to land use characteristics. Without current information on water quality and pollutant sources, the effects of land-based activities on water quality cannot be fully assessed, effective management and remediation programs cannot be implemented, and program success cannot be evaluated. Monitoring programs in the past have been the responsibility of state and federal agencies; recent local watershed management has shifted monitoring to the watershed or microscale environment. Some common types of water and environmental quality indicator parameters monitored in urban landscapes are shown in Table 2.

Not all urban runoff is polluted. The severity of the water quality problem in receiving water is directly related to the beneficial uses assigned to the waterbody by the regulating agency. Urban runoff quality is assessed against established standards and classifications. Many stream segments in urban areas have site-specific water quality standards and classifications. Consequently, streams or rivers flowing through urban centers with multiple segments can have differing sets of standards and classifications. Water quality parameters identified as pollutants in one segment may not be pollutants in the next contiguous downstream segment. This factor makes pollutant characterization, management, and control very difficult in urban areas.

There are many implementation methods or tools used in watershed planning and management. The method used for a specific watershed program depends on the problem and available resources. Best management practices (BMPs) are water quality management tools. BMPs used in urban water quality programs fall into two categories: (1) erosion control that prevents discharge

Table 2. Common Indicator Parameters of Urban Runoff and Landscape Alteration

	Common Indicator Parameters
Urban Runoff Pollutants	Nutrients (total phosphorus, *o*-phosphorus, total nitrogen, nitrite, nitrate, ammonia); chemical oxygen demand; total organic carbon; salts (deicer products like sodium chlorides, calcium chlorides); biochemical oxygen demand (caused by grass clippings, animal wastes, leaves, or other carbon compounds); pathogens (total coliform bacteria, fecal coliform bacteria, *E. coli* bacteria, viruses, or protozoa); solids (total suspended solids, total dissolved solids, settleable solids); pH; temperature; aluminum, arsenic, boron, barium, cadmium, chromium, chromium(VI), chlorine, copper, cyanide, iron, lead, manganese, magnesium, mercury, nickel, radium, selenium, silver, sulfate, zinc, uranium; phenols, BTEX (benzene and xylene), organic (oil, gasoline or grease), and inorganic compounds; and herbicides and other pesticides
Urban Hydrologic or Habitat Alteration	Changes in nitrogen and phosphorus concentrations (nutrient enrichment); increased chlorophyll concentrations in urban ponds, lakes, and reservoirs (measure of algal production); bioaccumulation of toxins in plant and animals; changes in species biodiversity; deposition of sediment (sand, soil, or silt); decreases in dissolved oxygen; increased runoff; increased flow velocity; increased or decreased temperatures in riparian corridors; increased sediment and bed-load transport; streambank and coastal erosion and deposition; hydromodification (dams and channels); increased impervious or paved surfaces; and loss of riparian habitat

of pollutants or provides improved water quality in runoff from construction sites or development areas and (2) urban stormwater practices intended to reduce loads in the built urban environment. Similar best management practices are applicable to both stormwater runoff in urban areas and construction site runoff. Many urban watershed programs rely on BMPs to help solve pollution and/or pollutant problems. BMPs include structural and nonstructural methods, measures, or practices that help prevent, reduce, or mitigate adverse water quality problems caused by urbanization.

Structural BMPs are constructed to treat runoff passively before it enters receiving waters. Such BMPs used on a construction or development site can be either temporary or permanent and are designed to reduce sediment load or other runoff waste products for the life of the project. Additionally, these practices should protect aquatic or riparian environments. Nonstructural BMPs include prevention and source control that minimize or eliminate a problem before it occurs. Source control BMPs are sometimes referred to as “good housekeeping” measures because a clean site produces fewer pollutants than a dirty site. Site planning and design of BMPs may, in and of itself, be considered a nonstructural BMP.

Even as treatment facilities have improved, water quality goals have become more difficult and costly to meet. The physical, biological, and ecological characterization of urban water resources intensified after the passage of the 1972 Clean Water Act, and substantial efforts are needed to resolve problems and find workable solutions. A locally defined *balanced ecological community* achievable through water quality protection and water resource management is often not a community goal because of resource limitations. Achieving urban water quality goals requires a balance between the natural environment and designated resource uses. Effective and cost-efficient water quality management and supply require an integrated resource management program.

Maintenance, improvement, and restoration of water resources at a specific, watershed or regional scale are issues of concern to local governments, special districts, state agencies, and federal agencies. Institutional responsibilities vary among these entities in the water quality management system. Many states have approached water quality planning and management through state, regional, or locally linked programs using management and operating agencies. In most states, local management agencies are responsible for implementing state water quality management plans. Local agencies can decide on the need for and specific characteristics of a wastewater treatment process and then implement activities to meet specified state and federal goals.

The protection of surface and groundwater resources from urban growth and development can best be managed by plans that limit runoff and reduce pollutant loads. Local communities can use management plans to address water quality problems on a site-specific (e.g., stream segment or drainage basin) and a larger watershed level. Management planning in the United States and in many places around the world has shifted from a governmental process to a local effort. Consequently, hundreds of new water quality management agencies, particularly in urban areas, have begun grassroots efforts to protect and enhance local water resources.

Local management agencies may be individual municipal governments, watershed associations and authorities, general-purpose governments that hold a state or federal discharge permit, or other special districts responsible for planning and approving permitted facilities. Local governments or affiliated agencies can enter into agreements to form watershed associations or authorities with a single management agency. A municipality can be defined as any regional commission, county, metropolitan district offering sanitation service, sanitation district, water and sanitation district, water conservancy district, metropolitan sewage disposal district, service authority, city and county, city, town, watershed association or authority, Indian tribe or authorized Indian tribal organization, or any two or more of them that act jointly in a domestic wastewater treatment works. Water quality management in urban landscapes is an evolving and dynamic effort of pollution and pollutant abatement involving a growing number of groups, agencies, and governments. Land use decision makers must consider growth and development impacts on urban water resources and identify offsetting mitigative practices.

BIBLIOGRAPHY

1. United States Environmental Protection Agency. (2000). *Water Quality Conditions in the U.S.—A profile from the 1998 National Water Quality Inventory Report to Congress*, EPA 841-F-00-006. Environmental Protection Agency, Office of Water.

WATER QUALITY MANAGEMENT IN THE U.S.: HISTORY OF WATER REGULATION

Absar Alum
Arizona State University
Tempe, Arizona

Water is the true lifeline of our planet because of the fundamental role it plays in all kinds of biological activities and its catalytic contribution/functions in major chemical and geologic processes. These fundamental roles make water the single most important factor determining the course of ecological, social, and economic changes on the earth. The sustained growth of any human society is intricately associated with the quality of freshwater resources. The Industrial Revolution in the early years of the last century resulted in much ecological pollution. Biological and chemical contamination of freshwater resources is the greatest threat to sustainability. Water quality management strategies used to prevent progression in water pollution are based on a decision-making process. Mandatory laws have traditionally been proved to be the most effective means to achieve the goals of any society. Historically, human societies have relied on some types of cultural norms, social standards, and ethical, cultural, or religious guidelines to maintain the integrity of water resources.

The Industrial Revolution of the early twentieth century and the environmental pollution quagmire resulting as its aftermath have acted as catalysts of change in water quality management. As a result, mandatory regulations have become a preferred option as opposed to voluntary guidelines to ensure proper water quality.

One of the hallmarks of the last century is increased public awareness of water quality, which has been augmented by plummeting water quality and steadily increasing population. Though water has historically been considered a commodity free for all, in the recent past, water quality and quantity have emerged as the bases of major conflicts on local, regional, national, and international levels. Under the new paradigm, water is perceived as a marketable commodity. Like any marketable commodity, the quality of water is dictated by the type of usage and is valued based on market laws of supply and demand. In the United States, it is estimated that one acre-foot of water used in agriculture and industry would generate revenue of $400 and $400,000, respectively. In the United States, the water quality management program has evolved over several decades. Based on the area of use and type of water, the history of water quality management can be divided into drinking water and wastewater regulations. Early legislation focused on ensuring the appropriate quality of drinking water, whereas the emergence of wastewater management is relatively new. Developing and implementing water quality standards and criteria achieve the appropriate quality of finished drinking water and reclaimed wastewater.

HISTORY OF DRINKING WATER REGULATIONS

The plight of public health during the nineteenth century acted as a driving force to understand waterborne disease outbreaks. By late nineteenth century, scientists established a link between gastroenteritis and drinking water contaminants, which resulted in realization of the need for water quality management. Therefore, the drinking water industry became the first sector with elaborate water quality assurance and management plans; however, there was no regulatory control over this industry. During the early 1900s, most drinking water plants built in the United States were driven by aesthetic need to reduce turbidity that resulted in an incidental reduction in disease causing agents. Slow sand filtration was commonly used by cities (such as Philadelphia) to reduce turbidity. In 1908, Jersey City (New Jersey) had the first water system to use chlorine as primary disinfectant. Other disinfectants (such as ozone) were not employed in the United States until several decades later.

The history of federal regulations for water quality goes back to the late nineteenth century. The River and Harbor Act of 1886 was the first federal regulation to address water quality. In 1893, the Interstate Quarantine Act, enacted by Congress, empowered the United States Public Health Service (PHS) to establish drinking water standards. The objective of the Interstate Quarantine Act of 1893 was "to prevent the introduction, transmission and spread of communicable diseases." The River and Harbor Act of 1886, recodified in 1899, empowered the Corps of Engineers to issue discharge permits.

In 1914, the U.S. Public Health Service (PHS) introduced the first drinking water quality standards. The standards for bacteriological (only contagious) quality of drinking water were applied only to interstate carriers like ships and trains. These standards were revised in 1925. The early 1940s was the beginning of the "Chemical Revolution" with the rapid introduction of synthetic and natural chemicals in the industrial, agricultural, and social sectors. Soon, it was realized that these chemicals were contaminating freshwater sources. Drinking water quality standards were revised again in 1946, and some new chemical standards were introduced.

The Federal Water Pollution Control Act of 1948 (FWPCA) was the first legislation to address the issue comprehensively. Principles of limited federal enforcement and limited federal financial assistance were adopted to develop state and federal cooperative programs. These principles provided the framework for the Federal Water Pollution Control Act of 1956 and the Water Quality Act of 1965. In 1962, the PHS introduced its last revision of drinking water standards. Although the emphasis was on microbial diseases, 28 chemicals were also regulated. These revisions included about 20 standards and, for the first time, recommended the use of qualified personnel. Some of the important recommendations were use of water from protected sources, control of pollution sources, and chlorination of water drawn from unprotected sources. Under the Water Quality Act of 1965, states were required to develop and establish water quality standards; by the early 1970s, all states had developed their standards.

Passage of the Water Quality Act of 1965 (October 2, 1965) also resulted in the creation of the Federal Water Pollution Control Authority (FWPCA) as a separate entity in the Department of Health, Education, and Welfare (HEW); however, FWPCA did not reside there long. In 1966, it was transferred to the Department of the Interior (Reorganization Plan No. 2 of 1966). Under the Water Quality Improvement Act of 1970 (84 Stat. 113, April 3, 1970), FWPCA was abolished and its functions were folded into the new United States Environmental Protection Agency (U.S. EPA).

The sporadic amendments of the FWPCA of 1948 and repeated restructuring and reorganization of the responsible agency resulted in inefficient implementation of these laws and regulations. During the early 1970s, increased awareness of environmental and health issues and emergence of environmental activism along with complacent environmental regulations resulted in an exigent need to overhaul the system and consolidate authority. To resolve these issues, FWPCA was amended in 1972 and Congress passed the Clean Water Act (CWA) of 1972. Under this act (the Clean Water Act of 1972), the U.S. EPA was given the mandate to develop regulations to restore and maintain the chemical, physical, and biological integrity of the nation's water. This act manifests a basic change in strategy, and the focus of enforcement was changed from implementing water quality standards to effluent regulations. The National Pollution Discharge

Elimination System (NPDES) was established under which each point source is required to obtain a discharge permit. Under the Clean Water Act, the U.S. EPA was given a mandate to establish technologically based effluent limitations.

The Clean Water Act of 1972 mandated that all publicly owned sewage treatment works use the "best practical control technology" for industrial wastes and secondary treatment by July 1, 1977, "best available treatment economically achievable" by July 1, 1983, and the elimination of all discharges of pollution by 1985. However, these deadlines were relaxed on the recommendations of the National Commission on Water Quality in 1977, and the zero discharge goals were eliminated. The commission recommended a comprehensive recycling and reuse program using alternative and innovative technologies to achieve the goals of the Clean Water Act.

Over the course of years, the CWA evolved into a multifaceted statute ensuring the achievement of its objectives through the construction of local wastewater treatment plants (construction grant program) and the creation of a major national wastewater discharge permitting system.

Local public agencies were required to submit proposals for each proposed construction project. These projects were funded through the federal grants allocated to the states. During the period 1972–1990, more than $60 billion in federal grant money was invested in municipal sewage treatment works, along with more than $20 billion from state and local governments.

Under the CWA amendments of 1987, the construction grant program was changed, and beginning in 1990, the method of municipal financial assistance was shifted from grants to loans provided by state revolving funds. Since 1990, all the states have created "revolving loan funds," which serve as low-cost financing sources for water quality infrastructural projects.

The Safe Drinking Water Act (SDWA), passed in 1974, established stringent regulatory standards for the quality of drinking water for all public water supply systems except the very smallest. Currently, the Safe Drinking Water Act regulates more than 175,000 public water systems in the United States.

The Surface Water Treatment Rule was passed in 1989. This rule covers all public water supply systems using surface water or groundwater under the direct influence of surface water and intends to protect against exposure to pathogens such as *Giardia*, viruses, and *Legionella*. It requires 4 and 3 log removal of enteric viruses and *Giardia* during water treatment. It also requires a 0.2 mg/L of residual disinfectant at the point of water entry and maintaining a detectable level of residual disinfectant throughout the distribution system. The Interim Enhanced Surface Water Treatment Rule (IESWTR) was promulgated in 1989. It covers all public water supply systems that serve 10,000 people or more and use surface water or groundwater under the direct influence of surface water. It requires 2 log removal of *Cryptosporidium* from filtered water systems.

The Stage I Disinfectants and Disinfection By-products (D-DBP) Rule of 1998 was aimed at providing public health protection to an additional 20 million households by reducing the levels of disinfectants and disinfection by-products in drinking water. This rule established the maximum residual disinfectant level goals (MRDLGs) for chlorine, chloramines, and chlorine dioxide; maximum contaminant level goals (MCLGs) for four trihalomethanes, two haloacetic acids, bromate, and chlorite; and National Primary Drinking Water Regulations (NPDWRs) for three disinfectants (chlorine, chloramines, and chlorine dioxide), two groups of organic disinfection by-products, and two inorganic disinfection by-products (chlorite and bromate). The NPDWRs consist of maximum residual disinfectant levels (MRDLs) or maximum contaminant levels (MCLs) or treatment techniques for these disinfectants and their by-products. The NPDWRs also include monitoring, reporting, and public notification requirements for these compounds.

In 2000, the U.S. EPA proposed a groundwater rule (GWR). It uses a multiple barrier strategy to reduce waterborne health risks. The following five components are the basis of this multiple barrier strategy:

- periodic sanitary survey of groundwater systems
- hydrogeologic assessment to identify wells sensitive to fecal contamination
- source water monitoring for systems drawing from sensitive wells without treatment or with other indications of risk
- correction of significant and fecal contamination
- compliance monitoring to ensure that disinfection treatment is reliably operated

The original Federal Water Pollution Control Act has had six major amendments in the intervening years. The passage of Federal Water Pollution Control Act Amendment of 1972, which is also known as the Clean Water Act (CWA), divides the history of water pollution control (including wastewater treatment) in the United States into two eras: the pre- and post-Clean Water Act periods. Before 1972, states had the regulatory authority for water pollution control, whereas federal involvement was limited to cases involving interstate waters. Due to unique socioeconomic and political circumstances, states have historically shown different levels of motivation and willingness to address water pollution control issues.

The history of water quality management and policy formulation is complicated (Table 1). The three most important lessons learned are institutional reform, improved processes for conflict resolution, and increased use of modern planning and decision-making procedures. All these lessons learned through the evolution of U.S. water quality management policy may serve as a guideline for other countries to avoid repeating mistakes and to ensure efficient use of water resources.

Table 1. Time Line of Water Quality Regulations in the United States

Year	Events	Comments
1890	Harbors and Rivers Act	Prohibited discharge of any refuse in to interstate waters
1893	Interstate Quarantine act	Intended to control spread of communicable diseases from one area to another
1899	Harbors and Rivers Act Amendment	Exempted the “refuse flowing from streets and sewers and passing therefrom in a liquid form”
1912	The Public Health Service Act	First water related regulation that prohibited use of a common cup on interstate carriers
1914	“Treasury Standards” or USPHS standards	First microbial standards applicable to water systems supplying water to interstate carriers. HPC count should be less than 100 per mL and not more than one of the five 10 mL portions of a sample can be positive for *E. coli*
1924	Oil and Pollution Control Act	To control discharge of oil damaging coastal waters
1925	Revised USPHS Standards	1 coliform/100 mL; first physical-chemical standards for Pb, Cu, Zn, and soluble mineral salts
1943	Revised USPHS Standards	Standards revised to include more contaminants
1946	Revised USPHS Standards	Standards revised to include more contaminants
1948	Federal Water Pollution Control Act	Public Law 8-845: first law to address comprehensively the issue based on the principle of limited federal enforcement and limited federal financial assistance
1948	Establishment of Division of Water Supply and Pollution Control (DWSPC),	DWSPC established in PHS, FSA, to administer the Water Pollution Control Act, June 30, 1948
1956	Federal Water Pollution Control Act Amendment	Authorized construction grants for municipal treatment works
1961	Water Pollution Control Act	Congress advocated 85% removal of pollutants using comprehensive programs and plans for water pollution abatement and control
1962	Revised USPHS Standards	Included 28 constituents
1965	Water Quality Act October 2	States are required to have water quality standards to receive grants
1965	Federal Water Pollution Control Administration (FWPCA)	FWPCA was established in the Department of Health, Education, and Welfare (HEW)
1969	Community Water Supply Study (CWSS) initiated	Initiated a survey of 969 public water systems (representing only 5% of nation’s water systems) for compliance with 1962 standards
1970	CWSS report released	41% of public water systems surveyed for compliance with 1962 standards
1970	Water Quality Improvement Act	No new provisions regarding standards. The act continued the authority of states to set water quality standards and the authority of FWPCA to approve such standards. Questioned the Environmental Protection Agency’s authority to require uniform treatment limitation for water discharge
1972	Louisiana (Mississippi River) Water Quality Report Published	36 organic compounds found in drinking water and source waters.
1972	“Clean Water Act” (CWA) or FWPCA amendment of 1972	Set discharge limits and required waste treatment prior to discharge and established National Pollution Discharge Elimination System (NPDES)
1973	GAO report	Out of 446 community water systems investigated, only 60 were found to comply with USPHS standards
1974	Consumer Reports (three part series) published	Organic contaminants in New Orleans
1974	Environmental Defense Fund	Organic contaminants in New Orleans
1974	Trihalomethanes (THMs) identified	A major public health concern
1974	National Organic Reconnaissance Survey	Survey of THM occurrence nationwide
1974	Safe Drinking Water Act (SDWA)	Established stringent regulatory standards for the quality of drinking water for all public water supply systems except the very smallest. Signed on Dec.16, 1974, as Public Law 93-523
1975	National interim primary drinking water regulations (NIPDWRs)	Based on 1962 USPHS standards, interim rules for 18 organic, inorganic, and microbial contaminants
1976	1st NIPDWRs amendment	Standards for four radionuclides
1976	Toxic Substances Control Act (TSCA)	Regulation of all compounds that have or may possess toxic properties
1977	Best practical control technology currently available (BPCTCSA)	Defined best available waste treatment for different industries
1979	2nd NIPDWRs amendment	Total THMs
1980	Comprehensive Environmental Restoration, Compensation, And Liability Act (CERCLA)	“Superfund” bill that established a detailed program to describe priority pollutants and first national database of pollutants
1980	3rd NIPDWRs amendment	Monitoring for corrosion and sodium
1983	4th NIPDWRs amendment	Identifies the best available mean to comply with THMs rule

Table 1. ***(Continued)***

Year	Events	Comments
1986	SDWA amendment	Major revisions introduced in SDWA. Set standards for pollutants in drinking water; authorizes contamination warning and alert system
1987	Volatile organic contaminants standards	Standards for eight contaminants
1988	SDWA amendment	Public Law 100-572: Lead Contamination Control Act
1989	Surface Water Treatment Rule	The Surface Water Treatment Rule regulates public water systems using a lake, stream, or pond water as a source of drinking water. It was intended to ensure the safety of drinking water
1989	Total Coliform Rule (TCR)	Establishes microbiological standards and related monitoring requirements for all community and noncommunity water systems
1991	National Primary Drinking Water Regulation for Radon-222	EPA proposed to regulate radon at 300 pCi/L. Public water supplies (systems serving over 25 individuals or with greater than 15 connections) are required to monitor, report, and notify requirements for radon (July 18, 1991) (56 FR 33050)
1991	Phase II SOCs and IOCs also lead and copper	Standards for 27 new contaminants
1992	Phase V SOCs and IOCs	Standards for 22 new contaminants
1996	SDWA amendment	Public Law 104-182 requires U.S. EPA to publish an MCLG and promulgate NPDWR for contaminants of public health significance
1996	Information Collection Rule	Monitoring only to provide information on the occurrence of disinfection by-products (DBPs) and pathogens, including *Cryptosporidium* (May 14, 1996)
	Analytical Methods for Radionuclides	Approval of additional (66) analytical methods to monitor gross alpha, gross beta, tritium, uranium, radium-226, radium-228, gamma emitters, and radioactive cesium, iodine, and strontium in drinking water (Mar. 5, 1997)
1997	Withdrawal of 1991 proposed rule on radon-222	Withdrawal of drinking water regulations proposed for radon-222 (Aug. 6, 1997)
1997	Small System Compliance Technology List for the Surface Water Treatment Rule	The list of technologies that small systems can use to comply with the Surface Water Treatment Rule (SWTR) (Aug. 6, 1997)
1998	Revisions to State Primacy Requirements	The regulations requiring states to obtain and retain primary enforcement authority (primacy) for the Public Water System Supervision (PWSS) program under Section 1413 of the Safe Drinking Water Act (SDWA) are amended (Apr. 28, 1998)
1998	Drinking Water Contaminant Candidate List	The list of contaminants that are known or anticipated in public water systems and may require regulations under the SDWA (Mar. 2, 1998)
1998	Variances and Exemptions Rule	The rule includes procedures and conditions under which a primacy state/tribe or the EPA Administrator may issue small system variances to public water systems serving less than 10,000 persons (Aug. 14, 1998)
1998	Consumer Confidence Reports	Community water systems are required to prepare annual consumer confidence reports on the quality of the water delivered by the systems and provide these reports to their customers (Aug. 14, 1998)
1998	Stage I D-DBP Rule	Standards for 11 new contaminants (Dec. 16, 1998)
1998	Interim Enhanced Surface Water Treatment Rule	Improve control of microbial pathogens, including specifically the protozoan *Cryptosporidium*, in drinking water; and address risk trade-offs for disinfection by-products (Dec. 16, 1998)
1999	Suspension of Unregulated Contaminant Monitoring Requirements for small public water systems	This rule exempts small and medium public water systems from monitoring unregulated contaminants during 1 year every 5 years, as required under UCMR (Jan. 8, 1999)
1999	Revisions to the Unregulated Contaminant Monitoring Rule.	This final rule includes a list of contaminants to be monitored. The data in the database will be used to identify contaminants on the Drinking Water Contaminant Candidate List (CCL) (Sept. 17, 1999)
1999	Analytical methods for chemical and microbiological contaminants and revisions to laboratory certification requirements	EPA has approved the updated versions of 25 American Society for Testing and Materials (ASTM), 54 Standard Methods for Examination of Water and Wastewater (Standard Methods or SM), and 13 Environmental Protection Agency (EPA) analytical methods for compliance determinations of chemical contaminants in drinking water (Dec. 1, 1999)
1999	Underground Injection Control Regulations for Class V Injection Wells	Class V Underground Injection Control (UIC) regulations are revised and new requirements are added for two categories of endangering Class V wells to ensure protection of underground sources of drinking water (Dec. 7, 1999)
1999	Lead and Copper Rule minor revisions	These minor revisions are intended to streamline requirements, promote consistent national implementation, and, in many cases, reduce the burden for water systems. The LCRMR do not change the action levels or MCLGs for lead and copper, established by the Lead and Copper Rule (Dec. 20, 1999)

(*continued overleaf*)

Table 1. **(*Continued*)**

Year	Events	Comments
2000	Analytical Methods for Perchlorate and acetochlor	This rule specifies the approved analytical methods for measuring of perchlorate and acetochlor in drinking water and also includes minor technical changes to correct or clarify the rule published on Sept. 17, 1999 (Mar. 2, 2000)
2000	Public Notification Rule	Owners and operators of public water systems are required to notify persons served when they fail to comply with the requirements of the National Primary Drinking Water Regulations (NPDWR), have a variance or exemption from the drinking water regulations, or are facing other situations posing a risk to public health (May 4, 2000)
2000	Removal of the MCLG for chloroform	In accordance with the order of the U.S. Court of Appeals for the District of Columbia Circuit, the zero MCLG for chloroform is removed from National Primary Drinking Water Regulations (NPDWRs) (May 30, 2000)
2000	Drinking Water State Revolving Fund Rule	This interim final rule gives each state considerable flexibility to determine the design of its DWSRF program and to direct funding toward its most pressing compliance and public health problem. It also explains the ways to receive a capitalization grant, how to use capitalization grant funds intended for infrastructure projects, and also the roles of both the states and EPA in managing and administering the program
2000	Radionuclides Rule	This rule establishes the MCLGs, MCLs, monitoring, reporting, and public notification requirements for radionuclides. Uranium is added to the list, and the monitoring requirements for combined radium-226 and radium-228, gross alpha particle radioactivity, and beta particle and photon radioactivity have been revised (Dec 7, 2000)
2001	Unregulated Contaminant Monitoring List 2 Rule	The analytical methods for thirteen chemical contaminants on List 2 are approved and also the schedule for *Aeromonas* monitoring is described (Jan. 11, 2001)
2001	Arsenic Rule	A health-based, nonenforceable MCLG for arsenic of zero and an enforceable MCL for arsenic of 0.01 mg/L (10 μg/L) is established for nontransient noncommunity water systems (Jan. 22, 2001)
2001	Filter Backwash Recycling Rule	The purpose of the FBRR is to further protect public health by requiring public water systems (PWSs), where needed, to institute changes to the return of recycle flows to a plant's treatment process that may otherwise compromise microbial control (June 8, 2001)
2002	Long Term 1 Enhanced Surface Water Treatment Rule (Jan. 14, 2002)	The rule requires systems to meet strengthened filtration requirements as well as to calculate levels of microbial inactivation
2002	Unregulated Contaminant Monitoring Regulation for Public Water Systems	Establishment of Reporting Date (Aug. 9, 2002) for water systems (serving more than 10,000 persons) to report all contaminant monitoring results they receive before May 13, 2002 for the Unregulated Contaminant Monitoring Regulation (UCMR) monitoring program; Direct Final Rule (Mar. 12, 2002)
2002	Guidelines establishing test procedures for the analysis of pollutants under the Clean Water Act; National Primary and Secondary Drinking Water Regulations	The test procedures (i.e., analytical methods) for the determination of chemical, radiological, and microbiological pollutants and contaminants in wastewater and drinking water are revised and updated; Final Rule (Oct. 23, 2002)
2002	Unregulated Contaminant Monitoring Regulation: Analytical methods for chemical and microbiological contaminants	Approval of the analytical method and an associated minimum reporting level (MRL) to support Unregulated Contaminant Monitoring Regulation's (UCMR) (for chemical and microbiological contaminants) List 2 *Aeromonas* monitoring; Final Rule (Oct. 29, 2002)
2003	Propose MCLG and National Primary Drinking Water Regulations	Regulations for any contaminant selected from contaminant candidate list (Aug. 2003)
2004	National Primary and Secondary Drinking Water Regulations	Approval of additional method (Colitag™) for detecting coliforms and *E. coli* in drinking water; Final Rule (Feb. 13, 2004)

READING LIST

U.S. National Archives and Records Administration. (2004). Records of the Federal Water Pollution Control Administration (FWPCA) (Record Group 382, Years 1948–68) www.archives.gov

U.S. EPA (1997). *The Safe Drinking Water Act—One Year Later—Success in Advancing Public Health Protection*. EPA 810-F-97-002.

U.S. EPA (1999). *25 Years of the Safe Drinking Water Act: History and Trends*, EPA 816-R-99-007.

U.S. EPA, Office of Water. http://www.epa.gov/water/.

WATER QUALITY MANAGEMENT IN A FORESTED LANDSCAPE

Thomas M. Williams
Baruch Institute of Coastal
Ecology and Forest Science
Georgetown, South Carolina

In 1955, Trousdell and Hoover published a paper in the *Journal of Forestry* that documented a rise in the water table following clear-cutting of a stand of loblolly pine on poorly drained soil. They warned that this rise may lead to saturated soils that inhibit regeneration and preclude using heavy equipment on the site. In the first century A.D., Pliny the Elder stated, "Often after woods have been cut down, springs on which the trees used to feed, emerge" (cited in Reference 1). Modern scientific methods can document truth that has been observed and debated for millennia. The influence of forests on water is such a subject. During two millennia of observation and nearly a century and a half of scientific investigation, the role of forests in water has been empirically demonstrated, but it is still not understood at a fundamental level.

Forested land produces the highest quality water of any land use (2), and forests have been used to protect drinking water supplies around the world. Forested streams also support food webs essential to the health of fish and other aquatic organisms. Rocky streambeds in forests are essential to aquatic macroinvertebrates that are used by some of our most desirable game fish. Protection of water was one of the objectives of the Organic Act of 1897, which formed the National Forest system, and the Weeks Act of 1911, which added 21 million acres of National Forests in the eastern United States (3).

In addition to protecting water, forests are also used by humans for harvest of timber products, harvest of game and fish, and outdoor recreation. Forests are often managed to enhance their usefulness for these purposes. Humans have used fire to alter ecosystems for hunting for at least 10,000 years, and exploitive logging has been practiced throughout the historical period. Management of forests to produce timber and fiber products has been widely practiced since the late nineteenth century. Much of our knowledge of the influence of forests on water has been gained by our efforts to evaluate the impact of forest management on water.

Modern forest hydrology research began in France in the nineteenth century. Widespread deforestation following the French revolution was blamed for a series of floods and landslides that plagued the country (1). During that debate, proponents of forests argued that reforestation would increase stream flow, regulate floods, and prevent erosion and landslides. Most of these claims were supported by observations of several romantic era writers and references to ancient Greek scholars. To resolve these disputes, the influence of forests became an area of active scientific inquiry. Surell (4) was one of the early researchers to collect numerical data on mountain watersheds and landslides. Observation of stream flow and rainfall began by 1850 and examination of watersheds with different land uses began by 1860. The controversy over the value of forests in regulating streams continued in the United States by the early 1900s. Pinchot (5) lamented that "the writing and talking on this branch of forestry has had little definite fact or trustworthy observation behind it." At that time, the U.S. Forest Service began efforts to collect trustworthy observations.

Studies at Wagon Wheel Gap, Colorado, introduced the most widely used technique in examining water in forested regions (6). The use of paired experimental watersheds (7) has produced statistically valid tests of many ideas that had been debated for centuries. This technique involves measuring the process of interest for several years in two closely matched watersheds, applying a treatment to alter the process in one watershed, and then continuing measures in both watersheds for several years. From a strong correlation between the data from the two watersheds prior to treatment, changes in the treated watershed can be determined with high confidence. Over the last century, paired watersheds have been used to study most aspects of forest management. A solid empirical basis is available for evaluating forest management impacts on water quality. We have a relatively clear understanding how forests produce clean water and which aspects of forest management have the potential to degrade water quality.

Water from undisturbed forests has very little suspended sediment. Estimates of total sediment loss from forested watersheds across the United States are generally less than 0.5 Mg/ha/yr (8). Very low rates of erosion are due to the mechanism of peak flow generation in forested watersheds. Between rains, flow in most streams is derived from groundwater, which has little associated sediment. Sediment is added to streams as rainwater travels across the soil surface and suspends soil particles. Horton (9) explained the generation of surface flow as rain that fell more quickly than water could enter the soil. If water could enter the soil at 3 mm/h and it rained 10 mm/h, then 7 mm of rain must flow across the surface and into streams. However, in temperate forests, measured infiltration rates may approach 1000 mm/h and any incoming rain will enter the soil (10). However, we know that forested streams rise quickly after rain begins. A number of different mechanisms have been found that deliver water quickly to streams in forested watersheds (11). These mechanisms do not include water moving very far along the surface of the ground. In forested areas, most of the sediment found in streams has been eroded from stream banks.

High rates of infiltration on forest soils result from two mechanisms. The tree canopy intercepts raindrops and, once it is wetted, allows drops to fall a short distance to the forest floor, thus protecting the forest floor from high energy impacts during intense rain. Litter covering the soil and organic matter incorporated into the soil create large pores and allow them to stay open throughout the storm. As long as these soil pores remain open, surface water does not erode the soil and add sediment to streams. Protection of the soil pores by litter and organics is the more important mechanism. High infiltration can occur if the litter and organic matter are not disturbed when the canopy is removed.

Disturbance of forest soil by harvest and replanting of trees in commercial forestry is the most likely way that forest management can degrade water quality. Removing forest products from the site requires cutting trees, skidding products to a landing, loading them on a truck, and hauling them to a mill. Each aspect can damage the forest floor. Sediment losses from skid trails and landings may be as high as 7000 Mg/ha and from roads can be as high as 1000 Mg/ha (12). On a watershed basis, sediment losses of 10 Mg/ha/yr have been found following logging and preparation of the watershed for replanting (13,14). This soil loss rate was similar to that considered acceptable to maintain agricultural productivity and occurs only for 3–5 years of a 25-year rotation. Since those studies were concluded, forestry has been challenged to improve water quality associated with harvesting and regeneration. Throughout the United States, foresters have examined techniques to minimize exposure of bare soil and to build roads and landings in a way that maintains a buffer of undisturbed forest soil for infiltration before runoff reaches a stream bank. These techniques, known collectively as best management practices (BMPs), are now being used across the United States (15). A number of studies have shown that BMPs reduce sediments by roughly 90% (16).

Cutting the forest does increase the amount of water that becomes stream flow and will cause an unavoidable increase in sediment in alluvial streams. More water is returned to the atmosphere by evaporation of intercepted rain and transpiration by trees than any other land cover. Examining runoff data from 94 paired watershed experiments, Bosch and Hewlett (17) concluded that there was a generally linear increase in steam flow with increasing percentage of forest removed. They also concluded that cutting conifers resulted in a greater increase than cutting deciduous hardwoods. Sahin and Hall (18) looked at 145 such experiments and estimated that stream flow increased 20–25 mm for each 10% reduction in the cover of conifers and 17–19 mm for each 10% reduction in hardwoods. Increases in stream flow tend to increase the size of alluvial channels, causing an unavoidable increase in suspended sediment. Van Lear et al. (19) found an increase from 0.040 Mg/ha/yr to 0.15 Mg/ha/yr in sediment loss from a watershed where the only source of sediment was a stream channel. However, the stream flow returns to preharvest values as the new forest grows. In some rapidly growing forests, water yield may actually decline below the pretreatment level when the growth rate is at a maximum (20).

An early experiment to determine if forest regeneration was responsible for the decline in stream flow revealed the importance of nitrogen cycling in forest harvesting. At Hubbard Brook, New Hampshire, all trees on a watershed were cut and all vegetation killed for three successive years to examine stream flow (21). Stream flow increased by more than 300 mm during the first year. This increase was maintained as long as herbicide was applied to the watershed and declined when vegetation was allowed to grow. In addition to an increase in flow, the concentration of nitrate, calcium, and magnesium increased greatly. Nitrate concentrations exceeded drinking water standards. The large pool of nitrogen in the vegetation and the forest floor was mineralized. Nitrification of this mineralized nitrogen yielded nitrate and hydrogen ions; nitrate was leached from the system, and hydrogen ions displaced cations, which were then lost.

Due to the concerns raised by the Hubbard Brook experiment, cycling and retention of nitrogen is now better understood than any other biogeochemical process in forested ecosystems. Nitrogen is tightly held in forested ecosystems and export of nitrate to streams is highly limited. Ten separate chemical mechanisms for retention of mineralized organic nitrogen have been identified in forest soils (22,23). The most common are ammonium uptake by bacteria that decompose high C:N material, ammonium or nitrate uptake by plants, and denitrification of nitrate to nitrous oxide or nitrogen gas.

When a forest is harvested, a great deal of organic nitrogen is deposited in the form of logging slash, and moisture and temperature increase in the forest floor. The logging slash is also high in carbon and decomposition bacteria outcompete nitrifiers for mineralized ammonium. As carbon is lost as carbon dioxide, the ratio of carbon to nitrogen declines and decomposition bacteria are less able to compete with nitrifiers. Unless regrowth is inhibited, plants become competitors for ammonium and also can use nitrate produced by nitrifiers. Finally, if nitrate is leached from the forest slope, it moves to the riparian zone where soils are often saturated. In these saturated zones, bacteria deplete oxygen from the soil and then use nitrate as an electron receptor in anaerobic respiration. Nitrate is reduced to nitrous oxide or to neutral nitrogen and lost from the system as a gas. Unless all these mechanisms are limited, normal forest harvesting does not result in nitrate concentrations of more than 1 mg/L. These mechanisms are most effective in moist, warm, temperate forests.

In addition to the original Hubbard Brook experiment, nitrate leaching has been occasionally found in other places, although stream concentrations have generally been found below 10 mg/L. Clear-cutting other watersheds in the mountains of New England also showed elevations of nitrate near 10 mg/L (24). Use of herbicide, especially hexazinone (25), in repeated treatments (26) may increase stream concentrations to the 2–5 mg/L range. In northwestern Europe, nitrogen input from air pollution appears to be increasing stream nitrate concentrations. Nitrate concentrations between 0.5 and 1 mg/L are found in undisturbed forested watersheds in Scandinavia (27). Lepistõ (28) found evidence that a watershed in Finland had become saturated with nitrogen and showed nitrate leaching even during the growing season. This occurred after atmospheric nitrogen input had increased from 2.9 kg/ha/yr in 1969–1979, to 6.2 kg/ha/yr in 1971–1979, to 8.5 kg/ha/yr in 1980–1990.

A number of other anions and cations are found in low concentrations in streams draining forested landscapes. The concentrations of these ions and ratios of one ion to another are related to location and subsurface geology. The ratios of various ions have been used widely to interpret pathways of water within the watershed. Mulholland (29) used ratios of sulfate and calcium to differentiate among vadose zone, shallow groundwater, and deep groundwater sources of stream flow in a watershed dominated by dolomite. In this case, atmospheric input of sulfate resulted in high concentrations in throughfall, but adsorption by iron and aluminum in the soil B horizon resulted in low sulfate in shallow groundwater. Likewise, calcium was very low in throughfall but very high in groundwater that had been in contact with the dolomite bedrock. Eisenbeer and Lack (30) used a similar technique with silica and potassium to examine surface runoff and pipe flow in an Amazonian watershed. Intense weathering of the soil and bedrock resulted in leaching of silica into the water passing through the soil. Intense weathering also isolated potassium into the living biomass; it was leached from leaves in throughfall and surface flow. The ratio of potassium to silica could be used in the same way that sulfate and calcium were used in the other geologic setting.

Hydrogen ions are generally not of great concern in forest streams with pH slightly acid, unless the geology is dominated by calcareous materials. Nitrification produces an H^+ ion for each NO_3^- ion and can be a source of acidification in stream water. Generally, that is not the case because H^+ is held more tightly in the soil cation exchange complex, and potassium, calcium, or magnesium is displaced to be carried to the stream. Afforestation acidified streams in the British Isles on former moorland (31).

As forest management strives to increase the growth of forest products, pesticides and fertilizer have been used on more forest plantations. Although these two activities are not closely related, contamination from both can be avoided in the same way. If rain dissolves fertilizer and pesticide from the soil surface, they will be quickly immobilized within the soil. Contamination from either can be eliminated by avoiding application in streams or near stream areas that may saturate in a large rain. Buffering both perennial and ephemeral streams when fertilizer or pesticide is applied will prevent contamination. Both materials are often aerially applied, so a number of management practices relating to atmospheric conditions are followed in development of best management practices. Most problems with these materials have been caused when aerial sprays are inadvertently applied to surface waters.

Fire is a tool that has been used to manage forest ecosystems since early man first learned to use it. Fire can have a profound impact on water quality. The impact of fire is closely related to the fate of the forest floor. If the forest floor is damp and fire intensity is low, there is little impact on water quality. As fire intensity increases, the impact on water also increases. If trees are killed and the forest floor remains intact, the result is similar to logging with an increase in flow but small changes in sediment or chemistry. However, fire effects can be much larger than those from logging if the forest floor is dry or the fire is intense enough to dry and consume the forest floor. As more bare soil is exposed, infiltration declines and water runs across the surface carrying sediment and cations from the ash. If the fire is highly intense, volatilized organic compounds move from the forest floor into the soil. As these compounds cool, soil particles become hydrophobic and prevent water from entering the soil (32). In this case, all rainwater runs across the surface, there is severe erosion, and streams are filled with sediment. Thus, fire may result in sediment losses from 1–2 Mg/ha to 500 Mg/ha, depending on fire severity and land slopes (33). In areas of high air pollution, severe fire may also result in nitrate concentrations that exceed drinking water standards (34). The goal of fire management is to prevent such intense fires by using prescribed, less intense fires, mechanical fuel reduction, and fire suppression.

Undisturbed forests produce the highest quality water of any land use. However, forested watersheds also lose more water to the atmosphere though evaporation and transpiration than other land uses. Human use of the forest can degrade the quality of water produced if care is not taken to protect the forest floor. Careless logging, especially poorly designed roads, can contribute to stream sedimentation and degrades water for both human use and as fish habitat. Most of these problems can be avoided if best management practices are used. There are few problems with chemical water quality, except for nitrate, which may become a problem if forest regrowth is prevented or in areas with prolonged high atmospheric nitrogen deposition. Such problems have been seen in northwestern Europe and after fire in the Los Angeles basin.

BIBLIOGRAPHY

1. Andréassian, V. (2004). Waters and forests: From historical controversy to scientific debate. *J. Hydrol.* **291**: 1–27.
2. Brooks, K. and Achouri, M. (2003). Sustainable use and management of freshwater resources: The role of forests. In: State of the Worlds Forests 2003. Food and Agricultural Organization, United Nations, Rome, Italy, pp. 74–85.
3. Sedell, J., Sharpe, M., Dravnieks Apple, D., Copenhagen, M., and Furniss, M. (2000). Water and the Forest Service. USDA, Forest Service, Washington Office, Washington, DC, p. 2000.
4. Surell, A. (2004). *Etude sur les torrents des Hautes Alps*. Carilian-Goeury et Victor Dalmont, Paris, p. 1841 (cited in Reference 1).
5. Pinchot, G. (1905). A primer of forestry. Bureau of Forestry,. U.S. Dept. of Agriculture, Government Printing Office, Washington, DC (cited in Reference 1).
6. Bates, C.G. and Henry, A.J. (1928). Forest and streamflow experiment at Wagon Wheel Gap, Colorado. *Mon. Weather Rev. Suppl.* No. 3. USDA Weather Bureau, Washington, DC.
7. Hewlett, J.D. (1982). *Principles of Forest Hydrology*. University of Georgia Press, Athens, GA, pp. 86–88.
8. Binkley, D. and Brown, T. (1993). Forest practices as nonpoint sources of pollution in North America. *Water Resour. Bull.* **29**: 729–740.
9. Horton, R.E. (1933). The role of infiltration in the hydrologic cycle. *Trans. Am. Geophys.* **14**: 446–460.
10. Hursh, C.R. and Brater, E.F. (1941). Separating stormhydrographs form small drainage areas into surface and subsurface flow. *Trans. Am. Geophys. Union.* **Part III**: 863–871.
11. Bonnel, M. (1993). Progress in understanding of runoff generation in forests. *J. Hydrol.* **150**: 217–275.
12. Elliot, W.J. (2000). Roads and other corridors. Drinking water from forests and grasslands: a synthesis of the scientific literature. G.A. Dissmeyer (Ed.). USDA Forest Service,

Asheville, NC. General Technical Report. SRS-39, P. 85–100. Chap. 9.

13. Hewlett, J.D. (1979). Forest water quality: an experiment in harvesting and regenerating Piedmont forests. School of Forest Resources, University of Georgia, Athens, GA.
14. Pye, J. and Vitousek, P.M. (1985). Soil and nutrient removals and windrowing on a Southeastern US Piedmont site. *Forest. Ecol. Manage.* **11**: 145–155.
15. Shepard, J. (2002). *For. Best Manage. Pract. Symp.*, Gainesville, FL. National Council on Air and Stream Improvement, CD-ROM.
16. Ice, G.G., Megahan, W.F., McBroom, W.M., and Williams, T.M. (2003). Opportunities to assess past, current and future impacts of forest management. In: *Total Maximum Daily Load (TMDL) Environmental Regulations II*. American Society of Agricultural Engineers, St, Joseph, MI, pp. 243–248.
17. Bosch, J.M. and Hewlett, J.D. (1982). A review of catchment experiments to determine the effect of vegetation changes on water yield. *J. Hydrol.* **55**: 3–23.
18. Sahin, V. and Hall, M.J. (1996). The effects of afforestation and deforestation on water yields. *J. Hydrol.* **178**: 293–309.
19. Van Lear, D.H., Douglass, J.E., Cox, S.K., and Augspurger, M.K. (1985). Sediment and nutrient export in runoff from burned and harvested pine watersheds in the South Carolina. Piedmont. *J. Environ. Qual.* **14**: 169–174.
20. Cornish, P.M. and Vertessy, R.A. (2001). Forest age induced changes in evapotranspiration and water yield in a eucalypt forest. *J. Hydrol.* **242**: 43–63.
21. Likens, G.E. et al. (1970). Effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook watershed-ecosystem. *Ecological Monogr.* **40**: 23–47.
22. Vitousek, P.M. and Melillo, J.M. (1979). Nitrate losses from disturbed forests: patterns and mechanisms. *Forest Sci.* **25**: 605–619.
23. Nadelhoffer K.J., Aber, J., and Melillo, J.M. (1984). Seasonal patterns of ammonium and nitrate uptake in nine temperate forest ecosystems. *Plant and Soil* **80**: 321–335.
24. Johnson, C.E., Johnson A.H., and Siccama, T.G. (1991). Whole-tree clear-cutting effects on exchangeable cations and soil acidity. *Soil Sci. Soc. Am. J.* **55**: 502–508.
25. Neary, D.G., Bush, P.B., and Grant, M.A. (1986). Water quality of ephemeral forest streams after site preparation with the herbicide hexazinone. *Forest Ecol. Manage.* **14**: 23–40.
26. Munson, A.D., Margolis, H.A., and Brand, D.G. (1993). Intensive silviculture treatment: impacts on soil fertility and planted conifer response. *Soil Sci. Soc. Am. J.* **57**: 246–255.
27. Arheimer, B., Andeson, L., and Lepistõ, A. (1996). Variation of nitrogen concentration in forests streams—influences of flow, seasonality and catchment characteristics. *J. Hydrol.* **179**: 281–304.
28. Lepistõ, A. (1995). Increased leaching of nitrate at two forested catchments in Finland over a period of 25 years. *J. Hydrol.* **172**: 103–123.
29. Mulholland, P.J. (1993). Hydrometric and stream chemistry evidence of three storm flowpaths in Walker Branch Watershed. *J. Hydrol.* **151**: 291–310.
30. Eisenbeer, H. and Lack, A. (1996). Hydrometric and hydrochemical evidence for fast flowpaths at La Cuenca, Western Amazonia. *J. Hydrol.* **180**: 237–250.
31. Neal, C., Smith, C.J., and Hill, S. (1992). Forestry impacts on upland water quality. Institute of Hydrology Report Series No. 30. Institute of Hydrology, Wallingford, UK.
32. Debano, L.F. (2000). The role of fire in soil heating and water repellency in wildland environments: a review. *J. Hydrol.* **231–232**: 195–206.
33. Landsburg, J.D. and Tiedemann, A.R. (2000). Fire management. In: Drinking water from forests and grasslands: a synthesis of the scientific literature. USDA Forest Service. Asheville, NC, General Technical Report SRS-39, Chap. 12, pp. 124–138.
34. Riggan, P.J. et al. (1994). Effects of fire severity on nitrate mobilization in watersheds subject to chronic atmospheric deposition. *Environ. Sci. Technol.* **28**: 369–376.

TRACE METAL SPECIATION

Jamie R. Lead
University of Birmingham
Birmingham,
United Kingdom

INTRODUCTION

Trace metals may be defined as those metals whose concentration is low enough so that their presence exerts a negligible influence on the overall chemistry of a water body. They can be contrasted with major ions that are present in sufficient amounts to affect the chemistry of the system. Major metal ions include the alkali and alkaline earth metals (Na, K, Ca, Mg) and perhaps Al, Fe, and Mn. Although there is no strict definition of a “trace metal,” we usually consider the transition metals, certain metalloids, and lanthanides. Radioactive elements are often considered separately. The potential effects on ecological and human health created by a certain number of trace metals mean that these metals are of particular importance and are the subject of intensive study. The effects of these metals may often be deleterious, but many metals are also micronutrients. In no order, some of the most important and most investigated trace metals are chromium (Cr), nickel (Ni), copper (Cu), zinc (Zn), arsenic (As), cadmium (Cd), tin (Sn), mercury (Hg), and lead (Pb).

Trace metals in natural waters can be present at different concentrations but can also be distributed between different physicochemical forms or species. This distribution of a metal between different forms is termed its speciation. Important examples of speciation in the environment are: (1) the oxidation state, for example, Fe(II) and Fe(III); (2) the nature of the ligand in a solution phase metal complex, for example, Cd^{2+} and $CdCl_4{}^{2-}$; and (3) the distribution of metal between dissolved, colloidal, and particulate phases. The exact form in which the trace metal exists substantially mediates the bioavailability and toxicity of the metal and its transport in the environment and has important implications for biomagnification (i.e., increased concentration of pollutant as it passes through the food chain), human and ecological health, and biogeochemical cycling. Speciation is thus intrinsically important and also relevant to understanding wider environmental processes. We can examine the three examples above and discuss some of the possible environmental implications of speciation.

1. *The Oxidation State.* [Fe(II)/Fe(III)]—Fe(II) occurs in reducing (oxygen-deficient) conditions, such as porewaters in sediments and bottom waters of stagnant and/or biologically active lakes, particularly in the summer months. Conversely, Fe(III) occurs in oxidizing (oxygen-rich) conditions, such as surface river, lake, or marine waters. Additionally, Fe(II) tends to be present in the dissolved phase, Fe^{2+} (aq), whereas Fe(III) tends to hydrolyze (react with water) very strongly and form solid colloidal or particulate oxide or hydroxide phases, $Fe(OH)_3$ (s), where the (aq) denotes the aqueous phase and (s) denotes solid. This has important implications for the behavior of Fe, which are discussed below in example 3. In addition, trace metals may bind to the surface of solid Fe(III) particles. Interconversion of Fe(II) and Fe(III) at interfaces between oxidizing and reducing conditions may have an impact on the environmental fate and behavior of any associated metals. See Davison and de Vitre (1) for an in-depth discussion of iron chemistry in aquatic systems.
2. *The Nature of Solution Phase Ligand (Cd^{2+} and $CdCl_4{}^{2-}$).* The salinity of water (the concentration of the chloride ligand) determines the extent to which metals such as cadmium are present as the positively charged hydrated Cd^{2+} or as the negatively charged chloro complex $CdCl_4{}^{2-}$. Intermediate numbers of chloride ions in the complex are also possible. This situation is most commonly encountered in estuarine systems, where freshwater meets seawater and the chloride concentration of the water gradually increases. The change in charge from the positive Cd to the negative $CdCl_4$ affects properties of the metal such as permeability through biological membranes and sorption/desorption to clay, organic, and oxide particles. These changes have implications for metal bioavailability and transport. In addition to chloride, other important ligands exist such as sulfate, amino acids, and EDTA.
3. *Dissolved, Colloidal, and Particulate Distribution.* The distribution of metals between these three phases has long been recognized as essential to the fate and behavior of trace metals. Particular prominence is given to the free (dissolved) metal ion, as it is predicted that the biological response is a function of the dissolved concentration in bulk water (2). The relationship between speciation and bioavailability may, however, be less straightforward; complex and poorly understood physicochemical and biological interactions mediate biological uptake of metal (3,4). Transport through the environment is also affected by the partitioning of metal (5); microbiological activity, aggregation of colloids, and sedimentation of particles are among the dominant processes. Somewhat simplistically, dissolved and colloid-bound metals tend to remain in the water column, and particle-bound metals tend to settle out to the sediments. This geochemical fractionation results in different transportation behavior and alters the ultimate fate of the metal.

Table 1. Approximate Concentrations of Trace Metals in Seawater and Freshwater[a]

Metal	Seawater, molar	Freshwater, molar
Mn	10^{-9}	$10^{-5} - 10^{-9}$
Fe	10^{-9}	$10^{-4} - 10^{-7}$
Cu	$10^{-8} - 10^{-9}$	$10^{-6} - 10^{-9}$
Zn	$10^{-8} - 10^{-10}$	$10^{-5} - 10^{-8}$
Ni	10^{-8}	$10^{-6} - 10^{-10}$
Cd	$<10^{-9}$	$10^{-7} - 10^{-10}$
Pb	$<10^{-10}$	$10^{-6} - 10^{-9}$
Hg	10^{-11}	$10^{-7} - 10^{-10}$

[a] Data from Reference 6.

As indicated briefly in these examples, the question of trace metal speciation is complex. Our understanding of these difficult areas is in part due to the physicochemical and biological complexity of natural waters. Natural waters contain all or almost all elements of the periodic table. Trace metals can be present at extremely low levels, often at or below analytical limits of detection; concentrations are as low as 10^{-12} M in some cases. See Table 1 for illustrative examples of element concentrations, although it must be kept in mind that a wide range of concentrations exists. Additionally, a range of dissolved, colloidal, and particulate organic matter and suspended particles exist that have complex and unknown structures and effects on metal speciation. These include a range of chemistries (silicates, oxides, carbonates, humic substances, etc.) and sizes (1 nm upward). Microorganisms also play an important (and again poorly understood) role in metal speciation, through biochemical processes, sorption to cell walls and membranes, and the production and excretion of complexing ligands into the environment.

Our understanding of speciation is further hindered by the variability of natural waters spatially (e.g., surface waters in lakes may have a higher oxygen content than bottom waters and have significant effects on metal speciation) and temporally (e.g., daily and seasonal cycles in chemistry, often mediated by changes in biological activity) often over very small scales.

CHEMICAL REACTIONS THAT AFFECT METAL SPECIATION

We can distinguish between two broad types of chemical reactions that may affect speciation (7). First is redox reactions where the oxidation state of the atoms is altered. An example of this is the Fe(II)/Fe(III) interconversions in the examples above. Second are reactions where changes in the "coordinative relationships" of participating atoms occur. An example of this is the change from Cd^{2+}, where the coordinative partner is water (i.e., Cd and water are chemically bound to each other) to $CdCl_4{}^{2-}$, where the coordinative partner is the chloride ions (Cd and Cl are chemically bound to each other). Both acid–base, precipitation and complexation reactions fit this category. Reactions occurring in environmental systems are often impossible to separate conceptually because of the system complexity (8). For instance, example 1 in the Introduction

[Fe(II) to Fe(III) interconversions] involves all of these types of reactions. Detailed discussion of these reactions can be found in chemistry textbooks and in relation to environmental systems in Stumm and Morgan (7) and Buffle and Stumm (6). In this review, we concentrate on complexation reactions.

A metal complex is formed by the donation of electrons from a ligand to a metal; this is a stable association of metal and ligand whose characteristics are different from the individual components. The simplest example is that of the hydrated metal, sometimes termed "free" or "dissolved". (Note: It is important to distinguish between the term dissolved in this sense and the operationally used term dissolved, which is used to denote metals passing through a standard membrane during filtration. In the latter definition, a fraction of colloid-bound metals is included.) The hydrated metal ion is often written simply by excluding the coordinated water (as in example 2 above with Cd). Other ligands will always be indicated, as when the Cd is bound to chloride ligands.

Typical ligand groups include

1. oxygen groups: water; various types of organic groups (carboxylic acids, ketones, etc.), which may be in solution or more importantly as part of colloids and surfaces; various types of inorganic groups (silicates, oxides) mostly occurring as solid-phase material;
2. sulfur groups: thiols, thioethers, and other compounds, which may either be simple solution compounds or as part of larger compounds such as humic substances; and
3. nitrogen groups: for instance, amino acids and humic substances.

Solid phases containing ligand groups capable of binding metals in a comparable way to solution phase ligands play a vital role in determining metal speciation. Often this role is dominant as solid material contains a much higher concentration of ligand groups and their associated metal ions compared with the solution phase.

Metal speciation (e.g., the distribution of metal between solid and solution phase) depends on both the solution conditions and the nature of the solid phase. For instance, pH (related to concentration of hydrogen ions, H^+) has a major impact on speciation because it competes directly with metal ions for ligand sites. Generally, as the pH is lowered (higher concentration of hydrogen ions), metal ions tend to occur in the solution rather than in the solid phase. Other factors, which influence metal speciation through complex mechanisms, include ionic strength and the concentration of dissolved organic matter. The role of solid-phase material is also essential in determining the solid: solution distribution; the roles of oxides and organic matter are particularly important.

MEASUREMENT OF METAL SPECIATION

Despite the importance of metal speciation and the major advances in the last few years, speciation measurement remains complicated because of the difficulty of unambiguously interpreting an analytical signal in terms of a chemical species. Ion-selective electrodes can specifically measure the free metal ion activity (concentration) of metals such as copper, but generally they cannot measure the low levels of metal in unpolluted waters (9). Other techniques capable of measuring labile metal complexes include voltammetric methods and various competition methods using ion exchange resins, solid phases or solvent extraction (10). In addition, there are a number of physical separation methods in use such as filtration, centrifugation, and dialysis. When combined with a method for measuring total metal concentration, these give size-based metal speciation data (5). However, the data extracted are, at least to some extent, operational.

MODELING METAL SPECIATION

Similarly to the measurement of metal speciation, modeling has made immense advances during the last 10–15 years. However, the complexity of the systems studied and the difficulties of speciation analysis mean that many fundamental difficulties have not been solved. For instance, speciation modeling assumes that the aquatic system modeled is at chemical equilibrium (11). For many environmental systems, this is not the case. For instance, redox reactions and many sorption reactions of metals to organic surfaces may not be at equilibrium. The difficulties in fully and rigorously modeling these systems can be understood if we consider some of the components present in natural waters. These include

- solution phase inorganic ligands (e.g., chloride, carbonate, sulfate)
- solution phase organic ligands (e.g., amino acids, EDTA)
- inorganic colloids and surfaces (e.g., oxides, silicates, carbonates)
- organic colloids and surfaces (e.g., humic substances, microbial exudates)
- mixtures of ligands and metals
- variety of solution conditions

A number of model types have been used successfully to interpret metal interactions in laboratory systems. Generally, solution phase speciation is relatively well understood. Additionally, much progress has been made in modeling metal interactions to oxide surfaces and humic substances (7,11), and development and validation continue (12). Individual models have been brought together in easy-to-use computer packages called speciation codes. There are now dozens of speciation codes that have been developed from the 1960s onward, often for very different purposes and with their own strengths and weaknesses. Despite this, the use of a valid, predictive speciation model for use in natural waters remains an elusive goal, toward which much effort is focused.

SUMMARY

Trace metal speciation is defined as the distribution of the metal among different physicochemical forms. Speciation has important implications for the bioavailability, toxicity, and environmental transport of metals. For instance, free and colloidal forms of the metal may tend to remain in solution; larger particle-bound metal tends to sediment out and is removed from the water column. This difference in behavior results in geochemical fractionation of different metal forms and results in differences in both the fate and behavior of the metal. Additionally, the size of a metal complex may be an important parameter in understanding metal uptake by organisms, although in more complex ways because of biological processes. However, despite major advances in our understanding, methods of measuring and modeling metal speciation require significant development before they can confidently provide rigorous understanding in this important area.

BIBLIOGRAPHY

1. Davison, W. and de Vitre, R.R. (1992). Iron particles in freshwaters. In: *Environmental Particles*. J. Buffle and H.P. van Leeuwen (Eds.). Lewis, Boca Raton, FL.
2. Campbell, P.G.C. (1995). Interactions between trace metals and aquatic organisms: a critique of the free-ion activity model. In: *Metal Speciation and Bioavailability in Aquatic Systems*. John Wiley & Sons, Chichester, UK.
3. Mirimanoff, N. and Wilkinson, K.J. (2000). Regulation of Zn accumulation by a freshwater gram-positive bacterium (*Rhodococcus opacus*). *Environ. Sci. Technol.* **34**: 616–622.
4. Tessier, A., Buffle, J., and Campbell, P.G.C. (1994). Uptake of trace metals by aquatic organisms. In: *Chemical and Biological Regulation of Aquatic Systems*. J. Buffle and R.R. de Vitre (Eds.). Lewis, Boca Raton, FL.
5. Lead, J.R., Davison, W., Hamilton-Taylor, J., and Buffle, J. (1997). Characterising colloidal material in natural waters. *Aquat. Geochem.* **3**: 219–232.
6. Buffle, J. and Stumm, W. (1994). General chemistry of aquatic systems. In: *Chemical and Biological Regulation of Aquatic Systems*. J. Buffle and R.R. de Vitre (Eds.). Lewis, Boca Raton, FL.
7. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters*. John Wiley & Sons, Hoboken, NJ.
8. Tessier, A. (1992). Sorption of trace element on natural particles in toxic environments.
9. Mota, A.M. and Correira dos Santos, M.M. (1995). Trace metal speciation of labile chemical species in natural waters and sediments: electrochemical approaches. In: *Metal Speciation and Bioavailability in Aquatic Systems*. John Wiley & Sons, Chichester, UK.
10. Apte, S. and Batley, G.E. (1995). Trace metal speciation of labile chemical species in natural waters and sediments: non-electrochemical approaches. In: *Metal Speciation and Bioavailability in Aquatic Systems*. John Wiley & Sons, Chichester, UK.
11. Turner, D. (1995). Problems in trace metal speciation modelling. In: *Metal Speciation and Bioavailability in Aquatic Systems*. John Wiley & Sons, Chichester, UK.
12. Vulkan, R. et al. (2000). Copper speciation and impacts on bacterial responses in the pore water of copper contaminated soils. *Environ. Sci. Technol.* 5115–5121.

METAL ION HUMIC COLLOID INTERACTION

Horst Geckeis
Institut für Nukleare
Entsorgung
Karlsruhe, Germany

INTRODUCTION

Humic and fulvic acids in natural water are well known to have significant impact on metal ion bioavailability and mobility in aquifer systems. In natural water, toxic heavy metals, rare earth elements, and radionuclides such as the actinide ions are frequently found to be attached to natural organic matter such as humic and fulvic acids. These complexes are known to be of high kinetic and thermodynamic stability under given geochemical conditions and therefore may control trace metal transfer to bioorganisms and their transport in soil and groundwater systems. For those reasons, the interaction of metal ions with such humic matter is a major issue in a number of research fields such as environmental monitoring, risk assessment of contaminated sites, and nuclear waste disposal. In order to understand and to assess the impact of humic and fulvic acids on the behavior of trace metals, we need to know the interaction mechanisms, complexation kinetics, and finally thermodynamic data that can be implemented into geochemical speciation codes. A large spectrum of experimental methods is used for obtaining (1) insight into the humic acid structure and its properties as a complexing ligand, (2) quantitative information on metal ion speciation in humic containing water, and (3) molecular insight into the metal–humic interaction mechanism.

The ongoing development of understanding and experimental methods still makes the research on humic matter quite a dynamic area. Besides this short overview of the current state of the art in humic–metal ion interaction research, more comprehensive reviews on various aspects of metal–humic interaction are available (e.g., see Reference 1).

BASIC PROPERTIES OF HUMIC ACID

Humic/fulvic acid is a class of organic substances in the natural carbon cycle with a high stability and comparably high residence time (2,3). It is found in different sources, such as dissolved in natural water and as constituent, especially in soil, sediments, and brown coal. Humic acid is used both as a collective term for humic substances with a sufficiently high content of hydrophilic groups that dissolve in pH neutral range and as a specific term for the fraction that is dissolved in pH neutral range and flocculates in the acidic range with the complementary fraction of fulvic acid, soluble also under acidic conditions. In the following, humic acid (HA) is used in the general sense (humic and fulvic acids).

The composition, mass distribution, and functional group content vary within limits, reflecting different origins and histories. Disregarding sulfur and nitrogen (present in varying low concentrations generally on the order of up to around 1% by weight), the atomic composition is dominated by carbon, oxygen, and hydrogen. The contributions of these substances vary around $CO_{0.5}H$; that is, the formal oxidation state of carbon is around zero.

It should be borne in mind that description of the metal ion–humic colloid (or humic acid and fulvic acid as humic colloid constituents) interaction has been hampered in the past by large uncertainty concerning the structure of humic and fulvic acids. For a very long time, the molecular mass was given by number from around 10,000 to 100,000 or higher (2). Recent application of advanced mass spectroscopic methods [matrix assisted laser desorption ionization-time-of-flight mass spectrometry (MALDI-TOF), electrospray ionization mass spectrometry (ESI-MS), and TOF–secondary ion mass spectrometry (TOF-SIMS)], Overhauser nuclear magnetic resonance (NMR), and the use of size/diffusion velocity-based methods [size exclusion chromatography (SEC) and flow-field flow fractionation (FFF)] show that the mass distribution maxima are considerably lower, centering at masses below 1000 daltons and ending at around 3000 mass units (4,5). This means that several assumptions made in the past, concerning, for example, assumed pronounced macromolecular polyelectrolyte character, have to be reassessed today. In dissolved form, the molecules are highly hydrated and possibly form associates. Therefore, the size distribution is generally determined to lie within a few nanometers in diameter. The proton exchange capacity can reach up to 14 meq/g (6), where protonation/deprotonation takes place from about pH 10 down into the acidic range (below pH 3 a considerable number of the groups are still ionized). Carboxylic types of groups are normally quantified with about one-third of the total capacity. The rest are mainly ascribed to phenolic-type groups. The carboxylic groups with pK_a values around 4 account for the humic negative charge at even low pH, while the phenolic sites tend to deprotonate in the high pH region only (pK_a around 10).

MODELING APPROACHES

Transformation of species distribution data to complexation constants is not trivial for humic acid. Today there is still no general agreement. Major challenges are the difficulty to take the inherent heterogeneity of humic acid functional group characteristics into account, the appropriate selection of an electrostatic model, and finally the applicability to complex multicomponent real systems with a broad variability in water chemistries. Different model approaches have been developed.

A relatively simple nonelectrostatic approach has been applied to estimate metal ion speciation in a humic-containing groundwater (7). The authors found the accessibility of functional groups for interaction with metal ions to vary with water chemistry (pH, ionic strength) and the origin of the humic acid under consideration. Taking such variability into account by an experimentally determined value for the capacity of metal ion loading under given conditions allowed the determination of complexation constants for a variety of metal ions that are invariant with metal ion concentration, ionic strength, and pH. The hitherto most widely applied humic acid complexation models (8) are the more sophisticated WHAM (Windermere humic aqueous model) (9) and NICA (nonideal competitive adsorption)–Donnan model (10). Both approaches take metal ion interaction with carboxylate and phenolic groups into account. Heterogeneity of functional groups is considered in WHAM by a discrete binding model postulating four different subsites for each functional group type, that is, carboxylic and phenolic sites. In the NICA–Donnan approach, a continuous site distribution is implemented. Both models use Donnan-type expressions for taking electrostatic effects into account. The model assumption of a Donnan gel phase varying with ionic strength is certainly debatable for the very small humic acid molecules but appears to be able to mimic the surrounding diffuse double layer (11). A comprehensive data set on generic values on deprotonation and metal ion complexation constants for a variety of metal ions is available for both codes (12,13) and at least allows for a first estimation of metal ion speciation in complex systems containing humic matter even though parameter adjustment turns out necessary in many specific cases.

Figure 1 shows trends for complexation constants derived from different model approaches for several metal ions. Complexation constants taking an experimentally obtained loading capacity into account are shown in Fig. 1a (14). The bars reflect the uncertainty of the data obtained by applying a multitude of different speciation methods ranging from ultrafiltration to fluorescence spectroscopy. Log ß* values describing the complexation reaction $Me^{z+} + HA(z) \rightleftharpoons MeHA(z)$ increase—as one may expect—more or less with increasing charge of the metal cations and their tendency to hydrolysis in solution. $HA(z)$ corresponds to the accessible site concentration. The fact that the derived complexation constants do not exactly follow the trend in hydrolysis constants suggests the possible influence of steric hindrances. A similar picture can be obtained by plotting generic data obtained from reanalysis of literature data by using the NICA–Donnan approach (Fig. 1b) (12). Those constants K_1 and K_2 represent median affinities of metal ions to HA proton exchanging groups, which are modeled by a continuous site distribution with two maxima.

SPECIATION METHODS

Experimental approaches to differentiate humic-bound metal ions from other species and to characterize the metal–humic compound can be roughly subdivided in two groups:

1. Physical or chemical separation of bound humic from molecular metal ions (invasive speciation).
2. Spectroscopic methods (noninvasive speciation).

Different experimental methods are used for obtaining the different types of information. For the purpose of determining interaction constants, quantification of the noncomplexed and complexed ions is required.

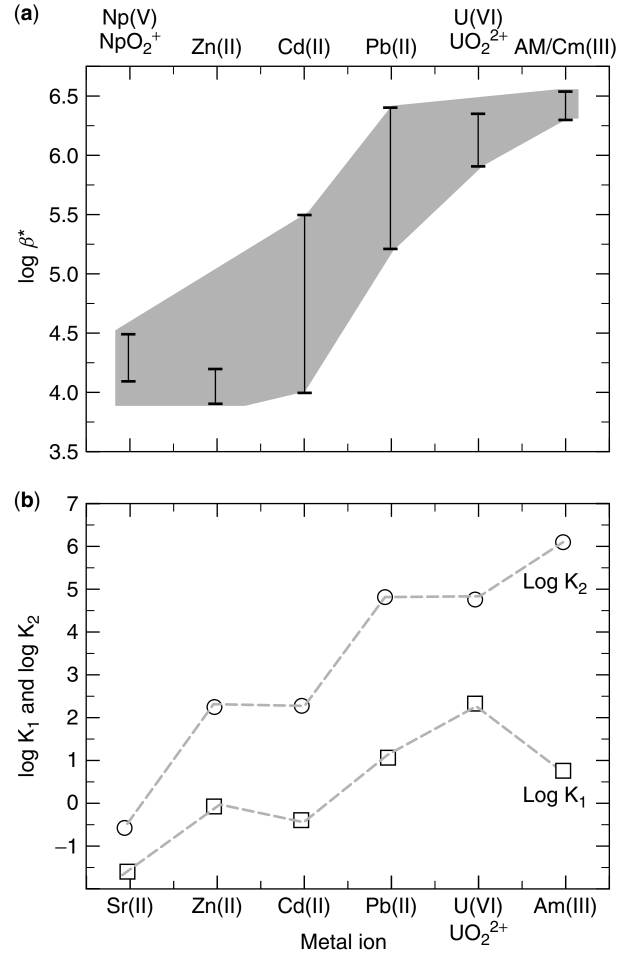

Figure 1. HA complexation constants for various metal ions obtained by using different model approaches. Data in part (**a**) are taken from Reference 14; those in part (**b**) correspond to the NICA–Donnan model approach (12). For explanation see text.

Experimental methods frequently used for determination of the species distribution in the metal ion–humic acid system are given in Table 1. Species distribution can be obtained by separation of the metal ion and the humic acid (complex) by size, density, or charge. Application of such methods is associated with uncertainty with respect to loss of substance, incomplete or uncertain species separation, possible kinetic rearrangements during separation, and problems related to changes in chemical composition. For this reason, methods are preferred where equilibrium is maintained. Frequently applied speciation tools for a number of trace metal ions such as Cu, Cd, Zn, and Pb are the electrochemical methods. Notably, anodic stripping voltammetry (ASV) achieves very low detection limits but intervenes in chemical equilibria and in many cases provides some kind of "operational" speciation. It is possible to differentiate "labile" metal ion species (i.e., free ions and readily dissociating complexed species) from stable or inert complexes.

Spectroscopic methods shown in Table 1 can give the species distribution without disturbing the equilibrium conditions. In general, absorption spectroscopy requires relatively high metal ion concentrations exceeding the ranges relevant to natural conditions. Fluorescence spectroscopy using appropriate fluorescent metal ions as probes [e.g., U(VI), Eu(III), Cm(III)] usually provide higher sensitivity for both excitation and emission. The humic acid itself is also fluorescent. The humic acid fluorescence intensity is quenched by metal ion complexation and thus the degree of fluorescence quenching can be used to study the metal ion–HA interaction and to quantify the complexation reaction. Probing the coordination environment for insight into the structure of the complex is specifically possible by X-ray absorption spectroscopy methods such as EXAFS and XANES as well as multinuclide nuclear magnetic resonance (NMR) spectroscopy. Examples for all these methods are given below.

METAL ION–HA INTERACTION MECHANISMS

Interaction of metal ions with humic/fulvic acids in principle may proceed via different pathways (see Table 2). Those individual mechanisms cannot be differentiated just by physical and/or chemical separation. In many cases spectroscopy can help to identify the binding mode and provide even details of the metal ion coordination environment. The following section is certainly not complete and only aims to give a rough overview into recent investigations on metal binding to humic/fulvic matter.

Complexation to Humic/Fulvic Acid

Various spectroscopic methods have been applied to gain insight into the coordination of the metal–humic/fulvic acid binding. Cu(II) is found to interact with four oxygen or two oxygen and two nitrogen atoms in a tetragonal coordination environment as shown by ESR (27). Later studies combining EXAFS and XANES spectroscopic experiments seem to verify the contribution of N ligands to the Cu(II) coordination (28). Studies on Cd(II) complexation with humic acid by EXAFS resulted in coordination via six oxygen atoms (distances 2.297 Å) (29). It seems to be clear that only carboxylate groups coordinate with Cd(II) as suggested by ^{113}Cd-NMR (30). Similar findings are recorded for the Ni(II) coordination with soil fulvic acid by EXAFS (31). Different to those *d*-transition metal ions, organomercury ions such as $Hg(CH_3)^+$ mainly bind to organic sulfur groups, for example, the thiol or disulfane groups. After saturation of those groups, coordination to oxygen atoms is found by EXAFS (32). Inorganic Hg(II) ions appear to form bidentate complexes with one thiol and one oxygen containing functional group (33). Al(III)–FA interaction studied by ^{27}Al-NMR suggests a coordination by carboxylic acid groups very similar to that observed in the oxalate complex (34). Formation of binary humate and ternary hydroxo and carbonato humate complexes are found for the interaction of trivalent actinides such as Am(III) and Cm(III) by UV/Vis absorption and notably the time-resolved laser fluorescence spectroscopy (TRLFS) (35,36). Such ternary species have rarely been taken into account by most model approaches up to now, but they appear to be even dominant

Table 1. Examples of Experimental Methods Used for Determination of Species Distribution (Complexed and Noncomplexed Metal Ions) in the Metal Ion–Humic Acid System and Methods Providing Analytical Information on the Nature of Such Complexes

Methods	Principle
Invasive	
Ultrafiltration, dialysis, ultracentrifugation, flow-field flow fractionation (FFF), size exclusion chromatography (SEC)	Size- or density-based separation (15–18,48)
Ion exchange, electrophoresis, Donnan membrane	Charge-based separation (19–22)
Ion selective electrodes (potentiometry), anodic stripping voltammetry (ASV)	Separate determination of noncomplexed or “labile” metal ion species (23,24)
Noninvasive (or Minimum Invasive)	
Fluorescence quenching, time-resolved laser fluorescence spectroscopy (TRLFS), UV/Vis, IR absorption	Detection and quantification of the complexation reaction via spectral shifts, emission intensity variations, fluorescence lifetime variations (provides to some extent information on coordination environment) (25,26) (see section below)
X-Ray absorption spectroscopy (EXAFS, XANES)	Absorption spectrum fine structure provides information on coordination environment (see section below)
Nuclear magnetic resonance spectroscopy (NMR)	Chemical shift provides chemical information on coordination environment (see section below)

in the near-neutral range in typical groundwater. It is noteworthy that TRLFS studies can be carried out at metal ion trace concentrations. For Cm(III) concentrations as low as 10^{-8} mol/L are sufficient. Such low metal ion concentrations are not accessible by most other spectroscopic methods. The tetravalent actinide Th(IV) studied by XPS and EXAFS turned out to mainly interact with carboxylic groups and was found to be coordinated by 9–10 oxygens at a distance of 2.43 Å (37), similar to trivalent ions such as Tb(III) (38). The coordination of trivalent metal ions like Eu(III) via carboxylate groups has been unequivocally verified by studying the C1s NEXAFS spectra of the humic acid (39). The $\pi^*_{C=O}$ transition of the carboxylate groups is clearly depressed upon complexation with Eu(III) ions. Hexavalent $UO_2{}^{2+}$ ions are also found to interact only with carboxylate groups as shown by EXAFS study. Phenolic groups appear to be of minor importance as a coordinating ligand (40).

Metal Ion Binding to Inorganic–Humic Agglomerates

Metal ion–HA interaction in natural systems tends to be complicated by the existence of both mineral surfaces and nanosized mineral particles, both partly coated with HA. Simple additive approaches—that is, assuming the metal complexation to be accomplished by interaction with the mineral surface and independently with sorbed and dissolved humic acid—often tend to underestimate the sorbed metal ion fraction (41,42). One possibility for the discrepancy may be the formation of ternary complexes such as ≡S–O–Me–HA (where ≡S–O stands for a deprotonated mineral surface hydroxyl group and Me for the metal ion). A recent study claims, on the evidence of fluorescence spectroscopic experiments, that such ternary complexes indeed exist in the γ-Al_2O_3/HA/Cm(III) system (43).

Formation of Inorganic Colloids Stabilized by Humic/Fulvic Acid

In many cases, redox reactions are involved in the metal ion–humic interaction. Such processes can also be followed by EXAFS and XPS as has been shown for the reactions of Np(V), Pu(VI), and Tc(VII) (44–46). Humic bound Fe(II) and/or quinone groups are thought to act as electron donors. In all those cases it turned out that precipitating oxyhydroxides of the reduced tetravalent species to some extent are kept in solution as colloidal species stabilized by humic acid coatings. The differentiation of Me–HA complexes and colloid–HA agglomerates becomes challenging in such systems because pronounced long-term kinetic effects are involved in such redox reactions. Those studies were performed to investigate the possible behavior of radionuclides tentatively released from a nuclear waste repository. But inorganic colloid–humic aggregates have also been found to dominate the transport of trace elements like iron (47). The binding state of Fe(III) in a river colloidal matter was scrutinized by combined electron paramagnetic resonance (EPR), transmission electron microscopy (TEM), and FFF experiments. It turned out that the total Fe(III) was distributed between humic acid complexes, FeOOH colloids, and structurally

Table 2. Possible Binding Modes of Metal Ion Interaction with Humic Acid (HA)

	Interaction Mechanism	Dissociation Rates^a
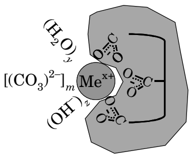	Formation of inner-sphere complexes with functional groups of the HA macromolecules	↑ Increasing
	or	
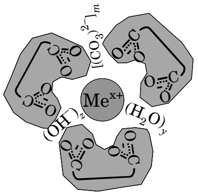	Interaction with more than one HA entity. In both cases the formation of ternary complexes with HA and other dissolved ligands are possible.	
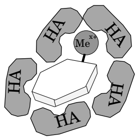	Surface sorption to inorganic colloids (e.g., aquatic clay colloids), which in turn are stabilized by humic/fulvic acid coating.	
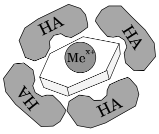	Inclusion of the metal ion into inorganic colloids stabilized by humic/fulvic coating due to formation of oxyhydroxide colloids as a consequence of redox reactions or formation of solid solutions.	

^{a}In a qualitative way, the dissociation rates of metal ions from colloidal species increase from bottom to top

incorporated in clay colloids, the last two agglomerated with humic matter. Similar observations have also been made for trace elements such as the naturally abundant lanthanides, uranium, and thorium in a humic-rich groundwater (48).

It is obvious that the metal ion interaction mode with humic matter is of importance for the assessment of HA influence on the geochemical cycle of trace elements. Notably, the kinetics of the metal ion desorption from the individual HA compounds varies considerably depending on the interaction mechanism (see Table 2). Even though there is still uncertainty concerning the exact coordination environment of HA bound metal ions, the coordination of most metal ions with HA carboxylate groups is now verified by a number of spectroscopic studies.

KINETICS

The metal ion interaction shows a sequence of reaction/exchange velocities. Primary association is rapid, followed by rearrangement to kinetically more stable binding. Dissociation is considerably slower and rates are decreasing with increasing aging time. Dissociation shows different kinetic modes, varying from seconds to weeks (49,50). The metal ion is believed to either migrate to stronger binding sites (51) or move to inner sites of the humic macromolecular structure (52). In view of the now generally accepted relatively small size of the humic acid molecules, the existence of a significant inner space becomes somewhat implausible. It is more likely that intermolecular bridging by polyvalent cations leading to the interaction of metal ions with more than one humic acid entity (see Table 2) become responsible for the observed time-dependent decrease of dissociation rates. The existence of such metal-induced humic agglomeration even at trace metal concentrations has been visualized, for example, by flow-field flow fractionation studies (48) or by using fluorescence probes (53). The metal-induced aggregation of natural organic matter is furthermore well known from studies in the marine environment (54).

Investigating humic matter in a natural groundwater revealed that part of the highly charged metal ion inventory of the humic acid shows a very slow dissociation, basically irreversible for time scales that can be studied under laboratory conditions. These metal ions show a different behavior from that observed for metal ion–humate complexes synthesized under laboratory conditions (48). The exact nature of binding for this inventory is not yet understood. One possibility is association with nanosized inorganic mineral presumably as solid solution. Various support for that assumption is found in analyzing various groundwater humic colloid size fractions: while humic acid is found mainly at sizes <10 nm, a major fraction of natural colloid-borne polyvalent elements such as thorium and rare earth elements are found in larger colloids >30 nm. Since metal ion dissociation from such types of colloids is very slow, their colloid-mediated transport becomes very relevant. Such processes are of interest in areas such as colloid-facilitated radionuclide transport in nuclear waste disposal research. A final conclusion also with regard to the relevance of such mechanisms to the toxic metal transport will require further investigations.

OUTLOOK

Metal ion interaction with humic acid is still a challenging field. A lot of new insight into the metal ion coordination could be obtained by the application of sensitive spectroscopic and other analytical methods. Knowledge about the specific binding state at least allows the estimation of humic-bound metal behavior in a semiquantitative way. There are chemical models available that—with some drawbacks—are able to predict metal ion speciation to some extent in humic-containing media. Open questions still remain concerning the “real” structure of humic/fulvic substances under *in situ* conditions and regarding the “real” configuration of metal ion–humic complexes. Application of newly emerging advanced analytical methods will certainly drive the research field and improve both the insight into the “real” metal–humic interaction mechanisms and the accuracy of model predictions.

BIBLIOGRAPHY

1. Clapp, C.E., Hayes, M.H.B., Senesi, N., Bloom, P.R., and Jardine, P.M. (Eds.). (2001). *Humic Substances and Chemical Contaminants*. Soil Science Society of America, Madison, WI, Section 3, p. 303.
2. Stevenson, F.J. (1994). *Humus Chemistry*, 2nd Edn. John Wiley & sons, Hoboken, NJ.
3. Buckau, G. et al. (2000). Origin and mobility of humic colloids in the Gorleben aquifer system. *Appl. Geochem.* **15**: 171–179.
4. Simpson, A.J. et al. (2002). Molecular structures and associations of humic substances in the terrestrial environment. *Naturwissenschaften* **89**: 84–88.
5. Leenheer, J.A. and Croue, J.-P. (2003). Characterizing aquatic dissolved organic matter. *Environ. Sci. Technol.* **37**: 18A–26A.
6. Elkins, K.M. and Nelson, D.J. (2002). Spectroscopic approaches to the study of the interaction of aluminum with humic substances. *Coord. Chem. Rev.* **228**: 205–225.
7. Kim, J.I. and Czerwinski, K.R. (1996). Complexation of metal ions with humic acid: metal ion charge neutralization model. *Radiochim. Acta* **73**: 5–10.
8. Dudal, Y. and Gerard, F. (2004). Accounting for natural organic matter in aqueous chemical equilibrium models: a review of the theories and applications. *Earth Sci. Rev.* **66**: 199–216.
9. Tipping, E. and Hurley, M.A. (1992). A unifying model of cation binding by humic substances. *Geochim. Cosmochim. Acta* **56**: 3627–3641.
10. Benedetti, M.F., Milne, C.J., Kinniburgh, D.G., van Riemsdijk, W.H., and Koopal, L.K. (1995). Metal ion binding to humic substances: application of the non-ideal competitive adsorption model. *Environ. Sci. Technol.* **29**: 446–457.
11. Benedetti, M.F., van Riemsdijk, W.H., and Koopal, L.K. (1996). Humic substances considered as a heterogeneous Donnan gel. *Environ. Sci. Technol.* **30**: 1805–1813.
12. Milne, C.J., Kinniburgh, D.G., van Riemsdijk, W.H., and Tipping, E. (2003). Generic NICA-Donnan model parameters for metal-ion binding by humic substances. *Environ. Sci. Technol.* **37**: 958–971.
13. Tipping, E. (1998). Humic ion-binding model VI: an improved description of the interactions of protons and metal ions with humic substances. *Aquatic Geochem.* **4**: 3–48.
14. Marquardt, C.M. (Ed.). (2000). Influence of humic acids on the migration behaviour of radioactive and non-radioactive substances under conditions close to nature. *Forsch. Wiss. Ber. FZKA* **6557**: 14.
15. Kim, J.I., Rhee, D.S., and Buckau, G. (1991). Complexation of Am(III) with humic acids. *Radiochim. Acta* **52/53**: 49–55.
16. Czerwinski, K.R., Buckau, G., Scherbaum, F., and Kim, J.I. (1994). Complexation of the uranyl ion with aquatic humic acid. *Radiochim. Acta* **65**: 111.
17. Glaus, M.A., Hummel, W., and van Loon, L.R. (1995). Equilibrium dialysis–ligand exchange: adaptation of the method for determination of conditional stability constants of radionuclide–fulvic acid complexes. *Anal. Chim. Acta* **303**: 321.
18. Hasselöv, M., Lyven, B., Haraldsson, C., and Sirinawin, W. (1999). Determination of continuous size and trace element distribution of colloidal material in natural water by on-line coupling of flow field-flow fractionation with ICPMS. *Anal. Chem.* **71**: 3497–3502.
19. Burba, P., Rocha, J., and Klockow, D. (1994). Labile complexes of trace metals in aquatic humic substances: investigation by means of an ion exchange-based flow procedure. *Fresenius J. Anal. Chem.* **349**: 800–807.
20. Sonke, J.E. and Salters, V.J.M. (2004). Determination of neodymium–fulvic acid binding constants by capillary electrophoresis inductively coupled plasma mass spectrometry (CE-ICP-MS). *J. Anal. At. Spectrom.* **19**: 235–240.
21. Seibert, A. et al. (2001). Complexation behaviour of neptunium with humic acid. *Radiochim. Acta* **89**: 505–510.
22. Temminghoff, E.J.M., Plette, A.C.C., Van Eck, R., and van Riemsdijk, W.H. (2000). Determination of the chemical speciation of trace metals in aqueous systems by the Wageningen Donnan membrane technique. *Anal. Chim. Acta* **417**: 149–157.
23. Florence, T.M. (1982). The speciation of trace elements in waters. *Talanta Rev.* **29**: 345–364.
24. Filella, M., Buffle, J., and van Leeuwen, H.P. (1990). Effect of physico-chemical heterogeneity of natural complexes—Part I: Voltammetry of labile metal–fulvic complexes. *Anal. Chim. Acta* **232**: 209.
25. Cook, R.L. and Cooper, H.L. (1995). Metal ion quenching of fulvic acid fluorescence intensities and lifetimes: nonlinearities a possible three-component model. *Anal. Chem.* **67**: 174.
26. Kim, J.I., Rhee, D.S., Wimmer, H., Buckau, G., and Klenze, R. (1993). Complexation of trivalent actinide ions (Am^{3+}, Cm^{3+}) with humic acid: a comparison of different experimental methods. *Radiochim. Acta* **62**: 35–43.
27. Senesi, N., Bocian, D.F., and Sposito, G. (1985). Electron spin resonance investigation of copper(II) complexation by soil fulvic acid. *Soil Sci. Soc. Am. J.* **49**: 114–115.
28. Frenkel, A.I., Korshin, G.V., and Ankudinov, A.L. (2000). XANES study of Cu^{2+}-binding sites in aquatic humic substances. *Environ. Sci. Technol.* **34**: 2138–2142.
29. Liu, C., Frenkel, A.I., Vairavamurthy, A., and Huang, P.M. (2001). Sorption of cadmium on humic acid: mechanistic and kinetic studies with atomic force microscopy and X-ray absorption fine structure spectroscopy. *Can. J. Soil Sci.* **81**: 337–348.

30. Otto, W.H., Carper, W.R., and Larive, C.K. (2001). Measurement of cadmium(II) and calcium(II) by fulvic acids using ^{113}Cd-NMR. *Environ. Sci. Technol.* **35**: 1463–1468.
31. Strathmann, T.J. and Myneni, S.C.B. (2004). Speciation of aqueous Ni(II)-carboxylate and Ni(II)-fulvic acid solutions: combined ATR-FTIR and XAFS analysis. *Geochim. Cosmochim. Acta* **68**: 3441–3458.
32. Qian, J. et al. (2002). Bonding of methyl mercury to reduced sulfur groups in soil and stream organic matter as determined by X-ray absorption spectroscopy and binding affinity studies. *Geochim. Cosmochim. Acta* **66**: 3873–3885.
33. Haitzer, M., Aiken, G.R., and Ryan, J.N. (2003). Binding of mercury(II) to aquatic humic substances: influence of pH and source of humic substances. *Environ. Sci. Technol.* **37**: 2436–2441.
34. Lee, N.C.Y. and Ryan, D.K. (2004). Study of fulvic–aluminum(III) ion complexes by ^{27}Al solution NMR. In: *Humic Substances—Nature's Most Versatile Materials*. E. Ghabbour and G. Davies (Eds.). Taylor and Francis, New York, pp. 219–228.
35. Buckau, G., Kim, J.I., Klenze, R., Rhee, D.S., and Wimmer, H. (1992). A comparative spectroscopic study of the fulvate complexation of trivalent transuranium ions. *Radiochim. Acta* **57**: 105–111.
36. Panak, P., Klenze, R., and Kim, J.I. (1996). A study of ternary complexes of Cm(III) with humic acid and hydroxide or carbonate in neutral pH range by time-resolved laser fluorescence spectroscopy. *Radiochim. Acta* **74**: 141–146.
37. Denecke, M.A., Bublitz, D., Kim, J.I., Moll, H., and Farkes, I. (1999). EXAFS investigation of the interaction of hafnium and thorium with humic acid and Bio-Rex70. *J. Synchrotron Rad.* **6**: 394–396.
38. Monsallier, J.M. et al. (2003). Spectroscopic study (TRLFS and EXAFS) of the kinetics of An(III)/Ln(III) humate interaction. *Radiochim. Acta* **91**: 567–574.
39. Plaschke, M., Rothe, J., Denecke, M.A., and Fanghänel, Th. (2004). Soft X-ray spectromicroscopy of humic acid europium(III) complexation by comparison to model substances. *J. Electron Spectrosc. Relat. Phenom.* **135**: 53–62.
40. Denecke, M.A. et al. (1998). *Radiochim. Acta* **82**: 103–108.
41. Vermeer, A.W.P., McCulloch, J.K., van Riemsdijk, W.H., and Koopal, L.K. (1999). Metal ion adsorption to complexes of humic acid and metal oxides: deviations from the additivity rule. *Environ. Sci. Technol.* **33**: 3892–3897.
42. Rabung, T., Geckeis, H., Kim, J.I., and Beck, H.P. (1998). The influence of anionic ligands on the sorption behavior of Eu(III) on natural hematite. *Radiochim. Acta* **82**: 243–248.
43. Wang, X.K. et al. (2004). Effect of humic acid on the sorption of Cm(III) onto γ-Al_2O_3 studied by the time-resolved laser fluorescence spectroscopy. *Radiochim. Acta*, **92**: 691–734.
44. Zeh, P., Kim, J.I., Marquardt, C.M., and Artinger, R. (1999). The reduction of Np(V) in groundwater rich in humic substances. *Radiochim. Acta* **87**: 23–28.
45. Marquardt, C.M. et al. (2004). The redox behaviour of plutonium in humic rich groundwater. *Radiochim. Acta* **92**: 619–662.
46. Maes, A. et al. (2004). Evidence for the interaction of technetium colloids with humic substances by X-ray absorption spectroscopy. *Environ. Sci. Technol.* **38**: 2044–2441.
47. Benedetti, M.F., Ranville, J.F., Allard, T., Bednar, A.J., and Menguy, N. (2003). The iron status in colloidal matter from the Rio Negro, Brasil. *Coll. Surf. A* **217**: 1–9.
48. Geckeis, H., Rabung, Th., Ngo Manh, T., Kim, J.I., and Beck, H.P. (2002). Humic colloid-borne natural polyvalent metal ions: dissociation experiment. *Environ. Sci. Technol.* **36**: 2946–2952.
49. Langford, C.H. and Gutzman, D.W. (1992). Kinetic studies of metal ion speciation. *Anal. Chim. Acta* **256**: 183–201.
50. Schüssler, W., Artinger, R., Kienzler, B., and Kim, J.I. (2000). Conceptual modeling of the humic colloid-borne americium(III) migration by a kinetic approach. *Environ. Sci. Technol.* **34**: 2608–2611.
51. Rate, W., McLaren, R.G., and Swift, R.S. (1993). Response of copper(II)–humic acid dissociation kinetics to factors influencing complex stability and macromolecular conformation. *Environ. Sci. Technol.* **27**: 1408–1414.
52. Choppin, G.R. and Clark, S.B. (1991). The kinetic interactions of metal-ions with humic acids. *Mar. Chem.* **36**: 27–38.
53. Engebretson, R. and Von Wandruszka, R. (1998). Kinetic aspects of cation enhanced aggregation in aqueous humic acid. *Environ. Sci. Technol.* **32**: 488–493.
54. Chin, W.C., Orellana, M.V., and Verdugo, P. (1998). Spontaneous assembly of marine dissolved organic matter into polymer gels. *Nature* **391**: 568–572.

HEAVY METAL UPTAKE RATES AMONG SEDIMENT DWELLING ORGANISMS

Jos P.M. Vink
Institute for Inland Water
Management and Waste Water
Treatment–RIZA
Lelystad, The Netherlands

INTRODUCTION

Sediments of rivers, river banks, and floodplains are commonly characterized by elevated levels of heavy metals such as Cd, Cu, Cr, Ni, Pb, Zn, and As (1). Organisms that occur in these sediments are known to accumulate these metals (2,3), which, in turn, are passed on through the foodweb. However, the development of well-defined quality criteria for (aquatic) sediments is seriously hindered mainly because of three reasons:

1. Laboratory exposure tests that are conducted with organisms (bioassays) do not consider the geochemical composition of the sediment *in situ* and do not link temporal changes in chemical speciation with toxicological effects.
2. It is usually unclear which routes of exposure contribute to accumulation in organisms and dictate the adverse effects.
3. There is a general uncertainty about the relative contributions of the various metal species that take part in biological uptake.

The general concept of (bio)availability of metals becomes clear when body concentrations of either aquatic or terrestrial organisms collected from the field are compared with those performed in exposure tests using disturbed, oxygenated, and spiked sediments or waters. In the latter case, concentrations are mostly significantly higher, both in the close environment and in the organism itself. This phenomenon is explained by chemical speciation, which is based on the lability of metal species. The most labile form is the free dissolved ion, followed by dissolved inorganic metal–ion pairs (e.g., M–OH, $M-CO_3$, M–Cl) and organically complexed forms (M–DOM).

This concept of lability, however, has in recent years been more often theorized than actually shown by empirical data. This is mainly due to the fact that only very few techniques have become available to actually measure free ion activities, let alone to measure these periodically during exposure of organisms. The concept of the free ion activity model (FIAM) assumes that free aqueous metal ion concentrations, rather than the total or dissolved concentration, largely determine the toxicological or biological effect that is observed in organisms that are exposed to water or sediment containing heavy metals.

In a critical literature review, Campbell (4) concluded that studies that are suitable for testing the FIAM concept in the presence of natural dissolved organic matter (DOM) are extremely scarce. There are numerous studies that report on the effects of DOM on metal bioavailability, but virtually all of these are qualitative in nature (speciation is underdefined). The few quantitative studies (5–9,11) are, however, not in agreement, and the applicability of FIAM to natural waters (with DOM) remains to be demonstrated. It is concluded that “future progress would be greatly aided by the development of methods capable of measuring the free ion concentration of metals in the presence of DOM.”

Recently, a new technique was introduced aimed at tackling the difficulties described above. This technique was introduced as Sediment Or Fauna Incubation Experiment, SOFIE in short (12). With this technique, it is possible to quantify very labile metal fractions, including free metal ions, in pore water, at low oxygen and reduced conditions, repeatedly, and in a nondestructive manner to the sample. Bioassays are conducted simultaneously in the same setting. Sediment samples are obtained in an undisturbed manner, including the overlying surface water.

Here, we deal with the gradients and speciation of metal species, including free ion activities, over water–sediment interfaces. The effect of metal uptake by sediment dwelling organisms on metal speciation, and vice versa, is discussed.

METHODS

Materials

This study was performed with a novel experimental technique (EU patent 1018200/02077121.8, October 2001, J. Vink, Rijkswaterstaat), which was introduced as Sediment Or Fauna Incubation Experiment, or SOFIE®. The method was described in detail by Vink (12). In short, this “cell” consists of a core, 190 mm radius, 200 mm height, which is used as a sampling device to obtain undisturbed sediment, and the corresponding surface water. After sampling, this water–sediment system becomes part of the cell. Specific pore water probes, constructed from permeable PES polymers, are attached at 5-mm layer increments. Probes are directly connected to preconditioned ion-exchange microcolumns (MIC). A MIC contains a metal chelating polymer (styrene divinylbenzene copolymers, RI-718, containing paired iminodiacetate ions that act as chelating groups), which is highly selective to metals. Retained metals are liberated from the polymer with an acid-extraction procedure. In this setup, ion exchange between pore water and exchange polymer is practically instantaneous and is therefore executed at the reigning geochemical status, including redox potential, of the probed sediment layer.

With these cells, water–sediment interfaces from six different river systems in The Netherlands were sampled (Table 1). Sites were selected for their degree of contamination, physical properties, or their general representativeness for sediments commonly found in the river. Aqueous steady-state conditions were determined by probing the water/sediment on three different occasions. Location MB was probed every 5 mm, but the other locations were probed in intervals that coincided with the layering of the sediment—which could be observed through the transparent cell wall. After the introduction of test organisms, chemical speciation was monitored in time and depth.

Exposure Tests with Benthic Organisms

Species of the oligochaete *Limnodrilus* (family Tubificidae) and the mosquito larvae *Chironomus riparius* were chosen for accumulation test because of their wide abundance in aquatic systems and their importance for the ecological food chain. Both species are sediment dwellers, which means that they live in close contact with the sediment. Populations were bred in the laboratory under contaminant-free conditions. Oligocheate populations consisted mainly of *L. claparedeanus* and *L. hoffmeisteri* in equal amounts. Individuals with corresponding life stages (L2) were introduced in the cell in field-realistic densities (approximately 1250 oligochaetes and 220 chironomids, respectively). For locations MB and SW, only exposure tests with oligochaetes were performed. Body concentrations were measured periodically by sampling individuals from the population on seven occasions at 2–3 day intervals. They were allowed to void their gut for 48 h, which, based on dissection, was an adequate period. The organisms were freeze dried for 48 h and the dry weight was determined. Digestion was done in 500 μL of 14.9 M HNO_3 (Ultrex) at 180 °C. Metal concentrations in the digests were determined using inductively coupled plasma mass spectrometry (ICP-MS, Perkin Elmer Elan 6000). Dolt-2 (certified by the Community Bureau of Reference, BCR, Brussels, Belgium) was used as biological reference material.

Chemical Analysis

Aqueous concentrations of NO_3^-, NO_2^-, SO_4^{2-}, PO_4^{3-}, and Cl^- were measured with segmented ion chromatography (Dionex AS50 chromatography compartment, ED_{50} electrochemical detector). Limits of detection (LOD) ranged from 0.025 $mg{\cdot}L^{-1}$ for NO_2^- to 0.05 $mg{\cdot}L^{-1}$ for PO_4^{3-}. As, Al, Ca, Cd, Co, Cu, Cr, Fe, Mg, Mn, Ni, Pb, and Zn analyses were carried out with ICP-MS (LOD min/max: 0.003 $\mu g{\cdot}L^{-1}$ for Cd to 6.0 $\mu g{\cdot}L^{-1}$ for Zn). NH_4^+ analyses were performed by Skalar segmented flow analyzer SAN+

6250 matrix fotometer (LOD: 0.022 mg·L^{-1}). Dissolved organic carbon was determined with a Shimadzu 5000 TOC analyzer (LOD: 0.45 mg·L^{-1}).

RESULTS AND DISCUSSION

Table 1 shows the solid phase characteristics of the six sites. Steady-state concentrations in pore water and overlying surface water are shown in Fig. 1.

Typically, steep redox gradients occur over the water–sediment interface, more or less pronounced for the various sites. Redox-sensitive compounds follow these gradients: denitrification precedes sulfate reduction, and anaerobic conditions were reached in most cases within 15 mm from the water–sediment interface. The increasing reduction of ferric(III)oxyhydroxides to Fe(II), and manganese(V)oxides to Mn(II) results in increasing concentrations with depth of aqueous Fe^{2+} and Mn^{2+}. Pore water concentrations of Cd, Cu, Ni, Pb, and Zn in most cases tend to decrease with depth and coincide with sulfate reduction patterns. These metals most probably associate with sulfides and therefore become insoluble. Heavy metals, however, compete for binding opportunities between the reactive sulfide phase and relatively large amounts of dissolved organic matter, which keeps them in the aqueous phase. There is a pH range between the sites of almost 2 units.

Time-Varying Exposure Concentrations

Immediately after introduction of the test organisms to the water–sediment system, steady-state concentrations shifted. Oligochaetes and chironomids are sediment dwellers. They live in burrows and irrigate these with oxygen-rich surface water. Consequently, their immediate environment is oxidized, which in turn directly affects the chemical reactivity and bioavailability of metals that remain in the sediment. In Fig. 2, an example is shown of how aqueous concentrations of copper species change with time and depth during the exposure test (site MB).

Dissolved metal concentrations change with time, and so do specific metal species such as M^{2+}. These time-varying changes may follow first-order reaction kinetics:

$$dC/dt = kC$$

so the external concentration, or exposure concentration, at a given time $C(t)$ is written as

$$C(t) = C_i e^{k_0 t} \tag{1}$$

in which C_i is the initial concentration in (pore) water, and k_0 is a rate term that describes the increase or decrease rate of the initial steady-state concentration. In analogy, the free ion activity concentration at a given time, M^{2+} (t), may be written as

$$[M^{2+}](t) = [M_i^{2+}]e^{k_{0,\mathrm{act}} t} \tag{2}$$

The rates at which these concentration shifts occurred proved to highly affect bioaccumulation patterns, which is shown later.

Table 1. Sediment Properties

	Location					
	MB	GA	KV	HD	AS	SW
River System	Rhine	Waal	IJssel South	Hollands Diep	Meuse	IJssel North
Depth (m)	3.5	0.3	1	5.2	0.4	0.4
Dry weight (%)	55	80	73	34	45	43
<63 μm (%)	72.6	4.1	15.7	21.8	54.2	65.5
<16 μm (%)	59.3	3.1	9.7	17.8	23.3	49
<2 μm (%)	35.2	1.8	5.65	10.4	14	29.55
Organic C (%)	2.4	0.2	0.55	1.14	2.25	4.15
Inorganic C (%)	1.1	0.6	1.3	0.87	0.55	0.56
pH	6.9	8.1	7.4	7.6	7.2	7.6
E_h (mV)	−448	−163	−149	23	−147	−171
Al (g/kg)	23.4	5.6	3.5	17.3	9.2	11
As (mg/kg)	nd	5.6	4	6.5	8.5	22
Ca (g/kg)	24.1	10.7	19.6	35.1	23.2	32
Cd (mg/kg)	7.49	0.37	0.14	2.59	1.57	2.35
Co (mg/kg)	15.3	7.8	3.3	7.7	8.6	12
Cr (mg/kg)	102.1	32.4	16.8	45.9	27.7	57.5
Cu (mg/kg)	93.1	11.4	4.9	26.5	25.7	48
Fe (g/kg)	29.3	7.6	8.0	18.2	22.2	24.5
Hg (mg/kg)	1.5	0.2	0.055	0.31	0.205	1.20
Mg (g/kg)	7.79	2	3.06	5.24	3.48	8
Mn (mg/kg)	654	241.7	196.5	430.5	524.5	530
Ni (mg/kg)	35.6	6.6	8.35	19.4	20.8	31.5
Pb (mg/kg)	703.6	28	11.5	39.8	79	91.5
Zn (mg/kg)	1091	133	33	231	320.5	365

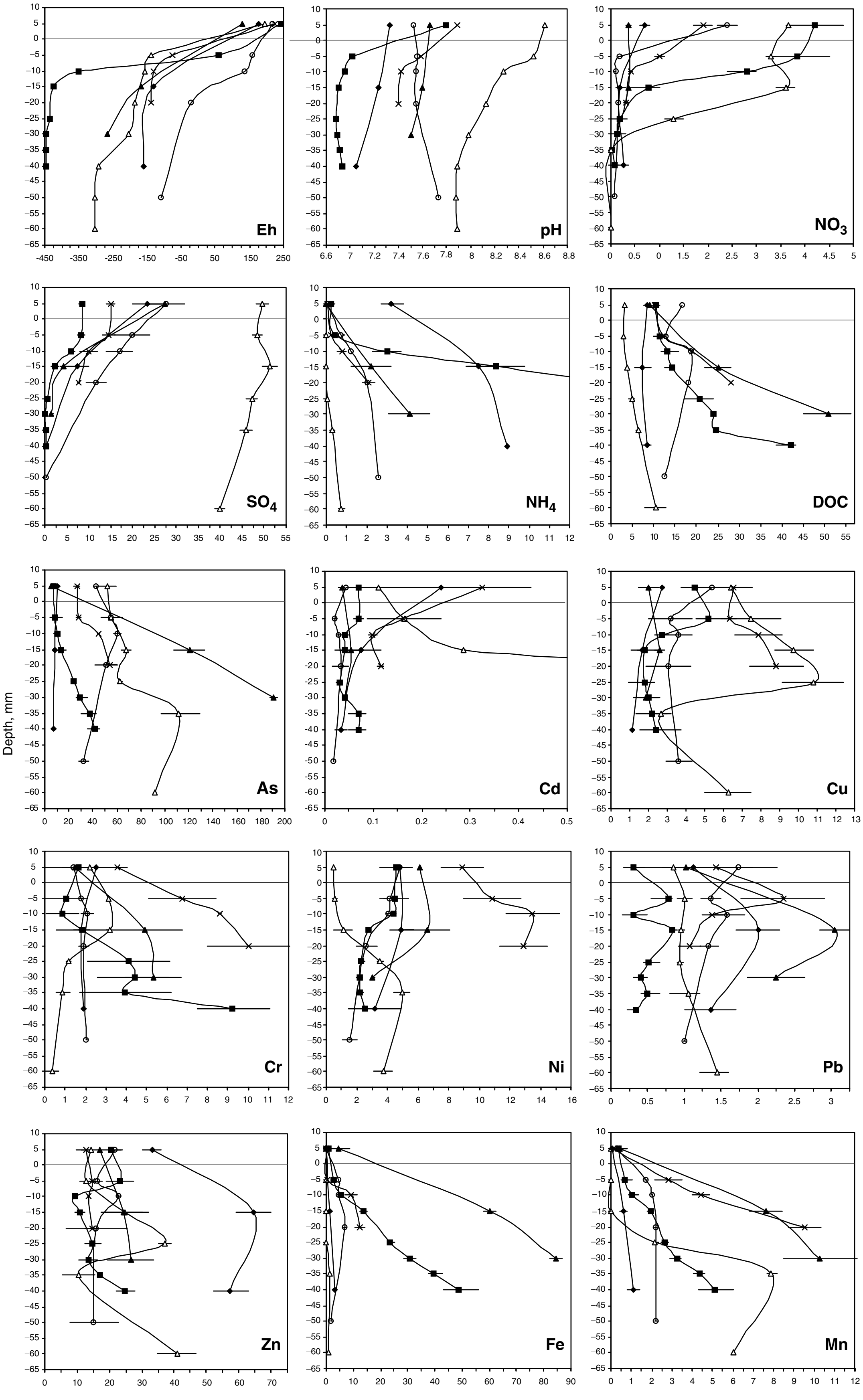

Figure 1. Aqueous steady-state characteristics for the six sites. MB (■); GA (△); KV (×); HD (O); AS (♦); SW (▲). E_h in mV; Fe, Mn, NO_3, SO_4, NH_4, DOC in mg/L; As, Cd, Cu, Cr, Ni, Pb, Zn in μg/L.

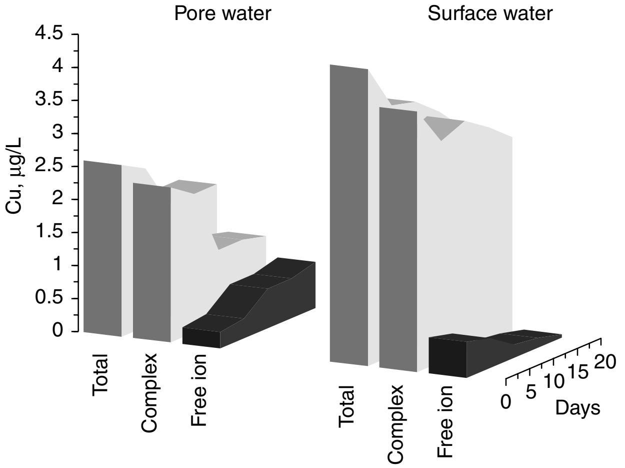

Figure 2. Example of time-dependent speciation shifts in overlying surface water and pore water (20 mm below water–sediment interface) during an exposure test. The free ion concentration in pore water increases, through oxidation, and is exhausted from the overlying water due to uptake by biota.

Uptake and Elimination by *Limnodrilus* and *Chironomus*

Metal uptake and excretion rates were approximated by a time/concentration-dynamic, two-compartment model. It is assumed that body concentrations vary with time and depend on uptake from the water phase with a certain rate. Uptake is at the same time accompanied by elimination (excretion) from the organism's body. Therefore,

$$dQ/dt = k_1 C(t) - k_2 Q(t)$$

with Q being the internal body concentration ($\mu g \cdot g^{-1}$ dw), k_1 and k_2 being uptake and elimination rate constants, respectively (d^{-1}), and t being time. This yields, for $Q(0) = 0$,

$$Q(t) = \frac{k_1 C_0}{k_2 - k_0}(e^{-k_0 t} - e^{-k_2 t}) \tag{3}$$

where C and k_0 are derived from Equations 1 and 2.

In analogy with the free ion activity,

$$Q(t) = \frac{k_1 [M^{2+}]}{k_2 - k_{0,\mathrm{act}}}(e^{-k_{0,\mathrm{act}} t} - e^{-k_2 t}) \tag{4}$$

With the introduction of k_0, time-dependent concentrations determine the overall exposure of organisms, and thus $Q(t)$. Note that elimination may partly be attributed to growth of the organisms during the exposure period. This separate loss term is not considered in the model.

To determine the consequence of *not* considering time-dependent concentrations, a one-compartment model was included for comparison. With this model, a constant uptake rate, a, is assumed, as opposed to concentration-dependent uptake. Body concentrations are then described as

$$dQ/dt = a - kQ$$

which yields

$$Q(t) = C_0 e^{-k \cdot t} + \frac{a}{k}(1 - e^{-k \cdot t}) \tag{5}$$

Figure 3 shows bioaccumulated amounts of some priority elements and heavy metals by oligochaetes and chironomids (site GA). Typically, metals are taken up at a certain rate (k_1), leading to increased body concentrations. After some time, elimination (k_2) exceeds uptake, and concentrations decline to reach a steady state. In general, bioaccumulated amounts can be ranked according to Ca > Fe > Mg > Zn > Al > Mn > Cu > Ti ~ Pb > Cr > As ~ Cd ~ Ni. Of both species, chironomids appear the most sensitive since they tend to take up these elements faster than oligochaetes and accumulate larger amounts (normalized to dry weight).

Examples of accumulation patterns for oligochaetes and chironomids for all locations are shown in Fig. 4. To test whether these sediment dwellers actually show different patterns at different locations (i.e., to determine location specificity), the accumulation data of oligochaetes and chironomids were statistically compared (two-tailed F-test, Mann–Whitney, $p < 0.05$). Results are summarized in Table 2 and show that locations do indeed matter. When distinguishing locations, both test species show the lowest sensitivity for chromium and nickel, which seem nondiscriminating. On the other hand, the combinations cadmium–chironomids and zinc–oligochaetes show a high discriminating potential. The fact that accumulation patterns on various locations are indeed different—also on locations with comparable dissolved concentrations—indicates that there is a difference in the actual bioavailability of metals.

Relation Between Time-Dynamic Speciation and Body Concentrations

To analyze the performance of the uptake and elimination models with varying exposure concentrations and speciation, the overall data set was divided into four subsets: two "environment compartments" (overlying surface water, and sediment pore water), each with two chemical "species" (total dissolved metal concentration, and free ion activities). Performance of the uptake–elimination models (Eqs. 3, 4, and 5) were tested with each subset. Results of measurements and corresponding model parameters C_i and k_0 were subsequently subjected to statistical analyses. As a descriptor for the goodness of prediction for the uptake–elimination model, the standard error of each model parameter was used rather than the correlation coefficient of the model with the measurements. The standard error is a measurement of the variability of a data point Y around the predicted value Y_p. It represents information about the goodness of fit in the same manner as the standard deviation does about the spread around the mean. The standard error is written as

$$sy \cdot x = \sqrt{\frac{\Sigma(Y - Y_p)^2}{n}}$$

By minimizing the sum of squares, values for k_1 and k_2 and corresponding standard errors were derived. Considering only Cu, Cd, Ni, Pb, Zn, and As, the data set consisted of over 1800 chemical and 608 biological measurements.

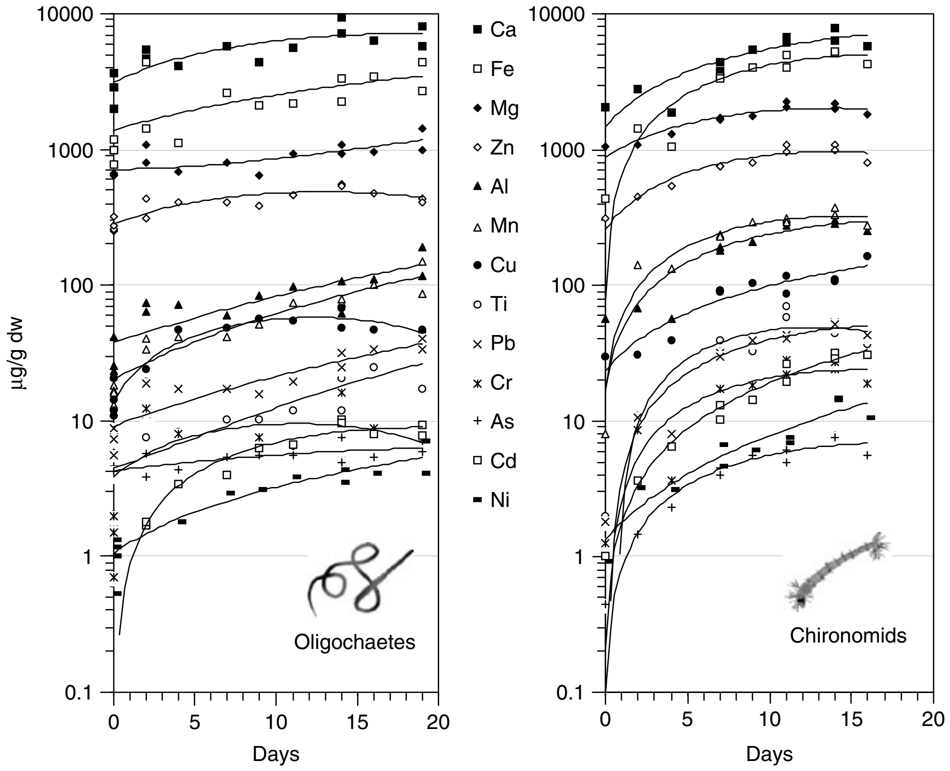

Figure 3. Bioaccumulation patterns of priority elements and heavy metals in two benthic species.

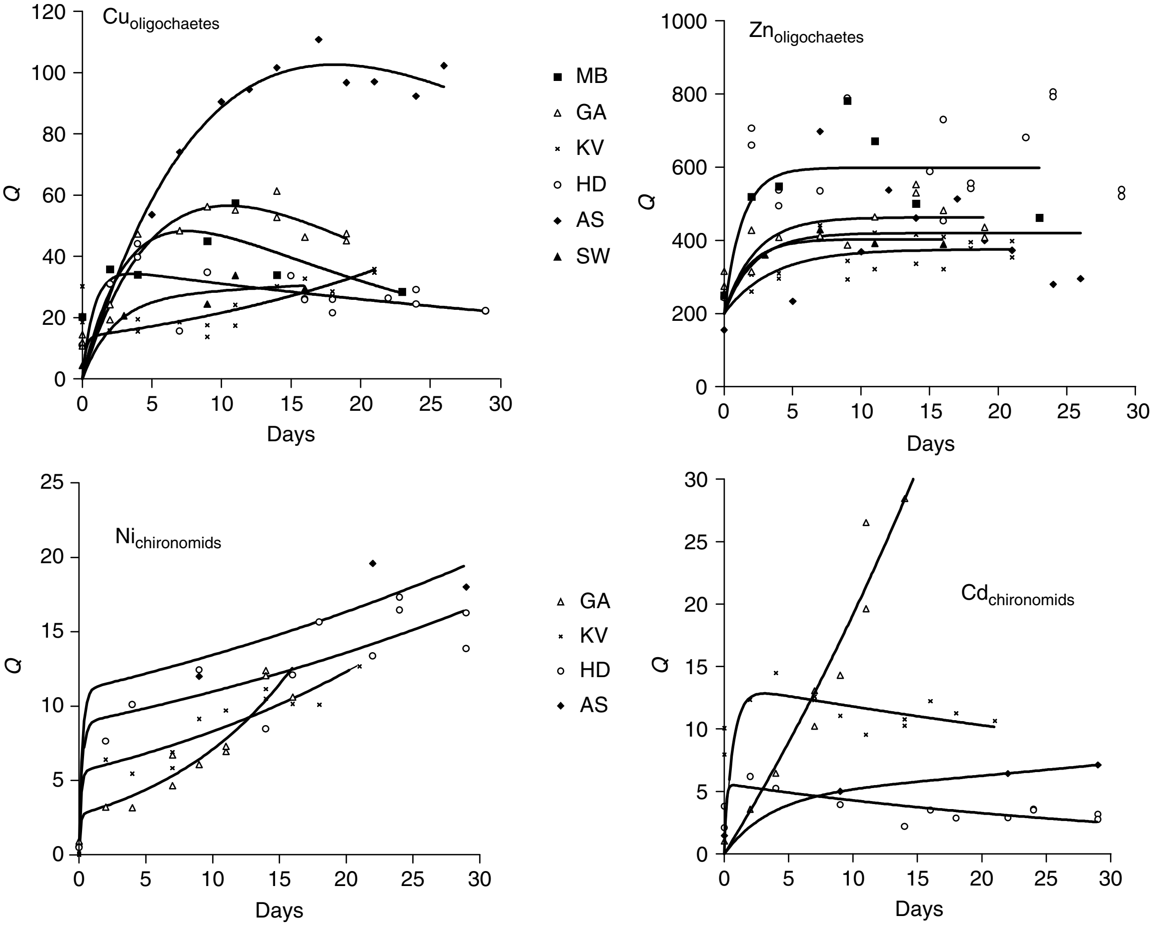

Figure 4. Examples of body concentrations in oligochaetes and chironomids with time. The solid line is the model prediction based on Equation 4 (Cd, Cu, Ni) and Equation 5 (Zn).

Table 2. Variance Between Locations: Accumulation Patterns in Oligochaetes[a] and Chironomids

Cd	MB	GA	KV	HD	AS	SW
MB		D	ns	ns	ns	ns
GA			D	D	D	D
KV		D		ns	ns	ns
HD		D	D		ns	ns
AS		D	D	ns		ns
SW						

Cu	MB	GA	KV	HD	AS	SW
MB		ns	D	ns	D	ns
GA			D	ns	D	D
KV		D		ns	D	ns
HD		D	ns		D	ns
AS		ns	ns	ns		D
SW						

Ni	MB	GA	KV	HD	AS	SW
MB		ns	ns	ns	D	ns
GA			ns	ns	ns	ns
KV		ns		ns	D	ns
HD		ns	D		D	ns
AS		ns	ns	ns		D

Pb	MB	GA	KV	HD	AS	SW
MB		D	D	D	D	D
GA			D	ns	ns	ns
KV		D		D	D	ns
HD		ns	D		ns	ns
AS		ns	ns	ns		ns
SW						

Zn	MB	GA	KV	HD	AS	SW
MB		D	D	ns	ns	D
GA			D	D	ns	D
KV		ns		D	ns	D
HD		D	D		D	D
AS		ns	ns	ns		D
SW						

As	MB	GA	KV	HD	AS	SW
MB		D	D	D	D	ns
GA			ns	D	D	ns
KV		D		D	ns	ns
HD		D	ns		ns	ns
AS						ns
SW						

Cr	MB	GA	KV	HD	AS	SW
MB		ns	ns	ns	D	ns
GA			ns	ns	D	ns
KV		ns		ns	ns	ns
HD		ns	D		D	ns
AS		ns	ns	ns		D
SW						

[a] Data for oligochaetes shaded gray; D, significantly different; ns, not significantly different.

Only those combinations with standard errors that were of the same magnitude or smaller than the values of k_1 and k_2 themselves were considered significant. This does not mean that other relationships between compartments, chemical species, and accumulation do not exist; this statistical criterion is necessary to identify the best model predictor, and therefore the most significant source of metals to biota.

The outcome of the modeling exercise and the statistical comparisons are summarized in Table 3. Chemical speciation and concentration shifts for the two compartments are shown on the left side. These were used as input for the biotic uptake models, for which results are shown on the right side. Only combinations where the statistical criterion was met ($sy \cdot x < k_1, k_2$) are shown. Obviously, the other combinations are thefore of less significance.

Negative k_0 values indicate a decrease in concentrations, positive values indicate an increase. Combinations of measured C_i and k_0 values were used as initial constants in the iteration procedure. The advantage of this approach is that variables that have a relatively large certainty or reliability do not participate in the iteration process as such. They do not need to be included, since $C(t)$, M^{2+}, and k_0 were actually measured with time. This is a large advantage, since the optimization procedure is now focused on the accurate estimation of the remaining variables k_1 and k_2, that is, the uptake and elimination rates.

The most important conclusion from Table 3 is that both benthic species appear to obtain their metals primarily from the overlying surface water and not from the sediment pore water compartment. Total concentrations in the sediment (mg/kg) are not indicative and do not reflect the magnitude of total body concentrations. Both test species accumulated the largest amounts of cadmium at site GA (chironomids even faced lethal concentrations), although sediment contents are second-lowest. In all cases, and therefore conclusively, the most important source for cadmium and nickel for oligochaetes is the free ion activity in the overlying surface water. In 100% of the cases, the measured C_{act} and k_0 combination for free ions in the surface water compartment yielded the best descriptors for accumulation patterns based on the two-compartment model. For copper, this seems valid in half of the cases. On some occasions, free ion concentrations dropped to a point where the contribution as a relevant source probably becomes negligible (see also Fig. 2). Potential sources of free Cu ions, such as humic complexes, have mostly slow dissociation kinetics that cannot make up for Cu-uptake rates by biota. Chemical dissociation is therefore rate-limiting.

Uptake from the sediment pore water was observed only occasionally, and solely, for lead. The source of Pb appears more variable, which is in agreement with the fact that the physiological mechanisms of Pb uptake are still unclear. It is occasionally suggested that uptake and elimination of Pb obey different mechanisms than in the case of the other heavy metals (i.e., an interaction between the free metal ion and a channel and/or carrier transport system in the external and/or internal epithelium), but this is yet to be demonstrated.

For zinc and arsenic, the performance of the one-compartment model was in 80% of the cases better than the model that used two compartments. External concentrations (i.e., outside the organism) obviously never became rate-limiting for uptake. In river sediments, Zn is mostly present in fair amounts, and the fraction of labile species, among them free ions, is mostly the largest of all metals. Most organisms regulate their Zn body concentrations very efficiently, and steady state is often quickly reached (see also Fig. 4). Dissolved arsenic does not yield significant free ion concentrations, being an oxyanion; the predominant species are arsenate or arsenite, and these are negatively charged.

It is also shown that uptake rates are not "universal constants." The variation of k_1 values over the locations may be significant. Kinetic processes therefore do indeed matter. The fact that the overlying water plays such a significant role in metal uptake should actually not be very surprising. Benthic organisms are often considered to be in close contact with sediment pore water, but species that live in burrows are actually in closer contact with overlying water because (1) the burrow inhibits direct contact between the organism and pore water, and (2) organisms exchange the water in their burrows with overlying water.

These presented findings provide a significant contribution to our understanding of chemical speciation of metals in water–sediment systems and subsequent metal

Table 3. Chemical Speciation in and Uptake from Different Compartments[a]

		Speciation in Surface Water					Speciation in Pore Water					Oligochaetes											Chironomids										
													Uptake from Surface Water				Uptake from Pore Water							Uptake from Surface Water				Uptake from Pore Water					
Metal, Site	Total, mg/kg	C_{tot}	k_0	C_{act}	k_0	% act	C_{tot}	k_0	C_{act}	k_0	% act	Q_{max}	k_1-C	$sy \cdot x$	k_1-act	$sy \cdot x$	k_1-C	$sy \cdot x$	k_1-act	$sy \cdot x$	k_2	r^2	Q_{max}	k_1-C	$sy \cdot x$	k_1-act	$sy \cdot x$	k_1-C	$sy \cdot x$	k_1-act	$sy \cdot x$	k_2	r^2
Cd																																	
MB	7.49	0.08	0.037	0.025	−0.277	32.1	0.05	0.033	0.01	0.157	9.6	2.2			20.16	0.01					0.03	0.47	nd										
GA	0.37	0.20	−0.003	0.050	0.360	25.0	0.30	0.038	0.07	0.010	23.3	10.0			18.45	6.65					0.03	0.94	29.0	8.20	0.01							−0.03	0.95
KV	0.14	0.23	0.005	0.060	−0.005	26.7	0.09	0.001	0.03	−0.001	34.8	0.7			0.31	0.11					0.05	0.58	11.0	401.0	4.65							8.37	0.13
HD	2.59	0.05	0.002	0.014	0.0003	30.0	0.03	0.001	0.01	0.001	28.8	2.0			48.07	0.01					0.20	0.71	6.3	208.0	26.30							1.34	0.37
AS	1.57	0.39	−0.004	0.053	0.0003	13.6	0.12	0.005	0.06	−0.001	53.4	0.6			53.50	0.05					8.18	0.99	7.4	2.42	1.17							0.13	0.88
SW	2.35	0.04	−0.001	0.003	0.0001	7.3	0.04	0.001	0.02	−0.001	51.2	0.9			10.27	3.15					3.47	0.85	nd										
Cu																																	
MB	93.1	4.48	−0.180	0.45	−0.380	10.0	2.51	−0.095	0.33	0.012	13.1	57.5			45.96	0.01					0.03	0.85	nd										
GA	11.4	6.00	0.094	0.70	0.088	11.7	5.38	0.095	2.00	0.096	37.2	55.0			20.01	5.34					0.09	0.84	140.0	3.25	2.27							−0.05	0.82
KV	4.9	6.56	−0.043	0.57	−0.036	8.7	10.17	−0.043	0.79	−0.043	7.8	35.0	37.25	0.01							16.76	0.41	51.0	2.85	0.92							0.18	0.31
HD	26.5	5.41	−0.100	1.23	−0.041	22.7	3.07	−0.070	0.12	−0.012	3.8	39.7	6.06	1.61							0.76	0.71	48.0	3.40	1.82							0.55	0.71
AS	25.7	2.74	−0.040	0.07	−0.003	2.6	1.90	−0.066	0.12	−0.006	6.3	103.0	21.25	17.71							1.24	0.55	152.0	6.65	4.82							0.33	0.78
SW	48.0	2.55	−0.082	0.49	−0.027	19.2	2.87	−0.063	0.70	−0.041	24.3	32.0			31.08	24.9					0.68	0.88	nd										
Ni																																	
MB	35.6	4.57	0.020	1.02	−0.042	22.3	2.69	0.039	0.50	−0.602	18.6	12.0			1.77	2.19					0.15	0.87	nd										
GA	6.60	1.00	0.081	0.05	0.024	5.0	3.00	0.200	0.05	0.021	1.7	5.1			7.34	2.44					0.01	0.75	14.7			14.96	4.71					−0.03	0.81
KV	8.35	10.27	−0.066	1.16	−0.012	11.3	16.29	−0.086	5.86	−0.086	36.0	11.0			0.44	0.36					−0.01	0.54	13.0			11.01	5.68					4.31	0.94
HD	19.4	4.82	0.016	0.73	−0.031	15.1	2.65	0.053	0.73	−0.012	27.4	7.6			6.41	2.56					0.42	0.51	17.3			8.60	3.05					0.57	0.94
AS	31.5	5.44	0.094	0.58	0.009	10.7	5.70	−0.048	1.97	−0.028	34.6	3.9			0.47	0.24					0.07	0.52	21.0	9.32	0.01							4.36	0.90
SW		6.17	0.027	1.27	0.055	20.6	5.16	−0.031	0.30	0.018	5.7	9.0			1.28	0.29					0.09	0.91	nd										
Pb																																	
MB	703.6	0.30	0.024	0.00	0.047	1.0	0.83	−0.015	0.14	0.080	16.8	135.6					27.20	0.01			0.21	0.10	nd										
GA	28.0	2.00	−0.040	0.10	−0.001	5.0	1.00	−0.028	0.05	0.016	5.0	35.0					10.68	5.88			0.54	0.71	51.0							105.3	18.73	0.05	0.92
KV	11.5	1.93	0.100	0.64	−0.007	33.2	1.39	−0.007	0.25	−0.005	17.8	21.0	0.53	0.18							−0.06	0.66	13.5	0.75	0.18							0.09	0.94
HD	39.8	1.74	−0.080	0.20	−0.009	11.8	1.33	−0.073	0.20	−0.019	14.7	26.4			74.82	18.75					0.44	0.68	32.7			38.70	7.25					0.29	0.96
AS	79.0	2.60	−0.081	0.17	0.000	6.5	2.51	−0.004	0.11	0.014	4.3	26.1							18.52	3.61	0.05	0.81	97.0							138.8	7.96	0.12	0.99
SW	91.5	3.70	0.135	0.56	−0.030	15.2	4.68	−0.155	0.44	−0.062	9.4	20.3	1.21	0.29							0.05	0.86	nd										
Zn																																	
MB	1091	20.60	−0.020	7.83	−0.034	38.0	10.96	0.114	2.78	0.050	25.4	781			30.57	0.01					0.39	0.82	nd										
GA	133.0	14.17	−0.144	3.01	−1.117	21.2	37.00	−0.262	19.00	−0.261	51.4	480	*0.26*	*0.15*								0.72	1100	*0.13*	*0.07*								0.87
KV	33.0	11.27	−0.006	4.47	−0.006	39.7	13.28	−0.006	4.00	−0.006	30.1	395	*0.19*	*0.12*								0.56	710	*8.69*	*1.36*								0.70
HD	231.0	21.70	−0.125	4.12	−0.043	19.0	15.96	0.048	4.51	−0.157	28.3	730	*29.23*	*0.01*								0.67	512	*5.08*	*7.77*								0.35
AS	320.5	33.40	0.091	4.80	0.018	14.4	64.70	−1.526	11.30	−0.396	17.5	530	*0.17*	*0.03*								0.95	765			17.83	4.78					0.06	0.95
SW	365	23.96	0.539	0.14	−0.017	0.6	18.81	0.216	0.43	0.377	2.3	429	*7.27*	*2.52*								0.72	nd										
As																																	
MB	nd	8.61	4.865				23.95	7.262				23.0	0.10	0.06								0.90	nd										
GA	5.6	44.27	0.949				71.57	0.518				7.4	*0.05*	*0.12*								0.54	7.5	*0.06*	*0.05*								0.87
KV	4.0	23.95	1.633				46.0	−0.072				10.0	*−0.09*	*0.05*								0.78	3.3	*0.67*	*0.15*								0.96
HD	6.5	42.44	−0.319				50.76	−0.378				10.1	*42.40*	*0.01*								0.12	4.0	0.73	0.55							12.42	0.91
AS	8.5	12.70	−0.323				5.90	0.599				14.6	*0.01*	*0.01*								0.96	nd										
SW	22.0	4.60	0.215				107.8	0.589				23.1	*−0.25*	*0.02*								0.99	nd										

[a] C, concentration (µg/L); tot, total dissolved; act, free ion activity; k_0, k_1, k_2 are rate constants (d^{-1}); Q_{max}, maximum C in biota (µg/g dw); r^2, correlation coefficient. Values printed in italic are derived with the one-compartment uptake model.

uptake by organisms. It is demonstrated that the presence of benthic organisms in a water–sediment system has a significant effect on chemical speciation, and this has to be considered when performing bioassays and chemical tests. Body concentrations are regulated by temporal variations in metal concentrations and chemical speciation. The measurements illustrate that biological kinetics (uptake) are for most metals faster than chemical kinetics (dissociation of less labile or complexed metals). Exposure can therefore only be estimated and modeled when time-related changes in metal concentrations are taken into account. The construction of generic models that are capable of transferring chemical information to a biological response has undoubtedly a high priority. Knowledge of the free metal ion activity is a significant step forward. Still, it may in cases be insufficient to predict the biological response, since one must also consider potential interactive effects of hydrogen ions (4,10) or effects of other macrochemicals (13). In any case, it is concluded that an integrated cell such as SOFIE provides the necessary experimental tool to support, in a mechanistic way, environmental risk assessments of contaminants. This knowledge may in time be incorporated in the derivation of better-founded quality criteria for sediments.

BIBLIOGRAPHY

1. Chapman, P.M. and Wang, F. (2001). *Environ. Toxicol. Chem.* **1**: 3–22.
2. DiToro, D.M. et al. (1992). *Environ. Sci. Technol.* **26**: 96–101.
3. Hare, L., Tessier, A., and Warren, L. (2001). *Environ. Toxicol. Chem.* **4**: 880–889.
4. Campbell, P.C.G. (1995). *Metal Speciation and Bioavailability in Aquatic Systems*. John Wiley & Sons, Hoboken, NJ.
5. Sunda, W.G. and Lewis, J.A.M. (1978). *Limnol. Oceanogr.* **23**: 870–877.
6. Giesy, J.P., Newell, A., and Leversee, G.J. (1983). *Sci. Total Environ.* **28**: 23–28.
7. Borgmann, U. and Charlton, C.C.J. (1984). *J. Great Lakes Res.* **10**: 393–401.
8. Daly, H.R., Jones, M.J., Hart, B.T., and Campbell, I.C. (1990). *Environ. Toxicol. Chem.* **9**: 1013–1018.
9. Meador, J.P. (1991). *Aquat. Toxicol.* **19**: 13–32.
10. Brown, P.L. and Markich, S.J. (2000). *Aquat. Toxicol.* **51**: 177–194.
11. Davison, W. and Zhang, H. (1994). *Nature* **367**: 546–548.
12. Vink, J.P.M. (2002). *Environ. Sci. Technol.* **36**: 5130–5138.
13. DeSchamphelaere, K.A.C. and Janssen, C.R. (2002). *Environ. Sci. Technol.* **36**: 48–54.

METHEMOGLOBINEMIA

Daniel Shindler
UMDNJ
New Brunswick, New Jersey

INTRODUCTION

Methemoglobinemia is a disorder of the oxygen-carrying hemoglobin in red blood cells. One cause of methemoglobinemia is contaminated well water. There are also many non-water-related causes. Individuals are particularly susceptible to the development of methemoglobinemia during early infancy and during pregnancy. Exposure to certain foods, drugs, and chemicals may cause methemoglobinemia. Methemoglobinemia can also occur in the absence of environmental exposure as part of several inherited disorders.

BLUE BABY SYNDROME AND WELL WATER

In 1945, Dr. Hunter Comly from Iowa City, Iowa described the blue baby syndrome. Two infants developed a bluish-gray skin discoloration (medically referred to as cyanosis) as a result of being fed infant formula diluted with well water.

The water from these wells was found to have high concentrations of nitrate. Once nitrate contamination was recognized as a cause of methemoglobinemia in infants, an additional 278 cases in 14 different states were reported by Walton in 1951. Yet, despite awareness and well water surveillance for nitrate and nitrite content, cases of methemoglobinemia continue to be reported.

High concentrations of nitrate in drinking water can cause high levels of methemoglobin in the body. Nitrite (converted from the ingested nitrate) oxidizes iron in the hemoglobin of red blood cells to form methemoglobin, which lacks hemoglobin's oxygen-carrying ability. Oxidation occurs when oxygen-carrying ferrous (Fe^{2+}) iron loses an electron and becomes ferric (Fe^{3+}). If unrecognized and untreated, rising levels of methemoglobin progressively decrease the ability of red blood cells to carry oxygen.

SUSCEPTIBLE INDIVIDUALS

Infants are particularly prone to develop methemoglobinemia from nitrate in drinking water until about 6 months of age. In early life, their ratio of fluid intake to body weight is greater than in adults. The stomach bacteria of infants are more capable of changing ingested nitrate to methemoglobin-begetting nitrite. Gastrointestinal problems that frequently occur in infants, such as diarrhea and vomiting, seem to further increase their predisposition to develop clinically significant methemoglobinemia.

Pregnant women also have an increased predisposition to develop methemoglobinemia. Increased methemoglobin levels during pregnancy may cause spontaneous abortion. During March 1993, the LaGrange County (Indiana) Health Department identified three women who reported a total of six spontaneous abortions during 1991–1993 and who resided in proximity to each other; each had obtained drinking water from nitrate-contaminated private wells. A fourth woman from another part of the county had two spontaneous abortions after she had moved into a new home with a nitrate-contaminated private well.

NITRATES IN DRINKING WATER

Nitrate contamination of well water may come from easily identifiable sources such as sewage disposal systems or nearby livestock facilities. Other times the cause of water contamination is less obvious, originating from distant landfills, fertilized cropland, parks, golf courses, lawns, and gardens. Episodes of methemoglobinemia continue to occur in rural farm settings especially with shallow (less than 100 feet deep) domestic wells in a setting where nitrogen-based fertilizers are heavily used. These wells may not be tested frequently enough for fluctuating concentrations of nitrate, which (according to the Environmental Protection Agency) should be no greater than 10 parts per million (ppm) at all times. Runoff of contaminated surface water from cultivated fields into a well may happen after a flood. This mode of well water contamination may be especially common when the well is too shallow, poorly located, or badly constructed.

OTHER WATER-RELATED CAUSES

Contaminated well water is not the only water-related cause of methemoglobinemia. Nor is the danger of undiagnosed methemoglobinemia limited to infants. Nitrite containing boiler fluid additives contaminating the faucet water can be responsible. An outbreak in New Jersey grammar school children in October 1992 was caused by dilution of canned soup with hot tap water. The water boiler had been serviced 5 months earlier with commercial conditioner fluid containing nitrite, and it had not been started until the morning of the incident. In another incident in 1996, six adult office workers developed methemoglobinemia after drinking coffee prepared with water from a hot water faucet connected to a boiler with a defective backflow-prevention valve. The conditioning fluid in this boiler also contained nitrites. Boiling water before using it does not decrease the level of existing nitrate or nitrite in the water. Prolonged boiling may actually increase the concentration of existing nitrate and nitrite in the water.

As water is almost invariably involved in food preparation, it is also important to recognize dietary causes of methemoglobinemia. Although vegetables are washed and cooked in water, nitrates may not be originating from the water used for rinsing, cooking, or serving the vegetables, but rather from the vegetables themselves. Broccoli, carrots, cauliflower, collard greens, onions, and spinach all contain some nitrate. Fortunately, the amount of vegetable nitrate in the normal diet is insufficient to cause significant methemoglobinemia in adults. Yet, ingestion of concentrated vegetables (carrot juice, carrot soup, spinach water) has resulted in methemoglobinemia in susceptible infants. Methemoglobinemia caused by excessive vegetable ingestion is not confined to humans. A herd of cows consumed approximately 20 kg of onions per cow per day for 6 weeks, which caused methemoglobinemia in the cows. Five cows died, and two had spontaneous abortions.

Diaper washing and reuse has also been implicated in methemoglobinemia. Anilines are dyes that are used in laundry ink. Aniline labeled diapers have caused infant methemoglobinemia by direct absorption through the skin and possibly by spreading to other poorly rinsed diapers in the same laundry water. Children's crayons no longer contain methemoglobin-producing dyes, but commercial markers such as red wax crayons may still contain them. Skin contact with dyed blankets and freshly dyed shoes can also potentially cause methemoglobinemia.

NON-WATER-RELATED CAUSES

Nitrates are found in the normal diet, not just in vegetables but also in meat and fish preservatives, and as meat color enhancers. Sodium and potassium nitrate have both caused life-threatening methemoglobinemia when they were mistaken for salt or used inappropriately in meat preparation.

Certain medications contain nitrates. Silver nitrate is used on the skin to treat burns. Nitroglycerin is administered orally, intravenously, under the tongue (sublingual), and on the skin in patients with heart disease. Amyl nitrite is an inhaled medication that is abused as a recreational drug. Amyl nitrite and isobutyl nitrite may be sold as "room deodorizers" to avoid drug laws. Nitrite can also be inhaled accidentally. Firefighters responding to a fire where isobutyl nitrite was stored developed methemoglobinemia because of exposure to the products of combustion.

Many non-nitrate medications have caused methemoglobinemia. Several antibiotics have caused methemoglobinemia during routine administration. Topical anesthetics are common during medical and dental procedures. There are many reports of methemoglobinemia after their routine administration at all ages. They may be given to teething infants and consequently get absorbed into the bloodstream through the mucous membranes of the mouth. Topical anesthetic ointments can also cause methemoglobinemia if they are absorbed through the skin. Anesthetic ointments given for diaper rash should be discontinued if the treated infant develops methemoglobinemia. Infants have also developed methemoglobinemia by accidentally ingesting a small number of tablets of a urinary tract analgesic called phenazopyridine.

Some industrial chemicals can cause methemoglobinemia. They may be ingested accidentally by infants and children when found in the household. Examples include naphthalene in mothballs, nitrobenzene in solvents such as gun cleaners, and copper sulfate in plant fungicides.

HEREDITARY CAUSES

Methemoglobinemia can also occur because of an inherited lack of an enzyme in red blood cells. The enzyme is called NADH diaphorase, or cytochrome b_5 reductase. Several mutations of the enzyme have been identified. The mode of transmission is autosomal recessive, which means that the gene abnormality has to be present in both parents to cause methemoglobinemia. When both parents are carriers of the trait, there is a 25% chance of a child inheriting abnormal genes from both parents, and therefore of developing methemoglobinemia. There is a 50% chance of a child inheriting one abnormal gene and becoming a carrier without actually developing methemoglobinemia. Many reports of such cases describe siblings derived from inbred populations.

Also, rare abnormalities of hemoglobin are inherited in an autosomal dominant fashion—a 50–50 chance of passing the disorder to the children. These inherited abnormal forms of hemoglobin are unstable and, as a result, are more readily transformed to methemoglobin.

DIAGNOSIS

Methemoglobinemia is potentially lethal if it is left undiagnosed. Transient blue or grayish skin discoloration of an infant may be mistakenly dismissed as being caused by changes in temperature. Familiarity with methemoglobinemia and a high index of suspicion by the physician are potentially lifesaving. Untreated infants can become irritable and lethargic. Some may develop seizures. The heart rate may become slow. The blood pressure may drop. They may eventually lapse into a coma and die.

Isolated cases of methemoglobinemia are rare. The diagnosis of this potentially life-threatening disorder can only be made if a physician is familiar with the clinical clues and manifestations. A patient who is visibly cyanotic will initially be suspected to have a primary cardiac or a respiratory problem. Consequently, oxygen is routinely administered to patients with cyanosis. The lack of a response to the oxygen provides a valuable clue to the diagnosis. Patients with methemoglobinemia will remain blue or ashen even after oxygen administration.

Pulse oximetry will be inaccurate and unreliable in these patients. The pulse oximeter is a simple finger-clip device containing a light sensor. It was designed to measure the percentage of oxygen saturation of hemoglobin by measuring the amount of light passing through the finger or earlobe. Methemoglobin absorbs light at wavelengths similar to normal hemoglobin, and as a result, oximetry cannot reliably distinguish the two.

The blood from a patient with methemoglobinemia will be dark red when drawn and remain dark even when exposed to room air. Dark blood drawn from a vein of a patient with heart or lung disease will turn bright red on exposure to the oxygen in room air or after bubbling pure oxygen through the blood.

Methemoglobinemia should also immediately come to mind if multiple cases of cyanosis are discovered. An environmental exposure may then be a likely cause.

As soon as the diagnosis of methemoglobinemia is suspected, a blood methemoglobin level should be measured to confirm the diagnosis. Patients will appear cyanotic when the methemoglobin level is approximately 10% or greater.

TREATMENT

Normal methemoglobin levels in the blood range from 1% to less than 3%. If blood methemoglobin levels are only mildly elevated (less than 20%), patients are observed and given oxygen to saturate the remaining normal hemoglobin.

In more serious cases, where methemoglobin levels are greater than 30%, patients are given intravenous methylene blue. Patients with underlying conditions such as heart problems or severe anemia may get treated with methylene blue even with a lower methemoglobin level.

Although at first glance giving blue dye to a blue patient seems counterintuitive, the treatment is simple and can be life saving. Methylene blue acts as a nonenzymatic electron carrier. It accelerates the rate with which red blood cells convert methemoglobin back to oxygen-carrying hemoglobin. Large doses of methylene blue can indeed cause methemoglobinemia, so if there is no response to the first dose, it is usually only repeated once.

In some patients, however, even a single dose of methylene blue may pose a danger. Patients with a deficient glucose-6-phosphate dehydrogenase enzyme system will not respond to the methylene blue, and their red cells may get destroyed (hemolyzed). In such cases, other more specialized treatments such as hyperbaric oxygen, or packed red blood cell transfusions, may be necessary.

In cases of environmental exposure, such as contaminated well water, it is equally important to identify the excess nitrate source and to prevent repeat exposure to it by the susceptible individual.

READING LIST

Well Water

Comly, H. (1945). Cyanosis in infants caused by nitrates in well water. *JAMA* **129**: 112–116.

Reprinted as a landmark article in *JAMA* **257**: 2788–2792 (1987).

Walton, G. (1951). Survey of literature relating to infant methemoglobinemia due to nitrate-contaminated water. *Am. J. Public Health* **41**: 987–995.

Bucklin, R. and Myint, M.K. (1960). Fatal methemoglobinemia due to well water nitrates. *Ann. Intern. Med.* **52**: 703–705.

Carlson, D.J. and Shapiro, F.L. (1970). Methemoglobinemia from well water nitrates: a complication of home dialysis. *Ann. Intern. Med.* **73**: 757–759.

Gastrointestinal Problems

Yano, S.S., Danish, E.H., and Hsia, Y.E. (1982). Transient methemoglobinemia with acidosis in infants. *J. Pediatr.* **100**: 415–417.

Lanir, A., Borochowitz, Z., and Kaplan, T. (1986). Transient methemoglobinemia with acidosis and hyperchloremia in infants. *Am. J. Pediatr. Hematol. Oncol.* **8**: 353–355.

Smith, M.A., Shah, N.R., Lobel, J.S. et al. (1988). Methemoglobinemia and hemolytic anemia associated with *Campylobacter jejuni* enteritis. *Am. J. Pediatr. Hematol. Oncol.* **10**: 35–38.

Kay, M.A., O'Brien, W., Kessler, B. et al. (1990). Transient organic aciduria and methemoglobinemia with acute gastroenteritis. *Pediatrics* **85**: 589–592.

Murray, K.F. and Christie, D.L. (1993). Dietary protein intolerance in infants with transient methemoglobinemia and diarrhea. *Pediatrics* **122**: 90–92.

Hanukoglu, A. and Danon, P.N. (1996). Endogenous methemoglobinemia associated with diarrheal disease in infancy. *J. Pediatr. Gastroenterol. Nutr.* **23**: 1–7.

Pregnancy

MMWR (1996). Spontaneous abortions possibly related to ingestion of nitrate-contaminated well water—LaGrange County, Indiana, 1991–1994. *Morb. Mortal. Wkly. Rep.* **45**: 569–572.

Boiler Fluid

Askew, G.L., Finelli, L., Genese, C.A. et al. (1994). An outbreak of methemoglobinemia in New Jersey in 1992. *Pediatrics* **94**: 381–384.

Nitrate in Vegetables

Smeenk, C. and van Oudheusden, A.P. (1968). Hemoglobinemia caused by spinach water. *Ned. Tijdschr. Geneeskd.* **112**: 1378–1379.

Keating, J.P. et al. (1973). Infantile methemoglobinemia caused by carrot juice. *N. Engl. J. Med.* **288**: 824–826.

Chan, T.Y. (1996). Food-borne nitrates and nitrites as a cause of methemoglobinemia. *Southeast Asian J. Trop. Med. Public Health* **27**: 189.

Sanchez-Echaniz, J., Benito-Femandez, J., and Mintegui-Raso, S. (2001). Methemoglobinemia and consumption of vegetables in infants. *Pediatrics* **107**: 1024–1028.

Bryk, T., Zalzstein, E., and Lifshitz, M. (2003). Methemoglobinemia induced by refrigerated vegetable puree in conjunction with supraventricular tachycardia. *Acta Paediatr.* **92**: 1214–1215.

Rae, H.A. (1999). Onion toxicosis in a herd of beef cows. *Can. Vet. J.* **40**(1): 55–57.

Dyes and Crayons

Graubarth, J., Bloom, C.J., Coleman, F.C., and Solomon, H.N. (1945). Dye poisoning in the nursery: a review of seventeen cases. *JAMA* **128**: 1155.

Howarth, B.E. (1951). Epidemic of aniline methaemoglobinaemia in newborn babies. *Lancet* **1**: 934–935.

Rieders, F. and Brieger, H. (1953). Mechanism of poisoning from wax crayons. *JAMA* **151**: 1490.

Kearney, T.E., Manoguerra, A.S., and Dunford, J.V. Jr. (1983). Chemically induced methemoglobinemia from aniline poisoning. *West. J. Med.* **140**: 282.

Poisoning from Food Preparation

Roueche, B. (1947). *Eleven Blue Men*. Little Brown & Co., Boston, MA.

Orgeron, J.D. et al. (1957). Methemoglobinemia from eating meat with high nitrite content. *Public Health Rep.* **72**(3): 189–193.

Singley, T.L. III. (1963). Secondary methemoglobinemia due to the adulteration of fish with sodium nitrite. *Ann. Intern. Med.* **57**: 800–803.

Silver Nitrate for Burns

Cushing, A.H. and Smith, S. (1967). Methemoglobinemia with silver nitrate therapy of a burn. *J. Pediatr.* **74**: 613–615.

Inhalation

Guss, D.A., Normann, S.A., and Manoguerra, A.S. (1985). Clinically significant methemoglobinemia from inhalation of isobutyl nitrite. *Am. J. Emerg. Med.* **3**: 46–47.

Hovenga, S., Kcenders, M.E., van der Werf, T.S. et al. (1996). Methemoglobinemia after inhalation of nitric oxide for treatment of hydrochlorothiazide-induced pulmonary edema. *Lancet* **348**: 1035–1036.

Nitroglycerin and Amyl Nitrite

Gibson, G.R. et al. (1982). Methemoglobinemia produced by high-dose intravenous nitroglycerin. *Ann. Intern. Med.* **96**: 615.

Paris, P.M., Kaplan, R.M., Stewart, R.D., and Weiss, L.D. (1986). Methemoglobin levels following sublingual nitroglycerin in human volunteers. *Ann. Emerg. Med.* **15**: 171.

Forsyth, R.J., Moulden, A. (1991). Methaemoglobinaemia after ingestion of amyl nitrite. *Arch. Dis. Child* **66**: 152.

Antibiotics

Damergis, J.A., Stoker, J.M., and Abadie, J.L. (1983). Methemoglobinemia after sulfamethoxazole and trimethoprim. *JAMA* **249**: 590.

Koirala, J. (2004). Trimethoprim-sulfamethoxazole—induced methemoglobinemia in an HIV-infected patient. *Mayo. Clin. Proc.* **79**: 829–830.

Moreira, V.D. et al. (1998). Methemoglobinemia secondary to clofazimine treatment for chronic graft-versus-host disease. *Blood* **92**: 4872.

Reiter, W.M. and Cimoch, P.J. (1987). Dapsone-induced methemoglobinemia in a patient with P. carinii pneumonia and AIDS. *N. Engl. J. Med.* **317**: 1740.

Coleman, M.D., Rhodes, L.E., Scott, A.K. et al. (1993). The use of cimetidine to reduce dapsone-dependent methaemoglobinaemia in dermatitis herpetiformis patients. *Br. J. Clin. Pharmacol.* **34**: 244.

Wagner, A., Marosi, C., Binder, M. et al. (1995). Fatal poisoning due to dapsone in a patient with grossly elevated methaemoglobin levels. *Br. J. Dermatol.* **133**: 816.

Sin, D.D. and Shafran, S.D. (1996). Dapsone- and primaquine-induced methemoglobinemia in HIV infected individuals. *J. Acquir. Immune. Defic. Syndr. Hum. Retrovirol.* **12**: 477–481.

Ward, K.E. and McCarthy, M.W. (1998). Dapsone-induced methemoglobinemia. *Ann. Pharmacother.* **32**: 549–553.

Mandrell, B.N. and McCormick, J.N. (2001). Dapsone-induced methemoglobinemia in pediatric oncology patients. *J. Pediatr. Oncol. Nurs.* **18**: 224–228.

Brabin, B.J., Eggelte, T.A., Parise, M., and Verhoeff, F. (2004). Dapsone therapy for malaria during pregnancy: maternal and fetal outcomes. *Drug. Saf.* **27**: 633–648.

Topical Anesthetics

Duncan, P.G. and Kobrinsky, N. (1983). Prilocaine-induced methemoglobinemia in a newborn infant. *Anesthesiology* **59**: 75.

Collins, J.F. (1990). Methemoglobinemia as a complication of 20% benzocaine spray for endoscopy. *Gastroenterology* **98**: 211.

Postiglione, K.F. and Herold, D.A. (1992). Benzocaine-adulterated street cocaine in association with methemoglobinemia. *Clin. Chem.* **38**: 596.

Brown, C.M., Levy, S.A., and Sussan, P.W. (1994). Methemoglobinemia: life-threatening complication of endoscopy premedication. *Am. J. Gastroenterol.* **89**: 1108–1109.

Guerriero, S.E. (1997). Methemoglobinemia caused by topical benzocaine. *Pharmacotherapy* **17**: 1038.

Cooper, H.A. (1997). Methemoglobinemia caused by benzocaine topical spray. *South Med. J.* **90**: 946.

Brisman, M. et al. (1998). Methaemoglobin formation after the use of EMLA cream in term neonates. *Acta Paediatr.* **87**: 1191.

Balicer, R.D. and Kitai, E. (2004). Methemoglobinemia caused by topical teething preparation: a case report. *Sci. World J.* **4**: 517–520.

Moore, T.J., Walsh, C.S., and Cohen, M.R. (2004). Reported adverse event cases of methemoglobinemia associated with benzocaine products. *Arch. Intern. Med.* **164**: 1192–1196.

Phenazopyridine Overdose

Nathan, D.M., Siegel, A.J., and Bunn, H.E. (1977). Acute methemoglobinemia and hemolytic anemia with phenazopyridine: possible relation to acute renal failure. *Arch. Intern. Med.* **137**: 1636–1638.

Gavish, D. et al. (1986). Methemoglobinemia, muscle damage and renal failure complicating phenazopyridine overdose. *Isr. J. Med. Sci.* **22**: 45.

Christensen, C.M., Farrar, H.C., and Kearns, G.L. (1996). Protracted methemoglobinemia after phenazopyridine overdose in an infant. *J. Clin. Pharmacol.* **36**: 112.

Gold, N.A. and Bithoney, W.G. (2003). Methemoglobinemia due to ingestion of at most three pills of pyridium in a 2-year-old: case report and review. *J. Emerg. Med.* **25**: 143–148.

Other Causes of Methemoglobinemia

Banning, A., Harley, J.D., O'Gorman Hughes, D.W. et al. (1965). Prolonged methaemoglobinaemia after eating a toy-shaped deodorant. *Med. J. Aust.* **2**: 922–925.

Chilcote, R.R., Williams, B., Wolff, L.J. et al. (1977). Sudden death in an infant from methemoglobinemia after administration of "sweet spirits of nitre." *Pediatrics* **59**: 280–282.

Sherman, J.M. and Smith, K. (1979). Methemoglobinemia owing to rectal-probe lubrication. *Am. J. Dis. Child* **133**: 439–440.

Shessler, R., Dixon, D., Allen, Y. et al. (1980). Fatal methemoglobinemia from butyl nitrite ingestion. *Ann. Intern. Med.* **92**: 131–132.

Ellis, M., Hiss, Y., and Shenkman, L. (1992). Fatal methemoglobinemia caused by inadvertent contamination of a laxative solution with sodium nitrite. *Isr. J. Med. Sci.* **28**: 289.

Mallouh, A.A. and Sarette, W.O. (1993). Methemoglobinemia induced by topical hair oil. *Ann. Saudi. Med.* **13**: 78–88.

Osterhoudt, K.C., Wiley, C.C., Dudley, R. et al. (1995). Rebound severe methemoglobinemia from ingestion of a nitroethane artificial fingernail remover. *J. Pediatr.* **126**: 819–821.

Methylene Blue Treatment

Clifton, J. II and Leikin, J.B. (2003). Methylene blue. *Am. J. Ther.* **10**: 289–291.

Goluboff, N. and Wheaton, R. (1961). Methylene blue induced cyanosis and acute hemolytic anemia complicating the treatment of methemoglobinemia. *J. Pediatr.* **58**: 86–89.

Beutler, E. and Baluda, M.C. (1963). Methemoglobin reduction. Studies of the interaction between cell populations and of the role of methylene blue. *Blood* **22**: 323.

Ng, L.L., Naik, R.B., and Polak, A. (1982). Paraquat ingestion with methaemoglobinaemia treated with methylene blue. *BMJ* **284**: 1445.

Harvey, J.W. and Keitt, A.S. (1983). Studies of the efficacy and potential hazards of methylene blue therapy in aniline-induced methaemoglobinaemia. *Br. J. Haematol.* **154**: 29.

Fitzsimons, M.G., Gaudette, R.R., and Hurford, W.E. (2004). Critical rebound methemoglobinemia after methylene blue treatment: case report. *Pharmacotherapy* **24**: 538–540.

Khanduri, U., Sharma, A., and Kumar, P. (2004). Unique red cell abnormality in a case of fatal methylene blue induced haemolysis. *Hematology* **9**: 239–241.

Rosen, P.J., Johnson, C., McGehee, W.G., and Beutler, E. (1971). Failure of methylene blue treatment in toxic methemoglobinemia. Association with glucose-6-phosphate dehydrogenase deficiency. *Ann. Intern. Med.* **75**: 83–86.

Hibbard, B.Z. Jr. et al. (1978). Severe methemoglobinemia in an infant with glucose-6-phosphate dehydrogenase deficiency. *J. Pediatr.* **93**: 816–818.

Red Blood Cell Exchange

Golden, P.J. and Weinstein, R. (1998). Treatment of high-risk, refractory acquired methemoglobinemia with automated red blood cell exchange. *J. Clin. Apheresis.* **13**: 28–31.

MICROBIAL ACTIVITIES MANAGEMENT

Peter D. Franzmann
Luke R. Zappia
CSIRO Land and Water
Floreat, Australia

INTRODUCTION

Microbial activity can be either detrimental or beneficial to the quality of water in water treatment and distribution systems. The growth of microbial community biofilms on the pipe walls in potable water distribution systems, and in water supply reservoirs, can significantly reduce the aesthetic, microbiological, and chemical quality of the water. Microbial processes can also be beneficial in processes such as biofiltration, to prevent reactions that occur in the distribution system that are otherwise detrimental to the water quality.

Normally more than 99.9% of microbial cells in water distribution systems occur in biofilms rather than in the planktonic form. Pipe wall biofilm communities can act as reservoirs of opportunistic pathogenic microorganisms as bacteria embedded in biofilm matrices are protected from water disinfectants (1). Biofilm communities may also impact the chemical and aesthetic quality of distributed waters by concentrating, oxidizing, and releasing metals that can result in "dirty water" problems (2), producing taste and odor compounds (3), enhancing corrosion (4), and reducing disinfectant residuals in chlorinated (5) and chloraminated distribution systems (6).

GENERAL CONTROL OF BIOFILM GROWTH AND ACTIVITY

Many water treatment processes are focused on the management of biofilm growth and activity in distribution systems generally, and more specifically on the control of biofilm processes such as nitrification and manganese oxidation. The parameters that can manage biofilm growth and activity exert their control collectively in management strategies, but their contributions to management of microbial activity are best understood by examination of their individual effects. In a comparative study on the control of microbial activity in four different distribution systems in the United States and France, the authors concluded that it was always difficult to make definitive conclusions from full-scale systems where no operating parameter was truly controlled (7); however they emphasized that maintenance of a disinfectant residual was the strongest controller of bacterial growth in their systems. On the other hand, other studies have suggested that with careful management and appropriate treatments, a disinfectant residual is not required (8).

Control of Biodegradable Organic Carbon in Distributed Water

Control of microbial activity in distribution systems is required so that the water quality meets community health and aesthetic expectations. It can be achieved first by distribution of "biologically stable water," that is, water that will not promote microbial growth and activity in the distribution system, and second by avoiding the use of pipe materials that leach substrates that can support microbial growth (9). Biologically stable water can be produced without the use of a disinfectant by reducing the concentration of biodegradable organic carbon (10) and other growth-promoting compounds such as methane, ammonium, or reduced sulfur compounds, in the water that leaves the treatment plant.

Types of Organic Carbon in Drinking Water. For the distribution of biologically stable water, water treatment plants aim to reduce dissolved organic carbon (DOC), specifically the fractions of the DOC that are consumed by microbes. In practice, two types of DOC bioassay are used at treatment plants to determine if biostable water is being produced, and each assay measures different but related fractions of the bioreactive DOC.

The first of these fractions, the biodegradable dissolved organic carbon (BDOC), is measured by the decrease of DOC over time, usually up to 10 days, after the water has been inoculated with washed biomass from a sand filter (11) or a microbial population that has been grown on glass beads. BDOC is the measure of the change in DOC because of its consumption by a mixed microbial community until that consumption effectively stops. The remaining DOC is considered refractory to biodegradation in reasonable time. Methods for the measurement of BDOC have been detailed by Joret and Lévy (11) and Servais et al. (12), and rapid flow-through methods for the determination of BDOC have been proposed (13).

The other bioreactive fraction of the DOC is the assimilable organic carbon (AOC) and is the fraction used by cells to build biomass. Most water treatment plants measure this parameter by the method developed by van der Kooij (8). In this method, AOC is measured by the increase in viable numbers of two strains of bacteria, *Pseudomonas fluorescens* strain P17 and *Spirillum* sp. strain NOX, over time in water that was pasteurized at 60 °C for 30 min before inoculation. The amount of growth is compared with a standard curve of growth of both strains in AOC-free water to which standard additions of a carbon source, usually acetate, are added. In most water samples, growth reaches the stationary phase in about 10 days. The growth of each strain is taken as the difference between the number of viable cells in the water at completion of growth and the number of cells in the water just after inoculation. AOC is mainly composed of organic matter with a molecular weight of <1000 (14) and is expressed in the units μg (acetate-C equivalents)$\cdot L^{-1}$. Other methods for AOC measurement have been described, using increase in turbidity or ATP as indicators of biomass production (13). As the water industry is increasingly demanding standardized methods so that parameters from different plants can be compared, most AOC determinations now use the van der Kooij method, which is the method detailed in the entry Standard Methods for the Examination of Water and Wastewater.

AOC and BDOC can be measured in chlorine-containing waters only if the chlorine is reduced with thiosulfate by the application of about 2 moles of thiosulfate per mole of chlorine. The ratio of BDOC or AOC to DOC varies in untreated surface and aerated groundwaters, but BDOC usually represents about 10% to 30% of the DOC, and AOC usually accounts for 1% to 3% of the DOC (15). In treated waters, the correlation of AOC or BDOC to DOC concentration is poor (16). van der Kooij (8) showed that distribution of water with an AOC concentration below $10\ \mu g \cdot L^{-1}$ without a disinfectant limited the mean heterotrophic plate count to below 100 colony-forming units$\cdot mL^{-1}$, a guideline value considered by many regulators to indicate water of good microbiological quality. Biologically stable water does not promote growth or activity of microorganisms to a significant extent. Various parameter values have defined biological stability of distributed waters, and some of these are compared in Table 1.

For measures of biostability, Escobar and Randall (19) argued that both AOC and BDOC measures should be used. BDOC has a detection limit of about $100\ \mu g \cdot L^{-1}$, whereas for AOC, the detection limit is about $10\ \mu g \cdot L^{-1}$.

Table 1. Various Parameter Values that Define Biologically Stable Water

Attribute	Value	Comment	Reference
AOC	$10\ \mu g \cdot L^{-1}$	AOC uptake by biofilm limited at $<10\ \mu g \cdot L^{-1}$	9
AOC	$<50\ \mu g \cdot L^{-1}$	For AOC in a system with a chlorine residual	17
BDOC	$200\ \mu g \cdot L^{-1}$	In water with no disinfectant residual	18
BDOC	$500\ \mu g \cdot L^{-1}$	In water <15 °C with no disinfectant residual	18
BDOC	$<500\ \mu g \cdot L^{-1}$	In water with a chlorine residual of 0.1 mgd$\cdot L^{-1}$	18
BDOC	$\leq 150\ \mu g \cdot L^{-1}$	Water without a disinfectant	10

Some treatments like nanofiltration can reduce BDOC almost completely, but considerable AOC can remain in the permeate. Measurement of AOC alone does not give information on the pool of hydrolyzable BDOC than can form AOC in the distribution system; thus, measurement of AOC alone can give an overestimate of the biostability of distributed water (19).

The Effect of Treatment Processes on AOC and BDOC. As the quantity and type of dissolved organic carbon in water is the major factor that promotes bacterial growth (and hence activity) in distribution systems, modern water treatment plants usually investigate the effect of different water treatment methods on the quantity and type of DOC in their treated water. Some treatments, such as the use of oxidants for destruction of pesticides in treatment plant influent, or chlorine for disinfection, can partially oxidize the DOC to form biologically available AOC or BDOC. Other processes, notably those that use methods of biological treatment such as slow sand filtration and biofiltration using a granulated activated carbon support (GAC), can produce water with low AOC and BDOC.

A range of treatments and their effects on AOC and BDOC concentrations are shown in Table 2. The data in Table 2 oversimplify the outcomes of the different treatment options, as combinations of treatment options can have different outcomes, individual treatment options can be modified in many ways, and the success of each treatment type is dependent on the nature of the preceding treatments that were used. For example, in one study, there was 57% less AOC in finished waters from GAC filters than for anthracite sand filters when using polymer as the coagulant, whereas when using ferric ions as the coagulant, that improvement was only 13% [derived from the data of Table 4 of Volk and LeChevallier (20)]. The data in Table 2 provides a list of site-specific examples of the effect of different treatments on AOC and BDOC.

Table 2. Some Examples of the Effects of Water Treatments on DOC, BDOC, and AOC Concentrations. The Examples are Site Specific, as the Outcome from the Use of any Treatment is Dependent on Many Water Parameters, Especially the Nature of the Natural Organic Matter in the Receiving Water, and on the Pretreatment Processes Employed

Treatment	Effect	Reference
Ozonation	↑AOC 560%	23
Enhanced coagulation	↓29% DOC, ↓38% BDOC, no change AOC	24
Ultrafiltration	↓85% DOC, ↓76% to ↑8.7% AOC[a]	24
Alum coagulation	↓16% AOC, ↓49% BDOC	25
Two-stage GAC biofilter	↓51–72% AOC for ozonated water	26
Two-stage GAC biofilter	↓37–70% AOC for non-ozonated water	26
Biofiltration	↓80 to 90% AOC depending on concentration	27
Slow sand filtration	↓48% AOC	21
Chlorination	↑80% AOC	21
MIEX treatment	↓>88% BDOC, ↓24% AOC, ↓> 95% when filtered	Our data

↑= increase, ↓= decrease.

[a] Dependent on water source.

Choice of a treatment option to obtain a desired water quality will require testing by each treatment plant using its receiving water, and its implementation will require some optimization of the process parameters that can be manipulated. For substantial removal of AOC and BDOC, some form of biological treatment is generally required (21), although initial results with a magnetic ion exchange treatment for DOC removal, MIEX (WesTech Engineering, Salt Lake City, Utah), has shown exceptionally good removal of BDOC, and AOC when subsequently filtered (Table 2), in addition to removal of 60–70% of the DOC (22).

AOC concentrations in the distribution system are dependent not only on the water type and treatment system but also on the choice of materials used. Materials in contact with the water can leach considerable amounts of AOC and affect the biofilm growth on pipe surfaces. For example, nonplastic polyvinyl chloride supports an order of magnitude less biofilm than does high-density polyethylene (9).

Although the concentration of the biologically available DOC fraction is arguably the most important determinant of microbial activity in distribution systems, it does not act in isolation, and its effects are modulated by other parameters that control microbial activity in the distribution system, such as disinfectant type and concentration, temperature, residence time, and concentrations of nonorganic oxidizable microbial substrates such as ammonium or manganese. Mathematical models that attempt to explain those interactions have been developed (e.g., Ref. 28), as have relatively simple, descriptive, decision-support criteria (16).

Disinfection to Control Microbial Activity

The maintenance of a disinfectant residual is probably the strongest inhibitor of microbial activity through a distribution system (7). High concentrations of disinfectant, even in the presence of relatively high concentrations of dissolved organic carbon or AOC, can limit microbial activity such as biofilm growth in distribution systems, e.g., $160\ \mu g \cdot L^{-1}$ AOC with a chloramine residual of $>2\ mg \cdot L^{-1}$, or $107\ \mu g \cdot L^{-1}$ AOC with $>0.8\ mg \cdot L^{-1}$ free chlorine [in Pine Hills Distribution System; (7)]. However, as high concentrations of disinfectant are the most common cause of customer taste or odor complaints (29), and as there are increasing concerns about disinfection by-products (30), the goal of water treatment utilities is often to deliver water with a minimal amount of disinfectant residual, often with reduced DOC, to control microbial activity and maintain microbiologically safe water.

As with treatment methods for removal of AOC and BDOC, many different concentrations of disinfectant have been shown to prevent biofilm growth, or specific microbial activities depending on the water quality, temperature, residence time, or type of disinfectant used. As already discussed, if the water quality is high (i.e., biologically stable), in some cases, no disinfectant is required to deliver safe and aesthetically acceptable water to the consumer (9).

Table 3. Disinfectant Residual Concentrations and Their Effects on Microbial Activity in Pilot-Scale and Actual Water Distribution Systems

Treatment: Effect	Reference
Free chlorine 0.05 $mg{\cdot}L^{-1}$: no biofilm growth initiation in pipes	33
Free chlorine > 0.5 $mg{\cdot}L^{-1}$ or chloramine > 1.0 $mg{\cdot}L^{-1}$: required to decrease probability of coliform occurrence	16
Chloramine 2.1 $mg{\cdot}L^{-1}$: inhibited bacterial activity	
Total chlorine 0.3 $mg{\cdot}L^{-1}$: prevents establishment of biofilms	Our data
Total chlorine < 0.02 $mg{\cdot}L^{-1}$: allows establishment of biofilms	Our data
Chlorine 1.0–3.0 $mg{\cdot}L^{-1}$: disrupts established biofilms	34
Free chlorine residual 3 $mg{\cdot}L^{-1}$: no effect on biofilms in iron pipes	17
Chlorine or chloramine 1 $mg{\cdot}L^{-1}$: inactivates bacteria on copper PVC or galvanized pipes	17
Free chlorine 0.2 $mg{\cdot}L^{-1}$: reduced biofilm concentration	35

Although site and water specific, some examples of recommended disinfectant levels to prevent biofilm growth or microbial activity in distribution systems are given in Table 3. In reality, the target level of a disinfectant residual should be decided only after appropriate testing. Once established, biofilms can exert a chlorine demand that increases linearly with the BDOC concentration in the water and with the ratio of the surface area to volume of the biomass, whereas the chlorine demand of detached biomass is related to the amount of biomass (5). The effectiveness of a disinfectant residual in controlling microbial activity is also markedly affected by pipe material, retention time of the water in the distribution system, and water chemistry (17).

Chloramine has been shown to provide better control of microbial activity than has chlorine (31) which may be related to slower decay and greater penetration into biofilms (32).

Temperature and Microbial Activity

Temperature has a marked effect on microbial activity. Generally, microbial growth and microbial processes such as nitrification in distribution systems are limited at temperatures of less than 15 °C, and deterioration of water quality in colder climates can be temperature limited (16).

Control of Nitrification in Distribution Systems

Nitrification, the oxidation of ammonium through nitrite to nitrate, is a microbial activity that is problematic for many water utilities that distribute water with chloramine as the disinfectant. The activity depletes the residual chloramine and can make conditions more conducive to growth of heterotrophic bacteria in the distribution system (36). Ammonium in source waters can lead to biological instability in distribution systems if not treated. For example, ammonium concentrations throughout Illinois averaged 0.62 $mg\ N{\cdot}L^{-1}$, which equates to an oxygen demand of 2.8 $mg\ O_2{\cdot}L^{-1}$ (15).

For distribution systems that do not use chloramine as the disinfectant, ammonium can be removed economically from source waters by nitrifying bacteria. Attached biofilm processes such as GAC-biofiltration, fluidized bed filtration, or rapid sand filtration of nonchlorinated water are generally used because of the slow growth rates of nitrifying bacteria (15). In this way, microbial activity that would otherwise occur in the distribution system is used in a treatment process at the plant to convert ammonium to nitrate. Larger support media (1 mm diameter) should be used in nitrifying biofilters to lessen sheer stresses on slow-growing populations of nitrifying bacteria (15).

In chloraminated systems, nitrification needs to be controlled to prevent the loss of the disinfectant with subsequent increase in biofilm activity. In cooler climates, nitrification may be controlled for much of the year without intervention, as nitrification usually occurs when water temperatures are above 15 °C (6). Active methods of control of nitrification in distribution systems include maintenance of a chloramine residual of 2.5 $mg{\cdot}L^{-1}$ with a chlorine-to-ammonia-nitrogen ratio of 4:1, which limits biofilm development in the distribution system by reduction of the organic carbon of the water, active cleaning of the distribution system by flushing or pigging, reducing the detention time of the water in the system, eliminating “dead ends” in the mains (36), and limiting reactive surfaces on pipes by coating with carbonate or polymer additives (37).

Control of Microbial Activity Involved in “Dirty Water” Formation

“Dirty water” caused by iron and manganese precipitates is a common problem for water supply utilities (38). Dirty water precipitates result from abiotic or biological oxidation of reduced iron or manganese, or organic materials with which they may be complexed. Iron and manganese deposits and sludges can limit the penetration of disinfectants into biofilms. Microbial manganese oxidation in a distribution system has been controlled by maintaining a chlorine residual of $>0.2\ mg{\cdot}L^{-1}$ and reducing manganese concentrations to below 0.02 $mg{\cdot}L^{-1}$ (2).

Removal of manganese and iron before distribution usually involves chemical oxidation to their insoluble forms Fe^{3+} and Mn^{4+} by aeration, chlorine, chlorine dioxide, potassium permanganate, or ozone, with subsequent removal by flocculation, sedimentation or filtration (38), or combined chemical and biological oxidation and removal by biofiltration (15,39). Biofilters are usually slow to develop and generally require the development of $MnO_2(s)$ precipitates to help catalyze the oxidation of soluble Mn^{2+} (15). Biofilters can be used for the simultaneous removal of AOC, BDOC, ammonium, iron, and manganese (40).

CONCLUSIONS

Microbial activity in water distribution systems affects the health compliance and aesthetics of the distributed water. Control of microbial activity in distribution systems is never complete, but it can be sufficiently achieved by limiting microbial energy and food sources (i.e., by removing as much AOC, BDOC, DOC, ammonia, reduced metals, and reduced sulfur compounds as is possible), in most cases by maintaining a reasonable residual disinfectant in the distributed water, and by careful selection and maintenance of distribution infrastructure, so that minimal microbiological growth and chemical reactions can occur on surfaces.

Water utilities have many treatment and disinfection choices for the control of microbial activity in distribution systems, and those choices will be guided by the quality of the waters they receive, their geography, the retention time of water in the system, and the expectations of the community. The most appropriate treatment options for control of microbial activity by any one utility must be guided by analysis of the performance of treatment methods in other treatment plants, and extensive testing and optimization of treatment methods alone, and in combination with the full treatment system.

BIBLIOGRAPHY

1. LeChevallier, M.W., Cawthon, C.D., and Lee, R.G. (1988). Factors promoting survival of bacteria in chlorinated water supplies. *App. Environ. Microbiol.* **54**: 649–654.
2. Sly, L.I., Hodgkinson, M.C., and Arunpairojana, V. (1990). Deposition of manganese in a drinking water distribution system. *Appl. Environ. Microbiol.* **56**: 628–639.
3. Franzmann, P.D., Heitz, A., Zappia, L.R., Wajon, J.E., and Xanthis, K. (2001). The formation of malodorous dimethyl oligosulphides in treated groundwater: the role of biofilms and potential precursors. *Water Res.* **35**: 1730–1738.
4. Wagner, D. and Chamberlain, A.H.L. (1997). Microbiologically influenced copper corrosion in potable water with emphasis on practical relevance. *Biodegradation* **8**: 177–187.
5. Lu, W., Kiéné, L., and Lévi, Y. (1999). Chlorine demand of biofilms in water distribution systems. *Water Res.* **33**: 827–835.
6. Wolf, R.L., Lieu, N.I., Izaguirre, G., and Means, E.G. (1990). Ammonia-oxidising bacteria in a chloraminated distribution system: seasonal occurrence, distribution, and disinfection resistance. *Appl. Environ. Microbiol.* **56**: 451–462.
7. Najm, I. et al. (2000). *Case Studies of the Impacts of Treatment Changes of Biostability in Full Scale Distribution Systems*. AWWA Research Foundation, Denver, CO.
8. van der Kooij, D. (1992). Assimilable organic carbon as an indicator of bacterial regrowth. *J. Am. Water Wastewater Assoc.* **84**: 57–65.
9. van der Kooij, D., van Lieverloo, J.H.M, Schellart, J., and Heimstra, P. (1999). Maintaining quality without a disinfectant residual. *J. Am. Water Wastewater Assoc.* **91**: 55–64.
10. Servais, P., Laurent, P., and Randon, G. (1993). Impact of biodegradable dissolved organic carbon (BDOC) on bacterial dynamics in distribution systems. *Proceedings AWWAWQTC Conference*, Miami, FL, 7–10 November.
11. Joret, J.C. and Lévi,Y. (1986). Méthode rapide d'évaluation du carbone éliminable des eaux par voie biologique. *Trib. Cebedau* **39**: 3–9.
12. Servais, P., Billen, G., and Hascoët, M.C. (1987). Determination of the biodegradable fraction of dissolved organic matter in waters. *Water Res.* **21**: 445–450.
13. Frías, J., Ribas, F., and Lucena, F. (1992) Comparison of methods for the measurement of biodegradable organic carbon and assimilable organic carbon in water. *Water Res.* **29**: 2785–2788.
14. Hem, L.J. and Efraimsen, H. (2001). Assimilable organic carbon in molecular weight fractions of natural organic matter. *Water Res.* **35**: 1106–1110.
15. Rittmann, B.E. and Snoeyink, V.L. (1984). Achieving biologically stable drinking water. *J. Am. Water Wastewater Assoc.* **76**: 106–114.
16. Volk, C.J. and LeChevallier, M.W. (2000). Assessing biodegradable organic matter. *J. Am. Water Wastewater Assoc.* **92**: 64–76.
17. LeChevallier, M.W., Schulz, W., and Lee, R.G. (1991). Bacterial nutrients in drinking water. *Appl. Environ. Microbiol.* **57**: 857–862.
18. Piriou, P., Dukan, S., and Kiene, L. (1998). Modelling bacteriological water quality in drinking water distribution systems. *Water Sci. Technol.* **38**: 299–307.
19. Escobar, I.C. and Randall, A.A. (2001). Assimilable organic carbon (AOC) and biodegradable dissolved organic carbon (BDOC): complementary measurements. *Water Res.* **18**: 4444–4454.
20. Volk, C.J. and LeChevallier, M.W. (2002). Effects of conventional treatment on AOC and BDOC levels. *J. Am. Water Wastewater Assoc.* **94**: 112–123.
21. Easton, J. (1993). *Measurement and Significance of Assimilable Organic Carbon*. Foundation of Water Research FR0373, Marlow, UK.
22. Slunjski, M. et al. (2002). MIEX—good research commercialised. *Water* **29**: 42–47.
23. Huck, P.M. (2000). Measurement of biodegradable organic matter and bacterial growth potential in drinking water. *J. Am. Water Wastewater Assoc.* **82**: 78–86.
24. Volk, C.J. et al. (2000). Impact of enhanced and optimized coagulation on removal of organic matter and its biodegradable fraction in drinking water. *Water Res.* **34**: 3247–3257.
25. Owen, D.M., Amy, G.L., and Chowdhury, Z.K. (1993) *Characterization of Natural Organic Matter and Its Relationship to Treatability*. AWWA Research Foundation, Denver, CO.
26. Vahala, R., Ala-Peijari, T., Rintala, J., and Laukkanen, R. (1998). Evaluating ozone dose for AOC removal in two-step GAC filters. *Water Sci. Technol.* **37**: 113–120.
27. Hacker, P.A., Paszkokolva, C., Stewart, M.H., Wolfe, R.L., and Means, E.G. (1994). Production and removal of assimilable organic-carbon under pilot-plant conditions through the use of ozone and peroxone. *Ozone-Sci. Eng.* **16**: 197–212.
28. Gagon, G.A., Ollos, P.J., and Huck, P.M. (1997). Modelling BOM utilisation and biofilm growth in distribution systems: review and identification of research needs. *J. Water Supply Res. Technol. Aqua* **46**: 165–180.
29. Suffet, I.H., Corado, A., Chou, D., Butterworth, S., and McGuire, M.J. (1994). AWWA Taste and Odour Survey. *Proceedings of the 1993 Water Quality Technology Conference*, Miami, FL November 7–11.
30. Richardson, S.D. (2003). Disinfection by-products and other emerging contaminants in drinking water. *Trends Anal. Chem.* **22**: 666–684.
31. Neden, D.G., Jones, R.J., Smith, J.R., Kirmeyer, G.J., and Frost, G.W. (1992). Comparing chlorination and chloramination for controlling bacterial regrowth. *J. Am. Water Wastewater Assoc.* **84**: 80–88.

32. Momba, M.N.B., Kfir, R., Venter, S.N., and Cloete, T.E. (2000). An overview of biofilm formation in distribution systems and its impact on the deterioration of water quality. *Water South Africa* **26**: 59–66.
33. Lund, V. and Ormerod, K. (1995). The influence of disinfection processes on biofilm formation in water distribution systems. *Water Res.* **29**: 1013–1021.
34. Armon, R., Arbel, T., and Green, M.. (1998). A quantitative and qualitative study of biofilm disinfection on glass, metal and PVC surfaces by chlorine, bromine, and bromochloro-5,5-dimethylhydantoin (BCDMH). *Water Sci. Technol.* **38**: 175–179.
35. Hallam, N.B., West, J.R., Forster, C.F., and Simms, J. (2001). The potential for biofilm growth in water distribution systems. *Water Res.* **35**: 4063–4071.
36. Kirmeyer, G.J., Odell, L.H., Jacangelo, J., Wilczak, A., and Wolfe, R. (1995). *Nitrification Occurrence and Control in Chloraminated Systems*. AWWA Research Foundation and American Water Works Association, Denver, CO.
37. Woolschlager, J.E., Rittmann, B.E., Piriou, P., and Schwartz, B. (2001). Developing a effective strategy to control nitrifier growth using the comprehensive disinfection and water quality model (CDWQ). In: *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges: "World Water Congress CD-ROM". Proceedings of the World Water and Environmental Resources Congress*, Orlando, FL, May 20–24.
38. Larson, A.L. (1993). Case pilot study: Organic fouling of a coated media pressure filter. *Proceedings AWWAWQTC Conference*, Miami, FL, 7–10 November.
39. Sly, L.I., Arunpairojana, V., and Dixon, D.R. (1993). Bioloigcal removal of manganese from water by immobilized manganese-oxidising bacteria. *Water* **20**: 38–40.
40. Korth, A., Bendinger, B., Czekalla, C., and Wichmann, K. (2001). Biodegradation of NOM in rapid sand filters for removing iron and manganese. *Acta Hydrochim. Hydrobiol.* **29**: 289–295.

MICROBIAL DYNAMICS OF BIOFILMS

Baikun Li
Pennsylvania State University
Harrisburg, Pennsylvania

BIOFILM FORMATION

The solid–liquid interface between a surface and an aqueous medium (e.g., water, blood) provides an ideal environment for the attachment and growth of microorganisms. Bacteria have a marked tendency to attach to surfaces exposed to an aqueous environment and initiate biofilm formation, characterized as a matrix of cells and cellular products attached to a solid substratum. The biofilm systems that result can be beneficial, as exemplified by fixed-film wastewater treatment processes (e.g., trickling filters and rotating biological contactors). However, this biological matrix can increase operational costs and/or decrease product quality in a variety of industries, including oil and paper production, semiconductor manufacture, and drinking water distribution. In water distribution systems and heat transfer equipment, for example, biofilms can cause substantial energy losses resulting from increased fluid frictional resistance and increased heat transfer resistance. Moreover, the pathogens trapped in microbial biofilm in a water supply pipeline have caused public health concerns.

Microorganisms grow on surfaces enclosed in biofilm and form microcolonies of bacterial cells, which are encased in adhesive polysaccharides excreted by the cells. Biofilm can provide advantages for microorganisms' survival in natural environment, especially under oligotrophic condition, because biofilms trap nutrients for growth of the microbial population and help prevent detachment of cells on surfaces present in flowing systems. In drinking water systems (oligotrophic environment), the percentage (70%) of positive (viable) cells was significantly higher at surface-adhered biofilm cells than that (40%) in the planktonic cells (1). It has been postulated that bacteria attached to surfaces are more active because of enrichment of nutrients at the surface. Therefore, the growth of cells at the surface is predominant over deposition and has advantage over planktonic (free swimming) in an oligotrophic environment.

Biofilms may form on a wide variety of surfaces, including living tissues, indwelling medical devices, industrial or potable water system piping, or natural aquatic systems. The variable nature of biofilms has been observed from an industrial water system and a medical device (Figs. 1 and 2). The water system biofilm is highly complex, containing corrosion products, clay material, freshwater diatoms, and filamentous bacteria. The biofilm on the medical device, on the other hand, seems to be composed of a single, coccoid organism and the associated extracellular polymeric substance (EPS) matrix.

It has been found that microbial motility may play a role in structure formation in biofilms (2). After the initial attachment to the substratum, *Pseudomonas aeruginosa* evidently moves on the substratum by means of twitching motility, which suggests that the initial microcolonies are formed by aggregation of bacteria. Moreover, microcolonies of bacterial cells in biofilms are the basic structural units in sessile communities, different from that of planktonic cells

Figure 1. Scanning electron micrograph of a native biofilm that developed on a mild steel surface in an 8-week period in an industrial water system. Rodney Donlan and Donald Gibbon, authors. Licensed for use, American Society for Microbiology Microbe Library. Available from: http://www.microbelibrary.org/.

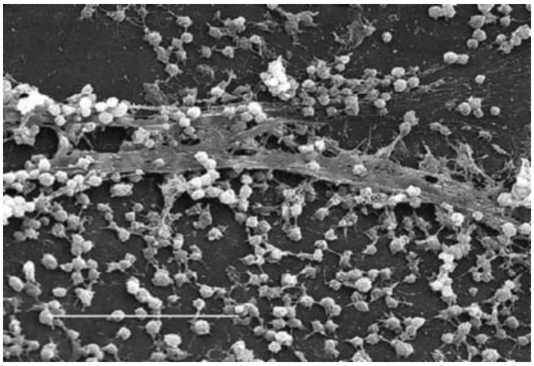

Figure 2. Scanning electron micrograph of a staphylococcal biofilm on the inner surface of an indwelling medical device. Bar, 20 μm. Used with permission of Lippincott Williams & Wilkins. Photograph by Janice Carr, CDC (http://www.cdc.gov/ncidod/eid/vol8no9/02-0063.htm).

and there is different gene expression in sessile bacteria. A large fraction of the biofilm bacteria (*P. aeruginosa, Discosoma* sp.) occasionally swim from one microcolony to another. In the initial phase of biofilm formation, bacteria are nonmotile and sessile inside the microcolonies. After several days of biofilm formation, biofilm bacteria start to swim rapidly in circles, the compact microcolonies are dissolved, and loose structures containing bacteria from different microcolonies are formed, along with flow channels. These results suggest that biofilms contain both sessile populations and planktonic populations.

BIOFILM STRUCTURE HETEROGENEITY

Biofilm Physical Structure

It had long been assumed that biofilm had a homogeneous structure, with microorganisms uniformly distributing throughout the entire biofilm at both structure and dynamics levels (3). Recently, biofilm was discovered heterogeneous in structure and function by microelectrode and confocal laser scanning microscope (CLSM). Numerous microbiology and microbial ecology of microbial ecosystem studies revealed that biofilms typically contain many layers with microbial ecology variation in each layer. Biofilms are not composed of bacterial cells randomly distributed in a homogeneous matrix, but they are composed of microcolonies of matrix-enclosed cells with intervening open-water channels. This complex and fascinating organization of biofilm communities was seen in flow-cell biofilms of several *Pseudomonas* species and in mixed-species natural biofilms formed on rock surfaces (4).

Microorganisms tend to accumulate in the bottom layer inside biofilm toward substratum, which leads to higher biomass density in the depth of biofilm. With the higher biomass density, the effective diffusivity distribution in heterogeneous biofilm decreased toward the bottom of a biofilm. It has been measured that the ratio of effective diffusivity in biofilm to the bulk solution diffusivity (Df/D) decreased from 0.68–0.81 in the top layer to 0.38–0.45 in the bottom layer (5). The highest biomass density, 120 g/L, was near the bottom of the biofilm, whereas the lowest density was measured near the top layers of biofilm, 20–30 g/L (6). It is anticipated that in the bottom layer of biofilm, metabolite and other by-products accumulate and have difficulties to diffuse out of biofilm, which lead to high mass.

Besides the mass density and diffusivity, other physiological features also vary along the depth of biofilm, which offer the heterogeneous structure in biofilm property (7). The solid densities (with units of mg-TS (total solid)/cm^3 total biofilm) in the bottom layers were 5–10 times higher than were those in the top layers; the ratio of living cells to total biomass decreased from 72–91% in the top layers to 31–39% in the bottom layers; the porosities of biofilms changed from 89–93% in the top layers to 58–67% in the bottom layers. In contrast, the mean pore radius of biofilms decreased from approximately 1.7–2.7 μm in the top layers to 0.3–0.4 μm in the bottom layers. Thus, bacterial cells living in a dense multilayer matrix face the problem of limited transport of substrate and nutrients into the film inner layers as well as export of waste products out of these deep regions.

Extracellular Polymeric Substances (EPSs)

A biofilm is an assemblage of microbial cells that is irreversibly associated with a surface and enclosed in a matrix of primarily polysaccharide material. The microcolonies in a biofilm are surrounded by large amounts of EPS. It has been found that EPSs play an important role in gene transformation and cell–cell contact in biofilm. Although it is not clear that EPSs have a role in shaping the spatial structure in biofilms, evidence has been provided that EPSs can facilitate storage of nutrients in biofilms for subsequent intake during periods of carbon limitation (8–10). The EPSs play several important roles for the biofilm, including (1) protection from naturally occurring or manmade biocides and antibiotics, (2) stabilization of bacterial adhesion to the substratum and to other cells in the biofilm, and (3) entrapment and concentration of soluble nutrients in proximity to biofilm members.

EPSs may account for 50–90% of the total organic carbon of biofilms (11) and can be considered the primary matrix material of the biofilm. EPSs synthesized by microbial cells may vary greatly in their composition, but the major component is polysaccharide. Some are neutral macromolecules, but most are polyanionic because of the presence of either uronic acids (D-glucuronic acid, D-galacturonic, D-mannuronic acids) or ketal-linked pyruvate. Inorganic residues, such as phosphate or rarely sulfate, may also confer polyanionic status (10). A few EPSs may even be polycationic, as exemplified by the adhesive polymer obtained from strains of *Staphylococcus epidermidis* associated with biofilms (12).

The amount of EPS synthesis within the biofilm is known to depend greatly on the availability of carbon substrates (both inside and outside the cell) and on the balance between carbon and other limiting nutrients (e.g., nitrogen, calcium) (10). Moreover, different organisms produce differing amounts of EPSs, and the amount of EPSs increases with age of the biofilm. Because EPSs are highly hydrated, they prevents desiccation in some natural biofilms. EPSs may also contribute to the antimicrobial resistance properties of biofilms by impeding the mass transport of antibiotics through the biofilm, probably by binding directly to these agents (13). This process can lead to the high antibiotics resistance of pathogens growing in biofilm.

Channel Structures in Biofilm

Biofilms are highly hydrated open structures containing a high fraction of EPSs and large void spaces between microcolonies. Most of the bacteria in the nutrient-sufficient parts of the biosphere grow in distinct microniches in highly structured multispecies communities. Bacterial growth zones form as discrete microcolony structures separated by channel boundaries, and these structures are maintained over time (Fig. 3). Also, open spaces extend from the top of the biofilms to the deep inner regions, which forms a channel-like network. Mushroom-shaped structures, separated by channels and voids, were observed in *P. fluorescens* biofilm (14).

It has been found that when biofilm grew from 20 to 300 μm in thickness, many vertical and horizontal channels permeated the film (15). These channels extended from the surfaces into at least one-half of the depth of the film. Moreover, on anaerobic fixed-bed reactor communities, integrative networks of channel-like structures were found throughout the microbial granules (15), which has been interpreted as playing a role in transport of cooperative substrates, nutrients, and gases (16). This porous architecture of biofilms may represent a more optimal arrangement for the influx of substrate and nutrients and transfer of wastes.

The biofilm heterogeneous structure of microbial clusters of cells and EPSs was also verified by microelectrode measurement (17). The microbial clusters of cells and EPSs were separated by interstitial voids. The cell clusters were 300 μm wide, and the voids were 100 μm wide. Biofilms are not flat homogeneous layers of cells, in which transport is exclusively diffusional. The transport of metabolites involved both convection and diffusion, which offers a relative importance depending on the hydraulic regime. The velocity of the bulk liquid had a strong influence on the oxygen concentration and mass distribution in the biofilm, with high flow velocity increasing oxygen transferring inside biofilm and concentration differences and thus increasing convectional transport.

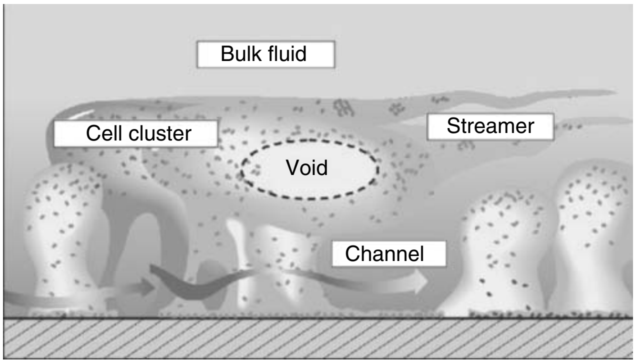

© 1996 Center for biofilm engineering montana state university-bozeman

Figure 3. The diagram of cluster-channel structure in biofilm (Image from http://www.erc.montana.edu/Res-Lib99-SW/Image_Library/Structure-Function/default.htm.)

BIOFILM METABOLIC HETEROGENEITY

Metabolic heterogeneity may account for the phenomenal success of microbial biofilms in nature and in device-related and other chronic bacterial infections, and it may be a major factor in the inherent resistance of medical biofilms to antibiotics.

Heterogeneity within a population (differing activity of single cells associated with increasing depth in a biofilm) might provide useful insight into aspects of substrate and mass transport limitation. The established biofilm population was composed of cells having a variety of growth rates. Such differences might reflect different positions of cells within the biofilm matrix; for example, cells located near the surface of the biofilm may not be as limited by mass transport of, or competition by surrounding cells for, nutrients (18).

Cell ribosome content changes at different growth phases, with higher ribosome content in new cells and lower ribosome content in old cells. The variation in ribosome content can provide a useful tool for evaluating the status of a great variety of microorganism growing at different generation times inside biofilm (19). Quantifying the fluorescent light emitted from *in situ* rRNA hybridized cells indicated that sulfate-reducing bacteria (SRB) in a young multispecies biofilm were more active than were the sulfate-reducing bacteria in an established biofilm (20). When microcolonies of a *P. putida* strain growing in a biofilm reaches a critical size, the light emitted by the cells decreased in the center of the microcolonies and eventually throughout the microcolonies, which indicates that the cells displayed different levels of growth activity correlating with their location in the biofilm and with the biomass of biofilm (18).

MICROBIAL STRUCTURE AND DYNAMIC HETEROGENEITY IN BIOFILM IN WATER/WASTEWATER TREATMENT

The heterogeneity of microbial structure and dynamics is very important in biological wastewater treatment. Trickling filter, which depends on a variety of microorganisms growing in different layers in biofilm, is a major process in nutrient removal. Wastewater biofilms are complex multispecies biofilms, displaying considerable heterogeneity with respect to both the microorganism present and their physicochemical microenvironment. Successive vertical zonations of predominant respiratory processes occurring simultaneously in close proximity have been found in aerobic wastewater biofilms with a typical thickness of only a few millimeters (21). Numerous environmental microbiology and molecular biology studies have revealed the correlation between the microbial structure and dynamic heterogeneity and the availability of oxygen and substrate in two biofilm systems: sulfur-reducing biofilm and nitrification/denitrification biofilm.

Sulfate-Reducing Biofilm

The reduction of sulfate (SO_4^{2-}) or sulfite (SO_3^{2-}) to sulfide (S^{2-}) occurs in an anaerobic zone deeper inside biofilm with organic carbon as the electron donor. It is carried out by anaerobic SRB. The presence of oxygen and nitrate in the deep layer of biofilm has been found to inhibit the anaerobic SRB growth, because oxygen and nitrate compete with sulfate or sulfite to be the electron acceptor in anaerobic biofilm. The correlation of the vertical distribution of SRB with the gradients of oxygen and hydrogen sulfide (the product of the reducing sulfate by SRB) in a trickling filter revealed that hydrogen sulfide was predominant in the anoxic zone of the biofilm, and most of the SRB were found in the anoxic layer as dense clusters of assemblies (22). Oxygen respiration occurred in the upper 0.2–0.4 mm of the biofilm, whereas sulfate reduction occurred in deeper, anoxic parts of the biofilm (below 1.5 mm inside biofilm) (23). Moreover, the presence of sulfate in biofilm significantly affects the stratification of microbial community inside biofilm, because sulfate is the main electron acceptor for SRB in anoxic reaction. Total sulfate reduction rates (μmol S/cm^3/h) increased when the availability of SO_4^{2-} or organic substrate increased as a result of deepening the sulfate reduction zone or an increase in the sulfate reduction intensity.

SRB can even survive in the aerobic zone in biofilm and conduct good sulfur-reducing ability. In one study, SRB (approximately 10^9–10^{10} cell per cm^3 of biofilm) was evenly distributed throughout the biofilm, even in the oxic surface (21). Oxygen penetrated only about 100 μm from the surface in a biofilm approximately 1000 μm thick. Sulfide was produced from SRB in a narrow zone of 150–300 μm below the surface at a maximum specific rate of 21 μmol H_2S /cm^3/h.

Nitrification/Denitrification Biofilm

Nitrogen is removed in biological wastewater treatment by two types of bacteria: aerobic autotrophic nitrifier (oxidize NH_4^+ to NO_3^-) and anoxic heterotrophic denitrifer (reduce NO_3^- to N_2). Two major groups of nitrifiers exist: ammonia oxidizing bacteria (AOB, oxidize NH_4^+ to NO_2^-) and nitrite oxidizing bacteria (NOB, oxidize NO_2^- to NO_3^-). In biofilm, aerobic nitrifier and anoxic denitrifer dominate in the different zones because of the transfusion of oxygen and organic and inorganic substrates. The nitrifying and denitrifying bacterial populations and their activities in biofilms have been investigated by the application of fluorescent *in situ* hybridization (FISH) and microsensor technology. A dense layer of ball-shaped clusters of *Nitrosomonas* cells (ammonium oxidizers) in close physical association with surrounding smaller clusters of *Nitrobacter* cells (nitrite oxidizers) was found in the aerobic zone of a nitrifying biofilm from the trickling filter of an aquaculture (17). The nitrifying zone (measured with microelectrode) was restricted to the outer 100–150 μm of the biofilm, and the nitrifiers were found in this zone as a dense layer of clusters. The central part of the biofilm was found anoxic and inactive for nitrification, and significantly fewer nitrifiers were found there.

The dynamic studies of nitrifers have revealed their stratification and activity inside biofilm. The active ammonia-oxidizing zone was located in the outer part of a biofilm, whereas the active nitrite-oxidizing zone was located just below the ammonia-oxidizing zone (24). The number of nitrite-oxidizing bacteria (NOB) is low compared with the number of ammonia-oxidizing bacteria (AOB). Oxygen penetration depth was about 1200 μm in the wastewater biofilm because of the loose structure and low microbial activity. Nitrification occurred in the top 200 μm, whereas denitrification occurred in the deeper anoxic layers (below a depth of about 1200 μm). The nitrification layer is 150–175 μm, AOB activity is 0.13–0.08 μmol NH_4^+/cm^2/h, and average specific nitrifying rate is 8.7–4.6 μmol NH_4^+/cm^3/h. NOB activity is 0.22–0.3 μmol NO_2^-/cm^2/h, and the specific rate is 14.7–17.1 μmol NO_2^-/cm^3/h. AOB (*Nitrosomonas*) were detected throughout the biofilm, whereas NOB (*Nitrospira*-like bacteria) were found mainly in the inner parts of biofilm and always found in vicinity of AOB clusters, which may have reflected the synthophic association between AOB and NOB (25).

Besides the microbial stratification and activity in biofilm, some parameters were examined to indicate the different microbial processes inside biofilm. The stratification of microbial processes and the associated redox potential (ORP) changes in biofilms were studied by microelectrodes (26). In aerobic/sulfate-reducing biofilm, a clearly stratified structure with depth was observed: aerobic oxidization took place only in a shallow layer near the surface, and sulfate reduction occurred in the deeper anoxic zone. The boundary between these two processes was well defined by redox potential (ORP increased from −100 mV to +100 mV). In the aerobic/nitrifying biofilm, a stratified structure with depth was also observed, but it was less well defined (ORP increased from +300 mV to +330 mV). In this biofilm, aerobic oxidation took place throughout the biofilm depth, and more nitrification occurred in the deeper section of the biofilm.

The competition occurs between anoxic denitrifer and SRB in biofilm (27). In SRB biofilm, oxygen was depleted in the first 0.2 mm of the biofilm (1.5 mm thickness). SRB occurred in the anoxic zones inside biofilm. However, when nitrate is added in the bulking solution, the SRB zone withdrew deeper into the biofilm. Nitrate penetrated into 0.4 mm depth inside biofilm, which indicates that denitrification occurred. Nitrate decreased the sulfide content either by oxidation of produced sulfide or by inhibition of sulfate reduction.

Oxygen Limitation in Biofilm

Growth rate limitation occurs in biofilms presumably as a result of restriction for a particular nutrient (such as oxygen and substrate) that fails to fully penetrate the biofilm. The expression of alkaline phosphatase (APase) could be a physiological indicator. It was found that the band of APase (to indicate enzyme activities) on the *P. aeruginosa* biofilm surface decreased from 46 μm thick with biofilms grown in pure oxygen to 2 μm thick with biofilms grown in pure nitrogen (28). The expression of APase on the biofilm surface revealed marked spatial physiological heterogeneity within biofilms in which active protein synthesis was restricted by oxygen availability inside the biofilm. This heterogeneity has implications for microbial ecology and for understanding the reduced susceptibilities of biofilms to antibiotics agents under limited nutrients in the establishment of physiological gradients.

The direct observation of living biofilms has found many different microniches occupied by individual cells within these complex communities. Cells near the surface of the microcolony would obtain high concentrations of the oxygen delivered by the water in the channels while oxygen is rapidly depleted below the microcolony surface, and the centers of the microcolonies that comprise biofilms growing in air are virtually anaerobic (29). The steep oxygen gradient is paralleled by gradients of differing steepness for every ion and molecule needed or produced by cells in biofilms, and the interaction of these gradients produces a different microniche for every individual cell in the complex heterogeneous community.

BIOFILM DETACHMENT

Biofilm development is a balance between new biofilm growth and old biofilm detachment. Biofilm detachment is the removal of cells from the existing biofilm to the bulk fluid. It is a significant factor that balances cell growth in the biofilm and brings the system to steady state. Biofilm cells may be dispersed either by shedding of daughter cells from actively growing cells, detachment as a result of nutrient levels, or shearing of biofilm aggregates (continuous removal of small portions of the biofilm) because of flow effects.

The mechanisms underlying the process of shedding by actively growing cells in a biofilm are not well understood. It has been recognized that several physical forces are involved in biofilm detachment (30): erosion (continuous removal of small particles of biofilm as a result of shear forces), sloughing (random removal of large portions of biofilm, often occurring in old and thick biofilms or when exposed to dramatic environmental changes), abrasion (removal caused by collisions of solid particles with biofilms particularly in fluidized bed reactors), and predator grazing (consumption of biofilm by larger organisms such as protozoa).

Besides these physical forces, the microbial metabolic dynamic inside biofilm also contributes to biofilm detachment. As biofilm matures, oxygen will be depleted in the upper layer of biofilm and anoxic/anaerobic zones will dominate in the deeper layers of biofilm. The anoxic/anaerobic reactions produce a variety of biogas (N_2, H_2S, CO_2, CH_4, H_2, etc.) and might increase up to 20% in the established biofilm (31). The biogas contributes to the biofilm detachment. Moreover, the amount of EPSs in biofilm also changes during biofilm formation and maturation. The content of EPSs decreased within the biofilm in the course of biofilm maturation. As a consequence, the adhesion strength of biofilm adjacent to substratum becomes more weakened, which leads to biofilm detachment (32).

BIBLIOGRAPHY

1. Manz, W. et al. (1993). *In situ* identification of bacteria in drinking water and adjoining biofilming by hybridization with 16S and 23S rRNA-directed fluorescent oligonucleotide probes. *Appl. Environ. Microbiol.* **59**(7): 2293–2298.
2. Nielson, T.T. et al. (2000). Development and dynamics of *Pseudomonas* sp. biofilm. *J. Bacteriol. Microbiol.* **192**(22): 6482–6489.
3. Dalsgaard, T. and Revsbech, N.P. (1992). Regulation factors of denitrification in trickling filter biofilms as measured with the oxygen/nitrous oxide microsensor. *FEMS Microbiol. Ecol.* **101**: 151–164.
4. Costerton, J.W. et al. (1995). *Microbial Biofilms Annu. Rev. Microbiol.* **49**: 711–745.
5. Beyenal, H., Tanyolaç, A., and Lewandowski, Z. (1998). Measurement of local effective diffusivity in heterogeneous biofilms. *Water Sci. Technol.* **38**(8–9): 171–178.
6. Zhang, T.C. and Bishop, P.L. (1994). Evaluation of tortuosity factors and effective diffusivities in biofilms. *Water Res.* **28**(11): 2279–2287.
7. Zhang, T.C. and Bishop, P.L. (1994). Density, porosity and pore structure of biofilms. *Water Res.* **28**(11): 2267–2277.
8. Hussain, M., Wilcox, M.H., and White, P.J. (1993). The slime of coagulase-negative staphylococci: biochemistry and relation to adherence. *FEMS Microbiol. Rev.* **104**: 191–208.
9. Leriche, V., Sibille, P., and Carpentier, B. (2000). Use of an enzyme-linked lectinsorbent assay to monitor the shift in polysaccharide composition in bacterial biofilms. *Appl. Environ. Microbiol.* **66**: 1851–1856.
10. Sutherland, I.W. (2001). Biofilm exopolysaccharides: a strong and sticky framework. *Microbiology* **147**: 3–9.
11. Flemming, H-C., Wingender, J., Griegbe, and Mayer, C. (2000). Physico-chemical properties of biofilms. In: *Biofilms: Recent Advances in Their Study and Control*. L.V. Evans (Ed.). Harwood Academic Publishers, Amsterdam, pp. 19–34.
12. Mack, D. et al. (1996). The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear ß-1,6-linked glucosaminoglycan: purification and structural analysis. *J. Bacteriol.* **178**: 175–183.
13. Donlan R.M. (2000). Role of biofilms in antimicrobial resistance. *ASAIO J.* **46**: 47–52.
14. Korber, D.R., Lawrence, J.R., and Caldwell, D.E. (1994). Effect of motility on surface colonization and reproductive success of *Pseudomonas fluorescens* in dual-dilution continuous culture and batch culture systems. *Appl. Environ. Microbiol.* **60**: 1421–1429.
15. Massol-deya, A.A., Whallon, J., Hickey, R., and Tiedjie, J.M. (1995). Channel structure in aerobic biofilms of fixed-film reactors treating contaminated groundwater. *Appl. Environ. Microbiol.* **61**(2): 769–777.
16. MacLeod, F.A., Giot, S.R., and Costerton, J.W. (1990). Layered structure of bacterial aggregates produced in an upflow anaerobic sludge bed and filter reactor. *Appl. Environ. Microbiol.* **56**: 1598–1607.
17. De Beer, D. and Stoodley, P. (1995). Relation between the structure of anaerobic biofilm and transport phenomena. *Water Sci. Technol.* **32**(8): 11–18.
18. Tolker-Nielsen, T. et al. (2000). Development and dynamics of *Pseudomonas* sp. biofilms. *J. Bacteriol.* **182**: 6482–6489.
19. Poulsen, L.K., Ballard, G., and Stahl, D.A. (1993). Use of rRNA fluorescence *in situ* hybridization for measuring the activity of single cells in young and established biofilms. *Appl. Environ. Microbiol.* **59**(5): 1354–1360.
20. Ito, T., Okabe, S., Satoh, H., and Watanabe, Y. (2002). Successional development of sulfate-reducing bacterial populations and their activities in a wastewater biofilm growing under

microaerophilic conditions. *Appl. Environ. Microbiol.* **68**(3): 1392–1402.

21. Okabe, S., Itoh, T., Satoh, H., and Watanabe, Y. (1995). Analyses of spatial distributions of sulfate-reducing bacteria and their activity in aerobic wastewater biofilms. *Appl. Environ. Microbiol.* **65**(11): 5107–5116.
22. Ramsing, N.B., Juhl, M., and Jorgensen, B.B. (1993). Distribution of sulfate-reducing bacteria, O2, and H2S in photosynthetic biofilms determined by oligonucleotide probes and microelectrodes. *Appl. Environ. Microbiol.* **59**(11): 3840–3849.
23. Kuhl, M. and Jorgensen, B.B. (1992). Microsensor measurements of sulfate reduction and sulfide oxidation in compact microbial communities of aerobic biofilms. *Appl. Environ. Microbiol.* **58**(4): 1164–1174.
24. Okabe, S., Satoh, H., and Watanabe, Y. (1999). *In situ* analysis of nitrifying biofilms as determined by *in situ* hybridization and the use of microelectrodes. *Appl. Environ. Microbiol.* **65**(7): 3182–3191.
25. Gieseke, A. et al. (2001). Community structure and activity dynamics of nitrifying bacteria in a phosphate-removing biofilm. *Appl. Environ. Microbiol.* **67**(3): 1351–1362.
26. Bishop, P.L. and Yu, T. (1999). A microelectrode study of redox potential change in biofilms. *Water Sci. Technol.* **39**(7): 179–185.
27. Santegoeds, C.M., Muyzer, G., and de Beer, D. (1998). Biofilm dynamics studied with microsensors and molecular techniques. *Water Sci. Technol.* **37**(4–5): 125–129.
28. Xu, K.D. et al. (1998). Spatial physiological heterogeneity in *Pseudomonas aeruginosa* biofilm is determined by oxygen availability. *Appl. Environ. Microbiol.* **64**(10): 4035–4039.
29. Lewandowski, Z. (2000). Structure and function of biofilms. In: *Biofilms: Recent Advances in Their Study and Control*. L.V. Evans (Ed.). Harwood Academic Publishers, Amsterdam, pp. 1–17.
30. Peyton, B.M. and Characklis, W.G. (1993). A statistical analysis of the effect of substrate utilization and shear stress on the kinetics of biofilm detachment. *Biotechnol. Bioeng.* **41**(7): 728–735.
31. Ohashi, A. and Harada, H. (1994). Characterization of detachment mode of biofilm developed in an attached-growth reactor. *Water Sci. Technol.* **30**(11): 35–45.
32. Horan, N.J. and Eccels, C.R. (1986). Purification and characterization of extracellular polysaccharide from activated sludgs. *Water Res.* **20**: 1427–1432.

MICROBIAL ENZYME ASSAYS FOR DETECTING HEAVY METAL TOXICITY

Oladele A. Ogunseitan
University of California
Irvine, California

THE PERSISTENT PROBLEM OF TOXIC METALS IN THE WATER ENVIRONMENT

Throughout the world, industrial development has been associated with increased mining, refining, and use of potentially toxic metals in large-scale manufacturing processes. Industrialization is also tightly linked to increases in the combustion of fossil fuels such as coal and petroleum, which produce considerable amounts of toxic metals such as mercury and lead, respectively (1). Inevitably, many industrial activities generate environmental pollution that changes the natural biogeochemical cycling of elements by modifying the size and flux of metal reservoirs. The primary consequence of metal pollution in natural water systems is toxic effects on aquatic organisms, but ultimately human exposure occurs from consumption of contaminated water, seafood, and plant crops that are cultivated in land irrigated with polluted water.

The United States Environmental Protection Agency (EPA) has established maximum contaminant levels (MCL) for eleven metals that dominate the inorganic contaminants most commonly encountered in drinking water and groundwater sources (Table 1). The symptoms associated with these toxic metals range from skin disease from chromium (MCL = 0.1 mg/L) to more serious conditions such as kidney disease from cadmium (MCL = 0.005 mg/L) and mercury (MCL = 0.002 mg/L). In addition, nonenforceable standards have been established for secondary pollutants such as Al, Cu, Fe, Mn, Ag, and Zn because they are regarded as nuisance contaminants with the assumption of negligible health consequences at the established concentrations permitted in water (Table 2).

Metals are also ranked third among the top 100 contaminant categories most commonly found in violation of established water quality standards such as the total maximum daily load (TMDL; see Table 3). Metals represent the largest proportion (1977 or 20.6%) of the total number of TMDLs approved for contaminant categories in water systems (Table 4). Mercury, copper, and lead are the three metals found most frequently in impaired water systems. These issues make metal pollution one of the most pressing water quality problems. In addition, there is a lot of uncertainty about metal doses that can produce disease because of the variability of individual genetic and behavioral susceptibility factors. Therefore, new techniques are continuously sought to produce consistent and reliable measures of ecological and human health risks from metal exposure. Microorganisms are involved in three of the multifaceted strategies designed to deal with metal contamination of water resources: first, in the control of acid mine drainage; second, in the bioremediation of contaminated water systems; and third, as indicators of metal bioavailability in toxicity testing. This chapter focuses on the third aspect of microbial interactions with metals in water, the use of microbial enzymes to indicate bioavailability and toxicity. A case study of microbial enzyme response to Pb is presented as an example.

MICROBIAL METALLOENZYMES

Several enzymes used by microorganisms to catalyze key reactions of the major biogeochemical cycles are metalloenzymes; to function properly, they require metal ions positioned at or near the active site. At least a third of all proteins in the biosphere are metalloproteins, and several metals are commonly found in physiological systems in small concentrations. For example Ca, K, Na, Fe, Mo, Mn, Mg, Cu, Zn, Co, Ni, W, and Al are commonly

Table 1. Inorganic Chemicals Category of Water Contaminants with Established Maximum Contaminant Levels (MCLs)

Contaminant	MCLG[b], mg/L[c]	MCL or TT[b], mg/L[c]	Potential Health Effects from Ingestion of Water	Sources of Contaminant in Drinking Water
Antimony	0.006	0.006	Increase in blood cholesterol; decrease in blood sugar	Discharge from petroleum refineries; fire retardants; ceramics; electronics; solder
Arsenic	0[d]	0.010 as of 01/23/06	Skin damage or problems with circulatory systems and may have increased risk of cancer	Erosion of natural deposits; runoff from orchards; runoff from glass and electronics production wastes
Asbestos (fiber >10 μm)	7 million fibers per liter	7 million fibers per liter	Increased risk of developing benign intestinal polyps	Decay of asbestos cement in water mains; erosion of natural deposits
Barium	2	2	Increase in blood pressure	Discharge of drilling wastes; discharge from metal refineries; erosion of natural deposits
Beryllium	0.004	0.004	Intestinal lesions	Discharge from metal refineries and coal-burning factories; discharge from electrical, aerospace, and defense industries
Cadmium	0.005	0.005	Kidney damage	Corrosion of galvanized pipes; erosion of natural deposits; discharge from metal refineries; runoff from waste batteries and paints
Chromium (total)	0.1	0.1	Allergic dermatitis	Discharge from steel and pulp mills; erosion of natural deposits
Copper	1.3	TT[e]; action level = 1.3	Short term exposure: gastrointestinal distress Long term exposure: liver or kidney damage People with Wilson's disease should consult their personal doctors if the amount of copper in their water exceeds the action level	Corrosion of household plumbing systems; erosion of natural deposits
Cyanide (as free cyanide)	0.2	0.2	Nerve damage or thyroid problems	Discharge from steel/metal factories; discharge from plastic and fertilizer factories
Fluoride	4.0	4.0	Bone disease (pain and tenderness of the bones); children may get mottled teeth	Water additive that promotes strong teeth; erosion of natural deposits; discharge from fertilizer and aluminum factories
Lead	zero	TT[e]; action level = 0.015	Infants and children: delays in physical or mental development; children could show slight deficits in attention span and learning abilities Adults: kidney problems; high blood pressure	Corrosion of household plumbing systems; erosion of natural deposits
Mercury (inorganic)	0.002	0.002	Kidney damage	Erosion of natural deposits; discharge from refineries and factories; runoff from landfills and croplands
Nitrate (measured as nitrogen)	10	10	Infants below the age of six months who drink water containing nitrate in excess of the MCL could become seriously ill and, if untreated, may die. Symptoms include shortness of breath and blue-baby syndrome.	Runoff from fertilizer use; leaching from septic tanks, sewage; erosion of natural deposits

Table 1. (*Continued*)

Contaminant	MCLG[a], mg/L[b]	MCL or TT[a], mg/L[b]	Potential Health Effects from Ingestion of Water	Sources of Contaminant in Drinking Water
Nitrite (measured as nitrogen)	1	1	Infants below the age of six months who drink water containing nitrite in excess of the MCL could become seriously ill and, if untreated, may die. Symptoms include shortness of breath and blue-baby syndrome.	Runoff from fertilizer use; leaching from septic tanks, sewage; erosion of natural deposits
Selenium	0.05	0.05	Hair or fingernail loss; numbness in fingers or toes; circulatory problems	Discharge from petroleum refineries; erosion of natural deposits; discharge from mines
Thallium	0.0005	0.002	Hair loss; changes in blood; kidney, intestine, or liver problems	Leaching from ore-processing sites; discharge from electronics, glass, and drug factories

[a] Asbestos, cyanide, nitrate, and nitrite are the few nonmetallic chemicals in this category. They are included here for comparison of standards and also because their presence in water may affect the biological availability of metals and the nature of microbial interaction with available metal ions.

[b] Definitions: **Maximum contaminant level (MCL)**: The highest level of a contaminant that is allowed in drinking water. MCLs are set as close to MCLGs as feasible using the best available treatment technology and taking cost into consideration. MCLs are enforceable standards.

Maximum contaminant level goal (MCLG): The level of a contaminant in drinking water below which there is no known or expected risk to health. MCLGs allow for a margin of safety and are nonenforceable public health goals.

Treatment Technique (TT): A required process intended to reduce the level of a contaminant in drinking water.

[c] Units are in milligrams per liter (mg/L) unless otherwise noted. Milligrams per liter are equivalent to parts per million.

[d] MCLGs were not established before the 1986 Amendments to the Safe Drinking Water Act. Therefore, there is no MCLG for this contaminant.

[e] Lead and copper are regulated by a treatment technique that requires systems to control the corrosiveness of their water. If more than 10% of tap water samples exceed the action level, water systems must take additional steps. For copper, the action level is 1.3 mg/L, and for lead it is 0.015 mg/L.

Source: U.S. EPA (http://www.epa.gov/safewater/mcl.html#mcls).

associated with diverse proteins in microorganisms (see Table 5). For each of these “essential” metals, there are at least two other metals that can compete for their interaction with proteins, but these competing “nonessential” metals are likely to be very toxic because the proteins to which they bind do not function properly (2). To survive in environments with fluctuating metal concentrations and parameters of metal bioavailability, many microorganisms have evolved specific uptake systems for essential metals, whereas specific detoxification systems are evolved to resist the detrimental effects of nonessential metals. For example, the uptake of iron is facilitated by the production of pyoverdins in fluorescent *Pseudomonas*, whereas many members of this genus also produce enzymes such as mercuric reductase which has the only known function of mercury detoxification (3). Therefore, measurement of metalloenzyme kinetics in the presence and absence of toxic metal ions can provide a direct assay for the biological availability of metallic contaminants in water systems.

Metal bioavailability is influenced by many factors, including pH, particulate matter that can adsorb metals, and inorganic elements and ions such as sulfur, chloride, and carbonate that can affect metal speciation (4,5). Therefore, monitoring metal concentrations in water based solely on chemical extraction and physical analysis typically generates “total concentration” data, but the biologically available concentration remains unknown. This is the main reason for investigating enzyme bioassays to estimate metal toxicity. When cell-based assays are used, enzyme assays reflect the quantity of metals which must have entered the cell to provoke physiological responses. Ogunseitan (3) and Ogunseitan and colleagues (6,7) have used such responses to monitor the biological availability of Hg and Pb, respectively. The following section describes the use of a microbial enzyme to assess the toxicity of biologically available Pb in water systems.

Table 2. Permissible Concentrations of Six Metals and Some Nonmetal Components[a]

Contaminant	Secondary Standard
Aluminum	0.05 to 0.2 mg/L
Chloride	250 mg/L
Color	15 (color units)
Copper	1.0 mg/L
Corrosivity	Noncorrosive
Fluoride	2.0 mg/L
Foaming agents	0.5 mg/L
Iron	0.3 mg/L
Manganese	0.05 mg/L
Odor	3 threshold odor number
pH	6.5–8.5
Silver	0.10 mg/L
Sulfate	250 mg/L
Total dissolved solids	500 mg/L
Zinc	5 mg/L

[a] The concentrations of the metals are monitored according to National Secondary Drinking Water Regulations (NSDWRs or Secondary Standards), which are nonenforceable guidelines regulating contaminants that may cause cosmetic effects (such as skin or tooth discoloration) or aesthetic effects (such as taste, odor, or color) in drinking water. EPA recommends secondary standards to water systems but does not require systems to comply. However, states may choose to adopt them as enforceable standards. The co-occurrence in water of nonmetallic components of this list may affect the biological availability of metal ions, which is a major factor in toxicity assessments based on microbial enzymes.

Table 3. Metals Are Ranked Third Among the Top 100 Impairments from Water Pollution; Hg, Cu, and Pb are the Most Frequently Cited Pollutants

Impairment	Number of Impairments
Chromium VI	1
Contaminated sediments (zinc)	1
Zinc	351
Contaminated sediments (alum)	1
Contaminated sediments (lead)	2
Molybdenum	1
Thallium	4
Arsenic	177
Contaminated sediments (cadmium)	1
Contaminated sediments (copper)	1
Contaminated sediments (silver)	1
Heavy metals	3
Iron	154
Lead	442
Manganese	52
Aluminum	48
Cadmium	227
Contaminated sediments (gold)	1
Silver	48
Trace elements	2
Boron	23
Chromium	75
Contaminated sediments (chromium)	1
Copper	528
Metals	2597
Metals—mercury	3
Contaminated sediments (mercury)	2
Mercury	697
Nickel	44
Selenium	173

Source: U.S. EPA (http://oaspub.epa.gov/pls/tmdl/waters_list.impairments?p_impid=13).

Table 4. Approved TMDLs by Metallic Pollutants Since January 1, 1996

Pollutant	Number of TMDLs
Aluminum	415
Arsenic	13
Boron	9
Cadmium	69
Chromium	12
Chromium VI	3
Contaminated sediments (copper)	1
Contaminated sediments (mercury)	2
Contaminated sediments (nickel)	1
Copper	79
Iron	439
Lead	95
Manganese	409
Mercury	146
Metals	49
Nickel	29
Selenium	65
Silver	29
Thallium	1
Zinc	104

Source: U.S. EPA (http://oaspub.epa.gov/pls/tmdl/waters_list.tmdl_pollutants?p_polid=13).

MICROBIAL AMINOLEVULINATE DEHYDRATASE AS A BIOSENSOR FOR Pb BIOAVAILABILITY

delta-Aminolevulinate dehydratase (ALAD; porphobilinogen synthase; EC: 4.2.1.24) is a 128-kDa metalloenzyme that catalyzes the first common step in the biosynthetic pathway for all tetrapyrroles, including oxygen carriers (e.g., heme), light gathering arrays (e.g., chlorophyll), vitamin B_{12}, and cofactor F430. The enzyme catalyzes the reaction between two molecules of aminolevulinic acid (ALA), which engage in a Knorr-type condensation reaction to produce porphobilinogen, the universal precursor of all pyrroles (Fig. 1) (7). ALAD activity has been demonstrated in all microorganisms that have been investigated across several phylogenetic groups. Microbial ALAD requires either zinc, magnesium, or both for optimum functioning, although these metals can be substituted by heavy metals such as Pb and Hg (Fig. 2).

In the proximity of the enzyme active site is the metal-binding domain, which for zinc consists of three cysteines at positions 133, 135, and 143, and a solvent molecule, presumably a hydroxide ion. Asparagine at positions 135 and 143, it has been postulated, is essential for Mg-dependent ALAD (8). Few microbial ALAD systems have been described, and the metallic component of ALAD differs among species. For example, Zn is required for ALAD activity in yeasts, whereas Mg is required in *Bradyrhizobium japonicum* and *Pseudomonas aeruginosa*, and both Zn and Mg are required by the *Escherichia coli* ALAD, although the requirement for Mg in *E. coli* is not stringent (8). Moreover, the sensitivity of ALAD to toxic metals depends on the identity of the metallic cofactor (8). ALAD activity in aquatic organisms has long been used as a biomarker of lead exposure (9,10), and a polymorphism in the ALAD locus has been described in humans that has implications for susceptibility to lead poisoning (11).

ALAD activity in an environmental strain of *Pseudomonas putida* ATCC 700097 originally isolated from an urban wastewater stream was sensitive to Pb and other toxic metals (12). There is a statistically significant dose–response relationship between ALAD activity in cells of this organism and Pb (Pearson correlation coefficient $= -0.985$; $r^2 = 0.97$; and $p < .001$). The highest level of inhibition of ALAD activity was approximately 74% of the normal level when cells were incubated with $Pb^{2+} > 500\ \mu M$. The relationship between Pb and ALAD activity was statistically described by $\log_{10}$ [Pb] $= 3 : 68 - 1.41$ [ALAD activity]. ALAD activity is measured in μmol of porphobilinogen synthesized per mg protein per minute. The direct exposure of protein extracted from *P. putida* showed a stronger inhibitory effect on ALAD activity, represented by a reduction of up to 85% in response to 500 μM of Pb. A higher concentration of Pb is needed to produce a comparable level of ALAD inhibition in *P. putida* cells seeded into natural freshwater; this suggests that Pb is not completely biologically available in freshwater. Therefore, an immobilized enzyme biosensor can be expected to be an even more sensitive tool for assessing biologically reactive forms of Pb in water, although this measurement precludes the barrier represented by cellular integrity (7).

The reduced bioavailability of Pb in water systems can be attributed to the presence of ligand chemicals, which

Table 5. Examples of Metals Required by Key Enzyme Functions[a]

Metal	Enzyme Examples	Relevant Organisms and/or Ecological Function
Calcium	Collagenase Calpain protease	*Clostridium histolyticum* Pathogenesis
Cobalt	Halomethane methyltransferase	Facultative methylotrophs Chloromethane degradation
Copper	Copper amine oxidases	*Arthrobacter globoformis* Deamination of primary amines to corresponding aldehydes
Iron	Cytochrome P450 Soluble methane Monooxygenase	*Pseudomonas putida*; P450 cam; polyaromatic hydrocarbon biodegradation
Magnesium	Magnesium chelatase	*Rhodobacter sphaeroides* Chlorophyll synthesis
Manganese	Manganese-dependent peroxidases	*Phanerochaete chrysosporium* Carbon cycling–lignin degradation
Molybdenum	Xanthine dehydrogenase Nitrate reductase; nitrogenase	*Pseudomonas putida* *Rhizobium* Purine (caffeine) degradation Nitrogen fixation
Selenium	Selenocysteine (UGA codon) enzymes; glutathione peroxidase; hydrogenases	*Escherichia coli* *Methanococcus voltae* Nitrogen respiration
Tungsten	“True W-enzymes” Aldehyde ferredoxin oxidoreductase	*Thermophilic archaea; Pyrococcus furiosus* Molybdenum antagonist
Zinc	Aminolevulinate dehydratase (inhibited by Pb)	Several bacteria and nearly all Archaea Porphyrin synthesis

[a] The biological availability of metal ions affects enzyme function, potentially constituting the basis for the development of protein-based biosensors.

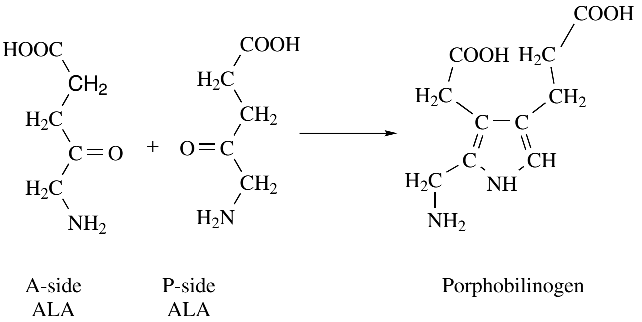

Figure 1. The metalloenzyme, delta-aminolevulinate dehydrate, catalyzes the condensation of two molecules of aminolevulinic acid to produce porphobilinogen, the universal precursor of tetrapyrroles.

are anions or molecules that form coordination compounds or complexes with metals, and commonly occur in natural waters (5). The most important inorganic ligands that form complexes with metals in natural waters are F^-, Cl^-, $SO_4{}^{2-}$, OH^-, $HCO_3{}^-$, $CO_3{}^{2-}$, $HPO_4{}^{2-}$, and NH_3; HS^- and S^{2-} are influential in anoxic waters. The covalent bonding tendency ($\Delta\beta$) and the ionic bonding tendency (Z^2/r) of metals determine the types of ligands that could reduce bioavailability (5). Metals such as Pb are classified as borderline (for Pb, $2 > \Delta\beta > 0$; and $2.5 < Z^2/r < 7.0$). As such, not all of these ligands are relevant to the bioavailability of Pb, but clearly many commonly detected regulated and unregulated inorganic ions in water systems affect Pb bioavailability. This limitation of bioavailability is likely to exist for other toxic heavy metals.

In conclusion, the pervasiveness of metal contamination of water systems indicates a great need to develop sensitive measures of biologically available concentrations of toxic heavy metals in water. The use of microbial enzymes assays for detecting heavy metal toxicity is an emerging technology with abundant resources represented by the physiological diversity inherent in microbial communities. However, further research is required to produce easily interpreted information on enzyme responses to environmental metals in a manner that can supplement the monitoring of other water quality parameters.

Acknowledgment

The Program in Industrial Ecology at the University of California, Irvine, supported this work.

BIBLIOGRAPHY

1. Pang, D.J. (1995). Lead in the environment. In: *Handbook of Ecotoxicology*. D.J. Hoffman, B.A. Rattner, G.A. Burton, and J.C. Cairns Jr. (Eds.). Lewis, Boca Raton, FL, pp. 356–391.

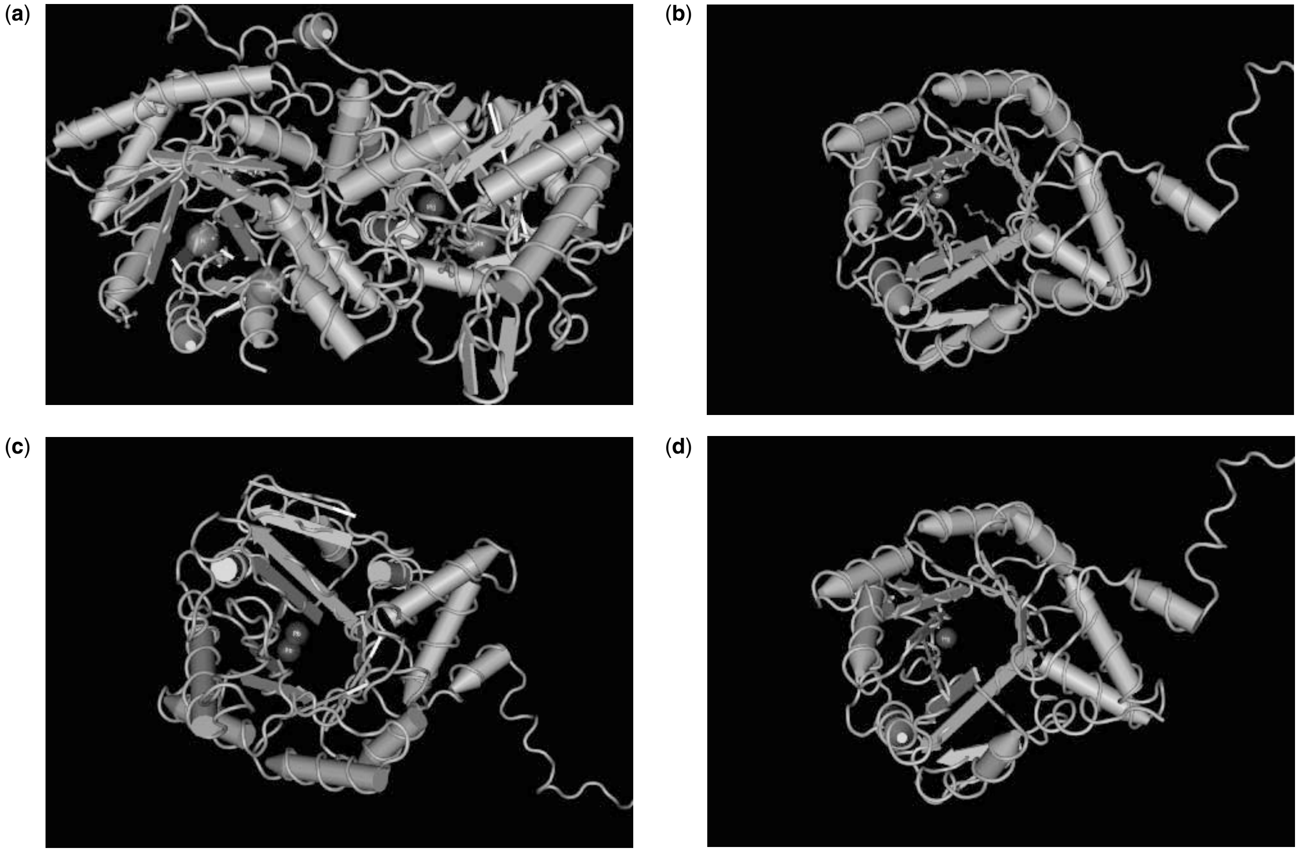

Figure 2. The three-dimensional structure of aminolevulinate dehydratase (ALAD) protein showing the metal-binding domains is presented in (**a**) and (**b**). (**a**) The Mg-dependent ALAD of *P. aeruginosa* with the position of Mg indicated by the arrow. The N-terminal arms of the *P. aeruginosa* ALAD do not project out of the octamer, and the dimers are asymmetric, unlike in Zn-dependent ALAD found in *Archaea* microbes ((**b**) Arrow indicates metal-binding domain with zinc. Phylogenetic analysis of the amino acid sequence of the metal-binding domain of ALAD indicates that the occurrence of various metal ions in ecosystems inhabited by microorganisms exerts a strong selective pressure on the evolution of preference for specific metal ions. For example, the metal-binding domain of ALAD accepts Pb (**c**) and Hg (**d**). Protein structures were created and manipulated by the Cn3D software, version 4.1.

2. Ogunseitan, O.A. (1999). Microbial proteins as biomarkers for ecosystem health. In: *Integrated Assessment of Ecosystem Health* K. Scow, G. Fogg, D. Hinton, and M. Johnson (Eds.). Lewis, Boca Raton, FL, pp. 207–222.
3. Ogunseitan, O.A. (1998). Protein method for investigating mercuric reductase gene expression in aquatic environments. *Appl. Environ. Microbiol.* **64**: 191–204.
4. Hsieh, K.M., Murgel, G.A., Lion, L.W., and Shuler, M.L. (1994). Interactions of microbial biofilms with toxic trace metals: 2. Prediction and verification of an integrated computer model of lead (II) distribution in the presence of microbial activity. *Biotechnol. Bioeng.* **44**: 232–239.
5. Newman, M.C. and Jagoe, C.H. (1994). Ligands and the bioavailability of metals in aquatic environments. In: *Bioavailability: Physical, Chemical, and Biological Interactions* J.L. Hamelink, P.F. Landrum, H.L. Bergman, and W.H. Benson (Eds.). Lewis, Boca Raton, FL, pp. 39–62.
6. Ogunseitan, O.A., Yang, S., and Scheinbach, E. (1999). The delta-aminolevulinate dehydratase activity of *Vibrio alginolyticus* is resistant to lead (Pb). *Biol. Bull.* **197**: 283–284.
7. Ogunseitan, O.A. Yang, S. and Ericson, J. (2000). Microbial delta-aminolevulinate dehydratase as a biosensor of lead (Pb) bioavailability in contaminated environments. *Soil Biol. Biochem.* **32**: 1899–1906.
8. Chauhan, S. and O'Brian, M.R. (1995). A mutant *Bradyrhizobium japonicum* delta-aminolevulinic acid dehydratase with an altered metal requirement functions *in situ* for tetrapyrrole synthesis in soybean root nodules. *J. Biol. Chem.* **270**: 19, 819–823, 827.
9. Conner, E.A. and Fowler, B.A. (1994). Biochemical immunological properties of hepatic delta-aminolevulinate dehydratase in channel catfish (*Ictalurus punctatus*). *Aquatic Toxicol.* **28**: 37–52.
10. Overman, S.R. and Krajicek, J.J. (1995). Snapping turtles (*Chelydra serpentina*) as biomonitors of lead contamination of the big river in Missouri's old lead belt. *Environ. Toxicol. Chem.* **14**: 689–695.
11. Claudio, L., Lee, T., Wolff, M.S., and Wetmur, J.G. (1997). A murine model of genetic susceptibility to lead bioaccumulation. *Fundam. Appl. Toxicol.* **35**: 84–90.
12. Ogunseitan, O.A. (1996). Removal of caffeine in sewage by *Pseudomonas putida*: implications for water pollution index. *World J. Microbiol. Biotechnol.* **12**: 251–256.

MICROBIAL FORMS IN BIOFOULING EVENTS

Baikun Li
Pennsylvania State University
Harrisburg, Pennsylvania

BIOFOULING FORMATION: NUTRIENT ADSORPTION

The term *biofouling* refers to organic debris as a result of the accumulation of inorganic and organic particulate matter along with the growth of various organisms on a surface. The development of a biofouling layer depends on the initial attachment of bacteria to this surface, their eventual increase in number, and the development of a protective slime layer, or glycocalyx.

Biofouling leads to many undesirable problems in industry, such as decreased heat transfer in cooling towers, regrowth in drinking water distribution systems (Fig. 1), and deterioration of materials. In the medical arena, biofilms are responsible for dental plaque and persistent infections on medical implants. It has long been observed that biofilms are much less susceptible to antimicrobial agents than are their planktonic (free-swimming) counterparts.

Most microorganisms, especially heterotrophic bacteria, can easily access nutrients previously deposited on a surface. The attachment of microorganisms to surfaces has a survival function, because the surface in oligotrophic conditions served as sites of nutrients for microorganisms' growth even though they cannot multiply in the bulk aqueous phase. One possible reason for the rapid bacterial growth at surfaces in aquatic environments was the evolvement of microbially produced extracellular enzymes (1). These enzymes can metabolize the macromolecular or particulate substrates. In the bulk aqueous phase, the contact between enzyme and substrate was rare, and products of enzymatic hydrolysis would diffuse away and be unavailable to the bacterium. At the surface, on the other hand, the enzyme and substrate would be in close contact, and provided the enzyme remained active in the adsorbed state, the bacterium should benefit from the end products of the reaction.

Figure 1. The cut section of a cast-iron pipe shows a severe iron bacteria buildup on the pipe walls.

The adhesion of microorganisms to a solid surface is a strategy for access to nutrients. If the bulk water had very few dissolved organic nutrients in it, the growth and reproduction of bacteria would be a very slow process. However, the attachment to a solid surface can substantially accelerate bacterial growth. The bacterial attachment to surfaces shows a great advantage: attached bacteria, known as *sessile or adherent bacteria,* become exposed to a plentiful supply of fresh water containing fresh nutrients that they can extract continuously. Another advantage that sessile bacteria have is their ability to adsorb dissolved organic molecules (food sources for many bacteria) from the water layer. This attachment or adsorption facilitates the growth of microorganisms, especially in oligotrophic environment.

Because of the easy access to nutrients on the surface, the attached bacteria normally have a bigger size than do free-swimming bacteria. When acetate was a substrate, the attached bacteria were larger (average vol. = $0.271\ \mu m^3$) than were planktonic bacteria (average vol. = $0.118\ \mu m^3$), had a higher frequency of dividing cells, and were responsible for most of the acetate uptake (2). These larger attached cells suggest that many of the planktonic (free-swimming) cells were starving and, hence, remained small, whereas the attached cells were growing on substrate available at the surface (3).

MICROORGANISM FORMS IN BIOFOULING FORMATION

One advantage of attachment to a surface is that microorganisms can easily get access to nutrients on the surface and survive in the adverse environment. Starved cells of the marine *Vibrio* DW1 are more adhesive than are unstarved cells, which leads to the suggestion that adhesion may be a tactic in the survival strategy of these organisms under starvation conditions (4). It has been widely accepted that microorganisms attach to available surfaces and initiate the formation of monolayer covering on surfaces; then they develop into thick layers of extensive biofilms, enveloped within extracellular polymeric matrices (5). Generally, there are five diversities in bacterial growth patterns adhered onto surfaces (Fig. 2). These diversities are as follows:

1. *Mother–Daughter or Shedding Cells*. After bacteria adhere to a surface, the small starved cells ($<0.5\ \mu m$ diameter) grow to normal size ($>1.0\ \mu m$ diameter), presumably with substrate accumulating on the surface. The growing cells (now the mother cell) attach to the surface in a perpendicular orientation, and as each gowing cell begins to divide, the daughter cells exhibit a rapid spinning motion before release into the aqueous phase. The mother cell, which remains on the surface, repeats the processes of cellular growth and reproduction. Many daughter cells return to colonize the surface further to get nutrients, and the extensive secondary colonization of the surface results in nutrient uptake exceeding

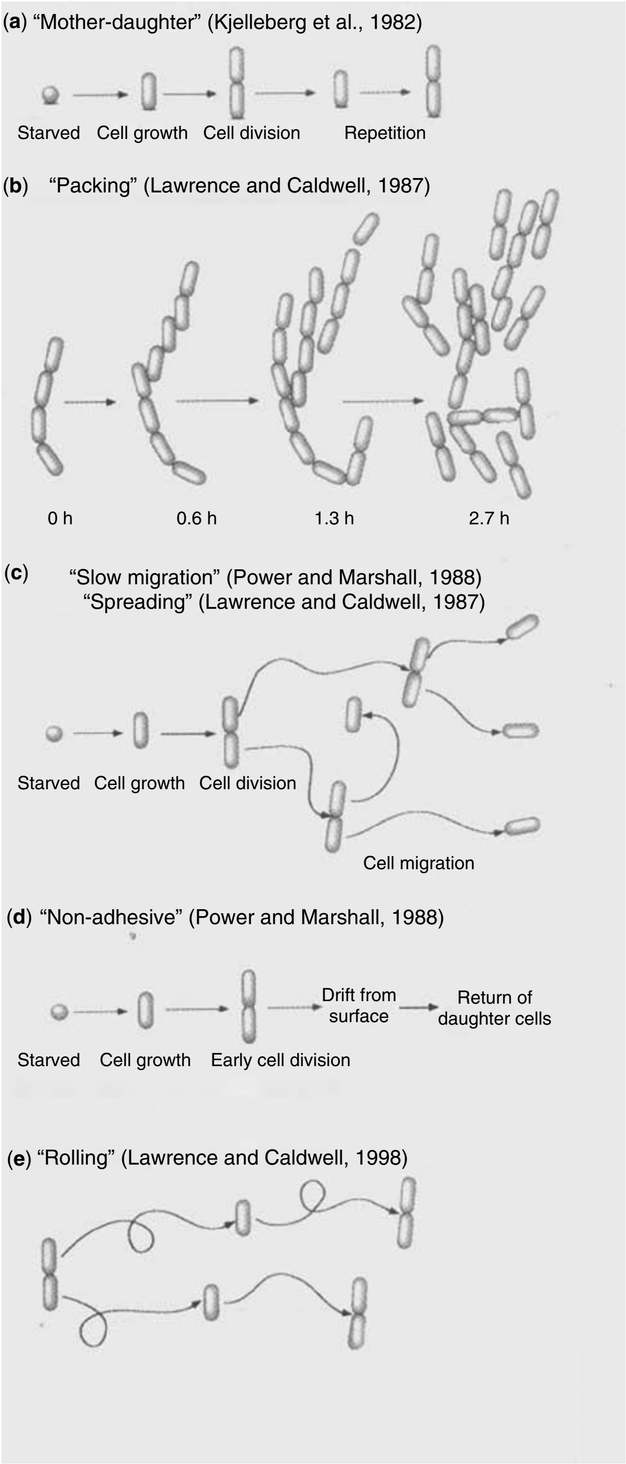

Figure 2. Diversity in growth patterns in bacteria associated with surface. (**a**) Mother–daughter or shedding cells, (**b**) packing cells, (**c**) slow migratory or spreading cells, (**d**) reversibly adhering cells, and (**e**) rolling cells. (Courtesy of Ref. 6.)

nutrient input to the surface, and hence, the cells cease dividing and revert to typical small starvation forms (6).

2. *Packing Cell*. Bacteria divide in a manner that results in the daughter cells aligning adjacent to each other, which eventually leads to the formation of a contiguous monolayer (7).
3. *Slow Migratory or Spreading Cells*. When *Pseudomonas* JD8 cells are in a medium lacking an energy source, the adhering, small, and starved cells grow to normal size and begin to divide, and then the daughter cells migrate very slowly (0.15 μm/min) and at random for variable distances (5–50 times their cell length) across the surface before the next division occurs (8).
4. *Reversibly Adhering or Rolling Cells*. *Vibrio* MH3 is capable of metabolizing surface-bound static acid, the starved cells grow to normal size, and cellular division begins near the surface, but the dividing cells return to the aqueous phase for the final production of daughter cells. That cell division proceeds to completion in the aqueous phase was also described by Lawrence and Caldwell (7) as rolling; that is, cells loosely attached and somersaulting across the surface as they grow and divide similarly may develop as a result of the bacteria reversibly adhering to the surface.
5. *Gliding Cells*. Many cytophage-like gliding bacteria can use complex polysaccharide substrata as an energy substrate. They have a unique form of adhesion that allows them to remain attached to a solid substrate in aqueous condition, and they can glide (translocate) across the surface (9). After the initial attachment to the substratum, *P. aeruginosa* evidently moves on the substratum by means of twitching motility, which suggests that the initial microcolonies are formed by aggregation of bacteria. A large fraction of the biofilm bacteria (*P. aeruginosa, Discosoma* sp.) occasionally swim from one microcolony into another (10). This characteristic certainly makes for efficient use of the solid substrates they metabolize.

THE RELATIONSHIP OF MICROORGANISMS INSIDE BIOFOULING LAYERS

In the biofouling events, different types of microorganisms form either competition or syntrophic correlation. In the mixed biofilms formed by fungi and bacteria, the interactions vary with different microbial species and forms. In the fungal biofouling, *P. aeruginosa* (bacteria) cells readily adhere to *Candida albicans* (a pathogenic fungi) hyphae, but not to the yeast form of *Candida*. Over a 24–48 hour time course, bacteria attach to the *C. albicans* hyphae and form biofilms consisting of densely packed cells surrounded by extracellular matrix (ECM). These findings suggest that the bacteria in biofilm extract nutrients from the underlying *Candida* hyphae, eventually killing the hyphal form of the yeast, and several virulence genes of *P. aeruginosa* that are involved in the killing of *Candida* hyphae in mixed biofilms (11).

Although *P. aeruginosa* is harmful to *Candida* in the biofouling formation, *Streptococcus mutans* can make *C. albicans* strongly adhere to the surface. It was found that *C. albicans* is detached very easily when growing alone, even by a slight movement, when it adheres to plastic rods in a medium containing sucrose (12). However, in mixed cultures of *C. albicans* and *S. mutans, Candida* tend to grow superficially on the surface of a polysaccharide layer formed by the cocci that have colonized the rod surface and are strongly adhered. These findings are confirmed in a subsequent study showing that firm adhesion of *C. albicans* to acrylic resin occurs when the yeasts are incubated simultaneously with *S. mutans* in the presence of glucose. (13). Other studies also found that *S. mutans* grow under and above the yeast and hyphal layers and are clearly adherent to both morphological forms of *C. albicans* (14). This association in the mixed species biofilm makes the biofilm develop very strongly and the fouling is not easy to remove.

Moreover, in a two-species bacterial biofouling event, different metabolic interactions are formed in chemostats (bulk water) and biofouling layers (15). When *P. putida* R1 and *Acinetobacter* C6 strains were grown together in chemostats with limiting concentrations of benzyl alcohol as substrate, they competed for the primary carbon source, which gave *Acinetobacter* C6 a growth advantage. *Acinetobacter* strain C6 outnumbered *P. putida* R1 (cell number ratio: 500:1). However, under similar growth conditions in biofilms, *P. putida* R1 was present in higher numbers than was *Acinetobacter* strain C6 (cell number ratio: 5:1). In addition, these two organisms compete or display commensal interactions depending on their relative physical positioning in the biofilm. In the initial phase of biofilm development, the growth activity of *P. putida* R1 was shown to be higher near microcolonies of *Acinetobacter* strain C6. After a few days, *Acinetobacter* strain C6 colonies were overgrown by *P. putida* R1 cells and new structures developed, in which microcolonies of *Acinetobacter* strain C6 cells were established in the upper layer of the biofilm. In this way, the two organisms developed structural relationships allowing *Acinetobacter* strain C6 to be close to the bulk liquid with high concentrations of benzyl alcohol and allowing *P. putida* R1 to benefit from the benzoate leaking from *Acinetobacter* strain C6.

In sticky biofouling layers, a variety of microorganisms (e.g., hetetrotrophic, autotrophic, aerobic, and anoxic) attach to each other and form a stable and complex microecological community, so the slime layer increases in size and the bacteria become firmly attached to the surface. With the biofilm becoming thicker, oxygen depletion (anaerobiosis) occurs in the lower layers of the biofouling layers, caused by the mass transfer resistance inside the biofilm and the metabolic activity of an aerobic microorganism on the surface layer of the biofilm. Under these conditions, bacteria that require oxygen-free conditions to grow can start to develop. These bacteria include those that can grow in both oxygenated and deoxygenated conditions (known as facultative anaerobic bacteria). Examples are the coliform bacteria, such as *E. coli,* and certain species of *Pseudomonas.* Other bacteria are strictly anaerobic, growing well only in the absence of oxygen. These bacteria include certain species of sulfate reducing bacteria (SRB) and methanogenic bacteria, which are predominant in the deeper layer of biofilm near substratum. The biogas (N_2, H_2S, CO_2, CH_4, etc.) produced by these anoxic and anaerobic bacteria may cause the detachment of the thick biofilm from the surface.

BIOFOULING IN A COOLING TOWER

Biofouling in a cooling tower has been a long-standing problem in industry (Fig. 3). The inorganic and organic contaminants in a cooling tower may precipitate from the flowing water and lead to fouling problems (16). Several major types of microorganisms are involved in cooling tower biofouling, for example, iron bacteria, sulfur reducing/oxidizing bacteria, algae, and fungi.

Major aerobic bacteria that may cause problems in cooling systems are iron-oxidizing bacteria. This group of bacteria are commonly found in well waters that contain soluble iron and large amounts of carbon dioxide. The iron bacteria oxidize the soluble ferrous (Fe^{2+}) to insoluble ferric (Fe^{3+}) as a source of energy, and they use carbon dioxide as a source of carbon for cell growth. They generally produce slimy red-brown flocs of hydrated ferric oxide that can cause plugging of filters and heat exchangers.

Another group of aerobic bacteria in cooling systems is sulfur-oxidizing bacteria, which produce yellow stringy slimes. These bacteria oxidize and reduce sulfur compounds (H_2S) to sulfuric acid ($SO_4{}^{2-}$) in the aerobic environment, which reduces the pH of the water to even lower than 2. The sulfuric acid produced by the sulfur bacteria may cause slime production and corrosion of concrete basins. This group of bacteria is mainly found in refinery cooling towers contaminated with reduced sulfur compounds.

Many types of algae contain the green pigment chlorophyll, which traps the energy from sunlight. Algae synthesize carbon from carbon dioxide from air and require energy from sunlight. Growth of algae primarily occurs in the tower deck area because of ample sunlight. Having been introduced from the atmosphere as spores to water, algae readily become established in areas where plenty of sunlight exists, such as the top of the deck and at the sides of the tower. Algae produce mainly green or blue–green gelatinous slimes, which may be long and filamentous or

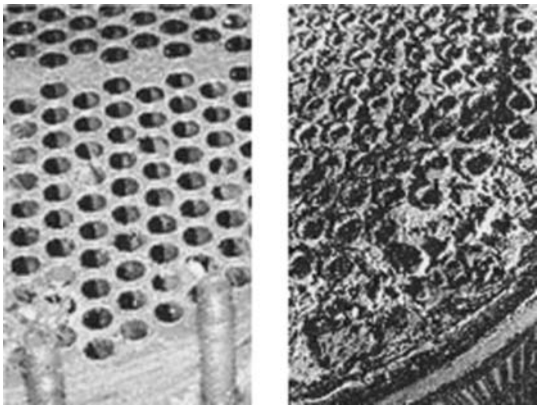

Figure 3. Cooling tower corrosion caused by biofouling. Left: cooling tower wall without biofouling. Right: cooling tower wall with biofouling.

loose and flocculent. The slime species most commonly found is the green algae, the cyanobacteria (originally the blue–green algae) and the diatoms. Prolific algal growth may produce green or blue–green scums in some cooling tower sumps. Algae slimes can plug distribution nozzles and troughs in the cooling tower deck, which causes poor water distribution across the tower and hence reduces cooling efficiency. In the biofouling development, the growth of algae may provide a food source that encourages the growth of other organisms, such as bacteria and fungi.

Fungi are another major group of microorganism that causes problems in the cooling tower. The most insidious problem caused by the growth of fungi is the attack on cooling tower wood, which can go unnoticed for many years. Molds and yeasts can break down the cellulose in the wood, which leads to structural weakening.

BIOFOULING IN MEMBRANE WATER TREATMENT SYSTEM

Biological contamination, known as biofouling, occurs most often during nanofiltration and reverse osmosis in water treatment systems. The attachment of inorganic/organic matters and microorganisms has damaging, often irreversible effects on nanofiltration and reverse osmosis systems (Fig. 4). The major impact of biofouling is the substantial reduction of the flux through the membrane systems, which results in a higher pressure, that causes higher operation cost (18). In some cases, membrane materials are suitable environments for microorganisms to grow, which will cause the membrane to be completely destroyed in a short period of time.

The types of microorganisms, their growth factors, and their concentration in a membrane system greatly depend on critical factors, such as temperature, the presence of sunlight, pH, dissolved oxygen concentrations, and the presence of organic and inorganic nutrients. Both aerobic (oxygen-dependent) and anaerobic (oxygen-independent) bacteria can attach onto the membrane surface and develop biofilm. Aerobic bacteria usually live in an environment of warm, shallow, and sunlit water, with a high dissolved oxygen content, a pH of 6.5 to 8.5, and an abundance of organic and inorganic nutrients. Anaerobic bacteria, on the other hand, are usually present in closed systems with little to no dissolved oxygen and become active when a sufficient amount of nutrients (e.g., organic matter or the debris of dead microorganisms) is present.

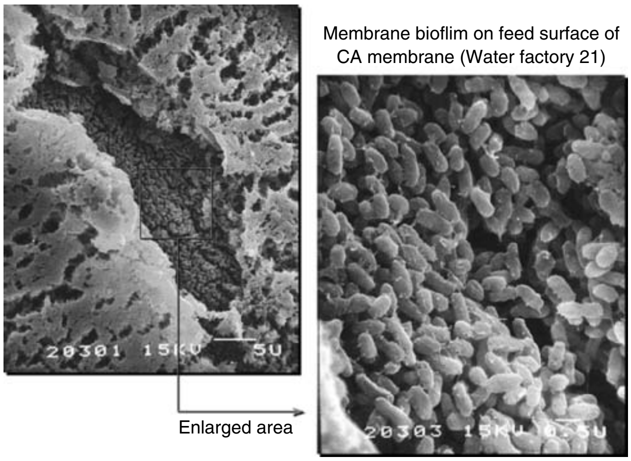

Figure 4. Scanning electron micrographs of mature biofilms formed on cellulose acetate (CA) reverse osmosis (RO) membranes. The RO membranes were fed with a pretreated municipal effluent at Water Factory 21 in Orange County, California. (Courtesy of Ref. 17.)

Another major type of microorganism in membrane biofouling is algae, which can grow on membrane parts exposed to sunlight. The amount of sunlight determines the amount of oxygen that is produced by algae. Aerobic bacteria, which are oxygen-dependent, need the oxygen produced by algae when the dissolved oxygen content in the feed water is insufficient for metabolism. Therefore, the growth of algae boosts the aerobic bacteria growth and further causes biofouling. While the algae die off, they become a food source for bacteria, because they release organic nutrients that bacteria need for growth in a membrane system, especially in an oligotrophic water environment.

METHODS OF BIOFOULING CONTROL

The most efficient approach to solve biofouling is to kill living attached microorganisms by using biocides, or to inhibit the life processes of living organisms without actually killing them by using biostat. It is possible for a chemical to be a biostat to certain organisms at a low concentration and a biocide to the same organisms at a higher concentration. Both biocides and biostats solve the slime problem by affecting various parts of the microbial cell.

Two types of biocides exist: organic biocides, classified as nonoxidizing biocides because they do not oxidize other chemicals or oxidize them only very slightly, and inorganic biocides, classified as oxidizing biocides (e.g., chlorine) because they have a very strong oxidation capacity for many organic and inorganic matters. Nonoxidizing (organic) molecules react with various parts of the microbial cell, generally at a specific site. The most common groups of nonoxidizing biocides are aldehydes, organo-sulfur biocides, chlorinated isocyanurates, organo-nitrogen biocides, amine salts, and heavy metals.

Examples of useful oxidizing biocides are halogens such as chlorine (Cl_2), chlorine dioxide (ClO_2) (19), hydrogen peroxide (H_2O_2), ozone, and chlorine, bromine donors such as chlorinated isocyanurates and hydantoins. The main drawback with the oxidizing biocides is their high level of reactivity to reduce compounds such as iron and H_2S and organic matter during the penetration through biofilm (little residual is left to control microorganisms). So they do not effectively penetrate and disperse microbial slimes. Moreover, some oxidizing biocides (e.g., chlorine) are not effective at solving clogging on the membrane surface. At low chlorine levels, biofilm bacteria can produce extracellular material faster than chlorine can diffuse through biofilm, so they are shielded in slime and make the slime problem worse (20). Therefore, chlorine is usually supplemented by the addition of nonoxidizing biocides in many membrane systems.

Next Page

BIBLIOGRAPHY

1. ZoBell, C.E. and Anderson, D.Q. (1943). The effect of solid surfaces upon bacterial activity. *J. Bacteriol* **46**: 39–56.
2. Pedros-Alio, C. and Brock, T.D. (1983). The importance of attachment to particles for planktonic bacteria. *Arch. Hydrobiol.* **98**: 354–379.
3. Morita, R.Y. (1982). Starvation-survival of heterotrophs in the marine environment. *Adv. Microbiol. Ecol.* **6**: 171–198.
4. Dawson, M.P., Humphrey, B.A., and Marshall, K.C. (1981). Adhesion: a tactic in the survival strategy of a marine *Vibrio* during starvation. *Curr. Microbiol.* **6**: 195–1999.
5. Costerton, J.W. et al. (1987). Bacterial biofilms in nature and disease. *Annu. Rev. Microbiol.* **41**: 435.
6. Fletcher, M. (1996). *Bacterial-Adhesion-Molecular and Ecological Diversity*. Wiley-Liss, New York, p. 71.
7. Lawrence, J.R. and Caldwell, D.E. (1987). Behavior of bacterial stream populations within the hydrodynamic boundary layers of surface microenvironments. *Microbiol. Ecol.* **14**: 15–27.
8. Power, K. and Marshall, K.C. (1988). Cellular growth and reproduction of marine bacteria on surface-bound substrate. *Biofouling* **1**: 163–174.
9. Humphrey, B.A., Dickson, M.R., and Marshall, K.C. (1979). Physicochemical and *in situ* observations on the adhesion of gliding bacteria to surface. *Arch. Microbiol.* **120**: 231–238.
10. Nielson, T.T. et al. (2000). Development and dynamics of *Pseudomonas* sp. biofilm. *J. Bacteriol. Microbiol.* **192**(22): 6482–6489.
11. Hogan, D.A. and Kolter, R. (2002). *Pseudomonas–Candida* interactions: an ecological role for virulence factors. *Science* **296**: 2229–2232.
12. Arendorf, T.M. and Walker, D.M. (1980). *Candida albicans:* its association with dentures, plaque and the oral mucosa. *J. Dent. Assoc. South Africa.* **35**: 563–569.
13. Branting, C., Sund, M.L., and Linder, L.E. (1989). The influence of *Streptococcus mutans* on adhesion of *Candida albicans* to acrylic surfaces *in vitro*. *Arch Oral Biol.* **34**: 347–353.
14. Adams, B., Baillie, G.S., and Douglas, L.J. (2002). Mixed species biofilms of *Candida albicans* and *Staphylococcus epidermidis*. *J. Med. Microbil.* **51**: 344–349.
15. Christensen, B.B., Haagensen, J.A.J., Heydorn, A., and Molin, S. (2002). Metabolic commensalism and competition in a two-species microbial consortium. *Appl. Environ. Microbiol.* **68**(5): 2495–2502.
16. Rogers, J. et al. (1994). Influence of temperature and plumbing material selection on biofilm formation and growth of *Legionella pneumophila* in a model potable water system containing complex microbial flora. *Appl. Environ. Microbiol.* **60**(5): 1585–1592.
17. http://www.desertwildlands.com/AquaMem/AquaMem/biofouling/biofilm_links.html.
18. Vrouwenvelder, J.S. et al. (2003). Tools for fouling diagnosis of NF and RO membranes and assessment of the fouling potential of feed water. *Desalination* **157**(1–3): 361–365.
19. Walker, J.T. et al. (1995). Control of *Legionella pneumophila* in a hospital water system by chlorine dioxide. *J. Ind. Microbiol.* **15**: 384–390.
20. Saeed, M.O. (2002). Effect of dechlorination point location and residual chlorine on biofouling in a seawater reverse osmosis plant. *Desalination* **143**(3): 229–235.

MICROBIOLOGICAL QUALITY CONTROL IN DISTRIBUTION SYSTEMS

Markku J. Lehtola
Ilkka T. Miettinen
National Public Health Institute
Kuopio, Finland

Pertti J. Martikainen
University of Kuopio
Kuopio, Finland

MICROBES IN DRINKING WATER

Good drinking water should not be harmful for health or materials of the distribution system. Undesired health effects can be caused by chemicals or biological agents, including viruses, bacteria, fungi, and protozoa (1–7).

Most of bacteria found in drinking water cannot cause disease in the general population (8). Usually waterborne epidemics are caused by accidental contamination of drinking water, for example, by flooding, surface runoff, or leakage of wastewater pipelines (6). Pathogenic fecal microbes like enteric viruses, protozoan parasites, *Campylobacter*, enterohemorragic *Escherichia coli* (EHEC), *Yersinia enterocolitica*, *Microsporidia*, *Helicobacter pylori*, *Salmonella*, and *Shigella* from contaminated raw water sources may also enter the distribution system as a result of inadequate water treatment (2–6,9,10). These microbes probably do not multiply in the drinking water environment, but they may survive there, especially in biofilms (9,11).

The growth of native nonpathogenic microbes in the distribution system can have several undesired effects on the water quality: the microbial growth can complicate bacteriological water quality monitoring, iron bacteria can precipitate iron-producing iron flocs, the growth of actinomycetes and fungi gives an unpleasant taste and odor to the water, and microbes can destroy the piping materials resulting from their biocorrosion (9,12,13). There are some pathogens or opportunistic pathogens like *Legionella, Pseudomonas aeruginosa, Mycobacteria*, and *Aeromonas*, which can grow in drinking water distribution systems (9,10,14,15). Bacteria and fungi also promote the occurrence of protozoa and higher animals by serving as food for them (13,16).

Microbial cells are attached to almost any surface in aquatic environments (17). Bacteria in biofilms represent the most important part of the bacterial biomass in drinking water distribution networks (18,19), and detachment of bacteria from biofilms accounts for most planktonic cells present in water (20). Biofilms promote the survival and growth of pathogenic bacteria and survival of viruses and parasites (4,9,11,21–23), and they increase the disinfection resistance of bacteria (9,21,24,25).

MICROBIAL QUALITY CONTROL

The growth of microbes in a drinking water distribution system is affected by factors like disinfection, availability of nutrients, temperature, and water residence time.

BIBLIOGRAPHY

1. ZoBell, C.E. and Anderson, D.Q. (1943). The effect of solid surfaces upon bacterial activity. *J. Bacteriol* **46**: 39–56.
2. Pedros-Alio, C. and Brock, T.D. (1983). The importance of attachment to particles for planktonic bacteria. *Arch. Hydrobiol.* **98**: 354–379.
3. Morita, R.Y. (1982). Starvation-survival of heterotrophs in the marine environment. *Adv. Microbiol. Ecol.* **6**: 171–198.
4. Dawson, M.P., Humphrey, B.A., and Marshall, K.C. (1981). Adhesion: a tactic in the survival strategy of a marine *Vibrio* during starvation. *Curr. Microbiol.* **6**: 195–1999.
5. Costerton, J.W. et al. (1987). Bacterial biofilms in nature and disease. *Annu. Rev. Microbiol.* **41**: 435.
6. Fletcher, M. (1996). *Bacterial-Adhesion-Molecular and Ecological Diversity.* Wiley-Liss, New York, p. 71.
7. Lawrence, J.R. and Caldwell, D.E. (1987). Behavior of bacterial stream populations within the hydrodynamic boundary layers of surface microenvironments. *Microbiol. Ecol.* **14**: 15–27.
8. Power, K. and Marshall, K.C. (1988). Cellular growth and reproduction of marine bacteria on surface-bound substrate. *Biofouling* **1**: 163–174.
9. Humphrey, B.A., Dickson, M.R., and Marshall, K.C. (1979). Physicochemical and *in situ* observations on the adhesion of gliding bacteria to surface. *Arch. Microbiol.* **120**: 231–238.
10. Nielson, T.T. et al. (2000). Development and dynamics of *Pseudomonas* sp. biofilm. *J. Bacteriol. Microbiol.* **192**(22): 6482–6489.
11. Hogan, D.A. and Kolter, R. (2002). *Pseudomonas–Candida* interactions: an ecological role for virulence factors. *Science* **296**: 2229–2232.
12. Arendorf, T.M. and Walker, D.M. (1980). *Candida albicans:* its association with dentures, plaque and the oral mucosa. *J. Dent. Assoc. South Africa.* **35**: 563–569.
13. Branting, C., Sund, M.L., and Linder, L.E. (1989). The influence of *Streptococcus mutans* on adhesion of *Candida albicans* to acrylic surfaces *in vitro*. *Arch Oral Biol.* **34**: 347–353.
14. Adams, B., Baillie, G.S., and Douglas, L.J. (2002). Mixed species biofilms of *Candida albicans* and *Staphylococcus epidermidis*. *J. Med. Microbil.* **51**: 344–349.
15. Christensen, B.B., Haagensen, J.A.J., Heydorn, A., and Molin, S. (2002). Metabolic commensalism and competition in a two-species microbial consortium. *Appl. Environ. Microbiol.* **68**(5): 2495–2502.
16. Rogers, J. et al. (1994). Influence of temperature and plumbing material selection on biofilm formation and growth of *Legionella pneumophila* in a model potable water system containing complex microbial flora. *Appl. Environ. Microbiol.* **60**(5): 1585–1592.
17. http://www.desertwildlands.com/AquaMem/AquaMem/biofouling/biofilm_links.html.
18. Vrouwenvelder, J.S. et al. (2003). Tools for fouling diagnosis of NF and RO membranes and assessment of the fouling potential of feed water. *Desalination* **157**(1–3): 361–365.
19. Walker, J.T. et al. (1995). Control of *Legionella pneumophila* in a hospital water system by chlorine dioxide. *J. Ind. Microbiol.* **15**: 384–390.
20. Saeed, M.O. (2002). Effect of dechlorination point location and residual chlorine on biofouling in a seawater reverse osmosis plant. *Desalination* **143**(3): 229–235.

MICROBIOLOGICAL QUALITY CONTROL IN DISTRIBUTION SYSTEMS

Markku J. Lehtola
Ilkka T. Miettinen
National Public Health Institute
Kuopio, Finland

Pertti J. Martikainen
University of Kuopio
Kuopio, Finland

MICROBES IN DRINKING WATER

Good drinking water should not be harmful for health or materials of the distribution system. Undesired health effects can be caused by chemicals or biological agents, including viruses, bacteria, fungi, and protozoa (1–7).

Most of bacteria found in drinking water cannot cause disease in the general population (8). Usually waterborne epidemics are caused by accidental contamination of drinking water, for example, by flooding, surface runoff, or leakage of wastewater pipelines (6). Pathogenic fecal microbes like enteric viruses, protozoan parasites, *Campylobacter*, enterohemorragic *Escherichia coli* (EHEC), *Yersinia enterocolitica*, *Microsporidia*, *Helicobacter pylori*, *Salmonella*, and *Shigella* from contaminated raw water sources may also enter the distribution system as a result of inadequate water treatment (2–6,9,10). These microbes probably do not multiply in the drinking water environment, but they may survive there, especially in biofilms (9,11).

The growth of native nonpathogenic microbes in the distribution system can have several undesired effects on the water quality: the microbial growth can complicate bacteriological water quality monitoring, iron bacteria can precipitate iron-producing iron flocs, the growth of actinomycetes and fungi gives an unpleasant taste and odor to the water, and microbes can destroy the piping materials resulting from their biocorrosion (9,12,13). There are some pathogens or opportunistic pathogens like *Legionella, Pseudomonas aeruginosa, Mycobacteria*, and *Aeromonas*, which can grow in drinking water distribution systems (9,10,14,15). Bacteria and fungi also promote the occurrence of protozoa and higher animals by serving as food for them (13,16).

Microbial cells are attached to almost any surface in aquatic environments (17). Bacteria in biofilms represent the most important part of the bacterial biomass in drinking water distribution networks (18,19), and detachment of bacteria from biofilms accounts for most planktonic cells present in water (20). Biofilms promote the survival and growth of pathogenic bacteria and survival of viruses and parasites (4,9,11,21–23), and they increase the disinfection resistance of bacteria (9,21,24,25).

MICROBIAL QUALITY CONTROL

The growth of microbes in a drinking water distribution system is affected by factors like disinfection, availability of nutrients, temperature, and water residence time.

Waterworks normally disinfect drinking waters by chlorination, ozonation, UV-radiation, or some combination of them. Nutrients, temperature, and residence time are important in microbial growth control when water is distributed without disinfection residuals (ozonation, UV-disinfection) or when disinfection residual is low (chlorine in peripheral parts of distribution system). Also, the pipe material and water hydraulics affect microbial growth in biofilms.

Disinfection

Chlorine is added to water as free chlorine, chlorine dioxide, or chloramines. Free chlorine destroys bacteria by reactions with bacterial enzymes (26), whereas chloramine reacts with nucleic acids, tryptophan, and sulfur-containing amino acids (27). The following factors affect the disinfection efficiency: (1) nature of disinfectant, (2) concentration of disinfectant, (3) length of the contact time, (4) temperature, (5) type and concentration of organisms, and (6) pH (26). One major disadvantage of using chlorine as a disinfectant is the formation of harmful by-products in reactions with organic matter (28,29). Chlorine also degrades organic matter and increases the content of easily available organic carbon for bacteria (AOC) (13,30,31).

There are many advantages to using chloramines instead of free chlorine. Chloramines have a good residual effect and the ability to destroy microbes in biofilms (27). Also, the formation of unwanted chlorinated organic by-products is low with chloramines (32,33). However, the addition of chloramines may enhance growth of ammonia oxidizing bacteria and thus accumulation of nitrite in drinking water (34,35).

Ozonation is a commonly used technique for removing pathogenic microbes, taste, and odor from water (36,37). Ozone is an unstable gas, which has to be generated onsite in the waterworks. The disinfection mechanism is based on the reaction with the double bonds in fatty acids of bacterial cell walls and cell membranes and the protein capsid of viruses (36). In water, ozone is decomposed rapidly back to oxygen and has therefore no residual effects in the distribution system. As a strong oxidant, ozone effectively degrades natural organic matter (30,38). Reactions with natural organic matter increase the content of easily available organic carbon (AOC) and microbially available phosphorus (MAP) in water and thus microbial growth in the distribution network (30,39–41). Ozonation can also produce ozonation by-products; of particular concern is carcinogenic bromate, which is produced when water containing bromide is ozonated (42,43).

UV-radiation ($\lambda = 253.7$ nm) is effective in destroying bacteria and viruses in water (44–47). The destructive effect of UV-radiation is based on DNA and RNA damage such as thymine dimer formation, hydrate formation in the DNA, denaturation of the DNA double strand, and polymerization between nucleic acids and proteins. The absorption maximum of DNA and RNA 255–265 nm is near the wavelengths emitted from mercury low-pressure lamps (253.7 nm). Most microorganisms are inactivated by relatively low UV_{254} dosages—usually in the range of 2–6 mWs/cm^2, although viruses tend to be more resistant to UV_{254} radiation than bacteria (45). There is no evidence on the formation of undesirable by-products or increase in the contents of microbial nutrients like biodegradable organic carbon or microbially available phosphorus by UV-disinfection (45,48,49). Disadvantage of UV-disinfection is similar to ozonation: the lack of residual activity in the distribution system (45).

Biofilms. Some countries use phosphate-based anticorrosion (50–52) chemicals in their distribution systems. The efficiency of phosphates to control biofilm formation is based on neutralization of positively charged corrosion products, such as goethite (α-FeOOH), to negatively charged $FePO_4$, which lowers the adhesion of bacteria to pipe surfaces (50,52). Some studies have shown that these chemicals have not increased microbial growth (50,51,53,54). However, there are geographical regions where use of phosphate against corrosion should be considered carefully because there availability of phosphorus is a limiting factor for microbial growth in drinking waters (see later).

It is difficult to prevent formation of biofilms by chlorine disinfection because residual concentration >1–2 mg/L is needed (20,21,55). Inactivation of fixed bacteria in biofilms needs even a higher concentration of chlorine (>3 mg/L) (25). The inactivation efficiency depends on the composition of the pipe material, the disinfection agent (chloramine vs. chlorine) (27,55), temperature, and water velocity (17).

Microbial Nutrients

Organic Carbon. Except photoautotrophs and chemoautotrophs, organisms need organic compounds for their carbon and energy sources. Organic compounds are partially assimilated into the cell material and partially oxidized to provide energy (56). Bacteria can use a wide range of substrates, but some substrates like amino acids, carboxylic acids, and carbohydrates are more readily usable than others (57).

Aromatic humus substances in drinking water are difficult to use as microbial substrates; usually only a small part of total organic carbon is easily used by microbes (39). There are two commonly used approaches to analyze the microbial usability of aquatic natural organic matter: determination of assimilable organic carbon (AOC) or biodegradable organic carbon (BDOC). AOC is that part of organic matter that can be converted to cell mass. It is expressed as a carbon concentration by means of a conversion factor (39). BDOC is the part of organic carbon that can be mineralized by heterotrophic microbes (58,59).

The first method for AOC assay was published in 1982 by Van der Kooij et al. (39). The test is based on the growth of certain bacterial strains in a water sample, and AOC is calculated using empirical yield values for these bacteria. The bacterial strains used in the test are *Pseudomonas fluorescens* strain P-17 and *Spirillum* strain NOX (60). In regions with high content of organic matter and low amount of phosphorus in water, a modification of the test is required. Miettinen et al. (61) modified the test by adding inorganic nutrients to the water sample to ensure that only organic carbon limits the bacterial growth in the water and suggested the term $\mathrm{AOC}_{\mathrm{potential}}$. They used the term $\mathrm{AOC}_{\mathrm{native}}$ for the assimilable carbon determined without addition of inorganic nutrients.

In the BDOC method, organic carbon is mineralized by the natural microbial community in water, and BDOC is measured as a difference in content of dissolved organic carbon (DOC) before and after (>10 days) incubation of the inoculated water sample (58). There also is a biofilm reactor application for the method. In this application, a water sample is filtered through a column where microbes are attached in biofilm (e.g., glass/sand matrix), and the difference in DOC between inlet and outlet is analyzed (62,63).

Phosphorus. Phosphorus, as a macronutrient, is an essential nutrient for microbes. Bacteria need phosphorus for biosynthesis of nucleic acids, lipopolysaccharides, and phospholipids (56). Phosphate is a vital component of the intracellular energy-transferring ATP system (56). The optimum C:N:P ratio (carbon:nitrogen:phosphorus) for bacterial growth is 100:10:1.

Phosphorus occurs in nature only in the form of chemical compounds, either as inorganic orthophosphate ($\mathrm{HPO_4}^{2-}$, $\mathrm{H_2PO_4}^{-}$) or in organic compounds. In a humus-rich environment, phosphorus is associated with high-molecular-weight humic materials, especially in the presence of iron or manganese (64,65).

Soluble inorganic phosphate is considered to be entirely biologically available (66). All living organisms possess the enzyme, alkaline phosphatase, to convert organic phosphorus to inorganic phosphorus, but only microbes and fungi can excrete the enzyme outside of their cells (exoenzymes) to mineralize and dissolve organic phosphates (66).

Chemical phosphorus analysis has two steps: (1) conversion of phosphorus compounds to dissolved orthophosphate, and (2) colorimetric determination of the dissolved orthophosphate. With these colorimetric standard methods, the detection limit for phosphorus in water is 2–10 μg P/L (60).

Another approach for phosphorus analysis is to measure microbially available phosphorus (MAP) with a bioassay (67). There inorganic nutrients (except phosphorus) and organic carbon are added to the water samples to ensure that inorganic nutrients or organic carbon do not restrict microbial growth. Samples are inoculated with *Pseudomonas fluorescens* bacteria, and the maximum growth of *Pseudomonas fluorescens* is transformed to the amount of microbially available phosphorus with a conversion factor taken from the standardization of the method. The MAP bioassay is very sensitive, and the detection limit is 0.08 μg P/L. MAP represents the part of total phosphorus that is easily available for bacteria. It is important that MAP does not correlate to the content of total phosphorus in water (67).

Nutrient Limitations of Bacterial Growth in Drinking Water. Microbial growth in drinking water is generally limited by AOC or BDOC (39,68–71). In many cases, organic carbon limits microbial growth also in biofilms (18,53–55,72).

The importance of AOC over phosphorus is based on the high ratio of carbon to phosphorus. For example, in an American study, the C:P ratio in drinking water was in the range 100:250 to 100:43, whereas the typical ratio for optimal microbial activity is 100:1 (73). LeChevallier et al. (74) proposed that to limit the growth of coliform bacteria in drinking water, AOC levels should be below 50 μg/L. According to Van der Kooij (13), the AOC concentration should be less than 10 μg/L to limit the microbial aftergrowth. For BDOC, the guideline value of 0.15 mg BDOC/L for biologically stable water has been proposed (69).

In some regions, the correlation between AOC and heterotrophic growth response is weak (73,75) or there is no correlation at all (76,77). Kerneis et al. (78) found no correlation between BDOC and growth of heterotrophic microbes in a drinking water distribution network. It was first found in Finland that microbial growth in drinking water is limited by the availability of phosphorus (79). Subsequently, phosphorus limitation has been reported for drinking waters in Japan (80,81).

In regions where microbial growth in drinking water is limited by phosphorus, a major increase in microbial growth has been achieved with additions of only 1–5 μg/L of $\mathrm{PO_4}$-P (80,82). Sathasivan and Ohgaki (81) reported that phosphorus could become a limiting nutrient at concentrations of 1–3 μg/L. In phosphorus-limited drinking water, the formation of biofilms was also affected by the availability of phosphorus (83).

The reasons for the phosphorus limitation in treated drinking waters are high organic carbon concentration in raw waters and more effective phosphorus removal than that of AOC (31). In fact, in many cases, the AOC content even increases during the drinking water purification process, especially if the waterworks uses ozonation.

BIBLIOGRAPHY

1. Gerba, C.P., Wallis, C., and Melnick, J.L. (1975). Viruses in water: The problem, some solutions. *Environ. Sci. Technol.* **9**(13): 1122–1126.
2. Lippy, E.C. and Waltrip, S.C. (1984). Waterborne disease outbreaks—1946-1980: a thirty-five-year perspective. *J. AWWA* **77**(2): 60–67.
3. Lahti, K. and Hiisvirta, L. (1995). Causes of waterborne outbreaks in community water systems in Finland: 1980–1992. *Water Sci. Technol.* **31**(5–6): 33–36.
4. Ford, T.E. (1999). Microbiological safety of drinking water: United States and global perspectives. *Environ. Health Perspect.* **107**(Supplment I): 191–206.
5. Kukkula, M., Maunula, L., Silvennoinen, E., and von Bonsdorf, C-H. (1999). Outbreak of viral gastroenteritis due to drinking water contaminated by norwalk-like viruses. *J. Infect. Dis.* **180**: 1771–1776.
6. Miettinen, I.T., Zacheus, O., von Bonsdorff, C-H., and Vartiainen, T. (2001). Waterborne epidemics in Finland in 1998–1999. *Water Sci. Technol.* **43**(12): 67–71.
7. Kramer, M.H., Quade, G., Hartemann, P., and Exner, M. (2001). Waterborne diseases in Europe-1986-96. *J. AWWA* **1**: 48–53.

8. Hunter, P. (2002). Epidemiological evidence of disease linked to HPC bacteria. In: *Heterotrophic Plate Counts and Drinking-water Safety*, J. Bartram et al. (Eds.), WHO/IWA Publishing, London, UK, pp. 119–136.
9. Percival, S.L. and Walker, J.T. (1999). Potable water and biofilms: A review of the public health implications. *Biofouling* **42**(2): 99–115.
10. Szewzyk, U., Szewzyk, R., Manz, W., and Schleifer, K.-H. (2000). Microbiological safety of drinking water. *Annu. Rev. Microbiol.* **54**: 81–127.
11. Storey, M.V. and Ashbolt, N.J. (2001). Persistence of two model enteric viruses (B40-8 and MS-2 bacteriophages) in water distribution pipe biofilms. *Water Sci. Technol.* **43**(12): 133–138.
12. LeChevallier, M.W. and McFeters, G.A. (1985). Interactions between heterotrophic plate count bacteria and coliform organisms. *Appl. Environ. Microbiol.* **49**(5): 1338–1341.
13. Van der Kooij, D. (1990). Assimilable organic carbon (AOC) in drinking water. In: *Drinking Water Microbiology, Progress and Recent Developments*, G.A. McFeters (Ed.). Springer-Verlag, Michigan, pp. 57–87.
14. Rusin, P.A., Rose, J.B., Haas, C.N., and Gerba, C.P. (1997). Risk assessment of opportunistic bacterial pathogens in drinking water. *Rev. Environ. Contam. Toxicol.* **152**: 37–38.
15. Norton, C.D., LeChevallier, M.W., and Falkinham, J.O. III (2004). Survival of *Mycobacterium avium* in a model distribution system. *Water Res.* **38**(6): 1457–1466.
16. Sibille, I., Sime-Ngando, T., Mathieyu, L., and Block, J-C. (1998). Protozoan bacteriovory and *Escherichia coli* survival in drinking water distribution systems. *Appl. Environ. Microbiol.* **64**(1): 197–202.
17. Characklis, W.G. and Marshall, K. (1990). Biofilms: A basis for an interdisciplinary approach. In: *Biofilms*, W.G. Characklis and K.C. Marshall (Eds.). John Wiley & Sons, Hoboken, NJ.
18. Laurent, P., Servais, P., and Randon, G. (1993). Bacterial development in distribution networks. study and modelling. *Water Supply* **11**(3/4): 387–398.
19. Zacheus, O.M., Lehtola, M.J., Korhonen, L.K., and Martikainen, P.J. (2001). Soft deposits the key site for microbial growth in drinking water distribution networks. *Water Res.* **35**(7): 1757–1765.
20. Van der Wende, E., Characklis, W.G., and Smith, D.B. (1989). Biofilms and bacterial drinking water quality. *Water Res.* **23**(10): 1313–1322.
21. LeChevallier, M.W., Babcock, T.M., and Lee, R.G. (1987). Examination and characterization of distribution systems biofilms. *Appl. Environ. Microbiol.* **53**(12): 2714–2724.
22. Keevil, C.W. (1989). Chemostat models of human and aquatic corrosive biofilms. In: *Proc. Recent Advances in Microbial Ecology, ISME 5*, T. Hattori et al. (Eds.). Japan Scientific Societies Press, Tokyo.
23. Buswell, C.M. et al. (1998). Extended survival and persistence of *Campylobacter* spp. in water and aquatic biofilms and their detection by immunofluorescent-antibody and—rRNA staining. *Appl. Environ. Microbiol.* **64**(2): 733–741.
24. LeChevallier, M.W., Lowry, C.D., and Lee, R.G. (1990). Disinfecting biofilms in a model distribution system. *J. AWWA* **82**(7): 87–99.
25. Gilbert, P., Allison, D.G., and McBain, A.J. (2002). Biofilms *in vitro* and *in vivo*: do singular mechanisms imply cross-resistance? *J. Appl. Microbiol. Symp. Suppl.* **92**: 98S–110S.
26. White, G.C. (1986). *The Handbook of Chlorination*, 2nd Edn., Van Nostrand Reinhold Company, New York.
27. LeChevallier, M.W., Cawthon, C.D., and Lee, R.G. (1988). Inactivation of biofilm bacteria. *Appl. Environ. Microbiol.* **54**: 2492–2499.
28. Rook, J.J. (1974). Formation of haloforms during chlorination of natural waters. *Water Treatm. Exam.* **23**: 234–243.
29. Boorman, G.A. et al. (1999). Drinking water disinfection byproducts: Review and approach of toxicity evaluation. *Environ. Health Perspect*. **107**(Supplement I): 207–217.
30. Miettinen, I.T., Vartiainen, T., and Martikainen, P.J. (1998). Microbial growth in drinking waters treated with ozone, ozone/hydrogen peroxide or chlorine. *Ozone Sci. Eng.* **20**: 303–315.
31. Lehtola, M.J., Miettinen, I.T., Vartiainen, T., and Martikainen, P.J. (2002). Changes in content of microbially available phosphorus, assimilable organic carbon and microbial growth potential during drinking water treatment processes. *Water Res.* **36**(15): 3681–3690.
32. Trussel, R.R. (1998). An overview of disinfectant residuals in drinking water distribution systems. In: *Preprint of Specialized Conference on Drinking Water Distribution With of Without Disinfectant Residual*, M. Gerlach and R. Gimbel (Eds.). IWW Rheinisch-Westfälisches Institut fur Wassenforschung, Mulheim an der Ruhr, Germany.
33. Nissinen, T.K., Miettinen, I.T., Martikainen, P.J., and Vartiainen, T. (2002). Disinfection by-products in Finnish drinking waters. *Chemosphere* **48**: 9–20.
34. Odell, L.H. et al. (1996). Controlling nitrification in chloraminated systems. *J. AWWA* **88**(7): 86–98.
35. Wilczak, A. et al. (1996). Occurrence of nitrification in chloraminated distribution systems. *J. AWWA* **88**(7): 74–85.
36. Anselme, C., Suffet, I.H., and Mallevialle, J. (1988). Effects of ozonation on tastes and odor. *J. AWWA* **October**: 45–51.
37. Langlais, B., Reckhow, D.A., and Brink, D.R. (Eds.). (1991). *Ozone in Water Treatment, Application and Engineering*. CRC Press/Lewis Publishers, Boca Raton, FL.
38. Glaze, W.H. (1987). Drinking water treatment with ozone. *Environ. Sci. Technol.* **21**(3): 224–230.
39. Van der Kooij, D., Visser, A., and Hijnen, W.A.M. (1982). Determination the concentration of easily assimilable organic carbon in drinking water. *J. AWWA* **74**: 540–545.
40. Van der Kooij, D. and Hijnen, W.A.M. (1984). Substrate utilization by an oxalate-consuming *Spirillum* species in relation to its growth in ozonated water. *Appl. Environ. Microbiol.* **47**: 551–559.
41. Lehtola, M.J., Miettinen, I.T., Vartiainen, T., Myllykangas, T., and Martikainen, P.J. (2001). Microbially available organic carbon, phosphorus and microbial growth in ozonated drinking water. *Water Res.* **35**(7): 1635–1640.
42. Fielding, M. and Hutchison, J. (1995). Formation of bromate and other ozonation by-products in water treatment. *Water Supply* **13**(1): 71–84.
43. Myllykangas, T., Nissinen, T., and Vartiainen, T. (2000). Bromate formation during ozonation of bromide containing drinking water—a pilot scale study. *Ozone Sci. Eng.* **22**: 487–499.
44. Anghern, M. (1984). Ultraviolet disinfection of water. *Aqua* **2**: 109–115.
45. Wolfe, R.L. (1990). Ultraviolet disinfection of potable water. *Environ. Sci. Technol.* **24**(6): 768–773.
46. Oppenheimer, J.A., Jacangelo, J.G., Laine, J-M., and Hoagland, J.E. (1997). Testing the equivalency of ultraviolet

light and chlorine for disinfection of wastewater to reclamation standards. *Water Environ. Res.* **69**(1): 14–24.

47. Parrotta, M.J. and Bekdash, F. (1998). UV-disinfection of small groundwater supplies. *J. AWWA* **90**(2): 71–81.
48. Shaw, J.P., Malley, J.P. Jr., and Willoughby, S.A. (2000). Effects of UV irradiation on organic matter. *J. AWWA* **92**(4): 157–167.
49. Lehtola, M.J. et al. (2003). Impact of UV-disinfection on microbially available phosphorus, organic carbon, and microbial growth in drinking water. *Water Res.* **37**(5): 1064–1070.
50. Abernathy, C.G. and Camper, A.K. (1998). The effect of phosphorus based corrosion inhibitors and low disinfection residuals on distribution biofilms. *Proc. Water Quality Technology Conference. AWWA*, CD-Rom. San Diego, CA.
51. Rompre, A., Prevost, M., Coallier, J., Brisebois, P., and Lavole, J. (2000). Impacts of implementing a corrosion control strategy on biofilm growth. *Water Sci. Technol.* **41**(4–5): 287–294.
52. Appenzeller, B.M., Duval, Y.B., Thomas, F., and Block, J-C. (2002). Influence of phosphate on bacterial adhesion onto iron oxyhydroxide in drinking water. *Environ. Sci. Technol.* **36**(4): 646–652.
53. Appenzeller, B.M.R. et al. (2001). Effect of adding phosphate to drinking water on bacterial growth in slightly and highly corroded pipes. *Water Res.* **35**(4): 1100–1105.
54. Chandy, J.P. and Angles, M.L. (2001). Determination of nutrients limiting biofilm formation and the subsequent impact on disinfectant decay. *Water Res.* **35**(11): 2677–2682.
55. Block, J-C. (1992). Biofilms in drinking water distribution systems. In: *Biofilms, Science and Technology*, L.F. Melo et al. (Eds.), Kluwer Academic Publishers, The Netherlands.
56. Schlegel, H.G. (1997). *General Microbiology*, 7th Edn., Cambridge University Press, Cambridge, UK.
57. Van der Kooij, D., Visser, A., and Oranje, J.P. (1982). Multiplication of *fluorescent pseudomonads* at low substrate concentrations in tap water. *Antonie van Leeuwenhoek* **48**: 229–243.
58. Servais, P., Billen, G., and Hascoet, M-C. (1987). Determination of the biodegradable fraction of dissolved organic matter in waters. *Water Res.* **21**(4): 445–450.
59. Huck, P.M. (1990). Measurement of biodegradable organic matter and bacterial growth potential in drinking water. *J. AWWA* **82**(7): 78–86.
60. Clescerl, L.S., Greenberg, A.E., and Eaton, A.D. (Eds.). (1999). *Standard Methods for Examination of Water and Wastewater*, 20th Edn. American Public Health Association, Washington, DC.
61. Miettinen, I.T., Vartiainen, T., and Martikainen, P.J. (1999). Determination of assimilable organic carbon in humus-rich drinking waters. *Water Res.* **33**(10): 2277–2282.
62. Lucena, F., Frias, J., and Ribas, F. (1990). A New dynamic approach to the determination of biodegradable dissolved organic carbon in water. *Environ. Technol.* **12**: 343–347.
63. Frias, J., Ribas, F., and Lucena, F. (1992). A method for the measurement of biodegradable organic carbon in waters. *Water Res.* **26**(2): 255–258.
64. Jones, R.I., Salonen, K., and De Haan, H. (1988). Phosphorus transformations in the epilimnion of humic lakes: abiotic interactions between dissolved humic materials and phosphate. *Freshwater Biol.* **19**: 357–369.
65. Shaw, P.J., Jones, R.I., and De Haan, H. (2000). The influence of humic substances on the molecular weight distributions of phosphate in epilimnic lake waters. *Freshwater Biol.* **45**(4): 383–393.
66. Jones, R.D. (1997). Phosphorus cycling. In: *Manual of Environmental Microbiology*, J.C. Hurst, G.R. Knudsen, M.J. McInerney, L.D. Stetzenbach, and M.V. Walter (Eds.). ASM Press, Washington, DC, pp. 343–348.
67. Lehtola, M.J., Miettinen, I.T., Vartiainen, T., and Martikainen, P.J. (1999). A new sensitive bioassay for determination of microbially available phosphorus in water. *Appl. Environ. Microbiol.* **65**: 2032–2034.
68. Joret, J.C., Levi, Y., and Volk, C. (1991). Biodegradable dissolved organic carbon (BDOC) content of drinking water and potential regrowth of bacteria. *Water Sci. Technol.* **24**(2): 95–101.
69. Servais, P., Laurent, P., and Randon, G. (1995). Comparison of the bacterial dynamics in various French distribution systems. *J Water SRT-Aqua* **44**(1): 10–17.
70. Prévost, M. et al. (1998). Suspended bacterial biomass and activity in full-scale drinking water distribution systems: impact of water treatment. *Water Res.* **32**(5): 1393–1406.
71. Niquette, P., Servais, P., and Savoir, R. (2001). Bacterial dynamics in the drinking water distribution system of Brussels. *Water Res.* **35**(3): 675–682.
72. Van der Kooij, D., Veenendaal, H.R., Baars-Lorist, C., Van der Klift, D.W., and Drost, Y.C. (1995). Biofilm formation on surfaces of glass and teflon exposed to treated water. *Water Res.* **29**(7): 1655–1662.
73. Zhang, W. and DiGiano, F.A. (2002). Comparison of bacterial regrowth in distribution systems using free chlorine and chloramine: a statistical study of causative factors. *Water Res.* **36**: 1469–1482.
74. LeChevallier, M.W., Schultz, W., and Lee, R.G. (1991). Bacterial nutrients in drinking water. *Appl. Environ. Microbiol.* **57**(3): 857–862.
75. Noble, P.A., Clark, D.L., and Olson, B.H. (1996). Biological stability of ground water. *J. AWWA* **88**(5): 87–96.
76. Gibbs, R.A., Scutt, J.E., and Croll, B.T. (1993). Assimilable organic carbon concentrations and bacterial numbers in a water distribution system. *Water Sci. Technol.* **27**(3–4): 159–166.
77. Miettinen, I.T., Vartiainen, T., and Martikainen, P.J. (1997). Microbial growth and assimilable organic carbon in Finnish drinking water. *Water Sci. Technol.* **35**(11–12): 301–306.
78. Kerneis, A., Nakache, F., Deguin, A., and Feinberg, M. (1995). The effects of water residence time on the biological quality in a distribution network. *Water Res.* **29**(7): 1719–1727.
79. Miettinen, I.T., Vartiainen, T., and Martikainen, P.J. (1996). Contamination of drinking water. *Nature* **381**: 654–655.
80. Sathasivan, A., Ohgaki, S., Yamamoto, K., and Kamiko, N. (1997). Role of inorganic phosphorus in controlling regrowth in water distribution system. *Water Sci. Technol.* **35**(8): 37–44.
81. Sathasivan, A. and Ohgaki, S. (1999). Application of new bacterial regrowth potential method for water distribution system—a clear evidence of phosphorus limitation. *Water Res.* **33**(1): 137–144.
82. Miettinen, I.T., Vartiainen, T., and Martikainen, P.J. (1997). Phosphorus and bacterial growth in drinking water. *Appl. Environ. Microbiol.* **63**: 3242–3245.
83. Lehtola, M.J., Miettinen, I.T., and Martikainen, P.J. (2002). Biofilm formation in drinking water affected by low concentrations of phosphorus. *Can. J. Microbiol.* **48**(6): 494–499.

WATER QUALITY MODELS FOR DEVELOPING SOIL MANAGEMENT PRACTICES

HARRY X. ZHANG
Parsons Corporation
Fairfax, Virginia

INTRODUCTION

Water quality models are used extensively in water quality planning and pollution control. Models are applied to answer a variety of questions, support watershed planning and analysis, and develop total maximum daily loads (TMDLs). Watershed loading models include techniques designed primarily to predict pollutant movement from the land surface to waterbodies. TMDLs result in allocating the quantity of pollution that can be discharged from point sources (wasteload allocation) and nonpoint sources (load allocation) to ensure that water quality standards are achieved within a specified margin of safety (1). Management measures can be applied to achieve all of the pollution control needed by pollutant sources to achieve the load allocation. Best management practice (BMP) is generally defined as a practice or combination of practices that are the most effective and practicable (including technological, economic, and institutional considerations) means of controlling point and nonpoint pollutants at levels compatible with economic and environmental quality goals (2).

It is important to estimate the collective impacts of all management activities in a watershed to evaluate whether water quality goals will be achieved. In watersheds that have easily characterized problems, it may be easy to project that water quality benefits will be achieved by implementing management measures such as erosion and sediment control and runoff minimization. However, in a watershed that has multiple land uses where agriculture is considered to contribute groups of pollutants, it is more complicated to estimate the combined impacts of a variety of management measures and practices on a fairly large number of diverse farming operations. In this type of situation, computer modeling may be needed. A variety of models exist to help assess the relative benefits of implementing practices at the field and watershed level. Managers can use models to develop goals and objectives for watershed and waterbody use and protection, to develop and to evaluate the effectiveness of various management and mitigation alternatives to improve existing water quality, and to determine the response time of a waterbody after implementing a management alternative. In particular, models can help understanding key “cause-and-effect” processes within the natural environment, developing waste load allocations and TMDL analysis, and evaluating the effectiveness of various contaminant reduction scenarios.

BRIEF OVERVIEW OF NONPOINT SOURCE WATER QUALITY MODELS

Watershed loading models range from simple loading rate assessments in which loads are a function of land use type only, to complex simulation techniques that more explicitly describe the processes of rainfall, runoff, sediment detachment, and transport to receiving waters. Some loading models operate on a watershed scale, integrating all loads within a watershed, and some allow subdividing the watershed into contributing subbasins. Given the complex nature of the processes and interactions that determine the fate and transport of a pollutant or chemical in nature, mathematical models can attempt to represent only the principal components of the environment that influence a given water quality variable. It is virtually impossible to incorporate fully all of the mechanisms and processes that occur in nature. Therefore, there are varying degrees of simplicity or complexity embodied in watershed and water quality modeling.

Many computer models are available to assess soil management practices. Based on their complexity, operation, time step, and simulation technique, watershed-loading models can be grouped into three categories: simple methods, midrange models, and detailed models.

Simple Models

The major advantage of simple methods is that they can provide a rapid means of identifying critical areas with minimal effort and data requirements. Simple methods are typically derived from empirical relationships between physiographic characteristics of the watershed and pollutant export. Simple methods are often used when data limitations and budget and time constraints preclude using complex models. They are used to diagnose nonpoint source pollution problems, when information is relatively limited. They can be used to support an assessment of the relative significance of different sources, guide decisions for management plans, and focus continuing monitoring efforts.

Typically, simple methods rely on a large-scale aggregation of characteristics and neglect features of small patches of land. They rely on generalized sources of information and therefore have low to medium requirements for site-specific data. Default values provided for these methods are derived from empirical relationships that are based on regional or site-specific data. The estimates are usually expressed as mean annual values. Simple methods provide only rough estimates of sediment and pollutant loadings and have limited predictive capability.

In addition, field-scale models, which have traditionally specialized in agricultural systems, are loading models that are designed to operate on a smaller, more localized scale. Field-scale models have often been employed to aid in selecting management measures and practices.

Midrange Models

The advantage of midrange watershed loading models is that they evaluate pollution sources and impacts over broad geographic scales and, therefore, can assist in defining target areas for pollution mitigation programs in a watershed. The midrange model compromises between the empiricism of simple methods and the complexity of detailed mechanistic models (e.g., relating pollutant loading to hydrologic and erosion processes). Mechanistic models usually include detailed input–output features.

Detailed Models

Detailed models best represent the current understanding of watershed processes that affect pollution generation. Detailed models are best able to identify causes of problems rather than describe overall conditions. If properly applied and calibrated, detailed models can provide relatively accurate predictions of variable flows and water quality at any point in a watershed. The additional precision they provide, however, comes at the expense of considerable time and resources for data collection and model application.

Model Comparison

Detailed models incorporate the manner in which watershed processes change over time continuously rather than relying on simplified terms for rates of change. Algorithms in detailed models more closely simulate the physical processes of infiltration, runoff, pollutant accumulation, in-stream effects, and groundwater–surface water interaction. The input and output of detailed models also have greater spatial and temporal resolution. Moreover, the manner in which physical characteristics and processes differ over space is incorporated within the governing equations. However, input data file preparation and calibration of detailed models require professional training and adequate resources. Their added accuracy might not always justify the amount of effort and resources required.

In comparison, simple methods use large simulation time steps to provide longer term averages or annual estimates. Their accuracy may decrease in estimating seasonal or storm loading because they cannot capture the large fluctuations of pollutant loading or concentration usually observed when smaller time steps are employed. Neither the simple nor the midrange models consider degradation and transformation processes, and few incorporate detailed representations of pollutant transport within and from the watershed. Although model applications might be limited to relative comparisons, they can often provide useful information to water quality managers for watershed-level planning decisions. If adequate site-specific data are not available, values can be estimated and assumptions can be made, but these estimates will increase the uncertainty in the model predictions. It will be very difficult to validate whether simple or complex models provide better estimates. Therefore, if data availability does not reach the level that a detailed model requires, it is recommended that a simpler model be employed.

REVIEW OF WATER QUALITY MODELS FOR DEVELOPING SOIL MANAGEMENT PRACTICES

Hydrologic Simulation Program–FORTRAN (HSPF) and WinHSPF (in BASINS 3.0)

Developed and maintained by the EPA (http://www.epa.gov/ost/basins/) and the USGS (http://water.usgs.gov/software/hspf.html), the Hydrological Simulation Program–FORTRAN (HSPF) is one of the most comprehensive modeling systems for simulating watershed hydrology, point and nonpoint loading, and receiving water quality for both conventional pollutants and toxicants. The HSPF model was developed in the early 1960s as the Stanford Watershed Model. In the 1970s, water quality processes were added. The development of a Fortran version incorporating several related models using software engineering design and development concepts was funded by the EPA Research Laboratory (Athens, Ga.) in the late 1970s. In the 1980s, the USGS and EPA jointly developed preprocessing and postprocessing software, algorithm enhancements, and use of the USGS Watershed Data Management (WDM) system.

HSPF simulates hydrologic and water quality processes on land surfaces, streams, and impoundments. HSPF is generally used for a watershed-based analysis of the effects of land use, point and nonpoint source treatment alternatives (e.g., soil management practices), and flow diversions. Time series of the runoff flow rate, sediment yield, and user-specified pollutant concentrations can be generated at any point in the watershed. Compared with many other watershed loading models, HSPF has several advantages. First, HSPF is both a watershed model and a receiving water model; it can simulate both the pollutant loads on land surfaces and in-stream water quality in a complex watershed. Second, HSPF can be run under both continuous and storm situations. Data requirements for HSPF are extensive, and calibration and verification are strongly recommended. Because of its comprehensive nature, the HSPF model requires highly trained personnel. A team effort is recommended to apply it to real case studies.

WinHSPF is a new interface to HSPF Version 12 that is replacing Nonpoint Source Model (NPSM) from Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) Version 2.0. Within BASINS V3.0 system, GenScn is a model postprocessing and scenario analysis tool that is used to analyze output from HSPF and SWAT. WDMUtil is used to manage and create the WDM files that contain the meteorologic data and other time series data used by HSPF. BASINS 3.0 includes several new data and functions including 1° DEM grids, an automatic delineation tool that creates watershed boundaries based on DEM grids, and a new watershed report function for land use, topography, and hydrologic response units.

Applications. HSPF and BASINS are widely used to for develop TMDLs for bacteria and many other wet weather pollutants. HSPF and BASINS produce pollutant concentration time series output that make it easy to determine the frequency and duration of water quality standard exceedances. Control scenario alternatives (e.g., soil management practices) can be evaluated for effectiveness in addressing water quality standard exceedances and the extent to which they would bring excessive costs to a community. The water quality impact of land use changes can also be conveniently evaluated. In addition, HSPF and BASINS can support source water assessments to examine the potential pollution risks to public water supplies.

The latest BASINS V3.0 provides a range of watershed models from which watershed managers can select the one that best meets their needs for an appropriate level of model complexity, control scenario detail, uncertainty, and analysis cost. The three watershed models in BASINS 3.0 are HSPF v12, Soil and Water Assessment Tool (SWAT), and the simple PLOAD model. SWAT and PLOAD models will be further discussed.

Watershed Modeling System (WMS)

The Watershed Modeling System (WMS) (http://www.emrl.byu.edu/wms.htm) is a comprehensive simulation model for hydrologic analysis. It was developed by the Environmental Modeling Research Laboratory of Brigham Young University in cooperation with the U.S. Army Corps of Engineers Waterways Experiment Station (USACE-WES). WMS merges information obtained from terrain models and GIS with industry-standard, lumped-parameter, hydrologic analysis models such as HEC-1, TR-20, and HSPF. WMS V6.1 provides tools for all phases of watershed modeling, including automated watershed and subbasin delineation, geometric parameter computation, hydrologic parameter computation, and result visualization. WMS is a powerful tool for analyzing and visualizing watersheds, which includes interfaces to several industry-standard hydrologic models. Each model can be set up and run within WMS. Model schematics are created and maintained automatically by WMS as you delineate a watershed from Map, Triangulated Irregular Networks (TINs), or Digital Elevation Models (DEMs) data. Terrain models can obtain geometric attributes such as area, slope, and runoff distances. WMS contains several calculators to aid in hydraulic/hydrologic design and analysis. These include an interface to HY-8 for culvert design/analysis, an open channel calculator, a detention basin calculator, a curb and gutter calculator, and a weir calculator. In addition, a calculator for Lag Time and/or Time of Concentration is included in WMS to aid in hydrologic model input. This calculator allows using any basin or flow path equation to compute travel times in the watershed and evaluate the hydrologic impact of management practices.

Soil and Water Assessment Tool (SWAT)

The Soil and Water Assessment Tool (SWAT) (http://www.brc.tamus.edu/swat/) is a watershed-scale model developed by Dr. Jeff Arnold for the USDA Agricultural Research Service (ARS). SWAT simulates the effect of agricultural management practices such as crop rotation, conservation tillage, residue, nutrient, pesticide management, and improved animal waste application methods on water quality. SWAT incorporates features of several ARS models and is a direct outgrowth of the SWRRB model (Simulator for Water Resources in Rural Basins) (3,4). The SWRRB model has been used on several watersheds to assess management practices and to test its validity (5,6). Specific models that contributed significantly to the development of SWAT were CREAMS (Chemicals, Runoff, and Erosion from Agricultural Management Systems) (7), GLEAMS (Groundwater Loading Effects on Agricultural Management Systems) (8), and EPIC (Erosion-Productivity Impact Calculator) (9).

SWAT is a continuous time model, that is, a long-term yield model. The model is not designed to simulate detailed, single flood routing. Instead, SWAT was developed to predict the impact of land management practices on water, sediment, and agricultural chemical yields in large complex watersheds where soils, land use, and management conditions vary over long periods of time. To satisfy this objective, the model is physically based. Rather than incorporating regression equations to describe the relationship between input and output variables, SWAT requires specific information about weather, soil properties, topography, vegetation, and land management practices in the watershed. The physical processes associated with water movement, sediment movement, crop growth, and nutrient cycling, are directly modeled by SWAT using this input data. The benefits of this approach are (1) watersheds without monitoring data (e.g., stream gauge data) can be modeled, (2) the relative impact of alternative input data (e.g., changes in management practices, climate, vegetation) on water quality or other variables of interest can be quantified, and (3) uses readily available inputs. Furthermore, SWAT is computationally efficient. Very large basins or a variety of management strategies can be simulated without excessive investment of time or money. SWAT enables users to study long-term impacts.

A new watershed model, the SWAT 2002 has been added to the EPA BASINS V3.0 system. SWAT is a watershed-scale model developed to predict the impact of land management practices on water, sediment, and agricultural chemical yields in large complex watersheds where soils, land use, and management conditions vary over long periods of time.

Applications. SWAT is ideally suited for simulating agricultural watersheds where fertilizer and pesticide application rates are known and where altering agricultural management practices is viewed as a viable and necessary way to meet water quality standards. SWAT is ideally suited for nitrogen, phosphorous, and pesticide TMDLs due to its use of prepackaged nutrient and pesticide application loads in standard forms (e.g., by type of manure or fertilizer and by the common name of the pesticide) as well as other agriculture specific inputs including tillage practices and crop rotation. SWAT requires a considerable amount of information on agricultural management practices but readily houses and internally processes this information within the interface. Management options are already included, so there is a direct translation to soil conservation practices and BMP-based controls. SWAT can be used to determine the likely reduction in sediment loads from installing small dams in critical areas in watersheds or how employing other land-based best management practices, such as terraces, vegetative buffer strips, or other cropland BMPs, would cut silt runoff/sediment loads to lakes and reservoirs.

CREAMS and GLEAMS

CREAMS, first developed in 1970s by the USDA, is a field-scale model designed to simulate the impact of land management on water, sediment, nutrients, and pesticides leaving the edge of a field. A number of other USDA models such as GLEAMS, SWRRB, AGNPS and EPIC, trace their origins to the CREAMS model.

GLEAMS is a nonpoint source model that focuses on pesticide and nutrient groundwater loadings. GLEAMS (10) simulates the effects of management practices and irrigation options on edge of field surface runoff, sediment, and dissolved and sediment-attached nitrogen, phosphorus, and pesticides. The model considers the effects of crop planting date, irrigation, drainage, crop rotation, tillage, residue, commercial nitrogen and phosphorus applications, animal waste applications, and pesticides on pollutant movement. The model has been used to predict the movement of pesticides (11), nutrients, and sediment from various combinations of land uses and management (12,13).

Erosion–Productivity Impact Calculator (EPIC)

The Erosion–Productivity Impact Calculator (EPIC) (http://www.brc.tamus.edu/epic/index.html) was originally developed to simulate the impact of erosion on crop productivity and has now evolved into a comprehensive agricultural management, field-scale, nonpoint source loading model. In the early 1980s, teams of USDA Agricultural Research Service (ARS), Soil Conservation Service (SCS) (now Natural Resources Conservation Service, or NRCS), and Economic Research Service (ERS) scientists developed EPIC to quantify the costs of soil erosion and the benefits of soil erosion research and control in the United States. Developed in late 1980s, the model uses a daily time step to simulate weather, hydrology, soil temperature, erosion–sedimentation, nutrient cycling, tillage, crop management and growth, pesticide and nutrient movement with water and sediment, and field-scale costs and returns. EPIC (14) simulates the effect of management strategies on edge of field water quality and nitrate nitrogen and pesticide leaching to the bottom of the soil profile. The model considers the effect of crop type, planting date, irrigation, drainage, rotations, tillage, residue, commercial fertilizer, animal waste, and pesticides on surface and shallow groundwater quality. The EPIC model has been used to evaluate various cropland management practices (15,16).

AGNPS

Agricultural Nonpoint Source Pollution Model 2001 (AGNPS 2001) (http://www.sedlab.olemiss.edu/agnps.html) is a joint USDA–ARS and NRCS system of computer models developed to predict nonpoint source pollutant loadings within agricultural watersheds. AGNPS is a midrange model. It contains a continuous simulation, surface runoff model designed to assist in determining BMPs, the setting of TMDLs, and for risk and cost/benefit analyses.

AGNPS 2001 is a tool for evaluating the effect of management decisions that impact a watershed system. The AGNPS 2001 system is a direct update of the AGNPS 98 system of modules containing many enhancements. These enhancements have been included to improve the capability of the program and to automate many of the input data preparation steps needed for use with large watershed systems. The capabilities of RUSLE, used by USDA-NRCS to evaluate the degree of erosion on agricultural fields and to guide development of conservation plans to control erosion, have been incorporated into AnnAGNPS. The AnnAGNPS model has been applied to many field and watershed size areas to estimate pollutant runoff from various land uses and management practices (6,15,17–19).

Revised Universal Soil Loss Equation (RUSLE)

The Revised Universal Soil Loss Equation (RUSLE) (http://www.sedlab.olemiss.edu/rusle/) is an easily and widely used computer program that estimates rates of soil erosion caused by rainfall and associated overland flow. The most current version of RUSLE is Version 1.06b released on January 2001 (20). In the United States, the USDA–NRCS is the principal user of RUSLE and has implemented this approach in most of its local field offices. The NRCS is the major source for data needed to apply RUSLE and is the leading authority on field application of RUSLE. RUSLE is used by numerous government agencies, private organizations, and individuals to assess the degree of erosion, identify situations where erosion is serious, and guide development of conservation plans to control erosion. RUSLE has been applied to cropland, rangeland, disturbed forest lands, landfills, construction sites, mining sites, reclaimed lands, military training lands, parks, land disposal of waste, and other land uses where mineral soil material is exposed to the erosive forces of raindrop impact and overland flow. RUSLE estimates average annual soil loss, expressed as mass per unit area per year, which is defined as the amount of sediment delivered from the slope length assumed in the RUSLE computation.

Water Erosion Prediction Project (WEPP)

The Water Erosion Prediction Project (WEPP) model (http://topsoil.nserl.purdue.edu/nserlweb/weppmain/) is a process-based, distributed parameter and continuous simulation model for erosion prediction. WEPP is expected to become a major component of the “conservation tool kit.” It is a computer program designed to be employed by the same personnel currently using USLE (Universal Soil Loss Equation) or RUSLE. By analyzing how farming and land use affect soil erosion and sediment delivery, WEPP promises better conservation planning, better project planning, and optimum resource inventory and assessment. It will help users select the best erosion control practices, aid in choosing optimum locations for future project sites, and evaluate erosion and sedimentation over specified areas. It is applicable to small watersheds (field-sized) and can simulate small profiles up to large fields. It mimics the natural processes that are important in soil erosion.

WEPP can be applied on the field scale to simulate single hillslope erosion or more complex watershed-scale erosion: Hillslope Applications and Watershed Applications. WEPP (21) simulates water runoff, erosion, and sediment delivery from fields or small watersheds. Management practices, including crop rotation, planting and harvest date, tillage, compaction, strip-cropping, row arrangement, terraces, field borders, and windbreaks, can be simulated. The WEPP model has been applied to various land use and management conditions (22,23).

Nitrate Leaching and Economic Analysis Package (NLEAP)

The Nitrate Leaching and Economic Analysis Package (NLEAP) (http://gpsr.ars.usda.gov/products/nleap/nleap.htm) is a field-scale computer model developed to provide a rapid and efficient method for determining potential nitrate leaching from agricultural practices. It uses basic information concerning on-farm management practices, soils, and climate to project nitrogen budgets and nitrate leaching indexes. NLEAP calculates potential nitrate leaching below the root zone and to groundwater supplies. NLEAP has three levels of analysis to determine leaching potential: an annual screening, a monthly screening, and an event-by-event analysis.

The NLEAP model was designed to answer questions regarding potential leaching of nitrate. The processes modeled include movement of water and nitrate, crop uptake, denitrification, ammonia volatilization, mineralization of soil organic matter, nitrification, and mineralization–immobilization of crop residue, manure, and other organic wastes. The limitations of NLEAP include modeling soil nitrogen processes in organic soils that is not currently available. NLEAP does not predict yield reductions caused by pests or nutrient deficiencies. However, the user should consider the effects of these problems when estimating crop yield. The model should not be used where rapid water infiltration, leaching, denitrification, and ammonium volatilization require time steps smaller than one day. NLEAP (24) evaluates the potential of nitrate nitrogen leaching due to land use and management practices. The NLEAP model has been used to predict the potential for nitrogen leaching under various management options (25,26). Most recently, it has been applied to compare the impact of TMDL allocation scenarios on nitrate leaching (27).

Root Zone Water Quality Model (RZWQM)

The Root Zone Water Quality Model (RZWQM) (http://gpsr.ars.usda.gov/products/rzwqm.htm) was developed during the past 10 years by a team of USDA–ARS scientists. A majority of the team members are part of the present Great Plains Systems Research Unit in Fort Collins, Colorado. RZWQM is a one-dimensional process-based model that simulates the growth of a plant and the movement of water, nutrients, and agrochemicals over, within, and below the crop root zone of a unit area of an agricultural cropping system under a range of common management practices. The primary use of RZWQM will be as a tool for assessing the environmental impact of alternative agricultural management strategies on the subsurface environment. These alternatives may include conservation plans on a field-by-field basis; tillage and residue practices; crop rotations; planting date and density; and irrigation, fertilizer, and pesticide scheduling (method of application, amounts, and timing). The model predicts the effects of these management practices on the movement of nitrate and pesticides to runoff and deep percolation below the root zone.

Pesticide Root Zone Model (PRZM)

The EPA's Pesticide Root Zone Model (PRZM) (http://www.epa.gov/ceampubl/gwater/przm3/index.htm) is a one-dimensional, dynamic, compartmental model that can be used to simulate chemical movement in unsaturated soil systems within and immediately below the plant root zone. It has two major components: hydrology/hydraulics and chemical transport. The hydrologic component for calculating runoff and erosion is based on the Soil Conservation Service curve number technique and USLE. The chemical transport component can simulate pesticides or organic and inorganic nitrogen species.

PRZM can simulate multiple zones. This allows combining different root zone and vadose zone characteristics into a single simulation. Zones can be visualized as multiple land segments joined together horizontally. PRZM (28) simulates the movement of pesticides in unsaturated soils within and immediately below the root zone. Several different field crops can be simulated and up to three pesticides are modeled simultaneously as separate parent compounds or metabolites. The PRZM model has been used under various conditions to assess pesticide leaching under fields (11,13).

Generalized Watershed Loading Function (GWLF)

The Generalized Watershed Loading Function (GWLF) model (http://www.avgwlf.psu.edu/) was developed by Haith and Shoemaker (29). The GWLF model can simulate runoff, sediment, and nutrient loadings from a watershed, given variable-size source areas (i.e., agricultural, forested, and developed land). ANSWERS (30), which has been used in GWLF, is a spatially distributed watershed model. The model is primarily a runoff and sediment model because soil nutrient processes are not simulated. The ANSWERS model has been applied to several small field-sized areas using various management practices (6,31).

During the last 5–10 years, the Pennsylvania Department of Environmental Protection (PADEP) has recognized the indispensability of GIS technology and endeavored to integrate it into all of the agency's internal program areas. Pennsylvania State University has been assisting PADEP in developing and implementing various GIS-based watershed assessment tools, including one called AVGWLF that facilitates using the GWLF model via a GIS software (ArcView) interface. The comprehensive modeling approach was based on this GWLF/ArcView interface that enables accurate prediction of nutrient and sediment loads in watersheds throughout the state of Pennsylvania, particularly for those watersheds for which historical stream monitoring data do not exist. This methodology relies on using statewide data sets to derive reasonably good estimates of various critical model parameters that exhibit significant spatial variability within the state. GWLF is a continuous simulation model that uses daily time steps for weather data and water balance calculations. Monthly calculations are made for sediment and nutrient loads, based on the daily water balance accumulated to monthly values.

Source Loading and Management Model (SLAMM)

The Source Loading and Management Model (SLAMM) (http://wi.water.usgs.gov/slamm/) was originally developed in the late 1970s to understand better the relationships between sources of urban runoff pollutants and runoff quality. Runoff is calculated by a method developed by the University of Alabama at Birmingham for small-storm hydrology. SLAMM is strongly based on actual field observations, and relies minimally on theoretical processes that have not been adequately documented or confirmed in the field. SLAMM is used mostly as a planning tool to understand better sources of urban runoff pollutants and their control. SLAMM incorporates unique process descriptions to predict more accurately the sources of runoff pollutants and flows for the storms of most interest in water quality analyses and quantification of the impacts of management practices.

Regression Methods and Simple Spreadsheet Models

USGS Regression Method. The USGS has developed equations for determining pollutant loading rates based on regression analyses of data from sites throughout the country (76 gauging stations in 20 states). The regression approach is based on a statistical description of historic records of storm runoff responses on a watershed level (32,33). This method may be used for rough preliminary calculations of annual pollutant loads when data and time are limited. Input required for this level of modeling includes drainage data, percent imperviousness, mean annual rainfall, general land use pattern, and mean minimum monthly temperature. Application of this method provides mean planning loads and corresponding confidence intervals for storms.

Simple Method by MWCOG. The “simple method,” as its name implies, is an easy-to-use empirical equation for estimating pollutant loadings of an urban watershed by the Metropolitan Washington Council of Governments (MWCOG). The method is applicable to watersheds less than 1 square mile in area, and can be used for analyzing a smaller watershed or site planning. The method was developed using the database generated during a Nationwide Urban Runoff Program (NURP) study (34) in the Washington, DC, area and the national NURP data analysis. The equations, however, may be applied anywhere in the country. Some precision is lost as a result of the effort to make the equation general and simple. The method is adequate for decision-making at the site-planning level.

FHWA Model. The Federal Highway Administration (FHWA) has developed a simple statistical spreadsheet procedure to estimate pollutant loading and impacts to stream and lakes that receive highway stormwater runoff (35). The FHWA model uses a set of default values for pollutant event mean concentrations that depends on traffic volume and whether the setting of the highway’s pathway is rural or urban. The FHWA uses this method to identify and quantify the constituents of highway runoff and their potential effects on receiving waters and to identify areas that might require controls. The FHWA model is well suited for screening application and can evaluate lake and stream impacts of highway stormwater discharges. However, it assesses seasonal variability in a limited manner and does not consider the soluble fraction of pollutants or precipitation and settling of pollutants in lakes.

PLOAD Model. PLOAD is a simplified GIS-based watershed loading model. It can model combined point and nonpoint source loads in either small urban areas or in rural watersheds of any size. The user may calculate the nonpoint source loads using either the export coefficient or the EPA’s simple method approach. Best management practices (BMPs), which reduce both point and nonpoint source loads, may also be included in computing total watershed loads. PLOAD broadly addresses pollutant loading by land use categories and subwatersheds but does not get at individual nonpoint sources or at actual pollutant fate and transport processes.

PLOAD was designed to be generic, so that it can be applied as a screening tool in a wide range of applications, including NPDES stormwater permitting, watershed management, and reservoir protection projects. The uncertainty of export coefficients is high, but the PLOAD model provides a good fit to the regulatory concept of the phased TMDL. Uncertainty can be reduced by calibrating to local conditions based on multiyear datasets and can be addressed in later phases of the TMDL as more detailed source identification and seasonal data become available (36).

SUMMARY AND CONCLUSION

Water quality models of different complexity level (ranging from simple, midrange, to detailed) for developing soil management practices have been reviewed. In general, complex models potentially have greater accuracy in their predictions than simple models because of the comprehensive simulation of environmental processes. However, more site-specific data are required to use these complex models fully. Application of complex models is usually more expensive than using simple models for developing management practice. Therefore, the trade-off among these models should be one of the key considerations during model selection.

Although guidance is still needed for determining nonpoint source loadings from urban and rural areas (37), the common approaches used currently for estimating wet weather nonpoint pollution loads are (38) (1) unit load concepts (long-term statistical average loads are related to land uses), (2) event-oriented models (water and pollutant loads are calculated for a single storm), and (3) continuous models (water and pollutant loads are calculated using daily or shorter time intervals and complex, mostly deterministic hydrologic models).

The realization that management of nonpoint sources of pollutants is as important an issue as management of the more readily controllable point sources has led to the concept of watershed management and planning. Increased computational power, as well as a desire for watershed planning, has resulted in the advancement of mathematical modeling tools. One of the trends is developing integrated modeling systems that link the models, data, and user interface within a single system. New developments in modeling systems have increasingly relied on a geographic information system (GIS) and a database management system to support modeling and analysis (1). U.S. EPA's BASINS offers a comprehensive modeling system that brings key data and analytical components together "under one roof" (2). USDA's benchmark watershed model SWAT has been incorporated into the BASINS V3 system. The Watershed Modeling System (WMS) is another useful tool for watershed modeling and management practice development. These multipurpose environmental analysis systems are expected to be extensively used by federal, regional, state, and local agencies in watershed and water quality studies in the future.

BIBLIOGRAPHY

1. U.S. EPA. (1997). *Compendium of Tools for Watershed Assessment and TMDL Development*. EPA 841-B-97-006, Washington, DC.
2. U.S. EPA. (2000). *National Management Measures to Control Nonpoint Source Pollution from Agriculture (Draft)*. Washington, DC.
3. Williams, J.R., Nicks, A.D., and Arnold, J.G. (1985). Simulator for water resources in rural basins. *J. Hydraul. Eng.* **111**(6): 970–986.
4. Arnold, J.G., Williams, J.R., Nicks, A.D., and Sammons, N.B. (1990). *SWRRB: A Basin Scale Simulation Model for Soil and Water Resources Management*. Texas A&M University Press, College Station, TX.
5. Arnold, J.G. and Williams, J.R. (1987). Validation of SWRRB—simulator for water resources in rural basins. *J. Water Resour. Plann. Manage.* **113**: 243–256.
6. Bingner, R.L., Murphree, C.E., and Mutchler, C.K. (1987). *Comparison of Sediment Yield Models on Various Watersheds in Mississippi*. American Society of Agricultural Engineers Paper 87–2008, St. Joseph, MI.
7. Knisel, W.G. (1980). *CREAMS, a Field Scale Model for Chemicals, Runoff and Erosion from Agricultural Management Systems*. USDA Conservation Research Report. No. 26.
8. Leonard, R.A., Knisel, W.G., and Still, D.A. (1987). GLEAMS: groundwater loading effect on agricultural management systems. *Trans. ASAE* **30**(5): 1403–1428.
9. Williams, J.R., Jones, C.A., and Dyke, P.T. (1984). A modeling approach to determining the relationship between erosion and soil productivity. *Trans. ASAE* **27**(1): 129–144.
10. Knisel, W.G., Leonard, R.A., and Davis, F.M. (1991). Water balance components in the georgia coastal plain: a GLEAMS model validation and simulation. *J. Soil Water Conserv.* **46**(6): 450–460.
11. Zacharias, S., Heatwole, C.D. Dillaha, T., and Mostighimi, S. (1992). *Evaluation of GLEAMS and PRZM for Predicting Pesticide Leaching Under Field Conditions*. American Society of Agricultural Engineers Paper 92–2541, St. Joseph, MI.
12. Knisel, W.G. and Leonard, R.A. (1989). Irrigation impact on groundwater: model study in humid region. *J. Irrig. Drainage Eng.* **15**(5): 823–837.
13. Smith, M.C., Bottcher, A.B., Campbell, K.L., and Thomas, D.L. (1991). Field testing and comparison of the PRZM and GLEAMS models. *Trans. Am. Soc. Agric. Eng.* **34**(3): 838–847.
14. Sharply, A.N. and Williams, J.R. (1990). *EPIC—Erosion/ Productivity Impact Calculator: 1.* Model Development. USDA Bulletin No. 1768, p. 235.
15. Sugiharto, T., McIntosh, T.H., Uhrig, R.C., and Lavdinois, (1994). Modeling alternatives to reduce dairy farm and watershed nonpoint source pollution. *J. Environ. Qual.* **23**: 18–24.
16. Edwards, D.R. and Daniel, T.C. (1994). Quality of runoff from fescue grass plots treated with poultry litter and inorganic fertilizer. *J. Environ. Qual*. **23**: 579–584.
17. Bingner, R.L. and Theurer. F.D. (2001). AnnAGNPS: estimating sediment yield by particle size for sheet & rill erosion. In: *Proceedings of the Sediment: Monitoring, Modeling, and Managing, 7th Federal Interagency Sedimentation Conference*. Reno NV, 25-29 March 2001. pp. I-1–I-7.
18. Line, D.E., Coffey, S.W., and Osmond, D.L. (1997). WATERSHEDSS—GRASS AGNPS modeling tool. *Trans. Am. Soc. Agric. Eng.* **40**(4): 971–975.
19. Young, R.A., Onstad, C.A., Bosch, D.D., and Anderson, W.P. (1994). *Agricultural Nonpoint Source Pollution Model, Version 4.03 AGNPS User's Guide*. USDA Agricultural Research Service, Morris, MN.
20. USDA. (2001). Revised Universal Soil Loss Equation, http://www.sedlab.olemiss.edu/rusle.
21. Flanagan, D.C. and Nearing, M.A. (1995). *USDA-Water Erosion Prediction Project: Hillslope Profile and Watershed Model Documentation*. NSERL Report No.10, West Lafayette, IN.
22. Tiscareno-Lopez, M., Lopes, V.L., Stone, J.J., and Lane, L.J. (1993). Sensitivity analysis of the WEPP watershed model for rangeland applications: I. Hillslope processes. *Trans. ASAE* **36**: 1659–1672.
23. Liu, B.Y., Nearing, M.A., Baffault, C., and Ascough, J.C. (1997). The WEPP watershed model: three comparisons to measured data from small watersheds. *Trans. Am. Soc. Agric. Eng.* **40**: 945–952.
24. Follet, R.F., Kenney, D.R., and Cruse, R.M. (1991). *NLEAP (Nitrogen Leaching and Economic Analysis Package)*. Soil Science of America, Inc., Madison, WI.
25. Wylie, B.K. et al. (1994). Predicting spatial distributions of nitrate leaching in northeastern colorado. *J. Soil Water Conserv.* **49**: 288–293.
26. Wylie, B.K., Shaffer, M.J., and Hall, M.D. (1995). Regional assessment of NLEAP NO_3-N leaching indices. *Water Resour. Bull.* **31**: 399–408.
27. Culver, T.B. et al. (2002). Case study of impact of total maximum daily load allocations on nitrate leaching. *ASCE J. Water Resour. Plann. Manage.* **128**(4): 262–270.
28. Mullens, J.A., Carsel, R.F., Scarbrough, J.E., and Ivery, I.A. (1993). *PRZM-2, A Model for Predicting Pesticide Fate in the Crop Root and Unsaturated Soil Zones: User Manual for Release 2.0*. Office of Research and Development, U.S. EPA, Athens, GA, EPA Report 600/R-93/046.
29. Haith, D.A. and Shoemaker, L.L. (1987). Generalized watershed loading functions for stream flow nutrients. *Water Resour. Bull.* **23**(3): 471–478.

30. Beasley, D.B., Huggins, L.F., and Monke, E.J. (1980). ANSWERS: a model for watershed planning. *Trans. Am. Soc. Agric. Eng.* **23**: 938–944.
31. Griffin, M.L., Beasley, D.B., Fletcher, J.I., and Foster, G.R. (1988). Estimating soil loss on topographically non-uniform field and farm units. *J. Soil Water Conserv.* **43**: 326–331.
32. Tasker, G.D. and Driver, N.E. (1988). Nationwide regression models for predicting urban runoff water quality at unmonitored sites. *Water Resour. Bull.* **24**(5): 1091–1101.
33. Driver, N.E. and Troutman, B.M. (1989). Regression models for estimating urban storm-runoff quality and quantity in the United States. *J. Hydrology* **109**: 221–236.
34. U.S. EPA. (1983). *Results of the Nationwide Urban Runoff Program*. Vol. Final Report, Washington, DC.
35. Federal Highway Administration (FHWA). (1996). *Evaluation and Management of Highway Runoff Water Quality*. Publication No. FHWA-PD-96-032.
36. Cocca, P.A. (2001). *BASINS 3.0: New Tools for Improved Watershed Management*. Proc. ASCE World Water Congr. 2001, Orlando, FL.
37. U.S. EPA. (1998). *Report of the Federal Advisory Committee on the Total Maximum Daily Load (TMDL) Program*. EPA-100-R-98-006, Washington, DC.
38. Novotny, V. (1999). Integrating diffuse/nonpoint pollution control and water body restoration into watershed management. *J. Am. Water Resour. Assoc.* **35**(4): 717–727.

READING LIST

American Society of Civil Engineers (ASCE). (2001). National Stormwater Best Management Practices (BMPs) Database, http://www.bmpdatabase.org/.

Arnold, J.G., Williams, J.R. and Maidment, D.R. (1995). Continuous-time water and sediment-routing model for large basins. *J. Hydraul. Eng.* **121**(2): 171–183.

Bicknell, B.R., Imhoff, J.C., Kittle, J.L Jr., and Donigian, A.S. Jr. (1996). *Hydrological Simulation Program—FORTRAN, User's Manual for Release 11*. Environmental Research Laboratory, U.S. EPA, Athens, GA.

Carrubba, L. (2000). Hydrologic modeling at the watershed scale using NPSM. *J. Am. Water Resour. Assoc.* **36**(6): 1237–1246.

Cooper, A.B., Smith, C.M., and Bottcher, A.B. (1992). Predicting runoff of water, sediment, and nutrients from a New Zealand grazed pasture using CREAMS. *Trans. Am. Soc. Agric. Eng.* **35**(1): 105–112.

Deliman, P.N., Glick, R.H., and Ruiz, C.E. (1999). *Review of Watershed Water Quality Models*. U.S. Army Corps of Engineers, Technical Report W-99-1.

Donigian, A.S. and Huber, W.C. (1991). *Modeling of Nonpoint Source Water Quality in Urban and Non-urban Areas*. EPA 600/3-91/039, U.S. EPA, Athens, GA.

Evans, B.M., White, R.A., Petersen, G.W., Hamlett, J.M., Baumer, G.M. and McDonnell, A.J. (1994). *Land Use and Non-point Pollution Study of the Delaware River Basin*. Environmental Resources Research Institute, Penn State University, Publication No. ER94-06, p. 76.

Haith, D.R., Mandel, R., and Wu, R.S. (1992). *GWLF: Generalized Watershed Loading Functions User's Manual, Version 2.0*. Cornell University, Ithaca, NY.

Novotny, V. and Olem, H. (1994). *Water Quality: Prevention, Identification, and Management of Diffuse Pollution*. Van Nostrand Reinhold, New York.

Water Environment Research Foundation (WERF). (2001). *Water Quality Models: A Survey and Assessment*, Final Report on Project 99-WSM-5.

Whittemore, R.C. and Beebe, J. (2000). EPA's BASINS model: good science or serendipitous modeling. *J. Am. Water Resour. Assoc.* **36**(3): 493–499.

Yoon, K.S., Yoo, K.H., Wood, C.W., and Hall, B.M. (1994). Application of GLEAMS to predict nutrient losses from land application of poultry litter. *Trans. Am. Soc. Agric. Eng.* **37**(2): 453–460.

Young, R.A., Onstad, C.A., Bosch, D.D., and Anderson, W.P. (1989). AGNPS: a nonpoint source pollution model for evaluating agriculture watersheds. *J. Soil Water Conserv.* **44**: 168–173.

WATER QUALITY MODELING—CASE STUDIES

Zhen-Gang Ji
Minerals Management Service
Herndon, Virginia

This entry focuses on case studies of water quality modeling. Information on water quality models is provided in Water Quality Models.

MODEL SELECTION

Although a variety of models are available for studying water quality processes, it is still a complex task to select a model (or models) that best meets all (or most) of the study objectives. Model selection should be balanced among competing demands. Time and resources are always limited, to one degree or another, so the goal should be to identify the model(s) that addresses all of the important processes that affect the waterbody. Selecting a model that is too simple can result in lack of accuracy and certainty that are needed in decision making; selecting an overly complex model may result in misdirected resources, study delays, and increased cost.

It is helpful to ask the following questions during model selection:

1. What are the key hydrodynamic processes?
2. What are the water quality concerns?
3. What spatial and temporal scales are adequate for resolving these processes?
4. How will the model be used in supporting management decision making?

It might be desirable to select the simplest model that can meet the study needs. However, in real practice, comprehensive models are often preferred to simple models for a variety of reasons. Typical features of comprehensive models include:

1. three-dimensional (3-D) in space and time dependent
2. turbulent scheme for vertical mixing
3. hydrodynamic, thermal, sediment, and eutrophication processes.

Generally speaking, comprehensive models should (but not always) have better mathematical, physical, chemical, and biological representations of natural waters than simple models. It is often advantageous to adopt a more detailed model to address various scientific and engineering applications than to switch models continuously from one phase of a project to another or from one project to another (1). As discussed later and in the entry WATER QUALITY MODELS: CHEMICAL PRINCIPLES, the Environmental Fluid Dynamics Code (EFDC) model is a comprehensive model (2).

In many applications, comprehensive models have advantages compared to simple models for the following:

1. *The Evolving Understanding of the Waterbody Studied.* A waterbody is understood gradually as the modeling effort progresses. Detailed understanding of the system is often achieved after the modeling study is finished. At the beginning of a study, it is sometimes difficult to know which simple model should be able to describe the system adequately. Therefore, it is helpful to have a model that has the most capability (and flexibility for further model enhancement) to describe important processes (known and to be known) in the system.
2. *The Management Needs.* A primary goal of a modeling study is often to support management decision making, which might change with time and cannot always be foreseeable. It is desirable to choose a model that can meet these "expanding" needs. It is generally more cost-effective to choose a comprehensive model that can address current and future needs than to use a simple model that is later found inadequate and has to be replaced by one more advanced.
3. *The Falling Costs of Computing.* The progress in computer technology has dramatically reduced the costs of computing and makes comprehensive models more affordable.
4. *The Modelers.* A modeler often needs to study hydrodynamic, sediment, toxic, and/or water quality processes in a variety of waterbodies and, therefore, needs models and tools that are versatile enough for these applications. It is common for a modeler to have several modeling studies during any given period and to complete tens of modeling projects over the years. Rather than learning and using a variety of models for different applications, it is advantageous to stick with one (or a few) comprehensive model, know it well, and apply it to modeling rivers, lakes, and estuaries.

Comprehensive models can be applied appropriately only by well-trained and experienced modeling professionals and often need extensive measured data for model calibration and verification. The progress in computer technology, the enrichment of measured data, and the enhancement of water quality models are making comprehensive models powerful tools for water quality management.

Two cases studies are introduced here: the first is on sediment and toxic metals in a shallow river, and the second is on eutrophication in a deep reservoir.

CASE STUDY I: BLACKSTONE RIVER

Many waterbodies are polluted by sediments and toxic metals. This case study illustrates the modeling of hydrodynamics, sediment transport, and toxic metals in the Blackstone River, Massachusetts (3).

Background

The Blackstone River Basin (Fig. 1) consists of approximately 1657 square kilometers encompassing 30 cities and towns. The Blackstone River flows from Worcester, Massachusetts, to Pawtucket, Rhode Island. The river is 77 km long and has 133 meters of total fall. The distances shown in Fig. 1 are river kilometers from Slaters Mill Dam. There are presently 14 dams and impoundments on the mainstem of the Blackstone that are significant to the hydrodynamic and water quality processes in the river. The Blackstone has been the largest source of pollutants discharging into Narragansett Bay, principally from industrial discharge of metals and resuspension of contaminated sediments behind the low head dams in the river. Sediment and metals in certain sections of the riverbed can be traced back 200 years to the American Industrial Revolution.

This study is intended to provide the United Sates Environmental Protection Agency (EPA) with a relatively simplified example of how comprehensive models can be used to evaluate the distributions of heavy metals in a shallow, narrow, urban, industrial river.

Hydrodynamic, Sediment, and Metals Simulations

The Blackstone River model is developed within the framework of the EFDC model (2). The EFDC model is a public-domain modeling package used for simulating three-dimensional flow, transport, and biogeochemical processes in surface water systems, including rivers, lakes, estuaries, reservoirs, wetlands, and coastal regions. The EFDC model has been extensively tested and documented in more than 80 modeling studies and is currently supported by the EPA. The model is used by a number of organizations, including universities, governmental agencies, and environmental consulting firms. Representative applications of the EFDC model include modeling of sediment and metals transport in the Blackstone River (3), wetting and drying simulation of Morro Bay (4,5), and simulating hydrodynamic, thermal, and sediment processes in Lake Okeechobee (6–8). More information on the EFDC model is also provided in the entry WATER QUALITY MODELS: CHEMICAL PRINCIPLES.

The Blackstone River model includes a hydrodynamic submodel, a sediment submodel, and a toxicant submodel. They are coupled together and are executed simultaneously. The hydrodynamic model simulates velocity, water elevation, and turbulent mixing for sediment modeling. The output of the hydrodynamic and sediment models are linked to the toxicant model to simulate the five metals in the river: cadmium (Cd), chromium (Cr), copper (Cu), nickel (Ni), and lead (Pb).

The typical width of the Blackstone River is around 25 meters and varies from less than 10 m upstream to

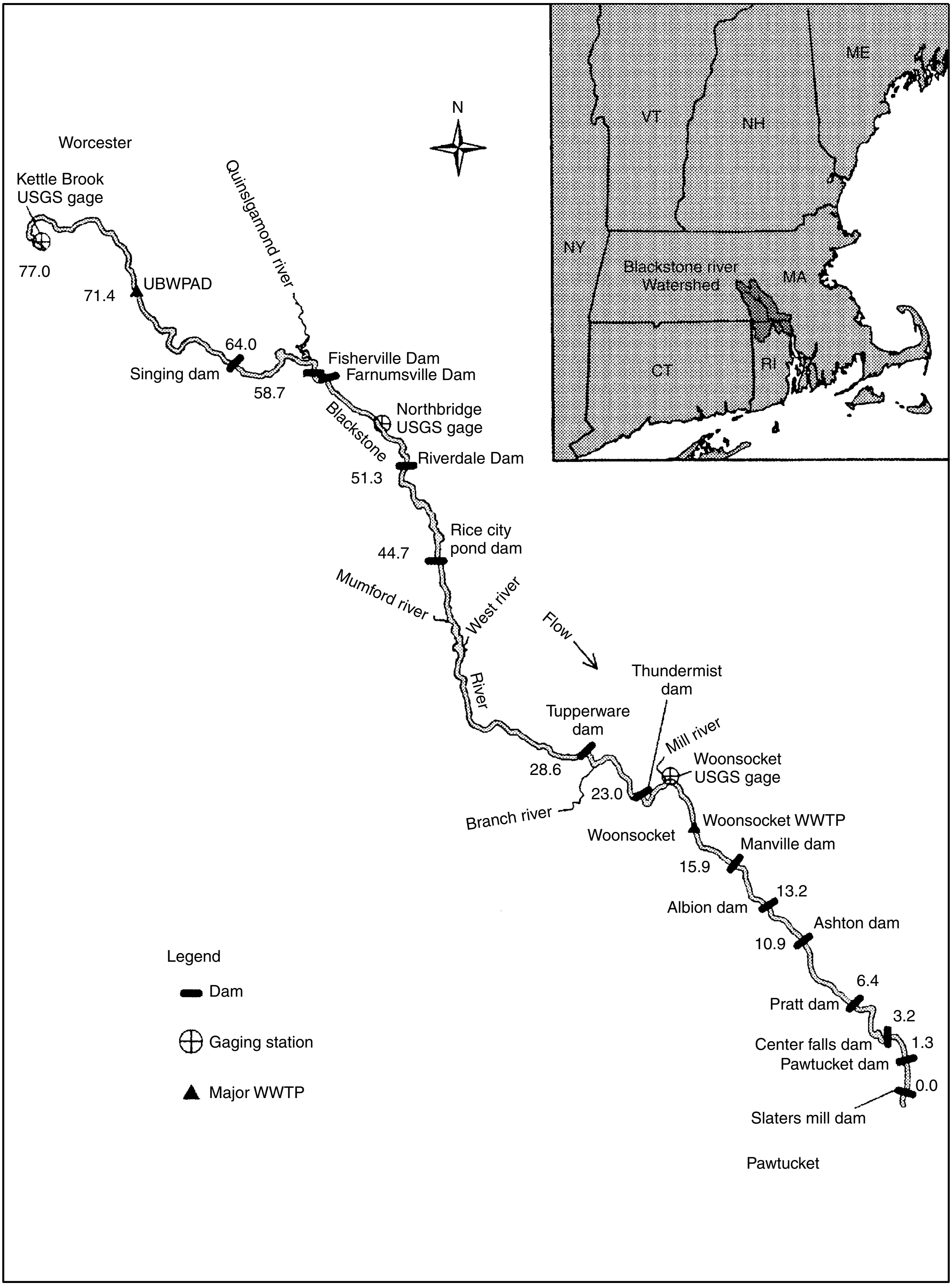

Figure 1. Blackstone River study area. The distances are river kilometers from Slaters Mill Dam.

more than 35 m downstream. It has an average drop of 1.73 m/km, so the Blackstone is a gravity-driven river. The grid of the Blackstone River model has one cell across the river and one layer in the vertical. Along the river, there are 256 grid cells of varying cell widths and a uniform cell length of 300 m. A time step of 30 seconds is used throughout the simulation.

Figure 2 presents the time series of the modeled and measured concentrations of total suspended solids (TSS) along the Blackstone River during a storm (November 2–6, 1992). The horizontal axis is in days from November 2, 1992, and the vertical axis is in milligram/liter. The closed circles represent the measured TSS, and the solid line represents the model results. The river kilometers are also shown in the plots. The model realistically simulates the sediment resuspension processes at dams, including Singing Dam (km = 64.0), Fisherville Dam (km = 58.4), Riverdale Dam (km = 51.3), and Rice City Pond Dam (km = 44.7). The high sediment concentration at km = 37.3 is caused by flows from upstream and from tributaries. Figure 3 is similar to Fig. 2, except for Pb concentrations and shows that the model simulates Pb processes well.

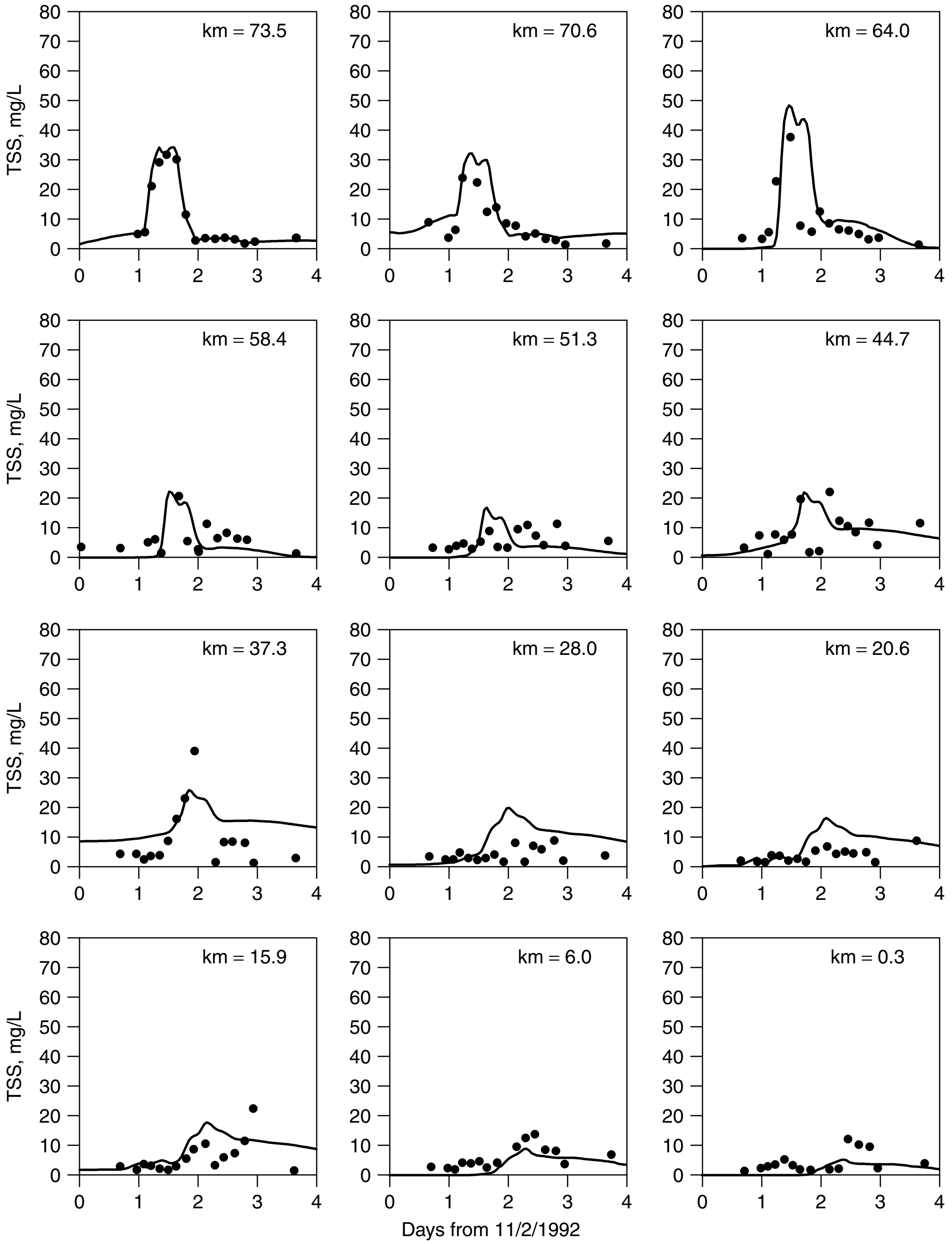

Figure 2. Measured and modeled total suspended solids (TSS) concentration along the Blackstone River from November 2–6, 1992.

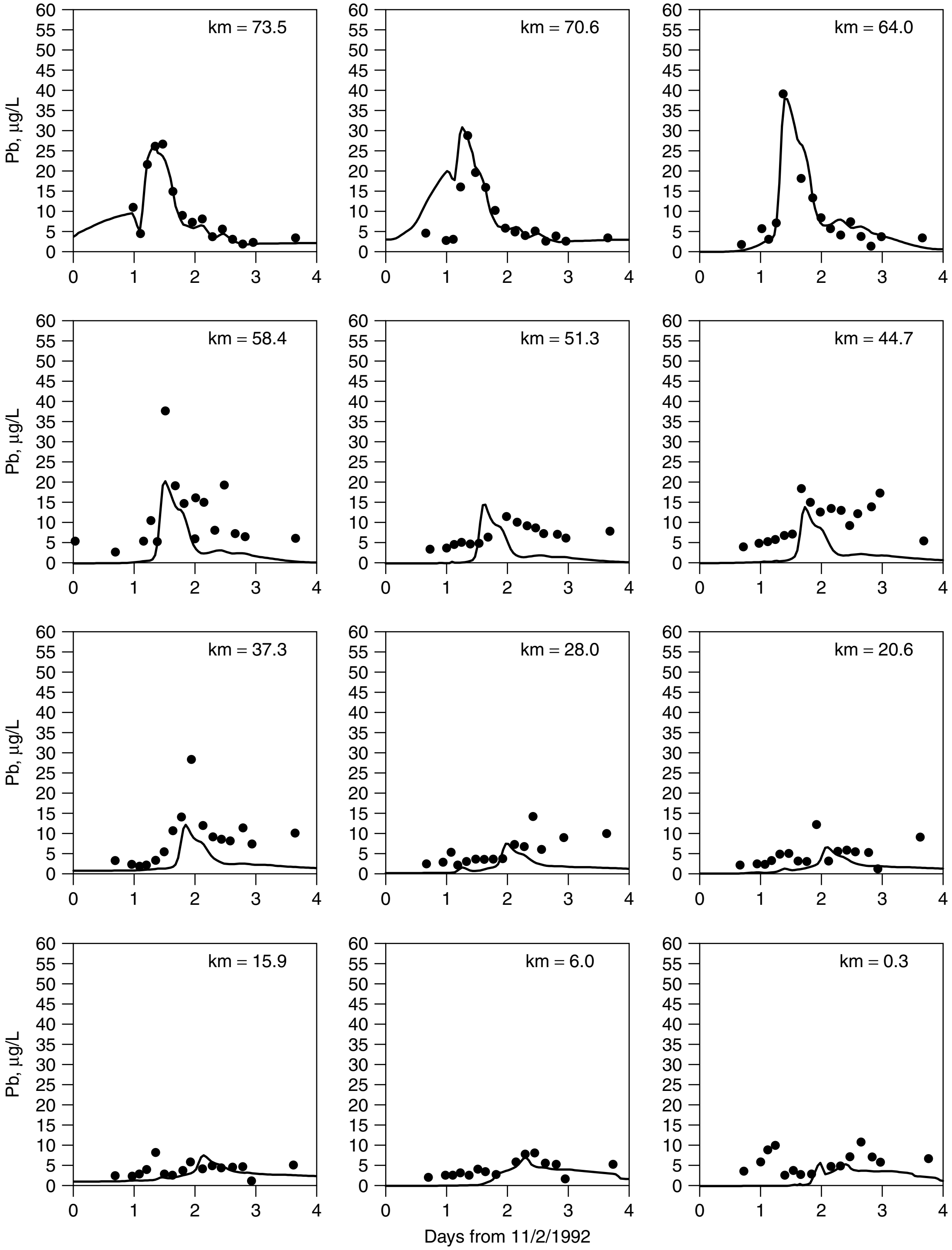

Figure 3. Measured and modeled lead (Pb) concentration along the Blackstone River from November 2–6, 1992.

To conduct a statistical analysis, the model results are also saved at the same locations and times at which the data were collected. Table 1 summarizes the error analyses of seven variables: river flow rate (Q), TSS, Cd, Cr, Cu, Ni, and Pb. The table presents the values of observed mean, modeled mean, absolute mean error, root mean square (rms) error, observed change, and the percentage of relative rms error. The relative rms error is defined as the rms error divided by the observed change. These analyses show that the simulations agree well with the observed data. The relative rms errors range from 6.17% (Q) to 17.27% (TSS), and the relative rms errors for the five metals are no more than 15.03%.

Conclusions

It is a challenge to apply coupled hydrodynamic, sediment, and contaminant fate and transport models to studies of surface water systems. Contaminant transport depends

Table 1. Error Analysis of Observed and Modeled River Flow Rate (*Q*), Total Suspended Sediment (TSS), Cadmium (Cd), Chromium (Cr), Copper (Cu), Nickel (Ni), and Lead (Pb)

Variable	Obs. Mean	Modeled Mean	Mean abs. err.	rms err.	Obs. Change	Relative rms err. (%)
Q (m^3/s)	9.53	10.95	1.86	3.42	55.39	6.17
TSS (mg/L)	6.80	8.20	4.54	6.53	37.80	17.27
Cd (μg/L)	0.88	1.22	0.52	0.71	5.12	13.95
Cr (μg/L)	4.08	1.99	2.27	3.17	27.08	11.71
Cu (μg/L)	12.27	12.09	5.77	8.69	57.80	15.03
Ni (μg/L)	7.92	7.73	2.31	3.52	25.70	13.68
Pb (μg/L)	7.52	4.57	4.00	5.43	37.80	14.38

on hydrodynamic and sediment conditions. Heavy metals and toxic chemicals can preferentially adsorb and desorb with solids in the water column and sediment bed. High flow events, such as storms, increase solid loadings from the watershed, increase river flow velocity, reintroduce previously deposited chemicals back into the water column via resuspension, and transport the resuspended contaminants further downstream until they settle out in deposition zones.

So far, there are few published modeling studies on sediment and metal transport in rivers that simulated storms on an hourly basis and used comprehensive data sets for model input and model calibration. This study can simulate the sediment and metal transport processes in detail. Statistical analysis and graphic presentation indicate that the model results are in very good agreement with the data. For shallow and narrow rivers like the Blackstone, the EFDC model with a 1-D grid can represent the hydrodynamic, sediment, and metals processes reasonably well.

CASE STUDY II: LAKE TENKILLER

Water quality management requires scientifically credible numerical models for evaluating pollution control options. This case study summarizes the eutrophication and water quality modeling of Lake Tenkiller, Oklahoma (9).

Lake Tenkiller (Fig. 4) is a reservoir located in northeastern Oklahoma. The lake is 48 km long, up to 3 km wide, and 70 km^2 in area. Its depth varies from more

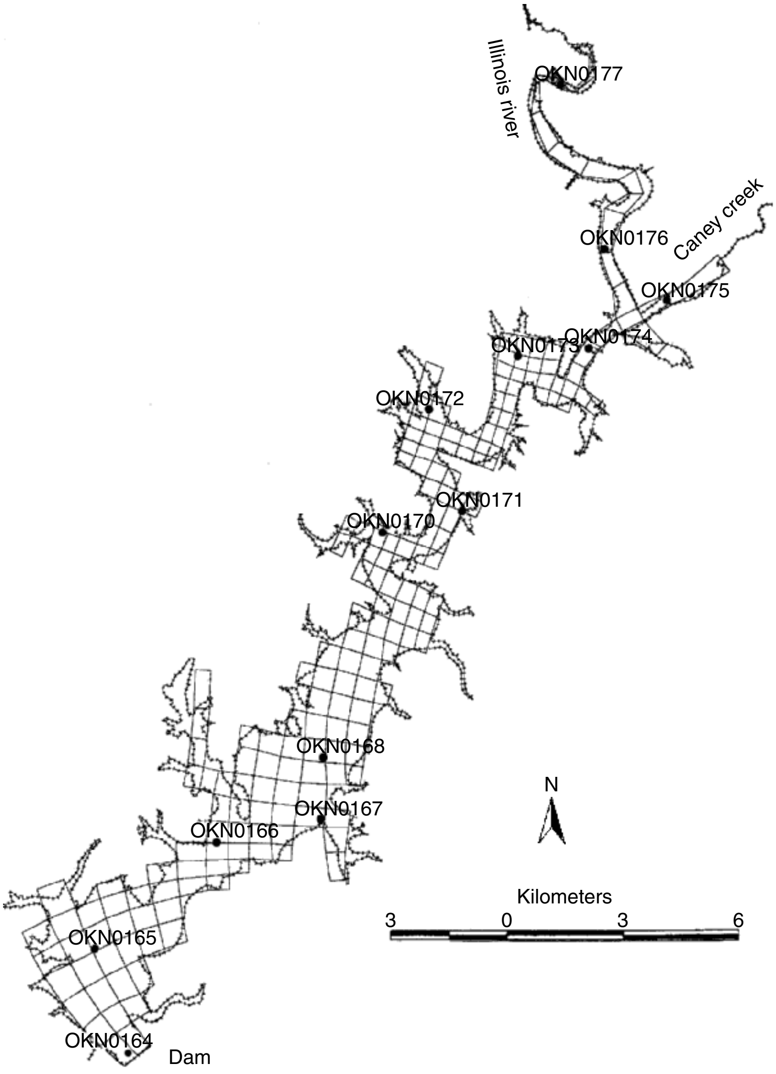

Figure 4. Lake Tenkiller study area, model grid, and data stations.

than 45 m near the dam to less than 6 m in the northern section. The lake has a retention time of 1.76 years. Major water quality issues include nutrient enrichment, eutrophication, and hypolimnetic dissolved oxygen (DO) depletion. Because of the 3-D variability of the lake, it is critical to simulate the hydrodynamic and water quality processes using a 3-D model, so that water quality parameters in the lake can be described in detail and cost-effective water management approaches can be proposed and evaluated.

Water Quality Simulations

The Lake Tenkiller model is also developed within the framework of the EFDC model (2) and includes the hydrodynamic, sediment, and water quality submodels. The water quality parameters considered, as presented in the entry WATER QUALITY MODELS: CHEMICAL PRINCIPLES, include algae, DO, organic nitrogen, ammonia, nitrite and nitrate, organic phosphorus, and phosphates.

The study area is divided into a grid of discrete cells. Figure 4 shows the model grid overlaying an outline of

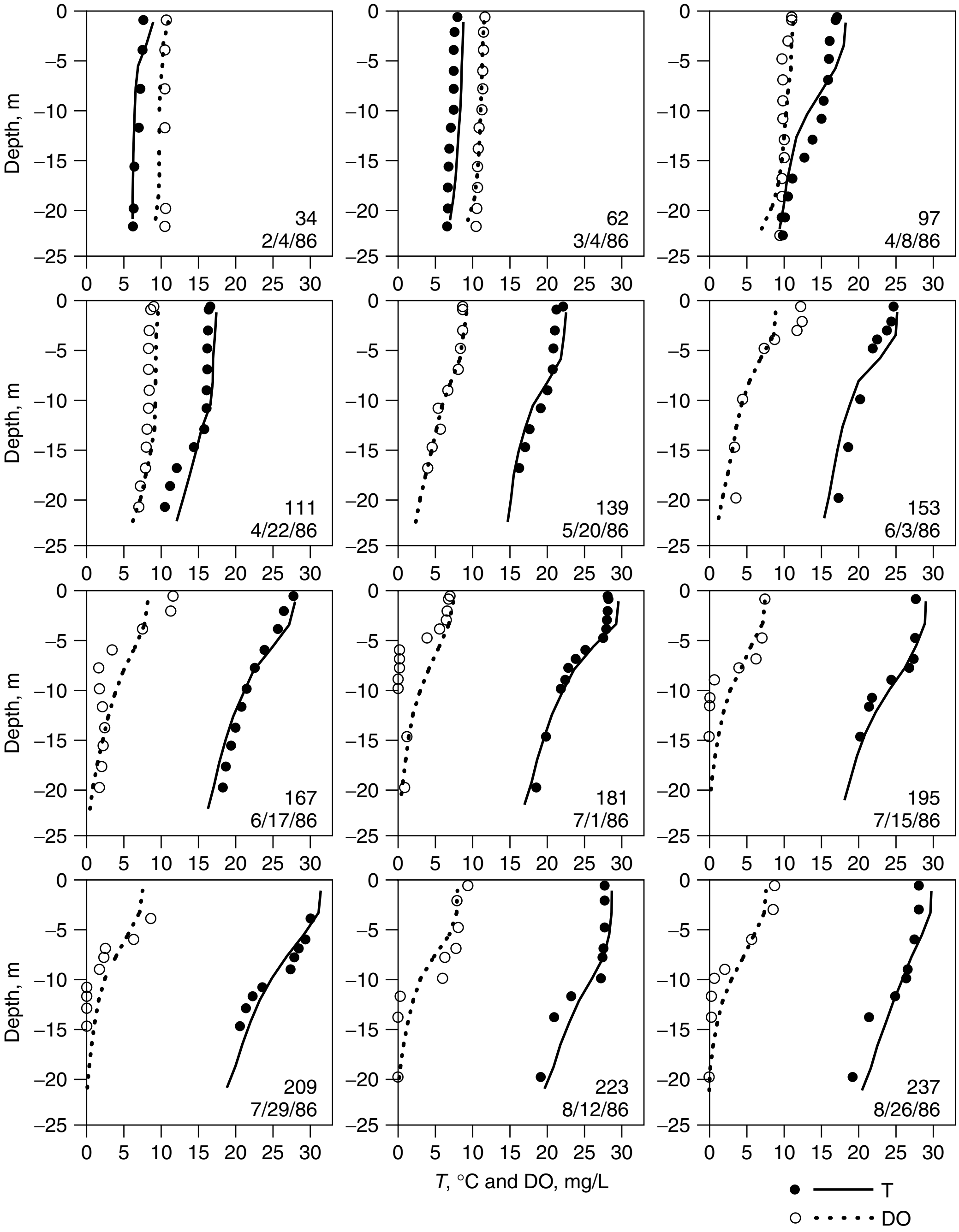

Figure 5. Vertical profiles of temperature and dissolved oxygen (DO) at Station OKN0166. Solid line = modeled temperature, dashed line = modeled DO, closed circles = measured temperature, and open circles = measured DO.

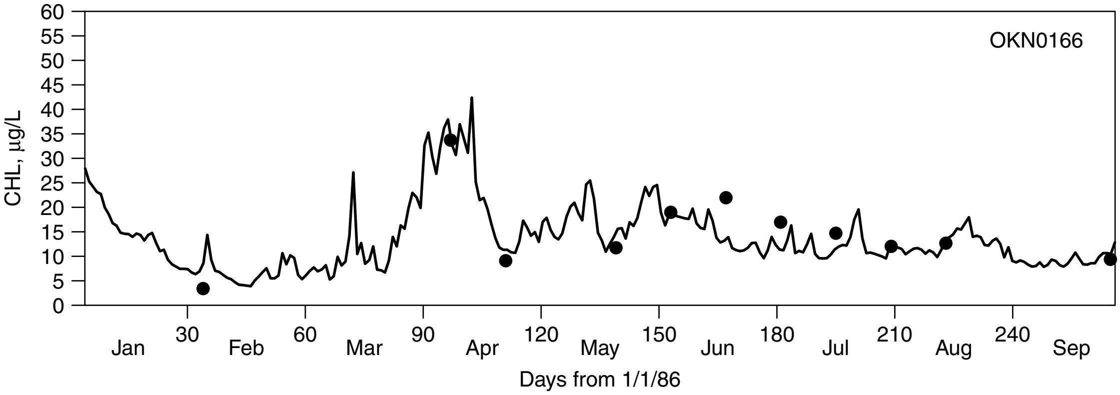

Figure 6. Model-data comparison of chlorophyll-*a* (CHL) at Station OKN0166. Solid line = modeled surface CHL, and closed circles = measured surface CHL.

Lake Tenkiller. The numerical grid consisted of 198 cells in the horizontal plane and 10 vertical sigma layers. The 10 vertical layers are necessary and important for resolving the vertical temperature and DO profiles in the lake. The time period for model calibration is 262 days from January 5, 1986 to September 24, 1986. Solutions of the model are obtained using a 50-second time step. Using a 2.4-Ghz Pentium IV PC, approximately 6 CPU hours are required for a 262-day simulation.

The 12 plots in Fig. 5 show the vertical profiles of modeled and measured water temperature (T) and DO at Station OKN0166. The horizontal axis represents T in °C and DO in mg/L. The vertical axis represents water depth in meters. The solid curve is the modeled T, and the closed circles are the measured T. The dashed curve is the modeled DO, and the open circles are the measured DO. Overall, the modeled T and DO are consistent with the measured data. Both the model and the data indicate that DO is vertically mixed at values around 10 mg/L in the winter and is very stratified at 8 mg/L or more on the surface and almost 0 mg/L at the bottom in the summer. The lake started to become stratified in April and exhibited strong thermal stratification in the summer; the surface-bottom T difference was up to 15 °C, and the DO difference up to 10 mg/L. It also appears that the model might need more vertical layers to resolve DO stratifications better in the summer.

Algae levels are often measured by chlorophyll-*a* concentrations (CHL) in water quality models. Figure 6 gives the modeled CHL (solid line) and the measured CHL (closed circle) at OKN0166, the same station whose T and DO are displayed in Fig. 5. Figure 6 reveals that algae in the lake vary from less than 10 μg/L in the winter to around 25 μg/L in the summer, both the model and the data indicate a strong algae bloom in the spring.

Conclusions

The dynamic simulation of eutrophication in a deep water system is a very complicated and computationally intensive endeavor because a large number of chemical, biological, and biochemical processes interact and the reaction rates and external inputs vary with time. Despite the progress in 3-D hydrodynamic, water quality, and sediment diagenesis models and their successful applications in estuaries and bays (10), few similar 3-D modeling studies on lakes and reservoirs have been published.

The primary purpose for developing the Lake Tenkiller model is to use the model as a tool for proposing and testing load-management strategies aimed at limiting eutrophication processes in the lake. The Lake Tenkiller model is calibrated and represents existing hydrodynamic and water quality processes in the lake satisfactorily and is a useful tool for managing eutrophication in the lake (11).

BIBLIOGRAPHY

1. Nix, J.S. (1990). Mathematical modeling of the combined sewer system. In: *Control and Treatment of Combined Sewer Overflows*. P.E. Moffa (Ed.). Van Nostrand Reinhold, New York, pp. 23–78.
2. Hamrick, J.M. (1992). *A Three-Dimensional Environmental Fluid Dynamics Computer Code: Theoretical and Computational Aspects*. The College of William and Mary, Virginia Institute of Marine Science, Special Report 317.
3. Ji, Z.-G., Hamrick, J.H., and Pagenkopf, J. (2002). Sediment and metals modeling in shallow river. *J. Environ. Eng.* **128**: 105–119.
4. Ji, Z.-G., Morton, M.R., and Hamrick, J.M. (2000). Modeling hydrodynamic and sediment processes in Morro Bay. In: *Estuarine and Coastal Modeling: Proc. 6th Int. Conf.* M.L. Spaulding and H.L. Butler (Eds.). ASCE, pp. 1035–1054.
5. Ji, Z.-G., Morton, M.R., and Hamrick, J.M. (2001). Wetting and drying simulation of estuarine processes. *Estuarine Coastal Shelf Sci.* **53**: 683–700.
6. Jin, K.R. and Ji, Z.G. (2003). Modeling of sediment transport processes and wind-wave impact in a large shallow lake. *J. Hydraul. Eng.* ASCE (tentatively accepted).
7. Jin, K.R. and Ji, Z.G. (2003). Application and validation of a 3-D model in a shallow lake. *J. Waterway Port Coastal, Ocean Eng.* (accepted).
8. Jin, K.R., Ji, Z.G., and Hamrick, J.M. (2002). Modeling winter circulation in Lake Okeechobee, Florida. *J. Waterway Port Coastal Ocean Eng.* **128**: 114–125.
9. Ji, Z.-G., Morton, M.R., and Hamrick, J.M. (2003). Three-dimensional hydrodynamic and water quality modeling in a reservoir. In: *Estuarine and Coastal Modeling: Proc. 8th*

Int. Conf. M.L. Spaulding and H.L. Butler (Eds.). ASCE (accepted).

10. Cerco, C.F. (1999). Eutrophication models of the future. *J. Environ. Eng.* **125**(3): 209–210.
11. Tetra Tech. (2000). *Water Quality Modeling Analysis in Support of TMDL Development for Tenkiller Ferry Lake and the Illinois River Watershed in Oklahoma*. Technical report to United States Environmental Protection Agency, Region 6, and Department of Environmental Quality, State of Oklahoma. Tetra Tech, Inc., Fairfax, VA.

FIELD SAMPLING AND MONITORING OF CONTAMINANTS

Timothy J. Downs
Clark University
Worcester, Massachusetts

INTRODUCTION

Field sampling and monitoring of water/wastewater contaminants are strategically important for more sustainable integrated water resources management (IWRM). One of the vital first steps in IWRM is the collection of an inventory that characterizes existing watershed conditions. Relevant features include geographic and environmental aspects (e.g., land use, topography, wetlands), infrastructure (e.g., supply, sewerage and drainage), municipal data (e.g., numbers of people, growth rate, regulations, by laws), pollution sources/discharges (point and nonpoint sources, e.g., landfills, underground storage tanks), and the characteristics of receiving surface waters and groundwaters (indicators of flow rates and physical, chemical, biological, and microbiological quality).

Water quality problems that require characterization through sampling and monitoring are a combination of the effects of human use and degradation of water in domestic, industrial, commercial, and agricultural activities (and its discharge as wastewater), and the impact of human activities that change land use/land cover on the water quality of runoff. Wastewater discharge can be handled by treatment technologies, but during rainstorms, combined sewer overflows (CSOs) and sanitary sewer overflows (SSOs) can degrade water quality. In urban areas, domestic wastewaters contain sanitary sewage composed of oxygen-demanding wastes, nutrients, and pathogens. Industrial wastewaters contain metals and organic pollutants. Urban runoff from paved surfaces contains hydrocarbons, runoff from eroded land surfaces produces sediment-laden flows to receiving waters.

Chief among inventory data are contaminant-related data (1):

- pollutants of concern and effects on water resources;
- a base map for locating pollution sources and controls;
- areas of concern where pollutant loadings pose high public health and ecological health risks and where control efforts should focus;
- areas of high water quality where protection should focus; and
- information for developing water quality models, if needed.

The inventory of existing data reveals gaps to be filled by sampling and monitoring. Chambers Dictionary defines sampling as: the “taking of, testing of a sample; the examination and analysis of data obtained from a random group in order to deduce information about the population as a whole”; to monitor is to “watch, check, supervise” (2). The inventory also provides information resources for setting clear concise IWRM goals. The goals and data gaps determine sampling and monitoring regimes.

Six basic practical considerations of environmental sampling/monitoring are as follows

- Which environmental media to sample (surface water, groundwater, air, soil, sediment, biota), and which standardized sampling method to use to obtain the sample?
- Where in the contaminant-environment cycle to take samples (physical location)?
- Which indicators/parameters of environmental quality to measure and which laboratory analysis methods to apply to detect them?
- When to take samples (times of the year, e.g., wet and dry seasons, cold and hot seasons)?
- How often to sample (frequency of sampling, subhourly, hourly, daily, weekly, monthly)?
- How to report contamination results (implied risks, impacts, data and modeling limitations and assumptions)?

Figure 1 shows stages in sampling/monitoring as an adaptive process and its input data.

Knowing the following greatly informs water–contaminant system sampling and characterization:

- physicochemical characteristics of the contaminant of interest especially its degree of water phase affinity (hydrophilicity) or water phase aversiveness (hydrophobicity),
- dominant removal processes—decay by chemical reaction (e.g., hydrolysis, aqueous oxidation or reduction), sunlight degradation (photolysis), and/or microbial (bio-) degradation;
- dominant transport processes: sediment transportation in runoff, soil erosion and washing into rivers and lakes, leaching into aquifers, dry deposition, or wet deposition from air into water;
- contaminant toxicity and risk: human cancer and noncancer health effects from sufficient exposure by ingestion, inhalation, and/or dermal/ocular contact; and nonhuman health effects (ecotoxicology).

SPATIAL CONSIDERATIONS—WHERE TO SAMPLE AND MONITOR

Conservation of mass (see Eq. 1) always applies, so scientific sampling and monitoring of water/wastewater quality

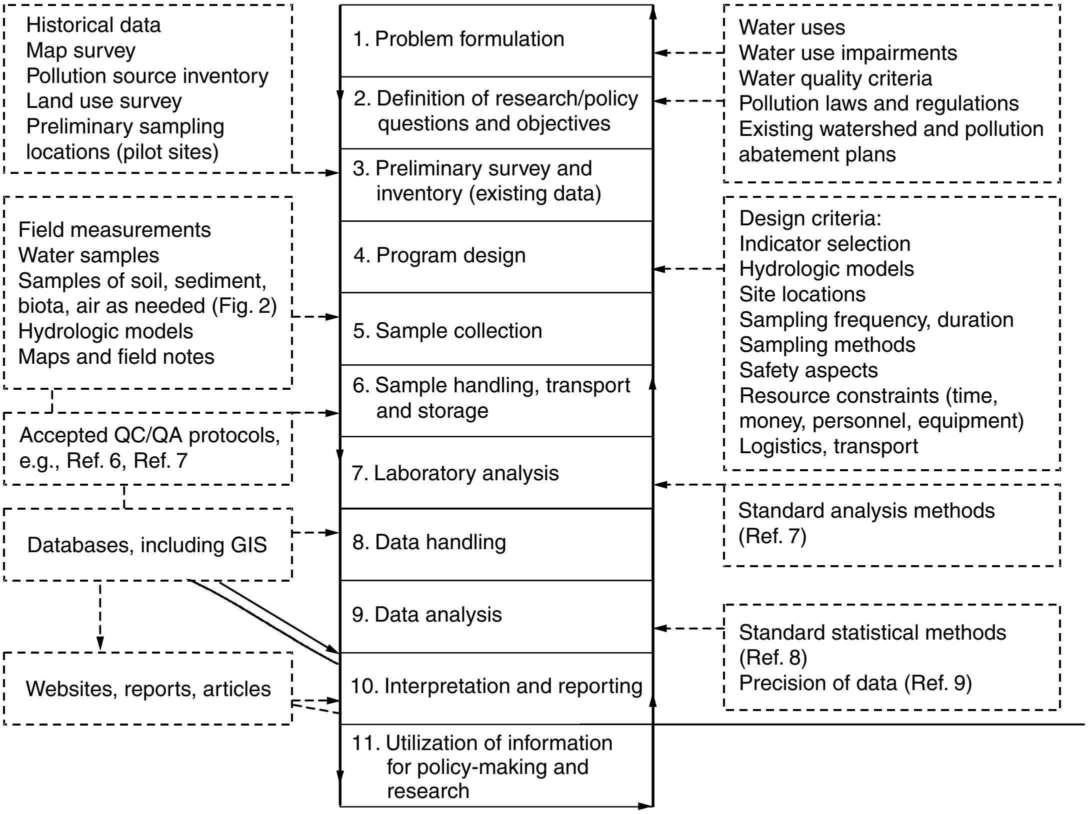

Figure 1. Stages of water quality monitoring. Shows determining factors, sources of data, and standard methods (expanded from Ref. 5). Stages 1–11 can be thought of as an adaptive feedback loop, responding to changing conditions, interests, capabilities, and priorities over time. In this way, the sampling/monitoring becomes sustainable. A key dimension to the relative sustainability is the strengthening of capacities/resources (human, informational, policy-making, financial, and equipment).

seek to take measurements that provide information about the following aspects of the contaminant source stream and the receiving water stream (4):

- transport processes of contaminants from sources to receiving waters (advection, diffusion, volatization, atmospheric deposition)
- contaminant transformation processes from pollution source to receiving waters
- the timescale on which dominant processes operate
- the spatial scale on which dominant processes operate
- the factors governing these time and space rates
- the existing data providing information about the above
- the missing data needed to be collected
- the models that can be used to represent the processes and to explore contaminant control and water management change scenarios.

$$\begin{aligned}&\text{Rate of contaminant mass increase in water}\\&\text{(or any bounded space)} = \text{rate of mass input}\\&- \text{rate of mass output} + \text{rate of mass created}\\&\text{internally} - \text{rate of mass lost internally}\end{aligned}\tag{1}$$

Under different hydrologic conditions, the relative contribution to pollution loading by nonpoint versus point sources can change dramatically. For example, agricultural sources of nutrients and urban runoff sources of hydrocarbons may become dominant during storms compared to factory effluents and sewage treatment effluents that dominate pollution loading in dry conditions. Figure 2 provides a useful schematic framework for sampling/monitoring water and wastewater flows.

TEMPORAL CONSIDERATIONS—WHEN AND HOW OFTEN TO SAMPLE

Watershed assessments can be simple or detailed; the latter require more data collected over an extended period of time. Simple assessments (taking weeks) can be used to design more detailed ones (taking months or years) or are used when there are extreme constraints on resources that can be applied (the case in many poor countries). The following conditions require a detailed assessment (4):

- high variability in water quality and quantity in time and/or space;
- receiving waters for pollution are complex (e.g., large lakes, rivers, nonuniform aquifers), requiring two- or three-dimensional sampling over time;

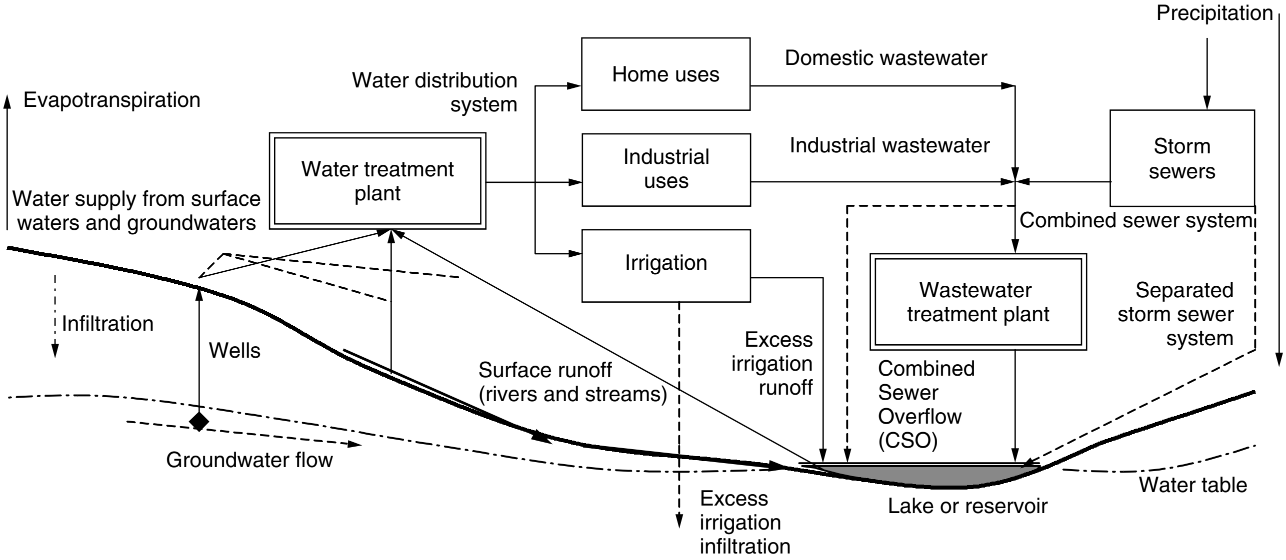

Figure 2. Water and wastewater flows to and from receiving waters. A full hydrologic picture is needed to assess where sampling of water quality and quantity should occur for baseline (preaction) characterization and where sampling should be done to monitor the performance of contaminant controls and water management actions (expanded from Ref. 3).

- existing data are insufficient to model the impacts of management alternatives on the water resources; and/or
- complex supply and sanitation systems need to be modeled to optimize their performance under varying conditions (e.g., treatment plants that have multiple unit process stages between influent and effluent).

Contaminants are sampled and monitored to determine the nature and magnitude of contamination and the sources of it. Simple assessments rely largely on existing data; detailed water quality assessments involve considerable field sampling to estimate the hydrograph (flow vs. time) and its associated pollutograph (pollution levels vs. time).

Sampling design is a statistical question. A hypothesis is posed, for example, "Groundwater contamination by arsenic (a highly toxic, naturally occurring substance that has potential cancer and noncancer health effects) represents a potential health risk (if enough water is consumed) because levels exceed health standards on the average." To test such hypotheses requires taking samples that allow us to estimate the 'real' variability of the values. Samples allow us to determine the sample mean (x_m) and sample variance (s^2), that are used to estimate the population/complete system's mean (μ) and variance (σ^2). The challenge, as in any sampling, is to characterize population variability using a representative sample. Sample statistics and a knowledge of the way values are distributed by probability (normally, log normally, or otherwise), allows us to accept or reject a hypothesis at a given confidence level. Detailed sampling methods for specific situations are given in Reference (10).

The minimal water quality sample requirements are three samples, representative of low, moderate, and high contaminant concentrations. But to gauge "representativeness" requires knowledge of variability. The goal of a good sampling program is to provide information to scientists and policy-makers on the following (4):

- probable maximum and minimum values;
- variability on different timescales (hours, days, months, years, decades as required);
- correlations with associated variables (hydrological, meteorological, pollutant production rates); and
- evidence about sources of contamination, environmental transport, and degradative/transformative processes.

Point-pollution sources, such as factories and sewage treatment plants, can be highly variable in time for the level (concentration as mass of contaminant per unit volume) and type (microbial, organic chemical, inorganic chemical, conservative, or nonconservative) of contaminants. Adding natural variability in environmental conditions and receiving waters (flow rates, temperatures, pH, other contaminant levels), it is evident that data quality and reliability are highly sensitive to the sampling regime selected.

Samples are either discrete or composite. If sampling seeks merely to estimate average levels of contaminants in water or wastewater, composite samples can save time by combining subsamples (aliquots) of discrete samples together. However, information about variability (mean, median, mode, variance, ranges, 90–95% confidence intervals, probability distributions) is lost, and it is these data that are most relevant to risk assessments and management policy. There are important trade-offs between sampling/monitoring costs and information value. Clear sampling/monitoring objectives must guide the optimization of this trade-off.

Rapidly changing conditions of water quantity and quality over time often call for flow- or quality-sensitive sampling: the higher the variability, the higher the sampling frequency. Under such conditions, it is desirable to monitor the qualities and quantities of contaminant sources (wastewater flows) and receiving waters simultaneously.

WATER QUALITY INDICATORS—CRITERIA FOR SAMPLING/MONITORING PARAMETERS

It is helpful to classify water contaminants into eight groups: pathogens, oxygen-demanding wastes, nutrients, salts, thermal pollution, heavy metals, pesticides, and volatile organic compounds (11). Each water use has quality requirements monitored by these indicators.

Pathogens (Disease-Causing Organisms)

These belong to four main groups:

1. bacteria—causing diseases like cholera, bacillary dysentery (shigellosis), typhoid, and paratyphoid fever;
2. viruses—diseases including infectious hepatitis, poliomyelitis;
3. protozoa—causing amoebic dysentery, giardiasis, and cryptosporidiosis;
4. helminths (parasitic worms)—causing schistosomiasis, and dracunculiasis (guinea worm)

There are four main types of water pollution diseases: waterborne (by ingestion); water-washed (lack of clean water for washing); water-based (involve water contact, no ingestion); and water-related (pathogen host relies on water habitat; humans need no water contact).

Oxygen-Demanding Wastes (ODWs)

Dissolved oxygen (DO) is a primary water quality indicator (8 mg/L needed for coldwater fish; 5 mg/L often used for others) and a saturation value 8–15 mg/L (depending on temperature, salinity). ODWs oxidize in receiving waters as bacteria use oxygen to degrade them. As DO drops, aquatic life is threatened, and odors, tastes, and color become undesirable. ODWs are mainly biodegradable wastes from households and livestock and effluents from food processing and paper production. The wastewater indicators are chemical oxygen demand (COD)—the amount of oxygen needed to oxidize wastes chemically—and biochemical oxygen demand (BOD)—the amount needed by microbes. The BOD of a diluted wastewater sample after 5 days (BOD_5) is used to measure the performance of wastewater treatment plants (WWTPs).

Nutrients

The chemicals essential to life—nitrogen (N), phosphorus (P), carbon (C), calcium (Ca), potassium (K), iron (Fe), manganese (Mn), boron (B), and cobalt (Co)—are viewed as contaminants when their levels allow excessive algal growth and compromise drinking uses, recreation, and habitat for other species. Nutrient enrichment occurs by *eutrophication*, important in lakes. C, N, and P are the most important nutrients. The limiting nutrient controls growth, and it is usually N in seawaters and P in freshwaters. Major N sources are WWTPs, animal feedlot runoff, chemical fertilizers, and N deposits from polluted air. NO_3 (nitrate) is the most common form in water. P is limited in nature, but human actions contribute it: runoff of fertilizers and domestic sewage (feces and detergents).

Salts

Dissolved solids, cations, anions—total dissolved solids (TDS)—is the indicator used: TDS <1500 mg/L is freshwater, 1500–5000 mg/L is brackish water, and saline water is >5000 mg/L. Seawater has 30–34,000 mg/L TDS. Drinking water levels must be <500 mg/L for humans but can be higher for livestock. Irrigation waters with TDS >500 mg/L and poor drainage can lead to salt buildup in soils that affect crops. Water is unsuitable for irrigation above 2100 mg/L.

Thermal Pollution

Cooling waters from power plants change water temperature, affecting aquatic species. Warming benefits some species over others, but power outages create sudden drops in temperature. Increases in temperature increase organism metabolic rates so oxygen demand goes up, waste degradation rates also go up, and the amount of DO that water can hold drops. All these factors mean that DO is drawn down.

Heavy Metals

These elements give up one or more electrons in aqueous solution to form cations (X^{n+}). "Heavy" means that the specific gravity is >4–5. Those most important for environmental impact are mercury (Hg), lead (Pb), cadmium (Cd), and arsenic (As). Metals are nondegradable and become toxic above trace levels.

Pesticides

Insecticides, herbicides, rodenticides, and fungicides are of three main chemical groups: organochlorines (OCPs or chlorinated hydrocarbons, e.g., DDT, Kepone®, aldrin, dieldrin), organophosphates (OPPs, e.g., parathion, malathion, diazinon), and carbamates (e.g., aldicarb). OCPs are persistent and tend to bioaccumulate in fatty tissue. OPPs are not persistent, but they are more acutely toxic to humans than OCPs. Carbamates are also short-lived and of high acute toxicity to humans.

Volatile Organic Compounds (VOCs)

Their high volatility means we find low levels in surface waters (a few μg/L), but when released as solvent waste into soil, they migrate to groundwater, and levels can be 100–1000 times surface water levels. Five VOCs are very toxic: vinyl chloride (VC), tetrachloroethylene (PCE), trichoroethylene (TCE), 1,2-dichloroethane (metal degreaser), and carbon tetrachloride (household cleaner, grain fumigant, solvent). VC is a known human carcinogen that is used in the production of PVC resins. TCE, a solvent cleaner used on engines, electronics, even septic tanks, is the most common groundwater contaminant in the United States.

WATER QUALITY STANDARDS—U.S., EUROPEAN UNION, AND WORLD HEALTH ORGANIZATION

Water quality (and treatment) standards depend on the usage of the water—drinking, recreational, industrial, agricultural, hydroelectric, or ecological.

World Health Organization (WHO) Standards

WHO standards are used as the reference for all national standards (12,13). There are five types of sampling/monitoring parameters:

1. Bacteriological: *E. coli* or thermotolerant coliform bacteria and total coliform bacteria must not be detectable in drinking water anywhere in the distribution system (once leaving the treatment plant).
2. Inorganic chemical (levels in mg/L): 19 inorganics.
3. Organic chemical (levels in μg/L): 11 classes of organic substances: (i) chlorinated alkanes (five parameters, e.g., dichloromethane); (ii) chlorinated ethenes (five, e.g., vinyl chloride); (iii) aromatic hydrocarbons (six, e.g., benzene); (iv) chlorinated benzenes (five, e.g., 1,2-dichlorobenzene); (v) miscellaneous organics (nine, e.g., acrylamide); (vi) pesticides (30, e.g., lindane); chlorophenoxy herbicides (six, e.g., 2,4-DB); (vii) disinfectants (five, e.g., chlorine); (viii) disinfectant by-products (12, e.g., bromoform); (ix) chlorinated acetic acids (five, e.g., dichloroacetic acid); (x) haloacetonitriles (six parameters, e.g., chloropicrin).
4. Radionuclide (Bq/L): gross alpha-particle and gross beta-particle activity screening levels.
5. Secondary physicochemical "nuisance" parameters: 18 inorganic/physical parameters (e.g., iron, copper, color, turbidity); and 13 organics.

Standards for carcinogens, such as the pesticide alachlor, are given as levels for a given acceptable risk, based on the precautionary assumption that any exposure to a carcinogen has a finite risk.

European Union Standards

Domestic Supply. In the European Union (EU), raw surface waters intended for domestic supply (the most stringent quality) are placed into three categories—A1, A2, A3—based on the sampling of 46 parameters from pH, odor, through metals, organics, nitrate, nitrogen, ammonia, to total and fecal coliforms, fecal streptococci and *Salmonella*. Each category has a guide limit and a (higher) mandatory limit, and each must undergo different water treatment, ranging from simple physical treatment and disinfection (A1), through normal physicochemical treatment and disinfection (A2), to intensive physicochemical treatment and disinfection (A3) (14). Once the waters have been treated, drinking water must be sampled and monitored to meet the new (1998) EU drinking water standard grouped as follows (15):

- Microbiological monitoring: Tap water—*E. coli* and enterococci (both 0/100 mL); water for sale in bottles and containers—*E. coli* and enterococci (both 0/250 mL), bacterial colony counts at 22 °C (100/mL), 37 °C (20/mL), *Pseudomonas aeruginosa* (0/250 mL).
- Chemical monitoring: 26 parameters—acrylamide, antimony, arsenic, benzene, benzo(a)pyrene, boron, bromate, cadmium, chromium, copper, cyanide, 1,2-dichloroethane, epichlorohydrin, fluoride, lead, mercury, nickel, nitrate, nitrite, specified pesticides, total specified pesticides, specified polyaromatic hydrocarbons (PAHs), selenium, tetrachloroethene plus trichloroethane, total specified trihalomethanes, and vinyl chloride.
- Indicator monitoring parameters: aluminum, ammonium, chloride, *Clostridium perfringens*, color, conductivity, pH, iron, manganese, odor, sulfate, sodium, taste, bacterial colony count, coliform bacteria, total organic carbon, turbidity, tritium (radioactivity), and total indicative radiation dose.

Sampling is done exclusively at consumers' taps under the EU Directive (98/83/EEC).

Surface Waters Not Used Primarily for Supply. The EU Freshwater Fish Directive specifies 14 physicochemical parameters to protect salmonid and cyprinid waters: temperature (where there is a thermal discharge), dissolved oxygen, pH, suspended solids, biochemical oxygen demand, total phosphorus, nitrite, phenolic compounds, petroleum hydrocarbons, nonionized ammonia, total ammonium, total residual chlorine, total zinc, and dissolved copper. Guide and mandatory values are given. The EU also has a Bathing Water Directive (76/110/EEC) for inland and coastal waters. Coastal and estuarine waters are monitored using these physicochemical and microbial water quality parameters, most sampled every two weeks or more frequently when water quality is compromised (15):

- Microbiological: total coliforms, fecal coliforms, fecal streptococci, *Salmonella*, entero viruses;
- Physicochemical: pH, color, mineral oils, surfactants that react with methylene blue, phenols, transparency, dissolved oxygen (% saturation), tarry residues and floating debris, ammonia, nitrogen (Kjeldahl), plus any other substances chosen as appropriate indicators of pollution (site specific).

Wastewaters. The EU monitoring of wastewaters is broken into toxic, persistent, and potentially bioaccumulating substances (black list) and less toxic substances or those whose effects are confined to a limited area depending on the nature of the receiving water (gray list). The Dangerous Substances Directive provides for monitoring the following parameters in wastewater streams:

- List 1 ("black" list): organohalogen compounds (or substances that may form them in aquatic environments), organophosphorus compounds, organotin compounds, carcinogens in the aquatic environment (from gray list), mercury and its compounds, cadmium and its compounds, persistent mineral oils and petroleum hydrocarbons, and persistent synthetic substances.
- List 2 ("gray" list): metalloids/metals and their compounds, biocides and their derivatives, substances that adversely affect taste and odor, toxic/persistent silicon compounds (or substances that may form them in aquatic environments), inorganic compounds of phosphorus and elemental phosphorus, nonpersistent mineral oils and petroleum hydrocarbons, cyanides, fluorides, and oxygen-depleting substances such as ammonia and nitrites.

Groundwaters. EU List 1 substances must be prevented from entering groundwaters, and pollution by List 2 substances must be strictly regulated and monitored. Strict hydrologic and environmental impact must be assessed before any listed substance can be licensed for disposal to groundwater in the EU.

U.S. Standards

U.S. Chemical Standards. In the United States, drinking water standards are of two types: (1) primary standards, specified as maximum contaminant levels (MCLs) derived from health concerns; and (2) unenforceable secondary standards based on taste, odor, color, hardness, and corrosivity. Maximum contaminant level goals (MCLGs) are levels that represent no anticipated health effects, including a margin of safety, regardless of attainment cost and feasibility. Inorganic substances that must be monitored (and for which there are MCLs in the U.S.) include toxic heavy metals (e.g. arsenic, cadmium, lead, mercury); nitrites (NO_2) and nitrates (NO_3); fluoride; and asbestos fibers. Organic chemical contaminants that have MCLs in the United States can be put in three groups (16):

1. Synthetic organic chemicals (SOCs)—used in the manufacture of agricultural and industrial products, such as insecticides and herbicides.
2. Volatile organic chemicals (VOCs)—synthetic substances that vaporize at room temperature. Examples include degreasing agents, paint thinners, glues, dyes. Specific examples include benzene, trichloroethylene (TCE), and vinyl chloride.
3. Trihalomethanes (THMs)—by-products of water chlorination, including chloroform, dibromochloromethane, and bromoform.

The levels are measured in milligrams per liter (mg/L). For the MCLs of primary drinking water contaminants regulated in the United States see (17,18).

U.S. Microbiological Standards. Instead of testing routinely for single pathogenic organisms, simpler tests are used that measure indicator organisms, principally coliform bacteria (e.g., *E. coli*), that indicate the contamination by fecal material. Levels of coliform bacteria in feces and untreated domestic wastewater are very high. In the United States, microbial standards require that for 'large' water supplies (serving more than 1000 people), no more than 5% of test samples can show any coliforms in smaller systems, testing less than 40 samples/month, no more than one sample can test positive (16). This method works when pathogen survival rates in water (outside their host) are less than those of coliforms. But nonbacterial pathogens—especially viruses, *Cryptosporidium* and *Giardia* cysts—often outlive coliforms and can be present when no coliform are detected. Treatment techniques for these contaminants are specified instead of MCLs in the United States. Coliform sampling is also used to monitor water-contact recreational water quality; typical standards are 1000 coliforms per 100 milliliters.

U.S. Radionuclide Standards. Naturally occurring radionuclide contaminants include radon and radium-226, often found in groundwater. Others, such as strontium-90 and tritium contaminate surface water after nuclear fallout. The levels are measured in picocuries per liter (pCi/L). Dissolved radon gas is the most prominent groundwater contaminant of this type. Colorless, odorless, and tasteless, the risk derives not from ingestion but from inhalation as it outgases from heated shower water.

BIBLIOGRAPHY

1. Bingham, D.R. (1998). Watershed management in U.S. urban areas. In *Watershed Management—Practice, Policies and Coordination*. R.J. Reimold (Ed.). McGraw-Hill, New York, pp. 181–189.
2. Kirkpatrick, E.M. (Ed.). (1983). *Chambers 20th Century Dictionary*. W & R Chambers Ltd., Edinburgh.
3. Masters, G.M. (1996). *Introduction to Environmental Engineering and Science*, 2nd Edn. Prentice-Hall, Englewood Cliffs, NJ.
4. Heathcote, I.W. (1998). *Integrated Watershed Management—Principles and Practice*. John Wiley & Sons, Hoboken, NJ, pp. 196–206.
5. Gray, N.F. (1999). *Water Technology—An Introduction for Environmental Scientists and Engineers*. Arnold, London, p. 135.
6. ISO (1990). *Water Quality Sampling—International Standard 5667*. International Organization for Standardization, Geneva.
7. APHA (1989). *Standard Methods for the Examination of Water and Wastewater*. American Public Health Association, Washington, DC.
8. Demayo, A. and Steel, A. (1996). Data Handling and Presentation. In: *Water Quality Assessments*. D. Chapman (Ed.). E and FN Spon, London, pp. 511–612.
9. Caulcutt, R. and Boddy, R. (1983). *Statistics for Analytical Chemists*. Chapman and Hall, London.
10. Ostler, N.K. and Holley, P.K. (Eds.) (1997). *Sampling and Analysis*, *Environmental Technology Series*, Vol. 4. Prentice-Hall, Upper Saddle River, NJ.

11. Masters, G.M. (1996). *Introduction to Environmental Engineering and Science*, 2nd Edn. Prentice-Hall, Englewood Cliffs, NJ, pp. 175–184.
12. WHO (1993). *Revision of WHO Guidelines for Drinking Water Quality*. World Health Organization, Geneva.
13. WHO (1993). *Guidelines for Drinking Water Quality,* Vol. 2, *Health Criteria and Other Supporting Information*. World Health Organization, Geneva.
14. Gray, N.F. (1999). *Water Technology—An Introduction for Environmental Scientists and Engineers*. Arnold, London, pp. 1–5.
15. Gray, N.F. (1999). *Water Technology—An Introduction for Environmental Scientists and Engineers*. Arnold, London, pp. 16–18.
16. Masters, G.M. (1996). *Introduction to Environmental Engineering and Science*, 2nd Edn. Prentice-Hall, Englewood Cliffs, NJ, pp. 269–274.
17. Hammer, M.J. and Hammer, M.J. Jr. (1996). *Water and Wastewater Technology*, 3rd Edn. Prentice-Hall, Englewood Cliffs, NJ.
18. USEPA (1995). *National Primary Drinking Water Standards*, U.S. Environmental Protection Agency, Washington, DC.

WATER QUALITY MODELS: CHEMICAL PRINCIPLES

Zhen-Gang Ji
Minerals Management Service
Herndon, Virginia

Water quality represents the physical, chemical, and biological characteristics of water resources and is used to measure the ability of a waterbody to support beneficial uses. Eutrophication is a process of nutrient overenrichment of a waterbody, resulting in accelerated biological productivity (growth of algae and weeds). Symptoms of eutrophication include algal blooms, reduced water clarity, and oxygen depletion. In modeling studies, water quality and eutrophication are sometimes used interchangeably to represent the processes of waterbody enrichment by nutrients. Key state variables in water quality models include algae, nutrients, and dissolved oxygen (DO).

This entry focuses on water quality models. Information on applications of water quality models is provided in Water Quality Modeling—Case Studies.

ALGAE AND PHYTOPLANKTON

Algae are a group of aquatic microscopic plants that contain chlorophyll and grow by photosynthesis. Algae takeup nutrients, including phosphate, ammonium, nitrate, silica, and carbon dioxide, from water or benthic sediments and release oxygen to the water. Algae may be free-floating or rooted on the bottom of a waterbody. An overabundance of algae in a waterbody is known as eutrophication.

Phytoplankton are a group of tiny, free-floating plants transported by currents. In water quality modeling studies, phytoplankton are frequently referred to as free-floating algae and are often more important than rooted aquatic vegetation in the basic food production of an ecosystem. They are the most biologically active plants in aquatic ecosystems and generally have greater influences on water quality than other plants. Phytoplankton is also the most important element and the first stage in the food chain: it provides food for zooplankton, fish, and small aquatic animals.

NUTRIENTS

Nutrients are chemical elements or compounds necessary for algal growth. The essential nutrients include nitrogen, phosphorus, carbon dioxide, and silica. A host of other micronutrients, such as iron, manganese, potassium, sodium, copper, zinc, and molybdenum, are also required for algal growth. However, these minor nutrients are generally not considered in water quality modeling because they are needed only in trace amounts, and they are usually present in quantities sufficient for algal growth.

Excessive nutrient levels can be harmful to ecosystems by causing eutrophication and unwanted algal blooms. Major processes affecting nutrient concentrations include

1. algal uptake
2. mineralization and decomposition of dissolved organic nutrients
3. hydrolysis converting particulate organic substances into dissolved organic form
4. chemical transformations of nutrients
5. sediment sorption and desorption
6. settling of particulate matter
7. nutrient fluxes from benthic sediments
8. external nutrient loadings

The two nutrients of greatest concern are nitrogen and phosphorus. Nitrogen, one of the most abundant elements on the earth, makes up 78% of the earth's atmosphere as a gas. Nitrogen exists in these different forms:

1. organic nitrogen (ON)
2. ammonia (NH_3)
3. nitrite and nitrate ($NO_2 + NO_3$)
4. nitrogen gas (N_2)
5. algae
6. zooplankton and aquatic animals

Phosphorus is one of the vital nutrients for algal growth. However, too much phosphorus can cause excessive algal growth and can lead to eutrophication. Compared to nitrate (NO_3), phosphates dissolve less readily and tend to attach to sediment particles. Phosphorus exists in different forms, including

1. organic phosphorus
2. phosphates (largely orthophosphates—PO_4)
3. algae
4. zooplankton and aquatic animals

In modeling water quality processes, these are the major differences between nitrogen and phosphorus:

1. Unlike nitrogen, phosphorus does not have a gas phase, which can be a major nitrogen loss mechanism in certain aquatic systems.
2. Forms of inorganic nitrogen are easily dissolved in water; phosphates are often strongly sorbed to sediment. Therefore, phosphates can settle with sediment solids to the bottom of a waterbody, such as a lake or a reservoir, and become a phosphorous source to the waterbody. In many waterbodies, bottom sediments contain enough phosphorus to accelerate eutrophication, even after external sources have been terminated.

Figure 1 illustrates processes that affect nutrients in an aquatic system:

1. Physical transport: Nutrients are advected and dispersed within the water column and are transported into the system by inflow and out of the system by outflow.
2. Exchange with the atmosphere: Atmospheric deposition adds nutrients to a waterbody, and volatilization removes gaseous nutrients from the waterbody. Reaeration adds DO to the waterbody.
3. Sorption and desorption: Exchanges between particulate nutrients and dissolved nutrients are affected by the total suspended solid concentrations and the partition coefficients.
4. Reaction and algal uptake: Chemical reactions and algal uptake transform nutrients and reduce the concentrations of dissolved nutrients.
5. Exchange on the bed–water interface: Dissolved nutrient is exchanged between the sediment bed and the water column via diffusion, and the particulate nutrient can be settled onto or be resuspended from the bed, depending on flow conditions.
6. Sediment diagenesis: In the sediment bed, sediment diagenesis can be a significant factor in nutrient cycling and oxygen balance in the water column.

Note that not all processes shown in Fig. 1 are essential to every nutrient. For instance, volatilization is insignificant to phosphorus cycling, and sorption and desorption are not essential to nitrogen transformation.

DISSOLVED OXYGEN

Dissolved oxygen, the amount of oxygen that is dissolved in water, is one of the most important parameters of water quality. Most fish and aquatic insects need DO to survive. Fish, especially larvae, die when DO levels get too low. Low DO is a sign of possible pollution in a waterbody. As DO levels in water drop below 5.0 milligrams/liter (mg/L), aquatic life is put under stress. The lower the concentration, the greater the stress. Low DO levels for a prolonged period can result in large fish kills.

Oxygen enters water by reaeration from the atmosphere and by plant photosynthesis. Oxygen concentrations in the water column fluctuate under natural conditions, but severe oxygen depletion usually results from human pollution. Figure 2 shows the measured DO for 20 months from November 1992 to June 1994 in Lake Wister, Oklahoma (1). The lake exhibits strong seasonal DO variations. During the winter months (November, December, and January), DO in the lake is high; values are around 10 mg/L, and it is well mixed vertically. In the summer, DO is very stratified; values are less than 1.0 mg/L in the lower portion of the lake (water depth from 9 to 18 m), whereas the surface DO can still be 8 mg/L or higher. This large DO gradient in the vertical is caused primarily by temperature stratification of the lake, biochemical processes in the water column, and benthic oxygen demand at the lake bottom.

WATER QUALITY MODELS

Variations in water quality variables, such as algae, nutrients, and dissolved oxygen, are often described by

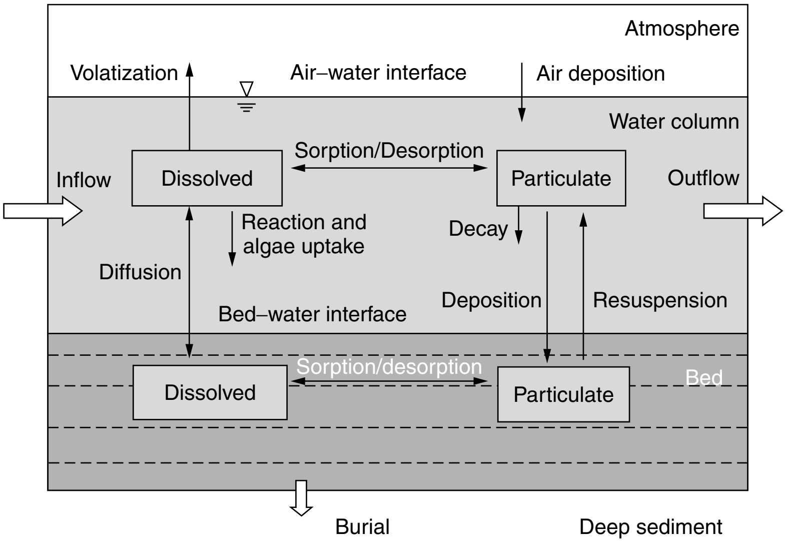

Figure 1. Processes of nutrient transformation.

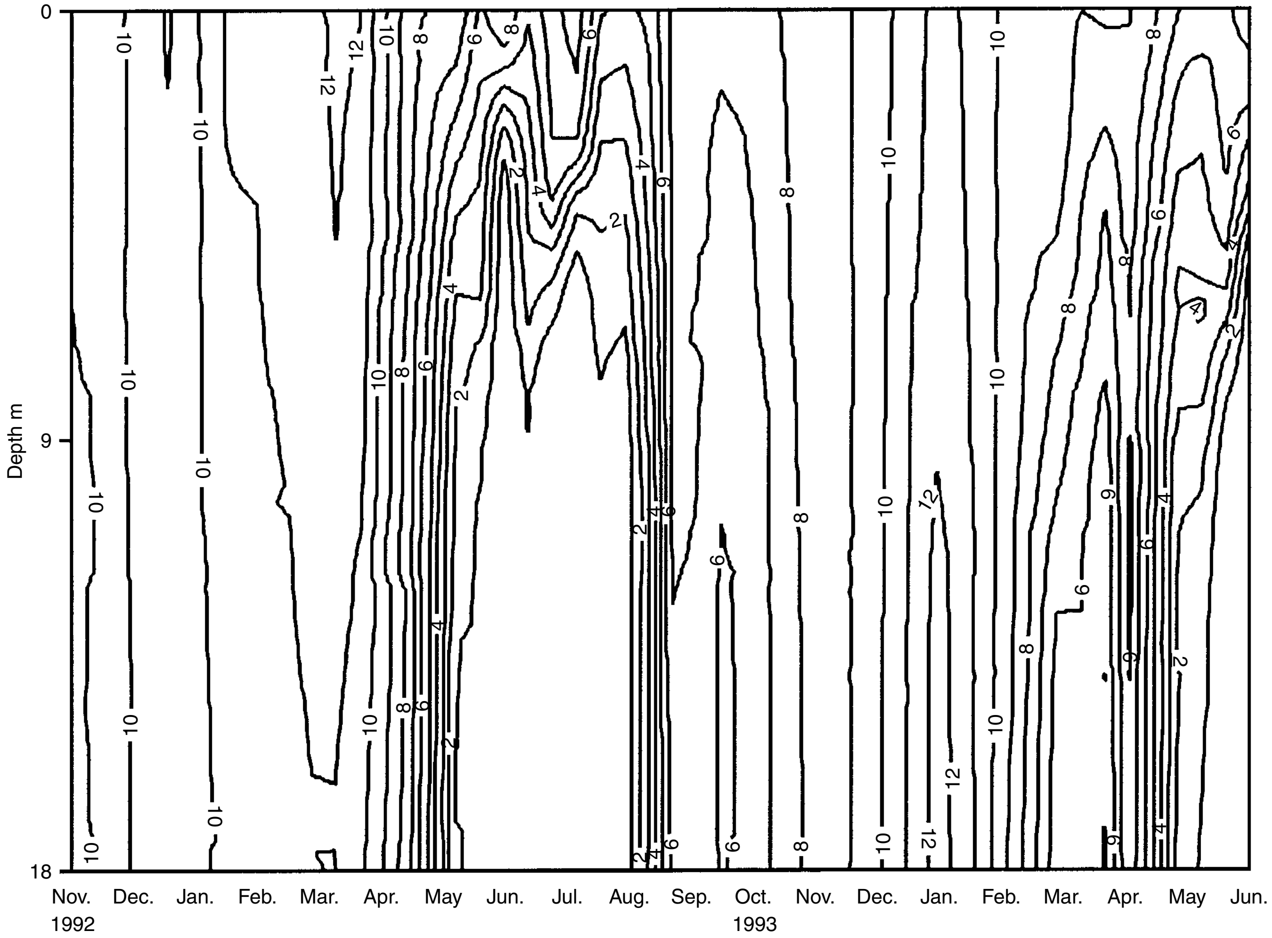

Figure 2. Dissolved oxygen (mg/L) in Lake Wister (1).

using a set of mass conservation equations in a water quality model. The conservation of mass accounts for all material entering or leaving a waterbody; transport of material within the waterbody; and physical, chemical, and biological transformations of the material. Therefore, all governing equations for water quality processes have a form similar to

$$\frac{\partial C}{\partial t}+\frac{\partial (uC)}{\partial x}+\frac{\partial (vC)}{\partial y}+\frac{\partial (wC)}{\partial z}=\frac{\partial}{\partial x}\left(K_x\frac{\partial C}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y\frac{\partial C}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z\frac{\partial C}{\partial z}\right)+S_{\mathrm{C}} \tag{1}$$

where C = concentration of a water quality state variable

u, v, w = velocity components in the x, y, and z directions, respectively

K_x, K_y, K_z = turbulent diffusivities in the x, y, and z directions, respectively

S_{C} = internal and external sources and sinks per unit volume

Equation 1 incorporates transport due to flow advection and dispersion, external pollutant inputs, and kinetic interaction among water quality variables. The last three terms on the left-hand side of Equation 1 account for advection transport. The first three terms on the right-hand side of Equation 1 account for diffusion transport. The term S_{C} represents kinetic processes and external loads.

In water quality models, kinetic processes are often decoupled from physical transport processes. By linearizing S_{C} with respect to C, the equation for kinetic processes and external loadings, called the kinetic equation, is

$$\frac{\partial C}{\partial t}=S_{\mathrm{C}}=k\times C+R \tag{2}$$

where k = kinetic rate (time^{-1})

R = source/sink term due to external loadings and/or internal reactions (mass volume^{-1} time^{-1}).

The governing equations, Equations 1 and 2, are widely used in water quality models. Water quality and eutrophication processes are very complicated. Even though all kinetic equations are based on mass-balance equations, empirical formulations are often used as approximations for specifying model parameters, such as k and R. Therefore, for the same water quality process, there might be a variety of ways to describe it mathematically. The major differences among water quality models are

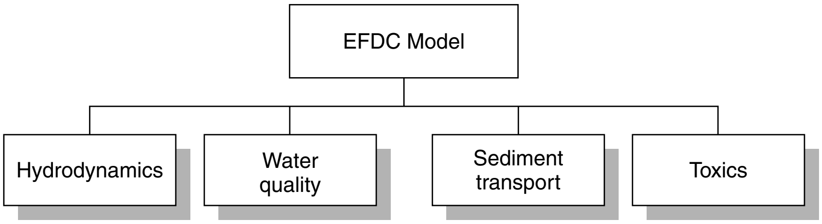

Figure 3. Primary modules of the EFDC model.

primarily in the way the kinetic equations are specified, how many nutrients are considered, and how many state variables are used to describe each nutrient cycle. Mathematical expressions of these terms can vary from model to model (such as References 2 and 3).

A water quality model typically has the following water quality variable groups:

1. algae
2. organic carbon
3. phosphorus
4. nitrogen
5. silica
6. other water quality variables

Each group may consist of several state variables, representing different components of the group. For example, a nutrient can be dissolved, particulate, refractory, or labile.

A water quality model is a discretized version of a set of differential equations, such as Equations 1 and 2, that describe processes in a waterbody. The discretized set of equations is converted into computer code, so that numerical solutions of the model can be derived by computer.

Numerical models span a wide variety of approaches and levels of sophistication and are typically categorized in terms of their representation of space and time:

1. steady-state or time-dependent (dynamic)
2. zero-, one-, two, or three-dimensional

The temporal characteristics indicate whether the model is steady-state (inputs and outputs constant over time) or time-dependent (dynamic). A steady-state model uses constant values of input variables to produce time-independent results, in which the state variables are independent of time. In contrast, a time-dependent model describes the temporal variability of a waterbody.

Since the 1990s, three-dimensional hydrodynamic and water quality modeling has matured from a research subject to a practical analytical technology. The rapid development of the computer industry provides more and more powerful computing ability for numerical simulation. Computational requirements for realistic three-dimensional modeling have changed from supercomputers and high-end workstations to desktop personal computers. Water quality models are more and more widely used in water resource management (4,5).

A variety of water quality models have been developed for modeling rivers, lakes, estuaries, and coastal areas. They can be categorized as

1. One-dimensional and steady-state models. They are one-dimensional in space and do not vary with time, such as the QUAL2E model (2).
2. Two-dimensional and time-dependent models. They are two-dimensional in space and time-dependent, such as the laterally averaged W2 model (6).
3. Comprehensive models. They are (1) three-dimensional in space and time-dependent, (2) with a hydrodynamic submodel to provide physical transport to the water quality submodel, and (3) coupled with a sediment diagenesis submodel to calculate sediment fluxes from the benthic sediments. The Environmental Fluid Dynamics Code (EFDC) (7,8), supported by the U.S. Environmental Protection Agency, is in this category. The primary modules of the EFDC model are shown in Fig. 3, which indicates submodels for hydrodynamics, water quality process, sediment transport, and toxic processes.

BIBLIOGRAPHY

1. OWRB. (1996). *Final Report for Cooperative "Clean-Lakes" Project, Phase I: Diagnostic and Feasibility Study on Wister Lake*. Oklahoma Water Resources Board, United States Army Corps of Engineers, and Oklahoma State University, Stillwater, OK.
2. Brown, L.C. and Barnwell, T.O. (1987). *The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS: Documentation and User Manual*. U.S. Environmental Protection Agency, Athens, GA. EPA/600/3-87-007.
3. Cerco, C.F. and Cole, T.M. (1994). *Three-Dimensional Eutrophication Model of Chesapeake Bay: Volume 1, Main Report*. Technical Report EL-94-4, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.
4. Ji, Z.-G., Hamrick, J.H., and Pagenkopf, J. (2002). Sediment and metals modeling in a shallow river. *J. Environ. Eng.* **128**: 105–119.
5. Ji, Z.-G., Morton, M.R., and Hamrick, J.M. (2004). Three-Dimensional Hydrodynamic and Water Quality Modeling in a Reservoir. In: *Estuarine and Coastal Modeling: Proc. 8th Int. Conf.* M.L. Spaulding and H.L. Butler (Eds.). ASCE (accepted).
6. Cole, T.M. and Buchak, E.M. (1995). *CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model, Version 2.0*. Instruction Report EL-95, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

7. Hamrick, J.M. (1992). *A Three-Dimensional Environmental Fluid Dynamics Computer Code: Theoretical and Computational Aspects*. The College of William and Mary, Virginia Institute of Marine Science, Special Report 317.
8. Park, K., Kuo, A.Y., Shen, J., and Hamrick, J.M. (1995). *A Three-Dimensional Hydrodynamic–Eutrophication Model (HEM3D): Description of Water Quality and Sediment Processes Submodels*. The College of William and Mary, Virginia Institute of Marine Science, Special Report 327.

WATER QUALITY MODELS: MATHEMATICAL FRAMEWORK

K.W. Chau
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

INTRODUCTION

In recent decades, accompanying agricultural and industrial development and population growth, it has often been observed that an increasing tendency exists for some water bodies, for example, the North Sea (1), along the coast of Brittany to marine waters (2), some rivers in western North America (3), the Chesapeake Bay in the United States (4), and the coastal waters in Hong Kong (5–7), to exhibit increases in the severity and frequency of algae blooms and growth of aquatic weeds apparently as a result of elevated levels of nutrients, including nitrogen and phosphorus. Many of these problems, including bottom-water anoxia, decline in fisheries, and loss of submerged aquatic vegetation, are associated with the deteriorated water quality caused by eutrophication.

Water quality models are employed to describe the balance of mass and energy within an aquatic ecosystem. The problem can be defined as the input of polluting materials including organic and inorganic nutrients into a water body that stimulate the growth of algae or rooted aquatic plants resulting in the interference with desirable water uses of aesthetics, recreation, fish maintenance, and water supply. This topic has been discussed extensively, and many studies, including water quality control and modeling, have been conducted (4,8). Initially, water quality modeling focused on seasonal steady-state conditions (9,10) and, later, time-varied modeling for an entire year was explored with reasonable success (11). Nowadays, the forefronts of water quality modeling are the integration of hydrodynamic and water quality models and the development and incorporation of sediment layers to investigate the long-term recovery of an aquatic ecosystem (4,12,13).

TYPICAL MODEL FRAMEWORK

Typically, in an ecosystem under an aquatic environment, several physical, chemical, biochemical, and biological processes can affect the transport and interaction among the nutrients, dissolved oxygen (DO), phytoplankton, zooplankton, and carbonaceous material (10). The mass and energy transformations are regulated by processes such as growth, respiration, mortality, and decomposition; these in turn are governed by environmental quality parameters such as temperature, toxicity, and nutrient concentrations. The principal variables of importance then in the analysis of eutrophication are as follows:

1. Solar radiation at the surface and with depth, water temperature, and chloride
2. Geometry of water body
3. Flow, velocity, diffusion, and dispersion
4. Nutrients: phosphorus, nitrogen, silica, and so on
5. Phytoplankton, zooplankton, and so on

As shown in Fig. 1, the system is highly coupled, and energy and mass balance for individual constituents are invariably linked to several others. A system of state variables, including organic parameters (carbon, nitrogen and phosphorus), inorganic parameters (DO, ammonia nitrogen, nitrite + nitrate nitrogen, and orthophosphate), and biological constituents (phytoplankton and zooplankton), is used to develop an unsteady multilayered two-dimensional or even three-dimensional water quality model. It should be recognized that the nutrients above are present in several forms in the water body, and not all forms are readily available for uptake by phytoplankton nor are limited nutrients for growth of phytoplankton. In the system, ammonia, nitrite + nitrate, and orthophosphate are taken as the available nutrients uptaken by phytoplankton (13). Moreover, available historical field data of solar radiation intensity, layer-averaged water temperature, and salinity are often taken as the known variables that will mainly affect some parameters and the saturated DO.

GOVERNING EQUATIONS

Because of the difficulty in obtaining regular, high-frequency spatial and temporal field data with reasonable accuracy, it is almost improbable to select a unique correct formulation for a specific process in an ecosystem. Such difficulties can partly be overcome by examination of the actual processes in the laboratory, but this is only, as indicated, partly a solution, because the processes in the ecosystem can never be isolated from their environments and it is impossible to take into account all natural coupling mechanisms. Consequently, each process has several alternative mathematical formulations that, at this stage of ecological modeling, are equally valid to a certain extent.

A water quality model is established to solve the partial differential equations describing conservation of mass and momentum of incompressible fluid and pollutants over the depth of each layer. The layer-averaged water quality variables and transport equations of pollutants solved in the model are defined and derived in a similar way to those of the hydrodynamic model. The hydrodynamic and water quality models are run simultaneously. The water quality algorithms are integrated directly into the hydrodynamic

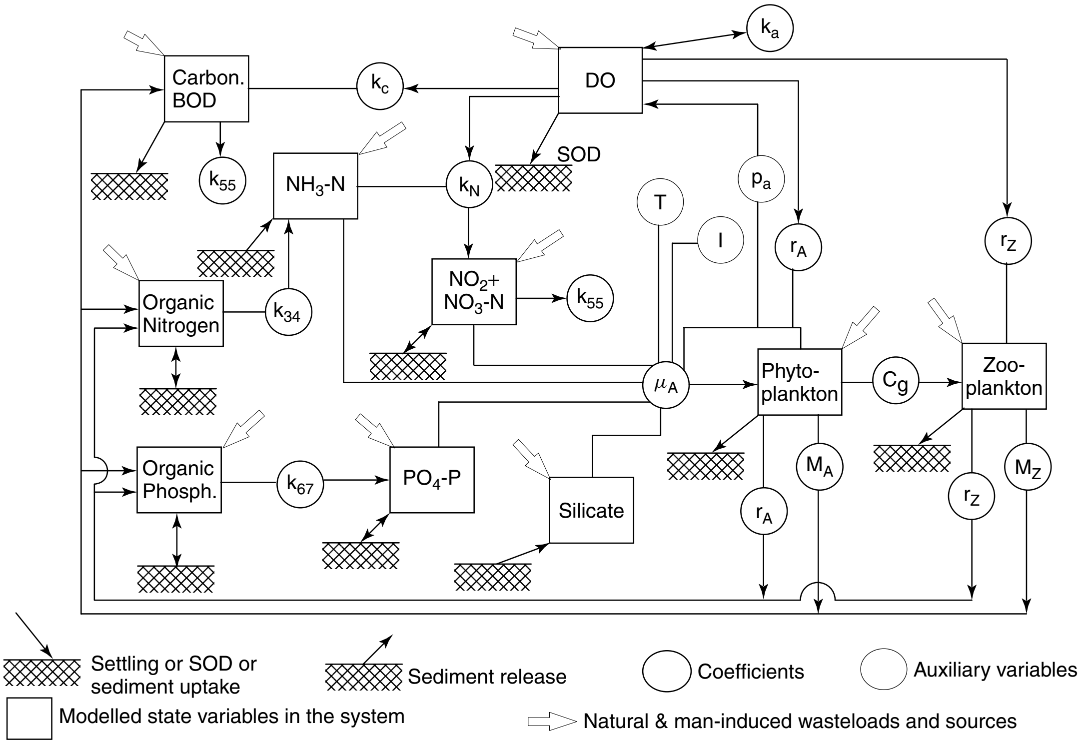

Figure 1. State variable interactions in a water quality model.

model, so that they share the same computational grid and time steps. The water body can be very irregular on the lateral topography and includes some small side cove. To adapt the complicated boundaries of the water body, a boundary-fitted orthogonal curvilinear grid system can be employed. It can be generated numerically by solving two elliptical equations (14). A grid "block" technique (15,16) has been introduced to overcome the computational difficulty caused by the unsteady fluctuation of the water surface level.

In the following sections, subscript "u" and "l" of all variables mean the value in the upper and the lower layer, respectively, and subscript "0", "s", and "b" denote the value at the layer interface, at the free water surface, and at the bed, respectively. Subscript k is the index of layer: $k = u$ or l. Notation $X_{y,20}$ designates the value of X_y at temperature 20 °C, and θ is the corresponding temperature correction coefficient. In this way, the general differential transport equation for the layer-averaged concentration of each state variable in the boundary-fitted orthogonal curvilinear coordinate system is expressed as follows:

$$\frac{\partial(h_k\varphi_k)}{\partial t}+\frac{1}{g_{11}g_{22}}\left\{\frac{\partial}{\partial\xi}\left[g_{22}h_k\left(u_k^*\varphi_k-\frac{\Gamma_{\xi,k}}{g_{11}}\frac{\varphi_k}{\partial\xi}\right)\right]\right.$$
$$\left.+\frac{\partial}{\partial\eta}\left[g_{11}h_k\left(v_k^*\varphi_k-\frac{\Gamma_{\eta,k}}{g_{22}}\frac{\varphi_k}{\partial\eta}\right)\right]\right\}=\Phi_{\varphi,k} \qquad (1)$$

$$\Phi_{\varphi,k}=w_0\varphi_0+\mathrm{Flux}_{s,\varphi}-\mathrm{Flux}_{0,\varphi}+S_{\varphi,k} \qquad (2)$$

where u_k^* and v_k^* = kth-layer-averaged velocity components in $\xi - \eta$ orthogonal curvilinear coordinates; h_k = depth of kth layer; φ_k = kth-layer-averaged concentration of a state variable, and it may represent phytoplankton, DO, and so on; $\Gamma_{\xi,\kappa}$ and $\Gamma_{\eta,\kappa}$ = longitudinal and transverse dispersive coefficients, respectively, in the kth layer (17); g_{11} and g_{22} = coordinate transformation coefficients between Cartesian and orthogonal curvilinear coordinates (ξ,η); $S_{\varphi,k}$ represents reaction kinetics, settling, sediment release, and external sources and/or sinks in the kth layer, and its expression for each water quality variable is listed in Table 1; w_0 = vertical velocity at the layer interface (positive upward); and φ_0 = concentration of φ at the layer interface.

In this context, $\mathrm{Flux}_{s,\varphi}$, $\mathrm{Flux}_{b,\varphi}$, and $\mathrm{Flux}_{0,\varphi}$ are the vertical diffusive fluxes of φ across the water surface, the bed, and the layer interface, respectively. $\mathrm{Flux}_{s,\varphi}$ is assumed to be zero. Density stratification affects the mixing across the layer interface, and $\mathrm{Flux}_{0,\varphi}$ is derived as follows:

$$\mathrm{Flux}_{0,\varphi}=\left(\Gamma_m\frac{\partial\varphi}{\partial Z}\right)_{z=Z_0}\approx\Gamma_m\frac{\varphi_u-\varphi_l}{\delta},$$
$$\Gamma_m=\Gamma_{m0}(1+\beta_C\mathrm{Ri})^{\alpha_C},\ \Gamma_{m0}=\frac{\varepsilon_{m0}}{\sigma_t}, \qquad (3)$$

in which Γ_m = vertical flux diffusive coefficient, Γ_{m0} and $\varepsilon_{m0}(\approx \kappa\ h_l(1-h_l/h)u_*)$ = values of Γ_m and turbulent viscosity coefficient for the case without density stratification, respectively; h = total water depth; u_* = friction velocity; δ = mixing layer thickness at the interface and is approximated by $\delta \approx \kappa\ h_l(1-h_l/h)^{1/2}$. κ ($\approx$0.4) and σ_t ($\approx$0.9) are the von Kármán constant and turbulent Schmidt number, respectively; β_C (= 10/3) and α_C (= −1.5) are the empirical coefficients (18). Ri is a gradient Richardson number

Table 1. List of Notations and Source Terms for Water Quality Variables

Description	Notation	Sources and/or Sinks (S_φ)
(1) Phytoplankton (chlorophyll-*a*)	A (μg Chl-*a*/L)	$h(\mu_A - r_A - M_A)A - w_{\mathrm{SA}}A - hC_gAZ + S_A$
(2) Zooplankton (organic carbon equivalence)	Z (mg ZooplC/L)	$h(\alpha_2 \frac{K_Z}{K_Z + A} C_g \alpha_{12} A - r_Z - M_Z)Z - w_{\mathrm{SZ}}Z + S_Z$
(3) Organic nitrogen	N (μg N/L)	$h[f_{\mathrm{ON}}\alpha_{13}(r_A + M_A)A + f_{\mathrm{ON}}\alpha_{23}(r_Z + M_Z)Z$ $-k_{34}N] - w_{\mathrm{SN}}(1 - f_{\mathrm{DON}})N + \mathrm{SEDI}_N + S_N$
(4) Ammonia nitrogen	$\mathrm{NH_3}$ (μg N/L)	$h\{(1 - f_{\mathrm{ON}})[\alpha_{13}(r_A + M_A)A + \alpha_{23}(r_Z + M_Z)Z] + k_{34}N -$ $f_{\mathrm{pref}}\alpha_{14}\mu_A A - k_N\mathrm{NH_3}\} + \mathrm{SEDI}_{\mathrm{NH_3}} + S_{\mathrm{NH_3}}$
(5) Nitrite + nitrate nitrogen	$\mathrm{NO_{23}}$ (μg N/L)	$h[k_N\mathrm{NH_3} - (1 - f_{\mathrm{pref}})\alpha_{15}\mu_A A - k_{55}NO_{23}] + \mathrm{SEDI}_{\mathrm{NO_{23}}} + S_{\mathrm{NO_{23}}}$
(6) Organic phosphorus	P (μg P/L)	$h\{f_{\mathrm{OP}}[\alpha_{16}(r_A + M_A)A + \alpha_{26}(r_Z + M_Z)Z] - k_{67}P\} - w_{\mathrm{SP}}(1 - f_{\mathrm{DOP}})P +$ $\mathrm{SEDI}_P + S_P$
(7) Orthophosphate phosphorus	$\mathrm{PO_4}$ (μg P/L)	$h\{(1 - f_{\mathrm{OP}})[\alpha_{16}(r_A + M_A)A + \alpha_{26}(r_Z + M_Z)Z] - \alpha_{17}\mu_A A + k_{67}P\} -$ $w_{\mathrm{SPO_4}}(1 - f_{\mathrm{DOP_4}})PO_4 + \mathrm{SEDI}_{\mathrm{PO_4}} + S_{\mathrm{PO_4}}$
(8) Carbonaceous BOD	BOD (mg $\mathrm{O_2}$/L)	$h[2.67(\alpha_{18}M_A A + \alpha_{28}M_Z Z) - (k_C\mathrm{BOD} + \frac{5}{4}\alpha_{\mathrm{ON}}k_{55}\mathrm{NO_{23}})] -$ $w_{\mathrm{SBOD}}\mathrm{BOD} + S_{\mathrm{BOD}}$
(9) Dissolved oxygen	DO (mg $\mathrm{O_2}$/L)	$h\{k_a(\mathrm{DO}^s - \mathrm{DO}) - \alpha_{19}r_A A - \alpha_{29}r_Z Z + \mu_A[\alpha_{19} + \frac{48}{14}\alpha_{13}(1 - f_{\mathrm{pref}})]A -$ $(k_C\mathrm{BOD} + \alpha_{49}k_N\mathrm{NH_3})\} - \mathrm{SOD} + S_{\mathrm{DO}}$

Note: A in the equations for zooplankton, carbonaceous BOD, and DO is in mg Chl-*a*/L; Z in the equations for organic nitrogen, ammonia, organic phosphorus, and orthophosphate is in μg ZooplC/L; and $\mathrm{NH_3}$ in the equation for DO and $\mathrm{NO_{23}}$ in the equation for CBOD are in mg N/L.

at the layer interface and is calculated by

$$\mathrm{Ri} = -\frac{g}{\rho}\frac{\frac{\partial\rho}{\partial z}}{\left(\frac{\partial u}{\partial z}\right)^2} \approx \frac{\rho_l - \rho_u}{\rho}\frac{g\delta}{[(u_u^* - u_l^*)^2 + (v_u^* - v_l^*)^2]} \tag{4}$$

in which ρ_u and ρ_l = density of water in the surface and the bottom layer, respectively; ρ = averaged density of water column; and $g(= 9.81\ \mathrm{m/s^2})$ = acceleration due to gravity. Temporal water temperature and salinity are estimated from the field data.

The quantities of water quality variables decrease because of the corresponding settling—w_{SA}, w_{SZ}, w_{SN}, w_{SP}, $w_{\mathrm{SPO_4}}$, and w_{SBOD} (m/day). The algal growth relates to many factors, including available nutrients, water temperature, solar radiation, zooplankton grazing, and tidal flushing. Organic nitrogen undergoes a bacterial decomposition whose end product is ammonia ($\mathrm{NH_4}$-N). Organic phosphorus converts to $\mathrm{PO_4}$-P by mineralization. Ammonia nitrogen, in the presence of nitrifying bacteria and oxygen, is converted to $\mathrm{NO_3}$-N (k_N). Denitrification (k_{55}) occurs under anaerobic conditions and CBOD decreases because of stabilization (k_C). Both $\mathrm{NH_4}$ and $\mathrm{NO_2 + NO_3}$ are available for algae uptake, although the preferred form is $\mathrm{NH_4}$-N for physiological reasons. The ammonia preference is described by a factor f_{pref} (9):

$$f_{\mathrm{pref},k} = \frac{\mathrm{NH}_{4,k}}{K_{\mathrm{mNN}} + \mathrm{NO}_{23,k}} \times \left(\frac{\mathrm{NO}_{23,k}}{K_{\mathrm{mNN}} + \mathrm{NH}_{4,k}} + \frac{K_{\mathrm{mNN}}}{\mathrm{NH}_{4,k} + \mathrm{NO}_{23,k}}\right) \tag{5}$$

A by-product of photosynthetic carbon fixation is the production of DO. An additional source of oxygen from algal growth occurs when the available ammonia nutrient source is exhausted and the phytoplankton begins to use the available nitrate. For nitrate uptake, the initial step is a reduction to ammonia, which produces oxygen. DO also increases by atmospheric re-aeration (k_a, day^{-1}) because of the deviation from the saturation concentration DO^s (mg $\mathrm{O_2}$/L) (19). The re-aeration rate in natural waters depends on internal mixing and turbulence caused by velocity gradients and fluctuation, temperature, wind mixing, waterfalls, dams, rapids, surface films, and so on. It is expressed in the following forms:

$$k_a = k_{a,20}\theta^{T-20}, \theta = 1.005 \sim 1.030 \approx 1.024 \tag{6}$$

$k_{a,20}$ represents re-aeration as a function of velocity and depth, with the following formula (19):

$$k_{a,20} = \begin{cases} \dfrac{3.9\ V^{0.5}}{h^{1.5}} & \text{when } W < 6.0\ \mathrm{m/s} \\ \dfrac{3.9\ V^{0.5}}{h^{1.5}} + \dfrac{0.728W^{0.5} - 0.317W + 0.0372W^2}{h} & \text{when } W \geq 6.0\ \mathrm{m/s} \end{cases} \tag{7}$$

$$\ln \mathrm{DO}^s = -139.3441 + \frac{1.5757 \times 10^5}{T + 273.15} - \frac{6.6423 \times 10^7}{(T + 273.15)^2} + \frac{1.2438 \times 10^{10}}{(T + 273.15)^3} - \frac{8.6219 \times 10^{11}}{(T + 273.15)^4} - S\left[1.7674 \times 10^{-2} - \frac{1.0754 \times 10}{T + 273.15} + \frac{2.1407 \times 10^3}{(T + 273.15)^2}\right] \tag{8}$$

where V is velocity of flow in meters/second; W is wind speed in meters/second at 10 m above water surface (it will be taken into account when wind speed $W \geq 6$ m/s); and T ($^\circ$C) and S (ppt) are the relevant water temperature and salinity. A minimum value of 1.6/h is imposed on $k_{a,20}$ (9).

Moreover, an adsorption–desorption interaction occurs between dissolved inorganic phosphorus and suspended particulate matter in the water column. The subsequent settling of the suspended solids together with the sorbed inorganic phosphorus can act as a significant loss mechanism in the water column and is a source of phosphorus to the sediment.

Table 2 lists typical kinetic parameters used in water quality models from the literature (9,19,20). The growth and proliferation of phytoplankton is a result of the utilization and conversion of inorganic nutrients into organic plant material through photosynthesis. μ_A depends on three principal components: (1) temperature—T, (2) solar radiation—I, and (3) nutrients—N_{nutri}. Multiplicative effects are assumed, that is, $\mu_A = f(T, I, N_{\text{nutri}}) = g(T)g(I)g(N_{\text{nutri}})$.

The relationship with temperature is as follows:

$$g(T)_k = \mu_{A,k}^{\max} = \mu_{A,20}^{\max}\theta^{T_k - 20} \tag{9}$$

where $\mu_{A,20}^{\max}$ = maximum growth rate of phytoplankton at 20 °C under optimal light and nutrient conditions in the kth layer. The light-limitation function, represented by $g(I)$ over a given water depth of layer (h_k), is approximately integrated as follows (19):

$$g(I)_k = \frac{1}{h_k}\int_{h_u}^{h_u+h_k} \frac{I}{I_s} e^{(1-I/I_s)} dz = \frac{2.718}{\gamma_k h_k}[e^{-\alpha_2} - e^{-\alpha_1}]$$

$$\alpha_2 = \alpha_1 e^{-\gamma_k h_k}, \alpha_0 = \frac{I_0}{I_s}, \alpha_1 = \alpha_0 e^{-\gamma_u h_u} \tag{10}$$

$$\frac{dI}{dz} = -\gamma_k I, \quad I = I_0 \text{ @ } z = 0$$

and z-axis is upward vertically

where I, I_0, and I_s = light intensity at depth z, incoming solar radiation intensity just below the surface, and saturating light intensity (the optimum light intensity at which the relative photosynthesis is a maximum), respectively; and γ_κ = overall extinction coefficient in m^{-1}. Phytoplankton may adjust its chlorophyll composition to adapt to the changes in solar radiation. Therefore, I_s is determined to be the weighted average of the light intensity for the previous 3 days as follows: $I_s = 0.7I_1 + 0.2I_2 + 0.1I_3$, where $I_i = 0.5\times$ [daily average visible light intensity beneath the surface] i days earlier (20). The annual average of daily solar radiation intensity is used in the modeling.

The minimum value of nutrient limitations computed by Michaelis–Menten-type expression is chosen for

Table 2. Typical Kinetic Parameters Used in Water Quality Models

Parameters	Description	Value (and θ)
$\mu_A^{\max}$	Maximum phytoplankton growth rate	2.10 (1.066)
K_{mNN}	Half-saturation constant for nitrogen uptake	15.0
K_{mNP}	Half-saturation constant for phosphorus uptake	1.50
r_A	Endogenous respiration rate of phytoplankton	0.05 (1.08)
M_A	Nonpredatory mortality rate of phytoplankton	0.10
C_g	Grazing (filtering) rate of zooplankton	0.30 (1.066)
r_Z	Endogenous respiration rate of zooplankton	0.02 (1.045)
M_Z	Mortality rate of zooplankton	0.10
α_{12}	Assimilated carbon per unit algae mass ingested (average CCHL)	112.5
α_2	Zooplankton assimilation efficiency	0.6
K_Z	Half-maximum-efficiency food level for zooplankton filtering	12.0
k_{34}	Conversion efficiency of organic nitrogen to ammonia nitrogen	0.05 (1.08)
α_{13}	Stoichiometric ratio of cell nitrogen to algae chlorophyll-a	10.0
α_{23}	Stoichiometric ratio of cell nitrogen to zooplankton carbon	α_{13}/CCHL
k_N	Nitrification rate of ammonia to nitrate via nitrite nitrogen	0.05 (1.08)
k_{55}	Denitrification rate	0.09 (1.045)
k_{67}	Conversion efficiency of organic phosphorus to inorganic one	0.03 (1.08)
α_{16}	Stoichiometric ratio of phosphorus to algae chlorophyll-a	1.00
α_{26}	Stoichiometric ratio of phosphorus to zooplankton carbon	α_{16}/CCHL
k_C	Deoxygenation or decay rate of carbonaceous BOD	0.23 (1.047)
BOD_{u-5}	Ratio of the ultimate to 5-day carbonaceous BOD	1.54
α_{18}	Stoichiometric ratio of phytoplankton to organic carbon	CCHL
α_{28}	Stoichiometric ratio of zooplankton to organic carbon	1.00
α_{ON}	Oxygen to nitrogen ratio	2.28
α_{19}	Stoichiometric ratio of phytoplankton to oxygen	2.67CCHL
K_{mpc}	Half saturation constant for phytoplankton affects mineralization	10.0
K_{NIT}	Half saturation DO constant for oxygen limitation of nitrification	2.0
K_{NO}	Half-max. DO constant for oxygen limitation of denitrification	0.1
K_{BOD}	Half saturation DO constant for CBOD deoxygenation.	0.5
w_{SA}	Settling rate of phytoplankton	1.10
w_{SN}	Settling rate of particulate organic nitrogen	0.30

$g(N_{\text{nutri}})$. Thus, the full expression for layer-averaged phytoplankton growth rate is expressed as follows (19):

$$\mu_{A,k} = \mu_{A,k}^{\max} g(I)_k \min \times \left\{ \frac{\text{NH}_{4,k} + \text{NO}_{23,k}}{K_{\text{mNN}} + \text{NH}_{4,k} + \text{NO}_{23,k}}, \frac{\text{PO}_{4,k} \cdot f_{\text{DPO}_4}}{K_{\text{mNP}} + \text{PO}_{4,k} \cdot f_{\text{DPO}_4}} \right\} \quad (11)$$

in which $\text{NH}_{4,k}$ (μg N/L), $\text{NO}_{23,k}$ (μg N/L), and $\text{PO}_{4,k}$ (μg P/L) = concentrations of NH_4-N, $\text{NO}_2 + \text{NO}_3$-N, and PO_4-P in kth layer, respectively; and K_{mNN} (μg N/L) and K_{mNP}(μg P/L) = Michaelis constants for nitrogen and phosphorus uptake by algae, respectively.

Depending on the history of the algal cells, the CCHL ratio (mg C/mg Chl-a) is affected by light intensity, temperature, and nutrient availability as follows (21):

$$\text{CCHL} = \frac{\alpha I_s}{2.718 \mu_{A,k}^{\max}} \quad (12)$$

where α can be obtained from laboratory results.

A saturating recycle equation is used for hydrolysis and bacterial decomposition of organic nitrogen to ammonia (k_{34}) and the mineralization of organic phosphorus to inorganic phosphorus (k_{67}):

$$k_{34} = k_{34,20} \theta^{T_k - 20} \frac{A_k}{K_{\text{mpc}} + A_k}$$

$$k_{67} = k_{67,20} \theta^{T_k - 20} \frac{A_k}{K_{\text{mpc}} + A_k} \quad (13)$$

This expression is a compromise between the conventional first-order temperature-corrected mechanism and a second-order recycle mechanism, with the recycle rate being directly proportional to the amount of phytoplankton biomass. It approaches second-order dependency at low phytoplankton concentrations ($A \ll K_{\text{mpc}}$), where K_{mpc} is the half-saturation constant for recycle, and it approaches first-order recycle when the phytoplankton greatly exceeds the half-saturation constant. This mechanism slows the recycle rate if the algal population is small, but it does not permit the rate to increase continuously as phytoplankton increases.

The following processes, namely, nitrification (k_N) of NH_4 to NO_3 via NO_2-N, denitrification (k_{55}) of NO_3-N, and deoxygenation (k_C) of CBOD, are considered temperature and oxygen dependent:

$$k_N = k_{N,20} \theta^{T_k - 20} \frac{\text{DO}_k}{K_{\text{NIT}} + \text{DO}_k}$$

$$k_{55} = k_{55,20} \theta^{T_k - 20} \frac{K_{\text{NO}}}{K_{\text{NO}} + \text{DO}_k}$$

$$k_C = k_{C,20} \theta^{T_k - 20} \frac{\text{DO}_k}{K_{\text{BOD}} + \text{DO}_k} \quad (14)$$

where $\text{DO}_k(\text{mg O}_2/\text{L})$ = concentration of DO; K_{NIT} and $K_{\text{BOD}}(\text{mg O}_2/\text{L})$ = half-saturation DO constants for oxygen limitation of nitrification and of organic carbon stabilization; and $K_{\text{NO}}(\text{mg O}_2/\text{L})$ = half-maximum DO constant for denitrification.

Primary production by phytoplankton in surface waters is a major source of labile organic carbon to coastal sediment. Particles from the euphotic zone sink to the sediment–water interface, where benthic organisms rapidly degrade the labile organic compounds present in the settled materials (22). Sediment algal carbon is expressed as follows (23):

$$\frac{dC_{\text{sedi}}}{dt} = \alpha_{\text{sediCA}} w_{\text{SA}} A_l - \frac{v_{\text{sedi}}}{h_{\text{sedi}}} C_{\text{sedi}} - k_{\text{Csedi}} C_{\text{sedi}} \quad (15)$$

where $C_{\text{sedi}}(\text{g C/m}^2)$ = sediment algal carbon; α_{sediCA} (g C/g Chl-a) = sediment algal carbon per unit algal mass settled; $v_{\text{sedi}}(\text{m/day})$ = sediment accumulation rate; $h_{\text{sedi}}(m)$ = thickness of the sediment layer; $k_{\text{Csedi}}(\text{day}^{-1})$ = oxidation rate of sediment algal carbon.

SUMMARY

Temporal simulation of phytoplankton growth in a coastal water system is complicated, resource demanding, and computationally intensive because several chemical, biological, and biochemical processes interact, and the reaction rates and external inputs vary with time. Furthermore, the flow and associated circulation are also functions of time, with time scales ranging from minutes to months, even years. Flow and pollutant transport in a natural water body commonly interact with density stratification that resulted from salt water intrusion and solar radiation, which may be described as a layered system. Nowadays, with the advances in computing resources, water quality models are capable of addressing this problem efficiently and flexibly because of their grid system and good numerical performance. In the near future, more work should be undertaken on seeking more accurate pollution source data, including the point/nonpoint sources and atmospheric loads carried directly to the water surface by rain and wind, and for more information on the sediment release of nutrients to obtain much better simulations by the present model.

BIBLIOGRAPHY

1. van Eden, A. (1993). Marine pollution and the environment of the North Sea. *Eur. Water Pollut. Contr.* **3**(5): 16–24.
2. Fera, P. (1993). Marine eutrophication along the Brittany coasts—origin and evolution. *Eur. Water Pollut. Contr.* **3**(4): 26–32.
3. Bothwell, M.L. (1992). Eutrophication of rivers by nutrients in treated kraft pulp mill effluent. *Water Pollut. Res. J. Canada* **27**(3): 447–472.
4. Cerco, C.F. and Cole, T. (1993). Three-dimensional eutrophication model of Chesapeake Bay. *J. Environ. Eng. ASCE* **119**(6): 1006–1025.
5. Chau, K.W. and Jiang, Y.W. (2002). Three-dimensional pollutant transport model for the Pearl River estuary. *Water Res.* **36**(8): 2029–2039.
6. Chau, K.W. and Sin, Y.S. (1992). Correlation of water quality parameters in Tolo Harbour, Hong Kong. *Water Sci. Technol.* **26**(9-11): 2555–2558.
7. Sin, Y.S. and Chau, K.W. (1992). Eutrophication studies on Tolo Harbour, Hong Kong. *Water Sci. Technol.* **26**(9–11): 2551–2554.

8. Lung, W.S., Martin, J.L., and McCutcheon, S.C. (1993). Eutrophication analysis of embayments in Prince William Sound, Alaska. *J. Environ. Eng. ASCE* **119**(5): 811–824.
9. Ambrose, R.B., Wool, T.A., Connolly, J.P., and Schanz, R.W. (1988). *WASP4, a hydrodynamic and water quality model–model theory, user's manual, and programmer's guide.* U.S. Environmental Protection Agency Report, EPA/600/3-87/039.
10. Orlob, G.T. (1982). *Mathematical Modeling of Water Quality: Streams, Lakes, and Reservoirs.* John Wiley & Sons, New York.
11. Lung, W.S. (1993). *Water Quality Modeling, Vol.III: Application to Estuaries*. CRC Press, Boca Raton, FL.
12. Chau, K.W. and Jiang, Y.W. (2003). Simulation of transboundary pollutant transport action in the Pearl River delta. *Chemosphere* **52**(9): 1615–1621.
13. Chau, K.W. and Jin, H.S. (1998). Eutrophication model for a coastal bay in Hong Kong. *J. Environ. Eng. ASCE* **124**(7): 628–638.
14. Thompson, J.F., Warsi, Z.U.A., and Mastin, C.W. (1985). *Numerical Grid Generation*. Elsevier Science Publishing, New York.
15. Chau, K.W. and Jin, H.S. (1995). Numerical solution of two-layer, two-dimensional tidal flow in a boundary fitted orthogonal curvilinear coordinate system. *Int. J. Numer. Methods Fluids* **21**(11): 1087–1107.
16. Chau, K.W. and Jiang, Y.W. (2001). 3d numerical model for Pearl River estuary. *J. Hydraul. Eng. ASCE* **127**(1): 72–82.
17. Fischer, H.B., List, E.J., Koh, R.C.Y., Imberger, J., and Brooks, N.H. (1979). *Mixing in Inland and Coastal Waters.* Academic Press, New York.
18. Rodi, W. (1980). *Turbulence Models and Their Application in Hydraulics, State-of-the-Art.* IAHR Publications, Delft, The Netherlands.
19. Thomann, R.V. and Mueller, J.A. (1987). *Principles of Surface Water Quality Modeling and Control*. Harper & Row Publishers, New York.
20. Lee, J.H.W., Wu, R.S.S., Cheung, Y.K., and Wong, P.P.S. (1991). Forecasting of dissolved oxygen in marine fish culture zone *J. Environ. Eng. ASCE* **117**(6): 816–833.
21. Jin, H.S., Egashira, S., and Chau, K.W. (1998). Carbon to chlorophyll-*a* ratio in modeling long-term eutrophication phenomena. *Water Sci. Technol.* **38**(11): 227–235.
22. Di Toro, D.M., Paquin, P.R., Subburamu, K., and Gruber, D.A. (1990). Sediment oxygen demand model: methane and ammonia oxidation. *J. Environ. Eng. ASCE* **116**(5): 945–986.
23. Chau, K.W. (2002). Field measurements of SOD and sediment nutrient fluxes in a land-locked embayment in Hong Kong. *Adv. Environ. Res.* **6**(2): 135–142.

ENVIRONMENTAL APPLICATIONS WITH SUBMITOCHONDRIAL PARTICLES

Diana J. Oakes
John K. Pollak
University of Sydney
Lidcombe, Australia

It is now recognized that the supply of sufficient clean and safe water for drinking and growing food as well as for supplying adequate volumes of water for the maintenance of ecosystems will be one of the most significant problems facing the population of the twenty-first century. Shortages of clean water have already a direct effect on human health, causing many deaths as well as severely damaging plant and animal life and thus affecting the ecology as well as the production of agricultural crops. Hence, it is important that the quality of water is routinely evaluated.

The classic toxicological methods that have been developed and used by scientists over many years for risk assessments often evaluate the outcomes of exposures to single chemicals. However, in the present environment, in which exposures to mixtures of chemicals have been significantly increased, it is inappropriate to only evaluate the toxic effects of individual chemicals or toxins. It is now widely recognized that the toxic effects caused by exposures to chemical mixtures may even be caused by extremely low concentrations of individual chemicals because of the additive or even synergistic actions of other chemicals in a water sample. Thus, it was demonstrated in 105 extracts of sediments collected from aquatic environments that many of the extracts with low or barely detectable levels of priority pollutants were highly toxic when evaluated with the Microtox bioassay (1). Jacobs et al. (1) suggested that "... chemical data alone provides no direct indication of the potential effects that contaminant mixtures may have on the aquatic biota. Toxicity bio-assays furnish reasonable estimates concerning the biological impacts of sediment contaminants; both chemical and biological analyses should be performed." Although some of these toxic samples did not contain any significant concentrations of priority pollutants, it is likely that these extracts contained unidentified and possibly unregulated pollutants at low concentrations.

Even nontoxic chemicals have the potential to increase the toxicity of other chemicals in a mixture. An example of such an effect is the increased toxicity of phenytoin by the presence of lactose in the mixture (2). In this context, it is questionable whether the guidelines for permissible concentrations of individual chemicals (such as ADI—acceptable daily intake; MRL—minimal risk level; NOAEL—no observable adverse effect level) provide meaningful safety levels either for the ecology or for human health, because these parameters focus only on individual chemicals and do not consider the potential effects of other chemicals on the overall toxicity of a chemical mixture. It is, therefore, important that we use bioassays for the evaluation of the toxic effects of aquatic samples.

For the past 24 years, the Organization for Economic Co-operation and Development (OECD) has provided standardized test protocols to evaluate aquatic toxicity with laboratory-based representative single-species toxicity tests, such as fish, Daphnia spp., and algae (3). These bioassays provide useful data of the overall toxicities of chemical mixtures that are present in aqueous samples.

An example of a subcellular bioassay is the submitochondrial particle (SMP) test. With this test, we use the cellular organelle, the mitochondrion, which is present in all eukaryotic (nucleated) cells. Mitochondria are effectively the powerhouses of the cell, providing over 90% of all energy produced in all eukaryotic organisms. Mitochondria catalyze reactions in which molecules with a low energy content are oxidized (lose electrons) to give rise to "energy-rich" molecules, such as adenosine-triphosphate (ATP). ATP is required for synthetic metabolic processes or for mechanical or heat energy. The metabolic activities of mitochondria are so fundamental to aerobic life that their functional organization, their proteins and enzymes, are very similar in all aerobic organisms (from amoeba, plants, or even fungi to humans). Hence, the evaluation of cellular toxicities by mitochondria from one species is likely to apply to other species. Stable SMPs are produced by the disruption of the internal membranes of isolated bovine heart mitochondria (4,5). Once prepared, the SMPs are stable for at least 2 years when kept at −80 °C. SMPs are now commercially available in the United States from Harvard Bioscience, Inc., Madison, WI. SMPs consist of inverted, vesicular portions of the inner mitochondrial membranes that can perform the integrated enzymatic functions of electron transport and oxidative phosphorylation (6).

The test measures inhibitory effects that chemicals present in contaminated water may exert on SMP enzyme activity. Two different assay procedures have been devised for SMPs (7,8) as illustrated in Fig. 1. Figure 1a illustrates the complexes of the inner mitochondrial membrane that are involved in the so-called forward electron transfer reaction (ETR), whereas Fig. 1b illustrates the reverse transfer reaction (RET) in which access to complex III and complex IV is blocked with the inhibitor antimycin A.

The ETR assay measures the flow of electrons from NADH via complexes I, II, III, and IV to molecular oxygen. Measurements of the oxidation are carried out in six 1-mL cuvettes. Five cuvettes receive different concentrations of the samples that are tested, whereas one cuvette serves as the control. The extent of the inhibition of the oxidation of NADH provides an indication of the toxic effects of the aqueous sample that is being evaluated. The assay buffer consisted of 0.25 M sucrose, 20 mM Hepes buffer (pH 7.4), 6 mM $MgCl_2$, and 200 nM NADH (25 μL). The decrease in the absorption at 340 nm measures the oxidation of NADH to NAD with a Beckman DU650 spectrophotometer (Fullerton, CA), with the automatic six-cell holder that is maintained at 30 °C.

In the RET assay, succinate is added as a substrate and because access to complexes III and IV is blocked (see Fig. 1b), the NAD that has been added is reduced to NADH by the SMPs. Hence, the RET assay measures the reduction of NAD to NADH leading to the increased absorption at 340 nm. The assay buffer consisted of 0.18 M sucrose, 50 mM Tris buffer (pH 7.5), 0.1 M sodium succinate, 0.7 μg/mL antimycin A, 30 mM $MgCl_2$, and 1 mM NAD. The NAD reduction was measured at 340 nm with a Beckman DU650 spectrophotometer, with an automatic six-cell holder maintained at 30 °C. This test evaluates the inhibitory effect exerted by chemicals in the aqueous samples on the oxidation of succinate that leads to the reduction of NAD to NADH. For further details of the SMP assay protocols see Oakes and Pollak (9).

The SMP test has evaluated the toxicities of 113 different chemicals (10,11). SMP test procedures have also been validated in field trials through the participation in several ongoing U.S. Environmental Protection Agency

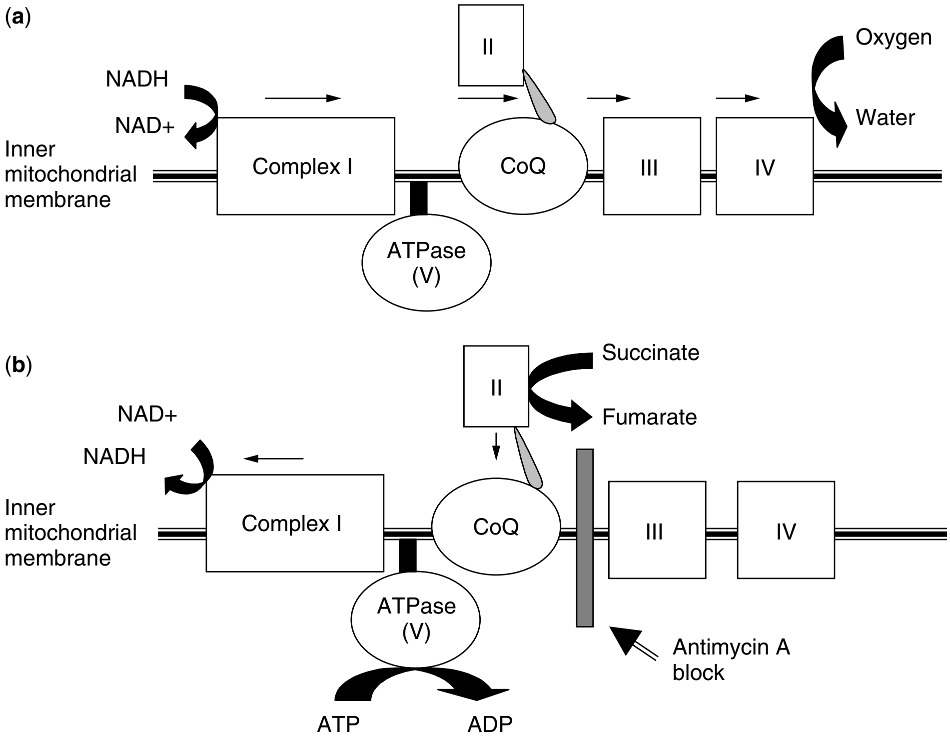

Figure 1. (**a**) Diagram of the ETR adapted from Ref. 7. (**b**) Diagram of the RET adapted from Ref. 7.

Geological Surveys and by the Wisconsin Department of Natural Resources Bio-Monitoring Projects, as well as in studies in the Illinois and Minnesota River Basins (11). Other studies have shown that the results obtained by the SMP test correlate well with the results obtained with *in vitro* and *in vivo* bioassays, such as Microtox, mammalian cell cultures, nematode, Daphnia spp., and fish assays or sea urchin fertilization and larval development tests (6,10,12–15). A further example of the useful information that the SMP test can provide was established when the toxicity of the herbicide formulation Tordon 75D was investigated (9). The herbicide Tordon 75D consists of a mixture of the triisopropanolamine salts of 2,4-dichlorophenoxy acetic acid (2,4-D) and 4-amino-3,5,6-trichloropicolinic acid (Picloram) as well as other components, such as the solvent triisopropanolamine and diethyleneglycol monoethyl ether, a silicone defoamer, and a proprietary surfactant (polyglycol 26-2). Using both the ETR and the RET assays and testing the complete formulation as well as the individual components, researchers have shown that although the proprietary surfactant by itself severely inhibited the oxidative functions of the SMPs, the mixture of only the active components had no significant effects on the oxidative functions of the SMPs. These results suggested that the surfactant damaged the inner mitochondrial membrane and was the primary component of Tordon 75D that inhibited the oxidative functions of the SMPs. A later study assessed the potential toxicity of three related herbicide formulations containing the ester derivatives of 2,4,5-T and 2,4-D. It was shown that in these formulations, the “other components” such as diesel oil and surfactants contributed significantly (approximately 50%) to the overall toxicity in an additive manner (16). More recently, the SMP test evaluated the toxicities of leachates that emanated from the Sydney Olympic Site. The leachates were also evaluated with several other bioassays as well as by chemical analyses. The SMP test, although less sensitive than other bioassays, such as the sea urchin fertilization and development tests, provided useful information as a rapid prescreening tool (15). Of the two SMP assays, only the RET was suited for evaluation of the leachate toxicity because it is known that samples containing significant levels of divalent cations can activate the SMP activity when measured with the ETR assay, effectively masking any inhibitory effect (6,15).

Overall, these studies have demonstrated that the SMP test is a useful prescreening tool for evaluating the toxic effects of aqueous media or chemical formulations. The SMP test is less sensitive than the Microtox assay to changes in pH and to the presence of solvents; hence, it can evaluate a greater range of aqueous samples. The major limitation of the SMP test as well as other cellular and subcellular *in vitro* assays is that they cannot account for pharmacokinetic and pharmacodynamic processes that may affect the actions of chemicals *in vivo*, nor can the SMP test evaluate the action of chemicals on specific receptors or enzymes (e.g., acetyl choline esterase) other than mitochondrial enzymes.

BIBLIOGRAPHY

1. Jacobs, M.W. et al. (1993). Comparison of sediment extract Microtox toxicity with semi-volatile organic priority pollutant concentrations. *Arch. Environ. Contam. Toxicol.* **24**: 461–468.
2. Tyrer, J.H., Eadie, M.J., Sutherland, J.M., and Hooper, W.D. (1970). Outbreak of anticonvulsant intoxication in an Australian city. *Br. Med. J.* **4**: 271–273.
3. Organization for Economic Co-operation and Development (OECD). (1981). *Guidelines for Testing of Chemicals*. OECD, Paris.
4. Blondin, G.A., Knobeloch, L.M., Read, H.W., and Harkin, J.M. (1987). Mammalian mitochondria as *in vitro* monitors of water quality. *Bull. Environ. Contam. Toxicol.* **38**: 467–474.
5. Knobeloch, L.M., Blondin, G.A., and Harkin, J.M. (1990). Use of submitochondrial particles for prediction of chemical toxicity in man. *Bull. Environ. Contam. Toxicol.* **44**: 661–668.
6. Read, H.W., Harkin, J.M., and Gustavson, K.E. (1997). Environmental applications with sub-mitochondrial particles. In: Wells, P.G., Lee, K., and Blaise C. (Eds.). *Microscale Testing in Aquatic Toxicology—Advances, Techniques and Practice*. CRC Press, Boca Raton, FL.
7. Knobeloch, L.M., Blondin, G.A., and Harkin, J.M. (1994). A rapid bio-assay for toxicity assessment of chemicals: forward electron transport assay. *Environ. Toxicol. Water Quality* **9**: 79–82.
8. Knobeloch, L.M., Blondin, G.A., and Harkin, J.M. (1994). A rapid bio-assay for toxicity assessment of chemicals: reverse electron transport assay. *Environ. Toxicol. Water Quality* **9**: 231–234.
9. Oakes, D.J. and Pollak, J.K. (1999). Effects of a herbicide formulation, Tordon 75D, and its individual components on the oxidative functions of mitochondria. *Toxicology* **136**: 41–52.
10. Knobeloch, L.M., Blondin, G.A., Read, H.W., and Harkin, J.M. (1990). Assessment of chemical toxicity using mammalian mitochondrial electron transport particles. *Arch. Environ. Contam. Toxicol.* **19**: 828–835.
11. Harkin, G.M., Blondin, G.A., Knobeloch, L.M., and Read, H.W. (1991). Mitochondrial bio-assay for toxic substances in water. *Gov. Rep. Announce. Index (U.S.)* **91**(16). Abstr. No. 143,481 (only seen in abstract form in *Chem. Abstr.* **116**: 417).
12. Blondin, G.A., Knobeloch, L.M., Read, H.W., and Harkin, J.M. (1989). An *in vitro* Submitochondrial Bio-assay for predicting acute toxicity in fish. In: *Aquatic Toxicology and Environmental Fate*. G.W. Suter and M.A. Lewis (Eds.). American Society for Testing and Materials, Vol. 11, pp. 551–563.
13. Degli-Esposti, D.M., Ngo, A., and Myers, M.A. (1996). Inhibition of mitochondrial complexI may account for IDDM induced by intoxication with the rodenticide Vacor. *Diabetes* **45**: 1531–1534.
14. Argese, E. et al. (1998). Comparison of *in vitro* submitochondrial particle and Microtox assays for determining the toxicity of organotin compounds. *Environ. Toxicol. Chem.* **17**: 1005–1012.
15. Byrne, M., Pollak, J., Oakes, D., and Laginestra, E. (2003). Comparison of the submitochondrial particle test, Microtox and sea urchin fertilization and development tests: parallel assays with leachates. *Aust. J. Ecotox.* **9**: 19–28.
16. Oakes, D.J. and Pollak, J.K. (2000). The *in vitro* evaluation of the toxicities of three related herbicide formulations containing ester derivatives of 2,4,5-T and 2,4-D using submitochondrial particles. *Toxicology* **151**: 1–9.

INTEREST IN THE USE OF AN ELECTRONIC NOSE FOR FIELD MONITORING OF ODORS IN THE ENVIRONMENT

Jacques Nicolas
University of Liege
Arlon, Belgium

INTRODUCTION

The increasing number of complaints related to malodors generated by agricultural, landfill, and wastewater treatment facilities give rise to a growing interest in the measurement of olfactive annoyance. Instrument manufacturers have sought to provide suitable environmental monitoring solutions. Because the need for such devices is generally driven by legislation, compliance with the standardized validation of techniques is often mandatory. So, current standard methods remain the references to measure odor annoyance.

Two procedures are generally proposed: either olfactometry, based on odor assessment by human panels, or chemical analysis by chromatographic techniques. However, neither meets the requirements of on-line monitoring for environmental applications.

New emerging technologies such as electronic noses, based on nonspecific sensor arrays, may offer objective and on-line instruments for assessing environment odors. But, despite the recent appearance of commercial devices that can be implemented in the field, a number of limitations can be highlighted, both from the technology itself and its application to environmental studies.

BASIS OF THE CHIEF ODOR MEASUREMENT METHODS

Olfactometry is the scientific measurement of odor concentration using a system of sampling and a methodology regulated to European (EN 13 725) (1) and American (ASTM E679-91) (2) standards. The odor concentration of a gaseous sample is determined by presentation to a panel of observers, who have known acuity to odor, in varying dilutions. The procedure aims at determining the dilution at which only half the panel can detect the odor. The odor concentration is then the number of dilutions required for the sample to reach this threshold. It is expressed in multiples of one odor unit per m^3 (ou/m^3). This technique provides directly comparable data for different odor types. Although this technique gives the right human sense evaluation and is now based on standard methodology, it remains strongly influenced by subjectivity and is time-consuming, labor-intensive, and expensive. Olfactometry laboratories are often remote from the odor source: it is obviously not appropriate for real-time and continuous operation on-site.

Alternatively, chemical analysis of the odorous mixture by gas chromatographic techniques (such as GC-MS) provide an accurate concentration of specific compounds in the sample (3). Some instruments can be used on-site and for continuous assessment, but analytic methods never provide the global olfactive perception, as most environmental odors are complex mixtures of compounds. Moreover, they do not take into account the interactions between different odorants and interfering background substances, which may lead to synergistic or antagonistic effects.

The electronic nose is a more recent technique based upon a fixed array of nonspecific gas sensors, each of which responds differently to an odorant and the global pattern of responses that is characteristic of that odorant (4).

The heart of the electronic nose is an array of chemical sensors (metal oxide semiconductors, conductive polymers, quartz microbalances). When exposed to a gaseous atmosphere, these individual sensors provide individual signals. Subsequently, a signal pattern is deduced by analyzing these signals by suitable statistical and mathematical methods. The diagnosis is thus based on the “fingerprint” of the gas mixture classified by pattern recognition techniques (artificial neural networks, discriminant function analysis). Before applying such an instrument on a real atmosphere for on-line recognition of odors in the environment, it must be trained with typical gas mixtures for this environment. After the training, the system can identify one of the learned odors by establishing similarities between the actually observed pattern of signals and those previously observed. Hence, the electronic nose is an analytical instrument that can characterize an odor without reference to its chemical composition. The signal provided is the odor, considered as a whole. That makes this method particularly attractive and shows its great potential.

STATE OF THE ART

Reported developments in complex odor analysis are focused chiefly on quality assessment in the food, drink, and perfume industries. So far, few attempts have been made to characterize and monitor complex odor in the environment. Although original work carried out with laboratory-based systems and more recent field investigations have shown promising results, the number of trials carried out under realistic environmental conditions is relatively limited.

Some papers review the current status of sensor array technology and discuss its potential application to the assessment of olfactive annoyance in the environment by referring to exhaustive literature surveys (5–8).

Early investigations include the analysis of single substances, but some works deal with more complex odors. In a review of odor measurement techniques for sewage treatment works, Gostelow et al. (9) listed some examples of electronic nose application to environmental odor problems. However, the use of electronic noses to monitor water quality in the field remains a virtually unexplored domain with only few applications reported so far.

TYPICAL USER NEEDS

In the field of environmental monitoring and more particularly of water quality, the typical needs of the final user of electronic nose require both qualitative (classification) and quantitative (regression) approaches.

From a qualitative point of view, different annoyance odors must be discriminated, and the identity of unknown samples must be predicted using a previously calibrated learning model. A range of data processing techniques may be used to analyze sensor array data, but due to the large number of variables (i.e., number of sensors) and samples, pattern recognition techniques (such as multivariate statistics and artificial neural networks) are employed to reduce the dimensionality of the sensor array data. The relationships between the samples can then be compared and correlated using simple scatter plots. The choice of analysis technique depends on the amount and nature of information available and the type of information required from the analysis.

Figure 1 could refer to a typical result of a discriminant function analysis (DFA) of observations made at a landfill site by an electronic nose. Three types of odor tonalities can be perceived: the fresh refuse, the odor of landfill gas, and one of the leachates. The scatter plot of the observations in the plane of the two calculated roots shows clustering into the three expected groups. The model so calibrated by DFA can further be used by the landfill manager to classify a new observation. For example, the emergence of an odor identified as "leachates" can let us suspect a problem in the leachate treatment system.

This ability of an electronic nose to discriminate rapidly between slight variations in complex mixtures makes the technique ideal for on-line process diagnostics and screening across a wide range of applications.

Quantitative approaches imply regression procedures by correlating sensor responses with some quantitative parameter obtained either by another instrument or by human perception. In some cases, reasonable fits are obtained when the sensor responses are compared with odor concentration resulting from olfactometry measurements. However, this approach is impeded by the subjectivity of the olfactive evaluation, the nonlinear relation between the perceived intensity of an odor and its concentration, as well as by adaptation phenomena (perception) as a result of extended exposure to that odor (8). This complicates the elaboration of quantitative models. Moreover, the model so calibrated is often site specific, and its validity over long periods of time must be demonstrated before sensor array systems can be used for quantitative analysis.

In the field of water quality, the overall electronic nose output is often compared with classical water characterization variables, such as volatile suspended solids (10) or global organic parameters (biochemical oxygen demand and total organic carbon) (11). Studies generally conclude that a number of different wastewater quality relationships could be formulated from the electronic nose analysis of a sewage liquid. In all cases, the principal advantages put forward by the authors to justify an electronic nose with respect to traditional analytical methods are the rapidity of the response and the noninvasive nature of the measurement.

The model obtained by regression could be used as an on-line monitoring tool for a process. For example, Fig. 2 highlights a "stress" episode in a composting process, caused by the absence of aeration of the compost pile. A specific indicator is constructed by relating the sensor signals to the four families of compounds emitted during that "stress" phase (nitrogen compounds, carboxylic acids, ammonia, and chlorinated compounds). It is constructed by a canonical correlation analysis. The figure shows the temporal evolution of that combination of the sensor signals applied to the whole data set obtained by continuous signal monitoring for 11 days. The root value exhibits a peak during the "stress" period. Of course, such a global odor indicator could be calculated for other environmental processes, such as wastewater treatment, and it could be monitored as a global process control variable.

More generally, the purposes of continuous, *in situ* monitoring of odorous emissions by an electronic nose could be

— to predict the rise of malodor in the background before it becomes an annoyance for the surroundings;

— to use the odor as a process variable, aiming at a better understanding of the odor release and

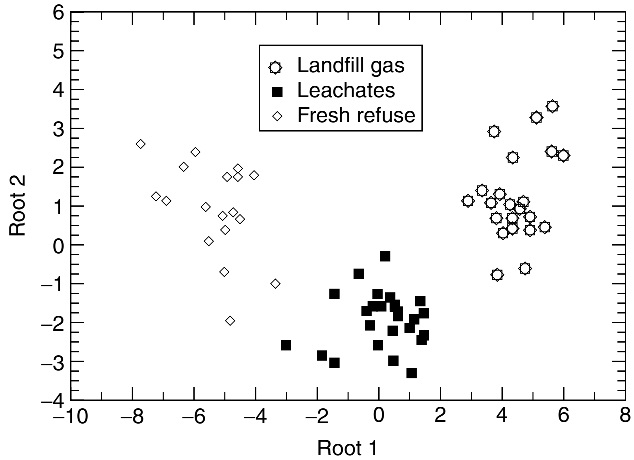

Figure 1. Possible scatter plot of electronic nose observations at a landfill site in the plane of the two first roots of a discriminant function analysis.

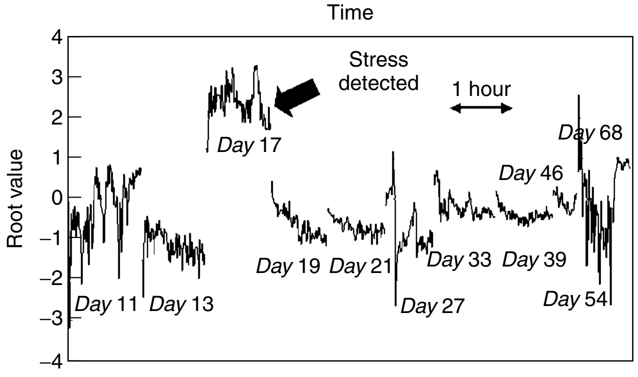

Figure 2. Global odor signal obtained from the response of a sensor array to detect a stress episode in a composting process.

relating this emission to the process phase, or the problem that caused the emission;
- — to control odor abatement techniques in real time, such as atomization of neutralizing agents.

Simple sensory early warning devices that can detect sudden changes in a process stream obviously present great interest.

APPLICATIONS RELATED TO WATER QUALITY

The use of electronic noses to monitor water quality in the field may involve the agricultural and industrial effluents allowed to seep into groundwater or to flow into streams or rivers (6). The electronic nose can be used in these applications by collecting samples of the effluent. Boreholes can also be employed to collect samples to test groundwater contamination.

A very promising area is the detection of hydrophobic and highly volatile organic compounds (VOCs) in water systems (8). The recent developments in on-line detection of oils and petroleum hydrocarbons (12) have arisen from the recent advances in sensor technology as well as from progress in computing and pattern recognition. Studies aim also at real-time control of wastewater treatment plants (13) and at the quality tests of potable water (14).

A specific obstacle when dealing with liquid samples is to generate or collect a headspace gas that can be reliably and safely measured by the sensor array. Because parameters such as flow, temperature, or suspended solids constantly change, the major difficulty lies in drawing a headspace sample that is representative of the liquid phase (8). But several applications involve also monitoring odors emitted in the air from wastewater (9). Odor abatement and control are major issues facing municipal sewage treatment facilities.

SPECIFIC PROBLEMS FOR ENVIRONMENTAL APPLICATIONS

Much of the original work on environmental applications has used sensor arrays in laboratory-based conditions. However, to understand the effects of environmental parameters on localized odor pollution, it is necessary to translate these laboratory-based experiences into formats that can be applied to measurements under variable conditions (15). Although new sensor materials and designs are continuously being reported, the major limitation of currently available sensors remains their sensitivity to changes in temperature, humidity, and flow rate.

The solution to this obstacle is generally to work under fixed experimental conditions, that is, to incorporate the effects of variations in ambient temperature and humidity into the design of instruments by using sample pretreatment systems. That was demonstrated for headspace pretreatment in on-line wastewater monitoring (13). However, when considering the practical application of a sensor array, this can make the overall instrument more complex and expensive and can also affect its portability or limit sample throughput.

A second approach is to measure these parameters and calibrate the sensors under varying humidity levels to compensate for changes in subsequent data analysis (16). Yet, the method has still to be validated for real-life environmental applications where pollutants must be analyzed rapidly under ever changing background conditions and in the presence of interfering compounds.

Alternatively, the development of a classification model under a wide range of operating and meteorologic conditions makes those conditions “neutral” for odor recognition (17). It is an effective but sometimes fastidious solution.

Sensor drift is another serious impairment of chemical sensors. They alter over time and so have poor repeatability because they produce different responses for the same odor. That is particularly troublesome for electronic noses for which the sensor signals can drift during the learning phase, requiring repeated exposure to the same type of odor.

The consequences of sensor drift on classification results exclude the possibility of further odor identification (18). Though drift correction algorithms could be applied for each individual sensor, the usual way to counteract the global drift of a sensor array is to correct the whole pattern, using multivariate methods. First, the main direction of the drift is determined in the first component space of a multivariate method, such as principal component analysis (PCA). The drift component can then be removed from the sample gas data, thus correcting the final score plot of PCA (19).

Moreover, the combination of multiple sensors, cross-selectivity, and pattern recognition analysis makes the usual techniques of instrument calibration difficult to apply to electronic noses. As a result of this issue there are, to date, no international standards that directly refer to electronic nose measurement.

An additional problem occurs when the concentration of volatiles is low, which is often the case in the environment: the limit of detection of the sensors and the limits of recognition of the electronic nose are reached. A possible solution to this problem should be to improve the sample uptake, for example, by preconcentrating the analytes prior to investigation (20).

Other requirements of field monitoring include low cost, low maintenance, small size, low response time, and low power consumption where autonomous, portable instruments are concerned.

TRENDS AND FUTURE WORK

Electronic nose technology exhibits several possibilities for environmental monitoring. There is a true potential for sensor array in this area, provided that some conditions are respected. With the technology used so far, it is unrealistic to envisage a universal electronic nose to cope with any odor type. Specific data processing and sometimes specific instruments must be designed for each application. In the field of water quality, sewage odor profiles are specific for individual treatment works and for different unit processes within a work. These different sewage odor compositions induce a scatter of observations, which must be removed before trying to find a relationship with odor concentration determined by other methods (21).

Results to date have been based mainly on assessing environmental odors measured near emission sources. This is due to the constraints of using commercial chemical sensors that exhibit a limit of detection often higher than the threshold of smell (generally on the order of 1 ppmv). So, environmental monitoring by an electronic nose must be envisaged only in the surroundings of the emission.

Potential applications in odor assessment by electronic noses are numerous, but, before these specific applications can become a reality, a number of challenges still need to be overcome. Programs of experimental work should be undertaken to assess various characteristics of electronic nose performance, including drift, comparability of sensors before and after replacement, environmental influence, sensitivity, and signal-to-noise ratio.

There are not, so far, any possible comparisons between different electronic nose results. Even an instrument whose sensors were replaced is no longer the same as the original: the learning phase must be started again. In the frame of environmental applications, that particularly excludes the possibility of using the electronic nose to test compliance with standards.

So, it is necessary to develop calibration procedures and to realize suitable calibration artifacts. Any new “calibration standards” will need to consist of generic mixtures that reflect the gas components typically found in the environment and have a significant effect on the various detector responses. It is also essential to validate the quantitative assessment of sensor array responses against other methods, such as olfactometry measurements, to confirm comparisons with human perception.

There is a need to develop internationally acceptable methodologies for harmonizing electronic nose characterization and performance testing. This requirement can be best met by developing standardized procedures and protocols that enable making quantitative and objective comparisons between different types of instruments. These procedures could then form the basis of future international standards in this area.

BIBLIOGRAPHY

1. European Standard EN13725. (2003). *Air Quality-Determination of Odour Concentration by Dynamic Olfactometry*.
2. ASTM - E679-04. (2004). *Standard Practice for Determination of Odor and Taste Thresholds By a Forced-Choice Ascending Concentration Series Method of Limits. ASTM International*.
3. Ramel, M. and Nomine, M. (2000). Physicochemical characterisation of odours. *Analusis* **28**: 171–179.
4. Gardner, J.W. and Bartlett, P.N. (1999). *Electronic Noses: Principles and Applications*. Oxford University Press, Oxford.
5. Pearce, T.C., Schiffman, S.S., Nagle, H.T., and Gardner, J.W. (Eds.). (2002). *Handbook of Machine Olfaction: Electronic Nose Technology*. Wiley-VCH, Weinheim.
6. Nagle, H.T., Gutierrez-Osuna, R., Kermani, B.G., and Schiffman, S.S. (2002). Chapter 22: Environmental monitoring. In: Ref. 5, pp. 419–444.
7. Stuetz, R.M. and Nicolas, J. (2001). Sensor arrays: an inspired idea or an objective measurement of environmental odours? *Water Sci. Technol.* **44**: 53–58.
8. Bourgeois, W., Romain, A.C., Nicolas, J., and Stuetz, R.M. (2003). The use of sensor arrays for environmental monitoring: Interests and limitations. *J. Environ. Monitoring* **5**: 852–860.
9. Gostelow, P., Parsons, S.A., and Stuetz, R.M. (2001). Odour measurements for sewage treatment works. *Water Res.* **35**: 579–597.
10. Dewettinck, T., Van Hege, K., and Verstraete, W. (2001). The electronic nose as a rapid sensor for volatile compounds in treated domestic wastewater. *Water Res.* **35**: 2475–2483.
11. Stuetz, R.M., Georges, S., Fenner, R.A., and Hall, S.J. (1999). Monitoring wastewater BOD using a non-specific sensor array. *J. Chem. Technol. Biotechnol.* **74**: 1069–1074.
12. Ueyama, S., Hijikata, K., and Hirotsuji, J. (2001). *Water monitoring system for oil contamination using polymer-coated quartz crystal microbalance chemical sensor. Instrum. Control Automation*, Malmo, Sweden, pp. 287–292.
13. Bourgeois, W. and Stuetz, R.M. (2000). Measuring wastewater quality using a sensor array: prospects for real-time monitoring. *Water Sci. Technol.* **41**: 107–112.
14. Gardner, J.W., Shin, H.W., Hines, E.L., and Dow, C.S. (2000). An electronic nose system for monitoring the quality of potable water. *Sens. Actuators B* **69**: 336–341.
15. Flint, T.A., Persaud, K.C., and Sneath, R.W. (2000). Automated indirect method of ammonia flux measurement for agriculture: Effect of incident wind angle on airflow measurements. *Sens. Actuators B* **69**: 389–396.
16. Orts, J., Llobet, E., Vilanova, X., Brezmes J., and Correig, X. (1999). Selective methane detection under varying moisture conditions using static and dynamic sensor signals. *Sens. Actuators B* **60**: 106–117.
17. Nicolas, J., Romain, A-C., and Maternova, J. (2001). Chemometrics methods for the identification and monitoring of an odour in the environment with an electronic nose. In: M.T. Ramirez (Ed.), *Sensors and Chemometrics*. Research Signpost, Kerala, pp. 75–90.
18. Romain, A.C., Nicolas, J., and Andre, Ph. (2002). Three years experiment with the same tin oxide sensor arrays for the identification of malodorous sources in the environment. *Sens. Actuators B* **84**: 271–277.
19. Artursson, T. et al. (2000). Drift correction for gas sensors using multivariate methods. *J. Chemometrics* **14**: 711–723.
20. Nakamoto, T. (2002). Chapter 3: Odor handling and delivery system. In: Ref. 5.
21. Stuetz, R.M., Fenner, R.A., and G. Engin, G. (1999). Assessment of odours from sewage treatment works by an electronic nose, H_2S analysis and olfactometry. *Water Res.* **33**: 452–461.

OIL-FIELD BRINE

Dusan P. Miskovic
Northwood University
West Palm Beach, Florida

INTRODUCTION

From the beginnings of oil production, oil producers started to separate the oil field brine (OFB) mixture of oil and salt water from crude oil. The basic saline constituents of an OFB are as follows: chlorides, sulfates, hydrocarbonates, and sulfides of Na, Ca, Mg, ferrous Fe, and Sr. Much smaller concentrations of K, Li, B, Zn, Cu, bromides, and iodides can be attributed to total salinity as well. The actual salt content (qualitative and quantitative) may vary in a broad range of appearance and concentrations (1,2). Very often, the total concentration of salts in OFB can exceed several times the salinity of sea water. The separation of oil and OFB is a logical consequence of the oil production process due to the fact that much more OFB than oil is more often produced, especially in older oil wells. However, until recently neither the oil drillers nor any governmental agency paid much attention to disposal of OFB as a by-product of crude oil production. A simple procedure for OFB was diluting it in streams, rivers, and natural fresh and marine waterbodies. That practice took place until interest in crude oil production and consequently consumption of derivatives dramatically increased. Widespread oil production worldwide is constantly increasing, but the pollution aspects came to a focal point of environmental concern. To keep a high profit margin in the industry and satisfy strict environmental quality regulations, more sophisticated technologies for OFB treatment are becoming viable alternatives to the conventional OFB management practice.

HISTORICAL REVIEW OF OFB CONTROL PRACTICE

A chronology of OFB control was presented recently in the context of a concurrent increase in oil production efficiency and environmental quality self-regulation (3). It was pointed out through ethical consideration that technically improved efficiency of oil production did not successfully resolve pollution problems caused by OFB.

Generally, since the beginning of commercial oil production in the 1920s, a relatively small amount of OFB was lifted from formations. In most cases, oil producers did not realize that they had a disposal problem until someone objected. There are few records about early OFB control practices worldwide. In most inland oil fields, OFB was dumped into shallow evaporation pond (pits).

Since the first lawsuits were filed, simple OFB control practices were used. During the 1930s in California, for example, up to 200,000 barrels/day were simply discharged into the ocean (3). A common way to "control" OFB was in the context of early wastewater control practice; "dilution can be a solution for pollution." Saline water from oil fields would be released into dry channels when rain turned these into fast moving streams. Other practical modes to cope with the lack of systematic regulation in a particular industry were an oil producer's acceptance of the costs associated with storage reservoirs, evaporation ponds, and pipes to the ocean. These were ways to prevent damage suits.

The period of OFB management, as a method of pollution prevention, started in the 1930s. In Texas, for example, limited production from most wells was introduced. A more technical approach to the problem was addressed to prevent the field's hydrostatic pressure from dropping too quickly. Reducing and stopping production when the water-to-oil ratio exceeded a certain value was recommended. Also, it was concluded that evaporation ponds that created more concentrated brine did not resolve the disposal problem. A new alternative for OFB disposal in the form of brine injection wells became more suitable.

The brine injection technology seemed to be very promising as a way of disposing of OFB and also as a way to boost oil production. It was their own technology, which was being improved until now (see next section). In the beginning, an "environmentally friendly" method had been put to additional good use. Not long after the practice was widely accepted, some abandoned injecting wells came to life either as artesian salt springs or as steady salty streams polluting surface freshwater bodies.

Summarizing this brief historical review of OFB pollution control practice, it is possible to point out that a positive shift from increasing industrial efficiency to environmental quality improvement occurred after long-term experience in monitoring. Finally, a new approach to OFB control in the form of wastewater treatment technology emerged.

OFB USE FOR IMPROVING OIL PRODUCTION

Crude oil produced by natural reservoir pressure is referred to as primary production. Any additional production of oil from the introduction of artificial energy into a reservoir is considered enhanced recovery. As mentioned before, OFB injection technology was introduced primarily to recover as much as 50–70% of the original remaining petroleum. A comprehensive historical development emphasizing the environmental issues was presented (3), and a concise review of oil field water injection technology is given in Reference 2. In addition, there is much information scattered in oil production literature (4–6). Specific data focused on improving the quality of OFB before reinsertion is found elsewhere (7–10).

A whole industry supporting well injection was established. It deals with the conditioning of the physico-chemical properties of OFB, to remove remaining oil, as well as to prevent scaling (11). It is of great importance to establish injection water (OFB) quality criteria to prevent interruption or a decrease in oil production. There are three categories of substances affecting reinsertion of OFB: inorganic (dissolved gases, TDS, and TSS), organic (residual oil), and microorganisms. More important qualitative features of OFB and their pollution consequences after well injection will be discussed later.

To understand the many sources of wastewater originating in oil production and some of the polluting characteristics most relevant for OFB management as its major component, the schematic presented in Fig. 1 will be discussed.

Assuming that up to 90% of all fluid lifted from an oil well is practically OFB, the main source of wastewater in oil production is OFB itself. The initial volume of OFB is somehow decreased after two steps of crude oil dehydration, free water knock out (FWKO)

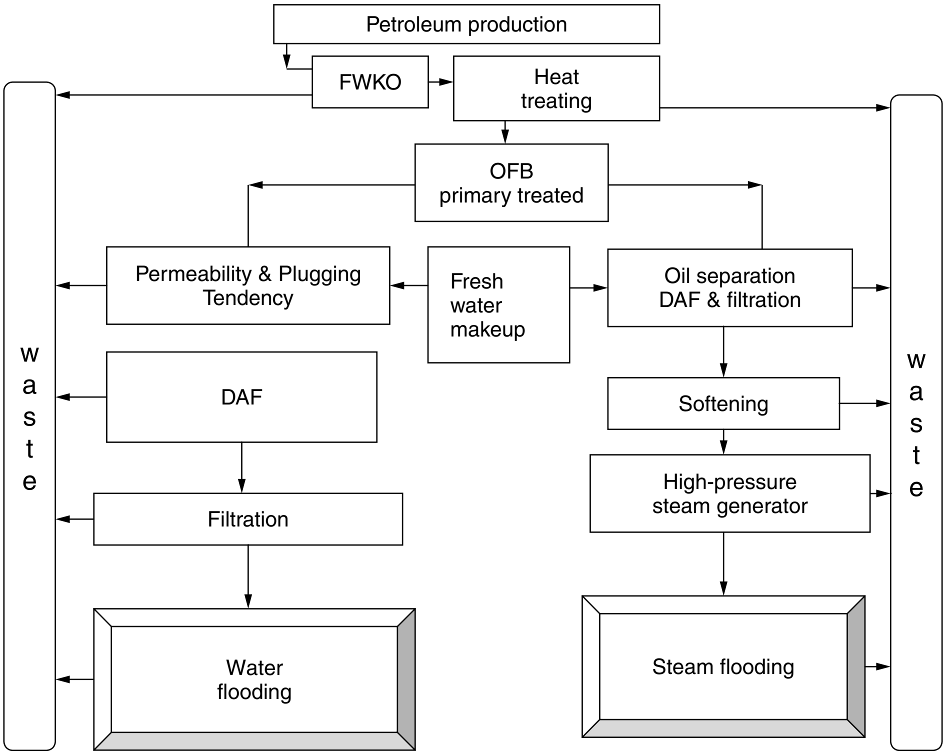

Figure 1. Schematic of OFB use for improving oil production.

and heat treating. Following the removal of oil and gas from the fluid produced, primary treated OFB is now prepared for additional treatment in the course of enhanced recovery. Water and steam flooding are the most common processes for secondary recovery of crude oil. Several steps of additional preparation/purification of the produced water (OFB) have to be conducted, regardless of the type of process applied. In flooding, specific tests have to be done to evaluate permeability and the plugging tendency of a core subjected to a particular floodwater (3). The produced water has to be clarified after mixing with supplemental fresh makeup water. A combination of dissolved air flotation (DAF) and filtration methods can be applied to produce brine physicochemically compatible with the formation rock (12). Also, a more sophisticated method such as radially entrained vapor extraction (REVEX) was considered for removing oil and solids from brine (13). This technology uses well-accepted principles of gravity and centrifugal sedimentation, as well as flotation.

On the other hand, a more complex train of consecutive stepwise processes have to be performed to prepare OFB for steam flooding. Here are three steps of OFB treatment:

- physicochemical removal of oil by gravity separation, primary/secondary DAF, and filtration;
- two-stage softening to provide water suitable for a high-pressure boiler;
- high-pressure steam generation based on previously deoiled and clarified OFB.

The serious disadvantage of steam flooding is condensed water, which contains all solids (salts) concentrated several times compared to the feed water. Practically, very concentrated brine as a disposal stream must be reinserted. An improvement in the existing process was achieved recently in German and Dutch oil fields (14). This new process allows the production of high quality boiler feed water by using mechanical vapor compression, without salt content for steam generation. It was accomplished by ensuring a closed water cycle, thus preventing the excessive volume of injected steam from brine, by continuously adding fresh water instead.

It is possible to conclude that no matter which process is chosen to improve oil production, a far from ideal elimination of excess OFB by reinsertion is a reality. Generally speaking, the OFB used in crude oil production by reinserting it into wells still remains a major source of groundwater pollution (see next section).

POLLUTION ASPECTS OF OFB

The extensive use of injection wells for disposal of a variety of chemical wastes, including OFB, started in the late 1960s. The first concrete action to protect drinking water supplies from possible contamination by injection wells came from the Safe Drinking Water ACT of 1974 (3). The same practice continued, and during the 1990s, there were approximately 170,000 active and more than 2 million abandoned injection wells in the United States. Generally, to estimate the total quantitative or qualitative budget of OFB even on a single oil field requires a very systematic analysis of a long-term record. That was also the problem in pollution control treatment (15).

For example, the concentration range of TDS contained in an OFB varies from 10 to more than 350 g/L, of which sodium chloride contributes about 80% or more. In contrast to the extreme high levels of TDS, TSS values are at incomparably lower levels, in most cases around 100 mg/L, even though, three categories of constituents affect OFB well injection: inorganic, organic, and microorganisms. The crucial impact on freshwater pollution is due to TDS content.

A concise review of the impact of OFB liquidation on water quality in reverse chronological order is presented in Table 1. Based on the available data, it is possible to distinguish two types of research activity: the water management approach and specific characterization of OFB in selected regions with different implications for water quality. In most cases, two categories of water sources were investigated. Because deep well injection is the most common OFB liquidation technique, particular attention was paid to groundwater quality change (16–21). After many years in use, OFB pits and injection wells have started to impact the quality of streams and shallow aquifers too (22–30). Some cases applied an integrated approach to a combined quality control management of groundwater and surface water affected by OFB disposal (31).

In addition to the two main categories of water sources affected by OFB contamination, there is a third group of more specific and very diversified implications of OFB management practice. Some illustrative examples are briefly described below.

An interesting result was obtained from a simulating study of the use of OFB as a deicer on a gravel roadbed (16). Constituents of brine were detected in water samples from groundwater monitoring wells. The map of an OFB plume was generated based on groundwater quality and surface geophysical methods.

A recently published article illustrates research in progress that quantifies human health risks from naturally occurring radioactive materials (NORM) related to OFB streams, pipe scale, and sludge (33). In another risk-based approach, the status of the water quality of a former OFB disposal pit, whose operation was discontinued in the mid-1940s, was evaluated (34). After a thorough water analysis, it was found that a long-term

Table 1. Impact of OFB on Water Quality

Year	Research Status (Region)	Impact/Implication	Ref.
	Management/Modeling	*Groundwater*	
1992	Use of a OFB as a deicer on roads (OH)	OFB plume in groundwater	16
1987	Groundwater management model for OFB plume containment (KS)	The potential for improving groundwater quality affected by OFB	17
1985	An inventory of long-term (1930s–1950s) disposal of OFB by ponds, wells, and pits (KS)	The progressive deterioration of groundwater quality during the time	18, 19
1985	As simplified model of differentiating and mixing of groundwater; OFB and natural brine (OK)	Potential for contaminating a freshwater aquifer	20
1984	General study on OFB contamination (NM)	OFB pit and deep groundwater quality (10,000 ft)	21
	Monitoring	*Surface Water, Shallow Aquifers*	
2001	Sodium adsorption ratios and salinity levels in aquifers (TX)	Agriculture and irrigation	22
1997	Nitrate, chloride, and bromide concentrations in OFB increased as well depth decreased (TX)	Cropland, pasture, and groundwater quality	23
1989	OFB pits, oil-exploration holes, and watershed (TX)	The salinization of aquifers and hydrocarbon concentration increase in artesian wells	24
1980	Colorado River water quality (TX)	Salinization of a river flow	25
1976	OFB recharge and groundwater quality fluctuation (OH)	Shallow groundwater aquifer contamination	26
1976	OFB and river water quality (KS)	An extremely high chloride concentration in river water	27
1975	Water supply and OFB pits and injection wells (OH)	Salinization up to 28 g/L of surface water	28
1975	Safe future development of water supply (OK)	Salinization of shallow alluvial aquifers	29
1975	The estimate of the quantities of a waste by deep well injection; general and OFB	Potential impact on water quality in the context of the Water Pollution Control Act of 1972	30
	Integrated Management	*Surface/Groundwater*	
1985	Study on the interrelationship of surface and groundwater quality management (IL)	OFB as a quality impact parameter	31
	Specific Monitoring	*Diversified*	
2000	Isolation and characterization of bacteria; nitrate-reducing; sulfide-oxidizing (Canada)	Characterization of OFB biogeochemical cycle	32
1999	Naturally occurring radioactive materials (NORM) in OFB (TX, LA, OK, MS, CA)	Human health risk	33
1997	Remediation of water quality of a former disposal pit (TX)	Surface water and aquatic life risk assessment for recreation	34
1994	Isolation of anaerobic bacteria (OK)	New specific anaerobic bacteria identified in OFB	35
1981	Chemical characterization of brines in salt mines (LA)	Potential impact of OFB on salt mines	36

natural remediation process gradually turned brine into freshwater that sustained aquatic plants. The pond was categorized for potential recreational use.

Specific microbiological investigation has proven that OFB can be a good substrate for the growth of strictly anaerobic bacteria (35). In another microbiologically focused study, two novel nitrate-reducing and sulfide-oxidizing strains of bacteria were isolated from production fluid (32).

An interesting study was performed on the impact of OFB on salt mines (36). Based on isotopic studies, it was hypothesized that active leaks, stalagmites, and stalactites in salt mines, which contained high levels of Ca, K, and bromides, arose from OFB formation water.

Generally, it is possible to conclude that a very broad spectrum of monitoring research data clearly indicates the significant impact of OFB management practice on the deteriorating quality of different water resources.

OFB CONTROL BY WASTEWATER TREATMENT METHODS

The previously considered aspects of OFB management such as environmental monitoring, socioeconomic aspects and improvement of crude oil production by using OFB were based on a broad pool of literature sources. In contrast, reports on the control of OFB by wastewater treatment methods are rather limited and sporadic.

It is worthwhile mentioning that there are some attempts to treat oil field effluents in a different engineering context. Cationic polyelectrolytes were developed to partially treat effluents containing a high percentage of residual free oil, as well as stable emulsions (37). Also, there are reports of treating OFB by using configurations similar to those employed for oil refinery wastewater (38–41). Two case studies of systematic approaches will be presented below. In both studies the OFB was considered wastewater that needed to be purified and discharged or reused.

Table 2. Alternatives of Microbiological[a] and Advanced Treatment[b]

No.	Description
Alternative I	[a]Activated sludge [b]Advanced treatment with GBAC
Alternative II	[a]Activated sludge + 25% dilution with freshwater [b]Advanced treatment with GBAC
Alternative III	[a]Activated sludge + PAC [b]Advanced treatment with GBAC
Alternative IV	[a]Activated Sludge + 25% dilution with freshwater + PAC [b]Advanced treatment with GBAC
Alternative V	[a]Activated sludge + 50% dilution with freshwater + PAC [b]Advanced treatment with GBAC

Various aspects of an innovative technology for OFB treatment were investigated on a laboratory and pilot scale (15,42–44). Both free and dispersed oily matter were separated by gravitation and sedimentation. Apart from physicochemical oil removal processes, special attention was paid to different variants of improved microbiological treatment (see Table 2). In that regard, two approaches were applied; dilution with freshwater and powdered activated carbon (PAC). The aim was to intensify the microbiological process and neutralize its inhibitors by adsorption on PAC. Advanced treatment was carried out on granular biological activated carbon (GBAC). It was found that the GBAC removed up to three times more organic matter than its adsorption capacity, and thus its adsorption power was not exhausted. This was a clear indication that the microorganisms present in the biofilm on GBAC oxidize the adsorbed pollutants and regenerate it at the same time. A technological scheme for complete treatment was proposed (see Fig. 2). It was concluded that the complete treatment employed here was very efficient. The organic matter content in the effluent did not exceed

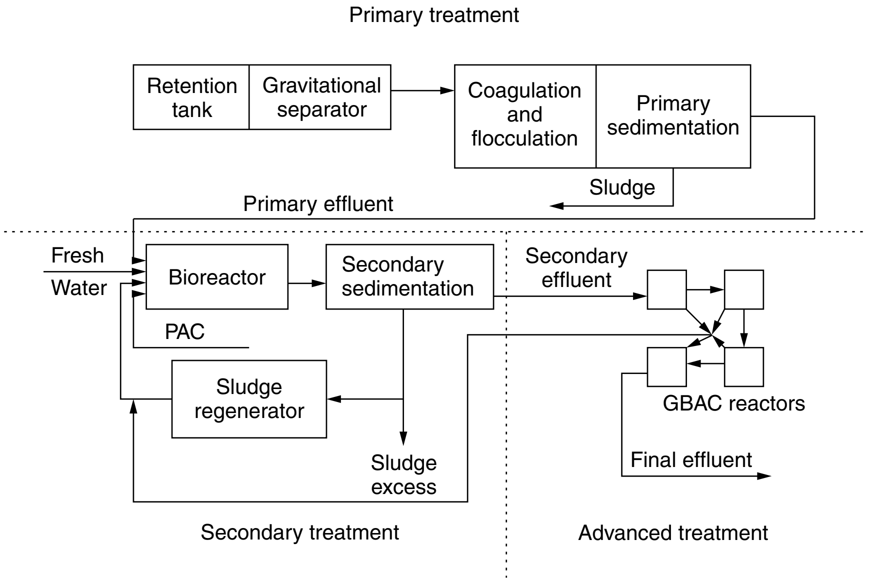

Figure 2. Block schematic of a complete OFB treatment.

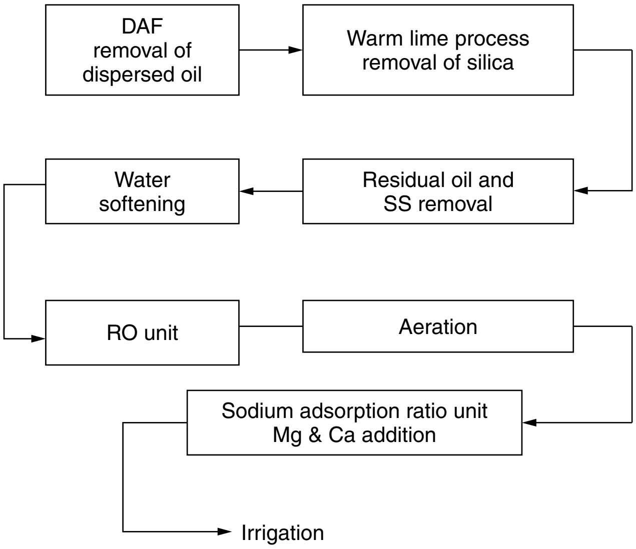

Figure 3. Block schematic of RO pilot unit for treating OFB.

2.5 ppm of BOD. Also, a final decrease of salinity was achieved by mixing effluent in an adequate proportion with the recipient water (44). Thus the treated OFB can be discharged into recipients of low self-purification capacity. Based on a preliminary cost analysis, it was estimated that a complete OFB treatment would cost 10% more than a conventional PAC/GBAC treatment of municipal wastewater.

The second case study of complete OFB treatment represents a state-of-the-art process based on reverse osmosis (RO). The configuration applied converts an oil field effluent into freshwater (45). The simplified block schematic of the process is presented in the Fig. 3. The input to the process contained a relatively low level of salinity (about 7 g/L).

The concentration of a soluble oil was also at a moderate level, below 200 ppm. Still, due to the passage of oil through the filters, RO membrane fouling was observed. Despite that, the overall results showed clearly successful conversion of OFB into freshwater of irrigation quality. The operating cost was less then 10¢/bbl.

Concluding this section, it is possible to draw out some positive outcomes. Crude oil still represents the most important energy source, so it is crucial to continue further improvement of OFB treatment technologies. The case studies presented clearly show the feasibility and the cost-effectiveness of the complete OFB treatment. It is quite obvious that OFB pollution control by wastewater treatment methods is superior to the liquidation–reinsertion practice of the past.

BIBLIOGRAPHY

1. Stoessell, R.K. (1997). *Ground Water* **35**(3): 409–418.
2. Kemmer, F.N. (Ed.). (1988). *The Nalco Water Handbook*, 2nd Edn. McGraw-Hill, New York, pp. 43.1–43.23.
3. Gorman, H.S. (1999). *Business History Review* **73**(4): 601–640.
4. American Petroleum Institute (1983). *Primer of Oil and Gas Production*.
5. Collins, A.G. (1975). *Geochemistry of Oilfield Waters*. Elsevier, Amsterdam.
6. American Petroleum Institute (1974). *Treating Oil Field Emulsions*, 3rd Edn.
7. Wright, C.C., Cross, D., Ostroff, A.G., and Stanford, J.R. (1977). *Oil Field Subsurface Injection of Water*. ASTM, Philadelphia.
8. Muvrin, B. and Salopek, B. (1992). *Nafta* **43**(7): 349–362.
9. Ariman, T. and Veziroglu, T.N. (Eds.). (1987). *Particulate and Multiphase Processes*. Hemisphere. Vol. 2, pp. 535–550.
10. Hoke, S.H. and Collins, A.G. (1981). *Proceedings of the Second ASTM Symposium on Water for Subsurface Injection*, Ft. Lauderdale, FL, pp. 34–48.
11. M-I Drilling Fluids Co. (1992). 1992–93 General Catalog, Houston.
12. Miskovic, D. and Sears, B. (1994). *Private Communication*. Lenox Institute of Water Technology, MA.
13. Advanced Separation Technologies, Inc. (1994). *REVEX Technology*, Houston.
14. Koren, A. and Nadov, N. (2002). *Mechanical Vapor Compression to Treat Oil Field Produced Water*. I.D.E. Technologies Ltd., Ra'anana, Israel.
15. Miskovic, D. et al. (1989). *Water Sci Technol.* **21**: 1849–1852.
16. Chapman, M.J. and Bair, E.S. (1992). *Ground Water Monitoring Rev.* **12**(3): 203–209.
17. Heidari, M. et al. (1987). *Water Resour. Bull. WARBAQ* **23**(2): 325–335.
18. Whittemore, D.O. et al. (1985). *Proceedings of a Symposium on Groundwater Contamination and Reclamation*, Tucson, AZ, pp. 109–116.
19. Sophocleus, M.A. (1984). *J. Hydrology* **69**(1–4): 197–222.
20. Mast, V.A. (1985). *Ground Water Monitoring Rev.* **5**(3): 65–69.
21. Stephens, D.B. and Spalding, C.P. (1984). *Proc. Ogallala Aquifer. Symposium II*, Lubbock, TX, pp. 440–450.
22. Hudak, P.F. (2001). *Int. J. Environ. Stud.* **58**(3): 331–341.
23. Hudak, P.F. and Blanchard, S. (1997). *Environ. Int.* **23**(4): 507–517.
24. Dutton, A.R. et al. (1989). *Ground Water GRWAAR* **27**(3): 375–383.
25. Rawson, J. (1980). *Geological Survey Austin TX*, 80–224, p. 66.
26. Pettyjohn, W.A. (1976). *Ground Water* **14**(6): 472–480.
27. Leonard, R.B. and Kleinschmidt, M.K. (1976). *Kansas Geological Survey*. Lawrence Chemical Quality Series 3, September. p. 24.
28. Pettyjohn, W.A. (1975). *Ground Water* **13**(4): 332–339.
29. U.S. EPA-660/3-74-033, April (1975).
30. Conrad, E.T. and Hopson, N.E. (1975). *Water Resour. Bull.* **11**(2): 370–378.
31. Crandall, D.A. (1984). *Proc. 20th Annu Conf. Ame Water Resour Assoc. Symp.*, Washington, DC, pp. 193–199.
32. Gevertz, D. et al. (2000). *Appl Environ Microbiol.* **66**(6): 2491–2501.
33. Anonymous (1999). *Oil Gas J.* **97**(1): 29–32.
34. Alvillar, S.B. and Hahn, R.W. (1997). *Proc. Air Waste Manag Assoc 90*. Annual Meeting, Toronto, 8–13.
35. Bhupathiraju, V.K. et al. (1994). *Int. J. Systematic Bacteriol.* **44**(3): 565–572.

36. Kumar, M.B. and Martinez, J.D. (1981). *J. Hydrology* **54**(1–3): 107–140.
37. Mahanta, D. (1994). *Chem. Ind. Dig.* **7**(2): 125–127.
38. Roev, G.A. (1981). *Ochistnye Sooruzheniya Gazonefte-Perekachivayushchih Stancii I Neftebaz*. Nedra, Moskva.
39. Jones, H.R. (1973). *Pollution Control in the Petroleum Industry*. Noyes Data Corp., Park Ridge, NJ, pp. 175–202.
40. Barnes, D. et al. (1984). *Surveys in Industrial Wastewater Treatment 2-Petroleum and Organic Chemicals Industries*. Pitman, London, pp. 29–52.
41. Miskovic, D. et al. (1986). *Water Sci. Technol.* **18**(9): 105–114.
42. Dalmacija, B. et al. (1995). *Water Res.* **30**(2): 295–298.
43. Dalmacija, B. et al. (1996). *Water Res.* **30**(5): 1065–1068.
44. Miskovic, D. *Private communications with local oil company*. Novi Sad, 1983–1986.
45. Tao, F.T. (1993). *Oil Gas J.* **91**(38): 88–91.

OIL POLLUTION

Kusum W. Ketkar
Rajiv Gandhi Chair
Jawaharlal Nehru University
New Delhi, India

The word oil refers to hydrocarbons—crude oil, gasoline, heating oil, fuel oil, diesel, jet fuel, and other refined oil-based products. Oil pollution refers to input of hydrocarbons into a water body, either on the surface of the earth or underground. Oil may directly enter surface water bodies like ponds, rivers, streams, and oceans, and it may leach into underground water. Recognizing that moving water cleans itself, oil pollution occurs when more oil enters water than natural, physical, and chemical processes can absorb it.

Accidental oil spills by oil tankers are only one of many causes of oil pollution, but they attract a lot of media attention. Media attention is justified because accidents spill a lot of oil in a short period of time. For example, oil tanker *Exxon Valdez* spilled 11 million gallons of crude oil in Prince William Sound, Alaska, in 1989. In addition to accidental oil spills, oil may enter marine environment through natural seepage, off-shore oil drilling, illegal disposal of waste water containing oil from tankers, tank barges and other vessels, land-based petrochemical industrial complexes, runoff containing oil from roads, oil-contaminated stormwater, and sludge from municipalities and numerous other land-based activities. Initially, oil-contaminated water may enter local streams, rivers, lakes, or ponds, but ultimately it is dumped into the oceans. According to Oceana, 80% of oil dumped into the sea is from inland operators, of which 44% is the result of direct dumping or coastal drainage, 33% is transported through the atmosphere, and the remaining 20% comes from accidental or deliberate spills from vessels and marine facilities (see Fig. 1) (1).

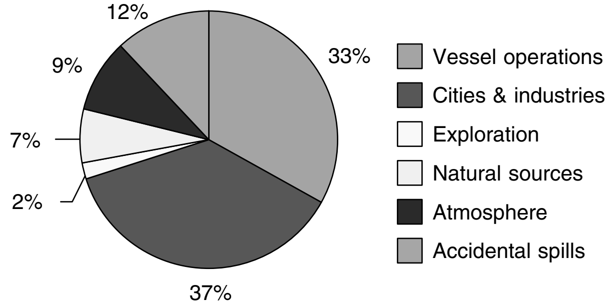

Figure 1. Causes of oil pollution. *Source*: Euoropa Oceana.

Worldwide, accidental oil spills account for about 12% of the total oil input into our oceans, whereas deliberate dumping from vessels and illegal dumping from tankers on the high seas account for 33%. Every year, over 6000 oil tankers use oceans as highways to transport crude oil. In 1983, 1.21 billion tons of crude oil was transported, and by 2002, that quantity had risen to 2 billion tons of crude. Table 1 gives estimated world maritime operational and accidental sources of oil entering marine environment.

On average, 25 major accidents and many smaller ones spill over 700 tons of oil into the seas every year (see Table 2). Overall, 2 to 10 million tons of oil leaks into the oceans annually.

Table 3 lists oil spills over a million gallons (3367 tons) in U.S. waters. After 1990, no spill in U.S. waters has been over a million gallons, and overall, there has been a sharp decline in large oil spills since the passage of the Oil Pollution Act of 1990. Most oil spill incidents in the United States occur along Alaska and the western and southern coasts. Alaska, Texas, and Louisiana are the major oil-producing states in the United States.

Since World War II, unprecedented increases in the world demand for energy and the rapid economic growth of the world economy have gone hand in hand. Oil is an indispensable source of energy in the production of goods and services in general and particularly in transportation. Most countries are poorly endowed with this valuable product also known as “black gold.” The result is that oil-poor countries have to import it from oil-rich countries. Petroleum and petroleum products are transported in tank ships or tank barges. According to Oceana, over 6000 tankers traverse world oceans carrying crude oil and oil products. The United States is the world’s largest consumer of oil, and it imports over 50% of the oil that it needs. Oil comes from both short-hauls like Mexico, Venezuela, and Nigeria and long hauls like Saudi Arabia,

Table 1. World Maritime Operational and Accidental Sources of Oil Entering Water Bodies (Million Tons Annually)

Source	1990	1981–85	1973–75
Bilge and fuel oil	0.25	0.31	—
Tanker operational losses	0.16	0.71	1.08
Accidental spillage			
Tanker accidents	0.11	0.41	0.20
Nontanker accidents	0.01		0.10
Marine terminal operations	0.03	0.04	0.50
Dry-docking	0.01	—	—
Scrapping of ships			

Source: U.S. Coast Guard, OPA-90, Double Hull Tanker Study, 1991.

Table 2. World's 25 Biggest Oil Spills Since 1970

Number	Quantity (10^3 Tons)	Location	Year	Accident
1	1,000	Arab Gulf	1991	Gulf War
2	816	Mexico	1979	Platform Ixtoc I
3	476	South Africa	1983	Castillo de Bellver
4	267	France	1978	Amoco Cadiz
5	234	Italy	1991	Haven
6	158	Canada	1988	Odessey
7	146	Trinidad	1979	A. Empress & A. Captain
8	145	Barbados	1979	Atlantic Empress
9	141	Greece	1980	Irene Seranade
10	135	UK	1967	Torrey Canyon
11	130	Oman	1972	Sea Star
12	115	North Sea	1971	Texaco
13	110	North Pacific	1977	Hawaiian Patriot
14	103	Kuwait	1981	Tank
15	100	Spain	1976	Urquiola
16	95	Turkey	1999	Independentza
17	92	Portugal	1969	Julius Schindler
18	85	Portugal	1975	Jacob Maersk
19	84	Nigeria	1979	Tank
20	83	UK	1993	Brear
21	81	Spain	1992	Aegean Sea
22	80	Morocco	1989	Kharg S
23	77	Spain	2002	Prestige
24	72	UK	1996	Sea Empress
25	72	Mozambique	1992	Kalina

Source: Reference 1.

Russia, the North Sea, and Indonesia. Over 4000 tankers make port calls every year, and the ports experiencing the highest tanker traffic are the ports of New York-New Jersey, Houston, Los Angeles/Long Beach, and San Francisco Bay. As the New York seaboard is shallow and has a soft bottom, only smaller tankers make port calls there. Crude from the Middle East and other long-haul countries is transported in very large crude carriers (VLCCs) and ultra-large crude carriers (ULCCs) to New York and other shallow ports. Lightering operation is used offshore to bring crude in smaller tankers to the shallow ports. These long-haul tankers (VLCCs) that bring crude to the United States have a capacity of 200,000 tons or more. They make calls to ports on the West Coast and the Gulf of Mexico because the coastal waters there are deep with rocky bottoms. Also, tank vessels use U.S. coastal waters, rivers, and lakes to distribute over 90% of petroleum and petroleum products, which results in over 20,000 port calls annually. As the vessel traffic increases, the risk of casualties and oil spills increases. Navigational risk in coastal waters is significantly influenced by vessel traffic, width of shipping channels, and the number of course changes. But the risk rises less than proportionately as vessel traffic increases (2). In addition to transportation of oil and refined products, there is offshore drilling and the crude is transported via pipelines to the mainland refineries. Alaskan crude is transported in tankers to the refineries. Thus, oil drilling, transfer, processing, and distribution make intensive use of the country's marine resources. With this intensive use comes the risk of oil input into the marine environment. The consumption of crude-based products generates waste, which releases oil into the natural environment. Hydrocarbons that are disposed of on land ultimately enter oceans.

The question is as follows: What is the impact of input of hydrocarbons on the marine environment? Scientists seem to agree that, in the long run, little adverse impact occurs on the marine environment. Take the Persian Gulf, for instance. Oil-producing Middle Eastern countries use ports on the Persian Gulf coast to ship crude oil and refined products. But biologists have observed that the Persian Gulf has experienced sharp increases in the concentration of oil-eating bacteria as oil input in the Gulf has increased (3). The environmental impact of oil input depends on the spill amount, its type—heavy or light crude or refined oil—the time of the year, location, previous exposure, and current length of exposure (4). The 1989 *Exxon Valdez* spill of 37,000 tons of crude in winter in Prince William Sound did more damage to the marine environment and the beaches than did the spill of 83,000 tons of crude by *Brear* on the high seas off the coast of Scotland. Temperate climates tend to have fewer oil-eating bacteria than do the tropics, and natural degradation takes much longer in the former climates than in the latter.

Generally, industrial, urban, and heavily populated areas are located near waterways and ports. Inputs of stormwater, sewerage, and industrial waste containing oil and oil-based products and input of oil from operational and accidental spills causes severe damage to the marine environment of the coastal waterways. For instance,

Table 3. Oil Spills in U.S. Coastal Waters

Number	Quantity (10^3 Tons)	Location	Year	Accident
1	37.00	Prince William Sound, Alaska	1989	Exxon Valdez
2	25.30	Nantucket, Massachusetts	1976	Argo Merchant
3	13.00	Gulf of Mexico, Texas	1990	Mega Borg
4	12.60	Pilot Town, Louisiana	1981	Apex Houston
5	9.30	Off Cameron, Louisiana	1984	Alvenus
6	6.90	South of Semidi Islands, Arkansas	1988	UMTB 283
7	4.24		1985	Exxon Barge No. 283
8	4.20	San Francisco, California	1984	Puerto Rican
9	3.54	Mississippi River Montx, Louisiana	1982	Arkas
10	3.53	Donaldsville, Louisiana	1983	Barge No. 218

Source: U.S. Coast Guard.

coastal waters around New Orleans, Houston, Galveston, New York, and many other ports are dead in the sense that they can support no marine life. These waters continue to violate the mandate of the Clean Water Act of 1977. Coastal waters around New York–New Jersey ports are neither suitable for fishing nor swimming. To go fishing or swimming in New Jersey, one has to drive about 50 miles south from New York City. Similarly, the Hudson and East Rivers around New City have not been able to support any marine life until recently. The ecological impact of oil trapped in environmentally sensitive areas like marshes, wetlands, estuarine and freshwater communities, and intertidal and subtidal zones is the heaviest, although these effects appear to last from a few months to 15–20 years. But there can be exceptions. For instance, long-term and concentrated exposure of oil has killed colonies of coral reefs in Indonesian coastal waters and is suspected to have adversely affected invertebrates and plants in the coral reef community. The number 2 fuel oil that oil barge *Bouchard* 65 spilled in the west entrance of the Cape Cod canal in Buzzards Bay, Massachusetts, in 1974 entered salt marshes of Winsor Cove; and these mashes had not fully recovered by 2000 (5). The National Research Council observed that the most damaging effects of petroleum are oiling and tarring of beaches through the modification of benthic communities along polluted coastlines and the endangering of seabird species (6). Neither human health nor the food chain seems to be endangered by oil spills. The images of blackened beaches and dead birds are eye-catching, and television and the print media have effectively used them to win ratings and earn advertising revenues (7).

Large oil tankers crisscrossing the world's oceans and coastal waters pose a major threat for accidents leading to the spillage of large quantities of oil in a short period of time. Oil spreads quickly and evaporates. Heavy oils take longer to evaporate than do lighter and refined oils. If the spill is not quickly contained, it can trap sea birds and choke fish and other sea animals to death. If the spill is near a coastal area, it can kill life and damage property onshore and adversely affect commercial and subsistence fishing. The oil spill from the *Exxon Valdez* spread over hundreds of miles into numerous and remote coves, inlets, and bays from Prince William Sound through the Southern Kenai Peninsula and further south to Kodiak Island. Public outcry over the spoilage of pristine, rugged Alaskan environment cost Exxon-Mobile Corporation several billion dollars in cleanup costs. Organisms under rocks where hot water washing was used are taking longer to return to normal levels than those under those rocks that were left for nature to clean. In the aftermath of the accidental oil spill by the *Exxon Valdez*, the U.S. Congress passed the Oil Pollution Act in 1990 with the goal of preventing future oil spills and minimizing damage from accidental oil spills from operations and transportation of crude or refined petroleum products. The European Union followed the American example and passed a similar act after oil spills off the Spanish and British coasts. Among the preventive measures, the most controversial measures are the replacement of single hull tankers by double hull tankers and tanker owners' liability in case of a spill. As a result of various regulatory measures and increased vigilance by tanker ship owners, the number of large oil spills world wide has declined.

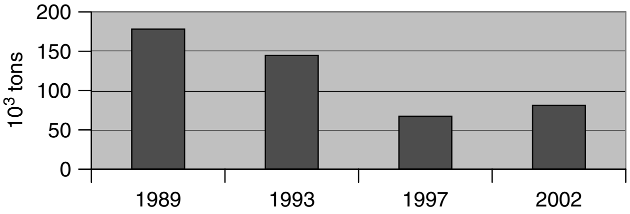

Figure 2. The world maritime operational oil spills. *Source*: International Tanker Owners Pollution Federation.

According to INTERTANKO, the quantity of oil spilled due to maritime operations has declined worldwide since 1989 because of the Oil Pollution Act of 1990 and the adoption of International Convention MARPOL 73/78 by the International Maritime Organization. The European Union followed the U.S. example and passed a law to prevent future oil spills after accidental spills by *Brear* off the coast of Scotland and by the *Aegean Sea* off the coast of Spain in the early 1990s. Figure 2 gives the quantity of oil spills from maritime operations.

BIBLIOGRAPHY

1. OCEANA. Available: http://www.euoropa.oceana.org.
2. Ketkar, K.W. (2003). The Oil Pollution Act of 1990—a decade later. *Spill Sci. Technol.* **7** (1–2).
3. Holloway, M. (1991). Soiled shores. *Sci. Am.* **October**.
4. Mielke, J.E. (1990). *Oil in the Ocean: The Short - and Long-Term Impacts of Spills*. Congressional Research Service, the Library of Congress, Washington, DC.
5. Hampson, G. Destruction and recovery of the Winsor Cove, Cataumet, MA Salt Marsh from a #2 fuel oil spill: a 25 year history. *Environment Cape Cod* **3** (1).
6. National Research Council. (1985). *Oil in the Sea: Inputs, Fates and Effects*. National Academy Press, Washington DC.
7. Ketkar, K.W. (1995). The US Oil Pollution Act of 1990 and the future of the maritime industry. *Marine Policy* **15**(5).

INDICATOR ORGANISMS

Gary P. Yakub
Kathleen Stadterman-Knauer
Allegheny County Sanitary Authority
Pittsburgh, Pennsylvania

INTRODUCTION

Assessing biological contamination has always been an important part of any water quality study. Sources of fecal contamination include natural or agricultural sources and infrastructure problems that result in the release of untreated sewage into natural waters. This contamination can result in releasing dangerous human pathogens into recreational waters and source waters used to supply potable drinking water. People who contact contaminated water through bodily contact or ingestion are at risk for a variety of infections.

Since 1920, 1884 waterborne disease outbreaks, 882,144 cases of illness, and 1169 deaths in the United States have been from contaminated drinking and recreational waters (1). The causative agents responsible for these diseases could be bacteria, fungi, viruses, helminthes (worms), protozoa, or other pathogenic organisms. An attempt to characterize the biological quality of a water sample by investigating all of the known causative agents of waterborne disease would be impractical due to the large number of possible agents and the time-consuming methodology required. In addition, some agents can cause disease even at very low concentrations, and other causative agents can be difficult, if not impossible, to culture in the laboratory. These factors point to the need for a system to monitor biological water quality quickly, easily, and cost-effectively.

Indicator bacteria are the most common way of assessing biological water quality. Indicator bacteria are bacteria that can predict the presence of pathogenic organisms. To be a successful indicator, the following criteria must be met:

1. Indicator bacteria must be easy to cultivate in the laboratory using simple, cost-effective procedures.
2. Indicator bacteria must be present in high enough concentrations to be readily detectable in a manageable sample volume.
3. Indicator bacteria must effectively predict the risk of illness from exposure to the waterborne pathogens present in contaminated water.

Indicator bacteria are used in epidemiological studies to establish a correlation between bacteria levels and the rate of disease in a given population. From this information, water quality standards can be developed for a particular indicator bacteria based on acceptable risk limits. These water quality standards can then be used for quantitative evaluation of biological water quality.

Indicator bacteria are currently used to assess water quality for a variety of applications. Wastewater treatment facilities use fecal coliforms or *Escherichia coli* as indicators of proper disinfection. Drinking water plants monitor total coliforms as an indicator of water quality in their distribution systems. Recreational water quality standards are based upon fecal coliforms, *E. coli*, or enterococcus bacteria levels depending upon water type and usage.

THE COLIFORM GROUP

The coliform group is the group of bacteria that has received the most attention for application as an indicator of fecal pollution. Coliform bacteria are found in the intestinal tract of humans and other animals; therefore, their presence in ambient water indicates fecal pollution and the potential presence of pathogens (2).

During the late 1940s and early 1950s, the United States Public Health Service began a series of studies based on coliform bacteria levels (3). Its goal was to determine if coliform concentrations in recreational water were correlated with the incidence of gastroenteritis in swimmers. From these studies, total coliform was used as the indicator bacteria for recreational waters. In the mid 1960s, total coliforms were replaced by fecal coliforms, which showed a greater specificity for mammalian fecal pollution and less variation due to environmental conditions. Fecal coliforms made up about 18% of total coliforms (3). In 1976, the U.S. EPA recommended fecal coliforms to the regulatory community as indicator organisms for the presence of pathogens in recreational waters (2). In 1986, based on new research data, U.S. EPA recommended new monitoring guidelines, including the use of *E. coli* as an indicator to replace fecal coliforms. The presence of *E. coli*, it was found, correlated with swimming-associated gastrointestinal illness better than fecal coliform levels (2).

THE ENTEROCOCCUS GROUP

In its 1986 Bacterial Water Quality Criteria report, the U.S. EPA also concluded that members of the enterococcus group of bacteria were acceptable as an indicator of fecal pollution due to the high correlation between enterococcus bacteria levels and swimming-related gastrointestinal illness. The strong correlation was noted in both freshwater and saltwater, and enterococci have been proposed as an indicator of water quality for both environments (3).

Enterococcus bacteria are a subgroup of the fecal streptococci that includes *S. faecalis, S. faecium, S. gallinarum,* and *S.avium*. Enterococci are differentiated from other streptococci by their ability to grow in 6.5% sodium chloride, at pH 9.6, and at 10 °C and 45 °C.

There has been some attempt to use the fecal coliform to fecal strep ratio to differentiate the source of fecal pollution. A ratio higher than four was proposed to suggest human sources, and a ratio of less than 0.7 was proposed to suggest nonhuman sources. However, factors such as a difference in die-off rates among the enterococci and a variability to chlorine disinfection resulted in a recommendation that the fecal coliform to fecal strep ratio should not be used for source differentiation (4).

THE FUTURE OF INDICATOR ORGANISMS

There are many organisms that do not correlate with the indicators used to determine fecal contamination in surface waters. The ecology, prevalence, and resistance to stress of indicator organisms may differ from many of the pathogenic microorganisms for which they serve as proxies. For example, fecal coliform does not indicate many disease-causing viruses, such as hepatitis A or E, Coxsackie viruses, adenoviruses, and Norwalk viruses or indigenous bacteria such as *Helicobacter* or *Legionella*. *Giardia* and *Cryptosporidium*, two parasitic protozoa, also do not correlate with fecal coliform levels (5). A

2001 American Academy of Microbiology report advocates the advancement of molecular techniques for detecting waterborne pathogens in the context of a risk-based analysis (5). Molecular tools could allow rapid detection of an organism, and also determine its virulence, prevalence, and source of contamination. However, many of these techniques are in development and may not be available to utilities for many years. Therefore, indicator organisms will continue to play a role in water quality analysis until these new techniques become available and cost-effective.

BIBLIOGRAPHY

1. American Water Works Association. (1999). *Waterborne Pathogens*. Manual of Water Supply Practice M48, Denver, CO.
2. U.S. EPA. (1999). *Action Plan for Beaches and Recreational Waters*. EPA/600/R-98/079. Washington, DC.
3. U.S. EPA. (1986). *Ambient Water Quality Criteria for Bacteria—1986*. EPA/440/5-84-002. Washington, DC.
4. American Public Health Association, American Water Works Association, and Water Environment Federation. (1992). *Standard Methods for the Examination of Water and Wastewater*. 18th Edn. Washington, DC.
5. Rose, J.B. and Grimes, D.J. (2001). *Reevaluation of Microbial Water Quality: Powerful New Tools for Detection and Risk Assessment*. A Report from American Academy of Microbiology, Washington, DC.

pH

Michael P. Dziewatkoski
Mettler-Toledo Process
Analytical
Woburn, Massachusetts

pH THEORY

Some chemical compounds, when dissolved in water, dissociate to yield hydrogen ions (H^+, e.g., protons), as in Equation 1, and are called acids. Other compounds called bases yield hydroxide ions (OH^-), when dissolved in water, as in Equation 2.

$$HA = H^+ + A^- \tag{1}$$

$$MOH = M^+ + OH^- \tag{2}$$

The chemical species in Equations 1 and 2 represented by A^- and M^+ are anions (negatively charged ions) and cations (positively charged ions), respectively. Acids and bases can be organic (containing two or more of the following atoms: carbon, hydrogen, oxygen, and nitrogen), such as acetic acid ($HC_3H_2O_2$), or inorganic (consisting of a metal and nonmetal), such as the base sodium hydroxide (NaOH). In addition to classification as organic or inorganic, acids and bases can also be ranked according to their strengths. Some common acids and bases are listed in Table 1.

The relative strength of an acid or base is related to the degree of dissociation that the compound undergoes according to either Equation 1 or 2. A strong acid is one that dissociates completely to yield a relatively large number of hydrogen ions, whereas a strong base is a compound that yields a relatively large number of hydroxide ions.

For strong acids and strong bases, the reactions described by Equations 1 and 2 proceed from left to right to form products and the reverse reactions are negligible. The converse is true for weak acids and weak bases; only a relatively small number of either hydrogen ions or hydroxide ions are produced. In the case of a weak acid or base, the reverse reactions described by Equations 1 and 2 are significant. Weak acids and bases are characterized by an acid dissociation constant (K_a) or a base dissociation constant (K_b) (see Table 1). The larger the value of either K_a or K_b, the greater is the degree of dissociation of the acid or base.

The hydrogen ion concentration in a solution of 1 mol/L hydrochloric acid would be approximately 1 mol/L. However, for a 1 mol/L sodium hydroxide solution (a base), the hydrogen ion concentration is approximately

Table 1. Common Strong and Weak Acids and Bases

Strong Acids		Strong Bases	
HCl, Hydrochloric acid		LiOH, Lithium hydroxide	
H_2SO_4, Sulfuric acid		NaOH, Sodium hydroxide	
HNO_3, Nitric acid		KOH, Potassium hydroxide	
$HClO_4$, Perchloric acid		$Ca(OH)_2$, Calcium hydroxide	
HBr, Hydrobromic acid		$Sr(OH)_2$, Strontium hydroxide	
HI, Hydroiodic acid		$Ba(OH)_2$, Barium hydroxide	
Weak Acids	K_a@25 °C[a]	**Weak Bases**	K_b@25 °C[b]
HF, Hydrofluoric acid	6.8×10^{-4}	$(CH_3)_2$ NH, Dimethylamine	5.1×10^{-4}
$HCHO_2$, Formic acid	1.7×10^{-4}	NH_3, Ammonia	1.8×10^{-5}
$HC_7H_5O_2$, Benzoic acid	6.3×10^{-5}	N_2H_4, Hydrazine	1.7×10^{-6}
$HC_2H_3O_2$, Acetic acid	1.7×10^{-5}	NH_2OH, Hydroxylamine	1.1×10^{-8}
HClO, Hypochlorous acid	3.5×10^{-8}	C_6H_5 NH_2, Aniline	4.2×10^{-10}
H_3BO_3, Boric acid	5.9×10^{-10}	NH_2CONH_2, Urea	1.5×10^{-14}

[a] Reference 1, p. 672.
[b] Reference 1, p. 680.

10^{-14} mol/L. In pure water, the hydrogen ion concentration is 10^{-7} mol/L. In the preceding examples, the hydrogen ion concentration varies by a factor of 10^{14} in going from acid to base. Therefore, it is more convenient to express the relative acidity of a solution by using a logarithmic scale based on hydrogen ion concentration and referred to as pH.

pH is expressed mathematically as

$$\mathrm{pH} = -\log[\mathrm{H}^+] \tag{3}$$

where $[\mathrm{H}^+]$ is the hydrogen ion concentration in the solution expressed in mol/L or molarity (M). Using this definition, the pH values for 1 M hydrochloric acid, water, and 1 M sodium hydroxide are 0, 7, and 14, respectively. Although the pH scale is technically defined from 0 to 14, pH values below 0 and above 14 are possible.

To contrast, consider the pH of a 1 M HCl solution and a 1 M acetic acid solution. From the previous example, we know that the pH of the 1 M HCl solution is 0, whereas the 1 M acetic acid solution will have a pH of 2.4. The difference in pH between the two solutions lies in the fact that acetic acid is a "weak" acid, and hence it does not completely dissociate in aqueous solution. For this reason, a simple relationship between pH and weak acid or weak base concentration does not hold true.

In our previous example, we stated that the pH of a 1 M NaOH is 14. To arrive at this result, one must first calculate the pOH:

$$\mathrm{pOH} = -\log[\mathrm{OH}^-] \tag{4}$$

and then calculate the pH:

$$\mathrm{pH} = 14 - \mathrm{pOH} \tag{5}$$

pH BUFFERS

The reaction shown in Equation 1 will proceed from left to right to a limited extent for a weak acid. In this case, the reverse reaction (formation of HA) can also occur to a significant extent. The same is true for weak bases (Equation 2). As an example, consider the weak acid acetic acid (Equation 6) and the weak base ammonia (Equation 7):

$$\mathrm{HC_2H_3O_2} = \mathrm{H}^+ + \mathrm{C_2H_3O_2}^- \tag{6}$$

$$\mathrm{NH_3} + \mathrm{H_2O} = \mathrm{NH_4}^+ + \mathrm{OH}^- \tag{7}$$

Acetic acid dissociates to form hydrogen ion and acetate ion ($\mathrm{C_2H_3O_2}^-$). Acetate ion is referred to as the conjugate base to acetic acid because it can acquire a hydrogen ion and reform acetic acid. Ammonia, on the other hand, acquires a hydrogen ion from a water molecule to form ammonium ion ($\mathrm{NH_4}^+$) and hydroxide ion. Ammonium ion is referred to as the conjugate acid of ammonia because it can lose a hydrogen ion to reform ammonia.

Solutions that can resist pH changes upon the addition of limited quantities of an acid or base are called buffers. Buffer solutions may contain either an acid and conjugate base or a base and conjugate acid. The presence of both species (acid/conjugate base or base/conjugate acid) permits the unique properties of these solutions. Buffers can be acidic, neutral, or basic.

Buffers can be prepared in the laboratory by combining the acid or base species with a salt consisting of the corresponding conjugate acid or conjugate base. In the case of acetate buffer, acetic acid is combined with an acetate salt such as sodium acetate. The quantity of each component used can be obtained from reference literature or calculated (1). Standard buffer "recipes" are available for numerous buffer compositions and pH ranges (2).

Alternatively, buffers can be purchased commercially. These are typically available in pH values from 1 to 14. The accuracy with which a buffer is prepared can depend upon the manufacturer and composition of the buffer. The National Institute of Standards and Technology (NIST) sells salts for preparing buffer standards with an uncertainty of ±0.005 pH units. Other manufacturers typically sell buffers having uncertainties ranging from 0.01 to 0.02 pH units. Most buffers will state that they are traceable to NIST standards. This means that during the manufacture of the buffer, its value is standardized using NIST buffers.

The pH values of buffer solutions depend on temperature. For example, the pH value for the standard NIST borax buffer at 25°C is 9.180. At 50°C, the pH decreases to 9.011 (3). Commercially available buffers usually state the pH value at 25°C. pH values for other temperatures can be obtained from the buffer manufacturer.

pH MEASUREMENT

There are two well-accepted methods for determining solution pH: colorimetry and sensor-based technology. Each method has advantages and disadvantages, and each has utility based primarily on the application and desired accuracy of the determination.

Colorimetry

The colorimetric determination of pH employs dye compounds, called indicators, which change color depending upon solution pH. The dyes used exhibit colors in the visible spectrum and can undergo one or more color changes as a function of pH. The indicator can be immobilized on a paper substrate (pH paper) or can be prepared as a solution using water or water combined with an organic solvent. The simplest indicators are those that undergo only a single color change such as litmus paper or phenolphthalein solution. For example, acidic solutions cause the color of litmus paper to change from blue to red whereas basic solutions cause phenolphthalein solution to change from colorless to red (or pink). Table 2 lists several acid–base indicators along with their color changes as a function of pH.

The pH paper method offers greater utility because the immobilized indicators can undergo several different color changes across the full pH scale. Some papers are designed to indicate only a single unit pH change; others can indicate changes of 0.1 or 0.2 pH units.

Sensor-Based Measurement

These technologies employ sensors that either generate a potential (glass pH sensors and solid-state sensors) or use

Table 2. Compounds Used to Indicate Solution pH

Indicator	Transition pH	Acid Color	Base Color
Methyl violet	0.0–1.6	Yellow	Blue
Cresol purple	1.2–2.8	Red	Yellow
Methyl orange	3.1–4.4	Red	Orange
Congo red	3.0–5.0	Violet	Red
Methyl red	4.8–6.0	Red	Yellow
Litmus	5.0–8.0	Red	Blue
Phenolphthalein	8.0–9.6	Colorless	Red
Thymolphthalein	8.3–10.5	Colorless	Blue
Alizarin yellow	10.1–12.0	Yellow	Orange red

optical principles to determine the pH. Signal conditioning electronics are required in addition to the sensor for measurement, display, and data transmission. Both types of sensor technology can be used in the laboratory or in the field.

Glass pH Sensor. A glass pH electrode consists of two electrode elements; a silica glass membrane that selectively responds to hydrogen ions and a reference element. The two elements can be separate or combined as one into a single sensor, as shown in Fig. 1. When the elements are combined, the unit is referred to as a combination pH electrode. The combined electrode and the separate pH and reference elements are available commercially.

pH Glass Element. A pH glass element consists of a short section (<1 cm) of special glass that produces a voltage in the presence of hydrogen ions. This section of glass is attached to a shaft of lead glass that can be 10 or more centimeters long. The bulk of the membrane glass matrix usually consists of silicon dioxide. Other compounds, such as oxides of lithium, calcium, lanthanum, cesium, and uranium, are added to provide unique characteristics such as low resistance, high selectivity to hydrogen ions, and high strength. When immersed in aqueous solution, the glass matrix forms a hydrated layer, often called a gel layer, which is approximately 5 to 20 nm thick (4). The processes responsible for the sensing function of the glass can be viewed as an equilibrium binding of hydrogen ions within the gel layer (Fig. 2). The inside of the pH glass element is filled with a buffer saturated with silver chloride. Electrical contact is completed through a silver chloride coated silver wire immersed in the buffer. Usually the buffer is designed to have a pH of 7. Therefore, when the electrode is immersed in a pH 7 buffer, the potential difference across the glass membrane will be 0 volts versus the reference element. The shaft glass plays no role in the function of the electrode, aside from providing a substrate that is a good electrical insulator to which the membrane glass can be attached.

Reference Element. The reference element provides a reference potential against which the potential of the pH element is measured. The pH element and reference element form a complete electrical circuit when they are immersed together in the same solution. The reference is characterized by its ability to maintain a constant potential regardless of solution conditions. Additionally, the potential of the reference must not drift with time and should exhibit little or no hysteresis when the element undergoes a change in temperature. The reference element is simply a tube filled with an electrolyte into which a porous material called a frit is imbedded (Fig. 1). The frit allows the solution being tested to come in contact with the reference electrolyte. Usually, references are designed so that a small amount of the reference electrolyte will flow through the frit into the test solution. Ideally, there is no bulk flow of test solution into the reference

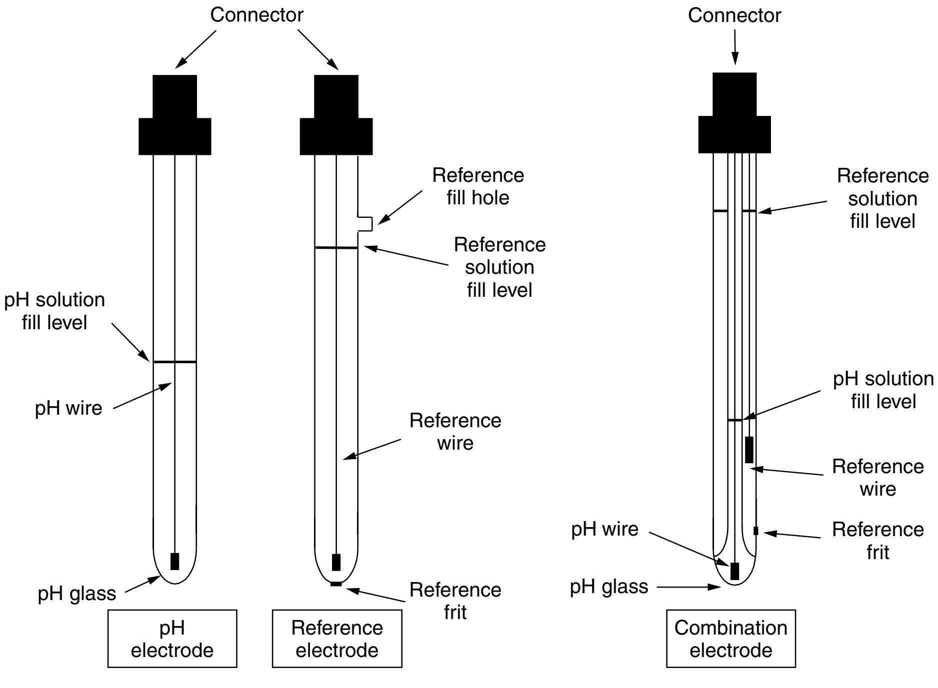

Figure 1. Separate pH and reference electrodes and the combination pH electrode.

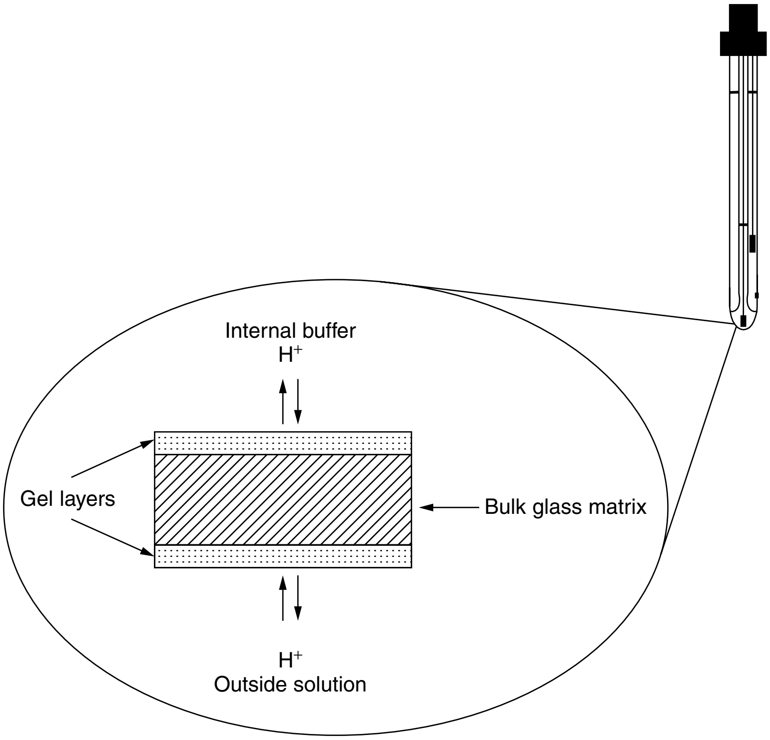

Figure 2. Representation of a cross section of a pH glass membrane showing gel layers and bulk glass membrane.

element. However, some test solution components may diffuse across the frit and contaminate the reference electrolyte after exposure for an extended period of time. The most common reference is based on a silver–silver chloride electrochemical cell. The system consists of a silver chloride coated silver wire immersed in a solution containing a high concentration of potassium chloride, usually either 3.5 M or saturated potassium chloride.

In use, the pH and reference elements are electrically connected to a high input impedance measuring device called a meter. Both the pH and reference elements are immersed to a depth great enough to ensure that the test solution completely covers the pH glass and the reference frit. The potential measured and displayed by the meter (E_{meter}) is the difference in potential between the two elements:

$$E_{\text{meter}} = E_{\text{pH}} - E_{\text{ref}} + E_J \tag{8}$$

where E_{pH} and E_{ref} are the potentials of the pH and reference elements, respectively. The other potential in the equation, E_J, is the reference junction potential. This potential results from the charge separation occurring due to differences in the migration rates of ions across the reference electrolyte/test solution interface. Reference electrodes are designed to keep this term small and constant, regardless of solution conditions.

The potential E_{meter} as a function of solution pH is given by the following expression;

$$E_{\text{meter}} = \left(\frac{2.303\,RT}{F}\right)(7 - \text{pH}) \tag{9}$$

where R is the ideal gas law constant (8.314 J/mol·K), T is the temperature in K, and F is Faraday's constant (96485 C/mol). At 25 °C, the slope term in ( ) has a value of 59.2 mV. This is the operational form of the Nernst equation applied to a glass pH electrode (see Reference 4 for a discussion of the Nernst equation). Equation 9 yields a plot that has a negative slope, as shown in Fig. 3.

When the pH is measured at temperatures different from the calibration temperature, a measurement error can result. This is due to the change in electrode response as a function of temperature. For this reason, practically all commercially available pH meters provide automatic temperature compensation (ATC). This feature corrects only for the effect of temperature on electrode response, not for the true temperature coefficient of a solution resulting from chemical equilibria shifts with temperature. For ATC, the temperature at calibration and measurement are either entered into the meter manually, or a temperature sensor sends this information to the meter automatically. The temperature sensor can be built into the pH probe, or it may be separate.

Calibration. The probe and meter are calibrated as a pair. During calibration, the meter relates the potential produced by the electrode immersed in a buffer to the pH value of the buffer. Calibration is performed either using standard buffer solutions or a sample solution of known pH value. In laboratory use, the calibration is generally done either daily or just before the system is used for a series of measurements. The time interval between calibrations in an industrial process depends upon how quickly solution conditions degrade electrode performance. The interval can be as often as every few days or as little as a month or more. Detailed procedures for the proper use and calibration of glass pH electrodes can be found in reference sources (5,6).

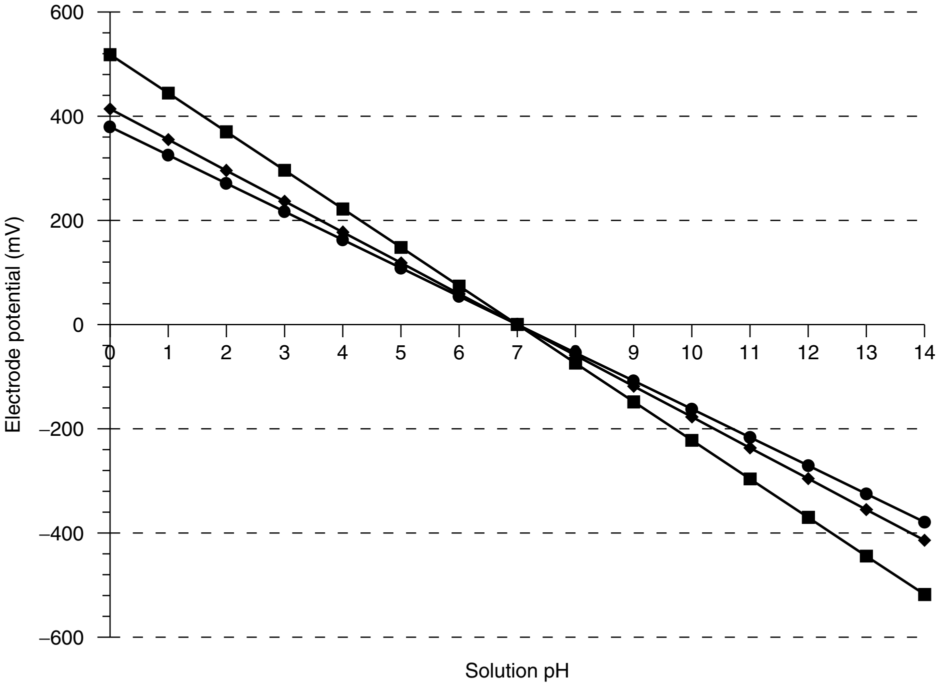

Figure 3. pH electrode potential as a function of solution pH at 0 °C (●), 25 °C (♦), and 100 °C (■).

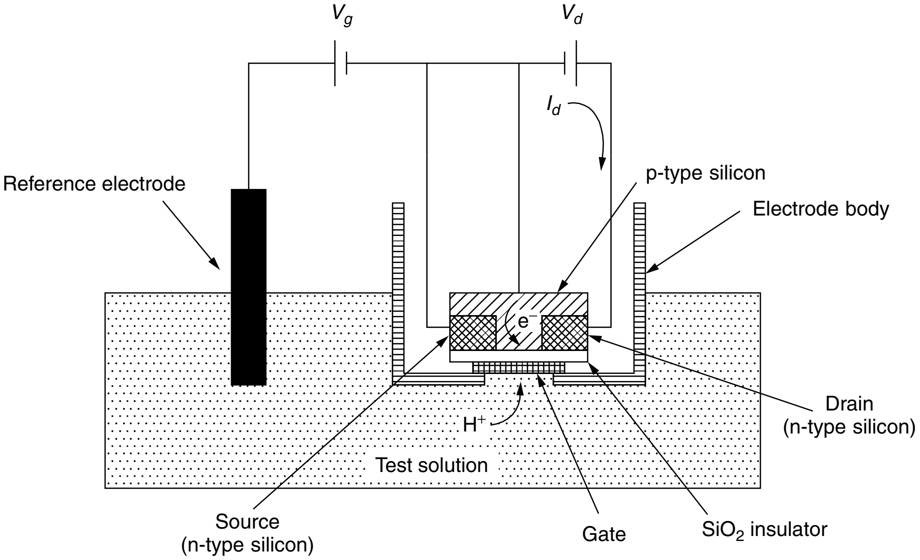

Figure 4. Typical arrangement of an ISFET sensor used for pH measurement. V_g and V_d are the gate and drain voltages, respectively. I_d is the drain current.

Solid State Sensors. More recently, pH sensitive electrodes have been developed based on metal-oxide semiconductor field effect transistor (MOSFET) technology. The sensors are similar in construction and operation to a traditional MOSFET, except that the gate is modified by an ion selective coating. The sensor is then referred to as an ion-selective field effect transistor (ISFET) (Fig. 4) (7–9).

The ISFET is enclosed in the electrode body so that only the gate is exposed to the solution. The sensors can be made to respond selectively to many different ions. For pH measurement, the gate is coated with either an oxide of silicon, aluminum, or tantalum or a nitride of silicon. In operation, a bias voltage is applied to the gate and the drain. Protons adsorbed to the gate alter the conductance of the p-type silicon above the gate. This results in current flow in the n-p-n region between the source and drain. Under these conditions, the drain current (I_d) is approximated by the following equation (10):

$$I_d = AV_d[V_c + V_g - E_{\mathrm{N}}(\mathrm{pH})] \tag{10}$$

where A is a constant, V_d is the drain potential, V_c is a combination of the threshold potential of the p-n junction and phase boundary potentials, V_g is the potential applied to the gate, and E_{N} is the Nernst potential (or slope, equal to 59.2 mV at 25 °C). Either the voltage or current generated at the ISFET can be monitored by the meter. If the gate voltage is monitored, then the drain current is kept constant, or conversely, if the drain current is monitored, the gate voltage is kept constant.

The ISFET sensor uses the same type of reference element as the glass electrode. Attempts have been made to modify the ISFET to act as a reference electrode (11). Thus far, work has not produced significant improvement over a conventional reference system.

Optical pH Sensors. The concept of pH determination using colorimetry and absorbance spectrophotometry has been integrated into a sensor that is the size of a conventional laboratory glass pH electrode (12). The measurement is based on using indicator dyes immobilized in a polymer film. Several different indicator films are available, depending upon the pH range desired for the measurement.

BIBLIOGRAPHY

1. Ebbing, D.D. (1993). *General Chemistry*, 4th Edn. Houghton Mifflin Company, Boston, pp. 695–696.
2. Perrin, D.D. and Dempsey, B. (1974). *Buffers for pH and Metal Ion Control*. Chapman and Hall, London.
3. Bates, R.G. (1964). *Determination of pH: Theory and Practice*. John Wiley & Sons, Hoboken, NJ, p. 76.
4. Galster, H. (1991). *pH Measurement: Fundamentals, Methods, Applications and Instrumentation*. VCH, Weinheim, p. 117.
5. ASTM Standard E70-97. (1998). *Standard Test Method for pH of Aqueous Solutions with the Glass Electrode*.
6. Eaton, A.D., Clesceri, L.S. and Greenberg, A.E. (Eds.). (1995). *Standard Methods for the Examination of Water and Wastewater*, 19th Edn. American Public Health Association, Washington, DC, pp. 4-65–4-69.
7. Ref. 4, pp. 189–195.
8. Zemel, J.N. (1975). *Anal. Chem.* **47**: 255A.
9. Wohltjen, H. (1984). *Anal. Chem.* **56**: 87A.
10. Ref. 4, p. 192.
11. Ref. 4, pp. 193–194.
12. Ocean Optics, Inc. (2002). Product literature. Avialable: www.oceanoptics.com.

PERCHLOROETHYLENE (PCE) REMOVAL

Enos C. Inniss
University of Texas
San Antonio, Texas

INTRODUCTION

Perchloroethylene (C_2Cl_4) or PCE is also known by other names, tetrachloroethylene and perc. Often in chemistry, one might see the -ethylene replaced with -ethene. PCE is a chemically synthesized compound that has been used as a solvent since the 1940s because of its nonflammable nature. This characteristic eliminated many accidental fires in the workplace, which resulted in increased employee safety and reduced cost of operation. Since it is a volatile organic compound (VOC), PCE was considered "clean" because it quickly volatilizes when exposed to air. Appropriate applications of this chlorinated solvent include dry cleaning, metal degreasing, brake cleaning (replacing 1,1,1-trichloroethane), and production of HFC134a (a hydrofluorocarbon alternative to the chlorofluorocarbons or CFCs phased out as refrigerants in the early 1990s) (1).

By the early 1980s widespread use of PCE coupled with historically poor handling and disposal practices led to its regulation under the Safe Drinking Water Act Amendments of 1986 (SDWAA–1986) (2). Now a common groundwater contaminant and a priority pollutant, PCE has been observed to degrade both biologically and chemically.

The biological transformation process is reductive dechlorination, an anaerobic process where microorganisms use the chlorinated ethylenes as electron acceptors (3). Once reduced, the reduction products may also be further transformed by cometabolic oxidation, a gratuitous metabolic reaction under aerobic conditions where the microorganisms are growing on another substance and gain no energy from the cometabolized substance (4).

The chemical transformations may also be either reductions or oxidations. The reductions are achieved through interactions with zero valent metals (ZVMs). Oxidative processes include potassium permanganate or photochemical oxidation.

REDUCTIVE DECHLORINATION

Microbial metabolism requires that energy and carbon sources be readily available. Most sugars, a common hydrocarbon, have the ability to give electrons by simply breaking the carbon–hydrogen bond and are therefore referred to as electron donors. The highly oxidized nature of PCE (with all of the chlorine substitutions) makes it nearly impossible to further oxidize. This quality promotes its recalcitrance or persistence in the environment. Therefore, any transformations of this compound require that a reduction (with the chlorinated ethylene serving as the electron acceptor) first take place. In the presence of microorganisms, this process is referred to as reductive dechlorination (Fig. 1). Under anaerobic conditions a number of identified microorganisms with the proper dehalogenase enzyme are expected to be able to sequentially reduce PCE to trichloroethylene (TCE), then to several isomers of dichloroethylene (DCE), then to vinyl chloride (VC), and finally to ethylene (ETH), where it could be taken up by microorganisms and plants as part of regular metabolic pathways (7).

Because of the significant energy required to push this reaction pathway to the innocuous ethylene product, the process is more likely to halt at either DCE or VC production. Note that the higher the potential ($E^{\circ\prime}$) the more likely the compound will be reduced. All of the chlorinated ethylenes (PCE, TCE, DCE, and VC) are considered priority pollutants under SDWAA–1986. So an additional transformation step is usually required to achieve removal of all contaminants of concern.

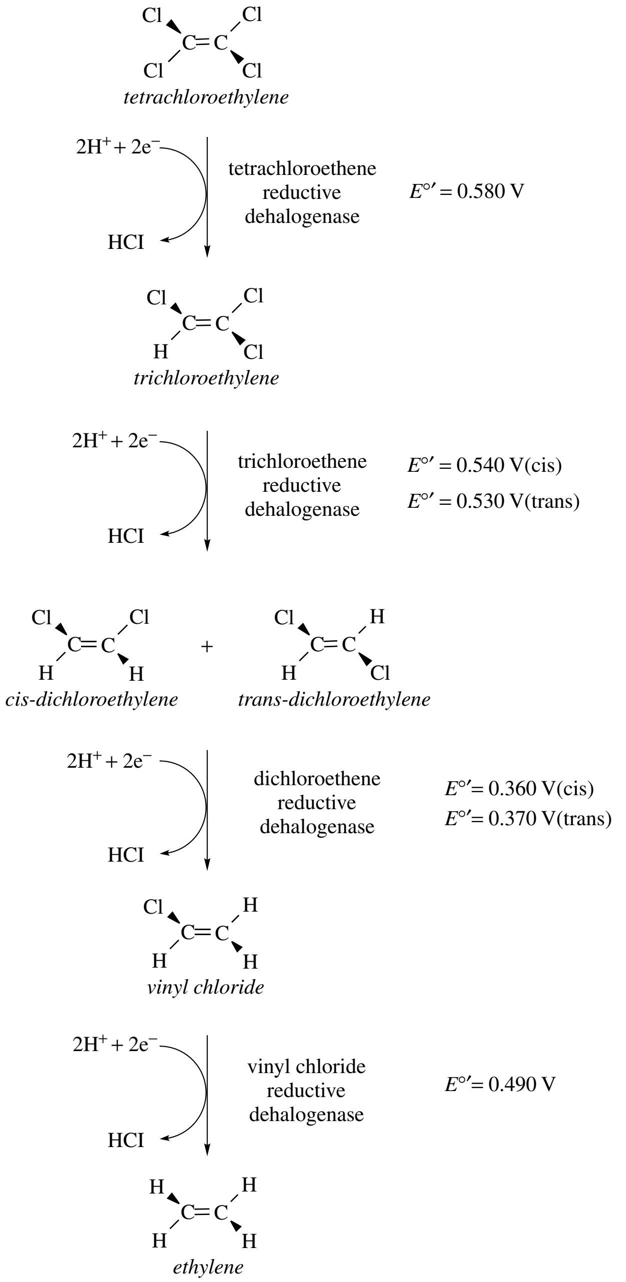

Figure 1. The biologically mediated reductive dechlorination process from perchloroethylene all the way to (the nonchlorinated) ethylene (5) with the corresponding standard reduction potentials, $E^{\circ\prime}$, for each half-reaction (6).

COMETABOLIC OXIDATION

Concerns for the incomplete reduction of the chlorinated ethylenes led to consideration of additional transformation pathways/mechanisms. Biologically, the reductive dechlorination may be forced to try to get beyond the VC production or a different pathway may be followed. Once PCE is initially reduced to TCE, then this and the remaining reduction products may more readily be cometabolically oxidized. The pathway shown for TCE (Fig. 2) oxidation is assumed to be similar for DCE and VC oxidation (8). The energetics of this process also suggest that the lesser chlorinated products are more easily oxidized by microorganisms possessing the appropriate oxygenase enzymes.

The commonality of the enzymes is that they do chemistry on the carbon nucleus of the compound, weakening the carbon–carbon double bond. These transformations are further considered cometabolic because the enzymes are produced to do chemistry on compounds with similar chemical structures to that of the chlorinated ethylenes (Fig. 3). Sustainability of this process therefore requires the addition of both the similar target compound (e.g., phenol) and reducing equivalents [e.g., NAD(P)H] as cometabolites (9).

The resulting products from this cometabolic oxidation are innocuous compounds that can readily be incorporated into typical metabolic function of microorganisms.

METAL REDUCTION

Similar to the biologically mediated reductive dechlorination, zero valent metals (ZVMs, typically granular iron) when they corrode either produce hydrogen or serve as an electron donor directly (10). ZVMs are used in permeable reactive barriers (PRBs) as well as in canister filters to allow for on-site treatment.

CHEMICAL OXIDATION

Photocatalytic oxidation uses a titanium dioxide catalyst, which is excited by the UVA spectrum (315–400 nm) of ultraviolet light to produce hydroxyl radicals. The process is used with volatilized contaminants at ambient temperature and pressure. Therefore, photocatalytic oxidation is an appropriate treatment option to use in conjunction with either air stripping or soil vapor extraction (11).

Potassium permanganate is another stronger oxidizer that rapidly degrades chlorinated ethylenes either in solution or in soils. This compound, similar to the biological oxygenase enzymes, attacks and weakens the carbon–carbon double bond (12).

CONCLUSION

There are both chemical and biological treatment options available for removal of PCE. Depending on cost, availability, and preference of options, one can choose from biologically mediated reductive dechlorination of PCE and its reduction products. Recognizing the likely limitation in achieving complete transformation to innocuous compounds, one could switch to biologically mediated cometabolic oxidation. Or depending on the form the chlorinated ethylenes are in, solubilized versus volatilized, a chemical reduction (ZVMs) or oxidation process (photocatalytic or potassium permanganate) may be chosen.

BIBLIOGRAPHY

1. HSIA (Halogenated Solvents Industry Alliance, Inc.) (1999). Perchloroethylene white paper. Available at http://www.hsia.org/white_papers/perc.htm.

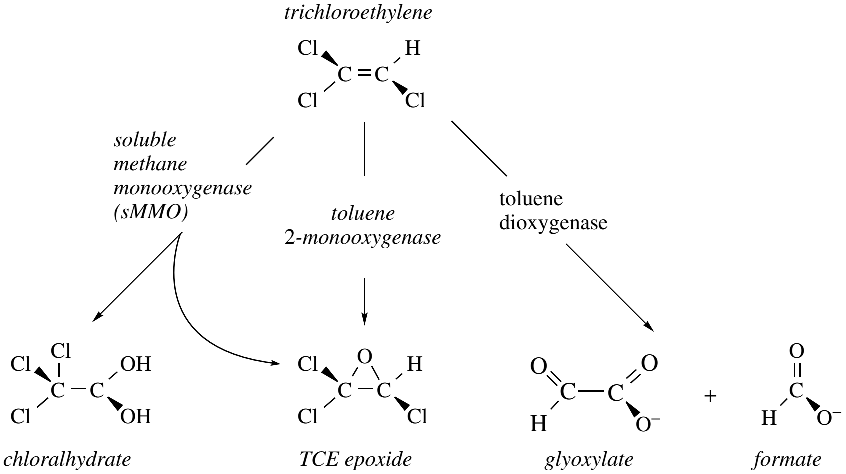

Figure 2. The first step of biologically mediated chlorinated ethylene oxidation (8).

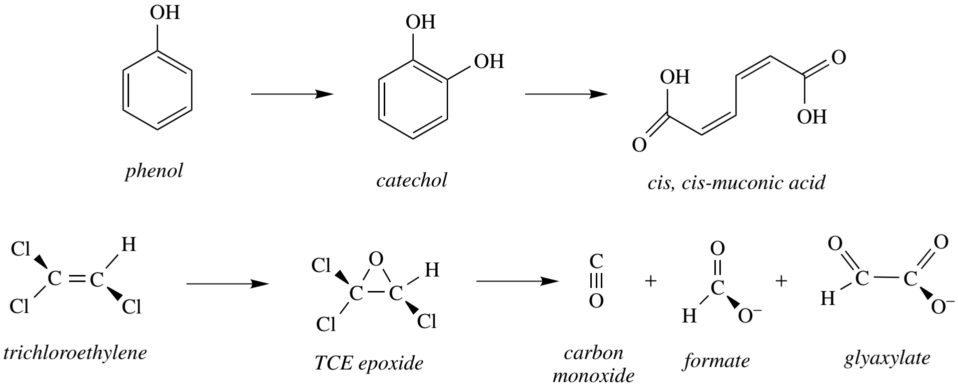

Figure 3. Illustration of the cometabolic nature of chlorinated ethylene oxidation where chemistry similar to that performed on the target compound (phenol) is also done on the chlorinated compound.

2. Safe Drinking Water Act Amendments (1986). Conference report (to accompany S.124), United States Congress (99th House, 2nd Session:1986), U.S. Government Printing Office, Washington, DC.
3. McCarty, P.L. (1994). An overview of anaerobic transformation of clorinated solvents. In: *Symposium on Intrinsic Bioremediation of Ground Water*, U.S. Environmental Protection Agency, Washington, DC, EPA 540/R-94/515, pp. 135–142.
4. Atlas, R.M. and Bartha, R. (1993). *Microbial Ecology: Fundamentals and Applications*, 3rd Edn. Benjamin/Cummings Publishing, Redwood City, CA.
5. Ellis, L. (updated April 2000). Anaerobic tetrachloroethylene graphic pathway map, UMBBD (The University of Minnesota Biocatalysis/Biodegradation Database). Available at http://www.labmed.umn.edu/umbbd/tce2/tce2_image_map.html.
6. Vogel, T.M., Criddle, C.S., and McCarty, P.L. (1987). Transformations of halogenated aliphatic compounds. *Environ. Sci. Technol. Crit. Rev.*, 722–736.
7. Ellis, L. (updated April 2000). Tetrachloroethylene pathway map (anaerobic), UMBBD (The University of Minnesota Biocatalysis/Biodegradation Database). Available at http://www.labmed.umn.edu/umbbd/tce2/tce2_map.html.
8. Jung Oh, D. (updated April 2000). TCE graphical pathway map, UMBBD (The University of Minnesota Biocatalysis/Biodegradation Database). Available at http://www.labmed.umn.edu/umbbd/tce/tce_image_map.html.
9. Whittaker, M., Monroe, D., and Jung Oh, D. (updated April 2000). Trichloroethylene pathway map, UMBBD (The University of Minnesota Biocatalysis/Biodegradation Database). Available at http://www.labmed.umn.edu/umbbd/tce/tce_map.html.
10. Tratnyek, L. (updated September 2002). Chemistry of contaminant reduction with zero-valent iron metal, CGR (Center for Groundwater Research). Available at http://cgr.ese.ogi.edu/iron/.
11. Hager, S., Bauer, R., and Kudielka, G. (2000). Photocatalytic oxidation of gaseous chlorinated organics over titanium dioxide. *Chemosphere* **41**: 1219–1225.
12. Huang, K.-C. et al. (2001). Oxidation of chlorinated ethenes by potassium permanganate: a kinetics study. *J. Hazardous Mater.* **B87**: 155–169.

A PRIMER ON WATER QUALITY

Gail E. Cordy
U.S. Geological Survey

WHAT IS IN THE WATER?

Is it safe for drinking? Can fish and other aquatic life thrive in streams and lakes that are affected by human

---

activities? What is the water quality? To answer these questions, it is helpful to understand what "water quality" means, how it is determined, and the natural processes and human activities that affect water quality.

WHAT DO WE MEAN BY "WATER QUALITY"?

Water quality can be thought of as a measure of the suitability of water for a particular use based on selected physical, chemical, and biological characteristics. To determine water quality, scientists first measure and analyze characteristics of the water such as temperature, dissolved mineral content, and number of bacteria. Selected characteristics are then compared to numeric standards and guidelines to decide if the water is suitable for a particular use.

HOW IS WATER QUALITY MEASURED?

Some aspects of water quality can be determined right in the stream or at the well. These include temperature, acidity (pH), dissolved oxygen, and electrical conductance (an indirect indicator of dissolved minerals in the water). Analyses of individual chemicals generally are done at a laboratory.

WHY DO WE HAVE WATER-QUALITY STANDARDS AND GUIDELINES?

Standards and guidelines are established to protect water for designated uses such as drinking, recreation, agricultural irrigation, or protection and maintenance of aquatic life. Standards for drinking-water quality ensure that public drinking-water supplies are as safe as possible. The U.S. Environmental Protection Agency (USEPA) and the States are responsible for establishing the standards for constituents in water that have been shown to pose a risk to human health. Other standards protect aquatic life, including fish, and fish-eating wildlife such as birds.

HOW DO NATURAL PROCESSES AFFECT WATER QUALITY?

Natural water quality varies from place to place, with the seasons, with climate, and with the types of soils and rocks

through which water moves. When water from rain or snow moves over the land and through the ground, the water may dissolve minerals in rocks and soil, percolate through organic material such as roots and leaves, and react with algae, bacteria, and other microscopic organisms. Water may also carry plant debris and sand, silt, and clay to rivers and streams making the water appear "muddy" or turbid. When water evaporates from lakes and streams, dissolved minerals are more concentrated in the water that remains. Each of these natural processes changes the water quality and potentially the water use.

WHAT IS NATURALLY IN THE WATER?

The most common dissolved substances in water are minerals or salts that, as a group, are referred to as dissolved solids. Dissolved solids include common constituents such as calcium, sodium, bicarbonate, and chloride; plant nutrients such as nitrogen and phosphorus; and trace elements such as selenium, chromium, and arsenic.

In general, the common constituents are not considered harmful to human health, although some constituents can affect the taste, smell, or clarity of water. Plant nutrients and trace elements in water can be harmful to human health and aquatic life if they exceed standards or guidelines.

Dissolved gases such as oxygen and radon are common in natural waters. Adequate oxygen levels in water are a necessity for fish and other aquatic life. Radon gas can be a threat to human health when it exceeds drinking-water standards.

HOW DO HUMAN ACTIVITIES AFFECT WATER QUALITY?

Urban and industrial development, farming, mining, combustion of fossil fuels, stream-channel alteration, animal-feeding operations, and other human activities can change the quality of natural waters. As an example of the effects of human activities on water quality, consider nitrogen and phosphorus fertilizers that are applied to crops and lawns. These plant nutrients can be dissolved easily in rainwater or snowmelt runoff. Excess nutrients carried to streams and lakes encourage abundant growth of algae, which leads to low oxygen in the water and the possibility of fish kills.

Chemicals such as pharmaceutical drugs, dry-cleaning solvents, and gasoline that are used in urban and industrial activities have been found in streams and ground water. After decades of use, pesticides are now widespread in streams and ground water, though they rarely exceed the existing standards and guidelines established to protect human health. Some pesticides have not been used for 20 to 30 years, but they are still detected in fish and streambed sediment at levels that pose a potential risk to human health, aquatic life, and fish-eating wildlife. There are so many chemicals in use today that determining the risk to human health and aquatic life is a complex task. In addition, mixtures of chemicals typically are found in water, but health-based standards and guidelines have not been established for chemical mixtures.

WHAT ABOUT BACTERIA, VIRUSES, AND OTHER PATHOGENS IN WATER?

The quality of water for drinking cannot be assured by chemical analyses alone. The presence of bacteria in water, which are normally found in the intestinal tracts of humans and animals, signal that disease-causing pathogens may be present. Giardia and cryptosporidium are pathogens that have been found occasionally in public-water supplies and have caused illness in a large number of people in a few locations. Pathogens can enter our water from leaking septic tanks, wastewater-treatment discharge, and animal wastes.

HOW CAN I FIND OUT MORE ABOUT MY WATER QUALITY?

Contact your local water supplier and ask for information on the water quality in your area. The USEPA requires public-water suppliers to provide water-quality data to the public on an annual basis in an understandable format. State agencies that deal with health, environmental quality, or water resources also can provide information on the quality of your water. Additional resources can be found on the Internet at:

http://water.usgs.gov/nawqa
http://www.epa.gov/safewater

OVERVIEW OF ANALYTICAL METHODS OF WATER ANALYSES WITH SPECIFIC REFERENCE TO EPA METHODS FOR PRIORITY POLLUTANT ANALYSIS

Ralph J. Tella
Lord Associates, Inc.
Norwood, Massachusetts

INTRODUCTION

This article is an overview of the analytical chemistry behind the analyses of common environmental contaminants. Specific reference is made to the U.S. Environmental Protection Agency (EPA) "priority pollutants." The list of priority pollutants was first published in the *Federal Register* in 1980 as the list of toxic pollutants for which categorically defined industries were required to analyze their effluent streams to obtain wastewater discharge permits under the National Pollution Discharge Elimination System (NPDES).

To require industry to analyze for pollutants, the EPA was compelled to develop standardized test procedures. As scientists and engineers became familiar with these procedures, the priority pollutant list developed into a commonly requested set of analyses used to characterize and evaluate the toxicity of various environmental media including soils, groundwater, and air.

The priority pollutant list is comprised of several classes of compounds or "fractions," grouped according to similar chemical and physical characteristics. A complete list of the priority pollutants is provided in Table 1. These fractions include 111 organic compounds, and 15 inorganic elements and compounds, a total of 126 parameters.

Note that these 126 parameters are not the only toxic pollutants in the world. Rather, they are only a convenient list of common pollutants for which there are standardized test methods. Many more pollutants can be analyzed for, some use the same analytical techniques. However, to gain this information from the laboratory, you must specifically ask for it. Many commercial laboratories have developed economical techniques, which "screen" for these nonpriority pollutants, such as "library searches," or "extended runs." Many of these techniques are, however, only semiquantitative and qualitative.

Because petroleum products in the aqueous environment comprise a large focus of attention in the field of environmental restoration, this article also describes various methods of analyses for petroleum products. Note that only some components of petroleum are regulated as EPA "priority pollutants."

GAS CHROMATOGRAPHY

Gas chromatography (GC) is a powerful and highly versatile instrumental method of analysis that was developed in the early 1950s. The number of technological advances in the technique since then has been so tremendous that a complete and detailed discussion of the technique is beyond the scope of this article. Rather, this discussion is limited to those techniques that are commonly used in environmental analyses.

Theory

This is the basic process by which a GC works. A sample is introduced into the instrument in a vapor state, and the various gaseous components are separated so that they can be individually identified and quantitatively measured. A schematic diagram of this process is provided in Fig. 1. The basic components are a gas cylinder with a reducing valve, a constant pressure regulator, a port for injecting the sample, a chromatographic column, a detector, an exit line, and a data collection device such as a strip chart recorder or computer.

The gas cylinder contains a carrier gas such as hydrogen, helium, or nitrogen, which is continuously swept through the chromatographic column at a constant temperature and flow rate. A small sample for analysis is injected, usually with a syringe, into the sample port, where it is flash-evaporated to convert its components into a gas. The constantly flowing stream of carrier gas carries the gaseous constituents through the chromatographic column. The gases travel through at different rates so that they emerge from the column at different times. Their presence in the emerging carrier gas is detected by chemical or physical means, and the response of the detector is fed to the data collection device. A typical chromatogram produced by this process is shown in Fig. 2. Each peak represents a specific chemical compound or mixture of compounds that has the same rate of movement through the column. The time for each compound to emerge from the column is a characteristic of the compound known as its retention time. The area under the peak is proportional to the concentration of the compound in the sample.

Standard packed chromatographic columns are generally glass or metal tubes varying from about 1–10 m in length and about 3–6 mm in diameter. The tube is packed with a inert solid impregnated with a nonvolatile liquid, such as silicone oil or polyethylene glycol. As a gaseous sample passes through the column, its components are partitioned between the stationary liquid phase of the column and the moving gas phase. Components that are relatively soluble in the liquid phase move through the column at slower rates than components that are not so soluble. Most gaseous materials can be separated by selecting suitable stationary phases and column lengths. More recently, capillary columns with diameters of 0.2–0.4 mm and lengths of 20–30 m have replaced standard packed columns. These columns contain the solid phase on the inside wall of the capillary tube. The gas containing the organic components passes through the center of the capillary tube, and the organics partition themselves between the gas and the stationary phase, emerging from the column at different times.

Table 1. 126 EPA Priority Pollutants

A. Chlorinated benzenes
Chlorobenzene
1,2-Dichlorobenzene
1,3-Dichlorobenzene
1,4-Dichlorobenzene
1,2,4-Trichlorobenzene
Hexachlorobenzene

B. Chlorinated ethanes
Chloroethane
1,1-Dichloroethane
1,2-Dichloroethane
1,1,1-Trichloroethane
1,1,2-Trichloroethane
1,1,2,2-Tetrachloroethane
Hexachloroethane

C. Chlorinated phenols
2-Chlorophenol
2,4-Dichlorophenol
2,4,6-Trichlorophenol
Parametachlorocresol
(4-chloro-3-methyl phenol)

D. Other chlorinated organics
Chloroform (trichloromethane)
Carbon tetrachloride
(tetrachloromethane)
Bis(2-chloroethoxy)methane
Bis(2-chloroethyl)ether
2-Chloroethyl vinyl ether (mixed)
2-Chloronaphthalene
3,3-Dichlorobenzidine
1,1-Dichloroethylene
1,2-*trans*-Dichloroethylene
1,2-Dichloropropane
1,2-Dichloropropylene
(1,3-dichloropropene)
Tetrachloroethylene
Trichloroethylene
Vinyl chloride (chloroethylene)
Hexachlorobutadiene
Hexachlorocyclopentadiene
2,3,7,8-Tetrachloro-dibenzo-*p*-dioxin
(TCDD)

E. Haloethers
4-Chlorophenyl phenyl ether
2-Bromophenyl phenyl ether
Bis(2-chloroisopropyl) ether

F. Halomethanes
Methylene chloride (dichloromethane)
Methyl chloride (chloromethane)
Methyl bromide (bromomethane)
Bromoform (tribromomethane)
Dichlorobromomethane
Chlorodibromomethane

G. Nitrosamines
N-Nitrosodimethylamine
N-Nitrosodiphenylamine
N-Nitrosodi-*n*-propylamine

H. Phenols (other than chlorinated)
2-Nitrophenol
4-Nitrophenol
2,4-Dinitrophenol
4,6-Dinitro-*o*-cresol
(4,6-dinitro-2-methylphenol)
Pentachlorophenol
Phenol
2,4-Dimethylphenol

I. Phthalate esters
Bis(2-ethylhexyl)phthalate
Butyl benzyl phthalate
Di-*n*-butyl phthalate
Di-*n*-octyl phthalate
Diethyl phthalate
Dimethyl phthalate

J. Polnuclear aromatic hydrocarbons
(PAHS)
Acenaphthene
1,2-Benzanthracene (benzo(*a*)
anthracene)
Benzo(*a*)pyrene (3,4-benzo-pyrene)
3,4-Benzofluoranthene (benzo(*b*)
fluoranthene)
11,12-Benzofluoranthene (benzo(*k*)
fluoranthene)
Chrysene
Acenaphthalene
Anthracene
1,12-Benzoperylene (benzo(*ghi*)
perylene)
Fluorene
Fluoranthene
Phenanthrene
1,2,5,6-Bibenzanthracene (dibenzo(*ah*)
anthracene)
Indeno (1,2,3-*cd*) pyrene
(2,3-*o*-phenylene pyrene)
Pyrene

K. Pesticides and metabolites
Aldrin
Dieldrin
Chlordane (technical mixture and
metabolites)
Alpha-endosulfan
Beta-endosulfan
Endosulfan sulfate
Endrin
Endrin aldehyde
Heptachlor
Heptachlor epoxide
(BHC-hexachlorocyclohexane)
Alpha-BHC
Beta-BHC
Gamma-BHC (Lindane)
Delta-BHC Toxaphene

L. DDT and metabolites 4,4-DDT
4,4-DDE (p,p-DDX)
4,4-DDD (p,p-TDE)

M. Polychlorinated biphenyls (PCBS)
PCB-1242 (Arochlor 1242)
PCB-1254 (Arochlor 1254)
PCB-1221 (Arochlor 1221)
PCB-1232 (Arochlor 1232)
PCB-1248 (Arochlor 1248)
PCB-1260 (Arochlor 1260)
PCB-1016 (Arochlor 1016)

N. Other organics
Acrolein
Acrylonitrile
Benzene
Benzidine
2,4-Dinitrotoluene
2,6-Dinitrotoluene
1,2-Diphenylhydrazine
Ethylbenzene
Isophorone
Naphthalene
Nitrobenzene toluene

O. Inorganics
Antimony
Arsenic
Asbestos
Beryllium
Cadmium
Chromium
Copper
Cyanide
Total lead
Mercury
Nickel
Selenium
Silver
Thallium
Zinc

The absence of packing minimizes the dispersion of the contaminants so that excellent resolution of the different compounds is obtained, even when there are 20 or more in the sample. This makes it possible to identify a greater number of compounds in highly complex mixtures.

The temperature of the sample and column must be held sufficiently high during movement through the column and subsequent detection to maintain them in the gaseous state. Only materials that can be volatilized or converted to volatile compounds can be detected by gas chromatography.

Sample Introduction Techniques

Samples may be introduced into the GC system in only two phases: liquid and gaseous. For successful liquid phase introduction, the injection port and column temperature

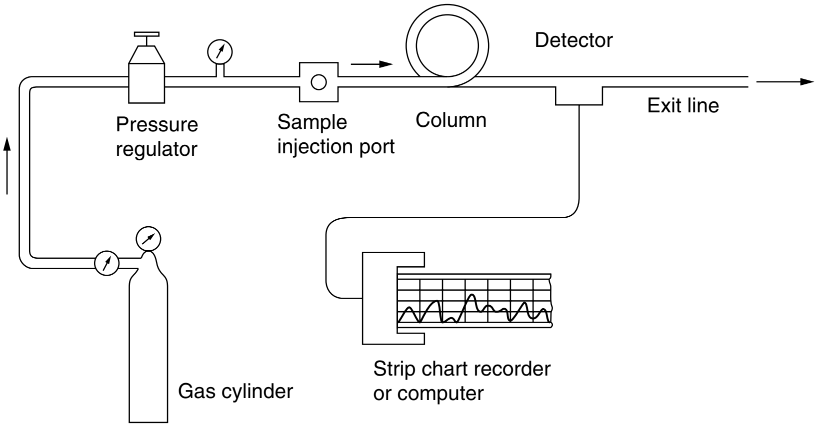

Figure 1. Schematic of a gas chromatograph system.

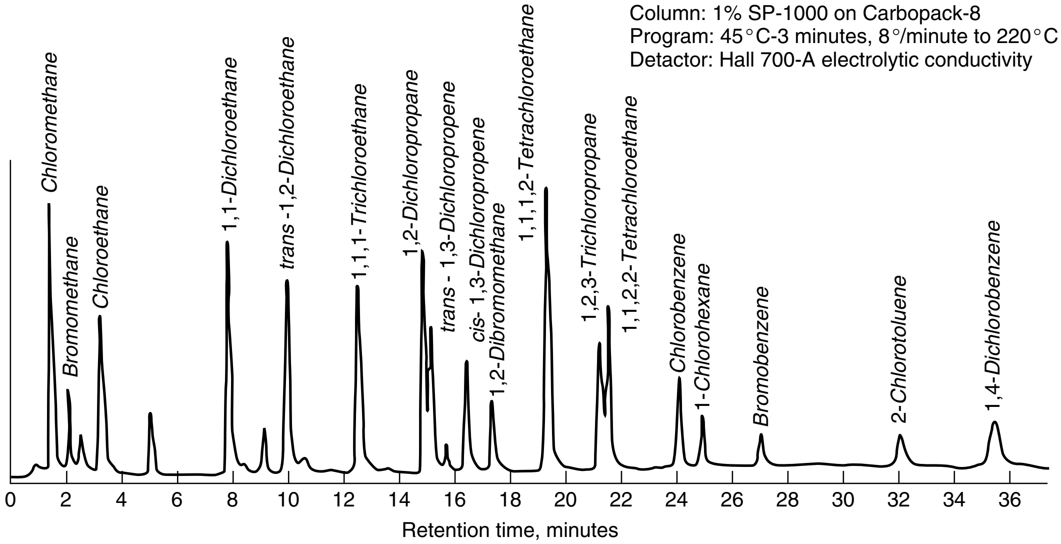

Figure 2. Gas chromatogram of purgeable halocarbons.

must be sufficiently high to permit complete volatilization of the material to be analyzed. A solid sample may be introduced by dissolving the sample in a low boiling solvent, thereby treating the sample as a liquid. Gas samples may be either directly injected via a gastight syringe or by using a gas sampling valve that uses the principle of having the carrier gas purge, under pressure, a calibrated volume loop within which the gas sample was trapped.

A major consideration in sample analysis is the required level of sensitivity or limit of detection. A common way of increasing a procedure's detection limit is to concentrate the substance being analyzed (the analyte). Three common procedures are used in gas chromatography: purge and trap, headspace, and solvent extraction.

Purge and Trap. The purge and trap procedure is commonly used in environmental analysis to determine volatile organic compounds or "purgeables," such as in EPA Methods 624, 601, and 602. In each of these methods, the purge and trap apparatus must be similar (although trap materials differ). The purge and trap device consists of three separate pieces of equipment: the sample purger, the trap, and the desorber.

An inert gas is bubbled through the sample in a specially designed purging device at ambient temperature. The purgeables are efficiently transferred from the aqueous phase to the vapor phase. The vapor is swept through a sorbent column where the purgeables are trapped. After purging is completed, the sorbent column is heated and backflushed with the inert gas to desorb the purgeables into a GC column.

Headspace. The headspace technique is commonly used for analyses of solid samples, although it can as easily be used for liquids. In this technique, EPA Method 5020, a known volume or weight of sample is collected in a sealed glass container and allowed to equilibrate at 90 °C for 1 hour. A sample of the headspace gas is then withdrawn for analysis. The advantages of this technique are its speed of sample preparation and its ability to leave potentially interfering semivolatile compounds behind.

Solvent Extraction. Solvent extraction is any procedure by which a solvent is used to extract an analyte into a liquid phase for direct injection (with or without preconcentration) into a GC. Solvent extraction techniques include both liquid/liquid and solid/liquid procedures. Although these procedures work best for semivolatile and nonvolatile compounds, the EPA has established similar procedures for analyzing volatiles in solid materials.

A popular liquid/liquid technique is used in EPA Method 625 (or 8270) for acid/base/neutral extractable organics. In this method, a liquid sample is separately extracted under both acidic and base/neutral conditions. A variety of glassware and apparatus has been developed for use in this method, although the separatory funnel is probably the simplest and most commonly used.

Solid/liquid techniques typically employ either Soxhlet extraction glassware or ultrasound. A Soxhlet extractor is a device that ensures intimate contact of the sample matrix with the extraction solvent by continuously deluging the sample with solvent. Alternatively, the EPA has approved ultrasound (Method 3550) for analyzing certain solid matrices.

Air Sampling Techniques. A variety of techniques are used to extract and concentrate air pollutants. The two most common procedures use either solid or liquid trap materials. Liquid traps, called "impingers," are used when the analyte can be easily absorbed into a stable liquid phase, such as cyanide into sodium hydroxide.

Solid trap materials may range from simple paper filters to exotic polymers. The difficulty in using solid trap materials is getting the analyte off of the trap back into a gaseous phase. This can be accomplished by either eluting the trap material with a solvent or by thermal adsorption.

If the analysis does not require concentration, air samples may be collected in any inert gastight device of known volume such as gastight syringes, evacuated glass bulbs, inert bags, or canisters. The methods most commonly used in the environmental industry are the EPA "TO" series of methods or others adapted from NIOSH approved procedures.

Detectors

The photoionization detector (PID) has become well known as a useful tool in the environmental industry. It is a nondestructive detector that lends itself to series operation with other detectors; thereby it enhances the selectivity and sensitivity of the GC.

The PID is the required detector for EPA Method 602 (or 8020/8021) for purgeable aromatics. Because it does not require the use of a fuel gas, it has become a popular detector for handheld portable organic vapor analyzers.

The PID operates on the principle of photoionization. The fact that certain molecules can be ionized by light is related to their characteristic molecular structure. The PID consists of a high-energy ultraviolet lamp located adjacent to an ionization chamber that contains a pair of electrodes. When a positive potential is applied to one electrode, the field created collects any ions formed by the absorption of UV light at the other electrode. The measured current is proportional to the concentration.

A PID responds to all compounds whose ionization potentials are less than or equal to the energy of the UV light source. Typical UV lamp energies in use are 10.2 eV and 11.7 eV. The magnitude of the ionization potential does not by itself dictate the sensitivity of the PID response. The degree of ionization depends on molecular structure. The number, type, and structure of the carbon bond, therefore, dictate the sensitivity of the detector to the compound. In general, the PID shows the greatest response to condensed-ring aromatic compounds (i.e., polyaromatic hydrocarbons), followed by aromatic compounds. These are other general selectivity rules:

- aromatics > alkanes > alkenes
- ketones > aldehydes > esters > alcohols > alkanes
- cyclic compounds > noncyclic
- branched compounds > nonbranched
- iodine-substituted > bromine-substituted > chlorine-substituted> fluorine-substituted

Because of the PID's response to aromatic compounds, it is the detector of choice for EPA Method 602 (or 8020/8021). The aromatic compounds in petroleum-based products such as benzene, toluene, and xylene are easily detected by this procedure. The aliphatic compounds in petroleum-based products such as methane, pentane and octane are suppressed. The flame ionization detector responds well to these compounds.

The Flame Ionization Detector. The flame ionization detector (FID) is probably the oldest and most generally used GC detector. The FID responds to most organic compounds, which can be both a benefit and liability in its application. As a benefit, the FID can be nearly universally used in a number of applications. This property lends itself to "fingerprinting" analysis of complex mixtures such as petroleum products. As a liability, the FID may respond to so many compounds that individual compound separation in a complex mixture, may become difficult, if not impossible.

The FID operates on the principle of flame ionization. As gases are eluted from a column, they are burned in a tiny hydrogen flame. A device continuously monitors the electrical conductivity of this flame. As organic compounds ionize in the flame, this conductivity is altered, proportional to the concentration, and measured. Most FIDs have sufficient energy to ionize any organic species with an ionization potential (IP) of 15.4 eV or less.

The use of FIDs in portable instrumentation has been hampered by the problem of maintaining a stable flame and fuel gas requirements. However, now there are many popular field instruments that have overcome these obstacles and are frequently used where gases such as methane are of interest.

The Electron Capture Detector. The electron capture detector (ECD) is used to detect compounds that have halogen atoms or polar functional groups such as pesticides, herbicides, and PCBs. EPA Method 608 (or 8080) uses an ECD for organochlorine pesticides and PCBs. The ECD uses a small radioactive source such as nickel 63 or tritium to emit beta particles. The ECD measures the ability of compounds to capture these particles from the ion current. The reduced current is measured and converted into an electrical signal for detection.

Mass Spectroscopy. Mass spectroscopy is the most rapidly developing analytical technology today. When coupled with a gas chromatograph (GC/MS), it is a powerful tool for positive identification and quantification of a large number of organic compounds. Recent advances in computer technology have dramatically expanded these systems' capabilities, speed of analyses, and cost efficiency. EPA Methods 624 or 8260 (purgeables) and 625 or 8270 (acid/base/neutral extractables) use this technique.

The mass spectrometer sorts charged gas molecules or ions according to their masses. The substance to be analyzed is vaporized in the GC and converted to positive ions or ion fragments by bombardment with rapidly moving electrons. These ions or ion fragments are pulled from the gas stream by an electrical field, accelerated, and separated by their mass-to-charge ratio. A computer then records the fragmentation pattern of the molecule, which is characteristic of a particular compound. The relative retention time and abundances of key ions and secondary ions are used to identify the compound qualitatively.

Most commercial labs have access to spectral libraries with information on more than 50,000 compounds. Compounds identified in this manner are referred to as "tentatively identified compounds" or "TICs." As the name implies, the data collected in this manner should be used with caution because an actual analytical reference standard has not been used. The quanitation of TICs may be orders of magnitude in error.

Other Detectors

Microcoulometric Detectors. Microcoulometric detectors such as the "Hall" detector are used for halide specific analyses such as EPA Method 601 (or 8010) for purgeable halocarbons. In this detector, halide-containing compounds are catalytically reduced using hydrogen, and the resulting "HX" is measured by a highly sensitive conductivity meter.

Thermal Conductivity. Thermal conductivity detectors use a heated wire that is placed in the stream of eluting gases. Each gas has a different thermal conductivity or ability to carry away heat from the wire and thus changes the temperature of the wire. These changes alter the resistivity of the wire from which a signal is generated.

Thermal conductivity detectors respond to a wide range of compounds but are relatively insensitive to low concentrations. For these reasons, they are usually used in the environmental industry for permanent gas analysis such as automobile and stack gas emissions. Many portable combustible gas meters also operate on the principle of thermal conductivity, such as units known as "explosimeters."

Flame Photometric Detectors. The flame photometric detector uses the optical emission of a gas burned in a flame to detect pollutants, typically sulfur and phosphorus containing compounds such as H_2S and phosphorus-based pesticides. The EPA has several interim pesticide and herbicide procedures that use this detection system.

Dry Electrolytic Conductivity Detectors. The dry electrolytic conductivity detectors (DELCD) are halogen specific. Often paired with other detectors such as FIDs, the DELCD can be used for highly selective identification of halogen containing compounds in gas streams.

HIGH PERFORMANCE LIQUID CHROMATOGRAPHY

High performance liquid chromatography (HPLC) is another analytical tool used for compounds not readily accessible by GC. As the name implies, HPLC uses a liquid carrier instead of a gas to sweep the sample through the column for separation. Compounds that are high in molecular weight (such as polynuclear aromatic hydrocarbons) or are thermally liable (i.e., decompose when heated) are good candidates for this technique. Many detectors are available to analyze both inorganic and organic species.

METALS ANALYSES

There are thirteen priority pollutant metals (see Table 1), and two EPA accepted methodologies used to analyze for them, atomic absorption spectroscopy (AA) and inductively coupled plasma (ICP). The former method uses the technique of absorption spectroscopy and the latter emission spectroscopy. Each technique has its advantages and disadvantages relative to selectivity, sensitivity, interferences, cost, and speed of analysis.

Atomic Absorption Spectroscopy

Atomic absorption spectroscopy (AA) is the most commonly used technique to analyze for metals. This technique has been around since the 1950s and was the first instrumental method of metal analysis accepted to replace what were previously colorimetric or gravimetric methods of analyses.

Because instrumental operating conditions vary from metal to metal, the EPA has published specific methods for each metal. Collectively, these methods are known as the "200 series" in water or the "7000" series in solid materials.

Theory. Atomic absorption is based on the principle that all atoms absorb light at particular wavelengths corresponding to the energy requirements of a particular atom. For example, copper atoms absorb light strongly at 324.7 nm because this particular wavelength has the right energy to alter the atom's electronic state from a ground state to an excited state. Other elements have their own specific wavelengths. Because each atom may have several different electronic states, a spectrum of these characteristic wavelengths can be created. The wavelength lines that are used in AAS are those that originate in the ground state, called resonance lines.

The basic AA system consists of a stable light source of particular wavelength, a flame or other excitation device, a

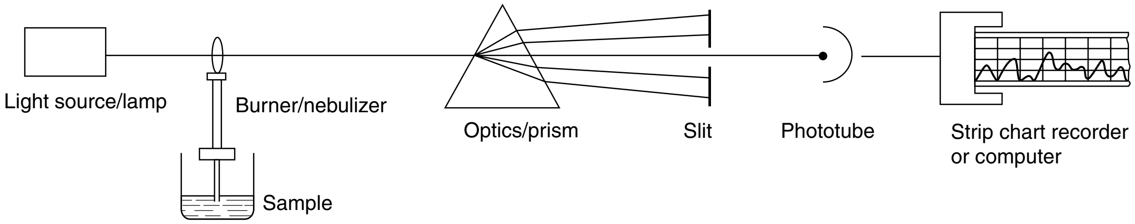

Figure 3. Schematic diagram of atomic absorption spectrometer.

prism to disperse and isolate emission lines, and a detector with appropriate amplifiers (see Fig. 3). The light source is directed into the flame or other excitation device, and its intensity is measured with and without the introduction of a sample. The decrease in intensity or absorption is proportional to the concentration. Two different excitation devices are commonly used. Flame AA is used to detect concentrations in the approximate parts per million range and flameless AA or a heated graphite furnace is used to detect in the parts per billion range. Flameless AA or a heated graphite furnace uses an electrically heated graphite tube to concentrate the atoms, instead of a flame.

Because mercury is unique in that it can be liberated in its ground state at room temperature a special cold vapor technique, EPA Method 245.1 should be used to measure low (ppb) concentrations. This technique uses a chemical reaction to liberate mercury atoms instead of a heated excitation source as in flame or graphite furnace techniques.

Inductively Coupled Plasma

Inductively coupled plasma (ICP) is another commonly used technique for metals analyses. The main attraction of ICP is that it can analyze many elements at once, and for certain refractory elements such as beryllium, it is much more sensitive.

However, AA must still be used to detect low levels of mercury and other metals. For these reasons, many laboratories that have ICP systems use both techniques to analyze for the 13 priority pollutant metals. Because of its multielement capabilities, ICP is also very useful in screening unknown samples. The EPA method for multielement ICP analysis is Method 6010.

Theory. The technique by which ICP operates is emission spectroscopy. Using a radio-frequency generated, argon sustained plasma, atoms are excited to a much higher energy state that can be achieved in normal AA systems. As the electrons fall back to their original state, they emit radiation at element specific or characteristic wavelengths. This radiation is in turn sensed by a detector; the signal is amplified and recorded. Prior to recent computer advances, this signal processing was often extremely time-consuming and subject to interferences.

Metal Sample Preparation

Nearly all samples require some preparation prior to metals analysis. This is required to isolate selectively the element you are analyzing for (the "analyte") from the rest of the sample (its "matrix") and to put it into a more readily analyzable state. This preparation usually begins in the field with the addition of nitric acid as a preservative. Nearly all aqueous samples for metals analysis should be preserved with nitric acid if they cannot be delivered to the laboratory on the same day.

The exception to this rule is discriminating between valence states of certain metals such as chromium. Samples of hexavalent chromium should not be preserved with acid. Although the chemistry of hexavalent chromium is not completely understood, it is believed that acid may promote the reduction of the hexavalent to the trivalent state.

Acidification of the sample helps to (1) promote digestion of the sample matrix and dissolution of the metal and (2) inhibits its precipitation or "plating" out on the walls of the sample container. Preserving metals samples with acid is not intended to inhibit microbial activity. It does not matter if microbes ingest or otherwise alter the metallic compounds because the sample will be digested and analyzed for total metals, regardless of valence state. For this same reason, samples for total metals analyses need not be refrigerated.

Samples submitted for dissolved metals should not be acidified until they have been filtered. Acidification would promote dissolution of the larger particles and bias the results. Note that "dissolved" metals is an empirically defined term. Specifically, "dissolved" metals mean all metals that are filterable (pass through) a 0.45-micron membrane filter.

"Total" metals mean the entire element irrespective of valence or physical state. To enhance the analysis and reduce the chance of matrix interferences, most metals samples undergo rigorous acid digestion. Through this procedure, solids and semisolid samples may be broken down into a dissolved aqueous state that allows the nebulizer of the instrument to aspirate or pass the sample into the flame.

Metals samples may also be treated or prepared through various solvent extraction or preconcentration techniques to enhance sensitivity and/or eliminate matrix effects. However, most of these techniques are not required by the modern instrumentation manufactured today.

CYANIDE AND TOTAL PHENOLICS

Cyanide and total phenolics are typically analyzed by colorimetry. These techniques involve lengthy wet chemical digestions and distillations to eliminate interference.

Once these steps are taken, the analyte is reacted chemically to produce a color, whose intensity is proportional to concentration. Various procedures exist that cover the variety of instrumentation manufactured to help automate these analyses. However, the most commonly used methods for these analyses are EPA Method 335.1 for cyanide and EPA Method 420.1 for total phenolics.

Sample Preparation

Cyanide. It is extremely important to preserve all aqueous samples for cyanide analysis with sodium hydroxide at a pH greater than 12 to ensure that the sample is not volatilized to hydrogen cyanide, an extremely toxic gas. In addition, the sample should be kept refrigerated.

Total Phenolics. Approved preservatives for total phenolics have varied in recent years. However, the generally agreed upon preservative is either phosphoric or sulfuric acid. Although not specifically required, copper sulfate may be added to inhibit microbial growth and initiate the analysis.

ASBESTOS

Although asbestos is a priority pollutant, it is rarely analyzed in water during routine or initial site investigations because its occurrence in groundwater is unusual. When it does occur, it is typically the result of localized dumping of asbestos insulation. Exceptions to this occur where asbestos fibers are locally mined, are part of an industrial process effluent, or when an asbestos cement pipe has deteriorated.

Because the focus of concern with asbestos is its previously common use as an insulator, it has been largely considered an industrial hygiene or indoor air pollution problem. As such, it has been the responsibility of the National Institute for Occupational Safety and Health (NIOSH) to promulgate standardized testing procedures for the material.

Generally, asbestos-containing material (ACM) contains chrysotile or amosite asbestos in a matrix of other fibers or particles such as glasswool or gypsum. This fibrous matrix makes positive identification of asbestos difficult, and it is best left to the experts. State laboratory certification for asbestos analysis is generally not available. Therefore, one should ascertain if the laboratory satisfactorily passed the latest EPA "round-robin" asbestos proficiency exam.

Bulk asbestos can be efficiently identified by phase-contrast microscopy. Because its effective range is for fibers approximately 5 microns or more in length, more sensitive microscopic methods are often employed for water analyses or other surface (filter) counting techniques. Transmission electron microscopy (TEM) is often specified for full-spectrum fiber size evaluations or comparisons of fiber abundances. Scanning electron microscopy (SEM) is also a powerful tool and can provide semiqualitative and quantitative analysis when coupled with energy dispersive X-ray instrumentation.

NONPRIORITY POLLUTANT PETROLEUM HYDROCARBON ANALYSES

Petroleum hydrocarbon products such as gasoline, diesel, fuel oil, and asphalt are complex mixtures of hundreds of individual compounds. They are manufactured according to performance-based standards, not chemical composition per se; therefore, their composition is not fully known and also varies based on the crude source. These petroleum compounds are primarily alkanes, although depending upon the distillation fraction (as well as the original parent crude), various amounts of unsaturated hydrocarbons (aromatics), nitrogen and sulfur compounds may also be present. In addition to these "base" compounds, various manufacturers add different proprietary additives such as tetraethyl lead, methyl *tert*-butyl ether, ethyl *tert*-butyl ether, ethanol, detergents, lubricants, and antifreezes.

Complicating matters for the environmental scientist, as soon as many of these compounds enter the environment, they begin transforming partitioning into the atmosphere, water, and biosphere. This all happens at highly variable rates depending on temperatures, pressures, redox potentials, available bacteria, and nutrients. No single analytical method can be used to quantify or qualify all petroleum hydrocarbons. The scientist must instead carefully decide which compounds are of most significance and whether or not an empirically derived method or surrogate parameter will suffice to describe conditions.

The lighter molecular weight volatile aromatic compounds, benzene, toluene, ethylbenzene, and xylenes (BTEX) can be resolved in any of the EPA purgeable VOC methods, such as 624 (or 8260), 602 (or 8021). The PID detector for a 602 analysis is highly sensitive to aromatics and where only petroleum hydrocarbons are a concern, is the cost-effective method of choice. There are also specialized petroleum analysis labs that routinely analyze for additives, other petroleum hydrocarbon background components, and can match ("fingerprint") and date petroleum types.

Heavier molecular weight polynuclear aromatic compounds associated with petroleum compounds may be detected via EPA Method 625 (or 8270). These compounds include benzo(*a*)pyrene, anthracene, and pyrene. As previously described, the PNA fraction may also be resolved by HPLC techniques.

Due to the difficulty in analyzing petroleum products as a "whole," it has become common to use various surrogate parameters for the different petroleum hydrocarbon fractions within different product types. Some of these surrogate parameters include total ionizable compounds (TIC); total oil and grease; total petroleum hydrocarbon oil and grease (TPH); total petroleum hydrocarbons by FID (TPH); volatile petroleum hydrocarbons (VPH); extractable petroleum hydrocarbons (EPH); total organic compounds (TOC); total volatile solids (TVS); and except for aromatics, chemical oxygen demand (COD). A discussion of each of these procedures is beyond the scope of this article, but because national regulatory guidelines have been established for total petroleum hydrocarbon oil and grease, a discussion of these procedures is in order.

Total Oil and Grease

Total oil and grease by EPA Method 418.1 or Standard Method 503 includes any compound that is extractable by Freon® and may be quantified either gravimetrically (503 A) or by infrared spectroscopy (503 B). More recently, Freon has been replaced by hexane due to Freon's ozone depletion properties. Because the solvent is not specific (no solvent is), a variety of "interfering" compounds which may not be of particular importance to the scientist, such as sulfuric compounds, humic substances, and chlorophyll, may also be extracted. To help discriminate between oil and grease of petroleum origin and that of animal or vegetable origin, one should request total petroleum hydrocarbon oil and grease (TPH), Method 503 E. This method takes the additional step of passing the Freon extract through silica gel. Silica gel preferentially adsorbs polar compounds such as fatty acids. The material that passes through the silica gel is defined as TPH. This separation is by no means absolute. Many complex aromatic hydrocarbons, sulfur and nitrogen compounds may also be adsorbed by the silica gel.

The scientist must carefully choose between the procedure and the relative benefits and drawbacks of each before requesting laboratory services. In general, gravimetric procedures should be used only when a higher detection limit is adequate (approximately 5 ppm for water) and the "oil and grease" does not volatilize below 700 °C. This would include most components of gasoline and No. 2 fuel oil. When lower detection limits (approximately 0.1 ppm for water) are required and the "oil and grease" is a lighter fraction, the IR procedure is the method of choice.

Gas chromatography largely replaced infrared spectroscopy in the analysis of TPH in the 1990s. A modification of EPA Method 8100 lends itself well to producing a FID chromatogram of the various petroleum hydrocarbons from C_6 through higher than C_{38}. By simply altering the temperature program of the GC, gasoline and diesel range organics can be selectively identified. Moreover, by combining various detector arrangements with selective resins to presort and separate alkane fractions from aromatic fractions, more selective groupings of various hydrocarbon fractions can be made. These selective groupings have been useful as surrogate parameters for assessing the toxicity of petroleum in environmental risk assessments.

SOURCE-WATER PROTECTION

National Water-Quality
Assessment (NAWQA)
Program—U.S. Geological
Survey

Not all water resources are equally vulnerable to contamination. Even areas having similar land uses and sources of contamination can have different degrees of vulnerability and, therefore, different response rates to protection and management strategies. NAWQA findings clearly demonstrate that natural features—such as geology, soils, and hydrology—and land-management practices—such as tile drainage and irrigation—can affect the movement of chemicals over land or to aquifers. Effective management of nonpoint source pollution may, therefore, require targeted strategies based on different degrees of vulnerability rather than uniform treatment of contaminant sources. Linking knowledge on natural features with the use, occurrence, and transport of chemicals through the watershed makes it easier to set priorities in streams and aquifers most vulnerable to contamination and increase the cost-effectiveness of strategies designed to protect water resources in diverse settings.

State of Washington. The Washington State Department of Health, in concert with USGS, assessed the vulnerability of public water-supply wells to pesticide contamination based on geology, well characteristics, land-use activities, and low levels of detection. NAWQA information on pesticide contamination enabled the health department to identify wells with low vulnerability to contamination and obtain waivers for quarterly monitoring required under the Federal Amendments to the Safe Drinking Water Act, 1996. By using the information to meet USEPA requirements for safe drinking water, Washington State was able to protect their drinking-water source while saving at least $6 million in costly additional monitoring. This is an annual savings of as much as $70 per household on small public supply systems that were granted full monitoring waivers.

State of New Jersey. NAWQA data on organic compounds are used heavily in New Jersey's source-water assessment. USGS and the New Jersey Department of Environmental Protection are developing models to assess the vulnerability of public water supplies (including surface-water intakes and ground-water community and non-community wells) in the State to contamination by regulated compounds.

I coordinate the Upper Mississippi River Source Water Protection Initiative, an effort that will lead to the development of source-water protection plans for public water suppliers within the upper Mississippi River basin. The water quality data that have been generated and documented through NAWQA will figure prominently in the preparation of these plans. The information on certain contaminant levels in various settings within the basin, and the information describing the sources of contaminants provide documentation and solid rationale for identifying source water protection strategies, priorities, and protection measures for public water suppliers. In my opinion, more than any single information source, the Upper Mississippi River NAWQA provides an extremely valuable substantive basis for source water protection in the Upper Mississippi River basin (Mr. David Brostrom, Coordinator, Upper Mississippi River Source Water Protection Initiative, March 2001).

In addition to the examples cited above, more than 30 other states use USGS information to develop source-water protection plans for drinking-water sources. The collaborative projects in these states address nearly 40 percent of the nation's public water supply, serving more than 90 million people.

USGS information is also used widely by states to develop management plans for constituents, such as pesticides, nutrients, and MTBE. Specifically, state environmental and natural resource agencies prioritize streams and ground-water areas for assessment of these constituents on the basis of vulnerability concepts, contaminant occurrence data, and quality-assurance protocols of the NAWQA Program.

State of Kansas. NAWQA findings on elevated concentrations of atrazine (frequently approaching or exceeding the USEPA drinking-water standard) in water-supply reservoirs in the Lower Kansas River Basin were used by the Kansas State Board of Agriculture as the basis for establishing a pesticide management area in northern Kansas (Delaware River Basin). Within this management area, the State of Kansas called for both voluntary and mandatory restrictions on pesticide usage on cropland to improve water quality. The management area was the first in the nation to focus on reducing atrazine in runoff to streams and reservoirs.

State of Washington. The Washington State Department of Ecology created a Ground Water Management Area to protect ground water from nitrate contamination. The management area covers Grant, Franklin, and Adams counties, located in an intensive agricultural region of the Central Columbia Plateau. NAWQA information and communication of those findings in the USGS publication "Nitrate Concentrations in Ground Water of the Central Columbia Plateau" provided the scientific basis for implementing the management area. As follow-up to the NAWQA findings, USGS works with the Department of Ecology to (1) identify areas with lower nitrate concentrations, which could potentially serve as sources of future drinking-water supplies, (2) statistically correlate nitrate concentrations with natural features and human activities to better assess vulnerability; and, (3) design a long-term monitoring strategy for assessment of changes in nitrate concentrations over time.

State of California. USGS works with the California State Water Resources Control Board and Department of Health Services to assess the vulnerability of public supply wells to contamination. The State uses USGS ground-water-age-dating analyses as one indicator of vulnerability. In addition, on the basis of NAWQA findings on the occurrence of industry-related and petroleum-based chemicals in ground water, the State has included the collection and analysis of VOCs in their vulnerability assessment. More than 200 wells have been sampled in southern California, and these efforts will be extended to northern California and the Central Valley.

State of Idaho. NAWQA information formed the framework for predictive models and maps showing the vulnerability of ground water to contamination by the widely used herbicide atrazine in Idaho. The maps are used by the Idaho State Department of Agriculture to develop its State Pesticide Management Plan. Atrazine data from the NAWQA study in the Upper Snake River Basin were used to calibrate and verify the predictive models, which showed that significant factors associated with elevated atrazine concentrations in ground water were atrazine use, land use, precipitation, soil type, and depth to ground water. These modeling tools aid in the design of cost-effective programs for monitoring and protecting ground-water resources throughout the State.

State of Pennsylvania. The Pennsylvania Department of Environmental Protection works with USGS as a follow-up to NAWQA findings on the prevalence of MTBE in ground water and its potential to contaminate public drinking-water supplies. Through the partnership, consistent and quality-assured data will be compiled, and a qualitative vulnerability rating for MTBE will be developed for different hydrogeologic settings throughout the State of Pennsylvania. The State will use the results to prioritize areas where MTBE should be assessed and where public-supply wells should be tested, and to target inspections of gasoline storage tanks.

National scale. At the request of the USEPA's Office of Ground Water and Drinking Water, the NAWQA Program published a national map that shows the patterns of risk for nitrate contamination of shallow ground water (available in *Nitrate in Ground Waters* *of the United States—Assessing the Risk*, USGS FS-092- 96). By targeting regions with the highest risk of nitrate contamination, resources can be directed to areas most likely to benefit from pollution-prevention programs and long-term monitoring. Use of risk guidelines to locate areas for prevention of contamination also can result in cost-effective management. Once ground water is contaminated, it is expensive and, in many cases, virtually impossible to clean up.

VULNERABILITY CAN CHANGE OVER TIME

NAWQA findings show that the vulnerability to contamination of streams and ground water can differ seasonally in nearly every basin. For example, in streams that drain agricultural areas in many parts of the nation, the highest levels of nutrients and pesticides occur during spring and summer when recently applied chemicals are washed away by spring rains, snowmelt, and irrigation. Excessive amounts of contaminants can also enter streams during storm events. For example, sampling of nutrients and pesticides through a large storm event on the Potomac River in 1996 showed that concentrations and total amounts of nutrients and atrazine can increase during localized large storms, sometimes with overwhelming effects on receiving waters, such as the Chesapeake Bay. In this case, concentrations of individual compounds exceeded USEPA drinking-water standards during and following the extreme storm events. Such information helps water suppliers better understand the role of the short-term and seasonal events, and raises considerations related to timing of withdrawals, mixing, and storage to most effectively deliver high quality water at a minimum cost.

PROTOZOA IN WATER

Christine L. Bean
University of New Hampshire
Durham, New Hampshire

Protozoa are eukaryotic unicellular animals that possess membrane-bound genetic material and other cellular organelles that carry out various life functions (1). They are colorless and lack cell walls. These motile Eukarya obtain food by ingesting other organisms or organic particles. The uptake of these macromolecules occurs by a process called pinocytosis (2), whereby fluid is sucked into a channel formed by cell membrane invagination and enclosing the fluid within a membrane-bound vacuole. Particulate matter such as bacteria may also be ingested by phagocytosis.

Protozoa may be found in a variety of freshwater and marine habitats and in soil. Pathogenic protozoa include *Giardia lamblia, Entamoeba histolytica*, and *Cryptosporidium parvum. Giardia* and *Cryptosporidium* have been frequent causes of waterborne outbreaks in the United States. Many protozoa are parasitic in animals as well as humans. Reservoirs of protozoa include wild and domestic animals and humans. *Giardia* may be found in cats, dogs, sheep, pigs, goats, and cattle. *Cryptosporidium* has long been known as a pathogen of calves.

Morphologic differences in groups of protozoa can identify and differentiate them microscopically. They move and reproduce in various ways that can categorize them into groups or phyla such as amoeba (move by means of pseudopodia), flagellates, and ciliates. Most protozoa that can be found in water multiply by binary fission and are transmitted to humans through the ingestion of the infective stage of the parasite.

Free living protozoa such as flagellates, ciliates, and thecamoebae may feed on bacterial communities and actually multiply in water distribution systems (3).

FLAGELLATES

The taxonomic group of protozoa that use flagella are called Mastigophora. Motility is performed with the flagellum by either pushing or pulling the organism. Flagella can occur singly, in pairs, or in large numbers. Nearly one-half of the known species of protozoa are flagellates (4). Many of these are free-living, but several are parasitic for both humans and animals. Free-living protozoa include the green flagellate *Euglena*, which may be found in freshwater ponds (4). *Giardia lamblia* is an example of a parasitic freshwater flagellate that has a cyst and a trophozoite form. The cyst is the resting, environmentally resistant, infective form, and the trophozoite is the motile, feeding stage.

AMOEBAS

The taxonomic group of protozoa that use pseudopods for motility are called Sarcodina. Pseudpods may capture algae, bacteria, and other protozoa in the process of phagocytosis (4). Amoebic dysentery is caused by *Entamoeba histolytica*, which results in ulcerations of the intestinal tract and bloody diarrhea. Exposure may be from fecally contaminated freshwater or food supplies as well as person-to-person transmission of the cyst form. *Acanthamoeba* spp. is free-living amoebas that normally live in soil or in fresh water. They are capable of causing human infection in the brain, skin, and eye after exposure takes place through the respiratory tract, skin, or mucosal ulceration or open wounds (1).

CILIATES

The taxonomic group of protozoa that move by beating of cilia are called Ciliophora. Cilia are hair-like structures arranged in longitudinal or spiral rows that beat and result in an organized motion that moves the organism in one direction. Ciliates have two kinds of nuclei: the micronucleus and the macronucleus. The micronucleus is concerned only with inheritance, and the macronucleus is involved in the production of mRNA for cell growth and function (2). The genus *Paramecium* is a common protozoan in water. These are huge ciliated cells measuring about 50 μm in diameter. *Balantidium coli* is an example of a pathogenic ciliate. It is primarily a parasite of domestic animals such as pigs, but it has been shown to parasitize humans also.

SPOROZOANS

The taxonomic group of protozoa once believed to be spores is called Apicomplexa or Sporozoa. These protozoa have complex life cycles in which sexual and asexual reproductive phases occur sometimes in different hosts (4). Sporozoa do not have cilia, flagella, or pseudopods. Microsporidia have recently been recognized as a waterborne pathogen. The term "Microsporidia" describes protozoans belonging to the phylum *Microspora* (5). These protozoa are very small (<2 μm in diameter), and detection has been problematic. *Toxoplasma gondii* is another sporozoan highly associated with immunocompromised patients. The oocyst is shed in cat feces and is highly resistant to environmental degradation.

BIBLIOGRAPHY

1. AWWA (American Water Works Association). (1999). Waterborne Pathogens, Manual of Water Supply Practices. AWWA M48 177–182.
2. Brock, T., Madigan, M., Martinko, K., and Parker, J. (1994). *Biology of Microorganisms*, 7th Edn. Prentice-Hall, Englewood Cliffs, NJ, pp. 845–856.
3. Sibille, I., Sime-Ngando, T., Mathiew, L., and Block, J. (1998). Protozoan bacterivory and *Escherichia coli* survival in drinking water distribution systems. *Appl. Environ. Microbiol.* **64**(1): 197–202.
4. Pommerville, J.C. (2004). *Fundamentals of Microbiology*, 7th Edn. Jones and Bartlett Publishing, Sudbury, MA, pp. 594–630.
5. Dowd, S., Gerba, C., and Pepper, I. (1998). Confirmation of the human-pathogenic Microsporidia Enterocytozoon bieneusi, Encephalitozoon intestinalis, and Vittaforma corneae in water. *Appl. Environ. Microbiol.* **64**(9): 3332–3335.

WATER QUALITY

Kenneth M. Mackenthun
Arlington, Virginia

To most people, water quality means water whose characteristics represent a high degree of purity for the use to which the water is put. For drinking, this means cool, clear water, free from harmful contaminants and conditions that produce off-flavors. Other uses require a quality whose characteristics best support the particular use.

The Water Quality Act of 1965 first designated broad water uses when requiring the establishment of water quality standards for "public water supplies, propagation of fish and wildlife, recreational purposes, and agricultural, industrial, and other legitimate uses." The current Clean Water Act maintains those uses in Section 303(c)(2)(A) and adds language to take into consideration the use and value for navigation.

CHRONOLOGY OF EVENTS

In the decades of the 1950s and 1960s, many waterways in the United States showed signs of gross pollution. The Potomac River near the nation's capital was too dirty for swimming; the Cuyahoga River near Cleveland burned because of ignited oil on its surface; the Androscoggin River in Maine carried logs to pulp and paper mills and was treated with sodium nitrate to enhance dissolved oxygen through nitrification; the city of St. Louis collected its garbage, ground it, and barged it to the center of the Mississippi River for dumping; and the Flambeau River in Wisconsin carried so much black liquor from sulfite paper pulping that its bottom was covered with wood fibers, the river would rumble from escaping gases of decomposition, and boils of wood fibers and slime would rise to the river surface 6 inches thick and 10 feet in diameter. Many lakes of that era were overcome with algal and other vegetation growths from excessive nitrogen and phosphorus nutrients. These lakes included Lake Sebasticook in Maine, Lake Okeechobee in Florida, the Madison Lakes in Wisconsin, Detroit Lakes in Minnesota, and Lake Washington in Washington. Boston Harbor emitted sufficient hydrogen sulfide to blacken white-painted shoreside houses because of decomposing marine algae stimulated by nitrogen and phosphorus from the discharge of inadequately treated sewage.

The decade of the 1960s awakened the nation to the fact that water pollution must be better controlled. Prior to the 1970s, water quality went through a period of field investigation, investigative equipment development, and understanding of the relationship among aquatic organisms, chemistry, and physical conditions. Lake and stream sampling equipment was developed by Ekman (1), Petersen (2), and Surber (3). The writings of Kolkwitz and Marsson (4), Forbes and Richardson (5), Birge and Juday (6), Bartsch (7), and Hutchinson (8) influenced relationship understanding. The North American Benthological Society, formed in 1953, had its first meeting in Havana, Illinois, to share investigative information among the state biologists and coordinate taxonomic difficulties associated with stream benthic organisms.

In the decade of the 1970s, federal regulatory controls were instituted. Initial laws and regulations were refined in later years, increased emphasis was placed on identifying and controlling toxic pollutants, and efforts were made to manage nonpoint source pollutants and stormwater runoff. Water quality standards and point source discharge permits were refined and enhanced. The North American Lake Management Society held its first meeting in 1979.

LAWS AND REGULATIONS

The basic law controlling freshwater quality, with implementing regulations, dates to the River and Harbor Act of 1899 (33 USC 401). There have been many amendments along the way, and the present law was rewritten in the Federal Water Pollution Control Act Amendments of 1972 (33 USC 1251), which was significantly amended by the Clean Water Act of 1977 to emphasize toxic pollutant control. The latter laws now are commonly referred to as the Clean Water Act. The basic laws controlling marine water quality include the Marine Protection, Research and Sanctuaries Act (33 USC 1401), the Coastal Zone Management Act (16 USC 1451), the Act to Prevent Pollution from Ships (33 USC 1901), and the Magnuson–Stevens Fishery Conservation and Management Act (16 USC 1801). Other federal laws that impact actions that may affect water quality include the Endangered Species Act (16 USC 1531), the Fish and Wildlife Conservation Act (16 USC 2901), the Marine Mammal Protection Act (16 USC 1361), the National Environmental Policy Act (42 USC 4321), the National Invasive Species Act (16 USC 4701), and the Oil Pollution Act (33 USC 2701).

PROGRAMS

Defining water quality has a long record of accomplishments. Ellis (9) reviewed the literature for 114 substances and, in a 72-page document for the U.S. Bureau of Fisheries, listed lethal concentrations found by the authors. California published a 512-page book on water quality criteria that contained 1369 references. This publication was revised and enhanced the following year to include 3827 references. In 1968, the Secretary of the Interior appointed nationally recognized scientists to a national technical advisory committee to develop water quality criteria for the five recognized uses of water. This resulted in the Green Book, which constituted the most comprehensive documentation of its time on water quality requirements for defined water uses. The Blue Book followed in 1974 as the result of a contract with the National Academy of Sciences and the National Academy of Engineering. Section 304 of the Clean Water Act mandates that the Environmental Protection Agency (EPA) develop and issue Federal water quality criteria, which resulted in the Red Book, "Quality Criteria for Water," in 1976. The EPA published 65 separate water criteria documents for the 65 toxic pollutants in 1980. The Gold Book summaries followed in 1986. Presently, water quality criteria are defined in 63 *Federal Register* 68354, 10 December 1998.

Section 303 of the Clean Water Act requires that water quality standards be developed by the states and approved by EPA for all waters of the United States. Water quality standards along with national pollutant discharge elimination system permits are the hallmarks of water quality controls in the Clean Water Act. All permits required by the Act must be written in a manner that does not violate water quality standards, and Section 401 of the Act provides for state certification to that fact.

The discharge of any pollutant to the waters of the United States is controlled in Section 402 of the Clean Water Act by a national pollutant discharge elimination system (NPDES) permit. Section 301 (b)(1)(C) of the Act requires that such permits meet water quality standards. The elements of a NPDES permit include identification, description, and a unique number for each discharge; effluent limits based upon effluent guidelines, new source performance standards, toxic effluent standards, applicable water quality standards, and the best professional judgment of the permit writer; requirements for monitoring and reporting; delineation of best management practices; and a compliance schedule.

A Clean Lakes Program was established by the EPA in 1972, pursuant to Section 314 of the Clean Water Act, to provide financial and technical assistance to states in restoring publicly owned lakes. The program has funded approximately $145 million in grant activities, but there have been no appropriations since 1994.

The storage and retrieval of water quality data are vested in the EPA's STORET system. Begun in 1962, STORET contains more than 206 million analyses of chemical and physical water quality constituents at more than 761,000 locations across the United States, including coastal and international waters. Biological data are also included.

States are required by Section 305(b) of the Clean Water Act to submit a biennial report to the EPA that describes the water quality of the state. The EPA, in turn, prepares a National Water Quality Inventory, which is submitted to the Congress biennially. According to these reports, the United States has 3.5 million miles of rivers and streams; 41 million acres of lakes, ponds, and reservoirs; 34,000 square miles of estuaries; 58,000 miles of ocean shoreline; and 277 million acres of wetlands such as marshes, swamps, bogs, and fens. Rivers are polluted by bacteria, siltation, nutrients, oxygen-depleting substances, metals, habitat alterations, and suspended solids. Lakes are polluted by nutrients, siltation, oxygen-depleting substances, metals, suspended solids, pesticides, and organic toxic pollutants.

WATER QUALITY ETHIC

Nationally, water quality is significantly improved over the conditions of the 1950s and 1960s. Fish now inhabit waterways where once they were restricted from doing so because of unsuitable water quality. There is evidence, also, that an environmental ethic is gaining ground among the general population. An environmental ethic means exercising respect and consideration for the integrity, stability, and beauty of the water and its inhabitants. Such an ethic needs to be nurtured and expanded in its application. Regulations are required to keep societal goals focused on future attainment and to protect and preserve the environment for humans. A dedicated environmental ethic, coupled with viable regulations allowing flexibility for adjustment to a particular circumstance, will foster the Clean Water Act's Section 101 national goal that, wherever attainable, water quality shall provide for the protection and propagation of fish, shellfish, and wildlife and provide for recreation in and on the water.

BIBLIOGRAPHY

1. Ekman, S. (1911). Neue apparate zur qualitativen und quantitativen erforschung der bordenfauna der seen. *Int. Rev. Hydrobiol.* **7**: 164–204.
2. Petersen, C.G.J. (1911). Valuation of the sea. *Danish Biol. Sta. Report I*, **20**: 1–76.

3. Surber, E.W. (1936). Rainbow trout and bottom fauna production in one mile of stream. *Trans. Am. Fish. Soc.* **66**: 193–202.
4. Kolkwitz, R. and Marsson, M. (1909). Oekologie der tierischen saprobien. *Int. Rev. Gesamten Hydrobiol.* **2**: 126–152.
5. Forbes, S.A. and Richardson, R.E.. (1913). Studies on the biology of the Upper Illinois River. *Ill. Nat. Hist. Surv. Bull.* **9**(10): 481–574.
6. Birge, E.A. and Juday, C. (1922). *The Inland Lakes of Wisconsin. The Plankton. I. Its Quantity and Chemical Composition. Wis. Geol. and Natural History Bull. No. 64, Scientific Series No. 13*. 1–222.
7. Bartsch, A.F. (1948). Biological aspects of stream pollution. *Sewage Works J.* **20**(2): 292–302.
8. Hutchinson, G.E. (1957). *A Treatise on Limnology*. John Wiley & Sons, New York.
9. Ellis, M.M. (1937). *Detection and Measurement of Stream Pollution. U.S. Bureau of Fisheries Bull. No. 48*. 365–437.

WATER QUALITY

U.S. Geological Survey

Water quality is a term used to describe the chemical, physical, and biological characteristics of water, usually in respect to its suitability for a particular purpose. Although scientific measurements are used to define a water's quality, it's not a simple thing to say that "this water is good," or "this water is bad." After all, water that is perfectly good to wash a car with may not be good enough to serve as drinking water at a dinner party for the President! When the average person asks about water quality, they probably want to know if the water is good enough to use at home, to play in, to serve in a restaurant, etc., or if the quality of our natural waters are suitable for aquatic plants and animals.

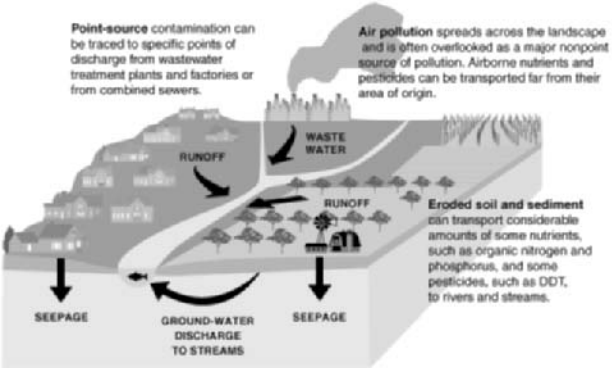

As the diagram above shows, assessment of the occurrence of chemicals that can harm water quality, such as nutrients and pesticides in water resources, requires recognition of complicated interconnections among surface water and ground water, atmospheric contributions, natural landscape features, human activities, and aquatic health. The vulnerability of surface water and ground water to degradation depends on a combination of natural landscape features, such as geology, topography, and soils; climate and atmospheric contributions; and human activities related to different land uses and land-management practices.

More and more nowadays we are hearing about situations where the quality of our water is not good enough for normal uses. Bacteria and microorganisms have gotten into drinking-water supplies, sometimes causing severe illness in a town; chemical pollutants have been detected in streams, endangering plant and animal life; sewage spills have occurred, forcing people to boil their drinking water; pesticides and other chemicals have seeped into the ground and have harmed the water in aquifers; and runoff containing pollutants from roads and parking lots have affected the water quality of urban streams.

Yes, water quality has become a very big issue today, partly because of the tremendous growth of the Nation's population and urban expansion and development. Rural areas can also contribute to water-quality problems. Medium- to large-scale agricultural operations can generate in animal feed, purchased fertilizer, and manure, more nitrogen and phosphorus than can be used by crops or animals. These excess nutrients have the potential to degrade water quality if incorporated into runoff from farms into streams and lakes All this growth puts great stress on the natural water resources, and, if we are not diligent, the quality of our waters will suffer.

EMERGING AND RECALCITRANT COMPOUNDS IN GROUNDWATER

Tom Mohr
Santa Clara Valley Water District
San Jose, California

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

During the last two decades, the dominant groundwater contamination issues have focused on petroleum hydrocarbons (including gasoline, diesel, motor oil, lubricants, and jet fuel), chlorinated solvents, metals, and pesticides. Recently, attention has shifted to a variety of newly discovered compounds in groundwater. Some of these compounds were additives to fortify or stabilize the previously mentioned products. As additives, these compounds were present at low ratios, ranging from 10% to less than 1% of the total volume of the product. Though widely used in industry for decades, these emerging compounds were, until recently, not analyzed or regulated. Emerging and recalcitrant compounds include, but are not

---

limited to, methyl tertiary-butyl ether (MtBE), a gasoline oxygenate additive; perchlorate, the major component of solid rocket motors and a minor ingredient of highway safety flares, and various stabilizers of chlorinated solvents such as 1,4-dioxane. Many of these recalcitrant compounds are also suspected to be carcinogenic or to cause other adverse health effects. Consequently, advisory drinking water action levels have been adopted for many emerging contaminants; regulatory agencies have added these compounds to their list of Compounds of Interest at remediation sites.

In the past, when confronted by new and emerging contaminants, consultants in the environmental industry generally did not have all the required remedial information because it was not immediately available. Consequently, many of these contaminants, when first encountered, are described in the literature as recalcitrant. Some compounds that were thought to be recalcitrant years ago, such as PCBs and TNT, have since been cleaned up using a variety of biological or chemical methods, following years of laboratory testing.

In approximately 1979, MtBE was added to gasoline as a fuel oxygenate to reduce air emissions and improve performance. MtBE was originally thought recalcitrant by many investigators. As an ether compound, MtBE was recognized as a compound significantly different from other petroleum hydrocarbon compounds and was not thought to be easily treated using existing hydrocarbon remedial technologies. As remedial solutions emerged, the cost of MtBE remediation was considered prohibitive. The state of the art of MtBE remediation has continued to improve in effectiveness and affordability, following years of research in the laboratory, as well as failures and successes in the field.

MtBE is highly soluble in groundwater and has low affinity for organic carbon in the soil. As a result of these physical and chemical characteristics, MtBE retardation by organic matter in soil is minimal. The solubility of MtBE in water is 43,000 mg/L, compared with the 1780 mg/L solubility of benzene. Consequently, MtBE is often found at the leading edge of gasoline plumes, followed by benzene (1).

Over the years, it has been determined that MtBE does break down using chemical oxidation (ozone, Fenton's reagent) as well as aerobic and anaerobic *ex situ* bioremediation methods. The chemical characteristics of MtBE require prompt source control of the fuel leak to minimize groundwater impacts and remediation costs (2). Successful management of MtBE contamination at fuel release sites requires adequate site characterization, selection of the appropriate remedial technology, sound engineering remedial design plans and implementation, proper operation and maintenance of the remedial system, and a cooperative approach with the regulatory agencies. Although MtBE moves quickly in groundwater, when detected soon after a spill, MtBE can be cleaned up at a cost equal to or slightly higher than that for other hydrocarbons, such as benzene. However, at sites where fuel has been leaking for some time, the MtBE plume can be significantly larger than the corresponding benzene plume, and the MtBE cleanup can cost much more.

The widespread release of chlorinated solvents in industrial and commercial cleaning and degreasing operations allowed solvents to impact groundwater throughout the nation from the 1960s through the 1980s. One place where solvent groundwater plumes are abundant is in Silicon Valley, south of San Francisco, California, where printed circuit board and semiconductor manufacturers using various solvents in degreasing operations created numerous large (1 mile or longer) chlorinated solvent plumes in shallow groundwater.

There has been innovation in restoring aquifers impacted with chlorinated solvents, and progress has been made at many of Silicon Valley's solvent release sites. The additives to solvents sometimes have the potential to be as problematic as the solvents themselves. Numerous additives are routinely included with most industrial solvents to ensure that the solvents perform as needed in their intended degreasing application. These additives are collectively known as *solvent stabilizers*, or inhibitors. They mitigate or prevent reactions in the degreaser with water, acids, and metals and inhibit degradation from heat, light, and oxygen. Stabilizers are generally added to TCE and PCE at volumetrically small proportions, generally totaling 1% or less. A few stabilizers, however, were added in the percent range. For example, 1,1,1-trichloroethane (TCA) was stabilized with 1,4-dioxane at 2–5% by volume, and some citations list as much as 8%. 1,4-Dioxane is a cyclic ether that inhibits reactions with metals, particularly aluminum salts.

The stabilizers most commonly associated with the four main solvents, TCA, trichloroethylene (TCE), tetrachloroethylene (PCE), and dichloromethane (DCM), are listed in Table 1.

Once released to the subsurface, the relative rates of migration of TCA and 1,4-dioxane are markedly different, governed by their physicochemical properties, as contrasted in Table 2. 1,4-Dioxane is resistant to both abiotic degradation and biotransformation, and owing to its infinite solubility and low affinity for sorption to soil organic matter, moves through the subsurface relatively unimpeded. Among 123 organic compounds ranked for their subsurface mobility, 1,4-dioxane is ranked first; it is deemed the most mobile among the compounds ranked (3).

Column experiments in the literature and modeling exercises have shown that 1,4-dioxane is extremely mobile in the subsurface. Accordingly, site investigation and remedial designs may require modification to account

Table 1. Solvents and Their Stabilizers

TCA	1,4-Dioxane	1,3-Dioxalane
	Nitromethane	1,2-Butylene oxide
TCE	1,2-Butylene oxide	Cresol
	Tetrahydrofuran	Epichlorohydrin
	Diisopropylamine	Alkyl pyrroles
DCM	Cyclohexane	Cyclohexene
PCE	Cyclohexene	Amines
	Butoxymethyloxirane	Phenols

Table 2. Physicochemical Properties of TCA and 1,4-dioxane

Property	1,4-Dioxane CASRN 123-91-1 $C_4H_8O_2$	TCA CASRN 71-55-6 $C_2H_3Cl_3$
Molecular weight	88.10	133.4
H_2O solubility mg/L at 20 °C	Miscible	1,360
Boiling point at 760 mmHg	101.1 °C	74.1 °C
Vapor pressure mmHg at 20 °C	37 mmHg at 25 °C	96 mmHg at 20 °C
Vapor density	3.03	5.45
Henry's constant, atm·m^3/mol	3×10^{-6}	1.5×10^{-2}
Log K_{OW}	−0.42	2.49
Log K_{OC}	0.54	2.85
Specific gravity	1.03 at 20 °C	1.34 at 20 °C

for the presence of 1,4-dioxane. Where 1,4-dioxane has been discovered at solvent release sites, regulators found it necessary to require that monitoring well networks be expanded, due to the larger footprint of 1,4-dioxane occurrence, and capture zones increased. Similarly, it has been necessary to add remediation equipment capable of removing 1,4-dioxane.

Existing groundwater treatment systems designed to remove chlorinated solvents are generally ineffective for remediating 1,4-dioxane, due to its K_{OC} and low Henry's law constant. In El Monte, California, a liquid granular activated carbon treatment system consisting of two 20,000-pound carbon vessels and treating 500 gallons per minute of solvent-contaminated groundwater was ineffective in reducing influent 1,4-dioxane concentrations at 14 μg/L to the treatment target of 3 μg/L (Fig. 1). In the City of Industry, California, an air stripper designed to remove 1.2 mg/L chlorinated solvents at 70 gallons per minute reduced 610 μg/L influent 1,4-dioxane to 430 μg/L in effluent. Conventional activated sludge and other common municipal wastewater treatment technologies have also proved ineffective in removing 1,4-dioxane.

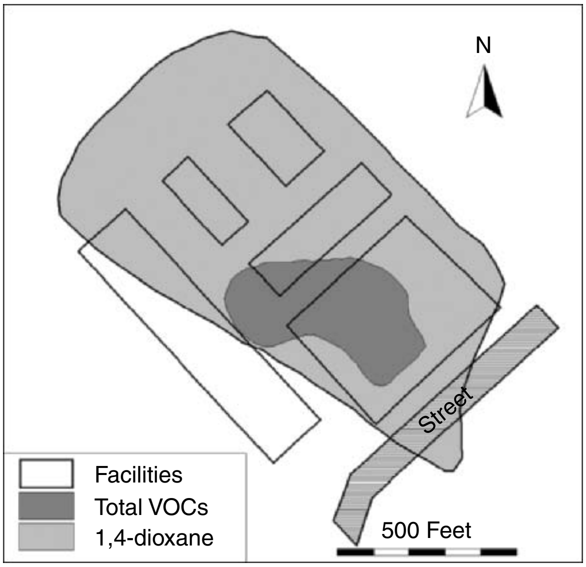

Figure 1. Silicon Valley solvent recycling facility plumes (10 μg/L contour).

The remedial technologies most commonly employed in removing 1,4-dioxane from groundwater *ex situ* are advanced oxidation processes (AOP), often in combination with ultraviolet light. AOP processes include ultraviolet light with ozone, hydrogen peroxide with ultraviolet light, ozone and hydrogen peroxide in combination, and Fenton's reagent (hydrogen peroxide and ferrous iron). Ultraviolet light causes release of hydroxyl radicals from hydrogen peroxide added to influent contaminated water. The hydroxyl radicals can react with 1,4-dioxane to oxidize the molecule to harmless reaction products (water, carbon dioxide, and residual chloride).

Biodegradation of 1,4-dioxane *in situ* is not presently considered a viable remediation option. The ether bond is a highly stable linkage and not readily biodegraded under ambient conditions. Recent research has established that there is promise for engineered bioreactors treating 1,4-dioxane *ex situ* and for *in situ* bioremediation using propane as a substrate. These methods of enhanced biodegradation show promise, whereas the more passive approach of monitored natural attenuation remains a poor candidate for effective management of 1,4-dioxane in groundwater.

SUMMARY

As many sites near closure or switch over to monitored natural attenuation, the specter of finding new and emerging compounds that are very mobile and recalcitrant is real. Discoveries of previously unknown contaminants may delay completion of site cleanup and require expensive revisions to remedial infrastructure to address their presence. Laboratory methods should be employed to obtain a full appreciation of the compounds likely to be present. Many of these emerging recalcitrant compounds are not presently subject to legal standards such as a maximum contaminant level, so there is often confusion regarding how the detection of these compounds affects progress at cleanup sites. Better regulatory guidance for the full gamut of compounds likely to be present at degreasing, fuel release, defense, and other sites is needed to ensure that time, money, and effort are well spent to complete the remedial investigation/feasibility study (RIFS) process once instead of repeating it upon the discovery of each new contaminant.

A more detailed treatment of this topic, with complete reference listings, may be found in the revised *Solvent Stabilizers White Paper* (4). The presentation binder from the Groundwater Resources Association symposium titled *Characterization and Remediation of Emerging and Recalcitrant Contaminants* in San Jose on June 14–15, 2001, and the December 10, 2003 symposium on *1,4-Dioxane and Other Solvent Stabilizers in the Environment* also contains additional information and articles on recalcitrant compounds (5).

BIBLIOGRAPHY

1. Morrison, R., Beers, R., and Hartman, B. (1998). *Petroleum Hydrocarbon Contamination: Legal and Technical Considerations*. Argent Communications Group, Foresthill, CA.
2. Wilson, J. (2003). Fate and Transport of MtBE and Other Gasoline Compounds. In: *MtBE Remediation Handbook*. E. Moyer and P. Kostecki (Eds.). Amherst Scientific, Amherst, MA.
3. Roy, W.R. and Griffin, R.A. (1985). Mobility of organic solvents in water-saturated soil materials. *Environ. Geol. Water Sci.* **7**: 241–247.
4. Mohr, T.K.G. (2001). *Solvent Stablizers White Paper*. Presented to *Groundwater Resour. Assoc. 2nd Symp. Groundwater Contaminants, Characterization and Remediation of Recalcitrant and Emerging Contaminants*, San Jose, CA June 14–15, 2001. Available for download at: http://www.valleywater.org/Water/Water_Quality/Protecting_your_water/_Lustop/_Publications_ and_Documents/_PDFs_Etc/SolventStabilizers.pdf
5. Groundwater Resources Association. (2001). *Characterization and Remediation of Emerging and Recalcitrant Contaminants. 2nd Symp. GRA Contaminant Ser.*, June 14–15, 2001, San Jose, CA.

ROAD SALT

M. Eric Benbow
Richard W. Merritt
Michigan State University
East Lansing, Michigan

INTRODUCTION

Maintenance of safe winter roadways and walkways is frequently accomplished by the application of chemical deicers and anti-icing compounds before or during snowfall and ice cover (Fig. 1). Road salt has been used since the 1940s and continues to be extensively applied throughout much of North America and Europe. It is estimated that 10 million metric tons of road salt are spread on American roadways each year, while another 3 million metric tons are used in Canada (1,2) (Fig. 2). The total annual cost of applying road salt can be greater than $62 billion in the United States because of damaged roads, bridges, and motor vehicles (1). By far, the most regularly used deicing agent is sodium chloride (NaCl), although several other inorganic salts are available (3). For instance, in the winter of 1997–1998, approximately 5 million tons of NaCl were applied to Canadian roadways compared with 120,000 tons of calcium chloride ($CaCl_2$) (4).

The widespread and increasing use of road salt has raised concerns about its effects on natural resources and drinking water. During roadway application, the components of road salt are left in the environment, typically on the soils adjacent to treated roads or into aquatic ecosystems, such as streams, rivers, wetlands, and lakes. Some salt enters groundwater during its movement. This large-scale input of salt into aquatic ecosystems is thought to have an impact on biological communities. However, the magnitude of these effects on specific ecosystem components is relatively unknown compared with other pollutants, and recent research is beginning to provide important information. In an extensive review by Environment Canada (4), it was found that: "Based upon the available data, it is considered that road salts that contain inorganic salts with or without ferrocyanide salts are entering the environment in a quantity or concentration or under conditions that have or may have an immediate or long-term harmful effect on the environment on which life depends. Therefore, it was concluded that road salts... are 'toxic' as defined in Section 64 of the Canadian Environmental Protection Act, 1999." Concerns exist about the effects of road salt on the natural environment that are balanced by transportation safety and economic issues.

Figure 1. A road salt storage facility open to the elements that can increase runoff during heavy rain effects. This facility is on a parking lot near Lansing, MI (photo by M.E. Benbow).

Figure 2. Residual road salt on the pavement of a parking lot near Lansing, MI (photo by M.E. Benbow).

For this topic on road salt, we will focus on the most commonly used form, NaCl. We briefly discuss its basic physical and chemical properties, ice control alternatives to NaCl, and the potential impacts on surface water biological communities.

DEFINITION AND PROPERTIES

Sodium chloride is the most common road salt used for melting snow and ice on pavements (3,4). Road salt can also be defined as calcium chloride, magnesium chloride ($MgCl_2$), and potassium chloride (KCl_2). In addition, an anticaking agent is often added to road salt, which is usually sodium ferrocyanide ($Na_4Fe[CN]_6 \cdot H_2O$), but is sometimes ferric ferrocyanide ($Fe_4[Fe[CN]_6]_3$), a component that has been listed as a toxic substance in the Clean Water Act (5). These anticaking agents are typically added in amounts ranging from 20 to 100 ppm (6).

The common physical and chemical properties of NaCl are given in Table 1. Pure sodium chloride is composed of 60.7% elemental chlorine and 39.3% sodium that form cubic crystals. Rock salt NaCl usually has other trace elements, including metals such as phosphorus, sulfur, nitrogen, zinc, and copper (4). By lowering the eutectic freezing point of water, the crystalline structure of pure NaCl provides anti-icing properties that results in melting when salt crystals combine with water (H_2O) molecules. The eutectic freezing point is the lowest freezing temperature for water after combining with a salt. Thus, pure NaCl melts snow and ice down to a temperature of −21.12 °C (Table 1). However, NaCl eliminates freezing when temperatures are above −11 °C, and as the difference between ambient temperature and the eutectic temperature decreases, the melting rate of ice and snow slows down (4). It has been reported that the lowest effective deicing temperature for NaCl is from 0 °C to −15 °C (4). When temperatures go below this range, other inorganic road salts are sometimes mixed with NaCl for more effective treatment on roadways, or used in place of NaCl. In addition, abrasives such as sand are mixed with rock salt during application when temperatures drop below the optimal effective temperatures. They also are added to help the salt stick to the pavement surface. Sand and other abrasives can have negative effects on roadside communities (4,7,8).

Table 1. Physical and Chemical Properties of Sodium Chloride

Property	Value or Description
Chemical formula	NaCl
Molecular weight	58.45
Atomic weight of Na	22.9898
Atomic weight of Cl	35.4527
Appearance	Clear or white solid
Solubility	35.7 g/100g water at 0 °C
Density	2.2×10^3 kg/m^3
Melting point	801 °C
Boiling point	1465 °C
Eutectic freezing point	−21.12 °C

ALTERNATIVES TO NACL

Historically, NaCl has been the preferred deicing agent because of its economic cost, easy storage, quick melting properties, and ease of handling and application. However, with recent concerns about its impact on the natural environment (see below), several alternatives have been proposed and tested (1). Table 2 gives a summary of the most common deicing and anti-icing alternatives to NaCl by comparing their eutectic temperatures and relative cost differences. The other inorganic salts ($CaCl_2$, $MgCl_2$, KCl) are more expensive than NaCl and are usually more corrosive to vehicle and roadway structures such as bridges and overpasses (1,6). They are often used in mixtures for snow and ice melting at higher rates and lower temperatures than using NaCl alone, but they also are more expensive (1). Calcium magnesium acetate (a mixture of calcium and magnesium acetates) is even more expensive, but can work well at temperatures (e.g., −9.44 °C) lower than those most effective for NaCl, and has been reported to be the least toxic of the alternatives (1). Other alternatives have been tested, such as methanol and urea. However, methanol is considerably more volatile, toxic, and flammable than NaCl, and urea is a nutrient fertilizer that can quickly contribute to eutrophication of surface waterbodies. The overwhelming argument by government agencies for the continued use of NaCl is based on economic cost–benefit comparisons to the alternatives, despite the potential effects on the environment.

IMPACTS ON AQUATIC ENVIRONMENTS

Road salt enters the air, soil, surface, and groundwater through vehicle splashing and direct runoff into adjacent areas of roadways. During windy conditions, splash and spray from melted road salt slurries can be transported

Table 2. Economic Cost and Melting Temperature Comparisons Among Chemical Alternatives to NaCl for Controlling Ice and Snow

Chemical	Relative Economic Comparison	Eutectic Temperature, °F	Eutectic Temperature, °C	Eutectic Concentration, % w/w
Sodium chloride (NaCl)	1	−6	−21	23.3
Calcium chloride ($CaCl_2$)	7	−67	−55	29.8
Magnesium chloride ($MgCl_2$)	7	−28	−33	21.6
Potassium chloride (KCl)	4	13	−11	19.5
Calcium acetate	35	5	−15	44.0
Magnesium acetate	35	−22	−30	31.0
Methanol	10	−144	−98	100.0
Urea	7	11	−12	32.5

Source: Data were taken and modified from Table 1 in the Salt Institute (6).

several meters from the side of the road and can cover roadside vegetation (4). As NaCl quickly dissociates into its component ions Na^+ and Cl^- when in solution, it is the concentrations of these ions that can affect communities adjacent to treated roadways. Although Na^+ ions readily bond to negatively charged soil particles and are typically consumed in biological processes in the sediments, Cl^- ions are more mobile and tend to be transported in surface and groundwater, where concentrations can accumulate and persist if the geological conditions are favorable (e.g., clay-lined wetlands). However, if substantial ground–surface water exchange exists, Cl^- will move with the direction of water flow in the system.

ROAD SALT LOADING AND CHLORIDE LEVELS

Studies have shown that deicing salts accumulate in rivers, streams, wetlands, and lakes during application, during snow removal and disposal, and from losses during handling and storage (4,9–11). For instance, it was estimated that road salt input into the Great Lakes was nearly two million tons by 1960: about 700,000 tons in Lake Erie; 500,000 in Lake Ontario; 400,000 in Lake Michigan; 190,000 in Lake Huron; and 90,000 in Lake Superior (12). Environment Canada (4) estimated, from a report by Morin and Perchanok (13), that the total chloride road salt loading for Canada between 1997 and 1998 was nearly 5 million tons. In 1992, 17 states in the United States reported that road salt application was a significant source of groundwater contamination, and four reported degraded wetland integrity because of salinity (14).

Increases have been documented in sodium and chloride concentrations in surface and groundwater over the past 50 years (10,15–17). In New York, it was reported that chloride transport rates increased by 145% from the 1950s to the 1970s, and that road salt application accounted for 69% of that increase (16). Other studies have shown similar trends in other parts of the United States and Canada (18–20). However, Blasius and Merritt (7) reported rather low chloride concentrations from two Michigan streams during winter months, but explained that increased discharge during spring thaw acted to dilute the chloride additions from winter road salt application. This inverse relationship of discharge and chloride levels also has been reported in other studies (21,22).

Environment Canada (4) summarized various chloride studies of urban lakes and retention ponds and reported chloride levels from 200 to 5910 mg/L. In the interstitial peat mat waters of an Indiana wetland, Wilcox (23) reported chloride levels up to 1215 mg/L. Benbow and Merritt (24) measured chloride concentrations from 43 water bodies in 37 counties of Michigan and reported a range from 18 to 2700 mg/L. From a wetland adjacent to a sand-salt storage facility, chloride was reported in another study to be 12,463 mg/L (25).

INDIRECT EFFECTS OF NACL

Residual road salt runoff is known to directly impact aquatic flora and fauna from both elevated chloride levels, and potentially from the anticaking chemicals routinely mixed for application (26,27). Sodium ferrocyanide is one such agent reported to release cyanide ions via photolysis (28); however, little information exists on the amount of ferrocyanide that might be entering impacted waterbodies during road salt application. Another indirect effect is the potential release of heavy metals already present in benthic (stream, wetland, or lake bottom) sediments through ion exchange mechanisms (29–31). For example, Doner (29) found that sodium chloride increased the mobility of Ni(II), Cu(II), and Cd(II) in soil leach columns. Furthermore, the release of mercury-adsorbed sediment particles (exchanged with Na^+) has been documented to increase with chloride ion concentrations in freshwater sediments (32). In laboratory column experiments, nearly 90% of mercury sorbed to sediments was mobilized when salt was introduced (33). Road salt also has been suggested as an agent of lake and stream acidification via ion exchange of Na^+ for H^+ (34), and is suspected to reduce soil fertility through accelerated leaching of calcium and magnesium (31).

Road salt accumulation affects small lake mixing through the establishment of short- and long-term chemoclines that prevent mixing and leading to anoxic conditions (19,35,36). Dense saline layers accumulate at lake bottoms, reducing the likelihood of mixing and prolonging anaerobic conditions of the lower layers, thus resulting in fish kills and destruction of the benthic fauna on which fish feed. Road salts entering surface waters may also trigger excessive blue-green algae growth, stimulating nuisance algal blooms (37). The mechanism of this growth involves facilitated phosphate absorption by algae in the presence of inorganic ions (38). Seasonal influxes of road salt runoff have been reported to maintain high chloride concentrations of lakes so that salt outflow is abnormally high throughout the year (39), thereby creating long-term water chemistry changes downstream.

These studies provide evidence that road salt can have indirect effects on aquatic biological communities beyond that of direct toxicity; however, direct mortality of aquatic animals can have cascading effects on the rest of the ecosystem. As macroinvertebrates commonly serve as biological indicators of aquatic ecosystems and are a dominant food source for higher consumers (40,41), we have provided a more thorough review of road salt toxicity on these organisms. Only a brief summary for zooplankton and fish is given. We refer readers to Environment Canada (4) for more information on these organisms and protozoans, bacteria, and fungi.

DIRECT EFFECTS ON AQUATIC MACROINVERTEBRATES

Before direct toxicity effects can be evaluated, the natural salinity tolerance for invertebrates should be reviewed; and this has been done in several studies. Williams et al. (42) found relationships between macroinvertebrate community structure and chloride levels of Canadian springs. In that study, they attributed high chloride levels to groundwater contamination by road salt (42). Several authors have reported benthic macroinvertebrate salinity tolerances along natural salt gradients in estuaries (43,44), salt marshes (45), and more recently in hyporheic habitats along a salinity gradient from a freshwater stream into an estuary (46). Overall, these studies have shown that distinct invertebrate communities are often associated with salinity ranges. But what about macroinvertebrates typically found in freshwater that are exposed to elevated salt concentrations from residual road salt?

Road salt toxicity studies on aquatic macroinvertebrates have shown variable results (4,7,24). Early studies found compromised osmoregulatory processes for certain aquatic insects (47–49), and species-specific tolerance variation is common in the literature (Table 3). Most acute (24–96 h) and chronic toxicity (>7 d) studies report LC_{50} values that vary among species and test conditions. The LC_{50} is the lethal concentration at which 50% of the test

Table 3. Review of Field and Laboratory Studies of NaCl Effects on Aquatic Macroinvertebrates

Endpoint Effect	NaCl, Cl^-, or Road Salt[a] Concentrations, mg/L	Reference
LC_{50} for *Culex* sp. (Culicidae)	10,254	Dowden and Bennett (50)
100% Mortality for 12-h exposure for *Chironomus attenatus* (Diptera: Chironomidae)	9995	Thornton and Sauer (47)
100% Mortality for 48-h exposure for *Cricotopus trifascia* (Diptera: Chironomidae)	8865	Hamilton et al. (51)
100% Mortality for 48-h exposure for *Hydroptila angusta* (Trichoptera: Hydroptilidae)	10,136	Hamilton et al. (51)
100% Mortality for 48-h exposure for *Nais variabilis* (Oligochaeta: Naididae)	3735	Hamilton et al. (51)
No effect on drift or mortality for *Gammarus pseudolimnaeus* (Amphipoda: Gammaridae)	800	Crowther and Hynes (27)
No effect on drift or mortality for *Hydropsyche betteni* and *Cheumatopsyche analis* (Trichoptera: Hydropsychidae)	1650	Crowther and Hynes (27)
Some drift activity of macroinvertebrates observed in field experiments	>1000	Crowther and Hynes (27)
No effect on composition or diversity of Ephemeroptera, Plecoptera, Trichoptera and Coleoptera in the field	50–67	Molles (8)
Mortality of *Hydropsyche betteni, H. bronta*, and *H. slossonae* (Trichoptera: Hydropsychidae) unaffected after 10 d	800	Kersey (52)
80% Mortality for 6-day exposure for *H. betteni*	6000	Kersey (52)
Similar growth rates among treatments for *Hexagenia limbata* (Ephemeroptera: Ephemeridae)	0, 2000, 4000, 8000	Chadwick (53)
96-h LC_{50} at 28 °C and 18 °C, respectively, for *H. limbata*	2400 and 6300	Chadwick (53)
96-h LC_{50} for *Tricorythus* sp. (Ephemeroptera: Tricorythidae)	2200–4500	Goetsch and Palmer (54)
Approximate LC_{50} for *Hydropsyche betteni* (Trichoptera: Hydropsychidae); however, no dose-response was tested	13,308	Kundman (55)
No significant drift observed for *H. betteni* in lab	2000–8000	Kundman (55)
96-h LC_{50} for *Lepidostoma* sp. (Trichoptera: Lepidostomatidae)	6000	Williams et al. (10)
70% mortality 96-h exposure for *Nemoura trispinosa* (Plecoptera: Nemouridae)	6000	Williams et al. (10)
100% mortality 96-h exposure *Gammarus pseudolimnaeus*	6000	Williams et al. (10)
No effect on macroinvertebrate functional feeding group assemblages in two rivers	8–16 (Cl^-)	Blasius and Merritt (7)
Drift effects on *Gammarus pseudolimnaeus* (Amphipoda): from 55%—68% in laboratory experiments over 24 h	2500–10,000	Blasius and Merritt (7)
Variable and no significant drift of Perlidae, Limnephilidae, Heptageniidae, Tipulidae in laboratory experiments over 24 h	1000–10,000	Blasius and Merritt (7)
Significantly lower leaf pack decomposition, but due to sand not salt	—	Blasius and Merritt (7)
96 h LC_{50} for *Gammarus* sp. (Amphipoda)	7700	Blasius and Merritt (7)
96 h LC_{50} two limnephilid caddisflies (Trichoptera)	3526	Blasius and Merritt (7)
No significant mortality for two Perlidae (Plecoptera) and one Tipulidae (Diptera) species in lab	>10,000	Blasius and Merritt (7)
100% mortality after 7 days for *Hexagenia limbata* (Ephemeroptera)	12,000	Chadwick et al. (56)
Loss of osmoregulation and increased mortality	>8000	Chadwick et al. (56)
96 h LC_{50} estimated for *Callibaetis fluctuans* (Ephemeroptera) and *Physella integra* (Gastropoda) under different test conditions (lab and field)	5000–10,000 (road salt)	Benbow and Merritt (24)
96 h LC_{50} estimated for *Hyallela azteca* (Amphipoda) and *Chaoborus americanus* (Diptera: Chaoboridae) under different test conditions (lab and field)	>10,000 (road salt)	Benbow and Merritt (24)

[a] When Cl^- or road salt was used instead of NaCl in a study, it is indicated in parentheses.

Source: The table is modified and expanded from Blasius and Merritt (7).

organisms die. In sublethal tests, the EC_{50} is reported, which is defined as the effective concentration where the tested variable (e.g., number of eggs or growth rate) is reduced by 50%. We provide Table 3 as a summary of tolerance levels of macroinvertebrates reported in the literature. This table was modified and expanded from Blasius and Merritt (7).

From the literature in Table 3, it is apparent that salt tolerance for aquatic macroinvertebrates is relatively high, with highest acute toxicity reported from mayflies (Ephemeroptera: *Hexagenia limbata*). Acute toxicity studies reported LC_{50} values (or greater mortality) from 2400 to >13,000 mg/L NaCl (or road salt). Variable mortality was evident among taxa with *Culex* mosquitoes (Culicidae) having an LC_{50} of 10,254 mg/L NaCl and caddisfly (Trichoptera) taxa with values from 3526 to 13,308. Nonlethal effects also were variable. Drift effects for various taxa (amphipods, caddisflies, mayflies, stoneflies, and craneflies) ranged from 2500 to 10,000 mg/L NaCl, with some drift effects found at concentrations >1000 mg/L NaCl. Growth rates for a mayfly (*H. limbata*) did not differ among treatments from 0–8000 mg/L NaCl; however, osmoregulation was lost at higher concentrations. This variation in lethal and nonlethal road salt tolerance may be a product of artificial testing conditions that do not represent natural environmental changes in temperature and other variables.

In addition, many taxonomic groups have not been tested, indicating that unstudied, yet sensitive species may exist. Most studies did not evaluate the background chloride levels of waterbodies typically occupied by the invertebrates tested, and when background chloride levels were measured, they were often below the invertebrate toxicity levels. In several studies, it was noted that chloride levels were sometimes diluted with snow melt runoff at the time of the highest expected impact (early spring). Nonlethal effects (e.g., drift, fecundity, production) have been understudied and may be important to invertebrates impacted by residual road salt runoff. Additional data and studies can be found in the review by Environment Canada (4).

EFFECTS ON ZOOPLANKTON AND FISH

Road salt tolerance variability has been reported for several species of zooplankton, with *Daphnia pulex, D. magna*, and *Ceriodaphnia dubia* being the most commonly tested taxa (4). Results from short-term acute toxicity studies ($\leq$24 to 96 h) reported LC_{50} values ranging from 2308 to 7754 mg/L NaCl (4). In studies lasting 7–10 d, LC_{50} values ranged from about 2000 to 6000 mg/L NaCl, whereas EC_{50} for zooplankton mean brood size, number of broods, or total progeny ranged from about 1400 to 6000 mg/L NaCl (4).

Fish road salt toxicity was low during studies <24 h, with LC_{40} and LC_{50} values ranging from 20,000 to 50,000 mg/L NaCl (4). In 24 h studies, road salt toxicity was higher, but LC_{50} values were >7000 mg/L NaCl for bluegill (*Lepomis macrochirus*), indian carp fry (*Cirrhinius mrigalo, Labeo rohoto*, and *Catla catla*), brook trout (*Salvelinus fontinalis*), and rainbow trout (*Oncorhynchus mykiss*) (4). Longer acute toxicity studies reported LC_{50} values for bluegill, rainbow trout, and indian carp fry to be >9500, 11,112, and 4980 mg/L NaCl, respectively (4). The American eel (*Anguilla rostrata*) was reported to have LC_{50} values from about 18,000 to 22,000 mg/L NaCl in one study, and fathead minnow (*Pimephales promelas*) values were from >7500 to 10,831 mg/L NaCl (4). In seven to ten day acute toxicity studies on fathead minnow and frog (*Xenopus laevis*) embryos, eggs, and larvae, LC_{50} values ranged from 1440 to 5490 mg/L NaCl depending on life stage (4).

CONCLUSIONS

The sustained and increasing use of road salt for winter roadway maintenance is becoming a popular issue as a result of raised awareness of potential ecological effects on natural waterways. Balancing human transportation safety with potential ecological repercussions drives this issue, and it is clear from the literature that the answer will be complicated, if ever resolved. The economics associated with using deicer alternatives suggest that, at present, road salt is the best practical agent for winter snow and ice removal. Chlorides are known to have effects on organisms in nature, and in many reports, it is evident that chloride concentrations are rising in waterbodies associated with roadways that are heavily treated with road salt. However, short-term toxicity studies have shown that mortality among aquatic zooplankton, macroinvertebrates, and fish is highly variable, species-dependent, and related to test conditions. Many toxicity studies are done under laboratory conditions that are far removed from the natural variation of climatic and biological components of natural waterbodies, which may influence true mortality of certain taxa. Apart from toxicity studies, larger scale studies that evaluate complex community and ecosystem level responses are needed for a better understanding of road salt effects in nature.

BIBLIOGRAPHY

1. D'Itri, F.M. (1992). *Chemical Deicers and the Environment*. Lewis Publishers, Chelsea, MI.
2. Salt Institute. (2000). Deicing salt facts: a quick reference. *SI Pub.* **2**(77): 83.
3. Field, R. and O'Shea, M.L. (1992). The USEPA research program on the environmental impacts and control of highway deicing salt pollution. In: *Chemical Deicers and the Environment*. F.M. D'Itri (Ed.). Lewis Publishers, Chelsea, MI, pp. 117–133.
4. Environment Canada. (2001). *Priority Substances List Assessment Report—Road Salts*.
5. Environmental Protection Agency. (2003). *Administrative Determination: EPA Clarifies that Ferric Ferrocyanide is one of the "Cyanides" in the Clean Water Act's List of Toxic Pollutants*. Environmental Protection Agency, Washington, DC, EPA-821-F-03-012.
6. Salt Institute. (2004). *Highway Salt and our Environment*. Salt Institute, Alexandria, VA.
7. Blasius, B.J. and Merritt, R.W. (2002). Field and laboratory investigations on the effects of road salt (NaCl) on

stream macroinvertebrate communities. *Environ. Pollut.* **120**: 219–231.

8. Molles, M.C. Jr. (1980). *Effects of Road Salting on Aquatic Invertebrate Communities*. U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, p. 10.
9. Scott, W.S. (1976). The effect of road deicing salts on sodium concentration in an urban water-course. *Environ. Pollut.* **10**: 141–153.
10. Williams, D.D., Williams, N.E., and Cao, Y. (2000). Road salt contamination of groundwater in a major metropolitan area and development of a biological index to monitor its impact. *Water Res.* **34**(1): 127–138.
11. Scott, W.S. (1980). Road salt movement into two Toronto streams. *J. Environ. Eng. Div.*
12. Struzeski, E. (1971). *Environmental Impact of Highway Deicing*. United States Environmental Protection Agency, Washington, DC.
13. Morin, D. and Perchanok, M. (2000). *Road Salt Loadings in Canada. Supporting Document for the Road Salt PSL Assessment*. Commercial Chemicals and Evaluation Branch, Environment Canada, Hull, Quebec.
14. United States Environmental Protection Agency. (1994). *National Water Quality Inventory: 1992*. U.S. EPA, Washington DC.
15. Bubeck, R.C. et al. (1971). Runoff of deicing salt: effects on Irondequoit Bay, Rochester, New York. *Science* **172**(3988): 1128–1132.
16. Peters, N.E. and Turk, J.T. (1981). Increases in sodium and chloride in the Mohawk River, New York, from the 1950s to the 1970s attributed to road salt. *Water Res. Bull.* **17**(4): 586–598.
17. Mayer, T., Snodgrass, W.J., and Morin, D. (1999). Spatial characterization of the occurrence of road salts and their environmental concentrations as chlorides in Canadian surface waters and benthic sediments. *Water Qual. Res.* **34**(4): 545–574.
18. Environment Canada (2001). *Priority Substances List Assessment Report—Road Salts*. Environment Canada, Hull, Quebec.
19. Judd, J.H. (1970). Lake stratification caused by runoff from street deicing. *Water Res.* **4**: 521–532.
20. Lathrop, R.C. (1975). *A Limnological Study of Two Neighboring Lakes: First Sister Lake and Second Sister Lake, Washtenaw County, Michigan* [MS]. Natural Resources, The University of Michigan, Ann Arbor, MI.
21. Emmett, W.W. (1975). *The Channels and Waters of Upper Salmon River, Idaho.* U.S. Department of Interior, Washington, DC.
22. Drever, J.I. (1982). *The Geochemistry of Natural Waters*. Prentice Hall, Englewood Cliffs, NJ.
23. Wilcox, D.A. (1986). The effects of deicing salts on water chemistry in Pinhook Bog, Indiana. *Water Res. Bull.* **22**(1): 57–65.
24. Benbow, M.E. and Merritt, R.W. (2004). Road-salt toxicity of select Michigan wetland macroinvertebrates under different testing conditions. *Wetlands* **24**(1): 68–76.
25. Ohno, T. (1990). Levels of cyanide and NaCl in surface waters adjacent to road salt facilities. *Environ. Pollut.* **67**: 123–132.
26. Demers, C.L. (1997). Effects of road deicing salt on aquatic invertebrates in four Adirondack streams. In: *Chemical Deicers and the Environment*. F.M. D'Itri (Ed.). Lewis Publishers, Ann Arbor, MI, pp. 245–251.
27. Crowther, R.A. and Hynes, H.B.N. (1977). The effect of road deicing salt on the drift of stream benthos. *Environ. Pollut.* **14**: 113–126.
28. Hsu, M. (1984). *Anticaking Agents in Deicing Salt*. Materials and Research Division, Maine Department of Transportation, p. 84–4.
29. Doner, H.E. (1978). Chloride as a factor in mobilities of Ni(II), Cu(II), and Cd(II) in soil. *Soil Sci. Soc. Am. J.* **42**: 882–885.
30. Jones, P.H., Jeffrey, B.A., Watler, P.K., and Hutchon, H. (1992). Environmental impact of road salting. In: *Chemical Deicers and the Environment*. F.M. D'Itri (Ed.). Lewis Publishers, Chelsea, MI, pp. 1–116.
31. Shanley, J.B. (1994). Effects of ion exchange on stream solute fluxes in a basin receiving highway deicing salts. *J. Environ. Qual.* **23**: 977–986.
32. Wang, J.S., Huang, P.M., Liaw, W.K., and Hammer, U.T. (1991). Kinetics of the desorption of mercury from selected freshwater sediments as influenced by chloride. *Water Air Soil Poll.* **56**: 533–542.
33. Macleod, C.L., Borcsik, M.P., and Jaffe, P.R. (1996). Effect of infiltrating solutions on the desorption of mercury from aquifer sediments. *Environ. Technol.* **17**: 465–475.
34. Norton, S.A., Henriksen, A., Wright, R.F., and Brakke, D.F. (1987). Episodes of natural acidification of surface waters by natural sea salts. In: *Geomon, Int. Workshop on Geochemistry and Monitoring in Representative BasinsóExtended Abstracts*. B. Moldan and T. Paces (Eds.). Geological Survey, Prague, Czechoslovakia, pp. 148–150.
35. Hoffman, R.W., Goldman, G.R., Paulson, S., and Winters, G.R. (1981). Aquatic impacts of deicing salts in the Central Sierra Nevada Mountains, California. *Water Res. Bull.* **17**(2): 280–285.
36. Kjensmo, J. (1997). The influence of road salts on the salinity and the meromictic stability of Lake Svinsjoen, southeastern Norway. *Hydrobiologia.* **347**: 151–158.
37. Sharp, R.W. (1971). *Road Salt as a Polluting Element*. State University College of Forestry at Syracuse University, Syracuse, NY.
38. Shapiro, J., Chamberlain, W., and Barrett, J. (1969). Factors influencing phosphate use by algae. Paper presented at *Proc. 4th Int. Conf. Wat. Pollut. Res.* Prague.
39. Cherkauer, D.S. and Ostenso, N.A. (1976). The effect of salt on small, artificial lakes. *Water Res. Bull.* **12**(6): 1259–1267.
40. Rosenberg, D.M. and Resh, V.H. (1993). *Freshwater Biomonitoring and Benthic Macroinvertebrates*. Chapman and Hall, New York.
41. Merritt, R.W. and Cummins, K.W. (Eds.). (1996). *An Introduction to the Aquatic Insects of North America*, 3rd Edn. Kendall/Hunt, Dubuque, IA.
42. Williams, D.D., Williams, N.E., and Cao, Y. (1997). Spatial differences in macroinvertebrate community structure in springs in southeastern Ontario in relation to their chemical and physical environments. *Can. J. Zool.* **75**: 1404–1414.
43. Williams, D.D. and Williams, N.E. (1998). Seasonal variation, export dynamics and consumption of freshwater invertebrates in an estuarine environment. *Estuarine, Coastal Shelf Sci.* **46**: 393–410.
44. Williams, D.D. and Williams, N.E. (1998). Aquatic insects in an estuarine environment: densities, distribution and salinity tolerance. *Freshwater Biol.* **39**: 411–421.
45. Giberson, D.J., Bilyj, B., and Burgess, N. (2001). Species diversity and emergence patterns of Nematocerous flies (Insecta: Diptera) from three coastal salt marshes in Prince Edward Island, Canada. *Estuaries* **24**: 862–874.

46. Williams, D.D. (2003). The brackishwater hyporheic zone: invertebrate community structure across a novel ecotone. *Hydrobiologia* **510**: 153–173.
47. Thornton, K.W. and Sauer, J.R. (1972). Physiological effects of NaCl on *Chironomus attenuatus* (Diptera: Chironomidae). *Ann. Entomol. Soc. Am.* **65**(4): 872–875.
48. Sutcliffe, D.W. (1961). Studies on salt and water balance in caddis larvae (Trichoptera): I. Osmotic and ionic regulation of body fluids in *Limnephilus affinis* Curtis. *J. Exp. Biol.* **38**: 501–519.
49. Sutcliffe, D.W. (1961). Studies on salt and water balance in caddis larvae (Trichoptera): II. Osmotic and ionic regulation of body fluids in *Limnephilus stigma* Curtis and *Anabolia nervosa* Leach. *J. Exp. Biol.* **38**: 521–530.
50. Dowden, B.F. and Bennett, H.J. (1965). Toxicity of selected chemicals to certain animals. *J. Water Pollut. Control Fed.* **37**: 1308–1316.
51. Hamilton, R.W., Buttner, J.K., and Brunetti, R.G. (1975). Lethal levels of sodium chloride and potassium chloride for an oligochaete, a chironomid midge, and a caddisfly of Lake Michigan. *Environ. Entomol.* **4**(6): 1003–1006.
52. Kersey, K. (1981). *Laboratory and Field Studies on the Effects of Road De-icing Salt on Stream Invertebrates*. University of Toronto, Institute for Environmental Studies, Snow and Ice Control Working Group, Toronto, Canada, SIC-9.
53. Chadwick, M.A.J. (1997). *Influences of Seasonal Salinity and Temperature on Hexagenia limbata (Serville) (Ephemeroptera: Ephemeridae) in the Mobile River*. Masters Thesis. Auburn University, Auburn, AL.
54. Goetsch, P.-A. and Palmer, C.G. (1997). Salinity tolerances of selected macroinvertebrates of the Sabie River, Kruger National Park, South Africa. *Arch. Environ. Contam. Toxicol.* **32**: 32–41.
55. Kundman, J.M. (1998). *The Effects of Road Salt Runoff (NaCl) on Caddisfly (Hydropsyche betteni) Drift in Mill Run, Meadville, Pennsylvania*. Senior Thesis. Department of Environmental Science, Allegheny College, Meadville, PA.
56. Chadwick, M.A., Hunter, H., Feminella, J.W., and Henry, R.P. (2002). Salt and water balance in *Hexagenia limbata* (Ephemeroptera: Ephemeridae) when exposed to brackish water. *Florida Entomol.* **85**(4): 650–651.

REVIEW OF RIVER WATER QUALITY MODELING SOFTWARE TOOLS

Anne Ng
B.J.C. Perera
Swinburne University of Technology
Hawthorne, Victoria, Australia

INTRODUCTION

Rivers supply valuable water resources for humans, and many aquatic ecosystems. However, due to population increase and its adverse effects on the rivers, and other adverse activities, the water quality in rivers has generally declined. Therefore, appropriate river water quality management strategies aimed at controlling and improving water quality should be seriously considered. To manage river water quality in the most effective and efficient way, the cause and effect relationships of the river system must first be investigated. River water quality modeling tools are extensively used in water quality management to identify these cause and effect relationships.

River water quality modeling software is designed to model the water quality in a river system. Many generic water quality modeling software tools are widely available, and most of them are in the public domain and available at no cost. The applicability of these software tools depends on the study objectives. Therefore, it is necessary to review available water quality software modeling tools, so that the most appropriate software tool can be selected for the specific application.

Here, we mainly concentrate on reviewing available public domain river water quality modeling software tools, although a brief review of catchment water quality modeling software tools is also presented. A case study is also included to identify and select the most suitable software for modeling of water quality in the Yarra River, Australia.

WATER QUALITY MODELING SOFTWARE TOOLS

In general, the water quality modeling software tools can be categorized into three broad groups, namely, catchment, river, and integrated software tools. Under these groups, the available public domain software is shown in Fig. 1. Since the focus here is on river water quality software tools, they are further divided into two groups, namely, steady and unsteady state modeling software tools.

Catchment Water Quality Modeling Softwares

Catchment water quality modeling software tools are used to estimate the amount of pollutant loadings generated from different land surfaces in catchments, which affects the water quality in streams and rivers. A listing of commonly used catchment water quality modeling software is shown in Fig. 1.

The most commonly used catchment water quality modeling software is the Agricultural Nonpoint Source Pollution Software (AGNPS) (1), which was developed by the United States Agricultural Research Services. AGNPS can be used in both event and continuous simulation modes, to estimate sediment and nutrient loads from agricultural areas. The catchment is divided into a number of cells to determine pollutant loadings. Tim et al. (2) used AGNPS (linked with a GIS system) to examine the effect of varying widths of vegetated buffer strips on sediment yield of the bluegrass catchment in Iowa, USA. AGNPS was also used to predict pollutants generated from site-specific catchment characteristics in Missouri, USA (3)

Storm Water Management Model (SWMM) is an urban stormwater quantity and quality software tool that was developed by the United States Environmental Protection Agency (U.S. EPA) (4), which can be used in both event and continuous simulation modes. Data required by SWMM are relatively intensive. It has the most versatile hydrological and hydraulic simulation modules, while the water quality simulation is relatively weak

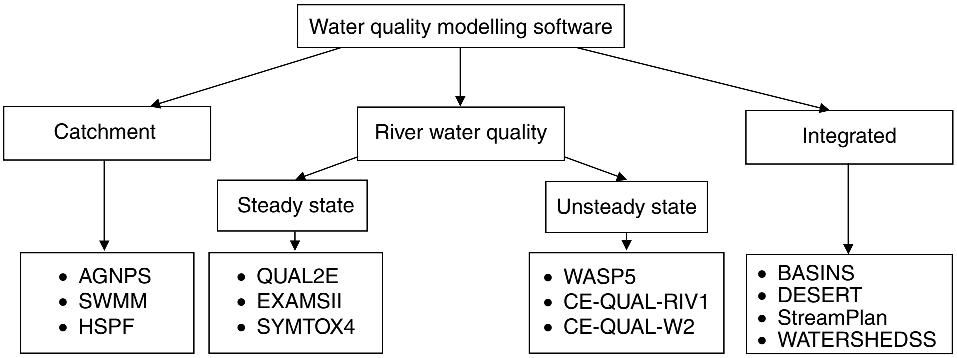

Figure 1. Summary of available public domain water quality software.

in representation of the true physical, biological, and chemical processes (5).

The Hydrological Simulation Program–FORTRAN (HSPF) is one of the comprehensive software tools, which simulates the catchment runoff processes together with a minor component on river water quality (6). This software tool can be used to simulate nonpoint source runoff. Major conventional water quality constituents such as dissolved oxygen (DO), biochemical oxygen demand (BOD), and all forms of nitrogen and phosphorus can be simulated. As discussed in References 5 and 7, HSPF is a highly complex software tool, which requires extensive resources and data. Moore et al. (8) used HSPF on the North Reelfoot Creek catchment, located in the northwest corner of Tennessee, USA to examine several best management practices in reducing erosion and sedimentation. Runoff and sediment load in the Upper Changjiang River Basin in China was simulated using HSPF (9). They found that HSPF underestimated the suspended solid concentration by up to 72% of the actual. The error in the model may have resulted from lack of data to produce a well-calibrated model.

River Water Quality Modeling Software Tools

River water modeling software tools are used to simulate the effects of pollutants generated from catchment and point sources (stormwater/treated sewage discharge outfall), on river and stream water quality. They do not estimate the nonpoint source pollutant load from the catchments. This category is reviewed in detail, since the focus is on review of river water quality modeling software tools. Two types of software tools exist, namely, steady and unsteady state.

Steady-State Software Tools. The steady-state river water quality modeling software assumes that the magnitude of flow and pollutant entering the stream do not vary with time (10). Therefore, in these software, the "average" inputs of flows and pollutants are considered for the flow event, giving "average" values for output water quality concentrations. Although these software tools cannot assess the water quality for time varying conditions, they can be useful in determining the critical water quality concentrations under design conditions. The results obtained from steady-state software are always more conservative than the results obtained with the unsteady-state software (11). The steady-state software tools are commonly used because they are less complex, easy to use, and require less input data. Below is a review of some commonly used public domain steady-state river water quality modeling software tools and their applications.

The Enhanced Stream Water Quality Model (QUAL2E) is a one-dimensional steady-state river water quality simulation software tool, which was developed and is supported by U.S. EPA (12). In 1995, QUAL2E was upgraded with a Windows interface to enhance its user-friendliness. Although QUAL2E is a steady-state flow software tool, it can also account for diurnal variation (difference in temperature between the warmest and coolest parts of the day) in temperature or algal photosynthesis and respiration. QUAL2E can model all conventional water quality constituents, as well as three other user-defined water quality constituents. QUAL2E can be applied to different waterbody types and allows modeling of multiple waste discharges and diversions. It also includes components that allow implementation of uncertainty analysis of model parameters using first-order error analysis, one-at-a-time, and Monte Carlo simulation (see Uncertainty Analysis in Watershed Modeling in this encyclopedia).

QUAL2E has been extensively used in many applications. The majority of these applications used this software tool to simulate major conventional water quality constituents in rivers. Some recent applications on modeling conventional water quality constituents in rivers include the work of Ghosh and McBean (13), Cvitanic and Kompare (14), and Ning et al. (15). In comparison, limited applications were cited in the literature using QUAL2E to simulate microbial water quality (16).

Ghosh and McBean (13) used QUAL2E to develop a water quality model of the Kali River in India. Although they experienced data limitations in their application, they commented that provided adequate data were available, QUAL2E can be effectively used to model water quality in rivers that have water pollution problems. This is consistent with the findings of Barnwell et al. (17), who stated that it is important to have site-specific data representing the properties of the actual system to model river water quality successfully.

Cvitanic and Kompare (14) applied QUAL2E to simulate and predict the possible changes in water quality in the River Sava in Croatia with the construction of impoundments. However, they found that QUAL2E was not suitable for their application, because the model prediction could not be validated during the validation stage. They concluded that a two-dimensional model is more suitable to predict water quality in impoundments, since large variations of river water quality exist in the river (laterally and vertically) during summer periods. They could have simply avoided this problem by selecting the most suitable software tool for their study, since they had a good prior knowledge of lateral and vertical water quality variations of the river.

Ning et al. (15) developed and calibrated a model for the Kao-Ping River Basin in Taiwan using QUAL2E. They successfully used the model as a simulation tool to assess the water quality standard requirements downstream by hypothetically eliminating pig farming activities and constructing a sewer system in the upstream areas. Although QUAL2E has been successfully used as a simulation tool in this study, they commented that an economic instrument for controlling and reducing the wasteload allocations would also be needed in the long term.

Steynberg et al. (16) used QUAL2E in an effort to simulate the effect of various management strategies, which can result in the desired level of fecal coliform in the Rietspruit catchment in South Africa. It was found that the model was unsuccessful in predicting the measured fecal coliform level. Due to the time varying discharge of effluent and different levels of wastewater treatment, the use of the steady-state model QUAL2E was inadequate for their study.

Reviewing the river water quality modeling software, Shanahan et al. (18) reported that QUAL2E has become the standard modeling software tool and has shown to be most applicable in situations where point source pollutants are dominant. Therefore, QUAL2E has been integrated and linked into a number of other modeling software tools. For example, QUAL2E has been integrated into decision support systems (DSSs), as in BASIN (5). Mulligan and Brown (19) and Ng (20) linked QUAL2E with genetic algorithm optimization software, GENESIS (21). As stated in Yang et al. (22), the use of remotely sensed water quality data with QUAL2E in the Te-Chi Reservoir in Taiwan can accurately interpret the spatial variation in water quality and monitor the water quality. De Azevedo et al. (23) linked QUAL2E with a water quantity network flow allocation model MODSIM of Labadie (24) to assess and evaluate six management alternatives for strategic river basin planning.

Exposure Analysis Modelling System (EXAMSII) (25) can be used in modeling of streams, rivers, and reservoirs in one-, two-, and three-dimensional modes. It accounts for many water quality transformation processes, such as photolysis, hydrolysis, oxidation, and sorption with sediments and biota (5).

SYMTOX4, the Simplified Method Program–Variable Complexity Stream Toxics Model (26), is a one-dimensional software tool that can be used to simulate the water column and benthic toxicity caused by point sources discharged into rivers. Major conventional water quality constituents can be simulated using SMYTOX4. This software is Windows based and has the capability to perform uncertainty and sensitivity analyses. One example was citied in Reference 5 using SMPTOX 4 on the Flint River, Michigan, in the United States.

Unsteady-State Software Tools. Unsteady-state (or dynamic) water quality software tools can be used to simulate water quality response in rivers whose flow and water quality characteristics change with time. All natural rivers and streams have unsteady-state flow characteristics, especially during high flow period, and therefore unsteady-state modeling software tools are more realistic. However, they require more data inputs compared to steady-state software tools. Below is a review of the available public domain unsteady-state water quality software tools (which are also listed in Fig. 1), together with their applications.

Water Quality Analysis Simulation Program (WASP5) is a well-known unsteady-state water quality simulation software tool supported by U.S. EPA (27). This software has flexible compartments such as hydrodynamics, eutrophication (DO/nutrients/algal/carbon), and toxins. The user can use these compartments selectively or all compartments simultaneously. This software tool can be used to model rivers and streams in one-, two-, and three-dimensional modes.

Lung and Larson (28) successfully used WASP5 to predict the impact of eutrophication under steady state in the Upper Mississippi River and Lake Pepin in the United States. They recognized that the unsteady-state mode should be used to study algal growth; however, relevant data inputs for algal growth dynamics were not available to them for unsteady-state flow modeling. They justified the use of steady-state mode after ascertaining that the phytoplankton population did not vary greatly from hour to hour.

Suárez et al. (29) developed an unsteady-state water quality model of the Nalón River in Spain using WASP5. This model was used to assess the impact of combined sewer flow (with daily fluctuations) on river quality and its effect on the aquatic system. The majority of the water quality related inputs required in WASP5 (e.g., reaeration rate methods, decay rates) were obtained from preliminary water quality modeling using QUAL2E. The WASP5 model was successfully calibrated, which adequately simulated the water quality and activities occurring in the river accounting for time variation. However, the decay rates were required to be the same for all reaches of the river in WASP5, which they found to be one of the main deficiencies.

A eutrophication model was developed for the Tolo Harbour in Hong Kong using WASP5 by Lee and Arega (30). This model accounts for sediment water interaction together with time and spatial variation in water quality. They successfully simulated DO and chlorophyll-*a* concentrations and matched them with observations. The model was developed to study the long-term trends of eutrophication in the harbor.

As reported by the World Bank (31), WASP5 is not appropriate for basins with large catchment areas, since it is complex and time consuming to calibrate and simulate water quality conditions of rivers and streams associated with these large basins.

The Hydrodynamic and Water Quality Model for Streams (CE-QUAL-RIV1) is a one-dimensional software tool developed by The U.S. Army Experiment Waterways Experiment Station (32). It has two separate compartments: hydrodynamics and water quality. The results obtained from the hydrodynamics compartment are used as input to the water quality compartment. Many conventional water quality constituents can be modeled, including the effects of algae and macrophytes. One advantage of using this model is that it allows modeling of river structures such as dams. This model is less widely used compared to QUAL2E and WASP5 (33). CE-QUAL-RIV1 has been applied to the Cumberland River, the Chattahoochee River, and the lower Ohio River in the United States (5).

CE-QUAL-W2, also developed by The U.S. Army Waterways Experiment Station (32), is a two-dimensional experiment, laterally averaged, hydrodynamic and water quality model. It contains one module, which models both hydrodynamics and water quality. It can model DO, nutrients, and algae interactions. Since this software accounts for variations in longitudinal and vertical directions (not in lateral direction), this software is best used in situations where large variations in lateral velocities and water quality concentrations do not occur (5). Martin (34) used CE-QUAL-W2 for DeGray Lake in Arkansas (USA) and demonstrated its usefulness.

Integrated Water Quality Modeling Software Tools. The integrated software tools consist of several stand-alone tools in one package. For example, catchment and river modeling software tools can be integrated into one package to analyze both flow and water quality in rivers and associated catchments. When some form of decision support is available in integrated software, they are called decision support systems (DSSs). There is an increased use of DSSs in river water quality management in recent times. The purpose of a DSS is to effectively allow decision-makers to simulate the whole process of decision-making, related to the particular application (e.g., improving river water quality), to investigate and simulate alternative decision management scenarios, and to improve the effectiveness of decision-making. Below are descriptions of four public domain integrated water quality software tools (or DSSs) found in the literature.

Better Assessment Science Integrating Point and Nonpoint Sources (BASINS), developed by the U.S. EPA Office of Water (35), consists of a catchment water quality modeling software tool (NPSM) and a river water quality modeling software tool (QUAL2E). The Nonpoint Source Model (NPSM) is a Windows interface that works with the catchment model HSPF (6). The graphical system in BASINS uses Arc-View GIS software. One disadvantage of this system is that the data management module is less useful to countries other than the United States, since all relevant information and data are only applicable for basins in the United States, which are updated annually.

Decision Support System for Evaluating River Basin Strategies (DESERT) is a flexible, Microsoft Windows-based tool for decision support for water quality management at the catchment scale. DESERT was developed by two organizations jointly: International Institute for Applied Systems Analysis in Austria (IIASA) and the Institute for Water and Environmental Problems in Russia (36). This software tool provides a powerful instrument for developing least-cost river catchment policies, and for assessing these policies under conditions that are deviating from the design scenario (37). Fan (38) used DESERT to identify the most efficient water quality management strategy in terms of wastewater treatment alternatives for the Veszprémi-Séd River in Budapest, Hungary.

StreamPlan (Spreadsheet Tool for River Environmental Assessment Management and Planning) was developed at IIASA in 1996 (39). It is a DSS that allows decision-makers to evaluate river and catchment water quality policies considering local and regional water quality goals, effluent standards, costs, financing, economic instruments, municipal water management issues, and generation of wastewater treatment plant alternatives. StreamPlan was developed for use on a Microsoft Excel platform, which is familiar to most model users. Jolma et al. (40) discussed the use of StreamPlan in three degraded river catchments in Central and Eastern Europe: Narew (Poland), Morava (The Czech Republic), and Nitra (Slovak Republic).

Water, Soil and Hydro-Environmental Decision Support System (WATERSHEDSS) is similar to BASINS, except WATERSHEDSS is more focused on nonpoint source pollution (41). The U.S. EPA Office of Research developed this system in 1994, with the cooperation of North Carolina State University water quality group and the Department of Biological Engineering of Pennsylvania State University (42).

EVALUATION AND SELECTION OF WATER QUALITY SOFTWARE FOR THE YARRA RIVER, AUSTRALIA

To manage river quality in the most effective and efficient way, the cause and effect relationships of the river system must first be identified. A river water quality modeling tool is required for the Yarra River, Australia, for this purpose mainly to identify the cause and effect relationships in the river from different settings of effluent license limits from sewage treatment plants (STPs). Furthermore, this modeling should be able to simulate and study the effects of various “what-if” management strategies prior to implementation.

The catchment water quality software tools mainly deal with the generation and transport of overland pollution and do not directly consider river water quality. Therefore, they are not suitable for modeling Yarra River water quality, although quantifying overland pollutant runoff into streams is an important component in the modeling of river water quality. However, modeling of overland pollutant runoff is outside the scope of modeling in-stream water quality in the Yarra River. The river water quality software tools deal with river water quality and therefore are relevant to modeling water quality in the Yarra River. Although the integrated water quality modeling software (or DSS) tools are very efficient simulation and management systems as complete decision-making tools, they require extensive data, which were not available for the Yarra River at the time of selection of the appropriate modeling tools and, therefore, were not considered. Thus, the river water quality software tools were further investigated for modeling the Yarra River, its tributaries, and associated STPs.

Four criteria, as listed below, were used in selecting the river water quality software for use in the Yarra River, from the computer software tools listed in Fig. 1.

- Data availability for use in the modeling software and the purpose of using the model
- Ability to simulate major conventional water quality constituents such as DO, BOD, and nutrients
- Ability to produce longitudinal profiles (upstream/downstream) of water quality concentrations
- Wider successful usage of the software

These broader criteria will certainly reduce the number of river water quality software tools that can be used for the Yarra River. QUAL2E and WASP5 were the only two software tools that fitted the above criteria and therefore can be used for development of the Yarra River water quality model. Both software tools can simulate conventional water quality constituents, can produce longitudinal water quality profiles, and have been used successfully on many applications.

As stated in References 5 and 18, both QUAL2E and WASP5 software tools are well known and credible, with extensive capabilities and wide usage. These software tools were further evaluated for modeling the Yarra River. A summary of the evaluation results is given in Table 1. Three main categories, namely, fundamentals, water quality, and others, were considered in classifying the attributes of these software tools. Some of these attributes were considered in a comparative study by Ambrose et al. (27).

The first category consists of attributes, which forms the basic structural framework of the software tools. Most attributes are self-explanatory. Hydrodynamics is an extremely important attribute in water quality modeling, because the movement of water affects the fate of the water quality constituents. WASP5 has an independent compartment for simulating hydrology of the water body system, whereas QUAL2E requires hydrology (or flows)

Table 1. Evaluation Summary of QUAL2E and WASP5 Attributes

Fundamentals		QUAL2E	WASP5
Operational requirements	Documentation	Y	Y
	Support	Y	Y
	Credibility	Y	Y
Water body type	River, stream	Y	Y
	Estuary	N	N
	Lake	N	Y
	Reservoir	N	Y
Dimension	one	Y	Y
	Two	N	Y
	Three	N	Y
Transport	Advection	Y	Y
	Dispersion	Y	Y
Hydrodynamics	Input	Y	Y
	Simulated	N	Y
Steady/unsteady	Steady state	Y	Y
	Unsteady state	N	Y
Discretization	—	Y	Y
Hydraulic structures	—	Y	N
Water quality constituents			
DO	Reaeration/(built-in equations)	Y/Y	Y/Y
	CBOD	Y	Y
	NH_3	Y	Y
	NO_2	Y	Y
	SOD	Y	Y
	Algae	Y	Y
Nitrogen forms	Org-N	Y	Y
	NH_3	Y	Y
	NO_2	Y	N
	NO_3	Y	Y
	Algae	N	Y
Phosphorus forms	Org-P	Y	Y
	Diss−P	Y	Y
	Algae	Y	Y
Temperature		Y	N
Settling/benthos		Y/Y	Y/Y
Toxicity		N	Y
Others			
Uncertainty and sensitivity analysis		Y	N

as input. Both software can be operated in a steady-state environment, which is the most common water quality modeling application, although WASP5 can also be used in an unsteady-state environment. A one-dimensional longitudinal process can be modeled with both software tools and is considered as the dominant transport process in most river systems, since the water quality process in the river is considered well mixed in both lateral and vertical directions (10). WASP5 can simulate river water quality in two and three dimensions. The possible increase in DO concentration by water quality structures such as dams and weirs can only be considered in QUAL2E.

The second category is related to water quality constituents. Both software tools account for most sinks on DO processes. Built-in reaeration formulas are available in both QUAL2E and WASP5. QUAL2E accounts for the four forms of nitrogen in the nitrogen cycle: Org-N, NH_3, NO_2, and NO_3 (10). WASP5 combines NO_2 and NO_3, in the overall nitrification process from NH_3 to NO_3. However, as stated in Reference 7, lumping of NO_2 and NO_3 does not cause any significant effect in the overall result, since the transformation of NO_2 to NO_3 is rapid. All phosphorus processes including the algae cycle can be accounted for in both software tools. Both software tools have the ability to simulate both settling and benthic activity, which are important for streams with low velocity. QUAL2E is the only software tool that can simulate temperature using the atmospheric heat balance equation.

The third category deals with the additional attributes. QUAL2E provides a built-in uncertainty and sensitivity analysis module, which is useful in determining the sensitivity of input parameters to output water quality. The uncertainty and sensitivity analysis of model parameters is a major component in the overall model development.

Based on the above evaluation, WASP5 is considered to be “over qualified” for the development of the Yarra River water quality model. Steady-state simulation is considered sufficient for this study, because it can be used to determine the critical water quality concentrations under design conditions. This is necessary when the model is used to simulate and study the effect of different effluent license limits on river water quality. Furthermore, only grab sample water quality data of the river and at sewage treatment plants are available for the Yarra River. This data limitation can at best be considered as steady state and only suitable to model water quality in one dimension. QUAL2E is also less complex and provides all the essential elements that are required for modeling Yarra River water quality. These elements include modeling of interaction of conventional water quality constituents and the built-in uncertainty and sensitivity analysis. The use of QUAL2E is also supported by a large number of applications in river water quality modeling (15,19).

CONCLUSION

Management of river water quality has become increasingly important due to a decline in water quality caused by human activities. Successful implementation of efficient management strategies requires the development of river water quality models. These water quality models can be used to simulate and assess the cause and effect relationships of river water quality and then to study various management strategies to improve water quality, before their implementation.

Many water quality software tools are available in the public domain, which can be used to develop river water quality models. They were reviewed here. After a detailed evaluation of the river water quality modeling tools, QUAL2E was considered as the most suitable tool for use in the Yarra River model development. The major purpose of the development of the Yarra River water quality model is to assess the effect of various sewage treatment plants (STPs) effluent license limits on river water quality and this can be adequately done with QUAL2E software. The major evaluation criteria were the appropriateness of available data for use in the modeling software and the purpose of using the model.

BIBLIOGRAPHY

1. Bingner, R., Theurer, F., Cronshey, R., and Darden, R. (2001). AGNPS 2001 Web Site. Available at http://www.sedlab.olemiss.edu/agnps.html.
2. Tim, U., Jolly, R., and Liao, H. (1995). Impact of landscape feature and feature placement on agricultural non-point source pollution control. *ASCE J. Water Resour. Plan Manage.* **121**(6): 463–470.
3. Trauth, K. and Adams, S. (2004). Watershed-based modeling with AGNPS for storm water management. *ASCE J. Water Resour. Plan. Manage.* **130**(3): 206–214.
4. Huber, W.C. and Dickinson, R. (1988). *Storm Water Management Model User's Manual, Version 4*, EPA/600/3-88/001a (NTIS PB88-236641/AS). Environmental Protection Agency, Athens, GA. Available at http://ccee.oregonstate.edu/swmm/.
5. U.S. EPA (1997). *Compendium of Tools for Watershed Assessment and TMDL Development*, Report No. EPA/841/B/97/006. Office of Water, Washington, DC.
6. Bicknell, B., Imhoff, J., Kittle, J., Donigian, A., and Johnson, R. (1997). *Hydrological Simulation Program–FORTRAN (HSPF): User's Manual for Release 11.0*, Report No. EPA 600SR97080. National Service Center for Environmental Publication, Cincinnati, OH. Available at http://www.epa.gov/ceampubl/swater/hspf/index.htm.
7. U.S. EPA (1997). *Technical Guidance Manual for Developing Total Maximum Daily Loads—Book 2: Streams and Rivers*, Report No. EPA/823/B/97/002. Office of Water, Washington, DC.
8. Moore, L., Chew, C., Smith, R., and Sahoo, S. (1992). Modeling of best management practices on North Reelfoot Creek, Tennessee. *Water Environ. Res.* **64**(3): 241–247.
9. Hayashi, S., Murakami, S., Watanabe, M., and Hua, X. (2004). HSPF simulation of runoff and sediment loads in the Upper Changjiang River Basin, China. *ASCE J. Environ. Eng.* **130**(7): 801–815.
10. Thomann, R. and Mueller, J. (1987). *Principles of Surface Water Quality Modeling and Control*. HarperCollins Publishers, New York.
11. U.S. EPA (1987). *Selection Criteria for Mathematical Models Used in Exposure Assessments*, Report No. EPA/600/8/87/042). Office of Health and Environmental Assessment, Washington, DC.

12. Brown, L.C. and Barnwell, J.O. (1987). *The Enhanced Stream Water Quality Models QUAL2E and QUAL2E - UNCAS: Documentation and User Manual*, Report No. EPA/600/3-87/OC. U.S. EPA, Athens, GA. Available at http://www.epa.gov/docs/QUAL2E_WINDOWS/index.html.
13. Ghosh, N.C. and McBean, E. (1998). Water quality modeling of the Kali River, India. *Water, Air Soil Pollut.* **102**: 91–103.
14. Cvitanic, I. and Kompare, B. (1999). Lessons learned by modelling impounded river water quality with QUAL2E. In: *Water Pollution V—Modelling, Measuring and Prediction*, WIT Press, Greece.
15. Ning, S., Chang, N., Yang, L., Chen, H., and Hsu, H. (2001). Assessing pollution prevention program by QUAL2E simulation analysis for the Kao-Ping River Basin, Taiwan. *J. Environ. Manage.* **61**, 61–76.
16. Steynberg, M.C. et al. (1995). Management of microbial water quality: new perspectives for developing areas. *Water Sci. Technol.* **32**(5–6): 183–191.
17. Barnwell, T., Brown, L., and Whittemore, C. (2004). Importance of field data in stream water quality modeling using QUAL2E-UNCAS. *ASCE J. Environ. Eng.* **130**(6): 643–647.
18. Shanahan, P., Henze, M., Koncsos, L., Rauch, W., Reichert, P., Somlyódy, L., and Vanrolleghem, P. (1998). River water quality modelling: problems of the art. In: *IAWQ 19th Biennial International Conference,* 21–26 June 1998, Vancouver, Canada, pp. 253–260.
19. Mulligan, A.E. and Brown, L.C. (1998). Genetic algorithm for calibrating water quality models. *ASCE J. Environ. Eng.* **124**(3): 202–211.
20. Ng, A.W.M. (2001). *Parameter Optimisation of River Water Quality Models Using Genetic Algorithms [Ph.D. thesis].* Victoria University, Melbourne, Australia.
21. Grefenstette, J. (1995). *A User's Guide to GENESIS (Version 5.0)*. Available at http://www. cs.purdue.edu/coast/archive/clife/Welcome.html.
22. Yang, M., Merry, C., and Sykes, R. (1999). Integration of water quality modeling, remote sensing, and GIS. *J. Am. Water Resour. Assoc.* **35**(2): 253–263.
23. De Azevedo, L. et al. (2000). Integration of water quantity and quality in strategic river basin planning. *ASCE J. Water Resour. Plan. Manage.* **126**(2): 85–97.
24. Labadie, J. (1994). MODSIM-DSS: generalized river basin decision support system and network flow model. Department of Civil Engineering, Colorado State University. Available at http://modsim.engr.colostate.edu/.
25. Burns, L. (2004). *Exposure Analysis Modelling System: User's Guide For EXAMSII Ver 2.98*, Report No. EPA/600/R-00/081. U.S. EPA, Athens, GA. Available at http://www.epa.gov/ceampubl/swater/exams/exams29804.htm.
26. CEAM (1993). SMPTOX3 Simplified Method Program—Variable Complexity Stream Toxics Model Version 2.01. Available at http://www.epa.gov/ceampubl/swater/smptox3/index.htm.
27. Ambrose, R., Wool, T., and Martin, J. (1988). *The Water Quality Analysis Simulation Program, WASP5; Part A: Manual Documentation*, Report No. EPA/600/3-87/039. Environmental Research Laboratory, U.S. EPA, Athens, GA. Available at http://www.epa.gov/ceampubl/swater/wasp/index.htm.
28. Lung, W. and Larson, C. (1995). Water quality modeling of Upper Mississippi River and Lake Pepin. *ASCE J. Environ. Eng.* **121**(10): 691–699.
29. Suárez, J., Ascorbe, A., Liaño, A., Sáinz, J., Temprano, J., and Tejero, I. (1995). Dynamic simulation of water quality in rivers. WASP5 application to the River Nalón (Spain). In: Wrobel, L. and Latinopoulos, P. (Eds.), *Water Pollution III: Modelling, Measuring and Prediction*, Computational Mechanics Publications, UK, pp. 178–188.
30. Lee, J. and Arega, F. (1999). Eutrophication dynamics of Tolo Harbor, Hong Kong. *Mar. Pollut. Bull.* **39**(1–12): 187–192.
31. World Bank. (1997). *Pollution Prevention and Abatement Handbook, Part II, Water Quality Models*. Available at http://www.esd.worldbank.org/pph/home.html.
32. USACE (1990). *CE-QUAL-RIV1: A Dynamic One-Dimensional (Longitudinal) Water Quality Model for Streams*, Instruction Report E-90-1. U.S. Army Engineer Waterway Experiment Station, Vicksburg, MI. Available at http://www.wes.army.mil/el/elmodels/index.html#wqmodels.
33. Gore, J.A. and Petts, G.E. (1989). *Alternatives in Regulated River Management*. CRC Press, Boca Raton, FL.
34. Martin, J. (1988). Application of two-dimensional water quality model. *ASCE J. Environ. Eng.* **114**(2): 317–336.
35. U.S. EPA (2001). Better Assessment Science Integrating point and Nonpoint Sources (BASINS). Available at http://www.epa.gov/waterscience/basins/basinsv3.htm.
36. Ivanov, P., Masliev, I., De Marchi, C., Kularathna, M., and Somlyódy, L. (1996). *DESERT User's Manual: Decision Support System for Evaluating River Basin Strategies*. International Institute for Applied System Analysis, Austria. Available at http://www.iiasa.ac.at/Research/WAT/docs/getit.html.
37. Perera, B.J.C. and Somlyódy, L. (1998). Decision support systems for river water quality management. In: *Second International Conference on Environmental Management*, February 10–13, 1998, Wollongong, Australia, pp. 677–684.
38. Fan, D. (1996). *Water Quality Management of the Veszprémi-Séd River and Malom and Nádor Channel System in Hungry*. International Institute for Applied System Analysis, Austria.
39. De Marchi, C., Jolma, A., Masliev, I., Perera, B.J.C., Smith, M.G., and Somlyódy, L. (1996). *StreamPlan User's Manual: Spreadsheet Tool for River Environment Assessment, Management and Planning*. International Institute for Applied System Analysis, Austria. Available at http://www.water.hut.fi/wr/research/StreamPlan/.
40. Jolma, A. et al. (1997). StreamPlan: a support system for water—quality management on a river basin scale. *Environ. Software Model.* **12**(4): 275–284.
41. Osmond, D. et al. (1997). WATERSHEDSS: a decision support system for watershed-scale nonpoint source water quality problems. *J. Am. Water Resour. Assoc.* **33**(2): 327–341.
42. NCSU (1994). WATERSHEDSS—a decision support system for nonpoint source pollution control. Available at http://www.water.ncsu.edu/watershedss/.

RIVER WATER QUALITY CALIBRATION

Anne Ng
B.J.C. Perera
Swinburne University of Technology
Hawthorne, Victoria, Australia

CALIBRATION OF RIVER WATER QUALITY MODELS

Introduction

River water quality models play an important role in water quality management. Through these models, the cause and effect relationships and the river assimilative capacity can be determined so that the appropriate strategies can be implemented in managing river water quality.

The process of the model development involves data collection and model selection, model assembly, calibration, validation, and uncertainty and sensitivity analysis. Once the river water quality model is developed using the appropriate river water quality modeling software (refer to REVIEW OF RIVER WATER QUALITY MODELING SOFTWARE TOOLS in this encyclopedia) and data, it is necessary to calibrate the model before it can be confidently used as a decision-making tool. Model calibration is necessary when the model parameters cannot be measured physically. One such parameter (which cannot be physically measured) is the decay rate of biochemical oxygen demand (BOD), and there are other such parameters in river water quality models. Model calibration is frequently referred to as parameter estimation (1), because the calibration yields the model parameters.

Here, methods and techniques used in model calibration (specifically in water related applications) are reviewed first. Then the overall methodology used in calibrating the river water quality model developed for the Yarra River, Australia, using genetic algorithms is presented.

Model Calibration Methods

Model calibration techniques can broadly be divided into two categories: manual and automatic, as shown in Fig. 1. The manual method is a trial and error method, where the model output due to different parameter sets is compared with observations visually and the parameter set that best matches the model output with observations is selected as the optimum parameter set (2,3). This method is subjective and time consuming. It can also miss the optimum parameter set. It may even lead to unrealistic parameter sets (4,5).

The automatic calibration method provides some measure of objectivity to parameter estimation and generally is conducted through optimization. Parameter optimization is achieved in the automatic calibration method through an objective function, which minimizes or maximizes a user-defined function. Several objective functions have been used in the past to assess modeled and observed responses in mathematical models in determining the optimum parameter set. The minimization of the sum of the squared difference of modeled output (due to different model parameters) and observations residual sum of squares between the actual and modeled values has been commonly used as the objective function in many hydrological studies (5–9). The use of this objective function is known as the least squares method. There are two variations from the least squares method, and they are the simple least squares method (5) and the weighted least squares method (8,9). The difference between these two methods is that the weighted least squares method requires different weights to be attached to each data point in the objective function, whereas in the simple least squares method, the weights are assumed to be equal. As stated by Sorooshian and Gupta (4), the selection of the objective function can be subjective and can produce different results for different objective functions. For example, the optimum model parameters obtained from a catchment model using two different objective functions—one considering peak flows and the other considering runoff volumes—can produce different results.

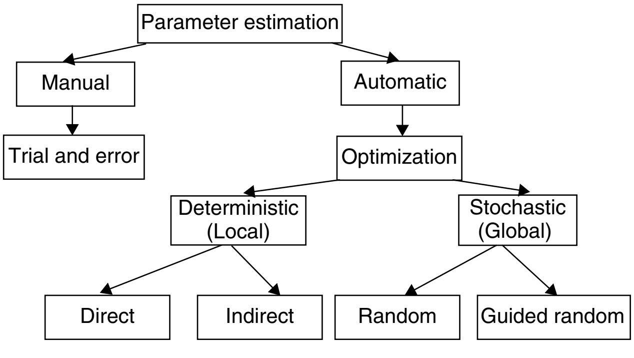

Figure 1. Broad methods in model parameter estimation.

In parameter optimization, an optimization technique is used to determine the optimum parameter set within a prescribed parameter space. As Fig. 1 shows, the optimization techniques can be divided into two broad categories: deterministic and stochastic. The deterministic techniques (also defined as local search methods) determine the optimum parameter set through a systematic search. They are designed to locate the optimum parameter set when the response surface defined by the user-defined function is unimodal (i.e., a single peak/trough). However, if the response surface is multimodal, the parameter set obtained from deterministic methods may not produce the global optimum, since the solution can be trapped at a local optimum point. Starting the optimization with different "seeds" (i.e., starting with different parameter sets for the optimization) may alleviate this problem to a certain extent. Sorooshian and Gupta (4) stated that in calibration of hydrologic models, the optimum parameter set is very rarely found through deterministic methods, since the hydrological problems contain multimodal response surfaces. Duan et al. (10) showed that there were more than hundreds of local optimum solutions in their rainfall and runoff model.

Direct and indirect search methods are two deterministic optimization methods. The direct method seeks the optimum by "hopping" around the search space of a predefined grid, where each grid point defines a parameter set, and assessing the objective function at each of these grid points. Basically, these methods use the objective functions of previous two points to determine the next point to be considered in the optimization. Sorooshian and Gupta (4) listed the most common direct search methods used in hydrologic models: Rosenbrock (11), pattern search (12), and Nelder–Mead downhill simplex methods (13). The indirect method (also known as the gradient search method) seeks the optimum solution by defining the next search point, considering both the objective function value and its gradient. Steepest decent, Hessian matrix, and Newton method are examples of indirect methods (4). These three methods have the common feature that they start from a user-defined starting point, but they differ from each other in the direction of moves and the lengths of moves.

As compared to deterministic methods, the stochastic methods are more efficient in locating the optimum parameter set, when the response surface is multimodal. However, they can also be used when the response surface is unimodal. Stochastic methods are also known as global search methods, since they are designed to produce the global optimum parameter set. They can be subdivided into two main categories, namely, random and guided random. The random search method generates the parameter sets randomly from the parameter range and optimizes the parameter sets. Generally, the random search method generates a parameter set from a uniform distribution of parameters. It does not consider the history of previous solutions in terms of optimality to determine the next parameter set, and hence the method can be inefficient. On the other hand, the guided random search method provides guided information for the next search based on the history of previously considered points (14).

Several guided random search methods exist, such as simulated annealing, adaptive random search, shuffled complex algorithms, and evolutionary algorithms (EAs) (10,14). Of these four methods, EAs have recently been used by many researchers successfully and have attracted wide attention from diverse fields, such as different areas of engineering, computer science, operations research, mathematics, and political science. The growing number of these applications is due to their ease of interfacing, simplicity, and extensibility (15).

Evolutionary Algorithms (EAs)

Evolutionary algorithms (EAs) are stochastic optimization methods that utilize the natural process of evolution (16). It has been demonstrated that EAs are robust search techniques that outperform the traditional optimization methods in many applications, in particular, when the response surface is discontinuous, noisy, nondifferentiable, and multimodal (9,17,18). Bäck and Schwefel (19) stated that EAs have become common and successful techniques in model parameter optimization and have been successfully used in model parameter optimization by Mulligan and Brown (9) and Seibert et al. (20).

There are three main forms of EAs: evolutionary programming (21), evolutionary strategies (22), and genetic algorithms (GAs) (17). Apart from these three forms of EAs, there are two other forms that were originally derived from GAs: classifier system (17) and genetic programming (23). These five forms share the common conceptual principle of EAs. That is, they repeatedly apply sequential evolutionary operators that simulate the evolution of parameter sets from the search space. These evolutionary operators are parameter representation, parameter initialization, selection, crossover, and mutation to yield offsprings (or new parameter sets) for the next generation. Depending on the type of EAs, these evolutionary operators are applied simultaneously or selectively. Of these five forms, GAs have proved to provide robust search in complex parameter search spaces. It is also the only form of EAs that utilizes all above five evolutionary operators (i.e., parameter representation, population initialization, selection, crossover, and mutation) (24).

Genetic Algorithms (GAs)

Genetic algorithms (GAs) are the most prominent and powerful optimization techniques that have been applied successfully recently in many disciplines (25). They are robust search techniques that are based on concepts of natural selection and genetics. For this reason, the terminology used in GAs is borrowed from natural genetics. The reader is referred to Goldberg (17) for details of GAs. The overall GA process as applicable to parameter optimization of river water quality models (or any mathematical model) is described below.

Mathematical models have their own model parameters. According to the genetics terminology, each model parameter is a gene, while a complete set of model parameters is a chromosome. The process of a GA begins with an initial population of a number of model parameter sets, which are chosen at random or are heuristic (requires some prior knowledge of the likely “optimum” parameter set) within a specified parameter range. This is the first generation of a number of generations (generally with a constant population size for all generations) in a GA run. Each model parameter set is then evaluated via an objective function to yield its fitness value.

The second and subsequent populations (or generations) are generated by combining model parameter sets with high fitness values from the previous population (i.e., parent) through selection, crossover, and mutation operations to produce successively fitter model parameter sets (i.e., offspring). The selection GA operator favors those parent parameter sets with high fitness value over those of lower fitness value in producing offspring. The crossover operator exchanges model parameter values from two selected model parameter sets. The mutation operator adds variability to randomly selected model parameter sets by altering some of the values randomly. Several generations are considered in one GA run, until no further improvement (within a certain tolerance) is achieved in the objective function.

The proper selection of GA operators is important for efficient optimization of model parameters. Franchini and Geleati (26) and Ng and Perera (27) investigated the effect of these GA operators on model parameters and found that they were insignificant in achieving the optimum parameter set provided that the widely used GA operators methods and values are used. Wardlaw and Sharif (28), on the other hand, found that different GA operators could produce different optimal solutions.

CALIBRATION OF RIVER WATER QUALITY MODEL USING GENETIC ALGORITHMS

River Model Development

The procedure of calibrating a river water quality model is discussed in this section. The case study discussed is of Yarra River in Australia. No attempt is made to discuss technical results; rather, the overall methodology of calibration using GAs is presented to give readers an overview of the calibration procedure. GAs were selected because they have proved to provide a robust search in complex parameter search space. Interested readers can refer to Ng (29) for complete details of the calibration and the results of the case study.

The QUAL2E (30) river water quality modeling software tool was selected and used to develop the Yarra River Water Quality Model (YRWQM) (refer to REVIEW OF RIVER WATER QUALITY MODELING SOFTWARE TOOLS in this encyclopedia). A precalibration uncertainty/sensitivity analysis was first undertaken to identify the sensitive and insensitive model parameters. This step was necessary because more effort could then be given to calibration of sensitive parameters during the model calibration. Monte Carlo simulation (MCS) was used for precalibration uncertainty/sensitivity analysis, since it has the ability to consider many different input parameter sets sampled from their distributions and to analyze the output response probability distributions. Interested readers can refer to PARAMETER UNCERTAINTY AND SENSITIVITY ANALYSIS IN RIVER WATER QUALITY MODELS (in this encyclopedia) for a review of different methods used for uncertainty/sensitivity analysis of model parameters. The precalibration uncertainty/sensitivity analysis of YRWQM model parameters and results were discussed in detail in Ng and Perera (31).

Linking of YRWQM and GA Software

A public domain GA software, namely, GENESIS (32), was used in calibration of YRWQM. GENESIS was selected for this study, since it has been used successfully for different applications (9,33,34) in the past. However, it was necessary to link YRWQM and GENESIS in order to perform the model calibration, since the calibration requires several simulation runs of YRWQM with model parameters generated from GENESIS. Linking was done through input and output files of YRWQM and GENESIS for each calibration event.

Procedure in River Water Quality Calibration

Several flow events were available for modeling of Yarra River water quality. Each event had information on flow in the river and tributaries, emissions from sewage treatment plants (STPs), and water quality measurements. Three flow events were used in calibration (35), and a further three events were used in validation.

Eleven decay rates, which were responsible for nitrogen, phosphorus, and dissolved oxygen, were considered in this calibration (Table 1). This table also shows the influence and relationships of these decay rates on water quality output responses. To use GAs, an appropriate decay rate range is required. These ranges for BOD_d, NH_3, NO_2, Org-N_d, Org-P_d, BOD_s, Org-N_s, and Org-P_s reaction rates were derived by applying the standard first-order reaction equation (36) using field data (35). The decay rate range for SOD, NH_3 benthos, and Diss-P benthos could not be estimated, because of lack of data. Therefore, the range of decay rates for the three parameters were obtained from Bowie et al. (37), who had compiled results from previous studies.

Four water quality output responses, total kjeldahl nitrogen (TKN), total nitrogen (TN), total phosphorus (TP), and dissolved oxygen (DO), were used in calibration and compared with respective observations at six water quality monitoring stations located along the Yarra River. These four output responses require the estimation of eleven water quality decay rates (Table 1). There is some interaction between these output responses (i.e., TKN affects TN, and both TN and TP affect DO) as shown in Fig. 2.

Table 1. Water Quality Decay Rates Considered in Calibration

Decay Rates	Symbols Used in Text	Influence on Output Responses
Org-N decay	Org-N_d	TKN
Org-N settling	Org-N_s	TKN
NH_3 decay	NH_{3-d}	TKN
NH_3 benthos	NH_3ben	TKN
NO_2 decay	NO_{2-d}	TN
Org-P decay	Org-P_d	TP
Org-P settling	Org-P_s	TP
Diss-P benthos	Diss-Pben	TP
CBOD decay	$CBOD_d$	DO
CBOD settling	$CBOD_s$	DO
SOD (sediment oxygen demand)	SOD	DO

The procedure for estimating decay rates was done in a systematic way as shown in Fig. 2 and also stated by McCutheon (38), Wesolowski (39), and U.S. EPA (40). First, the parameters of the water quality constituents that were not affected by other water quality constituents were estimated. Then, these parameters were kept constant, and the parameters of other water quality constituents were estimated as previously. In this study, the objective function based on simple least squares was used, since no information was available on the weights. The simple least squares objective function is given in Equation 1. This equation considers the minimization of squared difference between observed and modeled water quality concentrations at the water quality stations. This squared difference in Equation 1 is known as the fitness in GA. The lower the value, the fitter is the parameter set.

$$\mathrm{Min} \sum_{i=6} (\mathrm{OBS}_i - \mathrm{MOD}_i)^2 \qquad (1)$$

where OBS_i is the observed water quality concentration at water quality station i, and MOD_i is the modeled water quality concentration at water quality station i.

The first set of parameters considered was Org-N_d, Org-N_s, NH_{3-d}, and NH_3 benthos, which are responsible for TKN. Then, these parameters were kept constant as discussed previously and the second parameter set, NO_{2-d}, was optimized considering the output response of TN. The third set of parameters includes Org-P_d, Org-P_s, and Diss-P benthos and was optimized using the output response of TP. Optimization of decay rates for phosphorus can be done in parallel with TKN and/or TN, since TKN and TN are not influenced by TP and vice versa. The last set of parameters of $CBOD_d$, $CBOD_s$, and SOD was “optimized” using DO, keeping all other parameters at their optimized values.

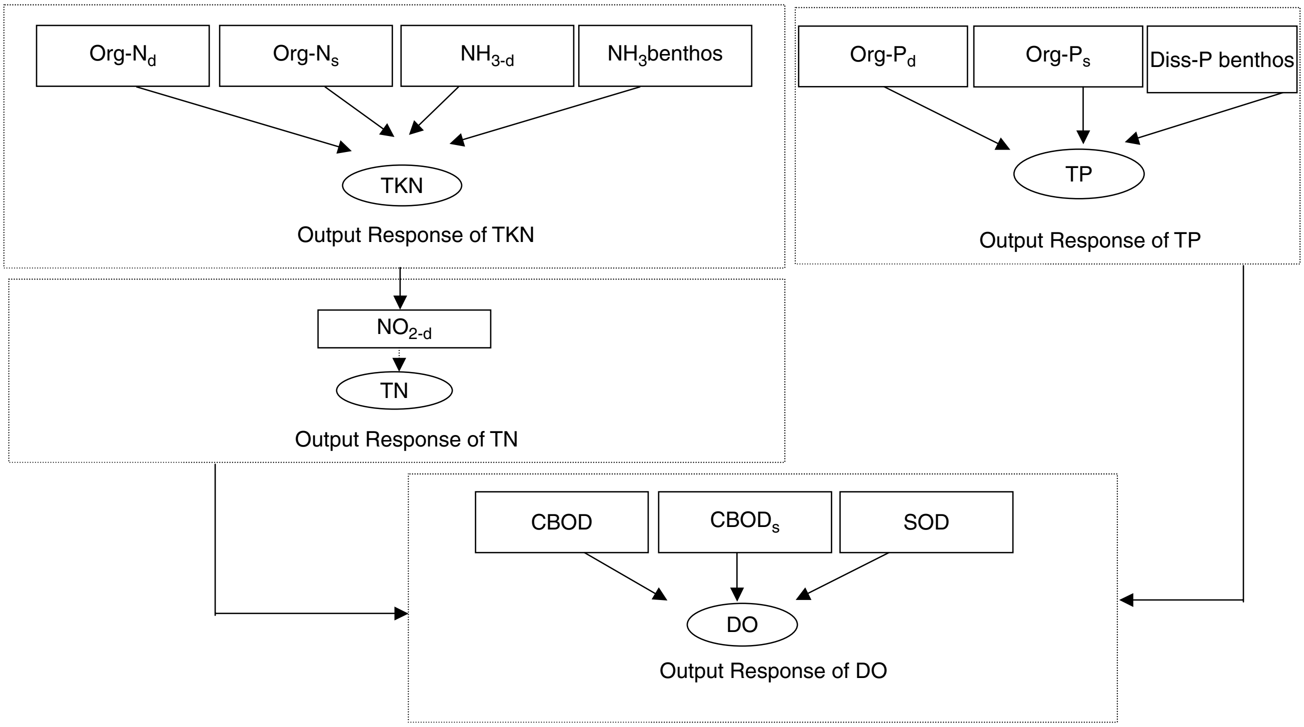

Figure 2. Systematic process used in calibration of YRWQM.

Selection of a Single Optimized Parameter Set

Three flow events used in calibration produced three sets of optimum model parameter sets. Selection of a single optimum parameter set from these sets for use in YRWQM requires subjective judgment. An attempt was made to select the single "optimum" parameter set, which models the observed water quality with reasonable accuracy for all three events. The standard student t-test (41) showed the three optimal decay rate sets (obtained from three events) produced equally good water quality predictions that match the observations at 95% significant level. In other words, the use of the three respective optimal decay sets in YRWQM has produced output water quality prediction that does not show any significant difference between modeled and observed water quality. Therefore, another statistical test, known as cumulative absolute relative error (CARE) cited in Reference 40, was used to quantitatively assess which parameter set had produced the overall lowest error with respect to all three events. The CARE value was determined using the following function:

$$\mathrm{Min} \sum_{i=1}^{3} \sum_{j=1}^{4} \sum_{k=1}^{6} \left(\frac{\mathrm{OBS}_{i,j,k} - \mathrm{MOD}_{i,j,k}}{\mathrm{MOD}_{i,j,k}} \right) \qquad (2)$$

where OBS is observed water quality, MOD is modeled water quality, i is event used in the calibration (three events), j is output water quality constituents of TKN, TN, TP, and DO (four constituents), and k is water quality sampling stations (six stations).

Once the single optimum parameter set was obtained, this set was then used in the validation process. Validation is a process that assesses the predictability of the model once it has been calibrated. This was done using three independent events, which were not used in the calibration. The model validation will enhance the confidence in using YRWQM for analysis for various management schemes in improving water quality. The student t-test results also showed that the observed and modeled water quality concentrations were not significantly different from each other at 95% significance level for all three validation events.

CONCLUSION

Simulation models are used to assess various management scenarios in improving river water quality. In order to use these simulation models confidently, the models must be well calibrated. Model calibration (or often referred to as parameter estimation) yields a set of model parameters that best estimate conditions that match with the observations. The calibrated model can then be used to simulate various management scenarios so that the implementation of water quality policy can be done in the most efficient way.

Model calibration can be done using manual and automatic methods. The manual methods use trial and error approaches, which are time consuming and require subjective judgment in defining the optimum parameter set. They can often miss the optimum parameter set. The automatic calibration methods provide some measure of objectivity in calibrating the mathematical models and obtaining the optimum model parameters. Genetic algorithms (GAs) are widely used stochastic search methods, based on the concepts of natural selection and genetics, which have proved to be successful and efficient in identifying the optimum parameter set in many water-related applications.

Due to the proven success of GAs as a calibration method, this method was used here to demonstrate the overall methodology in calibrating a river water quality model using the case study of the Yarra River in Australia. The methodology, including model development and linkage with GAs, the process of water quality calibration, and validation, was discussed.

BIBLIOGRAPHY

1. Beck, M. (1987). Water quality modeling: a review of the analysis of uncertainty. *Water Resour. Res.* **23**(8): 1393–1442.
2. Tsihrintzis, V., Fuentes, H., and Rodriguez, L. (1995). Calibration and verification of QUAL2E water quality model in sub-tropical canals. In: *First International Conference*, Aug. 14–18, 1995, Texas, USA, pp. 214–217.
3. Janssen, P. and Heuberger, P. (1999). Calibration of process-oriented models. *Ecol. Model.* **83**: 55–66.
4. Sorooshian, V. and Gupta, K. (1995). Model calibration. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*, Water Resources Publications, Littleton, CO, pp. 26–68.
5. Mohan, S. (1997). Parameter estimation of non-linear Muskingum models using genetic algorithm. *ASCE J. Hydraul. Eng.* **123**(2): 137–142.
6. Johnston, P. and Pilgrim, D. (1976). Parameter optimisation for watershed models. *Water Resour. Res.* **12**(3): 477–485.
7. Wang, Q. (1991). The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. *Water Resour. Res.* **27**(9): 2467–2471.
8. Little, K. and Williams, R. (1992). Least squares calibration of QUAL2E. *Water Environ. Res.* **64**(2): 179–185.
9. Mulligan, A.E. and Brown, L.C. (1998). Genetic algorithm for calibrating water quality models. *ASCE J. Environ. Eng.* **124**(3): 202–211.
10. Duan, Q., Sorooshian, S., and Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* **28**(4): 1015–1031.
11. Rosenbrock, H. (1960). An automatic method for finding the greatest or least value of a function. *Comput. J.* **3**: 175–184.
12. Hooke, R. and Jeeves, T. (1961). Direct search solutions of numerical and statistical problems. *J. Assoc. Comput. Machine* **8**(2): 212–229.
13. Nelder, J. and Mead, R. (1965). A simplex method for function minimisation. *Comput. J.* **7**: 308–313.
14. Filho, J., Alippi, C., and Treleaven, P. (1994). Genetic algorithm programming environments. *IEEE Comput. J.* **27**(6): 28–63.
15. Dasgupta, D. and Michalewicz, Z. (1997). Evolutionary algorithms—an overview. In: Dasgupta, D. and Michalewicz, Z. (Eds.), *Evolutionary Algorithms in Engineering Applications*, Springer, New York, pp. 3–28.
16. De Jong, K., Fogel, D., and Schwefel, H. (1997). A history of computation. In: Bäck, T., Fogel, D., and Michalewicz, Z. (Eds.), *Handbook on Evolutionary Computation* (Vol. Part A, Sec. 2.3), Institute of Physics Publishing, Oxford University Press, New York, pp. 1–12.
17. Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimisation and Machine Learning*. Addison-Wesley Publishing, Reading, MA.
18. Schwefel, H. (1997). Advantages and disadvantages of evolutionary computation over other approaches. In: Bäck, T., Fogel, D., and Michalewicz, Z. (Eds.), *Handbook on Evolutionary Computation* (Vol. Part A, Sec. 1.3), Institute of Physics Publishing, Oxford University Press, New York, pp. 1–2.
19. Bäck, T. and Schwefel, H. (1993). An overview of evolutionary algorithms for parameter optimisation. *Evol. Comput.* **1**(1): 1–23.
20. Seibert, J., Uhlenbrook, S., Leibundgut, C., and Halldin, S. (2000). Multiscale calibration and validation of a conceptual rainfall-runoff Model. *Phys. Chem. Earth* **25**(1): 59–64.
21. Fogel, D.B. (1993). Evolving behaviors in the iterated prisoner's dilemma. *Evol. Comput.* **1**: 77–97.
22. Schwefel, H. (1981). *Numerical Optimisation of Computer Models*. John Wiley & Sons, Chichester, UK.
23. Kinnear, J., Smith, R., and Michalewicz, Z. (1997). Derivative methods. In: Bäck, T., Fogel, D., and Michalewicz, Z. (Eds.), *Handbook on Evolutionary Computation* (Vol. Part B, Sec. 1.5), Institute of Physics Publishing, Oxford University Press, New York, pp. 1–15.
24. Eshelman, L. (1997). Genetic algorithms. In: Bäck, T., Fogel, D., and Michalewicz, Z. (Eds.), *Handbook on Evolutionary Computation* (Vol. Part B, Sec. 1.2), Institute of Physics Publishing, Oxford University Press, New York, pp. 1–11.
25. Paz, E. (1998). A survey of parallel genetic algorithms. Available at http://www.illigal.ge.uiuc.edu.
26. Franchini, M. and Galeati, G. (1997). Comparing several genetic algorithm schemes for the calibration of conceptual rainfall-runoff models. *Hydrol. Sci.* **42**(3): 357–379.
27. Ng, A.W.M. and Perera, B.J.C. (2003). Selection of genetic algorithm operators for water quality model calibration. *Eng. Appl. Artificial Intell. Elsevier* **16**(5–6): 529–541.
28. Wardlaw, R. and Sharif, M. (1999). Evaluation of genetic algorithms for optimal reservoir system operation. *ASCE J. Water Resour. Plan. Manage.* **125**(1): 25–33.
29. Ng, A.W.M. (2001). *Parameter optimisation of river water quality models using genetic algorithms* [Ph.D. thesis]. Victoria University, Melbourne, Australia.
30. Brown, L.C. and Barnwell, J.O. (1987). *The Enhanced Stream Water Quality Models QUAL2E and QUAL2E -UNCAS: Documentation and User Manual*, Report No. EPA/600/3-87/OC. U.S. EPA, Athens, GA.
31. Ng, A.W.M. and Perera, B.J.C. (2001). Uncertainty and sensitivity analysis of river water quality model parameters. In: *First International Conference on Water Resources Management*, Halkidiki, Greece, 24–26 Sep., pp. 175–184.
32. Grefenstette, J. (1995). *A User's Guide to GENESIS (Version 5.0)*. Available at http://www. cs.purdue.edu/coast/archive/clife/Welcome.html.
33. Whitley, D. (1989). The genitor algorithm and selection pressure: why rank-based allocation of reproductive trials is best. In: *Proceedings of the Third Conference on Genetic Algorithms*, June 4–7, 1989, George Mason University, Virginia, pp. 116–121.
34. Liong, S.Y., Chan, W.T., and ShreeRam, J. (1995). Peak-flow forecasting with genetic algorithm and SWMM. *ASCE J. Hydraul. Eng.* **121**(8): 613–617.
35. Ng, A.W.M. and Perera, B.J.C. (2000). *Data analysis and preliminary estimates of parameters for river water quality model development*. In: *Proceedings of the 26th Hydrology and Water Resources Symposium*, Perth, Australia, pp. 773–778.
36. Thomann, R. and Mueller, J. (1987). *Principles of Surface Water Quality Modeling and Control*. HarperCollins Publishers, New York.

37. Bowie, G., Mills, W., Porcella, D., Campbell, C., Pagenkopf, J., Rupp, G., Johnson, K., Chan, P., Gherini, S., and Chamberlin, C. (1985). *Rates, Constants and Kinetic Formulations in Surface Water Quality Modeling*, Report No. EPA/600/3-85/040. U.S. EPA, Athens, GA.
38. McCutcheon, S.C. (1989). *Water Quality Modeling, Volume I: Transport and Surface Exchange in Rivers*. CRC Press, Boca Raton, FL.
39. Wesolowski, E. (1994). *Calibration, Verification, and Use of a Water Quality Model to Simulate Effects of Discharging Treated Wastewater to the Red River of the North at Fargo, North Dakota*, Report No. 94-4058. U.S. Geological Survey, North Dakota.
40. U.S. EPA (1997). *Technical Guidance Manual for Developing Total Maximum Daily Loads—Book 2: Streams and Rivers*, Report No. EPA/823/B/97/002. Office of Water, Washington, DC.
41. Haldar, A. and Mahadevan, S. (2000). *Probability, Reliability and Statistical Methods in Engineering Design*. John Wiley & Sons, Hoboken, NJ, pp. 97–99.

SALMONELLA: MONITORING AND DETECTION IN DRINKING WATER

Hemant J. Purohit
Atya Kapley
National Environmental
Engineering Research Institute,
CSIR
Nehru Marg, Nagpur, India

INTRODUCTION

Providing safe drinking water is one of the challenges faced by most countries due to increasing population and industrialization. The Water Quality Guidelines stated that the most simple waterborne disease risk management involves, among other factors, identifying potential sources of contamination (1). The World Health Organization (WHO) report of March 2001 states that more than three million people die annually from waterborne related diseases. Unsafe drinking water and inadequate sanitation are two of the reasons for this problem. The WHO, conducting a workshop on hazard characterization of pathogens in food and water, launched a program of work with the objective of providing expert advice on risk assessment of microbiological hazards. *Salmonella* has been identified as one of the key pathogens for risk analysis.

Salmonellosis and typhoid fever are the main diseases caused by infection of *Salmonella*. Salmonellosis is a major public health problem because of its large and varied animal reservoir, the existence of human and animal carrier states, and the lack of a concerted nationwide program to control *Salmonella*. However, the epidemiology of typhoid fever and other enteric fevers primarily involves person-to-person spread because these organisms lack a significant animal reservoir. Contamination with human feces is the major mode of spread, and the usual vehicle is contaminated water. Typhoid fever is a public health problem of which there are an estimated 33 million cases, resulting in 500,000 deaths each year worldwide (2).

Different culture media and enrichment methods have been proposed for the isolation of *Salmonella* species from environmental samples (3,4). Conventional methods for testing *Salmonella* include the most probable number (MPN) technique (5). Additionally, selective growth media like tetrathionate broth (TB), tetrathionate broth with brilliant-green (TTBG), bismuth sulfide (BS) agar, *Salmonella*-Shigela (SS) agar, and xylose lysine deoxycholate (XLD) agar are used for enumeration of *Salmonella* (3). However, these organisms are sometimes difficult to detect or enumerate from the natural ecosystems (6). This is due to *Salmonella* often entering into the viable and nonculturable state, when exposed to environmental stresses. The routine surveillance of drinking water relies on the growth and biochemical properties of the microorganisms, which is time consuming (3,4). These drawbacks are overcome with the use of molecular tools that can detect even the viable but nonculturable forms with sensitivity and specificity.

Molecular tools involve the use of gene probes, enzyme linked assays, and the polymerase chain reaction (PCR). Use of molecular tools for detection of *Salmonella* has mainly been demonstrated in food samples. There are not many reports that describe detection methods in water samples. Purohit and Kapley (7) have proposed the use of PCR as an option in microbial quality control of drinking water. It is estimated that there are over 40, 000 references to PCR that describe its use in various applications (8). The bottleneck is in getting the PCR as a diagnostic tool from the laboratory to the field, but this needs to be done as a case-specific solution.

METHODOLOGY FOR DETECTION USING MOLECULAR TOOLS

Sampling and Generation of DNA

Methods recommended by the United States Food and Drug Administration require the culture of a sample prior to testing. An aliquot of the overnight grown sample is to be inoculated onto selective media for *Salmonella*. The suspected colonies are identified by a set of biochemical tests and, if required, further confirmed by conventional serology for *Salmonella* detection (9). Use of DNA probes and immunodetection systems also require culturing or pre-enrichment of the test sample for a few hours on selective media prior to testing with antibody or DNA probe. Immunofluorescence methods (10) and enzyme immunosorbent assays (11,12) have been successfully used to enrich *Salmonella* in primary enrichment broths. An electrochemical enzyme-linked immunosorbent assay (ELISA) coupled with flow injection analysis (ELISA-FIA) was used for detecting *Salmonella* in meat after only 5 h of incubation of pre-enrichment broth (13). Meckes and MacDonald (14) demonstrated the use of a commercially available molecular probe system to isolate and enumerate *Salmonella* spp. in sludge in less time than cultural techniques with biochemical confirmation. Nucleic acid sequence-based amplification (NASBA) results demonstrated the detection of *Salmonella* in food samples after 18 h of pre-enrichment at initial inoculum levels of 10^2 and 10^1 CFU per 25 g food sample. The primer and probe set were based on mRNA sequences of the *dnaK* gene of *Salmonella* (15).

Direct detection of *Salmonella* spp. in water samples has also been demonstrated without the use of any pre-enrichment steps using molecular methods (6,16,17). These methods have the advantage of being able to detect *Salmonella* that are in viable but nonculturable form. Bacteria enter into the nonculturable state on prolonged exposure to river or seawater, due to environmental stresses and limiting nutrient conditions (16,18,19). Such bacteria can still work as etiological agents when they come across the suitable host. Knight et al. (6) demonstrated the use of gene probe for detection from 500–1500 mL seawater samples. Kapley et al. (16) have shown the use of multiplex PCR to detect *Salmonella* from river water. In this case 5-liter samples were collected and filtered on-site through a 0.2 μm filter. The filter paper with residue was suspended in phosphate buffered saline and brought to the lab on ice within 2 h for analysis. The filter paper with buffer was thoroughly mixed and the cells were harvested from buffer. The total DNA was extracted using proteinase K treatment. The lysed cell preparation was used directly as template for PCR for detection of *Salmonella*.

PCR Based Monitoring

The polymerase chain reaction (PCR) is an established molecular technique, which reliably identifies a segment of DNA; and it uses a set of specific subsequences to amplify target segment DNA from a mixed template population. There are some loci reported for amplification target *Salmonella* along with other members of the Enterobacteriaceae family to detect the main waterborne pathogens but some are highly specific to *Salmonella*. Way et al. (20) have used multiplex PCR to detect *Salmonella* and other coliform bacteria. *PhoP* primers, specific to the phoP/phoQ loci of coliform pathogenic bacteria such as *Salmonella*, *Shigella*, *Escherichia coli*, and *Citrobacter* species served as presumptive indicators of enteric bacteria. In addition to the *phoP* primers, the *Hin* primers, 5′CTAGTGCAAATTGTGACCGCA 3′ and 5′ CCCCATCGCGCTACTGGTATC 3′, targeting a 236-bp region of hin/H2 and the *H-1i* primers, 5′ AGCCTCGGCTACTGGTCTTC 3′ and 5′ CCGCAGCAAGAGTCACCTCA 3′, amplifying a 173-bp region of the H-1i flagellin gene, were used. Both *Hin* and *H-1i* primers are specific to motile *Salmonella* species and are not present in *Shigella*, *E. coli*, or *Citrobacter* species. Cohen et al. (21) evaluated the suitability of the *fimA* gene amplification by PCR as a specific method for detection of *Salmonella* strains. *Salmonella typhimurium* and other pathogenic members of the Enterobacteriaceae family produce morphologically and antigenically related, thin, aggregative, type 1 fimbriae. Waage et al. (22) demonstrated the detection of low numbers of *Salmonella* in environmental water by a nested polymerase chain reaction assay. The target loci selected were the conserved sequences within a 2.3 kb randomly cloned DNA fragment from the *Salmonella typhimurium* chromosome. The nested PCR assay correctly identified 128 of a total of 129 *Salmonella* strains belonging to subspecies I, II, IIIb, and IV with a sensitivity of the assay was 2 CFU. No PCR products were obtained from any of the 31 non-*Salmonella* strains examined. Riyaz-Ul-Hassan et al. (23) also describe a PCR reaction using *Salmonella* enterotoxin gene (*stn*) as target loci. The protocol describes the detection of less than 10 cells of *Salmonella* in 250 mL of blood and approxmately 1 cell in 1 mL of water without any enrichment.

Duplex or multiplex reactions further improve over the limitation of the locus-specific detection in PCR or gene probes. A duplex PCR described by Kapley and co-workers demonstrates the use of *phoE* primers to detect *Salmonella* in drinking water where *Vibrio* specific primers were also used in the same reaction (24). Multiplex PCR to detect *Salmonella* along with different pathogens in a single reaction has also been demonstrated (17,25). While Kapley and co-workers have demonstrated the use of the gene locus that has been shown to be essential for the invasion of *Salmonella* into epithelial cells of the host's intestine (inv), Kong and co-workers have used the invasion plasmid antigen B (*ipaB*) gene of *Salmonella typhimurium*. The primers for *ipaB*, forward primer 5′ GGACTTTTTAAAAGCGGCGG 3′ and reverse primer 5′ GCCTCTCCCAGAGCCGTCTGG 3′, amplified a 314 bp region. A multiplex PCR that amplifies five different loci in a single reaction has been reported for detection of *Salmonella* from river water samples (16). The gene markers used were *invA*, *SpvA*, *Spv*, and *phoE* for pathogenic *Salmonella, and* 16S rRNA specific primers to assess the total eubacterial load in the water samples. The *inv* locus was chosen as its expression has been shown to be essential for the invasion of *Salmonella* into epithelial cells of the host's intestine; causing the gastrointestinal infections (26). The *invA* gene codes for a protein, which is necessary for virulence of the bacterium. The primers used, upper 5′-CCTGATCGCACTGAATATCGTACTG 3′ and lower 5′GACCATCACCAATGGTCAGCAGG 3′, amplified a 598 bp fragment. The other pathogenic determinants in *Salmonella* are the plasmid coded virulence *Spv* genes; *SpvA*, *SpvB*, and *SpvC* (27). More than 80% of all *Salmonella* isolated from food and clinical specimens contain the virulence factors coded through plasmid (28). The primers reported for *spvA* upper 5′ TGTATGTTGATACTAAATCC 3′ and lower 5′ CTGTCATGCAGTAACCAG 3′, amplify a 470 bp product, while the primers reported for *spvB*, upper 5′ ATGAATATGAAT CAGACCACC 3′, and lower 5′ GGCGTATAGTCG GCGGTTTTC 3′ amplify a 669 bp product (16). The *phoE* gene encodes for a phosphate limitation-inducible outer membrane pore protein and has been shown to be *Salmonella* specific. The primers, upper 5′ AGCGCCGCGGTACGGGCGATAAA 3′ and lower 5′ ATCATCGTCATTAATGCCTAA CGT 3′, from the *phoE* locus have already been tested for 133 different *Salmonella* strains (29). The key to this detection tool has been provided through the thermocycling steps used in the M-PCR. The protocol has used gradient temperature steps, which has provided the selective annealing of primers. The developed multistep program could coamplify the target gene markers, used in the study. The specificity of the program has been proved with the DNA template derived from even the river water samples, which represents the heterogeneous microbial population.

The most widely used selective primer set for detection of *Salmonella* has been reported, which are designed from *invA* locus (30). An extended determination of selectivity by using 364 strains showed that the inclusivity was 99.6% and the exclusivity was 100% for the *invA* primer set (31).

FUTURE OPTIONS IN DETECTION OF *SALMONELLA* IN WATER SAMPLES

With the increasing incidences of spread of waterborne diseases, there is a great need to be able to detect waterborne pathogens like *Salmonella* (31). The key to this problem is to monitor subtle impacts that may have long-term effects that are hazardous. Genomic tools provide an option for analyzing basic resources at the nucleic acid level of an organism or population of microorganisms. Purohit et al. (32) have reviewed the use of genomic tools in monitoring and assessment processes that can help evaluate the environment impact. These tools can be extrapolated for assessing the mixing of raw sewage waters with water resources. A shortened PCR time with quantification and assured selectivity could be achieved by real time PCR. Based on this concept, an advanced nucleic acid analyzer (ANNA) has been reported that can detect bacteria within 7 min (33). This battery-operated device can be used in the field and has software that can be used by first timers. Inclusion of fluorescent dyes to the DNA during amplification has led to fresh advances in detection methods. Use of molecular beacons, oligonucleotide probes that become fluorescent upon hybridization, have been used in real time PCR assay to detect as few as 2 colony forming units (CFU) per reaction of *Salmonella* species. This method uses a 122 bp section of the *himA* as the amplification target. This method could also discriminate between amplicons obtained from similar species such as *E. coli* and *C. freundii*. The assay could be carried out entirely in sealed PCR tubes, enabling fast and direct detection of *Salmonella* in a semiautomated format (34). Detection of 1 colony forming unit/mL in food products was also demonstrated by real time PCR using *SipB* and *SipC* as target loci (35). Beyond the real time PCR, protocols need to be developed that can directly detect the pathogen using a biosensor. Kramer and Lim (36) have demonstrated a rapid and automated fiberoptic biosensor assay for the detection of *Salmonella* in sprout rinse water. Alfalfa seeds contaminated with various concentrations of *Salmonella typhimurium* were sprouted. The spent irrigation water was assayed 67 h after alfalfa seed germination with the RAPTOR (Research International, Monroe, WA), an automated fiberoptic-based detector. *Salmonella typhimurium* was identified in spent irrigation water when seeds were contaminated with 50 CFU/g. Viable *Salmonella typhimurium* cells were also recovered from the waveguides after the assay. This biosensor assay system has the potential to be directly connected to water lines and can assist in process control to identify contaminated water.

BIBLIOGRAPHY

1. Deere, D., Stevens, M., Davison, A., Helm, G., and Dufour, A. (2001). Management strategies. In: *Water Quality—Guidelines, Standards and Health: Assessment of Risk and Risk Management for Water-Related Infectious Disease*. World Health Organization, Geneva, Switzerland, World Health Organization, TJ International (Ltd), Padstow, Cornwall, UK.
2. Garmory, H.S., Brown, K.A., and Titball, R.W. (2002). *Salmonella* vaccines for use in humans: present and future. *FEMS Microbiol. Rev.* **26**: 339–353.
3. Hussong, D, Enkiri, N.K., Burge, W.D. (1984). Modified agar medium for detecting environmental *Salmonella* by the most-probable number method. *Appl. Environ. Microbiol.* **48**: 1026–1030.
4. Morinigo, M.A., Borrego, J.J., and Romero, P. (1986). Comparative study of different methods for detection and enumeration of *Salmonella* spp. in natural waters. *J. Appl. Bacteriol.* **61**: 169–176.
5. Russ, C.F. and Yanko, W.A. (1981). Factors affecting salmonellae repopulation in composted sludges. *Appl. Environ. Microbiol.* **41**: 597–602.
6. Knight, I.T., Shults, S., Kaspar, C.W., and Colwell, R.R. (1990). Direct detection of *Salmonella* spp. in estuaries by using a DNA probe. *Appl. Environ. Microbiol.* **56**: 1059–1066.
7. Purohit, H.J. and Kapley, A. (2002). PCR as an emerging option in the microbial quality control of drinking water. *Trends Biotechnol.* **20**: 325–326.
8. White, T.J. (1996). The future of PCR technology: diversification of technologies and applications. *Trends Biotechnol.* **14**: 478–483.
9. Durango, J., Arrieta, G., and Mattar, S. (2004). Presence of *Salmonella* as a risk to public health in the Caribbean zone of Colombia. *Biomedica* **24**: 89–96.
10. Thomason, B.M. (1981). Current status of immunofluorescent methodology for salmonellae. *J. Food Prot.* **44**: 381–384.
11. Anderson J.M. and Hartman P.A. (1985). Direct immunoassay for detection of salmonellae in foods and feeds. *Appl. Environ. Microbiol.* **49**: 1124–1127.
12. Mattingly, J.A. and Gehle, W.D. (1984). An improved enzyme immunoassay for the detection of *Salmonella*. *J. Food Sci.* **49**: 807–809.
13. Croci, L., Delibato, E., Volpe, G., De Medici, D., and Palleschi G. (2004). Comparison of PCR, electrochemical enzyme-linked immunosorbent assays, and the standard culture method for detecting *Salmonella* in meat products *Appl. Environ. Microbiol.* **70**: 1393–1396.
14. Meckes, M.C. and MacDonald, J.A. (2003). Evaluation of a DNA probe test kit for detection of salmonellae in biosolids. *J. Appl. Microbiol.* **94**: 382–386.
15. D'Souza, D.H. and Jaykus L.-A. (2003). Nucleic acid sequence based amplification for the rapid and sensitive detection of *Salmonella enterica* from foods. *J. Appl. Microbiol.* **95**: 1343–1350.
16. Kapley, A., Lampel, K., and Purohit, H.J. (2001). Rapid detection of *Salmonella* in water samples by multiplex PCR. *Water Environ. Res.* **73**: 461–465.
17. Kong, R.Y.C., Lee, S.K.Y., Law, T.W.F., Law, S.H.W., and Wu, R.S.S. (2002). Rapid detection of six types of bacterial pathogens in marine waters by multiplex PCR. *Water Res.* **36**: 2802–2812.

18. Roszak, D.B., Grimes, D.J., and Colwell, R.R. (1984). Viable but non recoverable stage of *Salmonella enteritidis* in aquatic systems. *Can. J. Microbiol.* **30**: 334–338.
19. Roszak, D.B. and Colwell, R.R. (1987). Survival strategies of bacteria in the natural environment. *Microbiol. Rev.* **51**: 365–379.
20. Way, J.S. et al. (1993). Specific detection of *Salmonella* spp. by multiplex polymerase chain reaction. *Appl. Environ. Microbiol.* **59**: 1473–1479.
21. Cohen, H.J., Mechanda, S.M., and Lin, W. (1996). PCR amplification of the *fim*A gene sequence of *Salmonella typhimurium*, a specific method for detection of *Salmonella* spp. *Appl. Environ. Microbiol.* **62**: 4303–4308 (1996).
22. Waage, A.S., Vardund, T., Lund, V., and Kapperud, G. (1999). Detection of low numbers of *Salmonella* in environmental water, sewage and food samples by a nested polymerase chain reaction assay. *J. Appl. Microbiol.* **87**: 418–428.
23. Riyaz-Ul-Hassan, S., Verma, V., and Qazi, G.N. (2004). Rapid detection of *Salmonella* by polymerase chain reaction. *Mol. Cell Probes* **18**: 333–339.
24. Kapley, A., Lampel, K., and Purohit, H.J. (2000). Development of Duplex PCR for *Salmonella* and *Vibrio*. *World J Microbiol. Biotechnol.* **16**: 457–458.
25. Kapley, A., Lampel, K., and Purohit, H.J. (2000). Thermocycling steps and optimization of multiplex PCR. *Biotechnol. Lett.* **22**: 1913–1918.
26. Clark, C.G., MacDonald, L.A., Ginocchio, C.C., Galan, J.E., and Johnson, R.P. (1996). *Salmonella typhimurium* invA expression probed with a monoclonal antibody to the C-terminal peptide of InvA. *FEMS Microbiol. Lett.* **136**: 263–268.
27. Suzuki, S. et al. (1994). Virulence region of plasmid pNL2001 of *Salmonella enteritidis*. *Microbiology* **140**: 1307–1318.
28. Williamson, C.M., Baird, G.D., and Manning, E.J. (1998). A common virulence region on plasmids from eleven serotypes of *Salmonella*. *J. Gen. Microbiol.* **134**: 975–982.
29. Spierings, G., Elders, R., vanLith, B.M., Hofstra, H., and Tommassen, J. (1992). Characterization of the *Salmonella typhimurium* phoE gene and development of *Salmonella* specific DNA probes. *Gene* **122**: 45–52.
30. Rahn, K. et al. (1992). Amplification of an invA gene sequence of *Salmonella typhimurium* by polymerase chain reaction as a specific method of detection of *Salmonella*. *Mol. Cell Probes* **6**: 271–279.
31. Malorny, B., Hoorfar, J., Bunge, C., and Helmuth R. (2003). Multicenter validation of the analytical accuracy of *Salmonella* PCR: towards an international standard. *Appl. Environ. Microbiol.* **69**: 290–296.
32. Purohit, H.J., Raje, D.V., Kapley, A., Padmanabhan, P., and Singh, R.N. (2003). Genomics tools in environmental impact assessment. *Environ. Sci. Technol.* **37**: 356A–363A.
33. Belgrader, P. et al. (1999). PCR detection of bacteria in seven minutes. *Science* **284**: 449–450.
34. Chen, W., Martinez, G., and Mulchandani, A. (2000). Molecular beacons: a real-time polymerase chain reaction Assay for detecting salmonella. *Anal. Biochem.* **280**: 166–172.
35. Ellingson, J.L.E., Anderson, J.L.A., Carlson, S.A., and Sharma, V.K. (2004). Twelve hour real-time PCR technique for the sensitive and specific detection of *Salmonella* in raw and ready-to-eat meat products. *Mol. Cell. Probes* **18**: 51–57.
36. Kramer, M.F. and Lim D.V. (2004). A rapid and automated fiber optic-based biosensor assay for the detection of *Salmonella* in spent irrigation water used in the sprouting of sprout seeds. *J. Food Prot.* **67**: 46–52.

LYSIMETER SOIL WATER SAMPLING

Ken'ichirou Kosugi
Kyoto University
Kyoto, Japan

INTRODUCTION

The solute concentration of soil water provides important information regarding spatial and temporal distribution of plant nutrients, salinity, and trace elements (1); biological and chemical reactions in soil; and soil and groundwater contamination by industrial wastes and pesticides. For quantifying convective chemical transport in soil, both the water flux and the solute concentration of flowing water should be measured (2). In such a case, a sampling methodology for soil water needs to provide quantitative information on soil water flow as well as to maintain the physical soil environment of the water-sampling profile in a state similar to that of the natural soil profile, as the soil environment affects biological and chemical reactions.

TENSION-FREE LYSIMETER

One of the traditional techniques for sampling soil water is the tension-free lysimeter, which is also referred to as the zero-tension lysimeter. In this method, a horizontally buried pan intercepts infiltrating water, forming a temporary saturated zone above the pan, and the water drains into a sampling bottle. The tension-free lysimeter collects water only when the soil immediately above the pan has a positive pressure (3). Therefore, the soil in the sampling profile above the lysimeter is wetter than the soil surrounding the lysimeter. With this matric pressure gradient, water will tend to flow from the lysimeter into the surrounding dry-soil region, resulting in an underestimation of the natural water flux (e.g., 4,5).

CAPILLARY LYSIMETER

In the capillary lysimeter, a wick made of glass or nylon fibers is attached to the base of the water-collecting pan (Fig. 1) in order to establish drier conditions above the lysimeter and to lessen the problem of bypass flow around the lysimeter (6,7). Thus, this device is also referred to as the wick lysimeter. In order to make the water-sampling rate the same as the natural water flux, the length, number, and material of the wicks of the capillary lysimeter must be designed to match the local soil hydraulic properties as well as to respond to the range of fluxes to be encountered (8). However, Gee et al. (3) pointed out that capillary lysimeters tend to undersample when the soil water flux is less than 1000 mm/yr, even if the wick length is optimized for a given soil type. In a laboratory test, Kosugi (9) showed that the capillary lysimeter resulted in wetter conditions in the sampling soil profile than in a natural soil profile, given heavy irrigation of 18.6 mm/h. Recently, Gee et al. (3) proposed extending an impermeable pipe, of about the same diameter as the water-collecting pan, from the base of the pan to a height of

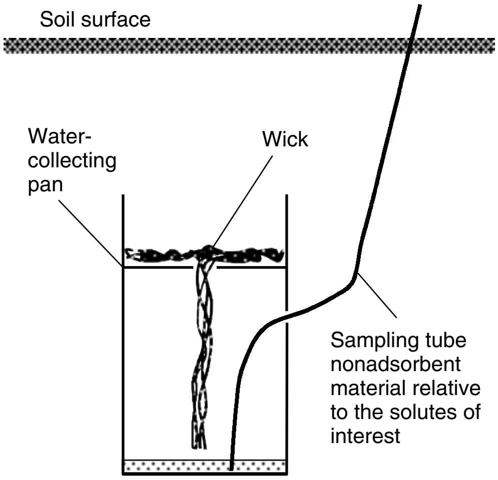

Figure 1. Schematic diagram of the capillary lysimeter.

15 cm or more, in order to control water divergence around the lysimeter. Although they succeeded in reducing the flux divergence to less than 10% for coarse sand by using a 15-cm-long pipe, they concluded that finer-textured soils might require a pipe that was at least 60 cm long. Such a long pipe might disturb the water-sampling profile and might reduce root water uptake by cutting root systems. Detailed descriptions of the capillary lysimeter are found in Selker (10).

TENSION LYSIMETER

A generalized version of the capillary lysimeter is the tension lysimeter, which is one of the most frequently used water-sampling techniques. In this method, water is collected through a porous cup or plate by applying suction with a vacuum source instead of a wick. In most cases, porous cup samplers (also referred to as suction cups) are used because of their ease of installation, simplicity of design, and low cost (Fig. 2). However, owing to their small sphere of measurement, the suction cups provide only "point information," which does not adequately integrate spatial variability (11,12).

With the tension lysimeter method, the suction for water extraction is fixed at an empirically decided value between about 20 and 85 kPa (1,13). As a result, the soil moisture condition in the water-sampling profile can be altered depending on the suction applied. In addition, as the volume of soil that is sampled by the tension lysimeter depends on the soil moisture condition at the time of sampling, the soil hydraulic properties, and the flow properties of the lysimeter (12), the water-sampling rate is not necessarily the same as the natural water flux. Moreover, it is reported that the solute concentrations in the sampled water depend on the duration and degree of the sampler vacuum because water extracted from large pores at low suctions may have a different composition from that extracted from micropores at high suctions (14).

Corwin (1) provides more information on the tension lysimeter with regard to equipment preparation, installation procedure, and problems.

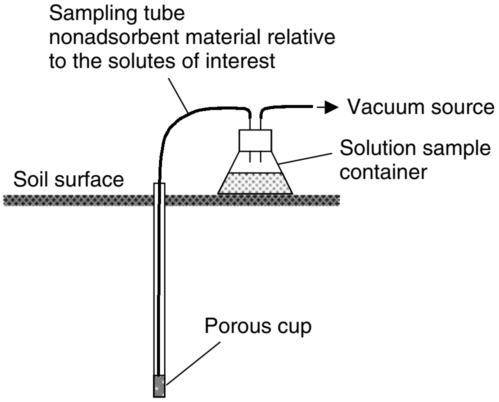

Figure 2. Schematic diagram of the tension lysimeter.

CONTROLLED-TENSION LYSIMETER

In contrast to the above-mentioned techniques, some recent studies have proposed controlling the suction for extracting soil water by referring to matric pressure observations made in the surrounding natural soil profile. The suction control is intended to make the rate of water extraction by the lysimeter the same as the unsaturated water flux in the natural soil profile. Such a "controlled-tension lysimeter" seems to be the most accurate alternative to the methods traditionally used to measure water and convective chemical fluxes in soil.

In the controlled-tension lysimeter, the suction is manually or automatically adjusted to a target value determined from tensiometer observations in the natural soil profile (2,15–17), or the suction control is set so that the soil matric pressure immediately above the porous plate should be similar to the matric pressure at the same depth in the natural soil profile (9,18–20). When the two matric pressures are the same, the rate of water extraction by the lysimeter is expected to be similar to the unsaturated water flux in the natural soil profile, because the water-sampling profile has the same upper (i.e., the flux boundary condition defined by rain and irrigation rate) and lower (i.e., the hydraulic head boundary condition) boundary conditions as the natural profile. Thus, the controlled-tension lysimeter method is soundly in accordance with unsaturated flow theory.

Figure 3 shows a schematic of a type of controlled-tension lysimeter proposed by Kosugi (9) and Kosugi and Katsuyama (20). The equipment consists of two tensiometers (TE) and a ceramic porous plate (PP) connected to a suction system by a sampling tube. The porous plate is buried horizontally in the water-sampling profile. One tensiometer monitors the soil matric pressure immediately above the plate, ψ_a. The other tensiometer monitors the matric pressure, ψ_b, at the same depth in a natural soil profile, adjacent to the sampling profile. In the sampling profile, infiltrating water is extracted through the porous plate by applying suction so that $\psi_a = \psi_b$. The suction system consists of a water collection container (WC) connected to a vacuum pump (PU), a valve for releasing the suction (RV), and a pressure transducer (PT) for monitoring the air pressure, p_c, in the water collection

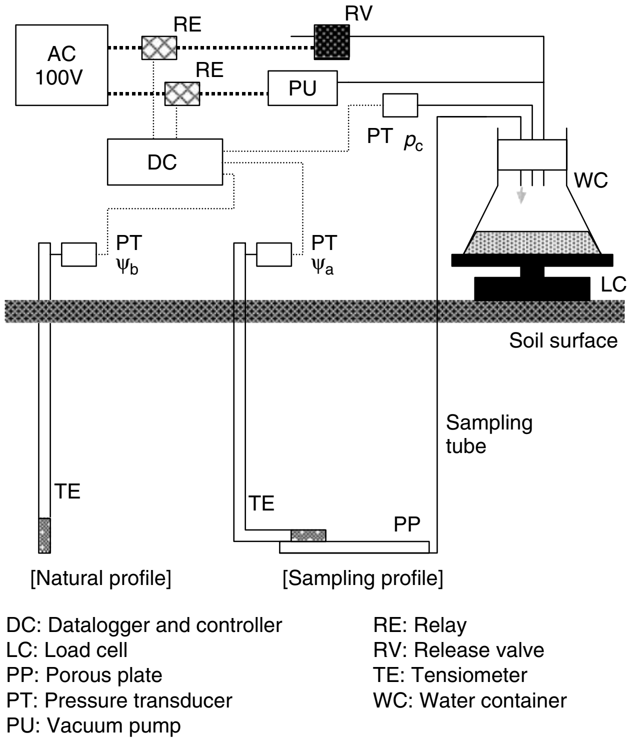

Figure 3. Schematic diagram of the controlled-tension lysimeter (modified from Ref. 9).

container. The values of ψ_a, ψ_b, and p_c are continuously monitored at 3-second intervals. When $\psi_a > \psi_b$ (i.e., the sampling profile is wetter than the natural profile), the vacuum pump is turned on in order to extract the water above the porous plate. When $\psi_a < \psi_b$ (i.e., the sampling profile is dryer than the natural profile), the pump is turned off and the valve is opened in order to immediately stop water extraction by releasing the suction in the water collection container. During a field test for more than 400 d, the lysimeter maintained the soil moisture condition in the sampling profile similar to that in the natural soil profile and extracted a reasonable amount of water (20).

COMMENTS

Problems common to the capillary, tension, and controlled-tension lysimeters include sample biases caused by contaminations and adsorption as the solution passes through wicks and porous cups or plates. Sample contamination can be significantly reduced by pretreatment of the sampler with an acid wash (1 M HCl solution) and repeated rinsing with a salt solution similar to the soil solution that will be sampled (21). Ceramic, Teflon, and stainless steel are commercially available as materials for porous cups and plates. From among these, ceramic is the greatest in trace metal adsorption (22). Various studies have detailed the adsorption of NH_3, P, and K by ceramic cups. Nonetheless, ceramic is most commonly used for porous cups and plates because it has the lowest cost. Although Teflon is the least reactive, it is hydrophobic and has poor flow characteristics. Hence, Teflon is not necessarily a suitable material for porous cups and plates (1).

Another complicating issue is the change in pH that results from CO_2 degassing (23). To reduce CO_2 degassing, it is effective to minimize the gas/liquid ratio in the sample container. Collected samples should be tightly capped with the smallest airspace possible remaining in the sample bottle.

BIBLIOGRAPHY

1. Corwin, D.L. (2002). Suction cups. In: *Methods of Soil Analysis, Part 4—Physical Methods.* J.H. Dane and G.C. Topp (Eds.). Soil Science Society of America, Madison, WI, pp. 1261–1266.
2. Brye, K.R., Norman, J.M., Bundy, L.G., and Gower, S.T. (1999). An equilibrium tension lysimeter for measuring drainage through soil. *Soil Sci. Soc. Am. J.* **63**: 536–543.
3. Gee, G.W., Ward, A.L., Caldwell, T.G., and Ritter, J.C. (2002). A vadose zone water fluxmeter with divergence control. *Water Resour. Res.* **38**: 10.1029/2001WR000816.
4. Jamison, J.M. and Fox, R.H. (1992). Estimation of zero-tension pan lysimeter collection efficiency. *Soil Sci.* **153**: 85–94.
5. Chiu, T.F. and Shackelford, C.D. (2000). Laboratory evaluation of sand under-drains. *J. Geotech. Geoenviron. Eng.* **126**: 990–1001.
6. Holder, M. et al. (1991). Capillary-wick unsaturated zone soil pore water sampler. *Soil Sci. Soc. Am. J.* **55**: 1195–1202.
7. Maeda, M., Liyanage, B.C., and Ozaki, Y. (1999). Water collection efficiency of wick samplers under steady state flow conditions. *Soil Sci. Plant Nutr.* **45**: 485–492.
8. Knutson, J.H. and Selker, J.S. (1994). Unsaturated hydraulic conductivities of fiberglass wicks and designing capillary pore-water samplers. *Soil Sci. Soc. Am. J.* **58**: 721–729.
9. Kosugi, K. (2000). A new sampling method of vertical infiltration water in unsaturated soil without disturbing soil moisture condition. *J. Jpn. Soc. Hydrol. & Water Resour.* **13**: 462–471 (in Japanese with English summary).
10. Selker, J.S. (2002). Passive capillary samplers. In: *Methods of Soil Analysis, Part 4—Physical Methods.* J.H. Dane and G.C. Topp (Eds.). Soil Science Society of America, Madison, WI, pp. 1266–1269.
11. Biggar, J.W. and Nielsen, D.R. (1976). Spatial variability of the leaching characteristics of a field soil. *Water Resour. Res.* **12**: 78–84.
12. Hart, G.L. and Lowery, B. (1997). Axial-radial influence of porous cup soil solution samplers in a sandy soil. *Soil Sci. Soc. Am. J.* **61**: 1765–1773.
13. Tokuchi, N. (1999). Sampling method of solute in soil. In: *Methods of Forest Environment Measurement*. Hakuyusha, Tokyo, Japan, pp. 186–188 (in Japanese)*.
14. Rhoades, J.D. and Oster, J.D. (1986). Solute content. In: *Methods of Soil Analysis, Part 1: Physical and Mineralogy Methods.* A. Klute (Ed.). Soil Science Society America, Madison, WI, pp. 985–1006.
15. Ozaki, Y. (1999). Bury-type lysimeter to collect infiltrated soil water. In: *Proc. Symp., Study on practical problems on monitoring of NO_3-N leaching from crop field*. Tsukuba, Japan. Nov. 4–5. National Agriculture Research Center, Tsukuba, Japan, pp. 9–19 (in Japanese)*.

*The title is tentative translation from the original Japanese title by the authors of this paper.

16. Ciglasch, H., Aoungsawad, K., Amelung, W., Pansombat, K., Sombuttun, T., Tarn, C., Totrakool, S., and Kaupenjohann, M. (2002). Fluxes of agrochemicals in an upland soil in Northern Thailand as investigated with novel techniques. Paper 1566. In: Abst., 17th World Congress of Soil Science. Bangkok, Thailand, Aug. 14–21.
17. Lentz, R.D. and Kincaid, D.C. (2003). An automated vacuum extraction control system for soil water percolation samplers. *Soil Sci. Soc. Am. J.* **67**: 100–106.
18. Duke, H.R. and Haise, H.R. (1973). Vacuum extractors to assess deep percolation losses and chemical constituents of soil water. *Soil Sci. Soc. Am. Proc.* **37**: 963–964.
19. van Grinsven, J.J.M. et al. (1988). Automated *in situ* measurement of unsaturated soil water flux. *Soil Sci. Soc. Am. J.* **52**: 1215–1218.
20. Kosugi, K. and Katsuyama, M. (2004). Controlled-suction period lysimeter for measuring vertical water flux and convective chemical fluxes. *Soil Sci. Soc. Am. J.* **68**: 371–382.
21. Creasey, C.L. and Dreiss, S.J. (1988). Porous cup samplers: cleaning procedures and potential sample bias from trace element contamination. *Soil Sci.* **145**: 93–101.
22. McGuire, P.E., Lowery, B., and Helmke, P.A. (1992). Potential sampling errors: trace metal adsorption on vacuum porous cup samplers. *Soil Sci. Soc. Am. J.* **56**: 74–82.
23. Suarez, D.L. (1986). A soil water extractor that minimizes CO_2 degassing and pH errors. *Water Resour. Res.* **22**: 876–880.

REGULATORY AND SECURITY REQUIREMENTS FOR POTABLE WATER

Suzanne Du Vall Knorr
Ventura County Environmental
Health Division
Ventura, California

INTRODUCTION

Water is essential for the nutrients, metabolic processes, and cellular activity that all living bodies need to survive and thrive. One adult human requires approximately 2.5 liters of potable water from food or liquid sources every day (1). Potable water is water that is fit to drink, free of pathogenic microorganisms or chemicals that are harmful to human health, and free of offensive taste and odors. Drinking water sources include surface waters such as lakes and rivers, and groundwater accessed by wells. Desalination can produce potable water from ocean water, though cost often prohibits widespread use of this method. Public water systems include community, noncommunity, and transient water systems. A water system includes the source water, treatment steps, storage, and the distribution network. Contamination can occur at any point in the water system.

Standard treatment methods typically involve several filtration steps to remove larger particles and objects from the source water and sedimentation to allow smaller particles to settle out. Coagulation and flocculation follow sedimentation to remove unsettleable particles. Disinfection to inactivate pathogens commonly involves chlorine, bromine, or ozone, and fine filtration to remove protozoa. The treated water may also be pH-stabilized. These treatment methods control most pathogens and inorganic contaminants, such as arsenic and lead, and the secondary drinking water standards of taste, odor, color, and balance. The most common types of industrial chemicals, volatile organic compounds (VOCs), such as benzene (a known carcinogen), toluene, and methyl-tertiary-butyl ether, have been routinely found in groundwater. Removing VOCs requires more extensive treatment methods.

Known waterborne disease pathogens can infect healthy adults as well as sensitive populations (young children, pregnant women, elderly adults, and immunocompromised individuals). The Centers for Disease Control (CDC) estimated that over 900,000 incidents of waterborne intestinal disease occur annually, resulting in 900 deaths per year (2). More recently, 39 waterborne disease outbreaks occurred between 1999 and 2000, affecting over 2000 people and resulting in two deaths. Twenty-eight of the 39 were from drinking water that used groundwater sources.

Treatment and protection of drinking water sources will not prevent contamination after delivery (point-of-use). Illness resulting from point-of-use contamination is not considered a waterborne disease incident by the CDC. Household point-of-use treatment methods typically do not remove many organic chemicals (3). Drinking water requirements change as the federal Environmental Protection Agency (EPA) and the CDC further discover the causes of waterborne disease. This article briefly reviews selected contaminants and current regulatory and security requirements that seek to ensure safe and reliable sources of drinking water.

SELECTED BIOLOGICAL CONTAMINANTS

Cryptosporidium

Cryptosporidium parvum oocysts (cysts) are the infective form of a protozoan smaller than *Giardia* that is highly resistant to chlorine disinfection and can cause symptoms of acute watery diarrhea (cryptospiridiosis) from as few as 10 cysts. The surface water treatment rule in effect in 1993 did not require the fine filtration needed to remove *Cryptosporidium* cysts. It is estimated that over 80% of surface water sources are contaminated with *Cryptosporidium* or *Giardia* (4). *Cryptosporidium* can spread in fully chlorinated water, even in the absence of detectable coliforms (5), and medical therapy to treat cryptospiridiosis is limited.

The most notable outbreak of waterborne disease caused by *Cryptosporidium parvum* happened in Milwaukee, Wisconsin, in 1993. This outbreak affected an estimated 403,000 people, resulting in 4400 hospitalizations and 104 deaths (5). It cost city residents an estimated $54 million in medical costs and lost wages (2). Two weeks passed before the waterborne nature of the outbreak was recognized and boil water notices were issued to the public (6). Widespread absenteeism, increased emergency room visits for diarrhea, and a citywide shortage of over-the-counter antidiarrheal medicines led to the detection of the cryptospiridiosis outbreak (5).

Investigation of the outbreak revealed that the public water supply came from Lake Michigan and was treated and chlorinated before distribution. Decreased raw source water quality and decreased coagulation and flocculation effectiveness led to increased turbidity in the treated water. The turbidity standards in effect in 1993 allowed daily turbidity measurements in a month not to exceed 1.0 NTU, though periodic spikes above 1.0 NTU were allowed. The highest turbidity measurement was 1.7 NTU, and the treated water had met all state and federal standards at that time (6). The investigation did not find the original source of contamination, but it was determined that the water quality standards "were not adequate to detect the outbreak" (6). This outbreak led to the Interim Enhanced Surface Water Rule in 1998 and later adoption of a primary regulatory standard requiring 99% cyst removal by filtration. *Cryptosporidium* contaminated water can be disinfected with ozone (4). Point-of-use treatment includes boiling water for 1 minute or filtration with a pore size of 1 micron or smaller (4), which is capable of 99% removal.

Escherichia coli O157:H7

Escherichia coli is a fecal coliform, found in the intestines of both humans and animals, that may contaminate water through contact with sewage discharges, leaking septic systems, or water runoff from animal feedlots. *Escherichia coli* O157:H7 is an enterohemorrhagic strain of one of the four classes of virulent pathogenic *E. coli* that cause gastroenteritis in humans. It produces a potent verotoxin that damages the intestinal lining, often accompanied by severe bloody diarrhea, particularly in sensitive populations. Severe cases of *E. coli* O157:H7 can cause renal failure and loss of kidney function (hemolytic uremic syndrome). Unlike other fecal coliforms, *E. coli* O157:H7 can cause illness from ingesting as few as 10 organisms. One study of a waterborne *E. coli* O157:H7 outbreak involved 243 patients who used an unchlorinated groundwater supply (32 hospitalizations, 4 deaths) in Burdine township, Missouri, between 12/15/89 and 1/20/90 (7). The case study determined that the largest number of cases of bloody diarrhea was in the municipal water supply service area (7). More recently, four waterborne outbreaks of *E. coli* O157:H7 occurred between 1999 and 2000, including one in Albany, New York, in 1999. The suspected source in Albany, New York, was contaminated well water consumed at a fair that caused 127 cases of illness, 71 hospitalizations, including at least 10 children with bloody diarrhea; 14 cases of hemolytic uremic syndrome; and two deaths (8,9).

The detection of coliforms during routine sampling leads to detecting the presence of fecal coliforms, including *E. coli*. Standard treatment methods inactivate or remove *E. coli* O157:H7 or other fecal coliforms. These and other waterborne disease pathogens are included in Table 1.

SELECTED CHEMICAL CONTAMINANTS

Methyl-*tertiary*-butyl Ether

Methyl-*t*-butyl ether (MTBE) was introduced in the 1970s as a substitute for lead in gasoline, to reduce emissions and maintain the oxygen level required by the Clean Air Act. MTBE is water-soluble and moves rapidly through soil into groundwater. It is the second most frequently detected chemical in monitoring wells, according to the U.S. Geological Survey (4). Contamination can come from leaking underground fuel

Table 1. Selected Waterborne Pathogens and Parasites[a]

Pathogens	Disease	Incubation Period	Estimated Infective Dose, # Organisms	Survival in Water, Days
Hepatitis A virus	Type A infectious hepatitis	Up to 50 days	10–50[b]	Up to 140 days in groundwater
Salmonella typhi	Typhoid fever	7–21 days	10,000–100,000	Unknown[c]
E. coli O157:H7	Hemorrhagic colitis; hemolytic uremic syndrome	2–8 days	As few as 10[d]	Up to 20 days in wastewater[e]
E. coli (all others)	Enteropathogenic diarrhea	12–72 hours	10^8	Up to 45 days in groundwater
Vibrio cholerae	Cholera	1–5 days	10^6–10^9	5–16 days in surface water
Shigella spp.	Bacillary dysentery	1–7 days	10–100	35 days in groundwater; 24 months in surface water
Cryptosporidium spp.	Cryptosporidiosis	2–21 days	10–100 cysts	Up to 6 months in a moist environment (ova)
Giardia lamblia	Giardiasis	6–22 days	5–100 cysts	Up to 4 months in surface water (ova)
Leptospira spp.	Leptospirosis, hemorrhagic jaundice	4–19 days	Unknown	3–9 days in surface water
Entamoeba histolytica	Amebic dysentery	2–4 weeks	10–20 cysts	Up to 1 month in surface water (ova)
Ascaris lumbricoides	Ascariasis	2 months	1 cyst, egg or larvae	Up to 7 years in soil

[a] Information compiled from Reference 10, unless otherwise indicated.

[b] Based on contaminated food consumption (11).

[c] 240 days survival in ice (12).

[d] The infective dose is unknown, but the FDA estimates that the dose may be similar to that of *Shigella* spp., based on outbreak data and the organism's ability to transmit person-to-person (13).

[e] Based on information in agricultural research (14).

tanks, leaking gasoline pipelines, and surface spills. It was detected in groundwater in 55% (3180 out of 5738 sites) of underground leaking gasoline tank sites under investigation in 1998 (15).

The oral LD_{50} for MTBE is approximately 4000 mg/kg, and MTBE was tentatively classified as a possible carcinogen (16). MTBE use has been banned in 11 states. In 2000, the California Senate approved an executive order requesting the ban of MTBE and a proposed phase-out by December 31, 2003 (17). This proposed phase-out has not yet occurred. California has set both primary and secondary standards for MTBE. At the federal level, MTBE is an unregulated contaminant, subject to the Unregulated Contaminant Monitoring Rule (UCMR), and is on the EPA drinking water contaminant candidate list (18). The UCMR requires that all large and selected small public water systems monitor for MTBE. MTBE can be removed in varying amounts through granular activated carbon, air stripping, or advanced oxidation.

Perchlorate

Perchlorate is an ion with four chlorine attachments, most commonly used in industrial applications. Ammonium perchlorate is used as an energetics booster in explosives, pyrotechnics, rocket fuel, and highway safety flares (19). It has been detected in groundwater in several states, including the Colorado River which supplies water to millions of people (20), and was found most recently in milk and lettuce (21,22). The concern is that perchlorate interferes with iodide uptake by the thyroid gland, leading to a disruption of thyroid hormones that regulate metabolism and growth; continuous thyroid disruptions could cause a hormone imbalance, particularly in pregnant women, developing fetuses, and infants. Several studies show conflicting results, though two studies indicated adverse health effects in susceptible populations at levels as low as 0.01 μg/kg (20,23). Perchlorate can be removed by treatment with ion exchange, ultraviolet light and peroxide, or peroxide and ozone (24).

A maximum contaminant level (MCL) has not been established by EPA. The Perchlorate Community Right to Know Act of 2003 mandated an enforceable national perchlorate contaminant standard by July 1, 2004; however, the projected research completion date set by the EPA, necessary for setting a standard, extends beyond this deadline (20,25). The EPA set a reference dose (RfD) of 1 ppb (or 2.1 μg/kg for a 70 kg adult) (26) as part of the draft Health Assessment on Perchlorate in 2002, currently under review by the National Academy of Sciences (NAS) (18). NAS estimates that it will complete the review by December 2004 (13). The RfD is an estimate of the daily dose below which health risks would be considered negligible based on lifetime exposure.

The California Department of Health Services (CDHS) was required by law to establish a maximum contaminant level (MCL) for perchlorate by January 1, 2004, although one has not yet been established (27). CDHS set an action level of 6 ppb (28), equivalent to the public health goal (PHG) set by the California Environmental Protection Agency (Cal-EPA) Office of Environmental Health Hazard Assessment (OEHHA). The studies used to establish the PHG were criticized by environmental organizations (20). The action level and the PHG are advisory levels, not regulatory requirements. If an action level is exceeded in groundwater, the water system must notify the local governing body (city council or board of supervisors), and CDHS recommends public notification. If an action level is exceeded by ten times, CDHS recommends discontinuing consumption until treated. Eight other states have set advisory levels for perchlorate, ranging from 1 to 18 ppb (29).

REGULATORY DRINKING WATER STANDARDS

Drinking water must meet the requirements of the Safe Drinking Water Act (SDWA), with monitoring and reporting as specified in the regulation. Individual states may either adopt the federal standard (SDWA) or stricter standards. California adopted stricter standards through the Calderon-Sher Safe Drinking Water Act. Other local drinking water requirements can be found at http://www.epa.gov/safewater/dwinfo.htm. Drinking water sources that include surface waters must be disinfected prior to distribution to the consumer. Current laws and regulations do not require routine treatment of groundwater, unless blended with surface water, although EPA has proposed the Ground Water Rule (GWR) which would require disinfection (8,18). Primary drinking water standards are legally enforceable standards applicable to public drinking water systems, defined as water systems with 15 or more service connections, serving 25 or more people, or operating at least 60 days per year (30). CDHS regulates water systems with five or more service connections. These standards address both biological and chemical contamination and are summarized in Table 2. The maximum contaminant levels (MCL), or specific treatment techniques, such as filtration, used in place of the MCL, as required by the SDWA, are established through evaluation and research to protect public health. Ensuring a safe potable water supply requires vigilance by wholesale and retail water purveyors and water treatment facilities.

SDWA requires testing for coliforms, a group of aerobic or facultative anaerobic bacteria commonly found in soil, vegetation, and the intestinal tracts of warm-blooded animals. The presence of coliforms indicates potential fecal coliform contamination and the probability that other pathogens or parasites are also present. Some fecal coliforms survive in water as long as 10 weeks (32). The maximum contaminant level for coliforms is less than or equal to 5% of the water samples that are positive for coliforms when 40 or more samples are collected per month, or one positive sample when less than 40 samples are collected per month (33). The largest public water systems, serving millions of people, must take at least 480 samples per month (34). Smaller systems must take at least five samples per month, unless the state conducted a sanitary survey during the previous 5 years (34). Under certain circumstances, the smallest systems, those serving less than 1000 people, could take one sample per month, not counting repeat sampling (34). Any samples positive for coliforms must also be tested for fecal coliforms. In California, analytical results must be

Table 2. Current Primary Regulatory Standards for Selected Contaminants[a]

Contaminant	Potential Health Effect	EPA, MCL in mg/L[b]	CDHS, MCL in mg/L[b]
Cryptospiridium spp.	Gastrointestinal illness	99% removal	99% removal
Giardia lamblia	Gastrointestinal illness	99.9% removal	99.9% removal
Total coliforms	Pathogen indicator organism	5.0%[j]	5.0%[j]
Viruses, enteric	Gastrointestinal illness	99.99% removal or inactivation	99.99% removal or inactivation
Arsenic	Circulatory problems, skin damage, increased risk of cancer	0.010	0.05[k]
Asbestos	Intestinal polyps from ingestion	7 MFL	7 MFL
Benzene	Anemia, reduced blood platelets, increased risk of cancer	0.05	0.001
Chromium[e]	Variable	0.1	0.05
Copper[a]	Gastrointestinal illness, liver and kidney damage	1.3	1.3
Free cyanide	Nervous system and thyroid problems	0.2	0.2
Ethylene dibromide (EDB)	Liver, stomach, kidney, and reproductive problems; increased risk of cancer	0.00005	0.00005
Fluoride	Bone disease; dental discoloration and pitting	4.0[f,g]	2.0
Lead[c]	Developmental effects in children	0.015	0.015
Inorganic mercury	Kidney damage	0.002	0.002
MTBE	Tentatively classed as possible carcinogen	None	0.013[g]
Nitrate (as nitrogen)	Blue baby syndrome	10	10
Nitrite(as nitrogen)	Blue baby syndrome	1	1
Polychlorinated biphenyls (PCBs)	Immune, thymus, reproductive and nervous system damage, increased risk of cancer	0.0005	0.0005
Styrene (vinyl benzene)	Liver and kidney damage; circulatory problems	0.1	0.1
Toluene(methyl benzene)	Nervous system; liver and kidney damage	1.0	0.15
Trichloroethylene (TCE)[h]	Liver damage; increased risk of cancer	0. 005	0.005
1,1,1-Trichloroethane (1,1,1-TCH)	Liver, nervous system, and circulatory system problems	0.2	0.2
Trihalomethanes (TTHMs)	Disinfection by-product; liver, kidney, central nervous system effects; increased risk of cancer	0.08[f]	0.10
Uranium	Kidney toxicity, increased risk of cancer from radionuclides	30 μg/L	20 pCi/L
Vinyl chloride	Increased risk of cancer	0.002	0.0005
Xylenes (total)	Nervous system damage	10	1.750[d]

[a] Information from U.S. EPA and California Department of Health Services websites, unless otherwise indicated.

[b] MFL = million fibers greater than 10 μm in length per fluid liter; pCi/L = picocuries per liter; 1 pCi = 37 becquerels (Bq) per cubic meter. Food intake is generally measured in Bq/kg or Bq/L. EPA uses 0.9 pCi/μg as a conversion factor (11).

[c] U.S. EPA has a maximum contaminant level goal (MCLG) of 1.3 mg/L for copper and 0.015 mg/L for lead based on the number of samples taken at point-of-use. Additional monitoring, corrosion control, and treatment are required if the regulatory action level is exceeded in more than 10% of samples. Public education is required if the regulatory action level is exceeded for lead.

[d] As a single isomer or the sum of all isomers.

[e] All chromium levels. Hexavalent chromium causes severe diverse health affects, including cancer. It was the groundwater contaminant affecting public health in Hinkley, California, portrayed in the movie "Erin Brockovich."

[f] EPA standard effective January 1, 2004 for all water systems.

[g] Also has a secondary standard.

[h] TCE was the groundwater contaminant involved in a large leukemia cluster (28 cases in 20 years, four times the national average) in Woburn, Massachusetts, portrayed in the movie, "A Civil Action" (31).

[j] Based on 40 or more samples per month, or one positive sample when less than 40 samples are collected per month.

[k] California must adopt the federal standard, or stricter, by 2006. EPA handles enforcement until the new standard is adopted.

reported electronically by the tenth day of the following month. The water supplier must retain bacteriological analysis records for at least 5 years, and chemical analysis records for at least 10 years. Wholesale and retail water suppliers are required to provide an annual water quality report (AWQR) to their customers. In California, the Consumer Confidence Report (CCR) acts as the AWQR, but must adhere to CDHS regulations, which requires contaminant reporting in easy to read table format (35).

A violation of a primary drinking water standard requires that the water system take corrective action and notify its customers. The notification must include a description of the violation, what it means using appropriate health effects language, when it occurred, what action they should take, and who to contact for more information. Violations are classified by tiers under the Public Notification Rule (36,37). A Tier 1 violation requires public notification within 24 hours by radio, television, hand delivery, or other approved means, and concurrent consultation with the appropriate implementing agency (CDHS or EPA where it directly implements the program). Examples of Tier 1 violations are exceeding the MCL for nitrates or nitrites, failure to test for fecal coliforms or *E. coli*, a repeat fecal coliform positive sample, a waterborne disease outbreak, or other emergency. A Tier 2 violation requires public notification within 30 days. An example of a Tier 2 violation is exceeding the MCL for total coliforms or turbidity. Water system customers must receive notification of a Tier 3 violation within 12 months of the violation. Examples of Tier 3 violations are testing procedure violations or nonfecal coliform monitoring violations. The implementing agency must receive a copy of the public notification within 10 days of initiation for Tier 2 and 3 violations. The water system must send out repeat notifications every 3 months for Tier 2 violations, or every 12 months for Tier 3 violations, for as long as the violation exists. Copies of public notices for Tier 1, 2, and 3 violations must be retained for at least 3 years. The water system must also conduct public notification within 12 months of available monitoring results for unregulated contaminants (36).

The implementing agency has formal and informal enforcement actions available to address a failure on the part of the water system to take corrective action or to notify the public or the implementing agency properly. Formal enforcement includes administrative orders, penalties, and civil and criminal actions. States may also refer a recalcitrant water system to the EPA for further enforcement action.

Secondary drinking water standards are nonenforceable guidelines for secondary contaminants that may cause cosmetic effects, such as tooth discoloration from high fluoride levels, or aesthetic effects such as a disagreeable taste, odor, or color. Secondary contaminants are defined as any contaminant that adversely affects taste, odor, or appearance, or that may cause a substantial number of people to discontinue use, or that may otherwise adversely affect public welfare. Selected secondary drinking water standards are summarized in Table 3. The water purveyor must notify the public when a secondary drinking water standard is exceeded. Treatment is at the option of the purveyor or the customer at point-of-use.

Table 3. Current Secondary Drinking Water Standards for Selected Contaminants[a]

Constituent	U.S. EPA[b]	CDHS[b]
Color	15 color units	15 color units
Chloride	250 mg/L	250 mg/L
Fluoride	2.0 mg/L	None[c]
Iron	0.3 mg/L	0.3 mg/L
Odor	3 threshold odor units	3 threshold odor units
pH	6.5–8.5	6.5–8.5
Turbidity	0.5 NTU	0.5 NTU
Sulfates	250 mg/L[d]	250 mg/L[d]
Total dissolved solids	500 mg/L	500 mg/L
MTBE	None	0.005 mg/L

[a] Information from U.S. EPA and California Department of Health Services websites, unless otherwise indicated.
[b] NTU = nephelometry turbidity units. 0.5 NTU is required for 95% of the daily measurements taken in a monthly period, with no spikes above 1.0 NTU.
[c] Primary standard in California.
[d] 250 mg/L is the lower limit; 500 mg/L is the upper limit.

ENHANCED SECURITY REQUIREMENTS

The danger of a contaminated water supply to large populations became one focus of the Department of Homeland Security as part of a national directive to protect critical infrastructures. An intentionally introduced biological or chemical contaminant may have a limited effect unless introduced in sufficient quantity. However, the 1993 Milwaukee *Cryptosporidium* outbreak demonstrated that sometimes a small quantity is all that is necessary to affect a significant portion of the population adversely.

Two federal laws involving two agencies were passed in 2002: the Homeland Security Act and the Public Health Security and Bioterrorism Preparedness and Response Act (Bioterrorism Act). Both laws pertain to public water systems. The Homeland Security Act requires reporting knowledge of a potential threat to a critical infrastructure, including public water systems, to the Department of Homeland Security. The Bioterrorism Act amended the SDWA to address terrorist and intentional or malevolent acts (threats) against public drinking water supplies. It requires that each public water system serving more than 3300 people conduct a vulnerability assessment (VA), to submit the VA to EPA by deadlines mandated by system size, to prepare an emergency response plan (ERP), and to implement any necessary security enhancements. Interaction between both laws and both agencies must be coordinated within a legal framework.

The purpose of the VA is to evaluate the susceptibility of the water system to potential threats and the risk to the community served by the water system and to plan for reducing risks. The VA must include six basic elements (38):

1. characterization of the water system, its mission, and objectives
2. identification and prioritization of adverse consequences to avoid
3. determination of critical assets that may be subject to a threat and potential undesirable consequences
4. assessment of the probability of different types of threats
5. evaluation of existing countermeasures
6. analysis of current identified risks and development of a prioritized plan to reduce these risks

The first element involves prioritizing the utility services and the facilities necessary to provide those services to the community it serves. Characterizing the water system may include a review of operating procedures and management practices for each element within the water system. The systems' objectives also depend on other utilities, such as power, and the need to evaluate the risk to the system if a threat affects another critical infrastructure.

The next two elements involve a review of each component within the water system in relation to any possible threat, including but not limited to physical barriers, water collection, treatment and pretreatment, storage facilities, distribution (including all piping, conveyances, and valves), automated systems, computer systems, and chemical use and storage areas. Preparing these two elements must consider how a threat could affect each element and the potential consequences, such as service disruption, illness, death, or economic impacts. For example, the introduction of a threat contaminant (TC) could involve concentration changes at every stage in the water system as it mixes with clean water or involve reactions with water or piping. Even a weak solution could adversely affect public health, either as an acute exposure to TC dangerous at low doses, or as a chronic health threat to low TC doses over long exposure periods. Each potential disruption in a system's ability to provide a safe and reliable water supply requires prioritization based on potential magnitude.

The fourth element involves identifying different types of threats the system has experienced in the past or that could occur in the future and the probability of each identified possibility. Specific types of threats such as a fire, vandalism, or cyberattacks, might already have specific countermeasures in place. The fifth element involves evaluating the capabilities and limitations of existing risk reduction methods. Examples of these include intrusion detection systems, delay mechanisms (locks, fencing, vehicle access points), and water quality monitoring alarms. The sixth element requires reviewing the vulnerabilities and countermeasures identified in the first five elements and providing recommendations for improvement.

VAs were to be completed and sent directly to EPA no later than 3/31/03 for water systems serving 100,000 or more people, no later than 12/31/03 for systems serving 50,000 to 99,000 people, and no later than 6/30/04 for water systems serving less than 50,000 people (39). Compliance with this requirement was voluntary for water systems serving less than 3300 people and updates to the VA are also voluntary (39). The EPA will determine system size based on data in the Safe Drinking Water Information System, submitted by each state in 2002 (30).

Some state agencies may require duplicate submission; neither CDHS nor Cal-EPA required a copy of the VA. The EPA plans to review each VA (38).

The next requirement of the Bioterrorism Act was for each water system to prepare an Emergency Response Plan (ERP), incorporating the results of the VA. Emergency response planning is essential for the water system to continue to provide a safe and reliable water supply in the event of an emergency. The ERP must include all procedures and equipment that will significantly reduce the impact of a threat, including basic information such as owner, address, emergency contacts, system components, population served, and number of service connections. The ERP must also identify alternative equipment and water sources, access control, emergency response, and incident specific procedures. Examples of incident specific procedures include policies for handling bomb threats, contamination incidents, or workplace violence.

The ERP is to be maintained securely on site. The water system is required to certify to EPA that the ERP was completed within 6 months of VA submission. The water system has the responsibility for ensuring and upgrading security procedures and equipment.

CONCLUSION

Drinking water laws and regulations were enacted to protect public health. Both the EPA and CDHS also provide information for individuals on private wells so they can protect their water quality. Some waterborne disease outbreaks demonstrate a need to evaluate water quality monitoring methods, and, as in the case of the 1993 Milwaukee cryptospiridiosis outbreak, could lead to a change in regulatory standards. Concerns over the security of critical infrastructures and evaluation of existing systems may also lead to changes that will ultimately enhance protecting public health. As science and technology advances and as more is known about the health affects of waterborne chemical and biological contaminants, drinking water laws and regulations will continue to change. The goal of both the regulatory agencies and the water purveyors should always be to ensure a safe and reliable drinking water supply.

BIBLIOGRAPHY

1. Whitney, E. and Boyle, M. (1984). *Understanding Nutrition*, 3rd Edn. West, pp. 7, 368, 369.
2. Lee, G.F. and Jones-Lee, A. (1993). "Public Health Significance of Waterborne Pathogens," Report to Cal-EPA Comparative Risk Project, December 1993, http://www.gfredlee.com/phealthsig_080801.pdf, accessed September 21, 2004.
3. Santa Clara Valley Water District (SCVWD). (2003). Frequently asked questions, perchlorate. May 20, 2003. http://www.valleywater.org/Water/Water_Quality/Protecting_your_water/_Perchlorate_Information/_pdf/Perchlorate FAQ 5-21-03.pdf −436.0 KB, accessed July 28, 2004.
4. Los Angeles Department of Water and Power (LADWP). (2003). MtBE, General Information. http://www.ladwp.com/ladwp/cms/ladwp001420.jsp, accessed July 26, 2004.
5. MacKenzie et al. (1994). A massive outbreak in Milwaukee of *Cryptospirodium* infection transmitted through public water supply. *N. Engl. J. Med.* **331**(3): 161–167.
6. Centers for Disease Control (CDC). (1996). Surveillance for Waterborne-Disease Outbreaks—U.S., 1993–1994, April 12, 1996. http://www.cdc.gov/epo/mmwr/preview/mmwrhtml/00040818.htm, accessed July 23, 2004.
7. Swerdlow et al. (1993). A waterborne outbreak in Missouri of *Escherichia coli* O157:H7, associated with bloody diarrhea and death. Centers for Disease Control, Enteric Diseases Branch. *Annu. Intern. Med.* **119**(3): 249–250.

8. Centers for Disease Control (CDC). (2002). Surveillance for Waterborne-Disease Outbreaks—U.S., 1999–2000, Surveillance Summaries. November 22, 2002, http://www.cdc.gov/epo/mmwr/preview/mmwrhtml/ss5108a1.htm, accessed September 21, 2004.
9. New York Department of Health (DOH). (2000). Health commissioner releases *E. coli* outbreak report. March 31, 2000. http://www.health.ny.us/nysdoh/commish/2000/ecoli.html, accessed September 27, 2004.
10. Salvato, J.A. (1992). *Environmental Engineering and Sanitation*, 4th Edn. John Wiley & Sons, Hoboken, NJ.
11. Food and Drug Administration (FDA). (2001). Center for Food Safety and Applied Nutrition. Outbreaks associated with fresh and fresh-cut produce, incidence, growth, and survival of pathogens in fresh and fresh-cut produce. September 30, 2001, http://vm.cfsan.fda.gov/~comm/1 ft3-4a.html, accessed September 20, 2004.
12. Health Canada. (2001). Office of Laboratory Security, "Material Safety Data Sheet-Infectious Substances, *Salmonella typhi*. May 15, 2001. http://www.hc-sc.gc.ca/pphb-dgspsp/msds-ftss/msds134e.html, accessed September 10, 2004.
13. Food and Drug Administration (FDA). (2003). Center for Food Safety & Applied Nutrition, Foodborne Pathogenic Microorganisms and Natural Toxins Handbook, Bad Bug Book, *Escherichia coli* O157:H7, updated January 7, 2003, http://vm.cfsan.fda.gov/~mow/chap15.html, accessed September 10, 2004.
14. McCaskey, T.A. et al. (1998). Constructed wetlands controlling *E. coli* O157:H7 and *Salmonella* on the farm. *Highlights Agric. Res.* **45**(1): http://www.ag.auburn.edu/aaes/communications/highlights/spring98/ecoli.html, accessed September 20, 2004.
15. Fogg, G.E. et al. (1998). Hydrologic Sciences, Department of Land, Air, and Water Resources, University of California Davis, Impacts of MTBE on California Groundwater. *Health and Environmental Assessment of MTBE*, Volume 4, Report to the Governor and Legislature of the State of California as Sponsored by SB 521, November 12, 1998. http://www.tsrtp.ucdavis.edu/mtberpt/vol4_1.pdf, accessed September 10, 2004.
16. Office of Environmental Health Hazard Assessment (OEHHA). (1999). California Environmental Protection Agency (Cal-EPA). MTBE in drinking water, California public health goal", March 12, 1999, p. 33. http://www.oehha.org/water/phg/pdf/mtbe_f.pdf, accessed September 27, 2004.
17. Office of the Governor. (2002). Governor Davis allows more time for ethanol solution. Press Release PR02:139, March 15, 2002. http://www.energy.ca.gov/releases/2002_releases/2002-03-15_governor_mtbe.html, accessed September 27, 2004.
18. Environmental Protection Agency (EPA). (2002). National Academy of Sciences' Review of EPA's Draft Perchlorate Environmental Contamination: Toxicological Review and Risk Characterization. http://cfpub2.epa.gov/ncea/cfm/recordisplay.cfm?deid=72117, accessed September 10, 2004.
19. Calgon Carbon Corporation. (2001). Groundwater Remediation Technologies Analysis Center, About Perchlorate. http://www.perchloratenews.com/about-perchlorate.html, accessed July 28, 2004.
20. Environmental Working Group (EWG). (2004). Rocket fuel in drinking water: Perchlorate pollution spreading nationwide. http://www.ewg.org/reports/rocketwater/, accessed July 28, 2004.
21. Lee, M. (2004). Perchlorate threat looms for farmers. *The Sacramento Bee*, July 26. http://www.sacbee.com/content/news/v-print/story/10140109p-11060888c.html, accessed July 26, 2004.
22. Bustillo, M. (2003). Tests link lettuce to toxin in water. *The Mercury News*, April 28. http://www.mercurynews.com/mld/mercurynews/news/5734263.htm, accessed July 28, 2004.
23. Silva, A. (2002). Perchlorate goes from obscurity to point of controversy. *San Bernardino County Sun*, December 14. http://www1800lawinfo.com/practice/printnews.htm?story_id=4009, accessed July 28, 2004.
24. Calgon Carbon Corporation. (2001). Groundwater Remediation Technologies Analysis Center, Case study, Calgon Carbon Corporation—ISEP (R) Continuous Ion Exchange. http://www.perchloratenews.com/case-study1.html, accessed July 28, 2004.
25. Spillman, B. (2003). Boxer's perchlorate bill doesn't give EPA enough time. *The Desert Sun*, March 10. http://www.thedesertsun.com/, accessed July 28, 2004.
26. Food and Drug Administration (FDA). (2004). Center for Food Safety and Applied Nutrition, Office of Plant and Dairy Foods. Perchlorate questions and answers. June 22, 2004. http://www.cfsan.fda.gov/~dms/clo4qa.html, accessed July 28, 2004.
27. Thompson, D. (2004). State a year tardy in water standards for rocket fuel. *Contra Costa Times*, January 29, 2004. http://www.contracostatimes.com/mld/cctimes/news/7823750.htm, accessed July 28, 2004, SB 1822 Statutes of 2002, available for review at http://www.leginfo.ca.gov.
28. California Department of Health Services (CDHS). (2004). Perchlorate in drinking water: action level July 1, 2004, http://www.dhs.ca.gov/ps/ddwem/chemicals/perchl/actionlevel.htm, accessed July 28, 2004.
29. Office of Environmental Health Hazard Assessment (OEHHA). (2003). California Environmental Protection Agency (Cal-EPA). Frequently asked questions about the public health goal for perchlorate. http://wwwoehha.org/public_info/facts/perchloratefacts.html, accessed July 28, 2004.
30. Environmental Protection Agency (EPA). (2003). Large water systems emergency response plan outline: Guidance to assist community water systems in complying with the public health security and bioterrorism preparedness and response act of 2002. Document EPA 810-F-03-007, July 2003. http://www.epa.gov/safewater/security, accessed September 21, 2004.
31. Durant, J., Chen, J., Hemond, H., and Thilly, W., (1995). Elevated incidence of childhood leukemia in Woburn, Massachusetts. *Environ. Health Perspect. Suppl.* **103**(56): September.
32. Kerr, M. et al. (1999). Survival of *Escherichia coli* O157:H7 in bottled natural mineral water. *J. Appl. Microbiol.* **87**(6): 833–849.
33. California Health and Safety Code, Section 111145, 111165, 116455 http://www.leginfo.ca.gov/accessed August 4, 2004.
34. Environmental Protection Agency (EPA). (2002). *E. Coli* O157:H7 in drinking water. http://www.epa.gov/safewater/ecoli.html, November 26, 2002, accessed September 10, 2004.
35. California Department of Health Services (CDHS). (2004). Consumer Confidence Reports February 13, 2004, http://www.dhs.ca.gov/ps/ddwem/publications/ccr/ccrguideca-02-13-04.pdf, accessed September 28, 2004.
36. California Department of Health Services (CDHS). (2002). Public notification requirements for drinking water regulation violations—proposed regulations. Document R-59-01, pp. 49–52, 58, November 1, 2002.

37. Environmental Protection Agency (EPA). (2002). Final state implementation guidance for the public notification (PN) rule. http://www.epa.gov/safewater/pws/impguide.pdf, accessed September 28, 2004.
38. Ware, P. (2002). EPA will check drinking water assessments for completeness but will not grade them. December 27, 2002. http://www.amsa-cleanwater.org/advocacy/security/articles02.cfm, accessed September 10, 2004.
39. Safe Drinking Water Act, 2002, Title IV—Drinking Water Safety and Security, Sections 1433–1435.

READING LIST

Environmental Protection Agency (EPA). (2002). MTBE in drinking water. http://www.epa.gov/safewater/mtbe.html, accessed September 14, 2004.

Environmental Protection Agency (EPA). (2002). Proposed ground water rule. EPA Document 815-F-00-003, April 2002. http://www.epa.gov/safewater/gwr.html, accessed September 27, 2004.

Environmental Protection Agency (EPA). (2002). Radionuclides in drinking water. November 26, 2002. http://www.epa.gov/safewater/radionuc.html, accessed September 27, 2004.

Los Angeles Department of Water and Power (LADWP). (2003). Cryptosporidium, General Information. http://www.ladwp.com/ladwp/cms/ladwp001423.jsp, accessed July 26, 2004.

A WEIGHT OF EVIDENCE APPROACH TO CHARACTERIZE SEDIMENT QUALITY USING LABORATORY AND FIELD ASSAYS: AN EXAMPLE FOR SPANISH COASTS

T.A. DelValls
M.C. Casado-Martínez
I. Riba
Facultad de Ciencias del Mar y Ambientales
Cádiz, Spain

M.L. Martín-Díaz
J.M. Blasco
Instituto de Ciencias Marinas de Andalucía
Cádiz, Spain

N. Buceta Fernández
Centro de Estudios de Puertos y Costas
Madrid, Spain

THE WEIGHT OF EVIDENCE APPROACH—DEFINITIONS AND LINES OF EVIDENCE

During recent years, different initiatives have been carried out to use multiple lines of evidence with the aim to assess sediment quality in aquatic ecosystems. One of the most useful and widely used methods was the design and application of integrative assessments to establish sediment quality (1). These methods comprise the synoptic use of different methodologies such as chemical analysis in sediments to establish contamination, in sediment toxicity to address the biological effects under laboratory conditions, chemical residue analyses in organism tissues to determine the bioavailability of the contaminants, and the structure of the benthic community or the histopathological lesions in resident organisms to determine the biological effects under field conditions (1,2). One of these integrative assessments widely used in the studies of sediment quality is the Sediment Quality Triad that was used synoptically for the first time in Spain and in Europe in the early 1990s (3,4). This method also permits using convenient statistical tools to derive sediment quality guidelines (SQGs), which can be used to derive concentrations associated and not associated with the biological effect in the overlapping area of the different lines of evidences (Fig. 1).

Recently, these integrative assessments, and especially the SQT, have suffered some modifications in their interpretation and the representation of their results. In this sense, new methodologies have been incorporated to establish the chronic effects under laboratory conditions, some new parameters were incorporated to avoid the influence of natural casuistry, and especially some new format or representation of the data have been proposed to avoid loss of information, which lacks association with the classic representations (1,6).

The main change suffered by the integrative assessment is the change in the names of these methods and they are more precisely known during recent years as weight of evidence approaches (WOE) and defined as concepts instead of as methods. The WOE approach is defined by Burton et al. (7) as "...a process used in environmental assessment to evaluate multiple lines-of-evidence concerning ecological conditions... and include

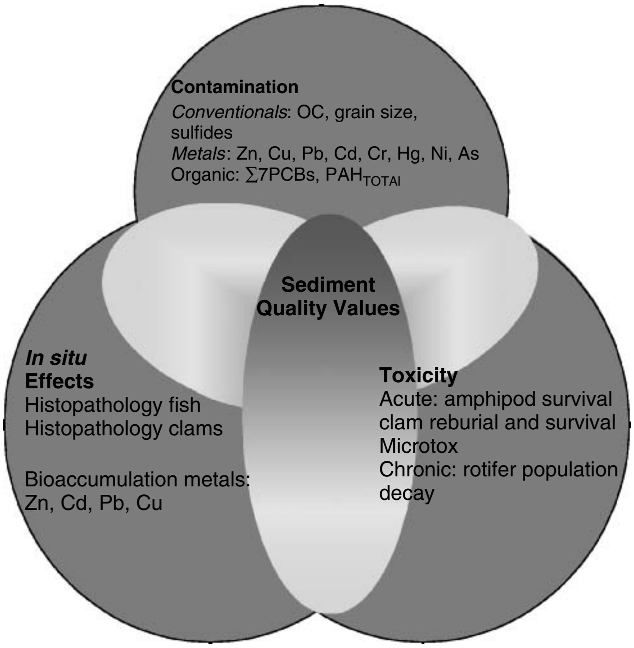

Figure 1. Schematic representation of a widely used weight of evidence approach such as the Sediment Quality Triad (SQT). Adapted from Riba et al. (5).

assessment of impairment, prioritization of site contamination, and decision-making on management actions." This definition goes farther than the classic definition related to the integrative assessment and the Sediment Quality Triad that were considered methods or concepts instead of processes, including the decision-making process.

In this sense, we have reviewed here the evolution of the different lines of evidence that have been growing during recent years associated with the use of this kind of WOE approach in Spain from the classic application of the Triad in 1990s and with emphasis on the potential new application for the decision-making processes in the management of sediment quality, including dredged material characterization. Basically, it describes the application of the different lines of evidence in new environmental problems in Spain that have been and are the marine accidental wastes, the management of dredged material, and the assessment of sediment quality in aquatic ecosystems. It describes the use of a conveniently designed battery of sediment toxicity tests to be incorporated in the tier testing approach for the management of the dredged material in Spain, the use of caged animals to help in the relationship between biological effects measured in laboratory and field conditions, the use of bioaccumulation studies either at field or at laboratory conditions to determine the bioavailability of contaminants, and finally, the use of histopathological methods as part of the weight of evidence approach and its use in combination with bioaccumulation of contaminants in different tissues to derive tissue quality values (TQVs) (5,8,9). Finally, some recommendations of the correct process to integrate these results under a weight of evidence approach using the sediment quality triad is proposed for further sediment quality assessment in Spain.

DESIGN OF ACUTE SEDIMENT TOXICITY TESTS FOR THE CHARACTERIZATION OF DREDGED MATERIAL IN SPAIN

The incorporation of the toxicity test in the management of dredged material disposal is part of a tier testing that is under final development by CEDEX and uses a battery of sediment toxicity tests as a complementary tool to the classic physicochemical analysis using a modification of the sediment quality guidelines (10). This battery of toxicity tests includes the commercial screening test Microtox® in the same samples in which chemical analyses are necessary. If doubts about hazardousness of these samples are found in the other two tests, the bioassay using amphipods and the bioassay using sea urchin larvae are conducted classifying the samples as hazardous if any of the tests show positive responses compared with the biological guidelines proposed.

THE USE OF CHRONIC, BIOACCUMULATION, AND SUBLETHAL BIOASSAYS TO CHARACTERIZE SEDIMENT QUALITY IN DREDGED MATERIAL

One of the recent improvements proposed by different agencies is the possibility to use chronic bioassays having endpoints different from mortality. The main objective is to compare the chronic and acute responses using bioassays to characterize the quality of sediments or dredged material.

The toxicity tests involve the comparison of the three proposed acute sediment toxicity tests versus different chronic bioassays using fish (*Solea senegalensis*) and clams (*Scrobicularia plana*) and using different sublethal endpoints such as the modification of biomarkers of exposure (methallotioneins, EROD, enzymatic activities of the oxidant stress, etc.) and biomarkers of effect (vitellogenin/vitellin; histopathology). Besides, the complete cycle of an amphipod species (*Corophium volutator*) is used to address the sediment quality of this material. Results suggest that the biomarkers are a powerful tool to identify the effect in the gray areas in which acute toxicity tests do not show positive responses. In this sense, the use of histopathological studies has permitted the identification of moderate toxicity associated with PCBs that showed absence of effect using the battery of acute toxicity tests (11).

The complexities at issue are even greater in the case of evaluating concerns for fish and wildlife at a management site. Assessing the potential for impacts of concern on such receptors requires information about the bioaccumulation potential of contaminants from the sediments or dredged material (e.g., results from a bioaccumulation test), how receptors use the management site, and how specific contaminants will move through the local food web. The proposed chronic tests can be used in the same design as bioaccumulation tests and different chemicals could be analyzed in the same tissues in which biomarker, including histopathological, analysis was performed. These studies will permit one to address the mobility of the different contaminants through the food chain and, specifically in some contaminants, to identify the hazardousness associated with the dredged material, such as in the case of PCBs that can suffer biomagnification processes.

The link between the chemical residues and the analysis of biomarker, especially including histopathological diseases in the same tissues by means of bioaccumulation-chronic tests, could permit one to derive tissue quality values (TQVs) in a similar approach to those used to derive sediment quality values or guidelines (SQGs) as reported by Riba et al. (5,9). These TQVs could be used to prevent the risk associated with the contaminants in human food of commercial species and contaminants that cannot be identified by the standardized acute toxicity tests.

USING *in situ* APPROACHES TO CHARACTERIZE DREDGED MATERIAL IN SPAIN—CAGING ANIMALS

Sediment toxicity tests provide information about toxicity or hazard to sediment-dwelling organisms. However, exposure conditions for benthos at a management site may differ significantly from those occurring in the laboratory. Understanding this difference will require information about the management site, including hydrodynamics (is the site dispersive or depositional?), the area extent of coverage once the material is deposited, how organisms at the disposal site interact with the sediment and overlying water, and so on. In this sense, different initiatives have been carried out in Spain to address the biological effects under field conditions funded by the Ministry of Science and Technology (DREDGED REN2002_01699/TECNO; TRIADA; VEM2003_−20563/INTER).

One of the studies designed to address part of the deficiencies associated with laboratory assays is the use of caged animals at the dredged or disposal site to the measurement of different endpoints, either lethal (12,13) or sublethal (8,13). Some designs of benthic chambers are used with this aim; an example is shown in (Fig. 2). The organisms, basically benthic, are confined in the chamber for ideally the same period of exposure as that used in the bioaccumulation-chronic laboratory assays. The tests need the anchorage of either field-collected or laboratory-cultured animals under field conditions exposed to the sediment of concern. The advantages compared with laboratory-exposed devices are that the exposure conditions are representatives of actual conditions at the site. The disadvantages of these field assays are related to the complexities of field-based experimentation, including the difficulties to discriminate the endpoints from noncontaminant variables.

In caged experiments, the same parameters as those registered in the laboratory assays are included in the bioassays under field conditions: biomarkers of exposure and effects, chemical concentration in the biological tissues, mortality, and other sublethal endpoints. Another advantage associated with these studies is the potential complementary use with the laboratory tests that permits one to determine the biological effect under field conditions and, by comparing with the laboratory responses, to discriminate some natural causality that is associated with other kinds of approaches to establish biological effect under field conditions, such as the macrobenthic community studies (10).

Figure 2. Schematic representation of the benthic chambers used to expose caged organisms under field conditions. The animals are transported from the laboratory in cold poliespan boxes to the boat. The chambers (50 cm by 25 cm by 15 cm) were immersed and anchored in each sampling station from every port by scuba divers.

FUTURE DEVELOPMENTS

The complexity of the matrix defined by sediments or dredged material usually located in highly dynamic ecosystems, such as the littoral areas and especially the estuaries, ensures that successfully using any analytical tool requires having an appreciation for the uncertainties associated with its use. No single test or evaluative tool is able to provide a complete and accurate picture of the complexities inherent in evaluating contaminated sediments. In Table 1, some of the sources of uncertainty associated with applying biological tests to manage dredged material are shown.

Notwithstanding the uncertainties included in Table 1, biological tests provide the most direct and certain means for assessing the potential for toxicity and contaminant bioavailability in sediments.

Other contaminants exist that are not usually considered in the studies of sediment quality, such as the presence of pathogens in sediments. Microbial pathogens in aquatic systems can contaminate drinking water supplies and/or shellfish and are responsible for hundreds of beach closure events annually. The proximity of sediments to be dredged to sources of pathogens (e.g., sewage outfalls and agricultural runoff) and the presence of conditions favorable for the long-term viability of microbes makes pathogens in sediments or dredged material a matter of growing concern. Unfortunately, standard microbial methods in use worldwide fail to provide adequate information on which to base management decisions. Some pathogens of concern are not associated with the widely used fecal coliform indicator. In addition, some pathogens cannot be cultured from environmental samples using standard media and conditions. Molecular methods are currently being developed to address the deficiencies in current approaches. The molecular approaches under development are not based on the need to culture the pathogens of concern but on extracting and analyzing their DNA. The use of gene-probe technology has the potential to provide a reliable and highly specific method for detecting a pathogen as well as its virulence potential, that is, its ability to cause disease. When such methods are refined and available for widespread use, more credible assessments of pathogens in sediments will be possible.

Another point of future research is related to the effects of two key environmental variables, such as pH and salinity in aquatic ecosystems. These two variables control most processes occurring in these environments, and they are especially of interest in the border of aquatic systems, such as those defined by estuaries between marine and freshwater environments. In this kind of ecosystem, these variables, and especially salinity, can play a significant role not only in the partitioning of the contaminants but in the bioavailability and thus in the toxic effects associated with the chemicals. Recently, Riba et al. (14,15) reported the high influence of these variables in the bioavailability of metals from a mining spill that occurred in Spain in April 1998 that affected the Guadalquivir estuary. Their results showed that the same sample tested at different salinity or pH values using a truly estuarine species (*Ruditapes phillipinarum*) showed significantly higher

Table 1. Summarized List of Most Common Uncertainties Associated with Biological Testing of Dredged Material. The Uncertainty Description and the Palliative Action to Minimize the Effects on the Final Results Are Included

Topic	Uncertainty	Palliative-Minimizing
Handling and sampling	How the sample is taken and handled prior to testing, and the conditions the sample is subjected to during testing, can affect the accuracy and precision of test results	Careful and standardized procedures conducted during sediment assessments are necessary to minimize the source of uncertainty in sediment assessment
Extrapolation and number of species in the battery	The battery of tests will be composed of a relatively small number of taxa in comparison with the number residing at the site of concern	Testing with multiple species that are closely associated with the sediment and have a demonstrated sensitivity to the contaminants of concern
Dynamism of the ecosystem	Provide only a snapshot of the processes affecting contaminant exposure and effects in sediments	Development of more significant tests including chronic tests (at least one cycle of life of the organism)
Space scale is limited	The laboratory tests are conducted under controlled conditions of space and may overestimate the extent of exposure and effects given that the movement of the organism are restricted to the material being tested, and water movement (i.e., flow) is minimal	Conduct bioassays under field conditions increasing the space scale in which they are performed to avoid the lack in the extent of exposure and effects associated with the mobility of the organisms
Extrapolation to ecological scales	Biological tests most commonly measure effects on individual organisms, whereas the ecological impacts of concern occur at the level of populations and communities	Selection of the appropriate battery of ecologically relevant organisms
Experimental variables	Test organisms will respond positively and negatively to different experimental variables that are unrelated to the degree of contamination present in a sediment (grain size, food frequency, etc.)	Developing a thorough experimental understanding of the biology and ecology of the test organisms used in biological tests
Results analysis (statistical)	Over or underestimate of the hypothesis used in the different tests	Ensuring that adequate replication is prescribed to detect a desired magnitude of effect

toxicity at salinity values of 10 or lower or at pH values of 6 or lower compared with the toxicity tested at higher salinity or pH values. They relate the effect based on the higher mobility of some metals from sediments to water and also in the speciation of them that changed when the pH and salinity values changed. It informs one of the necessity to incorporate this kind of environmental effect in the final design of the weight of evidence approach, taking into account the potential influence of the pH and salinity values, especially under estuarine conditions.

Finally, we recommend to integrative assessment or the more recently cited weight of evidence approaches to avoid or at least minimize most of the deficiencies and uncertainties associated with laboratory and field bioassays. Furthermore, specifically applied WOEs for sediments or dredged material quality assessment is the use of the sediment quality triad incorporating some modifications and improvements to assess the sediment quality in both the disposal and the dredged site. One premise of the method is the design of a tier testing approach based on the definition of the Triad (Fig. 3). In this sense, the incorporation of new chemicals of concern specifically detected in the area to be dredged should be taken into account in the first tiers together with some biological tests for screening of the biological effects. The next steps will lead one to conduct laboratory sediment toxicity and bioaccumulation (if contaminants of concern are present) tests. In this sense, design specific and low-cost chronic tests, including sublethal endpoints such as histopathology, to avoid false negatives of toxicity such as those associated with specific contaminants. In the last tiers, the use of *in situ* surveys or those assays under field conditions using caged animals to establish the overall pollution status of the sediment to be dredged are included. The integration of these measurements will permit one to address the environmental quality of the systems and, furthermore, to establish the correct decision-making when selecting the beneficial use and/or the disposal options of the dredged material.

These recommended methods are not unique and a more extensive list of potential lines of evidence (7) that can be incorporated into the final tier testing using the weight of evidence approaches, such as the Triad, are included in Table 2. Each of these lines of evidence (LOEs) has their own advantages and limitations, but what is clear is that only one LOE will not be sufficient to address the sediment quality to determine the different management options for dredged material, and the use of several of these LOEs under the final design of a tier testing is recommended using a weight of evidence approach, such as that shown in Fig. 3.

Acknowledgments

This work has been partially funded by grants funded by the Spanish Ministry of Development (BOE 13-12-02) and of Education and Science (REN2002_01699/TECNO and VEM2003_−20563/INTER). The design of the acute toxicity tests for dredged material characterization was conducted under joint contract between CEDEX and the University of Cádiz (2001 and 2003).

Table 2. Description of Different Lines of Evidence, Including Their Advantages and Limitations When Used to Assess Sediment Quality and Dredged Material Characterization (7)

Line of Evidence	Advantage	Limitation
Sediment chemistry	Relatively simple and standardized. Widely used and utility proved. Cause information	Lack in the effect of information. Assumed feasibility to measure all existing chemicals
Toxicity tests (laboratory)	Relatively simple and standardized. Utility proved and widely used. Effect information	Lack in the determination of cause. Difficulty in field extrapolation. Natural stressors not assessed
Tissue chemistry	Determination of bioavailability and measure of exposure. If biomagnifications proved: useful in food chain models (human risk)	Lack in the effect measurements. Difficult to eliminate influence from essential and nonessential chemicals, food chain measurements, acclimation/adaptation
In situ alteration (field studies)	Measurements under real environment conditions	No cause is identified. No standard method. Difficult to discriminate natural effects
Biomarkers of exposure (EROD, methallotioneins, etc.)	Indicators of exposure under either field or laboratory conditions, including natural variables	No clear identification of cause, although relative information. Relative relationship to effects. Influenced by natural variables
Biomarkers of effects (histology, etc.)	Indicators of effects under either field or laboratory conditions, including natural variables	Lack in the cause identification. Nonspecific to contaminants and affected by natural variables
Benthic community structure	Standard methods, *in situ* conditions. Long-term measure of effects, including natural variables	Highly affected by natural variables. Not predictive. Confounding factors different from stressors. Absence of organisms in highly disturbed (natural or not) areas
Laboratory TIE	Partitioning of chemicals under laboratory conditions	Not widely proved. Insensitive. Not available for all chemicals
Field TIE	Partitioning of chemicals under field conditions	Highly complex design on the device used. Problems with interpretation of the results. Not possible for all the chemicals. Not standardized
Bethic fluxes of nutrients and/or contaminants	Functional capability of the ecosystems	High complexity in the devised use. No standard methods. Not widely used for sediment quality but for biogeochemical cycles
Caging animals	Toxicity and chemical residues under field conditions. Direct exposure to sediments	Difficult to clearly extrapolate to real conditions. Complex to devise and casualty inherent
Tissue quality values (TQVs)	Relationship between sublethal and chemical residues in organisms tissues	Not standardized. Relatively new. Large and temporal data base. Not easily available for all chemicals

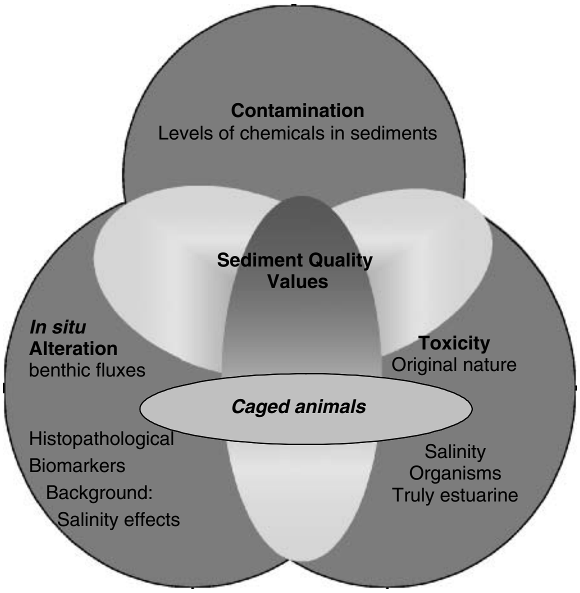

Figure 3. Synoptic representation of the weight of evidence approach recommended in the design of a convenient tier testing method either to assess sediment quality in littoral ecosystems or to manage dredged material. Each circle represents different areas including different lines of evidence (Table 2) to be integrated in the overall integrated approach. The overlapping area permits derivation of sediment quality values and tissue quality values for those overlapping areas, including histopathological and bioaccumulation studies under either field or laboratory conditions.

BIBLIOGRAPHY

1. Chapman, P.M. (2000). *Int. J. Environ. Pollut.* **13**: 1.
2. DelValls, T.A., Riba, I., and Martín-Díaz, M.L. (2004). *Using the triad to characterize dredged material in estuaries: lacks and improvements needed, 4th SedNet Workshop "Harmonization of Impact Assessment Tools for Sediment and Dredged Materials."* June 10–11, San Sebastian, Spain.
3. DelValls, T.A., Forja, J.M., and Gómez-Parra, A. (1998). *Environ. Toxicol. Chem.* **17**: 1073.
4. DelValls, T.A. and Chapman, P.M. (1998). *Cienc. Mar.* **24**: 336.
5. Riba, I. et al. (2004). *Chemosphere* (in press).
6. Riba, I., Forja, J.M., Gómez-Parra, A., and DelValls, T.A. (2004). *Environ. Pollut.* in press.
7. Burton, G.A. et al. (2002). *Hum. Ecol. Risk. Assess.* **8**: 1675.
8. Martín-Díaz, M.L. (2004). *Determinación de la calidad ambiental de sistemas litorales y de estuario de la Península Ibérica utilizando ensayos de campo y laboratorio*, Tesis Doctoral, Universidad de Cádiz.
9. Riba, I., Blasco, J., Jiménez-Tenorio, N., and DelValls, T.A. (2004). *Chemosphere* in press.
10. DelValls, T.A., Lubián, L.M., Forja, J.M., and Gómez-Parra, A. (1998). *Environ. Toxicol. Chem.* **16**: 2323.
11. DelValls, T.A. et al. (2005). *Track. Trend. Anal. Chem.* (submitted).
12. Riba, I., Casado-Martínez, M.C., Forja, J.M., and DelValls, T.A. (2004). *Environ. Toxicol. Chem.* **23**: 271.

13. Martín Díaz, M.L., Blasco, J., Sales, D., and Del Valls, T.A. (2003). *International Conference on Remediation of Contaminated Sediments*. Batelle Press.
14. Martín Díaz, M.L., Jiménez-Tenorio, N., Sales, D., and DelValls, T.A. (2005). *Mar. Pollut. Bull.* (submitted).
15. Riba, I., García-Luque, E., Blasco, J., and DelValls, T.A. (2003). *Chem. Speciation Bioavailab.* **15**: 101.

REMEDIATION AND BIOREMEDIATION OF SELENIUM-CONTAMINATED WATERS

Richard M. Higashi
Teresa A. Cassel
University of California
Davis, California

Joseph P. Skorupa
U.S. Fish and Wildlife Service

Teresa W.-M. Fan
University of Louisville
Louisville, Kentucky

GENERAL AND HISTORICAL BACKGROUND

The trace element selenium (atomic symbol Se) occurs naturally in soils, water, and biota, including food. It is nutritionally required due to various Se-bearing proteins that incorporate selenocysteine, now recognized as the 21st essential amino acid. In addition to being a required nutrient, Se in excess is toxic to biota (1). The first recorded cases of Se toxicity were penned by Marco Polo during his travels in western China, based on symptoms he observed in livestock (2). In the 1930s, the element Se was attributed to these symptoms, by then known variously as "alkali disease" and "blind staggers," which afflicted livestock in the western United States (3). However, in certain cases, such symptoms could have been due to other factors (4).

For Se, the margin between nutritional requirement and toxicity is unusually narrow and depends on the individual species and circumstance (3). This fact creates a quandary when setting environmentally protective Se criteria. Health benefits of consuming vegetables high in Se are widely touted (5), as cases of human toxicity and crop damage from Se are rare (6). Livestock face both deficiency and toxicity (7), while toxicity is the primary concern with fish, birds, and in particular various aquatic-associated wildlife (8). The latter issue roared to the news headlines as the aquatic bird disaster at California's Kesterson Reservoir surfaced in the early 1980s (9)—a harbinger of widespread Se problems such as fish mortality and deformities in Belews Lake (10) and numerous other cases internationally (8).

A closer look at the affected biota and circumstances surrounding Se toxicity reveals *water* as the transmitting medium of highest concern (8). Natural water processes (precipitation) and human activities relocate the naturally occurring Se in soils, rocks, and groundwater into collected water bodies that harbor and attract wildlife (8), which might also be used for growing forage and watering livestock (6). Human mobilization of Se mainly consists of irrigation, mine drainage, and surface discharge of groundwater; the latter includes petroleum processing, agriculture, and urban uses. We focus here on the factors important in Se-contaminated water; for the reader interested in the physiology and biochemistry of Se wildlife toxicity, there is an excellent recent review on this complex topic by Spallholz and Hoffman (11).

UNDERSTANDING THE PROBLEM: FUNDAMENTALS OF SELENIUM BIOGEOCHEMISTRY

In order to solve a problem, it is essential to first understand the nature of the problem. In the case of Se toxicosis of wildlife, exposure to Se primarily occurs through the diet, not by direct exposure to water (12). Therefore, it is vital to understand the "biogeochemistry" of Se: how it moves from contaminated water and sediment to work its way up the foodweb. Figure 1 illustrates the various paths comprising the biogeochemistry of Se. This figure illustrates the first four fundamental facts about wildlife exposure to Se:

1. The paths of Se are numerous and interrelated.
2. Most of the paths eventually lead up the foodweb.
3. Se changes chemical form from an inorganic salt to various organic forms.
4. Some of the organic forms lead to volatilization, a natural path for Se to leave an aquatic system altogether.

A fifth important fact, not evident in Fig. 1, is that Se "bioaccumulates" as it heads up the foodweb, generally becoming more concentrated in tissues. This concept may be familiar to many readers, as this phenomenon is discussed widely in the general press and health advisories regarding foodweb bioaccumulation of mercury, pesticides, polychlorinated biphenyls (PCBs), and many other contaminants. Table 1 (13) is an example from the scientific literature that documents this process for Se. Note the very large variation in bioaccumulation factors (BCFs), illustrating the effects of site-specific biogeochemistry in different aquatic systems. A comprehensive study illustrating the complex pathways of bioaccumulation has recently been published (14).

SETTING THE REMEDIATION TARGET: THE AQUATIC LIFE CRITERION

For remediation to occur, there must first be measurable targets to achieve. The complexity of Se biogeochemistry makes setting such a target extremely difficult to accomplish or even define. An earlier United States Environmental Protection Agency (U.S. EPA) Water Quality Criterion for Se, which set a maximum of 5 μg/L total Se, provided a clear target (15).

13. Martín Díaz, M.L., Blasco, J., Sales, D., and Del Valls, T.A. (2003). *International Conference on Remediation of Contaminated Sediments*. Batelle Press.
14. Martín Díaz, M.L., Jiménez-Tenorio, N., Sales, D., and DelValls, T.A. (2005). *Mar. Pollut. Bull.* (submitted).
15. Riba, I., García-Luque, E., Blasco, J., and DelValls, T.A. (2003). *Chem. Speciation Bioavailab.* **15**: 101.

REMEDIATION AND BIOREMEDIATION OF SELENIUM-CONTAMINATED WATERS

Richard M. Higashi
Teresa A. Cassel
University of California
Davis, California

Joseph P. Skorupa
U.S. Fish and Wildlife Service

Teresa W.-M. Fan
University of Louisville
Louisville, Kentucky

GENERAL AND HISTORICAL BACKGROUND

The trace element selenium (atomic symbol Se) occurs naturally in soils, water, and biota, including food. It is nutritionally required due to various Se-bearing proteins that incorporate selenocysteine, now recognized as the 21st essential amino acid. In addition to being a required nutrient, Se in excess is toxic to biota (1). The first recorded cases of Se toxicity were penned by Marco Polo during his travels in western China, based on symptoms he observed in livestock (2). In the 1930s, the element Se was attributed to these symptoms, by then known variously as “alkali disease” and “blind staggers,” which afflicted livestock in the western United States (3). However, in certain cases, such symptoms could have been due to other factors (4).

For Se, the margin between nutritional requirement and toxicity is unusually narrow and depends on the individual species and circumstance (3). This fact creates a quandary when setting environmentally protective Se criteria. Health benefits of consuming vegetables high in Se are widely touted (5), as cases of human toxicity and crop damage from Se are rare (6). Livestock face both deficiency and toxicity (7), while toxicity is the primary concern with fish, birds, and in particular various aquatic-associated wildlife (8). The latter issue roared to the news headlines as the aquatic bird disaster at California’s Kesterson Reservoir surfaced in the early 1980s (9)—a harbinger of widespread Se problems such as fish mortality and deformities in Belews Lake (10) and numerous other cases internationally (8).

A closer look at the affected biota and circumstances surrounding Se toxicity reveals *water* as the transmitting medium of highest concern (8). Natural water processes (precipitation) and human activities relocate the naturally occurring Se in soils, rocks, and groundwater into collected water bodies that harbor and attract wildlife (8), which might also be used for growing forage and watering livestock (6). Human mobilization of Se mainly consists of irrigation, mine drainage, and surface discharge of groundwater; the latter includes petroleum processing, agriculture, and urban uses. We focus here on the factors important in Se-contaminated water; for the reader interested in the physiology and biochemistry of Se wildlife toxicity, there is an excellent recent review on this complex topic by Spallholz and Hoffman (11).

UNDERSTANDING THE PROBLEM: FUNDAMENTALS OF SELENIUM BIOGEOCHEMISTRY

In order to solve a problem, it is essential to first understand the nature of the problem. In the case of Se toxicosis of wildlife, exposure to Se primarily occurs through the diet, not by direct exposure to water (12). Therefore, it is vital to understand the “biogeochemistry” of Se: how it moves from contaminated water and sediment to work its way up the foodweb. Figure 1 illustrates the various paths comprising the biogeochemistry of Se. This figure illustrates the first four fundamental facts about wildlife exposure to Se:

1. The paths of Se are numerous and interrelated.
2. Most of the paths eventually lead up the foodweb.
3. Se changes chemical form from an inorganic salt to various organic forms.
4. Some of the organic forms lead to volatilization, a natural path for Se to leave an aquatic system altogether.

A fifth important fact, not evident in Fig. 1, is that Se “bioaccumulates” as it heads up the foodweb, generally becoming more concentrated in tissues. This concept may be familiar to many readers, as this phenomenon is discussed widely in the general press and health advisories regarding foodweb bioaccumulation of mercury, pesticides, polychlorinated biphenyls (PCBs), and many other contaminants. Table 1 (13) is an example from the scientific literature that documents this process for Se. Note the very large variation in bioaccumulation factors (BCFs), illustrating the effects of site-specific biogeochemistry in different aquatic systems. A comprehensive study illustrating the complex pathways of bioaccumulation has recently been published (14).

SETTING THE REMEDIATION TARGET: THE AQUATIC LIFE CRITERION

For remediation to occur, there must first be measurable targets to achieve. The complexity of Se biogeochemistry makes setting such a target extremely difficult to accomplish or even define. An earlier United States Environmental Protection Agency (U.S. EPA) Water Quality Criterion for Se, which set a maximum of 5 μg/L total Se, provided a clear target (15).

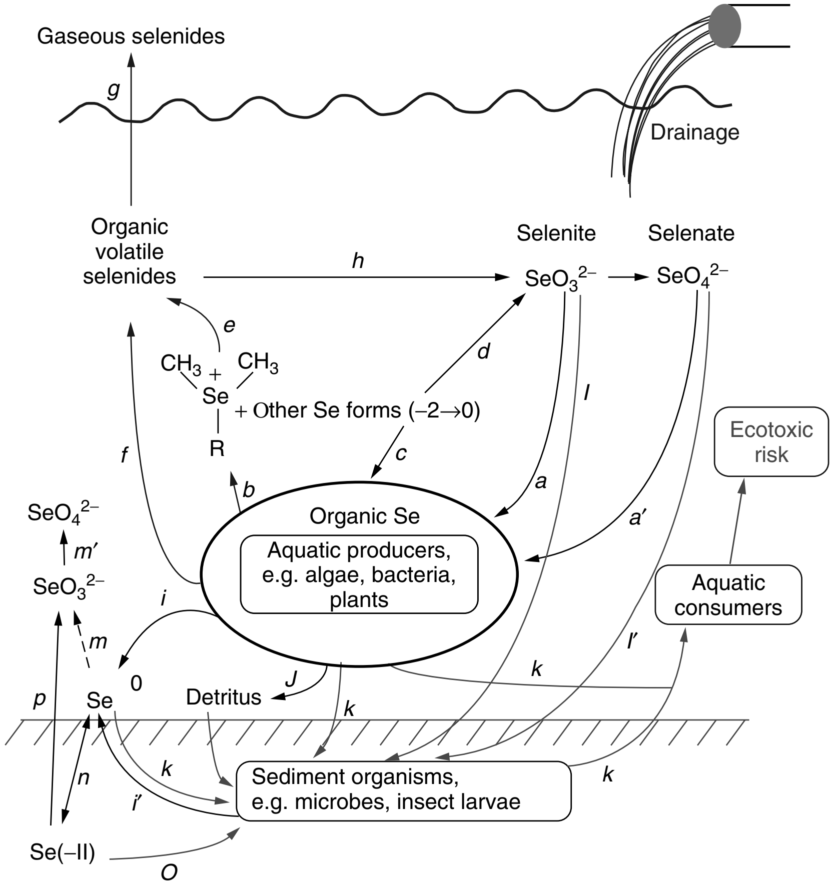

Figure 1. Biogeochemical cycling of Se in aquatic ecosystem. This scheme is modified from Reference 13. Arrows indicate processes that can lead to risk from foodweb accumulation of Se ("ecotoxic" risk). Other arrows trace the Se volatilization process by which Se can be lost from the aquatic system.
a, Uptake and transformation of Se oxyanions by aquatic primary and secondary producers; much of the biotransformation pathway is yet to be defined.
b, Release of selenonium and other organic Se metabolites by aquatic producers.
c, Uptake of organic Se compounds by aquatic producers.
d, Abiotic oxidation of organic Se compounds to Se oxyanions.
e, Release of alkylselenides from selenonium or other alkylated Se precursors through abiotic reaction.
f, Release of alkylselenides from selenonium or other alkylated Se precursors through aquatic producers.
g, Volatilization of alkyselenides into the atmosphere.
h, Oxidation of alkylselenides to Se oxyanions.
i, Formation of red amorphous Se element by aquatic and sediment producers.
j, Detrital formation from aquatic producers.
k, Se bioaccumulation into the foodchain with potential ecotoxic consequences; the toxic form(s) are yet to be defined.
l, Assimilation of waterborne selenium oxyanions into sediment biota.
m, Oxidation of sediment Se(0) to oxyanions.
n, Reduction of sediment Se(0) to Se(−II) or vice versa.
o, Assimilation of sediment Se(−II) into sediment biota.
p, Oxidation of sediment Se(−II) to selenite.

The setting of this temporary value spawned numerous attempts to apply the three traditional strategies of remediation, which are containment, removal, and treatment (16).

However, such a criterion for remediation remains very much a "moving target" due to the complexity of the biogeochemistry. In particular, there are currently attempts to develop a more scientifically sound U.S. EPA Aquatic Life Criterion for Se (17), for which there is now a draft (18). What follows is a very abbreviated description of the issues surrounding the setting of an Aquatic Life Criterion. For a more complete understanding, the reader should refer to the cited literature.

Several independent studies submitted to the U.S. EPA estimated that the 5-μg/L Se toxicity threshold was either too high or too low; interestingly, arguments for higher or lower thresholds appear to be closely associated with whether researchers were affiliated with the corporate-service (arguing that 5 μg/L is too low) or public-service (arguing that 5 μg/L is too high) scientific communities. A slightly different formulation of this observation has also been presented elsewhere (19). Additionally, the proposed California Toxics Rule of 5 μg/L Se in water (20) was judged by the U.S. Fish and Wildlife Service to be too high (21), jeopardizing 15 species protected by the Endangered Species Act (22), the majority of which were aquatic-dependent wildlife that do not actually live *in* the water. The lesson here is that, because of the biogeochemistry of Se, it is important to include consideration of the entire aquatic-based foodweb, not just aquatic organisms.

Recently, the U.S. EPA contracted the Great Lakes Environmental Center to derive chronic Se criteria on a fish-tissue basis, rather than the traditional water concentration basis. The results of this analysis were considered to have limited applicability because it produced a threshold (7.9 μg Se/g tissue) based on an LC_{20} (lethal concentration at which 20% mortality is expected), which is not adequately protective of the species of concern (22). Several other flaws were found in the study, which essentially compiled experimental data from 17 published studies (18). One of these (23) is qualitatively distinct because it incorporated a "winter-stress" design accounting for the increase in toxicity of dietary Se to birds, fish, and mammals under low temperatures. Adjustment of these data for the

Table 1. Excerpted Data from Reference 13 Illustrating Bioaccumulation of Se and the Great Variance in the Bioconcentration Factors of Se

	Salinity, ppth	Waterborne Se, μg/L	Body Burden, μg/g	BCF
Algae	76	13.5	11.8	873
Algae	14	7.9	14.8	1877
Algae	41	7.0	16.0	2281
Algae	136	13.1	9.7	745
Algae	80	13.3	22.6	1695
Algae	10	6.0	1.5	247
Algae	47	5.2	11.7	2273
Algae	66	5.1	9.9	1954
Algae	54	8.7	8.6	978
Average	58.2	8.8	11.8	1436
S.D.	38.1	3.5	5.8	738
Midge larvae	12	5.2	18.1	3489
Benthic composite	76	5.8	7.4	1288
Water column composite	76	5.8	12.3	2130
Corixid	39	7.5	8.9	1187
Corixid	90	8.8	6.8	778
Artemia + corixid	126	11.5	11.7	1012
Artemia	108	12.2	9.0	740
Corixid	108	12.2	10.1	834
Artemia	47	5.6	9.8	1749
Artemia	66	4.9	16.6	3372
Artemia	54	9.5	10.9	1145
Midge larvae	8.3	2.5	42.3	16886
Artemia	67	506.0	19.6	39
Average	67.5	46.0	14.1	2665
S.D.	35.8	138.3	9.4	4389

chronic-level protection needed (<5% mortality) would result in a limit of less than 5.8 μg/g on a whole-body fish-tissue basis.

Situations under which fish tissues would be contaminated to that level, however, may still be unacceptable to birds eating lower on the trophic level (e.g., invertebrate eaters) (24). Bioconcentration factors relating the transfer of Se from aquatic invertebrates to whole-body tissue of invertebrate-eating fish have been documented in the field to vary from 0.67 to 1.36, when whole-body fish tissue contains 5–10 μg/g Se. Applying those bioconcentration factors to the recently proposed fish-tissue criterion for Se of 7.9 μg/g (18) suggests that the corresponding Se content of aquatic invertebrates would range from 5.8 to 11.8 μg/g, which would result in an EC_{20}–EC_{85} range (effects concentrations, a range for which 20–85% of a population would show toxic effects) in mallard ducks (and other birds) (23) that also utilize the invertebrates as food (24).

STRATEGIES FOR REMEDIATION

From the preceding section, it is evident that there is no universally accepted water or tissue threshold concentrations for the protection of wildlife, and the issues are likely to remain unsettled for some time. Therefore, there is no fixed “target concentration” for remediation to achieve. This section on remediation strategies is written with this situation in mind.

Most of the past remediation attempts are mentioned only briefly here, as they were based on the obsolete 5-μg/L water criterion and do not incorporate the vital concepts of Se biogeochemistry. The following description attempts to discuss Se remediation efforts on the basis of the three strategies of remediation, which are containment, removal, and treatment (16). However, as the reader will see for Se, sometimes none of these three textbook categories are appropriate.

Containment

The first strategy of remediation, containment, remains difficult to achieve in many cases, as much Se contamination occurs in open lentic systems such as reservoirs, in lotic systems such as streams and rivers, and in receiving waters such as deltas and estuaries. To this end, the U.S. EPA has embarked on regulation based on the total maximum daily load (TMDL) for lotic and receiving waters, under Section 303(d) of the Clean Water Act (25). More recently, the best waterfowl protective estimates (based on empirical data from multiple environments) argue for reduction of the old 5-μg/L limits (8), forcing TMDL limits downward. This is an illustration of the “moving target” frustrating remediation efforts. Clearly, any remediation strategy must be compatible with changing regulations, but despite this clarion call, no traditional remediation strategies have met the challenge. For further description of TMDLs in general, the reader is referred to Reference 25.

The TMDL for Se basically collapses the complexity of Se biogeochemistry into a single box and a single total Se value, which may be protective of a given receiving water. However, this approach does not engender remediation solutions, since it does not account for Se biogeochemistry, provides little guidance for lotic system remediation, and provides no guidance at all for remediation of lentic systems.

In the United States, the lotic Se containment strategy is already implemented under TMDL regulations in California and elsewhere (25). For example, selected agricultural drainage from California's San Joaquin Valley is metered into the San Joaquin River, the allowed volume of which is based on quick turnaround time of Se analysis of the discharged water. For descriptions and issues surrounding the implementation of Se TMDL, the reader is referred further to Reference 25.

Simple containment of lotic Se into terminal water evaporation basins can avoid restrictive TMDL regulations since there is no discharge of water into lotic systems, and this strategy therefore has been suggested as worthy of further investigation (26). However, those authors also recognized that this simply obligates the discharger to use another remediation strategy, since the real problem is foodweb accumulation from either lotic or lentic systems. Therefore, while such total containment might "remediate" the downstream lotic systems, there is still a need to remediate the contaminated lentic system that has been created in the process.

Removal

The strategy of removal of Se is highly vulnerable to changing threshold regulations, as discussed earlier. Moreover, simple removal of Se is not necessarily "remediation" because the focus is put on waterborne concentrations of Se. In spite of this, several removal strategies have been tested in the past and tests of several more are currently underway. For the few that are in the peer-reviewed published literature, the reader is referred to a recent compilation (27). Removal of Se has been categorically difficult because of its typically low starting concentrations in the parts-per-billion range, its chemical similarity to sulfur (which can be present at more than a millionfold higher concentrations), the very high volumes of water for agricultural drainage, and especially because of the ever-lowering threshold target, as discussed previously.

Treatment

The third textbook strategy, treatment of Se, has much overlap with removal approaches, and again the focus here has been on reducing waterborne concentrations of Se, which does not equate to remediation (28).

An instructive example of a combined removal–treatment strategy is the algal–bacterial process (29) using the dissimilatory reduction capability of bacteria to reduce selenate into elemental Se, which has been developed to the point of large-scale trials. Although it can remove approximately 80% of the total Se from agricultural drainage water, the microbial and algal action on Se increased the levels of the more bioavailable selenite and organic Se, resulting in 2–4 times greater Se concentrations in the test invertebrates (30).

The lesson here is that waterborne Se is an inappropriate focus for remediation, one of the key reasons for drafting a new U.S. EPA Aquatic Life Criterion for Se based on tissue concentrations (18). The algal–bacterial process is also an example of how TMDL regulations, while written to be protective of wildlife in receiving waters, can inadvertently misguide major remediation efforts because it is based on waterborne Se.

Management for Mitigation

For Se remediation, there is a fourth strategy, that of management. This strategy is embodied by, for example, a California Waste Discharge Requirement under the Porter-Cologne Water Quality Act, which specified "a program of management actions to reduce, avoid, and mitigate for adverse environmental impacts to wildlife" (26). Management with this goal is a bona fide remediation strategy.

Along this vein, there has been much effort, especially in the western United States, to manage impacts on wildlife using a watershed-based approach, some of which are briefly outlined in Reference 6. In that paper, Engberg and co-workers have stated that "of the five management options presented, mitigation (option 5) may be the only long-term solution to managing selenium." A key tool in such mitigation is the construction of "alternative" and "compensation" habitats to build wetland habitats with clean water for impacted species (6,26). Note that this remediation strategy does not draw upon any of the textbook "remediation" strategies of containment, removal, or treatment.

The mitigation approach also has its downsides. In the western United States, water supply issues can complicate the Se remediation in water drainage. For example, the cost of land and water is a major obstacle to building and maintaining compensation and alternative habitats. In general, the management approach may require extraordinary effort: it has been stated that "managers must translate the complexities of selenium chemistry and biochemistry into cogent management and regulatory approaches, all the while understanding and blending financial, economic, and social constraints into the solutions" (6). Furthermore, since Se is a semi metal (or metalloid) with a wide variety of forms in many types of samples, it is exceptionally challenging to analyze.

Natural Bioremediation of Se

As mentioned above, the containment of waterborne Se into terminal water basins eliminates downstream impacts but creates a contaminated lentic system. Use of the mitigation approach is probably needed (26), but combinations of water treatment and management of the terminal basins have also been investigated. Only a few of these take advantage of the natural biogeochemistry to focus on removal of Se from the natural foodweb accumulation. Such an approach differs from the above strategies, which have an underlying assumption that the contaminated system presents only a problem that must be "fixed" somehow (containment, removal, treatment) or bypassed (management–mitigation). Rarely is consideration given that the aquatic system may be mostly functioning "right," and only a few particular processes that are awry need to be addressed.

"Biovolatilization" of Se is one such approach, since it takes advantage of the natural biogeochemistry to remove Se. The problem with biovolatilization of any type, as Fig. 1 outlines, is that the process also draws Se into the biota, and consequently up the foodweb. This tendency has been particularly troublesome in attempts to utilize aquatic vascular plants to volatilize Se. For example, vascular plants volatilize a relatively small amount of Se while sequestering it in bioavailable foodweb materials such as the shoots and roots. Although the shoots could be harvested and disposed, the Se is mostly contained in the below-ground portions of the plants (31), which are not practical for harvesting.

Fan and Higashi (28,32) have described natural algal Se volatilization as part of an alternative remediation process in terminal basins. The overall remediation process, outlined in Fig. 2, combines volatilization of Se with interrupting the foodweb accumulation of Se (33). In this strategy, photosynthetic algae function to volatilize Se while serving as food to macroinvertebrates (brine shrimp). The brine shrimp graze the algae, ideally preventing algal accumulation and participation in the detrital cycle. In turn, the brine shrimp are harvested as a marketable product, thereby intercepting the foodweb-accumulated Se before it can impact fish and birds. Both volatilization and harvesting of brine shrimp result in a net removal of Se from the aquatic system. The resulting blockages to the foodweb accumulation of Se are shown by the X's in Fig. 2.

A full-scale evaporation basin system has been monitored for over three years in this regard (33). The terminal evaporation basins are brine, which allows the flourishing of brine shrimp that has commercial value—in fact, at California's Tulare Lake Drainage District, there has been successful commercial harvesting of brine shrimp for over five years (34). The efficacy of the approach is clear: (1) the algae that volatilize Se and feed the brine shrimp grow naturally in these basins; (2) the brine shrimp also grow naturally in these basins to a high density; (3) much of the scheme utilizes water management that is familiar to drainage operators; and (4) costs of encouraging and managing brine shrimp growth is offset by marketing harvested materials.

Comparisons with analogous systems where brine shrimp harvesting is not implemented has demonstrated that, under active brine shrimp harvesting (33):

- Waterborne Se concentrations are not increased.
- Se volatilization—a removal process of Se—appears to be enhanced.
- Concentrations of Se in algae and macroinvertebrates—a risk indicator for fish and bird toxicity—are not increased.
- Benthic macroinvertebrate biomass—a key foodweb accumulation indicator—is greatly decreased.
- Total Se as well as organic Se deposited to the sediment (available for biogeochemical refluxing as in Fig. 2) is decreased >90%.

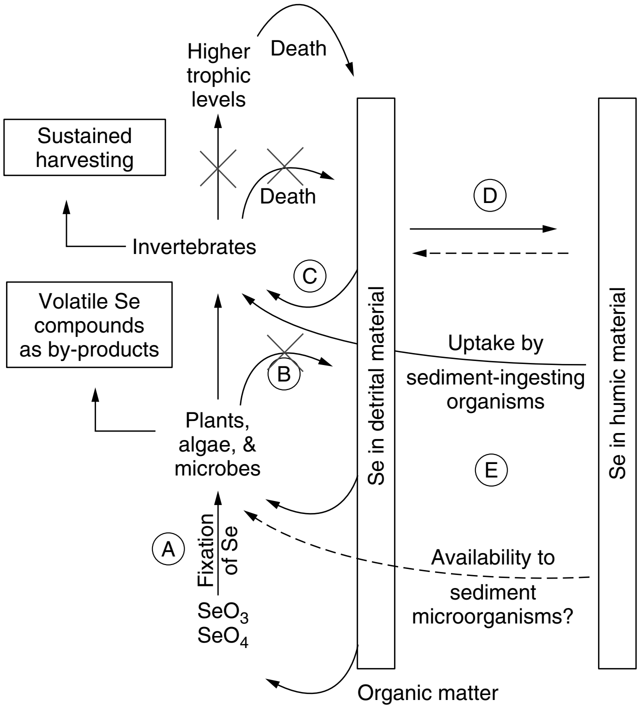

Figure 2. Bioremediation in drainage basins by reducing Se ecotoxic risk through invertebrate harvest and Se volatilization. In this "biogeochemical reflux" scheme, the drainage inorganic Se forms are initially biologically fixed by aquatic algae and microbes Ⓐ. The fixed Se does not directly head up the foodchain in the water column, as is often portrayed. Instead, a major fraction enters into organic matter, taking a detour through detritus (recently dead organic matter) and sediment Ⓑ, then reentering the foodchain at several trophic levels Ⓒ. Over longer periods, part of the detrital material is converted to recalcitrant humic material Ⓓ, locking up the Se until sediment-ingesting organisms reintroduce them to the foodchain Ⓔ. Through sustained harvesting (upper left box) of water-column invertebrates that consume algae and microbes, the bioavailable Se is removed from water, plus detrital formation resulting from the death of water-column organisms is also blocked. Both types of blockages are shown by the three X's. In turn, this would help minimize the sediment–detritus foodchain pathway for Se. In the meanwhile, additional Se can be removed by manipulating the algae/microbe community for optimal Se volatilization (lower left box).

Research is continuing to determine how this strategy can be applied to other systems and to achieve long-term sustainability. At present, the combined strategy of impounding Se-contaminated water in basins, applying the algal volatilization–foodweb interruption process, and constructing compensation/alternative habitats illustrates the advantages of an integrated management approach. In fact, such watershed-scale management appears to be required in order to remediate a contaminant with complex biogeochemistry such as Se.

BIBLIOGRAPHY

1. Cooper, W.C. and Glover, J.R. (1974). The toxicity of selenium and its compounds. In: *Selenium*. R. Zingaro and W.C. Cooper (Eds.). Van Nostrand Reinhold, New York.

2. Marsden, X. (1962). *The Travels of Marco Polo (1271–1285)*. The Heritage Press, Norwich, CT.
3. Combs, G.F. Jr., Levander, O.A., Spallholz, J.E., and Oldfield, J.E. (Eds.). (1987). *Selenium in Biology and Medicine*. Van Nostrand Reinhold, New York.
4. O'Toole, D. and Raisbeck, M.F. (1998). Magic numbers, elusive lesions: comparative pathology and toxicology of Selenosis in waterfowl and mammalian species. In: *Environmental Chemistry of Selenium*. W.T. Frankenberger and R.A. Engberg (Eds.). Marcel Dekker, New York, pp. 355–395.
5. Keck, A.-S. and Finley, J.W. (2004). Cruciferous vegetables: cancer protective mechanisms of glucosinolate hydrolysis products and selenium. *Integrative Cancer Therapies* **3**(1): 5–12.
6. Engberg, R.A., Westcot, D.W., Delamore, M., and Holz, D.D. (1998). Federal and state perspectives on regulation and remediation of irrigation-induced selenium problems. In: *Environmental Chemistry of Selenium*. W.T. Frankenberger, R.A. Engberg (Eds.). Marcel Dekker, New York, pp. 1–25.
7. Maas, J. (1998). Selenium metabolism in grazing ruminants: deficiency, supplementation, and environmental implications. In: *Environmental Chemistry of Selenium*. W.T. Frankenberger and R.A. Engberg (Eds.). Marcel Dekker, New York, pp. 113–128.
8. Skorupa, J.P. (1998). Selenium poisoning of fish and wildlife in nature: lessons from twelve real-world examples. In: *Environmental Chemistry of Selenium*. W.T. Frankenberger and R.A. Engberg (Eds.). Marcel Dekker, New York, pp. 315–354.
9. Ohlendorf, H.M. (2002). The birds of Kesterson Reservoir: a historical perspective. *Aquat. Toxicol.* **57**: 1–10.
10. Lemly, A.D. (2002). Symptoms and implications of selenium toxicity in fish: the Belews Lake case example. *Aquat. Toxicol.* **57**: 39–49.
11. Spallholz, J.E. and Hoffman, D.J. (2002). Selenium toxicity: cause and effects in aquatic birds. *Aquat. Toxicol.* **57**: 27–37.
12. Presser, T.S. and Ohlendorf, H.M. (1987). Biogeochemical cycling of selenium in the San Joaquin Valley, California, USA. *Environ. Manage.* **11**: 805–821.
13. Fan, T.W.-M., Teh, S.J., Hinton, D.E., and Higashi, R.M. (2002). Selenium biotransformations into proteinaceous forms by foodweb organisms of selenium-laden drainage waters in California. *Aquat. Toxicol.* **57**: 65–84.
14. Stewart, A.R., Luoma, S.N., Schlekat, C.E., Doblin, M.A., and Hieb, K.A. (2004). Food web pathway determines how selenium affects aquatic ecosystems: a San Francisco Bay case study. *Environ. Sci. Technol.* **38**: 4519–4526.
15. U.S. Environmental Protection Agency (1987). *Ambient Water Quality Criteria for Selenium*. EPA, Washington, DC, EPA440/5-87/083.
16. Brusseau, M.L. and Miller, R.M. (1996). Remediation. In: *Pollution Science*. I.L. Pepper, C.P. Gerba, and M.L. Brusseau (Eds.). Academic Press, San Diego, Chap. 11, p. 154.
17. Sappington, K.G. (2002). Development of aquatic life criteria for selenium: a regulatory perspective on critical issue and research needs. *Aquat. Toxicol.* **57**: 101–113.
18. Aquatic Life Water Quality Criteria for Selenium (2002). *Draft Document Prepared for the United States Environmental Protection Agency*, Washington, DC, by the Great Lakes Environmental Center, Traverse City, MI, p. 67.
19. Hamilton, S.J. (2004). Review of selenium toxicity in the aquatic food chain. *Sci. Total Environ.* **326**: 1–31.
20. EPA (1997). Water quality standards; establishment of numeric criteria for priority toxic pollutants for the State of California. *Fed. Reg.* **62**(150): 42159–42208.
21. *Final Biological Opinion on the Effects of the U.S. Environmental Protection Agency's "Final Rule for the Promulgation of Water Quality Standards: Establishment of Numeric Criteria for Priority Toxic Pollutants for the State of California"* (2000). United States Department of Interior: Fish and Wildlife Service and United States Department of Commerce: National Marine Fisheries Service, p. 323
22. United States Code [U.S.C.] (2000). *The Endangered Species Act of 1973. 16 U.S.C.*, Sections 1531–1544.
23. Lemly, A.D. (1993). Metabolic stress during winter increases the toxicity of selenium to fish. *Aquat. Toxicol.* **27**: 133–158.
24. Ohlendorf, H.M. (2003). Ecotoxicology of selenium. In: *Handbook of Ecotoxicology*, 2nd Edn. D.J. Hoffman, B.A. Rattner, G.A. Burton Jr., and J. Cairns Jr. (Eds.). Lewis Publishers, Boca Raton, FL, pp. 465–500.
25. The TMDL regulations and literature are very extensive—for more information, access the website http://www.epa.gov/owow/tmdl/.
26. Tanji, K. et al. (2003). Evaporation ponds as a drainwater disposal management option. *Irrigation Drainage Syst.* **16**: 279–295.
27. Frankenberger, W.T. and Engberg, R.A. (Eds.). (1998). *Environmental Chemistry of Selenium*. Marcel Dekker, New York.
28. Fan, T.W.-M. and Higashi, R.M. (1998). Biochemical fate of selenium in microphytes: natural bioremediation by volatilization and sedimentation in aquatic environments. In: *Environmental Chemistry of Selenium*. W.T. Frankenberger and R.A. Engberg (Eds.). Marcel Dekker, New York, pp. 545–563.
29. Gerhardt, M.B. et al. (1991). Removal of selenium using a novel algal-bacterial process. *Res. J. Water Pollut. Control Fed.* **63**: 799–805.
30. Amweg, E.L., Stuart, D.L., and Weston, D.P. (2003). Comparative bioavailability of selenium to aquatic organisms after biological treatment of agricultural drainage water. *Aquat. Toxicol.* **63**: 13–25.
31. Terry, N. and Zayed, A. (1998). Phytoremediation of selenium. In: *Environmental Chemistry of Selenium*. W.T. Frankenberger and R.A. Engberg (Eds.). Marcel Dekker, New York, pp. 633–655.
32. Fan, T.W.-M. and Higashi, R.M. (2000). Microphyte-mediated selenium biogeochemistry and its role in *in situ* selenium bioremediation. In: *Phytoremediation of Contaminated Soil and Water*. N. Terry and G.S. Banuelos (Eds.). CRC Press, Boca Raton, FL, pp. 283–302.
33. Higashi, R.M., Rejmankova, E.J., Gao, S., and Fan, T.W.-M. (2003). Mitigating selenium ecotoxic risk by combining foodchain breakage with natural remediation. *University of California Salinity/Drainage Program Annual Report*.
34. Rofen, R. (2001). Agricultural evaporation ponds—source of marketable brine shrimp. *Curr. Newlett. UC Center for Water Resources* **2**(2): 9–10.

SHELLFISH GROWING WATER CLASSIFICATION

Linda S. Andrews
Mississippi State University
Biloxi, Mississippi

Seafood-related illnesses in the United States are primarily associated with the consumption of bivalve molluscan shellfish (1,2). As these animals are water filter feeders and are often consumed raw, the water that they are grown in affects their safety and quality.

Bivalve molluscs feed by filtering microscopically sized particles out of the water. Via their inhalant siphon, they pump water through their gills. The gills sort out particles of the correct size and "feel" and direct them to their gut (3). The large volume of water that the bivalves pump through their siphons causes concern among public health officials. Consequently, it is imperative that water authorities, usually shellfish producer states, monitor the quality of water in and around the shellfish beds or reefs. For commercial markets, harvesting is only permitted from approved growing areas. The microbial acceptability of growing areas is classified on the basis of a sanitary survey of the shoreline to detect potential pollution sources and bacteriological analysis of water samples taken from the area (4).

Shellfish sanitation programs originated at the national level in the United States in 1925 after a series of typhoid epidemics spread by the consumption of raw shellfish threatened the collapse of the oyster industry (3). Today, through the National Shellfish Sanitation Program, NSSP (5), State and federal public health regulators have established guidelines for water quality in shellfish growing areas to decrease the risk of illness associated with these seafood products.

Sanitation surveys are normally performed by the state regulatory authority and include a written report of a shoreline survey, bacterial quality of the water, and an evaluation of the effect of any meteorological, hydrodynamic, and geographic characteristics of the growing area. Each surveying state maintains records of the surveys. For example, for Mississippi, survey results are filed with the Mississippi Department of Marine Resources (http://www.mississippiwebsite.com/deptmarinerec.htm). An analysis of the data from this survey determines the appropriate growing area classification. During shoreline surveys, the state authority identifies and evaluates all actual and potential sources of pollution to the growing area. If a pollution source exists, the distance from pollution to growing area is determined with an assessment of the potential impact. The effectiveness of sewage treatment and other wastewater treatment systems is evaluated for removal of the microbial and/or chemical contaminants.

Comprehensive sanitation surveys for each growing area must be performed at least once every 12 years with triennial reevaluation based on water quality analysis and potential for any new sources of pollution. If this triennial evaluation determines that conditions have changed from the 12-year survey, the area in question may then be reclassified. On an annual basis, the sanitary survey shall be updated to reflect changes in the conditions of the growing area based on "drive through" surveys, observations made during routine sample collection, or information from other sources.

In general, each growing area is classified as approved, conditionally approved, restricted, conditionally restricted, or prohibited based on the survey results (see Table 1). Status of a growing area is separate from its classification and may be either considered open or closed for harvesting of shellstock. All correctly classified growing areas are normally open for harvesting unless they are classified as prohibited. Closures may occur temporarily because of adverse weather conditions, the presence of biotoxins in concentrations of public health significance, or when the state authority fails to complete the written sanitary survey or triennial review evaluation report. A growing area may be placed into a remote status if the sanitation survey determines that the area has no human habitation and is not impacted by human pollution sources.

Table 1. Shellfish Harvest Area Classification Categories (5)

Classification	
Approved	Growing areas are classified as approved when the sanitation survey finds the area safe for the direct marketing of shellfish or is not subject to human or animal fecal pollution.
Conditional	Growing areas are classified as conditional when the area in question is in the open status for a reasonable period of time and when pollution factors are predictable. There may be direct potential for distribution of pollutants based on unusual conditions or specific times of the year when bacterial numbers are increased by heavy water runoff that affects wastewater treatment plant function.
Restricted	Growing areas are classified as restricted when the sanitary survey indicates a limited degree of pollution and when levels of fecal pollution, human pathogens, or poisonous or deleterious pollutants are at such levels that shellstock can be made safe through either relaying or depuration.
Prohibited	Growing areas are classified as prohibited when no current sanitary survey exists or when the survey determines that the growing area is adjacent to a sewage treatment plant outfall or other point source with public health significance or when the water is polluted because of previous or current sources of contamination.

The NSSP allows for a growing area to be classified using either a total or fecal coliform standard. Two sample collection strategies are used: adverse pollution condition and systematic random sampling. Each state authority may choose one of these sampling plans or use both depending on the location of the shellfish bed. Each state authority determines the number and location of sampling stations based on potential sources of pollution contamination as determined by the shoreline survey. Except for prohibited areas, the original or new classification of a shellfish growing area that has the potential to be impacted by a pollution source requires a minimum of 30 samples collected under various environmental conditions. For an area not impacted by a pollution source, 15 samples are required for initial classification. When using fecal coliforms (FC) as the indicator bacteria, the water quality must meet the following standards: FC median or geometric mean most-probable-number (MPN) may not exceed 14 per 100 mL, with not more than 10% of samples exceeding 43 MPN per 100 mL for 5 tube decimal dilution or 49 MPN per 100 mL for 3 tube decimal dilution. When classifying point source growing sites, the bacterial quality of every station in the growing area must meet the fecal coliform standards as described above. Sample stations must be located adjacent to actual or potential sources of pollution. Refer to the NSSP Model Ordinance Chapter IV for details on sampling to achieve statistically reliable results (5).

Separate standards are used for shellstock that will be processed by depuration posterior to harvesting. For the restricted classification of growing areas that are affected by point sources or nonpoint sources, FC median or geometric mean MPN of the water samples shall not exceed 88 per 100 mL and not more than 10% of the samples shall exceed an MPN of 260 per 100 mL for a 5 tube decimal dilution test or 300 MPN per 100 mL for a 3 tube decimal dilution test.

Numerous factors are involved in interpreting FC data. Fecal coliforms are common as indicators of human fecal pollution and potential human fecal pathogens. However, pathogenic microorganisms typically do not occur in direct proportion to the number of FC present (3). This is especially true with viruses that move through tidal water at a much faster rate than heavier coliforms. Also, FC may die off at a rate different from other pathogenic organisms, thus giving a false sense of security that the shellfish are safe for human consumption. Another problem with using FC as an indicator of potential human pathogens is that FC may originate from sources other than humans. In rural areas, wild animals are likely to be the primary source of FC found in runoff. Major human pathogens of concern in shellfish waters usually originate on land, but exceptions include sewage discharges from boats and sewage treatment facilities. Runoff from land varies with rainfall, soil type, and degree of saturation. Consequently, one or even a few samples collected periodically may indicate very little about the pollution potential. Numerous samples collected under all weather conditions (heavy rain and high tides) are the only way to properly evaluate and classify a particular growing site.

The restrictive rules for classification of shellfish growing waters have been established to provide the safest shellfish possible. All shellfish for human consumption are marked with a harvesters tag that indicates that they were harvested only from approved harvest areas. Today, illness outbreaks caused by the consumption of human pathogens that are present in raw molluscan shellfish are rare because of these rules for classification.

BIBLIOGRAPHY

1. Anonyomous. (1988). Seafood Safety: Seriousness of Problems and Efforts to Protect Consumers. GAP/RCED-88-135, United States General Accounting Office, Washington, DC.
2. CDC. (1999). *Morbidity and Mortality Reports*. Center for Disease Control, Atlanta, GA. Available: www.cdc.gov.
3. Croonenberghs, R.E. (2000). Contamination in shellfish-growing areas. Section VI, Seafood safety. In: *Marine and Freshwater Products Handbook*, pp. 665–696.
4. Cook, D.W. (1991). Microbiology of bivalve molluscan shellfish. In: *Microbiology of Marine Food Products*, pp. 19–39.
5. NSSP. (2002). Model Ordinance, Guide for the Control of Molluscan Shellfish, Chap. IV. U.S. Department of Health and Human Services, Public Health Service, Food and Drug Administration. Available: http://www.ISSC.org.

SORPTIVE FILTRATION

K.A. Matis
N.K. Lazaridis
Aristotle University
Thessaloniki, Greece

INTRODUCTION

Although filtration is one of the principal unit operations in the treatment of potable water, the filtration of effluents is less practiced; usual examples are supplemental removals of suspended solids from wastewaters of biological and chemical treatment processes and also the removal of chemically precipitated phosphorus (1). The available references dealing with filtration are quite voluminous—see, among others, the articles on filter presses, deep bed, cartridge, batch, variable volume, and continuous filters (2). Hence, the information provided in this section serves only as an introduction to the subject; for additional details, the literature should be consulted. In this article, no mention will be made of chromatographic applications (3).

The seepage of rainfall and runoff into porous solids and rocks and the storage and movement of groundwaters in open-textured geologic formations are important elements in the resource and quality management of water and wastewater. Although fine-textured granular materials remove pollutants, the water drawn from them is acceptable only when natural filtration together with a time lapse between pollution and use bar the transport of pollutants to springs, wells, and infiltration structures (4).

SORPTION

A typical recent example of sorption is the removal of copper by suspended particulate matter from river water (5). Iron, silicon, and aluminum oxides were used as adsorbents, serving as models of naturally occurring suspended particles. On the other hand, sorption upstream of filtration constitutes a solution; the combined process has been also termed adsorptive filtration (6). In this module (offered by Separation Technologies), fundamental aspects of the process are illustratively given, such as inertial impaction and diffusional interception; however, the latter is effective only in gas filtration. Sorption, in general, is defined as a surface process irrespective of mechanism, adsorption or precipitation (7).

Batch and column sorption studies of zinc, cadmium, and chromium were conducted on calcite to determine its retention capacity for these elements common in industrial effluents and to explore its behavior as a purification filter in a continuous flow system (8). Fibrous carbon materials were examined for heavy oil sorption (9); oil spill accidents have caused serious problems in the environment, including disasters in living systems. Elsewhere (10), biologically active carbon (followed by ultrafiltration) was used for simultaneous sorption and biodegradation of organic constituents. Naturally occurring diatomaceous earth was tested as a potential sorbent for lead ions; the intrinsic exchange properties were further improved by modification with manganese oxides (11). Today, effective low-cost adsorbents for toxic metals are available; for instance, the case study to remove (by Metsorb™ adsorptive filtration media) depleted uranium from contaminated water at a U.S. Army testing site (12).

In recent years, contamination of ground and surface waters with heavy metals has become a major concern. The knowledge of the oxidation state of pollutant ions is often a prerequisite for the application of efficient treatment methods, as in the case of arsenic. The inhibition of conventional metal precipitation due to the presence of chelating or complexing compounds, such as acetate, citrate, and tartrate ions, ammonia, and EDTA, which may be present in most real wastewater streams, is another problem for examination. Thermodynamic equilibrium diagrams and software packages (such as Mineql+) have been employed to construct aqueous speciation diagrams for the metals under investigation and, then, to interpret the removal mechanism involved. Sorption isotherm equations for equilibrium uptake, such as that of Freundlich and Langmuir, have been often used to fit the experimental data, and the activation energy of the process was calculated (13).

Figure 1a presents some typical kinetic results for hexavalent chromium ions in batch sorptive removal under different conditions by (uncalcined) hydrotalcite as the appropriate sorbent, noting that desorption, particularly in successive stages, has received much less attention in the literature. The data were effectively fitted by a second-order kinetics equation.

Another suitable (for metal cations) inorganic sorbent material is the group of natural or synthetic zeolites (Fig. 1b); the nonselectivity of the process is also apparent, sorbing zinc or calcium. Stress was given to electrokinetic measurements (expressed as zeta-potential) of the system for its surface charge under the applied conditions to predict its behavior (14). The solution pH is important and influences metal speciation. So, for example, at pH lower than 6.8, corresponding to the pristine point of zero charge of α-FeOOH, the surface of these particles is positively charged. Therefore, in the pH range examined, the adsorption of anionic metal species was more pronounced (15).

Fixed-Bed Sorption

Ferric hydroxides, such as goethite, are another good example extensively studied; fixed-bed operation is the appropriate configuration for large-scale applications such as wastewater treatment (see Fig. 2a). The adsorbent material was granulated by crystallization through controlled freezing. Column adsorbers, due to pressure drop, require a suitable shape and size of the bonding material used. The experimental results are presented, expressed typically by the breakthrough curve concept, service time versus breakthrough (column outlet concentration related to the initial metal concentration—$C_{\text{outlet}}/C_{\text{initial}}$), as a percentage. In the inplots, the isoremoval lines of treatment time versus the respective quantity of sorbent are also presented at various breakthrough points.

The major aim when sizing adsorptive columns is the ability to predict the service time until the column effluent exceeds a predefined solute (pollutant) concentration. The bed depth–service time model (abbreviated as BDST) relates the service time of a fixed-bed to the height of adsorbent in the bed, hence to its amount, because quantity is directly proportional to bed height. The measurement of sorbent quantity is more precise than the determination of the respective volume, especially for granules. Therefore, the sorbent quantity is preferably used instead of the bed height (15,16).

Surface complexation models were used to describe metal cation adsorption on adsorbent materials; models, such as constant capacitance, diffuse layer, and triple layer were applied, showing that they could simulate experimental results. The fundamental concepts on which all these models are based remain, more or less, the same (17).

The computer program PHREEQC (of the U.S. Geological Survey) permits a wide variety of aqueous geochemical calculations; surface-complexation modeling, applying mainly the diffuse double layer (DDL) model,

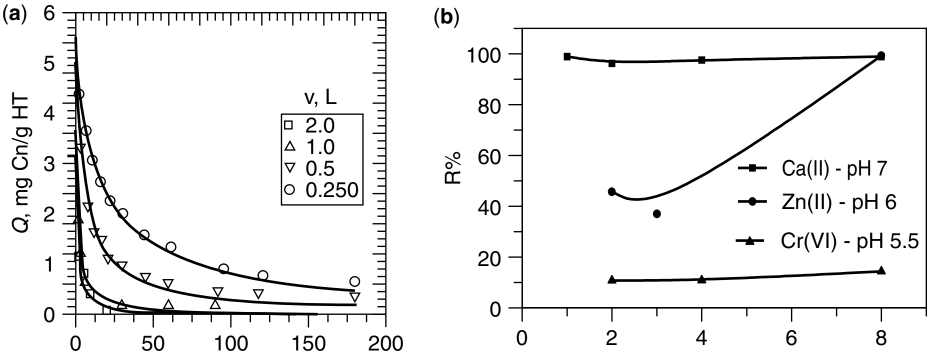

Figure 1. (**a**) Time variation of chromium(VI) loading on hydrotalcite (1 g/L) at various eluant volumes. Reprinted with permission from Reference 13; copyright (2004) American Chemical Society. (**b**) Removal of calcium, zinc, or chromate ions (50 mg/L) by zeolites as a function of their concentration, at different pH values. Reprinted with permission from Reference 14; copyright (2004) Elsevier.

is included in the program. The program uses a modification of the usually applied Newton–Raphson method iteratively to revise the values of the variables until a solution of the equation system is accepted within specified tolerances. A limitation of the program may be that uncertainties occur in determining the number of sorption sites, the surface area, the composition of sorbed species, and the appropriate surface-complexation constants. It was demonstrated that some of the surface complexation model (SCM) parameters might be incorrect due to an experimental artifact, though they give good simulations of data. This may be due to the success of least-squares fitting programs, such as FITEQL, in finding model parameters to describe experimental data sets.

The results of the model, according to surface-complexation reactions and constants (see Table 1), are presented in Fig. 2b, in comparison with the experimental data for arsenic(III) oxyanions; they show good agreement, although the diffusion effects (external, or liquid film, and intraparticle transport resistance) were not taken into consideration (18). Similar conclusions were obtained with As(V). The arsenic problem is of particular concern for small communities in rural areas around the world (a characteristic case is that of Bangladesh), where groundwater comprises the main drinking water source. Arsenite is favored under reducing (i.e., anaerobic) conditions. In this *Encyclopedia*, refer to another article on Sorption Kinetics in the chapter on Physics and Chemistry of Water.

Table 1. As(III) Species Distribution and Surface-Complexation Reactions and Their Reaction Constants[a]

Species Distribution	Reaction Constants, log *K*
$AsO_3^{3-} + 4H^+ = H_4AsO_3^+$	34.44
$AsO_3^{3-} + 3H^+ = H_3AsO_3$	34.74
$AsO_3^{3-} + 2H^+ = H_2AsO_3^-$	25.52
$AsO_3^{3-} + H^+ = HAsO_3^{2-}$	13.41
Surface-Complexation Reaction	
$Goeth\_OH + H_3AsO_3 = Goeth\_H_2AsO_3 + H_2O$	4.51

[a] Reference 18.

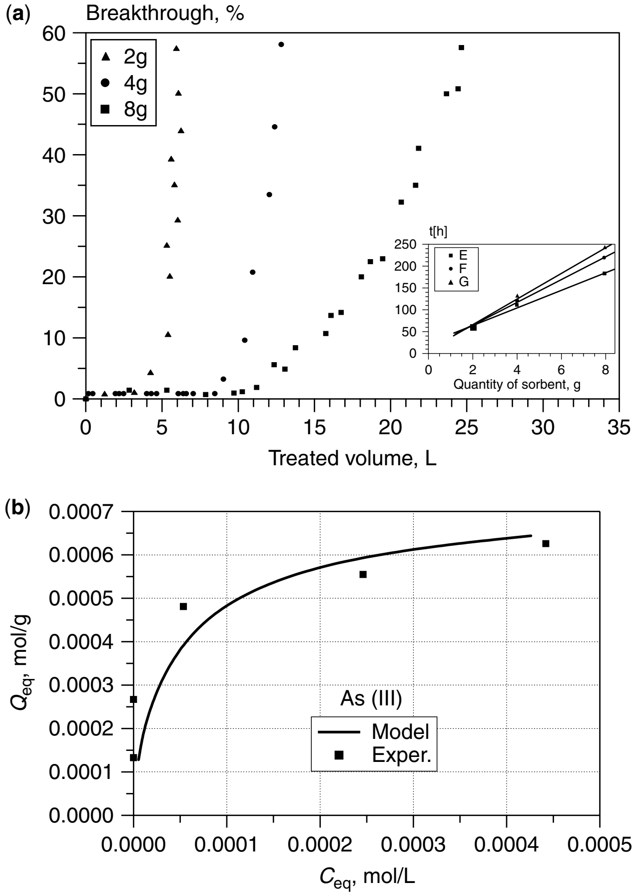

Figure 2. (**a**) Breakthrough curves for goethite (at different quantities) and chromates; as an inset, isoremoval lines for 20%, 35%, and 50% breakthrough. Reprinted with permission from Reference 15; copyright (2001) Elsevier. (**b**) Sorption of trivalent arsenic on goethite mineral: experimental vs. theoretical (DDL modeling) values. Reprinted with permission from Reference 18; copyright (1999) Kluwer.

For a number of ion exchangers of ultrafine particle size (for efficient use in packed column beds), a convenient binding polymer, such as modified polyacrylonitrile, has been used. The resulting sorbents were tested for separation and preconcentration of different contaminants, including radioactive wastes. The sorption characteristics of these composite materials were not affected by the binding polymer, whereas their physicochemical properties (hydrophilicity, mechanical strength, etc.) can be modified by the degree of cross-linking of the polymer, the use of suitable copolymers, or by changing the composition and temperature of polymerization (19). Certainly, an alternative to facilitate their solid/liquid separation downstream, particularly if fine adsorbents are to be dispersed, is to apply flocculation (20); note that many low-cost synthetic sorbents are produced as powders.

THE COMBINED PROCESS

Liquid-phase carbon adsorption is a full-scale technology in which groundwater is pumped through one or more vessels containing activated carbon to which dissolved organic contaminants adsorb (21). When the concentration of contaminants in the effluent from the bed exceeds a certain level, the carbon can be regenerated in place; removed and regenerated at an off-site facility; or removed and disposed of. Carbon used for explosive- or metal-contaminated groundwater probably cannot be regenerated and should be removed and properly disposed of. Adsorption by activated carbon has a long history of use in treating municipal, industrial, and hazardous wastes. Figure 3a illustrates a sorptive filtration system.

Another innovation has been proposed by introducing a two-stage process in a compact microfiltration (MF) hybrid cell (Fig. 3b). A large number of techniques have been used to limit membrane fouling; among them is air bubbling that also constitutes the transport medium in flotation, as applied in wastewater treatment; flotation is

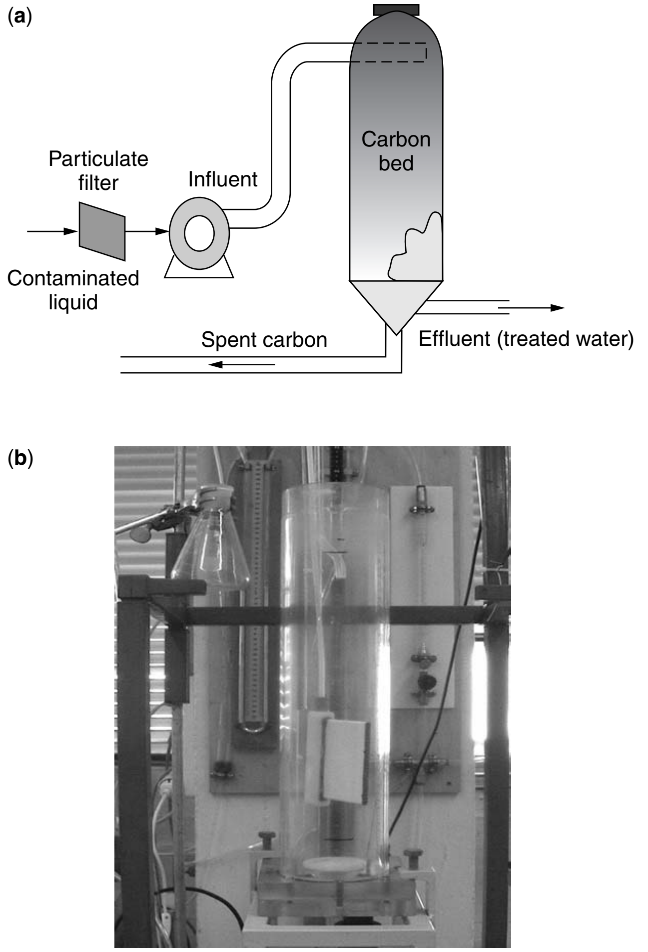

Figure 3. (**a**) Typical fixed-bed carbon adsorption system. From Reference 21. (**b**) A close-up photograph of the hybrid MF laboratory cell based on flotation (22).

suitable as a pretreatment stage for microfiltration. The former accounted for around 90% solids removal.

Ceramic flat-sheet membrane modules of multichannel geometry were used in this cell. The objective was to apply microfiltration by submerged membranes for efficient separation of metal-loaded zeolites. By this hybrid process, low residual metal concentration in treated water at maximum water yield and a higher concentration of metal bonding agent in the concentrates can be achieved (22). During the experimental series of continuous-flow separations by the membranes, all samples collected showed 100% zeolite recovery due to membrane filtration, hence, efficient solid/liquid separation. The metal (zinc) removal depended only on the amount of zeolite sorbent used; the remaining process parameters, examined in detail, had absolutely no influence on the former.

One of the early notable investigations of adsorptive filtration is that by the EPA (23) at Superfund sites, those sites classified as uncontrolled and abandoned places where hazardous waste is located, possibly affecting the local ecosystem or people. In this application, toxic metals were removed by attachment to a thin layer of iron oxides that were immobilized on the surface of an appropriate filter medium, such as sand grains.

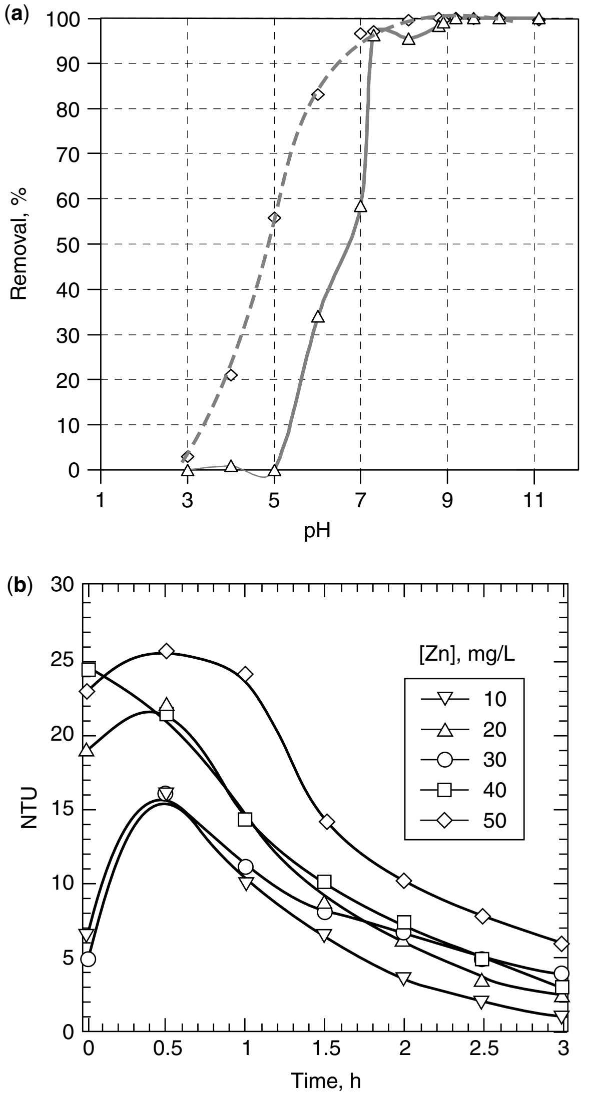

Figure 4. (**a**) Removal of copper by biosorption (rhombic symbols) or simply, precipitation (triangles): A comparison of the two processes. Modified from experimental results published by Zouboulis et al. (24). (**b**) Influence of linear velocity on residual concentration expressed as turbidity (pH 9.5). Reprinted with permission from Reference 26; copyright (2002) Dekker.

An alternative sorbent that has been tried quite extensively is microbial biomass (24). The reason is that thousands of tons of residual biomass are produced each year from the fermentation, pharmaceutical, and chemical (i.e., citric acid biosynthesis) industries and also from biological wastewater treatment plants. The bacterial cell surfaces are known to be anionic, due to ionized groups in the various cell wall polymers. From a comparison of biosorption and other metal separation methods, such as filtration and centrifugation, it was found that the former favors the removal efficiency and applicability at lower (acidic) pH. In these conventional processes, metal cation removal is mainly due to their precipitation as hydroxides from pH alteration; this is presented in Fig. 4a showing a pH front/edge moving to the right toward alkaline values; precipitation was followed by flotation with dodecylamine as a surfactant. Of course, in sorption, metals are bonded on dead biomass (*Streptomyces rimosus*). Oxidation processes leading to insoluble products can also be mediated by specific bacteria, for example, for Mn(II) and Fe(II) (25); in these papers, from our lab, upflow filtration units were tested. One of the objectives was to obtain low sludge volumes, which often depends upon the method of precipitation.

Upflow filtration has a theoretical advantage because coarse-to-medium filtration can be achieved by a single medium (such as sand) with almost perfect gradation of pore space and grain size from coarse to fine in the direction of flow. The bed is backwashed in the same direction but at higher flowrates, so the desired relative positions of fine media are maintained or reestablished by each run (1,4,26). The shortcomings of upflow filtration were overcome by using floating filter media. This type of filter can be backwashed with minimum water consumption and at lower rates. It also appeared from the laboratory experiments that metal removal was not limited by the solubility of the respective zinc hydroxide but by the effectiveness of the subsequent solid/liquid separation method (Fig. 4b). This semibatch unit, consisting of two columns in series and external recirculation of the filtered wastewater from the unit to the feed tank, was suitable for small- to medium-scale applications.

In conclusion, sorptive filtration presents a promising technique. Today, there is a tendency for combined and compact processes, offering both the capacity for effluent treatment for environmental reasons, plus the recovery of metals that otherwise would be lost and, more importantly, water reuse.

Acknowledgments

Thanks are due to Dr. A.I. Zouboulis (AUTh) and Prof. A. Grohmann (WABOLU, Berlin) for their help.

BIBLIOGRAPHY

1. Metcalf & Eddy Inc. (1991). *Wastewater Engineering: Treatment, Disposal, Reuse*. Revised. McGraw-Hill, Singapore, p. 248.
2. Purchas, D.B. (Ed.) (1977). *Solid/Liquid Separation Equipment Scale-Up*. Uplands Press, Croydon, U.K., pp. 289, 319, 365, 445.
3. Wankat, P.C. (1986). *Large-Scale Adsorption and Chromatography*. CRC Press, Boca Raton, FL.
4. Fair, G.M., Geyer, J.G., and Okun, D.A. (1968). *Water and Wastewater Engineering*. John Wiley & Sons, New York, Vol. 2, p. 27–1.
5. Grassi M.T., Shi B., Allen H.E. 1997. Sorption of copper by suspended particulate matter, *Coll. Surf. A* **120**: 199–203.
6. Univ. Pretoria, accessed 2004. Training module, Principles of Filtration, site: http://www.up.ac.za/academic/chemeng/Facilities/Tribology_presentations/05aSepTech.pdf.
7. Sposito, G.A. (1984). *The Surface Chemistry of Soils*. Oxford University Press, Oxford.
8. García-Sánchez, A. and Álvarez-Ayuso, E. (2002). Sorption of Zn, Cd, and Cr on calcite. Application to purification of industrial wastewaters. *Miner. Eng.* **15**: 539–547.
9. Inagaki, M., Kawahara, A., Nishi, Y., and Iwashita, N. (2002). Heavy oil sorption and recovery by using carbon fiber felts. *Carbon* **40**: 1487–1492.
10. Pirbazari, M., Ravindran, V., Badriyha, B.N., and Kim, S.-H. (1996). Hybrid membrane filtration process for leachate treatment. *Water Res.* **30**: 2691–2706.
11. Al-Degs, Y., Khraisheh, M.A.M., and Tutunji, M.F. (2001). Sorption of lead ions on diatomite and manganese oxide modified diatomite. *Water Res.* **35**: 3724–3728.
12. HydroGlobe, Inc., accessed 2004. Site www.hydrogobe.com
13. Lazaridis, N.K., Pandi, T.A., and Matis, K.A. (2004). Chromium(VI) removal from aqueous solutions by Mg-Al-CO_3 hydrotalcite: Sorption–desorption kinetic and equilibrium studies. *Ind. Eng. Chem. Res.* **43**: 2209–2215.
14. Zamboulis, D., Pataroudi, S.I., Zouboulis, A.I., and Matis, K.A. (2004). The application of sorptive flotation for the removal of metal ions. *Desalination* **162**: 159–168.
15. Lehmann, M., Zouboulis, A.I., and Matis, K.A. (2001). Modelling the sorption of metals from aqueous solutions on goethite fixed-beds. *Environ. Pollut.* **113**: 121–128.
16. McKay, G. (1996). *Use of Adsorbents for the Removal of Pollutants from Wastewaters*. CRC Press, Boca Raton, FL.
17. Dzombak, D.A. and Morel, F.M.M. (1990). *Surface Complexation Modelling: Hydrous Ferric Oxide*. John Wiley & Sons, New York.
18. Matis, K.A., Lehmann, M., and Zouboulis, A.I. (1999). Modelling sorption of metals from aqueous solution onto mineral particles: The case of arsenic ions and goethite ore. In: *Natural Microporous Materials in Environmental Technology*. P. Misaelides, F. Macašek, T.J. Pinnavaia, and C. Colella (Eds.). Kluwer Academic, The Netherlands, pp. 463–472.
19. Zouboulis, A.I., Matis, K.A., Loukidou, M., and Šebesta, F. (2003). Metal biosorption by PAN-immobilized fungal biomass in simulated wastewaters. *Coll. Surf. A* **212**: 185–195.
20. Matis, K.A., Zouboulis, A.I., Mandjiny, S., and Zamboulis D. (1997). Removal of cadmium from dilute solutions by hydroxyapatite. Part III. Flocculation studies, *Sep. Sci. Tech.* **32**: 2107–2128.
21. Federal Remediation Technology Roundtable. Ex Situ Physical/Chemical Treatment (assuming pumping), site: http://www.frtr.gov.
22. Lazaridis, N.K., Blöcher, C., Dorda, J., and Matis, K.A. (2004). A hybrid MF process based on flotation. *J. Membrane Sci.* **228**: 83–88; (b) Matis, K.A., Blöcher, C., Mavrov, V., Chmiel, H., and Lazaridis, N.K., Verfahren und vorrichtung zur membranunterstützten flotation, Patents DE 10214457 C1 & EP 1354632 A1.
23. EPA (1993). *Emerging technology summary: Metals treatment at Superfund sites by adsorptive filtration*. EPA/540/SR-93/515, Cincinnati, OH.
24. Zouboulis, A.I., Rousou, E.G., Matis, K.A., and Hancock I.C. (1999). Removal of toxic metals from aqueous mixtures: Part 1. Biosorption. *J. Chem. Tech. Biotechnol.* **74**: 429–436.
25. Katsoyiannis, I.A. and Zouboulis, A.I.: (a)(2002). Removal of arsenic from contaminated water sources by sorption onto iron-oxide-coated polymeric materials; (b) (2004). Biological treatment of Mn(II) and Fe(II) containing groundwater: Kinetic considerations and product characterization. *Water Res.* **36**: 5141–5155; **38**: 1922–1932.
26. Zouboulis, A.I., Lazaridis, N.K., and Grohmann, A. (2002). Toxic metals removal from waste waters by upflow filtration with floating filter medium. I. The case of zinc. *Sep. Sci. Tech.* **37**: 403–416.

QUALITY OF WATER IN STORAGE

Paulo Chaves
Toshiharu Kojiri
Water Resources Research Center
Kyoto University, Japan

Water is our most precious natural resource. It is not only indispensable for the survival of all living beings but also has social and economic importance. It is no coincidence that most early human settlements were located near sources of freshwater. However, available water resources are increasingly under pressure from human population growth, increased standard of living, urbanization, and development of extensive areas for agriculture. As a result, available water may, at times, become insufficient to fulfill all the needs of human living, particularly during dry periods, which are caused by the natural randomness of the hydrological cycle.

Therefore, it becomes necessary to find and to develop new sources of freshwater in order to secure a constant and appropriate supply of water. Certainly, many methods to maintain the stable supply of water exist, such as groundwater exploitation and water deviation from wet to dry areas. Undoubtedly, however, the most popular way worldwide to stabilize water supply is through the operation of stored water, particularly in lakes and manmade reservoirs. This concept was appreciated long ago, and reservoir construction has a long history, with the first dam believed to be built in ancient Egypt around 2800 B.C. (1).

In the last few decades, most of these water storage systems have suffered from poor water quality, mainly caused by human activities within their watersheds and inefficient operation of these reservoirs. Although a variety of techniques to improve water quality in storage exist, most are either very costly or not sustainable. Therefore, efforts to find new methods to achieve better water quality are extremely important. One way is by taking advantage of the close relationship between quality (limnological) and quantity (hydrological and climatological characteristics) of water to better manage stored volumes to yield optimal benefits for water use while at the same time improving its quality.

In the following sections, the main issues that impact on the quality of water in storage are presented. First, the important relationship between quantity and quality of water, and the potential improvement of water quality through the proper operation of storage volumes are highlighted. Then, the main issues related to evaluation of water quality are explained, showing the importance of proper water quality for different uses. Finally, all these topics are combined within the storage management and operation framework when water quality is considered.

WATER QUALITY AND STORAGE

Manmade reservoirs and lakes used for water storage, like other water bodies, are vulnerable to increasingly frequent problems of water quality. Some of the most common problems concerning the quality of water in storage are organic pollution, anoxia (deficiency in oxygen), eutrophication (increase of phytoplankton because of high nutrient concentrations), toxicity, turbidity, siltation, and waterborne diseases. In the same way, a variety of different possible causes of these problems exist, such as industrialization, increase of population, overexploitation of agricultural land, bad farming practices, poverty, lack of proper sanitation systems, runoff from cities, deforestation, lack of proper enforcement of the regulations for pollution control, and overuse of water resources with decrease of water volumes and levels.

Probably, the most effective method to control pollution and improve quality of water in storage is through comprehensive watershed management, which includes reduction and control of nonpoint and point sources of pollution, such as nutrient and organic matter loads. On the other hand, other techniques exist that may be applied directly to the water bodies themselves, such as artificial mixing, aeration, sediment removal, chemical manipulation, algaecides, and biomanipulation of species like fishes and aquatic plants. However, most of these techniques are costly and difficult to implement.

Basically, three main components of a reservoir system exist: inflow, storage, and release. Inflows may be defined as natural events, whereas storage and release may be considered under human control. Nevertheless, both natural variability and human decision can greatly influence the water quality of a reservoir. Reservoirs tend to present water of different quality during wet and dry hydrologic periods. With the increased runoff during periods of rain, higher nutrient loads would be expected in the reservoir, and these nutrients stimulate algae blooms and, therefore, may lower the water quality. Moreover, higher levels of inflow may also increase the inputs of organic matters resulting in dissolved oxygen (DO) depletion, which is defined as the decrease of dissolved oxygen concentrations, vital for sustaining aquatic life. On the other hand, higher inflows increase flushing rates and enhance mixing characteristics, decreasing stratification and potentially improving water quality. Hydrological factors, however, are not the only influences on water quality. Climatologic characteristics also play an important role, because during warm periods a higher probability of reservoir thermal stratification exists, because of a difference in temperature between water layers, leading to a difference in water density. The phenomenon of stratification, defined as the difference in density through the water column, which is commonly caused by variations of temperature, restricts the mixing of water between different layers. The isolated lower layer, called hypolimnion, can suffer from DO depletion by anaerobic processes and an increase in nutrients and other materials from bottom sediments.

Water quality is closely tied to storage volumes. Clearly an increase of water volume results in increased dilution of pollutants, but, in addition, other physical, chemical, and biological processes that influence the quality of water may be closely related to controllable hydrological factors, such as the variation of release discharge, storage volumes, and water levels. For example, an increased retention time (defined as the ratio between reservoir volume and discharge rate) can influence water quality in many ways, such as decreasing mixing proprieties, increasing phosphorus retention, increasing the difference between surface-bottom temperatures (stratification), and decreasing turbidity that can enhance chances of algae bloom. Moreover, fluctuation of water levels may also enormously affect water quality, such as increased erosion from the shores that may result in higher turbidity, increasing washout of margins, and decreasing reproduction of some fish species (that lay eggs on coastal vegetation), consequently changing the phytoplankton (algae) population of the aquatic system. The quality of stored water may also be affected by the release of water with poor quality from different levels, such as bottom-spillway discharge.

Zalewski et al. (2) emphasizes the importance of ecohydrology to use ecosystem properties as management tools for sustainability of water resources, such as through adequate management of quantities to improve water quality. Moreover, his paper presents an example of the relationship between storage levels and the eutrophication process for the Sulejow Reservoir. Straskraba and Tundisi (3) discuss, in a general way, many aspects related to the management of reservoir water quality, such as ecotechnologies, relation between storage and water quality, and the possible impacts of different volumes and quality of discharged water from reservoirs.

Therefore, the proper operation of release discharges and storage volumes considering single or multiple outlets and offtakes can also be considered an efficient method to improve the quality of stored water, by discharging water of undesirable quality or by increasing the storage volume with additional water, hence diluting water of poor quality (4,5). Although some general idea of the relationships between hydrologic and climatologic variables and water quality condition could be described, it is not possible to make a definitive statement about such relations. Specific studies for each reservoir system are needed to truly understand its limnology characteristics. Such studies need to include data collection and analysis, application of mathematical models, biological and hydrologic investigations, and, lastly, an understanding of the socioeconomic characteristics of the reservoir basin.

Mathematical models can play an important role in the decision-making process, as these tools may support the decision makers in making the “best” decisions. The outcome of different decisions may be simulated and explored virtually with relatively low costs and in a short period of time. For more detailed explanation behind the mathematics of physical, chemical, and biological concepts of the quality models, the readers are referred to Jorgensen (6), Orlob (7), and Thomann and Mueller (8).

WATER QUALITY EVALUATION

Water quality directly affects virtually all water uses. Fish survival, diversity, and growth; recreational activities; corrosion of turbines of hydropower plants; municipal, industrial, and domestic water supplies; agricultural uses like irrigation and livestock; waste disposal; and general aesthetics are all affected by the physical, chemical, biological, and microbiological conditions that exist in water bodies. Some water uses can affect water quality, which can influence other water uses. For instance, navigation may cause increased bank erosion leading to accelerated loss of storage volumes (reservoir sedimentation) and higher water turbidity. Water bodies used for wastewater assimilation may suffer from depletion of DO, which is required to sustain aquatic life, and an increase in total dissolved solids (TDSs), which is highly relevant to water supply, as higher TDSs increase treatment costs.

The minimum acceptable quality of water depends very much on the water use itself. Water for irrigation, for example, should be low in dissolved salts, but water intended for livestock should be low in bacteria. Water used in industrial processes should usually be of a much higher quality than water used for industrial cooling. As for municipal supply, water must not only be safe to drink, but ideally should contain low concentrations of materials such as calcium and iron, as they may cause costly infrastructure damage or add unpleasant characteristics to the water even after treatment (9).

A water quality index may be an efficient way to objectively measure water quality using a vast number of existing water quality parameters, helping decisions to be taken in a more efficient and less subjective way. Many of the water quality indices tend to be developed based on local characteristics and local expertise (10). Water quality indices may also be developed to express the conformity or net benefits of a water body with regard to a particular type of water use. In this scenario, different indices may be developed for the evaluation for a specific use or of the natural state of the system, such as for domestic drinking water supply, recreational use, wildlife preservation, trophic state situation (measure of eutrophication), and toxicity levels of water. Indices can be based on a single quality parameter or can even be presented as a combination of many quality parameters, such as the well-known Water Quality Index developed by the U.S. Environmental Protection Agency (EPA). Therefore, their use in the management and operation framework has to be closely studied and discussed among all stakeholders involved in the decision-making process. As a result of the vagueness involved in defining these indices, some new attempts have been made to evaluate water quality using fuzzy theory, such as in works by Norwick and Turksen (11), Dombi (12), and Chang et al. (13).

MANAGEMENT OF STORAGE WATER QUALITY

Water volume has always been a prime consideration in the design and construction of storage reservoirs, but quality has not. In the same way that quantity is regulated, water can be stored and released for the improvement of water quality and net benefits. As described in the previous sections, assessment, evaluation, and modeling of water quality are very complex tasks for water resource managers. Some of the reasons why water quality is still being left out from the decision-making framework process include:

- Multiple stakeholders and multiple objectives are present in most cases.
- A large number of physical, chemical, and biological water quality parameters exist to be considered.
- Quality conditions present great spatial and temporal variability.
- It is difficult to economically evaluate water and environmental quality.
- Quality may be highly influenced by human activities and natural conditions.
- Complete and representative spatial and temporal water quality data are seldom available.
- Water quality simulation is a complex task because of the high levels of uncertainty involved in this process.

The uncertainties when analyzing water quality are much greater than when dealing with water quantity. Uncertainty may strongly affect water quality management and the decision-making process of storage operation. However, from where does uncertainty come? Basically, uncertainty is inherent to many processes, such as measurement, natural randomness, and human perception. In a more concrete way, some sources of uncertainties assessing water quality include:

- Construction problems of treatment facilities
- Structures and parameters of a simulation model
- Uncertainty of mismeasurement and incomplete data sets
- Future and changing conditions (e.g., new socioeconomic activities in the watershed, global warming)
- Vagueness of objectives (e.g., “improvement” of water quality)
- Situations of accidents and disaster (e.g., traffic accidents or chemical leakage from factories)
- Natural randomness of climate and hydrological variability (e.g., extreme events such as floods or long drought periods)

Proper identification of the system uncertainties can support a more flexible and effective management framework, where components identified with greater uncertainty are considered for further investigation and the ones with less uncertainty may become more important decision tools, as the consequences can be more accurately predicted. Moreover, it may increase the public acceptance of a certain decision regarding reservoir management. Results after uncertainty analysis can be given in different ways, such as probability density function and confidence intervals, giving a range of possible outcomes avoiding the common “unexpected” results.

Moreover, information on cost is frequently needed to improve the water quality management prior to the water quality impact assessment. Such costs may be related to water treatment; penalties for excessive pollution; socioeconomic losses, such as associated with tourism activities; and decrease of benefits of certain objectives when a multipurpose operation is considered.

The operation of single and multiple outlet reservoir systems can present different objectives, such as improvement of water quality in stored volumes, released discharges, and intake water. Moreover, the management horizon can be specified to deal with different problems, such as related to the long-term planning, real-time operation, and emergency situations.

Management of water quality can be considered together with the various types of reservoirs that may present different water use, such as reservoirs for water supply, irrigation, hydropower generation, flood control, recreation, and, of course, multipurpose systems. As explained previously, for each type of reservoir, a different evaluation method may be necessary.

Besides the improvement of water quality condition itself, many advantages to considering water quality in the operation process exist, such as wider acceptance of reservoirs among the general public, decrease in maintenance costs, increase in net benefits, preservation of aquatic life, and achievement of sustainable policies.

Straskraba and Tundisi (3) present many management models that combine quality simulation with management tools. Research in this field is in constant development, and some of the new models are imbedded with complex components, such as remote sensing information, geographic information system (GIS), management and optimization techniques, artificial intelligence (AI), and decision support systems (DSSs).

It seems that much still exists to be accomplished and developed in the field of water quality in storage. The focus should be on integrated quality management, which must consider quantity and quality along with the three main components of a storage reservoir system: inflow, storage, and release. Moreover, other topics should be considered such as simultaneous combination of stochastic optimization and water quality simulation models, including uncertainty analysis, multiple water quality parameters in a multiobjective optimization framework, and multipurpose reservoirs considering the different stakeholders and water uses. The potential success in solving water quality related problems will depend on the ability of different professionals, including researchers, hydrologists, limnologists, and decision-makers, to work together in an interdisciplinary and holistic manner.

BIBLIOGRAPHY

1. Biswas, A.K. (1970). *A History of Hydrology*. North Holland Publishing, Amsterdam.
2. Zalewski, M. et al. (1997). *Ecohydrology: A New Paradigm for Sustainable Use of Aquatic Resources*. UNESCO, Paris, France.
3. Straskraba, M. and Tundisi, J.G. (1999). *Guidelines of Lake Management—Reservoir Water Quality Management*. Vol. 9. ILEC, Japan.
4. Welch, E.B. and Patmont, C.R. (1980). Lake restoration by dilution, Moses Lakes, Washington. *Water Res.* **14**: 1317–1325.

5. Chaves, P., Kojiri, T., and Yamashiki, Y. (2003). Optimization of storage reservoir considering water quantity and quality. *Hydrol. Processes* **17**: 2769–2793.
6. Jorgensen, S.E. (1983). *Application of Ecological Modeling in Environment Management*. Elsevier, Amsterdam.
7. Orlob, G.T. (1983). *Mathematical Modeling of Water Quality: Streams, Lakes and Reservoirs*. John Wiley & Sons, Chichester, UK.
8. Thomann, R.V. and Mueller, J.A. (1990). *Principles of Surface Water Quality Modeling and Control*. HarperCollins Publishers, New York.
9. Kiely, G. (1997). *Environment Engineering*. McGraw-Hill, Berkshire.
10. Thanh, N.C. and Biswas, A. (1990). *Environment—Sound Water Management*. Oxford University Press, Oxford, UK.
11. Norwick, A. and Turksen, I. (1984). A model for the measurement of membership and consequences of its empirical implementation. *Fuzzy Sets Syst.* **12**(1): 1–25.
12. Dombi, J. (1990). Membership function as an evaluation. *Fuzzy Sets Syst.* **35**: 1–21.
13. Chang, N.B., Chew, H.W., and Ning, S.K. (2001). Identification of river water quality using the fuzzy synthetic evaluation approach. *J. Environ. Manag.* **63**: 293–305.

QUALITY OF WATER SUPPLIES

Xinhao Wang
Diana Mitsova-Boneva
University of Cincinnati
Cincinnati, Ohio

Adequate quantity and quality of water are necessary for human living and activities. A broad definition of water supply includes any quantity of available water. A more restricted concept of water supply refers to water collected and distributed to the general public or to other public or private utilities for residential, commercial, and industrial use. An even narrower definition of water supply only includes the water collected and conveyed for use in a community or a region (1). In the discussion here, water supply refers to water from the process of collection and distribution to meet human demand.

WATER RESOURCES

Water resources are renewable but finite and scarce. As shown in Fig. 1, 97.22% of earth's water is captured in oceans. Its total volume is enormous—1.3 billion cubic kilometers—but inadequate for human consumption because of the salt content. Approximately 2% of water is locked up in polar icecaps and glaciers. Water found in land, including surface water and groundwater, makes up less than 1% of the earth's water resources (2). Groundwater is found in aquifers, moisture-laden strata where water fills the spaces between rock particles. Although only about half of groundwater can be economically withdrawn for human use, it represents more than 97% of the usable freshwater resources. In addition, groundwater is a major source of replenishment for surface water.

WATER CONSUMPTION

Human water consumption consists of three major categories—agricultural, industrial, and domestic water use (3). According to the Food and Agriculture Organization of the United Nations (FAO), agricultural use accounts for 70% of global water withdrawals, while industrial and municipal use account for another 20% and 10%, respectively (4). Farming withdraws large amounts of water from rivers, lakes, reservoirs, and wells in order to irrigate crop fields and sustain plant growth in agricultural and horticultural practices. Almost 60% of the world's freshwater withdrawals are used for irrigation purposes. In Asia, irrigated agriculture averages approximately 82% of all water consumption. The proportion in the United States is 41% and about 30% in Europe (5). Irrigation water is applied through irrigation systems for the purposes of field preparation, preirrigation, application of chemicals, weed elimination, dust reduction, plants cooling, frost protection, leaching salts from roots, and harvesting. Irrigation water is also used by golf courses, parks, nurseries, turf farms, and cemeteries.

Industry consumes the second largest amount of water. In heavily industrialized regions, the percentage could

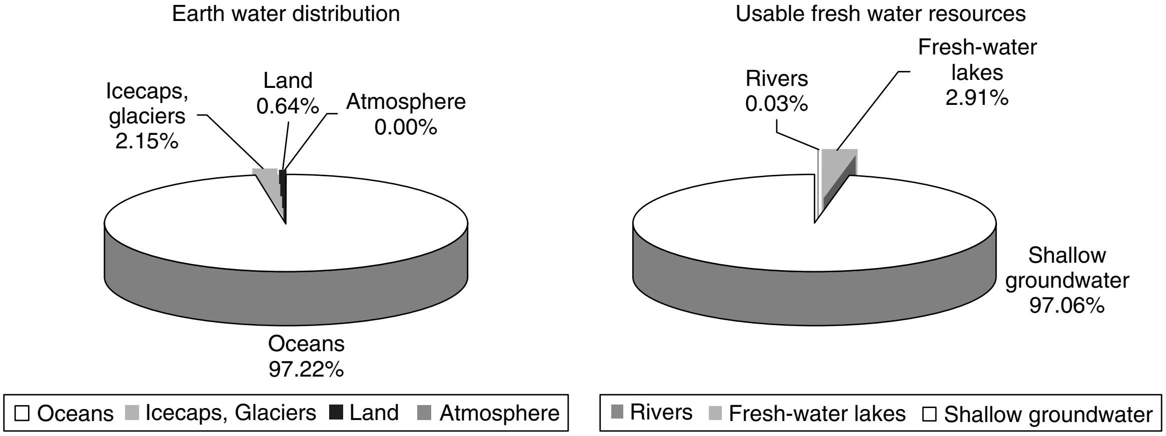

Figure 1. Earth's water resources.

be much higher. In Europe, industry accounts for 52% of the total water consumption and in North America, 44% (6). Almost one-half of the industrial withdrawals are utilized for generating thermoelectric power with steam-driven turbine generators. Much of the industrial water is used for purposes such as cooling and is returned to its source. Large amounts of water are required also by food processing and paper producing industries, oil refineries, and chemical and metallurgical plants.

Domestic water usage accounts for the last 10% of the total water consumption. This category refers to domestic consumption of water for drinking, hygiene, and other household uses, as well as commercial and institutional uses. The consumed domestic water is collected and treated in wastewater treatment plants before being discharged back to a natural water body.

TYPES OF WATER SUPPLY

Water supply systems refer to private (self-supplied, such as private wells) or public (may be operated by public agencies or private companies) withdrawals from ground and surface water designated to serve specific users. The most common is the public or municipal water supply system, which provides water to domestic, public, and commercial users and, occasionally, to industries, hydroelectric and thermoelectric plants, or for irrigation (7). The majority of the population throughout the world, especially urban population, is served by public water supply systems. Figure 2 illustrates the average percentage of the large city population served by each type of water supply service, by region. Household connection or yard tap and public tap are public water supplies. In the more developed regions, such as Europe and North America, more than 96% of the urban population has household connection to public water supply services. In Africa and Asia, considerable numbers of people are not served by public water supply systems.

QUALITY OF WATER SUPPLY

Water is an excellent solvent that can dissolve many elements. The amount of each element affects water quality, reflected in the chemical, physical, or biological condition of water. Therefore, water quality is often measured with what is contained in water, such as dissolved oxygen, suspended solids, pesticides, metals, oils, minerals, and nutrients. Pathogenic bacteria might affect human and animal health and so they are often measured. Physical conditions are another set of water quality measurement, such as water temperature, water color, and turbidity. Water quality also can be assessed with biological indicators, which measure the type and diversity of fish, macroinvertebrate species, and plants (9).

Some uses of water for industrial activities, such as mining, cooling water supply, washing, fire protection, or oil well repressurization, may not depend on water quality. Different industries may have different requirements for water quality, such as the turbidity or temperature. Normally, industrial water use does not require filtration and chlorine treatments.

Quality criteria for irrigation water include measurements of its salinity, alkalization, toxicity, and electrical conductivity. Salinity indicates the presence of high concentrations of chloride, sulfate, and bicarbonate compounds of sodium, calcium, and magnesium. In regions with significant rainfalls, those soluble salts are flushed out from the crop root zone on a regular basis. In arid and semiarid areas, water percolating through the soil is not sufficient to leach those salts. The increasing concentrations may adversely affect crops with low level salt tolerance thresholds. Therefore, irrigation waters must be controlled for their content of dissolved salts in order to prevent further deterioration of the soil quality and reduction of crop yields. There is a direct positive relationship between salt concentrations and the electrical conductivity of irrigation water. Alkalization occurs when the soil is more saturated with sodium than calcium and magnesium. Excess sodium results in high pH, reduced oxygen content, soil erosion, and less nutrient absorption by plants. Some constituents of the irrigation waters such as boron, chloride, and sodium are toxic to crops. Irrigation waters containing such chemicals are not recommended for use in watering soybeans, citrus, or grape plants (10).

Water quality of domestic water supplies has a direct effect on human health. Microbial pathogens that are found to cause enteric diseases in humans include bacteria (such as *Vibrio cholerae, Shigella, Salmonella, Campylobacter*, and *Escherichia coli*), viruses (such as hepatitis A virus and rotavirus), and parasites (such a *Giardia lamblia* and *Cryptosporidium*). Bacterial waterborne infections include cholera, dysentery, salmonellosis, typhoid fever, “traveler’s disease,” and *E. coli* outbreaks. The most common viral disease contracted through contaminated water is hepatitis A. Sewage-tainted drinking water can also cause certain parasitic protozoan diseases such as amebic dysentery and giardiasis. Those diseases are usually contracted by drinking water or eating food contaminated with pathogens of fecal origin. They are characterized by severe diarrhea, vomiting, and cramps, and in some cases, by high fever, headache, and muscle pains. The most common cause of death associated with those ailments is rapid dehydration and subsequent collapse of the vascular system (11).

Domestic water supply for drinking purposes is the most critical of all. The World Health Organization (WHO) estimates that each year approximately 2.5 million children all over the globe, ages 0–12, die of waterborne diarrhea illnesses. Waterborne diseases are the leading cause of death in most developing nations in Africa, Asia, and Latin America. Despite efforts to render safe drinking water to an increasing number of people, as a result of the unprecedented population growth, approximately 1.1 billion people still use unsafe drinking water (12). Water disinfection in the last century has largely reduced the incidence of infectious waterborne diseases. Health officials today and the general public, however, are more and more concerned about chronic health effects potentially caused by the presence of toxic chemicals in drinking water. Ingestion of even small amounts of some inorganic pollutants such as arsenic, lead, cadmium, chromium, mercury, asbestos, cyanide, or nitrates can

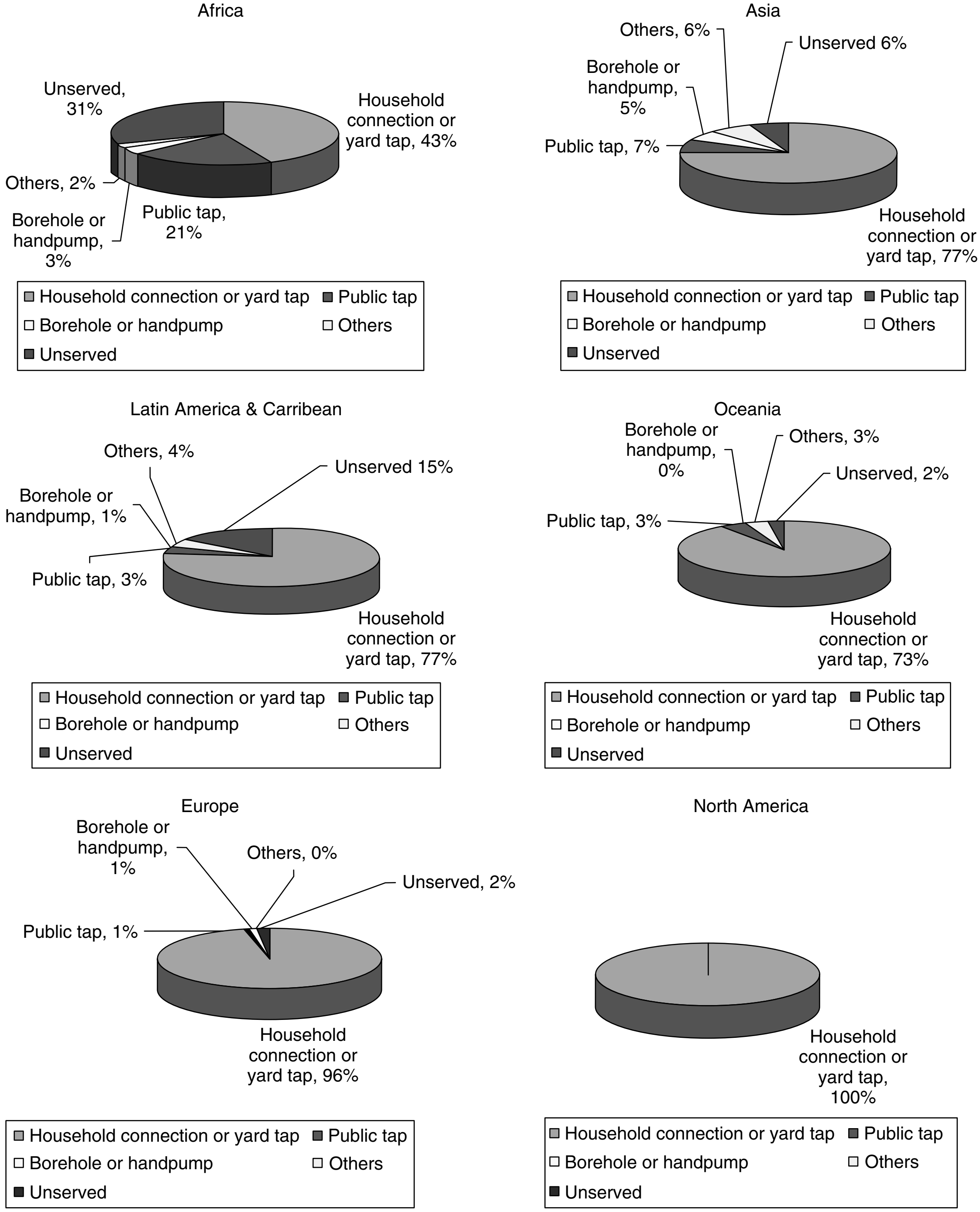

Figure 2. Water supply in the largest cities: average percentage of the population served by each type of service, by region (8).

cause a wide range of health problems. Especially vulnerable to lead poisoning are young children. Ingested lead can result in anemia, neurological disorders, kidney problems, sterility, and birth defects. Evidence suggests that exposure to even very low levels of asbestos in drinking water can result in cancers of the gastrointestinal tract. Although conclusive evidence about cause-and-effect link between chemical contaminants and cancers, birth defects, miscarriages, nervous and reproductive system disorders, and organ damage has not yet been established, there is little doubt that prolonged exposure to some organic and inorganic substances can result in chronic health conditions. Nitrates are another inorganic chemical with serious health effects. The intake of nitrate-containing water by infants under the age of six months may cause methemoglobinemia, known also as the "blue baby syndrome" (13).

The organic compounds detected in drinking water include a wide variety of herbicides and pesticides, disinfection by-products, PCBs, solvents, oils, and so on.

They are volatile and nonvolatile in nature. Many of them are suspected carcinogens (14). *Radionuclides* are also known to be present in some drinking water supplies.

The quality of the drinking water supplies in the United States is one of the highest on earth. With the adoption of the Safe Drinking Water Act in 1974, subsequently amended in 1986 and 1996, the U.S. Congress empowered the U.S. Environmental Protection Agency (U.S. EPA) to enforce regulations aimed at protecting the quality of drinking water supplies. The U.S. EPA has established two sets of standards for water quality: *primary* and *secondary*. The objective of the primary standards is to protect human health. They are based on toxicological, epidemiological, and clinical data. The *maximum contaminant level* (MCL) is the primary standard enforced by the U.S. EPA. Secondary drinking water standards refer to contaminants that create unpleasant odor, taste, and color or may be corrosive or staining. They can cause minor skin irritations or damage clothing and appliances (15). Table 1 displays the major contaminant categories, their potential health effects, and sources of contaminants in drinking water.

FACTORS AFFECTING THE QUALITY OF WATER SUPPLY

The quality of the water source is critical for the quality of water supply. Water pollution refers to the deteriorating quality of water due to the presence of pollutants in the water. Effluent discharges from industries and municipal sewage plants are considered point source pollution. They are relatively easy to identify, regulate, control, and monitor since they are usually located near factories, wastewater treatment plants, septic tanks, or stormwater outfalls. In the United States, industries discharging directly into rivers and lakes are required to apply for permits in accordance with the National Pollutant Discharge Elimination System (NPDES) program. The permit specifies the volume, type, and permissible concentrations of pollutants allowed to be released in the waterways. The other option for controlling industrial discharges is to reuse the treated wastewater instead of discharging it into receiving waters.

Nonpoint source pollutants consist of suspended solids, nutrients, petroleum, inorganic, and radioactive compounds. A major contributor of sediment to the waterways is the runoff from agricultural fields and construction sites. Silt and soil increase the turbidity of water, which impairs the productivity of photosynthetic plants, reduces water depth, endangers bottom-dwelling animals, and suffocates fish. Nutrients enter waterways in the form of fertilizer and sewage runoff, leaves and grass, or as runoff from livestock farms and pastures. Many hazardous chemicals such as asbestos, hydrocarbons, PCBs, pesticides, mercury, and lead enter the waterways

Table 1. The U.S. EPA's National Primary Drinking Water Standards

Contaminant	Potential Health Effects from Ingestion of Water	Sources of Contaminant in Drinking Water
Microorganisms	Gastrointestinal illness, legionnaire's disease	Human and fecal animal waste
Disinfection by-products	Increased risk of cancer, anemia; in infants and young children—central nervous system effects, liver and kidney problems; increased risk of cancer	Human and fecal animal waste
Inorganic chemicals	Increase in blood cholesterol; problems with circulatory system; increased risk of developing benign intestinal polyps; intestinal lesions; kidney damage; allergic dermatitis; nerve damage or thyroid problems; bone disease (pain and tenderness of the bones); delays in physical or mental development of infants and children	Discharge from petroleum refineries; ceramics; erosion of natural deposits; runoff from glass manufacturing sites; decay of asbestos cement in water mains; erosion of natural deposits; discharge from metal refineries; corrosion of galvanized pipes; discharge from steel and pulp mills, plastic and fertilizer factories; water additive that promotes strong teeth; corrosion of household plumbing systems; erosion of natural deposits
Organic chemicals	Nervous system or blood problems; increased risk of cancer; eye, liver, kidney, adrenal gland, or spleen problems; anemia; cardiovascular system or reproductive problems; circulatory system problems; thymus gland problems; immune deficiencies	Added to water during sewage/wastewater treatment; runoff from herbicide used on row crops; leaching from gas storage tanks and landfills; leaching from linings of water storage tanks and distribution lines; leaching of soil fumigant used on rice and alfalfa; discharge from chemical plants, agricultural chemical factories, and other industrial activities; residue of banned termiticide; runoff from herbicide used on rights of way; runoff/leaching from soil fumigant used on soybeans, cotton, pineapples, and orchards; emissions from waste incineration and other combustion; discharge from wood preserving factories; discharge from rubber and plastic factories; leaching from landfills; discharge from factories and dry cleaners; runoff/leaching from insecticide used on cotton and cattle; discharge from textile finishing factories; breakdown of heptachlor; residue of banned insecticide

through airborne deposition. Acid rainfall is a typical example of airborne pollution that affects the waterways.

In addition to water quality at sources, the contamination of distribution pipelines may be a major concern for the quality of water supply. One cause of contamination is cross-connection, where there is a physical connection between a drinking water supply line and a source of contamination, such as viruses, bacteria, and nitrate. The materials of pipes could also be the source of contamination, such as lead dissolved from pipes and plumbing fixtures.

TREATMENT OF WATER SUPPLIES

A water supply system normally consists of facilities for water withdrawal, storage, treatment, and distribution. After treatment, water is usually resent into the distribution system or to water storage reservoirs (16). Water suppliers apply different types of treatment according to the type of water use and the quality of their source water. Some groundwater supplies can meet the standard requirements without any additional treatment. In most cases, however, and especially when it comes to surface water supplies, a combination of treatment methods is applied to remove contaminants from raw water.

The first stage of the water treatment process is usually to settle the water in order to separate the large solid debris. During the process of c*hemical coagulation* the synthetic organic polymers (cationic and anionic) are aplied to coagulate colloidal material. During *flocculation*, often preceded by *rapid mixing*, tiny particles combine into larger and heavier “flocs.” S*edimentation* is used to reduce sediment loads in raw water both before and after flocculation. It allows for the removal of sand, silt, gravel, and alum floc. Some hard waters require an additional chemical treatment before filtration. Lime-soda softening is applied to raw waters rich in calcium and magnesium. As in flocculation, softening converts hardness-creating ions into insoluble precipitates, which are then removed through settling and filtration. Some raw water supplies containing iron, manganese, and volatile organic compounds need to be treated by *aeration*. This method allows for the oxidation of chemicals, formation of precipitates, and their subsequent removal by filtration. *Filtration* is mostly a physical process in which the water flows through beds of gravel, sand, anthracite, or diatomaceous earth. However, in order to increase its effectiveness, coagulants are also applied throughout the filtration process. *Adsorption* involves the application of activated carbon, which adsorbs nonvolatile organics and some taste- and odor-causing compounds. Ion exchange is used to remove inorganic chemicals (such as arsenic, chromium, or excess fluoride) as well as nitrates and radionuclides (17).

Chlorination, ozonation, and ultraviolet radiation are methods used during the proccess of *disinfection*. Chlorination is very effective in killing bacteria but less effective in eliminating viruses and protozoans. However, it has the advantage to continue the process of disinfection in the distribution pipes. Ozonation and UV radiation have powerful germicidal effect but both have disadvantages. Ozone has a very short half-life and must be generated electrically at the point and at the moment of use. UV disinfection is applied through UV lamps, which allows for contamination (18). Figure 3 illustrates the process in a domestic water treatment plant in Cincinnati, Ohio (19).

Desalination is not a typical water treatment method. However, in arid coastal areas or in regions with mounting pressure on existing freshwater resources, desalination is routinely used to transform seawater into potable water. Saudi Arabia is the leading producer of desalted water in the world. Two methods of desalination are usually applied. Smaller plants rely mainly on reverse osmosis, while larger producers use multistage flash distillation processes (20).

PROTECTION OF WATER RESOURCES

Rapid industrialization since the nineteenth century has been accompanied by reckless waste management practices, which had turned many surface water bodies into sewers or algae-coated cesspits containing microorganisms, toxic chemicals, heavy metals, organic pollutants, pesticides, fertilizers, and their by-products. In addition, these contaminants penetrate the soil, where they can

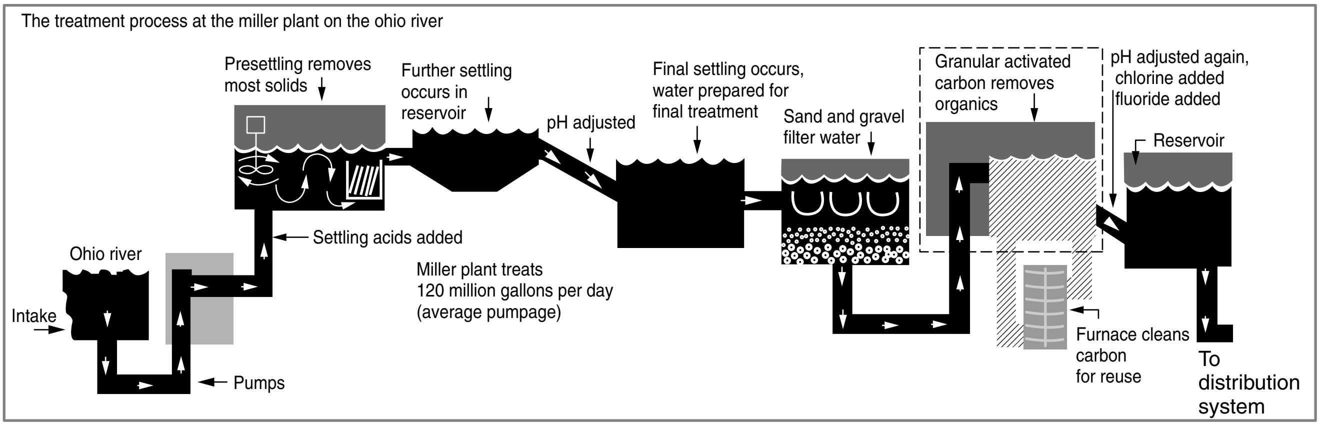

Figure 3. A domestic water treatment plant.

reach the water table and deteriorate the quality of groundwater. Since most of the freshwater supplies rely on groundwater and the slow flow of groundwater, it is critical to keep groundwater from being contaminated. Both the United States and the European Commission have come to the conclusion that setting chemical quality standards for groundwater is not appropriate, as it creates the impression of an allowed level of pollution.

The World Bank estimates that during the past century water consumption has risen sixfold, which is twice as fast as the rate of population growth. By the year 2025, four billion people, particularly in Africa, the Middle East, and South Asia, are expected to live under conditions of severe water stress, if appropriate actions are not taken (21). The total volume of water consumption also affects the quality of water supply. In order to keep the demand within the safe yield of the system, many cities are implementing water-saving policies and restrictions (22).

Three major irrigation methods—sprinkler, microirrigation, and surface (flood)—are applied (23). In general, only a small portion of the irrigation water is actually reaching its intended target. The majority is transformed into vapor through the processes of evaporation and transpiration or is lost in transit (24). Many conventional methods should be supported to increase field water use efficiency and the crop productivity of unit water consumption. Through more efficient irrigating practices such as the automated drip irrigation, alternate irrigation in controlled root zones, and water-deficit-regulated irrigation can substantially reduce water demand in agriculture without reducing crop yield (25). Research has shown that comprehensive water-saving farming technology can increase irrigation productivity 10–15% (26).

Water conservation approaches should also be applied in manufacturing and mining industries and in urban areas to establish a water-saving society. To do so, industrial wastewater should be treated and reused to increase the utilization ratio. Sustainable water resources will not only efficiently support the existing socioeconomic systems but will also provide the required quantity and quality of water at reasonable cost in the future (27).

BIBLIOGRAPHY

1. *The American Heritage Dictionary of the English language*, 4th Edn. (2000). Houghton Mifflin, Boston.
2. U.S Geological Survey (2003). Water science homepage. Available at http://ga.water.usgs.gov/edu/earthwherewater.html.
3. World Resources Institute Database (2000). Available at http://earthtrends.wri.org/searchable_db/index.cfm?theme=2.
4. Food and Agriculture Organization of the United Nations (FAO) (2003). Water Resources, Development and Management Service. *AQUASTAT Information System on Water and Agriculture: Review of World Water Resources by Country*. Available online at http://www.fao.org/waicent/faoinfo/agricult/agl/aglw/aquastat/water_res/index.htm.
5. Nadakavukaren, A. (2000). *Our Global Environment*, Waveland Press, Prospects Heights, NY.
6. World Resources Institute Database (2000). Available at http://earthtrends.wri.org/searchable_db/index.cfm?theme=2.
7. *USGS National Handbook of Recommended Methods for Water Data Acquisition*, Chap. 11. Available at http://water.usgs.gov/pubs/chapter11/chapter11C.html. Accessed 2004.
8. WHO Global Water Supply and Sanitation Assessment 2000 Report. Available at http://www.who.int/docstore/water_sanitation_health/Globassessment/Global4.3.htm. Accessed 2004.
9. U.S. EPA (2004). *Monitoring and Assessing Water Quality Guidelines*. Available at http://www.epa.gov/owow/monitoring/elements/elements.html#d. Accessed 2004
10. Hoffman, G.J. (2004). Water quality criteria for irrigation, EC97-782, University of Nebraska, Institute of Agriculture and Natural Resources, issued in cooperation with the U.S. Department of Agriculture. Available at http://ianrpubs.unl.edu/irrigation/ec782.htm#sum. Accessed 2004.
11. WHO drinking water quality and infectious diseases transmission and prevention. Available at http://www.who.int/water_sanitation_health/dwq/infectdis/en/print.html. Accessed 2004.
12. WHO Joint Monitoring Program Report (2004). Meeting the MDG drinking-water and sanitation target: a mid-term assessment of progress. Available at http://www.who.int/water_sanitation_health/monitoring/jmp2004/en/. 2004.
13. National Sanitation Foundation (1991). *Standard NSF 61-1991 Drinking water system components—health effects*.
14. National Sanitation Foundation (1991). *Standard NSF 61-1991 Drinking water treatment chemicals—health effects*.
15. U.S. Environmental Protection Agency (2002). National Primary Drinking Water Regulations. Available at http://www.epa.gov/safewater/mcl.html. EPA 816-F-02-013.
16. Litke, D.W. and Kauffman, L.F. (1993). Analysis of residential use of water in the Denver metropolitan area, Colorado, 1980–87. *U.S. Geol. Surv. Water Resour. Invest. Rep.* **92–4030**, p. 69.
17. Logsdon, G., Hess, A., and Horsley, M. (1999). Guide to selection of water treatment processes. In: *Water Quality and Treatment—A Handbook of Community Water Supplies*, 5th Edn. R.D. Letterman (Ed.). McGraw-Hill, New York, pp. 3.1–3.26.
18. Haas, C.N. (1999). Disinfection. In: *Water Quality and Treatment - A Handbook of Community Water Supplies*, 5th Edn. R.D. Letterman (Ed.). McGraw-Hill, New York, pp. 14.1–14.60.
19. Greater Cincinnati Water Works. Available at http://www.cincinnati-oh.gov/water/pages/-3283-/.
20. *Encyclopedia of Desalination and Water Resources* (1997). EOLSS Publishers Co. Ltd., UK.
21. World Bank Water Resources Strategy (2003).
22. Jenkins, M.W. et al. (2004). Optimization of California's water supply system: results and insights. *J. Water Resour. Plan. Manage.* **130**(4): 271–275.
23. U.S. Department of Agriculture, National Agricultural Statistics Service. Drought and Fire Survey 2000, Montana's 2000 Drought/Fire Survey Results: press release. Available at http://www.nass.usda.gov/mt/pressrls/misc/firesurv.htm. Accessed June 25, 2003.
24. Hutson, S.S., Barber, N.L., Kenny, J.F., Linsey, K.S., Lumia, D.S., and Maupin, M.A. (2004). *U.S. Geological Survey Estimated Use of Water in the United States in 2000*. USGS Circular 1268, released March 2004, revised May 2004.
25. Kang, S. and Li, Y. (1997). The water economic trends and countermeasures for the 21st century. *Acta Agric. Technol.* **18**: 32–36.

26. Feng, P.G. (1997). Study of the principles of support for water resource sustainability and water economical utilization. *Agric. Res. Arid Area* **15**: 1–7.
27. Loucks, D.P. (2000). Sustainable water resources management. *Water Int.* **25**(1): 2.

THE SUBMITOCHONDRIAL PARTICLE ASSAY AS A BIOLOGICAL MONITORING TOOL

Francis G. Doherty
AquaTox Research, Inc.
Syracuse, New York

BACKGROUND

While studies assessing the toxicity of environmental contaminants to aquatic organisms date back to the early 1900s, the use of toxicity testing as a tool for monitoring the environmental acceptability of discharges from wastewater treatment and manufacturing facilities to surface waters (i.e., rivers, lakes, and oceans) did not begin to evolve until the 1970s. One of the primary motivating factors behind the development of biological monitoring, or biomonitoring, was the manufacture of exponentially increasing numbers of new chemicals for which analytical procedures were generally lacking. Biomonitoring was viewed as an approach to detect the presence of potentially toxic chemicals in facility discharges without necessarily having to identify the contaminants at that stage of the assessment. Additional advantages included the reduced expense associated with biologically testing significantly greater numbers of samples in comparison with analytical procedures. This approach also focused on only biologically available forms of the chemicals in solution reflecting any interactive effects (synergistic or antagonistic) that might be occurring among contaminants within a complex sample.

Initial research efforts in the biomonitoring field tended to focus on identifying a "most sensitive species" that satisfied a variety of other criteria. These criteria included, but were not necessarily limited to, widespread geographic distribution and ease of culture in the laboratory so as to provide a steady supply of consistently healthy test organisms. A significant amount of effort was also invested in the development of standardized protocols that focused on the control of test variables to provide guidance to laboratories conducting aquatic toxicity tests, thereby permitting direct comparisons of toxicity data from multiple sources.

This preliminary round of research tended to use common endemic fish species (e.g., salmonids and cyprinids) and planktonic invertebrates (e.g., Cladocerans) that served as a food source for larval fish stages as the primary test species in whole organism assays. Lethality and growth or reproduction served as test endpoints. These efforts ultimately led to the development of acute ("short-duration") and chronic ("long-duration") standardized protocols most often employing the fathead minnow, *Pimephales promelas*, and the Cladoceran *Ceriodaphnia dubia* for discharges to freshwater systems (1,2). Companion protocols were also developed for discharges to brackish and marine environments employing saltwater organisms.

This initial focus, however, quickly and almost simultaneously expanded to include other less common fish and invertebrate species along with surrogate microorganisms and *in vitro* biochemical assays in a continuing attempt to identify that elusive "most sensitive species or assay." Alternately, these activities addressed refining or enhancing existing protocols with the primary, preferred species. Other objectives included in this broad research effort included accelerating the pace at which analyses were conducted by primarily reducing the exposure time for standardized tests, which also resulted in a reduction in the cost of testing.

SUBMITOCHONDRIAL PARTICLE ASSAY

One such surrogate assay developed during the early to mid-1980s was the *in vitro* submitochondrial particle (SMP) assay (3–6). This assay exposes processed mammalian mitochondria to samples of interest and spectrophotometrically monitors the concentration of NADH to assess the level of "toxicity" exerted by contaminants present in the test solution. SMPs are prepared from bovine heart mitochondria according to the procedure detailed in Hansen and Smith (7). Briefly, the patented SMP production process involves sonically disrupting isolated mitochondria and separating the inner and outer membranes by differential centrifugation. The inner-membrane isolates retain more than 60 fully functional interacting enzymes associated with cellular electron transport and oxidative phosphorylation. Since the test system is in direct contact with test material, there are no physical barriers interfering with the interaction between the test material and receptors as there are in whole organism tests, thereby creating a potentially ultrasensitive assay.

There are two distinct and functional complementary assays utilizing SMPs. The electron transfer (ETr) assay assesses forward electron transfer through the entire electron transport chain. The test endpoint is the loss of NADH monitored as a decrease in absorbance at 340 nm in a standard spectrophotometer using 1-cm path length disposable methyl acrylate cuvettes. The reverse electron transfer (RET) assay assesses effects on electron transfer and energy coupling processes. In this test, energy derived from the hydrolysis of ATP drives electrons in the reverse direction, up the electron transport chain. The test endpoint is the production of NADH also monitored at 340 nm on a standard spectrophotometer. Sample preparation and absorbance measurement patterns for the RET assay parallel that described for the ETr assay.

Processed SMP and associated reagents for conducting tests are commercially available. Reagents for conducting the assays include SMP diluent for diluting particles to 1.5 mg/mL; RET and ETr concentrated reaction mixtures (CRM) containing buffer, electron donors (RET CRM only), and cofactors required by the SMP, adenosine-5-triphosphate (ATP) for activating the RET reaction, and ß-nicotinamide adenine dinucleotide sodium salt, reduced

Table 1. Final Concentrations of Particles, Activating Agents, and Concentrated Reaction Mixture (CRM) Components for the Reverse Electron Transport (RET) and Electron Transport (ETr) Submitochondrial Particle (SMP) Assays

SMP	RET	ETr
Particles	0.05 mg/mL	0.0167 mg/mL
Activating agents	3.33 mM ATP	0.137 mM NADH
CRM components	50 mM HEPES (pH 7.5) 6 mM Mg^{2+} 5 mM K^+ succinate 1 mM NAD^+	50 mM HEPES (pH 7.5) 6 mM Mg^{2+}

form (NADH $2Na^+$) for activating the ETr reaction. Unpublished data generated in the author's laboratory has revealed that particles held at −80 °C for up to six months remain fully functional. In contrast, some vials of particles held at −20 °C for that same period of time exhibit a reduction to complete loss of activity that is not directly related to storage duration. The final concentrations of particles and reagents in the two assays are provided in Table 1.

The objective of an individual test is to determine the median effective concentration (EC_{50}): the concentration at which the rate of NADH loss (ETr) or production (RET) is twofold the rate observed in the control treatment. Individual rates of NADH loss or production (as expressed by the slopes of the lines-of-best-fit for each treatment) are determined by simple linear regression analysis for the relationship between time and the mean NADH concentration between replicates within a treatment. Individual slopes are then used to calculate percent inhibition values relative to the control treatment for each toxicant treatment. The EC_{50} for an individual test is determined from the linear portion of the regression of individual treatment percent inhibition values versus concentration. Determination of linearity (i.e., which points to include in the generation of the line-of-best-fit) is achieved through both qualitative and quantitative criteria. Plots are visually assessed to identify a linear portion to the curve and supplemented with the generation of coefficient of determination (R^2) values to confirm the most consistent relationship among points to include in the lines-of-best-fit. The statistical analyses can be conveniently performed by routine spreadsheet software.

EXISTING LITERATURE REVIEW

Published studies assessing the sensitivity of SMPs to test materials can be segregated into two broad classes—those dealing with pure chemicals and those dealing with environmental samples. Studies focused on assessing the toxicity of pure chemicals did not directly relate to the use of SMPs for environmental monitoring. Rather, these studies for the most part simply attempted to demonstrate the potential for use of SMPs for biomonitoring by invoking the limitations associated with existing methods and comparing the sensitivity of the SMP assay with other more established tests. For example, Bragadin and Dell'Antone (8) cited the cost and duration of acute fish tests as motivation for the development of alternatives such as the SMP assay. They reported the EC_{50} values for a few dozen toxic compounds using a modified approach to the assay described above and concluded that the results from their assay compared favorably with published values of toxicity for other species that included rainbow trout and the commercial bacterial luminescent assay, Microtox®.

Similar approaches were adopted using the RET variant by Argese et al. (9) with linear alkylbenzene sulfonates, nonylphenol polyethoxylates, and their derivatives; by Argese et al. (10) with chlorophenols; by Argese et al. (11) with heavy metals; and by Argese et al. (12) with chloroanilines. Each study compared their in-house generated toxicity values from the RET variant of the SMP assay with published literature values for other species, and in general concluded there were significant levels of correlation with many of the existing assays. Researchers associated with the development of the SMP assays also published their own results with a substantial set of diverse chemicals ($n = 162$) tested with one of the SMP assay variants while comparing toxicity values with published endpoints for other species (13). Their conclusions mirrored those of other researchers in this field while they also summarized other current issues related to the state of the SMP assays.

A second group of studies also reported relationships between toxicities of test materials (both pure chemicals and environmental samples) generated through the SMP assays and other species but relied on data generated within their own laboratories rather than from published literature values. This approach tends to be more reliable since each investigator has control over test variables associated with each assay conducted within their lab. Data drawn from multiple published studies containing data generated by a variety of individuals is inherently more variable due to inconsistencies in test protocols and investigator techniques from one laboratory to another.

Miana et al. (14) compared the toxicity of tributyltin to the green alga *Selenastrum capricornutum* and the Cladoceran *Daphnia magna* to submitochondrial particles and concluded that SMP assay was fast, reproducible, and easily handled. Bettermann et al. (15) assessed the toxicity of elutriate samples from sediments previously demonstrated to be toxic to the amphipod *Hyalella azteca*. They reported a significant correlation between ETr EC_{50} values and amphipod survival for two of three watershed sites. No such relationship was observed between results from the RET assay and sediment elutriates from any of the watersheds. Dutka et al. (16) used a multispecies test battery to assess water and sediment toxicity from two river basins in Chile. The SMP RET assay was among those suggestive of toxicity in several of the samples collected during this study.

Weideborg et al. (17) compared the RET variant with six other routine toxicity tests to assess the toxicity of 82 soluble and non-water-soluble chemicals. They reported similar EC_{50} values for both the RET variant and Microtox tests for low toxicity chemicals, with Microtox being the more sensitive of the two. However, the relationships among all of the tests for a diverse set of test substances was variable and may have been a function of test design rather than test sensitivity. Sherry et al. (18) assessed the toxicity of refinery effluents with Microtox, a variety of algae, invertebrate, and fish tests, and the RET variant of the SMP assay. They reported that acute toxicity was detected only by the Microtox and SMP tests. Argese et al. (19) compared the RET assay with Microtox for their sensitivity to 14 organotin compounds. They reported good agreement between the two assays with comparable sensitivity for organotins with an R^2 of 0.92. Sheesley et al. (20) used the RET variant, the invertebrate *Ceriodaphnia dubia*, and the green alga *S. capricornutum* to assess the toxicity of aqueous and organic solvent extracts of atmospheric particulate matter from the Lake Michigan airshed. Based on a direct comparison of LC_{50} and EC_{50} values, the RET SMP assay was the most sensitive of the three for aqueous extracts of particulate matter from three different sites.

ENVIRONMENTAL MONITORING WITH SMPs

As outlined above, studies assessing the potential for SMPs as a routine biomonitor generally attempted to compare the sensitivity of one or more of the SMP assay variants with that of an established assay. Investigators either ranked assays by comparing each assay's sensitivity to an individual chemical or environmental sample or generated regression analyses for pairs of assays over a range of toxicities from groups of test substances. However, while sensitivity should be a factor in the selection of surrogate assays, it is not the only parameter that needs to be considered.

Biological monitoring, as taken in the current context, was instituted for protection of natural aquatic systems, not wildlife or human health protection. To that end, the initial species selected for that task were considered representative of the aquatic environment. Consequently, the most appropriate surrogate species or assay for that purpose would be one that is equally, but not overly, sensitive to a broad range of chemicals as the primary biomonitoring species while incorporating advantages not provided by the primary species. Such advantages might include reduced test duration time or reduced expense associated with the conductance of those tests. Adopting a surrogate assay that exhibits increased sensitivity relative to the primary test species may prompt expensive remedial actions that are not necessary for the protection of the system under study. SMP assays have exhibited the potential to be among the most sensitive assays conducted with a group of chemicals or environmental samples.

One of the confounding issues related to sensitivity rankings is the change in rank among assays for different types of contaminants. The SMP assay variants are not immune to this type of variability as demonstrated by some of the studies cited above. The solution to this dilemma involves the development of multispecies test batteries with the components of those batteries exhibiting alternating sensitivities to varying chemical classes. Positive (toxic) responses by one or more components within the test battery should then be taken as an indication of contamination. SMP assays have exhibited the potential for being a constructive component of a multispecies or multiassay test battery.

As noted above, any assay adopted for biomonitoring purposes needs to be standardized so that results from different groups of investigators are compatible and comparable. In this regard, the SMP assays are only in a rudimentary stage of development. Manual test protocols only include an approach for generating EC_{50} values involving serial dilution of test solutions. There are no formalized screening procedures for whole environmental samples or order-of-magnitude dilutions for chemical samples. Additionally, the data analysis approach currently in use involves a subjective selection of data points used in the generation of an EC_{50}. This aspect of the assay will need to be evaluated so as to minimize differences in the test endpoints not necessarily related to the toxicity of the test material itself.

Included within the standardization process is the reproducibility of an assay. There have been no studies addressing between-laboratory reproducibility for either of the SMP assay variants. The only published study related to this point focused on the within-laboratory repeatability of the ETr and RET assay variants (21). They reported coefficients-of-variation that were comparable to more traditional assays using three different reference toxicants. Since the level of variability exhibited by an assay increases in between-laboratory round robin studies in comparison with in-laboratory results, such a study is necessary to further consider SMP assays as surrogates for traditional aquatic toxicity tests.

Lastly, SMPs are a manufactured test system in contrast with existing traditional assays that use whole organisms. A great deal of research has been invested in developing culture methods for standard test species so that tests are conducted with consistently healthy test organisms. There are no published studies available assessing variabilities associated with the manufacturing process aside from a preliminary study conducted within the author's laboratory.

CONCLUSION

The SMP assay variants have exhibited the potential to serve as surrogate biomonitoring assays in freshwater systems. Test results have demonstrated that the one or both variants are as sensitive or more sensitive than traditional assays. However, SMP assays are only in a rudimentary stage of development with respect to biomonitoring applications. Both the manufacturing and testing protocols need to be standardized and evaluated prior to serious consideration of SMP assays as a legitimate surrogate biomonitoring option.

BIBLIOGRAPHY

1. Weber, C.I. (1991). *Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms*, 4th Edn. Environmental Monitoring Systems Laboratory, U.S. Environmental Protection Agency, Cincinnati, OH, EPA/600/4-90/027.

2. Lewis, P.A., Klemm, D.J., Lazorchak, J.M., Nordberg-King, T.J., Peltier, W.H., and Heber, M.A. (1994). *Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms*, 3rd Edn. Environmental Monitoring Systems Laboratory, U.S. Environmental Protection Agency, Cincinnati, OH, EPA/600/4-91/002.
3. Blondin, G.A., Knobeloch, L.M., Read, H.W., and Harkin, J.M. (1987). Mammalian mitochondria as *in vitro* monitors of water quality. *Bull. Environ. Contam. Toxicol.* **38**: 467–474.
4. Blondin, G.A., Knobeloch, L.M., Read, H.W., and Harkin, J.M. (1989). An *in vitro* submitochondrial bioassay for predicting acute toxicity in fish. In: *Aquatic Toxicology and Environmental Fate: Eleventh Volume, ASTM STP 1007*. G.W. Suter, II and M.A. Lewis (Eds.). American Society for Testing and Materials, Philadelphia, pp. 551–563.
5. Knobeloch, L.M., Blondin, G.A., Lyford, S.B., and Harkin, J.M. (1990). A rapid bioassay for chemicals that induce prooxidant states. *J. Appl. Toxicol.* **10**: 1–5.
6. Knobeloch, L.M., Blondin, G.A., Read, H.W., and Harkin, J.M. (1990). Assessment of chemical toxicity using mammalian mitochondrial electron transport particles. *Arch. Environ. Contam. Toxicol.* **19**: 828–835.
7. Hansen, M. and Smith, A.L. (1964). Studies on the mechanism of oxidative phosphorylation. VII. Preparation of a submitochondrial (ETP_H), which is capable of fully, coupled oxidative phosphorylation. *Biochim. Biophys. Acta* **81**: 214–222.
8. Bragadin, M. and Dell'Antone, P. (1994). A new *in vitro* toxicity test based on the response to toxic substances in solutions of mitochondria from beef heart. *Arch. Environ. Contam. Toxicol.* **27**: 410–414.
9. Argese, E. et al. (1994). Submitochondrial particle response to linear alkylbenzene sulfonates, nonylphenol polyethoxylates and their biodegradation derivatives. *Environ. Toxicol. Chem.* **13**: 737–742.
10. Argese, E. et al. (1995). Submitochondrial particles as toxicity biosensors of chlorophenols. *Environ. Toxicol. Chem.* **14**: 363–368.
11. Argese, E. et al. (1996). Submitochondrial particles as *in vitro* biosensors of heavy metal toxicity. *J. Aquat. Ecosys. Health* **5**: 125–134.
12. Argese, E. et al. (2001). Assessment of chloroaniline toxicity by the submitochondrial particle assay. *Environ. Toxicol. Chem.* **20**: 826–832.
13. Read, H.W., Harkin, J.M., and Gustavson, K.E. (1998). Environmental applications with submitochondrial particles. In: *Microscale Testing in Aquatic Toxicology*. P.G. Wells, K. Lee, and C. Blaise (Eds.). CRC Press, Boca Raton, FL, pp. 31–52.
14. Miana, P., Scotto, S., Perin, G., and Argese, E. (1993). Sensitivity of *Selenastrum capricornutum, Daphnia magna* and submitochondrial particles to tributyltin. *Environ. Technol.* **14**: 175–181.
15. Bettermann, A.D., Lazorchak, J.M., and Dorofi, J.C. (1996). Profile of toxic response to sediments using whole-animal and *in vitro* submitochondrial particle (SMP) assays. *Environ. Toxicol. Chem.* **15**: 319–324.
16. Dutka, B.J. et al. (1996). Water and sediment ecotoxicity studies in Temuco and Rapel River Basin, Chile. *Environ. Toxicol. Water Qual.* **11**: 237–247.
17. Weideborg, M., Arctander Vik, E., Daae ϕfjord, G., and Kjϕnnϕ, O. (1997). Comparison of three marine screening tests and four Oslo and Paris Commission procedures to evaluate toxicity of offshore chemicals. *Environ. Toxicol. Chem.* **16**: 384–389.
18. Sherry, J., Scott, B., and Dutka, B. (1997). Use of various acute, sublethal and early life-stage tests to evaluate the toxicity of refinery effluents. *Environ. Toxicol. Chem.* **16**: 2249–2257.
19. Argese, E. et al. (1998). Comparison of *in vitro* submitochondrial particle and Microtox® assays for determining the toxicity of organotin compounds. *Environ. Toxicol. Chem.* **17**: 1005–1012.
20. Sheesley, R.J. et al. (2004). Toxicity of ambient atmospheric particulate matter from the Lake Michigan (USA) airshed to aquatic organisms. *Environ. Toxicol. Chem.* **23**: 133–140.
21. Doherty, F.G. and Gustavson, K.E. (2002). Repeatability of the submitochondrial particle assay. *Ecotox. Environ. Safety* **53**: 122–128.

MICROSCALE TEST RELATIONSHIPS TO RESPONSES TO TOXICANTS IN NATURAL SYSTEMS

JOHN CAIRNS, JR.
Virginia Polytechnic Institute
and State University
Blacksburg, Virginia

A primary purpose of ecotoxicology is to estimate toxicity thresholds below which no observable deleterious effects can be detected upon ecosystems. Since uncertainty is unacceptably high when extrapolations are made from single species (1), tests at higher levels of biological organization are much more effective (2). Moreover, such levels of testing permit increased opportunities of testing for alteration of functional attributes (3). Large-scale tests are both risky and expensive; consequently, microscale testing has been developed.

Microscale testing can be carried out with standardized communities, which are superb in furnishing a validated set of results. However, standardized communities do not closely resemble communities from natural systems. Communities from any two natural systems rarely contain the same species. However, any naturally derived community will contain species that span a range of tolerance to stress. This range of tolerance can be generalized from one natural community to another. This extrapolation provides predictive capabilities, which are the ultimate goal of ecotoxicological testing. Furthermore, culturing test organisms for use in microscale testing customarily consumes as much or more time than the testing itself.

Microscale toxicity testing with communities obtained from natural systems has many benefits, in addition to reducing culture problems typical of laboratory-assembled communities. Methods are even available for carrying out microscale toxicity testing in natural systems (4). Communities of organisms that have colonized artificial substrates can be collected. These collections are not artificial communities, although convenience sometimes dictates naming these communities "artificial" in order to distinguish them from collections made directly from natural substrates. Artificial substrates (5) have major advantages in microscale testing.

1. The age of the community is known.
2. A determination can be made of when the community has reached a dynamic equilibrium comparable to that of the system that furnished the colonizing organisms.
3. Natural systems are not disturbed by sampling efforts.
4. Communities can be moved to the microscale test system with minimal disturbance to the community that is the subject of the test.
5. Communities in the natural system can be monitored to determine if changes in either structure or function have occurred that were evident in the microscale system as a result of toxic stress.
6. Since microbial species have a cosmopolitan distribution, results from different geographic regions can be compared.
7. Since microorganisms and macroinvertebrates are much more abundant than larger organisms higher in the food chain, their use in toxicity testing is more likely to be accepted by animal rights groups.
8. Microscale tests allow a setup of many more concentrations in the same space that would be necessary for far fewer larger scale tests.
9. Automation is possible for the counting of numbers of individuals per species and total number of some species used in microscale toxicity tests (6).

Microscale testing units are not miniature ecosystems. Properly designed, they can mimic selected cause–effect pathways (e.g., nutrient cycling and prey relationships) (7) not observable at lower levels of biological complexity. However, they do not have all the cause–effect pathways found in natural ecosystems and landscapes. Some relevant discussions of complex system testing (8) and examples of microscale testing (9,10) are available in the literature. Suitability of the microscale test design can be confirmed by comparing measurements made in it to similar measurements in natural systems (11). Methods are available for the statistical analysis of this unique evidence (12–14). One of the most difficult attributes to include in microscale testing is the invasive pressure of organisms that are attempting to colonize (12). This invasion pressure is one of the major factors in succession of species in natural systems.

The major disadvantage to microscale toxicity testing is that sometimes a modest degree of taxonomic ability is essential. Taxonomy is often not taught in many American institutions of higher learning. However, staff members with no taxonomic ability can be instructed in just a few months in levels of identification necessary in some microscale tests, although years of experience are necessary for the more complex microscale toxicity tests.

BIBLIOGRAPHY

1. Cairns, J. Jr. (1983). Are single species toxicity tests alone adequate for estimating environmental hazard? *Hydrobiologia* **100**: 47–57.
2. Cairns, J. Jr. (Ed.) (1986). *Community Toxicity Testing*, STP920. American Society for Testing and Materials, Philadelphia.
3. Cairns, J. Jr. and Pratt, J.R. (Eds.) (1989). *Functional Testing of Aquatic Biota for Estimating Hazards of Chemicals*, STP988. American Society for Testing and Materials, Philadelphia.
4. Arnegard, M.E., McCormick, P.V., and Cairns, J. Jr. (1998). Effects of copper on periphyton communities assessed *in situ*using chemical-diffusing substrates. *Hydrobiologia* **385**: 163–170.
5. Cairns, J. Jr. (1982). *Artificial Substrates*. Ann Arbor Science Publishers, Ann Arbor, MI.
6. Cairns, J. Jr. and Niederlehner, B.R. (Eds.) (1995). *Ecological Toxicity Testing: Scale, Complexity and Relevance*. Lewis Publishers, Chelsea, MI.
7. Clements, W.H., Cherry, D.S., and Cairns, J. Jr. (1989). The influence of copper exposure on predator–prey interactions in aquatic insect communities. *Freshwater Biol.* **21**: 483–488.
8. Hoffman, D.J., Rattner, B.A., Burton, G.A. Jr., and Cairns, J. Jr. (2003). *Handbook of Ecotoxicology*, 2nd Edn. Lewis Publishers/CRC Press, Boca Raton, FL.
9. McCormick, P.V., Belanger, S.E., and Cairns, J. Jr. (1997). Evaluating the hazard of dodecyl alkyl sulphate to natural ecosystems using indigenous protistan communities. *Ecotoxicology* **6**: 67–85.
10. Orvos, D.R., Scanferlato, V.S., Lacy, G.H., and Cairns, J. Jr. (1989). Fate and effects of genetically-engineered *Erwinia carotovora* in terrestrial and aquatic microcosms. *89th Annual Meeting of the American Society for Microbiology*, New Orleans, Louisiana, p. 355.
11. Niederlehner, B.R. and Cairns, J. Jr. (1994). Consistency and sensitivity of community level endpoints in microcosm tests. *J. Aquat. Ecosys. Health* **3**: 93–99.
12. Pontasch, K.W. and Cairns, J. Jr. (1989). Establishing and maintaining laboratory-based microcosms of riffle insect communities: their potential for multispecies toxicity tests. *Hydrobiologia* **175**: 49–60.
13. Pontasch, K.W., Smith, E.P., and Cairns, J. Jr. (1989). Diversity indices, community comparison indices and canonical discriminant analysis: interpreting the results of multispecies toxicity tests. *Water Res.* **23**: 1229–1238.
14. Smith, E.P. and Cairns, J. Jr. (1993). Extrapolation methods for setting ecological standards for water quality: statistical and ecological concerns. *Ecotoxicology* **2**: 203–219.

TOXICITY IDENTIFICATION EVALUATION

Philip J. Markle
Whittier, California

Bioassays for aquatic chronic and acute toxicity testing using standardized methods have been used to measure toxic effects and in some cases can even estimate the magnitude of toxicity for many years. Their use has been incorporated into regulatory requirements for point and nonpoint discharges to both marine and freshwater environments. However, these standard methods provide virtually no information on which constituent(s) are responsible for the observed effect. Reliance on only chemical-specific analyses to determine the causative agent(s) in a water sample exhibiting toxicity or suspected of being toxic to aquatic organisms can be misleading. Initially, the list of potential toxicants that would need to be quantified is nearly limitless, not to mention costly.

Additionally, even GC/MS analyses may not be sensitive enough to detect low concentrations of some toxic substances. Furthermore, even if elevated levels of a particular compound are identified, the compound may or may not be bioavailable, and receiving water characteristics (i.e., total organic carbon, hardness, salinity, pH) may confound the observed toxicological sensitivity. Chemical quantification will also not identify any additive, synergistic, or antagonistic properties of many toxicant combinations. It was specifically because of these concerns that aquatic toxicity testing was developed and incorporated into many water pollution control plans.

The U.S. EPA has developed and published several manuals detailing specific methods for identifying toxic constituents (1–4). These methods are collectively referred to as "toxicity identification evaluations" (TIEs) and are often incorporated as part of a larger site-specific study identified as a "toxicity reduction evaluation" (TRE). The ultimate goal of a TRE is to identify toxic constituents, identify their sources, evaluate toxicity control options, and eventually eliminate the toxicity. The TIE is typically used to identify the toxic constituents and to evaluate the effectiveness of the toxicity control mechanisms. During the initial stages of a TIE (phase 1), toxicity is characterized through a series of sample manipulations followed by toxicity testing. These manipulations typically include pH adjustments, chemical additions (EDTA and sodium thiosulfate), chemical extractions (solid phase extraction), and physical manipulations, (filtration and aeration) followed by toxicity testing to assess any increases or decreases resulting from the manipulation.

The relative increases or decreases in observed toxicity in the subsequent toxicity tests conducted on the manipulated sample often characterize the properties of the chemical constituent(s). For example, a decrease in toxicity observed in the EDTA addition tests would indicate possible metal toxicity. A list of commonly applied phase 1 TIE manipulations and their interpretation is contained in Table 1. Other manipulations and result interpretations are specifically detailed in the U.S. EPA manuals. Once characterized in phase 1, the investigator would proceed to phase 2 to identify the individual constituent(s). The phase 2 TIE procedures incorporate additional toxicity testing and/or analytical methods. The testing and/or analytical measurements vary, depending on the phase 1 characterization interpretations. The last phase of the TIE procedure (phase 3) consists of a series of final confirmation steps designed to provide a "weight of evidence" approach in confirming the causative agent(s). Through the establishment of causal links and correlations, the approach is designed to ensure that suspected toxicants are conclusively identified and consistent from sample to sample.

Table 1. Summary of Commonly Applied Phase 1 TIE Manipulations and Associated Interpretation

TIE Manipulation	Purpose/Interpretation
Baseline—no manipulation	Conducted to determine if toxicity is present
EDTA addition	Controls most metal (cationic) toxicity
Sodium thiosulfate addition	Reduces easily oxidizable compounds such as chlorine or some metals
Piperonyl butoxide addition	Controls toxicity due to organophosphate pesticides An increase in toxicity would indicate possible pyrethroid pesticide toxicity
pH control at 6.5	Controls ammonia/metal toxicity A decrease in toxicity may indicate ammonia/metal toxicity
pH control at 8.5	Controls some metal toxicity An increase in toxicity would indicate ammonia and/or some metal toxicity
Aeration	Removes volatile compounds and some surfactants
Filtration	Removes solids and some surfactants
Solid phase extraction (SPE)	Removes nonpolar organic compounds and some surfactants Recovery of toxicity from the SPE column indicates nonpolar organic toxicity Failure to recover toxicity from the SPE column indicates surfactant toxicity

Identification of what is "toxic" in aquatic toxicity tests is most often determined quantitatively through the regulatory process as permit limits or objectives and no definitive standards have been defined or are universally used. Although no universal definitions exist, several measures of toxicity are more commonly employed and seem to have achieved greater acceptance. These include the use of the "no observed effect concentration" (NOEC) or "no observed effect level" (NOEL). The NOEC and NOEL use hypothesis testing to identify statistically significant differences relative to a nontoxic control. The NOEC or NOEL is defined as the highest concentration not significantly different from the control. Another commonly used analysis involves various point estimation techniques. These quantitative analyses use dose–response data to estimate "safe" or "nontoxic" concentrations or effects using metrics such as the LC_{50} (lethal concentration causing a 50% effect) or EC/IC_{25} (effective or inhibition concentration causing a 25% effect). Due to the large amount of toxicity testing required in the TIE process, reductions in testing replication and exposure volume and duration are often made to save on sample and staffing resources. Results are typically represented qualitatively as opposed to quantitatively that is, relative increases or decreases in toxicity compared to the no manipulation treatment or appropriately manipulated control were observed. Although fairly straightforward, several things can significantly complicate the TIE process. To identify the causative agent, a sample must first exhibit toxicity. For this reason, the persistence of toxicity is often a confounding factor. Multiple samples eventually need to be collected and tested during the course of the analysis. If toxicity is episodic or sporadic, collection and subsequent testing of a suitably toxic sample will also be more unlikely. Most TIEs are conducted during the course of 1 to 2 weeks requiring extended holding of samples. Constituents that rapidly degrade during this extended holding time can often confound identification due to a loss of toxicity. Furthermore, the numerous manipulations and increased holding time necessary in conducting a TIE may also alter basic water quality characteristics that may artificially affect the observed toxicity. For these reasons, successful completion of a TIE is much more likely in samples exhibiting a large amount of toxicity. Another common complication involves multiple constituents. If a sample contains two or more toxic constituents, the investigator will need to identify at least one of the constituents and selectively remove or control its toxicity without impacting the others prior to attempting to identify the other constituent(s).

Although much more work conducted using freshwater organisms has resulted in greater acceptance and standardization in the freshwater methods, the described TIE procedures have been successfully used to identify and ultimately control toxic compounds in effluents and freshwater and marine receiving waters. It would be impractical to detail all of the possible TIE manipulations and their interpretation in this overview. The referenced manuals provide specific details on all aspects of the conduct and interpretation of TIE results. Copies of the referenced manuals may be obtained directly from the U.S. EPA (http://www.epa.gov/waterscience/WET/).

BIBLIOGRAPHY

1. EPA/600/6-91/005F. (1992). *Toxicity Identification: Characterization of Chronically Toxic Effluents, Phase 1*. U.S. EPA Office of Research and Development, Washington, DC.
2. EPA/600/6-91/003. (1991). *Methods for Aquatic Toxicity Identification Evaluations-Phase 1 Toxicity Characterization Procedures*, 2nd Edn. U.S. EPA Office of Research and Development, Washington, DC.
3. EPA/600R-92/080. (1993). *Methods for Aquatic Toxicity Identification Evaluations-Phase 2 Toxicity Identification Procedures for Samples Exhibiting Acute and Chronic Toxicity*. U.S. EPA Office of Research and Development, Washington, DC.
4. EPA/600/3-88/036. (1989). *Methods for Aquatic Toxicity Identification Evaluations-Phase III Toxicity Confirmation Procedures*. U.S. EPA Environmental Research Laboratory, Duluth, MN.

WHOLE EFFLUENT TOXICITY CONTROLS

Philip J. Markle
Whittier, California

The selection of an appropriate control sample/water for whole effluent toxicity testing is a critical experimental design component. Whole effluent toxicity (WET) testing has been an important tool for toxicity compliance and research testing alike to identify acute and chronic toxicity in effluents and receiving waters. The U.S. EPA and others have developed several manuals detailing specific WET methods; many have been incorporated into the National Pollutant Discharge Elimination System (NPDES) permit program for compliance determination. The primary objective of these aquatic toxicity tests is to estimate a maximum "safe" concentration of toxic substances in a discharge or receiving water that will allow for normal growth and propagation of aquatic organisms. A typical WET test experimental design consists of exposing organisms under controlled laboratory conditions to a series of effluent, receiving water, and/or toxicant dosed samples and recording biological observations such as reproduction, growth, and survival.

A set of exposures is concurrently conducted using control water. The biological results obtained from the effluent/receiving water exposures are then qualitatively or quantitatively compared to the results obtained in the control water exposure to identify relative reductions in biological response as an indication of toxicity. Effects relative to the control are used to estimate toxicity in a WET test, so control water selection can be critical. Control water that stimulates an unusually large response or enhancement when compared to a more normal response in the effluent/receiving water is more likely to identify falsely that a sample is "toxic," when it was not. Likewise, control water that exhibits a suppressed response when compared to an effluent/receiving water that exhibits toxicity is more likely to be identified as nontoxic when in fact it was toxic. Nontoxicity related factors associated with control waters can elicit a stimulatory or inhibitory response in test organisms. These factors need to be identified by the investigator and carefully controlled to minimize any misidentification of toxicity or nontoxicity in a WET test.

In multiconcentration tests, the same control water is also used to dilute the effluent or receiving water to the appropriate concentration. For this reason, control water is often referred to as *dilution water*. Although the amount of sample increases as concentration increases, the amount of dilution water likewise decreases as concentration increases. The U.S.EPA whole effluent toxicity testing protocols define acceptable dilution water as water that is appropriate for the test objectives, supports adequate test organism response measured in the test, is consistent in quality, and is free of contaminants that could cause toxicity (1–3). Although test organisms may have a wide range of tolerance to many commonly measured water characteristics (pH, hardness, conductivity, salinity, etc.), it is likely that the organisms have a much narrower "optimal" range. Differences in these basic water characteristics between control water and effluent samples may account for an apparent observed difference in toxicological response in a WET test when in reality, the sample was actually nontoxic. Furthermore, control water/effluent sample interactions in multiple concentration tests could increase or decrease apparent toxicity and result in improper identification of toxicity.

Selection of an appropriate control water will depend in some part on the objectives of the experiment. If the primary objective of the test is to quantify the relative toxicity of a sample over time or location (i.e., is sample "A" more or less toxic than sample

"B," or is discharge "A" more toxic this month than it was last month), synthetic control water would be most appropriate. Dilution water characteristics such as hardness, pH, magnesium/calcium ratios, and many others can drastically effect the toxicological response in a test. It is impossible to understand or predict all of the dilution water/sample interactions that may effect the toxicological response, so the incorporation of standard synthetic control water allows the investigator to eliminate these concerns when the objective of toxicity testing is to measure relative toxicity. Synthetic dilution water is prepared by dissolving fixed amounts of reagent grade salts into high purity de-ionized water. For synthetic freshwater, $NaHCO_3$, $CaSO_4$, $MgSO_4$, and KCl are the most commonly incorporated reagent salts. Commercially available seawater reagents are available for testing with estuarine and marine species. The referenced U.S. EPA method protocols provide detailed preparatory procedures for many of these waters. The advantages of synthetic waters are that they are very consistent from batch to batch, would not be expected to contain toxic constituents, and are easily prepared. The disadvantages of synthetic waters are that they are typically nutrient poor and although adequate, may not be "optimal" water for the test organisms.

If the objective of the test is to estimate or predict ecological impacts on the receiving water, control water with similar receiving water characteristics should be used. The more similar the control water characteristics are to the receiving water, the more predictive the toxicity test will be. Characteristics most commonly simulated include pH, hardness, alkalinity, and conductivity. However, differences between receiving water and control water in other less understood characteristics such as total organic carbon and micronutrient concentration, to name only a few, may significantly compromise the predictive ability of the toxicity test and may require further investigation. An immediate upstream receiving water not influenced by the discharge may be a suitable source. The advantages to using receiving water is that sample/control water interactions are accounted for in the toxicity test design with little effort and surface waters tend to be higher in micronutrient content than synthetic waters. Disadvantages of using receiving water include the possibility that they may contain toxic constituents, they may contain pathogens or predatory organisms, water quality characteristics are not typically consistent over time, and upstream receiving water may not be available.

Additionally, any manipulations of a sample prior to or during a toxicity test should also be incorporated into the control/dilution water. For instance, if a sample requires aeration to maintain appropriate dissolved oxygen levels, the control should be similarly aerated to compensate for any "artifactual" increase or decrease in toxicological response that may occur from the manipulation. This becomes increasingly more important when conducting toxicity identification evaluations (TIEs). In a TIE, toxic effluent or receiving waters are manipulated to isolate or remove specific compounds or classes of compounds. Many of these manipulations involve chemical extractions and/or the addition of substances that could affect the response. Appropriately manipulated control water aliquots must also be tested and compared to identify any potentially stimulatory or inhibitory effects.

BIBLIOGRAPHY

1. U.S. Environmental Protection Agency. (2002). *Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms*, 4th Edn. EPA/821/R-02/013. Office of Water, Washington, DC.
2. U.S. Environmental Protection Agency. (2002). *Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms*. 3rd Edn. EPA/821/R-02/014. Office of Water, Washington, DC.
3. U.S. Environmental Protection Agency. (2002). *Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms*, 5th Edn. EPA/821/R-02/012. Office of Water, Washington, DC.

DEVELOPMENT AND APPLICATION OF SEDIMENT TOXICITY TESTS FOR REGULATORY PURPOSES

M.G.J Hartl
F.N.A.M. van Pelt
J. O'Halloran
Environmental Research Institute
University College Cork,
Ireland

INTRODUCTION

Traditional approaches to sediment toxicity assessment have employed chemical analysis to identify and quantify pollutants present. This approach, however, will only provide information on chemical classes that are analyzed and, when used alone, is of little value in ecotoxicological assessment, because toxicity cannot be determined on the basis of chemistry alone (1). Toxicity is ultimately defined as a measurable biological response to a particular substance or mixture of substances (2,3). Toxicity testing provides a more direct means of assessing the potential adverse effects of contaminants. In a complementary way, ecotoxicological assessment should provide a measure of the combined effects of the compounds in a complex sample, thereby taking into account any additive, antagonistic, or synergistic effects and include a degree of biological relevance.

An ecotoxicological sediment assessment necessitates a tiered approach using different endpoints and several test species representing different trophic levels, because the effect of pollutants may differ between species. Thus, a battery of bioassays rather than single species assays should be employed. Battery style approaches in the evaluation of sediment toxicity (both freshwater and marine) have been described (2,4–8). For example, the SED-TOX Index recommends the integration of multitrophic and multiexposure route tests (different sediment phases) in toxicity assessment of sediments (8). The ecotoxicological triad explores a similar concept (9). In this article, we define sediments, consider their ecotoxicological significance, and summarize some of the key sediment toxicity test systems in use and/or development.

WHAT ARE AQUATIC SEDIMENTS?

Sediments represent an open, dynamic, and heterogeneous biogeochemical system (10) that is formed by an accumulation of particulate matter derived from continental runoff, coastal erosion, or atmospheric deposition, which precipitates to the bottom of a water body. Typically, sediments are an accumulation of particulate mineral matter, inorganic matter of biogenic origin, organic matter in various stages of synthesis or decomposition, and water (11,12). Sediments normally consist of an inorganic matrix coated with organic matter (13), giving rise to a wide variety of physical, chemical, and biological characteristics. Control sediments can be formulated from particulate matter of known origin and characteristics for use in toxicity testing (12,14,15).

The ecotoxicological significance lies in the tendency of many pollutants, especially the less polar organic contaminants and trace elements, to show a strong affinity to suspended particulate matter (16,17). They are sequestered from the water column and incorporated into the sediment. Redox conditions influence the chemical speciation, sorption behavior, and partition coefficients of incorporated compounds and trace elements. Undisturbed sediments tend to accumulate many chemical compounds, and so act as sinks. The retention capacity of sediments for many pollutants is dependent on salinity, pH, E_{h}, and/or mechanical disturbance. Changes in these conditions can result in release of the contaminants, and therefore sediments may act not only as sinks but also as secondary sources, directing often highly concentrated pulses of pollutants at benthic organisms (i.e., organisms intimately associated with sediments). Fine-grained, organically rich sediments play a major role in the biogeochemical fate of chemicals, both of natural and anthropogenic origin, and, along with water quality, have increasingly become the focus of attention in assessing the state of the aquatic environment.

In situ sediment toxicity assessments are rarely performed because of logistics, the difficulties of identifying reference and control sites, and controlling or correcting for confounding environmental variables. Thus field-collected sediment samples are used in laboratory-based toxicity test models. Sediments vary on both a spatial and temporal scale and are structured systems of oxic and anoxic zones (18). These two zones display very different chemical conditions (19,20). Accordingly, during the collection of sediment samples, one must ensure that these zones are not mixed, as this may result in differences in redox status, which will affect the bioavailability of contaminants. The oxic layer of the sediment is preferably sampled and used for toxicity testing especially because this layer interfaces with the water column *in situ* (21,22).

Sediments are not homogeneous but are composed of the following phases: whole sediment, sediment–water interface, pore water, and elutriates. Examination of any single sediment phase may be insufficient to give an accurate ecotoxicological assessment (23,24).

Recent investigations using field-collected sediment samples have demonstrated that the whole sediment phase can be used in toxicity assessment under controlled laboratory conditions (see Tables 1–4).

TEST SYSTEMS

A comprehensive assessment of potential sediment toxicity requires the consideration of multiple exposure phases and multiple test models representing different trophic levels and sediment related habitats.

Primary criteria for test species selection for assessing sediment contamination and toxicity include the species' ecological and/or economical importance and its relative sensitivity to sediment contamination, predictable and consistent response of control organisms, ease of culture and maintenance, short duration, replicable, relatively inexpensive, comparable, and ecologically relevant (6,25,26). In addition to species selection, the endpoints in sediment toxicity tests depend on the question being addressed in the environmental risk assessment (2) and may include acute and long-term toxicity, endocrine, reproductive, and genotoxic effects.

The majority of test systems in use for regulatory purposes are commercial test kits assessing acute general effects on microorganisms using sediment extracts (Table 1). Sediment pore water extracts (also known as interstitial water) are defined as the water occupying the space between sediment particles (27). Contaminants in pore water represent the water-soluble, bioavailable fraction and, as a result, may be a major route of exposure to infaunal species (28–30). The use of elutriate extracts, as opposed to pore water extracts, provides information on the leaching potential of sediment-associated contaminants and may therefore yield important data on the potential adverse effects to benthic organisms, following disturbance of the underlying sediment (6,31). Methods applicable to whole sediment, sediment suspension, sediment elutriate, pore-water extracts, and/or sediment extracts from the marine and freshwater environment have been previously reviewed (32).

These test systems are well established, validated and reproducible, fast, cheap, and require little specialized training. In addition, *in vitro* models, using cells derived from a range of taxa, are currently being developed and validated (22). All these tests are suitable for screening purposes and initial hazard identification. However, they have limited ecological relevance because they are often restricted to nonspecific endpoints or a single trophic level. Therefore, the use and development of a multiple test system or test battery, using various endpoints for both general toxicity (see Tables 1 and 2) and specific toxicities (e.g., genotoxic and reproductive effects; see Tables 3 and 4) in sediment-associated organisms from several taxa representing different trophic levels, is desirable. These models may have a higher ecological relevance than the microbial test systems. In some cases, multiple endpoints for different toxic effects (acute, long term, and specific)

Table 1. Examples of Test Systems Used for Sediment Toxicity Assessment—Acute Toxicity Tests

Endpoint	Taxa	Chemicals Identified[a]	*In vivo*	*In vitro*	Test Phase Sediment	 Extracts[b]	Reference
Enzyme inhibition/ bioluminescence	**Bacteria**						
	Microtox©	PAHs, POPs, metals		+	+	p,e	6,34–38
	ToxiChromoPad©	PAHs, POPs		+	+		38
	LUMIStox©	unspecified		+	+	e	39
	BioTox©	Pesticides	+			p,e,o	3,40
	MetPAD©	Metals		+		e	41
	Toxi-Chromotest©	PAHs, POPs		+	+	e	6,38
Behavior							
Motility	**FW diatom**		+			e	42
	Invertebrates						
	(Daphtoxkit©)	Pesticides	+			p,e,o	3,40
Reburial		Resin-acid	+		+		43

[a] Major chemical classes identified in the sediments.

[b] p: porewater extracts; e: elutriates; o: organic solvent extracts; PAHs: polycyclic aromatic hydrocarbons; POPs: persistent organic pollutants.

Table 2. Examples of Test Systems Used for Sediment Toxicity Assessment—Subchronic Toxicity Tests

Endpoint	Taxa	Chemicals Identified[a]	*In vivo*	*In vitro*	Test Phase Sediment	 Extracts[b]	Reference
Survival							
	Invertebrates						
		PAHs, POPs, metals	+		+	e, p	32,44,45
	Vertebrates						
		PAHs, POPs, metals	+			e, o, p	32,46
Growth inhibition							
	FW Algaltoxkit©	PAHs, PCBs		+		o	38
	FW microalgae		+		+		47
	Marine microalgae	PAHs, PCBs	+		+		6
Behavior							
	Invertebrates						
		Organotins, metals	+		+		48–50
	Vertebrates						
		Metals	+		+		51,52
Enzyme induction							
EROD	**Invertebrates**						
		PAHs, PCBs	+		+		53
	Vertebrates						
		PAHs, OCPs, metals	+	+	+	o	54–67

[a] Major chemical classes identified in the sediments.

[b] p: porewater extracts; e: elutriates; o: organic solvent extracts; PAHs: polycyclic aromatic hydrocarbons; POPs: persistent organic pollutants.

or single endpoints in multiple organ systems have been used in the same test species in order to observe potential toxicity on various levels of biological organization (33). Unfortunately, this type of approach is currently not accepted for regulatory purposes, because many of the bioassays involved are generally less well validated.

Therefore, we recommend that sediment toxicity assessments should be further developed and evaluated using a tiered approached, consisting of screening using short-term general toxicity tests (Tier 1); hazard identification applying more specific (multiple) endpoints in multiorganism experiments, representing different trophic levels and habitats associated with sediments, as well as different modes of bioavailability by using both sediment extracts and whole sediments (Tier 2); and *in situ* ecosystem function, for example, lifetime reproductive success, and components of biodiversity (Tier 3).

BIBLIOGRAPHY

1. Heida, H. and van der Oost, R. (1996). *Water Sci. Technol.* **34**: 109–116.
2. Chapman, P.M. et al. (2002). *Mar. Pollut. Bull.* **44**: 271–278.
3. Fernandez-Alba, A.R., Guil, M.D.H., Lopez, G.D., and Chisti, Y. (2002). *Anal. Chim. Acta* **451**: 195–202.
4. Dutka, B.J., Tuominen, T., Churchland, L., and Kwan, K.K. (1989). *Hydrobiologia* **188**: 301–315.
5. Giesy, J.P. and Hoke, R.A.J. (1989). *Gt. Lakes Res.* **15**: 539–569.

Table 3. Examples of Test Systems Used for Sediment Toxicity Assessment—Genotoxicity Tests

					Test Phase		
Endpoint	Taxa	Chemicals Identified[a]	*In vivo*	*In vitro*	Sediment	Extracts[b]	Reference
Mutation							
	Bacteria						
	AMES test	PAHs, POPs		+		o	68
	Mutatox©	PAHs, POPs		+	+	p,e,o	69
Micronucleus							
	Vertebrates						
		PAHs, POPs	+		+		70
SSBs							
	Invertebrates						
		POPs	+		+		71
		PAHs, metals	+		+		33
	Vertebrates						
		PAHs, POPs	+		+		65,70–74
DNA adducts							
	Vertebrates						
		PAHs, PCBs	+		+		59,75

[a] Major chemical classes identified in the sediments.
[b] p: porewater extracts; e: elutriates; o: organic solvent extracts; PAHs: polycyclic aromatic hydrocarbons; POPs: persistent organic pollutants.

Table 4. Examples of Test Systems Used for Sediment Toxicity Assessment—Endocrine and Reproduction Tests

					Test Phase		
Endpoint	Taxa	Chemicals Identified[a]	*In vivo*	*In vitro*	Sediment	Extracts[b]	Reference
	Invertebrates						
Imposex		Organotins	+		+		76
Larval development		PAHs, metals	+			e,o,p	32,77
Spermiotoxicity		PAHS, POPs		+		p	44,78
Emergence		Unspecified	+		+		32
	Vertebrates						
Estrogen-like activity		PAHS, POPs	+			o	79
Fertility		PAHS, POPs		+	+		80
Vitellogenin		Unspecified	+		+		81,82

[a] Major chemical classes identified in the sediments.
[b] p: porewater extracts; e: elutriates; o: organic solvent extracts; PAHs: polycyclic aromatic hydrocarbons; POPs: persistent organic pollutants.

6. Cheung, Y.H. et al. (1997). *Arch. Environ. Contam. Toxicol.* **32**: 260–267.
7. Matthiessen, P. et al. (1998). *Mar. Environ. Res.* **45**: 1–15.
8. Bombardier, M. and Bermingham, N. (1999). *Environ. Toxicol. Chem.* **18**: 685–698.
9. Chapman, P.M. (2000). *Int. J. Environ. Pollut.* **13**: 351.
10. Chapman, P., Bishay, F., Power, E., Hall, K., Harding, L., McLeay, D., Nassichuk, M., and Knapp, W. (Eds.). (1991). 17th Annual Aquatic Toxicology Workshop. In: *Canadian Technological Report in Fisheries and Aquatic Science.* Vol. 1. Vancouver, BC, pp 323–330.
11. Knezovich, J.P., Harrison, F.L., and Wilhelm, R.G. (1987). *Water Air Soil Pollut.* **32**: 233.
12. ASTM. (1994). *American Society for Testing and Materials: Terminology*. ASTM, Provo, UT.
13. Rand, G.M., Wells, P.G., and McCarty, L.S. (1997). *Fundamentals of Aquatic Toxicology—Effects, Environmenatl Fate and Risk Assessment*, 2nd Edn. G.M. Rand (Ed.). Taylor and Francis, London, pp. 3–67.
14. Suedel, B.C. and Rodgers, J.H. (1994). *Environ. Toxicol. Chem.* **13**: 1163.
15. Quevauviller, P. and Ariese, F. (2001). Trac-Trends. *Anal. Chem.* **20**: 207–218.
16. IEEE. (1987). *Organotin Symposium Oceans'86*. Vol. 4. IEEE, Washington, DC, pp. 1370–1374.
17. Ragnarsdottir, K.V.J. (2000). *Geol. Soc.* **157**: 859–876.
18. Fenchel, T. (1969). *Ophelia* **6**: 1–182.
19. Machan, R. and Ott, J.A. (1972). *Limnol. Oceanogr.* **17**: 622–626.
20. Aller, R.C. (1978). *Am. J. Sci.* **278**: 1185–1234.
21. ASTM. (1990). *Standard guide for collection, storage, characterization and manipulation of sediments for toxicity testing*. ASTM, Provo, UT.
22. Ni Shuilleabhain, S., Davoren, M., Hartl, M.G.J., and O'Halloran, J. (2003). *In vitro Methods in Aquatic Toxicology*. C. Mothersill, B. Austin (Eds.). Springer Verlag, Berlin, pp. 327–374.
23. Burton, G.A. (1991). *Environ. Toxicol. Chem.* **10**: 1585–1627.
24. Bettermann, A.D., Lazorchak, J.M., and Dorofi, J.C. (1996). *Environ. Toxicol. Chem.* **15**: 319–324.
25. Traunspurger, W. and Drews, C. (1996). *Hydrobiologia* **328**: 215–261.

26. Doherty, F.G. (2001). *Water Qual. Res. J. Can.* **36**: 475–518.
27. Ingersoll, C.G. (1995). In: *Fundamentals of Aquatic Toxicology; Effects, Environmental Fate, and Risk Assessment*. G.M. Rand (Ed.). Taylor and Francis, New York.
28. Carr, R.S., Williams, J.W., and Fragata, C.T.B. (1989). *Environ. Toxcol. Chem.* **8**: 533–543.
29. Carr, R.S. (1998). In: *Microscale Testing in Aquatic Toxicology. Advances, Techniques and Practice*. Wells, P.G., Lee, K., and Blaise, C. (Eds.). CRC Press, Boca Raton, FL, pp. 523–538.
30. Nipper, M. et al. (2002). *Mar. Pollut. Bull.* **44**: 789–806.
31. Wong, C.K.C., Cheung, R.Y.H., and Wong, M.H. (1999). *Environ. Pollut.* **105**: 175–183.
32. Nendza, M. (2002). *Chemosphere* **48**: 865–883.
33. Coughlan, B.M. et al. (2002). *Mar. Pollut. Bull.* **44**: 1359–1365.
34. Kemble, N.E. et al. (2000). *Arch. Environ. Contam. Toxicol.* **39**: 452–461.
35. Cook, N.H. and Wells, P.G. (1996). *Water Qual. Res. J. Can.* **31**: 673–708.
36. Grant, A. and Briggs, A.D. (2002). *Mar. Environ. Res.* **53**: 95–116.
37. Ankley, G.T., Hoke, R.A., Giesy, J.P., and Winger, P.V. (1989). *Chemosphere* **18**: 2069–2075.
38. Côte, C. et al. (1998). *Environ. Toxicol. Water Qual.* **13**: 93–110.
39. Guzzella, L. (1998). *Chemosphere* **37**: 2895–2909.
40. Kahru, A., Pollumaa, L., Reiman, R., and Ratsep, A. (1999). *ATLA-Altern. Lab. Anim.* **27**: 359–366.
41. Bitton, G., Campbell, M., and Koopman, B. (1992). *Environ. Toxicol. Water Qual.* **7**: 323–328.
42. Cohn, S.A. and McGuire, J.R. (2000). *Diatom Res.* **15**: 19–29.
43. Hickey, C.W. and Martin, M.L. (1995). *Environ. Toxicol. Chem.* **14**: 1401–1409.
44. Ho, K.T. et al. (1997). *Environ. Toxicol. Chem.* **16**: 551–558.
45. Williams, T.D. (1992). *Mar. Ecol.-Prog. Ser.* **91**: 221–228.
46. U.S. EPA. (2001). Methods for Assessing the Chronic Toxicity of Marine and Estuarine Sediment-associated Contaminants with the Amphipod *Leptocheirus plumulosus*, 1st Edn. United States Environmental Protection Agency, Washington, DC.
47. Keddy, C.J., Greene, J.C., and Bonnell, M.A. (1995). *Ecotox. Environ. Safety* **30**: 221–251.
48. Stronkhorst, J., van Hattum, B., and Bowmer, T. (1999). *Environ. Toxicol. Chem.* **18**: 2343.
49. Phelps, H.J., Hardy, J.T., Pearson, W.H., and Apts, C.W. (1983). *Mar. Pollut. Bull.* **14**: 452–455.
50. Byrne, P.A. and O'Halloran, J. (2000). *Environ. Toxcol.* **15**: 456–468.
51. Byrne, P.A. and O'Halloran, J. (1999). *Mar. Pollut. Bull.* **39**: 97–105.
52. Weis, J.S. et al. (2001). *Can. J. Fish. Aquat. Sci.* **58**: 1442–1452.
53. Michel, X. et al. (2001). *Polycycl. Aromat. Compd.* **18**: 307–324.
54. Gagne, F., Blaise, C., and Bermingham, N. (1996). *Toxicol. Lett.* **87**: 85–92.
55. Gunther, A.J. et al. (1997). *Mar. Environ. Res.* **44**: 41–49.
56. Escartin, E. and Porte, C. (1999). *Mar. Pollut. Bull.* **38**: 1200–1206.
57. Myers, M.S. et al. (1998). *Mar. Pollut. Bull.* **37**: 92–113.
58. Willett, K.L. et al. (1997). *Environ. Toxicol. Chem.* **16**: 1472–1479.
59. Rice, C.A. et al. (2000). *Mar. Environ. Res.* **50**: 527–533.
60. Johnson, L. et al. (1993). *Mar. Environ. Res.* **35**: 165–170.
61. Besselink, H.T. et al. (1998). *Aquat. Toxicol.* **42**: 271–285.
62. Beyer, J.U. et al. (1996). *Aquat. Toxicol.* **36**: 75–98.
63. Mondon, J.A., Duda, S., and Nowak, B.F. (2001). *Aquat. Toxicol.* **54**: 231–247.
64. Wong, C.K.C., Yeung, H.Y., Woo, P.S., and Wong, M.H. (2001). *Aquat. Toxicol.* **54**: 69–80.
65. Digiulio, R.T. et al. (1989). *Environ. Toxicol. Chem.* **8**: 1103–1123.
66. Gadagbui, B.K.M. and Goksoyr, A. (1996). *Biomarkers* **1**: 252–261.
67. Zapata-Perez, O. et al. (2000). *Mar. Environ. Res.* **50**: 385–391.
68. Bombardier, M., Bermingham, N., Legault, R., and Fouquet, A. (2001). *Chemosphere* **42**: 931–944.
69. Johnson, B.T. and Long, E.R. (1998). *Environ. Toxicol. Chem.* **17**: 1099–1106.
70. Bombail, V., Aw, D., Gordon, E., and Batty, J. (2001). *Chemosphere* **44**: 383–392.
71. Everaarts, J.M. (1995). *Mar. Pollut. Bull.* **31**: 431–438.
72. Theodorakis, C.W., Dsurney, S.J., and Shugart, L.R. (1994). *Environ. Toxicol. Chem.* **13**: 1023–1031.
73. Pandrangi, R., Petras, M., Ralph, S., and Vrzoc, M. (1995). *Environ. Mol. Mutagen.* **26**: 345–356.
74. Kilemade, M.F. et al. (2004). *Environ. Mol. Mutagen.* **44**: 56–64.
75. Stein, J.E. et al. (1992). *Environ. Toxicol. Chem.* **11**: 701–714.
76. Diez, S., Abalos, M., and Bayona, J.M. (2002). *Water Res.* **36**: 905–918.
77. Del Valls, T.A., Forja, J.M., and Gomez-Parra, A. (1998). *Environ. Toxicol. Chem.* **17**: 1073–1084.
78. Carr, R.S. et al. (1996). *Environ. Toxicol. Chem.* **15**: 1218–1231.
79. Khim, J.S. et al. (2001). *Arch. Environ. Contam. Toxicol.* **40**: 151–160.
80. Nagler, J.J. and Cyr, D.G. (1997). *Environ. Toxicol. Chem.* **16**: 1733–1738.
81. Lye, C.M., Frid, C.L.J., and Gill, M.E. (1998). *Mar. Ecol. Prog. Ser.* **170**: 249–260.
82. McArdle, M.E., McElroy, A.E., and Elskus, A.A. (2004). *Environ. Toxicol. Chem.* **23**: 953–959.

ALGAL TOXINS IN WATER

B. Ji
M.H. Wong
R.N.S. Wong
Y. Jiang
Hong Kong Baptist University
Kowloon, Hong Kong

INTRODUCTION

Algal toxins have various adverse implications on the ecosystem directly and indirectly. Cyanotoxins, a group of algal toxins produced by toxic cyanobacteria, are of great concern. The understanding of their occurrence, properties, detection, and removal from water bodies is of importance for monitoring and management of cyanotoxins in freshwater bodies.

ECOLOGY OF TOXIC CYANOBACTERIA IN WATER

Over 40 genera of cyanobacteria are toxic, and the toxic species with particular public health concern generally inhabit freshwater bodies. They are classified into three groups according to the toxin produced: hepatotoxin-producing species (e.g., *Anabaena, Anabaenopsis, Microcystis, Nodularia, Oscillatoria*, and *Cylindrospermopsis*); neurotoxin-producing species (e.g., *Anabaena, Aphanizomenon*, and *Lyngbya*); and dermatotoxin-producing species (e.g., *Lyngbya*; and *Schizothrix*) (1,2).

The toxic cyanobacterial bloom usually occurs in the surface, near-shore area of slow-flowing freshwater bodies. The formation of cyanobacterial bloom and the associated occurrence of cyanotoxins are affected by a number of environmental parameters such as light intensity, temperature, nutrients, trace metals, and organic matters (3). The progressive eutrophication has been reported to be the major cause for the formation of cyanobacterial blooms (4).

CYANOTOXINS AND RELATED HEALTH PROBLEMS

Cyanotoxins are basically classified into three types according to their toxicological targets: hepatotoxins (e.g., microcystins, nodularins, cylindrospermopsins), neurotoxins (e.g., anatoxins, saxitoxins), and dermatotoxins (e.g., aplysiatoxin, lyngbyatoxin A) (5).

Hepatotoxins

The hepatotoxins mainly target the liver. The three major hepatotoxins are microcystins, nodularins, and cylindrospermopsins. Microcystins are the most implicated cyanotoxins in animal and human poisonings and can be produced by several genera, including *Microcystis, Anabaena, Oscillatoria, Nostoc*, and *Anabaenopsis*. Microcystins are cyclic peptides that are made up of five nonprotein amino acids and two protein amino acids (Fig. 1). The toxicity of microcystins is associated with the presence of Adda (3-amino-9-methoxy-2,6,8- trimethyl-10-phenyldeca-4,6-dienoic acid), which has a characteristic violet absorbance at 238 nm (6). Structural variants of microcystins differ in the presence of two protein amino acids (X and Z), two groups (R1 and R2), and two demethylated positions (3 and 7). The molecular weight of microcystins ranges from 800 to 1100 daltons (7).

Microcystins can inhibit hepatocyte protein phosphatases 1 and 2A (8,9). The contamination of microcystins in drinking water is associated with a high incidence of primary liver cancer. The LD_{50} of the extreme toxic microcystin, microcystin-LR, is 50 $\mu g\,kg^{-1}$ body weight in mice by intraperitoneal (i.p.) injection (10).

Nodularins are cyclic pentapeptides and structurally similar to microcystins (Fig. 2). Until now, only four variants of nodularin have been identified in *Nodularia spumigena*. Nodularin can also inhibit the activities of protein phosphatases 1 and 2A and potently promote tumor formation (11,12). The LD_{50} of nodularin in mice is 30 $\mu g\,kg^{-1}$ body weight by i.p. injection (10). In addition, long-term exposure to a low dose of nodularins or microcystins may cause progressive changes in liver tissue, leading to chronic inflammation and focal degeneration of hepatocytes (13).

Cylindrospermopsins are tricyclic alkaloids (Fig. 3) and are produced by *Cylindrospermopsis raciborskii, Aphanizomenon ovalisporum, Umezakia natans*, and *Raphidiopsis curvata*. Cylindrospermopsins are potent inhibitors of protein synthesis and their main target organ is the liver. Unlike the other hepatotoxins, cylindrospermopsins can also affect other organs (e.g., kidney, thymus, and heart). The LD_{50} of cylindrospermopsin in mice is 2100 $\mu g\,kg^{-1}$ body weight at 24 hours by i.p. injection (14).

Neurotoxins

The cyanobacterial neurotoxins are relatively unstable and less common when compared with hepatotoxins. They are, however, diverse in terms of the chemical structures and mammalian toxicities. The major neurotoxins are anatoxins and saxitoxins.

Anatoxins are produced mainly by *Anabaena* species. The two common anatoxins are anatoxin-a and anatoxin-a(s) (Fig. 4). Anatoxin-a is a cholinergic agonist that binds to nicotinic acetylcholine receptors in nerves and neuromuscular junctions. Anatoxin-a(s) mainly inhibits acetylcholinesterase acitivity. The values of LD_{50} of anatoxin-a

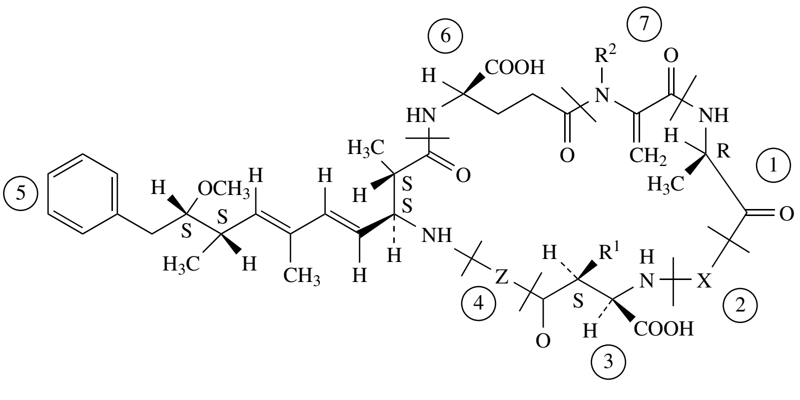

Figure 1. General structure of microcystins [cyclo-D-Ala^1-X^2-D-methyl-isoaspartic acid (MeAsp)^3-Z^4-Adda^5-D-isoglutamic acid (Glu)^6-N-methyldehydroalanine (Mdha)^7, R=H or CH_3].

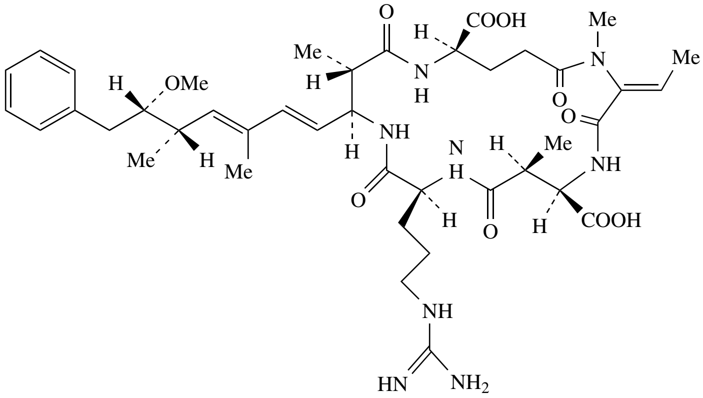

Figure 2. Structure of nodularin.

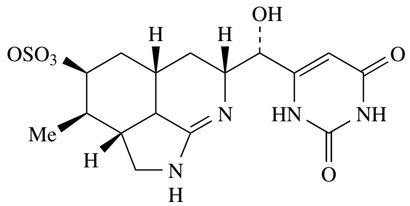

Figure 3. Structure of cylindrospermopsin.

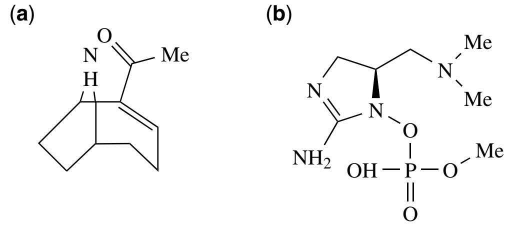

Figure 4. Structures of Anatoxin-a (**a**) and Anatoxin-a(s) (**b**).

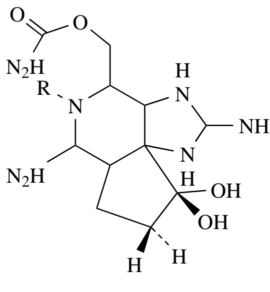

Figure 5. Structure of saxitoxin.

and anatoxin-a(s) are 200 $\mu g\,kg^{-1}$ and 20 $\mu g\,kg^{-1}$ body weight in mice by i.p. injection, respectively (15).

Saxitoxins (Fig. 5), known as marine paralytic shellfish poisons (PSPs), are recently found in freshwater cyanobacteria including *Anabaena circinalis, Lyngbya wollei, Cylindrospermopsis raciborskii*, and *Planktothrix* sp. The toxicity of saxitoxins in mice is different depending on their variants. Saxitoxin is the most potent PSP with LD_{50} of 10 $\mu g\,kg^{-1}$ body weight in mice by i.p. injection (5).

Dermatotoxins

Dermatotoxins are usually produced by benthic marine cyanobacteria (e.g., *Lyngbya majuscule*). Lyngbyatoxin A and aplysiatoxin are two specific dermatotoxins and are related to acute dermatitis and animal death. Long-term exposure to these toxins may lead to skin tumors (16).

DETECTION OF CYANOTOXINS IN WATER

The general methods developed for detection of cyanotoxins include chemical analyses, for example, high-performance liquid chromatography (HPLC), mass spectroscopy (MS), and nuclear magnetic resonance (NMR); biological assays, for example, mouse, invertebrate, and bacterial bioassays; and biochemical assays, for example, protein phosphatase inhibition assay and immunoassay (ELISA).

In chemical analyses, HPLC is widely used for the detection of most cyanotoxins (17). LC/MS is the best method for the detection of saxitoxins and its variants (18). To identify and differentiate the variants of microcystins, fast atom bombardment mass spectroscopy (FABMS) and FABMS/MS methods have been developed (19).

Mouse bioassay is the only method for the detection of all types of cyanotoxins compared with other methods (17). In invertebrate bioassay, brine shrimp, mosquito, fruitfly, house fly, and locust are commonly used target organisms (20,21). The invertebrate bioassay, however, is limited in the detection of some specific cyanotoxins. For example, house fly and locust are only sensitive to saxitoxins, whereas brine shrimp can be used only for detection of microcystins. Bacterial bioassay is seldom used because of its weakness on the reflection of the correlation between the actual concentrations of known cyanotoxin and the responses of the testing bacterium (17).

Protein phosphatase inhibition assay is the most sensitive biochemical method for the analysis of microcystins and nodularins. It uses ^{32}P-labeled glycogen phosphorylase and the colorimetric protein phosphatase inhibition assay. This assay has proved to be an efficient screening method for water samples because of its simple operation. The protein phosphatase inhibition assay is currently adopted for the analysis of microcystins in drinking water.

The detection limits for microcystin-LR in raw and finished drinking water are 0.87 $\mu g\,L^{-1}$ and 0.09 $\mu g\,L^{-1}$, respectively (22). The protein phosphatase inhibition assay, however, may be affected by various other noncyanobacterial toxins and metabolites (e.g., okadaic acid and tautomycin). Consequently, additional confirmation should be made to validate the presence of specific cyanobacterial hepatotoxins (23).

ELISA is a sensitive and specific method for the detection of cyanotoxins using polyclonal or monoclonal antibodies (24). It has been successfully employed for the quantitative detection of cyanobacterial hepatotoxins including microcystin-LR, -RR, -YR, and nodularins in domestic water supplies with the detection limits of 0.05 $\mu g\,L^{-1}$ for microcystin-LR in water samples, for instance (25–27).

STABILITY OF CYANOTOXINS IN WATERS

Cyanotoxins are normally present inside cyanobacterial cells, which enter into the water bodies during cell senescence and lysis. The release of cyanotoxins is influenced by several factors, including light intensity, application of algicide, chlorination, and the unfavorable conditions that accelerate cell death (28). Hence, the ratio of dissolved cyanotoxins to intracellular ones is different at different growth phases of cyanobacteria. In field study, the breakdown of a cyanobacterial bloom may lead to an increase of dissolved microcystins in water (29), although the concentrations detected are relatively low compared with cyanotoxins inside the cells.

The existence of cyanotoxins in the aquatic environment is also affected by their stability. Factors such as temperature, pH, light intensity, and the existence of other microorganisms and oxidants with respect to their effect on cyanotoxin stability have been investigated. Anatoxin-a, for example, is stable at pH 4 for at least 21 days in reservoir water, whereas less than 5% of the original concentration was detected after 14 days at pH 8–10 (30). Microcystins and cylindrospermopsins can be oxidized by strong oxidants, for example, ozone and sodium hypochlorite (NaOCl), and their degradation under intense UV light can be accelerated by the addition of titanium dioxide (TiO_2) (31–33). Similarly, anatoxin-a and saxitoxins are also unstable when exposed to ozone. Photolysis of microcystin may occur in the presence of humic substances and pigments of cyanobacterial cells (31,34). In addition, microcystin-LR and -RR can be degraded *in vitro* in bacterium *Sphingomonas* sp. by microcystinase (35). The degradation of nodularin can be performed with the presence of the extract of *Nodularia spumigena* (36). Anatoxin-a can be degraded micobiologically by *Pseudomonas* sp. in pure culture (37). On the other hand, microcystins and anatoxins can also adsorb onto sediments, suspended solids, and dissolved organic matters, which may correlate to the accumulation and persistence of these toxins in the environment.

REMOVAL OF CYANOTOXINS FROM WATER

Cyanotoxins may impose fatal, acute, or chronic effects on animals and humans and present a health problem to the public through contamination of drinking water, recreationally used water, and fish or shellfish. The extent of cyanotoxin poisoning can be decreased by reducing human exposure to cyanotoxins, which can be achieved through preventing toxic cyanobacterial bloom formation; monitoring of the number of cyanobacteria or the concentration of cyanotoxins in water; notifying the public of the possible hazards in the water; and providing technical and scientific advice for removal of cyanotoxins in water.

The World Health Organization (WHO) has suggested the value of 1 $\mu g\,L^{-1}$ for microcystin-LR in drinking water resources and provides guidelines on monitoring the number of cyanobacterial cells and the concentration of chlorophyll *a* in drinking and recreationally used waters, which indicates that the concentrations of microcystins can be forecasted from the densities of cyanobacterial cells if microcysin-producing cyanobacteria are dominant (38,39). Cyanobacterial blooms usually can be controlled in several ways, for example, prevention of eutrophication by reduction of external nutrient loading, control of cyanobacteria by raw water abstraction, and application of algicide. The conventional methods adopted in water treatment, including coagulation, flocculation, sand filtration, ultrafiltration, and microfiltration, are effective for removal of cyanobacterial cells from drinking water. However, the cyanotoxins may enter into water because of the lysis of cyanobacterial cells during the water treatment process. The efficiency of these methods to remove dissolved cyanotoxins in water is low (40).

The more effective and practical chemical and physical methods to eliminate the contamination of dissolved cyanotoxins in water are developed based on the adsorption and destruction of the toxins. Ozonation and potassium permanganate, for example, are effective for the destruction of dissolved microcystins. The treatments by chemical oxidants, however, require optimal doses, temperature, pH, and the optimal concentrations of dissolved organic carbon (41). Although the effect of chlorination on destruction of hepatotoxins has been investigated subsequently, the efficiency of conventional chlorination in water treatment on removal of microcystins is low, which might be because of the insufficient chlorine available for microcystin oxidation. The other pertinent techniques for removal of cyanotoxins from drinking water, for example, reverse osmosis, nanofiltration, powdered activated carbon adsorption, granular activated carbon adsorption, and biodegradation, have also been investigated (42–44). However, the efficiencies of these methods for removal of different cyanotoxins are variable. No sole method regarding the removal of all types of cyanotoxins from water bodies efficiently can be adopted. The selection of the appropriate treatment method depends on the target toxin. Further investigation is required to assess the optimization of these treatment methods.

BIBLIOGRAPHY

1. Carmichael, W.W. (1997). The cyanotoxins. In: *Advances in Botanical Research*. P. Callow (Ed.). Academic Press, London, pp. 211–256.

2. Whitton, B.A. (1992). Diversity, ecology and taxonomy of the cyanobacteria. In: *Photosynthetic Prokaryotes*. N.H. Mann and N.G. Carr (Eds.). Plenum Press, New York, pp. 1–51.
3. Kaebernick, M. and Neilan, B.A. (2001). Ecological and molecular investigations of caynotoxins production. *FEMS Microbiol. Ecol.* **35**: 1–9.
4. Reynolds, C.S. (1987). Cyanobacterial waterblooms. In: *Advances in Botanical Research*. P. Callow (Ed.). Academic Press, London, pp. 17–143.
5. Briand, J.F., Jacquet, S., Bernard, C., and Humbert, J.F. (2003). Health hazards for terrestrial vertebrates from toxic cyanobacteria in surface water ecosystems. *Vet. Res.* **34**: 361–377.
6. Rinehart, K.L. et al. (1988). Nodularin, microcystin and the configuration of Adda. *J. Am. Chem. Soc.* **110**: 8557–8558.
7. Carmichael, W.W. (1992). Cyanobacterial secondary metabolites—the cyanotoxins. *J. Appl. Bacteriol.* **72**: 445–459.
8. Runnegar, M. et al. (1995). *In vivo* and *in vitro* binding of microcystin to protein phosphatases 1 and 2A. *Biochem. Biophys. Res. Commun.* **216**: 162–169.
9. Fischer, W.J. et al. (2000). Microcystin-LR toxicodynamics, induced pathology, and immunohistochemical localization in livers of blue-green algae exposed rainbow trout (*Oncorhynchus mykiss*). *Toxicol. Sci.* **54**: 365–373.
10. Rinehart, K.L., Namikoshi, M., and Choi, B.W. (1994). Structure and biosynthesis of toxins from blue-green algae (cyanobacteria). *J. Appl. Phycol.* **6**: 159–176.
11. Ohta, T. et al. (1994). Nodularin, a potent inhibitor of protein phosphatases 1 and 2A, is a new environmental carcinogen in male F344 rat liver. *Cancer Res.* **54**: 6402–6406.
12. Carmichael, W.W., Eschedor, J.T., Patterson, G.M.L., and Moore, R.E. (1988). Toxicity and partial structure of a hepatotoxic peptide produced by the cyanobacterium *Nodularia spumigena* Mertens emend. L 575 from New Zealand. *Appl. Environ. Microbiol.* **54**: 2257–2263.
13. Elleman, T.C., Falconer, I.R., Jackson, A.R.B., and Runnegar, M.T. (1978). Isolation, characterization and pathology of the toxin from a *Microcystis aeruginosa* (*Anacystis cyanea*) bloom. *Aust. J. Biol. Sci.* **31**: 209–218.
14. Hawkins, P.R. et al. (1997). Isolation and toxicity of *Cylindrospermopsis raciborskii* from an ornamental lake. *Toxicon.* **35**: 341–346.
15. Falconer, I.R. (1998). Algal toxins and human health. In: *The Handbook of Environmental Chemistry, Part C: Quality and Treatment of Drinking Water II*. J. Hrubec (Ed.). Springer-Verlag, Berlin, pp. 53–82.
16. Osborne, N.J., Webb, P.M., and Shaw, G.R. (2001). The toxins of *Lyngbya majuscula* and their human and ecological health effects. *Environ. Int.* **27**: 381–392.
17. Lawton, L.A., Beattie, K.A., Hawser, S.P., Campbell, D.L., and Codd, G.A. (1994). Evaluation of assay methods for the determination of cyanobacterial hepatotoxicity. In: *Detection Methods for Cyanobacterial Toxins. Proceedings of the First International Symposium on Detection Methods for Cyanobacterial Toxins*. G.A. Codd, T.M. Jeffries, C.W. Keevil, and E. Potter (Eds.). The Royal Society of Chemistry, Cambridge, UK, pp. 111–116.
18. Lawrence, J.F., Menard, C., and Cleroux, C. (1995). Evaluation of prechromatographic oxidation for liquid chromatographic determination of paralytic shellfish poisons in shellfish. *J. AOAC Int.* **78**: 514–520.
19. Harada, K.I. (1996). Chemistry and detection of microcystins. In: *Toxic Microcystis.* M.F. Watanabe, K.I. Harada, W.W. Carmichael, and H. Fujiki (Eds.). CRC, Boca Raton, FL, pp. 103–148.
20. Falconer, I.R. (1993). Measurement of toxins from blue-green algae in water and foodstuffs. In: *Algal Toxins in Seafood and Drinking Water*. I.R. Falconer (Eds.). Academic Press, London, pp. 165–175.
21. McElhiney, J., Lawton, L.A., Edwards, C., and Gallacher, S. (1998). Development of a bioassay employing the desert locust (*Schistocerca gregaria*) for the detection of saxitoxin and related compounds in cyanobacteria and shellfish. *Toxicon.* **36**: 417–420.
22. Lambert, T.W., Boland, M.P., Holmes, C.F.B., and Hrudey, S.E. (1994). Quantitation of the microcystin hepatotoxins in water at environmentally relevant concentrations with the protein phosphatase bioassay. *Environ. Sci. Technol.* **28**: 753–755.
23. Metcalf, J.S., Bell, S.G., and Codd, G.A. (2001). Colorimetric immuno-protein phosphatase inhibition assay for specific detection of microcystins and nodularins of *Cyanobacteria*. *Appl. Environ. Microbiol.* **67**: 904–909.
24. Nagata, S. et al. (1995). Novel monoclonal antibodies against microcystin and their protective activity for hepatotoxicity. *Nat. Toxins.* **3**: 78–86.
25. An, J. and Carmichael, W.W. (1994). Use of a colorimetric protein phosphatase inhibition assay and enzyme linked immunosorbent assay for the study of microcystins and nodularins. *Toxicon.* **32**: 1495–1507.
26. Chu, F.S., Huang, X., and Wei, R.D. (1990). Enzyme-linked immunosorbent assay for microcystins in blue-green algal blooms. *J. Assoc. Anal. Chem.* **73**: 451–456.
27. Chen, G., Chen, C.W., Yu, S.Z., and Ueno, Y. (1998). Evaluation of different water treatment to remove microcystins by using a highly sensitive ELISA. In: *Mycotoxins and Phycotoxins—Developments in Chemistry, Toxicology and Food Safety*. M. Miraglia, H. van Egmond, C. Bera, and J. Gibert (Eds.). Alaken Inc, Fort Collins, CO.
28. Welker, M., Steinberg, C., and Jones, G.J. (2001). Release and persistance of microcystins in natural waters. In: *Cyanotoxins—Occurrence, Causes and Consequences*. I. Chorus (Ed.). Springer-Verlag Telos, Berlin, pp. 85–103.
29. Jones, G.J., Bourne, D.G., Blakeley, R.L., and Doelle, H. (1994). Degradation of the cyanobacterial hepatotoxin microcystin by aquatic bacteria. *Nat. Toxins* **2**: 228–235.
30. Stevens, D.K. and Krieger, R.I. (1991). Stability studies on the cyanobacterial nicotinic alkaloid anatoxin-a. *Toxicon.* **29**: 167–179.
31. Tsuji, K. et al. (1995). Stability of microcystins from cyanobacteria II: effect of UV light on decomposition and isomerization. *Toxicon.* **33**: 1619–1631.
32. Chiswell, R.K. et al. (1999). Stability of cylindrospermopsin, the toxin from the cyanobacterium, *Cylindrospermopsis raciborskii*: effect of pH, temperature, and sunlight on decomposition. *Environ. Toxicol.* **14**: 155–161.
33. Senogles, P.J., Scott, J.A., Shaw, G., and Stratton, H. (2001). Photocatalytic degradation of the cyanotoxin cylindrospermopsin using titanium dioxide and UV irradiation. *Water Res.* **35**: 1245–1255.
34. Welker, M. and Steinberg, C. (2000). Rates of humic substance photosensitized degradation of microcystin-LR in natural waters. *Environ. Sci. Technol.* **34**: 3415–3419.
35. Bourne, D.G. et al. (1996). Enzymatic pathway for the bacterial degradation of the cyanobacterial cyclic peptide toxin microcystin-LR. *Appl. Environ. Microbiol.* **62**: 4086–4094.

36. Twist, H. and Codd, G.A. (1997). Degradation of the cyanobacterial hepatotoxin, nodularin, under light and dark conditions. *FEMS Microbiol. Lett.* **151**: 83–88.
37. Kiviranta, J., Sivonen, K., Lahti, K., and Niemelasi, R. (1991). Production and biodegradation of cyanobacterial toxins—a laboratory study. *Arch. Fur. Hydrobiol.* **121**: 281–294.
38. WHO. (1998). *Guidelines for Drinking Water Quality*. World Health Organization, Geneva.
39. WHO. (1999). *Toxic Cyanobacteria in Water: A Guide to their Public Health Consequences, Monitoring and Management*. World Health Organization, Geneva.
40. Falconer, I.R. (1999). An overview of problems caused by toxic blue-green algae (cyanobacteria) in drinking and recreational water. *Environ. Toxicol.* **14**: 5–12.
41. Newcombe, G. and Nicholson, B. (2004). Water treatment options for dissolved cyanotoxins. *J. Water Sup. Res. Technol.—Aqua.* **53**: 227–239.
42. Newcombe, G. et al. (2003). Treatment options for microcystin toxins: similarities and differences between variants. *Environ. Technol.* **24**: 299–308.
43. Falconer, I.R. et al. (1989). Using activated carbon to remove toxicity from drinking water containing cyanobacterial blooms. *J. Am. Water Works Assoc.* **81**: 102–105.
44. Svrcek, C. and Smith, D.W. (2004). Cyanobacteria toxins and the current state of knowledge on water treatment options: a review. *J. Environ. Eng. Sci.* **3**: 155–184.

GROUND WATER QUALITY IN AREAS ADJOINING RIVER YAMUNA AT DELHI, INDIA

Chakresh K. Jain
Mukesh K. Sharma
National Institute of Hydrology
Roorkee, India

The groundwater quality in areas adjoining the River Yamuna at Delhi (India) has been assessed to determine the suitability of groundwater for domestic use. Thirty-eight groundwater samples from shallow and deep aquifers were collected each during pre- and postmonsoon seasons in the year 2000 and analyzed for various physicochemical and bacteriological parameters and trace elements. The study indicated concentrations of total dissolved solids, nitrate, sulfate, and sodium higher than water quality standards. The presence of total coliforms indicates bacterial contamination in the groundwater. The grouping of samples according to their hydrochemical facies indicated that the majority of the samples fall into Na-K-Cl-SO_4 followed by Na-K-HCO_3 and Ca-Mg-Cl-SO_4 hydrochemical facies. The qualitative analysis of data depicted higher concentrations of various physicochemical and bacteriological parameters on the western side of River Yamuna, even in deep aquifers.

INTRODUCTION

Water is an essential and vital component of our life support system. In tropical regions, groundwater plays an important role within the context of fluctuating and increasing contamination of surface water resources. Groundwater has unique features (excellent natural quality, usually free from pathogens, color, and turbidity), which render it particularly suitable for a public water supply. Groundwater also plays an important role in agriculture for watering of crops and for irrigation of dry season crops. It is estimated that about 45% of the irrigation water requirement is met from groundwater sources. Unfortunately, the availability of groundwater is not unlimited nor is it protected from deterioration. In most instances, the extraction of excessive quantities of groundwater has resulted in drying up of wells, damaged ecosystems, land subsidence, saltwater intrusion, and depletion of the resource. Groundwater quality is being increasingly threatened by agricultural, urban, and industrial wastes. It has been estimated that once pollution enters the subsurface environment, it may remain concealed for many years, disperses over wide areas of the groundwater aquifer, and renders groundwater supplies unsuitable for consumption and other uses. The rate of depletion of groundwater levels and the deterioration of groundwater quality are of immediate concern in major cities and towns of the country.

The National Capital Territory (NCT) of Delhi is facing severe problems in managing groundwater quality and quantity. Surface water bodies play a significant role in groundwater flow. The hydraulic gradient has a significant role in lateral and vertical migration of contaminants in groundwater aquifers. Therefore, the present study has been carried out to assess the suitability of groundwater for domestic uses in areas adjoining the River Yamuna at Delhi and to examine the likely impact of Yamuna River water quality on groundwater. The suitability of each well for drinking has been reported in an earlier report (1).

STUDY AREA

Delhi generates about 1900 MLD of sewage against installed capacity of 1270 MLD of sewage treatment. The balance of untreated sewage along with a significant quantity of partially treated sewage is discharged into the River Yamuna every day. The river receives sewage and industrial wastes through various drains, which join the River Yamuna between Wazirabad and Okhla. Thus Delhi is the largest contributor of pollution to the River Yamuna, which receives almost 80% of its pollution load through these drains.

The climate of Delhi is influenced mainly by its inland position and the prevalence of continental type air during the major part of the year. Extreme dryness with an intensely hot summer and cold winter are the characteristics of the climate. Only during the monsoon months does air of oceanic origin penetrate to this area and cause increased humidity, cloudiness, and precipitation. The normal annual rainfall in the National Capital Territory of Delhi is 611.8 mm. The rainfall increases from southwest to northeast; about 81% of the annual rainfall is received during the three monsoon months of July, August, and September. The balance of annual rainfall is received as winter rains and as thunderstorm rain during pre- and postmonsoon months.

Thirty-eight groundwater samples from shallow and deep aquifers were collected from both sides of the River Yamuna at Delhi (Fig. 1). The details of sampling locations are given in Table 1.

METHODOLOGY

The groundwater samples were collected in polyethylene bottles during pre- and postmonsoon seasons in the year 2000 from open wells, hand pumps, and tube wells and were preserved by adding an appropriate reagent (2,3). The water samples for trace element analysis were collected in acid-leached polyethylene bottles and preserved by adding ultrapure nitric acid (5 mL/L). Samples for bacteriological analysis were collected in sterilized high-density polypropylene bottles. All samples were stored in sampling kits maintained at 4 °C and brought to the laboratory for detailed chemical and bacteriological analysis. The details of sampling locations are given in Table 1. The physicochemical and bacteriological analyses were performed following standard methods (2,3).

RESULTS AND DISCUSSION

The National Capital Territory of Delhi has the peculiar feature of infiltration of surface water to groundwater from the River Yamuna and from various drains in addition to customary recharge from rainfall. Groundwater recharge also occurs through stagnant water pools in low-lying areas, where surface runoff water collects. The quartzite ridge, which is the prolongation of the Aravalli mountain range, forms the principal watershed in the south, southeast, and southwest parts of Delhi. Because of it, the eastern surface runoff and drainage join the River Yamuna, whereas the runoff from the western part of Delhi goes into the Najafgarh drain. The ever increasing discharge of domestic and industrial wastes into improperly lined sewage drains in Delhi leads to a high risk of contaminating the groundwater. The excessive groundwater uplift also has an adverse impact on water quality in groundwater aquifers of limited thickness.

General Characteristics

The hydrochemical data for the two sets of samples collected from the areas adjoining the River Yamuna at Delhi are given in Table 2. The pH in the groundwater

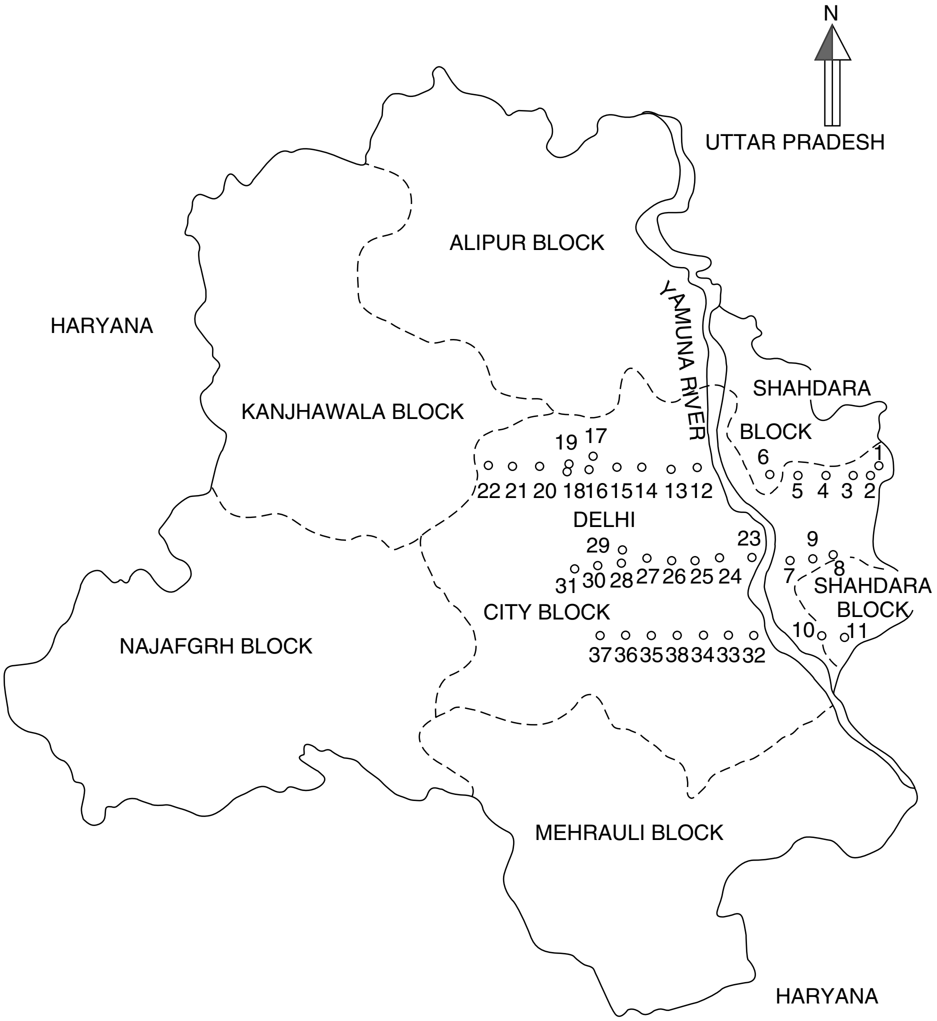

Figure 1. Study area showing sampling locations.

Table 1. Description of Groundwater Sampling Locations

Sample No.	Location	Type of Well (depth in ft.)[a]
1	Bhagwanpur Khera	HP (25)
2	Loni Road	BW (20)
3	Kabul Nagar	OW (35)
4	Naveen Shahdara	HP (20)
5	Seelampur	HP (30)
6	Shastri Park	HP (25)
7	Lakshmi Nagar	HP (100)
8	Prit Vihar	TW (40)
9	Shankar Vihar	HP (200)
10	Pratap Nagar	TW (100)
11	Himmatpuri	HP (80)
12	Civil Lines	TW (250)
13	Rajpur road	OW (25)
14	Malka Gunj	HP (20)
15	Tripolia	BW (100)
16	Gulabi Bagh	BW (40)
17	Gulabi Bagh	BW (160)
18	Shastri Nagar	BW (100)
19	Shastri Nagar	BW (60)
20	Lekhu Nagar	HP (40)
21	Ram Pura	HP (40)
22	Punjabi Bagh West	HP (40)
23	Rajghat	BW (50)
24	JLN Marg	HP (60)
25	GB Pant Hospital	BW (20)
26	Panchkuin Marg	TW (450)
27	Panchkuin Marg	HP (60)
28	Rajendra Nagar	BW (100)
29	Rajendra Nagar	BW (300)
30	Shankar Road	OW (20)
31	IARI, Pusa	TW (120)
32	Zoological Park	HP (60)
33	Golf Course	BW (80)
34	Rabindra Nagar	TW (150)
35	Teen Murti Chowk	HP (60)
36	Malcha Marg	BW (100)
37	Sardar Patel Road	TW (150)
38	Janpath	BW (250)

[a] HP: Hand pump; OW: open well; BW: bore well; TW: tube well.

of areas adjoining the River Yamuna is mostly within the range 6.7 to 8.3 during the premonsoon season and 6.6 to 8.2 during the postmonsoon season; most of the samples point toward the alkaline range in both seasons. The pH of all the samples is within the limits prescribed by BIS (4) and WHO (5) for various uses of water, including drinking and other domestic supplies. The conductivity varies from 628 to 3540 μS/cm during the premonsoon season and from 620 to 3250 μS/cm during postmonsoon; the conductivity of more than 70% of the samples was above 1000 μS/cm during both pre- and postmonsoon seasons. The maximum conductivities of 3540 and 3250 μS/cm were observed at Golf Course during the pre- and postmonsoon seasons, respectively. Higher values of conductivity in the areas nearby the River Yamuna indicate high mineralization of the groundwater.

The total dissolved solids (TDSs) in the groundwater vary from 402 to 2266 mg/L during the premonsoon season and from 397 to 2080 mg/L during the postmonsoon season. The TDSs of only about 10% of the samples analyzed were within the desirable limit of 500 mg/L. The TDSs of more than 80% of the samples were above the desirable limit of 500 mg/L. An almost similar trend was observed during the postmonsoon season. Water containing more than 500 mg/L of TDSs is not considered desirable for drinking water, though more highly mineralized water is also used where better water is not available. For this reason, 500 mg/L as the desirable limit and 2000 mg/L as the maximum permissible limit has been suggested for drinking water (4). Water containing TDSs of more than 500 mg/L causes gastrointestinal irritation (4). One sample at Golf Course (BW, 80) even crosses the maximum permissible limit of 2000 mg/L.

Carbonates, bicarbonates, and hydroxides are the main cause of alkalinity in natural waters. Bicarbonates represent the major form because they are formed in considerable amount from the action of carbonates on basic materials in the soil. The alkalinity value in the groundwater varies from 116 to 380 mg/L during the premonsoon season and from 106 to 310 mg/L during postmonsoon. About 60% of the samples from the study area fall within the desirable limit of 200 mg/L, and about 40% of the samples cross the desirable limit but are within the maximum permissible limit of 600 mg/L. No sample of the study area crosses the maximum permissible limit of 600 mg/L. The high alkalinity in the study area may be attributed to the action of carbonates upon basic materials in the soil. Such water has an unpleasant taste.

Calcium and magnesium along with their carbonates, sulfates, and chlorides make the water hard, both temporarily and permanently. A limit of 300 mg/L has been recommended as the desirable limit for drinking water (4). The total hardness in the study area ranges from 116 to 871 mg/L during the premonsoon season and from 114 to 792 mg/L during postmonsoon. More than 80% of the samples were found well within the desirable limits for domestic use during both pre- and postmonsoon seasons. However, one sample from Golf Course even crosses the maximum permissible limit of 600 mg/L during both seasons.

The desirable limits for calcium and magnesium in drinking water are 75 and 30 mg/L, respectively (4). In groundwater in the study area, the levels of calcium and magnesium vary from 25 to 240 and from 9 to 64 mg/L, respectively, during the premonsoon season. Slightly lower levels of calcium and magnesium were observed during the postmonsoon season. In groundwater, the calcium content generally exceeds the magnesium content in accordance with their relative abundance in rocks. The increase in magnesium is proportionate to calcium in both seasons.

The concentration of sodium in the study area varies from 55 to 340 mg/L during the premonsoon season and from 51 to 322 mg/L during postmonsoon. The sodium concentration of more than 50 mg/L makes the water unsuitable for domestic use. The sodium concentration was higher at all sites in the study area. The high sodium values may be attributed to base-exchange phenomena. Groundwater with high sodium is unsuitable for irrigation due to the sodium sensitivity of crops/plants.

Table 2. Hydrochemical Data of Groundwater Samples of Delhi (Pre- and Postmonsoon, 2000)[a]

Parameter	Minimum	Maximum	Mean
pH	6.7 (6.6)	8.3 (8.2)	7.1 (7.1)
Conductivity, μS/cm	628 (620)	3540 (3250)	1463 (1370)
TDSs, mg/L	402 (397)	2266 (2080)	936 (879)
Alkalinity, mg/L	116 (106)	380 (310)	201 (194)
Hardness, mg/L	116 (114)	841 (792)	235 (225)
Chloride	17 (17)	400 (390)	86 (86)
Sulfate, mg/L	43 (48)	690 (680)	180 (177)
Nitrate, mg/L	1.0 (ND)	286 (287)	78 (68)
Phosphate, mg/L	0.07 (0.03)	0.28 (0.21)	0.16 (0.08)
Fluoride, mg/L	0.33 (ND)	1.31 (0.76)	0.80 (0.41)
Sodium, mg/L	55 (51)	340 (322)	160 (157)
Potassium, mg/L	4.2 (2.2)	121 (103)	26 (20)
Calcium, mg/L	25 (21)	240 (227)	59 (59)
Magnesium, mg/L	9.0 (7.0)	64 (56)	21 (19)
MPN Coliform, per 100 mL	Nil (Nil)	2400 (150)	— (—)
Total count, per 100 mL	10 (12)	1850 (540)	— (—)
Copper, mg/L	0.006 (0.003)	0.178 (0.085)	0.023 (0.012)
Iron, mg/L	0.390 (0.128)	5.740 (5.842)	1.960 (1.216)
Manganese, mg/L	0.009 (0.008)	0.944 (0.837)	0.162 (0.183)
Cobalt, mg/L	0.005 (Nil)	0.034 (0.037)	0.016 (0.011)
Nickel, mg/L	0.005 (0.009)	0.043 (0.113)	0.028 (0.028)
Chromium, mg/L	0.006 (0.003)	0.033 (0.078)	0.013 (0.011)
Lead, mg/L	0.010 (0.012)	0.064 (0.098)	0.033 (0.037)
Cadmium, mg/L	0.003 (0.005)	0.010 (0.021)	0.006 (0.011)
Zinc, mg/L	0.021 (0.012)	1.110 (0.732)	0.320 (0.190)

[a] Values in parentheses represent postmonsoon data.

The concentration of potassium in groundwater varies from 4.2 to 121 mg/L during the premonsoon season and from 2.2 to 103 mg/L during postmonsoon. Potassium is an essential element for humans, plants, and animals and is derived in the food chain mainly from vegetation and soil. The main sources of potassium in groundwater include rainwater, weathering of potash silicate minerals, potash fertilizers, and surface water used for irrigation. It is more abundant in sedimentary rocks and commonly present in feldspar, mica, and other clay minerals. The Bureau of Indian Standards has not included potassium in drinking water standards. However, the European Economic Community has prescribed a guideline level for potassium at 10 mg/L in drinking water. Though potassium is found extensively in some igneous and sedimentary rocks, its concentration in natural waters is usually quite low because potassium minerals are resistant to weathering and dissolution. A higher potassium content in groundwater is indicative of groundwater pollution.

The concentration of chloride varies from 17 to 400 mg/L during the premonsoon season. Almost the same trend was observed during the postmonsoon season. The maximum chloride content in groundwater was recorded at Golf Course. No sample in the study area crosses the maximum permissible limit of 1000 mg/L. The limits of chloride have been laid down primarily from taste considerations. A limit of 250 mg/L chloride has been recommended for drinking water supplies (4,5). However, no adverse health effects on humans have been reported from intake of waters containing a higher chloride content.

The concentration of sulfate in the study area varies from 43 to 690 mg/L during the premonsoon season and from 48 to 680 mg/L during postmonsoon. Most of the samples fall within the permissible limit (400 mg/L) for drinking water supplies. Only two samples from Himmat Puri and Golf Course exceeded the maximum permissible limit. In groundwater, sulfate generally occurs as soluble salts of calcium, magnesium, and sodium. The sulfate content of water may change significantly with time during infiltration of rainfall and groundwater recharge, which takes place mostly from stagnant water pools, puddles, and surface runoff water collected in low-lying areas.

The nitrate content of drinking water is considered important for its adverse health effects. The occurrence of high levels of nitrate in groundwater is a prominent problem in NCT-Delhi. The nitrate content in the groundwater of areas adjoining the River Yamuna varies from 1 to 286 mg/L during the premonsoon season. Almost the same trend was observed during postmonsoon. Of the 38 samples analyzed, 20 samples (52.5%) had nitrate content less than 45 mg/L, whereas in about 20% of the groundwater samples, the nitrate content exceeded even the maximum permissible limit of 100 mg/L. The higher level of nitrate at certain locations may be attributed to the surface disposal of domestic sewage and runoff from agricultural fields. It has also been observed that the groundwater samples collected from hand pumps at various depths have a high nitrate content, which may be attributed to wellhead pollution.

Nitrate is an effective plant nutrient and is moderately toxic. A limit of 45 mg/L has been prescribed by WHO (5) and BIS (4) for drinking water supplies. A concentration above 45 mg/L may be detrimental to human health. In higher concentrations, nitrate may produce a disease known as methemoglobinemia (blue baby syndrome), which generally affects bottle-fed infants. Repeated heavy doses of nitrates by ingestion may also cause carcinogenic diseases.

Fluoride is present in soil strata due to the presence of geological formations such as fluorspar and fluorapatite and amphiboles such as hornblende, tremolite, and mica. Weathering of alkali, silicate, igneous, and sedimentary rocks, especially shales, contributes a major portion of fluorides to groundwaters. In addition to natural sources, considerable amounts of fluoride may be contributed by human activities. Fluoride salts are commonly used in the steel, aluminum, brick, and tile industries. Fluoride containing insecticides and herbicides may be contributed by agricultural runoff. Phosphatic fertilizers, which are extensively used, often contain fluorides as an impurity, and they may increase the levels of fluoride in soil. The accumulation of fluoride in soil eventually results in leaching by percolating water, thus increasing the fluoride concentration in groundwater. However, in the groundwater of the area adjoining the River Yamuna, the fluoride content was within the maximum permissible limit of 1.5 mg/L in all samples.

Bacteriological Parameters

The presence of coliforms in water is an indicator of contamination by human or animal excrement. The presence of fecal coliforms in groundwater is a potential public health problem, because fecal matter is a source of pathogenic bacteria and viruses. Groundwater contamination from fecal coliform bacteria is generally caused by percolation from contamination sources (domestic sewage and septic tank) into the aquifers and also by poor sanitation. Shallow wells are particularly susceptible to such contamination. The bacteriological contamination of groundwater in Delhi is mostly attributable to indiscriminate dumping of waste and garbage without any precautions or scientific disposal practices. In Delhi, most of the hand pumps withdraw groundwater from upper strata, which is most susceptible to contamination from polluted surface water.

The groundwater samples collected from the area adjoining the River Yamuna in Delhi have significantly high total coliform. Bacteriological analysis indicates the presence of coliforms in more than 75% of samples during both pre- and postmonsoon seasons. The presence of coliforms was reported mostly from hand pumps. More than 30% of the samples have MPN coliforms >100 per 100 mL during the premonsoon season. Inadequate maintenance of hand pumps and unhygienic conditions around the structures may be responsible for bacterial contamination. Indiscriminate land disposal of domestic waste on the surface, improper disposal of solid waste, and leaching of wastewater from landfills further heighten the chances of bacterial contamination.

Trace Elements

Most of the trace metals are of immediate concern because of their toxicity and nonbiodegradable nature. Cadmium, chromium, and lead are highly toxic to humans even in low concentrations. The concentrations of heavy metals in groundwater except iron, which is present in appreciable concentration, were below the prescribed maximum permissible limits in most of the samples. The concentration of iron varies from 0.39 to 5.74 mg/L during the premonsoon season and from 0.128 to 5.842 mg/L during postmonsoon. The concentration ranges of copper and zinc were well below maximum permissible limits.

The study clearly indicated higher concentration of total dissolved solids, electrical conductivity, nitrate, sulfate, and sodium. The presence of total coliforms indicates bacterial contamination in groundwater. The presence of heavy metals in groundwater was recorded in many samples, but the levels were not significant. Water quality standards have been violated for TDSs, nitrate, sulfate, and sodium at a few places.

Chadha's Diagram for Hydrochemical Classification

Chadha's diagram (6) is a somewhat modified version of the Piper trilinear diagram (7). In this diagram, the difference in milliequivalent percentage between alkaline earths (calcium plus magnesium) and alkali metals (sodium plus potassium), expressed as percentage reacting values, is plotted on the x axis and the difference in milliequivalent percentage between weak acidic anions (carbonate plus bicarbonate) and strong acidic anions (chloride plus sulfate) is plotted on the y axis. The resulting field of study is a square or rectangle depending on the size of the scales chosen for x and y coordinates. The milliequivalent percentage differences between alkaline earth and alkali metals and between weak acidic anions and strong acidic anions would plot in one of the four possible subfields of the diagram. The main advantage of this diagram is that it can be made simply on most spreadsheet software packages.

The square or rectangular field describes the overall character of the water. The diagram has all the advantages of the diamond-shaped field of the Piper trilinear diagram and can be used to study various hydrochemical processes, such as base cation exchange, cement pollution, mixing of natural waters, sulfate reduction, saline water (end product water), and other related hydrochemical problems.

The chemical analysis data of all the samples collected from the area adjoining the River Yamuna in Delhi have been plotted on Chadha's diagram (Fig. 2) and results have been summarized in Table 3. It is evident that during the premonsoon season, most of the samples fall in Group 7 (Na-K-Cl-SO_4) followed by Group 8 (Na-K-HCO_3 type) and Group 6 (Ca-Mg-Cl-SO_4). An almost similar trend was observed during the postmonsoon season.

Impact of River Water Quality on Groundwater

Surface waterbodies play an important role in groundwater flow. The infiltration of surface water to groundwater usually occurs in a recharge geographical area; base flow from groundwater to surface waterbodies may occur in a discharge geographical area. In a discharge area, the hydraulic head increases with depth, and net saturated flow is upward toward the water table, but in a recharge area, the water table lies at a considerable depth beneath a thick unsaturated zone. The relationship of surface water

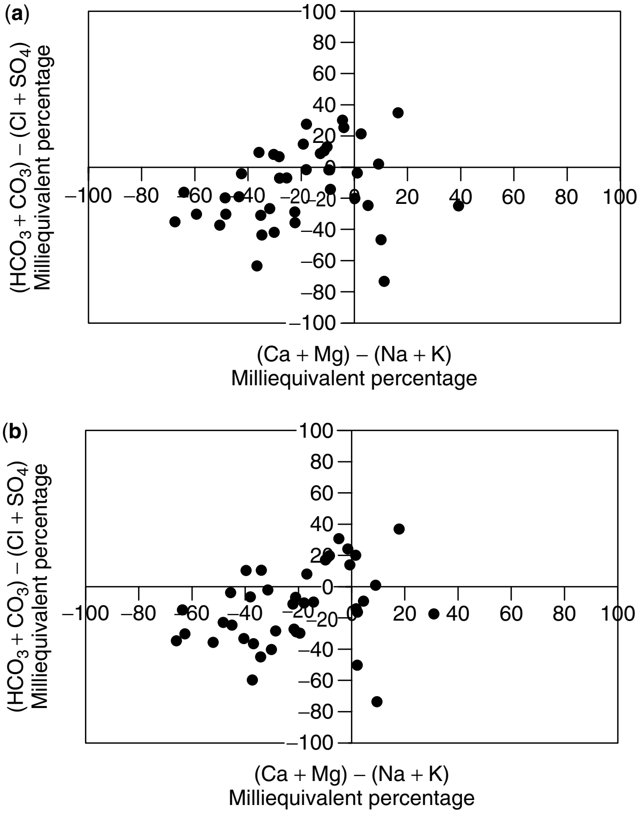

Figure 2. (**a**) Chadha's diagram for hydrochemical classification (premonsoon 2000); (**b**) Chadha's diagram for hydrochemical classification (postmonsoon 2000).

to groundwater and its recharge/discharge characteristics may change seasonally or once during a longer time span. In deep groundwater aquifers, the movement of water from a recharge to a discharge area may take place during several years, but in shallow aquifers, recharge and discharge may be much closer and even adjacent to each other. The hydraulic gradient plays a significant role in lateral and vertical migration of contaminants in groundwater aquifers.

During the present study, 38 groundwater samples were collected within 10 km of the eastern and western banks of the River Yamuna to ascertain the impact of river water on groundwater. The qualitative analysis of data showed higher concentrations of various physicochemical and bacteriological parameters on the western side of River Yamuna, even in deep aquifers. However, due to a paucity of hydrochemical, geologic, and water level data, no specific inferences could be drawn. Further studies are being planned to investigate the impact of Yamuna River water quality on the groundwater system.

CONCLUSIONS

The suitability of groundwater in the area adjoining the River Yamuna in Delhi has been examined per BIS and WHO standards. The quality of the groundwater varies from place to place with the depth of the water table. It also shows significant variation from one season to another. Only about 10% of the total samples analyzed were within the desirable limit of 500 mg/L for TDSs, and more than 80% of the samples were above the desirable limit but within the maximum permissible limit of 2000 mg/L. From the viewpoint of hardness, more than 80% of the samples were well within the desirable limits for domestic use during both pre- and postmonsoon seasons. More than 50% of the samples had a nitrate content of less than 45 mg/L, whereas in about 20% of the groundwater samples, the nitrate content even exceeded the maximum permissible limit of 100 mg/L. The higher level of nitrate at certain locations may be attributed to the surface disposal of domestic sewage and runoff from agricultural fields. The grouping of samples according to their hydrochemical facies clearly indicates that the majority of the samples fall in Na-K-Cl-SO_4 followed by Na-K-HCO_3 and Ca-Mg-Cl-SO_4 hydrochemical facies. The qualitative analysis of data showed higher concentrations of various physicochemical and bacteriological parameters on the western side of River Yamuna, even in deep aquifers. However, due to the paucity of hydrochemical, geologic, and water level data, no specific inferences could be drawn. More detailed studies including contaminant transport modeling studies are needed to understand better the impact of Yamuna River water quality on the groundwater aquifer.

BIBLIOGRAPHY

1. Jain, C.K. and Sharma, M.K. (2001). *Groundwater Quality in Adjoining Areas of River Yamuna at Delhi*. Technical Report, National Institute of Hydrology, Roorkee, India.
2. APHA. (1992). *Standard Methods for the Examination of Water and Waste Waters*, 18th Edn. American Public Health Association, Washington, DC.
3. Jain, C.K. and Bhatia, K.K.S. (1988). *Physico-chemical Analysis of Water and Wastewater*, User's Manual, UM-26. National Institute of Hydrology, Roorkee, India.
4. BIS. (1991). *Specifications for Drinking Water*, IS:10500:1991, Bureau of Indian Standards, New Delhi, India.

Table 3. Summarized Results of Water Classification

	Sample Numbers	
Classification/Type	Premonsoon 2000	Postmonsoon 2000
Na-K-HCO_3	1,2,3,4,6,7,8,14,20,21	1,2,3,6,7,13,20,21
Ca-Mg-HCO_3	5,13	4,5
Ca-Mg-Cl-SO_4	16,17,19,29,30,33	16,17,19,28,29,33
Na-K-Cl-SO_4	9,10,11,12,15,18,22,23,24,25,26,27,28, 31,32,34,35,36,37,38	8,9,10,11,12,14,15,18,22,23,24,25,26, 27,30,31,32,34,35,36,37,38

5. WHO. (1996). *Guidelines for Drinking Water, Recommendations*, Vol. 2, World Health Organization, Geneva.
6. Chadha, D.K. (1999). A proposed new diagram for geochemical classification of natural waters and interpretation of chemical data. *Hydrogeol. J.* **7**(5): 431–439.
7. Piper, A.M. (1944). A graphical procedure in the geochemical interpretation of water analysis. *Trans. Am. Geophys. Union.* **25**: 914–923.

CHLORINE RESIDUAL

Linda S. Andrews
Mississippi State University
Biloxi, Mississippi

Chlorine is a disinfectant added to drinking water to control microbial contamination. Both commercial water bottlers and public water systems use chlorine for this purpose. Levels added are calculated to maintain a disinfectant (chlorine) residual throughout the distribution system from the treatment facility to the end user.

Historically, the disinfection and sanitation of water, food products, and food processing equipment have used halogen-containing disinfectants such as chlorine (1). The worldwide use of chlorine has improved our quality of life and prevented many disease outbreaks caused by waterborne pathogens. Chlorine was first used as a disinfectant in 1897 to treat polluted water mains after an outbreak of typhoid fever in England, and later, in 1912, it was used after a typhoid fever outbreak in Niagara Falls, New York (2). Chlorine was introduced to the food processing industries in 1935. Today, disinfection is considered to be the primary mechanism for the inactivation/destruction of many pathogenic organisms to prevent the spread of waterborne diseases to downstream users and the environment (3).

Three common methods of disinfection exist in the United States and Canada: chlorination, ozonation, and ultraviolet (UV) disinfection. Chlorine, the most widely used disinfectant for municipal wastewater, destroys target organisms by oxidation of the cellular material (3). Chlorine may be applied as chlorine gas, hypochlorite solutions, or other chlorine compounds in solid or liquid form. Hypochlorous acid present in aqueous chlorine solutions is the biocidal “active chlorine” (4).

Advantages for use of chlorination include the following (3):

- Chlorination is an established method for disinfection.
- Chlorination is inexpensive compared with other methods (UV and ozone).
- Chlorine residuals remaining in the wastewater effluent will enhance the long-term effects of disinfection after the initial treatment and can be measured to determine the effectiveness.
- Chlorine is very effective against a wide variety of microorganisms, especially enteric pathogens.
- Chlorine has the ability to oxidize certain undesirable organic and inorganic compounds.
- Levels of chlorine can be adjusted according to biological load.
- Oxidative capacities of chlorine can eliminate off odors during disinfection.

In recent years, some health concerns have developed over the discovery of potentially carcinogenic by-products generated during chlorination in the presence of organic material. These by-products, termed “disinfection by-products,” include such compounds as trihalomethanes, chloroform, and chlorophenols. Tests of chlorine-treated water have identified more than 250 compounds, approximately 10% of which have been classified as potentially mutagenic and/or carcinogenic (5). Both disinfectants and by-products can have adverse health effects (6).

Other disadvantages for use of chlorination include the following (3):

- Chlorine residuals, even in low concentrations, can be toxic to aquatic life and thus may require a dechlorination process.
- Shipping, storage, and handling of chlorine chemicals can pose a safety risk to workers and to the environment.
- Levels of dissolved solids are increased in chlorine-treated effluent.
- Chlorine residuals are unstable in the presence of heavy biological load to the chlorination system, requiring increased dose applications.
- Certain parasitic species known to cause problems in water supplies are relatively resistant to chlorine oxidation treatments. These species include the oocysts of *Cryptosporidium parvum* and the cysts of *Entamoeba histolytica* and *Giardia lamblia* as well as the eggs of some parasitic worms.
- Other long-term effects of chlorination and the presence of chlorine residuals are yet unknown.

Continuous chlorination is a necessity for surface water supplies from lakes, springs, ponds, or cisterns. The effectiveness of chlorine as a disinfectant is a function of contact time, the chlorine solution used, the temperature of the water and environment, and the level of residual chlorine from first introduction to the water system and continuing to the end user. Continuous chlorination typically uses a chlorine residual of 3–5 ppm. Municipalities, however, use lower levels of 0.2–0.5 ppm because they represent larger distribution systems that provide a longer contact time. Also, the higher levels of chlorine residual may cause an objectionable flavor and odor (7).

Several types of chlorine residuals can be measured in treated water. Free chlorine residual is the measure of the disinfectant safety margin and is in the form of hypochlorous acid. The combined chlorine residual includes chlorine as chloramines and chlororganics that are produced when chlorine reacts with ammonia products that may be present because of biological load. The total chlorine residual is the sum of free and combined residuals (7). A growing number of water systems have naturally occurring ammonia and organic contaminants in their water supplies. These contaminants react with their chlorine treatment to form the combined chlorine residuals (8). These combined residuals require up to 100 times the contact time or 25 times the chlorine concentration to be effective disinfectants in water (9).

It is estimated that 3–4% of the total chlorine produced in the United States is used in potable water treatment, wastewater treatment, swimming pools, and cooling water biocide applications. Many health risks are associated with water that has not been disinfected, and health risks are associated with total residual chlorine. Residual chlorines are produced from water treatment facilities and in processing plants. Large power producing plants that use chlorinated water to prevent biofouling of heat exchangers in once-through cooling towers are also a major source of residual chlorine. Approximately 90% of such plants chlorinate cooling water on a periodic basis as compared with a continuous basis for water and wastewater treatment. Typically, chlorine is dosed at 1–2 mg/L for 20–30 minutes two to three times daily. To minimize chlorine residuals in once-through cooling systems, chlorinated effluent from one group of condensers can be blended with nonchlorinated cooling water, or targeted chlorination can be practiced (10).

DECHLORINATION

Chlorination has been used widely for disinfection of wastewater before discharge. Before the passage of the 1972 Federal Water Pollution Control Act and for several years after, significant levels of residual chlorine were routinely discharged into the environment. As the impact of the toxic by-products became known, dechlorination was instituted to remove residual chlorine from wastewater before discharge into sensitive aquatic waters (11). Dechlorination effectively removes free or total combined chlorine residuals remaining after chlorination. A common method of dechlorination is to expose the effluent to sulfur dioxide or sulfite salts. Although expensive, carbon absorption has also been used when total dechlorination is desired. State regulatory agencies set the policy for allowable levels of chlorine residuals, which effectively makes dechlorination essential for certain industries. Dechlorination protects aquatic life from the toxic effects of residual chlorine and prevents formation of harmful chlorinated compounds in drinking water. However, chemical dechlorination can be difficult to control with the potential for significant overdosing with sulfite, which suppresses the dissolved oxygen content and lowers the pH (11).

BIBLIOGRAPHY

1. Biocide Int., Inc. (1994). *ECO-Benefits of Oxine-A Chlorine Dioxide Based Antimicrobial*. Biocide Intl., Norman, OK.
2. Sconce, J.S. (1962). *Chlorine-Its Manufacture, Properties and Uses*. R.E. Dreiger Publ., Malabar, FL, pp. 28–37; 514–537.
3. Solomon, C., Casey, P., Mackne, C., and Lake, A. (1998). *Chlorine Disinfection*, a Technical Overview. Environmental Technology Initiative, U.S. EPA.
4. Andrews, L.S., Key, A., Martin, R., Grodner, R., and Park, D. (2002). Chlorine dioxide wash of shrimp and crawfish an alternative to aqueous chlorine. *Food Micro.* **19**: 261–267.
5. Sen, A.C., Owuso-Yaw, J., Wheeler, W., and Wei, C. (1989). Reactions of aqueous chlorine and chlorine dioxide with tryptophan, *n*-methyltryptophan and 3-indolelactic acid: kinetic and mutagenicity studies. *J. Food Sci.* **54**: 1057–1060.
6. *Federal Register* 66FR 16858. March 28, 2001.
7. Quality Water Solutions, Chlorination. (2004). Available at: http://www.qh20.com/chlorination.
8. Spon, R. (2002). *Do You Really Have a Free Chlorine Residual: How to Find Out and What You Can Do About It*. RR Spon & Associates, Roscoe, IL.
9. Disinfection with Chlorine. (2004). Available at: http://www.geocities.com.
10. Lew, C.C., Mills, W., and Loh, J. (1999). Power plant discharges of total residual chlorine and trihalomethanes into rivers: Potential for human health and ecological risks. *Hybrid Methods Eng.* **1**: 19–36.
11. USEPA (2000). *Wastewater Technology Fact Sheet, Dechlorination. EPA 832-F-00-022.*

SOURCE WATER QUALITY MANAGEMENT

LAUREL PHOENIX
Green Bay, Wisconsin

Water utilities serving water authorities or municipalities can have many levels of control over the quality of their source waters. The more control they have over protecting the water quality of their sources, the less money they have to spend on water treatment, and the lower their risk when dealing with unforeseen emergencies. When they have no control of changing water quantities from drought, flood, or industrial pollution, they must spend more money to access backup supplies. Here, the phrase “water utilities” will be used interchangeably with water authorities or municipalities to refer to any local government entity providing public drinking water.

HIGH CONTROL OF SURFACE WATER SOURCES

Water authorities with the highest control own the entire watershed of a surface water source and protect the water quality by keeping people and development out. Others may rely on a river surrounded on all sides by a pristine forest with little or no human activity (e.g., hiking, camping) and would try to extract their water at a point just below this pristine area rather than farther downstream where other land uses could degrade the water. For utilities that cannot practicably deny access to the entire watershed or the waterbody itself, strict guidelines for land use in that area can be locally imposed instead. These guidelines can include such things as restricting the amount of impervious surfaces that can be added through development (e.g., buildings, driveways, roads), forbidding dumping of hazardous materials or use of pesticides, or restricting recreational activities on land and water to prevent animal waste, sediment, or fossil fuels from watercraft from contaminating the water supply. Some states grant water authorities extraterritorial land use control over land in a different municipality in order to create and enforce restrictions that will protect the water supply.

HIGH CONTROL OF GROUNDWATER SOURCES

Groundwater sources can be highly protected if the utility has total control over land use in the aquifer's recharge zone. This type of control occurs when the utility owns all the land in the recharge zone, or where state or local laws require recharge zones to be fenced with signs directing people to stay out. Some states or local governments require this absolute protection of recharge zones, no matter who owns the land or what municipality has political jurisdiction. When total denial of human access or development is impossible, guidelines like the ones aforementioned can be mandated for any land lying over the recharge zone. The distance from the land surface to the water table and the substrate geology can then guide how strict the guidelines are; more restrictions being placed on land parcels overlying more vulnerable aquifers.

LITTLE CONTROL OF WATER SOURCES

Few water utilities have the high degree of control previously mentioned. This is because protecting surface or groundwater supplies can only be done through land use controls, and land use decisions have traditionally been the right of each local government. Therefore, most utilities may have to use a variety of strategies to deal with issues over which they have no control. This section discusses four common problems: (1) when source water is degraded, (2) when supply is uncertain, (3) when drinking water regulations are tightened, or (4) when emergency situations arise.

Degraded Source Water Quality

Surface waters can easily be polluted from increased drainage area development, farm runoff, or heavy rains. Development in the watershed increases erosion and causes sediment loading to surface waters. Farm runoff includes pesticides, sediment, fertilizer or manure, and silage leachate (nitrogen-rich liquid that drips from stored grains in silos). Heavy rains mobilize nonpoint source pollutants from urban landscapes, mines, farms, or silvicultural areas into water bodies. In fact, this "first flush" of runoff into water bodies carries high concentrations of pollutants. This multitude of diverse pollutants—sediment, chemicals, metals, pet waste, and so on—causes different kinds of problems for water treatment plants.

Groundwater can be polluted for the same reasons just stated, because, except for sediment, these nonpoint source pollutants can pollute groundwater by percolating down to the water table. Groundwater can also be degraded as the water table drops, because the quality of the water from that level may be higher in total dissolved solids (TDSs), be briny, or have arsenic or radon. Groundwater can also become unfit for use if compounds such as MTBE, perchlorate, solvents, or methane and diesel fuel from hydraulic fracturing migrate into the aquifer.

Even desalting ocean water does not protect utilities from pollution problems. As coastal waters become more polluted, the costs of the reverse osmosis plants increase. Therefore, source water pollution from poor land use practices, type of geological substrate, or industrial pollutants often leaves municipalities vulnerable to forces they cannot control.

Variability in Surface and Groundwater Supplies—Quantity

Surface water supply quantities can vary for several reasons. First, precipitation patterns can vary from drought conditions that reduce the amount of water available to very stormy conditions that cause flooding and subsequently degrade those floodwaters with everything they come in contact with. In states with riparian water law, the drought conditions will typically cause all utilities using the same water source to equally share the burden of insufficient water. Second, well pumping next to a river or lake can reduce surface water levels for anyone next to or downstream of the pumping. If state water law has no conjunctive use rules to prevent well owners from unfairly "stealing" water from surface supplies, then a utility may receive less water than it expects. Third, a utility may lose access to a water source if it was leased and will not be renewed when the lease is up. And finally, in states with prior appropriation water law, a utility must consider the problems that come along with junior water rights on a stream or lake. In times of drought, a very junior right on a stream may mean it gets little or no water from that stream. If they have water rights on several streams, the different adjudication dates for each stream and the amount of water in each stream will help it calculate which streams will be able to give them how much "wet" water that year. In prior appropriation law, "first in time is first in right." The year the right is obtained determines the order of use, so owners of earlier (and older) water rights will get their turn to use their full amount of allotted water before someone with a more junior right may divert their water.

Groundwater can become inaccessible the moment other well-pumpers over that aquifer draw down the water table to a level below the water utility's well.

Consequently, the uncertainty of water supplies due to the weather, leasing versus owning a water right, or state water law rules means that utilities best serve their public by having numerous water sources available to them. In this way, diverting water from an alternate or several alternate sources protects them from inadequate water available in drought times or highly degraded water resulting from flood. If they do as Phoenix does and purchase a "water ranch" (a local ranch with attached groundwater rights), then they are buying not only a water source, but a clean and nonevaporative storage site as well.

New Regulation

If new federal drinking water regulations are promulgated with lower allowable concentrations of some pollutant, then a utility's water supply becomes automatically degraded with respect to the new regulations. Arsenic is a good example of a pollutant that recently had its maximum allowable concentrations in drinking water reduced. Utilities either must stop using that water, spend enormous amounts to clean that pollutant out of the water, or mix that water with water from another source to "dilute" it down to the new, acceptable drinking water standard.

Planning for Emergency Situations

Preparing for emergencies by planning strategies is always good to do whenever possible. Water source emergencies can be a pollutant spill (e.g., fuel truck falls off highway into a river, accidental release of industrial materials into a river) or a pipeline break. These types of emergencies can only be efficiently dealt with when a utility has several sources from which to draw or sources designated as emergency backup supplies.

THE FUTURE OF SOURCE WATER QUALITY MANAGEMENT

Blending different source waters to upgrade quality, having multiple sources to give a utility options for accessing better quality or larger amounts, or instituting conservation to "create" new water are only a few strategies that utilities must employ to prepare for the future and attenuate rising costs.

Future strategies to manage source waters will need a foundation of federal protection first. Not only is a significant strengthening of the Water Quality Act and Clean Water Act necessary to reduce discharges into water sources, but also much higher, and therefore effective, fines for industrial spills or NPDES permit violations need to be mandated and enforced. After this foundation is laid, there are a few additional strategies that utilities can use. Cooperative interlocal agreements can result in larger, regional water treatment systems that are more efficient. A community downstream of its surface water source or having its groundwater recharge area under a neighboring community can craft agreements to protect the supplies. For example, one community can pay another to reduce development over a recharge area by buying some of their land outright or convincing them to only use conservation subdivision designs with reduced impervious surfaces and installation of water gardens. An upstream community can be paid to improve its stormwater pollution controls or reduce land use practices that create nonpoint source pollution. For source water quality management to improve in the future, it needs not only more cooperative interlocal strategies recognizing land use regulations as the key problem to address, but also effective water pollution controls at the federal level.

DOSE-RESPONSE OF MUSSELS TO CHLORINE

S. Rajagopal
G. van der Velde
Radboud University Nijmegen
Toernooiveld, Nijmegen, The Netherlands

V.P. Venugopalan
BARC Facilities
Kalpakkam, India

H.A. Jenner
KEMA Power Generation and Sustainables
Arnhem, The Netherlands

Chlorine, like any other biocide, produces a dose-dependent response in organisms. However, the response is dependent on a number of parameters. Chlorination of cooling water leaves residuals and by-products that are potential pollutants in the receiving water body and can impact nontarget organisms. Therefore, it is imperative to generate data on the chlorine dose–response relationships of important fouling organisms, such as various mussel species, for efficient but environmentally acceptable biofouling control. Data are presented to show the effects of mussel size (shell length), season of sample collection (spawning vs. nonspawning season), nutritional status (fed vs. nonfed), and acclimation temperature (5–30 °C) on the mortality pattern of different mussel species under continuous chlorination ($0.5–5\ \mathrm{mg\,L^{-1}}$). From the data, it can be concluded that although various factors can influence the dose–response relationships of mussels, generalization is not possible because of species specificity. Among the various parameters, mussel species, mussel size, status of byssal attachment, spawning season, and acclimation temperature have significant effect on chlorine tolerance of mussels, whereas nutritional status shows very little effect.

INTRODUCTION

Chlorination is the most commonly used disinfectant and biocide for treatment of drinking and industrial cooling water (1,2). Depending on the organisms involved and the end result desired, chlorine may be dosed either intermittently or continuously (3,4). A third important criterion that often has a bearing on the applied dose is the concentration of the chlorine (and its reaction products) that reaches the consumer (in the case of drinking water chlorination) or the environment (in the case of cooling water chlorination) (2,5). A fourth factor is chlorine demand, which exhibits variations that are both seasonal and geographical (1). Accordingly, in an actual chlorination program, the chlorine dose is modulated in such a way that the measured level of chlorine residuals is sufficient to bring about the desired result, whether it be microbiological control in drinking water or biofouling control in cooling water (6,7).

For any given organism, one can generate a dose−response curve on the basis of how the organism responds to a given dose. However, in the case of chlorine, it is appropriate to represent this in the form of a concentration−response curve, because dosing of a certain amount of chlorine results in a variable quantity of chlorine that is available for biocidal action. In other words, the available quantity of chlorine may vary depending on the chlorine demand of the water (8).

FACTORS INFLUENCING THE CHLORINE TOXICITY

The response of an organism to chlorine would vary depending on a number of factors. Among the various factors, the type of organism is an important criterion. Shelled organisms (such as mussels) can withstand relatively long-term exposure to chlorinated water (9,10), when compared with soft-bodied organisms, such as hydroids or ascidians (11). A review of literature clearly indicates all other fouling organisms succumb faster than mussels. Therefore, chlorine regime targeted against mussels would also eliminate other fouling organisms. In the present review, chlorine dose−response relationships of mussels are discussed in relation with various factors, including mussel species, mussel size, spawning season, byssus attachment, nutritional status, and acclimation temperature, which can influence the chlorine tolerance of mussels.

MUSSEL SPECIES

Efficacy of chlorine as an antifoulant depends on various parameters, most importantly residual levels of chlorine and contact time (6,9). A survey of existing literature shows that at residual levels commonly employed ($1\ \mathrm{mg\,L^{-1}}$) in power station cooling circuits, mortality takes several days. For example, at $1\ \mathrm{mg\,L^{-1}}$ continuous chlorination, *Mytilus edulis* (blue mussels) takes about 480 h for 100% mortality (12). On the other hand, in tropical marine mussels, *Perna viridis* (green mussels) takes about 816 h for 100% mortality when $1\ \mathrm{mg\,L^{-1}}$ residual chlorine is applied continuously (13). In *Dreissena polymorpha* (zebra mussel), 95% mortality is observed after about 552 h exposure to $1\ \mathrm{mg\,L^{-1}}$ residual chlorine (14). In comparison, *Mytilopsis leucophaeata* (dark false mussel) takes about 1104 h to achieve 100% mortality at $1\ \mathrm{mg\,L^{-1}}$ residual chlorine (Fig. 1). The exposure time required for 100% mortality of *M. leucophaeata* at different chlorine concentrations are much higher than that required for *D. polymorpha* (588 h) and *Mytilus edulis* (966 h).

MUSSEL SIZE

In the case of common fouling organisms such as mussels and barnacles, it is often seen that the size (or age) of the organism is an important factor that influences its sensitivity to chlorine. It has been shown that, for several organisms, a size-dependent variation in the response exists, with larger organisms showing increased tolerance (Fig. 1). However, such size-dependent nature of toxicity is not universal and there are organisms which exhibit uniform sensitivity to chlorine, irrespective of size. An example for mussel size influencing tolerance to chlorine is the mussel *M. leucophaeata* (9). Here, the tolerance is maximum in medium-size mussels (about 10 mm), whereas smaller (2 mm) and larger (20 mm) mussels show greater sensitivity (Fig. 1). This pattern is in contrast with the results reported for *Mytilus edulis*, where the tolerance linearly increases with shell size (15). In the case of *D. polymorpha*, shell size has no effect on the sensitivity of the organism to chlorine (10,16). Therefore, the relationship between mussel size and chlorine toxicity is not similar among different mussel species, and generalizations regarding the size effect should be made after careful observation.

BYSSUS ATTACHMENT

Mussels use byssus threads to attach themselves to hard substrata. The status of attachment is also an important criterion that determines the response of mussels to chlorine (15,17). It has been experimentally shown that mussels, which normally are attached with the help of their byssus threads, become more sensitive to chlorine when they are exposed to it under unattached condition (Fig. 2). Once detached from a substratum, the mussel tries to reattach itself by producing new byssus threads, for which it has to open its bivalve shell and extend its “foot” outside. This kind of enhanced byssogenic activity increases the exposure of the soft tissues of the mussel to chlorine, thereby increasing the toxic effect. On the other hand, attached mussels are byssogenically less active, and in a chlorinated environment their shells remain mostly closed, thereby protecting their soft body from chlorine (12,18).

SPAWNING SEASON

Physiological status of the organism is also an important factor that influences chlorine toxicity. Research using a number of organisms has shown that chlorine toxicity is significantly higher during breeding seasons than during nonbreeding seasons. Mussels collected during the spawning season and those collected during nonspawning season behave quite differently with respect to their sensitivity to chlorine (Fig. 3). Mussel species collected during their spawning season were less tolerant to chlorine, whereas those collected during the nonspawning season were more tolerant. The difference in tolerance between the two groups was nearly 29% (9). Kilgour and Baker (16) and Jenner et al. (1), who reported similar results for *D. polymorpha,* attribute the greater tolerance of mussels during the nonspawning season to low metabolic rates and reduced filtration rates, which would result in reduced exposure to the toxicant. Lower energetic demands during nonbreeding seasons may be the reason for reduced toxicant uptake. On the other hand, mussels tend to be weaker after spawning when they have little energy reserves in the body (19), with the result that

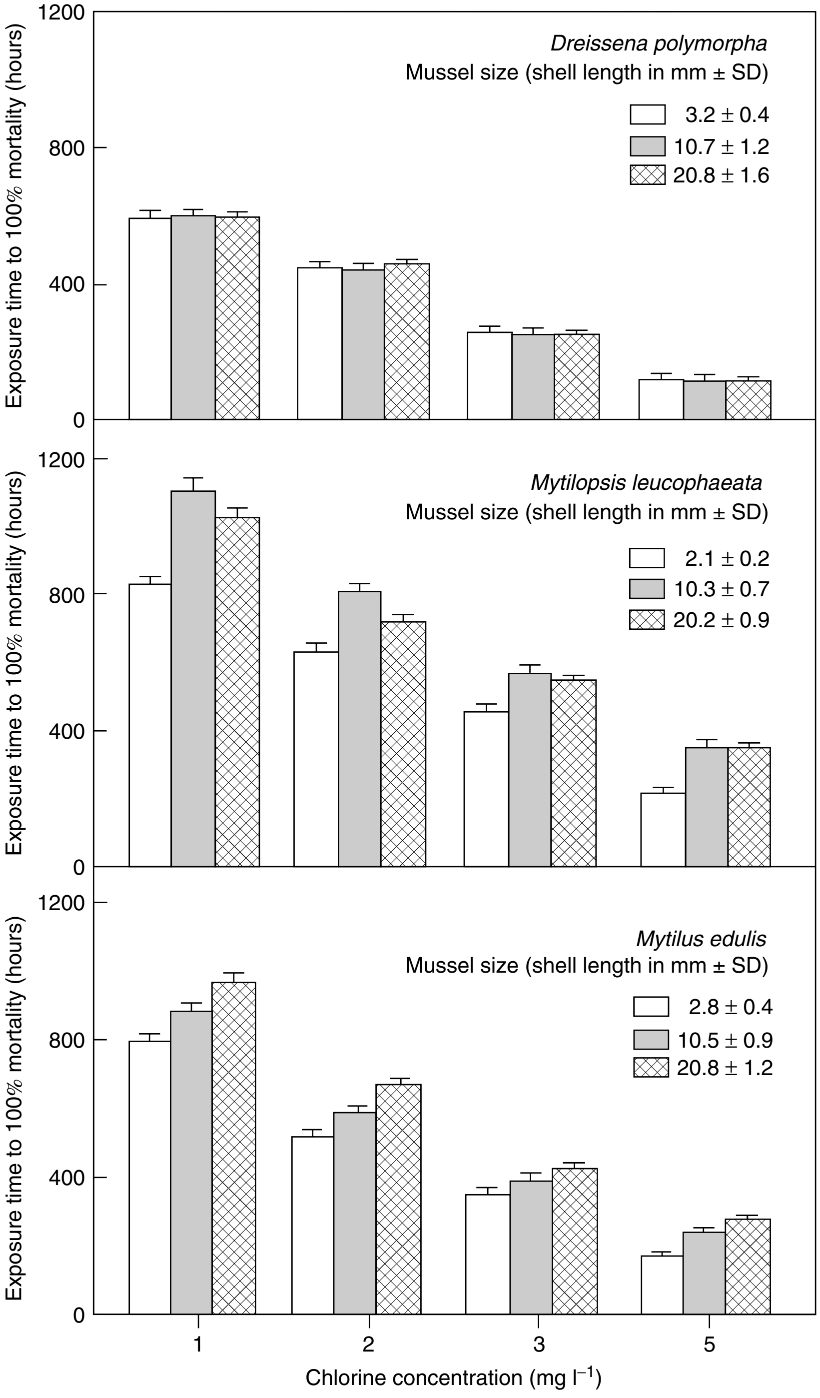

Figure 1. Comparison of exposure times to reach 100% mortality of different size groups of *Mytilopsis leucophaeata, Dreissena polymorpha*, and *Mytilus edulis* at different chlorine concentrations. Mortality data are expressed as mean ±SD ($n = 80$) of four replicate experiments ($n = 20$ individuals in each experiment). Test methods and mortality determinations were similar in all toxicity studies of species.

they are less tolerant to biocide. The data point to the importance of judicious sampling while carrying out toxicity experiments using seasonal breeders.

FED VS. NONFED MUSSELS

Status of feeding may have an effect on the toxicity of chlorine to organisms. Kilgour and Baker (16) showed that mussels *D. polymorpha*, when maintained on a diet of *Chlorella*, were consistently more sensitive to hypochlorite than starved mussels. The effect was attributed to an increased tendency of the fed mussels to filter water, which incidentally increases the exposure of their body parts to chlorine. Mussels that are fed with microalgae are likely to filter more water than those that are unfed. On the other hand, Rajagopal et al. (13) showed that, in the case of the mussel *Perna viridis*, fed and starved individuals showed similar mortality rates when exposed to chlorine. It must be kept in mind that chlorine may act as a strong suppressant of filtration activity in bivalve mussels, which has been shown by Rajagopal et al. (18) using Mussel-Monitor®, an automated instrument with which one can

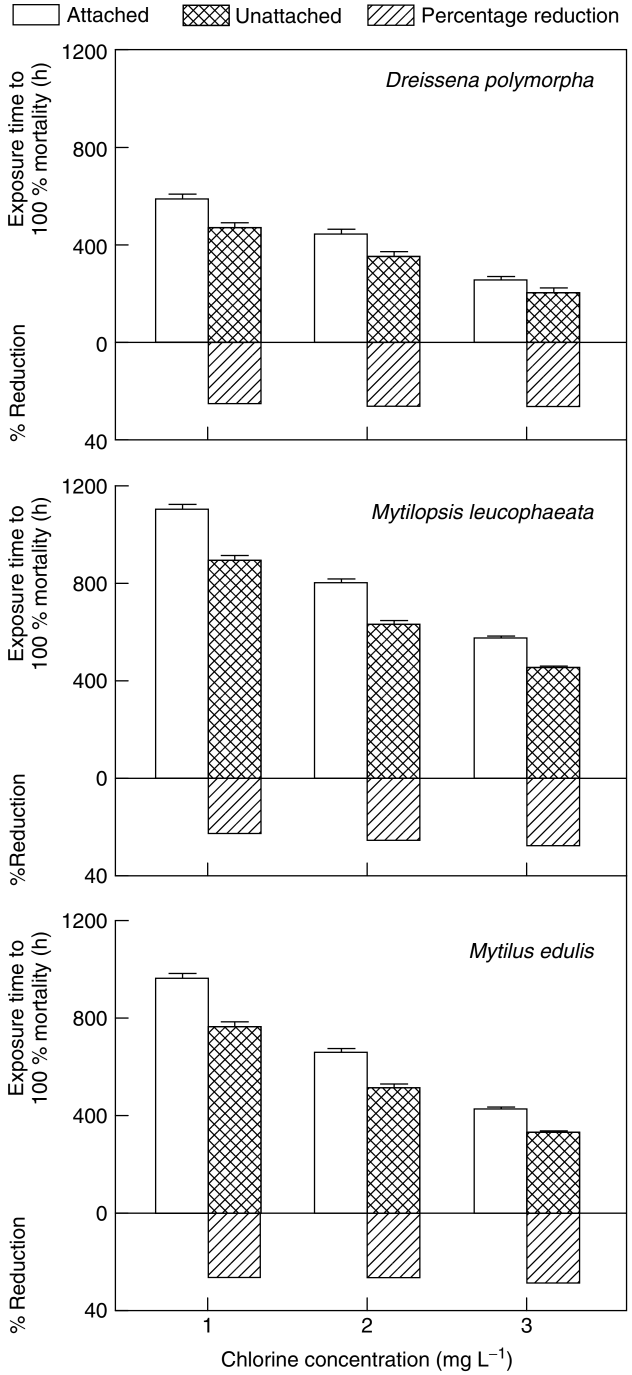

Figure 2. Time difference (%) between byssally attached and unattached *Dreissena polymorpha* (17), *Mytilopsis leucophaeata*, and *Mytilus edulis* (15) for 100% mortality at different chlorine concentrations.

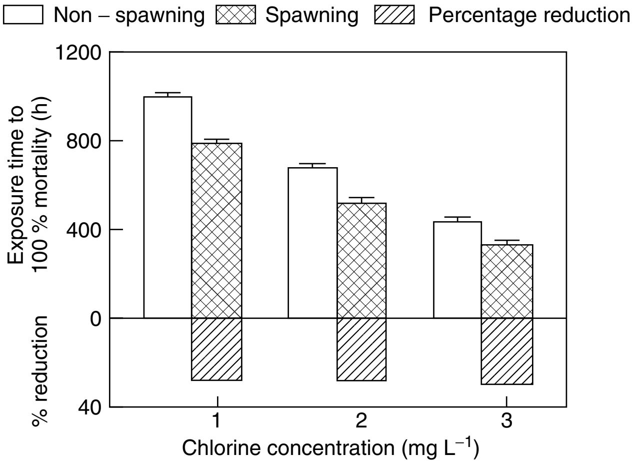

Figure 3. Cumulative mortality (%) of spawning and nonspawning *Mytilopsis leucophaeata* at different chlorine concentrations (TRC = total residual chlorine). Eighty mussels were used at each chlorine dose.

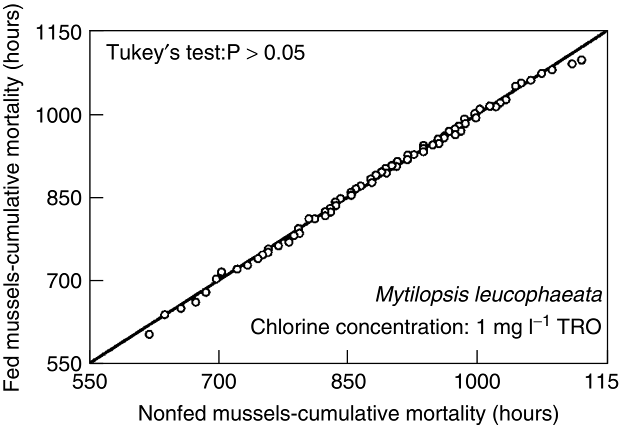

Figure 4. Cumulative mortality (%) of fed and nonfed *Mytilopsis leucophaeata* at different chlorine concentrations (TRO = total residual oxidant).

monitor the opening and closing of mussel shells (1,4). Shell valve movement of *M. leucophaeata* tested with unfiltered brackish water from the Noordzeekanaal in The Netherlands showed little or no filtration in presence of 1 mg L^{-1} residual chlorine. Obviously, presence of microalgae would have no significant effect on *M. leucophaeata* at a residual chlorine concentration of 1 mg L^{-1} (Fig. 4). Rajagopal et al. (20) have also shown that filtration activity in *D. polymorpha* stops almost completely at residual chlorine level of 0.5 mg L^{-1} and higher.

ACCLIMATION TEMPERATURE

Temperature is yet another important factor that influences the sensitivity of organisms to chlorine. Mussels acclimated to different temperatures show significantly different tolerances to chlorine (Fig. 5). Decrease in acclimation temperature from 30 °C to 5 °C increases chlorine tolerance (0.5 mg L^{-1} residual chlorine) of *M. leucophaeata* by 52 days. Such increase in chlorine tolerance at lower acclimation temperatures has also been reported for other mussel species, such as *D. polymorpha* (11,14) and *Mytilus edulis* (1,12). However, at acclimation temperatures above 35 °C, temperature has overriding effects when compared with chlorine. Harrington et al. (21) showed that at 36 °C, combined use of temperature and chlorine resulted in mortality

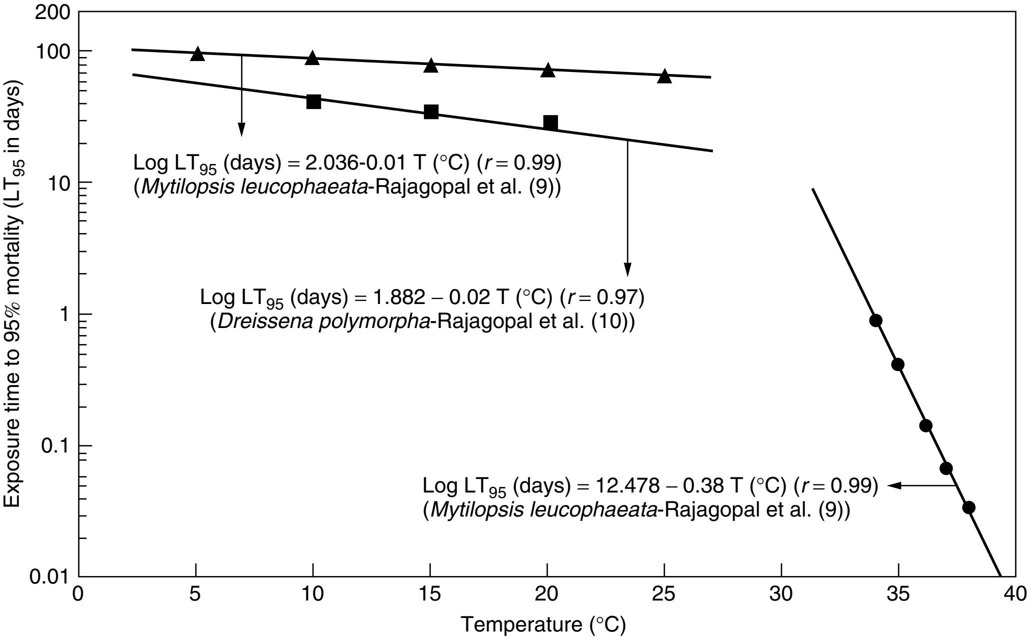

Figure 5. Comparison of exposure times to reach 95% mortality of *Mytilopsis leucophaeata* and *Dreissena polymorpha* at 0.5 mg L^{-1} residual chlorine depending on the acclimation temperature. Triangles and circles: data of Rajagopal et al. (9); Rectangles: data of Rajagopal et al. (10). Lines are linear regressions.

of *D. polymorpha* at rates similar to that obtained with heat alone.

CONCLUSION

Data available in the literature show that various factors can influence the sensitivity of organisms estimated using chlorine bioassays. Among the parameters, mussel species, mussel size, spawning season, acclimation temperature, and status of attachment seem to have significant influence on chlorine tolerance. Therefore, chlorine bioassays using mussels or such organisms need to be carried out after taking the above factors into consideration.

Acknowledgments
We thank M.G. Versteeg and M. Van der Gaag for assistance with field studies. KEMA Power Generation and Sustainables, Arnhem and Schure-Beijerinck-Popping Fonds, Amsterdam, The Netherlands financially supported this research.

BIBLIOGRAPHY

1. Jenner, H.A., Whitehouse, J.W., Taylor, C.J.L., and Khalanski, M. (1998). Cooling water management in European power stations: biology and control. *Hydroecol. Appl.* **1–2**: 1–225.
2. White, G.C. (1999). *Handbook of chlorination and alternative disinfectants*. John Wiley & Sons, Hoboken, NJ, p. 1569.
3. Claudi, R. and Mackie, G.L. (1994). *Practical Manual for Zebra Mussel Monitoring and Control*. Lewis Publishers, London, p. 227.
4. Rajagopal, S., Van der Velde, G., Van der Gaag, M., and Jenner, H.A., (2003). How effective is intermittent chlorination to control adult mussel fouling in cooling water systems? *Water Res.* **37**: 329–338.
5. Allonier, A.S., Khalanski, M., Camel, V., and Bermond, A. (1999). Characterization of chlorination by-products in cooling effluents of coastal nuclear power stations. *Mar. Pollut. Bull.* **38**: 1232–1241.
6. Mattice, J.S. and Zittel, H.E. (1976). Site-specific evaluation of power plant chlorination. *J. Water Pollut. Cont. Fed.* **48**: 2284–2308.
7. Rajagopal, S. (1997). *The ecology of tropical marine mussels and their control in industrial cooling water systems*. Ph.D. thesis, University of Nijmegen, The Netherlands, p. 184.
8. Rajagopal, S., Azariah, J., Nair, K.V.K., Van der Velde, G., and Jenner, H.A. (1996). Chlorination and mussel control in the cooling conduits of a tropical coastal power station. *Mar. Environ. Res.* **41**: 201–221.
9. Rajagopal, S., Van der Gaag, M., Van der Velde, G., and Jenner, H.A. (2002). Control of brackish water fouling mussel, *Mytilopsis leucophaeata* (Conrad) with sodium hypochlorite. *Arch. Environ. Contam. Toxicol.* **43**: 296–300.
10. Rajagopal, S., Van der Velde, G., Van der Gaag, M., and Jenner, H.A. (2002). Laboratory evaluation of the toxicity of chlorine to the fouling hydroid *Cordylophora caspia*. *Biofouling* **18**: 57–64.
11. Rajagopal, S., Van der Velde, G., and Jenner, H.A. (2002). Effects of low-level chlorination on zebra mussel, *Dreissena polymorpha*. *Water Res.* **36**: 3029–3034.
12. Lewis, B.G. (1985). *Mussel control and chlorination*. Report No TPRD/L/2810/R85, Central Electricity Research Laboratories, Leatherhead, Surrey, England, p. 33.
13. Rajagopal, S., Venugopalan, V.P., Van der Velde, G., and Jenner, H.A. (2003). Tolerance of five species of tropical marine mussels to continuous chlorination. *Mar. Environ. Res.* **55**: 277–291.

14. Van Benschoten, J.E., Jensen, J.N., Harrington, D.K., and DeGirolamo, D. (1995). Zebra mussel mortality with chlorine. *J. Am. Water Works Assoc.* **87**: 101–108.
15. Rajagopal, S., Van der Velde, G., Van der Gaag, M., and Jenner, H.A. (2004). Byssal detachment underestimates tolerance of mussels to toxic compounds. *Mar. Pollut. Bull.* (in press).
16. Kilgour, B.W. and Baker, M.A. (1994). Effects of season, stock, and laboratory protocols on survival of zebra mussels (*Dreissena polymorpha*) in bioassays. *Arch. Environ. Contam. Toxicol.* **27**: 29–35.
17. Rajagopal, S., Van der Velde, G., and Jenner, H.A. (2002). Does status of attachment influence survival time of zebra mussel, *Dreissena polymorpha* exposed to chlorination? *Environ. Toxicol. Chem.* **21**: 342–346.
18. Rajagopal, S., Van der Velde, G., and Jenner, H.A. (1997). Shell valve movement response of dark false mussel, *Mytilopsis leucophaeta*, to chlorination. *Water Res.* **31**: 3187–3190.
19. Bayne, B.L., Thompson, R.J., and Widdows, J. (1976). Physiology. In: *Marine Mussels: Their Ecology and Physiology*. B.L. Bayne (Ed.). Cambridge University Press, Cambridge, pp. 121–206.
20. Rajagopal, S., Van der Velde, G., Van der Gaag, M., and Jenner, H.A., (2002). Sublethal responses of zebra mussel, *Dreissena polymorpha* to low-level chlorination: an experimental study. *Biofouling* **18**: 95–104.
21. Harrington, D.K., Van Benschoten, J.E., Jensen, J.N., Lewis, D.P., and Neuhauser, E.F. (1997). Combined use of heat and oxidants for controlling adult zebra mussels. *Water Res.* **31**: 2783–2791.

METALLOTHIONEINS AS INDICATORS OF TRACE METAL POLLUTION

Trifone Schettino
Maria Giulia Lionetto
Università di Lecce
Lecce, Italy

INTRODUCTION

Metallothioneins (MTs) are low molecular weight, cysteinrich metal-binding proteins that are involved in detoxification and homeostasis of heavy metals. Since their discovery in the horse kidney by Margoshes and Vallee (1), metallothioneins have been further identified in most living organisms, vertebrates, invertebrates, algae, fungi, and plants. Today, the term “metallothionein” is used to designate a series of well-known molecules showing a large degree of structural and functional similarities to those first described for horse kidney metallothionein. A nomenclature system for metallothionein was adopted in 1978 (2) and then extended by introducing a subdivision of all MTs into three classes (3). By this convention Class I includes metallothioneins with locations of cysteines closely related to those in the horse kidney metallothionein; Class II comprises metallothionein with locations of cysteines only distantly related to those in horse kidney metallothionein; and Class III subsumes low molecular weight metalloisopolypeptides containing gammaglutamyl-cysteinyl units resembling in their features mammalian MTs. They are called phytochelatin and occur predominantly in plants and fungi. As the number of known MT sequences has grown, this subdivision has become inadequate. In order to better differentiate the several known MTs, in 1999 (4,5) a new classification system was proposed based on sequence similarities and phylogenetic relationships. This system subdivides the MT superfamily into families, subfamilies, subgroups, and isolated isoforms and alleles. As before, the metallothionein superfamily is defined phenomenologically as comprising all polypeptides that resemble equine renal metallothionein in several of their features (2,3). Such general features are low molecular weight, high metal content, characteristic amino acid composition (low content of aromatic amino acid residues, high cystein content, which accounts for MT heavy metal affinity and binding capacity), unique amino acid sequence with characteristic distribution of Cys (i.e., Cys–X–Cys, where X stands for an amino acid residue other than cysteine), and spectroscopic manifestations characteristic of metal thiolate clusters, which provide the protein with a highly stable tertiary structure. The metal affinity for the binding sites follows the general order found for inorganic thiolates: Hg(II) > Ag(I) > Cu(I) > Cd(II) > Zn(II). The MT superfamily is subdivided into several families, each of them including MTs that share a particular set of sequence specific characters. Members of a family can belong to only one family and are thought to be evolutionarily related. The inclusion of a MT in a family presupposes that its amino acid sequence is alignable with that of all members. A common and exclusive sequence pattern, a profile, and a phylogenetic tree can therefore be connected with each family. Each family is identified by its number and its taxonomic range. To date, 15 MT families are known: Family 1, vertebrate MTs; Family 2, mollusc MTs; Family 3, crustacean MTs; Family 4, echinodermata MTs; Family 5, diptera MTs; Family 6, nematoda MTs; Family 7, ciliata MTs; Family 8, fungi-I MTs; Family 9, fungi-II MTs; Family 10, fungi-III MTs; Family 11, fungi-IV MTs; Family 12, fungi-V MTs; Family 13, fungi-VI MTs; Family 14, prokaryota MTs; and Family 15, planta MTs.

BIOLOGICAL FUNCTION OF METALLOTHIONEINS

The ubiquitous distribution of MTs in virtually all types of organisms studied to date attests to the conserved nature of MTs and their function. The biological function of MTs is likely related to the physiologically relevant metals that these proteins bind. In mammals, MT is found to bind zinc and copper under normal physiological conditions. Both zinc and copper are trace metals that are essential for life. Recent studies have produced strong evidence to support the idea that MT functions as a metal chaperone for the regulation of gene expression and for synthesis and functional activity of proteins, such as metalloproteins and metal-dependent transcription factors (6–10). MT could thus serve as a reservoir of essential metals.

MTs are inducible proteins. Exposure of the organisms to high levels of heavy metals (e.g., Zn, Cu, Cd, and Hg) and the following increase of heavy metal cations in the cells stimulates metalloprotein neosynthesis by enhancing MT gene transcription. *Cis*-actin sequences, termed metal response elements (MREs), located in multiple copies along the promoter, allow heavy metal ion induction of MT transcripts (11,12). The MT mRNA is translated by cytosolic free ribosomes, leading to an increase of apometallothioneins that will rapidly react with free metal cations, sequestering them and protecting cell structures from nonspecific interaction with heavy metal cations. This process circumvents cellular damage preventing metal toxicity under overload conditions. MT induction can be measured as the concentration or rates of formation of the responsible mRNA, MT, and levels of MT-bound metals. Each of them provides different information on the inductive process and may display differential dynamics.

METALLOTHIONEINS AS BIOMARKERS

The importance of metallothionein (MT) in toxicologic responses to heavy metals was early recognized for potential application as a "biomarker" of organism exposure to heavy metals in aquatic environments. A biomarker is a pollutant-induced variation in cellular or biochemical components or processes, structures, or functions that is measurable in a biological system or sample (13). The biomarker approach in environmental monitoring has been increasingly used in the last 20 years for the ecotoxicological assessment of aquatic ecosystems. Since the harmful effects of pollutants are typically manifested at lower levels of biological organization before disturbances are realized at the population, community, or ecosystem levels (14), the use of biomarkers measured at the cellular level has been proposed as sensitive "early warning" tools for biological effect measurement in environmental quality assessment (15).

One aspect of environmental degradation in aquatic environments is pollution from heavy metals, which are persistent and accumulable by aquatic organisms. A great number of metal pollution events have been reported in fresh waters, coastal waters, and groundwaters, where abnormal metal levels occurred as the consequence of natural and, especially, of manufactured sources of pollution. A crucial aspect of heavy metal pollution, determining the actual ecological risk, is the bioavailability of heavy metals in aquatic environments. Bioavailable heavy metals represent that portion of the total environmental metal load that is of direct ecotoxicological relevance (16). Apart from chemical analytical techniques, some other new approaches have recently been used to determine the availability of toxic metals for living organisms in aquatic environments. One of the most relevant contributions in this field is that provided by MT induction determination.

Due to the interest in MT application in environmental monitoring, different techniques and methodologies have been developed in the last few years for the quantification of total MTs, including chromatographic separation of soluble cytosolic MT-containing fraction associated with the evaluation of the metal concentration, HPLC-AAS (17,18), HPLC-ICP (19–21), metal substitution assays (22,23), radioimmunological techniques (24–26), electrochemical analysis (27,28) and a spectrophotometric method recently developed for routine application in biomonitoring (29). Molecular biology techniques are employed for the analysis of MT mRNA. However, it is important to point out that the expression of MT levels is not implicit from the presence of endogenous MT mRNA levels in tissues.

Thanks to the growing knowledge about MT quantification, induction of MT following heavy metal exposure has been reported in different aquatic species and tissues (30,31), especially in common bioindicator species such as *Mytilus galloprovincialis, Mytilus edulis, Littorina littorea, Ostrea edulis, Crassostreä* virginic*a*, *Dreissena polymorpha*, and *Macoma balthica*.

Recent field studies using *Mytilus galloprovincialis* as a bioindicator have also demonstrated that measurements of MTs can provide an accurate indication of subtle environment increases in metal contamination (32,33), confirming their usefulness as a biomarker of trace metal exposure in aquatic environmental monitoring.

Therefore, these proteins have been proposed by the European Commission and other international scientific organizations to be included in environmental monitoring programs as a biomarker to assess metal pollution in aquatic environments.

It is known that biotic and abiotic factors such as seasonal variation, sex, age, size of the animal, and dietary factors affect MT levels in the aquatic organisms (30,34). Such factors are likely to interfere with MT synthesis in response to metal occurrence (toxic or excess of essential metals) in the aquatic ecosystems. Therefore, it is very important to know how these factors affect MT expression in the utilized bioindicator species before MT can be used in a monitoring program.

Ideally, information obtained by quantification of MT in indicator species could be used to assess adverse effects on the species themselves, on other components of the ecosystem and on humans. This information could be used in routine biomonitoring programs as early warning of potential human health effects, to make decisions about possible cleanup, and to evaluate the efficacy of past cleanup of hazardous waste.

BIBLIOGRAPHY

1. Margoshes, M. and Vallèe, B.L. (1957). *J. Am. Chem. Soc.* **79**: 4813–4814.
2. Nordberg, M. and Kojima, Y. (1979). *Experientia Suppl.* **34**: 48–55.
3. Fowler, B.A., Hildebrand, C.E., Kojima, Y., and Webb, M. (1987). *Experientia Suppl.* **52**: 19–22.
4. Binz, P.A. and Kägi, J.H.R. (1999). In: C. Klaassen (Ed.). Birkhäuser Verlag, Basel, pp. 7–13.
5. Kojima, Y., Binz, P.A., and Kägi, J.H.R. (1999). In: C. Klaassen (Ed.). Birkhäuser Verlag, Basel.
6. Maret, W. (1995). *Neurochem. Int.* **27**: 111–117.
7. Zeng, J., Vallee, B.L., and Kagi, J.H. (1991). *Proc. Natl. Acad. Sci. U.S.A.* **88**: 9984–9988.
8. Zeng, J., Heuchel, R., Schaffner, W., and Kagi, J.H. (1991). *FEBS Lett.* **279**: 310–312.
9. Jacob, C., Maret, W., and Vallee, B.L. (1998). *Proc. Natl. Acad. Sci. U.S.A.* **95**: 3489–3494.
10. Maret, W., Larsen, K.S., and Vallee, B.L. (1997). *Proc. Natl. Acad. Sci. U.S.A.* **94**: 2233–2237.

11. Stuart, G.W., Searle, P.F., and Palmiter, R.D. (1985). *Nature* **317**: 828–831.
12. Hamer, D.H. (1986). *Annu. Rev. Biochem.* **55**: 913–951.
13. NRC (National Research Council) (1987). *Environ. Health Perspect.* **74**: 3–9.
14. Adams, S.M. (1990). *Am. Fish. Soc. Symp.* **8**: 1–8.
15. McCarthy, F. and Shugart, L.R. (1990). Lewis Publishers, Chelsea, MI.
16. Rainbow, P.S. (1995). *Mar. Pollut. Bull.* **31**: 183–192.
17. Suzuki, K.T. (1980). *Anal. Biochem.* **160**: 160–168.
18. Lehman, L.D. and Klaassen, C.D. (1986). *Anal. Biochem.* **102**: 305–314.
19. Sunaga, H., Kobayashi, E., Shimojo, N., and Suzuki, K.T. (1987). *Anal. Biochem.* **160**: 160–168.
20. Mason, A.Z., Storms, S.D., and Jenkins, K.D. (1990). *Anal. Biochem.* **186**: 187–201.
21. Mazzucotelli, A., Viarengo, A., Canesi, L., Ponzano, E., and Rivaro, P. (1991). *Analyst* **116**: 605–608.
22. Scheuhammer, A.M. and Cherian, M.G. (1986). *Toxicol. Appl. Pharmacol.* **82**: 417–425.
23. Eaton, D.L. (1982). *Toxicol. Appl. Pharmacol.* **66**: 134–142.
24. Nolan, C.V. and Shaikh, Z.A. (1986). *Anal. Biochem.* **154**: 213–223.
25. Roesijadi, G., Morris, J.E., and Unger, M.E. (1988). *Can. J. Fish. Aquat. Sci.* **45**: 1257–1263.
26. Hogstrand, C. and Haux, C. (1990). *Toxicol. Appl. Pharmacol.* **103**: 56–65.
27. Olafson, R.W. and Sim, R.G. (1979). *Anal. Biochem.* **100**: 343–351.
28. Thompson, J.A.J. and Cosson, R.P. (1984). *Mar. Environ. Res.* **11**: 137–152.
29. Viarengo, A., Ponzano, E., Dondero, F., and Fabbri, R. (1997). *Mar. Environ. Res.* **44**: 69–84.
30. Bordin, G. (2000). *Cell. Mol. Biol.* **46**(2): special issue.
31. Langston, W.J., Bebianno, M.J., and Burt, G. (1998). In: W.J. Langston and M.J. Bebianno (Eds.). Chapman and Hall, London, pp. 219–284.
32. Bebianno, M.J. and Machado, L.M. (1997). *Mar. Pollut. Bull.* **34**: 666–671.
33. Lionetto, M.G. et al. (2001). *Aquat. Conserv. Mar. Freshwater Ecosys.* **11**: 305–310.
34. Serafim, M.A. and Bebianno, M.J. (2001). *Environ. Toxicol. Chem.* **20**: 552–554.

AMPHIPOD SEDIMENT TOXICITY TESTS

Augusto Cesar
Maria del Carmen Casado Martínez
Inmaculada Riba López
Tomás Ángel Del Valls Casillas
Universidad de Cadiz
Cadiz, Spain

INTRODUCTION

General

Amphipod sediment toxicity tests are technically developed and are widely accepted as useful tools for a wide variety of research and regulatory purposes (1–3). For example, they can be used to determine the sediment toxicity of single chemicals and chemical mixtures, the chemical bioavailability, the potential adverse effects of dredged material, and the magnitude and spatial and temporal distribution of pollution impacts in the field (4).

Various methods have been developed to evaluate sediment toxicity and these procedures range in complexity from lethal to sublethal tests that measure effects of chemical mixtures on the amphipod species. The evaluated sediment phase may include whole sediment, suspended sediment, elutriates, or sediment extracts (5–7).

The test organisms include amphipods, algae, macrophytes, fishes, and other benthic, epibenthic, and pelagic invertebrates (8). However, amphipod toxicity tests can provide rapid and effective information on the potential effects of contaminants in sediments.

Historical Background

Historically, the assessment of sediment quality has often been limited to chemical characterizations (9). However, quantifying contaminant concentrations alone cannot always provide enough information to adequately evaluate potential adverse effects that arise from interactions among chemicals, or that result from time-dependent availability of sediment-associated contaminants to aquatic organisms (7).

The evaluation of contaminant has primarily emphasized surface waters and effluents, not sediments, and the first incentive for sediment testing was dredged material (7). In 1977, the U.S. EPA and the U.S. Army Corps of Engineers recommended a series of 10-d toxicity and bioaccumulation tests with amphipods, clams, polychaetes, shrimps, and fishes to evaluate proposed discharge of dredged material into estuarine and marine waters.

Significance of Use

Sediment provides habitat for many benthic organisms and is a major repository for many of the more persistent chemicals that are introduced into surface waters. In the aquatic environment, most anthropogenic chemicals and waste materials including toxic organic and inorganic chemicals eventually accumulate in sediment (3).

The objective of an amphipod sediment toxicity test is to determine whether contaminants in sediment are harmful to amphipod species. The tests can be used to measure interactive toxic effects of complex contaminant mixtures in sediment.

The purpose of the amphipod solid-phase toxicity test is to determine if test sediment samples reduce survival (growth, reproduction, etc.) of exposed organisms relative to that of organisms exposed to control and reference sediment. Test results are reported as treatment (station) or combination of treatments (sites or chemicals) that produce statistically significant reduced survival (growth, reproduction, etc.) from control or reference sediments.

Sediment amphipod tests (3) can be used to (1) determine the relationship between toxic effects and bioavailability, (2) investigate interactions among contaminants, (3) compare the sensitivities of different organisms,

(4) determine spatial and temporal distribution of contamination, (5) evaluate hazards of dredged material, (6) measure toxicity as part of product licensing or safety testing or chemical approval, (7) rank areas for clean up, and (8) set cleanup goals and estimate the effectiveness of remediation or management practices.

Scope and Application

Procedures are described for testing amphipod crustaceans in the laboratory to evaluate the toxicity of contaminants associated with sediments. Sediments may be collected from the field or spiked with compounds in the laboratory. A toxicity method is outlined for diverse species of estuarine, marine, and freshwater sediment amphipods found within coastal and fresh waters. Generally, the methods described may be applied to all species, although acclimation procedures and some test conditions (temperature, salinity, etc.) will be species specific. Procedures described here are principally based on References 1–3. Although it is recognized that a variety of other nonstandardized toxicity test methods are used in ecotoxicologic research, emphasis is placed on standardized protocols provided by the U.S. EPA and ASTM, because these are the tests most commonly used in regulatory applications (10). Other countries such as The Netherlands have adapted and developed standardized protocols using local species for regulatory purposes (11).

Selection of Test Organisms

The choice of a test organism has a major influence on the relevance, success, and interpretation of a test. Test organism selection should be based on both environmental relevance and practical concerns (3). The species should be selected based on sensitivity to contaminant behavior in sediment and feeding habitat, ecological relevance, geographic distribution, taxonomic relation to indigenous organisms, acceptability for use in toxicity assessment (e.g., a standardized method), availability, and tolerance to natural geochemical sediment characteristics (7).

Amphipods have been used extensively to test the toxicity of marine, estuarine, and freshwater sediments (1–3). Ideally, a test organism should have a toxicological database demonstrating relative sensitivity to a range of contaminants of interest in sediment, have a database for interlaboratory comparisons of procedures, be in direct contact with sediment, be readily available year-round from culture or through field collection, be easily maintained in the laboratory, be easily identified, be ecologically or economically important, have a broad geographical distribution, be indigenous (either present or historical) to the site being evaluated or have a niche similar to organisms of concern (e.g., similar feeding guild or behavior to the indigenous organisms), be tolerant of a broad range of sediment physicochemical characteristics (e.g., grain size), and be compatible with selected exposure methods and endpoints.

The sensitivity of an organism is related to route of exposure and biochemical response to contaminants (3). Generally, benthic organisms can receive exposure via from three primary sources: interstitial water, whole sediment, and overlying water. Because benthic communities contain a diversity of organisms, many combinations of exposure routes may be important (3). Therefore, behavior and feeding habits of a test organism can influence its ability to accumulate contaminants from sediment and should be considered when selecting test organisms for sediment testing (3). Table 1 lists some commonly used amphipod species for sediment toxicity testing and some useful information when selecting the proper amphipod species for sediment toxicity assessment.

Amphipods

Amphipods are ecologically important members of benthic infaunal communities and are a primary food resource for a number of marine invertebrate, fish, and bird species worldwide. In general, crustacea are among the most sensitive members of benthic communities to anthropogenic disturbance, including pollution (10).

Table 1. Some Commonly Used Amphipod Species for Sediment Toxicity Testing

Amphipods	Test and Point(s)[a]	Test Period, d	Habitat[b]	References
Fresh water				
Diporeia sp.	S	10–28	B, I	1,12
Hyalella azteca	S, G, R	10–28	B, E	1,13
Salt water				
Ampelisca abdita	S, G, R	10	T, I	2,14
Ampelisca brevicornis	S, G, R	10–28	T, I	15,16
Corophium voluntator	S, I	10	T, I	17,18
Eohaustorius estaurius	S	10	B, I	2,14
Gammarus aequicauda	S	10	E	19,20
Gammarus locusta	S, R	10–28	E	21,22
Grandidierella japónica	S, G	10	T, I	2,14
Hyalella azteca	S, G, R	10–28	B, E	2,23
Leptocheirus plumulosus	S, G, R	10–28	B, I	2,24
Microdeutopus gryllotalpa	S	10	T, I	25,26
Repoxynius abronius	S	10	B, I	2,14
Tiburonella viscana	S	10	B, I	27,28

[a] S = survival, G = growth, I = immobilization, R = reproduction.

[b] B = burrow, E = epibenthic, I = infaunal, T = tube dweller.

Amphipods are a group of small crustaceans that can live in very different habitats. They are found in the sea, in estuaries, in continental waters, and even in certain terrestrial wetlands (29). Their biogeographical distribution embraces the polar waters all the way to the tropics. As for their bathymetric distribution, they extend from humid atmospheres of some forests all the way to abyssal depths. At the present time, more than 6000 species have been registered worldwide (29). Amphipods are widely distributed and common in unpolluted lotic and lentic systems; however, they are less common in hydrodynamics zones and are a primary food source for fish and voracious feeders of animal, plant, and detrital material (30).

Infaunal amphipods are excellent organisms for toxicity tests with sediment (Fig. 1) and are strongly recommended as appropriate test species for toxicity bioassays (14,30,31). Amphipods are often chosen for ecological and ecotoxicological studies due to their ecological relevance, sensitivity to environmental disturbance, and amenability for culture and experimentation (2,21,32). Overall, infaunal amphipods are excellent bioassay organisms for toxicity tests with whole sediment (31).

METHOD DESCRIPTION AND EXPERIMENTAL DESIGN

The test system described by Swartz et al. (14) for the amphipod *Rhepoxynius abronius* is recommended for bioassays with this and other amphipod species. This section describes a general laboratory method to determine the toxicity of contaminated sediments using marine, estuarine, and freshwater amphipod crustaceans, following compiled standards procedures (1–3).

Test sediments may be collected from marine, estuarine, and freshwater environments or spiked with compounds in the laboratory. The toxicity test usually is conducted in 1-L glass chambers containing 2 cm (175 mL) of sediment and 800 mL of overlying seawater (1:4 ratio of sediment to water). For 10-d acute tests the exposure is static, and the organisms are not fed during this period. Survival is most frequently used as the endpoint in studies, and reburial of surviving amphipods is an additional measurement that can be used. However, in chronic exposure, renewal systems can be static and organisms are fed over the 28-d exposure period. The endpoints commonly used are growth and reproduction.

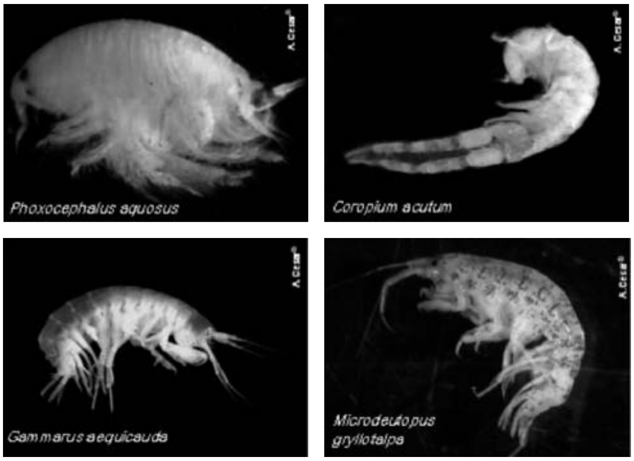

Figure 1. Some species used in amphipod sediment toxicity tests.

A summary of general parameters and conditions to follow when developing an amphipod sediment toxicity test in the laboratory is included in Table 2.

Quality Control

If more than 10% mean mortality occurs in the control, the test must be repeated (3). However, in chronic tests 20% mean mortality in the control is acceptable. Unacceptably

Table 2. Summary of Some Conditions to Develop the Test Using Amphipods in the Laboratory

Parameter	Conditions
Test type	Whole sediment, suspended sediment, elutriates, or sediment extracts, static or renewal (depends on test type)
Temperature	Species dependent (10–25 °C)
Salinity	Species dependent (2–38 psu)
Light quality	Broad-spectrum fluorescent lights
Illuminance	500–1000 lux
Photoperiod	24 light: 0 dark
Test chambers	1-L glass beakers, recommended 10 cm Ø
Sediment volume	175–200 mL (1:4 sediment/water ratio)
Overlying water volume	600–800 mL (1:4 sediment/water ratio)
Water renewal	Not necessary or renewal (depends on test type)
Size and life stage of amphipods	Species dependent, 2–6 mm (no mature males or females)
Number of organisms per chamber	10–20
Number of replicates	At a minimum, 4 to 5 replicates must be used
Feeding regime	Species and test type dependent
Aeration	Trickle flow (<100 bubbles/min)
Overlying water	Clean natural or reconstituted water
Test exposure	10–28 d
Endpoints	Survival, growth, immobilization, reproduction, etc.
Test acceptability	Greater than or equal to 80–90% survival in controls

high control mortality indicates that the organisms are being affected by important stress factors other than contamination in the material being tested, and the test has to be reevaluated. These stresses may be due to injury or disease, unfavorable physical or chemical conditions, improper handling or acclimation, or possibly unsuitable sediment grain size. Species selection and the potential effects of these and other variables should be carefully reexamined in an attempt to reduce unacceptably high mortality.

Reference toxicant tests should be performed routinely in parallel on all populations of organisms in sample testing. When data for a particular reference toxicant have been generated on at least five populations of organisms of a species, two standard deviations above and below the mean are established as the bounds of acceptability. Reference toxicant tests should be conducted at least monthly on each species cultured in house and should be performed on each lot purchased or field collected organisms (3,33). Reference toxicant tests are most often acute lethality tests performed in the absence of sediment (3,7).

Statistical Analysis

The statistical comparisons and interpretations of the results should be appropriate to the experimental design. Study design depends on its objective and each study design has specific statistical design and analytical considerations.

Most amphipod sediment toxicity tests are based on the calculation of a medium effect or lethal and effective concentration or dose (LD_{50}, EC_{50}, IC_{50}, etc.) when sediments are spiked with individual contaminants, or upon the magnitude of response for field-collected sediments. However, the most useful information for predictive purposes is an estimation of the dose–response curve and the subsequent ranking of toxicity.

A variety of techniques are used to calculate these responses or endpoints: probit, moving angle average, and Spearman–Karber are examples. Most of these methods allow calculation of a confidence interval. These NOEC calculations depend on the number of replicates used in the toxicity test and the power of the ANOVA and other multiple comparison techniques used. When sediment samples are independently replicated, results can be statistically compared with control or reference sediments by *t*-test, analysis of variance (ANOVA), or regression analysis. An ANOVA is used to determine differences among treatments. When the assumptions of normality or homogeneity of variance are not met, transformations of the data may remedy the problem, so that the data can be analyzed by parametric procedures, rather than by nonparametric methods. Detailed statistical hypothesis testing, for making decisions about the appropriate method used, is discussed in Reference 3.

Confounding Factors

Interferences are characteristics of a sediment or sediment test system that can potentially affect test organism survival aside from those related to sediment-associated contaminants (3). These interferences can potentially confound interpretation of test results in two ways: (1) toxicity is observed in the test when contamination is not present, or there is more toxicity than expected; and (2) no toxicity is observed when contaminants are present at elevated concentrations, or there is less toxicity than expected.

There are three categories of interfering factors: those characteristics of sediments affecting survival independent of chemical concentration (i.e., noncontaminant factors); changes in chemical bioavailability as a function of sediment manipulation or storage; and the presence of indigenous organisms. For example, survival of some marine amphipods can be significantly affected by the grain size of the sediment (34) or ammonia concentration (35). Other confounding factors within sediment habitats include both abiotic (i.e., salinity, temperature, pH, dissolved oxygen, hydrogen sulfide, alkalinity, hardness, light) and biotic parameters (i.e., food limitation, predators, competitors, disease, biogenic disturbance).

Because of the heterogeneity of natural sediments, extrapolation from laboratory studies to the field can sometimes be difficult. Sediment collection, handling, and storage may alter the bioavailability and concentration by changing the physical, chemical, or biological characteristics of the sediment. Maintaining the integrity of a field-collected sediment during removal, transport, mixing, storage, and testing is extremely difficult and may complicate the interpretation of effects.

FUTURE RESEARCH

Research is continuing in the areas of chronic sediment toxicity methods, field validation of laboratory toxicity and bioaccumulation tests, toxicity identification and evaluations (TIE), and the multispecies experiments. Additional research is also needed to evaluate factors controlling the partitioning or sorption of a compound between water, colloids, and sediment, including aqueous solubility, pH, redox, affinity for sediment organic carbon and dissolved organic carbon, grain size of the sediment, sediment mineral constituents (oxides of iron, manganese, and aluminum), and the quantity of acid volatile sulfides in sediment. Additional research is also needed on aspects of fluid mechanics to quantify the physical processes of colloidal growth and transport and physical mass-transfer processes over a wide spectrum of solute–solid interfaces. The goal of this research would be to assimilate the biological, hydrological, geological, and ecological information necessary for the assessment and management of contaminated sediments. The end result of this research would be an improved understanding of the importance of contaminants in sediment relative to other factors influencing ecosystems.

Acknowledgments

A. Cesar thanks CAPES/MEC-Brasil for a postdoctoral scholarship (2558/03-3). M.C. Casado-Martínez has been funded by the Spanish Ministry of Education and Science (REN2002-01699/TECNO). This work was conducted under grants funded by the Spanish Ministry of Development (BOE 13-12-02) and of Education and Science (VEM2003-20563/INTER).

BIBLIOGRAPHY

1. American Society for Testing and Materials (1995). *Standard Test Methods for Measuring the Toxicity of Sediment-Associated Contaminants with Freshwater Invertebrates*. ASTME 1706–95a.
2. American Society for Testing and Materials (2000). *Standard Guide for Conducting 10-Day Static Sediment Toxicity Tests with Marine and Estuarine Amphipods*. ASTME 1367–99.
3. U.S. Environmental Protection Agency (1994). *Methods for Assessing the Toxicity of Sediment-Associated Contaminants with Estuarine and Marine Amphipods*. EPA/600/R-94/025.
4. Ferraro, S.P., and Cole, F.A. (2002). A field validation of two sediment–amphipod toxicity tests. *Environ. Toxicol. Chem.* **21**: 1423–1437.
5. Chapman, P.M. (1986). Sediment quality criteria for the sediment quality triad—an example. *Environ. Toxicol. Chem.* **5**: 957–964.
6. Swartz, R.C. et al. (1994). Sediment toxicity, contamination, and amphipod abundance at a DDT and dieldrin-contaminated site in San Francisco Bay. *Environ. Toxicol. Chem.* **13**: 949–962.
7. Ingersoll, C.G. (1995). Sediment tests. In: *Fundamentals of Aquatic Toxicology*. G.M. Rand (Ed.). Taylor & Francis, London.
8. Burton, G.A. Jr., Nelson, M.K., and Ingersoll, C.G. (1992). Freshwater benthic toxicity test. In: *Sediment Toxicity Assessment*. G.A. Burton Jr. (Ed.). Lewis, Ann Arbor, MI, pp. 213–240.
9. McCauley, D.J., DeGraeve, G.M., and Linton, T.K. (2000). Sediment quality guidelines and assessment: overview and research needs. *Environ. Sci. Pollut.* **3**: 133–144.
10. Anderson, B., Nicely, P., Gilbert, K., Kosaka, R., Hunt, J., and Phillips, B. (2004). *Overview of Freshwater and Marine Toxicity Tests*. California Environmental Protection Agency.
11. RIKZ (1999). *Standard Operating Procedure Specie-01 Marine Amphipod Corophium volutator Mortality Sediment Toxicity Test*. Riijksinstituut voor Kust en Zee. RIKZ/AB-99.114x, The Netherlands.
12. Lamdrum, P.F. (1989). Bioavailability and toxicokinetics of polycyclic aromatic hydrocarbons sorbed to sediments for the amphipod, *Pantoporeia hoyi*. *Environ. Sci. Technol.* **23**: 588–595.
13. Ingersoll, C.G. and Nelson, M.K. (1990). Testing sediment toxicity with *Hyalella azteca* (Amphipoda) and *Chironomus riparius* (Diptera). In: *Aquatic Toxicology and Risk Assessment*. W.G. Landis and W.H. van der Schalie (Eds.). ASTM, Washington, DC, pp. 93–109.
14. Swartz, R.C., DeBen, W.A., Jones, J.K.P., Lamberson, J.O., and Cole, F.A. (1985). Phoxocephalid amphipod bioassay for marine sediment toxicity. In: *Aquatic Toxicology and Hazard Assessment*. R.D. Cardwell, R. Purdy, and R.C. Bahner (Eds.). ASTM, Washington, DC, pp. 284–307.
15. Riba, I., DelValls, T.A., Forja, J.M., and Gómez-Parra, A. (2003). Comparative toxicity of contaminated sediments from a mining spill using two amphipod species: *Corophium volutator* and *Ampelisca brevicornis*. *Bull. Environ. Toxicol. Chem.* **71**: 1061–1068.
16. Riba, I., Zitko, V., Forja, J.M., and DelValls, T.A. (2003). Deriving sediment quality guidelines in the Guadalquivir estuary associated with the Aznalcóllar mining spill: a comparison of different approaches. *Cien. Mar.* **29**: 261–264.
17. van den Hurk, P., Chapman, P.M., Roddie, B., and Swartz, R.C. (1992). A comparison of North American and West European infaunal amphipod species in a toxicity test on North Sea sediments. *Mar. Ecol. Prog. Ser.* **91**: 237–243.
18. Brils, J.M. et al. (2002). Oil effect in freshly spiked marine sediment on *Vibrio fischeri, Corophium volutator* and *Echinocardium cordatum*. *Environ. Toxicol. Chem.* **21**: 2242–2251.
19. Cesar, A., Marín-Guirao, L.R., Vita, R., and Marín, A. (2002). Sensitivity of Mediterranean amphipods and sea urchins to reference toxicants. *Cien. Mar.* **28**: 407–417.
20. Cesar, A., Marín-Guirao, L.R., Vita, R., and Marín, A. (2004). Amphipod and sea urchin tests to assess the toxicity of Mediterranean sediments: the case of Portmán Bay. *Sci. Mar.* **68**: 205–213.
21. Neuparth, T., Costa, F.O., and Costa, M.H. (2002). Effects of temperature and salinity on life history of the marine amphipod *Gammarus locusta*. Implications for ecotoxicological testing. *Ecotoxicology* **11**: 61–73.
22. Costa, F.O., Correia, A.D., and Costa, M.H. (1996). Sensitivity of a marine amphipod to non-contaminant variables and to copper in the sediment. *Écologie* **27**: 269–276.
23. Nebeker, A.V. and Miller, C.E. (1988). Use of the amphipod crustacean *Hyalella azteca* in freshwater and esturine sediment toxicity tests. *Environ. Toxicol. Chem.* **7**: 1027–1033.
24. Schlekat, C.E., McGee, B.L., and Reinharz, E. (1991). Testing sediment toxicity in Chesapeake Bay using the amphipod *Leptocheirus plumulosus*: an evaluation. *Environ. Toxicol. Chem.* **11**: 225–236.
25. DelValls, T.A. and Chapman, P.M. (1998). The use of multivariate analysis to link the sediment quality triad components to site-specific sediment quality values in the Gulf of Cádiz (Spain) and in San Francisco Bay (USA). *Cien. Mar.* **24**: 313–336.
26. DelValls, T.A., Forja, J.M., and Parra, A.G. (1998). Integrated assessment of sediment quality in two littoral ecosystems from the gulf of Cádiz, Spain. *Environ. Toxicol. Chem.* **17**: 1073–1084.
27. Nipper, M. (2000). Current approaches and future directions for contaminant-related impact assessments in coastal environments: Brazilian perspective. *Aquat. Ecosys. Health Manage.* **3**(4): 433–447.
28. Abessa, D.M.S., Sousa, E.C.P.M., Rachid, B.R.F., and Mastroti, R.R. (1998). Use of the burrowing amphipod *Tiburonella viscana* as a tool in marine sediments contamination assessment. *Braz. Arch. Biol. Technol.* **41**: 225–230.
29. Ortiz, M. and Jimeno, A. (2001). Guía ilustrada para la identificación de las familias y los géneros de los anfípodos del suborden Gammaridea de la península Ibérica. *Graellsia* **57**: 3–93.
30. Burton, G.A. Jr. (1991). Assessing the toxicity of freshwater sediments. *Environ. Toxicol. Chem.* **10**: 1585–1627.
31. Swartz, R.C., DeBen, W.A., and Cole, E.A. (1979). A bioassay for the toxicity of sediment to marine macrobenthos. *Water Pollut. Control Fed.* **51**: 944–950.
32. DeWitt, T.H., Redmond, M.S., Sewall, J.E., and Swartz, R.C. (1992). *Development of A Chronic Sediment Toxicity Test for Marine Benthic Amphipods*. U.S. EPA, Chesapeake Bay Program.
33. Lee, D.R. (1980). Reference toxicants in quality control of aquatic bioassays. In: *Aquatic Invertebrate Bioassays*. A.L. Burkema Jr. and J. Cairms Jr. (Eds.). ASTM, Washington, DC, pp. 188–199.
34. De Witt, T.H., Ditsworth, G.R., and Swartz, R.C. (1988). Effects of natural sediments features on the phoxocephalid

amphipod, *Rheposynius abronius*: implications for sediment toxicity bioassays. *Mar. Environ. Res.* **25**: 99–124.

35. Ankley, G.T., Katko, A., and Arthur, J.W. (1990). Identification of ammonia as an important sediment-associated toxicant in the lower Fox River and Green Bay, Wisconsin. *Environ. Toxicol. Chem.* **9**: 313–322.

CILIATED PROTISTS AS TEST ORGANISMS IN TOXICITY ASSESSMENT

Ana Nicolau
Nicolina Dias
Nelson Lima
Centro de Engenharia Biológica
Braga, Portugal

TOXICITY ASSESSMENT, CYTOTOXICITY, AND TEST ORGANISMS

Increasing environmental pollution and the continuous development of new drugs have led to an ever growing concern about the potential effects of these compounds on human and/or environmental health, directly or indirectly. Toxicity data collected in the laboratory are the primary source of information that can be used for regulatory or risk assessment proposes.

In vitro cytotoxicity techniques are being increasingly developed and evaluated as alternatives to the use of vertebrates in testing environmental pollutants. Besides the reduction of laboratory animals, these lower cost tests allow for the use of specific endpoints to determine the targets of toxic effects with great precision and reproducibility (1). Because of the long-standing interest in aquatic toxicity testing, a myriad of test species, types of tests, and endpoints have been developed and used.

In recent years, much research has been done to assess the toxicity of various compounds in a series of biotests using several organisms. The appeal of these tests lies in their simplicity and high degree of reproducibility (2). On the other hand, test organisms for assessing environmental risk and impact must possess a lot of desirable features: they must be eukaryotic, their biology and general responses must be well known, the laboratory handling must be relatively easy, and a short generation time is desirable whenever studies of long-term effects are necessary (3).

Ciliated protists fulfill all those requirements. Furthermore, as they have a wide distribution and ecological significance, performing key functions in energy flow and elementary cycling in aquatic foodwebs, they can be ideal early warning indicators of aquatic ecosystem deterioration. Recently, the use of ciliates in toxicological tests has been investigated and their potential in standard bioassays has been demonstrated in aquatic environments and even elsewhere (4–6). Nevertheless, there are only a few works suggesting new test organisms in this group (7–9).

Toxicokinetics and toxicodynamics govern the response of exposed organisms (10). Single-celled organisms are useful test subjects because toxicokinetics is, in this case, very simplified and one can consider that it is proportional to water–lipid partitioning (11). Among protists, the ciliate *Tetrahymena*, and among the several species, *Tetrahymena pyriformis*, is the most commonly ciliated model used for laboratory research: it has been, for more than five decades, the organism of choice in analyses, evaluation of protein quality, and determination of effects of several toxic substances. It was the first protist to be cultivated axenically, that is, in a standard medium, free from bacteria or other organisms (12). This fact could explain the extensive use of this ciliate as a model cell system since the addition of a test compound can be, in principle, the only change in culture conditions (3). Therefore, under axenic culture conditions, standardization and guarantee of reproducibility of assays are obtained (9).

Beyond the rare reviews on this subject (3,13–16), some works demonstrate the potential of ciliates as test organisms, especially *Tetrahymena pyriformis* and *T. termophila*, but also *Colpidium campylum* and, more recently, *Spirostomum teres* and *S. ambiguum*.

Silverman (17) used the movement of *T. termophila*, in axenic cultures, to assess the toxicity of several compounds using the Draise test, concluding its validity. Later, Noever et al. (18) suggested another test to substitute for the Draise test, this time with *T. pyriformis*, using the mobility of the ciliate in only 30 s.

Yoshioka et al. (19) criticized the lack of rigorous establishment of experimental conditions in assays with *T. pyriformis* and validated a growth impairment assay of this ciliate in axenic culture, to assess the toxicity of a series of compounds, comparing the results with the calculation of LD_{50} of *Orizias latipe* and concluding for the validity of *T. pyriformis* in the assessment of toxicity in aquatic environments.

Dive et al. (7) presented, for the first time in a standardized form, a bioassay method using the ciliate *Colpidium campylum* as a test organism. Le Dû et al. (20) compared this bioassay to tests with *Daphnia* or the MICROTOX® test (Microbics) in the detection and modeling of the aquatic toxicity of soluble residuals, concluding for its validity as a complementary test or even a substitute for the other tests.

Nalecz-Jawecki et al. (8) developed a biotest using *S. ambiguum*, one of the biggest protists, 2–3 mm long, to assess water quality. Later, Nalecz-Jawecki and Sawicki (21) implemented a miniaturized and standardized bioassay based on the former and called it the SPIROTOX test. The development of the SPIROTOX test for the toxicity of volatile compounds enabled the creation of a database of toxicity of organic pollutants to *S. ambiguum* (22). Futhermore, comparison of the sensitivity of this test to other bioassay systems used worldwide indicated that the SPIROTOX test can be an important element of an assessment battery as the representative ciliated protists component (23,24).

Sauvant et al. (25) evaluated the toxicity of bottled water using two models in a series of *in vitro* assays: the mammalian fibroblasts L-929 and *T. pyriformis*. They concluded that both models are adequate and that *T. pyriformis* was more sensitive than the fibroblasts in samples bottled for longer periods.

Larsen et al. (26) proposed an international pilot ring test in order to standardize a growth inhibition test protocol with *T. pyriformis*. They concluded that the *Tetrahymena* test could be recommended for ecotoxicological screening purposes, being an ecologically relevant supplement to aquatic toxicity testing. In recent years, the *T. pyriformis* multigeneration growth impairment assay or TETRATOX (26,27) has been used quite often. It uses data from 40-h growth of a population of *T. pyriformis*.

Twagilimana et al. (9) suggested a new indicator, *Spirostomum teres*, a ciliated protist tolerant to low levels of dissolved oxygen; the assay was based on the calculation of the LD_{50} in nonaxenic cultures exposed to the toxic compounds for 24 h. The authors assured the high sensitivity of the assay, especially to heavy metals, in relation to the MICROTOX (Microbics) and to other conventional bioassays.

Schlimme et al. (28) suggested an assay to determine the toxicity of several bacteria strains: in the BACTOX test, these bacteria were given to *T. pyriformis* as prey. The authors highlighted the lack of toxicological assays with whole bacteria and, based on the results of their work, which seem to indicate that toxicity is specific to strains and not to species, they recommend that every bacteria strain should pass the BACTOX test before being produced in large scale.

TOXICOLOGICAL ASSESSMENT WITH PROTISTS

Growth rate and morphological changes have been used in the evaluation of toxicity for some decades. Population growth impairment is an often-used sublethal toxic endpoint for organic and inorganic compounds and does not require special technical expertise. It is considered the most sensitive sublethal parameter, reflecting the global state of a series of parcel effects. Other parameters such as cell motility, swimming patterns, cytoskeleton analysis, phagocytosis rate, and biochemical parameters can be assessed and have been proposed to determine the physiological and energetic state of protists when in contact with toxic compounds.

Tetrahymena pyriformis has been used extensively in studies on effects of heavy metals. Although the mechanisms by which heavy metals affect the microorganisms is not clear, authors have described copper as an essential metal that accumulated in the mitochondria (29) with its distribution observable as small electron-dense dots (3,30). Alteration of *Tetrahymena* morphology after aluminum exposure was reported by Sauvant et al. (31), who observed the presence of large and irregular-shaped monster cells in treated cultures. These morphological abnormalities were detected in zinc-treated cells but not in copper cultures. Most of the works focus on one metal only.

There are also several studies that use protists as test organisms to assess the toxicity of several compounds, both inorganic and organic. Some of these works are collected in Table 1.

Beyond these, Rogerson et al. (70) determined the effects of 17 hydrocarbons on the mortality of *Colpidium colpoda* and *T. ellioti* and Otsuka et al. (71) assessed the toxicity of diphenyl, *o*-phenylphenol, and 2-(4-thiazol)benzimidazol on the growth of *T. pyriformis*. Dias et al. (72) and Nicolau et al. (73) studied the toxicity of metals, Triton X-100, and cycloheximide on biochemical parameters of *T. pyriformis*. Nicolau et al. (74) had already assessed the toxicity of these compounds on the growth, mortality, and grazing of *T. pyriformis*.

Table 1. Toxicity Assessment of Different Compounds with Protists (*T—Tetrahymena* sp.; *C—Colpidium* sp.; *S—Spirostomum*)

Compound	Organism	References
Metals		
Aluminum	*T*	31
Cadmium	*T*; *T*; *T*; *T*; *T*; *C*	32; 33; 34; 35; 36; 37
Lead	*T*; *T*	38; 39
Copper	*T*; *T*; *T*; *C*	30; 29; 35; 37
Iron	*T*	40
Mercury	*T*; *T*	41; 42
Vanadate	*T*	43
Zinc	*T*; *T*	32; 33
Other		
Ceramide analogues	*T*	44
Cycloheximide	*T*	45
Chloramphenicol	*T*	46
Chloroquinone	*T*	47
Concanavalin A	*T*	48
Cocaine and derivates	*T*; *T*	49; 50
Dichloroisoproterenol	*T*	51
2,4-Dinitrophenol	*T*	52
DMSO (dimethylsulfoxide)	*T*; *T*; *T*	53; 54; 55
Ethidium bromide	*T*	56
Histamine	*T*	57
Histidinol	*T*	58
Insulin	*T*; *T*	59; 60
Lofepramine	*T*	61
Methotrexate	*T*	62
Octanol	*T*	63
Opioids	*T*	64
Penicillin	*T*	65
Phenols	*T*	66
Surfactants	*T*; *T*; *S*; *T*	24; 67; 68; 69

NEW APPROACHES IN THE ASSESSMENT OF TOXICITY WITH PROTISTS

Population growth impairment is, as mentioned above, the most frequently used sublethal toxic endpoint. A growth/mortality bioassay using microscopic observation of morphological changes at low magnification is a simple reproducible technique that does not require technical expertise or operational expenses. Nevertheless, some limitations detected in such an assay can make it unsuitable with several toxicants. Recognizing dead cells was sometimes ambiguous, since nonmotile cells as well as altered shape cells were counted as dead cells. It was found that light microscopy observations underestimated the true number of viable cells (i.e., false-negative counting), since several abnormal cells were not in fact dead. Indeed, they recovered their normal shape immediately after toxicant removal and began growing in the recovery assay (75). A novel cellular cytotoxicity assay using two fluorescent dyes was developed by Dias and Lima (75) as an alternative method to standard direct counting of viable protozoa under light microscopy. The compound calcein AM is a nonfluorescent substance that diffuses passively into cells and is converted by intracellular esterases to green fluorescent calcein in viable cells. EthD-1, which binds to DNA-stained red dead cells, was also added. The live/dead assay was found to be a very sensitive method for toxicants that damaged the cell membrane, given that all green cells were established as viable cells even with an altered shape. Recently, an improvement of this approach was developed by Dayeh et al. (69) using *T. termophila* and similar sensitivity was found between the ciliate and piscine and mammalian cells.

Another problem that the toxicologist must deal with is that conventional toxicological assays are often slow and labor-intensive and become impractical when many compounds and/or concentrations are being tested rapidly. This has led to a greater interest in colorimetric and fluorimetric assays that can be miniaturized in 96-well microtiter plates and assessed using an ELISA spectrophotometric microtiter plate reader.

Assays of ATP content have been widely used to characterize biomass viability and to detect potential spoilage microorganisms in the beer and food industry (76,77) assuming that it is possible to use the concentration of ATP to measure the viable cells of a certain species. In cultures of *T. pyriformis*, ATP content was determined to give information about the general energetic state of the culture when submitted to the toxicants (73).

Previous works have been devoted to studying the acid phosphatase (ACP) activity of other hydrolases to detect digestive activity in protozoa (78–81), since there is an intimate relationship between lysosomal function and intracellular digestion in *Tetrahymena* and other ciliates. The ACP activity was used also as an indicator of the metabolic state of the cultures, namely, of the intracellular digestive function, when the ciliates were exposed to toxicants (73,82).

Other colorimetric assays are the assays based on tetrazolium dye reduction, like MTT, which have been extensively used in the *in vitro* evaluation of cellular proliferation and cytotoxicity as well as screening for additives, mycotoxins, or anticancer drugs. The MTT assay is a quantitative method that consists of metabolic reduction of the tetrazolium dye, a group of water-soluble quaternary ammonium compounds, by dehydrogenases of viable immobilized or suspended cells. Production of intensely colored formazan water-insoluble crystals is the result of the tetrazolium reduction assay. Formazan crystals can be either observed microscopically in the cell cytoplasm or extracted and dissolved with organic solvents, such as DMSO, enabling spectrophotometric quantification in cultures of *T. pyriformis* (72).

As mentioned previously, bioactivity of various chemicals can be assessed by means of *Tetrahymena*'s grazing rate determination. For this purpose, fluorescent-labeled latex beads (FLLB) can be used instead of labeled bacteria, assuming that ciliates do not discriminate between different particles on the basis of properties other than size and shape. Therefore, direct counting of fluorescent beads ingested through a period less than 30 min allows an estimate of the xenobiotic's influence on the grazing activity (68,73).

Counting beads one by one inside all digestive vacuoles of a protist is not only tedious but also time consuming. Recently, an automated method of FLLB counting was proposed by Dias et al. (68) to improve the FLLB ingestion toxicity test. It was found that the image analysis is superior to direct FLLB counting, when speed of analysis is the concern.

A different toxicological test approach using *T. pyriformis* was based on calorimetric measurements assuming that protists produce heat by metabolism and movement (83). The addition of toxic substances should result in a reduction of heat production. It was found that the calorimetric measurement showed some advantages compared to the turbimetry.

QSARS: PREDICTING TOXICITY

With the increase of chemical usage, more and more newly manufactured chemicals are created by industry. Today, it is impossible to test all these new products and the already existing ones. Consequently, gaps in toxicity data exist and predictive models have become a means of filling these gaps. Relating chemical structure and toxic potency is a way to predict toxicity. The result of such an approach is known as quantitative structure–activity relationships, that is, QSARs.

There are several ways to do it: the traditional one relies on the use of a set of homologous compounds, each having a functional group common to them all and an alkyl substituent of varied length. In such a set, chemical reactivity, due to the functional group, is ignored because it is the same for all compounds; the hydrocarbon moiety alters hydrophobicity and therefore toxic potency. The models obtained in this way are usually simple regressions of toxicity versus hydrophobicity.

A more complex but innovative approach uses analogous groups of compounds with multiple function groups. The models obtained in this case are often multiple regressions of toxicity versus hydrophobic and stereoelectronic properties (84). It must be assumed, however, that one model predicts only one mechanism of toxic action (85).

Acute toxicity can fall into one of two categories: reversible nonspecific toxicity or narcosis and irreversible specific toxicity (11). Over 70% of organic industrial chemicals cause narcosis (86), which is reversible, physical alterations in the membrane, because they are nonbiding to the macromolecules (10). The remainder of industrial organic chemicals elicited toxicity that is nonreversible (87). These cause reactions that are typically covalent, resulting in changes in the biological systems, and are known as electro(nucleo)philic. The electro(nucleo)philic mechanism of toxic action seems to be dependent on the chemical class of the compound (84).

CONCLUSIONS

Conventional toxicological assays are often labor intensive and become impractical when many compounds and/or concentrations are being tested rapidly. This has led to a greater interest in colorimetric and fluorimetric assays that can be miniaturized and, on the other hand, do not require technical expertise or operational expenses that are needed in bioassays carried out by regulatory agencies at higher trophic levels such as in fish. The use of protists in miniaturized assays seems an excellent option.

All that was referred emphasizes the importance of protists as test organisms and encourages further work in this area, namely, in the establishment of new response patterns to toxicants and in the research of new species to be validated as test organisms, especially species with high indicator value or significant species in certain specific environments (activated sludge, eutrophic waters, amended soils, etc.).

BIBLIOGRAPHY

1. Olabarrieta, I., L'Azou, B., Yuric, J., Cambar, J., and Cajaraville, M.P. (2001). *In vitro* effects of cadmium on two different animal cell models. *Toxicol. In Vitro* **15**: 511–517.
2. Cairns, J. Jr. and Pratt, J.R. (1987). Ecotoxicological effect indices. *Water Sci. Technol.* **19**: 1–2.
3. Nilsson, J.R. (1989). *Tetrahymena* in cytotoxicology: with special reference to effects of heavy metals and selected drugs. *Eur. J. Protozool.* **25**: 2–25.
4. Benitez, L. et al. (1994). The ciliated protozoa *Tetrahymena termophila* as a biosensor to detect micotoxins. *Lett. Appl. Microbiol.* **19**: 489–491.
5. Eckelund, F., Ronn, R.E., and Christensen, S. (1994). The effect of three different pesticides on soil protozoan activity. *Pestic. Sci.* **42**: 71–78.
6. Campbell, C.D., Warren, A., Cameron, C.M., and Hope, S.J. (1997). Direct toxicity assessment of two soils amended with sewage sludge contaminated with heavy metals using a protozoa (*Colpoda steinii*) bioassay. *Chemosphere* **34**: 501–514.
7. Dive, D. et al. (1989). A bioassay using the measurement of the growth inhibition of a ciliate protozoan: *Colpidium campylum* Stokes. *Hydrobiologia* **188/189**: 181–188.
8. Nalecz-Jawecki, G., Demkowicz-Dobrzanski, K., and Sawicki, J. (1993). Protozoan *Spirostomum ambiguum* as a highly sensitive bioindicator for rapid and easy determination of water quality. *Sci. Total Environ. Suppl.* 1227–1234.
9. Twagilimana, L., Bohatier, J., Groliere, C.-A., Bonnemoy, F., and Sargos, D. (1998). A new low-cost microbiotest with the protozoan *Spirostomum teres:* culture conditions and assessment of sensitivity of the ciliate to 14 pure chemicals. *Ecotoxicol. Environ. Safety* **41**: 231–244.
10. Seward, J.R., Sinks, G.D., and Schultz, T.W. (2001). Reproducibility of toxicity across mode of toxic action in the *Tetrahymena pyriformis* population growth impairment assay. *Aquat. Toxicol.* **53**: 33–47.
11. Schultz, T.W., Sinks, G.D., and Cronin, M.T.D. (1997). Quinone-induced toxicity to *Tetrahymena*: structure–activity relationships. *Aquat. Toxicol.* **39**: 267–278.
12. Lwoff, M.A. (1923). Sur la nutrition des infusoires. *C.R. Acad. Sci. Paris* **176**: 928–930.
13. Persoone, G. and Dive, D. (1988). Toxicity tests on ciliates: a short review. *Ecotoxicol. Environ. Safety* **2**: 105–114.
14. Lynn, D.H. and Gilron, G.L. (1992). A brief review of approaches using ciliated protozoa to assess aquatic ecosystem health. *J. Aquat. Ecosys. Health* **1**: 263–270.
15. Sauvant, M.P., Pepin, D., and Piccinni, E. (1999). *Tetrahymena pyriformis*: a tool for toxicological studies. A review. *Chemosphere* **38**: 1631–1669.
16. Nicolau, A., Dias, N., Mota, M., and Lima, N. (2001). Trends in the use of protozoa in the assessment of wastewater treatment. *Res. Microbiol.* **152**: 621–630.
17. Silverman, J. (1983). Preliminary findings on the use of protozoa (*Tetrahymena termophila*) as models for ocular irritation testing in rabbits. *Lab. Anim. Sci.* **33**: 56–59.
18. Noever, D.A., Matsos, H.C., Cronise, R.J., Looger, L.L., and Relwani, R.A. (1994). Computerized *in vitro* test for chemical toxicity on *Tetrahymena* swimming patterns. *Chemosphere* **29**: 1373–1384.
19. Yoshioka, Y., Ose, Y., and Sato, T. (1985). Testing for the toxicity of chemicals with *Tetrahymena pyriformis*. *Sci. Total Environ.* **43**: 149–157.
20. Le Dû, A., Dive, D., Guebert, M., and Jouany, J.M. (1993). The protozoan biotest *Colpidium campylum*, a tool for toxicity detection and toxic interaction modelling. *Sci. Total Environ. Suppl.* 809–815.
21. Nalecz-Jawecki, G. and Sawicki, J. (1999). Spirotox—a new tool for testing the toxicity of volatile compounds. *Chemosphere* **38**: 3211–3218.
22. Nalecz-Jawecki, G. and Sawicki, J. (2002). The toxicity of tri-substituted benzenes to the protozoan ciliate *Spirostomum ambiguum*. *Chemosphere* **46**: 333–337.
23. Nalecz-Jawecki, G. and Sawicki, J. (2002). A comparision of sensitivity of Spirotox biotest with standard toxicity tests. *Arch. Environ. Contam. Toxicol.* **42**: 389–395.
24. Nalecz-Jawecki, G., Grabinska-Sota, E., and Narkiewicz, P. (2003). The toxicity of cationic surfactants in four bioassays. *Ecotoxicol. Environ. Safety* **54**: 87–91.
25. Sauvant, M.P., Pepin, D., Bohatier, J., Groliere, C.A., and Veyre, A. (1994). Comparative study of two *in vitro* models (L-929 fibroblasts and *Tetrahymena pyriformis* GL) for the cytotoxicological evaluation of packaged water. *Sci. Total Environ.* **156**: 159–167.
26. Larsen, J., Schultz, T.W., Rassmussen, L., Hoofman, R., and Pauli, W. (1997). Progress in an ecotoxicological standard protocol with protozoa: results from a pilot ring test with *Tetrahymena pyriformis*. *Chemosphere* **35**: 1023–1041.
27. Schultz, T.W. (1997). TETRATOX: the *Tetrahymena pyriformis* population growth impairment endpoint—a surrogate for fish lethality. *Toxicol. Methods* **7**: 280–309.
28. Schlimme, W., Marchiani, M., Hansekmann, K., and Jenni, B. (1999). BACTOX, a rapid bioassay that uses protozoa to assess the toxicity of bacteria. *Appl. Environ. Microbiol.* **65**: 2754–2757.
29. Wakatsuki, T., Tazaki, Y., and Imahara, H. (1986). Respiratory inhibition by copper in *Tetrahymena pyriformis* GL. *J. Fermentol. Toxicol.* **64**: 119–127.
30. Nilsson, J.R. (1981). Effects of copper on phagocytosis in *Tetrahymena*. *Protoplasma* **109**: 359–370.
31. Sauvant, M.P., Pepin, D., Bohatier, J., and Groliere, C.A. (2000). Effects of chelators on the acute toxicity and bioavailability of aluminum to *Tetrahymena pyriformis*. *Aquat. Toxicol.* **47**: 259–275.
32. Chapman, G. and Dunlop, S. (1981). Detoxification of zinc and cadmium by the freshwater protozoan *Tetrahymena pyriformis*. I. The effect of water hardness. *Environ. Res.* **26**: 81–86.
33. Chapman, G. and Dunlop, S. (1981). Detoxification of zinc and cadmium by the freswater protozoan *Tetrahymena pyriformis*.

II. Growth experiments and ultrastructural studies on sequestration of heavy metals. *Environ. Res.* **24**: 264–274.

34. Houba, C., Remacle, J., and De Parmentier, F. (1981). Influence of cadmium on *Tetrahymena pyriformis* in axenic culture. *Eur. J. Appl. Microbiol. Biotechnol.* **11**: 179–182.
35. Piccinni, E., Irato, P., Coppellotti, O., and Guidolin, L. (1987). Biochemical and ultrastructural data on *Tetrahymena pyriformis* treated with copper and cadmium. *J. Cell Sci.* **88**: 283–293.
36. Lawrence, S.G., Holoka, M.H., and Hamilton, R.D. (1989). Effects of cadmium on a microbial food chain, *Chlamidomonas reinhardii* and *Tetrahymena vorax*. *Sci. Total Environ.* **87/88**: 381–395.
37. Sekkat, N., Le Dû, A., Jouany, J.M., and Guebert, M. (1992). Study on the interaction between copper, cadmium, and ferbam using the protozoan *Colpidium campylum* bioassay. *Ecotoxicol. Environ. Safety* **24**: 294–300.
38. Nilsson, J.R. (1978). Retention of lead within the digestive vacuole in *Tetrahymena*. *Protoplasma* **95**: 163–173.
39. Nilsson, J.R. (1979). Intracellular distribution of lead in *Tetrahymena* during continuous exposure to the metal. *J. Cell Sci.* **39**: 383–396.
40. Rasmussen, L., Toftlund, H., and Suhr-Jessen, P. (1984). *Tetrahymena*. Adaptation to high iron. *Exp. Cell Res.* **153**: 236–239.
41. Thrasher, J.D. and Adams, J.F. (1972). The effects of four mercury compounds on the generation time and cell division in the *Tetrahymena pyriformis*, WH14. *Environ. Res.* **5**: 443–450.
42. Tingle, L.E., Pavlat, W.A., and Cameron, I.L. (1973). Sublethal cytotoxic effects of mercuric chloride on the ciliate *Tetrahymena pyriformis*. *J. Protozool.* **20**: 301–304.
43. Nilsson, J.R. (1999). Vanadate affects nuclear division and induces aberrantly-shaped cells during subsequent cytokinesis in *Tetrahymena*. *J. Eukaryot. Microbiol.* **46**: 24–33.
44. Kovács, P. and Csaba, G. (1998). Effect of ceramide-analogues on the actin cytoskeleton of *Tetrahymena pyriformis* GL. A confocal microscopic analysis. *Acta Protozool.* **37**: 201–208.
45. Ricketts, T.R. and Rappit, A.F. (1975). The effect of puromycin and cycloheximide on vacuole formation and exocytosis in *Tetrahymena pyriformis* GL-9. *Arch. Microbiol.* **102**: 1–8.
46. Nilsson, J.R. (1986). Effects of chloramphenicol on the physiology and fine structure of *Tetrahymena pyriformis* GL: correlation between diminishing inner mitochondrial membrane and cell doubling. *Protoplasma* **135**: 1–11.
47. Nilsson, J.R. (1992). Does chloroquine, an antimalarial drug, affect autophagy in *Tetrahymena pyriformis*? *J. Protozool.* **39**: 9–16.
48. Scott, S.M. and Hufnagel, L.A. (1983). The effect of concanavalin A on egestion of food vacuoles in *Tetrahymena*. *Exp. Cell Res.* **144**: 429–441.
49. Stefanidou, M.E., Georgiou, M., Maravelias, C., and Koutselinis, A. (1990). The effects of morphine, cocaine, amphetamine and hashish on the phagocytosis of the protozoon *Tetrahymena pyriformis* strain W. *Toxicol. In Vitro* **4**: 779–781.
50. Stefanidou, M.E., Alevisopoulos, G., Maravelias, C., Loutsidis, C., and Koutselinis, A. (1999). Phagocytosis of the protozoon *Tetrahymena pyriformis* as an endpoint in the estimation of cocaine salt and cocaine freebase toxicity. *Addict. Biol.* **4**: 449–452.
51. Fok, A.F. and Shockley, B.U. (1985). Processing of digestive vacuoles in *Tetrahymena* and the effects of dichloroisoproterenol. *J. Protozool.* **32**: 6–9.
52. Nilsson, J.R. (1995). pH-dependent effects of 2,4-dinitrophenol (DNP) on proliferation, endocytosis, fine structure and DNP resistance in *Tetrahymena*. *J. Eukaryot. Microbiol.* **42**: 248–255.
53. Lee, D. (1971). The effect of dimethylsulphoxide on the permeability of lysosomal membrane. *Biochim. Biophys. Acta* **233**: 619–623.
54. Nilsson, J.R. (1974). Effects of DMSO on vacuole formation, contractile vacuole function, and nuclear division in *Tetrahymena pyriformis* GL. *J. Cell Sci.* **16**: 39–47.
55. Nilsson, J.R. (1980). Effects of dimethyl sulphoxide on ATP content and protein synthesis in *Tetrahymena*. *Protoplasma* **103**: 189–200.
56. Meyer, R.R., Boyd, C.R., Rein, D.C., and Keller, S.J. (1971). Effects of ethidium bromide on growth and morphology of *Tetrahymena pyriformis*. *Exp. Cell Res.* **70**: 233–237.
57. Darvas, Z., Madarász, B., and Lászlo, V. (1999). Study of histamine effects on phagocytosis and enzyme secretion of *Tetrahymena pyriformis*. *Acta Biol. Hung.* **50**: 325–334.
58. Ron, A. and Wheatley, D.N. (1984). Stress-proteins are induced in *Tetrahymena pyriformis* by histidinol but not in mammalian (L-929) cells. *Exp. Cell Res.* **153**: 158–166.
59. Fülöp, A.K. and Csaba, G. (1994). Insulin pretreatment (imprinting) produces elevated capacity in the insulin binding of *Tetrahymena*. Different binding by the cilia and the oral field. *Biosci. Rep.* **14**: 301–308.
60. Fülöp, A.K. and Csaba, G. (1997). Accumulation of insulin-gold particles in the oral apparatus of *Tetrahymena* after insulin pretreatment (imprinting). *Microbios* **90**: 123–128.
61. Darcy, P., Kelly, J.P., Leonard, B.E., and Henry, J.A. (2002). The effect of lofepramine and other related agents on the motility of *Tetrahymena pyriformis*. *Toxicol. Lett.* **128**: 207–214.
62. Nilsson, J.R. (1983). Methotrexate permits a limited number of cell doublings and affects mitochondrial substructure in *Tetrahymena*. *Protoplasma* **117**: 53–61.
63. Bearden, A.P. and Schultz, T.W. (1998). Comparision of *Tetrahymena* and *Pimephales* toxicity based on mechanism of action. *SAR QSAR Environ. Res.* **9**: 127–153.
64. Wu, C., Fry, C.H., and Henry, J.A. (2002). Membrane toxicity of opioids measured by protozoan motility. *Toxicology* **117**: 35–44.
65. Szablewski, L. (1981). The effect of selected antibiotics on *Tetrahynmena pyriformis*. *Acta Protozool.* **20**: 309–322.
66. Jaworska, J.S. and Schultz, T.W. (1991). Comparative toxicity and structure–activity relationships in *Chlorella* and *Tetrahymena*: monosubstituted phenols. *Bull. Environ. Contam. Toxicol.* **47**: 57–62.
67. Dias, N., Mortara, R.A., and Lima, N. (2003). Morphological and physiological changes in *Tetrahymena pyriformis* for the *in vitro* cytotoxicity assessment of Triton X-100. *Toxicol. In Vitro* **17**: 357–366.
68. Dias, N., Amaral, A.L., Ferreira, E.C., and Lima, N. (2003). Automated image analysis to improve bead ingestion toxicity test counts in the protozoan *Tetrahymena pyriformis*. *Lett. Appl. Microbiol.* **37**: 1–4.
69. Dayeh, V.R., Chow, S.L., Schirmer, K., Lynn, D.H., and Bols, N.C. (2004). Evaluating the toxicity of Triton X-100 to protozoan, fish and mammalian cells using fluorescent dyes as indicators of cell viability. *Ecotoxicol. Environ. Safety* **57**: 375–382.

70. Rogerson, A., Shiu, W.Y., Huang, G.L., Mackay, D., and Berger, J. (1983). Determination and interpretation of hydrocarbon toxicity to ciliate protozoa. *Aquat. Toxicol.* **3**: 215–228.
71. Otsuka, K., Yoshikawa, H., Sugitani, A., and Kawai, M. (1988). Effect of diphenyl, *o*-phenylphenol and 2-(4-thiazol)benzimidazole on growth of *Tetrahymena pyriformis*. *Bull. Environ. Contam. Toxicol.* **41**: 282–285.
72. Dias, N., Nicolau, A., Carvalho, G.S., Mota, M., and Lima, N. (1999). Miniaturization and application of the MTT assay to evaluate metabolic activity of protozoa in the presence of toxicants. *J. Basic Microbiol.* **39**: 103–108.
73. Nicolau, A., Mota, M., and Lima, N. (2004). Effect of different toxic compounds on ATP content and acid phosphatase activity in axenic cultures of *Tetrahymena pyriformis*. *Ecotoxicol. Environ. Safety* **57**: 129–135.
74. Nicolau, A., Mota, M., and Lima, N. (1999). Physiological responses of *Tetrahymena pyriformis* to copper, zinc, cycloheximide and Triton X-100. *FEMS Microbiol. Ecol.* **30**: 209–216.
75. Dias, N. and Lima, N. (2002). A comparative study using a fluorescence-based and a direct-count assay to determine cytotoxicity in *Tetrahymena pyriformis*. *Res. Microbiol.* **153**: 313–322.
76. Dowhanick, T.M. and Sobczak, J. (1994). ATP bioluminescent procedure for viability testing of potential beer spoilage microorganisms. *J. Am. Soc. Brew. Chem.* **52**: 19–23.
77. Gamborg, G. and Hansen, E.H. (1994). Flow-injection bioluminescent determination of ATP on the use of the luciferin–luciferase system. *Anal. Chem. Acta* **285**: 321–328.
78. Rothstein, R.L. and Blum, J.J. (1974). Lysosomal physiology in *Tetrahymena*. III. Pharmacological studies on acid hydrolase. *J. Cell Biol.* **62**: 844–859.
79. Rothstein, R.L. and Blum, J.J. (1975). Lysosomal physiology in *Tetrahymena*. II. Effect of culture age and temperature on the extracellular release of 3 acid hydrolases. *J. Protozool.* **21**: 163–168.
80. Ricketts, T.R. and Rappit, A.F. (1975). The adaptative acid phosphatase of *Tetrahymena pyriformis* GL-9. Effect of endocytosis of various nutrients. *Protoplasma* **85**: 119–125.
81. Williams, J.T. and Juo, P.S. (1976). Release and activation of a particulate bound acid phosphatase from *Tetrahymena pyriformis*. *Biochim. Biophys. Acta* **422**: 120–126.
82. Martin, A. and Clydes, M. (1991). Acid phosphatase endpoint for *in vitro* toxicity tests. *In Vitro Cell. Dev. Biol.* **27A**: 183–184.
83. Beermann, K., Buschmann, H.-J., and Schollmeyer, E. (1999). A calorimetric method for the rapid evaluation of toxic substances using *Tetrahymena pyriformis*. *Termochim. Acta* **337**: 65–69.
84. Akers, K.S., Sinks, G.D., and Schultz, T.W. (1999). Structure–toxicity relationships for selected halogenated aliphatic chemicals. *Environ. Toxicol. Pharmacol.* **7**: 33–39.
85. Cronin, M.T.D. and Dearden, J.C. (1995). QSAR in toxicology. 1. Prediction of aquatic toxicity. *Quant. Struct. Act. Relat.* **14**: 1–7.
86. Bradbury, S.P. and Lipnick, R.L. (1990). Introduction: structural properties for determining mechanisms of toxic action. *Environ. Health Perspect.* **87**: 181–182.
87. Mekenyan, O.G. and Veith, G.D. (1994). The electronic factor in QSAR: MO-parameters, competing interactions, reactivity and toxicity. *SAR QSAR Environ. Res.* **2**: 129–141.

SOFIE: AN OPTIMIZED APPROACH FOR EXPOSURE TESTS AND SEDIMENT ASSAYS

Jos P.M. Vink
Serge Rotteveel
Cornelis J.H. Miermans
Institute for Inland Water
Management and Waste Water
Treatment–RIZA
Lelystad, The Netherlands

INTRODUCTION

Regulatory quality standards are primarily developed to indicate, scale, or rank the environmental risks of contaminants. For soils and sediments, most countries legally prescribe the chemical extraction of total amounts of contaminants as a first step for risk assessment. Contaminants have mostly been sequestered over a long period of time and may only become available to organisms to a limited extent. The introduction of the bioassay, as a second step, generally increased the understanding of the concept of chemical versus biological availability. This is illustrated in Fig. 1. Aquatic oligochaetes that have been exposed to a number of sediments with a varying degree of cadmium content appear to be particularly susceptible to dissolved fractions. Although location L1 is ranked as seriously contaminated, based on total extracted amounts of cadmium from sediment, *Limnodrilus* spp. accumulated the most cadmium at the “clean” location L2.

The thought that chemical speciation generally determines the adverse toxic effects in biota is well accepted [reviewed by, among others, Campbell (1) in 1995]. In recent years, some techniques have become available to actually measure the availability (i.e., labile metal species) in “natural” solutions. These techniques make use of the (combined) electrochemical, diffusive, or competitive properties of metal ions. Table 1 summarizes some of the most promising techniques. However, only a few of these may find their use in exposure tests because of their disturbance (or destruction) of the sample, or the inability to measure speciation under anoxic conditions.

The purpose of bioassays is typically to determine the toxicity of ambient waters and sediments. Some assays focus on the determination of the exact concentration at which a chemical becomes toxic to an organism, so a biological response is used as a sensor. In order to be useful, this response should be repeatable. Ideally, bioassays are used to make predictions of environmental toxicity required for contaminant management (9,10) and should therefore at all times aim at extrapolating results to the field. Sediment bioassays are a relatively new approach to determine the environmental effects of sequestered metals in sediments. This type of assay has been discussed widely (11,12) in terms of limitations and advantages. Sediments typically consist of oxidized layers, overlying reduced ones. It is well known and documented how redox processes affect metal chemistry (3) and consequently the outcome of the test. Nevertheless, sediments are generally collected as bulk samples that contain this initial redox

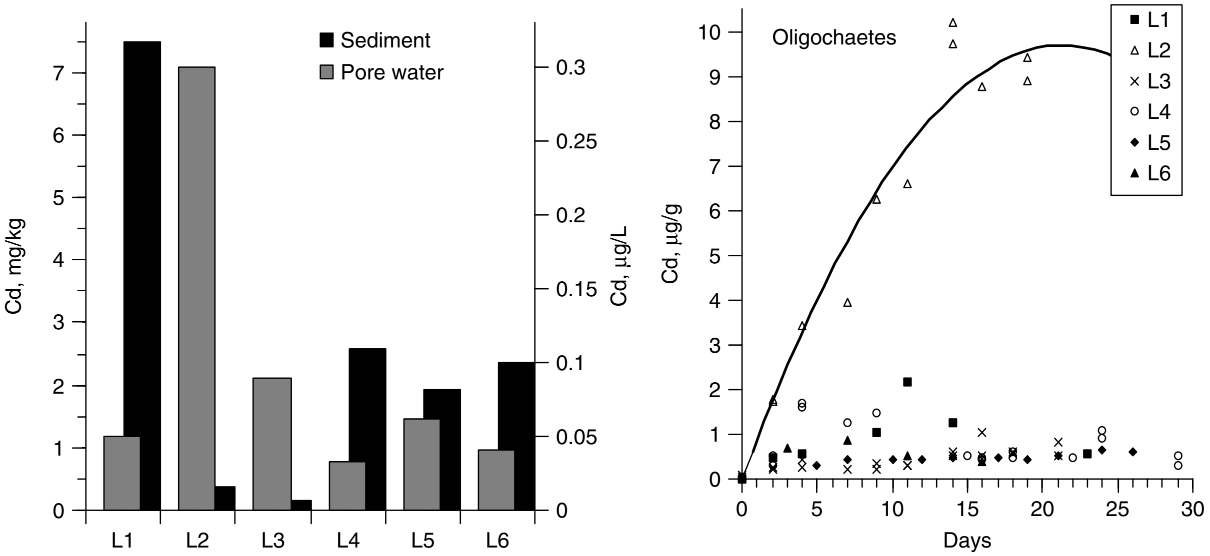

Figure 1. The availability concept in practice: cadmium in sediment, pore water, and oligochaetes at various locations (L1–L6).

zonation. Mostly, these samples are further manipulated by sieving, mixing, and oxygenation in order to increase homogeneity and repeatability. Indeed, variability is likely to be increased if sediments are not homogenized after collection, but these types of handling may destroy important reaction zones that are typical for the sediment in question and its environmental effect.

Data from standardized bioassays (i.e., protocols) generally provide the clear-cut answers that managers find useful as the basis for decisions, regardless of the accuracy to natural conditions. Therefore, these data continue to be the principal tool in the development of regulatory quality standards (12). If a bioassay is conducted in order to meet the objective of field-extrapolation, that is, to assess the environmental risks under natural conditions, a method of minimal sample manipulation should be considered.

Recently, a new approach was introduced that allows for a nondestructive handling and true-to-nature assessment of water/sediment systems. This approach is the "sediment or fauna incubation experiment," or SOFIE, and was developed to measure chemical speciation of metals over redox zones in undisturbed systems, while simultaneously

Table 1. Operational Metal Speciation Sensors

Method		Chemical Separation	Measured Species	Response Time	Remark	Reference
AdSV	Adsorptive stripping voltammetry	Electrochemical	Free metal + labile species	Minutes	Complex detection	2
CLE	Competing ligand exchange	Exchange reaction with ligands	Free metal (+ very labile species?)	Seconds	Critical modification and handling	3
DET	Diffusive equilibrium in thin film	Porous gel	Free metal + penetrable complexes	Days	Ionic strength limitations	4
DGT	Diffusive gradient in thin film	Porous gel	Free metal + labile penetrable complexes	Days	Critical gel composition	5
DMT	Donnan membrane technique	Charged pore membrane	Free metal + part of cationic penetrable complexes	Days	Critical pore size	
GIME	Gel impregnated microelectrode	Porous gel + electrochemical	Free metal + labile penetrable complexes	Minutes	Surface/lability interaction	6
ISE	Ion selective electrode	Liquid partition without countercharge transport	Free metal	Seconds	Interference from other metals and organic matter	7
PLM	Permeation liquid membrane	Liquid partition with countercharge transport	Free metal + labile complexes	Minutes	Diffusion-controlling step	8

conducting a single or multiple species bioassay. Here, we discuss the performance of this approach and compare it with the outcome of a standardized protocol for sediment or surface water bioassays. Tests were conducted with the sediment dweller *Chironomus riparius* (mosquito larva) and *Daphnia magna* (water flea) for a range of contaminants. Since this study aims at the straightforward comparison between the outcome of both methods, the speciation measurements conducted with SOFIE are not reported here.

METHODS OF EXPOSURE

The Standard Bioassay

The widely applied protocol for water and sediment testing for single species is what we call here the "standard bioassay." This protocol is based on the TRIAD guidelines of 1993 and describes the general procedure of the test, including handling of samples and test species. From this protocol, two test procedures were followed.

Standard Test with Sediment Dwellers. For tests using *Chironomus riparius* as a test organism, use is made of a water/sediment system. A bulk sample of the top 10 cm of a sediment is collected from the field and mechanically homogenized. The sample is then mixed 1:4 v/v with DSW ("standard water"; demineralized water with mineral additives). The water/sediment mixture is shaken for 24 h at ambient atmosphere. Next, 100 mL of the suspension is decanted into dishes to settle, after which test species are introduced. After 28 d of exposure, the organisms are collected by sieving, allowed to void their gut content for 24 h, and freeze dried.

Standard Test with Water Dwellers. For tests using *Daphnia magna*, the protocol prescribes the use of pore water, which is separated from the homogenized sediment by centrifugation (2500 g, 30 min) and stored in bottles. Exposure is carried out in 100-mL glass beakers and pore water is refreshed twice a week. Individuals are periodically collected, rinsed with Elendt medium, and analyzed for metals (not PAHs because of mass and detection limitations).

The Sediment or Fauna Incubation Experiment (SOFIE)

This study was performed with a novel experimental technique (EU patent 1018200/02077121.8, October 2001, J. Vink, Rijkswaterstaat), which was introduced as sediment or fauna incubation experiment, in short SOFIE. This device or "cell" is shown in Fig. 2.

SOFIE is based on the competing ligand exchange technique, which is integrated in a (pore) water probe and combined with a bioassay setting. Full and detailed experimental possibilities have been described earlier (3). The cell consists of a circular core, which is used as a sampling device to obtain undisturbed water/sediment systems. Field samples are taken including the overlying surface water in such a way that the physical and geochemical integrity of the sample (e.g., bulk density, redox status) is guaranteed. After sampling, the core is closed at the bottom with a shutter plate and is mounted on a base socket. The sample does not leave the body core but is now part of the cell. A top socket is attached, ensuring a gas-tight fitting with the body core with an internal silicone seal. The top plate of this socket has four gas-tight connectors for any commercial electrode to probe sediment or overlying water. For experimental scenario purposes, aerobic or anaerobic conditions in the water–sediment system may be manipulated by flushing the atmospheric head space with any desirable gas. The cell wall contains fifteen gas-tight connectors for probes. These probes (diameter 1 mm, length 50 mm) consist of a semipermeable polyethersulfon, which acts as a membrane to discriminate colloidal fractions greater than 0.1 μm. This mesh size is impermeable to bacteria, so pore water that passes the polymer is sterile. The probes can be directly coupled onto micro ion-exchange columns (MICs), as was described in detail by Vink (3), to separate free metal ions from other forms of organic and inorganic metal ligands in solution at the reigning geochemical status of the probed sediment layer.

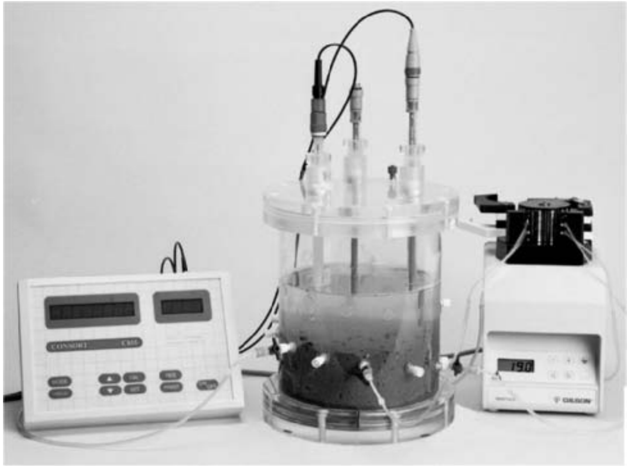

Figure 2. Sediment or fauna incubation experiment (SOFIE).

For a simple, straightforward comparison between the standard bioassay and SOFIE, sediment samples were collected from the river Meuse, The Netherlands, and analyzed for a large amount of properties. An undisturbed water–sediment interface was taken with the cell's body core at 0.4 m water depth. After steady-state metal concentrations were determined, 125 individuals of *Chironomus riparius* were introduced into the SOFIE cell and sampled at the same time intervals as in the standard bioassay. Tests with *Daphnia magna* were performed in the same way.

The Presence of Food During Exposure

In an additional test with *Daphnia magna,* the effect of feeding during exposure was investigated. This test was performed in a two-compartment SOFIE cell, which has two separate chambers that divide the sample while it is cored. In both compartments, daphnids were introduced in comparable amounts as in the standard bioassay. Daphnids in compartment A were fed with the alga *Chlorella,* their natural food source, while those in compartment B were not fed. Organisms and water samples were collected from the cell compartments on several occasions.

Metal concentrations in chironomids and daphnids were determined by digestion in 500 μL of 14.9 M HNO_3 (Ultrex) at 180 °C in a microwave. The digest was analyzed using inductively coupled plasma mass spectrometry (ICP-MS, PerkinElmer Elan 6000). Polycyclic aromatic hydrocarbons (PAHs) in chironomids were determined by refluxing for 6 h with hexane. Analyses were done with an Agilent HP liquid chromatograph using a fluorescence detector and a Vydac 201TP54 reverse phase C18 column. Dolt-2 (certified by the Community Bureau of Reference, BCR, Brussels, Belgium) was used as biological reference material.

RESULTS AND DISCUSSION

The sediment was composed of 14% < 2 μm, 2.3% organic C, and 0.55 inorganic C, at pH = 7.2, and was moderately contaminated with metals (As, 8.5 mg/kg dw; Cd, 1.57; Cr, 27.7; Cu, 25.7; Hg, 0.21; Ni, 20.8; Pb, 79; Zn, 321) and various PAHs (ranging from 100 to 800 ng/g dw).

Collection of pore water from the sediment by homogenization and centrifugation, according to the protocol, generally increased metal concentrations, most significantly for Pb and Cd. Table 2 shows the concentrations that were measured at the start of the exposure test in the standard bioassay and in SOFIE's surface water.

The collected time data for chironomids are, for reasons of survey ability, summarized in Fig. 3, which shows body concentrations of polycyclic aromatic hydrocarbons (PAH) and heavy metals after four weeks of exposure. The data quite clearly show that, in the standard bioassay, chironomids take up PAHs in significantly larger amounts than in the SOFIE environment. Although we did not perform or aim at a mechanistic study here, it is likely that sample handling (and in particular stirring and homogenization) has liberated PAHs from the organic sediment matrix by unlocking and exposing organic surface areas. Many authors have described this phenomenon.

Bioaccumulated metals are shown in Figs. 3 and 4. Accumulated amounts for both test species are lower in the standard bioassay than those collected from SOFIE (note the log scale, differences are significant). This is valid for all metals. Although this observation may appear somewhat counterintuitive [oxidation leads to liberation of metals from sulfides and should therefore increase the (bio)available fraction of metals], the explanation is quite straightforward. Indeed, oxidation leads to enhanced metal concentrations, as is shown in Table 3. Following chemical thermodynamics, this primarily yields large free ion activities. Simultaneously, iron and manganese are oxidized from the reduced state into highly reactive oxyhydroxides (either precipitated or as colloids). Free metal ions have a high affinity for these newly formed sorption sites and are quickly adsorbed and immobilized. In the case of the standard *Daphnia* bioassay, where

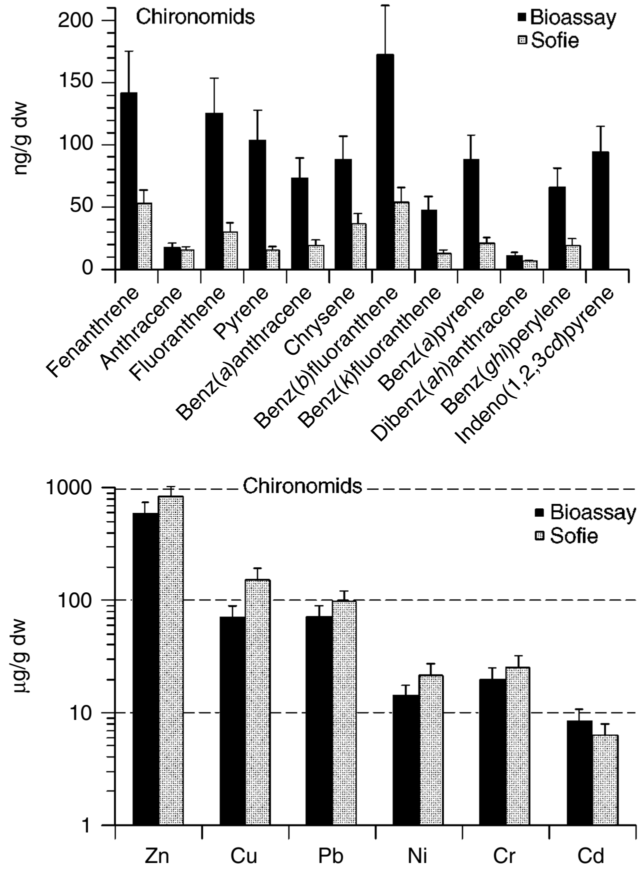

Figure 3. PAHs (top) and heavy metals (bottom) in chironomids after 28 d. Error bars denote the variation in analytical recovery.

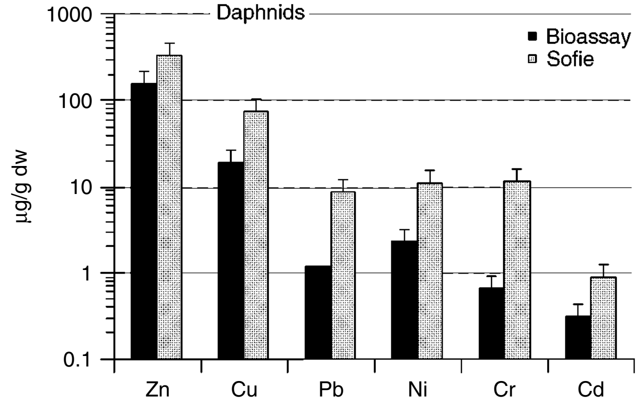

Figure 4. Metals in daphnids after 16 d of exposure.

Table 2. Pore Water Concentrations (μg/L) After Sample Handling (Standard Bioassay) and in Sediment of SOFIE

	Standard Bioassay[a]		SOFIE[a]	
Zn	23.3	(4.7)	28.1	(5.6)
Cu	6.5	(0.7)	2.7	(0.2)
Pb	17.8	(3.6)	1.3	(0.3)
Ni	6.4	(0.7)	2.0	(0.2)
Cr	4.2	(0.3)	2.4	(0.2)
Cd	0.26	(0.01)	0.03	(0.002)

[a] Values in parentheses are analytical standard deviations.

Table 3. Initial Concentrations C_i (μg/L) in the Exposure Medium, Model Parameters k_0, k_1, k_2 (1/d), and Correlation Coefficient

	Standard Bioassay						SOFIE					
	Zn	Cu	Pb	Ni	Cr	Cd	Zn	Cu	Pb	Ni	Cr	Cd
Chironomids												
C_i	9.2	3.7	3.8	2.8	2.0	0.08	45.8	2.7	2.5	4.9	1.0	0.39
k_0	−0.040	−0.11	−0.120	−0.030	−0.050	−0.100	−0.007	−0.015	−0.018	0.018	0.045	0.001
k_1	44.05	7.41	5.96	18.67	9.56	10.19	1.55	4.53	47.53	7.52	2.04	3.39
k_2	1.15	6.51	9.71	8.84	3.72	1.07	0.09	0.11	1.98	1.61	0.04	0.08
r^2	0.80	0.38	0.54	0.68	0.44	0.14	0.90	0.79	0.96	0.55	0.67	0.53
Daphnids												
C_i	23.3	6.5	17.8	6.4	4.2	0.26	45.8	2.7	2.5	4.9	1.0	0.39
k_0	−0.114	−0.258	−0.350	−0.215	−0.151	−0.272	−0.040	0.047	0.101	0.023	−0.062	0.106
k_1	10.54	0.94	0.005	0.03	0.26	0.31	0.66	1.44	0.01	0.01	0.11	0.07
k_2	8.31	12.94	12.72	2.24	12.80	12.67	0.11	−0.06	−0.35	−0.32	−0.17	0.01
r^2	0.92	0.13	0.71	0.89	0.58	0.47	0.97	0.99	1.00	1.00	0.99	0.85

sediment is absent, there is obviously no delivery from the sediment (e.g., by diffusion and desorption) to the aqueous phase. Daphnids take up metals from solution and eventually face declining exposure concentrations to a point where concentration levels may become inhibiting to uptake. Figure 5 shows copper for an example. This trend was observed for all metals in the standard bioassays, but not in the SOFIE cells where concentrations remained relatively constant during the exposure period or showed some release from the sediment.

Figure 6 shows the effect of feeding daphnids during exposure. The effect of adding the alga *Chlorella* is profound (note the log scale). In most cases, bioaccumulated amounts of metals are reduced some orders of magnitude compared to the nonfeeding test. Since alga are always present in surface waters in varying amounts, it is obvious that this test is closer to exposure conditions as they occur in the field.

Results show that dissolved metal concentrations change during exposure. These time-varying changes may follow first-order reaction kinetics; that is, $dC/dt = kC$, so the external concentration, or exposure concentration, at a given time $C(t)$, is written as

$$C(t) = C_i e^{k_0 t}$$

where C_i is the initial exposure concentration, and k_0 is a rate term describing the increase or decrease of the initial concentration during the test. A time/concentration-dynamic, two-compartment model approximated metal bioaccumulation patterns of both test species. Body concentrations vary in time and relate to uptake from the water phase at a certain rate. At the same time, uptake is accompanied by elimination, which is kinetically directed by the organism itself. Hence, one may write

$$dQ/dt = k_1 C(t) - k_2 Q(t)$$

where Q is the internal body concentration ($\mu g \cdot g^{-1}$), k_1 and k_2 are uptake and elimination rate constants, respectively (d^{-1}), and t is time. This yields, for $Q(0) = 0$,

$$Q(t) = \frac{k_1 C_0}{k_2 - k_0}(e^{-k_0 t} - e^{-k_2 t})$$

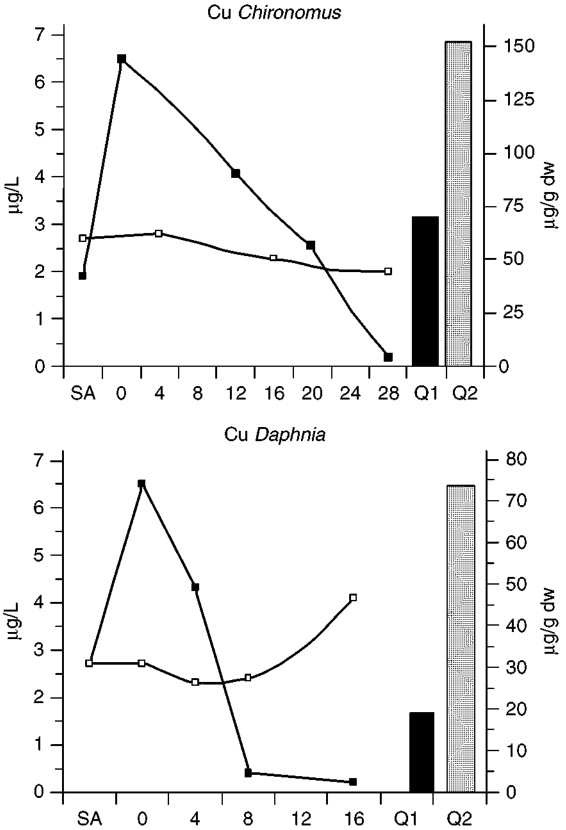

Figure 5. Copper concentrations in overlying water at time of sampling (SA) and during exposure (days) in the standard bioassay (■) and SOFIE (□) for chironomids (top) and daphnids (bottom). Q1 and Q2 are body concentrations from the standard bioassay and SOFIE, respectively. While concentrations in the undisturbed water–sediment system remain relatively constant or increase slightly, those in the standard bioassay show large variation due to the handling and manipulation of the sample, followed by exhaustion from solution.

With the introduction of k_0, time-dependent concentrations determine the overall exposure of organisms, and thus $Q(t)$.

Table 3 summarizes the results of the uptake model. In the standard bioassay, initial concentrations differ

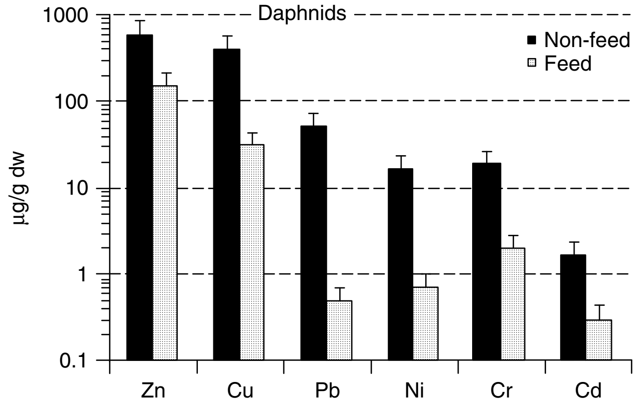

Figure 6. Heavy metals in *Daphnia magna* in two feeding scenarios.

between chironomids and daphnids. This is the result of sample handling and the absence of sediment in the daphnid test. From the viewpoint that bioassays should deliver generic, true-to-nature outcomes, this appears an undesirable starting point. Although *Chironomus* is a sediment dweller and *Daphnia* is not, both species encounter contaminant exposure in an environment that consists of both sediment and surface water, where bioavailability is a combined action of the two compartments.

The rate by which exposure concentrations change during the tests is represented by k_0. Positive values denote an increase, negative values a decrease in concentrations. The outcome clearly shows that in the standard bioassay, k_0 values are in all cases negative and larger than those observed in the assays that were performed with SOFIE. The effect is most profound in the daphnid tests, where k_0 values diverge both in direction and in magnitude. For all metals, and both test species, the two-compartment model performed better with the data that were acquired from the SOFIE setup.

CONCLUSIONS

Based on the outcome of water–sediment testing of two aquatic species with the standard protocol and with the sediment or fauna incubation experiment, SOFIE, the following conclusions are drawn:

- Sample handling such as homogenization and oxygenation affects the chemical speciation of the exposure medium for both organic and inorganic contaminants. Disadvantages of this manipulation (disrupting chemical speciation and bioavailability) outweigh the advantages (homogenization to increase reproducibility) by far and should therefore seriously be reconsidered or applied with great caution.
- Tests conducted with the standard protocol tend to *overestimate* uptake of PAHs by chironomids and *underestimate* uptake of heavy metals by chironomids and daphnids. This may generate both false-positive and false-negative results, respectively.
- Aquatic bioassays should be conducted in sediment/water systems, regardless of whether the tested species are sediment dwellers or not. The physical, chemical, and biological interaction between sediment and the water phase is a site-specific characteristic that determines the ultimate risk of contaminants.
- SOFIE provides the necessary tool to conduct risk assessments in a close-to-nature setting. It may therefore contribute to establish better-founded quality criteria for natural waters.

BIBLIOGRAPHY

1. Campbell, P.C.G. (1995). *Metal Speciation and Bioavailability in Aquatic Systems*. John Wiley & Sons, Hoboken, NJ.
2. Kaldova, R.A. and Kopanica, M. (1989). *Pure Appl. Chem.* **61**: 97–112.
3. Vink, J.P.M. (2002). *Environ. Sci. Technol.* **36**: 5130–5138.
4. Davison, W., Grime, G.W., Morgan, J.W., and Clarke, K. (1991). *Nature* **352**: 545.
5. Zhang, H. and Davison, W. (1995). *Anal. Chem.* **67**: 3391.
6. Tercier, M.L. and Buffle, J. (1996). *Anal. Chem.* **68**: 3670–3678.
7. Berner, R.A. (1963). *Geochim. Cosmochim. Acta* **27**: 563–575.
8. Buffle, J., Parthasarathy, N., Djane, N.K., and Matthiasson, L. (2000). *In Situ Monitoring of Aquatic Systems*. John Wiley & Sons, Hoboken, NJ.
9. Chapman, P.M. (1989). *Environ. Sci. Technol.* **8**: 589–601.
10. Blum, D.J.W. and Speece, R.E. (1990). *Environ. Sci. Technol.* **24**: 284b.
11. Schwartz, R.C. (1987). *Fate and Effects of Sediment Bound Chemicals in Aquatic Systems*. Pergamon Press, New York.
12. Luoma, S.N. (1995). *Metal Speciation and Bioavailability in Aquatic Systems*. John Wiley & Sons, Hoboken, NJ.

PASSIVE TREATMENT OF ACID MINE DRAINAGE (WETLANDS)

Arthur W. Rose
Pennsylvania State University
University Park, Pennsylvania

Starting in about 1985, passive treatment has been increasingly used as a method of remediating acid mine drainage (AMD). A passive treatment facility uses natural reactions such as microbial oxidation or reaction with a bed of crushed limestone to improve water quality, with only minimal human attention after construction. The initial passive method was a simple wetland, which was found to improve water quality for some types of AMD. However, for other waters, wetlands were not effective, but a variety of other methods have been developed, as indicated in Fig. 1. Passive treatment contrasts with active treatment in which a chemical such as sodium hydroxide, lime, or similar alkaline agents is added to the water. Active methods require frequent human attention and continued replenishment of the reagents.

Passive methods applicable for a specific site are selected using a flowsheet, as in Fig. 2. The first step is to determine the flow and water chemistry. Key chemical

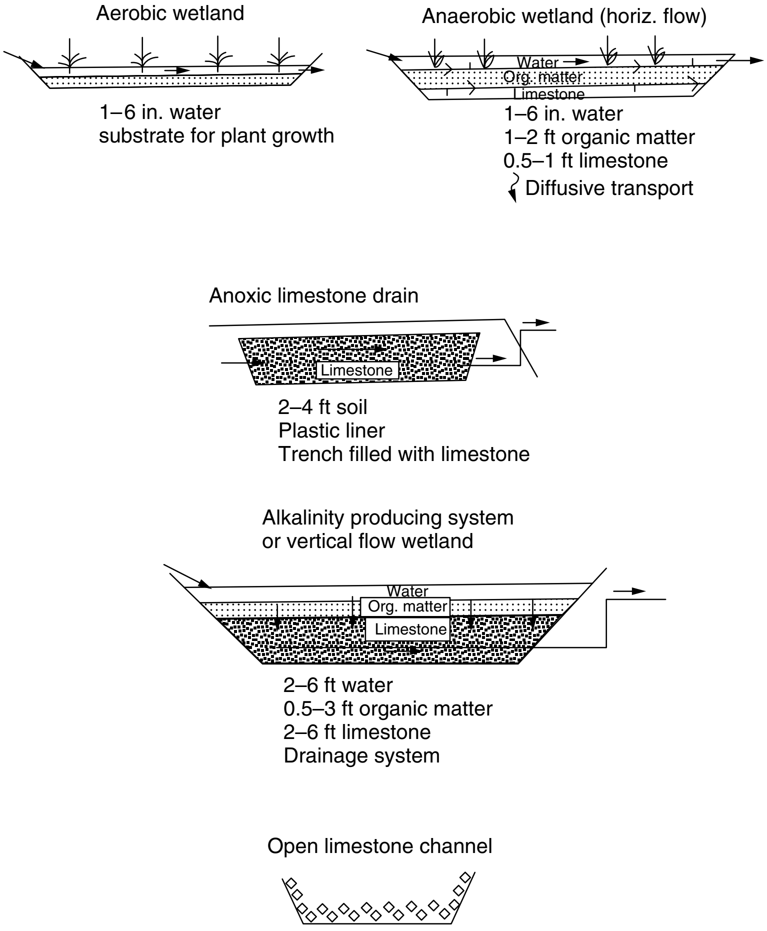

Figure 1. Schematic diagrams for passive treatment systems (after Ref. 1).

variables are the acidity, alkalinity, Fe, Al, Mn, and SO_4. The acidity should be a hot peroxide acidity such as presented in References 2 and 3. In these methods, the Fe and Mn are oxidized so that the total acidity to pH 8.3 is measured and the alkalinity is subtracted during the procedure so the result is a net acidity (4).

If the water is net alkaline (negative acidity), the water can easily be treated by a wetland. Sizing of wetlands is discussed by Hedin et al. (5). In the wetland, Fe and Mn are oxidized in the shallow water, precipitate as hydroxides, and settle in the slowly flowing water. The alkalinity is adequate to maintain a pH above 4.5, usually higher.

If the water is net acid (positive acidity), then the dissolved O_2, ferric iron, and aluminum must be measured. If these are negligible (less than about 1 mg/L each), then an anoxic limestone drain (ALD) can be successful. An ALD is a trench filled with crushed limestone ($CaCO_3$) and covered to eliminate access of air. Acidic water from a subsurface source flows into one end and passes along the limestone bed. The water acquires alkalinity and increases in pH by reaction with the limestone. On emerging from the limestone bed, Fe oxidizes and precipitates in an oxidation-settling pond.

If the water is net acid and contains appreciable dissolved O_2, ferric iron, or aluminum, then a common passive method is a vertical flow pond (also known as a SAPS, successive alkalinity producing system) (6,7). In this method, the water first flows into a settling pond to remove sediment and possibly some Fe. It then passes into the vertical flow pond that drains through the bottom. The top layer in the bottom of the pond is compost that consumes dissolved O_2 and reduces ferric iron to ferrous iron. Some Fe and acidity may be removed by microbial reduction of SO_4 in the compost. The water then enters the lower layer of crushed limestone, where acidity is consumed, alkalinity is added, and pH is increased. The water leaves the limestone layer through perforated pipes and flows to an oxidation-settling pond for Fe precipitation and removal. A wetland may be used to further clean up the water.

A problem in some vertical flow systems is dissolved Al, which can precipitate in the limestone bed and plug the system, or coat limestone leading to inhibited reaction. Some units are constructed to be flushed, so that precipitates are flushed out of the limestone bed. Recent measurements indicate that this method is only partially effective. Trials are underway with back-flushing systems, upflow limestone ponds, mixing of fine limestone into the compost, and other variations.

On relatively steep slopes, an open limestone channel can be successful in neutralizing acidity. To be successful, the flow must be vigorous enough to erode the Fe and Al coatings off the limestone.

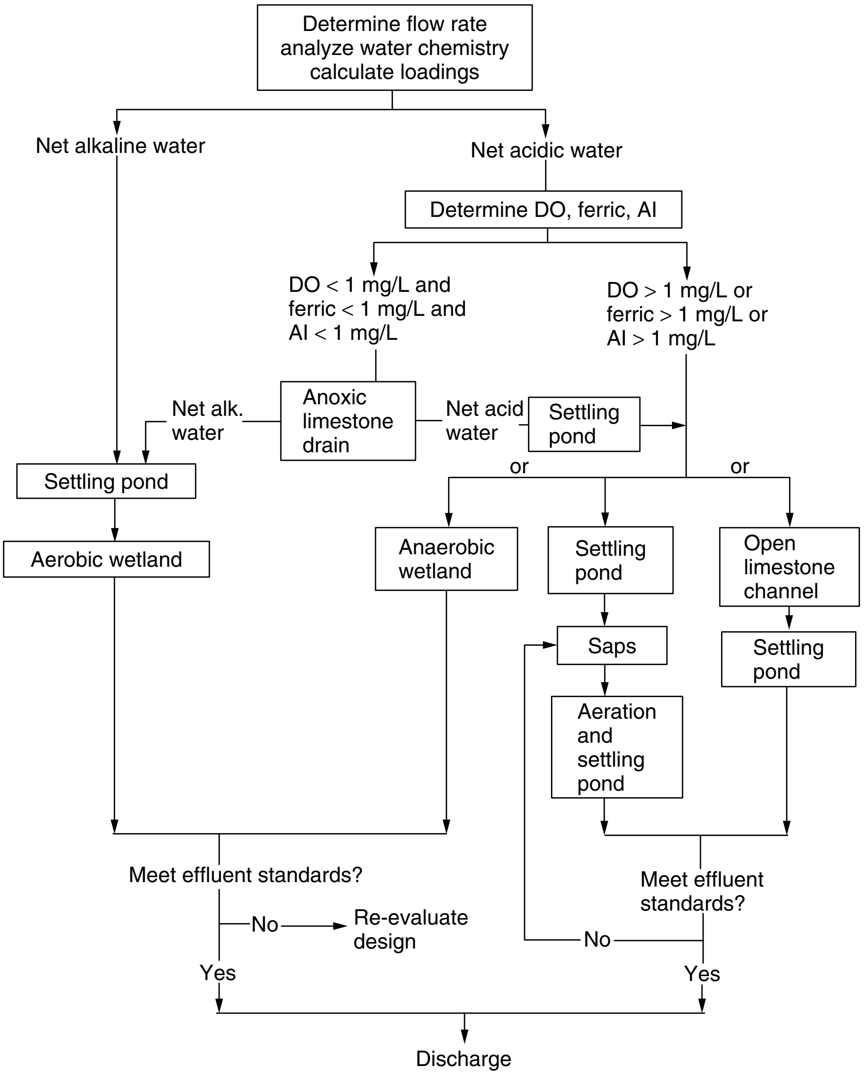

Figure 2. Flowsheet for selection of passive treatment method (after Ref. 1).

Anaerobic wetlands, in which the substrate for the wetland is compost and/or limestone, can be used to treat net acid water. However, the area of such wetlands must be much larger than aerobic wetlands or vertical flow ponds, so these are not common.

Discharge standards for mine waters specify removal of most manganese, which can be removed passively in aerobic limestone beds (8). The influent waters must be oxidizing and have negligible concentrations of Al and Fe. In these beds, Mn-oxidizing bacteria oxidize Mn and precipitate Mn oxides and hydroxides at pH 6.5 and higher. In some such facilities, the limestone bed is inoculated with specific Mn-oxidizing bacteria (9).

Recent publications indicate that periodic inspection, maintenance, and occasional rebuilding are necessary for most passive systems, especially those treating net acid water. Problems include washout during high flow, muskrat penetration of the dams, plugging of pipes by precipitates or vegetation, plugging of limestone by Al or Fe precipitates, accumulation of Fe oxides on top of compost, and channeling of flow through limited portions of the system.

BIBLIOGRAPHY

1. Skousen, J., Rose, A., Geidel, G., Foreman, J., Evans, R., and Hellier, W. (1998). *A Handbook of Technologies for Avoidance and Remediation of Acid Mine Drainage*. National Mine Land Reclamation Center, West Virginia University.
2. American Public Health Association (1998). Acidity (2310)/ titration method. In: *Standard Methods for the Examination of Water and Wastewater*, 20th Edn. L.S. Clesceri et al. (Eds.). American Public Health Association, Washington, DC, pp. 2.24–2.26.
3. U.S. Environmental Protection Agency (1979). Method 305.1, acidity (titrimetric). In: *Methods for Chemical Analysis of Water and Wastes*, U.S. Environmental Protection Agency Report EPA/600/4-79-020 Available at (www.nemi.gov).
4. Cravotta, C.A. and Kirby, X. (2004). Acidity and alkalinity in mine drainage: practical considerations. In: *Proceedings, American Society of Mining and Reclamation*, April 18–22, 2004, Morgantown, WV, pp. 334–365.
5. Hedin, R.S., Watzlaf, G.R., and Kleinmann, R.L.P. (1994). *Passive Treatment of Coal Mine Drainage: U.S. Bureau of Mines*, Information Circular IC 9389.
6. Kepler, D.A. and McCleary, E.C. (1994). Successive alkalinity producing systems (SAPS) for the treatment of acidic mine

drainage. In: *Proceedings, International Mine Drainage and Land Reclamation Conference*, April 24–29, 1994, Pittsburgh, PA, U.S. Bureau of Mines Special Publication SP06A-94, pp. 195–204.

7. Rose, A.W. (2004). Vertical flow systems—effects of time and acidity relations. In: *Proceedings, American Society of Mining and Reclamation*, April 18–22, 2004, Morgantown, WV, pp. 1595–1616.
8. Rose, A.W., Means, B., and Shah, P.J. (2003). Methods for passive removal of manganese. In: *Proceedings, West Virginia Surface Mine Drainage Task Force Symposium*, Morgantown, WV, April 15–16, 2003, pp. 71–82.
9. Vail, W.J. and Riley, R.K. (2000). *The Pyrolusite Process, A Bioremediation Process for the Abatement of Acid Mine Drainage*, Vol. 30, Greenlands, pp. 40–46.

BIOMARKERS AND BIOACCUMULATION: TWO LINES OF EVIDENCE TO ASSESS SEDIMENT QUALITY

M.L. Martín-Díaz
Instituto de Ciencias Marinas
de Andalucía
Cádiz, Spain

M.C. Morales-Caselles
N. Jiménez-Tenorio
I. Riba
T.A. DelValls
Facultad de Ciencias del Mar y
Ambientales
Cádiz, Spain

INTRODUCTION

Contaminated sediments pose a risk to aquatic life, human health, and wildlife throughout the world. There is an overwhelming amount of evidence that chemicals in sediments are responsible for toxicological (1) and adverse ecological effects (2). Frequently, the chemicals causing these effects are present in the sediment as mixtures of organic, metal, and other types of contaminants.

Identification of toxicants in sediments is useful in a variety of contexts. Adverse environmental effects from contaminated sediment resulted in international treaties and protocols for the environmental management of these dredged sediments. In this sense the OSPAR and Helsinki Conventions (North Sea, northeast Atlantic, Baltic Sea) proposed guidelines to control the disposal of sediment.

Once a toxicant is identified in the sediment, steps can also be taken to link a toxicant to a discharger and prevent further discharge. In addition, identification of major causes of toxicity in sediments may guide programs such as the development of environmental sediment guidelines and, retrospectively, aid regulators in determining the type of pesticide or manufactured chemical that may cause toxicity in the field. Here, we review the use of biomarkers and bioaccumulation in sediment quality assessment. These different tools provide a more sensitive measure of bioavailability and the effects of the different contaminants in the sediment.

SEDIMENT QUALITY ASSESSMENT

Sediments can act as deposits of contaminants entering the environment and, as a consequence, constitute a source of contamination. Today, remediation and management of contaminated sediments is becoming more and more economically and technologically demanding. Because of this, the scientific community has been developing science-based tools to identify sediments that are impaired and, ultimately, to support effective management decisions and priorities for dealing with contaminated sediments (3). When discussing the adverse effects of pollutants, some consider the accumulation of pollutant residues in the tissues of organisms to be adverse. Others consider an effect injurious only if changes occur in physiological processes in organisms, such as alterations in cellular morphology, metabolic activity, or physiological rates. Ecologists might restrict this definition still further to only those pollutant-induced effects that give rise to ecologically significant changes, that is, those at the population level (4). It is important to take into account the bioindicator species, the toxicity assay, and the different biological measurements that could show the exposure to a contaminant and the effect.

Different sensitive toxicity tests have been developed to assess sediment toxicity. The different toxicity tests can group pore water and whole sediment exposure tests. These tests are conducted using benthic organisms such as amphipods, polychaetes, algae, and juvenile fish (5). Use of these organisms allows one to estimate the effect of sediment ingestion and measure bioaccumulation and different biomarkers of exposure and effect. For example, the role of sediment ingestion, which should not be underestimated, has been identified as an important route of uptake of polycyclic aromatic hydrocarbons (PAHs) by deposit-feeding benthic organisms, as shown in experiments with polychaete worms (6,7).

More appropriately, three ecological life styles (filter feeding, deposit feeding, and burrowing) can be represented by the species used to assess sediment effects (8). Table 1 shows that the most appropriate phyla for potential impact testing in sediment is crustaceans, followed by annelids, mollusks, insects, fish, and echinoderms. Some of the species belonging to the taxonomic groups of crustaceans, mollusks, and fish are estuarine individuals that tolerate salinity fluctuation, which could affect the toxicity evaluation response in an environment affected by variation of physicochemical variables (9).

Bioaccumulation

Bioaccumulation values from different contaminant studies can help identify the bioavailability of chemicals in marine sediments and waters since, unlike chemical analysis, bioaccumulation provides a measurement of bioavailable contaminants (15).

Bioaccumulation is the uptake and retention of a bioavailable chemical from any possible external source (water, food, substrate, air). It is the net result of the uptake, distribution, and elimination of a substance in an organism due to exposure in water, food, sediment, and air. For bioaccumulation to occur, the rate of uptake from

Table 1. Summary of Species Appropriate for Testing Potential Impacts of Contaminated Sediment (10–14)

Taxonomic Group	Name	Test Species	Salinity Tolerance
Crustaceans	Mysid shrimp	*Americamysis* sp.	M
		Neomysis	M
		Americana	M
		Holmesimysis costata	
	Grass shrimp	*Palaemonetes* sp.	M, E
	Sand shrimp	*Crangon* sp.	M, E
	Shrimp	*Farfante penocus*	M
		Pandalus sp.	M
		Sicyonia ingentis	M
	Crab	*Callicnetes sapidus*	M, E
		Cancer sp.	M, E
		Carcinus maenas	M, E
	Amphipod	*Ampelisca* sp.	M
		Rhepoxynius sp.	M
		Eohaustarius sp.	M, E
		Grandiderella japonica	M, E
		Corophium insidiosum	M, E, F
		Leptocheirus plumulosus	M, E
		Hyalella azteca	M, E
		Corophium sp.	M
		Ampelisca brevicornis	M
		Mycrodentopus	M
		Gryllotalpa	M
	Cladoceran	*Daphnia magna*	F
		Ceriodaphnia dubia	F
Insects	Midge	*Chironomus tentans*	F
		C. riparius	F
	Mayfly	*Hexagenia limbata*	F
Molluscs	Mussel	*Anodonta imbecillis*	F
	Clam	*Yoldia limatula*	M
		Protothaca staminea	M
		Tapes japonica	M, E
		Ruditapes philippinarum	M
	Burrowing polychaete	*Nereis* sp.	M
		Neanthes arenaceodentata	M
		Nephthys sp.	M
		Glycera sp.	M
		Arenicola sp.	M
		Abarenicola sp.	M
Annelids	Oligochaete	*Pristina leidyi*	F
		Tubifex tubifex	F
		Lumbriculus variegates	F
Echinoderms	Sea urchin	*Strongylocentratus purparatus*	M
		Lytechinus pietus	M
		Echinocardium cordatum	M
		Paracentrotus lividus	M
Fish	Sheepshead minnow	*Cyprinodon variegates*	M, E
	Arrow gobi	*Clevelandia ios*	M
	Rainbow trout	*Oncorhynchus mykiss*	F, E
	Flat fish	*Solea senegalensis*	M
	Sea bream	*Sparus aurata*	M
	Turbot	*Scophthalmus maximus*	M

all sources must be greater than the rate of loss of the chemical from the tissues of the organism (16) (Fig. 1).

Toxic effects of anthropogenic compounds in biota and ecosystems are regarded in relation to their chemistry and fate in the environment. The bioavailability of chemicals, which is dependent on biogeochemical and physiological processes, is an important factor, often neglected in ecotoxicological evaluation and hazard assessment. The bioavailable fraction is critical for uptake and, ultimately, for the concentration at the target site in the organisms. The bioavailability of contaminants in sediments depends on several factors: physical (grain size of the sediment and suspended particulate materials), chemical (solubility, reactivity of compounds, complexing agents), and biological (benthic or pelagic organisms, mode of exposure) (17). Nevertheless, not all bioavailable

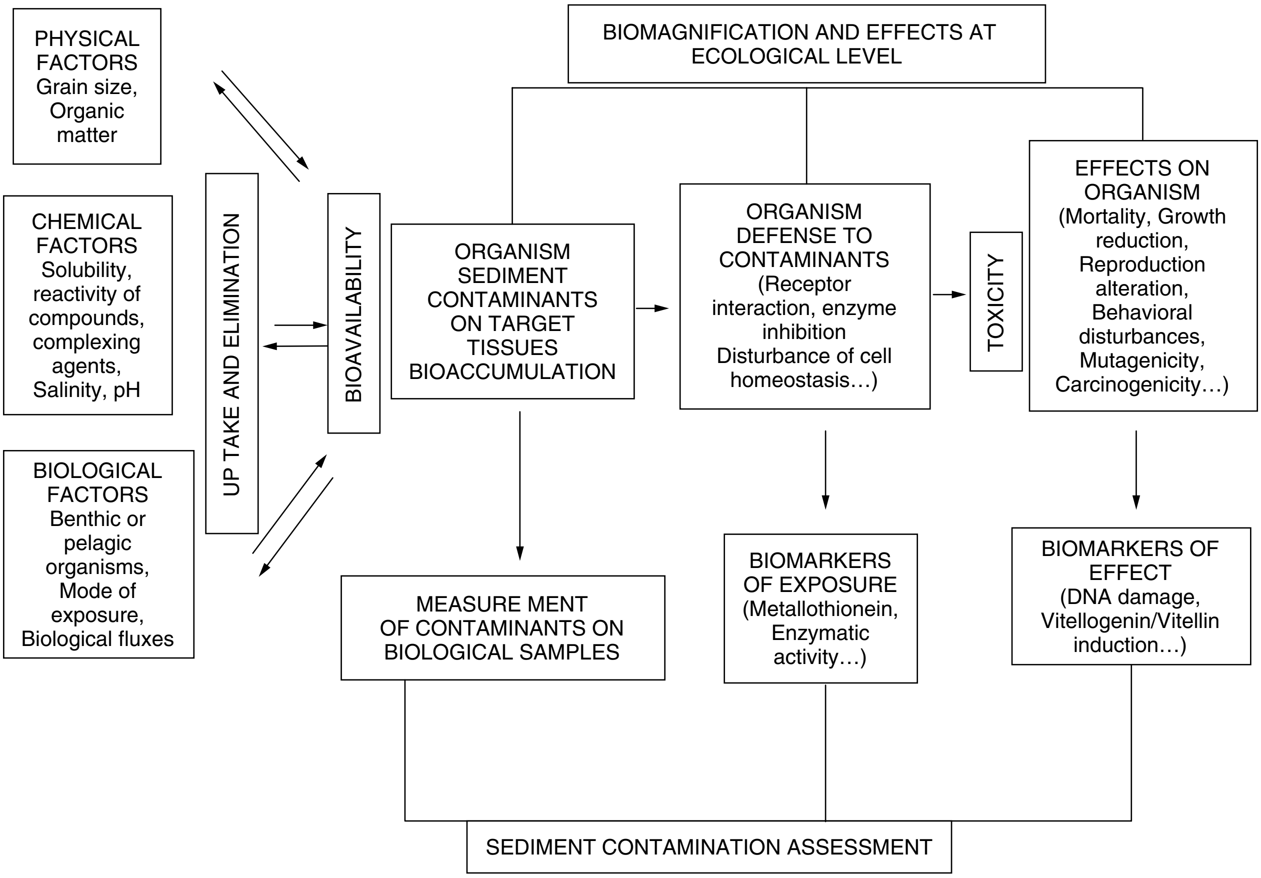

Figure 1. Ecotoxicological effects and assessment of the bioavailable fraction of pollutants in sediment.

chemicals can be bioaccumulated by the organisms in their tissues. For example, highly soluble chemicals, such as ammonia and some inorganic ions, are bioavailable and rapidly penetrate permeable tissues of marine organisms. However, they are not retained and are lost just as rapidly from the tissues by diffusion, metabolic transformation, or active transport. Other bioavailable chemicals are also taken up rapidly but are transformed and/or excreted rapidly by metabolic processes of the organism and are not bioaccumulated (16). In this sense, it is important to take into account the bioaccumulation factor (BAF), which is the ratio of the sum of the uptake rate constants of the chemical from all environmental compartments accessible to the organism to the sum of the release rate constants by active and passive mechanisms from the organism (18). The bioaccumulation factor allows one to predict if a bioavailable chemical is being bioaccumulated.

Once a chemical is bioaccumulated, the concentration of the contaminant at the target tissue in the organism induces molecular manifestations of defense and the effects can be determined (Fig. 1).

Persistent chemicals may accumulate in aquatic organisms through different mechanisms: via the direct uptake from water by gills or skin (bioconcentration), via uptake of suspended particles (ingestion), and via the consumption of contaminated food (biomagnification). Even without detectable acute or chronic effects in standard ecotoxicity tests, bioaccumulation should be regarded as a hazard criterion in itself, since some effects may only be recognized in a later phase of life, are multigeneration effects, or manifest only in higher members of a food web (e.g., impact of PCBs on the hatching success of eggs) (19).

Bioaccumulation of certain persistent environmental contaminants in animal tissues may be considered a biomarker of exposure (20,21). However, based on the definitions provided by Van Gestel and Van Brummelen (22), body burdens are not considered to be biomarkers or bioindicators since they do not provide information on deviations related to “health.” To avoid confusion, Van der Oost et al. (23) proposed that analytical–chemical indicators (body burdens) be referred to as bioaccumulation markers, while all biological (biochemical, physiological, histological and morphological) indicators measured inside an organism or its products be referred to as biomarkers.

The relationship between concentration of contaminants in tissues and toxic effects measured in organisms has received increased attention during the last few decades (1). These relationships should allow one to derive tissue quality guidelines (TQGs), defined as the concentrations of the chemicals measured in the different tissues that are associated or not associated with the biological effects measured using the sediment quality guidelines (SQGs) widely used around the world (24). To identify possible toxic agent(s), one must have body burden data collected from the same organisms exhibiting toxicity.

Biomarkers

Biomarkers are specific biological responses related to metabolism, detoxification, or toxicity induced by pollutants and associated with disposed materials. In this respect, an important advantage of biomarkers in assessing the impact of contaminated sediment is the inherent capability to detect early occurrence of various stress conditions within the organism and monitor the temporal progression (or regression) of the disturbance at various levels of biological organization (25).

Laboratory studies have documented the use of biomarkers to provide rapid quantitative predictions of toxicity on individual organisms. At this time, the use of biomarkers is not a replacement for traditional monitoring techniques, but it can be a useful supplementary approach to demonstrate links between sublethal biochemical exposure and decreases noted in field population studies (26).

A biomarker is defined as a change in a biological response (ranging from molecular though cellular and physiological responses to behavioral changes) which can be related to exposure to or toxic effects of environmental chemicals (27). Van Gastel and Van Brummelen (22) redefined the terms “biomarker,” “bioindicator,” and “ecological indicator,” linking them to different levels of biological organization. They considered a biomarker as any biological response to an environmental chemical at the subindividual level, measured inside an organism or its products (urine, feces, hair, feathers, etc.), indicting a deviation from the normal status that cannot be detected in the intact organism. A bioindicator is defined as an organism giving information on the environmental conditions of its habitat by its presence or absence or by its behavior. An ecological indicator is an ecosystem parameter, describing the structure and functioning of ecosystems.

One of the most important features of biomarkers is that they have the potential to anticipate changes at higher levels of the biological organization (i.e., population, community, or ecosystem). Thus, these “early warning” biomarkers can be used in a predictive way, allowing the initiation of bioremediation strategies before irreversible environmental damage of ecological consequences occurs. Biomarkers are then defined as short-term indicators of long-term biological effects.

According to the National Research Council (20) and the World Health Organization (21), biomarkers can be subdivided into three classes: biomarkers of exposure, biomarkers of effects, and biomarkers of susceptibility. The biomarkers of exposure cover the detection and measurement of an exogenous substance or its metabolite or the product of an interaction between a xenobiotic agent and some target molecule or cell that is measured in an organism. On the other hand, the biomarkers of effect show an established or possible health impairment or disease through a measurable biochemical, physiological, or other alteration in the organism. Finally, biomarkers of susceptibility indicate the inherent or acquired ability of an organism to respond to the challenge of exposure to a specific xenobiotic substance, including genetic factors and changes in receptors which alter the susceptibility of an organism to that exposure.

New techniques allow one to detect the effects of complex mixtures of contaminants. Many are diagnostic of causes, provide information on the bioavailability of contaminants, and allow more accurate assessments of potential ecological damage. Cellular and molecular indicators provide the greatest potential for identifying individuals and populations for which conditions have exceeded compensatory mechanisms and which are experiencing chronic stress, which, if unmitigated, may progress to severe effects at the ecosystem level.

The classification of biomarkers in the literature is very diffuse since biomarkers of exposure and those of effect are distinguished by the way they are used, not by an inherent dichotomy (28).

Good biomarkers are sensitive indices of both pollutant bioavailability and early biological responses (23,29–31). Recently, the biomarker approach has been incorporated into several pollution monitoring programs in Europe and the United States. Likewise, different methods for biological effect measurement have been evaluated in a series of practical workshops organized by the International Council for the Exploration of the Sea (ICES) and the Intergovernmental Oceanographic Commission (IOC), such as those in the North Sea (32). The United Nations Environment Programme has founded a biomonitoring program in the Mediterranean Sea including a variety of biomarkers (10). Biomarkers have also been included in the Joint Monitoring Programme of the OSPAR convention, where Portugal and Spain are members. Nevertheless, the biomarkers approach has not been included in the guidelines for the management and monitoring of dredging and disposal activities. The current guidelines for the control of these activities are based on several approaches, which take into account chemical measurements, analysis of benthic communities, and toxicity tests. Very few studies have been done on the utility of biomarkers.

In this sense, different studies on the use of biomarkers to assess the impact of contaminants on sediment (Table 2) are being carried out. Most of the studies focus on the determination of the activity of biotransformation enzymes, antioxidant enzymes, and biochemical indices of oxidative damage, DNA damage, and metallothioneins. Today, promising biomarker tools are the “genomics” and “proteomics.” In some cases genomic is used as a broad term—including proteomic—and it includes genomic sequencing, functions of specific genes, genome architecture, gene expression at transcriptome level, and protein expression at proteome level and metabolite flux (metabolomics).

USING BIOMARKERS AND BIOACCUMULATION FOR REGULATION PURPOSES

Frequently, the chemicals causing these effects are present in the sediment as mixtures of organic, metal, and other types of contaminants. The foundation for sediment quality assessment in the context of environmental risk assessment (ERA) is the weight-of-evidence (WOE), whose objective is to integrate multiple lines-of-evidence (LOE), both chemical and biological.

Table 2. Biomarkers Used for Sediment Toxicity Assessment in the Laboratory, *In Situ* (Caged Individuals), and in the Field (11–14,26,33–36)

Parameter	Description	Toxicity Assay	Pollutant Response
Metallothionein	Induction of this protein indicates exposure to metals	Field, caged individuals, laboratory	Cd, Cu, Zn, Hg, Co, Ni, Bi, Ag
GSH	Assay that determines the total glutathione content, a natural antioxidant	Field	PAHs, PCBs
DNA damage	Assay that detects single-strand breaks in DNA, a measure of DNA damage	Laboratory, caged individuals	PAHs, PCBs
EROD	Assay for ethoxyresorufin-*O*-deethylase, a phase I detoxification enzyme	Laboratory, caged individuals, field	PAHs, PCBs
CAT	Assay for catalase, an antioxidant enzyme	Caged individuals, field	PAHs, PCBs
SOD	Assay for superoxide dismutase, an antioxidant enzyme	Field	PAHs, PCBs
GR	Assay for glutathione reductase, an antioxidant enzyme	Laboratory, field	PAHs, PCBs
GPX	Assay for glutathione peroxidase, an antioxidant enzyme	Field	PAHs, PCBs
GST	Assay for glutathione-*S*-transferase, a phase II detoxification enzyme	Laboratory, field, caged individuals	PAH, PCB, BNF
LPO	Assay to determine the level of thiobarbituric reactive substances from lipid peroxide breakdown	Laboratory, caged individuals	Cd, PCB
Vitellogenin/vitellin	Induction of this protein indicates the exposure to substances that could perturb endocrine function	Laboratory, caged individuals	Cd, Zn, PAHs, PCBs
TOSC	Total oxyradical scavenging capacity	Field	PAHs, PCBs
Imposex, intersex	Sex change	Field, laboratory	TBT

Different tools are proposed in order to obtain multiple LOE in sediment quality assessment: (1) sediment chemistry including numeric sediment quality guidelines (SQGs), (2) toxicity tests, (3) bioaccumulation tests, (4) biomarkers, and (5) resident aquatic community structure. These tools should provide adequate estimation of the influence of the physical, chemical, and biological factors in the level of exposure and bioavailability of the different xenobiotics in the sediment (37).

These tools, expressing different lines of evidence, are integrated in environmental risk assessment methodologies and utilized in monitoring and assessment sediment programs. In this regard, a tiered approach to testing is recommended. At each tier, it will be necessary to determine whether sufficient information exists to allow a management decision to be taken or whether further testing is required (8).

The toxicological significance and complexity increase with each tier (Fig. 2). On Tier 0, a recompilation of information that already exists is developed. Tier I testing involves physical–chemical studies and ecotoxicological screening of the dredged material. If toxic effects are observed at Tier I (major hazard concern), the material may be evaluated on this basis. If no toxic effects are observed at this stage, but chemical analyses indicate that there may be reasons for concern, Tier II is entered and further biotests are requested. Tier II testing covers a wide range of effects parameters, including long-term and sublethal toxicity. In most cases, the Tier II results will allow a comprehensive evaluation of the (dredged) sediment. Only if the results are still inconclusive will it be necessary to add further tests on specific effects depending on evidence for particular contaminants, assessment of biomarkers of exposure and effect, and/or verification of laboratory measurements in the field (Tier III). Tier III testing may be used to monitor possible impacts of dredging operations in the field. The available techniques comprise nonspecific monitoring methods that respond to a wide range of environmental contaminants, biomarkers (fish, shellfish) of specific exposures depending on evidence for particular (bioaccumulation) contaminants, and assessment of long-term effects on benthos community structure and function (33).

FUTURE RESEARCH NEEDS

The evaluation of sediment quality in the environmental risk assessment context requires the use of more sensitive tools for sediment evaluation and the extrapolation to the field. The use of bioaccumulation and biomarkers in different bioindicator species is being researched for the assessment and management of sediment. There is an increasing consensus and interest in the valuable information that they could provide. Nevertheless, although they are more sensitive tools, they show higher variability with respect to the physical–chemical characteristics of the sediment, their extrapolation from laboratory to the field, and their intercalibration.

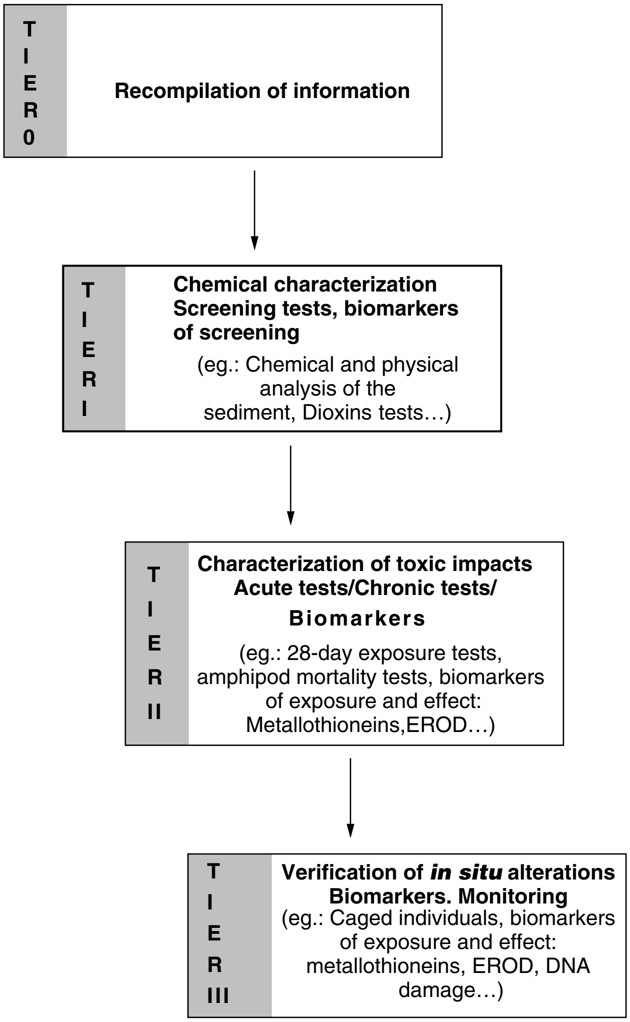

Figure 2. Schematic representation of the tier testing steps including biomarkers and bioaccumulation measurements.

In summary, some recommendations are proposed for the incorporation of biomarkers and bioaccumulation in research programs for the evaluation of dredged material quality using a weight of evidence approach:

1. Determination of biomarker screening responses to contaminated and control sediments for different sites and organisms in order to determine the advantages and disadvantages of the use of different biomarkers for the identification of different sources of contamination.
2. Study of the bioavailability of contaminants in sediment and biomarker responses through the exposure–dose–response triad methodology (exposure–chemical/biomarkers of exposure and bioaccumulation; dose/response—biomarkers of effect). It could allow the analysis of the sensitivity of different biomarkers to different availabilities of contaminants in sediment and to different contamination.
3. Measurement of biomarkers and bioaccumulation response to natural environmental conditions (pH, temperature, salinity). It could allow the application of biomarkers to environments that support high variability of environmental conditions, avoiding the “noise” that could produce this variability in a biomarker’s response.
4. Determination of biomarkers and bioaccumulation response over time through a toxicokinetic study of these responses. It could provide knowledge of biomarker and bioaccumulation behavior in the different organisms.
5. Intercalibration of bioaccumulation and biomarker responses; determination of biomarkers and bioaccumulation of the same contaminated sites in different laboratories. It could provide knowledge of the biomarkers and bioaccumulation response variation due to handling of organisms and methodology of measurement.
6. Biomarkers and bioaccumulation extrapolation from laboratory to field conditions through the determination of biomarkers and bioaccumulation response to laboratory and field contaminated toxicity exposure. These measurements could allow the analysis of responses due to realistic environmental conditions. It should be developed for three species: filter feeding, deposit feeding, and burrowing. The link between both sets of data will allow derivation of tissue quality values that need to be validated for human health and regulatory purposes.

Acknowledgments
This research has been partially supported by Plan Nacional I + D + I, REN2002-01699/TECNO, Ministerio de Fomento and the Spanish Ministry of Development (BOE 13-12-02) and of Education and Science (VEM2003-20563/INTER).

BIBLIOGRAPHY

1. Chapman, P.M. (1998). Death by mud: amphipod sediment toxicity test. In: *Microscale Testing in Aquatic Toxicology*. P.G. Wells, K. Lee, and C. Blaise (Eds.). CRC Press, Boca Raton, FL.
2. Andersen, B. et al. (1998). Comparision of marine sediment toxicity test protocols for the amphipod *Rhepoxonius abronius* and the polychaete worm *Nereis (Neanthes) arenaceodentata*. *Environ. Toxicol. Chem.* **17**: 859–866.
3. Environment Canada (2004). *Quality Guidelines*. Environmental Protection Publications.
4. Depledge, M.H. and Hopkin, S.P. (1995). In: R.A. Linthurst, P. Bourdeau, and R.J. Tardiff (Eds.). John Wiley & Sons, Chichester, UK.
5. Carr, R.S. (1998). Marine and estuarine porewater toxicity testing. In: *Microscale Testing in Aquatic Toxicology*. P.G. Wells, K. Lee, and C. Blaise (Eds.). CRC Press, Boca Raton, FL, pp. 523–538.
6. Forbes, T.L., Forbes, V.E., Giessing, A., Hansen, R., and Kure, L.K. (1998). Relative role of pore water versus ingested sediment in bioavailability of organic contaminants in marine sediments. *Environ. Toxicol. Chem.* **17**: 2453–2462.
7. Weston, D.P. and Mayer, L.M. (1998). *In vitro* digestive fluid extraction as a measure of the bioavailability of sediment-associated polycyclic aromatic hydrocarbons: sources of variation and implications for partitioning models. *Environ. Toxicol. Chem.* **17**: 820–829.

8. U.S. Environmental Protection Agency and U.S. Army Corps of Engineers (U.S. EPA and U.S. ACE) (1998). *Evaluation of Dredged Material Proposed for Discharge in Waters of the US: Testing Manual*, EPA/823-B-94/002. EPA, Washington, DC.
9. DelValls, T.A., Casado-Martínez, M.C., Martín-Díaz, M.L., and Riba, I. IV (2004). Workshop SedNet.
10. UNEP (1997). *Report of the Meeting of Experts to Review the MEDPOL Biomonitoring Programme*, Athens, Greece. UNEP- (OCA)/MED WG, pp. 132–137.
11. CEDEX (2003). Recomendaciones para la gestión del material de dragado en los puertos Españoles. Centro de Estudios y Experimentación de Obras Públicas, Puertos del Estado, Madrid.
12. Meussy, J.J. and Payen, G.G. (1988). Female reproduction in malacostracan crustacea. *Zool. Sci.* **5**: 217.
13. Lafontaine, Y. et al. (2000). Biomarkers in zebra mussels (*Dreissena polymorpha*) for the assessment and monitoring of water quality of St. Lawrence River (Canada). *Aquat. Toxicol.* **50**: 51–71.
14. Regoli, F. et al. (2002). Application of biomarkers for assessing the biological impact of dredged materials in the Mediterranean: the relationship between antioxidant responses and susceptibility to oxidative stress in the red mullet (*Mullus barbatus*). *Mar. Pollut. Bull.* **44**: 912–922.
15. Furley, T.H. and Niencheski, L.F. (1993). Temporal and spatial variability of Cd, Pb, Zn, Cu and Mn concentration in mussels (*Perna perna*, Linne, 1758) from the coast of Rio Grande do Sul, Brazil. In: *Proceedings of the International Symposium on Perspectives for Environmental Geochemistry in Tropical Countries*. J.J. Abrao, J.C. Wasserman, and E.V.F. Silva (Eds.). Universidade Federal Fluminense, Niteroi, R.J., Brazil, pp. 39–40.
16. Neff, M. (2004). *Bioaccumulation in Marine Organisms*. Elsevier Ltd., Oxford, UK.
17. Borgmann, U. (2000). Methods for assessing the toxicological significance of metals in aquatic ecosystems: bioaccumulation–toxicity relationships, water concentrations and sediment spiking approaches. *Aquat. Ecosyst. Health Manage.* **3**: 277–289.
18. Farrington, J.W. and Westal, J. (1986). Organic chemical pollutants in the oceans and groundwater: A review of fundamental chemical properties and biogeochemistry. In: *The Role of the Oceans as a Waste Disposal Option*. G. Kullenberg (Ed.). Reidel Publishing, New York, pp. 361–425.
19. Tillit, D.E. et al. (1992). Polychlorinated biphenyl residues and egg mortality in double-crested cormorants from the Great Lakes. *Environ. Toxicol. Chem.* **11**: 1281–1288.
20. NRC: Committee on Biological Markers of the National Research Council (1987). Biological markers in environmental health research. *Environ. Health Perspect.* **74**: 3–9.
21. WHO: International Programme on Chemical Safety (IPCS) (1993). *Biomarkers and Risk Assessment: Concepts and Principles*, *Environmental Health Criteria* 155. World Health Organization, Geneva.
22. Van Gestel, C.A.M. and Van Brummelen, T.C. (1996). Incorporation of biomarker concept in ecotoxicology calls for a redefinition of terms. *Ecotoxicology* **5**: 217–225.
23. Van der Oost, R., Beyer, J., and Vermeulen, N.P.E. (2003). Fish bioaccumulation and biomarkers in environmental risk assessment. *Environ. Toxicol. Pharmacol.* **13**: 57–147.
24. Riba, I., Forja, J.M., Gómez-Parra, A., and Del Valls, T.A. (2004). Sediment quality in littoral regions of the Gulf of Cádiz: a triad approach to address the influence of mining activities. *Environ. Pollut.* 341–353.
25. Maycock, D.S. et al. (2003). Incorporation of *in situ* and biomarker assays in higher tier assessment of the aquatic toxicity of pesticides. *Water Res.* **37**: 4180–4190.
26. Hyne, R.V. and Maher, W.A. (2003). Invertebrate biomarkers: links to toxicosis that predict population decline. *Ecotoxicol. Environ. Safety* **54**: 366–374.
27. Peakall, D.W. (1994). Biomarkers: the way forward in environmental assessment. *Toxicol. Ecotoxicol. News* **1**: 55–60.
28. Suter, G.W. (1993). *Ecological Risk Assessment*. Lewis Publishers, Boca Raton, FL.
29. Stegeman, J.J., Brouwer, M., Richard, T.D.G., Förlin, L., Fowler, B.A., Sanders, B.M., van Veld, P.A. (1992). In: *Biomarkers: Biochemical, Physiological and Histological Markers of Anthropogenic Stress*. R.J. Huggett, R.A. Kimerly, P.M. Mehrle Jr., and H.L. Bergman (Eds.). Lewis Publishers, Chelsea, MI, p. 235.
30. Pedersen, S.N., Lundebye, A.K., and Depledge, M.H. (1997). Field application of metallothionein and stress protein biomarkers in the shore crab (*Carcinus maenas*) exposed to trace metals. *Aquat. Toxicol.* **37**: 183–200.
31. Fossi, M.C., Savelli, C., and Casini, S. (1998). Mixed function oxidase induction in *Carcinus aestuarii*. Field and experimental studies for the evaluation of toxicological risk due to Mediterranean contaminants. *Comp. Biochem. Physiol.* **121C**: 321–331.
32. Stebbing, A.R.D. and Dethlefsen, V. (1992). Introduction to the Bremerhaven workshop on biological effects of contaminants. *Mar. Ecol. Prog. Ser.* **91**: 1–8.
33. OSPAR (1995). *Report of the workshop on biological effects monitoring techniques*. Abeerden, Scotland, 2–6 October.
34. Fredrickson, H.L. et al. (2001). *Sci. Total Environ.* **274**: 137.
35. Martín Díaz, M.L., Sales, D., and Del Valls, T.A. (2004). Influence of salinity in hemolymph vitellogenin of the shore crab *Carcinus maenas* to be used as biomarker of contamination. *Bull. Environ. Contam. Toxicol.*, in press.
36. Martín Díaz, M.L., Blasco, J., Sales, D., and DelValls, T.A. (2004). *International Conference on Remediation of Contaminated Sediments*. Batelle Press, Columbus, OH.
37. Wenning, R.J. and Ingersoll, C.G. (2002). Summary booklet of a SETAC Pellston workshop on use of sediment quality guidelines and related tools for the assessment of contaminated sediments; Fairmont, Montana. Society of Environmental Toxicology and Chemistry (SETAC), Pensacola, FL.

LEAD AND ITS HEALTH EFFECTS

Shahida Quazi
Dibyendu Sarkar
Rupali Datta
University of Texas
San Antonio, Texas

INTRODUCTION

Lead (Pb) is a bluish white, lustrous, naturally occurring element that people have used almost since the beginning of civilization. Because of its malleability, Pb was one of the first metals that humans began to use for industrial purposes. Plumbism, or Pb poisoning, has been recognized for many centuries and is one of the oldest ailments afflicting humans. Nikander (200 B.C.) was first to describe the symptoms of Pb poisoning, which included colic and pallor. The English aristocracy of the seventeenth and eighteenth centuries suffered from widespread Pb poisoning from consumption of Portuguese wine, transported with submerged Pb bars to enhance flavor and to prevent spoilage. Lancereaux provided the first description of kidney disease and interstitial nephritis by postmortem examination of a Pb poisoned artist.

Anthropogenic activities during the past century have resulted in increased accumulation of toxic metals in soils and natural waters (1). Lead contamination is widespread in urban areas due to industrial emissions, extensive use of alkyl-lead compounds as antiknock additives in gasoline in the past, and lead-based paints and pipes (2), resulting in soil contamination on a global scale with adverse effects on human and environmental health (3).

Lead occurs naturally in low concentrations in all rocks, soils, and dust, usually ranging from 2 to 200 ppm (4). The periodic position of Pb favors the formation of covalent rather than ionic bonds in Pb^{4+} compounds. Lead is similar to certain alkaline earth metals, particularly calcium. Both calcium and Pb form insoluble carbonates and phosphates. However, Pb carbonates and phosphates are less soluble compared to those of calcium. This insolubility of Pb together with its high affinity to biological donors (enriched in oxygen and possibly nitrogen) accounts for its toxicological properties (5).

Here, we provide an overview of the various sources of Pb, its impact on the ecosystem and its effects on human health. Future directions in research related to the effects of Pb on human health are also discussed.

SOURCES OF LEAD

Natural Sources

The average concentration of Pb in the earth's crust is estimated to be 16 mg/kg (4). Lead is mainly found in the form of its sulfide, galena. Some other ores of Pb include cerussite and anglesite. In general, granitic rocks that have high acidic content tend to have more Pb than basaltic rocks. Lead concentrations in soils (farming and city) range from 2 to 200 mg/kg, with most soils having Pb in the range of 5–25 mg/kg (6). Moreover, soils that are underlain by granitic rocks with high acid content have higher Pb concentrations than those overlying basaltic and other alkaline rocks (6). It has been estimated that global natural emissions of Pb to the atmosphere are in the range of $18.6–29.5 \times 10^6$ kg/yr.

Anthropogenic Sources

The major anthropogenic sources of Pb are industrial activities such as mining operations (nonferrous metal, secondary Pb smelters), vehicle emissions, coal burning, refuse incineration, and industrial outdoor paints applied to structural surfaces. However, with the phasing out of Pb in gasoline (which began in the 1970s), Pb in soils and house dusts originating from lead-based paints have become the principal sources of exposure in the United States (7).

In rivers or water bodies, the concentration of Pb mainly depends on local inputs. Since most Pb compounds have low solubility, high concentrations are very rare. However, they are found in acid mine drainage in high concentrations mostly in areas of Pb mineralization (8). Lead shots used for hunting and fishing can also result in significant Pb input to both soil and water and contribute to inputs to surface waters (9), ultimately adding to the Pb burden in the sediments.

MOVEMENT OF LEAD IN THE ENVIRONMENT

The passage of metals through the atmosphere is integral to their biogeochemical cycling. Because of the dynamic nature of the atmosphere, metals can be deposited in areas remote from their initial source. Air is the primary medium for the transport of Pb. Fine particles of Pb generated anthropogenically are transported over long distances until they settle through wet, dry, or cloud deposition (5). Lead transport in the environment depends on the size of the particles. While larger particles (comprising 95% of the emission) settle down after sometime, the smaller Pb particles can travel hundreds of kilometers before they are deposited (4). Lead in the soil and dust is relatively unavailable biologically. It normally reacts with the soil to form insoluble salts. Moreover, Pb also complexes with components of organic matter such as humic acid and fulvic acid, thus making Pb more unavailable (6). The downward movement of Pb from soil by leaching is very slow under most natural conditions (10). Mobility of Pb in the soil is dictated by a number of geochemical process, including oxidation/reduction, precipitation/dissolution, adsorption/desorption, and complexation/chelation. Metallic Pb can be oxidized to more labile forms, the rate of oxidation and the oxidation products being dependent on the site. Once Pb is oxidized, it can be precipitated in the form of hydroxides, sulfates, sulfides, carbonates, and phosphates. The solubility of these precipitates is dependent on factors such as pH and oxidation–reduction conditions. In general, Pb is more soluble under acidic conditions compared to neutral and alkaline conditions. Most of the Pb accumulated in the soil may be directly taken up by grazing animals and microorganisms or may be ingested by children through hand-to-mouth activity. From the grazing animals Pb may enter the food chain (5).

Lead in rivers comes from runoff, erosion, and direct deposition from air. Fresh water generally contains more inorganic and organic suspended material than marine water, and this suspended material has a tendency to adsorb any dissolved Pb. In the case of the oceans, most of the Pb comes from atmospheric deposition rather than from rivers (4).

ECOTOXICOLGICAL EFFECTS OF LEAD

The uptake of Pb in both aquatic and terrestrial ecosystems is determined by its bioavailability. The term "bioavailability" reflects the premise that, for some heavy metals, organisms may be exposed to less than the total amount present in their habitat (11). The extent to which heavy metals react with and affect biological components is immensely influenced by the physicochemical factors of the specific environment (12). In general, the systemic availability of Pb is matrix and chemical species dependent (13–15).

Toxicity to Aquatic Organisms

Dose-related increases in lead concentrations of submerged and floating plants have been observed upon their exposure to dissolved Pb, accumulation being the greatest in the roots, followed by stems and leaves (16). In a study conducted by Kay et al. (17) water hyacinth (*Eichhornia crassipes*) was exposed to solutions containing lead nitrate concentrations of 0–5 mg/L for 6 weeks. The accumulation of Pb was dose-dependent and in the order of roots > stems > leaves. Uptake of Pb from shot-contaminated sediments has been observed in aquatic plants. In an investigation of a skeet target shooting range located on the shores of Lake Merced, California, sediments contained Pb levels up to 1200 μg/g in the shotfall zone, and tule seedheads and coontails growing within these sediments exhibited Pb concentrations averaging 10.3 and 69.2 μg/g dry weight, respectively, compared to concentrations of 2.3 and 11.9 μg/g dry weight, respectively, at control sites (18). However, Van der Werff and Pruyt (19) reported that application of lead nitrate with a Pb concentration of up to 10^{-5} mol/L had no observable toxicity on four aquatic plants, *Elodea gibba, Callitiche platycarpa, Spirodela polyrhiza*, and *Leman gibba*. In the case of fishes, it has been found that *in vivo* sublethal dietary doses of Pb (II) and tributyltin (TBT) induced toxic effects in *H. malabaricus* (20). Davis et al. (21) conducted long-term bioassays with rainbow trout to establish a maximum acceptable toxicant concentration (MATC) limit for inorganic Pb. Fingerling trout were exposed to nominal total Pb concentrations of 0, 40, 120, 360, 1080, or 3240 mg/L. Actual dissolved Pb was measured and results were expressed in terms of dissolved salt. They reported that MATC values between 0.018 and 0.032 mg/L caused the "black tail effect." A similar bioassay with soft water in fishes hatched from exposed eggs suggested MATC values between 0.041 mg/L (where no black tails occurred) and 0.076 mg/L (where 4.7% of fish showed the black tail effect). When fingerlings from nonexposed eggs were used in a soft water bioassay, the MATC for the black tail effect was between 0.072 mg/L (where no black tails occurred) and 0.146 mg/L (where 41.3% of fish had black tails). There were no significant differences between the measured dissolved Pb concentrations in the two tests, indicating that fish from exposed eggs were more sensitive to the effects of Pb than those from nonexposed eggs.

Maddock and Taylor (22) investigated the uptake of organolead by shrimp, mussel, and dab (a flat fish) in short-term experiments and in mussel and dab in long-term experiments. Bioconcentration factors were higher for tetraalkyl Pb than for trialkyl Pb in the short-term exposure. In the long-term experiment conducted on mussels, the greatest tissue concentration of Pb occurred in the gills; with the digestive glands, gonads, and feet containing progressively less Pb (17).

Toxicity to Terrestrial Organisms

Toxicity to Plants. Although Pb is considered to be toxic to plants just like other life forms, sensitivity of plants to Pb and responses of different plant species to Pb vary according to their genetic and physiological makeup (23). The metal remains largely as a superficial deposit or topical aerosol coating on plant surfaces (24,25). Toxicity of Pb to a large extent depends on its absorption, transport, and cellular localization (23). Plants of *Cassia tora* and *C. occidentalis* growing by the roadside accumulated up to 300 mg/g dry weight of Pb (26). Furthermore, accumulated Pb content generally increases with the increase in Pb in the environment, as has been reported for maize and pea leaves and Vigna and sesame roots (23) and leaves (27, 28). Photosynthesis is considered as one of the most sensitive metabolic processes to Pb toxicity (23). Moreover, nitrate reduction is inhibited drastically in roots by the metal, but in the leaves a differential effect is observed in various cultivars (23). Nitrate assimilation is inhibited by 10–100 mM Pb^{2+} in sorghum leaves (29). Lead also inhibits nodulation, N fixation, and ammonium assimilation in the root nodules (23).

Toxicity to Invertebrates. Long-term exposure of terrestrial invertebrates to lead-contaminated soil can result in high tissue concentrations, but one of the detoxification mechanisms within their cells is the subcellular compartmentalization of Pb. This results in only a fraction of the total body Pb burden actually contributing to toxicity (30). Doelman et al. (31) incubated a mixed culture of bacteria in lead nitrate solutions and also grew the fungus *Alternaria solani* on malt agar to which lead nitrate had been added. The cultures were used as food for the nematodes *Mesorhabditus monohystera* and *Aphelenchus avenae,* which were reared for up to 22 days on bacteria and fungus, respectively. Lead was taken up by bacteria to give a range of doses to the nematodes of between 7.6 and 110 μg/g of food. All these exposures had a significant inhibitory effect on the reproduction of *Mesorhabditus monohystera*. However, when woodlice *(Porcellio scaber*) were exposed to treated soil litter containing between 100 and 12,800 mg/kg dry weight of Pb as lead oxide over 64 weeks, no significant effect was found on adult survival, number of young produced, or the survival of the young at exposures up to 6400 mg lead/kg (32).

Toxicity to Birds. In the case of birds, metallic Pb is highly toxic when it is given in the form of Pb shots. Lead poisoning through the ingestion of spent gunshots is a widely recognized waterfowl mortality factor (33). Water fowl ingest spent gunshot when searching for grit (and possibly for food) particles. Once in the bloodstream, Pb is rapidly (but reversibly) deposited in soft tissues, mainly liver and kidney, and (relatively irreversibly) into bone (34). If a large quantity of Pb (e.g., >10 shots) is ingested within a short time, acute Pb poisoning may occur, resulting in death within a few days. However, most of the mortality results from the ingestion of a smaller number of shots, with birds dying of chronic poisoning 2–3 weeks following ingestion (34). The U.S. Fish and Wildlife Service (35) reported that Pb poisoning from ingested sinker and jigs accounts for 10–15% mortality of common loons (*Gavia immer*) with a study revealing that 27% of adult loons had fishing tackle in their stomachs and high Pb levels in their blood. Signs of chronic Pb poisoning include green and watery feces, dropping wings, anemia, weight loss, and atypical behavior; the affected birds die approximately 2–3 weeks after ingesting the shot, often in a very weak condition (36). Secondary Pb shot poisoning can occur when a predator or scavenger consumes animals that have been shot with Pb shot shell ammunition or consumes the gizzard of a bird that has ingested Pb shot (36). Lead also causes a decrease in clutch and egg size, mortality of embryos, depression of growth, and deficits in behavior, thus affecting the survival of birds (37).

EFFECTS ON HUMANS

Lead is present in various forms in the environment, such as free hydrated ions, ion-pair salts/complexes, and organic complexes/chelates (38). Humans are most likely to be exposed to the inorganic forms of Pb such as halides, oxides, sulfides, carbonates, and chromates. Since Pb was phased out as a gasoline additive (tetraethyl Pb) in the 1970s and its use in paint and food containers (e.g., ceramic ware and tin cans) was curtailed, blood Pb concentrations (BLLs) have decreased significantly. However, other sources of Pb and its unknown threshold of subclinical toxicity continue to make Pb an issue of public health concern (39). Lead exposure in the general population (including children) occurs primarily through ingestion, although inhalation also contributes to Pb body burden and may be the major contributor for workers in Pb-related occupations (40). Children's hand-to-mouth activity, increased respiratory rates, and increased intestinal absorption of Pb make them more susceptible to Pb exposure than adults (41,42). Exposure to Pb can result in significant adverse health effects to multiple organ systems. Once Pb is absorbed into the blood plasma, it rapidly equilibrates with the extracellular fluid. More slowly, but within minutes, Pb is transferred from plasma into blood cells (43,44). From the blood plasma, the absorbed Pb is transferred to the liver and the kidneys. Lead tends to accumulate in areas with high levels of calcium; hence, the highest body burden of Pb is found in the bones. In a typical adult, 95% of the Pb is stored in the bones, whereas in children, it amounts to 70% (45,46). During childhood, Pb accumulates predominantly in the trabecular bones. But in the case of adults most of the Pb is found in both cortical and trabecular bones. In both the trabecular and cortical bones there are two physiological compartments, one labile and the other inert for Pb (7). While Pb may be exchanged between the blood and the bone through the labile component, Pb may be stored for years in the inert component (7).

The distribution of Pb in tissues reflects a state of constant, dynamic equilibrium. Any situation that mobilizes the very large, relatively stable pools of Pb within the body, particularly those in the bones, will lead to the redistribution of Pb to a variety of tissues. Redistribution is known to occur during pregnancy, which results in increased risk to the fetus, particularly in women with prior Pb poisoning (5).

The various detrimental human health effects of Pb include:

- Neurological effects including effects on the central and peripheral nervous systems.
- Effects on both male and female reproduction.
- Renal effects.
- Effects on vitamin D metabolism.
- Cardiovascular effects.
- Hematopoietic effects.

Neurological Effects

Lead affects both the central and peripheral nervous systems. An extensive database has provided a direct link between low level Pb exposure and deficits in the neurobehavioral–cognitive performance evidenced in childhood through adolescence (47). Many of the biological aberrations produced by Pb appear to be related to the ability of this heavy metal to either inhibit or mimic the action of calcium. Calcium ions play a special role in the release of neurotransmitters from presynaptic nerve endings (48). A large number of experimental studies have investigated the effects of prenatal and neonatal Pb exposure on central nervous system development and behavior. These studies have shown variable results, forming a spectrum that includes delayed nervous system development, deficits in visual motor function, abnormal social and aggressive behavior, hyperactivity and hypoactivity, as well as no changes in activity (49). The brain and the nervous systems of children and fetuses are generally vulnerable to Pb intoxication because of the incomplete blood–brain barrier and also because both the brain and the nervous system are still developing (40).

Effects on Central Nervous System

In the central nervous system, Pb toxicity is more common in children than adults and may produce either overt symptoms of acute encephalopathy, such as ataxia, headache, convulsions, and coma, or lesser deficits including learning disorders and hyperactive behavior (50,51). Lead disrupts the main structural components of the blood–brain barrier through primary injury to astrocytes with a secondary damage to the endothelial microvasculature. Within the brain, lead-induced damage occurs preferentially in the prefrontal cerebral cortex, hippocampus, and cerebellum (47) Lead enhances the spontaneous or basal release of neurotransmitters from presynaptic nerve endings at low concentrations (48). In addition to enhancing the spontaneous release of neurotransmitters, Pb blocks the release of neurotransmitters normally produced by depolarization of nerve endings. It modulates neurotransmitter release by altering calcium metabolism either by competing with calcium for entry into the cell or by increasing intracellular calcium levels (48). Moreover, high dose exposure to Pb (i.e., BLLs in excess of

4 μM) disrupts the blood–brain barrier (48). In general, blood–brain barrier excludes plasma proteins and most organic molecules and limits even the passage of ions such as sodium and potassium (52). However, at high BLLs large molecules such as albumin, which are normally excluded, freely enter the brain of immature animals exposed to high concentrations of Pb (53,54). Ions and water follow and edema is produced, resulting in the rise of intercranial pressure because of the physical restraint of the skull. When the intercranial pressure approaches the systemic blood pressure, cerebral perfusion decreases and brain ischemia occurs (48).

Children exposed to high levels of Pb suffer from encephalopathy and hyperirritability, ataxia, convulsions, stupor, and coma or death. In general blood, BLLs of about 70–80 μg/dL or greater in children pose a threat (7). However BLLs as low as 10 μg/dL or less can adversely affect the developing nervous system of a child (40). Studies have shown that for every 10 μg/dL increase in the BLL, there is a decrease of children's IQ by about four to seven points (55–60). In the case of adults, Pb encephalopathy may occur at BLLs of about 460 μg/dL (61).

Effects on Peripheral Nervous System

Peripheral neuropathy is a manifestation of Pb toxicity in adults with excessive occupational exposure. Children exposed to high levels of Pb may also suffer from peripheral neuropathy. While in adults the most common manifestation of peripheral neuropathy is wrist drop, in the case of children, general weakness and foot drop are more common (62).

Effects of Lead on Reproduction

Male Reproductive System. An adverse effect of Pb on the male reproductive system has been reported in several epidemiologic studies (63). It has been reported that BLLs below the currently accepted working protection criteria adversely affect spermatogenesis (64). Earlier studies on Pb workers had shown androgenic dysfunction including asthenospermia, hypospermia, and teratospermia that could be produced by direct toxicity to the testis (65). It has been shown that significant levels of asthenospermia and teratospermia were found in workers having BLLs over 500 μg/L (66). The effect of Pb poisoning on the male reproductive system can be seen at BLLs of about 40 μg/dL, with long-term Pb exposure resulting in diminishing sperm counts, sperm concentrations, and total sperm motility (7). Moreover, it was found that the risk of stillbirth or birth defects was elevated for preconception employment in a high Pb exposure environment compared with low Pb exposure jobs (67).

Female Reproductive System. There have been few well-documented reports on reproductive effects of Pb in humans (68). Exposure may result in increased preterm delivery. It was found that preterm delivery was statistically significantly correlated with maternal BLLs in a dose-responsive manner (69). Moreover, due to changes in the bone physiology and mineral metabolism during pregnancy, Pb may move from the bone into the maternal circulation. Lead does cross the human placenta as early as the 12th week of gestation (70). Cord-blood content of Pb shows a high correlation with maternal-blood samples, suggesting that infants are exposed in uterus to BLLs equivalent to those of the mother (71). Recent isotopic speciation studies have demonstrated that the skeletal contribution to BLLs increases from 9% to 65% during pregnancy (72). The maternal Pb burden has been reported to be negatively associated with the birth weight of infants (73). It has been found that prenatal exposure to low Pb levels (e.g., maternal BLLs of 14 μg/dL) may increase the risk of reduced birth weight and premature birth (7).

Effects on the Renal System

Lead affects the renal system in three stages. In the first stage, there is proximal tubular dysfunction (Fanconi's syndrome) manifested by aminoaciduria, glycosuria, and phosphaturia (74). Then chronic exposure results in the second stage that is characterized by gradual tubular atrophy and interstitial fibrosis. There is reduced incidence of inclusion bodies, and glomerular filtration is impaired (75). At the third stage, renal failure occurs and is characterized by renal tubular neoplasia or adenocarcinoma (76). In general, individuals with Pb levels exceeding 60 μg/dL are at a risk of developing renal failure (77,78).

It has been reported that Pb inhibits rBAT-induced amino acid transport in a noncompetitive, allosteric fashion. This blockade of rBAT-induced amino acid transport may be involved in aminoaciduria following Pb intoxication (79). Lead exposure also results in the onset of "saturnine gout" as a result of lead-induced hyperuricemia due to decreased renal excretion of uric acid. Most documented renal effects for occupational workers are observed in acute and high to moderate chronic exposures with BLLs of more than 60 μg dL (40).

Endocrine Effects/Effects on Vitamin D Metabolism

From the studies of children exposed to high Pb levels, it has been found that a correlation exists between BLLs and the level of vitamin D. Lead interferes with the conversion of vitamin D to its hormonal form, 1,25-dihydroxyvitamin D, which is responsible for the maintenance of extracellular and intracellular calcium homeostasis. Thus, impairment by Pb results in impaired cell growth, maturation, and tooth and bone development (40).

Cardiovascular Effects

Lead exposure is one of the factors that may contribute to the onset of hypertension. Other factors contributing to the development of hypertension include old age, increased weight, poor diet, and excess alcohol intake (40). Studies have shown that greater exposure to Pb in occupational environments may increase the risk of hypertensive heart disease and cardiovascular disease. There are several reports that have associated Pb exposure to elevation in blood pressure (80–83).

Hematopoietic Effects

Lead inhibits the body's ability to make hemoglobin by interfering with several enzymatic steps in the heme pathway (40). Heme is the prosthetic group of hemoglobin, myoglobin, and cytochromes. In mammals, heme is synthesized from succinyl-CoA and glycine in eight enzyme-mediated steps (Fig. 1) (84). Enzyme studies of the heme biosynthetic pathway have shown that Pb is an inhibitor of δ-aminolevulinate dehydratase (ALAD), corporphyrinogen oxidase, and ferrochelatase (85). However, the metal has the greatest influence on ALAD, and measurement of ALAD activity can be used as an indicator of Pb levels in the blood (84). Heme synthesis occurs partly in mitochondria and partly in cytoplasm. The three important steps in the synthesis of heme that are influenced by Pb intoxication are:

1. The condensation reaction of two molecules of δ-aminolevulinic acid that is catalyzed by PBG synthase (porphobilinogen synthase) or ALA dehydratase to form porphobilinogen (PBG). ALA-D is the enzyme that is most sensitive to Pb. Inhibition of the enzyme results in the prevention of the utilization of ALA and subsequently a decline in heme synthesis (68).
2. The rate-limiting step in heme biosynthesis is the ALA synthase catalyzed step. The oxidation product of heme is hemin, which acts as a negative feedback inhibitor of ALA synthase and inhibits the transport of ALA synthase from the cytosol to the mitochondria and also represses the synthesis of the enzyme. Lead interferes with the negative feedback control of heme synthesis; ALA synthetase activity is depressed, resulting in increased activity of the enzyme and increased synthesis of ALA (68).
3. The third reaction that is influenced by the intoxication of Pb is the incorporation of the ferrous iron into the porphyrin ring structure that is catalyzed by the enzyme ferrochelatase. Lead inhibits the enzyme ferrochelatase, thus preventing the introduction of iron into the protoporphyrin IX to form heme. The porphyrin chelates with Zn nonenzymatically to form ZnPP, which in turn is incorporated into hemoglobin. It has been found that ZnPP containing hemoglobin has a much lower oxygen capacity than Fe-containing hemoglobin (68). Moreover, iron in the form of ferritin and ferruginous micelles also accumulate in the mitochondria of bone marrow reticulocytes (68) (Fig. 1). In general, Pb induces two types of anemia. While acute, a high Pb level results in hemolytic anemia; chronic Pb exposures result in anemia that is induced by the interference of Pb with heme biosynthesis and also by diminished red blood cell survival. According to the EPA, in the case of occupationally exposed adults, the threshold BLL for a decrease in hemoglobin is 50 µg/dL, whereas in children, the threshold BLL has been observed to be 40 µg/dL (7).

FUTURE DIRECTION

Lead is ubiquitous in nature, but anthropogenic activities have further exacerbated Pb levels in the environment. Although lead-based petroleum has been phased out, Pb is still being used in the industry (such as lead-acid battery manufacturing industry) due to the absence of viable alternatives. Blood lead levels of 10 µg/dL have been associated with adverse health effects in children (86). The treatment of Pb poisoning involves chelation, which accelerates the process of reducing the Pb levels in the circulating blood. In children, chelation therapy is recommended only when the BLLs exceed 45 µg/dL (86). There is an abundance of evidence demonstrating that dietary calcium decreases gastrointestinal Pb absorption,

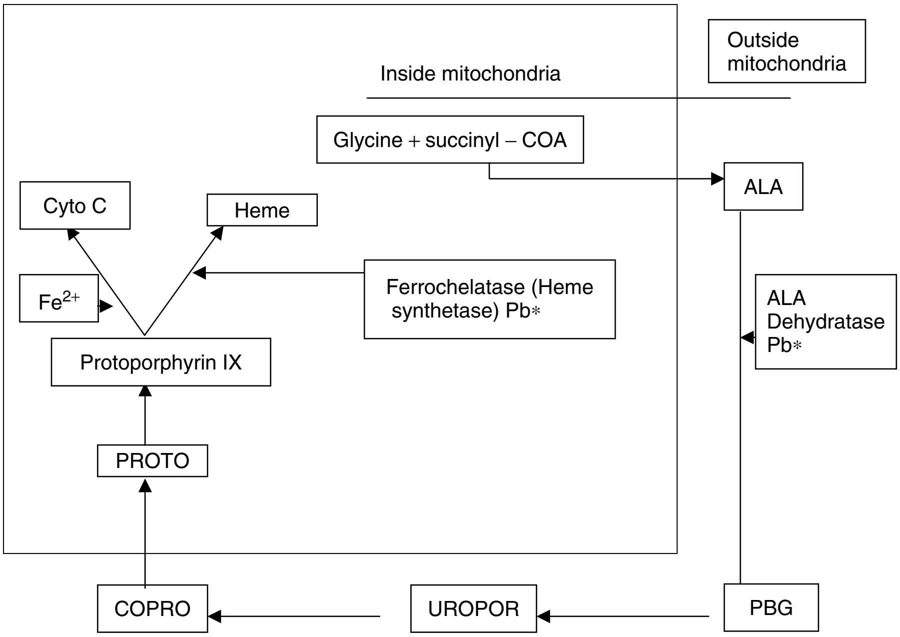

Figure 1. Scheme of heme synthesis showing sites where lead has an effect. PGB, phorbobilinogen; UROPOR, uroporphyrinogen; COPRO, coproporphrinogen; PROTO, protophorphyrinogen; Co A, coenzyme A; ALA, δ—aminolevulinic acid; Pb* site for Pb effect (from Ref. 68).

thereby reducing Pb toxicity (87). Experiments carried out on rats suggest that bone Pb concentrations increased by about fourfold in rats fed with a low calcium diet, compared to rats on a normal calcium diet, although the amounts of Pb ingested were equal (88). Short-term intake of calcium does not alter BLLs, but reports suggest that under conditions of physiological stress, skeletal minerals can be mobilized for calcium and that Pb will also be released along with calcium during this mineral mobilization (72,89,90). Ongoing research is now addressing whether calcium intake will reduce this mobilization.

Epidemiological studies have shown correlations between low levels of Pb contamination in the environment and subtle neurobehavioral effects in children. However, the biochemical mechanisms forming the basis of these subtle neurotoxic effects of low concentrations of Pb in the environment have not been clearly established (5). This is another area where research needs to be focused, to find out the biochemical interactions that may be primarily responsible for the neurotoxic effects of Pb.

BIBLIOGRAPHY

1. Pinheiro, J.P., Mota, A.M., and Benedetti, M.F. (1999). Lead and calcium binding to fulvic acids: salt effect and competition. *Environ. Sci. Technol.* **33**: 3398–3404.
2. Callender, E. and Rice, K.C. (2000). The urban environmental gradient: anthropogenic influences on the spatial and temporal distribution of lead and zinc in sediments. *Environ. Sci. Technol.* **34**(2): 232–238.
3. Markus, J. and McBratney, A.B. (2001). A review of the contamination of soils with lead II. Spatial distribution and risk assessment of soil lead. *Environ. Int.* **27**: 399–411.
4. Thorton, I., Rautia, R., and Brush, S. (2001). Lead the facts. Prepared by IC Consultants Ltd., London, UK. Available at http://www.ldaint.org/factbook/.
5. Mahaffey, K.R., McKinney, J., and Reigart, J.R. (2000). Lead and compounds. In: *Environmental Toxicants: Human Exposures and Their Health Effects*. M. Lippmann (Ed.). John Wiley & Sons, Hoboken, NJ, pp. 481–553.
6. Lansdown, R. and Yule, W. (1986). *Lead Toxicity, History of Environmental Impacts*, The John Hopkins University Press, Baltimore.
7. Agency for Toxic Substances and Disease Registry (ATSDR) (1999). *Toxicological Profile for Lead*. U.S. Department of Health and Human Services, Atlanta.
8. Alloway, B.J. and Ayres, D.C. (1997). *Chemical Principles of Environmental Pollution*. Blackie Academic and Professional, London.
9. Roorda, A.A.H. and Van Der Ven, B.L. (1999). *Lead sheet and the environment*, TNO Institute of Environmental Sciences, Energy Research and Process Innovation, Amsterdam, The Netherlands, Report no. TNOMEP-R98/503.
10. National Science Foundation (NSF) (1977). *Lead in the Environment*. W.R. Bogess (Ed.). NSF/RA-770214. NSF, Washington, DC.
11. Schmitt, C.S., Finger, S.E., May, T.W., and Kaiser, M.S. (1987). Bioavailability of lead and cadmium from mine tailings to the pocket book mussel (*Lampilis ventriosa*) In: *Proceedings of the Die-offs of Freshwater Mussel in the United States*. R. Neves (Ed.). U.S. Fish and Wildlife Service and the Upper Mississippi River Conservation Committee, pp. 115–122.
12. Erten-Unal, M., Wixson, B.G., Gale, N., and Pitt, J.L. (1998). Evaluation of toxicity, bioavailability and speciation of lead, zinc and cadmium in mine/mill wastewaters. *Chem. Speciation Bioavail.* **10**(2): 37–46.
13. LaVelle, J.M. et al. (1991). Bioavailability of lead in mining wastes: an oral incubation study of young swine. *Chem. Spec. Bioavail.* **3**: 105–111.
14. Davis, A., Ruby, M.V., and Bergstrom, P.D. (1992). Bioavailability of arsenic and lead in soils from the Butte, Montana, mining district. *Environ. Sci. Technol.* **26**: 461–468.
15. Gulson, B.L. et al. (1994). Lead bioavailability in the environment of children; blood lead levels in children can be elevated in a mining community. *Arch. Environ. Health* **49**: 326–331.
16. Fleming, S. (1994). *Scientific Criteria Document for Multimedia Environmental Standards Development—Lead*. PIBS 2832, Ontario Ministry of Environment and Energy, March.
17. Kay, S.H., Haller, W.T., and Garrard, L.A. (1984). Effects of heavy metals on water hyacinths (*Eichhornia crassipes* (Mart.) Solms). *Aquat. Toxicol.* **5**: 117–128.
18. Peterson, S., Kim, R., and Moy, C. (1993). *Ecological risks of lead contamination at a gun club: waterfowl exposure via multiple dietary pathways*. Prepared for Society of Environmental Toxicology and Chemistry, by Ecology and Environment Inc., San Francisco.
19. Van Der Werff, M. and Pruyt, M.J. (1982). Long-term effects of heavy metals on aquatic plants. *Chemosphere* **11**: 727–739.
20. Rabitto, I.S. et al. (2004). Effects of dietary Pb(II) and tributyltin on neotropical fish *Hoplias malabaricus*: histopathological and biochemical findings. *Ecotoxicol. Environ. Safety* **60**: 147–156.
21. Davis, P.H., Goettl, J.P., Sinley, J.R., and Smith, N.F. (1976). Acute and chronic toxicity of lead to rainbow trout *Salmo gairdneri*, in hard and soft water. *Water Res.* **10**: 199–206.
22. Maddock, B.G. and Taylor, D. (1980). The acute toxicity and bioaccumulation of some lead alkyl compounds in marine animals. In: *Lead in the Marine Environment*. M. Branica and Z. Konrad (Eds.). Pergamon Press, Oxford, pp. 233–261.
23. Singh, R.P. et al. (1997). Response of higher plants to lead contaminated environment. *Chemosphere* **34** (11): 2467–2493.
24. Schuck, E.A. and Locke, J.K. (1970). Relationship of automotive lead particulates to certain consumer crops. *Environ. Sci. Technol.* **4**: 324–330.
25. Zimdahl, R.L. and Faster, J.M. (1976). The influence of applied phosphorus, manure or lime or uptake of lead from soil. *J. Environ. Qual.* **5**: 31–34.
26. Krishnayappa, N.S.R. and Bedi, S.J. (1986). Effect of automobile lead pollution of *Cassia tora* and *Cassia occidentalis* l. *Environ. Pollut.* **40**: 221–226.
27. Kumar, G., Singh, R.P., and Sushila, X. (1993). Nitrate assimilation and biomass production in *Sesamum indicum* L. Seedlings in a lead enriched environment. *Water, Air Soil Pollut.* **66**: 163–171.
28. Bharti, M.B. and Singh, R.P. (1993). Growth and nitrate reduction by *Sesamum indicum* cv PB-I respond differentially to lead. *Phytochemistry* **33**: 531–534.
29. Venkataraman, S., Veeranianeyulu, K., and Ramdas, V.S. (1978). Heavy metal inhibition of nitrate reductase. *Ind. J. Exp. Biol.* **16**: 615–616.

30. Jones, R., Inouye, L., Bednar, A., and Boyd, R. (2004). Lead toxicity and subcellular compartmentalization in earthworms exposed to contaminated soil. The Society of Environmental Toxicology and Chemistry, 25th Annual Meeting, North America. Available at http://abstracts.co.allenpress.com/pweb/setac2004/document/?ID=42021.
31. Doelman, P., Nieboer, G., Scrooten, J., and Visser, M. (1984). Antagonistic and synergistic toxic effects of Pb and Cd in a simple food chain: nematodes feeding on bacteria or fungi. *Bull. Environ. Contam. Toxicol.* **32**: 717–723.
32. Beyer, W.N. and Anderson, A. (1985). Toxicity to woodlice of zinc and lead oxides added to soil litter. *Ambio* **14**: 173–174.
33. Sanderson, G.C. and Bellrose, F.C. (1986). *A Review of the Problem of Lead Poisoning in Waterfowl*. Illinois Natural History Survey, Champaign, Illinois, Special Publication 4. Northern Prairie Wildlife Research Center Jamestown, ND. Available at http://www.npwrc.usgs.gov/resource/othrdata/pbpoison/pbpoison.htm.
34. Pain, D.J. (1990). Lead shot ingestion by water birds in the Camargue, France: an investigation of levels and interspecific differences. *Environ. Pollut.* **66**: 273–285.
35. United States Fish and Wildlife Service (USFWS) (1999). Establishing "lead free fishing areas" and the prohibition of the use of certain fishing sinkers and jigs made with lead on specific units of the National Wildlife Refuge System. *Fed. Reg.* **64**(70).
36. Lead poisoning report (2000). 3rd International Update Report on Lead Poisoning in water birds. Available at http://www.wetlands.org/pubs&/Lead_P_Report.htm.
37. Burger, J. (1995). A risk assessment for lead in birds. *Toxicol. Environ. Health* **45**(4): 369–396.
38. Boline, D.R. (1981). Some speciation and mechanistic aspects of trace metals in biological systems. In: *Environmental Health Chemistry*. J.D. McKinney (Ed.). Science, Ann Arbor.
39. Kelada, S.N., Shelton, E., Kaufmann, R.B., and Khoury, M.J. (2001). δ-Aminolevulenic acid dehydratase genotype and lead toxicity: a huge review. *Am. J. Epidemiol.* **154**(1): 1–13.
40. Agency for Toxic Substances and Disease Registry (ATSDR) (2003). Case studies in environmental medicine. Lead toxicity. Available at http://www.atsdr.cdc.gov/HEC/CSEM/lead/physiologic_effects.html#Pregnancy%20Outcomes.
41. Lin–Fu, J.S. (1973). Vulnerability of children to lead exposure and toxicity: part one. *N. Engl. J. Med.* **289**: 1229–1233.
42. Ziegler, E.E. et al. (1978). Absorption and retention of lead by infants. *Pediatric Res.* **12**: 29–34.
43. Chamberlain, A.C. (1983). Effect of airborne lead on blood lead. *Atmos. Environ.* **17**: 677–692.
44. Simons, T.J.B. (1986). Passive transport and binding of lead by human red blood cells. *J. Physiol.* **378**: 267–286.
45. Barry, P.S. and Mossman, D.B. (1970). Lead concentrations in human tissues. *Br. J. Ind. Med.* **27**: 339–351.
46. Schroeder, H.A. and Tipton, I.H. (1968). The human body burden of lead. *Arch. Environ. Health* **17**: 965–978.
47. Finkelstein, Y., Markowitz, M.E., Rosen, J.F. (1998). Low-level lead-induced neurotoxicity in children: an update on central nervous system effects. *Brain Res. Rev.* **27**: 168–176.
48. Bressler, J.P. and Goldstein, G.W. (1991). Mechanisms of lead neurotoxicity. *Biochem. Pharmacol.* **41** (4): 479–484.
49. Damstra, T. (1977). Toxicological properties of lead. *Environ. Health Perspect.* **19**: 297–307.
50. Blackman, S.A. (1937). The lesions of lead encephalitis in children. *Bull. Johns Hopkins Hosp.* **61**: 1–61.
51. Needleman, H.L., Schell, A., Bellinger, D., Leviton, A., and Allred, E.N. (1990). The long-term effects of exposure to low doses of lead in childhood. *N. Engl. J. Med.* **322**: 83–88.
52. Bradbury, M.W.B. (1984). The structure and function of blood–brain barrier. *Fed. Proc.* **43**: 186–190.
53. Clasen, R.A. et al. (1973). Electron microscopic and chemical studies of the vascular changes and edema of lead encephalopathy. *Am. J. Pathol.* **74**: 215–240.
54. Goldstein, G.W., Asbury, A.K., and Diamond, I. (1974). Pathogenesis of lead encephalopathy. Uptake of lead and reaction of brain capillaries. *Arch. Neurol.* **31**: 382–389.
55. Yule, X., Lansdown, R., Miller, I., and Urbanowicz, M. (1981). The relationship between blood lead concentrations, intelligence, and attainment in a school population: a pilot study. *Dev. Med. Child Neurol.* **23**: 567–576.
56. Schroeder, Sr., Hawk, B., Otto, D.A., et al. (1985). Separating the effects of lead and social factors on IQ. *Environ. Res.* **38**: 144–154.
57. Fulton M., Raab, G., Thomson, G., et al. (1987). Influence of blood lead on the ability and attainment of children in Edinburgh. *Lancet* **1**: 1221–1226.
58. Lansdown R., Yule, W., Urbanowicz, M.A., and Hunter, J. (1986). The relationship between blood–lead concentrations, intelligence, attainment and behavior in a school population: the second London study. *Int. Arch. Occup. Environ. Health* **57**: 225–235.
59. Hawk, B.A., Schroeder, S.R., Robinson, G., et al. (1986). Relation of lead and social factors to IQ of low SES children: a partial replication. *Am. J. Ment. Defic.* **91**: 178–183.
60. Winneke, G., Brockhaus, A., Ewers, U., et al. (1990). Results from the European multicenter study on lead neurotoxicity in children: implications for risk assessment. *Neurotoxicol. Teratol.* **12**: 553–559.
61. Kehoe, R.A. (1961). The metabolism of lead in man in health and disease: present hygienic problems relating to the absorption of lead: the Harben 1960. *J. Inst. Public Health. Hyg.* **24**: 177–203.
62. Seto, D. and Freeman, J.M. (1964). Lead neuropathy in childhood. *Am. J. Dis. Child.* **107**: 337–342.
63. Xu, Bo., Chia, S.E., Tsakok, M., and Ong, C. (1993). Trace elements in blood and seminal plasma and their relationship to sperm quality. *Reprod. Toxicol.* **7**: 613–618.
64. Alexander B.H. et al. (1996). Paternal occupational lead exposure and pregnancy outcome. *Int. J. Occup. Environ. Health* **2** (4): 280–285.
65. Lancranjan, I. et al. (1985). Reproductive ability of workmen occupationally exposed to lead. *Arch. Environ. Health* **30**: 396–401.
66. Lerda, D. (1992). Study of sperm characteristics in persons occupationally exposed to lead. *Am. J. Ind. Med.* **22**: 567–571.
67. Alexander, B.H. et al. (1996). Semen quality of men employed at a lead smelter. *Occup. Environ. Med.* **53**: 411–416.
68. Chisolm, J.J., Jr. and O'Hara, D.M. (1982). *Pb Absorption in Children: Management, Clinical and Environmental Aspects*. Urban & Schwarzzenberg, Baltimore.
69. McMichael, A.J. et al. (1986). The Port Pirie Cohort study: maternal blood lead and pregnancy outcome. *J. Epidemiol. Community Health* **40**: 18–25.
70. Barltrop, D. (1969). Transfer of lead to the human foetus. In: *Mineral Metabolism in Paediatrics*. D. Barltrop and W.L. Burland (Eds.). Blackwell, Oxford.

71. Gershanik, J.J., Brooks, G.G., and Litle, J.A. (1974). Blood lead values in pregnant women and their offspring. *Am. J. Obstet. Gynecol.* **119**: 508–511.
72. Gulson, B.L. et al. (1997). Dietary lead intakes for mother–child pairs and relevance to pharmacokinetic models. *Environ. Health Perspect.* **105**: 1334–1342.
73. Sanin, L.H. et al. (2001). Effect of maternal lead burden on infant and weight gain at one month of age among breastfed infants. *Pediatrics* **107**(5): 1016–1023.
74. Nolan, C.V. and Shaikh, Z.A. (1992). Lead nephrotoxicity and associated disorders: biochemical mechanisms. *Toxicology* **73**: 127–146.
75. Cramer, K., Goyer, R.A., Jagenburg, R. and Wilson, M.H. (1974). Renal ultra structure, renal function and parameters of lead toxicity in workers with different periods of lead exposure. *Br. J. Ind. Med.* **31**: 113–119.
76. Dobryszycka, W. et al. (1984). Morphological and biochemical effects of chronic low level cadmium and lead feeding rats. *Acta Pol. Pharm.* **41**: 111.
77. Lilis, R. et al. (1980). Kidney function and lead: relationships in several occupational groups with different levels of exposure. *Am. J. Ind. Med.* **1**: 405–412.
78. Nuyts, G.D. et al. (1991). Does lead play a role in the development of chronic renal disease? *Nephrol. Dial. Transplant.* **6**: 307–315.
79. Waldegger S. et al. (1995). Heavy metal inhibition of rBAT—induced amino acid transport. *Kidney Int.* **47**(6): 1667–1681.
80. Victery W., Throler, H.A., Volpe, R., et al. (1988). Summary of discussion sessions: symposium on lead blood pressure relationships. *Environ. Health Perspect.* **78**: 139–155.
81. Schwartz, J. (1995). Lead, blood pressure, and cardiovascular disease in men. *Arch. Environ. Health* **50**: 31–37.
82. Korrick, S.A., Hunter, D.J., Rotnitzky, A., et al. (1999). Lead and hypertension in a sample of middle-aged women. *Am. J. Public Health* **89**(3): 330–335.
83. Hu, H. et al. (1996). The relationship of bone and blood lead to hypertension. The normative aging study. *JAMA* **275**: 1171–1176.
84. Warren, M.J., Cooper, J.B., Wood, S.P., and Shoolongin-Jordan, P.M. (1998). Lead poisoning, haem synthesis and 5-aminolaevulinic acid dehydratase. *TIBS* **23**: 217–221.
85. Kappas, A., Sassa, S., Galbraith, R.A., and Nordmann, Y. (1995). In: *The Metabolic Basis of Inherited Diseases*, 7th Edn. Scriver, C.R., A.L. Beaudet, W.S. Sly, and D. Valle (Eds.). McGraw Hill, New York, pp. 2103–2159.
86. Centers for Disease Control (1991). *Preventing Lead Poisoning in Young Children. A Statement by the Centers for Disease Control*, United States Department of Health and Human Service, United States Public Health Service, Atlanta.
87. Bruening, K. et al. (1999). Dietary calcium intake of urban children at risk of lead poisoning. *Env. Health Perspect.* **107**: 431–435.
88. Mahaffey, K.R. (1995). Environmental health issues—commentary on nutrition and lead: strategies for public health. *Environ. Health Perspect.* **103**(S6): 191.
89. Gulson, B.L. et al. (1997). Pregnancy increases mobilization of lead from maternal skeleton. *J. Lab. Clin. Med.* **130**: 51–62.
90. Gulson, B.L. et al. (1998). Mobilization of lead from the skeleton during the postnatal period is even larger than during pregnancy. *J. Lab. Clin. Med.* **131**: 324–329.

MICROBIAL DETECTION OF VARIOUS POLLUTANTS AS AN EARLY WARNING SYSTEM FOR MONITORING OF WATER QUALITY AND ECOLOGICAL INTEGRITY OF NATURAL RESOURCES, IN RUSSIA

GALINA DIMITRIEVA-MOATS
University of Idaho
Moscow, Idaho

GENERAL INTRODUCTION

General

The U.S. Environmental Protection Agency (U.S. EPA) persistently supports programs that monitor water resources using bioassessments and that develop biocriteria for different water bodies throughout the country in accordance with the Clear Water Act, the Watershed Protection Approach, and the National Water-Quality Assessment (NAWQA) Program. Currently, assessments of biological integrity and water quality are based on habitat measurements and also on biosurveys of the status of indicator aquatic communities. In order to evaluate an integrative ecological situation in water resources, the U.S. EPA has established protocols for bioassessments and for developing biocriteria applicable to selected biological assemblages, such as phytoplankton, zooplankton, periphyton, macrophytes, macroinvertebrates, and fish. As the U.S. EPA literature mentions, at the present time, no established protocols exist for other groups of organisms, including bacteria, because the data about those organisms are not sufficiently numerous. However, bioassessments based on the use of currently selected indicator groups of hydrobionts do not indicate the cause of water resource impairment. Also, usually habitat measurements monitor already known impairing factors; therefore, if a new pollutant has appeared, it may remain undetected.

The author of this paper had long-term extensive experience working with water quality control and ecological integrity monitoring using microbial indices. This practice includes eight years of microbial bioassessments and the development of biocriteria implications for the detection of various organic pollutants, a number of heavy metals, and biological contamination in seawaters of Russia, which provided impressive results, demonstrating the advantages of such an approach for monitoring ecological integrity and water quality of estuaries and shallow and pelagic regions of marine coastal zones. The composition alteration of aquatic sediment or plankton microbial assemblages subjected to pollution resulted in an increase of pollutant-resistant (indicator) groups that detected expected and unexpected pollutants and their levels. The analysis of the proportion between all indicator groups assisted us in drawing conclusions about the natural or anthropogenic process or factors causing impairment of water resources.

As monitoring of water quality using microbial indices and habitat metrics can detect many different pollutants at a number of monitoring stations, we gathered sample point data via a Global Positioning System (GPS) and displayed the distribution of each pollutant over the water body on maps employing a Geographic Information System (GIS). Several pollutants have been monitored using microbial assessments and habitat metrics, including measurements of toxic compounds at the reference stations and at the monitoring stations. Microbiological assessments for the evaluation of the biological contamination included controlling fecal and pathogenic bacteria, such as *E. coli, Salmonella, Shigella, and Listeria*, detecting phenol-resistant bacteria as indicators of fecal phenols, and also evaluation of the proportion of proteolytics and lipolytics bacteria in microbial assemblages. In order to monitor the pollution of oil and oil by-products, we studied the percent of microorganisms resistant to those pollutants as well as the concentration of those compounds. The portion of lipolytic bacteria that might contribute to oil biodegradation was also detected and evaluated. We also applied microbiological assessments and habitat measurements in order to detect the presence and the origin of phenols that can appear in the water from point and nonpoint sources. The choice of monitored heavy metal pollution using microbial assessments, and also habitat metrics, was based on the environmental role of heavy metals and their impact. As a primer pool of elements, we controlled Cu, Pb, and Ni as indicators of industrial impact; Cd and Zn as indicators of anthropogenic impact; Co and Cs as detectors of radioactive contamination; and Fe as a marker of terrestrial influence.

Historical Background

Application of Microbial Assessments for Detection of Various Pollutants in the Environment

Response of Indicator Groups of Microorganisms to Pollution in the World Ocean. The concept of microbial indication of organic pollutants in the world ocean has been developed most actively during the last two decades. Progressively increasing anthropogenic load on the marine environment creates tense ecological situations, primarily in coastal areas. As had been established earlier, an increase of anthropogenic pressure on the marine environment influences structural and functional properties of hydrobiont communities (1). Pollutants cause various changes in marine organisms. Organisms that have adapted to new chemical compounds, or increased concentrations of pollutants in the environment, begin to dominate in the structure of biocenosis. The ability to quickly adapt to changing environmental conditions because of high rates of reproduction and growth and versatile enzymatic activity is a unique feature of microorganisms. For this reason, microorganisms are unique tools that scientists have started using to control marine pollution because they have the most direct connections to the surrounding environment, they react quickly by altering population size and composition, and they reflect any environmental changes. On this basis, microbial assemblages are considered indicators of physical, chemical, and biological processes in the world ocean (2–7). As is known, microorganisms are indicators of contamination of various organic substances, because they are destroyers and consumers of those compounds at the same time. The existence and sensitivity of the individual response of the plankton community to variations in the concentrations of metals in contaminated waters, with a complex of metals, are confirmed on the basis of microbiological data and the results of their comparison with the data of chemical analysis (8–11). A number of microorganisms, the composition of main physiological groups, and distinct genera in seawaters and sediments are characteristics of microbial assemblages that researchers used for the purposes of biomonitoring of pollutants in the marine environment (12–15). Biological contamination criteria were developed earlier, historically for fresh waters more than for seawaters. Microbial assessments of other kinds of pollution have longer records in the marine environment, particularly, in Russia.

Detection of Pathogenic Bacteria and Biological Contamination in Surface Waters. Indication of pathogenic microorganisms in surface waters is under the control of environmental services and laboratories. Observed indicator bacteria of biological contamination exist, as well as saprobe indices of waters in freshwater and seawater environments. Freshwater saprobe indices are established on the basis of the ratio between the number of bacteria grown on reach nutrient agar and the number of bacteria grown on pour agar at starving conditions. For barely contaminated sites, the ratio is 1.2–2.8, and for references, the ratio appears as 0.1–0.7 (16). Also, most of the microbial indicators of biological contamination have been studied in fresh waters. In the 1960s, seawaters in Russia were investigated and classified according to their saprobes (17). In the 1970s, indicator microorganisms for biological contamination were determined, and sanitary criteria were established according to international agreement (18,19). According to that classification, in secure waters, the number of *Escherichia coli* cells must not exceed 10 per mL. In 1 mL of dangerous waters that can cause infectious diseases, more than 24 cells were found. The control of biological contamination is also a very highly developed field at the present time.

As seawaters accept an increasing amount of domestic and industrial wastewaters, the biological contamination of seawaters is a subject of the continuing investigation. Distribution of fecal and pathogenic bacteria has been studied in the waters of the Atlantic Ocean (20) and in the benthos coastal zone of the Antarctic Ocean (21). Other regions of the world ocean have been studied in estuarial (22), coastal (13), and pelagic (23) zones. Among the indicators of fecal contamination indicated are such microorganisms as *E. coli, Nitrobacteria* spp., *Clostridium perfringens, Pseudomonas aeruginosa, Vibrio* spp., *Salmonella* spp., *Candida albicans, Klevsiella* spp., *Bacterioides* spp., enterococci, bifid-bacteria, and fecal streptococci (22). Intestinal viruses are also the subject of study in the water environments.

Use of Microorganisms as Indicators of Oil Pollution in Seawaters. Microorganisms were beginning to be used many years ago as detectors of oil pollution and of water quality in the marine environment. One of the first marine microbiologists studying this subject was ZoBell and his coauthors (24). These scientists have found more than 100 species of marine microbes representing 30 genera, which have been able to use hydrocarbons. Those microorganisms have been considered indicators of oil pollution in seawaters. By the present time, researchers have isolated more than 1000 strains of microorganism-oxidizing hydrocarbons (25–28). The most active hydrocarbon-oxidizing marine bacteria were determined to be *Mycobacterium* and *Arthobacter* (29).

In Russia, the first publications devoted to indication of oil contamination in seawaters were published in the 1950s (30,31). The evaluation criteria for oil pollution were based on the ratio of oil-oxidizing heterotrophic bacteria and saprotrophic bacteria or on the ratio of oil-oxidizing bacteria and the total bacterial quantity (29,32). Bacterial criteria for determining acceptable standards for clean water were also established. More recently, data appeared about microbial degradation of not only aliphatic hydrocarbons, but also of polycyclic hydrocarbons (33–35). Their results have demonstrated that indicator groups of microorganisms are developed weakly in noncontaminated pelagic seawaters. In oil-contaminated areas of the world ocean, those groups were very numerous. For this reason, researchers use the activity and a number of hydrocarbon-oxidizing microorganisms as an index of oil pollution in seawaters. Now this is the most heavily investigated field in the study of marine microbiology in Russia and in the world at large.

Microbial Detection of Phenols in Seawaters. Contamination of surface waters by phenols can have different origins. Fecal sterols, polyaromatic hydrocarbons, biphenyls, and Cl-phenols enter water environments from domestic wastewaters, from industrial oil pollution, from oil spills, and from cellulose factories. Contrary to microbial indication of oil pollution in seawaters, indication of phenols does not have long-term records. Over more than twenty years, some articles devoted to microbial detection of phenols in seawaters were published (7,36,37). The microbial detection of fecal sterols (38–40), Cl-phenols, and biphenyls has been studied in more detail (1,15). Microorganisms and degrading aromatic phenols are considered as indicators of phenol pollution in deep-water regions of the ocean (7,36,41). Remarkably, the composition of microbial assemblages from phenol-contaminated seawaters and marine sediments has been found to be close to the composition of the microbial community from active silts of wastewater treatment facilities, which were adapted to phenol by-products (42). Representatives from genera *Bacillus* (25%), *Pseudomonas* (12.5%), *Mycobacterium* (24%), *Micrococcus* (25%), and *Actinomyces* (12.5%) were found there. However, no indices were established to confirm the phenol contamination in the marine environment.

Bacteria-Degrading Lipids, Proteins, and Starch Are Additional Indicators of Organic Contamination in Coastal Waters. Seawater and marine neuston organisms contain considerable amounts of various lipids. The concentration of lipids in seawater varies from 0.01 to 0.12 $\mathrm{mg \cdot mL^{-1}}$, and concentration of lipids in marine plankton appears as 2–25%. Also, three gycerides, fatty acids, and phospholipids are oil by-products during the oil transformation process in the marine environment and bacteria synthesize them as well. As lipids are much lighter than water, they are concentrated on the margin of the surface of seawater and interfere with the gas and temperature exchange process between the ocean and the atmosphere, causing different negative consequences. Many marine bacteria can use lipids as a source of carbon. They possess lipolytic activity and provide self-clearance of lipids from seawaters. Bacterial neuston is particularly important for the biodegradation of lipids in the sea (43). For this reason, some authors suggest considering the appearance of lipolytic activity in the dominant part of neuston microbial assemblages, as confirmation of the presence of oil contamination in seawater.

Determination of a number of lipolytic microflora and measurement of lipolytic activity in marine bacteria serves as a test, characterizing the capability of the marine coastal environment for self-clearance (43,44). Nitcowski et al. (12) presented data in evaluating a portion of the physiological groups of lipolytic microorganisms in aquatic microbial association, as well as a portion of proteolytics, as an indicator of biological contamination. The authors have found that a number of bacteria from those groups were four times as numerous at contaminated sites. Among those physiological groups of microbes, 80% of the population has been identified as pseudomonades and vibrions.

Some marine bacteria, yeasts, and fungi are capable of degrading starch. Such microorganisms were detected among microbial assemblages living on the surface of marine algae and plants (45–47), their remains (44,46), and in the water (45). In our opinion, such microorganisms may also indicate contamination of seawaters by carbohydrates.

Application of Microbial Biosurveys for Detection of Heavy Metal Contamination in the Environment

Practical use of the microbial approach for detection of the presence of metals started in the 1950s in geological survey practice for the metal ore search in Russia. The method was based on the ability of certain bacteria to stimulate or to inhibit their growth in response to the presence of metals of interest in the ground and in the surface waters (48,49). In order to prove that the method of microbial detection can be implicated for metal deposit searches, scientists used the station grid approach, where habitat metrics included measurements of the concentrations of metals. Such field investigations covered extensive areas in different states and republics of the former Soviet Union in Eastern Siberia, Kazakhstan, Ural, and the Caucasus regions. Researchers have found a direct connection between the distribution of metals (Mo, Zn, Cu) and a number of cultured, metal-resistant, or metal-sensitive bacteria. All other similar applications in geological microbiology were reviewed later (50).

In the 1970s, the geochemical ecology of microorganisms was established as a science in Russia. The subject of geochemical ecology of microorganisms is the study of interactions between natural and industrial factors, including metals, in the environment (51,52). At the present time, environmental microbiology studies mechanisms of the interaction between microorganisms and heavy metals. Microorganisms absorb metals for their physiological needs, as do many other organisms, because metals are an important part of biological molecules, including enzymes, hormones, vitamins, pigments, and lipids.

Generated data demonstrate that microorganisms also have a unique system of cell organization and function, responsible for selective binding and absorption of metals by bacteria. The major element of this system is the cell wall. It is remarkable that the adsorptive capability of bacterial cell walls is so high and that it is comparable with the most efficient resins used for ion-exchange chromatography (53). As now established, the selective adsorption of distinct metals by cell walls varies significantly in different bacteria (52–54). In firmicutes, an affinity of cell walls in binding different metals is presented as follows: $La^{3+} > Cd^{2+} > Sr^{2+} > Ca^{2+} > Mg^{2+} > K^{+} > Na^{+} > Li^{+}$ (53). As a result of such high capability of the selective binding of metals, bacteria accumulate metals that greatly exceed the natural concentrations of metals in the environment.

The absorption process involves active and passive transport of metals through the plasma membrane. Specificity of active transport mechanisms is based on ligands selective to individual metals. Those minor membrane protein structures are a second element of the system responsible for selective binding and absorption of metals by bacteria. For example, it has been demonstrated that *E. coli* has three specific systems for transport of iron (55).

Bacteria have a high resistance to heavy metals because of different mechanisms of their detoxification. For this reason, microorganisms are the only live organisms on the Earth that are able to live in environments containing very high concentrations of heavy metals. The detoxification of heavy metals is the system that provides bacterial survival in heavy metal environments. In general, those mechanisms include: limitation or complete blocking of metal internalization into bacterial cells, active removal of absorbed metals from cells, and detoxification of metal ions by binding and transforming them into nontoxic forms or by the resulting intracellular accumulation (56,57). Thus, bacteria possess various mechanisms of self-protection against toxic effects of heavy metals in the environment.

On the basis of presented data, we may conclude that the variety of systems of selective binding and transport of heavy metals, and also the system of detoxification of heavy metals, are real cell functions that provide survival and the specific response of the microbial community to the presence of toxic metals in the aquatic and terrestrial environments. Of course, passive transport and nonspecific transport of metals also take place in microbial cells. Other mechanisms of interaction of bacteria with metals exist as well, but it is not an aim of this entry to discuss all aspects of the interaction of bacteria with metals, including details of detoxification of metals by microbial cells. Those aspects are presented in other papers (56–61). Also, many environmental factors influence the efficiency of the interaction of metal ions with microbial cells, including the pH of various environments, the state of metal ions, and the presence of other ions (56). The information presented above aims to demonstrate a general knowledge that may help in the understanding of the true foundation for the prospective use of microbial assessments for monitoring heavy metals in water environments. However, this powerful tool for environmental monitoring is not yet widely used. Also, since recent times, the technique of microbial survey, employed to indicate different pollutants, their concentrations, and also the ecological state of marine or freshwater environments, was not sufficiently used in Russia, or in the remainder of the world for that matter.

Monitoring Marine Environmental Quality in Russia

The anthropogenic load and its influence on the condition of the world ocean, particularly in the marine coastal zones, continues to grow as a result of population growth and increasing human activity. Such pressure influences qualitative development of sea life and the structural and functional characteristics of hydrobiont assemblages in the ecological, morphological, physiological, biochemical, and genetic levels (1). One of the most important consequences of the pollution of the marine environment is an appearance of indicator forms that can possess harmful characteristics, which is one of the reasons why the methods for biological control of marine environmental quality were developed in the 1960s. Two major different strategic approaches have been applied for biomonitoring of pollution in the marine environment in Russia in the past. These approaches are bioassay and bioindication. Microbial detection of marine pollutants combines the advantages of both methods and gives new possibilities for efficient tracking of pollutants issued from point and nonpoint sources in surface waters.

Bioassay Approach. The bioassay approach is based on the ability of marine organisms to react through their longevity, by the reduction of some life functions, and by alteration of cell structure at the appearance and concentration of toxic compounds in the seawater. These variations are immediately observable and provide information about the degree of toxicity of the environmental samples containing the complex of pollutants for certain kinds of organisms. However, this method does not provide information about the concentration of pollutants or an indication of which pollutant causes this negative effect. The most suitable subjects used for such control are microalgae, larvae of marine invertebrates, young generations of fishes (62,63), sea urchins and molluscs, and also sexual cells of marine organisms and marine polychaete (64). Usually, those subjects are single cell or simply organized forms, which are very sensitive to unfavorable conditions and have short lifespans.

Bioindication Approach. Another technique widely used for the marine environmental quality control is bioindication, which pertains to the study of mature, long-term life forms and requires that organism monitors must be immobile or attached to the marine substrate, and that they must be resistant to pollution and also able to accumulate toxic compounds at a concentration exceeding environmentally acceptable levels to several significant digits. The organisms most often used for bioindication in the marine environment are macroalgae and molluscs (65). These organisms are used particularly to monitor heavy metals in seawaters. In such cases, chemical analysis of the organism's tissues displays every distinct pollutant and establishes the dynamic of its behavior in the marine environment over time. The disadvantage of this approach is that no information about how toxic a certain compound may be and what the concentration of that chemical was during a certain period may be determined.

Use of a Combination of the Bioassay and Bioindication Methods in Russia. The use of both the bioassay and bioindication methods in Russia is mostly applicable to scientific research at the present time. Currently, no state regulation of these methods exists in environmental policy. However, because such approaches are numerous and multidisciplinary, Russian Scientific Institutes always are afforded opportunities to submit results of expertise and to make strong recommendations to any federal government or local state and city organizations and agencies to influence their economical or political solutions. Such scientific investigations are always a subject of consideration for the Environmental Protection Committees and Ministry, and also for the Health Protection Ministry and the local Sanitary Committees of Russia.

It is fortunate to note that most of the time their final decisions are taken with consideration of the recommendations given by Scientific Organizations monitoring the marine environmental quality and condition of marine habitats. For example, several scientific marine institutions conducted long-term and multidisciplinary investigations of the ecological state of the marine environment in the important international economical zone near Vladivostok, Russia during 1996–1998. Named The Tumen River Economic Development Area Project (TREDA), it was planned by the United Nations Organization to help economical and political development of that Pacific Region, involving China, North Korea, Japan, and Russia. The first initiative was considered the only economic side of that project. No detailed information existed about the current complex ecological situation of the area and how the project application could change the surrounding marine environment. The existence of the unique Marine National Reservation in Russia was not considered at all. Subsequently, scientific investigations showed that nobody controlled the flow of wastewater into the Tumen River from China and North Korea; consequently, the pollution at this site has increased tremendously during recent years. The increased flow of domestic, agricultural, and industrial wastewaters into the Sea of Japan caused the appearance of many unhealthy forms among the marine habitats in the Marine State Reservation area. The expertise conducted made responsible persons and organizations plan additional efforts and funding to protect not only the Russian National Marine Reservation but also the marine environment of the Sea of Japan. I am proud to announce that microbial detection results of those dramatic changes were among the first data obtained through these efforts.

MICROBIAL BIOSURVEYS IN FAR EASTERN SEAS OF RUSSIA

Areas of Field Investigations

We began our investigation in this field 12 years ago. The areas that have been investigated in the Russian Pacific region have all involved international, economically, and politically important areas (Fig. 1), such as the Southeast Pacific Ocean coast of the Kamchatka peninsula (1995, 1997, 1999), the southwest (1995, 1997) and the northeast (1996, 1998) of Sakhalin island, and the North (1992, 1996–2000) and South parts of Primorie (1995–2001) (8–11,37,47,66–68). As a matter of background, the Far Eastern seas are influenced by global natural factors such as the Pacific ore "belt," volcanic activity, grand streams, upwelling, and by the action of bringing chemical elements to the surface water in active geochemical zones. As a result, the Far Eastern seas are distinguished as containing the highest level of toxic material, primarily heavy metals, which can provoke ecological stress and cause pollution.

Every water body chosen for our investigation had specific characteristics and was distinguished in the degree of anthropogenic impact, geochemical, geographic, and ecological conditions. In the past, Avachinskaya Guba Bay, Krasheninnikov Bight (Kamchatka peninsula) was known as the biggest Russian Navy base, and because of this status, the area became contaminated with radionuclides. Seawaters surrounding Sakhalin Island are

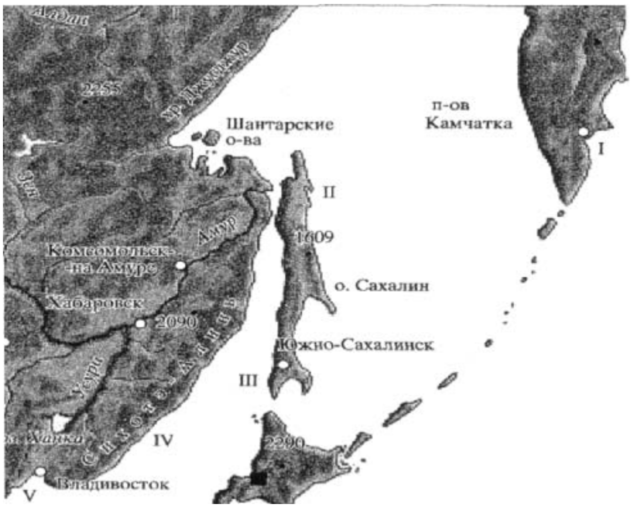

Figure 1. Areas of the studies on microbial indication of pollution of the near-shore waters by heavy metals: I—Kamchatka coast, area of Avachinskaya Guba; II—northeastern coast of Sakhalin Island, Nyiskii, Nabil', and Chaivo Bays; III—southwestern coast of Sakhalin Island, off Kholmsk; IV—Northern Primorie, Rudnaya Bight, and Lidovka Bights; V—Peter the Great Bay, the area of the mouth of the Tumen River Amur and Ussuri Bays, and Nakhodka Bight.

the most biologically productive in the Russian Far East and are used as one of the major international fishery regions. However, this region is affected by a strong anthropogenic impact. The waters near the Kholmsk area (South-West Sakhalin) is under the influence of the Tsushima Current, which carries oil, phenol, and heavy metal pollution from the coast of Japan. The shallow waters of bays of Northeast Sakhalin (Nyiski, Chaivo, and Nabil Bays) undergo oil extraction.

The Russian coast of the Sea of Japan (the Primorie region) is also an important local fishery and an international transportation and trade zone. Rudnaya Bay (Northern Primorie) lies within the region of production and processing of polymetal ores rich in lead, zinc, and also containing a great admixture of copper, cadmium, or rare-earth metals. Traditionally, that aquatic area has been investigated because of its particular characteristic of extensive heavy metal contamination. Kit Bay (Northern Primorie), one of the cleanest parts of the Primorie coast, was used as a reference site for the investigation.

Peter the Great Bay (Southern Primorie) presented itself as a rare type of habitat for aquatic life possessing unique geographic conditions, where the cold water of the Primorie stream mixes with the warm waters of the South Eastern Korean stream (a branch of the Tsushima Current) (69). The bay includes a highly productive shelf with a unique cenosis of aquatics. This area is composed of the only State Marine Reservation of Russia, the mouth of the Tumen River, Amur Bay, Ussuri Bay, Nakhodka Bight, and Golden Horn Bight (1995–2001). At the northwest end of Peter the Great Bay lie Golden Horn Bight, Amur Bay, Ussury Bay, and Nakhodka Bight, which are affected by the severe influence of domestic and industrial wastewaters from Vladivostok, Nakhodka, and other settlements. The ecological situation in the southwest part of Peter the Great Bay, during the longer warm period of the year, is determined by the influence of Tumen River running through industrial zones of China and North Korea. The State Marine Reservation is severely impacted by the pollution carried by that big river.

Parameters Analyzed

In selected aquatoria, a seasonal and monthly microbiological and chemical monitoring of the water quality was conducted. The following parameters were analyzed: (1) the number of colony-forming heterotrophic microorganisms and the proportion of metal-resistant groups of bacteria in the assemblages of colony-forming heterotrophic microorganisms in the near-surface and near-bottom water layers; (2) long-term variations in the number of metal-resistant colony-forming hetertrophic microorganisms; (3) seasonal dynamics in the number of metal-resistant colony-forming hetertrophic microorganisms; and (4) morphological, physiological, biochemichal, and molecular genetic peculiarities of metal-resistant heterotrophic microorganisms. All listed parameters have been investigated at reference sites and at impact sites.

Microbiological monitoring and chemical analyses were performed simultaneously on water samples from areas that had been subjected to a pronounced anthropogenic impact. Special selective media were developed in laboratory experiments and corrected during field observations for the selection of heterotrophic microorganisms that were highly resistant to contaminants. The composition of the pollutant-resistant groups of marine microbial communities was determined under conditions differing in the type of the anthropogenic impact and geochemical features.

Monitoring of Heavy Metal Contamination in Seawaters Using Microbial Biosurveys

The composition of microbial assemblages and also of physiological, biochemical, morphological, and molecular genetic characteristics of bacterial representatives were distinct in the contaminated and reference sites. However, because coastal seawaters usually contained a complex of pollutants, it was difficult to conclude which compound or factor caused that alteration. Only a number of bacteria in the planktonic microbial communities were resistant to each pollutant and seemed to be a clear indicator to characterize the presence of the distinct polluting compound and the degree of contamination. On the basis of the comparisons between the results from the synchronously conducted microbiological and chemical monitoring of the pollution of the near-shore seawaters, a positive correlation was established between the microbial indices of the abundance of resistant forms and the concentrations of each of the metal-treated and organic compounds. The microbial monitoring of heavy metal pollution was the least developed method. Therefore, as an example, I chose to demonstrate the approach of developing a microbial assessment for heavy metal monitoring in seawaters.

Design of Selective Media

The first step needed for the implication of microbial biosurveys was to design selective media that would allow calculation of a number of pollutant-resistant microorganisms in the water. A number of representative bacteria that commonly occur in coastal seawaters have been chosen in order to define the sensitivity of microorganisms to heavy metals. Bacteria of genera *Escherichia, Pseudomonas*, and *Bacillus* from laboratory collections that did not demonstrate a special resistance to metals were grown in nutrient media, containing metal salt at different concentrations. Every species has an individual sensitivity to the toxic metals. The concentration of a compound that completely inhibited any bacterial growth was considered as a selective factor and then used in selective media in order to detect the appearance and the proportion of metal-resistant microorganisms.

As Fig. 2 demonstrates, the reverse correlation took place between the content of metal in the media and the growth of bacteria, if microorganisms did not have the resistance to the introduced toxic compound. However, when metal contamination is present in water, an abundance of microorganisms resistant to contaminating factors appear among the aquatic microbial communities. The illustration below demonstrates a negative correlation between the concentration of heavy metals and number

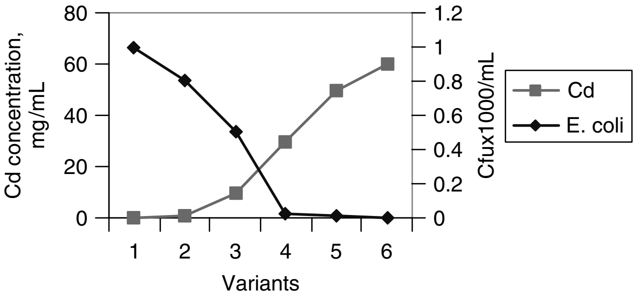

Figure 2. Dependence of bacterial growth on metal concentration in the growth medium.

of surviving microorganisms in the same samples of seawaters.

Chemical and Microbiological Monitoring of Heavy Metals in Seawaters of Peter the Great Bay

The synchronous sampling of heavy metal contents in the northern part of Peter the Great Bay allowed us to compare the results obtained. As graphs demonstrate (Fig. 3a–c), the alteration of a number of the bacteria resistant to heavy metals is an agreement with the concentration of dissolved forms of corresponding metal in the water. We can observe individual response of microbial community to the content of metals in different concentrations in the water under complex pollution conditions in Peter the Great Bay. It points to the existence and sensitivity of the individual response of microbial populations to the content of metals in different concentrations in the water under complex pollution. Moreover, in the case of each heavy metal, the dynamic of that alteration occurs differently within certain intervals of metal concentration. The boundaries between gradual and drastic variations of microbial indices were very clear, which allowed us to formulate microbial criteria that might adequately detect the concentration range of heavy metals in the water.

Average value of the percentage of the metal-resistant forms of bacteria in planktonic and benthal microbial assemblages allows comparison of the water quality from different parts of the large body of water. Moreover, microbial indices indicate from which part of the aquatorium the pollutant enters, from the surface or from the bottom. The next illustration demonstrates such a comparison made on the basis of microbial assessments conducted in several bights and secondary bays of Peter the Great Bay during the summer months of June–August of 1999 (Fig. 4a–d).

Morphological Features of Marine Bacteria Inhabiting Contaminated Waters of Peter the Great Bay

Golden Horn Bight is one of the most impaired water bodies among other bights of Peter the Great Bay. Fishery fleets, trade fleets, the military ship yard, and the city ferry transportation system are among industrial polluters of that aquatorium. Analysis of the morphological characteristics of bacteria in the planktonic assemblages can serve as an indirect characteristic of water pollution. It is a well-known fact that in heavily contaminated marine areas, rod-shaped (over 87%) and gram-negative (over 60%) microflora prevail (70). We conducted microscopic examinations of 117 strains isolated from the near-surface water layer in the central part of Golden Horn Bight, Vladivostok, Russia. Among all microorganisms cultivated, rod-shaped bacteria made up 92%, and 72% of the bacteria were gram-negative.

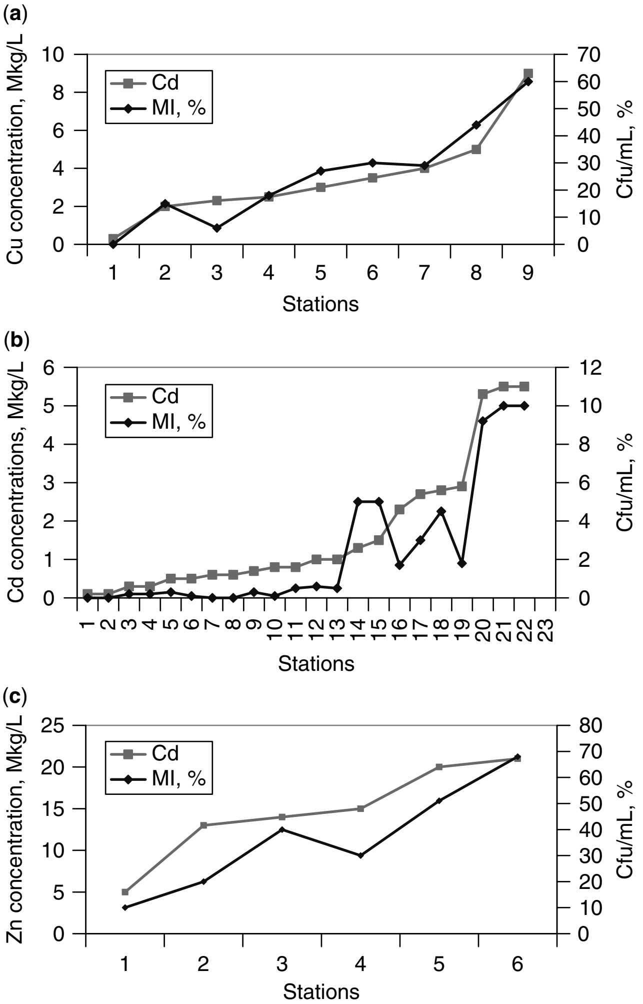

Figure 3. Changes in the values of microbial indices (line MI) and concentration of dissolved forms of metals in the water samples taken from Ussuri Bay: (**a**) copper, (**b**) cadmium, and (**c**) zinc. Along the horizontal axis, the station numbers were sorted with respect to increasing concentrations of heavy metal in the water.

The isolated microorganisms were resistant to several heavy metals at a time, but the combination of metals and the range of bacterial resistance were different, which once again confirms the individual character of microorganism response to the pollutant concentrations in the environment. The use of a transmission electron microscope demonstrated typical symptoms of absorbing specific substances from the environment in both bacilli and cocci and their accumulation both inside the cells and in the cell wall (11). Although these substances were not identified, but based on the electron density of the substances and the detection of multiple metal resistances

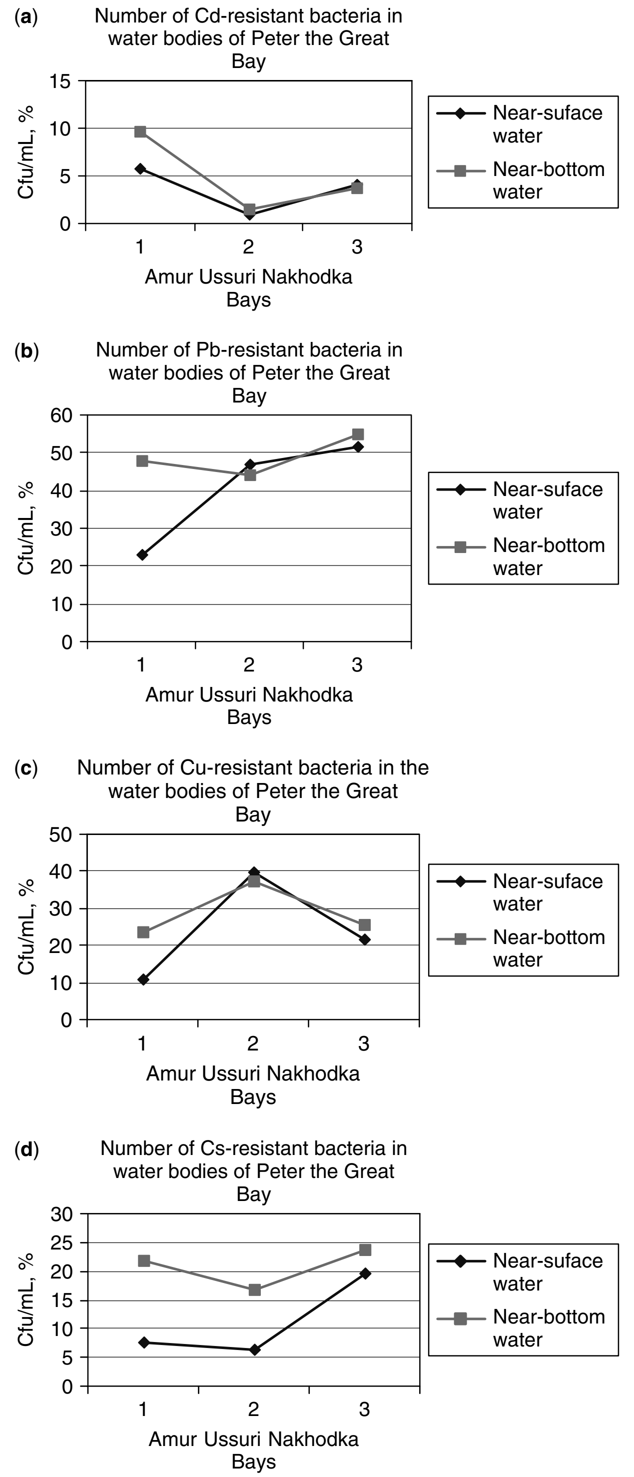

Figure 4. Average value of the number of metal-resistant bacteria in near-surface and near-bottom water samples taken from different water bodies of Peter the Great Bay in June–August of 1999.

in the test bacteria, we assumed that we had observed accumulation of heavy metals by the marine bacteria.

Study of the molecular genetics of Cd-resistant bacteria was also carried out. It was established that all Cd-resistant strains also displayed resistance to antibiotic ampicillin in the concentration of 250 mg/mL (71). Electrophoresis analysis of plasmid DNA from eight strains detected the presence of a small plasmid with a molecular mass of 4 MD in the form of a monomer, dimer, or trimer that were identical to marker plasmid p2728, containing transposon Tn 501 and designed from pUC19 and encoding resistance to Hg and ampicillin (72). Thus, we have demonstrated that marine bacteria, as well as terrestrial bacteria, possess cell mechanisms providing resistance to heavy metals. We assume that numerous bacteria with such morphological and molecular genetic features could serve as indicators of environmental disturbance and contamination in the investigated areas. The presence of such bacteria also suggests the active participation of microbial communities in the process of transformation of substances finding their way into the environment. Depending on the phenomena in question, marine microbial cenoses can act both as indicators of physicochemical and biological processes and as a powerful biotic factor promoting the elimination of pollutants from the marine environment.

Microbial Assessments for Evaluation of Heavy Metal Pollution Level in the Water of Far East Seas

Comparison of Environmental Conditions in Different Regions of Far Eastern Seas Using Microbial Biosurveys. Investigations carried out in a number of regions of Far Eastern seas, having distinguished impacts and ecological situations, resulted in data that confirmed the informative efficiency of the use of microbial assessments for the evaluation of water quality and ecological integrity in the marine environments over geographic and environmental variations taking place in different regions (Table 1).

For example, the data from Table 1 on the microbial detection of heavy metal pollution of Avachinskaya Guba show the presence of a significant portion of cobalt (Co)- and cesium (Cs)-tolerant microorganisms in the microbial associations in the areas where an increased content of radionuclides, such as Co-60 and Cs-137, had formerly been found (73). Krasheninnikov Bay (Fleet harbor), which has been exposed to a pronounced technogenic and anthropogenic impact, is the most contaminated by a complex of heavy metals in this aquatic area. Near the southwestern seaboard of Sakhalin Island (off Kholmsk), the microbial indication confirms the known intense influence of the Tsushima Current on the content of heavy metals in the water (65). In northern Primorie, in Rudnaya Bight, the microbial characteristics reflect the features in the composition of the lead ores mined and the intensity of the technogenic load. The distribution of the metal-resistant groups of microorganisms is the least uniform in Peter the Great Bay, which is caused by the active hydrodynamics and the presence of a large number of sources of heavy metal pollution of both natural and anthropogenic character. Nakhodka Bight, where major commercial, oil, and fishery ports, together with several dockyards are concentrated, is the region most contaminated by complex heavy metals.

Application of Microbial Assessments for Long-Term Monitoring and Evaluation of Water Quality in the Marine Environment. Long-term observation can give very important

Table 1. Relative Abundance of Metal-Resistant Forms (%) with Respect to the Total Amount of the Colony-Forming Heterotrophic Microorganisms in the Near-Shore Waters for Selected Aquatic Areas of the Far Eastern Seas

	Kamchatka Peninsula, Avachinskaya Guba July 1999		Sakhalin Island, Southwest, May 1997	Northern Primorie, Polymetal Ore Mining, July 1999		Peter The Great Bay, July 1999	
Bacteria, Resistant to Metal	Fleet Harbor	Malaya Sarannaya Bay, Reference	Kholmsk	Ore River Mouth	8 km to the South, Reference	Nakhodka Bight, Next to the Shipyard	Big Peles Island, Marine Reservation
Co	13	0	3	6.4	0	2.5	0
Cs	25	2.3	16	19	9.8	10.6	4.7
Cd	0	0	7	0.5	0	4.4	0
Zn	57.3	0.6	1.6	14.3	0.6	62.6	2.3
Ni	4.6	0	15	25.3	5.8	3.9	0.02
Cu	3	0	15.9	7.4	0	40.1	0
Pb	24.5	0.08	10	0.12	0	87.2	0
Fe	0	0	16.1	0	0	41.2	0

environmental information about the intensity of anthropogenic and industrial load on water bodies. Table 2 demonstrates the dynamic of the presence of metal-resistant forms of bacteria in seawater of the port, where Pb-Zn concentrate is transported by sea along the coast, which is located next to the lead smelter in Rudnaya Bight (see Fig. 5). Once again, when microbial indices are compared with the data of chemical analysis, they were in agreement. Moreover, both data of chemical and microbiological monitoring reflected the intensity of the ore processing and the production of the lead concentrate over time.

Years 1997 and 2000 are known as a period of intensive ore processing and the operation of the lead smelter at full production. On the contrary, the situation during 1998–1999 resulted in a drop in the volume of extracted and processed ore because of the general economic recession and long-term forced inactivity of the plant. The dynamics of the microbial assessments remarkably accurately reflect the environmental situation in Rudnaya Bight during that period.

Use of GIS Technology to Indicate Concentrations of Heavy Metals Found in Marine Sediments and to Show the Distribution of Metal-Resistant Bacteria in the Water. In the southwestern part of the bay, the environmental situation is mainly determined by the effect of the Tumen River flowing through the industrialized regions of China and North Korea. From the data of the microbiological analysis of this area, we found that stations located in the zone of riverine runoff waters were characterized in a state where the entire association of microorganisms cultivated showed an increased resistance to the presence of Ni, Zn, and Fe in the environment. In the communities mentioned, there exists a considerable portion of bacteria with an increased resistance to the presence of Cd, Pb, Cu, Co, and Cs in the environment. The data from the chemical analysis and the determination of the content of heavy metals in the tissues of hydrobionts confirm the contamination of the aquatic area by the pollutants listed in the scientific report (74). With a sufficient density in the network of monitoring stations, the data from microbial indication may be applied to preliminary biogeochemical

Table 2. A Number of the Highly Phenol-Resistant and Oil-Resistant Colony-Forming Heterotrophic Microorganisms (10^3 per mL) in the Near-Shore Waters for Selected Aquatic Areas of the Far Eastern Seas

Region	Kamchatka Peninsula, Avachinskaya Guba Bay, April 1996		Sakhalin Island, Southwest, Kholmsk, May 1997		Northwest Sakhalin, Nyi Bay, August 1996		Peter the Great Bay, July 1999	
Microbial Groups	Fleet Harbor	Ocean Waters, 5 km from Coast, Reference	Port	Destroyers Pulp and Paper Plant	Exit into Sea	Oil Drill	Amur Bay, Next to Domestic Wastewater	Big Peles Island, Reservation
Heterotrophic bacteria	800 ± 130	500 ± 110	600 ± 3	400 ± 3	3000 ± 300	1000 ± 80	24000 ± 307	60 ± 5
Destroyers of:								
Oil	40 ± 8	2 ± 0.3	100 ± 20	20 ± 0.1	1 ± 0.1	40 ± 4	130 ± 15	3 ± 0
Fuel oil	90 ± 7	3 ± 0.3	4 ± 0.7	30 ± 0.2	0.1 ± 0.01	40 ± 1	100 ± 16	0.8 ± 0.001
Engine fuel	7 ± 1	10 ± 2	2 ± 0.1	0	10 ± 0.8	0.7 ± 0.08	20 ± 4	0
Phenol	0	0	0.05 ± 0	20 ± 0	0.003 ± 0	0	3 ± 0.25	0

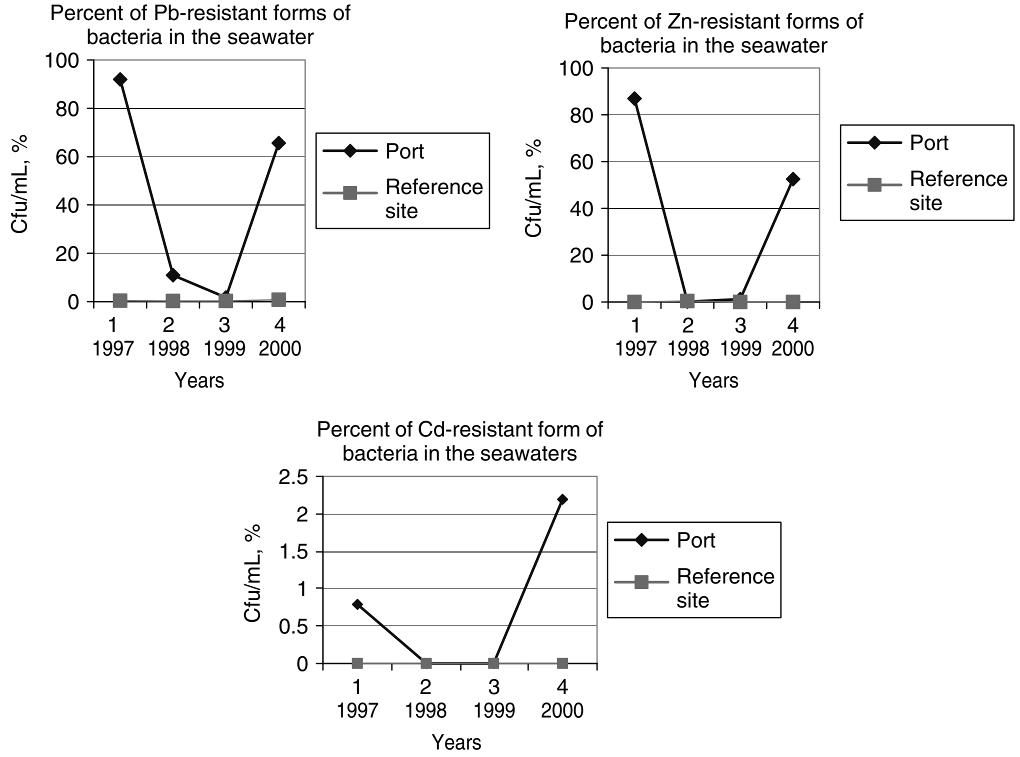

Figure 5. Long-term dynamics of the percent of metal-resistant forms of bacteria in the near-surface water layer of the Rudnaya Bight. Reference station was 2 miles away from the lead smelter.

mapping of the near-shore aquatic areas, which was demonstrated by the example in the area off the mouth of the Tumen River for two of the eight heavy metals analyzed (Fig. 6).

Use of Microbial Biosurveys for Assessments of Organic Pollution in Seawater. Microbial assessments proved to be even more efficient for monitoring organic pollution. This statement is based on the ability of microorganisms to use organic compounds as a source of carbon and energy. As it was reviewed above, the presence of such chemicals results in an increase of bacteria using hydrocarbons for their living needs.

As Table 3 demonstrates, microbial criteria at reference sites are significantly less than at impact areas. Also, microbial assessments give information about sources of the contamination. For example, phenols can be a part of fecal contamination or they can be the result of the waste from cellulose processing. Seawater samples contained phenol destroyers taken either from places where a number of ships are present (Kholmsk, Port; Nyi Bay, exit to the sea), near the domestic wastewater dumping site (in Amur Bay), or next to plants manufacturing products from cellulose (Kholmsk, paper plant). Aside from the case of the paper plant, the number of phenol destroyers was ten times as much in comparison to fecal contaminated sites and correlated with concentrations of phenol in contaminated waters.

A number of microorganisms destroying oil and oil by-products also had a connection with the content of those chemicals in the water. The presence of point or nonpoint contaminating sources close to the beach explains the reason for distribution of those specific microorganisms in the coastal waters. Moreover, the proportion between bacteria resistant to different oil by-products reflected the presence of those particular chemicals in seawaters, which may even explain the origin of the contaminant detected far away from land. For example, it is understandable why the water samples taken from port aquatoria have a significant number of oil and fuel oil microbial destroyers (Kamchatka, Fleet harbor). However, isolation of engine fuel microbial users from the surface of the open ocean water far away from Fleet harbor may cause us to think that bacteria trace the passage of ships that use that kind of fuel.

Example of the Implications of Microbial Criteria for Water Quality Control of Seawater in the Seas of Far East Russia. After years of monitoring the water quality in a large number of water bodies using microbial assessments and chemical habitat metrics, we worked out criteria that can be used for monitoring ecological situations and the quality of seawaters (Table 4). Those criteria can be used during the first evaluation of water samples in order to receive general information about the ecological status of the water body. Then, criteria should be adjusted with respect to all other metrics measured in the water.

Use of criteria, covering a wide number of polluting factors, gives information about the incredible value of

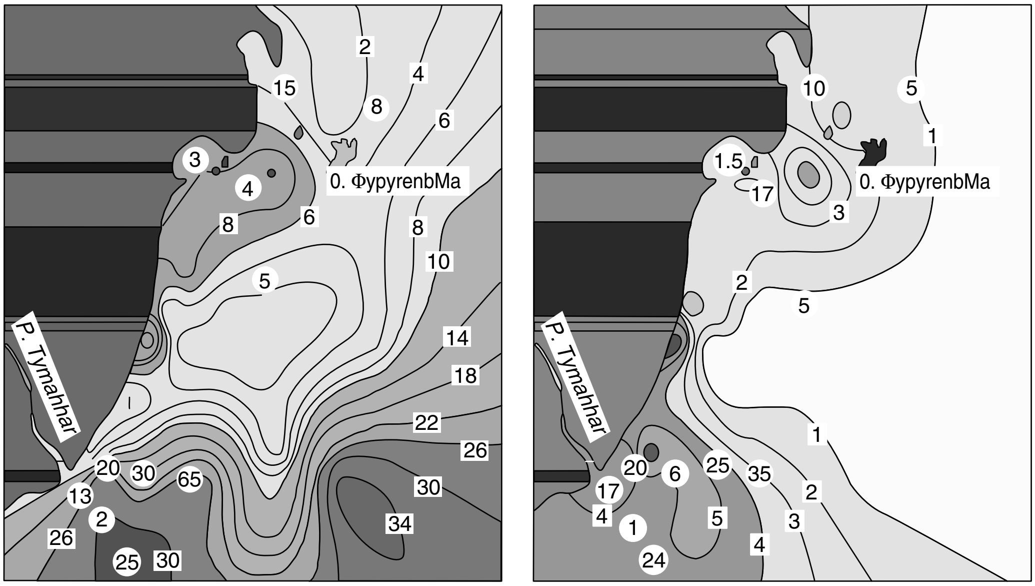

Figure 6. Distribution of the metal-resistant forms of bacteria over the aquatic area adjacent to the mouth of the Tumen River: (left) Cu- and (right) Co-resistant forms. Concentrations of metals in the bottom sediments (in μg/kg of dry mass) are given within the squares; microbial indices (%) are given in the circles.

this global approach, because we can present data on the ecological integrity of a larger system and complete an environmental situation report about the marine or freshwater water body.

Table 4 demonstrates the use and advantage of the microbial assessments for the evaluation of the environmental water quality in the example of data, presented above. The use of symbols better assists in presenting the data on ecological integrity of the marine environment at the moment.

On the basis of data received by a number of scientists over a long-term period, we can conclude that application of microbial criteria and microbial assessments is shown to serve as an operative method of monitoring and short-term forecasting of changes in the environmental conditions of the coastal waters of the sea. This method allows us to

Table 3. Criteria of Water Area Contamination on the Basis of Microbial Characteristics in Absolute Figures or in Percentage to the Maximum Number of Colony-Forming Microorganisms per 1 mL of Seawater

	Level of Contamination				
Pollutant	Background (Below the MAC)[a]	Insignificant (Around 1 MAC)[a]	Perceptible (1–3 MAC)	Considerable (Over 3 MAC)	Strong (Over 10 MAC)
Heterotrophic bacteria	$<10^3$	$10^3 < 10^4$	$10^4 < 10^5$	10^5–10^6	$>10^7$
Oil by-products	10–10^3	$>10^3 < 10^4$	10^4–10^5	10^4–10^5	$>10^5$
Phenol	0	1–9	10–$<10^2$	10^2–$<10^3$	$>10^3$
Pathogenic bacteria	0	—	—	various	various
Intestinal bacteria	0	0	1–9	10–24	>24
Proteins	<10%	—	10–46%	>46%	—
Polysaccharides	0	1–9	10–40%	>40%	—
Lipids	<10%	—	10–46%	>46%	—
Metals					
Cadmium	<0.01%	<0.5%	0.5–6%	>6%	—
Cobalt	<10%	10–46%	>46%	—	—
Cesium	<10%	10–46%	>46%	—	—
Lead	<0.1%	<10%	10–46%	>46%	—
Copper	<0.1%	<10%	10–46%	>46%	—
Zinc	<1%	<20%	20–46%	>46%	—
Nickel	<1%	<20%	20–46%	>46%	—
Iron	<20	<50%	>50%	—	—
Symbol	**O**	✻	□	▲	■

[a] MAC is a maximum admissible concentration of chemical in the water (for fishery water bodies).

Table 4. Evaluation of Environmental Quality of Krasheninnikov Bay (Kamchatka Peninsula, Avacha Bay, 1996)

Stations Pollution	Military Depot	Fleet Harbor	Rybachi, Navy Base Residences	Ferry-Boat Station	Ocean Waters, 3 miles Away from Coast
Heterotrophic bacteria	▲	▲	▲	▲	□
Intestinal bacteria	▲	O	■	O	■
Pathogenic bacteria	O	O	■	O	O
Oil	□	□	□	□	O
Fuel oil	□	□	□	□	O
Engine oil	✻	□	□	□	□
Lipids	O	O	O	□	O
Peptides	O	O	O	O	O
Polysaccharides	O	O	O	O	O
Cd	O	O	O	O	□
Cu	✻	□	✻	✻	✻
Ni	✻	✻	✻	✻	✻
Zn	O	✻	O	O	
Co	O	✻	O	O	O
Cs	✻	✻	O	O	O

O—Background (below the MAC); ✻—insignificant (around 1 MAC); □—perceptible (1–3 MAC); ▲—considerable (over 3 MAC); ■—strong (over 10 MAC); MAC is maximum admissible concentration of heavy metals in the water (for fishery water bodies in Russia).

obtain preliminary information, to create biogeochemical maps, preceding labor-intensive and expensive chemical analysis, and, e.g., at the initial stage when we select environmental quality monitoring stations. The microbial biosurvey does indeed permit control of the spread of pollutants in the near-shore waters.

BIBLIOGRAPHY

1. Izrael, Yu.A. and Tsiban, A.V. (1985). *Anthropogenic Ecology of the Ocean*. Hydrometeoizdat, Leningrad, p. 528.
2. Korneeva, G.A. (1996). Utilizing of enzyme test-systems for monitoring of sea waters condition of Black Sea. *Izvestiya RAN, Ser. Biol.* **5**: 589–597.
3. Novozhilova, M.I. and Popova, L.E. (1984). Bacterioplankton of meddle and south Kaspi. In: *Ecological Aspects of Aquatic Microbiology*. Nauka, Novosibirsk, pp. 35–43.
4. Sorokin, Yu.I. (1990). *Ecosystems of Coral Rife*. Nauka, Moskva, p. 503.
5. Teplinskaya, N.G. (1990). *Processes of Bacterial Production and Destruction of Organic Compounds in North Seas*. KNTS AN SSSR, Apatiti, p. 105.
6. Tsiban, A.V. (1970). *Bacterioneiston and Bacterioplankton of Shelf Zone of Black Sea*. Naukova Dumka, Kiev, p. 224.
7. Tsiban, A.V., Panov, G.V., and Barinova, S.P. (1990). Indicator microflora in Baltic Sea. In: *Investigation of Ecosystem of Baltic Sea*. Hydrometeoizdat, Leningrad, Issue 3, pp. 69–83.
8. Dimitrieva, G.Yu., Bezverbnaya, I.P., and Cemikina, G.I. (2000). Microbiological monitoring heavy metal pollution in coastal waters of Peter the Great Bay. *Proc. Pacific Institute Fishery and Oceanology (TINRO)*. Vol. 127, Vladivostok, Russia, pp. 657–676.
9. Dimitrieva, G.Yu., Bezverbnaya, I.P., and Khristoforova, N.K. (2001). Microbial indication is a possible approach for monitoring of heavy metal in Far Eastern Seas. *Proc. Pacific Institute Fishery and Oceanology (TINRO)*. Vol. 128, Vladivostok, Russia, pp. 719–736.
10. Dimitrieva, G.Yu. and Bezverbnaya, I.P. (2002). Microbial indication is the efficient tool for monitoring of heavy metal pollution in coastal sea waters. *Oceanology* **42**: 1–8.
11. Bezverbnaya, I.P., Dimitrieva, G.Yu., Tazaki, K., and Vatanabe, H. (2003). Evaluation of sea water quality in the coastal zone of Primirie using the method of microbial indication. *Water Res.* **30**: 199–208.
12. Nitcowski, F., Dudley, Sh., and Graikowski, J.T. (1977). *Mar. Poll. Bull.* **8**: 276–279.
13. Zoffman, M., Rodrigues-Valeria, F., and Peres-Filloi, M. (1989). *Mar. Poll. Bull.* **20**: 74–81.
14. Aubert, M. (1990). *Mar. Poll. Bull.* **21**: 24–29.
15. Tsiban, A.V., Panov, G.V., and Barinova, S.P. (1992). Heterotrophic saprotrophic microflora. Distribution, ecological and physiological properties. In: *Investigations of Ecosystems of Bering and Chukotsk Seas*. Hudrometeoizdat, St. Petersburg, Issue 3, pp. 143–165.
16. Gavrishina, N.A. (1980). O komplekse mikrobiologicheskih pokazatelei pri harakteristike kachestva void. In: *Smoochischenie I Bioindikatsiya Zagryaznennih Vod*. Nauka, Moskva, pp. 74–78.
17. Telitchenko, M.M. and Kokin, K.A. (1968). *Sanitary Hydrobiology*. MGU, Moskva, p. 102.
18. Kriss, A.E. (1976). *Microbiology of the Ocean*. Nauka, Moskva, p. 286.
19. Loransaki, D.N., Raskin, B.M., and Alhimov, N.N. (1975). *Sanitary Protection of the Sea*. Meditsina, Moskva, p. 168.
20. Fisher, B. (1984). Untersuchungen. *Ergebnisse der Plankton-Expedition der Humboldt-Stiftung* **4**: 1–83.
21. Edwards, D.D., McFeters, G.A., and Venkatesan, M.L. (1998). *Appl. Environ. Microbiol.* **64**: 2596–2600.
22. Eliot, E.L. and Colwell, R.R. (1985). *FEMS Microbiol. Rev.* **32**: 61–79.
23. Tabor, P.S. and Colwell, R.R. (1976). Initial investigations with a deep ocean *in situ* samples. In: *Proc. Oceanology, Marine Technol. Soc.*, Washington, DC, pp. 13D–1–13D-4.
24. ZoBell, C.T., Granth, C., and Haas, H. (1943). *Bull. Am. Assoc. Petrol. Geol.* **24**: 75.
25. Atlas, R.M. and Bartha, R. (1973). *Environ. Sci. Technol.* **7**: 538–541.
26. Guilott, M.L. (1983). *Sci. Techn.* **94**: 35–36.
27. Guerin, W.F. and Boyd, S.A. (1995). *Appl. Environ. Microbiol.* **61**: 4061–4068.
28. Iriberri, J. and Hernbel, G.J. (1995). *Micribiologia (Madrid)* **11**: 309–322.
29. Gusev, M.V., Koronelli, T.V., and Sentsova, O.Yu. (1985). Using of microorganisms as bioindicators to study ecological consequences of water pollution. In: *Ecological Consequences of the Ocean Pollution*. Hydrometeoizdat, Leningrad, pp. 113–127.
30. Izyurova, A.I. (1950). The rate of oil by-products destruction in water and soil. *Hygiene and Sanitary* **5**: 23–28.
31. Voroshilova, A.A. and Dianova, E.V. (1952). Oil-oxidizing bacteria as detectors of intensity of biochemical oxidizing of oil at natural conditions. *Microbiology* **24**: 64–73.
32. Mironov, O.G. (1980). Biological indication of carbohydrates in the sea. In: *Self-Purification and Bioindication of Contaminated Waters*. Nauka, Moskva, pp. 129–134.

33. Kong, H.-L. and Saylor, G.S. (1983). *Appl. Environ. Microbiol.* **46**: 666–672.
34. Hartmans, S., Bont, J.A., and Harder, W. (1989). *FEMS Microbiol. Rev.* **63**: 235–264.
35. Sigoillot, J.-C. and Nguen, M.-H. (1992). *Appl. Environ. Microbiol.* **58**: 1308–1312.
36. Divavin, I.A. and Kopilov, Y.P. (1977). Investigation of some biochemical aspects for biodegradation of carbohydrates at the condition of combined pollution. Abstracts of Workshop *Problems of the Marine Environment Protection*. Kaliningrad, Russia, pp. 31–33.
37. Dimitrieva, G.Yu. (1995). Microorganisms as bioindicators of phenol pollution of the marine coastal environment. *Russian J. Mar. Biol.* **21**: 407–411.
38. Dutka, B.J., Chay, A.S., and Corburn, J. (1974). *Water Res.* **8**: 1047.
39. Ayres, P.A. (1977). *Mar. Poll. Bull.* **8**: 283–285.
40. Goodfelow, R.M., Cardoso, J., and Eglinton, G. (1977). *Mar. Pol. Bull.* **8**: 272–276.
41. Janase, H., Zuzan, K., and Kita, K. (1992). *J. Ferm. Biol.* **74**: 297–300.
42. Kalmikova, G.Ya., Lazareva, M.F., Rogovskaya, Ts.I., Skirdov, I.V., and Tihaya, N.B. (1980). Metabolizm aromaticheskih soedinenii adaptirovannoi I neadaptirovannoi mikrofloroi aktivnih ilov. In: *Samoochischenie I Bioindikatsiya Zagryznennih Vod*. Nauka, Moskva, pp. 171–175.
43. Tsiban, A.V. (Ed.). *Rukovodstvo po Metodam Biologicheskogo Analiza Morskoi Void I Donnih Otlozhenii*. Hydrometeoizdat, Leningrad, p. 193.
44. Kondratyeva, L.M. (1996). Ekologicheskii potentsial morskih mikroorganizmov. Ph.D. (Biol.) Dissertation, HKI, Habarovsk, p. 412.
45. ZoBell, C.T. and Hittle, L.L. (1969). *J. Oceanogr. Soc. Japan*. **25**: 36–47.
46. Uchida, M. (1995). *Mar. Biol.* **123**: 639–644.
47. Dimitrieva, G.Yu. and Dimitriev, S.M. (1996). Symbiotic microflora of brown algae of genus *Laminaria* as a bioindicator of ecological condition of coastal laminarian biocenoses. *Russian J. Marine Biol.* **22**: 300–305.
48. Slavina, G.P. (1957). O vozmozhnosti ispol'zovaniya mikroorganizmov pri poiske rudnih mestorozhdenii. *Trudi 1 Vsesoyuznogo Coveschaniya po Geohimicheskim Metodam Poiska Rudnih Mestorozhdenii*. Gosgeoltehizdat, Moskva, pp. 302–305.
49. Belyakova, E.E., Reznikov, A.A., Kramarenko, L.E., Nechayeva, A.A., and Kronidova, T.F. (1962). *Gidrohimicheskii Method Poiskov Rudnih Mestorozhdenii v Aridnih I Poluaridnih Oblastyah*. Gosgeoltehizdat, Moskva, pp. 173–228.
50. Glazovskaya, M.A. and Dobrovolskaya, N.G. (1984). *Geochemical Functions of Microorganisms*. MGU, Moskva, p. 152.
51. Vernadski, V.I. (1967). *Biosphere*. Nauka, Moskva, p. 100.
52. Letunova, S.V. and Koval'ski, V.V. (1978). *Geohimicheskaya Ecologiya Mikroorganismov*. p. 147.
53. Gromov, B.V. (1985). *Stroenie Bacterii*. LGU, Leningrad, p. 190.
54. Mullen, M.D. et al. (1989). *Appl. Environ. Microbiol.* **55**: 3143–3149.
55. Grinyus, L.L. (1986). *Transport Molekul u Bacterii*. Nauka, Moskva, pp. 6–13.
56. Gadd, G.M. (1990). In: *Metal Tolerance*. C. Edwards (Ed.). Alden Press, Oxford, UK, pp. 178–211.
57. Gadd, G.M. (1993). *Phytology* **124**: 25–60.
58. Hingman, J.P., Sadler, P.L., and Scawen, M.D. (1985). *J. Gen. Microbiol.* **131**: 2539–2544.
59. Petterson, A., Kint, D., Bergman, B., and Roomans, Y.M. (1985). *J. Gen. Microbiol.* **131**: 2545–2548.
60. Schutt, C. (1989). *Microbial. Ecol.* **1**: 59–62.
61. Silver, S. (1990). *Forum. Microbiol.* **13**: 36.
62. Galaktionov, S.G. and Yurin, V.M. (1984). *Alga Danger Signals*. Nauka, Moskva, p. 356.
63. Kosolapov, V.N. and Sinyova, L.N. (1982). The action of phenol and waste waters from wood factory on motility of hydrobionts. In: *Toxicological and Ecological Aspects of the Environmental Pollution*. Irkutskii State University, Irkutsk, pp. 91–95.
64. Flerov, B.A. (1974). About using of animal behavior investigations in aquatic toxicology. *Hydrobiological J.* **5**: 114–118.
65. Khristoforova, N.K. (1985). *Bioindication of Seawater Contamination With Heavy Metals*. Ph.D. (Biol.) Dissertation, the Pacific Institute of Geography, Vladivostok, p. 394.
66. Dimitrieva, G.Yu. (1999). Plankton and epiphyte microorganisms: detection and stabilization of the marine coastal environment. Ph.D. (Biol.) Dissertation, Far Eastern State University, Vladivostok, Russia, p. 408.
67. Dimitrieva, G.Yu. (1999). The role of microorganisms in control and reservation of marine coastal environment. *Proc. Intern. Symp. "Earth—Water—Humans*. May 30–June 1, 1999, Kanazawa University, Kanazawa, Japan, pp. 22–35.
68. Dimitrieva, G.Yu., Dimitriev, S.M., Bezverbnaya, I.P., and Khristoforova, N.K. (2000). The role of microorganisms for control and restoration of the marine coastal environment. In: *Waste Recycling and Resource Management in the Developing Word*. B.B. Jana et al. (Ed.). Univ. of Kalyani, India and Intern. Ecol. Eng. Soc. Switzerland, pp. 437–447.
69. Bogdanov, D.V. (1981). *Short Physical and Geographic Characterization of Main Seas. The Pacific Ocean*. Nauka, Leningrad, pp. 129–143.
70. Izrael, Yu.A. and Tsyban, A.V. (1992). *Studies of the Bering and Chukotka Seas*. Gidrometeoizdat, Petersburg, Issue 3.
71. Dimitrieva, G.Yu. (2003). Microbial Detection as an effective technique for monitoring of heavy metal and organic pollution of coastal seawater. *Proc. Int. Workshop on Bioassay of the Marine Environment in the Northwest Pacific Region, Toyama, Japan*. March 15–16, 2003, Northwest Pacific Region Environmental Cooperation Center, Toyama, Japan, pp. 24–31.
72. Yanish-Perron, C., Viera, T., and Nessing, J. (1985). Improved M13 phage cloning vectors and in host strains: nucleotide sequence of the M13, mp 18, and pUC19 vectors. *Gen.* **33**: 103–119.
73. *Estimation of the Pollution Condition of the Area of Avachinskaya Inlet*. City of Petropavlovsk-Kamchatskii, and Primorskii and Rybachii Settlements. (1993). Report, Karpinskii Russian Geol. Inst., p. 346.
74. *Multidisciplinary Ecological Estimate of the Influence of the Tumen River Runoff on the Coastal Waters of Russia*. (1998). Report, Ross. Akad. Nauk, Vladivostok: IBM DVO, Part I.

READING LIST

Dimitrieva, G.Yu. et al. (1999). Phenol detoxification by microorganisms of the sea coastal zone. *Microbiology* **68**: 107–113.

LUMINESCENT BACTERIAL BIOSENSORS FOR THE RAPID DETECTION OF TOXICANTS

Maria B. Davoren
Dublin Institute of Technology
Dublin, Ireland

INTRODUCTION

The need to rapidly screen and assess the potential toxicity of aquatic contaminants has led to the development of several short-term bioassays utilizing bacteria, algae, protozoa, invertebrates, and cell cultures. Bioassays with microorganisms are frequently employed as ecotoxicological screening tools to monitor the hazards of chemical contaminants in the environment. In addition to their ecological importance as decomposers, microbial test systems have several advantages over traditional bioassays, including rapid response times, reproducibility of test conditions, amenability to genetic manipulations, increased sensitivity, and cost saving. Increasingly, luminescent microbial biosensors are being developed and employed for environmental monitoring (1–4). A biosensor integrates a biological sensing component (e.g., bacteria) with a transducer or detection system, which produces a measurable signal (e.g., luminescence) upon exposure of the sensing component to a specific or nonspecific analyte (5). Here the focus is on the use of both natural and genetically engineered luminescent bacterial biosensors for the assessment of toxicity and bioavailability.

BIOLUMINESCENT REPORTER GENES

Bioluminescence, which has been observed in a variety of organisms, is the production of light from living cells by the oxidation of a luciferin (a pigment) in the presence of the enzyme luciferase (6,7). Two frequently exploited bioluminescent reporter systems are the bacterial (the marine *Vibrio fischeri* and terrestrial *Photorhabdus luminescens*) and insect (the firefly, *Photinus pyralis*, and the click beetle, *Pyrophorus plagiophthalamus*) luciferase reactions.

Bacterial bioluminescence is encoded for by *lux* genes and light is produced through the oxidation of fatty aldehydes (8–10). The complete *lux*-gene cassette consists of five structural genes, *lux*CDABE. The *lux*AB genes encode the luciferase enzyme and the *lux*CDE genes encode an enzyme complex that synthesizes the aldehyde substrate for luciferase (3). Bacterial biosensors have been created using either the *lux*AB or the entire *lux*CDABE reporter system, but where the shorter gene construct is employed, an aldehyde substrate (usually *n*-decanal) must be added exogenously before bioluminescence can be measured (11,12).

The *luc*FF and *luc*GR genes encode for firefly and click beetle bioluminescence, respectively. In these eukaryotic systems, light is produced when the luciferase enzyme catalyzes the reaction between the substrate luciferin (which must be added exogenously), adenosine triphosphate (ATP), and oxygen (1,11). Rapid and sensitive detection and quantification of bioluminescence can be achieved using luminometers, liquid scintillation counters, or photographic films (4,13).

BACTERIAL BIOSENSORS BASED ON NATURAL BIOLUMINESCENCE

The naturally luminescent marine bacterium *V. fischeri* (formerly *Photobacterium phosphoreum*) has been extensively employed to generate acute toxicity data. Kaiser and Palabrica (14) published *P. phosphoreum* toxicity data for more than 1300 chemicals. A rapid screening test based on bioluminescence inhibition in *V. fischeri* (NRRL B 11177) has been commercialized as the Microtox test system (Azur Environmental, CA, USA) and is now an internationally adopted bioassay (15,16). The test system measures the light output of the luminescent bacteria after they have been challenged by a sample and compares it to the light output of a control (reagent blank). As bacterial bioluminescence is directly linked to cellular respiration, a difference in light output (between the sample and control) can be attributed to the toxic effect of the sample. The Microtox test has also been widely employed to generate acute toxicity data for environmental samples such as wastewaters (17), industrial effluents (18), and sediment samples (19). For a comprehensive review of the use of Microtox for assessing the toxicity of sediments and soils, the reader is referred to Doherty (20).

In addition to the well established Microtox test system, other toxicity bioassays have been developed which utilize *V. fischeri*. These include BioTox (Aboatox Oy, Finland), LUMIStox (Dr. Lange GmbH, Düsseldorf, Germany), and two portable systems ToxAlert (Merck Ltd., UK) and Deltatox (Azur Environmental, CA, USA). These test systems have been used either individually or in combination with each other in toxicity assessments (21–28). The ToxScreen-II assay (Checklight Ltd., Israel) employs another naturally luminescent marine bacterium, *Photobacterium leiognathii*, as a test organism (29). This test enables discrimination between cationic heavy metals and organic toxicants and has the added advantage of being able to be run at ambient temperatures.

ECOLOGICALLY RELEVANT BACTERIAL BIOSENSORS CONSTRUCTED USING LUMINESCENT-MARKED GENES

Despite the widespread use of Microtox and other test systems utilizing *Vibrio* sp., one frequently cited shortcoming of these assays is the lack of ecological relevance for the testing of freshwater and terrestrial environments (30,31). *Vibrio fischeri* is a marine bacterium and as such is sensitive to pH and osmotic conditions. The genes for bacterial bioluminescence (*lux*AB and *lux*CDABE) have therefore been cloned and inserted using plasmids into environmentally relevant terrestrial and freshwater bacterial test strains (32–35). As with the *V. fischeri* based biosensors, luminescence is dependent on metabolic activity; therefore, any chemical causing metabolic stress will result in a reduction in bioluminescence. Luminescent marking of terrestrial bacteria offers scope for environmentally relevant, whole-cell biosensors that can either be toxin specific or general indicators of pollutant toxicity. These genetically modified bacteria have the added advantage of not being limited by a narrow pH range.

RECOMBINANT LUMINESCENT BACTERIAL BIOSENSORS

Recent advances in reporter gene technology have led to the development of many luminescent marked bacterial biosensors, which have been employed as rapid bioassays to assess the toxicity of a wide range of pollutants. Some of the bacterial biosensors that have been constructed are used for general toxicity testing as they constitutively express luminescence and respond to toxicants in a "lights off" manner, with a decrease in light output related to exposure of the cells to pollutants (Table 1). These bacteria are not toxicant specific and cannot be used to determine the type or the actual concentration of a given toxicant, but rather are general indicators of adverse conditions. As no single biosensor will have a universal sensitivity to all toxicants, multiple luminescent bacteria are increasingly being employed, in a battery style approach to better characterize the toxicity associated with a chemical (50–52). Constitutively expressing biosensors have been used to assess the toxicity of a wide range of pollutants including heavy metals (32), chlorobenzenes (31), chlorophenols (37), industrial effluents (43), and wastewater (53).

In addition to constitutive expression, other biosensors have been developed that react to toxicants in a "lights on" manner, where light production is induced by bioavailable concentrations of toxicants. These recombinant bacteria have been constructed by the molecular fusion of two elements inside a bacterial host, a promoter gene regulated by specific chemical agents or their pathway intermediates, and a promoterless reporter gene (e.g., *luc*FF or *lux*CDABE) inserted as a plasmid. In these toxicant-specific bacterial sensors, a genetic regulatory unit controls the expression of the reporter gene. In the presence of a toxicant this regulatory unit is induced and generates a measurable signal relative to toxicant level, that is, receptor–reporter concept (54). These microbial biosensors offer a powerful new approach to environmental monitoring, providing an indication of the bioavailabilty of specific pollutants in complex environments, rather than the total concentrations obtained by traditional analytical techniques. Biosensors constructed for the detection of specific pollutants in the environment can complement analytical methods by distinguishing bioavailable from immobilized, unavailable forms of contaminants, and frequently an integrated approach using both methods is employed (36,55). In order to assess the overall toxicity of the sample, constitutively expressing reporter biosensors (non-toxicant-specific bioassays) are often run in parallel with toxicant-specific inducible biosensors (36,48,56). Many luminescent bacterial biosensors have been developed for assessing pollutant bioavailability in environmental samples, selected examples of which are presented in Table 2.

Table 1. Selected Examples of Non-Toxicant-Specific Luminescent Biosensor Strains

Bacterial Biosensor Strain	Plasmid	Luminescent Reporter Gene	Tested Compounds[a]	Reference
Bacillus subtilis BR151	pCSS962/pBL1	*luc*FF	Metal-contaminated soils	36
Escherichia coli HB101	pUCD607	*lux*CDABE	2,4-DCP; atrazine, simazine, propazine, mecoprop, MCPA, diuron, paraquat; As, $CuSO_4$, 2,4-DCP, 3,5-DCP, bronopol, TTAB, $ZnSO_4$	37–39
Escherichia coli MC1061	pCSS810	*luc*FF	Metal-spiked water samples	40
Pseudomonas putida F1	pUCD607	*lux*CDABE	Atrazine, simazine, propazine, mecoprop, MCPA, diuron, paraquat; oil-polluted soil	38,41
Pseudomonas fluorescens 10586r	pUCD607	*lux*CDABE	Atrazine, simazine, propazine, mecoprop, MCPA, diuron, paraquat	38
Pseudomonas fluorescens 10586s	pUCD607	*lux*CDABE	Copper in whiskey distillery effluent; papermill effluent; contaminated groundwater; benzene, catechol, phenol; chlorobenzenes; 2,4-DCP; metal-contaminated soils	31,37,42–46
Pseudomonas fluorescens OS8	pNEP01	*luc*GR	Soil-associated arsenite and mercury	47
Ralstonia eutropha ENV307	pUTK2	*lux*CDABE	Polychlorinated biphenyls	48
Staphylococcus aureus RN4220	pCSS810	*luc*FF	Metal solutions	49
Staphylococcus aureus RN4220	pTOO02	*luc*FF	Metal-contaminated soils	36

[a] As = sodium arsenite; bronopol = 2-brom-2-nitro-1, 3-propandiol; 2,4-DCP = 2,4-dichlorophenol; 3,5-DCP = 3,5-dichlorophenol; MCPA = 2-methyl-4-chlorophenoxyacetic acid; TTAB = tetradecyltrimethylammonium bromide.

Table 2. Selected Examples of Toxicant-Specific Luminescent Biosensor Strains

Bacterial Host Strain	Sensor Plasmid	Luminescent Reporter Gene	Inducing Compound(s)	Reference
Bacillus subtilis BR151	pTOO24	*luc*FF	Cadmium, antimony, zinc, tin; metal-contaminated soils	36,49
Escherichia coli AW3110	pTOO31	*luc*FF	Arsenic	40
Escherichia coli MC1061	pTOO31	*luc*FF	Arsenic; arsenic-contaminated sediments; arsenic-contaminated soils	40,47,57
Escherichia coli MC1061	pTOO11	*luc*FF	Mercury	40,58
Escherichia coli MC1061	pzntRluc	*luc*FF	Zinc, cadmium, mercury in spiked soil	59
Escherichia coli MC1061	pmerBR$_{BS}$luc	*luc*FF	Mercury; mercury, cadmium, zinc in spiked soil	59,60
Escherichia coli MT102-PIR	pUT-mer-lux	*lux*CDABE	Mercury	61
Staphylococcus aureus RN4220	pTOO24	*luc*FF	Cadmium, lead, antimony, tin; metal-contaminated soils	36,49
Staphylococcus aureus RN4220	pTOO21	*luc*FF	Arsenic	62
Pseudomonas fluorescens OS8	pTPT31	*luc*GR	Arsenite; arsenite-spiked soil; arsenic-contaminated sediments	47,56,63,64
Pseudomonas fluorescens OS8	pTPT11	*luc*GR	Mercury; mercury-spiked soil	56,63,64
Ralstonia eutropha AE104	pchrBluc	*luc*FF	Chromate	59
Ralstonia eutropha AE2515	pMOL1550	*lux*CDABE	Nickel, cobalt, bioavailable nickel in soils	65
Ralstonia eutropha ENV307	pUTK60	*lux*CDABE	Polychlorinated biphenyls	48

LUMINESCENT BACTERIAL BIOSENSORS FOR ASSESSING POTENTIAL GENOTOXICITY

Bioluminescent assay systems have also been employed to assess the potential genotoxicity of both chemicals and environmental samples. The Mutatox test (Azur Environmental, CA, USA) employs a dark mutant of *V. fischeri* (strain M169) and works in an opposite manner to the Microtox test, in that this strain exhibits increased light production (restoration of bioluminescence) when grown in the presence of sublethal concentrations of genotoxic agents. The Mutatox test is considered a promising tool for detection of genotoxic compounds in environmental samples (66–68). Both the Microtox and Mutatox test have been used in tandem to assess the potential acute toxicity and genotoxicity, respectively, of lipophilic contaminants in aquatic ecosystems (69).

Another test that utilizes a luciferase reporter gene system is the Vitotox assay (Thermo Electron Corporation, USA). This assay, which measures both genotoxicity and cytotoxicity, is based on recombinant *Salmonella typhimurium* TA104 bacteria containing the complete *V. fischeri lux* gene cassette under the control of the *rec*N promoter (70).

Vollmer et al. (71) have also described the construction and initial characterization of bacterial biosensors for the detection of DNA damage. In their research they constructed plasmids in which the DNA damage-inducible promoters *rec*A, *uvr*A, and *alk*A from *Escherichia coli* were fused to the *V. fischeri lux*CDABE operon. They reported that the fusion of the *rec*A promoter to *lux*CDABE yielded the most sensitive response. This recombinant strain (*E.coli* DPD2794) was later employed by Min et al. (72) to detect the genotoxicity of various chemicals and was demonstrated to be capable of distinguishing distinct bacterial responses for direct and indirect DNA damaging agents.

FUTURE DEVELOPMENT OF BIOSENSORS AS RAPID AND ROBUST TOOLS FOR ENVIRONMENTAL MONITORING

There is a growing demand for sensitive and easy to use biological assays for the rapid and cost effective detection of environmentally relevant concentrations of pollutants in soil, sediment, and water ecosystems. Recombinant luminescent bacterial biosensors have recently emerged as promising tools in environmental monitoring for the sensitive and specific quantification of both toxicity and bioavailability. These biosensors possess several of the characteristics of an ideal bioassay system, in that they are relatively inexpensive to conduct, easy to use, ecologically relevant, and amenable for on-line or *in situ* monitoring. In addition, previous research has demonstrated that freeze-drying of luminescent bacterial biosensors has only moderate effects on the performance with respect to sensitivity or induction coefficients (49,59). The availability of bacterial biosensors in a freeze-dried state would negate the need and cost of continuous culturing while ensuring a clonal strain, making reagent-like usage a viable option. Integration of the biosensor constructs into the host chromosome and the creation of stable constructs, which can be maintained even under nonselective conditions, have also been demonstrated (61,73). Optimization of immobilization techniques will advance the exploitation of the commercial potential of these biosensors, by providing a stable and practical system for use in real-time pollutant detection (74–77). A novel automated continuous toxicity test system using a recombinant bioluminescent freshwater bacterium has recently been described (78), giving further promise to the routine utilization of these biosensor systems for real-time biomonitoring of water toxicity in the near future.

BIBLIOGRAPHY

1. Steinberg, S.M., Poziomek, E.J., Engelmann, W.H., and Rogers, K.R. (1995). A review of environmental applications of bioluminescence measurements. *Chemosphere* **30**(11): 2155–2197.
2. Ramanathan, S., Ensor, M., and Daunert, S. (1997). Bacterial biosensors for monitoring toxic metals. *Trends Biotechnol.* **15**: 500–506.
3. D'Souza, S.F. (2001). Microbial biosensors. *Biosensors Bioelectron.* **16**: 337–353.
4. Keane, A., Phoenix, P., Ghoshal, S., and Lau, P.C.K. (2002). Exposing culprit organic pollutants: a review. *J. Microbiol. Methods* **49**: 103–119.
5. Simpson, M.L. et al. (1998). Bioluminescent-bioreporter integrated circuits form novel whole-cell biosensors. *Trends Biotechnol.* **16**: 332–338.
6. Grayeski, M.L. (1987). Chemiluminescence analysis. *Anal. Chem.* **59**: 1243A–1256A.
7. Wilson, T. and Hastings, J.W. (1998). Bioluminescence. *Annu. Rev. Cell Dev. Biol.* **14**: 197–230.
8. Hastings, J.W. et al. (1985). Biochemistry and physiology of bioluminescent bacteria. *Adv. Microbial Physiol.* **26**: 235–291.
9. Stewart, G.S.A.B. and Williams, P. (1992). *Lux* genes and the applications of bacterial luminescence. *J. Gen. Microbiol.* **138**: 1289–1300.
10. Meighen, E.M. and Dunlap, P.V. (1993). Physiological, biochemical and genetic control of bacterial luminescence. *Adv. Microbiol. Physiol.* **34**: 1–67.
11. Lampinen, J., Virta, M., and Karp, M. (1995). Use of controlled luciferase expression to monitor chemicals affecting protein synthesis. *Appl. Environ. Microbiol.* **61**: 2981–2989.
12. Tom-Petersen, A., Hosbond, C., and Nybroe, O. (2001). Identification of copper-induced genes in *Pseudomonas fluorescens* and use of a reporter strain to monitor bioavailable copper in soil. *FEMS Microbiol. Ecol.* **38**: 59–67.
13. Burlage, R.S. (1997). Emerging technologies: bioreporters, biosensors, and microprobes. In: *Manual of Environmental Microbiology*. C.J. Hurst, G.R. Knudsen, M.J. McInerney, L.D. Stetzenbach, and M.V. Walter (Eds.). ASM Press, Washington, DC, pp. 115–123.
14. Kaiser, K.L.E. and Palabrica, V.S. (1991). *Photobacterium phosphoreum* toxicity data index. *Water Pollut. Res. J. Can.* **26**: 361–431.
15. Bulich, A.A., Greene, M.W., and Isenberg, D.L. (1981). Reliability of the bacterial luminescence assay for the determination of the toxicity of pure compounds and complex effluents. In: *Aquatic Toxicology and Hazard Assessment: Fourth Conference*. D.R. Branson and K.L. Dickson (Eds.). STP 737, American Society for Testing and Materials, Philadelphia, pp. 338–347.
16. Vasseur, P., Ferrard, J.F., Rast, C., and Larbaight, D. (1984). Luminescent marine bacteria in ecotoxicity screening tests of complex effluents. In: *Toxicity Screening Procedures Using Bacterial Systems*. D.L. Liu and B.J. Dutka (Eds.). Marcel Dekker, New York, pp. 22–36.
17. Pardos, M. et al. (1999). Acute toxicity assessment of Polish (waste) water with a microplate-based *Hydra attenuata* assay: a comparison with the Microtox® test. *Sci. Total Environ.* **243/244**: 141–148.
18. Jung, K. and Bitton, G. (1997). Use of Ceriofast™ for monitoring the toxicity of industrial effluents: comparison with the 48-h acute *Ceriodaphnia* toxicity test and Microtox®. *Environ. Toxicol. Chem.* **16**: 2264–2267.
19. Cheung, Y.H. et al. (1997). Assessment of sediment toxicity using different trophic organisms. *Arch. Environ. Contam. Toxicol.* **32**: 260–267.
20. Doherty, F.G. (2001). A review of the Microtox® toxicity test system for assessing the toxicity of sediments and soils. *Water Qual. Res. J. Can.* **36**: 475–518.
21. Guzzella, L. (1998). Comparison of test procedures for sediment toxicity evaluation with *Vibrio fischeri* bacteria. *Chemosphere* **37**: 2895–2909.
22. Lappalainen, J., Juvonen, R., Vaajasaari, K., and Karp, M. (1999). A new flash method for measuring the toxicity of solid and colored samples. *Chemosphere* **38**: 1069–1083.
23. Lappalainen, J.O. et al.(2000). Comparison of the total mercury content in sediment samples with a mercury sensor bacteria test and *Vibrio fischeri* toxicity test. *Environ. Toxicol. Chem.* **15**: 443–448.
24. la Farre, M. et al. (2001). Wastewater toxicity screening of non-ionic surfactants by Toxalert® and Microtox® bioluminescence inhibition assays. *Anal. Chim. Acta* **427**: 181–189.
25. Jennings, V.L.K., Rayner-Brandes, M.H., and Bird, D.J. (2001). Assessing chemical toxicity with the bioluminescent photobacterium (*Vibrio fischeri*): a comparison of three commercial systems. *Water Res.* **35**(14): 3448–3456.
26. Mowat, F.S. and Bundy, K.J. (2001). Correlation of field-measured toxicity with chemical concentration and pollutant availability. *Environ. Int.* **27**: 479–489.
27. Fernández-Alba, A.R., Guil, M.D.H., Lopez, G.D., and Chisti, Y. (2002). Comparative evaluation of the effects of pesticides in acute toxicity luminescence bioassays. *Anal. Chim. Acta* **451**: 195–202.
28. Wang, C. et al. (2003). Ecotoxicological and chemical characterisation of selected treatment process effluents of municipal sewage treatment plant. *Ecotoxicol. Environ. Safety* **56**: 211–217.
29. Ulitzur, S., Lahav, T., and Ulitzur, N. (2002). A novel and sensitive test for rapid determination of water toxicity. *Environ. Toxicol.* **17**(3): 291–296.
30. Paton, G.I., Rattray, E.A.S., Campbell, C.D., Cresser, M.S., Glover, L.A., Meussen, J.C.L., and Killham, K. (1997). Use of genetically modified microbial biosensors for soil ecotoxicity testing. In: *Biological Indicators of Soil Health*. C.E. Pankhurst, B.M. Doube, and V.V.S.R. Gupta (Eds.). CAB International, Wallingford, UK, pp. 397–418.
31. Boyd, E.M., Meharg, A.A., Wright, J., and Killham, K. (1998). Toxicity of chlorobenzenes to a *lux*-marked terrestrial bacterium, *Pseudomonas fluorescens*. *Environ. Toxicol. Chem.* **17**: 2134–2140.
32. Paton, G.I., Campbell, C.D., Glover, L.A., and Killham, K. (1995). Assessment of bioavailability of heavy metals using lux modified constructs of *Pseudomonas fluorescens*. *Lett. Appl. Microbiol.* **20**: 52–56.
33. Layton, A.C., Gregory, B., Schultz, T.W., and Sayler, G.S. (1999). Validation of genetically engineered bioluminescent surfactant resistant bacteria as toxicity assessment tools. *Ecotoxicol. Environ. Safety* **43**: 222–228.
34. Chaudri, A.M. et al. (2000). Response of a *Rhizobium*-based luminescence biosensor to Zn and Cu in soil solutions from sewage sludge treated soils. *Soil Biol. Biochem.* **32**: 383–388.
35. Bundy J.G., Campbell, C.D., and Paton, G.I. (2001). Comparison of response of six different luminescent bioassays to

bioremediation of five contrasting oils. *J. Environ. Monitor.* **3**: 404–410.

36. Ivask, A. et al. (2004). Recombinant luminescent bacterial sensors for the measurement of bioavailability of cadmium and lead in soils polluted by metal smelters. *Chemosphere* **55**: 147–156.
37. Sinclair, G.M., Paton, G.I., Meharg, A.A., and Killham, K. (1999). *Lux*-biosensor assessment of pH effects on microbial sorption and toxicity of chlorophenols. *FEMS Microbiol. Lett.* **174**: 273–278.
38. Strachan, G. et al. (2001). Use of bacterial biosensors to interpret the toxicity and mixture toxicity of herbicides in freshwater. *Water Res.* **35**: 3490–3495.
39. Turner, N.L. et al. (2001). A novel toxicity fingerprinting method for pollutant identification with *lux*-marked biosensors. *Environ. Toxicol. Chem.* **20**: 2456–2461.
40. Tauriainen, S., Virta, M., Chang, W., and Karp, M. (1999). Measurement of firefly luciferase reporter gene activity from cells and lysates using *Escherichia coli* arsenite and mercury sensors. *Anal. Biochem.* **272**: 191–198.
41. Bundy, J.G., Paton, G.I., and Campbell, C.D. (2004). Combined microbial community level and single species biosensor responses to monitor recovery of oil polluted soil. *Soil Biol. Biochem.* **36**: 1149–1159.
42. Paton, G.I. et al. (1995). Use of luminescence-marked bacteria to assess copper bioavailability in malt whiskey distillery effluent. *Chemosphere* **31**: 3217–3224.
43. Brown, J.S. et al. (1996). Comparative assessment of the toxicity of a papermill effluent by respirometry and a luminescence-based bacterial assay. *Chemosphere* **32**: 1553–1561.
44. Boyd, E.M. et al. (1997). Toxicity assessment of xenobiotic contaminated groundwater using *lux* modified *Pseudomonas fluorescens*. *Chemosphere* **35**: 1967–1985.
45. Boyd, E.M., Meharg, A.A., Wright, J., and Killham, K. (1997). Assessment of toxicological interactions of benzene and its primary degradation products (catechol and phenol) using a *lux*-modified bacterial bioassay. *Environ. Toxicol. Chem.* **16**: 849–856.
46. McGrath, S.P. et al. (1999). Assessment of the toxicity of metals in soils amended with sewage sludge using a chemical speciation technique and a *lux*-based biosensor. *Environ. Toxicol. Chem.* **18**: 659–663.
47. Petänen, T. et al. (2003). Assessing sediment toxicity and arsenite concentration with bacterial and traditional methods. *Environ. Pollut.* **122**: 407–415.
48. Layton, A.C., Muccini, M., Ghosh, M.M., and Sayler, G.S. (1998). Construction of a bioluminescent reporter strain to detect polychlorinated biphenyls. *Appl. Environ. Microbiol.* **64**: 5023–5026.
49. Tauriainen, S., Karp, M., Chang, W., and Virta, M. (1998). Luminescent bacterial sensor for cadmium and lead. *Biosensors Bioelectron.* **13**: 931–938.
50. Gu, M.B. and Gil, G.C. (2001). A multi-channel continuous toxicity monitoring system using recombinant bioluminescent bacteria for classification of toxicity. *Biosensors Bioelectron.* **16**: 661–666.
51. Wiles, S., Whiteley, A.S., Philp, J.C., and Bailey, M.J. (2003). Development of bespoke bioluminescent reporters with the potential for *in situ* deployment within a phenolic-remediating wastewater treatment system. *J. Microbiol. Methods* **55**: 667–677.
52. Ren, S. and Frymier, P.D. (2005). Toxicity of metals and organic chemicals evaluated with bioluminescence assays. *Chemosphere* (in press).
53. Ren, S. (2004). Assessing wastewater toxicity to activated sludge: recent research and developments. *Environ. Int.* **30**: 1151–1164.
54. Lewis, J.C. et al. (1998). Applications of reporter genes. *Anal. Chem.* **70**: 579A–585A.
55. Bontidean, I. et al. (2004). Biosensors for detection of mercury in contaminated soils. *Environ. Pollut.* **131**: 255–262.
56. Petänen, T. and Romantschuk, M. (2003). Toxicity and bioavailability to bacteria of particle-associated arsenite and mercury. *Chemosphere* **50**: 409–413.
57. Turpeinen, R., Virta, M., and Häggblom, M.M. (2003). Analysis of arsenic bioavailability in contaminated soils. *Environ. Toxicol. Chem.* **22**: 1–6.
58. Virta, M., Lampinen, J., and Karp, M. (1995). A luminescence-based mercury biosensor. *Anal. Chem.* **67**: 667–669.
59. Ivask, A., Virta, M., and Kahru, A. (2002). Construction and use of specific luminescent recombinant bacterial sensors for the assessment of bioavailable fraction of cadmium, zinc, mercury and chromium in the soil. *Soil Biol. Biochem.* **34**: 1439–1447.
60. Ivask, A., Hakkila, K., and Virta, M. (2001). Detection of organomercurials with whole-cell bacterial sensors. *Anal. Chem.* **73**: 5168–5171.
61. Hansen, L.H. and Sørensen, S.J. (2000). Versatile biosensor vectors for detection and quantification of mercury. *FEMS Microbiol. Lett.* **193**: 123–127.
62. Tauriainen, S., Karp, M., Chang, W., and Virta, M. (1997). Recombinant luminescent bacteria for measuring bioavailable arsenite and antimonite. *Appl. Environ. Microbiol.* **63**: 4456–4461.
63. Petänen, T., Virta, M., Karp, M., and Romantschuk, M. (2001). Construction and use of broad host range mercury and arsenite sensor plasmids in the soil bacterium *Pseudomonas fluorescens* OS8. *Microbial Ecol.* **40**: 360–368.
64. Petänen, T. and Romantschuk, M. (2002). Use of bacterial sensors as an alternative method for measuring heavy metals in soil extracts. *Anal. Chim. Acta* **456**: 55–61.
65. Tibazarwa, C. et al. (2001). A microbial biosensor to predict bioavailable nickel in soil and its transfer to plants. *Environ. Pollut.* **113**: 19–26.
66. Thomas, K.V. et al. (2002). Characterisation of potentially genotoxic compounds in sediments collected from United Kingdom estuaries. *Chemosphere* **49**: 247–258.
67. Isidori, M. et al. (2003). *In situ* monitoring of urban air in Southern Italy with the tradescantia micronucleus bioassay and semipermeable membrane devices (SPMDs). *Chemosphere* **52**: 121–126.
68. Isidori, M., Lavorgna, M., Nardelli, A., and Parrella, A. (2004). Integrated environmental assessment of Volturno river in South Italy. *Sci. Total Environ.* **327**: 123–124.
69. Johnson, B.T. and Long, E.R. (1998). Rapid toxicity assessment of sediments from estuarine ecosystems: a new tandem *in vitro* testing approach. *Environ. Toxicol. Chem.* **17**: 1099–1106.
70. van der Lelie, D. et al. (1997). The Vitotox® test, an SOS bioluminescence *Salmonella typhimurium* test to measure genotoxicity kinetics. *Mutat. Res.* **389**: 279–290.
71. Vollmer, A.C. et al. (1997). Detection of DNA damage by use of *Escherichia coli* carrying *rec*A′::*lux*, *uvr*A′::*lux*, or *alk*A′::*lux* reporter plasmids. *Appl. Environ. Microbiol.* **63**: 2566–2571.

72. Min, J., Kim, E.J, LaRossa, R.A., and Gu, M.B. (1999). Distinct responses of a *rec*A::*lux*CDABE *Escherichia coli* strain to direct and indirect DNA damaging agents. *Mutat. Res.* **442**: 61–68.
73. Hay, A.G., Rice, J.F., Applegate, B.M., Bright, N.G., and Sayler, G.S. (2000). A bioluminescent whole-cell reporter for detection of 2,4-dichlorophenoxyacetic acid and 2,4-dichlorophenol in soil. *Appl. Environ. Microbiol.* **66**: 4589–4594.
74. Corbisier, P. et al. (1999). Whole cell- and protein-based biosensors for the detection of bioavailable heavy metals in environmental samples. *Anal. Chim. Acta* **387**: 235–244.
75. Tauber, M., Rosen, R., and Belkin, S. (2001). Whole cell detection of halogenated organic acids. *Talanta* **55**: 959–964.
76. Sagi, E. et al. (2003). Fluorescence and bioluminescence reporter functions in genetically modified bacterial sensor strains. *Sensors Actuators* **B 90**: 2–8.
77. Timur, S., Pazarlioğlu, N., Pilloton, R., and Telefoncu, A. (2003). Detection of phenolic compounds by thick film sensors based on *Pseudomonas putida*. *Talanta* **61**: 87–93.
78. Cho, J.-C. et al. (2004). A novel continuous toxicity test system using a luminously modified freshwater bacterium. *Biosensors Bioelectron.* **20**: 338–344.

DEVELOPMENT AND APPLICATION OF SEDIMENT TOXICITY TEST FOR REGULATORY PURPOSES

Thomas Ángel del Valls Casillas
University of Cadiz
Puerto Real, Spain

INTRODUCTION

Since 1950 there is a growing concern about sediment contamination, and international conventions have been established for the protection of the marine environment (i.e., GIPME, 2000; London Convention, www.Londonconvention.org). Each country has developed particular guidelines for the protection of the environment but even these recommendations have no legislative force; they have been routinely used for characterizing contaminated sediments and for practical purposes such as dredging.

The first approaches to this issue were centered on sediment chemistry and thus the classification was made by comparing the concentration of determined contaminants of concern and sediment quality values established to effectively protect the aquatic environment (1). All these approaches follow the same schema and the same principles starting with an evaluation of the existing information. Contaminants of concern should be evaluated according to the information regarding physical, chemical, and biological properties of the sediments. The principal advantage of using this information is the cost effectiveness of the data collection and, if the data are consistent, decisions can be made to protect the environment without any further assessment.

For those sediments where existing information is not enough for decision-making, further assessments are needed. Traditionally, further assessment included the analysis of a list of contaminants, provided by regulators and assessors, that are to be considered in the evaluation. Concentrations are corrected for the particular characteristics and compared to sediment quality values defined as limit values to classify the sediments. These types of chemical guidelines are used worldwide and can be divided into two different categories depending on how they are derived: the empirical approach is based on statistical analyses of co-occurring chemical and biological effect data (2,3), while the mechanistic approach is based on calculation of equilibrium partitioning between a bound phase and a dissociated phase within pore water (4). Other guidelines use only a measure of the natural background condition or the contaminant concentration that satisfies the different concerns. The use of each one is under debate and new approaches are under development to meet existing regulatory requirements (5,6). One general limitation that applies for all matrices using numerical sediment quality guidelines is that they only address concerns regarding those chemicals for which guidelines have been developed. In addition, the actual availability and thus the effects depend on other factors such as the grain size distribution, composition of the organic matter, or concentration of sulfides in the sediment.

Since the mid-1970s research has focused on the development of sediment toxicity bioassays that could effectively predict potential biological adverse effects and could be used as an effective tool to characterize sediment toxicity. It is widely accepted that there are no chemical measurements that reliably predict sediment toxicity. In addition, sediments are a complex matrix that require complex tests for estimating the actual bioavailability of contaminants (7). Here we review the evaluation process that has been applied in the different initiatives carried out to introduce this new approach in a regulatory context. The design of toxicity tests for the characterization of dredged material in Spain is used as an example.

TEST DEVELOPMENT FOR USE IN REGULATORY PURPOSES

Once the new approach was established and widely accepted, much research was done regarding the selection of the proper battery of tests to include for regulatory purposes. The suitability of a toxicity test for decision-making is determined by the development degree of these tests through research (under academic or private initiatives) and the quantity and quality of the available scientific information. Involvement of the regulatory agencies is fundamental to evaluate the usefulness of the resulting information; they are the entities that should address the uncertainties associated with the different analytical tools used for the decision-making process. Before a bioassay can be included for regulatory purposes, some basic characteristics should be considered. Here we include some of the factors that should be taken into account when evaluating a bioassay for regulatory purposes:

Selection of Organisms

Test species should be autochthon and ecologically relevant in the areas of concern. This extends to those species widely distributed and adapted to the regional aquatic ecosystems naturally or by means of cultures. Their importance in the trophic chain should be established in relation to the geographical zone where the bioassay is going to be applied. It is preferable that they are going to be directly affected by contaminated sediments. Test species are more valuable if they are tolerant over a wide range of environmental variables (i.e., pH, salinity, grain size distribution, or volatile sulfides) and easy to maintain in the laboratory or purchased from commercial cultures. Information regarding the life cycle of the test species, information about its growth and reproduction, and information about the test species' sensitivity for the contaminants of concern are valuable tools for the final selection of the proper test organism. Easy taxonomical identification of the species must be taken into account. For example, the manila clam can be selected as a test organism for sediment toxicity assessment instead of *Macoma nasuta* because although it is not an autochthon species, it is widely cultured and meets all the criteria for its selection.

Selection of Endpoints

Criteria for selecting endpoints in environmental risk assessment processes (ERAs) include susceptibility to the stressor, ease of measurement, unambiguous definitions, and societal and ecological relevance (8). To link sediment quality with ERAs, similar criteria should be employed to select endpoints to assess sediment quality (9). The selection of the endpoint measured in the test must be related to a receptor, process, or effect of concern. The design of the test should pose and answer some questions of the decision-making process and its use should be justified in the followed assessment framework. As stated by Chapman et al. (10), the primary response is mortality, followed by reproductive or growth effects, and finally other sublethal endpoints such as behavior or biomarkers. However, these last ones only provide an indication of exposure and have not yet been linked directly to impacts at the organism level except using histopathology (11). The recommended tests measure mortality and reproductive or growth effects and use ecologically significant taxa similar to or related to resident taxa, which are likely to be exposed and which are appropriately sensitive. Even taking into account all these factors for the selection of the best measurable endpoint, sediment toxicity tests for predicting the effects of contaminants on natural populations, communities, and ecosystems are criticized because they do not take indirect effects into account.

Experimental Design

Two main conditions must be met to properly develop a laboratory toxicity test for regulatory purposes: the test should be designed with a specific goal for the decision-making process and the test must be consistently applied, with its experimental conditions perfectly defined, and should account for the influence of the experimental conditions on test organisms. Standard protocols, including technical reports from government or private organizations, or those endorsed by a standardization group [e.g., American Society of Testing and Materials (ASTM) or Environment Canada (EC)], or published papers in peer-reviewed journals, including interpretive guidance, are available for different species [e.g., amphipods (12,13) or polychaetes (14)]. Another factor affecting the experimental design and application of a test is adherence to standards to ensure the good health of test organisms and to limit the total costs of conducting the test. For example, the commercial bioassay Microtox® has been widely recommended and applied. As a commercial test it is highly standardized after a long period of research for its characterization. Although modifications to the initial protocols have been developed, the oldest design is recommended for use until the new one has been standardized and characterized.

Ruggedness

Test ruggedness is related to the sensitivity of test methods to variations in laboratory conditions such as the feeding regime, light and photoperiod, or grain size distribution. It is important to know the precision in defining features and to establish good quality assurance/quality control guidance.

Ring Tests

Standard methods are required in regulatory programs. In addition to standardization, other important prerequisites are good reproducibility, acceptable interlaboratory variability (defined as multilaboratory precision), and repeatability and intralaboratory precision. While intralaboratory precision reflects the ability of trained laboratory personnel to obtain consistent results repeatedly when performing the same test on the same organisms using the same toxicant (repeatability can include variability related to different operators), interlaboratory precision (also referred to as round-robin or ring tests) is a measure of how reproducible the method is when conducted by a large number of laboratories using the same method, organism, and samples (13). This information is a measure of confidence in the consistency of test performance and application.

Different studies have been successfully completed to evaluate sediment toxicity bioassays (including whole sediment and pore water or elutriate), for example, the whole sediment toxicity test using amphipods (15,16), the bioassay of inhibition of bioluminescence (17,18), and the liquid phase bioassay using sea urchin embryos (17,19).

Validation of the Tests

The validation process of a test can be accomplished through studies to find relationships between sediment toxicity and contaminant concentrations, using a different range of contamination, a different range of sediment conditions, or by means of comparison with other standardized tests. Test validation focuses on comparing data of field impacts to ensure the predictive value of such tools.

Data Management Procedures

When test results are accepted, they must be easily interpreted. The role of the test in the regulatory process is to answer the assessment questions formulated during the sample evaluation. Different approximations have been made to classify sediments according to the biological effects produced. In some countries and for some perspectives, strict guidelines have been established, although it seems more desirable to integrate multiple lines of evidence (terms that define categories of information) and to individuate in this way the process for each particular situation. This approach integrates different categories of information, for example, chemical, toxicological, and benthic community structure. The way in which this information is integrated is going to depend on the project itself and on the information available for each line of evidence.

APPLICATION OF SEDIMENT TOXICITY TEST FOR REGULATORY PURPOSES

In Fig. 1 a general tiered testing scheme to be followed when characterizing contaminated sediments and dredged materials is included as an example of the multiple approaches available. The first tier is accomplished if the available information, including geographical, geological, and pollutant sources, ascertains that the sediment contamination has no effects. If we find fine sediments and important sources of pollutants that result in high concentrations of contaminants, the characterization proceeds to the next tier and information regarding sediment chemistry and the characterization of other parameters, such as microbiology, should be accomplished. Some variables involved in the process are sampling, handling, and storage of the sediment, which influence the bioavailability of contaminants. Each regulatory agency should establish standard procedures that can guarantee the accuracy of the results. The regulatory agencies should also establish the parameters to be characterized, which normally include metals and metalloids, PCBs, PAHs, pesticides, and other compounds such as TBTs. Then the results are compared to the selected sediment quality guidelines for a single species. Such sediment chemical benchmarks have been used as rough limit values or regulatory levels. Depending on how the guidelines were derived (if they were intended to predict toxicity or the absence of toxicity) and what effects were to be assessed, one can judge if the regulatory requirements are met.

For the next tier, a secondary assessment regarding biological tests is used for the characterization. Three different categories can be distinguished according to the exposure routes: tests addressed to measure direct water column effects, those addressed to measure direct effects on benthic organisms, and those that measure the bioaccumulation effect. Because potential effects should include all possible exposure routes, water column effects are characterized by means of liquid phase bioassays and benthic effects by using benthic organisms exposed to the whole sediment. Differences are found when bioaccumulation tests are used. Some countries have included these last bioassays (i.e., Canada) as an important part of the battery of tests. Other countries only recommend them when contaminants accumulated by organisms are present and thus apply these more cost-effective tools to the particular cases where high concentrations of determined substances are present.

Some regulatory agencies have included a list of suitable or recommended bioassays but the final set of tools used for the characterization process is open. Others just include a list of recommendations and factors in selecting the proper battery of tests. The wise use of sediment bioassays is to design a battery of tests that will give a wide range of effects. A good example is the management of dredged materials in Canada (20). In this country, test methods currently used include an acute toxicity test using amphipods, a fertilization assay using echinoids, a solid phase toxicity test using the photoluminescent bacteria, and a bedded sediment bioaccumulation test using bivalves.

During the last decades different sediment toxicity tests have been carried out in Spain as part of a weight of evidence approach using the sediment quality triad (21). These studies used basically acute sediment toxicity tests that covered organisms from a bacterial population to juveniles of commercial species of fish (Table 1). Today, some of these acute toxicity tests shown in the table have been adapted for use in characterizing dredged material quality assessment in Spanish ports and are being considered for incorporation in the new recommendations for the management of dredged material in ports of Spain (RMDM). These tests include the next types and bioassays: screening, the Microtox STP, whole sediment, the amphipod, the polychaeta, the irregular sea urchin tests, and, using elutriates, the sea urchin larvae test (Table 1).

The design and selection of the different tests was carried out under a joint contract between CEDEX and the University of Cádiz and after a practical phase to validate the potential use of standardized tests by other countries but using autochthonous species of organisms in Spain. During this practical phase, different ports were sampled and analyzed using chemical and ecotoxicological tests. Results showed that the mentioned bioassays can be used to characterize dredged material in Spain and some biological guidelines or criteria were developed. These guidelines are similar to others designed and applied for the same purposes by different regulatory agencies, such as the U.S Environmental Protection Agency, Environment Canada, the Dutch Ministry for Commerce and Navigation, and Environment Australia (Table 2). Biological guidelines derived for Spain are shown in Table 3. These values are similar to those proposed by other agencies, although it should be noted that there is a double criteria for classification of toxicity in one sample based on an absolute value together with the demonstration of significant differences by means of statistical tools using a control or reference sediment.

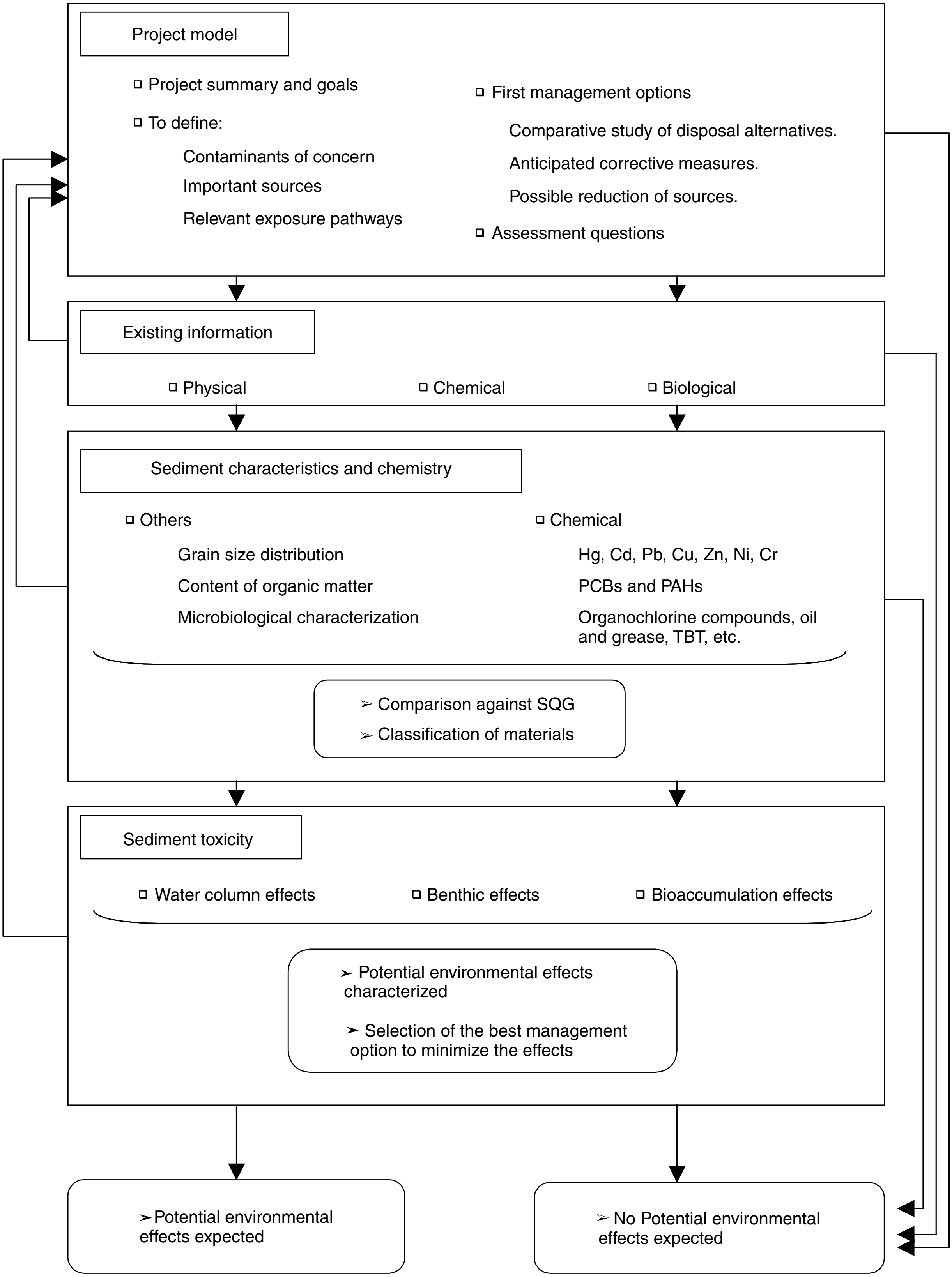

Figure 1. Tiered testing scheme to characterize contaminated sediments and dredged material. It has been adapted from different examples (Basque country, The Netherlands, Canada, or the United States).

Table 1. List of Different Bioassays Checked for Potential Uses in the Battery of Sediment Toxicity Tests Routinely Used to Characterize Dredged Material Disposal Options in Spain

Organisms	Exposure Route	Duration	Temperature, °C	Endpoints	Checked[a]	Acceptable[b]	References
Microtox® (SPT) (*Vibrio fischeri)*	Screening	15 min to 1 h	15	Luminescence inhibition	Yes	Yes	23
Echinoderms (*P. lividus*)	Elutriates	48 h	20	Survival (%)	Yes	Yes	23,24
Rotifer population (*B. plicatilis)*	Elutriates	7d	25	Survival (%)	Yes	?	25
Amphipod (*Corophium volutator, A. brevicornis*)	Whole sediment	10 d	20	Survival (%)	Yes	Yes	21,23
Benthic microalgae (*Cylindrotheca closterium*)	Whole sediment	24–72 h	20	Growth inhibition	0	?	26
Echinoderm (*Echinocardium cordatum*)	Whole sediment	14 d	15	Survival (%)	Yes	Yes	23,24
Polychaetae (*Arenicola marina*)	Whole sediment	10 d	15	Survival (%), bioaccumulation	Yes	Yes	14,23
					Yes	?	
Bivalve (*Ruditapes philippinarum*)	Whole sediment	48 h to 15 d	20	Survival (%); burrowing (TB_{50}) bioaccumulation; histological damage; biomarkers; vitellogenin	Yes	?	27
Crab (*Carcinus maenas*)	Whole sediment	21 d to 3 m	15	Survival (%); burrowing (TB_{50}) Bioaccumulation; histological damage; biomarkers; vitellogenin	Yes/?	Yes/?	28
Fish: benthic and pelagic (*Solea senegalensis; Sparus aurata*)	Whole sediment	1–3 m	20	Survival; bioaccumulation; histological damage; biomarkers	Yes	Yes/?	11

[a] Yes = standardized; ? = in development; 0 = not checked.

[b] Yes means useful for dredged material characterization.

Source: Adapted from Reference 22.

Another important aspect is the use of autochthonous species from Spanish coasts.

FINAL REMARKS AND FUTURE RESEARCH

Research and development are focused on reducing the uncertainties and costs associated with assessing dredged sediment material. Both the number and complexity of environmental questions being posed to port and dredged material managers have steadily increased in recent years, and this trend will likely continue in the future. Ongoing research is being used to provide answers to these questions and to address sources of uncertainty in the assessment and decision-making process. Other confounding factors already discussed must be taken into account when selecting the battery of tests to be applied for characterizing toxicity. Although the effect of some of these factors can be estimated and corrected using a proper reference sediment, others are difficult and the lack of control on some of these variables can have serious implications if test results are used for regulatory purposes.

Acknowledgments
This work has been partially funded by grants from the Spanish Ministry of Development (BOE 13-12-02) and of Education and Science (REN2002_01699/TECNO and VEM2003_-20563/INTER). The design of the acute toxicity tests for dredged material characterization was conducted under joint contract between CEDEX and the University of Cádiz (2001 and 2003).

BIBLIOGRAPHY

1. CEDEX (1994). *Recommendations for the Management of Dredged Material in the Ports of Spain*. Puertos del Estado, Madrid.
2. Long, E.R., Field, L.J., and MacDonald, D.D. (1998). *Environ. Toxicol. Chem.* **17**: 714–727.
3. Riba, I., Casado-Martínez, M.C., Forja, J.M., and DelValls, T.A. (2004). *Environ. Toxicol. Chem.* **23**: 271.

Table 2. Summarized Description of Different Bioassays Selected for Dredged Material Characterization in Spain[a]

Type	Test[b]	Country	Criteria	References
Screening	Microtox	The Netherlands	100 toxic units	17
		Canada	1000 mg · L^{-1} dw	29,30
		Australia	—	30
		Italy	Toxicity assigned according to STI units after correction for fine fraction	
Solid phase	Amphipods	The Netherlands	35%	17
		Canada	+20% mortality compared to reference sediment or +30% compared to control sediment and statistical difference	20
		EEUU	+20% mortality compared to reference sediment or +30% compared to control sediment and statistical difference	31
		UK	40% mortality	32
		Italy	—	—
		Australia	—	30
	Benthic algae viability	Australia	—	30
Liquid phase	Sea urchin embryo development	Canada	+25% abnormality compared to control seawater and statistical difference	20
		Italy	—	—
	Sea urchin fertilization and larval development	Australia	—	30
	Bivalve larval development	Australia	—	30
	Tiger prawn survival (postlarvae)	Australia	—	30
	Algal growth inhibition test	Australia	—	30
Bioaccumulation	Bivalves	Canada	Statistical significant difference in tissue concentrations between test sediment and reference or control sediment	29

[a] The biological guidelines proposed for each test in different countries are also included and are used in the management of disposal options of these materials in each country.

[b] At least two different acute toxicity tests required, preferably three.

Table 3. Biological Guidelines for the Sediment Toxicity Tests Recommended in Spain to Be Part of the New Recommendations for the Management of Dredged Material (RMDM) in Ports of Spain that Are Under Discussion for Final Adoption by the Ministry of Development

Type	Test	Criteria	Comments
Screening	Microtox	1000 mg · L^{-1} dw	SPT protocol; correction for the fine fraction
Liquid phase	Sea urchin embryo development	+25% abnormality compared to control seawater and statistical difference	Test species: *Paracentrotus lividus*
Solid phase	Amphipods	+20% mortality compared to reference sediment or +30% compared to control sediment and statistical difference	*Corophium* sp. *Ampelisca* sp.

4. DiToro, D.M. et al. (1991). *Environ. Toxicol. Chem.* **10**: 1541.
5. Fairey, R. et al. (2001). *Environ. Toxicol. Chem.* **20**: 2276–2286.
6. McCauley, D.J., DeGraeve, G.M., and Linton, T.K. (2000). *Environ. Sci. Pollut.* **3**: 133–144.
7. O'Connor, T.P. and Paul, J.F. (1999). *Mar. Pollut. Bull.* **40**: 59–64.
8. Suter, G.W. (1993). *Ecological Risk Assessment*. Lewis, Chelsea, MI.
9. Clements, W.H. (1995). Ecological significance of endpoints used to assess sediment quality. In: *Ecological Risk Assessment of Contaminated Sediments*. C.G. Ingersoll, T. Dillon, and G.R. Biddinger (Eds.). SETAC Press, Pensacola, FL.
10. Chapman, P.M., Ho, K.T., Munns, W.R. Jr., Solomon, K., and Weinstein, M.P. (2002). *Mar. Pollut. Bull.* **44**: 271.
11. DelValls, T.A. et al. (1998). *Ecotoxicol. Environ. Safety* **41**: 157.
12. ASTM (1991). *Standard guide for conducting 10 day static sediment toxicity test with marine and estuarine amphipods*. In: *Annual Book of ASTM Standards*, Vol. 11.04, E 1367-90. American Society of Testing and Materials, Philadelphia, pp. 1052–1076.

13. U.S. EPA (1994). *Methods for assessing the toxicity of sediment-associated contaminants with estuarine and marine amphipods*. United States Environmental Protection Agency, EPA/600/R-94/025.
14. Thain, J.E. and Bifield, S.A. (1993). *A sediment bioassay using the polychaete Arenicola marina*. Test guideline for the PARCOM sediment reworker ring test. Unpublished report, MAFF Fisheries Laboratory, Burham-on-Crouch, UK.
15. Schlekat, C.E. et al. (1995). *Environ. Toxicol. Chem.* **14**: 2163.
16. Casado-Martínez, M.C. et al. Inter laboratory precision study of a whole sediment toxicity test using crustacean amphipods. *Cien. Mar.* (submitted).
17. Stronkhorst, J. (2003). *Ecotoxicological effects of Dutch harbour sediments*. The development of an effects-based assessment framework to regulate the disposal of dredged material in coastal waters of The Netherlands [Ph.D. thesis]. Vrije UniversiteitSuter.
18. Casado-Martínez, M.C. et al. Inter laboratory study of precision: the bioluminescence inhibition test for rapid sediment toxicity assessment. *Cien. Mar.* (submitted).
19. Casado-Martínez, M.C. et al. *Cien. Mar.* (submitted).
20. Environment Canada (2000). *Report On Biological Toxicity Tests Using Pollution Gradient Studies: Sydney Harbour*. Minister of Public Works and Government Service Report, Canada, EPS 3/AT/2.
21. DelValls, T.A., Forja, J.M., and Gómez-Parra, A. (1998). *Environ. Toxicol. Chem.* **17**: 1073.
22. DelValls, T.A., Casado.Martínez, M.C., and Buceta, J.L. (2003). Proposal of a tiered testing schema to characterize dredged material from Spanish ports. In: *Remediation of Contaminated Sediments—2003. Proceedings of the Second International Conference on Remediation of Contaminated Sediments*. M. Pellei and A. Porta (Eds.). Battelle Press, Columbus, OH.
23. DelValls, T.A. et al. (2001). *Investigación conjunta sobre la viabilidad de utilizar ensayos ecotoxicológicos para la evaluación de la calidad ambiental del material de dragado*, Technical Report for CEDEX, Puerto Real. Cádiz.
24. Casado-Martínez, M.C., DelValls, T.A., Fernández, N., Riba, I., and Buceta, J.L. (2004). *The use of acute sediment toxicity test to characterize dredged material disposal in Spanish Ports*. 4th SedNet Workshop Harmonization of Impact Assessment Tools for Sediment and Dredged Materials, San Sebastian, Spain 10–11 June.
25. DelValls, T.A., Lubián, L.M., Forja, J.M., and Gómez-Parra, A. (1997). *Environ. Toxicol. Chem.* **16**: 2323.
26. Moreno-Garrido, I., Hampel, M., Lubián, L.M., and Blasco, J. (2003). *Ecotoxicol. Environ. Safety* **54**: 290.
27. DelValls, T.A., Forja, J.M., and Gómez-Parra, A. (2002). *Chemosphere* **46**: 1033.
28. Martín-Díaz, M.L. (2004). *Determinación de la calidad ambiental de sistemas litorales y de estuario de la Península Ibérica utilizando ensayos de campo y laboratorio [Tesis doctoral]*, Universidad de Cádiz.
29. Environment Canada (2002). *Biological Test Method: Reference Method for Determining the Toxicity of Sediment Using Luminescent Bacteria in a Solid-Phase Test*. Report EPS 1/RM/42.
30. Environment Australia (2002). *National Ocean Disposal Guidelines for Dredged Material*. May.
31. U.S. EPA/U.S. ACE (1998). *Evaluation of Dredged Material Proposed for Discharge in Waters of the US. Testing manual* (The Inland Testing Manual). EPA-823-F-98-005.
32. ICES (2000). *Biological assessment of toxicity of marine dredged materials*. In: *Report of the ICES Working Group on Biological Effects of Contaminants*, Nantes, France, 27–31 March, Chap. 7.

E_h

Suduan Gao
USDA–ARS
Parlier, California

Kenneth K. Tanji
University of California
Davis, California

A reduction–oxidation (redox) reaction involves the transfer of electrons. Thus redox reactions change the oxidation state and affect the speciation, solubility, mobility, and toxicity of elements in the aquatic environment. Major elements involved in redox reactions are O, C, N, S, Mn, and Fe, and the major sensitive redox components include O_2/H_2O, $NO_3^-/(N_2, NH_4^+)$, Mn(III, IV)/(II), Fe(III)/Fe(II), SO_4^{2-}/H_2S, and CO_2/CH_4. In polluted waters, trace elements could include As(V)/(III), Cr(VI)/(III), Hg (II)/(0, −II), Se(VI)/(IV, 0, −II), and U(VI)/(IV). Strong oxidants such as chlorine and ozone are used in water treatment processes too. The theoretical aspects of redox potential (E_h) are based on chemical equilibrium. Redox reactions in nature are driven by chemical and, more importantly, biological processes toward equilibrium, but it is seldom achieved. The traditional method of E_h measurement in water is simple but it is difficult to interpret because redox systems are seldom at equilibrium and multiple redox components coexist. Here we review and summarize theoretical predictions, measurement, and the applications of E_h in aqueous chemistry.

WHAT IS E_h?

E_h is an intensity measure of the reducing or oxidizing conditions in a system, that is, the tendency of a solution to donate to or accept electrons from a chemical species or electrode into the solution (1). E_h is conventionally defined in terms of the potential of a cell composed of two half-reactions: the half-cell of particular interest and the standard hydrogen half-cell. The potential of the hydrogen half-cell $[H^+ + e^- = \frac{1}{2} H_2(g)]$ at standard-state conditions, that is, $(H^+) = 1.0$ M and $P_{H_2} = 1.0$ atm, is set to 0.00 V.

Any redox reaction in nature is completed by two simultaneous half-reactions involving reduction and oxidation of two different chemical species and electrons are transferred from one species (reductant) to another (oxidant). The most important redox reaction, for example, is the microbially mediated oxidation of organic matter (represented by CH_2O) by oxygen (O_2) in an oxic environment (where oxygen is present) (2):

$$CH_2O + O_2 = CO_2 + H_2O \qquad (1)$$

where CH_2O loses and O_2 accepts the electrons, although the number of electrons transferred do not show up in this overall reaction.

Any of the half-reactions in a redox reaction can be generalized as a reduction reaction:

$$\mathrm{Ox} + n\,\mathrm{e}^- + m\,\mathrm{H}^+ = \mathrm{Red} + \mathrm{H_2O} \qquad (2)$$

where Ox and Red represent the oxidized and the reduced species, respectively. Most reduction reactions consume protons while involved in transferring electrons (Table 1).

The reduction potential of Equation 2 is expressed by the Nernst equation,

$$E_\mathrm{h} = E_\mathrm{h}^0 - \frac{2.303RT}{nF}\log\frac{(\mathrm{Red})}{(\mathrm{Ox})(\mathrm{H}^+)^m} \qquad (3)$$

where E_h = reduction potential for the half-reaction (V), E_h^0 = standard-state reduction potential for the half-reaction (V), n = moles of electrons transferred in the reaction, m = moles of protons involved in the reaction, R = natural gas constant ($8.314\ \mathrm{J\cdot K^{-1}\cdot mol^{-1}}$), T = absolute temperature, F = Faraday constant ($96484.56\ \mathrm{C\cdot mol^{-1}}$), and (P) = activity. In dilute solutions, one may assume that activity approximately equals concentration.

The standard state is defined when all activities of redox species are at unity, that is, when $E_\mathrm{h} = E_\mathrm{h}^0$. At 25 °C, Equation 3 becomes

$$E_\mathrm{h} = E_\mathrm{h}^0 - \frac{0.05916}{n}\log\frac{(\mathrm{Red})}{(\mathrm{Ox})(\mathrm{H}^+)^m} \qquad (4)$$

or

$$E_\mathrm{h} = E_\mathrm{h}^0 - \frac{0.05916}{n}\left(m\mathrm{pH} + \log\frac{(\mathrm{Red})}{(\mathrm{Ox})}\right) \qquad (5)$$

Thus, the reduction potential is determined by the ratio of the redox couple and the pH. A high ratio of the reducing species to the oxidizing species yields a low E_h value, indicating the abundance of reducing species (an electron-rich system) and a high tendency of the solution to donate electrons. Thus, E_h is an intensity parameter representing the electron activity of a system.

The standard-state reduction potentials of half-reactions (E_h^0) encountered commonly in aquatic environments are listed in Table 1. The free energy change related to the reduction potential of the reaction is

$$\Delta G_r = -nFE_\mathrm{h} \qquad (6)$$

or

$$\Delta G_r{}^0 = -nFE_\mathrm{h}^0 \qquad (7)$$

The more positive E_h values indicate stronger tendency for the reaction to proceed for the reaction to occur from left to right as it relates to more free energy release (Eqs. 6 and 7). The redox couples with higher positive E_h^0 indicate stronger oxidants (e.g., O_2) and those with lower E_h^0 indicate stronger reductants (e.g., CH_4, H_2S), However, the standard-state conditions do not normally occur in natural systems, where activities of most dissolved redox species including protons are much below unity. The reduction potentials of the important reactions at pH = 7 and at 25 °C, $E_\mathrm{h}^0(\mathrm{w})$, are calculated using Equation 5, and are given in Table 1. These values indicate redox reactions may proceed sequentially in a closed system when electron acceptors are consumed from the strongest to the weakest. It implies that in natural aquatic systems, where organic matter serves as the most abundant reductant, oxidation of organic matter may be accompanied by the consumption of electron acceptors in the sequence of oxygen, nitrate and Mn oxides, Fe oxides, sulfate and carbonate reductions. In reality, these reductions often overlap because microbially mediated reactions can be slow.

In aquatic systems, water reduction to H_2 or oxidation to O_2 sets the boundaries and limits the range of E_h^0. E_h can exceed these boundaries if stronger oxidants than oxygen are added to water, such as chlorine and ozone that are used in drinking water and wastewater treatment. The most important redox couples in natural waters are O_2/H_2O, $NO_3{}^-/N_2$, Mn (III/IV)/Mn(II), Fe

Table 1. Selected Standard-State Reduction Potentials of Half-Reactions Involving Important Elements in Water, Sediments, and Soils

Reaction	E_h^0 (V)[a]	E_h^0(w) at pH = 7.0[b]	pe^0(w) at pH = 7.0[b]
$\frac{1}{4}O_2(g) + H^+ + e^- = \frac{1}{2}H_2O$	1.230	0.816	13.79
$\frac{1}{5}NO_3{}^- + 6/5\ H^+ + e^- = \frac{1}{10}N_2(g) + \frac{3}{5}H_2O$	1.248	0.751	12.70
$\frac{1}{8}NO_3{}^- + 5/4\ H^+ + e^- = \frac{1}{8}NH_4{}^+ + \frac{3}{8}H_2O$	0.881	0.363	6.14
$\frac{1}{2}MnO_2\ (s) + 2\ H^+ + e^- = \frac{1}{2}Mn^{2+} + H_2O$	1.230	0.579	9.79
$\gamma\text{-}MnOOH(s) + 3\ H^+ + e^- = Mn^{2+} + 2\ H_2O$	1.503	0.616	10.41
$Fe(OH)_3(s) + 3\ H^+ + e^- = Fe^{2+} + 3\ H_2O$	0.935	0.048	0.81
$FeOOH(s) + 3\ H^+ + e^- = Fe^{2+} + 2\ H_2O$	0.769	−0.118	−1.99
$\frac{1}{8}SO_4{}^{2-} + \frac{5}{4}H^+ + e^- = \frac{1}{8}H_2S + \frac{1}{2}H_2O$	0.308	−0.210	−3.54
$\frac{1}{8}SO_4{}^{2-} + \frac{9}{8}H^+ + e^- = \frac{1}{8}HS^- + \frac{1}{2}H_2O$	0.254	−0.212	−3.58
$\frac{1}{8}CO_2(g) + H^+ + e^- = \frac{1}{8}CH_4(g) + \frac{1}{4}H_2O$	0.172	−0.242	−4.09
$H^+ + e^- = \frac{1}{2}H_2(g)$	0.000	−0.414	−7.00

[a] Data are from Reference 3.

[b] Calculated based on the E_h^0 and agreeable with data in Stumm and Morgan (6). E_h^0(w) and pe^0(w) apply to the electron activity for unit activities of oxidant and reductant in neutral water (pH = 7.0) at 25 °C; except for the Fe and Mn couples that are based on Mn^{2+} or Fe^{2+} activity of 10^{-6} M because these values are more realistic than the unit activities of standard-state conditions.

(III)/Fe(II), $SO_4{}^{2-}$/HS^-, and CO_2/CH_4. Highly reducing conditions result in methanogenesis and fermentation. Multiple reaction paths result in an array of E_h values for some elements that depend on a particular biological or chemical reaction path (e.g., N, Mn, S, and C). For example, nitrate reduction to N_2 (denitrification) has a higher E_h^0 than reduction to ammonium ($NH_4{}^+$). Various forms of solid phase Mn and Fe oxyhydroxides result in varying reduction potentials.

USE OF pe

In lieu of E_h, electron activity of a system is defined as a pe analogous to pH: $\mathrm{pe} = -\log(\mathrm{e}^-)$.

From Equation 2, at equilibrium, we have

$$K = \frac{(\mathrm{Red})}{(\mathrm{Ox})(\mathrm{e}^-)^n(\mathrm{H}^+)^m} \tag{8}$$

$$\mathrm{e}^- = \frac{(\mathrm{Red})}{n(\mathrm{Ox})(\mathrm{H}^+)^m K} \tag{9}$$

Take the logarithmic form on both sides and rearrange:

$$\mathrm{pe} = \mathrm{pe}^0 - \frac{1}{n}\log\left[\frac{(\mathrm{Red})}{(\mathrm{Ox})(\mathrm{H}^+)^m}\right] \tag{10}$$

or

$$\mathrm{pe} = \mathrm{pe}^0 - \frac{1}{n}\left(m\mathrm{pH} + \log\left[\frac{(\mathrm{Red})}{(\mathrm{Ox})}\right]\right) \tag{11}$$

where

$$\mathrm{pe}^0 = \frac{1}{n}\log K^0 \tag{12}$$

Both Equations 10 and 4 indicate that the activity of electrons can be quantitatively expressed by and directly related to the distribution of redox species and pH. Use of E_h or pe is a matter of preference. There are some advantages in using pe to examine redox phenomenon in some aspects. For example, pe + pH is considered a true variable in representing redox status as it can be used to differentiate the regions of redox species stability (3). When dealing with measured potential, however, it is more straightforward to use E_h than pe; although the conversion shown below is straightforward.

Conversion between E_h and pe at 25 °C is simplified as

$$E_h = 0.05916\ \mathrm{pe}, \tag{13}$$

where E_h is in volts. E_h and pe values change in the same direction and higher values indicate electron-poor systems and lower values indicate electron-rich systems. Table 1 lists the correspondence between $\mathrm{pe}^0(\mathrm{w})$ and $E_h^0(\mathrm{w})$.

E_h MEASUREMENT

Redox potential (E_h) of a solution is the difference in electrical potential measured between a noble metal (e.g., Pt or Au) electrode (also called indicator electrode) that is sensitive to redox couples in solution and a reference electrode that has a fixed potential in reference to the standard hydrogen electrode (SHE). Both electrodes are immersed in the solution and the potential is read on a potentiometer (e.g., pH meter). The noble electrodes are inert and resistant to chemical reactions; that is, they are not involved in the redox reactions and function only to conduct current. Because a standard hydrogen reference electrode ($E_h^0 = 0$) is not convenient for practical use, reference electrodes such as Ag/AgCl or calomel (Hg/Hg_2Cl_2) are commonly used in E_h measurement. Combination ORP (oxidation–reduction potential) electrodes consisting of both the indicator and reference electrodes are commercially available. These electrodes employ a platinum element to measure the potential due to the ratio of two oxidation states or species of an element in solution and each contains its own reference electrode. A commonly used Pt electrode system (electrochemical cell) is illustrated (3) as follows:

Pt electrode **Solution**
Pt | Ox (aq) → Red (aq) |

Reference electrode
KCl (aq, saturated) | Hg(l) : Hg_2Cl_2(s)

In this system, the Pt electrode conducts the electrons from the reduction reaction. The calomel reference electrode consists of a gel composed of liquid Hg^0 and solid Hg_2Cl_2 bathed in a saturated KCl solution. An asbestos fiber connects the saturated KCl solution to the sample by allowing K^+ and Cl^- ions to diffuse from the reference electrode into the sample solution to maintain electrical neutrality. This connection is called a salt bridge or liquid junction and electrically connects the reference electrode to the sample solution. A small and unpredictable liquid junction potential is encountered due to the differences between K and Cl ion diffusion rates, which are affected by the types of ions and the ion concentrations. The junction potential can be minimized by using a concentrated salt bridge and can be corrected to some degree with calibration against a solution with a known redox potential (e.g., ZoBell solution).

ZoBell solution, which contains 0.1 M KCl and an equal amount (e.g., 3.33×10^{-3} M) of $K_4Fe(CN)_6$ and $K_3Fe(CN)_6$, is typically used for testing redox instruments. It is more stable than quinhydrone solution above 30 °C. Quinhydrone solution, however, is much easier to prepare and use. The potential of the ZoBell solution measured by a Pt electrode versus a saturated KCl, Ag/AgCl reference electrode is a linear function of temperature, up to 55 °C (4):

$$E_{h,\,\mathrm{ZoBell}} = 0.428 - 0.0022\,(T - 25), \tag{14}$$

where T is the solution temperature (in °C).

A solution E_h can be obtained from the measurements of the solution and the ZoBell solution by an ORP combination electrode (5):

$$E_{h,\,\mathrm{system}} = E_{\mathrm{observed}} + E_{\mathrm{reference}}$$

where E_{observed} is the sample potential relative to the reference electrode used and

$$E_{\mathrm{reference}} = E_{h,\,\mathrm{ZoBell/reference}} - E_{h,\,\mathrm{ZoBell/observed}}$$

Table 2. Potential of Pt Electrode Versus Selected Reference Electrodes at 25 °C in Standard Zobell Solution[a] (mV)

Calomel Hg/Hg_2Cl_2	Silver/Silver Chloride (Ag/AgCl)			Standard Hydrogen
Saturated KCl	1.00 M KCl	4.00 M KCl	Saturated KCl	
+183	+192	+228	+229	+428

[a] ZoBell solution contains 1.4080 g potassium ferrocyanide ($K_4Fe(CN)_6 \cdot 3H_2O$), 1.0975 g potassium ferricyanide ($K_3Fe(CN)_6$), and 7.4555 g potassium chloride (KCl) in 1 L. The solution should be stored in a dark plastic bottle in a refrigerator.
Source: Reference 5.

where $E_{h,\,ZoBell/reference}$ is the theoretical E_h of the reference electrode and ZoBell's solution, relative to SHE (Eq. 14 for a Pt electrode and a saturated KCl, Ag/AgCl system; others are listed in Table 2) and $E_{h,\,ZoBell/observed}$ is the observed potential of ZoBell's solution relative to the reference electrode.

Care must be taken in measurement of water sample E_h. Proper instrument calibration and electrode testing must be performed. Freshwater samples must be used. Minimizing exposure to the atmosphere during E_h reading is necessary for low E_h samples. Flow-through cells are recommended for use for groundwater or a system with dynamic flow. Detailed procedures for measurement of E_h are available in Clesceri (5) and Nordstrom and Wilde (7).

INTERPRETATION

The concept of E_h measurement is derived based on chemical equilibrium, while the E_h measured in water does not reflect this aspect because redox reactions are seldom at equilibrium in natural waters. Thermodynamics only determines if a reaction is possible. Most redox reactions are mediated by microorganisms with slow kinetics. Lack of equilibrium is particularly prevalent among the various redox couples. Furthermore, many redox species (e.g., O_2, N_2, $NO_3{}^-$, $SO_4{}^{2-}$, $HCO_3{}^-$, and CH_4) are not electroactive; that is, they do not readily take up or give off electrons at the surface of the Pt electrode to contribute to the potential meant to be measured. The only exception is the couple Fe(III)/Fe(II), which are electroactive at low pH and where dissolved Fe is in certain concentrations (e.g., 10^{-5} M). Therefore, the measured E_h using a Pt electrode does not necessarily reflect the distribution of redox couples or species. Because it is a measurement of potential, the Pt electrode responds to changes in pH and other potentials. Comparison among E_h measurements should be based on the same pH and other conditions. Due to the presence of multiredox couples and nonequilibrium status in water, the measured potential is a mixed potential from only those electroactive redox species. Thus, E_h normally does not provide a quantitative measurement of the electron activity of the system measured; that is, it is normally only qualitative. Some authors believe that the measured E_h has no valuable theoretical meaning.

Other factors contributing to the difficulty in using E_h measurement to reflect electron activity of a solution include low concentrations of redox species, which cannot provide enough electron transfers that are detectable by the Pt electrode. The value of junction potential may differ greatly from that in the redox buffer solutions used to calibrate the cell. The Pt electrode can become contaminated by oxides or other coatings; thus, proper cleaning and polishing of the probe are needed before use. Permanently installed probes can be fouled either in high dissolved O_2 solution or strongly reducing conditions. Oxygen can react with Pt to form $Pt(OH)_2$ on the electrode surface. $Pt(OH)_2$ is electroactive and develops a potential with Pt ($\frac{1}{2}Pt(OH)_2 + e^- + H^+ \rightarrow \frac{1}{2}Pt + H_2O$; $E_h^0 = 0.982$ V) and masks the electrode response from electron transport between redox-sensitive solution species (3). Poising of permanent electrodes was also observed in reducing soils (8). The surface of the electrodes became dull after immersing into soil reducing solution for a few days and unable to respond to increased E_h values when nitrate was introduced into the system, while temporary electrodes detected the increase. Thus, a cleaning procedure was necessary to recover the function of the electrodes.

As a result of the discussion above, a large discrepancy is often observed between measured E_h values and predicted E_h values based on the redox couples present in the solution. Figure 1 shows an example of the theoretical E_h values and those observed in soil solutions at neutral pH for some important redox couples (compiled from Refs. 1 and 3). The large discrepancy for the O_2–H_2O, $NO_3{}^-$–N_2, and Mn oxides–Mn^{2+} are partially attributed to the inert behavior of O_2, $NO_3{}^-$, and N_2 at the Pt electrode and inappropriate choice of MnO_2 as electron acceptor as various forms of Mn oxides are present in a number of oxidation states (II, III, and IV) in nature. Similar reasons can apply for the $SO_4{}^{2-}$–H_2S and CO_2–CH_4 couples. The predicted and measured E_h values are closer only for the $Fe(OH)_3$–Fe^{2+} couple. In fact, wider E_h ranges for the same redox reactions than those shown in Fig. 1 were reported under various conditions. This should not be surprising when considering all the factors influencing the E_h measurement as discussed above.

APPLICATION

Although there are practical problems and difficulties in interpretation, E_h is a parameter that can be easily obtained and can be useful for better understanding of water chemistry. A decrease of E_h showed correspondence to sequential reduction of nitrate, Mn, and Fe in soil incubation experiments (9). The critical E_h at which all of the $NO_3{}^-$ was reduced and Mn^{2+} appeared in solution was about 200 mV and the critical E_h at which Fe^{2+} appeared was 100 mV. The qualitative E_h measurement has been used to classify redox regions or zones into different categories in relation to specific redox reactions. For examples, the following zones in soils at pH 7 were defined: oxidized ($E_h > 414$ mV, O_2 present), moderately reduced (120–414 mV, nitrate reduction and Mn reduction), reduced (−120 to 120 mV,

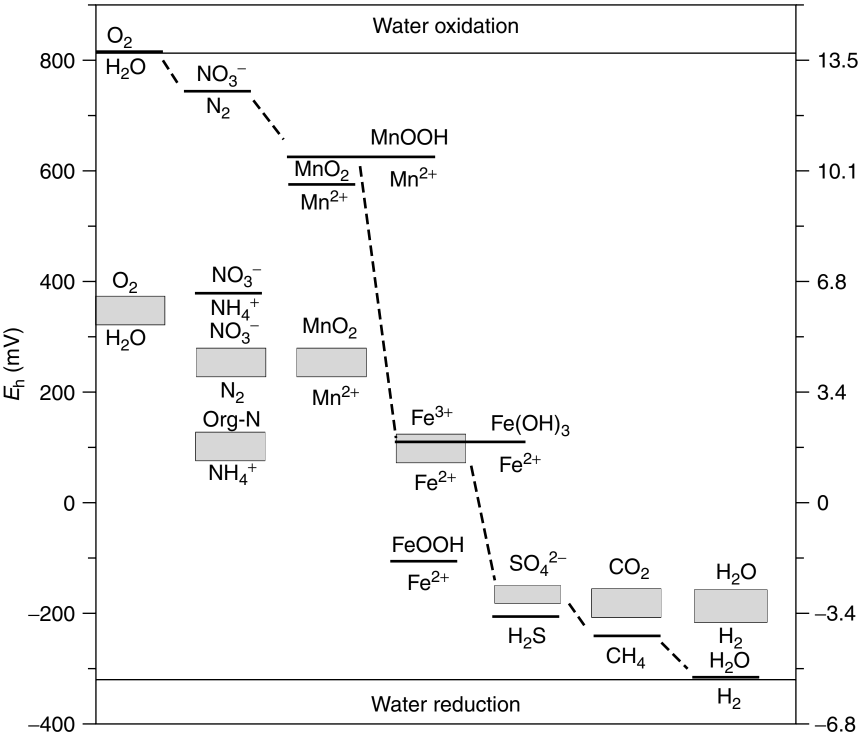

Figure 1. Theoretical and observed E_h for reduction sequences under neutral soil pH solutions. (Modified from Refs. 1 and 3.) The theoretical E_h values for the redox transformations to occur at pH 7.0 are indicated by the horizontal solid line. These values are based on equal activities of reduced and oxidized species unless noted. The activities of dissolved Fe^{2+} and Mn^{2+} are assumed at levels of 10^{-6} M and the pressure of H_2 is arbitrarily set at 10^{-3} atm. The observed E_h ranges for transformation from oxidized to reduced species are indicated by the shaded area. Corresponding pe values are shown on the right y axis.

Fe reduction), and highly reduced ($E_h < -120$ mV, sulfate reduction) (10). The following zones are further proposed at pH 7: oxic ($E_h > 414$ mV, oxygen and nitrate reduction), suboxic (120–414 mV, Mn and Fe oxide reductions), and anoxic ($E_h < 120$ mV, sulfate reduction) (11). Ranges of E_h for redox reactions may vary greatly from one observation to another. In fact, a wide range of E_h for the same redox couple has been reported from the literature (12). These variations originate from the differences among systems observed. Overlapping of redox processes is common especially in an organic matter-rich environment for a complex system such as paddy soil with straw amendment. E_h decrease corresponded to the progressing reducing environment developed upon flooding (13).

Redox buffering capacity (poising) in analogy to pH buffering capacity controls the E_h changes of a system. It is the capability of a system to maintain a redox potential upon addition of oxidizable or reducible materials. In a closed system with abundant organic matter or water near the organic matter-rich sediment (assuming originally equilibrated with oxygen in the atmosphere), oxygen will be quickly depleted due to oxidation of organic matter resulting in an abrupt drop in E_h (13). The E_h level may quickly decrease to and maintain the level indicative of sulfate reduction in most natural waters because nitrate concentrations are usually low and denitrification adds little buffering capacity to the water. In soils or groundwater in contact with oxidizing sediments, however, Mn and Fe oxides in the solid phase can add significant amount of buffering capacity to maintain a relatively higher E_h level for a period of time. The E_h ranges of these sequential redox reactions can be inferred from (Fig. 1).

In conclusion, E_h measurement in water is only a qualitative indication of the electron richness or poorness of a system. E_h measured in a solution is a mixed potential and usually does not quantitatively relate to specific redox reactions or redox couples. The easy measurement of E_h can have some practical use if its determination if handled carefully by following proper procedures (5,7). It can indicate the progressing of a system toward reduction or oxidation and can be used to predict potential redox reactions when comparing similar type systems. This qualitative measurement can assist in further quantitative determinations of redox reactions and chemical species. Nordstrom and Wilde (7) concluded that E_h data can be useful for gaining insights on the evolution of water chemistry and for estimating the equilibrium behavior of multivalent elements relative to pH for an aqueous system. E_h can delineate qualitatively strong redox gradients, such as in stratified lakes and rivers, in an oxidized surface flow that becomes anaerobic after passing through stagnant organic-rich systems, and in mine-drainage discharges.

OTHER APPROACHES

Many other alternative methods in describing and examining redox phenomenon of soil and water have been examined and proposed. A capacity parameter, oxidative capacity (OXC), was defined by Scott and Morgan (14). It integrates all the oxidized and reduced species into a single descriptive parameter. By setting an electron reference level, comparisons among oxidative capacities of various components results in identification of redox classes such as oxic (oxygen present), postoxic (nitrate, Mn, and Fe reduction), sulfidic (sulfate reduction), and methanogenesis. Challenges remain in quantifying the capacity of individual components especially those involving solid phase and organic matter. A nonequilibrium approach to defining the dominant terminal electron-accepting processes (TEAPs) has been applied to groundwater systems (15). This method considers simultaneously the consumption of electron acceptors [dissolved O_2, $NO_3{}^-$–N, Fe(III), $SO_4{}^{2-}$–S, and CO_2], final products [Mn(II), Fe(II), H_2S, and CH_4], and the intermediate product (dissolved hydrogen gas, H_2). A high degree of confidence can be achieved if a combination of all three indicators yields a positive identification of the predominant TEAPs. This is difficult to achieve sometimes because of the incorporation with the solid phase for some electron acceptors and final products. Dissolved H_2 is considered as a master variable in characterizing organic matter degradation or redox chemistry in anoxic conditions (16,17). Molecular hydrogen can be produced by many fermentative microorganisms during metabolism of organic compounds and at the same time consumed by respiratory microorganisms that use different electron acceptors. Molecular hydrogen is a result of the energy threshold for the organisms catalyzing the terminal electron-accepting process. The more electrochemically positive electron acceptors [e.g., nitrate and Mn(IV)] give a greater potential energy yield and reducers of these compounds can obtain their threshold energy yield at lower H_2 concentrations. Thus, a correlation between H_2 concentrations and predominant TEAPs was applied (12,15).

BIBLIOGRAPHY

1. McBride M.B. (1994). *Environmental Chemistry of Soils*. Oxford University Press, New York.
2. Manahan, S.E. (1994). *Environmental Chemistry*, 6th Edn. CRC Press, Boca Raton, FL.
3. Essington, M.E. (2004). *Soil and Water Chemistry—An Interactive Approach*. CRC Press, Boca Raton, FL.
4. Wood, W.W. (1976). Guidelines for collection and field analysis of ground water samples for selected unstable constituents. In: *Techniques of Water-Resources Investigations of the United States Gelological Survey, Book 1*. U.S. Geological Survey, Washington, DC, Chap. D2.
5. Clesceri, L.S., Greenberg, A.E., and Eaton, A.D. (1998). *Standard Methods for Examination of Water and Waste Water: Including Bottom Sediments and Sludges*, 20th Edn. American Public Health Association, Washington, DC.
6. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry*. John Wiley & sons, Hoboken, NJ.
7. Nordstrom, D.K. and Wilde, F.D. (1998). Reduction–oxidation potential (electrode method). In: Wilde, E.D. and Radtke, D.B. (Eds.), *National Field Manual for the Collection of Water-Quality Data*, U.S. Geological Survey Techniques of Water-Resources Investigations, Book 9, Chap. A6, Sec. 6.5. Available at http://water.usgs.gov/owq/FieldManual/Chapter6/6.5_contents.html.
8. Bailey, L.D. and Beauchamp, E.G. (1971). Nitrate reduction and redox potential with permanently and temporarily placed platinum electrodes in saturated soils. *Can J. Soil Sci.* **51**: 51–58.
9. Patrick, W.H. Jr. and Jugsujinda, A. (1992). Sequential reduction and oxidation of inorganic nitrogen, manganese, and iron in flooded soil. *Soil Sci. Soc. Am. J.* **56**: 1071–1073.
10. Sposito, G. (1981). *The Thermodynamics of Soil Solution*. Oxford University Press, New York.
11. Sposito, G. (1989). *The Chemistry of Soils*. Oxford University Press, New York.
12. Lovley, D.R. and Goodwin, S. (1988). Hydrogen concentrations as an indicator of the predominant terminal electron-accepting reactions in aquatic sediments. *Geochim. Cosmochim. Acta* **52**: 2993–3003.
13. Gao, S., Tanji, K.K., Scardaci, S.C., and Chow, A.T. (2002). Comparison of redox indicators in a paddy soil during rice-growing season. *Soil Sci. Soc. Am. J.* **66**: 805–817.
14. Scott, M.J. and Morgan, J.J. (1990). Energetics and conservative properties of redox systems. In: *Chemical Modeling of Aqueous Systems II*. X. Melchior and R.L. Barrett (Eds.). *CACS Symposium Series* 416. American Chemical Society, Washington, DC, Chap. 29, pp. 368–378.
15. Chapelle, F.H. et al. (1995). Deducing the distribution of terminal electron-accepting processes in hydrogeologically diverse groundwater systems. *Water Resour. Res.* **31**: 359–371.
16. Hoehler, T.M., Alperin, M.J., Albert, D.B., and Martens, C.S. (1998). Thermodynamic control on hydrogen concentrations in anoxic sediments. *Geochim. Cosmochim. Acta* **62**: 1745–1756.
17. Jakobsen, R. et al. (1998). H_2 concentrations in a landfill leachate plume (Grindsted, Denmark): *in situ* energetics of terminal electron acceptor processes. *Environ. Sci. Technol.* **32**: 2142–2148.

WATER RESOURCE DEVELOPMENT AND MANAGEMENT

WATER RESOURCES CHALLENGES IN THE ARAB WORLD

M.E. Young
Conwy, United Kingdom

THE PROBLEMS

The water shortage in the Arab World (Fig. 1) is severe and worsening. The region receives only 1% of the world's renewable water resources yet contains more than 5% of the world's population. In most Arab states at the start of the twenty-first century, the gross volumetric human and environmental demand for water exceeds natural replenishment. Per capita water availability is worsened by a high regional population growth; a 3% annual population growth rate halves the per capita availability every 24 years.

The paramount need, common to all states, therefore, is to find a sustainable balance between use and replenishment by a combination of measures to increase supply and to reduce demand. Both approaches involve technical (engineering) developments, legal and financial measures, improved water resources management through institutional interventions, and adjustments to national policies of water allocation.

All states need, as a priority, to develop coherent national water resources policies and strategies for the integrated management of water in all sectors—agricultural, municipal, industrial and environmental—in the context of the level of supply that is sustainable and according to sectoral priority. Institutional capacities to assess and manage water resources require development in many countries.

Many states of the region share resources from river flows or aquifers. Protocols that provide for equitable, sustainable and secure sharing must be agreed upon.

Specific water resources problems vary from state to state and include

- low levels of access to safe water supplies;
- inadequate or obsolete supply systems in growing cities;
- waterlogging and salinization of agricultural areas by irrigation (e.g., along the Nile, Tigris, and Euphrates);
- unsustainable groundwater abstraction
- pollution of aquifers and watercourses by human, industrial, and agricultural contamination;
- pollution of overpumped coastal aquifers by seawater intrusion (e.g., Oman, the Gulf States, and Libya); and
- drought and seasonal variability of water supplies.

Most freshwater in the region is used for irrigated agriculture. However, the economic value of water used for irrigation, in terms of the value of food produced, is lower than its value for municipal or industrial use. This situation has been exacerbated in some states by now outdated national policies restricting production to low-value staple food crops that have high water requirements.

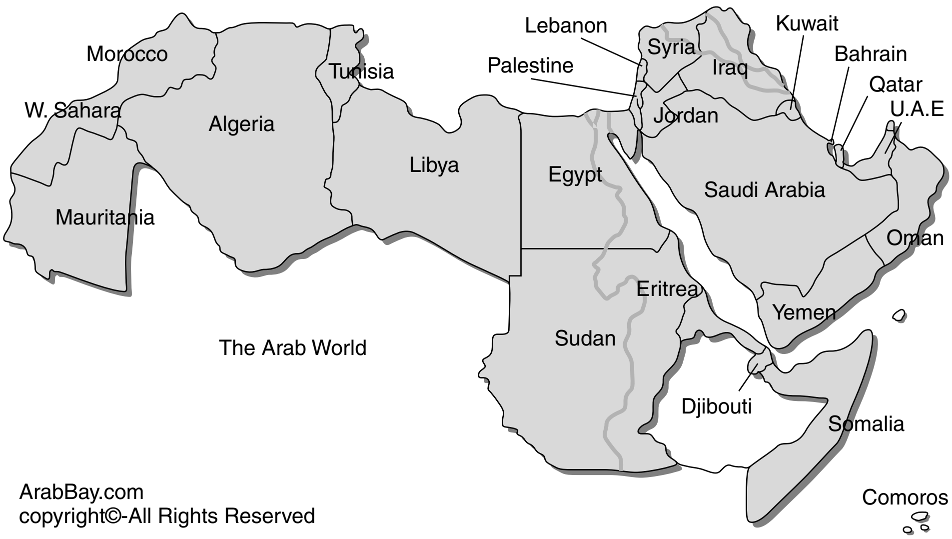

Figure 1. The Arab World.

Reducing water use in irrigation would be the most effective means of reducing the water deficit in most countries. At the same time, reallocating water to higher value sectors would enhance the GDP of all countries of the region. How to effect this transformation with the minimum social disruption in the agricultural sector, while maintaining food supplies, is the central issue of water resources management in the region.

SUPPLY AND DEMAND AT THE START OF THE TWENTY-FIRST CENTURY

Data on water supply and demand from Arab states vary in extent, reliability, and detail. Inventories of water resources are still in progress in most countries, and data on consumption, particularly, are generally inferred rather than measured.

Table 1 shows the estimated volumes of naturally renewable water for the region by state and per capita. Renewable resources comprise internally renewable resources from groundwater recharge and drainage to rivers and the net flows (inflows less outflows) of surface and groundwater from neighboring states. Inflows and outflows agreed to by treaty, or established by political circumstances, are implicit in these data; in the eastern Mediterranean states, particularly, such arrangements are complex and politically uncertain (3). The proportion of the actual renewable resource that comes from outside a state's borders is shown by the dependency ratio.

By 1995, resources per capita had fallen below the level of 1000 m^3 pcpa (per capita per annum) in four states and below 500 m^3 pcpa in 10 more. (An average level of 1000 m^3 is regarded as the level above which a state can be self-sufficient in food; less than 500 m^3 pcpa indicates severe water stress.) At the current rate of population growth, there would be four and 14 states, respectively, in these categories by 2020.

Figure 2 shows how much of this renewable resource is provided by recharge to groundwater, for those states where data exist. Those states that have the highest levels of water stress rely heavily on groundwater.

In many states, groundwater is abstracted faster than it is replenished. Figure 3 shows the proportion of the total annual water supply that is drawn unsustainably ("mined") from nonrenewable ("fossil") groundwater. Figure 4 shows groundwater abstraction as a percentage of renewal; in some states, groundwater is being mined several times faster than it is replenished.

This continued depletion of freshwater resources results from the increasing food demands of growing populations and from agricultural policies that seek national or local food self-sufficiency through irrigation, regardless of the unsustainability of the water supply. Figure 5 shows the proportion of national water supplies allocated to irrigation.

In many states, private well owners hold unrestricted traditional rights to groundwater for irrigation, regardless of the unsustainability of the resource (e.g., Oman, Yemen). Some governments have sought to exploit fossil resources to supply large-scale agriculture (e.g., the Great

Table 1. Renewable Water Resources[a]

	Internal, Million m^3	Actual Total, Million m^3	Actual, pcpa (1995), m^3/a	Dependency Ratio, %
Mauritania	400	11,400	5,013	96
Iraq	35,200	75,420	3,688	53
Sudan	35,000	88,500	3,150	77
Eritrea	2,800	8,800	2,480	68
Comoros	1,020	1,020	2,100	0
Syria	7,000	26,260	1,791	80
Somalia	6,000	15,740	1,702	62
Lebanon	4,800	4,407	1,465	1
Morocco	30,000	30,000	1,110	0
Egypt	1,800	58,300	926	97
Oman	1,650	1,650	775	0
Djibouti	300	300	520	0
Algeria	13,900	14,300	512	3
Tunisia	3,520	4,120	463	15
Yemen	4,100	4,100	283	0
Bahrain	4	116	206	97
Jordan	780	880	161	23
Saudi Arabia	2,400	2,400	134	0
Libya	600	600	111	0
Palestinian Terr.	600	225	100	0
Qatar	51	53	96	4
UAE	150	150	79	0
Kuwait	0	20	13	100

[a] References 1 and 2.

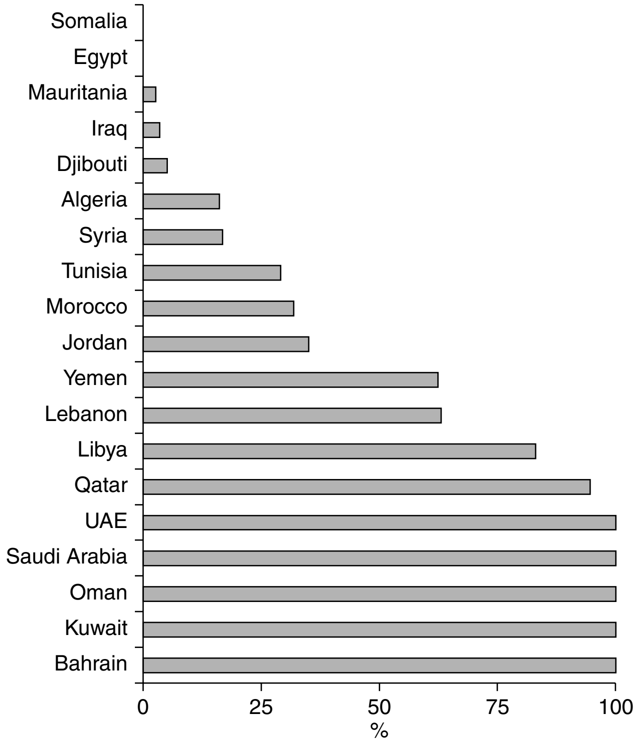

Figure 2. Groundwater recharge as a % of total renewable water resources (1,4).

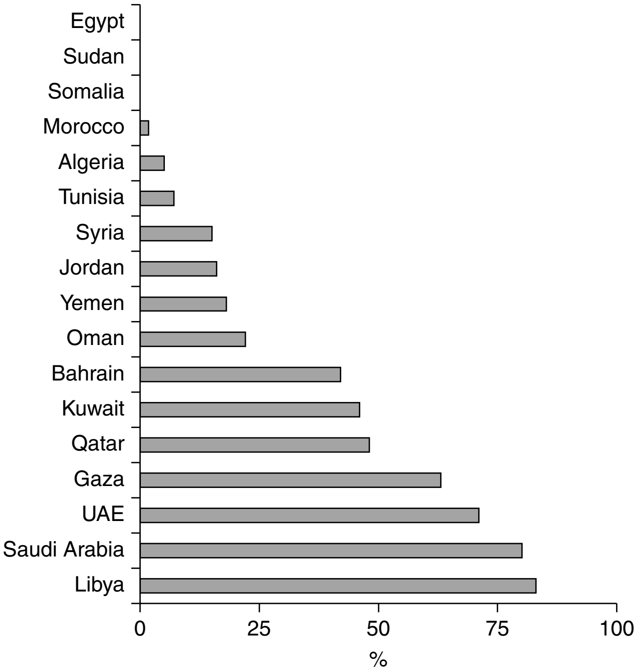

Figure 3. Groundwater overdraft (withdrawal less renewal) as a percentage of total supply (4).

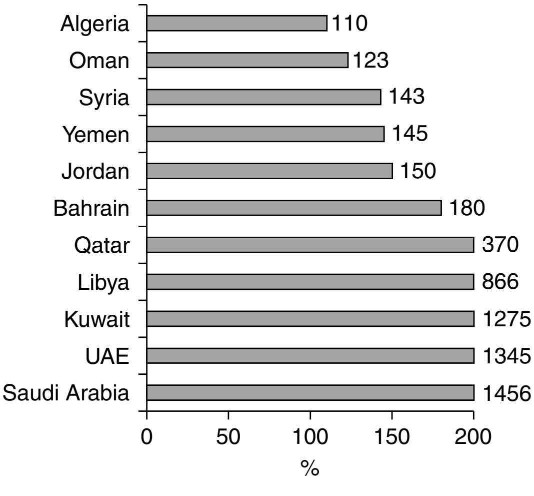

Figure 4. Groundwater withdrawal as a percentage of renewal (1,4).

Man-made River Project in Libya or for grain production in Saudi Arabia). Access to groundwater has been easier since about 1970, following the introduction of engine-driven groundwater pumps and the availability of deep drilling technology.

Figure 6 shows the volumes of “virtual” water already imported either as processed food, grain, or as live food animals, as a percentage of the total renewable water resource. The high volumes of virtual water show clearly that food self-sufficiency is an unrealistic target for many states.

These data show that the reason underlying the growing shortage of water in the Arab region is the increasing use of water, mainly for irrigation, at rates that exceed rates of replenishment. Historically, governments in the region have sought to address this problem by augmenting water supplies through engineering methods, rather than by encouraging conservation. Augmentation measures considered or implemented include interstate and interbasin water transfer, aquifer development, desalination, and minor processes such as cloud seeding and fog collection. However, the full production and delivery cost of water by all these methods typically exceeds the economic value of the food produced, and most water supplied by these means is now used only to supply higher value urban or industrial demands. The scope for further cost-effective augmentation of supply is limited by technology.

CONSERVATION AND DEMAND MANAGEMENT

Conservation methods used in the region include rainwater harvesting using small structures, recharging flood flows to aquifers, recycling of wastewater to irrigation, and leak reduction in urban networks. The storage of flood flows and treated wastewater in aquifer zones for later recovery and use needs to be further researched and promoted in the region. There is also substantial potential for increasing the conservation and re-use of treated wastewater. However, water can be most effectively conserved by improving farming and irrigation practices, in two ways: by improving food productivity per unit of water used and by reducing water losses in irrigation systems. The former approach includes the introduction of higher yielding and more water efficient (including salt tolerant) crop species and the use of scientific irrigation strategies specific to the crop type and conditions. Water reduction methods include the introduction of drip and sprinkler technology to replace spate irrigation and reduction in transmission and evaporation losses. In most countries of the region, significant savings can still be made by these technical means.

Despite these helpful measures, water resource deficits can be reversed only by fundamental changes in national water and food policies that reduce the water demands of the agricultural sector and gradually divert water to more economically productive sectors, while, at the same time, supporting rural communities in this transition.

POSSIBLE MANAGEMENT SOLUTIONS

It is clearly necessary for each country to develop and implement policies and strategies to produce a new and sustainable balance between supply and demand in irrigation. Most countries continue to pursue supply augmentation by various methods and to promote and subsidize conservation through engineering methods. Nonengineering interventions to promote conservation of water fall into four classes:

- Political
  - Policy revision
  - Water resources assessment, monitoring, and control
  - Pursuit of favorable international agreements for food importation
  - Promoting efficient management of centralized water supplies
- Legal
  - Quotas (limits) on abstraction
  - Water rights—possibly transferable or tradable within a basin
  - Restrictions on water use
- Fiscal
  - Tariffs on water use
  - Taxes on pump fuel
  - Removal of tariffs on cheap agricultural imports
  - Subsidy programs for irrigation and agricultural modernization
- Institutional
  - Expert consultancy
  - Social pressure and public awareness
  - Development of integrated management on a river basin or aquifer scale
  - Outreach and participation, including water users' associations

The combination of approaches that is appropriate for each country depends on the extent and effectiveness of national and local skills and institutions.

Political

Revision of government policy and coherency of policies between ministries and internationally, is paramount. A lack of coordination may result in development plans that cause unsustainable depletion of resources. Some states in the region still actively pursue policies of food self-sufficiency despite the lack of sustainable water resources.

The task of government is first to establish an enabling environment in which government, industry, and non-governmental entities work coherently toward water resource conservation, informed by a good understanding of the extent and quality of the resources. Second, governments and the international community must seek favorable international trade agreements for importing food.

Practical political interventions, whereby centralized governments operate at least primary systems of water-lifting and distribution, are common throughout the region. Great investment is required to upgrade infrastructure and improve efficiency, particularly for the delivery of municipal and rural drinking water. The capital needed is beyond the capacity of some countries but private sector financial participation through public–private partnerships is an alternative. Full privatization of water supply is widely opposed in the region and governments need to develop partnership models that are locally acceptable. To reduce risk and encourage investment, governments need to establish clear water law and robust fiscal regulations, and improve transparency.

International treaties to define shared resources and protocols for sharing information are essential for securing transboundary supplies. The Neil Basin Initiative and the Nubian Sandstone Aquifer System are two vehicles through which several countries consult and collaborate on managing the surface and underground resources of the major resources of northeast Africa. Similar processes are needed for the shared resources of the northern Arabian Peninsula.

International cooperation at a wider level to promote the development and sharing of water technology and information is essential. Recent valuable initiatives and foundations include the Dialogue on Water and Climate in the Mediterranean Region, organized by the World Conservation Union; the Middle East Desalination Research Center, in Oman; and the International Center for Biosaline Agriculture, in Dubai.

Legal and Fiscal

Economists believe that a sustainable water demand will be achieved only by charging the full cost of water to all users. The principle of charging tariffs for water delivery is widely accepted in urban areas of the region, but costs are only partially recovered. Stepped tariff structures that subsidize low-volume users are common. Uncharged-for water losses are abnormally high in some cities of the region.

Other legal and fiscal measures, such as irrigation quotas, water rights, and full-cost pricing are applicable only where water law is clearly defined and where established economic and financial frameworks facilitate the operation of market-based systems. This is not yet the case in most of the Arab region, where volume restrictions on irrigators or fiscal restrictions have been and are likely to be opposed. A cultural perspective that water is a free natural resource is a major obstacle to the introduction of innovative legal and fiscal instruments.

Institutional

Decentralization of water resources management to regional and local management systems is widely regarded as the most effective means of improving water supply services at the level of use, coupled with a well defined water law and integrated water resources management policy at the national level (5).

Institutional approaches are most appropriate where organizations and networks of civil society are amenable to development through training, leadership, and incentives. For example, Egypt, Morocco, and Tunisia have developed water users' associations and basin management organizations that seek to bring all stakeholders together in cooperative and participative operations. They assess the availability and sustainability of the resources; allocate water-use rights based on need, environmental demand, social contingencies, and economic value; control and monitor water distribution and quality; provide training; and collect fees to finance administration and system maintenance.

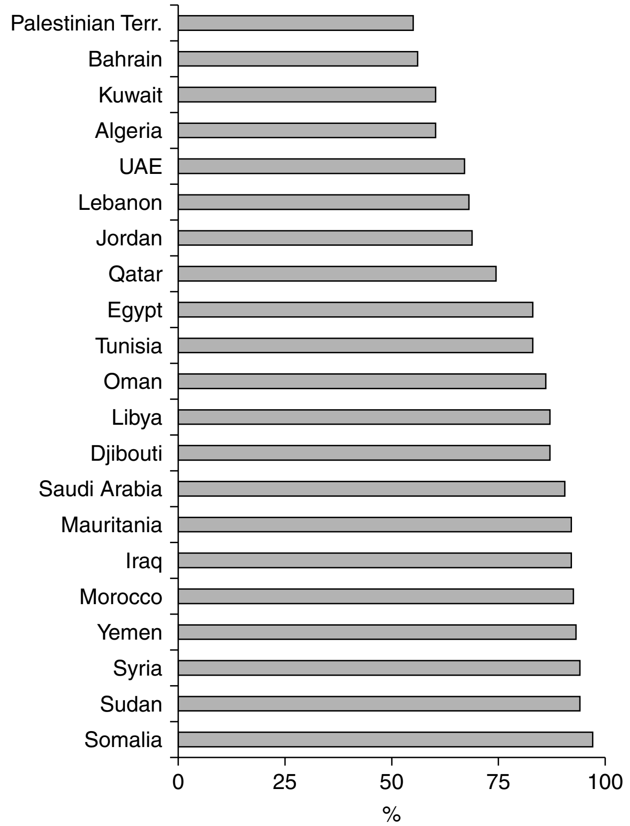

Figure 5. Share of water allocated to agriculture (1).

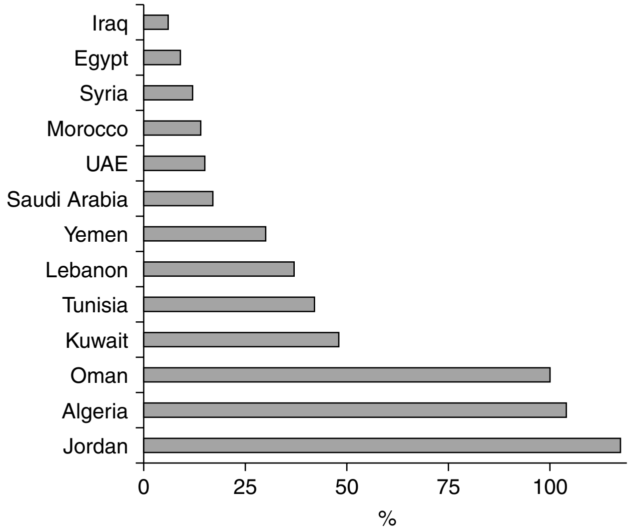

Figure 6. Water embedded in net food imports as a percentage of total renewable resource (4).

The Influence of Islam

Recent international prescriptions for water management, such as the Dublin Principles of 1992 and subsequent pronouncements, have emphasized economic value as a driving criterion for allocating water among sectors. This approach tends to ignore social priorities in rural Arab states and underestimates the practicalities of implementing water management changes in the contexts of Islamic law and, in some states, weak or inflexible political systems.

The holy writings from which Islamic law is derived contain many prescriptions relevant to water management and concur with the Dublin principles regarding community participation. Water is seen as a social good that must be managed sustainably and for the good of all; concepts of payment for delivery, the use of treated wastewater, and privatization of delivery are regarded as permissible (6).

Islam plays a prominent, central, and universally respected role throughout the region and, as an institution distinct from but working with both government and people, wields great influence. Islam has the potential to educate and lead consumers and drive the necessary changes in social attitudes toward water management. Islamic institutions can, therefore, provide leadership in promoting water resource conservation where political and fiscal interventions may fail.

BIBLIOGRAPHY

1. Food and Agriculture Organization (FAO), Water Resources Development and Management Service. (2001). AQUASTAT Information System on Water in Agriculture. World Wide Web address: http://www.fao.org/.
2. Inter-Islamic Network on Water Resources Development and Management (INWRDAM). (1999). Table of water scarcity in OIC states. INWRDAM, Amman, Jordan. World Wide Web address: http://www.nic.gov.jo/inwrdam/.
3. Brooks, D.B. and Mehmet, O. (Eds.). (2000). *Water Balances in the Eastern Mediterranean*. International Development Research Centre, Ottawa, Canada.
4. Lajaunie, M.-L. (1999). *MENA/MED regional water data report*. In: *Proceedings, Second Regional Seminar of the MENA/MED Initiative on Water Policy Reform, Amman, Jordan, May 1999*. World Bank, Washington, DC.
5. 3rd World Water Forum, March (2003), Kyoto, Japan. *Thematic and Regional Statements*. World Wide Web address: http://www.world.water-forum3.com/.
6. Faruqui, N.I., Biswas, A.K., and Bino, M.J. (Eds.). (2001). *Water Management in Islam*. United Nations University Press.

READING LIST

Guerquin, F., Ahmed, T., Hua, M., Ikeda, T., Ozbilen, V., and Schuttelar, M. (2003). *World Water Actions: Making Water Flow for All*. World Water Council. Earthscan Publications, UK.

Mahdi, K. (Ed.). (1998). *Water in the Arabian Peninsula: Problems & Policies*. Exeter University, UK.

Rogers, P. and Lydon, P. (Eds.). (1994). *Water in the Arab World*. Harvard University Press, Cambridge, MA.

Schiffler, M. (1998). *The Economics of Groundwater Management in Arid Countries*. Frank Cass, London, UK.

Shady, A. (2000). *Summary: Arab countries session*. In: *Conference on Water Security in the 21st Century*. The Hague, Netherlands, March 2000. World Wide Web address: www.worldwaterforum.net/index2.html

World Bank. (1998). *From Scarcity To Security—Averting a Water Crisis in the Middle East*. World Bank, Washington, DC.

EFFLUENT WATER REGULATIONS IN ARID LANDS

RICHARD MEYERHOFF
CDM
Denver, Colorado

ROBERT GENSEMER
Parametrix
Corvallis, Oregon

JAMES T. MARKWEISE
Neptune and Company, Inc.
Los Alamos, New Mexico

This article addresses the regulation of water resources in arid environments and focuses on the special considerations of effluent-driven ecosystems. Effluent serves as an increasingly important aquatic resource in the arid West, and the application of existing regulations for natural systems to effluent-based bodies of water is receiving greater scrutiny. Issues affecting regulation of effluent include defining appropriate attainable beneficial uses for these created ecosystems, net ecological benefits of effluent discharge, and impacts from the modification of extant flow regimes resulting from effluent addition. The practical application of regulations in effluent-driven ecosystems depends on recognizing these key issues.

EFFLUENT WATERS IN ARID ENVIRONMENTS

Effluent-based waterbodies are an important aquatic resource in arid environments because of the limited availability of alternate sources of water. In addition to general scarcity, water availability is sporadic and unpredictable in dry lands; consequently, effluents from residential, commercial, or agricultural activities are frequently one of the more stable sources of water in the arid West (1). Effluent-driven ecosystems have proliferated, as dams and diversions, groundwater pumping, and the need to dispose of treated effluent in metropolitan areas have modified natural flow regimes.

Three basic natural flow regime types are perennial, intermittent, and ephemeral. Perennial waters flow continuously, intermittent waters flow continuously only at certain times of the year (e.g., because of spring snowmelt), and ephemeral waters flow only in direct response to precipitation. Effluent-based ecosystems are either “dependent” on (effluent is discharged to an ephemeral stream channel) or “dominated” by (effluent is discharged to an intermittent stream channel) treated wastewater effluent discharge.

The discharge of effluent to intermittent or ephemeral streams represents a discontinuity resulting in a disruption to the natural equilibrium as it exists at the time the discharge begins. The natural tendency for the created stream ecosystem is to restructure itself so that a new equilibrium is achieved. This restructuring will take some time; the length will depend on local factors and whether additional stressors are placed on the system (e.g., construction of physical structures in, across, or along the stream, or an increase in effluent flow) (2).

Riparian systems that develop as a result of wastewater discharged into normally dry channels may stand in stark contrast to the adjacent upland vegetation that is not influenced by discharge. In addition, the terrestrial community downstream of the discharge point can be distinctly different from the terrestrial community upstream of the discharge. The expected characteristics of an effluent-based waterway downstream of the effluent discharge can be modeled conceptually (Fig. 1). This conceptual model, which represents a created stream ecosystem where the effluent discharge is the only discontinuity, is based on the physical, chemical, and biological characteristics of 10 effluent-dependent and dominant ecosystems in the arid West (2). From a regulatory standpoint, created or modified stream ecosystems are a class of aquatic environments that are

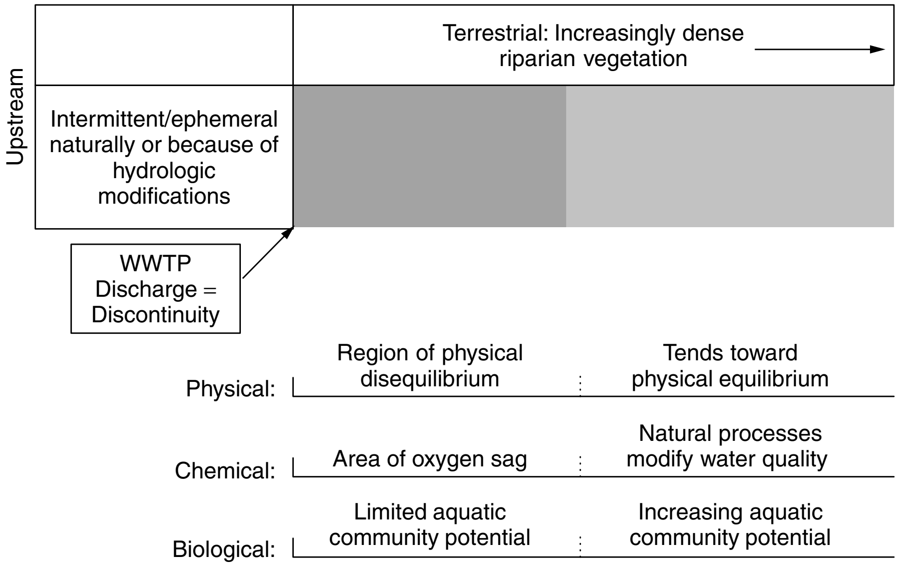

Figure 1. Conceptual model of an effluent-driven stream ecosystem.

typically not distinguished from natural ecosystems. Yet, these ecosystems may have unique attributes that result in different expectations for a "functioning" ecosystem.

Given the unique nature of effluent-based waterbodies, the challenge is to apply the existing regulatory framework in a way that is protective of this resource and still biologically meaningful. Below we address key regulatory issues for protecting effluent-derived ecosystems in arid lands. These considerations include defining "attainable uses" in effluent-driven ecosystems, using net ecological benefit as an alternative end point to measure use attainability, and implications of modifying the flow regime—is wetter better?

ATTAINABLE USES IN EFFLUENT-DRIVEN ECOSYSTEMS

States and tribes are required to establish beneficial uses for surface waters under their jurisdiction. At a minimum, existing uses must be protected (uses actually attained in the waterbody on or after November 28, 1975). States and tribes may go beyond the protection of existing uses and protect designated uses, which are expressed as goals for the waterbody. Designated uses must be protected only if the uses are actually attainable. The EPA defines attainable uses as those beneficial uses that can be achieved (1) when effluent limits under the Clean Water Act are imposed on point source dischargers and (2) when cost-effective and reasonable best management practices (BMPs) are imposed on nonpoint source dischargers. If all appropriate water quality control activities are implemented, then the pertinent question is, "what is the best beneficial use that can be achieved?" Determining what is attainable in a natural ecosystem can itself be difficult; however, when effluent creates the aquatic ecosystem, determining what is attainable is an even greater challenge.

The attainable use may be equal to the designated use or it may be something less. Consider a dry stream bed characterized as an ephemeral flow system. At the point of effluent input (Fig. 1), the water becomes perennial. The addition of water changes the potential of the system and thus what is attainable, especially with regard to the aquatic community. The effects of water addition, physical habitat, water quality, and other stressors define the potentially attainable aquatic community. The result of the interaction of these variables differs with time and also with distance downstream from the point of discharge (Fig. 1). Often, the water quality changes over time as a new technology is installed, new regulations are implemented, or the stream channel is modified. Each change modifies the potential outcome or potential aquatic community. In other words, defining what is attainable is a moving target.

Determining potential or what is attainable is typically defined in the context of the effluent quality, which is related to the degree/type of wastewater treatment. A general belief exists that improved treatment will result in improved aquatic community characteristics. This is likely to be true at low levels of treatment. For example, with the change from primary to secondary treatment, major improvements in the aquatic community are expected and do occur. It is not necessarily realistic, however, to expect the same degree of improvement with more advanced treatment changes, such as with the addition of nutrient removal or filtration, if the limiting factor is habitat-related.

Failing to establish appropriate end points for beneficial uses has significant regulatory and cost implications because the Clean Water Act mandates that states assess water quality in surface waters every 2 years and determine if any of these waters is "impaired." A waterbody is considered impaired if a determination is made that, even after implementing all required effluent limitations (technology and water-quality-based) and BMPs, the water quality still does not meet applicable water quality standards. This analysis presumes that beneficial uses and criteria are appropriately established. However, if the uses and criteria are inappropriate, the basis for the assessment is inaccurate, and the waterbody may be incorrectly assessed as impaired.

Impaired waters must have a total maximum daily load (TMDL) established under the Clean Water Act. A TMDL is the sum of the wasteload allocations for point sources and load allocations for nonpoint sources (including natural background), plus a margin of safety (uncertainty factor). TMDLs are implemented through permits, BMPs, or other water quality management methods. Pollutant allocations are based on meeting the applicable end point or water quality standard. TMDLs can require substantial resources to develop and implement; thus, it is very important to use appropriate end points as the basis for assessing waters correctly.

Under the current method for regulating created or effluent-based ecosystems, states typically establish the same water quality standards or end points for effluent-based waters as established for natural perennial waters. This approach is based on the assumption that the aquatic community in effluent-based waters should be the same as the aquatic community in natural perennial waters; that is, the attainable end point in a natural ecosystem should be the same as the attainable end point in a created ecosystem. When this assumption forms the basis for implementing water quality control programs, the same requirements for effluent limitations are applied to both types of waters.

From a regulatory perspective, establishing appropriate end points in effluent-based ecosystems is especially problematic. The approach outlined above can result in applying substantially more stringent effluent limitations on facilities discharging to effluent-based ecosystems than facilities that discharge to a perennial water. The reason for this difference is that when a facility discharges to a perennial water, credit is given to allow for mixing the effluent with in-stream water. This credit results in less stringent effluent limitations. No such credit is available when the receiving waterbody is dry. This approach results in regulation that is strictly discharge or end-of-pipe based. No consideration (positive or negative) is given to the ecosystem of the receiving water.

The discharge of effluent to dry riverbeds creates an aquatic ecosystem. This ecosystem may be "new," or it may replace an aquatic ecosystem that has been lost because of dams, diversions, groundwater withdrawals, etc. If effluent-based ecosystems are created and if they function differently from natural perennial waters such that what is attainable in the created ecosystem differs from what is attainable in the natural ecosystem, then the use of an alternative regulatory approach to implement a water quality control program might be appropriate.

NET ECOLOGICAL BENEFIT

A concept, called "net ecological benefit," addresses some of the concerns associated with establishing attainable end points for effluent-based waters (3,4). The net ecological benefit methodology is a tool for evaluating whether an alternative end point could be established if it can be demonstrated that the ecological benefits created by the discharge outweigh any identified detriments. A natural progression of this thought process is to consider the use of a "performance-based" approach that considers the functionality of the ecosystem. This approach could lead to establishing alternative success criteria or end points for evaluating ecosystem protection. Examples of net ecological benefit include enhancement of habitat/food for terrestrial native or threatened and endangered (T&E) species, enhancement or restoration of riparian values, preservation of existing habitat that could not be supported without effluent flow, restoration of aquatic and riparian values lost due to human activities, enhancement of water quality resulting in conditions conducive to ecosystem restoration and/or preservation, improvement or creation of habitat capable of supporting fish or allowing migration of anadramous species, and restoration of species diversity in aquatic ecosystems.

EPA (3) guidance notes key issues that should be considered when evaluating whether or not a net ecological benefit exists as a result of an effluent discharge. Foremost among these is that effluent discharges should not produce or contribute to concentrations of pollutants in tissues of aquatic organisms or wildlife that are likely to be harmful to humans or wildlife through food chain concentration. This requirement addresses EPA and USFWS policies to minimize persistent, bioaccumulative chemicals in the environment. Also, a continued discharge to the waterbody should not create or be likely to cause or contribute to violations of other end points, for example, downstream water quality standards or groundwater quality. Finally, all practicable pollution prevention programs, such as pretreatment and source reduction, are operational and the discharger has to be appropriately responsive to previous and ongoing compliance actions. These considerations are designed to ensure that (1) a minimum level of wastewater management and treatment is implemented and (2) pollutants are not exported from the water column into the ecosystem through the food chain.

Use of a net ecological benefit approach as an alternative for setting end points would change the regulatory focus in effluent-driven ecosystems from end-of-pipe based to ecosystem-based. Establishing this nontraditional approach would allow recognizing the environmental benefits of water flow in an otherwise dry channel. Moreover, alternative end points that include terrestrial elements would support efforts to implement permitting on a watershed basis, support ecological restoration efforts, address concerns about the increasing loss of riparian habitat in the arid West, support increased interest in habitat restoration in urban rivers, and potentially benefit threatened and endangered species by maintaining wildlife migration corridors.

MODIFYING THE FLOW REGIME—IS WETTER BETTER?

The addition of effluent to a riverbed has been portrayed as a benefit in the context that any water in a riverbed is better than no water in a riverbed. For sites where the effluent replaces historically natural flows, this thought process makes sense. However, when the effluent discharge creates a flowing river where none previously existed, the question can be asked, "what has been lost or changed by the addition of effluent?" This question is relevant because naturally ephemeral streams have important biological attributes that are as distinct as the biological attributes of a natural perennial river.

The addition of an artificial perennial flow makes the system clearly different from historical conditions (Fig. 1). Biological attributes such as aquatic community richness and diversity are likely to be greater. The increased biological productivity of the aquatic community will provide additional food resources for terrestrial organisms. In addition to these changes in the aquatic community, the terrestrial community will be substantially different, especially in terms of the types of organisms supported. Are these biological changes good? Is having a wetter channel better biologically? These questions have no simple answer. In fact, the answer will depend on public values and local needs. One can easily argue that the number of ephemeral channels, especially in arid regions, far exceeds the number of naturally perennial channels, and thus the creation of a perennial stream in a previously ephemeral stream is a positive benefit. However, in some areas, especially in rapidly developing urban environments, the number of lost ephemeral channels can be significant, and the loss of habitats as a result of effluent discharge can be an important issue for the public to consider.

The creation of effluent-driven ecosystems creates a relatively stable aquatic environment, at least from the standpoint of temperature and flow. Certainly, stormwater runoff will still occasionally cause spikes in flow, but the norm is a relatively constant flow with reduced temperature variation (compared to what might be expected in a natural environment). Native fish species are adapted to the relatively episodic flow regime of natural arid waters. These species have evolved where abiotic factors are the dominant driver in establishing biological communities. Accordingly, these species are better adapted to this environment than nonnative species. Modification of the flow regime removes this driver and if nonnative fish become established (as they have throughout the arid West), they may outcompete the native species.

Similar to fish, the richness and diversity of native riparian vegetation is influenced by the flow regime. For example, the selection of native cottonwoods over nonnative salt cedar can be driven by the distribution and frequency of flows across the floodplain during floods. Under a stable flow regime, native species may be outcompeted. In the end, ecological benefits may result from the discharge of effluent to an otherwise dry or intermittent riverbed, but the modified flow regime is not necessarily a positive benefit for native species.

FINDING A BALANCE IN REGULATING ARID LAND WATERS

Among the general public, a common desire exists to protect beneficial uses, establish water quality criteria to protect uses, and protect riverine habitat and associated wildlife. However, basic agreement on these fundamentals exists, but substantial disagreement may exist over how attainable end points should be defined and how uses should be protected. This disagreement often stems from differences of opinion over how end points should be calculated, what are appropriate assumptions, and how much conservatism or how much of a safety factor should be built into the regulations.

The EPA (4), states, and tribes often take the simplest, most conservative approach when establishing attainable end points because this approach requires the least amount of scientific data. Regulatory institutions cannot be faulted for this. Given the number of different types of aquatic systems (natural or created) and the various and often competing beneficial uses of water (from protection of native species to use of water for industry), the number of factors that could be considered in establishing end points for each waterbody is substantial. Resources are simply not available to establish site-specific uses and criteria for each waterbody. Although regulators are limited in what they can reasonably consider when they establish criteria for an entire state or area, options are available for modifying uses and criteria and establishing appropriate site-specific end points. These tools should be used, when appropriate, because the importance of establishing correct end points cannot be understated. If water quality control efforts are based on unattainable uses or criteria, substantial costs may be incurred with little or no improvement in water quality.

BIBLIOGRAPHY

1. Ryti, R.T. and Markwiese, J.T. (in press). Assessment of ecological effects in water-limited environments (WR-3). *Encyclopedia of Water*, J. Lehr (Ed.). John Wiley & Sons, New York.
2. Pima County Wastewater Management Department. (2002). *Habitat Characterization Study*. Prepared by URS Corporation and CDM for Arid West Water Quality Research Project, PCWMD, Tucson, AZ.
3. EPA. (1992). *Guidance for Modifying Water Quality Standards and Protecting Effluent-Dependent Ecosystems*. Interim Final. EPA Region 9.
4. EPA. (1993). *Supplementary Guidance on Conducting Use Attainability Analyses on Effluent Dependent Ecosystems*. EPA Region 9.

CALIFORNIA—CONTINUALLY THE NATION'S LEADER IN WATER USE

William E. Templin
U.S. Geological Survey
Sacramento, California

In 1950, the U.S. Geological Survey began publishing a series of water-use circulars entitled "Estimated Use of Water in the United States." Every 5 years since then the report has been updated and now provides a valuable, long-term data set of national water-use estimates. Since the inception of this series, California has always reported the largest total fresh and saline withdrawals of all states in the Nation (Fig. 1). The most recent update in this series (1) reports 1995 conditions. California again accounts for the largest withdrawal of water for off-stream uses of all states, which is about 45.9 billion gallons per day, followed by Texas (29.6 Bgal/d), Illinois (19.9 Bgal/d), and Florida (18.2 Bgal/d, Figs. 2 and 3).

The poster demonstrates WHY and HOW California continues to lead the Nation in many withdrawal categories (Figs. 4 and 5). Withdrawals include water removed from the ground or diverted from a surface-water source for use.

WHY? California continues to be the most populous state in the United States, accounting for 12% of its people (more than 32 million in 1995) followed by Texas and New York with about 7% each (Fig. 6). In 1995, California accounted for 5.62 Bgal/d, 14% of the Nation's public supply freshwater use (Fig. 7). Public supply withdrawals are directly related to population supplied, but they also are influenced by withdrawals to supply industrial and commercial water users.

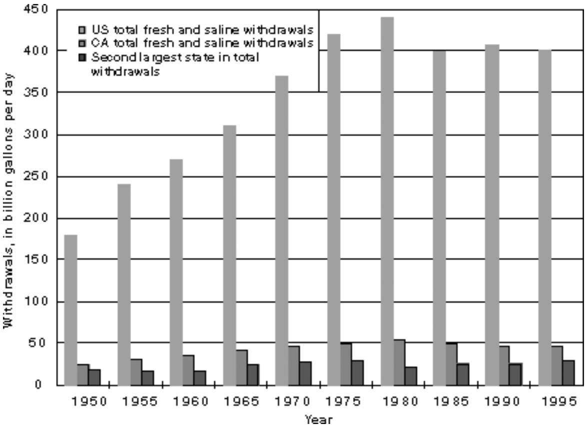

Figure 1. Trends in total fresh and saline withdrawals, 1950–1995.

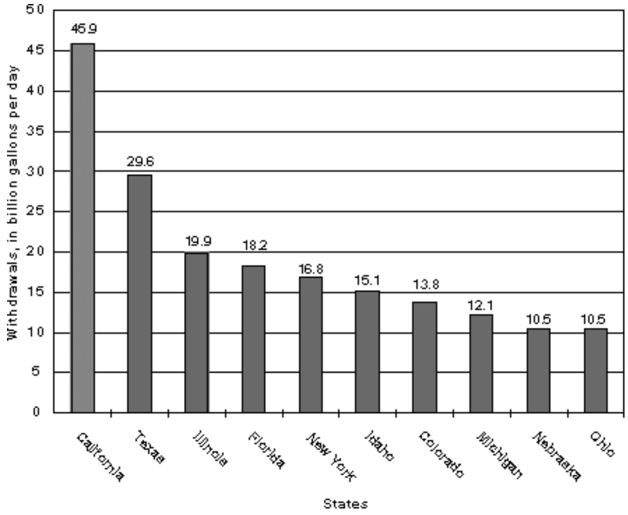

Figure 2. Total fresh and saline withdrawals by top 10 states, 1995.

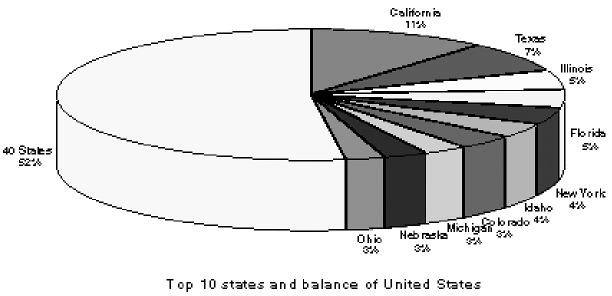

Figure 3. Total fresh and saline withdrawals in the United States during 1995, 402 Bgal/d.

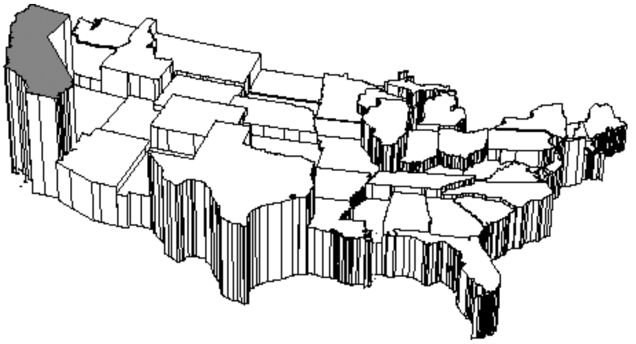

Figure 4. California led all states in freshwater withdrawals followed by Texas, Illinois, and Idaho, 1995.

HOW? Los Angeles County's population exceeded 9.3 million in 1995 and contributed to the county's large public supply withdrawals. Only eight states have more people than Los Angeles County, which accounts for 29% of California's population (Fig. 8).

WHY? California accounted for 28.9 Bgal/d in total irrigation withdrawals in 1995, with 22% of the Nation's total irrigation withdrawals (Fig. 9).

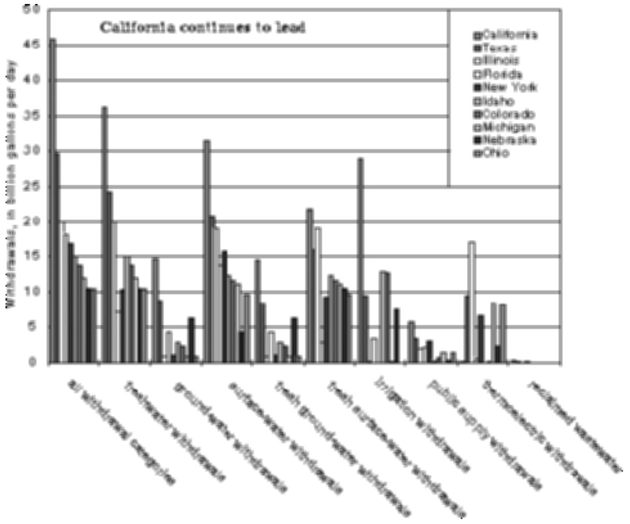

Figure 5. Categories of total withdrawals in the top 10 states, 1995.

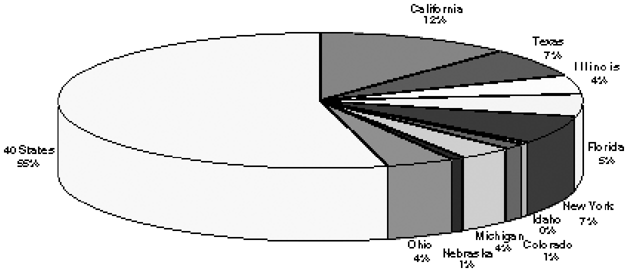

Figure 6. 1995 U.S. population—267 million.

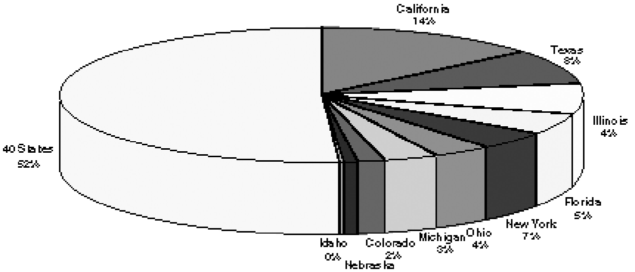

Figure 7. U.S. public supply withdrawals, 40.2 billion gallons per day, 1995.

- GOT MILK? California is the Nation's largest dairy state, producing 26 billion pounds of milk and cheese (2).
- GOT COTTON CLOTHING? California is the Nation's second largest cotton producer.
- HUNGRY? More than half of the Nation's fruits, vegetables, and nuts are grown in California.

HOW? The total withdrawals in the Tulare-Buena Vista Lake Hydrologic Cataloging Unit accounted for more than 8330 Mgal/d in withdrawals for irrigation uses in 1995 (Fig. 10) and 8800 Mgal/d of total freshwater. Only 13 states had more total freshwater withdrawals. This area

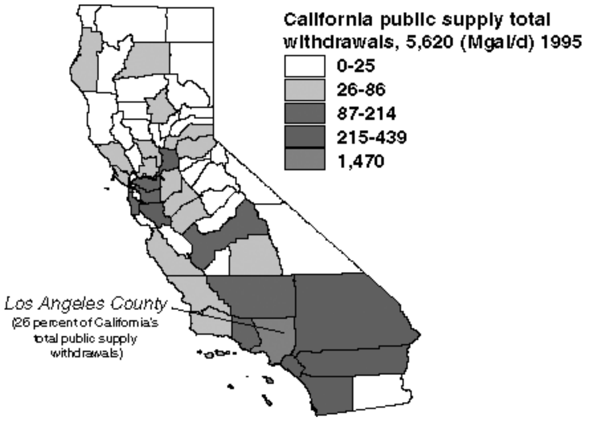

Figure 8. Public supply total withdrawals by county, in California, 1995.

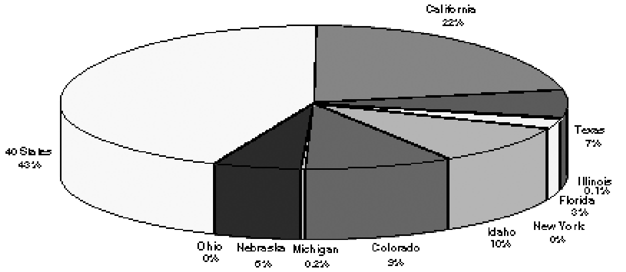

Figure 9. Irrigation withdrawals in the United States in 1995 totals 134 billion gallons per day.

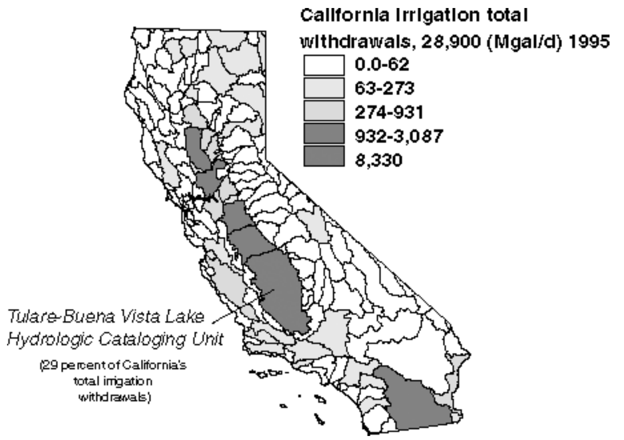

Figure 10. Irrigation total withdrawals by hydrologic cataloging unit, in California, 1995.

contains many of California's dairies and much of our irrigated acreage.

Even though California leads all states in total withdrawals, California's per-capita use rate for total offstream use (1130 gal/d) ranks below the national average (1280 gal/d, Fig. 11).

Acknowledgments
The author acknowledges the assistance of Michael B. Roque, Illustrator, in the layout and preparation of the original poster presented at the American Water Resources Association Annual Conference on Water Resources, November 16–19 1998, Point Clear, Alabama.

The author also acknowledges the assistance of Tammy Shelton, Management Analyst, in the development of the web page version of this information.

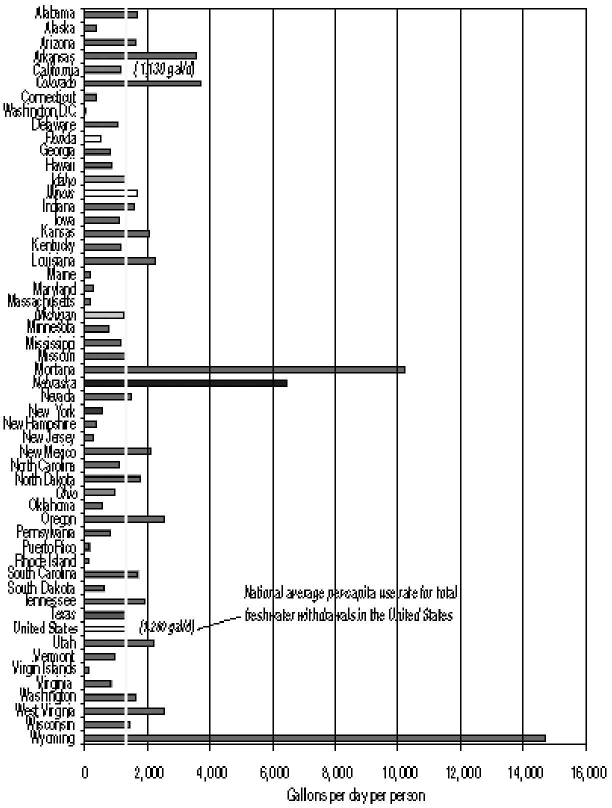

Figure 11. Per-capita use rates for total freshwater withdrawals for the United States and each state, 1995.

BIBLIOGRAPHY

1. Solley, W.B., Pierce, R.R., and Perlman, H.A., 1998, *Estimated use of water in the United States in 1995: U.S. Geological Survey Circular 1200*, p. 71. http://water.usgs.gov/public/watuse/
2. California Department of Food and Agriculture. Home Page. October 1998. http://www.cdfa.ca.gov/

LESSONS FROM THE RISING CASPIAN*

Igor S. Zonn

INTRODUCTION

The Caspian Sea is the biggest inland body of water in the world. Its surface area is roughly equivalent to the combined area of The Netherlands and Germany (about 400,000 km^2, or 144,000 mi^2). The surface water inflow into the sea is formed by the flow of the Volga, Ural, Terek,

---

*From *Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts*, T.D. Potter and B.R. Colman (Eds.). John Wiley & Sons, Hoboken, NJ. pp. 885–892.

Sulak, Samur, and Kura Rivers, and small Caucasian and Iranian rivers. The watershed area of the Caspian Sea is 3.5 million square kilometers. The basin of the Volga River makes up nearly 40% of the territory of the catchment of the Caspian Sea, and it supplies about 80% of the total volume of annual water flow into the sea. All components of the Caspian ecosystem, directly or indirectly, to a greater or lesser extent, are influenced by river flow.

The Caspian Sea basin falls into three morphologically different parts: (1) the northern (25% of the sea area), a shallow area (less than 10 m deep; about 20% with depths less than 1 m) extending to a conventional line passing from the Terek River to the Mangyshlak Peninsula; (2) the medium (35%), with an average depth of 170 m (the maximum being 790 m); and (3) the southern (39%), the deepest area, with a maximum depth of 1025 m and an average depth of 325 m. Deep depressions in the northern and southern parts of the sea are divided by an underwater threshold running from the Apsheron Peninsula to Turkmenbashi (formerly Krasnovodsk) (1).

Before the breakup of the Soviet Union in December 1991, the USSR and Iran were the only two independent nations occupying the shores of the Caspian. With the breakup, three additional newly independent nations emerged along the coast: Azerbaijan, Kazakstan, and Turkmenistan. The Russian Federation's Caspian coastline is shared by three of its political units: Astrakhan Oblast, the Republic of Kalmykia, and the Republic of Dagestan.

NATURE OF SEA-LEVEL CHANGES IN CASPIAN SEA

The Caspian Sea is a closed basin in the inland part of Eurasia and this sea's water level is below that of the world ocean. The sea basin stretches almost 1200 km from north to south and its width varies between 200 and 450 km. The total length of the coastline is about 7000 km. Its water surface area is about 390,600 km^2 (as of January 1993). Water salinity in the northern part is 3–6‰ and reaches 12‰ in the middle and southern parts.

Fluctuations in sea level for various lengths of time can be found in the data of geomorphological and historical studies of the record of the Caspian Sea (Fig. 1). Within the last 10,000 years, the amplitude of fluctuations of Caspian Sea level has been 15 m (varying from −20 to −35 m). During the period of instrumental observations (from 1830 onward), this value was only about 4 m, varying from −25.3 m during the 1880s to −29 m in 1977. Annual increases in the level during this period met or exceeded 30 cm on three occasions (in 1867, 30 cm; in 1979, 32 cm; and in 1991, 39 cm). The mean annual increment in the level in the 1978–1991 period was 14.3 cm.

Natural factors are the primary cause of recent Caspian Sea level fluctuations (but not the only cause). Scientists have identified three distinct periods of level changes: 1830–1930, 1931–1977, and 1978 to the present. The first period of 100 years saw sea-level fluctuations not exceeding 1.5 m (5 ft). Researchers considered this period to have been relatively stable. The second period, from 1931 to 1977, is identified by a constant decline in level by 2.8 m (9.1 ft), and in 1977 the Caspian Sea reached its lowest level since the beginning of instrumental record-keeping in the 1830s.

As the sea level declined throughout the 1950s, 1960s, and early 1970s, Soviet scientists forecast that the decline would continue for at least a few decades into the future. Scientists have linked the reason for the decline to the regulation of Volga River flow. During these decades, major engineering activities were undertaken along the Volga, such as the construction of water diversion canals, reservoirs, and dams. The construction of such engineering facilities diverted water away from the Caspian.

In response to this major drop in sea level, human settlements bordering the sea coast began to move toward the receding coastline. Fields and pasturelands were prepared for use, roads and rail lines were constructed, and housing and factories were built on the newly exposed seabed. During the Soviet era, many people emigrated from other parts of the region to settle along the border of the sea. Development of infrastructure along the coast took place to support the increasing population.

In an attempt to save the Caspian from drying out, Soviet scientists and engineers proposed the construction of a dam to block the flow of Caspian water to Kara-Bogaz-Gol Bay, a large desert depression in Turkmenistan adjacent to the sea's eastern shore. Political decisions made in the mid-1970s ordered the construction of the dam, but due primarily to bureaucratic inertia, the dam was not completed until the early 1980s. This was a few years after the Caspian's sea-level change had reversed direction. Before the dam was constructed, the bay took in 40 km^3 (8.6 mi^3) of Caspian water annually. It served as a huge evaporation pond, as well as a natural location for the accumulation of commercially useful mineral salts.

Another Soviet government response to the decline in the Caspian's sea level was a diversion of water into the Volga River from other Soviet rivers that flowed northward into the Arctic Ocean. River water flowing into the Arctic was viewed as wasted and without value to the Soviet Union because it was unused by human activity.

THE CASPIAN RISES

To the surprise of Soviet scientists, the level of the sea began to rise suddenly in 1978, the beginning of its third period of level changes. Since then, the Caspian has risen

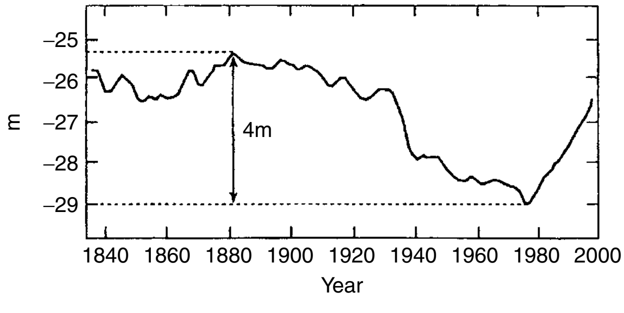

Figure 1. Observed Caspian Sea level, 1835–1999.

steadily by more than 2.5 m. One of the first actions the newly independent government of Turkmenistan took in 1992 was to tear down the dam in order to allow great amounts of water to flow into Kara-Bogaz-Gol Bay again and to replenish the supply of salts.

Scientists have proposed a variety of hypotheses about why the Caspian Sea level had increased so rapidly. These can be clustered into the following categories: tectonic plate movement on the seabed, climate fluctuations and change, and hydraulic construction along the Volga River, or some combination of these factors.

Tectonic Plate Movement Hypothesis

Tectonic movements over periods such as centuries and millennia have been the cause of many geologic changes in the Caspian basin. The region has been subjected to uplift, subsidence, overthrust of landforms, seabed mud-volcanic activity, and landslides, in addition to erosion processes and the accumulation on the Caspian seabed of river-transported sediments. However, it is difficult to see how tectonic movements could cause such sharp fluctuations in the Caspian's sea level over relatively short periods. Thus, it appears that such movements have had an insignificant impact on recent sea-level fluctuations.

Climate Change Hypothesis

Today, most Russian scientists believe that climatic factors are the real cause of the Caspian Sea level rise. Studies by Golitsyn (2) and Golitsyn and McBean (3) indicate that recent changes of the Caspian Sea level are 90% associated with corresponding changes in the water balance components of the sea, as opposed to possible tectonic activity. The volume of inflow from rivers to the sea increased sharply after 1978. During certain years (e.g., 1979, 1985, and 1990), more than 350 km^3 of river water entered the sea. From 1978 until 1990, Volga River flow exceeded 260 km^3/yr. At present, no arguments have challenged the view that the main contribution to seasonal and annual level fluctuations of the Caspian is accounted for by surface inflow and evaporation levels. Within recent decades, the sea's fluctuations have been subjected to anthropogenic impacts as well.

In this regard, climate has two dimensions: climate fluctuations and climate change. Climate fluctuations occur on various time scales, with those of interest to present-day society being on the order of decades and perhaps centuries. Climate-related fluctuation refers to the increase and decrease of sea level over the course of decades. During the past two centuries, the sea has undergone several fluctuations. Those of the twentieth century have adversely affected socioeconomic activities and infrastructure along the sea's coastline.

The view that climatic processes in the Volga basin are the dominant cause of sea-level fluctuations has been recently reinforced. Droughts in this basin and sharply reduced Volga flow into the Caspian from mid-1995 until early 1997 have been associated with a 25-cm (10-inch) drop in Caspian level. Nevertheless, Russian scientists still suggest that the sea level will continue to rise into the first decades of the twenty-first century.

Climate change associated with global warming induced by human activities has also been proposed as the forcing factor behind the Caspian's rise since 1978. Those who see global warming as the forcing factor suggest that the most recent sea-level rise can be associated with intensification of the hydrologic cycle (i.e., more active precipitation-producing processes), an intensification that some scientists have linked to the human-induced global warming of the atmosphere. An increase in precipitation within the Volga River basin would translate into increased sea level.

Hydraulic Construction Hypothesis

Some observers have argued that the recent fall and rise in sea level were the result of human activities. They suggest that the widespread development of hydraulic structures (e.g., dams, reservoirs, irrigation systems) in the Volga River basin, beginning in the 1950s, led to a sharp decline in Volga flow. The filling of many reservoirs built along the rivers flowing into the Caspian, the increase in industrial and municipal water use by several times, and changes in the water regime of the floodplains led to a decrease of streamflow into the sea. Such a hypothesis could be tested by constructing a water budget model for the Caspian. Such a model would need to identify all the inflows into the Caspian Sea (such as from rivers and groundwater) and all outflow from the sea (such as evaporation and water diversions). While it is a seemingly straightforward task, identifying all the sources and sinks of Caspian water is not easy.

There is also a hypothesis about an Aral Sea connection. Yet another suggestion that seems to be made at just about every Aral or Caspian Sea conference is that the decline in the level of the Aral Sea is linked to the rise in level of the Caspian. The reasoning is that water diverted from the Aral basin to the Caspian basin to irrigate the desert sands for cotton production in Turkmenistan ends up either being evaporated into the air or seeping into the groundwater, which eventually makes its way into the Caspian. However, it is important to point out that *both* the recent fall and rise in the Caspian Sea level occurred during three and a half decades of a constant decline in the Aral's level.

SOCIETAL IMPACTS OF SEA-LEVEL RISE

According to a UN Environment Programme estimate, the cost of the impact of the sea-level rise of the Caspian, as of 1994, was $30 to $50 billion (U.S.). Coastal ecosystems have been destroyed, villages inundated and populations evacuated, sea banks eroded, and buildings destroyed. Coastal plains have been invaded by subsurface seawater or have become waterlogged. Fauna have changed, and pasturelands and sturgeon spawning grounds have been destroyed.

Each of the five countries sharing the coasts of the Caspian Sea has suffered losses, and those losses increased until the mid-1990s. They suffer from the different impacts of sea-level rise because the territory along its coastline is neither uniformly settled nor uniformly developed economically. Economic losses in the big cities and villages have been higher than in the rural areas. More specifically, in Astrakhan Oblast (equivalent to an American state), about 10% of its agricultural land was out of production by 1995 because of sea-level rise. The coastline of the Republic of Dagestan (also part of Russia) was affected by the flooding of at least 40 factories in its cities of Makhachkala, Kaspiysk, Derbent, and Sulak. Nearly 150,000 hectares (370,000 acres) of land have been inundated, with a loss of livestock production and breeding facilities. Much of the 650-km (390-mile) Caspian coastline of Turkmenistan is made up of low-lying sandy beaches and dunes that are vulnerable to coastal flooding and erosion. In fact, some Turkmen villages that were once several kilometers from the sea are now coastal communities. Similar adverse impacts of sea-level rise on human settlements and ecosystems are found in Kazakstan, Azerbaijan, and Iran.

The Caspian has been referred to as a “hard currency sea” because of its large oil and natural gas reserves and because of its highly valued caviar-producing sturgeon. Regional reserves contain upward of 18 billion metric tons of oil and 6 billion cubic meters (215 billion cubic feet) of natural gas. Experts suggest that the Caspian is second only to the Persian Gulf with respect to the size of its oil and gas reserves, and that Turkmenistan is a “second Kuwait.” If the sea level were to continue to rise, a large part of the oil and gas mains along the Turkmen coast would become submerged and would also be subjected to corrosion by seawater. Coastal settlements, which include the greater part of Turkmenistan’s oil, gas, and chemical enterprises, would also be threatened. Similar environmental problems would certainly affect other Caspian coastal countries as well (4).

The Caspian Sea is unique in yet another respect: it contains about 90% of the sturgeon that produce the lucrative prized black caviar for export to foreign markets. Sturgeon roe is often referred to as “black gold.” Today, however, Caspian sturgeon is at risk of extinction from overexploitation by illegal poachers and by destitute fishermen desperately seeking funds to buy food for their families. The sea-level rise, with its destruction of sturgeon spawning grounds, adds yet another threat to the endangered Caspian sturgeon.

Poachers hunt sturgeon only for its caviar. Today, they catch sturgeon directly in the open sea. However, in the early 1960s, prohibition was introduced by the former USSR against catching sturgeon in the open sea. Since that time, catching sturgeon has been carried out in the river deltas. Sturgeon reproduce very slowly: the fish do not spawn for the first time until they reach the age of 20–25 years. In 1990, the permissible catch of sturgeon in the USSR was set at 13,500 metric tons. In 1996, permissible (legal) catch was only 1200 metric tons (5).

SEA-LEVEL CHANGE AS A GLOBAL PROBLEM

Given the growing concern about, and possible evidence of, global warming, there has been considerable speculation about the potential impacts on coastal areas of a sea-level rise related to global warming. Scientists who participated in the 1995 Intergovernmental Panel on Climate Change (IPCC) Report (6) suggested that global sea level may well increase by an additional 15–70 cm (6–27 inches) by the end of the twenty-first century. The exact amount of rise would depend on the actual increase in global temperatures. Clearly, any additional increase in sea level could have devastating consequences for coastal communities.

All states that border bodies of water, whether along the global oceans or inland seas, should pay attention to fluctuations in sea level as well as to the rise in sea level linked to global warming. Inland seas, for example, can be viewed as living bodies in the sense that they can expand and can shrink. These changes can occur on different time scales: from daily to seasonally, from a year to a decade, or a century, or a millennium. In fact, they fluctuate and change on all these scales. The same can be said of the open oceans, but they tend to fluctuate on much longer time scales than do the inland seas, over periods of many decades and centuries. Such time scales are difficult to factor into the thinking of economic development planners, whose time frames are on the order of years to a few decades at most.

In essence, one can consider the Caspian as a laboratory of sea-level change and its potential societal and environmental consequences. For the Caspian to serve as a true “laboratory,” its environmental-monitoring network, which collapsed with the breakup of the Soviet Union, must be restored and maintained by regional cooperation among the Caspian states. Impacts on ecosystems that are managed (farms and pastures) and unmanaged (wetlands, forests, deserts) can be identified. Effective human responses to changes in the coastal zone (both land and sea) can also be identified and assessed; environmental engineering proposals to deal with sea-level changes (such as seawall construction, higher oil platforms in the sea, diversion of water from the Caspian to the drying Aral Sea) can be evaluated for effectiveness, taking into consideration the scientific uncertainties surrounding sea-level fluctuations.

Whether the global climate gets warmer, cooler, or stays as it has been for the last several decades, the level of inland seas will likely continue to fluctuate (the mean ocean level has already gone up by 5–6 inches in the twentieth century alone). Societies must learn to cope with both short- and long-term fluctuations. In the Middle Ages, people in the Caspian region were not allowed to settle too close to the sea’s shore, under the threat of death. Apparently, leaders were then aware of the dangers that the Caspian’s fluctuating levels posed to their citizens. Today’s leaders would be well advised to pay attention to traditional wisdom.

BIBLIOGRAPHY

1. Kosarev, A.N., and Yablonskaya, E.A. (1994). *The Caspian Sea*. SPB Academic Pub., The Hague.
2. Golitsyn, G.S. (1989). Once more about the changes of the Caspian Sea level. *Vestnik AN SSR* **9**: 59–63.
3. Golitsyn G.S., and McBean, G.A. (1992). Changes of the atmosphere and climate. *Proc. Russian Acad. Sci. Geogr. Ser.* **2**: 33–43.

4. Ragozin, A.L. (1995). Synergistic effects and consequences of the Caspian Sea level rise. In: *Proceedings of the International Scientific Conference: Caspian Region: Economics, Ecology, Mineral Resources*. Geocenter-Moscow, pp. 120–121 (in Russian).
5. Rosenberg, I. (1996). Catching sturgeon. *Itogi Magazine* 4 June (in Russian), pp. 48–50.
6. IPCC. (1996). *Climate Change 1995: The Science of Climate Change*. Cambridge University Press, Cambridge, UK.

INSTITUTIONAL ASPECTS OF WATER MANAGEMENT IN CHINA

Rongchao Li
Delft University of Technology
Delft, The Netherlands

INTRODUCTION

Located in East Asia, on the western shore of the Pacific Ocean, the People's Republic of China (PRC) has a land area of about 9.6 million km^2. China, the most populous country in the world, had 1.25909 billion people at the end of 1999, about 22% of the world's total. The population density in China is 130 people per sq km. A map of China is shown in Fig. 1.

China has a marked continental monsoonal climate characterized by great variety. The summer monsoons last from April to September. The warm and moist summer monsoons last from April to September, bringing abundant rainfall and high temperatures from the oceans, and little difference in temperature between the south and the north. China's complex and varied climate results in a great variety of temperature belts and dry and moist zones.

China has 50,000 rivers that cover catchment areas of at least 100 square kilometers, and 1500 of them cover catchment areas of more than 1000 square kilometers. Most of the rivers flow from west to east to empty into the Pacific Ocean. The main rivers include the Yangzi (Changjiang), the Yellow River (Huanghe), Heilongjiang, the Pearl River, Liaohe, Haihe, Qiantangjiang, and Lancang Rivers. At 6300 kilometers long, the Yangzi is the longest river in China. The second longest is the Yellow River at 5464 kilometers. The Grand Canal from Hangzhou to Beijing is a great water project of ancient China, 1794 kilometers in length. It is the longest canal in the world (*Source*: http://www.chinatour.com/countryinfo/countryinfo.htm#RIVER).

INSTITUTIONAL ASPECTS OF WATER MANAGEMENT IN CHINA

Legislative and Executive Structure

The agencies involved in WRM can be divided into legislative and executive organizations. The legislative structure is organized, in descending order, as follows:

- state level: state congress
- provincial level: provincial congress

Figure 1. Map of China (*Source*: http://www.maps-of-china.com/china-country.shtml).

- city level: municipal congress
- county level: county congress

The legislative structure in China is composed of four levels, shown in Fig. 2.

The executive structure closely relevant to water management is shown in Fig. 3.

Large water projects are managed by Basin Administrative Institutions that are subordinations of the MWR, such as the Yellow River Water Conservation Commission (YRCC). Medium and small water projects are managed by provincial, municipal, and county water administrative institutions, respectively, on the basis of ownership.

FUNCTIONS OF ADMINISTRATIVE INSTITUTIONS

In China, the water resource management organization has traditionally been organized along narrow, subsectoral lines. The national government determines the general legislative and policy framework for water management. The Ministry of Water Resources (MWR) is responsible for the state water management and issues the national water legislation and policies. Other ministries, such as land and agriculture and electronic, are also involved in water management.

The following competent departments under the State Council and local government and their respective functions based on the duty division of the State Council compose the administrative structure of water resource management in China (Table 1).

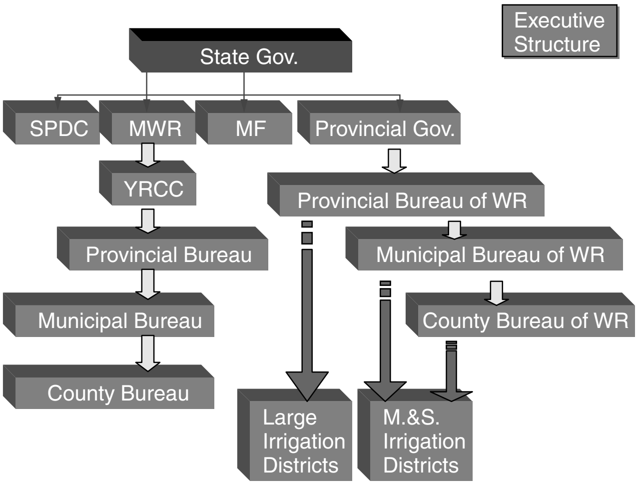

Figure 3. Executive structure of water management.

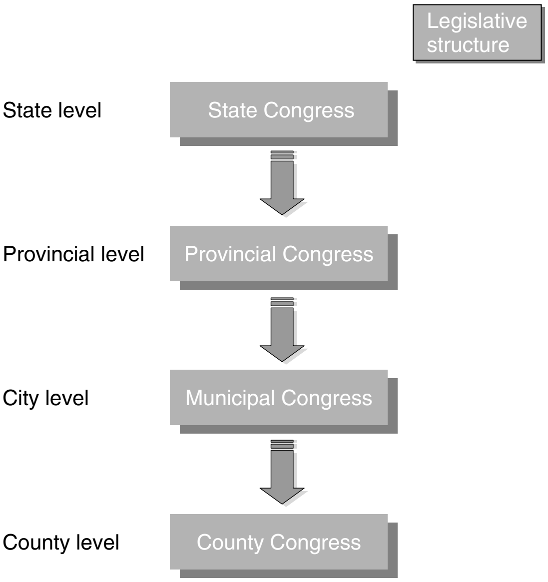

Figure 2. Legislative structure in China.

Ministry of Water Resources

The Ministry of water resources (MWR) is the main department of water resources management under the state council. Its functions are :

- to establish respective regulations of water resources management on water law

Table 1. The 14 Sectors Involved in Water Management in China

Index	Sector	Task
1	Ministry of water resources	Surface water resource management
2	National environmental protection bureau	Water environmental protection
3	Ministry of geomineral	Underground water management
4	Ministry of construction	Construction of city water resource development and protection
5	Ministry of agriculture	Agricultural water management
6	Ministry of energy	Water and electronic construction
7	National planning and development commission	Permits for conservancy projects
8	Ministry of sanitation	Monitor and protect drinking water
9	Ministry of finances	Permits for flood control finance capital
10	National scientific commission	Management of water resource research
11	National meteorological bureau	Prediction and management of precipitation
12	Ministry of transportation	Management of inland navigation
13	Bureau of national land management	Project management of protected basins
14	Ministry of forestry	Protection of river basin forestry

- to carry out a national comprehensive scientific survey and investigation as well as an assessment of water resources
- to draw up the overall planning for the main river basins in China
- to formulate a long-term water supply plan on national and regional scales
- to issue permits for water use
- to construct and maintain water projects such as reservoirs, dikes, and weirs
- to manage the waterways, including rivers, lakes, canals, and flood discharge areas

Generally, there is a department or bureau of water resources under each level of local government which is responsible for local water resources development and management.

To strengthen the development and conservation of water resources in the large rivers in China, the Yangtze, Yellow, Pearl, Huai, Hai, Songhua, and Liao, there are seven river basin commissions established under the leadership of the MWR. These river basin commissions as agencies of the MWR are responsible for the tasks mentioned above in the basin concerned.

National Environmental Protection Agency

The National Environmental Protection Agency, the environmental protection department of the State Council, exercises unified supervision and management for the prevention and control of water pollution. It is in charge of the following responsibilities of water quality management:

- to establish the implementation rules and regulations of water quality management on the basis of prevention and control in the water pollution law
- to establish national water environmental quality standards
- to establish national pollutants discharge standards in accordance with national water environmental quality standards and the country's economic and technological conditions
- to amend national water environment quality standards and national pollutants discharge standards in due time
- to examine and approve the environmental impact statement and inspect water pollution prevention and control facilities of new projects, extensions, or reconstruction projects which discharge pollutants into waterbodies directly or indirectly
- to carry out the task of registration and application for permits for water pollution discharge
- to carry out the task of payment for water pollutant discharge
- to implement elimination and control of water pollution within a certain period
- to exercise the punishment authority of warning and fines on violators of prevention and control in the water pollution law
- to be in charge of pesticide registration

Ministry of Geology and Mineral Resources

The functions of the Ministry of Geology and Mineral Resources are:

- general investigation and exploitation of groundwater
- assessment and unified planning for the development and use of groundwater
- supervision and management of groundwater

Ministry of Agriculture

This ministry is in charge of the management and conservation of farmland irrigation water and fishery waterbodies.

Ministry of Public Health

- supervision and management of drinking water quality in urban and rural areas
- formulation of sanitary standards of drinking water
- monitoring drinking water quality
- definition of sources of drinking water in accordance with other competent departments
- participation in the management of protection zones for drinking water

Ministry of Construction

Construction and management of the water supply in urban, and rural areas and sewerage engineering as well as domestic wastewater treatment plants in cities.

Ministry of Transportation

Supervision and management of water quality for navigation and navigation ways in inland rivers.

In addition to the responsibility of the various Ministries under the State Councils as mentioned above, the responsibility for flood control belongs directly to the State Councils and local governments at various levels as well as the seven large river basin commissions.

LEGAL SYSTEM

In 1988, the Water Law of China was enacted and issued, representing the legal basis for water resources development and management. At present, other relevant laws or regulations have been enacted for effective water pollution control, water and soil conservation, the water-drawing permit system, water conservation in urban areas, flood control, safe management of dams, tariff collection, riverbank management, navigation management, land use for water works, and resident resettlement. A series of laws and regulations related to water resources development and management is also being drafted and processed for enactment, concerning integrated use of water resources, watershed management, flood control, water resources fee collection and management, water pricing of water supply projects, totaling more than 20 (see below).

List of Chinese Laws Related to WRM

Laws

- Constitution of the People's Republic of China (DOE*: 1982, RA**: National People's Congress)
- Forest Law of the People's Republic of China (DOE: 1984, RA: National People's Congress)
- Fishery Law of the People's Republic of China (DOE: 1987, RA: National People's Congress)
- Water Law of the People's Republic of China (DOE: 1988, RA: National People's Congress)
- Environmental Protection Law of the People's Republic of China (DOE: 1989, RA: National People's Congress)
- Water and Soil Conservation Law of the People's Republic of China (DOE: 1991, RA: National People's Congress)
- Mineral Resources Law of the People's Republic of China (DOE: 1994, RA: National People's Congress)
- Water Pollution Control Law of the People's Republic of China (DOE: 1996, RA: National People's Congress)
- Flood Control Law of the People's Republic of China (DOE: 1997, RA: National People's Congress)
- Land Administration Law of the People's Republic of China (DOE: 1998, RA: National People's Congress)

Regulations and Policies

- Working Regulation of Soil Conservancy (DOE: 1982, RA: State Council)
- Price Management Regulation (DOE: 1987, RA: State Council)
- Law on Land Administration (DOE: 1987, RA: State Council)
- Water Extraction Permit System Regulation (DOE: 1993, RA: State Council)
- Urban Water Supply Regulation (DOE: 1994, RA: State Council)
- Method for Management of Irrigation Districts (DOE: 1981, RA: Ministry of Water Resources)
- Method for Construction and Management of Irrigation and Drainage Projects (DOE: 1981, RA: Ministry of Water Resources)
- Technical and Economical Index of Irrigation Management (DOE: 1996, RA: Ministry of Water Resources)
- Industry Policies for the Water Sector (DOE: 1997, RA: Ministry of Water Resources)
- Urban Water Conservation Management Regulation (DOE: 1989, RA: Ministry of Construction)

*DOE: date of enactment;
**RA: responsible agency.
(*Source*: The government of the PRC and the Asian Development Bank, 1999)

FINANCIAL SOURCES FOR WATER MANAGEMENT IN CHINA

The main financial source for water management is composed of three parts: the infrastructure fund of the water conservancy construction fund, maintenance of projects, and institutional management fees. The sources of the infrastructure fund for water conservancy are:

- the infrastructure fund of the general budget
- the water conservancy construction fund
- loans from domestic banks and other financial organizations
- funds from the sale of bonds
- loans from foreign financial organizations
- other sources approved by state governments

Both the sources for maintenance of projects and institutional management fees are the general budget and the water tariff. Spending of the infrastructure fund for water conservancy is on flood control, drainage, irrigation, water supply, hydropower, navigation, and others. The Financing chart for water management of the YRCC is shown in Fig. 4.

POLICY, LAWS AND REGULATIONS

The Central Government lays the utmost stress on the Yellow River. A series of principles, policies, general water laws, water administration regulations, and basin-oriented water laws have been developed. They involve the water administration of the Yellow River, the management of water resources, management of a water conservancy project and engineering affairs of the Yellow River, management of flood control, water and soil conservation, and protection of water resources. Examples are given below.

Water Laws such as the "Water Laws of the People's Republic of China," "Laws of Water Pollution Prevention of the People's Republic of China," "Laws of Water and Soil Conservation of the People's Republic of China," and "Laws of Flood Prevention of the People's Republic of China"

Laws and Regulations of Water Administration: "Working Regulations of Water and Soil Conservation," "Management Regulations of River Channels of the People's Republic of China," "Management Means of Ratification and Collecting of Water Fees for a Water Conservancy Project," "Regulations of Water and Soil Conservation in Developing the Contiguous Areas of Shanxi, Shaanxi, and Inner Mongolia," "Collecting and Management Means of Water Fees for the Intake of the Irrigation Canal," "Means of Fee Collecting and Management of Sand Mining in River Channels," "Implementation Means of Permit Regulation of Abstracting Water," and "Means of Control and Management of Water Quantity"

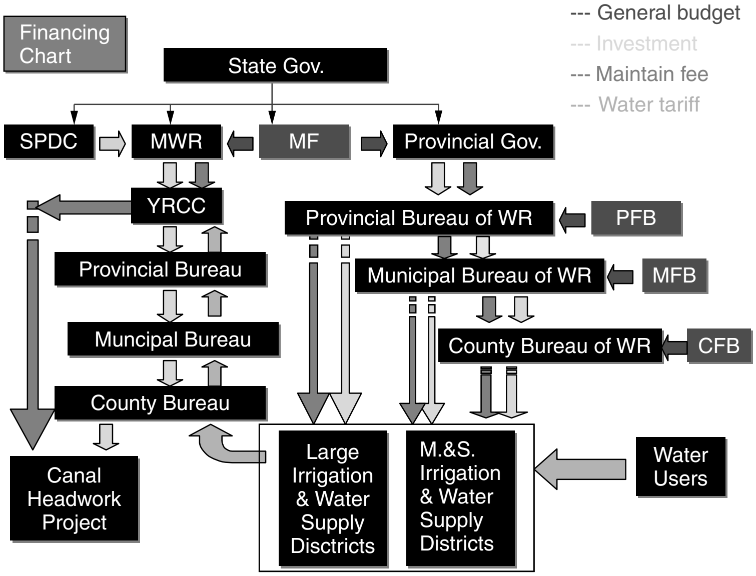

Figure 4. Financial chart for water management of the YRCC.

In addition, there are many local laws and regulations relevant to the management and development of the Yellow River. For example, "Management Means of Water Conservancy Project of the Yellow River of Shandong Province," and "Regulations for Certain Problems in River Channels and Flood Detention Areas in Henan Province."

READING LIST

General Information of the People's Republic of China: http://www.chinatoday.com/general/a.htm#POPU

WILL WATER SCARCITY LIMIT CHINA'S AGRICULTURAL POTENTIAL?

Bryan Lohmar
Economic Research Service,
U.S. Department of Agriculture

Water shortages in important grain-producing regions of China may seriously compromise China's agricultural production potential. Rapidly increasing industrial and domestic water consumption and expanding irrigation have drawn down groundwater tables and disrupted surface-water deliveries. The problem is most severe in north-central China, where most of China's wheat and cotton is produced and irrigation is essential to maintaining high yields. The situation may worsen unless effective water conservation policies can be put into place rapidly.

China is responding to these concerns on several levels. At the national level, the Ministry of Water Resources began promoting water conservation through various measures in the late 1990s, such as strengthening the authority of National River Basin Commissions to enforce water withdrawal limits and promoting irrigation management reforms. Provincial and other local officials are mediating conflicts between users to improve overall water management. In villages, local water managers and farmers are adopting water management reforms and water-saving techniques, such as forming water user associations and alternating wet-dry irrigation for rice. In addition, reforms in the pricing and fee collection system may provide farmers with better incentives to conserve water. Pricing water deliveries to farms based on volume could improve efficiency, but would be costly to monitor since China has over 200 million farm households, each tending several tiny plots of land.

As water for agriculture becomes more scarce, changes in the pattern of crops are more likely than a reduction in cultivated acreage. Wheat is most likely to suffer declines, since wheat is irrigated in much of north China and brings low returns to water. Production of a variety of crops—corn, cotton, and high-value fruits and vegetables—may increase as farmers switch from irrigated wheat. High-value fruits and vegetables are often more water intensive, but are also more suited to water-saving irrigation technologies, such as drip irrigation and greenhouse production.

The success of current efforts to encourage water conservation in China will depend on a variety of factors. Policy reforms will depend critically on the enforcement of withdrawal limits both from surface-water systems and from ground water. Also important is the extent to which policies and local management practices motivate water users and water managers to conserve water resources.

WATER AND COASTAL RESOURCES

U.S. Agency for International
Development (USAID)

Every ecosystem, society, and individual on Earth depends on water. Food security and human health, energy supplies, and industrial production are all dependent on water to a large degree. Water plays an important role in regulating global climate. It is essential for plants and wildlife and ecosystems they inhabit. Water shortages and poor water management can lead to loss of biodiversity and agricultural production, increase in malnutrition and disease, reduced economic growth, social instability, and conflict.

Reliable water supply has been a fundamental component of the increased agricultural productivity achieved by the green revolution. Irrigated agriculture currently uses 70% of the world's developed water supplies. With increasing competition for water to meet domestic and industrial needs, and for servicing essential ecosystems, agriculture is faced with producing more food with less water.

Water resources are not distributed uniformly across the globe—nor are they necessarily located where the largest concentrations of people reside. Demand for water outstrips supply in a growing number of countries, and the quality of that supply is rapidly declining: 450 million people in 31 countries already face serious shortages of water. These shortages occur almost exclusively in developing countries, which are ill-equipped to adopt the policy and technology measures needed to address the crisis. By the year 2025, one-third of the world's population is expected to face severe and chronic water shortages.

Exacerbating the problem, human activities often contaminate the world's limited freshwater resources, making them unavailable for further human use and threatening the health of the lake, river, and wetland ecosystems they support. Likewise, coastal and ocean systems are under threat from the impact of a broad range of human activities. Coastal systems are particularly vulnerable to degradation from land-based activities, climate change, over-fishing, and damage to coral reefs, and they require active intervention to ensure their continued survival.

The U.S. Agency for International Development has made the preservation and environmentally sound development of the world's water resources a top priority. The Agency's investments have helped improve access to safe and adequate water supply and sanitation, improve irrigation technology, enhance natural environments, and develop better institutional capacity for water resources management in countries around the world (Fig. 1). This has supported the Agency's underlying goals of reducing conflict and improving the welfare of people across the globe.

Figure 1. USAID worked with local partners to improve water resources management in Morocco's Nakhla region of the Rif Mountains.

WATER USE CONSERVATION AND EFFICIENCY

Mario O. Buenfil-Rodriguez
National University of Mexico
Cuernavaca, Morelos,
Mexico

WATER USE AND USERS ON GLOBAL AND REGIONAL SCALES

Water has many uses and users, ranging from fluvial navigation to employing water as a component in a chemical experiment. On a global level, the four most water demanding human activities are energy—hydroelectricity—generation, irrigation for agriculture, urban supply, and nonurban industrial production. Each of them requires huge quantities of water; and in its usage at least some water characteristic is affected, which limits the options of the remaining users to profit from that same water. This causes a competition in which each user would like to have exclusive rights to manage and exploit the resource. Hydroelectricity affects the potential energy of water (topographic elevation), and the other three basically affect water quality. There is a fifth user, frequently neglected by humans, which has the most ancient right to benefit from water, nature. The natural world needs water in lakes, forest, woods, floods, and climate regulation and has been greatly affected by the human desire to conquer and transform and to grow and expand.

Worldwide, the agricultural sector is the biggest water consumer, although this balance is not equal in every continent or country. For instance, nations that have benign and frequent rains that allow good direct agricultural production may have no need for irrigation; others that are very dry or have an uneven distribution

---

of water and rain require huge irrigation works to get reasonable crop yields. Figure 1 depicts the distribution of volumes among the three human uses that most affect water quality.

WATER SCARCITY, CONFLICTS, TECHNOLOGY, AND CONSCIOUSNESS

Water scarcity, which is rapidly growing in most countries, creates serious conflicts, competition for water, and evidently great unfairness in many cases. Figure 2 is a clear and simple illustration of why there is growing scarcity in most nations. Of course, its severity differs widely, depending on regional natural resources and demographic pressure. In general, there are growing crises in the whole world.

Legal rules establishing some priorities for water usage, or simply the legal support to historical privileges, sometimes help in attenuating real or apparent conflicts among users. Other regulatory mechanisms are water markets, where users can trade their water rights to other users. This last has advantages, but also great disadvantages when the market and the world economy (usually a totally artificial issue subject to political fancies and fashions) send erroneous signals about the value of things.

Competition among users is not absolute; water used upstream can later be freely used downstream, provided the residual water discharges from the first consumers are not dangerous enough to inhibit their later usage. Anyway, there always is some technological possibility of treating and reusing water. If there were enough money to apply technology, conflicts would be eliminated, but technology is costly, and most of the world is poor. This is worsened by the low value given to water by traditional economies and by the unequal competition for globalized trade in agricultural products. So unfairness and crises still have a long way to go.

An often misconceived economics paradigm suggests that water should be used where it renders the greater rate of return. This is a kind of taboo to be seriously questioned and reviewed because it leads us to accept the idea that uses in services such as tourism, electronics, and industry in general are more important than use in agriculture and food. A better approach is to accept that the latter use is greatly undervalued by the market and the others are exaggeratedly overvalued. The same situation occurs in ecological water use.

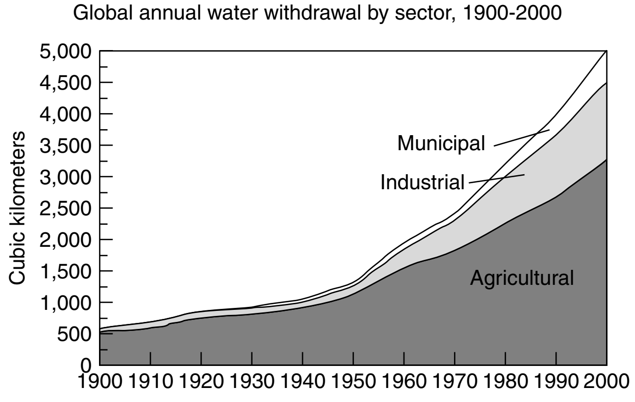

Figure 1. Rising water use. Global annual water withdrawal by sector, 1900–2000 (from Ref. 1).

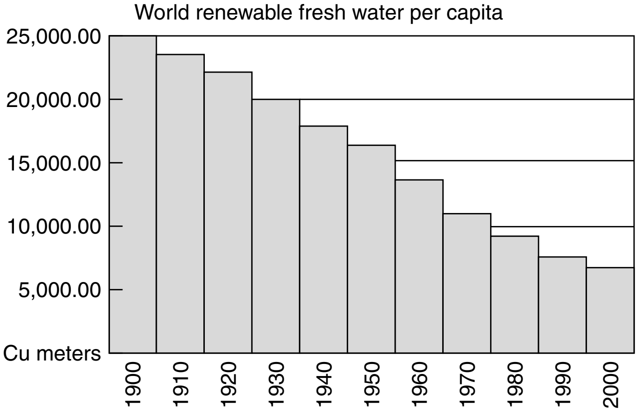

Figure 2. A century of changes in water availability: clear signs of scarcity and conflicts.

The mid-twentieth century vision of dealing with water needs and scarcity problems was to increase the supply by means of big dams and similar huge engineering works. They have proved insufficient and merely made the demand disproportionately grow, so the vision for this century must be radically changed. The approach now must be "water demand management" respecting natural limits and boundaries. Scarcity must be directly dealt with, by conservation and "efficient water use." These options also have a technological component but rather rest mainly on human habits and consciousness. They are radically different from options that rely on huge engineering works for water transfers and treatment.

WATER USERS AND DEMAND ON AN URBAN SCALE

Water users in cities are mostly the people who live and work there; they can be represented by existing houses, buildings, shops, offices, parks, hotels, schools, and factories. There are many types of cities and ample variations in these components. Tourist and recreational cities have lots of hotels and almost no factories; in industrial cities, it is the opposite. Housing for families usually is the greater component in any city and evidently the sector demanding more water.

The water demand in a city can be expressed in liters/person/day; it is comprised of the water used directly or indirectly by an average citizen and the leaks in the piping network. People use water directly in their houses, but also in their jobs or at school, and indirectly by others serving them, as in restaurants where they eat, by shops or services visited, and in general by every industry in the city. Leaks unavoidably occur so the water may arrive in the consumption area. The previous reasons explain why if the average in-house *consumption* is 300 L/capita/day, the full *demand* may be around 500 L/capita/day. Figure 3

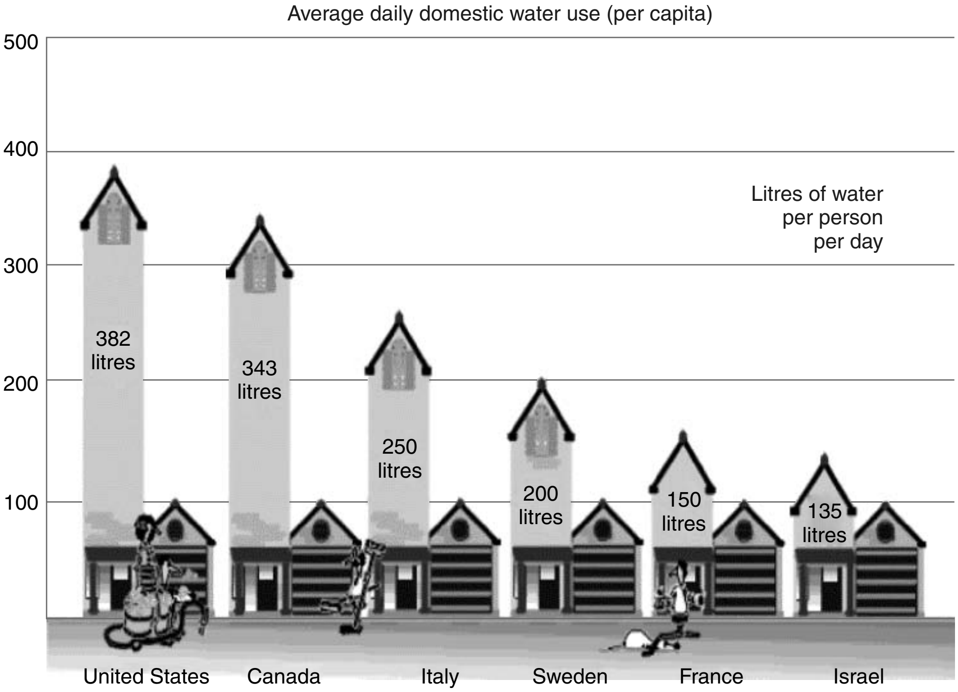

Figure 3. House water consumption in some countries (from Ref. 2).

contrasts some international statistics about in-house consumption. There are big variations, of course, even within a nation and even among different zones and persons in the same city. Figure 4 is a sample of in-house water usage distribution.

CONSERVATION OF WATER AND OTHER RESOURCES

Water emphasizes many conflicts and scarcities, although water may not be the real issue considering that there is plenty of water on our planet. Water quantities may not be the problem, but its quality, or its location in time or in space. Infrastructure, investments, and technology are always essential for actions such as water treatment, storage for other seasons, or transport to other places. Therefore, the real restrictions may not be water volumes or qualities, but those regarding energy resources, financial sources, and managerial capabilities to transport it or to preserve its purity. Energy is one important inhibitor; if it were freely available, water could be pumped from far away flood zones to deserts (this does not consider possible negative ecological impacts).

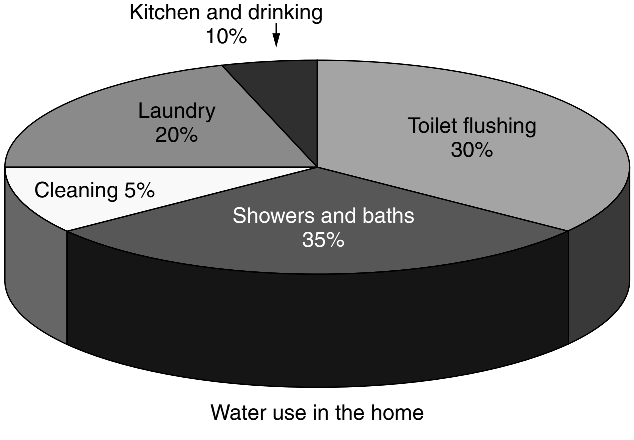

Figure 4. Typical household water use in Canada (from Ref. 2).

When institutions request people to save water, they do not merely think of saving water volumes, but rather of saving energy expenditures by eliminating the need for larger pipes and, mainly, by diminishing pollution caused by the water usage and discharge. Many propose conservation plans, so the volume saved is used to support demographic and demand growth. In this case, there are no real savings and compensations for previous damage and overexploitation, but merely variants of the twentieth century backing of human expansionism.

The need for conservation and for “efficient water use” has a much broader scope than merely using less water. The real purposes and issues may be protecting the whole environment, our economy, and the quality of life of future generations. They can have a rather dark side when used as a political and demagogic argument simply to maintain inertias for growth and expansionism, which require less money investments but possibly greater human and environmental sacrifices and discomforts that are not taken into account.

MEASURING THE EFFICIENCY OF URBAN SERVICES

Even in something as specific as public water supply, efficiency has many angles and goals. Simultaneously, it is important to save water volumes, to avoid pollution, to diminish unnecessary expenditures, and to distribute water more fairly among different users. For instance, if there were an optimal amount of water to be given to people, considering health, welfare, and other aspects (for example, 150 L/person/day), anything above or below such a quantity could be considered inefficient.

One common type of inefficiency occurs in water transportation because there always is some degree of leakage caused by breaks in the piping system. This efficiency is computed as the ratio of the volume delivered to clients, measured by micrometers, divided by the total volume extracted from the water source. The operating or "revenue efficiency" is the total money collected through bills paid by clients, compared to the total amount in issued bills. The "unaccounted for water" (UFW) indicator is a measure of the overall inefficiency of a water supplier, considering physical and commercial losses. Another strategic indicator is represented by the ratio of total expenditures to total income during a certain period.

Yet another important way to assess performance is by comparing the initial water quality at the source against the quality returned after discharge. An increasing degree of recycled or reused water within a city evidently is also a desirable way to demonstrate improvements.

It would be extremely expensive and totally impractical to have a system without inefficiencies, that is, 100% efficiency in all aspects. What is done by developed water utilities is to work simultaneously using an ample set of "performance indicators" and to apply several benchmarking rules and comparisons with other utilities, as suggested by the International Water Association (3).

Just as any water utility may have various types of efficiencies to be confronted routinely and evaluated, the private home owner, or any factory or hotel manager, may have many parameters by which to assess performance in water usage and its impact on the environment. Water audits (Fig. 5) are appropriate as systematic procedures for assessing performance and to detecting failures as well as opportunities for improvement.

CONSERVATION AND EFFICIENCY, THE NEED AND THE HOW

The need for conservation and efficient water use comes from scarcity and competition for the resource, each time more evident in many places. In this regard, many national and international institutions advise how to save water. The most classical are to take shorter showers, not to leave the faucet open while shaving, or changing the old toilet for one of newer design (maximum 6 liters/flush, see Fig. 6). There are lots of these recommendations, which are not bad at all, but maybe at sites of severe crises and scarcities, recommendations should be even more daring and definitive, to attack the roots of problems, not their symptoms. Some of these newer counsels are: use dry latrines instead of toilets, shower just once a week, and use perfumes; be a vegetarian and not a meat eater (meat production demands much more water than vegetables); family planning; bicycles instead of cars; reuse, recycle, rethink; charge "green taxes" for many environmentally damaging products or packaging.

The importance of systematically assessing and measuring efficiency lies in finding improvement opportunities. Performance or efficiency indicators should be compared against reference values to detect failures, changes, and trends. The reference standards should be carefully discussed and selected, because if not, some incongruence and evident mistakes may occur. One pitiable flaw would be to report efficiencies greater than 100% when abstaining from using water in certain industrial process or eliminating a superfluous water need. For instance, when comparing the change to dry latrines after using traditional toilets, it evidently does not mean infinite efficiency, although doubtlessly it is great environmental progress.

Excessive, wasteful, and polluting water uses should be restricted. However, how these restrictions should be

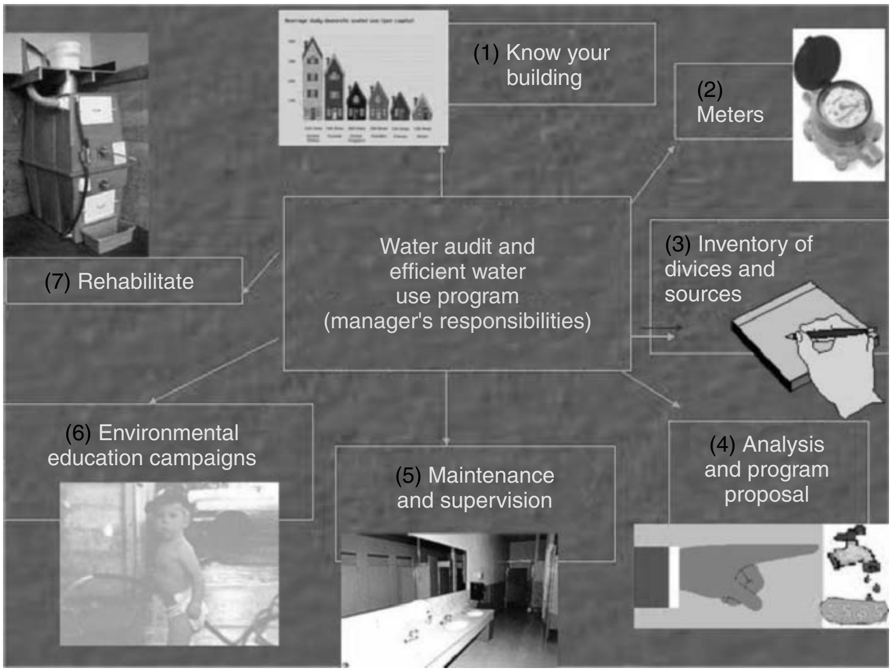

Figure 5. Seven stages for water-saving programs in buildings.

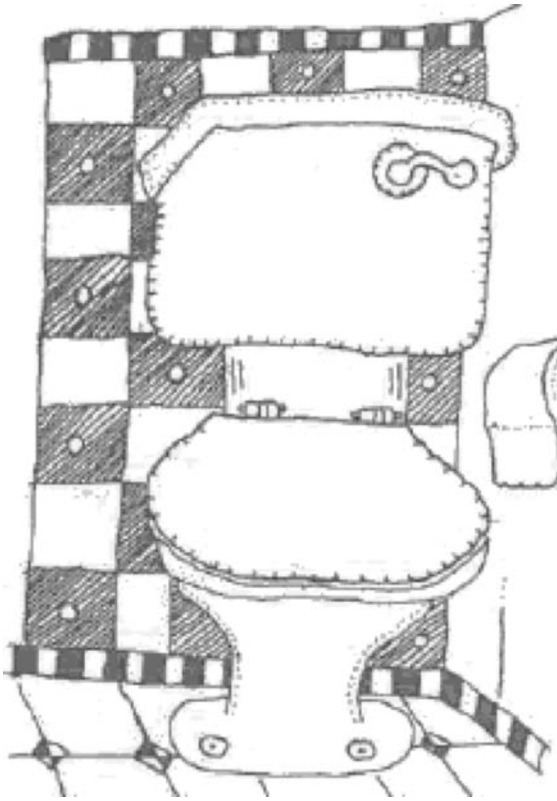

Figure 6. Toilet bowl: friend and foe. The most frequent cause of water consumption, leakages, and pollution.

attained is an interesting issue. The least desirable options would be laws for direct prohibitions. There are better choices, for instance, measuring water and contaminant loads in discharges and applying tariffs high enough to discourage those uses and encourage the search for better technologies and processes.

Besides metering and tariffs, there are many programs and actions that any water supplier must consider, which contribute to efficient water use in a city. Among them are user education and water culture, macrometering, information systems, client census, demand forecasting, automation of operations, leak detection and control, materials testing and quality control, employee training, improved maintenance procedures, bylaws, aquifer recharge, rain harvesting, water audits, improved domestic water-saving devices, water treatment, water reuse and recycling, and desalination. Each of these programs and tasks has different degrees of importance depending on the local conditions, and together make up a desirable set of simultaneous actions to be considered and used.

This article supplies only a list of these programs without further details. Each one deserves a specific and detailed explanation, which hopefully may appear in other sections of this Encyclopedia of Water. All of them are important and strategic for efficient water use and resource conservation.

LIMITS TO GROWTH AND EFFICIENCY

The options for solving water scarcity, along with the crises and conflicts it causes, rest mainly in the social, moral, and economic realm, rather than in the technical one, although at first glance, it may appear the other way.

Most likely, technology can always find a way to improve efficiency, but also it implies that costs sometimes are not apparent. Solutions often have clear and immediate economic components but often also may have vague, but possibly important, environmental, quality of life, and long-term consequences. Unfortunately, traditional economic and technical models merely deprecate and do not consider such impacts because they do not know how to deal with them.

Technological options for improving water efficiency sometimes are clear and simple, and can be adopted without much questioning and doubt; however, in many other cases, it is important to have broader vision. Here is a brief list of arguments to reflect on and consider:

— Efficient water use is not simply “using less water.” There are many other options for managing demands. Reuse is an important one.

— Efficient water use depends more on the efficient use of resources other than merely on water saved. Energy, money, technology, people’s willingness, forests, and climate are some of these other resources.

— Urban water scarcity is not a water availability problem. It may be a symptom and reflection of other problems so that, when confronted, the scarcity issue will be immediately relieved.

— Scarcity problems could be “easily” solved by stopping consumption and population growth.

— Urban water scarcity is a “moral and ecological problem,” not a technical one.

— More severe and urgent laws are required for sites suffering extreme scarcity, particularly, where aquifers present continued water table drawdown.

— Scarcity is referred to as a problem of shortages but is a situation of “longages” (excesses) of bad habits and population or industrial growth in inappropriate zones.

Technology can succeed in allocating population increases by using and reusing the fixed available water volume more intensely. However, this style of technology application merely justifies expanding the coverage, without a clear improvement in the quality of life. Reuse, treatment, and desalination, are desirable options up to certain limits, but should never be used indiscriminately to justify further growth. On the contrary, they must promote diminishing human impacts to make more natural space available. Theoretically, it is possible to have houses or factories with zero water discharge, and almost zero water demand by intensive treatment and reuse.

Water desalination is an interesting option for increasing the supply of water. It has a history as ancient as the hydrologic cycle, its basis. Humans have used desalination techniques for survival on long ocean journeys and for supplying isolated islands for centuries. Nowadays, desalination processes use reverse osmosis that demands huge energy input, which mainly comes from thermoelectric power plants; so the process is equivalent to exchanging petrol for water. In that sense, besides being a rather expensive choice not for every one, it is not a sustainable option because fossil energy is a nonrenewable resource. It would be more desirable and sound to research and promote technology using solar energy for these processes.

Now that the failures and flaws of traditional economic models that cheer growth and expansion are evident, there is need for a new paradigm. Efficient water use and conservation should be seen as a route to reach this new paradigm, not as a means for reinforcing the old one.

When continuous, long lasting, and growing sacrifices or "improved efficiencies" are the alternative to make available water to increasing populations and when there is not a clear end to these situations, serious questioning must be done about the risk of failure of this route (plagued by decreasing safety levels) and of the worth of such sacrifices. It could be that opting to limit growth and technology would be more appropriate, sound, and simple.

BIBLIOGRAPHY

1. Hinrichsen, D., Robey, B., and Upadhyay, U.D. (1997). Solutions for a Water-Short World. *Populations Reports, Series M*, No. 14 Baltimore, Johns Hopkins School of Public Health, Population Information Program, http://www.jhuccp.org/pr/m14edsum.stm.
2. Environment Canada (1998). *The Management of Water*. National Water Issues Branch, Environment Canada, http://www.ec.gc.ca/water/en/manage/e_manag.htm.
3. Alegre, H., Baptista, J. M., Hirner, W., and Cabrera, E. (2000). *The IWA system of performance indicators for water supply services* International Water Association IWA, manual of best practices (performance indicators). http://www.iwapublishing.com/template.cfm?name=isbn1900222272.

READING LIST

Arreguín, C. F. and Buenfil, R. M. (1997). *64 Recomendaciones Para Ahorrar Agua*, IMTA- Ministerio de Desarrollo Económico de Colombia. http://www.cepis.ops-oms.org/bvsadiaa/e/noticias/colombia/agua.pdf.

Asian Development Bank (McIntosh A., Yñiguez, C.) (1997). *Second Water Utilities Data Book. Asian and Pacific Region*. Asian Development Bank.

Barker, P. (1977). *Cost-Based Water Prices*, 23 WEDC conference, Durban South Africa, http://www.lboro.ac.uk/departments/cv/wedc/papers/23/groupf/barker.pdf.

Brown, L. (2001). *Eco-Economy: Building an Economy for the Earth*, Earth Policy Institute, http://www.earth-policy.org./Books/Eco_contents.htm.

Brown, L. (2001). *Worsening Water Shortages Threaten China's Food Security*. Earth Policy Institute, http://www.earth-policy.org/.

Buenfil, M. (2001). *Selfishness and Efficiency (in Water Use)*, 4th Information Water Summit, Panama. http://www.waterweb.org/wis/wis4/cd/Buenfil-S&E%20water%20use.pdf.

Environment Canada (1993). *Manual for Conducting Water Audits and Developing Water Efficiency Programs at Federal Facilities*. Environment Canada, water issues branch, http://www.ec.gc.ca/water/en/info/pubs/manual/e_contnt.htm.

EPA. (2001). *Residential Water Conservation Techniques*. Environmental Protection Agency, USA, http://www.epa.gov/seahome/watcon.html.

Gleick, P.H. (2000). *The Changing Water Paradigm. A Look at Twenty-First Century Water Resources Development*. International Water Resources Association. http://www.iwra.siu.edu/win/win2000/win03-00/gleick.pdf.

Gleick, P.H. (2001). Making every drop count. *Sci. Am.* http://www.scientificamerican.com/2001/0201issue/0201gleick.htm.

Hardin, G. (1968). Tragedy of the commons. *Science* http://www.sciencemag.org/cgi/content/full/162/3859/1243.

Hawken, P., Lovins, A., and Lovins, L. H. *Aqueous Solutions, Chapter 11 of Book "Natural Capitalism: Creating the Next Industrial Revolution*. Rocky Mountain Institute http://www.natcap.org, http://www.natcap.org/images/other/NCchapter11.pdf, http://www.rmi.org/sitepages/pid172.php, http://www.rmi.org/images/other/W-AqueousSol.pdf.

Hinrichsen, D., Robey, B., and Upadhyay, U.D. (1997). *Solutions for a Water-Short World*. Population Reports, Series M, No. 14. Johns Hopkins School of Public Health, Baltimore, Population Information Program, http://www.jhuccp.org/pr/m14edsum.stm.

James, W. (1999). *Urban Water Systems*. Guelph University, Canada. http://www.eos.uoguelph.ca/webfiles/james/homepage/Teaching/661/wj661Modules.html.

Mello, E.J. (2001). *Conservação e Medição da Água* various links and articles related to metering and water conservation, http://www.geocities.com/hidrometro/.

Postel, S. (1984). *Water: Rethinking Management in an Age of Scarcity*, Worldwatch Institute paper 62, Worldwatch Institute, Washington USA, http://isbn.nu/0916468623.

Postel, S. (1999). *Pillar of Sand. Can the Irrigation Miracle Last?* WorldWatch book.

Reimold, R.L. and Bloetscher, F. (1996). An innovative opportunity for water reuse. *Fla. Water Resour. J.* http://www.fwrj.com/articles/9607.pdf.

Vickers, A. (2000). *Handbook of Water Use and Conservation*. California Urban Water Conservation Council, http://www.cuwcc.com/publications/action.lasso?-database=cuwcc_catalog&-layout=CDML&-response=detailed_results.html&-recordID=33056&-search.

WaterMagazine (Cayford, Joel). (2002). Water services—regulation & pricing—article links. *Water Mag.* New Zealand, http://www.watermagazine.com/secure/pricereg.htm.

World Bank (Yepes, G. and Dianderas, A.). (1996). *Water and Waste Water Utilities Indicators 2nd Edition*, Water and Sanitation Division, TWUWS, The World Bank, http://www.worldbank.org/html/fpd/water/pdf/indicators.pdf.

CONSERVATION OF WATER

Mukand Singh Babel
Asian Institute of Technology
Pathumthani, Thailand

Water is the main component of the natural ecosystem and an essential element for human life. It is also fundamental for the social and economic development of a country. Water is needed for drinking, washing, and sanitation at personal, domestic, and municipal levels; for agricultural purposes such as irrigation, livestock, and fisheries; for industrial development, navigation, hydropower production, recreation, and so on. In addition, there is increased realization that water is needed for maintaining ecosystems.

The higher rate of population growth associated with several anthropogenic activities leading to increased pollution of water resources has augmented critical situations worldwide related to scarcity of water resources either in quantity or quality to meet the vital necessities of humankind.

When the ratio of water availability to demand exceeds certain limits, the stress on resources becomes noticeable, and the conflicts among users intensify and increase in frequency. A region that have water availability of about 1700 m^3 per capita/year faces periodic water scarcity, and this is considered a warning if the population continues to grow. Below this limit, “water stress” is periodic. If this limit is even lower, about 1000 m^3 per capita/year, the region is under “chronic water scarcity.” Economic development, well-being, and health are affected at these levels. The situation is “extremely scarce” if the per capita availability is below 500 m^3 per year. Many countries that used to have plentiful water resources are facing water scarcity, that is likely to be aggravated in coming years.

On one hand, there is the problem of scarce availability in relation to increasing demands; on the other hand, there is poor resources management and wasteful and inefficient use of freshwater. To address these water problems, it is necessary to control the overuse and pollution of water resources by improved management of water resources through conservation and water saving concepts and practices.

The main water-use sectors are agricultural, industrial, and domestic. At a global level, agricultural water use accounts for 70%, industrial use 21%, and 9% for domestic consumption. Though industrial water demand is much smaller than agricultural demand in low- and middle-income countries, it exceeds domestic and agricultural demands in developed and industrialized countries, as shown in Table 1.

Water is wasted nearly everywhere; there is great potential for better conservation and management, no matter how water is used. The term water conservation is associated mainly with the policy, managerial measure, or user practice aiming at conserving or preserving water resources and taking care of the degradation of water resources; water saving aims to limit or control water demand and avoid wastage and misuse. Both perspectives are complementary and interrelated in managing scarce water resources.

Table 1. Distribution of Water Use[a]

	Agricultural Use	Industrial Use	Domestic Use
World	70%	22%	8%
High-income countries	30%	59%	11%
Low- and middle-income countries	82%	10%	4%

[a] Reference 1.

AGRICULTURAL SECTOR

Most water is used for agriculture (Table 1). Efficient agricultural water management, therefore, is a major conservation option. It is common to find that more than half of the water drawn from a source does not even reach the fields being irrigated. Generally, less than 40% system efficiency prevails in many irrigation schemes. Such an appalling situation should be tackled by pursuing better systems of farm water management—improving irrigation efficiency by reducing water distribution losses, changing cropping patterns, improving irrigation scheduling, and adopting irrigation efficient technologies. Deterioration of irrigation network systems due to lack of proper and adequate maintenance are the main causes of increased losses through leakage and percolation. Rehabilitation of existing irrigation schemes is, therefore, required to alleviate the situation. It is also equally important, when implementing new irrigation schemes, to give sufficient attention to adequate operation and maintenance. Active participation by the users is vital for sustainable management of irrigation systems. Several governments are in the process of transferring management functions to farmers or developing joint (government agency and water users) management systems.

In gravity irrigation methods such as basin, furrow, and flooding, a large quantity of water is lost by leaching down the root zone and as runoff at the end of the field. Land leveling, along with other technical measures, provides conditions for reducing the advance time and water volumes required to complete the advance, for better water distribution as well as application efficiencies. For moderate to high infiltration soils, where system automation is possible, water can also be saved by adopting surge irrigation which consists of intermittent cycling of water applied to furrows and borders, thereby producing changes in the soil surface conditions which favor a faster advance of water.

Sprinkler irrigation methods have higher system efficiencies than gravity irrigation. Water in sprinkler irrigation can be conserved by improving the distribution uniformity—optimizing the overlapping of sprinkler jets, minimizing discharge variations, minimizing wind drift and evaporation losses, maximizing infiltration of applied water, and avoiding surface runoff losses. Microirrigation has the least loss of water among irrigation methods. It can be classified in two general categories: drip or trickle irrigation, where water is slowly applied through small emitter openings from plastic tubing; and microspray irrigation or microsprinkling, in which water is sprayed over the soil surface. Water conservation in microirrigation can be improved by using a single drip line for a double crop row, using microsprayers in high infiltration soils, adjusting the duration and timing to soil and crop characteristics, controlling pressure and discharge variations, adopting automation and fertigation, and chemigation (application of fertilizers and herbicides or other chemicals with irrigation water).

Water conservation measures are also very important in dry-land agriculture because crop growth depends solely on the moisture retained by the soil from rain. Conservation measures in rain-fed agriculture are carried out mainly by appropriate crop and soil management practices. Crop management techniques consist of managing crop risks, such as selecting the cropping pattern to suit the available seasonal rainfall; controlling water stress, such as adaptation of crop rotation to environmental constraints, including covered fallow for grazing; and water conservation cropping techniques, such as by employing conservation tillage, windbreaks to decrease wind impacts on evaporation, etc. Soil management practices are broadly categorized as runoff control, such as contour tillage, using mulches, furrow diking, etc; improvement of soil infiltration rates, such as application of organic matters and mulching; increasing soil water storage holding capacity, such as deep tillage; control of evaporation from soil by crop residue and mulching; runoff control in sloping areas by terracing, contour ridges, and strip cropping; and water harvesting.

Reuse of water, such as using wastewater (graywater) for agriculture, is considered a conservation option to increase the availability of water either currently or in the future. More than 500,000 ha of farmland are currently being irrigated with treated wastewater in 15 countries around the world. However, the use of such wastewater, if not treated well, may threaten the public health by spreading bacteria, viruses, and parasites. For example, the outbreak of typhoid fever in the mid-1980s in Santiago, Chile, was attributed to the use of wastewater, with unsatisfactory sanitary quality, to irrigate 16,000 ha of horticultural farms (2).

The major challenge to water planners and managers is that physical availability of water is fixed, but demand will increase to cater to the needs of an ever increasing population. The tendency today is to shift from water policies based on water supply management to those based on demand management. Agricultural water demand management consists of reducing crop irrigation requirements, adopting irrigation practices that lead to higher irrigation performance and water saving, controlling system water losses, and increasing yields and income per unit of water used. It works with agronomic, economic, and technical management decisions for decreasing irrigation water demand and recovering the costs of irrigation water supply. An agricultural water demand management program should consider the interaction between the quantity, quality, and biological aspects of both groundwater and surface waters; the sustainability of irrigation and drainage schemes; and environmental sustainability. Water demand management necessitates establishing structural incentives, regulations, and restrictions that help to bring about efficient water use in irrigation by farmers while encouraging innovations and saving technology.

INDUSTRIAL SECTOR

Water is used in industries in power generation, temperature control, manufacturing processes, and washing. The principal water users in industry are thermal and atomic power generation, chemistry and petroleum chemistry, ferrous and nonferrous metallurgy, wood pulp and paper industry, and machine building.

The quantity and quality of water required by industry vary with the type of industry and the industrial processes. The effluents produced by the industry are also of different qualities, which may require different treatment processes. Water recycling and reuse are the main water saving strategies in industry. Industrial water recycling, which began on a large scale in the 1970s to help cope with antipollution regulations, has also become an effective conservation measure. American steelmakers, which once consumed 280 tons of water for every ton of steel made, later needed to use only 14 tons of new water; the rest was recycled. Similarly, in Japan and western Germany, for example, a large increase in the number of factories has been accommodated without a considerable increase in total industrial water use (2).

The most common conservation measures in industrial water use are water-use monitoring, recycling, reuse, cooling tower use, equipment modification, and employee education. Of all industrial water use, cooling water accounts for a very large share. Replacement of once-through cooling processes with recirculating systems can provide some of the largest water savings. Often, cooling water can be reused for landscape irrigation and cleanup. The physical and chemical properties of water may change after use, requiring some simple treatment before being used again. Treated water can be reused for heating processes; water from washing raw food material can be reused for cooling, heating, or indoor washing.

DOMESTIC SECTOR

In urban areas, water supplies are used in households and for services such as city washing, fire fighting, and maintenance of swimming pools or recreational parks and lakes. A considerable amount of water is lost by leakage in the water distribution network. The unaccounted for water (UFW) comprises system leakages, illegal tapping, theft, incorrect meter reading and billing, and overflowing reservoirs. System monitoring and metering should be done in appropriate intervals to obtain information on the system state variables to prepare plans for water conservation measures as well as for estimating water losses, their location, and causes so as to prepare for appropriate corrective measures. Typical water losses as a percent of water supplied are 8–24% for developed countries, 15–24% for newly industrialized countries; for developing countries, it is 25–45%, as surveyed in 1991 by the International Water Services Association (IWSA). Examples of typical water losses for developing countries are well portrayed by Arequipa, Peru, with a 45% loss, and Hanoi, Vietnam, with a 68% loss (43% nonbilled, 20% leakage, 5% company use) (3). Between 20 to 40% of the supply was UFW in Indian cities and towns (4). On the municipal level, cities have saved from 10 to 25% of their water by repairing leaky pipes (2). Preventive as well as corrective maintenance can play a major role in water conservation.

Water pricing and billing can induce water saving habits in users if price increases with the volume used, particularly when the differential increases in price are large enough to encourage water savings. The reclamation of wastewater for nonpotable purposes has become a proven measure to help meet increasing demand. Where extreme water scarcity exists, a feasible but expensive solution is to have a dual distribution network for high quality water which is needed for uses such as drinking, food preparation, and bathing, and for treated reusable water which can be used for landscape and recreational ground irrigation, toilet flushing, and floor washing. Such reusable water can be available from, for example, treated urban drainage stormwater. Such examples of dual systems can be found in Australia, Hong Kong (China), Japan and the United States, to name a few. In Japanese cities, for example, such reclaimed water finds greatest use in toilet flushing (40%), compared with only 15% for irrigation in urban landscaping; in Hong Kong, seawater has been used extensively for decades for toilet flushing in commercial and multifamily residential buildings. However, a dual distribution system is difficult to implement because it requires very large investments from the public sector as well on the part of users in their homes. Financial incentives such as subsidies may be required for such investments in additional infrastructures and to implement other high-cost water conservation technologies.

In the domestic sector, most water is used for gardens and for bathrooms and toilets. In a typical household that has outdoor facilities, nearly 50% of the water is used for gardening; of the total indoor water use, 33% is spent on showers and bath, 24% for toilet flushing, and 22% in laundry (5). Water saving plumbing fixtures and appliances are cost-effective and provide permanent long-term economic advantages. Replacing traditional toilets which use 16 to 20 liters per flush with low-flow toilets (such as 6 liters per flush) or placing containers or bags filled with water in the tanks of large volume toilets can save a lot of water by reducing their capacity. A minor repair to control toilet leaks can be of immense help in conserving water. These leaks can be detected by adding a few drops of food coloring to water in the toilet tank and observing if the colored water appears in the bowl. Showers take second place in domestic water demand. Reducing the time for showering and preferring showers to immersion bathing can be very effective in conserving water. Low-flow showerheads and faucet aerators save water as well as energy used to heat water (where warm water is needed).

When washing clothes and dishes by machine, operating clothes washers and dishwashers with full loads can economize on water used. Outside the home, water can be saved by avoiding unnecessary spillage of water. Simple practices like using a broom, not a hose, to clean driveways, steps, and sidewalks; washing the car with water from a bucket; and irrigating plants either in the early morning or the late evening can save a lot of water. In landscaping, use of native plants that require less care and water than ornamental varieties can save a large quantity of water. While mowing the lawn, the lawn mower should be adjusted to a higher setting to provide natural ground shade and to promote water retention by the soil. In swimming pools, water can be continuously recycled, clarified, and purified by means of portable equipment and appropriate chemicals. Evaporation losses can be minimized by covering the pool when not in use.

INSTITUTIONAL MEASURES

Educational programs are very important in instilling water conservation awareness among users. Conservation literature can be distributed along with regular bills to water users. Local service organizations can also play a big role in disseminating water conservation information. An excellent time to provide this information to customers is just prior to the summer season when demand normally peaks. Local newspaper articles, radio and television public service announcements, information centers at local fairs and shopping centers, etc. can also be used to promote conservation awareness and education. Local schools can be a very good place to offer conservation knowledge to young people by means of water education activities such as water conservation posters, slogans, essays, or exhibit contests for children. High-use facilities such as schools and colleges, hospitals and institutions, country clubs, and health clubs are appropriate places to initiate water conservation programs.

Water laws provide a general framework for water use and conservation and are complemented by regulations. They can help water conservation by restricting excessive water use, for example, establishing standards for indoor plumbing fixtures, maximum volumes per flush in toilets, or making it compulsory for filtering, treating, and recycling the water in swimming pools, etc. The provision of incentive and penalties helps in enforcement of regulations as well as in the adoption of water saving technology by customers.

BIBLIOGRAPHY

1. World Bank. (2001). World Development Indicators (WDI). Washington, DC, Available in CD-Rom.
2. Beekman, G.B. (1998). Water conservation, recycling and reuse. *Water Resour. Dev.* **14**(3): 353–364.
3. WHO. (2001). Leakage Management and Control—A Best Practice Training Manual. http://www.who.int/docstore/water_sanitation_health/leakage/begin.html.

4. Venugopalan, V. (1997). Water conservation methods. *Water Supply* **15**(1): 51–54.
5. Stanger, G. (2000). Water conservation: what can we do? In: *Water Management, Purification & Conservation in Arid Climates*. M.F.A. Goosen and W.H. Shayya (Eds.). Technomics, Lancaster, PA.

THE DEVELOPMENT OF AMERICAN WATER RESOURCES: PLANNERS, POLITICIANS, AND CONSTITUTIONAL INTERPRETATION

Martin Reuss
Office of History Headquarters
U.S. Army Corps of Engineers

To understand the development of American water resources, one must first look at American political and social values and at American governmental institutions. Even a cursory examination shows the lasting influence of decisions and attitudes molded as the country took its first hesitant steps as a republic. Historian Joyce Appleby has argued that the first generation of Americans bequeathed "open opportunity, an unfettered spirit of inquiry, [and] personal liberty" to future generations—qualities, we might note, that often introduce an element of uncertainty into public administration. But if we extend the analysis a bit, we might not only gain an appreciation of the many challenges facing water resource developers, but also illuminate a fundamental question facing democratic nations: To what extent should human liberty be constrained in order to provide and manage a human necessity—water?

Beyond Appleby's observations, one notes at least pervasive elements woven into American political behavior. The first, the inescapable, element is distrust of powerful governments. Power corrupts, the first Americans agreed without much hesitation, and the challenge was how to minimize that corruption, how to ensure that good men will not be enticed to do evil, and how to disperse power to minimize oppression. Loudly over the years, Americans continue to proclaim their distrust of big government; even popular presidents generate skepticism when they appear to reach for increased power and authority.

Only as a last resort, and then with resignation, not enthusiasm, as during the Great Depression, do Americans turn to the national government to solve their problems. The result can be truly impressive: Grand Coulee and Bonneville dams, locks and dams on the Upper Mississippi, the California Central Valley Project, and the Los Angeles flood control system all came out of depression-era politics, but arguably all are aberrations in the story of American water resources.

The second element, almost as pervasive as the first, is that power and liberty are fundamental antagonists. The dispersion of power among the three branches of government purposely sets power at war with itself rather than with "life, liberty, and the pursuit of happiness." Each branch would be allowed only sufficient power to discharge official duties, and a system of checks and balances would guard against abuse. Recoiling from British monarchism, the constitutional drafters took special care to try to prevent executive branch intrusions into the duties of the other two branches. This was a system that, regardless of its merits, made implementation of rational planning enormously difficult, as water developers soon appreciated.

Political attitudes one thing; government structure was another. And here the Founding Fathers developed a system that guaranteed further complications. They fashioned a republican form of governments within the government. A century later, young political scientist Woodrow Wilson thought that this structure posed the principal challenge to American administration. Few water resource planners would disagree. Republican government, it must be remembered, began in the states, not in the new national capital; delegates to the Continental Congress delayed business so they could go home and participate in state constitutional conventions. The formation of these state governments may have excited Americans more than the latter formation of the union itself, and the American Constitution explicitly guaranteed to each state a republican form of government (Article IV, Section 4). Once the United States achieved its independence, many Americans pondered how citizens could owe allegiance to two governments, two legislatures, simultaneously. Were the states and national government partners or were the states meekly to accept national supremacy? No one at the Constitutional Convention quite knew what to expect from this layer-of-powers (or was it a marble-cake, twentieth century political scientists later debated), and numerous, contrary explanations emerged of what the delegates had actually achieved. In no area did the confusion become more manifest or disruptive than in internal improvements, especially in water projects that crossed state lines.

The term "internal improvements" came to mean many things to the citizens of the young republic. It included roads, canals, schools, lighthouses, fortifications, and even technological innovations-most anything that seemed to provide security and promote the economy. Gradually, it came to mean something a bit more specific, though still covering (pardon the pun) a large amount of ground: it applied to what we now call "infrastructure," and water transportation was a central concern. Benjamin Franklin had proposed at the Constitutional Convention that Congress have the power to construct canals, but opponents won the day, fearing that Congress would become too powerful. In fact, the term "internal improvements" cannot be found in the American Constitution, an obstacle for those seeking affirmative authority for federal involvement in public works. But neither did the Constitution proscribe the activity, which meant to internal improvement advocates that the function lay legitimately within federal authority. This ambiguity not only produced a constitutional quagmire for internal improvements, but it provided a platform upon which larger issues of the role of government and the nature of liberty could be debated. In short, the internal

---

improvements issue amplified and sharpened the debates about the very nature of American republicanism. By any other name, it continues to serve that function to the present day.

Given Americans' distrust of government and emphasis on personal liberty, America's first politicians, and all the generations following, confronted the difficulty of promoting economic growth without expanding governmental authority. One answer was the corporation, a device that actually predated the Constitution but in the age of internal improvements became much favored. As presumed promoters of the public good, they effectively became agencies of government. In this way, legislatures could support economic and political development without necessarily involving tax monies. The fact that individual incorporators might thereby profit aroused little concern. The more important point was that corporations brought together sufficient capital to launch an enterprise, whether a canal or a municipal water system. Even if a number of these ventures brought forth charges of corruption, internal improvement advocates ceaselessly trumpeted the moral and intellectual gifts stemming from public works, as though canals were spiritual as well as economic enterprises. To complaints that corporations disenfranchised people and led to the inequitable distribution of wealth, champions argued-somewhat quaintly in light of what subsequently emerged corporations were nothing more than little republics eminently suited for the United States. For better or worse, the victory of the corporation in American life was almost as revolutionary as the victory of republicanism itself, and the alliance between government and corporations became a hallmark of American economic development. Government was not to replace business, but was to support and, within certain limits, protect it.

George Washington and other Federalists had ardently hoped that corporations might provide the capital and means to build internal improvements to bind the nation together and transcend local interests, perhaps leaving overall planning to the national government. But the chance slipped through their hands. The structure of Congress assured that state interests in internal improvements would prevail over national interest. There would be no national board, no national planning. Rather, Congress would periodically pass rivers and harbors acts that generally reflected parochial politics. To stimulate states and the private sector, Congress also provided a percentage of funds obtained from the sale of public lands in new states to finance roads and canals (the three and five percent funds dating back to 1802) and voted to turn over certain lands to states for reclamation (Swampland acts of 1849 and 1850). In a few cases, too, Congress might vote to subscribe to canal stock or even grant land to a company-a practice that presaged the enormous land grants given to railroad companies as they extended their presaged the enormous land grants given to railroad companies as they extended their lines across the continent later in the century.

Caught in a congressional quagmire that appeared to offer no rational plan for the development of the country's infrastructure, succeeding presidents attempted to develop some orderly process, but at the same time they worried over possibly unconstitutional intrusions into areas beyond federal authority. We turn our attention to the Executive Branch in part two of this series.

This article is based on Martin Reuss, "The Development of American Water Resources, Planners, Politicians, and Constitutional Interpretation," in Paul Slack and Julie Trottier (Eds.), *Managing Water Resources, Past and Present: The Twelfth Annual Linacre Lectures*. Oxford University Press, 2004. Used by permission of Oxford University Press.

The Development of American Water Resources: Planners, Politicians, and Constitutional Interpretation.

The Constitution and Early Attempts at Rational Water Planning.

The Expansion of Federal Water Projects.

WATER MARKETS: TRANSACTION COSTS AND INSTITUTIONAL OPTIONS

Terence R. Lee
Santiago, Chile

The costs involved in the transfer of property, or transaction costs, can significantly affect the capacity of any market to operate efficiently. If water marketing is to achieve its full potential, markets must be designed to minimize transaction costs. Water marketing may lead, however, to efficiency gains, even if transaction costs are high. Moreover, increasing water scarcity will raise the welfare gains from trade relative to these costs, as the value of water rights increases. Transaction costs can be lowered by technological advance and by institutional design. It can also be expected that transactions costs will fall as the market matures, as learning by doing has its effect.

TRANSACTION COSTS

Transaction costs can be defined as those resources dedicated to establish, operate, and enforce a market system. They may take one of two forms, the services which buyers or sellers must provide from their own resources; the differences (margins) between the buying price and the selling price of a water right, for example, due to the direct financial cost of any brokerage services.

In a water market, there are three potential sources of transaction costs:

the costs of finding trading partners, of verifying ownership of water rights, and of describing the right for purposes of the proposed transfer;

the costs involved in reaching an agreement to trade, such as negotiating the price; arranging financing and other terms of transfer; drawing up contracts; consulting with lawyers and other experts; paying fees for brokerage, legal and insurance services; and transferring legal titles;

the costs of setting up a legal, regulatory and institutional framework, mitigating possible third-party effects, and ensuring compliance with applicable laws.

Water markets tend to be highly regulated, so many transaction costs arise from public policies governing water transfers. The burden imposed by policy-induced transaction costs comes from the direct costs of a transfer and also from the delays due to waiting for regulatory approval. These costs can be high. For example, in the western United States, waiting periods for state agency decisions, vary from 4–6 months in New Mexico, to 5–9 months in Utah, and to an astonishing 20–29 months in Colorado (1,2). It is possible to place institutional restraints on such regulatory induced delays. In Chile, the Water Code regulates the maximum time periods permitted for each stage of any recourse to regulatory intervention (3). This type of institutional requirement, at least, reduces the arbitrariness inherent in regulatory delays.

Transaction costs prevent markets from operating efficiently and reduce the overall economic benefits of water marketing. They increase the cost of water rights and decrease the incentive to trade. Transaction costs introduce inefficiencies, which block equalization of marginal values among different uses, users, and locations. In extreme cases, transaction costs can prevent markets from forming altogether.

These considerations imply that transaction costs usually reduce welfare and that public policies to reduce costs will be welfare improving. This does not mean, however, that all regulation of water markets is bad. Some costs arise from justifiable efforts to protect third-party interests. Such policy-induced transaction costs are not necessarily wasteful or inefficient if they provide real protection for third parties that may be affected by water transfers.

Transaction costs can be independent of the quantity of water rights transferred. The title search, filing fees, and other similar costs are often fixed. This penalizes smaller trades and favors larger trades. Other transaction costs, such as brokers' commissions, may be proportional to the amounts or values traded. More extensive water infrastructure and highly developed user institutions can lower costs. A recent empirical study indicates that, in Chile, transaction costs are particularly low in the areas that have more modern infrastructure and well-developed water users associations (4).

Variations in transaction costs can affect the spatial concentration of market activity. In Texas, for example, the major market activity is concentrated in the Lower Rio Grande Valley where transaction costs are low (5). Transaction costs can also affect the choice of the type of transaction. For example, usually the transaction costs for a permanent transfer exceed those for a lease.

MARKET-RELATED TRANSACTION COSTS

The efficiency of any competitive markets rests on the assumption that good, reliable, and easily accessible information is available. In practice, such information is not always available for water markets, partly because producing and disseminating information is often costly and difficult and, partly, because the means to redistribute risk are incompletely developed due to uncertainty about the nature of the risk and the asymmetry of information between market participants. Altogether, however, the information requirements for efficient water marketing are no greater than those needed for effective administrative allocation of water.

A study of water markets in the western United States concluded that all markets studied were characterized by varying degrees of uncertainty and incomplete access to market information on water commodities, prices, and market opportunities (6). Where there is no ready means for buyers and sellers to obtain information, they face legal, hydrologic, and economic uncertainties. These constitute a substantial cost disincentive to engage in water marketing.

INSTITUTIONAL OPTIONS

Efficient construction of any market requires the existence of the necessary conditions for trading to occur: (1) well-defined property rights, (2) public information on the supply of and the demand for water rights, and (3) the physical and legal possibility for trading to take place. Of these three necessary conditions, by far the most important is the existence of well-defined property rights. In the case of water, property rights define and limit the rights and duties of their holders relative to one another and to the rest of society to the use of a certain amount of water, which may be defined either volumetrically or in terms of shares of a stream or canal flow. If rights are poorly defined, market processes cannot be relied upon to allocate water resources efficiently. It is a basic responsibility of governments, as far as markets are concerned, to define, allocate, and enforce property rights in water. Government policies play a critical role in defining the institutional setting for market operation and provide the basis for market activity by defining, allocating, and enforcing water rights.

The way property rights are defined will structure the incentives and disincentives that members of society face in their decisions regarding water ownership, use, and transfer. For market participants to estimate the value of a water right, they must be able to form secure expectations about the benefits and costs of owning and transferring it and the degree to which it is protected from impairment by others (7).

So, there is a need for a reliable and trusted means of registering rights. Hydrologic information is also required to permit the right to be defined. Various types of information are essential for rational decision-making by water rights holders: the legal and hydrologic characteristics of water rights and the costs of alternative means of obtaining water. This implies the existence of good data and monitoring systems, which can be provided only by public agencies.

The public availability of information on the supply of and the demand for water rights must also include the means to identify willing buyers, sellers and intermediaries or brokers and the means for entering into enforceable contracts. Such information is better provided by the market rather than by public agencies.

In addition, a clearly defined set of transfer rules is necessary to permit market transactions to take place when buyers and sellers determine. Transactions should be contingent only upon compliance with a known set of trading rules or transfer criteria. This is an essential prerequisite for a continuous water market. There must also be information ensuring that the physical possibilities for the transfers produced through trade can actually happen. This may require the legal possibility of easements or the purchase of rights of way across the property of third parties.

BIBLIOGRAPHY

1. Colby, B.G., McGinnis, M.A., and Rait, K.A. (1989). Procedural aspects of state water law: transferring water rights in the western states. *Ariz. Law Rev.* **31**.
2. MacDonnell, L. (1990). *The Water Transfer Process as a Management Option for Meeting Changing Water Demands*. Natural Resources Law Center, University of Colorado, Boulder.
3. Lee, T.R. (2000). *Water Management in the 21st Century: the Allocation Imperative*. Elgar, Cheltenham.
4. Hearne, R.R. and William, E.K. (1995). *Water allocation and water markets: an analysis of gains-from-trade in Chile*, Technical Paper Number 315, The World Bank, Washington, DC.
5. Chang, C. and Griffin, R.C. (1992). Water marketing as a reallocative institution in Texas. *Water Resour. Res.* March, No. 3.
6. Saliba and Colby B. (1987). Do water markets 'work'? Market transfers and trade-offs in the Southwestern states. *Water Resour. Res.* **7**.
7. Colby, B.G. (1988). Economic impacts of water law—state law and water market development in the Southwest. *Nat. Resour. J.* October, No. 4.

AVERTING WATER DISPUTES

Jody W. Lipford
PERC
Bozeman, Montana, and
Presbyterian College
Clinton, South Carolina

INTRODUCTION

At midnight on August 31, 2003, time ran out on a proposed agreement among the states of Alabama, Florida, and Georgia to allocate water in the Apalachicola-Chattahoochee-Flint (ACF) River Basin. The deal had been 13 years in the making, but it ended in failure. "It's a true shame that we were as close as we were and couldn't get an agreement," said Alabama's chief negotiator (1).

It was a shame. The collapse of these lengthy negotiations sends the matter to the courts, and the Supreme Court may ultimately decide how the disputed water will be divided. More broadly, the failure of the state governments to reach agreement reveals that water, long considered plentiful in the southeastern United States, is in danger of becoming a subject of intractable conflict. The failure signals that a water crisis may well emerge in the region unless new approaches to allocating water are adopted.

As the population of the Southeast increases, competing demands for water—for municipal use, for recreation, and for hydropower, to name just a few—are growing. Today the problem surfaces in the form of occasional interstate disputes such as this one, but the failure to resolve them casts an ever-longer shadow over the future of water resources in the region. When demands of competing users outstrip supply, there must be ways to ensure that water goes to the users who value it most and that the waterways of the Southeast are not roiled by unending conflict.

This article explains the reasons behind the conflict in the Apalachicola-Chattahoochee-Flint River Basin, why attempts at resolving it failed, and what alternatives should be considered. It explains how to allocate water to its most productive uses, restore peace to the areas around these waterways, and avert other conflicts that are emerging, not only in these states but elsewhere in the South.

THE BACKGROUND

As shown on the map in Fig. 1, the ACF Basin drains an area of 19,800 square miles in the states of Georgia, Alabama, and Florida. The Basin starts in the headwaters of the Chattahoochee River in northern Georgia, above Atlanta. The Chattahoochee flows through Georgia's Piedmont before turning sharply south, forming the southern half of Georgia's border with Alabama and a notch in Georgia's border with Florida. At the border, it meets the Flint River to form the Apalachicola River, which flows through the Florida panhandle into the Gulf of Mexico (2).

Historically, in the ACF basin, as in most of the southeastern United States, water has been abundant and has met the many demands for it. The demands include water for domestic, commercial, industrial, hydroelectric, navigational, and recreational uses.

Under riparian water rights—the system of water rights in the eastern part of the United States—landowners can use water that flows adjacent to their property as long as they do not appreciably diminish the quantity or quality of water available to downstream users. But riparian rights were effectively overridden in 1946, when Congress authorized the U.S. Army Corps of Engineers to construct dams for flood control, navigation, and hydroelectricity along the Chattahoochee River.[1] Later,

---

[1] The Army Corps of Engineers has long played a role in the ACF River Basin. To facilitate commercial traffic, the Corps began dredging the Chattahoochee River in the 1880s. At present, five Army Corps dams dot the ACF River Basin: the Buford Dam, which forms Lake Sidney Lanier; the West Point Dam, which forms Lake West Point; the Walter F. George Dam, which forms Lake Walter F. George; the George W. Andrews Dam; and the Jim Woodruff Dam, which forms Lake Seminole. Authorization and construction of these dams began in the 1940s, 1950s, and 1960s, and by the 1970s they were all operational (3, pp. 151, 155).

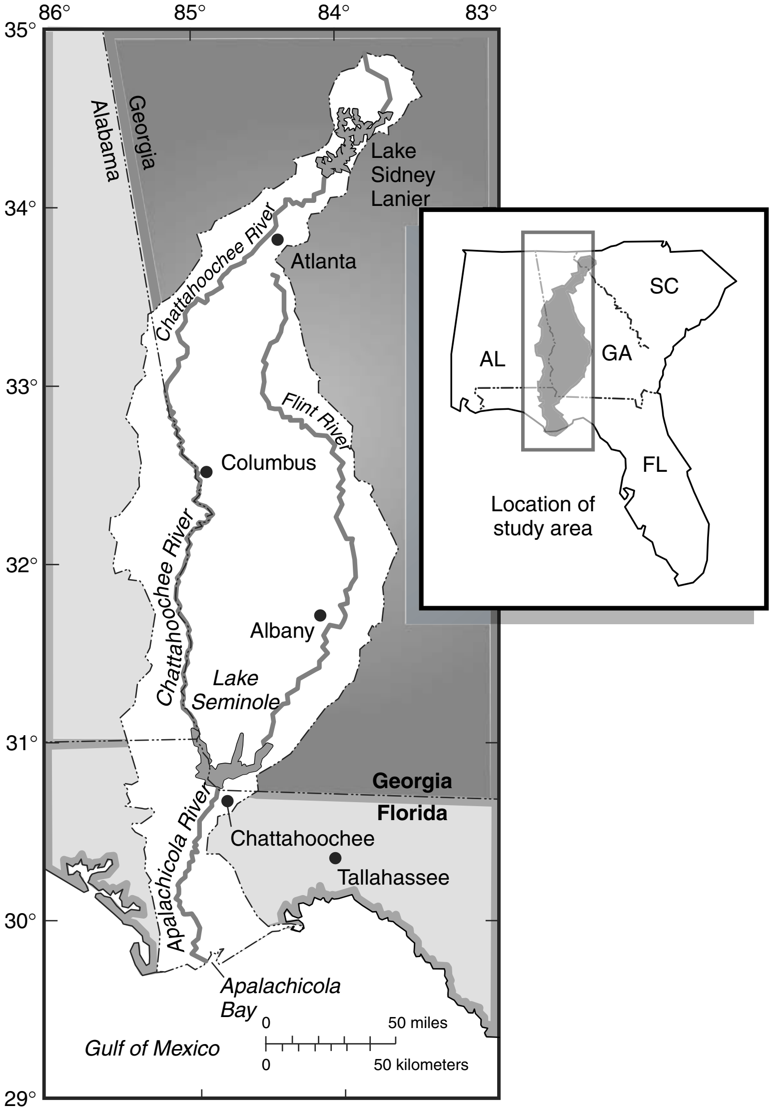

Figure 1. Apalachicola-Chattahoochee-Flint (ACF) River Basin. Courtesy of Mandy-Scott Bachelier.

the Corps added water supply and recreation as purposes of the dams and the reservoirs they created (4,p.2; 5). These projects transformed the waters of the basin from private property governed by the riparian doctrine to public property.

Today, the waters of the ACF River Basin continue to be owned and managed by the federal government through the Army Corps of Engineers. The Corps' managers meet weekly to consider various water needs, such as hydroelectric production, recreation, navigation, and environmental quality (6). In addition, all requests to increase water withdrawals must be approved by the Corps (7, p. 993).

Except for flood control, each of these purposes requires a minimal lake level or river flow rate. Electricity cannot be produced, nor can barges navigate, without sufficient water. Similarly, fish populations require stable lake levels during spawning season. Sufficient flow also dilutes pollution, helping to ensure water quality (6). The Corps also commonly provides recreational facilities, such as parking areas, boat ramps, and public restrooms. Response to the recreational amenities of the lakes has been heavy; millions of user-days are tallied each year (3, p. 158). The Basin also supplies water for public use. All these competing demands limit the amount of withdrawals that can be made.

Conflicts Over Increasingly Scarce Water

For a long time, the system of riparian doctrine and public management through the Army Corps of Engineers worked well. In the 1980s and 1990s, however, rapid population growth, particularly in metropolitan Atlanta, combined with recurrent drought led to increased pressure on the ACF River Basin's resources.

Atlanta's population grew from 2.2 million in 1980 to 3.0 million in 1990, and then to 4.1 million in 2000. Of 126 metropolitan statistical areas listed by the Census Bureau, only 17 had higher population growth rates from 1980 to 1990, and only eight had higher population growth rates from 1990 to 2000.[2]

---

[2] See Table 30 of U.S. Census Bureau (8, pp. 32–34).

Demand for water to satisfy this growing population increased dramatically. Metropolitan Atlanta's water use increased from 289 million gallons per day in 1980 to 459 million gallons per day in 1990, and then to 606 million gallons per day in 2000.[3] Metro Atlanta relies almost exclusively on surface water; over 70% comes from the Chattahoochee River and Lake Sidney Lanier, the lake formed north of Atlanta by the dam at the headwaters of the Chattahoochee.

Yet, the Chattahoochee River and Lake Lanier are ill suited to supplying Atlanta's water needs. The Chattahoochee is the smallest watershed in the country to supply a metropolitan area with the majority of its water (9). The largest share of Atlanta's water use—53.8%—is for residential use; commercial, government, and industrial users take 22.8%, 5.6%, and 4.2%, respectively (10).

A preliminary study by the Army Corps of Engineers indicates that Atlanta is already approaching, and at times exceeding, water use levels that were not expected until 2030. Whether these findings are accurate or not—the Georgia Environmental Protection Division says that the Chattahoochee River and Lake Lanier are sufficient to supply Atlanta through 2030—there is no doubt that future demands on the Chattahoochee River and Lake Lanier will be heavy and growing (11).

Drought has worsened this pressure on the River Basin's waters. During the 1980s and 1990s, the Southeast experienced recurrent and severe droughts. In the most severe drought, rainfall in Atlanta fell by as much as 25%, and annual average streamflows along the Apalachicola River fell to less than half their historical norms.[4]

The Crisis Begins

In 1989, recognizing that Atlanta's "finite supply of clean water is looming as a barrier to growth" (12, p. 68), Atlanta and the Army Corps of Engineers proposed to approximately double the water Atlanta drew from Lake Lanier, to 529 million gallons per day. Without sufficient water, Georgia officials feared the loss of 680,000 jobs and $127 billion in wages through 2010 (13, p. 26).

The proposal by Georgia and the Army Corps alarmed the citizens of Alabama. Increased withdrawals from Lake Lanier would reduce flows along the segment of the Chattahoochee River that forms the southern half of Alabama's border with Georgia and would stunt economic development there. So, in 1990, Alabama sued the Army Corps to keep it from allocating more of the ACF River Basin's waters to Atlanta.

The state of Florida quickly joined the lawsuit on the side of Alabama, fearing that reduced water flows would harm the oyster-rich Apalachicola Bay farther downstream and lead to deterioration of the Apalachicola's water quality. The state of Georgia then joined the lawsuit on the side of the Army Corps of Engineers to defend its withdrawal. The stage was set for 13 years of studies, proposals, counterproposals, and extended deadlines.

These actions reflected unique concerns in each state. Alabama officials worried that Atlanta's withdrawals would stifle Alabama's economic development by limiting water needed for domestic, industrial, and commercial use. Water quality would also suffer because reduced downstream flows would mean less dilution of polluted upstream water. Atlanta not only uses a large amount of water, but also discharges heavily polluted water back into the Chattahoochee (7, p. 996; 12, p. 68–69).[5] If dirtier water came from Atlanta, Alabama might have to raise water cleanup standards for industrial and municipal users, which would be costly and would put Alabama at a competitive disadvantage in attracting economic development (14, p. 3).

From Florida's perspective, the chief problem was oyster beds. Reduced flows, especially at critical times, and heavier pollution could threaten the Apalachicola Bay's oyster industry, which supplies approximately 10% of the country's oysters and employs over 1000 people. The river is also a commercial source of shrimp, blue crab, and finfish, as well as the home of an important sport fishery and the only commercial source of Tupelo honey. It has been recognized as an Outstanding Florida Water (15, p. 34; 16).

Other factors exacerbated the conflict. For example, recreational users want lakes kept full or nearly full; however, the competing objectives of hydropower and flood control require that lakes be drawn down, typically during summer and winter. Similarly, navigation requires minimal flows that reduce the water in lakes when river levels become too low for barge traffic. Finally, ecosystem preservation requires a pattern of flows that mimics nature's seasonal cycle and may conflict with other demands (4, p. 2; 14, p. 3).

Despite years of negotiation, the states never agreed how to allocate the basin's waters. Because these negotiations failed, the matter is now likely to be settled by the Supreme Court.

Constitutional Paths to Solution

Once Alabama took the Corps to court in 1989, the states had several constitutional options to choose from to settle the conflict over water allocation. They could go to Congress for a decision, their suits could reach the Supreme Court, or they could negotiate their own agreement or compact. Neither Congress nor the Supreme Court likes to get involved in interstate water disputes. Moreover, suits reaching the Supreme Court are costly, usually require lengthy negotiations, and yield uncertain outcomes.

For these reasons, the states opted for an interstate water compact. Alabama agreed to deactivate its lawsuit during the negotiation period, the Army Corps of Engineers agreed not to allocate additional water from Lake Lanier to Atlanta, and all parties agreed to a comprehensive study of the water resources in the basin (4,

---

[3] Data supplied by Julia Fanning, U.S. Geological Survey, Atlanta, e-mail correspondence, November 7, 2003.

[4] For data on streamflow of the Apalachicola River at Chattahoochee, FL, see waterdata.usgs.gov/fl/nwis/annual/calendar_year/?site_no=02358000.

[5] In response to federal consent decrees to stop spills of untreated wastewater into the Chattahoochee and to comply with the Clean Water Act, Atlanta is currently trying to raise over $3 billion to renovate its antiquated sewage treatment system.

pp. 3–4).[6] In early 1997, all three state legislatures ratified legislation authorizing the negotiation of an ACF River Basin Compact to allocate the basin's waters. These acts were subsequently signed by the three state governors. The U.S. Constitution requires congressional approval of interstate compacts, so in November 1997, Congress approved and President Clinton signed federal authorizing legislation. The goal of the compact was to assign property rights to water fairly and avoid future conflicts. This goal proved elusive, however.

The Failure of the Compact

The federal legislation set an initial deadline of December 31, 1998 for compact negotiations, unless the states agreed unanimously to extend that deadline.[7] This deadline proved much too optimistic, as each state presented proposals reflecting its parochial interests. To begin with, Alabama and Florida wanted consumptive uses of water defined and limited.[8] Georgia resisted this approach in favor of one that focused on reservoir levels. Specifically, Georgia wanted the ACF River Basin's reservoirs to be kept full or nearly full (19, p. 8), so that Georgia would have enough water to supply Atlanta (4, p. 14). In return, Georgia offered Alabama and Florida minimum flow guarantees. But Alabama and Florida rejected Georgia's proposal, fearing that the minimum flows might become the norm, in essence reducing the water flowing downstream.

To ensure adequate flows along its border, Alabama argued that the ACF River Basin's water should be allocated to meet the original objectives of dam construction. These included navigation (i.e., barge traffic), flood prevention, and hydroelectric production—but not water supply to municipalities or recreation (4, p. 2; 19, p. 9). Florida agreed with Alabama in opposing minimum flows, but also wanted downstream flows to be adjusted to mimic natural flow cycles. Additional problems plagued the negotiations, from definitional questions (e.g., how to define "severe drought") to the choice of the computer model for forecasting river flows and lake levels (19, pp. 9–10).

No agreement was forthcoming, so the states agreed to extend the deadline till January 1, 2000. Yet, the passing of another year did not appreciably advance the negotiations. Once again, the states set a 1-year deadline, establishing a pattern of deadline extensions that continued until July 22, 2003. At that point, progress seemed to have been made. The three governors signed a Memorandum of Understanding that set a blueprint for water allocation in the ACF River Basin. The memorandum authorized water supply for Atlanta from Lake Lanier at 705 million gallons per day and left open the possibility of greater future withdrawals. The memo also established minimum flow requirements downstream from Atlanta; the most important was a flow of 5,000 cubic feet per second on the Apalachicola River at Chattahoochee, Florida. The deadline for final agreement on the memorandum was August 31, 2003 (20).

Although Florida Governor Jeb Bush signed the memorandum, he and other Florida officials had reservations, which they expressed in an accompanying statement (21). They insisted that minimum flows must not become targets, that Lake Lanier must be managed to deal effectively with drought, and that the governing ACF Commission must have authority to approve any withdrawals from Lake Lanier that exceeded the amount specified in the memorandum.[9]

Ultimately, the states could not agree. Florida feared that minimum flows, which had been less than 5000 cubic feet per second only twice during the recent droughts (on a mean monthly basis), might become the norm. Thus, Florida officials again raised the issue of Georgia's withdrawals from Lake Lanier. Georgia officials responded by agreeing either to limit Atlanta's withdrawals or to promise minimum flows through the basin, but not both, and accused Florida of trying to micromanage its waters (23,24). As a result, Florida officials refused to accept the Memorandum of Understanding, and the deal collapsed. The states have now reactivated their original lawsuits (1,5,24).[10]

THE ALLOCATION OF WATER

Economists recognize the scarcity of valuable resources. Without prices on these resources, there is not enough to satisfy all who want to use them. In most settings, however, market prices allocate resources and allow them to move to those users who value them the most. This market process allocates resources to their most productive uses and maximizes society's wealth. In the case of the ACF River Basin water, however, market prices do not currently allocate water; it is allocated politically.

Economists also recognize that resources have multiple uses. Water, for example, may be used to generate

---

[6] Until the compact was completed, the states agreed to "freeze" water at current use levels. Should increased withdrawals be needed, the states agreed to notify other states in advance (17, p. 202). For further details, see Public Law 105-104, Article VII (c).

[7] [Public Law 105-104, Article VIII (3)].

[8] Consumptive use, also known as water consumed or water depleted, may be defined as the "part of water withdrawn that is evaporated, transpired, incorporated into products or crops, consumed by humans or livestock, or otherwise removed from the immediate water environment." The definition also includes "any water withdrawn in the basin and transferred out of the basin for use" (18).

[9] Sixty-four percent of the ACF River Basin's reservoir capacity is held in Lake Lanier (22).

[10] Complicating the legal proceedings is a deal struck by Georgia and the Army Corps of Engineers in January 2003 in which metropolitan Atlanta governments agreed to pay the Corps $2.5 million per year toward the operation of Lake Lanier's Buford Dam in exchange for greater withdrawals from the lake. With this deal, metropolitan Atlanta sought not only to obtain additional water, but also to mollify hydropower customers who pay for the dam and who had filed suit against the Corps in 2000 because the Corps had already allocated water from hydropower to supply metropolitan Atlanta. Georgia and the Corps negotiated this deal without informing Alabama or Florida and set it aside only when Alabama and Florida found out about it and threatened to withdraw from the compact negotiations. This deal, like the initial lawsuits between the states and the Corps, has been reactivated and will have to be settled by the courts (5,25).

electricity; aid oyster production; provide wildlife habitat; produce industrial products; provide channel depth for barges; provide recreational opportunities for boaters, skiers, and fishers; or supply households with water for drinking, watering lawns, or filling swimming pools. In the ACF Basin, some water is used in each of these ways.

Whether allocation occurs through market prices or other methods, it is rarely all-or-nothing. In the Southeast (and the United States generally), there is ample water to supply basic human needs, such as drinking water. Where conflicts occur, as in the ACF Basin, it is over shifting some water, not all water, from one use to another. Simply put, the ACF Basin issue is whether more water should be allocated to Atlanta, presenting Alabama and Florida with the prospect of less water but not complete deprivation.

Although compacts have some advantages over congressional or judicial apportionment, they are poorly suited to allocate water in ways that maximize water's productive value to society. Compacts are highly political and confront intractable information problems, and such was the case with the negotiations over the ACF Basin's waters.

Interest Groups

Groups with a vested interest in the outcome of the compact influenced the ACF River Basin negotiations. Each tried to get more water allocated in its favor, irrespective of water's most productive uses. The influence of these groups introduced conflict, making a workable agreement difficult to achieve. Industrial, environmental, municipal, and political interests all made their voices heard (19, p. 8). The *Atlanta Journal-Constitution* identified political and business leaders of metropolitan Atlanta, environmentalists, and Florida's shellfish and fishing industries as uncompromising interest groups who refused to yield to the demands of other users (1).

The Apalachicola Bay's oyster industry serves as an example of a small, well-organized interest group with strong influence; its employment of approximately 1000 people is minimal in a state with total employment of approximately 7.2 million.[11] Note, however, that the industry had support throughout the state of Florida from citizens who wanted the environmental amenities of their state's river preserved.

After the agreement failed, some interests, such as the Atlanta Regional Commission and homeowners and businesses on Lakes Lanier and West Point, seemed just as happy. They feared that Georgia had compromised too much already.

Informational Problems

Even if negotiators could be insulated from interest group influence, they would still face important informational questions. If their goal is to allocate the water to its most productive uses, negotiators must first know how much total water can be allocated and how that will vary over years of normal rainfall and drought. Perhaps most important, they need to determine whether society will benefit more from allocating water to Atlanta's developers, say, or to Florida's oyster producers. And, if they can decide objectively to allocate more water to Florida's oyster producers, they would still have to decide whether the extra water should come from Lake Lanier (thereby maintaining levels at downstream reservoirs) or from downstream reservoirs (thereby maintaining levels at Lake Lanier).

By making these decisions, policy makers are implicitly choosing who will benefit and who will be harmed. A decision to allocate more water to Atlanta lessens development in eastern Alabama and reduces Florida's seafood production. A decision to allocate more water to Alabama and Florida benefits the economies of these states, but curbs Atlanta's economic development. Similarly, a decision to supply downstream users from Lake Lanier diminishes recreational opportunities for users of that lake and maintains those opportunities for users of downstream reservoirs; the opposite decision would benefit Lake Lanier's recreational users but harm those who use downstream reservoirs. And even if policy makers could determine water's most productive uses, their decisions would soon be rendered obsolete by changes in the total supply of water, changes in the total demand for water, and marginal changes in allocation necessary to maximize the total productive value of the basin's water resources.

Negotiators did try to obtain answers to some of the technical questions through the use of computer software that forecast future river flows and reservoir levels based on consideration of "historic rainfall patterns over the last fifty-five years" and "anticipated water uses within the basin in a future year, typically 2030 or 2050" (19, p. 8). But repetition of historical rainfall patterns is not guaranteed. Nor are anticipated water uses easily forecast. Atlanta's rapid population growth and commensurate water use have been dramatically underpredicted by the experts.[12] To compound matters further, different software programs give different estimates, and, not surprisingly, the states have used different modeling programs (19, p. 10).

Practical Problems

Negotiators confronted two other factors that made agreement less likely: (1) The ACF River Basin's waters were already fully allocated, and (2) the drought was expected to end soon. That the water wealth of the ACF River Basin is already fully allocated made bargaining more contentious because changes will force redistribution of existing allocations. In contrast, for example, the country's first interstate water compact, the Colorado River Compact of 1922, was negotiated in the arid West, and more water was anticipated from the Boulder Canyon Project Act, which authorized the Hoover Dam and created Lake Mead. Moreover, negotiators knew that the 1998–2002 drought was unprecedented and would likely come to an end, reducing pressures on the ACF River Basin's waters. These expectations were borne out;

---

[11] The employment figure is taken from table 602 of U.S. Census Bureau (8, p. 393).

[12] The *State and Metropolitan Area Data Book, 1997–98* predicted that metropolitan Atlanta's 2000 population would be 3.682 million (26). The actual figure was 4.112 million, an error of 430,000 for a prediction published only 2 years in advance.

abundant rain fell during the latter half of 2002 and during 2003.[13]

PROPOSALS FOR MARKETING WATER

With compact negotiations now in disarray, policy makers must look to other alternatives. An obvious proposal is for Atlanta (and other municipalities in the basin) to charge a price for its water that at least approximates its market value. Ample evidence shows that higher water prices reduce consumption (27, p. 12–13). At present, water is underpriced in Atlanta, leading to overuse. Even in the most recent drought, the city of Atlanta raised its price to residential users by only 3%. For the average residential user, the monthly water bill rose from $16.55 to $17.05. The city plans additional rate hikes through 2004, but these will raise average residential bills by a mere $1.00 per month (28). During the drought, rather than raising the price of water further, officials imposed restrictions on outdoor water use that continue in effect (29,30).

Raising water prices to market levels is apparently not politically feasible. This means that the states of Alabama, Florida, and Georgia and the federal government should consider basinwide proposals to create water markets. Experience with markets in water has shown that they can overcome some of the most difficult challenges of water allocation. They can ensure that water is allocated to its most productive uses and can prevent conflicts among users.

To some, marketing water is still a strange idea. Long accustomed to the notion that water is a commonly owned resource, many readers may doubt that it is feasible to "trade" water and thereby satisfy various interests more readily than through political negotiations. Yet, there is a strong precedent for marketing of water. Much of the American West is arid; it receives less than 17 inches of precipitation a year. With water always precious, the West evolved a system of private property rights to water, and with it, water markets. This system, known as the prior appropriation doctrine, resulted from the need to divert water for mining and agriculture. In simplest terms, it allowed a person to divert water from a river or stream on the basis of seniority (or "first in time, first in right"); the right remained as long as the individual continued to use the water ("use it or lose it"). Water users could transfer their rights to others. The private provision of water flourished (31, p. 31–45), and continues to do so, although laws that guide the transfers of water are somewhat antiquated, and restrictions on transfers that made sense in the past do not necessarily encourage efficient use today.

In spite of these restrictions, water trades occur among agricultural users, between agricultural users and cities, and between agricultural users and environmentalists. Even interbasin and interstate trades are common (27, p. 14–21). In Texas, where both riparian and prior appropriation doctrines are recognized, a system of marketable permits similar to that described below allocates water along the Rio Grande River (32).

In a fully functioning water market, users pay a market price for water consumed, and that price serves as a rationing mechanism. Those who can put the water to the most productive use and demonstrate this by willingness to pay, will purchase the water, be they residential developers or oyster producers. Market prices motivate those with relatively less productive opportunities to sell the water to more productive users. Through markets, groups can work out slight or marginal changes that maximize the total value from all uses.

Markets also yield peace among transacting parties. In markets, only the parties considering buying or selling a resource take part in the negotiations. Outside influences from politicians, bureaucrats, or interest groups do not guide the negotiations, as they do in the political process. The terms of exchange, such as the price, must be voluntarily agreed upon for exchange to occur. Any would-be buyer or seller who does not like the price does not have to buy or sell.

In contrast, in the political sphere, resource users often do not pay a price for a resource they consume, or they pay less than the resource's market value. This encourages them always to want more and leads to conflicts among users and a state of perpetual unhappiness for all.

As economists often point out, the foundation for markets is private property rights that are defined—that is, rights with a clearly specified ownership claim; enforced—that is, rights with a claim that is secure; and transferable—that is, rights that may be sold to others. Clearly defined, enforced, and transferable property rights are necessary for exchange. Buyers will not purchase resources if the rights to those resources are uncertain or insecure, but when rights to property are certain, secure, and transferable, markets flourish. Market-based allocation of the ACF River Basin's waters would encourage allocation of the basin's waters to their most productive uses and foster peace among users.

The Army Corps' Role

To understand how markets might work, it is appropriate to begin with the Army Corps of Engineers, which is the effective owner of the water in the ACF Basin. At present, the Corps almost entirely depends on congressional appropriations. In the Mobile District in which the ACF River Basin is located, the Corps receives some fees for its services, but they represent a small part of the Corps' budget. The fees it receives are either insufficient to cover the costs of its services (as in recreation fees), the Corps does not retain the revenues (as in the case of revenues from hydropower), or it simply does not charge for the services it provides (as for navigational services, although commercial vessels do pay a fuel tax that is used to fund inland waterway projects).

The Corps is supported by taxpayer dollars and cannot receive financial benefit from the services it provides, so it has no incentive to determine which competing uses are most productive and thus to adopt market exchange as the

---

[13] From July 2002 to August 2003, rainfall exceeded normal levels in Atlanta, Columbus, and Albany, Georgia, by 9.64 inches, 12.21 inches, and 6.57 inches, respectively (data supplied by Pam Knox, assistant state climatologist, Georgia State Climatology Office, University of Georgia, Athens, e-mail correspondence, October 15, 2003).

way to allocate water in the Basin. In an ideal world, the Corps' financing and function would be changed to give it an incentive to allocate scarce water resources to their most productive uses, thereby raising the total wealth generated from the Basin's waters. This would happen if the Corps were to retain property rights and management authority over the Basin's waters, but taxpayer support of the Corps and its projects were reduced. In exchange, the Corps would be given the authority to charge whatever fees it deemed appropriate for the services it provided and to retain the revenues. For example, the Corps could implement or change fees it charges for hydropower, dredging, water supply, and recreational services. If drought or increased demand raised the relative scarcity of water, the Corps would have the authority to raise fees. Some taxpayer support is justified because the Corps also provides flood control that benefits all users.

Although the Corps is extremely powerful in Congress and therefore such a change in the financing of the Corps is unlikely, there is some precedent for this kind of institutional reform of a public agency. In 1996, the Fee Demonstration Program allowed the National Park Service, the Forest Service, the Bureau of Land Management, and the Fish and Wildlife Service each to choose 100 sites that could raise or implement new fees and retain 80% of the revenues. Although the Fee Demonstration Program does not intend self-sufficiency for the participating agencies or individual sites, the results from changed incentives are evident, as these agencies have improved services to visitors of public lands by allocating more funds to badly needed repair and maintenance of some of the country's most-valued natural and recreational resources (33). Like the Fee Demonstration program, reform of the Army Corps of Engineers could begin on a short-term, experimental basis.

With a mandate to balance its budget and the authority to set fees and retain revenues, the Corps would have an incentive to allocate water resources to their most productive uses. If Atlanta developers wanted more water, they would have to pay a price that reflected the value of the water to other users. If it did not, those other users would outbid it. The Corps would also have to take into account the costs of its services. If barge traffic was insufficient to generate revenues to cover the costs of dredging, the Corps would cease to dredge the basin. Through this system, those with the most productive opportunities for the water would be the ones to obtain it. Such allocation would also maximize the Corps' net revenues. Unfortunately, this outcome is not very likely to occur in a political setting.

A System of Marketable Permits

Under current political arrangements, marketable permits seem to be the most promising approach to creating a water market. Marketable permits depend on the assignment of property rights to water. To implement them in the ACF River Basin, the Army Corps of Engineers could first establish a daily "water budget," consisting of the total net withdrawals allowed from the basin, based on average daily withdrawals from some past period of consumption. After this global budget is established, the Army Corps of Engineers could grant water allocations to each user based on average daily use, again from some period of past consumption. Even though hydropower producers, barges, and oyster producers do not strictly divert water, the water they use is "diverted" from the Basin into the Gulf of Mexico and therefore should be measured for the allocation. By making the allocations daily, the Corps would allow for seasonal variations in demand and flood control. Permit allocations would be divisible and transferable. And, of course, under no condition could water be allocated in a way that violates federal water use laws.

When the supply of rainfall was abundant, so that water in the Basin exceeded the global daily budget, all users could be satisfied without the need to transfer water among users. However, in the case of drought, the Army Corps of Engineers could cut daily permit allowances by an equal percentage for all users. The Corps could then serve as a water broker, facilitating transactions among users by matching buyers and sellers and helping to negotiate terms of exchange, while charging a fee to cover administrative costs. Similarly, if the demand for water rises to the point that it exceeds the global daily budget, users who want more water would have to purchase that water from other users.

To see how this might work, consider a simplified example with two users, a lake, and a river running downstream from the lake. Suppose the two users are Atlanta developers and Florida oyster producers, the lake named Lake Lanier, and the river the Chattahoochee. Suppose that for a given day, the water budget for this river basin is 1000 gallons, allocated between 800 gallons for Atlanta developers and 200 gallons for Florida oyster producers. (Actual quantities would, of course, be in the millions of gallons per day.)

If rainfall allows greater net withdrawals, say to 1100 gallons, each user's allotment can rise by 10%. A drought, however, might reduce net withdrawals to 900 gallons, forcing cuts in permitted allotments to 720 gallons for Atlanta developers and 180 gallons for Florida oyster producers. This is where trading comes in. If Atlanta's developers want to restore their allocation, they must offer to purchase an additional 80 gallons from Florida's oyster producers. If the contracting parties agree, the Army Corps will release less water from Lake Lanier, increasing the amount available to Atlanta's developers and reducing the downstream flow for Florida's oyster producers.

If, in contrast, Florida's oyster producers want to retain a flow of 200 gallons, they will offer to purchase an additional 20 gallons from Atlanta developers. If the parties agree, this time the Army Corps of Engineers will release more water from Lake Lanier, reducing the amount available for Atlanta's developers but increasing the flow for Florida's oyster producers.

One can envision associations of users with similar wants, such as an upstream association of developers and recreational users and a downstream association of hydroelectric utilities, barges, environmentalists, and oyster producers. At times, association members would benefit by combining funds and sharing costs.

Purchasing water allotments to retire (i.e., not use) them should also be allowed. For example, if electric utilities want to increase downstream flows to generate electricity at the same time that recreational users want lake levels held high, as on a summer holiday weekend, the recreational users could purchase water rights from the electric utilities, if the utilities agreed, and retire those rights. Instead of producing revenues through hydropower, the electric utilities would receive payments from recreational users. Similarly, environmentalists might want to purchase and use or retire rights during seasons when fish spawn.[14]

In each of the exchanges described, the amount of water traded would be a small portion, not all of the total allowances. For example, recreational users would be likely to purchase some, but not all, of the electric utilities' water. Lake levels would fall enough to generate some electricity, but not as much as they would fall if recreational users did not purchase some of the water rights.[15]

To be effective, these marketable permits must have the key characteristics of property rights: They must be clearly defined, so that each user knows its allocation for each day; enforced, with the Army Corps of Engineers serving as enforcer of the permit allowances through its monitoring of lakes and dams; and transferable, with transfer facilitated by the Army Corps of Engineers serving as broker. With defined, enforced, and transferable property rights, a water market could develop that would ensure an allocation of water to its most productive uses and peace among contracting parties.

As an alternative to water transfers among users at mutually agreed upon prices, the Corps could advance market allocation by establishing a water bank. The Corps could serve as an underwriter that buys and sells water at specified prices, with the spread between these prices used to cover the costs of administering the bank. Such banks have been used in times of drought. For example, water banks were used successfully in 1977 and 1991 in California to cope with drought. In 1991, California offered to purchase water at a price of $125 per acre-foot and to sell water at a price of $175 per acre-foot. The state purchased and sold 400,000 acre-feet of water, mostly to municipal and agricultural users (31, pp. 11–12, 102–103).

In the ACF River Basin, the Army Corps of Engineers could assign to users daily property rights to flows of water, based on historic use patterns, and then serve as a water banker, standing ready to buy and sell water at specified prices. Depositors could leave water in the Basin, and withdrawers could buy it. With price playing an allocative role and with voluntary transactions, the Basin's waters would be allocated more efficiently, and relations among the ACF River Basin's users would become more harmonious.

Implications for the Southeast

Making these kinds of changes in the ACF River Basin is critical because water conflicts are brewing throughout the Southeast. Along Georgia's border with South Carolina, a request by Habersham County, Georgia, to withdraw 12.5 million gallons per day from the Savannah River Basin provoked the South Carolina state legislature to introduce resolutions calling on Congress to stop the Army Corps of Engineers from granting the request, which would have transferred water from the Savannah River Basin to the ACF Basin.[16] In addition, Georgia is involved in a dispute with Alabama over water in the Alabama-Coosa-Tallapoosa River Basin. And North Carolina and South Carolina have disputed the flow of water in the Yadkin-Great Pee Dee River Basin (35–37). By establishing water markets in the ACF River Basin, the states of Alabama, Florida, and Georgia could serve as an example to other southeastern states to help them avoid the conflicts that have for so long plagued the attempts to allocate that Basin's waters.

WHAT CHANCE FOR CHANGE?

Market reforms offer great potential, but when government is in control, change typically comes about only with crisis. Higher rainfall since the summer of 2002 has diminished the sense of crisis. Does this mean that all hope is lost for market allocation of water in the ACF River Basin or elsewhere in the Southeast? Not at all.

By failing to achieve compact resolution, the states of Alabama, Florida, and Georgia have embarked down the risky path of judicial apportionment. The risk is that the Supreme Court could allocate the ACF River Basin's waters in a way that is unsatisfactory to each or all of the states (17, p. 226). Because these allocations are not transferable, states with an unsatisfactory allocation would have no way, short of further litigation, to change the allocation. As the states contemplate this possibility, they may find it prudent to drop their lawsuits and pursue a means of allocating water that relies on markets, such as one of the proposals offered in this article.

Once demand permanently outstrips supply under current arrangements, water in the Southeast will be rationed. The question will be how. Will water be rationed by markets, which promote productive use and harmony among users? Or will it be rationed by political processes that are likely to result in misallocation and conflict? As economic development continues, perhaps plagued by drought, the citizens of the Southeast may choose the efficiency and harmony of markets over the misallocation and contention of politics.

BIBLIOGRAPHY

1. Shelton, S. (2003). Water talks a washout—states point fingers. *Atlanta Journal-Constitution* September 6, G1.

---

[14] Retiring rights requires some ranking among users. If hydropower users have the higher ranking, recreational users would have to purchase rights from them to keep lake levels up. On the other hand, if recreational users had the higher ranking, hydropower users would have to purchase rights from them to drop lake levels. Coase (34) argues that clearly defined property rights and sufficiently low transaction costs will lead to allocating resources to their most productive uses.

[15] With less water for hydroelectric production, utilities might have to raise prices to their customers.

[16] Before this conflict could escalate, Habersham County withdrew its permit request.

2. U.S. Geological Survey. (2000). Apalachicola-Chattahoochee-Flint River Basin NAWQA study: Description of the ACF River Basin study area, 2000 (last modified August 23). Online: ga.water.usgs.gov/nawqa/main.description.html (cited Sept. 24, 2003).
3. Jeane, D.G. (2002). *A History of the Mobile District Corps of Engineers, 1815–1985*. Online: www.sam.usace.army.mil/MobileDistrictHistory (cited December 10, 2003).
4. Carriker, R.R. (2000). *Water Wars: Water Allocation Law and the Apalachicola-Chattahoochee-Flint River Basin*. University of Florida, Cooperative Extension Service, Institute of Food and Agricultural Services. Online: edis.ifas.ufl.edu/BODY_FE208 (cited Sept. 24, 2003).
5. Shelton, S. (2003). Water deal may buy Atlanta some time. *Atlanta Journal-Constitution* September 12, C3.
6. U.S. Army Corps of Engineers, Mobile District. (2001). How the river systems are managed (Dec. 2, 2001). Online: water.sam.usace.army.mil/narrativ.htm (cited Sept. 30, 2003).
7. Beaverstock, J.U. (1998). Learning to get along: Alabama, Georgia, Florida and the Chattahoochee River Compact. *Ala. Law Rev.* **49** (Spring): 993–1007.
8. Census Bureau, U.S. (2002). *Statistical Abstract of the United States: 2002*. Government Printing Office, Washington, DC.
9. Metro Atlanta Chamber of Commerce and the Regional Business Coalition. (2003). *A Primer on Water Resources in the Metro Atlanta Region*. Online: www.cleanwaterinitiative.com/background/primer.htm (cited October 28, 2003).
10. Atlanta Regional Commission. (2003). *Regional Water Supply Plan*. Online: atlreg.com/water/supplyplan.html (cited October 28, 2003).
11. Seabrook, C. (2002). Atlanta guzzling water; metro thirst exceeds projections. *Atlanta Journal-Constitution* May 15.
12. Walker, B.P. (2001). Using geographic information system mapping and education for watershed protection through better-defined property rights. In: *The Technology of Property Rights*, T.L. Anderson and P.J. Hill. (Eds.). Rowman & Littlefield, Lanham, MD, pp. 57–78.
13. *Economist*. (1991). River rivalry. March 30.
14. Hull, J.W. (2000). The War Over Water. *Regional Resource*. Council of State Governments, Atlanta.
15. American Rivers. (2002). *America's Most Endangered Rivers of 2002*. Online: www.amrivers.org/mostendangered/2002report.htm (cited December 10, 2003).
16. Apalachicola Bay Chamber of Commerce. (2003). *Apalachicola*. Online: www.apalachicolabay.org (cited Sept. 25, 2003).
17. Erhardt, C. (1992). The battle over "The Hooch": The federal-interstate water compact and the resolution of rights in the Chattahoochee river. *Stanford Environ. Law J.* **11**: 200–228.
18. Marella, R.L., Fanning, J.L., and Mooty, W.S. (1993). *Estimated Use of Water in the Apalachicola-Chattahoochee-Flint River Basin During 1990 with State Summaries from 1970 to 1990*. Water-Resources Investigation Report, 93–4084. U.S. Geological Survey, Tallahassee, FL.
19. Moore, C.G. (1999). Water wars: Interstate water allocation in the Southeast. *Nat. Resour. Environ.* **14**(Summer): 5–10; 66–67.
20. Alabama, Florida, and Georgia. (2003). *Memorandum of Understanding Regarding Initial Allocation Formula for the ACF River Basin (July 22)*. Online: www.dep.state.fl.us/secretary/comm/2003/july/0722_acf.htm (cited Sept. 29, 2003).
21. Struhs, D.B. (2003). *Statement of Intent to Accompany the Memorandum of Understanding Regarding Initial Allocation Formula for the ACF River Basin, July 22*. Online: www.dep.state.fl.us/secretary/comm/2003/july/0722_acf.htm (cited Sept. 29, 2003).
22. Ritchie, B. (2003). River pact moves closer. *Tallahassee Democrat* July 23, A1.
23. Ritchie, B. (2003). Florida willing to take battle to court. *Tallahassee Democrat* August 27, B3.
24. Ritchie, B. (2003). High court may hear water fight. *Tallahassee Democrat* September 2, A1.
25. Seabrook, C. (2003). Water Costs Likely to Rise. *Atlanta Journal-Constitution* Sept. 8, F1.
26. U.S. Census Bureau. 1998. *State and Metropolitan Area Data Book, 1997–98*. 5th Edn. Washington, DC.
27. Anderson, T.L. and Snyder, P.S. (1997). Priming the invisible pump. *PERC Policy Series*, PS-9. PERC, Bozeman, MT.
28. Atlanta Bureau of Water. (2003). *Rates, Fees & Meter Price*. Online: www.atlantaga.gov/citydir/water/index.htm (cited Oct. 28, 2003).
29. Atlanta Regional Commission (2003). *Current Water Restrictions (May 9)*. Online: atlreg.com/water/waterrestrictions.html (cited November 6, 2003).
30. Judd, A. (2000). State slaps water limits on 15 metro counties: outdoor water use banned from 10 a.m. to 10 p.m. *Atlanta Journal-Constitution* June 2.
31. Anderson, T.L. and Snyder, P.S. (1997). *Water Markets: Priming the Invisible Pump*. Cato Institute, Washington, DC.
32. Yoskowitz, D.W. (2001). Markets, mechanism, institutions, and the future of water. *Environ. Law Rep.* **31**(February).
33. Fretwell, H.L. (1999). Paying to play: The fee demonstration program. *PERC Policy Series*, PS-17. PERC, Bozeman, MT.
34. Coase, R. (1960). The problem of social cost. *J. Law Econ.* **3**(October): 1–44.
35. Henderson, Bruce. (2002). Who gets the water? *Charlotte Observer* Dec. 29.
36. Libaw, O.Y. (2000). Water wars: Drought-ridden Southeast battles over use of rivers (August 14). ABC News. Online: more.abcnews.go.com/sections/us/dailynews/water000811.html (cited Sept. 16, 2003).
37. Pompe, J.J. and Franck, D.P. (2003). *The Economic Impact of Water Transfer: Options for Policy Reform*. Working paper, Department of Economics, Francis Marion University, Florence, SC.

WATER SUPPLY AND WATER RESOURCES: DISTRIBUTION SYSTEM RESEARCH

Mark C. Meckes
Roy C. Haught
U.S. Environmental
Protection Agency

Two water distribution system simulators (DSS) are in operation at the U.S. EPA Test and Evaluation (T&E) Facility in Cincinnati, Ohio. The T&E Facility is a multifaceted research resource in which a wide variety

---

of water treatment and other environmental protection technologies are conceived, designed, and evaluated in the laboratory, bench, and pilot plant scale. The EPA designed and fabricated the DSS systems to evaluate and understand the dynamics which influence water quality within water distribution infrastructure systems in the United States and abroad. The first distribution system simulator (DSS-1) has continuous flow conditions. There are six individual 75 ft (25 meter) lengths of 6 in (15 cm) diameter ductile iron pipe arranged into "pipe loop" configurations to simulate a distribution system. This pipeloop system can be configured to operate as; independent/individual loops, collectively as one unit, or in various experimental configurations to complement EPA's experimental research. DSS-1 is equipped with two 1500 gallon reservoir tanks to simulate a comprehensive distribution infrastructure system. This unique engineering design permits operating any combination and configuration of the six loops under various experimental operating parameters (Table 1). Each loop is insulated and fitted with a heat exchanger to maintain constant temperature conditions during operation. The DSS is interfaced with a Supervisory Control and Data Acquisition (SCADA) system, which is used to monitor, control, and archive operating conditions and collected data continuously. Bioflim samples are collected on coupons [View1–View2] which are set flush with the interior surface of the pipe wall. The coupons may be removed for sampling and analysis without disrupting water flow.

The second distribution system (DSS-2) is over 300 feet long and is a once through system composed of six inch diameter PVC pipe. This unit is being utilized to evaluate water quality in a dead end branch of a distribution system and to develop design models for water distribution. Both DSS units are located above ground to permit easy access to the entire pipe network. Experimental studies (Table 2) are ongoing to understand the physical, chemical, and biological activities that occur in drinking water distribution systems.

Table 1. Operating Parameters Of The Distribution System Simulator

Parameter	Normal Operation	Experimental Test Conditions
Distribution system simulator	Parallel (6 individual distribution system simulators)	Parallel or series in groups of 6, 3, 2, or 1
Housing	Ductile iron (Non-lined)	Ductile iron (Non-lined)
Flow	88 gpm or 1 ft/s	No Flow to 140 gpm or 0 to 1.6 ft/s
Temperature	60 °F (15.5 °C)	35 °F (1.6 °C) to Ambient
Chemical control	Free chlorine 1.0 ppm	Chemical control as needed
pH	7.0 to 7.5	Control/monitor as needed
Turbidity	< 0.5 NTU	Control/monitor as needed
Water supply	Cincinnati—tap water (Chlorinated)	Dechlorinated, deionized, tanked, surface water (River)

OBJECTIVES

The DSS provides researchers with a mechanism to study how water quality is affected during distribution. Results from research studies (Table 2) will be used to provide guidance on how to maintain a high level of water quality during distribution. A secondary objective to this project is to develop, evaluate, and demonstrate real time monitoring of water quality parameters within distribution systems using remote telemetry. Results from research will be used to provide guidance on how to utilize remote monitoring of water quality to detect changes in water quality within distribution systems.

RELEVANCE

Throughout the world there are millions of miles of water distribution pipe lines which provide drinking water for use by individuals and industry. Although these distribution systems provide drinking water to the world, very little is known about the physical, chemical, and biological activities that occur within them. Some of these water distribution systems have been in service well over one hundred years. DSS-1 has been in operation for over three years, as water moves through distribution systems, it comes into contact with a wide range of material some of which can cause significant changes to the quality of the finished water supply. Suspended solids in finished water can settle out under low flow conditions and can be suspended during high flow. Various disinfection agents and water additives react with organic and inorganic materials within a distribution system generating by-products which may be undesirable in the water supply. Oxidant resistant microorganisms colonize pipe surfaces producing a complex micro environment known as "biofilm". Biofilms are highly resistant to many disinfection methods and techniques. Our research looks at the various experimental test parameters which influence biofilm growth. This work will also help EPA

Table 2. Proposed DSS Studies

No	Title of Proposed Study
1	Preliminary Studies of Biofilm Formation in Pilot-Scale Distribution Systems
2	Opportunistic Pathogens in Biofilms
3	Effect of a Pollution Event on a Simulated Water Distribution System
4	Impact of Nutrient Removal on Growth Potential for Bacteria
5	Impact of Alternative Treatment on Biofilm Growth
6	Real-Time Monitoring and Control of Distribution Systems
7	Effects of pH Changes on Biofilm Growth in a Distribution System
8	Bacterial Growth in Distribution Systems

develop a better understanding of the dynamics inside distribution systems.

RESEARCH GOALS

Fabricate aboveground water distribution system simulator's which permit easy access and can be operated under controlled conditions. Conduct studies to develop a better understanding of the dynamics that occur in drinking water distribution systems. Determine what physical, chemical, and biological factors influence biofilm growth within such systems. Develop and test mechanisms for the enhancement and control of biofilm growth within a simulated distribution system. Additionally, develop and evaluate real time monitoring, data collection, and archiving of water quality parameters within water distribution systems using remote telemetry. Results will be used to develop and provide guidance on ways to maintain high levels of water quality through distribution systems.

DROUGHT IN THE DUST BOWL YEARS

National Drought Mitigation Center

INTRODUCTION

In the 1930s, drought covered virtually the entire Plains for almost a decade (1). The drought's direct effect is most often remembered as agricultural. Many crops were damaged by deficient rainfall, high temperatures, and high winds, as well as insect infestations and dust storms that accompanied these conditions. The resulting agricultural depression contributed to the Great Depression's bank closures, business losses, increased unemployment, and other physical and emotional hardships. Although records focus on other problems, the lack of precipitation would also have affected wildlife and plant life and would have created water shortages for domestic needs.

Effects of the Plains drought sent economic and social ripples throughout the country. For example, millions of people migrated from the drought areas, often heading west, in search of work. These newcomers were often in direct competition for jobs with longer-established residents, which created conflict between the groups. In addition, because of poverty and high unemployment, migrants added to local relief efforts, sometimes overburdening relief and health agencies.

Many circumstances exacerbated the effects of the drought, among them the Great Depression and economic overexpansion before the drought, poor land management practices, and the areal extent and duration of the drought. [Warrick et al. (2) and Hurt (3) discuss these issues in greater detail.] The peculiar combination of these circumstances and the severity and areal coverage of the event played a part in making the 1930s drought the widely accepted drought of record for the United States. To cope with and recover from the drought, people relied on ingenuity and resilience, as well as relief programs from state and federal governments. Despite all efforts, many people were not able to make a living in drought-stricken regions and were forced to migrate to other areas in search of a new livelihood. It is not possible to count all the costs associated with the drought of the 1930s drought, but one estimate by Warrick (1) claims that financial assistance from the government may have been as high as $1 billion (in 1930s dollars) by the end of the drought. Fortunately, several lessons were learned that were used in reducing the vulnerability of the regions to future droughts.

THE GREAT DEPRESSION AND OVERCAPITALIZATION

In the early 1920s, farmers saw several opportunities for increasing their production. New technology and crop varieties were reducing the time and costs-per-acre of farming, which provided a great incentive for agricultural expansion. This expansion was also necessary to pay for expensive, newly developed equipment (such as listers and plows) that was often purchased on credit, and to offset low crop prices after World War I.

When the national economy went into decline in the late 1920s because of the Great Depression, agriculture was even more adversely affected. In addition, a record wheat crop in 1931 sent crop prices even lower. These lower prices meant that farmers needed to cultivate more acreage, including poorer farmlands, or change crop varieties to produce enough grain to meet their required equipment and farm payments.

When drought began in the early 1930s, it worsened these poor economic conditions. The depression and drought hit farmers on the Great Plains the hardest. Many of these farmers were forced to seek government assistance. A 1937 bulletin by the Works Progress Administration reported that 21% of all rural families in the Great Plains were receiving federal emergency relief (4). However, even with government help, many farmers could not maintain their operations and were forced to leave their land. Some voluntarily deeded their farms to creditors, others faced foreclosure by banks, and still others had to leave temporarily to search for work to provide for their families. In fact, at the peak of farm transfers in 1933–34, nearly 1 in 10 farms changed possession, with half of those being involuntary (from a combination of the depression and drought).

POOR LAND MANAGEMENT PRACTICES

A number of poor land management practices in the Great Plains region increased the vulnerability of the area before the 1930s drought. Some of the land use patterns and cultivation methods in the region can be traced back to the settlement of the Great Plains nearly 100 years earlier. At that time, little was known of the region's

---

climate. Several expeditions had explored the region, but they were not studying the region for its agricultural potential, and, furthermore, their findings went into government reports that were not readily available to the general public (5). Misleading information, however, was plentiful. "Boosters" of the region, hoping to promote settlement, put forth glowing but inaccurate accounts of the Great Plains' agricultural potential. In addition to this inaccurate information, most settlers had little money or other assets, and their farming experience was based on conditions in the more humid eastern United States, so the crops and cultivation practices they chose often were not suitable for the Great Plains. But the earliest settlements occurred during a wet cycle, and the first crops flourished, so settlers were encouraged to continue practices that would later have to be abandoned. When droughts and harsh winters inevitably occurred, there was widespread economic hardship and human suffering, but the early settlers put these episodes behind them once the rains returned. Although adverse conditions forced many settlers to return to the eastern United States, even more continued to come west. The idea that the climate of the Great Plains was changing, particularly in response to human settlement, was popularly accepted in the last half of the 19th century. It was reflected in legislative acts such as the Timber Culture Act of 1873, which was based on the belief that if settlers planted trees they would be encouraging rainfall, and it was not until the 1890s that this idea was finally abandoned (6). Although repeated droughts tested settlers and local/state governments, the recurrence of periods of plentiful rainfall seemed to delay recognition of the need for changes in cultivation and land use practices.

Several actions in the 1920s also increased the region's vulnerability to drought. Low crop prices and high machinery costs (discussed above) meant that farmers needed to cultivate more land to produce enough to meet their required payments. Since most of the best farming areas were already being used, poorer farmlands were increasingly used. Farming submarginal lands often had negative results, such as soil erosion and nutrient leaching. By using these areas, farmers were increasing the likelihood of crop failures, which increased their vulnerability to drought.

These economic conditions also created pressure on farmers to abandon soil conservation practices to reduce expenditures. Furthermore, during the 1920s, many farmers switched from the lister to the more efficient one-way disc plow, which also greatly increased the risk of blowing soil. Basically, reductions in soil conservation measures and the encroachment onto poorer lands made the farming community more vulnerable to wind erosion, soil moisture depletion, depleted soil nutrients, and drought.

DROUGHT DURATION AND EXTENT

Although the 1930s drought is often referred to as if it were one episode, there were at least four distinct drought events: 1930–31, 1934, 1936, and 1939–40 (7). These events occurred in such rapid succession that affected regions were not able to recover adequately before another drought began. Historical maps of U.S. climate divisions and graphs of U.S. river basins reflect this situation.

COPING WITH AND RECOVERING FROM DROUGHT

During the 1930s, many measures were undertaken to relieve the direct impacts of droughts and to reduce the region's vulnerability to the dry conditions. Many of these measures were initiated by the federal government, a relatively new practice. Before the 1930s drought, federal aid had generally been withheld in emergency situations in favor of individual and self-reliant approaches. This began to change with the development of the Great Depression in the late 1920s and the 1933 inauguration of President Franklin Delano Roosevelt. The depression helped "soften deep-rooted, hard-line attitudes of free enterprise, individualism, and the passive role of the government", thus paving the way for Roosevelt's New Deal programs, which in turn provided a framework for drought relief programs for the Great Plains (1).

Warrick et al. (2) describe these drought relief programs, which are credited with saving many livelihoods throughout the drought periods. The programs had a variety of goals, all of which were aimed at the reduction of drought impacts and vulnerability:

- Providing emergency supplies, cash, and livestock feed and transport to maintain the basic functioning of livelihoods and farms/ranches.
- Establishing health care facilities and supplies to meet emergency medical needs.
- Establishing government-based markets for farm goods, higher tariffs, and loan funds for farm market maintenance and business rehabilitation.
- Providing the supplies, technology, and technical advice necessary to research, implement, and promote appropriate land management strategies.

As important as these programs may have been, the survival of a majority of the families and enterprises undoubtedly rested solely with their perseverance and integrity. Whether they stayed or moved into the drought regions or migrated to other areas in hopes of a better life, families encountered new hardships and obstacles that would require ingenuity, resilience, and humility. Those who remained in the drought regions were forced to endure severe dust storms and their health effects, diminished incomes, animal infestations, and the physical and emotional stress over their uncertain futures. Humor helped; tales about birds flying backward to keep from getting sand in their eyes, housewives scouring pots and pans by holding them up to keyholes for a sandblasting, and children who had never seen rain were among the favorite stories of Dust Bowl inhabitants. In the end, it was a combination of willpower, stamina, humor, pride, and, above all, optimism that enabled many to survive the Dust Bowl. These qualities are succinctly expressed in the comments of one contemporary Kansan: “We have faith in the future. We are here to stay” (quoted in Ref. 3).

The 1930s drought and its associated impacts finally began to abate during spring 1938. By 1941, most areas of the country were receiving near-normal rainfalls. These rains, along with the outbreak of World War II, alleviated many of the domestic economic problems associated with the 1930s. In fact, the new production demands and positive climatic conditions brought the United States into a rapid economic boom.

Even though short-term conditions seemed to be relatively stable, there were some drawbacks to this production growth. One drawback (described in Ref. 3) was that the start of World War II shifted remaining funds and priorities away from drought-related programs. Men were taken off work programs to enter the armed forces and to produce for the war effort. Moreover, items such as gasoline and replacement parts were redirected from federal drought and conservation programs to the war efforts. This meant that conservation programs and research were significantly reduced during this period. Another drawback was that with the return of the rains, many people soon forgot about conservation programs and measures implemented during the 1930s droughts. This led to a return to some of the inappropriate farming and grazing practices that made many regions so vulnerable to drought in the 1930s.

1930s DROUGHT COSTS

Although the 1988–89 drought was the most economically devastating natural disaster in the history of the United States (7), a close second is undoubtedly the series of droughts that affected large portions of the United States in the 1930s. Determining the direct and indirect costs associated with this period of droughts is a difficult task because of the broad impacts of drought, the event’s close association with the Great Depression, the fast revival of the economy with the start of World War II, and the lack of adequate economic models for evaluating losses at that time. However, broad calculations and estimates can provide valuable generalizations of the economic impact of the 1930s drought.

Overall Drought Costs

In 1937, the Works Progress Administration (WPA) reported that drought was the principal reason for economic relief assistance in the Great Plains region during the 1930s (4). Federal aid to the drought-affected states was first given in 1932, but the first funds marked specifically for drought relief were not released until the fall of 1933. In all, assistance may have reached $1 billion (in 1930s dollars) by the end of the drought (1).

According to the WPA, three-fifths of all first-time rural relief cases in the Great Plains area were directly related to drought, with a disproportionate amount of cases being farmers (68%) and especially tenant farmers (70% of the 68%). However, it is not known how many of the remaining cases (32%) were indirectly affected by drought. The WPA report also noted that 21% of all rural families in the Great Plains area were receiving federal emergency relief by 1936 (4); the number was as high as 90% in hard-hit counties (1). Thus, even though the exact economic losses are not known for this time period, they were substantial enough to cause widespread economic disruption that affected the entire nation.

LESSONS LEARNED: THE LEGACY OF THE 1930s DROUGHT

The magnitude of the droughts of the 1930s, combined with the Great Depression, led to unprecedented government relief efforts. Congressional actions in 1934 alone accounted for relief expenditures of $525 million (8); the total cost (social, economic, and environmental) would be impossible to determine.

If the Roosevelt era marked the beginning of large-scale aid, it also ushered in some of the first long-term, proactive programs to reduce future vulnerability to drought. It was in these years, for example, that the Soil Conservation Service (SCS)—now the Natural Resources Conservation Service—began to stress soil conservation measures. Through their efforts, the first soil conservation districts came into being, and demonstration projects were carried out to show the benefits of practices such as terracing and contouring (for a discussion of the activities of the SCS during this period, see Ref. 3).

Warrick et al. (2) note that the proactive measures continued in the years following the drought: conservation practices and irrigation increased, farm sizes grew larger, crop diversity increased, federal crop insurance was established, and the regional economy was diversified. Many other proactive measures taken after the 1930s drought also reduced rural and urban vulnerability to drought, including new or enlarged reservoirs, improved domestic water systems, changes in farm policies, new insurance and aid programs, and removal of some of the most sensitive agricultural lands from production (7).

Problems remained, but these programs and activities would play a fundamental role in reducing the vulnerability of the nation to the forthcoming 1950s drought. Although a larger area was affected during the 1950s drought, the conservation techniques that many farmers implemented in the intervening years helped prevent conditions from reaching the severity of the 1930s drought.

BIBLIOGRAPHY

1. Warrick, R.A. 1980. Drought in the Great Plains: A Case Study of Research on Climate and Society in the USA. In J. Ausubel and A.K. Biswas (eds.). *Climatic Constraints and Human Activities*, pp. 93–123. *IIASA Proceedings Series*, Vol. 10. Pergamon Press, New York.
2. Warrick, R.A.; P.B. Trainer; E.J. Baker; and W. Brinkman. 1975. *Drought Hazard in the United States: A Research Assessment*. Program on Technology, Environment and Man Monograph #NSF-RA-E-75-004, Institute of Behavioral Science, University of Colorado, Boulder.

3. Hurt, D.R. 1981. *An Agricultural and Social History of the Dust Bowl*. Nelson Hall, Chicago.
4. Link, I.; T.J. Woofter, Jr., and C.C. Taylor. 1937. *Research Bulletin: Relief and Rehabilitation in the Drought Area*. Works Progress Administration, Washington, DC.
5. Fite, G.C. 1966. *The Farmer's Frontier, 1865–1900*. Holt, Rinehart and Winston, New York.
6. White, R. 1991. *"It's Your Misfortune and None of My Own": A History of the American West*. University of Oklahoma Press, Norman.
7. Riebsame, W.E.; S.A. Changnon, Jr.; and T.R. Carl. 1991. *Drought and Natural Resources Management in the United States: Impacts and Implications of the 1987–89 Drought*. Westview Press, Boulder, Colorado.
8. United States House of Representatives. 1934. *Relief of the drought area*. Communication from the President of the United States, 73d Congress, 2d Session, Document No. 398, Washington, DC.

OTHER RESOURCES

Dust Storms and Their Damage. Photos from the collection of the Wind Erosion Research Unit of USDA's Agricultural Research Service at Kansas State University.

The Dust Bowl (photos). A collection of 20 photos for sale from The Social Studies Web.

American Memory from the Library of Congress. A multi-format look at daily life in Farm Security Administration migrant work camps in central California, 1940 and 1941.

Weedpatch Camp. The story of the Arvin Federal Government Camp, immortalized as "Weedpatch Camp" in Steinbeck's Grapes of Wrath.

Dust storm footage. A short film clip of a dust storm during the 1930s. From the website of the USDA/ARS Wind Erosion Research Unit at Kansas State University.

Surviving the Dust Bowl. The website of the 1998 Public Broadcasting Service (PBS) film "Surviving the Dust Bowl". Provides a good background on the people and events of the 1930s Dust Bowl.

1930s Dust Bowl. Excerpts from "The Dust Bowl, Men, Dirt and Depression" by Paul Bonnifield. On the Cimarron Heritage Center (Boise City, Oklahoma) website.

Woody Guthrie Dust Bowl Ballads. Discussion and analysis of Guthrie's "Dust Bowl Ballads" recording.

America from the Great Depression to World War II. Photos from the Farm Security Administration—Office of War Information Collection, 1935–1945.

The Plow that Broke the Plains. The story behind the 1936 documentary made for the Farm Security Administration on the destruction and suffering associated with drought in the early 1930s.

DROUGHT MANAGEMENT PLANNING

Neil S. Grigg
Colorado State University
Fort Collins, Colorado

It is impossible to "manage" a naturally occurring drought, so "drought management" refers to managing resources, including water, to mitigate the adverse effects of droughts, which are certain to occur. Other resources that can be managed are food supplies, including animal feed, public facilities of various kinds, and economic resources to aid businesses, farmers, and citizens.

As "creeping disasters," droughts are easy to ignore until too late, thus presenting a significant risk to water resources managers. The dilemma is that to prepare for them might be expensive and require seldom-used facilities, but without preparation, water supplies might run out on rare occasions. For this reason, drought management planning presents a challenge to water agencies, as well as to other governmental and private sector organizations.

Drought has two forms: a period without enough rain and a period of shortage: each requires drought management planning. A period without enough rain is a "meteorologic drought," and leads to shortages, which are called by names such as "hydrologic," "agricultural," and "socioeconomic" drought.

Drought management planning is a form of contingency planning that answers the question "what do we do if a drought occurs?" As in other forms of contingency planning, the process is to assess the threat, identify the vulnerable parts of systems, and take measures to prepare, mitigate, respond, and recover from the impact of a drought, should one occur.

Failures in water supply can have serious consequences for cities, industries, and other water users such as irrigation, hydropower, recreation, and wildlife. The primary impact of drought is due to real or feared interruption of supplies because water supply is critical to the economy and the natural environment. This risk increases with the interdependence and vulnerability of water systems, environmental stakes are also high. Drought is a serious threat to food supplies and farm income, especially in nations where food supplies are marginal.

Assessing the threat of drought requires knowledge of the security of supplies or the probability that a raw water supply system will run out of water. This is usually estimated for individual water agencies in terms of the return period of the drought planned for or the annual probability of running short. The concepts of return period and failure probability are in wide use, but for droughts they are complex because drought duration, system yield, and return period must be considered.

The concept of "safe yield" is used to describe the reliability of a water system, another measure of security. To estimate safe yield requires analysis of components and the systems that deliver water to users. Vulnerability analysis requires assessing of all possible modes of failure, not only hydrologic failure. For example, if a dam requires emptying for repair, supplies may run short from lack of storage, even when they are available from precipitation.

Examples of measures to prepare, mitigate, respond, and recover from drought impacts are shown in Table 1 (1).

Drought response plans are usually prepared by water management organizations to anticipate drought and plan activities to take place after a drought occurs. These are custom-tailored to the needs of each

Table 1. Drought Planning

Supply Augmentation	Demand Reduction	Impact Minimization
New sources	Legal restrictions	Forecasting
Water storage	Pricing	Mutual aid
Reuse of water	Devices to limit use	Insurance

organization but generally include the usual contents of emergency response plans, such as assessment of the hazard, identification of vulnerable components, mitigation measures, plans to meet critical water needs, arrangements for mutual aid, team organization and roles, and special conditions.

BIBLIOGRAPHY

1. Grigg, N.S. and Vlachos, E.C. (1993). Drought and water-supply management: roles and responsibilities. *J. Water Resour. Plann. Manage. American Society of Civil Engineers*. September/October.

READING LIST

Wilhite, D.A., Easterling, W.E. and Wood, D.A. (Eds.) (1987). *Planning for Drought: Toward a Reduction in Societal Vulnerability*, Westview Press, Boulder, CO.

DROUGHT AND WATER SUPPLY MANAGEMENT

Neil S. Grigg
Colorado State University
Fort Collins, Colorado

Drought is a normal, albeit extreme, aspect of hydrologic variation, so it should be considered when making plans for water. Planning for water supply management involves both supply augmentation and demand management. Supply augmentation can involve new supplies or extending use of existing supplies.

During drought, making water available to meet requirements of people, business, and the environment is critical. As explained in the article on "Drought Management Planning," preparing for drought requires contingency planning to assure adequacy of water supplies, and management of available water supplies during drought is the key element of drought response. Thus, effective water resources management requires actions that assure water availability during drought as well as normal times.

Technologies and methods for providing water supplies go back to early Rome, which brought in water supplies by aqueducts. Water supply systems in the United States date from early settlements, which used spring water pumped through bored logs. Some early pumps were horse-driven, followed later by steam power.

Water supply systems have three subsystems: source of supply, treatment and distribution, and delivery of enough water of high quality at sufficient pressure for intended uses. A percentage of the water is normally unaccounted for, mainly through leakage and other losses. Water for fire fighting must also, be provided. Providing enough water requires volumes of delivery to meet needs reliably during peak demand periods and during drought.

The required quantities of water can be estimated from statistics of water use for domestic, commercial, and industrial uses. These vary widely, depending on many socioeconomic factors. Average per capita use of water in the United States, including all uses, varies from about 100–200 gallons per capita per day (gpcd) (378–756 liters per capita per day). The highest rates are normally in dry regions, where water supplies are used for lawn irrigation. In highly concentrated urban areas, water uses may be mostly inside apartment buildings and in commercial facilities and may be close to 100 gpcd (378 liters per capita per day), without any requirement for lawn irrigation. In other countries, water demands may be less, depending on socioeconomic variables. In a developing country city, use might be much less than levels in the United States, and water may not be available to all residents.

Sources of freshwater supply include surface water, groundwater, and reclaimed waters. Surface water can include stored water in reservoirs or direct diversions from streams. Typical infrastructure components include dams, tunnels, outlet tubes, canals, gates and controls, spillways, and support structures. Groundwater sources include springs, wells, infiltration galleries, and aquifers that store recharged water. Infrastructure components include wells, casings, pumping systems, piping, housing, and other support facilities. Given the increasing restrictions on the development of surface water, groundwater development is receiving more attention, including the implementation of aquifer storage and recovery (ASR) systems.

In special cases, rainwater can be caught on roofs and stored in cisterns. This type of local supply system is found mostly in developing countries and in special, remote locations where centralized systems are not possible. Bottled water is an increasingly popular source, and point-of-use treatment systems are becoming more popular.

Demand reduction can be legal, physical, or voluntary. Drought plans normally contain provisions to restrict use during times of shortage. Physical controls can include reduced pressures and requirements for devices to impose water conservation. Voluntary systems rely primarily on education and calls for citizenship.

As population increases and environmental water needs are more recognized, it becomes more difficult to find new sources of supply. Therefore, innovative approaches being studied include

- dual use of water where reclaimed and impaired waters are used for nonpotable applications
- conservation systems, where "new" sources are created by saving water
- innovative storage systems, such as aquifer–storage–recovery (ASR) systems

- conjunctive use, where water from different sources, such as surface and groundwater, are managed jointly
- reuse, where wastewater is treated and used again in one form or another
- point-of-use treatment systems
- bottled water

ASSESSMENT OF ECOLOGICAL EFFECTS IN WATER-LIMITED ENVIRONMENTS

Randall T. Ryti
James T. Markweise
Neptune and Company, Inc.
Los Alamos, New Mexico

We consider that water-limited environments include arid lands that have less than 250 millimeters (10 inches) of annual precipitation and semiarid lands that have annual precipitation between 250 and 500 millimeters (1). This paper discusses the characteristics of arid and semiarid environments, approaches for evaluating the potential ecological impacts from anthropogenic contaminant sources, and the issues that make such assessments complex in arid and semiarid environments.

Precipitation defines water-limited environments, and plants and animals have developed specialized adaptations to the paucity/lack of water. For example, avian adaptation to desert environments is characterized by decreased metabolic rates and minimized water loss (2). Organisms have adapted to more efficient use of limited water quantities by behaviorally and/or physiologically optimizing their ability to acquire and retain water. For example, some beetles meet all their water needs by drinking the dew that collects on their bodies. Another strategy for surviving in dry environments includes adaptations to exploit periodic monsoons that are characteristic of arid and semiarid regions and involves rapid population growth to exploit periods of abundant resources (3).

Precipitation can be large, but it is infrequent. The spatial scale for precipitation also varies from thunderstorms that may impact a portion of a watershed to larger storm systems that impact an entire region. The predictability of storms varies widely between seasons or years for water-limited environments. Terms like intermittent or interrupted have been used to describe flow in water-limited environments. The pattern and intensity of flow change as a result of prolonged periods of wet or dry conditions in these environments.

Effluent and surface water runoff from residential, commercial, or agricultural activities are frequently more predictable sources of water in dry environments. Effluent/runoff is typically laden with metals and organic chemicals, and, it has been demonstrated, affects the composition and abundance of aquatic organisms in receiving streams (4). Wastewater effluent from anthropogenic sources may be inserted as a point source (i.e., an outfall pipe), or runoff may originate from nonpoint sources (e.g., distributed spatially from agricultural application). These types of contaminant sources frequently overlap and confound the interpretation and mitigation of ecological effects.

Directly related to effluent/runoff is the source of water used by people in water-limited environments. Two common water sources are impoundments along major river systems or pumping of groundwater. When dams are introduced, flow regimes in these water-limited environments change and lead to dramatic changes in the species in areas of modified flow regimes (5).

In addition to anthropogenic influences on water availability and water quality, the physical environment of arid/semiarid land potentially affects water resources. In the arid southwest United States, closed hydrologic basins exist; salt flats are one resultant landscape feature of closed basins. Surface depressions may also accumulate water during intense precipitation, and evaporation concentrates chemicals in these ponds during dry periods. This is in contrast to moist environments, where water transport of contaminants generally leads to diluted concentrations. Wind-driven contaminant transport is also more important in dry environments where the soil and rock lie exposed to wind erosion because of limited plant coverage.

Dry environments are also more susceptible to disturbance due to several factors. One is decreased resilience to stressors because of the extreme nature of the environment. Aridity in itself can be a stressor for plants and animals, and, it has been shown, influences species composition (6) and population characteristics (7). Adding anthropogenic stressors to this environment (e.g., through pollution) may decrease a population's viability if it is already near the individual tolerance limits for surviving in the dry environment. For example, Sjursen and others (8) showed that exposure to organic chemicals could reduce drought tolerance in soil invertebrates. In addition, populations depending on rainfall for reproduction could be locally exterminated if their source of water (e.g., an ephemeral rain pool) were impacted with chemical toxicants. Soils in arid environments also develop more slowly than in moist areas, recover from stress slowly (9), and are more susceptible to erosion after physical disturbance.

Adverse effects contributing to impacts can be defined as "Changes that alter valued structural or functional attributes of ecological entities ..." (10). Effects can be assessed for ecological entities on any scale of biological organization, including genetic, individual, population, community, and ecosystem levels. "An evaluation of adversity may include a consideration of the type, intensity, and scale of the effect as well as the potential for recovery" (10). These evaluations are typically categorized as either stressor-based or effects-based assessments. Stressor-based assessments are initiated to evaluate physical, chemical, or biological entities that can induce an adverse response (11). Effects-based assessments are initiated to determine the cause(s), once an adverse impact is observed.

Tools to evaluate and manage environmental effects come from a number of disciplines. Resource management fields (e.g., forestry, wildlife management) develop plans to manage or mitigate effects of logging, hunting, and fishing. Conservation biology develops plans to mitigate more general human impacts on the environment. For regulatory agencies, ecological risk assessment has been the tool used to evaluate ecological effects of planned or historic actions (10).

The U.S. Environmental Protection Agency (EPA) Office of Solid Waste and Emergency Response (a.k.a. Superfund) started developing guidance on ecological risk assessment in the late 1980s. This process culminated in the Risk Assessment Forum process for ecological risk assessment (12). Parallel to these efforts was the Ecological Risk Assessment Guidance for Superfund or ERAGS (11).

ERAGS starts by identifying the ecological entity that needs to be protected (the assessment endpoint). By defining an adverse effect, the measures used to evaluate effects will be identified. ERAGS encourages using multiple lines of evidence, including field information, laboratory studies, and literature studies for each measure. Effects are characterized through lines of evidence (13,14), and the weight of the evidence dictates how one manages adverse effects.

The complications from assessments in water-limited environments are often related to the spatial and temporal scales of the assessment. For example, one may be concerned about potential adverse ecological impacts on aquatic and terrestrial receptors of effluent releases to a stream from an industrial facility. One set of considerations is the physical environment and other anthropogenic impacts on the stream. As discussed before, flow in the stream may be expected to change dramatically as seasonal and annual changes in precipitation occur. Altered flow regimes and other anthropogenic impacts on this stream are also expected. These considerations may help select appropriate assessment endpoints.

The abundance of bats in roosts near a stream near an industrial facility could be selected as an assessment endpoint. However, bats have large foraging ranges and are likely to be impacted by contaminants from multiple sources. In addition, bats are migratory and could be exposed to contaminants from distant locations. Another assessment end point for this example could be abundance of benthic macroinvertebrates. This choice has the advantage of greater site fidelity and thus minimizes confounding effects from neighboring anthropogenic releases. Benthic macroinvertebrates also lend themselves to direct experimentation as a line of evidence without some of the ethical and logistical problems of vertebrates.

Another complication of adverse effects assessments in water-limited environments is the availability of appropriate toxicity bioassays for taxa commonly found in such environments (15). This problem is more obvious for terrestrial receptors. For example, earthworms are the most commonly used soil animals in soil toxicity bioassays. However, earthworms are not representative of the more important detritivores in arid/semiarid environments (mites). This is an area of active research for more biologically relevant toxicity testing organisms for water-limited environments.

Last, variation in timing and amount of precipitation complicates some empirical evaluation of ecological effects. For example, the phenology of most organisms in arid and semiarid environments is tied to precipitation, and most studies of adverse effects are tied to government funding cycles. This is why some empirical studies of biota abundance and diversity are dismissed for practical reasons and laboratory studies of toxicity are emphasized. In some cases, a field experiment where precipitation is artificially simulated is another option, if time represents a logistical constraint.

Acknowledgments

We thank two peer reviewers, R. Mirenda and M. Tardiff, for their constructive comments on an earlier draft of this paper.

BIBLIOGRAPHY

1. Walker, A.S. (1997). *Deserts: Geology and Resources*. Online Edition. United States Geological Survey, Washington, DC.
2. Tieleman, B.I. and Williams, J.B. (2000). The adjustment of avian metabolic rates and water fluxes to desert environments. *Physiol. Biochem. Zool.* **73**: 461–479.
3. van der Valk, H. (1997). Community structure and dynamics in desert ecosystems: potential implications for insecticide risk assessment. *Arch. Environ. Contam. Toxicol.* **32**: 11–21.
4. Ganasan, V. and Hughes, R.M. (1998). Application of an index of biological integrity (IBI) to fish assemblages of the rivers Khan and Kshipra (Madhya Pradesh), India. *Freshwater Biol.* **40**: 367–383.
5. Saunders, D.L., Meeuwig, J.J., and Vincent, A.C.J. (2002). Freshwater protected areas: strategies for conservation. *Conserv. Biol.* **16**: 30–41.
6. Tsialtas, J.T., Handley, L.L., Kassioumi, M.T., Veresoglou, D.S., and Gagianas, A.A. (2001). Interspecific variation in potential water-use efficiency and its relation to plant species abundance in a water-limited grassland. *Functional Ecol.* **15**: 605–614.
7. Spinks, A.C., Bennett, N.C., and Jarvis, J.U.M. (2000). A comparison of the ecology of two populations of the common mole-rat, *Cryptomys hottentotus hottentotus*: the effect of aridity on food, foraging and body mass. *Oecologia* **125**: 341–349.
8. Sjursen, H., Sverdrup, L.E., and Krogh, P.H. (2001). Effects of polycyclic aromatic compounds on the drought tolerance of *Folsomia fimerata* (Collembola, Isotomidae). *Environ. Toxicol. Chem.* **20**: 2899–2902.
9. Belnap, J. (1993). Recovery rates of cryptobiotic crusts: inoculant use and assessment methods. *Great Basin Naturalist* **53**: 89–95.
10. Committee on Environment and Natural Resources. (1999). *Ecological Risk in the Federal Government*. National Science and Technology Council, Washington, DC, CENR/5-99/001.
11. United States Environmental Protection Agency. (1997). *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments*. U.S. Environmental Protection Agency Interim Final Report. Emergency Response Team, Edison, NJ.
12. United States Environmental Protection Agency. (1998). *Guidelines for Ecological Risk Assessment*. Risk Assessment Forum, EPA/630/R-95/002F, Final. U.S. Environmental Protection Agency, Washington, DC.

13. Menzie, C. et al. (1996). Report of the Massachusetts weight-of-evidence workgroup: a weight-of-evidence approach for evaluating ecological risks. *Hum. Ecol. Risk Assessment* **2**: 277–304.
14. United States Environmental Protection Agency. (2000). *Stressor Identification Guidance Document*. U.S. Environmental Protection Agency, Office of Water and Office of Research and Development, Washington, DC.
15. Markwiese, J.T., Ryti, R.T., Hooten, M.M., Michael, D.I., and Hlohowskyj, I. (2001). Toxicity bioassays for ecological risk assessment in arid and semiarid ecosystems. *Rev. Environ. Contam. Toxicol.* **168**: 43–98.

REACHING OUT: PUBLIC EDUCATION AND COMMUNITY INVOLVEMENT IN GROUNDWATER PROTECTION

Jennifer Nelson
The Groundwater Foundation
Lincoln, Nebraska

INTRODUCTION

Throughout its existence, The Groundwater Foundation (TGF) has identified public education and community involvement as vital ingredients to successful groundwater protection activities. TGF is an international nonprofit organization based in Lincoln, Nebraska, with a mission to educate and motivate the public to care about and for groundwater. Groundwater provides drinking water to nearly half of the U.S. population (1) and is critical for other economic uses such as irrigation, agriculture, and industry. Unfortunately, because it is a hidden resource, the general public knows comparatively little about groundwater or how to protect it. Providing education to stakeholders about the value of groundwater is an excellent way to illustrate how individual and group decisions and activities can have a direct impact on the quality and quantity of local groundwater resources.

TGF was founded in 1985 and since has held that people could learn about groundwater in ways that were both scientifically accurate and user friendly. TGF works to make groundwater science accessible to and understandable by people everywhere, allowing these educated people to become active partners in protecting the environmental and economic vitality of their communities by protecting local and regional groundwater resources.

Groundwater Guardian is a program of TGF that provides a framework for community action and groundwater protection by providing recognition, support, and lessons learned. Over 300 communities from the United States and Canada have participated in Groundwater Guardian since the program began in 1994. Groundwater Guardian relies on voluntary efforts spearheaded at the local level to effectively address local groundwater-related concerns and issues through the creation of a diverse local team, which focuses the interests and resources of a community on the importance of long-term groundwater protection.

Groundwater Guardian serves as both an organizing and planning tool for communities interested in groundwater protection. A community within the program is broadly defined—it can be a city, county, watershed area, school, or other area where there are people committed to learning about and protecting groundwater. To enter the program, communities first form a team of diverse stakeholders, including citizens, local government, educators, business, and/or agriculture. The team is the heart of the program and is designed to be broad-based to draw on the different resources and expertise of various sectors of the community. These stakeholders who represent key local groups are in an excellent position to identify a community's groundwater problems, develop an education/action plan to address these issues, and document their progress and success.

After a team is formed, local issues of concern are identified and then addressed through activities addressing public education and awareness, pollution prevention, public policy, conservation, and best management practices. Teams submit plans for these activities in the beginning of the year and report back to TGF at the end of the year on progress made. Groundwater Guardian designation is awarded based on the information provided in these plans and reports.

USING EDUCATION AS A STARTING POINT FOR GROUNDWATER PROTECTION

Groundwater Guardian teams generally start with some sort of public education campaign to build awareness for future protection activities. This education may take the form of a youth festival, newspaper column, water bill insert, public service announcement, website, brochure, seminar, public meeting, or other means. Education is a popular starting point for most Groundwater Guardian teams because:

- Education is viewed as an accessible, common activity that invites broad community participation.
- Goals and objectives can easily be set and achieved and give community efforts valuable success to build upon.
- Awareness and understanding are the basis for future activity, for adults and children alike.
- Children are likely to communicate with and involve parents, relatives, neighbors, and friends with successful and memorable education experiences (2).

Public education allows a progressionary approach to groundwater resources protection. It first engages community residents to know that their individual actions and practices, like disposing of motor oil properly, using only appropriate amounts of fertilizer, and taking shorter showers, can have an impact on local groundwater quality and quantity. It then also builds a base of knowledge for future support of protection activities. Residents who know that their drinking water is supplied by groundwater, pumped from wells in a specific area, are more likely to support efforts such as wellhead protection and the implementation of best management practices that protect that source than a resident who simply thinks that drinking water comes from the tap. Education creates a foundation for action.

VALUE OF COMMUNITY INVOLVEMENT IN GROUNDWATER PROTECTION

In TGF's experiences, community involvement is necessary to successful, meaningful, and long-lasting groundwater protection, which provides numerous benefits to all community residents, including:

- *Protecting Natural Resources*. Groundwater is sometimes referred to as the environmental "bottom line" as an indicator of how successful all environmental protection activities are. By monitoring groundwater quality and quantity, communities can see the impacts of their groundwater and other natural resource protection activities. By protecting groundwater resources, communities are often also protecting a variety of other natural resources, such as wildlife habitat and wetlands.
- *Safeguarding Public Health*. Safe groundwater is directly related to good health. Contaminated drinking water can cause illness, disease, or even death. By protecting groundwater supplies that are used for drinking water, communities are protecting their own health.
- *Developing Economic Vitality*. Communities depend on local groundwater supplies not only as a source of drinking water but to provide businesses and agricultural producers with an input to create a market good. Without clean, safe and plentiful groundwater supplies, communities across the country would face serious challenges to their long-term sustainability.
- *Building Community Capacity*. Communities that make the commitment to groundwater are the very same communities who are working to be innovative and strive for long-term growth and sustainability of their community. Because of the variety of effective groundwater protection activities and options available, communities working to protect groundwater are also building their capacity to address other community issues, such as growth, strategic planning, solid waste, transportation, and economics.
- *Sustaining Supplies*. As populations continue to grow, increased demand is placed on groundwater supplies. Communities that are educated about the value of local groundwater resources are much more likely to manage those resources in a sustainable way.
- *Sharing Responsibility*. Whether it is for drinking or as an economic input, all those within a community use and benefit from groundwater. Consequently, protecting groundwater supplies is best done by a diverse team of community representatives who share a variety of viewpoints but are working toward a common goal. By sharing the responsibility for groundwater protection, strong communities will protect groundwater supplies for generations to come.

STRATEGIES FOR PUBLIC EDUCATION AND COMMUNITY INVOLVEMENT IN GROUNDWATER PROTECTION

There are numerous options for public education and community involvement for local groundwater protection. Protection efforts implemented by Groundwater Guardian communities often begin with public education activities, including:

- Organizing a community or school water festival
- Launching a public awareness campaign
- Distributing educational materials about local groundwater issues
- Holding public meetings
- Planning field trips

Communities often find success with these relatively easy to implement, yet high impact, public education activities and can then move on to more challenging efforts, such as:

Conservation

Endorse and encourage the use of water-saving devices
  Encourage sustainable lawn care and gardening practices
  Work with the local water department to foster conservation awareness

Pollution Prevention

Inventory local pollution sources
  Properly close abandoned wells
  Recognize local pollution sources and identify possible solutions

Public Policy

Establish and manage wellhead protection areas
  Support compliance with the Safe Drinking Water Act
  Encourage federal, state, and local interagency coordination
  Develop a community comprehensive plan

Best Management Practices (BMPs)

Encourage the adoption of BMPs in rural, urban, and commercial areas
  Adopt land use protection measures
  Provide voluntary management options

PUBLIC EDUCATION + COMMUNITY INVOLVEMENT = GROUNDWATER PROTECTION

Groundwater protection in small, groundwater-dependent communities is reliant on a high level of citizen involvement to be successful. These communities may lack the financial resources and/or professional expertise to ensure the protection of groundwater resources that are vital to the human and economic health of the community. Instead, they rely on an educated and empowered citizenry acting responsibly on behalf of the community. When citizens understand the fundamental importance of groundwater to their environmental and economic future, they are motivated and able to develop innovative and cost-effective strategies for its protection.

TGF has found that no matter how diverse communities are, those that are successful begin the groundwater protection process in generally the same way: by involving local stakeholders. The involvement of a diverse group of stakeholders is the most important part of the Groundwater Guardian program because a credible, broad-based team effectively represents community interests and therefore shares the responsibility for groundwater protection among community groups.

As is the case for any issue that depends on community involvement for its success, groundwater protection needs a long-term commitment that is continuously rejuvenated by creativity, new ideas, and leadership. The same ingredients that contribute to sustaining meaningful groundwater education and protection activities also apply to other beneficial community activities. By starting with public education and stressing diverse community involvement, communities are on the path to sustainable groundwater protection. As TGF has learned, “it is because of people that groundwater must be protected, but it is only through people that we can do so.”

BIBLIOGRAPHY

1. Alley, W.M., Reilly, T.E., and Franke, O.L. (1999). *Sustainability of Ground-Water Resources*. U.S. Geological Survey Circular 1186.
2. Herpel, R., Jackson, J., Kuzelka, R., Rodenberg, S., and Seacrest, S. (1998). *Let's Make a Difference: Mobilizing for Community Action*. The Groundwater Foundation, Lincoln, NE.

INTEGRATION OF ENVIRONMENTAL IMPACTS INTO WATER RESOURCES PLANNING

Mehmet Ali Yurdusev
Celal Bayar University
Manisa, Turkey

Nature is so complicated that we need to idealize or simplify to understand, interpret, and benefit from it. Simplifications are carried out by modeling natural events. Models give us the opportunity to input values representing the observed values of natural phenomena to see what the possible outcome would be. Water is one of the few natural substances without which human beings cannot survive. Water use, therefore, is much more complicated because it has to reconcile an irregular or random natural phenomenon and humans' attitudes toward it, which is not less complicated than water in nature itself. Therefore, it has to be managed by employing complicated computer-based mathematical models. The spatial and temporal dimensions of this complicated issue can then be handled. These modeling studies are generally referred to as water resources planning (WRP). WRP can be defined as matching future demands to potential resources, satisfying some preset objectives such as cost-effectiveness or environmental quality. The end product of water resources planning is a development plan for some future period, normally 20 to 30 years. As such, WRP mainly encompasses three types of activities:

- assessment of available water resources,
- assessment of future water requirements, and
- matching between available resources and forecast demands.

In the third phase of water resources planning, what is primarily involved is identifying the most appropriate development strategy for meeting projected demands at minimum cost, over the planning horizon. On a national scale, the planning process is characterized by the need to screen a large number of potential options (1) to determine

- which resources should be developed,
- the timing and order of that development, and
- the areas of demand to which each new resource should be assigned.

Typically, this has required information relating to future demand, the performance of different size sources at various locations, the possible links between sources and demand centers, together with all associated construction and operating costs. Faced with an enormous choice of possibilities for matching resources to needs, a systematic search procedure is normally required to achieve an objective assessment. The objective function of the procedure would be minimization of total discounted costs. To that end, various forms of mathematical programming have been used (2) to identify the minimum net present value of a development program. Alternatively, heuristic search techniques can be used which, although not as rigorous, provide a practical solution (3).

In the past, scant attention would have been given to environmental considerations within the water resources planning process until a scheme had already been selected. Then, detailed environmental assessment would be undertaken. At that stage, some attempts would be made to identify and ameliorate damaging environmental impacts. In contrast, the environmental considerations can be incorporated from the outset and included in selecting the scheme to be promoted. This is achieved by weighting the costs of the various options (both construction and operating) to reflect their environmental impacts, before including them in the economic evaluation. The objective function of the combined methodology is to minimize the total monetary and environmental discounted cost of meeting the projected demands. Quantification of such considerations is achieved by running an appropriate model with and without the environmental factors. Such a model should integrate an environmental impact assessment (EIA) and economic planning (EP) models. The EIA model develops what might be called environmental impact factors (EIF) to weigh the costs of schemes, and the EP model is to undertake the scheme selection and timing process using environmentally adjusted costs. The EIF is a weighting function whose role is to raise the cost of a water resource option if its environmental impact assessment indicates that it is environmentally damaging and to decrease it otherwise. Using such a factor in an economic planning model has the effect of influencing the solution toward selecting environmentally friendly schemes at the expense of environmentally damaging ones. In mathematical terms, an EIF is

$$0 < \text{EIF} < 1 \text{ and } 1 < \text{EIF} < 2$$

where 0 is the best possible outcome and 2 is the worst. Using such factors, environmental gain can be expressed by the former and the environmental loss by the latter: a value of 1 indicates neutrality.

A multicriterion decision-making technique referred to as composite programming (CP) can be used to derive the EIFs. The output from CP is a measure of the composite distance resulting from the aggregation of a series of basic indicators (4). For a given L, the composite distance from a so-called ideal point, where 0 represents the best possibility and 1 represents the worst possibility, the main features of CP are (1) $0 < L < 1$ and (2) the larger L, the worse the associated scheme. The real values of basic indicators are weighed to obtain a further indicator value. The procedure continues until it reaches a final indicator value that will represent the performance of the option according to the indicators considered. A simple mathematical function which deliberately exaggerates the extreme values can be used to transform L into EIF as follows:

$$\mathrm{EIF} = 1.4907\sqrt{L} \qquad \text{if } 0 < L < 0.45 \tag{1}$$

$$\mathrm{EIF} = 2 - 1.4907\sqrt{1-L} \qquad \text{if } 0.45 < L < 1 \tag{2}$$

In this way, it is possible to convert a series of basic indicators covering both detrimental and beneficial impacts to a corresponding EIF within the range of 0 to 2. When evaluating an impact, assessment values range from negative significant to positive significant with a central point at neutral. The impacts of water projects can be grouped according to their resource use, quality, ecosystem, and social implications. For example, the aggregation structure given in Table 1 can be considered as assessing the overall construction impact of a groundwater development scheme. The assessment values in Table 1 are assigned within a range between 1 and 8; 1 is the worst situation, 8 is the best, and 4 is neutral. The values in parenthesis show the normalized ones.

The values of second level indicators are calculated from the weighted sum of the associated basic impacts. The value of resource implications is directly taken as neutral because there are no associated basic impacts from which its value can be obtained. The overall value is, then, calculated as 0.41, and the EIF as 95% from Equation 1. This value indicates that the project is somewhat environmentally friendly and, therefore, the associated cost will be reduced accordingly during the optimization process.

Clearly, different types of resources/links will affect the environment in different ways, and therefore, each has its own specific set of considerations. Similarly, there are different considerations for each type of resource/link during the construction and operational phases. All these different impacts can be accommodated within the EIA model (5).

The economic planning model is intended to undertake the scheduling process to determine when the resources are to be built. Therefore, an existing model should be adopted for this purpose. Such a model would perform the planning process using the costs of water schemes contained in the case study under consideration. The objective is to find the least-cost solution for the whole exercise. As previously mentioned, the idea is to use the outputs of an EIA model (EIFs) in the economic planning model selected to reflect the environmental performance of each scheme. The two-way usage of an economic planning model would give the user the opportunity to compare results with and without environmental impacts. This may even be used to quantify the cost of environmental impacts in the form of additional investment required to apply environmentally friendly solutions (with environmental impact solutions). There are several economic water resources planning models that can be used for this purpose. The models can be based on unit costs (6), integer linear programming (2,7), heuristic programming (8), and genetic algorithms (9). The selection of an economic planning model depends on the availability of models, the amount to be invested, the nature of the exercise, and the capability of the models. Unfortunately, most of the models mentioned may not be commercially available. However, the owners may lend the models on request.

Combining economic planning and EIA models produces an overall environmentally-influenced economic planning system in which EIFs are used to weigh the real construction and operating costs for various development options to reflect their positive or negative environmental

Table 1. Constructional Environmental Impact Assessment for a Groundwater Scheme

Basic Impacts	Assessment Values	Weights	Second-Level Indicators	Indicator Values	Weights
			Resource use implications	0.50	0.18
Lack of visual intrusion	6(0.18)	1.00	Quality implications for physical system	0.18	0.29
Effect of construction on surface ecosystem	4(0.50)	0.44			
Pipeline laying impact on vegetation and wildlife	3(0.66)	0.56	Ecosystem implications	0.59	0.25
Impact of infrastructure construction; noise, mud, etc.	4(0.50)	0.27			
Landscape/land use impact of pipeline/structures	4(0.50)	0.27	Social implications	0.35	0.29
Implementation benefits	6(0.18)	0.45			

impacts. The coupling can be achieved simply by an economic model using the cost figures modified by the EIFs produced by the EIA model.

Such a methodology was applied to the whole of England and Wales using the National Rivers Authority's (NRA) data (3). The exercise was undertaken using the same economic planning model used by NRA coupled with an independent code for the EIA model (8). The environmentally sensitive solution was repriced in real terms by the model and therefore, the additional cost of preferring such a plan was estimated (5).

BIBLIOGRAPHY

1. Jamieson, D.G. (1986). An integrated multi-functional approach to water-resources management. *J. Hydrol. Sci.* **31**(3): 14.
2. O'Neill, P.G. (1972). A mathematical programming model for planning a regional water-resource system. *J. Inst. Water Engs.* **26**: 47–61.
3. National Rivers Authority. (1994). *Water-Nature's Precious Resource: An Environmentally Sustainable Water Resources Development Strategy for England and Wales*. HMSO, London, England.
4. United Nations Educational Social and Cultural Organisation (UNESCO). (1987). *Methodological Guidelines for the Integrated Environmental Evaluation of Water Resources Development*. United Nations Environmental Programme(UNEP), Paris, France.
5. Yurdusev, M.A. (2002). *Environmental Impacts for Water Resources Planning*. WIT Press, Southampton, UK.
6. Armstrong, R.B. and Clarke, K.F. (1972). Water resource planning in South East England. *J. Inst. Water Eng.* **26**: 11–46.
7. Schrage, L. and Cunningham, K. (1991). *LINGO Optimisation Modelling Language*. Lindo Systems Inc., Chicago.
8. Anglian Water Services. (1993). *Economic Planning Model of Water Resources: RESPLAN User Guide*. Anglian Water Services Ltd., Cambridge, England.
9. Dandy, G. and Connarty, M. (1995). Use of generic algorithms for project sequencing. In: *Integrated Water Resources Planning for the 21st Century*. M.F. Domenica (Ed.). Proc. 22nd Annu. Conf., ASCE, Cambridge, Massachusetts, USA, pp. 540–543, May 7–11.

THE EXPANSION OF FEDERAL WATER PROJECTS

Martin Reuss
Office of History Headquarters
U.S. Army Corps of Engineers

Instead of national planning, Congress settled on a piecemeal approach to public works development in the early nineteenth century and, with rare exceptions, has clung to this approach ever since. In May, 1824, President Monroe signed legislation appropriating $75,000 to improve navigation on the Ohio and Mississippi rivers—major routes to the western part of the country. The act empowered him to employ "any of the engineers in the public service which he may deem proper" and to purchase the "requisite water craft, machinery, implements, and force" to eliminate various obstructions. Along with the General Survey Act, signed a month before, the Mississippi-Ohio rivers legislation initiated the permanent involvement of the Army Corps of Engineers in rivers and harbors work. However, each act focused on one activity: the General Survey Act on planning; the other on construction. Two years later, Congress combined surveys and projects in one act, thus establishing a pattern that lasts until the present. The 1826 act, therefore, can be called the first true rivers and harbors legislation.

By the time of the Civil War, the federal contribution to river, harbor, and canal improvements amounted to about $17 million in appropriated monies. Some 4.6 million acres of public lands were given for canal improvements and another 1.7 million acres for river improvements. Land grants under the 1849 and 1850 Swamp Land acts and the 1841 land grant act totaled some 73 million acres. While these grants and appropriations were significant, they represented a modest amount of aid compared with state and private sector contributions, which by 1860 totaled well over $185 million for canals alone. Corporations and public agencies spent many millions more on the construction of urban water systems.

Many of the canal companies incorporated by the states ran into trouble. The 1837 depression had driven a number into bankruptcy; others survived, but only with a healthy influx of state money, state guaranteed bonds, and occasional federal and state land grants. Often, too, the national story was repeated at the state level, with rationally planned canal routes sacrificed to local political pressures to extend canals to uneconomical out-of-the-way villages. The one major exception to this sad story was the Erie Canal, whose success had spurred the canal boom that increasingly appeared more like a dismal bust, especially with new competition from the railroads. From Pennsylvania to Ohio to Indiana to Illinois and on into the states of the old Northwest, canal fever turned to canal panic, and the public lost faith in both the companies and the politicians who had supported the enterprises.

The American Civil War (1861–1865) also affected the development of water projects. Military action and wartime budgetary constraints took their toll on many of the nation's ports and navigable waterways, and after the war commercial development accelerated demands for waterway improvements. A business-oriented Republican Congress responded by authorizing a great deal more money for rivers and harbors. The federal government also took over many of the bankrupt canal companies, and the Corps of Engineers became the custodian of former private or state waterways. This, as one author put it, was the "Golden Age of the Pork Barrel". Between 1866 and 1882, the Presidents signed 16 rivers and harbors acts. The 1866 act appropriated $3.7 million for 49 projects and has been described as the first omnibus bill, so called because like a horse-drawn omnibus of the time,

---

the legislation provided room for a great many people boosting various projects. Sixteen years later, though, the 1882 act appropriated five times more money. By that year, the federal government had spent over $111 million on rivers and harbors projects. "Willingness to pay"—the primary test of project implementation before the Civil War—now included unprecedented federal largess. In the so-called "Gilded Age," lack of federal or non-federal funds was about the only thing that prevented construction.

By the 1880s, the basic working relationship between Congress and the Army Corps of Engineers was set. Congress directed the Corps to survey potential projects, make recommendations, and provide cost estimates. Rivers and Harbors acts funded both the surveys and the projects that Congress chose to authorize. Also in the early 1880s, Congress mandated that the Corps of Engineers use more contractors and less hired labor. By the end of the century, contractors did nine-tenths of all waterways construction, and no Corps officer could use hired labor without the express authority of the Chief of Engineers. Increasingly, then, the Corps became a funding conduit to the private sector. This pattern did not stop private sector engineers from calling for the complete elimination of the Corps from public works, but Congress rejected all bills that leaned in that direction.

Fear of railroad competition and questions about federal aid to projects of apparently local benefit moved the Senate in 1872 to create a Select Committee on Transportation Routes to the Seaboard. Composed of nine senators, the committee was headed by Senator William Windom of Minnesota and known popularly as the Windom Committee. Its 1873 report promoted waterway over railway transportation wherever waterways were properly located. Of more relevance here is the committee's conclusion (on a five to four vote) that the sum of local rivers and harbors projects contributed to the national interest. Generally accepted by Congress, this conclusion justified federal contributions for waterway improvements. The result was the authorization of dozens of dubious projects. By 1907, the cumulative total for rivers and harbors appropriations was more than four times the 1882 figure; the federal role in navigation improvements continued to grow.

Meanwhile, the issue of constitutional authority had somewhat changed focus. In 1870, the Supreme Court ruled in The Daniel Ball case that the common law doctrine that navigability depended on tidal influence, a doctrine accepted in British courts, did not fit the American situation. However, the definition the Court substituted was extraordinary. The test of navigation was to be the river's "navigable capacity." That meant, the Court went on:

> Those waters must be regarded as public navigable rivers in law which are navigable in fact. And they are navigable in fact when they are used, or are susceptible of being used, in their ordinary condition, as highways for commerce over which trade and travel are or may be conducted in the customary modes of trade and travel on water.

In short, American rivers were navigable if they were, are, or could be navigable. This decision, in combination with the earlier 1824 Gibbons v. Ogden Court ruling, made the federal government the clear guardian and ultimate decision-maker on tens of thousands of miles of waterways in the United States. In practice it sufficed to show that a stream had the capacity to float logs to declare it navigable. However, with this issue more or less settled, another appeared: flood control.

Rivers always flood, but the floods do not always damage life and property. In the United States, we can trace floods as far back as 1543, when Mississippi River floods stopped Hernando De Soto's expedition. Naturally, as settlers moved into the floodplains and built villages, then cities, the damages increased. By the mid-nineteenth century, the problem was becoming critical along the lower Mississippi. Most people put their faith in technology to protect them. Indeed, the then popular term "flood prevention" testifies to an extraordinarily unrealistic idea when one thinks about it a bit. In the twentieth century, the term became "flood control," a somewhat more modest formulation. Nowadays we speak of "flood damage reduction," which probably comes closest to the mark. In any case, in the 1870s calls came for repairing and raising the levees on the Mississippi River. In 1879, Congress created a joint military—civilian Mississippi River Commission to develop and implement plans to improve navigation and flood control on the lower Mississippi. However, once again some Congressmen raised constitutional objections, expressing doubts that flood control was an appropriate federal activity. Until 1890, no appropriation could be used for repairing or constructing any levee in order to prevent damage to lands from overflow, or for any purpose other than deepening and improving the navigation channel. The 1890 floods along the lower Mississippi resulted in the removal of this restriction, which, in any event, had had little practical effect other than satisfying congressional scruples.

Floods in 1912, 1913, and 1916 along the Ohio and Mississippi rivers eventually led to passage of the 1917 Flood Control Act, the nation's first act dedicated solely to flood control. It provided funds on a cost-shared basis for levee construction along the lower Mississippi and another appropriation to improve the Sacramento River in California. While an important step towards federal involvement in flood control, it was comparatively modest compared to what followed in the coming decades, when flood control became intertwined with multipurpose development—the subject of the next essay.

This article is based on Martin Reuss, "The Development of American Water Resources, Planners, Politicians, and Constitutional Interpretation," in Paul Slack and Julie Trottier (Eds.), *Managing Water Resources, Past and Present: The Twelfth Annual Linacre Lectures*. Oxford University Press, 2004. Forthcoming. Used by permission of Oxford University Press.

The Development of American Water Resources: Planners, Politicians, and Constitutional Interpretation.

The Constitution and Early Attempts at Rational Water Planning.

The Expansion of Federal Water Projects.

FLOOD CONTROL HISTORY IN THE NETHERLANDS

RONGCHAO LI
Delft University of Technology
Delft, The Netherlands

The traditional flood control approach in The Netherlands is fighting against natural dynamics or, restricting the natural dynamics of a river system by canalization and embankments. However, such an approach results in destruction of river scenery, damage to nature and cultural values, as well as continuous dike raising and strengthening. Changing societal views in The Netherlands have aroused discussion of alternative flood management strategies. Around the beginning of the for twenty-first century, new visions appeared such as "living with floods" and the "room for the river" policy.

FLOOD CONTROL HISTORY IN THE NETHERLANDS

The Netherlands is a small country in Western Europe. The delta of the Rhine, Meuse, and Schelde Rivers lie within the country. One-third of The Netherlands needs and has artificial protection against floods from the sea or major rivers (see Fig. 1). Millions of people live, and large industries have settled in this area. In the western coastal area, the major cities such as Amsterdam, The Hague, and Rotterdam are situated. Large parts of the western coastal area lie below sea level. Also in the area along the rivers, cities and industries have settled. Therefore, flood risk management is an important issue for The Netherlands.

Figure 1. Almost 75% of The Netherlands is threatened by floods (1).

The Rhine and Meuse Rivers are the largest and most important rivers in The Netherlands. The Rhine River has a large catchment stretching from the high Alps in Switzerland, Germany, and France. The Meuse River flows through France and Belgium. Only the most downstream parts of these rivers are situated in the Netherlands. Therefore, flood risk management focuses on the discharge of water. Nowadays, after a long history of coping with floods, a high safety standard has been reached.

Before 1000 A.D., floods occurred very frequently. Mainly the higher parts in The Netherlands were inhabited, and in the lower parts, people lived on mounds, so no real disasters occurred in that period. Because people had no technological options to prevent flood then, they had no other option but to live with the floods and to avoid the floodwaters. They adapted their lives and land use to the rivers and sea.

About 1000 A.D., the inhabitants of The Netherlands started to build the first dikes around relatively small polders. The flood risk management strategy changed from living with floods to a strategy where the river and its floods were more controlled. This trend of increasing the control of the river continued, and by 1400 A.D., an almost completely closed dike system already existed along the rivers (2). Furthermore, huge changes in the course of the river and the riverbed in the form of regulation and canalization took place (3). The river was harnessed into a small area and now the people and land use did not adapt to the river any longer, but the river was adapted according to the wishes of the inhabitants. The main flood management strategy consisted of preventing floods. This strategy continued unchanged until 1953. After each dike breach, the broken dikes were reconstructed, raised, and improved resulting in increasingly higher and stronger dikes. The last great flood of the Rhine River occurred in 1926, after which the dikes were improved again (2). Since then, the rivers seemed controlled, and people feel safe.

In 1953, a major flood from the North Sea took place that killed more than 1800 people. This flood resulted in new safety regulations for the coast and also for the river area. Previously, dike heights were based on the maximal recorded water level, but after 1953, a more scientific base has been used. The optimal level of safety was defined as the accepted probability of flooding for the different areas in The Netherlands. To be able to use this new norm, it was simplified to the demand that dike levels should exceed water levels related to a discharge with a chosen return time. After several years of discussion, one safety level was chosen for the whole area threatened by river floods: a discharge with a probability of once in 1250 years (the design discharge). Flooding by causes other than overtopping of dikes and uncertainties in nature and in the calculations were considered by adding 0.5 m to the required height and some regulations for the design of the dikes. Every 5 years, the design discharge is calculated anew based on the recorded discharges.

Nowadays, the Dutch rivers are thus strongly human-influenced. Figure 2 shows a cross section of a typical

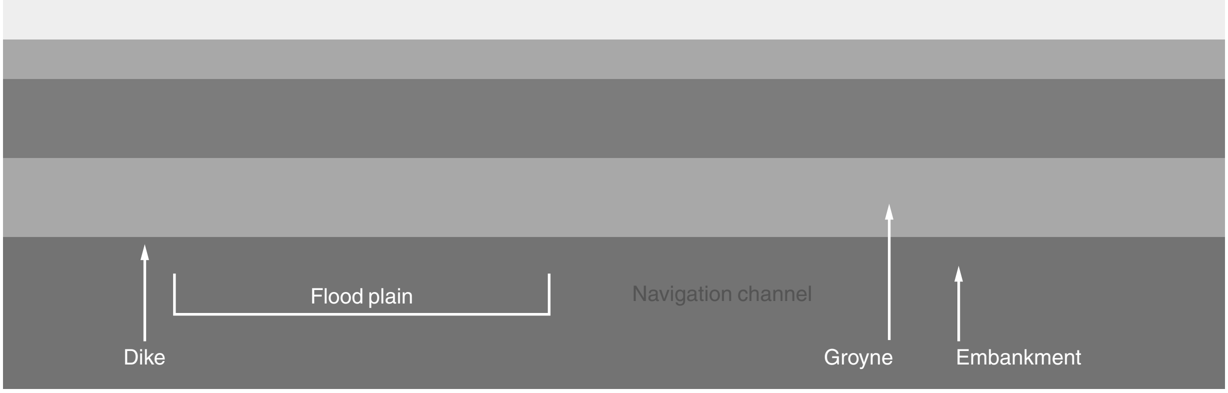

Figure 2. A cross section of a typical river in The Netherlands (4).

Dutch part of a river. The rivers consist of a small mainstream that has to enable navigation all year round. Next to this mainstream, are floodplains which carry water during peak flows, about twice each year in wintertime. These floodplains are used for agriculture and recreation in the summertime during lower discharges. The area outside these floodplains is protected from floods by huge embankments. The unprotected floodplains sometimes lie higher than the surrounding area due to sedimentation in the floodplains and subsidence of the surrounding area. To prevent the area along the rivers from becoming flooded at once, this area is subdivided into dike circles, which are areas surrounded by embankments or natural higher areas.

In 1993 and 1995, extremely high river discharges occurred in the Meuse and Rhine Rivers. Because of these extreme peak discharges, the design discharge with a probability of once in 1250 years increased from 15.000 m^3/s to 16000 m^3/s at the Dutch–German border. Using the traditional approach of flood risk management, this would lead to a further increase of dike heights. However, new solutions have developed nowadays creating room for the rivers (5).

NEW VISIONS AND IDEAS IN FLOOD RISK MANAGEMENT

Due to changes in the values of society and available technology in The Netherlands, flood risk management is changing. Some of these changes and the resulting ideas on flood risk management are mentioned here (1):

- The current strategy based on *one design discharge for the whole* area threatened by river floods leads to the same level of protection for different areas with different potential flood damage. Cities, agricultural areas, and nature reserves all have the same probability of flooding in The Netherlands. The question arises whether this is economically sensible. It also means that it is unknown which area will become flooded first when an extreme discharge occurs. Because all areas theoretically have the same probability of flooding, a large area has to be evacuated.
- Due to the fact that this strategy *focuses on preventing floods*, little attention has been paid to the consequences of possible floods. Relatively little attention is paid to plans for emergency situations, evacuation plans, flood mitigation measures, etc. which also might lower flood risks. New technology has increased the possibilities for good flood forecasts, of communication technology, and flood warnings. Also the possibility to anticipate peak flows by lowering a structure and using a detention pond has grown. These technologies make it easier to allow controlled floods and therefore reduce the need to prevent floods in all circumstances.
- By looking only at the probability of a design discharge, risks are not clearly visualized. The uncertainties in, for example, the design discharge, the translation of the discharge to water levels, the diversion of the peak flow through the different river branches, the strength of dikes and structures, and human behavior are not clearly considered. This results in a false sense of safety in the area and thus in rapid economic development and increasing potential flood damage.
- Another disadvantage of the strategy is that it includes an endless need for raising and improving the water defense structures. *Restricting* the natural dynamics of a river system by canalization and embankments requires continuous maintenance and improvements, otherwise the river dynamics will damage these works, and the river will try to return to its own natural behavior.
- The current strategy results in destruction of the river scenery and damages nature and cultural values. Recently, these values have become more important. A new strategy should consider these values and not focus only on reaching a “safe” situation by technical solutions.
- In these modern times, disasters are not accepted any longer. The government and water managers would be held responsible for a disaster, even if the event that caused the disaster is extreme and rare.

In The Netherlands, consciousness has increased that absolute safety cannot be guaranteed and that floods will always occur. However, disasters are not accepted any longer. Therefore, and because of the disadvantages of the current strategy, as mentioned above, new ways to deal with uncertainties and flood risks have to be found.

In 1990, the policy report "Room for Rivers" was published. This policy wants to create room for rivers to stop the continuous cycle of dike raising, increased feelings of safety, increased investments, increased flood risks, and needs for further dike raising. By moving the dikes away from the river, digging out the floodplains, removing obstacles from the floodplains and sometimes constructing bypasses for peak flows, room for the river is created. The purpose of this policy is to interrupt the continuous dike raising and lower the water levels in the river by giving the river more space. In reality, this policy has been applied at different locations. Numerous obstacles, such as roads to bridges, roads to ferries, and brick factories in the floodplains have been removed or have changed to permit water flow. At Nijmegen near the Dutch–German border, a dike has been moved to increase room for the river. The inhabitants of the area normally opposed these projects because the room for the river is taken from other functions. However, they are mostly also against dike raising because it deteriorates their scenery and sometimes it destroys houses and restaurants built on or near the dike.

Another new policy is the indication of so-called "emergency areas." These are areas that must be flooded first to protect other, more vulnerable areas. Officially these areas should be used only for discharges that exceed the design discharge, thus with a frequency of less than once in 1250 years. Official research has been published, but until now, no governmental decision has been made about the indication of such areas, the location of these areas, or the regulations for the inhabitants of these areas. It is unclear whether the villages in these areas will be protected by dikes and what damage compensation they will receive. The purpose of this policy is to limit the emergency measures to smaller areas and to know what area will become flooded and not leave the choice to the river.

Other new concepts are introduced in the National Policy on Water Management. According to this policy we should "aim at resilient and healthy water systems." Resilient water systems are systems that may be disturbed or show temporary change but will easily recover. To discuss resilience and resistance, a systems approach to flood risk management has to be adopted. The system on which flood risk management focuses can be defined as the whole area threatened by floods including the society, ecosystems, and the river itself in this area. Optimal flood risk management depends on the society and culture of the inhabitants, not only on the discharge regime of the river. Furthermore, the whole system should be studied, not just a small river stretch, because measures as well as floods downstream and upstream interact with each other. Measures may transport the flood problems downstream, and upstream floods may prevent downstream inundations. In resilience strategies for flood risk management, floods are not necessarily prevented, but the flood impacts are limited, and recovery after the flood is enhanced. Resilience strategies focus on the whole discharge regime, not on a certain design discharge or design event. These strategies are still only a subject of research and an option for the long term. For more information, see De Bruijn and Klijn (1) or Vis et al. (6).

CONCLUSIONS

One-third of The Netherlands needs and has artificial protection against floods from the sea or the major rivers. In this area, millions of people live and large industries have settled. This proves the importance of flood management for The Netherlands. After centuries of coping with floods, research and reconsideration of flood management strategies are still needed because society is changing and thus the optimal flood management strategy must also change.

BIBLIOGRAPHY

1. De Bruijn, K.M., Klijn, F. (2001). *Resilient Flood Risk Management Strategies*. Proc. IAHR Congress September 16–21, Beijing, Tsinghua University Press, Beijing, China, pp. 450–457.
2. Commissie Rivierdijken (1977). *Rapport commissie rivierdijken*. (In Dutch) Hoofddirectie van de Waterstraat, s-Gravenhage.
3. Janssen, J.P.F.M. and Jorissen, R.E. (1997). Flood management in The Netherlands; recent development and research needs. In: *Ribamod, River Basin Modelling, Management And Flood Mitigation Concerted Action*. R. Casale, K. Havnø, and P. Samuels (Eds.). pp. 89–104.
4. Silva, W., Klijn F., and Dijkman, J. (2000). *Ruimte Voor Rijntakken. What The Research Taught Us. Delft*, WL|Delft Hydraulics, The Netherlands.
5. Min.VROM En V&W (Ministeries VROM En V&W) (1997). *Beleidslijn Ruimte Voor De Rivier*. Den Haag (In Dutch).
6. Vis, M., Klijn, F., and Van Buuren, M. (Eds.). (2001). *Living with Floods, Resilience Strategies for Flood Risk Management and Multiple Land Use in the Lower Rhine River Basin*. Summary report. NCR, Delft, The Netherlands, de Bruijn, 2001.

FOOD AND WATER IN AN EMERGENCY

Federal Emergency
Management Agency

INTRODUCTION

If an earthquake, hurricane, winter storm or other disaster strikes your community, you might not have access to food, water and electricity for days, or even weeks. By taking some time now to store emergency food and water supplies, you can provide for your entire family. This brochure was developed by the Federal Emergency Management Agency in cooperation with the American Red Cross and the U.S. Department of Agriculture.

WATER

Having an ample supply of clean water is a top priority in an emergency. A normally active person needs to drink at least two quarts of water each day. Hot environments can double that amount. Children, nursing mothers and ill people will need even more.

You will also need water for food preparation and hygiene. Store a total of at least one gallon per person, per day. You should store at least a two - week supply of water for each member of your family.

If supplies run low, never ration water. Drink the amount you need today, and try to find more for tomorrow. You can minimize the amount of water your body needs by reducing activity and staying cool.

HOW TO STORE WATER

Store your water in thoroughly washed plastic, glass, fiberglass or enamel-lined metal containers. Never use a container that has held toxic substances.

Plastic containers, such as soft drink bottles, are best. You can also purchase food-grade plastic buckets or drums. Seal water containers tightly, label them and store in a cool, dark place. Rotate water every six months.

EMERGENCY OUTDOOR WATER SOURCES

If you need to find water outside your home, you can use these sources. Be sure to purify the water according to the instructions listed below before drinking it.

- Rainwater
- Streams, rivers and other moving bodies of water
- Ponds and lakes
- Natural springs

Avoid water with floating material, an odor or dark color. Use saltwater only if you distill it first. You should not drink floodwater.

THREE WAYS TO PURIFY WATER

In addition to having a bad odor and taste, contaminated water can contain microorganisms that cause diseases such as dysentery, typhoid and hepatitis. You should purify all water of uncertain purity before using it for drinking, food preparation or hygiene.

There are many ways to purify water. None is perfect. Often the best solution is a combination of methods. Two easy purification methods are outlined below. These measures will kill most microbes but will not remove other contaminants such as heavy metals, salts and most other chemicals. Before purifying, let any suspended particles settle to the bottom, or strain them through layers of paper towel or clean cloth.

1. *Boiling*
   Boiling is the safest method of purifying water. Bring water to a rolling boil for 1 minute, keeping in mind that some water will evaporate. Let the water cool before drinking.
   Boiled water will taste better if you put oxygen back into it by pouring the water back and forth between two clean containers. This will also improve the taste of stored water.
2. *Disinfection*
   You can use household liquid bleach to kill microorganisms. Use only regular household liquid bleach that contains 5.25 percent sodium hypochlorite. Do not use scented bleaches, color-safe bleaches or bleaches with added cleaners.
   Add 16 drops of bleach per gallon of water, stir and let stand for 30 minutes. If the water does not have a slight bleach odor, repeat the dosage and let stand another 15 minutes.
   The only agent used to purify water should be household liquid bleach. Other chemicals, such as iodine or water treatment products sold in camping or surplus stores that do not contain 5.25 percent sodium hypochlorite as the only active ingredient, are not recommended and should not be used.
   While the two methods described above will kill most microbes in water, distillation will remove microbes that resist these methods, and heavy metals, salts and most other chemicals.
3. *Distillation*
   Distillation involves boiling water and then collecting the vapor that condenses back to water. The condensed vapor will not include salt and other impurities. To distill, fill a pot halfway with water. Tie a cup to the handle on the pot's lid so that the cup will hang right-side-up when the lid is upside-down (make sure the cup is not dangling into the water) and boil the water for 20 minutes. The water that drips from the lid into the cup is distilled.

HIDDEN WATER SOURCES IN YOUR HOME

If a disaster catches you without a stored supply of clean water, you can use the water in your hot-water tank, pipes and ice cubes. As a last resort, you can use water in the reservoir tank of your toilet (not the bowl).

Do you know the location of your incoming water valve? You'll need to shut it off to stop contaminated water from entering your home if you hear reports of broken water or sewage lines.

To use the water in your pipes, let air into the plumbing by turning on the faucet in your house at the highest level. A small amount of water will trickle out. Then obtain water from the lowest faucet in the house.

To use the water in your hot-water tank, be sure the electricity or gas is off, and open the drain at the bottom of the tank. Start the water flowing by turning off the water intake valve and turning on a hot-water faucet. Do not turn on the gas or electricity when the tank is empty.

Food

Short—Term Supplies. Even though it is unlikely that an emergency would cut off your food supply for two weeks, you should prepare a supply that will last that long.

The easiest way to develop a two-week stockpile is to increase the amount of basic foods you normally keep on your shelves.

Storage Tips

- Keep food in a dry, cool spot-a dark area if possible.
- Keep food covered at all times.
- Open food boxes or cans carefully so that you can close them tightly after each use.
- Wrap cookies and crackers in plastic bags, and keep them in tight containers.
- Empty opened packages of sugar, dried fruits and nuts into screw-top jars or air-tight cans to protect them from pests.
- Inspect all food for signs of spoilage before use.
- Use foods before they go bad, and replace them with fresh supplies, dated with ink or marker. Place new items at the back of the storage area and older ones in front.

Nutrition Tips

During and right after a disaster, it will be vital that you maintain your strength. So remember:

- Eat at least one well-balanced meal each day.
- Drink enough liquid to enable your body to function properly (two quarts a day).
- Take in enough calories to enable you to do any necessary work.
- Include vitamin, mineral and protein supplements in your stockpile to assure adequate nutrition.

Food Supplies

When Food Supplies Are Low. If activity is reduced, healthy people can survive on half their usual food intake for an extended period and without any food for many days. Food, unlike water, may be rationed safely, except for children and pregnant women.

If your water supply is limited, try to avoid foods that are high in fat and protein, and don't stock salty foods, since they will make you thirsty. Try to eat salt-free crackers, whole grain cereals and canned foods with high liquid content.

You don't need to go out and buy unfamiliar foods to prepare an emergency food supply. You can use the canned foods, dry mixes and other staples on your cupboard shelves. In fact, familiar foods are important. They can lift morale and give a feeling of security in time of stress. Also, canned foods won't require cooking, water or special preparation. Following are recommended short-term food storage plans.

SPECIAL CONSIDERATIONS

As you stock food, take into account your family's unique needs and tastes. Try to include foods that they will enjoy and that are also high in calories and nutrition. Foods that require no refrigeration, preparation or cooking are best.

Individuals with special diets and allergies will need particular attention, as will babies, toddlers and elderly people. Nursing mothers may need liquid formula, in case they are unable to nurse. Canned dietetic foods, juices and soups may be helpful for ill or elderly people.

Make sure you have a manual can opener and disposable utensils. And don't forget non-perishable foods for your pets.

SHELF-LIFE OF FOODS FOR STORAGE

Here are some general guidelines for rotating common emergency foods.

- Use within six months:
  - Powdered milk (boxed)
  - Dried fruit (in metal container)
  - Dry, crisp crackers (in metal container)
  - Potatoes
- Use within one year:
  - Canned condensed meat and vegetable soups
  - Canned fruits, fruit juices and vegetables
  - Ready-to-eat cereals and uncooked instant cereals (in metal containers)
  - Peanut butter
  - Jelly
  - Hard candy and canned nuts
  - Vitamin C
- May be stored indefinitely (in proper containers and conditions):
  - Wheat
  - Vegetable oils
  - Dried corn
  - Baking powder
  - Soybeans
  - Instant coffee, tea and cocoa
  - Salt
  - Non-carbonated soft drinks
  - White rice
  - Bouillon products
  - Dry pasta
  - Powdered milk (in nitrogen-packed cans)

IF THE ELECTRICITY GOES OFF

FIRST, use perishable food and foods from the refrigerator.

THEN, use the foods from the freezer. To minimize the number of times you open the freezer door, post a list of freezer contents on it. In a well-filled, well-insulated freezer, foods will usually still have ice crystals in their centers (meaning foods are safe to eat) for at least three days.

FINALLY, begin to use non-perishable foods and staples.

HOW TO COOK IF THE POWER GOES OUT

For emergency cooking you can use a fireplace, or a charcoal grill or camp stove can be used outdoors. You can also heat food with candle warmers, chafing dishes and fondue pots. Canned food can be eaten right out of the can. If you heat it in the can, be sure to open the can and remove the label first.

Disaster Supplies

Emergency Supplies. It's 2:00 a.m. and a flash flood forces you to evacuate your home-fast. There's no time to gather food from the kitchen, fill bottles with water, grab a first-aid kit from the closet and snatch a flashlight and a portable radio from the bedroom. You need to have these items packed and ready in one place before disaster strikes.

Pack at least a three-day supply of food and water, and store it in a handy place. Choose foods that are easy to carry, nutritious and ready-to-eat. In addition, pack these emergency items:

- Medical supplies and first aid manual n Money and matches in a waterproof
- Hygiene supplies container
- Portable radio, flashlights and n Fire extinguisher extra batteries n Blanket and extra clothing
- Shovel and other useful tools
- Infant and small children's needs (if appropriate)
- Household liquid bleach to purify drinking water.
- Manual can opener

Learn More

The Federal Emergency Management Agency's Community and Family Preparedness Program and the American Red Cross Community Disaster Education Program are nationwide efforts to help people prepare for disasters of all types. For more information, please contact your local emergency management office and American Red Cross chapter.

WATER DEMAND FORECASTING

Steven J. Renzetti
Brock University
St. Catharines, Ontario, Canada

INTRODUCTION

People who are involved with planning, constructing, and operating water supply facilities frequently must have an idea of the current state of water use and also what future water demands are likely to be. This need is made all the more important when water use is rising rapidly and when the cost of major water infrastructure projects is considered. These factors imply that there is a need for accurate forecasts of water demands so that infrastructure projects are correctly sized and scheduled. The purpose of this article is to examine briefly the concepts and methods for forecasting water demands (for a more detailed discussion, see Refs. 3–5). We begin by considering the general nature of water demand forecasts.

ISSUES IN FORECASTING WATER DEMANDS

There are several issues surrounding the general practice of water demand forecasting. Boland (1) points out that any forecast has two essential components: explanation and prediction: "Explanation of water use usually takes the form of a model that relates the past observed level of water use to various variables. Replacing past values of the explanatory variables with those expected in the future produces a prediction of future water use" (pp. 162–163). Historically, the first of these two tasks—explaining water use—has been done by using 'fixed-coefficient' models. In this approach, it is assumed that water use is related to a single explanatory factor such as population and, further, that the relationship between water use and population is a fixed, proportional one. Dziegielewski (2), for example, provides a brief review of the history of urban water demand forecasting and demonstrates that, in the 'traditional' method of forecasting, total future demand is predicted as the product of expected population growth and a fixed per capita water use coefficient. This method was subsequently refined by disaggregating total water use by user classes, area, and time period. The fundamental forecasting method remained the same, however; 'unit water use coefficients' is multiplied by the projected growth in a particular user group in a specific location. The fundamental shortcoming of the fixed coefficient approach is that it fails to anticipate changes in the relationship between water use and the dominant explanatory variable that may arise from changes in other, neglected variables. For example, forecasts of residential water use that are based on the number of households in an area may overstate future water demands if they neglect the impacts of rising water prices. More recently, more complex models have sought to estimate statistically the relationship between water use and a set of explanatory variables, including water prices.

Estimated relationships form the basis for demand forecasts. Dziegielewski (2) describes how a computer model commonly used to forecast urban water demands (IWR-MAIN) is evolving to incorporate more complex models of water use.

Another feature of water demand forecasts is that they have different time horizons and that these different time horizons reflect, in part, the different uses to which water demand forecasts are put. These may be roughly categorized into short-, medium-, and long-term forecasts (3,6). Short-run forecasts are conducted by water utility managers and water supply agencies to anticipate daily peak loads and other seasonal features of water use. Medium-term forecasts often examine user group investment decisions (such as the locations of new housing developments and manufacturing facilities) to establish anticipated changes in water demands during the next 5–10 years. These forecasts are then used to guide water agency investment decisions. Finally, long-term forecasts can investigate the impacts of structural and technological changes in the economy (such as the shift in employment from manufacturing to service-based industries) as well as the impacts of major policy changes (such as water quality legislation).

CASES OF WATER DEMAND FORECASTING

As mentioned earlier, water demand forecasts can be conducted on a variety of geographic scales. This section draws on chapter 11 of Renzetti (5) and considers examples of water demand forecasts conducted on municipal, regional, and national scales. Boland (1) summarizes a forecasting exercise where the IWR-MAIN water demand model was combined with predictions of climate change to predict water use in Washington, DC. Boland combines the predictions of five climate change models with forecasts of other explanatory variables such as prices, incomes, and housing stock. These are then fed into IWR-MAIN, and the model provides predicted water use levels for 2030 under different climate and policy scenarios. In the absence of any change in climate patterns, summer water use in the region is expected to grow 100% by 2030. Once the impacts of climate change are incorporated, the predicted growth ranges from 74% to 138%, assuming no change in policies. However, implementing relatively modest nonprice conservation efforts (such as adopting a revised plumbing code) reduces the range of forecasted growth in summer water use to between 40 and 92%. Adding a 50% increase in real water prices reduces expected water use growth even further: predicted growth in summer water use between 1985 and 2030 in this case ranges from 26.6% to 73.6%.

Planners often also forecast the water demands of major water-using sectors within a region. Representative of this type of exercise is the California Department of Water's recent effort to forecast urban, agricultural, and in-stream water use in California to the year 2020 (7). Urban water use is assumed to be the product of population growth and changes in per capita water use the latter is modeled as determined by a number of factors such as income, water prices, and conservation measures. Statewide urban per capita water use in 1995 was 229 U.S. gallons/day; the forecasted level of per capita water use in 2020 is 243 gallons/day without new conservation measures and 215 gallons/day using new conservation measures. The California Department of Water adopts a sophisticated and multifaceted approach to forecasting irrigation water demands. Each crop's aggregate irrigation water use is calculated as the product of the crop's water "requirements" and the statewide irrigated acreage for the crop. Individual crop's water requirements are estimated using agronomic and climate data. Expected statewide acreage for each crop is a function of forecasted market conditions and expected government policies. The final type of water use to be factored into the Californian forecasts is in-stream needs. It turns out that these are largely determined by legislation and future climate conditions. After summing across sectors, total Californian water use in 1995 was 79,490 thousand acre-feet and, assuming a continuation of past climatic conditions, total Californian water use in 2020 is forecast at 80,500 thousand acre-feet. The prediction of constant total water use masks the fact that several sectors and regions are expected to exhibit different rates of growth. For example, total projected urban water use will rise from 8770 thousand acre-feet in 1995 to 12,020 thousand acre-feet in 2020, and agricultural water use is expected to decline from 33,780 thousand acre-feet to 31,500 thousand acre-feet for the same period.

Finally, there have been considerable efforts expended to forecast water demand growth on a national and even global scale. Gleick (8) critically surveys researchers' efforts to project global water use into the twenty-first century. Gleick points out that early projections were essentially extrapolations of current water use patterns under 'business as usual' assumptions regarding technology, water pricing, and water use efficiency. The extrapolations are then based on projections of expected growth in population, agricultural, industrial, and energy output. Gleick contends, however, that almost all of the projections based on fixed-coefficient models of water use have significantly overstated actual water use in the past because they cannot anticipate the extent to which water's growing scarcity induces technological and institutional innovation, changes in sectoral composition, and slowing rates of population growth. For example, Gleick compares a number of projections of global water use in 2000 with actual recorded water use and finds that actual withdrawals in the mid-1990s were only half of what most forecasters in the 1960s and 1970s projected.

A representative example of this type of effort is Seckler et al. (4). The authors project per capita and total withdrawals for 118 countries for the period 1990–2025. Projections of per capita residential withdrawals are based on population or GDP growth, and per capita irrigation water use is assumed constant ("business as usual") or declines due to improved technology ("high efficiency"). The impact of improved efficiency is dramatic. In the first scenario, global irrigation water use grows 60% for the period; in the second scenario, irrigation water use grows by only 13%. Similarly, global water withdrawals increase by 56% in the "business as usual" scenario and by 26% in the 'high efficiency' scenario.

CONCLUSIONS

The rapid growth in water demands that has been observed in many regions and sectors and the costs of servicing these demands highlight the need for accurate water demand forecasts. This article has demonstrated that researchers and utility operators have developed techniques to conduct forecasts for different time periods and different geographic scales. Historically, the fixed-coefficient method of explaining and predicting water demands has dominated, but experience with this model has demonstrated that it neglects many factors (such as prices, incomes and, water conservation programs) that influence demands. This shortcoming has typically led fixed-coefficient models to overstate expected water demand growth. It is expected that more sophisticated models that account for a wider range of factors that influence water use and that allow for feedback between water use and economic conditions will provide more accurate forecasts of water demands.

BIBLIOGRAPHY

1. Boland, J. (1997). Assessing urban water use and the role of water conservation measures under climate uncertainty. In: *Climate Change and Water Resources Planning Criteria*. K. Frederick, D. Major, and E. Stakhiv (Eds.). Kluwer Academic, Dordrecht, pp. 157–171.
2. Dziegielewski, B. (1996). Long-term forecasting of urban water demands. In: *Marginal Cost Rate Design and Wholesale Water Markets*. D. Hall (Ed.). *Advances in the Economics of Environmental Resources* vol. 1, JAI Press, Greenwich, Connecticut, pp. 33–47.
3. Billings, R.B. and Jones, C.V. (1996). *Forecasting Urban Water*. American Water Works Association, Demand Denver.
4. Seckler, D., Barker, R., and Amarasinghe, U. (1999). Water scarcity in the twenty-first century. *Int. J. Water Resour. Dev.* **15**(1/2): 29–42.
5. Renzetti, S. (2002). *The Economics of Water Demands*. Kluwer Academic, Norwell, Massachusetts.
6. Baumann, D. (1998). Forecasting urban water use: theory and principles. In: *Urban Water Demand Management and Planning*. D. Baumann, J. Boland and W.M. Hanemann (Eds.). McGraw-Hill, New York, pp. 77–94.
7. California Department of Water Resources (1998). *The California Water Plan Update*. Bulletin 160–98, Sacramento, California.
8. Gleick, P. (2000). *The World's Water 2000–2001*. Island Press, Washington, DC.

REMOTE SENSING AND GIS APPLICATION IN WATER RESOURCES

SANJAY KUMAR JAIN
National Institute of Hydrology
Roorkee, India

Water resources are essential for human life, agriculture, and hydroelectric power generation. Remote sensing and the Geographic Information System (GIS) help in creating an appropriate information base for efficiently managing these resources. The synoptic view provided by satellite remote sensing and the analytical capability provided by the GIS offer a technologically appropriate method for studying various features related to land and water resources.

Remote sensing is a tool that permits accurate and real-time evaluation, continuous monitoring or surveillance, and forecasts of inland water resources. Remote sensing systems are used to observe the earth's surface from different levels of platforms, such as satellites and aircraft, and make it possible to collect and analyze information about resources and environment across large areas. Remote sensors record electromagnetic energy reflected or emitted from the earth's surface. Different kinds of objects or features such as soils, vegetation, and water reflect and emit energy differently. This characteristic makes it possible to measure, map, and monitor these objects and features using satellite or aircraft-borne remote sensing systems. Satellite imagery offers a number of advantages over conventional survey techniques:

- areal synoptic coverage (gives areal information as against point information through conventional techniques)
- repetitive global coverage (for monitoring change)
- real-time processing
- sensing of surrogates rather than the desired specific observation
- multispectral coverage
- more automation, less human error

For many water related studies, remote sensing data alone are not sufficient; they have to be merged with data from other sources. Hence a multitude of spatially related (i.e., climatic and geographic) data concerning rainfall, evaporation, vegetation, geomorphology, soils, and rocks have to be considered. In addition, information is also required such as locations and type of tube wells and rain and river gauges. Thus, the fast storage, retrieval, display, and updating of map contents are important functions. A system that can store data, select and classify stations, and perform mathematical and sorting operations is called a database, and information can be extracted from it for a given purpose. If this information can also be displayed in the form of maps, we can speak of geographic information. So this complete set of information forms the Geographic Information System (GIS). The GIS is an effective tool for storing, managing, and displaying spatial data (Fig. 1) often encountered in hydrology and water resources management studies. The GIS technology integrates common database operations, such as queries and statistical analysis, with the unique visualization and benefits of geographic analysis that maps and spatial databases offer. The use of remote sensing data in the GIS is shown in Fig. 2.

One of the capabilities of the GIS most important to water resources management studies is describing the topography of a region. A digital elevation model (DEM) provides a digital representation of a portion of the earth's terrain across a two-dimensional surface. DEMs have proved to be a valuable tool for topographic parameterization of hydrologic models, especially for drainage analysis, hillslope hydrology, watersheds, groundwater flow, and contaminant transport.

Problems of water resources development and efficient use such as frequent floods and droughts, waterlogging and salinity in command areas, an alarming rate of reservoir sedimentation due to deforestation, and deteriorating water quality and environment are varied and numerous in nature. All must be tackled by systematic approaches involving a judicious mix of conventional methods and remotely sensed data in the GIS environment. The availability of remotely sensed data and the use of the GIS has provided significant impetus to hydrologic analysis and design and their use in water resources

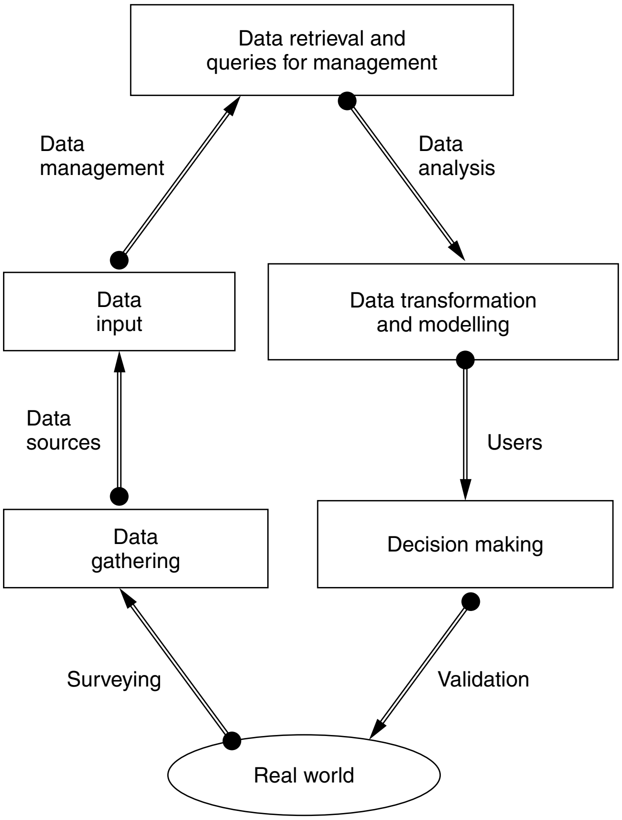

Figure 1. The GIS, a management tool.

planning and development. Remote sensing and GIS techniques have been extensively used in various areas of water resources development and management, such as land use/cover classification, precipitation, snow cover, floodplain management, command area studies, waterlogging and soil salinity, sedimentation in reservoirs, and hydrologic modeling, which are defined below in brief.

LAND USE/LAND COVER CLASSIFICATION

Multispectral classification of land cover types was one of the first well-established remote sensing applications for water resources. Numerous investigators have used a classification of land cover from satellite data as input to various water resources studies. Land use features can be identified, mapped, and studied on the basis of their spectral characteristics. Healthy green vegetation has considerably different characteristics in the visible and near-infrared regions of the spectrum, whereas dry bare soil has a relatively stable reflectance in both regions of the spectrum. Water shows very low reflectance in the visible part of the spectrum and almost no reflectance in the infrared part of the spectrum. Thus, by using multispectral data suitably, different ground features can be differentiated from each other, and thematic maps depicting land use can be prepared from satellite data. Land use data devised from remote sensing are an ideal database for distributed hydrologic modeling changes. In many hydrologic models, the parameters of vegetation related model components (e.g., for interception and evapotranspiration) are chosen, dependent on land use.

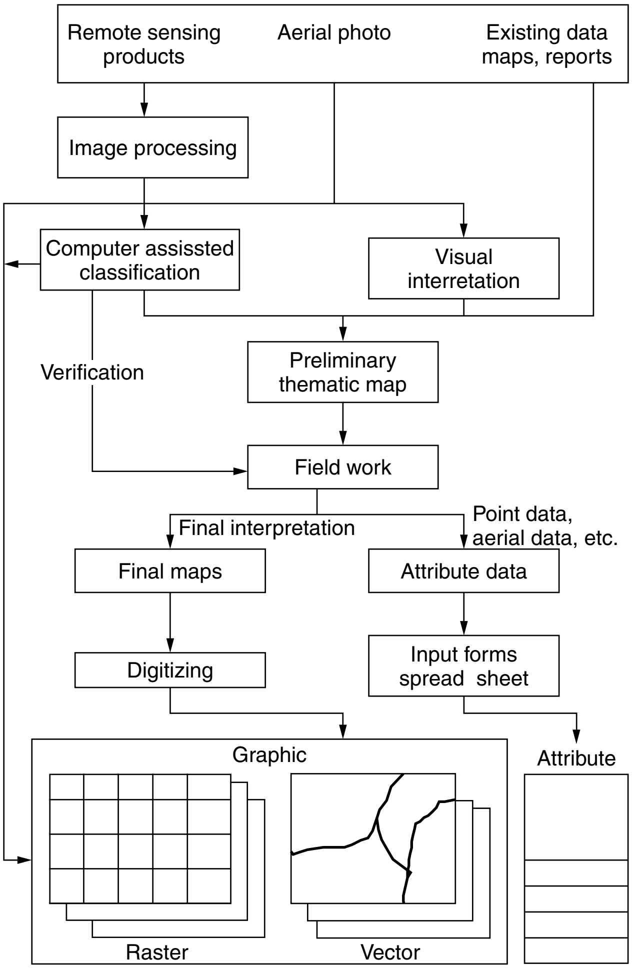

Figure 2. Remote sensing products in a GIS.

PRECIPITATION STUDIES

Precipitation, one of the driving forces in the hydrologic cycle, is a prime candidate variable for remote sensing measurement. There are many remote sensing possibilities, including ground based radar, visible and thermal infrared satellite imagery, and microwave satellite data. The use of ground based radar has been successful especially when used with an integrated rain-gauge network and in areas of low relief. Visible and thermal infrared techniques using measurements of cloud top reflectance and temperature have been used in a variety of ways by meteorologists and other scientists to estimate monthly, daily, and storm precipitation totals. Visible and infrared measurements represent observations only of the upper surface of clouds. In contrast, it is often believed that microwave radiation is not affected by the presence of clouds. Measuring rain from a spaceborne radar is very attractive because it can discriminate in range (altitude when operated from space), and radar reflectivity is directly related to rainfall rate.

SNOW COVER STUDIES

Measurement of the area of a basin covered by snow is similar to land cover measurement. Conventional methods have limitations in monitoring the snow covered area in mountainous basins because of inaccessibility.

Very little information on snow is collected regularly in mountainous areas, so remote sensing remains the only practical way of obtaining some relevant information on the snow cover in the large number of mountain basins. At present the visible, near IR, and thermal IR data from various satellite (Landsat, IRS, NOAA) are being used operationally for mapping the areal extent of snow cover in mountainous basins. Visible and near-infrared wavelengths do not penetrate far into the snowpack, so they provide information mainly about the surface of the snowpack. Microwave remote sensing is promising because of its ability to penetrate the dry snowpack and its capability of acquiring data in cloudy or nighttime conditions. Snow cover data extracted by satellite remote sensing is immediately usable in snowmelt runoff models. In addition to the extent of snow cover, satellite data are useful in computing snow water equivalent from microwave data.

SOIL MOISTURE ASSESSMENT

Remotely sensed data have great potential for providing areal estimates of soil moisture rather than point measurements. Quantitative assessment of soil moisture regime is essential for water balance models, irrigation scheduling, crop management, and surface and subsurface flow predictions. Remote sensing of soil moisture can be accomplished to some degree in all regions of the electromagnetic spectrum. The use and application of remotely sensed data for soil moisture depends on measuring the electromagnetic energy that is reflected or emitted from the surface. Thus, either direct observations of the soil by remote sensors or indirect estimates using plant condition are useful in soil moisture observations. The soil moisture beneath microwave remote sensing can directly measure the dielectric properties of the earth's surface, which in turn is strongly dependent on moisture content.

FLOODPLAIN MANAGEMENT

Reliable data on river morphology, river meandering, the extent of flooding, and duration are required for proper planning of flood control projects. In conventional methods of flood risk zoning, the flood discharge is routed through a river reach to estimate the likely inundation due to spilling over banks/embankments based on topographic contour maps available and the configuration of the river geometry obtained from land surveys. Continuous availability of satellite based remote sensing data has made understanding the dynamics of floods much easier. Satellite remote sensing techniques provide wide area coverage, repetitiveness, and consistency, which enables collecting reliable information on all major floods.

Remote sensing can provide information on flood inundated areas for different magnitudes of floods, so that the extent of flooding in floodplains can be related to flood magnitude. The duration of flooding can be estimated with the help of multiple coverage satellite imagery of the same area within 2–3 days by satellites. High-resolution satellite data provide information on the floodplain and effectiveness of flood control works. The extent of inundation for specific flood return periods can be estimated. Using close contour information, the extent of inundation for a given elevation can be estimated, this is a vital input for risk zone mapping.

The GIS provides a broad range of tools for determining the area affected by floods and for forecasting areas that are likely to be flooded due to the high water level in a river. Spatial data stored in the digital database of the GIS, such as a digital elevation model (DEM), can be used to predict future floods. The GIS database may also contain agricultural, socioeconomic, communication, population, and infrastructural data. This can be used in conjunction with flooding data to adopt an evacuation strategy, rehabilitation planning, and damage assessment in a critical flood situation.

MAPPING AND MONITORING OF WATERSHEDS

Proper planning of watersheds is essential for conservation of water and land resources and their productivity. Characterization and analysis of watersheds are a prerequisite for this. Watershed characterization involves measuring parameters of geologic, hydrogeologic, geomorphological, hydrologic, soil, and land cover/land use. Remote sensing via aerial and spaceborne sensors can be used for watershed characterization, assessing management requirements, and periodic monitoring. The various physiographic measurements that can be obtained from remotely sensed data are size, shape, topography, drainage pattern, and landforms. Quantitative analysis of drainage networks enables relationships among different aspects of the drainage pattern of a basin to be formulated as general laws and to define certain useful properties/indexes of the drainage basin in numerical terms. The laws of stream numbers, stream length, and stream slopes can be derived from measurements made in the drainage basin. Remote sensing along with ground based information in the GIS mode can used for broad and reconnaissance level interpretations of land capability classes, irrigation suitability classes, potential land users, responsive water harvesting areas, monitoring the effects of watershed conservation measures, correlation for runoff and sediment yields from different watersheds, and monitoring land use changes and land degradation.

COMMAND AREA STUDIES

Water management in command areas requires serious attention in view of the disappointing performance of our irrigation projects, despite huge investments. The command area is the total area lying between drainage boundaries, which can be irrigated by a canal system.

Remote sensing can play a useful complementary role in managing the land and water resources of command areas to maximize production. Management of water supplies for irrigation in command areas is a critical problem to tackle with limited quantities. This requires information on total demand and the distribution of demand for irrigation in command areas. Moreover, the vastness of the areas involved, time constraints, and yearly changes demand fast inventories of conditions. As more area is brought under irrigation, crop monitoring also becomes essential for estimating agricultural production and efficient planning of water management. In all these, remote sensing can be looked upon as an aid in planning and decision making. The usefulness of remote sensing techniques in inventorying irrigated lands; identifying crop types, their extent, condition; and production estimation has been demonstrated in various investigations. Conjunctive use planning of surface and groundwaters can be done using remotely sensed information on surface water in conjunction with ground based data on groundwater availability. This would permit development of conjunctive use models for water allocations in a GIS environment.

WATERLOGGING AND SOIL SALINITY

Waterlogging and soil salinity are some of the major land degradation processes that restrict the economic and efficient use of soil and land resources in command areas. To assess waterlogging in command areas, multispectral and multitemporal remote sensing data are very useful. The satellite data thus provide a quick and more reliable delineation of waterlogged areas and standing water. The spatial distribution of soil affected by “positional waterlogging” (i.e., that due to its location in the landscape) can be modeled from digital topographic data using the concept of contributing area. This waterlogging depends on two topographic factors: (1) the local slope angle and (2) the drainage area. The probability of waterlogging increases with the contributing drainage area and decreases with increasing local slope angle. The waterlogging phenomenon is related to topography, so digital terrain modelling (DTM) can aid in detecting waterlogged areas. DTM provides information regarding slope and aspect, which in turn provide information about the areas that are susceptible to waterlogging.

One of the common practices for observing waterlogged areas in command areas is to take observations in existing open wells at regular intervals, twice a year in the pre- and postmonsoon seasons. Data are also collected on water quality. The information thus collected is used to draw hydrographs and the depth of the water table to prepare maps subsequently. So preparation of these maps can help a lot in identifying waterlogged zones in a command area. Using field data, which is available as point data, groundwater depth distributional maps can be prepared in the GIS. With the help of these maps, shallow GW areas, areas susceptible to waterlogging, can be identified. The areas falling within the 0–1.5 m range generally indicate waterlogged or salt-affected patches depending primarily on the soil characteristics, particularly texture.

RESERVOIR SEDIMENTATION

Most common conventional techniques for sedimentation quantification are (1) direct measurement of sediment deposition by hydrographic surveys and (2) indirect measurement of sediment concentration by the inflow–outflow method. Both methods are laborious, time-consuming, and costly and have their own limitations. Sampling and measurement of suspended sediments is a tedious and expensive program for either *in situ* or laboratory work.

The introduction of remote sensing techniques in the recent past has made it very cheap and convenient to quantify sedimentation in a reservoir and to assess its distribution and deposition pattern. Remote sensing techniques, offering data acquisition over a long time period and broad spectral range, are superior to conventional methods for data acquisition. The advantage of satellite data over conventional sampling procedures include repetitive coverage of a given area every 16–22 days, a synoptic view which is unobtainable by conventional methods, and almost instantaneous spatial data across the areas of interest. Remote sensing techniques provide a synoptic view of a reservoir in a form very different from that obtained by surface data collection and sampling.

WATER QUALITY STUDIES

Recently, the alarming proportions of water quality deterioration have necessitated rapid monitoring for efficient checks to prevent further deterioration and to cleanse our polluted water resources. Moreover, surveillance of water quality is an important activity for multiple uses such as irrigation and water supply. Water quality is a general term used to describe the physical, chemical, thermal, and/or biological properties of water.

A combination of ground (water) and remote sensing measurements are required to collect the data necessary to develop and calibrate empirical and semiempirical models and validate the more physically based models. Water samples analyzed for substances of interest (i.e., suspended sediment, chlorophyll) should be collected at the same time (or on the same day) that the remote sensing data are collected. Location of sample sites should be determined by GPS (or other available technique) so that the correct data (pixel information) can be extracted from remote sensing for comparison. Remote sensing applications to water quality are limited to measuring those substances or conditions that influence and change the optical and/or thermal characteristics of the apparent surface water properties. Suspended sediment, chlorophyll (algae), oil, and temperature are water quality indicators that can change the spectral and thermal properties of surface waters and are most readily measured by remote sensing techniques.

GROUNDWATER STUDIES

A remote sensing system is quite helpful in groundwater exploration because remotely sensed data provide a synoptic view of high observational density. The common current remote sensing platforms record features on the surface. Most of the information for groundwater, as yet, has to be obtained by qualitative reasoning and semiquantitative approaches. Remotely sensed information is often surrogate and has to be merged with geohydrologic data to become meaningful. Vegetation can be used as an indicator if local knowledge is available and the types can be identified on the satellite data. Apart from the contribution that remote sensing can make to understanding regional hydrogeology—necessary for managing groundwater resources—perhaps the strongest application for management is the evaluation of recharge, groundwater drafts for irrigation, and identification of flow systems in areas where there is a paucity of geohydrologic data. Surface conditions, soils, weathered zones, geomorphology, and vegetation determine recharge, suitability for artificial recharge, and soil and water conservation measures that can affect recharge.

Groundwater vulnerability to pollution is also directly related to surface conditions. Indexing methods for group depth to water table, net recharge, topography, impact of vadose zone media, and hydraulic conductivity of the aquifer (leading to the acronym DRASTIC) into a relative ranking scheme use a combination of weights and ratings to produce numerical values.

HYDROLOGIC MODELING

In the early days, the GIS was used mainly as a hydrologic mapping tool. Nowadays, it plays a more important role in hydrologic model studies. Its applications span a wide range from sophisticated analyses and modeling of spatial data to simple inventories and management tools. The GIS has evolved as a highly sophisticated database management system to put together and store the voluminous data typically required in hydrologic modeling. The application of the GIS has enhanced the capacity of models in data management, parameter estimation, and presentation of model results, but the GIS cannot replace hydrologic models in solving hydrologic problems.

Due to its data handling and manipulation capabilities, the GIS is increasingly being used as an interface and data manager for hydrologic models. There are four levels of linkage of hydrologic model with the GIS. These levels vary from considering the GIS and the model as separate systems to fully integrating the model and the GIS.

One typical application of a GIS in watershed analysis is predicting the spatial variability of surface erosion from spatial data sets obtained from maps of the vegetative cover, soils, and slope of the area. Solutions of a surface erosion (soil loss) prediction model, for example, the universal soil loss equation (USLE) or its modifications, combine spatial data sets, their derivatives, and other information necessary to predict the spatial variability of surface erosion on a watershed. This analysis can determine areas of potentially severe surface erosion and provide an initial step in the appraisal of surface erosion problems.

GIS FOR DECISION MAKING

The GIS is derived from multiple sources of data of different levels of accuracy. Though a single piece of data can be assigned an accuracy value, information derived from multiple sources of inaccurate data can also be assigned a level of accuracy. In any pictorial representation of data, uncertainty can be brought in as one of the dimensions to guide final decision making. Any decision today has to depend on a variety of factors, which are available in an information system like the GIS. However, weighting as well as the proper use of such data is still problematic.

The decision support system (DSS) can be designed as an interactive, flexible, and adaptable computer based information system (CBIS), especially developed to support the solution of a management problem for decision making. It uses data, provides an easy interface, and allows for decision-makers' own insights to illustrate the objectives and information characteristics of the various levels of decision making. DSS provides a framework for incorporating analytical modeling capabilities with a database to improve the decision-making process. Currently available DSS provides primarily for well-structured problems. It are also being used for decision making in water resources planning and management.

CONCLUSIONS AND FUTURE NEEDS

Many water resources applications can profit from the use of remote sensing and the GIS. The reason for adopting GIS technology is that it allows displaying spatial information in integrative ways that are readily comprehensible and visual. Remote sensing is now being widely regarded as a layer in the GIS. Although remote sensing is a specialized technique, it is now being accepted as a basic survey methodology and as a means of providing data for a resource database. The GIS provides a methodology by which data layers can be interrelated to arrive at wider decisions (1).

The research needs in the area of water resources are as follows:

- A first and very important aspect is data availability and compatibility in any GIS related study. Spatial information required for water resources studies should be readily available for timely execution. The data banks should provide digitized maps and their spatial data compatible with various systems. Such data availability could significantly speed up the analysis.
- One difficult task in incorporating the GIS in water resources modeling is interfacing water system models with the GIS. Automation of interfacing tasks is one of the areas to be researched in incorporating the GIS and available models.
- The recent development of the decision support system (DSS) to assist in water resources decision making holds the key for integrating the GIS and water resources models.

- Another area of potential research to enhance the modeling process further is the integration of expert systems and the GIS. Expert GIS systems can be used to provide regulatory information by linking regulatory facts stored in a database to sites located in a GIS through an expert system query interface.
- Further research is needed for comparing the GIS packages available on the market and their positive and negative aspects, providing check lists for GIS users.

BIBLIOGRAPHY

1. Tsihrintzix, V.A., Hamid, R., and Fuentes, H.R. (1996). Use of GIS in water resources: a review. *Water Resour. Manage.* **10**: 251–277.

READING LIST

Schultz, G.A. and Engment, E.T. (Eds.). (2000). *Remote Sensing in Hydrology and Water Management*. Springer Verlag, New York.

Rango, A. (1990). Remote sensing of water resources: accomplishments, challenges and relevance to global monitoring. *Proc. Int. Symp. Remote Sensing Water Resour.*, Enschede, The Netherlands, August 20–24.

GLOBALIZATION OF WATER

A.Y. Hoekstra
UNESCO–IHE Institute for
Water Education
Delft, The Netherlands

WHY WATER IS A GLOBAL RESOURCE

People in Japan can affect the hydrological system in the United States. People in Europe can affect regional water systems in Thailand or Brazil. There are basically two mechanisms that make global connections between seemingly local water systems. First, the climate system connects different places on earth, because evaporation in one place results in precipitation in another place. The climate system is inherently global: local emissions of greenhouse gases contribute to a changing global climate, thus affecting temperature, evaporation, and precipitation patterns elsewhere. In this way, human activities in the economic centers of the world affect base and peak flows in rivers throughout the world. There is, however, a second mechanism through which people can affect water systems in other parts of the world. A European consumer of Thai rice raises rice demand in Thailand and subsequently the use of water for rice irrigation in Thailand. Globally, roughly one-fifth of the water used in agriculture is applied in areas used for producing export commodities. This fraction is increasing in line with the increase of global trade. Scientists have known for several decades that human activities within a river basin strongly impact on the water flows and water quality in such a river basin. It is more recent that scientists have started to realize that local water problems are often, in their roots, problems that cannot be solved locally or regionally, because the driving forces lay outside the region. The gradual disappearance of the Aral Lake in Uzbekistan–Kazakhstan, for instance, is directly linked to the global demand for cotton.

EFFICIENT USE OF THE GLOBAL WATER RESOURCES

According to international trade theory, countries with abundant water resources might have a comparative advantage in producing water-intensive goods if compared to water-poor countries. Since water resources are quite unevenly distributed over the world, one should not be surprised that under conditions of free trade water-intensive production processes get concentrated in the water-rich parts of the world. A water-scarce country can thus aim at importing products that require a lot of water in their production (water-intensive products) and exporting products or services that require less water (water-extensive products). This is called *import of virtual water* (as opposed to import of real water, which is generally too expensive) and will relieve the pressure on the nation's own water resources. For water-abundant countries an argument can be made for *export of virtual water*. Import of water-intensive products by some nations and export of these products by others results in international "virtual water flows."

The idea of achieving "global water use efficiency" through virtual water trade between countries is relatively new. Traditionally, water managers have focused on achieving "local water use efficiency" at user level and "water allocation efficiency" at river basin level. "Local water use efficiency" can be increased by creating awareness among the water users, by charging prices based on full marginal cost, and by stimulating water-saving technology. "Water allocation efficiency" can be improved by allocating the water within a catchment or river basin to those types of use where water creates the highest added value. A key question today is which level—the user, river basin, or global level—will be most relevant in increasing overall water use efficiency and reducing overall impact on the globe's water systems.

THE CONCEPT OF "VIRTUAL WATER"

The volume of virtual water "hidden" or "embodied" in a particular product is defined as the volume of water used in the production process of that product (1,2). Not only agricultural products contain virtual water: most studies to date have been limited to the study of virtual water in crops. Industrial products and services also contain virtual water. As an example of virtual water content, one often refers to the virtual water content of grains. It is estimated that for producing 1 kg of grain, grown under rain-fed and favorable climatic conditions, we need about 1–2 m^3 of water, which is 1000–2000 kg of water. For the same amount of grain, but growing in an arid country, where the climatic conditions are not favorable (high temperature, high evapotranspiration), we need up to 3000–5000 kg of water.

If one country exports a water-intensive product to another country, it exports water in virtual form. In this way, some countries support other countries in their water needs. Worldwide both politicians and the general public increasingly show interest in the pros and cons of "globalization" of trade. The tension in the debate relates to the fact that the game of global competition is played with rules that many see as unfair. Knowing that economically sound water pricing is poorly developed in many regions of the world, this means that many products are put on the world market at a price that does not properly include the cost of the water contained in the product. This leads to situations in which some regions in fact subsidize export of scarce water.

METHOD TO CALCULATE VIRTUAL WATER FLOWS

Calculation of the Virtual Water Content of a Product

The virtual water content of a product is a function of the water volumes used in the different stages of the production process. The virtual water content of a particular crop in a particular country can, for instance, be calculated on the basis of country-specific crop water requirements and crop yields (3):

$$\mathrm{VWC}[n, c] = \mathrm{CWR}[n, c]/Y[n, c] \qquad (1)$$

Here, VWC denotes the virtual water content (m^3/ton) of crop c in country n, CWR the crop water requirement (m^3/ha), and Y the crop yield (ton/ha). The crop water requirement can be calculated from the accumulated crop evapotranspiration ET_c (in mm/d) over the complete growing period. The crop evapotranspiration ET_c follows from multiplying the "reference crop evapotranspiration" ET_0 with the crop coefficient K_c:

$$\mathrm{ET}_c = K_c \times \mathrm{ET}_0 \qquad (2)$$

The crop coefficient accounts for the actual crop canopy and aerodynamic resistance relative to the hypothetical reference crop. The crop coefficient serves as an aggregation of the physical and physiological differences between a certain crop and the reference crop. The concept of "reference crop evapotranspiration" was introduced by FAO to study the evaporative demand of the atmosphere independently of crop type, crop development, and management practices. The only factors affecting ET_0 are climatic parameters. The reference crop evapotranspiration ET_0 is defined as the rate of evapotranspiration from a hypothetical reference crop with an assumed crop height of 12 cm, a fixed crop surface resistance of 70 s/m, and an albedo of 0.23. This reference crop evapotranspiration closely resembles the evapotranspiration from an extensive surface of green grass cover of uniform height, actively growing, completely shading the ground, and with adequate water (4). Reference crop evapotranspiration can be calculated on the basis of the FAO Penman-Monteith equation (4,5).

The calculation of the virtual water content of livestock products is a bit more complex. First, the virtual water content (m^3/ton) of live animals is calculated, based on the virtual water content of their feed and the volumes of drinking and service water consumed during their lifetime (6). Second, the virtual water content is calculated for each livestock product, taking into account the product fraction (ton of product obtained per ton of live animal) and the value fraction (ratio of value of one product from an animal to the sum of the market values of all products from the animal).

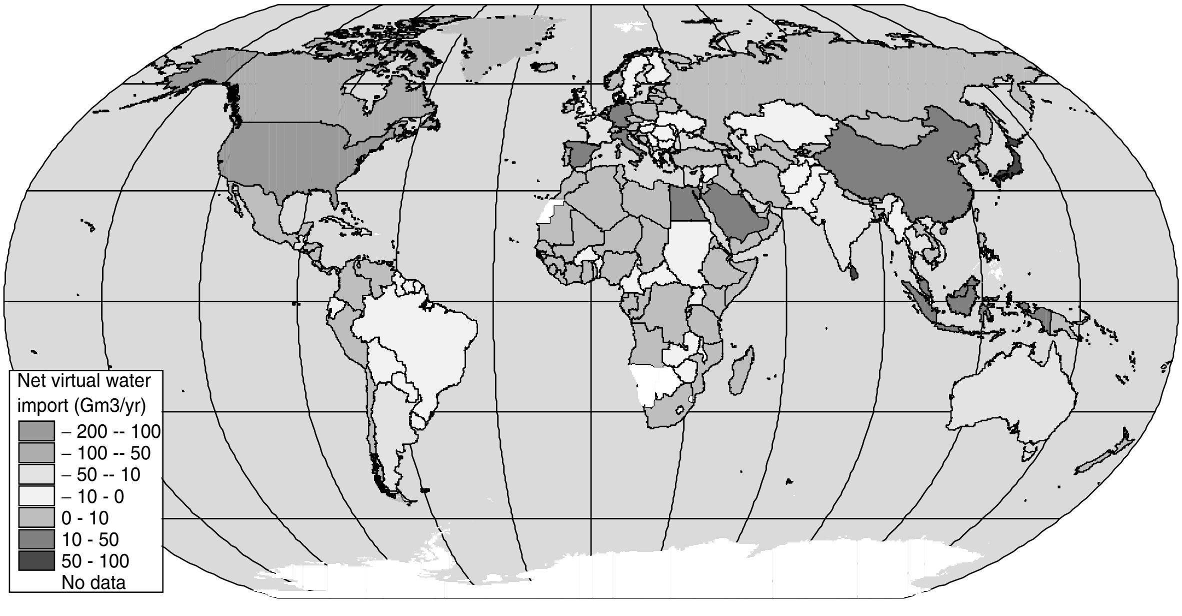

Figure 1. National virtual water balances over the period 1995–1999.

Table 1. Water Withdrawals, Virtual Water Import and Export, Water Scarcity, Water Self-sufficiency, and Water Dependency of Nations (1995–1999)

Country	Population	Water Availability, 10^6 m^3/yr	Water Withdrawal, 10^6 m^3/yr	Gross Virtual Water Import, 10^6 m^3/yr	Gross Virtual Water Export, 10^6 m^3/yr	Net Virtual Water Import, 10^6 m^3/yr	Water Scarcity, %	Water Self-sufficiency, %	Water Dependency, %
Afghanistan	25765766	65000	23261	58.5	287.5	−229	36	100	0
Albania	3387574	41700	1714	277.7	14.4	263.2	4	87	13
Algeria	29959010	14320	6074	9810.6	6.9	9803.7	42	38	62
Angola	12771448	184000	343	173.9	5.5	168.4	0	67	33
Argentina	36577450	814000	29072	1486.9	46755.4	−45268.4	4	100	0
Armenia	3798845	10529	2954	316.2	5.5	310.6	28	90	10
Australia	18963804	492000	23932	1011	30130.3	−29119.3	5	100	0
Austria	8095446	77700	2112	1281	976.1	304.9	3	87	13
Azerbaijan	7979460	30275	17247	1004.4	30	974.4	57	95	5
Bahrain	666956	116	299	137.4	0.2	137.1	258	69	31
Bangladesh	128837760	1210644	79394	8304.6	2562.6	5742	7	93	7
Belarus	10039496	58000	2789	1255.9	35	1220.9	5	70	30
Belgium-Lux	10227060	21400	8980	14412.4	2497	11915.4	42	43	57
Benin	6112575	24800	250	472.6	1077.8	−605.3	1	100	0
Bhutan	782229	95000	420	26.5	0	26.5	0	94	6
Bolivia	8139894	622531	1387	674.8	1732.1	−1057.3	0	100	0
Bosnia Herzegovina	3865576	37500	1354	238	63.8	174.2	4	89	11
Brazil	168220660	8233000	59298	23161.6	32161.8	−9000.2	1	100	0
Bulgaria	8213543	21300	10498	288.4	759.7	−471.4	49	100	0
Burkina Faso	11005226	12500	780	68.6	973.4	−904.7	6	100	0
Burundi	6677800	3600	234	3.7	0.3	3.4	6	99	1
Cambodia	11755836	476110	4091	130.1	27.7	102.4	1	98	2
Cameroon	14557762	285500	985	175.3	187.9	−12.6	0	100	0
Canada	30498614	2902000	45974	4814.4	59308.4	−54494	2	100	0
Central African Republic	3657263	144400	22	1.8	2.9	−1.1	0	100	0
Chad	7492965	43000	234	1.4	0	1.4	1	99	1
Chile	15013962	922000	12539	3262.6	1211.2	2051.4	1	86	14
China	1252042000	2896569	630289	30550.4	10114.9	20435.6	22	97	3
Colombia	41543956	2132000	10711	7535.4	865.2	6670.2	1	62	38
Comoros	544534	1200	13	39.3	0	39.3	1	25	75
Congo	2764600	832000	39	93.4	6.8	86.6	0	31	69
Costa Rica	3731672	112400	2677	1947.7	690	1257.7	2	68	32
Cote d'Ivoire	15580058	81000	931	773.5	83.5	690	1	57	43
Croatia	4395695	105500	1760	569.9	346	223.9	2	89	11
Cuba	11150144	38120	8204	1082.5	1304	−221.5	22	100	0
Czech Republic	10283004	13150	2566	1245.7	761.2	484.5	20	84	16
Denmark	5318089	6000	1267	1382.6	1843.6	−461	21	100	0
Djibouti	620352	300	8	109.5	0.2	109.3	3	7	93
Dominican Republic	8237523	20995	3386	731.8	2663.9	−1932.1	16	100	0
Ecuador	12409904	432000	16980	1594.2	2184.4	−590.2	4	100	0
Egypt	62782964	86800	68653	16937.1	901.6	16035.5	79	81	19
El Salvador	6155042	25230	1273	1142.3	94.6	1047.7	5	55	45
Eritrea	3988805	6300	304	74.8	0.3	74.6	5	80	20
Estonia	1388705	12808	163	631	100.4	530.6	1	24	76
Ethiopia	62782412	110000	2648	349.1	22.7	326.4	2	89	11
Fiji	802087	28550	69	174.6	0	174.6	0	28	72
Finland	5164368	110000	2478	918.9	1091.8	−172.9	2	100	0
France	58656600	203700	39959	9376.3	27051.4	−17675.1	20	100	0
Gabon	1198661	164000	128	100.7	0.7	100	0	56	44
Gambia	1263370	8000	32	319.3	164.2	155.1	0	17	83
Georgia	5188007	63330	3607	308.4	103	205.4	6	95	5
Germany	82109980	154000	47052	23260.4	9671.3	13589.1	31	78	22
Ghana	18875980	53200	520	671.2	217.5	453.8	1	53	47
Greece	10537058	74250	7759	3121.4	5088	−1966.6	10	100	0
Guatemala	11095762	111270	2005	1195.1	15536.6	−14341.5	2	100	0
Guinea Bissau	1174665	31000	110	8.3	5.4	2.9	0	97	3
Guyana	757015	241000	1642	67.9	226.6	−158.7	1	100	0
Haiti	7803032	14025	984	389	0	389	7	72	28
Honduras	6257825	95929	860	799.1	331.7	467.4	1	65	35
Hungary	10221682	104000	7641	635.6	4589.6	−3954	7	100	0
Iceland	277700	170000	153	64.8	1.6	63.2	0	71	29
India	997775760	1907760	645837	2413	34612.3	−32199.3	34	100	0
Indonesia	207029780	2838000	82773	21366.2	1139.2	20227	3	80	20
Iran	62762116	137510	72877	6623.1	803.4	5819.7	53	93	7
Iraq	22797032	96420	42702	1100.7	3.3	1097.4	44	97	3

Table 1. (*Continued*)

Country	Population	Water Availability, 10^6 m^3/yr	Water Withdrawal, 10^6 m^3/yr	Gross Virtual Water Import, 10^6 m^3/yr	Gross Virtual Water Export, 10^6 m^3/yr	Net Virtual Water Import, 10^6 m^3/yr	Water Scarcity, %	Water Self-sufficiency, %	Water Dependency, %
Ireland	3752276	52000	1129	945.8	201.9	743.9	2	60	40
Israel	6100032	1670	2041	5188.1	589.9	4598.2	122	31	69
Italy	57627528	191300	44372	19625.8	6762.1	12863.7	23	78	22
Jamaica	2604246	9404	409	392.8	137.3	255.5	4	62	38
Japan	126624200	430000	88432	59632	188.4	59443.6	21	60	40
Jordan	4742815	880	1016	4536	55	4481	115	18	82
Kazakhstan		109610	35008	41.8	7876	−7834.2	32	100	0
Kenya	29402552	30200	1576	970.2	169.7	800.5	5	66	34
Korea, Democratic People's Republic	22141004	77135	9024	643	2.2	640.8	12	93	7
Korea, Republic	46839720	69700	18590	22582.6	69	22513.6	27	45	55
Kuwait	1925635	20	445	497.8	0.1	497.7	2227	47	53
Kyrgyzstan	4844973	46450	10080	192.5	145.2	47.3	22	100	0
Lao	5159165	333550	2993	94.2	1.7	92.5	1	97	3
Latvia	2408205	35449	293	301.4	53.4	248	1	54	46
Lebanon	4267969	4837	1372	776.2	29.4	746.8	28	65	35
Liberia	3046804	232000	107	67.5	1.7	65.7	0	62	38
Libya	5176657	600	4811	789.1	45.4	743.7	802	87	13
Lithuania	3531820	24900	267	500.4	383.9	116.5	1	70	30
Macedonia	2020714	6400	847	149	97.9	51.1	13	94	6
Madagascar	15057966	337000	14970	320	131.7	188.3	4	99	1
Malawi	10096722	17280	1005	25.9	786.6	−760.8	6	100	0
Malaysia	22724518	580000	9016	11508.3	1255.9	10252.4	2	47	53
Mali	10588286	100000	6930	79.8	14.9	65	7	99	1
Mauritania	2579964	11400	1698	375.7	0.6	375.1	15	82	18
Mauritius	1173176	2210	612	564.7	274.9	289.7	28	68	32
Mexico	96615488	457222	78219	24361.3	15374.7	8986.7	17	90	10
Moldova Republic	4291104	11650	2308	82.8	455.5	−372.7	20	100	0
Mongolia	2377183	34800	444	24.6	10.1	14.5	1	97	3
Morocco	28240226	29000	12758	5617.8	87.4	5530.4	44	70	30
Mozambique	17331232	216110	635	337	85.1	251.9	0	72	28
Myanmar	47134402	1045601	33224	21.1	3501.3	−3480.2	3	100	0
Nepal	22507210	210200	10177	47.6	19	28.6	5	100	0
Netherlands	15812200	91000	7944	35002.3	5462.6	29539.7	9	21	79
New Zealand	3808760	327000	2111	1000.6	113.1	887.5	1	70	30
Nicaragua	4940828	196690	1300	583.6	333.1	250.5	1	84	16
Niger	10478080	33650	2187	309.5	107.9	201.6	6	92	8
Nigeria	123837060	286200	8004	5796.4	934.4	4862	3	62	38
Norway	4461300	382000	2185	2214.7	11.1	2203.6	1	50	50
Oman	2350640	985	1350	1228.1	119.6	1108.5	137	55	45
Pakistan	134871900	233770	169384	2547.1	2556.8	−9.8	72	100	0
Panama	2810118	147980	824	539.8	331.1	208.7	1	80	20
Papua New Guinea	5006703	801000	75	48.9	20.4	28.5	0	72	28
Paraguay	5358929	336000	489	343	8768.1	−8425.1	0	100	0
Peru	25230198	1913000	20132	5566.3	143.5	5422.8	1	79	21
Philippines	74178100	479000	28520	8206.7	7242	964.7	6	97	3
Poland	38654642	61600	16201	4210.1	452.4	3757.7	26	81	19
Portugal	10028200	77400	11263	6758	529.9	6228.1	15	64	36
Qatar	563710	53	294	59.3	0	59.3	554	83	17
Romania	22469358	211930	23176	877.7	2701.2	−1823.5	11	100	0
Russian Federation	146180880	4507250	76686	14534.5	12079.6	2454.9	2	97	3
Rwanda	8304804	5200	76	93	0.2	92.9	1	45	55
South Africa	42043988	50000	15306	6927.6	2558.3	4369.3	31	78	22
Saudi Arabia	20239432	2400	17320	11313.3	435.2	10878.1	722	61	39
Senegal	9279048	39400	1591	2680.4	43.6	2636.8	4	38	62
Sierra Leone	4932139	160000	380	83.2	0.6	82.6	0	82	18
Singapore	3957913	600	211	3839.2	435.2	3404.1	35	6	94
Slovakia	5395677	50100	1818	386.6	977.4	−590.8	4	100	0
Slovenia	1986239	31870	762	1062.9	21.8	1041.1	2	42	58
Somalia	8480576	13500	3298	299.7	22.7	277.1	24	92	8
Spain	39415552	111500	35635	22124.6	5621	16503.6	32	68	32
Sudan	30534126	149000	37314	561.3	1712.3	−1151.1	25	100	0
Suriname	415105	122000	665	27.5	114.5	−86.9	1	100	0
Sweden	8864128	174000	2965	737.3	1577.2	−839.9	2	100	0
Switzerland	7145332	53500	2571	2098.9	162.5	1936.5	5	57	43
Syria	15798242	46080	19947	884.5	5263.2	−4378.6	43	100	0
Tajikistan	6138744	99730	11962	46.7	83.7	−37.1	12	100	0
Tanzania	32902714	91000	1996	1211.5	283.5	928.1	2	68	32

(*continued overleaf*)

Table 1. (*Continued*)

Country	Population	Water Availability, 10^6 m^3/yr	Water Withdrawal, 10^6 m^3/yr	Gross Virtual Water Import, 10^6 m^3/yr	Gross Virtual Water Export, 10^6 m^3/yr	Net Virtual Water Import, 10^6 m^3/yr	Water Scarcity, %	Water Self-sufficiency, %	Water Dependency, %
Thailand	60275202	409944	87065	4098	50763.3	−46665.4	21	100	0
Togo	4392474	14700	166	851.6	214.8	636.8	1	21	79
Trinidad Tobago	1293248	3840	305	679.6	90.6	589	8	34	66
Tunisia	9448461	4560	2726	3925.1	57.8	3867.4	60	41	59
Turkey	64341266	231700	37519	10297.6	8244.4	2053.1	16	95	5
Turkmenistan	5057637	60860	24645	57.4	0.6	56.9	40	100	0
Uganda	21616208	66000	295	208.8	294.8	−86	0	100	0
Ukraine	49904874	139550	37523	468.9	6832.6	−6363.8	27	100	0
United Arab Emirates	2800073	150	2306	2109.4	418.1	1691.2	1538	58	42
United Kingdom	59481556	147000	9541	14204.1	15174.5	−970.4	6	100	0
Uruguay	3312629	139000	3146	821.8	3223.3	−2401.5	2	100	0
United States	278035840	3069400	479293	29264.3	180924.3	−151660	16	100	0
Uzbekistan	24394002	72210	58334	532.7	123.7	409	81	99	1
Venezuela	23705676	1233170	8368	6250.8	1325.2	4925.6	1	63	37
Viet Nam	77508750	891210	71392	153.8	18185.7	−18031.9	8	100	0
Yemen	17056736	4100	6631	1448.9	11.5	1437.4	162	82	18
Zambia	9872326	105200	1737	34.8	132.9	−98.1	2	100	0
Zimbabwe	12382668	20000	2612	115.8	633.3	−517.5	13	100	0

Source: Reference 3.

Calculation of International Virtual Water Flows and Drawing a National Virtual Water Balance

Virtual water flows between nations can be calculated by multiplying international product trade flows by their associated virtual water content. The latter depends on the water needs in the exporting country where the product is produced. Virtual water trade is thus calculated as

$$\mathrm{VWT}[n_e, n_i, p, t] = \mathrm{PT}[n_e, n_i, p, t] \times \mathrm{VWC}[n_e, p, t] \qquad (3)$$

in which VWT denotes the virtual water trade (m^3/yr) from exporting country n_e to importing country n_i in year t as a result of trade in product p. PT represents the product trade (ton/yr) from exporting country n_e to importing country n_i in year t for product p. VWC represents the virtual water content (m^3/ton) of product p in the exporting country.

The gross virtual water import (GVWI) to a country is the sum of all product-related virtual water imports. The gross virtual water export (GVWE) from a country is the sum of all product-related virtual water exports. The virtual water balance of country x for year t can be written as

$$\mathrm{NVWI}[x, t] = \mathrm{GVWI}[x, t] - \mathrm{GVWE}[x, t], \qquad (4)$$

where NVWI stands for the net virtual water import (m^3/yr) to the country. Net virtual water import to a country has either a positive or a negative sign. The latter indicates that there is net virtual water *export* from the country.

GLOBAL VIRTUAL WATER FLOWS

The global volume of international virtual water flows is roughly 1000 billion m^3/yr (7). For comparison, the total water use by crops in the world has been estimated at 5400 Gm^3/yr. These figures show that a substantial volume of water serves the global market. Countries with a large virtual water export are the United States, Canada, Thailand, Argentina, and India. Large importers of virtual water are Europe (except France), Japan, China, and Indonesia.

National virtual water balances over the period 1995–1999 are shown in Fig. 1. Full data on virtual water imports and exports are provided in Table 1. Some countries, such as Brazil, Syria, Pakistan, Tajikistan, and Uganda, have net export of virtual water over the period 1995–1999, but net import of virtual water in one or more particular years in this period. There are also countries that show the reverse, such as the Philippines, the Russian Federation, Uzbekistan, Kyrgyztan, Mongolia, Nicaragua, and Mexico. Developed countries generally have a more stable virtual water balance than the developing countries.

Countries that are relatively close to each other in terms of geography and development level can have a rather different virtual water balance. While European countries such as The Netherlands, Belgium, Germany, Spain, and Italy import virtual water in the form of crops, France exports a large amount of virtual water. In the Middle East we see that Syria has net export of virtual water related to crop trade, but Jordan and Israel have net import. In Southern Africa, Zimbabwe and Zambia had net export in the period 1995–1999, but South Africa had net import. (It should be noted that the balance of Zimbabwe has recently turned due to the recent political and economic developments.) In regions of the former Soviet Union, countries such as Kazakhstan and the Ukraine have net export of virtual water, but the Russian Federation has net import.

Net virtual water flows between thirteen world regions in the period 1995–1999 are shown in Fig. 2. The largest virtual water flows have been indicated with arrows. Regions with a significant net virtual water import are

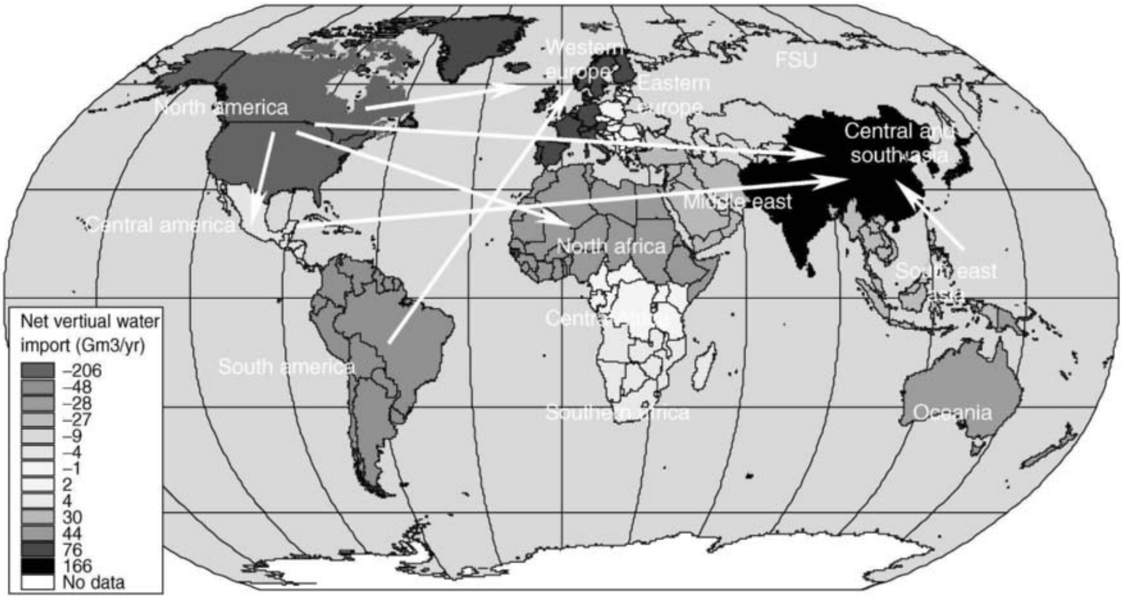

Figure 2. Virtual water balances of thirteen world regions over the period 1995–1999. The arrows show the largest net virtual water flows between regions (>20 Gm^3/yr).

Central and South Asia, Western Europe, North Africa, and the Middle East. Two other regions with net virtual water import, but less substantial, are southern Africa and Central Africa. Regions with substantial net virtual water export are North America, South America, Oceania, and Southeast Asia. Three other regions with net virtual water export, but less substantial, are the former Soviet Union, Central America, and Eastern Europe. North America is by far the biggest virtual water exporter in the world, while Central and South Asia is by far the biggest virtual water importer. Central and South Asia is the largest region in terms of population, so food demand is higher than in the other regions. This explains why the region is the biggest virtual water importer. The virtual water flows between countries within the region are also high, thus the countries within the region highly depend on both countries outside and countries within the region. Western Europe is the region with the biggest volume of internal virtual water flows. Besides, the volume appears to be rather stable over the years.

BIBLIOGRAPHY

1. Allan, J.A. (1997). *Virtual Water: A Long Term Solution for Water Short Middle Eastern Economies?* Occasional Paper 3, School of Oriental and African Studies (SOAS), University of London.
2. Hoekstra, A.Y. (1998). *Perspectives on Water: A Model-Based Exploration of the Future*. International Books, Utrecht, The Netherlands.
3. Hoekstra, A.Y. and Hung, P.Q. (2002). *Virtual Water Trade: a Quantification of Virtual Water Flows Between Nations in Relation to International Crop Trade*. Value of Water Research Report Ser. No. 11, UNESCO-IHE, Delft, The Netherlands.
4. Smith, M., Allen, R.G., Monteith, J.L., Perrier, A., Pereira, L.S., and Segeren, A. (1992). *Report on the Expert Consultation on Revision of FAO Methodologies for Crop Water Requirements*. FAO, Rome, Italy, 28–31 May 1990.
5. Allen, R.G., Pereira, L.S., Raes, D., and Smith, M. (1998). *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements*. FAO Irrigation and Drainage Paper 56, FAO, Rome, Italy.
6. Chapagain, A.K. and Hoekstra, A.Y. (2003). *Virtual Water Flows Between Nations in Relation to Trade in Livestock and Livestock Products*. Value of Water Research Report Ser. No. 13, UNESCO-IHE, Delft, The Netherlands.
7. Hoekstra, A.Y. (Ed.) (2003). Virtual water trade. In: *Proceedings of the International Expert Meeting on Virtual Water Trade*. Value of Water Research Report Ser. No. 12, UNESCO-IHE, Delft, The Netherlands.

WATER SCIENCE GLOSSARY OF TERMS

U.S. Geological Survey

Here's a list of water-related terms that might help you understand our site better. It is compiled from a number of sources and should not be considered an "official" U.S. Geological Survey water glossary. A detailed water glossary is kept by the Water Quality Association, and an **extremely detailed** water dictionary is offered by the Nevada Division of Water Planning.

A

acequia–acequias were important forms of irrigation in the development of agriculture in the American Southwest. The proliferation of cotton, pecans and green

---

chile as major agricultural staples owe their progress to the acequia system.

acid–a substance that has a pH of less than 7, which is neutral. Specifically, an acid has more free hydrogen ions (H^+) than hydroxyl ions (OH^-).

acre-foot (acre-ft)–the volume of water required to cover 1 acre of land (43,560 square feet) to a depth of 1 foot. Equal to 325,851 gallons or 1,233 cubic meters.

alkaline–sometimes water or soils contain an amount of alkali (strongly basic) substances sufficient to raise the pH value above 7.0 and be harmful to the growth of crops.

alkalinity–the capacity of water for neutralizing an acid solution.

alluvium–deposits of clay, silt, sand, gravel, or other particulate material that has been deposited by a stream or other body of running water in a streambed, on a flood plain, on a delta, or at the base of a mountain.

appropriation doctrine–the system for allocating water to private individuals used in most Western states. The doctrine of Prior Appropriation was in common use throughout the arid west as early settlers and miners began to develop the land. The prior appropriation doctrine is based on the concept of "First in Time, First in Right." The first person to take a quantity of water and put it to Beneficial Use has a higher priority of right than a subsequent user. Under drought conditions, higher priority users are satisfied before junior users receive water. Appropriative rights can be lost through nonuse; they can also be sold or transferred apart from the land. Contrasts with Riparian Water Rights.

aquaculture–farming of plants and animals that live in water, such as fish, shellfish, and algae.

aqueduct–a pipe, conduit, or channel designed to transport water from a remote source, usually by gravity.

aquifer–a geologic formation(s) that is water bearing. A geological formation or structure that stores and/or transmits water, such as to wells and springs. Use of the term is usually restricted to those water-bearing formations capable of yielding water in sufficient quantity to constitute a usable supply for people's uses.

aquifer (confined)–soil or rock below the land surface that is saturated with water. There are layers of impermeable material both above and below it and it is under pressure so that when the aquifer is penetrated by a well, the water will rise above the top of the aquifer.

aquifer (unconfined)–an aquifer whose upper water surface (water table) is at atmospheric pressure, and thus is able to rise and fall.

artesian water–ground water that is under pressure when tapped by a well and is able to rise above the level at which it is first encountered. It may or may not flow out at ground level. The pressure in such an aquifer commonly is called artesian pressure, and the formation containing artesian water is an artesian aquifer or confined aquifer. See *flowing well*

artificial recharge–an process where water is put back into ground-water storage from surface-water supplies such as irrigation, or induced infiltration from streams or wells.

B

base flow–streamflow coming from ground-water seepage into a stream.

base–a substance that has a pH of more than 7, which is neutral. A base has less free hydrogen ions (H^+) than hydroxyl ions (OH^-).

bedrock–the solid rock beneath the soil and superficial rock. A general term for solid rock that lies beneath soil, loose sediments, or other unconsolidated material.

C

capillary action–the means by which liquid moves through the porous spaces in a solid, such as soil, plant roots, and the capillary blood vessels in our bodies due to the forces of adhesion, cohesion, and surface tension. Capillary action is essential in carrying substances and nutrients from one place to another in plants and animals.

commercial water use–water used for motels, hotels, restaurants, office buildings, other commercial facilities, and institutions. Water for commercial uses comes both from public-supplied sources, such as a county water department, and self-supplied sources, such as local wells.

condensation–the process of water vapor in the air turning into liquid water. Water drops on the outside of a cold glass of water are condensed water. Condensation is the opposite process of evaporation.

consumptive use–that part of water withdrawn that is evaporated, transpired by plants, incorporated into products or crops, consumed by humans or livestock, or otherwise removed from the immediate water environment. Also referred to as water consumed.

conveyance loss–water that is lost in transit from a pipe, canal, or ditch by leakage or evaporation. Generally, the water is not available for further use; however, leakage from an irrigation ditch, for example, may percolate to a ground-water source and be available for further use.

cubic feet per second (cfs)–a rate of the flow, in streams and rivers, for example. It is equal to a volume of water one foot high and one foot wide flowing a distance of one foot in one second. One "cfs" is equal to 7.48 gallons of water flowing each second. As an example, if your car's gas tank is 2 feet by 1 foot by 1 foot (2 cubic feet), then gas flowing at a rate of 1 cubic foot/second would fill the tank in two seconds.

D

desalinization–the removal of salts from saline water to provide freshwater. This method is becoming a more popular way of providing freshwater to populations.

discharge–the volume of water that passes a given location within a given period of time. Usually expressed in cubic feet per second.

domestic water use–water used for household purposes, such as drinking, food preparation, bathing, washing clothes, dishes, and dogs, flushing toilets, and watering lawns and gardens. About 85% of domestic water is delivered to homes by a public-supply facility, such as a county water department. About 15% of the Nation's population supply their own water, mainly from wells.

drainage basin–land area where precipitation runs off into streams, rivers, lakes, and reservoirs. It is a land feature that can be identified by tracing a line along the highest elevations between two areas on a map, often a ridge. Large drainage basins, like the area that drains into the Mississippi River contain thousands of smaller drainage basins. Also called a "watershed."

drip irrigation–a common irrigation method where pipes or tubes filled with water slowly drip onto crops. Drip irrigation is a low-pressure method of irrigation and less water is lost to evaporation than high-pressure spray irrigation.

drawdown–a lowering of the ground-water surface caused by pumping.

E

effluent–water that flows from a sewage treatment plant after it has been treated.

erosion–the process in which a material is worn away by a stream of liquid (water) or air, often due to the presence of abrasive particles in the stream.

estuary–a place where fresh and salt water mix, such as a bay, salt marsh, or where a river enters an ocean.

evaporation–the process of liquid water becoming water vapor, including vaporization from water surfaces, land surfaces, and snow fields, but not from leaf surfaces. See transpiration

evapotranspiration–the sum of evaporation and transpiration.

F

flood–An overflow of water onto lands that are used or usable by man and not normally covered by water. Floods have two essential characteristics: The inundation of land is temporary; and the land is adjacent to and inundated by overflow from a river, stream, lake, or ocean.

flood, 100-year–A 100-year flood does not refer to a flood that occurs once every 100 years, but to a flood level with a 1 percent chance of being equaled or exceeded in any given year.

flood plain–a strip of relatively flat and normally dry land alongside a stream, river, or lake that is covered by water during a flood.

flood stage–The elevation at which overflow of the natural banks of a stream or body of water begins in the reach or area in which the elevation is measured.

flowing well/spring–a well or spring that taps ground water under pressure so that water rises without pumping. If the water rises above the surface, it is known as a flowing well.

freshwater, fresh water–water that contains less than 1,000 milligrams per liter (mg/L) of dissolved solids; generally, more than 500 mg/L of dissolved solids is undesirable for drinking and many industrial uses.

G

gage height–the height of the water surface above the gage datum (zero point). Gage height is often used interchangeably with the more general term, stage, although gage height is more appropriate when used with a gage reading.

gaging station–a site on a stream, lake, reservoir or other body of water where observations and hydrologic data are obtained. The U.S. Geological Survey measures stream discharge at gaging stations.

geyser–a geothermal feature of the Earth where there is an opening in the surface that contains superheated water that periodically erupts in a shower of water and steam.

giardiasis–a disease that results from an infection by the protozoan parasite Giardia Intestinalis, caused by drinking water that is either not filtered or not chlorinated. The disorder is more prevalent in children than in adults and is characterized by abdominal discomfort, nausea, and alternating constipation and diarrhea.

glacier–a huge mass of ice, formed on land by the compaction and recrystallization of snow, that moves very slowly downslope or outward due to its own weight.

greywater–wastewater from clothes washing machines, showers, bathtubs, hand washing, lavatories and sinks.

ground water–(1) water that flows or seeps downward and saturates soil or rock, supplying springs and wells. The upper surface of the saturate zone is called the water table. (2) Water stored underground in rock crevices and in the pores of geologic materials that make up the Earth's crust.

ground water, confined–ground water under pressure significantly greater than atmospheric, with its upper limit the bottom of a bed with hydraulic conductivity distinctly lower than that of the material in which the confined water occurs.

ground-water recharge–inflow of water to a ground-water reservoir from the surface. Infiltration of precipitation and its movement to the water table is one form of natural recharge. Also, the volume of water added by this process.

ground water, unconfined–water in an aquifer that has a water table that is exposed to the atmosphere.

H

hardness–a water-quality indication of the concentration of alkaline salts in water, mainly calcium and magnesium. If the water you use is "hard" then more soap, detergent or shampoo is necessary to raise a lather.

headwater(s)–(1) the source and upper reaches of a stream; also the upper reaches of a reservoir; (2) the water upstream from a structure or point on a stream; (3) the small streams that come together to form a river. Also may be thought of as any and all parts of a river basin except the mainstream river and main tributaries.

hydroelectric power water use–the use of water in the generation of electricity at plants where the turbine generators are driven by falling water.

hydrologic cycle–the cyclic transfer of water vapor from the Earth's surface via evapotranspiration into the atmosphere, from the atmosphere via precipitation back to earth, and through runoff into streams, rivers, and lakes, and ultimately into the oceans.

I

impermeable layer–a layer of solid material, such as rock or clay, which does not allow water to pass through.

industrial water use–water used for industrial purposes in such industries as steel, chemical, paper, and petroleum refining. Nationally, water for industrial uses comes mainly (80%) from self-supplied sources, such as a local wells or withdrawal points in a river, but some water comes from public-supplied sources, such as the county/city water department.

infiltration–flow of water from the land surface into the subsurface.

injection well–refers to a well constructed for the purpose of injecting treated wastewater directly into the ground. Wastewater is generally forced (pumped) into the well for dispersal or storage into a designated aquifer. Injection wells are generally drilled into aquifers that don't deliver drinking water, unused aquifers, or below freshwater levels.

irrigation–the controlled application of water for agricultural purposes through manmade systems to supply water requirements not satisfied by rainfall. Here's a quick look at some types of irrigation systems.

irrigation water use–water application on lands to assist in the growing of crops and pastures or to maintain vegetative growth in recreational lands, such as parks and golf courses.

K

kilogram–one thousand grams.

kilowatt-hour (KWH)–a power demand of 1,000 watts for one hour. Power company utility rates are typically expressed in cents per kilowatt-hour.

L

leaching–the process by which soluble materials in the soil, such as salts, nutrients, pesticide chemicals or contaminants, are washed into a lower layer of soil or are dissolved and carried away by water.

lentic waters–ponds or lakes (standing water).

levee–a natural or manmade earthen barrier along the edge of a stream, lake, or river. Land alongside rivers can be protected from flooding by levees.

livestock water use–water used for livestock watering, feed lots, dairy operations, fish farming, and other on-farm needs.

lotic waters–flowing waters, as in streams and rivers.

M

maximum contaminant level (MCL)–the designation given by the U.S. Environmental Protection Agency (EPA) to water-quality standards promulgated under the Safe Drinking Water Act. The MCL is the greatest amount of a contaminant that can be present in drinking water without causing a risk to human health.

milligram (mg)–One-thousandth of a gram.

milligrams per liter (mg/l)–a unit of the concentration of a constituent in water or wastewater. It represents 0.001 gram of a constituent in 1 liter of water. It is approximately equal to one part per million (PPM).

million gallons per day (Mgd)–a rate of flow of water equal to 133,680.56 cubic feet per day, or 1.5472 cubic feet per second, or 3.0689 acre-feet per day. A flow of one million gallons per day for one year equals 1,120 acre-feet (365 million gallons).

mining water use–water use during quarrying rocks and extracting minerals from the land.

municipal water system–a water system that has at least five service connections or which regularly serves 25 individuals for 60 days; also called a public water system

N

nephelometric turbidity unit (NTU)–unit of measure for the turbidity of water. Essentially, a measure of the cloudiness of water as measured by a nephelometer. Turbidity is based on the amount of light that is reflected off particles in the water.

non-point source (NPS) pollution–pollution discharged over a wide land area, not from one specific location. These are forms of diffuse pollution caused by sediment, nutrients, organic and toxic substances originating from land-use activities, which are carried to lakes and streams by surface runoff. Non-point source pollution is contamination that occurs when rainwater, snowmelt, or irrigation washes off plowed fields, city streets, or suburban backyards. As this runoff moves across the land surface, it picks up soil particles and pollutants, such as nutrients and pesticides.

O

organic matter–plant and animal residues, or substances made by living organisms. All are based upon carbon compounds.

osmosis–the movement of water molecules through a thin membrane. The osmosis process occurs in our bodies and is also one method of desalinizing saline water.

outfall–the place where a sewer, drain, or stream discharges; the outlet or structure through which reclaimed water or treated effluent is finally discharged to a receiving water body.

oxygen demand–the need for molecular oxygen to meet the needs of biological and chemical processes in water. Even though very little oxygen will dissolve in water, it is extremely important in biological and chemical processes.

P

pH–a measure of the relative acidity or alkalinity of water. Water with a pH of 7 is neutral; lower pH levels indicate increasing acidity, while pH levels higher than 7 indicate increasingly basic solutions.

View a diagram about pH.

particle size–the diameter, in millimeters, of suspended sediment or bed material. Particle-size classifications are:

[1] *Clay* 0.00024–0.004 millimeters (mm);
[2] *Silt* 0.004–0.062 mm;
[3] *Sand* 0.062–2.0 mm; and
[4] *Gravel* 2.0–64.0 mm.

parts per billion–the number of "parts" by weight of a substance per billion parts of water. Used to measure extremely small concentrations.

parts per million–the number of "parts" by weight of a substance per million parts of water. This unit is commonly used to represent pollutant concentrations.

pathogen–a disease-producing agent; usually applied to a living organism. Generally, any viruses, bacteria, or fungi that cause disease.

peak flow–the maximum instantaneous discharge of a stream or river at a given location. It usually occurs at or near the time of maximum stage.

per capita use–the average amount of water used per person during a standard time period, generally per day.

percolation–(1) The movement of water through the openings in rock or soil. (2) the entrance of a portion of the streamflow into the channel materials to contribute to ground water replenishment.

permeability–the ability of a material to allow the passage of a liquid, such as water through rocks. Permeable materials, such as gravel and sand, allow water to move quickly through them, whereas unpermeable material, such as clay, don't allow water to flow freely.

point-source pollution–water pollution coming from a single point, such as a sewage-outflow pipe.

polychlorinated biphenyls (PCBs)–a group of synthetic, toxic industrial chemical compounds once used in making paint and electrical transformers, which are chemically inert and not biodegradable. PCBs were frequently found in industrial wastes, and subsequently found their way into surface and ground waters. As a result of their persistence, they tend to accumulate in the environment. In terms of streams and rivers, PCBs are drawn to sediment, to which they attach and can remain virtually indefinitely. Although virtually banned in 1979 with the passage of the Toxic Substances Control Act, they continue to appear in the flesh of fish and other animals.

porosity–a measure of the water-bearing capacity of subsurface rock. With respect to water movement, it is not just the total magnitude of porosity that is important, but the size of the voids and the extent to which they are interconnected, as the pores in a formation may be open, or interconnected, or closed and isolated. For example, clay may have a very high porosity with respect to potential water content, but it constitutes a poor medium as an aquifer because the pores are usually so small.

potable water–water of a quality suitable for drinking.

precipitation–rain, snow, hail, sleet, dew, and frost.

primary wastewater treatment–the first stage of the wastewater-treatment process where mechanical methods, such as filters and scrapers, are used to remove pollutants. Solid material in sewage also settles out in this process.

prior appropriation doctrine–the system for allocating water to private individuals used in most Western states. The doctrine of Prior Appropriation was in common use throughout the arid West as early settlers and miners began to develop the land. The prior appropriation doctrine is based on the concept of "First in Time, First in Right." The first person to take a quantity of water and put it to beneficial use has a higher priority of right than a subsequent user. The rights can be lost through nonuse; they can also be sold or transferred apart from the land. Contrasts with riparian water rights.

public supply–water withdrawn by public governments and agencies, such as a county water department, and by private companies that is then delivered to users. Public suppliers provide water for domestic, commercial, thermoelectric power, industrial, and public water users. Most people's household water is delivered by a public water supplier. The systems have at least 15 service connections (such as households, businesses, or schools) or regularly serve at least 25 individuals daily for at least 60 days out of the year.

public water use–water supplied from a public-water supply and used for such purposes as firefighting, street washing, and municipal parks and swimming pools.

R

rating curve–A drawn curve showing the relation between gage height and discharge of a stream at a given gaging station.

recharge–water added to an aquifer. For instance, rainfall that seeps into the ground.

reclaimed wastewater–treated wastewater that can be used for beneficial purposes, such as irrigating certain plants.

recycled water–water that is used more than one time before it passes back into the natural hydrologic system.

reservoir–a pond, lake, or basin, either natural or artificial, for the storage, regulation, and control of water.

return flow–(1) That part of a diverted flow that is not consumptively used and returned to its original source or another body of water. (2) (Irrigation) Drainage water from irrigated farmlands that re-enters the water system to be used further downstream.

returnflow (irrigation)–irrigation water that is applied to an area and which is not consumed in evaporation or transpiration and returns to a surface stream or aquifer.

reverse osmosis–(1) (Desalination) The process of removing salts from water using a membrane. With reverse osmosis, the product water passes through a fine membrane that the salts are unable to pass through, while the salt waste (brine) is removed and disposed. This process differs from electrodialysis, where the salts are extracted from the feedwater by using a membrane with an electrical current to separate the ions. The positive ions go through one membrane, while the negative ions flow through a different membrane, leaving the end product of freshwater. (2) (Water Quality) An advanced method of water or wastewater treatment that relies on a semi-permeable membrane to separate waters from pollutants. An external force is used to reverse the normal osmotic process resulting in the solvent moving from a solution of higher concentration to one of lower concentration.

riparian water rights–the rights of an owner whose land abuts water. They differ from state to state and often depend on whether the water is a river, lake, or ocean. The doctrine of riparian rights is an old one, having its origins in English common law. Specifically, persons who own land adjacent to a stream have the right to make reasonable use of the stream. Riparian users of a stream share the streamflow among themselves, and the concept of priority of use (Prior Appropriation Doctrine) is not applicable. Riparian rights cannot be sold or transferred for use on nonriparian land.

river–A natural stream of water of considerable volume, larger than a brook or creek.

runoff–(1) That part of the precipitation, snow melt, or irrigation water that appears in uncontrolled surface streams, rivers, drains or sewers. Runoff may be classified according to speed of appearance after rainfall or melting snow as direct runoff or base runoff, and according to source as surface runoff, storm interflow, or ground-water runoff. (2) The total discharge described in (1), above, during a specified period of time. (3) Also defined as the depth to which a drainage area would be covered if all of the runoff for a given period of time were uniformly distributed over it.

S

saline water–water that contains significant amounts of dissolved solids.

Here are our parameters for saline water:

Fresh water. Less than 1,000 parts per million (ppm)
Slightly saline water. From 1,000 ppm to 3,000 ppm
Moderately saline water. From 3,000 ppm to 10,000 ppm
Highly saline water. From 10,000 ppm to 35,000 ppm

secondary wastewater treatment–treatment (following primary wastewater treatment) involving the biological process of reducing suspended, colloidal, and dissolved organic matter in effluent from primary treatment systems and which generally removes 80 to 95 percent of the Biochemical Oxygen Demand (BOD) and suspended matter. Secondary wastewater treatment may be accomplished by biological or chemical-physical methods. Activated sludge and trickling filters are two of the most common means of secondary treatment. It is accomplished by bringing together waste, bacteria, and oxygen in trickling filters or in the activated sludge process. This treatment removes floating and settleable solids and about 90 percent of the oxygen-demanding substances and suspended solids. Disinfection is the final stage of secondary treatment.

sediment–usually applied to material in suspension in water or recently deposited from suspension. In the plural the word is applied to all kinds of deposits from the waters of streams, lakes, or seas.

sedimentary rock–rock formed of sediment, and specifically: (1) sandstone and shale, formed of fragments of other rock transported from their sources and deposited in water; and (2) rocks formed by or from secretions of organisms, such as most limestone. Many sedimentary rocks show distinct layering, which is the result of different types of sediment being deposited in succession.

sedimentation tanks–wastewater tanks in which floating wastes are skimmed off and settled solids are removed for disposal.

self-supplied water–water withdrawn from a surface- or ground-water source by a user rather than being obtained from a public supply. An example would be homeowners getting their water from their own well.

seepage–(1) The slow movement of water through small cracks, pores, Interstices, etc., of a material into or out of a body of surface or subsurface water. (2) The loss of water by infiltration into the soil from a canal, ditches, laterals, watercourse, reservoir, storage facilities, or other body of water, or from a field.

septic tank–a tank used to detain domestic wastes to allow the settling of solids prior to distribution to a leach field for soil absorption. Septic tanks are used when a sewer line is not available to carry them to a treatment plant. A settling tank in which settled sludge is in immediate contact with sewage flowing through the tank, and wherein solids are decomposed by anaerobic bacterial action.

settling pond (water quality)–an open lagoon into which wastewater contaminated with solid pollutants is placed and allowed to stand. The solid pollutants suspended in the water sink to the bottom of the lagoon and the liquid is allowed to overflow out of the enclosure.

sewage treatment plant–a facility designed to receive the wastewater from domestic sources and to remove materials that damage water quality and threaten public health and safety when discharged into receiving streams or bodies of water. The substances removed are classified into four basic areas:

[1] greases and fats;
[2] solids from human waste and other sources;
[3] dissolved pollutants from human waste and decomposition products; and
[4] dangerous microorganisms.

Most facilities employ a combination of mechanical removal steps and bacterial decomposition to achieve the desired results. Chlorine is often added to discharges from the plants to reduce the danger of spreading disease by the release of pathogenic bacteria.

sewer–a system of underground pipes that collect and deliver wastewater to treatment facilities or streams.

sinkhole–a depression in the Earth's surface caused by dissolving of underlying limestone, salt, or gypsum.

Drainage is provided through underground channels that may be enlarged by the collapse of a cavern roof.

solute–a substance that is dissolved in another substance, thus forming a solution.

solution–a mixture of a solvent and a solute. In some solutions, such as sugar water, the substances mix so thoroughly that the solute cannot be seen. But in other solutions, such as water mixed with dye, the solution is visibly changed.

solvent–a substance that dissolves other substances, thus forming a solution. Water dissolves more substances than any other, and is known as the "universal solvent."

specific conductance–a measure of the ability of water to conduct an electrical current as measured using a 1-cm cell and expressed in units of electrical conductance, that is, Siemens per centimeter at 25 degrees Celsius. Specific conductance can be used for approximating the total dissolved solids content of water by testing its capacity to carry an electrical current. In water quality, specific conductance is used in ground water monitoring as an indication of the presence of ions of chemical substances that may have been released by a leaking landfill or other waste storage or disposal facility. A higher specific conductance in water drawn from downgradient wells when compared to upgradient wells indicates possible contamination from the facility.

spray irrigation–a common irrigation method where water is shot from high-pressure sprayers onto crops. Because water is shot high into the air onto crops, some water is lost to evaporation.

storm sewer–a sewer that carries only surface runoff, street wash, and snow melt from the land. In a separate sewer system, storm sewers are completely separate from those that carry domestic and commercial wastewater (sanitary sewers).

stream–a general term for a body of flowing water; natural water course containing water at least part of the year. In hydrology, it is generally applied to the water flowing in a natural channel as distinct from a canal.

streamflow–the water discharge that occurs in a natural channel. A more general term than runoff, streamflow may be applied to discharge whether or not it is affected by diversion or regulation.

subsidence–a dropping of the land surface as a result of ground water being pumped. Cracks and fissures can appear in the land. Subsidence is virtually an irreversible process.

surface tension–the attraction of molecules to each other on a liquid's surface. Thus, a barrier is created between the air and the liquid.

surface water–water that is on the Earth's surface, such as in a stream, river, lake, or reservoir.

suspended sediment–very fine soil particles that remain in suspension in water for a considerable period of time without contact with the bottom. Such material remains in suspension due to the upward components of turbulence and currents and/or by suspension.

suspended-sediment concentration–the ratio of the mass of dry sediment in a water-sediment mixture to the mass of the water-sediment mixture. Typically expressed in milligrams of dry sediment per liter of water-sediment mixture.

suspended-sediment discharge–the quantity of suspended sediment passing a point in a stream over a specified period of time. When expressed in tons per day, it is computed by multiplying water discharge (in cubic feet per second) by the suspended-sediment concentration (in milligrams per liter) and by the factor 0.0027.

suspended solids–solids that are not in true solution and that can be removed by filtration. Such suspended solids usually contribute directly to turbidity. Defined in waste management, these are small particles of solid pollutants that resist separation by conventional methods.

T

tertiary wastewater treatment–selected biological, physical, and chemical separation processes to remove organic and inorganic substances that resist conventional treatment practices; the additional treatment of effluent beyond that of primary and secondary treatment methods to obtain a very high quality of effluent. The complete wastewater treatment process typically involves a three-phase process: (1) First, in the primary wastewater treatment process, which incorporates physical aspects, untreated water is passed through a series of screens to remove solid wastes; (2) Second, in the secondary wastewater treatment process, typically involving biological and chemical processes, screened wastewater is then passed a series of holding and aeration tanks and ponds; and (3) Third, the tertiary wastewater treatment process consists of flocculation basins, clarifiers, filters, and chlorine basins or ozone or ultraviolet radiation processes.

thermal pollution–a reduction in water quality caused by increasing its temperature, often due to disposal of waste heat from industrial or power generation processes. Thermally polluted water can harm the environment because plants and animals can have a hard time adapting to it.

thermoelectric power water use–water used in the process of the generation of thermoelectric power. Power plants that burn coal and oil are examples of thermoelectric-power facilities.

transmissibility (ground water)–the capacity of a rock to transmit water under pressure. The coefficient of transmissibility is the rate of flow of water, at the prevailing water temperature, in gallons per day, through a vertical strip of the aquifer one foot wide, extending the full saturated height of the aquifer under a hydraulic gradient of 100-percent. A hydraulic gradient of 100-percent means a one foot drop in head in one foot of flow distance.

transpiration–process by which water that is absorbed by plants, usually through the roots, is evaporated into the atmosphere from the plant surface, such as leaf pores. See evapotranspiration.

Tributary–a smaller river or stream that flows into a larger river or stream. Usually, a number of smaller tributaries merge to form a river.

turbidity–the amount of solid particles that are suspended in water and that cause light rays shining through the water to scatter. Thus, turbidity makes the water cloudy or even opaque in extreme cases. Turbidity is measured in nephelometric turbidity units (NTU).

U

unsaturated zone–the zone immediately below the land surface where the pores contain both water and air, but are not totally saturated with water. These zones differ from an aquifer, where the pores are saturated with water.

W

wastewater–water that has been used in homes, industries, and businesses that is not for reuse unless it is treated.

wastewater-treatment return flow–water returned to the environment by wastewater-treatment facilities.

water cycle–the circuit of water movement from the oceans to the atmosphere and to the Earth and return to the atmosphere through various stages or processes such as precipitation, interception, runoff, infiltration, percolation, storage, evaporation, and transportation.

water quality–a term used to describe the chemical, physical, and biological characteristics of water, usually in respect to its suitability for a particular purpose.

water table–the top of the water surface in the saturated part of an aquifer.

water use–water that is used for a specific purpose, such as for domestic use, irrigation, or industrial processing. Water use pertains to human's interaction with and influence on the hydrologic cycle, and includes elements, such as water withdrawal from surface- and ground-water sources, water delivery to homes and businesses, consumptive use of water, water released from wastewater-treatment plants, water returned to the environment, and instream uses, such as using water to produce hydroelectric power.

watershed–the land area that drains water to a particular stream, river, or lake. It is a land feature that can be identified by tracing a line along the highest elevations between two areas on a map, often a ridge. Large watersheds, like the Mississippi River basin contain thousands of smaller watersheds.

watt-hour (Wh)–an electrical energy unit of measure equal to one watt of power supplied to, or taken from, an electrical circuit steadily for one hour.

well (water)–an artificial excavation put down by any method for the purposes of withdrawing water from the underground aquifers. A bored, drilled, or driven shaft, or a dug hole whose depth is greater than the largest surface dimension and whose purpose is to reach underground water supplies or oil, or to store or bury fluids below ground.

withdrawal–water removed from a ground- or surface-water source for use.

X

xeriscaping–a method of landscaping that uses plants that are well adapted to the local area and are drought-resistant. Xeriscaping is becoming more popular as a way of saving water at home.

More on xeriscaping: Texas Agricultural Extension Service

More on xeriscaping: Texas Natural Resource Center

Y

yield–mass per unit time per unit area

Some of this information is courtesy of the Nevada Division of Water Planning.

HARVESTING RAINWATER

Ganesh B. Keremane
University of South Australia
Adelaide, Australia

INTRODUCTION

Water is the genesis and continuing source of all life on the earth, and we all completely depend on this renewable source. According to the United Nation's World Water Development Report (1), "Mankind's most serious challenge in the 21st century might not be war or hunger or disease or even the collapse of civic order; it may be the lack of fresh water," which justifies the present scenario that freshwater is a scarce resource. Although it is the commonest and the most critical stuff on the earth, water has been always an undervalued resource. There are more than 1.4 billion cubic kilometers (km^3) of water—enough to give every man, woman, and child more than 230 million cubic meters (m^3) each if we were to divide it evenly (2). However, more than 98% of the world's water is saltwater, and for our basic vital needs, we have to depend on freshwater, most of which is locked in the polar ice caps. This leaves less than 1% of the earth's freshwater to be accessed in lakes, rivers, and groundwater aquifers. The major portion of the water diverted for human needs is drawn from this renewable, readily accessible part of the world's freshwater resources, thus making it a scarce resource.

NEED FOR WATER HARVESTING

The prevailing water scarcity experienced by most of the water-scarce countries (arid and semiarid) of the world can be attributed to high water stress and unsustainable rates of withdrawal. Dupont (3) attributes this widening gap between freshwater demand and supply to three main factors: a limited global supply of freshwater, increased water consumption, and poor water management. The global supply of freshwater is limited and unevenly distributed as against the demand for freshwater, which increased sixfold between 1900 and 1995 (4,5). Moreover, the exponential increase in the consumption of water from population growth, rapid industrial development, and the expansion of irrigated farmland (6) contributes to added pressure on the water supply. Of these, the trend toward irrigating farmland, which provides one-third of the world's harvest from just 17% of its cropland, is responsible for much of the pressure on the water supply and has risen from 50 million hectares to 250 million hectares since 1990 (7). Poor water management and conservation practices have contributed to the decline in the quality and availability of water in addition to overlogging and deforestation that are destroying water tables and causing siltation and salinization. Wasteful and environmentally destructive farming techniques, overuse of pesticides and chemical fertilizers, and expansion of urban areas are other factors that have further reduced freshwater supplies. The statistics provided in the United Nations World Water Development Report (1) support this:

- World population is likely to increase by half as much again, to about 9.3 billion by 2050.
- One liter of wastewater is sufficient to pollute about eight liters of freshwater, and every day about 2 million tons of wastes are dumped into rivers, lakes, and streams.
- As of today, it is estimated that across the world there are about 12,000 cubic kilometers of wastewater, which is more than the total amount in the world's 10 largest river basins at any given moment.
- At present, 48% of the earth's population lives in towns and cities; this will increase to 60% by 2030.

For these reasons, water has become a scarce resource in many parts of the world. About 80 countries containing 40% of the world's population suffer from serious shortages, mostly in the developing world (4). Despite widely available evidence of the crisis, ways to help ensure adequate water supplies for household, agricultural, and other uses are available to farms and communities. Rudimentary conservation measures, application of new technology, and sensible conservation practices could do much to stabilize water withdrawals (3). Water harvesting is one such technology.

WATER HARVESTING—TO MAKE EVERY DROP COUNT

Providing irrigation to crops in hot, dry countries (arid and semiarid) accounts for 70% of the water use in the world (8). Further, the rise in land pressure has resulted in using more and more marginal areas for agriculture. Much of this land is located in the arid and semiarid belts characterized by scanty, uneven distribution of rainfall, where much of the precious water is soon lost as surface runoff. Recent droughts have highlighted the risks from the failure of monsoons to all life on the earth. Irrigation may be the most obvious response to drought but has proved costly and can benefit only a fortunate few (9). Lately this has gained popularity for an old technology in a new way, and that technology is water harvesting.

Water harvesting is no new concept for humankind. Though its importance in mitigating the water scarcity problem has been highlighted much in the recent past, it traces its history to biblical times. Extensive water harvesting apparatus existed 4000 years ago in Palestine and Greece (10). Critchley and Chapman (9), while discussing the history of water harvesting systems, mentioned that the earliest evidence of the use of water harvesting is the well-publicized systems used by the people of the Negev Desert, where the hillsides were cleared of vegetation and smoothed to provide as much runoff as possible; the water was then channeled in contour ditches to agricultural fields and/or to cisterns. In the New World, about 400–700 years ago, people living in Colarado (North America) and Peru (South America) also employed relatively simple methods of water harvesting for irrigation (11). The practice of collecting water from rooftops is also a very ancient type of water harvesting, practiced since the earliest times up to the present. However, renewed interest in the technology of water harvesting arose in the 1950s in Israel, Australia, and the United States. Since then, this technology has developed manifold through experimentation and demonstration projects.

WHAT IS WATER HARVESTING?

Prinz (12) opines, "There is no generally accepted definition of water harvesting." But the term water harvesting in its crudest form can be understood as, *catching rain as it falls*. However, many experts have defined water harvesting and some of them are mentioned for reference: Critchley and Chapman (9), in their manual on water harvesting, define water harvesting in its broadest sense as the collection of runoff for its productive use. Thames (11) defines water harvesting systems as "artificial methods whereby precipitation can be collected and stored until it is beneficially used. The system includes a catchment area and a storage facility for the harvested water, unless the water is to be immediately concentrated in the soil profile." "Water harvesting is collecting and using precipitation from a catchment surface" (10).

Though defined differently, it is clear from the definitions that water harvesting is the practice of collecting and storing water from various sources for beneficial use; the catchment area, runoff, and the storage area are the central elements or components of any water harvesting system. These elements are briefly explained in Box 1.

Box 1: Components of Water Harvesting[a]

Catchment Area. An area that is reasonably impermeable to water which can be used to produce runoff. It may include natural surfaces such as rock outcrops, surfaces developed for other purposes such as paved highways or roof tops, and/or surfaces prepared with minimal cost and effort such as those cleared of vegetation.

Runoff. That portion of precipitation which is not intercepted by vegetation, absorbed by the land surface, or evaporated, and thus flows overland into a depression, stream, lake, or ocean. Runoff may be classified as direct runoff or base runoff according to the speed of appearance after rainfall or melting snow, and as surface runoff, storm interflow, or groundwater runoff, according to its source.

Storage Area. Based on where the water is stored, the storage area can be the soil profile, excavated ponds, or any underground structure.

1. *The Soil Profile*: Simple arrangements are made to direct water from hillsides onto cultivated areas to store the water immediately in the soil for plant use. Whether sufficient water can be stored to offset a prolonged drought is a disadvantage.
2. *Excavated Ponds*: Ponds excavated in the ground surface are used for storing large quantities of water. These are often economical means of storage, but evaporation and seepage are serious drawbacks.
3. *Underground Structures*: Can be in the form of aquifers (an underground layer of porous rock, sand, or gravel containing large amounts of water) or cisterns (man-made underground construction to store water). In the light of the disadvantages/drawbacks of the other two types, it is an interesting alternative.

^{a}*Source*: Reference 13.

FORMS OF WATER HARVESTING

As mentioned before, water harvesting has been practiced for ages and is still relevant worldwide. A great number of forms exist with various names. Prinz and Singh (14) group water harvesting techniques into three broad groups: rainwater harvesting, floodwater harvesting, and groundwater harvesting. A brief description of these techniques along with subtypes is given below.

Rainwater Harvesting

Defined as a method for inducing, collecting, storing, and conserving local surface runoff for agriculture or other beneficial use (15). This can be further grouped into three subtypes:

- Water collected from roof tops, courtyards, and similar compacted or treated surfaces that is used for domestic purposes or garden crops.
- Microcatchment water harvesting is a method of collecting surface runoff from a small catchment area (catchment length usually between 1 and 30 meters) and storing it in the root zone of an adjacent infiltration basin. The basin is planted with a tree, a bush, or with annual crops.
- Macrocatchment water harvesting, also called harvesting from external catchments, is done where runoff from a hillslope catchment (catchment length usually between 30 and 200 meters) is conveyed to a cropping area located at the hill foot on flat terrain.

Floodwater Harvesting

Defined as the collection and storage of creek flow for irrigation use. According to Critchley and Chapman (9), floodwater harvesting is also known as “large catchment water harvesting” or “spate irrigation” and may be further classified as

- *Floodwater harvesting within a steambed* in which case the water flow is dammed and, as a result, inundates the valley bottom of a floodplain. The water is forced to infiltrate and the wetted area can be used for agriculture or pasture improvement.
- *Floodwater diversion* wherein the wadi water is forced to leave its natural course and is conveyed to nearby fields.

Groundwater Harvesting

A rather new term employed to cover traditional as well as unconventional ways of groundwater extraction. Subsurface dams and sand storage dams are some fine examples of groundwater harvesting. They obstruct the flow of temporary or short-lived streams in a riverbed, and store the water in the sediment below the ground surface, which is useful for aquifer recharge.

WATER HARVESTING STRUCTURES

All the previously mentioned forms of water harvesting are done with the aid of different permanent or temporary structures. There are several structures by which water can be harvested for aquifer recharge. A few most widely used and cheapest structures are described below. Barton (16) ranks them among the most popular methods for two main reasons: low construction cost and ease of operation and maintenance.

- *Contour trenches* are excavations parallel to contours on slopes to conserve water and prevent soil erosion and thus help water to infiltrate.
- *Gully plugs* are stone barriers built across gullies and deep rills that trap sediments eroded from higher up a slope and impound runoff, encouraging infiltration.
- *Check dams* are temporary structures constructed from locally available material such as brushwood, loose rock, or woven wire to impede the soil and water removed from the catchment. The impeded water collects behind the dam and infiltrates the soil, recharging the aquifer. The cost of construction ranges from US $200–400 (for temporary structures) to US $1000–2000 (for permanent structures), depending on the material used, the size of the gully, and the height of the dam (16).

- *A percolation tank* is a dam built on permeable ground so that floodwater is held back long enough to percolate into the ground, hence, the name "percolation tank." This is constructed by excavating a depression to form a small reservoir or by constructing an embankment in a natural ravine or gully. The construction cost is estimated at approximately US $5,000–10,000.
- *Groundwater dams* are structures that intercept or obstruct the natural flow of groundwater. These are often built within riverbeds to obstruct and detain groundwater flow so as to sustain the storage capacity of the aquifer and meet demands during high periods. These may be either a subsurface dam that is constructed within the aquifer itself or a sand storage/silt trapping dam that is constructed above ground to trap sand transported with floodwater.
- *Rooftop harvesting* is a method that can be adopted by individuals. Several techniques to harvest rainwater from the roof tops are in use: abandoned dug well, abandoned/running hand pump, gravity head recharge, and recharge pit (16). Though in use mainly to harvest water for household and domestic use, there are evidences of groundwater recharge using this technique. CGWB (10) has observed additional recharge to groundwater by dug wells to the extent of 6.6 TCM with a benefited area of 1.3 hectare in the Jalgaon district of Maharashtra, and, also, a rise in water level from 1.43 to 2.15 m has been recorded in Delhi from the adoption of rooftop harvesting techniques.

CHOICE OF WATER HARVESTING AREA AND TECHNIQUES

Planning is an important process involved in implementing any water harvesting project by an individual or community. It starts with the identification of the area for water harvesting, followed by selection of the techniques/forms/structures. While doing so, it is of utmost importance to consider the following parameters (11,12).

- *Amount and Seasonal Distribution of Rainfall*: Knowledge of the intensity and distribution of rainfall for a given area is necessary for designing a water harvesting system.
- *Topography*: Knowledge of the landform along with the slope is the next important parameter to determine the system for harvesting water and also the structures to be included.
- *Soil Type and Depth*: Important for designing a water harvesting system, the characteristics of soils determine the water movement and the quantity of water that can be stored and also influence the rainfall runoff process.
- *Provisions for Maintenance*: A maintenance program, even when the water collected is not being used (off season), is a must. Failure to provide for maintenance will result in early failure of the system.
- *Acceptance by the Local Community*: Last but not the least, to be successful water harvesting requires local capacity building, cooperation, and extensive participation. The users must believe that the system proposed is the best for their needs. Hence, before implementing any system, acceptance by the end or final user is a must to ensure the sustainability of the system. Only then can it be a success.

CONCLUSIONS

The water harvesting concept used to be more applicable to arid and semiarid regions where rainfall is either not sufficient or precipitation is erratic. But, lately, due to ever increasing population, growing urbanization, unsustainable withdrawal of groundwater accompanied by erratic monsoons, other regions of the world are also facing the problem of water scarcity. Hence, water harvesting is being looked into as an effective method to mitigate this problem of water scarcity all over the world because it offers a method of effectively developing scarce water resources by concentrating the rainfall/runoff in parts of the total area. These systems have the potential to increase the productivity of arable and grazing land by increasing the yields and by reducing the risk of crop failure. As studied by Kakade et al. (17), they also facilitate afforestation, fruit tree planting, and agroforestry. Using harvested rainwater helps decrease the use of other valuable water sources such as groundwater, and helps in groundwater recharge (18). It is also a relatively inexpensive method of water supply that can be adapted to the resources and needs of the rural poor. As such, there is no "the best" system of water harvesting. However, the system can be designed to fit best within the constraints of a given location taking into account technical, social, physical, and economic factors. Thus, learning from the past experiences and adopting new technological developments that take place, any water harvesting system can be considered a successful or an ideal system if it is

- technically sound, properly designed and maintained,
- economically feasible for the users, and
- capable of being integrated into the social traditions and abilities of the users.

BIBLIOGRAPHY

1. United Nations World Water Development Report. (2003). Accessed at http://www.unesco.org/water/wwap/wwdr/table_contents.shtml.
2. Asian Development Bank. (1999). *Annual Report—Special theme-Water in the 21st Century*. Philippines. [http://www.adb.org/].
3. Dupont, A. (2003). Will there be water wars? *Dev. Bull.* **63**: November 16–20.
4. United Nations Environmental Programme. (1991). *Freshwater Pollution*. UNEP, Nairobi.
5. Gleick, P.H. (1998). *The Worlds Water 1998–1999: The Biennial Report on Freshwater Resources*. Island Press, Washington, DC, pp. 40–47.

6. Wallensteen, P. and Swain, A. (1997). *International Freshwater Resources Conflict or Cooperation?* Stockholm Environment Institute, Stockholm, Sweden.
7. Gleick, P.H. (1993). An introduction to global fresh water issues. In: *Water in Crisis: A Guide to the World's Fresh Water Resources*. P.H. Gleick (Ed.). Oxford University Press, New York, p. 6.
8. McCarthy, M. (2003). *Water scarcity could affect billions: Is this the biggest crisis of all?* Accessed at http://www.commondreams.org/headlines03/0305-05.htm.
9. Critchley, W. and Chapman, C. (1991). *Water Harvesting—A Manual for the Design and Construction of Water Harvesting Schemes for Plant Production*. Food and Agriculture Organization of the United Nations, Rome.
10. Central Ground Water Board. (2000). *Rainwater Harvesting and Artificial Recharge to Groundwater-A Guide to Follow*. CGWB and United Nations Educational, Scientific and Cultural Organisation, New Delhi, India.
11. Thames, J.L. (1985). *Water Harvesting*. Proceedings of the FAO Expert Consultation on the Role of Forestry in Combating Desertification, Saltillo, Mexico June 24–28, 1985, Thematic paper 3.11.
12. Prinz, D. (1994). Water harvesting: past and future. *Sustainability of Irrigated Agriculture*. Proc., NATO Advanced Research Workshop, Vimerio, March 21–26, 1994, Balkema, Rotterdam, pp. 135–144.
13. Nevada Division of Water Resources. *Nevada Water Words Dictionary*. Carson City, NV [http://www.water.nv.gov/].
14. Prinz, D. and Singh, A. (2000). *Technological Potential for Improvements of Water Harvesting*. Contributing paper to the World Commission on Dams final report Dams and Development: A new Framework for Decision Making, Earthscan, London, UK.
15. Boers, T.M. and Ben-Asher, J. (1982). A review of rainwater harvesting. *Agric. Water Manage.* **5**: 145–158.
16. Barton, J. (2002). *Water Harvesting for Groundwater Recharge—Is It Effective?* Seminar presented at the student seminars on water harvesting in drought prone areas, University of Wales, Bangor, April 2002.
17. Kakade, B.K., Neelam, G.S., Petare, K.J., and Doreswamy, C. (2001). *Revival of Rivulets Through Farm Pond Based Watershed Development*. BAIF Development Research Foundation, Pune, India.
18. Chandrakanth, M.G. and Diwakara, H. (2001). *Synergistic Effects of Watershed Treatments on Farm Economy Through Groundwater Recharge—A Resource Economic Analysis*. Department of Agricultural Sciences, Bangalore, India.

URBAN WATER RESOURCE AND MANAGEMENT IN ASIA: HO CHI MINH CITY

Hiep N. Duc
Environment Protection
Authority, NSW
Bankstown, New South Wales
Australia

INTRODUCTION

Ho Chi Minh City, the largest city in Vietnam, has a current population of about 5 million that is increasing rapidly due to internal migration and natural growth. The water demand and the stress on the environment due to increasing waste and pollution pose a serious challenge to the people and the authority administering the city.

The Ho Chi Minh City Service of Communication and Public Works (SCPW) admitted in September 1998 that the city's decades-old water supply system cannot meet the growing demand. According to the SCPW, the water consumption is estimated at 800,000 m^3 per day, but the combined daily capacity of its two water plants at Thu Duc and Hoc Mon and underground wells stops short at 750,000 m^3. But just only 500,000 m^3 is available to users each day after water loss in delivery is taken into account (the proportion of water loss is as high as 32% in some areas). However, in 2000, the daily demand jumped to 1,250,000 m^3/day but the Water Supply Company can provide only 850,000 m^3/day.

Ho Chi Minh City is situated next to the Saigon River, which joins the Dong Nai River to form the northern and eastern boundaries of the city. This means that the waterways can be both potential water sources and receivers of polluted wastes.

The Dong Nai River has a minimum flow of approximately 100 m^3/s. It originates in the Central Highlands of Vietnam and flows through Dong Nai and Ho Chi Minh City with tributaries from other provinces. The total catchment area is 42,665 km^2, and the total flow volume is 30.6 km^3/year (1). Forests cover approximately 30% of the basin. The upstream area in the lowland part of the river is the present Cat Tien National Park. The Tri An reservoir and hydropower plant are located downstream nearby.

WATER ENVIRONMENT AND POLLUTION

One of the important aspects of water resource usage management is the prevention and mitigation of sources of pollution going into the watercourses.

Sources of Pollution

The majority of industries are located in industrial zones around Ho Chi Minh City. In addition to old industrial zones such as Bien Hoa zone 1 built in 1963, a large number of new industrial zones were set up recently in the neighboring provinces of Dong Nai, Song Be, and Ba Ria-Vung Tau. They were established to attract foreign investment to set up factories to produce goods for domestic and export markets. The industrial zone in Bien Hoa (Dong Nai) is upstream from Ho Chi Minh City. The foremost concern is that the wastewater discharge from this zone to the Dong Nai River is near the Hoa An pumping station that supplies water to the residents of Ho Chi Minh City (2).

Altogether, it was estimated that about 200,000 m^3/day of wastewater are discharged into the Dong Nai–Saigon River from the industrial zones. Industrial source discharges are located mostly in the Thu Duc and Bien Hoa areas, but they can also be found scattered along many canals in and around Ho Chi Minh City. There are five main canal systems that receive direct wastes from various sources: Nhieu Loc-Thi Nghe, Tau Hu-Ben Nghe, Doi-Te, Tan Hoa-Lo Gom, and Tham Luong-Vam Thuat. The total length of these canal systems is about 56 km with about

100 km of smaller tributaries. According to statistical data from the Water Supply Company, there is 461.291 m^3 of wastewater discharge into the canal systems daily (2).

Water pollution is more visible during the rainy season when the inadequate drainage system cannot cope with the water volume. In many areas of the city, traffic cannot move due to extensive flooding after a rain.

Pollution Management

Wastewater treatment facilities are required for new and established industrial zones. In Binh Duong province, a wastewater treatment plant was built in 1998 for the Vietnam–Singapore Industrial Park and adjacent areas. The system includes an 8-km canal to drain rain and treated wastewater from the zone into the Saigon River. In Bien Hoa II Industrial Park, a plant for treating wastewater from factories was put into operation in 1999. The plant has an initial daily capacity for treating 4,000 m^3, and 20 of the 76 factories had been connected to the plant since it started operating. The 2-hectare plant has four reservoirs for biological, chemical, and physical treatment of liquid waste before it is conveyed to another reservoir for final treatment.

Another issue regarding the management of water pollution is the storm water system. In 2001, the municipal Communication and Public Works will spend VND101 billions to upgrade and improve the storm water drainage system. It is part of a 5-year plan to reduce flooding during the rainy season. To partly finance the maintenance and upgrading of the drainage system, the city authority decided in early 2001 to allow the Water Supply Company to collect a water drainage levy from water users starting in July 2001. The rest of the capital is in the form of loans from international bodies such as foreign governments or the World Bank.

The most significant plan for improving water quality and the environment is the Urban Drainage and Sewer Master Plan. This plan was prepared and developed in 1999 with assistance from the Japanese Government. The Nhieu Loc–Thi Nghe canal basin has been identified and given the highest priority to improve the water environment and associated infrastructure (3). The shoreline improvement via house clearance, relocation, and parkland development as well as canal dredging has been done since 1996 but without capital investment in the drainage and sewer infrastructure. Street flooding and traffic jams during the rainy season are a frequent phenomenon. The Nhieu Loc–Thi Nghe Project was started in early 2001; part of the financing came from a World Bank loan. The Project is managed and implemented by the Urban Drainage Company.

The city Urban Drainage and Sewer Master Plan conforms to the national Urban Wastewater Collection and Sanitation Strategy Study and the Government Decision No. 35/1999/QD-TTg of March 5, 1999 (3).

WATER SUPPLY AND DISTRIBUTION

Government and City Organizations

In 1977, the city People's Committee established the Environmental Protection Council to oversee the environment. In 1992, the Environment Committee was formed. In 1994, the Department of Science, Technology and Environment (DOSTE) was established as part of national body run by the Ministry of Science, Technology and Environment (MOSTE) when the Environment Law was passed by the National Assembly.

The Department of Transport Communication and Public Works is the city government body in charge of planning the water supply network and other infrastructures. HCM Water Supply Company is the city operating enterprise in charge of supplying water to the city. The Urban Water Drainage Company is another city authority that oversees the operation of the wastewater and storm water network. The Waste Treatment Company of Ho Chi Minh City is concerned with the disposal of solid wastes, including household and hospital wastes, produced in the city. It owns several landfill waste sites in and around the city. These three public service enterprises report to the city's Director of Transport Communication and Public Works.

All the above companies are city government enterprises. The first private water treatment plant in Vietnam was built on a build-operate-transfer (BOT) basis under a 20-year contract with the city People's Committee and the Water Supply Company of Ho Chi Minh City. The plant was built and run by Binh An Water Corporation, a consortium of Malaysian companies that has a loan from the International Finance Corporation (IFC) (4). The Binh An treatment plant will provide 100,000 m^3/day (about 10% of the current demand) to the Ho Chi Minh City and Bien Hoa industrial zone (5).

Water Supply Network

The main water supply plant for Ho Chi Minh City is the Thu Duc Water Station. It currently supplies 650,000 m^3/day to the city using water from the Dong Nai River. The Dong Nai Water Supply Project will increase the plant capacity to 750,000 m^3/day (6). Part of the project cost is funded by a loan from the Asian Development Bank (ADB). Lyonaise des Eaux is building another water plant at the existing Thu Duc Water station to supply 300,000 m^3 each day to districts 2, 7, 9, Thu Duc, and Nha Be. This is a 25-year joint BOT venture with the Water Supply Company.

In 1991, there was a project planned to build a new water supply plant using water from the Saigon River with funding aid from the Italian government. The planned supply plant has a capacity of 300,000 m^3/day and would provide water to District 12, 6, Tan Binh, and Binh Chanh. But as the project was underway in 1995, the aid was stopped by the Italian government due to faults in timing management. Since then, the HCM People Committee has tried to use various schemes to get it restarted, first as a joint venture with foreign companies, then as a BOT project, and finally in April 2000 as a total sale package including equipment and project works to private investors. But so far none was successful.

In surrounding areas of the city, the water demand was alleviated by some new projects. In 1997, the Hoc Mon Underground Water plant was upgraded to a capacity of 50,000 m^3/daily using a loan from the Asian Development

Bank (ADB). In Dong Nai province, construction of a new water supply plant is being carried out at Thien Tan. The US $25 million Thien Tan water project is funded by aid from the government of the Republic of Korea to provide clean water to more residents of the province. The plant will have a capacity of 100,000 m^3/day.

BIBLIOGRAPHY

1. Talaue-McManus, L. (2000). *Transboundary Diagnostic Analysis for the South China Sea.* EAS/RCU Technical Report Series No. 14, UNEP, Bangkok, Thailand.
2. Nguyen Thi Lan (6/3/1997). Bao Dong o nhiem moi truong nuoc Tp HCM, Tuoi Tre.
3. World Bank—Vietnam-Ho Chi Minh Environmental Sanitation Project. http://www.worldbank.org/pics/pid/vn52037.txt
4. IFC Press Release (1/7/1998). *IFC invests US $25 million in first private water treatment facility in Vietnam.*
5. World Bank—Vietnam Binh An Water Corp. Project, http://www.worldbank.org/pics/ifcers/vne08948.txt.
6. *Saigon Times Daily* (March 22, 2001). City seeks capital scaledown for water supply project.

HYDROPOWER—ENERGY FROM MOVING WATER

Energy Information Administration—Department of Energy

Of the renewable energy sources that generate electricity, hydropower is the most often used. It accounted for 6 percent of U.S. generation and 42 percent of renewable generation in 2001. It is one of the oldest sources of energy and was used thousands of years ago to turn a paddle wheel for purposes such as grinding grain. Our nation's first industrial use of hydropower to generate electricity occurred in 1880, when 16 brush-arc lamps were powered using a water turbine at the Wolverine Chair Factory in Grand Rapids, Michigan. The first U.S. hydroelectric power plant opened on the Fox River near Appleton, Wisconsin, on September 30, 1882. Until that time, coal was the only fuel used to produce electricity. Because the source of hydropower is water, hydroelectric power plants must be located on a water source. Therefore, it wasn't until the technology to transmit electricity over long distances was developed that hydropower became widely used.

Mechanical energy is derived by directing, harnessing, or channeling moving water. The amount of available energy in moving water is determined by its *flow* or *fall*. Swiftly flowing water in a big river, like the Columbia River along the border between Oregon and Washington, carries a great deal of energy in its flow. So, too, with water descending rapidly from a very high point, like Niagara Falls in New York. In either instance, the water flows through a pipe, or *penstock*, then pushes against and turns blades in a turbine to spin a generator to produce electricity. In a *run-of-the-river system*, the force of the current applies the needed pressure, while in a *storage system*, water is accumulated in reservoirs created by dams, then released when the demand for electricity is high.

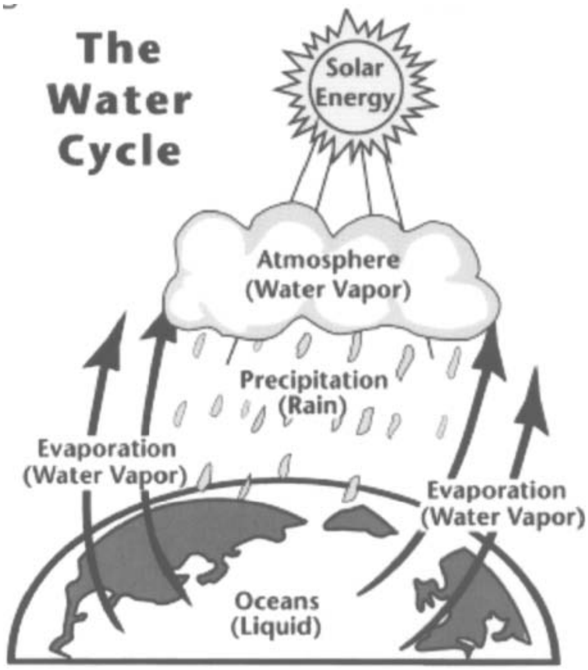

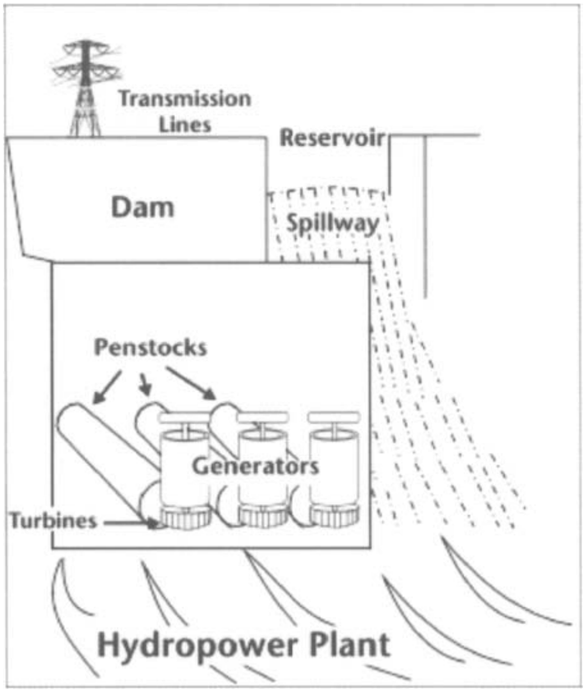

Meanwhile, the reservoirs or lakes are used for boating and fishing, and often the rivers beyond the dams provide opportunities for whitewater rafting and kayaking. Hoover Dam, a hydroelectric facility completed in 1936 on the

---

Colorado River between Arizona and Nevada, created Lake Mead, a 110-mile-long national recreational area that offers water sports and fishing in a desert setting.

Over one-half (52 percent) of the total U.S. hydroelectric capacity for electricity generation is concentrated in three States (Washington, California and Oregon) with approximately 27 percent in Washington, the location of the Nation's largest hydroelectric facility—the Grand Coulee Dam.

It is important to note that only a small percentage of all dams in the United States produce electricity. Most dams were constructed solely to provide irrigation and flood control.

Some people regard hydropower as the ideal fuel for electricity generation because, unlike the nonrenewable fuels used to generate electricity, it is almost free, there are no waste products, and hydropower does not pollute the water or the air. However, it is criticized because it does change the environment by affecting natural habitats. For instance, in the Columbia River, salmon must swim upstream to their spawning grounds to reproduce, but the series of dams gets in their way. Different approaches to fixing this problem have been used, including the construction of "fish ladders" which help the salmon "step up" the dam to the spawning grounds upstream.

WATER MARKETS IN INDIA: ECONOMIC AND INSTITUTIONAL ASPECTS

Halanaik Diwakara
University of South Australia
Adelaide, Australia

BACKGROUND

Historically, India's water resources have been perceived as an important input to production in agriculture for regional and economic development. India is an agrarian economy. Therefore, water is the elixir of life for millions of farmers in the country. Water markets are not akin to commodity markets. Because of ambiguous property rights to both surface and groundwater, informal selling and buying are popping up in different parts of India. These are not legal; surface water structure is built by government funding, so farmers have no right to sell it. In addition, electricity pumping costs are not borne by farmers and a major portion of groundwater pumping is through electricity; so farmers have no right to sell groundwater also.

The water markets are prominent because of the profit from the sale of water for nonagricultural purposes. Electricity is virtually free, so selling groundwater for agriculture is not as profitable as it is for nonagricultural purposes. Even though nonagricultural purposes such as domestic water needs are not as large as for agriculture, consumer demand for domestic water is apparent, and consumers are willing to pay more than farmers who buy water for agriculture. Groundwater is reckoned as a poverty reduction tool in rural India, and its contribution to rural wealth is more than that of surface water (1).

Increasing demand for groundwater in India on account of the Green Revolution and commercial farming has led to enormous pressure on groundwater resources. Because of the growing population and expanding economy, the demand for water is surpassing the supply. Further, development of water resources is constrained by investment bottlenecks, environmental concerns, and political and legal snags inherent in interregional water transfers (3). Though the onus of achieving effective use of water falls on water markets, unless the true cost (value) of water is reflected in terms of actual electricity cost, any attempt by the government would be futile.

Reallocation of water through marketing is a response to growing demands on limited supplies of water (4). Neoclassical economic theory tells us that a market increases efficiency by allocating resources to their most valuable use. However, in the case of natural resources, the economic theory of markets may not always be true, for example, in central California, where the water market is immature (5). This is precisely true for India as well.

This article explores the nature and extent of water markets in India and focuses on groundwater markets and relevant economic and institutional (legislative) aspects. In the next section, the nature and extent of water markets in India are provided, followed by some economic aspects of water markets. In the following section, institutional aspects focusing on regulation of groundwater resource are discussed in the context of water markets. Finally, some water policy (institutional) issues under which water markets can be improved by a water reform process in India are discussed.

NATURE OF WATER MARKETS

Recent literature on water markets in India indicates that markets for groundwater are one of the ways of accessing irrigation (6). These water markets are niche markets, highly location-specific and endemic, at least in south India. In India, water markets are informal; the transfer of water takes place among farmers (between irrigators). In informal water markets, enforcement of contracts does not occur by recourse to legal and institutional measures but rather on personal trust (4). Some markets for water resemble close to perfect competition. Each seller is offering an identical product, so there is free entry and exit of farmers, no single buyer and seller can influence the price, and there is enough information as well (for instance, water markets in Gujarat). These water markets are seasonal and not permanent. Though the water transfer takes place mainly between irrigators, in some pockets like in periurban areas of Bangalore, a southern city of India, water is traded between irrigator and nonirrigators, water markets are localized, and the nature of competition resembles an oligopoly (7). Similar findings are reported by Palanisami (8) and Janakarajan (9). Nevertheless, the demand for groundwater remains large for agriculture.

EXTENT OF WATER MARKETS

The extent of the area irrigated through water markets, which is often considered a proxy of water trading, varies across regions as well as over time, depending on a number of factors such as rainfall, groundwater supply, cropping patterns, and the cost and availability of electricity (3).

It is estimated that 20% of the owners of the 14.2 million pump sets in India are likely to be involved in water trading (10); this indicates that water markets are providing water for about 6 million hectares, or 15% of the total area irrigated by groundwater (4). This is on the high side and therefore it cannot be a reliable estimate. Still, there is a lack of systematic estimates of the magnitude of water trading on the macrolevel. Using a methodology based on pump set rental data, Saleth (10) estimated that 6 million hectares or 15% of the total area under groundwater irrigation is on account of water trading. Further, by a net addition to output of $230/ha/year (based on the difference between the average irrigated and rain-fed yields, as reported by Government of India), the total value of output due to water sales is estimated at $1.38 billion per year (11). According to the irrigation and electricity acts, water selling is illegal, but it persists in niches.

ECONOMIC ASPECTS

The thesis behind the economics of groundwater use is that *water is used most efficiently when it is extracted at rates that maximize net benefits over time*. Hence, the benefit–cost calculus matters. The cost components include the cost of extracting and delivering the groundwater, and the opportunity or the uses to which the water is put determines user cost and benefits. For example, Diwakara and Nagaraj (7) indicated that the benefit–cost ratio for groundwater sales to an urban area in an informal setting is 2.67:1 for the construction and hotel industries compared to 1.35:1 for paddy crops. This signals that water transfer to nonagriculture is a lucrative business. The costs of extraction are primarily a function of pumping technology, the depth from which the groundwater can be pumped, and the costs of energy (mainly electricity in India). These costs increase with pumping depth and costs of energy and decrease as pump efficiency is improved. Thus the water market for agriculture is different from the water market for nonagriculture. Farmers selling water for nonagricultural purposes are relatively more efficient than farmers selling for agriculture because the marginal value product (MVP) of the former is higher than that of the latter.

In India, the cost of energy used to pump the water is charged on a flat rate (fixed) basis, $30.55 per horsepower (in Gujarat), indicating that the marginal extraction cost of pumping groundwater is zero; this encourages farmers to use water to the point where the marginal value of production is close to zero.

The appropriators and policy makers largely ignore the user pays principle, so water markets are inefficient and illegal because user costs are not borne either by the buyer or seller. This has encouraged water sellers to use electric motors to their full capacity at no marginal cost. Further, the cost of extraction also includes the value of the opportunity foregone by extracting and using the water immediately rather than at sometime in the future (opportunity cost). The user cost is a measure of the economic consequences of pumping now and thereby lowering the water table and increasing the cost of extraction for all future periods, that is, increasing negative economic effects that are not internalized. Therefore, sellers are more into a neoclassical return through water rates and interlocked markets.

INSTITUTIONAL (LEGAL) ASPECTS

So far in India, there has been little or no effort to devise appropriate water institutions to cope with the increasing demand for water in a dynamic agrarian economy. Solutions to the technical problems of groundwater resources are well known and are enforceable. However, efficient allocation and use of water resources depend on institutions, which have received lip service. An influential article, “The Tragedy of the Commons” by Garrett Hardin (12) encouraged scholars to hold the view that lack of private property rights to the resource has led to degradation and overexploitation of common pool resources (13). The solution is to promote self-governing institutions (14,15).

Albeit there has been an attempt to regulate groundwater resources through the “Model Bill to Regulate and Control the Development of Groundwater” in 1992, but it failed to address the groundwater extraction and use issues; rather, more emphasis was given to administrative control mechanisms. This misses the potential management opportunities represented by water markets (16). The National Water Policies of 1987 and 2002 have also failed to address regulatory issues to govern the private extraction of groundwater resources.

The issue of property rights has fueled the recent debate on groundwater ownership. It is believed that assigning property rights to groundwater and permitting competitive markets to function can finesse most of the inefficiencies of the common pool problem (17), but this neglects the fact that the appropriative activities by which individuals create effective property rights are themselves an alternative use of scarce resources (18). The priority for exclusive private ownership may be an obstacle to incompletely defining resources situations where some collective ownership may be more appropriate (19). The main function of property rights is to encourage greater internalization of effects (20,21). The rights to extract groundwater are by and large in the hands of private individuals who own land. So, the landlords are also water lords in India, provided the groundwater is available from the vulnerable aquifers. Groundwater is a common pool resource unit, so its yield is subtractable and its exclusion is not an issue of rivalry (22). Hence, defining the rights to use is a challenging task.

Current energy pricing (flat rate) is not desirable for India's groundwater sustainability. Even though some big farmers use high-energy diesel pumps to lift groundwater, the water extraction will not be and cannot be as much as that in the use of electrical pumpsets because farmers have to pay for diesel and not for electricity. Hence, groundwater mining and overuse of energy are interlinked and necessitate a high opportunity cost in water-scarce regions. In Gujarat State, energy pricing is US $37 per year per pump horsepower (23). Therefore, marginal cost pro-rata pricing for electricity could be one of the welfare-improving institutional measures that could be considered. However, as Zilberman (24) notes, these welfare measures would be felt only when the efficiency benefits from reform are more significant and apparent.

On the one hand, groundwater is treated as a "free good," and on the other, it is exploited competitively. This underscores the importance of well-defined, clearly enforceable rights to extract groundwater. Groundwater resources are largely extracted by individual well owners in India. In the absence of well-defined and enforceable rights to extract water, the availability of groundwater is determined and subject to the law of capture, that is, whoever taps the groundwater first gets to use it. In such a situation, pumpers have an incentive to extract as much water as possible, subject to the constraints imposed by pumping costs (25).

Access to groundwater, a fugitive (dynamic) resource, is largely a function of rights to extract. Though a legal framework that ties rights to groundwater to land ownership governs groundwater use in India, there is no legal limit on the amount of water a landowner can draw. Hence, markets for groundwater in India are informal and illegal and fluctuate in pricing as well as in pricing methods. For example, if a farmer is selling groundwater to his neighbor, it is also in return for some labor shortage situations or may be some interest payments and so on. These, too, are subsumed in the water charge paid by the buyer.

In the absence of a limit on the volume of groundwater extracted, farmers are tapping the resource with myopic behavior ignoring the fact that one's extraction is a function of the neighboring well's extraction at a time and over time leading to cumulative interference of wells that affects the lives of the wells (26). The lack of quantity regulation of groundwater extraction might lead to the *tragedy of the water law*. Hence, quantity regulations are common practices around the world and are preferred to taxes (27).

CONCLUDING REMARKS

Though water policy has generally been considered an important issue, its rational formulation and implementation have received lip service, and the water policies of the twenty-first century have to be significantly different from the past policies because of several changes that have already occurred in the water sector (28). India is gearing up to meet the challenges of the millennium with a fast developing economy. Throughout history, institutions have been a panacea for managing natural resources. Without suitable water institutions, the dream of achieving sustainable use of water resources will be in vain. Recognizing water markets as informal institutions that tackle the problem (access) and efficiency (high value use), appropriate measures through regulation of groundwater are the need of the hour for India. It is indicative that markets and states are now seen not as competing but as complementary institutions in the quest to "*get the rules right*," and many formulations see a broader range of institutions of economic governance as essential in this task, including small scale communities-neighborhoods, nongovernmental associations, and the like (29). In this regard, a blend of regulatory and economic measures applicable to Indian situation is suggested herewith.

There is a lack of explicit policy on water markets in India. Keeping in view the reckoned benefits of formal water markets, an attempt could be made to devise a set of rules focusing on property rights (tradable), quantity regulation, energy pricing, well-spacing norms, water laws to promote self-governing institutions that manage groundwater, and establishing an oversight commission to settle disputes between users and water authorities.

PROPERTY RIGHTS

Groundwater in India is largely developed by private investment, as in many parts of the world. However, the rights to extract water are still chattel to the land. The potential benefit of well-defined property rights is that it reduces transaction costs, and assuming access to credit, small farmers can have water rights. Considerable consensus exists that, for markets to be effective, legally defining property rights to water as private and transferable should be in place (4,30,31) along with a declaration of rights for efficient use of the resource (32). As the sovereign power and guardian of the nation's resources, as well as the protector of the law and rights, the government has the power to define necessary property rights and then dispose them to individuals (33). However, the third party effects of water markets need attention because the effects of water transfers and the costs of administering efficient and fair solutions to these third party effects could lead to restricting trade (34). In this regard, quantity regulation would be a suitable strategy.

QUANTITY REGULATION

The groundwater resource is extracted in unlimited quantities in India without regard to its impact on neighboring wells and environmental consequences. Evidence indicates that, in a variety of situations, governments regulate undesirable activities by controlling quantity rather than charging taxes (27). So, simply defining property rights is not a sufficient condition. It should be blended with quantity regulation. If quantity regulation is opted for and implemented, still, the access to groundwater resources is a function of number of users across time and space. Hence, to facilitate efficiency (physical and economic), the resource has to be priced to reflect its true value. This could be done through marginal cost pricing of electricity and groundwater.

MARGINAL COST PRICING

The energy used to lift the groundwater from the deeper layers of the aquifer in many parts of India is electricity and diesel to some extent. Electricity is priced on a fixed (flat rate) basis. Hence, the marginal cost of extraction is zero, and there is no incentive for users to conserve the resource. In a true economic sense, users extract groundwater until the marginal value product becomes zero. The value of groundwater needs to be reflected through marginal cost pricing of electricity and groundwater as well. This could be achieved by pricing electricity on a pro-rata basis.

WELL-SPACING NORMS

In the absence of well spacing norms, the well interference problem poses serious threats to sustainability and equity in well irrigation, especially in water-scarce regions in India. The withdrawal of groundwater without regard to recharge efforts goes unabated, and the resource itself becomes unsustainable by reducing the lives of wells. So, to avoid the well interference problem, which underscores sustainability, there is a need for spacing norms between wells.

PROMOTING SELF-GOVERNING INSTITUTIONS

The community should accept a law to allocate water rights, regulate quantities of water, implement marginal cost pricing, and impose well-spacing norms. Evidence indicates that, if the economic rents from water trading are concentrated in the hands of a few individuals or the negative effect on third-party users is large and unmitigated, the community is not likely to obey the law (4). Therefore, self-governing institutions, such as water users' cooperatives, associations, or small private well organizations need to be promoted by enacting a suitable water law. In any society, there is no guarantee that all stakeholders obey the law. To settle disputes between violators of the rules, there is a need for an oversight commission with representatives from users, local, state, and federal governments.

In conclusion, emphasis should be on collective action through user groups and cooperatively managed irrigation systems for sustainable use of water resources. Therefore, a combination of local water rights and user organizations should be encouraged. Nevertheless, the institutional linkage between water markets and rental markets for irrigation assets needs attention because the rental market allows farmers to irrigate their farms by renting irrigation assets from neighbors, and they contribute both to equity in water use and better use of irrigation assets, especially when groundwater rights are not defined (35).

These are some of the measures that are believed to be possible in the Indian scenario and could be incorporated into legislation, provided that the sociopolitical economy realizes the need and accepts them. Further, there shall not be any underenforced water law, especially in the context of the economic and environmental damage that has been created and unaccounted for largely due to policy failure.

BIBLIOGRAPHY

1. International Water Management Institute. (2002). The socio-ecology of groundwater in India. *Water Policy Briefing* **4**: 1–6.
2. Bhatia, B. (1992). Lush fields and parched throats: Political economy of groundwater in Gujarat. *Economic and Political Weekly* **XXVII**(51–52): A12–170.
3. Saleth, R.M. (1994). Groundwater markets in India: A legal and institutional perspective. *Indian Econ. Rev.* **29**(2): 157–176.
4. Easter, K.W., Rosegrant, W.M., and Dinar, A. (1999). Formal and informal markets for water: Institutions, performance, and constraints. *World Bank Res. Observer* **14**(1): 99–116.
5. Carey, J., Sunding, L.D., and Zilberman, D. (2002). Transaction costs and trading behaviour in an immature water market. *Environ. Dev. Econ.* **7**: 733–750.
6. Dubash, K.N. (2000). Ecologically and socially embedded exchange: Gujarat model of water markets. *Economic and Political Weekly* **April 15**: 1376–1385.
7. Diwakara, H. and Nagaraj, N. (2003). Negative impacts of emerging informal groundwater markets in peninsular India: Reduced local food security and unemployment. *J. Soc. Econ. Dev.* **5**(1): 90–105.
8. Palanisami, K. (1994). M.W. Rosegrant and R.G. Schleyer (Eds.). Tradable Water Rights Experiences in Reforming Water Allocation Policy, USAID, Arlington, VA.
9. Janakarajan, S. (1994). M. Moench (Ed.), *Selling Water: Conceptual and Policy Debates over Groundwater Markets in India, VIKSAT*, Ahmedabad.
10. Saleth, R.M. (1998). Water markets in India: Economic and institutional aspects. In: *Markets for Water: Political and performance*, K.E. William, M.W. Rosegrant, and A. Dinar (Eds.). Kluwer Academic, Boston, MA.
11. Mohanty, N. and Gupta, S. (2002). Breaking the gridlock in water reforms through water markets, Working Paper Series, Julian L. Simon Centre for Policy Research.
12. Hardin, G. (1968). The tragedy of the commons. *Science* **162**: 1243–1248.
13. Sethi, R. and Somanathan, E. (1996). The evolution of social norms in common property resource use. *Am. Econ. Rev.* **86**(4): 766–788.
14. Ostrom, E. and Gardner, R. (1993). Coping with asymmetries in the commons: Self-governing irrigation systems can work. *J. Political Econ.* **7**(4): 93–112.
15. Bardhan, P. (1993). Analytics of the institutions of informal cooperation in rural development. *World Dev.* **21**(4): 633–639.
16. Moench, M. (1998). M. Moench (Ed.), *Groundwater Law: The growing debate*, VIKSAT, Ahmedabad pp. 35–68.
17. Gisser, M. (1983). Groundwater: Focusing on the real issue. *J. Political Econ.* **91**(6): 1001–1027.
18. Grossman, I.H. (2001). The creation of effective property rights. *Am. Econ. Rev.* **91**(2): 347–352.
19. Dragun, K.A. (1987). Property rights in economic theory. *J. Econ. Issues* **XXI**(2): 859–868.

20. Demsetz, H. (1967). Toward a theory of property rights. *Am. Econ. Rev.* **57**(2): 347–359.
21. Dragun, K.A. (1983). Externalities, property rights, and power. *J. Econ. Issues* **XVII**(3): 667–680.
22. Walker, J.M. and Gardner, R. (1992). Probabilistic destruction of common pool resources: Experimental evidence. *Econ. J.* **102**: 1149–1161.
23. Gupta, K.R. (2002). Water and energy linkages for groundwater exploitation: A case study of Gujarat State, India. *Water Resour. Dev.* **18**(1): 25–45.
24. Zilberman, D. (2003). Water marketing in California and the West. *Int. J. Public Adm.* **26**(3): 291–315.
25. Committee on Valuing Groundwater. (1997). *Valuing Groundwater: Economic Concepts and Approaches*. National Academy Press, Washington, DC.
26. Chandrakanth, M.G. and Arun, V. (1997). Externalities in groundwater irrigation in hard rock areas. *Indian J. Agric. Econ.* **52**(4): 761–771.
27. Glaeser, L.E. and Shleifer, A. (2001). A reason for quantity regulation. *Am. Econ. Rev.* **91**(2): 431–435.
28. Biswas, K.A. (2001). Water policies in the developing world. *Water Resour. Dev.* **17**(4): 489–499.
29. Bowles, S. and Gintis, H. (2000). Walrasian economics in retrospect. *Q. J. Econ.* **November**: 1411–1439.
30. Bauer, J.C. (1997). Bringing water markets down to earth: The political economy of water rights in Chile, 1976–95. *World Dev.* **25**(5): 639–656.
31. Hearne, R.R. and Easter, K.W. (1995). Water Allocation and Water Markets: An Analysis of Gains-from-Trade in Chile, World Bank Technical Paper Number 315. The World Bank, Washington, DC.
32. Schlager, E. and Ostrom, E. (1992). Property rights regimes and natural resources: A conceptual analysis. *Land Econ.* **68**(3): 249–262.
33. Dragun, K.A. (1985). Property rights and Pigovian taxes. *J. Econ. Issues* **XIX**(1): 111–122.

READING LIST

Brennan, D. and Scoccimarro, M. (1999). Issues in defining property rights to improve Australian water markets. *Aust. J. Agric. Resour. Econ.* **43**(1): 69–89.

Richards, A. and Singh, N. (2001). No easy exit: property rights, markets, and negotiations over water. *Water Resour. Dev.* **17**(3): 409–425.

Saleth, R.M. (2004). Strategic analysis of water institutions in India: applications of a new research paradigm, Research Report, 79. International Water Management Institute, Colombo, Sri Lanka.

Shah, T. (1993). *Groundwater Markets and Irrigation Development: Political Economy and Practical Policy*. Oxford University Press, Bombay.

WATER RESOURCES OF INDIA

Sharad K. Jain
National Institute of Hydrology
Roorkee, Uttranchal, India

India occupies an important place in Southeast Asia. Its geographic area of 3.29 million square kilometers extends from 8° 4′ N and 37° 6′ N latitude and between 68° 7′ E and 97° 25′ E longitude. The country is bounded by the Himalayas in the north and has a large peninsular region tapering toward the Indian Ocean. The coastline of India extends for more than 6000 km. This unique combination of location and topography, a large peninsula, very high mountains followed by wide river plains, and the ocean in the south, create a complex climate and hydrology in the country.

From the physiographical point of view, India can be divided into seven broad regions: (1) northern mountains; (2) Plains of Indus and Ganga Basin; (3) Central Highlands; (4) peninsular plateau, the east coast region; and (5) the west coast region. Two chains of islands, the Andaman and Nicobar in the Bay of Bengal and Lakshadweep in the Arabian Sea, are also part of India.

The Himalayas in the north are one of the major mountain ranges of the world. In India, they extend over 2500 km from east to west, and the width varies from 250 to 400 km. Topographically, the Himalayas can be subdivided into three major ranges, the Greater Himalayas, the Middle Himalayas, and the Shivaliks whose average heights are 6000, 4000, and 1000 meters, respectively. The Himalayas are home to many of the highest peaks of the world and have large permanently snow-covered areas. Geologically, the Himalayas are comparatively young mountain chains that are formed by sedimentary rocks. This coupled with steep slopes and intense monsoon rainfall yields high sediment loads in the rivers. The other prominent mountains include the Aravalis, the Vindhyachals, the Satpuras, the Eastern Ghats, and the Western Ghats. The mountains are the primary source of rivers that derive their flow from rainfall and snow and glacier melt (Himalayan rivers). Plateaux, another striking feature of the topography in India, range in elevation from 300 to 900 meters. The major plateaux are the Malwa, the Vindhyas, the Chota Nagpur, the Satpura, and the Deccan.

India is traversed by a number of mighty rivers that can be broadly divided into two groups: the rivers of the Himalayan region and the peninsular rivers. The snow-fed Himalayan rivers are mostly perennial, whereas many small rivers of the peninsular region remain dry during the summer months.

CLIMATE OF INDIA

The mighty mountains, extensive plateaux, and the ocean in the south have an important bearing on the climate of India that can be described as a tropical monsoon climate. India is a country with extremes of climate. The climate varies from extremely hot to extremely cold, from extremely arid to extremely humid, and from continental to oceanic. Here, temperature varies from more than 47° C at some places in summers to below −40° C at many places in the Himalayas; rainfall varies from almost negligible to about 1100 cm. The eastern state of Meghalaya is the home to Cheerapunji, which is famous for the most rainfall in the world—on the order of 1142 cm in a year. Rainfall amounting to nearly 104 cm has been recorded at this place in a day. The potential evapotranspiration over the country varies from 150 to 350 cm. This wide range of climatic conditions working in conjunction with a range of topographic and soil/rock properties produces a complex and interesting pattern of water resource distribution across the country.

In India, rainfall is received through southwest and northeast monsoons, cyclonic depressions, and western disturbances. Most of the rainfall in India takes place under the influence of the southwest monsoon between June and September, except in Tamil Nadu where it occurs under the influence of the northeast monsoon from October–November. The rainfall in India shows great variations, unequal seasonal distribution, unequal geographic distribution, and frequent departure from the normal. The flow of monsoon winds is significantly influenced by high mountains in the north which are an effective barrier. The peninsular shape of the southern part that is close to the ocean provides a big source of moisture.

Monsoons

The southwest monsoon starts from the equatorial belt and hits the Indian subcontinent in two distinct currents. The Bay of Bengal branch sets in the northeastern part of the country, and the Arabian sea branch hits at the southern part of the peninsula. The first branch moves westward, and the second northward; together they cover the whole country. Normally, the monsoon sets in over the entire country by the beginning of July. The monsoon withdraws in September.

As the southwest monsoon withdraws, a northeasterly flow of air begins. This air picks up moisture from the low-pressure areas in the Bay of Bengal, hits the coastal areas of Orissa and Tamil Nadu, and causes rainfall. During this period, severe tropical cyclones are also formed in the Bay of Bengal and the Arabian Sea. These cyclones are responsible for intense rainfall in the coastal areas.

The water vapor carried by the monsoon from June–September amounts to about 11,100 km^3. About 3000 km^3 (= 27% of the total) of this moisture precipitates as rainfall. During the remaining 8 months of the year, precipitation is of the order of 1000 km^3. In South Indian states such as Tamil Nadu, Andhra Pradesh, and Kerala, the northeastern monsoon also contributes significantly to precipitation.

At the beginning of December, weather disturbances originating in extratropical regions enter India from Afghanistan and Pakistan. These are known as western disturbances and cause moderate to heavy rain and snowfall in the northern mountainous region. Light to moderate rainfall is also experienced in the northern plains.

The annual precipitation over the country, including snowfall, is about 4000 km^3 which is equivalent to about 120 cm of depth. The variation of average annual rainfall over the country is shown in Fig. 1. As seen from the figure, there is a considerable amount of areal variation in the annual rainfall in India. The annual rainfall varies from about 10 cm in western deserts to about 1100 cm in northeastern parts of the country. More than half of the precipitation takes place in about 15 days and that too in about 100 hours in a year. The number of rainy days varies from about five in western deserts to 150 in the northeast. Such a wide variation makes the task of water resource engineers very challenging.

River Basins of India

A river basin, the natural context in which water occurs, is the most appropriate unit for planning, development, and management of water resources. The Ganga–Brahmaputra–Meghna basin, the largest basin in India, receives waters from an area that comprises about one third of the total area of the country. The second largest basin, that of Godavari, covers about 10% of the total area of India. The rivers are well spread over the entire country except for the Thar desert in Rajasthan. The catchment areas of the major river basins and the states through which they flow are given in Table 1. A map of India showing important rivers is given in Fig. 2.

Based on topography, the river systems of India can be classified into four groups: (1) Himalayan rivers, (2) Deccan rivers, (3) coastal rivers, and (4) rivers of the inland drainage basin. The Himalayan rivers receive input from rain as well as snow and glaciers and, therefore, are perennial. During the monsoon months, the Himalayas receive very heavy rainfall, and these are the periods when the rivers carry about 80% or more of the annual flow. This is also the period when floods are commonly experienced. The Deccan rivers are rain-fed and therefore have little or no flow during the nonmonsoon season. The coastal streams, especially on the west coast, have limited catchment areas and are not lengthy. Most of them are nonperennial. The streams of the inland drainage basin of Western Rajasthan are few and far between. They flow for a time only during the monsoon.

The main river systems in the Himalayas are those of the Indus and the Ganga–Brahmputra–Meghna. The River Indus rises near Mansarovar Lake in Tibet. Flowing through Kashmir, it enters Pakistan and empties into the Arabian Sea near Karachi. A number of important tributaries of the Indus, the Sutlej, the Beas, the Ravi, the Chenab, and the Jhelum, flow through India. The Bhagirathi and Alakhnanda are two important rivers that originate in the Garhwal Himalayas. These join at Dev Prayag to form the Ganga, the most sacred river of India. It traverses Uttranchal, Uttar Pradesh, Bihar, and West Bengal and thereafter enters Bangladesh. Some important tributaries of the Ganga are the Yamuna, the Ramganga, the Ghaghra, the Gandak, the Kosi, and the Sone. Many of these tributaries are mighty rivers themselves. For example, River Yamuna has its own mighty tributaries, Chambal and Betwa. The Brahmaputra rises in Tibet (where it is known by the name Tsangpo). It enters India in Arunachal Pradesh and after traversing Assam, enters Bangladesh. Its important tributaries are the Dibang, Luhit, Subansiri, Manas, and Tista. The Ganga and Brahmputra Rivers meet at Goalundo in Bangladesh. The Barak river, the head waters of the Meghna rises in the hills in Manipur. The Meghna is part of the Ganga–Brahmaputra–Meghna system. The combined Ganga–Brahmaputra River meets the Meghna in Bangladesh, and this combined system flows into the Bay of Bengal.

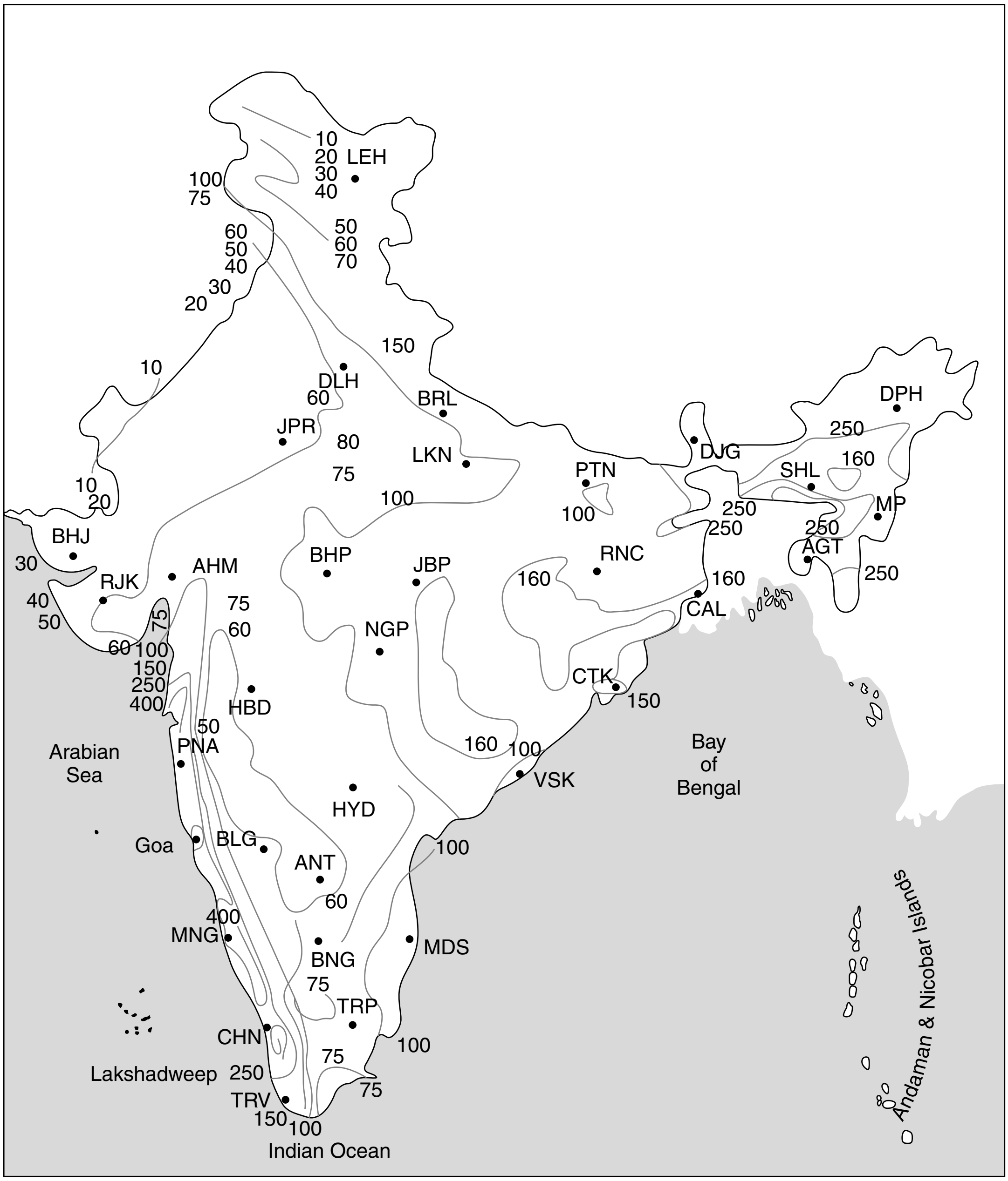

(i) Based upon survey of India map with the permission of the surveyor general of India

(ii) © Government of India copyright 1990.

(iii) The territorial waters of India extend into the sea to a distance of twelve nautical miles measured from the appropriate base line.

Figure 1. Annual rainfall across India (1).

The Deccan rivers can be further classified in two groups: west flowing rivers and east flowing rivers. The Narmada and the Tapi Rivers flow westward into the Arabian Sea. The important east flowing rivers are the Brahmani, the Mahanadi, the Godavari, the Krishna, the Pennar, and the Cauvery. These rivers flow into the Bay of Bengal. The rivers on the west coast are important because they contain as much as 14% of the country's water resources and drain only 3% of the land.

SURFACE WATER RESOURCES OF INDIA

The water resources of a basin are based on stream flow measured at a terminal site on the river. The average annual flow at the terminal point of a river is normally denoted as the water resource of the basin. Note that this refers to availability of water with a probability of 50%. To plan a water resource project, dependability at other levels such as 75% and 90% is needed. Most Indian rivers

Table 1. Major River Basins of India[a]

Sl. No.	River Basin	Catchment Area, km^2	States in the Catchment
1	Ganga–Brahmaputra–Meghna Basin		
1a	Ganga Subbasin	862,769	Uttranchal, Uttar Pradesh, Himachal Pradesh, Haryana, Rajasthan, Madhya Pradesh, Bihar, West Bengal, and Delhi UT.
1b	Brahmaputra subbasin	197,316	Arunachal Pradesh, Assam, Meghalaya, Nagaland, Sikkim, and West Bengal
1c	Meghna (Barak) subbasin	41,157	Assam, Meghalaya, Nagaland, Manipur, Mizoram, and Tripura
2	Indus	321,289	J&K, Punjab, Himachal Pradesh, Rajasthan, and Chandigarh
3	Subernarekha	29,196	Bihar, West Bengal, and Orissa
4	Brahmani–Baitarani	51,822	Madhya Pradesh, Bihar, and Orissa
5	Mahanadi	141,589	Madhya Pradesh, Maharashtra, Bihar, and Orissa
6	Godavari	312,812	Maharashtra, Andhra Pradesh, Madhya Pradesh, Orissa, and Pondicherry
7	Krishna	258,948	Maharashtra, Andhra Pradesh, and Karnataka
8	Pennar	55,213	Andhra Pradesh and Karnataka
9	Cauvery	87,900	Tamil Nadu, Karnataka, Kerala, and Pondicherry
10	Tapi	65,145	Madhya Pradesh, Maharashtra, and Gujarat
11	Narmada	98,796	Madhya Pradesh, Maharashtra, and Gujarat
12	Mahi	34,842	Rajasthan, Gujarat and Madhya Pradesh
13	Sabarmati	21,674	Rajasthan and Gujarat
14	West flowing rivers of Kachchh, Saurashtra, and Luni	334,390	Rajasthan, Gujarat, Daman, and Diu
15	West flowing rivers south of Tapi	113,057	Karnataka, Kerala, Goa, Tamil Nadu, Maharashtra, Gujarat, Daman, Diu, and Nagar Haveli
16	East flowing rivers between Mahanadi and Godavari	49,570	Andhra Pradesh and Orissa
17	East flowing rivers between Godavari and Krishna	12,289	Andhra Pradesh
18	East flowing rivers between Krishna and Pennar	24,649	Andhra Pradesh
19	East flowing rivers between Pennar and Cauvery	64,751	Andhra Pradesh, Karnataka, and Tamil Nadu
20	East flowing rivers south of Cauvery	35,026	Tamil Nadu and Pondicherry Union Territory
21	Area of North Ladakh not draining into the Indus	28,478	Jammu and Kashmir
22	Rivers draining into Bangladesh	10,031	Mizoram and Tripura
23	Rivers draining into Myanmar	26,271	Manipur, Mizoram, and Nagaland
24	Drainage areas of Andaman, Nicobar and Lakshadweep Islands	8,280	Andaman, Nicobar, and Lakshadweep
	Total	3,287,260	

[a] Reference 1.

carry large flows during the monsoon season (June to September). The average annual flow in important Indian rivers is given in Table 2.

Usable Surface Water Resources

Though the estimated surface water availability is of the order of 1869 km^3, the amount of water that can be put to beneficial uses is much less because India experiences monsoon climate, implying that nearly 80–90% of the annual runoff occurs during monsoon months. As enough storage capacity is not available, a large part of this flow goes to waste, and some of it also causes flood damage.

As is clear from Table 2, the storage space available from the projects completed by 1995 is about 174 km^3; the projects under construction would provide 76 km^3, a total of about 250 km^3 of storage space, ignoring small structures. This is about 36% of the possible storage space of 690 km^3. To harness the usable surface water, about 400 km^3 of additional live storage needs to be created. Thus, the development of surface water has a long way to go to achieve full potential.

GROUNDWATER RESOURCES ASSESSMENT

Groundwater resources can be classified as static and dynamic. The static resource can be defined as the amount of groundwater available in the permeable portion of the aquifer below the zone of water level fluctuation. The dynamic resource is the amount of groundwater available in the zone of water level fluctuation. The usable groundwater resource is essentially a dynamic resource which is recharged annually or periodically

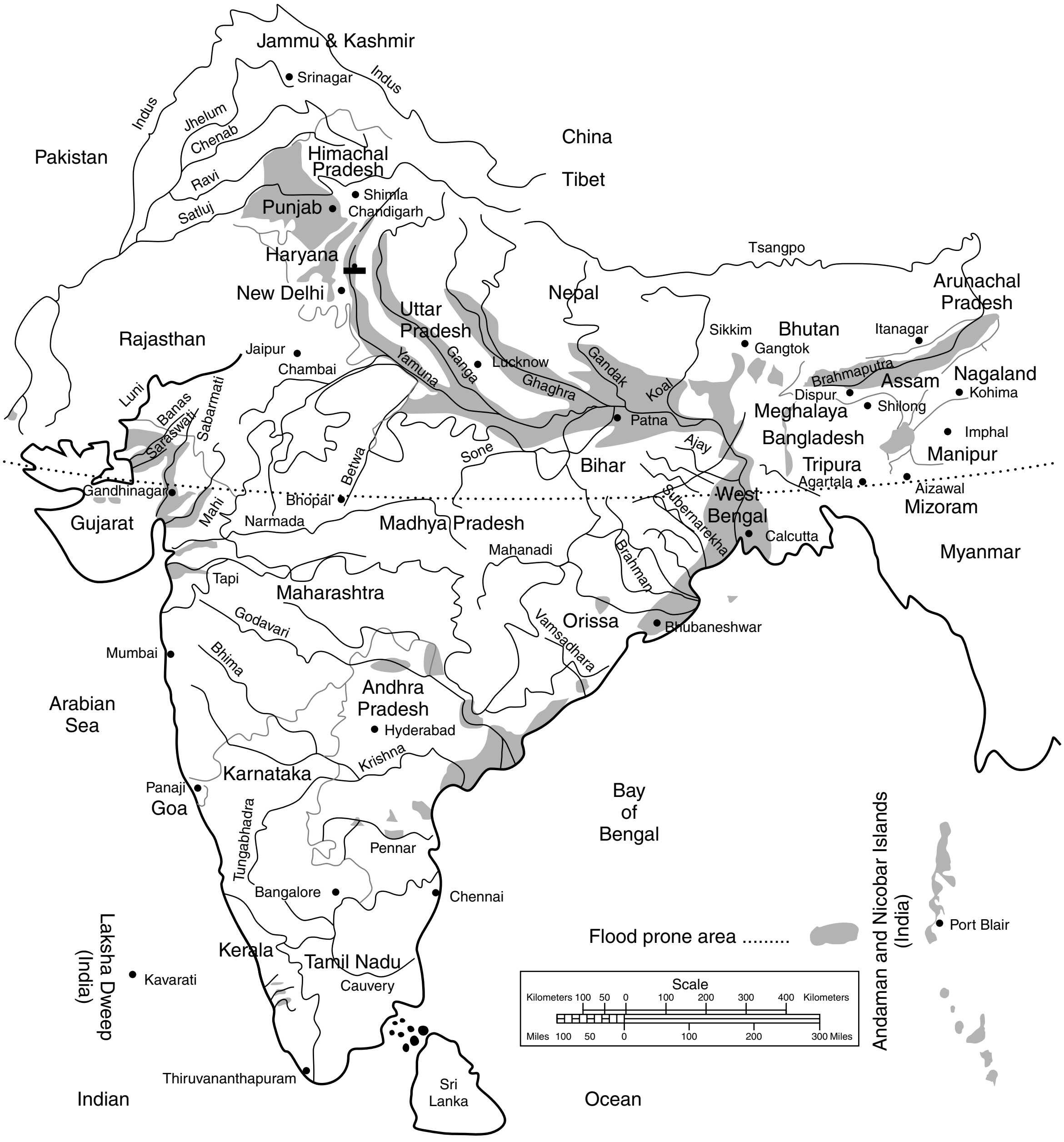

Figure 2. Important rivers of India and flood-prone areas (1).

by rainfall, irrigation return flows, canal seepage, and influent seepage.

Due to diversified meteorologic, topographic, and geologic features, the groundwater distribution is quite divergent across the country. Rock formations vary from Archean to Recent, and the topography ranges from the mountain terrain of Himalayas to the flat alluvial plains and coastal tracts. The northern plains in the Indus–Ganga–Brahmputra Basins that extend over a distance of 2400 km from east to west constitute one of the largest groundwater reservoirs in the world. Almost the entire peninsular part is covered by a variety of hard and fissured rock formations. The coastal and deltaic tracts form a narrow strip around the peninsula; some of them contain highly productive aquifers. The basin-wise groundwater potential of the country is given in Table 3.

The groundwater resources available for irrigation are about 360.3 km^3; out of this, the usable quantity (90% of potential) is about 324.27 km^3. Thus the total usable groundwater resource is $324.27 + 71.08$(domestic and other uses) $= 395.35$ km^3. The present level of development is about 32%.

The water that is present in the aquifer below the zone of water level fluctuation is known as static groundwater. The static groundwater resource of the country is estimated at 10, 812 km^3.

To summarize, the usable volume of water from surface resources is 690 km^3, and from groundwater sources 396 km^3, giving total annual usable water resources of

Table 2. Surface Water Resource Potential of the River Basins of India, km^3

Sl. No.	Name of the River Basin	Average Annual Potential in the River	Estimated Usable Flow, Excluding Groundwater	Live Storage		
				Completed Projects	Ongoing Projects	Proposed Projects
1	Indus (area in Indian territory)	73.31	46.00	13.83	2.45	0.27
2	(a) Ganga	525.02	250.00	36.84	17.12	29.56
	(b) Brahmaputra, Barak, and others	585.60	24.00	1.10	2.40	63.35
3	Godavari	110.54	76.30	12.51	10.65	8.28
4	Krishna	78.12	58.00	34.48	7.78	0.13
5	Cauvery	21.36	19.00	7.43	0.39	0.34
6	Pennar	6.32	6.86	0.38	2.13	—
7	East flowing and rivers from Mahanadi to Godavari and Krishna to Pennar	22.52	13.11	1.63	1.45	0.86
8	East flowing rivers between Pennar and Kanyakumari	16.46	16.73	1.42	0.02	—
9	Mahanadi	66.88	49.99	8.49	5.39	10.96
10	Brahmani and Baitarani	28.48	18.30	4.76	0.24	8.72
11	Subarnarekha	12.37	6.81	0.66	1.65	1.59
12	Sabarmati	3.81	1.93	1.35	0.12	0.09
13	Mahi	11.02	3.10	4.75	0.36	0.02
14	West flowing rivers of Kutch and Saurastra, including Luni	15.10	14.98	4.31	0.58	3.14
15	Narmada	45.64	34.50	6.60	16.72	0.46
16	Tapi	14.88	14.50	8.53	1.00	1.99
17	West flowing rivers from Tapi to Tadri	87.41	11.94	7.10	2.66	0.84
18	West flowing rivers from Tadri to Kanyakumari	113.53	24.27	10.25	2.31	1.70
19	Area of inland drainage in Rajasthan desert	Negligible		—	—	—
20	Minor rivers draining to Myanmar (Burma) and Bangladesh	31.00		0.31	—	—
	Total	1869.00	690.00	173.73	75.42	132.30

1086 km^3. In addition, about 10–15% of this quantity will be available as return flow from irrigation and domestic uses. As against this, the total withdrawal for all uses in the year 1990 was 552 km^3. The irrigation sector accounted for about 83% of this use.

Most of the groundwater development in India is up to a depth of about 50 meters and is financed mainly by institutional and private sources. In the shallow region, the wells include dug wells, dug and bore wells, and shallow tube wells. The exploitation in the deeper zone (50 m and more) is usually through the public sector for community irrigation, water supply, or industrial purposes. Groundwater development has been intensive in the plains of the Indus–Ganga basin of Punjab, Haryana, and Uttar Pradesh.

WATER RESOURCES REQUIREMENT

The increase in population and expansion of economic activities inevitably lead to rising demands for water for diverse purposes. Surface water is either used in-stream for hydropower and navigation or is diverted for off-stream use. A small part of diverted water returns to streams or aquifers, and the rest is consumed. Groundwater is used mostly for irrigation or for domestic requirements.

The Indian economy has traditionally been agriculturally based. At independence, it was of crucial importance to develop irrigation to increase agricultural production to make the country self-sustaining and to alleviate poverty. Accordingly, the irrigation sector was assigned a very high priority in the 5-year plans. Giant schemes such as the Bhakra Nangal, Hirakud, Damodar Valley, Nagarjunasagar, and Rajasthan Canal project were undertaken to increase irrigation potential and maximize agricultural production.

Long-term planning has to take into account the growth of population. A number of individuals and agencies have estimated the likely population of India by the year 2025 and 2050. According to the estimates adopted by NCIWRD(1), by the year 2025, the population is expected to be 1,333 million in the high growth scenario and 1,286 million in the low growth scenario. In the year 2050, the high rate of population growth is likely to result in about 1,581 million people, whereas the low growth projections place the number at nearly 1,346 million. Keeping in view the level of consumption, the estimated food grain requirements per capita per year would be 218 kg for the year 2025 and 284 kg for the year 2050. After considering factors such as feed requirement, losses in storage and transport, seed requirements, and buffer stock, the projected food-grain and feed demand for 2025 would be 320 million tonnes (high demand scenario) and 308 million tonnes (low demand scenario). The requirements for the year 2050 would be 494 million

Table 3. Groundwater Potential in the River Basins of India (Pro Rata Basis) in km^3/yr[a]

Sl. No.	Name of the Basin	Total Replenishable Groundwater Resources	Provision for Domestic, Industrial, and Other Uses	Available Groundwater for Irrigation	Net Draft	Balance of Groundwater Potential	Level of Groundwater Development, %
1	Brahmani with Baitarni	4.05	0.61	3.44	0.29	3.16	8.45
2	Brahmaputra	26.55	3.98	22.56	0.76	21.80	3.37
3	Chambal Composite	7.19	1.08	6.11	2.45	3.66	40.09
4	Cauvery	12.30	1.84	10.45	5.78	4.67	55.33
5	Ganga	170.99	26.03	144.96	48.59	96.37	33.52
6	Godavari	40.65	9.66	30.99	6.05	24.94	19.53
7	Indus	26.49	3.05	23.43	18.21	5.22	77.71
8	Krishna	26.41	5.58	20.83	6.33	14.50	30.39
9	Kutch & Saurashtra Composite	11.23	1.74	9.49	4.85	4.64	51.14
10	Chennai and South Tamil Nadu	18.22	2.73	15.48	8.93	6.55	57.68
11	Mahanadi	16.46	2.47	13.99	0.97	13.02	6.95
12	Meghna	8.52	1.28	7.24	0.29	6.95	3.94
13	Narmada	10.83	1.65	9.17	1.99	7.18	21.74
14	Northeast Composite	18.84	2.83	16.02	2.76	13.26	17.20
15	Pennar	4.93	0.74	4.19	1.53	2.66	36.60
16	Subarnarekha	1.82	0.27	1.55	0.15	1.40	9.57
17	Tapi	8.27	2.34	5.93	1.96	3.97	33.05
18	Western Ghat	17.69	3.19	14.50	3.32	11.18	22.88
	Total	431.43	71.08	360.35	115.21	245.13	31.97

[a] Reference 2.

tonnes (high demand scenario) and 420 million tonnes (low demand scenario).

Irrigation

The irrigated area in the country was only 22.6 million hectare (Mha) in 1950–1951. Food production was much below the requirement of the country, so due attention was paid to expansion of irrigation through surface and groundwater projects. The ultimate irrigation potential of India has been estimated at 140 Mha. Out of this, 76 Mha would come from surface water and 64 Mha from groundwater sources.

The quantity of water used for irrigation in the last century was of the order of 300 km^3 of surface water and 128 km^3 of groundwater. The estimates indicate that by the year 2025, the water requirement for irrigation would be 561 km^3 for a low demand scenario and 611 km^3 for a high demand scenario. These requirements are likely to increase to 628 km^3 for a low demand scenario and 807 km^3 for a high demand scenario by the year 2050.

Domestic Use

Water for the community water supply is the most important requirement, and it is about 5% of total water use. It is estimated that about 7 km^3 of surface water and 18 km^3 of groundwater are being used for community water supply in urban and rural areas. Along with the increase in population, another important change in water supply is a higher rate of urbanization. Per the projections, the higher the expected growth, the higher would be urbanization. It is expected that nearly 61% of the population will be living in urban areas by the year 2050 in a high growth scenario as against 48% in a low growth scenario.

Different organizations and individuals have given different norms for water supply in cities and rural areas. The figure adopted by NCIWRD (1) was 220 liters per capita per day (lpcd) for class I cities. For cities other than class I, the norms are 165 for the year 2025 and 220 lpcd for the year 2050. For rural areas, 70 lpcd and 150 lpcd have been recommended for the years 2025 and 2050. Based on these norms and the projection of population, it is estimated that by the year 2050, the water requirements per year for domestic use will be 90 km^3 for a low demand scenario and 111 km^3 for a high demand scenario. It is expected that about 70% of the urban water requirement and 30% percent of the rural water requirement will be met by surface water sources and the remaining from groundwater.

Hydroelectric Power

The hydropower potential of India has been estimated at 84,044 MW at a 60% load factor. At independence (1947), the installed capacity of hydropower projects was 508 MW which was about 37% of the total installed capacity. By the end of 1998, the installed hydropower capacity was about 22,000 MW which was 24.85% of the total installed capacity of 88,543 MW. The status of hydropower development in major basins is quite uneven. For example, the hydropower potential of Brahmputra Basin is about 34,920 MW which is 41.5% of the total potential of the country. Only 1.3% of this was developed by the year 1998. The percentage development in Indus was about 14.7% and 17.26% in Ganga. The major hydroelectric projects (above 750 MW) in India are Bhakra, Dehar, Koyna, Nagarjunasagar, Srisailam, Sharavarthy, Kalinadi, and Idukki.

Table 4. Annual Water Requirement for Different Uses in km^3

Uses	Year 1997–1998	Year 2010			Year 2025			Year 2050		
		Low	High	%	Low	High	%	Low	High	%
				Surface Water						
Irrigation	318	330	339	48	325	366	43	375	463	39
Domestic	17	23	24	3	30	36	5	48	65	6
Industries	21	26	26	4	47	47	6	57	57	5
Power	7	14	15	2	25	26	3	50	56	5
Inland navigation		7	7	1	10	10	1	15	15	1
Flood control		—	—	0	—	—	0	—	—	0
Environment (1) afforestation		—	—	0	—	—	0	—	—	0
Environment (2) ecology		5	5	1	10	10	1	20	20	2
Evaporation losses	36	42	42	6	50	50	6	76	76	6
Total	399	447	458	65	497	545	65	641	752	64
				Groundwater						
Irrigation	206	213	218	31	236	245	29	253	344	29
Domestic	13	19	19	2	25	26	3	42	46	4
Industries	9	11	11	1	20	20	2	24	24	2
Power	2	4	4	1	6	7	1	13	14	1
Total	230	247	252	35	287	298	35	332	428	36
				Total Water Use						
Irrigation	524	543	557	78	561	611	72	628	807	68
Domestic	30	42	43	6	55	62	7	90	111	9
Industries	30	37	37	5	67	67	8	81	81	7
Power	9	18	19	3	31	33	4	63	70	6
Inland navigation	0	7	7	1	10	10	1	15	15	1
Flood control	0	0	0	0	0	0	0	0	0	0
Environment (1) afforestation	0	0	0	0	0	0	0	0	0	0
Environment (2) ecology	0	5	5	1	10	10	1	20	20	2
Evaporation losses	36	42	42	6	50	50	6	76	76	7
Total	629	694	710	100	784	843	100	973	1180	100

Industrial Water Requirement

Rough estimates indicate that present water use in the industrial sector is of the order of 15 km^3. The water use by thermal and nuclear power plants with installed capacities of 40,000 MW and 1,500 MW (1990 figures), respectively, is estimated at about 19 km^3. In view of the shortage of water, the industries are expected to switch over to water efficient technologies. If the present rate of water use is continued, the water requirement for industries in the year 2050 would be 103 km^3, and if water saving technologies are adopted on a large scale, the requirement is likely to be nearly 81 km^3.

Water Quality

Water quality issues are gaining recognition as river waters are getting heavily polluted in many places and groundwater quality at many places is beginning to deteriorate. Among the main concerns for water quality are (1) arsenic in drinking water in West Bengal (affecting about 5 million persons); (2) fluoride levels are high in Andhra Pradesh, Gujarat, Haryana, Karnataka, Punjab, Rajasthan, Tamil Nadu, and Madhya Pradesh (affecting 14 million persons); (3) iron levels are high in the northeastern and eastern parts of the country (affecting 29 million persons); and (4) salinity is high in Gujarat, Haryana, Karnataka, Punjab, Rajasthan, and Tamil Nadu.

Total Water Requirements

Based on the various studies, the total annual requirement for freshwater from various sectors is estimated at about 1050 billion m^3 by 2025 A.D. The breakdown is shown in Table 4. Estimates show that this demand will be met by harnessing 700 billion m^3 of surface water and 350 billion m^3 of groundwater.

FLOODS AND DROUGHTS

Floods and drought are two water-related natural calamities that are recurrent phenomena in India.

Floods

Floods are the most frequent natural calamities faced by India in different magnitudes, year after year. About 80–90% of the annual precipitation in India takes place during 4 months of monsoon; this is also the season when floods are mostly experienced. The main causes of floods in India are inadequate capacity of river sections to contain high flows, silting of river beds, and drainage congestion. Floods are also caused by cyclones, cloud bursts, and snow/glacier melt. Based on the causes of floods, the country can be divided into four basins/regions. In the Brahmaputra and Barak Basin, the main problems in the state of Assam are inundation caused by overflow of the Brahmaputra and its tributaries and erosion along the river bank. In the Ganga Basin, the flooding problem is confined to small portions in the middle and terminal reaches. In general, the severity of the problem increases from west to east and from south to north. The worst flood affected states in the Ganga Basin are Uttar Pradesh, Bihar, and West Bengal. In Uttar Pradesh, flooding is confined to the eastern districts: the rivers that cause flooding include the Sarada, the Ghagra, the Rapti, and the Gandak. The major causes here are drainage congestion and bank erosion. North Bihar is in the grip of floods almost every year due to spillage of rivers. In West Bengal, floods are caused by drainage problems as well as tidal effects. Flood problems are not very severe in the rivers in the northwest part of India, and here the cause of floods is inadequate drainage capacity. Flash floods are also experienced here. Flash floods are characterized by a sharp rise and recession of river flow, and the damage is mainly due to the sudden rise of the river, leaving a very short time for evacuation. In the central and southern part of the country, floods occur in the Narmada, Tapi, Godavari, Mahanadi, and Krishna Rivers. The coastal regions of Orissa, Andhra Pradesh, and Tamil Nadu also face problems due to cyclones.

According to the estimates, the average area annually affected by floods is 7.52 Mha of which the agricultural area is 3.52 Mha. According to NCIWRD (1), during the second half of the twenty-first century, on average, 1515 lives and 95,285 head of cattle were lost every year. The area liable to floods was assessed at 40 Mha which is about one-eighth of the geographical area of the country.

Drought

Normally, the term drought is used for a shortage of precipitation or water at a place when it is expected. Meteorologic drought is said to occur at a place when there is more than a 25% decline in rainfall from the normal value. Meteorologic drought over an extensive area triggers hydrologic drought which is the condition of significant depletion of water resources in rivers, reservoirs, lakes, and springs. A marked depletion of soil moisture and precipitation leads to agricultural drought. Based on detailed drought studies for the country, 99 districts have been identified as drought prone. The total geographic area of these districts is about 1,081,131 sq. km.

BIBLIOGRAPHY

1. NCIWRD. (1999). *Integrated Water Resources Development—A Plan for Action*. Report of *The National Commission for Integrated Water Resources Development.* Govt. of India, Ministry of Water Resources, New Delhi.
2. IWRS. (1998). *Theme Paper "Five decades of water resources development in India"*. Indian Water Resources Society, Roorkee.

READING LIST

CGWB. (1995). *Ground Water Resources of India, Central Ground Water Board*, Govt. of India, New Delhi.

IWRS. (1999). *Theme Paper "Water: Vision 2050"*. Indian Water Resources Society, Roorkee.

Rao, K.L. (1973). *India's Water Wealth*. Orient Longman Limited, New Delhi.

WATER INFRASTRUCTURE AND SYSTEMS

NEIL S. GRIGG
Colorado State University
Fort Collins, Colorado

Water infrastructure is the set of constructed elements that operate within water resources systems and interact with environmental elements, or natural systems, such as watersheds, streams, lakes, and aquifers (see article on WATER RESOURCES MANAGEMENT). Table 1 gives examples of structures used in constructed systems for different purposes of water management.

In a watershed, sets of these structures may control water for different purposes. Most apparent might be a multiple-purpose reservoir, that has a dam and hydroelectric plant connected to electric transmission lines. Below might be a diversion dam to enable a canal to take irrigation water from the stream, if it is located in a dry region. There might be a re-regulating reservoir with locks for navigation. Dams are a special class of hydraulic structure, and they provide barriers to the flow to enable reservoirs to store water.

Dams have a number of special components, including service and emergency spillways, outlet structures,

Table 1. Structures for Water Resources Management

Purpose	Conveyance	Storage	Treatment[a]	Pump or Generate	Control
Water supply	Water pipes	Reservoir	WTP	Pumps	Valves
Water quality	Sewer pipes	Tanks, ponds	WWTP	Sewage pumps	Gates
Flood control	Channels	Detention ponds	SWTP	Flood pumps	Inlets
Hydropower	Penstocks	Reservoir	–	Turbines	Gates
Navigation	Locks	Reservoir	–	–	Valves
Environment	Rivers	Natural storage	Wetland	–	–

[a] WTP = water treatment plant; WWTP = wastewater treatment plant; SWTP = stormwater treatment plant.

and drains. Outlet works might discharge to a municipal water supply treatment plant, and further downstream a wastewater treatment plant might discharge to the stream.

A reservoir is a lake where water is stored, either natural or dammed artificially. Reservoirs require special management attention because they control flow throughout river systems. As a consequence, they tend to be both strategically important and controversial. Great dams such as Hoover Dam on the Colorado River, Aswan Dam on the Nile River, and Three Gorges Dam on the Yangtze River in China influence whole regions and even nations.

To control flows, there will be valves, gates and spillways, but they are mostly invisible. Wells might provide water for users such as farming operations or irrigation of golf courses.

Diversion structures on a stream change its direction or flow patterns. They might include intakes, boat chutes, or river training structures. A fish ladder is a special kind of diversion structure meant to provide migrating fish species with methods for swimming upstream, usually around a dam.

The structures and components that are associated with conveyance systems include natural open channels and canals, pipelines, pipe networks and sewers, bridges, and levees. Open channels are either natural or man-made. A river is an example of a natural open channel. Levees are part of conveyance systems that form the banks of channels to protect land areas from flooding. A canal or lined drainage ditch would be an example of a man-made channel. A lock raises or lowers boats and ships up or down a river.

Pipes, or closed conduits, can be classified as tunnels, transmission pipelines, pressure pipe networks, or sewer networks.

A bridge is part of a conveyance system in the sense that it provides a method for separating a stream bed from a road, a rail line, or some other structure crossing a stream. A culvert has a function somewhat like a bridge that enables a separation of grades between a small stream and a roadway or other embankment above.

Hydroelectric plants are facilities that generate electric energy from water discharge. They include turbines, generators, and associated hydraulic components.

Pumps impart energy to water and raise it in elevation or pressure. There are several different types of pumps, built mainly around the centrifugal or turbine pump format, and they have numerous applications.

Valves, gates and spillways are control devices for conveyance systems or dams. Spillways are used as emergency overflow devices to protect dams. There are a number of types, normally classified as the service spillway or emergency spillway. Water flowing over a spillway normally flows into some type of energy dissipation device to avoid erosion downstream.

Coordinating the operation of these components places the focus on the water resources *system*, as opposed to individual pieces. A water resources system is a combination of constructed water control facilities and environmental elements that work together to achieve water management purposes (see article on WATER RESOURCES MANAGEMENT). A constructed system consists of structural facilities that control water flow and quality, and a natural water resources system comprises environmental elements such as watersheds, stream channels, and groundwater systems.

OVERVIEW AND TRENDS IN THE INTERNATIONAL WATER MARKET

RANDY T. PIPER
Dillon, Montana

OVERVIEW AND TRENDS

The term "water market" is a slippery concept. Like water, the term water market is free flowing and can take a variety of forms, depending on the cultural environment and who invokes the term. In "The Importance of Getting Names Right: The Myth of Markets for Water," Joseph Dellapenna argues for greater precision when using the language of water markets (1). Consequently, when we speak of water markets, we need to identify specifically the rules that governments employ and the roles that governments play in defining, divesting, and defending water markets. In most geographic settings, water markets do not exist without some formal role for government administration.

In Brazil, local water markets have evolved with minimal government oversight. Dating from 1854 in the Cariri region of northeast Brazil, sugar cane farmers established a water market. The farmers themselves defined the amount of water for each farm. Moreover, farmers developed water measuring devices and monitoring rules independent of formal government administration. However, this water market limits who can own the water, that is, water is neither auctioned to the highest bidder nor sold to nonfarmers who would divert the water for fisheries or for nonagricultural urban consumption (2).

Technology can and does play an important role in water markets. For example, technology can make bottled water more affordable to consumers and more aesthetically appealing to consumers. In 2001, Norwegian water in American markets gained fame and market share largely because of the package design. Technology can improve the quality of water, from wastewater treatment plants to desalination. For example, 120 countries now employ some form of desalting equipment (3). At the 2001 Bahrain meeting of the World Congress of International Desalination Association, the primary focus was on the technologies of water markets, not the institutions of water markets (4). Technology can improve the institutions of water markets and make it more cost-effective for governments and nongovernments to define and enforce the boundaries of water markets and to monitor the consumption and transactions of water markets (5).

Technology in the narrow sense can and does improve the product and package of water and water markets. Technology in the broad sense can and does improve the institutions of water markets. However, the technology and science of water markets per se are not the paramount challenges in many international markets. In 2000, the 10th Stockholm Water Symposium reached this conclusion: “Long-term challenges in water management are not so much linked to classical scientific and technical aspects as to institutional innovations” (6).

In “Water Challenge and Institutional Response,” R. Maria Saleth and Ariel Dinar reviewed the institutional innovations of 11 countries: Australia, Brazil, Chile, China, India, Israel, Mexico, Morocco, South Africa, Spain, and Sri Lanka. These World Bank researchers found several commonalities among the 11 countries. For example, countries have reached a supply limit of water resource development and are now focusing on the rules of allocation of their supplies. Countries have also begun to experiment with decentralization and privatization. Among the countries (Australia, Israel, Mexico, and Spain) that have national water plans, Israel and Mexico delegate much of the allocation administration to subnational levels that are organized primarily along hydrogeologic lines in the form of basins (7). These river basin organizations then may develop rules of participation and the roles of market-oriented institutions for allocation. According to R. Maria Saleth and Ariel Dinar, “Countries have begun to recognize the functional distinction between centralized mechanisms needed for coordination and enforcement and decentralized arrangements needed for user participation and local level solutions” (8).

In the urban water subsector, another institutional innovation is privatization or corporatization of water and water infrastructures. Privatization can take a number of forms, from management, service, and lease contracts to build, own, operate, and transfer contracts to full private ownership under a regulatory regime. In the 1990s, several transnational corporations acquired water firms and/or formed subsidiaries to compete in international markets. This trend is illustrated by noting the home country’s parent company followed by its water subsidiary: England’s Kelda and Alcontrol, France’s Vivendi and Vivendi Environment, Germany’s RWE and Bovis Thames, and America’s Enron and Azurix. In comparison to the 1990–1994 period, during the 1995–1999 period, the monetary value of water and sanitation privatization projects increased by 100% to approximately $9 billion in Latin America and approximately $9 billion in East Asia and Pacific regions, by more than 1000% to approximately $1 billion in sub-Saharan Africa and approximately $2 billion in Europe and Central Asia regions, and by more than 2000% to approximately $5 billion in the Middle East and North Africa regions (9). The ownership and names of the transnational corporations will change based on the ebb and flow of supply–demand interactions. However, the role of any newly named corporation will remain the same: to maximize its returns on equity for various privatization projects.

Some institutional analysts consider the two trends of decentralization and privatization innovative and preferable to centralized national ownership and management. However, other institutional analysts view these trends as less than innovative and undesirable. A statement from the conclusions of the 10th Stockholm Water Symposium captures this sentiment: “The economically based approach that dominates at present, and has the market as a recommended mode of interpretation, was seen as having severe limitations when it comes to acceptable ways of coping with the emerging water crisis” (10).

Even when countries can overcome the cultural and linguistic barriers to crafting water market-oriented institutions, these countries still encounter other barriers, primarily cost-effective technology. Cost-effective technology exists for the defining, defending, and divesting of water *quantity*. For example, affordable geographic information systems (GIS) allow governments and nongovernments to monitor whether water users are “rustling water” and consuming illegal quantities. However, when the focus turns to water *quality*, the costs of technology for measuring and monitoring pollution trespassers increase dramatically (11). The wealthier countries of the European Union, South America, Asia, and Middle East can afford some of these emerging technologies—both for water quantities and qualities—but less affluent countries cannot allocate sufficient resources to create water markets, whether for quantities or qualities.

SUMMARY

Generalizations remain precarious. Institutional successes at one national level may not transfer to other national levels within the same region. Institutional successes at one national level may not transfer to subnational levels within the same country. The barriers to diffusion of institutional innovations are numerous and varied, flowing from economic, political, and social to technological and linguistic barriers.

Governments will continue with cross-country agreements that more clearly define their respective rights, rules, and roles. National governments may play a more active role in achieving hydrosolidarity among themselves, but this external hydrosolidarity does not preclude internal hydrosolidarity in the form of denationalization and water-market-oriented institutions. In a few countries, governments may well continue to finance, manage, and own water resources and view water as a nationalized “means of production.” However, as the World Bank’s R. Maria Saleth and Ariel Dinar observe, “But nowhere, even in China, is the state’s absolute ownership of water established to exclude private use rights” (8).

Many governments will still continue to denationalize and privatize water resources. In urban areas, private corporations will continue to build and manage water systems, while the government finances the construction and capital costs, and consumers finance the operating costs via user fees. When urban consumers cannot finance the full capital and operating costs, governments will experiment with full-cost pricing mechanisms, transparency of subsidies, and “water vouchers.” In rural areas, national governments may allocate some of the financing, managing, and owning to local governments. In turn, localized governments may experiment with a variety of market-oriented institutions. The structure of these institutions could include allowing current owners and users to sell their water for nonagricultural uses.

During the next few decades, it is unlikely that many governments will permit the following scenario for international water markets to develop. Both national and local governments give full private property rights to agriculture and/or urban "owners." With full rights, these owners could then sell their water to whomever they wanted. Willing higher and highest bidding buyers and consumers would include rural and urban residents, transnational corporations, environmental groups, and even water speculators (12). Once governments and nongovernments have experimented with internal hydrosolidarity and learned the lessons from the multiple forms of water-market-oriented institutions such as water banks and water vouchers, governments and nongovernments may well transfer this knowledge to the realm of external hydrosolidarity.

Finally, when we interpret "international water markets" and we more precisely speak in terms of water-market-oriented institutions, our language and logic must accurately reflect the multiple roles that governments play in the creation and evolution of international water markets.

BIBLIOGRAPHY

1. Dellapenna, J.W. (2000). The importance of getting names right: the myth of markets for water. *William and Mary Environ Law Policy Law Rev.* **25**(Winter): 317–378.
2. Kemper, K.E. (1999). Jose Yarley de Brito Goncalves, and Francisco William Brito Bezerra, Water allocation and trading in the Cariri region—Ceara, Brazil. In: *Institutional Frameworks in Successful Water Markets: Brazil, Spain, and Colorado, USA*. M. Marino and K.E. Kemper (Eds.). World Bank Technical Paper 427, February, Washington, DC, pp. 1–9.
3. Gleick, P.H. (2000). *The World's Water, 2000–2001: The Biennial Report on Freshwater Resources*. Island Press, Covelo, CA, p. 96.
4. International Desalination Association. Accessed at http://www.ida.bm on August 24, (2001).
5. Landry, C.J. (2001). The role of geographic information systems in water rights management. In: *The Technology of Property Rights*. L.A. Terry and P.J. Hill (Eds.). Rowman & Littlefield Publishers, Lanham, MD, p. 23.
6. Stockholm International Water Institute. *The Final Overall Conclusions: The 10th Stockholm Water Symposium*. August 14–17, 2000, p. 1. Accessed at http://www.siwi.org/sws2000/eng/sws2000conclusions.html on September 24, 2001.
7. Maria, R.S. and Ariel, D. (1999). *Water Challenge and Institutional Response: A Cross-Country Perspective*. World Bank Policy Research Working Paper 2045, January, Washington, DC, p. 36.
8. Saleth and Dinar, p. 37.
9. The Water Page. Accessed at http://www. thewaterpage.com/ppp_new_main.htm on October 22, 2001.
10. Stockholm International Water Institute. *The Final Overall Conclusions: The 10th Stockholm Water Symposium*. August 14–17, 2000, p. 3. Accessed at http://www.siwi.org/sws2000/eng/sws2000conclusions.html on September 24, 2001.
11. Yandle, B. (2001). Legal foundations for evolving property rights technologies. In: *The Technology of Property Rights*. T.L. Anderson and P.J. Hill (Eds.). Rowman & Littlefield Publishers, Lanham, MD, pp. 12–18.
12. For a review of the dynamics of domestic and international water markets, see *Water Marketing—The Next Generation*. T.L. Anderson and P.J. Hill (Eds.). 2000. *The Political Economy of Water Pricing Reforms* Ariel Dinar (Ed.). 2000, and Water Bank at http://www.WaterBank.com.

BEST MANAGEMENT PRACTICES FOR WATER RESOURCES

Mohammad N. Almasri
An-Najah National University
Nablus, Palestine

Jagath J. Kaluarachchi
Utah State University
Logan, Utah

INTRODUCTION

Best management practices (BMPs) for water resources are affordable, practical, and effective methods for eliminating or reducing the movement of pollutants from the ground surface to groundwater or surface waterbodies. Such pollutants include nutrients and pesticides, among others, that may have adverse effects on human health. In addition, BMPs can be viewed as methods to ensure a sustainable and safe yield of water resources through optimal conjunctive management of these resources. The main premise in developing BMPs is to achieve a balance between the goal of protecting water resource quality, for instance, and the economic and social ramifications of adopting these BMPs. Before adopting BMPs for a specific site, it is essential to understand the BMPs thoroughly and to ensure their technical feasibility through research, experience, field pilot studies, and surveys. In addition to technical feasibility, BMPs should be economically well founded through cost-effectiveness analysis and acceptable via adopting and maintaining them. BMPs may consist of single practices or combinations of them based on the problem at hand and the social and economic ramifications. BMPs for groundwater resources are important because they preserve and protect groundwater quality and quantity. Preservation of groundwater resources minimizes public health problems and maintains the social and economic values of groundwater. For instance, the cost of mitigating polluted groundwater can be extremely high and by adopting effective BMPs, such mitigation costs can be reduced dramatically.

This document focuses on BMPs to protect and preserve groundwater resources from nitrate contamination.

NITRATE CONTAMINATION OF GROUNDWATER

Groundwater is the primary source of drinking water in many parts of the world and the sole supply of potable water in many rural communities (1). There is increasing could include allowing current owners and users to sell their water for nonagricultural uses.

During the next few decades, it is unlikely that many governments will permit the following scenario for international water markets to develop. Both national and local governments give full private property rights to agriculture and/or urban "owners." With full rights, these owners could then sell their water to whomever they wanted. Willing higher and highest bidding buyers and consumers would include rural and urban residents, transnational corporations, environmental groups, and even water speculators (12). Once governments and nongovernments have experimented with internal hydrosolidarity and learned the lessons from the multiple forms of water-market-oriented institutions such as water banks and water vouchers, governments and nongovernments may well transfer this knowledge to the realm of external hydrosolidarity.

Finally, when we interpret "international water markets" and we more precisely speak in terms of water-market-oriented institutions, our language and logic must accurately reflect the multiple roles that governments play in the creation and evolution of international water markets.

BIBLIOGRAPHY

1. Dellapenna, J.W. (2000). The importance of getting names right: the myth of markets for water. *William and Mary Environ Law Policy Law Rev.* **25**(Winter): 317–378.
2. Kemper, K.E. (1999). Jose Yarley de Brito Goncalves, and Francisco William Brito Bezerra, Water allocation and trading in the Cariri region—Ceara, Brazil. In: *Institutional Frameworks in Successful Water Markets: Brazil, Spain, and Colorado, USA*. M. Marino and K.E. Kemper (Eds.). World Bank Technical Paper 427, February, Washington, DC, pp. 1–9.
3. Gleick, P.H. (2000). *The World's Water, 2000–2001: The Biennial Report on Freshwater Resources*. Island Press, Covelo, CA, p. 96.
4. International Desalination Association. Accessed at http://www.ida.bm on August 24, (2001).
5. Landry, C.J. (2001). The role of geographic information systems in water rights management. In: *The Technology of Property Rights*. L.A. Terry and P.J. Hill (Eds.). Rowman & Littlefield Publishers, Lanham, MD, p. 23.
6. Stockholm International Water Institute. *The Final Overall Conclusions: The 10th Stockholm Water Symposium*. August 14–17, 2000, p. 1. Accessed at http://www.siwi.org/sws2000/eng/sws2000conclusions.html on September 24, 2001.
7. Maria, R.S. and Ariel, D. (1999). *Water Challenge and Institutional Response: A Cross-Country Perspective*. World Bank Policy Research Working Paper 2045, January, Washington, DC, p. 36.
8. Saleth and Dinar, p. 37.
9. The Water Page. Accessed at http://www. thewaterpage.com/ppp_new_main.htm on October 22, 2001.
10. Stockholm International Water Institute. *The Final Overall Conclusions: The 10th Stockholm Water Symposium*. August 14–17, 2000, p. 3. Accessed at http://www.siwi.org/sws2000/eng/sws2000conclusions.html on September 24, 2001.
11. Yandle, B. (2001). Legal foundations for evolving property rights technologies. In: *The Technology of Property Rights*. T.L. Anderson and P.J. Hill (Eds.). Rowman & Littlefield Publishers, Lanham, MD, pp. 12–18.
12. For a review of the dynamics of domestic and international water markets, see *Water Marketing—The Next Generation*. T.L. Anderson and P.J. Hill (Eds.). 2000. *The Political Economy of Water Pricing Reforms* Ariel Dinar (Ed.). 2000, and Water Bank at http://www.WaterBank.com.

BEST MANAGEMENT PRACTICES FOR WATER RESOURCES

Mohammad N. Almasri
An-Najah National University
Nablus, Palestine

Jagath J. Kaluarachchi
Utah State University
Logan, Utah

INTRODUCTION

Best management practices (BMPs) for water resources are affordable, practical, and effective methods for eliminating or reducing the movement of pollutants from the ground surface to groundwater or surface waterbodies. Such pollutants include nutrients and pesticides, among others, that may have adverse effects on human health. In addition, BMPs can be viewed as methods to ensure a sustainable and safe yield of water resources through optimal conjunctive management of these resources. The main premise in developing BMPs is to achieve a balance between the goal of protecting water resource quality, for instance, and the economic and social ramifications of adopting these BMPs. Before adopting BMPs for a specific site, it is essential to understand the BMPs thoroughly and to ensure their technical feasibility through research, experience, field pilot studies, and surveys. In addition to technical feasibility, BMPs should be economically well founded through cost-effectiveness analysis and acceptable via adopting and maintaining them. BMPs may consist of single practices or combinations of them based on the problem at hand and the social and economic ramifications. BMPs for groundwater resources are important because they preserve and protect groundwater quality and quantity. Preservation of groundwater resources minimizes public health problems and maintains the social and economic values of groundwater. For instance, the cost of mitigating polluted groundwater can be extremely high and by adopting effective BMPs, such mitigation costs can be reduced dramatically.

This document focuses on BMPs to protect and preserve groundwater resources from nitrate contamination.

NITRATE CONTAMINATION OF GROUNDWATER

Groundwater is the primary source of drinking water in many parts of the world and the sole supply of potable water in many rural communities (1). There is increasing awareness that groundwater is vulnerable to contamination from domestic, industrial, and agricultural wastes (2). Sources of groundwater contamination are widespread and include accidental spills, landfills, storage tanks, pipelines, agricultural activities, and many other sources. Of these sources, agriculture activities produce nonpoint source pollution in small to large watersheds especially due to nitrogen in fertilizers and various carcinogenic substances found in pesticides. Due to its high mobility, nitrate is the primary nitrogen species lost from soils by leaching (3). Groundwater contamination by nitrate has been confirmed throughout the United States (4,5) and worldwide (6). Elevated nitrate concentration in drinking water can cause *methemoglobinemia* in infants and stomach cancer in adults (7–9). Therefore, the U.S. Environmental Protection Agency (U.S. EPA) has established a maximum contaminant level (MCL) of 10 mg/L NO_3-N (10). Nitrate may indicate the presence of bacteria, viruses, and protozoa in groundwater if the source of nitrate is animal waste or effluent from septic tanks. Likewise, nitrate contamination of surface water has health and environmental effects. Transport of nitrate to surface water occurs mainly via discharge of groundwater during baseflow conditions (11). Hence, prevention of groundwater contamination protects surface water quality as well.

Agricultural activities are probably the most significant anthropogenic source of nitrate contamination in groundwater. Nitrogen is a vital nutrient for plant growth. Nevertheless, when nitrogen-rich fertilizer application exceeds the plant demand and the denitrification capacity of the soil, nitrogen can leach to groundwater usually in the form of nitrate (12). Many studies have shown high correlation between agricultural land use and nitrate concentration in groundwater (6,13,14). Agricultural practices result in nonpoint source pollution of groundwater and the effects of these practices accumulate over time (7). Nonpoint sources of nitrogen from agricultural practices include fertilizer, dairy farms, manure application, and leguminous crops (15). Point sources of nitrogen such as septic tanks and dairy lagoons contribute to nitrate pollution of groundwater.

BMPs TO MINIMIZE NITRATE CONTAMINATION OF GROUNDWATER

Identification of areas with heavy nitrogen loadings from point and nonpoint sources is important for land use planners and environmental regulators. Once such high-risk areas have been identified, BMPs can be implemented to minimize the risk of nitrate leaching to groundwater. The knowledge of the spatial distribution of nitrogen loading can identify the areas where groundwater needs to be protected. This assessment is of great importance in designating areas that can benefit from pollution prevention and monitoring programs. In large areas such as large watersheds or basins, proper understanding of on-ground nitrogen loading from different sources and corresponding transformations in the soil are needed before areas can be designated for future protective measures (16).

BMPs developed to protect groundwater quality are improvements in agricultural practices and land use patterns (17). Improving agricultural practices focuses on tillage practices, crop rotation, rational fertilizer and manure applications, vegetation cover, and irrigation management (18). For instance, on-ground application of manure should be in accordance with agronomic requirements and the nitrogen mass in the soil (19). Land use changes include reevaluation of land use distribution and changes in the land use in areas that are vulnerable to contamination. However, this option may not be feasible in many instances due to competing stakeholder concerns arising from economic constraints and these constraints may, sometimes, override environmental concerns. BMPs imply conflicting objectives. On the one hand, the main aim of a BMP is to reduce nitrate concentrations in groundwater below the MCL. On the other hand, a protection alternative should consider the minimization of economic losses incurred from implementing the protection alternatives. Therefore, a multicriteria decision analysis might be developed to determine the most viable BMP that accounts for the different decision criteria.

Fertilizer and manure applications are required to replace crop land nutrients that have been consumed by previous plant growth. They are essential for economical yields. However, excess fertilizer and manure use and poor application methods can lead to nitrate pollution of groundwater and surface waterbodies. Therefore, it is important to match nitrogen applications to crop uptake to minimize nitrate leaching. This is the core of the BMPs that address the problem of nitrate contamination of groundwater resources.

Effective BMPs limit movement by minimizing the quantity of nutrients available for leaching below the root zone. This is achieved by developing an inclusive nutrient management plan that uses field studies, soil and groundwater models, uses only the types and amounts of nutrients necessary to produce the crop, applies nutrients at the proper times and by proper methods, implements additional farming practices to reduce nutrient losses, and follows proper procedures for fertilizer storage and handling. In the following, the main BMPs pertaining to fertilizer and manure applications for reducing nitrate occurrences in the groundwater are summarized based on the work provided by Waskom (20), the EPA (21), Almasri (22), and Almasri and Kaluarachchi (23), among other studies listed in the references.

Application Rates and Fertilizer Types

One component of a comprehensive nutrient management plan is determining proper fertilizer application rates (24). The objective is to limit fertilizer application to an amount sufficient to achieve a realistic crop yield. Allowing for other nitrogen sources in the soil is also part of the concept. Previous legume crops, irrigation water, manure, and soil organic matter all contribute nitrogen to the soil. Along with soil samples and fertilizer credits from other sources, nitrogen fertilizer recommendations are based on yield goals established by crop producers. Yield expectations are established for each crop and field based on soil properties, available moisture, yield history, and management level.

Applying the appropriate form of nitrogen fertilizer can reduce leaching.

Nitrate forms of nitrogen fertilizer are readily available to crops but are subject to leaching losses. Nitrate fertilizer use should be limited when the leaching potential is moderate to high. In these situations, ammonium nitrogen fertilizers should be used because they are not subject to immediate leaching. However, ammonium nitrogen transforms rapidly into nitrate via the nitrification process, especially when soils are warm and moist. More slowly available nitrogen fertilizers should be used in these conditions. Nitrification inhibitors can also delay the conversion of ammonium to nitrate under certain conditions (23).

Fertilizer Application Timing

Nitrogen fertilizer applications should be timed in periods of maximum crop uptake. Fertilizer applied in the fall causes groundwater quality degradation. Partial application of fertilizer in the spring, followed by small additional applications as needed, can improve nitrogen uptake and reduce leaching risk.

Fertilizer Application Methods

Fertilizer application equipment should be checked at least once a year and should be properly calibrated to ensure that the recommended amount of fertilizer is applied. Fertilizer placement in the root zone can greatly enhance plant nutrient uptake and minimize leaching losses. Subsurface applied or incorporated fertilizer should be used instead of surface broadcast fertilizer. An efficient application method for fertilizer is to place dry fertilizer into the ground in bands closer to the seed, so that it can be recovered by the crop very efficiently. To reduce losses through surface runoff and volatilization, all fertilizers should be mechanically incorporated into the soil. Fertilizer application to frozen ground should be avoided and limited on slopes and areas with high runoff or overland flow.

Irrigation Water Management

Irrigation water should be managed to maximize efficiency and minimize runoff or leaching. Irrigated crop production has the greatest potential for source water contamination because of the large amount of water applied and the movement of nutrients via irrigation flux. Nitrate can leach into groundwater when excess water is applied to fields. Irrigation systems, such as sprinklers, low-energy precision applications, surges, and drips, allow producers to apply water uniformly and with great efficiency. Efficiency can also be improved by using delivery systems such as lined ditches and gated pipe, as well as reuse systems such as field drainage recovery ponds that efficiently capture sediment and nutrients. Gravity-controlled irrigation or furrow runs should be shortened to prevent overwatering at the top of the furrow before the lower end is adequately watered.

Manure Application Reduction

Manure is an excellent source of plant nutrients, especially nitrogen. Nevertheless, excessive applications of manure lead to nitrogen buildup in the soil that eventually leads to groundwater pollution. Prior to application, manure should be analyzed to determine its nutrient content. The application of increasing quantities of manure to the same land area may result in groundwater quality problems and adverse environmental consequences. If it turns out that the land base is insufficient for manure application, then BMPs should be introduced. In the following subsections, BMPs pertaining to dairy manure such as reducing herd size, composting/exporting manure, or implementing feeding strategies to reduce nutrient content in the excrement are summarized (25).

Dairy Herd Size Reduction. Manure production and loading are functions of dairy herd size, so downsizing the dairy herd is apparently the most straightforward and effectual BMP that minimizes manure loading (25). Nonetheless, such an option has serious economic ramifications that may prohibit the adoption of this alternative.

Manure Composting/Exporting. Manure exporting is a viable alternative because it does not involve herd size reduction. For manure to be exported, it should be composted. Composting is a biological process that converts organic manure to a more stable material such as humus. Composting is the aerobic decomposition of organic matter by certain microorganisms such as bacteria that consume oxygen and use nutrients such as carbon, nitrogen, phosphorus, and potassium as they feed on the organic waste. They produce heat, carbon dioxide, and water vapor. Heat destroys pathogens and weed seeds. The resulting composted manure is humus-like organic material, fine-textured, low in moisture, and has a nonoffensive earthy odor. Because of the carbon dioxide and water vapor that escape during the process, the resulting compost can be approximately half the volume and weight of the original material. Efficient composting requires maintaining proper temperatures and oxygen levels in the composting material (26). Warm temperatures help microorganisms to grow best. Elevated temperatures are needed to destroy pathogens and weed seeds. Bulking materials can be mixed with manure to provide structural support when manure solids are too wet to maintain air spaces within the composting pile. Composting is best accomplished by placing the manure in windrows. The windrows can be turned and mixed periodically to maintain oxygen levels for proper composting. The windrows may be 3 to 6 feet high and 10 to 20 feet wide, depending mostly on the type of machinery used in turning them. The width of the windrow must allow air movement into the manure to introduce oxygen. Composting reduces spreading costs, facilitates land application, eliminates manure odors, and promotes manure transporting and marketing. Composted manure can be used mainly on lawns and gardens to keep and add organic matter in the soil. Organic matter provides good aeration porosity, holds more water, and releases important nutrients.

Improve Dairy Cow Diets. The future challenge for dairy producers and nutritionists will be to formulate rations properly for high milk production levels while simultaneously minimizing the environmental impact of excessive nitrogen content in the excrement. A properly formulated ration that precisely meets the cow's requirements for milk production, maintenance, and growth minimizes excessive nitrogen in the manure. Many dairy producers overfeed crude protein to support high levels of milk production. However, this results in excessively high nitrogen in the excrement. The protein that is not used for milk production or maintenance and growth is excreted as urea or organic-N. This practice has an adverse effect on the environment. Feeding strategies can be adopted to minimize manure nitrogen content while maintaining the same milk production levels. Van Horn (27) showed that when cows are precisely fed to meet the National Research Council recommendations (28), the nitrogen content in manure dropped by 14%.

BIBLIOGRAPHY

1. Solley, W.B., Pierce, R.R., and Perlman, H.A. (1993). Estimated use of water in the United States in 1990. U.S. Geological Survey Circular 1081.
2. Almasri, M.N. and Kaluarachchi, J.J. (2004). Assessment and management of long-term nitrate pollution of ground water in agriculture-dominated watersheds. *J. Hydrol.* **295**: 225–245. doi:10.1016/j.jhydrol.2004.03.013.
3. Tesoriero, A., Liecscher, H., and Cox, S. (2000). Mechanism and rate of denitrification in an agricultural watershed: Electron and mass balance along ground water flow paths. *Water Resour. Res.* **36**(6): 1545–1559.
4. Spalding, R.F. and Exner, M.E. (1993). Occurrence of nitrate in groundwater—a review. *J. Environ. Qual.* **22**: 392–402.
5. Nolan, B.T. (2001). Relating nitrogen sources and aquifer susceptibility to nitrate in shallow ground waters of the United States. *Ground Water* **39**(2): 290–299.
6. Hudak, P.F. (2000). Regional trends in nitrate content of Texas groundwater. *J. Hydrol.* **228**: 37–47.
7. Addiscott, T.M., Whitmore, A.P., and Powlson, D.S. (1991). Farming, fertilizers and the nitrate problem. CAB International, Wallingford, United Kingdom.
8. Lee, Y.W. (1992). Risk assessment and risk management for nitrate-contaminated groundwater supplies. Unpublished Ph.D. Dissertation. University of Nebraska, Lincoln, NE.
9. Wolfe, A.H. and Patz, J.A. (2002). Reactive nitrogen and human health: Acute and long-term implications. *Ambio* **31**(2): 120–125.
10. U.S. Environmental Protection Agency. (1995). Drinking water regulations and health advisories. Office of Water, Washington, DC.
11. Bachman L.J., Krantz, D.E., and Böhlke. J, (2002). Hydrogeologic framework, ground-water, geochemistry, and assessment of N yield from base flow in two agricultural watersheds, Kent County, Maryland. U. S. Environmental Protection Agency, EPA/600/R-02/008.
12. Meisinger, J.J. and Randall, G.W. (1991). Estimating N budgets for soil-crop systems. In: *Managing N for groundwater quality and farm profitability*. R.F. Follet, D.R. Keeney, and R.M. Cruse (Eds.). Soil Science Society of America, Madison, WI, pp. 85–124.
13. Ator, S.W. and Ferrari, M.J. (1997). Nitrate and selected pesticides in ground water of the Mid-Atlantic Region. United States Geological Survey. Water-Resources Investigation Report 97–4139.
14. Harter, T., Davis, H., Mathews, M., and Meyer, R. (2002). Shallow groundwater quality on dairy farms with irrigated forage crops. *J. Contam. Hydrol.* **55**: 287–315.
15. Cox, S.E., and Kahle, S.C. (1999). Hydrogeology, groundwater quality, and sources of nitrate in lowland glacial aquifer of Whatcom County, Washington, and British Columbia, Canada. USGS Water Resources Investigation Report 98–4195, Tacoma, Washington.
16. Schilling, K.E. and Wolter, C.F. (2001). Contribution of base flow to nonpoint source pollution loads in an agricultural watershed. *Ground Water* **39**(1): 49–58.
17. Latinopoulos, P. (2000). Nitrate contamination of groundwater: Modeling as a tool for risk assessment, management and control. In: *Groundwater Pollution Control*. K.L. Katsifarakis (Ed.). WIT Press, Southampton, United Kingdom, pp. 7–48.
18. Goss, M.J. and Goorahoo, D. (1995). Nitrate contamination of groundwater: Measurement and prediction. *Fert. Res.* **42**: 331–338.
19. Livingston, M.L. and Cory, D.C. (1998). Agricultural nitrate contamination of ground water: An evaluation of environmental policy. *J. Am. Water Resour. Assoc.* **34**(6): 1311–1317.
20. Waskom, R.M, (1994). Best Management Practices for Nitrogen Fertilization. Colorado State University Cooperative Extension, (XCM-172).
21. U.S. Environmental Protection Agency. (2001). Source water protection practices bulletin: Managing agricultural fertilizer application to prevent contamination of drinking water. Office of Water, 4606, EPA 916-F-01-028, Washington, DC.
22. Almasri, M.N. (2003). Optimal management of nitrate contamination of ground water. Ph.D. Dissertation, Utah State University, Logan, UT.
23. Almasri, M.N. and Kaluarachchi, J.J. (2004). Implications of on-ground nitrogen loading and soil transformations on ground water quality management. *J. Am. Water Resour. Assoc.* **40**(1): 165–186.
24. Yadav, S.N. and Wall, D.B. (1998). Benefit-cost analysis of best management practices implemented to control nitrate contamination of groundwater. *Water Resour. Res.* **34**(3): 497–504.
25. Davis, J., Koenig, R., and Flynn, R. (1999). Manure best management practices: A practical guide for dairies in Colorado, Utah, and New Mexico. Utah State University Extension, AG-WM-04.
26. Pace, M.G., Miller, B.E., and Farrell-Poe, K.L. (1995). The composting process. Utah State University Extension, AG-WM 01.
27. Van Horn, H.H. (1992). Recycling manure nutrients to avoid environmental pollution. In: *Large Dairy Herd Management*. H.H. Van Horn and C.J. Wilcox (Eds.). American Dairy Science Association, Champaign, IL, pp. 640–654.
28. National Research Council. (1989). *Nutrient Requirements of Dairy Cattle*, 6th Revised Edn. National Academic Sciences, Washington, DC.

INTEGRATED WATER RESOURCES MANAGEMENT (IWRM)

MUKAND SINGH BABEL
Asian Institute of Technology
Pathumthani, Thailand

TODAY'S WATER PROBLEMS

Population and economic growth has led to rising demand for water for human consumption, food production, industrial uses, and other development activities, and opportunities for increasing the supply are becoming prohibitively expensive. The growing concern for maintaining natural ecosystems is exerting further pressure on water resources. The threatening environmental problems originating from unsustainable use of water resources are of great concern. Pollution of water is connected with human activities. Proliferation of development activities and modern living styles has also contributed to the worsening of water quality. Adverse conditions such as social inequity, economic marginalization, and lack of poverty alleviation programs in many countries also force people living in extreme poverty to overexploit soil and forestry resources, which often results in negative impacts on water resources. Deteriorating water quality caused by pollution makes safe use difficult for downstream users and threatens human health and the functioning of aquatic ecosystems.

In most countries, water policies have been largely dominated by a supply-oriented and compartmentalized sector mentality. Top-down approaches have prevailed in development and management of water resources without participation by stakeholders in decision making. In these traditional approaches, the important distinction between managing water as a commonly owned natural resource and providing services for water users has not been fully recognized, leading to a confusion of roles and responsibilities of different organizations dealing with water. Thus, inefficient governance and increased competition for finite water resources are the root cause of water problems.

Balancing and compromising the need for water by people, different economic sectors, and nature is the main challenge faced by the water sector. About 1 billion people in the world are without access to safe drinking water, and about 3 billion people do not have access to adequate sanitation. Securing water for people is a prime concern in many countries. According to the UN medium projections, the population on earth in 2025 will reach about 7.8 billion, a 38% increase over present levels. Providing water to meet the needs of food and fiber for the world population in years ahead is another main challenge. Even with improved irrigation management and water productivity, it is estimated that 17% more water will be required for this population (1). Ecosystems are sources of various benefits to society such as providing timber, fuel, wildlife habitats, aquatic resources, etc. Ecosystem security through maintenance of the natural environment and proper management of land and water resources is therefore essential.

Variation of precipitation over time and space leads to the classical problem of too much or too little water. The variations in water flow and groundwater recharge, whether of climatic origin or due to land mismanagement, can add to drought and floods, which can have catastrophic effects in terms of large-scale loss of human life and damage to economic, social, and environmental systems. Managing the variability of water in time and space and the associated risks are other challenges.

CONCEPTS OF IWRM

Compared to traditional approaches to water resources problems, integrated water resources management (IWRM) takes a broader holistic view and examines a more complete range of solutions. It looks outside the narrow water sector for policies and activities to achieve sustainable water resources development. It also considers how different actions affect and can reinforce each other. IWRM has attracted particular attention since the International Conferences on Water and Environmental in Dublin in 1992. IWRM can be defined as follows:

"IWRM is a process which promotes the coordinated development and management of water, land and related resources in order to maximize the resultant economic and social welfare in an equitable manner without compromising the sustainability of vital ecosystems" (2).

IWRM principles place its overriding importance over economic, social, and natural facets. It draws its inspiration from the Dublin principles, culminating from the International Conference on Water and the Environment in Dublin, 1992. These are the four Dublin principles:

- Freshwater is a finite and vulnerable resource, essential to sustain life, development and the environment.
- Water development and management should be based on a participatory approach, involving users, planners, and policy makers at all levels.
- Water has an economic value in all its competing uses and should be recognized as an economic good.
- Women must play a central part in the provision, management and safeguarding of water.

In the traditional approach, the planning and operation of water systems is usually fragmented, causing a lack of coordination, wastage of resources, and conflict among stakeholders. Water issues are also generally neglected when decisions are made about crop patterns, trade and energy policies, and urban design and planning, all of which are critical determinants of water demand. For the sustainable use of water resources, there is a need to create institutions and frameworks that can transcend these traditional boundaries and involve a variety of users and stakeholders. IWRM promotes a holistic view and looks at the entire hydrologic cycle and the interaction of water with other natural and socioeconomic systems. IWRM focuses on integration, participation, consultation, gender awareness, and consensus.

Integration among different stakeholders and different domains of water resources is one of the highlights of IWRM. At its most fundamental level, IWRM is as concerned with the management of water demand as with its supply. Thus, integration can be considered in two basic categories—the natural system and the human system. The natural system puts critical importance on resource availability and quality; the human system fundamentally determines the resource use, waste production, and pollution of the resource. Integration of the natural system implies a concern with upstream–downstream water-related interests, integration of land and water management, integration of freshwater management and coastal zone management, a unified management of surface water and groundwater, a shift to management at a river basin level, integration of quantity and quality, and matching water management with other sectoral policies with a collateral impact (trade, housing, energy, agriculture, etc.). Human system integration involves mainstreaming of water resources, cross-sectoral integration in national policy development, analyzing macroeconomic effects of water development on overall economic development, influencing economic sector decisions, and integrating all stakeholders in the planning and decision-making process.

As one of the principles of the IWRM, the concept of water as an economic good introduces a wholly new model to water resources management. Therefore, it differs from the traditional approach in that it invariably associates economic aspects of water with water resources management issues. It distinguishes between the value of water and its pricing. The concept of the value of water is used in assisting allocation processes; the pricing concept is used as part of the cost recovery issue. The economic value of water is highlighted through greater stress on demand management rather than supply-side management, a recognition (and estimation, where possible) of the economic value of water in different uses, acceptance of the notion of opportunity cost (what is lost to other uses from taking it for a particular purpose), and attention to cost recovery, though with concern for affordability and securing access for the poor.

Water conservation is another critical aspect of IWRM. Freshwater ecosystems face enormous threats, directly and indirectly, from human activities. Conserving freshwater systems is essential for the future survival of all living species on the earth. Water conservation and management call for the protection, improvement, and use of water according to principles that will assure their highest economic and social benefits. Maintaining the delicate and sustainable balance of demand and supply of water is a critical element of any sound water strategy.

To fulfill desirable social, economic, and natural conditions, IWRM places its overriding criteria on economic efficiency in water use, equity and access for all, and sustainability of vital ecosystems. The increasing scarcity of water as well as financial resources, under increasing pressure from the excessive demands upon it, demands that water be used with maximum possible efficiency. The "social equity and access for all" criteria recognize the basic human right to access to water of adequate quantity and quality. The third criterion (ecological sustainability) is about ensuring that the present use of the resource does not undermine ecological stability and that future generations are not adversely affected.

Three basic elements highlight the IWRM framework: the enabling environment, the institutional roles, and the management instruments, as shown in Fig. 1.

The enabling environment refers to the general framework of national policies, legislation, and regulations. The

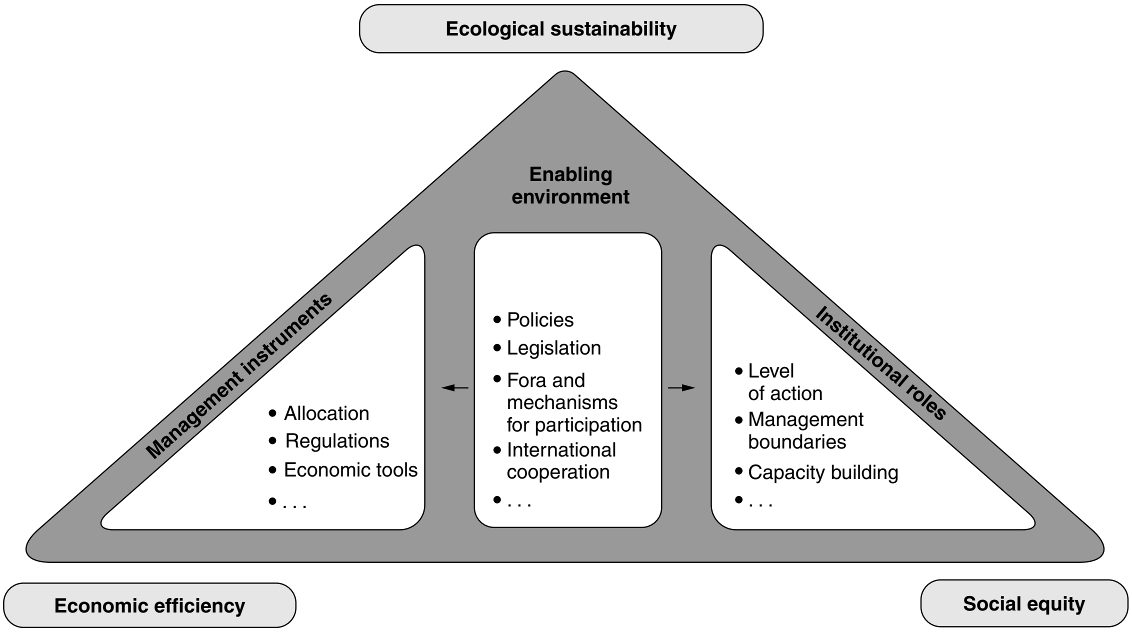

Figure 1. General framework for IWRM (2).

institutional framework defines the roles and responsibilities of government agencies at different levels and other stakeholders. The management instruments include water resources assessment, communication and information systems, allocation, economic tools, technology, and so on.

In spite of considerable advantages in IWRM, its implementation requires a much greater commitment from all stakeholders. There are many challenges, which have to be overcome for successful implementation. The short-term costs and disadvantages may seem more apparent than tangible benefits in the pursuit of longer term advantages. Integration is difficult to implement at first and may not guarantee equity for all; simply involving a wider range of stakeholders in decision making may not automatically ensure fair treatment for them.

BIBLIOGRAPHY

1. Seckler, D., Amarasinghe, U., Molden, D., de Silva, R., and Barker, R. (1998). *World Water Demand and Supply, 1990 to 2025: Scenarios and Issues, Research Report 19*. International Water Management Institute, Colombo, Sri Lanka.
2. GWP. (2000). *Integrated Water Resource Management, TAC Background Paper No. 4*. Global Water Partnership Technical Advisory Committee, Stockholm, Sweden.

MANAGEMENT OF WATER RESOURCES FOR DROUGHT CONDITIONS*

William R. Walker
Virginia Water Resources
Research Center

Margaret S. Hrezo
Radford University
Virginia

Carol J. Haley
Virginia Water Resources
Research Center

INTRODUCTION

Droughts have been a part of our environment since the beginning of recorded history, and humanity's survival may be testimony only to its capacity to endure this climatic phenomenon. According to Yevjevich and others (1, p. 41), one of the earliest records of efforts to plan for droughts is found in the biblical story of Joseph. The pharaoh of Egypt gave Joseph the authority to plan and implement the world's first recorded national drought strategy. One can ask if improvements in society's planning capability over the past several millenniums have increased our ability to deal with droughts in other than catastrophic terms. Certainly, droughts will continue to be a randomly occurring climatic phenomenon. Have we built on any past successes pertaining to the planning for droughts and managing water resources during droughts; have we learned from past actions or inactions; or are we doomed to repeat our mistakes? The evidence indicates that society tends to be unwilling to plan for droughts.

To better understand why society is still reacting to droughts rather than planning for their eventuality, the activities of various levels of government in planning for droughts must be examined. Until very recently, the policies of most State governments for dealing with drought have been to "wait 'til it rains" and in the meantime provide some form of emergency assistance to localities and hope a catastrophe can be avoided. Many States also, as a matter of implicit policy, place primary responsibility for drought action with the Federal Government and local governments who, throughout the Nation's history, have been the primary levels of government involved when droughts occur. The policies of both the Federal Government and the local governments also have been to provide emergency relief and to try to reduce water demand to match the available supply. This governmental approach, however, has not reduced the economic losses or the level of inconvenience and suffering of the Nation's citizens. With each succeeding drought, the cycle repeats itself. As a result, the following questions arise: Do the States have a responsibility for planning for droughts? If the States do have such a responsibility, what type of planning should take place? Is there a possibility that drought-related activities of the Federal Government and the local governments are likely to change in the future, and if so, how and why?

In strictly climatic terms, a drought can be defined as an interval of time, generally months or years in duration, during which the actual moisture supply at a given place consistently is less than the climatically expected or climatically appropriate moisture supply (2, p. 3). Drought generally is defined as being meteorological, hydrological, or agricultural. However, the ultimate consequences of droughts have to be placed in the context of the effects on the social and economic activities of a given region (Evan C. Vlachos, Colorado State University, written commun., January 1988). Thus, the climatic attributes of drought also need to be defined in social and economic terms because it is in these contexts that water management becomes important (Evan C. Vlachos, Colorado State University, written commun., January 1988). Yevjevich and others (1, p. 32) have suggested that we think in terms of "sociological drought," which is defined as the meteorological and hydrological conditions under which less water is available than is anticipated and relied on for the normal level of social and economic activity of the region.

---

*This article is reproduced in its entirety from Walker, W.R., Hrezo, M.S., and Haley, C.J. (1991). Management of water resources for drought conditions. In: *National Water Summary 1988-89—Hydrologic Events and Floods and Droughts*. Paulson, R.W., Chase, E.B., Roberts, R.S., and Moody, D.W., (Compilers). U.S. Geological Survey Water-Supply Paper 2375, pp. 147–156.

Due in part to our agrarian heritage and because water deficiencies can develop rather quickly in the root zone, most droughts are thought of as agricultural droughts. This characterization can be confusing, however, because an agricultural drought can occur in the midst of a hydrological wet period. The distribution of precipitation during a year can be such that there is a shortage of moisture (drought) during a critical growing period for a variety of crops, yet the total precipitation for the year can be greater than the historical yearly average. In this discussion of institutional and management aspects, droughts are considered as moisture shortages that seriously affect the established economy and the quality of life of a region.

The following analysis of the roles of government in planning for droughts begins with an examination of the traditional activities of government during a drought. If the traditional approach by government seems less than desirable, then the obstacles to greater involvement need to be identified and evaluated to determine if they can be removed or circumvented. The suggestions by King and others (3) in their "Model Water Use Act" for managing water resources during droughts are examined in some detail to determine if they provide a minimum approach for States to consider in developing a plan for managing water resources for drought conditions. To better assess the status of State involvement, the drought programs and drought-related actions of the 48 conterminous States are reviewed in terms of the minimum criteria recommended by King and others (3). Lastly, the expanded role that the Federal Government might have in planning for and mitigating the effects of droughts pursuant to existing Federal water-quality legislation is considered.

TRADITIONAL GOVERNMENT APPROACH TO DROUGHTS

Droughts have two components—climatic (decrease in precipitation) and demand (use of water). In responding to droughts, governments tend to concentrate most of their activity on reducing the demand for water, although they might have limited options for controlling the climatic component. Cloud seeding, for example, might make increasing the snowpack possible and, in some instances, obtaining a limited summer rainfall. Because climatic control, however, generally is less reliable, most measures focus on management, reallocation, and distribution of existing water sources and on establishing priorities accordingly for different uses (4, p. 41).

LOCAL GOVERNMENT

Traditionally, managing water resources during droughts has been based on immediate reactions to a current crisis. The focus of most action is to reduce the daily demand for water, and local governments usually are responsible for reducing water demand within their jurisdictions. Broader issues that impinge on their programs to reduce water demands are not within their purview; for example, local governments are not permitted to allocate surface water and groundwater among competing users. Local governments, however, do not covet the existing authority for reducing water demands. Even in times of impending crisis, there is a great reluctance to impose water-conservation measures, if there is any hope that rain will fall in time to save officials from having to do so. Decision makers are not popular when they must halt or reduce industrial activity, curtail domestic use, or prohibit "nonessential" services. As a result, timely action rarely is achieved (Evan C. Vlachos, Colorado State University, written commun., January 1988).

The local approach to the management of water resources during droughts is not responsive to other drought-induced issues such as minimum instream flows because these issues are not within the purview of local government. Yet water for wildlife, increased contamination due to low flows, and decreased navigation and hydroelectric-power generation are issues equally as important as the ones being addressed by local government. Water shortages cause low streamflows that have an adverse effect on fish and wildlife habitat. If there are no limits as to how much flows can decrease or for how long (the longer the period of low flow, the greater the stress), the recovery time for a habitat can be extremely long, or a habitat can be lost permanently. Low flows can increase salt-water intrusion, increase health hazards because of increased concentration of toxic substances and pathogens, decrease hydroelectric-power generation, and curtail recreational opportunities. Decreased precipitation also increases the potential for brush and forest fires and wind erosion of topsoil.

With a few exceptions, the response of government at any level to the shortages caused by decreased precipitation has been to react rather than to adopt a proactive approach to minimize the effect of droughts. Drought planning at the local level in many areas appears to be given a low priority because of the randomness of droughts, the limited resources for planning, the limited jurisdiction (local government might not be able to control streamflow levels), and the programs of the Federal Government to provide disaster relief in time of crisis. As a result, local governments are encouraged to accept an implicit policy of doing only what can be done after a crisis has occurred.

A notable exception to inaction at the local level to plan for drought conditions occurred in the Washington, DC, area. The leadership role in this case was taken by the Interstate Commission on the Potomac River Basin (ICPRB), but the implementation of the water resources plan was by local government. The plan finally adopted saved between $200 million and $1 billion compared to longer scale structural solutions previously proposed (5, p. 106). Implementation of the plan was through eight separate but interlocking contracts executed in 1982 (5, p. 106).

STATE GOVERNMENT

With the exception of eight States—Arkansas, California, Connecticut, Delaware, Florida, Minnesota, New Jersey, and South Carolina—the activities of the State governments in managing water resources during droughts have been minimal. Most State governments have not passed legislation providing for additional drought planning beyond slight modifications in their water laws. Governors, on occasion, will declare counties or designated areas as disaster areas in order to make individuals eligible for Federal relief.

FEDERAL GOVERNMENT

The Federal Government generally has limited its activities to providing direct relief to drought victims and to farmers in general. For example, the Federal response to the 1976-77 drought, which affected about two-thirds of the country, was to enact the Emergency Drought Act of 1977 (Public Law 95-18), the Community Emergency Relief Act of 1977 (Public Law 95-31), and certain provisions of the Supplemental Appropriations Act of 1977 (Public Law 95-26) to bolster existing emergency-assistance programs. As the result of these laws, 40 Federal programs, administered by 16 agencies, offered drought relief in the form of loans, grants, indemnity payments, and other forms of assistance to State and local governments, households, farms, and private businesses (6, p. 1–2). The overriding objective was to reduce impending damage by implementing short-term actions to augment existing water supplies. Other Federal programs available to drought victims were designed to provide assistance after damage had occurred (6, p. 2). Typical of these were the disaster loan programs of the Farmers Home Administration. Four Federal agencies—the Departments of Agriculture, Commerce, and Interior and the Small Business Administration—were responsible for implementing emergency drought programs at a cost of $5 billion, which included an additional $1 billion for short-term emergency actions to augment existing water supplies (6, p. 11). The Comptroller General of the United States (6, p. 11) concluded that:

- Some drought programs were enacted or implemented too late to have much effect in augmenting water supplies.
- Inadequate standards for determining the worthiness of projects meant that many projects were funded that had little, if any, effect in mitigating the effects of the drought.
- Drought victims were treated in an inconsistent, inequitable, and confusing manner.
- Inadequate coordination among the agencies resulted in inefficient and inequitable distribution of funds.

The report (6, p. 21) also recommended that a national plan be developed for providing assistance in a more timely, consistent, and equitable manner. Issues to be considered in the development of such a plan are:

- Identification of respective roles of agencies involved to avoid overlap and duplication of activities.
- Need for legislation to more clearly define agency roles and activities.
- Need for standby legislation to permit more timely response to drought-related problems.

In 1988, in response to another severe drought, Congress passed the Disaster Assistance Act of 1988 (Public Law 100–387). This legislation is less comprehensive than the 1977 legislation in that its primary purposes are to protect farm income in an efficient and equitable manner, protect the economic health of rural communities affected by the drought, and help assure a continued adequate supply of food for American consumers. Little in this legislation reflects the major recommendations of the Comptroller General of the United States (6) regarding the development of a national plan for providing future assistance in a more timely, consistent, and equitable manner. Although the Secretary of the U.S. Department of Agriculture is authorized to make grants and provide other assistance to combat water shortages, the Disaster Assistance Act of 1988 is primarily an agricultural-relief act. Thus, to date, the Federal Government has limited its involvement in droughts to the provision of water-resources information, technical assistance, and financial relief to mitigate the costs incurred once a drought has occurred.

OBSTACLES TO EFFECTIVE PLANNING FOR DROUGHTS

Although governments can plan effectively for droughts, fundamental problems that deter action need to be examined and understood before drought planning can become a reality. Five obstacles to planning for droughts—specificity, randomness, drought phenomenon, cost of droughts, and political considerations—are discussed here.

SPECIFICITY

Planning for and management of hazardous events presuppose that those events are well defined and discernible to all. The planning necessary to reduce the effects of most natural hazards is difficult because the intensity and frequency of the events are unknown, although there is never any question as to their eventual occurrence. Although no technical expertise is required to determine when floods, volcanic eruptions, or earthquakes have been experienced, considerable uncertainty exists as to when droughts start and end. A drought is almost a “non-event” (Evan C. Vlachos, Colorado State University, written commun., January 1988). Any discussion about planning for drought conditions and management of water resources, therefore, requires some definition as to what constitutes a drought. This lack of specificity can be a major contributing factor, although unstated, as to why the planning for droughts and the managing of water resources during droughts have received less attention than they deserve.

RANDOMNESS

Droughts, when placed in historical perspective, have not received much attention from governments at any level until severe water shortages occur. What most inhibits the planning for water shortages associated with droughts is their random nature. It is this inherent variability that makes "reacting" to a water-shortage crisis and instituting relief efforts when the shortages occur appear to be more rational than does "planning" for droughts. People tend to overlook that droughts are a normal part of the climatic regime and that they will recur. Droughts remain a certainty; only their frequency and severity are unknown. Thus, it makes good sense to plan to reduce both the costs that result from droughts and the associated personal hardships.

DROUGHT PHENOMENON

Another obstacle to planning for droughts is inherent in the drought phenomenon. Although droughts do affect individuals, in reality droughts are a community problem having characteristics of Hardin's famous "Tragedy of the Commons" (1, p. 34). The self-interest of each individual using communal property is to maximize it for immediate gain. The net result can well be the destruction or deterioration of the communal property. In Hardin's example, the overgrazing and destruction of the common pastures occurred because each person sought to graze all the animals possible. The sum of the actions by each individual was not "best" for the sum of the individuals. This phenomenon makes the best policy infeasible unless the individuals reach a consensus themselves or are compelled to do so by a government.

Droughts produce the same type of situation. For individuals experiencing a drought, options to deal effectively with water shortages are limited. Collective action by all the individuals provides the best solution. The use of groundwater during times of drought in many Eastern States provides a comparable situation to the one described by Hardin in the case of common pastures. The common law in many Eastern States deals with groundwater as a common resource that is appropriated under the Rule of Capture (7, p. 115). What you capture is what you get, and those having the deepest wells and the largest pumps get the most water. Given these circumstances, the solution for an individual during drought conditions might be to drill a well. If all individuals act in the same manner, a variety of consequences can occur to the detriment of each. A shallow aquifer eventually can be depleted, the individuals can be competitors for water and cause larger and larger cones of depression as deeper wells and larger pumps are utilized, or the increased pumping can cause saltwater intrusion, which will destroy the quality of the water in the aquifer for all. The best solution for all parties might be an agreement between the individual well owners, or restriction by the local government, to curtail the time and rate of pumping. Under these circumstances, an individual who pumps water from a common aquifer cannot plan effectively for droughts. If an individual well owner seeks to conserve groundwater or plan for a water shortage, he or she needs to be aware that the water he or she does not pump will probably be pumped by others.

Because droughts affect larger geographic areas than those occupied by single communities, the ability of an individual community to respond effectively is affected by the actions of similar communities in the drought area. The position of each community in this larger arena can be analogous to that of the individual in the community. Individual actions by each community can be counterproductive to the policy best for the region as a whole. For one community, the solution may be the building of a reservoir on a stream that is the water source for other communities downstream. As each community opts to resolve its water needs without regard to its neighbors, the stream can become an inadequate water source for all. The development of a regional water supply for all communities might be the best solution, but this will require the consensus of all the communities.

COST OF DROUGHTS

The lack of information about the cost of droughts is another reason why only marginal interest exists in planning for droughts, especially at the State level. The magnitude of drought costs is assumed to be less than that of other natural hazards because the losses associated with other natural hazards are more evident and generally are incurred during short periods of time. In contrast, drought losses generally are distributed over longer time periods. When the true costs of drought are known, drought losses can dwarf the losses from other natural hazards. For example, Australia determined that, for the period 1945–1975, the costs of droughts were four times the costs of other natural hazards (8, p. 226). In addition, all the costs associated with droughts are not clearly defined. The social effects of droughts and the associated costs, how the effects propagate throughout society, and who is ultimately affected need to be better understood. Human suffering is less likely to be factored into the cost assessment even though these costs are real and can continue for years, long after other costs have been absorbed. The aggregated indirect costs probably are far greater than the direct costs, but because of their diffused nature they are difficult to identify and quantify and, thus, generally go unrecognized (1, p. 32). These indirect costs, which are disbursed among large groups and throughout large geographic areas, nevertheless constitute a major proportion of the total costs resulting from droughts (1, p. 32). Again, the random nature of droughts, coupled with the rapid decrease of public interest in droughts after normal precipitation resumes and the limited resources available for planning, make the determination of in-direct costs associated with the droughts less urgent. As long as these indirect, diffused costs remain undisclosed, decision makers will have incomplete knowledge of the costs of drought. If the past is any guide, the total costs of droughts probably never will be reliably assessed.

The length of a drought also has a significant effect on the total costs; long droughts are more costly than shorter droughts. A sustained drought, such as the one in the 1930s, can have economic and social costs that are never quantified. During the 1930s drought, for example, agriculture was abandoned in some sections of the Nation; this abandonment, in turn, caused dislocation of people and severe impairment of the economic substructure that supported agriculture. The 1930s drought had an effect on a whole generation of Americans, wherever they lived and however they made their living (4, p. 34). Even if the value of human life is ignored, the total economic losses from droughts can be staggering.

POLITICAL CONSIDERATIONS

Lastly, political considerations affect action that might lessen the effects of droughts through better planning and management. The randomness of droughts induces the public to believe that little can be done to reduce the costs of droughts before they occur. In addition, the public's memory of past tragedies usually is short, and political attention shifts quickly to new political problems. The public, lacking an analysis of the total costs associated with a drought, has the illusion that droughts are affordable, although inconvenient. Thus, decision makers lack the public support needed to take aggressive action in planning for droughts and managing water resources during droughts. This is not strictly an American phenomenon, as witnessed by the inaction of the British Parliament during the mid-1970s when the country experienced its worst drought in 500 years (9, p. 51). Early in the drought, efforts were made to have Parliament enact legislation to extend the responsibilities of the river authorities to mitigate the effects of the drought. Before responding, Parliament waited until there was widespread public awareness of the need for the legislation; thus, the damage factor was increased substantially compared to what it would have been had Parliament acted earlier.

In contrast to a lack of public support, special-interest groups at the State level might oppose activities that are essential for an effective water-management plan applicable to droughts. For example, farm groups in Virginia strongly oppose any Federal, State, or regional water-management plan because they believe there should be no regulation of water apart from the Riparian Doctrine (Mark Tubbs, Virginia Farm Bureau, commun., 1981). Water management, of necessity, must be at the core of any program to mediate the effects of water shortages that occur during droughts, but political factors can substantially dampen the interest in managing water resources even during droughts.

FRAMEWORK FOR STATE ACTION

The "Tragedy of the Commons" phenomenon, which characterizes the problems associated with any management plan to mitigate the costs associated with droughts, illustrates that, in the absence of agreement among all the parties affected by the drought, the management responsibility needs to be at the lowest possible level of government that will permit the attainment of management's objectives (1, p. 34). The State, in most situations, represents the unit of government that has the authority to allocate water, to set policy objectives that are concerned with water-use efficiency and equity, to consider interboundary issues and externalities associated with matters such as minimum instream flows, and to coordinate the activities of local governments in meeting water-supply needs during times of severe water shortages.

The responsibility for managing water resources during droughts, once assumed by the State, needs to be vested in such a manner as to require timely action and not be vulnerable to legal challenges by groups who do not favor an approach taken by the State. Although expanding the Governor's powers to deal with disasters by including droughts might be expeditious, the action taken by most States generally is to group management activities with the authority primarily designed to respond to disasters after they have occurred rather than to undertake planning activities to reduce the cost of droughts in advance of their occurrence. Colorado, Delaware, New Jersey, and North Carolina are examples of States that have used executive power to develop statutory guidelines that define droughts and delineate interaction among State agencies responsible for water resources (10, p. 162–163).

PLANNING TOOLS

The authors of the "Model Water Use Act" (3, 1958) developed five planning tools—identification of drought indicators, designation of government authority, notification of the public, curtailment of water use and maintenance of revenues, and monitoring of water-user compliance—to cope with planning for water shortages associated with droughts. These tools addressed the following fundamental questions:

- How does a State know when there is a drought?
- If there is a drought, who is in charge?
- How is the public informed?
- How are current allocations and uses of water to be modified?
- How is compliance assured?

The answers or responses to these fundamental questions need not be identical for each State having a functioning water-management plan applicable to droughts. Each State, however, needs to address each question in terms of its own circumstances. Failure to address each of the questions will detract seriously from the effectiveness of a water-management plan. Each of these planning tools is now examined in detail.

IDENTIFICATION OF DROUGHT INDICATORS

Because of the difficulty of deciding when droughts start and end, specific drought indicators must be used to decide when to implement a water-management plan.

When such indicators have been identified, water users can formulate contingency plans and make decisions on future economic investments (11, p. 47). The drought indicators must be precise and susceptible to little, if any, subjective decision making. The latter makes the indicators vulnerable to court action by those who oppose advanced planning.

A variety of drought indicators can be used, including the Palmer Index (a drought-severity index), in stream flows, historical data on the present and anticipated needs for water, the degree of subsidence or saltwater intrusion, the potential for irreversible adverse effects on fish and wildlife, and reservoir or groundwater conditions relative to the number of days of water supply remaining (11, p. 47). Usually it is desirable to select a number of drought indicators to reflect the seasonal relation of supply versus demand. The Delaware River Basin Commission (11, p. 51) relies on five drought indicators—precipitation, groundwater levels, reservoir storage, streamflow, and the Palmer Index. Ranges of values for each of these indicators are assigned to one of four drought stages—normal, drought watch, drought warning, and drought emergency (11, p. 51). To activate any one of the drought stages, three of the five drought indicators must indicate a given drought stage (12, p. 34). The drought indicators should not be so complex as to cause uncertainty about whether some stage of the water-management plan should be activated; for example, if precipitation and reservoir storage are two drought indicators, the decision to activate the plan will be unclear if the precipitation is less than normal while reservoir storage is normal.

When there is only one source of water supply, one drought indicator may be sufficient; for example, when a city's only source of water is a reservoir, the water-management plan can be activated if reservoir storage, expressed as a percentage of normal seasonal capacity, decreases below a specified percentage. The phasing criteria used by Manchester, Connecticut, are an example (Table 1): a drought watch goes into effect when reservoir storage is at 70 percent of normal seasonal capacity, and stage I of the water-management plan becomes operational when reservoir storage is 57 percent of normal seasonal capacity.

A sliding scale for drought indicators also can be shown graphically in terms of the storage in a reservoir, as shown by the operation curves for three reservoirs in New York that also are part of the Delaware River basin (Fig. 1). When the actual reservoir level drops below the drought-warning zone or drought zone, schedule of reduced diversions from the basin to the various localities takes effect. The Pennsylvania Drought Contingency Plan for the Delaware River Basin is based on these criteria. Where groundwater is one of the main sources of water supply, drought indicators based on groundwater levels can be used; the Alameda County Water District in California has such a plan (12, p. 38).

Most water-management plans have correlated successive stages of a drought strategy to certain deficit-reduction goals (12, p. 39). Fewer than three stages in a plan can result in marked differences in the actions to be implemented between the first and the second stages. More than five stages in a plan, however, can cause frequent transitions between stages, which can decrease the effectiveness of the plan. An example of a workable plan is the Seattle Water Department five-stage plan for reducing water use (Table 1). Agencies having water-management

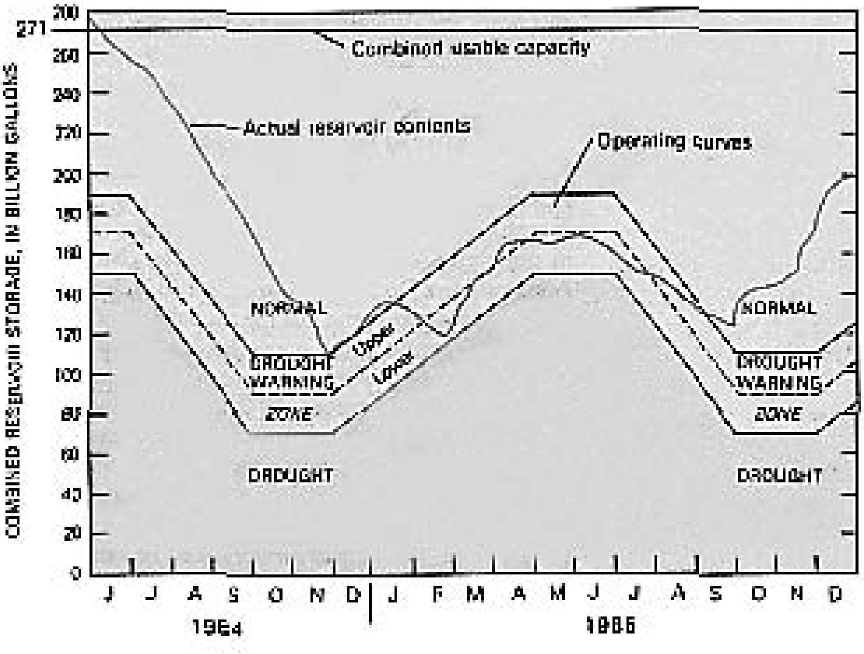

Figure 1. Operation curves for three New York City reservoirs (Cannonville, Pepacton, and Neversink) in the Delaware River Basin (from Ref. 13, p. 31).

Table 1. Drought-Contingency-Plan Phasing Criteria Used By Two Cities in the United States

MANCHESTER, CONNECTICUT
(South System)

Drought Stage	Drought Stage Initiating Conditions	Demand-Reduction Objective
Drought watch	Reservoir water levels at 70 percent of normal seasonal capacity.	Informational only, raise public awareness
Stage 1	Reservoir water levels at 57 percent of normal seasonal capacity.	Cut back withdrawals from reservoirs by 5 percent or reduce total system use by 3.8 percent.
Stage 2	Reservoir water levels at 40 percent of normal seasonal capacity.	Cut back withdrawals from reservoirs by 30 percent or reduce total system use by 20 percent.
Stage 3	Reservoir water levels at 0 percent of normal seasonal capacity	Eliminate withdrawals from reservoirs and reduce total system use by 70 percent.

Source: Data from Reference 12, pp. 33, 35, and 36.

SEATTLE, WASHINGTON, WATER DEPARTMENT
(Deficit reduction objective is based on demand levels in a l-in-50-year
drought, which are expected to be above normal averages due to warm, dry conditions)

Summer-Shortage Response Plan

		Demand-Reduction Amount, in Million Gallons Per Day	
Stage 1 Minor shortage potential	Total system storage is not filled to capacity as of June 1. Streamflow and snowmelt forecasts indicate that inflows will be inadequate to fill storage facilities before the beginning of the peak-use season.	Water-system management Customer Total	3.0 0.0 3.0
Stage 2 Moderate shortage potential	Total system storage is predicted to fall below the level required to meet expected demands during a 1-in-50-year drought. System inflows continue to be low. Weather forecasts predict a continuing trend of warmer, drier than normal conditions.	Water-system management Customer Total	3.0 4.7 7.7
Stage 3 Serious shortage	Total system storage drops below the level required to meet expected demands during a 1-in-50-year drought. System inflows continue to be low. Weather forecasts predict a continuing trend of warmer, drier than normal conditions.	Water-system management Customer Total	5.0 15.0 20.0
Stage 4 Severe shortage	Conditions described for stage 3 occur near the end of the peak-use season.	Water-system management Customer Total	5.0 16.1 21.1
Stage 5 Critical emergency	Customer demands and system pressure requirements cannot be met.	Not applicable.	

Fall-Shortage Response Plan

		Demand-Reduction Amount, in Million Gallons Per Day	
Stage 1 Minor shortage potential	Total system-storage levels are dropping due to the increased use associated with a warm, dry summer. Weather forecasts predict a continuing trend of warmer, drier than normal conditions.	Water-system management Customer Total	3.0 0.0 3.0
Stage 2 Moderate shortage potential	Total system storage is expected to fall below the level required to meet expected demands during a 1-in-50-year drought.	Water-system management Customer Total	5.0 2.4 7.4
Stage 3 Serious shortage	System inflows continue to be low.	Water-system management Customer Total	5.0 6.2 .11.2
Stage 4 Severe shortage	Weather forecasts predict a continuing trend of warmer, drier than normal conditions.	Water-system management Customer Total	5.0 19.6 24.6
Stage 5 Critical emergency	Customer demands and system pressure requirements cannot be met.	Water-system management Customer Total	5.0 52.4 57.4

plans have determined that fall droughts have a lesser probability of occurrence, but if they do occur, they are likely to develop more quickly and be more severe.

DESIGNATION OF GOVERNMENT AUTHORITY

The designation of a governmental unit or agency having specific drought-planning authority in advance of a drought is one of the critical aspects of providing for managing water supplies during droughts (10, p. 162). For a water-management plan to be effective, the designated agency needs to have authority to declare that a drought exists and to alter water-use patterns. The location of a unit or agency that has this authority within the State's administrative structure will vary. Statutory authority needs to be detailed and specific. If discretionary authority is given, well-defined guidelines for its use need to be given. Failure to adequately delineate the limits wherein action needs to be taken can create situations where administrators postpone action in order to avoid conflict with user groups. This postponement lessens the protection afforded to both water quantity and quality (10, p. 164).

NOTIFICATION OF THE PUBLIC

When drought conditions activate the implementation of the water-management plan, the public needs to be notified. The notice needs to contain information about the provisions to curtail use, when conservation measures become effective, the availability of variances, and the procedures for obtaining a variance. The exact notification procedure can be developed to reflect local conditions. Examples of States having well-defined notification procedures include Florida, Georgia, North Carolina, and South Carolina (10, p. 165).

CURTAILMENT OF WATER USE AND MAINTENANCE OF REVENUES

A system of priorities of water-use categories needs to be in place before droughts occur, so that each user knows, in advance of a drought, in what order water restrictions will be applied. If industries and commercial establishments know, in advance of a drought, the procedures to be used in reducing water availability, they can establish their own contingency plans for those reductions; for example, they can arrange for alternative water supplies or plan reductions in production schedules. Unlike most Western States, the majority of Eastern States, which allocate water according to the riparian doctrine, do not have a water-use-priority system to deal with water shortages caused by droughts. Some attempt, however, has been made in most States to show a preference for certain uses (10, p. 162).

A reduction in water use will cause a reduction in the revenues of the water suppliers. The reduced revenues come at a time when costs are greater because of expenditures made to deal with the drought. In the absence of a water-revenue reserve or a drought-emergency account, water suppliers need to either increase water rates or impose a drought surcharge (12, p. 49). The use of a drought surcharge has several advantages compared to a simple increase in water rates. For example, the drought surcharge is easier to administer, and the amount of revenue to be generated is more predictable. This surcharge probably is more acceptable to the customer because it is a one-time charge that is understandable and allays the fear that a water-rate increase to make up revenues lost during droughts will continue when the drought is over.

MONITORING OF WATER-USER COMPLIANCE

Experience has indicated that reductions in water use greater than 20–25 percent cannot be obtained with a request for voluntary conservation (12, p. 29). There does not seem to be general agreement as to whether mandatory conservation regulations or water-rate increases and drought surcharges are the most effective means of reducing water use to a volume less than that obtained by voluntary conservation. Utah determined that price increases to reduce water use were not as effective as mandatory conservation regulations for a short drought (12, p. 28). New Jersey, in contrast, recently enacted a "water emergency price schedule" in preference to mandatory restrictions (New Jersey Administrative Code, title 7, section 19B-1.5(a)). Some utilities managers argue that the availability of enforcement mechanisms is the important feature of the plan and that their application is rare (12, p. 29). Much of the monitoring of customers for compliance comes from peer-group pressure, but governmental employees, such as supervisors of streets and wastewater departments and inspectors for buildings, plumbing, electric, construction, and health services, can be empowered to issue citations (12, p. 29). This is an effective method of monitoring a service area with a minimum of expense and with minimal disruption of employees' regular work schedules.

SYNOPSIS OF PLANNING AND MANAGEMENT ACTIVITIES OF THE STATES

The planning tools suggested by King and others (3) in their "Model Water Use Act" provide a minimum set of criteria for evaluating the activities of States in the area of water management during droughts. A survey in 1983 and its update in 1986 by the Virginia Water Resources Research Center (10) determined that, of the 48 conterminous States, only 8 States have comprehensive water-shortage plans, 27 States have emergency drought provisions within their water-rights system, and 13 States do not appear to have plans in place for managing water resources during droughts (Fig. 2).

The eight States that have comprehensive water-management plans incorporate all of the basic concepts suggested in the "Model Water Use Act." However, the approach by each State has been different and reflects individual State needs and existing water-allocation systems. The drought indicators that are used

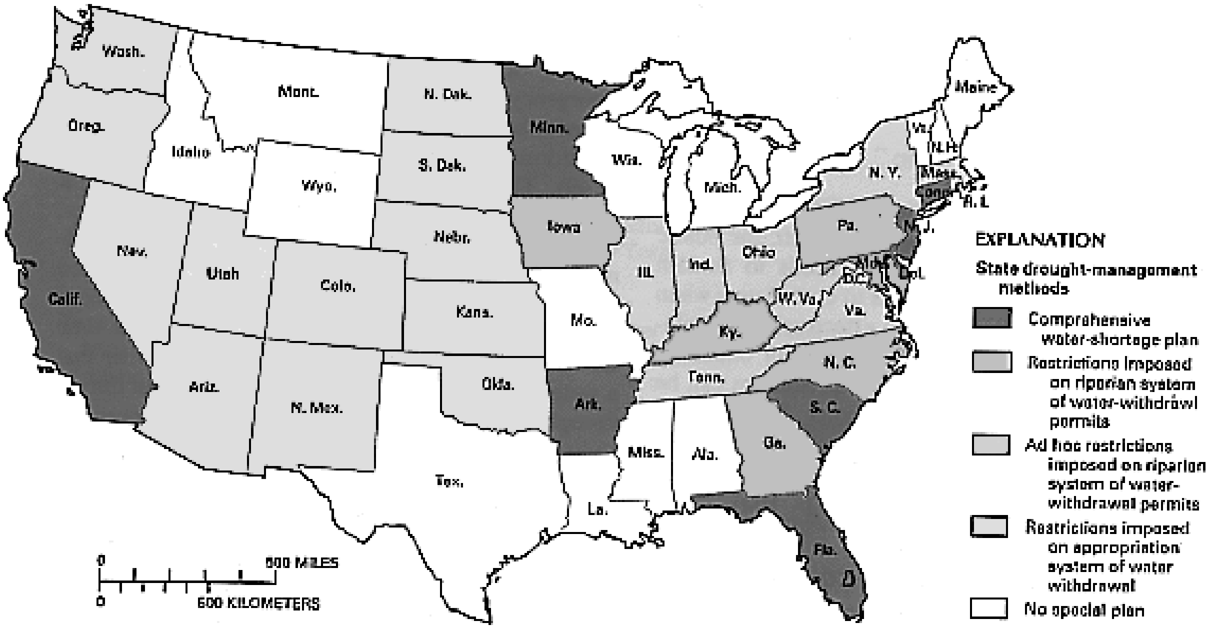

Figure 2. State-level methods for managing water resources during drought conditions (From Ref. 10, p. 146).

to determine when the plans are to begin and to end probably were the least precise component in the plan of each of the eight States. All the water-management plans are activated at the State level except in Florida, where activation is the responsibility of existing water-management districts, and in California, where activation is in the charge of the intrastate districts mandated by statute.

Water laws of 27 States have been modified to establish systems of water-use categories (Fig. 2). Although having these systems in place eases the management of water supplies once a drought occurs, most of the changes are designed to meet emergency water shortages rather than to provide a strategy to mitigate the effects of droughts before an emergency develops (10, p. 149). These 27 States can be subdivided into those that have modified the riparian doctrine with some form of a permit system; that use a modified appropriation system; or that rely on civil-defense, disaster, or emergency legislation that incorporates droughts within its definitions (10, p. 148). A description of these subdivisions follows.

Six States (Fig. 2) in the East have modified then common-law riparian doctrine by requiring water-use permits for some uses. Many of the permits are subject to water-use restrictions during times of severe water shortages. Georgia has water-use restrictions that are implemented during water emergencies. Iowa has made all use of surface and groundwater subject to a statewide permit system; permit users are denied water whenever the minimum instream-flow standard is reached. In Kentucky, permits give no guarantee of a water right during droughts. Conservation programs are required under all permits in Maryland when the safety is threatened by existing or projected water demand. North Carolina's permit system is limited to designated capacity-use areas; other laws in North Carolina grant water-emergency powers that affect the capacity-use areas and the rest of the State. Pennsylvania requires permits only of public water suppliers using surface water; permits are conditioned on there being an emergency water plan in place. All these States have taken some initial steps to manage their water resources during droughts, but they utilize few of the planning tools suggested by King and others (3) in their "Model Water Use Act" for water management (10, p. 155).

Twelve States (Fig. 2) in the West have attempted to deal with severe water shortages by modifying the appropriation doctrine for allocating water. Under this doctrine only the most senior appropriation in time will receive water during water shortages. All these States have modified the appropriation doctrine to some degree to accommodate water shortages caused by droughts. New Mexico, for example, provides for changing the place of diversion, storage, or use of water if an emergency exists (10, p. 152). In Oregon, the Director of Water Resources can order State agencies and political subdivisions to develop water-conservation and water-use-curtailment plans that encourage conservation, reduce nonessential water use, prevent waste, provide for reuse of water, and allocate or rotate the supply to domestic, municipal, and industrial uses (10, p. 152). Utah allows its State Engineer to use regulatory authority to prevent waste in order to mandate rotation of irrigation water where no use will benefit from a diversion of the water supply (10, p. 152). In these 12 States, legislation provides some management tools to assist in decreasing the effects of a water-shortage crisis. Little of the modifying legislation includes the planning tools suggested by King and others (3) in their “Model Water Use Act” to manage water shortages due to droughts in a comprehensive way.

Nine States (Fig. 2) in the Midwest and Mid-Atlantic region have left the management of droughts, from a State perspective, almost entirely to the Governor. These States perceive water shortages as primarily the responsibility of local government. The State has a role when the shortage becomes extensive in terms of time and scope and affects such a large area of the State that the water shortage is perceived as a severe crisis approaching a disaster. These States choose to deal with water shortages due to droughts on an ad hoc basis. New York passed a law that took effect in 1990 and requires a conservation plan for surface withdrawals.

INTRASTATE REGIONAL AUTHORITIES AND INTERSTATE COMMISSION COMPACTS

Intrastate regional authorities can perform the water-management function during droughts if droughts are less than statewide in scope. Politically, such authorities are difficult to create because of the rivalry that exists among local units of government. If intrastate regional authorities are established, such regional grouping must be done carefully to avoid possible constitutional challenge as unlawful delegation of legislative authority (10, p. 162). The water-management districts in Florida are examples of this approach.

When droughts affect more than one State but are not national in scope, interstate commissions and compacts can provide the management function to mitigate drought effects. Their effectiveness is predicated on having well-publicized plans and specific rules for planning purposes so that all users know how they will fare when the river flows cannot accommodate all the withdrawal demands (10, p. 164). The Delaware River Basin Compact (Public Law 87–328, 75 Stat. 688, 1961) and the 1978 Potomac River Low Flow Agreement are examples of interstate compacts that have been used to plan for the problems associated with droughts. Such commissions or compacts require, however, the approval of all of the involved State legislatures, which could be a difficult task politically. The fact that the number of such entities is small is ample evidence of the difficulty and time required to establish them.

FEDERAL RESPONSIBILITIES FOR WATER QUALITY AS RELATED TO DROUGHTS

Since 1972, with the passage of the Water Pollution Control Act Amendments (Public Law 92–500), the Federal Government has assumed a more dominant role on water-quality issues related to surface water. Because droughts and the resulting low flows have a substantial effect on water quality, it is possible, on the basis of existing legislation and regulations, that the Federal Government might become more interested in drought management. For example, Section 208(b)(2)(1) of the Clean Water Act (Public Law 92–500) provides:

> Any plan prepared under such process shall include, but not be limited to, (1) a process to (i) identify, if appropriate, salt water intrusion into rivers, lakes, and estuaries resulting from reduction of fresh water flow from any cause, including irrigation, obstruction, groundwater extraction, and diversion, and (ii) set forth procedures and methods to control such intrusion to the extent feasible where such procedures and methods are otherwise a part of the waste treatment management plan.

The legislative history amplifies on this point (1972 U.S. Code Congressional and Administrative News, p. 3706) (emphasis added):

> Salt water intrusion no less than point sources of discharge, alters significantly the character of the water and the life system it supports. Salt water intrusion often devastates the commercial shellfish industry. It must be accounted for and controlled in any pollution control program. It makes no sense to control salts associated with industrial or municipal waste point sources and allow, at the same time, similar effects to enter the fresh water as a result of intrusion of salt water. *Fresh water flows can be reduced from any number of causes. The bill requires identification of those causes and establishment of methods to control them so as to minimize the impact of salt water intrusion.*

Droughts can be one of the major causes for reduced freshwater flows, and the law requires that methods be established to control or to minimize the causes of reduced freshwater flows that allow saltwater intrusion.

The Siting Requirement under the Safe Drinking Water Act (Public Law 93–523) contains language(emphasizes added) that can be construed to include drought conditions and, thus, impose on States the requirement to adapt siting criteria that include drought conditions:

> Before a person may enter into a financial commitment for or initiate construction of a new public water system or increase the capacity of an existing public water system, he shall notify the State and, to the extent practicable, avoid locating part or all of the new or expanded facility at a site which: (a) Is subject to a significant risk from earth-quakes, floods, fires *or other disasters which could cause a breakdown of the public water system or a portion thereof.*

Droughts could conceivably be “other disasters” provided for in the Safe Drinking Water Act. A combination of these water-quality considerations might, in the future, cause the Federal Government to give greater consideration to managing the effects of droughts, at least with respect to those effects that impinge on water-quality issues.

Lastly, the quality of surface-water bodies is affected markedly by the runoff that occurs when precipitation increases after a drought. During the drought, pollutants accumulate on the land surface and on other surfaces, such as pavement and structures. It is not uncommon for droughts to be followed by a period of abnormally high precipitation that tends to aggravate the already existing water-quality problems by rapidly flushing large loads of pollutants into surface-water bodies. After the drought in England in the 1970s, the nitrate concentration in the Thames River increased to the point where the public-supply intakes had to be closed (9, p. 54). This kind of post drought problem may be reflected in what the Federal Government requires States to do to meet water-quality standards. Some drought planning may occur at the State level as a byproduct of the action taken to address this water-quality problem.

CONCLUSIONS

The planning for and the management of the effects of droughts appear to have a low priority in all but a few States, although all have experienced severe water shortages. For the most part, accommodating the inconvenience caused by droughts is considered a local-government responsibility. The Federal Government's role has been to provide financial assistance to citizens after the droughts have occurred. Water-quality legislation may cause the Federal Government to take a more proactive approach to managing the effects of droughts. Several factors will have to coexist before many States will undertake development of plans to mitigate the effects of droughts. Such factors may be the occurrence of a drought that is long and extensive, thereby increasing demands on a fixed water supply, and a public awareness of the economic costs of droughts. As water demands continue to grow, even minor droughts will become more serious, and States will be compelled to become leaders in developing water-management plans.

BIBLIOGRAPHY

1. Yevjevich, V., Hall, W.A., and Salas, J.D. (Eds.). (1978). *Drought Research Needs-Conference on Drought Research Needs, Colorado State University, Fort Collins, Colorado*, December 12–15, 1977, Proceedings: Fort Collins, CO, Water Resources Publications, p. 288.
2. Palmer, W.C. (1965). *Meteorological Drought*. U.S. Weather Bureau Research Paper 45, p. 64.
3. King, D.B., Lauer, T.E., and Zieglar, W.L. (1958). Model water use act with comments. In: *Water Resources and the Law*. University of Michigan, Ann Arbor, pp. 533–614.
4. Harrison, R. (1977). Response to Droughts. *Water Spectrum* **9**(3): 34–41.
5. Sheer, D.P. (1986). *Managing Water Supplies to Increase Water Availability, in National Water Summary 1985—Hydrologic Events and Surface-Water Resources*. U.S. Geological Survey Water-Supply Paper 2300, pp. 101–112.
6. Comptroller General of the United States. (1979). *Federal Response to the 1976-77 Drought-What Should Be Done Next*. Government Printing Office, Washington, DC, p. 25.
7. Cox, W.E. (1982). Water law primer. *American Society of Civil Engineers, Water Resources Planning and Management Division, Proceedings*, v. 18 (WRI), pp. 107–122.
8. Heathcode, R.L. (1986). Drought mitigation in Australia. *Great Plains Quarterly* **6**: 225–237.
9. Blackburn, A.M. (1978). Management strategies—dealing with drought. *Am. Water Works Assoc. J.* **1978**: 51–59.
10. Hrezo, M.S., Bridgeman, P.G., and Walker, W.R. (1986). Integrating drought planning into water resources management. *Nat. Resour. J.* **26**: 141–167.
11. Hrezo, M.S., Bridgeman, P.G., and Walker, W.R. (1986). Managing droughts through triggering mechanisms. *Am. Water Works Assoc. J.* **1986**: 46–51.
12. California Department of Water Resources. (1988). *Urban Drought Guidebook*. California Department of Natural Resources, Department of Water Resources, Office of Water Conservation, Water Conservation Guidebook 7, p. 144.
13. U.S. Geological Survey. (1986). *National Water Summary 1985-Hydrologic Events and Surface-Water Resources*. U.S. Geological Survey Water-Supply Paper 2300, p. 506.

WATER RESOURCES MANAGEMENT

Neil S. Grigg
Colorado State University
Fort Collins, Colorado

Water is an essential asset for all life and economic activity, and the term "water resources management" describes activities necessary to balance supplies and demands of water. It involves applying structural and nonstructural measures to control natural and man-made water resources systems to achieve beneficial human and environmental purposes, which may be listed as water supply, wastewater service, flood control, hydropower, recreation, navigation, and environmental protection (1). Along with these, water resources management deals with mutual impacts between water and other systems, such as land, wildlife, and economic systems.

Water resources management controls *water resources systems*, that are combinations of constructed water control facilities and natural, or environmental elements that work together to achieve water management. A constructed water resources system, consisting of structural facilities, provides control of water flow and quality and includes facilities for water supply and wastewater management; for drainage of land and control of floods; and for water control in rivers, reservoirs, and aquifers. Examples include conveyance systems (channels, canals, and pipes), diversion structures, dams and storage facilities, treatment plants, pumping stations and hydroelectric plants, wells, and all appurtenances.

Natural water resource systems comprise sets of environmental or hydrologic elements in nature that include the atmosphere, watersheds, stream channels, wetlands, floodplains, aquifers and groundwater systems, lakes, estuaries, seas, and the ocean.

In addition to structural measures to control water flow and quality, nonstructural measures are used for programs or activities that do not require constructed facilities. Examples of nonstructural measures include pricing schemes, zoning, incentives, public relations, regulatory programs, and insurance.

Water resources management takes place within a "water industry" that consists of water service organizations, regulators, coordinators, and support organizations. This industry places great value on public involvement in decisions because water affects all people and activity in interdependent relationships.

Water supply serves four categories of water users: people, industries, farms, and the general environment. Thus, we speak of water supply for people (domestic water supply), for cities (urban water supply), for farms (irrigation), for industries (industrial water supply), and for cities and industries (municipal and industrial or M&I water), or we can speak of water for the environment (water for natural systems and habitat). Wastewater management serves the same categories, as urban wastewater, industrial wastewater, and drainage for farms.

In-stream uses of water include hydropower generation, navigation, recreation, and sustenance of fisheries and ecological systems.

Stormwater and flood control are different types of water management activities that handle excessive water. As "protective" services, they do not provide water, but remove or store the excess water.

Given its many facets, water resources management should be comprehensive, coordinated, and integrated. Comprehensive water resources management includes all purposes and stakeholders in its activities. When water management is coordinated, there should be linkages between activities so that they occur with due consideration of each other. Integrated water resources management is a term that includes aspects of comprehensive and coordinated water management, and it has been explained in several ways; usually the goal is describing a holistic management process.

For example, a working group of the Awwa Research Foundation (2) used the term "Total Water Management" to explain the integrated nature of water resources management:

> "Total Water Management is the exercise of stewardship of water resources for the greatest good of society and the environment. A basic principle of Total Water Management is that the supply is renewable, but limited, and should be managed on a sustainable use basis. Taking into consideration local and regional variations, Total Water Management
>
> - encourages planning and management of natural water systems through a dynamic process that adapts to changing conditions;
> - balances competing uses of water through efficient allocation that addresses social values, cost-effectiveness, and environmental benefits and costs;
> - requires the participation of all units of government and stakeholders in decision-making through a process of coordination and conflict resolution;
> - promotes water conservation, reuse, source protection, and supply development to enhance water quality and quantity; and
> - fosters public health, safety, and community good will."

In the past, water resources management was primarily an engineering arena for building dams, laying pipelines, installing pumps, and operating systems. Now, managing water resources requires skills and approaches that go beyond pure engineering, science, management, and law. Water resources managers must deal with complexity and conflict to unravel interdependency of systems and resolve political and legal dilemmas. Skills from several disciplines are required, including law, finance, and public administration, along with engineering and science.

BIBLIOGRAPHY

1. Grigg, N.S. (1996). *Water Resources Management: Principles, Cases, and Regulations*. McGrawHill, New York.
2. Awwa Research Foundation. (1996). *Minutes of Workshop on Total Water Management*. August 1996, Seattle, Denver.

NASA HELPING TO UNDERSTAND WATER FLOW IN THE WEST

Lindsay Renick Mayer
Krishna Ramanujan
Goddard Space Flight Center
Greenbelt, Maryland

To do their jobs, water resource managers in the Columbia River Basin have mostly relied on data from sparsely located ground stations among the Cascade Mountains in the Pacific Northwest. But now, NASA and partnering

Figure 1. Map of Rio Grande and Columbia River Basins. *Credit*: Image by Robert Simmon, NASA GSFC Earth Observatory, Michael Tischler, NASA/GSFC.

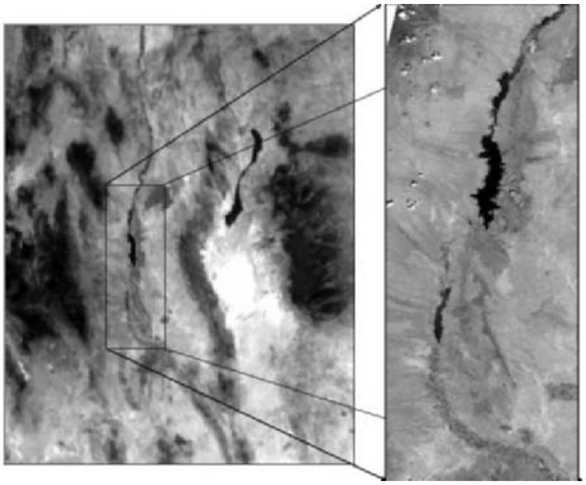

Figure 2. Rio Grande River Basin, SeaWiFS Image with MODIS higher resolution inset. This graphic shows a larger view (1 kilometer resolution) of the Rio Grande River Basin taken by the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) on the OrbView-2 satellite, with an inset of the same area (250 meter resolution) taken by the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument aboard the Terra satellite. In the inset, the largest dark area indicates Elephant Butte Reservoir. *Credit*: Image by Andrew French of NASA GSFC/NRC, Kristi Arsenault, NASA GSFC and Univ. of Maryland-Baltimore County.

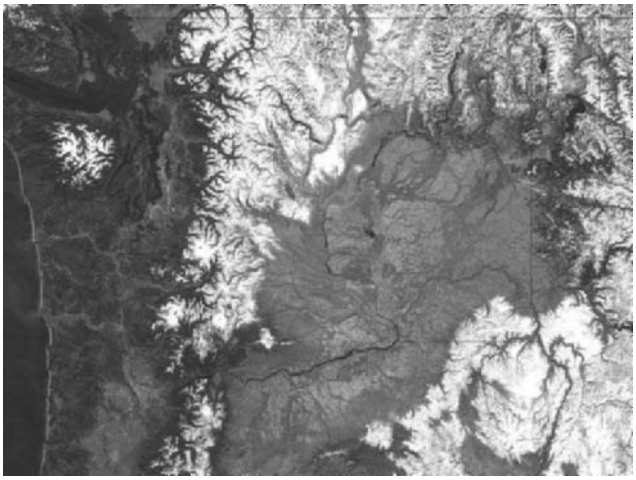

Figure 3. MODIS image of Columbia River Basin snowcover, February 24, 2003. This image from the Moderate Resolution Imaging Spectroradiometer (MODIS) shows snowcover for the Columbia River Basin in the Cascade Mountains of Washington State, taken on February 24, 2003 (250 meter resolution). *Credit*: Jeff Schmaltz MODIS Land Rapid Response Team, NASA/GSFC.

agencies are going to provide United States Bureau of Reclamation water resource managers with high resolution satellite data, allowing them to analyze up-to-date water-related information over large areas all at once.

The pilot program is now underway with the Rio Grande and Columbia River basins where water is scarce while demands range from hydropower, to farming, fishing, boating and protecting endangered species (Fig. 1). Water resource managers in these areas grapple with the big money stakes of distributing a finite amount of water to many groups. NASA satellite data offer to fill the data gaps in mountainous and drought-ridden terrain, and new computer models let users quickly process that data.

Land Surface Models (LSMs) from NASA, other agencies and universities, and NASA satellite data can be used to determine snowpack, amounts of soil moisture, and the loss of water into the atmosphere from plants and the soil, a process known as evapotranspiration. Understanding these variables in the water cycle is a key to managing water in such resource-limited areas.

"The latest satellites provide so much up-to-date and wide-ranging data, which we can use in the models to monitor and better understand what is happening with the water cycle in these areas," said Kristi Arsenault, research associate for the Land Data Assimilation System (LDAS) team at NASA's Goddard Space Flight Center, and Research Associate at University of Maryland, Baltimore County.

"These efforts are designed to improve the efficiency of the analysis and prediction of water supply and demand using the emerging technologies of the Land

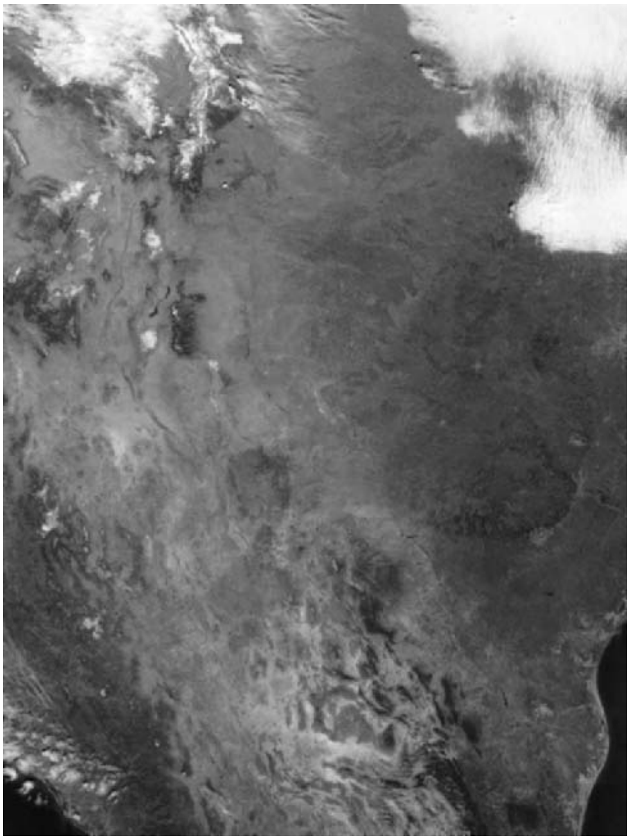

Figure 4. MODIS image of Rio Grande River Basin, February 22, 2003. This image from the Moderate Resolution Imaging Spectroradiometer (MODIS) shows the Rio Grande River Basin taken on February 22, 2003 (250 meter resolution). *Credit*: Jeff Schmaltz MODIS Land Rapid Response Team, NASA/GSFC.

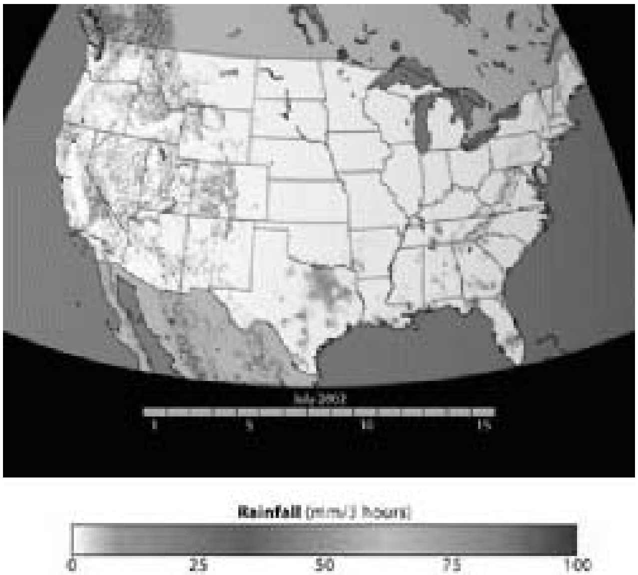

Figure 5. U.S. rainfall (July 1 to 15, 2002). This animation shows 3-hourly rainfall from July 1 through 15, 2002. Rainfall is one of the main variables that drive land surface models used by water resource managers and scientists. *Credit*: Images by Robert Simmon, NASA GSFC Earth Observatory, based on data provided by Kristi Arsenault, NASA GSFC and Univ. of Maryland-Baltimore County.

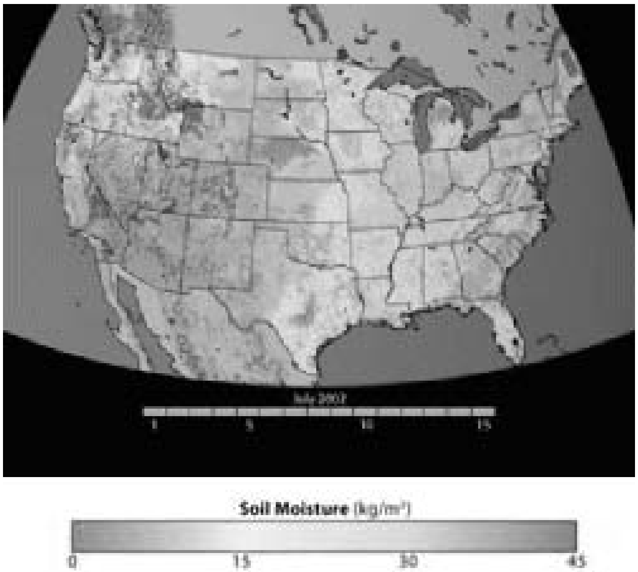

Figure 6. U.S. soil moisture (July 1 to 15, 2002). This animation shows changes in soil moisture from July 1 through 15, 2002, derived from a land surface model. Soil moisture is important for knowing how much water is contained in the soil, which is vital for crop production, flood prediction, and evapotranspiration estimates. *Credit*: Images by Robert Simmon, NASA GSFC Earth Observatory, based on data provided by Kristi Arsenault, NASA GSFC and Univ. of Maryland-Baltimore County.

Data Assimilation System," said Dr. Dave Matthews, manager of the River Systems and Meteorology Group of the Technical Services Center, U.S. Bureau of Reclamation (Reclamation). Computer models, known as decision support systems, that factor in ecological, human, and legal restrictions are vital to managing and allocating water, Matthews added. These systems will incorporate NASA satellite and model data.

NASA's tools may be of vital use in the Rio Grande and Columbia River basins where the disparate and numerous water demands have enormous economic implications. In the Rio Grande Basin, for example, water managers dole out water to farmers so they can irrigate their land. At the same time, under the Endangered Species Act, states are required by law to maintain river water levels to protect the habitat of the endangered silvery minnow. A recent seven-year drought has exacerbated these demands.

Similarly, the Columbia River Basin provides water for the Coulee Dam, the largest concrete dam in North America, and a means for controlling floods. This hydroelectric dam is the third largest producer of electricity in the world. At the same time, the basin is a source of water for a billion dollar agricultural area.

To help make big decisions of allocating water, NASA's special technologies can provide a unique perspective from space. For example, satellites can classify vegetation, a task that is essential to calculating evapotranspiration, which accounts for up to 60 percent of water loss into the air in a region like the Rio Grande Basin. Some managers have been relying on vegetation maps that dated back to 1993, in areas where wild-lands, crops and farming practices are subject to change.

Landsat data can provide highly detailed spatial information, but these images may only be available once a month, and are very expensive. The newer technologies of the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on the Terra and Aqua satellites provides more frequent passes and day-to-day and week-to-week changes in vegetation production (Fig. 2). In addition, other variables of interest, like snow cover and land surface temperatures, are updated more regularly by MODIS, which can aid in identifying areas with potential flooding and help with the daily management of the water resources (Figs. 3 and 4).

LDAS has also begun to evaluate soil moisture data from NASA's Advanced Microwave Scanning Radiometer (AMSR-E) aboard the Aqua satellite and 3-hour rainfall estimates from NASA's Tropical Rainfall Measuring Mission (Figs. 5 and 6). All this data helps determine how much water is being absorbed into the ground, versus how much is evaporating into the atmosphere. These observations will then be assimilated into Land Surface Models so that water managers can assess flood risks and other factors and act accordingly in a timely manner.

Reclamation brings water to more than 31 million people and provides one out of five Western farmers with irrigation water for 10 million acres of farmland.

One mission of NASA's Earth Science Enterprise is to expand and accelerate the realization of economic and societal benefits from Earth science information and technology.

TRANSBOUNDARY WATER CONFLICTS IN THE NILE BASIN

Rongchao Li
Delft University of Technology
Delft, The Netherlands

WATER SCARCITY IN THE NILE

The Nile, the longest river in the world, has nourished livelihoods, supported a vast array of ecosystems, and played a central role in a rich diversity of cultures. As shown in Fig. 1, from its major source, Lake Victoria in east central Africa, the White Nile flows generally north through Uganda and into Sudan where it meets the Blue Nile at Khartoum, which rises in the Ethiopian highlands. From the confluence of the White and Blue Nile, the river continues to flow northward into Egypt and on to the Mediterranean Sea. From Lake Victoria to the Mediterranean Sea, the length of the Nile is 5584 km (3470 mi). From its remotest headstream, the Ruvyironza River in Burundi, the river is 6671 km (4145 mi) long. The river basin has an area of more than 3,349,000 sq km (1,293,049 sq mi) (source: http://www.nilebasin.org/IntroNR.htm).

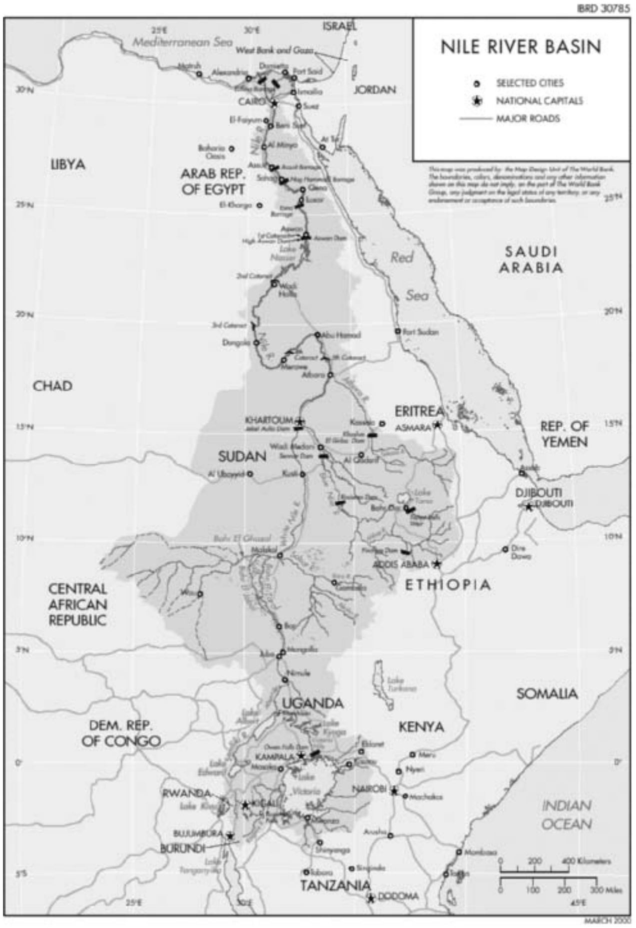

Figure 1. Nile basin map. (*Source*: http://www.nilebasin.org/nilemap.htm.)

As a typical transboundary case, the Nile River Basin is shared by ten countries: Burundi, Democratic Republic of Congo, Egypt, Eritrea, Ethiopia, Kenya, Rwanda, Sudan, Tanzania, and Uganda.

Since the early 1970s, water scarcity has become a matter of concern for the downstream states of the Nile. The engagement of the two downstream states (Egypt and the Sudan) in effective development of irrigated agriculture marked the beginning of real scarcity. The 1960s, in particular, marked the beginning when the two countries engaged in constructing large-scale water projects; by the end of the decade, three dams (Roseires and Khashm Al-Girba in Sudan and the High Dam in Egypt) were constructed. Large-scale agricultural schemes have become dependent on these dams.

In addition to damming the river and increasing demand for irrigation, population growth is considered a major cause of water scarcity in the Nile Basin. Since the late 1940s, probably all Nile Basin countries started to witness steady population increases. In the 1990s, the Nile Basin region experienced one of the highest population increments among the world's regions. Striking evidence of this is that 6 out of the top 20 countries in the world in population increments in 1991 were Nile riparians. The population of Ethiopia, Sudan, and Egypt will double from 157 million to 388 million by the year 2050. The total population of the Nile Basin will more than double when it increases from its 1992 figure of 259 million (1, p. 82) to 607 million by the year 2025. The Nile riparians will remain, well into the first half of this century, among those with the highest growth rate in the world; their rates range from 2.4% to 3.6% (1, p. 82). The most important conclusion to be derived is that more significant increases are taking place in upstream than in downstream countries (2, p. 2). Moreover, it is observed that the countries that have large populations are increasing their population faster than those that have smaller populations. Thus, except for the Democratic Republic of Congo, all large-populated countries will be in dire need of Nile waters as irrigated agriculture increases.

The 1980s witnessed yet another driving force in the region that contributed to water scarcity, the recurrent droughts and their consequential population concentration along the banks of the Nile and its tributaries. The 1980s and 1990s seem to provide new causes that have turned migration into "displacement." Referring to the Nile, Mageed (3, p. 159) stated that "The arid and semi-arid regions of the Basin are experiencing a serious breakdown of the environmental fabric and the spread of desertification along with the collapse of socio-economic systems." The part of the Nile Basin that is crossed by the Sudano–Sahelian zone has scored the highest rate (88%) of desertification in the early 1980s compared to the African average (84%) and the world average (61%). The consequences of drought and land degradation surfaced soon after the 1980s decade started and reached their peaks in famines in Ethiopia and the Sudan simultaneously in 1984. The effect of these, however, has hit all the Nile riparians in different degrees. These consequences have led/will lead to radical changes in the population map in Ethiopia and the Sudan. These famines have pushed large groups of population to the Nile River zone and thus contributed to real population concentration along the banks of the Nile and cultivable irrigable lands in its vicinity. Computations from population census data reveal that the river zone in central Sudan has increased its relative population size by 5.61% between 1983 and 2000, and similarly, in Ethiopia, the relative size of the population inside the Nile Basin increased by 6.5% in just a decade between 1984 and 1994. The same census data show that the population of urban areas in both countries is increasing along the banks and tributaries of the Nile/inside the Nile Basin. The population of Egypt has always been concentrated along the borders of the Nile except for some small oases in the deserts. Although the government of Egypt is developing urban areas outside the Nile valley, most of these developments still depend on Nile water for their water supply.

Unlike population concentration on a regional scale, population concentration in urban areas generates scarcity of a much higher magnitude (4, p. 90). This concentration generates water scarcity in several ways, including high demand for safe water for drinking and sanitation and high pollution rates, which render some water useless, or even harmful. However, as no further facilities have been installed, the concentrating of population has negatively affected the provision of safe water for urban dwellers; in Sudan, the rates decreased from 100% in 1980 to 86% in 2001. In all Nile Basin countries, except for Egypt, provision of safe water has been very limited. Water scarcity caused by population concentration is increased by the in-stream demand for hydropower generation, by the evaporation of reservoirs developed for hydropower generation, by increasing demand due to agricultural development, and by climate change due to global warming.

WATER ALLOCATION AND DISPUTES IN THE NILE BASIN

In reality, water scarcity already causes violence and conflict within and between nations, which is a threat to social and political stability. Peter Gleick, who makes an ongoing effort to understand the connections between water resources and international conflict and security issues, points out that "disputes over control of water resources may reflect either political power disputes, disagreements over approaches to economic development, or both. It is evolving as international and regional politics evolves..." (5). Water disputes typically are erupting within countries in the downstream regions of stressed river basins. Even if water disputes between countries do not lead to war directly, they have fueled decades of regional tensions and thwarted economic development before eventually giving way to cooperation. The Nile is an example in this respect.

Legal arrangements for the Nile in the nineteenth century were much more comprehensive than today in terms of legal obligations abided by all governments ruling over almost all the territory of the Nile Basin. Britain had been the main party to all these agreements; it signed agreements with Italy, Congo Free State, Ethiopia and Egypt. These agreements remained intact until independence. The decolonization process left the Nile with colonial agreements that were not respected by all newly independent riparians other than Egypt.

Immediately after acquiring its independence, the Sudan sought a new deal with Egypt over the Nile water to replace the 1929 agreement signed between the Sudan government (British administration) and Egypt. Conflicts arising at the postindependence stage, unlike the previous conflict, which was basically for securing the flow of rivers, were about the allocation of specified quotas of the now known as large quantities of water, yet scarce to meet the heightening demand. Hence, the Sudan started to confront the unfair division of water; it considered that Egypt's share was more than it actually should be, given that all the Nile water flows into Egypt, save 4 billion m^3 allocated to the Sudan by the 1929 agreement. For the Sudan, a new deal was thus seen as necessary. The challenge of the Sudan did not stem from preserving rights to the future realization of its agricultural potential; rather it was the obstruction of its already developed policies for bringing on board this competent potential. "The Sudan, to prove its case, embarked on drawing up the 'Nile Valley Plan'. The main object of the plan was to provide for the irrigation needs of Egypt and the Sudan and the assumed needs of the other riparian states for irrigation, together with the full development of hydro-electric potential in Uganda, Ethiopia, the Sudan and Egypt" (6, p. 73). The whole exercise was challenging Egypt's monopoly of the Nile by adding an element of egalitarian multilateralism. What made this grow to a challenge was that Sudan was seeking to mobilize the other riparians behind its plan. Egypt, however, was striving to start the implementation of the High Dam at Aswan. The Sudan, while opposing the Aswan Dam project in principle, had put forward three principles for resettling the water question. First, that Sudan's share should be determined before work began on the High Dam. Second, that Egypt must provide for the resettlement and "adequate alternative livelihood" of the people of the Wadi Halfa town and district who would be removed by the High Dam construction. Finally, that the Sudan should have the right in the future to build whatever works were deemed necessary for using her share of the Nile waters (7, pp. 173–174).

Upon accepting these conditions, the conflict between Egypt and the Sudan was bilaterally settled by the Full Utilisation of the Nile Waters Agreement in (1959). The 1959 agreement "established a joint committee (PJTC) and specifies quantitative allocations for the two countries: assuming an inflow of 84 billion m^3 from the Blue Nile and the White Nile sources, the agreement allocates 55.5 billion m^3 to Egypt and 18.5 billion m^3 to Sudan" (8, p. 17; 9).

The deliberations leading to the signing of the 1959 agreement, however, had created a serious conflict with Ethiopia. In 1957, Ethiopia sent a note to Egypt and Sudan in which it asserted its "natural rights" to use the water originating in her territory and in which it also referred to Ethiopia's projected water needs and requirements. These projected needs would certainly decrease the discharge of the Nile (10). The 1959 agreement was considered to favor the Sudan (11,12), dismember Ethiopia (12), and neglect the other upstream countries. The outcome of the 1959 Agreement had effectively put Egypt and the Sudan in one camp against the camp of all other riparians. The newly independent upstream states—individually, but sharing the same view—started to criticize Egypt and the Sudan's monopoly of the Nile waters. Egypt seemed to have accepted the idea that the upstream states are bound by agreements signed during the colonial time, but Sudan's position was different, namely, recognizing the rights of upstream states (10, p. 12).

In the eyes of the Egyptians, Ethiopia has always represented a potential threat to the Nile water. According to Hultin (13, p. 37): "It is not so much what Ethiopian governments—or other riparian governments for that matter—have done with regard to the waters of the Nile, but rather what they might be doing, that is the cause of anxiety in Cairo." Repeatedly, the Ethiopian governments asserted that they are not bound by any agreements on the Nile and that those agreements that were signed during colonial rule were either not signed by them or they were forced to sign by a colonial power. An important element in the dispute is the difference between the states in their starting argument. Egypt and the Sudan, bound by an agreement, based their contest for the Nile waters on real needs, but Ethiopia is driven by the awareness of its future needs and, therefore, is not ready to risk her unused resources (12). Egypt and the Sudan see that their 1959 Full Utilisation of the Nile Waters agreement can be the foundation for any future agreement among all riparians, but the other riparians, especially Ethiopia, want new arrangements.

FORMULATION OF A BASIN-WIDE WATER SHARING FRAMEWORK

The current environmental condition in general and the water scarcity issue in particular require new regimes of thinking and social organization or reorganization. As stated above, there are a variety of driving forces in water scarcity. For instance, until recently, the quarrel over the Nile waters was driven by the states' ambitions to develop economically. In the last two decades, this necessity to develop, in the Nile riparian, is being replaced by a much more urgent need—food security. Two conditions in this respect can be illustrated: one is a condition of abundance that allowed states to expand their irrigated large-scale agriculture endlessly to engage more and more in the market and the second is the condition of scarcity which brought new pressures that affect political stability and necessitate responses from states to groups demands. The condition of scarcity makes it a necessity that states adopt certain regulations to cater to the pressures arising.

The first attempt at cooperation among the Nile countries took place in 1960 through what is referred to as the Hydromet Survey Project, basically aiming to build a database for the Basin. In 1992, Hydromet was replaced by TECCONILE (Technical Cooperation Committee for the Promotion of the Development and Environmental Protection of the Nile Basin). TECCONILE continued until 1999 when it was replaced by the Nile Basin Initiative (NBI) as a wider framework with different operating organs to pave the way for final legal and institutional arrangements.

When competition reaches the crisis stage that the likely win–lose balance turns into a lose–lose balance, the contestants are likely to cooperate to survive the ordeal (14). Developments in Nile Basin politics in the last half of 1999 indicate that the co-Basin states have already reached the situation indicated above. Unlike the first half of the 1990s of national assertiveness and unilateral approaches to use of the Nile waters, the co-Basins now have (in the second half of the 1990) a new initiative, using the Technical Advisory Committee as its technical arm. The objective of the initiative is to reach a solution through the equitable use of the Nile waters. The conference of the Nile riparians in Addis Ababa in May 1999 moved further in this direction, and the Nile co-riparians have emphasized the sustainable development of the river. Thus, already the issue of equitable water use has been expressed together with the issue of sustainability.

Out of these initiatives, the shared vision of the NBI arose that is meant "to achieve sustainable socio-economic development through the equitable utilization of, and benefit from, the common Nile Basin water resources." The initiative as a comprehensive cooperative framework provides for significant changes in the old formula that would protect the co-riparians from inflicting harm on each other. Significant harm is always the consequence of competition, whereas equity and reasonableness can be attained through cooperation. The first is often defined by ideological and political imperatives, whereas the latter is for calculating mutual benefits or avoiding mutual losses—the economic imperatives. In this regard, the request of the Nile-COM for the World Bank and its partners to host a consultative group (ICCON) as a forum for seeking funds is a significant move toward emphasizing the economic imperatives. The Ethiopia entry into the debate on legal arrangements is likely to cool down the previous tension and create a condition conducive to cooperation.

The emphasis on sustainability is what makes the NBI different from the previous initiatives. Moreover, unlike the previous ones, it is a vision shared by all riparians. It means, in short, that awareness of the current environmental problems that affect all riparians is finally on its way toward institutionalization. It is the emphasis on sustainability that gives way to the newest part in the Initiative: the subsidiary action subprogram. The principle of subsidiary is considered "an important approach to cooperative action within a Basin-wide framework" and is meant "to take decisions at the lowest appropriate level, to facilitate the development of real action on the ground." It is this part which gives place to alternative developments by decentering the definition-making and decision-making about the Nile waters, which has been always the monopoly of central governments or bodies representing them. Building the pillars of the temple of cooperative framework (confidence building and stakeholder involvement; socioeconomic, environment and sectoral analysis; development and investment planning; and applied training) in itself involves a necessary learning process in which different levels will contribute. It, thus, informs the new priorities to be set for the Nile Basin countries as it translates to "capacity building and human resource development." As well, it makes place for priorities depending on the needs and specific problems facing each country or a sub-Basin that combines more that one country. It caters to the communities dwelling in the Basin that would be affected by the grand policies to be launched.

Another important aspect of NBI is that it will also consider measures outside the basin when this will ease the problems in the basin, in particular, measures with respect to population and agriculture. Their approach will be based on equitable regional economic development to ensure the balance between upstream and downstream *and* between inside and outside the basin.

What complicates the issues of water scarcity in the future is that the Nile Basin countries are far from providing water efficient alternatives. Thus, the development of the river, necessarily sustainable, is an urgent necessity.

PLAUSIBLE SOLUTIONS

Despite the fact that the Nile River is viewed as the archetype for any future conflict over international water, recent developments show that through cooperation, the Nile riparians can overcome many of the hurdles and establish a peaceful and sustainable solution to the problems of water scarcity in the Nile. What is needed, therefore, is:

1. Principle of equity. Egypt and the Sudan must accommodate the position and demands of Ethiopia and other upstream countries in an institutional manner (15). Ethiopia which views itself as being dismembered by the 1959 Nile water agreement will likely offer greater cooperation should the downstream riparians recognize its right to development of its part of the Basin and facilitate it. If Ethiopia is allowed to go ahead with its Blue Nile basin plan, Egypt and the Sudan would benefit from it (see also Infrastructure Solutions). Governments and international organizations must act early and constructively. Tensions among the Nile basin countries are finally easing, thanks in part to unofficial dialogue among scientists and technical specialists that have been held since the early 1990s and more recently at a ministerial level in the NBI. The key is establishing a process of early cooperation before serious hostilities erupt that make it difficult for nations to sit around a negotiating table together.
2. Strong institutions. Treaties that provide for effective monitoring and enforcement are often remarkably resilient, holding even when the signatories are engaged in hostilities over nonwater issues. Long-term programs of joint fact-finding, technical cooperation, and other initiatives that establish a climate of cooperation among countries can pave the way for resolving disputes when they do arise.
3. Effective rural development. To our understanding, the major problems that the Nile Basin countries are facing now are not merely the need for sharing water; rather they have to do with what makes the quest for sharing more feverish. In other words, environmental degradation and population concentration makes the quest for more water endless. So dealing with these problems and similar ones is what is urgent for all Nile riparians. The NBI can be seen as a framework which allows overcoming some of the factors contributing to water scarcity, including rural development. The multiple billions of dollars that are expected for the development of the Nile basin should be spent on large-scale hydraulic constructions and should also partly be spent on rural development in the territory of the co-riparians—not necessarily inside the Basin.
4. The import of "virtual water" represents one important element in how states could solve their water scarcity problems. In the Nile, this virtual water importation can be thought of in a more innovative way. The Nile Basin countries differ in climate; therefore, crops produced as well as the natural hazards, solutions, and therefore cooperation should consider these differences. Ethiopia faces severe soil erosion due to deforestation, and this very soil erosion, among others, is what pushes Ethiopia toward using Nile water. A fund for forestation and then benefiting from forest products (service for the ecosystem, fruits and other products which become part of the livelihood of rural population, and products for exportation) would thus produce a double advantage. Downstream as well as upstream Nile riparians can import virtual water from Ethiopia necessary for the rehabilitation of the latter. The Sudan, viewed as the breadbasket of Africa and the Middle East, could import virtual water from Ethiopia and export virtual water too. The difference, however, is that virtual water from Ethiopia, in the form of afforestation products, would help rehabilitate this country's highlands, therefore, decrease its demand for Nile water. Similarly Egypt can import virtual water from Ethiopia and the Sudan, through a formula of less virtual water from downstream and more virtual water from upstream—less water-consuming products versus more water-consuming products (rice).
5. Infrastructure solutions. Establishment of reservoirs on the Blue Nile in Ethiopia is suggested as "the greatest opportunity over the long term for dramatic improvement in the overall management of Nile resources" (16, p. 152). What is interesting in this regard is that the water savings so made, which could be of the order of 12–21.4 billion m^3 per year (17), would quadruple Ethiopia's irrigated area without reducing supplies to Egypt and the Sudan. Some other solutions are transfer of storage from reservoirs in Egypt and the Sudan to Lake Tana and the Blue Nile basin combined with water conservation in the Ethiopian highlands (by way of reducing evaporation and annual flooding losses) and will provide regulated flow to downstream countries. Similar schemes could be implemented in the equatorial lake regions of the upper White Nile riparian countries,

resulting in additional water saving construction of a bypass canal (the Jonglei diversion canal) in southern Sudan reducing the enormous evaporation losses and increasing the contribution of the White Nile to the main Nile appreciably. The Machar Marsh and Bahr El Ghazal in southern Sudan can contribute 11 billion m^3 (18, p. 139; also see 1, p. 92).

6. Water saving technologies. In all cases, maximum use should be made of water saving technologies (more crop per drop). This includes water saving options in the surface water system (reduction of losses of seepage and evaporation), more efficient irrigation techniques (land leveling, drip irrigation, etc.), and other management techniques (reuse, nightirrigation, etc.).

BIBLIOGRAPHY

1. Shapland, G. (1997). *Water in The Middle East and North Africa: The Political Economy and Ethics of Scarce Resource Use*, compiled and Edited by Timothy C. Weiskel, For the Harvard Seminar on Environmental Values (1997–98) “Water—Symbol and Substance of Life”.
2. Omar, A. and Takele, T. (1998). *Egypt and the Horn of Africa: That Serpent of Old Nile*, Ethiopia Tribune (26.06.1998).
3. Mageed, Y.Abdel (1994). The Nile Basin: Lessons from the Past. In: *International Waters of the Middle East: From Euphrates-Tigris to Nile*. A.K. Biswas (Ed.). Oxford University Press, Bombay.
4. Falkenmark, M. Lindh (1993). *Global Changes and Water Resources*, www.water.hut.fi/wr/research/glob/egloshow/egloshow2.pdf.
5. Gleick, P.H. (2000). The world’s water 2000–2001. *The Biennal Report on Freshwater Resources*. Island Press (Ed.). Island Press, Washington, DC, p. 315.
6. Mageed, Y.A. (1981). *Integrated River Basin Development: A Challenge to the Nile Basin Countries*, in Sudan Notes and Records No. 62, Vol. LXII.
7. Holt, P.M. and Daly M.W. (1979). *The History of the Sudan. From the Coming of Islam to the Present Day*, 3rd Edn. Weidenfeld and Nicolson, London.
8. Tadesse, H. (1998). The Nile river basin. *Ethioscope* **4**(1).
9. Waterbury, J. (1979). *Hydropolitics of the Nile Valley*, Syracuse, Syracuse University Press.
10. Khalid, M. (1984). The Nile waters: the case for an integrated approach. In: *The Nile Valley Countries: Continuity and Change* (Vol. 1), M.O. Beshir (Ed.). Institute of African and Asian Studies—University of Khartoum—Sudanese Library Series (12), Khartoum.
11. Davies, H.R.J. (1984). “Continuity and change in the nile valley” a geographical viewpoint. In: *The Nile Valley Countries: Continuity and Change*, Vol. 2, M.O. Beshir (Ed.). Institute of African and Asian Studies—University of Khartoum—Ethiopiaese Library Series (12), Khartoum.
12. Tilahun, W. (1979). *Egypt’s Imperial Aspirations over Lake Tana and the Blue Nile*. United Printers Ltd, Addis Ababa.
13. Hultin, J. (1995). The Nile: source of life, source of conflict. In: *Hydropolitics: Conflicts over Water as Development Constraint*, Leif Ohlsson (Ed.). University Press, Dhaka and Zed Books, London.
14. Wolf, A.T. (1997). *Water Wars’ and Water Reality: Conflict and Cooperation Along International Waterways*. NATO Advanced Research Workshop on Environmental Change, Adaptation, and Human Security. Budapest, Hungary, 9–12 October.
15. Tafesse, T. (2001). The Nile question: hydropolitics, legal wrangling, modus vivendi and perspectives. Münster, Lit. pp. 131–145.
16. Whittington, D. and McClelland, E. (1992). Opportunities for regional and international cooperation in the Nile Basin. *Water Int.* **17**: 144–154.
17. Kliot, N. (1994). *Water Resources and Conflict in the Middle East*. Routledge, London.
18. Naff, T. and Matson, R. (Eds.) (1984). *Water in the Middle East: Conflict or Cooperation?* Westview Press, Boulder, CO.

PLANNING AND MANAGING WATER INFRASTRUCTURE

Neil S. Grigg
Colorado State University
Fort Collins, Colorado

Water infrastructure consists of the constructed elements that control water for water supply, wastewater service, flood control, hydropower, recreation, navigation, and environmental protection (see articles on Water Resources Management and Water Infrastructure and Systems). Generally, these constructed elements involve conveyance systems, including channels, canals, pipes, and bridges; diversion structures; dams; reservoirs; locks; treatment plants for water supply and wastewater management; pumping stations; hydroelectric plants; spillways, valves, and gates; and wells. Planning and managing water system capital assets require provisions for capital investment, operations, and maintenance.

Capital investment in water facilities requires a careful process that begins with planning them, preparing the designs, and constructing the systems. The goal of capital investment is to obtain a maximum return over the project’s “life cycle.” This means considering costs and management needs over the full life of the facility and requires that operations and maintenance, including corrective maintenance, be carefully conducted to gain maximum usage and life from the investment.

The planning process involves a series of steps that include identifying the need; searching out alternative ways to meet the need; evaluating economic, environmental, and social costs and benefits; making decisions about the alternatives to pursue; and implementing the plan. The steps in planning drive the process, but they involve many complex issues and substeps. Each project has unique features and stakeholders, political issues in gaining approval for infrastructure often control the pace and agenda. Traditional planning processes now incorporate sophisticated environmental, social, and economic investigations and are now much more complex than in the past. For example, the United States’ experience with environmental impact statements required by the National Environmental Policy Act shows that it costs more to add environmental impact analysis to planning, and the decision process becomes more complex.

In an attempt to rationalize complex criteria, the concept of "multiobjective evaluation" was developed to consider economic development, social benefits, and environmental enhancement when planning for water purposes such as water supply, power, and flood control. This concept was included in the "Principles and Standards" that emerged from the Water Resources Planning Act. These later were changed to "Principles and Guidelines."

For a minor component, such as a new pump station, the process may be straightforward and short term. For major facilities such as dams, the process may take many years and, in some cases, ultimately not succeed at all. In addition to the planning steps cited, practical assessments such as how to pay for facilities and identifying legal obstacles are required. Also, stakeholders must be involved at each step of the process through public involvement and consultation.

The design process involves creative decision-making about the configuration and details of projects. Design includes drawings, documents, and plans necessary to initiate construction. Construction begins with preparing contract documents and involves bidding, review, award, organization, construction itself, inspection, and acceptance. Keeping all of the phases going in the proper direction and order is the task of the "project management process." Designing and constructing major water resources projects require high levels of skill and organization involving risk to the public and environment. For example, dam construction may require control of an entire river over several seasons when floods might occur. A wastewater treatment plant might involve large-scale and sensitive biological processes.

After construction is completed, the project enters the "operations and maintenance" or "O&M" phase. Large water systems such as dams and reservoirs are subject to special operating rules. Water treatment plants require trained operators who can monitor chemical and biological parameters while adjusting operating controls. Far-flung systems extending over several watersheds may have sophisticated supervisory control and data acquisition (SCADA) systems, consisting of instruments, wireless and wired communications, computers, and control systems.

Decisions about water system operation are controlled by complex sets of rules and guidelines. In some cases, regulatory controls are very stringent, and violations result in heavy penalties. Operations may be based on computer simulations and multiobjective criteria, as in following rule curves for reservoir operation. A science of *water resources systems analysis* has emerged to provide tools and techniques for operating complex systems.

Maintenance management is also required in a capital management system. Without adequate maintenance, water resources infrastructure systems will usually not perform well and will wear out sooner than planned. The maintenance management system brings activities together for investment, organization, scheduling, and monitoring. Its general functions are inventory, condition assessment, preventive maintenance, and corrective maintenance.

An inventory of facilities, components, and equipment is the cornerstone of maintenance management. Inventory methods differ for real property, fixed assets, and equipment. An inventory can be as simple as, for example, a set of drawings indicating where sewer pipes are located in a section of a city. These drawings, with annotations, are used by maintenance forces to locate and service pipes. On a more sophisticated level, the drawings could indicate other, nearby facilities, such as water and electricity lines, and can also be used to coordinate between services. Even more sophisticated would be coordinated data in a GIS and database format, available on a common basis to different sections of the organization whose work involves shared data, processes, facilities, and staff.

Condition assessment requires inspection and analysis of condition so that maintenance activities can be planned and scheduled. It is used to plan and schedule capital management activities, such as maintenance, needs assessments, budgeting, and capital improvement programs.

Preventive maintenance (PM) is the ongoing program of care given to equipment or components. In general, PM requires consistent, timely completion of tasks prescribed by documented procedures according to set schedules that include regular follow-up. Information sources for PM are O&M manuals, product information, and experience of workers. PM records include equipment data, the preventive maintenance record, the repair record, and a spare parts stock card.

Generally speaking, corrective maintenance means to repair equipment or components that have failed or deteriorated. It can range from minor to major repair and drives the "3 R's" of infrastructure—repair, rehabilitation, and replacement. Corrective maintenance requires a decision if the deficiency is minor or major enough to require capital budgeting. If the problem is major, the capital plan and budget incorporate information about new standards and growth forecasts to lead to decisions about rehabilitation and/or replacement.

When a water resources facility has reached the end of its useful life, it must be removed, rehabilitated, or replaced. Many dams around the world are now 50 or more years old and will face these decisions in the future. Wastewater treatment plants built during the 1970s in the United States are now three decades old. In addition to new investment to handle population growth, a discussion will occur in the future about strategies for reinvestment and modernization of these key infrastructure components.

APPLICATION OF THE PRECAUTIONARY PRINCIPLE TO WATER SCIENCE

Leonard I. Sweet
Engineering Labs Inc.
Canton, Michigan

Discussions of applying the precautionary principle have received new momentum from several highly contentious policy issues, including the bovine spongiform encephalopathy crisis, blood donor screening, genetically modified foods, antibiotic feed additives, and phthalates in baby toys. The level of precautionary interpretation and its application to "decision-making under ambiguity" has wider ramifications for waste policy, ecological toxicology, and especially water science. A balanced debate is warranted over the merits of the two seemingly rival environmental risk management paradigms—risk assessment and the precautionary principle.

The basic definition of the precautionary principle has three elements: threat of harm, scientific uncertainty, and the precautionary action. As such, the principle questions whether scientific knowledge is strong enough to permit reasonably confident decision-making, and it favors preventive action in the face of uncertainty as well as a participatory public in decision-making. The precautionary principle can be applied in general on two fronts, the *ex ante* side that reflects a desire to avoid or prevent undesirable adverse effects in policy frameworks and legislation, and on the *ex post* side, reflecting the fact of having to live with consequences of past decisions, side effects, or accidents (e.g., historical liability or natural resource damage assessments).

The relative constraints and uncertainties of risk assessment suggest to some that risk managers should abandon the risk framework in favor of the precautionary approach. In water science, it is not prudent to propose one as an alternative to the other. An immediate concern is that the precautionary-based approach fails to account for risk transfers, priority substances, and costs versus benefits. It is proposed that a variety of precautionary initiatives and decision analysis tools in risk management (e.g., life cycle assessment, weight of evidence, cost–benefit), as well as the open exchange of information and rational thinking, will aid in bridging the dichotomy between risk assessment and the precautionary principle.

INTRODUCTION

The momentum toward a paradigm shift to precautionary approaches will be faced globally by government and industry, and those in water science will have to rethink radically their current safety, risk management, and regulatory control schemes. The premise of precautionary decision-making is that environmental or human health threats are challenging to comprehend and resolve and that it is necessary to establish an environmental management paradigm to regulate suspect activities before they cause harm. Although originally intended for environmental management issues, the precautionary principle has now expanded to address human health and safety issues (i.e., toxicology, epidemiology, industrial hygiene, medicine). For example, it is likely that a precautionary approach was used in addressing the health and safety risks of using thalidomide because the precise mechanism by which thalidomide impacts the fetus was not characterized until the late 1980s.

In general, precautionary approaches instruct environmentally protective decisions before any potentially harmful effects of a given action or technology are fully proven. There are varying degrees of interpretation of the precautionary principle, including the most stringent or maximally precautionary versus the most permissive or minimally precautionary: (1) strong precaution (where the scientific evidence is uncertain but the public is informed and the evidence indicates that the worst case is the most probable outcome), (2) strong precaution (where the evidence is too weak to justify abrupt changes, given costs, but policymakers take a holding pattern and halt further intervention until more risk evidence is available), and (3) weak precaution (where the evidence is overly weak to justify change in public policy, given costs and other factors, and the public is informed of possible risk).

It can be seen that a regime based strictly on the precautionary principle places the burden of proving the harmlessness of a given action on the party who would engage in that action. This version of the burden of proof policy is a departure from the traditional toxic tort approach in which no harm is presumed to result from the activity of another until a party can demonstrate damage and causation. In the environmental arena, the tort-oriented nonprecautionary approach postpones the regulation of activities until harm occurs and compels a response to potential degradation. Consistent with this continuum of approaches is the presumption that precautionary action is acceptable as an interim measure because, as the scientific evidence accumulates, it is only a matter of time until an adequate level of certainty is reached.

It is apparent that in the governance of risk, the risk-based and precautionary approaches are divergent with regard to scope, risk, and how best to handle uncertainties and values (Table 1).

Rather than seeing that precaution is in conflict with science-based regulations, some suggest that the precautionary approach is consistent with sound scientific practice: (1) in acknowledging and responding to difficult problems in risk assessment such as ignorance (we don't know what we don't know) and incommensurability (we have to compare apples and pears) and (2) in critically thinking about the "use" of scientific knowledge.

GLOBAL SCOPE

Reflecting the divergent values of risk management, the precautionary principle has been defined, and the relative importance weighted differently by various nations (e.g., lack of firmly established international or legal meanings of the precautionary principle). The inconsistency in the precautionary approach is further complicated because actions within one state may cause identifiable harm to another state's environment, which in turn may create international responsibilities.

In a global economy, there are certainly different country-specific circumstances for interpreting the precautionary principle, including the level of industrialization, governance, economic and technical contexts, and human resources available for monitoring and evaluating problems connected with chemical risks. The current perception is that Europe believes that the earlier one exercises precaution, especially when cumulative pollution and diverse chemical risks do not respect boundaries, the more effectively it can be done.

Table 1. Highlights of Risk-Based and Precautionary Frameworks

Risk Framework	Precautionary Framework
Scope: • origins in quantification and hypothesis testing • premise of threshold and assimilative capacity • focus on individual and population risk • methods (i.e., hazard identification and characterization, appraisal of exposure, risk characterization) • adheres to tenets of toxicology (i.e., Paracelsus, Pare's, Haber's, Arndt–Schulz)[a]	Scope: • origins in subjectivity and emotion • does not support assimilative capacity or threshold in environmental/human health • focus on values and a unifying perspective of precaution • methods not well defined; no cause–consequence chain • adheres to caveat of *primum non nocere* or "first do no harm"
Chemical innocent until proved guilty: • "prove harm" regulatory approach • burden of proof on regulatory agency, customer, or public • pollution prevention/minimization opportunity if established threat • focus on preventing false positives	Chemical guilty until proved innocent: • "prove safety" regulatory approach • burden of proof on innovator or manufacturer • ensure safety or ban agent/activity before established threat • focus on avoiding false negatives
Optimistic view of • best available data • default values • analytical techniques • modeling methods • margins of error • statutory time constraints • reasonable estimates	Pessimistic view of • data (inaccuracies, gaps, unsuitability) • modeling (overly complex, nonrandom) • analytical (detection limits, interference) • timescales (inadequate for decision-making and legislation) • uncertainty (e.g., smokescreen vs. scientific indeterminacy)
Risk recommendations viewed as • entrenched in model and method uncertainty • carried out on product versus process • subject to cost/benefit analysis	Risk recommendations viewed as • proactive and adaptive in the face of new information or the unforeseen • not subject to cost/benefit analysis

[a] Paracelsus: dose–response contains the notion of threshold of effect; Arndt/Arndt–Schulz: predicts continuum of physiological mode of action dependent on strength of stimuli, that is, from hormesis/stimulation at low levels to retardation at high levels; Haber's: includes the variables of time and dose as a function of exposure to compare hazards better and formulate thresholds of toxicity; Pare's: the specificity and toxicity of chemicals are due to their unique structure and the laws of biology that govern the response interaction.

Regardless of the geographical complexities, however, there appear to be two consistent global elements of the precautionary principle: first, insufficient scientific information and second, significant but uncertain risks.

Consider the European policy on chemicals and the environment. Europe regards the precautionary principle highly, and erring on the side of caution when knowledge is limited and prediction speculative is viewed as a sensible and sound approach to good governance. The European community has also adopted a variety of principles that are directly or indirectly also designed to improve implementation of the precautionary principle, including the following: (1) polluter pays, (2) preventive principle, (3) sustainable development, and (4) inter-/intragenerational equity. Taken together, the core direction of these principles suggests that liability, at a minimum, should enhance incentives for more conservative, proactive, preventive, future-oriented, and responsible actions.

The European community's approach to the "polluter pays" principle could lead to a structure of environmental liability based on increased levels of precaution and prevention. Such an approach may also introduce new damage assessments, such as environmental damages to nature or biological diversity (e.g., due to the discharge of hazardous substances into waterways) when polluters are identifiable, where the damages are quantifiable and concrete, and where a causal link is established between polluters and damage.

Although there is no fully agreed upon version of the precautionary principle, it is being embraced extensively in international articles, treaties (International Joint Commission on Great Lakes; EU Institutional Treaty), summits (Rio Earth Summit), laws (e.g., German Vorsorgeprinzip legislation; Dutch wildlife laws), conventions (e.g., United Nations' Convention on Law of the Sea, Climate Change, and Biodiversity; UN General Assembly World Charter for Nature), conferences (e.g., 1998 Wingspread Conference Center), and other policy-making formats (e.g., Communication from the European Commission about guidelines for EU research activities; Expert Advisory Board on Children's Health and the Environment). The precautionary principle is clearly viewed as a logical corollary of the established international law norm that no state has the right to engage in activities within its borders that cause harm to other states.

Although the various interpretations of the role and application of the precautionary principle revolve around science and policy, the issue is steadily subjected to political motivations in administration and legislation. In particular, precautionary action has elevated such issues as genetically modified organisms, acid rain, radio-frequency fields, toxic waste disposal, and climate change. However, the precautionary principle is currently lacking in providing evidentiary standards for safety, (e.g., with medicines it revolves around quality, efficacy, and dose–response constancy), as well as criteria for obtaining regulatory approval. This is a serious shortcoming because it is conceivable that practitioners of precaution could be overcautious or unscientific in considering whether new technologies pose unique, extreme, or unmanageable risks, without properly characterizing the many new aspects that may confer net benefits and total hazard reduction. Such a discriminatory policy could inflate the costs of research or remedies, result in high taxation of products, waste resources, inhibit the development of new technologies, restrict consumer choice, and result in zero hazard reduction, all potentially for theoretical risks.

DRIVERS TO ADOPT

Proponents of the precautionary principle believe that existing environmental regulations and other decisions based on risk appraisal techniques have failed to protect environmental and industrial health. Recognized drivers for adopting the precautionary principle include the following:

- risk aversion
- reduced reliance on the quantitative risk in favor of qualitative techniques and participative deliberation
- strong commonsense appeal to general public
- working to safe minimum standards
- adopting sustainability efforts
- following a win–win/least regrets regime with regard to costs and benefits
- primary prevention
- data deficiency or uncertainty challenges, modeling difficulties, analytical difficulties, and insufficient time for adequate research
- commonsense appeal of having the polluter pay
- those in risk governance are frequently asked to provide the final word of advice on safety or risk before they have sufficient proof
- concern that too much focus on risk assessment could delay needed protections
- provides additional rationale for environmental and public health surveillance efforts, to use indicators, and detect the probability of adverse consequences as early as possible.

But in practice or in policy implementation, the precautionary approach is not very popular among experts in the risk arena because the potential benefits from the drivers may not warrant the cost and it is a challenge to make the principle a unified perspective that addresses many of the concerns.

ANTIDRIVERS

Clearly, however, there are antidrivers to adopting the precautionary principle. The main ones include:

1. the reverse onus that shifts the duty of care to be placed on the developer's actions to "prove safety," which may ignore the tenets of toxicology that demonstrate a need to regulate chemicals based on both the beneficial as well as adverse; the other concern is that this is potentially counter to notions of liberty and freedom by shifting the burden of proof from government onto innovators
2. the apparent advocating of controversial measures, including bans or phaseouts, clean production, alternatives assessments, organic agriculture, ecosystem management, health based occupational exposure limits, and a requirement for premarket testing
3. the apparent lack of cost–benefit analyses, in that for many cases that call for precautionary action, the marketplace does not necessarily play a critical role and good cost–benefit information is not available
4. the belief that the precautionary principle allows controlling important resource allocation issues by fear, emotion, and politics and that becoming more "safe" in some areas may involve spending resources that cannot be used doing more 'good' in other areas
5. the continuing challenge for proponents to frame an adequate and harmonized version of the principle. This is a reflection that the principle is still evolving—to some, it is already part of the risk appraisal arsenal, to others it is not yet on the radar screen, and to still others it is the antithesis of risk assessment.

Other concerns in adopting precautionary approaches revolve around the implications for the relationship between environmental protection and economic development, the economic consequences. Technologists and scientists are concerned that the precautionary principle is irrational and potentially based on nonobjective science and that it is anti innovation/antibusiness. The concern is that the precautionary principle is biased toward noncorporate and nonindustrial opinions, is susceptible to manipulation with regard to trade and equality, and may become an excuse for arbitrary restrictions that may stifle progress, prevent the use of newer and safer technologies, and possibly create more risk or uncertainty. Not every action or technology that has associated uncertainty should trigger precautionary action, and furthermore, every chemical should not be a presumed or preestablished carcinogen or endocrine disruptor.

The reticence to implement precautionary approaches fully is likely to remain until a more wise and balanced approach to chemical risk management is established. Most would agree that uncertainties in major decisions need to be weighed, analyzed, and communicated to impacted parties, but a combination of options in addition to the precautionary principle may prove most effective in managing risk (e.g., perhaps prioritized precautionary thinking in tandem with risk characterization and the use of current and better/best control technologies). It is also clear that the growing interest in precautionary approaches is more likely a reflection of shifting values on risk tolerance and governance versus a preference for subjective versus objective problem-solving techniques. Herein lies much of the controversy around its implementation—what constitutes sufficient proof, proof of safety, or proof of negligence?

RISK TOLERANCE

One certainly must consider that the tolerability of individual risk and the regulation of societal risk are integral to our government and culture. Yet the concepts of safety and risk are relative and subjective. In some areas of regulation, an explicit treatment of risk is warranted, whereas in other areas, risk appraisal is more implicit and reliant on best professional judgment. But how safe is safe enough with regard to personal, national, or international risk criteria in water science?

To address this question practically, one must characterize the time frame and magnitude of the hazard and decide on a common language—the language of risk analysis or the precautionary principle.

- *Risk Analysis.* Risk-based target levels, health-based criteria, standards as low as reasonably practicable, reasonable probability of no harm, an adequate margin of safety, proportionality, weighing costs and benefits, virtually safe doses, allowable daily intakes, benchmarks, maximum allowable tissue concentrations
- *Precautionary Principle.* Common sense, an ounce of prevention is worth a pound of cure; better safe than sorry; a stitch in time saves nine; if in doubt, don't pump it out; sentence first, verdict afterward

Apparently lacking in the language of the precautionary principle are a cost–benefit analysis, risk communication issues (e.g., regarding whether the hazard is imposed upon, entertained voluntarily, controllable, inequitable, and to what level the public is averse to the issue), and a good characterization of the boundaries and dimensions of uncertainty (i.e., ignorance, systemic, indeterminacy).

Risk informed decision-making must involve some level of cost–benefit analysis to focus resources and set priorities so that the greatest social benefits are achieved at the lowest cost. Some practitioners may strive to make a formal cost–benefit analysis the determining factor for risk tolerability. However, a formal cost–benefit analysis seemingly contradicts the precautionary principle, as it assumes certainty where, by definition, certainty does not exist. Furthermore, there are some concerns that requiring a cost–benefit analysis would render the precautionary principle less effective, cumbersome, and bureaucratic. Cost–benefit analysis may not be an appropriate methodology for managing uncertainty, as it can be applied only if there is full knowledge of the cost of that damage—and this is challenging because precautionary actions are not based on demonstrated risk, but rather on anticipated risks that are considered plausible.

UNCERTAINTY

Uncertainty refers to the failure to address fully a standard of proof/safety required by decision-makers and to the recognition that one cannot fully characterize the extent and potential seriousness of the consequences of not meeting that standard. In the face of inherent and irreducible uncertainty, it is clear that risk management and precautionary principle paradigms need to be adaptive and consider statistical significance (i.e., power, probability of type I or type II error), as well as environmental significance.

What level of causation, threshold of confidence, or lack of proof is required to trigger use of the precautionary principle to link an exposure to some possible harm? Various trigger criteria have emerged, including reasonable suspicion, potential damages are serious or irreversible, scientific certainty, cost-effectiveness, a single complaint, a single case, a single animal study, a combination, only where the probability and cost of impacts are completely unknown, or some gradient or level lower than the balance of probabilities or prevailing probability. Whatever the trigger, the precautionary principle should be subject to review and assigning responsibility, as well as subject to the European Commission criteria of proportionality, nondiscrimination, consistency, and cost–benefit analysis. The precautionary measures must also be proportionate to the benefits to be achieved. Hence, there is a need for an open and transparent procedure to identify the best options for harm avoidance.

Uncertainty in detecting and evaluating chemical hazards leads to challenges when conducting risk analyses (e.g., robust estimates of exposure and risk for pharmaceuticals in drinking water are not available). In the absence of a scientifically rigorous approach to evaluating the risk potential, actions may be guided by societal perceptions of risk or safety (e.g., refractory pharmaceuticals in wastewater) and may lead to malignment, restriction, or banning of chemical agents. One has to balance carefully the perceived reduction in health risks to society from precautionary action and the denial of public access to materials or technologies that, under appropriate conditions of safe handling and use, would not result in notable risks and may have substantial net benefits.

In the end, one must recognize that the scientific basis of modern water policy involves varying degrees of uncertainty. But it is more certain that risk analysis and precaution converge at the level where they are issue driven and that, regardless of the paradigm or language agreed upon, risk informed decision-making in water science must be efficient and rational, even when data are uncertain and the lines of evidence are imperfect.

POTENTIAL RELEVANCE TO WATER AND ENVIRONMENTAL SCIENCE ISSUES

The precautionary principle can be applied to emerging public concerns, such as large-scale environmental and health issues, and to the associated changes in regulatory frameworks. Most notably, issues include major industrial accidents, agricultural and food sector (e.g., mad cow disease), dioxin exposure and long-range transport, the irreversibility of global issues (e.g., ozone layer depletion, enhanced greenhouse effect), and biotechnologies (e.g., cloning and genetic modification). For watersheds, management decisions based on the relationship between stressors and resources may include act now, protect, restore, or watch.

These are current examples of the precautionary principle in action in the regulatory and environmental arenas:

- *DDT Use to Combat Malaria.* The precautionary principle demands no risk taking when there are unpredictable consequences—the potential environmental and health risks posed by future use of DDT—and hence there is a strong push not to allow it to be used or manufactured. The millions of people suffering from malaria that could have benefited from using DDT simply do not understand the argument of focusing only on potential future risks versus present and clear benefits: the economic benefits of the chemical if the threat of malaria were reduced, health care costs reduced, needless morbidity and mortality reduced, that is, the need to balance the potential risks against the certainties and benefits.
- *Global Climate Change.* With climate change, proof will be validated only after many years of data collection. Given the scope, potential injury, and irreversible consequences, the available information should be given great deference.
- *Biodiversity and Wetlands Loss.* Consider unforeseen circumstances with regard to the Endangered Species Act in the United States; it does not require landowners to protect unlisted but declining species on their lands.
- *Fisheries Sustainability (Disease, Transgenics, Overfishing, etc.).* The use of transgenic salmonid fish that are reproductively sterile. Aquatic organism risk analysis is proposed to evaluate the environmental safety of transgenic aquatic organisms, but its primary weakness is that the probabilities and consequences are based on subjective evaluations and opinions. There is no requirement for scientific analysis or collection of experimental data in support of the risk analysis process. The role of risk analysis as a basis for fish health management is questioned, given the uncertainties, including unrealistic expectations of the ability to identify the right fisheries management decision, stakeholder participation, inadequate predictive ability of models, lack of integration of management and science, and a tendency to overestimate possible versus likely hazards. There are also considerations of failing to account for the unknown, unexpected, or synergistic impacts.
- *Substances of Concern.* California's Proposition 65 assumes that a listed chemical of concern is released to waterways and that it is "guilty before proven innocent."

  Banning persistent, bioaccumulative, toxic substances. Sweden is considering banning substances that persist in the environment and accumulate in living tissue, regardless of whether they are proven significantly toxic such as lead in leaded glass. Under the current system, regulations are aimed at controlling rather than preventing end of pipe versus upstream pollution.

  Restricting pesticide use. Canada is moving closer to adopting the precautionary principle based on a year 2001 Supreme Court ruling that upholds the right of local governments to restrict pesticide use to essential uses to protect human health and the environment. Northern Ireland is banning many pesticides used for food production amid rising consumer concerns about impacts on human health and the environment due to chemical residues. This, of course, involves a ban on any pesticide where there is doubt about its safety, even if the weight of scientific evidence is insufficient or inconclusive.

  Potential bans on antibiotics, given the uncertainties and risk of microbial resistance in the environment.
- *Ocean Shipments or Dumping.* Ocean shipments of radioactive materials and the reprocessing of nuclear wastes and acquiring stockpiles of plutonium. Proponents may want to cite explicitly the precautionary principle as a framework to regulate such shipments. The principle might require users of the ocean to exercise precaution, alertness, and effort by undertaking relevant research, developing nonpolluting technologies, and avoiding activities that present uncertain risks to the marine ecosystem—and further it would reject the current notion of a measurable ability of the environment to assimilate wastes.

  Proposals to dump sewage sludge on the seemingly barren areas of the deep seabed. There is little known about the ecology of this region, the harm to the seabed and interconnected ecosystems. A traditional tort approach would allow dumping until harm is recognized, but a precautionary regime would require the dumping state to prove that the dumping is harmless before engaging in that activity.
- *Water Reuse, Reclamation, and Safety.* A precautionary approach might stipulate that climate change and increased urbanization may threaten to dry up wetlands, pushing species of those habitats to extinction. The fix may be to increase the efficiency of water use and stop wasteful irrigation practices where water losses are highest. The principle might also stipulate that all types of water abstraction and water transfers should be controlled. Certainly economic instruments and regulation can provide a flexible means of improving water resource issues.

  The events of 9/11 may warrant conducting a security sweep of the nation's drinking water systems or for facilities to develop specific security plans based on vulnerability assessments.
- *Sewage Sludge.* In the United States, the biosolids rule imposes limits on molybdenum and alkylphenol. Some critics claim that the risk assessment process is inadequate and does not address all potential or foreseeable risks. Environmentalists, farmers, and scientists have encouraged the U.S. EPA to adopt a more protective program that uses the precautionary principle to determine the safety of applying water to land.
- *Habitat Conservation.* Involves dealing with data deficiencies in species status and magnitude of impacts in a manner that guards target species from irreversible habitat loss yet does not preclude development.
- *Biotechnology.* Consider genetically modified organisms or foods. The benefits of genetically modified

plants include reduced use of fertilizer and pesticides, increased yields, better adaptation to extreme environmental conditions, and reduced allergenicity versus the cost of impacts that may include allergenicity and exotic species that outcompete native and nonmodified organisms.

- *Antidegradation.* Antidegradation and antibacksliding rules are likely to become increasingly important in water policy because the intention of the Clean Water Act is to prevent perceptible downgrading of designated water uses. The protection of high-quality waters under the antidegradation rule might be interpreted differently by risk analysis versus the precautionary principle: For example, the precautionary principle might interpret the rule as an absolute prohibition on lowering water quality. The risk-based approach would consider that avenue a no-growth interpretation and propose a structure that provides systematic evaluation of all available information regarding the social, environmental, and economic impacts of lowering water quality.
- *Sustainable Development.* Some would argue that to achieve sustainable development, policies must be based on the precautionary principle. Environmental measures must anticipate, prevent, and attack the causes of environmental degradation. Science alone cannot determine the appropriate focus of the management inquiry or how to define a just or politically acceptable solution. Setting standards is the confluence of science and policy determination.

Consider the example of applying the precautionary approach to the risks of contaminants in surface and groundwater, including biological, chemical, and radiological agents. Certainly, on a global scale, pathogenic contamination of drinking water can be considered a significant health risk due to disease outbreaks, but there are also exposures to nonpathogenic toxicants. Water pollution comes from many routes, sources, and in various forms; in each case, there are incomplete or lacking data, certainly in the exposure, probability, and effects stages.

In contrast to the precautionary principle, risk-based objectives in water science are typically technology-based (e.g., best available technology, or best available technology economically achievable). One clear advantage of an approach based on technology includes national consistence in treatment level, predictable economics, and relative ease of enforcement—although it does not guarantee improved environmental performance because some areas may be under or overmanaged.

For some environmental practitioners, the precautionary principle appears to focus only on the risk from the agent or technology versus the clear benefits that may result from its use or from the risk trade-offs or transfers that may result from an alternative. Because the risks of a new technology are often not apparent until the technology is in use, some suggest it is then too late to undo the harm (e.g., various case studies in the Great Lakes have suggested to some investigators that the length of time between the introduction of a new technology and the discovery of its deleterious effects and the regulatory action to reduce exposure is typically greater than 25 years). Certainly, unpredictable consequences are not new to water science, considering the unanticipated challenges caused by the use of asbestos and PCBs. The concepts of risks and benefits in water science are not as clearly etched as some may think.

RECOMMENDATIONS FOR BALANCING PROOF VERSUS PRECAUTION

Proponents should proceed cautiously with the precautionary principle. Potential problems in abandoning the risk assessment framework in favor of following the precautionary principle are highlighted in Table 2:

There are certainly some refractory questions to consider, such as:

1. What grounds justify applying the precautionary principle, and what will be the trade-offs? Stifle new technology and associated benefits (e.g., genomic understanding of microbes, plants, animals) from which there is little experience and lots of uncertainty; increase the costs of developing and marketing; threaten freedom and trust of basic research.
2. Who will trigger use of the precautionary principle, and how will they interpret the volume and breadth of data generated from such areas as genomics and proteomics in a positive manner? Risk related data are increasingly in the form of biological markers, gene expression, and protein level change, for which there is an abundance of data generated; yet these are possibly laden with uncertainty and lack of validation. How will this data be processed by the precautionary principle in characterizing new products and assessing efficacy, safety, and so on?
3. What is "sustainable" about the precautionary principle? If we apply the recognized working definition that sustainability is "a challenge that economic growth and development must take place and be maintained over time within the limits set by the physical, chemical, biological, social, and political environments," it follows that environmental protection, product stewardship, and economic development are complementary rather than antagonistic processes, as the precautionary principle would suggest.
4. What is the role of scientific evidence, and to what extent can the precautionary regulatory approach be thought of as without factual or scientific justification, arbitrary and irrational, defensible or enforceable in court, and vulnerable to scientific, legal, and constitutional challenges?
5. To what extent can the precautionary regulatory approach for setting risk standards be flexible enough to consider site-specific conditions in which it is applied (i.e., natural and anthropogenic factors, chemical agent, receptor type and size, exposure conditions, bioavailability, and cost)?

In the end, the relative strength and quality of the precautionary policy will depend on legitimacy, degree of

Table 2. Issues and Risks

Issue	Recommendation
• Process and structure not costless and likely has high expenditures associated given uncertainties	• Utilize cost–benefit and regulatory flexibility analyses • Rational decisions should account for the impact of benefit reduction
• Chemical use does not equal chemical risk	• Account for the way chemical is used, materials handling, exposure sources, potency to receptors • Adopt guidelines to ensure chemical risk and substitution are balanced and not harmful
• Potential risk transfers	• Do not trade a known manageable risk for an uncertain alternative • Actions to remove hazard may remove benefits; loss of benefits introduces new risks • Brainstorm unforeseen risk potentials, such as danger of responding to wrong threat • Utilize a "total hazard" approach
• Undefined and arbitrary methods or interpretation	• Rigorous, prioritized; avoid overregulation that may lead to loss of potential benefits
• Technological stagnation	• Cessation or phaseout of regulatory prescription threatens to kill the patient while inventing a cure for a disease that does not necessarily exist; it may lead to arbitrary unscientific rejection of beneficial technologies
• Lack of prioritization	• Final assessment should involve prioritization of substances and relative risk; should consider impacts of cessation or phaseout on international legislation, production and use, economics, and alternatives
• No zero risk in a technology-based society	• Neither risk assessment nor the precautionary principle can eliminate all risks; these paradigms can only be refined to reduce the relevant risks to a perceived acceptable level; proving no risk is a difficult if not impossible task
• Risks of no action	• The precautionary principle must include a full range of options, including no action
• Tools for precautionary thinking	• Life cycle analysis, waste-free product goals, cleaner production, green building, environmental justice, ecoefficiency, benchmarking, best practices
• Tolerability and meaningful restraints	• Need to define and quantify risk acceptability criteria • The margins of safety for descriptions of low risk must be defined and should consider factors of uncertainty, cost–benefits, and lack of knowledge • Goal is to introduce reasonable/plausible conservatism
• Politicization	• Balance technical, scientific, economic, and political
• Codes of ethics	• Practitioners should adhere to a code of ethics that re-iterates the need for validity of data and sound scientific approaches that are transparent, systematic, subject to peer review, accountable, and independent

public trust, and actual quality of products and decisions. Whether by traditional risk analysis or by precautionary thinking, risk management should be determined case by case and population by population, with a full and transparent accounting of the bigger picture, the consequences of different actions including no action, and the public values surrounding those actions; even if the details are obscure.

READING LIST

Burke, M. (2000). Promoting a greener European Union. *Environ. Sci. Technol.* **17**: 388A–393A.

Chapman, P.M. (1999). Does the precautionary principle have a role in ecological risk assessment? *Hum. Ecol. Risk Assessment* **5**: 885–888.

Charnley, G. and Donald Elliott, E. (2002). Risk versus precaution: environmental law and public health protection. *The Environ. Law Report* **32**: 10363–10366.

European Commission. (2000). *Communication from the Commission on the Precautionary Principle*. COM 02.02.2000, Brussels.

Fairbrother, A. and Bennett, R.S. (1999). Ecological risk assessment and the precautionary principle. *Hum. Ecol. Risk Assessment* **5**: 943–949.

Foster, K.R., Vecchia, P., and Repacholi, M.H. (2000). Science and the precautionary principle. *Science* **288**: 979–981.

Gilbertson, M. (2002). The precautionary principle and early warnings of chemical contamination of the Great Lakes. In: *EU Communication "Late Lessons from Early Warnings: The Precautionary Principle 1896–2000*.

Graham, J.D. (1999). Making sense of the precautionary principle. *Risk Perspect.* **7**: 1–6.

Graham, J.D. (2000). Decision-analytic refinements of the precautionary principle. *J. Risk Res.* in press.

Hansson, S.O. (1999). Adjusting scientific practices to the precautionary principle. *Hum. Ecol. Risk Assessment* **5**: 909–921.

Mckinney, W.J. and Hammer Hill, H. (2000). Of sustainability and precaution: The logical, epistemological, and moral problems of the precautionary principle and their implications for sustainable development. *Ethics Environ.* **5**: 77–87.

Pittinger, C.A. and Bishop, W.E. (1999). Unraveling the chimera: a corporate view of the precautionary principle. *Hum. Ecol. Risk Assessment* **5**: 951–962.

Raffensperger, C. and Tickner, J. (Eds.). (1999). *Protecting Public Health and the Environment: Implementing the Precautionary Principle*, Island Press.

Raffensperger, C. and deFur, P.L. (1999). Implementing the precautionary principle: rigorous science and solid ethics. *Hum. Ecol. Risk Assessment* **5**: 933–941.

Rasmussen, F.B. (1999). Precautionary principle and/or risk assessment. *Environ. Sci. Pollut. Res.* **6**: 188–192.

Rogers, M.D. (2001). Scientific and technological uncertainty, the precautionary principle, scenarios and risk management. *J. Risk Res.* **4**: 1–15.

Rogers, M.F., Sinden, J.A., and De Lacy, T. (1997). The precautionary principle for environmental management: a defensive-expenditure application. *J. Environ. Manage.* **51**: 343–360.

Sandin, P. (1999). Dimensions of the precautionary principle. *Hum. Ecol. Risk Assessment* **5**: 889–907.

Santillo, D. and Johnston, P. (1999). Is there a role for risk assessment within precautionary legislation? *Hum. Ecol. Risk Assessment* **5**: 923–932.

Santillo, D., Stringer, R.L., Johnston, P.A., and Tickner, J. (1998). The precautionary principle: protecting against failures of scientific method and risk assessment. *Mar. Pollut. Bull.* **36**: 939–950.

Shipworth, D. and Kenley, R. (1999). Fitness landscapes and the precautionary principle: the geometry of environmental risk. *Environ. Manage.* **24**: 121–131.

Stijkel, A. and Reijnders, L. (1995). Implementation of the precautionary principle in standards for the workplace. *Occup. Environ. Med.* **52**: 304–312.

Tucker, A. (2002). Life-cycle assessment and the precautionary principle. *Environ. Sci. Technol.* **36**(3): 70A–75A.

Underwood, A.J. (1997). Environmental decision-making and the precautionary principle: what does this principle mean in environmental sampling practice? *Landscape Urban Plann.* **37**: 137146.

Wiener, J.B. and Rogers, M.D. (2001). *Comparing Precaution in the United States and Europe*. Working Paper August 2001–01.

WATER PRICING

Steven J. Renzetti
Brock University
St. Catharines, Ontario, Canada

INTRODUCTION

The price of water is measured by the value of the economic resources that must be sacrificed to acquire potable water. In many cases, this price is easily understood as the number of dollars and cents that must be paid to obtain a specific volume of water. In other cases, often in low-income countries, the price of water is best understood as the combination of out-of-pocket expenses plus the time required for travel to and from a water source.

The purpose of this article is twofold. First, it describes the nature and range of water prices around the world. Second, it explains the significance of the price of water for decisions regarding the supply of water and the demand for water. The article concludes that, despite the potentially important role of water prices, a lack of attention to water pricing has been a major contributing factor to worldwide problems of water allocation and water quality.

WATER PRICES AROUND THE WORLD

The level and form of water prices vary widely around the world. Figure 1 gives an indication of the range of residential water prices across a number of countries. The form or structure of water prices also varies across countries and across sectors (agricultural, residential, commercial, and industrial). On the one hand, there are many examples where users face a price of zero for their water. This occurs in municipalities where household water use is not metered and in rural areas where households rely on, but do not pay for, groundwater. There is also another way in which the price of using water can be said to be zero in many cases. If a factory or sewage treatment plant discharges wastes into a waterbody without paying for that right, then these facilities are using water free. It is important to remember, however, that a zero price for water does not mean that there is no cost for using it. This cost may come in the form of the infrastructure needed to purify and transport the water to users, in the form of reduced water quality, or it may come in the form of foregone benefits that would have been enjoyed by other water users had they access to another's water.

On the other hand, water prices can be quite complex and have several components: a one-time connection charge, a recurring fixed charge, a volumetric charge, and supplemental charges for sewage treatment and environmental protection. Further, the volumetric component of the price of water can, itself, vary in structure. It can be a constant amount (such as $0.76 U.S. per cubic meter in Toronto, Canada in 2002) or it can be variable in the sense that the price charged on the last unit of consumption either rises or falls with consumption (these are termed increasing and decreasing block rate structures, respectively). The City of Los Angeles, for example, has a very complex rate structure for residential consumers; the volumetric component varies by season, the household's location, and the quantity of water consumed (2).

Table 1 provides a snapshot of water prices in different sectors and countries. There are a number of noteworthy features of the information presented. First, the information is somewhat fragmentary because information for a number of sectors is not included. Most national governments do not collect data on water prices in their countries. Second, in many sectors and countries, the price of water is very low. The implications of this observation for the viability of water supply systems and for the level of water use are discussed in the next

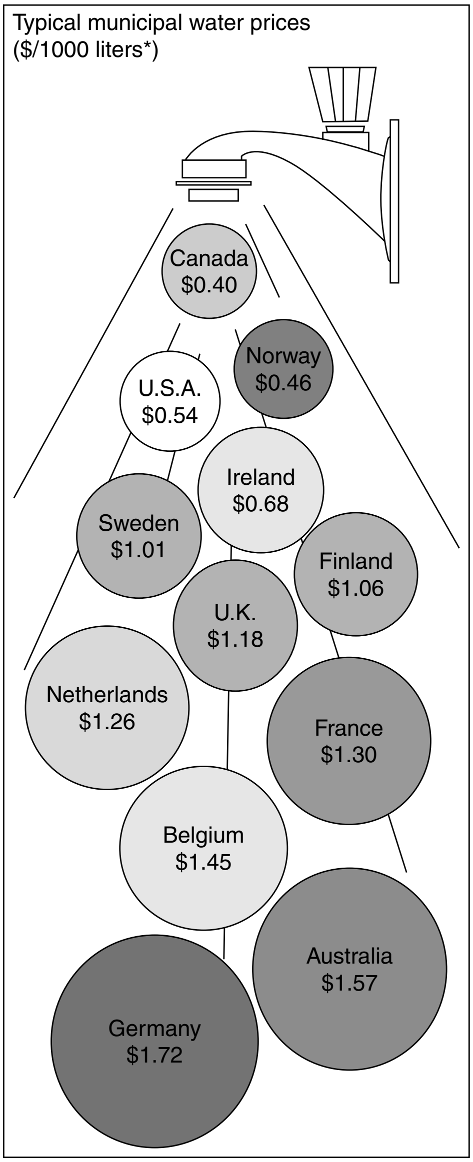

*All amounts are in 1992 Canadian dollars. In most countries, these prices are increasing from year to year. These figures do not include the cost of waste treatment.

Source: *National Utility Service International Water Price Surveys* as published in World Water and Environmental Engineer.

Figure 1. International comparison of water prices. (*Source*: Reference 1. Note that 1 Canadian dollar in 1992 is the approximate equivalent of $0.75 U.S. in 2001).

section. As Dinar and Subramanian (3) point out, within any sector, prices may vary by the quantity consumed, by the type of water use (e.g., what type of crop is grown), by the time of consumption, and by location. Prices also vary significantly across sectors; industry typically faces the highest prices. Interestingly, prices appear to differ less across countries than they do across sectors within a given country. This may be due to some sectors' political power that has allowed them to avoid facing the full costs of their water use. Finally, there does not appear to be a strong connection between a country's water prices and the scarcity of water in that country. Canada's low water prices might be explained, in part, by its abundant water supplies, but low water prices are also observed in water-scarce countries such as Australia and Spain.

THE FUNCTIONS OF WATER PRICES

In principle, the price of water can play several important roles. From the viewpoint of society as a whole, the price of water informs users and suppliers of the value and scarcity of water. From the viewpoint of the agency supplying water, water prices generate the revenues needed to ensure the financial viability of the agency. From the viewpoint of an individual household, farm, or business, water prices inform decision-makers what they must give up in exchange for getting water. Thus, it can be seen that water prices play many roles and that they have the potential for influencing the demand for water, the supply of water, and the allocation of water among competing users (4,5).

To see how the price of water plays these multiple roles, consider the impact of rising water prices (for

Table 1. Water Prices[a] for Various Sectors and Countries (1996 U.S.$/m^3)[b]

Country	Agriculture[c]	Domestic[c]	Industry[c]
Algeria	0.019–0.22	0.057–2.7	4.64
Australia	0.0195	0.23–0.54	7.82
Botswana	NA	0.28–1.48	NA
Brazil	0.0042–0.032	0.4	NA
Canada	0.0017–0.0019	0.34–1.36	0.17–1.52
France	0.11–0.39	0.36–2.58	0.36–2.16
India	NA	0.0095–0.082	0.136–0.290
Israel	0.16–0.26	0.36	0.26
Italy	NA	0.14–0.82	NA
Madagascar	NA	0.325–1.75	NA
Namibia	0.0038–0.028	0.22–0.45	NA
New Zealand	NA	0.31–0.69	NA
Pakistan	NA	0.06–0.10	0.38–0.97
Portugal	0.0095–0.0193	0.153–0.529	1.19
Spain	0.0001–0.028	0.0004–0.0046	0.0004–0.0046
Sudan	NA	0.08–0.10	0.08–0.10
Tanzania	0.260–0.398	0.062–0.241	0.261–0.398
Tunisia	0.02–0.078	0.096–0.529	0.583
Uganda	NA	0.38–0.59	0.72–1.35
United Kingdom	NA	0.0095–0.0248	NA
United States	0.0124–0.438	NA	NA

[a] Prices refer only to the volumetric component. They do not include any connection fees or other charges not related to the volume of water consumed.

[b] A cubic meter is approximately 220 Imperial gallons or 264 U.S. gallons.

[c] "NA" means that no number was reported in the original source.

Source: Adapted from Table 1.2 in Reference 3.

example, that might be brought on by the increased energy needed to pump water from shrinking aquifers). Water users see increased prices as a signal of the increased costs that result from using an increasingly scarce (and, thus, valuable) resource. Users can then be expected to reconsider their planned water uses with an eye to conservation and/or substitution toward less expensive alternatives. These types of responses can include changing landscaping practices, retrofitting plumbing fixtures, introducing water recycling, and investing in more efficient water-using equipment. It is important to notice that changing water prices, then, can have an influence on water-use practices and also on the use of other goods and services.

From the perspective of water suppliers, rising prices are a signal that alternative sources of supply are now more attractive options than previously thought. Securing additional water can take the form of reducing system losses, investing in new technologies (perhaps more efficient pumps or water meters), extending the water collection network, or purchasing or leasing water rights.

Rising water prices also provide valuable signals to other decision-makers in society. Rising water prices (and their signal of growing water scarcity) indicate to government that there is an increasing possibility of heightened conflicts among competing water users. As a result, that there may be a need to review the adequacy of water rights laws and water allocation regulations. In addition, rising water prices are a signal to those firms who design and build equipment that delivers or uses water. These firms will see that there is a stronger incentive to innovate and design improved forms of water-related technologies and equipment.

ASSESSING WATER PRICES

The preceding section indicates that water prices, in principle, play a variety of roles, including raising revenue for water utilities and signaling relative scarcity to both suppliers and demanders. Water prices are best able to play these roles when they are free to rise and fall according to changes in supply and demand. Furthermore, the signal they send will be most informative when prices accurately reflect the full costs to society of water use (including any environmental damage). This section briefly examines how well prices perform these roles.

Many researchers and analysts (6,7) are critical of the history and current state of water pricing and argue that, in most cases throughout the world, water prices fail either to reflect accurately the full social costs arising from water use or to fulfill their other functions. Renzetti (8) demonstrates that municipal water prices in Ontario, Canada, often understate the incremental cost of supply by a factor of 2. Munasinghe (9) examines water pricing in Manila and finds that it underestimates the cost of supply by failing to reflect declining groundwater levels. Underpricing water is hardly limited to municipal water supply systems. Research indicates that low prices for irrigation water have historically provided significant subsidies to the agricultural sector. Roodman (7) asserts that "The U.S. government spent an estimated $45–93 billion more than it earned on public irrigation projects between 1902 and 1986 (in 1995 dollars)" (p. 37). These researchers believe that inadequate water prices are a major factor in many of the problems relating to water management (6,10). These problems include overuse of water, poor system reliability, reduced water quality, and the diversion of productive capital to subsidized water projects.

It is sometimes thought that low water prices benefit low-income farmers and households. Unfortunately, this is unlikely to be the case. First, low water prices are an inefficient way to assist the poor because water generally makes up a relatively small share of total household expenditures. In any case, a better option is to design water prices so that families who use relatively little water can be given a free allowance of a specific quantity of water. Second, low water prices often translate into water agencies that do not have adequate financial resources to maintain water quality and system reliability. Middle and high-income households may be able to shield themselves by finding alternative sources of water supply, but the poor can rarely do this. Thus, the general principle is to set prices so that they may reflect, as fully as is politically and administratively feasible, the forces of supply and demand.

CONCLUSIONS

This article has examined the nature of water prices and has demonstrated that the price of water can play an important role in many decisions related to the supply and demand of water and to the allocation of other scarce resources such as investment capital. Despite this potential, water prices have not played an important part in the history of water management. For example, many people remember the building of the Grand Coulee and High Aswan dams, but how many know the prices charged for water in these projects? This lack of attention to water pricing has had important consequences. These include prices that are far below the cost of supply, overuse of water, reduced water quality, and the delayed development of water-conserving technologies. Perhaps because of these conditions, a growing amount of attention is being directed toward water pricing.

BIBLIOGRAPHY

1. Environment Canada (2002). Freshwater website http://www.ec.gc.ca/water/e_main.html. Accessed February 27, 2002.
2. City of Los Angeles Department of Water Services (2002). Water Rates For Service In The City And County Of Los Angeles http://www.ladwp.com/water/rates/sch_b.htm. Accessed February 27, 2002.
3. Dinar, A. and Subramanian, A. (Eds.). (1997). *Water Pricing Experiences: An International Perspective*. World Bank Technical Paper 386, The World Bank, Washington, DC.
4. Hall, D. (Ed.). (1996). *Marginal Cost Rate Design and Wholesale Water Markets*, *Advances in the Economics of Environmental Resources*. Vol. 1. JAI Press, Greenwich, CT.
5. Renzetti, S. (2002). *The Economics of Water Demands*. Kluwer Academic, Boston.
6. Postel, S. (1996). Forging a sustainable water strategy. In: *State of the World 1996*. L. Brown (Ed.). Worldwatch Institute, Washington, DC, pp. 40–59.
7. Roodman, D.M. (1996). *Paying the Piper: Subsidies, Politics and the Environment*. Research paper 133, WorldWatch Institute, Washington, DC.
8. Renzetti, S. (1999). Municipal water supply and sewage treatment: costs, prices and distortions. *Can J. Econ.* **32**(2): 688–704.
9. Munasinghe, M. (1992). *Water Supply and Environmental Management: Developing World Applications*. Studies in Water Policy and Management. Westview Press, Boulder, CO.
10. Easter, K.W., Feder, G., Le Moigne, G., and Duda, A. (1993). *Water Resources Management: A World Bank Policy Paper*. The World Bank, Washington, DC.

SPOT PRICES, OPTION PRICES, AND WATER MARKETS

Terence R. Lee
Santiago, Chile

Markets in water rights do not operate only for buying and selling permanent water rights. Using a market makes available a variety of other tools for increasing the efficiency of water use. Among the most interesting of these tools is the use of spot markets and water use options or future markets. Both alternatives lead to a broadening of the market and to an increase in the number of participants on both the buyer and the seller sides of the market. Their use also results in more stable demand and supply conditions as the market has a larger number of buyers and sellers. They also permit water users to hedge against changes in the price and availability of water rights and thus eliminate the need to hoard as a hedge against this risk. On the other hand, the use of spot and option or future markets requires developing storage facilities and interconnected conveyance systems if their full benefits are to be obtained.

Spot markets involve the leasing, not the sale of water rights. The sale of rights is replaced by the sale of water. In a spot market, the water right remains with the owner. The temporary sale of water in a spot market is a preferred market response to short-term changes in demand and supply.

Water users may find it advantageous to engage to engage in a spot market for a variety of reasons. The seller has an opportunity to earn revenue in the temporary trade of surplus water but does not give up any water rights. This is particularly useful for accommodating a short-term demand for additional water for any use that has a predictable and fixed life span or for a use of uncertain duration. For example, a farmer facing an unexpected variation in normal supplies may buy on a spot market when it is not economical to transport water in periods of sufficient water supply. There are, however, examples of the longer term use of spot markets when a water user is unable or unwilling to commit the resources necessary to buy a water right.

One example of a permanent spot market, where farmers always pay spot prices for water, is found in the community of Huerta in Alicante, Spain. Here the ownership of water is separate from the ownership of land. Water is distributed by rotation at a fixed rate, approximately the same quantity of water in each successive rotation, and the proportion of water available to any water right holder varies for each rotation, depending on the water rights acquired. Before each rotation, a notice is posted that announces the date on which the rotation will commence and informs water rights holders that, within a prescribed period, they should claim their "*albalaes*" or tickets for this rotation. Once allocated, tickets, available in twelve denominations for a constant supply of water from 1 hour to 1/3 minute, are freely tradable in a public auction and an informal market. The community makes a genuine effort to provide information to farmers so they can buy and sell water intelligently; there are brokers who facilitate trading. A simulation model comparison of this system with those found elsewhere in Spain, where trading is not permitted, indicated that the spot market approach adopted in Alicante is the most efficient in terms of net increases in regional income. The differences are not great for moderate water shortages but are significant in conditions of severe water shortage (1).

In a typical spot market contract, the buyer pays the owner of the water right, generally in periodic installments, but there can also be an up-front payment to initiate the sale. Such contracts are often renewable. The length of a contract in irrigation districts is typically a single season. Contracts can be, however, very short, sometimes for as little as a few hours. In the short run, spot prices usually provide a source of water cheaper than the purchase of a permanent water right, but water on a spot market can and will fluctuate in price. The supply is secure only for the period of the contract. There is also an expense in the constant renewal costs for those who depend on short-term purchases in a spot market.

By providing immediate access to water, spot markets can, however, accommodate the most varying needs. The flexibility of spot markets makes them an attractive option for many users, and they are often very active, particularly among neighboring water rights holders in irrigation districts, even where markets in permanent rights have not been established. Spot purchases, often informal, are usually the predominant form of market transactions, whether or not a formal market exists.

In Chile, seasonal spot purchases of water have been a much more active form of water reallocation than water rights sales (2). The most common transaction is an arrangement between neighboring farmers whose water requirements differ through the cropping cycle (3,4). Spot trading between neighbors is easier because it does not require investment in storage facilities or elaborate systems of interconnection between canal systems. In California, as well, water marketing is characterized by an emphasis on seasonal spot markets, although there are good facilities for transferring water over very large distances (5).

Option prices are based on longer term agreements to lease, or sometimes, but less commonly, to sell a water right. The option is taken up when a given event occurs, typically a drought. There are examples of the use of options in both the United States and in Chile. A typical arrangement in Chile is the payment by a fruit farmer of a prenegotiated fee to a farmer growing annual crops for an option on water supply in the case of drought (6).

Options can be and are also used to transfer water from agricultural to nonagricultural users; this has been common practice in the western United States. Dry year option contracts are an attractive alternative when supplies are normally adequate. Options provide supplies during droughts at a cost lower than purchases or leases of water rights.

Options are particularly attractive because the buyer secures long-term water supplies and the seller receives compensation for the option, including the income lost when the option is exercised. For the buyer of an option, the contract also provides a means of obtaining additional water supply, under predetermined conditions, and at a specified price.

Option contracts are, however, complex. The contract must address the risk to the buyer that the water right will not be available when the time comes to exercise the option. The establishment of an option contract limits the rights of the owner of the water right to sell. This issue can be addressed by including in the contract the "right of first refusal" which allows the seller to retain the option of selling the right, but gives the option holder the right to match the offered price (7). Moreover, to produce the greatest benefits, options require long-term contractual commitments, up to 20 or more years, which can introduce many uncertainties into the arrangement.

Option contracts can include payments adjusted over time to allow for changes in water use, production costs, technology, and other conditions. Alternative methods of charging can be used; a lump sum, annual payments, or a combination of annual payments with a lump sum when the option is exercised. The latter is a particularly attractive alternative, because neither party needs to anticipate fully the number of times the option will be exercised over the contract period (8). The holders of water rights on which option contracts have been agreed to can also be compensated in kind, for example, by lower rates for the option holder's production; as in irrigation to hydroelectricity generation transfers, farmers can be compensated by lower electricity rates.

BIBLIOGRAPHY

1. Maass, A.R. and Anderson, R.L. (1978). *. . . and the Desert Shall Rejoice: Conflict, Growth, and Justice in Arid Environments*. MIT Press, Cambridge, MA.
2. Chile, *Trabajo de Asesoría Económica al Congreso Nacional, 'Nuevo proyecto de modificación al Código de Aguas'*, (1996). Trabajo de Asesoría Económica al Congreso Nacion-0al, No. 74, December, Programa de Postgrado en Economía, ILADES/Georgetown University.
3. Gazmuri, R. (1994). Chile's market-oriented water policy: institutional aspects and achievements. In: *Water Policy and Water Markets*. Selected papers and proceedings from the World Bank's Ninth Annual Irrigation and Drainage Seminar, Annapolis, Maryland, December 8–10, 1992. G. Le Moigne, K.W. Easter, W.J. Ochs, and S. Giltner (Eds.). Technical Paper Number 249, The World Bank, Washington, DC.
4. Gazmuri, R. and Rosegrant, M.W. (1994). *Chilean Water Policy: The Role of Water Rights, Institutions, and Markets*.

Tradable water rights: experiences in reforming water allocation policy by M.W. Rosegrant and R. Gazmuri, Irrigation Support Project for Asia and the Near East.

5. Howitt, R.E. (1997). *Initiating Option and Spot Price Water Markets: Some Examples from California*, Seminar on Economic Instruments for Integrated Water Resources Management: Privatization, Water Markets and Tradable Water Rights. Proceeding, Inter-American Development Bank, Washington, DC.
6. Thobani, M. (1997). *Formal Water Markets: Why, When and How to Introduce Tradable Rights in Developing Countries*, Seminar on Economic Instruments for Integrated Water Resources Management: Privatization, Water Markets and Tradable Water Rights. Proceedings, Inter-American Development Bank, Washington, DC.
7. Michelsen, A.M. and Young, R.A. (1993). Optioning agricultural water rights for urban water supplies during droughts. *Am. J. Agric. Econ.* **73**: 4.
8. Hamilton, J.R., Whittlesey, N.K., and Halverson, P. (1989). Interruptible water markets in the Pacific Northwest. *Am. J. Agric. Econ.* **71**: 1.

WATER MANAGED IN THE PUBLIC TRUST

Laurel Phoenix
Green Bay, Wisconsin

Managing water in the public trust requires considering the water demand of current and future citizens as well as that of ecosystems, fish, and wildlife. This presumes that water will be managed sustainably for human and ecosystem benefit, despite the fact that ecosystems have no voice in the political process. To understand how complicated it is to accomplish this, the definition and history of the public trust is examined. Next, water management is briefly described to contrast current water management methods with the finesse that managing water in the public trust would require. Finally, some of the difficulties of achieving this in a world of rising demand and finite resources are highlighted, followed by a set of principles describing government obligations to uphold human rights.

WHAT IS THE PUBLIC TRUST?

The public trust is an enduring doctrine obligating government to manage natural resources in the best interests of its citizens. Because some vital natural resources like oceans, air, and wildlife cannot be owned in the same way a person can hold a property interest in land, it is the responsibility of government to hold ownership over these common resources in trust for its citizens, so that their use may be regulated to ensure universal access and prevent their overuse or degradation.

HISTORY OF THE PUBLIC TRUST

The common law doctrine in European history is first attributed to the Roman Emperor Justinian, who declared in 533–534 B.C.E. that running water, seas, and wildlife were part of a "negative community" of things that could not be owned but could be used under regulation to ensure vital common resources were not exploited (1). A later example is the Magna Carta signed in 1215, granting guaranteed liberties to the people that not even King John or future kinds could overrule. Because of its importance to fisheries, commerce, and navigation, the public trust doctrine in America was initially used at federal and state levels only to apply to waterways, but court rulings since then have expanded the public trust to include many other environmental issues, such as protection of ecosystems, recreation like swimming and boating, preservation of views (scenic values), habitat (wildlife values), open space, and beaches (2). Some states have taken a step further to apply the public trust to nonenvironmental issues, such as historical sites; civic sites or economically important areas, such as aging downtown districts; or economic development such as tearing down old, low value single homes to be replaced with many more valuable units, such as expensive condos.

Consequently, the public trust has evolved from rules for equitably sharing the natural "commons" of the land to case law expanding the multiple uses of water resources, as well as rules looking at any social or economic good that could be said to advance the public welfare.

HOW IS WATER "MANAGED"?

Water resources management has been defined as "the application of structural and nonstructural measures to control natural and man-made water resources systems for beneficial human and environmental purposes" (3). Structural facilities include wells, dams, diversion structures, pipes, pumps, treatment plants, and hydroelectric plants. Nonstructural measures are economic incentives, regulations, public education, and planning. Numerous individuals, agencies, and jurisdictions are involved in these structural and nonstructural components. Beyond this complexity, water managers must concomitantly address needs for navigation, municipal water supply, irrigation, fishing, industrial use, hydropower, recreation, public access, and traditionally sacred sites, among others. Water must also be managed to protect ecosystems, fish, and wildlife and should not be unsustainably used to the detriment of future generations.

In America, water resources management is a fragmented collection of federal, state, and local laws managed by numerous and sometimes competing agencies, water districts, ditch companies, acequias, and so on. There are a haphazard mix of treatment and pollution laws at the federal level, water allocation laws at the state level, and sporadic conservation laws at some local levels when supplies run low. It is complicated further by outdated international treaties and interstate compacts. Some states permit, measure, and track individual water use, and others do not.

Two reasons for this lack of a more organized and efficient water management system are (1) a long history, until recently, of having enough water to meet the needs of most users and (2) a variety of political influences on water management decisions. First, competing human uses will have great influence in politics, whereas ecosystems or future generations have few advocates. Second, of the current competing uses, some interest groups have more political power than others, so water may be inequitably or inefficiently managed, allocated, and used. Third, of the current competing uses, some are mutually exclusive in that they cannot happen at the same time in the same location, causing more conflict and subsequent pressure on water managers. For example, building a dam on a prime salmon run damages their ability to spawn, reducing their populations by 80–90%. Another incompatible use of a stream or lake is jet skiing and fishing. Fourth, political influences can cause the best science to be ignored, so, for example, irrigators will get the water they want even after biologists predict large fish kills will result (e.g., Klamath River fish kills of 2002–2003). Fifth, politicians and other decision-makers will be pressured to make decisions with short-term benefits, ignoring long-term solutions. Finally, land use decisions, usually made at a local level, rarely are made with the goal of preserving/protecting water resources. Thus, there are small but cumulative negative effects on surface and ground water everywhere.

MANAGING WATER IN THE PUBLIC TRUST

Managing water in the public interest, therefore, would be a complex and deliberate allocation of water resources to meet the requirements of as many uses as possible, consistent with protecting the resources from pollution and allocating some water for users with no political voice (e.g., ecosystems and future generations)—a Sisyphean task. (Punished for his misdeeds, the Greek Sisyphus was condemned by the gods to rolling a large boulder up a hill for eternity, only to have it roll back down again and force him to start over.) This is complicated enough, but how is this to be done in a world with a fixed amount of water but rapidly growing and often impoverished populations? Already there is evidence of a current global crisis of not enough water to meet human demand. If over 1.1 billion people have no access to safe drinking water today (4), and water conflicts have already been mapped over a large part of the world (5), how can water be wisely managed to provide for the larger populations predicted for the future? There are 6.4 billion people on the Earth now, and by 2050 there will be 9.3 billion (6). Can technological and regulatory improvements produce enough water for these future generations? Besides the inevitable greater demand from additional people, water managers will also have to prepare for climate changes that could alter the historic precipitation and evaporation rates of different regions. Set with such a task, where do we start?

GOVERNMENTAL OBLIGATIONS TO MANAGE WATER IN THE PUBLIC INTEREST

The United Nations has set some guidelines for how governments can fulfill their obligations to their people to support their human rights. These directives of "respect, protect, fulfill" were applied to water to give examples of these governmental obligations (7):

1. The obligation to respect means no interference with anyone's current access to water. Nations could no longer be:

   "engaging in any practice or activity that denies or limits equal access to adequate water;
   artibrarily interfering with traditional arrangements for water allocation;
   unlawfully polluting water;
   limiting access to, or destroying, water services and infrastructure."

2. The obligation to protect means protecting citizens' access to water from third parties, such as warring clans or multinational water companies. Nations would need to:

   "adopt the necessary and effective legislative and other measures to restrain third parties from denying access to adequate water and from polluting and inequitably extracting from water resources;
   prevent third parties from compromising equal, affordable, and physical access to sufficient and safe water where water services are operated or controlled by third parties."

3. The obligation to fulfill means government would need to institute laws and policies to ensure that *every* citizen finally had equal access to water. This could be, for example, "legislative implementation, adoption of a national water strategy and plan of action to realize this right while ensuring that water is affordable and available for everyone."

This framework of respect, protect, and fulfill would use structural and nonstructural measures, but the goals and responsibilities would be more clearly defined than they are now, driving the laws, policies, and choices for governmental expenditures.

FUTURE CHALLENGES

A world where water is managed in the public trust will require major changes. First, it will need governments responsive to and under the control of their citizens, where there is full public participation and transparency in decision-making regarding equitable allocation of water and effective pollution control. Prioritization will then need to be established in advance to deal with a triaged allocation strategy in times of scarcity. Nonpoint and point sources of pollution will need to be far more rigorously monitored to prevent drinking water and habitat degradation. We will no longer be able to pollute with impunity. And finally, inefficient consumption across all human uses, domestic, municipal, industrial, and agricultural will need to be replaced with the most careful consumption, as water becomes increasingly precious.

BIBLIOGRAPHY

1. Getches, D.H. (1997). *Water Law in a Nutshell*, 3rd Edn. West Publishing, St. Paul.
2. Bray, P. (2004). *An Introduction to the Public Trust Doctrine*. Government Law Center, Albany Law School, 80 New Scotland Ave., Albany, NY 12208. Available at http://www.responsiblewildlifemanagement.org/an_introduction_to_public_trust_doctrine.htm.
3. Grigg, N.S. (1996). *Water Resources Management: Principles, Regulations, and Cases*. McGraw-Hill, New York.
4. Intergovernmental Council of the IHP. *IHP contribution to the World Water Assessment Programme (WWAP)*. Document # SC.2004/CONF.203/CLD.32; IHP/IC-XVI/INF.19.
5. Basins at Risk Project. A study by Dr. Aaron Wolf with maps. Available at http://www.transboundarywaters.orst.edu/projects/bar/.
6. The world population highlights (2004). Available at http://www.prb.org/.
7. Scanlon, J., Cassar, A., and Nemes, N. (2004). *Water as a Human Right?* IUCN Environmental Policy and Law Paper # 51. International Union for Conservation of Nature and Natural Resources (IUCN), p. 22.

WATER RECYCLING AND REUSE: THE ENVIRONMENTAL BENEFITS

Environmental Protection
Agency

The Experience at Koele Golf Course, on the Island of Lanai, has used recycled water for irrigation since 1994. The pond shown is recycled water, as is all the water used to irrigate this world-class golf course in the state of Hawaii.

> "Water recycling is a critical element for managing our water resources. Through water conservation and water recycling, we can meet environmental needs and still have sustainable development and a viable economy."
>
> —*Felicia Marcus, Regional Administrator*

WHAT IS WATER RECYCLING?

Recycle: verb 1.a. To recover useful materials from garbage or waste, b. To extract and reuse.

While recycling is a term generally applied to aluminum cans, glass bottles, and newspapers, water can be recycled as well. Water recycling is reusing treated wastewater for beneficial purposes such as agricultural and landscape irrigation, industrial processes, toilet flushing, and replenishing a groundwater basin (referred to as groundwater recharge). Water is sometimes recycled and reused onsite; for example, when an industrial facility recycles water used for cooling processes. A common type of recycled water is water that has been reclaimed from municipal wastewater, or sewage. The term water recycling is generally used synonymously with water reclamation and water reuse.

Through the natural water cycle, the earth has recycled and reused water for millions of years. Water recycling, though, generally refers to projects that use technology to speed up these natural processes. Water recycling is often characterized as "unplanned" or "planned." A common example of unplanned water recycling occurs when cities draw their water supplies from rivers, such as the Colorado River and the Mississippi River, that receive wastewater discharges upstream from those cities. Water from these rivers has been reused, treated, and piped into the water supply a number of times before the last downstream user withdraws the water. Planned projects are those that are developed with the goal of beneficially reusing a recycled water supply.

HOW CAN RECYCLED WATER BENEFIT US?

Recycled water can satisfy most water demands, as long as it is adequately treated to ensure water quality appropriate for the use. Figure 1 shows types of treatment processes and suggested uses at each level of treatment. In uses where there is a greater chance of human exposure to the water, more treatment is required. As for any water source that is not properly treated, health problems could arise from drinking or being exposed to recycled

The Palo Verde Nuclear Generating Station, located near Phoenix, Arizona, uses recycled water for cooling purposes.

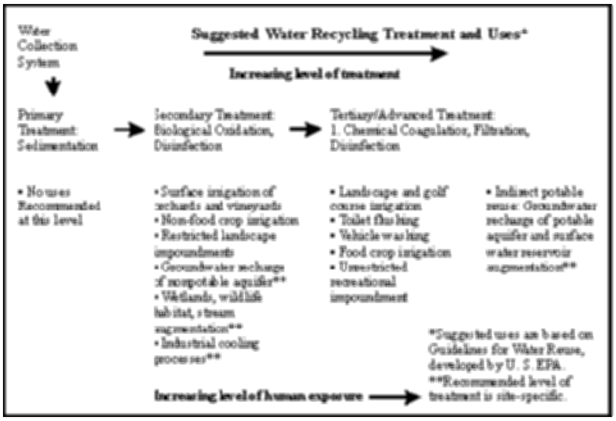

Figure 1. While there are some exceptions, wastewater in the United States is generally required to be treated to the secondary level. Some uses are recommended at this level, but many common uses of recycled water such as landscape irrigation generally require further treatment.

The Irvine Ranch Water District provides recycled water for toilet flushing in high rise buildings in Irvine, California. For new buildings over seven stories, the additional cost of providing a dual system added only 9% to the cost of plumbing.

water if it contains disease-causing organisms or other contaminants.

The U.S. Environmental Protection Agency regulates many aspects of wastewater treatment and drinking water quality, and the majority of states in the United States have established criteria or guidelines for the beneficial use of recycled water. In addition, in 1992, EPA developed a technical document entitled "Guidelines for Water Reuse," which contains such information as a summary of state requirements, and guidelines for the treatment and uses of recycled water. State and federal regulatory oversight has successfully provided a framework to ensure the safety of the many water recycling projects that have been developed in the United States.

Recycled water is most commonly used for nonpotable (not for drinking) purposes, such as agriculture, landscape, public parks, and golf course irrigation. Other nonpotable applications include cooling water for power plants and oil refineries, industrial process water for such facilities as paper mills and carpet dyers, toilet flushing, dust control, construction activities, concrete mixing, and artificial lakes.

Although most water recycling projects have been developed to meet nonpotable water demands, a number of projects use recycled water indirectly for potable purposes. (Indirect potable reuse refers to projects that discharge recycled water to a water body before reuse. Direct potable reuse is the use of recycled water for drinking purposes directly after treatment. While direct potable reuse has been safely used in Namibia (Africa), it is not a generally accepted practice in the United States.) These projects include recharging groundwater aquifers and augmenting surface water reservoirs with recycled water. In groundwater recharge projects, recycled water can be spread or injected into groundwater aquifers to augment groundwater supplies, and to prevent salt water intrusion in coastal areas. For example, since 1976, the Water Factory 21 Direct Injection Project, located in Orange County, California, has been injecting highly treated recycled water into the aquifer to prevent salt water intrusion, while augmenting the potable groundwater supply.

While numerous successful groundwater recharge projects have been operated for many years, planned augmentation of surface water reservoirs has been less common. However, there are some existing projects and others in the planning stages. For example, since 1978, the upper Occoquan Sewage Authority has been discharging recycled water into a stream above Occoquan Reservoir, a potable water supply source for Fairfax County, Virginia. In San Diego, California, the Water Repurification Project is currently being planned to augment a drinking water reservoir with 20,000 acre-feet per year of advanced treated recycled water.

WHAT ARE THE ENVIRONMENTAL BENEFITS OF WATER RECYCLING?

In addition to providing a dependable, locally controlled water supply, water recycling provides tremendous environmental benefits. By providing an additional source

For over 35 years, in the Montebello Forebay Groundwater Recharge Project, recycled water has been applied to the Rio Hondo spreading grounds to recharge a potable groundwater aquifer in south-central Los Angeles County.

In California, Mono Lake's water quality and natural resources were progressively declining from lack of stream flow. In 1994, the Los Angeles Department of Water and Power was required to stop diverting one-fifth of the water it historically exported from the basin. The development of water recycling projects in Los Angeles has provided a way to partially offset the loss of Mono Basin water, and to allow the restoration of Mono Lake to move ahead. Copyright 1994, Mono Lake Committee.

of water, water recycling can help us find ways to decrease the diversion of water from sensitive ecosystems. Other benefits include decreasing wastewater discharges and reducing and preventing pollution. Recycled water can also be used to create or enhance wetlands and riparian habitats.

WATER RECYCLING CAN DECREASE DIVERSION OF FRESHWATER FROM SENSITIVE ECOSYSTEMS

Plants, wildlife, and fish depend on sufficient water flows to their habitats to live and reproduce. The lack of adequate flow, as a result of diversion for agricultural, urban, and industrial purposes, can cause deterioration of water quality and ecosystem health. Water users can supplement their demands by using recycled water, which can free considerable amounts of water for the environment and increase flows to vital ecosystems.

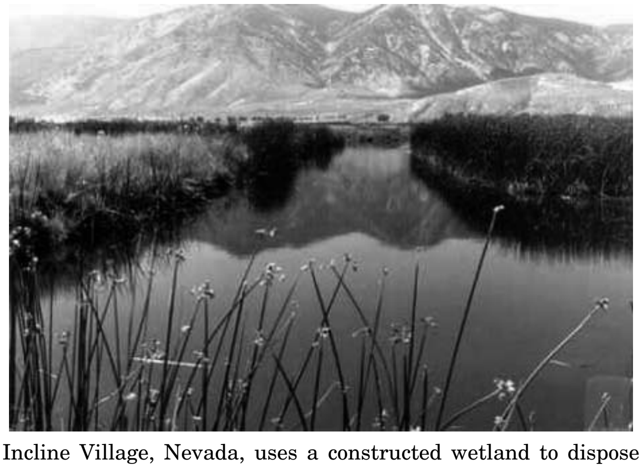

Incline Village, Nevada, uses a constructed wetland to dispose of wastewater effluent, expand the existing wetland habitat for wildlife, and provide an educational experience for visitors.

Recycled water has been used for a number of years to irrigate vineyards at California wineries, and this use is growing. Recently, Gallo Wineries and the City of Santa Rosa completed facilities for the irrigation of 350 acres of vineyards with recycled water from the Santa Rosa Subregional Water Reclamation System.

WATER RECYCLING DECREASES DISCHARGE TO SENSITIVE WATER BODIES

In some cases, the impetus for water recycling comes not from a water supply need, but from a need to eliminate or decrease wastewater discharge to the ocean, an estuary, or a stream. For example, high volumes of treated wastewater discharged from the San Jose/Santa Clara Water Pollution Control Plant into the south San Francisco Bay threatened the area's natural salt water marsh. In response, a $140 million recycling project was completed in 1997. The South Bay Water Recycling Program has the capacity to provide 21 million gallons per day of recycled water for use in irrigation and industry. By avoiding the conversion of salt water marsh to brackish marsh, the habitat for two endangered species can be protected.

RECYCLED WATER MAY BE USED TO CREATE OR ENHANCE WETLANDS AND RIPARIAN (STREAM) HABITATS

Wetlands provide many benefits, which include wildlife and wildfowl habitat, water quality improvement, flood diminishment, and fisheries breeding grounds. For streams that have been impaired or dried from water diversion, water flow can be augmented with recycled water to sustain and improve the aquatic and wildlife habitat.

WATER RECYCLING CAN REDUCE AND PREVENT POLLUTION

When pollutant discharges to oceans, rivers, and other water bodies are curtailed, the pollutant loadings to these bodies are decreased. Moreover, in some cases, substances that can be pollutants when discharged to a body of water can be beneficially reused for irrigation. For example, recycled water may contain higher levels of nutrients, such as nitrogen, than potable water. Application of recycled water for agricultural and landscape irrigation can provide an additional source of nutrients and lessen the need to apply synthetic fertilizers.

At West Basin Wastewater Treatment Plant in California, reverse osmosis, an advanced treatment process, is used to physically and electrostatically remove impurities from the wastewater.

WHAT IS THE FUTURE OF WATER RECYCLING?

Water recycling has proved to be effective and successful in creating a new and reliable water supply, while not compromising public health. Nonpotable reuse is a widely accepted practice that will continue to grow. However, in many parts of the United States, the uses of recycled water are expanding in order to accommodate the needs of the environment and growing water supply demands. Advances in wastewater treatment technology and health studies of indirect potable reuse have led many to predict that planned indirect potable reuse will soon become more common.

While water recycling is a sustainable approach and can be cost-effective in the long term, the treatment of wastewater for reuse and the installation of distribution systems can be initially expensive compared to such water supply alternatives as imported water or groundwater. Institutional barriers, as well as varying agency priorities, can make it difficult to implement water recycling projects. Finally, early in the planning process, agencies must implement public outreach to address any concerns and to keep the public involved in the planning process.

As water demands and environmental needs grow, water recycling will play a greater role in our overall water supply. By working together to overcome obstacles, water recycling, along with water conservation, can help us to conserve and sustainably manage our vital water resources.

STATE AND REGIONAL WATER SUPPLY

Timothy J. Downs
Clark University
Worcester, Massachusetts

INTRODUCTION

Integrated watershed management (IWM) seeks to combine interests, priorities, and disciplines as a multistakeholder planning and management process for natural resources within a watershed ecosystem, centered on water quantity and water quality. Driven bottom-up by local needs and priorities and top-down by regulatory responsibilities, it must be adaptive, evolving dynamically as conditions change (1). Water resource spatial and political scales range from local/municipal, subregional, state, regional (multistate), national, multinational (“regional” in a global sense) to continental and global. Large watersheds occupy state and regional spatial scales, presenting considerable challenges for IWM and water supply, in particular, because of the number and diversity of human consumers and biodiverse habitat needs.

A watershed spatial scale is often used by ecologists to analyze species distribution and biogeochemical cycling. Watersheds can be characterized by the proportions of land area taken up by five main land use/land cover categories (2): urban land, agricultural land, rangeland, forestland, and water. It is arguable whether a watershed or ecoregion scale (region of relative homogeneity based on soil, climate, and vegetative criteria) is most appropriate for managing water quality and aquatic habitats (2). Even in terms of water stocks and flows, the differences between precipitation and evapotranspiration in one watershed that determine the runoff and infiltration that can be used for water supply are influenced by regional climatic and hydrological processes.

Water has many users and uses: drinking, industrial processing, irrigation, hydroelectricity, cooling water for power plants, and recreation (swimming and fishing). It also satisfies the needs of nonhuman species and participates in hydrogeochemical cycles and climate. Competing uses can be sources of conflict, especially when supplies are scarce or degraded by pollution. For example, using a river for industrial processing upstream may conflict with its use as a drinking water supply downstream if toxic discharges are significant. On a state and regional scale, the number of users and competing uses increases, and successful water supply depends on effective communication and negotiation among constituencies. Just as in small watersheds, conflicts in one portion can affect other portions.

The state and regional hydrographic area becomes an operational unit of analysis when water supply and wastewater sanitation are administered on this scale; this means that policies and practices are carried out on a geopolitical scale larger than a municipality or a subregional watershed. Such is the case in many countries, like France, Mexico, parts of the United States, and the United Kingdom. Examples of regional approaches are (1) in France, six water agencies administer the six large hydrographic regions of the country; (2) in the United States, examples of regional management include the Great Lakes Commission and the Ohio River Valley Water Sanitation Commission; and (3) in Mexico, Watershed Councils administer 26 hydrographic regions; several have portions of two or three states.

Much of the emphasis on regional watershed protection in the United States focuses on pollution and water quality, which directly impacts water supply sources and calls for an approach based on risk assessment. For example, in the Mid-Atlantic region (consisting of 18 U.S.

Geological Survey hydrologic units), the Environmental Protection Agency (EPA) is coordinating efforts to identify watersheds that make the most significant contribution to pollution (2). The GIS is employed to model the relationships between water quality indicators and land use/land cover. Water resources management on a state or regional scale is sensitive to the resolution of data gathered and the spatial scale on which predictive models are run and boundaries for analysis set. The pollution of lakes and rivers by nutrients and sediment from nonpoint sources (NPS), it has been shown, is well predicted by land use/land cover patterns (2). Two management strategies are employed to manage NPS pollution: (1) planning and source reduction, for example, limiting impervious surfaces, keeping grazing animals out of streams; and (2) structural and best management practices (BMPs), for example, vegetative borders, conservation tillage.

Water supply on a state and regional hydrographic scale requires drawing up plans that address the following (3): inter- and intra-agency coordination, technical assistance to local agencies and the public, public participation, and program boundaries (e.g., coastal zones must include watersheds contributing NPS pollution). And specifically NPS pollution control programs must address (3) urban sources of pollution, marinas and boats, agricultural sources of pollution, forestry, hydromodification, and wetlands.

INSTITUTIONAL STRUCTURE

Howe (4) has identified key water management requirements for a watershed/multiwatershed scale: (1) coordinated management of surface water and groundwater resources, (2) coordinated management of both water quantity and water quality, (3) provision of incentives for greater economic and technical efficiencies in water use, and (4) protection of public values associated with water service (e.g., reliable, safe, clean, affordable supply). To achieve this, the water institutions must develop the following characteristics (after Ref. 4):

- the capacity to coordinate water plans with other agencies (e.g., urban planning, agriculture, public health, environment, industrial);
- the capacity to solve water problems creatively using a variety of options and approaches (e.g., laws, pricing, taxes, tradeable supply and/or pollution permits, subsidies);
- the foresight to separate roles and responsibilities for water resource planning and management activities from construction activities (i.e., avoid conflicts of interest);
- the multidisciplinary capacity to undertake multicriteria/multiobjective planning and evaluation of alternatives;
- devolve decision-making power to the lowest level—national, state/regional (provincial), local or municipal—consistent with the scale of the water issue;
- capacity and willingness to use appropriate participatory methods involving different stakeholders at different stages of a project (preplanning/conceptual, planning and design, implementation, maintenance, monitoring, and evaluation);
- ability to reward innovation and adapt to changing conditions and priorities.

Federal, state, and regional laws and regulations are crucial instruments that help us respond to the challenge of water sustainability. Such responsive legal instruments should do the following (after Ref. 5):

- Encourage administration at the appropriate hydrological scale: watershed, multiwatershed or aquifer system.
- Foster internalization of the values and ethics of sustainable resource development.
- Encourage integrated approaches to water supply and sanitation, ecology, and public health.
- Prevent water allocation and usage policy making from fragmenting among agencies.
- Promote integrated appraisals, notably environmental, economic, and sociopolitical impacts assessment of alternative actions.
- Encourage integrated capacity strengthening of governmental institutions, NGOs, community associations, and businesses to transition into more sustainable policies and practices and work collaboratively.
- Enforce reward and penalty incentives that encourage sustainability.

WATER SUPPLY PLANNING AND MANAGEMENT

These are the five categories of human consumption (6):

1. domestic (in-house and out-of-house);
2. industrial (e.g., factories, power stations), commercial (e.g., shops and hotels), and institutional (e.g., hospitals, schools);
3. agricultural (e.g., farms, crops, livestock);
4. public (e.g., parks, fire fighting, sewer flushing;
5. Losses—distribution losses, consumer wastage, metering errors, and/or unrecorded consumption.

Table 1 shows the water supply categories used by the U.K. National Rivers Authority for watershed and regional multiwatershed planning and the principal demand considerations.

For state and regional water supply that may include several watersheds and aquifers, coordinating the forecasting of demand is essential for sustainable supply. Forecasting beyond 15 years is highly uncertain; so for longer time frames, we employ scenario projections that ask the question: If these conditions or assumptions apply (e.g., a certain population growth rate, unit demand, supply losses), then what will the future water demand be? The estimation of demand involves (6):

Table 1. Regional Water Use Categories Employed by the U.K. National Rivers Authority and Demand Considerations

Use Category	Typical Uses	Demand Considerations
Drinking water supply	Municipal water supply (surface and/or groundwater; private wells)	Population levels, growth rates and consumption patterns, pricing and wastage
Industrial water supply	Process water supply; cooling waters	Production process technologies and efficiencies
Agricultural water supply	Irrigation waters; livestock watering; milkhouse wash water; livestock housing wash water	Cropland area and crop types and cultivation methods; livestock populations; dairy production
Flood control	Impoundment of high flows for controlled release; construction of dams, reservoirs, levees, channel protection	Flood risks, especially in the floodplain; seasonality and intensity of rainfall and snowmelt; vegetative buffers
Thermal power generation	Cooling waters; settling pond waters; water for pipe flushing and maintenance	Thermal cycle efficiency; power demand and scheduling; capacity
Hydroelectric power generation	Impoundment of water for power generation; construction of dams and reservoirs; pumping and drawdown of water levels	Minimum head required for expected power generation
Navigation	Recreational boating; commercial transport shipping; commercial tourist shipping	Minimum depths and flows required for safe passage
Water-based recreation	Recreational fishing; boating; swimming; hiking, picnicking; nature enjoyment (e.g., bird watching); aesthetic enjoyment	Cultural values, including water odor and color, nature appreciation; seasonality of recreation; numbers of recreators; laws
Fish and wildlife habitat	Aquatic and riparian habitats; protection of community structure; protection of rare and endangered species	Sensitivity of species to disturbance; ecosystem type; types and intensities of stressors (e.g., water loss, pollution)
Water quality management	Protection of minimal flows for water quality preservation; low-flow augmentation from reservoirs; assimilation of waste discharges from municipalities and industries; assimilation of storm and combined sewer discharges	Magnitude, duration and frequency of waste flows; point or nonpoint sources; types and toxicities of contaminants; health risks to humans and other species

Source: After References 7 and 8.

1. plotting the population trend for the past 10–20 years and estimating the proportion likely to be due to immigration and that due to natural increase (births minus deaths);
2. dividing the supply area into different socioeconomic classes of domestic use;
3. estimating the typical domestic consumption per capita in each class;
4. seeking values of future immigration and natural increase for the different classes of housing;
5. Estimating distribution pipe loss rates, consumer wastage, and unsatisfied demand;
6. Estimating the growth in industrial, commercial, and service demands (a function of population growth);
7. Estimating growth in agricultural demand, hydroelectric demand, and public and ecosystem maintenance demand.

Equation 1 is the general supply equation for any bounded area (6).

$$\begin{aligned}\text{Total supply} &= \text{total legitimate potential demand} \\ &\quad + \text{consumer wastage} + \text{distribution} \\ &\quad \text{losses} - \text{unsatisfied demand} \\ &= \text{water supply unreturned} \\ &\quad + \text{wastewater return} \end{aligned} \tag{1}$$

Water can be supplied, and demand can be met by one or a combination of the following options: (1) surface freshwater withdrawal (lakes, reservoirs, rivers and streams), (2) saltwater desalination, (3) groundwater withdrawal (springs and aquifers), (4) rainfall harvesting, (5) wastewater reuse, and (6) demand management (efficiency measures and absolute reduction). For the latter, pricing water equitably is an important consideration, and it has been postulated that there is a inverse power relationship between demand (Q) and price (P), such that (6)

$$Q = k\,P^{-x} \tag{2}$$

where k is a constant and x is the elasticity between Q and P.

In addition, in developing countries, the price of water must be tiered according to ability to pay, such that the poorest people (those at a subsistence level without surplus income) are fully subsidized, whereas high-end consumers (more than 300–400 liters/person/day) pay the most and have a financial incentive to reduce their consumption.

REGIONAL WATER TREATMENT AND DISTRIBUTION

Water treatment technologies seek to produce adequate and continuous supplies of water that are of sufficient quality—chemically, microbiologically, and aesthetically—for their intended uses. Treatment consists of a series of physical and chemical unit processes that are designed as a function of the source and quality of the raw water, the required water quality produced, and the throughput (or flow rate). A typical full water treatment system consists of eight stages (9): (1) intake, (2) pretreatment (e.g., screening, neutralization, aeration); (3) primary treatment (e.g., coagulation, flocculation, sedimentation); (4) secondary treatment (rapid and slow sand filtration); (5) disinfection (e.g., chlorination or ozonation); (6) advanced treatment (activated carbon, membrane diffusion); (7) fluoridation; and (8) distribution.

On a state and regional scale, sufficient water treatment capacity must be built and planned for using regional demand forecasting. The location of water treatment plants depends on the location of the source of raw water, the location of the users, and whether or not it is more cost-effective to build a few large plants to serve the region or a network of smaller local plants. This treatment plant spatial configuration will also determine, and be determined by, the distribution system design. The distribution system consists of a network of pipes (9): (1) trunk mains that bring large flows from the source to the treatment plant and then to storage reservoirs or towers; and (2) distribution mains that deliver water from the service storage to the users, a highly branched network. Service reservoirs and water towers cope with the diurnal variation in demand by satisfying peak demand, supplementing the average daily flows. Service reservoirs must be able to satisfy the population for at least 24–36 hours and provide enough pressure to reach the storage tanks of individual users, perhaps more than 50,000 households (9). The hydraulic head for fire fighting is at least 30m, and maximum head is about 70m (above this, leaks and bursts are common). In flat areas, water is fed by gravity from water towers, which can incur high pumping costs. High-rise buildings require their own pumping. Typically, each user has a supply pipe that connects to the communication pipe at the property boundary, which in turn connects to the service main. Water distribution networks serve different zones, each one supplied by a service reservoir and/or tower (9). On a regional and state scale, the planning and management of such an infrastructure must be carefully coordinated to meet the growing demands of a diverse set of users. In times of drought, regional water supply is severely tested, and the risk of drought must be addressed with a careful regional provision of adaptive reserve storage and aggressive water conservation strategies that penalize nonessential use.

WATER SUPPLY IN DEVELOPING COUNTRIES

Table 2 summarizes the sources, decision criteria, delivery methods, and levels of service for water supply in developing countries. Three main factors affect the choice of water supply in developing countries (11):

- institutional—land tenure, administrative structure, finance level;
- population—demography, size, growth rate, socioeconomic levels, attitudes, and behaviors;
- technical—scale, geophysical conditions, urban layout, alternatives, resources.

For the rural subsistence poor, the main water supply problem is stable access to safe water. For the urban poor and periurban marginalized communities, the main problem of water supply is that the connection and supply costs are too high.

On a state or regional scale, these water supply issues become magnified, and it is no surprise that the most controversial aspect of water supply in developing countries is selling of water administration concessions to private water companies by a central government. Recognizing water as an economic good is essential for efficient management, but it should be priced equitably according to users' ability to pay. For subsistence communities, a full

Table 2. Key Aspects of Water Supply in Developing Countries

Sources	Selection Factors	Delivery Methods	Levels of Service
• Springs • Rivers • Lakes • Aquifers • Rainfall • Snowmelt • Recycled wastewater • Desalinated saltwater	• Distance from source to user • Water quality at source • Cost of raw water and delivery • Scale of supply • Alternatives available • Sanitation method • Return drainage method • Preferences of users • Land and water rights • Risk of shortfall, outage, pollution	• Hand transport in containers • Vehicular transport • Pipelines (gravity, pumped) • Pipelines to troughs, reservoirs, standpipes, communal ablutions, households, businesses, public places, fields of crops • Channels	1. Simple source with disinfection 2. Pump and supply line to reservoir or standpipe 3. Household standpipes or connections 4. Individual household supplies

Source: After Reference 10.

public subsidy is appropriate for basic needs (1): overprivatization risks occur when water services are run purely for profit and marginalized communities unable to pay incur crippling water debts (12). This has occurred in South Africa and touches a nerve in many other places where water is viewed as both a public good and a human right that should not be controlled by commercial interests. We should also recognize that the problem with user willingness to pay is cultural (habits of free water) and also one of poor service quality: Why should a person pay for water that is not clean and comes only intermittently?

WATER POLLUTION AND HEALTH RISKS

Pollution produced in one part of a watershed or regional hydrographic area poses potential health risks to other parts. Microbiological water pollution causes many persistent water-related diseases worldwide, whereas chemical water pollution adversely affects humans and other species, especially sensitive aquatic organisms. Much of the serious chemical pollution stems from runoff of agricultural chemicals (fertilizers containing nutrients and pesticides such as DDT) and industrial wastes such as solvents and heavy metals. For example, many tributaries of the Amazon Basin have been contaminated by mercury from mining waste, resulting in methylmercury levels in fish well above the World Health Organization (WHO) standard (13). Another dramatic example of regional mercury contamination occurred in Canada in 1984, where 64% of all Cree Indian residents of Chiasabi had levels above the WHO standard (13). Nash (14) claims that the main obstacles to effective water quality management are cost and a lack of information. Downs argues that the problems go much deeper to cultural obstacles and weak societal capacity (15).

AQUATIC ECOSYSTEMS

Freshwater lakes, ponds, rivers, streams, wetlands, and aquifers are inextricably linked by hydrologic cycles acting on watershed, multiwatershed, regional hydrographic, and continental scales. A mere 1% of land surface is covered by freshwater, but it contains 12% of the animal species (16). Conservationists claim that freshwater biota are being destroyed by anthropogenic impacts faster than they can be studied or protected: overconsumption by human populations, cultural eutrophication, acidification, dam construction and hydromodification, the introduction of invasive species, habitat loss from development, and climate change. Examples are all too plentiful (16): (1) the invasion by lampreys (an ocean parasite) of the Laurentian Great Lakes after the construction of the Erie Canal, resulting in the decimation of salmonid species; (2) the invasion of the same system by the zebra mussel and spiny water flea, upsetting native invertebrates and zooplankton; and (3) the massive diversion of river water to the Aral Sea on the border of Kazakhstan and Uzbekistan to irrigate cropland, leading to a 65% decrease in volume from 1926–1990, destruction of the fisheries, economic collapse, and health problems.

The Aral Sea is a dramatic example of the need to take a regional approach to water supply and wastewater sanitation and to account for the coupled needs of humans and other species. Covitch (16) argues that technology alone—such as GIS monitoring and remote sensing—cannot hope to address the pace or complexity of freshwater ecosystem degradation. Rather, institutional reforms must provide water to aquatic habitats and minimize adverse impacts, while enabling multidisciplinary research is done on regionally representative biotas and hydrologic regimes (16).

WATER SUPPLY FOR AGRICULTURAL PRODUCTION

The twentieth century's Green Revolution boosted agricultural production and was made possible in large part by doubling the amount of irrigated land from 1950–1980 (17). In 1989, a third of the world's food harvested came from the 17% of the cropland that was irrigated, and two-thirds of the freshwater supply is used for irrigation (17). Clearly, on a state and regional scale, for example, California, agricultural demands are often the largest of all demands, and so savings derived from more efficient irrigation technology and/or diversions to other uses can have a huge impact on water supply sustainability. There is concern that growing populations, especially in water-scarce developing countries, will overwhelm the capacity of irrigated cropland to provide sufficient food and that water stress will become even more acute. The cost of adding to land currently irrigated is high in places such as Africa because of weak capacity and the need to build supporting infrastructure. In addition, many rivers are seasonal.

There is great scope for improving the efficiency and management of irrigation water: (1) using much less water per hectare [often less than half the water applied benefits the plant (17)] and (2) ensuring that water is delivered at the right time in sufficient quantities. Many schemes that involve massive irrigation cause fields to become waterlogged and soils to become salinized, leach nutrients, and lose fertility. Grossly inefficient schemes also contribute directly to the overexploitation of aquifers and surface waters and the degradation of aquatic ecosystems. Postel (17) has estimated that about 20% of the irrigated land in the United States relies on overexploited groundwater and causes major state and regional water supply impacts. Similar problems exist in many other countries, such as China, India, and the Middle East. Already, in the Great Plains of the United States, farms have been abandoned because of falling water tables; in the Texas High Plains Region, irrigated land area shrunk by 34% from 1974–1989 because the cost of overpumping the Ogallala Aquifer System exceeds the value of crops grown there (17).

Diversions of irrigation water to other uses, especially urban areas, is becoming a more common practice; some farmers sell water for more than they would earn from using it to grow crops. Los Angeles, Mexico City, and Beijing are just three examples of megacities that rely on importing water from rural farming zones. Unless laws and regulations adequately compensate the regions that lose the water to thirsty cities, serious issues of environmental justice exacerbate rural poverty.

WATER SUPPLY FOR ENERGY PRODUCTION

Water supply and energy supply are closely related, and conflicts over access, costs, pricing, and the overarching sustainability challenges for both resources invariably run parallel (18). Energy production relies on reliable water supplies: fossil fuel, geothermal, and nuclear power plants draw vast amounts of water from lakes, rivers, or the ocean for cooling. In water-scarce regions, either withdrawals alter hydrologic regimes and impact ecosystems, water scarcity places absolute limits on energy production, or scarcity forces interbasin water transfers. Droughts in Africa and Asia and even California have significantly impacted hydroelectric production, causing shortfalls that either go unmet or must be met by scaling up other energy sources (18). For example, a decade of drought in Egypt severely reduced the capacity of the Aswan Dam that supplied half of Egypt's power. The massive investment worldwide in dams, aqueducts, and reservoirs was made to overcome both the uneven distribution of freshwater for water supply and to provide hydroelectricity for development. Cheap energy has allowed water resources to be superexploited, that is, pumping of deep fossil groundwater and pumping of interbasin water transfers over mountains in water-stressed regions. In this way, just as coal and oil allowed industry to escape the solar budget, cheap energy also allows settlements to escape the immediate constraints of local hydrology. But on a regional scale, such constraints are not escaped, merely reconfigured temporarily. Sustainable regional water and energy supply must be considered as coupled resources and managed accordingly.

WATER SUPPLY POLITICS AND LAW

All regional issues impacting water supply—urban demand, public health, ecosystem integrity, agricultural demand, energy needs—can be effectively addressed only by state and regional institutions that design and apply the aforementioned responsive regulatory instruments and allow the participation of regional stakeholders in water supply planning and management. McCaffrey (19) states that there are two major principles that underpin most international water treaties: (1) the principle of equitable utilization—the apportionment of uses and benefits of a shared watercourse should be made in an equitable way; and (2) that a country, through actions that affect an international watercourse, may not significantly harm other countries. The key to successful comanagement is sharing of information and communication, especially about development plans. These same ideas should be applied to state and regional water supply within a country.

CLOSING REMARKS

Water supply on a state and regional scale poses special challenges for planning and management because of the large number and wide diversity of users and habitats to support; the related uncertain forecasting of demand; the need to provide a reliable, affordable service; and the number of potential multisectoral priority issues (energy, agriculture, health, ecology) to address. In practice, the success of this depends on communication and coordination on three levels: (1) between each village, town, city, farm and industry of the region; (2) among the water, industrial, energy and agricultural development sectors of the state or region; and (3) among the water utility companies, their customers, and local and state government regulators. Responsive institutions and regulations are vitally important to achieve this, to mitigate the risks of drought and pollution, and to meet the growing demand, especially in water-scarce/water-stressed regions where the competition for safe water is increasingly intense.

BIBLIOGRAPHY

1. Downs, T.J. (2002). A participatory integrated capacity building approach to the theory and practice of sustainability—Mexico and New England Watershed Case Studies. In: *International Experiences on Sustainability*. W.L. Filho (Ed.). Peter Lang, Frankfurt am Main, pp. 179–205.
2. Hunsaker, C.T., Jackson, B.L., and Simcock, A. (1998). Regional assessment for watershed management in the Mid-Atlantic States. In: *Watershed Management—Practice, Policies and Coordination*. R.J. Reimold (Ed.). McGraw-Hill, New York, pp. 11–29.
3. Smith, J.P., Carlisle, B.K., and Donovan, A.M. (1998). Watershed protection and coastal zone management. In: *Watershed Management—Practice, Policies and Coordination*. R.J. Reimold (Ed.). McGraw-Hill, New York, p. 99.
4. Howe, C. (1995). *Guidelines for the Design of Effective Water Management Institutions Utilizing Economic Instruments*. Report presented at Workshop on the use of economic principles for the integrated management of freshwater resources, United Nations Environment Programme (UNEP), Nairobi, Kenya.
5. Smith, D., and Rast, W. (1998). Environmentally sustainable management and use of internationally shared freshwater resources. In: *Watershed Management—Practice, Policies and Coordination*. R.J. Reimold (Ed.). McGraw-Hill, New York, p. 294.
6. Twort, A.C., Ratnayaka, D.D., and Brandt, M.J. (2000). *Water Supply*, 5th Edn. Arnold and IWA, London, pp. 1–35.
7. NRA (National Rivers Authority) (1993). *National Rivers Authority Strategy*. NRA Corporate Planning Branch, Bristol, UK.
8. Heathcote, I.W. (1998). *Integrated Watershed Management, Principles and Practice*. John Wiley & Sons, Hoboken, NJ, p. 74.
9. Gray, N.F. (1999). *Water Technology, An Introduction for Water Scientists and Engineers*. Arnold, London, pp. 205–226.
10. Stephensen, D. (2001). Problems of developing countries. In: *Frontiers in Urban Water Management*. C. Maksimovic and J.A. Tejada-Guibert (Eds.). UNESCO and IWA, London, p. 284.
11. Lyonnaise de Eaux. (1998). *Alternative Solutions for Water Supply and Sanitation in Areas with Limited Financial Resources*. Paris.
12. Gleik, P., Wolff, G., Chalecki, E.L., and Reyes, R. (2002). *The New Economy of Water: The Risks and Benefits of*

Globalization and Privatization of Fresh Water. Pacific Institute, Oakland, CA.

13. Gleik, P. (1993). An introduction to global fresh water issues. In: *Water in Crisis, a Guide to the World's Fresh Water Resources*. P. Gleik (Ed.). Oxford University Press, New York, pp. 1–12.
14. Nash, L. Water quality and health. In: *Water in Crisis, a Guide to the World's Fresh Water Resources*. P. Gleik (Ed.). Oxford University Press, New York, pp. 26–37.
15. Downs, T.J. (2001). Changing the culture of underdevelopment and unsustainability. *J. Environ. Plann. Manage.* **43**(5): 601–621.
16. Covitch, A.P. (1993). Water and ecosystems. In: *Water in Crisis, a Guide to the World's Fresh Water Resources*. P. Gleik (Ed.). Oxford University Press, New York, pp. 40–55.
17. Postel, S. (1993). Water and agriculture. In: *Water in Crisis, a Guide to the World's Fresh Water Resources*. P. Gleik (Ed.). Oxford University Press, New York, pp. 56–66.
18. Gleik, P. (1993). Water and energy. In: *Water in Crisis, a Guide to the World's Fresh Water Resources*. P. Gleik (Ed.). Oxford University Press, New York, pp. 67–79.
19. McCaffrey, S.C. (1993). Water, politics and international law. In: *Water in Crisis, a Guide to the World's Fresh Water Resources*. P. Gleik (Ed.). Oxford University Press, New York, pp. 92–104.

RIVER BASIN DECISIONS SUPPORT SYSTEMS

Charles D.D. Howard
Water Resources
Victoria, British Columbia,
Canada

The interacting components of water resource systems incorporate physical, biological, chemical, and institutional constraints and objectives (Fig. 1). Water projects are difficult to manage efficiently within potentially conflicting objectives and constraints on water quality, river flows, and lake levels. There are many types of decisions, and these must be made on a timely and reliable basis. Decisions vary from timescales of minutes for assigning loads to hydroelectric generating units to weeks and months for managing water quality, reservoir levels, and pumping from groundwater. On shorter timescales, decisions are based on deterministic analyses. On longer timescales, decisions are dominated by consideration of probabilities.

Planning can benefit from decision support systems developed for actual operations. Sound plans require realistic analysis that these systems provide for changes

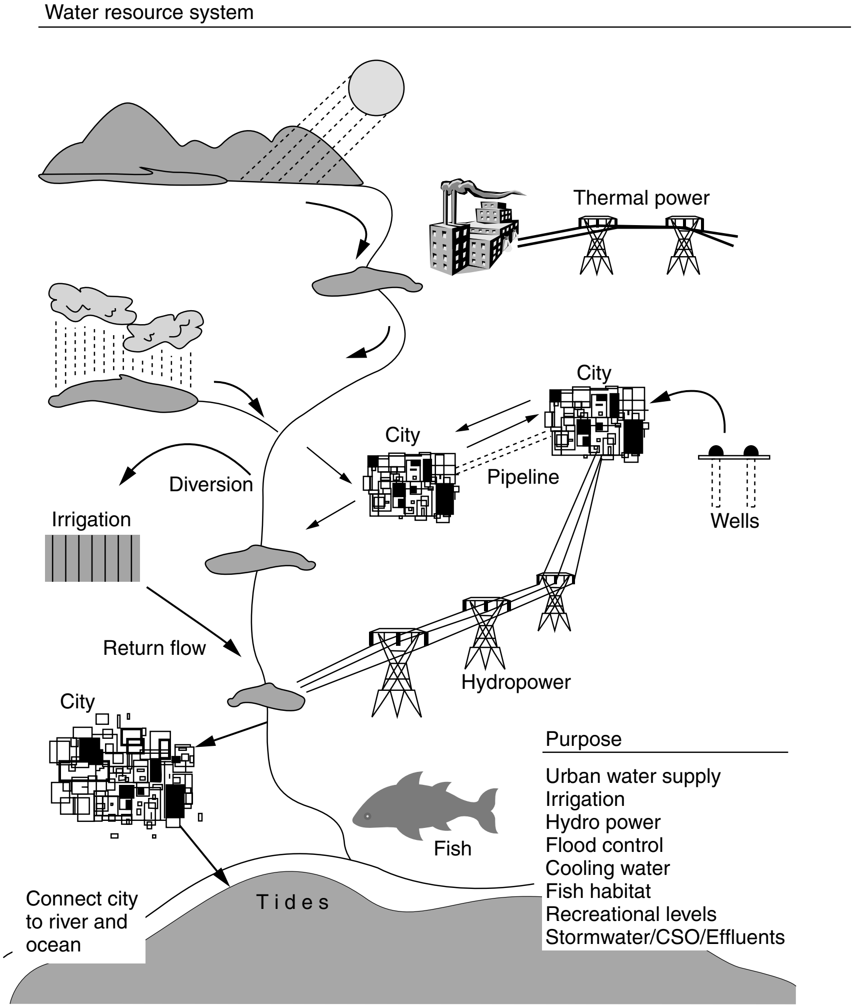

Figure 1. Water resource systems are complex.

in operating parameters, capacity expansion, and the benefits and costs of potential new environmental objectives. Figure 2 illustrates how many of the same program modules and data are used for planning studies and for actual operations.

Until recently, the data and computer simulation analysis behind water management decisions simply supported experience and judgments. In the latter half of the twentieth century, *decision support systems* based on optimization techniques were developed to recommend the best possible decisions. These software systems develop recommendations by continuous analysis of constantly changing data and evolving priorities. The key components of decision support systems are data, computer models of processes, and trained people.

The purpose of a decision support system is to acquire data and to increase its informational content. To be effective, a decision support system should execute rapidly in response to changing data. Computer input and output should be in a form that is most convenient for the person making the decisions. Modern computer systems minimize the manual input required and automatically suggest optimal operating decisions that recognize all current factors.

COMPUTER MODELS FOR WATER MANAGEMENT

There are two related, but fundamentally different types of water resource models—scientific models and management models. Scientific models are used primarily to investigate physical relationships in some detail. They provide a surrogate for extensive field monitoring by using available field data and scientific principles to interpolate and to extrapolate. They in-fill missing information on how the physical system is operating. A secondary purpose for scientific models is to test how different assumptions about future or past physical conditions would affect the system.

Management models are used primarily to suggest water management decisions—they are the main tools of decision support systems. They include a description of the scientific, institutional, and economic factors that bear on the decision-making process. To accomplish this, they must be comprehensive in modeling the interactions among components of these systems. For some types of decisions, an approximate description of the physical processes is adequate; for most types of decisions, the institutional and economic factors must be comprehensively described because they make up the key objectives in water management decisions.

Scientists should not expect management models to be absolutely accurate in scientific matters. Managers should be prepared to have recommendations from management models reality-checked by scientists.

SPREADSHEETS IN DECISION SUPPORT SYSTEMS

Spreadsheets can be the engine of a decision support system for planning studies but are unsuitable for supporting short-term operating decisions for data-intensive water resource systems. They are used routinely for hydrologic studies and for economic analysis of water resource systems. A spreadsheet can model operation of reservoirs, aqueducts, treatment plants, and water quality. The sensitivity of decisions to data

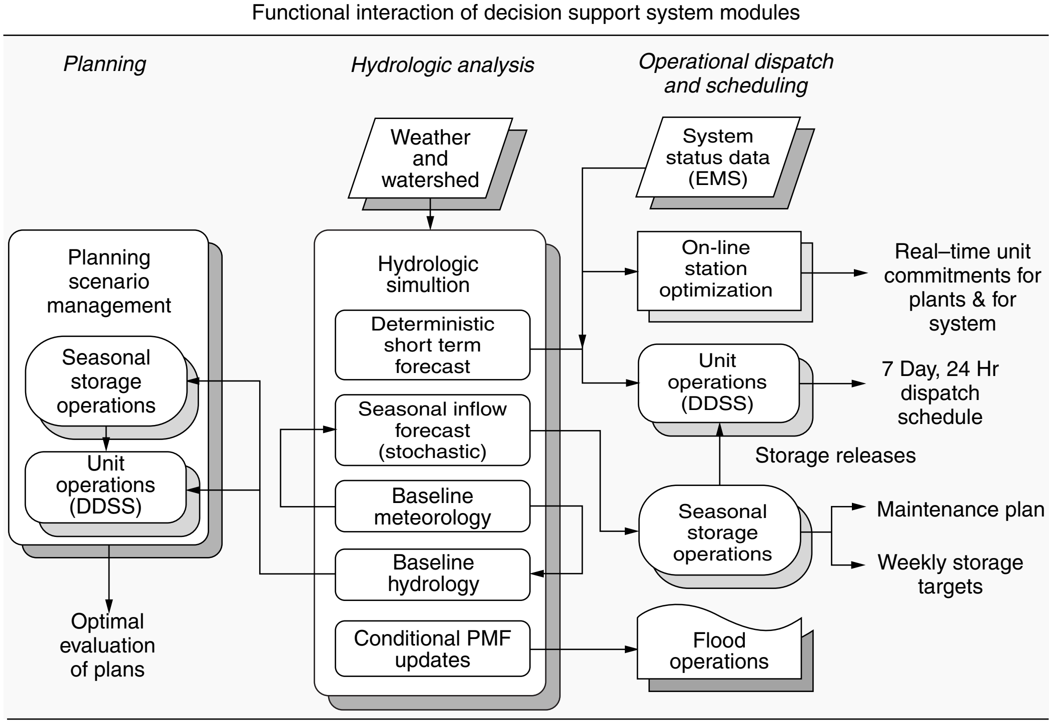

Figure 2. A decision support system for hydroelectric planning and operations.

can be examined, and specific recommendations for improvements in other models can be developed with the aid of spreadsheets.

Macro capabilities in spreadsheets are used for managing data files, displaying and printing graphs and tables, and for eliminating the need for others to be experts in programming. Macros can readily merge spreadsheets and acquire data from servers without user interaction. The detailed logic for water planning can be examined to determine how it might be structured for optimization routines in a more complex model.

Spreadsheet add-ons provide routines that can be very useful for optimizing water management decisions. A spreadsheet SOLVER function deals with nonlinear constraints and automatically drives the spreadsheet to maximize objectives within specified constraints. This eliminates trial and error in using the spreadsheet, reduces the number of explicit logical operators, and simplifies technology transfer and user training. Aftermarket optimization routines from third-party vendors extend this capability to larger more complex management issues.

A spreadsheet offers the following advantages: It can be a platform for quickly trying out new ideas, other computer models and their output can be conveniently evaluated, spreadsheets are widely used and upgradeable as computer technology advances, and output tables and graphics can be readily tailored to different needs. A spreadsheet is an easy link with past methods of operating a water resource system, and it can be carried forward into the future with minimal risk of obsolescence.

Spreadsheets are not a good choice for large complex systems that require comprehensive data management and quality control, a high degree of automation, and rapid execution. Special purpose water management computer programs are more effective for large water resource systems if they are structured conveniently for the specific application. But, as part of an overall decision support system, the spreadsheet format offers advantages in rapid implementation, capabilities for data manipulation, flexibility in presentation of results, and ease of understanding.

DEALING WITH UNCERTAINTY IN HYDROLOGIC DATA

On longer timescales, allowances for uncertainty may be intuitive—primarily colored by recent experience or based on hydrologic ensembles. A hydrologic ensemble is a set of inflow records used in a computer model to test river and reservoir operations across a range of possible inflow conditions. The ensemble members are input to simulation or optimization routines that model the operating constraints and objectives and provide information that supports actual decisions.

In some applications, hydrologic ensembles are developed with a calibrated hydrologic forecasting model from initial hydrologic conditions and historical weather data. Meteorologic ensembles generated by medium-range weather forecasting models offer the potential for conditioning future weather on current atmospheric conditions, and they eventually may replace historical weather data for forecasting hydrologic ensembles.

Hydrologic ensembles also may consist of the records of river discharge or synthetic streamflow sequences based on the autocorrelation and cross correlation characteristics of basin discharge records. Ensembles based on historical discharges are missing the atmospheric and basin moisture conditioning information that limits the range of likely future discharges. It can be observed that historical stream flow records contain a wider range of inflow possibilities than it is realistic to expect for the current conditions that are faced during operations. Heuristic methods may consider that each member of an ensemble is equally likely or some members of the ensemble may be given more weight, for example, if an El Niño condition is relevant or there is a heavy snow cover. Uncertainty of hydrology is reduced to some extent in forecasts that realistically deal with the initial watershed state and with improved future medium-range weather forecasts.

The realities of water project operations are often avoided in project planning. Postconstruction project evaluations and benefits seldom, if ever, relate back to the projections made during planning. Part of this difficulty can be avoided if the planning methodology simulates how actual operating decisions will be made. It is unrealistic to plan a system without considering the forecasted inflows during actual operations and how the operators will respond. Deterministic optimization does not consider uncertainty at all, and simulation studies based on rule curves and historical flow data lack both the finesse of optimization and a realistic expectation of near-term inflows. Water resource planning studies should include some measure to simulate the forecasts of water supply and demand that are updated at each decision point during actual operations.

EXAMPLE OF A HYDROELECTRIC DECISION SUPPORT SYSTEM

With the advent of market competition, there are clear advantages to being in position to provide quickly an optimum generation schedule based on projections of market prices and the availability and value of water. For a cascade of dams, this is a complex decision process. When water is in short supply, maximizing revenue or ensuring that all generating stations are loaded so that overall system generation is as efficient as possible will optimize the 168-hour generating schedule. This requires allocating load so that plants operate at the relatively few discrete points of maximum efficiency determined by optimal use of the generating units within each plant while meeting operational constraints (Fig. 3). Water movement between plants is considered in the optimization so that head is maximized, spill is minimized, and generation takes advantage of time of day variations in loads and energy prices.

Additional considerations include very short-term adjustments to generation and longer term plans for the most effective use of reservoir storage. Planning for maintenance of the machinery considers availability of crews, seasonal variability of hydrology, environmental constraints, and energy prices. The decision support

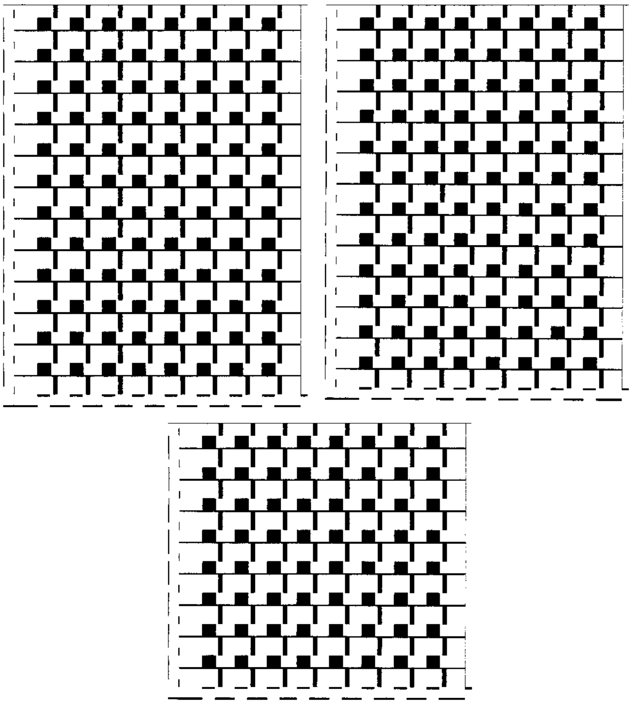

Figure 3. Hydroelectric plants have discrete points of maximum efficiency.

system recommends the optimum decisions by considering hydrologic probabilities and risks.

Each decision support system is slightly different, depending on the requirements for water management and the date when the software was developed. For a hydroelectric decision support system, the following functions are now supported by software (1):

- water management through reservoir release scheduling,
- systemwide maintenance planning for generators,
- forecasting hourly and seasonal local unregulated inflows,
- near-real-time unit loading optimization,
- hourly schedules for all generators in the system, and
- convenient data acquisition, quality control, archiving, and retrieval

Each of these functions requires access to a common pool of consistently reliable data. Static data describe the physical facilities and operational constraints such as minimum river flows and allowable ramping of river and reservoir levels. Dynamic data include hydrometerologic data, weather forecasts, discharge throughout the river system, projected hourly value of energy, the state of generating machinery, reservoir and tailwater levels, spill, current system electrical load, and special operating goals and constraints.

Some of these data are current, some are not—some have been previously checked, and some are in a raw form and may contain errors. The software provides quality control that takes data from several sources, on different time steps, of different quality, and concatenates it to provide currently useful information.

Recent installations are PC-based; local workstations use the Windows operating system. Data are stored on dedicated central servers with an SQL or Oracle database. The database sends and receives data from various servers and workstations around the system via an internal network. An archive of historical operating data supports projections and comparisons with past conditions.

REQUIREMENTS FOR A WATER RESOURCES MANAGEMENT DECISION SUPPORT SYSTEM

A customized modeling and database system will provide convenient graphical current and historical information and operating recommendations that are automatically optimized by the system. Data handling must be convenient with minimal opportunities for introducing human error and maximum flexibility for editing storing, retrieving, and displaying data and model results. There should be an emphasis on realistically designed graphical user interfaces and interactive graphical editing that reflect the preferences of the intended operators.

The data management model consists of a database system and site-specific capability for editing and displaying data in a manner that is keyed to the basin mapping. The model will be informative and practical for routine quality control, storage, retrieval, and graphic display. The database will provide a home for the monitoring data and an integrated backbone for all of the modeling software. The data will probably come from more than one source, so the database must be compatible with the communications facilities.

The hydrologic forecasting model establishes local inflow forecasts for all of the reservoirs. Two types of forecasts are necessary: deterministic hourly forecasts for one week that support daily operating decisions and weekly conditional probability forecasts for 1 year that support longer term management strategies. Both types are driven by weather data and weather forecasts. Initial conditions for each forecast reflect the computed or measured current moisture state of the basin, including soil moisture and water in storage as snow. Conditional probabilities are derived from historical weather data and the current initial moisture conditions of the basin. Deterministic forecasts include hydrodynamic routing in river systems.

An effective implementation often uses a combination of a calibrated conceptual model of the hydrologic cycle and field monitoring of river discharges, soil moisture, and snow cover. River discharge measurements provide forecast verification and information for model readjustment during operation. The local inflow forecasts are routinely updated daily and more frequently when there is concern for floods.

The water demand forecasting model considers the planning objectives, past demands, current demands and forecasts, and hydrologic forecasts. The model includes individual subarea requirements that place significant demands on the system. This is an important component for ongoing water management planning in response to changing circumstances, including water and energy demand management during drought.

The reservoir operations models must be practical, fast, and provide comprehensive optimizations that reflect complex site-specific physical and institutional details and prioritized or multiple objectives. River basin operation may be simulated by suitably constraining the optimization model(s) or by a special simulation model. Generic river and reservoir models used for planning are unsuitable for operational scheduling within a decision support system if they lack features that minimize manual data input and have excessive execution time during optimization.

The river and reservoir operations optimization model(s) use forecasts of local unregulated inflows to recommend the best long-term and short-term strategies. The objective function may include water supply for irrigation and cities, hydroelectric production, and flood management within environmental and facilities management constraints, specific discharge and water level targets, and physical limitations of the system of reservoirs and river channels. All components and operating constraints of important facilities are represented in the optimization modeling system. A simulation capability can be used to test the viability of operating plans developed by judgments that do not follow the recommendations from deterministic and probabilistic optimization models.

When there are several operating objectives, optimization can be accomplished in the software by using a set of prioritized objectives. For example, in flood management, the first priority might be to protect downstream property from flooding. With this as the objective, the software determines an operating sequence for spillway gate operations. Then, with the resulting minimized downstream flood damage as a constraint, a new objective function is used to determine the gate operations that will minimize upstream flooding. Finally, with both the upstream and downstream previously optimized flood damages as constraints, a third optimization minimizes the number of nighttime gate changes. This last step determines the sequence of gate operations actually recommended for managing the flood with available reservoir storage. With a suitable optimization model, the method is quick because each successive optimization uses the feasible solution from the previous analysis as a starting point.

The method of prioritized objectives also may be useful as a practical procedure for defining trade-offs for other types of water management decisions that require considering multiple objectives. Revising the priorities and reoptimizing can examine trade-offs. This application of a decision support system has the advantage of avoiding the difficult and often confusing task of assigning relative weights to objectives. The significance of priorities may be more readily grasped than the relative value of objectives that may be incommensurate.

BASIC REQUIREMENTS

The entire system of software must run on an operating system that is compatible with standard word processing and spreadsheet software. Maintenance for the entire system should be minimal. The software and data should be capable of being maintained and upgraded by computer specialists readily available to the owner of the decision support system. The water management models, operating system software, and computer hardware should have sufficient redundancy to provide a high level of reliability. The entire modeling system must run on the owner's computer network and must provide remote access to the database by users who are not connected to the network. The system must have levels of password security to control access to various portions of the database and modeling modules.

WHY HAVE A DECISION SUPPORT SYSTEM?

A common bottom line question among water resource system managers is: What is the value of a decision support system? The benefits are not uniform for all systems because of the relative importance of uncertainties and the multitude of institutional and physical constraints on operation. The value of a specific decision support system can be estimated by reconstructing the actual operating history and optimizing at each step as each operating decision was made. The decision support software can be designed to automate such an audit as a routine business practice aimed at raising overall efficiency within an enterprise.

Since 1987, a decision support system has been used to guide weekly reservoir release decisions at two hydroelectric plants in the coastal mountains of British Columbia, Canada. Studies of 1970–1974 operations (a period before the decision support system became operational) showed that, compared to operation with perfect foresight, as determined by a deterministic optimization model, the rule curve based operation had produced 83.4% of the maximum attainable energy compared to 95.1% with the full decision support system. Without the hydrologic forecast component, the optimization component would have produced 92.8% by simply using long-term average monthly inflows in place of the forecast. The actual energy produced by operating with the DSS in each year between 1989 and 1993 was 100, 93, 98, 94, and 96% compared to the maximum possible. The decision support system provides accessible data and a consistent framework for improving operating decisions.

Decision support systems can be expensive, but they bring improvements through more effective management to achieve objectives, reduced personnel requirements, fewer violations of environmental constraints, more effectively trained personnel, and objective technical performance evaluations. For a complex river system with many control points and reservoirs, a team of operators is trained to use the software modules and to interpret their output. For systems with only two or three major facilities, one or two trained operators can manage a highly automated, complex decision support system.

BIBLIOGRAPHY

1. Livingstone, A., Smith, D., Van Do, T., and Howard, C.D.D. (1999). Optimizing the hydro system. *Hydro Rev.* **18** (5).

READING LIST

Van Do, T., and Howard, C.D.D. Hydro-power stochastic forecasting and optimization. *Proc. ASCE 3rd Water Resour. Operations Manage. Workshop*, Ft. Collins, CO, June 1988.

Howard, C.D.D. (1999). Death to rule curves. *26th ASCE Conf., WRPM*, Tempe, AZ, June.

Stedinger, J.R. and Howard, C.D.D. (1996). Control room III: finding work for humans. *Hydro Rev.* August.

Stedinger, J.R., and Howard, C.D.D. (1995). The control room of the future II. *Hydro Rev.* July.

Howard, C.D.D., Shamir, U., and Crawford, N. (1995). Hydrologic forecasting for reservoir operations. *ASCE WRPM Div. Conf.* Boston, May.

Howard, C.D.D. (1994). Optimal integrated scheduling of reservoirs and generating units. *Leading Edge Technol. Hydro-Vision Conf.*, Phoenix, AZ, August.

Howard, C.D.D. (1993). Sewer and treatment plant operations control for receiving water protection. *Proc. 6th Int. Conf.* Urban Drainage Syst., Niagara Falls, Ont., Sept.

Stedinger, J.R. and Howard, C.D.D. (1993). A look at the control room of the not-too-distant future. *Hydro Rev.* August.

Howard, C.D.D. (1992). Experience with probabilistic forecasts for cumulative stochastic optimization. *AWRA Annu. Conf.*, Reno, NV., Nov.

Howard, C.D.D. (1992). Optimal operation of reservoirs and hydro plants during floods. *4th Water Resour. Operations Manage. Workshop, ASCE Water Resour. Plann. Manage. Div.*, Mobile, AL, Mar.

Howard, C.D.D. (1992). Optimization methodology for hydropower and instream flow operations. *Annu. Conf. Northwest Hydropower Assoc.*, Bellevue, WA., Jan.

Howard, C.D.D. (1990). Achieving effective technology transfer. In: *Decision Support Systems in Water Resources Planning*, Loucks et al. (Eds.). NATO sponsored workshop in Ericiera, Portugal, Springer-Verlag, Berlin.

Howard, C.D.D. (1988). Expert systems analysis of hydropower scheduling with forecasting uncertainty. *15th* Specialty Conf., *ASCE Water Resour. Plann. Manage. Div.*, Norfolk, VA, June.

Howard, C.D.D. (1984). Very short term forecasts for optimal flood control and hydroelectric peaking operations. *AWRA Symp.* Forecasting Water Manage., Seattle, WA, June.

Howard, C.D.D. and Flatt, P.E. (1976). Multiple objective optimization model for real-time operation of a water supply system. *Int. Symp. Large Scale Syst.*, Univ. of Manitoba, Winnipeg, Aug.

WATER RESOURCE SUSTAINABILITY: CONCEPTS AND PRACTICES

Timothy J. Downs
Clark University
Worcester, Massachusetts

INTRODUCTION

The most famous definition of sustainable development is "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (1). This unifying concept was the theme of the 1992 Earth Summit in Rio de Janeiro that produced the so-called Blueprint for Sustainable Development, *Agenda 21* (2). At the core of the sustainability paradigm is the desire to reframe the perceived competition between development and environment (economy vs. ecology) into new forms of development ethics, thought, practice, and policy, reconciling enhancement of the quality of (human) life and social well-being with the conservation of ecological integrity. Underpinning this lofty goal was the ethical pursuit of a more equitable society both today (intragenerational equity) *and* in the future (intergenerational equity). The term *sustainable development* is somewhat of an oxymoron because it implies stability, yet change, and begs the coupled questions: What to sustain, for how long, and for whom? This has led to confusion and a plethora of debate over the conceptual validity of the model. A useful working definition of a sustainable project, program, practice, or policy is one that yields a steady stream of benefits (or positive impacts) that exceed costs (negative impacts) over intergenerational timescales (25 years and beyond), while conserving ecological integrity.

Too much talk of "achieving" sustainable development has led to some disenchantment with the concept when it is not delivered, or even approximated. The real power of the sustainable development idea lies in its ability to change the way we think about intra- and intergenerational equity and our responsibility as custodians of biodiversity (3). It is much more helpful to view sustainability as a *relative*, dynamic state of development practice, one that can be degraded or improved substantially, but never reached absolutely. Sustainability becomes very useful when we work with stakeholders to explore development alternatives that are compared against the status quo, business-as-usual baseline state. So, one key criterion for comparison is the relative sustainability of an option compared to others, for example, one water supply alternative versus others. In practice, this criterion is broken down into social, economic, and ecological subcriteria, the comparative evaluation of alternatives is done by a multicriteria method such as environmental and social impacts assessment (ESIA).

Viewing development in terms of helping people help themselves, finding alternative policies, plans, programs, and actions, and comparing them for positive and negative impacts is a useful simplifying approach that cuts through skepticism about sustainability and its inherent complexity. This is exactly what lawmakers had in mind when they enacted the landmark U.S. National Environmental Policy Act (NEPA) in 1969, the law that gave birth to environmental impact assessment (EIA). But in few cases has compliance with this law led to more sustainable solutions, and the public is skeptical that it is merely a technical process that endorses preconceived actions (3).

The fact that Earth Summit 2002 in Johannesburg was oriented strongly toward understanding why the plans and lofty goals of Earth Summit 1992 have not been reached, even approximated,—prospects are worsening for sustainability according to global trends (4–6)—adds considerable impetus to our quest to understand what makes the sustainability concept operational. What are the underlying concepts and philosophies that describe

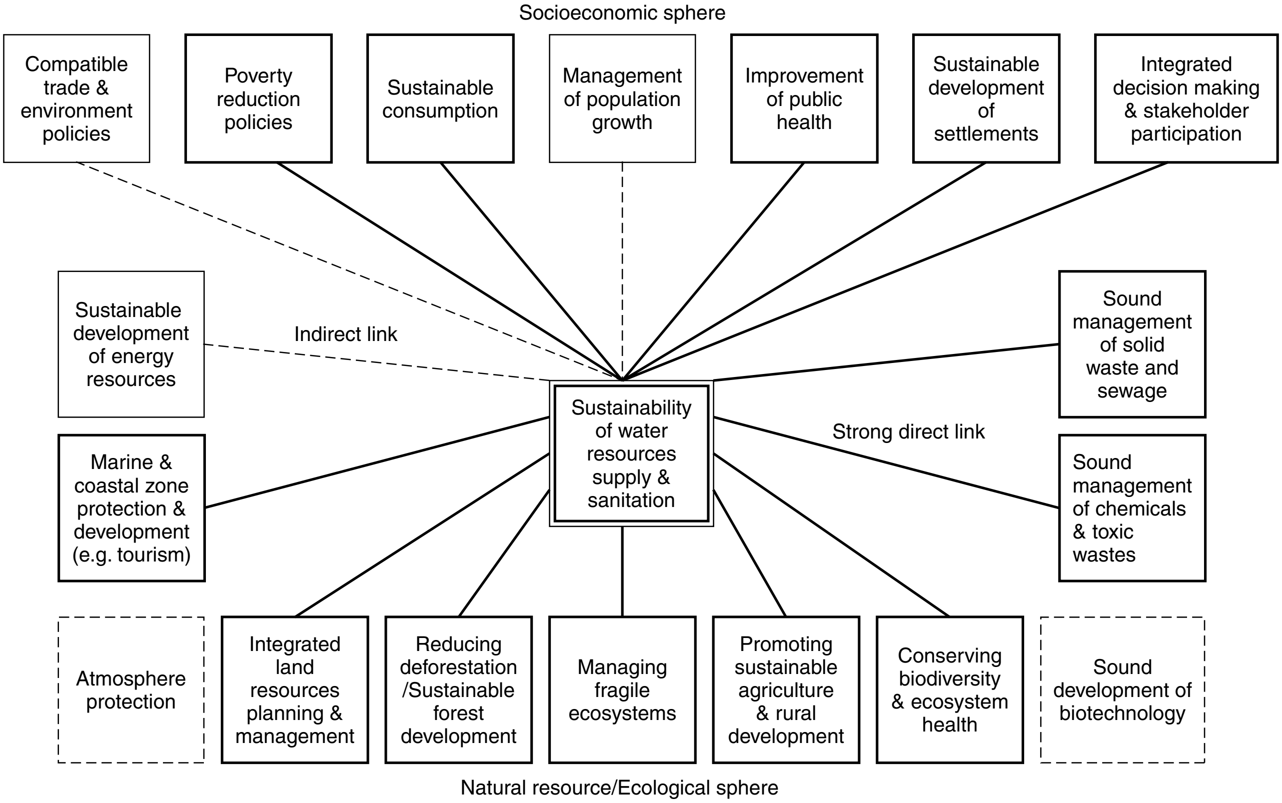

Figure 1. Agenda 21 components of sustainable development. Water resource sustainability influences many sectors and is a strategic operational focal point (7) (permission of University of Newcastle–upon-Tyne).

sustainability, and how can these be put into practice? Notably, freshwater resource sustainability is of great strategic importance to sustainable development as a whole because it has by far the highest degree of structural and functional interdependency with the other sectors/topics considered in *Agenda 21* (Fig. 1). This means, in theory, if we can solve water problems, many others will follow.

SUSTAINABILITY CONCEPTS

Human activities—agriculture, industry, urban development, and the like—stress environment and ecological health (including human health) in three main ways that often overlap and magnify each others' effects (3):

- through overexploitation of natural capital/resources (soils, water, plants, animals),
- through overproduction of wastes beyond the assimilative/attenuative capacity of ecosystems, and
- through land use/land cover changes that alter ecological structure and function.

The literature of the late 1980s and early 1990s is replete with discussions of sustainability as a concept, model, idea, paradigm, and philosophy. Hardin's seminal article called "The Tragedy of the Commons" in 1968 (8) was an important work that provoked discussion and raised awareness of the risk of overexploiting common natural resources without property rights. Hardin's theory held that individuals who have access to a common resource would seek to overexploit the resource to maximize their benefits and that the aggregate cost of overexploitation borne by the population as a whole would exceed the benefits. Such a scheme would be neither optimal for the society nor sustainable. The field of Common-Pool Resources (CPR) seeks to find governance solutions to sustainable natural resource management (9,10), investigating institutional predisposing conditions for successful governance of common natural resources, a partial refutation of Hardin's theory. Critics of CPR state that it is predicated on a single resource concept isolated from the ecosystem context of many resources, that studies are about self-governance that excludes the state, and that it is based on assumptions that resources and people do not change (11).

Another seminal publication was *The Limits to Growth* (12), which predicted the collapse of natural resources such as fossil fuel, metals, timber, and fish from overexploitation. This "crisis" scenario proved largely erroneous because of a gross underestimation of available reserves. Initially, it heightened concern over the sustainability of development, but then it undermined these same concerns by exaggerating the risks of collapse and losing legitimacy.

In 1987, the United Nations Commission on Environment and Development produced a report called *Our*

Common Future (1), widely known as the Bruntland Report after the chairperson, Norwegian Prime Minister Gro Harlem Bruntland. Focusing on satisfying basic human needs—water, food, shelter, energy, health—the Commission promoted seven strategic imperatives for sustainability and seven preconditions (Table 1). The report went on to suggest that the transition to sustainability will be driven by industrial wealth and that wealth creation be made more equitable and environmentally responsible. Its value has been to stimulate activity, though it has been criticized for failing to fully consider the dependence of industrial economic growth on natural resource exploitation (13).

The physical thermodynamic aspects of sustainability have been addressed by the Natural Step philosophy, set up in Sweden, based on a set of four systemic principles (14):

1. Substances from the earth's crust may not be extracted at a rate faster than their slow redeposit into the earth's crust.
2. Substances must not be produced by society faster than they can be broken down in nature or deposited into the earth's crust.
3. The physical basis for nature's productivity and diversity must not be allowed to deteriorate.
4. There must be fair and efficient use of energy and other resources to meet human needs.

Despite the obvious difficulty of meeting the first three principles, the World Business Council for Sustainable Development (WBCSD), a coalition of 125 companies worldwide, endorses *Natural Step*, as do conservation groups such as the International Union for the Conservation of Nature (IUCN) (14). The interests of the business sector are addressed by the concept of *triple bottom line*, the three dimensions being economic, social, and environmental. The main issues among these three are as follows (after Ref. 14):

- Economy–environment issues: How to promote ecoefficiency, clean production, and waste prevention and minimization? How to employ ecotaxation and create a market for tradable environmental permits? How to account for nonmarket costs and benefits?
- Society–environment issues: How to raise awareness and environmental literacy? How to counter environmental injustice and foster generational equity?
- Economy–society issues: How companies consider social costs and benefits of their investments? How wealth generation can address human rights and business adhere to an ethical code of practice.

Pearce and Turner (15) have defined sustainable development as "maximizing the net benefits of economic development subject to maintaining the services and quality of natural resources over time." They state several sustainability rules: (1) use renewable resources at rates less than or equal to the natural rate at which they can regenerate; (2) always keep waste flows to the environment at or below the assimilative capacity of the environment; and (3) optimize the efficiency with which nonrenewable resources are used, subject to substitutability between resources and technological progress. The first two are absolute physical thermodynamic sustainability criteria readily applicable to water quantity and water quality, whereas the third is an economic criterion. Examples of violations of rule 1 and 2 are the depletion of aquifers by withdrawing more water than is being recharged and the degradation of water quality, respectively.

Table 1. Sustainability Imperatives and Preconditions

Strategic Imperatives	Preconditions[a] (Capital Needs)
1. Economic growth must be revived in developing nations to alleviate poverty and reduce pressure on the environment.	1. A responsive political decision-making process (S).
2. Equity and nonmaterial value must be included in the consideration of growth.	2. More efficient economic systems that use less natural resources (N, S, H, F, P)
3. Basic human needs for food, water, shelter, energy must be met, accepting that changing (more equitable) patterns of consumption are needed.	3. Responsive social systems that redistribute the costs and benefits of development (S, H, F).
4. Reduce population growth by reducing economic pressures to have children.	4. Production processes that operate within ecological limits (N, S, H).
5. Conserve and enhance the natural resource base.	5. Technology development that supports efficient energy and resource solutions (N, S, H, F, P).
6. Environmental risk management technology must be developed and made available to the developing world.	6. International order that maintains cohesion globally (S).
7. Decision making should consider ecological and economic criteria.	7. Responsive, flexible, self-correcting government institutions (S, H).

[a] S: social capital; H: human capital; F: financial capital; P: physical capital; N: natural capital.

Source: After Reference 1.

Sustainability fundamentally challenges cultural norms, values, and behaviors. Using a historical cultural evolution argument, Downs (16) argues strongly that a combination of ethics, productive social interaction, and knowledge integration are cultural prerequisites for improving sustainability. Professionals and academics in the environmental field are now (and for the foreseeable future) centrally concerned with the pursuit of 'sustainable solutions' to priority problems of natural resource degradation. Reflecting the triple bottom line, sustainable solutions are those that are at the same time ecologically viable, economically feasible, and socially desirable (Fig. 2).

The International Union for the Conservation of Nature (IUCN) has actively promoted *green accounting*,

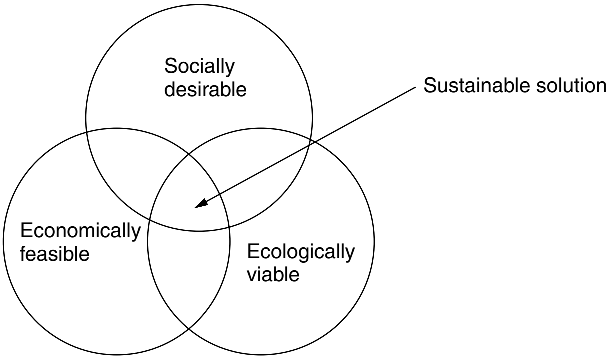

Figure 2. Sustainable solutions must satisfy three sets of criteria: social, ecological, and economic (17).

a process by which the economic costs and benefits of natural environment exploitation are measured (14). Norway has institutionalized this process to track the state of natural resource stocks and pollution rates. The United States has developed a supplement to normal economic accounting called the Integrated Economic and Environmental Satellite Account (IEESA) that seeks to account for the economic impact of environmental changes in water quantity, water quality, mineral stocks, forest stocks, and fish stocks (14). To do this will require new concepts, methods, and data, and impetus that can counter lobbying against such policies by special interests.

Spearheading what has been called the *Blue Revolution* for water sustainability (14) is the method of integrated watershed management (IWM)/integrated water resources management (IWRM). IWM/IWRM seeks to combine interests, priorities, and disciplines as a multistakeholder planning and management process for natural resources within the watershed ecosystem, centered on water quantity and water quality. Driven bottom-up by local needs and priorities and top-down by regulatory responsibilities, it must be adaptive, evolving dynamically as conditions change (3). Two global institutions play a key role in strengthening the social capital to sustain water resources: (1) the Global Water Partnership (GWP, founded 1996) that seeks to translate principles into practice through partnerships based on integrating knowledge and experience and (2) the World Water Council (WWC, founded 1996) that seeks to raise awareness of water problems and promote solutions (14).

IWM/IWRM is new and still evolving: “IWRM has neither been unambiguously defined nor has the question of how it is to be implemented been fully addressed. What has to be integrated and how is it best done? Can the broad principles of IWRM be operationalized in practice—and, if so, how?” (18). There are four broadly accepted IWRM principles, as set down by the 1992 *International Conference on Water and the Environment* in Dublin (19):

- Fresh water is a finite and vulnerable resource, essential to sustain life, development, and environment.
- Water development and management should be based on a participatory approach, involving users, planners, and policy-makers at all levels.
- Women play a central part in the provision, management, and safeguarding of water.
- Water has an economic value in all its competing uses and should be recognized as an economic good.

Building on these principles, water resource analysis, planning and management must be (after Ref. 3)

- holistic—incorporate the governing aspects of the “whole ecosystem” (watershed links to other watersheds, aquifer systems, bays, and estuaries), including interrelated natural resources (air, water, soil, biota), and humans as an integral, supermodifier of the ecosystem;
- integrated and participatory—incorporate relevant knowledge from the natural sciences, engineering, social sciences, and humanities, and accommodate the diverse set of interests, concerns, roles, and responsibilities of different stakeholders (including marginalized groups like periurban poor and rural subsistence farmers);
- strategic—identify and focus primarily on governing dynamics (the few drivers and responses that govern the way the system behaves), gather priority data, identify and address priority problems most cost-effectively;
- adaptive—adapt to changing geophysical and sociopolitical conditions; and
- sustainable—provide a steady stream of use benefits (or positive impacts) to present and future populations of humans and other species that exceed costs (negative impacts), while conserving ecosystem integrity and environmental quality.

Laws and regulations are crucial instruments that help us respond to the challenge of water sustainability. Such responsive legal instruments should do the following (after Ref. 20):

- Encourage administration at the appropriate hydrological scale: watershed, multiwatershed, or aquifer system.
- Foster internalization of the values and ethics of sustainable resource development.
- Encourage integrated approaches to water supply and sanitation, ecology, and public health.
- Prevent water allocation and usage policy making from fragmenting among agencies.
- Promote integrated appraisals, notably environmental, economic, and sociopolitical impacts assessment of alternative actions.
- Encourage integrated capacity strengthening of governmental institutions, NGOs, community associations, and businesses to transit into more sustainable policies and practices, and work collaboratively.
- Enforce reward and penalty incentives that encourage sustainability.

SUSTAINABILITY PRACTICES

The numerous rules and principles that populate the concepts of sustainability can be translated into indicators that can be measured quantitatively (e.g., waste flows are being assimilated/waste accumulation is zero) or qualitatively (e.g., the quality of stakeholder collaboration). These ecological, economic, and social sustainability indicators should be used to assess existing conditions and identify priorities for action, comparing action alternatives (impact assessment), and monitoring/evaluating the performance of the chosen alternative.

Earth Summit 2002 in Johannesburg tried to come to grips with ways to make the concept operational, how to put into practice the aspirations of Rio 1992. Howe (21) identified key water management requirements for a watershed/multiwatershed scale: (1) coordinated management of surface water and groundwater resources, (2) coordinated management of both water quantity and water quality, (3) provision of incentives for greater economic and technical efficiencies in water use, and (4) protection of public values associated with water service (e.g., reliable, safe, clean, affordable supply). And to achieve this, water institutions must develop the following characteristics (21):

- the capacity to coordinate water plans with other agencies (e.g., urban planning, agriculture, public health, environment, industrial);
- the capacity to solve water problems creatively using a variety of options and approaches (e.g., laws, pricing, taxes, tradable supply and/or pollution permits, subsidies);
- the foresight to separate roles and responsibilities for water resource planning and management activities from construction activities (i.e., avoid conflicts of interest);
- the multidisciplinary capacity to undertake multicriteria/multiobjective planning and evaluation of alternatives;
- devolve decision-making power to the lowest level—national, state/regional (provincial), local, or municipal—consistent with the scale of the water issue;
- the capacity and willingness to use appropriate participatory methods involving different stakeholders at different stages of a project (preplanning/conceptual, planning and design, implementation, maintenance, monitoring, and evaluation).
- the ability to reward innovation and adapt to changing conditions and priorities.

It is clear that to do most of what sustainability challenges society to do requires us to strengthen our *capacity to respond* to those challenges and opportunities. “Capacity building is the sum of efforts needed to develop, enhance and utilize the skills of people and institutions to follow a path of sustainable development” (22). A UNDP program *Capacity 21* (22), seeks to build capacity to implement *Agenda 21*.

Considerable community-based water management experience worldwide provides empirical evidence of what works on the community scale and why (see 23–26 among many others). IRC (23) found that the following are major factors for success in community water projects: baseline community capacity, community demand, donor and government support (financial and as policy), sufficient water resources, and the capacity of implementation agencies. But work has tended to focus on one population context alone (e.g., rural subsistence), and on one aspect (water supply and/or sanitation, or irrigation, or land degradation), instead of the more integrated watershed management approach advocated for sustainability. A considerable body of knowledge exists on participatory methods, methods include participatory rural appraisal (PRA), participatory action research (PAR), rapid rural appraisal (RRA), and participatory action development (PAD); all have been applied to strengthening community water management. But calls have been made to move beyond traditional participatory methods to more integrated analysis and planning methods with community and government stakeholders (27).

How do we make sustainability progress in a field that builds on successes, learns from failures? From 1998–2000, local working groups in collaboration with the Mexican National Water Commission (CNA), coordinated by Downs (7), developed a participatory integrated capacity-building (PICB) approach to water sector sustainability. Following an analysis comparing relatively sustainable development projects worldwide during the past 10–15 years (those yielding a steady stream of benefits *after* external support was removed) with a much larger number of *un*sustainable projects, six broad synergistic levels of capacity building emerged as critical components for success: (1) strengthening political and financial commitment; (2) strengthening human resources, including education, training, and awareness-raising; (3) strengthening information resources for policymaking (e.g., monitoring and GIS tools for data integration); (4) strengthening policies, regulations, enforcement, and verification; (5) applying appropriate technology and basic infrastructure (e.g., for water and wastewater treatment); and (6) stimulating local enterprise development (i.e., support products and services providing socioeconomic sustainability). Each one builds on those before it with positive feedback. Crucially, we see that operational sustainability is a function of participation and integrated capacity building (7).

Figure 3 shows a desirable participatory project sequence. On the analytical (science) side, we need strategic information on watershed conditions, especially water quantity, water quality, and use requirements (present and future users are humans and other species). On the policy and management sides, we need strategic planning methods that accommodate stakeholder inputs, evaluate options, choose a preferred option (by multiple criteria), and support the implementation and sustainable operation of this option. PICB is used to make the selected preferred management option (PMO) sustainable (for more

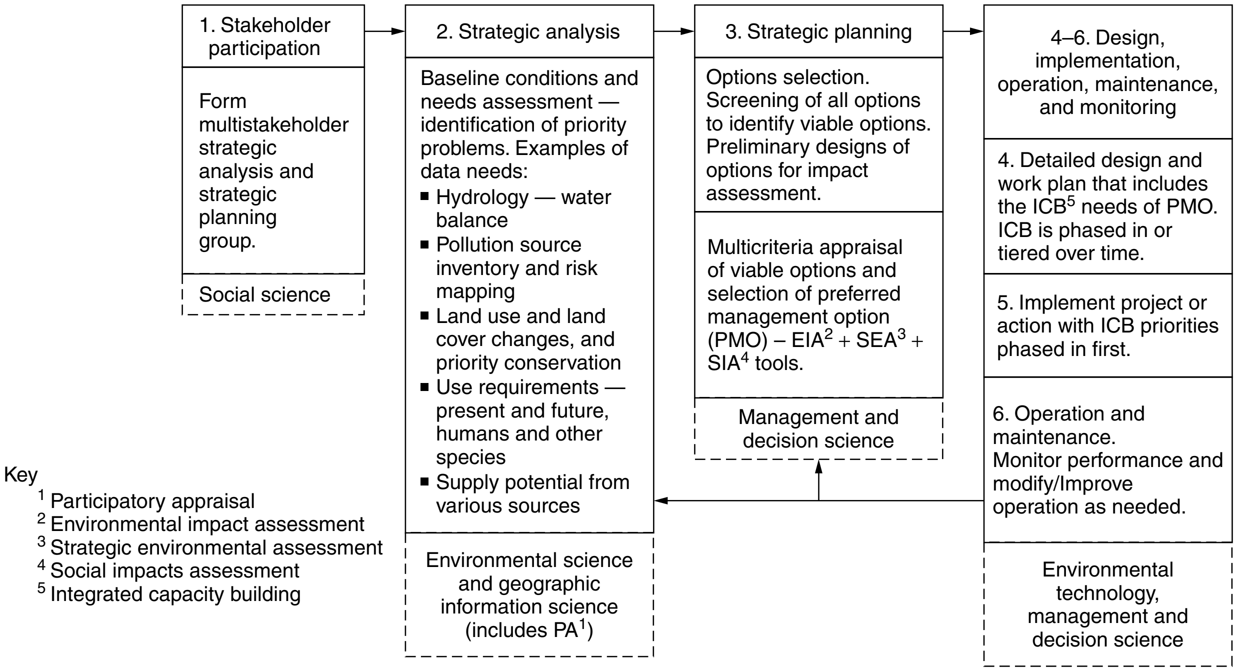

Figure 3. Six stages of operational water sustainability. Strategic participatory planning selects an alternative using sustainability criteria (impact assessment), and integrated capacity building sustains it (after Ref. 3).

information, see the article INTEGRATED CAPACITY BUILDING NEEDS FOR WATER SUPPLY AND WASTEWATER SANITATION).

An integrated, sustainable watershed management plan must be based on reliable information (PICB Level 3) about baseline geophysical, hydrologic, water quality, pollution source, ecological and socioeconomic conditions. Identifying key indicators, integrating data on them, interpreting and disseminating information to watershed–bay stakeholders informs policy and raises awareness of priority issues. It also facilitates collaborative actions to mitigate those issues. This strategic data can be fed into a GIS database that can be manipulated for decision-making. The GIS format can be used to engage community, NGO and government stakeholders in the *strategic analysis* process that must underpin the watershed–bay planning, design, implementation, and maintenance/monitoring stages (Fig. 3). A web-accessible GIS can also provide actors with a powerful means to understand how their interests fit into the wider multi-interest context of a watershed ecosystem. Such visualization and information is conducive to multistakeholder participation and dispute mediation: Actors internalize the stakes of others and can better negotiate compromises to meet common goals (3). Such information resource/planning tools do not yet exist, data are scattered, of varied formats, and have significant gaps that flag primary data gathering needs. Close collaboration with watershed NGOs engaging in community capacity building and state and local environmental agencies will ensure that results have the greatest impact and utility (3).

The term *strategic* is stressed: We must choose to collect data of the highest information value for our objectives and only this data (3). In this way, what appears to be a huge project scope is reduced to its critical components, is manageable, and most importantly, effective. This also keeps volunteer data gatherers engaged because they see the impacts of their efforts on policy. We already know what most of these key data are (priority pollutants, use data, hydrologic balance, etc.); often some exists, though dispersed and must be integrated. Other priority data needs emerge as part of the analysis.

Empirical evidence points strongly to community-based natural resource management as a key to social sustainability. But how can this be done on the watershed scale required for hydrologic sustainability? The answer lies in strengthening the social capital of the watershed by forming collaborative networks of communities. Often in developing countries, marginalized rural agroforesters occupy the upper reaches of the watershed and become, *de facto*, producers of water resources for downstream users. Evidently, both producers and users share the common interest of watershed sustainability, and both must be involved in strategic planning. By networking and through capacity building coordinated by NGOs, marginalized communities can gain the power they need to negotiate with influential user groups and government agencies.

CLOSING REMARKS

Integrated watershed management and integrated capacity building seem to confirm that the challenges of sustainability are primarily cultural and sociopolitical, not economic, scientific, or technical. Echoing the cultural prerequisites of Downs (16), any sustainability process, method, or approach—certainly IWM, PICB, EIA—will ultimately depend for its success on two cultural determinants:

- *an ethical core*—values and attitudes respectful of intra- and intergenerational equity and natural resource conservation; and
- *a participatory culture*—swapping "win–lose" for "win–win"; choosing the philosophy of collaboration and mutual gains over conflict negotiation and trade-offs.

Providentially, international case studies reveal that water sustainability challenges and opportunities are similar across cultures and geographical contexts and yield significant economies of scale in the development of theory and practice.

BIBLIOGRAPHY

1. WCED (World Commission on Environment and Development). (1987). *Our Common Future*, Oxford University Press, New York.
2. UNCED (United Nations Conference on Environment and Development). (1992). *Earth Summit 1992—the United Nations Conference on Environment and Development*, The Regency Press, London.
3. Downs, T.J. (2002). A participatory integrated capacity building approach to the theory and practice of sustainability—Mexico and New England Watershed Case Studies. In: *International Experiences on Sustainability*. W.L. Filho (Ed.). Peter Lang, Frankfurt am Main, pp. 179–205.
4. UN. Earth Summit 2002 Briefing Documents, available at: http://www.earthsummit2002.org/, consulted 2002.
5. UN. Earth Summit +5: Programme of action adopted by the Assembly, Special Session of the Assembly to review and appraise the implementation of Agenda 21, UN Department of Economic and Social Affairs, New York, available at: http://www.un.org/esa/earsummit/, authored 1997, consulted 2000.
6. World Watch. (2002). *State of the World 2002*. WW Norton, New York.
7. Downs, T.J. (2001). Making sustainable development operational: Capacity building for the water supply and sanitation sector in Mexico. *J. Environ. Plann. Manage.* **44**(4): 525–544.
8. Hardin, G. (1968). The tragedy of the commons. *Science* **162**: 1243–1248.
9. Ostrom, E. (1999). Coping with the tragedy of the commons. *Annu. Rev. Political Sci.* **2**: 493–535.
10. Dietz, T. et al. (2002). The drama of the commons. In: *The Drama of the Commons*. E. Ostrum (Ed.). National Academy Press, Washington, DC.
11. Steins, N., Roling, N., and Edwards, V. (2000). Redesigning the principles: an interactive perspective to CPR theory. *Proc., 8th Conf. Int. Assoc. Study Common Property*, Bloomington, IN, June 1–4.
12. Meadows, D.H., Meadow, D.L., Randers, J., and Behrens, W.W. (1972). *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*. Universe Books, New York.
13. Clayton, M.H. and Radcliffe, J. (1996). *Sustainability, a Systems Approach*. Earthscan, London, p. 65.
14. Calder, I.R. (1999). *The Blue Revolution, Land Use and Integrated Water Resources Management*. Earthscan, London, pp. 62–71.
15. Pearce, D. and Turner, R.K. (1990). *Economics of Natural Resources and the Environment*, Johns Hopkins University Press, Baltimore, pp. 24–44.
16. Downs, T.J. (2001). Changing the culture of underdevelopment and unsustainability. *J. Environ. Plann. Manage.* **43**(5): 601–621.
17. Doyle, K. (2002). *Environmental Career Trends*. Environmental Careers Organization, Boston.
18. GWP (Global Water Partnership). (2000). *Framework for Action: Responding to the Forum*. Report reflecting on the Framework for Action presented at the Second World Water Forum, Stockholm, Sweden, March.
19. GWP/TAC (Global Water Partnership/Technical Advisory Committee). (1999). *The Dublin Principles for Water*. TAC Background Papers, No. 3, Global Water Partnership, Stockholm, Sweden.
20. Smith, D. and Rast, W. (1998). Environmentally sustainable management and use of internationally shared freshwater resources. In: *Watershed Management—Practice, Policies and Coordination*. R.J. Reimold (Ed.). McGraw-Hill, New York, p. 294.
21. Howe, C. (1995). *Guidelines for the Design of Effective Water Management Institutions Utilizing Economic Instruments*. Report presented at Workshop on the use of economic principles for the integrated management of freshwater resources, United Nations Environment Programme (UNEP), Nairobi, Kenya.
22. UNDP (United Nations Development Programme), About Capacity 21, Available at: www.sdnp.undp.org/c21, consulted 2001.
23. IRC. (2001). *Community Management: The Way Forward*. Report of the workshop, November 19–27, Rockanje, Netherlands.
24. IRC. (2001). *From System to Service—Scaling-up Community Management*. Report of the conference, The Hague, Netherlands, December 12–13.
25. World Bank. (1994). *World Development Report: Infrastructure for Development*, Oxford University Press, Oxford, UK.
26. WSSCC. (2000). *Vision 21: Water for People—a Shared Vision for Hygiene, Sanitation and Water Supply*. Water Supply and Sanitation Collaborative Council, Geneva, Switzerland.
27. Hailey, J. *Beyond the PRA Formula*, Paper presented at the *Symposium* on *Participation: the New Tyranny?* Institute for Development Policy and Management, Manchester, UK, November 3, 1998.

THE PROVISION OF DRINKING WATER AND SANITATION IN DEVELOPING COUNTRIES

Terence R. Lee
Santiago, Chile

Few die from a lack of water, but many die from water. Despite all the resolutions made at international congresses and all the pledges from governments and international organizations, more than 3 million people die each year from water-related diseases (Fig. 1). These deaths occur in circumstances where safe water and sanitation would have cut by at least one-third the number of diarrhea cases (currently at 4 billion worldwide every year) which result in 2.2 million of those deaths (1).

The U.S. National Academy of Engineering includes drinking water supply as one of the great engineering achievements of the twentieth century. For those fortunate to live in a developed country, it was. However, the failure to provide universal safe drinking water and adequate sanitation is amongst the gravest contradictions of our world at the beginning of the twenty-first century. The necessary technology has existed for at least two millennia, but there remain serious obstacles to its application. The result is that at least 1 billion people remain without access to a reliable and safe source of water, and more than 2.5 billion lack safe sanitation. This population is concentrated in countries at the lowest levels of development.

The deficits in services are greatest in Africa where 38% of its population remains without safe water and 40% without sanitation; Asia has 19% of its population without access to safe water, 52% without adequate sanitation; and in Latin America and the Caribbean, 15% of the population is without a potable water supply and 22% without adequate sanitation. Although, it has been estimated that a large number of additional people obtained access to services in the 1990s (800 million to water and 750 million to sanitation), population migrations and growth have meant that the proportion of the urban population with access to a safe water supply actually decreased, whereas the absolute number of people without access to water and sanitation remained the same (3).

An example of what these statistics can mean in reality for the population of a city can be illustrated by the situation in Dakar, the capital of Senegal in West Africa. In Dakar, the urban area can be subdivided into three zones:

- The first zone is completely laid out in lots and endowed with water supply and sanitation. It corresponds to the area of the colonial city founded in the second half of the nineteenth century. Today, it is the commercial and administrative center.
- The second zone is the formal suburban area, which is partly parceled in legally constituted lots. However, only part of this area has public water supply and sanitation. In the rest, where the majority of houses have been built on public land without permission, the inhabitants eliminate domestic wastes *in situ*. Of the inhabitants, 40% use latrines, 40% use abandoned wells, and 20% use dug holes.
- The third zone is entirely without any public water supply or sanitation (4).

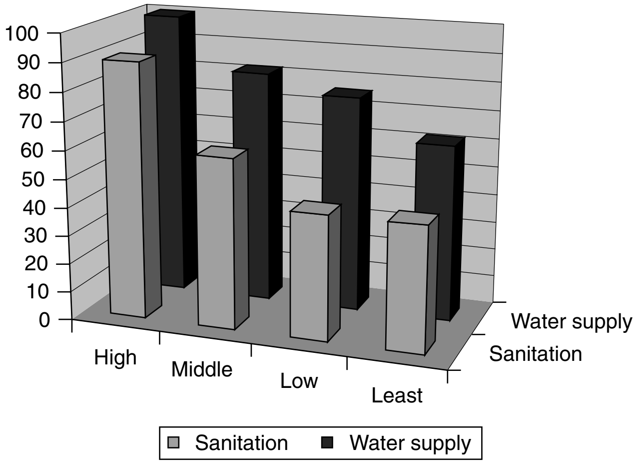

Figure 1. Percentage of population with access to improved water source and sanitation, by level of development. (*Source*: Reference 2.)

In the suburbs of Dakar, only 40% of the population receives drinking water from a protected source, usually public taps; the remainder of the population relies on water from shallow wells, running the risk of contracting diseases as the groundwater quality is menaced through the existence of poor quality latrines and of excreta deposited directly on the soil.

The menace of well contamination is present even in rural areas that have much lower population densities than those of the suburbs of Dakar. The problem is the reliance on shallow wells when excreta are also disposed of into the shallow aquifer. Problems due to contamination are compounded by the restrictions on water consumption imposed by the need to carry water over long distances (Table 1). People will go a very long distance to get water, even more than 2000 meters, but consumption drops drastically when water has to be carried so far. It has been estimated by the World Health Organisation that something of the order of 50 liters per person a day is needed to live safe from disease, which, the data in the table suggest, means a tap in the dwelling.

For the rural population, water may be free, even if not close by. In urban areas, even for low levels of supply, water has to be paid for. It has long been known that the poor pay more when the source is not a public utility (Table 2).

Rural water supply service can be extended through the successful change to a demand-led provision of both technologies for the individual household (hand dug wells, hand pumps, roof catchments) and community technologies (gravity flow and simple pumped systems manageable by community members). The main issue is to overcome the operations and maintenance problems, which, in the past, have made up to one-third of even such

Table 1. Water Consumption in Rural Areas

Type of Water Supply	Average Daily Consumption, Liters per Person
Communal source	
Well or tap farther than 1000 meters	7
Well or tap between 250 and 1000 meters	12
Well less than 250 meters	20
Tap less than 250 meters	30
Tap in yard	40
Tap in house	
Single tap	50
Multiple taps	150

Source: Reference 5.

Table 2. Ratio Between the Price Charged by Water Vendor and by Public Utilities, Selected Cities

City	Ratio
Dacca, Bangladesh	12–25
Cali, Colombia	10
Guayaquil, Ecuador	20
Port-au-Prince, Haiti	17–100
Tegulcigalpa, Honduras	16–34
Jakarta, Indonesia	4–60
Adidjan, Ivory Coast	5
Nairobi, Kenya	7–11
Lagos, Nigeria	4–10
Karachi, Pakistan	28–83
Lima, Peru	17
Istanbul, Turkey	10
Kampala, Uganda	4–9

Source: Reference 6.

relatively low technology systems out of order soon after installation (7). Any improvement will require continuing support once the systems are installed. It is now accepted that increased private sector involvement and the transfer of many water points from community to private household management can improve performance. For example, a World Bank survey in the 1980s showed that 60% of the public hand pumps in rural India were out of order, compared to only 10% of privately owned pumps (8).

Rural sanitation normally takes the form of on-plot sanitation, which does not impose high external investment costs or significant operation and maintenance responsibilities outside the household. Public support is required for the critical tasks of health and sanitation education and promotion, and there is a need for effective provision of small-scale credits to allow the purchase of components, such as latrine slabs. Some of these activities could be successfully delegated to the private sector, as has been done with both rural water supply and sanitation installation and maintenance support in Chile (9).

In urban areas, self-help solutions have limited application. Urban water supply systems, even where they exist, commonly do not cover the entire population, have not expanded to keep pace with population growth, and fail to provide either adequate pressure or a continuous supply. The shortfall in coverage always means the absence of service for lower income groups. In addition to the failure of systems to expand or provide continuous supply, demand for water grows as industrial and commercial demands rise with economic growth. At the same time, increased household demand for drinking water is in part driven by higher living standards resulting from economic growth and by the use of conventional waterborne sewerage, which significantly increases per capita water use.

Because of economies of scale and the low value of the product supplied, urban water supply has been considered a 'natural' monopoly. It is certainly not practicable or economical to have direct competition in the provision of network services at the household level. Traditionally, providing water supply and sewerage services has been a local responsibility where services are overwhelmingly run through municipal governments. One consequence of local municipal ownership and management of services has been the creation of serious difficulties for the transfer of successful experience. In the nineteenth and early twentieth centuries, technological and administrative transfers, even if restricted, did occur through colonial governments. Now, transfer has been largely left to international aid organizations which, whether governmental or nongovernmental, have failed to provide continuity in commitment over the long term.

Even where historically services have been provided by private companies, these too have tended to be small and local. Water supply and sewerage services have largely resisted globalization and even today foreign operation and investment in services is very limited. There has been, therefore, little transfer of the ways and means of providing service from the successful to the rest of the world.

Where transfers have occurred, advances and innovation in water supply and sewerage have been related mainly to the technical aspects of the operation and administration of large organized urban systems. There has been little or no innovation for the provision of services to the informal suburbs of the large cities in developing countries, and this is why the use of the shallow well for water supply and the latrine for excreta disposal remain the methods of choice.

Moreover, it has been the conventional wisdom that a large and relatively sophisticated institution is required to plan, organize, and control this monopoly provision. However, there is considerable potential for private sector involvement in managing, or supporting the management of, all or parts of a system, as well as in contributing to the capital requirements in middle-income metropolitan areas.

Dar-es-Salaam, Tanzania, presents an example of one possible alternative approach. Residents of Dar-es-Salaam face chronic water shortages and often have to combine several sources of water at different costs and quality to fill their needs. The Dar-es-Salaam Water and Sewerage Authority provides most of the bulk water supplies to the city. The Authority is relatively efficient; some 80% of the water produced is billed, although only 50% of the bills is collected. However, only one-third of households receive their water directly from the utility, creating a large niche for the private sector in water distribution. Large parts of the distribution system—both the piped system and nonpiped supply—are in private hands (10). This "privatization by default" is manifested in two ways:

- through the private redistribution of water, that is, water vending, ranging from home-based, quasi-legal resellers and informal low-income pushcart vendors to more formalized and regulated distributors delivering water to people's homes by tank truck; and
- by the "spaghettization" of the piped water network, where privately financed individual service lines are, in practice, directing investments and expansion of the distribution system.

Urban sanitation needs may be met through on-plot systems, communal facilities, or traditional sewerage systems. Privatization by default of network sewerage services is rarely, if ever, found. On-plot systems involve households in providing their own facilities through different types of latrines, but contamination of the groundwater results in densely populated communities. Communal facilities economize on investment and water but may not be acceptable in all cultures, however. Alternative systems to traditional sewerage systems can reduce costs and could be provided communally. The provision of simplified sewers is one alternative. Simplified sewerage is essentially conventional sewerage without any of its conservative design requirements. It can be considered as the latter stripped down to its hydraulic basics. It is also called shallow sewerage, and its in-block variant is often called backyard or condominial sewerage. Under this system, the sewers are laid on private rather than public land, household labor is accepted as partial payment for the connection, and smaller bore sewers are used. It is estimated that the costs of service delivery can be reduced by 30% using this system.

The provision of sewerage through conventional waterborne sewers has considerable economies of scale leading to monopolistic provision and needs complex institutional support. However, smaller scale provision is possible with simplified sewerage systems. This means that the private sector and local communities could become more involved in supplying and servicing such systems as well as in providing on-plot sanitation, where this is possible, and also in providing and managing communal sanitation facilities.

The provision of services to the rural population, or even to less dense urban areas, should not and does not require such high levels of investment, and the same technology can be applied that has existed for milennia. However, a serious obstacle to improving water supply and sanitation in large third-world cities is the cost of modern centralized water supply and waterborne sewage systems. These require a capital investment of, at least, US$ 1000 per connection even without sewage treatment. It is in the cities that innovation is required and alternative systems must be sought.

BIBLIOGRAPHY

1. World Health Organization (WHO), UNICEF and the Water Supply and Sanitation Collaborative Council. (2001). Global Water Supply and Sanitation Assessment 2000 Report. Available at: www.who.int/water-sanitation-health/Globassesment/GlasspdfToC.htm.
2. United Nations Development Programme. (2003). *Human Development Report*, New York.
3. Report of the World Panel on Financing Water Infrastructure. (2003). *Financing Water for All*. Presentation to the 3rd World Water Forum, Kyoto, Japan.
4. Tandia, A.A., Gaye, C.B., and Faye, A. (1998). *Origin, Process and Migration of Nitrate Compounds in the Aquifer of Dakar Region, Senegal*. *IAEA-TECDOC-1046*, Application of isotope techniques to investigate groundwater pollution, Joanneum Research, Vienna, pp. 67–80.
5. International Referral Centre. (1981). *Small Community Water Supplies*. International Reference Centre for Community Water Supply and Sanitation, Technical Paper No.18, The Hague.
6. Bhatia, R. and Falkenmark, M. (1992). *Water Resource Policies and the Urban Poor: Innovative Approaches and Policy Imperatives*. Background paper ICWE, Dublin.
7. Franceys, R. (1997). *Private Waters?—A Bias Towards the Poor*. Department for International Development, Occasional paper No. 3, London.
8. Cairncross, S. (1989). Water supply and sanitation: An agenda for research. *J. Trop. Med. Hyg.* **92**: 571–577.
9. Chile, Ministerio de Obras Públicas. (1999). *Memoria*, Santiago, Chile.
10. Kjellén, M. (2001). *Water Provisioning in Dar-es-Salaam, Tanzania: The Public—Private Interface.* Paper presented to the UNESCO symposium: Frontiers in Urban Water Management, Marseilles, France.

SUSTAINABLE MANAGEMENT OF NATURAL RESOURCES

Lucas Reijnders
University of Amsterdam
Amsterdam, The Netherlands

INTRODUCTION

It is easier to agree on what is *un*sustainable than on what is sustainable. The once great Newfoundland cod fishery collapsed in 1992, causing 20,000 people to lose their jobs (1). It is uncontroversial to call the fishing practices that led to this collapse unsustainable. Similarly, it is uncontroversial to call mining of fossil water for corn production in Saudi Arabia unsustainable. However, by now there are probably more than 100 definitions of sustainability. Such definitions include environmental and often also social and economic aspects of human activities. They may either stress matters that can in principle be objectively established (such as fish stocks) or be social constructivist, emphasizing divergent views of stakeholders, different world views, or procedures to come to agreement about what sustainability means in a specific social context.

Here sustainability will be defined according to the World Conservation Strategy (2) that introduced the concept among policy makers. This definition fits a steady-state economy (3)—an equilibrium relation between human activities and the physical environment. The definition is also in line with long-standing approaches to safeguard the long-term productivity of forestry and fisheries (4). The main underlying principle of so defined sustainability is that of justice between the generations. This principle comes back in the most famous of later definitions of sustainability, given in the report *Our Common Future* (5): "meeting the needs of the present without compromising the ability of future generations to meet their needs."

"Natural resources" is a concept less controversial than sustainability. It includes all that we use or take from nature to provide for our existence. Mineral oil, river water, North Atlantic cod, and iron ore all belong to our natural resources. Efforts are needed to obtain these resources. They can be considered natural capital and are not technologically replaceable by other forms of capital, such as social or financial capital (6). There is also a category of services of nature that we use, whether or not we are aware of it. These include degradation of toxic substances, maintenance of atmospheric composition, sequestration and cycling of minerals, and structuring services. Many of these services are ecosystem services. There is no reason to believe that financial or social capital can be an effective substitute for nature in producing ecosystem services (7).

In this article, sustainable management of natural resources comes down to maintaining natural capital so that it can be used by all future generations of humans. This is based on a production level that can be sustained indefinitely without threatening the living conditions of future generations (8).

First, sustainability is discussed in the context of four different types of natural resources. Thereafter, three aspects of sustainability are discussed in more detail:

- sustainable management of fresh water resources,
- sustainable management of fisheries, and
- sustainable management of ecosystems.

TYPES OF NATURAL RESOURCES AND THEIR SUSTAINABLE USE

In the context of resource management, several important types of natural resources can be distinguished. First, there are flux-type natural resources. These can be used without any impact on future availability. Examples of such flux-type natural resources are wind and solar radiation. Second, there are renewable natural resources. In this case, there are currently substantial additions to the stock. Such resources include freshwater in rivers and fish in the seas. Aquifers that are replenished, mostly at a rate of 0.1–0.3%/year (9), also belong to the category of renewable resources.

A third type of natural resource is the product of slow geological processes. This is the category of the (virtually) nonrenewables and includes metal ores, fossil water, and fossil carbon compounds. Compared with the total stock, there are no or relatively small additions to these resources. Total stocks may vary strongly. For instance, stocks of iron ore are large, but the stocks of, for instance, Ag (silver), Sn (tin), or Pt (platinum) ores and fossil water are relatively small. A fourth type of resource covers living nature, which provides for ecosystem services.

From the viewpoint of sustainable resource management, flux-type resources may be freely used. This is not so for the other three types of resources.

To conserve renewables over long periods of time, the usage of renewables should not exceed the addition to stock. Moreover, the quality of the stock should be maintained. For instance, as to groundwater, the condition should be met that water tables do not fall, and also pollution that negatively affects the future value of the resource should be prevented.

The concern for (virtually) nonrenewable natural resources is mainly with geochemically scarce natural resources that are currently being depleted rapidly. In this category belong fossil water resources in North Africa and Southwest Asia that are currently "mined" (10). Fossil carbon compounds and ores for phosphates and metals such as silver, platina, tin, and nickel are other examples. If the stocks that are formed by slow geological processes are to last during the normal lifetime of a primate species ($\sim 10^6$ years), losses from stocks (including stocks in the economy) due to dissipation should be kept small. When substitutes that are dependent on flux-type or renewable resources are made available in a sustainable way, relatively large reductions of the stock that was formed by geological processes are defensible (11).

In ecosystem functions that are important to us, both the area available to natural ecosystems and species composition ("biodiversity") of ecosystems matter. Moreover, for proper functioning, a number of conditions regarding the abiotic aspects of the ecosystem should be met. Requirements as to the quantity and quality of water available function, often prominently, among such conditions (12).

SUSTAINABLE MANAGEMENT OF FRESHWATER RESOURCES

Freshwater is a vital natural resource for humankind. Only 3% of the world's water is fresh, and of that over 98% occurs as groundwater and less than 2% in surface water such as rivers and lakes. Desalinized seawater is used to a very limited extent as a resource. This is closely related to its high price (9).

To meet the criterion of sustainability, as defined here, the usage of renewable freshwater resources should not exceed additions to stocks. Withdrawal of seawater is by itself insignificant in view of sustainability, but desalinization should be powered by energy from flux-type resources (wind solar) or by sustainable biomass chains. The use of fossil freshwater stocks may be considered to bridge gaps temporarily on the way to sustainability.

Traditionally, management of freshwater resources has been strongly focused on the supply side: trying to supply adequate amounts of good-quality water and limiting the negative impacts of excessive water. Interest in sustainable management has shifted the focus. Demand-side management, reuse of water, and water conservation have become important. Though there is substantial scope for improvement at current price levels, it has rightly been argued that sustainable management of water resources is dependent on major changes in economic arrangements for dealing with water resources. A number of proposals for such changes have focused on rationing, where the sum of rations reflects sustainable management. Most proposals, however, argue in favor of water pricing that reflects the true full cost, including external costs.

In demand-side management, providing for freshwater-based services with less water ("improving efficiency") matters. As agriculture is the most important consumer of freshwater, requiring an estimated 65–75% of all water use, agriculture is a useful primary focus for efforts to improve water efficiency and reuse (9,12,13). This is especially so in semiarid areas that are currently "water starved" or where current practices cannot be sustained. Moreover, the share of agriculture in water withdrawals in such areas is often >85% (14). Water losses from agricultural systems in semiarid areas due to leakage during storage and conveyance (in case of irrigation), and to runoff, drainage, and evaporation are usually of the order of 70–85%. In some parts of sub-Saharan Africa, the loss for rain-fed agriculture may be 95% (13). Though a substantial part of these losses is unavoidable, there is also scope for improvement. For instance, when current water efficiency in sub-Saharan Africa is 5%, an increase to 10% (doubling vegetation yield) can often be achieved by the introduction of dwarf shrubs and improved tillage and nutrient management (13).

In irrigated agriculture, now responsible for about one-third of worldwide food and fiber harvests, better efficiencies are also possible. Options for efficiency improvement include better provisions for storage and conveyance, better scheduling, furrow diking, direct seeding, land leveling, microirrigation systems (drip irrigation and microsprinklers), and better regulation of the groundwater table (12,13).

Reuse of water in agriculture is a way to limit water withdrawals. In practice, both water from wastewater treatment plants and drainage water are reused (10,13–16). In the latter case, such water is often used to irrigate relatively salt-resistant crops. In managing reuse, the quality of water is important. The salinity and sodicity of water may negatively affect the future productivity of agricultural land. Excessive levels of nitrate, pesticides, and the presence of agents causative of infectious diseases may pose a direct threat to animal and human health. Receiving soils may, furthermore, act as a "sink" for hazardous compounds that may subsequently enter human and animal food chains. When persistent hazardous pollutants enter deep groundwater, they can result in long-term contamination of this important freshwater stock. Negative aspects of the secondary water supply may be influenced by pollution prevention ("source reduction") and treatment. Often there are possibilities for source reduction that are profitable at current prices (18).

Though both improvements in agricultural water efficiency and reuse offer scope for a more sustainable management of resources, it should be realized that in an increasing number of countries, especially due to increasing population pressure, such improvements will not be sufficient to feed the population in a sustainable way. This leads to the need for such countries to import "water-intensive" food from countries that have no water shortages.

Improvements in efficiency and reuse are also possible outside agriculture. An obvious object for improved efficiency is excessive loss and illegal withdrawal from distribution systems. In industrialized countries, losses of the order of 6% to more than 25% are common (17,19). Higher percentages have been noted elsewhere. In Jordanian cities, for instance, 55% and more of the input in distribution systems is "unaccounted for water" (10).

Water-efficiency improvement in industry and households has advanced. In industry, much of the improvement has come from internal reuse and process optimization, and there is still scope for further reductions in industrialized countries. For instance, a detailed study in the United Kingdom (20) suggested that industrial water consumption can be cut by 30% in a profitable way at current price levels. When water prices for industry increase to reflect true costs, profitable efficiency gains may increase considerably (17,21). Studies pertinent to developing and emerging industrial countries also tend to show a large potential for profitable water efficiency improvements (22–24).

Substantial improvements are underway in household water efficiency. Water use by water closets has been reduced by 75% or more, and major improvements in water efficiency have also been achieved in showers and appliances such as (dish and clothes) washing machines. Such efficiency improvements are usually profitable if current lifetime costs of appliances are considered (12,25). Internal reuse of household water, for instance, by using gray water for flushing water closets is practiced to a limited extent.

Finally, stocks of water may be improved by better water conservation. An eye-catching example of water conservation is in Sun Valley in the Los Angeles basin (USA). This valley was beset by floods from periodic winter rains. This led to regular press coverage showing vehicles driving in heavy (rain) water. To solve this problem, the County's Public Works Department initially opted for a storm drain to carry surplus water to the ocean. However, public discussion led to another solution. Water from winter rain is now increasingly collected in Sun Valley to recharge groundwater. In Europe, new urban developments increasingly rely on keeping rainwater out of the sewer systems and allowing it to infiltrate to add to groundwater stocks.

SUSTAINABLE MANAGEMENT OF FISHERIES

Fisheries catch renewable natural resources. When catches systematically exceed additions to stock, such resources can be depleted. If catches that exceed additions to stock are stopped, in a number of cases, the population rebounds. However, there are also cases where stocks do not rebound and remain at low equilibrium populations.

To be sustainable, catches should not exceed additions to stock in quality and quantity. Regarding cod, for instance, it is not only important to maintain a constant number of fish but also to safeguard that such fish are in the right age and size classes to maintain predation and reproduction (8). Moreover, for economic, biological, and social reasons, fisheries should manage stocks toward the higher end of abundance (1).

Sustainable management of fisheries is a long-standing issue. At least since the thirteenth century, European kings and other people in high places are on record opposing unsustainable fishing in their realms. The earliest recorded government-sponsored efforts to manage fisheries in a sustainable way in Europe date from the same time (26).

There is evidence that a number of traditional management systems have operated sustainably while managing stocks toward the higher end of abundance. Such systems have usually been characterized by ownership and controlled access; closed seasons during which fishing was prohibited; a taboo on fishing in refuge areas; and gear, size, and species restrictions (27,28). By present standards, fishing effort has been low under such traditional arrangements. When efforts increase (usually linked to financial pressures, population pressure, and technological development), stocks tend to decrease—especially when there is open access. The unregulated dynamic of fisheries is to deplete a stock as far as markets and technology will allow and then move on to the next stock (1,29).

A variety of approaches have been tried to counter the tendency of fisheries to deplete stocks. In many cases, such efforts have failed. Consequently, prices paid for the produce of fisheries are now often much higher than they would have been under conditions of sustainable management (8). Moreover, it has been estimated that the revenues of the worlds fishing fleets are approximately US\$20–40 $\times 10^9$ below their operating costs (30,31).

Against this sobering background, there is no shortage of suggestions for improvement, all focusing on the supply side. Some have argued in favor of a revival of traditional arrangements that have worked well, such as the community-based fisheries management systems in the Pacific Islands and Japanese inshore fisherman's traditional user rights (28,32). Whether such a revival can be successful is dependent on meeting a number of conditions (28). Current developments are not necessarily conducive to this (27).

Recently, based on an initiative of the Worldwide Fund for Nature (WWF) and the company Unilever, a nongovernmental initiative emerged to certify sustainable fisheries. This is the Marine Stewardship Council that, by March 2004, had certified eight fisheries as sustainable. There is considerable favorable consumer interest in this approach (33), but it is too early to offer a verdict on the success of this scheme.

Most proposals for improving management practices have focused on government intervention. The use of incentive-based price instruments, such as landings and vessel taxes, has been advocated (31). Though, in the case of freshwater, adapting prices is increasingly popular as a means to attain sustainable management, this is different for fisheries. To the extent that efforts have been undertaken by governments, the focus has been firmly on trying to manage "physical" aspects of fisheries by measures such as limitation of effort, minimum allowable sizes for individual species, maximum allowable catches, and protected areas.

There have been both failures and cases of successful governmental management of fisheries focusing on physical aspects of fisheries. From the successes and failures so far, a number of lessons can be drawn (1,32,34–37).

- The problem of illegal catches exists, especially if stocks are close to the market. When the policing effort is not up to the ingenuity of illegal fishing, management efforts may well become a failure.
- With few exceptions ("dolphin safe" tuna fishing, the use of turtle excluders, and the Norwegian no-discard policy), management systems have been poor at limiting the negative impacts on stocks of by-catch and discards. Globally, discarding may be about 30% of landed tonnage.
- Many management strategies have opted for limitation of effort, for instance, by regulating the efficiency of fishing gear; engine size; fishing season; and/or limiting the number, size, and storage capacity of fishing boats. Only in a limited number of cases (e.g., the rock lobster fishery in West Australia) have such approaches been successful.
- A number of efforts have opted for allocating total allowable catches to political entities such as countries. In some cases, these efforts have been successes and in other cases, they have been failures. An important reason for failure, apart from illegal catches, has been that conservation goals were not separated from allocation, which resulted in losing conservation goals in allocation battles. Electoral considerations favoring perceived (short-term) interests of the fishing industry have also contributed to failure.
- More recently, there has been the tendency to base management on the allocation of long-term or even perpetual individual rights or individual transferable rights for their fishermen. In most cases, these rights have been based on past catch histories ("grandfather rights"). In some cases, such individual rights have been auctioned. There have been failures of such systems, such as the abalone fishery of British Columbia and the Southern school sharks fishery of Australia. But in other cases, this approach has worked relatively well. Success may be partly linked to the fact that individual rights that are allocated for a long time become more valuable if fishing practices are sustainable.
- The establishment of marine protected areas that are protected from fishing (which are aimed at conservation of fish stock but may also conserve or restore ecosystems) has so far not been practiced to an extent that allows a verdict on its effectiveness. The same holds for the allocation of fishing rights to cooperatives.

SUSTAINABLE MANAGEMENT OF ECOSYSTEMS

Natural ecosystems provide for a wide variety of services. Part of the services of natural ecosystems relate to water. Natural ecosystems are important in determining levels of greenhouse gases, and thereby climate, that in turn is a determinant of the global water cycle. Such ecosystems are also a major determinant of worldwide evapotranspiration. There are, furthermore, local quantitative and qualitative links between ecosystems and water. For instance, deep-rooted natural vegetation along watercourses catches minerals and returns these to topsoils. Thereby, it reduces the mineral content of surface waters, while increasing primary productivity on topsoils (38,39). Tropical rain forests are characterized by high levels of water recycling. When there is clear felling of such forests at low levels of externally added precipitation, the area may be converted into a savanna (40).

Marine, freshwater, and wetland ecosystems generate extensive services. It is estimated that the value of a number of ecosystem services in the entire biosphere is in the range of US$16–54 $\times 10^{12}$; the average (US $33 $\times 10^{12}$) is in excess of the gross world product (41). Of the average value of US $33 $\times 10^{12}$, somewhat over US $27 $\times 10^{12}$ comes from marine, freshwater, and wetland ecosystem services. There is even a case to attribute infinite monetary value to ecosystem services provided by living nature because human life would be impossible without them.

For sustainability, as defined here, ecosystem services should be maintained at their present level. To attain this, the area allocated to nature and the abiotic conditions under which nature functions should be such that ecosystem services do not change.

A debate has developed regarding the question, how much biodiversity can be reduced without reducing ecosystem services. Loss of a species should not lead to loss of function when there is functional redundancy. However, experimental evidence suggests that the statistical relation between species diversity and ecosystem functions important to humans may well be strong, though it is likely that actual contributions may show large differences between species (42–45). Even when a species seems functionally redundant, it has been found that the actual loss of apparent functional redundancy may come at great cost (46). Also a “redundant species” may still have an insurance function in case of ecosystem perturbation (47). So sustainable management is presumably strongly dependent on maintaining species diversity in ecosystems as that has developed.

A main problem in attaining criteria for sustainability is that many ecosystem services in practice do not have a price. There are economic incentives that may be conducive to the sustainable management of ecosystems. These include user fees for national parks, royalties from bioprospecting, conservation easements, and mitigation banking. So far, however, the application and effect of economic incentives are very limited (35). Legal systems for protecting nature are, with a few exceptions, rather focused on preventing the extinction of specific species than on ecosystem services. A major effort is needed to base nature conservation on the true value of ecosystem services.

BIBLIOGRAPHY

1. Hilborn, R. et al. (2003). State of the world’s fisheries. *Annu. Rev. Environ. Resour.* **28**: 359–399.
2. IUCN, UNEP, WWF. (1980). *The World Conservation Strategy*. WWF, Gland, Switzerland.
3. Daly, H. (1973). *Toward a Steady State Economy*. Freeman, San Francisco.
4. Becker, B. (1998). *Sustainability Assessment: A Review of Values, Concepts and Methodological Approaches*. World Bank, Washington, DC.
5. World Commission on Environment and Development. (1987). *Our Common Future*. Oxford University Press, Oxford, UK.
6. Ayres, R.U. (1996). Statistical measures of unsustainability. *Ecol. Econ.* **16**: 239–255.
7. Batabyal, A.A., Kahn, J.R., and O’ Neill, R.V. (2004). On the scarcity value of ecosystem services. *J. Environ. Econ. Manage.* in press.
8. Hueting, R. and Reijnders, L. (2004). Broad sustainability contra sustainability: the proper construction of sustainability indicators. *Ecol. Econ.*, in press.
9. Pimentel, D. and Houser, J. (1997). Water resources, agriculture, the environment and society. *Bioscience* **47**: 97–107.
10. Haddadin, M.J. (2002). Water issues in the Middle East. Challenges and opportunities. *Water Policy* **4**: 205–222.
11. Reijnders, L. (2000). A normative strategy for sustainable resource choice and recycling. *Resour. Conserv. Recycling* **28**: 121–133.
12. Gleick, P.H. (2003). Water use. *Annu. Rev. Environ. Resour.* **28**: 275–314.
13. Quadir, M. et al. (2003). Agricultural water management in water-starved countries. Challenges and opportunities. *Agric. Water Manage.* **62**: 165–185.
14. Araus, J.L. (2004). The problems of sustainable water use in the Mediterranean and research requirements for agriculture. *Ann. Appl. Biol.* **144**: 259–272.
15. Crites, R., Reed, S., and Bastian, R. (2001). Applying treated wastewater to land. *Biocycle* **April**: 32–36.
16. Tsagarakis, K.P., Dialynas, G.E., and Angelakis, A.N. (2004). Water resources management in Crete (Greece) including water recycling and reuse and proposed quality criteria. *Agric. Water Manage.* **66**: 35–47.
17. Levin, R.B. et al. (2002). US drinking water challenges in the twenty-first century. *Environ. Health Perspect.* **110**: 43–52.
18. Reijnders, L. (1996). *Environmentally Improved Production Process and Products*. Kluwer, Dordrecht, The Netherlands.
19. Antolioni, B. and Filipppini, M. (2001). The use of a variable cost function in the regulation of the Italian water industry. *Utilities Policy* **10**: 181–187.
20. Holt, C.P., Philips, P.S., and Bates, M.P. (2002). Analysis of the role of waste minimization clubs in reducing industrial water demand in the UK. *Resour. Conserv. Recycling* **30**: 315–331.
21. Braeken, L., van der Bruggen, B., and Vandecasteele, C. (2004). Regeneration of brewery waste water using nanofiltration. *Water Res.* **38**: 3073–3082.
22. Van Berkel, C.W.M. (1996). *Cleaner Production in Practice*. IVAM Environmental Research, Amsterdam, The Netherlands.
23. Rao, J.R. et al. (2003). Recouping the wastewater: A way forward for cleaner leather processing. *J. Cleaner Prod.* **11**: 591–599.
24. Puplampu, E. and Siebel, M. (2005). Minimisation of water use in a Ghanaian brewery: Effects of personnel practices. *J. Cleaner Prod.* in press.
25. Allen, J. (2004). Water efficiency. *Environ. Design Constr.*, May, pp. 65–66.
26. Hoffmann, R.C. (1996). Economic development and aquatic ecosystems in medieval Europe. *Am. Hist. Rev.* **101**: 631–699.
27. Evans, S.M. et al. (1997). Traditional management practices and the conservation of the gastropod (*Trochus nilitocus*) and

fish stocks in the Maluku Province (eastern Indonesia). *Fish. Res.* **31**: 83–91.
28. Ruddle, K. (1998). The context of policy design for existing community-bases fisheries management systems in the Pacific Islands. *Ocean Coastal Manage.* **40**: 105–126.
29. Kirby, M.X. (2004). Fishing down the coast: historical expansion and collapse of oyster fisheries along continental margins. *Proc. Natl. Acad. Sci. USA* **101**: 13096–13099.
30. Cochrane, K.L. (2000). Reconciling sustainability, economic efficiency and equity in fisheries: the one that got away. *Fish Fish.* **1**: 2–21.
31. Sanchirico, J.N. (2003). Managing marine capture fisheries with incentive based price instruments. *Public Finance Manage.* **3**: 67–93.
32. Caddy, J.F. (1999). Fisheries management in the twenty-first century: Will new paradigms apply? *Rev. Fish Biol. Fish.* **9**: 1–49.
33. Jaffry, S. et al. (2004). Consumer choices for quality and sustainability labeled seafood products in the UK. *Food Policy* **29**: 215–228.
34. National Research Council. (1999). *Sharing the Fish, Toward a National Policy on Individual Fishing Quota*. National Academy Press, Washington, DC.
35. Leva, C.E.di (2001). The conservation of nature and natural resources through legal and market-based instruments. *Reciel* **11**: 8495.
36. Baskaran, R. and Anderson, J.L. (2004). Atlantic sea scallop management: an alternative rights-based cooperative approach to resources sustainability. *Mar. Policy* in press.
37. Daw, T. and Gray, T. (2004). Fisheries science and sustainability in international policy: a study of failure in the European Union's Common Fisheries Policy. *Mar. Policy*, in press.
38. Schlesinger, W.H. (1991). *Biogeochemistry*. Academic Press, San Diego.
39. Christinaty, L., Kimmins, J.P., and Maily, S. (1997). Without bamboo, the land dies. *Forest Ecol. Manage.* **91**: 83–91.
40. Silveira, L.da and Sternberg, L. (2001). Savanna-forest hysteresis in the tropics. *Global Ecol. Biogeography* **10**: 369–378.
41. Constanza, R. et al. (1997). The value of the world's ecosystem services and natural capital. *Nature* **387**: 253–260.
42. Tilman, D., Wedin, D., and Knops, J. (1996). Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* **379**: 718–720.
43. Duarte, C.M. (2000). Marine biodiversity and ecosystem services: an elusive link. *J. Exp. Mar. Biol. Ecol.* **250**: 117–131.
44. Hector, A. et al. (2001). Community diversity and invasion resistance. *Ecol. Res.* **16**: 819–831.
45. Symstad, A.J. et al. (2003). Long term and large scale perspectives on the relationship of biodiversity and ecosystem functioning. *Bioscience* **53**: 89–98.
46. Bellwood, D.R., Hughes, T.P., Folke, C., and Nystrom, M. (2004). Confronting the coral reef crisis. *Nature* **429**: 827–833.
47. Walker, B. (1995). Conserving biological diversity through ecosystems resilience. *Conserv. Biol.* **9**: 747–752.

SUSTAINABLE WATER MANAGEMENT ON MEDITERRANEAN ISLANDS: RESEARCH AND EDUCATION

Manfred A. Lange
Dörte Poszig
Antonia A. Donta
University of Münster
Centre for Environmental Research
Münster, Germany

INTRODUCTION

The availability of water in adequate quality and sufficient quantity is indispensable for human beings and the functioning of the biosphere. Shortcomings in integrated water management practices aimed at linking natural, social, and economic needs in a given region, wasteful and inadequate use of water, and inadequate administrative practices have often resulted in depleted water resources. Pollution has exacerbated these problems.

Addressing issues of water quality and water supply, the European Union-Water Framework Directive (WFD) came into force on December 22, 2000 (1). Two basic goals are to be achieved by the parties within the following 10 to 15 years: (1) users will have to bear the full cost of the water supplied (1, Annex III, WFD), and (2) water will have to be provided based on "good ecological status" by adhering to practices that do not harm the environment. Moreover, the WFD strives toward implementing the "polluter pays principle."

Considering the long-term perspectives for sustainable water use in the European Union, the Mediterranean in general and its islands in particular deserve special attention for a number of reasons:

- Island water resources are largely influenced by the (non)existence, timing, and amount of precipitation. In some parts of the Mediterranean, the amount of (winter) precipitation has decreased (2), and the main precipitation period has shortened during the last few decades. This trend may continue as a result of regional climate change, thus exacerbating the problem (2,3).
- Overabstraction of aquifers, in many instances, has led to lowering of the water table and the deterioration of water quality, primarily through saltwater intrusion in coastal areas. In some cases, this has resulted in exploitation of lower lying aquifers, thereby further reducing the remaining reserves of the islands.
- The demand for water is expected to increase on Mediterranean islands because of, among others, an expected increase in the population and the number of tourists. Conflicts between competing users of water will inevitably intensify.
- Agriculture is the predominant consumer of water on most Mediterranean islands (Fig. 1). Agricultural activities threaten the availability (quantity) but

also the quality of water due to the use of fertilizers, pesticides, and the release of olive-oil-mill wastes. During the summer months, tourism also becomes a major water consumer, and competition between agriculture and tourism can lead to serious stakeholder conflicts.

- The rational distribution and use of water requires resolving conflicts among the different end users of water. Such conflicting interests have contributed to the unsustainable use of water in the past.
- Given the need for the protection and the sustainable use of island water resources, coupled with the need to satisfy the increasing water demand, the islands have to formulate strategic policies based on integrated water management and should pay proper attention to the objectives of the WFD.

The WFD may introduce an effective avenue to address these problems, but at the same time its implementation presents a challenge. The problems that govern the availability of water in the Mediterranean have specific manifestations on each of the islands, but the islands also share a number of general characteristics and patterns. Thus, learning how to implement the WFD best on selected Mediterranean islands through comparisons, mutual learning, and experience sharing offers the prospects of deriving a more generic set of recommendations.

The overriding characteristic that governs most of the problems mentioned lies in their complex and multifaceted nature, which requires a holistic approach involving the integration of natural and social sciences. Thus, in considering water availability and water use and deriving strategies for sustainable and equitable distribution of water, interdisciplinarity is a necessary condition.

Moreover, lasting solutions for sustainable water management in the Mediterranean will be found only through recommendations and/or regulations that are based on mutually agreed principles among the stakeholders. This requires a stakeholder-based, participatory process that builds on the results of scientific investigations, on the one hand, and on the consent of major stakeholders, on the other (5).

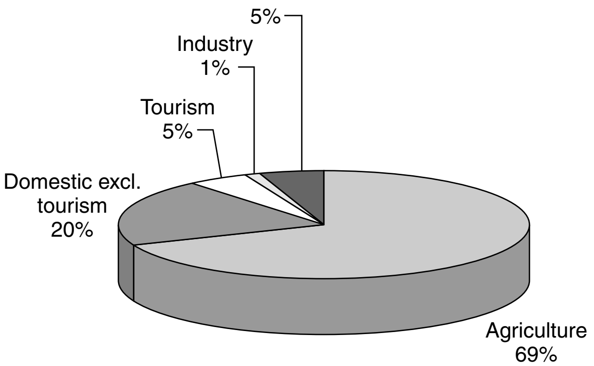

Figure 1. Distribution of water consumers on Cyprus in 2000 (after Reference 4). The predominance of agriculture as the major user is typical of most Mediterranean islands.

TOWARD SUSTAINABLE, STAKEHOLDER-BASED WATER MANAGEMENT PRACTICES: THE MEDIS PROJECT

To address the problems of water management on Mediterranean islands in accordance with the criteria just introduced and by observing the need to implement the WFD swiftly and effectively on these islands, an EU-funded research project, MEDIS (toward sustainable water use on Mediterranean islands: addressing conflicting demands and varying hydrological, social, and economic conditions), has been implemented. MEDIS is carried out by a consortium of 12 institutions from seven countries.

Goals and Objectives

The overall goal of MEDIS is to contribute to sustainable and equitable use of water on Mediterranean islands. This is to be achieved through a set of recommendations to implement sustainable, equitable water management regimes on each island that reflect the current and possible future (climatic) conditions on each island. These recommendations will comply with the WFD and will build on the conclusions reached in consultation with stakeholders from all five islands involved. More specifically, work in MEDIS

- concentrates on a synthesis of results from previous studies and integrating information from various disciplines;
- is carried out on a catchment scale on major islands of the Mediterranean stretching from the west to the east: Majorca, Corsica, Sicily, Crete, and Cyprus;
- seeks to develop an infrastructure for participatory stakeholder involvement that will enable establishing sustainable and equitable water management schemes;
- considers (possibly drastically) altered conditions that may arise due to changed climatic conditions through a set of what-if scenarios;
- concentrates on agriculture as the predominant consumer of water on Mediterranean islands; and
- undertakes comparative analyses between the islands to derive generic conclusions/recommendations and to use possible common solutions to water management problems in compliance with the WFD.

Methodology

The major stages of the research strategy applied in MEDIS, which are carried out during a 4-year period, are depicted in Fig. 2. The initial phase of the project comprises a thorough assessment of the major characteristics that determine the availability and the demand of water on each island. This includes (current and future) climatic conditions; (geo) physical, pedological, and hydrologic characteristics as well as the vegetative cover on each island; major sources and consumers of water; agricultural practices in irrigation and agrochemical use; conditions governing water management and administration; basic demographic and social conditions; and major economic indicators, particularly related to water-dependent economic sectors. This assessment leads to a typology of water management regimes for each island, based on a number

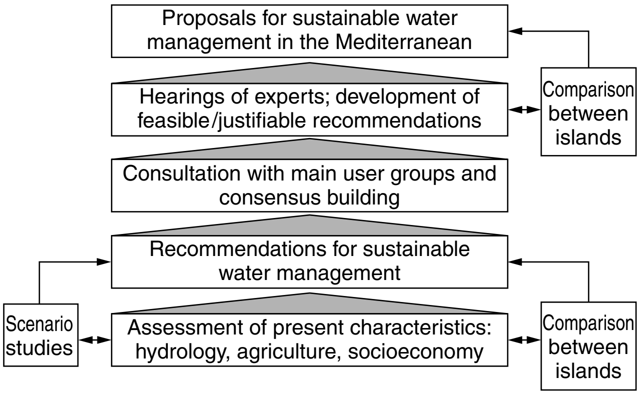

Figure 2. Schematic of major stages of the MEDIS research strategy (for details, see text).

of well-defined parameters, that share similar quantitative or qualitative ranges for each of the types specified. The final result of this project stage is a first set of recommendations for improved water management practices geared to each of the management types identified.

An integral part of the work during this stage of the project is involving stakeholders and gaining their help in specifying some of the previously mentioned parameters. They are either part of the supply network for water and/or consumers of water, so their experience and expertise are indispensable for the initial assessment. Therefore, an important activity during the first project stage is identifying relevant stakeholders and involving them in Intra-Island Workshops. In addition, Inter-Island Workshops among the islands under consideration are conducted.

Current physical/climatic conditions and possible future variations (i.e., scenario studies) are also considered. For climate, the results of the Intergovernmental Panel on Climate Change (IPCC) scenarios (6) are used in a downscaled format (or through a dedicated regional climate model). When considering prospective situations on each of the islands, future scenarios of other major determinants of water supply and demand are taken into account accordingly.

The spatial scales of the investigations depend on the specific questions asked. MEDIS is being carried out on a local (hundreds of m to km) and/or catchment scale, and inferences for all islands are drawn by appropriate upscaling. The Mediterranean scale is being addressed by comparing islands. As for temporal scales, MEDIS focuses on two time slices of 20 years centered at 2020 and 2050, corresponding to the IPCC assessments.

The second stage in the project comprises intensive consultation with the stakeholders on each island. Typical stakeholders include regional or national water administrative bodies, representatives of institutions or enterprises representing major water users, agricultural or industrial enterprises that develop and/or maintain water management schemes, nongovernmental organizations concerned with water issues, farmers and farming cooperatives, tourist enterprises and operators, and communities/residents in need of water.

The major goal comprises a consensus among the different stakeholder groups concerning their major objectives for an improved water management regime. At the same time, the stakeholders should bear in mind the recommendations from the first stage of MEDIS and specify possible alterations and amendments to these recommendations. Ultimately, a new set of recommendations will be obtained that represents both the factual basis and the mutually agreed upon principles for an improved water management regime on each island.

During the third stage of the project the second set of recommendations will be subject to scrutiny by legal and institutional experts. The objective here is to ensure that the recommendations comply with regional and national laws and regulations, and agree with the goals and principles of the WFD. This may result in a third set of recommendations that reflect the factual basis, the stakeholder consensus, and the legal and institutional requirements on national and European scales.

In the final stage of the project, island-specific conclusions will be discussed in the broader context of the Mediterranean Basin through Inter-Island Workshops and attempts to synthesize conclusions. The common solutions and possible deviations in the conclusions among the islands will be discussed and resolution attempted. Ultimately, this results in desired recommendations for improved, sustainable and equitable water management in compliance with the WFD and national rules and regulations. It is expected that these recommendations will significantly enhance the implementation of the WFD in the Mediterranean Basin.

First Results

Essential information was compiled in a series of Island Reports. In summary, the islands

- comprise areas between 3600 km^2 (Majorca) and 28,000 km^2 (Sicily),
- have permanent populations between 0.26 million (Corsica) and 5.6 million (Sicily),
- receive mean annual precipitation between 500 (Cyprus) and 1000 mm (Corsica, Sicily),
- attain a total annual water supply (through precipitation) from 2400×10^6 m^3 (Majorca) to $14{,}000 \times 10^6$ m^3 (Sicily; these numbers have partly been derived from the mean annual precipitation and the total area of the islands under consideration and are thus only rough estimates),
- have increasing water demand trends in the foreseeable future, and
- a water consumption structure clearly dominated by agricultural needs.

Though climatically similar, the amount of water supplied to various users differs greatly among the islands. This will have obvious consequences for the measures to be taken to implement the WFD.

The Island Reports and their synthesis led to recommendations aimed at reducing the vulnerability to water scarcity and implementing a sustainable and stakeholder-driven water management regime on the islands. In considering vulnerabilities, three different dimensions of water management have to be taken into account: a physical/environmental dimension, an economic/regulatory dimension, and a social/institutional/political dimension. Details of current and future vulnerabilities to water scarcity are described elsewhere (7), but some of the recommendations are presented in the following. These recommendations are interrelated and overlapping in many ways. They require a holistic, interdisciplinary approach to implement them.

Physical/Environmental Dimensions. The following recommendations have been formulated with regard to the physical and environmental dimensions of water scarcity:

- *Reduce Water Consumption*: The most promising strategies embrace (1) water pricing, an increase in consumer prices or alternatively/additionally quotas on water extraction may be imposed (possibly differentiated between different user groups); and (2) establishment of incentives for reduced water consumption, for example, subsidies for water saving and not—as has often been the case—for water consumption.
- *Change in Water Allocation*: Support of economic sectors accounting for maximum gross-domestic-product generation and employment through government programs. However, care must be taken to avoid possible one-sided economic advantages for specific sectors. Incentives for saving water in water-intensive sectors should be introduced in parallel.
- *Reduce Losses*: (1) Loss of water to the sea either as riverine discharge or through subsea groundwater discharge should be reduced (8), though care must be taken to avoid adverse impacts on biogeochemical cycles and on marine ecosystems in near-coastal waters; (2) reduce losses and contamination of water in distribution networks. On Cyprus, annual losses of drinking water in the distribution network account for 40×10^6 m^3, corresponding to 15% of the total demand and 23% of the total domestic demand (G. Constantinou, Geological Survey of Cyprus, pers. comm., March 2003). Improvement and/or renewal of distribution networks should be pursued where appropriate.
- *Increase Use of Additional Water Resources*: Though largely neglected for a long time, increasing attention is paid to this possible remedy. In particular, potable or irrigation water can be obtained through wastewater recycling, use of brackish water, and rainwater harvesting.

Economic/Regulatory Dimensions. The following actions addressing vulnerability related to economic/regulatory dimensions are recommended:

- Support sectors with high economic potential and small water needs.
- Change agricultural practices: (1) Consult and inform stakeholders about possible alternatives to current cropping patterns with regard to water requirements of plants, watering schedules, and the amount of irrigation water applied per plant (for a summary of possible alternatives, see Reference 9), optimize tillage systems as well as weeding and harvest controls to minimize irrigation; (2) optimize the application of agrochemicals to avoid adverse effects on groundwater quality. The use of agrochemicals and changes in agricultural techniques should be considered holistically to arrive at best practices that maximize yield and also minimize water use and the possible impact of agrochemicals on groundwater quality.
- Eliminate/reduce subsidies for water prices (in the past, subsidies have often contributed either to excessive water use and/or—in the case of agriculture—to cropping patterns that are economically and environmentally unsustainable).
- Promote cultivation of crops that have a high potential on the domestic and the foreign market (eliminate wasting products and water).
- Provide assistance in capacity building of farmers and for investments in modern irrigation technology.
- Provide economic incentives for rational water use in all sectors (e.g., specific water tariffs or relaxation of quotas and limitations in water consumption to those sectors that strive for a more rational use of potable water).

Social/Institutional/Political Dimensions. Reducing vulnerabilities related to social, institutional, and political aspects may be pursued through the following actions:

- Increase public awareness (water use, ownership, conflicts) to enhance the capacity to deal with water scarcity problems.
- Implement comprehensive monitoring of water extraction.
- Improve enforcement of existing rules and regulations.
- Simplify administrative mechanisms to enhance the efficiency of water administration.
- Transfer power from central government institutions to regional and local decision makers that seem better suited to become partners in the EU-WFD than individual users or single municipalities (e.g., TOEBs, Farmers Irrigation Organizations, in Greece).
- Encourage stakeholder-controlled water management structures.
- Ensure/improve an adequate factual basis for political decision making through input from science and stakeholders.

THE NEED FOR INTERDISCIPLINARY WATER STUDIES: THE ADVANCED STUDY COURSE SUSTAINIS

In addressing water management through research and technology projects, one is often faced with a problem different from those outlined. This is the lack of well-trained and motivated young professionals able to face the challenge of interdisciplinary water studies and the implementation of management options. The underlying reason for this problem is in the prevalent focus of traditional universities on specialized and discipline-focused curricula. Though quite adequate and successful in many instances, this kind of training leaves graduates ill-prepared for the type of work required in cross-disciplinary investigations. To resolve this problem at least partly, an Advanced Study Course, SUSTAINIS (Sustainable Use of Water on Mediterranean Islands: Conditions, Obstacles and Perspectives), was devised and implemented. SUSTAINIS was funded by the European Commission and was closely linked to the MEDIS project and its objectives.

Goals and Rationale

The course aimed at introducing 24 graduate students from various disciplinary and national backgrounds to the patterns of water supply and demand, to obstacles preventing rational use of water resources, and to perspectives for sustainable, equitable, and stakeholder-based use of water on Mediterranean islands. To provide them with an opportunity to learn about the specific environmental and socioeconomic features of Mediterranean islands, the course was carried out in Cyprus.

Curriculum and Structure

The emphasis in the teaching lay on demonstrating the complex interrelationships between environmental and societal systems that govern water regimes on Mediterranean islands and the importance of stakeholder-based strategies.

The course comprised lectures by 23 researchers from a variety of professional and national backgrounds, field trips, and discussion forums in the form of role games. The discussion forums were prepared by small, interdisciplinary, international groups of students and exposed the participants to the challenges but also to the advantages of interdisciplinary work.

The curriculum included eight modules:

Module 1: Introduction, Historical and Cultural Background of Water Use in the Mediterranean and Basic Concepts of Environmental Ethics. Cultural practices of management of freshwater resources from antiquity to the present, the recent history of water management and its importance for health, well-being, and prosperity of inhabitants of Mediterranean islands, and concepts of environmental ethics addressing normative issues in water management.

Module 2: Perspectives on Sustainable, Equitable, and Stakeholder-Based Water Management on Mediterranean Islands. Principles and tools for supporting participative water resource planning and management, design and implementation of an integrated water management approach and portrayal of research projects on integrated and sustainable water resources management in the Mediterranean.

Module 3: Essential Geologic and Hydrologic Characteristics of Mediterranean Islands. Characteristics and properties of the geology and hydrology of Mediterranean islands, description of subsurface aquifer conditions and problems related to overabstraction of groundwater resources (saltwater intrusion), and application of models in water management.

Module 4: Current and Possible Future Climatic Conditions and Their Impacts on the Hydrologic Regime. Climatic conditions in the Mediterranean and their effects on the timing and magnitude of precipitation, global and regional climate models and their strengths and weaknesses, review of expected future trends of climate development and their repercussions for water availability in the Mediterranean, and basic concepts and methodologies for integrated climate impact assessments.

Module 5: Current Patterns of Water Consumption, Major Stakeholders, Existing Conflicts, and Resource Policies. Conditions of water resources and common trends of water use with case studies from Cyprus, Malta, and Majorca; irrigation scheduling and water quality for agriculture; threat of deteriorating water quality through the application of agrochemicals and wastewater practices; water and tourism; examples of conflicts between users and possible solutions; and basic resource policies and patterns of political decision making.

Module 6: A Review of Nonconventional Technologies Aimed at Rational Water Management. Optimization of water consumption by employing appropriate and innovative technology like novel irrigation practices in agriculture, water recycling, and reuse for domestic and agricultural purposes, desalinization.

Module 7: Socioeconomic and Administrative Conditions of Water Use on Mediterranean Islands. Estimation of the value of water in its different uses, cost–benefit analysis of water projects, water consumption, public awareness, and alternative models of water management.

Module 8: Political Options, Initiatives, and Conditions for Sustainable Water Management, Specifically in the Context of the EU-WFD. Description of the EU Water Framework Directive, water pricing policies and the implications of the WFD, economic approaches in Integrated River Basin Management, and examples of integrated water management from European countries.

Results of the discussion forums as well as the lectures are published as course proceedings. In addition, a documentary film about the course and the issue of sustainable water use on Mediterranean islands for universities and passing on to stakeholders has been produced.

Student Feedback and Perspective

The student feedback showed that exposure to a wide variety of disciplines and research fields, and interacting with a group of individuals from a wide spectrum of disciplinary and cultural backgrounds were the most important contributing factors for the success of SUSTAINIS. It is expected that the SUSTAINIS course succeeded in teaching students about sustainable water use and management in the Mediterranean and in motivating them to pursue integrated regional studies in their careers.

Acknowledgments

MEDIS and SUSTAINIS were funded by the European Commission (contract numbers EVK1-CT-2001-00092 and EVK1-CT-2002-60001, respectively). We would like to thank Dr. Panagiotis Balabanis, the responsible officer at the Commission, for his support. Thanks are also extended to all partners in MEDIS, all colleagues who served as lecturers in SUSTAINIS, as well as the participants who contributed substantially to the success of the course.

BIBLIOGRAPHY

1. Commission of the European Communities. (2000). *Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy*, Off. J. Eur. Communities **L 327**: L 327/1–L 327/72.
2. Xoplaki, E. (2002). *Climate variability over the Mediterranean*, Ph.D. thesis, University of Bern, Switzerland.
3. Bolle, H.-J. (Ed.). (2003). *Mediterranean Climate—Variability and Trends*, Vol. 1 of Series in Regional Climate Studies, H.-J. Bolle, M. Menenti, and I. Rasool (Eds.), Springer-Verlag, New York.
4. Savvides, L., Dörflinger, G., and Alexandrou, K. (2001). *The assessment of water demand of Cyprus*. Ministry of Agriculture, Natural Resources and Environment, and Food and Agriculture Organization of the United Nations report TCP/CYP/8921.
5. Lange, M.A., Cohen, S.J., and Kuhry, P. (1999). Integrated global change impacts studies in the Arctic: the role of the stakeholders. *Polar Res.* **18** (2): 389–396.
6. Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K., and Johnson, C.A. (Eds.). (2001). *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, UK.
7. Lange, M.A. and MEDIS Consortium. (2004). Water management on Mediterranean islands: current issues and perspectives. In *Proc. IDS Water 2004*, http://www.idswater.com/Admin/Images/Paper/Paper_47/Manfred%20Lange-Paper.pdf, 15 pp., May 10–28 2004.
8. Burnett, W.C., Bokuniewicz, H., Huettel, M., Moore, W.S., and Taniguchi, M. (2003). Groundwater and pore water inputs to the coastal zone. *Biogeochemistry* **66**: 3–33.
9. Olesen, J.E. and Bindi, M. (2002). Consequences of climate change for European agricultural productivity, land use and policy, *Eur. J. Agron.* **16**: 239–262.

MEETING WATER NEEDS IN DEVELOPING COUNTRIES WITH TRADABLE RIGHTS

Terence R. Lee
Santiago, Chile

The most serious current issue in water management is not water scarcity, but the allocation of water among competing uses and users. The issue of allocation overshadows all other aspects of water management. Unfortunately, almost everywhere, the methods used for water allocation are woefully inadequate to achieve an effective, efficient, and timely distribution of water among uses and users. This failure to allocate water efficiently is the root cause of the widespread perception that water is becoming so scarce as to amount to a crisis. Almost everywhere, water is allocated either according to tradition or by bureaucratic processes subject to political pressures. Economists have long argued that water should be treated as an economic good. If it is treated as an economic good, then its use can be allocated by creating water rights through a market.

In most developing countries, where agriculture accounts for around 80% of water use, the reallocation of the use of water is crucial if water is not to be a limiting resource for future economic growth. It has been convincingly argued that permitting trade in water rights will increase available supplies because consumption will be reduced as water is used more efficiently by all users. This will not only alleviate, but probably even reverse, what many perceive as the coming water crisis.

In any country, however, the decision to introduce a system of tradable water rights requires considering of many issues so that the proposed water market can function smoothly and equitably to meet water needs:

1. In the initial distribution of rights, consideration must be given to both the rights of existing water users and the need to limit any windfall gains that they may receive as the owners of rights.
2. Once rights are allocated, the holder must be assured that the right is clearly and securely defined and appropriately registered.
3. The establishment of individual rights must not give the holders benefits at the expense of society as a whole.
4. Whatever decisions are taken and policies adopted, the system must be as simple as possible.
5. Once a system of water rights is clearly defined and the market is used as the mechanism for water allocation, then the intervention of government should be kept to a minimum.

The approach taken in the initial distribution of water rights is very important if trading in water rights is, in the future, to be the main means for water allocation. The size of any political opposition to the introduction of a water rights market is usually proportionate to the fairness seen in the initial distribution of water rights.

PRINCIPAL ISSUES

Under free market conditions, the trading of rights would result ultimately in the same allocation of water rights regardless of their initial distribution. However, in the expected presence of significant costs in any exchange of rights, the nature of the initial distribution of rights can affect the future efficiency of the market. Information, bargaining, contracting, and enforcement are required for a market to function, and they are not without cost. The initial distribution of water rights can also affect the quantity of transactions, the equilibrium allocation of rights, and the aggregate benefits of water marketing. If the costs of trading in rights are large enough to deter trading, then exchanges will not occur. There will always be a trade-off between promoting efficiency and equity considerations in the initial distribution of water rights.

Alternative initial assignments of water rights among users, whether individuals, ethnic groups, local governments, or environmental protection agencies, will also result in entirely different future sets of bargaining relationships and potentially different patterns of transfers.

Basing the initial distribution of water rights on historic water use is the easiest and the most commonly used system. The rule that has been usually adopted is the historic record of possession of licences or permits for water use under the previous allocation system, or, where such evidence is lacking, according to other benchmarks, such as land holdings. This is usually called "grandfathering." However, it does represent a transfer of wealth, the new water rights, to current water users.

Once a market is established, water rights will be reallocated through voluntary transfers between willing buyers and sellers in the same way as other goods are traded. It must be emphasized that the social impact of market transactions will depend on the initial distribution of water rights and bargaining power. The initial distribution will determine how the benefits from owning water rights are distributed, who has the protection of the state to use water, and who must pay to obtain water rights in the future.

An issue often raised in opposition to establishing markets in water rights is the possibility of windfall gains if existing water users are favored in the initial distribution of rights. Proposed solutions to this issue include taxing transfers, prohibiting transfers, and other restrictions on trading. Such restrictions, however, defeat the purpose of assigning rights and establishing a market. It might be simpler to ignore the problem, which will facilitate the reallocation of water to higher value uses but provide windfall gains to original users. One mechanism that has been used for the initial distribution of rights to avoid windfall gains is auctions, but these can raise the problem of rights going only to those who can pay.

In the initial distribution, care needs to be taken to respect the rights of the disadvantaged, such as poor farmers and indigenous peoples. In Chile, in introducing the market allocation of water, the government put in place and maintains a program to facilitate the legalization of property titles to water rights by poorer farmers; it has been spending more than US$320, 000 annually for this purpose during the last 20 years.

In the initial distribution of water rights, consideration should also be given to establishing minimum flows to protect aquatic and riparian habitats and other uses which, because of strong public goods characteristics from which everyone benefits, cannot compete in the market for water. Where historic uses have preempted the total supply, an argument can be made for a one-time reallocation of water. In Chile, although no provision of this kind is included in the law, the interpretation of the protection of the rights of third parties has been expanded to include environmental protection and ecological flows.

The main advantage of adopting "grandfathering" is that it avoids conflicts and reduces the opposition of existing water users, typically farmers, to the introduction of a market. Irrigation farmers usually argue that they are entitled to receive water rights without charge because they have already paid for the rights implicitly in the purchase price of the land. Under the auction approach, although payments for rights do not represent real economic costs to society as a whole but merely transfers from one group to another, to the users, the payments constitute a financial burden.

THE MARKET ALLOCATION OF WATER

A water market is a water management tool. It is, moreover, a tool that spreads the burden and difficulties of water management among a larger population; it permits greater popular participation in management decisions and can introduce greater flexibility into management systems. The creation of a water market will also demand new skills and attitudes from public administration, judicial systems, and water users, as well as investment in the registration of rights, monitoring and measurement systems and, possibly, in improving water distribution and transportation systems.

A water market will allocate water rights at a price set by the free exchange of the right either for a limited period of time through a lease or in perpetuity by a sale. For a market to work, it is necessary only that there be a tradable margin in water rights, even if the number of rights traded is marginal to the total supply. Water markets can normally be expected to be relatively small or "thin," with few transactions.

Any water allocation system should be both flexible and secure. At the same time, all costs and benefits should be reflected in the decisions that participants make; otherwise, these decisions will be inefficient from an overall social perspective. The system must also be predictable, equable, and fair. This is particularly important in developing societies where economic growth and water use efficiency require achieving a balance in the allocation of water between flexibility and security. Water markets are flexible because markets are by their very nature a decentralized and incentive-oriented institution, rather than centralized and regulatory.

SECURITY OF TENURE

Markets require security of ownership to function. Therefore, it is important that rights be clearly defined and publicly registered. Security of ownership of the right, in turn, helps encourage efficient use, resource conservation, and capital investment. It also helps strengthen and consolidate the autonomy of water user organizations. Security of ownership and the possibility of acquiring water rights in the market, it has been shown, encourage investment and growth in activities that require secure water supplies. Moreover, the reallocation of rights by voluntary exchanges allows market systems to defuse potential political conflicts over water allocation.

The political dimension of a system of secure and transferable water rights arises from the definition and clarification of property rights, which the introduction of such a system requires. The market transfer of water rights reduces political conflicts, as these transfers are always voluntary transactions in which owners of rights will participate only if they believe that it is in their best interest, given the alternative opportunities available. Administrative allocation, in contrast, often generates intense conflicts because granting a water right to one user necessarily precludes another.

Markets can, therefore, reduce conflicts among environmental interests, water suppliers, and polluters by providing natural economic incentives for water conservation and wastewater treatment. Environmental economists often prefer property-right systems to pricing systems because property-right systems can be ecologically more dependable than pricing systems.

WATER MARKETS IN CHILE

Chile is the only country where tradable rights are universal and all water allocation is the result of market transactions. The introduction of water markets in Chile coincided with a major increase in agricultural production and productivity. It is reasonable to conclude that the introduction of tradable, and particularly secure, property rights in water made a noticeable contribution to this overall growth in the value of Chile's agricultural production. This increase occurred within an agricultural sector largely dependent on irrigation, without a significant increase in either the supply of water or the area under irrigation. There has been, however, considerable private investment in improving the existing irrigation infrastructure both on and off the farm. The influence of water markets, however, cannot be fully separated from the effects of other economic factors, especially stable economic policies, trade liberalization, and secure land rights. This notwithstanding, it is recognized that trading in water rights reduced the need for new hydraulic infrastructure, improved overall irrigation efficiency, and has reduced the number of conflicts over water allocation. It also appears to have facilitated the shift from low-value, water-intensive crops to higher value, less-water-intensive crops.

In addition, market transfers of water rights have produced substantial economic gains from trade in some river basins. These gains occur both in trades between farmers and in trades between farmers and other sectors. The economic gains from trade tend to be large between farmers but relatively modest in intersector trade because water has been transferred from profitable farmers to urban drinking water supply; even though the financial gain to the farmer who sells is large, the overall economic gains of the reallocation are relatively small, because both uses have similar social benefits. In Chile, as in other areas of irrigation agriculture, if, prior to the sale, water was not being taken from river or canal by the owner of the right, other farmers would have used it downstream. Transfers among farmers, however, lead to using the water for higher value crops.

READING LIST

Anderson, T.L. and Snyder, P. (1997). *Water Markets*. Cato Institute, Washington, DC.

Holden, P. and Thobani, M. (1995). Tradable water rights: a property rights approach to improving water use and promoting investment. *Cuadernos de Economía* **32**(97): 263–316.

Frederick, K.D. (1993). *Balancing Water Demands with Supplies. The Role of Management in a World of Increasing Scarcity*. Technical Paper No. 189, The World Bank, Washington, DC.

Lee, T. and Jouravlev, A. (1998). *Prices, Property and Markets in Water Allocation*. United Nations, Economic Commission for Latin America and the Caribbean, Santiago.

Lee, T. (2000). *Water Management in the 21st Century: The Allocation Imperative*. Elgar, Cheltenham.

Peña, H. (1996). Water markets in Chile: what they are, how they have worked and what needs to be done to strengthen them. *Paper Presented at the Fourth Annual World Bank Conference on Environmentally Sustainable Development*. Washington, DC, pp. 25–27.

Rosegrant, M.W. and Binswanger, H.P. (1994). Markets in tradable water rights: potential for efficiency gains in developing country water resource allocation. *World Dev.* **22**(11): 1613–1625.

WATER USE IN THE UNITED STATES

Frédéric Lasserre
Université Laval
Ste-Foy, Québec, Canada

A STABILIZATION IN TOTAL WITHDRAWALS?

Total water withdrawals in the United States are estimated at 402 billion gallons per day (Bgal/day) in 1995. This is 1.5% less than withdrawals for 1990 and confirms an overall stabilization of withdrawals throughout the country since 1975. That was the first time that withdrawals growth stopped outpacing population increases, implying an intensification of the use of water withdrawn. All major users, thermoelectric, irrigation, and industrial uses, saw their withdrawals stabilize; rural (domestic and livestock) and urban use, on the contrary, witnessed an ongoing trend of withdrawal increase.

The decline in overall water use does not translate into the same pattern for all water sources (Table 1). Saline surface water use grew rapidly after 1965, to reach 71 Bgal/d in 1980, but has declined since then to 59.7 Bgal/d; groundwater use also declined after the 1980 peak. After a similar trend from 1980 to 1990, surface freshwater use, however, started growing again by 1.9% between 1990 and 1995.

On the other hand, after a decrease between 1980 and 1985 and a small increase from 1985 to 1990, consumptive use was back to growing faster than that of population. This trend could be linked to increased urban water use in suburbs. Particularly as urban sprawl takes place in the West, that implies a large share of consumption, such as lawn watering and pools, as will be seen below.

Table 1. Evolution of Water Use in the United States, 1950–1995[a]

	1950	1955	1960	1965	1970	1975	1980	1985	1990	1995
Population, in millions	150.7	164	179.3	193.8	205.9	216.4	229.6	242.4	252.3	267.1
Variation, %		8.8	9.3	8.1	6.2	5.1	6.1	5.6	4.1	5.9
Total withdrawals, Bgal/d	180	240	270	310	370	420	440	399	408	402
Variation, %		33.3	12.5	14.8	19.4	13.5	4.8	−9.3	2.3	−1.5
Of which:										
Thermoelectric	40	72	100	130	170	200	210	187	195	190
Other industrial	37	39	38	46	47	45	45	30.5	29.9	29.1
Irrigation	89	110	110	120	130	140	150	137	137	134
Rural domestic and livestock	3.6	3.6	3.6	4	4.5	4.9	5.6	7.8	7.9	8.9
Public supply	14	17	21	24	27	29	34	36.5	38.5	40.2
Source of water										
Ground freshwater	34	47	50	60	68	82	83	73.2	79.4	76.4
Surface, fresh	140	180	190	210	250	260	290	265	259	264
Surface, saline	10	18	31	43	53	69	71	59.6	68.2	59.7
Consumptive[b] use	NA	NA	61	77	87	96	100	92.3	94	100
Variation, %				26.2	13	10.3	4.2	−7.7	1.8	6.4

[a] Water-related data in billion gallons per day, Bgal/d.

[b] Consumptive use as of 1970: freshwater only.

Source: Adapted from Reference 1.

Table 2. Total Water Withdrawals by Water-Use Category Among Selected States, 1995, in Million Gallons per Day

State	Public, Domestic, and Commercial	Agriculture	Industrial, Incl. Mining	Thermoelectric
Alabama	880	268	753	5200
Arizona	867	5702	197	62
Arkansas	519	6294	187	1770
California	6125	29359	801	9655
Colorado	741	12759	192	114
Connecticut	475	29	12	3940
Delaware	104	52	64	1326
Florida	2417	3526	649	11636
Georgia	1295	770	677	3073
Idaho	560	14460	76	0
Illinois	2053	236	2407	17100
Kansas	399	3489	77	1270
Maine	146	29	16	135
Maryland	931	97	331	6360
Michigan	1535	241	1913	8370
Montana	161	8602	80	22
New Jersey	1144	127	486	4360
New York	3344	64	320	13060
Ohio	1628	54	650	8190
Texas	3464	9765	2916	13460
Utah	511	3638	252	55
Vermont	92	9	12	453

Source: Adapted from Reference 1.

THE DRY WEST STILL THE LARGEST USER AND CONSUMER

Regional Water Use

Wide discrepancies show up immediately when the structure of water use in the United States is examined.

Patterns do appear: they highlight very industrialized and urbanized states such as New Jersey, New York, and Michigan, where water used is concentrated in domestic, industrial, and power generating uses (Table 2). States that have rain-fed agriculture and a small urban population, such as Maine and Vermont, use water mainly for their power and their cities. In highly urbanized and industrialized Texas and California, all uses are strong, except industrial use in California because of an industrial base that does not use a large amount of water; other Western States do show a large dominance of agricultural use, such as in Arizona, Colorado, Idaho, and Montana.

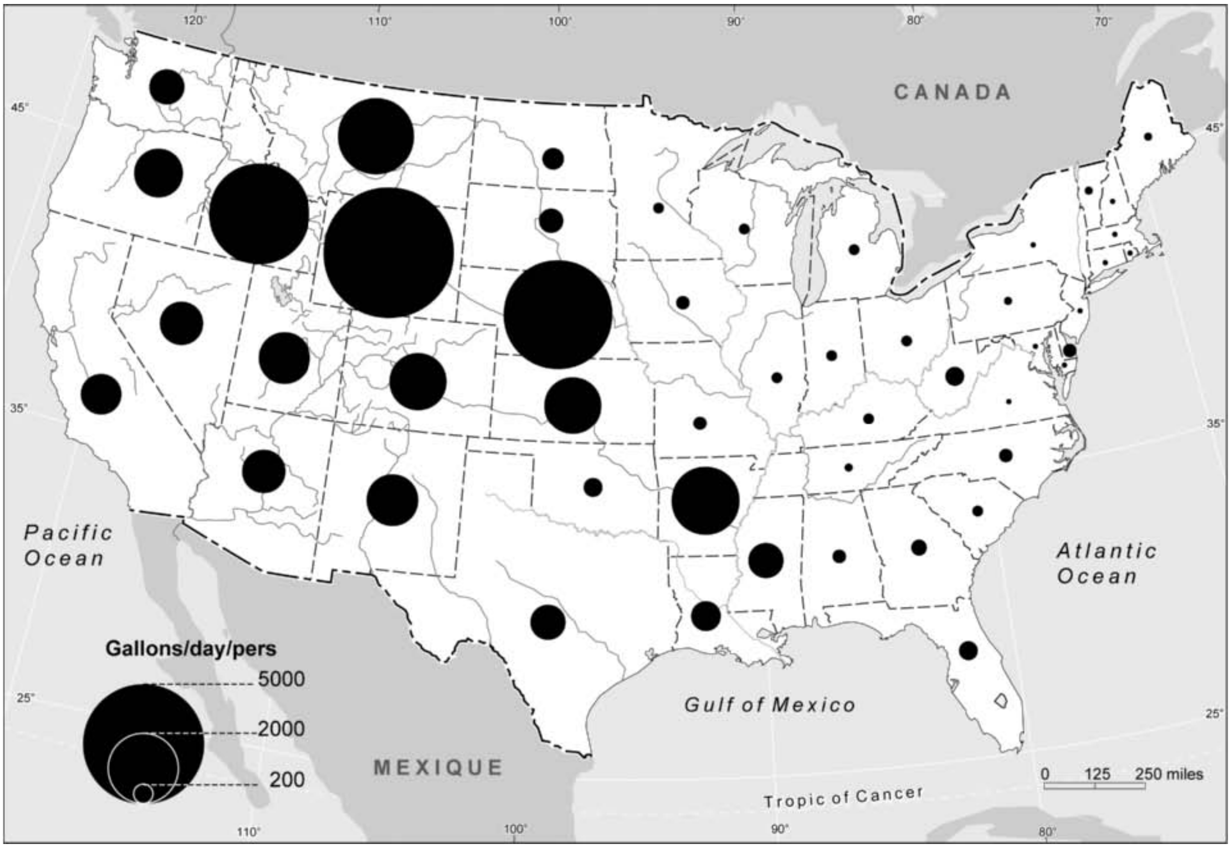

Figure 1. Water consumptive use by state, 1995 (adapted from Reference 2).

Consumptive Use

Regionally, states in the West do show much greater per capita consumptive use, up to 4288 gallons per day per person (gal/d/p) for Nebraska, but it is not the urban use that is responsible for their high consumption (Table 3). Highly urbanized California displays a more moderate figure of 795 gal/d/p (Fig. 1). Large withdrawals for irrigation are required in a dry climate.

AMERICAN WATER USE STRUCTURE COMPARED TO OTHER COUNTRIES

Agriculture remains the main user and consumer of water in the United States, as it is, by a large margin, among developing countries (Table 4). The picture is somewhat subtler among developed countries. Drier countries such as Spain and Greece, unsurprisingly, use a large share of water for agriculture, but water can also be largely used in more humid places that have traditionally rain-fed agriculture, such as Belgium, Germany, and Denmark (Table 5). This trend is explained, as observed in the United States, by farmers turning to irrigation to avoid all risk of yield decrease.

Table 3. Consumptive Use in Selected States, in gal/d/p, 1995

Arizona	California	Colorado	Florida	Idaho
908	795	1396	196	3732
Illinois	Maine	New York	Ohio	Rhode Island
72	39	26	71	19

Source: Adapted from Reference 1.

WATER USE MANAGEMENT : RECONCILING DIFFERENT USE STRUCTURES ACROSS THE COUNTRY

The geography of withdrawals and consumption, illustrated by the consumptive use map shown above, is all the harder to manage globally as Western states consume more and more (states west of the 100th meridian accounted for 75.1% of total consumption in 1995), but draw on a smaller share of the resource; water is far more plentiful in the eastern part of the country. From the end of the Civil War, this relative scarcity of water never prevented development schemes and led to large engineering designs that were built to harness the West's major rivers and make up for nature's pattern of groundwater, surface water, and precipitation distribution.

On the other hand, urbanized states in the eastern United States display small consumptive use structures, as low as 19 gal/d/p for Rhode Island and 26 for New York. It is the sheer size of the urban area that is responsible for New York's freshwater supply, not the profligacy of its inhabitants. However, New York and

Table 4. Share of Withdrawals in Selected Countries, by Economic Sector

		Withdrawals, Share of Total, %		
Country	Year	Agriculture	Industry	Domestic and Commercial
Developing countries				
India	1992	91	4	5
Uzbekistan	1994	94	2	4
Thailand	1992	90	6	4
Tanzania	1994	89	2	9
Vietnam	1992	78	9	13
China	1993	77	18	5
Developed countries				
Greece	1990	63	29	8
Spain	1994	62	26	12
Italy	1990	53	33	14
Portugal	1990	48	37	15
Denmark	1990	43	27	30
United States	1995	35	54	11
Germany	1990	18	68	14
France	1994	15	69	16
Ireland	1990	10	74	16
Norway	1985	8	72	20
Austria	1991	8	73	19
Belgium	1990	4	85	11
United Kingdom	1994	3	77	20

Source: Reference 3, pp. 208–211.

Table 5. Comparison of Withdrawals and Consumption by Main Sector Between France and the United States, 1995

In % of Total	Domestic/Commercial	Industrial	Thermoelectric	Agriculture
United States :				
Withdrawals	12.2	8.2	38.7	40.9
Consumption	8	4.1	3.3	84.6
France :				
Withdrawals	15	10	62.5	12.5
Consumption	24	5	3	68

Source: Reference 4, p. 102.

Los Angeles authorities, for instance, have aggressively promoted water-saving programs with effective results.

Local authorities are acutely conscious of serious water problems in the country. These are twofold. First, although withdrawals are on a stabilizing trend, and consumption was up until a few years ago, withdrawal levels are often locally unsustainable in the long term, especially in the West. Aquifers are being overpumped. The now famous Ogallala aquifer, in the Midwest, is expected to run unusable—not totally dry—in 50 to 70 years at present withdrawal rates. California saw its water consumption increase by 41.5% between 1980 and 2000, still untapped available resources are dwindling rapidly. As other upstream states withdrew small amounts of water for small populations, even Colorado water that California long took for granted is now being reallocated to the “New West” states under the Colorado Compact provisions because Nevada, Arizona, and New Mexico boast extremely fast population growth (Table 6). Between 1980 and 2000, consumptive uses also grew by 42% in Texas, 58.2% in Florida, and 70% in Arizona. Groundwater withdrawals in California exceed its aquifer renewal capacity by 15% (5).

Table 6. Population of Selected States, 1970–2000, in Millions

	1970	1980	1990	2000	Average Annual Growth, 1970–2000, %
Florida	6.79	9.75	12.94	16	+2.9
California	19.95	23.67	29.76	33.9	+1.8
Arizona	1.77	2.72	3.67	5.13	+3.6
Nevada	0.49	0.8	1.2	2	+4.8
Colorado	2.21	2.89	3.3	4.3	+2.2
New Mexico	1.02	1.3	1.52	1.82	+1.9
Utah	1.06	1.46	1.72	2.23	+2.5

Source: Adapted from Reference 6. Average population growth in the United States 1970–2000: +1.1%.

Second, the American society changes, and its geographical setting and consumption structures evolve. The

Table 7. Reclaimed Wastewater Use, 1995, in Gallons/Days/Person, Top 10 States

State	Arizona	California	Colorado	Florida	Maryland
Reclaimed water	42.7	10.4	2.9	16.7	13.9
State	Missouri	Nevada	Texas	Utah	Wyoming
Reclaimed water	2.1	15.7	5.8	7.2	19

Source: Reference 8.

Table 8. Desalination Capacity by Country, January 1999

Country	Total Capacity, Mgal/day	Total Capacity,[a] Mm^3/day	Country	Total Capacity, Mgal/day	Total Capacity,[a] Mm^3/day
Saudi Arabia	1.344	5.1	Japan	204.7	0.78
United States	851	3.2	Libya	185	0.7
United Arab Emirates	574.7	2.18	Qatar	149	0.57
Kuwait	338.4	1.29	Italy	137.1	0.52
Spain	210	0.8	Iran	115.3	0.44

[a] $1\ m^3 = 263.16$ gal.
Source: Reference 3, p. 288.

past 20 years have witnessed several phenomena that will question the water use structure: the acceleration of suburban growth, especially in the West; and the persistence of population shifts toward the Sun Belt in the West, Texas, and Florida; and more particularly toward what some geographers call the "New West," interior states such as Nevada, Arizona, Colorado, and New Mexico, where water resources are even more delicate to manage than in California.

This fast increasing population fuels the suburban sprawl that poses a great challenge to water management and distribution authorities; the newcomers are quick to adopt a new lifestyle and behavior that embodies, in their eyes, economic and societal achievement—green lawns, pools, frequent car washes. As a whole, Americans consume about 40 gal/d/p; in the Southwest, per capita daily consumption averages 120 gallons. In cities such as Phoenix and Las Vegas, residents consume more than 300 gal/d/p. In Denver, more than half of the water consumed is attributable to outdoor landscaping (7).

CONSUMING LESS

Water saving measures have thus been enforced by local authorities in various ways. Higher water pricing, for instance, led to a sharp rationalization of industrial water use, water had become a substantial cost for large water users, although trends in industrial restructuring led to a decline in water-intense processes. The same basic economic mechanism is also expected to prevent citizens from using too much water.

Total irrigation withdrawals decreased since 1980 to stabilize around 136 Bgal/d, although risk-averse farmers are more prone to irrigating their fields for fear of a decrease in their yield, even in more humid areas (4, p. 86); irrigated acreage remained stable at 58 million acres since 1980. This stabilization does not mean that there is no room for improvement, especially in the West, where extensive irrigation still often uses flood irrigation. In Idaho, porous soils require constant watering, up to 12 feet of water an acre per season.

This decrease in withdrawals stems from pricing, though it remains very low for farmers compared to urban and industrial users. This discrepancy fuels "water wars" in the West, as urban communities are confronting the first appropriation rights of rural irrigation districts to increase their share of a relatively scarce resource. It also stems from the idea the main resource would become too scarce. These two factors drove farmers to adopt improved irrigation systems and techniques.

DEVELOPING NEW SOURCES

Using and consuming less is an option more and more favored in the United States due to the transition from water-supply management to water-demand management. This shift in focus is attested to by changes in pricing policies as well as in recurring awareness campaigns targeted at citizens, industry managers, and farmers. But increasing supply is not a forgotten option. Grand schemes of damming Canada's untamed rivers and diverting them toward the thirsty Midwest were part of some engineering firms' and government agendas, notably in Québec, up until 1986; large-scale water exports were again contemplated in Canada from 1997 to 2001 but are now banned by a federal law. These projects no longer have the favor of an environmentally aware public, especially in a decade that saw large dams become the focus of environmentalists' anger and campaigning in the United States and throughout the Western world and even against some large projects in developing countries, such as the Narmada Dam project in India.

Reclaimed Wastewater

Reclaimed wastewater, for instance, was long thought of as irrelevant because of cost and safety issues. In 1995, total wastewater use amounted to 1.02 Bgal/d, a 36% increase since 1990 (Table 7). Water recycling is mainly a feature of the dry West. Florida, confronted with acute freshwater supply problems, has invested in extensive reuse programs. Arizona, Wyoming, and Florida appear to be the most effective states in recycling wastewater.

Desalting

Desalting also used to be considered an irrelevant solution because it produced very expensive water. During the last 15 years, improved technological processes have led to a fast decrease in the cost per gallon of water produced, although only urban or industrial users can afford it. Local governments are keenly interested in this technology, as attested to by recent major contracts in Florida, where a plant of 95,000 m^3/ day (25 million gal/d) approved by regional water officials in March 1999, is to be located near Tampa Bay. The water produced should cost between 45 and 55 ¢/m^3, or $1.71 to $2.08/1000 gal.

Desalination is proving very popular in the United States, where built capacity in early 1999 ranked second in the world (Table 8).

BIBLIOGRAPHY

1. USGS. (1995). *Water Use in the United States*.
2. USGS data. (2001).
3. Gleick, P. (Ed.). (2000). *The World's Water 2000–2001*. Island Press, Washington, DC, pp. 208–211, 288.
4. Lasserre, F. (2002). L'eau rare? Des solutions pour assurer l'approvisionnement. In: Lasserre and Descroix, *Eaux et territoires: tensions, coopérations et géopolitique de l'eau*. Presses Universitaires du Québec, Québec, p. 102.
5. Mayrand, K. (1999). *Les marchés internationaux de l'eau: exportations d'eau douce et marché des infrastructures et des services urbains*. Ministry of International Relations, Québec, pp. 8–9.
6. US Census Bureau. (2001).
7. Riebsame, W.E. (Ed.). (1997). Water for the west. *Atlas of the New West*. WW Norton & Company, New York, p. 81.
8. USGS. (1995). *Offstream Use*.

HOW WE USE WATER IN THESE UNITED STATES

U.S. Environmental Protection
Agency

Water use is usually defined and measured in terms of withdrawal or consumption—that which is taken and that which is used up. Withdrawal refers to water extracted from surface or ground water sources, with consumption being that part of a withdrawal that is ultimately used and removed from the immediate water environment whether by evaporation, transpiration, incorporation into crops or a product, or other consumption. Conversely, return flow is the portion of a withdrawal that is actually not consumed but is instead returned to a surface or ground water source from a point of use and becomes available for further use.

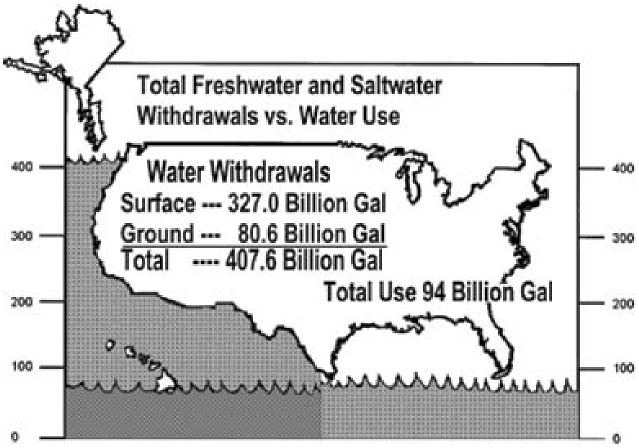

Figure 1. Total daily offstream water use in the United States.

Water use can also be divided into offstream and instream uses. Offstream water use (see Fig. 1) involves the withdrawal or diversion of water from a surface or ground water source for

- Domestic and residential uses
- Industrial uses
- Agricultural uses
- Energy development uses

Instream water uses are those that do not require a diversion or withdrawal from the surface or ground water sources, such as:

- Water quality and habitat improvement
- Recreation
- Navigation
- Fish propagation
- Hydroelectric power production

NATIONAL TRENDS IN WATER USE

National patterns of water use indicate that the largest demand for water withdrawals (fresh and saline) is for thermoelectric generation (47 percent), followed by irrigation (34 percent), public supply (9 percent), industrial (6 percent), mining (1 percent), livestock (1 percent), domestic (1 percent), and commercial uses (1 percent). While thermoelectric generation represents the largest demand for fresh and saline withdrawals, irrigation represents the largest demand for freshwater withdrawal

---

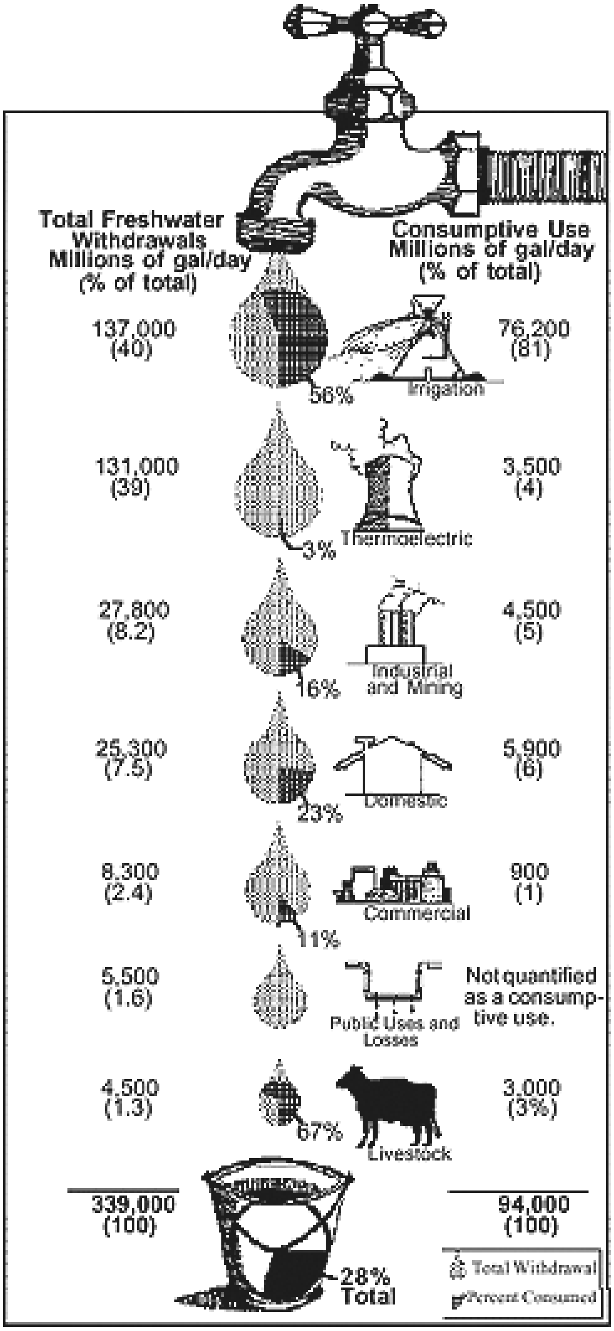

Figure 2. Comparison of freshwater consumptive use in the United States for 1990 by category.

alone (see Fig. 1). Activities that reduce the need to withdraw surface and ground water will lead to many of the beneficial effects of conserving water.

NATIONAL CONSUMPTION PATTERNS

Water consumption varies by water use category, with irrigation consuming the highest percent (81 percent) and commercial the lowest (1 percent) (see Fig. 2). The difference between the volume of water withdrawn and that consumed is the return flow. As more good-quality water is available in return flows, more water is available for other beneficial uses.

Some categories of water use, such as irrigation and livestock watering, consume a high percentage of water that is withdrawn from surface and ground water sources. Thus, less water is available for return flows from these high-consumption activities. Other categories of use like thermoelectric power consume only a small fraction of the water they withdraw.

CATEGORIES OF WATER USE

With several different ways to categorize water use in the United States, this article separates offstream uses into

- Municipal/public supply
- Domestic and commercial
- Industrial and mining
- Agricultural
- Thermoelectric power

MUNICIPAL/PUBLIC WATER SUPPLY

While water withdrawals for public use can be applied to street cleaning, fire fighting, municipal parks, and public swimming pools, keep in mind that municipalities and private suppliers might also provide water for other purposes (domestic/commercial, agricultural, thermoelectric power).

Per capita (per person) use of public water supplies in the United States (1990) averaged 183 gallons per day (gal/d). The average per capita use can vary greatly between communities for any number of reasons, including, but not limited to:

- Climate differences
- The mix of domestic, commercial, and industrial uses
- Household sizes
- Lot sizes
- Public uses
- Income brackets
- Age and condition of distribution system

For instance, per capita use of public water is about 50 percent higher in the West than the East mostly due to the amount of landscape irrigation in the West (see Fig. 3). However, per capita use can also vary greatly within a single state. For example, in 1985 the demand for municipal water in Ancho, New Mexico, totaled 54 gallons per capita per day (gal/cap/day) while in Tyrone, New Mexico, municipal demand topped off at 423 gal/cap/day. Rural areas typically consume less water for domestic purposes than larger towns.

In 1990, water withdrawn nationwide for public supplies totaled 38,530 million gallons per day (Mgal/d) (see Table 1). Although this withdrawal rate represents a 5 percent increase over 1985 amounts, the number of people supplied with water distributed through public systems also increased 5 percent during that same 5-year period. Again in 1990, surface water supplied about

Figure 3. Average use per person (gal/day) of public water in the United States by USGS water region.

Table 1. Fate of Water in Public Water Supplies of the United States, 1990

Receiving Category	Volume, Mgal/day	Percentage of Total
Domestic	21,900	57
Commercial	5,900	15
Public Use Losses	5,460	14
Industrial	5,190	13
ThermoelectricPower	80	<1
Total	**38,530**	**100**

61 percent of the public water supply, with ground water supplying the other 39 percent.

Of the total water withdrawn in 1990 for public supplies representing 11 percent of total U.S. offstream freshwater withdrawals 72 percent went to domestic and commercial uses, 13 percent to industrial uses, and 0.2 percent to thermoelectric power. The remaining 14 percent went to public uses such as fire protection or was lost during distribution (usually due to leaks).

DOMESTIC/COMMERCIAL

Domestic water use includes everyday uses that take place in residential homes, whereas commercial water uses are those which take place in office buildings, hotels, restaurants, civilian and military institutions, public and private golf courses, and other nonindustrial commercial facilities. Combined freshwater withdrawals for domestic and commercial use in 1990 totaled 33,600 Mgal/d, or 10 percent of total freshwater withdrawals for all offstream categories (see Fig. 2).

Typical categories of residential water use include normal household uses such as

- Drinking and cooking
- Bathing
- Toilet flushing
- Washing clothes and dishes
- Watering lawns and gardens
- Maintaining swimming pools
- Washing cars

When divided into indoor uses and outdoor uses, the amount of indoor water use remains fairly constant throughout the year, with the breakdown of typical indoor water uses depicted in Fig. 4. By far the largest percentage of indoor water use occurs in the bathroom, with 41 percent used for toilet flushing and 33 percent for bathing.

Outdoor residential water use, however, varies greatly depending on geographic location and season. On an annual average basis, outdoor water use in the arid West and Southwest is much greater than that in the East or Midwest. Figure 5 compares the national average for residential outdoor water use with that of Pennsylvania and California, with landscape irrigation the primary application. While average outdoor water use in Pennsylvania represents only approximately 7 percent of the total residential demand, in California average outdoor use climbs to about 44 percent of the demand.

INDUSTRIAL AND MINING

Industrial water uses, estimated to be 8 percent of total freshwater use for all offstream categories, include cooling in factories and washing and rinsing in manufacturing processes. Some of the major water-use industries include mining, steel, paper and associated products, and chemicals and associated products.

Water for both industrial and mining uses comes from public supplies, surface sources, and ground water. During the 5-year span from 1985 to 1990, industrial water use in the United States decreased approximately 13 percent. In the same period, mining water use increased about 24 percent.

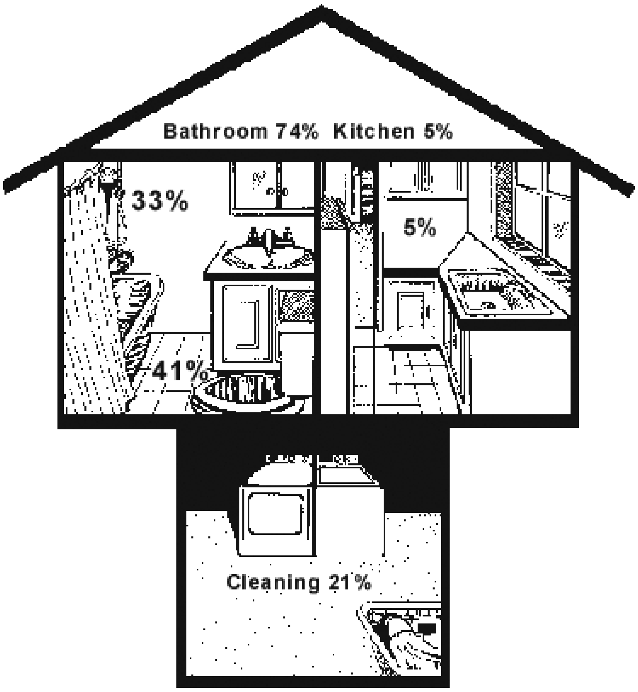

Figure 4. Typical breakdown of interior water use.

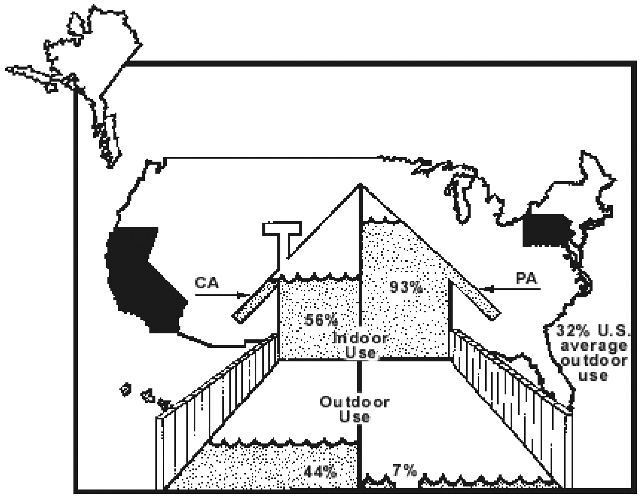

Figure 5. Comparison of average national residential outdoor water use with that of Pennsylvania and California.

AGRICULTURAL

Agricultural water use can be divided between irrigation and livestock. Irrigation includes all water applied to farm or horticultural crops; livestock incorporates water used for livestock, dairies, feedlots, fish farms, and other farm needs.

Estimated annual water use for irrigation remained at about the same level between 1985 and 1990, with approximately 63 percent of the water used for irrigation in 1990 coming from surface water. Approximately 60 percent of the water used for livestock came from ground water sources and the remaining 40 percent from surface water sources. Combined water use for irrigation and livestock represents about 41 percent of total offstream freshwater use for 1990, (see previous Fig. 1) with 40 percent going to irrigation and the lone 1 percent to livestock uses.

Not only can the loss of water from irrigation conveyance systems be significant, but the percentage of consumptive water use for agriculture is high as well an estimated 54 percent consumption in 1985. By 1990 this had climbed to an estimated 56 percent consumption for irrigation uses and 67 percent for livestock uses (see Fig. 2).

THERMOELECTRIC POWER GENERATION

This final category includes water used for the production of energy from fossil fuels, nuclear energy, or geothermal energy. Most water withdrawn for thermoelectric power production is used for condenser and reactor cooling. While 1990 estimates of freshwater withdrawals remained constant from 1985, nearly half again as much saline water was also used.

More than 99 percent of the water used for thermoelectric power production comes from self-supplied surface water, less than 0.2 percent from public supplies. In 1990, water used for thermoelectric power production represented close to 39 percent of total offstream freshwater use in the United States, but only about 3 percent was consumed.

The Mid-Atlantic, South Atlantic Gulf, Ohio, and Great Lakes water resource regions use the largest amounts of water for thermoelectric production. The eastern United States uses about five times more water than the West to produce about twice as much thermoelectric power.

VALUING WATER RESOURCES

Diane Dupont
Steven J. Renzetti
Brock University
St. Catharines, Ontario,
Canada

INTRODUCTION

Suppose that consideration is being given to dam a river. The dam will provide benefits in the form of hydroelectric power, water for irrigation, and protection from floods. Conversely, building the dam will also destroy wetlands, fish spawning grounds, and impair recreational opportunities. In deciding whether to build the dam, one important type of information is the value to society of the benefits and damages that will arise. Obtaining estimates of these values requires an understanding of the services that society derives from water, how those services will change with the building of the dam, and the value that society attaches to each service. It is the last of these tasks that this article addresses. Specifically, the purpose of this article is to review briefly how economists estimate the value of water. Because of the importance of water and the increasing costs of maintaining clean water, there has been a large amount of research into this topic (1–5). This article reviews the methods developed by economists to estimate water's value in the next section. Empirical estimates of the various uses of water are discussed later.

THE ECONOMIC PERSPECTIVE ON VALUATION

The starting point for most economic methods for measuring the value of water is an understanding of the structure of household tastes or preferences and firms' productive technologies. This understanding can then be used to develop measures of the benefits of water use and the change in household welfare or firm profits that arise from changes in the price, quantity, or quality of water. These measures, in turn, allow the analyst to determine the maximum willingness to pay (in the case of a change leading to an improvement in welfare) for water. This idea of willingness to pay forms the cornerstone of economic measures of value. (This approach to understanding and measuring value is not without its critics; see the discussion in Reference 1.)

These measures provide estimates of the value of water to the individual user (i.e., the "private" value). This may differ from the social value of water use (the value to all of society) by an individual household, firm, or farm.

There are a number of reasons that social and private values of water use may differ. First, an individual's water use may result in lowered water quality from pollution, and this, in turn, would lower the value of that water for other users and, thus, to society. Second, a farm may use water to increase production of crops that are supported by government subsidies. In this case, the farmer's valuation of the increased output will be higher than society's because of the cost of the subsidies to society.

Another issue complicating the measurement of the value of water is that, like many environmental goods, the value of some of the uses to which water is put are reflected in market transactions and others are not. An example of the former is a farmer purchasing water to irrigate crops—an analyst may be able to infer the value of water to the farmer by using market information on the increase in crop yield brought on by irrigation. An example of the latter is a family enjoying a day of swimming at a municipal beach. In this case, analysts must use "nonmarket" methods to estimate the family's valuation of its recreation.

As Table 1 indicates, economic valuation methods may be categorized into indirect and direct methods. Indirect valuation methods are those that infer a household or firm's valuation of water from its observed market behavior. There are a number of indirect valuation techniques. The derived demand method applies statistical or mathematical programming techniques to observations on market transactions in water to estimate a household's or firm's demand for water. The estimated demand relationship can, in turn, be used to calculate the user's willingness to pay for water. The avoided cost method assesses the value of water by examining the costs that are avoided because water is available. (A related method is the averting behavior method. If households purchase water filters to remove waterborne impurities, then the value to the household of removing the impurities can be estimated by the cost of water filtration.) For example, if a firm is able to release effluents into a lake rather than treating them, then the avoided cost of treatment provides an estimate of the value of the lake's assimilative capacity to the firm. The residual imputation method calculates the value of water to a firm or farm as the difference between the operation's revenues and the cost of all of its nonwater inputs. The hedonic method uses statistical techniques to determine the contribution of water to the total value of marketed commodities. For example, a house of a given size is usually more valuable, the closer it is to a body of water, and hedonic methods can estimate the increase in value that is related to the proximity to water. Finally, the travel cost method is used to estimate households' valuation of water-related recreation experiences by examining the time traveled to a recreational site and out-of-pocket expenses incurred by a family. A greater time traveled implicitly raises the cost of recreation because it usually means more time away from work.

Direct valuation methods are those that survey households to obtain their valuations. The most important direct valuation technique is contingent valuation. In this approach, individuals are presented with information concerning a hypothetical or constructed market and asked to indicate their willingness to pay to achieve a desired good or service. For example, a respondent might be asked to value a potential improvement in water treatment that reduces the risk of illness. This method is valuable for assessing projects before they are undertaken and for assessing households' "nonuse" values of water. For example, some households may be willing to pay a positive amount to ensure access to water some time in the future (6).

EMPIRICAL ESTIMATES OF THE VALUE OF WATER

The preceding section demonstrates that economic researchers have developed the conceptual framework and measurement methods for estimating the value of water. This section briefly points to a number of examples of the results of estimation studies. (The interested reader

Table 1. Economic Valuation Methods

Method	Applications	Comments
	Indirect	
Derived demand	Residential, industrial	Estimated demand curve yields user's valuation of water
Avoided cost	Residential, industrial	Access to clean water allows firms and homes to avoid costs such as water filters
Residual imputation	Agricultural, industrial	Requires information on revenues and costs of nonwater inputs
Hedonics	Residential, recreation	Good for assessing the impacts of changes in water quality
Travel cost	Recreation	One of the earliest methods—in use since the 1960s
	Direct	
Contingent valuation	Recreation, changes in water quality, ecosystems	Only method that can estimate nonuse values (values that arise from an aesthetic appreciation of water or a desire to hold an option to use water in the future)

Table 2. Representative Valuation Studies

Author	Sector/Use	Method	Estimate/Year
		Indirect	
Moore (8)	Agriculture	Derived demand	\$42–70/acre-foot (1989 US\$)
Young and Gray (9)	Industrial waste disposal	Avoided cost	\$0.07–1.28/acre-foot (1972 US\$)
Gibbons (2)	Industrial water intake	Avoided cost	\$6–78/acre-foot (1980 US\$).
Abdalla (10)	Residential	Avoided cost	\$125–330/household (1987 US\$)
Smith, Desvousges and McGivney (11)	Recreation	Travel cost	\$10–20/household (1977 US\$)
		Direct	
Whitehead and Blomquist (12)	Wetland preservation	Contingent valuation	\$6–21/household (1992 CAN\$).
Carson and Mitchell (13)	Recreation	Contingent valuation	\$70–90/household (1990 US\$)

should consult the valuable compilation of empirical estimates of water's value in Reference 7.) The results of a number of representative studies are reported in Table 2 and are discussed briefly below.

Indirect Valuation Studies

Moore (8) compares estimates of the value of irrigation water using the derived demand and residual imputation methods. The author finds that the two methods yield quite different results (the derived demand method provides higher estimated values). This is an important finding as the price paid by some American farmers for irrigation water is based, in part, on their estimated valuation of that water.

Young and Gray (9) and Gibbons (2) employ the avoided cost approach to value industrial water use. In both cases, the authors calculate the costs that firms are able to avoid by being able to dispose of effluents in waterbodies and by using increased water intake rather than in-plant water recirculation. Young and Gray find that the value of water used in industrial waste disposal is relatively low as few costs are actually avoided. Conversely, Gibbons finds a wider range of values for increased intake water substituting for water recirculation because, in some applications, in-plant recirculation must also address reduced water quality.

Abdalla (10) examines household valuation of improvements in the quality of their drinking water supplies by measuring the cost savings of those improvements. These cost savings arise from decreased purchases of water filters and bottled water and from less time needed to boil water. The authors also find that households with children place a higher value on water quality improvements than those without.

Smith et al. (11) address the change in the valuation of sport fishing from changes in water quality at specific sites. The authors do this by investigating the role of water quality in approximately 20 published studies of water-based recreation that use the travel cost method. The principal finding of the Smith et al. study is that increases in water quality increase the demand for, and valuation of, water-based recreation. For example, when applied to data derived from users of the Monongahela River, the average consumer surplus (a measure of willingness to pay net of fishing expenses) for water quality improvements from boatable to game fishing conditions was \$9.96 per household per season (1977 US\$).

Direct Valuation Studies

One of the most important advantages of direct valuation methods is that they can be employed to consider household 'nonuse' values of water. These values arise not from household use of water but rather because households may have an interest in knowing that a particular waterbody has been preserved (possibly for their use in the future). For example, Whitehead and Blomquist (12) employ a contingent valuation survey to examine household willingness to pay for wetlands preservation and the role that information provision plays in determining that value. The authors find that, while many households are willing to contribute toward the preservation of wetlands, as more information is provided about the availability of substitutes for an endangered wetland, respondents' willingness to pay for preservation falls off.

Finally, Carson and Mitchell (13) report on an influential study of American household valuation of enhanced water-based recreational opportunities that could arise from improved water quality. A particularly important feature of this study's methodology is the contingent valuation survey that employs the water quality "ladder" to set out for respondents the various levels of water quality that they are asked to consider. In doing so, the authors frame or present differ levels of water quality ("boatable, fishable, and swimmable", p. 2447) in a way that is comprehensible for most respondents. The results indicate that the sample of households places significant value on improving water quality.

CONCLUSION

The growing scarcity of water creates the need to develop and implement rational methods for allocating water across competing needs. Part of this process of decision-making requires understanding the value that individual households, firms, and farms obtain from their various uses of water. This article has briefly outlined the methods developed by economists to estimate those values.

BIBLIOGRAPHY

1. Donahue, J. and Johnston, B. (Eds.). (1998). *Water, Culture and Power: Local Struggles in a Global Context*. Island Press, Washington, DC.
2. Gibbons, D. (1986). *The Economic Value of Water*. Resources for the Future, Baltimore.
3. Young, R. (1996). *Measuring Economic Benefits for Water Investments and Policies*. Technical Paper No. 338. The World Bank, Washington, DC.
4. Bergstrom, J., Boyle, K., and Poe, G. (2001). Economic value of water quality: introduction and conceptual background. In: *The Economic Value of Water Quality*. J. Bergstrom, K. Boyle, and G. Poe (Eds.). Edward Elgar, Cheltenham, UK, pp. 1–17.
5. Renzetti, S. (2002). *The Economics of Water Demands*. Kluwer Academic Press, Norwell, MA.
6. Desvousges, W., Smith, V.K., and Fisher, A. (1987). Option price for water quality improvements: a contingent valuation study for the monongahela river. *J. Environ. Econ. Manage.* **14**: 248–267.
7. Frederick, K., VandenBerg, T., and Hanson, J. (1997). *Economic Values of Freshwater in the United States*. Discussion Paper 97-03, Resources for the Future, Washington, DC.
8. Moore, M. (1999). Estimating irrigators' ability to pay for reclamation water. *Land Econ.* **75**(4): 562–578.
9. Young, R., Gray, S., Held, R., and Mack, R. (1972). *Economic Value of Water: Concepts and Empirical Estimates Report to the National Water Commission*, NTIS no. PB210356, National Technical Information Service, Springfield, MA.
10. Abdalla, C. (1994). Groundwater values from avoidance cost studies: implications for policy and future research. *Am. J. Agric. Econ.* **76**: 1062–1067.
11. Smith, V.K., Desvousges, W., and McGivney, M. (1983). Estimating water quality benefits: an econometric analysis. *South Econ. J.* **50**: 422–437.
12. Whitehead, J. and Blomquist, G. (1991). Measuring contingent valuation for wetlands: effects of information about related environmental goods. *Water Resour. Res.* **27**(10): 2523–2531.
13. Carson, R.T. and Mitchell, R.C. (1993). The value of clean water: the public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resour. Res.* **29**(7): 2445–2454.

WATER—HERE, THERE, AND EVERYWHERE IN CANADA

Environment Canada

How much water is there in the world? Scientists estimate over one billion cubic kilometers (one cubic kilometer of water would fill 300 Olympic-sized swimming pools). Water covers nearly three quarters of the earth's surface in oceans as well as rivers, lakes, snow and glaciers. There is water in the atmosphere and water underground. Water evaporates and returns to the land surface in what is known as the *hydrologic cycle*.

In the hydrologic cycle, water evaporates from the ocean into the atmosphere, from there it can precipitate back into the ocean, or onto the land surface. From the land, it can evaporate or transpire back into the atmosphere, or flow overland or percolate underground before flowing back into the ocean. The distribution of the water around the globe varies from season to season and year to year, but the total quantity of water on the earth's surface remains essentially constant. The hydrologic cycle is discussed in detail in Freshwater Series No. A-1, "Water—Nature's Magician."

Although water exists in other forms in the hydrologic cycle, this issue in the Freshwater Series focuses on surface water as it is this water which we see in our everyday lives. Most of the earth's water is salty or permanently frozen. Figure 1 illustrates the proportion of fresh water that is available to us from the world's water supply.

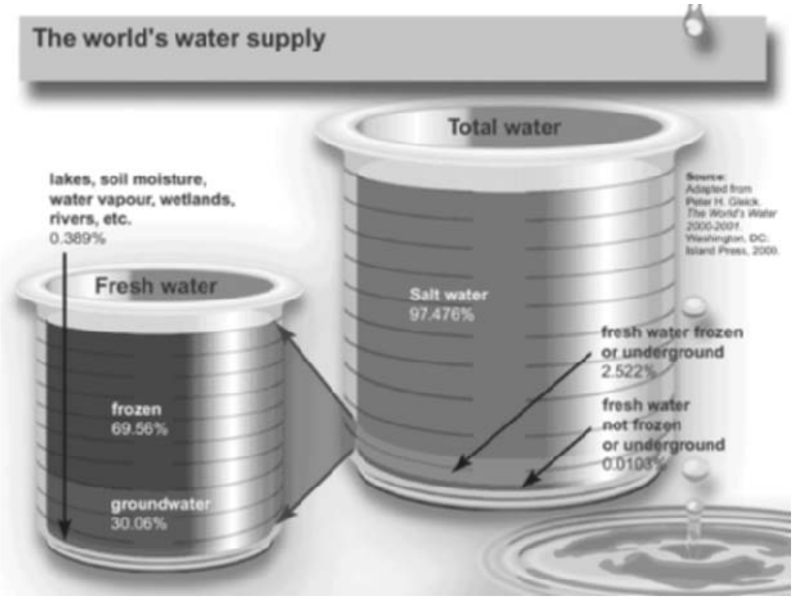

Source: Adapted from Peter H. Gleick, The World's Water 2000-2001. Washington, DC: Island Press, 2000.

Figure 1.

This article is produced by Environment Canada and is reproduced with permission.

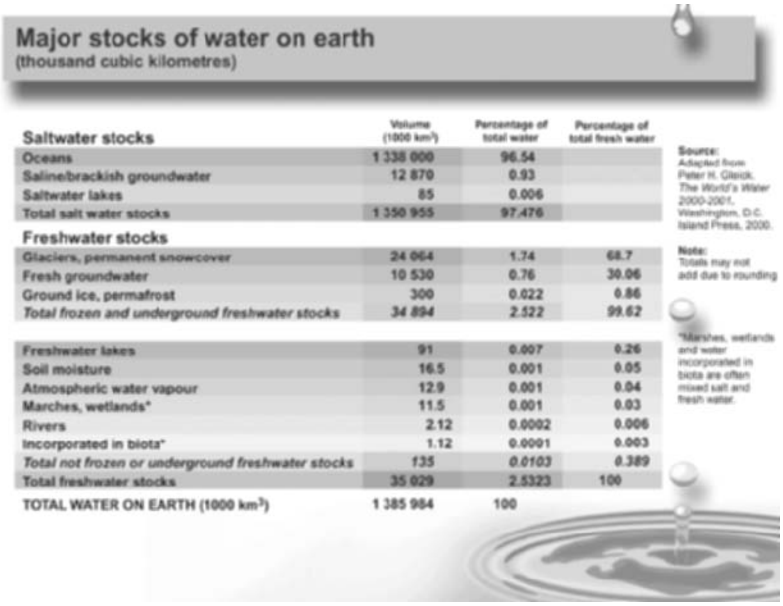

Major stocks of water on earth
(thousand cubic kilometres)

Saltwater stocks	Volume (1000 km³)	Percentage of total water	Percentage of total fresh water
Oceans	1 338 000	96.54	
Saline/brackish groundwater	12 870	0.93	
Saltwater lakes	85	0.006	
Total salt water stocks	1 350 955	97.476	
Freshwater stocks			
Glaciers, permanent snowcover	24 064	1.74	68.7
Fresh groundwater	10 530	0.76	30.06
Ground ice, permafrost	300	0.022	0.86
Total frozen and underground freshwater stocks	*34 894*	*2.522*	*99.62*
Freshwater lakes	91	0.007	0.26
Soil moisture	16.5	0.001	0.05
Atmospheric water vapour	12.9	0.001	0.04
Marches, wetlands*	11.5	0.001	0.03
Rivers	2.12	0.0002	0.006
Incorporated in biota*	1.12	0.0001	0.003
Total not frozen or underground freshwater stocks	*135*	*0.0103*	*0.389*
Total freshwater stocks	35 029	2.5323	100
TOTAL WATER ON EARTH (1000 km³)	1 385 984	100	

Source: Adapted from Peter H. Gleick, *The World's Water 2000-2001*, Washington, D.C. Island Press, 2000.

Note: Totals may not add due to rounding

*Marshes, wetlands and water incorporated in biota are often mixed salt and fresh water.

Figure 2.

World's largest rivers

Rank	Based on drainage area		Based on length		Based on average annual total discharge	
	Name	Drainage area (1000 km²)	Name	Length (km)	Name	Average annual total discharge (km³/yr)
1	Amazon	6 915	Nile	6 670	Amazon	6 923
2	Congo	3 680	Mississippi*	6 420	Ganges	1 386
3	Murray	3 520	Amazon	6 280	Congo	1 320
4	La Plata	3 100	Yangtze	5 520	Orinoco	1 007
5	Ob	2 990	Mackenzie*	5 472	Yangtze	1 006
6	Mississippi*	2 980	La Plata	4 700	La Plata	811
7	Nile	2 870	Hwang Ho	4 670	Yenisei	618
8	Yenisei	2 580	Mekong	4 500	Lena	539
9	Lena	2 490	Lena	4 400	Mississippi*	510
10	Niger	2 090	Congo	4 370	Mekong	505
11	Amur	1 855	Niger	4 160	Chutsyan	430
12	Yangtze	1 800	Ob	3 650	Ob	404
13	Mackenzie*	1 790	Yenisei	3 490	Amur	360
			Murray	3 490		
14	Ganges	1 730	Volga	3 350	Mackenzie*	325
15	Volga	1 380	Indus	3 180	St. Lawrence*	318
16	Zambezi	1 330	St. Lawrence*	3 060	Niger	302
17	St. Lawrence*	1 030	Ganges	3 000	Volga	255
			Yukon*	3 000		

Source: Adapted from *World Water Resources and Their Uses*, Joint SHI/UNESCO Product, prepared by Prof. Igor A. Shiklomanov, 1999.

* Partly or entirely in Canada

Figure 3.

Figure 2 lists quantities that scientists have estimated for the various types of water that make up the world's supply. These amounts should be regarded as indicators of the relative quantities of water on earth. Owing to the difficulties in estimating volumes of water on a global scale, especially water underground, estimates can vary considerably. What is important is the overall picture that these estimates give.

RIVERS

A river's watershed or drainage basin—the area supplying it with water—is separated from the watersheds of neighboring rivers by higher lands called *drainage divides*. The map shows Canada's continental watersheds, one to each surrounding ocean: the Pacific, the Arctic, and the Atlantic as well as to Hudson Bay and to the Gulf of Mexico. Small watersheds combine to make up regional watersheds, which in turn join others to form continental watersheds.

The world's largest rivers are shown in Fig. 3.

Sculpting the Earth

As a swiftly flowing river, water can erode the underlying terrain. Where the river slope is flatter, the river slows down and deposits materials. This usually occurs in the lower reaches and especially near the mouth of the river, either at a lake or an ocean. A river can carve steep valleys, especially in higher parts of the drainage basin.

In the lower parts of the basin, deposits may create deltas at the river's mouth.

The volume of water flowing in a river, together with the speed and timing of the flows, determines how a river shapes the surrounding landscape and how people can use its waters. Rainfall, snowmelt, and groundwater all contribute to the volume of flow, producing variations from season to season and year to year.

In Canada, most high flows are caused by spring snowmelt. This is the season when floods are most likely to occur. Rainstorms can also cause high flows and floods, especially on small streams. The effects of floods and storms can be much less severe on rivers with large drainage basins. The lowest flows for rivers in Canada generally occur in late summer, when precipitation is low and evaporation along with consumption by plants is high, and in late winter, when rivers are ice covered and the precipitation is stored until spring in the form of ice and snow.

LAKES

Canada has more lake area than any other country in the world (Fig. 4), with 563 lakes larger than 100 square kilometers. The Great Lakes, straddling the Canada–U.S. boundary, contain 18% of the world's fresh lake water.

How is Water Measured?

The Water Survey of Canada, Environment Canada, along with many contributing agencies, measure the rate of flow (discharge) in rivers and record the levels of lakes and rivers at more than 2,600 locations in Canada. Typical river flows are listed in Fig. 5.

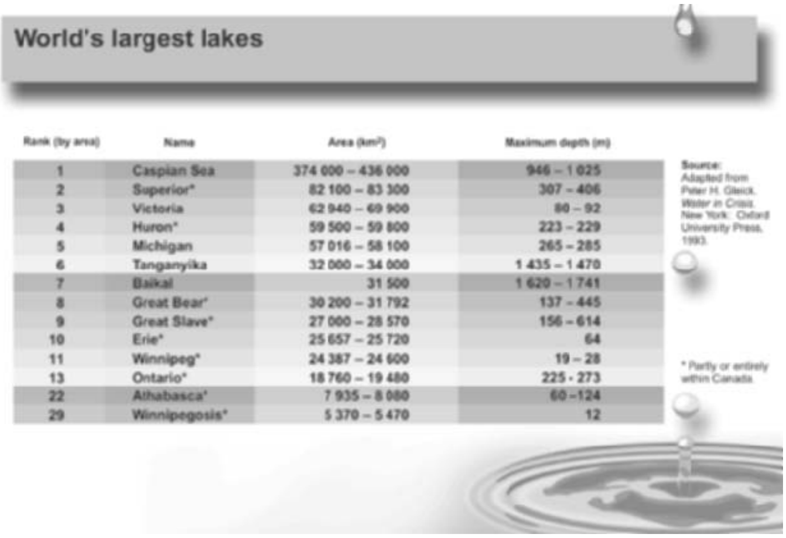

World's largest lakes

Rank (by area)	Name	Area (km²)	Maximum depth (m)
1	Caspian Sea	374 000 – 436 000	946 – 1 025
2	Superior*	82 100 – 83 300	307 – 406
3	Victoria	62 940 – 69 900	80 – 92
4	Huron*	59 500 – 59 800	223 – 229
5	Michigan	57 016 – 58 100	265 – 285
6	Tanganyika	32 000 – 34 000	1 435 – 1 470
7	Baikal	31 500	1 620 – 1 741
8	Great Bear*	30 200 – 31 792	137 – 445
9	Great Slave*	27 000 – 28 570	156 – 614
10	Erie*	25 657 – 25 720	64
11	Winnipeg*	24 387 – 24 600	19 – 28
13	Ontario*	18 760 – 19 480	225 - 273
22	Athabasca*	7 935 – 8 080	60 –124
29	Winnipegosis*	5 370 – 5 470	12

Source: Adapted from Peter H. Gleick, Water in Crisis. New York: Oxford University Press, 1993.

* Partly or entirely within Canada.

Figure 4.

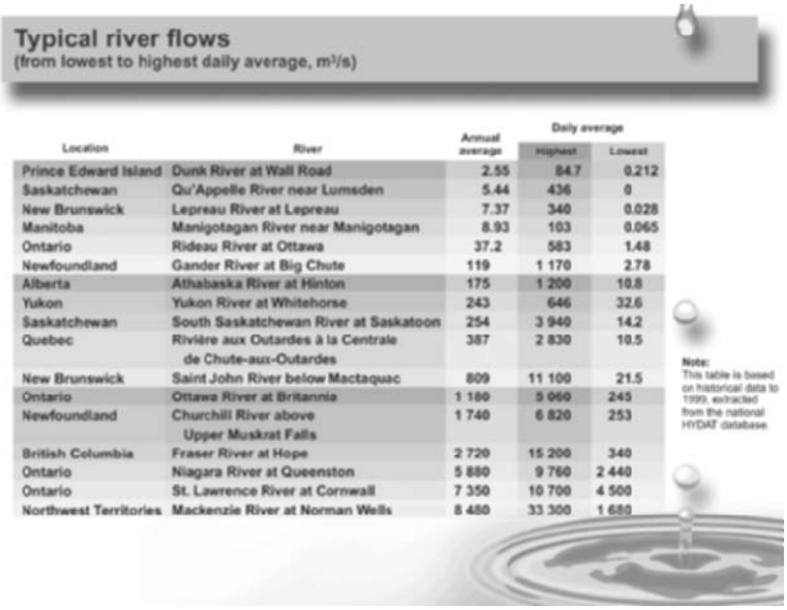

Typical river flows
(from lowest to highest daily average, m³/s)

Location	River	Annual average	Daily average Highest	Daily average Lowest
Prince Edward Island	Dunk River at Wall Road	2.55	84.7	0.212
Saskatchewan	Qu'Appelle River near Lumsden	5.44	436	0
New Brunswick	Lepreau River at Lepreau	7.37	340	0.028
Manitoba	Manigotagan River near Manigotagan	8.93	103	0.065
Ontario	Rideau River at Ottawa	37.2	583	1.48
Newfoundland	Gander River at Big Chute	119	1 170	2.78
Alberta	Athabaska River at Hinton	175	1 200	10.8
Yukon	Yukon River at Whitehorse	243	646	32.6
Saskatchewan	South Saskatchewan River at Saskatoon	254	3 940	14.2
Quebec	Rivière aux Outardes à la Centrale de Chute-aux-Outardes	387	2 830	10.5
New Brunswick	Saint John River below Mactaquac	809	11 100	21.5
Ontario	Ottawa River at Britannia	1 180	5 060	245
Newfoundland	Churchill River above Upper Muskrat Falls	1 740	6 820	253
British Columbia	Fraser River at Hope	2 720	15 200	340
Ontario	Niagara River at Queenston	5 880	9 760	2 440
Ontario	St. Lawrence River at Cornwall	7 350	10 700	4 500
Northwest Territories	Mackenzie River at Norman Wells	8 480	33 300	1 680

Note: This table is based on historical data to 1999, extracted from the national HYDAT database.

Figure 5.

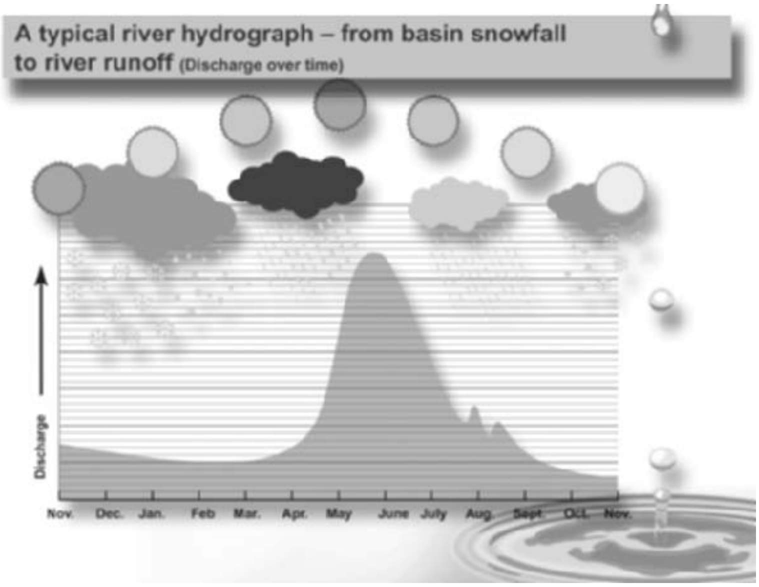

Figure 6.

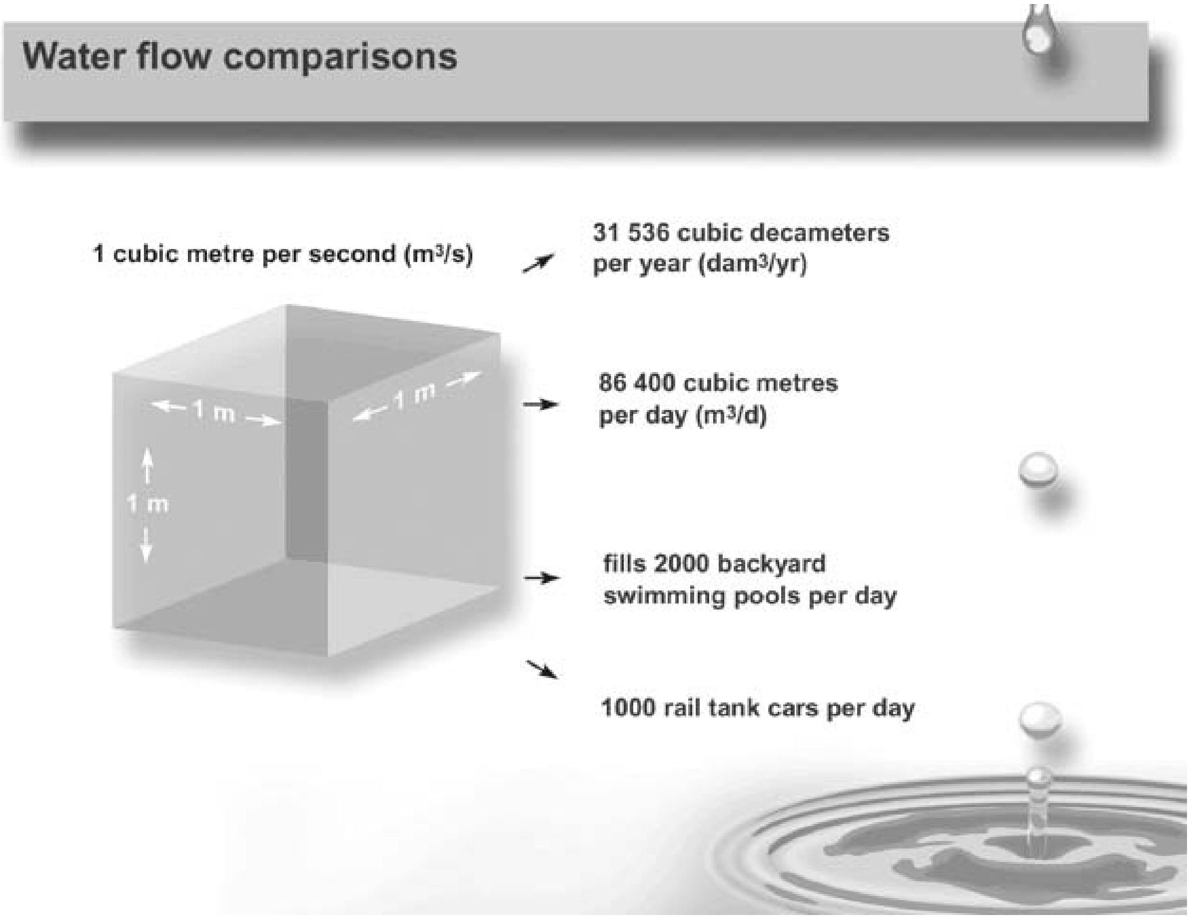

- Water levels are read manually by gauge readers or continuously recorded either digitally or on graph paper.
- Rate of flow (or discharge) requires multiple measurements of channel depth, width, and flow velocity to yield the average discharge in the stream crossing for a given water level. Measurements can be made from a bridge, by wading in a stream, by boat or from a cableway strung across the river. In winter, the measurements are made through the ice.

- With sufficient measurements of flow over a variety of water levels (including extreme lows and highs), a water level-discharge relationship is established at each location. The discharge rate can then be computed from measured water levels.
- Historical records from 5,000 active and discontinued sites permit the estimation of streamflow at ungauged locations.

Creating a Balance—Naturally

The importance of lakes lies in their ability to store water during times of plenty and release it gradually. Thus lakes perform an extremely valuable task in balancing the flow of the rivers on which they are located. For example, the Saskatchewan River, with few lakes, has a maximum recorded flow of 59 times its minimum flow. On the other hand, the St. Lawrence River, which drains the Great Lakes, has a maximum flow of only twice its minimum flow. The difference in flow patterns in these two rivers is partly due to precipitation differences, but results mainly from the vast storage provided by the Great Lakes for the St. Lawrence River compared with the negligible lake storage on the Saskatchewan River.

Creating a Balance—Artificially

Since ancient times, people have built dams to control the outflow from existing lakes or to create new lakes. Dams and their reservoirs have provided:

- A stable source of inexpensive energy
- A more dependable water supply throughout the year
- Flood control downstream
- Recreation

THE UNDERGROUND RESERVOIR

Beneath the surface of the earth is a huge reservoir of fresh water. Groundwater does not rest; it moves continuously, but at a snail's pace, from its point of entry to areas of natural discharge. Groundwater moves so slowly that its speed is measured in meters per day, and even per year. (Surface water velocities are described in meters per second.) Wells intercept some groundwater but most of it continues until it reappears naturally in a spring or a seepage area and joins a watercourse.

Groundwater contributes to Canada's water supply by:

- Feeding streams, producing the entire flow of some streams during dry periods
- Replenishing wells—a valuable source of supply for individuals, communities, industries, and irrigated farms
- Supporting important ecological systems such as wetlands
- Moderating the adverse impacts of acid rain on surface water systems

Additional information on groundwater can be found in: Freshwater Series No. A-5, *Groundwater—Nature's Hidden Treasure*.

NATURE'S FROZEN RIVERS

A huge quantity of fresh water is frozen in polar ice caps and in high mountain glaciers. Snow that is packed down over many years at high elevations becomes glacial ice, which slowly proceeds downslope like a frozen river, under the pull of gravity, and eventually melts to become part of streamflow at lower elevations. If the rate of melting is greater than the rate of accumulation, the glacier recedes; if it is less, the glacier advances.

Glaciers exert a direct influence on the hydrologic cycle by slowing the passage of water through the cycle. Like lakes and groundwater reservoirs, glaciers are excellent natural storehouses, releasing water when it is needed most. Glaciers, however, can release water when you need it least. Glacier-outburst floods, called *jökulhlaups*, can be devastating. Glacier-fed rivers reach their peak during hot summer weather.

SNOWFALL

Much of Canada's annual precipitation comes as snow: in the North, 50%; in the Prairies, 25%; and on both coasts and in southern Ontario, as little as 5%. Snow exerts a marked effect on the distribution of streamflow throughout the year. Instead of immediately infiltrating the soil or running off into stream channels as rainfall does, this water is first stored in the snowpack for several months.

The relatively quick melting of snow in spring causes peak flows, sometimes resulting in floods. Some of the worst and most unpredictable flooding occurs when ice that has not yet melted is carried along in the swollen rivers until it jams, blocking the flow of water and creating a lake behind the jam with attendant flooding. When the ice jam breaks, a tremendous amount of water is suddenly released downstream, and more flooding may result.

FRESHWATER SERIES A-2

Note: A resource guide, entitled *Let's Not Take Water For Granted*, is available to help classroom teachers of grades 5–7 use the information from the Water Fact Sheets.

WATER CONSERVATION—EVERY DROP COUNTS IN CANADA

Environment Canada

The importance of protecting our water resource cannot be overstated. In economic terms, the measurable contribution of water to the Canadian economy is difficult to estimate. In environmental terms, water is the lifeblood of the planet. Without a steady supply of clean, fresh water, all life, including human, would cease to exist.

This article is produced by Environment Canada and is reproduced with permission.

The perception that Canada is blessed with an abundance of fresh water has led to misuse and abuse of the resource: from household toilets that use 18 liters per flush where 6 liters would do, to industrial plants—and some municipalities—that use water bodies as convenient sewers.

In 1999, the average Canadian daily domestic use of fresh water per capita was 343 liters. The quantity, quality and economic problems we face as a result of our use of water are complex but, at least *one* of the causes of these problems is easy to manage—the way we waste water. And, the solution is straight forward—*water conservation*. Simply stated, water conservation means doing the same with less, by using water more efficiently or reducing where appropriate, in order to protect the resource now, and for the future. Using water wisely will reduce pollution and health risks, lower water costs, and extend the useful life of existing supply and waste treatment facilities.

And it's easy. With little change to the way we do things now, or the equipment we use, we can reduce water consumption in the home, and in business, by 40% or more. These pages outline the role of water conservation in addressing problems related to water use and water quality. It also shows us what part we can play as residential consumers in finding solutions.

SUSTAINING OUR WATER SUPPLY

Water is considered a renewable resource: "renewable" referring to that portion which circulates back and forth in the hydrological cycle (described more fully in Freshwater Series A-1, *Water—Nature's Magician*). However, pressures on the resource are growing. For example, between 1972 and 1996, Canada's rate of water withdrawals increased by almost 90%, from 24 billion m^3/yr (cubic meters per year) to 45 billion m^3/yr. But, our population increased by only 33.6% over the same period, illustrating the growth in our thirsty lifestyles. As the readily available supplies of fresh water are being used up, we begin to see that there are real limits to how much water we can count on.

LIMITING FACTORS

- Although Canada has a significant amount of fresh water, we possess only 7% of the world's *renewable* freshwater supply.
- In Canada, 84% of the population lives in a narrow southern band, while 60% of our water supply flows north to the Arctic Circle.
- Our growing population, and our growing thirst for water, are being concentrated in expanding metropolitan areas, and are forcing water regulators and policy makers to find ways to stretch available supplies even further.
- Increasing pollution of surface and groundwater is further reducing the supplies of readily available, clean water.
- Because our water use almost always leads to some degree of deterioration in water quality, the less water we withdraw, the less we upset the natural balance of our aquatic ecosystems. And, the less we upset the ecosystem, the less we have to spend to restore the water quality to an acceptable standard for public use.
- Finally, financing by municipal governments for the treatment of water supplies and wastewater is becoming increasingly constrained.

We can, however, make a significant contribution to solving these problems by reducing unnecessary levels of water use. To do so requires that we identify the areas within our homes, businesses, buildings and processes where we waste water and then make appropriate changes, either in our fixtures, or in our water-using habits.

HOW WATER IS USED

We use water in many ways, and assign different values to those uses. *Instream* uses (e.g., for transportation and recreation) are valued highly, but it has proven difficult to give them a dollar value that has any real meaning. For example, just what would the average consumer be willing to pay to swim in a clean lake or for a chance to catch fish in a clean, unpolluted river?

By far the greater number and variety of water uses occur on land. These are called *withdrawal* uses and, although important to our daily lives, they have tended to be assigned a low value. Water is withdrawn, used and then discharged. Most withdrawal uses "consume" some of the water, which means less is returned to the source than was taken out. And, after it has been used, the quality of the water that is returned is often diminished, which has a negative impact on both the environment and recreational instream uses.

In 1996, five main withdrawal uses accounted for a total annual water intake (extractive uses) of 44.6 billion m^3. These uses are described more fully in Freshwater Series A-4, *Water Works*!.

Thermal power generation includes both conventional and nuclear power generating plants, which withdrew slightly more than 64% of the total water intake in 1996.

Manufacturing accounted for 14% of water withdrawals in 1996. Paper and allied products, primary metals and chemicals were the main industrial users.

Agriculture accounted for nearly 9% of total withdrawals, with the semi-arid Prairie region of Canada accounting for 75% of this total. Agriculture consumes a large portion of what it uses, returning less than 30% to its source where it can be used again. Irrigation is the largest agricultural consumer of water.

Municipal use accounted for 10% of all water withdrawals in 1996, or 12% when similar rural uses were included (excludes industrial uses and large-scale agriculture). In the municipal sector, more than half of the water demand is a result of residential use.

Mining use, including metal mining, non-metal mining, and the extraction of coal, accounted for 1% of all water withdrawals in 1996. Water is used by the mining industry to separate ore from rock, to cool drills, to wash the ore during production, and to carry away unwanted material.

Solutions—the Municipal Challenge

Municipal governments across Canada are beginning to take action to manage the demand for water, instead of seeking new sources of supply. Demand management, incorporating water efficient applications, is rapidly gaining popularity as a low cost, effective way to get more service out of existing systems, thus delaying or deferring the need for constructing new works. The benefits of water efficient techniques apply equally well to rural, private wells and septic disposal systems, as they do to central water and sewer systems in the city.

The wide range of water efficiency initiatives currently being undertaken, can be grouped under four principal categories:

1. Structural,
2. Operational,
3. Economic, and
4. Socio-political.

Most of these water conservation activities fall within the jurisdiction of municipal governments and/or public utilities.

Structural

- metering
- water recycling systems
- wastewater re-use
- flow control devices
- distribution system pressure reduction
- water saving devices (efficient fixtures, appliances and retrofits)
- drought resistant landscaping (xeriscaping)
- efficient sprinkling/irrigation technology
- new process technologies
- plant improvements

Operational

- leak detection and repair
- water use restrictions
- elimination of combined sanitary/storm sewers to reduce loadings on sewage treatment plants
- plant improvements

Economic

- rate structures
- pricing policies
- incentives through rebates and tax credits
- other sanctions (fines)

Socio-political

- public education
- information transfer and training
- regulatory (legislation, codes, standards and by-laws)

SUSTAINING OUR INFRASTRUCTURE

While many communities have access to an abundant water supply, the costs of the infrastructure that provides homes and industry with water and sewer services are straining the available municipal financial resources.

By *infrastructure*, we mean the *water treatment plants* that purify our water, the *water mains* in the ground that transport water, and the *towers* and *reservoirs* that store water. The term includes the *sewer pipes* that carry away wastewater and the *sewage treatment plants* that treat wastewater before returning it to the environment where it often becomes the source of water for communities downstream.

Experts are predicting a growing problem involving municipal water and sewer infrastructure in Canada. In 1991, the value of this investment was estimated to be worth over 90 billion dollars, of which a significant amount is deteriorating with age.

An increasing number of Canadian municipalities are considering water conservation as the key to keeping expansion needs to a minimum. Water conservation also optimizes plant efficiency, while assisting municipalities in financing the replacement of infrastructure that may be over 50 years old in some communities and up to 100 years old in several others.

Communities with older systems in need of extensive repairs or replacement face the most difficult problems.

With all levels of government adopting policies of realistic water pricing and user pay principles, many municipalities have instituted *full cost pricing* to recover the total cost of providing both water and sewer services—including the costs of financing the replacement of older systems and the upgrading of overloaded treatment plants. Higher municipal costs, in turn, mean higher water—and sewer—bills.

The problem of stressed treatment systems is not restricted to communities with piped water and sewer systems. Over the past 25 years, there has been a substantial migration of urban dwellers to the countryside. City-bred water using habits and attitudes are, in many instances, lowering the water table. And, the flood of wastewater produced is stressing the soil's ability to treat septic effluent adequately.

For both urban and rural communities, water conservation can extend the life of this over-stressed infrastructure.

Metering

Tied to price increases, metered households generally show reductions in water use, with the greatest savings occurring during the summer months, when water use is usually much higher due to frequency of lawn watering, car washing and other outdoor uses. In 1999, water use was 70% higher when consumers faced flat rates rather than volume-based rates. And yet, only about 56% of Canada's *urban* population was metered in 1999.

Metering of industry has been common for some time. What's new is the metering of the return flow to the sewer system, particularly as it relates to the industrial sector. Case studies show that including sewage treatment in rate calculations generates greater water savings. An increasing number of municipalities are applying sewer surcharges to residential water bills.

Water Efficient Residential Technology

More than half of municipal water is used by the residential sector. As a consequence, the residential sector represents a logical target for demand management activities. Depending on the nature of the water efficiency program developed, each household can reduce water use by 40% or more.

Leak Detection and Repair

Up to 30% of the total water entering supply-line systems is lost to leaking pipes. In most cases, if unaccounted for water in a municipal system exceeds 10 to 15%, a leak detection and repair program is cost-effective. For example, studies have shown that for every $1.00 spent in communities with leak detection programs, up to $3.00 can be saved.

Rates, Pricing and Public Education

About 55% of Canadians served municipal water pay in ways that do not promote conservation. A 2001 study of rate structures by Environment Canada showed that in 1999, 43% of the population was under a *flat rate* structure (where the charge or assessment is fixed, regardless of the amount of water used). Another 12% were under a *declining block rate* structure (where the consumer's bill rises at a slower rate as higher volumes of water are used); i.e., the more you use, the less you pay per unit.

Only about 45% of the population served was found to be under a rate structure that provided a definite incentive to conserve water: 36% were under a *constant rate* structure (where the bill to the consumer climbs uniformly with the volume used); and 9% were under an *increasing block* rate structure (where a successively higher price is changed as larger volumes of water are used).

Introducing conservation-oriented pricing or raising the price has reduced water use in some jurisdictions, but it must be accompanied by a well articulated public education program that informs the consumer what to expect.

SUSTAINING OUR WATER QUALITY

In addition to water supply and infrastructure issues, water quality is a problem in many Canadian communities. Generally speaking, the decline in water quality is a function of the way we use water. Even something as simple as rinsing dishes in the kitchen creates wastewater that is contaminated to some degree. Once this water enters the sewer system, it must be treated in a sewage treatment plant. These facilities are never 100% effective, which means that some water quality deterioration remains after the treatment process.

Specific causes of impaired water quality are numerous, including: agricultural runoff containing the residues of fertilizers, pesticides and other chemicals, industrial pollution, either directly from the facility, or indirectly from the leaching of chemicals from landfills, or pollution from average households in the form of improperly treated municipal sewage (refer to Freshwater Series A-3, *Clean Water—Life Depends on It*). Nearly 75% of Canadians are serviced by municipal sewer systems. In 1999, 97% of the Canadian population on sewers received some form of wastewater treatment. The remaining 3% of Canadians served by sewage collection systems were not connected to wastewater treatment facilities in 1999 and discharged their untreated sewage directly into receiving water bodies.

For the roughly 25% of the Canadian population served by private wells and septic disposal systems, the news is not much better. These systems were originally designed for houses that were widely separated from their nearest neighbor, such as farmhouses and the occasional rural residence. Yet, today, in many parts of the country, individual private wells are being installed in subdivisions at suburban densities. The primary danger here is that too many wells may pump too much water for the aquifer to sustain itself.

Septic treatment systems associated with these developments can stress the environment in a number of other ways. They are often allowed in less than satisfactory soil conditions and are seldom maintained properly. They are also unable to treat many household cleaners and chemicals which, when flushed down the drain or toilet, often impair or kill the bacterium needed to make the system work (The same applies in urban systems). The end results are improper treatment of wastewater—if not outright failure of the system—and the contamination of adjacent wells with septic effluent containing bacterium, nitrates and other pollutants.

Once these contaminants are in the groundwater, they eventually reach rivers and lakes. In other words, once we have a *pollution* problem, we may be only a step away from a *water supply* problem.

Solutions—in the City and in the Countryside

The irony in all of this is that water quality impacts from overloaded or poorly maintained and operated municipal and private sewage disposal systems are the number one preventable type of pollution in Canada. The answer lies in better, more thorough treatment. And, one of the ways to enhance the treatment process is to limit the amounts of wastewater entering the wastewater stream. Again, water conservation is one of the easiest and cheapest ways to reduce the volume of wastewater flows and improve water quality.

Following a few common sense rules, it should be possible to safeguard your water supply while extending the life of your sewage disposal system, regardless of whether you live in the country or in the city.

Think carefully about the quantities of wastewater your household or business produces, as well as the quality of the wastewater. Do you make it a habit of discarding solvents, cleaners and related chemicals down your drains? If you do, you may be introducing substances that are toxic to the bacterium and other organisms that play a vital role in the treatment of sewage. This statement applies equally well to urban and rural households and businesses.

INDIVIDUAL ACTION—CONSERVING WATER IN THE HOME, COMMUNITY AND AT WORK

As we have seen, water quality and quantity are two sides of the same coin. How does saving water help water quality? Because water saved is water that does not end up in the wastewater stream requiring treatment. This, in turn, reduces municipal pumping and treatment costs and frees up monies that can be used for infrastructure renewal and replacement and protection of supply sources. Less wastewater in the sewage treatment plant also means that the plant has a better chance of doing the job it was intended to do.

So where do we start? The first step is to identify where we use water in the home. Then we need to decide on what to do to reduce the amount of water we use, either by eliminating wasteful practices and habits, or by improving the efficiency of our water using fixtures and devices. Since we waste so much, this should be a relatively easy and painless process. The prime area to target is the bathroom, where nearly 65% of all indoor water use occurs.

What follows are some suggestions for how to get your house or business in order. Based on the three rules of water conservation—*reduce, repair* and *retrofit*—a typical household can reduce water consumption by 40% or more, with or no effect on lifestyle.

REDUCE

Much of the water “consumed” in our daily activities is simply wasted. Taps are left running while we brush our teeth. Dishwashers and laundry machines are operated without full loads. Really, everywhere we use water there is room for improvement. Here are just a few examples for both indoor and outdoor water use.

- Don’t use the toilet as a wastebasket or flush it unnecessarily.
- A quick shower uses less hot water than a full tub (and saves energy too).
- Keep a bottle of drinking water in the refrigerator rather than letting your tap run to get cold water when you want a drink. (Rinse the bottle every few days.)
- More than 50% of the water applied to lawns and gardens is lost due to evaporation, or run-off because of overwatering. Find out how much water your lawn really needs. As a general rule, most lawns and gardens require little more than 2 to 3 cm (1 inch) of water per week.
- To reduce loses due to evaporation, water early in the morning (after the dew has dried).
- When washing a car, fill a bucket with water and use a sponge. This can save about 300 liters of water.

REPAIR

Leaks can be costly. A leak of only one drop per second wastes about 10,000 liters of water per year. Most leaks are easy to find and to fix, at very little cost.

- Leaking faucets are often caused by a worn out washer that costs pennies to replace. Most hardware stores will have faucet repair kits with illustrations showing how to replace a washer.
- A toilet that continues to run after flushing, if the leak is large enough, can waste up to 200,000 liters of water in a single year! To find out if your toilet is leaking, put two or three drops of food coloring in the tank at the back of the toilet. Wait a few minutes. If the colour shows up in the bowl, there’s a leak.
- Toilet leaks are often due to a flush valve or flapper valve that isn’t sitting properly in the valve seat, bent or misaligned flush valve lift wires, or a corroded valve seat. All of these can be fixed easily and inexpensively. To get at the valve seat, which surrounds the outlet hole at the bottom of the tank,

you must first empty the tank. This is accomplished by turning off the inlet tap under the tank and flushing the toilet, making sure to keep depressing the flush lever until no more water drains out of the tank. Then, holding the valve out of the way, sand the corroded or warped valve seat smooth with a piece of emery cloth, if, however, the leak is around the base of the toilet where it sits on the floor, call a professional.

RETROFIT

Retrofit means *adapting* or *replacing* an older water-using fixture or appliance with one of the many water-efficient devices now on the market. While these solutions cost more, they also save the most water and money. Retrofitting offers considerable water saving potential in the home and business, so this issue in the Freshwater Series is devoting considerable space to it.

Toilet Retrofits

When it comes to retrofitting, the prime fixture to target is the toilet. You can: (i) *adapt* your existing toilet in a number of ways, by installing certain water-saving devices inside the tank at the back of the toilet; or, (ii) if the toilet is more than fifteen years old—which means it probably uses about 18 or more liters of water per flush—you can *replace* it with one of the growing number of ultra-low-volume (ULV) toilets, that can be ordered from most plumbing outlets, and use only 6 liters or less per flush.

There are many *toilet adaptations* you can install in the tank of an existing toilet to reduce the amount of water used in a flush cycle. These devices fall into three generic categories:

- water retention devices;
- water displacement devices; and,
- alternate flushing devices.

The most common water retention device available is the toilet dam. A set will save about 5 liters per flush when installed properly. Their main attraction is their low cost (under $10.00 per set) and the fact that they are easy to distribute and install for example, as part of a wider municipally-sponsored retrofit program. Their main disadvantage is that they tend to leak over time by slipping out of adjustment and can slip free and interfere with the moving parts inside the toilet tank, if not routinely checked.

Toilet dam and displacement bag:

The water displacement devices familiar to most people are the *plastic bags* or *bottles* filled with water which are suspended inside the toilet tank. As the name implies, these devices displace several liters of water, saving an equivalent amount during each flush. Like the toilet dam, most displacement devices are inexpensive and easy to install. Their chief disadvantage is that they don't save as much water as other devices and, if they are not installed carefully, they can interfere with the proper operation of the toilet.

One displacement device to stay away from is the *brick*! It can disintegrate inside the toilet tank, leading to excessive leakage at the flapper valve and may even be heavy enough to actually crack the tank.

There are essentially two types of alternative flush devices: *early-closure* and *dual-flush*. They are usually attached to the overflow tube inside the toilet tank. In both cases, they close the flush valve or flapper after the tank is only partially emptied. In theory, this interruption in the flush cycle occurs after the bowl has been cleared. In the case of the dual-flush mechanism, the amount of water saved is dependent upon how long the flush lever is activated—a partial flush for light duty or full flush or heavy duty.

While all of the above toilet adaptations appear to work as intended when first installed, their performance may vary considerably, depending on the toilet design. The best advice is to monitor the performance of the devices periodically. If you discover that it becomes necessary to double flush the toilet, something is in need of adjustment or replacement. Remember that double flushing defeats the purpose of your water conservation efforts and is costing you money.

If you decide that it is time for a *toilet replacement* in your home or business, you are well on your way to significant water savings that you can bank on over the life of the toilet. Replacing a 18 liter per flush toilet with an ultra-low-volume (ULV) 6 liter flush model represents a 70% savings in water flushed and will cut indoor water use by about 30%. Keep in mind that 18 liters per flush, assuming 4.5 flushes per person per day, translates into nearly 30,000 liters of clean, fresh water per year just to get rid of 650 liters of body waste. A 6 liter flush toilet only uses about 10,000 liters to do the same task. Low flush toilets are available for less than $150.00 at most plumbing and supply stores.

Remember, the ULV toilet not only uses less water, *it produces less wastewater*. If your municipality applies a *sewer surcharge* on your water bill, the investment in the better toilet could translate into a 50% reduction in your combined water/sewer bill. If you are on a private well and septic system, you are significantly reducing the loading on your tile field while extending its useful life. To a lesser degree, the same applies to the other water-saving devices described here.

Showerheads and Faucets

After the toilet, the shower and bath consume the most water inside the home. Conventional showerheads have flow rates up to 15 to 20 liters per minute. A properly designed *low-flow showerhead* can reduce that flow by half and still provide proper shower performance. Low-flow showerheads can be purchased in most plumbing supply outlets.

Depending on your preference for finish and appearance, you can select a serviceable low-flow showerhead starting at around ten dollars. Consider one with a shut-off button. The advantage of the shut-off button is that it allows you to be really water efficient if you so choose, by being able to interrupt the flow, while you lather up or shampoo, and then resume at the same flow rate and temperature.

Beware of the type of showerheads that produce such a fine mist that the water is quite cool by the time it reaches your feet. And, stay away from so-called *flow restrictors* that are inserted inside your existing showerhead. They look like a small plastic washer and can produce a fierce, stinging spray pattern which may significantly reduce the enjoyment of taking a shower.

Conventional faucets have an average flow rate of 13.5 liters of water per minute. Install *low flow aerators* to reduce this flow. In the bathroom, a flow rate of 2 liters per minute should do the trick, and in the kitchen a flow rate of 6 to 9 liters per minute is sufficient. Don't bother retrofitting the tap in the utility sink; it is intended to provide large volumes of water quickly, for example, for cleaning or washing, such that low flows will only inconvenience the user.

Outdoors

During the growing season water use can increase by as much as 50%. While lawns require a lot of water, much of this water is wasted—lost due to overwatering and evaporation (see Reduce).

Watering equipment also plays a part in how much water is saved and lost. Ideally, sprinklers should be suited to the size and shape of the lawn. That way, you avoid watering driveways and sidewalks. Installing timers on outdoor taps can be a wise investment.

Sprinklers that lay water down in a flat pattern are better than oscillating sprinklers which lose as much as 50% of what they disperse through evaporation. Drip irrigation systems which apply water only to the roots zone are the most efficient—and the most expensive—alternative.

The water you use to water your lawn doesn't have to come out of a tap. A cistern, which captures and stores rainwater, can be used as a source of irrigation water. A rain barrel can adequately fulfil this function.

Finally, consider a low-maintenance landscape—one which requires little more water than nature provides. Often called xeriscaping, the principles of a low-maintenance landscape are as follows:

- a reduced amount of lawn;
- proper plan selection making use of native grasses, shrubs and trees;
- mulching to reduce evaporative losses around shrubs and trees;
- the use of rain barrels and roof drainage;
- improvements to soils;
- a proper irrigation system; and
- planned maintenance.

The most significant savings of course, come from a reduction in lawn area and switching from exotic plant forms to native species which require less water. In general, lawn areas should not exceed what is useful for play and social activities, and should be limited to the backyard where the family spends the majority of its time.

Solutions—At Work and in the Community

Many of the suggestions made for reducing water use in the home have wider application, both in the workplace, and in the community at large. Low-flow equipment are available for most commercial and toilet applications, instituting them may mean taking a leading role yourself, for example, forming and leading a committee that would address the following questions:

- do your workplace bathrooms, kitchens, etc. have water-efficient toilets, faucets, etc. similar to those discussed for the household?
- if your workplace uses water in its production process or for washing goods or equipment, is this being done efficiently?
- does your community have a water-efficiency assistance program that helps households and business improve their water-use efficiency?
- is the water distribution system properly maintained so that no pollution leaks into it and so that no water is wasted through leaky mains?

THE BOTTOM LINE

Water conservation. The message is clear. If we each save a little, it can add up to major savings in water, energy and money. For the average household, reductions in water use as high as 40% or more are feasible, just by following the steps outlined in this issue in the Freshwater Series.

The benefits don't stop at the household or business. The municipal water and sewer department gets a break on the amount of water it has to pump to our homes and businesses *and* on the amount of wastewater it has to treat in sewage treatment plants. Water conservation can extend the useful life of municipal water supply and treatment plants, and will benefit the operating efficiency—and life expectancy—of private septic disposal systems.

And, finally, water conservation can generate *significant* environmental benefits. It can reduce water diverted and the pollution loadings on our lakes and rivers by reducing the volumes of wastewater which we have to treat. This can help to protect our drinking water and the ecological balance in sensitive aquatic ecosystems.

If we all practice water conservation, everyone—and everything—benefits.

FRESHWATER SERIES A-6

Note: A resource guide, entitled *Let's Not Take Water For Granted*, is available to help classroom teachers of grades 5–7 use the information from the Water Fact Sheets.

ECOREGIONS: A SPATIAL FRAMEWORK FOR ENVIRONMENTAL MANAGEMENT

MD. SALEQUZZAMAN
Khulna University
Khulna, Bangladesh

DEFINITION

An ecoregion may be defined as a scale for planning and analysis in the hierarchy of the earth's ecosystem. It has broad applicability for modeling and sampling, strategic planning and assessment, and international planning of domain, division, and provincial ecological units.

Ecoregionization is a process of delineating and classifying ecologically distinctive areas of ecological land. Each area can be viewed as a discrete system that has resulted from the mesh and interplay of the geologic, landform, soil, vegetative, climatic, wildlife, water, and human factors where ecological functions and processes are continuing. The dominance of any one or more of these factors varies with the given ecological land unit. This holistic approach to land classification can be applied incrementally on a scale-related basis from site-specific ecosystems to very broad ecosystems (1). Ecological processes, evolutionary mechanisms, and geological forces are continually reshaping landscapes across various scales of time and space and result in distinctive but dynamic ecoregions. All of the world's food and most medicines and raw materials are derived from these processes and associated biodiversity. Thus ecoregions gain their identity through spatial differences in a combination of landscape characteristics. Several factors such as topography, hydrology, and nutrients are important to identify these characteristics that may vary from one place to another in an ecoregion.

OBJECTIVE OF ECOREGIONIZATION

The main objective of using the ecoregion as a spatial framework for environmental management is to maximize ecoefficiency by minimizing resource-use conflicts and harmonizing sociocultural needs in the ecoregion. Ecoregions are also used to isolate areas for the interpretation of environmental quality (2). Countries such as Canada and the United States are now using the ecoregion as the coarsest scale for evaluating representative areas (3). This process needs lots of information. The baseline information on ecosystem properties and changing conditions of ecosystems in any specific ecoregion may be collected through a series of monitoring, which affects the management practices in that ecoregion. Knowledge of the ecoregion facilitates; the allocation or linking of sites to a standard ecological hierarchy minimizes the sampling variance and increases the ability to extrapolate results for areas with similar properties.

IMPORTANCE OF THE STUDY OF AN ECOREGION

Ecological integrity, natural capital, and biodiversity are three ecological requirements for environmental management in any ecosystem (4). These three factors should be present as a prerequisite for a healthy ecoregion because an ecoregion itself is a big ecosystem. Among these factors, biodiversity has global importance and combines wildlife value with a rich cultural heritage. Thus each ecoregion has an immense natural value with a colorful and rich cultural heritage bound up with the unique natural heritage of age-old cultural traditions. As each ecoregion is ecologically distinct and strikingly beautiful, an ecoregion attracts the ecotourism that has a great potential to improve the local economy as well as develop a global network for ecobusiness and globalization.

SOME IMPORTANT ISSUES FOR ENVIRONMENTAL MANAGEMENT IN THE ECOREGIONS AROUND THE WORLD

Ensure a Balanced Ecosystem

Every organism, including animals and plants, no matter how small they are, plays a vital role in keeping the systems of the earth as healthy and functioning ecoregions. Removing any species from the complex web of life disturbs the balance of the entire system, resulting in the impoverished life zones left today. Scientific studies have made it clear that a balanced ecosystem is teeming with abundant life in quality and quantity, such as a wetland ecosystem or the ocean. Historical evidence clearly suggests that once the oceans were filled with whales, sea turtles, fish, and other forms of life that seem more identical to a scientific fantasy than reality. Many of them are now extinct. The reality is that the actual populations were exponentially greater than they are today (before human predation began). Once humans began hunting, their appetites were insatiable, and a pattern of imbalance started that may culminate in the complete collapse of some ecosystems of the ecoregion. Thus one of the important aspects of ecoregionization is to keep the ecosystem balance in any ecoregion.

Ensure the Protection of the Natural Heritage and the Environment

The natural heritage and its environment are the identity of any ecoregion, which are also the natural wealth that sustained human communities in comfort and plenty for millennia, such as the rain forest in many countries of the world, for example, Alaska. Recently, the Alaskan rain forest is being destroyed in such a way that the region faces a graver, more permanent threat: the same razed-earth logging that has already devastated the Pacific Northwest. Rain forest logging is perfectly in step with the boom-and-bust rhythm of the Alaskan economy due to poorly planned development spread out over wildlands, slashing of shorelines for beach homes and resorts, limiting public access, and destroying fragile barrier beaches and wetlands. Besides this destruction, it is also destroying a lot of wildlife habitats as well.

Ensure the Protection of the Ethnic Community of the World

Each ecoregion has a special characteristic of the specific group of ethnic community people. They preserve a sustainable livelihood from generation to generation. Now, this sustainability livelihood is challenged for survival. The ethnic community people attempt overexploitation of natural resources such as fish, oil, gas, forest products, and minerals to satisfy their mounting needs induced by globalization and modernization. Recently, some nations have agreed to monitor environmental damage in their ecoregion, establish an emergency response program, protect the sensitive environment, and conserve endemic flora and fauna. However, huge chunks of the ecoregion still lack essential wilderness protections. Oil and gas development threatens the coastal plain and millions of acres at the heart of biological refuges. Recently, a large portion of coral bleaching (hot spot) in coral reefs was discovered throughout the world (5). Coral bleaching is the loss of color from the living coral animal and can cause their death. The color of coral comes from microalgae living in the coral's tissue; when the water surrounding the coral becomes too warm, corals become stressed and expel the microalgae. These microalgae also provide the coral with food, and as the algae fade away, corals starve to death.

Therefore, environmental management and conservation should be adopted first to quickly protect the ecoregion, including its organisms, people, and its total effective surroundings from misuse or damage by human activities; second, to ensure that life support systems and renewable resources are used in an ecologically sustainable manner while supplying the conditions for a satisfactory human existence. Community-based projects will demonstrate the benefits for local people and the environment combining sustainable development and conservation and building up coordination among various activities in the ecoregion. In addition, ecologically sustainable management should be adopted in every ecoregion that will be used to assess the ecological sustainability of the biological resources in the ecoregion.

Environmental justice often plays itself out as a disparity between the rich and the poor, between the "haves" and the "have-nots." Just as environmental justice speaks to issues of poor people and people of color who are disproportionately subject to environmental pollution and public health risks, it also should speak of issues where poor people and people of color do not equally derive opportunities to obtain benefits and values associated with environmental resources.

CONCLUSION

Sustaining ecological processes, the landscape function, and the biological diversity of any ecoregion are important globally, nationally, and locally. Pressure to accommodate a rapidly increasing human population together with increased provision of goods and services has been growing for decades. This has impacted each ecoregion in its sociocultural, economic, and environmental characteristics. Therefore, a multiple scale, transdisciplinary integrated approach, and strategic framework for sustainable environmental planning and management of an ecoregion are needed. In addition, pressure to take action to halt land degradation, breakdown in production systems, decreasing water quality, loss of biodiversity, and global climate change is also needed for any government to improve the way it conserves, protects, and manages ecological processes, habitats, and species in the ecoregion. Furthermore, it must integrate planning and management across multiple scales, nesting the functional requirements of ecological systems and social systems to uphold and restore resilience in the ecological and social systems of the ecoregion for long-term sustainability.

BIBLIOGRAPHY

1. Wiken, E.B. (1986). Terrestrial ecozones of Canada. Ecological Land Classification Series No. 19. Environment Canada, p. 26.
2. Jarvis, I.E., MacDonald, K.B., and Betz, T. (1995). Development and application of a Canadian ecological framework. In: *Proc. GIS 95, 9th Ann. Symp. Geogr. Inf. Syst.*, Vancouver, British Columbia. GIS World Inc., Ft. Collins, CO, pp. 605–612.
3. Gauthier, D. (1992). Framework for developing a nationwide system of ecological areas: Part 1-A Strategy. Occasional Paper No. 12, Canadian Council on Ecological Areas. Secretariat, Canadian Wildlife Service, Environment Canada, p. 47.
4. Brunckhorst, D.J. (2000). *Bioregional Planning: Resource Management- Beyond the New Millennium*. Harwood Academic, Australia, p. 5.
5. Greg, R. (2002) Coral Threat Renews Push on Protocol, Wednesday 9 January 2002, http://www.theage.com.au/news/national/2002/01/09/FFX715DK6WC.html (access on 12 September 2002).

FLOOD OF PORTALS ON WATER

Dick de Jong
IRC International Water and
Sanitation Centre
Delft, The Netherlands

Marta Bryce
CEPIS/PAHO
Delft, The Netherlands

INTRODUCTION

The Internet has been inundated with new gateways and portals on water over the last couple of years, as witnessed by information specialists at the IRC International Water and Sanitation Centre. The New on the Net section in Source Water and Sanitation News (http://www.irc.nl/source/section.php/12) carries 54 news items since July 2002, of which seven refer to "gateway" or portal. Unfortunately, many of them only serve to confuse instead of help people who search the web for water-related information. Part of the problem resides in defining what constitutes a "portal" and what doesn't. Another issue is that most portals do not seem to have a clear notion of their intended visitors—and their motives for visiting. As a result, the portals lack much-needed focus.

Here, we give some recent examples, try to create more clarity by providing definitions and types, discuss key lessons learned, and raise key issues for further discussion and research. This is based on various Internet searches and an annotated selection of various gateways/portals by IRC staff.

Here are five examples of sites, none of which have all the functionality of a portal yet.

1. The closest example of a portal is the BBC Water Portal—The Water Debate (http://news.bbc.co.uk/2/hi/in_depth/world/2003/world_forum/water/default.stm) that was launched early June 2003. This is not only an impressive interactive site for the general public with video, audio, expert views, fact files, and water stories; it also provides fun options such as online opinion polls and a water quiz. However, it does not contain an option for online collaboration.
2. The IRC International Water and Sanitation Centre (established in 1968) family of websites (http://www.irc.nl) provides news and information (including advocacy/communication, publications, and bibliographic database) on low cost water supply and sanitation in developing countries. It has a number of topic sites. It also provides a starting point for the Source Water and Sanitation News service and the InterWATER guide to more than 650 organizations, in collaboration with the Water Supply and Sanitation Collaborative Council. The search engine (Atomz) provides access to all these related sites. The IRC DOC online database contains 15,000 bibliographical entries on developing country water, sanitation and hygiene documents, books, periodicals, videos, and so on. The (www.worldwaterday.org) site maintained by IRC contains information on one topic per year that is selected by the UN. Visitors can add comments and events.
3. Sanitation Connection (http://www.sanicon.net/) is an Internet-based resource that gives access to accurate, reliable, and up-to-date information on technologies, institutions, and financing of sanitation systems around the world. Institutions of international standing contribute to the information base by providing and maintaining a topic of their specialization.
4. Water Portal of the Americas (http://www.Water-Portal-Americas.org) is a combination of a gateway (entryway or portal) and a community of practice. It is basically a search engine, a structure and system for other organizations to copy or fit in. It is still a prototype that may grow into a collaborative portal. "The goal is to provide both an entryway (portal) to water information and to create a water information network, community, and resource that will provide qualified, trusted, and verifiable information and contacts."

   At the March 2003 Third World Water Forum the prototype of a "Water Portal for the America's" was shown by UNESCO/World Water Assessment Programme and the Waterweb Consortium, a year and a half after it was first announced at the Fifth Water Information Summit. The prototype gave an impression of what will be on the site, with only limited content. This is the first regional water portal that is scheduled to become part of the UNESCO WWAP/Waterweb Global Water Portal, as was announced in Japan. Funding for this started in April 2002.
5. The U.S.-based WaterTechOnline (http://www.watertechonline.com/news.asp?mode=4&N_ID=31857) is one of the commercial portals addressing business needs. It contains breaking daily news stories about the water and wastewater treatment industry, as well as industry bulletin boards, searchable supplier directories and article archives, industry opinion polls, new product announcements, and other industry-specific information.

These examples illustrate the proliferation of water portals, gateways, and websites. It also raises questions about the definitions, users, and relevance of these internet-based initiatives.

DEFINITIONS AND TYPOLOGIES

General information on portals can be found on the Portals Community site. This advertises on Google.com as "The Definite Enterprise Portal Resource" and offers Portals information, White papers, networking, and research (http://www.portalscommunity.com/registration/member.cfm?code=TDXTL05).

Defining Water Portals

From About.com and our own sources, a portal is a kind of website and provides a single point of access to aggregated information—a virtual front door. The term originated with large, well-known Internet search engine sites that expanded their features to include e-mail, news, stock quotes, and an array of other functionality. Some corporations took a similar approach in implementing their intranet sites, which then became known as enterprise information or corporate portals.

There are general portals and specialized or niche portals. Some major general portals include Yahoo, CNET, AOL, and MSN. Examples of niche portals that are accessible to the public include Garden.com (for gardeners), Fool.com (for investors), and DPReview.com (for photographers). These are sometimes also called vortals—vertical portals. Private niche portals are those that are used by employees of a company.

Typical services offered by public portal sites include a start page with rich navigation, a directory of websites, a facility to search for other sites, news, a collection of loosely integrated features (some of which may be provided by partners or other third parties) like weather information, stock quotes, phone, and map information, e-mail, and sometimes a community forum. Private portals often include access to payroll information, internal phone directories, company news, and employee documentation.

Here, we define a portal as "a one-stop, client-oriented website that offers visitors a broad array of interactive resources such as news, databases, discussion forums, search options, space to collaborate online, and links on water-related topics."

Searches

IRC staff did various online literature searches:

- On “Water Portals” on all the web search engines in April, May, and June 2003. Of the 45 results shown, 32 were relevant. On Google.com, 22 out of 42 were relevant. The other results were mainly related to game sites.
- On the specialized ICT site Whatis.com there is a lot of information on corporate portals through Best web links, news headlines expert technical advice, and web results.

CONCLUSIONS

Our main conclusions are:

1. There is a wide variety of “portals” on water, definition wise, topic wise, and language wise.
2. Even in the world of corporate portals the concept is subject to confusion in the marketplace (Peter A. Buxbaum, Nov 2001, searchEBusiness). Buxbaum quotes:

   > There are several different types of corporate portals, each sporting different capabilities and each requiring different technologies and technology providers. Experts segment the enterprise portal arena into at least three components. *Information portals* emphasize data and document retrieval and management. *Knowledge portals* enable collaboration and information sharing as well as expertise and knowledge capture. *Application portals* provide end-users with interfaces to enterprise systems. What do all these have in common? As IDC defines it, they all manage end-user access to multiple applications and information sources on the corporate intranet.

   Looking at the software giant SAP’s recent steps in its portal strategy, Buxbaum leaves us with a nagging question: If a software giant like SAP can’t explain portals clearly, how are its potential customers supposed to understand them?
3. Most of the sites that claim to provide a water “portal” can be more correctly described as providing a start page or gateway to web information. There is in fact a flood of gateways, but too many providers are drowning in ambition.
4. An interesting paper that describes pros and cons of a portal from a university point of view in easily understandable language is Kevin Lowey’s University Portals FAQ, *University Web Developers’ Mailing List 2002* (http://www.usask.ca/web_project/uwebd/portals_faq.html).
5. Some interesting findings emerge from a user survey of electronic information services (EIS) among 25 U.K. institutions of higher education. The effect of the Internet on information seeking by staff and students is hugely significant; search engines and known sites are the first resort for most academic queries, as well as for many personal domestic queries. There was a wide range of engines used and indications of haphazard searching.
6. Another major conclusion from the same survey is that subject gateways are notable only for their lack of mention among students and academic staff, although there is some use among library and information staff.

LESSONS LEARNED

The majority of the gateways/portals described here focus on the broadest possible water topic: water resources and management.

The claims for “The Water Portal” cannot be justified. At best we can help clarify what the sites we have selected are focusing on, what they are, what they deliver to visitors (who?), and how interactive they are.

Two good examples of the different approaches are The Netherlands-based start page on water (www.waterpagina.nl) that provides four pages of clickable names divided over 32 categories, compared to the recently established www.h2o-scanner.com. This has a start page from where searches can be done combining countries, segments, issues, and content.

With a portal the emphasis is shifting from a public website showing online pamphlets to a user-oriented website that provides tools, reports, and services specifically designed for that individual.

WHAT NEXT?

By offering its new website architecture to interested partners, IRC started since mid-2003 to develop a family of related websites on water sanitation and hygiene, using the same technical foundation: *A Water Portal Is Born*. The STREAMS coalition of resource centers is going to be the first partner using this opportunity to contribute content through its own segment of this shared platform. More partners will be added.

In 2004 the Portal functionality will be expanded to also contain login functionality to provide Intra- and Extranet opportunities, moderated lists to enable E-conferences, workflow management, and others.

INTERNET ACCESS GAP

A few data on Internet connectivity according to CIA’s World Factbook (http://cyberatlas.internet.com/big_picture/geographics/article/0,,5911_151151,00.html) show the enormous gap between Internet use in the developed and developing world.

The percentage of the population of Internet users in the world in 2002 was around 9%. In The Netherlands and Denmark it was around 61%, in the United States some 59%, in Canada 50%, and in Belgium 36%. In Colombia it was about 2.8%, in Bolivia 0.9%, in India 0.7%, in Nepal 0.23%, in Mozambique and Bangladesh 0.1%, and in Ethiopia 0.03%.

Kamel from Nepal Telcom indicates that there are 8 phone lines per 1000 persons in Nepal. E-mail was introduced in 1993 and Internet in 1994. They have 50 cyber cafes in Kathmandu and 20 in Pokhara. Government official Vandya indicates that IT development is important and Nepal has to continue to develop it. This needs to be done in parallel with other media such as wall newspapers, taking into account that basic literacy is 55% and English literacy 2%.

In Burkina Faso the literacy is only 20% and they have 60 spoken local languages in a population of 2.6 million. A local priest is making a very interesting effort by publishing a quarterly farmers magazine in some of the local languages. He indicates that Internet is useful to find information, but a considerable part of the information in the magazine comes from local farmers.

In terms of difficulties it is clear that the number of users is very low in the South and even more so in rural areas. Another difficulty is that Internet performance gives problems. Downloading in the South, for example, is slow and the cost of access is high. It is often a struggle to access a telephone line let alone high bandwidth facilities. In the preparation for a knowledge management workshop at the Sixth Water Information Summit in September 2003, it was already very clear that files over 1 MB gave a lot of problems. Only after simplifying and zipping them did they become more manageable.

LATIN AMERICA: VIRTUAL LIBRARY ON ENVIRONMENTAL HEALTH

A good example from the achievements and challenges of web use in the developing world is the Virtual Library on Environmental Health (VLEH) of the Pan American Information Network on Environmental Health, REPIDISCA, which has been operating since 1982 in 21 countries of Latin America and the Caribbean with 370 Cooperating Centres. The Pan American Center for Sanitary Engineering and Environmental Sciences, CEPIS/PAHO, of the Pan American Health Organization, PAHO/WHO, is in charge of the regional coordination.

The VLEH disseminates information on water supply, sanitation, and related topics (http://www.cepis.ops-oms.org/indexeng.html). It uses the same methodology as the Latin American and Caribbean Health Sciences Information Centre, BIREME, another information system sponsored by PAHO/WHO. Its architecture is based on six types of information sources and services: electronic publications in full text, secondary sources, teaching material, SDI service, news and discussion lists, and integrating components. The bibliographic database contains more than 134,000 records with abstracts, of which 7553 are full text.

The principles of the VLEH are:

- Equitable access to environmental health information.
- Alliances and consortia for maximizing resource sharing.
- Cooperative work and exchange of experiences.
- Decentralized development and operation at all levels.
- Development based on local conditions.
- Integrated mechanisms for evaluation and quality control.

The VLEH architecture (Fig. 1) is based on the characteristics of its information sources. An information source is any resource that responds to information needs, including information products and services, persons or networks of

Pan American Center for Sanitary Engineering and Environmental Sciences

Portals

- Air Quality
- Appropriate Technology
- Citizen Participation
- Disasters
- Drinking water and Sanitation Sector Information
- Epidemiology
- Hazardous Waste
- Health and Environment Impact Assessment
- Health in Housing
- Healthy Environments: Healthy Children
- Information: REPIDISCA
- Institutional Development
- Laboratory
- Native Populations
- Primary Environmental Care
- REPAMAR
- Solid Waste
- Toxicology
- Wastewaters
- Water Quality
- Water Resources - Management
- Water Treatment
- Workers' Health

Figure 1. The VLEH main page.

Table 1. Number of Records in Each Information Source of the VLEH

Number	Source	1981–1997	1998	1999	2000	2001	2002	5 June 2003
	Secondary							
1	Bibliographic	96,467	106,577	117,925	119,321	120,784	130,087	134,972
2	Institutions	1,953	1,986	2,043	1,914	1,779	2,106	2,283
3	Specialists					103	103	136
4	Events (current)		186	475		1,065	1,005	1,100
5	Academic programs					129	284	298
6	Journals			451	475	483	505	516
7	Videos			398	404	475	494	500
	Primary							
8	Full texts		404	927	1,568	3,159	6,071	7,553
9	Legislation			165	547	1,079	1,235	1,004
10	HDT	70	73	76	79	82	88	88
11	Repindex	63	66	69	72	75	78	78
12	Indicators					ALC	ALC	ALC
	Teaching Material							
13	Teaching material					607	829	879
14	Self-instruction courses						9	9
	SDI Service							
15	SDI data base		452	519	654	765	2,020	2,114
	News/Discussion Lists							
16	News			119	263	322	973	1356
17	Open discussion lists					7	11	11
18	Restricted discussion lists					12	11	13
	Integrating Components							
19	LISA (www links)			175	314	1,448	1,682	1,754
20	Thesaurus in five languages					3,000	3,143	3,259

persons, and computer programs. The VLEH architecture is organized as follows (Table 1):

1. *Secondary Sources*. Composed of the following databases: bibliographic (more than 134,000 records); institutions (includes the address and names of authorities); events (includes seminars, congresses, and short courses); academic programs (includes diplomas, master, and doctoral programs); journals (includes the titles, electronic availability, and access to full texts); videos (includes material available at CEPIS and Cooperating Centres).
2. *Primary Sources*. Composed of the following databases: full text documents (more than 7500 documents in Acrobat); legislation (more than 1000 laws and regulations in full text related to water quality); HDT (includes 78 issues of the series Technical Dissemination Sheets, published by REPIDISCA); Repindex (includes 88 issues of this series also published by REPIDISCA); indicators (includes data from 45 countries of the Latin American and Caribbean region based on the global assessment of water and sanitation services, 2000).
3. *Teaching Material*. The material training database contains electronic and multimedia sources with added value for teaching purposes; the database on self-instruction courses contains nine courses prepared by PAHO/CEPIS which are also available on CD-ROM.
4. *SDI Service*. Selective dissemination of information (SDI) to update users based on their specific interest profiles. It is also a mechanism for providing information through e-mail to users that lack or have communication constraints via the Internet.
5. *News and Discussion Lists*. Fosters communication among persons, including discussion lists, forums, and virtual communities.
6. *Integrating Components*. Ensure the integration of decentralized information sources, like the REPIDISCA Thesaurus (more than 3000 terms in five languages: Spanish, English, Portuguese, French, and German, with its corresponding hierarchical structure and synonyms) and the Health and Environment Information Locator (gathers other Internet sources on environmental health, based on the Global Information Locator Service, GILS, adopted by the Global Program of the Information Society and the Dublin Core).

Building the virtual library implies the development, adoption, and adaptation of tools to operate information sources according to its architecture.

PAHO's Headquarters, in Washington, DC, hosts the virtual library because its server has higher connection speed. Statistics show that the number of visitors increases every year: 549,469 users have visited our website up to June 18, 2003 (Fig. 2).

Figure 3 shows the number of users per information source from January to June 2003. The source of full-text documents had the highest number of users, indicating that efforts should be concentrated on this information resource.

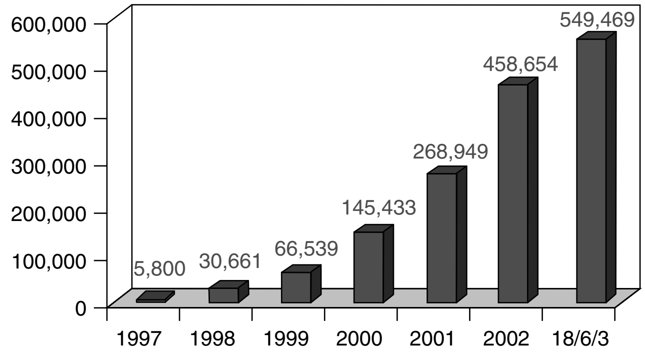

Figure 2. VLEH visitors, 1997–2003 (cumulative).

Every year, comparisons of statistics are done to analyze the trend of the VLEH use, as shown in Fig. 4 regarding the number of visitors per water-related subject.

The VLEH includes 17 additional topics related to CEPIS/PAHO working scope and that of the Sustainable Development and Environmental Health Area of PAHO/WHO. In 2002, the topic with the highest number of visitors was Workers' Health, in 2001 it was the Assessment in Drinking Water and Sanitation, and in 2002 it was the Solid Wastes Assessment.

Throughout its 20 years, REPIDISCA has carried out several user surveys. The last one in 2000 focused on web users and 424 persons replied to the questionnaire. Figure 5 shows the percentage of users per sector.

Regarding the type of user, 35% of them were professionals and technicians, as shown in Fig. 6. It is worthwhile to note that the use of the VLEH increases during weekends.

DIFFICULTIES

The difficulties that CEPIS/PAHO faces now with VLEH are:

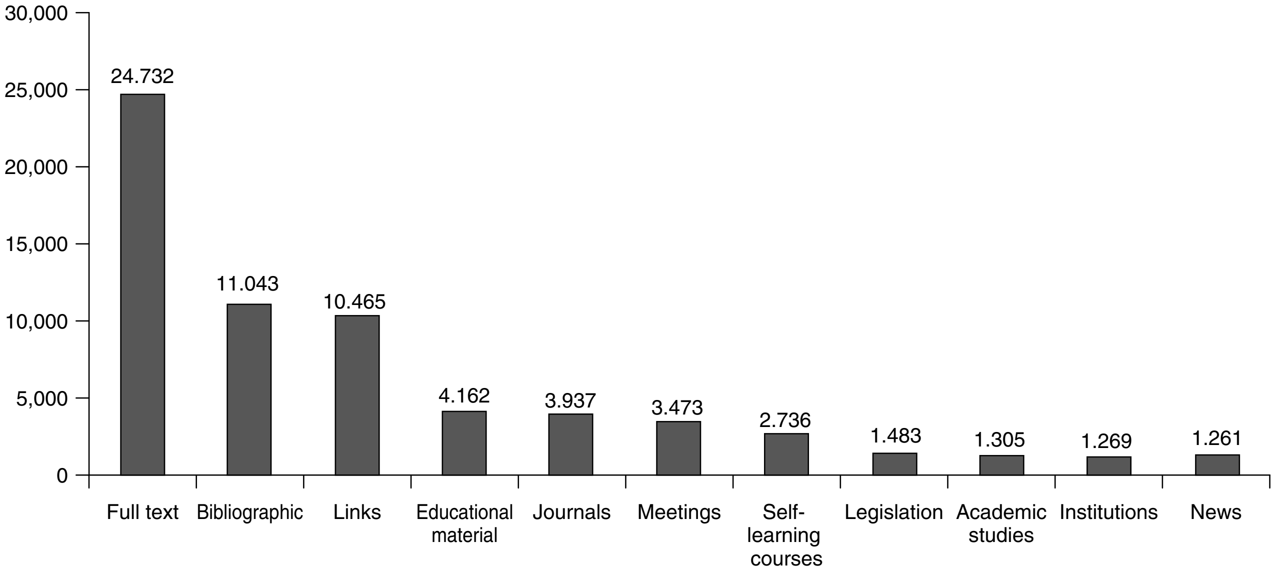

Figure 3. VLEH visitors per information source, January–June 6, 2003.

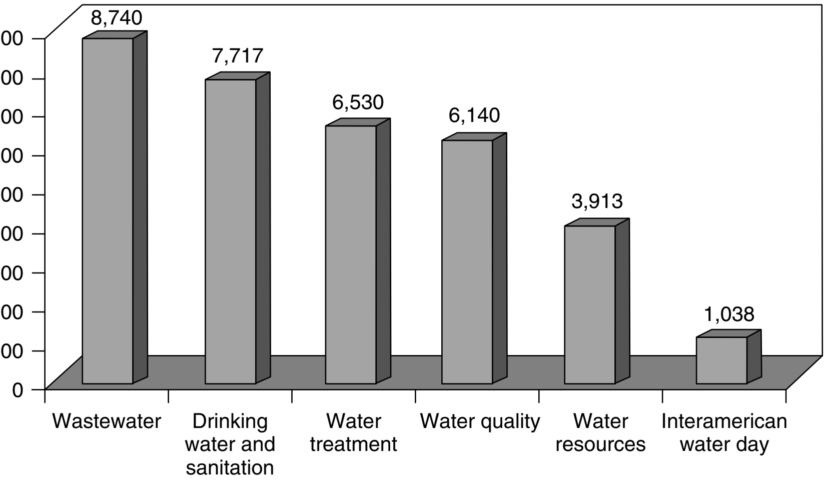

Figure 4. Visitors per water-related subject, January–June 2003.

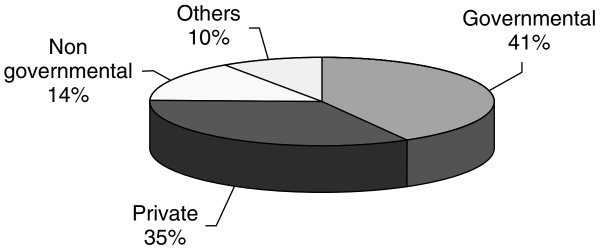

Figure 5. Users by sectors.

- Weakness of national institutions; in many countries water and sanitation utilities have undergone a privatization process.
- Staff reduction in public and private organizations; lack of investment in technical information and lack of resource sharing.
- Poor budgets for acquisition of information.
- Use of commercial software and methodologies in large universities that would not be willing to change their information system to follow the VLEH procedures; thus they need to duplicate the recording of the information they want to record into the VLEH.
- Inadequate updating of the VLEH at the regional and national levels.

Now CEPIS/PAHO has a great deal of information, but the VLEH cannot be just a space to make links with other websites. Users need a place where information is gathered, organized, and contained. Toward this end, CEPIS/PAHO and BIREME are implementing *Shared*, a Dutch technology that allows the production of "fingerprints" of every piece of information that is physically located in another web. These "fingerprints" are kept in a centralized file and weights may also be assigned to the topics of every information item. Institutions using the VLEH methodology will do this work more easily. In the future, it will be necessary to sign agreements to extract information automatically using *Shared*.

If *Shared* is implemented successfully, CEPIS would be able to enter into the Water Web Portal and other relevant sites updating their server of "fingerprints" and repeating the operation every month. The same procedure may be applied by each Cooperating Centre; thus national VLEH would be updated automatically. This will also avoid their regular export and sending of new information to CEPIS and the subsequent follow-up and control of duplicates that is done currently. Finally, regular training should be considered as a must to enable national institutions the use of the VLEH methodology. Workshops have been carried out in eleven countries and as a result they are already operating their own national VLEH.

FUZZY CRITERIA FOR WATER RESOURCES SYSTEMS PERFORMANCE EVALUATION

Slobodan P. Simonovic
The University of Western Ontario
London, Ontario, Canada

INTRODUCTION

One of the main goals of engineering design is to ensure that a system performs satisfactorily under a wide range of possible future conditions. This premise is particularly true of large and complex water resources systems. Water resources systems usually include conveyance facilities such as pipes and pumps, treatment facilities such as sedimentation tanks and filters, and storage facilities such as reservoirs and tanks. These elements are interconnected in complicated networks serving broad geographical regions. Each element is vulnerable to temporary disruption in service due to natural hazards or human error whether unintentional, as in the case of operational errors and mistakes, or due to intentional causes such as a terrorist act. Most of the hazards cannot be controlled or predicted with an acceptable degree of accuracy. Uncontrollable external factors also affect the capacity and the performance of water resource systems. The determination of the load pattern presents unique challenges. Ang and Tang (1) point out that there is uncertainty in all engineering-based systems because these systems rely on the modeling of physical phenomena that are either inherently random or difficult to model with a high degree of accuracy.

The sources of uncertainty are many and diverse and, as a result, impose a great challenge to water resources systems design, planning, and management. The goal to ensure failsafe system performance may be unattainable. Adopting high safety factors is one means to avoid the uncertainty of potential failures. However, making

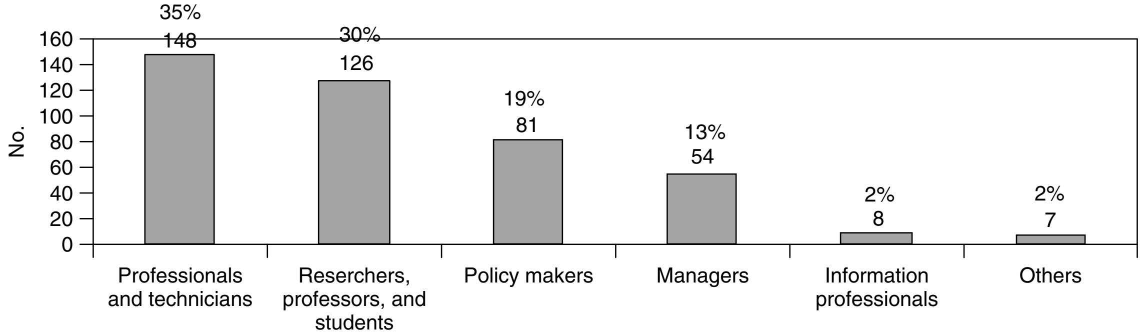

Figure 6. Types of user.

safety the first priority may render the system solution infeasible. Therefore, known uncertainty sources must be quantified. Engineering risk and reliability analysis is a general methodology for the quantification of uncertainty and the evaluation of its consequences for the safety of engineering systems (2). The first step in any risk analysis is to identify the risk, clearly detailing all sources of uncertainty that may contribute to the risk of failure. The quantification of risk is the second step, where the effects of the uncertainties are measured using different system performance indices and figures of merit. The early works of Hashimoto et al. (3,4) are the basis for the use of performance indices to evaluate the risk and reliability of water resources systems. They suggest *reliability, resiliency, vulnerability*, and *robustness* indices as criteria for evaluating the performance of water resources systems. Prior knowledge of the probability density functions of both resistance and load and/or their joint probability distribution function is a prerequisite to the probabilistic approach. In practice, data on previous failure experience is usually insufficient to provide such information. The subjective judgment of the decision-maker to estimate the probability distribution of a random event [subjective probability approach of Vick (5)] is another approach to deal with data insufficiency. The third approach is Bayes's theory, where engineering judgment is integrated with observed information. Until now the probabilistic approach was the only approach for water resources systems reliability analyses. But it fails to address the problem of uncertainty that goes along with human input, subjectivity, and lack of history and records. There is a practical and urgent need to investigate new approaches that can compensate for the ambiguity or uncertainty of human perception.

This contribution investigates the different approaches used to handle the problem of system reliability. Three new fuzzy reliability measures are presented (6): (1) the combined reliability–vulnerability index, (2) the robustness index, and (3) the resiliency index. They are developed to evaluate the operational performance of water resources systems. These measures could be useful decision-making aids in a fuzzy environment where subjectivity, human input, and lack of previous records impede the decision-making process.

FUZZY SETS

Zadeh (7), the founder of the theory of fuzzy sets, defines it as a formal attempt to capture, represent, and work with objects with unclear or ambiguous boundaries. This concept, although relatively new, has its origins in the early application of a multivalued logic notion to overcome the difficulties faced by the dual-logic representation in set theory. Water resources systems are potentially vulnerable to a wide variety of hazards that could limit their satisfactory performance. As a result, risks of future systems failure are often unavoidable (1). The engineering risk index characterizes the safety of water resources systems. The uncertainties associated with various sources of risk undermine the efficiency of this index. Because of these inevitable inefficiencies, all properties of water resources systems are subject to unavoidable and uncertain risk conditions. Different performance measures have been used in water resources systems reliability analyses to describe the system performance under extreme loading conditions (2). The majority of engineering reliability analyses rely on the use of a probabilistic approach. Both resistance and load are considered random variables. However, the characteristics of resistance and/or load cannot always be measured precisely or treated as random variables. Therefore, the fuzzy representation of either must be examined. The case of both fuzzy resistance and fuzzy load is rarely addressed in studies on the subject (6,8).

FUZZY CRITERIA FOR SYSTEM PERFORMANCE EVALUATION

Key Definitions

Failure. The calculation of performance indices depends on the exact definition of unsatisfactory system performance. It is difficult to arrive at a precise definition of failure because of the uncertainty in determining system resistance, load, and the accepted unsatisfactory performance threshold. Figure 1 depicts a typical system performance (resistance time series), with a constant load during the operation horizon. According to the classical definition, the failure state is the state when resistance falls below the load, margin of safety $M < 0.0$ or safety factor $\Theta < 1.0$, which is represented by the ratio between the system's resistance and load, shown in Fig. 1 by the dashed horizontal line.

Sometimes water resources systems fail to perform their intended function. For example, the available resistance from different sources in the case of water resources systems is highly variable. The actual load may also fluctuate significantly. Consequently, in the design of water resources systems, certain periods of water shortage may be a given. The precise identification of failure is neither realistic nor practical. It is more realistic to build in the inevitability of partial failure. A degree of acceptable system failure was introduced using the solid horizontal line, as shown in Fig. 1.

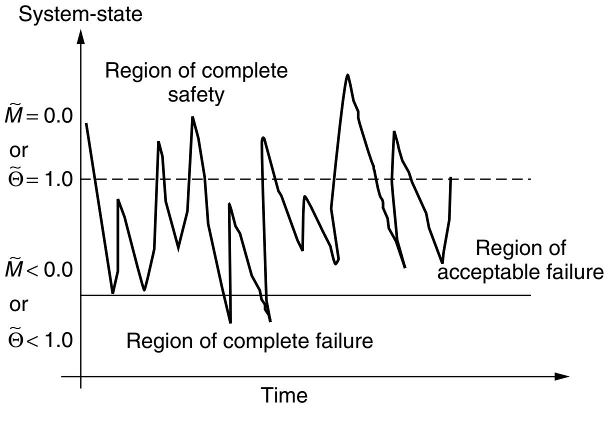

Figure 1. Variable system performance.

The boundary of the acceptable failure region is ambiguous and varies from one decision-maker to the other depending on personal perception of risk. Boundaries cannot be determined precisely. Fuzzy sets, on the other hand, are capable of representing the notion of imprecision better than ordinary sets. The acceptable level of performance can be represented as a fuzzy membership function in the following form:

$$\tilde{M}(m) = \begin{cases} 0, & \text{if } m \leq m_1 \\ \phi(m), & \text{if } m \in [m_1, m_2] \\ 1, & \text{if } m \geq m_2 \end{cases}$$

or

$$\tilde{\Theta}(\theta) = \begin{cases} 0, & \text{if } \theta \leq \theta_1 \\ \phi(\theta), & \text{if } \theta \in [\theta_1, \theta_2] \\ 1, & \text{if } \theta \geq \theta_2 \end{cases} \qquad (1)$$

where $\tilde{M}$ is the fuzzy membership function of the margin of safety; $\phi(m)$ and $\phi(\theta)$ are functional relationships representing the subjective view of the acceptable risk; m_1 and m_2 are the lower and upper bounds of the acceptable failure region, respectively; $\tilde{\Theta}$ is the fuzzy membership function of the factor of safety; and θ_1 and θ_2 are the lower and upper bounds of the acceptable failure region, respectively.

Figure 2 is a graphical representation of the definition presented in Equation 1. The lower and upper bounds of the acceptable failure region are introduced in Equation 1 by m_1 (or θ_1) and m_2 (or θ_2). The value of the margin of safety (or factor of safety) below m_1 (or θ_1) is definitely unacceptable. Therefore, the membership function value is zero. The value of the margin of safety (or factor of safety) above m_2 (or θ_2) is definitely acceptable and therefore belongs in the acceptable failure region. Consequently, the membership value is one. The membership of the in-between values varies with the subjective assessment of a decision-maker. Different functional forms may be used for $\phi(m)$ [or $\phi(\theta)$] to reflect the subjectivity of different decision-makers' assessments.

High system reliability is reflected through the use of high values of the margin of safety (or factor of safety), that is, high values for both m_1 and m_2 (or θ_1 and θ_2). The difference between m_1 and m_2 (or θ_1 and θ_2) inversely affects the system reliability: that is, the higher the difference, the lower the reliability. Therefore, the reliability reflected by the definition of an acceptable level of performance can be quantified in the following way:

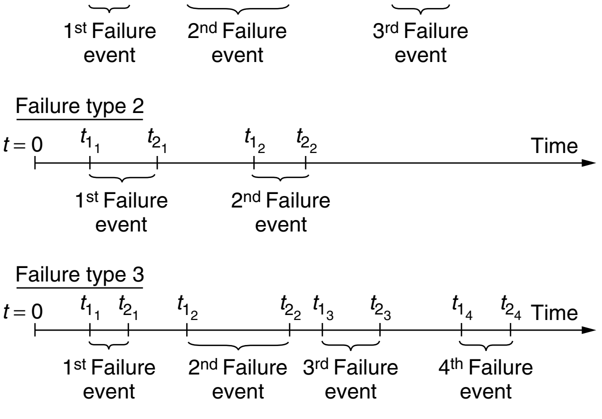

Figure 2. Fuzzy representation of acceptable failure region.

$$LR = \frac{m_1 \times m_2}{m_2 - m_1} \quad \text{or} \quad LR = \frac{\theta_1 \times \theta_2}{\theta_2 - \theta_1} \qquad (2)$$

where LR is the reliability measure of the acceptable level of performance.

The subjectivity of decision-makers will always result in a degree of ambiguity of risk perception. This alternate definition of failure allows for a choice among the lower bound, upper bound, and the function $\phi(m)$ [or $\phi(\theta)$]. This approach also provides an easy and comprehensive tool for risk communication.

Fuzzy System-State. System resistance and load can be represented in a fuzzy form to capture the uncertainty inherent in the system performance. The fuzzy form allows for the determination of the membership function of the resistance and load in a straightforward way even when there is limited available data. Fuzzy arithmetic can be used to calculate the resulting margin of safety (or factor of safety) membership function as a representation of the system-state at any time:

$$\tilde{M} = \tilde{X}(-)\tilde{Y} \quad \text{and} \quad \tilde{\Theta} = \tilde{X}(/)\tilde{Y} \qquad (3)$$

where $\tilde{M}$ is the fuzzy margin of safety; $\tilde{X}$ is the fuzzy resistance capacity; $\tilde{Y}$ is the fuzzy load requirement; (−) is the fuzzy subtraction operator; (/) is the fuzzy division operator; and $\tilde{\Theta}$ is the fuzzy factor of safety.

Compatibility. The purpose of comparing two fuzzy membership functions is to illustrate the extent to which the two fuzzy sets match. Several classes of methods are available, none of which can be described as the best method. The reliability assessment, presented here, involves a comparative analysis of the system-state membership function and the predefined acceptable level of the performance membership function. Therefore, the compliance of two fuzzy membership functions can be quantified using the fuzzy compatibility measure.

Possibility and necessity lead to the quantification of the compatibility of two fuzzy sets. The possibility measure quantifies the overlap between two fuzzy sets, while the necessity measure describes the degree of inclusion of one fuzzy set into another fuzzy set (9). However, in some cases, high possibility and necessity values do not reflect clearly the compliance between the system-state membership function and the acceptable level of performance membership function. For example, let's consider the two system-state functions A and B with the same possibility and necessity values. However, assume that system-state A has a larger overlap with the performance membership function than the system-state B. The overlap area between the two membership functions, as a fraction of the total area of the system-state, illustrates compliance more clearly than the possibility and necessity measures; that is,

$$C_{S,L} = \frac{OA_{S,L}}{A_S} \tag{4}$$

where $C_{S,L}$ is the compliance between the system-state membership function (S) and the acceptable level of performance membership function (L); $OA_{S,L}$ is the overlap area between the system-state membership function (S) and the acceptable level of performance membership function (L); and A_S is the area of the system-state membership function (S). An overlap in a high significance area (area with high membership values) is preferable to an overlap in a low significance area. The compatibility measure can be calculated using

$$CM_{S,L} = \frac{WOA_{S,L}}{WA_S} \tag{5}$$

where $CM_{S,L}$ is the compatibility measure between the system-state membership function (S) and the acceptable level of performance membership function (L); $WOA_{S,L}$ is the weighted overlap area between the system-state membership function (S) and the acceptable level of performance membership function (L); and WA_S is the weighted area of the system-state membership function (S).

Combined Fuzzy Reliability–Vulnerability Criteria

Reliability and vulnerability were used to provide a complete description of system performance in case of failure and to determine the magnitude of the failure event. Once an acceptable level of performance is determined in a fuzzy form, the anticipated performance in the event of failure as well as the expected severity of failure can be determined.

When certain values are specified for the lower and upper bounds [m_1 and m_2 (or θ_1 and θ_2) in Equation 1], thus establishing a predefined acceptable level of performance, the anticipated system failure is limited to a specified range. Systems that are highly compatible with the predefined acceptable level of performance will yield a similar performance; that is, the expected system failure will be within the specified range ([m_1, m_2] or [θ_1, θ_2]). In order to calculate system reliability, several acceptable levels of performance must be defined to reflect the different perceptions of the decision-makers. A comparison between the fuzzy system-state membership function and the predefined fuzzy acceptable level of performance membership function provides information about both system reliability and system vulnerability at the same time (Fig. 3).

The system reliability is based on the proximity of the system-state to the predefined acceptable level of performance. The measure of proximity is expressed by the compatibility measure suggested in Equation 5. The new combined fuzzy reliability–vulnerability index is formulated as follows:

$$RE_f = \frac{\max_{i\in K}\{CM_1, CM_2, \ldots, CM_i\} \times LR_{\max}}{\max_{i\in K}\{LR_1, LR_2, \ldots, LR_i\}} \tag{6}$$

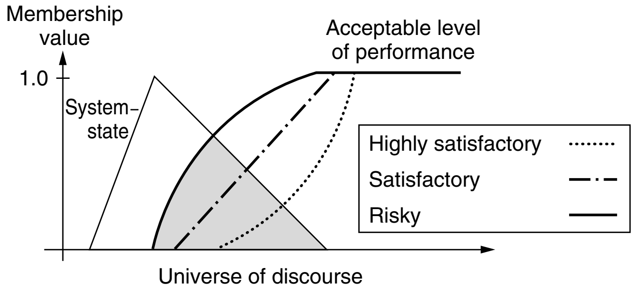

Figure 3. Compatibility with different level of performance membership functions.

where RE_f is the new combined fuzzy reliability–vulnerability index; $LR_{\max}$ is the reliability measure of acceptable level of performance corresponding to the system-state with maximum compatibility value; LR_i is the reliability measure of the ith acceptable level of performance; CM_i is the compatibility measure for the system-state with the ith acceptable level of performance; and K is the total number of the defined acceptable levels of performance. The reliability–vulnerability index is normalized to attain a maximum value of 1.0, by the introduction of the $\max_{i\in K}\{LR_1, LR_2, \ldots, LR_i\}$ value as the maximum achievable reliability.

Fuzzy Robustness

Robustness measures the system's ability to adapt to a wide range of possible future load conditions, at little additional cost (4). The fuzzy form of change in future conditions can be obtained through a redefinition of the acceptable level of performance and a change in the system-state membership function. As a result, the system's robustness is defined as the change in the compatibility measure:

$$RO_f = \frac{1}{CM_1 - CM_2} \tag{7}$$

where RO_f is the new fuzzy robustness index; CM_1 is the compatibility measure before the change in conditions; and CM_2 is the compatibility after the change in conditions.

Equation 7 reveals that the higher the change in compatibility, the lower the value of the fuzzy robustness. Therefore, high robustness values allow the system to better adapt to new conditions.

Fuzzy Resiliency

The time required to recover from the failure state can be represented as a fuzzy set. Because the reasons for a failure may be different, system recovery times will vary depending on the type of failure, as shown in Fig. 4. A series of fuzzy membership functions can be developed to allow for various types of failure. The maximum recovery time is used to represent the system recovery time (10):

$$\begin{aligned}\tilde{T}(\alpha) = (&\max_{j\in J}[t_{1_1}(\alpha), t_{1_2}(\alpha), \ldots, t_{1_J}(\alpha)] \\ &\times \max_{j\in J}[t_{2_1}(\alpha), t_{2_2}(\alpha), \ldots, t_{2_J}(\alpha)])\end{aligned} \tag{8}$$

where $\tilde{T}(\alpha)$ is the system fuzzy maximum recovery time at α level; $t_{1_J}(\alpha)$ is the lower bound of the jth recovery time at α level; $t_{2_J}(\alpha)$ is the upper bound of the jth recovery time at α level; and J is total number of failure events.

The center of gravity of the maximum fuzzy recovery time can be used as a real number representation of the system recovery time. Therefore, system resilience is determined to be the inverse value of the center of gravity:

$$RS_f = \left[\frac{\int_{t_1}^{t_2} t\,\tilde{T}(t)\,dt}{\int_{t_1}^{t_2} \tilde{T}(t)\,dt} \right]^{-1} \quad (9)$$

where RS_f is the new fuzzy resiliency index; $\tilde{T}(t)$ is the system fuzzy maximum recovery time; t_1 is the lower bound of the support of the system recovery time; and t_2 is the upper bound of the support of the system recovery time.

The inverse operation can be used to illustrate the relationship between the value of the recovery time and the resilience. The longer the recovery time, the lower the system's ability to recover from the failure, and the lower the resilience.

CONCLUSIONS

Water resources systems are vulnerable to a wide variety of hazards that could potentially limit their ability to perform satisfactorily. The diversity of uncertainty sources presents a great challenge to water resources systems design, planning, and management. The probabilistic approach usually fails to address the problems of human error, subjectivity, and the lack of system performance history and records. The fuzzy set approach addresses those issues. A fuzzy system reliability analysis is ideally based on the comparison between the fuzzy sets representing both the system's state of safety and the potential for system failure.

A fuzzy reliability–vulnerability measure clearly quantifies the reliability and the vulnerability of multicomponent systems. The quantification is based on the use of appropriate fuzzy compatibility measures to illustrate the relationship between the fuzzy system's state of safety and the fuzzy failure events. The contribution also shows a fuzzy measure, capable of determining the system's ability to adapt to changing conditions. This measure of fuzzy robustness provides a vital tool to assess the system's behavior through the introduction of a wide variety of uncertain conditions. A fuzzy resiliency measure was also presented to capture the system's response to uncertain future failure events. This measure is able to incorporate all types of system responses to potential failure events throughout the life of the design.

The presented fuzzy performance criteria were evaluated using two simple hypothetical cases by El-Baroudi and Simonovic (6).

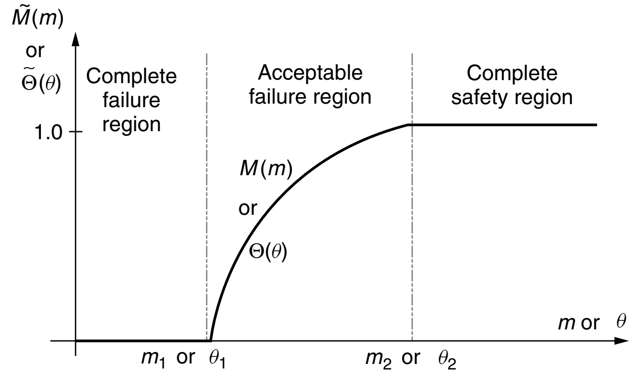

Figure 4. Recovery times for different types of failure.

BIBLIOGRAPHY

1. Ang, H.-S. and Tang, H. (1984). *Probability Concepts in Engineering Planning and Design*. John Wiley & Sons, Hoboken, NJ.
2. Ganoulis, J.G. (1994). *Engineering Risk Analysis of Water Pollution: Probabilities*. VCH, Weinheim, The Netherlands.
3. Hashimoto, T., Stedinger, J.R., and Loucks, D.P. (1982). Reliability, resiliency, and vulnerability criteria for water resources system performance evaluation. *Water Resour. Res.* **18**(1): 14–20.
4. Hashimoto, T., Loucks, D.P., and Stedinger, J.R. (1982). Robustness of water resources systems. *Water Resour. Res.* **18**(1): 21–26.
5. Vick, S. (2002). *Degrees of Belief: Subjective Probability and Engineering Judgment*. ASCE Press, Reston, VA.
6. El-Baroudy, I. and Simonovic, S.P. (2004). Fuzzy criteria for the evaluation of water resources systems performance. *Water Resour. Res.* **40**(10): W10503.
7. Zadeh, L.A. (1965). Fuzzy sets. *Inf. Control* **8**: 338–353.
8. Sherstha, B. and Duckstein, L. (1998). A fuzzy reliability measure for engineering applications. In: *Uncertainty Modelling and Analysis in Civil Engineering*. CRC Press, Boca Raton, FL, pp. 120–135.
9. Pedrycz, W. and Gomide, F. (1998). *An Introduction to Fuzzy Sets*. MIT Press, Cambridge.
10. Kaufmann, A. and Gupta, M. (1985). *Introduction to Fuzzy Arithmetic: Theory and Applications*. Van Nostrand Reinhold, New York.

PARTICIPATORY MULTICRITERIA FLOOD MANAGEMENT

Slobodan P. Simonovic
The University of Western Ontario
London, Ontario, Canada

INTRODUCTION

Flood management in general comprises different water resources activities aimed at reducing the potentially harmful impact of floods on people, the environment, and the economy of a region. Sustainable flood management decision-making requires integrated consideration of economic, ecological, and social consequences of disastrous flood. While economic consideration gets priority in traditional approaches to decision-making, empowerment of stakeholders is an issue that is demanding increasing attention today in many decision-making processes. Flood management activities (i.e., disaster mitigation, preparedness, and emergency management) may be designed and achieved without the direct participation of stakeholders; however, it cannot be implemented without them (1). In order to decide what flood control measures to adopt in a floodplain, the decision-making process should include different stakeholders.

Government policy-makers and professional planners are the first to consider. However, others such as the general public, communities affected by the decision outcomes, nongovernmental organizations, and different interest groups should be included as well. This contribution describes a multicriteria decision-making methodology for participatory flood management (2,3). The methodology should (1) evaluate potential alternatives based on multiple criteria under uncertainty; (2) accommodate the high diversity and uncertainty inherent in human preferences; and (3) handle a large amount of data collected from stakeholders.

METHODOLOGY

Flood management decision-making problems are complex due to their multicriteria nature. For a given goal, many alternative solutions may exist that provide different levels of satisfaction for different issues, such as environmental, social, institutional, and political. These concerns naturally lead to the use of multicriteria decision-making techniques, in which trade-offs among the single objectives can lead to the most desirable solution. Multicriteria decision-making becomes more complicated with the increase in the number of individuals/groups involved in the decision-making process. In reality, the decision-making process often involves multiple stakeholders/decision-makers. Multiple stakeholders' participation introduces a great deal of complexity into the analysis. The decision problem is no longer limited to the selection of the most preferred alternative among the possible solutions according to a single set of preferences. The analysis must also be extended to account for the conflicts among different stakeholders with different objectives. Most of the existing approaches in multicriteria decision-making with a single stakeholder/decision-maker consist of two phases (4): (1) the aggregation of the judgments with respect to all criteria and per decision alternatives; and (2) the ranking of the decision alternatives according to the aggregated judgment. In the case of multiple stakeholders, an additional aggregation is necessary with respect to the judgments of all the stakeholders.

Consider a multicriteria multiparticipant decision-making problem where m alternatives are to be evaluated by n decision-makers, who are using p objectives. The general conceptual decision matrix for the discrete multicriteria multiparticipant problem is shown in Table 1.

Table 1. Conceptual Decision Matrix for a Discrete Multicriteria Multiparticipant Decision Problem

	O$_1$	...	**O**$_p$
A$_1$	a_{11}	...	A_{1p}
...	...	...	...
A$_m$	a_{m1}	...	a_{mp}
DM$_1$	w_{11}	...	w_{1p}
...	...	...	...
DM$_n$	w_{n1}	...	w_{np}

In Table 1, **A** denotes the alternative, **O** is the criterion, and **DM** is the decision-maker/stakeholder. The preference of the decision-maker k $(k = 1, \ldots, n)$ for the objective j $(j = 1, \ldots, p)$ is expressed by w_{jk}, and a_{ij} is the performance evaluation of the alternative i $(i = 1, \ldots, m)$ for each objective j.

The classical outcome of the decision matrix is the ranking of the alternatives. To obtain that, a number of steps are necessary, such as establishing the preference structure, the weights, and also the performance evaluations. All these can be termed as the inputs for the decision matrix. These inputs come from the stakeholder/decision-maker. The decision matrix shows that the inputs can be for the preference of criteria as well as for the performance evaluations. The decision-maker might also have a preference structure for the alternatives. In the case of a multiparticipant decision-making problem, these inputs are to be collected from all the stakeholders.

Equation 1 is a general mathematical formulation of this multicriteria, multiparticipant problem. A payoff matrix can be obtained for the problem where m alternatives are to be evaluated by n stakeholders/decision-makers, who are using p criteria:

$$A^k = (a_{ij})^k = \begin{bmatrix} a_{11} & \ldots & \ldots & a_{1p} \\ a_{21} & \ldots & \ldots & a_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & \ldots & \ldots & a_{mp} \end{bmatrix}, \quad (k = 1, \ldots, n) \quad (1)$$

Here $A^k_{i.} = (a_{i1}, \ldots, a_{ip})^k$ means that alternatives i are being evaluated by criteria from 1 to p by decision-maker k. The symbol $A^k_{.j} = (a_{1j}, \ldots, a_{mj})^k$ means that the objective j is being used by decision-maker k to evaluate all alternatives from 1 to m.

The solution to this problem is to have each alternative evaluated by all the decision-makers using all criteria. The process can be summarized as the following mapping function:

$$\Psi : \{A^k | k = 1, \ldots, n\} \to \{G\} \quad (2)$$

where G is a collective weighted agreement matrix. It is crucial that this mapping function represents all criteria that the decision-makers use in judging all the alternatives.

Flood management decision-making is always associated with some degree of uncertainty. This uncertainty could be categorized into two basic types: uncertainty caused by inherent hydrologic variability and uncertainty due to a lack of knowledge (5). Uncertainty of the first type is associated with the spatial and temporal changes of hydrologic variables such as flow, precipitation, and water quality. The second type of uncertainty occurs when the particular value of interest cannot be assessed exactly because of the limitation in the available knowledge. The second type of decision uncertainty is more profound in the area of public decision-making such as in the case of flood management. Capturing the views of individuals presents the problem of uncertainty. The major challenge while collecting these views is to find out the technique that will capture those uncertainties and also will be usable in a multicriteria tool.

Participation of Multiple Stakeholders

An aggregation procedure is one of the ways to include information from the participating decision-makers into the decision matrix. The available methods do not seem to be appropriate for flood management for two reasons. The methodology (3) includes representation of inputs from a large number of participants and the analysis of inputs to make them usable for application to various multicriteria decision-making methods. Fuzzy set theory and fuzzy logic are used to represent the uncertainties in stakeholders' opinions. Three possible types of fuzzy input have been considered to capture the subjectivity of the responses from stakeholders. When a stakeholder is asked to evaluate an alternative against a particular criterion, the answer may take one of the following forms: (1) a numeric scale response; (2) a linguistic answer (e.g., poor, fair, good, very good); or (3) an argument (e.g., "if some other condition is satisfied then it is good"). For the first type, the input is quite straightforward. For a type (2) answer, it will be necessary to develop the membership functions for the linguistic terms. Type (3) input can be described by using a fuzzy inference system, which includes membership functions, fuzzy logic operators, and the if–then rule. For this, the membership functions for the input arguments need to be developed first. Then fuzzy operator and fuzzy logic are applied to obtain the output. It should be noted that the interpretation of type (2) and type (3) input values is highly dependent on the shape of the membership functions and the degree of severity chosen by the expert for a particular application.

After receiving the inputs from all stakeholders, the next step is to aggregate those inputs to find a representative value. It is obvious that for all input types considered above, the responses are sure to be influenced by a number of repetitions. This means many respondents can provide the same response. This implies that the general methodologies of fuzzy aggregation cannot be applied for deriving the resultant input from a large number of decision-makers. The fuzzy expected value (FEV) method can be used instead to get the resulting opinion of the stakeholders. Following is the definition of the fuzzy expected value: Let χ_A be a B-measurable function such that $\chi_A \in [0,1]$. The FEV of χ_A over the set A, with respect to the fuzzy measure μ, is defined as

$$\mathrm{FEV}(\chi_A) = \sup_{T\in[0,1]} \{\min[T, \mu(\xi_T)]\} \tag{3}$$

where

$$\xi_T = \{x|\chi_A(x) \geq T\} \tag{4}$$

and

$$\mu\{x|\chi_A(x) \geq T) = f_A(T) \quad \text{is a function of the threshold } T \tag{5}$$

Figure 1 provides a geometric interpretation of the FEV. Performing the minimum operator, the two curves create the boundaries for the remaining triangular curve. The supremum operator returns the highest value of $f_A(T)$, which graphically represents the highest point of the triangular curve. This corresponds to the intersection of the two curves where $T = H$.

The FEV can be computed for all three types of inputs mentioned earlier in this section. For type (1) input, the resultant FEV should be a numeric value between 0 and 1. For both type (2) and type (3) inputs, the resultant FEVs are membership functions. The crisp numeric equivalents of these membership functions can be obtained by applying a defuzzification method and can then be compared with type (1) answers. The centroid of area defuzzification method has been used to return a value obtained by averaging the moment area of a given fuzzy set. Mathematically, the centroid, $\overline{x}$, of a fuzzy set, A, is defined as

$$\overline{x} = \frac{\int_0^1 x \cdot \mu_A(x)dx}{\int_0^1 \mu_A(x)dx} \tag{6}$$

where $\mu_A(x)$ is the membership function of the fuzzy set A.

The resultant FEVs are now the aggregated evaluation of the alternatives from all the stakeholders. They can now be used as the input value in the decision matrix for the multicriteria analysis.

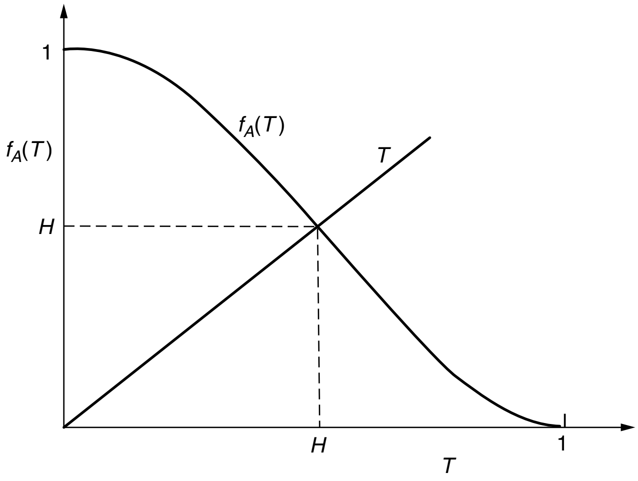

Figure 1. Participatory multicriteria flood management.

Participatory Multicriteria Decision-making Under Uncertainty

Here an innovative modification has been made to the compromise programming multicriteria decision-making technique to accommodate participatory flood decision-making under uncertainty. Bender and Simonovic (6) fuzzified compromise programming entirely and thus formulated fuzzy compromise programming (FCP). The driving force for the transformation from a classical to a fuzzy environment is that there is a need for accurate representation of subjective data in the flood decision-making. It is the theory of fuzzy sets that can represent the subjective data well. Thus, instead of using crisp numbers in the compromise programming distance metric equation, fuzzy numbers are used; instead of using classical arithmetic, fuzzy arithmetic is applied; instead of simply sorting distance metrics, fuzzy set ranking methods must be applied to sort the fuzzy distance metrics. In other words, the fuzzy transformation complicates the interpretation of the results but, on the other hand, models the decision-making process more realistically. Mathematically, the compromise programming distance metric in its discrete form can be presented as

$$L_j = \left[\sum_{z=1}^{t} \left\{ w_z^p \left(\frac{f_z^* - f_z}{f_z^* - f_z^-} \right)^p \right\} \right]^{1/p} \tag{7}$$

where $z = 1, 2, 3, \ldots, t$ and represents t criteria; $j = 1, 2, 3, \ldots, n$ and represents n alternatives; L_j is the distance metric of alternative j; w_z corresponds to a weight of a particular criterion; p is a parameter ($p = 1, 2, \infty$); f_z^* and f_z^- are the best and the worst value for criterion z, respectively (also referred to as positive and negative ideals); and f_z is the actual value of criterion z. The parameter p is used to represent the importance of the maximal deviation from the ideal point. Varying the parameter p from 1 to infinity allows one to move from minimizing the sum of individual regrets (i.e., having a perfect compensation among the criteria) to minimizing the maximum regret (i.e., having no compensation among the criteria) in the decision-making process. The choice of a particular value of this compensation parameter p depends on the type of problem and desired solution. The weight parameter, w_z, characterizes decision-makers' preference concerning the relative importance of criteria. Simply stated, the parameter places emphasis on the criteria that the decision-maker deems important. The parameter is needed because different participants in the decision-making process have different viewpoints concerning the importance of a criterion.

Bender and Simonovic (6) fuzzified compromise programming and thus formulated fuzzy compromise programming (FCP). The driving force for the transformation from a classical to a fuzzy environment is that there is a need for accurate representation of subjective data in the flood decision-making. It is the theory of fuzzy sets that can represent the subjective data well. Thus, instead of using crisp numbers in the distance metric, Equation 7, fuzzy numbers are used; instead of using classical arithmetic, fuzzy arithmetic is applied; instead of simply sorting distance metrics, fuzzy set ranking methods must be applied to sort the fuzzy distance metrics. In other words, the fuzzy transformation complicates the interpretation of the results but, on the other hand, models the flood decision-making process more realistically.

In fuzzy compromise programming, obtaining the smallest distance metric values is not easy, because the distance metrics are also fuzzy. To pick out a smallest fuzzy distance metric, from a group of distance metrics, fuzzy set ranking methods have to be used. A study by Prodanovic and Simonovic (7) compared fuzzy set ranking methods for use in fuzzy compromise programming and recommended using the method of Chang and Lee (8). This recommendation was founded on the fact that Chang and Lee's (8) method gave the most control in the ranking process—with degree of membership weighting and the weighting of the subjective type. The overall existence ranking index (OERI) suggested by Chang and Lee (8) has the following mathematical form:

$$\mathrm{OERI}(j) = \int_0^1 \omega(\alpha)[\chi_1 \mu_{jL}^{-1}(\alpha) + \chi_2 \mu_{jR}^{-1}(\alpha)] d\alpha \tag{8}$$

where the subscript j stands for alternative j, while α represents the degree of membership; χ_1 and χ_2 are the subjective type weighting indicating neutral, optimistic, or pessimistic preferences of the decision-maker, with the restriction that $\chi_1 + \chi_2 = 1$; parameter $\omega(\alpha)$ is used to specify weights, which are to be given to certain degrees of distance metric membership (if any); and μ_{jL}^{-1} (α) represents an inverse of the left part, and μ_{jR}^{-1} (α) the inverse of the right part of the distance metric membership function.

For χ_1 values greater than 0.5, the left side of the membership function is weighted more than the right side, which in turn makes the decision-maker more optimistic. Of course, if the right side is weighted more, the decision-maker is more of a pessimist (this is because he/she prefers larger distance metric values, which means the solution farther from the ideal solution). In summary, the risk preferences are: if $\chi_1 < 0.5$, the user is a pessimist (risk averse); if $\chi_1 = 0.5$, the user is neutral; and if $\chi_1 > 0.5$, the user is an optimist (risk taker). Simply stated, Chang and Lee's (8) overall existence ranking index is a sum of the weighted areas between the distance metric membership axis and the left and right inverses of a fuzzy number.

FLOOD MANAGEMENT IN THE RED RIVER BASIN, MANITOBA, CANADA

The proposed methodology is applied to flood management in the Red River Basin (9–11). One of the flood management problems at the planning stage in the Red River Basin is the complex, large-scale problem of ranking potential flood management alternatives. During the evaluation of alternatives, it is necessary to consider multiple criteria that may be quantitative and qualitative. The flood management process in the basin also involves numerous stakeholders. They include different levels of government, different agencies, private organizations, interest groups, and the general public. They all have different and specific

Table 2. Resultant FEVs

Alternative Type	Structural			Nonstructural			Combination		
	A	B	C	A	B	C	A	B	C
Question Number	FEV	FEV	FEV	FEV	FEV	FEV	FEV	FEV	FEV
Community Involvement									
1	0.600	0.650	0.544	0.647	0.650	0.544	0.600	0.625	0.544
2	0.529	0.517	0.500	0.500	0.517	0.491	0.500	0.570	0.544
3	0.618	0.700	0.529	0.559	0.625	0.529	0.600	0.625	0.544
4	0.600	0.650	0.544	0.657	0.650	0.559	0.686	0.650	0.544
5	0.700	0.700	0.559	0.629	0.650	0.544	0.700	0.650	0.544
6a	0.800	0.825	0.677	0.704	0.770	0.588	0.800	0.825	0.647
6b	0.771	0.770	0.588	0.714	0.717	0.574	0.743	0.770	0.574
6c	0.700	0.700	0.574	0.629	0.650	0.574	0.686	0.700	0.574
7	0.800	0.825	0.735	0.829	0.850	0.718	0.857	0.825	0.718
8	0.700	0.717	0.574	0.700	0.650	0.574	0.700	0.700	0.574
Personal Loss									
1	0.800	0.770	0.718	0.700	0.700	0.574	0.700	0.717	0.671
2	0.588	0.570	0.544	0.600	0.650	0.544	0.600	0.625	0.574
3a	0.500	0.570	0.574	0.559	0.625	0.574	0.559	0.570	0.574
3b	0.700	0.717	0.625	0.700	0.717	0.588	0.706	0.717	0.588
4	0.771	0.770	0.574	0.700	0.650	0.574	0.700	0.717	0.544
5	0.500	0.570	0.529	0.700	0.570	0.544	0.571	0.570	0.544

needs and responsibilities during all stages of flood management—planning, emergency management, and flood recovery. There has been increasing concern by the general public about the decisions to be taken on the selection of flood control measures. During the 1997 flood, it was indicated that certain stakeholders in the basin, particularly the floodplain residents, did not have adequate involvement in flood management decision-making. Dissatisfaction has been observed among the stakeholders about evacuation decisions during the emergency management and about compensation decisions during the postflood recovery (10).

The methodology presented in the previous section has been used to collect information from the stakeholders across the Canadian portion of the Red River Basin. In order to evaluate the utility of the methodology, a generic experiment was considered for the study to evaluate three alternative options for improved flood management. Three generic options considered are (1) structural alternatives, (2) nonstructural alternatives, and (3) a combination of both. The selection of criteria against which the alternatives are ranked is one of the most difficult but important tasks of any multicriteria decision analysis. The following two social objectives have been considered in our case study: level of community involvement and amount of personal losses (include financial, health, and psychological losses). A detailed survey has been conducted in the basin to collect the information on the two selected social criteria (12). All three types of inputs obtained from all the stakeholders were processed using the fuzzy expected value method. Table 2 summarizes the results of all three types of inputs (scale, linguistic, and conditional types, which are termed A, B, and C, respectively, in the table) as the evaluation of three alternatives (structural, nonstructural, combination) against two criteria (community development, personal loss). Obtained results show good correlation between the numeric scale type and linguistic type of inputs with an average difference of only 0.029.

The FEVs obtained in Table 2 are used further to rank the three generic alternatives. All questions are considered to carry the same weight. A set of ranking experiments has been conducted to evaluate the impact of different stakeholder groups on the final rank of alternatives: (a) experiment 1—all stakeholders interviewed; (b) experiment 2—stakeholders from the city of Winnipeg; (c) experiment 3—stakeholders from the Morris area (south of Winnipeg); and (d) experiment 4—stakeholders from the Selkirk area (north from Winnipeg).

The final results of four ranking experiments with three generic alternatives and two social criteria are shown in the Table 3 (defuzzified distance metric value and the rank in parentheses). It is obvious that the final rank varies with the experiment, therefore confirming that preferences of different stakeholders are being captured by the developed methodology.

CONCLUSIONS

The analyses of flood management options in the Red River Basin, Manitoba, Canada (3) show the applicability of the methodology for a real flood management decision-making problem. The stakeholders can now express their concerns regarding flood hazard in an informal way, and

Table 3. Final Rank of Flood Management Alternatives

Participants	Alternative 1	Alternative 2	Alternative 3
All stakeholders	13.22 (1)	13.72 (3)	13.29 (2)
Morris	15.43 (2)	16.09 (3)	13.63 (1)
Selkirk	14.63 (3)	14.42 (1)	14.58 (2)
Winnipeg	13.74 (1)	15.25 (3)	13.92 (2)

that can be incorporated into the multicriteria decision-making model. The application of methodology helps in solving the problem of incorporating a large number of stakeholders in the flood decision-making process.

BIBLIOGRAPHY

1. Affeltranger, B. (2001). *Public Participation in the Design of Local Strategies for Flood Mitigation and Control*, UNESCO IHP-V, Technical Documents in Hydrology, Vol. 48.
2. Simonovic, S.P. (2004). Sustainable floodplain management—participatory planning in the Red River Basin, Canada. In: *Proceedings of the Workshop on Modelling and Control for Participatory Planning and Managing Water Systems*, IFAC, CD-ROM.
3. Akter, T. and Simonovic, S.P. (2004). Aggregation of fuzzy views of a large number of stakeholders for multiobjective flood management decision making. *J. Environ. Manage.* **22**(1–2): 57–72.
4. Zimmerman, H.J. (2001). *Fuzzy Set Theory—and Its Application*, Academic Publishers, Boston.
5. Simonovic, S.P. (2000). Tools for water management: one view of the future. *Water Int. IWRA* **25**(1): 1–8.
6. Bender, M.J. and Simonovic, S.P. (2000). A fuzzy compromise approach to water resources planning under uncertainty. *Fuzzy Sets Syst.* **115**(1): 35–44.
7. Prodanovic, P.P. and Simonovic, S.P. (2002). Comparison of fuzzy set ranking methods for implementation in water resources decision-making. *Can. J. Civil Eng.* **29**: 692–701.
8. Chang, P.T. and Lee, E.S. (1994). Ranking of fuzzy sets based on the concept of existence. *Comput. Math. Appl.* **27**: 1–21.
9. Simonovic, S.P. (1999). Decision support system for flood management in the Red River Basin. *Can. Water Res. J.* **24**(3): 203–223.
10. International Joint Commission (2000). *Living with the Red.* Ottawa, Washington. Available at http://www.ijc.org/php/publications/html/living.html.
11. Simonovic, S.P. and Carson, R.W. (2003). Flooding in the Red River Basin—lessons from post flood activities. *Nat. Hazards* **28**: 345–365.
12. Salonga, J. (2004). Aggregation methods for multi-objective flood management decision making with multiple stakeholders: Red River Basin case study [Civil Engineering thesis]. The University of Western Ontario, London.

WATER RESOURCES SYSTEMS ANALYSIS

Slobodan P. Simonovic
The University of Western Ontario
London, Ontario, Canada

INTRODUCTION

Water resources planning and management process is a search for the solution of how to meet the needs of a population with the available resources (1). Water resources planning and management is as old as humanity. However, with knowledge and technology development, a change in the living standard of people, and further economic development, the analysis procedure changes.

Principal Objectives for Industrialized Countries

An example of the principal water resources planning and management objectives for industrialized countries is based on the *Principles and Standards for Planning Water and Related Land Resources* used in the United States and introduced by the U.S. Water Resources Council in 1973 and modified in 1979 and 1980 (2). According to them the overall purpose of water resources planning and management is improvement of the quality of life through contributions to national economic development, environmental quality, regional economic development, and other social effects. Under these "principles and standards," the water resources planning process consists of six major steps: (1) specification of the water and related land resources problems and opportunities; (2) inventory, forecast, and analysis of water and related land resource conditions within the planning area relevant to the identified problems and opportunities; (3) formulation of alternative plans; (4) evaluation of the effects of the alternative plans; (5) comparison of alternative plans; and (6) selection of a recommended plan based on the comparison of alternative plans.

Principal Objectives for Developing Countries

Water resources planning and management in developing countries should correspond to a set of criteria distinct from those used in industrialized countries. These criteria should reflect prevailing constraints on physical, financial, and human resources and the need to allocate critically sparse resources to programs that correspond to short- and long-term sociopolitical objectives and promise to be the most cost effective.

The United Nations Industrial Development Organization (UNIDO) (3) recommended that planning analysis consider the following objectives: (1) aggregate consumption, (2) income redistribution, (3) growth rate of national income, (4) employment level, (5) self-reliance, and (6) merit wants. The UN publication treats the problems of evaluating the extent to which projects advance each of the objectives and presents their combination as a measure of "aggregate economic profitability." The UNIDO guidelines, in developing a system of objectives, do not lay any stress on the quality of the environment or other intangible descriptors applied to the quality of human life.

SYSTEMS ANALYSIS

Systems analysis is the use of rigorous methods to help determine preferred plans and designs for complex, often large-scale systems. It combines knowledge of the available analytic tools, understanding of when each is more appropriate, and skill in applying them to practical problems. It is both mathematical and intuitive, as is all planning and design.

Systems analysis is a relatively new field. Its development parallels that of the computer, the computational power of which enables us to analyze complex relationships, involving many variables, at reasonable cost. Most of its techniques depend on the use of the computer for practical applications. Systems analysis may be thought of as the set of computer-based methods essential for the planning of major water resources projects. It is thus central to a modern water resources engineering curriculum.

Systems analysis covers much of the same material as operations research, in particular, linear and dynamic programming and decision analysis. The two fields differ substantially in direction, however. Operations research tends to be interested in specific techniques and their mathematical properties. Systems analysis focuses on the use of the methods.

Systems analysis emphasizes the kinds of real problems to be solved; considers the relevant range of useful techniques, including many besides those of operations research; and concentrates on the guidance they can provide toward improving plans and designs. Use of systems analysis instead of the more traditional set of tools generally leads to substantial improvements in design and reductions in cost. Gains of 30% are not uncommon. These translate into an enormous advantage when one is considering projects worth tens and hundreds of millions of dollars.

DEFINITIONS

There are many variations in the definition of what a system is, but all of the definitions share many common traits. Some kind of system is inherent in all but the most trivial water resources engineering planning and design problems. To understand a problem, the engineer must be able to recognize and understand the system that surrounds and includes it. Some of the reasons for poor system definition in former projects include poor communications, lack of knowledge of interrelationships, politics, limited objectives, and transportation difficulties.

What, then, is a system? The dictionary definition of the term "system" is a mass of verbiage providing no less than 15 ways to define the word. In the most general sense, a system may be defined as a collection of various structural and nonstructural elements that are connected and organized in such a way as to achieve some specific objective by the control and distribution of material resources, energy, and information.

A more formal definition of a system can be stated as

$$\mathrm{S}: \mathbf{X} \to \mathbf{Y} \qquad (1)$$

where **X** is an input vector and **Y** is an output vector. So, a system is a set of operations that transforms input vector **X** into output vector **Y**.

The usual representation of the system definition is presented in Fig. 1. The system's objects are input, output, process, feedback, and a restriction. Input energizes the operation of a given process. The final state of the process is known as the output. Feedback performs a number of operations to compare the actual output with an objective and identifies the discrepancies that exist between them.

To avoid any misunderstanding, let's define some terms often in use: *mathematical model*—a set of equations that describes and represents the real system; *decision variables*—the controllable and partially controllable inputs; *policy*—resulting set of decision variables, when each decision variable is assigned a particular value; *objective function*—quantity used to measure the effectiveness of a particular policy, expressed as a function of the decision variables; *constraints*—physical, economic, or any other restrictions applied to the model; *feasible policy*—a policy that does not violate any constraints; and *policy space*—the subset consisting of all possible feasible policies.

WATER RESOURCES SYSTEM

General formulation of water resources system (4) may be presented as the transformation of available water resources,

$$W_A = \{Q_A, K_A, L_A\} \qquad (2)$$

into required water resources,

$$W_D = \{Q_D, K_D, L_D\} \qquad (3)$$

taking into consideration water quality protection, flood control, and regional and national development plans. Notation used in Equations 2 and 3 includes: Q_A—available quantity of water; Q_D—required quantity of water; K_A—quality of available water; K_D—required water quality; L_A—location of available resources; and L_D—location of demand points.

Mathematical expression of the transformation problem can be given as

$$W_D = T \times W_A \qquad (4)$$

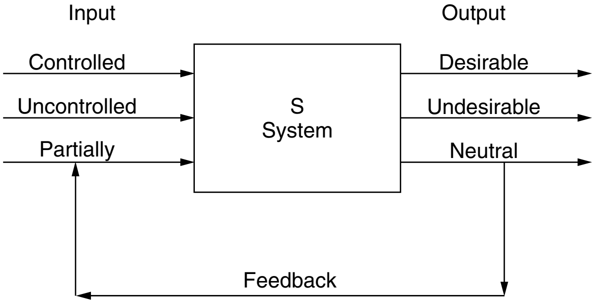

Figure 1. System presentation.

This transformation may be performed through the following three processes:

1. *Water Resources Planning*—formulation of goals and objectives that are consistent with political, social, environmental, economic, technological, and aesthetic constraints; and the general definition of procedures designed to meet those goals and objectives.
2. *Water Resources Design*—selection of a particular combination of resources and a way to use them.
3. *Water Resources Operations*—implementation of procedures in order to achieve preselected goals and objectives.

SYSTEMS APPROACH

The systems approach is a general problem-solving technique that brings more objectivity to the water resources planning/design processes. It is, in essence, good design; a logical and systematic approach to problem solution in which assumptions, goals, objectives, and criteria are clearly defined and specified. Emphasis is placed on relating system performance to these goals. A hierarchy of systems, which allows handling of a complex system by looking at its component parts or subsystems, is identified. Quantifiable and nonquantifiable aspects of the problem are identified, and immediate and long-range implications of suggested alternatives are evaluated.

The systems approach establishes the proper order of inquiry and helps in the selection of the best course of action that will accomplish a prescribed goal by broadening the information base of the decision-maker; by providing a better understanding of the system and the interrelatedness of the system and its component subsystems; and by facilitating the prediction of the consequences of several alternative courses of action.

The systems approach is a framework for analysis and decision-making. It does not solve problems, but it does allow the decision-maker to undertake resolution of a problem in a logical, rational manner. While there is some art involved in the efficient application of the systems approach, other factors play equally important roles. The magnitude and complexity of decision processes require the most effective use possible of the scientific (quantitative) methods of systems analysis. However, one has to be careful not to rely too heavily on the methods of systems analysis. Outputs from simplified analyses have a tendency to take on a false validity because of their complexity and technical elegance.

The steps in the systems approach include (5): definition of the problem; gathering of data; development of criteria for evaluating alternatives; formulation of alternatives; evaluation of alternatives; choosing the best alternative; and final design/plan implementation. Often several steps in the systems approach are considered simultaneously, facilitating feedback and allowing a natural progression in the problem-solving process.

The systems approach has several defining characteristics. It is a repetitive process, with feedback allowed from any step to any previous step. Frequently, because systems analysis takes such a broad approach to problem-solving, interdisciplinary teams must be called in. Coordination and commonality of technique among the disciplines is sometimes hard to achieve. Close communication among the parties involved in applying the systems approach is essential if this understanding is to be achieved.

Definition of the Problem

Problem definition may require iteration and careful investigation, because problem symptoms may mask the true cause of the problem. A key step of problem definition requires the identification of any systems and subsystems that are part of the problem, or related in some way to it. This set of systems and interrelationships is called the environment of the problem. This environment sets the limit on factors that will be considered when analyzing the problem. Any factors that cannot be included in the problem environment must be included as inputs to, or outputs from, the problem environment.

Gathering of Data

Gathering of data to assist in planning and design decision-making through the systems approach will generally be done in conjunction with several steps. Some background data will have to be gathered at the problem definition stage and data gathering and analysis will continue through the final plan/design and implementation stage. Data that are gathered as the approach continues will help to identify when feedback to a previous step is required.

Data will be required at the problem definition stage to evaluate if a problem really exists; to establish what components, subsystems, and elements can reasonably be included in the delineation of the problem environment; and to define interactions between components and subsystems. Data will be needed during later steps to establish constraints on the problem and systems involved in it, to increase the set of quantifiable variables and parameters (constants) through statistical observation or development of measuring techniques, to suggest what mathematical models might contribute effectively to the analysis, to estimate values for coefficients and parameters used in any mathematical models of the system, and to check the validity of any estimated system outputs. When feedback is required, the data previously acquired can assist in redefining the problem, systems, or system models.

Development of Evaluative Criteria

Evaluative criteria must be developed to measure the degree of attainment of system objectives. These evaluative criteria will facilitate a rational choice of a particular set of actions (from among a large number of feasible alternatives), which will best accomplish the established objectives. Some evaluative criteria will provide an absolute value of how good the solution is, such as the cost of producing one unit of some product. Other evaluative criteria will only produce relative values that can be compared among the alternatives to rank them in order of preference, as in economic comparisons such as benefit/cost analysis.

In most complex real-world water problems, more than one objective can be identified. A quantitative or qualitative analysis of the trade-offs between the objectives must be made. For many problems, cost effectiveness would be the primary objective. Cost effectiveness can be defined as the lowest possible cost for a set level of control of a system, or the highest level of system control for a set cost.

Formulation of Alternatives

Formulation of alternatives is essentially the development of system models that, in conjunction with evaluative criteria, will be used in later analysis and decision-making. If at all possible, these models should be mathematical in nature. However, it should not be assumed that mathematical model building and optimization techniques are either required or sufficient for application of the systems approach. Many problems contain unquantifiable variables and parameters that would render results generated by even the most elegant mathematical model meaningless. If it is not practical to develop mathematical models, subjective models that describe the problem environment and systems included can be constructed. Models allow a more explicit description of the problem and its systems and facilitate the rapid examination of alternatives. Effective model building is a combination of art and science. The science includes the technical principles of mathematics, physics, and engineering science. The art is the creative application of these principles to describe physical or social phenomena.

Evaluation of Alternatives

To evaluate the alternatives that have been developed, some form of analysis procedure must be used. Numerous mathematical techniques are available, including the simplex method for linear programming models, the various methods for solving ordinary and partial differential equations or systems of differential equations, matrix algebra, various economic analyses, and deterministic or stochastic computer simulation. Subjective analysis techniques may be used for multiobjective analysis, or subjective analysis of intangibles. The appropriate analysis procedures for a particular problem will generate a set of solutions for the alternatives that can be tested according to the established evaluative criteria. In addition, these solution procedures should allow efficient utilization of personnel and computational resources.

Choosing the Best Alternative

Choice of the best alternative from among those analyzed must be made in the context of the objectives and evaluative criteria previously established, but also must take into account nonquantifiable aspects of the problem such as aesthetic and political considerations. The chosen alternative will greatly influence the development of the final plan/design and will determine in large part the implementability of the suggested solution. Preferably, the best alternative can be chosen from the mathematical optimization within feasibility constraints. Frequently, however, a system cannot be completely optimized. Near optimum solutions can still be useful, especially if sensitivity analysis has shown that the solution (and thus the objective function) is not sensitive to changes in the decision variables near the optimum point.

Final Plan/Design Implementation

Actual final planning/design is primarily a technical matter that is conducted within the constraints and specifications developed in earlier stages of the systems approach. One of the end products of final planning or design is a report that describes the recommendations made.

MATHEMATICAL MODELING

In general, to obtain a way to control or manage a water resources system we use a mathematical model that closely represents the physical system. Then the mathematical model is solved and its solution is applied to the physical system. Models, or idealized representations, are an integral part of everyday life. Mathematical models are also idealized representations, but they are expressed in terms of mathematical symbols and expressions. Such laws of physics as $F = ma$ and $E = mc^2$ are familiar examples. Similarly, the mathematical model of a business problem is the system of equations and related mathematical expressions that describe the essence of the problem. Thus, if there are n related quantifiable decisions to be made, they are represented as decision variables (say, $x_1, x_2, \ldots, x_n$) whose respective values are to be determined. The appropriate measure of performance (e.g., profit) is then expressed as a mathematical function of these decision variables (e.g., $P = 3x_1 + 2x_2 + \cdots + 5x_n$). This function is called the objective function. Any restrictions on the values that can be assigned to these decision variables are also expressed mathematically, typically by means of inequalities or equations (e.g., $x_1 + 3x_1x_2 + 2x_2 \leq 10$). Such mathematical expressions for the restrictions often are called constraints. The constants (coefficients or right-hand sides) in the constraints and the objective function are called the parameters of the model. The mathematical model might then say that the problem is to choose the values of the decision variables so as to maximize the objective function, subject to the specified constraints.

Mathematical models have many advantages over a verbal description of the problem. One obvious advantage is that a mathematical model describes a problem much more concisely. This tends to make the overall structure of the problem more comprehensible, and it helps to reveal important cause-and-effect relationships. In this way, it indicates more clearly what additional data are relevant to the analysis. It also facilitates dealing with the problem in its entirety and considering all its interrelationships simultaneously. Finally, a mathematical model forms a bridge to the use of high-powered mathematical techniques and computers to analyze the problem. Indeed, packaged software for both microcomputers and mainframe computers is becoming widely available for many mathematical models.

The procedure of selecting the set of decision variables that maximizes/minimizes the objective function, subject to the systems constraints, is called the optimization procedure. The following is a general optimization problem. Select the set of decision variables $x_1^*, x_2^*, \ldots, x_n^*$ such that

$$\text{Min or Max}\, f(x_1, x_2, \ldots, x_n)$$

subject to

$$\begin{aligned} g_1\ (x_1, x_2, \ldots, x_n) &\leq b_1 \\ g_2\ (x_1, x_2, \ldots, x_n) &\leq b_2 \\ g_m\ (x_1, x_2, \ldots, x_n) &\leq b_m \end{aligned} \tag{5}$$

where $b_1, b_2, \ldots, b_m$ are known values.

If we use matrix notation, Equation 5 can be rewritten as

$$\text{Min or Max}\, f(x) \tag{6}$$

subject to

$$g_j\ (\mathbf{x}) \leq b_j, \quad j = 1, 2, \ldots, m$$

When optimization fails, due to system complexity or computational difficulty, a reasonable attempt at a solution may often be obtained by simulation. Apart from facilitating trial and error design, simulation is a valuable technique for studying the sensitivity of system performance to changes in design parameters or operating procedure.

According to Equations 5 and 6, our main goal is the search for an optimal or best solution. However, it needs to be recognized that these solutions are optimal only with respect to the model being used. Since the model necessarily is an idealized rather than an exact representation of the real problem, there cannot be any utopian guarantee that the optimal solution for the model will prove to be the best possible solution that could have been implemented for the real problem. There just are too many imponderables and uncertainties associated with real problems. However, if the model is well formulated and tested, the resulting solution should tend to be a good approximation to the ideal course of action for the real problem. Therefore, rather than be deluded into demanding the impossible, the test of the practical success of an operations research study should be whether it provides a better guide for action than can be obtained by other means.

The eminent management scientist and Nobel Laureate in Economics, Herbert Simon, points out that satisficing is much more prevalent than optimizing in actual practice. In coining the term *satisficing* as a combination of the words *satisfactory* and *optimizing*, Simon is describing the tendency of water resources managers to seek a solution that is "good enough" for the problem at hand. Rather than trying to develop various desirable objectives (including well-established criteria for judging the performance of different segments of the organization), a more pragmatic approach may be used. Goals may be set to establish minimum satisfactory levels of performance in various areas, based perhaps on past levels of performance or on what the competition is achieving. If a solution is found that enables all of these goals to be met, it is likely to be adopted without further ado. Such is the nature of satisficing. The distinction between optimizing and satisficing reflects the difference between theory and the realities frequently faced in trying to implement that theory in practice.

BIBLIOGRAPHY

1. Goodman, A.S. (1984). *Principles of Water Resources Planning*. Prentice-Hall, Englewood Cliffs, NJ.
2. U.S. Water Resources Council (1973). *Principles and Standards for Planning Water and Related Land Resources*, September 10, 1973; Rev. December 14, 1979; Rev. September 29, 1980.
3. Dasgupta, P., Sen, A., and Marglin, S. (1972). *Guidelines for Project Evaluation*. United National Industrial Development Organization, Vienna.
4. Simonovic, S.P. (1989). Application of water resources systems concept to the formulation of a water master plan. *Water Int.* **14**: 37–50.
5. Jewell, T.K. (1986). *A Systems Approach to Civil Engineering Planning and Design*. Harper and Row, New York.

SURFACE WATER HYDROLOGY

ACIDIFICATION—CHRONIC

Tim Sullivan
E & S Environmental
Corvallis, Oregon

The objective here is to describe and discuss in general terms, processes that acidify freshwaters (including lakes, rivers, and streams), agents of acidification, watershed features associated with sensitivity to acidification, and associated chemical and biological effects. This is a complicated topic. A variety of acidifying agents and a number of physical and chemical watershed characteristics can make a particular body of water susceptible to acidification. Chronic, long-term acidification of water is addressed, not the short-term (episodic) acidification that sometimes accompanies rainstorms or snowmelt.

First, it is necessary to define a few terms. *Acidification* is generally a decrease in the acid-neutralizing capacity (ANC) of water. It can also be defined as a decrease in pH. ANC refers to the capacity of a solution to neutralize strong acids. It can be measured in a laboratory, usually by the Gran titration procedure. It can also be defined in different ways, based on the measured values of various ion concentrations in the water. Many mathematical models of acid–base chemistry define ANC as the sum of the base cation concentrations ($Ca^{2+} + Mg^{2+} + Na^{+} + K^{+} + NH_4{}^{+}$ [termed SBC]) minus the sum of the mineral or strong acid anion concentrations ($SO_4{}^{2-} + NO_3{}^{-} + Cl^{-}$ [termed SAA]).

$$\mathrm{ANC} = \mathrm{SBC} - \mathrm{SAA} \tag{1}$$

where all ions are in units of microequivalents per litre (μeq/L). ANC defined in this way is approximately equal to laboratory measurements (Gran titrated) of ANC only if the solution contains relatively low concentrations of dissolved organic carbon (DOC) and Al. Both of these latter constituents cause Gran ANC to differ from ANC defined as SBC − SAA (1). ANC reflects the extent to which added strong acids can be neutralized, or buffered, by the nontoxic base cations in solution such as Ca^{2+} rather than by H^{+} and Al^{n+}, which are toxic to some species of aquatic animals and other life forms. The term *acidic* is used to describe a lake, river, or stream (surface water) that has ANC below zero. In other words, if the water is acidic, then the sum of the concentrations of the strong acid anions exceeds the sum of the concentrations of the base cations (SAA > SBC). Thus, a body of water can become *acidified* by increasing the concentration of one or more of the SAA components, by decreasing the concentration of one or more of the SBC components, or a combination of both. There is a rather consistent relationship between pH and ANC, although varying levels of DOC contribute to scatter in this relationship. At ANC = 0, the pH is generally near 5 (Fig. 1).

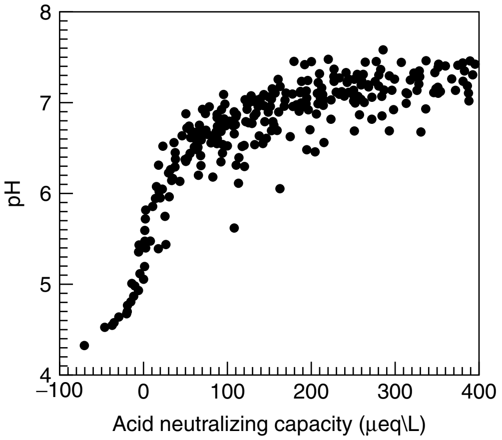

Figure 1. Relationship between pH and ANC in streams sampled by the U.S. EPA's Environmental Monitoring and Assessment Project (EMAP) in the Mid-Atlantic Appalachian Mountains.

Weak acids in solution, such as carbon dioxide, aluminum hydroxides, and organic acids, act as buffer systems that prevent dramatic changes in water pH upon adding small amounts of strong acid. When CO_2 from the atmosphere dissolves in water, it forms carbonic acid (H_2CO_3), which readily dissociates into hydrogen and bicarbonate ($HCO_3{}^{-}$) ions. Bicarbonate constitutes the most important buffer system in most freshwaters because it undergoes the following reaction upon adding strong acid:

$$HCO_3{}^{-} + H^{+} \rightarrow CO_2 + H_2O \tag{2}$$

At low pH (<5.5), the bicarbonate buffering system becomes less important, and the aluminum hydroxide buffering system becomes proportionately more important. In waters that have high DOC, the organic acid buffering system can dominate. In addition to providing pH buffering upon adding strong acid, organic acids also reduce the pH of the solution by up to a full pH unit or more.

Chronic acidification of fresh waters in North America is usually associated with an increased concentration of $SO_4{}^{2-}$. The source of this $SO_4{}^{2-}$ can be air pollution (by acidic deposition, also called acid rain) or geologic sources of sulfur. Acidic deposition is the process whereby S and/or N is transported in the atmosphere and deposited downwind of pollution sources onto water bodies and the terrestrial watersheds that drain into them. The major sources of atmospheric sulfur include coal-fired power plants and industrial facilities. Geologic S can be contributed to surface waters by acid mine drainage (AMD; 2) or other watershed disturbance and subsequent oxidation of sulfur-bearing minerals.

There are also examples of surface waters that have been acidified partially, or even mostly, by increased concentration of $NO_3{}^{-}$. In such cases, the source of the

NO_3^- is generally atmospheric. Nitrogen, in the form of nitrogen oxides and ammonia (which can be converted to NO_3^-), is emitted to the atmosphere largely from motor vehicles, industrial sources, and agriculture. NO_3^- is more likely to be a major contributor to short-term episodic acidification of surface waters, whereas SO_4^{2-} is more commonly associated with long-term chronic acidification. There are numerous exceptions to these general patterns, however.

It is also possible for waters to become chronically acidified due to Cl^- contamination, but this rarely occurs to any degree. Atmospheric deposition of Cl^- in sea spray can cause episodic acidification of surface waters in near-coastal areas, and road salt application can probably contribute in some cases to both chronic and episodic acidification, but there is little evidence that this is of any regional importance.

Thus, SO_4^{2-}, NO_3^-, and Cl^-, alone or in combination, can contribute to chronic acidification of surface waters. Where chronic acidification of North American waters occurs, it is most often from SO_4^{2-}.

Acid mine drainage is an important cause of sulfur-based chronic surface water acidification in mining regions of North America. This problem is especially prevalent in the northern Appalachian Mountains of West Virginia and Pennsylvania, where almost 10% of the streams are chronically acidic as a consequence of AMD (2). AMD results from the exposure to air and water of mine spoils containing sulfide minerals. The minerals are oxidized by a series of chemical reactions and microbial processes to yield high concentrations of sulfate and heavy metals in drainage water and low pH.

In many cases, the added SO_4^{2-} that contributes to chronic acidification is of atmospheric origin. However, it is important to note that, in most regions that receive high levels of S deposition, the primary determinant of whether, and to what extent, a body of water acidifies is not the amount of SO_4^{2-} that is added from the atmosphere and/or from geologic sources. Rather, the most important factor governing surface water acidification is usually the inherent sensitivity of the watershed. In most areas of North America, most watersheds are not sensitive to surface water acidification from atmospheric S deposition because abundant supplies of base cations in watershed soils usually exist. In such nonsensitive watersheds, increased contributions of SO_4^{2-} from the atmosphere are balanced by increased release of base cations from soils, and the drainage water does not acidify. However, some fraction of the surface waters in some regions are acid-sensitive, and increased concentrations of SO_4^{2-} can decrease surface water ANC. To understand this concept, it is helpful to consider the principle of electroneutrality.

The sum of the charges of all positively charged ions (cations) must equal the sum of the charges of all negatively charged ions (anions) in solution, so that

$$\sum \text{Cations} = \sum \text{Anions} \tag{3}$$

In relatively undisturbed freshwaters, these sums can be approximated as

$$\sum \text{Cations} = \text{SBC} + \text{H}^+ + \text{Al}^{n+} \tag{4}$$

and

$$\sum \text{Anions} = \text{SAA} + \text{HCO}_3^- + \text{Org}^- \tag{5}$$

where all ions are in units of μeq/L and Org^- represents a mixture of ill-defined organic acid anions. The valence of Al in Eq. 4 is expressed as n+ because inorganic ionic Al in solution can include a number of species, with varying charges. At a pH less than about 4.5, much of the Al in solution is present as Al^{3+}. At somewhat higher pH (about 4.5 to 5.5), proportionately more of the Al in solution is present as Al hydroxides and Al fluorides whose charges are +1 or +2. At a pH above 5.5, there is little Al in solution. Other ions, such as F^-, Fe^{n+}, Mn^{n+}, for example, are usually present in minor amounts in natural waters and have little impact on the charge balance.

If the concentration of one of the SAA anions, for example, SO_4^{2-}, increases in solution, then one of the other anions (e.g., HCO_3^- or Org^-) has to decrease in concentration and/or one or more of the cations has to increase in concentration to maintain the electroneutrality indicated in Eq. 3.

Combining Eqs. 3–5, we can express the charge balance as

$$\text{SBC} + \text{H}^+ + \text{Al}^{n+} = \text{SAA} + \text{HCO}_3^- + \text{Org}^- \tag{6}$$

Equation 6 can be rearranged to yield the following:

$$\text{SBC} - \text{SAA} = \text{HCO}_3^- + \text{Org}^- - \text{H}^+ - \text{Al}^{n+} \tag{7}$$

and therefore,

$$\text{ANC} = \text{HCO}_3^- + \text{Org}^- - \text{H}^+ - \text{Al}^{n+} \tag{8}$$

Thus, ANC can be defined in more than one way. If Equation 1 holds true, then Equation 8 must also hold true to satisfy the electroneutrality constraint (Eq. 3).

These equations and definitions presume that the mixture of organic acid anions in the water (Org^-) is made up entirely of weak acid anions that contribute ANC to solution. However, Driscoll et al. (3) showed that about one-third of the organic acid anions in Adirondack Mountain, New York, lakes actually exhibit strong acid characteristics and behave more like SO_4^{2-} and other strong mineral acid anions. The remainder of the organic acids act as weak acids. Thus, the contribution of organic acid anions to surface water acid–base chemistry is rather complicated. These substances contain a multitude of functional groups, some of which contribute ANC and some of which lower ANC.

As SAA concentration in surface water increases, the concentration of one or more of the other anions must decrease, and/or one of the cations must increase in concentration to maintain electroneutrality. For surface waters that are relatively high in ANC (greater than about 50 to 100 μeq/L), most of the compensating change is usually an increase in the concentration of Ca^{2+} and other base cations (SBC). In such cases, there is little or no water acidification from the increase in SAA; both the SAA and SBC increase, but the ANC(=SBC − SAA) stays about the same. For surface waters that are relatively low in ANC, in contrast, the compensating change generally involves multiple responses, including decreased HCO_3^- and increased SBC, H^+, and Al^{n+}. Increased H^+ and Al^{n+} can be toxic to aquatic life. In some cases, especially if DOC is relatively high, Org^- also decreases. This latter effect can influence the relative distribution of Al forms between nontoxic organic complexes and inorganic species, some of which are toxic.

The effects of acidification on aquatic life are due mainly to the changes that occur in HCO_3^-, H^+, and/or Al^{n+}. Any change in SBC to compensate for an increase in SAA does not result in acidifying water, although it can contribute to acidification of the soil. Because of the fundamental importance of these processes to water acidification, Henriksen (4) defined a factor, termed the F-factor, as the proportional change in SBC, as opposed to HCO_3^-, H^+, and Al^{n+}, relative to the change in SAA:

$$F = \frac{\Delta \text{SBC}}{\Delta \text{SAA}} \tag{9}$$

In watersheds that are not acid-sensitive, $F \approx 1.0$. In highly to slightly acid-sensitive Adirondack lake watersheds, F varies from about 0.4 to near 1.0.

Thus, for every 1 μeq/L increase in SO_4^{2-} concentration in an acid-sensitive lake or stream, the change in base cation concentrations will often balance much of that increase, and the resulting ANC change will generally be a rather small fraction (often less than half) of 1 μeq/L. There are, however, many highly sensitive watersheds in which ANC change constitutes an appreciable component of the increase in SAA. Within North America, such watersheds are especially prevalent in high mountainous areas of the western United States.

Waters sensitive to chronic acidification from acidic deposition generally had ANC less than about 50 μeq/L prior to the advent of the acidic deposition. In some cases, surface waters having somewhat higher initial ANC (∼50 to 100 μeq/L) have become acidified. Sensitive waters are usually found at moderate to high elevation, in areas of high relief (mountainous), rapid runoff of precipitation (flashy hydrology), and minimal opportunity for contact between drainage water and soils or geologic materials that may contribute weathering products (base cations) to solution. Sensitive streams generally occur in small watersheds (<10 km^2). These are generally the highest, smallest, coldest, and highest gradient streams and their associated lakes. Most acid-sensitive surface waters are underlain by bedrock that is resistant to weathering and/or contains only thin deposits of glacial till. Soils in the watersheds of acid-sensitive surface waters contain relatively small amounts of exchangeable base cations, and the soil base saturation is generally less than 10 to 15%. It is rare to find large lakes or large rivers that are acid-sensitive because large lakes and rivers have large watersheds and therefore, the probability increases that some of the bedrock and soils encountered by drainage water will be rich in base cations.

Chronic acidification of lakes and streams has occurred in areas throughout portions of the eastern United States, eastern Canada, and in parts of the upper Midwest. The areas most heavily impacted include the Adirondack and Catskill Mountains in New York, portions of the Appalachian Plateau in West Virginia, southeastern Canada, the mountains of western Virginia, and the Upper Peninsula of Michigan (5). In some portions of these regions, as many as 10 to 50% of the lake and/or stream resources may have experienced some degree of chronic acidification from sulfur deposition. Some of the most acid-sensitive watersheds in the world occur in high mountain areas of the western United States, including the Front Range of Colorado, northern Rocky Mountains, Sierra Nevada, and Cascade Mountains. However, because the levels of acidic deposition are generally low in most parts of the western United States, the amount of actual chronic acidification has been rather limited in most of these highly sensitive areas. Aquatic ecosystems in the Front Range of Colorado have probably been the most impacted of these western mountain ranges to date, and there the acidification has been more strongly from N than from S.

In general, most of the SO_4^{2-} that is deposited from the atmosphere to a watershed acts as a mobile anion and moves through soils into surface waters. In other words, SO_4^{2-} inputs to the watershed in the form of wet, dry, and cloud deposition are approximately equal to SO_4^{2-} outputs in drainage water. This steady-state condition is approximated in most areas that were previously glaciated, for example, throughout the northeastern United States and southeastern Canada. In the southeastern United States, however, which was not glaciated, soils tend to be much older and more highly weathered. Such soils adsorb some of the deposited sulfur before it reaches surface waters. This process effectively prevents water acidification, at least temporarily. Over time, however, the capacity of the soils to adsorb S becomes depleted as the S adsorption sites on the soil become saturated by continuing S inputs. Therefore, proportionately more of the SO_4^{2-} inputs reach surface waters, potentially contributing to acidification. At the present time, the soils in many areas in the southeastern United States, generally from Virginia to Tennessee and Georgia, are gradually losing their capacity to adsorb S. This will have a profound impact on watershed responses to future changes in S deposition. Model projections suggest that even if S deposition is reduced by more than 70% from 1995 levels, most streams in this region will continue to show an increase in streamwater SO_4^{2-} concentration, which will contribute to further acidification (6). Thus, southeastern U.S. watersheds exhibit a delayed response to S deposition due partly to gradual changes in the extent to which S is adsorbed on watershed soils.

Base cation depletion is another important process that is believed to contribute to a delayed watershed response to acidic deposition. This process is not confined to a particular region, but rather can be important in any watershed that has a low base cation supply and receives moderate to high levels of acidic deposition. Because ANC reflects the difference between the concentrations of SBC and SAA, a decrease in base cation concentrations can contribute to acidification just as an increase in one or more of the SAA constituents such as SO_4^{2-} or NO_3^-. It is believed that continuing levels of acidic deposition are depleting the exchangeable base cation stores in some watershed soils to the point that drainage water acidification is occurring in response to decreased base cation leaching from soils to drainage water. Such a process is not likely unless the soil base saturation (percent of the cation exchange capacity provided by base cations, rather than H or Al) is quite low, perhaps less than about 10%. Several factors can contribute to soils that have low base saturation, including long periods of acidic deposition, land use, (especially logging), and other disturbances that cause substantial erosion. In addition, some soils are naturally low in base cations as a consequence of the geologic makeup of their parent material and soil-forming processes. Recent decreases in the atmospheric deposition of base cations has also contributed to this problem.

In regions characterized by abundant wetlands, natural organic acidity from the breakdown of plant material can have a large impact on the acid–base chemistry of drainage waters. Lakes and streams in such regions are often characterized by relatively high concentrations of DOC (greater than about 5 or 6 mg C/L or 400–500 μM) and associated organic acids. Such lakes are often naturally acidic or low in ANC and show relatively little response of pH to changes in acidic deposition. For example, of the 560 Adirondack Mountain lakes considered most sensitive to acidic deposition (the thin till drainage and mounded seepage lake types), 45% had DOC higher than 6 mg C/L (500 μM) (7). Much of the acidity in these lakes is organic.

Chronic acidification can adversely impact many species of aquatic life, from algae to fish and fish-eating birds. Different species vary in their acid-sensitivity, as do different life forms within a given species. In lakes and streams of the eastern United States, much of the attention has focused on damage to native brook trout (*Salvelinus fontinalis*), which often would be expected in acid-sensitive lakes and streams. However, brook trout are relatively tolerant of acidity in comparison with some of the lesser known forage fish species such as blacknose dace (*Rhinichthys atratulus*). In most cases, brook trout are unable to live in waters that have chronic ANC less than about zero. Adverse episodic effects can occur at higher levels of chronic ANC in the range of 0 to 20 μeq/L or higher. In general, the eggs and larval stages of fish are more acid-sensitive than adults.

Acidification of streamwater also results in adverse impacts on several important orders of aquatic insects, especially mayflies. Acidified streams tend to contain fewer species or genera of these insects than streams that have not acidified. In lakes acidified by acidic deposition, several species of *Daphnia* (Cladocera) are often highly impacted.

Many species of algae, especially diatoms, are acid-sensitive. In fact, the occurrence and relative abundance of the various diatom species can be used to infer past lakewater chemistry. This technique has been used to estimate the acid–base characteristics of lakes during preindustrial times, as a way of quantifying how much acidification has occurred in response to a century of acidic deposition. Such paleoecological studies have involved collecting cores of lake sediment. The cores are then sliced into wafers, which are dated using radioisotopic dating techniques such as ^{210}Pb analysis. The remains of diatom species in each layer provide information about the lake conditions at the time that the sediment layer was deposited at the bottom of the lake (8). Such studies have shown, for example, that of the currently acidic Adirondack lakes, the median lake has acidified by about 37 μeq/L since preindustrial times (9).

Current research on the chronic effects of surface water acidification is focused largely on documenting and predicting the extent of chemical and biological recovery that will occur in response to recent and projected future large decreases in atmospheric sulfur deposition. An additional important research focus concerns identification of the critical load of S or N deposition required to protect against surface water acidification to harmful levels or to allow recovery of acidified waters to chemical conditions that are no longer harmful to aquatic life. Much of this work is focused on lakes and streams in wilderness areas, national parks, and other protected areas.

Additional information on this topic can be found in books by Charles (5) and Sullivan (9).

Acknowledgments

Helpful comments on an earlier draft were provided by A. Herlihy, J. Harrison, and L. Sullivan.

BIBLIOGRAPHY

1. Sullivan, T.J. et al. The influence of organic acid anions and aqueous aluminum on measurements of acid neutralizing capacity in surface waters. *Nature* **338**: 408–410.
2. Herlihy, A.T., Kaufmann, P.R., Mitch, M.E., and Brown, D.D. (1990). Regional estimates of acid mine drainage impact on streams in the Mid-Atlantic and Southeastern United States. *Water Air Soil Pollut.* **50**: 91–107.
3. Driscoll, C.T., Lehtinen, M.D., and Sullivan, T.J. (1994). Modeling the acid–base chemistry of organic solutes in Adirondack, New York, lakes. *Water Resour. Res.* **30**: 297–306.
4. Henriksen, A. (1984). Changes in base cation concentrations due to freshwater acidification. *Verh. Int. Verein. Limnol.* **22**: 692–698.
5. Charles, D.F. (Ed.). (1991). *Acidic Deposition and Aquatic Ecosystems: Regional Case Studies*. Springer-Verlag, New York.
6. Sullivan, T.J. et al. (2004). Regional model projections of future effects of sulfur and nitrogen deposition on streams in the Southern Appalachian Mountains. *Water Resour. Res.* **40**: WO2101, doi:10.1029/2003WR001998.
7. Driscoll, C.T. et al. (1998). The response of lake water in the Adirondack region of New York to changes in acidic deposition. *Environ. Sci. Pollut.* **1**: 185–198.
8. Charles, D.F., Battarbee, R.W., Renberg, I., van Dam, H., and Smol, J.P. (1989). Paleoecological analysis of lake acidification trends in North America and Europe using diatoms and chrysophytes, In: *Acid Precipitation*, *Soils, Aquatic Processes, and Lake Acidification*. Vol. 4. Advances in Environmental Sciences. S.A. Norton, S.E. Lindberg, and A.L. Page (Eds.). Springer-Verlag, New York, pp. 207–276.
9. Sullivan, T.J. (2000). *Aquatic Effects of Acidic Deposition*. Lewis, Boca Raton, FL.

EPISODIC ACIDIFICATION

Myron J. Mitchell
State University of New York
Syracuse, New York

Episodic acidification is the short-term decrease (≤ 0 μmol$_c$ L^{-1}) in acid neutralizing capacity (ANC) of surface waters due to both natural and anthropogenic factors. ANC is a measure of the ability of water or soil to neutralize added acids and is defined by the reaction of hydrogen ions with inorganic (e.g., HCO_3^-) or organic ions. Decreases in ANC generally result in a decline in pH, sometimes dropping to levels below 4.5 (1). Episodic acidification most commonly occurs during snowmelt or precipitation that results in rapid increases in the discharge rates of surface waters. Typically the decrease in ANC and pH of streams and lakes is accompanied by increases in Al^{n+} concentrations (2). These decreases in ANC and pH and increases in Al^{n+} concentrations have marked effects on water chemistry and freshwater biota.

RELATIONSHIP TO CHRONIC ACIDIFICATION

The importance of surface water acidification has been clearly recognized as an important environmental issue that has often been linked with atmospheric deposition of acidic compounds, such as sulfuric (H_2SO_4) and nitric acid (HNO_3). It has been suggested that acidic deposition has increased the extent and severity of episodic acidification. Both chronic and episodic acidification can be due to both natural (e.g., organic acids) and anthropogenic processes (e.g., strong mineral acids that are associated with SO_4^{2-} and NO_3^-). Often, acidification episodes are most marked in those surface waters near circumneutral pH (3). For surface waters at chronically low pH levels, episodic acidification is relatively less important and in some cases may not be evident (4).

CAUSES OF ACIDIFICATION

Hydrologic Flow Paths

The chemical characteristics and the relative magnitude of an acidic episode are linked to hydrologic flow paths. Hydrologic flow paths vary with landscape features and are also influenced by antecedent moisture conditions. Flow paths affect both the chemical attributes and the reaction time of the water as it drains from terrestrial to aquatic environments. Major hydrologic events such as snowmelt or large rainstorms can cause marked changes in discharge rates and rapid changes in surface water chemistry. Periods of high discharge are accompanied by dynamic changes in the size of water pools (e.g., riparian areas, soil water), flow paths and the relative contributions of different water sources to surface water. A considerable focus has been placed on evaluating the importance of snowmelt episodes in affecting surface water chemistry, especially in those areas of high amounts of acidic deposition and extended periods of snow accumulation. During snowmelt episodes, the most marked changes in discharge rates often occur during “rain on snow.” Such events can result in rapid movement of solutes to surface waters. In addition, flow paths (e.g., water flux via macropores) may develop during high soil moisture conditions that result in little contact of storm water in soil neutralizing processes and thus exacerbate acid pulses. In contrast, during dry conditions, H^+ and mobile anions (SO_4^{2-} and NO_3^-) can accumulate in soils and these ions will be flushed from the soil only during large rainstorms.

Chemical Relationships

Chemical factors affecting episodic acidification vary spatially and temporally among watersheds due to differences in geology, soils, climate, landscape features, and anthropogenic influences.

Dilution of Base Cations. In circumneutral streams that have relatively high concentrations of base cations (especially Ca^{2+}), the dilution of these base cations (as much as 100 to 280 μmol$_c$ L^{-1}) due to rapid increases in water discharge can contribute to rapid declines in ANC.

Sulfate. Pulses of SO_4^{2-} play a major role in episodic acidification in Europe, Canada, and some limited regions in the United States. The importance of SO_4^{2-} in contributing to acidic pulses has been most clearly shown for Europe, especially in those regions that have had much higher atmospheric inputs of sulfur than generally found in North America.

Nitrate. Fluxes of NO_3^- in contributing to acidification are most important in the northeast United States including the Adirondack and Catskill Mountains of New York State. Pulses of NO_3^- are mostly attributed to snowmelt episodes, but marked losses of NO_3^- can also occur during large precipitations that follow extended periods of low water availability. Some of the NO_3^- is derived directly from the snowpack, but most of NO_3^- comes from the soil which has accumulated NO_3^- from the microbially mediated processes of ammonification (N mineralization) and nitrification.

DOC. The contribution of organic acid (DOC) is important in many regions, and its relative role is most evident at those sites where wetlands are notable. Organic acids have been found especially important under saturated conditions in those catchments that have large proportions of peatlands (5).

Sea Salt Effect. The salt effect is an important contributor to chronic acidification in coastal areas of the northeast United States and western Europe. The deposition of neutral sea salts in coastal areas can cause pH and ANC depression (Fig. 1).

MODELING OF EPISODIC ACIDIFICATION

Various types of models have been used in the study of episodic acidification, including simulation and empirical approaches. A large number of physically and chemically

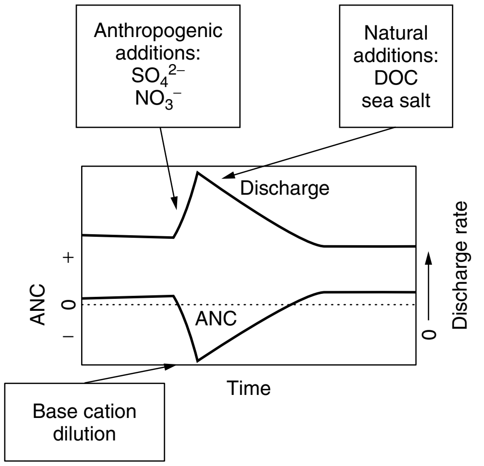

Figure 1. Schematic diagram of factors affecting episodic acidification and changes over time in acid neutralizing capacity (ANC) and discharge rate.

based simulation models have been used to predict acidification (6). Empirical models include four types: solute concentration versus stream discharge, mixing models (such as EMMA, end-member mixing analyses), time series relationships, and multiple linear regression models. The regression model approach has been particularly useful for regional predictions of which surface water sites may be most susceptible to acidic episodes (7).

REGIONAL DIFFERENCES

Episodic acidification has been well documented for the United States, Canada, and Europe. The causes of episodic acidification vary among geographical regions. Within the United States, most of the surface waters showing episodic acidification are in the Northeast and all regions document the importance of base cation dilution. Organic acid (DOC) contributions are most important in the Adirondack Mountains. The importance of NO_3^- is most evident in the Adirondack and Catskill Mountains. In Pennsylvania, changes in SO_4^{2-} concentrations are most important in affecting both episodic and chronic acidification (3). Similarly for Europe, SO_4^{2-} appears to be most often associated with acidic episodes, although in some sites, NO_3^-, DOC, and sea salt may also be important (1). Research in southeastern Canada has suggested the importance of base cation dilution in circumneutral waters compared to low ANC waters where SO_4^{2-} is the primary controller of acidic episodes.

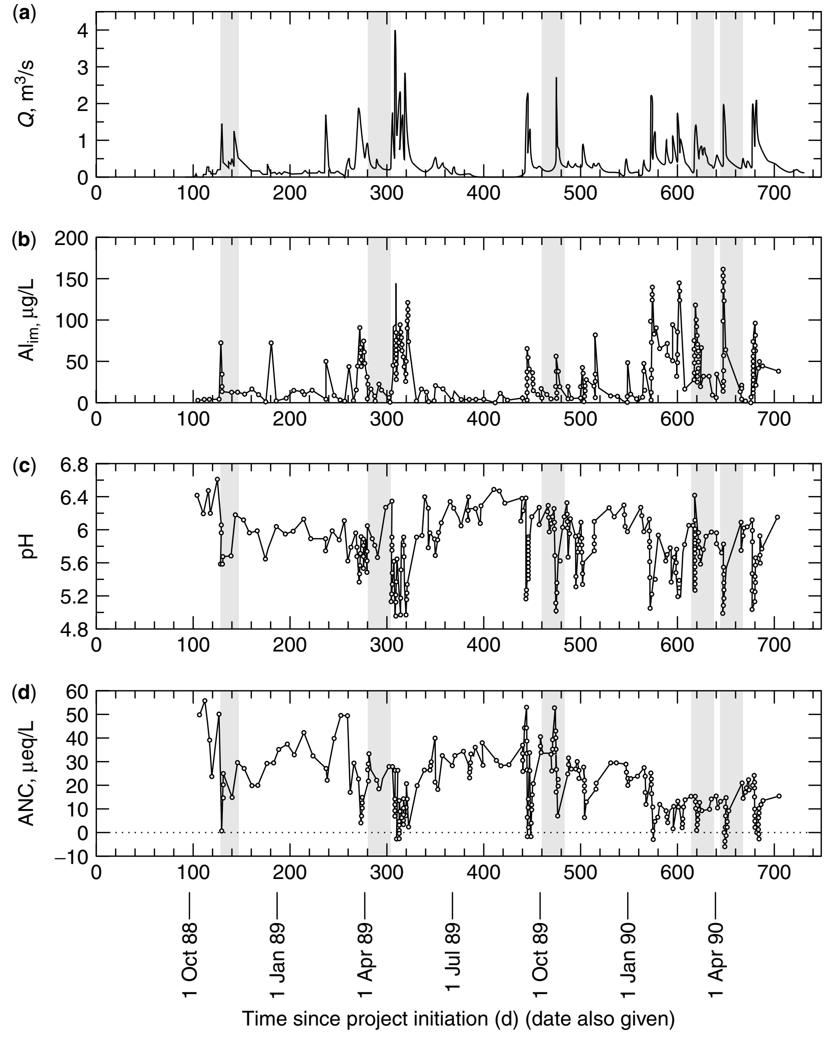

Figure 2. Effects of seasonal changes in discharge in affecting Al_{im}, pH and ANC. High discharge and low ANC coincide with spring snowmelt and fall storms. (**a**) Mean daily discharge (*Q*); (**b**) Al_{im} (inorganic monomeric aluminum); (**c**) pH; and (**d**) acid neutralizing capacity (ANC) for Biscuit Brook in Catskill Mountains, New York, U.S. Shaded areas are times when trout bioassays were conducted (3).

BIOLOGICAL IMPORTANCE

Episodic acidification associated with high concentrations of inorganic monomeric aluminum (Al_{im}), it has been shown, is deleterious to aquatic biota (8). In many regions, there is a general relationship between low pH, low ANC, and high concentrations of Al_{im}, but those surface waters that have lowest pH values do not necessarily have the highest Al_{im} values. The concentration of Al_{im}, it has been shown, is a good predictor of fish mortality. There are, however, major differences in the sensitivity of fish species to acidification, including acid pulses. Some acidic events may coincide with particularly sensitive periods of aquatic biota life cycles (e.g., spawning), so these events may have marked impacts on freshwater ecosystems. Other biotic impacts can include changes in the structure of aquatic communities and losses in biotic diversity (Fig. 2).

BIBLIOGRAPHY

1. Davies, T.D., Tranter, M., Wigington, P.J., and Eshleman, K.N. (1992). 'Acidic episodes' in surface waters in Europe. *Journal of Hydrology* **132**: 25–69.
2. Wigington, P.J., Jr. et al. (1996). Episodic acidification of small streams in the northeastern United States: episodic response project. *Ecological Applications* **62**: 374–388.
3. Wigington, P.J., Jr. et al. (1996). Episodic acidification of small streams in the northeastern United Sates: ionic controls of episodes. *Ecological Applications* **62**: 389–407.
4. Baker, J.P., Bernard, D.P., Sale, M.J., and Christensen, S.W. (1990). Biological effects of changes in surface water acid-base chemistry. NAPAP Report 13. In: *Acidic Deposition: State of the Science and Technology*. National Acid Precipitation Assessment Program, Washington, DC.
5. Wigington, P.J., Jr., Davies, T.D., Tranter, M., and Eshleman, K.N. (1992). Comparison of episodic acidification in Canada, Europe and the United States. *Environmental Pollution* **78**: 29–35.
6. Eshleman, K.N., Wigington, P.J., Jr., Davies, T.D., and Tranter, M. (1992). Modelling episodic acidification of surface waters: the state of the science. *Environmental Pollution* **77**: 287–295.
7. Davies, T.D. et al. (1999). Prediction of episodic acidification in the Northeastern USA: an empirical/mechanistic approach. *Hydrological Processes* **13**: 1181–1195.
8. Baker, J.P. et al. (1996). Episodic acidification of small streams in the Northeastern United States: effects on fish populations. *Ecological Applications* **62**: 422–437.

ACIDIFICATION OF FRESHWATER RESOURCES

Walter H. Geller
UFZ, Dept. of Inland Water Research
Magdeburg, Germany

GENERAL ASPECTS

All freshwater resources, groundwaters, lakes, and rivers, can be acidified. Inland waters receive acid inputs via the atmospheric path and by direct inflows from surface runoff and from groundwaters. Sources of acids are smoke from burning of fossil combustibles and from volcanoes, acidic wastewater from industrial plants, geogenic acids from weathering of sulfidic deposits at their natural sites, or from oxidation of sulfides with/after mining ores, hard coal, and lignite (1). Accordingly, the acid inputs into ground- and surface waters are either distributed by acid rain over long ranges and large areas, or they are restricted to local impacts by draining of mines, dumps, or sulfidic natural deposits. Both pathways are found in active volcanoes regionally releasing volatile acids into the atmosphere and being direct, local sources of highly mineralized acidic brines. Last but not least, freshwaters can be directly acidified by inflows of industrial wastewater. The effect of acidic inputs is larger in soft water because of the low content of ions and low buffering capacity.

The primary state of acidic waters is changed by secondary processes in soils and rock while seeping through the ground. Thereby, the ionic composition is modified by partial neutralization of the initial acidity and by increased mineralization at low pH. Due to the processes connected with ground passage, the waters usually contain additional dissolved substances, aluminum, iron, and other elements.

The atmospheric pathway distributes anthropogenic acids stemming from gaseous SO_2 and NO_x over large areas of land. During soil passage, this rain acidity can be neutralized in carbonate-rich areas, so that, here, ground- and surface waters are not affected. However, regions that are poor in carbonates (granitic bedrock, gneiss, sand, and sandstone) are sensitive. In a two-step process, the soils first become acidic after complete loss of carbonatic minerals; then, further acidic waters seeping through the soils dissolve further minerals, especially Al silicates.

Acids of geogenic origin can be set free by natural weathering or by weathering processes induced by mining. The first geogenic source of acidity is sulfur (metal sulfides: pyrite or other sulfidic ores) that can be oxidized after contact with air. Seeping water produces sulfuric acid and Fe (II), that later can be oxidized to Fe (III), thereby causing further acidity. The processes of sulfur and iron oxidation are accelerated by sulfur and iron oxidizing bacteria. This type of geogenic acidification is common in many mining areas where we find acidic open pit lakes, acidic streams, or acid mine drainages and tailings.

Another geogenic source of acidity is volcanic activity. Volatile acids are thermally set free as gaseous SO_2, HCl, and HF from the respective minerals. By mixing with steam of meteoric waters, strong acids are formed that dissolve volcanic rock. The resulting outflowing waters are extremely acidic and highly mineralized brines that often emerge in geothermal hot springs at the flanks and top of the volcanoes or collect in crater lakes. Because of the respective dominating buffering systems in the different acidic waters, the pH values are kept within typical ranges (2–5).

Ranges of Proton Concentration, Acidity, and Dissolved Solids

Generally, the content of dissolved minerals that are hydrolyzed during the passage of water through rock and soil increases with the acidity and proton concentration (−pH) of the resulting acidic waters. The total dissolved solids (TDS) increase, over the range from pH 4 to pH 0, from 200 mg/kg to 100 g/kg (Fig. 1).

The relationship between acidity and pH is shown in Fig. 2 for volcanic waters, for pit lakes, and for rain-acidified soft-water lakes. The log-log line is valid for strong acids completely dissociated into protons and anions $[\log_{10}(\text{acidity in eq/L}) = (-\text{pH})]$. The real values of the actual proton concentrations, however, are lower in the acidic volcanic waters where only 30% of the acidity is apparent from the pH. In mining waters of 10 to 30 meq/L acidity, only 12 to 30% of the protons that potentially can be set free are reflected by the pH values. This means that the acidification effect after mixing and dilution with fresh waters is three to eight times stronger than presumed from the actual pH values.

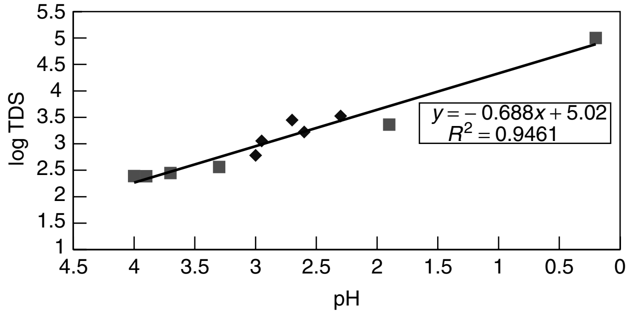

Figure 1. Relationship between the pH of acidic waters of different types and the content of total dissolved solids (TDS in ppm): diamonds: original values for pit lakes; squares: values for volcanic waters from Reference 6.

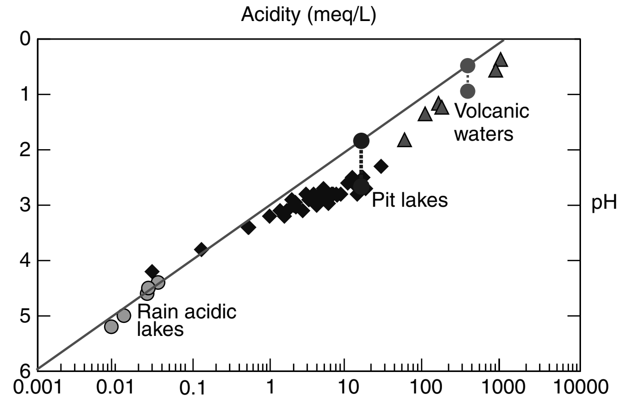

Figure 2. Acidities (measured as base neutralization capacity ($K_{B4.3}$) or calculated from an ion balance) and pH values for acidic volcanic waters (triangles: values from Reference 6), for acidic pit lakes (diamonds: original values and from Reference 7), and for rain-acidified softwater lakes (circles: from Reference 8). Vertical stippled lines show the deviation between the actual proton concentration (pH) and the acidity (or base neutralization capacity) of the waters: in pit lakes, 12 to 30%, and in volcanic lakes, 30% of the acidity is apparent as free H^+ ions.

ATMOSPHERIC ACIDIFICATION: ACID RAIN AND SENSITIVE WATERS

Acidification of Soil and Water and Chemistry of Acidified Soft Waters

Prior to anthropogenic emissions, the only atmospheric acid was carbon dioxide, resulting in weakly acidic rain of pH 5.6. This natural rainwater dissolves only carbonates from soils and weathering bedrock. Therefore, natural freshwaters contain mainly carbonates, HCO_3^-, and CO_3^{2-}, and most fresh inland waters in temperate climatic regions show nearly identical chemical compositions in which the carbonate buffering system dominates at pH 7. Carbonate-poor geological regions have "soft" waters of low mineral content, and carbonate-rich areas show higher mineralized "hard" waters. The relative (!) ionic composition of all natural freshwaters is similar, resulting in a "standard water composition" (2,9).

Only those geological regions that are poor in carbonates are sensitive to acid rain. The critical load is reached when the base saturation (percentage of Ca^{2+}, Mg^{2+}, K^+) decreases below 20% of the total cation exchange capacity (10) or below a limit of acid neutralization capacity (ANC) of 0 to 25 μeq L^{-1} (11). Estimates of the critical loads (11) showed a range between 300 and 800 mg m^{-2} of annual sulfur depositions in geologically sensitive areas (12–15).

Acid rain changes the chemical composition of seeping runoff, ground- and surface waters (8,11). Beyond the critical limits, the interaction between soil and rainwater leads to accumulation of acidity (N and S content). The carbonate and hydrogen carbonate anions in soil and surface waters are replaced by sulfate, and Ca^{2+} and Mg^{2+} cations by H^+ and Al^{3+} (16). Rain-acidified inland waters, after a large-scale process of titration, are carbonate-free, enriched in sulfate and aluminum, and acidic within a range between pH 4.5 and 5.5. In regional surveys of sensitive areas, the frequencies of pH of the respective lakes show bimodal distributions (3,17–20). Two buffering systems are involved. The nonacidified lakes are circumneutral and buffered by the carbon–acid system (CO_2–H_2CO_3–HCO_3^-–CO_3^{2-}). The rain-acidified soft waters are moderately acidic (pH 4.5–5.5) and are buffered by the aluminum system (Al^{3+}–$Al(OH)^{2+}$–$Al(OH)_2^+$–$Al(OH)_3$–$Al(OH)_4^-$) (2,11,21).

Regional Distribution of Rain-Acidified Waters

After the emissions of acid smoke, transport times of 15 to 21 hours were observed from the U.S. Midwest states, where the emissions occurred, to the Northeast states in New England and Canada, where the acid rain was precipitating. Accordingly, models use 15-h and 21-h back trajectories to refer the observed acid rain to the emission sources (22). Similar transport times and distances are known from England to southern Scandinavia. However, longer transports are also known that enable intercontinental atmospheric transport (23). The present global problem areas at risk of surface water acidification (14) are in mid- and northeast North America; in western, central, and northern Europe (24); in the eastern parts of China (25); and in different tropical countries (26–28). The long-term development and large-scale geographic distribution in comparison with the development in North America were described in a 9-year report of 11 EU countries (29). In Norway, 27% of the lakes were acidified, and 9% in both Sweden and Finland, corresponding to 6000 and 3000 lakes, respectively (30). Of a total number of 85,000 lakes (20%) in Sweden, 17,000 were assigned as affected until 1992 (31).

Biological Effects of Acidification on Physiological and Ecosystem Levels

The biological effects of rain acidification on individual groups of organisms and on the ecosystems were extensively studied (3,4,15,17–20,22,29,32–52). As acidity increases and pH decreases, the ecosystems lose major groups of organisms; fish, mollusks, cyanobacteria, and crustaceans. There is a shift to fewer species in all groups and to lower biodiversity. The simplified food web has some new dominating components, such as Dinophyceae, water bugs, sphagnum moss, or mats of filamentous green algae. A combination of stressors was identified: acidity, toxic species of aluminum, more intensive and deeper reaching UV irradiation, and loss of fish-food, such as macroinvertebrate species. In several thousand acidified lakes in southern Norway, the macroinvertebrates disappeared at a pH below 4.8. Many fish species disappeared at a pH below 5.7 in the acidic lakes of the La Cloche Mountains (NY), and no species survived at <pH 4.3 (17,20). About 50% of the populations of brown trout in southern Norway were lost in the respective area till 1975 (43,44). The losses increased with altitude in headwater lakes >1000 m above sea level to more than 80% (4,35). A survey showed till 1990, a total of 9,630 fish populations lost, and 5,405 populations were severely affected. The damaged area covered 25% of Norway (40). In a survey of benthic macroinvertebrates in neutral and acidified streams and in acidic lignite mining lakes in Germany, a maximum of about 50 species was found at pH 8, and the number of species decreased to zero at pH 2 (53,54).

Success of Countermeasures and Long-Term Developments

Countermeasures against acid rain and acidification of soils, lakes, and rivers started (1) by controlling the emissions of acidic smoke and (2) by direct measures applied to acidic waters, mainly liming. The emissions were reduced in North America and in Europe during the last two decades by technical improvements, as documented by monitoring networks (55) and by long-term observations at single sites (56,57). In Europe, the former Eastern countries followed after a one decade delay; the area of former East Germany could be identified by the atmospheric SO_2 content till 1992 (58). In the world's largest liming program in Sweden, 200,000 tons of lime were spread every year for 20 years. Of 6000 lakes treated, about half of the acidified area was restored till the early 1990s (31,59–61). The state of "Acidification of surface waters in Sweden—effects and counteracting measures" in 1993 is described in a series of 12 papers (AMBIO 1993, issue 12/5).

INDUSTRIAL ACIDIFICATION: LAKE ORTA (ITALY)

The input of industrial wastewater into fresh waters caused acidification and contamination. Between 1926 and 1982, prealpine Lake Orta (143 m deep, 1.3 km^3 volume) was heavily polluted by industrial wastewater containing ammonium sulfate and heavy metals Cu, Cr, Ni, and Zn. The development of the pollution and acidification and of the recovery are well documented (62). The annual input of copper reached 70 tons between 1950 and 1954, resulting in mean lake concentrations of more than 0.1 $mg\,Cu\,L^{-1}$. The ammonium input was 3,000 tons $NH_4{}^+$-N per year during the late 1960s. During the 1960s and 1970s, the lake water became acidic (pH 3.9 to 4.5) by oxidation of the ammonium. Fish and populations of animals disappeared nearly completely. After stopping the input of industrial wastewaters, the lake was treated from 1989–1990 by lime suspension. The lake was completely neutralized (pH 6.7–6.9) by using 18,000 tons of powdered limestone (63). To date, this was the world's biggest liming campaign applied to a single lake.

GEOGENIC ACIDIFICATION

Lakes and rivers can be acidified geogenically by volcanic acids and by natural acid tailings originating from ore and coal deposits, from surface-mining dumps, or from drainage of old underground pits. Direct acid input from mine tailings into lakes and rivers results mostly from sulfide minerals such as pyrite and related ores oxidized to sulfuric acid and iron hydroxides.

Natural Geogenic Acidification: Acidic Crater Lakes

In volcanic areas, natural pH values of acid crater lakes range from pH 0 to 5.5 (for volcanoes and their state of activities, see References: 64,65). Crater lakes of many active volcanoes contain highly acidic waters that originate from volatile mineral acids. Sulfuric acid, hydrochloric acid, and hydrofluoric acid, or their volatile forms, emerge from the hot active centers causing subsequent leaching and weathering of rock. These acid waters often collect in crater lakes with given constraints in their physical and chemical properties (66,67). For recent publications on crater lakes, see Reference 67. Several crater lakes and volcanic acid brines were chemically characterized (68–72) and showed a very broad spectrum of dissolved elements, many heavy metals, and high temperatures.

Natural Acidic Tailings from Geological Deposits

Natural acidification can originate from natural mineral sulfides weathering in soils and sediments (73). Examples of lake acidification from ore bodies or from metalliferous black shales in Japan (74) and Canada (75–77) were described. The near-surface weathering could have been initiated by glacial processes; the results are comparable with human mining activities (78). Furthermore, post-glacial isostatic land uplift can bring about long-term weathering of near-surface ores (gossans) (77). Examples of acidic hypersaline lakes in endorheic regions are located in Australia and Chile (79–81). The former existence of ancient lakes of this type was shown in the United States (73).

Acidic Rivers—Rio Tinto

The rivers Rio Tinto and Rio Odiel in southern Spain are natural acid drainages from the Iberian Pyrite Belt, the world's largest deposit of sulfidic ores. There are indications (82) that acid waters existed in the river basin 300,000 years ago (83). In addition to the natural acid drainages, human mining activities increased acidification for about 4,500 years and especially during the last 150 years. The Rio Tinto River is an extreme aquatic environment over its full length of 92 km, it has a mean pH of 2.2 and high concentrations of heavy metals (e.g., Fe 2.3 g/L, Zn 0.22 g/L, Cu 0.11 g/L) (82). The metal pollution is reflected in the estuarine and nearshore marine sediments that contain Fe (11.2%), Cu (0.93 g/kg), Zn (1.15 g/kg), Pb (0.73 g/kg), and Ba (0.66 g/kg) (84).

Acidic Mining Lakes

Lignite surface pits in east Germany (the former GDR) were largely closed after the unification of 1989. The voids that are left or will be filled within a few years will result in 200 pit lakes >1 ha. In Lusatia, the eastern lignite mining district of the former GDR, 63% of the existing pit lakes are acidic, have pH values of 2.3 to 3.5, and show high sulfate concentrations up to 2.5 g/L (85). In the United States, surface coal mining left hundreds of pit voids and lakes in the Appalachians and the Midwest during the first half of the last century (86). In addition, numerous new lakes will originate from 86 major ore mining plants (86). Thirty new pit lakes are expected in Nevada within the next 20 years. Existing acidic pit lakes are Berkeley Pit in Montana, Liberty Pit in Nevada, and Spencerville Pit in California (86). Presently, 19 metal mines are operating in Canada, 74 in Australia, 37 in Chile, 75 in Kazakhstan, and additional ones in other countries (Brazil, Peru, Mexico, Indonesia, Philippines, and in African and Asian countries) (86). Many of the voids left will become future lakes that often will be acidic and contaminated by toxic metals.

The present state of the literature does not permit one to work out satisfactorily the common chemical characteristics and a typology of the mining lakes of the different geological regions and climatic zones. A basic classification was found in regional surveys of lignite pit lakes, where multimodal frequency distributions of pH values similar to those of rain-acidified lakes were observed. The three modal groups of pit lakes are based on three buffering systems: bi-/carbonate (pH 6 to 8), $Al(OH)_x$ (pH 4.5 to 5.5), and $Fe(OH)_x$ [(pH 2 to 4) (87), and citations in (86)].

FRESH WATERS AFFECTED BY ACIDIC INFLOWS

Nonacidified waters, rivers, lakes, and groundwater are affected by mixing with acidic inflows such as acidic rivers from volcanic sources' acid drainages and tailings from natural deposits and from surface and underground mines, or acidic sludges that are accidentally set free by spills (89). In acidic drainages, heavy metals are often found as cocontaminants. Although the compounds are diluted, neutralized, and minerals are precipitated, acids and metals damage the freshwater biota (88). Unexpectedly diverse communities of extremophiles can develop in permanently acidic and metal-rich waters (82,83).

Acidification of Groundwaters by Acid Rain and Mining

The importance of the acidification of groundwater tables is related to the spatial scale of the affected aquifers. The effect of acid rain on springs and brooks in Germany was observed by a monitoring network whose areal distribution was screened by a 3-km grid (90). The results showed acidification of all near-surface groundwaters in regions sensitive because of soils poor in carbonates. The scale of lignite mining in the east German district of Lusatia also generally affects regional groundwater (91).

Acidification of Rivers and Lakes by Volcanic Water

Active volcanoes set free volatile acids (SO_2, HCl, HF) and form acidic brines. Outflows of the resulting acidic crater lakes and volcanic flank springs lead to highly acidic headwaters in catchment areas. Freshwater streams that receive the acidic and toxic volcanic inputs of the headwaters are heavily affected (65). Two examples from Java show the consequences of this sort of pollution. The crater lake of the volcano Kawah Ijen, East Java, contains 32×10^6 m^3 of hot and acidic brine (pH <0.4, TDS > 100 g/kg, SO_4^{2-} 70 g/kg, Cl^- 21 g/kg, F^- 1.5 g/kg). The brine seeps through the crater rim and mixes with the Banyupahit River, where the acidic inflow is only incompletely neutralized by dilution and results in the lack of biota. About 30 km downstream, rice fields are irrigated by the river, discharging 4,000 L/s; the daily load is 150 t SO_4, 2.8 t F, 50 t Cl, 10 t Al, 35 kg Ti, and 4 kg Cu (92). The age of the crater lake and its environmental impacts are estimated at more than 200 years old. Furthermore, the Patuha volcano in West Java has an acidic crater lake, Kawah Putih, and springs of acid brines of pH < 1 containing Al, As, B, Cl, Fe, Mn, and SO_4. The acid streamlets drain into the Citarum River that is contaminated by these potentially toxic elements from the crater lake and, too, is used for irrigation (93). In Argentina at Copahue Volcano, there is a network of acidic waters starting from an acid crater lake, Lago Copahue, and a 13 km-long, extremely acidic river, Upper Rio Agrio (pH 0.6 to 1.6). The river discharges into the glacial Lake Caviahue, where the highly acidic water is diluted to pH 2.5 (68,94). The outflow of the lake, Lower Rio Agrio, is further diluted in its course, but, presumably, damages downstream freshwater stretches by its acidic inputs.

Acidification of Surface Waters by Mine Drainage

In the United States, Kleinmann (1989) reported that 19,300 km of streams are affected by acid mine drainage (AMD). In Europe, a typical case is the Avoca River in southeastern Ireland which is affected by continuous AMD from an abandoned copper and sulfur mine area (88). The AMD of pH 2.7 discharges annually about 300 tons of metals into the river (108 t Zn, 276 t Fe, 6 t Cu, 0.3 t Cd).

In the contaminated river, whose pH is 5.8, macrophytes and fish are eliminated, and macroinvertebrates survive only for short periods outside the mixing zone. The damage to the indigenous biota is due to a combination of metal toxicity, sedimentation, acidity, and salinization (88).

Remediation of AMD and Affected Waters

The acidification of lakes and rivers by acid drainage can be prevented by preclusive measures and direct treatment of the polluting waters as (1) restrictions on running mining activities; (2) "active treatment" of the AMD by addition of alkaline chemicals, often in on-site reactor systems, neutralizing the waters; (3) "passive treatment" using natural geochemical and biological reactions to reduce acidity, pH, and ion concentrations; and (4) insitu treatment of acidic lakes.

Gray (95) developed a scheme to describe (1) the damage of AMD on affected river systems for the site of the Avoca tin mine in Ireland and (2) a management system to avoid damage by a combined application of prevention and remediation by passive treatment methods (88). Gazea et al. (96) published a comprehensive review on treating AMD by using passive systems, such as natural geochemical and biological processes to improve water quality with minimal requirements. The remediation of small-scale acid drainages was reached by sequential flow (1) through anoxic limestone drains for pretreating acidic waters by providing alkalinity, (2) through aerobic wetlands to remove iron through oxidation, and (3) through a series of anoxic cells in compost wetlands to raise the pH and to reduce concentrations of other metals and sulfate (96–100).

In extremely acidic pit lakes, the acidity and content of dissolved substances is one to three orders of magnitude higher than in rain-acidified waters. So, liming is far more expensive, and biological treatment becomes a promising alternative for neutralizing the acidity. The microbial processes (101) used for remediating lakes by insitu methods are summarized by Klapper et al. (97). Sulfate-reducing bacteria can use sulfate as an electron acceptor to reduce both iron and sulfur to form pyrite (102):

$$30\mathrm{CH_2O} + 8\mathrm{Fe(OH)_3} + 16\mathrm{SO_4}^{2-} + 32\mathrm{H}^+$$
$$\Rightarrow 30\mathrm{CO_2} + 58\mathrm{H_2O} + 8\mathrm{FeS_2}$$

Analogous to the process in anoxic cells of passive AMD treatments, the in situ reversion of pyrite oxidation to neutralize the acid water in mining lakes can be promoted by adding organic substances that lead to anoxic zones in and above the bottom sediments or in the hypolimnic waters. A pilot project using in situ remediation techniques on whole lakes was actually run (103,104).

DATA SOURCES

Data and actual reports on inland waters and relevant environmental issues are accessible for different regions and on a global scale from Internet sources (Table 1).

Table 1. Internet Access to Data, Reports, and Publications on Inland Waters of the World

Canada Centre of Inland Waters (CCIW)	http://www.cciw.ca/gems
European Environment Agency (EEA)	http://themes.eea.eu.int
European Union (EU)	http://europa.eu.int
Food and Agricultural Organisation of the United Nations (FAO)	http://www.apps.fao.org
Global Runoff Data Centre	http://www.bafg.de/grdc.htm
Global Water Partnership	http://www.gwpforum.org
Organisation for Economic Cooperation and Development (OECD)	http://www.oecd.org
World Hydrological Cycle Observing System	http://www.wmo.ch
U.S. Environmental Protection Agency (EPA)	http://www.epa.gov

BIBLIOGRAPHY

1. Nordstrom, D.K. and Alpers, C.N. (1998). The environmental geochemistry of mineral deposits. In: *Reviews in Economic Geology*. Vol. 7A. G.S. Plumlee and M.J. Logsdon (Eds.). p. 28.
2. Henriksen, A. (1980). In: *Ecological Impact of Acid Precipitation*. D. Drablos and A. Tollan (Eds.). Oslo-As. SNSF-Report, pp. 68–74.
3. National Research Council Canada. (1981). *Acidification in the Canadian Aquatic Environment: Scientific Criteria for Assessing the Effects of Acidic Deposition on Aquatic Ecosystems*. 18475. NRCC-Report. Ottawa, p. 369.
4. Overrein, L.N., Seip, H.M., and Tollan, A.T. (1981). *Acid Precipitation—Effects on Forests and Fish*. Oslo-As. SNSF, p. 175.
5. Geller, W., Klapper, H., and Schultze, M. (1998). In: *Acidic Mining Lakes*. W. Geller, H. Klapper, and W. Salomons (Eds.). Springer, Berlin, pp. 3–14.
6. Kempter, K.A. and Rowe, G.L. (2000). *J. Vulcanology and Geothermal Research* **97**: 143–159.
7. Nixdorf, B. et al. (2000). *Braunkohlentagebauseen in Deutschland Umweltforschungsplan BMU*. Umweltbundesamt, Berlin, p. 519.
8. Stumm, W., Morgan, J.J., and Schnoor, J.L. (1983). *Naturwissenschaften* 216–223.
9. Hem, J.D. (1992). *Study and Interpretation of the Chemical Characteristics of Natural Water*. U.S. Geological Survey Water-Supply Paper 2254, U.S. Geological Survey, Washington, DC, p. 264.
10. Driscoll, C.T., Baker, J.P., Bisogni, J.J., and Schofield, C.L. (1990). *Nature* **284**: 161–164.
11. Stumm, W. and Schnoor, J.L. (1995). In: *Physics and Chemistry of Lakes*, 2nd Edn. A. Lerman, D. Imboden, and J. Gat (Eds.). Springer, Berlin, pp. 185–215.
12. Brodin, Y-W. and Kuylenstierna, J.C. (1992). *Ambio* **21**: 332–338.
13. Lövblad, G. et al. (1992). *Ambio* **21**: 339–347.
14. Kämäri, J. et al. (1992). *Ambio* **21**: 377–387.
15. Stanners, D. and Bourdeau, P. (2000). *Europe's Environment: The Dobris Assessment*. Available: http://themes.eea.eu/showpage.php/state/water.

16. Cronan, C.S. and Schofield, C.L. (1990). *Environment. Sci. Technol.* **24**: 1100–1105.
17. Harvey, H.H. (1975). *Verh. Internat. Verein. Limnol.* **19**: 2406–2417.
18. Schofield, C.L. (1976). *Ambio* **5**: 228–230.
19. Wright, R.F. et al. (1977). *Regional Surveys of Small Norwegian Lakes*. IR 33/77. SNSF Project. SNSF, Oslo, p. 153.
20. Harvey, H.H. (1980). In: *Ecological Impact of Acid Precipitation*. D. Drablos and A. Tollan (Eds.). SNSF-proceedings. Oslo-As, pp. 93–98.
21. Dickson, W. (1980). In: *Ecological Impact of Acid Precipitation*. D. Drablos and A. Tollan (Eds.). Proc., Int. conf. ecol. impact acid precip. Oslo-As. SNSF-Report, pp. 75–83.
22. Driscoll, C.T. et al. (2001). *Bioscience* **51**: 180–198.
23. Benkovitz, C.M., Miller, M.A., Schwartz, S.E., and Kwon, O-U. (2001). *Geochem. Geophys. Geosyst.* **2**: (June).
24. Kristensen, P. and Hansen, H.O. (1994). *European Rivers and Lakes*. EEA Environmental Monographs; 1. European Environmental Agency, Copenhagen, p. 121.
25. Streets, D.G., Carmichael, G.R., Amann, M., and Arndt, R.L. (1999). *Ambio* **28**: 135–143.
26. Rhode, H. and Herrera, R. (1988). *Acidification in Tropical Countries*. John Wiley, Chichester, p. 405.
27. Rhode, H. (1989). *Ambio* **18**: 155–159.
28. GEMS/WATER. (2002). *The Annotated Digital Atlas of Global Water Quality*. Available: http://www.cciw.ca/gems/atlas-gwq/intro-e.html.
29. Luckewille, A. et al. (1997). *NIVA Report* serial no. 3637–97, p. 168.
30. Henriksen, A. et al. (1997). *NIVA Report* serial no. 3645–97. p. 41.
31. Leonard, J. and Crouzet, P. (1999). *Acidification*. 1/1999. Topic Report. Copenhagen, p. 85. <http://europa.eu.int>.
32. Steinberg, C.E.W. and Wright, R.F. (1994). *Acidification of Freshwater Ecosystems*. Environmental Sci. Res. Rep. ES 14. John Wiley, Chichester, p. 404.
33. Schindler, D.W. (1994). In: *Acidification of Freshwater Ecosystems*. C.E.W. Steinberg and R.F. Wright (Eds.). John Wiley, Chichester, pp. 153–164.
34. Arts, G.H. (2002). *Aquatic Botany* **73**(4): 373–393.
35. Drablos, D. and Tollan, A. (1980). *Ecological Impact of Acid Precipitation*. SNSF-proceedings. Oslo-As. p. 382.
36. Elwood, J. and Mulholland, P. (1989). In: *Acid Precipitation*. D. Adriano and A. Johnson (Eds.). Springer, Heidelberg, pp. 85–135.
37. Findlay, D.L. and Kasian, S.E.M. (1990). *Can. J. Fisheries Aquatic Sciences* **47**: 1378–1386.
38. Gorham, E. (1976). *Water, Air, and Soil Pollution* **6**: 457–481.
39. Haines, T.A. (1981). *Transactions American Fisheries Society* **110**: 669–707.
40. Hesthagen, T., Sevaldrud, I.H., and Berger, H.M. (1999). *Ambio* **28**: 112–117.
41. McKnight, D. and Feder, G. (1984). *Hydrobiologia* **119**: 129–138.
42. Minns, C.K., Moore, J.E., Schindler, D.W., and Jones, M.L. (1990). *Can. J. Fish. Aquat. Sci.* **47**: 821–830.
43. Mossberg, P. and Nyberg, P. (1979). Bottom fauna of small acid forest lakes. *Rep. Inst. Freshw. Res. Drottningholm* **58**: 77–87.
44. Muniz, I.P. and Leivestad, H. (1980). In: *Ecological Impact of Acid Precipitation*. D. Drablos and A. Tollan (Eds.). Oslo-As. SNSF-Report, pp. 84–92.
45. Murphy, K.J. (2002). *Aquatic Botany* **40**: in press.
46. Niyogi, D., McKnight, D., and Lewis, W. (1999). *Limnol. Oceanogr.* **44**: 804–809.
47. Rahel, F.J. and Magnuson, J.J. (1983). *Can. J. Fisheries Aquatic Sciences* **40**: 3–9.
48. Schindler, D.W. and Turner, M.A. (1982). *Water, Air, and Soil Pollution* **18**: 259–272.
49. Schindler, D.W. (1988). *Science* **239**: 149–157.
50. Schindler, D.W., Kasian, S.E.M., and Hesslein, R.H. (1989). *Environmental Monitoring and Assessment* **12**: 269–285.
51. Tipping, E. et al. (2002). *Environmental Pollution* **116**: 137–146.
52. Vinebrooke, R.D. et al. (2002). *Can. J. Fish. Aquat. Sci.* **59**: 483–493.
53. Braukmann, U. (1995). *Saurer Regen: Probleme für Wasser, Boden und Organismen*. Umweltministerium Baden-Württemberg (Eds.). Lech. Ecomed, pp. 205–224.
54. Rodrigues, G.G. (2001). *Benthic fauna of extremely acidic lakes (pH 2–3)*. Diss. TU Braunschweig, p. 131.
55. NADP. National Atmospheric Deposition Program. (2002). *Hydrogen Ion Concentration as pH from Measurements made at the Field Laboratories*. Available: http://nadp.sws.uiuc.edu.
56. Likens, G.E., Driscoll, C.T., and Buso, D.C. (1996). *Science* **272**: 244–246.
57. Likens, G.E., Butler, T.J., and Buso, D.C. (2001). *Biogeochemistry* **52**: 1–11.
58. UBA. (2000). *Daten zur Umwelt—Ausgabe 2000 Daten zur Umwelt*. E. Schmidt, Berlin.
59. Henrikson, L. and Brodin, Y-W. (1995). In: *Liming of Acidified Surface Waters*. L. Henrikson and Y-W. Brodin (Eds.). Springer, Berlin, pp. 1–44.
60. Olem, H. (1991). *Liming Acidic Surface Waters*. Levis Publ., Chelsea, p. 331.
61. Erikson, F., Hörnström, E., Mossberg, P., and Nyberg, P. (1983). *Hydrobiologia* **101**: 145–164.
62. Bonacina, C. (2001). *J. Limnology* **60**: 289–300.
63. Calderoni, A. and Tartari, G.A. (2001). *J. Limnology* **60**: 69–78.
64. Smithsonian Institution. (2002). *Global Volcanism Program. Volcan Activity Report*. Available: http://www.nmnh.si.edu/gvp/volcano.
65. Jensen, B. (2002). *World List of Volcanoes*. Avialable: http://users.bendnet.com/bjensen/volcano.
66. Pasternack, G.B. and Varekamp, J.C. (1997). *Bull. Volcanology* **58**: 528–538.
67. Varekamp, J.C., Pasternack, G.B., and Rowe, G.L., Jr. (2000). *J. Volcanology and Geothermal Research* **97**: 161–179.
68. Varekamp, J.C. et al. (2001). *Geology* **29**: 1059–1062.
69. Armienta, M.A., de la Cruz-Reyna, S., and Macias, J.L. (2000). *J. Vulcanology and Geothermal Research* **97**: 105–125.
70. Delmelle, P. et al. (2000). *J. Vulcanology and Geothermal Research* **97**: 31–53.
71. Martinez, M. et al. (2000). *J. Vulcanology and Geothermal Research* **97**: 127–141.

72. Sriwana, T. et al. (2000). *J. Vulcanology and Geothermal Research* **97**: 77–104.
73. Benison, C.K. and Goldstein, R.H. (2002). *Sedimentary Geology* **151**: 177–185.
74. Childs, C.W., Inoue, K., and Mizota, C. (1998). *Chemical Geology* **144**: 81–86.
75. Havas, M. and Hutchinson, T.C. (1983). *Nature* **301**: 23–27.
76. Kwong, Y.T. and Lawrence, J.R. (1998). In: *Acidic Mining Lakes*. W. Geller, H. Klapper, and W. Salomons (Eds.). Springer, Berlin, pp. 65–86.
77. Cameron, E.M. et al. (1998). *J. Geochemical Exploration* **64**: 197–213.
78. Loukola-Ruskeeniemi, K., Uutela, A., Tenhola, M., and Paukola, T. (1998). *J. Geochemical Exploration* **64**: 395–407.
79. Risacher, F., Alonso, H., and Salazar, C. (2002). *Chemical Geology* **187**: 39–57.
80. Alpers, C.N. et al. (1992). *Chemical Geology* **96**: 203–206.
81. Kittrick, J.A., Fanning, D.S., and Hossner, L.R. (1982). *Acid Sulfate Weathering*. Vol. 10, Soil Sciences Society of America Special Publication, Madison, WI.
82. López-Archilla, A.I., Marin, I., and Amils, R. (2000). *Microb. Ecol.* **41**: 20–35.
83. Amaral-Zettler, L.A. et al. (2002). *Nature* **417**: 137–137.
84. Borrego, J., Morales, J.A., de la Torre, M.L., and Grande, J.A. (2002). *Environmental Geology* **41**: 785–796.
85. Pietsch, W. (1998). In: *Acidic Mining Lakes*. W. Geller, H. Klapper, and W. Salomons (Eds.). Springer, Berlin, pp. 169–193.
86. Castro, J.M. and Moore, J.N. (2000). *Environmental Geology* **39**: 1254–1260.
87. Geller, W. et al. (2000). *Verh. Internat. Verein. Limnol.* **27**: 2475–2479.
88. Gray, N.F. (1998). *Water Research* **32**: 2122–2134.
89. van Geen, A., Takesue, R., and Chase, Z. (1999). *Sci. Total Environment* **242**: 221–229.
90. UBA. (1986). *Daten zur Umwelt 1986/87* Daten zur Umwelt. E.Schmidt, Berlin.
91. Grünewald, U. (2001). *Ecological Engineering* **17**: 143–152.
92. Delmelle, P. and Bernard, A. (2000). *J. Vulcanology and Geothermal Research* **97**: 55–75.
93. Sriwana, T. et al. (1998). *J. Geochemical Exploration* **62**: 161–182.
94. Pedrozo, F. et al. (2001). *Hydrobiologia* 129–137.
95. Gray, N.F. (1997). *Environmental Geology* **30**: 62–71.
96. Gazea, B., Adam, K., and Kontopoulos, A. (1996). *Minerals Engineering* **9**: 23–42.
97. Klapper, H., Geller, W., and Schultze, M. (1996). *Lakes & Reservoirs: Research and Management* **2**: 7–16.
98. Klapper, H. et al. (1998). In: *Acidic Mining Lakes*. W. Geller, H. Klapper, and W. Salomons (Eds.). Springer, Berlin, pp. 401–416.
99. Klapper, H. and Geller, W. (2001). *Acta Hydrochim. Hydrobiol.* **29**: 363–374.
100. Wendt-Potthoff, K. and Neu, T.R. (1998). In: *Acidic Mining Lakes*. W. Geller, H. Klapper, and W. Salomons (Eds.). Springer, Berlin, pp. 269–284.
101. Robbins, E.I. (1998). In: *Annual Meeting St. Louis 1998*. American Society for Surface Mining and Reclamation, pp. 174–191.
102. Peiffer, S. (2002). *Umweltgeochemische Bedeutung der Bildung und Oxidation von Pyrit in Gewässersedimenten*. Bayreuther Forum Ökologie, Bd. 47. Bayreuth, p. 103.
103. Frömmichen, R., Koschorreck, M., Wendt-Potthoff, K., and Friese, K. (2001). In: *Bioremediation of Inorganic Compounds 6(9)*. A. Leeson, B. Peyton, J. Means, and V.S. Magar (Eds.). Battelle Press, Columbus, pp. 43–51.
104. Wendt-Potthoff, K., Frömmichen, R., Herzsprung, P., and Koschorreck, M. (2002). *Water, Air, and Soil Pollution* **2**(3): 81–96.

GEOCHEMISTRY OF ACID MINE DRAINAGE

Roger C. Viadero
West Virginia University
Morgantown, West Virginia

Acid mine drainage (AMD) results when metal pyrites come in contact with water and/or air, to form dilute sulfuric acid. For instance, iron pyrite (FeS_2) is the major iron-sulfur impurity found in mined earth, particularly in the eastern United States where coal mining is prevalent. Typically, AMD waters contain elevated concentrations of SO_4, Fe, Mn, Al, and other metal ions. As an example, representative AMD water quality parameters from the Roaring Creek-Grassy Run Watershed, located in Elkins, West Virginia, include pH from 2.4 to 3.3, mineral acidities from 2.4 to 980 mg/L as $CaCO_3$, dissolved iron between 35 and 260 mg/L, and sulfate concentrations from 190 to 740 mg/L.

In contrast, copper and arsenic sulfide compounds associated with "hard rock" mining operations are common in the western United States; such compounds are generally much less prevalent in the east. In such instances, sulfide containing minerals such as pyrrhotite (FeS), arsenopyrite (FeAsS), and chalcopyrite ($CuFeS_2$) can produce acidic drainage when oxidized.

Stochiometrically, the interaction of iron pyrite with water is described by the following reactions:

$$4FeS_2(s) + 14O_2 + 4\,H_2O \longleftrightarrow 4Fe^{2+} + 8\,H^+ + 8SO_4^{2-} \tag{1}$$

$$4Fe^{2+} + 4\,H^+ + O_2 \longleftrightarrow 4Fe^{3+} + 2\,H_2O \tag{2}$$

$$4Fe^{3+} + 12\,H_2O \longleftrightarrow 4Fe(OH)_3(\downarrow) + 12\,H^+ \tag{3}$$

In this process, iron pyrite is oxidized to produce ferrous iron (Fe^{2+}), protons and sulfate ions. Ferrous iron is subsequently oxidized to ferric iron (Fe^{3+}) and precipitates from solution as $Fe(OH)_3(\downarrow)$, the sulfate ions can combine with protons to produce sulfuric acid. The iron solids, which form the characteristic red-yellow/yellow-brown coating on stream beds, can consist of a variety of ferric oxides, hydroxides, or oxyhydroxysulfates, including ferrihydrite [$Fe_3(OH)_4$, $Fe_5HO_8 4H_2O$, or $Fe_5O_3(OH)_9$] and schwertmannite [$Fe_8O_8(OH)_6SO_4$] (1–5). Further, ferrihydrite and schwertmannite are metastable and may eventually dehydrate and recrystallize to form hematite (Fe_2O_3).

Many factors determine the rate of AMD generation, including the presence and activity of bacteria, pH, temperature, and pyrite surface characteristics. The interactions of these factors are complex and readers are referred to the following for additional information on AMD geochemistry: Temple and Koehler (6), Singer and Stumm (7), Kleinmann et al. (8), Nordstrom (9), Hornberger et al. (10), Alpers et al. (11), and Evangelou (12). Further, approaches to managing/remediating various sources of AMD are largely functions of those parameters that govern the rate of AMD generation. Representative sources of information on the prevention and treatment of AMD include Sobek et al. (13), Faulkner and Skousen (14), Skousen and Ziemkiewicz (15), Caruccio et al. (16), and Meek (17).

Bacteria and pH

In many cases, the presence and activity of bacteria control the rate at which AMD forms. For example, in addition to the series of reactions presented previously, iron pyrite can interact with ferric iron according to the following relationship, in which additional acid and dissolved iron ions form:

$$\mathrm{FeS_2(s)} + 14\mathrm{Fe^{3+}} + 8\,\mathrm{H_2O} \xleftrightarrow{k} 15\mathrm{Fe^{2+}} + 2\mathrm{SO_4}^{2-} + 16\,\mathrm{H^+} \tag{4}$$

The rate law expression for the oxygenation of ferrous iron is given as

$$-\frac{d[\mathrm{Fe^{2+}}]}{dt} = k[\mathrm{Fe^{2+}}]p_{\mathrm{Ox}} \tag{5}$$

where p_{Ox} is the partial pressure of oxygen and k is the reaction rate constant. The reaction rate constant, k, is $\simeq 10^{-25}$ L/atm·min at 25 °C at pH < 5.5. Thus, iron oxidation is predicted to be very slow at pH $<$ 5.5. However, it is known through field and laboratory studies that pyrite is rapidly oxidized to $\mathrm{Fe^{3+}}$ at pH values of 2 to 3. In this case, the microorganisms *Thiobacillus thiooxidans, Thiobacillus ferrooxidans,* and *Ferrobacillus ferrooxidans* catalyze the oxygenation of ferrous iron (18). In particular, *Thiobacillus ferrooxidans* catalyzes or mediates the following oxidation–reduction reactions of $\mathrm{Fe^{2+}}$:

$$\mathrm{Fe^{2+}} \longleftrightarrow \mathrm{Fe^{3+}} + \mathrm{e^-} \tag{6}$$

$$4\,\mathrm{H_2O} + \mathrm{S^{2-}} \longleftrightarrow \mathrm{SO_4}^{2-} + 8\,\mathrm{H^+} + 8\mathrm{e^-} \tag{7}$$

Similarly, *Ferrobacillus ferrooxidans* mediates the $\mathrm{Fe^{2+}}$–$\mathrm{Fe^{3+}}$ reaction presented in Eq. 6.

To develop an understanding of the effects of pH on AMD formation, it is necessary to consider the solubility of $\mathrm{Fe^{3+}}$ hydroxide species in water. At equilibrium, the relationship between the $\mathrm{Fe(OH)_3}$ solid, soluble $\mathrm{Fe^{3+}}$ species [presented as—log(soluble iron species)] and pH is given in Fig. 1.

Kleinmann et al. (8) and Nordstrom (9) presented a three-step process of chemical and biologically mediated reactions through which AMD is generated. Note that this sequence is based on processes in the absence of appreciable amounts of alkaline material, in unsaturated systems, where there is an adequate oxygen supply. Thus, applying this rationale to systems that have substantially different characteristics must be done with caution. Stage I occurs at circumneutral pHs, where pyrite is oxidized chemically, as presented in Eqs. 1 and 2. Under such conditions, the $\mathrm{Fe^{3+}}$ concentration is limited by the low solubility of $\mathrm{Fe(OH)_3}$.

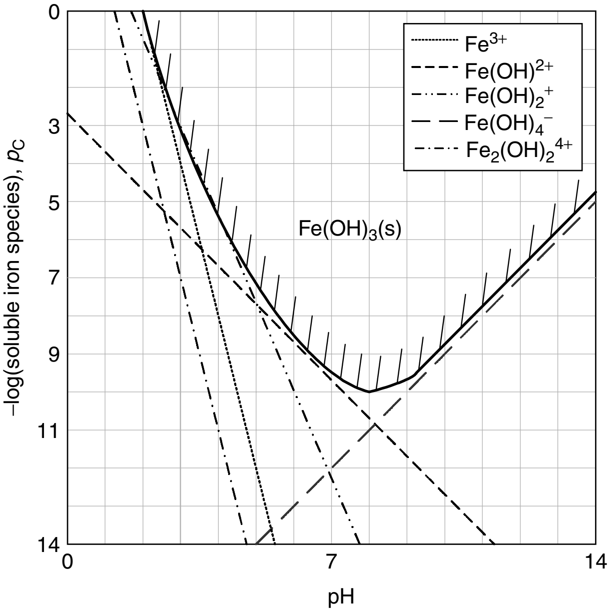

Figure 1. Log concentration versus pH for soluble $\mathrm{Fe^{3+}}$ species in the presence of $\mathrm{Fe(OH)_3(s)}$.

However, some limited oxidation of pyrite by bacteria may occur on the surface of pyrite grains. In Stage II, which occurs at pHs from 3.0 to ~4.5, $\mathrm{Fe^{2+}}$ chemical oxidation is slow; consequently, oxidation is mainly bacterially mediated. In Stage III, the relationship presented previously in Eq. 4 is principally responsible for acid production, which occurs much more rapidly than that in Stages I or II. Typically, Stage III occurs at pH $<$ 3 as $\mathrm{Fe^{3+}}$ becomes increasingly soluble, and the rate of pyrite oxidation can increase rapidly, as presented earlier. However, for pH $<$ 1.5 to 2, the effectiveness of *Thiobacillus ferrooxidans* as a catalyst for the oxidation of $\mathrm{Fe^{2+}}$ decreases (19,20).

One exception to the low pH conditions characteristic of most mine waters includes the case where carbonaceous minerals such as calcite and dolomite impart alkalinity to the water that equals or exceeds the acidity. Dissolution of carbonaceous minerals produces alkalinity, which neutralizes acidity and promotes the removal of Fe, Al, and other metal ions from solution. The result is drainage that is either neutral or net alkaline, though such waters can still have elevated concentrations of $\mathrm{SO_4}$, Fe, Mn, and other AMD-associated metal ions and may require further management.

PROPERTIES OF PYRITE

The rate of acid production depends on the surface area of pyrite exposed to a water solution or air; pyrites with higher surface areas produce more acid than pyrites with low surface areas (21). Similarly, rocks with a high percentage of pyrite produce acidity more rapidly than rocks with a low percentage of pyrite. Further, conclusions regarding the influence of surface defects, such as pitting, on pyrite oxidation rates have been studied by McKibben and Barnes (21). Other investigators have studied the crystal structure of various pyrite samples to explain differences observed in the rates of oxidation among multiple samples. For instance, Kitakaze et al. (22) observed vacant positions in the crystal lattice of pyrite and related variations in such physical properties to differences in oxidation rates. However, further studies are needed to elucidate the specific effect(s) that lattice structure, atomic spacing, and so on, have on pyrite oxidation rates.

BIBLIOGRAPHY

1. Chukhrov, F., Zvyagin, B., Ermilova, L., and Gorshkov, A. (1973). New data on iron oxides in the weathering zone. *Proc. Int. Clay Conf.* **1**: 397–404.
2. Eggleton, R. and Fitzpatrick, R. (1988). New data and a revised structural model for ferrihydrite. *Clays Clay Miner.* **36**: 111–124.
3. Brady, K.S., Bigham, J., Jaynes, W., and Logan, T. (1986). Influence of sulfate on Fe-oxide formation: Comparison with a stream receiving acid mine drainage. *Clays Clay Miner.* **34**: 266–274.
4. Murad, E., Schwertmann, U., Bigham, J., and Carlson, L. (1994). Mineralogy and characteristics of poorly crystallized precipitates formed by oxidation of Fe in acid sulfate waters. In: *Environmental Geochemistry of Sulfide Oxidation*. C. Alpers and D. Blowes (Eds.). American Chemical Society Symposium Series 550, pp. 190–200.
5. Bigham, J. et al. (1996). Schwertmannite and the chemical modeling of iron in acid sulfate waters. *Geochim. Cosmochim. Acta* **60**: 2111–2121.
6. Temple, K. and Koehler, W. (1954). *Drainage from Bituminous Coal Mines*. West Virginia University Bulletin, Engineering Experiment Station Research Bulletin 25, p. 35.
7. Singer, P. and Stumm, W. (1970). Acidic mine drainage—the rate-determining step. *Science* **167**: 1121–1123.
8. Kleinmann, R., Crerar, D., and Pacelli, R. (1981). Biogeochemistry of acid mine drainage and a method to control acid formation. *Min. Eng.* **33**: 300–303.
9. Nordstrom, D. (1982). Aqueous pyrite oxidation and the consequent formation of secondary iron minerals. In: *Acid Sulfate Weathering*. J. Kittrick, D. Fanning, and L. Hossner (Eds.). Soil Science Society of America, Madison, WI, pp. 37–63.
10. Hornberger, R., Smith, M., Friedrich, A., and Lovell, H. (1990). Acid mine drainage from active and abandoned coal mines in Pennsylvania. In: *Water Resources in Pennsylvania–Availability, Quality, and Management*. S. Majumdar, E. Miller, and R. Parizek (Eds.). The Pennsylvania Academy of Science, Harrisburg, PA, pp. 432–451.
11. Alpers, C., Blowes, D., Nordstrom, D., and Jambor, J. (1994). Secondary minerals and acid mine-water chemistry. In: *Environmental Geochemistry of Sulfide Mine-Wastes*. D. Blowes and J. Jambor (Eds.). Short Course Handbook, Mineralogical Association of Canada, **22**: 247–270.
12. Evangelou, V. (1995). *Pyrite Oxidation and its Control*. CRC Press, Boca Raton, FL.
13. Sobek, A., Schuller, W., Freeman, J., and Smith, R. (1978). *Field and Laboratory Methods Applicable to Overburdens and Minesoils*. EPA-600/2-78-054. U.S. Environmental Protection Agency, Cincinnati, OH.
14. Faulkner, B. and Skousen, J. (1994). Treatment of acid mine drainage by passive treatment systems. *Int. Land Reclamation Mine Drainage Conf.*, U.S. Bureau of Mines SP 06A-94, April 24–29, Pittsburgh, PA, p. 250.
15. Skousen, J. and Ziemkiewicz, P. (1996). *Acid Mine Drainage Control and Treatment*, 2nd Edn. West Virginia University, Morgantown, WV.
16. Caruccio, F., Geidel, G, and Sewell, J. (1976). The character of drainage as a function of the occurance of famboidal pyrite and ground water quality in eastern Kentucky, *6th Symposium on Coal Mine Drainage Research*, National Coal Association, Lousiville, KY, pp. 1–16.
17. Meek, A. (1994). Evaluation of acid prevention techniques used in surface mining, In: *International Land Reclamation and Mine Drainage Conference*. April 24–29, US Department of Interior, Bureau of Mines, Pittsburgh, PA, pp. 41–48.
18. Snoeyink, V. and Jenkins, D. (1980). *Water Chemistry*. John Wiley & Sons, New York.
19. Silverman, M.P. and Lundgren, D. (1959). Studies of the chemautotrophic bacterium *Ferrobacillus ferrooxidans*. II. Manometric studies. *J. Bacteriol.* **78**: 326–331.
20. Schnaitman, C., Korczynski, M., and Lundgren, D. (1969). “Kinetic studies of iron oxidation by whole cells of *Ferrobacillus ferrooxidans*. *J. Bacteriol.* **99**: 552–557.
21. McKibben, M. and Barnes, H. (1986). Oxidation of pyrite in low temperature acidic solutions—rate laws and surface textures. *Geochim. Cosmochim. Acta* **50**: 1509–1520.
22. Kitakaze, A., Graham, U., and Ohmoto, H. (1990). Anisotropy, crystal structure, density, and S/Fe ratios of hydrothermal pyrites. *Geol. Soc. Am. Abstr.* **22** (7): A362.

THE ARAL SEA DISASTER: ENVIRONMENT ISSUES AND NATIONALIST TENSIONS

Frederic Lasserre
Laval University
Quebec City, Canada

AN UNPRECEDENTED DISASTER UNFOLDING

The village of Muynak was located on an island in the Amu Daria delta in 1956. In 1962, the island had already become a peninsula; in 1970, the sea was 10 km away, then 40 km away in 1980, and 75 km in 1998. The sea bottom became a vast salted desert. In 1987, the retreating sea divided itself in two parts, a small Aral in the north, where the Syr flows, and a large Aral, where the Amu ends its journey when it reaches the sea (Figs. 1 and 2).

Between 1960 and 1998, the level of the small Aral dropped by 13 meters, and the level of the larger by 18 m. The total sea volume was cut by 78% (5,6). The vanishing of the sea constitutes an enormous ecological disaster that radically alters the environmental balance in the region:

- Vast deposits of salts, nitrates and pesticides spread in large quantities on the industrial crops upstream were formed, deposited by water of the Amu and the Syr. The wind now blows on these desiccated deposits, lifting nearly 40 million tons of toxic sediments

Figure 1. The watersheds of the Amu and Syr Rivers (1,2).

each year, and transporting them far inland: this phenomenon poisons the soil, the vegetation, and the inhabitants. The salt fallout can be as high as 500 to 700 kg/ha, which quickly worsens the problem of salt built-up due to poorly drained and excessive irrigation.

- As the Aral Sea, that acted as a climate regulator, gradually disappears, the local climate takes on distinct continental characteristics:
  - The frost-free period was shortened to about 170 days that undermines the profitability of various crops, including cotton, which needs about 200 frost-free days. As a consequence, some farmers have turned to growing rice, a crop that consumes even more water. Summers are shorter but also hotter.
  - Strong winds are more frequent, increasing the amount of toxic sediment that is blown inland. Aral dust has been detected as far away as Belarus, about 2000 km to the West.

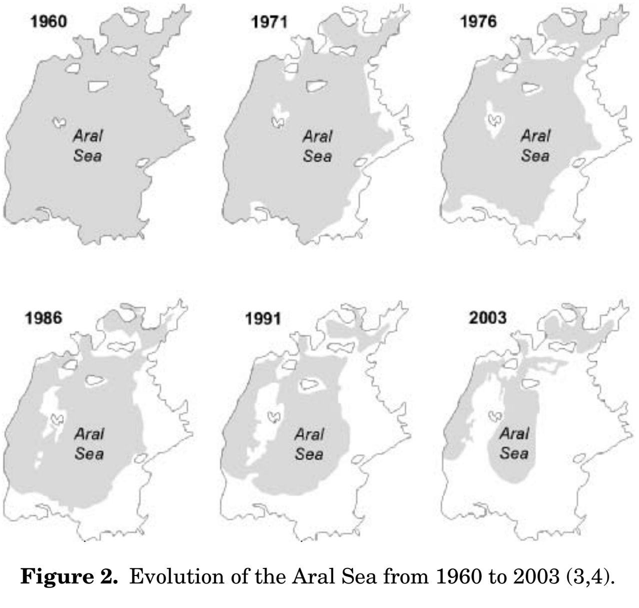

Figure 2. Evolution of the Aral Sea from 1960 to 2003 (3,4).

  - Local rain levels dropped by half in the sea area, and relative humidity dropped by about 30%.
- The salinity of the sea increased from 9.9 g/L in 1960 to 30 g/L in 1994, destroying most of marine life, then to 45 g/l in the larger Aral in 1998 (5). In 1977, the fish catch had already decreased by 75%; in 1982, all commercial fishing had stopped.
- The retreating sea level made the main aquifer level drop, thus drying up several oases around the sea area (7,8).
- Intensive and improperly managed irrigation led to salinization in as much as 95% of cropland (8 million hectares only for cotton), which drives the farmers to increase the amount of fertilizers and water (9,10), thus compounding the water stress.

The Aral Sea, the fourth lake in the world by its surface in 1960, went into a steady decline (Table 1).

On a geologic timescale, changes in the Aral Sea level have been noted: the phenomenon is not, in itself, radically new. The sea is naturally sensitive to variations as it is not very deep (16 m on average) and subject to strong evaporation (about 58 km^3/year, on average). But it is the speed with which the Aral Sea disappeared from 1960 on that hints of human causes, which probably were compounded by natural causes. A sharp decrease in precipitation was observed in the 1970s and 1980s; by itself, it was estimated that it would have led to a decrease in the sea level from 53.5 m in 1960 to 50.8 m in 1986. The rest can be accounted for by large withdrawals for irrigation, a demand all the stronger as rain became scarcer.

DEVELOPING THE VAST CENTRAL ASIAN PLAINS

The Russian Empire introduced cotton in Central Asia in the late nineteenth century. But it became intensively grown under the Soviet regime. In 1937, the Soviet Union, eager to increase its export revenues, became a net cotton exporter. There were social and political goals as well: through the collectivization and the mechanization that were put in place so as to expand large industrial cotton

Table 1. State of the Aral Sea from 1950 to 1998[a]

Year	River Discharge, km^3	Sea Level, m Large Aral After 1987	Surface, km^2	Volume, km^3	Salinity, g/L Large Aral After 1987	Fish Catch, t
1945	62				10	
1960	40	53.41	66,458	1090	10	43,300
1965	31	52.5	63,900	1030	10.5	31,040
1970	33	51.6	60,400	970	11.1	17,460
1975	11	49.4	57,200	840	13.7	2,940
1980	4	46.2	52,500	670	16.5	0
1985	0	42	44,200	470	23.5	0
1987	0	40.5	41,000	404	26.8	0
1989	6	38.6	36,900	330	30.1	0
1990	9	37.8	34,800	304	33.3	0
1993	14 (est.)	37.1	35,424	297	35	0
1998	21 (est.)	31.5	31,541	236	45	0

[a] Reference 3, pp. 182, 191; Reference 6, p. 15.

farms, traditional family farming was destroyed, and with it the remnants of resistance of Central Asian societies toward communism and a centrally planned economy.

Irrigation: The Main Culprit

Water was the key to a vastly expanding cotton production. Canals and dams were planned from 1950 on and built to take the water from the Amu and the Syr to the fields (Table 2). In 1956, the Karakum Canal was inaugurated. It was to irrigate the dry lands of Turkmenistan, when irrigation was considered the solution to humankind's development. In 1980, the highest dam in the world, the Nurek Dam (335 meters high), was commissioned in Tadjikistan.

Thanks to this infrastructural effort, irrigated land grew between 1965 and 1986 at an average rate of 2.1% annually (Table 3). During the same period, water demand doubled. A lot of it was lost through evaporation in open canals or through seepage in uncemented aqueducts. Water drained the fertilizers and the pesticides sprayed over the fields, but smaller amounts reached the sea as more canals were built to develop more acreage. In 1950, 50 km^3 of water flowed into the Aral; in 1990, this volume had dropped to about 7 km^3; the 1990s seem to have been wetter and, combined with the beginning of irrigation rationalization, contributed to an increase of river discharge into the sea. But the amount of water now reaching the Aral is not even enough to stop the sea from retreating, let alone to consider restoring the former Aral, for which at least 35 km^3 would be needed on a regular basis.

After the collapse of the Soviet Union, the Central Asian republics, which were not so eager to become independent, fell into a deep recession. Their main market, Russia, was itself in the throes of economic disarray. Trade was now

Table 2. Development of Irrigated Surfaces, Central Asian Republics, in Thousand Hectares[a]

	1950	1965	1970	1975	1980	1985	1987
Uzbekistan	2276	2639	2696	3006	3476	3930	4109
Kirghistan	937	861	883	911	956	1009	1028
Tadjikistan	361	468	518	567	617	653	675
Turkmenistan	454	514	643	819	927	1107	1224
Kazakhstan	1393	1368	1451	1648	1961	2172	2318
Total	5421	5850	6191	6851	7937	8871	9354

[a] Reference 3, p. 146.

Table 3. Irrigation in the Aral Sea Basin[a]

	Irrigated Surfaces (Million ha)	Withdrawals for Irrigation (km^3)	Withdrawals Per Surface Unit (m^3/ha)	Surfaces Sown with Cotton[b] (ha)	Cotton, % of Irrigated Surfaces[b]
1965	4.8	82	17,000	2287	47.6
1980	6.3	107–126	17,000–20,000	2869	45.5
1985	7	112–133	16,000–19,000	3051	43.6
1990	7.25	109	14,600–17,000	2909	40.1
1995	7.94	100	12,594	2574	32.4

[a] Reference 6, p. 28.

[b] Figures from 1993 for cotton production.

to be done at world market prices in convertible currencies, which was not favorable for them, and they realized they were highly dependent on one crop, cotton. Oil and gas reserves were already known at the time, but few pipelines existed to world markets, and all of them went through Russia, which levied a heavy transit fee (11). In 1992, cotton occupied 60% of the cropland in Uzbekistan. In 1998, the crop still represented a third of all Uzbek exports and the main source of hard currencies. Agriculture is 43% of the GDP in Uzbekistan, 46% in Turkmenistan and 33% in Kazakhstan, and employs 22% of the workforce in Uzbekistan. The Kazakh, Turkmen, and Uzbek governments, where most of the crops were grown, thus realized how difficult it would be to reform irrigation practices drastically as the more pressing problems were hardly environmental but the very survival of the economy.

Besides, available water is already insufficient to sustain all crops in the region. About 25% of the irrigated lands in Uzbekistan now receive only about 70% of the water needed for optimal cotton growth (12). But the worst is still to come. The cotton production levels induced the destruction of the Aral, but they are not even sustainable because of salt intrusion that is often fought back by spraying even more water. In 1994, 28% of the irrigated lands in the Aral basin, as against 23% in 1990, were plagued by salt deposits that reduced their productivity by 20 to 50%. Yields were adversely affected. Between 1981 and 1986, cotton production dropped by 21.3% in the Soviet Union, largely caused by salt buildup and poor management. To make up for this lost production, several countries opted to wash out the fields, which was temporarily effective but induced an even faster rise of the water table; irrigated surfaces were also increased, compensating the decreasing yield by larger surfaces, especially in Turkmenistan, where a rise of 31% occurred between 1990 and 1994.

Saving Water?

Irrigation is responsible for about 93% of the water consumption in the two Aral Sea basins. Solutions to the crisis have been known for long: they require a sharp increase in water yield and the reduction, if not the elimination, of the most consuming crops, such as rice, and their replacement with less demanding cereals, such as sorghum or millet (savings expected: 3 km^3). Unprofitable crops could be phased out (a water saving potential of about 15 km^3). The improvement of agricultural techniques in the most profitable cotton fields, notably by introducing more effective irrigation techniques such as sprinklers, could add an additional 8 km^3. The repair, lining, and tightening of canals could help save about 10 km^3.

But these solutions are all costly, whether financially or socially. Besides, they imply cooperation among five governments to implement water sharing and management schemes.

AFTER INDEPENDENCE, THE LONG ROAD TO AN AGREEMENT

When the Central Asian republics were federated states under the authority of the Soviet Union, arbitration and infrastructural development decisions were made in Moscow. The development of the Aral Sea basins through intensive irrigation was not questioned at the time. When independence took place, the resource sharing designed during Soviet times appeared all the more unjust as downstream users, which could benefit from irrigation, pressured the two upstream countries, Kirghizstan and Tadjikistan, not to interfere with water resources stored in reservoirs located in their territories; these very downstream users are also well endowed with hydrocarbon resources. Suddenly becoming independent in 1991 due to the collapse of the Soviet Union, the five republics were forced to face bitter disputes and look for cooperative schemes.

There were tensions, even before 1991. In 1989, clashes took place between Kirghiz and Tadjik farmers over water allotments. In the Osh area in Kirghizstan, another clash between Uzbeks and Kirghizs caused 300 deaths in 1990. Along more and more closed borders, patrols are organized to prevent border crossings and irrigation canal sabotage.

Bitter disputes have emerged between Turkmenistan and Uzbekistan, as the former keeps expanding its irrigated acreage, thanks to the Karakum Canal, consuming larger amounts of the Amu Daria. Tension is altering relations between Uzbekistan and Tadjikistan, as the latter is complaining about the heavy prices Uzbekistan is demanding for the gas it delivers, whereas water in the Amu flows freely from Tadjikistan to Uzbekistan; Tadjikistan threatened to bill Uzbekistan for the resource, on the basis that the water comes from its territory. Tension also disrupted relations between Uzbekistan and Kirghizstan over the use of the Toktogul Dam water. Built during the Soviet era in mountainous and water-rich Kirghizstan, it was mainly meant to store water during the winter to provide water for the irrigated lands of the Ferghana Valley downstream, mainly in Uzbek territory. In January 2001, the Kirghiz government defaulted on its gas payments to Uzbekistan; the Uzbek government suspended gas deliveries that are mainly used for urban heating. Kirghizstan opened the Toktogul Dam sluices to produce electricity. The induced flood destroyed several dikes in the Uzbek Ferghana region. The Kirghiz pressure peaked when, after 15 days of letting the reservoir gradually empty to generate electricity, Kirghizstan notified the Uzbek government that they would not be able to guarantee water deliveries for the next growing season. This diplomatic skirmishing over energy deliveries bore fruit inasmuch as an agreement was signed by both countries for more cordial relations, recognizing the interests of Kirghizstan in using the waters of the Syr River—but it also underlined the acute tensions over the use of this very water.

It is not that the five Central Asian republics have not undertaken negotiations. As early as 1993, they signed the *General Agreement on the Crisis of the Aral Sea*, also signed by Russia, which provided water sharing cooperative mechanisms and technical cooperation to improve agricultural yields. Because of poor funding and lack of political will, the agreement was never enforced. As of 1995, the European Union and USAID stepped up their cooperation by creating large funds to help the functioning of institutions such as the *Interstate Council* *for Addressing the Aral Sea,* or the *Optimization of the Use of Water and Energy Resources in the Syr-Darya Basin* project. So far, progress has been very slow.

THE TEMPTATION OF UNILATERALISM OR SOVIET STYLE MANAGEMENT

Soviet-style planning led to the idea that water was not particularly scarce; when it did become so, gigantic diversion of Siberian rivers, the Ob and the Ienissei, were designed by engineers to cope with the issue, only to be stopped in 1986 by Mikhail Gorbachev. It is all the more difficult for the Central Asian republics to reach an agreement as they lack financial resources to implement any conversion scheme in agriculture because cotton is a major export resource for the three downstream users. Besides, the economic recession that was triggered by the disintegration of cooperation within the former Soviet Union creates social unrest on which political opposition, particularly Islamic movements, capitalize to question the governmental authority, especially in Uzbekistan. In the latter, difficulties in reforming agriculture are compounded by the strong lobby of industrial cotton farmers.

Nationalist rhetoric, often associating water issues with ethnic tensions, developed in the region. This trend must be analyzed in the framework of tensions erupting between states as they close their respective borders and as governments have been charged with encouraging nationalist stances that help them stay in power in these troubled post-Soviet times. Besides, Uzbekistan is increasingly considered by its neighbors as pursuing a regional power policy (13).

Economic pressures and the slow progress in negotiations have encouraged several governments to act unilaterally to protect their interests. For instance, in 1994, Turkmenistan unilaterally resumed the extension of the Karakum Canal, thereby triggering strong protests from Uzbekistan. The annual volume diverted into the canal reportedly amounts to between 15 and 20 km^3 (14). Similarly, in 1998, Kazakhstan succeeded in building a dike between the Little (or Northern) and the Large Aral, trying to keep the flow of the Syr Daria in the Little Aral to restore it gradually. The World Bank showed interest in the project. The Little Aral level grew by several meters and a few fish species have been reintroduced. The Kazakh government hopes that this project can help save the Little Aral, but it also means that the Large Aral is doomed, as any conservation project must rely only on the Amu Daria. Uzbekistan strongly protested against the Kazakh scheme. Kazakh officials replied that Uzbekistan has not shown any real desire to reduce its water consumption.

CONCLUSION

The Aral Sea is disappearing because of the development of unsustainable levels of agricultural demand for irrigation. Cotton is a major crop in Central Asia. The Soviet government vastly expanded this culture for social, political, and industrial reasons. To do so, water development was planned regionally without considering the various federated republics.

When they became independent in 1991, the harsh economic recession and social unrest made any move to reform agriculture and implement water-saving measures difficult. Nationalist tensions, played upon by bureaucracies and governments, also played a role in stalling the negotiations that had begun as early as 1992. The Aral Sea keeps receding, and the temptation of unilateral decision-making remains strong in the five republics. Water withdrawals for irrigation have decreased since 1985, but they remain far too high to prevent the Aral Sea from disappearing.

Failure to reach an agreement could doom the disappearing sea, all the more if peace prevails in Afghanistan. A quarter of the drainage basin of the Amu Daria lies in that country. If peace is to return there, agriculture will resume, and with it, irrigation, thereby further reducing the available volume for Turkmenistan and Uzbekistan.

Acknowledgments
This research was made possible thanks to a grant from the FQRSC.

BIBLIOGRAPHY

1. Bethemont, J. (1999). *Les Grands Fleuves. Entre Nature et Société*. Armand Colin, Paris, p. 197.
2. Lasserre, F. and Descroix, L. (2003). *Eaux et Territoires: Tensions, Coopérations et Géopolitique de L'eau*. Presses de l'Université du Québec, Québec, p. 296.
3. Létolle, R. and Mainguet, M. (1993). *Aral*. Springer-Verlag, Paris, p. 190.
4. German Remote Sensing Data Center (DFD). (1997). *Chronology of the Dessication of the Aral Sea*. Earth Observatory, NASA.
5. Klötzli, S. (1994). The water and soil crisis in Central Asia—a source for future conflicts? *ENCOP Occasional Paper No. 11*, Center for Security Policy and Conflict Research, Swiss Peace Foundation, Zurich/Bern.
6. Micklin, P. (2000). *Managing Water in Central Asia*. Royal Institute of International Affairs, London, p. 13.
7. Perera, J. (1999). *The Aral Sea: Approaching Total Disaster*. Écodécision, September 1992, p. 60.
8. de Villiers, M. *Water*. Stoddart, Toronto, pp. 133–134.
9. Postel, S. (1999). *Pillar of Sand; Can the Irrigation Miracle Last?* WorldWatch Institute, Norton, New York, p. 94.
10. Coisne, S. (2000). Fleuves en sursis. *Sciences et Vie* **211**: 119.
11. Lasserre, F. (2002). Afghanistan: version moderne du Grand Jeu? In: *Espaces Et Enjeux: Méthodes d'une Géopolitique Critique*. F. Lasserre and E. Gonon (Eds.). L'Harmattan, Paris, pp. 180–190.
12. Belyaev, A. (1995). Water balance and water resources of the Aral Sea basin and its man-induced changes. *GeoJournal* **35**(1): 17–19.
13. Lasserre, F. (2004). Permanence et bouleversements en Asie centrale. L'insertion de l'étranger proche russe dans la dynamique stratégique internationale. In: *Asie Centrale et Caucase: une Sécurité Mondialisée*. G. Hervouet, T. Juneau, and F. Lasserre (dir.). Presses de l'Université Laval, Québec.

14. Glantz, M. (1998). Creeping environmental problems in the Aral Sea basin. In: *Central Eurasian Water Crisis. Caspian, Aral and Dead Seas*. I. Kobori and M. Glantz (dir.). United Nations University Press, Tokyo, p. 45.

LAKE BAIKAL—A TOUCHSTONE FOR GLOBAL CHANGE AND RIFT STUDIES

Deborah Hutchinson
Steve Colman
U.S. Geological Survey

> "The Lake Baikal rift system is a modern analogue for formation of ancient Atlantic-type continental margins. It tells us the first chapter in the story of how continents separate and ultimately develop into ocean basins like the Atlantic Ocean."
> —Deborah Hutchinson
>
> "Lake Baikal is a unique, nearly pristine environment for the study of global change. Nowhere else in the world can we go to study so long a record of such an important, but little known, part of the global climate system."
> —Steve Colman

THE SPECIAL ENVIRONMENTAL AND GEOLOGICAL SETTINGS OF LAKE BAIKAL PROVIDE UNPARALLELED OPPORTUNITIES FOR RESEARCH AND FOR INTER-NATIONAL COOPERATION

Lake Baikal is to Russia what the Grand Canyon is to the United States: a magnificent natural resource that instills national pride and awe. The Presidents of the United States and the Russian Federation recognized the uniqueness of Lake Baikal. In a recent Joint Statement, they affirm the need to conserve the environment of Lake Baikal and to use its potential for research in limnology, geology, and global climate change. The **U.S. Geological Survey's** (USGS) Office of Energy and Marine Geology and Office of International Geology, supported by the Coastal Geology Program and the Global Change and Climate History Program, are involved in broad collaborative programs to study Lake Baikal with the **Russian Academy of Sciences** and with a number of American universities through the (U.S.) **National Science Foundation**.

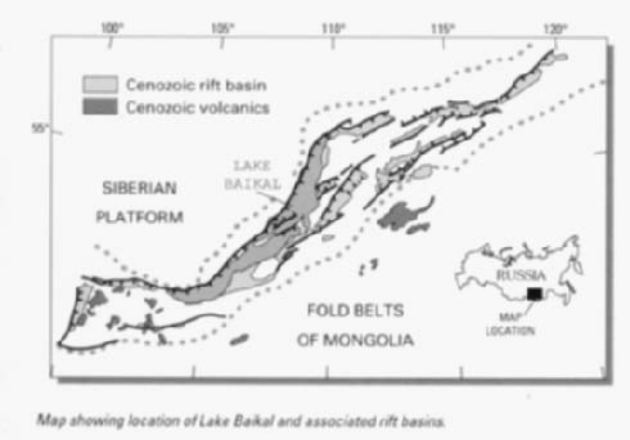

Map showing location of Lake Baikal and associated rift basins.

UNDERSTANDING THE ORIGIN OF THE LAKE BAIKAL RIFT CONTRIBUTES TO UNDERSTANDING ONE OF THE FUNDAMENTAL PHENOMENA BY WHICH THE HISTORY OF THE EARTH IS RECONSTRUCTED

Continental rifts, like the Lake Baikal rift, and their end products, such as passive continental margins like the east coast of the United States, are ubiquitous in the Earth's geologic record. They contain information from which a significant amount of the Earth's history has been interpreted. Due to their high sedimentation rates, large rift lakes like Lake Baikal have great potential for providing high-resolution information about both tectonic and climatic change. Significantly, sedimentary deposits of continental rifts are also associated with many of the Earth's hydrocarbon and mineral deposits.

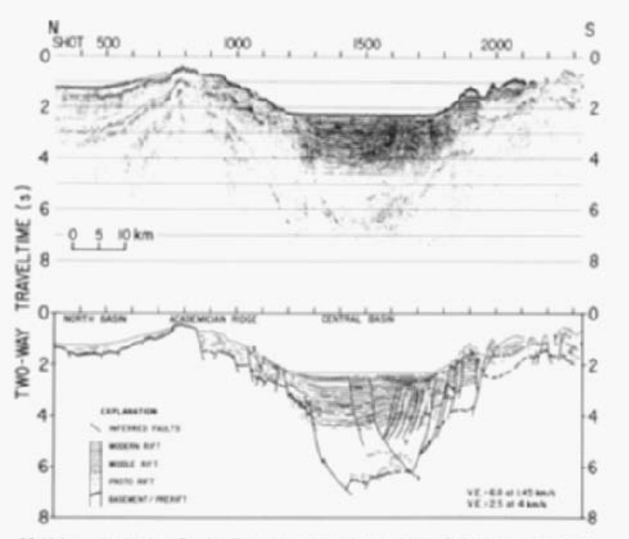

Multichannel seismic reflection line across central part of Lake Baikal showing seismic data (top) and interpretation (bottom). The thickest deposits are confined to a narrow trough that is 15 to 20 kilometers (9 to 12 miles) wide.

USGS AND RUSSIAN COOPERATIVE STUDIES HAVE BEGUN TO RESOLVE THE THREE-DIMENSIONAL GEOMETRY OF THE LAKE BAIKAL RIFT

Sediments of Lake Baikal reach thicknesses in excess of 7 kilometers (4 miles), and the rift floor is perhaps 8 to 9 kilometers (more than 5 miles) deep, making it one of the deepest active rifts on Earth. The shallowest sediments may contain the only known freshwater occurrence of natural gas hydrates. Maps of complex fault patterns and changing depositional environments provide the first opportunities to describe the development of the lake and to help explain its unique flora and fauna.

The RV Vereshchagin, a coring and high-resolution seismic vessel, docked on central Lake Baikal.

UNIQUE CHARACTERISTICS OF THE LAKE BAIKAL ENVIRONMENT COMBINE TO PRODUCE AN ESPECIALLY PROMISING SITE FOR STUDIES OF CLIMATE HISTORY

Lake Baikal is the largest freshwater lake on Earth containing 23,000 cubic kilometers of water, or roughly 20 percent of the world's total surface fresh water. It contains as much fresh water as the Great Lakes of North America combined. At over 1,600 meters (5250 feet), it is the deepest lake in the world, and at perhaps more than 25 million years old, the oldest as well. The water of Lake Baikal is so fresh that calcium carbonate does not survive in the fossil record. Despite the lake's great depth, its water is well-oxygenated throughout creating unique biological habitats.

THE HIGH LATITUDE OF LAKE BAIKAL MAKES IT PARTICULARLY SENSITIVE TO CLIMATIC VARIATIONS

Climate variations, including those resulting from atmospheric accumulation of carbon dioxide, are more pronounced at higher latitudes. Although the lake contains a record of glaciation of surrounding mountains in its sediments, it is unique among large, high-latitude lakes in that its sediments have not been scoured by overriding continental ice sheets.

Ice-based drilling rig being tested on a small marginal lake in North Baikal. Photo by D. Williams.

UNITED STATES AND RUSSIAN STUDIES OF SEDIMENT CORES TAKEN FROM LAKE BAIKAL PROVIDE A DETAILED RECORD OF CLIMATIC VARIATION OVER THE PAST 250,000 YEARS

Much attention is focused on numerical models of climate change but there have been few means for reliably testing or modifying boundary conditions of general circulation models. Studies of sedimentary environments in Lake Baikal provide important opportunities to establish ground truth for general circulation models. Very little data exist for long-term climate change from continental interiors; most of the data record derives from the marine or maritime environments. Finally, studies of past environments contribute to understanding the extent to which human activity affects natural conditions in the lake.

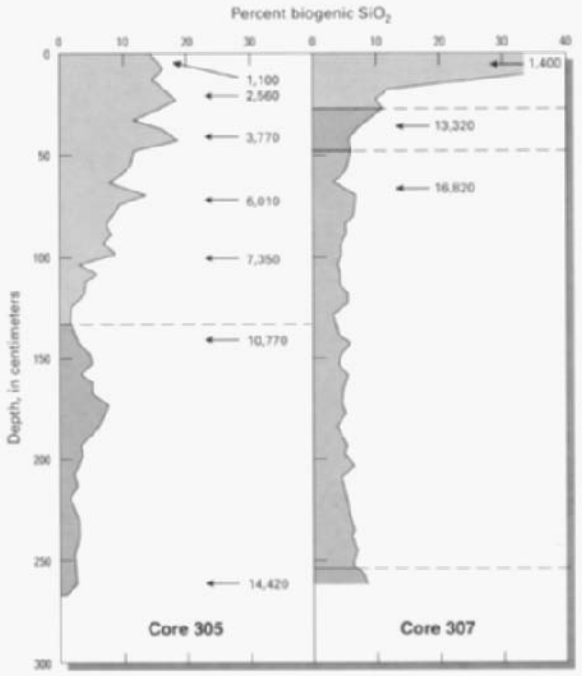

Graph showing percentage of biogenic silica in two Lake Baikal cores. Since diatoms and algae are the most important microorganisms in the lake, biogenic silica is a measure of primary productivity. Radiocarbon ages in years are shown by arrows. First-order changes are caused by glacial (blue)-interglacial (red) cycles, with transitions shown in purple. Second-order fluctuations, such as the peak at 3,770 years, are due to shorter climatic events.

SEISMIC AND SEDIMENT CORE ANALYSES ARE USED TO FIX FUTURE DRILLING SITES IN LAKE BAIKAL

Ice-based drilling operations begun in early 1993 are providing longer (over 100 meters in length) cores of Baikal sediments. Analyses of these cores are expected to reveal the climatic, environmental, and geological history of the region as far back as 5 million years. Seismic data will be tied to cores and drill samples to estimate rates of climate change and to map the history of the lake and rift. Very deep drilling in Lake Baikal remains technologically challenging; therefore, the deepest deposits of the rift are not likely to be sampled soon. However, the potentially very long record of sedimentation in Lake Baikal provides unique opportunities to understand the Cenozoic climate history of the Earth and to describe how continents begin to break apart, giving rise to new ocean basins.

BASE FLOW

Sharad K. Jain
National Institute of Hydrology
Roorkee, Uttranchal, India

Vijay P. Singh
Louisiana State University
Baton Rouge, Louisiana

Water may reach a river by flowing over the catchment surface, moving through the soil, or through groundwater. Based on the path taken by water, streamflow may be divided into surface flow, interflow, and base flow. The ASCE (1) defined base flow as the runoff that has reached the stream or river by passing first through the underlying aquifer, rather than by flowing directly on the ground surface. Thus, base flow is that portion of streamflow that is naturally and gradually withdrawn from groundwater storage or other delayed sources. The other names of base flow are groundwater flow, seepage flow, low flow, and fair weather flow.

Perennial streams depend on base flow for discharge between runoff producing events. The presence of base flow around the year indicates humid climate and a shallow water table that is hydraulically connected to the stream. Base flow is absent in (semi)arid climates and areas of deep groundwater. Of course, it is assumed that the streamflow is not influenced by storage or diversion.

FACTORS AFFECTING BASE FLOW

Base flow depends on precipitation, geologic conditions, and hydrogeologic controls that govern groundwater movement. Sources, such as lakes, marshes, snowpack, and channel bank, may also supply water for base flow. At a given time, more than one source may be supporting base flow. Climate influences recession through recharge and evapotranspiration.

Base flow contribution to streamflow varies widely according to the geologic nature of the water-table aquifer. A measure of the amount of water available from a water-table aquifer is the coefficient of storage, or specific yield. The lateral movement of groundwater is slower than vertical movement because the hydraulic gradient is smaller for lateral movement. For this reason, groundwater "mounds" and water tables beneath ridges may be at higher elevations than stream channels. This provides a gradient that directs base flow to the stream. The supply from groundwater to the channel will continue as long as the necessary gradient is present. If no additional infiltration recharges an aquifer, the hydraulic gradient decreases as water moves to the stream from higher elevations. Consequently, less water will be discharged to the stream with time. This process is called base flow recession.

Bank storage may also contribute to base flow depending upon the position of the water table, riverbank material, soil characteristics near the riverbank, and the rise and fall of the stream. Usually, bank storage is largely depleted by the time the base flow peak occurs. Other water-table aquifers such as fractured rocks and solution cavity rock openings also contribute base flow to streams. The important parameters related to groundwater outflow are the total length of perennial streams, average basin slope, and drainage density. Singh provides a detailed discussion on base flow (2).

BASE FLOW RECESSION

Streamflow recession represents withdrawal of water from storage without inflow. It constitutes the falling limb of a hydrograph to the right of the point of contraflexure. When all surface runoff has passed, the flow in the stream consists of base flow. Figure 1 shows components of a typical hydrograph.

Base flow is at maximum immediately after a storm because precipitation has just recharged the water table. As water is withdrawn from an aquifer, groundwater supplies are diminished and base flow decreases. This process continues until all groundwater available for base flow is exhausted or additional recharge occurs.

Many different forms of base flow recession curves have been proposed. These are discussed in what follows.

Simple Exponential

The basic differential equation that governs the flow in an aquifer was presented by Boussinesq in 1877 (3). Boussinesq solved this equation by assuming that the Depuit assumption of negligible vertical flow holds and the effect of capillarity above the water table can be neglected. This assumption leads to the solution,

$$Q_t = Q_0 \exp(-t/a) \tag{1}$$

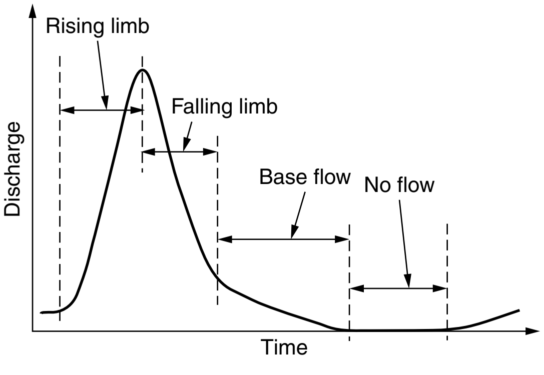

Figure 1. Sections of a streamflow hydrograph.

or its alternate forms

$$Q_t = Q_0 k_r^t \tag{2}$$

$$Q_t = Q_0(10)^{-t/b} \tag{3}$$

where Q_0 is flow at time $t = 0$; Q_t is flow at time t; and a, k, and b are constants. Parameter k_r is known as the recession constant or the depletion factor. It is a dimensionless quantity that depends on the unit of time selected (commonly, it is 1 day). It lies in the interval [0,1] and is normally greater than 0.7. Constants a and b, known as storage delay factors, have the dimension of time. Constant b denotes the time required for the flow to decrease by a factor of 10 or one log cycle (see Fig. 2). Similarly, constant a denotes the time required for the flow to decrease by a factor equal to e, or one natural log cycle, as shown in Fig. 3. In other words, a is the time elapsed between any discharge Q and Q/e of the recession and is related to the 'half flow period' by (4)

$$a = -t_{0.5}/\ln(1/2) \tag{4}$$

where $t_{0.5}$ is the time required to halve the streamflow. Parameters a and k_r are related by

$$a = -t/\ln k_r$$

$$k_r = (10)^{-l/b} \tag{5}$$

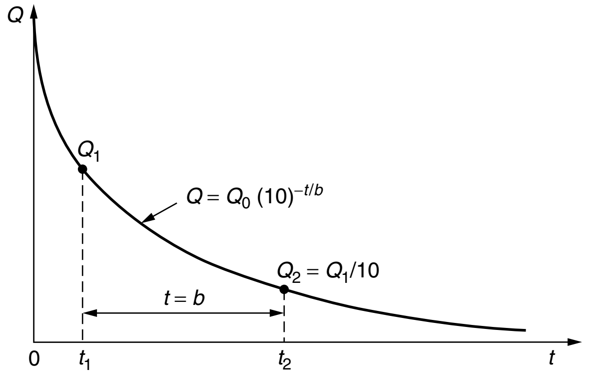

Figure 2. Base flow recession: determination of storage delay time b.

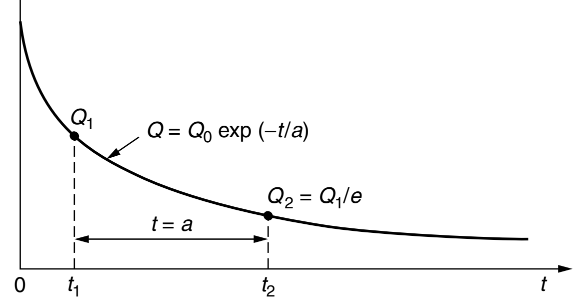

Figure 3. Base flow recession: determination of storage delay time a.

Horton (5) suggested a nonlinear recession equation, known as Horton's double exponential:

$$Q_t = Q_0 \exp(-a_2 t^m) \tag{6}$$

where a_2 and m are constants.

Singh (6) derived ideal base flow curves by solving the Boussinesq equation using a finite-difference equation for two cases: a shallow aquifer and fully penetrating stream and a deep aquifer and shallow-entrenched stream.

More on Exponential Recession

Equation 1 represents a first-order process or an exhaustion phenomenon expressed by

$$ds/dt = -Q, \quad Q(0) = Q_0 \tag{7}$$

and

$$s = aQ \tag{8}$$

where s is storage. The solution of Eqs. 7 and 8 is Eq. 1. The form of Eq. 1 does not change if discharges are replaced by volumes. The volume of flow during the time from $t - T$ to t is given by

$$\begin{aligned} V &= \int_{t-T}^{t} Q\,dt = \int_{t-T}^{t} Q_0 \exp(-t/a)\,dt \\ &= aQ_0[\exp(T/a) - 1]\exp(-t/a) \end{aligned} \tag{9}$$

For a fixed time interval T, such as 1 day,

$$V = V_0 \exp(-t/a) \tag{10}$$

in which

$$V_0 = aQ_0[\exp(t/a) - 1] \tag{11}$$

Thus, the recession constant can be estimated from Eq. 10. It may be preferable to use volumes instead of discharge because volume is less sensitive to errors than discharge.

Double Exponential

When base flow is plotted on semilog paper, the resulting curve is nonlinear which can be represented (5) by a double exponential of the form,

$$Q = Q_0 \exp(-mt^n) \tag{12}$$

where $\exp(-m) = k_r$ is the recession constant and n is a constant. Parameters m and n can be obtained either graphically or by using the method of least squares. By taking the log of both sides of Eq. 12 twice,

$$\ln(Q_0/Q) = mt^n \tag{13}$$

$$\ln[\ln(Q_0/Q)] = \ln(m) + n\ln(t) \tag{14}$$

Equation 14 represents a straight line.

Hyperbola

Another equation for a stream located on a horizontal impermeable lower boundary with an initial curvilinear water table and zero water-level elevation in the stream is of the form

$$Q = Q_0/(1 + ct)^2 \tag{15}$$

where c is a constant. This equation will hold where the groundwater storage is at its maximum after surface runoff stops or if no further precipitation or snowmelt occurred until the streamflow ceased. Parameter c can be estimated as

$$c = [(Q_0/Q)^{0.5} - 1]/t \tag{16}$$

The outflow from storage can be represented by a general non linear relation,

$$Q = kS^p \tag{17}$$

where S is storage and k and p are constants. The recession will be concave downward for $p > 1$ and upward for $p < 1$. Values of p greater than unity are typical of many catchments (4). The superposed exponential equation represents recession flow as a superposition of n exponential terms; each shows the response of a linear reservoir whose recession constant is c_i:

$$Q_t = \sum_{i=1}^{n} Q_{oi} \exp(-t/c_i) \tag{18}$$

Ice-Melt Recession

For snow and ice-melt conditions, base flow recession can be represented by (7)

$$Q = at^{-n} + b \tag{19}$$

where a, b, and n are constants. As time increases, Q asymptotically approaches a constant value b. This may typify base flow recession for permanent snow and ice. Constants a, b, and n can be determined either graphically or by using the method of least squares.

An exponential equation for base flow recession in watersheds that have permanent ice and snow is of the form (7),

$$Q = a + (Q_0 - a)k_r^t \tag{20}$$

where a and k_r are constants. This is similar to Eq. 19, where Q asymptotically approaches a constant value as t increases.

Base Flow Recession as Auto Regressive Processes

The base flow recession Eq. 2 can also be written as

$$Q_{t+1} = k_r Q_t \tag{21}$$

However, this equation is an approximation and contains errors due to a linearity assumption. The input data also contain some errors and, therefore, Eq. 22 can be written as

$$Q_{t+1} = k_r Q_t + \varepsilon_{t+1} \tag{22}$$

where ε_t are independent, normally distributed errors that have zero mean and a constant variance. This equation represents a first-order autoregressive process, denoted by AR(1), where k_r is the autoregressive parameter. If the errors are additive in log-space, Eq. 23 becomes

$$Q_t = k_r Q_{t-1} e^{\varepsilon t} \tag{23}$$

Vogel and Kroll (8) noted that the error structure of Eq. 23 is more representative of actual streamflow records. Base flow recession can also be interpreted as an AR(2) process:

$$Q_t = (\varphi_1 Q_{t-1} + \varphi_2 Q_{t-2})\varepsilon_t \tag{24}$$

Another way to look at recession is as an analogy with a linear reservoir where the storage–outflow relationship is assumed linear:

$$S = kQ \tag{25}$$

where S is storage, Q is outflow, and k is a constant whose dimension is time. The continuity equation for a reservoir is

$$I - Q = dS/dt \tag{26}$$

Assuming that the inflow is zero and letting $Q(t_0) = Q_0$ (the outflow at $t = t_0$) yields

$$Q_t = Q_0 \exp[-(t - t_0)/k] \tag{27}$$

Equation 2 can be written in discrete form as

$$Q_{j\Delta t} = Q_0 k_r^j = Q_{(j-1)\Delta t} k_r \tag{28}$$

where $t = j\Delta t$. Comparing Eq. 27 with Eq. 28,

$$k = -\Delta t/\ln k_r \tag{29}$$

Substituting from Eq. 27 in Eq. 26, letting $I = 0$, and integrating,

$$S_t = -Q_0/\ln k_r \tag{30}$$

DETERMINATION OF BASE FLOW RECESSION CONSTANTS

The recession constant k_r or parameters a and b can be determined in a number of ways.

Graphical Method

Constants a and b can be determined by constructing Figs. 2 and 3 and choosing any two values of discharge that satisfy the condition specified there. Note that these constants vary with the choice of discharges. Therefore, it is appropriate to obtain a number of values of a and b

and then take an average. Alternatively, one can take the logarithm of Eqs. 1–3,

$$\ln(Q) = \ln(Q_0) - t/a \tag{31}$$

$$\ln(Q) = \ln(Q_0) + t\ln k_r \tag{32}$$

$$\log(Q) = \log(Q_0) - t/b \tag{33}$$

Base flow discharge Q can be plotted against time on a semilog paper using any of Eqs. 31–33. The recession parameters can be determined by fitting a straight line to the plotted data. Alternately, one can also plot Q_{t-1} versus Q_t on a simple graph. For example, if Eq. 2 is valid, this plot will result in a straight line passing through the origin. Its slope will be k_r. Similar interpretations can be advanced for Eqs. 2 and 3. For example, Eq. 1 can be recast as

$$Q_t = Q_{t-1}\exp(-T/a) \tag{34}$$

where Q_t is the discharge at time t and T is the time interval between Q_{t-1} and Q_t. The plot of Q_{t-1} versus Q_t should give a straight line that has the slope $\exp(-T/a)$. Obviously, the slope depends on the time interval, which is commonly 1 day.

Least Squares Method

James and Thompson (9) determined the recession constants by using the least squares method. Any of Eqs. 1–3 can be used with this method. For illustration, Eq. 2 is recast as

$$Q_t = k_r Q_{t-1} \tag{35}$$

The recession constant can be estimated by minimizing the sum of the squares of the differences between the observed and computed values of discharge:

$$R = \sum_{t=1}^{N}(Q_t - k_r Q_{t-1})^2 \tag{36}$$

where N is the number of discharge values and Q_t corresponds to the observed values. To estimate k_r, the derivative of R with respect to k_r is set equal to zero to yield

$$k_r = \sum_{t=1}^{N} Q_t Q_{t-1} \Big/ \sum_{t=1}^{N} Q_{t-1}^2 \tag{37}$$

Statistical Methods

The recession constants can be estimated by the method of moments. Only one parameter is to be estimated, so the first moment about the origin will suffice. Again, any of Eqs. 1–3 can be used. To illustrate, consider Eq. 1. Let $M_1(Q)$ be the first moment of Q about the origin. Then,

$$M_1(Q) = \int_0^\infty tQ_0\exp(-t/a)dt \Big/ \int_0^\infty Q_0\exp(-t/a)dt = a \tag{38}$$

Thus, recession constant a for a given base flow sequence is equal to the first moment of that sequence about its origin.

If Eqs. 22 or 24 are treated as regression equations, the least-squares method can be used to estimate the recession constant. For the AR(1) model given by Eq. 22, the ordinary least squares estimate of k_r is

$$k_r = \frac{\sum_{t=1}^{n-1} Q_{t+1}Q_t}{\sum_{t=1}^{n-1} Q_t^2} \tag{39}$$

where n is the total number of consecutive observations for a site.

While comparing the various techniques to estimate recession constants, Vogel and Kroll (8) found that the recession constant obtained by treating the base flow as an AR(2) process that has an additive residual in log-space (e.g., Eq. 24) led to a regional regression model of low-flow statistics that has the highest R^2, lowest standard error, and highest t-ratios of model parameter estimates.

MASTER BASE FLOW RECESSION CURVE

The master base flow recession curve is a composite recession curve that represents mean recession behavior. Of course, the information about recession variability is lost in the process. This curve can be constructed in several ways (7); the most popular are experimentation and the correlation method.

Experimentation involves observing the discharge at a number of time intervals encompassing the entire dry-weather period. These are then plotted against time on semilog paper, and a best fit straight line is drawn through the plotted points. The resulting curve is a recession curve.

The correlation method involves plotting the discharge for the initial day of each segment on log-log paper, starting with the largest discharge, against discharge T time units later (where $T = 1/2$, 1, 2, 5, etc.). Plotting is continued for as many segments as possible until a good correlation is established for drawing a line. The points above the line represent surface flow and those below it represent recession.

Algorithms for Selecting Recession

The beginning of a recession segment can be a fixed discharge or a variable. A fixed value restricts the domain of recession analysis to flows less than that value. The beginning may also be specified as a certain time after the rainfall or after the peak of the hydrograph. Sometimes, some initial values are ignored to minimize the influence of surface flow, and some values near the end may be ignored to obviate the influence of the next storm.

Vogel and Kroll (8) describe an automatic hydrograph recession selection algorithm to separate the recession part from the discharge data. In this, a recession begins when a 3-day moving average begins to decrease and ends when the same begins to increase. Recessions longer than 10 days or more were identified by them.

Applications of Base Flow Recession

A major practical application of base flow recession is forecasting low flows in a river for drought management, planning drinking water supply, maintaining water quality in the river, or ecological planning. Other applications of base flow analysis are determining the relation between hydrologic and geologic parameters of a drainage basin, evaluating the effect of agricultural practices, locating suitable areas for induced infiltration, controlling withdrawal of groundwater for irrigation during low flow periods, and determining storage requirements to maintain minimum flow in rivers.

BASE FLOW SEPARATION

Base flow separation, also called hydrographic analysis, is the process of separating surface runoff from base flow. Even though such separation is somewhat arbitrary and subjective, it is useful in many analyses. Several techniques have been developed for base flow separation. Some popular techniques are described below.

Area Method

The area method of base flow separation is based on a nonlinear relation between time and area (10):

$$N = bA^{0.2} \tag{40}$$

where A is the drainage-basin area in km^2, b is a coefficient equal to 0.8, and N is the time in days from the hydrograph peak (see Fig. 4). This equation is not suitable for smaller watersheds. It generally gives a longer time base. For example, if $A = 1000$ km^2, then $N = 3.18$ days, if rainfall occurs for 6 hours, its effect will be felt for more than 3 days.

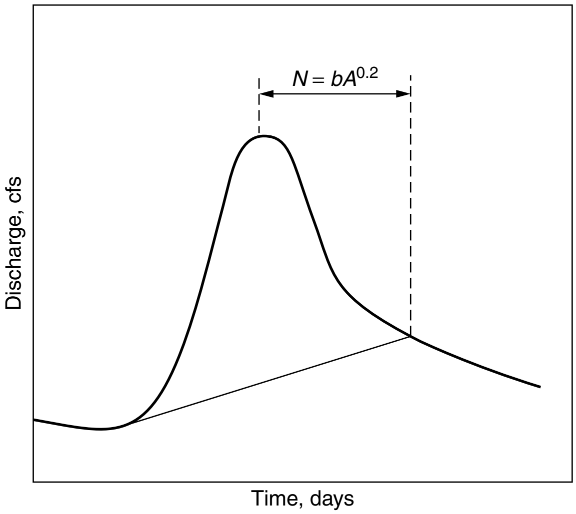

Figure 4. Base flow separation based on area.

It is convenient to draw a separation line directly from the chosen groundwater discharge on the receding limb to the point under the hydrograph peak. Although this linear separation does not represent the true boundary between direct runoff and groundwater runoff, the error may be acceptable in most cases.

Three-Component Separation

Three-component separation involves separating surface runoff, interflow, and base flow. The method, developed by Barnes (11), is based on Equation 2 and is illustrated in Fig. 5. First, streamflow recession is plotted on semilogarithmic paper. In Fig. 5, the groundwater recession plots approximately as a straight line, where $K_r = 0.992$. By extending this straight line under the hydrograph to the point directly under the point of inflection E and to B on line AB, points B and J are connected arbitrarily by a straight line. The area under the hydrograph above BJH is considered direct flow and that area below BJH is considered groundwater flow. The direct runoff is replotted, and a straight line IL where $K_r = 0.966$ is fitted and extended to point I directly under inflection point E and to the beginning point M. The line MIL divides the replotted hydrograph into surface runoff on top and interflow below.

Singh and Stall Method

Although Barnes' method has been used for quite some time, it is subjective and has many shortcomings, as pointed out by Kulandaiswamy and Seetharaman (12). Several methods have been developed to overcome these drawbacks and remove subjectivity. The method developed by Singh and Stall (13) is described here.

1. Select single peak hydrographs that are not closely followed by another stream rise.
2. Estimate the storage delay factor K_b and the range of base flow Q_b.

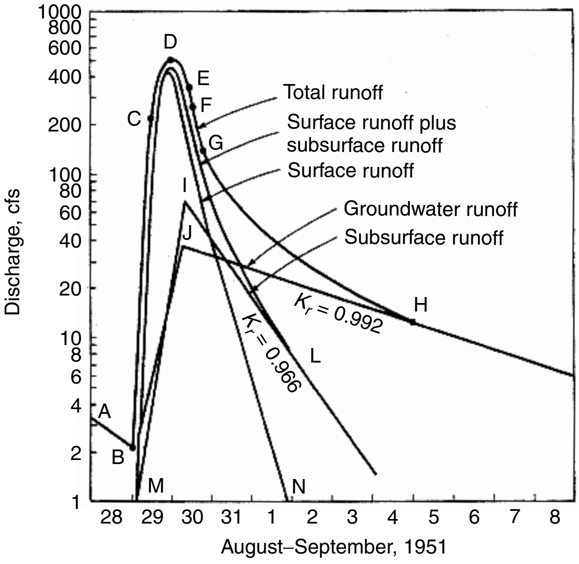

Figure 5. Three-component hydrograph separation.

3. Using an average value of K_b for the desired range of Q_b, begin at a suitable value of Q_b where the hydrograph shape suggests the end of flow from the surface runoff stored in the channel system, and compute daily values of Q_b by progressing backward in time to within 1 or 2 days of the inflection point on the hydrograph recession. Draw the straight-line base flow hydrograph.
4. Subtract Q_b from Q_T, the total discharge and plot the $(Q_T - Q_b)$ hydrograph on a semilog paper. Draw a best fit straight line.
5. Draw a few $(Q_T - Q_b)$ hydrographs by using slightly higher and lower values of K_b. The value of K_b that makes the $(Q_T - Q_b)$ plot nearest to a straight line is adopted. The straight line is denoted as the Q_s line, and the number of days for Q_s to drop in amount by one log cycle is the value of K_s in days.
6. Compute $(Q_T - Q_s)$ by starting a few days after the inflection point on the Q_T hydrograph, and obtain the base flow hydrograph. Note the time in days T_b by which the base flow peak lags behind the hydrograph peak.
7. Analyze a number of hydrographs to obtain sets of values of K_b, K_s, T_b, and the range of Q_b for which a particular K_b is suitable.
8. Draw base flow recession curves for both growing and dormant seasons.
9. Compute values of Q_s using different values of K_s. Add Q_b and Q_s to get Q_T. Draw traces of total hydrograph recessions by starting from various points on the base flow recession curve. The Q_T recessions will define the average limits within which actual Q_T recessions will lie.

Base Flow Separation using Digital Filters

Graphical methods of base flow separation are inconvenient for processing a long streamflow record. Digital filters are helpful in such situations. Basically, streamflow is partitioned into two components:

$$Q(i) = Q_l(i) + Q_h(i) \tag{41}$$

where $Q_l(i)$ is the low-frequency component and $Q_h(i)$ is the high-frequency component. The first can be interpreted as base flow and the second as overland and quick interflow.

A filter employing the two-parameter algorithm of Boughton (14) assumes that base flow at the current time is a weighted average of the direct runoff and the base flow at the previous time interval. Let $Q_b(i)$ and $Q(i)$ be the base flow and total streamflow for the time interval i. Then,

$$Q_b(i) = [p/(1 + C)]Q_b(i - 1) + [C/(1 + C)]Q(i) \tag{42}$$

subject to

$$Q_b(i) \le Q(i)$$

where p and C are the parameters of the model; p can be interpreted as a recession constant. The two-parameter algorithm was found the most satisfactory by Chapman (15). In this case, the parameters have less effect on the BFI (base flow index which is the ratio of the long-term ratio of base flow to total stream flow) than other models. Furey and Gupta (16) developed a filter from a physical point of view.

Chemical and Isotopic Techniques

Chemical and isotopic tracers provide another means to separate a hydrograph into components. An additional advantage of this method is that one can also identify the spatial source of water. The technique has been applied in many studies (17,18). However, its use in base flow separation requires elaborate field work and laboratory analysis that is time-consuming and expensive.

BIBLIOGRAPHY

1. ASCE. (1996). *Handbook of Hydrology*. American Society of Civil Engineers, New York.
2. Singh, V.P. (1992). *Elementary Hydrology*. Prentice Hall, Englewood Cliffs, NJ.
3. Hall, F.R. (1968). Base-flow recessions—A review. *Water Resources Research* **4**(5): 973–983.
4. Tallaksen, L.M. (1995). A review of base flow recession analysis. *Journal of Hydrology* **165**: 349–370.
5. Horton, R.E. (1933). The role of infiltration in the hydrologic cycle. *Transactions, American Geophysical Union* **14**: 446–460.
6. Singh, K.P. (1968). Some factors affecting base flow. *Water Resources Research* **4**(5): 985–999.
7. Toebes, C., Morrissey, W.B., Shorter, R., and Hendy, M. (1969). *Base Flow Recession Curves. Handbook of Hydrological Procedures: Procedure No. 8*. Water and Soil Divisions, Ministry of Works for the National Water and Soil Conservation Organization, Wellington.
8. Vogel, R.M. and Kroll, C.N. (1996). Estimation of base flow recession constants. *Water Resources Management* **10**: 303–320.
9. James, L.D. and Thompson, W.O. (1970). Least squares estimation of constants in linear recession model. *Water Resources Research* **6**(4): 1062–1069.
10. Linsley, R.K., Kohler, M.A., and Paulhus, J.L.H. (1982). *Hydrology for Engineers*. McGraw-Hill, New York.
11. Barnes, B.S. (1940). Discussion of 'analysis of runoff characteristics.' *Transactions, American Society of Civil Engineers* **105**: 104–108.
12. Kulandaiswamy, V.C. and Seetharaman, S. (1969). A note on Barnes' method of hydrograph separation. *Journal of Hydrology* **9**: 222–229.
13. Singh, K.P. and Stall, J.B. (1971). Derivation of base flow recession curves and parameters. *Water Resources Research* **7**(2): 292–303.
14. Boughton, W.C. (1993). A hydrograph-based model for estimating the water yield of ungauged catchments. *Hydrology and Water Resources Symposium*, Institution of Engineers (Australia), 317–324.
15. Chapman, T. (1999). A comparison of algorithms for stream flow recession and base flow separation. *Hydrological Processes* **13**: 701–714.

16. Furey, P.R., and Gupta, V.K. (2001). A physically based filter for separating base flow from streamflow time series. *Water Resources Research* **37**(11): 2709–2722.
17. Harris, D.M., McDonnell, J.J., and Rodhe, A. (1995). Hydrograph separation using continuous open system isotope mixing. *Water Resources Research* **31**(1): 157–171.
18. Ladouche, B. et al. (2001). Hydrograph separation using isotopic, chemical, and hydrological approaches (Strengbach catchment, France). *Journal of Hydrology* **242**: 255–274.

RIVER BASINS

Stefano Pagliara
Francesca Dellacasa
Università di Pisa
Pisa, Italy

RIVER BASIN DEFINITION AND DELINEATION

A river basin (also called drainage basin, watershed, or catchment) is the area that provides runoff to a given section of a stream and is separated from adjacent basins by a divide. All surface runoff produced by precipitation that falls within the area is discharged through the lowest point in the divide.

Basins are delineated on topographical maps, integrated if necessary with aerial photographs or field survey; a line is drawn that encloses all drainage lines and depressions in the basin and passes through the highest points. It is assumed that the movement of groundwater follows the surface divides, but in areas characterized by a particular geologic structure, it is possible that the phreatic divide does coincide with the topographic divide; large amounts of water could be transported below the surface from one basin to another. In this case, the groundwater divide needs to be determined by hydrogeologic prospection.

A river basin is generally considered the fundamental unit for geomorphologic and hydrological studies; it can be treated as an open system that receives energy from the climate (input) and loses energy by evapotranspiration to the atmosphere and by stream flow and sediment yield through the basin mouth (output) that determine storage changes in groundwater and soil moisture (Fig. 1).

A river basin may be described by a Geographic Information System (GIS) combined with a digital elevation model (DEM). A GIS is an electronic system of maps connected to tables of data that describe the features on the maps; a DEM represents a spatial distribution of elevations. The DEM can be structured by a square-grid network, a contour-based network, or a triangulated irregular network (TIN) (1). In a raster-based GIS, the basin is subdivided into uniform grid cells; to each of these cells, numerous attributes relative to basin characteristics can be attached. The GIS system is a useful tool for automating much of the geometric data acquisition necessary for hydrologic analyses (2). A raster-based GIS is well suited to the application of hydrologic rainfall–runoff models based on distributed parameters. Alternatively, the GIS procedure allows computing average basin parameters for the entire watershed that are necessary as input in lumped parameter hydrologic models.

BASIN CHARACTERISTICS

A river basin is composed of several subsystems: the drainage basin itself in its areal extent, and the channel system, in which by a further subdivision, we can distinguish the channel cross section, the channel reach, and the channel network. It is important to describe the river basin in quantitative terms to understand the processes and to analyze the interrelationships existing among these characteristics.

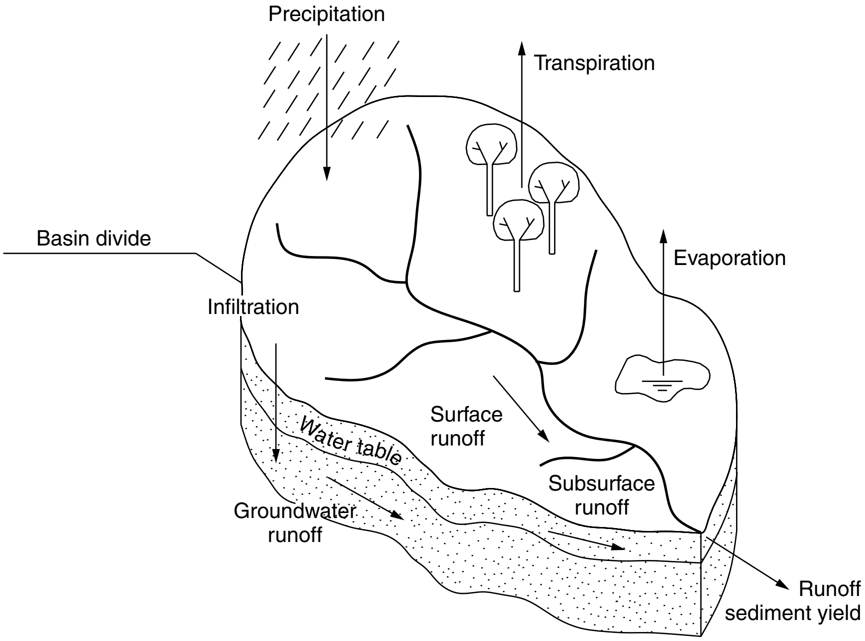

Figure 1. River basin.

Drainage Basin

A river basin's characteristics can be subdivided into topographic, rock and soil, vegetative and land use, and climatic characteristics.

Topographic Characteristics. Topographic characteristics directly influence the transport of water and sediment within a basin. Topographic attributes may be described by index of area, length, shape, and relief.

Basin Area. The area of a drainage basin is the area enclosed by the curve obtained by projecting the basin divide on a horizontal plane. This represents its most important physical parameter because it is correlated with almost every other characteristic, but for this reason it is not always easy to interpret its significance.

Empirical relations between stream discharge Q and basin area A assume the form,

$$Q = aA^b \tag{1}$$

where Q represents some index of stream flow such as the mean annual flood (3). The value of exponent b generally varies from 0.5 to 1.0. Stream flow per unit area is inversely proportional to area because intense precipitation tends to occur on a restricted area. Sediment yield per unit area in many regions, it has been found, is inversely proportional to basin area. Erosion is predominant in small basins, whereas in large basins, there are, in proportion, more locations for temporary sediment storage.

Basin Length. Basin length is usually defined as the length of the mainstream measured along the main channel from the outlet to the basin divide. Length is closely related to area, and a large number of rivers seem to satisfy the relation expressed by Hack:

$$L = 1.4A^{0.6} \tag{2}$$

where L is the stream length in miles and A is the area in square miles (4). The exponent 0.6 suggests that geometric similarity is not preserved as a river basin increases in area because A/L^2 decreases and the basin tends to elongate, becoming longer and narrower.

Basin Shape. The basin shape normally tends to a pear-shaped ovoid, but many substantial deviations from this shape due to geologic constraints have been observed. Several parameters have been proposed to describe basin shape. The form factor F is defined as what is F?

$$R_f = \frac{A}{L^2} \tag{3}$$

where A is the drainage basin area and L the basin length.

The basin circularity compares the basin area with the area of a circle that has the same perimeter p:

$$R_c = \frac{4\pi A}{p^2} \tag{4}$$

The basin elongation is defined as the ratio of the diameter of a circle that has the same area as the basin to the basin length:

$$R_e = \frac{2\sqrt{A}}{L\sqrt{\pi}} \tag{5}$$

R_e assumes values from 0.6 to 0.8 in regions of high relief, and in areas of low relief, values near one have been observed. The usefulness of these shape factors has been analyzed, and the elongation ratio has proved to be the parameter best correlated with hydrology (5). The basin shape affects the stream-flow hydrograph, in particular, the lag time and the time of rise.

Relief. The hypsometric area–altitude analysis expresses the relation between basin area and elevation. It is obtained by plotting the area above a certain elevation versus the elevation itself. The shape of the *hypsometric curve* gives an indication of the basin's geologic evolution any stage in the erosion cycle. The hypsometric curve is useful if hydrologic variables such as precipitation or evaporation, which vary with altitude, are studied.

The *mean basin elevation* can be obtained by integrating of the area–elevation curve and dividing the result by the drainage area; the median elevation is defined as the elevation that corresponds to 50% of the drainage area.

The *maximum basin relief* is the elevation difference between the highest point on the basin perimeter and the basin outlet.

The *basin slope* is a major factor in the overland flow process, and several methods have been developed for determining its average value because of the variation in land–surface slope throughout the basin. In the grid-square method, a grid of uniformly spaced lines is established over a map of the basin, and the slope of a short segment of line normal to the contours is determined at each grid intersection. The basin slope is then calculated as the mean of the resulting local slopes. In the random-coordinate method, the procedure is the same except that the points where the local slope is evaluated are randomly located over the basin (4). Basin slope can also be determined by measuring the total length L_{tot} along contours at contour intervals Δz and calculating the average value:

$$i = \frac{\Delta z L_{\mathrm{tot}}}{A} \tag{6}$$

Computations by this formula are time-consuming, but by introducing a DEM, they become more feasible. However, using a DEM, it is also possible to calculate the slope of each individual square or triangle in a raster or TIN model, respectively, and to use these values to compute the average basin slope.

The peak discharge per unit of basin area is strongly correlated with the mean basin slope. In basins that have a high relief ratio, the lag time and time of hydrograph rise are shorter, and the peak discharge rate is higher (Fig. 2A). Sediment erosion that takes place mainly during peak stream flow is also correlated with basin relief.

Rock and Soil Characteristics. Rock and soil characteristics determine the rate of rainfall infiltration into the soil

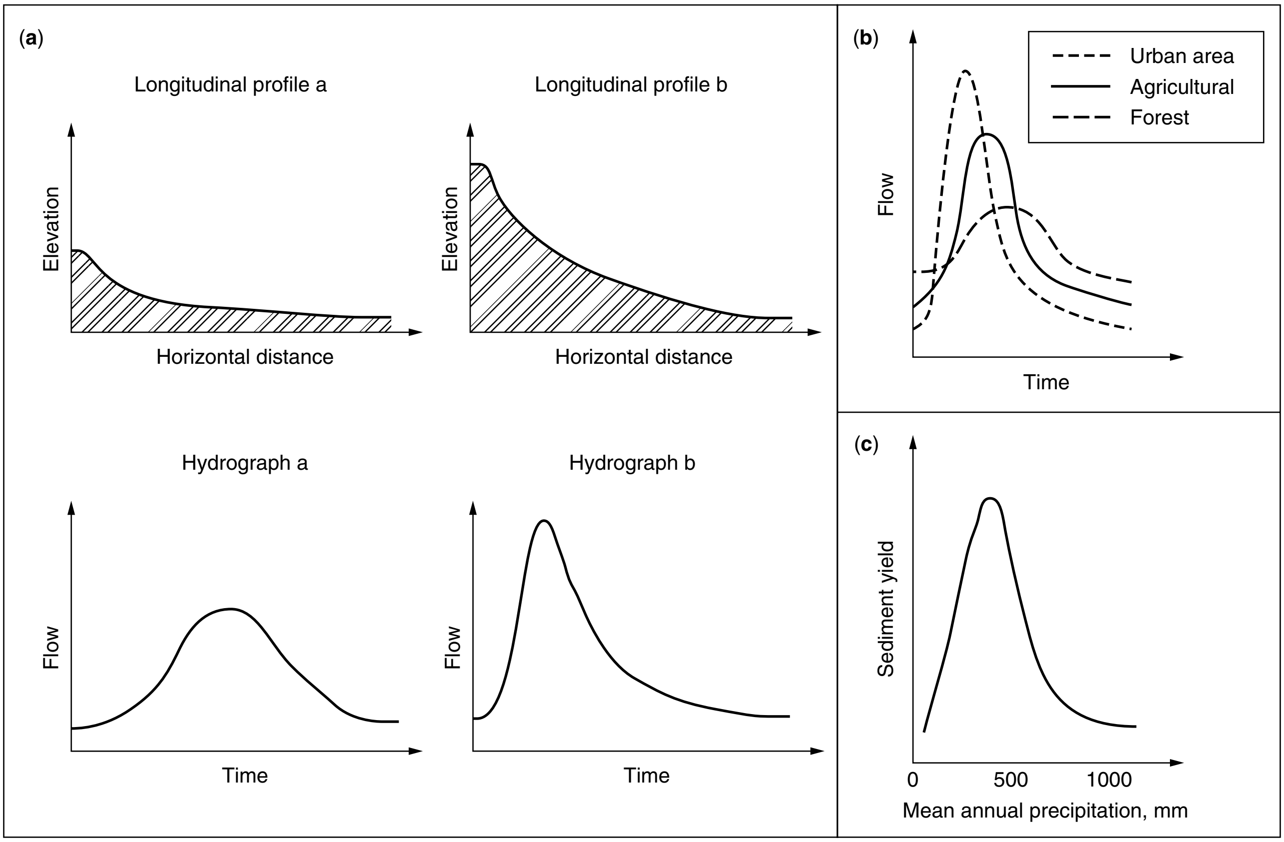

Figure 2. Influence of relief (**a**) and vegetative cover (**b**) on runoff and influence of climatic characteristics (**c**) on sediment yield.

and thus affect runoff. The character of the rock beneath the basin with regard to these processes can be expressed by porosity and permeability. *Porosity* is the percentage volume of voids in a material and is a measure of the amount of water that can be retained by a deposit; *permeability* provides an index of the ability of a material to transmit water and is a function of soil texture, grain size, density of grains, and particle shape. The Soil Conservation Service has classified most soils into four hydrologic soil groups A, B, C, and D, dependent on their infiltration rate (6):

— group A: deep, well-drained sand;
— group B: sandy loam;
— group C: clay loam;
— group D: clay soils of a high swelling potential.

The geology of a basin affects the type and the amount of material available for erosion and transport. The extent of removal and transport of sediment is influenced by the physical and chemical properties of the rock. The vulnerability of soils to weathering can be evaluated by several indexes of soil erodibility which depend upon particle size, moisture content, permeability, dispersion properties of the soil and shear resistance.

Vegetation and Land Use Characteristics. Vegetative cover is a significant factor influencing drainage basin dynamics because it affects the amount of net precipitation that reaches the ground surface through interception, evapotranspiration, and infiltration, and also the water and sediment rate produced by the basin.

Interception losses are given by the sum of precipitation retained on vegetative surfaces such as leaves, stems, and tree trunks. The amount of water captured by vegetation depends on the amount of leaf cover; the type, age and density of vegetative cover; seasonal variations; the amount, duration, and intensity of precipitation; and antecedent moisture conditions.

Vegetative cover also determines losses by evapotranspiration, which is the combination of evaporation from plant surfaces and transpiration of water from plants. Evapotranspiration is related to plant characteristics and environmental conditions, such as atmospheric pressure and moisture, air temperature, wind speed, soil moisture, and light intensity.

The effect of vegetative cover is also displayed by an increase in soil infiltration capacity due to the resistance afforded by vegetation, which slows down overland flow and thus allows the water to have more time to enter the soil surface and flow to the plant roots which make the soil more previous to infiltrating water. The type and density of vegetation, the presence of plant litter on the soil surface, and the depth and density of roots may influence infiltration dynamics.

The vegetative cover influences stream-flow runoff in a basin (3). As the vegetation extent decreases, passing from forested areas to agricultural land and further to urban areas, the range of flows between peak and low flows becomes wider, and the peak discharge is amplified and occurs earlier (Fig. 2B).

Sediment production by erosion depends on land use character, especially cover density. Vegetal cover has a protective influence against erosion processes. A dense cover such as grass reduces the flow velocity and operates simultaneously as a protection covering the underlying soil. Sediment production is maximum from cultivated land, decreases in pasture land, and is minimum in forest and wild land.

Climatic Characteristics. Climatic characteristics govern the amount of water received at the surface. Sediment yield is broadly correlated with mean annual precipitation (7). Maximum values occur in semiarid areas, where the proportion of precipitation available for surface runoff and therefore for erosion is greater. In more arid and more humid areas, sediment production decreases because of the reduced runoff potential and the increase in vegetative cover, respectively (Fig. 2C).

Channel System

Channel Cross Section. The shape and size of alluvial channel cross sections are strongly correlated with stream discharge. Numerous relationships that express the variation in channel characteristics, such as top width, mean depth, mean velocity, slope, with discharge at a particular section and between cross sections have been developed (9). As discharge increases downstream, surface width and mean depth tend to increase, whereas slope decreases. Downstream variations in channel form are influenced by local factors such as type and quantity of sediment load, rock type, superficial deposits, and vegetative cover. Thus, quantitative relationships between hydraulic geometry and discharge may be applied only locally in the same region where the data from which the equations are derived have been obtained, for example, to estimate stream flow from channel measurements in a site where no stream-flow records are available.

Channel Reach

Channel Slope. The slope of the principal drainage channel can be estimated by several methods (Fig. 3A). The simplest method consists of dividing the difference in elevation between the source and the mouth by the length of the channel (Definition 1). The slope can also be determined by calculating the slope of a line drawn on the channel profile through the lowest point such that the area under it equals the area under the profile curve (Definition 2). Channel slope affects flow velocity, travel time, and sediment transport capacity, so it influences water and sediment routing through the basin.

Channel Patterns. River channel patterns have been classified into meandering, braided, and straight (Fig. 3B).

A channel is meandering where it assumes the form of a number of loops or bends. The *sinuosity* is the ratio of channel length to valley length and has an average value of 1.5. Meander wavelength and amplitude range

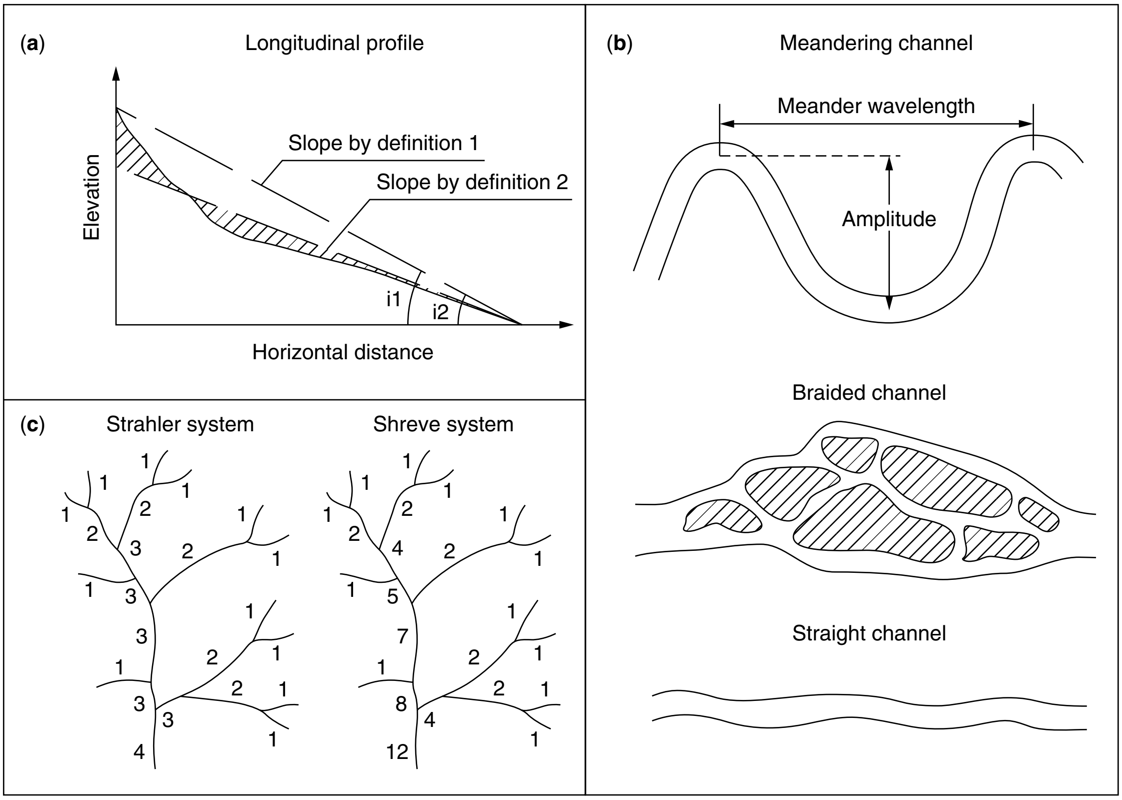

Figure 3. Channel system: (**a**) Channel slope definition; (**b**) channel pattern description; (**c**) stream order designation.

from 7 to 10 and from 3 to 20 times the channel width, respectively (7). Meander wavelength has proved to be strongly correlated with discharge. A channel is braided where the stream bifurcates into two or more intertwined channels separated by islands or bars. Braided channels are characterized by wide, shallow beds of coarse material. The factors that contribute to the development of braided channels are essentially bed-load availability, irregularity of stream flow, intense precipitation which causes high peak discharges, and high slopes (3). Perfectly straight channels are seldom found in nature, so the definition of straight channel has been extended to a channel whose sinuosity is less than 1.25.

Channel Network

Drainage Density. The density of a channel network is the total length of channels per unit area. Values of drainage density observed range from 1 to 248 km/km^2. High density is observed in basins that have weak and impermeable rocks, sparse vegetation, and a steep relief. Network density depends on climatic characteristics, and it varies with mean annual precipitation in the same way as sediment yield because high sediment delivery corresponds to a well-developed channel network. Drainage density is inversely related to precipitation-effectiveness Thornthwaites index P-E which adjusts monthly precipitation by a function of monthly average daily temperature to account for evapotranspiration losses (4). This relation shows the effectiveness of vegetative cover in erosion control. Other basin characteristics such as lithology and land use, and in particular, rock permeability, influence drainage density because they affect the infiltration–runoff rate. Network density is an important factor in the formation of flow. As it increases, peak discharge per unit area increases, and time-to-peak decreases, so that a basin that has a high drainage density is characterized by an efficient drainage system and rapid hydrologic responses.

Stream Order. A channel network may be described by designation of a stream order using with one of the systems developed by Strahler and Shreve to determine the position of a stream in the hierarchy of tributaries (Fig. 3C). In the Strahler system, all fingertip tributaries have order one; two first-order channels joining produce a second-order segment, and so on; if two channels of different order join, the resultant link has the order of the tributary of higher order. In the Shreve system, all fingertip tributaries have magnitude one, and each downstream channel segment has a magnitude equal to the sum of all the first-magnitude segments that are tributary to it. The basin order is a measure of the branching within the basin and is related to the size of channel network and to the amount of stream flow. Stream order is very sensitive to map scale which should always be specified together with the method of ordering used. The Strahler method is the most widely used method today, especially combined with the laws of drainage composition, whereas the Shreve method gives a description of a drainage network closer to physical reality with regard to stream discharge.

Horton's laws of drainage composition relate stream numbers, stream lengths, and drainage areas to stream order by simple geometric relationships (4). The bifurcation ratio R_B is defined by

$$R_B = \frac{N_u}{N_{u+1}} \tag{7}$$

where N_u is the number of streams of a given order and N_{u+1} is the number of the next higher order. Bifurcation ratios tend to be constant in a given basin and assume characteristic values between three and five. The law of stream numbers is expressed by

$$N_u = R_B^{k-u} \tag{8}$$

where k is the highest order within the basin. Similarly, the laws of stream lengths and drainage areas indicate a geometric progression of order:

$$\overline{L}_u = \overline{L}_1 R_L^{u-1}, \overline{A}_u = \overline{A}_1 R_A^{u-1}, R_L = \frac{\overline{L}_u}{\overline{L}_{u-1}}, R_A = \frac{\overline{A}_u}{\overline{A}_{u-1}} \tag{9}$$

where L_u is the mean length of streams of order u, A_u is the mean area of basins of order u, R_L is the length ratio, and R_A is the area ratio, whose ranges of values are 1.5–3.5 and 3–6, respectively.

An alternative model has been proposed by Shreve using the probabilistic–topological approach. It assumes that in the absence of climatic and geologic constraints, channel networks evolve at random; all topologically distinct channel networks (TDCN) of a given magnitude are equally probable (5). Shreve used the term "link" for stream segments between nodes (sources, junctions, or outlet) and distinguished exterior links, which extend from a source to the first junction, from interior links, which connect two successive junctions or the last junction with the outlet. Furthermore, he assumed that exterior and interior link lengths are random variables whose separate probability distributions are independent of position within the basin. Shreve showed that the most probable network in a random population closely conforms to Horton's law of stream numbers.

More recently a new approach has been developed for analyzing drainage network composition, which shows that river basins have a self-similar organization, that is, the same shape is found at another place in another size. This property can be interpreted by Mandelbrot's fractal theory which introduces fractals as irregular geometric shapes that have identical structure on all scales (8).

BIBLIOGRAPHY

1. Petrie, G. and Kennie, T.J.M. (1991). *Terrain Modelling in Surveying and Civil Engineering*. McGraw-Hill, New York.
2. Goodchild, M.F., Parks, B.O., and Steyaert, L.F. (1993). *Geographic Information Systems and Environmental Modelling*. Oxford University Press, New York.
3. Gregory, K.J. and Walling, D.E. (1973). *Drainage Basin Form and Process: A Geomorphological Approach*. Edward Arnold, London.

4. Strahler, A.N. (1964). Quantitative geomorphology of drainage basins and channel networks. In: *Handbook of Applied Hydrology*. V.T. Chow (Ed.). McGraw-Hill, New York, Section 4-II.
5. Morisawa, M. (1985). *Rivers Form and Process*. Longman, London.
6. Soil Conservation Service, Section 4. (1972). *Hydrology, SCS National Engineering Handbook, U.S. Department of Agriculture*. U.S. Govt. Printing Office, Washington, DC.
7. Leopold, L.B., Wolman, M.G., and Miller, J.P. (1964). *Fluvial Processes in Geomorphology*. Dover, New York.
8. Rodriguez-Iturbe, I. and Rinaldo, A. (1997). *Fractal River Basins: Chance and Self-Organization*. Cambridge University Press, New York.

RIVER BASIN PLANNING AND COORDINATION

Peder Hjorth
Lund University
Lund, Sweden

A river basin is a functional entity where water is the integrating factor. What happens in one location will impact any downstream location. Sometimes there are upstream impacts as well, for instance, backwater effects upstream of a dam. Thus, water management must be coordinated and concerted to result in desired outcomes in different parts of a drainage basin. To this end, there is a need for effective communication across institutional and disciplinary boundaries.

As stated in Agenda 21 [a wide-ranging blueprint for action to achieve sustainable development worldwide). Adopted by United Nations Conference on Environment and Development (UNCED), Rio de Janireo, June 3–14, 1992.], freshwater is a unitary resource. Long-term development of global freshwater requires holistic management of resources and recognition of the interconnectedness of the elements related to freshwater and freshwater quality. Failures in these respects have resulted so that there are few regions of the world that are still exempt from problems of loss of potential sources of freshwater supply, degraded water quality, and pollution of surface and groundwater sources.

In chapter 18 of the Agenda, current wisdom concerning water management is summarized, and it is stated that freshwater resources are an essential component of the earth's hydrosphere and an indispensable part of all terrestrial ecosystems. Water is needed in all aspects of life, and the general objective of water management is to make certain that adequate supplies of water of good quality are maintained for the entire population of this planet, while preserving the hydrologic, biological, and chemical functions of ecosystems.

It is pointed out that the extent to which water resources development contributes to economic productivity and social well-being is not usually appreciated. Thus, a spreading scarcity, gradual destruction, and aggravated pollution of freshwater is allowed to go on without much attention from responsible agents/institutions. A major problem is the fragmentation of responsibilities for water resources development among sectoral agencies such as those concerned with water supply and sanitation, agriculture, industry, urban development, hydropower generation, inland fisheries, transportation, recreation, and low- and flatland management.

In addition, different, but inherently linked aspects of catchment management, such as water quality, water quantity, and the processes of erosion and deposition, are frequently managed by different institutions. Often, there is also a division of responsibility for surface water issues and groundwater issues, respectively. These aspects need to be coordinated if we are to create the synergy needed to achieve the desired water status. Coordination is needed at both the political and technical levels.

An important starting point is that water should be managed at the river basin level and be seen as an integral part of the ecosystem, a natural resource, and a social and economic good, whose quantity and quality determine the nature of its use.

Efficient coordination is essential for the dynamic, interactive, iterative, and multisectoral approach called for. The holistic management of freshwater as a finite and vulnerable resource urgently calls for integrating sectoral water plans and programs within the framework of national economic and social policy. Various levels of complexity should be recognized when coordinating activities by different agents or organizations. Simplistic solutions should be avoided.

Four basic requirements can be identified:

1. management of the whole resource;
2. clear links between catchment functions;
3. emphasis on prevention and mitigation rather than responses;
4. adoption of multifunctional win–win solutions.

Water management should take into account long-term planning needs as well as those with narrower horizons. The long-term need should be related to the principle of sustainability; the more short-term objectives would be focused on the prevention and mitigation of water-related hazards.

The International Conference on Freshwater in Bonn, 2001, stated that water management arrangements should take account of climate variability and expand the capacity to identify trends, manage risks, and adapt to hazards such as floods and droughts. Anticipation and prevention are more effective and less expensive than having to react to emergencies. Early warning systems should become an integral part of water resources development and planning. It was also stated that systematic efforts are needed to revive and learn from traditional and indigenous technologies (for example, rainwater harvesting) around the world.

Closer links should be established between development and disaster management systems. Exposure to flood risks should be minimized through wetland and watershed restoration, better land use planning, and improved drainage. The greater fluctuation in resource availability associated with presumed climatic changes is causing concern, as many regions experience increasingly severe flooding and/or drought problems. The solutions to some of these problems can be sought locally, whereas others have to be dealt with in international forums.

Thus, the scale of issues must be understood to allow for a prioritized approach. Often, issues are perceived differently by different groups of people, and myths about their severity are legion. For instance, within basins embracing several nation states, the attitudes of those states in the headwaters is typically different from the attitudes of the states in the lower reaches of the river. Headwater states do not wish to be held to guarantees of water quantity or of water quality demanded by their downstream neighbors. Downstream countries are, very naturally, concerned about the prospects of any upstream dam construction which may affect the regime or total quantity of flow, especially by major diversions of water into or out of the basin (1).

With regard to existing uncertainty and lack of precise knowledge, it is essential to monitor continuously and evaluate what happens to water quality and quantity at key points and to make sure that the information gained is made available to those who need it, when they need it.

Strategies and programs need to be tied to indicators that enable actors to assess progress. As we are swamped by a wealth of information, a major problem is that we have difficulties in discriminating between information, useful knowledge, and science. Therefore, it is essential to have few but relevant indicators of high explanatory value. Jiménez-Beltrán (2) claims that efficient indicators should

1. show development over time and be policy relevant (there should be an explanation why a specific indicator has been developed);
2. be few in number and people should get used to their presentation and understand the message;
3. be closely linked to objectives to become efficient tools in decision-making.

In the United Kingdom, local Agenda 21 efforts show that indicators work best when they are developed in participatory ways. In communities of all kinds, "everyday experts" are getting involved in the monitoring and evaluation process, and results are starting to change policy (3).

On a more conceptual level, Brown (4) suggests that to be efficient, indicators should be

1. holistic—measuring if the catchment is worked with as a whole;
2. cooperative—measuring the degree of working together with shared responsibilities;
3. composite—measuring the degree of inclusion of the full range of diverse elements and the degree to which elements have been arranged to fit together;
4. coordinated—assessing the mechanisms for continuing cooperation; and
5. long term—assessing the durability of relationships and processes and assessing the progress toward a shared vision.

Basic to all activities is a baseline assessment. We have to know where we start from before we can decide how to go to the desired conditions. Such assessment, including the identification of potential sources of freshwater supply, comprises the determination of sources, extent, dependability, and quality of water resources and of the human activities that affect those resources. It is a prerequisite for evaluating the possibilities of water resources development. There is, however, growing concern that, in a time when more precise and reliable information is needed, hydrologic services and related bodies are less able to provide this information, especially concerning groundwater quantity and quality.

In most places, there is a need to strengthen the institutional arrangements for water assessment. This is not just about efficient collection of data, but also about processing, storage, retrieval, and dissemination to users of information about the quality and quantity of available water resources at the level of catchments and groundwater aquifers.

Prior to such strengthening, it is necessary to prepare catalogues of the water resources information collected and/or held by government services, the private sector, educational institutes, consultants, local water-use associations, and others. It is essential to build on what exists by strengthening and coordinating the different bodies involved in the collection, storage, and analysis of the relevant data.

There is a need to strengthen the technical support by installing additional observation networks, developing systems for data storage, and systems for data retrieval. However, strengthening of the human resource base in numbers and skills seems to be equally important.

BIBLIOGRAPHY

1. Young, G.J., Dooge, J.C.I., and Rodda, J.C. (1994). *Global Water Resource Issues*. Cambridge University Press, Cambridge, UK.
2. Jiménez-Beltrán, D. (2001). *Implementing the EU Sustainable Development Strategy.* http://org.eea.eu.int/documents/speeches/Speech%20Brussels%20July%209%2C%202001-en.
3. IDS. (1998). *Participatory Monitoring & Evaluation: Learning from Change.* IDS Policy Briefing, Issue 12, November 1998. http://www.ids.ac.uk/ids/bookshop/briefs/brief12.html.
4. Brown, V. (Ed.). (1997). *Managing for Local Sustainability: Policy, Problem-Solving, Practice and Place*. Commonwealth of Australia.

BIOACCUMULATION

Zhi-Qing Lin
Southern Illinois University at Edwardsville
Edwardsville, Illinois

Bioaccumulation is a process by which a chemical from the environment accumulates in an organism's body over time. Bioaccumulation occurs when there is a concentration increase of the chemical in the tissues of an organism.

Bioaccumulation results from mass equilibrium where the uptake of a chemical into the body of an organism exceeds the elimination of the chemical from the body. As such, a chemical that is eliminated immediately after uptake will not be bioaccumulated in an organism.

Uptake of a chemical means entrance of the chemical into an organism primarily from water or soil, directly or through consumption of foods that contain the chemical. A chemical can be taken up through roots or leaf surfaces in plants and by breathing, ingesting, or absorbing through the body surface in animals. Uptake of a chemical in plant cells can occur against its electrochemical potential gradient through the process of active transport. For example, selenate ($SeO_4{}^{2-}$) and arsenate ($AsO_4{}^{2-}$) are taken up via sulfate and phosphate transporters in the root plasma membrane, respectively.

The elimination of a chemical from an organism includes excretion and/or metabolism of the chemical. Elimination is the primary process by which an organism can reduce its chemical burden. Excretion is the removal of toxicants or excess chemicals from the body of plants or animals. A chemical taken up into the body of an animal can be eliminated by urination and defecation, whereas salt-tolerant halophytes can excrete sodium and other ions from their leaf surfaces.

A chemical can also be eliminated from an organism by transforming the original chemical into a new chemical, a process termed metabolism. Inorganic selenate can be taken up and biologically reduced or methylated into organic selenium compounds that can be volatilized into the atmosphere (1). Metabolism can also break down a large organic compound into small molecules or eventually mineralize the compounds into water, CO_2, and nutrients. Metabolism of an organic compound by an organism is termed biodegradation.

An organism's ability to take up or eliminate a chemical varies among species and depends on physiochemical properties of the chemical. Chemicals that are water-soluble will be taken up and readily eliminated by an organism. Persistent organic compounds are generally not water-soluble and are difficult for organisms to metabolize or degrade. These compounds tend to have a greater potential to accumulate in fat or lipid tissues.

Bioaccumulation can sometimes substantially increase the concentration of a chemical in the tissues of an organism. When the concentration of a chemical becomes higher in an organism than in the environment (e.g., water, soil, or air), the bioaccumulation process is specifically termed bioconcentration. The extent of bioconcentration can be expressed by the bioconcentration factor ($BF = C_{org}/C_{env}$), which is a ratio of the concentration of a chemical in an organism (C_{org}) to the concentration of the same chemical in the living environment (C_{env}). If a plant species can accumulate an exceptionally high level of a metal (e.g., about 1% dry weight for zinc and manganese or 0.1% for copper and nickel in shoots) without having harmful effects, the species is termed a metal hyperaccumulator. *Astragalus bisulcatus* is a selenium hyperaccumulator that can accumulate about 0.5% (dry weight) of selenium in shoots.

A chemical that is bioconcentrated/bioaccumulated is not homogeneously distributed in the body of an organism. Distribution of an accumulated chemical varies significantly among tissues/organs, species, and chemicals. For instance, fat-soluble chemicals (e.g., polychlorinated biphenyls—PCBs) accumulate primarily in fat, cadmium in kidneys, and mercury in livers. Such organ-specific accumulation of a chemical may impose a potential hazard to an organism because the chemical can reach a critical body burden in those target organs. This can even occur in organisms that are exposed only to low levels of a toxicant in the environment.

Plant roots generally accumulate higher concentrations of metals (e.g., lead and manganese) than shoots. The distribution of a chemical in a plant can be described by its translocation factor ($TF = C_{shoot}/C_{root}$), the ratio of the concentration of the chemical in shoots (C_{shoot}) to the concentration in roots (C_{root}). A larger translocation factor means that a greater proportion of the chemical is accumulated in shoots compared to that in the roots. Similarly, the ratio of the shoot concentration (C_{shoot}) to the soil concentration (C_{soil}) of a chemical is termed the phytoextraction coefficient ($FC = C_{shoot}/C_{soil}$). Phytoextraction is one of the phytoremediation approaches using plants to remove toxicants from contaminated waters and soils. Plants with high phytoextraction coefficients can remove a large amount of pollutant by harvesting the chemical-laden shoots at contaminated sites.

One of the important environmental concerns associated with bioaccumulation is biomagnification, whereby the concentration of a chemical increases at each higher trophic level through a food chain. Biomagnification can result in a concentration increase of two to three orders of magnitude between two trophic levels. Therefore, due to biomagnification through the entire food chain, the concentration of a chemical in top predators may reach tens of thousands times the concentration in the water column. For example, through a water–plankton/algae–fish food chain at the Kesterson reservoir in Central California, selenium in fish was biomagnified 35,000 times from selenium-contaminated drainage water. Very high concentrations of selenium result in reproductive deformities and death of fish and waterfowl (2).

The term bioaccumulation is sometimes used to refer to the total amount of a chemical accumulated in the body of an organism. The amount of a chemical accumulated in an organ or tissue depends on both the chemical concentration and the actual biomass in which the concentration is determined. For example, a higher bioconcentration of selenium in *Astragalus bisulcatus* does not necessarily mean that a greater total amount of selenium from the environment can be accumulated in the plant because the species grows slowly and the total biomass production is small. To overcome such biological limitations, scientists are currently applying new biotechnology to combine the genome of a tolerant, slow-growing, selenium hyperaccumulator, such as *Astragalus*, with that of a less tolerant, but fast growing nonhyperaccumulator, for example, *Brassica juncea*, to develop a somatic hybrid plant that could increase the phytoremediation potential (3).

BIBLIOGRAPHY

1. Lin, Z-Q. and Terry, N. (2003). Selenium removal by constructed wetlands: Quantitative importance of biological volatilization in the treatment of selenium-laden agricultural drainage water. *Environ. Sci. Technol.* **3**: 606–615.
2. Skorupa, J.P. and Ohlendorf, H.M. (1991). Contaminants in drainage water and avian risk thresholds. In: *The Economy and Management of Water and Drainage*. A. Dinar and D. Zilberman (Eds.). Kluwer Academic, Boston, MA, pp. 345–368.
3. Banuelos, G.S., Lin, Z-Q., Wu, L., and Terry, N. (2002). Phytoremediation of selenium-contaminated soils and waters: Fundamentals and future prospects. *Rev. Environ. Health* **4**: 291–306.

BIOTIC INTEGRITY INDEX TO EVALUATE WATER RESOURCE INTEGRITY IN FRESHWATER SYSTEMS

Lizhu Wang
Brian B.Weigel
Wisconsin Department of
Natural Resources
Monona, Wisconsin

Many varied and complex environmental problems, particularly row crop and grazing agriculture, urbanization, highway building, timber harvesting, and water projects such as dams, water withdrawals, and stream channel modifications have seriously affected the health of freshwater resources in many parts of the world. The health of aquatic resources can be measured in many ways. Among them, biotic (or biological) integrity is a concept most commonly used by the public, biologists, resource managers, and policy makers to measure the status of aquatic systems.

Biotic integrity is "the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of natural habitat of the region." This concept was first proposed by Frey (1) and later applied by Karr and Dudley (2) in discussing an ecological perspective on water quality goals.

Aquatic systems of high biotic integrity have a biological community in which composition, structure, and function have not been seriously altered by human activities. Such systems can withstand or rapidly recover from some perturbations imposed by natural environmental processes and survive many major disruptions induced by humans. Aquatic systems that lack integrity are often degraded and when further perturbed by natural or human-induced events are likely to change rapidly to an even more undesirable status.

Biotic integrity should not be confused with biotic diversity. Biotic diversity (or biodiversity) refers to "the variety and variability among living organisms and the ecological complexes in which they occur" (3). More thorough definitions include multiple organization levels (e.g., genes, species, and ecosystems) as fundamental components of biodiversity to distinguish it from the much simpler concept of species diversity (4).

Fundamental differences exist between biotic integrity and diversity. Biotic integrity refers to a system's wholeness, including the presence of all appropriate elements and occurrence of all processes at appropriate rates, whereas diversity is a collective property of system elements (5). Biotic integrity is also associated with conditions under which the biotic communities evolved, but diversity may not necessarily measure such an aspect. For example, introducing exotic species or genes from distant populations may increase local diversity, but it reduces biotic integrity. Some aquatic management activities for increasing harvestable products or services of economic value may not necessarily correlate with biotic integrity. A large amount of harvestable products in some systems may indicate low biotic integrity (6).

FACTORS AFFECTING BIOTIC INTEGRITY

Biotic communities in aquatic systems have evolved over millions of years. The environmental conditions under which biotic communities evolved can be classified into five major groups (6). The energy source group includes the type, amount, and particle size of organic materials entering an aquatic system from the surrounding terrestrial zones and internal primary production. The water quality group includes temperature, turbidity, dissolved oxygen, pH, nutrients, heavy metals, and natural and synthetic inorganic or organic chemicals. The habitat group includes substrate type, water depth, water current velocity, habitat complexity, and reproduction and hiding places. The flow regime group includes water volume and temporal distribution of floods and low flows. The biotic interactions group includes competition, predation, disease, and parasitism. Regional climate, landscape topography, geology, soil type, and watershed land-cover type largely determine the factors within these five groups. Without human disturbance, the factors in aquatic systems and their watershed conditions are in dynamic equilibrium, which operates under natural weather cycles. As such, many of these factors have natural patterns of temporal and spatial variation.

When watershed lands are converted from forest, prairie, or wetland to agriculture and urbanization, the dynamic equilibrium between the elements in an aquatic system and its watershed is broken, causing degradation in aquatic resources. Degradation is intensified by aggressive farming practices such as overapplying fertilizers, pesticides, and herbicides to improve crop yields, concentrating livestock in high densities to increase production efficiency, and channelizing streams and draining wetlands to expand agricultural acreage. Urban development further degrades aquatic resources by increasing toxicants, nutrients, and storm-water runoff, in turn, causing more frequent and severe flooding, accelerated channel erosion, and an altered stream channel and substrate composition. Agricultural and urban land-use impacts cause major changes in aquatic communities and thus, biotic integrity within the system.

WHY USE BIOTIC INTEGRITY TO MEASURE WATER RESOURCE HEALTH

The combined influences of human-induced stressors are often difficult to detect, particularly across diffuse urban areas or agricultural landscapes. For example, water quality impacts are often associated with increased surface runoff from rainstorms or snowmelt. These episodic events can be assessed only by continuous monitoring, which is expensive and time-consuming. Even from continuous monitoring, runoff volumes and characteristics tend to be highly variable, and long time-series of data are required to evaluate trends. If changes in concentrations of a pollutant are detected, its effects on the health of the aquatic ecosystem often remain unclear.

Chemical and physical criteria are insufficient to protect water resources. The earliest anthropogenic threats to water resources were often associated with human health, especially disease-causing organisms and oxygen-demand wastes. Early emphasis was on controlling these contaminants in urban areas, where effluents exceeded the natural waste assimilating capacities of water, by using chemical and physical indicators as primary regulatory tools to protect water resources; this has eliminated or greatly reduced the known-source problems. However, water resource quality and quantity continued to decline despite massive governmental regulatory efforts. Nonpoint-source pollution inputs, those originating from diffuse areas such as farm fields and parking lots, were not extensively reduced by chemical and physical criteria alone. For example, in 1986, nonpoint-source pollution affected 65% of the impaired stream length, 76% of the impaired lake area, and 45% of the impaired estuary area in the United States (7).

Biotic assemblages represent the end point of the combined influences of human-induced perturbations (8). As such, biotic indicators can provide measures of water resource health, although diffuse disturbances within a watershed, from agriculture or urbanization, have many physical and chemical effects on a waterbody. Direct measures of biotic assemblages, such as fish, aquatic insects, and algae, are cost-effective and ecologically relevant ways to assess human impacts on aquatic resources. Because the biotic assemblages consist of a variety of species that have different life histories, sensitivities to degradation, and functions in the ecosystem, they respond to a range of human disturbances. A few appropriate samples of the assemblage can provide unique insight into the condition of the aquatic system and the causes of degradation. Biotic assemblages are accurate and easily measured indicators of the overall quality or health of water resources (9).

BIOTIC INDICATORS

Numerous assemblage-level indicators are available for assessing human impacts on aquatic systems. The most commonly used and effective indicators for water resource health can be grouped into three categories: tolerance, taxa richness and diversity, and reproducing and feeding ecology. Most indicators were developed empirically, based on observed responses to a gradient of degradation, rather than experimental tests of sensitivity to a specific stressor. This approach is appropriate because human activities invariably have multiple interactive impacts on aquatic ecosystems that are often diffuse and cumulative. Although different types of indicators tend to be most sensitive to different environmental impacts, the best indicators are sensitive to most or all stressors that are typical of watershed disturbance.

Tolerance measures are based on the documented relative sensitivity of particular taxa to disturbance. The presence or abundance of taxa that have a known degree of sensitivity to a particular stress is a measure of the degree of that stress within the system. For example, Hilsenhoff (10) developed an aquatic macroinvertebrate index to assess organic pollution or the addition of excessive nutrients to a waterbody, which is a common consequence of agriculture and urbanization. Each taxon was assigned a tolerance value based on its sensitivity to organic pollution. The index was the weighted average, based on relative abundance, of the tolerance values of all of the taxa collected in a semiquantitative sample. If most of the taxa and individuals present were sensitive to organic pollution, then the index score was good, but if the assemblages were dominated by taxa tolerant of organic pollution, then the index score was poor. By their nature, tolerance measures tend to be relatively narrow in their sensitivity, but they have been strong indicators of water resource health.

Taxa richness is based upon the premise that the number of taxa is related to the amount and type of human disturbance. Diversity measures are based on the premise that both the number of taxa and evenness of the distribution of individuals among taxa are related to the amount and type of stress. Usually, richness and diversity are inversely proportional to stress, although this may not hold for “cold-water” streams where fish assemblages are dominated by salmonids. In warm-water streams, strong negative relations have been reported between watershed disturbance and both fish taxa richness and diversity, where the number of native, sensitive taxa decreased as stress increased. High natural variability in diversity scores has caused diversity measures to lose popularity, but taxa richness continues to be one of the most reliable and accepted indicators worldwide.

Reproducing and feeding ecology measures categorize organisms into groups that use similar reproduction and food resources. Multiple classification systems exist for classifying organisms by reproducing and feeding ecology, but most are based on combinations of what the animals eat, where they forage, how they acquire food, and what substrate they spawn on. Documenting the relative abundance of organisms that rely on each reproducing or feeding class reflects the habitat and food web conditions in which they live. Such measures also reflect the availability of essential life-cycle elements, energy flow, and nutrient dynamics, as well as the balance of these components between a water body and its surrounding terrestrial environment. Energy production and flow are difficult to measure directly, but biota provide surrogate measures.

MULTIMETRIC INDEXES OF BIOTIC INTEGRITY

A multimetric index of biotic integrity (IBI) combines a variety of different indicators, termed metrics, including tolerance, taxa richness, and reproducing and feeding ecology, into a single index that reflects structural, compositional, and functional attributes of an assemblage (9,11). By incorporating several different metrics, IBIs are sensitive to a wide range of human disturbances. Potential metrics are selected based on knowledge of aquatic systems, life history of the organisms, literature reviews, and historical data. The candidate metrics are evaluated and eliminated if they are not robust or show little relationship to human disturbances for the particular region and waterbody type of interest. The ability to interpret individual metrics is retained, even though the metrics are combined into a multimetric index.

An IBI must be regionally calibrated to take into account and correct for natural variation in assemblages owing to the biogeography or ecological conditions of the waterbodies. For example, it is well established that in the absence of human perturbations, larger streams tend to have more fish species than smaller streams and certain river basins have richer fish faunas than other basins because of geomorphic history. It is initially difficult to discern which factors are influencing the biotic metrics, natural environmental variability or human disturbance. Waterbodies that suffer relatively little human disturbance yet are influenced by the same regional set of natural environmental conditions (e.g., geology, climate, and waterbody type or size) are used as references to provide a baseline for comparison with waterbodies of unknown condition. Well-designed IBIs take natural differences into account and have metrics based on standards that are specific to the region and type of waterbodies for which they are intended.

Individual metric scores and final IBI values are relative to the set of reference waterbodies for a particular waterbody type and region. Test waterbody conditions are compared with a set of reference conditions to indicate their deviation from optimal conditions. Ranked scores, for example 5, 3, and 1, are assigned to each metric according to whether its value approximates, deviates somewhat from, or deviates strongly from the values expected at the reference condition. The overall index value is the sum of the ranked scores from all of the individual metrics. The highest possible score indicates a waterbody that is comparable to those that have the lowest stress within that region, and those of reduced quality have lower values. Based on the final value, one can qualitatively classify a water body as having good, fair, or poor health.

Fish Index of Biotic Integrity

Fish IBI is used worldwide and is especially popular in the United States. The IBI was first developed by Karr (12) for stream fish in the central United States. This version of IBI consists of 12 metrics that reflect basic structural and functional characteristics of fish assemblages: (1) number of native species; (2) number of darter species; (3) number of sucker species; (4) number of sunfish (excluding green sunfish) species; (5) number of individuals in sample; (6) number of intolerant species; (7) percent green sunfish; (8) percent hybrid individuals; (9) percent omnivores; (10) percent insectivorous cyprinids; (11) percent top carnivores; and (12) percent disease, tumors, fin damage, and other anomalies.

Because IBI is regionally specific, many versions of fish IBIs have been developed to meet the needs in several parts of the United States, Canada, France, Mexico, Australia, Belgium, Guinea, India, Namibia, and Venezuela. Many versions have also been adapted to specific types of waterbodies, including headwater streams, wadeable streams, nonwadeable rivers, cold-water streams, cool-water streams, lakes, wetlands, and estuaries.

During the past 20 years, fish IBIs have been broadly used to document impacts of watershed land uses on the health of streams, rivers, and lakes. One example of the studies that comprehensively evaluated the impacts of urbanization on streams using IBI was conducted in Wisconsin, United States (13). In this study, fish communities in 54 warm-water and 38 cold-water streams were sampled by electrofishing stream segments that were 35 times the mean wetted stream width or at least 100 m. Fish were sampled between late May and early August, when low stream flows facilitated sampling effectiveness and large-scale seasonal fish movements were unlikely. During sampling, efforts were made to collect all fish observed, and all captured fish were identified, counted, and then returned to the stream alive. The collected data were then entered into a computerized database, and an IBI value was calculated for each stream. The cold-water and warm-water IBI versions used here were specifically developed for the study region. The scores for both versions ranged from 0–100; higher scores indicated better stream health. One simple way to evaluate impacts of urban land use on stream health is to plot percentages of watershed urban land use against the IBI scores (Fig. 1). Some streams had very healthy conditions, and others did not. These plots indicated that at low percentages of urban land use, the stream conditions were varied. At low levels of urbanization (less than 12%), urban influences were weaker than other stressors, such as agriculture. As the proportion of urban land use increased in the watersheds, the fish IBI values decreased sharply. When land use exceeded certain levels, the stream health was consistently poor. Such a relation indicated that as urban land use increased, urbanization dominated over all other land uses and played the strongest role in influencing stream health.

Karr (12) noted several advantages of using fish IBIs to measure stream health: (1) life history information is extensive for most fish species; (2) fish communities generally include a range of species that represent a variety of trophic levels and include foods of both aquatic and terrestrial origins; (3) fish are relatively easy to identify; (4) both acute and cumulative effects can be evaluated; and (5) the general public can relate conditions of fish community to water resource health. The disadvantages of using fish as indicators include (1) seasonal and diel fish movements, (2) intensive field sampling effort, and (3) modifying the IBI for specific geographic regions and waterbody types.

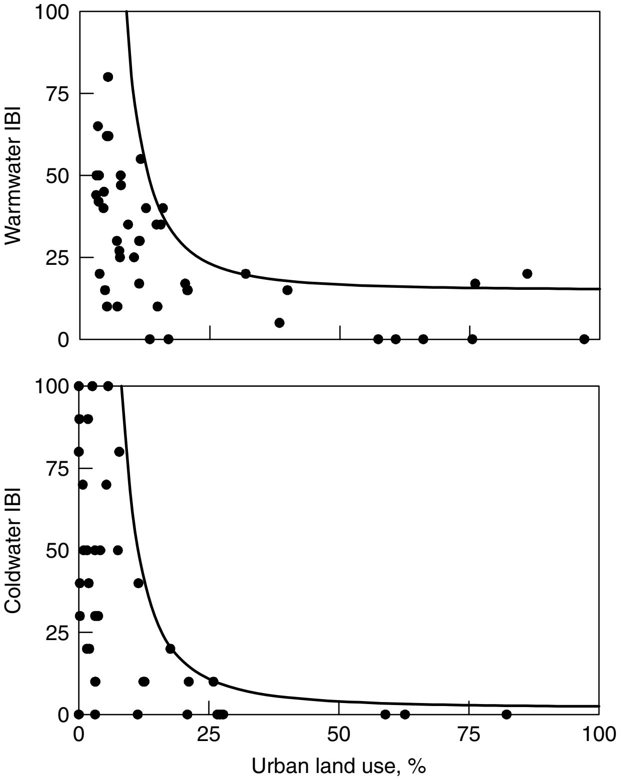

Figure 1. Relationships between watershed urban land use and fish index of biotic integrity for 54 warm-water and 38 cold-water streams in Minnesota and Wisconsin, United States. The land uses in the study watersheds are predominated by either agriculture or urban. At low level urbanization, urban influences of fish are weaker than other stressors, such as agriculture. As percentages of urban land in a watershed increase, fish IBI values decrease sharply and when urban land exceeds a certain level, the stream health is consistently poor.

Macroinvertebrate Index of Biotic Integrity

The macroinvertebrate IBI concept was extended from fish IBI and has been widely used in streams, rivers, and wetlands in North America. One of the early macroinvertebrate IBIs was developed for streams in Ohio, United States (14). This IBI, called the Invertebrate Community Index (ICI), consisted of 10 compositional and structural community metrics: (1) total number of taxa; (2) number of mayfly taxa; (3) number of caddisfly taxa; (4) number of dipteran taxa; (5) % mayfly composition; (6) percent caddisfly composition; (7) tribe tanytarsini midge composition; (8) percent other dipteran and noninsect composition; (9) percent tolerant organisms; and (10) number of Ephemeroptera–Plecoptera–Trichoptera taxa. Metric scoring criteria were developed through a quantitative calibration process in which reference values were plotted against a log-transformed watershed area.

Macroinvertebrate IBIs have also been broadly used to document human impacts on the health of streams, rivers, and wetlands in North America and other parts of the world. One example of such a study was also conducted in Wisconsin, United States, in which stream macroinvertebrate IBI values were empirically linked with watershed and local stressors (15). In this study, streams were stratified by region to account for natural environmental variation. Standardized methods were used to kick-sample macroinvertebrates from 397 independent streams and identify the samples in the laboratory. Kick-sampling is a method for dislodging macroinvertebrates by kicking the substrate and letting the current wash the organisms into a net; it is typically done in riffle or run habitats that have coarse substrates.

Quantifying human disturbance entailed using standardized methods to characterize the watershed land cover and assess the local in-stream and adjacent terrestrial (riparian) habitat upstream of the sampling point. Land cover categories included urban, row crop agriculture, forage crop agriculture, wetland, forest, and open water. Local habitat characterizations included assessment of the riparian vegetation, bank erodability, livestock grazing, macroinvertebrate food sources, channel morphology, and streambed. These local and watershed measures of human disturbance were summed to give an overall environmental condition value to each site. Likewise, several macroinvertebrate metrics were combined into a multimetric IBI to give an overall biotic integrity value to each site. A plot from 209 streams in the central–southeast region of Wisconsin shows how values of environmental condition were related to macroinvertebrate IBI values (Fig. 2). Streams of excellent environmental condition typically had relatively high proportions of forests, wetlands, and open water and a low percentage of urban land cover on the watershed scale. On the local scale, these excellent streams had a variety of food sources, relatively undisturbed riparian conditions, and a heterogeneous habitat structure, including riffles, meanders, and woody debris. The macroinvertebrate assemblages

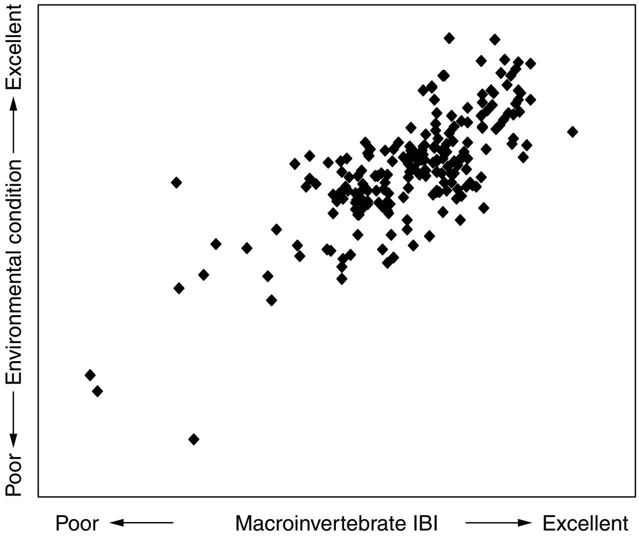

Figure 2. Macroinvertebrate index of biotic integrity tailored to small streams in central and southeast Wisconsin, United States ($n = 209$). The poorest streams had high proportions of urbanization or agriculture in their watersheds and local habitat stress. Macroinvertebrates in the poorest streams were tolerant to organic pollution, had no individuals from the relatively sensitive insect orders Ephemeroptera, Plecoptera, and Trichoptera, and had low taxa richness.

at these excellent streams had high species richness, were intolerant of organic pollution, and were comprised predominantly of intolerant taxa from the insect orders Ephemeroptera, Plecoptera, and Trichoptera, as opposed to Diptera. The empirical model indicated that extensive agriculture and urbanization impacted the biological community strongly and significantly.

There are several advantages of using macroinvertebrates to evaluate the health of aquatic systems. Macroinvertebrate assemblages consist of a variety of species whose different life histories are sensitive to a multitude of degradation types and play an important functional role in stream ecosystems. The assemblages can respond to a range of human-induced effects on streams. Macroinvertebrate assemblages represent the end point of the combined influences of hydrology, physical habitat, and water quality and it has been shown, respond predictably to these factors within specific geographical regions (9). Similar to the fish IBI, the disadvantages of using the macroinvertebrate IBI to measure the health of aquatic systems include (1) seasonal variation in their relative abundance, (2) intensive efforts for field sampling and taxa identification, and (3) modifying the IBI for particular geographic regions and waterbody types.

Periphyton Index of Biotic Integrity

The multimetric periphyton IBI was developed to measure the health of aquatic systems. Periphyton are algae attached to hard substrates like the rocks that comprise a streambed. The states of Montana and Kentucky were among the first to develop a multimetric periphyton IBI in the United States. Bahls' (16) periphyton IBI for Montana streams included three metrics based on soft-bodied taxa and four metrics based on diatom taxa. Metrics for soft-bodied taxa included dominant phylum, indicator taxa, and number of genera. Diatom metrics included the Shannon–Wiener diversity index, a pollution index, a siltation index, and a similarity index for comparison with a reference condition. Kentucky's stream periphyton IBI consisted of diatom species richness, species diversity, percent community similarity to reference sites, a pollution tolerance index, and percent sensitive species (17).

Periphyton respond to a variety of pollutants and can be used to diagnose the probable causes of health impairment in aquatic systems. Some periphyton can be sensitive to pollutants that other organisms tolerate relatively well, partly because periphyton cannot avoid pollutants due to their sedentary nature. Periphyton occur in most aquatic habitats, and typically have greater taxonomic richness than fish and macroinvertebrate. Although sampling periphyton is relatively easy, identification must be in a laboratory and it is labor-intensive. Typically, periphyton have rapid reproduction rates and short life cycles, and thus, they respond quickly to perturbation. However, this quick response to stress appears to increase their natural variation, which can make the detection of changes resulting from human perturbation more difficult.

PREDICTIVE MODELING

The River Invertebrate Prediction and Classification System (RIVPACS) and Australian River Assessment System (AusRivAs) are predictive models that quantify river health as the degree to which a waterbody supports the biota that would be expected there in the absence of human disturbance. These models were initially developed in the United Kingdom (18) and later modified for Australia (19), other parts of Europe, and North America.

In the process of assessing aquatic health, an empirical model of reference waterbodies is created that incorporates the natural environmental factors that are unlikely to be affected by human activities, yet influence the taxa there. This model predicts the taxa that occur under least-impacted, reference conditions. The observed biotic taxa (O) at a test waterbody are compared with the taxa that the predictive model expects (E) to find in the absence of human stress, and the deviation is expressed as a ratio (O/E). Impairment is inferred if the O/E values measured at a test site fall outside the error inherent in the predicted E and estimated O/E.

In the multimetric approach, the predictive model method can provide site-specific prediction of the composition of biota in test waterbodies, the assessment requires no assumptions regarding the specific types of stress that affect biota, and it uses independent data for matching test conditions with reference conditions. The predictive modeling approach is difficult to apply in regions where not many undisturbed reference waterbodies can be found. This approach has not been tested for organisms other than benthic macroinvertebrates.

SUMMARY

Biotic indexes for evaluating the health of freshwater systems are used increasingly as they are becoming more standardized and cost-effective, and presenting their results is easier. Many human-induced changes in the physical and chemical properties of water resources are difficult to detect because of their temporal and spatial variation. Even when water resources are monitored using broad temporal and spatial coverage, physical and chemical measurements are insufficient to protect water resources because many human influences cannot be measured by using a physicochemical approach. Furthermore, biotic assemblages represent the end point of the combined influences of human disturbance on aquatic environments, and thus, biotic indexes provide a measure of overall water resource health.

Many aquatic organisms, including fish, benthic macroinvertebrates, and periphyton, can be used as indicators of aquatic ecosystem health. During the past 20 years, many IBI versions for fish, benthic macroinvertebrates, and periphyton have been developed worldwide to meet the needs of localized climate, zoogeological zone, and thermal and hydrologic regimes. All multimetric IBIs share several features. The overall index score is the sum of the scores of several individual metrics. Each metric represents a different attribute of the structure, composition, or function of the biotic assemblage. Natural environmental or biogeographic factors that influence these attributes are taken into account in applying the index. Metrics are chosen and calibrated largely based on empirical data. Metric scores from references, those representing the least impacted waterbodies within a region, provide a standard for comparison. Each metric is sensitive to one or more types of environmental degradation. Each of the fish, benthic macroinvertebrate, and periphyton IBIs has its strengths and weaknesses. We need to use IBIs in combination with physical habitat and water chemistry assessments to establish criteria that direct human activities toward improvement and protection of water resource health.

The predictive modeling approach, such as RIVPACS, is different from the multimetric method. Instead of using multiple metrics, it uses the ratio between observed and expected for the occurring probability of aquatic organisms. However, both predictive model and multimetric approaches use regional reference sites and achieve similar goals in assessing the health of water resources.

BIBLIOGRAPHY

1. Frey, D.G. (1975). Biological integrity of water: an historical approach. R.K. Ballentine and L.J. Guarraia (Eds.). *The Integrity of Water: A Symposium*. US Environmental Protection Agency, Washington, DC., pp. 127–139.
2. Karr, J.R. and Dudley, D.R. (1981). Ecological perspective on water quality goals. *Environmental Management* **5**: 55–68.
3. OTA (Office of Technology Assessment). (1987). *Technologies to Maintain Biological Diversity*. Congress of the United States, OTA-F-330, Washington, DC.
4. Noss, R.F. (1990). Indicators for monitoring biodiversity: a hierarchical approach. *Conservation Biology* **4**: 355–364.
5. Angermeier, P.L. and Karr, J.R. (1994). Biological integrity versus biological diversity as policy directives. *BioScience* **44**: 690–697.
6. Karr, J.R., Fausch, K.D., Angermeier, P.L., Yant, P.R., and Schlosser, I.J. (1986). *Assessing Biological Integrity in Running Waters: A Method and its Rationale*. Illinois Natural History Survey, Special Publication 5, Champaign, IL.
7. GAO (General Accounting Office). (1989). *Water Pollution: More EPA Action Needed to Improve the Quality of Heavily Polluted Waters*. Report to the Chairman, Subcommittee on Regulation and Business Opportunities, Committee on Small Business, House of Representative. GAO/RCED-89-38.
8. Rosenberg, D.M. and Resh, V.H. (Eds.). (1993). *Freshwater Biomonitoring and Benthic Macroinvertebrates*. Chapman and Hall, New York.
9. Karr, J.R. and Chu, E.W. (1999). *Restoring Life in Running Water: Better Biological Monitoring*. Island Press, Washington, DC.
10. Hilsenhoff, W.L. (1987). An improved biotic index of organic stream pollution. *Great Lakes Entomologist* **20**: 31–39.
11. Barbour, M.T., Stribling, J.B., and Karr, J.R. (1995). Multimetric approach for establishing biocriteria and measuring biological condition. W.S. Davis and T.P. Simon (Eds.). *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*. Lewis Publishers, Boca Raton, FL, pp. 63–77.
12. Karr, J.R. (1981). Assessment of biotic integrity using fish communities. *Fisheries* **6**(6): 21–27.
13. Wang, L., Lyons, J., Kanehl, P., and Bannerman, R. (2001). Impacts of urbanization on stream habitat and fish across multiple spatial scales. *Environmental Management* **28**: 255–266.
14. Ohio EPA (Environmental Protection Agency). (1987). Biological criteria for the protection of aquatic life: Volume II: *Users manual for biological field assessment of Ohio surface waters. Surface Waters Section, Division of Water Quality Monitoring and Assessment*. Ohio Environmental Protection Agency, Columbus, OH.
15. Weigel, B.M. *In review*. Development of stream macroinvertebrate models that predict land cover and local stressors in Wisconsin. *Journal of the North American Benthological Society*.
16. Bahls, L.L. (1993). *Periphyton Bioassessment Methods for Montana Streams*. Water Quality Bureau, Department of Health and Environmental Sciences, Helena, MT.
17. Kentucky DEP (Department of Environmental Protection). (1993). *Methods for Assessing Biological Integrity of Surface Waters*. Kentucky Department of Environmental Protection, Division of Water, Frankfort, KY.
18. Wright, J.F., Furse, M.T., and Armitage, P.D. (1993). RIVPACS: a technique for evaluating the biological quality of rivers in the UK. *European Water Pollution Control* **3**: 15–25.
19. Smith, M.J. et al. (1999). AusRivAs: using macroinvertebrates to assess ecological condition of rivers in western Australia. *Journal of Freshwater Biology* **41**: 269–282.

REVERSAL OF THE CHICAGO RIVER

Richard Lanyon
Metropolitan Water
Reclamation District of Greater
Chicago
Chicago, Illinois

INTRODUCTION

Through the end of the nineteenth century, the rapidly growing city of Chicago was plagued with frequent epidemics of waterborne diseases. The city's death rate was one of the highest in the world. Despite the widespread installation and use of sewers beginning in 1855, the problem persisted. In fact, the sewers, while providing some adequate local drainage, only moved the public health hazard to the Chicago River and, inevitably, to Lake Michigan. As the lake also supplied the city with its drinking water, a vicious cycle existed. Successive efforts were undertaken to move the water intakes farther from shore. But, in time, the plume of pollution would reach farther out into the lake, enveloping the intakes. At this time, acceptable technology for the treatment of large quantities of potable water or sewage was nonexistent.

By the 1880s, it became clear that the only viable solution was to discharge the sewage across a subtle subcontinental divide 10 miles distant from the Lake Michigan shoreline. There, the Des Plaines River, a tributary of the Illinois and Mississippi Rivers, could be reached. A deep channel from Chicago to Joliet would be necessary. A new governmental entity, a sanitary district, was created and charged with the job of building the channel. When its work was finished, a mere decade later, this remarkable project resulted in the excavation of 42,230,000 cubic yards of rock and soil, construction of 460,000 cubic yards of masonry for channel walls and bridge abutments, and the erection of 31 bridges. The total cost was $33.5 million. Today, the 28-mile constructed channel is called the Chicago Sanitary and Ship Canal, although to its builders, it was referred to as the Main Channel.

COMMISSION FORMED

The City of Chicago Common Council responded to citizen demands in January 1886 by taking its first official action to solve the problem of the Chicago River: It formed the Commission on Drainage and Water Supply. The commission's charge was to outline a solution to the problem within 1 year, in time for legislative action by the Illinois General Assembly. The commission issued its 36-page preliminary report in January 1887, which presented three alternatives for sewage disposal: (1) discharge sewage into Lake Michigan away from the city, (2) dispose of sewage on land, and (3) discharge sewage into the Des Plaines River.

Cost estimates were based on a projected population of 2.5 million, roughly three times the then current population. Discharge to the lake was estimated at $37 million, disposal on land at $58 million, and discharge to the Des Plaines River at $28 million. Discharge to the Des Plaines River was the recommended solution based on cost and a belief that it was a more reliable technology.

The commission set the parameters for design of a large channel from Chicago to Joliet, using as input the plans of others as well as their own creative ideas and the results of their exhaustive investigations. Determining adequate capacity was based on meeting three fundamental needs of the city and region: storm flow, sanitation, and navigation. The commission recommended a channel cross-section area of 3600 square feet and a velocity of 3 feet per second. The resulting discharge capacity of 10,000 cubic feet per second was intended to serve a population of 2.5 million people. It was the commission's opinion that the channel capacity would prevent backflows into the lake and protect the water intake cribs located two miles offshore. The commission estimated the cost to build the channel at $20.3 to $24.5 million. It also recognized the potential for water power development because of the steep descent of the Des Plaines River near Lockport (see Fig. 1).

The scope of the recommendations and the size of the undertaking required the state legislature to authorize the formation of a new unit of local government. The city of Chicago was in debt to the legal limit, and the task at hand would require considerable financial resources. A new entity encompassing a larger area could borrow anew, could have a larger tax base than the city, and would have powers beyond the city limits. A legislative commission began work in May 1887, writing what would become an authorizing statute: "An act to create sanitary districts and to remove obstructions in the Des Plaines and Illinois rivers," effective July 1, 1889. The Governor of Illinois rapidly approved, opening the way for the sanitary district to be formed by referendum.

Typical of Chicago's political muscle, within 6 months, the proponents had established the boundaries of the new Sanitary District of Chicago (SDC), conducted the referendum, which passed by a whopping 70,958 to 242,

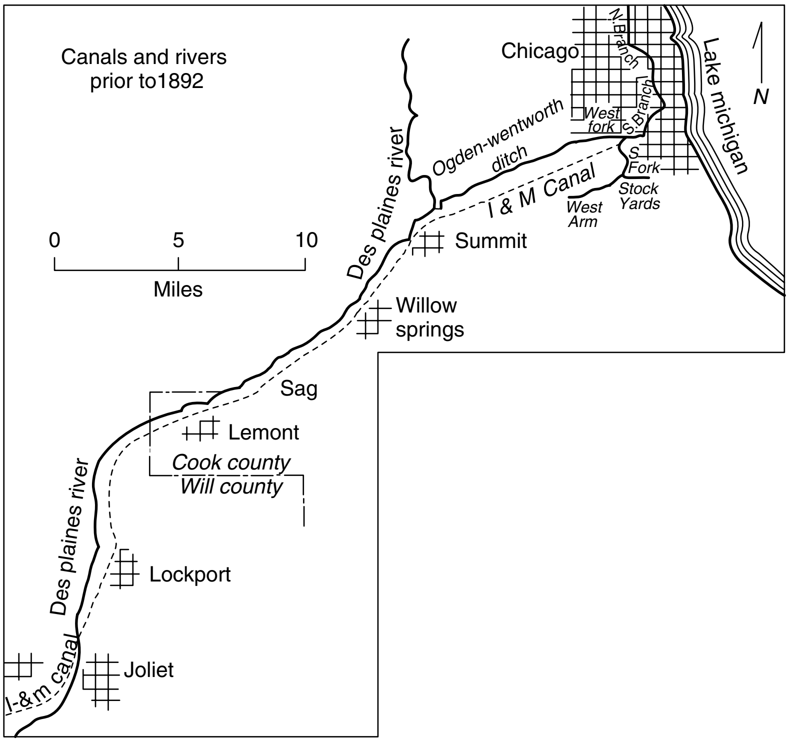

Figure 1. Canals and rivers prior to 1892.

elected nine trustees, and had them sworn in. The SDC was officially organized by January 1890. However, the new entity had to wait out two legal challenges, which went up to the Illinois Supreme Court. By June 1890, the court affirmed the authority of the SDC. Now the SDC could borrow money, approve a tax levy, and begin work.

MAIN CHANNEL WORK BEGINS

The SDC's first step was to select a Chief Engineer and direct him to present to the Board of Trustees (Board) not less than four routes of a channel to cross the 8 miles between the South Branch and Summit and to make investigations of the Des Plaines River valley. The intent was to build a channel and two pumping stations: one to lift the waters of the South Branch into the channel and the other to lift the water out of the channel into the Des Plaines River. It was also anticipated that the Des Plaines River would have to be enlarged to accept the additional flow. However, over a year went by with little progress as the board was divided and a succession of chief engineers did not survive the dissention. The third chief engineer was able to put together a plan that the board, with several new members, could agree on in January 1892.

In the spring of 1892, the board focused on getting the 14.8-mile rock section of the Main Channel designed and under construction. It was assumed that excavation in rock would be more difficult and time consuming than excavation in the earth. June 1893 saw the appointment of Isham Randolph as Chief Engineer (the fifth one), and this finally brought stability to the SDC Engineering Department as he remained in this position to 1907.

The rock section began at Willow Springs and ended at Lockport (see Fig. 2). New surveys and subsurface investigations were constantly revising knowledge of the route. The channel route was west of the I&M Canal and east of the Des Plaines River. The surface of Niagaran dolomite was very close to or formed the bed of the Des Plaines River throughout much of the reach from Sag to Lockport. From Willow Springs to Sag, the rock surface was high enough that a channel with a discharge capacity of 10,000 cubic feet per second (cfs) was required by the act. Throughout the rock section, the channel is rectangular with a width of 160 feet and a channel depth of 24 feet. In June, 14 contracts were awarded and work began on September 3, 1892, with a ceremony at the Cook-Will County Line. Some of the rock excavation contractors proceeded apace and their work was completed by 1895, 1 year earlier than anticipated. Other contractors were plagued with various problems, and the work was not completed until 1899.

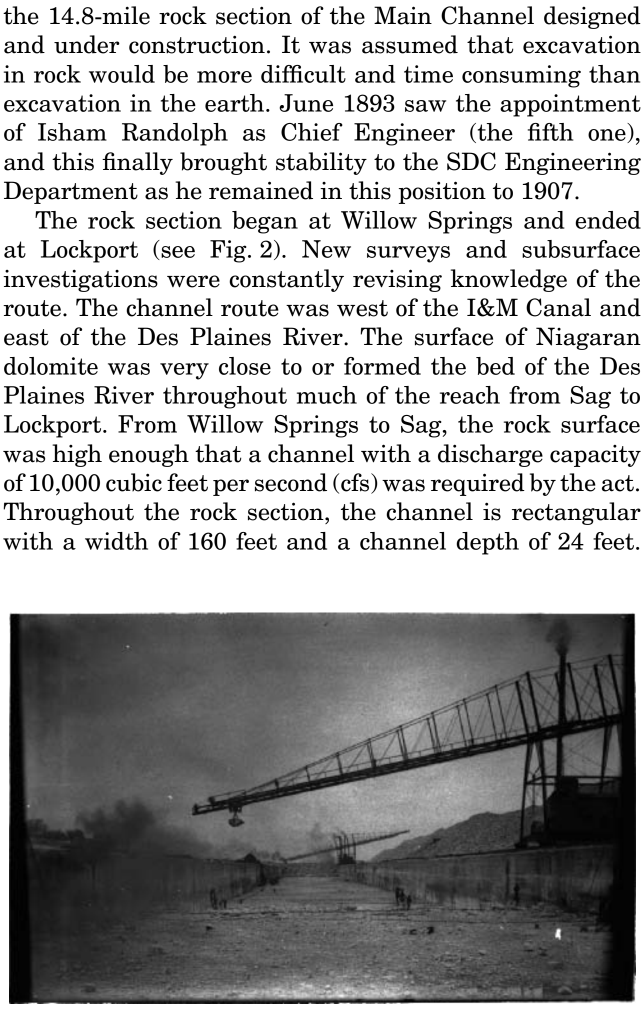

Figure 2. Excavation of the rock section.

To make room for the Main Channel between Summit and Lockport, the eastward meanders of the Des Plaines River had to be relocated; thus an aggregate of 13 miles of river diversion channels was constructed as well as a continuous levee separating the Main Channel and river. Keeping the river out of the construction area where the Main Channel was being constructed was a continuing problem for the contractors during flood periods.

Next to go under contract was the 6.2-mile earth and rock section from Summit to Willow Springs. Again, the route follows the I&M Canal and Des Plaines River. At Summit, the route of the Des Plaines River turns north, separating from the routes of the Main Channel and I&M Canal, which turn toward Chicago. Six contracts were awarded in January 1893 for channel excavation, and all work was completed by 1899. The channel was trapezoidal in cross section with a bottom width of 210 feet, side slopes of 2 horizontal to 1 vertical, and a depth of 24 feet. Because rock was encountered in this section, the act required the channel to have a discharge capacity of 10,000 cfs.

The earth section of the Main Channel extended from Summit to the West Fork, beginning just west of Robey Street (now Damen Avenue) in Chicago. As no rock was encountered in the channel cross section, the act allowed construction of a channel with a capacity of only 5,000 cfs. The act contemplated that as the population grew, the channel capacity would be increased to 10,000 cfs. This eventually happened in 1912. This lesser capacity channel had a bottom width of 110 feet, side slopes of 2 horizontal on 1 vertical, and a depth of 24 feet. This reach was notable because of the many railroad crossings. As a result of negotiations with railroads for one of the crossings, the SDC was obligated to also construct the one-third mile-long Collateral Channel connecting the Main Channel to the West Fork along the alignment of what is now Albany Avenue. The short bypass channel relieved the railroads of having two movable bridges within one-half mile of track right-of-way.

The 7.2-mile earth section reach was divided into eight contracts, the first six of which were awarded in December 1893. The last two contracts, awarded in May 1894, provided for dredging and transporting the spoil by scow to the lakefront to be used as fill for the creation of what is now Grant Park. Other spoil was deposited in what is now Douglas Park and used as fill for local streets and boulevards. Because of the delayed construction of a major railroad crossing, dredged channels extended from the West Fork to Western Avenue along the route of the Main Channel and to the Main Channel north embankment along the Collateral Channel. Work on these eight contracts was completed in either 1898 or 1899.

To control the discharge of water from the Main Channel, two more contracts were awarded in 1895 and 1896 for construction of the Lockport Controlling Works.

These works consisted of seven vertical gates, each 30 feet wide and 20 feet high, and a 160-foot long sector-type dam, called the Bear Trap Dam. The dam could be lowered to allow flow over its top, providing for sensitive discharge control. The gates were raised to provide for rapid increases in discharge over a short time period. These two types of control were necessary because the end of the Main Channel was 35 miles from Lake Michigan and the lake level could change rapidly on short notice. These control structures were among the largest in the world at the time, comparable with controls on the outlet of Lake Superior and on the Ohio River.

THE JOLIET PROJECT

Near Lockport, the Des Plaines River began a relatively steep descent to Joliet, at some places flowing over exposed dolomite. Once in Joliet, the river was joined by the I&M Canal where several successive dams created navigation or power pools. Planning for a channel through Joliet would require more study and dealing with the I&M Canal Commissioners and the City of Joliet. Complicating the matter was the physical setting north of Joliet. The river channel meandered from east to west across the valley floor; the I&M Canal and two railroads were on the east side of the valley, and several industries with water power developments were along the river.

The confluence of the new channel and the I&M Canal required modifications to the tow path, locks, and dams, but the canal commissioners were slow to come to terms with the SDC. The SDC proceeded with work and was sued by the canal commissioners. Three contracts were awarded in early 1898 for the 5.1 miles of work. Because of the lawsuit and other contract problems, the work was not completed until 1901. However, enlargement of the Des Plaines River channel capacity was completed by late 1899.

SOUTH BRANCH IMPROVEMENTS

By August 1895, 28 miles of the Main Channel were under contract, and rock excavation was nearly complete near Lemont. The SDC needed to improve the capacity of the Chicago River and the South Branch so that it could deliver the flow of water from Lake Michigan as required by the act. The river had many bends and constrictive bridge openings, was shallow in spots, and was always busy with boat traffic. Fortunately, no time would be consumed in debate over the route.

The U.S. Army Corps of Engineers (Corps) had plans to improve the Chicago River, so the SDC concerned itself with improvement of the 5-mile reach of the South Branch between Lake and Robey streets. All work in the South Branch was subject to permits issued by the Corps, but the SDC would often begin work before a permit was issued. One contract was awarded in May 1897 for removal and replacement of dock walls and dredging to pass 5000 cfs in a channel 200 feet wide and 20 feet deep. Other contracts were awarded in 1898 for replacement of two center pier bridges with restrictive openings and a large bypass conduit around a restrictive bridge opening that could not be enlarged. All work was sufficiently completed by 1899 to provide for the design capacity.

The act made reference in many locations to a navigable waterway, maximum velocities, and minimum depths and widths, all of which defined the conditions for safe commercial navigation. However, the act made no reference to bridges. Lacking statutory definition, the matter of bridges was discretionary to the SDC, which caused much debate and division among the members of the board. As there was no navigation lock at Lockport, navigation on the Main Channel was not an immediate priority. The bridges were put under contract late in the 1890s, and all substructure work was completed by late 1899 before the Main Channel was placed in service in January 1900. The SDC built or funded the construction of 31 bridges to effect the reversal of the Chicago River. The 13 bridges over the Main Channel were designed and constructed to eventually be made movable to allow for passage of boats.

REVERSING THE FLOW

To place the empty 28-mile-long Main Channel into operation, water would need to be added slowly so as not to cause damage by rapidly rising water levels or swift currents. Despite not having specific approval from the Governor, water was let in at the Chicago end beginning on January 2, 1900, through a wooden flume in the earth dike across the south end of the Collateral Channel. The filling continued to January 14 when the water level in the Main Channel reached the water level in the West Fork. The next day, the earth dam across the Main Channel west of Western Avenue was cleared away by dredges and the waters on each side came together. After 13 days of filling, the water level came to rest, to wait for the Governor's approval to discharge at Lockport.

A special commission had been appointed by the Governor to inspect the work and advise on satisfactory completion of the work. The SDC vowed to wait for the Governor's approval before releasing water from the Main Channel at Lockport. The Board and commission members traveled to Lockport on January 17, 1900, to be at the Lockport Controlling Works when the approval came. The Governor's approval was received by telegram, and the Bear Trap Dam was lowered slightly below the water level to allow a thin sheet of water to flow over its top (see Fig. 3). After a brief ceremony, the valves controlling the dam were opened and the massive 160-foot-long dam disappeared beneath the water. A torrent of water rushed out of the Main Channel over the dam toward the Des Plaines River. On this chilly day in the first month of the new century, slightly more than 10 years after passage of the authorizing act and after more than 7 years of construction, the Main Channel was now in operation to save Chicago from its own waste.

The reversal of the Chicago River has been supplemented with other works, and the diversion of water from out of the Great Lakes basin has been the subject of extensive litigation, but the reversal of the river has never been interrupted.

Figure 3. A total of 4250 cubit feet per second passes over the Bear Trap Dam.

READING LIST

Chicago Daily News. August 3, 17, and 28, 1885.

Currey, S.J. (1912). *Chicago: Its History and Its Builders*. Vol. III. S. J. Clarke Publishing Co., Chicago, IL.

SDC. (1928). *Engineering Works*.

Journal of the House of Representatives, 35th General Assembly, State of Illinois. (May 26, 1887).

Larson, J. (1979). *Those Army Engineers*. Chicago District, U.S. Army Corps of Engineers.

Laws of the State of Illinois, 1989, p. 125.

Preliminary Report of the Commission on Drainage and Water Supply, January 1887.

Proceedings of the Board of Trustees of the Sanitary District of Chicago, 1890 through 1901.

Report of the Citizens' Association, August 27, 1885.

FLOOD CONTROL IN THE YELLOW RIVER BASIN IN CHINA

Rongchao Li
Delft University of Technology
Delft, The Netherlands

GENERAL INFORMATION ON THE YELLOW RIVER

The Yellow River (see Fig. 1), the second longest river in China originates in the Yuguzonglie Basin in the Tibet Plateau in west China and enters the Bohai Sea in the east. It is 5464 km long, has a water level difference of 4480 m, and the basin area is 795,000 km^2, 8.28% of China. In the entire Yellow River Basin, 98 million people are settled.

The most important water resource in the Yellow River Basin is precipitation. The annual precipitation is concentrated in the period from July until October. The average annual runoff of the Yellow River is 58 billion m^3 (1987). About 10 to 20% of the runoff is discharged from March until June, and 60% is discharged in the flood season from July until October. The interannual runoff distribution is also not uniform. The ratio between the maximum and the minimum value can be up to 3.4. Since 1919, a dry period that was longer than 5 years occurred twice. The annual average runoff during these dry periods was 39 billion m^3. The total amount of groundwater in the basin is 40 billion m^3 (excluding the interior basin). The exploitable amount of groundwater is 12 billion m^3. The distribution of groundwater is not uniform in the basin (1).

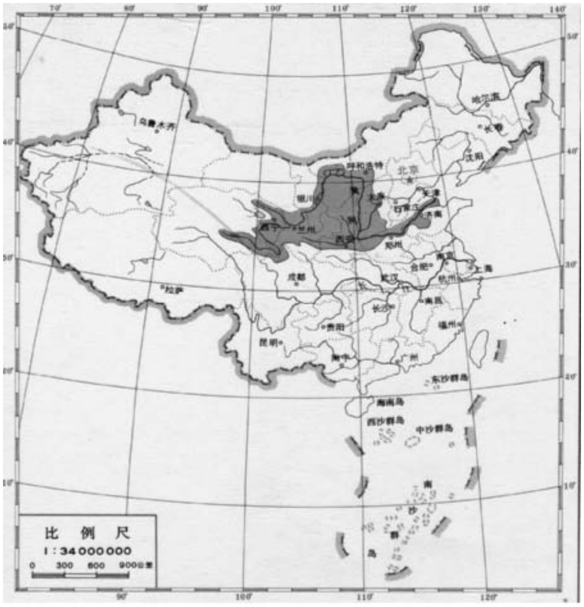

Figure 1. Location of the Yellow River Basin in China.

GENERAL CONDITIONS FOR FLOODING AND SEDIMENTATION

The frequent occurrence of extensive flood disasters is related to its geography (high in the west and low in the east) and to its uneven temporal and spatial distribution of precipitation. As recorded since the third millennium B.C., the Yellow River has overflowed its dikes 1590 times and changed its course 26 times. Nine of those were major course changes. The range of the course divagation covers an area of 250,000 km^2. Thousands of people were killed in many floods. The cycle of flooding and death earned the Yellow River the name "China's Sorrow."

Flooding in the Yellow River Basin

There are two kind of floods in the Yellow River: storm-caused floods that occur between July and October and ice-jam floods that normally occur in February (1). The main floods of the Yellow River occur in July and August. These are peak floods of short duration that rise and fall quickly. Historical investigations show that the largest flood occurred in 1843; the peak discharge was 36,000 m^3/s. The maximum flood measured occurred in 1958; the peak discharge was 22,300 m^3/s at Huayuankou; see Fig. 2 (1). This flood is used to determine the design level of the lower reach. The design level at Huayuankou hydrologic station is 22,000 m^3/s, and the probability is once in 1250 years. The construction of the Xiaolangdi dam is not included in

Figure 2. Main hydrology station in the Yellow River Basin (*Source*: Yellow River Water Conservancy Commission).

this design level. Table 1 shows the total flood frequency analysis for the lower reach of the Yellow River.

In spring, ice-jam floods often occur between Huayuankou and the estuary. The riverbed slope is shallow in this part, so the velocity of the water is low. The flow direction from west to east changes from a low latitude to a higher latitude at the Bohai Sea. This results in a higher temperature at Huayuankou than in the estuary. The average winter temperature near the estuary is 3.4 °C lower than at Huayuankou. For this reason, in 80% of the years, the river freezes, and because of this, in most of these years, ice-jam floods happen. According to the statistics, between 1950 and 1983, there were 29 years in which the river froze. In 1951 and 1955, serious ice floods occurred. The length of the frozen river was 550 km. and 623 km. respectively. In history, the ice-jam floods frequently broke dikes. According to rough statistics between 1883 and 1936, ice-jam floods broke dikes in 21 years.

The extreme floods measured at Huayuankou are from different geographical sources, and these floods also have different characteristics. A classification has been made for the different floods based on historical floods and for a few floods shown in Table 2. The classification is as follows:

- The inflow upstream of the Sanmenxia reservoir and comparatively small inflow downstream of the Sanmenxia reservoir mainly compose the flood. This type of flood has a high flood peak, a large flood discharge, and high sediment concentration.
- The flood is composed mainly of the inflow downstream of the Sanmenxia reservoir. This type of flood has a rapid flood rise, a high flood peak, low sediment concentration, and a short forecast period.
- Half the flood is composed of the inflow upstream of the Sanmenxia reservoir and half of the inflow downstream of the Sanmenxia reservoir. This type of flood has a small flood peak, a long duration, and low sediment concentration.

Sediment in the Yellow River Basin

The majority of the sediment sources in the basin are situated in the upper and the middle reaches. In the upper reach, the average sediment concentration is only 6 kg/m^3, and the average sediment transport is 142 million tons. In the middle reach, the river flows through a loess plateau. Due to the fine grain sizes and the restricted vegetation in the loess plateau, soil erosion can cause a huge sediment load, especially during storms. During a storm, the sediment concentration in the river can rise to more than 500 kg/m^3 (1), as illustrated in Fig. 3.

One of the important characteristics of the Yellow River is that water and sediment originate in different regions. In the upper reach, the annual sediment input is 8.7%, whereas the annual runoff is 54% of the annual input. In the middle reach, the annual runoff is 36%, whereas the sediment input is 89%. From Xiaolangdi to the Bohai Sea, the lower reach, the surface area is limited, and the annual runoff depends mostly on the tributaries, the

Table 1. Flood Frequency Analysis at Huayuankou[a]

Item	Unit	Average	$p = .01\%$	$P = 0.1\%$	$P = 1.0\%$
Flood peak	m^3/s	9780	55000	42300	29200
5d flood discharge	b m^3	26.5	12.5	9.84	7.13
12d flood discharge	b m^3	53.5	20.1	16.4	12.5
45d flood discharge	b m^3	153	41.7	35.8	29.4

[a] Reference 1.

Table 2. Composition of the Floods in Huayuankou[a]

	Huayuankou			Sanmenxia		
	Year	Discharge, m^3/s	12-d Discharge, billion m^3	Discharge, m^3/s	Huayuankou, m^3/s	12-d Discharge, billion m^3
Upstream of Sanmenxia	1843	33,000	13.6	36,000	30,800	11.0
	1933	20,400	10.1	22,000	18,500	9.18
Downstream of Sanmenxia	1761	32,000	12.0		6000	5.0
	1958	22,300	8.9		6400	5.15
Up- and down stream of Sanmenxia	1957	13,000	6.6		5700	43.1

[a] Reference 1.

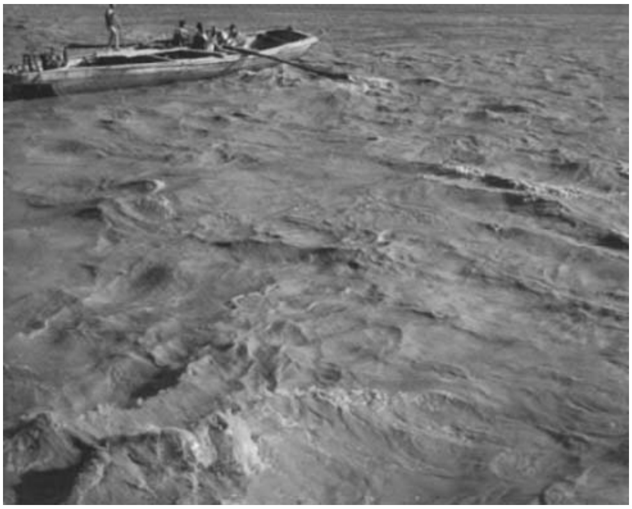

Figure 3. Yellow River: High sediment concentration flow (2).

Yilouhe and the Qinhe. The annual runoff of the lower reach is 10%, and the sediment input is 2% of the annual input. Due to severe soil loss in the middle reach, on average, 1.6 billion ton of sediment enters the river channel at Huayuankou annually, of which about 1.2 billion ton is carried to the estuary region, leaving behind a substantial amount to contribute to the silting of the river channel (1). This results in high deposition in the riverbed of the lower reach. Due to this, the riverbed fills up, on average, 0.1 meter every year. By the time, the Yellow River reaches Kaifeng, it runs 10 m above the surrounding plain. Some Ming and Song dynasties' artifacts 500 to 800 years old lie under 5 to 8 m of silt at Kaifeng City. The present riverbed above its floodplain is about the same as the old course that was breached and abandoned in 1855 (http://www-geology.ucdavis.edu/GEL115/115CHXXYellow.html). Therefore, nowadays, the Yellow River is called a "suspended" river. In Fig. 4, the *Q*/*h* relation for different years at Huayuankou is shown. This figure shows the impact of the decreasing flood conveyance due to the rise of the riverbed. Even for a normal peak flood of 5000 m^3/s the water level rose about 2 meters in the period 1958–1996. It is expected that the process of bed level raising will continue in the near future.

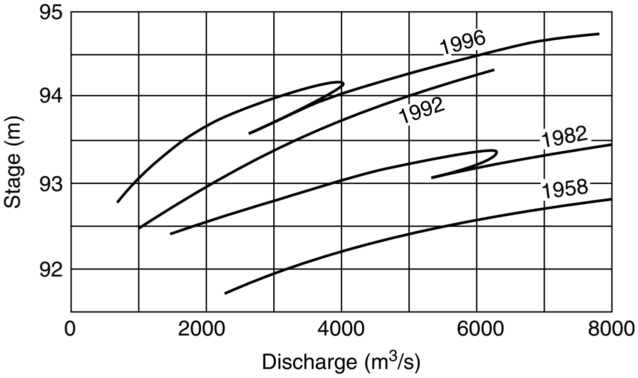

Figure 4. Water level-discharge relation for three floods (3).

FLOOD CONTROL IN THE YR BASIN

Historical Flood Management in the Yellow River

As the most sediment-laden river of the world, the extreme sediment load of the river causes problems in water resources management and flood protection (1). The first inhabitants of the Yellow River floodplains and the surrounding areas lived with floods. They escaped from flood inundation by living in highlands or moving far away from flood-prone areas. Legends say that before the first manager of Chinese waters, Yu The Great (twenty-first century B.C.), built dikes on the lower reaches, barriers to block or store the water in large holding areas were tried for a long time. Yu The Great was great because of his innovation to direct flow to the sea by dikes. The first recorded river training by construction of levees started probably in the Warring State periods (770–221 B.C.) when the Yellow River was not yet a suspended river.

During the Han Dynasty (206 B.C. –220 A.D), disasters of flood breaches and course shifting were more overwhelming than ever before along with sediment deposition in the lower reach as well as more intensive human activities in the loess plateau and in the floodplain. And the Yellow River lower reach has gradually become a suspended river. Because the heavy silt load cut the effectiveness of flood control, it is of key importance that enormous constant efforts are made to maintain equilibrium with the river. To tackle the flooding in the Yellow River, one prominent water manager Ja Rang stated his "three measures on

River treatment (flood management)" around 7 B.C. The first measure, the best option according to Jia Rang, was to return the river to the abandoned course. However, this idea proved to be impossible at that time because of the limited technological capability when the abandoned channel was seriously silted. The second measure was to dissipate the power of the river by draining off water for irrigation in the lower reach and diverting relatively silt-free streams into the Yellow River to increase the silt-carrying capacity. The last measure was to strengthen dikes to contain the river. This measure was his last choice. He believed that the endless need for dike raising was a waste of time, labor, and resources.

After Ja Rang, Wang Jing proposed and carried out river channel stabilizations in 69 A.D so effectively that there were no major dike breaches for the next nearly 1000 years. His methods included dredging, strengthening the levees at dangerous points, digging new channels for tributaries in rough terrain, and building numerous sluice gates. A thousand years later, Pan Jixun (1521–1595) was remarkable in the Ming Dynasty (1368–1644) because he advocated building strong close dikes to contain the river so that it would scour its own narrow channel. This was the first forceful statement against the ancient principle of dividing the flow to dissipate the river's power. This strategy was also followed by Jin Fu in the early Qing Dynasty (1644–1922 A.D) to stabilize the river for a certain period.

Present Flood Control in the YR

During the early decades of the twentieth century, flood control in the Yellow River basin has evolved as a combination of traditional management practices, influences from European-based industrial society, and new technical and organizational features from modern Chinese society. The views on flood management were determined by the famous Chinese flood managers, Li Yizhi, Shen Yi, Zhang Hanying, and Wang Huayun. In the mid-1950s, a multipurpose plan for permanent control of the river was initiated. This plan included the construction of more than 40 dams and projects to moderate the river's flow (and produce energy).

Today, on the upper and middle stream of the Yellow River, there are 173 large and medium-sized reservoirs whose total storage capacity is 55.2 billion m^3. Downstream of the Sanmenxia reservoir (see Figs. 2 and 5), two large flood retardation basins—the Beijinti and the Dongpinghu—were constructed. Their storage capacity is 2 billion m^3 each and detention area 2316 km^2 and 627 km^2, respectively. Altogether 5000 bank protection works with a total length of 585 km were built. In this way, the flood control works can resist peak discharges of 22,000 m^3/s (1958 flood) at the Huayuankou hydrologic station (1958 flood), corresponding to a 60-year return period. The Xiaolangdi Reservoir whose a storage capacity is 12.65 billion m^3 on the lower reach increased the protection standard there to a return period of 1000 years (4).

The flood control engineering system that has been gradually formed since 1950 is to retain water in the upper and middle reaches, drain water at the lower reach, and divert and detain water on both sides of the river, guided by the notion of "stabilizing the flow by widening the channel." The lower reach flood control system is shown in Fig. 5. The engineering works for retaining water include

- the Sanmenxia reservoir at the main course of the Yellow River, which controls 91.4% of the total basin area;
- the Luhun reservoir and the Guxian reservoir on the Yi River and Luo River, designed to reduce the flood risk for the lower reach of the Yellow River and the city of Luoyang, combined with the Sanmenxia and Xiaolangdi reservoirs; and
- the Xiaolangdi reservoir which controls 92% of the total basin area and has a total capacity of 12.65 billion m^3 and a long-term effective capacity of 5.1 billion m^3.

Currently, both structural and nonstructural measures are adopted in Yellow River flood control. Structural measures include reservoirs, diversion structures, retardation

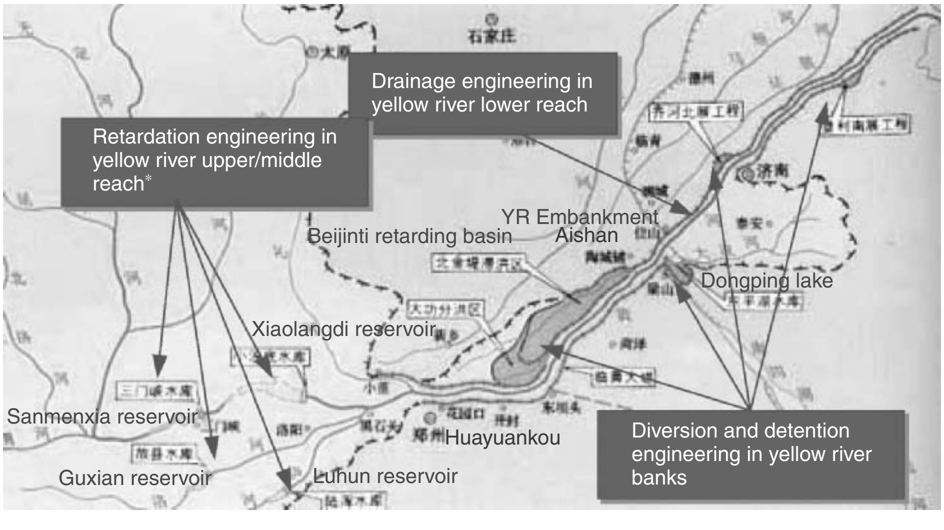

Figure 5. The flood control system of the Yellow River lower reach. Note: The upper reach's retardation engineering is not shown in this figure.

basins, embankment, dredging, and channel modifications. Nonstructural measures includes the flood control organization system, the flood control communication system, the hydrology monitoring and forecasting system, the flood regulation, command and decision making support system, the management of flood plain, and the diverting and detaining area.

Measures for Reducing Flood Risk

As stated above, the Yellow River delivers about 1.6 billion tons of sediment per year to the lower reach, resulting in an average content of 35 kg/m^3. In the reaches with the highest concentration rates, the river contains more then 10 million kg of sediment per kilometer of which 90% is from the loess plateau in the middle reach of the river between Longmen and Tongguan. Due to the heavy sedimentation, the bed level rises in the lower reach. Where in the past the river changed its course when the riverbed was too high for the surrounding area, at present embankments, buttresses, and control works harness the river and fix the river in its place. The sediment has no place to go but deposit in the lower reach. Due to human activities such as deforestation in the middle reach, soil erosion enhances the sedimentation process in the lower reach even more. At present, the riverbed is as high as 5 to 10 meters above ground level, creating a situation where in case of a dike breach during a flood, the damage can be enormous. Reducing this flood risk can be done in two ways, by reducing the peak flow and by increasing the channel capacity.

Reducing peak flow can be and presently is achieved by retention of water upstream, by detention in basins in the lower reach and by controlled flooding on floodplains. The problem of retention is difficult coordination with other water users, among whom are the Ministry of Power and farmers; when water is to be released to create a buffer and when water is to be retained to relieve downstream flooding, to increase the power generating capacity, and to conserve water for demand in the dry season.

Capacity can be increased by reducing sedimentation and by raising the embankments. Increasing capacity is currently attempted by four types of measures:

- dredging to enlarge flow cross section.
- decreasing erosion by soil conservation projects. Large afforestation projects in the middle reach have started to conserve the soil from eroding. The main focus is on simultaneous development of agriculture, forest, and animal husbandry.
- prevention of transport of sediment to the lower reach by trapping sediment in large basins or reservoirs. The recently built Xiaolangdi reservoir is constructed to trap sediment for 10 to 15 years, until the reservoir is as small as 30% of its original size. This will prevent about eight million cubic meters of sediment from flowing into the lower reach.
- increase capacity by raising the embankments, but there are financial as well as physical limits to keep doing so.

Despite all these measures, the problems remain. Even though the flood risk is momentarily decreased by Xiaolangdi, in 15 years, the risk will reach the same level again and will increase every year. Besides this, flooding in an area of 120,000 km^2 between Xiaolangdi and Huayuankou is still uncontrolled. A maximum of 13,000 m^3/s can be transported in this section, which corresponds to a 1 in 100 year event.

Analysis of Flooding in the Lower Reach of the Yellow River

Looking at the water resources system of the lower reach of the Yellow River, as described in the last sections, it can be concluded that 50 years have passed now without any dike-outbreak flooding in the lower reach. Two of the reasons are the three dike elevations on the lower reach and relatively low discharges. It cannot be concluded, however, that the danger of dike-outbreak flooding has been sufficiently prevented. In fact, the risk is still there, and a number of factors have even increased the risks. The following are reasons that the risk is still there (5):

1. Low design flood discharge. The present design flood discharge of the Yellow River on its lower reach is 22,000 m^3/s at Zhengzhou with a probability of 2.2% (without taking into account the Xiaolangdi reservoir), taking all the dikes and detention basins into account.
2. The construction of the Xiaolangdi reservoir. If flood control and silt reduction become the most important functions, the Xiaolangdi project can improve the situation for about 15 years. Around 2015, new measures have to be adopted to keep flood control in the lower reach at least at the same level.
3. Poor quality of flood control works. There are 832 km of dike sections whose height is lower than the design water level on account of the steady filling up of the riverbed. There are 340 km of weak dike sections.
4. The strongly varying width of the river between Xiaolangdi and Gaocun. The width varies between 5 and 20 km.
5. Increasing danger of earthquakes. The lower reach of the Yellow River is located in a strong faulting area. It is predicted that in a large part of the area an earthquake of strength 7–8 on the Richter scale may occur within the next 50 years, and there is even a possibility of earthquakes of strength 9 in some areas (6).

These are the reasons that the risk has increased and is still increasing:

1. Densely populated detention areas, whose population is still growing. Beijindi detention area and Dongping Lake house millions of people and are therefore not suitable any longer to store water in case of a flood with a short forecast period (less than 24 hours).
2. Inside the flood-prone areas, approximately 78 million people live. The industrial development is high in these areas, and the economic importance

of these areas is high. The potential for flood damage has increased enormously in recent years, so a dike breach would cause a disaster.

3. The increasing number of people who live in the floodplain. The total number of people who live outside the dikes in the lower reach is 1.8 million. In Henan province only, it is already 1.10 million. These people are very seriously threatened by floods.
4. Decreasing flood conveyance capacity of the river channel. Owing to the construction of reservoirs on the upper reach and irrigation development in the whole basin, discharge in the main channel of the Yellow River has decreased and sedimentation on its lower reach has increased considerably in recent years. As a result, the water level for the same discharge is much higher than before.
5. The role of the YRCC in the distribution decisions for the reservoirs. The Sanmenxia reservoir, for instance, is under the jurisdiction of the YRCC. The management of Xiaolangdi project is not yet known. If it is run under one set of operating priorities, it can improve the flood control in the lower reach. On the other hand, a different set of priorities may enable Xiaolangdi to deliver economically interesting hydropower.

Overall, it can be said that due to the decreasing flood conveyance and the low design flood discharge, the probability of a serious flood increases rapidly. If a flood occurs, the structural flood prevention measures such as dikes and detention areas do not meet the required standard or cannot be used. The increasing population and the economic development would cause much more economic damage in case of flooding than 10 years ago. Moreover, it can also be expected that the number of casualties will be much higher than 10 years ago. The laws and the executive authorities are not able to control the number of people living outside the dikes and related unspecified dike building inside the flood-plain. So still a lot has to be done to improve flood control in the lower reach.

Looking at the measures taken to control the river in the last 50 years, the approach focuses on flood prevention by raising and strengthening the embankments and the construction of dams like Sanmenxia and more recently Xiaolangdi. However, it seems that the traditional methods cannot find sustainable solutions and create a sustainable and safe WRS for the LYR and the flood-endangered areas around this part of the river. The danger of floods is still there and, for the reasons mentioned before, the danger is even increasing.

For the future, different research projects have started, and measures are proposed. Bypasses are also suggested. Theoretically, one could dig out an alternative route for the river, complete with levees. Some scholar even suggested where the bypass can be located and which region would benefit from the Yellow River discharge from a regional economic and ecological view point. However, such an artificial bypass would be incredibly expensive, and the consequences for social activities and ecosystems are uncertain.

BIBLIOGRAPHY

1. Li, R. (2001). *General Introduction to the Yellow River*. Draft version, Delft, The Netherlands.
2. Yellow River: *High sediment concentration flow*: http://users.ox.ac.uk/wolf1016/yellow_river_pics.htm.
3. Wang, Z-Y. and Wu, Y. (2001). Sediment-removing capacity and river motion dynamics. *Inf. J. of Sediment Res.* **16**(2): 105–115.
4. Zhang, H-L. and Kang, W. Flood Control And Management For Large Rivers In China, http://www.unescap.org/enrd/water_mineral/disaster/flood%20china.doc.
5. Kemink, E. (2002). *Flood Management in the Yellow River*. Draft version, Delft, The Netherlands.
6. Li, X. *Flood-Prone Areas and Their Management on the Lower Reach of the Yellow River*. Institute of Geography, Chinese Academy of Sciences, Beijing, China.

READING LIST

Yellow River: Geographic and historical settings: http://www.cis.umassd.edu/gleung/geofo/geogren.html.

Zusman, E. *A River Without Water: Examining Water Shortages in the Yellow River Basin* http://uts.cc.utexas.edu/journal/1998/river.html.

CHIRONOMIDS IN SEDIMENT TOXICITY TESTING

Paul K. Sibley
University of Guelph
Guelph, Ontario, Canada

INTRODUCTION

Historically, assessments of water quality have largely been based on the evaluation of water-borne contaminants and water quality criteria established by comparing aqueous concentrations of compounds to their toxicity determined in water-only exposures. It was not until the mid-1970s that it was realized that contaminants, particularly hydrophobic compounds, also occurred in sediments, often at high concentrations. With this new understanding came the realization that water quality criteria based solely on aqueous concentrations were insufficient to regulate contaminants that existed predominantly in sediments and would, therefore, not adequately protect benthic organisms. To develop appropriate and effective criteria for regulating sediment-associated contaminants, the need to develop new methods capable of assessing sediment toxicity became apparent. This need led to the birth of the field of sediment toxicology.

Since it was first recognized that sediments served as both a source and a sink for many contaminants, research in the field of sediment toxicology has largely focused on the development and validation of methods to assess toxicity. Detailed reviews that document the collection and manipulation of sediments, selection of test species, development of test methods, optimization of test conditions and experimental design, and validation of freshwater sediment toxicity test methods are provided in Nebeker et al. (1), Giesy and Hoke (2), Burton (3,4), Ankley et al. (5,6), Ingersoll et al. (7,8), and Diamond et al. (9). A critical consideration in the development of sediment toxicity tests is the selection of appropriate test species. Although several invertebrate species now play an integral role in the assessment of sediment toxicity in freshwater systems (10), arguably none have been used more extensively than the two midge species *Chironomus tentans* and *C. riparius*.

The genus *Chironomus* (Diptera: Chironomidae) includes a large number of species globally, many of which occur ubiquitously in aquatic habitats. To date, seven species of *Chironomus* have been used in sediment toxicity testing: *C. tentans* (nearctic populations of *C. tentans* are now known as *C. dilutus* [Shabanov et al. (11)], *C. riparius, C. plumosus, C. attenuatus, C. prasinus, C. crasssiforceps*, and *C. tepperi*. Of these, *C. tentans* and *C. riparius* have dominated the sediment toxicology literature, with the former being used predominantly in North America and the latter predominantly in Canada and Europe. Both species have been incorporated into standard test methods (12–14), and their application in sediment toxicity testing has been reviewed in detail (10,12,15). The adoption of these two midges in sediment toxicity assessments reflects their importance to the food webs of aquatic systems (e.g., ecological relevance), relative sensitivity to contaminants, tolerance of a wide range of sediment characteristics, and benthic mode of existence (8). From a practical standpoint, both midges have short generation times and are therefore easy to culture and available for year-round testing, are relatively insensitive to manipulation during culturing and toxicity testing, and have wide acceptance among regulatory agencies.

In light of the dominant application of both *C. tentans* and *C. riparius* in sediment toxicity testing, including the availability of standard protocols, this review will focus on these two species, with the goal of providing a pragmatic overview of their application in sediment toxicity assessment. For information on the application of the other species of *Chironomus* listed above in toxicity testing, the reader is referred to Sanchez and Tarazona (16) (*C. prasinus*), Wilson et al. (17) (*C. tepperi*), Peck et al. (18) (*C. crassiforceps*), Fargasova (19) (*C. plumosus*), and Darville and Wilhm (20) (*C. attenuatus*).

BIOLOGY OF *C. TENTANS* AND *C. RIPARIUS*

The biology of *C. tentans* and *C. riparius* has been described in detail elsewhere and is thus only briefly described here. For information of the biology of *C. tentans*, the reader is referred to Sadler (21), Hall et al. (22), Ineichen et al. (23,24), Sibley et al. (25,26), and Watts and Pascoe (27). For information on the biology of *C. riparius*, the reader is referred to Edgar and Meadows (28), Caspary and Downe (29), Downe (30), Rasmussen (31,32), Watts and Pascoe (27), and Hooper et al. (33).

C. tentans and *C. riparius* are holometabolous insects and thus have a life cycle that consists of an egg, larval, pupal, and adult stage. The larval stage in *C. tentans* lasts approximately 23 days at the recommended test temperature of 23 °C (34). In contrast, the larval stage of *C. riparius* lasts between 15 and 18 days at a temperature of 20 °C, depending on the life stage (e.g., egg versus newly hatched larvae) that initiates the test (35). However, in a study comparing the relative sensitivity of the two midges to selected contaminants, Watts and Pascoe (27) found that the duration of the larval stage for *C. riparius* and *C. tentans* was comparable at 22 °C. Larvae of both species pass through four instars and are generally comparable in size (based on head capsule width) up to the fourth instar. Thereafter, larvae of *C. tentans* grow considerably larger than those of *C. riparius* (27). Toward the end of the fourth instar, larvae of both species become inactive and cease to grow; this period is followed by pupation, which lasts 1–2 days. In both species, emergence follows a bimodal pattern (protandry) in which males emerge 2–5 days before females (26,35). Males can be distinguished from females by the presence of plumose antennae. Males of both *C. tentans* and *C. riparius* are capable of multiple matings, whereas females are not receptive to additional matings once they have been inseminated (26,30). Impregnated females of both species oviposit a single egg mass the day after insemination. Female *C. tentans* will occasionally produce a small secondary egg mass, but these eggs are usually not viable. Although *C. tentans* has been reported to produce egg masses containing up to 2300 eggs (22), 800–1000 is more typical (25,26). In contrast, egg masses of *C. riparius* typically contain between 400 and 600 eggs (36). Unmated females generally do not oviposit and resorb the egg mass. However, unmated females of *C. tentans* may occasionally produce an egg mass containing nonviable eggs (25). In *C. tentans*, adult females live up to 5 days, whereas males live up to 7 days.

ACUTE SEDIMENT TOXICITY TESTS

The most commonly applied standard test using both *C. tentans* and *C. riparius* is the 10-d survival and growth assay (1,7,10). The 10-d version using *C. tentans*, along with its 10-d counterpart using the amphipod *Hyallela azteca*, has been evaluated in an interlaboratory study of precision (3). The procedures and schedule for conducting a 10-d test are provided in Table 1. Although the test procedures described below and in Table 1 are those for *C. tentans*, the test conditions and requirements are similar for both species. The reader is referred to ASTM (13) and Environment Canada (14) for specific test methods for *C. riparius*.

The 10-d test method for *C. tentans* consists of exposing 10-d old larvae to a contaminated sediment at 23 °C with a light:dark photoperiod of 16:8 and an illumination of 500–1000 lux (7,12). Test chambers consist of 300 ml beakers containing 100 ml of sediment and 175 ml of overlying water. The number of replicates recommended for an assay is eight per treatment, although fewer (not less than 3) may be used. Each beaker is fed 1.5 ml of a 4 g/L Tetrafin fish food slurry each day. After the 10-d exposure period, larvae are sieved from the sediment and enumerated to determine survival. A test is considered to be acceptable if survival in the control or reference sediment is ≥70% (Table 2). If growth is to be measured,

Table 1. Test Setup and Activity Schedule of the *C. tentans* Acute Sediment Toxicity Test

Test Day	Activity
−1	Add 100 ml of homogenized sediment (or fine sand if water only exposure) to each replicate beaker and place in appropriate holding tank. Allow sediment to settle (about 1 hr), and then add 1.5 ml of Tetrafin fish food slurry (4 g/L) to each beaker.
0	Randomly allocate 10 larvae (50% must be in third instar) to each replicate beaker using a Pasteur pipette. Let beakers sit for 1 hr after addition of the larvae. After this period, gently immerse/place all beakers into the exposure system.
1-end	On a daily basis, check assay system and add 1.5 ml of food to each beaker. Dissolved oxygen should not fall below 2.5 mg/L. Reduce feeding rate or frequency if persistent low-dissolved oxygen occurs.
10	Sieve the sediment from each beaker to recover larvae for growth and survival determinations. Place recovered larvae from each replicate in preashed aluminum pans and oven dry for 24 hours at 90 °C.
11	Weigh dried larvae to the nearest hundredth milligram. If ash-free dry weight is to be determined, ash the pan and larvae at 550 °C for 2 hours. Allow pans to cool to room temperature in a desiccator and reweigh. Individual larval dry weight is determined as the difference between the total mean dry weight and total mean ashed weight divided by the number of surviving larvae per replicate.

Adapted from Ingersoll et al. (7).

larvae are placed in pre-ashed aluminum weigh pans, dried at 90 °C for 24 hours, and weighed once they have returned to room temperature. In instances where dry weight may be influenced by different sediment particle size characteristics, larval weight should be determined as ash-free dry weight (37). In this approach, the dried larvae in the aluminum pan are ashed at 550 °C for 2 hours and reweighed as above. The ash-free dry weight is determined as the difference between the dried weight and ashed weight. Larval weight is typically expressed on an individual basis (e.g., milligram dry weight per individual). However, Call et al. (38) showed that dry weight varied depending on the number of surviving larvae, with higher weights recorded in sediments with lower survival. They found that this bias was removed by expressing larval dry weights as total weight per replicate.

Once collected the data must be analyzed statistically. A detailed account of statistical procedures for analyzing data generated from both acute and chronic (see below) toxicity tests is beyond the scope of this document. For general guidance, the reader is referred to U.S. EPA (12) and Environment Canada (39).

Based on a comparison of LC50s, Watts and Pascoe (27) compared the relative performance and sensitivity of *C. tentans* and *C. riparius* in sediment toxicity tests with cadmium and lindane. They found no difference in sensitivity in exposures to cadmium but showed that *C. tentans* was significantly more sensitive than *C. riparius* to lindane. The latter result is consistent with the findings of Pauwels and Sibley (10) who compared the relative sensitivity of several sediment test species based on LC50 values obtained from a survey of the sediment toxicology literature. In most cases, *C. riparius* was found to be less sensitive than *C. tentans*.

CHRONIC SEDIMENT TOXICITY AND LIFE CYCLE TESTS

In response to the need for chronic evaluation of sediment toxicity, Benoit et al. (34) introduced a life cycle test using *C. tentans*. A parallel test was developed using the amphipod *Hyallela azteca* (40), and collectively, these two assays represent the only standardized tests currently available for evaluating the effects of sediment-associated contaminants over the full life cycle of a benthic organism (12,13). Although a long-term toxicity test using *C. riparius* has apparently been developed (41), this test has not been standardized and appears to have received little attention in recent years. Nonetheless, *C. riparius* has been applied in several long-term studies that have incorporated emergence and reproduction (42,43). Given its close similarity to *C. tentans* in most aspects of its life history, *C. riparius* could probably be readily adapted for use in life cycle testing following the standard procedure for *C. tentans*.

The recommended exposure system, required testing apparatus, and experimental design for conducting a life cycle test with *C. tentans* is described in detail in Benoit et al. (34) and U.S. EPA (12). Table 3 summarizes the activity schedule for a typical life cycle test. The test is initiated with first instar larvae that are <24 hours old. The egg cases are transferred to a glass petri dish (do not use plastic petri dishes as the larvae will quickly

Table 2. Test Acceptability Criteria (Survival and Growth) and Recommended Criteria (Emergence, Reproduction, and Percent Hatch) for Control Organisms in 10-d Tests with *C. riparius* and *C. tentans* and the *C. tentans* Life Cycle Test

Endpoint/Condition	10-d Test		Life Cycle Test
	C. tentans	*C. riparius*	
Survival	70%	70%	70%[a]
Growth	0.48 mg afdw[b]/individual	Not determined	0.60 mg dry weight/individual or 0.48 mg afdw/individual
Emergence	N/A	N/A	≥50%
No. of eggs/egg mass	N/A	N/A	≥800
Percent hatch	N/A	N/A	≥80%

Adapted from USEPA (12) and ASTM (13). N/A = Not applicable

[a] Measured at 20 days.

[b] afdw = ash free dry weight.

Table 3. Test Setup and Activity Schedule for Conducting a Life Cycle Chronic Sediment Toxicity Test with *C. tentans*

Test Day	Activity
−4	Start reproduction flask with cultured adults (10:30 male:female ratio).
−3	Collect egg cases (6−8) and incubate at 3 °C.
−2	Check egg cases for viability and development.
−1	Check egg cases for hatch and development. Add 100 ml of homogenized sediment (or fine sand if water only exposure) to each replicate beaker, and place in appropriate holding tank. Allow sediment to settle (about 1 hr), and then add 1 ml of Tetrafin fish food slurry (4 g/L) to each beaker.
0	Transfer egg cases to a glass crystallizing dish containing control water. Discard larvae that have already left the egg cases. Add 1 ml food to each beaker before the larvae are added. Randomly allocate 12 larvae to each replicate beaker using a Pasteur pipette. Let beakers sit for 1 hr after addition of the larvae. After this period, gently immerse/place all beakers into the exposure system.
1-end	On a daily basis, check assay system and add 1.5 ml of food to each beaker
6−10	Set up schedule for auxiliary male beakers as described for day −4 to day 0 above.
20	Randomly select four replicates from each treatment, and sieve the sediment to recover larvae for growth and survival determinations. Place recovered larvae from each replicate in preashed aluminum pans and oven dry for 24 hr at 90 °C. Install emergence traps on each of the remaining reproductive replicates.
21	Weigh dried larvae to the nearest hundredth milligram. Ash the pan and larvae at 550 °C for 2 hr. Allow pans to cool to room temperature in a desiccator and reweigh. Individual larval dry weight is determined as the difference between the total mean dry weight and total mean ashed weight divided by the number of surviving larvae per replicate.
23-end	On a daily basis, record emergence of males and females, pupal and adult mortality, and time to death for previously collected adults. Each day, transfer adults from each replicate to a corresponding reproduction/oviposition (R/O) chamber. Transfer each primary egg mass from the R/O chamber to a corresponding petri dish and monitor for incubation and hatch. Record each egg case oviposited, number of eggs produced (using a direct count or the ring method [Benoit et al. (34)]), and number of hatched eggs.
30	Place emergence traps on auxiliary male replicate beakers.
33-end	Transfer males emerging from the auxiliary beakers to individual inverted petri dishes. The auxiliary males are used for mating with females from corresponding treatments for which most males have already emerged (because of protandrous emergence).
40-end	After 7 d of no emergence in a given treatment, terminate that treatment by sieving the sediment to recover remaining larvae, pupae, and pupal exuviae.

Adapted from Benoit et al. (34).

attach to the bottom and become difficult to move), and only larvae that emerge from the egg cases after this transfer should be used to initiate the test. The larvae are transferred to the exposure beakers using a Pasteur pipette and the aid of a microscope. This stage of the assay can be difficult for those who lack practice in manipulating larvae of this size (typically <1 mm), and it is recommended that the user become familiar with this procedure before initiating an assay. However, starting a test with larvae at this life stage, which is typically the most sensitive to contaminants, may yield the most accurate estimate of toxicity. To aid in the successful transfer of these small larvae, the user can periodically verify the number of larvae in the pipette by holding it up to a light and against a black background. Typically each beaker receives 12 larvae, although tests using 10 larvae have been conducted successfully (44).

A unique aspect of the life cycle test is the addition of auxiliary beakers on day 10. These beakers are necessary to supply males to compensate for the protandrous emergence pattern of *C. tentans*; without the auxiliary beakers, many of the later emerging females would not have males with which to mate, diminishing the ability to adequately assess reproduction. The auxiliary male beakers are set up following the same procedures as those described for the beginning of the test.

The primary endpoints measured in the life cycle test are survival (at 20 days and at the end of the test), growth, emergence, and reproduction. If desired, additional measurements on pupal and adult survival may be recorded. Growth and survival are determined after 20 d in the same manner as described for the 10-d test. Emergence is assessed daily over a 20−30-day period using emergence traps placed on top of each beaker (34). The emergence traps double as reproductive chambers in which mating takes place and females can oviposit. Reproductive output is determined by counting the number of eggs in each egg case; this is achieved by counting the number of eggs in five equidistant rings along the egg rope and multiplying by the total number of rings in the egg case (34). The eggs are incubated until hatch is complete (approximately 6 days); at which time, remaining eggs are counted; the difference between the number of unhatched eggs and the total number of eggs from the original count provides an estimate of percent hatch. In situations where eggs are difficult to count because of poorly formed egg cases, they may be counted directly by dissolving the egg case in dilute acid (34). The life cycle test is terminated when no hatch has been observed in the control or reference sediment for a period of 7 days. Test acceptability criteria for survival and growth and recommended criteria for emergence and reproduction are provided in Table 2.

The *C. tentans* life cycle test has been evaluated in a round-robin precision evaluation, the results of which are summarized in U.S. EPA (12), and it has been used successfully in several evaluations of contaminated sediments and water-borne exposures (25,26,34,44−46). With the length of the *C. tentans* life cycle test ranging from 45 to 60 days, this test is resource-and labor-intensive relative to the 10-d test. However, because the life cycle test includes effects on survival, growth, emergence, and reproduction, it yields a substantial amount of information that can comprehensively evaluate potential risks of contaminated sediments to benthic invertebrates. Furthermore, a significant benefit of having detailed, synoptic information on survival, growth, emergence, and reproduction is that this can be applied in population modeling as has been done recently for both *C. tentans* and *C. riparius* (26,33,47,48).

FUTURE APPLICATIONS OF CHIRONOMUS SPP. IN SEDIMENT TOXICITY TESTING

With the availability of standard sediment testing procedures and clear guidelines for their use in sediment ecological risk assessment (ERA), recent efforts have been directed toward the development of sediment quality criteria/guidelines (SQGs) for the protection of benthic communities. A detailed description of the different types of SQGs and their basis is beyond the scope of this document; for information on SQGs, the reader is referred to Long et al. (49,50), Ingersoll et al. (8), MacDonald et al. (51), and U.S. EPA (52). Future refinements in sediment ERA, along with the development of appropriate SQGs, will continue to require sediment toxicity information, and it is likely that species of the genus *Chironomus* will continue to play a key role in this regard. Some areas for future methodological development in support of sediment ERA are briefly described below.

An area of recent activity in the field of sediment toxicology has been the development and validation of *in situ* test methods. Both *C. tentans* and *C. riparius* have played a central role in the development of *in situ* sediment testing methods (53–60). *In situ* tests involve the transplantation of organisms cultured under laboratory conditions to the field where exposure takes place in specialized exposure apparatus (53,54,59,60). In most cases, *in situ* tests have been developed and evaluated in relatively shallow habitats because of the difficulty in deploying organisms remotely in deeper waters. Although *in situ* assays involving *Chironomus* spp. have been evaluated, none have been developed into standardized protocols.

Another potentially important area for future development in sediment toxicity testing with both *C. tentans* and *C. riparius* will be in the area of endocrine disruption (ED). Few invertebrate assays are currently available to investigate sediment-based ED, but recent studies have shown that species of *Chironomus* may be useful in this application. From a physiological standpoint, *Chironomus* spp. undergoes complex biochemical changes, characterized by dramatic shifts in the titers of ecdysteroids, associated with both molting and pupation. Recent evidence indicates that the activity of certain ecdysteroids can be altered by endocrine disrupting compounds (61,62), which enhances the sensitivity of organisms during these sensitive life history stages. In this context, these developmental stages may serve as ideal focal points for the development of ED-based assays using these midges (63). Some research has already been conducted in this area. Watts et al. (43) used *C. riparius* growth and reproduction to evaluate ethynylestradiol (E2) and Bisphenol A but found little evidence of ED. Hahn et al. (61) found that mortality of males was twice as high as females in exposures of *C. riparius* to tebufenozide and hypothesized that this could reflect ED. Several studies have used mentum deformities in *C. riparius* as an indicator of ED. For example, Meregalli et al. (64) noted increased rates of mentum deformities after exposure to nonylphenol, whereas Meregalli and Ollevier (65) found that the rate of mentum deformities after exposure to E2 was not affected. Vermuelen et al. (66) found significant mentum deformities in *C. riparius* after exposure to lead and mercury but not ß-sitosterol, a known endocrine disrupting compound.

A common attribute of these studies was that ED was inferred secondarily through measurements on growth, reproduction, and deformations. No attempt was made to directly link changes in hormone levels to the biological effects observed. Recently, however, Hahn and Schultz (62) attempted to quantitatively link growth changes, determined by examining the development of imaginal discs in fourth instar *C. riparius* larvae exposed to tributyltin, to changes in the concentration of ecdysteroids in incubated prothoracic glands measured using radioimmunoassay. They showed that ecdysteroid synthesis decreased in females at 50 ng/L, with a corresponding decrease in the rate of imaginal disc development. In contrast, ecdysteroid synthesis increased in males at 500 ng/L, with a corresponding increase in the rate of imaginal disk development. These results may pave the way for the development of a standardized sediment bioassay using *Chironomus* spp. that can be applied in the assessment of sediment-associated ED.

Another potentially promising area in the development of sediment test methods with *Chironomus* spp. is the application of cell lines. In fact, epithelial cell lines have been used for over a two decades in relation to molecular-and genetic-based research with *C. tentans* (67–69). Recent studies have begun to elucidate hormone regulation and signaling pathways using these cell lines (70,71), whereas others have examined receptor binding affinity relationships using excised imaginal disks of *C. riparius* maintained in epithelial cell line media (72). The information derived from these studies may provide the necessary foundation for the development of an invertebrate-based cell line using *C. tentans* that could be applied individually to assess contaminants in aquatic environments or in conjunction with endocrine-based assays (see above) for assessing endocrine disruption.

A final area that may prove useful in the assessment of sediment toxicity is the development of behavioral endpoints. Behavioral changes often precede acute and chronic effects; yet behavioral endpoints have been sorely underused in the field of aquatic and sediment toxicology. Behavioral changes are likely to be strongly linked to acute and chronic responses (73), which enhance the ecological relevance of sediment bioassays using *Chironomus* spp. In *Chironomus* spp., activities such as ventilation, tube construction, and feeding could be developed into readily measured behavioral endpoints that may be more sensitive indicators of sediment contamination than many of the traditional measurement endpoints (e.g., survival, growth, etc.) (74).

BIBLIOGRAPHY

1. Nebeker, A.V. et al. (1984). Biological methods for determining toxicity of contaminated freshwater sediments to invertebrates. *Environ. Toxicol. Chem.* **3**: 617–630.
2. Giesy, J.P. and Hoke, R.A. (1989). Freshwater sediment toxicity bioassessment: Rationale for species selection and test design. *J. Great Lakes Res.* **15**: 539–569.
3. Burton, G.A. (1991). Assessment of freshwater sediment toxicity. *Environ. Toxicol. Chem.* **10**: 1585–1627.
4. Burton, G.A. (1992). *Sediment Toxicity Assessment*. Lewis Publishers, Boca Raton, FL.
5. Ankley, G.T. et al. (1993). Development and evaluation of test methods for benthic invertebrates and sediments: Effects of flow rate and feeding on water quality and exposure conditions. *Arch. Environ. Contam. Toxicol.* **25**: 12–19.
6. Ankley, G.T. et al. (1994). Evaluation of potential confounding factors in sediment toxicity tests with three freshwater benthic invertebrates. *Environ. Toxicol. Chem.* **13**: 627–635.
7. Ingersoll, C.G. et al. (1995). Toxicity and bioaccumulation of sediment-associated contaminants using freshwater invertebrates: A review of methods and applications. *Environ. Toxicol. Chem.* **14**: 1885–1894.
8. Ingersoll, C.G. et al. (1997). Workshop summary report on uncertainty evaluation of measurement endpoints used in sediment ecological risk assessments. In: *Ecological Risk Assessment of Contaminated Sediments*. C.G. Ingersoll, T. Dillon, and G.R. Biddinger (Eds.). SETAC Press, Pensacola, FL, pp. 271–298.
9. Diamond, J., Burton, A., and Scott, J. (2002). Methods for collecting, storing, and manipulating sediments and interstitial water samples for chemical and toxicological analyses. In: *A Handbook on Sediment Quality*. A special publication of the Water Environment Federation (WEF), Alexandria, VA, pp. 139–197.
10. Pauwel, S. and Sibley, P.K. (2002). Testing for toxicity and bioaccumulation in freshwater sediments. In: *A Handbook on Sediment Quality*. A special publication of the Water Environment Federation (WEF), Alexandria, VA, pp. 199–258.
11. Shabanov, N.A., Kiknadze, I.I., and Butler, M.G. (1999). Palearctic and Nearctic *Chironomus* (*Camptochironomus*) *tentans* Fabricius are different species (Diptera: Chironomidae). *Entomol. Scand.* **30**: 311–322.
12. United States Environmental Protection Agency. (2000). *Methods for Measuring the Toxicity and Bioaccumulation of Sediment-Associated Contaminants with Freshwater Invertebrates*, 2nd Edn. EPA 600/R-99/064. Duluth, MN.
13. American Society for Testing and Materials. (2004). Test method for measuring the toxicity of sediment-associated contaminants with freshwater invertebrates. Standard E1706-00e2, ASTM, Philadelphia, PA.
14. Environment Canada. (1998). Survival and growth in sediment using freshwater midge larvae *Chironomus tentans* or C. *riparius*. EPS 1/RM/32.
15. Burton, G.A., Denton, D.L., Jr., Ho, K., and Ireland, D.S. (2003). Sediment toxicity testing: Issues and methods. In: *Handbook of Ecotoxicology*, 2nd Edn. D.J. Hoffman, B.A. Rattner, G.A. Burton and J. Cairns, Jr. (Eds.). Lewis Publishers, Boca Raton, FL, pp 111–150.
16. Sanchez, P. and Tarazona, J.V. (2002). Development of a multispecies system for testing reproductive effects on aquatic invertebrates. Experience with *Daphnia magna, Chironomus prasinus*, and *Lymnaea peregra*. *Aquat. Toxicol.* **60**: 249–256.
17. Wilson, A.L., Stevens, M.M., and Watss, R.J. (2000). Acute and chronic toxicity of the herbicide benxofenap (Taipan 300) to *Chironomus tepperi* Skuse (Diptera: Chironomidae) and *Isidorella newcombi* (Adams and Angas) (Gastropda: Planorbidae). *Arch. Environ. Contam. Toxicol.* **48**: 7–13.
18. Peck, M.R., Klessa, D.A., and Baird, D.J. (2002). A tropical sediment toxicity test using the diptera Chironomus crassiforceps to test metal bioavailability with sediment pH change in tropical acid-sulfate sediments. *Environ. Toxicol. Chem.* **21**: 720–728.
19. Fargasova, A. (2001). Winter third- to fourth-instar larvae of *Chironomus plumosus* as bioassay tools for assessment of acute toxicity of metals and their binary combinations. *Ecotoxicol. Environ. Safety* **48**: 1–5.
20. Darville, R.G. and Wilhm, J.L. (1984). The effect of naphthalene on oxygen consumption and hemoglobin concentration in *Chironomus attenuatus* and on oxygen consumption and life cycle of *Tanytarsus dissimilis*. *Environ. Toxicol. Chem.* **3**: 135–141.
21. Sadler. (1935).
22. Hall, D.J., Cooper, W.E., and Werner, E.E. (1970). An experimental approach to the production dynamics and structure of freshwater animal communities. *Limnol. Oceanogr.* **16**: 839–928.
23. Ineichen, H., Riesen-Willi, U., and Fischer, J. (1979). Experimental contribution to the ecology of *Chironomus* (Diptera). II. The influence of the photoperiod on the development of *Chironomus plumosus* in the 4th instar. *Oecologia* **39**: 161–183.
24. Ineichen, H., Meyer, B., and Lezzi, M. (1983). Determination of the developmental stage of living fourth instar larvae of *Chironomus tentans*. *Develop. Biol.* **98**: 278–286.
25. Sibley, P.K., Benoit, D.A., and Ankley, G.T. (1997). The significance of growth in *Chironomus tentans* sediment toxicity tests: Relationship to reproduction and demographic endpoints. *Environ. Toxicol. Chem.* **16**: 336–345.
26. Sibley, P.K., Benoit, D.A., and Ankley, G.T. (2001). Factors affecting reproduction and the importance of adult size on reproductive output in *Chironomus tentans*. *Environ. Toxicol. Chem.* **20**: 1296–1303.
27. Watts, M.M. and Pascoe, D. (2000). A comparative study of *Chironomus riparius* Meigen and *Chironomus tentans* Fabricius (Diptera: Chironomidae) in aquatic toxicity tests. *Arch. Environ. Contam. Toxicol.* **39**: 299–306.
28. Edgar, W.D. and Meadows, P.S. (1969). Case construction, movement, spatial distribution, and substrate selection in the larva of *Chironomus riparius* Meigen. *J. Exp. Biol.* **50**: 247–253.
29. Caspary, V.G. and Downe, A.E.R. (1971). Swarming and mating of *Chironomus riparius* (Diptera: Chironomidae). *Can. Entomol.* **103**: 444–448.
30. Downe, A.E.R. (1973). Some factors influencing insemination in laboratory swarms of *Chironomus riparius* (Diptera: Chironomidae). *Can. Entomol.* **105**: 291–298.
31. Rasmussen, J.B. (1984). The life history, distribution, and production of *Chironomus riparius* and *Glyptotendipes paripes* in a prairie pond. *Hydrobiologia* **119**: 65–72.
32. Rasmussen, J.B. (1984). Comparison of gut contents and assimilation efficiency of fourth instar larvae of two coexisting chironomids, *Chironomus riparius* Meigen and *Glyptotendipes paripes* (Edwards). *Can. J. Zool.* **62**: 1022–1026.
33. Hooper, H.L., Sibly, R.M., Hutchinson, T.H., and Maund, S.J. (2003). The influence of larval density, food availability and

habitat longevity on the life history and population growth rate of the midge *Chironomus riparius*. *Oikos* **102**: 515–524.

34. Benoit, D.A., Sibley, P.K., Jeunemann, J.J., and Ankley, G.T. (1997). Design and evaluation of a life cycle test with the midge *Chironomus tentans* for use in assessing chronic toxicity of contaminated sediments. *Environ. Toxicol. Chem.* **16**: 1165–1176.
35. Watts, M.M. and Pascoe, D. (1998). Selection of an appropriate life-cycle stage of *Chironomus riparius* Meigen for use in chronic sediment toxicity testing. *Chemosphere* **36**: 1405–1413.
36. Downe, A.E.R. and Caspary, V.G. (1973). The swarming behaviour of *Chironomus riparius* (Diptera: Chironomidae) in the laboratory. *Can. Entomol.* **105**: 165–171.
37. Sibley, P.K., Monson, P.D., and Ankley, G.T. (1997). The effect of gut contents on dry weight estimates of *Chironomus tentans* larvae: implications for interpreting toxicity in freshwater sediment toxicity tests. *Environ. Toxicol. Chem.* **16**(8): 1721–1726.
38. Call, D.J. et al. (1999). Observations on the 10-day *Chironomus tentans* survival and growth bioassay in evaluating Great Lakes sediments. *J. Great Lakes Res.* **25**: 171–178.
39. Environment Canada (1999). Guidance document on application and interpretation of single-species tests in environmental toxicology. EPS 1/RM/34.
40. Ingersoll, C.G. et al. (1998). Use of sublethal endpoints in sediment toxicity tests with the amphipod *Hyalella azteca*. *Environ. Toxicol. Chem.* **17**: 1508–1523.
41. Streloke, M. and Kopp, H. (Eds.). *Long-term Toxicity Test with Chironomus riparius: Development and Validation of a New Test System*. Mitt. Biol. Bundesanst. Land-u Forstwirtschaft, Berlin.
42. Postma, J.F. and Davids, C. (1995). Tolerance induction and life cycle changes in cadmium-exposed *Chironomus riparius* (Diptera) during consecutive generations. *Ecotoxicol. Environ. Safety* **30**: 195–202.
43. Watts, M.M., Pascoe, D., and Carroll, K. (2001). Chronic exposure to 17 ∝-ethinylestradiol and bisphenol A—effects on development and reproduction in the freshwater invertebrate *Chironomus riparius* (Diptera: Chironomidae). *Aquat. Toxicol.* **55**: 113–124.
44. MacDonald, M. et al. (2004). Toxicity of perfluorooctanesulfonic acid (PFOS) to *Chironomus tentans* under field and laboratory conditions. *Environ. Toxicol. Chem.* **23**(9): 2116–2123.
45. Sibley, P.K., Ankley, G.T., Cotter, A.M., and Leonard, E.N. (1996). Predicting chronic toxicity of sediments spiked with zinc: An evaluation of the acid-volatile sulfide model using a life-cycle test with the midge *Chironomus tentans*. *Environ. Toxicol. Chem.* **15**: 2102–2112.
46. Kahl, M.D., Makynen, E.A., Kosian, P.A., and Ankley, G.T. (1997). Toxicity of 4-Nonylphenol in a Life-Cycle Test with the Midge *Chironomus tentans*. *Ecotox Environ. Safety* **38**: 155–160.
47. Pery, A.R.R. et al. (2002). A modeling approach to link food availability, growth, emergence, and reproduction for the midge *Chironomus riparius*. *Environ. Toxicol. Chem.* **21**: 2507–2513.
48. Pery, A.R.R. et al. (2003). A model to understand the confounding effects of natural sediments in toxicity tests with *Chironomus riparius*. *Environ. Toxicol. Chem.* **22**: 2476–2481.
49. Long, E.R., MacDonald, D.D., Smith, S.L., and Calder, F.D. (1995). Incidence of adverse effects within ranges of chemical concentrations in marine and estuarine sediments. *Environ. Manage.* **19**: 81–97.
50. Long, E.R. and MacDonald, D.D. (1998). Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems. *Human Ecol. Risk Assess.* **4**(5): 1019–1039.
51. MacDonald, D.D., Ingersoll, C.G., and Berger, T.A. (2000). Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. *Arch. Environ. Contam. Toxicol.* **39**: 20–31.
52. United States Environmental Protection Agency. (2003). *Technical Basis for the Derivation of Equilibrium Partitioning Sediment Benchmarks (ESBs) for the Protection of Benthic Organisms*. EPA-600-R-02-014. Office of Research and Development, Washington, DC.
53. Chappie, D.J. and Burton, G.A. (1997). Optimization of *in situ* bioassays with *Hyalella azteca* and *Chironomus tentans*. *Environ. Toxicol. Chem.* **16**: 559–564.
54. Sibley, P.K. et al. (1999). An exposure apparatus for the *in situ* assessment of sediment toxicity and bioaccumulation. *Environ. Toxicol. Chem.* **18**: 2325–2336.
55. Tucker, K.A. and Burton, G.A. (1999). Assessment of nonpoint source runoff in a stream using *in situ* and laboratory approaches. *Environ. Toxicol. Chem.* **18**: 2797–2803.
56. Crane, M., Delaney, P., Mainstone, C., and Clarke, S. (1995). Measurement by *in situ* bioassay of water quality in an agricultural catchment. *Wat. Res.* **29**: 2441–2448.
57. Crane, M. et al. (2000). An *in situ* system for exposing aquatic invertebrates to contaminated sediments. *Environ. Toxicol. Chem.* **19**: 2715–2719.
58. Castro, B.B., Guilhermino, L., and Ribeiro, R. (2003). *In situ* bioassay chambers and procedures for assessment of sediment toxicity with *Chironomus riparius*. *Environ. Pollut.* **125**: 325–335.
59. Meregalli, G., Vermeulen, A.C., and Ollevier, F. (2000). The use of chironomid deformation in an *in situ* test for sediment toxicity. *Ecotox Environ. Safety* **47**: 231–238.
60. den Besten, P.J., Naber, A., Grootelar, E.M.M., and van de Guchte, C. (2003). *In situ* bioassays with *Chironomus riparius*: Laboratory-field comparisons of sediment toxicity and effects during wintering. *Aquat. Ecosys. Health Manage. Soc.* **6**: 217–228.
61. Hahn, T., Liess, M., and Schultz, R. (2001). Effects of the hormone mimetic insecticide tebufenozide on *Chironomus riparius* larvae in two different exposure setups. *Ecotoxicol. Environ. Safety* **49**: 171–178.
62. Hahn, T. and Schultz, R. (2002). Ecdysteroid synthesis and imaginal disc development in the midge *Chironomus tentans* as biomarkers for endocrine effects of tributyltin. *Environ. Toxicol. Chem.* **21**: 1052–1057.
63. Ingersoll, C.G., Hutchinson, T., Crane, M., Dodson, S., DeWitt, T., Gies, A., Huet, M-C., McKenney, C.L., Oberdorster, E., Pascoe, D., Versteeg, D.J., and Warwick, O. (1999). Laboratory toxicity tests for evaluating potential effects of endocrine-disrupting compounds. In: *Endocrine Disruption in Invertebrates: Endocrinology, Testing, and Assessment*. P.L. DeFur, M. Crane, C.G. Ingersoll, and L. Tattersfield (Eds.). SETAC Press, Pensacola, FL, pp. 107–141.
64. Meregalli, G., Pluymers, L., and Ollevier, F. (2001). Induction of mouthpart deformities in *Chironomus riparius* larvae exposed to 4-n-nonylphenol. *Environ. Pollut.* **111**: 241–246.
65. Meregalli, G. and Ollevier, F. (2001). Exposure of *Chironomus riparius* larvae to 17 ∝-ethynylestradiol: effects on survival and mouthpart deformities. *Sci. Total Environ.* **269**: 157–161.

66. Vermeulen, A.C. et al. (2000). Exposure of *Chironomus riparius* larvae (Diptera) to lead, mercury, and B-Sitosterol: effects on mouthpart deformation and moulting. *Chemosphere* **41**: 1581–1591.
67. Wyss, C. (1982). *Chironomus tentans* epithelial cell line sensitive to ecdysteroids, juvenile hormone, insulin and heat shock. *Expl. Cell. Res.* **139**: 309–319.
68. Turberg, A. and Spindler, K-D. (1992). Properties of nuclear and cytosolic ecdysteroid receptors from an epithelial cell line from *Chironomus tentans*. *J. Insect. Physiol.* **38**: 81–91.
69. Turberg et al. (1992).
70. Rauch, P. et al.(1998). Ecdysteroid receptor and ultraspiracle from *Chironomus tentans* (Insecta) are phosphoproteins and are regulated differently by molting hormone. *Insect Biochem. Mol. Biol.* **28**: 265–275.
71. Grebe, M., Rauch, P., and Spindler-Barth, M. (2000). Characterization of subclones of the epithelial cell line from Chironomus tentans resistant to the insecticide RH 5992, a non-steroidal moulting hormones agonist. *Insect. Biochem. Mol. Biol.* **30**: 591–600.
72. Smagghe, G., Dhadialla, T.S., and Lezzi, M. (2002). Comparative toxicity and ecdysone receptor affinity of non-steroidal ecdysone agonists and 2-hydroxyecdysone in *Chironomus tentans*. *Insect Biochem. Mol. Biol.* **32**: 187–192.
73. Sibley, P.K., Benoit, D.A., and Ankley, G.T. (1998). Life cycle and behavioural assessments of the influence of substrate particle size on *Chironomus tentans* (Diptera: Chironomidae) in laboratory assays. *Hydrobiologia* **361**: 1–9.
74. Burton, G.A. (1996). Interlaboratory study of precision: *Hyalella azteca* and *Chironomus tentans* freshwater sediment toxicity assays. *Environ. Toxicol. Chem.* **15**(8): 1335–1343.

CIENEGA

Leo S. Leonhart
Hargis + Associates, Inc.
Tucson, Arizona

Cienega is both a hydrologic and ecological term used primarily in the southwestern United States and northwestern Mexico to designate a riparian marshland. The term cienega derives from the Spanish phrase *cien aguas* (literally "hundred waters" or more figuratively, "place of one hundred springs"). The modern Spanish equivalent is *ciénaga* ("marsh, swamp").

Most often, the term cienega refers to broad riparian areas of slowly flowing or ponded water, populated by dense vegetation. Such hydrologic conditions are unique in the Southwest because of the limited and declining number of perennially flowing stream reaches in the arid Southwest. Cienegas may also result from the discharge of springs or other conditions involving shallow or discharging groundwater. The term has also been applied to more diverse aquatic habitats throughout areas of Hispanic influence.

In South America, cienega generally refers to floodplain lakes that have open water (1). Thus, any marshy area within the Chihuahuan-Sonoran-Mohave desert region may be called a cienega. In upland environments in Arizona, the term cienega has been used to refer to marshy to bog-like meadowlands fed by seepage, precipitation, or moving waters bordering headwater streams (1). A database maintained by the Arizona Chapter of The Nature Conservancy (TNCA) lists 160 named cienegas in the Arizona region, including two in New Mexico; 34 in Sonora, Mexico; and one in Chihuahua, Mexico. Efforts are underway to verify and characterize each site (2).

Like many other wetland habitats, cienegas have been characterized as vanishing landforms, owing primarily to certain anthropogenic factors. However, increasing awareness of the significance and uniqueness of wetlands in general has resulted in various programs to maintain and restore such areas.

HYDROLOGIC CONDITIONS

As in all wetland environments, perennial moisture at shallow soil depths, if not standing water, is requisite for a cienega. These conditions, however, are rare in arid or semiarid landscapes, where surface drainage is typically ephemeral and flows only in response to direct surface runoff. In such dry climates, the water table is generally found at considerable depth below the base of the streambed, and groundwater discharge is thereby unable to contribute baseflow. Ephemeral streams are often incised due to high loading during high intensity, short duration precipitation. Historical information indicates that arroyo cutting and incisement of many stream reaches in the Southwest began near the end of the nineteenth century from both climatic and anthropogenic changes. In contrast to typical ephemeral drainages, groundwater beneath cienegas is generally at, or very near the surface, thereby supporting baseflow and a variety of wetland species. Additionally, the cross-channel profile in cienegas is comparatively shallow and broad, sometimes extending thousands of feet across.

Cienegas can form as a result of various hydrologic, geologic, and/or ecological conditions. To sustain indigenous wetland species, cienegas require replenishment of water at rates in excess of evaporation and transpiration rates. This replenishment is supplied from shallow or discharging groundwater that occurs due to structural and/or topographic factors underlying the streambed. Hendrickson and Minckley (1) have noted various in-stream conditions that may form cienegas along stream courses. These include alternating aggradation and degradation by floods, causing alternating concave/convex stream profiling; beaver damming; slope breaks, such as at the toes of bajadas or alluvial fans; upfaulted bedrock causing localized thinning of alluvial material; channel base level adjustments with time; and impoundments caused by debris such as from slumps and landslides. Efforts to restore former cienega areas have involved constructing various types of dams and impoundments.

ENVIRONMENTAL CONDITIONS

Cienegas are wetland ecosystems and thereby represent transitional environments from terrestrial to aquatic systems. Classified wetland ecosystems require three conditions: (1) the land environment supports predominantly aquatic flora at least periodically, (2) the substrate comprises mostly saturated soils, and (3) submergence and/or saturation of the substrate persists at least for some time during the annual growing season (3). Hendrickson and Minckley (1) defined three types of cienegas in the Southwestern United States initially based on elevation. Their elevation tiers included

- High-elevation (>2,000 m) cienegas described as marshy to bog-like alpine and cold temperate meadowlands.
- Midelevation (1,000 to 2,000 m) cienegas resulting from perennial springs and headwater streams, typically found in semidesert grasslands and evergreen woodlands.
- Low-elevation (<1,000 m) cienegas occurring as subtropical marshes in oxbows, behind natural levees, and along margins of major streams.

TNCA has noted features and characteristics common to all tiers, including groundwater at or very near the soil surface; perennial surface water; occurrence within basins, usually within unincised channels; and dominant herbaceous, perennial flora, including terrestrial, aquatic, or semiaquatic species, where scattered, less dominant woody species occur (4). These characteristics further demonstrate that cienegas meet the wetland classification criteria of Cowardin et al. (4).

Because cienegas represent comparatively lush, moist areas, they offer sharp contrast to adjacent dryland ecology. The combination of slow to stagnant water, the baffle effect of dense emergent flora, and the senescence and decay of this vegetation results in accumulation of rich organic detritus in the cienega basin. This detritus comprises a nutrient-rich substrate that enhances the overall productivity of these wetland areas.

A typical botanical progression from wettest to driest zones in a desert cienega might include emergent flora such as reeds (cattails, sedges, rushes) within the wettest areas. Salt-tolerant species such as salt grass might be found along the margins of the wetted areas, where remnant salts accumulate after evaporation. Broadleaf trees, such as various species of willow, cottonwood, and the invasive tamarisk, if present, might be found farther upland. These species adapt to less saline, less saturated soil conditions. Still farther upland, as depth increases to the water table, one might find a progression toward mesquite bosques. The availability of water and resultant luxuriance of forage in such semiarid regions often result in increases in herbivore populations (Fig. 1). In turn, this has been cited as a potential factor in the demise of these habitats (1).

Figure 1. San Pedro River National Conservation Area cienega near St. David, Arizona. Note vegetation transition from sacaton grasses to senescent bullrush to woody species (photo by B. Scully).

Examples

Cienega Creek, Pima County, Arizona. Cienega Creek, located in southern Arizona, represents a discontinuous stream that has perennial flow along certain reaches. Perennial flow along this reach supports a rich riparian habitat that has prompted the county to designate a portion of the reach as Cienega Creek Natural Preserve. The cienega exists primarily as a result of bedrock structural control over groundwater flow. Gravimetric investigations within the Cienega Creek basin have indicated that flow occurs as a result of upwardly dipping, low permeability bedrock that results in thinning of the overlying alluvium so as to force groundwater to daylight through the streambed. These sudden changes in depth to bedrock have been associated with high-angle faulting (5,6). A longitudinal profile of this regime is illustrated in Fig. 2.

Southern Arizona. Cienegas similar to Cienega Creek have been noted historically throughout southern Arizona including the San Simon, Sulphur Springs, Aravaipa, and San Pedro Valleys. Extensive archaeological research within these areas has shown that human settlements in the vicinity of such areas date back to 8000 B.C. and use continued through recent native American cultures (1,7,8). Accounts of the Mormon Battalion tell of grassy bottomlands with "... salmon trout, up to three feet long," "... profuse and luxuriant [cane grass up to six feet high]," along the San Pedro River (9). Accounts of stream conditions in the San Pedro River near St. David, Arizona, in the late nineteenth century suggest that the channel profile was flat and marshy, in contrast to its present-day incisement (10). The incisement has been attributed to a variety of anthropogenic factors including groundwater pumping, grazing, agriculture, and removal of beaver dams. There have been also been accounts of significant changes in the overall hydrologic conditions attributed to a 1887 earthquake centered to the south in Sonora, Mexico. Reports of liquefaction, new springs, disappearance of former surface water and springs, and changes in water levels in wells have all been noted (11).

Cienega de Santa Clara, Colorado River Delta. The Cienega de Santa Clara is a remnant of the former desert estuary of the Colorado River. This nearly 15,000-acre wetland habitat occupies an oxbow area of the Colorado River. Diversions from the river now limit its replenishment; the remaining flow represents irrigation drainage pumped

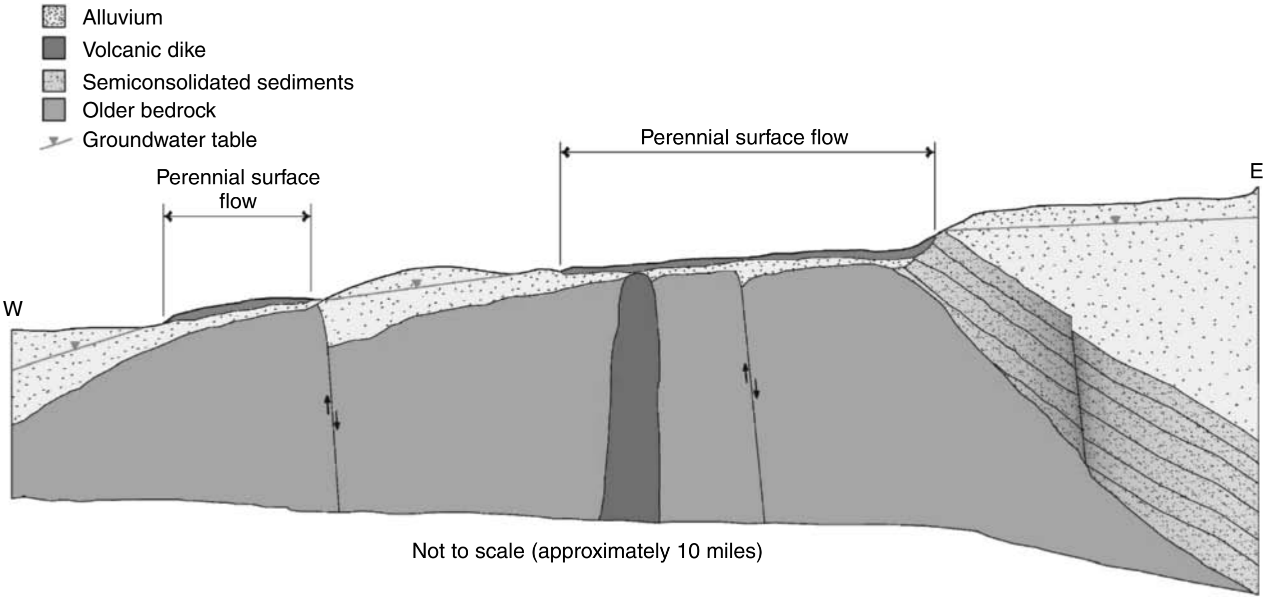

Figure 2. Conceptualization of structural components producing intermittent flow in Cienega Creek, Pima County, Arizona (5).

from agricultural land in the United States. Thus, the existence of the cienega is primary the result of anthropogenic water management. Owing to the high rate of evaporation as well as the salinity of the influent agricultural drainage, the Cienega is mildly brackish. Recent concerns over the survivability of the habitat have arisen from consideration of diverting the irrigation outflow through a desalination plant in the United States (12). This would have the combined impact of reducing the quantity of flow and increasing salinity as a result of splicing in saline blowdown from the desalination plant (13).

RELATED TERMS

Other equivalent or related terms are billabong, bog, bottomland, carr, fen, heath, marsh, pantano, riparian, mangrove, mire, moor, muskegs, oxbow, peatland, playa, pocosin, pothole, salt marsh, slough, swamp, vernal pool, and wetland. These terms relate to various types of wetland environments, but the hydrologic conditions that form them and the ecological communities that inhabit them may differ greatly. Many terms are local and regional.

Within the Southwestern United States and Northwestern Mexico, there have been numerous such water-related terms referred to in the journals of Spanish explorers, including those of Father Eusibio Kino, a late seventeenth century missionary, who made more than 150 *entradas* (“journeys” or “reconnaissance”) across the Sonoran desert during his 24-year residence. During such journeys along *El Camino del Diablo*, Kino made various references to *cienegas, tinajas*, or *tanques* (“water tanks” or “jars”), and *aguajes* (“watering holes”) along the way.

BIBLIOGRAPHY

1. Hendrickson, D.A. and Minckley, W.L. (1984). Ciénegas—vanishing climax communities of the American Southwest. *Desert Plants* **6**(3): [Special Issue].
2. TNCA. (2003). *Written Correspondence from Mr. Dale S. Turner, Conservation Planner, The Nature Conservancy, Arizona Chapter, to Dr. Leo S. Leonhart, Hargis + Associates, Inc.*, Tucson, Arizona, October 15, 2003.
3. Cowardin, L.M., Carter, V., Golet, F.C., and LaRoe, E.T. (1979). *Classification of Wetlands and Deepwater Habitats of the United States*. U.S. Department of the Interior, Fish and Wildlife Service. Washington, DC.
4. TNCA. (2003). *Cienegas of Arizona and the Apache Highlands Ecoregion, Including Portions of New Mexico, Sonora, and Chihuahua*. Submitted by Stephanie Weinstein to the Nature Conservancy of Arizona, June 20, 2003.
5. Ellett, W.J. (1994). *Geologic Controls on the Occurrence of Water in the Lower Cienega Creek Basin*. Master’s Thesis. University of Arizona, Department of Hydrology and Water Resources.
6. Roudebush, E.M. (1996). *The Influence of Bedrock on Perennial Streamflow in the Upper Cienega Creek Basin, Pima County, Arizona*. Master’s Thesis. University of Arizona, Department of Hydrology and Water Resources.
7. Eddy, F.W., Cooley, M.E., Martin, P.S., and Huckell, B.B. (1983). *Cultural and Environmental History of Cienega Valley in Southeastern Arizona*. Anthropological Papers No. 43. University of Arizona Press, Tucson, AZ.
8. Huckell, B.B. (1995). *Of Marshes and Maize: Preceramic Agricultural Settlements in the Cienega Valley, Southeastern Arizona*. Anthropological Papers No. 59. University of Arizona Press, Tucson, AZ.
9. Cooke, P.St.G., (1938). Cooke’s Journal of the March of the Mormon Battalion, 1846–1847. Cited in *Exploring Southwest Trails, 1846–1854. Southwest History Series*, V. 7. Cited in Reference 1.
10. Cooke, R.U. and Reeves, R.W. (1976). *Arroyos and Environmental Change in the American Southwest*. Oxford University Press. Cited in Reference 1.
11. DuBois, S.M. and Smith, A.W. (1980). *The 1887 Earthquake in San Bernardino Valley, Sonora: Historic Accounts and Intensity Patterns in Arizona*. Arizona Bureau of Geology and Mineral Technology, Special Paper No. 3.

12. Zengel, S.A. et al. (1995). Cienega de Santa Clara, a remnant wetland in the Rio Colorado Delta (Mexico): Vegetation distribution and the effects of water flow reduction. *Ecol. Eng.* **4**: 19–36.
13. Newcom, S.J. (1999). Deciding About the Colorado River Delta. Water Education Foundation, River Report. Accessed at http://www.sci.sdsu.edu/salton/DecidingAboutCoR%20Delta.html.

TIME-AREA AND THE CLARK RAINFALL-RUNOFF TRANSFORMATION

Bahram Saghafian
Soil Conservation and
Watershed Management
Research Institute
Tehran, Iran

INTRODUCTION

In 1945, Clark introduced a synthetic unit hydrograph whose variants have been widely applied in rainfall-runoff transformation. The Clark model, designed to consider the shape of the watershed, has two components. First, the excess rainfall is translated to the outlet based on a time-area (TA) histogram. Second, attenuation due to channel storage effects is achieved by routing the time-area hydrograph into a concentrated linear reservoir. The TA method is linear and may be considered an extension of the well-known rational method. Although Clark's original concept was semidistributed over space and invariant in time, some of its recent extensions are fully distributed (e.g., (1–4)) and time variable (4).

PRINCIPLES OF THE CLARK MODEL

Let us assume that the watershed of interest, subject to a given excess rainfall, can be divided into several zones ordered by the timing of their contributions to the outlet runoff (Fig. 1). Following the convolution rule of the unit hydrograph theory, the (excess) rainfall-runoff may be transformed by

$$Q_j = \sum_{k=1}^{j_0} E_k A_{j-k+1} \tag{1}$$

where j is the time step number, Q is the runoff outlet discharge, E is the excess rainfall intensity, and A is the area of the zone. In Eq. 1, $j_0 = j$ when j is less than or equal to the time step corresponding to rainfall duration (j_r), and $j_0 = j_\mathrm{r}$ otherwise. The application of Eq. 1 requires knowledge of the time-area histogram (TAH) which represents the time order of zone areas contributing to outlet runoff. Derivation of the TAH is a critical step in the Clark procedure and will be discussed later. Equation 1 is also known as the time-area method and constitutes the translational component of the rainfall-runoff transformation.

A fictitious linear reservoir comprises the second component of the Clark unit hydrograph which represents the attenuation effects. The storage–outflow relation of the linear reservoir is usually expressed by

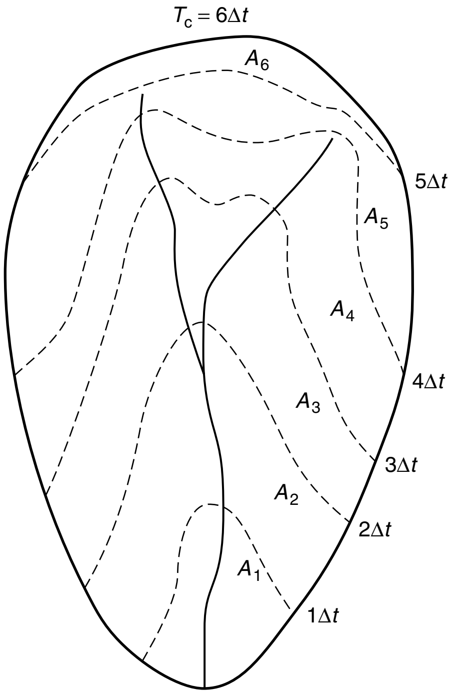

Figure 1. A watershed divided into zones (A_i) by isochrones (dashed lines).

$$S = KO \tag{2}$$

where S is the storage, K is the storage coefficient, and O is the outflow discharge. K is a characteristic of the reservoir and has the dimension of time. The Clark hydrograph is thus obtained by routing the time-area hydrograph of Eq. 1 into the reservoir of Eq. 2. One can write the continuity equation for the reservoir as follows:

$$Q - O = \frac{dS}{dt} \tag{3}$$

where Q is the inflow discharge of Eq. 1 and t is time. For a given time step Δt, Eq. 2 expands to:

$$\frac{Q_1 + Q_2}{2} - \frac{O_1 + O_2}{2} = \frac{S_2 - S_1}{\Delta t} \tag{4}$$

in which indexes 1 and 2 refer to the beginning and end of the time step. Combining Eqs. 2 and 4 results in the following form:

$$O_2 = C_0 Q_2 + C_1 Q_1 + C_2 O_1 \tag{5}$$

where

$$C_0 = \frac{\Delta t}{2K + \Delta t}$$

$$C_1 = \frac{\Delta t}{2K + \Delta t}$$

$$C_2 = \frac{2K - \Delta t}{2K + \Delta t} \tag{6}$$

Note that $C_0 = C_1$ and the summation of C coefficients in Eq. 6 is equal to one.

Example

The following example illustrates the calculation of a unit hydrograph and a total hydrograph based on the Clark model. Assume that the time-area histogram of a small 1000-ha watershed is given in Table 1. Also, $K = 30$ min and $\Delta t = 15$ min.

1. Determine the 15-min Clark unit hydrograph for 1 cm unit depth.
2. Determine the discharge hydrograph corresponding to an excess rainfall intensity of 10 mm/h over the (0–30) min period and 5 mm/h over the (30–45) min period.

Solution

1. Table 2 summarizes the unit hydrograph calculation based on Eqs. 1 and 5
2. There are two possible approaches for solving this part. One is to use the unit hydrograph, as shown in Table 3. The second is to bypass the unit hydrograph and calculate the total hydrograph directly from the time-area histogram, storage coefficient, and excess rainfall hyetograph. This is shown in Table 4.

Table 1. Time-Area Histogram

Time, min	0–15	15–30	30–45	45–60
Area, ha	100	300	500	100

Table 2. Unit Hydrograph Calculation

Time, min	E^*A, ha.cm/h	Q, cms	C_0Q_2, cms	C_1Q_1, cms	C_2O_1, cms	O, cms
0	0	0.00	0.00	0.00	0.00	0.00
15	400	11.11	2.22	0.00	0.00	2.22
30	1200	33.33	6.67	2.22	1.33	10.22
45	2000	55.56	11.11	6.67	6.13	23.91
60	400	11.11	2.22	11.11	14.35	27.68
75	0	0	0	2.22	16.61	18.83
90	0	0	0	0	11.30	11.30
105	0	0	0	0	6.78	6.78
120	0	0	0	0	4.07	4.07
135	0	0	0	0	2.44	2.44
150	0	0	0	0	1.46	1.46
165	0	0	0	0	0.88	0.88
180	0	0	0	0	0.53	0.53
195	0	0	0	0	0.32	0.32
210	0	0	0	0	0.19	0.19
225	0	0	0	0	0.11	0.11
240	0	0	0	0	0.07	0.07
255	0	0	0	0	0.04	0.04
270	0	0	0	0	0.02	0.02
285	0	0	0	0	0.01	0.01
300	0	0	0	0	0.01	0.01

Table 3. Total Hydrograph Derived from the Unit Hydrograph

Time,	Discharge, cms			
min	1st period	2nd period	3rd period	Total
0	0	0	0	0.00
15	0.56	0	0	0.56
30	2.56	0.56	0.00	3.11
45	5.98	2.56	0.28	8.81
60	6.92	5.98	1.28	14.18
75	4.71	6.92	2.99	14.62
90	2.82	4.71	3.46	10.99
105	1.69	2.82	2.35	6.87
120	1.02	1.69	1.41	4.12
135	0.61	1.02	0.85	2.47
150	0.37	0.61	0.51	1.48
165	0.22	0.37	0.31	0.89
180	0.13	0.22	0.18	0.53
195	0.08	0.13	0.11	0.32
210	0.05	0.08	0.07	0.19
225	0.03	0.05	0.04	0.12
240	0.02	0.03	0.02	0.07
255	0.01	0.02	0.01	0.04
270	0.01	0.01	0.01	0.02
285	0	0.01	0.01	0.01
300	0	0	0	0.01

CLARK MODEL PARAMETERS

Application of Clark's model requires knowledge of some key parameters that are physically or conceptually derivable from watershed characteristics, or, alternatively, from historic rainfall-runoff data. These parameters include the time of concentration (T_c), the time-area histogram (TAH), and the linear reservoir storage coefficient (K).

Several definitions have been proposed for the time of concentration. One of the oldest definition introduces T_c as the time for (a drop of) water to travel from the most distant part of the watershed to the outlet. Some literature emphasizes that the distance has a hydraulic base, not necessarily a geometric base. Another definition deals with the wave travel time from the most hydraulically remote point to the outlet of the watershed, which is essentially the time elapsed for all parts of the watershed to contribute to the direct runoff at the outlet. In graphical analysis of a combined hyetograph-hydrograph, T_c is also defined as the time between the end of excess rainfall and the inflection point on the falling limb of the hydrograph (Fig. 2). It is assumed that the inflection point indicates the end of direct runoff. Time of concentration is essentially a function of rainfall intensity and watershed characteristics. Refer to Saghafian (5) in the *Encyclopedia of Water* for further information about time of concentration and how it differs from time to equilibrium.

A fundamental step in applying the Clark model is constructing the TAH. But first, we have to turn our attention to the concept of an isochrone. An isochrone is a contour of equal travel time to the outlet or a line connecting all watershed points that have the same travel time to a common outlet. The isochrones start and end only at the watershed boundary. The time-area histogram, also known as the time–area graph,

Table 4. Total Hydrograph Derived Directly from the Time-Area Histogram

Time, min	E_1A, ha.cm/h	E_2A, ha.cm/h	E_3A, ha.cm/h	$Q = \Sigma EA$, ha.cm/h	Q, cms	C_0Q_2, cms	C_1Q_1, cms	C_2O_1, cms	O, cms
0	0	0	0	0	0	0	0	0	0
15	1000	0	0	1000	2.78	0.56	0	0	0.56
30	3000	1000	0	4000	11.11	2.22	0.56	0.33	3.11
45	5000	3000	500	8500	23.61	4.72	2.22	1.87	8.81
60	1000	5000	1500	7500	20.83	4.17	4.72	5.29	14.18
75	0	1000	2500	3500	9.72	1.94	4.17	8.51	14.62
90	0	0	500	500	1.39	0.28	1.94	8.77	10.99
105	0	0	0	0	0	0	0.28	6.60	6.87
120	0	0	0	0	0	0	0	4.12	4.12
135	0	0	0	0	0	0	0	2.47	2.47
150	0	0	0	0	0	0	0	1.48	1.48
165	0	0	0	0	0	0	0	0.89	0.89
180	0	0	0	0	0	0	0	0.53	0.53
195	0	0	0	0	0	0	0	0.32	0.32
210	0	0	0	0	0	0	0	0.19	0.19
225	0	0	0	0	0	0	0	0.12	0.12
240	0	0	0	0	0	0	0	0.07	0.07
255	0	0	0	0	0	0	0	0.04	0.04
270	0	0	0	0	0	0	0	0.02	0.02
285	0	0	0	0	0	0	0	0.01	0.01

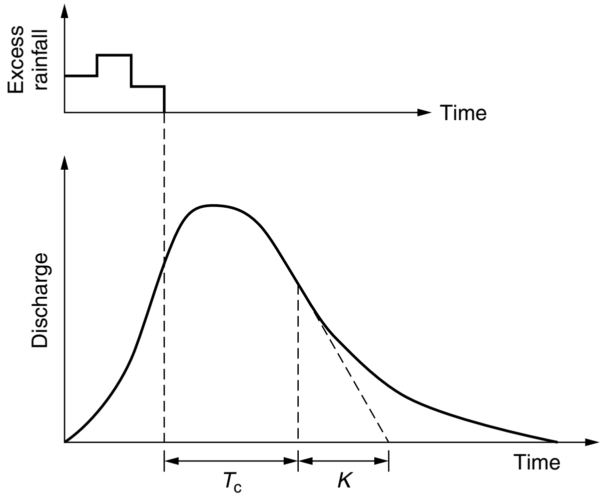

Figure 2. Graphical representation of time of concentration (T_c) and storage coefficient (K).

is a discrete time histogram of the incremental areas enclosed by adjacent isochrones, which are apart by a finite time interval Δt (Fig. 3a). TAH represents the time distribution of sequential areas participating in surface runoff generation at the outlet. To construct the TAH, the watershed time of concentration must be divided into a number of equal time intervals, which are the travel time differences between adjacent isochrones. The isochrone spacing and construction of a TAH will be discussed later.

Some literature also introduces a concept known as a time-area diagram. A time-area diagram is the plot of area, on the y axis, whose travel time is less than or equal to a given time, shown on the x axis (Fig. 3b). From the TA diagram, one can determine the area that contributes to the outlet runoff for any given time. A TA diagram is a continuous time function. A time-area curve (Fig. 3c), or TAC, is the derivative of a TA diagram and is equivalent to the TAH when Δt tends to zero.

The TAH can be looked upon as the basis for a unit hydrograph. When an excess rainfall of 1.0 unit depth over Δt time duration is applied to a TAH through Eq. 1, then a unit hydrograph of Δt duration is produced. For example, TAH's y axis, say, in m^2, multiplied by 0.01 m (1.0 cm) of rainfall depth and divided by the value of Δt, say, in seconds, yields a unit hydrograph whose ordinates represent discharge in m^3/s and its volume is equal to 1.0 cm or 0.01 m across the total watershed area. When both the rainfall duration and the TAH Δt tend to zero (i.e., the isochrone time difference becomes smaller and smaller so that the TAH turns into a continuous time-area curve), an instantaneous unit hydrograph is formed.

Clark's linear reservoir storage coefficient is a physical character of the watershed that cannot be directly measured. It is, however, graphically defined as the negative of the discharge divided by the slope of the hydrograph at the point of inflection on the recession limb.

$$K = \frac{-Q}{dQ/dt} \tag{7}$$

According to Fig. 2, K is equal to the time interval between the point of inflection and the point at which the tangent to the point of inflection intersects the time axis.

An alternate way of determining Clark parameters for gauged watersheds is to use recorded rainfall-runoff to calibrate a Clark-based model for unknown parameters. For ungauged watersheds, regional empirical relations may be developed using the rainfall-runoff data of gauged watersheds in the region. In different parts of Australia, for example, relations of the following type are commonly adopted (6,7):

$$T_c = m_0 L^{m_1} S^{m_2} \tag{8}$$

$$K = n_0 L^{n_1} S^{n_2} \tag{9}$$

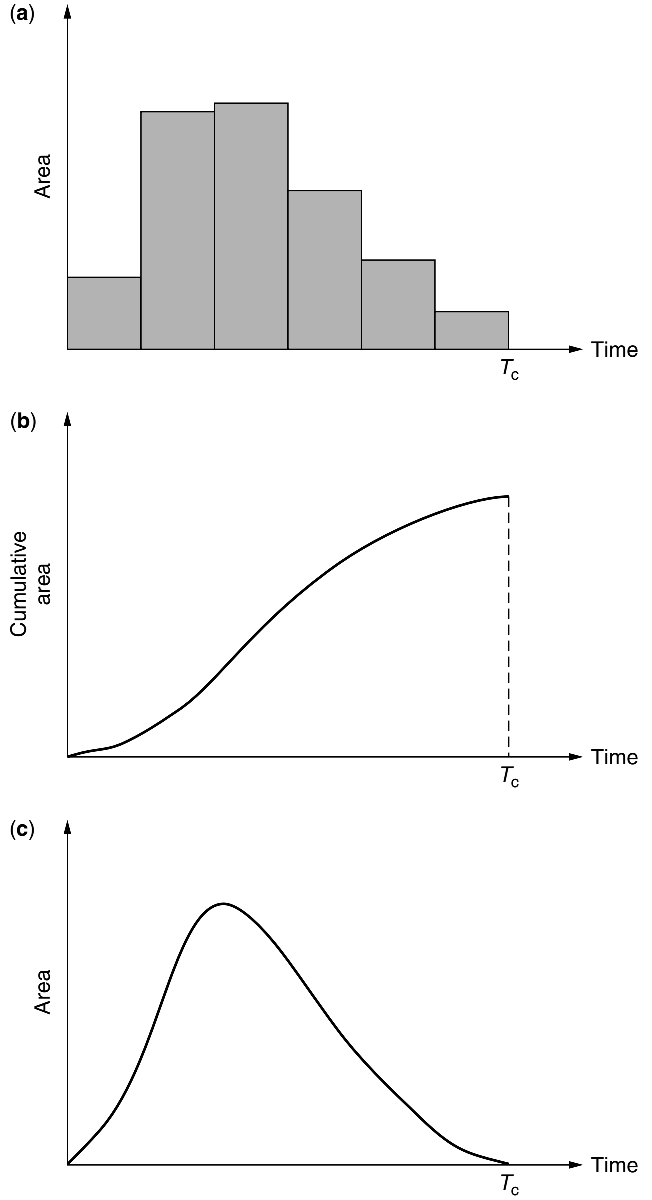

Figure 3. (**a**) Time-area histogram, (**b**) time-area diagram, and (**c**) time-area curve.

where L is the main stream length, S is the main stream slope, and m's and n's are constants for homogeneous regions.

A COMPREHENSIVE ISOCHRONE SPACING METHOD

Saghafian and Julien (8) derived a general formula for the travel time from the most hydraulically remote point to the watershed outlet:

$$T_{\mathrm{w}} = \int_0^L (1-\gamma)\left(\frac{a_1}{Q_e}\right)^{\gamma}\left(\frac{n}{a_2^{2/3}S_0^{1/2}}\right)^{1-\gamma} dx \qquad (10)$$

where T_{w} is the total travel time equal to the time to equilibrium, L is the total length along the hydraulically longest flow path; x is the distance measured along the flow path; Q_e is the equilibrium discharge; a_1, a_2, and γ are flow cross-sectional parameters; S_0 is the bed slope corresponding to the *kinematic time to equilibrium*; and n is the Manning roughness coefficient. T_{w} in Eq. 10 is also a theoretical form of time of concentration. Note that travel time in Eq. 10 can also be computed for points other than the watershed outlet by changing the upper integral limit. Therefore, the travel time may be computed for all locations in the watershed using Eq. 10, and isochrones can be derived.

The equilibrium discharge Q_{e} passing through a given cross section at location x, is equal to the spatially integrated excess rainfall rate over the drainage area of that cross section. In mathematical terms, Q_{e} may be expressed by

$$Q_{\mathrm{e}}(x) = \int_0^{A(x)} E\,dA \qquad (11)$$

where $A(x)$ is the upslope drainage area or flow accumulation at distance x and E is excess rainfall intensity at any point draining to x. To solve Eq. 10, the spatial variability of excess rainfall and that of the drainage area along the hydraulically longest flow path must be known.

The total travel time, T_{w}, can be separated into the travel time for overland flow, T_{wo}, and for channel flow, T_{wc}. If the cross-sectional area and the hydraulic radius can be expressed by relations of the form $A_x = a_1 h^{b_1}$ and $R = a_2 h^{b_2}$, where a_1, a_2, b_1, and b_2 are constants for a given cross section and h is the flow depth, then,

$$\gamma = \frac{2b_2}{2b_2 + 3b_1} \qquad (12)$$

when Manning's equation is used. For overland flow and flow in wide channels, a_2, b_1, and b_2 equal unity and $\gamma = 2/5$ (8). If the excess rainfall intensity E is uniform in space, then $Q_{\mathrm{e}} = EA$, and the term $1/E$ can be taken out of the integral. Thus, the travel time is proportional to $E^{-\gamma}$. For watersheds that have wide channels,

$$T_{\mathrm{w}} = CE^{(-2/5)} \qquad (13)$$

where C is a watershed hydrogeomorphologic index conglomerating spatially distributed characteristics such as surface roughness, slope, flow length and flow accumulation, drainage pattern, and channel geometry. Comparing Eq. 13 with the available empirical formulas of time of concentration, one observes that most empirical formulas ignore the effect of rainfall intensity.

SIMPLIFIED ISOCHRONE SPACING METHODS

Laurenson (9) proposed an early method of isochrone spacing. He assumed that the travel time from any point in the watershed to the outlet is proportional to $\sum(l_i/{S_i}^{1/2})$, where l_i is the length of the ith reach along the flow path of that point toward the outlet and S_i is the slope of the reach. The summation is over the entire flow path. The procedure starts by selecting a number of points on all elevation contours distributed uniformly. The flow paths to the outlet originating from all points are plotted. The flow reaches are bounded by adjacent contours that have uniform elevation differences of, say, H. Therefore, the value of ${l_i}^{3/2}/H^{1/2}$ may be computed for any reach, where the slope has been substituted by H/l_i. That is to say that the travel time is proportional to $l^{3/2}$, provided that the reaches are measured between adjacent elevation contours. Once the summation values are computed for all points, they are divided by the largest summation value which represents the maximum travel time. This yields the relative travel times for the points selected. Now the isochrones can be plotted by interpolating the point values. It is easy now to develop the TAH. Geographic information systems (GIS) can be deployed to facilitate and improve the procedure. A difficulty of this method is that an independent estimate of the absolute value of maximum travel time, or the time to equilibrium, is still required for further application of the time-area method.

Pilgrim (10) argues that, based on some tracer observations in a 96-acre watershed with a dense stream network in Australia, the ratios of velocity of different reaches to that of the longest reach vary markedly with discharge. This amounts to variable distribution of isochrones corresponding to different floods. This variability is weaker for medium to high discharges observed, however. The tracing data collected by Pilgrim (10) confirmed that flood velocities tend to increase slightly in a downstream direction through the watershed, despite decreasing slope.

Some methods rely on physiographic and geomorphological parameters to determine isochrone spacing. Pilgrim (10) examined several isochrone spacing parameters versus tracer data corresponding to two discharges collected in a small watershed. Four of the parameters were as follows:

$$\sum l_i \tag{14a}$$

$$\sum \frac{l_i}{{S_i}^{1/2}} \tag{14b}$$

$$\sum \frac{l_i}{{S_i}^{1/2}{R_i}^{2/3}} \tag{14c}$$

$$\sum \frac{n_i l_i}{{S_i}^{1/2}{R_i}^{2/3}} \tag{14d}$$

where l_i is the length of the ith reach, S_i is the slope, R_i is the hydraulic radius, and n_i is the Manning roughness. Hydraulic radius was estimated through field observations and geometric relations corresponding to bankful discharge. The first parameter assumes uniform velocity throughout the stream reaches. It was proposed by Clark (11) but was later neglected. Other parameters consider those reach characteristics believed to represent velocity. Parameter 14b is similar to that proposed by Laurenson (9). The last two more complex parameters gave the best overall estimates, and the commonly used second parameter was poorest. The first parameter, simply derived from topographic maps, was of accuracy comparable to that of more complex parameters. These parameters give the relative location where isochrones intersect with streams. The method must be extended somehow to overland parts of the watershed so that the isochrones can be extended to watershed boundaries. Even then, the value of time of concentration or the maximum travel time must be known so that the time-area transformation can be applied.

If no specific TAH is available, HEC (12) proposed a simple dimensionless cumulative time-area diagram as follows:

$$\begin{aligned} A^* &= 1.414(T^*)^{1.5} && \text{if } 0 < T^* < 0.5 \\ 1 - A^* &= 1.414(1 - T^*)^{1.5} && \text{if } 0.5 \le T^* < 1 \end{aligned} \tag{15}$$

where A^* is the cumulative area as a fraction of total watershed area and T^* is the fraction of T_c. This equation has been implemented in the HEC-1 model.

Kull and Feldman (3) assumed that travel time for each cell in a watershed is simply proportional to the time of concentration scaled by the ratio of travel length of the cell over the maximum travel length, that is, the average velocity of runoff traveling from any point to the outlet is assumed uniform and constant. Each cell's excess rainfall is then lagged to the outlet by the cell's travel length. Travel time in overland and in channels follows similar proportionality, and travel length and the watershed time of concentration must be determined a priori. The method is equivalent to the distributed form of the first parameter considered by Pilgrim (10). This approach has been implemented in the HMS model (13), a new version that replaces the HEC-1 model.

TIME VARIABLE ISOCHRONE METHOD

Rainfall intensity generally varies over time, so travel times will inevitably change. That is to say that the assumption of storm-independent unit hydrograph ordinates, represented by A's in Eq. 1, is not valid and time variability actually occurs in isochrone maps and TAHs. The use of the variable isochrone technique in rainfall-runoff transformation was first proposed by Saghafian et al. (4). In this technique, isochrone spacing and time-area histograms are updated at those time steps where rainfall intensity changes. Any TAH derived then corresponds to a given rainfall intensity and generates a partial hydrograph. The total hydrograph is derived by convoluting partial hydrographs. The numerical procedure is briefly described as follows.

The watershed is first discretized by a raster grid. The time-variable isochrone algorithm consists of four components. The first component derives terrain-based features by taking digital topographic data, usually in vector format, as input and generates raster terrain maps such as a depressionless digital elevation model (DEM), flow direction, and flow accumulation (i.e., an upslope drainage area). The temporal and spatial distributions of excess rainfall are determined by the second component, which requires soil infiltration parameters as well as the spatiotemporal variation of rainfall intensity. The third component uses maps derived in conjunction with the map of the roughness coefficient to develop differential (pixel scale) and total cumulative travel-time-to-outlet maps, based on Eq. 10, and a time series of isochrone maps corresponding to various excess rainfall intensities. The fourth component performs TA convolution to calculate the total runoff hydrograph. This hydrograph can then be routed into the linear reservoir.

The third component is a critical part of the computations. To generate a cumulative travel-time map, isochrones of equal travel times are derived, and the areas bounded by adjacent isochrones are determined. Aggregate time-area histograms may be sufficient for uniform excess rainfall, but the spatial extent of the areas in a given travel-time range is quite important in runoff computation when dealing with spatially variable excess rainfall. In the latter case, the equilibrium discharge map of Eq. 11 is directly substituted in Eq. 10 for travel-time calculations. Then, instead of TAH, a "time-discharge" histogram (TDH) is constructed based on the equilibrium discharge map divided by isochrones. TDH itself represents the runoff hydrograph for any period of constant excess rainfall intensity.

Note that the third component must be repeated N times, where N is the number of excess rainfall intensity maps. This is equal to the number of excess rainfall intervals of constant intensity. The final output is N isochrone maps. However, in watersheds where with wide channels subject to uniform excess rainfall, we need to execute the third module once. After derivation of the TAH for any value of excess intensity (say, E_1), the TAH of, say, E_2 intensity may be easily produced by scaling the time axis by the ratio $(E_2/E_1)^{-0.4}$. Resampling of the time axis of all TAHs may be necessary to achieve a common time interval.

In the fourth component, N incremental hydrographs corresponding to N isochrone maps are determined by convolution. These incremental hydrographs, delayed by their corresponding excess intensity times, are then superimposed to yield the total hydrograph.

GENERAL REMARKS

1. One of the major advantages of the Clark model is that the temporal rainfall pattern can be accounted for. Using the newly proposed Clark-based models discussed before, the spatial distribution of rainfall may also be considered in runoff computations.
2. It is possible to derive a discharge hydrograph directly from a TAH and an excess rainfall hyetograph. This bypasses the Clark unit hydrograph and is advantageous when only a single hydrograph is to be computed.
3. Clark's original idea has been upgraded to a fully distributed time-variant approach by Saghafian et al. (4). The proposed approach can handle the spatial distribution of excess rainfall, and it can also consider the variation of travel time due to the dynamics of excess rainfall.

BIBLIOGRAPHY

1. Maidment, D.R. (1993). Developing a spatially distributed unit hydrograph by using GIS. *Proc. HydroGIS 93*, IAHS Publ. no. 211.
2. Muzik, I. (1996). Lumped modeling and GIS in flood predictions. In: *Geographical Information Systems in Hydrology*. V.P. Singh and M. Fiorentino (Eds.). Kluwer Academic, Dordrecht, The Netherlands.
3. Kull, D.W. and Feldman, A.D. (1998). Evolution of Clark's unit graph method to spatially distributed runoff. *J. Hydrologic Eng.* **3**(1): 9–19.
4. Saghafian, B., Julien, P.Y., and Rajaie, H. (2002). Hydrograph simulation based on the variable isochrone technique. *J. Hydrol.* **261** (1–4): 193–203.
5. Saghafian, B. (2004). Time of concentration and travel time in watersheds. In: *Encyclopedia of Water*. J. Lehr (Ed.). John Wiley & Sons, New York.
6. Cordery, I., Pilgrim, D.H., and Baron, B.C. (1981). Validity of use of small catchment research results for large basins. *Civ. Eng. Trans. Inst. Eng. Aust.* **CE23**: 131–137.
7. Weeks, W.D. and Stewart, B.J. (1978). Linear and non-linear runoff routing for ungauged catchments. *Hydrol. Symp. Inst. Eng. Aust. Natl. Conf.* Publ. 78/9: 124–128.
8. Saghafian, B. and Julien, P.Y. (1995). Time to equilibrium for spatially variable watersheds. *J. Hydrol.* **172**: 231–245.
9. Laurenson, E.M. (1964). A catchment storage model for runoff routing. *J. Hydrol.* **2**: 141–163.
10. Pilgrim, D.H. (1977). Isochrones of travel time and distribution of flood storage from a tracer study on a small watershed. *Water Resour. Res.* **13**(3): 587–595.
11. Clark, C.O. (1945). Storage and the unit hydrograph. *Trans. Am. Soc. Civ. Eng.* **110**: 1419–1446.
12. Hydrologic Engineering Center (HEC). (1990). *HEC-1 Flood Hydrograph Package*. User's Manual.
13. Hydrologic Engineering Center (HEC). (2000). *Hydrologic Modeling System HEC-HMS*. Technical Reference Manual.

STREAM CLASSIFICATION

WAYNE D. ERSKINE
State Forests of New South Wales
Beecroft, New South Wales, Australia

Stream classification is important for characterizing the features of individual rivers or sections of river and also for determining the similarity of different streams. The latter is essential to assess which rivers are suitable for extrapolating of research results from elsewhere. Various schemes have been proposed for specific disciplines and purposes. There is no universally accepted stream classification scheme because they all require modification to apply to rivers different from those used to construct the scheme Therefore, classificatory schemes should be continuously revised and updated.

STREAM CLASSIFICATION HIERARCHY

A nested hierarchical scheme based on the channel network which places the stream within a drainage basin or watershed context is the best approach to stream classification. Table 1 defines the characteristics of, and range of scales for, each hierarchical level. Figure 1 clearly

Table 1. Hierarchical Framework of Stream Classification Proposed by References 1 and 2 and Illustrated in Fig. 1

Classification Level	Scale (River Length)[a] Small	Large	Essential Features
Watershed or drainage basin	$>10^3$ m	$>1,000$ km	Watershed/subwatershed scale boundary conditions
Valley segment	$>10^2$ m	100–1,000 km	Length of valley of essentially constant form and external conditions
Reach type	>10 m	10–100 km	Length of river exhibiting relatively homogeneous channel characteristics or a consistent pattern of repetitive/alternating characteristics
Channel unit or mesohabitat	$>10^0$m	0.1–10 km	Area of relatively homogeneous bed material, flow velocity, and flow depth
Microhabitat	10^{-1}m	<0.1 km	Patch of similar flow velocity, substrate, and cover

[a] Scale is only indicative and must be appropriate to basin or watershed size which is highly variable.

illustrates the relationships between each classification level, where each forms the environment of its subsystem at lower levels. The smallest linear spatial scale has been taken from Frissell et al. (1) who concentrated on small mountainous forested streams. However, the size of each level increases with watershed area and runoff. Each level of the classification scheme is now discussed.

Watershed

This is the highest level of the classification scheme (Table 1) and was proposed by Frissell et al. (1) for use in bioclimatic regions. Such regions have not been defined for many parts of the world. Nevertheless, the watershed can be characterized by other general watershed characteristics, such as the Koeppen–Gieger climatic region which provides an indicator of the natural hydrology of a river. In addition, geomorphic regions have been defined for most countries and provide a good indicator of the potential energy and bedrock confinement of a river. The nature of sediment supplied to rivers can be gauged from the generalized geology. Vegetative cover and land use partly determine sediment yields which exert a strong control on stream characteristics.

Valley Segment

Following Bisson and Montgomery (2), a valley segment refers to the valley form based on the dominant types of sediment input and transport processes (Table 1). The Frissell et al. (1) classes for valley segments have been expanded to allow for the full range of conditions likely to be experienced when attempting to classify streams. *Bedrock* valleys have high sediment transport capacities because of steep slopes and close lateral and vertical confinement. *Colluvial* valleys are partly filled with poorly sorted sediment supplied from the surrounding hillslopes. *Alluvial* valleys are partly filled with sediments deposited by the present or former river and by floodplain flows. *Lacustrine* valleys refer to those superimposed on lake sediments because many rivers discharge into terminal lakes or flow for some distance over exposed lake beds. *Aeolian* valleys are those where rivers cut through, or are dissipated in, sand dunes. *Artificial* valleys are those created by human construction such as urbanization, drainage, flood mitigation, and irrigation.

RIVER REACHES

River reaches are homogeneous lengths of channel within which hydrologic, geologic, and adjacent watershed

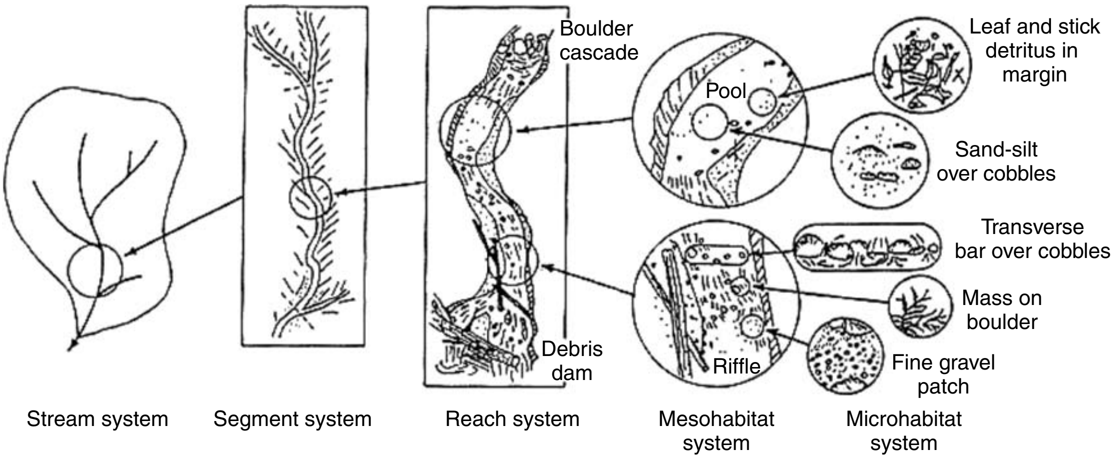

Figure 1. Hierarchy of a stream system proposed by Reference 1. The stream system is also called the watershed or drainage basin, the segment system is also called the valley segment, and mesohabitats are also called channel units (2).

surface conditions are sufficiently constant so that a uniform river morphology (3) or a consistent pattern of alternating river morphologies is produced (4). Bisson and Montgomery (2) proposed that channel reaches consist of relatively homogeneous associations of channel units, which distinguish them from adjoining reaches.

Reaches are typically 10–100 km long (Table 1). The core length of a reach is relatively easy to identify, but it is often more difficult to define precisely the boundaries of the reach due to its transitional nature (5). Bisson and Montgomery (2) proposed that provisional reach boundaries could be identified from topographic maps, vertical air photographs, and geology and soil maps, and then the channel units could be determined from selective field surveys. However, any method based on field surveys will be costly and time-consuming. There is also a problem of determining how to interpolate spatially from the sample site to the reach when the whole river is not sampled. Furthermore, a lot of detailed channel unit measurements often do not meaningfully define reach boundaries (4).

There is no agreement as to which criteria should be used to identify river reaches. Geomorphic criteria dominate most of the published schemes because reach identification is usually based on interpretation of vertical air photographs or other remotely sensed data. The schemes of References 3 and 6–10 used up to 25 different geomorphic criteria, and a similar trend is exhibited by ecologically based schemes.

Reach identification is an iterative process based on cross-checking between map and air photograph patterns and the channel units observed in the field. Therefore, periodic revisions of stream reaches should be expected as more data are collected, spatial variations in channel characteristics are better documented, and the controls on stream morphology are better understood (4).

Where appropriate, reaches can be given a formal name usually comprising three terms. The first term is based on a geographic name for a location within or near the reach; the second is a geomorphological or ecological descriptor for one of the dominant characteristics of the reach; and the third, when needed, is the term reach or zone. For example, the *Willis sand zone* on the Snowy River, Australia, is 93.5 km long and was named after Willis (the area at the New South Wales/Victorian border on the Snowy River, which is located within the reach) and the extensive sand storage that occurs in the long, relatively flat parts of this reach that are sandwiched between shorter, steeper bedrock sections (4,11).

Channel reaches can be placed into specific reach types such as those defined by Rosgen (8,9). However, Rosgen's classes must be expanded if they are to be applied universally so as to include reach types that are not currently recognized. For example, two such reach types are floodouts (12,13) and low-energy, straight channels (14,15) that are common in many parts of Australia. There are other reach types that also need to be added to the Rosgen scheme.

After identifying and naming channel reaches, each should then be allocated to one reach type for each valley segment. Each reach type can be given a simple descriptive name based on its geomorphic characteristics, such as sandstone gorge, anabranching river, or straight, low-energy river. This is more meaningful than the alphanumeric naming system of Rosgen (8,9). For example, "sandstone gorge" is a more informative name than "A1." Correlations between rivers should be based on reach type.

Channel Units

Hawkins et al. (16) defined channel units as "..quasi-discrete areas of relatively homogeneous depth and flow that are bounded by sharp physical gradients..". Recent research has demonstrated that each channel unit exhibits different physical characteristics and can also be associated with habitat-specific assemblages of fish species (17,18). Figure 2 outlines the channel units that are commonly recognized by geomorphologists and fish ecologists as useful for stream classification. A three-tiered system was proposed by Hawkins et al. (16) so that the user can select the level of channel unit resolution appropriate for the question being addressed. At the highest level, a bipartite division of channel units has been adopted, namely, slow water and fast water (Fig. 2). Each higher level channel unit has then been subdivided into two types, which are then further subdivided on the basis of hydraulic characteristics and the principal kind of habitat-forming structure or process (Refs. 2, 16, and 19).

Channel units are defined on the basis of low flow or baseflow conditions, and slow water habitats consist of various types of pools, which are formed by either erosion (*scour pools*) or damming (*dammed pools*). For the scour pool category (Fig. 2), additional classes have been added to those of Bisson and Montgomery (2) and Hawkins et al. (16) based on Webb and Erskine (19) and are outlined in Table 2. Each scour pool is now briefly defined. A *plunge pool* is a deep pool eroded at the base of a waterfall (2). Turner and Erskine (20) coined the term *"scour hole"* for large, deep holes eroded at the boundary of a bedrock gorge with alluvium. There is no waterfall at the transition from bedrock gorge to alluvium, and hence, the holes are not plunge pools because there is no vertical drop. Furthermore, scour holes are often sand-floored rather than lined by boulders or bedrock (20). Unlike Bisson and Montgomery (2), Hawkins et al. (16), and Grant et al. (21), *step pools* are differentiated from plunge pools because of the substantial difference in the surface area and depth of the pools and in the height of the fall at the upstream end of the pool. Step pools are usually orders of magnitude smaller and have smaller falls into them. They are small pools eroded by a fall over bedrock, boulders, and log steps in steep channels and are more closely spaced than free-formed pools associated with riffles (21). *Free-formed pools* are intended to cover standard pools eroded by high flows in many alluvial channels (22). They are rhythmically spaced five to seven channel widths apart and are closely associated with riffles, which are formed by the deposition of the sediment eroded from the pools by high bankfull flows (22). *Eddy pools* are eroded by turbulent eddies downstream of obstructions and are proportional to the size of the obstruction which is usually bedrock, boulders, or large woody debris. *Trench*

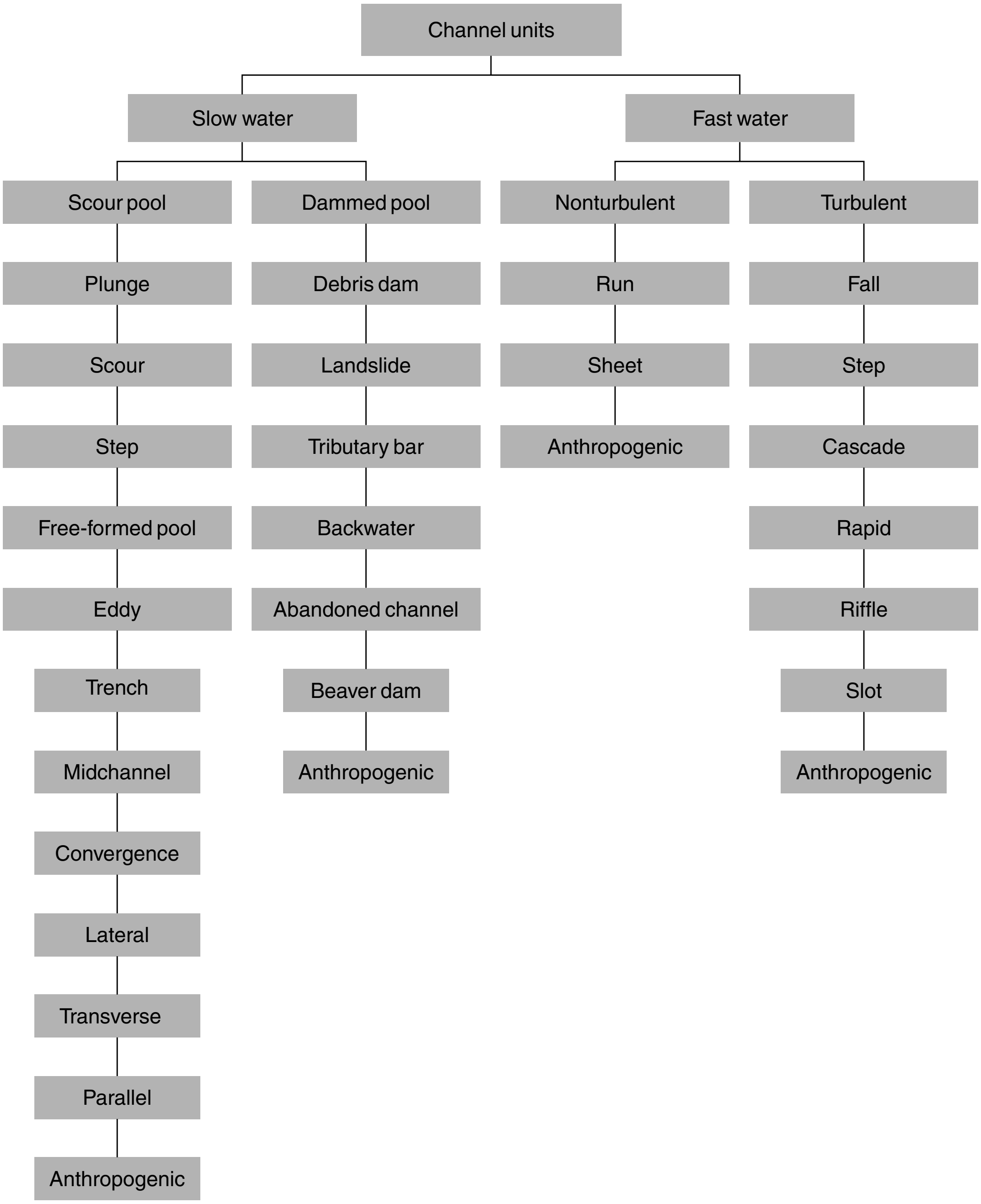

Figure 2. Channel units adapted from References 2, 16, and 19.

pools are uniformly eroded deep pools sandwiched between resistant, usually bedrock banks. They are a feature of the inner bedrock channels of Schumm et al. (23). *Midchannel pools* are formed by flow constrictions that concentrate scour in the center of a river. They differ from trench pools in that the constriction is oriented perpendicularly to the channel rather than parallel. *Convergence pools* form by the erosion of converging currents from two channels. These may be two independent channels (that is, a tributary joining the main stream) or two channels that form part of a braided or anabranching network. *Lateral* or *longitudinal scour pools* develop by erosion along or beside an obstruction or resistant material on one side of a river. They are common on bends where the river impinges against bedrock or large bank-attached pieces of large woody debris (19). *Transverse scour pools* are formed by scour under suspended, immobile pieces of bank-attached, channel spanning large woody debris or trees aligned transverse to the flow direction (19). *Parallel scour pools* are eroded by continuous scour at the head and around both sides of bed-attached large woody debris or large boulders aligned parallel to the flow direction and not attached to the banks (19). *Anthropogenic scour pools* are included for completeness to account for scour induced by human structures.

For the dammed pool category (Fig. 2), the classes of Bisson and Montgomery (2) and Hawkins et al. (16) have been slightly expanded (Table 2). *Debris dam pools*

Table 2. Types of Pools or Still-Water Channel Units

Pool Type	Formative Process	Common Locations
Plunge pool	Large-scale turbulent scour at base of waterfall	Immediately below high outcrops of resistant bedrock or the edge of escarpments
Scour hole	Large-scale turbulent scour at transition from resistant bedrock to alluvium	Where faults or monoclines cut across a river
Step pool	Small-scale turbulent scour below bedrock, boulder, or log steps	Steep upper reaches of channels or in reaches with high loadings of large woody debris
Free-formed pool	Large-scale, rhythmically spaced turbulent scour by flood flows in gravel and/or sand-bed streams	Alluvial reaches and often associated with bends
Eddy pool	Turbulent scour by eddies associated with obstructions, such as large boulders or large woody debris	Bedrock-confined reaches, where rivers flow across resistant bedrock or floodplain reaches with large riparian trees
Trench pool	Turbulent scour at channel constrictions between bedrock walls	Gorges, bedrock-confined reaches or localized outcrops of resistant bedrock
Midchannel pool	Turbulent scour at channel constrictions oriented at right angles to the river	Localized outcrops of thin resistant bedrock that strikes at right angles to the river
Convergence pool	Turbulent scour by converging currents from two channels	Tributary junctions, multichannel reaches (braided and anabranching rivers) or where flood chutes rejoin the main channel
Lateral longitudinal scour pool	Turbulent scour along an obstruction or resistant material	Wherever a river locally impinges against large woody debris or resistant bedrock; generally in higher energy reaches than transverse scour pools
Transverse scour pool	Turbulent scour under suspended, bank-attached, channel-spanning large woody debris or trees aligned transverse to flow	Wherever a river is flanked by a riparian forest and stream power is insufficient to remove suspended timber totally
Parallel scour pool	Continuous large-scale turbulent scour at the head and along both sides of bed-attached large woody debris or boulders aligned parallel to the flow	Wherever a river is flanked by a riparian forest, and stream power is insufficient to redistribute an obstruction from the center of the channel to either bank
Anthropogenic scour pool	Small- and large-scale turbulent scour induced by an engineering structure, such as on the downstream side of a weir or dam.	Essentially random depending on location of road crossings, sites with good foundations for a structure, etc.
Debris dam pool	Impoundment by large woody debris	Wherever a river is flanked by a riparian forest and stream power is insufficient to remove obstruction totally
Landslide dam pool	Impoundment by the debris transported by a landslide	Steep upland reaches where there are high rainfall intensities and cohesive soils
Backwater pool	Pools where one channel is dammed by another or where a local channel expansion prevents active flow	Can occur throughout river systems
Disconnected pool	Impoundment of a water-table window away from the main channel by a bar of sediment	Wherever bars are well developed
Abandoned channel pool	Ponding of a deep section of a former river course	Where floodplains are well developed
Beaver dam pool	Impoundment of water by a woody barrier constructed by beavers (*Castor canadensis; Castor fiber*)	Where beaver colonies are present
Anthropogenic dammed pool	Impoundment of water by an engineering structure, such as a weir or a causeway	Essentially random depending on location of road crossings, sites with good foundations for a structure

are impounded by one or more pieces of large woody debris that is anchored in the channel. *Landslide dam pools* are impounded by the debris deposited by a mass movement into the channel. Tributary-mouth bars are downstream elongated sediment bodies originating at the confluence of one, usually smaller channel with another (24). Such bars can accumulate in the channel, impounding a pool upstream (25). *Backwater pools* occur along the edges of channels at local expansions or at junctions with secondary channels. Flow separation envelopes with reverse or upstream-directed currents are often present in the expansions. *Disconnected pools* are completely separated from the main channel at low flows by a bar or mound but are deep enough to be a window in the water table. They may be associated with secondary channels. *Abandoned channel pools* are ponded sections of a cutoff or avulsion that have no surface water connection to the main stream. *Beaver dam pools* are formed by impoundment by a woody barrier constructed by beavers (*Castor canadensis; Castor fiber*). *Anthropogenic dammed pools* are included for completeness to account for structure-induced impoundment.

Fast water habitats are divided into *nonturbulent* and *turbulent* channel units (Fig. 2). A *run* is a deeper, slower, less steep fast water bedform than a riffle (see below) that exhibits no supercritical flow. A *sheet* is a section of uniform water flow over smooth bedrock of variable gradient. An *anthropogenic nonturbulent unit* is one constructed by humans such as a causeway.

Turbulent fast water channel units (Fig. 2) are similar to the classes of Bisson and Montgomery (2) and Hawkins et al. (16) and are characterized by supercritical flow and hydraulic jumps sufficient to entrain air and create localized patches of white water. However, different definitions are proposed below for some of their channel units. *Falls* are high, essentially vertical drops, spanning the whole channel. The distinction between a fall and a bedrock, boulder, or log step (see below) is height. Falls are higher than the bankfull channel depth and excavate plunge pools at their bases. *Steps* are low (less than bankfull channel depth), essentially vertical drops over bedrock boulders or logs that may not span the channel. *Cascades* are steep sections composed of a series of bedrock, boulder, and/or log steps that span the channel in a staircase fashion separated by closely spaced step pools (21). More than 50% of the stream exhibits supercritical flow. *Rapids* are less steep than cascades and consequently exhibit irregular bedrock, boulder, and/or log steps that partially or fully span the channel (21). Between 15 and 50% of the stream displays supercritical flow. *Riffles* are less steep than rapids and do not display bedrock, boulder, or log steps (21). They are characterized by shallow subcritical flow, and only 5–10% of the water surface area exhibits supercritical flow (21). The chutes of Bisson and Montgomery (2) and Hawkins et al. (16) have been renamed *"slots"* because chute already has an established, alternative meaning in geomorphology, namely, a secondary flood channel across the inside of a bend or parallel to the main channel. An *anthropogenic turbulent unit* is one constructed by humans such as a dam spillway.

Microhabitats

Microhabitats are patches within a particular channel unit that have relatively homogeneous substrates, water depth, flow velocity, and cover (Table 1 and Fig. 1). No attempt is made to define classes here because of the large number of classes involved.

BIBLIOGRAPHY

1. Frissell, C.A., Liss, W.J., Warren, C.E., and Hurley, M.D. (1986). A hierarchical framework for stream habitat classification: viewing streams in a watershed context. *Environ. Manage.* **10**: 199–214.
2. Bisson, P.A. and Montgomery, D.R. (1996). Valley segments, stream reaches, and channel units. In: *Methods in Stream Ecology*. F.R. Hauer and G.A. Lamberti (Eds.). Academic Press, San Diego, CA, pp. 23–52.
3. Kellerhals, R., Church, M., and Bray, D.I. (1976). Classification and analysis of river processes. *J. Hydract. Div. Proc. Am. Soc. Civ. Eng.* **102**: 813–829.
4. Erskine, W.D., Webb, A.A., Turner, L.M., Miners, B., Rose, T., and Bevitt, R. (2001). Benchmarking of present river condition after 34 years of large scale interbasin water transfers: Snowy River, Australia. In: *Proceedings of the Third Australian Stream Management Conference, The Value of Healthy Streams*, Volume 1, Brisbane, Queensland, 27–29 August 2001. I. Rutherfurd, F. Sheldon, G. Brierley, and C. Kenyon (Eds.). Cooperative Research Center for Catchment Hydrology, Clayton, pp. 205–210.
5. Erskine, W.D. (1996). Downstream hydrogeomorphic impacts of Eildon reservoir on the Mid-Goulburn River, Victoria. *Proc. R. Soc. Victoria.* **108**: 1–15.
6. Galay, V.J., Kellerhals, R., and Bray, D.I. (1973). Diversity of river types in Canada. In: *Proceedings of the Symposium on Fluvial Processes and Sedimentation*. Hydrology Symposium Series No. 9, 217–250. Subcommittee on Geodesy and Geophysics, National Research Council of Canada, Inland Waters Directorate, Canada.
7. Mollard, J.R. (1973). Airphoto interpretation of fluvial features. In: *Proceedings of the Symposium on Fluvial Processes and Sedimentation*. Hydrology Symposium Series No. 9, 341–380. Subcommittee on Geodesy and Geophysics, National Research Council of Canada, Inland Waters Directorate, Canada.
8. Rosgen, D.L. (1994). A classification of natural rivers. *Catena* **22**: 169–199.
9. Rosgen, D.L. (1996). *Applied River Morphology*. Wildland Hydrology, Pagosa Springs, CO.
10. Thorne, C.R. (1998). *Stream Reconnaissance Handbook*. John Wiley, Chichester, UK.
11. Erskine, W.D., Terrazzolo, N., and Warner, R.F. (1999). River rehabilitation from the hydrogeomorphic impacts of a large hydro-electric power project: Snowy River, Australia. *Reg. Rivers: Res. Mgt.* **15**: 3–24.
12. Tooth, S. (1999). Floodouts in central Australia. In: *Varieties of Fluvial Form*. A.J. Miller and A. Gupta (Eds.). John Wiley, Chichester, UK, pp. 219–247.
13. Melville, M.D. and Erskine, W.D. (1986). Sediment remobilization and storage by discontinuous gullying in humid southeastern Australia. *Int. Assoc. Hydrol. Sci. Publ. No.* **159**: 277–286.
14. Schumm, S.A. and Khan, H.R. (1971). Experimental study of channel patterns. *Nature* **233**: 407–9.
15. Bird, E.C.F. (1962). The river deltas of the Gippsland Lakes. *Proc. R. Soc. Victoria* **75**: 65–74.
16. Hawkins, C.P. et al. (1993). A hierarchical approach to classifying stream habitat features. *Fisheries* **18**(6): 3–12.
17. Peterson, J.T. and Rabeni, C.F. (2001a). Evaluating the physical characteristics of channel units in an Ozark stream *Trans. Am. Fish. Soc.* **130**: 898–910.
18. Peterson, J.T. and Rabeni, C.F. (2001b). The relation of fish assemblages to channel units in an Ozark stream *Trans. Am. Fish. Soc.* **130**: 911–926.
19. Webb, A.A. and Erskine, W.D. (2001). Large woody debris, riparian vegetation and pool formation on sand-bed forest streams in southeastern Australia. In: *Proceedings of the Third Australian Stream Management Conference. The Value of Healthy Streams*, Volume 2, Brisbane, Queensland, 27–29 August 2001, I. Rutherfurd, F. Sheldon, G. Brierley, and C. Kenyon (Eds). 659–664. Cooperative Research Center for Catchment Hydrology, Clayton.
20. Turner, L.M. and Erskine, W.D. (1997). Morphometry and stratification of the Bents Basin scour pool, Nepean River, NSW *Wetlands (Aust.)* **17**(1): 14–28.

21. Grant, G.E., Swanson, F.J., and Wolman, M.G. (1990). Pattern and origin of stepped-bed morphology in high-gradient streams, Western Cascades, Oregon. *Geol. Soc. Am. Bull.* **102**: 340–352.
22. Keller, E.A. and Melhorn, W.N. (1978). Rhythmic spacing and origin of pools and riffles. *Geol. Soc. Am. Bull.* **89**: 723–730.
23. Schumm, S.A., Mosley, M.P., and Weaver, W.E. (1987). *Experimental Fluvial Geomorphology*. John Wiley & Sons, New York.
24. Petts, G.E. (1984). Sedimentation within a regulated river. *Earth Surf. Proc. Landf.* **9**: 125–134.
25. Benn, P.C. and Erskine, W.D. (1994). Complex channel response to flow regulation: Cudgegong River below Windamere Dam, Australia. *Appl. Geogr.* **14**: 153–168.

COASTAL WETLANDS

Susan-Marie Stedman
NMFS F/HC
Silver Spring, Maryland

GENERAL

Coastal wetlands include a variety of wetland types such as salt marshes, bottomland hardwood swamps, fresh marshes, seagrass beds, and mangrove swamps. Although many people think of tidal salt marshes when they hear "coastal wetlands," there are many wetlands in coastal areas that are neither tidal nor salty. Coastal wetlands can be defined as all wetlands in coastal watersheds, that is, watersheds that drain to the ocean or to an estuary and extend landward to the limit of tidal influence (Fig. 1). Great Lakes wetlands are often considered coastal wetlands as well, but the remainder of this discussion focuses on coastal wetlands associated with the Atlantic and Pacific Oceans and the Gulf of Mexico. Coastal wetlands can be next to streams, behind barrier beaches, on deltas, in isolated depressions, or in numerous other places in coastal landscapes. Coastal wetlands such as seagrass beds also occur in bays, estuaries, or other shallow coastal waters. Whether a coastal wetland is tidal or nontidal, fresh or saline, it influences and is influenced by the entire coastal watershed.

Like all wetlands, coastal wetlands are sometimes easy to recognize, but they also can be very difficult to distinguish from uplands when obvious signs of water are absent. Coastal wetlands can occur in areas with standing water, tidal flooding, or only periodic or seasonal flooding. Coastal wetlands, like other coastal environments, are very dynamic, responding to the effects of waves, winds, currents, sea level changes, subsidence, and human alterations of the landscape and hydrology. For example, barrier beaches along the Atlantic coast of the United

Figure 1. Coastal watersheds based on U.S. Geological Survey hydrologic units.

States tend to "roll" toward the land in response to relative sea level rise (1), causing the salt marshes directly behind these beaches to be buried by the beach as new areas of salt marsh become established behind the beach on areas where wind or water carries sand. Coastal marshes around brackish ponds can change from salt marshes to fresh marshes if the pond's connection to salt water is blocked by drifting sand or a road. Coastal fresh marshes along streams may expand or shrink depending on the changing location of the river channel.

Coastal wetlands make up about 30% of the wetlands in the continental United States, or approximately 27 million acres. The Gulf of Mexico region includes 51% of the coastal wetlands in the lower 48 states. The south Atlantic region contains another 30%, which means that approximately 81% of coastal wetlands in the continental United States are in the southeast. The entire Pacific coast (excluding Alaska and Hawaii) contains less than 2%. Most coastal wetlands (77%) occur in or adjacent to estuaries. The Ten Thousand Islands estuary in southern Florida, known for its mangroves and extensive areas of forested, shrub, and fresh marsh in the Everglades, contains the largest acreage of coastal wetlands in the United States, followed by the Mississippi Delta estuary, and the Albemarle/Pamlico Sound estuary (2).

Only approximately 18% of coastal wetlands are truly estuarine (having water with a salinity of at least 0.5 ppt) and only about 25% are tidal. The most common coastal wetland type is forested and scrub-shrub, which comprises 63% of the coastal wetlands in the continental United States. The next most common coastal wetland type is fresh marsh (16%), followed by salt marsh (15%) and tidal flat (4%) (2).

ECOLOGICAL IMPORTANCE

Coastal wetlands are among the most productive ecosystems in the world, comparable to tropical rain forests and coral reefs. An immense variety of microorganisms, plants, insects, amphibians, reptiles, birds, fish, and mammals can be part of a wetland ecosystem. Coastal wetlands can be thought of as "biological supermarkets." They provide large volumes of food that attract and support many animals. The combination of shallow water, high levels of nutrients, and high primary productivity is ideal for the development of organisms that form the foundations of the foodweb, supplying food for many species of fish, shellfish, insects, amphibians, birds, and mammals. A small percentage of wetland vegetation is eaten directly by fish and invertebrates, but most is decomposed by bacteria, producing an organic soup that feeds organisms such as amphipods, shrimp, crabs, snails, shellfish, and finfish. These organisms in turn support a broad food chain that includes a variety of fish, birds, and mammal populations such as deer, muskrat, and manatee.

Coastal wetlands are extremely important to birds, which use them for feeding, resting, and/or nesting. Major migratory flyways follow coastal wetlands along the Pacific and Atlantic Oceans. Many shore birds are wetland-dependent, including gulls, terns, plovers, and sandpipers. Waterfowl such as bay ducks (canvasback, redhead, etc.) and sea ducks (bufflehead, harlequin duck) are dependent on coastal wetlands, and some declines in coastal waterfowl populations are thought to be linked to the decline in coastal wetlands (3).

Coastal wetlands are also important nurseries for juvenile marine organisms. The shallow water and often-abundant plant growth are ideal for sheltering animals from predators until the animals are big enough or fast enough to avoid predators outside the wetland. Some fish, including species of herring and bass, use coastal wetlands as both a spawning area and nursery, whereas others, such as some species of shrimp and the ladyfish (a popular sportfish) spawn in the open ocean, leaving the juveniles to migrate via coastal currents to coastal wetland nurseries. Even adult fish often use coastal wetlands as refuge from larger predators or environmental stresses such as storms or floods.

Coastal marshes and estuaries provide essential habitat for over 75% of the fish caught commercially and recreationally in the United States (4). Popular species such as spotted seatrout, scup, butterfish, mullet, spot, salmon, gag grouper, Pacific herring, white grunt, summer flounder, menhaden, pink shrimp, spiny lobster, Atlantic croaker, and blue crab all use coastal wetlands as juveniles or adults for feeding and refuge. Bluefish and striped bass, important recreational species, depend on salt marshes for the small fish that are an important part of their diet. Menhaden, one of the most important commercial species landed in the United States, has a particularly strong tie to salt marshes and their detrital food chain. All of these species, and many others too numerous to list here, depend on healthy abundant coastal wetlands to support their populations.

Coastal wetlands are also important for keeping coastal waters clean. Coastal ecosystems receive virtually all the water flowing off the upland landscape—water that may contain fertilizers, pesticides, sewage, sediment, or toxic chemicals. Coastal wetlands filter out sediment and some chemicals, reducing the amount of pollution that washes into bays and the ocean.

TRENDS

Coastal wetlands are among the most imperiled ecosystems in the world. By the mid-1970s, over half of all salt marshes and mangrove forests present in the precolonial United States had been destroyed (5). California, a large coastal state, has lost over 90% of its wetlands. Florida and Louisiana, two coastal states with the greatest acreage of wetlands, have lost about half of their original wetland area. Louisiana alone is losing between 16,000 and 25,000 acres of wetlands a year, which is the highest sustained wetland loss rate in the country (6). Most of that loss is occurring in coastal areas.

No long-term national data set is available for coastal wetlands, although some programs, such as the U.S. Fish and Wildlife's National Wetlands Inventory (NWI), track estuarine wetlands and the U.S. Department of Agriculture's Natural Resources Inventory (NRI) has been used to estimate loss of tidal wetlands (7) and wetlands in coastal counties (8). Coastal wetland losses were estimated by Gosselink and Bauman (9) as approximately 20,000 acres per year from 1922 to 1954 and approximately 46,000 acres per year from 1954 to 1974. NRI data for the period between 1982 and 1987 indicate that tidal wetland loss on nonfederal lands was approximately 20,000 acres per year. However, tidal wetlands make up only about 25% of all coastal wetlands, so the total coastal wetland loss for that time period would have been greater. Most recently, Brady and Goebel (8) estimate wetlands losses on nonfederal land in coastal counties from 1992 to 1997 to have been about 24,000 acres per year. The net loss of wetlands in coastal counties was nearly three times the net loss of wetlands from inland counties. Nearly all of these losses were to freshwater nontidal coastal wetlands.

According to the NWI, the rate of wetland loss of all wetlands nationally has decreased from 458,000 acres per year from the mid-1950s to mid-1970s (10) to approximately 58,000 acres per year from 1986 to 1997 (11). While the national wetland loss rate has decreased to less than one-eighth of what it once was, the rate of coastal wetland loss appears to be decreasing much less (about one-half of what it was) and much more slowly.

Many factors are responsible for coastal wetland loss. In the early part of this century, coastal wetlands were drained and used for farming or grazing. More recently, coastal wetlands have been filled or dredged for roads, houses, golf courses, marinas, and other development. In some parts of the world wetlands are being destroyed for aquaculture. In other places, such as Louisiana, wetlands are being lost to open water due to a combination of factors including canals dredged through the marshes, dams on the Mississippi River reducing sediment to the marshes, land subsidence, and sea level rise. Even wetlands that are not actually filled or dredged are becoming degraded due to pollution, changes in water flows, and invasion by weeds or other non-native plants and animals.

Coastal wetland losses can be directly traced to population pressures and other human changes occurring along the coast. Coastal populations have increased steadily since 1970. Currently, over half the population of the United States lives in coastal counties, at densities about five times greater than those of noncoastal counties (12). This trend of people moving to coastal areas is expected to continue in the coming decades. Expanding populations place enormous pressures on existing natural resources, particularly wetlands, which are very vulnerable to changes in water flow, pollution, and habitat fragmentation. During the period from 1992 to 1997, nonfederal land in coastal counties contained 20% of the wetlands in the continental United States but experienced 31% of the wetland losses. Sixty-six percent of wetlands lost in coastal counties was the result of development while only 41% of wetlands lost in inland counties was due to development (7).

Coastal wetland loss from anthropogenic factors is magnified by global considerations such as sea level rise. Along much of the U.S. coast, relative sea level is rising at a rate of between a few inches to a foot or more per century (1). In fairly flat coastal plains, a few inches of rise in sea level can mean that the coastline moves hundreds of feet inland. Coastal wetlands can and do move inland with rising sea level, but in developed areas roads, houses, parking lots, and other human structures interfere with this natural migration of coastal habitats.

Fortunately, the importance of coastal wetlands and their vulnerability have come to the attention of scientists, governments, and nongovernment organizations. Research into the restoration of coastal wetlands and on-the-ground application of this research have yielded promising results, allowing for the restoration of some types of coastal wetlands. Through programs like those funded under the Coastal Wetlands Planning, Protection, and Restoration Act, federal and state agencies in the United States are working together to reduce the loss of wetlands in particularly fragile areas such as coastal Louisiana. And in many places, local governments are developing land use plans to protect coastal wetlands while nongovernmental organizations develop support for these programs and advocate for additional funding for restoration. With these efforts as a starting point and additional protection and restoration efforts in the future, coastal wetlands will continue to provide enormous benefits to the ecology and economy of the world.

BIBLIOGRAPHY

1. Kraft, J.C., Chrzastowski, M.J., Belknap, D.F., Toscano, M.A., and Fletcher, C.H, III (1988). The transgressive barrier-lagoon coast of Delaware: morphostratigraphy, sedimentary sequences and responses to relative rise in sea level. Delaware Sea Grant publication DEL-SG-04-88.
2. Field, D.W., Reyer, A.J., Genovese, P.V., and Shearer, B.D. (1991). *Coastal Wetlands of the United States—An Accounting of a Valuable Resource*. National Oceanic and Atmospheric Administration, Silver Spring, MD.
3. Stewart, R.E., Jr. (1996). Wetlands as bird habitat. In: *National Water Summary on Wetland Resources*. Fretwell et al. (compilers). U.S. Geological Survey Water Supply Paper 2425, pp. 49–56.
4. Chambers, J.R. (1992). Coastal degradation and fish population losses. In: *Stemming the Tide of Coastal Fish Habitat Loss*. Stroud (Ed.). National Coalition for Marine Conservation, Washington, DC, pp. 45–51.
5. Dahl, T.E. (1990). *Wetland Losses in the United States 1780s to 1980s*. U.S. Department of the Interior, Fish and Wildlife Service, Washington, DC.
6. Johnston, J.B., Watzin, M.C., Barras, J.A., and Handley, L.R. (1995). Gulf of Mexico coastal wetlands: case studies of loss trends. In: *Our Living Resources: a Report to the Nation on the Distribution, Abundance, and Health of U.S. Plants, Animals and Ecosystems*. La Roe et al. (Eds.). U.S. Department of the Interior, National Biological Service, Washington, DC, pp. 269–272.
7. Brady, S.J. and Flather, C.H. (1994). Changes in wetlands on nonfederal rural land of the coterminous United States from 1982 to 1987. *Environ. Manage.* **18**(5): 693–705.
8. Brady, S.J. and Goebel, J.J. (2002). Geographical distribution of wetland changes in the U.S., 1992–1997. Paper presented at the 5th Wetlands Workshop, Atlantic City, NJ. Oct. 21–24.

9. Gosselink, J.G. and Baumann, R.H. (1980). Wetland inventories: wetland loss along the United States coast. *Z. Geomorph. N.F.* 173–187.
10. Frayer, W.E., Monahan, T.J., Bowdell, D.C., and Graybill, F.A. (1983). *Status and Trends of Wetlands and Deepwater Habitats in the Coterminous United States, 1950's to 1970's*. National Wetlands Inventory, Washington, DC.
11. Dahl, T.E. (2000). *Status and Trends of Wetlands in the Conterminous United States 1986 to 1997*. U.S. Fish and Wildlife Service, Washington, DC.
12. National Research Council (NRC). (2000). *Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution*. National Academy Press, Washington, DC.

FRESHWATER COLLOIDS

JAMIE R. LEAD
University of Birmingham
Birmingham, United Kingdom

INTRODUCTION

Colloids are defined on the basis of size; at least one dimension is in the range 1 nm–1 μm ($10^{-9} - 10^{-6}$ m). Below 1 nm is the dissolved or solution phase, and above 1 μm is the particle phase. In practice, this distinction is usually made by operational methods such as ultrafiltration and filtration (1). Other techniques such as centrifugation have also been used as separation techniques. Although useful and widely used, this operational definition is inadequate for a number of reasons. First, there are practical considerations such as experimental artifacts due to filtration and mass-based, rather than size-based, separation by centrifugation. Second, the terminology is not compatible with thermodynamic conceptions according to Stumm and Morgan (2), who defined colloids as species for which a chemical potential could not be defined. Gustaffson and Gschwend (3) delimited colloids from both the dissolved phase and the particle phase, based on the chemical and hydrologic properties of the system.

Dissolved and colloid species were distinguished on the basis that the colloid forms a distinct chemical phase whose interactions with other components is different from solution phase interactions. The distinction between colloid and particle phases was based on the idea that colloids have a tendency to remain in the water column whereas particles tend to settle to the sediments.

As noted, this distinction is in part based only on size and also includes colloid chemistry and mass and hydrologic conditions. In this conception, a colloid may be so defined in one waterbody but may not be classed as a colloid in another water-body. This is similar to the distinction made elsewhere in the literature (4,5), where the behavior of colloids in surface waters is dominated by aggregation and the behavior of particles is dominated by aggregation. In porous media, colloids and particles may also be governed by retention and migration in the pores, respectively (4,6). This idea of settling particles and nonsettling colloids has important implications for their transport and fate in the environment.

PHYSICAL, CHEMICAL, AND MICROBIOLOGICAL PROPERTIES OF COLLOIDS

To understand the role and properties of colloids (aggregation, sedimentation, pollutant binding), it is essential to understand that their nature is essentially cross-disciplinary, requiring knowledge of their chemical, physical, and microbiological properties. A number of the physicochemical properties of colloids are important in this respect: they include size, shape, density, chemical composition, and surface charge. Activities of microorganisms are also important in producing and removing colloids from freshwaters and in binding trace pollutants. Colloids are also not 'pure' compounds but are intrinsically polydisperse (containing a range of sizes) and chemically heterogeneous. This inhomogeneity in size and chemistry reflects the wide variety of different sources and sinks for forming and eliminating of colloids. For example, colloids are produced by the weathering of rocks and soils and by the death, lysis, and degradation of biological cells. Additionally, colloids are destroyed by their aggregation into larger particles and sedimentation and by their use as a food source by other organisms.

Typically, freshwater colloids consist of biological, organic, and inorganic materials. These include microorganisms such as bacteria and their debris; humic substances, which are the degradation products of plants and other biota; microbial exudates, primarily polysaccharides and proteins; and weathering products such as clays and oxides. Colloids are not pure phases of these different materials, but the phases are intimately linked.

Sampling, Fractionation and Analysis

Although colloids are intimately linked mixtures of phases, they are very unstable and easily modified. Indeed, even the act of sampling may alter their composition (1,6). This instability is linked to aggregation, microbial growth and death, modification of organic matter, sorption to container walls, changes in solution conditions, and temperature. Sampling and analytical procedures must therefore be carefully chosen, the prime requirement is that they are as minimally perturbing as possible. Ideally, analysis would be performed *in situ*, although this can rarely be achieved in practice. The sampling step itself should be quick with a minimum of handling, storage, and transport. Full details of sampling procedures are given in the two previous references.

Fractionation, usually performed by size, is often required to reduce the complexity of colloids. A number of methods have been used. The main method is filtration, although centrifugation, field-flow fractionation (FFF), and others are now employed. Again, the main criterion for successful use is nonperturbation of colloids during fractionation, one that is rarely, if ever fully met.

A number of standard techniques exist to characterize the microbiological (e.g., flow cytometry) and bulk chemical composition of colloids (e.g., inductively coupled plasma mass spectrometry, ICP-MS), although these are outside the scope of this paper. Other techniques of great importance for the direct analysis of freshwater colloids include electron and force microscopy and various light scattering techniques. Each technique has its own advantages and limitations for example, transmission electron microscopy allows direct visualization of nanometer-scale material and the analysis of major elements on a single particle basis. However, sample preparation may be problematic, despite the advances made in the past 10 years. In addition, quantification of results is difficult. Due to the different strengths of each technique and the complexity of the colloids, coupling of two or more methods on-line is advantageous (e.g., FFF-ICP-MS) as is the use of a number of complementary techniques to analyze the same sample.

ENVIRONMENTAL IMPORTANCE

Freshwater colloids have been studied primarily because of their role in chemically binding trace pollutants (metals and organics) and thus altering the transport and bioavailability of pollutants. Additionally, they have been studied in relation to water treatment processes and as a food source for aquatic microorganisms. The physical and chemical nature of colloids means that they can bind a large, often dominant, fraction of trace pollutants (7). In particular, they have high surface area. Consequently, a large number of functional groups, which can bind trace metals, are presented to the solution. Additionally, the functional groups themselves strongly bind the metals. Although less well-developed experimentally and theoretically, the binding of organic pollutants by colloids has also been investigated. It is generally thought that the organic component of colloids controls binding.

The chemical binding of metals by colloids has two implications for metal biogeochemistry. First, uptake and effects of the metals by biota are modified, and second, environmental transport of the metals is affected.

Colloids, it is thought, act to reduce biotic uptake to aquatic organisms such as microorganisms and fish. The widely used Free Ion Activity Model (8) assumes that bioavailability is directly related to the free ('dissolved') metal ion: related models (9) have explained reduced uptake by the decreased diffusive mobility and chemical lability of the colloid-bound metal. Despite the successes of these and other models (such as the Biotic Ligand Model), much remains to be understood about the interrelationships between colloids and biota and subsequent effects on metal bioavailability and toxicity.

Colloids affect transport and cycling of trace metals through a number of mechanisms. As mentioned in the definition, one definition of colloids is based on their transport behavior in water. Colloids are small enough that mixing processes maintain them in the water column, whereas particles are large enough so that they sediment out of the water column. This simple mechanism provides an important method for the geochemical fractionation of associated pollutants in surface waters. In lakes, settling of particles is one of the main 'self-purification' mechanisms for removing metals from the water column (10). Additionally, in groundwater, the flow of colloids through porous media may govern the transport of associated pollutants.

BIBLIOGRAPHY

1. Lead, J.R., Davison, W., Hamilton-Taylor, J., and Buffle, J. (1997). Characterising colloidal material in natural waters. *Aquatic Geochemistry* **3**: 219–232.
2. Stumm, W. and Morgan., J.J. (1996). *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters*. John Wiley & Sons, New York.
3. Gustafsson, O. and Gschwend, P.M. (1997). *Limnology and Oceanography* **42**: 519–528.
4. Buffle, J. and Leppard, G.G. (1995). Characterisation of aquatic colloids and macromolecules. 1. Structure and behaviour of colloidal material. *Environmental Science and Technology* **29**: 2169–2175.
5. Filella, M. and Buffle, J. (1993). *Factors Controlling the Stability of Sub Micron Particles*.
6. Buffle, J. and Leppard, G.G. (1995b). Characterisation of aquatic colloids and macromolecules. 2. Key role of physical structures on analytical results. *Environmental Science and Technology* **29**: 2176–2184.
7. Lead, J.R., Davison, W., Hamilton-Taylor, J., and Harper, M. (1999). Trace metal sorption by natural particles and coarse colloids. *Geochimica et Cosmochimica Acta* **63**: 1661–1670.
8. Campbell, P.G.C. (1995). Interactions between trace metals and aquatic organisms: a critique of the free-ion activity model. In: *Metal Speciation and Bioavailability in Aquatic Systems*. John Wiley, Chichester, UK.
9. Tessier, A., Buffle, J., and Campbell, P.G.C. (1994). Uptake of trace metals by aquatic organisms. In: *Chemical and Biological Regulation of Aquatic Systems*. J. Buffle (Ed.).
10. Sigg, L. (1994). *Chemical and Biological Regulation of Aquatic Systems*. J. Buffle and R.R. de Vitre (Eds.). CRC Press, Boca Raton, FL.

CULVERT DESIGN

Xing Fang
Lamar University
Beaumont, Texas

A culvert (Fig. 1) is a conduit placed under a road to convey surface water from one side to another. In addition to its hydraulic function, it must also carry construction and highway traffic and Earth loads; therefore, culvert design involves both hydraulic and structural design. Culvert design involves selection of the type of culvert and size to pass the design discharge without overtopping of the road and without erosion on either end of the culvert. For economy and hydraulic efficiency, engineers should design culverts to operate with the inlet submerged during flood flows, if conditions permit. Culverts are constructed from a variety of materials and are available in many different shapes and configurations. When selecting a culvert, engineers should consider roadway profiles, channel characteristics, flood damage evaluations, construction and maintenance costs, and estimates of service life. The most common shapes for culverts are circular, pipe-arch and elliptical, box (rectangular), modified box, and arch. Common culvert materials include concrete (reinforced and nonreinforced), steel (smooth and corrugated), aluminum

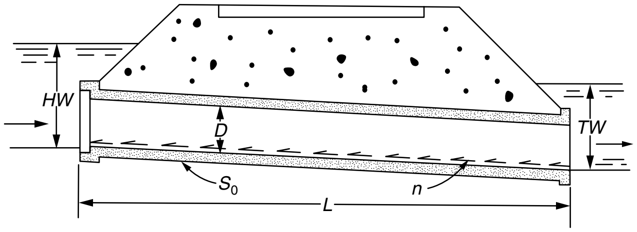

Figure 1. A typical culvert with submerged inlet and outlet.

(smooth and corrugated), and plastic (smooth and corrugated). The selection of material for a culvert depends on several factors that can vary considerably according to location. One should consider the following groups of variables: (1) structure strength, considering fill height, loading condition, and foundation condition; (2) hydraulic efficiency, considering Manning's roughness, cross-section area, and shape; (3) installation, local construction practices, availability of pipe embedment material, and joint tightness requirements; (4) durability, considering water and soil environment (pH and resistivity), corrosion (metallic coating selection), and abrasion; and (5) cost, considering availability of materials. Do not base culvert material selection solely on the initial cost. Replacement costs and traffic delay are usually the primary factors in selecting a material that has a long service life for a culvert design.

Energy is required to force flow through a culvert. This energy takes the form of an increased water surface elevation on the upstream side of the culvert. The depth of the upstream water surface measured from the invert at the culvert entrance is generally referred to as headwater depth (*HW* in Fig. 1). The flow through a culvert can take different forms and is a function of several variables, such as cross-sectional size (e.g., inside diameter *D* in Fig. 1) and shape, bottom slope (S_o), length (*L*), conduit roughness (usually in terms of Manning's *n*), entrance and exit design, and water depths at the upstream (headwater) and the downstream of the culvert (tailwater, *TW* from outlet invert in Fig. 1).

Both entrance flow and exit flow can be broadly classified as submerged (e.g., in Fig. 1) or free. Two types of flow controls are in culverts: inlet control and outlet control. For inlet control (Fig. 2), the control section is located at or near the culvert entrance, and the discharge through the culvert is dependent only on inlet geometry (size, shape, entrance type) and headwater depth (HW) because the conduit can convey more discharge than the inlet will allow. For the inlet control, if the outlet is not submerged, most flow in the conduit is open channel flow; if the outlet is submerged, part of the flow near the entrance can still be open channel flow, but flow near the outlet is pressurized pipe flow. For outlet control (Fig. 3), the control section is at or near the culvert outlet, and the discharge through the culvert is dependent on all hydraulic factors upstream from the outlet. For example, flow condition in Fig. 1 is an outlet control, and all hydraulic factors *HW, TW, L, D, n,* and S_o affect discharge through the culvert.

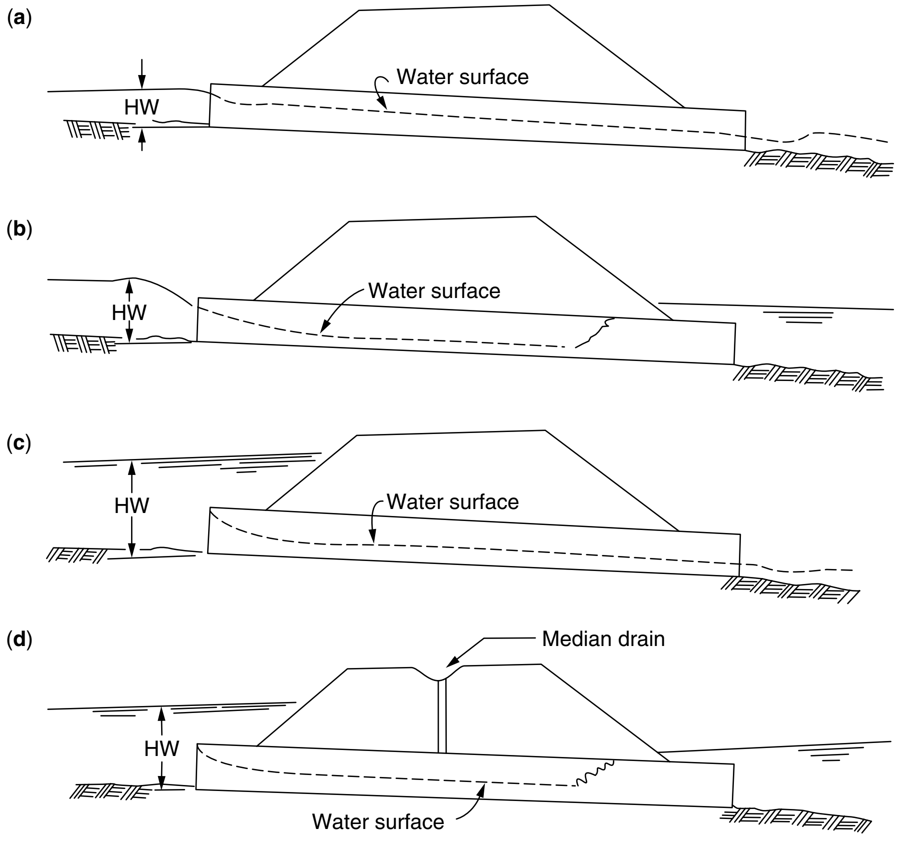

Figure 2. Types of inlet control for a culvert: (**a**) inlet and outlet unsubmerged; (**b**) inlet submerged, outlet unsubmerged; (**c**) inlet submerged only; and (**d**) inlet and outlet submerged [from Normann et al. (1)].

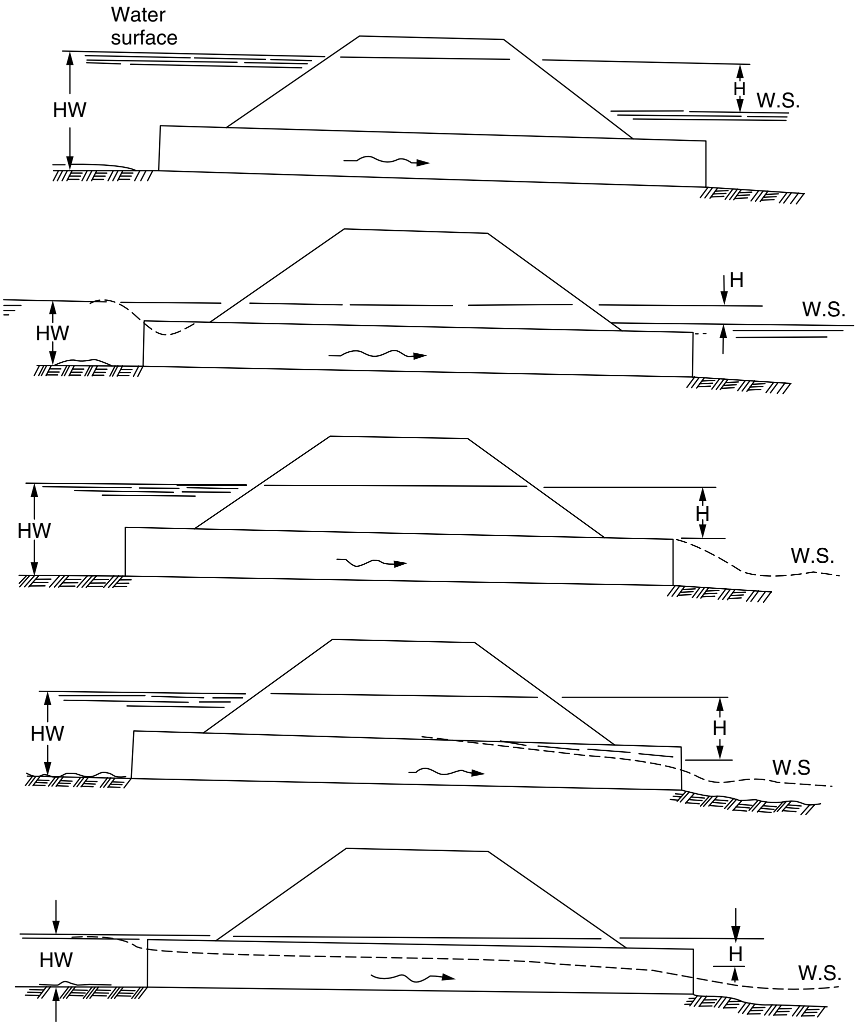

Figure 3. Types of outlet control for a culvert [from Normann et al. (1)].

The flow velocities from a culvert are likely to be higher than in the channel because a culvert usually constricts the available channel area. These increased velocities can cause streambed scour and bank erosion in the vicinity of the culvert outlet. Therefore, two basic culvert design criteria exist: allowable headwater (*HW*) and allowable velocity. Allowable headwater usually governs the overall configuration of the culvert. However, the allowable outlet velocity is the governing criterion in the selection and application of various downstream fixtures and appurtenances to slow the flow from the culvert. An engineer may improve culvert hydraulic capacity by selecting appropriate inlets (e.g., projecting, mitered, straight, flared, and parallel wingwalls). Many different inlet configurations including both prefabricated and constructed-in-place installations can be used by the engineer. When selecting various inlet configurations, one should also consider structural stability, aesthetics, erosion control, and fill retention.

The culvert design process typically includes the following basic steps:

1. Define the location, orientation, shape, and material for the culvert to be designed. One may consider more than a single shape and material.
2. With consideration of the site data, establish allowable outlet velocity and maximum allowable headwater depth.
3. Based on design discharges from hydrologic study of upstream watershed and associated tailwater levels, select an overall culvert configuration including culvert hydraulic length, entrance conditions, conduit shape, and material.
4. Perform hydraulic analysis to determine headwater depth and outlet velocity by using computer software or monographs.
5. If computed results do not satisfy the specified values at step 2, optimize the culvert configurations, and go back to step 3. Sometimes it may be necessary to treat any excessive outlet velocity separately (additional energy dissipation structures) from the headwater requirement.

Culvert design typically requires complex hydraulic calculations, especially to appropriately incorporate energy losses caused by entrance, friction, and exit. Equations, tables, and charts for culvert design can be found in many textbooks (2,3) or design manuals (1,4–6).

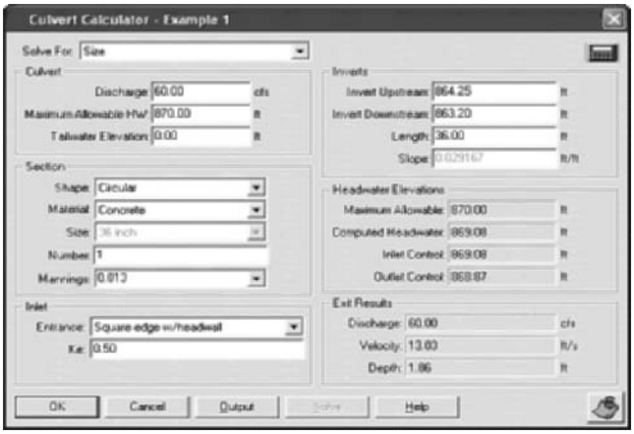

Solve For: Size

Culvert
Discharge: 60.00 cfs
Maximum Allowable HW: 870.00 ft
Tailwater Elevation: 0.00 ft

Section
Shape: Circular
Material: Concrete
Size: 36 inch
Number: 1
Mannings: 0.013

Inlet
Entrance: Square edge w/headwall
Ke: 0.50

Inverts
Invert Upstream: 864.25 ft
Invert Downstream: 863.20 ft
Length: 36.00 ft
Slope: 0.029167 ft/ft

Headwater Elevations
Maximum Allowable: 870.00 ft
Computed Headwater: 869.08 ft
Inlet Control: 869.08 ft
Outlet Control: 868.87 ft

Exit Results
Discharge: 60.00 cfs
Velocity: 13.03 ft/s
Depth: 1.86 ft

Figure 4. Culvert calculator in CulvertMaster (7).

Many computer models, e.g., HY8 by the Federal Highway Administration (FHWA) of the U.S. Department of Transportation, "CulvertMaster" developed by Haestad Methods, Inc. (7), and CULVERT2 (English units) and CULVERT3 (metric units) by the California Department of Transportation, are also available. Application of computer software is recommended because of complexity in culvert hydraulic computations. One must choose the shape and profile and establish the culvert dimensions by iterative application of the computation procedures until headwater and outlet velocity are reasonable. Sometimes it is possible to select culverts consisting of more than one box in wide channels to convey design discharge. The Culvert Calculator in CulvertMaster can solve for size, discharge, and headwater elevation for a culvert. The Calculator organizes information for a culvert into seven areas (Fig. 4): solve for, culvert, section, inlet, inverts, headwater elevation, and exit results. Figure 4 shows an example for sizing of a single, circular, concrete culvert (roughness n = 0.13) with square edge entrance and design discharge of 60 ft^3/sec and 36 ft long. The result for the culvert size is 36 inches; flow has inlet control because the headwater elevation for inlet control is greater than the one for outlet control.

BIBLIOGRAPHY

1. Normann, J.M., Houghtalen, R.J., and Johnson, W.J. (1985). *Hydraulic Design of Highway Culverts*. HDS No. 5, Federal Highway Administration (FHWA), U.S. Department of Transportation, Norfolk, VA.
2. Simon, L. and Korom, S.F. (1997). *Hydraulics*, 4th Edn. Prentice-Hall, Englewood Cliffs, NJ, pp. 142–143.
3. Roberson, J.A., Cassidy, J.J., and Chaudhry, M.H. (1997). *Hydraulic Engineering*, 2nd Edn. John Wiley & Sons, New York.
4. ACPA (American Concrete Pipe Association). (1987) *Concrete Pipe Design Manual*. ACPA, Vienna, VA.
5. AISI (American Iron and Steel Institute). (1995). *Modern Sewer Design*, 3rd Edn. AISI, Washington, DC.
6. Mays. (1999).
7. Haestad Methods, Inc. (2000). *Computer Applications in Hydraulic Engineering*. Halstad Press, Waterbury, CT.

DILUTION-MIXING ZONES AND DESIGN FLOWS

RAMAKAR JHA
K.K.S. BHATIA
C.S.P. OJHA
National Institute of Technology
Roorkee, Uttranchal, India

INTRODUCTION

The possibility that a pollutant will be discharged accidentally or intentionally into a stream/river is a matter of constant concern to those diverting and using water from streams and rivers. It is essential for water resources planners and managers to understand the physical processes involved in a stream/river prior to the implementation of any water resources projects. When a pollutant is discharged into a stream/river, it is subjected to initial dilution immediately due to the density difference between the receiving water and the pollutant.

The longitudinal, vertical and two-dimensional profiles of pollutant concentration indicate the density gradient between the pollutant and the receiving water. After initial dilution, the processes are governed by advection, reaction, and dispersion phenomena that tend to modify the initial pollutant concentration.

Advection phenomena represent the downstream transport of a discrete element of the waste load by the stream flow.

The reaction phenomena represent the decay of biodegradable materials in the waste under the action of naturally occurring bacteria in the stream. The dispersion phenomena represent that under the influence of turbulence, eddy currents, and similar mixing forces, a discrete element of the waste load tends not to remain intact, but mixes with adjacent upstream and downstream elements. In rivers and streams, the influence of dispersion phenomena is usually relatively small compared with advection and reaction phenomena; however, it can be important in some circumstances.

When a slug load results from a spill or accidental dump, dispersion effects can have an important influence on resulting peak concentrations, particularly at longer distances from the point of discharge (Fig. 1). Intermittent discharges, such as storm runoff, are also influenced by dispersion. However, for continuous discharges (e.g., from wastewater treatment plants) and steady-state flow conditions, dispersion effects are usually insignificant (Fig. 2).

As a result of these processes, the pollution level downstream from a wastewater disposal site varies in time and space. In the following sections, the importance of diluting pollutants downstream of their discharge based on various physical processes, the mixing length of pollutants (mixing zone), and design flows have been described.

IMPORTANCE OF DILUTION

The purpose of dilution is to reduce the concentration of pollutants in discharges below applicable standards. All

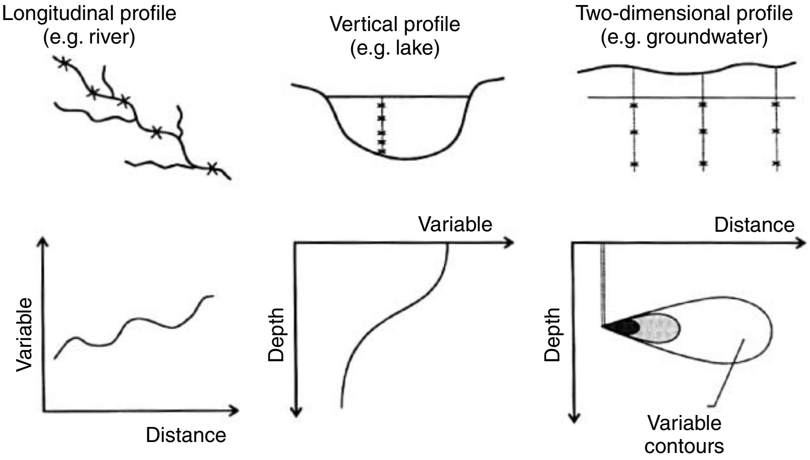

Figure 1. Spatial movement of water quality variables.

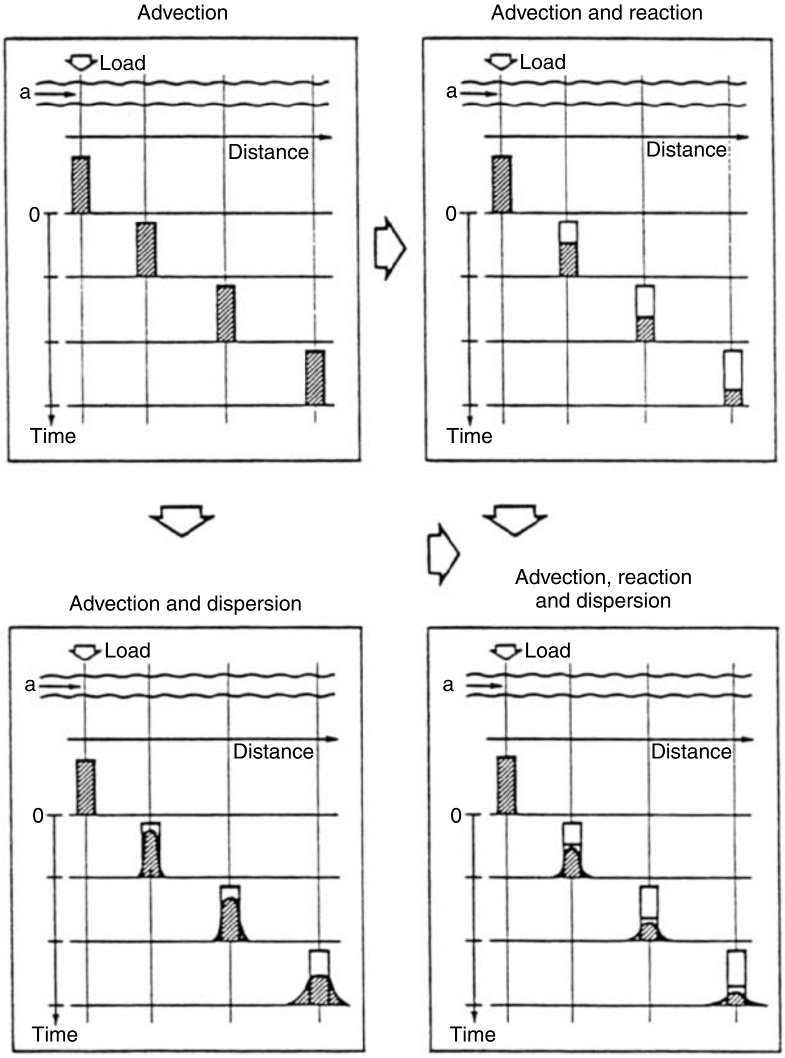

Figure 2. Transport mechanism for waste loads.

applicable water quality standards and nondegradation limits must be met at the end of a mixing zone. Through the first decades of the twentieth century, wastewater disposal practices were based on the premise that "the solution to pollution is dilution" (1). Dilution was considered the most economical means of wastewater disposal and as such was considered good engineering practice. It is one of the main processes for reducing the concentration of substances away from the discharge point. Dilution is more important for reducing the concentration of conservative substances (those that do not undergo rapid degradation, e.g., metals) than for nonconservative substances (those that do undergo rapid degradation, e.g., some organic substances).

Although dilution is a powerful adjunct to self-cleaning mechanisms of surface waters, its success depends upon discharging relatively small quantities of pollutants into large bodies of water. Growth in population and industrial activity and attendant increases in water demand and wastewater quantities preclude the use of many streams for diluting raw or poorly treated wastewaters or pollutants. In the United States, legal constraints further limit using waterbodies for wastewater dilution (1). Under present regulations, maximum allowable loads are set independently of dilution capacity. Only when the standard maximum loads result in violating in-stream water-quality standards, is dilution capacity considered, and then the increment of treatment necessary is determined.

To understand dilution phenomena due to increasing pollutant load in streams/rivers, different scenarios were obtained by plotting the changes in discharge versus simultaneous changes in the concentrations of various substances (2). As shown in Fig. 3, different curves may represent timescales that vary from a single storm to that of several years' duration. In the figure, curve (1) shows a general decrease in concentration with discharge, which implies increasing dilution of a substance introduced at a constant rate (e.g., major cations, particularly when concentrations are high). This is also characteristic of point source discharges such as municipal sewage and many industrial point sources. Curve (2) shows a limited increase in concentration generally linked to the flushing of soil constituents (e.g., organic matter, nitrogen species) during runoff. Curve (3) is basically the same as curve (2), but a fall in concentration occurs at very high discharges indicating dilution of soil-runoff waters. Curve (4) shows an exponential increase in concentration; this occurs for total suspended solids and all substances bound to particulate matter. The curve represents the increase in particulate matter due to sheet erosion and bed remobilization. Substances bound to such particulate include phosphorus, metals, and organic compounds, predominantly pesticides and herbicides. Curve (5) in the hysteresis loop is observed as time is introduced as an additional parameter to the sediment discharge relationship shown in curve (4). Such patterns can be seen for total suspended solids, dissolved organic carbon, and sometimes nitrate. The peak in sediment concentration occurs at X (advanced) before the occurrence of the peak discharge at Z. Curve (6) indicates a water source to the river of constant, or near constant concentration (e.g., chloride in rainfall, groundwater, and outlet from a lake).

From Fig. 3, it is found essential to determine the stream/rivers dilution capacity prior to the development of mathematical modeling and water quality simulations.

The dilution capacity of the receiving water can be defined as the effective volume of receiving water available for diluting the effluent. The effective volume can vary according to in-stream discharge and any other input source of discharge. In streams/rivers, in particular, the effective volume is much greater during the monsoon than during a lean period. It is important to consider concentrations of substances in the worst case scenarios (usually lean flow in summer, for example, when pollutants might be carried further into a sensitive location) when calculating appropriate discharge content conditions. The dilution capacity of a stream can be calculated by using the principles of mass balance. If the volumetric flow rate and the concentration of a given material are known in both the stream and waste discharge, the concentration after mixing can be calculated as follows:

$$C_sQ_s + C_wQ_w = C_mQ_m \tag{1}$$

where C represents the concentration (mass/volume) of the selected material, Q is the volumetric flow rate (volume/time), and the subscripts s, w, and m designate stream, waste, and mixture conditions.

MIXING ZONES

Wastewater effluents are discharged into streams and rivers to minimize any potential adverse effects on river water quality by suitable siting of outfalls. The stream zone between the outfall and the nearest cross section of uniform concentration distribution is known as a "mixing zone." The mixing of effluent with the river water takes place by the "jet effect" in the near-field region as shown in Fig. 4 (MixZon Inc., 2003).

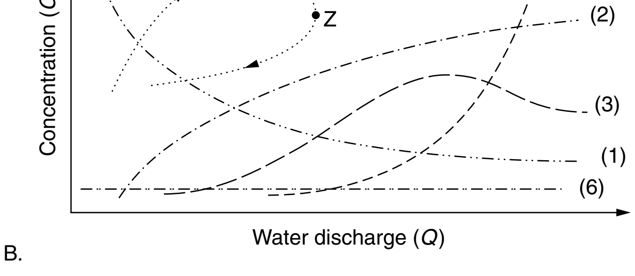

Figure 3. Patterns of concentration from discharges into streams/rivers.

Figure 4. A single port buoyant jet in cross flow deflected by the ambient current (near-field flow).

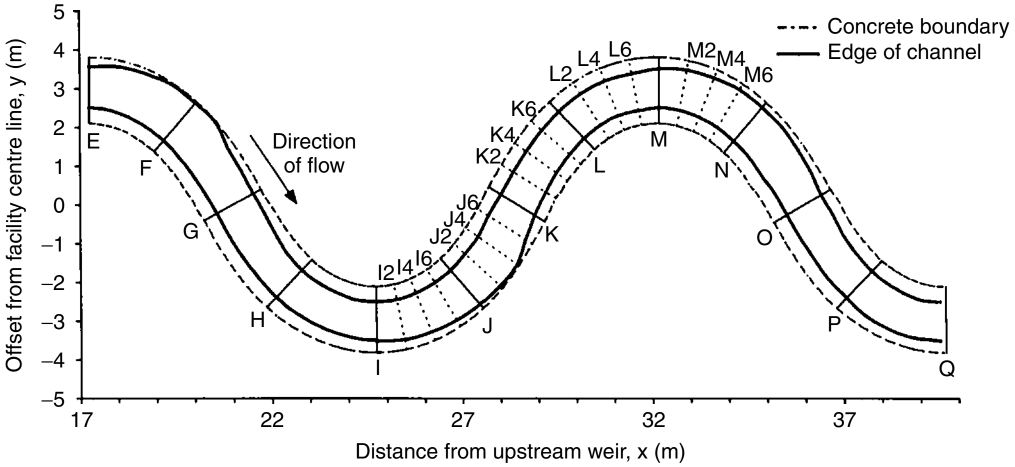

Figure 5. Typical channel geometry and section numbering.

The distance below the outfall where the effluent spreads across the channel width is termed "crossing distance." Within the mixing zone, apportion of the cross section termed "limited use zone" (LUZ) or "zone of noncompliance" occurs wherein the concentration of a pollutant may not comply with a specified water quality objective. The remaining portion of the cross section known as the "zone of passage" serves as a suitable habitat for fish and other desirable aquatic life. Note that in the far-field region, the mixing of effluent with the river water takes place due to lateral and longitudinal dispersion. A typical example of meandering given by Boxall et al. (3) is shown in Fig. 5. The cross-sectional profile and velocity pattern of flow at different sections of the channel are shown in Fig. 6. It can be seen from that figure that the maximum velocity profiles are always on the convex (right) portion of the channel looking downstream. When a pollutant is discharged at the bottom or on the surface of the different sections, the length of the mixing zone varies due to variation in velocity profiles. Also, the location of injection of pollutants (on the surface or at the bottom) changes the mixing phenomena, which can be well observed in Fig. 7.

It can be seen from the Fig. 7 that vertical mixing is usually complete in a short distance below the outfall or injection point, whereas lateral spreading is more gradual resulting in lateral concentration gradients. Ultimately, the cross-sectional concentration distribution attains uniformity at some distance below the outfall.

Mixing zones are prohibited in ephemeral waters or where there is no water for dilution (4). In a mixing zone, it is proposed to have high in-stream discharge for proper dilution and to meet standards at a boundary for the balance of the waterbody. It is the means for expediting mixing and dispersion of sewage, industrial waste, or other waste effluents in receiving waters.

Early workers (5–15) in the field devised mixing-zone concepts based on the lateral, vertical, and longitudinal dispersion characteristics of the receiving waters. Formulas predicting space and time requirements for diluting certain pollutants to preselected concentrations were also developed. The mixing behavior of a discharge depends largely on the depth of the ambient water, the momentum and buoyancy of the discharge, the spatial orientation of the discharge system, and the effects of many other factors. As a result, the mixing zone and mixing length are essential variables required for assessing the pollutant concentration/load downstream of the mixing zone.

Consideration of how to compute the distance from an outfall to compute mixing is a separate more complicated topic. However, the order of magnitude of the distance from a single point source to the zone of complete mixing is obtained from (16)

$$L_m = 2.6U\frac{B^2}{H} \tag{2}$$

for a side bank discharge, and from

$$L_m = 1.3U\frac{B^2}{H} \tag{3}$$

for a midstream discharge. In these equations,

L_m = distance from the source to the zone where the discharge has been well mixed laterally, in ft,

U = average stream velocity in fps,

B = average stream width in ft,

H = average stream depth in ft

DESIGN FLOW

The central problem in water quality management is assigning allowable discharges to a waterbody so that a designated water use and quality standard is met using basic cost–benefit principles. The following steps are required for proper design flow and waste load allocation in streams/rivers:

- evaluation of waste load input and in-stream discharge
- dilution phenomena and water quality response
- mixing length of pollutants in streams/rivers

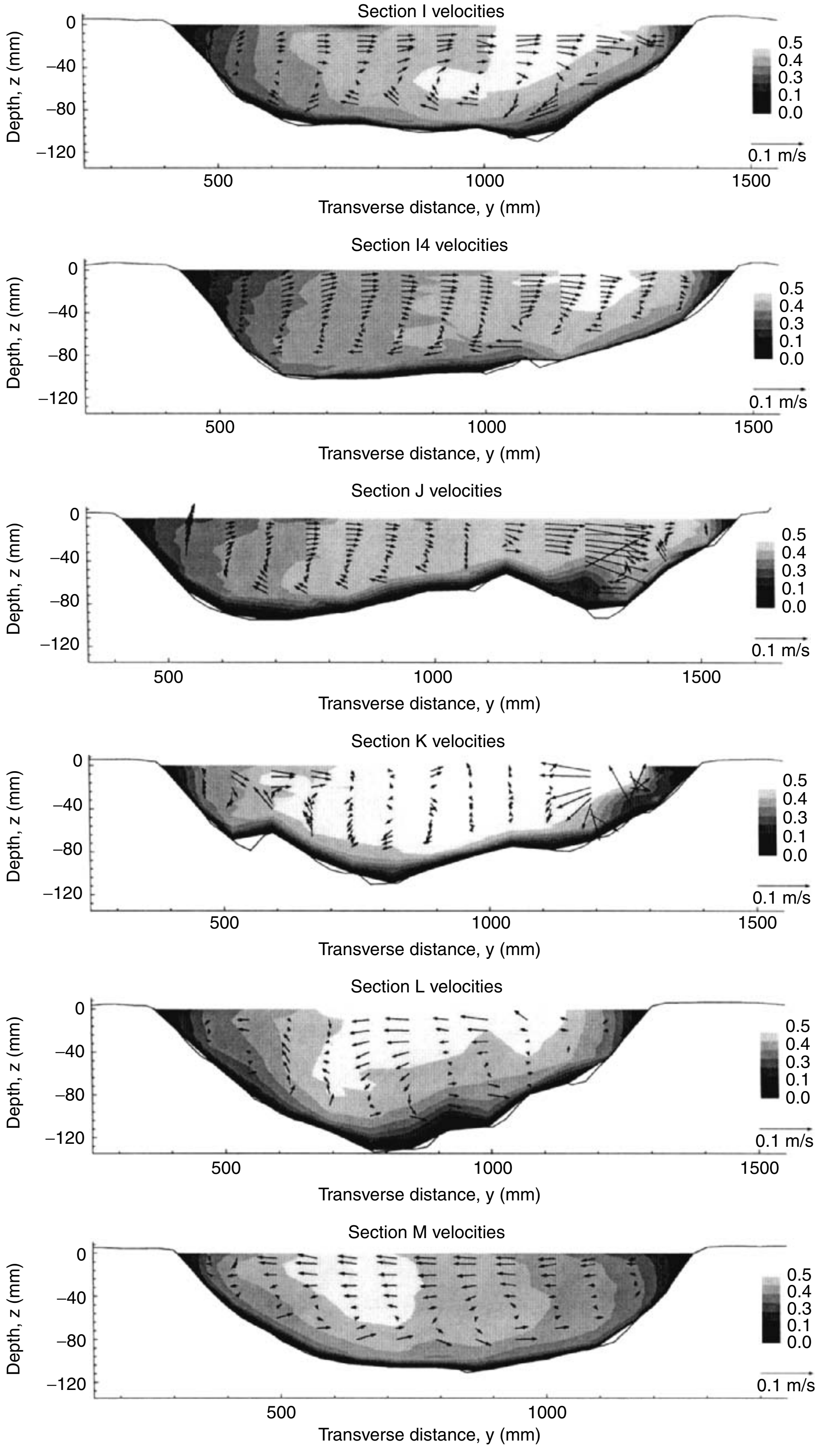

Figure 6. Cross-sectional and velocity profiles at different sections.

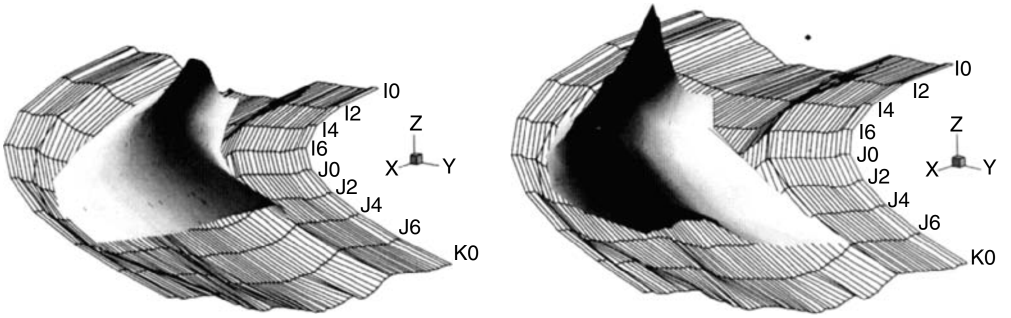

Bed injection at section I

Surface injection at section I

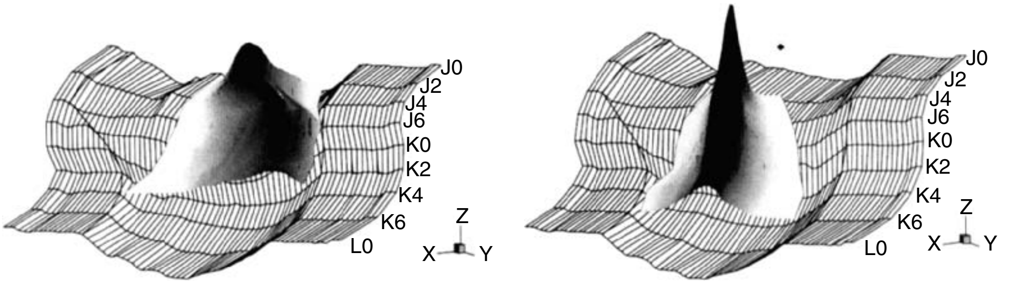

Bed injection at section J

Surface injection at section J

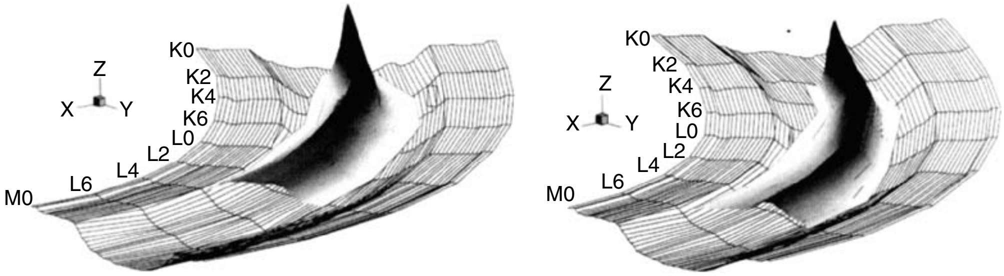

Bed injection at section K

Surface injection at section K

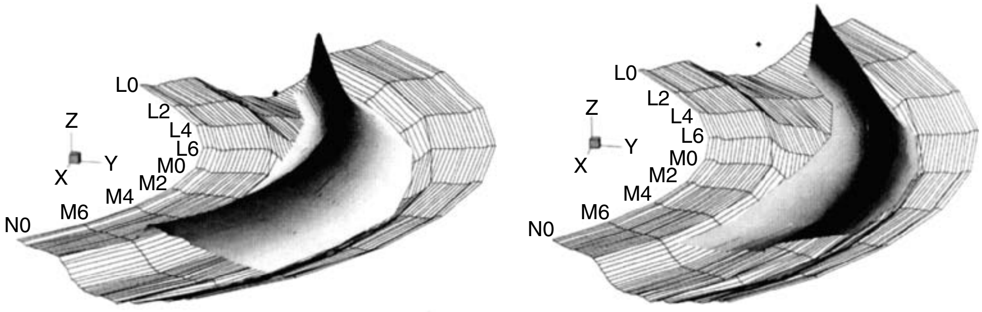

Bed injection at section L

Surface injection at section L

Figure 7. Three-dimensional representation of dye tracer measurements below different injection points.

- desirable water use and evaluation of water quality criteria that will permit such uses
- relationship between load and water quality and selection of projected conditions
- It is generally not sufficient to make a scientific engineering analysis of the effect of waste load inputs on water quality. The analytical framework should also include economic impacts which, in turn, must also recognize the sociopolitical constraints that are operative in the overall problem context.

To illustrate overall waste load allocation and a design flow problem for dissolved oxygen, a schematic diagram given by Thomman and Muller (17) is presented (Fig. 8).

It can be seen from Fig. 8 that for properly allocated waste loads in streams/rivers, DO values do not fall below the standard limits (as shown in the lower part of the figure). In the figure, there are several points at which

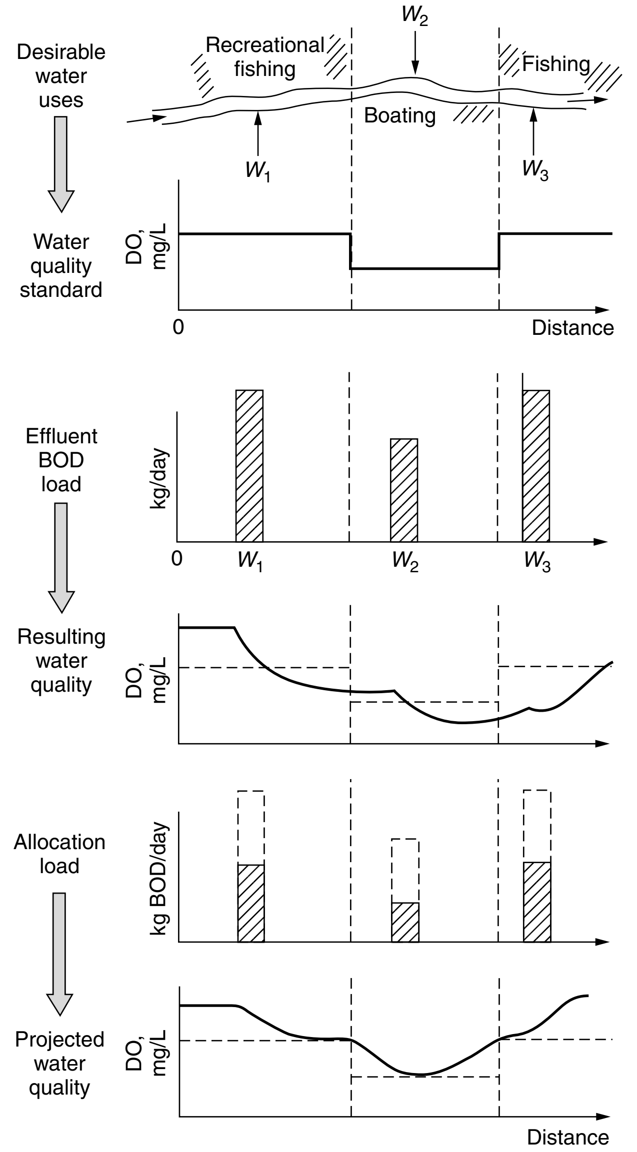

Figure 8. Representation of waste load allocation problem for dissolved oxygen.

careful judgments are required to provide a permissible waste load allocation. Determination of design conditions, including flow and parameters, must be evaluated for a waste load allocation. The specification or projection of flow and parameter conditions under a given design is a most critical step and is a combination of engineering judgment and sensitivity. The concentration at the outfall can be controlled by

(1) reducing the effluent concentration of the waste input by (a) wastewater treatment, (b) industrial in-plant process control, and (c) eliminating effluent constituents by pretreatment prior to discharge to the municipal sewer or by different product manufacturing for an industry;

(2) reducing the upstream concentration by upstream point and nonpoint source controls;

(3) reducing the effluent volume by (a) reduction in infiltration into the municipal sewer system, (b) reduction of direct industrial discharge volumes into the sewer system, and (c) reduction, for industry, of waste volume through process modifications;

(4) increasing the upstream flow by low flow augmentation, that is, release from upstream reservoir storage or from diversions from nearby bodies of water;

(5) increasing the environmental, in-stream degradation rate of the substance.

TYPICAL EXAMPLES OF DILUTION IN STREAMS/RIVERS OF INDIA

In India, all 14 major rivers including Ganga, Yamuna, Gomti, Cauvery, and Damodar are polluted. The Rivers Damodar and Yamuna are the most heavily polluted. Many parts of these rivers do not have any dissolved oxygen, and no wonder that they fail to support the growth of desirable aquatic fauna and flora. The River Ganges is heavily polluted at Allahabad, Kanpur, Varanasi, Patna, and Bhagalpur. The cross section and BOD/DO contours at Kanpur and Varanasi are shown in Figs. 9 and 10, respectively (18).

As can be seen from Figs. 9 and 10, the BOD concentrations found are very high at the pollutant outfall site. There is a significant amount of variation in BOD across the river. A similar situation has been observed in the longitudinal direction also due to dilution and mixing. The variation in dissolved oxygen (DO) downstream of the sewage outfall on the River Ganges, at Varanasi is shown in Fig. 11. As can be seen, the DO pattern at Assi Ghat in Varanasi, India, was monitored at different locations in longitudinal and transverse directions (19,20), and the river shows a nearly perfect dilution ratio. Similar studies were also, carried out for River Kali, a highly polluted river in India (18,21).

SUMMARY

Dilution, mixing zone, and design flows are interlinked issues that have been addressed here. Now, a better understanding of the mixing zone is available in the literature, and there are softwares like CORMIX (MixZon Inc., 2003) that can very elegantly perform mixing zone analysis. In the literature, near-field and far-field analysis of mixing of pollutants has also received attention. However, there are still challenges ahead, and these are related to the presence of a dead zone, adverse flow gradients, and curvilinear flows.

BIBLIOGRAPHY

1. Peavy, H.S., Rowe, D.R., and Techobanoglous, G. (1985). *Environmental Engineering*. McGraw-Hill, New York.
2. Chapman, D. (1992). *Water Quality Assessment*. E & FN Spon, An imprint of Chapman and Hall, New York.
3. Boxall, J.B., Guymer, I., and Marion, A. (2003). Transverse mixing in sinuous natural open channel flows. *J. Hydrau. Res.* **41**(2): 153–165.

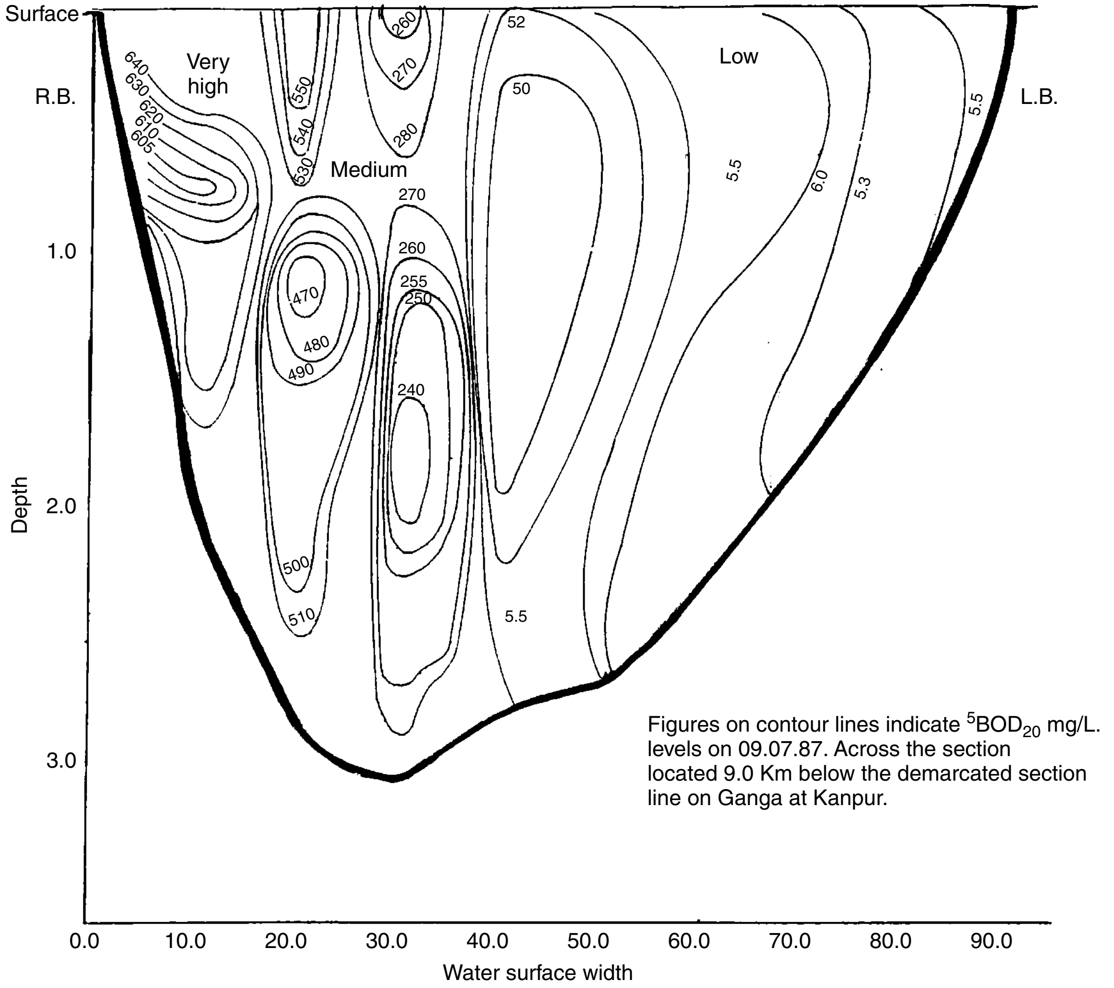

Figures on contour lines indicate $^5BOD_{20}$ mg/L. levels on 09.07.87. Across the section located 9.0 Km below the demarcated section line on Ganga at Kanpur.

Figure 9. Cross sectional profile of BOD levels in the Ganga at Kanpur.

4. Varga, C. (2003). *Environmental Programs*. Arizona Department of Environmental Quality.
5. Clark, J.W., Viessman, W., Jr., and Hammer, M.J. (1977). *Water Supply and Pollution Control*, 3rd Edn. Harper & Row, New York.
6. Velz, C.J. (1970). *Applied Stream Sanitation*. Wiley Interscience, New York.
7. Bathurst, J.C., Thorne, C.R., and Hey, R.D. (1979). Secondary flow and shear stress at river bends. *J. Hydrol. Div.* ASCE **HY10**: 1277–1295.
8. Chang, Y. (1971). *Lateral Mixing in Meandering Channels*. Ph.D. Thesis, submitted May 1971, University of Iowa, Ames, IA.
9. Chow, V.T. (1959). *Open Channel Hydraulics*. McGraw-Hill, New York.
10. Fisher, H.B. (1967). The mechanics of dispersion in natural streams. *J. Hydrol. Div.* ACSE **HY6**: 187–215.
11. Graf, W.H. and Kironoto, B.A. (1994). Turbulence characteristics in rough uniform open channel flow. *Proc. Inst. Civ. Eng., Water, Maritime & Energy* **106**: 333–344.
12. Guymer, I. and O'Brien, R.T. (2000). Longitudinal dispersion due to surcharged manholes. *J. Hydrol. Eng.* ASCE **126**(2): 137–149.
13. Krishnappen, B.G. and Lau, Y.L. (1977). *Transverse Dispersion in Meandering Channels*. Canada, Inland Waters Branch, Scientific Series, No. 75.
14. Lau, Y.L. and Krishnappen, B.G. (1981). Modelling transverse mixing in natural streams. *J. Hydrol. Eng.* ASCE **HY2**: 209–226.
15. Rutherford, J.C. (1994). *River Mixing*. John Wiley & Sons, New York.
16. Yotsukara. (1968). *As Referenced in Preliminary Report, Techniques of Water Resources Investigations of the U. S. Geological Survey, Measurement of Time of Travel and Dispersion by Dye Tracing*. Book 3, Chapter A9 by F.A. Kilpatrick, L.A. Martens, and J.F. Wilson, 1970.
17. Thomann, R.V. and Mueller, J.A. (1987). *Principles of Surface Water Quality Modelling and Control*. Harper & Row, New York.
18. Jha, R., Ojha, C.S.P., and Bhatia, K.K.S. (2003). A supplementary approach for estimating reaeration coefficients. *Hydrological Processes!* (accepted).
19. Choudhary, U.K. (1987). *The Ganga at Varanasi, DST Project Report*. Institute of Technology, Varanasi, India.
20. Choudhary, U.K., Jha, R., and Kumar, S. (1997). *Bed Profiles of Sediment Laden Flow in a Curved Open Channel, Int. Symp. Emerging Trends Hydrol*. Department of Hydrology, Roorkee, India.

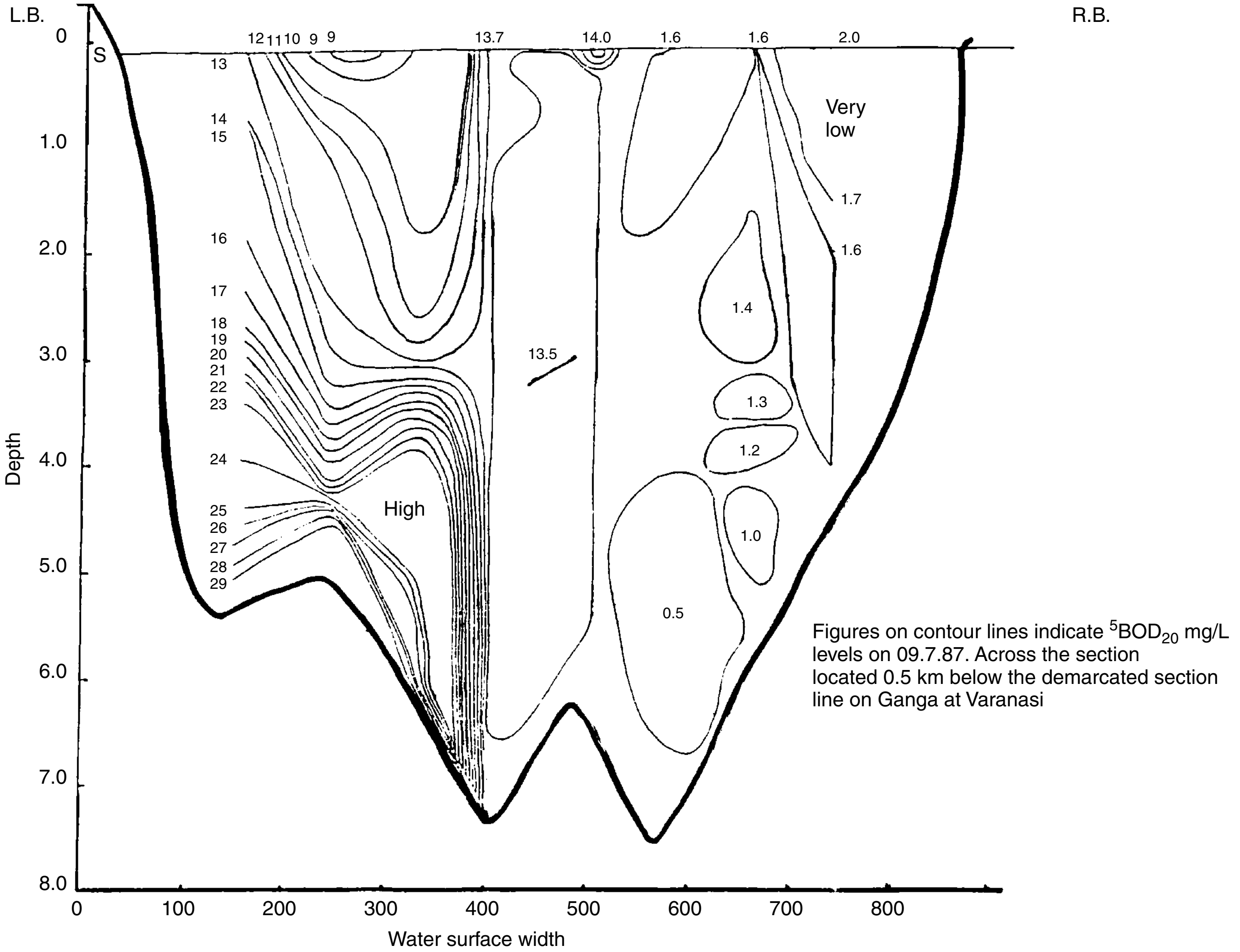

Figure 10. Cross-sectional profile of BOD levels in the Ganga at Kanpur.

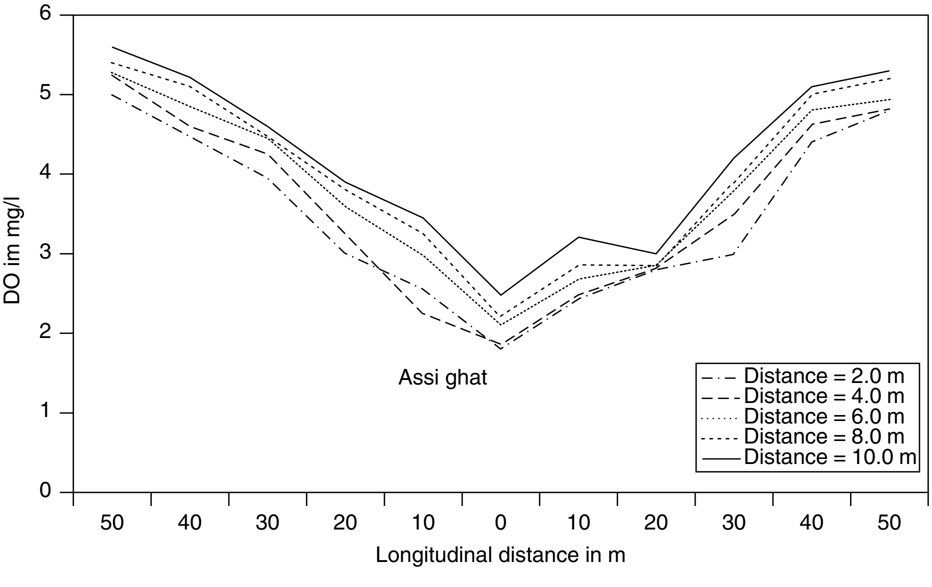

Figure 11. DO variations in the River Ganges at Assighat, Varanasi, India.

21. Jha, R., Ojha, C.S.P., and Bhatia, K.K.S. (2001). Refinement of predictive reaeration equations for a typical India river. *Hydrological Processes*. **15**(6): 1047–1060.

22. MixZon Inc. (2003). *CORMIX-GI v.4.2 and D-CORMIX v1.2 software*. USEPA/OHSU, Portland, OR.

READING LIST

Pawlow, J.R. (1983). *Technical Guidelines Manual for Performing Waste Load Allocations*. U.S. Environmental Protection Agency, Washington, DC.

Jobson, H.E. (1996). *Prediction of Travel Time and Longitudinal Dispersion in Rivers and Streams*. USGS Water-Resources Investigations Report 96–4013, Reston, VA.

DRAINAGE DITCHES

Matthew T. Moore
Charles M. Cooper
USDA Agricultural Research
Service National Sedimentation
Laboratory
Oxford, Mississippi

Jerry L. Farris
Arkansas State University
State University, Arkansas

OVERVIEW

Drainage ditches are common landscape features found in urban, industrial, and agricultural settings. Just as diverse as the environment in which they can be found, ditches have a wide array of physical and structural parameters. They may be lined by concrete, rock (rip rap), bare soil, or vegetation. Some ditches are small enough to step across, while others may require different methods of moving from one side to the opposite. Ditches, by their very nature, are water conveyance features. Natural ditches, formed as gullies, have been prominent landscape features in areas of topographical relief since the Earth was formed. Investigations of historical gully formation find ditches associated with primitive slash-and-burn agriculture. Drainage ditches have historically been efficient methods for transporting water away from moist or wet areas. The most obvious illustration is found in agricultural fields, where production acreage is surrounded by a myriad of drainage ditches (Figs. 1–3). Moving stormwater or excess irrigation water away from crops is the primary focus of agricultural drainage ditches.

In industrial areas, ditches are often the first postmanufacturing "treatment" area for discharges before the effluent reaches a receiving stream. The discharge pipe often empties into a vegetated drainage ditch, which will transport water to a river, lake, or stream. Urban settings are laced with interspersed ditches of varying sizes, again with the primary objective of removing stormwater from residential areas. Many urban ditches are lined with concrete, which further encourages water movement. Unfortunately, this also negates many positive water quality enhancement opportunities that drainage ditches possess.

Figure 1. Constructed ditch draining farm field.

Figure 2. Overgrown constructed ditch draining farm field.

Figure 3. Natural drainage canal in farm field.

Ditches may also serve to deliver water to areas for irrigation purposes. In western states where water supplies are closely managed, ditches provide an important conduit for water transportation. The California Gold Rush resulted in construction of over 5000 miles of canals to transport water to gold fields. Today, California supplies over half of the fruits, vegetables, and nuts for the United States (plus exports) on 3% of the nation's cropland because of its 9.5 million irrigated acres, due in part to the extensive network of canals and ditches. Thus, the topic of drainage ditches covers a broad spectrum of conveyance systems in time and space located from agricultural to urban settings. For our purposes, discussion is focused on ditches located in agricultural settings.

Engineering techniques associated with ditches have evolved so that an excellent level of confidence can be achieved in ditch design. Studies of open channel flow began in natural channels, but major knowledge breakthroughs have come through investigations of constructed channels. The use of laboratory flumes in the twentieth century greatly accelerated the knowledge base necessary for the current area of challenge—modeling. Energy, friction, channel roughness, Froude numbers, and critical depth of flow are some of the parameters where historical research provides the basis for aquatic habitat design.

Although drainage works have been in existence for over 2500 years, they are not without controversy (1). During the Middle Ages, ditches exsiccated English fens to increase agricultural production landscape. While beneficial for farmers, fen draining was seen as a direct threat to both fishermen and fowlers (1). In the early to mid-twentieth century, the U.S. government strongly favored drainage of wetlands and swamps to boost agricultural acreage, facilitate transportation, and combat mosquito and other vector-borne illnesses (2). As a result, fewer than half of the original U.S. wetlands remain (1). Without drainage, however, the fertile Midwest states of Ohio, Indiana, Illinois, and Iowa would still be largely swamps or land too wet for viable agriculture production (1). The last congressionally mandated decennial census of U.S. drainage (1985) indicated 110 million acres of agricultural land profited from drainage, and the net capital value of farm drainage work was estimated at $25 billion (1).

The U.S. Department of Agriculture's Natural Resource Conservation Service (USDA NRCS) promotes the use of conservation practices to improve soil and water quality. Under the broad topic of conservation are two standards for excess water removal, Practice 607 (Surface Drainage, Field Ditch) and Practice 608 (Surface Drainage, Main or Lateral) (3). Neither specifically addresses water quality enhancement.

Data in Tables 1–5 summarize available global drainage ditch research from the 1960s to 2004. As is common in progression of research, initial studies searched for basic information (e.g., surveys or nutrient loading). Later research focused more on current themes.

CURRENT AND FUTURE DIRECTIONS

Currently, experimental agricultural drainage ditch research is being primarily conducted with three institutions in the United States: Arkansas State University

Table 1. Ecological Research Conducted on Agricultural Drainage Ditches

Research	Location	Results	Reference
Floristic survey	Eastern Saudi Arabia	More than 100 plant species in canals	4
Flora	The Netherlands	78% of *Spirogyra* (algae) in Netherlands found in ditches and pools; some ditches had >20 species	5
Flora	United States	Maintenance of weeds on ditch banks	6–8
Flora	Mississippi Delta, United States	Ditches without surrounding riparian cover have increased nutrient concentrations	9
Macroinvertebrate	Rhône Delta, France	Description of copepod, *Cletocamptus retrogressus*	10
Macroinvertebrate	Germany	Discovery of aquatic beetle, *Dytiscus semisulcatus*	11
Macroinvertebrate	United Kingdom	Ephemoptera and Plecoptera present in nutrient-enriched ditches	12
Macroinvertebrate	Wicken Fen, United Kingdom	Rotation of ditch maintenance allowed Ephemeroptera establishment and survival	13
Macroinvertebrate	United Kingdom	Importance of vegetated refuges to survival of uncommon gastropods	14,15
Macroinvertebrate	Yakima Valley, Washington, United States	Successful overwintering of beneficial and harmful insects in ditches	16
Fish	Germany	Suggested management practices to offset fish population damages	17
Macrofauna	The Netherlands	Random distribution of macrofauna taxa in ditch sampling analyses	18
Fauna	Canada	90% of fish species utilizing drainage ditch habitat spawned in ditch	19
Fauna	United Kingdom	Ditch age and bank profile contributed to differences in individual ditch fauna	20
Flora/fauna	Elbe Estuary, Germany	Floral/faunal shifts as a result of pesticide applications and eutrophication in ditches	21
Plant and invertebrate interaction	Germany	*Anopheles* (mosquito) occurrence related to ditch plant succession	22
Floodplain diversity	Lower Frome floodplain, United Kingdom	Slow moving Rushton Ditch contributed largely to floodplain biodiversity	23

Table 2. General Research Conducted on Agricultural Drainage Ditches

Research	Location	Results	Reference
Design criteria	Malaysia	Integrate engineering and agricultural practices to maximize land productivity	24
Drain cleaning	Australia	Necessity of cleared drains in sugarcane production to maintain ditch carrying capacity	25
Maintenance	United Kingdom	Recommend dredging cycle of 3 years to maintain diverse floating and aquatic plants	26
Mechanical harvesting	The Netherlands	Determination of plant cutting times for maximum species richness on sand and peat soils	27
Upland ditches	United Kingdom	Little ecological advantage gained with ditches (grips) in uplands	28
Model	The Netherlands	Modeling determined critical nutrient loading dependence on ditch depth	29,30
Watershed hydrology	France	Differences seen between watersheds with and without ditches	31
Hydrology	United Kingdom	Physical-based hill slope model to understand changing catchment runoff regimes	32
Hydrology	United Kingdom	Drain interaction with channel hydrology (DITCH) model indicated higher ditch water levels did not improve ecology	33
Hydrology	United States	Correspondence between peak flows of ditches and rivers	34
Best management practice (BMP)	United States	Suggested use of ditches as BMP for nutrient and pesticide mitigation	35

Table 3. Nutrient Research Conducted on Agricultural Drainage Ditches

Research	Location	Results	Reference
Phosphorus	Australia	Decrease in phosphorus after traveling through bare drains; increase in phosphorus after passing through grass pasture	36
Phosphorus	New Zealand	Drainage sediments act as phosphorus sinks	37
Phosphorus	Delaware, United States	Ditch sediment biologically reactive phosphorus similar to that of topsoils of adjacent fields	38,39
Nitrate	River Eider Valley, Germany	Increase in retention time resulted in decrease in nitrate due to vegetation and organic debris accumulation	40
Nutrients	The Netherlands	No increase in phytoplankton in nutrient-enriched ditches	41
Nutrients	The Netherlands	90–95% phosphorus removal; low nitrogen removal	42
Nutrients	St. Lucie County, Florida, United States	Nutrient concentrations in citrus grove ditch water higher than concentrations in vegetable farm ditch water	43
Nutrients	United States	Ditches around baitfish farms are of benefit	44
Nutrients	Arkansas and Mississippi Deltas, United States	Ditches help remediate aquaculture effluent	45
Nutrient loading	United States	Seasonal variation of nutrients by distinctive ditch size	46,47
Nutrient standards	The Netherlands	Failure of general standards to support minimum water; need to be specific	48
Nutrients + pesticides	The Netherlands	Increase in macrophyte biomass and stored nitrogen and phosphorus	49
Herbicide + nutrient loading	South Saskatchewan River, Canada	Loadings diluted by storm; no significant increase in concentration	50

(ASU), Ohio State University (OSU), and the University of Minnesota (UM). ASU's research site includes a series of newly constructed vegetated drainage ditches, catchment ponds, small meanders, and riparian areas. Innovative research, in conjunction with the USDA Agricultural Research Service's National Sedimentation Laboratory, will highlight contaminant fate and transport, partitioning, mitigation, and ecological benefits of ditches alongside agronomic research. OSU continues to conduct research on the benefits of one-stage and two-stage ditches for longevity, water quality, and ecological enhancements (73). The University of Minnesota Southwest Research and Outreach Center recently constructed two 200-m open drainage ditches to study sediment and nutrient removal from drainage water (74).

Our literature review indicated ditch research is centered in the United States, The Netherlands, the United Kingdom, and Germany. Research thrusts range across topics of this decade, including greater understanding of basic hydrological and trapping processes, ditch contributions to pollution, contaminant processing by ditches, and maintenance. Future investigations must view drainage ditches holistically as water conveyance segments of a watershed. While basic studies on habitat, maintenance, and function are needed, research must define the ditch's role as a key facet of edge-of-field conservation

Table 4. Pesticide Research Conducted on Agricultural Drainage Ditches

Research	Location	Results	Reference
Herbicide retention	France	70% decrease in diflufenican with ditch surface contact	51
Herbicide transport	United Kingdom	99% of sulfosulfuron load found in first 4% of flow	52
Herbicide transport	United Kingdom	Mecoprop transport following storm event exceeded U.K. concentration guidelines	53
Herbicide (Linuron)	The Netherlands	Amendment decreased pH and dissolved oxygen; no effect on plant composition; diatoms and cryptophytes decreased; cladocera and copepods increased	54–56
Insecticide effects	The Netherlands	Vegetated systems (*Elodea*) sorbed the majority of applied chlorpyrifos; sediment was sink where no vegetation present; effects on macroinvertebrates occurred four times faster in open versus vegetated water	57–59
Lambda-cyhalothrin (insecticide)	The Netherlands	Sediment–pesticide sorption lower in areas with vegetation; alkaline hydrolysis main pesticide transformation process	60
Esfenvalerate	Mississippi Delta, United States	Toxicity measured in ditch; concentrations from 50-acre runoff mitigated to no observed effects levels in 510 m of vegetated ditch	61,62
Atrazine + lambda-cyhalothrin	Mississippi Delta, United States	Reduce pesticides from 5-acre runoff to no observed effects level in 50 m of vegetated ditch	63
Bifenthrin + lambda-cyhalothrin	Mississippi Delta, United States	Reduce pesticides from 50-acre runoff to no observed effects level in 650 m of vegetated ditch	64,65
Pesticides	California, United States	78% of ditch samples toxic—affected estuary	66
Spray drift buffer	The Netherlands	Buffer zone (3-m width) around ditch banks lowered spray drift by 95% in ditches	67
Insecticide + nutrients	The Netherlands	Increase in phytoplankton abundance; grazers effected by pesticide amendment	41
Insecticide + nutrients	The Netherlands	Increase in macrophyte biomass and stored nitrogen and phosphorus	49
Herbicide + nutrient loading	South Saskatchewan River, Canada	Loadings diluted by storm; no significant increase in concentration	50
Monitoring	Lower Fraser Valley, Canada	Consistent presence of diazinon and dimethoate in ditch water	68
Monitoring	Canada	4% of sampled ditches were toxic to *Ceriodaphnia dubia*; 14% impaired *C. dubia* reproduction; trace amounts of organophosphate insecticides	69

Table 5. Sediment Research Conducted on Agricultural Drainage Ditches

Research	Location	Results	Reference
Sediment	United States	Danger of Mississippi and Illinois River agricultural field ditches eroding	70
Sediment	United States	Cost is $0.45 per cropland acre to remove Ohio sediment from ditches; losses could be reduced by 25% with best management practices	71
Sediment	United States	Decreased sediment removal cost by implementing best management practices in Illinois, Indiana, Ohio, and Idaho	72

and pollution reduction systems. Given negative aspects of direct drainage, modifications that ameliorate ditches should blend habitat improvement with utilitarian design.

BIBLIOGRAPHY

1. Pavelis, G.A. (1987). *Farm Drainage in the United States: History, Status, and Prospects*. Miscellaneous Publ. No. 1455, Economic Research Service, U.S. Department of Agriculture, Washington, DC.
2. Grumbles, B.H. (1991). Section 404(f) of the Clean Water Act: trench warfare over maintenance of agricultural drainage ditches. *William Mitchell Law Rev.* **17**: 1021–1053.
3. USDA and NRCS. (2003). *National Handbook of Conservation Practices*. http://www.nrcs.usda.gov/technical/Standards/nhcp.html.
4. Shaltout, K. and El-Halawany, E.S. (1993). Vegetation analysis of the irrigation and drainage canals in eastern Saudi Arabia. *J. Univ. Kuwait* **20**(2): 261–274.
5. Simons, J. and van Beem, A.P. (1990). *Spirogyra* species and accompanying algae from pools and ditches in The Netherlands. *Aquat. Bot.* **37**: 247–269.
6. Millhollon, R.W. (1969). Control of johnsongrass on ditchbanks in sugarcane. *Weed Sci.* **17**(3): 370–373.
7. Millhollon, R.W. (1978). Controlling *Equisetum* in field drainage ditches of southern Louisiana. *Proc. Southern Weed Sci. Soc.* **31**: 230.

8. Wallis, R.L. and Turner, J.E. (1969). Burning weeds in drainage ditches to suppress populations of green peach aphids and incidence of beet western yellows disease in sugarbeets. *J. Econ. Entomol.* **62**: 307–309.
9. Bouldin, J.L., Farris, J.L., Moore, M.T., and Cooper, C.M. (2004). Vegetative and structural characteristics of agricultural drainages in the Mississippi Delta landscapes. *Environ. Pollut.* **132**: 403–411.
10. Mielke, W. (2001). *Cletocamptus retrogressus* (Copepoda, Harpacticoida) from irrigation and drainage ditches of the Rhône delta (Camargue, France). A redescription. *Vie Milieu* **51**(1–2): 1–9.
11. Braasch, D. and Bellstedt, R. (1991). *Dytiscus semisulcatus* Mueller, 1776—an inhabitant of drainage ditches. *Koleopterol. Rundsch.* **61**: 21–24.
12. Drake, C.M. (1991). Ephemeroptera and Plecoptera in freshwater and brackish ditch systems on British grazing marshes. *Entomol. Gazette* **42**(1): 45–59.
13. Painter, D. (1998). Effects of ditch management patterns on Odonata at Wicken Fen, Cambridgeshire, UK. *Biol. Conser.* **84**: 189–195.
14. Watson, A.M. and Ormerod, S.J. (2004). The microdistribution of three uncommon freshwater gastropods in the drainage ditches of British grazing marshes. *Aquat. Conser. Mar. Freshwater Ecosys.* **14**: 221–236.
15. Watson, A.M. and Ormerod, S.J. (2004). The distribution of three uncommon freshwater gastropods in the drainage ditches of British grazing marshes. *Biol. Conser.* **118**: 455–466.
16. Wallis, R.L. and Turner, J.E. (1972). Insects overwintering in the warm microenvironment of drainage ditches in central Washington. *Environ. Entomol.* **1**(1): 107–109.
17. Hinrichs, D. (1998). The influence of weed control measures on the fish fauna of drainage channels. *Wasser Boden* **50**(5): 22–25.
18. Rietveld, W. and Beltman, B. (1982). A qualitative analysis of microfauna sampling in ditches. *Hydrobiol. Bull.* **16**(2–3): 149–157.
19. Leslie, J.K. and Timmins, C.A. (1990). The community of young fish in drainage ditches in southwestern Ontario. *Arch. Hydrobiol.* **118**(2): 227–240.
20. Painter, D. (1999). Macroinvertebrate distributions and the conservation value of aquatic Coleoptera, Mollusca and Odonata in the ditches of traditionally managed and grazing fen at Wicken Fen, UK. *J. Appl. Ecol.* **36**: 33–48.
21. Caspers, H. and Heckman, C.W. (1981). Ecology of orchard drainage ditches along the freshwater section of the Elbe Estuary. Biotic succession and influence of changing agricultural methods. *Arch. Hydrobiol. Suppl.* **43**(4): 347–486.
22. Kuehlhorn, F. (1979). Succession of vegetation in neglected drainage ditches and the occurrence of larvae of *Anopheles*. *Arch. Hydrobiol.* **87**(2): 243–251.
23. Armitage, P.D., Szoszkiewicz, K., Blackburn, J.H., and Nesbitt, I. (2003). Ditch communities: a major contributor to floodplain biodiversity. *Aquat. Conser. Mar. Freshwater Ecosys.* **13**: 165–185.
24. Park, K.S. (1975). Design criteria and layout of farm irrigation and drainage ditches in paddy cultivation. *Bull. Minist. Agric. Rural Dev. Malays.* **139**: 140–151.
25. Linedale, A.I. (1970). Clean drains make better drains. *Queensland Bur. Sugar Exp. Stn. Cane Growers Q. Bull.* **34**(1): 13–14.
26. Milsom, T.P., Sherwood, A.J., Rose, S.C., Town, S.J., and Runham, S.R. (2004). Dynamics and management of plant communities in ditches bordering arable fenland in eastern England. *Agric. Ecosys. Environ.* **103**: 85–99.
27. Best, E.P.H. (1994). The impact of mechanical harvesting regimes on the aquatic and shore vegetation in water courses of agricultural areas of The Netherlands. *Vegetation* **112**(1): 57–71.
28. Coulson, J.C., Butterfield, J.E.L., and Henderson, E. (1990). The effect of open drainage ditches on the plant and invertebrate communities of moorland and on the decomposition of peat. *J. Appl. Ecol.* **27**(2): 549–561.
29. Janse, J.H. (1998). Model of ditch vegetation in relation to eutrophication. *Water Sci. Technol.* **37**(3): 139–149.
30. Janse, J.H. and Van Puijenbroek, P.J.T.M. (1998). Effects of eutrophication in drainage ditches. *Environ. Pollut.* **102**: 547–552.
31. Carluer, N. and De Marsily, G. (2004). Assessment and modeling of the influence of man-made networks on the hydrology of a small watershed: implications for fast flow components, water quality and landscape management. *J. Hydrol.* **285**: 76–95.
32. Dunn, S.M. and Mackay, R. (1996). Modeling the hydrological impacts of open ditch drainage. *J. Hydrol.* **179**: 37–66.
33. Gavin, H. (2003). Impact of ditch management on the water levels of a wet grassland in southeast England. *Agric. Ecosys. Environ.* **99**: 51–60.
34. Pierce, C. and Thompson, B. (1980). A survey of water flow in drainage ditches and streams in south central Minnesota. *J. Minn. Acad. Sci.* **46**(3): 10–12.
35. Moore, M.T., Cooper, C.M., Smith, S. Jr., Bennett, E.R., and Farris, J.L. (2001). Drainage ditches: new conceptual BMPs for non-point source pollution and TMDL development. In: *Proceedings of the Seventh Federal Interagency Sedimentation Conference*, Reno, Nevada, Vol. II, pp. 65–71.
36. Barlow, K., Nash, D., Turral, H., and Grayson, R. (2003). Phosphorus uptake and release in surface drains. *Agric. Water Manage.* **63**: 109–123.
37. Nguyen, L. and Sukias, J. (2002). Phosphorus fractions and retention in drainage ditch sediments receiving surface runoff and subsurface drainage from agricultural catchments in the North Island, New Zealand. *Agric. Ecosys. Environ.* **92**: 49–69.
38. Sallade, Y.E. and Sims, J.T. (1997). Phosphorus transformations in the sediments of Delaware's agricultural drainage ways. I. Phosphorus forms and sorption. *J. Environ. Qual.* **26**(6): 1571–1579.
39. Sallade, Y.E. and Sims, J.T. (1997). Phosphorus transformations in the sediments of Delaware's agricultural drainage ways. II. Effect of reducing conditions on phosphorus release. *J. Environ. Qual.* **26**(6): 1579–1588.
40. Scholz, M. and Trepel, M. (2004). Water quality characteristics of vegetated groundwater-fed ditches in a riparian peatland. *Sci. Total Environ.* **332**: 109–122.
41. Van Donk, E., Prins, H., Voogd, H.M., Crum, S.J.H., and Brock, T.C.M. (1995). Effects of nutrient loading and insecticide application on the ecology of *Elodea*-dominated freshwater microcosms: I. Responses of plankton and zooplanktivorous insects. *Arch. Hydrobiol.* **133**(4): 417–439.
42. Meuleman, A.F.M. and Beltman, B. (1993). The use of vegetated ditches for water quality improvement. *Hydrobiologia* **253**(1–3): 375.
43. Zhang, M., He, Z., Calvert, D.V., and Stoffella, P.J. (2004). Spatial and temporal variations of water quality in drainage ditches within vegetable farms and citrus groves. *Agric. Water Manage.* **65**: 39–57.

44. Frimpong, E.A., Lochmann, S.E., and Stone, N.M. (2003). Application of a methodology for surveying and comparing the prevalence of drainage ditches to baitfish farms. *North Am. J. Aquaculture* **65**(2): 165–170.
45. Stephens, W.W. and Farris, J.L. (2004). Instream community assessment of aquaculture effluents. *Aquaculture* **231**: 149–162.
46. Quade, H.W. (1978). *County Drainage Ditches in South Central Minnesota: A Unique Riparian Ecosystem*. General Technical Report, WO-USDA Forest Service, Washington, DC, pp. 400–410.
47. Quade, H.W. (1981). County drainage ditches in south central Minnesota—their nature and effects. In: *Selected Proceedings of the Midwest Conference on Wetland Values and Management*. B. Richardson (Ed.). pp. 165–179.
48. Peeters, E.T.H.M. and Gardeniers, J.J.P. (1998). Ecologically based standards for nutrients in streams and ditches in The Netherlands. *Water Sci. Technol.* **37**(3): 227–234.
49. Brock, T.C.M., Roijackers, R.M.M., Rollon, R., Bransen, F., and Van der Heyden, L. (1995). Effects of nutrient loading and insecticide application on the ecology of *Elodea*-dominated freshwater microcosms: II. Responses of macrophytes, periphyton and macroinvertebrate grazers. *Arch. Hydrobiol.* **134**(1): 53–74.
50. Cessna, A.J., Elliott, J.A., Tollefson, L., and Nicholaichuk, W. (2001). Herbicide and nutrient transport from an irrigation district into the South Saskatchewan River. *J. Environ. Qual.* **30**: 1796–1807.
51. Margoum, C., Gouy, V., Laillet, B., and Dramais, G. (2003). Retention of pesticides by farm ditches. *Rev. Sci. l'Eau* **16**(4): 389–405.
52. Brown, C.D., Bubus, I.G., Fogg, P., Spirlet, M., and Gustin, C. (2004). Exposure to sulfosulfuron in agricultural drainage ditches: field monitoring and scenario-based modeling. *Pest Manage. Sci.* **60**: 765–776.
53. Fletcher, C.A., Scrimshaw, M.D., and Lester, J.N. (2004). Transport of mecoprop from agricultural soils to an adjacent salt marsh. *Mar. Pollut. Bull.* **48**: 313–320.
54. Cuppen, J.G.M., Van den Brink, P.J., Van der Woude, H., Zwaardemaker, N., and Brock, T.C.M. (1997). Sensitivity of macrophyte-dominated freshwater microcosms to chronic levels of the herbicide linuron. *Ecotoxicol. Environ. Safety* **38**: 25–35.
55. Kersting, K. and Van Wijngaarden, R.P.A. (1999). Effects of a pulsed treatment with the herbicide Afalon (active ingredient linuron) on macrophyte dominated mesocosms: I. Responses of ecosystem metabolism. *Environ. Toxicol. Chem.* **18**(12): 2859–2865.
56. Van Geest, G.J., Zwaardemaker, N.G., Van Wijngaarden, R.P.A., and Cuppen, J.G.M. (1999). Effects of a pulsed treatment with the herbicide Afalon (active ingredient linuron) on macrophyte dominated mesocosms: II. Structureal responses. *Environ. Toxicol. Chem.* **18**(12): 2866–2874.
57. Brock, T.C.M. et al. (1992). Fate and effects of the insecticide Dursban 4E in indoor *Elodea*-dominated and macrophyte free freshwater model ecosystems: I. Fate and primary effects of the active ingredient chlorpyrifos. *Arch. Environ. Contam. Toxicol.* **23**(1): 69–84.
58. Brock, T.C.M. et al. (1992). Fate and effects of the insecticide Dursban 4E in indoor *Elodea*-dominated and macrophyte free freshwater model ecosystems: II. Secondary effects on community structure. *Arch. Environ. Contam. Toxicol.* **23**: 391–409.
59. Brock, T.C.M. et al. (1993). Fate and effects of the insecticide Dursban 4E in indoor *Elodea*-dominated and macrophyte free freshwater model ecosystems: III. Aspects of ecosystem functioning. *Arch. Environ. Contam. Toxicol.* **25**: 160–169.
60. Liestra, M. et al. (2003). Fate of the insecticide lambda-cyhalothrin in ditch enclosures differing in vegetation density. *Pest Manage. Sci.* **60**: 75–84.
61. Bouldin, J.L. et al. (2004). Evaluating toxicity of Asana XL (esfenvalerate) amendments in agricultural ditch mesocosms. *Chemosphere* **56**: 677–683.
62. Cooper, C.M. et al. (2004). Innovative uses of vegetated drainage ditches for reducing agricultural runoff. *Water Sci. Technol.* **49**(3): 117–123.
63. Moore, M.T. et al. (2001). Transport and fate of atrazine and lambda-cyhalothrin in an agricultural drainage ditch in the Mississippi Delta, USA. *Agric. Ecosys. Environ.* **87**: 309–314.
64. Cooper, C.M., Moore, M.T., Bennett, E.R., Smith, S., Jr., and Farris, J.L. (2002). Alternative environmental benefits of agricultural drainage ditches. *Verh. Int. Verein. Limnol.* **28**: 1678–1682.
65. Moore, M.T., Cooper, C.M., Bennett, E.R., Smith, S., Jr., Shields, F.D., Jr., and Farris, J.L. (2004). The role of vegetated drainage ditch research in the Mississippi Delta: current results and future directions. In: *Water Quality Assessments in the Mississippi Delta: Regional Solutions, National Scope*. M.T. Nett, M.A. Locke, and D.A. Pennington (Eds.). Oxford University Press, New York, pp. 194–203.
66. Hunt, J.W. et al. (1999). Patterns of aquatic toxicity in an agriculturally dominated coastal watershed in California. *Agric. Ecosys. Environ.* **75**: 75–91.
67. de Snoo, G.R. and de Wit, P.J. (1998). Buffer zones for reducing pesticide drift to ditches and risks to aquatic organisms. *Ecotoxicol. Environ. Safety* **41**: 112–118.
68. Wan, M.T., Szeto, S., and Price, P. (1994). Organophosphorus insecticide residues in farm ditches of the Lower Fraser Valley of British Columbia. *J. Environ. Sci. Health B Pesticides, Food Contam. Agric. Wastes* **B29**(5): 917–949.
69. McLeay, M.J. and Hall, K.J. (1999). Monitoring agricultural drainage ditches and the receiving water (Nicomekl River, Surrey, B.C.) for toxicity to *Ceriodaphnia dubia* and probable cause due to organophosphate contamination. *Water Qual. Res. J. Can.* **34**(3): 423–453.
70. Walker, R.D. (1982). Downstream problems from sediment: adverse effects, water quality, drainage ditches, lakes, reservoirs, and streams, Illinois. *Ill. Res. Ill. Agric. Exp. Stn.* **24**(2): 8–10.
71. Forster, D.L. and Abrahim, G. (1985). Sediment deposits in drainage ditches: a cropland externality. *J. Soil Water Conser.* **40**(1): 141–143.
72. Fletcher, J.J. (1985). Off-site erosion damages to drainage ditches and water conveyance structures. In: *Off-site Costs of Soil Erosion: Proceedings of a Symposium*. T.E. Waddell (Ed.). SCSA, Washington, DC, pp. 248–255.
73. Ward, A.D., Moore, M.T., Bouchard, V.L., Powell, K., Mecklenberg, D., Cooper, C.M., and Smith, S., Jr. (2004). Water quality benefits of grassed fluvial features in drainage ditches. In: *Proceedings of the 4th International Conference on Hydroscience and Engineering*. Brisbane, Australia. CDROM.
74. Strock, J.S. and Sands, G.R. (2001). Open ditches for water quality and water quantity. In: *Proceedings of the 1st Agricultural Drainage and Water Quality Field Day*. University of Minnesota Southwest Research and Outreach Center, Lamberton, MN, pp. 35–42.

DRAINAGE NETWORKS

Xiangjiang Huang
Jeffrey D. Niemann
Colorado State University
Fort Collins, Colorado

A drainage network is the configuration of the stream courses in a river basin. Drainage networks are the products of fluvial erosion processes. As erosion incises at a point on the land surface, that point becomes lower, which allows it to collect more flow. As the flow increases, the erosion rate also increases. This positive feedback ultimately produces a dendritic network of channels within a river basin. The drainage network is responsible for the surface transport of water, sediment, and contaminants out of the basin.

EXTRACTION OF DRAINAGE NETWORKS FROM DIGITAL DATA

Channel networks are commonly extracted from digital elevation models (DEMs), which are considered alternatives to traditional topographic maps. Before the drainage net is extracted, the DEM is usually processed to remove pits and flat areas that are artificially generated in the production of the DEM. To define the channel network, the direction that water flows on the surface must be determined for every grid point in the DEM. Then, the amount of flow passing through each point can be calculated and the channel heads can be identified.

The first task is the identification of the flow directions. The three prevailing methods are (1) single flow directions, (2) multiple flow directions, and (3) continuous flow directions. When single flow directions are used, water is assumed to travel from a grid point to the neighbor that produces the steepest downward slope (1). When multiple flow directions are used, the flow from a grid point is distributed to all neighbors that are lower than the grid point. The proportion of flow that each neighbor receives depends on the slope produced between the grid point and that neighbor (2). The method of single flow directions is usually adequate for convergent parts of the topography (e.g., valley bottoms), but multiple flow directions is more accurate for areas of divergent flow (e.g., hilltops). The method of continuous flow directions assigns a flow direction in any direction, not just the directions of the eight neighbors (3). Eight triangular facets are formed by connecting a grid point with its neighboring points, and the downslope vector is determined for each of these facets. The flow direction associated with the grid point is the direction of the steepest downslope vector from all eight adjoining facets. Once the flow directions are determined from any of these methods, the number of grid cells (or topographic area) whose flow would pass through a selected grid point can be determined. This accumulated area is known as the contributing area or drainage area.

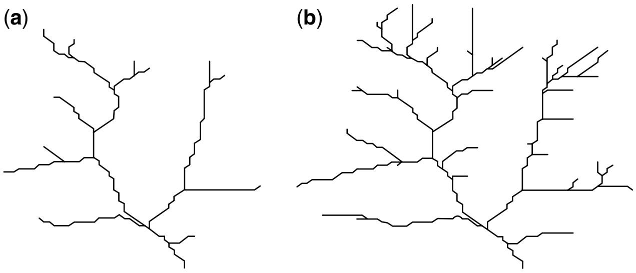

Figure 1. The effect of different critical contributing areas on drainage network extraction from a DEM: (**a**) with a threshold of 125 pixels and (**b**) with a threshold of 25 pixels.

One key issue in deriving channel networks from DEMs is the identification of channel heads. Flow directions and contributing areas are defined for all grid cells in a DEM, but not all grid cells contain a channel. Channel heads are widely defined as the points where fluvial processes become important agents of erosion. The simplest approach for channel head identification is to select a contributing area threshold (4). One way to select this threshold is to identify the contributing area at which “feathering” of the channel network occurs. When the threshold is too small, the smallest channels will often run parallel to each other making the network appear feathered. Figure 1 shows the effect of choosing different contributing area thresholds on the drainage net extracted from a DEM. A series of field studies (5) found that the location of channel heads is better described as a threshold that depends on both the contributing area and the local slope. Observations also suggest that real channel networks are usually much more extensive than the blue lines on USGS quadrangle maps.

STATISTICAL PROPERTIES OF DRAINAGE NETWORKS

The Strahler ordering system (6) is the most common method to identify distinct channels within a channel network. The rules of this ordering system are (1) streams beginning at a channel head are order 1; (2) when two streams of equal order ω meet, a stream of order $\omega + 1$ is created; and (3) when streams of unequal order meet (e.g., ω and $\omega - 1$), the channel segment immediately downstream has the higher order of the two joining streams (e.g., ω). Figure 2 shows an example of the Strahler ordering system. The largest stream order in a basin is common as a qualitative measure of basin size.

The relationship between streams of different order can be described by Horton’s laws (7), which include the laws of stream numbers, stream lengths, and basin areas. The law of stream numbers is expressed as

$$\frac{N(\omega)}{N(\omega + 1)} = R_B$$

where $N(\omega)$ is the number of streams of order ω and R_B is called the bifurcation ratio. The bifurcation ratio characterizes the propensity of the drainage network to branch. The law of stream lengths is expressed as

$$\frac{\overline{L}(\omega + 1)}{\overline{L}(\omega)} = R_L$$

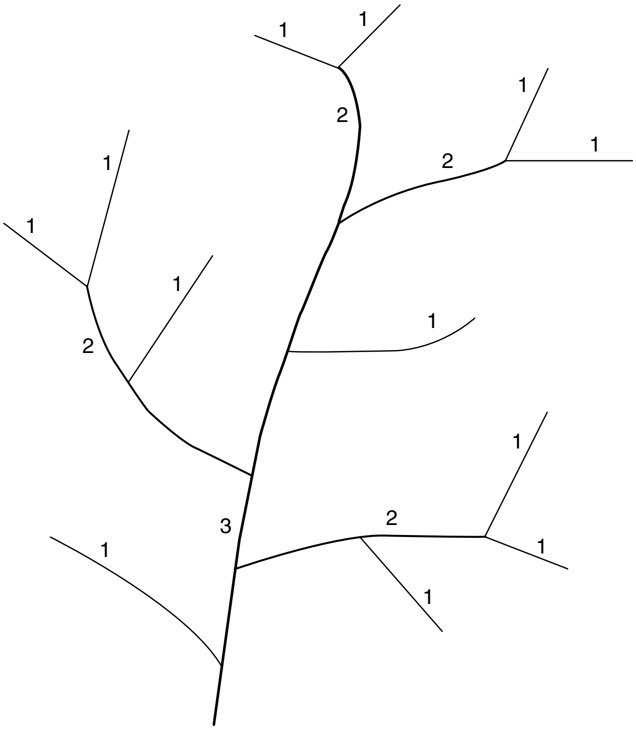

Figure 2. Strahler ordering of a third-order river basin.

where $\overline{L}(\omega)$ is the mean length of streams of order ω and R_L is called the length ratio. The length ratio describes the typical increase in stream length that occurs with increasing stream order. The law of basin areas is expressed as

$$\frac{\overline{A}(\omega+1)}{\overline{A}(\omega)} = R_A$$

where $\overline{A}(\omega)$ is the mean area contributing to streams of order ω and R_A is called the area ratio. Typical values for R_B, R_L, and R_A are 4, 2, and 5, respectively. The fact that Horton's laws hold for most real and simulated drainage networks has led some researchers to suggest that the laws are largely a product of the ordering system and thus not particularly useful for differentiating between different networks.

Natural drainage networks are also considered to be scaling-invariant or fractal objects. Loosely speaking, a fractal is an object that appears statistically similar when viewed up-close at a fine resolution or from a distance at a coarse resolution. Several empirical characteristics of river basins are products of this property. For example, Hack's law (8) describes the power-law relation between the length of a basin's main channel L and the basin's area A:

$$L \propto A^{0.57}$$

The main channel can be defined by identifying the longest channel in a basin when measured from the basin outlet to the channel head along the stream. The fact that the exponent in Hack's law is above 1/2 (the value one would expect from simple geometry) has been shown to be a result of the fractal sinuosity of the stream course. Another product of the scaling-invariance is the slope-area law, which can be written:

$$\overline{S} \propto A^{-\theta}$$

where A is contributing area (or equivalently basin area) and θ is a scaling exponent, which typically ranges from 0.3 to 0.7. $\overline{S}$ is the mean local slope for channel locations that drain basins with area A.

Some impacts of channel networks on the hydrologic response of a basin can be characterized by the drainage density and width function. The drainage density D_d is a measure of topographic texture and stream spacing and can be defined as

$$D_d = \frac{L_T}{A}$$

where L_T is the total length of stream channels in a basin. The drainage density is closely related to the average hillslope length and has been widely employed to predict characteristics of runoff production and the hydrologic response to precipitation events.

The width function is a graph that describes the probability distribution of flow path lengths in a basin. The x-axis of a width function spans the range of distances between the grid points in a basin and the basin outlet, where distance is measured along the flow paths. The width function's y-coordinate indicates the number of grid points that have the specified distance from the basin outlet (topographic area is often used instead of the number of grid points). Notice that the width function is closely related to the basin's unit hydrograph. In fact, if flow speed was constant throughout a basin, the width function would be nearly equivalent to the unit hydrograph.

The basin circularity ratio R_C (9) is a measure of the basin shape and can be defined as

$$R_C = \frac{4\pi A}{P^2}$$

where P is the length of the basin perimeter. It indicates how closely the basin resembles a circle. The closer the circularity is to 1, the greater the similarity to a circle. Another measure of basin shape is the form factor R_F, which can be written as

$$R_F = \frac{A}{L_x^2}$$

where L_x is the maximum length between the basin outlet and the opposite boundary (10). Numerous other measures of basin shape are available. More elongated basins tend to have hydrologic responses that are more distributed in time, whereas compact basins tend to exhibit more peaked hydrographs.

BIBLIOGRAPHY

1. O'Callaghan, J.F. and Mark, D.M. (1984). The extraction of drainage networks from digital elevation data. *Comput. Vis. Graph. Image Process.* **28**: 323–344.

2. Quinn, P.F., Beven, K.J., Chevallier, P., and Planchon, O. (1991). The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models. *Hydrol. Process.* **28**: 328–344.
3. Tarboton, D.G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resour. Res.* **33**: 309–319.
4. O'Callaghan, J.F. and Mark, D.M. (1984). The extraction of drainage networks from digital elevation data. *Comput. Vis. Graph. Image Process.* **28**: 323–344.
5. Montgomery, D.R. and Dietrich, W.E. (1989). Source areas, drainage density, and channel initiation. *Water Resour. Res.* **25**: 1907–1918.
6. Strahler, A.N. (1957). Quantitative analysis of watershed geomorphology. *Am. Geophys. Union Trans.* **38**: 913–920.
7. Horton, R.E. (1945). Erosional development of streams and their drainage density: hydrophysical approach to quantitative geomorphology. *Geol. Soc. Amer. Bull.* **56**: 275–370.
8. Hack, J.T. (1957). *Studies of Longitudinal Stream Profiles in Virginia and Maryland*. U.S. Geol. Surv. Prof. Pap., 294-B.
9. Miller, V.C. (1953). *A Quantitative Geomorphic Study of Drainage Basin Characteristics in the Clinch Mountain area, Virginia and Tennessee*. Department of Geology, Columbia University, Technical Report 3.
10. Horton, R.E. (1932). Drainage basin characteristics. *Am. Geophys. Union Trans.* **13**: 350–361.

DYES AS HYDROLOGICAL TRACERS

Jarai Mon
Markus Flury
Washington State University
Pullman, Washington

INTRODUCTION

Dye tracers have been used in hydrological investigations for more than a century. In 1877, uranine (fluorescein) was used as a tracer to test the hydraulic connection between the Danube River and the Ach Spring in southern Germany (1). In 1883, the French physician des Carriéres successfully proved the source of a typhus epidemic in the city of Auxerre by conducting a tracer experiment with the dye aniline (2). First systematic investigations on the suitability of dyes as tracers were conducted even before the turn of the century (3). Subsequently, the use of dyes as tracers became a common practice in hydrological investigations (4–6).

A classic example of the use of dyes in hydrology is the study of residence times and pathway connectivities in Karst (7). Further applications range from studying dispersion in streams and lakes, to determine sources of water pollution, and to evaluate sewage systems. In the vadose zone, dyes have been mainly used to visualize flow patterns (6).

Thousands of different dyes are commercially available (8), but only a few are suitable for hydrological investigations. Many dyes have been studied specifically for their suitability as hydrological tracers, and recommendations were made on "best" dye tracers (9–12). Depending on the specific applications, different chemical characteristics of a dye may be desirable. For instance, for visualization of water flow in soils, a dye should be clearly visible and trace the water movement accurately. In this particular case, the dye will preferably be blue, red, green, or fluorescent to contrast distinctly from the soil background. The accurate tracing of the water movement demands that the dye does not sorb too strongly to subsurface materials, which poses limitations on the chemical characteristics of a dye.

Here, we summarize tracer characteristics and the applications of dye tracers in surface and subsurface hydrology. We then discuss the limitations and potential problems in using dyes for tracing water flow and solute movement. Selection of an appropriate dye is critical for the success of a tracing study. We present a case study on the application of quantitative structure-activity relationships (QSARs) for screening, selecting, and designing optimal dye tracers for a specific use.

TRACER CHARACTERISTICS OF DYES

Dye tracers, particularly fluorescent dyes, are often preferred over several other types of tracing materials because of their unique characteristics. Many dyes (1) can be readily detected at a concentration as low as a few micrograms per liter, (2) can be quantified with simple and readily available analytical equipment, (3) are nontoxic at low concentrations, and (4) are inexpensive and commercially available in large quantity (7,12,13). In addition, because of their coloring properties, dyes allow us to visualize flow pathways in the subsurface. Many dye tracing studies conducted in the past 10 years have clearly demonstrated that flow patterns in the subsurface are often highly irregular; an example of a nonuniform infiltration front in a sandy soil is shown in Fig. 1.

Despite these desirable characteristics, important drawbacks exist in using dye tracers. Dye tracers are not conservative tracers; i.e., they sorb to subsurface media and do not necessarily move at the same speed as the water to be traced. The sorption behavior of dyes is influenced by the properties of the subsurface materials and the chemistry of the aqueous phase (9,12,14–16). Some dyes degrade when they are exposed to sunlight, e.g., uranine (10,17,18), and some can be degraded by microorganisms. Consequently, dye tracers may behave differently under different natural environments. Thus, the suitability of dye tracers should be tested before they are used in hydrological studies.

An "ideal" water tracer is a substance that (1) has conservative behavior (i.e., does not sorb to solid media, is resistant to degradation, and stable in different chemical environments); (2) does not occur naturally in high concentrations in the system to be investigated; (3) is inexpensive, (4) is easy to apply, sample, and analyze; and (e) is nontoxic to humans, animals, and plants (6). These requirements are difficult to meet for a single chemical. Different types of dyes have been proposed as best suitable water tracers, and these dyes are discussed below.

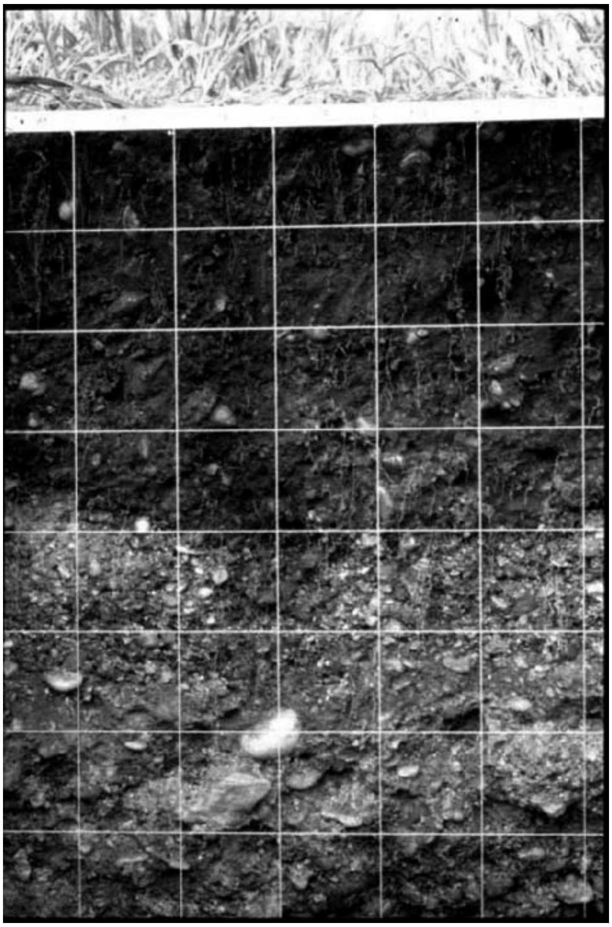

Figure 1. Visualization of flow patterns in soils using a dye tracer (Brilliant Blue FCF). Grid size is 10 cm.

SURFACE WATER, GROUNDWATER, AND VADOSE ZONE TRACERS

Fluorescent dyes are frequently used in surface and groundwater applications and, to some degree, in vadose zone hydrology. Certain dyes, such as rhodamine WT and uranine, are used for surface water, groundwater, and vadose zone applications, whereas others, such as Brilliant Blue FCF, are exclusively used as vadose zone tracers. Many of the common dye tracers (Table 1) belong to the chemical class of the xanthene dyes. The structures of commonly used dyes are shown in Fig. 2.

Dye tracers have been used in measuring flow velocity, travel time, and dispersion in rivers and streams (19,20). Among the dyes commonly used as surface water tracers (Table 1), the most frequently used one is rhodamine WT (20–23). Uranine has been recognized as a good hydrological tracer, but its susceptibility to photochemical decay (17) is of concern in tracing surface water.

Dye tracers have also been used to study groundwater flow velocity, flow direction, hydraulic connections, and aquifer characteristics (4,5). Uranine and rhodamine WT are the two most commonly used tracers in groundwater studies (Table 1). However, these two dyes should not be used as cotracers because rhodamine WT degrades to carboxylic fluorescein, which may confound tracer quantification (13). Rhodamine WT is highly water soluble, easily visible and detectable, photochemically more stable than uranine, and has a moderate tendency for sorption (13). Commercially available tracer-grade rhodamine WT contains two isomers (Fig. 3), which have different sorption properties (24). The para-isomer of rhodamine WT sorbs less to different aquifer materials than did the meta-isomer (24,25). Consequently, the two isomers travel with different velocities in subsurface media, which lead to chromatographic separation (24). In groundwater tracer studies, like in surface water tracing, dye tracers can be easily detected or quantified in water samples using fluorometers or spectrophotometers. Methods and software for designing and analyzing tracer tests are available (13,26–28).

In the vadose zone, dyes are mainly used to delineate water flow patterns. Flow pathways in soils, sediments, and fractured rock have been visualized using dye tracers (29–33). Many dyes have been tested in search for an

Table 1. Dyes Commonly Used As Hydrological Tracers

Commercial Name	C.I. Nr.	C.I. Name	Chemical Class	Fluorescence	Maximum Excitation (nm)[a]	Maximum Emission (nm)[a]	Major Uses
Brilliant Blue FCF	42090	Food Blue 2	Triarylmethane	No	None	630[b]	Vadose zone
Rhodamine WT	none	Acid Red 388	Xanthene	Yes	558[c]	583[c]	Surface water, groundwater, vadose zone
Sulforhodamine B	45100	Acid Red 52	Xanthene	Yes	560	584	Groundwater, vadose zone
Rhodamine B	45170	Basic Violet 10	Xanthene	Yes	555	582	Surface water, groundwater, vadose zone
Sulforhodamine G	45220	Acid Red 50	Xanthene	Yes	535	555	Groundwater
Uranine (Fluorescein)	45350	Acid Yellow 73	Xanthene	Yes	492	513	Groundwater, vadose zone
Eosine	45380	Acid Red 87	Xanthene	Yes	515	535	Groundwater, vadose zone
Methylene Blue	52015	Basic Blue 9	Thiazine	No	None	668[d]	Vadose zone
Lissamine Yellow FF	56205	Acid Yellow 7	Aminoketone	Yes	422	512	Groundwater, vadose zone
Pyranine	59040	Solvent Green 7	Anthraquinone	Yes	460	512	Groundwater

[a] *Source*: Field (27).

[b] Our own data.

[c] Sutton et al. (24) reported the excitation maximum for both the para- and meta-isomers as 555 nm, and the emission maximum as 585 nm for the para-isomer and 588 nm for the meta-isomer.

[d] *Source*: Merck (48).

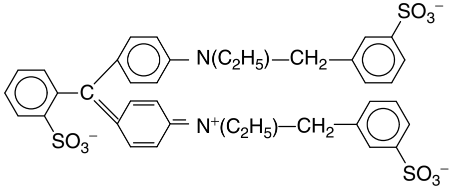

Brilliant blue FCF (C.I. food blue 2, C.I. 42090)
$pK_a = 5.8, 6.6$

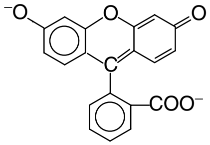

Fluorescein/Uranine (C.I. acid yellow 73, C.I. 45350)
$pK_a = 2, 4\text{-}5, 7$

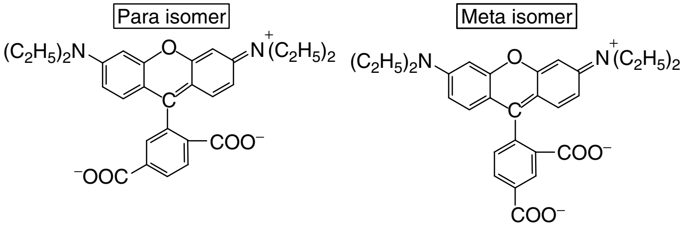

Rhodamine WT (C.I. acid red 388)
$pK_a = 5.1$

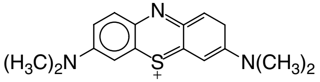

Methylene blue (C.I. basic blue 9, C.I. 52015)
$pK_a = 3.8$

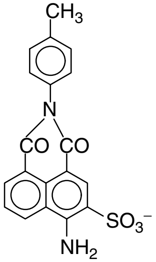

Lissamine yellow FF (C.I. acid yellow 7, C.I. 56205)

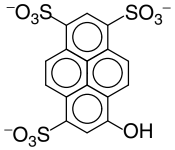

Pyranine (C.I. solvent green 7, C.I. 59040)
$pK_a = 7.3$

Figure 2. Structure of selected dye tracers. Dyes are shown in dissociated form. (Sources of the pK_a values are given in Ref. 6).

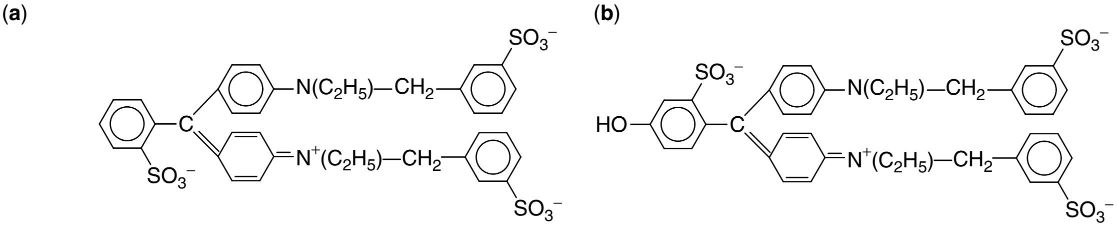

Brilliant blue FCF (C. I. food blue 2)

FD&C green No. 3 (C. I. food green 3)

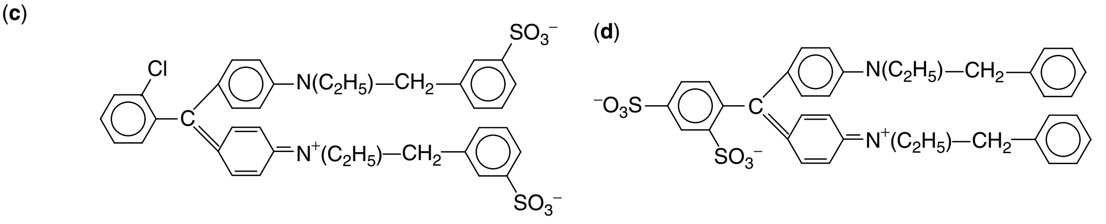

ORCOacid fast green B (C. I. acid green 9)

ORCOacid blue A 150% (C. I. acid blue 7)

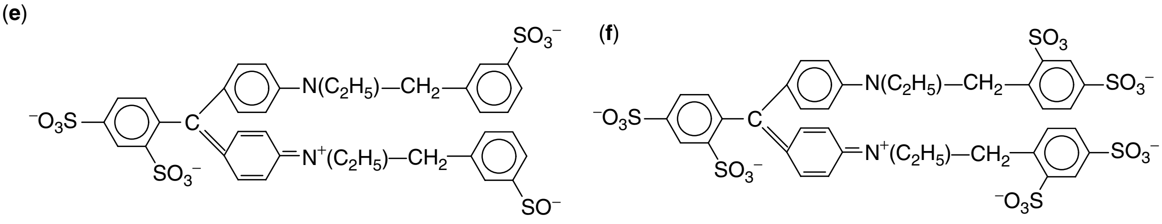

Hypothetical triarylmethane dye

Hypothetical triarylmethane dye

Figure 3. Test triarylmethane dyes that develop the QSAR model (**a**)–(**d**) and hypothetical structure of potential dye tracers (**e**)–(**f**). Dyes are shown in their anionic form.

optimal vadose zone dye tracer, and different dyes have been recommended (6). Most commonly used vadose zone tracers are listed in Table 1. Brilliant Blue FCF has gained acceptance as a good dye tracer for visualization of flow patterns (12,33) and solute transport in the vadose zone (34–36). In the vadose zone, dye tracer analysis is not as simple as in surface water or groundwater tracer studies, particularly if tracer concentrations are to be determined. Image analysis or fiber-optic spectroscopy can be used to measure tracer distributions in soil profiles (37,38).

LIMITATIONS IN USING DYES AS TRACERS

Most dyes are organic molecules, and their interactions with other materials in the subsurface are influenced by environmental conditions. Generally, dye tracers sorb to solid surfaces and the degree of sorption depends on surface properties and solution chemistry. Solubility, photochemical decay, absorption spectra, and fluorescence of dyes are often affected by environmental conditions, such as temperature, sunlight, acidity, and alkalinity. Thus, not only properties of the dyes but also of the environment in which dyes are to be applied often limit the use of dyes as tracers.

Sorption of dyes to subsurface media is one major limitation for using dyes to trace water flow pathways. Sorption causes dyes to move with a slower velocity than water. Some dyes can mimic the movement of certain chemicals rather than the flow of water. For instance, rhodamine WT mimiced the movement of atrazine (39).

Dyes selected as hydrological tracers often contain functional groups, such as carboxylic and sulfonic acids, which contribute to high water solubility and decrease sorption (9,40). However, the functional groups cause dyes to have pH dependent properties. The properties of mineral surfaces may also change with pH, i.e., negatively charged surfaces may become neutral or positively charged as pH decreases, and sorption of anionic dyes may increase. Therefore, the sorption of dyes should be tested before dyes are applied as tracers.

Fluorescence of dyes may change under different environmental conditions. For instance, fluorescence intensity of rhodamine B increases with decreasing temperature (17). The presence of electron donating ions, such as chlorine, bromine, and iodine, in water samples as well as changes in solution pH can cause fluorescence quenching (10,19).

SELECTION OF DYE TRACERS FOR SPECIFIC USES

QSAR Approach as an Alternative to Experimental Screening

Screening is a basic step for selection of the most suitable dye tracers for specific uses, but experimental screening of thousands of commercially available dyes is not practical. An efficient technique (accurate, simple, fast, and inexpensive) is necessary to find the most suitable dye tracer for a specific investigation. A promising screening technique is the use of QSAR.

QSARs relate the molecular structure of a chemical to its activity. Although this technique has been used extensively in pharmacology, it has also been applied to estimate environmental fate and risk of organic chemicals (41–44). QSAR models are based on calculated molecular descriptors and selected measured data that describe the property to be predicted. A statistical model then allows us to predict the properties of structurally similar chemicals that have not yet been experimentally tested.

QSAR Case Study Using Triarylmethane Dyes

We illustrate the use of QSAR for dye tracer screening using the example of the triarylmethane dyes. These dyes are often used as food dyes, and because they are highly water soluble, they have preferable characteristics as dye tracers (6). Brilliant Blue FCF, one member of this dye class, is commonly used as a vadose zone tracer. Other members, however, may be even better suited as dye tracers. We developed a QSAR model with triarylmethane dyes to predict their soil sorption characteristics. Four triarylmethane dyes were selected as a training set: Brilliant Blue FCF (C.I. Food Blue 2), FD&C Green No. 3 (C.I. Food Green 3), ORCOacid Blue A 150% (C.I. Acid Blue 7), and ORCOacid Fast Green B (C.I. Acid Green 9). These four dyes share the same molecular kernel but differ in numbers, types, and positions of functional groups [Fig. 3(a)–(d)].

We experimentally measured soil sorption parameters of the four dyes and used QSAR to relate these parameters to the structural properties of the dyes. Soil sorption was determined by batch sorption experiments similar to the ones described in German-Heins and Flury (16). A sandy soil (Vantage, WA), pH 8, and 0.01 M $CaCl_2$ solution were used for the sorption experiments. A Langmuir sorption isotherm was fitted to the experimental data to obtain the two adsorption parameters, the Langmuir coefficient K_L and the maximum adsorption A_m (Table 2), using a normal nonlinear least-squares method (45). The Langmuir isotherm describes the relation between sorbed (C_a) and aqueous concentrations (C_s) at equilibrium as (45):

$$C_a = \frac{A_m K_L C_s}{1 + K_L C_s} \tag{1}$$

Structural properties (molecular descriptors) of the dyes were calculated using the MDL QSAR (version 2.1, 2002, MDL Information System, Inc., San Leandro, CA). The MDL QSAR program converts molecular structures to structural properties, such as molecular connectivity indices (MCIs), molecular volume, and surface area. Stepwise linear regression analyses were applied to select the descriptors that are well correlated to the experimental parameters (46,47). The statistical significance was assumed at $p \leq 0.05$.

The cross validation technique tested the predictability of the models. Randomization tests were performed to check the probability that correlation occurred by chance. The models that achieved the best quality of statistics were selected for estimation of each sorption parameter. The two QSAR models, one for estimation of K_L and another for estimation of A_m, were established as follows:

1. Langmuir coefficient (K_L) model:

$$K_L = -54.47({}^9\chi_p) + 183.75 \tag{2}$$

where K_L has units of L/mmol and ${}^9\chi_p$ is the ninth-order simple path molecular connectivity index.

2. Maximum adsorption (A_m) model:

$$A_m = -45.72({}^9\chi_p^v) + 35.88 \tag{3}$$

where A_m has units of mmol/kg and ${}^9\chi_p^v$ is the ninth-order valence path molecular connectivity index.

Prediction of Soil Sorption Using QSAR Models

Approximately 70 hypothetical molecules were created based on the structure of Brilliant Blue FCF, and their sorption parameters were estimated using the QSAR models (Equations 2 and 3). These molecules all shared the same molecular kernel as Brilliant Blue FCF but were different in number and position of SO_3 groups. The effects of different numbers and positions of SO_3 groups on soil sorption parameters, i.e., K_L and A_m values, of the new compounds were examined.

Table 2. Comparison of Langmuir Coefficient (K_L) and Maximum Adsorption (A_m) for Test and Hypothetical Triarylmethane Dyes

Triarylmethane Dyes	C.I. Nr.	Number of SO_3 Groups	Langmuir Coefficient K_L (L/mmol)	Maximum Adsorption A_m (mmol/kg)
Test Triarylmethane Dyes			Experimental	
C. I. Food Blue 2	42053	3	5.29	0.42
C. I. Acid Blue 7	42080	2	10.1	2.99
C. I. Food Green 3	42090	3	3.94	0.30
C. I. Acid Green 9	42100	2	16.5	4.40
Hypothetical Triarylmethane Dyes			Predicted	
Dye set 1	none	1	20.9 to 37.8	5.8 to 11.2
Dye set 2	none	2	8.1 to 31.5	2.0 to 8.7
Dye set 3	none	3	−8.5 to 14.7	−2.9 to 4.1

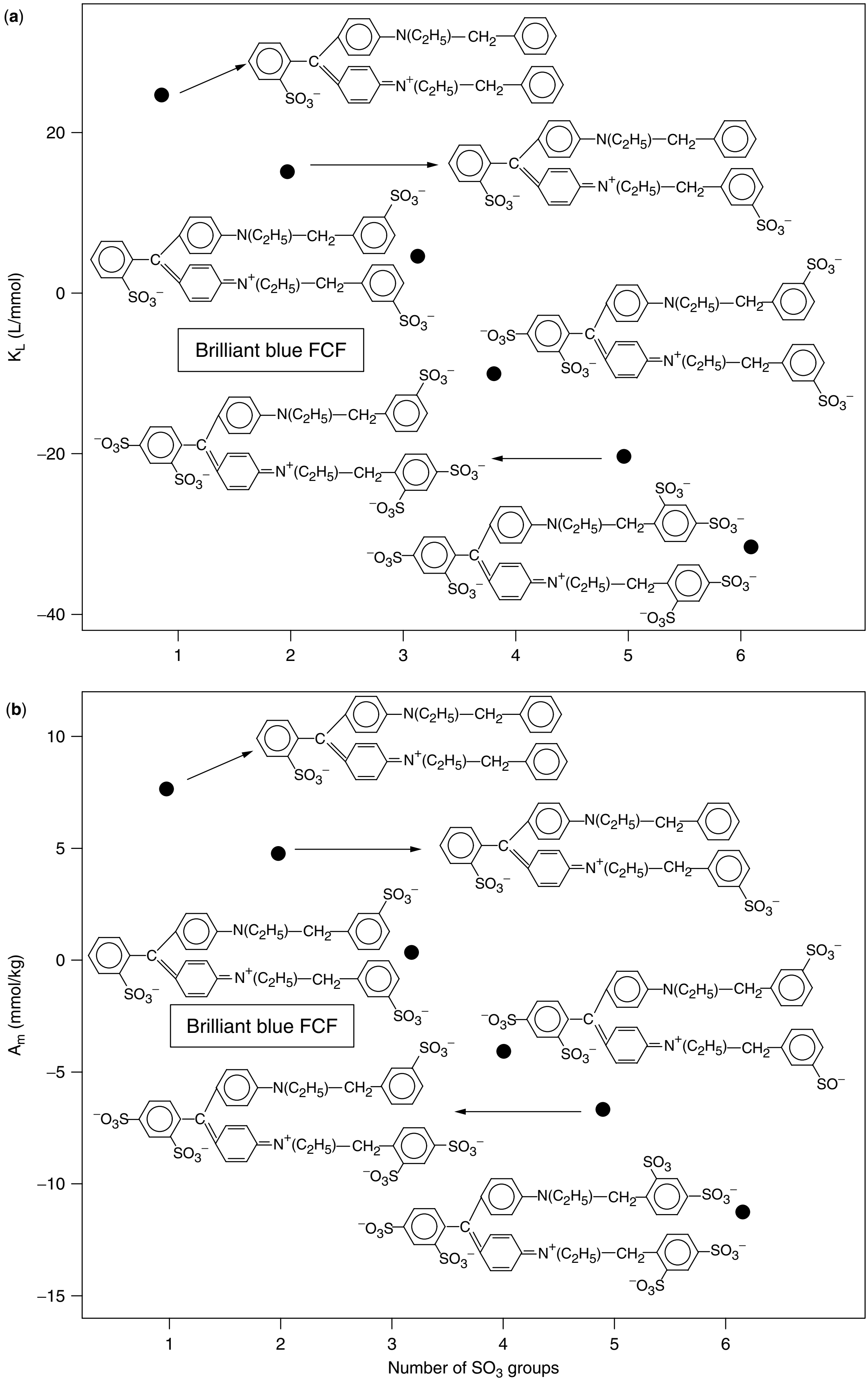

Figure 4. Changes in (**a**) Langmuir coefficient, K_L, and (**b**) adsorption maximum, A_m, as a function of the number of SO_3 groups on the molecular kernel of triarylmethane dyes.

The QSAR modeling indicates that the more SO_3 groups are attached to the molecular kernel, the smaller will be the soil sorption: Both K_L and A_m values decreased with the increasing number of SO_3 groups (Fig. 4). Negative K_L and A_m values were calculated by the models. Negative values are physically not possible, except in the case of ion exclusion. The predicted values should be considered as relative, rather than as absolute, measures for comparing the sorption of the chemicals.

QSAR modeling also to examined the effects of the positions of the SO_3 groups at the molecular kernel. Three sets of hypothetical molecules were created, which contained one, two, or three SO_3 groups attached at different positions at the benzene rings of the triarylmethane kernel. Set 1, which contained one SO_3 group, consisted of six molecules, set 2 (two SO_3 groups) consisted of 22 molecules, and set 3 (three SO_3 groups) consisted of 31 molecules. The range of the predicted K_L and A_m values is listed in Table 2. The large variation in K_L and A_m values within each group of chemicals showed that the sorption parameters were strongly influenced by the positions of the functional groups.

Many dyes in sets 2 and 3 had lower K_L and A_m values than did the four test dyes. The hypothetical dyes with four to six SO_3 groups attached to triarylmethane kernel [Fig. 3(e) and (f)] had considerably smaller K_L and A_m values than did the test dyes. These hypothetical dyes are likely better conservative tracers than are any of the test dyes.

The K_L and A_m values of C.I. Food Green 3 were lower than were those of C.I. Food Blue 2 (Brilliant Blue FCF). Between these two readily available dyes, C.I. Food Green 3 may be a better tracer than Brilliant Blue FCF for hydrological investigations in the vadose zone.

SUMMARY

Dye tracers are frequently used in hydrological investigations. Although dyes have unique tracer characteristics, some limitations and problems are associated with using dyes as hydrological tracers. Most dyes sorb to subsurface media, so that tracer characteristics of dyes should be tested under the specific conditions under which dye tracing is to be conducted.

An accurate and cost-effective screening technique is necessary for selection of optimal dye tracers. QSARs offer a powerful tool for screening of a large number of dyes in a short time. We conducted a QSAR case study using triarylmethane dyes. The results of the QSAR modeling indicate that many hypothetical triarylmethane dyes have considerably lower sorption characteristics than do the triarylmethane dyes currently used as tracers, and they likely are good tracer candidates.

BIBLIOGRAPHY

1. Knop, A. (1878). Über die hydrographischen Beziehungen zwischen der Donau und der Aachquelle im badischen Oberlande (Schluss). *Neues Jahrb. Mineral. Geol. Paleontol.* 350–363.
2. des Carrières, D. (1883). Étiologie de l'épidémie typhoïde qui a éclaté a Auxerre en septembre 1882, *Bull. et mém. Soc. méd. des hôpitaux de Paris 9, 2d ser*, pp. 277–284.
3. Trillat, M.A. (1899). Sur l'emploi des matières colorantes pour la recherche de l'origine des sources et des eaux d'infiltration. *C. R. Hebd. Seances Acad. Sci.* **128**: 698–702.
4. Davis, S.N., Thompson, G.M., Bentley, H.W., and Stiles, G. (1980). Ground-water tracers—A short review. *Ground Water* **18**: 14–23.
5. McLaughlin, M.J. (1982). A review on the use of dyes as soil water tracers. *Water SA* **8**: 196–201.
6. Flury, M. and Wai, N.N. (2003). Dyes as tracers for vadose zone hydrology. *Rev. Geophys.* **41**: 1002.
7. Aley, T. (1997). Dyes don't lie: Practical karst hydrology In: *Karst-Water Environment Symposium Proceedings*. T. Younos, T.J. Burbey, E.H. Kastning, and J.A. Poff (Eds.). Virginia Water Resources Research Center, Roanoke, VA, pp. 1–10.
8. Colour Index. (2001). *Colour Index International Fourth Edition Online. Society of Dyers and Colourists and the American Association of Textile Chemists and Colorists*. Available: http://www.colour-index.org, accessed September 2004.
9. Corey, J.C. (1968). Evaluation of dyes for tracing water movement in acid soils. *Soil Sci.* **106**: 182–187.
10. Smart, P.L. and Laidlaw, I.M.S. (1977). An evaluation of some fluorescent dyes for water tracing. *Water Resour. Res.* **13**: 15–33.
11. Smettem, K.R.J. and Trudgill, S.T. (1983). An evaluation of some fluorescent and non-fluorescent dyes in the identification of water transmission routes in soils. *J. Soil Sci.* **34**: 45–56.
12. Flury, M. and Flühler, H. (1995). Tracer characteristics of Brilliant Blue FCF. *Soil Sci. Soc. Am. J.* **59**: 22–27.
13. Field, M.S. (2002). The QTRACER2 Program for Tracer Breakthrough Curve Analysis for Tracer Tests in Karstic Aquifers and Other Hydrologic Systems. Office of Research and Development, U.S. Environmental Protection Agency, EPA/600/R-02/001, Washington, DC.
14. Reynolds, E.R.C. (1966). The percolation of rainwater through soil demonstrated by fluorescent dyes. *J. Soil Sci.* **17**: 127–132.
15. Kasnavia, T., Vu, D., and Sabatini, D.A. (1999). Fluorescent dye and media properties affecting sorption and tracer selection. *Ground Water* **37**: 376–381.
16. German-Heins, J. and Flury, M. (2000). Sorption of Brilliant Blue FCF in soils as affected by pH and ionic strength. *Geoderma* **97**: 87–101.
17. Feuerstein, D.L. and Selleck, R.E. (1963). Fluorescent tracers for dispersion measurements. *J. Sanit. Eng. Div. Am. Soc. Civ. Eng.* **89**: 1–21.
18. Viriot, M.L. and André, J.C. (1989). Fluorescent dyes: a search for new tracer for hydrology. *Analysis* **17**: 97–111.
19. Church, M. (1974). Electrochemical and Fluorometric Tracer Techniques for Streamflow Measurements; British Geomorphological Research Group, Technical Bulletin No. 12, Geo Abstracts Ltd., University of East Anglia, Norwich, UK.
20. Cox, M.H., Mendez, G.O., Kratzer, C.R., and Reichard, E.G. (2003). Evaluation of Tracer Tests Completed in 1999 and 2000 on the Upper Santa Clara River, Los Angeles and Ventura Counties, California; U.S. Geological Survey, Water Resources Investigations Report 03-4277, Sacramento, CA.
21. Kratzer, C.R., and Biagtan, R.N. (1998). Determination of Traveltimes in the Lower San Joaquin River Basin, California, from Dye-Tracer Studies during 1994-1995; U.S.

Geological Survey, Water Resources Investigations Report, 97-4018.

22. Imes, J.L. and Fredrick, B.S. (2002). Using Dye-Tracing and Chemical Analyses to Determine Effects of a Wastewater Drainage to Jam Up Creek on Water Quality of Big Spring, Southeastern Missouri, 2001; U.S. Geological Survey, USGS Fact Sheet, FS-103-02, Rolla, MO.
23. Gooseff, M.N., Wondzell, S.M., Haggerty, R., and Aderson, J. (2003). Comparing transient storage modeling and residence time distribution (RTD) analysis in geomorphically varied reaches in the Lookout Creek basin, Oregon, USA. *Adv. Water Resour.* **26**: 925–937.
24. Sutton, D.J., Kabala, Z.J., Francisco, A., and Vasudevan, D. (2001). Limitation and potential of commercially available Rhodamine WT as a groundwater tracer. *Water Resour. Res.* **37**: 1641–1656.
25. Vasudevan, D., Fimmen, R.L., and Francisco, A.B. (2001). Tracer-grade Rhodamine WT: Structure of constituent isomers and their sorption behavior. *Environ. Sci. Technol.* **35**: 4089–4096.
26. Field, M.S. (2002). Efficient hydrologic tracer-test design for tracer-mass estimation and sample-collection frequency, 1. Method development. *Environ. Geol. (NY)* **42**: 827–838.
27. Field, M.S. (2003). A review of some tracer-test design for tracer-mass estimation and sample-collection frequency. *Environ. Geology (NY)* **43**: 867–881.
28. Field, M.S. (2003). Tracer-Test Planning Using the Efficient Hydrologic Tracer-Test Design (EHTD) Program; Office of Research and Development, U.S. Environmental Protection Agency, EPA/600/R-03/034, Washington, DC.
29. Bouma, J., Jongerius, A., Boersma, O., Jager, A., and Schoonderbeek, D. (1977). The function of different types of macropores during saturated flow through four swelling soil horizons. *Soil Sci. Soc. Am. J.* **41**: 945–950.
30. Germann, P.F., Edwards, W.M., and Owens, L.B. (1984). Profiles of bromide and increased soil moisture after infiltration into soils with macropores. *Soil Sci. Soc. Am. J.* **48**: 237–244.
31. Flury, M., Flühler, H., Jury, W.A., and Leuenberger, J. (1994). Susceptibility of soils to preferential flow of water: A field study. *Water Resour. Res.* **30**: 1945–1954.
32. Hu, Q., Kneafsey, T.J., Trautz, R.C., and Wang, J.S.Y. (2002). Tracer penetration into welded tuff matrix from flowing fractures. *Vadose Zone J.* **1**: 102–112.
33. Nobles, M.M., Wilding, L.P., and McInnes, K.J. (2004). Pathways of dye tracer movement through structured soils on a macroscopic scale. *Soil Sci.* **169**: 229–242.
34. Vanderborght, J., Gähwiller, P., Wydler, H., Schultze, U., and Flühler, H. (2002). Imaging fluorescent dye concentrations on soil surfaces: Uncertainty of concentration estimates. *Soil Sci. Soc. Am. J.* **66**: 760–773.
35. Öhrström, P., Hamed, Y., Persson, M., and Berndtsson, R. (2004). Characterizing unsaturated solute transport by simultaneous use of dye and bromide. *J. Hydrol. (Amsterdam)* **289**: 23–35.
36. Zinn, B. et al. (2004). Experimental visualization of solute transport and mass transfer process in two-dimensional conductivity fields with connected regions of high conductivity. *Environ. Sci. Technol.* **38**: 3916–3926.
37. Forrer, I., Papritz, A., Flühler, H., and Luca, D. (2000). Quantifying dye tracers in soil profiles by image processing. *Eur. J. Soil Sci.* **51**: 313–322.
38. Aeby, P. et al. (2001). Fluorescence imaging of tracer distributions in soil profiles. *Environ. Sci. Technol.* **35**: 753–760.
39. Kanwar, R.S., Baker, J.L., and Singh, P. (1997). Use of chloride and fluorescent dye as tracers in measuring nitrate and atrazine transport through soil profile under laboratory conditions. *J. Environ. Sci. Health, Part A* **32**: 1907–1919.
40. Reife, A. and Freeman, H.S. (1996). Carbon adsorption of dyes and selected intermediates. In: *Environmental Chemistry of Dyes and Pigments*. A. Reife and H.S. Freeman (Eds.). John Wiley & Sons, New York, pp. 3–31.
41. Sabljić, A. (1989). Quantitative modeling of soil sorption for xenobiotic chemicals. *Environ. Health Perspect*. **83**: 179–190.
42. Nendza, M. (1998). *Structure Activity Relationships in Environmental Sciences*. Chapman & Hall, London, UK.
43. Sabljić, A. (2001). QSAR models for estimating properties of persistent organic pollutants required in evaluation of their environmental fate and risk. *Chemosphere* **43**: 363–375.
44. Worrall, F. (2001). A molecular topology approach to predicting pesticide pollution of groundwater. *Environ. Sci. Technol.* **35**: 2282–2287.
45. Schulthess, C.P. and Dey, D.K. (1996). Estimation of Langmuir constants using linear and nonlinear least squares regression analyses. *Soil Sci. Soc. Am. J.* **60**: 433–442.
46. Sekusak, S. and Sabjlić, A. (1992). Soil sorption and chemical topology. *J. Math. Chem.* **11**: 271–280.
47. Hall, L.H., Kier, L.B., and Hall, L.M. (2002). *The Guide for Development of QSAR with MDL QSAR*. MDL Information System, San Leandro, CA.
48. Merck. (1996). *The Merck Index. An Encyclopedia of Chemicals, Drugs, and Biologicals*, 12th Edn. Merck & Co, Inc., Whitehouse Station, NJ.

FLOW-DURATION CURVES

Robert M. Hordon
Rutgers University
Piscataway, New Jersey

The flow-duration curve is a cumulative frequency curve that shows the percentage of time that specified stream's discharges were equaled or exceeded during a period of record (1). It combines in one curve all of the flow characteristics of a stream, ranging from flood flows to drought situations.

In sharp contrast to a hydrograph where time is on the *x* axis, chronology is not shown in flow-duration curves. If the period of record on which the curve is based is sufficiently long, it may be considered a probability curve and therefore, is used to estimate the percentage of time that a specified stream's discharge will be equaled or exceeded in the future.

INTRODUCTION

The temporal sequence of flows is not included in the preparation of flow-duration curves, so one cannot tell from the curve itself whether periods of low or high flow occurred during one dry or wet period or were scattered over several years (2). However, the curve is very useful

Geological Survey, Water Resources Investigations Report, 97-4018.

22. Imes, J.L. and Fredrick, B.S. (2002). Using Dye-Tracing and Chemical Analyses to Determine Effects of a Wastewater Drainage to Jam Up Creek on Water Quality of Big Spring, Southeastern Missouri, 2001; U.S. Geological Survey, USGS Fact Sheet, FS-103-02, Rolla, MO.
23. Gooseff, M.N., Wondzell, S.M., Haggerty, R., and Aderson, J. (2003). Comparing transient storage modeling and residence time distribution (RTD) analysis in geomorphically varied reaches in the Lookout Creek basin, Oregon, USA. *Adv. Water Resour.* **26**: 925–937.
24. Sutton, D.J., Kabala, Z.J., Francisco, A., and Vasudevan, D. (2001). Limitation and potential of commercially available Rhodamine WT as a groundwater tracer. *Water Resour. Res.* **37**: 1641–1656.
25. Vasudevan, D., Fimmen, R.L., and Francisco, A.B. (2001). Tracer-grade Rhodamine WT: Structure of constituent isomers and their sorption behavior. *Environ. Sci. Technol.* **35**: 4089–4096.
26. Field, M.S. (2002). Efficient hydrologic tracer-test design for tracer-mass estimation and sample-collection frequency, 1. Method development. *Environ. Geol. (NY)* **42**: 827–838.
27. Field, M.S. (2003). A review of some tracer-test design for tracer-mass estimation and sample-collection frequency. *Environ. Geology (NY)* **43**: 867–881.
28. Field, M.S. (2003). Tracer-Test Planning Using the Efficient Hydrologic Tracer-Test Design (EHTD) Program; Office of Research and Development, U.S. Environmental Protection Agency, EPA/600/R-03/034, Washington, DC.
29. Bouma, J., Jongerius, A., Boersma, O., Jager, A., and Schoonderbeek, D. (1977). The function of different types of macropores during saturated flow through four swelling soil horizons. *Soil Sci. Soc. Am. J.* **41**: 945–950.
30. Germann, P.F., Edwards, W.M., and Owens, L.B. (1984). Profiles of bromide and increased soil moisture after infiltration into soils with macropores. *Soil Sci. Soc. Am. J.* **48**: 237–244.
31. Flury, M., Flühler, H., Jury, W.A., and Leuenberger, J. (1994). Susceptibility of soils to preferential flow of water: A field study. *Water Resour. Res.* **30**: 1945–1954.
32. Hu, Q., Kneafsey, T.J., Trautz, R.C., and Wang, J.S.Y. (2002). Tracer penetration into welded tuff matrix from flowing fractures. *Vadose Zone J.* **1**: 102–112.
33. Nobles, M.M., Wilding, L.P., and McInnes, K.J. (2004). Pathways of dye tracer movement through structured soils on a macroscopic scale. *Soil Sci.* **169**: 229–242.
34. Vanderborght, J., Gähwiller, P., Wydler, H., Schultze, U., and Flühler, H. (2002). Imaging fluorescent dye concentrations on soil surfaces: Uncertainty of concentration estimates. *Soil Sci. Soc. Am. J.* **66**: 760–773.
35. Öhrström, P., Hamed, Y., Persson, M., and Berndtsson, R. (2004). Characterizing unsaturated solute transport by simultaneous use of dye and bromide. *J. Hydrol. (Amsterdam)* **289**: 23–35.
36. Zinn, B. et al. (2004). Experimental visualization of solute transport and mass transfer process in two-dimensional conductivity fields with connected regions of high conductivity. *Environ. Sci. Technol.* **38**: 3916–3926.
37. Forrer, I., Papritz, A., Flühler, H., and Luca, D. (2000). Quantifying dye tracers in soil profiles by image processing. *Eur. J. Soil Sci.* **51**: 313–322.
38. Aeby, P. et al. (2001). Fluorescence imaging of tracer distributions in soil profiles. *Environ. Sci. Technol.* **35**: 753–760.
39. Kanwar, R.S., Baker, J.L., and Singh, P. (1997). Use of chloride and fluorescent dye as tracers in measuring nitrate and atrazine transport through soil profile under laboratory conditions. *J. Environ. Sci. Health, Part A* **32**: 1907–1919.
40. Reife, A. and Freeman, H.S. (1996). Carbon adsorption of dyes and selected intermediates. In: *Environmental Chemistry of Dyes and Pigments*. A. Reife and H.S. Freeman (Eds.). John Wiley & Sons, New York, pp. 3–31.
41. Sabljić, A. (1989). Quantitative modeling of soil sorption for xenobiotic chemicals. *Environ. Health Perspect*. **83**: 179–190.
42. Nendza, M. (1998). *Structure Activity Relationships in Environmental Sciences*. Chapman & Hall, London, UK.
43. Sabljić, A. (2001). QSAR models for estimating properties of persistent organic pollutants required in evaluation of their environmental fate and risk. *Chemosphere* **43**: 363–375.
44. Worrall, F. (2001). A molecular topology approach to predicting pesticide pollution of groundwater. *Environ. Sci. Technol.* **35**: 2282–2287.
45. Schulthess, C.P. and Dey, D.K. (1996). Estimation of Langmuir constants using linear and nonlinear least squares regression analyses. *Soil Sci. Soc. Am. J.* **60**: 433–442.
46. Sekusak, S. and Sabjlić, A. (1992). Soil sorption and chemical topology. *J. Math. Chem.* **11**: 271–280.
47. Hall, L.H., Kier, L.B., and Hall, L.M. (2002). *The Guide for Development of QSAR with MDL QSAR*. MDL Information System, San Leandro, CA.
48. Merck. (1996). *The Merck Index. An Encyclopedia of Chemicals, Drugs, and Biologicals*, 12th Edn. Merck & Co, Inc., Whitehouse Station, NJ.

FLOW-DURATION CURVES

Robert M. Hordon
Rutgers University
Piscataway, New Jersey

The flow-duration curve is a cumulative frequency curve that shows the percentage of time that specified stream's discharges were equaled or exceeded during a period of record (1). It combines in one curve all of the flow characteristics of a stream, ranging from flood flows to drought situations.

In sharp contrast to a hydrograph where time is on the *x* axis, chronology is not shown in flow-duration curves. If the period of record on which the curve is based is sufficiently long, it may be considered a probability curve and therefore, is used to estimate the percentage of time that a specified stream's discharge will be equaled or exceeded in the future.

INTRODUCTION

The temporal sequence of flows is not included in the preparation of flow-duration curves, so one cannot tell from the curve itself whether periods of low or high flow occurred during one dry or wet period or were scattered over several years (2). However, the curve is very useful for studying the flow characteristics of a stream for the entire range of discharge that can vary across several orders of magnitude.

Although flow-duration curves can be based on weekly or monthly averages, the greatest accuracy is reached in using the day as the unit of time. Therefore, it is recommended that flow-duration curves used in hydrologic analysis be based on mean daily discharges for the period of record.

As streamflow generally has a log-normal distribution, flow-duration curves are usually plotted on logarithmic probability graphs where the *x* axis represents the percentage probability that a given discharge was equaled or exceeded during a given period and the *y* axis indicates discharge in either cubic feet/second or cubic meters/second. A sample flow-duration curve for Bowie Creek near Hattiesburg, Mississippi is shown in Fig. 1. Note that the *x* axis ranges from 0.05 to 99.99% and the *y* axis is scaled logarithmically from 100 to 10,000 cubic feet/second (two orders of magnitude).

THE RELATIONSHIP BETWEEN CURVE SHAPE AND GEOHYDROLOGY

It is widely recognized that the shape of the flow-duration curve provides very useful information about the geohydrologic characteristics of watersheds. Steeply sloping curves indicate “flashy” streams where the flow is largely from direct runoff and where there is limited groundwater storage. For example, the flow-duration curves for the South Branch Raritan and Great Egg Harbor Rivers in New Jersey for the 1931–1960 time period are shown in Fig. 2. The South Branch Raritan River in North Jersey drains consolidated rock formations such as Precambrian granites and Paleozoic limestones. In contrast, Great Egg Harbor River in the unconsolidated formations of the Coastal Plain of South Jersey is underlain by thick seaward-dipping deposits of sand (3). The geohydrologic differences between the consolidated rocks of North Jersey and the unconsolidated sedimentary formations of the Coastal Plain are illustrated by the different slopes in Fig. 2. First, discharge on the *y* axis for both rivers has been equilibrated by using cubic feet per second/square mile values to make the flow independent of watershed size (4). Second, note in Fig. 2, that the discharge values are higher for the South Branch Raritan River at high flow conditions (0.01–20% exceedance values) and lower at low flow conditions (80–99.5% exceedance values). This difference is to be expected as the groundwater storage is much greater in the Coastal Plain than in the consolidated rock formations of northern New Jersey.

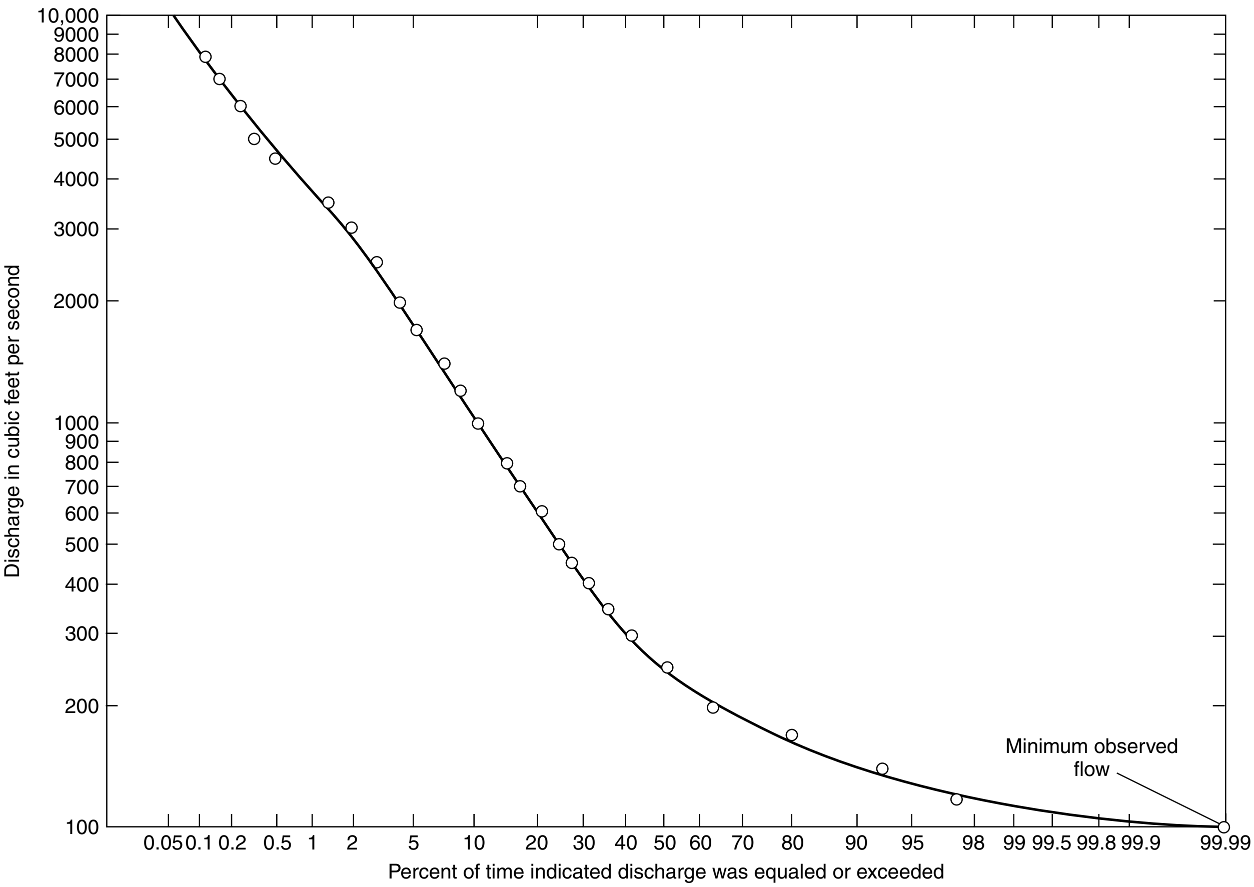
Manual of hydrology: part 2, low-flow techniques

Figure 1. Duration curve of daily flow, Bowle Creek near Hattiesburg, Mississippi, 1939–1948.

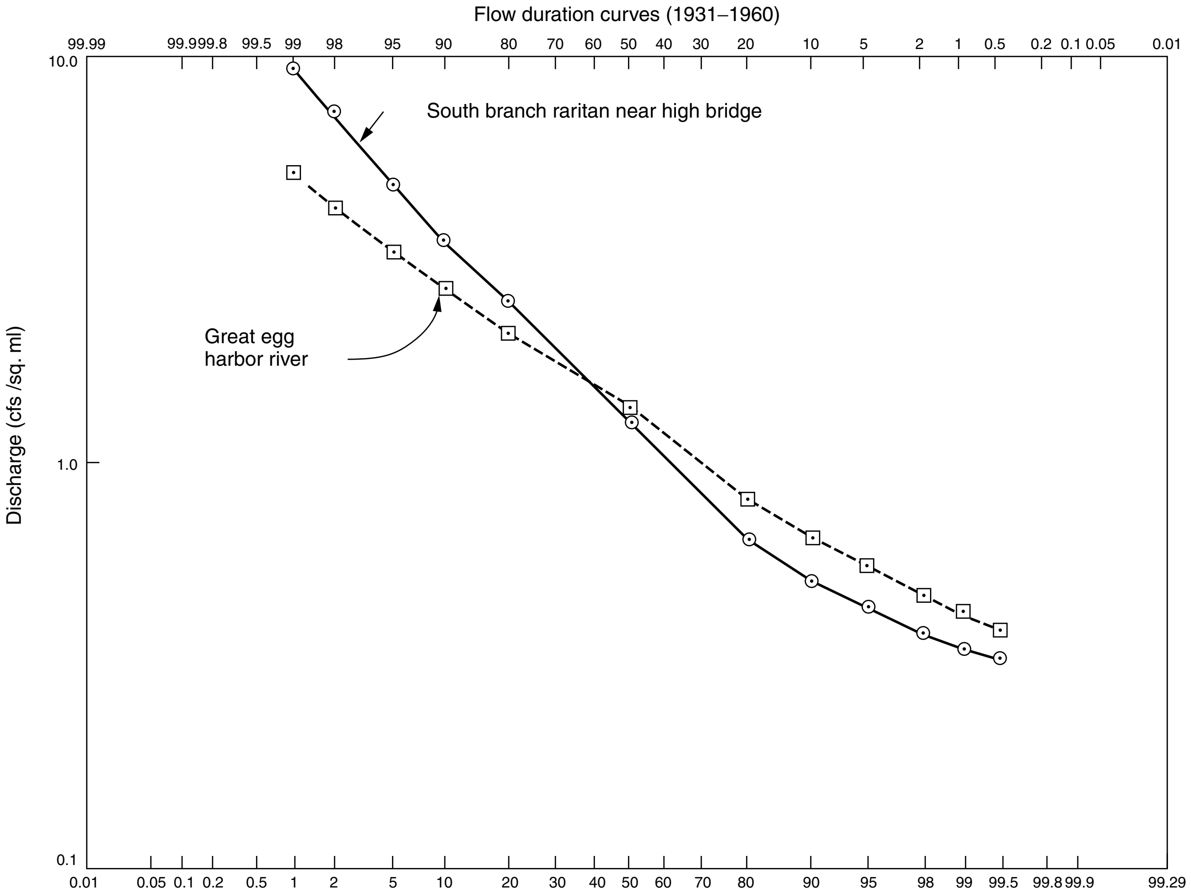

Figure 2. Percentage of time indicated discharge was equaled or exceeded.

An even better example of differential geohydrology and its effect on flow-duration curves is shown in Fig. 3. Both the South Branch Raritan River and Stony Brook at Princeton in New Jersey are underlain by consolidated rock formations. However, the geologic formations within the Stony Brook watershed include a large proportion of Lockatong argillite, a very tight, poorly fractured rock that is notorious for its poor groundwater yields and limited water storage. As in Fig. 2, discharge on the y axis has been equilibrated, in this case by using (mgd) million gallon per day/square mile values.

Note the precipitous decline of the slope of the curve for Stony Brook in Fig. 3, indicating the strong effects of the minimal amounts of groundwater in storage in the argillite formation. The flow-duration curve data were plotted on log-probability graphs, so the decline in expected flow is even more pronounced.

VARIABILITY INDEXES

One of the major characteristics of streamflow is its variability. This variability reflects the variability in precipitation as modified by the physical characteristics of the watershed. These physical characteristics include the differential geohydrology of consolidated and unconsolidated rock formations and the water storage available in lakes, swamps, and wetlands. Another factor that will become of increasing importance in streamflow variability is the anthropogenic factor as exemplified by the effects of urbanization on the hydrologic cycle. Streams are subject to more regulation in the form of water supply diversions, interbasin transfers of both raw water and treated effluent, and low flow release requirements, so the natural flow of a stream may become very difficult to measure, particularly in heavily urbanized smaller basins.

One quantitative measure of streamflow variability is the slope of the flow-duration curve itself. The steeper the slope, the greater the variability. A numerical index of variability was introduced by Lane and Lei (5) that is defined as the standard deviation of the logarithms of stream discharge at 10 points on the curve between the 5% and 95% exceedance values (5, 15, 25, ... 95%).

Miller (6) noted that many flow-duration curves tend to be nearly straight lines on log-probability paper between the 20 and 80% exceedance points. The curves above and below these points depart strongly from a straight line and would not be suitable for a numerical index. Thus, Miller (6) suggested using a simple index by dividing the discharge at the 20% point on the curve by the discharge at the 80% point. Flashy streams have high indexes, whereas streams of relatively uniform flows (such as those in the Coastal Plain) have low indexes. One additional benefit of this easily calculated index is that the resulting ratio is dimensionless, thereby facilitating the comparison of streamflow variability in varied locations.

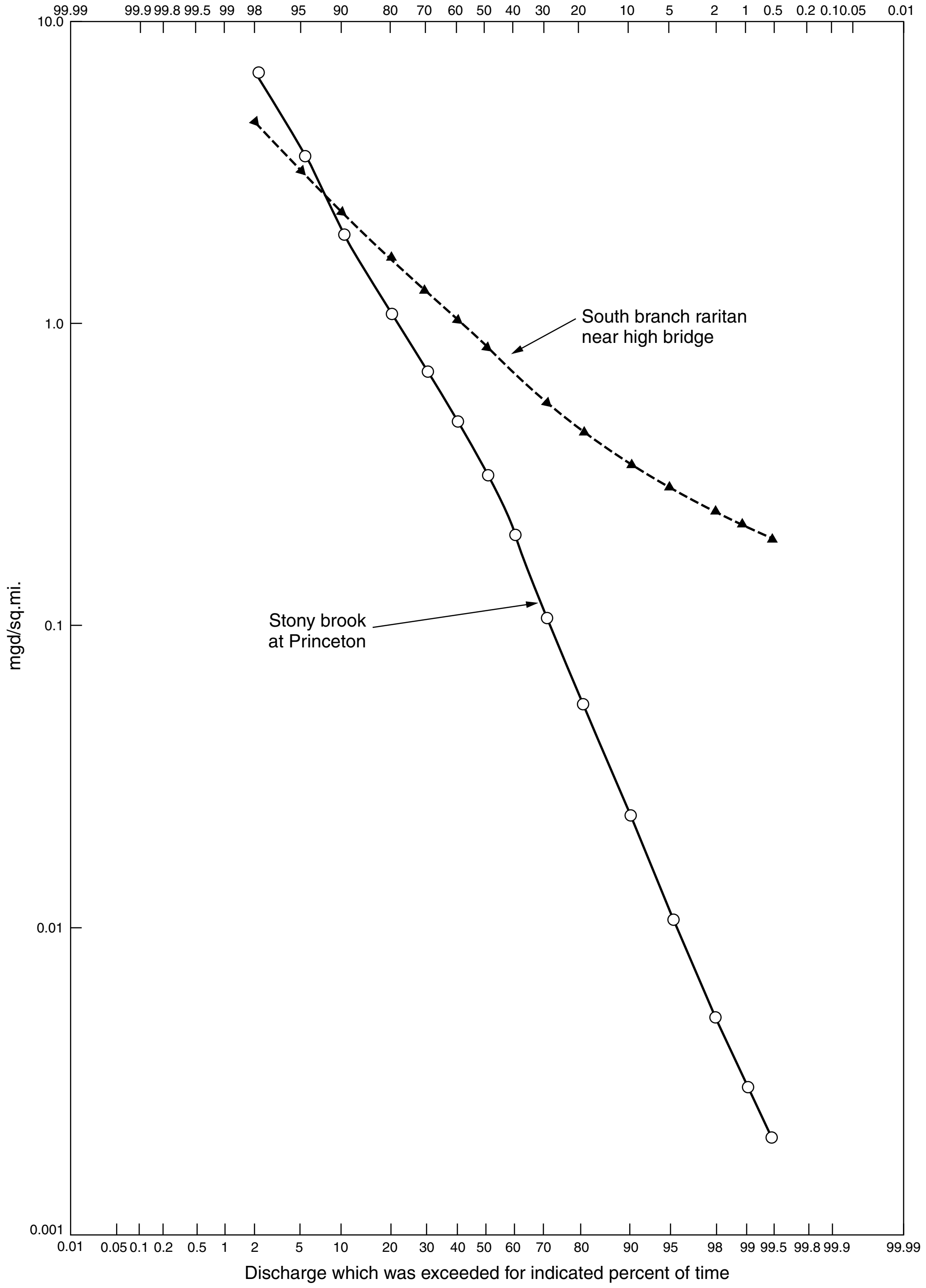

Figure 3. Flow-duration curves; Stony Brook at Princeton, 1954–1975; South Branch Raritan near High Bridge, 1919–1975.

SUMMARY

A flow-duration curve is a graph, usually plotted on log-probability paper, that shows the percentage of time that streamflow at a gaging station is either equaled or exceeded. Important information about the hydrologic and geologic characteristics of a watershed is revealed in the shape of the curve. Curves that slope steeply represent streams of very variable flows that reflect large inputs of overland runoff. Conversely, curves of gentler slopes are indicative of basins that have large amounts of groundwater in storage that result in flows of higher constancy.

BIBLIOGRAPHY

1. Searcy, J.K. (1959). *Flow-Duration Curves*. U.S. Geological Survey Water-Supply Paper 1542-A, Washington, DC.
2. Gillespie, B.D. and Schopp, R.D. (1982). *Low-Flow Characteristics and Flow Duration of New Jersey Streams*. U.S. Geological Survey Open-File Report 81–1110, Trenton, NJ.

3. Watt, M.K. (2000). *A Hydrologic Primer for New Jersey Watershed Management*. U.S. Geological Survey Water-Resources Investigation Report 00-4140. West Trenton, NJ.
4. Fetter, C.W. (1994). *Applied Hydrogeology*, 3rd Edn. Prentice-Hall, Englewood Cliffs, NJ.
5. Lane, E.W. and Lei, K. (1950). Stream Flow Variability. *Transactions of the American Society of Civil Engineers* **115**: 1084–1134.
6. Miller, E.G. (1966). *Flow Probability of New Jersey Streams*. New Jersey, Division of Water Policy and Supply, Water Resources Circular 15. (Prepared in cooperation with the U.S. Geological Survey).

ENVIRONMENTAL FLOWS

GEOFFREY PETTS
University of Birmingham
Birmingham, United Kingdom

The ecological integrity of riverine ecosystems depends on their natural dynamic character (1). Sustaining the natural functioning of aquatic and water-dependent ecosystems in the face of growing human demands is a major challenge (2). Its solution requires the allocation of water to protect aquatic ecosystems from the impacts of abstraction and river flow regulation (see Regulated Rivers), and this is one of five facets to freshwater management for the twenty-first century (see Water in History). It also requires acceptance of a multiuse ethic to provide flexible water budgets to support ecological functions in both dry and wet years (3).

Rivers are the arteries of fluvial hydrosystems comprising the river itself, riparian zone, floodplain, and alluvial aquifer. These four-dimensional systems are influenced by longitudinal processes, vertical and lateral fluxes, and strong temporal changes. Flow is the driver. Flow defines the environmental domains within which biological communities develop, including the vertical exchanges between surface and groundwater. High flows interact with the sediment load, and also with downed trees and driftwood (4), to shape the channel and floodplain morphology and to structure the complex mosaic of physical habitats that comprise a river corridor.

A major constraint to the advancement of tools for allocating flow to meet in-river needs was the lack of scientifically based models (5). In the 1960s and 1970s, early attempts to set *instream flows* for rivers focused on the annual minimum flow expressed as a hydrological statistic, commonly as either a flow duration statistic (such as the 95th percentile flow) or as a fixed percentage of the average daily flow (ADF), with several studies proposing 20% ADF to protect aquatic habitat in streams.

From the early 1980s, more complex approaches to determine instream flows were introduced that make the assumption that physical habitat attributes provide an index of suitability for biota. The most well known of these flow–habitat methodologies, Physical Habitat Simulation (PHABSIM), integrates the changing hydraulic conditions with discharge and the habitat preferences of one or more selected species. The method relies on three principles: The chosen species exhibits preferences within a range of habitat conditions that it can tolerate; these ranges can be defined for each species; and the area of stream providing these conditions can be quantified as a function of discharge and channel structure. In the majority of cases, these methodologies developed instream-flow guidelines that focused on the needs of a single species, usually a salmon or trout, although more advanced approaches considered the needs of different life stages. More recent advances of these approaches include, for example, assessment of flows (1) to encourage the upstream migration of salmon from the estuary to the spawning grounds, (2) to provide good quality habitat for spawning, and (3) to maintain the intragravel environment for egg development over winter for fry at emergence in late winter and for juvenile fish in spring and summer.

By the early 1990s, the science and management of *regulated rivers* had expanded from the determination of *instream flows* to *environmental flows*. Many schemes now applied more complex flow–habitat models to address wider issues than the instream needs, for example, the *hydraulic habitats* of a single species. These new approaches address the sustainability of communities and ecosystems. They address the access of aquatic biota to seasonal floodplain and riparian habitats as well as the need for high flows to sustain the geomorphological dynamics of the river corridor and floodplain habitats (6). They provide more sophisticated approaches to setting minimum flows and enable advancement of an ecologically acceptable flow regime concept (7,8), which recognizes that a set of minimum flow constraints does not provide sufficient protection for river ecosystems. Different life stages and different species benefit from different flows at different times of the year, and in different years. Rivers must be protected in wet years as well as drought years because high flows provide optimum conditions for some species and are also responsible for sustaining the quality and diversity of in-channel and riparian habitats.

The basic environmental principles needed to formulate policy decisions and management approaches on environmental flows have been summarized by Naiman et al. (9). These focus on the need to sustain *flow variability* that mimics the natural, climatically driven variability of flows at least from season to season and from year to year, if not from day to day. The two fundamental general principles are:

1. The natural flow regime shapes the evolution of aquatic biota and ecological processes.
2. Every river has a characteristic flow regime and an associated biotic community.

From these were developed four specific principles for advancing the provision of environmental flows (10):

1. Flow is a major determinant of physical habitat in rivers, which in turn is a major determinant of biotic composition.
2. Maintenance of the natural patterns of connectivity between habitats along a river and between a river and its riparian zone and floodplain is essential to the viability of populations of many riverine species.

3. Aquatic species have evolved life history strategies primarily in response to the natural flow regime and the habitats that are available at different times of the year and in both wet and dry years.
4. The invasion and success of exotic and introduced species along river corridors are facilitated by regulation of the flow regime, especially with the loss of natural wet–dry cycles.

In addressing the issues of environmental flows, water managers must become water-and-habitat managers, and holistic management strategies are being developed to support their activities (6). However, a continuing failure by policymakers to give due recognition to the array of goods, services, and other benefits provided by aquatic ecosystems (9), to the complexity of ownership and rights of access to these, and of how to integrate livelihood issues into water and ecosystem resource-management decisions remains a major issue.

BIBLIOGRAPHY

1. Poff, N.L. et al. (1997). The natural flow regime. *BioScience* **47**: 769–784.
2. Postel, S.L. and Richter, B. (2003). *Rivers for Life*. Island Press, Washington, DC.
3. Stalnaker, C.B. (1994). Evolution of instream flow habitat modelling. In: *Rivers Handbook*. P. Calow and G.E. Petts (Eds.). Vol. 2, Island Press, Washington, DC, pp. 276–288.
4. Gurnell, A.M. and Petts, G.E. (2002). Changing riverine landscapes. *Freshwater Biol.* **47**: 581–600.
5. Petts, G.E. and Maddock, I. (1994). Flow allocation for in-river needs. In: *Rivers Handbook*. P. Calow and G.E. Petts (Eds.). Vol. 2, Island Press, Washington, DC, pp. 289–307.
6. RRA. (2003). Environmental flows for river systems. *River Res. Appl. Special Issue* **19**(5–6), 375–681.
7. Petts, G.E. (1994). Water allocation to protect river ecosystems. *Regulat. Rivers* **12**: 353–365.
8. Petts, G.E. et al. (1999). Flow management to sustain groundwater-dominated stream ecosystems. *Hydrol. Proc.* **13**: 497–513.
9. Naiman, R.J. et al. (2002). Legitimizing fluvial ecosystems as users of water: an overview. *Environ. Manage.* **30**: 455–467.
10. Bunn, S.E. and Arthington, A.H. (2002). Basic principles and the ecological consequences of altered flow regimes for aquatic biodiversity. *Environ. Manage.* **30**: 492–507.

EUTROPHICATION (EXCESSIVE FERTILIZATION)

G. Fred Lee
Anne Jones-Lee
G. Fred Lee & Associates
El Macero, California

INTRODUCTION

"Eutrophication" is the process of a waterbodys becoming increasingly rich in aquatic plant life such as algae and aquatic macrophytes (water weeds). It is driven by the increasing input of aquatic plant nutrients, principally nitrogen and phosphorus, from natural and anthropogenic sources. Although natural eutrophication takes place over geologic time, activities of people that increase the aquatic plant nutrient inputs to waterbodies can rapidly accelerate this process and cause cultural eutrophication. Thus, the term "eutrophication" has become synonymous with "excessive fertilization" or the input of sufficient amounts of aquatic plant nutrients to cause the growth of excessive amounts of algae and/or aquatic macrophytes in a waterbody such that beneficial uses of the waterbody (i.e., water quality) are impaired. Beneficial uses of waterbodies that stand to be impaired by the presence of excessive amounts of aquatic plant life include domestic and industrial water supply, recreation, and fisheries.

Because of the public health and environmental quality significance of these water quality impairments, myriad strategies have been advanced to evaluate and regulate excessive fertilization and nutrient input to waterbodies, with varying degrees of technical validity and demonstrated effectiveness. This chapter reviews what should be considered in assessing the impacts of nutrients that are added to a waterbody on the waterbody's eutrophication-related water quality. References are provided to more detailed discussions of the issues covered.

IMPACTS OF EXCESSIVE FERTILIZATION ON WATER QUALITY

The excessive fertilization of waterbodies is a long-standing, well-recognized water quality problem throughout the United States and other countries. It is manifested as excessive growths of planktonic (suspended) algae, attached algae, and aquatic macrophytes (water weeds). Aquatic macrophytes can be floating forms such as water hyacinth or duckweed or attached-emergent forms. Water quality problems caused by these growths, discussed in detail by Lee (1), are summarized below.

Domestic Water Supplies

When raw water supplies contain large amounts of algae and some other aquatic plants, the cost of treatment increases and the quality of the product may be diminished. Planktonic algae can shorten filter runs. They can also release organic compounds that cause tastes and odors and, in some instances, serve as trihalomethane (THM) and haloacetic acid (HAA) precursors. THMs are chloroform and chloroform-like compounds; HAAs are low-molecular-weight chlorinated organic acids. These compounds are produced when the precursors react with chlorine during the disinfection process and are regulated as human carcinogens.

Violations of Water Quality Standards

Excessively fertile waterbodies can exhibit marked diel (over a 24-hr day) changes in pH and dissolved oxygen concentrations that can result in repeated short-term violations of water quality standards. During daylight, algal photosynthesis removes CO_2 from the water, which increases the pH; algal respiration in the night releases CO_2 and lowers the pH. In late afternoons, the pH of excessively fertile water can be found to exceed the water quality standard for pH. Similarly, algae produce oxygen during photosynthesis, but they consume it during respiration. Just before sunrise, after sufficient nighttime algal, bacterial, and other organism respiration, dissolved oxygen concentrations can be below water quality standards for protection of fish and other aquatic life. Excessively fertile waterbodies that thermally stratify (develop a thermocline) often exhibit dissolved oxygen depletion below the thermocline because of bacterial respiration of dead algae. Richards (2) showed that one phosphorus atom, when converted to an algal cell that subsequently dies, can consume 276 oxygen atoms as part of the decay process.

Toxic Algae

One major stimuli for the U.S. EPA's recently increased attention to excessive fertilization is the *Pfiesteria* problem in Chesapeake Bay (3); fish kills occurred there because of the presence of toxic algae. Fish kills associated with toxic algae have occurred in various waterbodies around the world, including off the west coast of Florida, for many years. In addition, blue-green algae at times excrete toxins that are known to kill livestock and other animals that consume the water.

Impaired Recreation and Aesthetics

Excessive growths of attached algae and aquatic macrophytes can impair swimming, boating, and fishing by interfering with water contact. Severe odor problems can also be caused by decaying algae, water weeds, and algal scums.

Water clarity—defined by the depth of the waterbody at which the bottom sediments can be seen from the surface—is an aesthetic quality that is compromised by eutrophication. Waterbodies with high degrees of clarity (i.e., the bottom can be seen at depths of 20 or more feet) have low planktonic algal content; in more eutrophic waterbodies, the sediments can only be seen at a depth of a few feet. The greenness of water, which contributes to diminished water clarity and is caused by the presence of algae, can be quantified by measurement of planktonic algal chlorophyll. Inorganic turbidity also diminishes water clarity and can influence the perception of greenness of a waterbody. Often, high levels of planktonic algal chlorophyll can be present in a shallow waterbody or river without the public's perceiving it to be excessively fertile, if the water is brown because of inorganic turbidity.

Impact on Fisheries

As illustrated in Fig. 1, fertilization increases total fish production (biomass). However, as Lee and Jones (4) discussed, it can adversely affect the production of desirable types of fish, especially at high fertilization levels. In stratified waterbodies, algae grow in surface waters, die, and settle to the hypolimnion (bottom layer) where they are decomposed. As noted above, the oxygen demand created by algal decomposition can be sufficient in eutrophic waterbodies to deplete the hypolimnetic oxygen, which means that the desirable coldwater fish (e.g., salmonids, trout) that normally inhabit the cooler hypolimnion cannot survive there because of insufficient oxygen. Thus, the higher fish production characteristic of highly eutrophic waterbodies is typically dominated by rough fish, such as carp, which can tolerate lower dissolved oxygen levels.

Shallow Water Habitat

Emergent aquatic vegetation in shallow waters provides important habitat for various forms of aquatic life. As

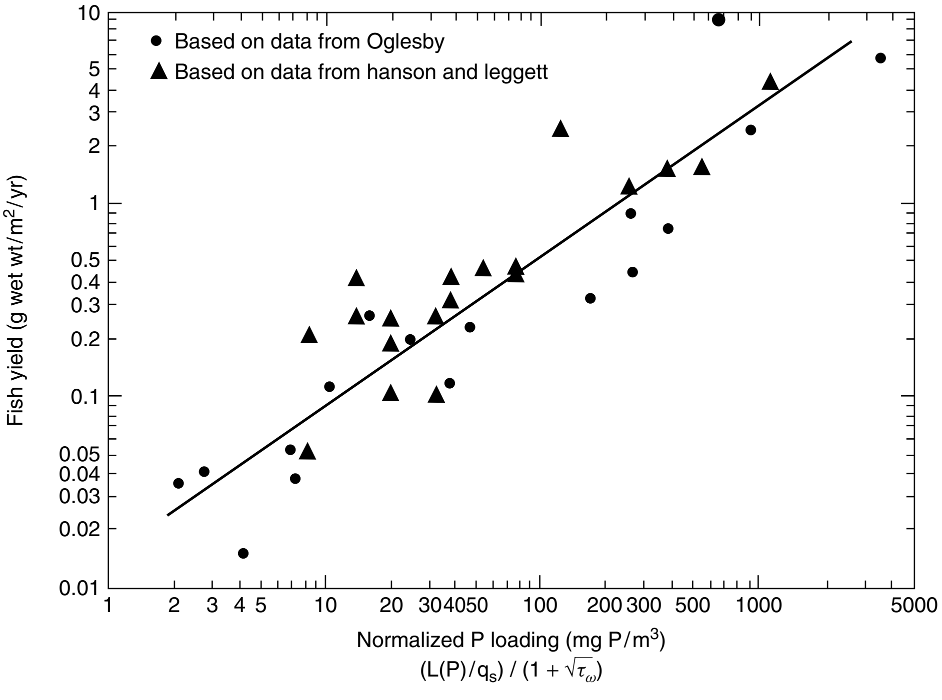

Figure 1. Relationship between normalized P load and fish yield [from Lee and Jones (4)].

discussed by Lee (1), increased planktonic algal growth reduces light penetration (water clarity), which in turn inhibits the growth of emergent vegetation. This process can result in loss of significant aquatic life habitat.

Overall Impacts

Excessive fertilization is one of the most important causes of water quality impairment of waterbodies. In its last National Water Quality Inventory, the U.S. EPA (3) listed nutrients as the leading cause of impairment of lakes and reservoirs.

CONTROLLING EXCESSIVE FERTILIZATION

Algae and other aquatic plants require a wide variety of chemical constituents, light, and appropriate temperatures to grow. Of those factors, however, only nutrient input is amenable to sufficient control to effect a meaningful decrease in algal and aquatic plant biomass to reduce the adverse impacts of excessive fertilization. The issues of which nutrient(s) should be controlled, sources of the nutrient, what type of and how much control is needed, and the positive impacts of the control must be addressed in a eutrophication management program.

Limiting Nutrient

For managing algal populations, the primary focus should be on control of the nutrient that is present in the least amount compared with algal needs, i.e., the limiting nutrient. Increasing or reducing the amount of that nutrient available to algae will affect an increase or a decrease in the algal biomass that can be sustained. This process is illustrated in Fig. 2, which shows that additional growth occurs in response to additional input of the limiting nutrient up to the point at which it is present in greater amounts than can be used. Nitrogen and phosphorus are the nutrients that typically limit algal growth. Phosphorus is more often the limiting nutrient in freshwater waterbodies, whereas nitrogen is often the limiting nutrient in marine waters. Although the potassium content of some soils can limit the growth of terrestrial plants, potassium is not an element that limits aquatic plant growth.

To determine which nutrient is limiting algal growth in a particular waterbody, some have relied on the comparison of the concentrations of nitrogen and phosphorus to the "Redfield" stoichiometric ratio of these elements in algae (16:1 atomic basis or 7.5:1 mass basis) shown in Eq. 1.

$$106\ CO_2 + 16\ N + 1\ P + \text{trace elements} \Rightarrow \text{algae} + O_2 \qquad (1)$$

It is presumed that if the ratio is smaller than this, N would be limiting, and vice-versa, which can give misleading results and lead to unreliable nutrient control measures because whatever the "ratio," either or both could be present in ample amounts for algal growth (5). Rather, it is the concentration of algal-available forms of nutrients at peak biomass—when the algal growth is being limited—that should be assessed. If the concentration of either has been decreased by its utilization to below growth-rate-limiting concentration, reasonable certainty exists that that nutrient is limiting algal growth.

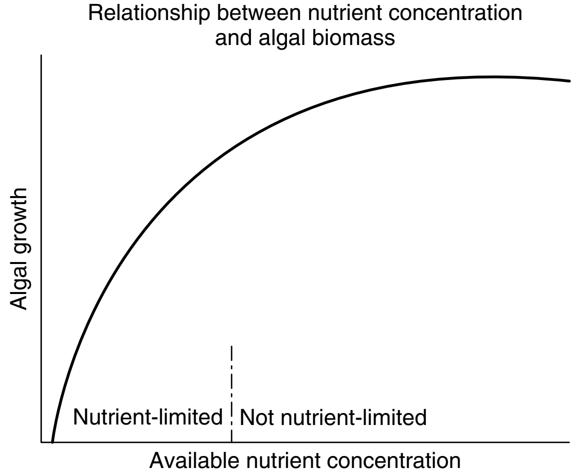

Figure 2. Relationship between nutrient concentration and algal biomass [from Lee and Jones-Lee (4)].

Typically, growth-rate-limiting concentrations for phosphorus are on the order of 2 to 8 μg/L available P, and for nitrogen, 15 to 20 μg/L available N. It is important to recognize, however, that even growth rate-limiting concentrations can support appreciable algal biomass if sufficient time is available for algal growth to occur. Furthermore, in many highly fertile waterbodies, neither nitrogen nor phosphorus is limiting algal growth. Both can be present above growth-rate-limiting concentrations—i.e., on the plateau of the algal growth-nutrient concentration relationship in Fig. 2.

AVAILABILITY OF NUTRIENTS

Nitrogen and phosphorus exist in aquatic systems in many different forms, only some of which can be used by algae and aquatic plants. Therefore, in assessing the limiting nutrient in a waterbody or evaluating the control of nutrient input to a waterbody, it is essential to consider the forms in which the N and P exist in the loading sources and waterbody. Algal available forms of nitrogen are nitrate, nitrite, ammonia, and after conversion to ammonia, some of the organic nitrogen. The fraction of the organic nitrogen that is available is site-specific and depends on its source and age. Under limited circumstances, some blue-green algae can fix (utilize) atmospheric nitrogen gas (N_2) that is dissolved in water and use it as a source of nitrogen for growth. Soluble orthophosphate is the form of phosphorus that is available to support algal growth. Most particulate phosphorus and organophosphorus compounds, and oxygen-phosphorus polymer chain and ring compounds (condensed phosphates), do not support algal growth.

In developing nutrient criteria, the U.S. EPA (6,7) has been focusing on total phosphorus rather than on algal-available forms. This approach can misdirect control programs to sources whose control will not result in cost-effective improvements in eutrophication-related water quality. For example, it was well established many years ago that most of the particulate phosphorus in agricultural and urban stormwater runoff is not available to support algal growth. Lee et al. (8) reported on their extensive research as well as on the findings of others on this topic in a review of these issues for the International Joint Commission for the Great Lakes. From both short-term and long-term (one-year) tests, they found that the algal available P in agricultural and urban runoff can be estimated as the sum of soluble ortho-P and about 20% of the particulate P. Thus, most particulate P in agriculture and urban stormwater runoff from a variety of sources is not available for algal growth.

The lack of availability of much phosphorus in soils is well known to the agricultural community, which finds that total P in soils is not a reliable measure of plant-available P. As discussed by Lee and Jones-Lee (9), nutrient criteria for regulating agricultural and urban stormwater runoff should be based on soluble orthophosphate and nitrate plus ammonia plus about 20% of the particulate P and N. However, if the source of the P and N is algae, then most of the total N and total P will be mineralized and in time will become available to support algal growth.

SOURCES AND CONTROL OF ALGAL NUTRIENTS

Domestic Wastewater Discharges

Lee and Jones (10) reviewed the North American experience in controlling the excessive fertilization of waterbodies. They reported that domestic wastewater discharges are one of the most significant and controllable sources of available nutrients contributing to eutrophication. To control phosphorus from this source, tertiary treatment of the wastewaters is commonly practiced. Chemical treatment using alum (aluminum sulfate) typically costs a few cents per person per day for the population served by the treatment plant. Enhanced biological treatment of domestic wastewaters may also significantly reduce the phosphorus content of domestic wastewaters. Typically, either chemical or enhanced biological treatment can reduce the phosphorus concentration in domestic wastewater effluent by 90% to 95%. The authors estimate that the domestic wastewaters of more than 100 million people in the world are treated for phosphorus removal to reduce the excessive fertilization of the waterbodies receiving the wastewater discharges.

Nitrogen can also be removed from domestic wastewaters although not as readily as phosphorus. Nitrogen removal generally involves nitrification of the ammonia and organic nitrogen to nitrate, followed by denitrification. The cost is typically five to ten times greater than for phosphorus removal. Although phosphorus control in domestic wastewaters is widely practiced, nitrogen control has only been implemented to a limited extent because of the higher cost and because for most freshwater waterbodies, phosphorus control is the more effective way to control excessive fertilization.

Land Runoff

Another source of nutrients for waterbodies is runoff from land. Based on the U.S. Organization for Economic Cooperation and Development (OECD) Eutrophication Study data for about 100 waterbodies' watersheds located across the United States, Rast and Lee (11) determined nutrient export coefficients for the main categories of land use. Shown in Table 1, these coefficients define the mass of N and P that runs off a unit area of watershed land annually.

Although the export coefficients for a given watershed depend on the particular setting, the values in Table 1 have shown reliability in several areas for estimating the potential significance of various types of land use in contributing nitrogen and phosphorus from a watershed. More specific nutrient export coefficients for agricultural lands should be evaluated based on soil characteristics, types of crops grown, and other factors that tend to influence the amount of nitrogen and phosphorus exported from the land. Although these coefficients are for total N and total P, when used in the Vollenweider–OECD eutrophication modeling approach discussed subsequently, the availability of the loading is taken into account.

Nutrient Runoff Control BMPs

Controlling nitrogen and phosphorus in runoff from rural land has not been highly successful. Sharpley (12) reviewed the experience in trying to achieve a 40% reduction in nitrogen and phosphorus loads from agricultural lands in the Chesapeake Bay watershed. He indicated that limited progress has been made toward achieving that goal after about 15 years of effort. Similarly, Logan (13) reported that little progress has been made in effectively controlling phosphorus from agricultural runoff in the Lake Erie watershed.

Sprague et al. (14) reviewed factors that affect nutrient trends in major rivers of the Chesapeake Bay watershed. They noted the difficulty discerning major changes in the contribution of nutrients from agricultural lands in the watershed caused by year-to-year variability in nutrient export. This variability is related to several factors, including climate. They indicated that one of the principal methods for nutrient reduction from agricultural lands has been land retirement, i.e., termination of agricultural activities on the land.

Various "best management practices" (BMPs) have been implemented to control nutrient export from

Table 1. Watershed Nutrient Export Coefficients [from Rast and Lee (11)]

	Export Coefficients ($g/m^2/y$)		
Land Use	Total Phosphorus	Total Nitrogen	
Urban	0.10	0.5	0.25*
Rural/Agriculture	0.05	0.5	0.20*
Forest	0.01	0.3	0.10*
Other:			
Rainfall	0.02	0.8	
Dry Fallout	0.08	1.6	

*Describe nitrogen loadings for waterbodies in Western United States.

agricultural activities, including grassy strips, buffer lands, altering fertilizer applications, and so on. The U.S. EPA (15) discussed the current information on BMPs to control potential pollutants derived from agricultural lands. Although claims are made as to their effectiveness, it is evident from the U.S. EPA review and the authors' experience that there is a lack of quantitative understanding of the cost-effectiveness of BMPs for control of nutrients from agricultural activities (16). Quantitative studies are urgently needed to determine how various BMPs influence phosphorus and nitrogen export from the land, efficacy for controlling eutrophication, as well as costs associated with controlling phosphorus export to various degrees (e.g., 25%, 50%, and 75%). This information then needs to be viewed in the context of what agricultural interests of various types can afford relative to market prices, including issues of foreign competition. Maintaining agriculture through subsidies is a long-standing tradition in the United States. The control of nutrients from agricultural lands for the benefit of downstream waterbody users may also become one of the subsidy issues that will need to be considered to keep agriculture viable (although subsidized) in many parts of the United States.

Importance of Light Penetration

Algal growth in almost all waterbodies is light-limited to some extent. Turbidity and natural color diminish the penetrability of light into a waterbody, which affects the extent to which algae can use available nutrients. In fertile waterbodies, where the presence of abundant planktonic algae reduces the penetration of light further by self-shading, algae can photosynthesize only in the upper few feet of water. It is important to understand the influence of inorganic turbidity and natural color on the coupling between nutrient loads and eutrophication-related water quality. Although erosion from a waterbody's watershed may increase the nutrient load, it also increases the turbidity in the waterbody, which in turn decreases light penetration and thereby slows algal growth. Thus, control of erosion in a waterbody's watershed can result in greater algal growth for the same nutrient concentration than would occur if the waters were still turbid from erosion in the watershed.

Issues That Need to be Considered in Developing Appropriate Nutrient Control Programs

Several key issues need to be considered and evaluated in formulating nutrient control programs, the most important of which is the relationship between nutrient load and eutrophication-related water quality in the waterbody of concern. Each waterbody has its water quality-related load—a response relationship that needs to be defined.

First, the nature of the water quality impairment needs to be defined, which includes defining what the problem is (e.g., recreation impairment, aesthetics, tastes, and odors), when the water quality problems occur (e.g., summer, fall, winter, and spring), how eutrophication is manifested (planktonic algae, attached algae, and macrophytes), and the desired eutrophication-related water quality characteristics. Next, the limiting nutrient during the period of concern and the primary sources of that nutrient should be determined. Each source should be evaluated for the availability of nutrients, the controllability of the available nutrients, and the cost of implementing and maintaining the control strategy. Finally, a reliable modeling approach needs to be applied to estimate the improvement in eutrophication-related water quality that would be effected by the estimated expenditures for the potentially viable control options.

Desired Nutrient-Related Water Quality

The first step in developing appropriate nutrient load criteria is to identify the eutrophication-related water quality problem as well as the desired outcome of management for the waterbody. Types of problem/solution goals that may be identified include, as discussed above, preventing violations of average or worst-case-diel DO or pH standards, controlling algae-caused domestic water supply raw water quality problems (e.g., controlling tastes and odors, lengthening filter runs, reducing THMs, etc.), or increasing water clarity (Secchi depth). This evaluation should be done through a public process conducted by the regulatory agency because the public's perception of eutrophication-related water quality can be site-specific. In those areas where there are numerous waterbodies with marked differences in lake water clarity, for example, the public has the opportunity to compare waterbodies that are green with those that are clearer. There, the public's perception of high water quality is different from that in areas where all waters have the same general greenness because of planktonic algae.

Nutrient control must be undertaken with appropriate consideration of factors that govern how the nutrient loading is used within the specific waterbody. Eutrophication modeling can integrate these factors to relate nutrient load to eutrophication-related water quality response. Basically two types of eutrophication models exist:

- An empirical, statistical model, such as the Vollenweider–OECD eutrophication model discussed subsequently herein, developed from a large database quantifying how nutrient concentrations or loads relate to the nutrient-related water quality characteristics of the waterbody.
- Deterministic models, in which differential equations can describe the primary rate processes that relate nutrient concentrations/loads to algal biomass.

Deterministic models have several drawbacks for use in eutrophication management. Because of the number of equations incorporated into a deterministic model, no unique solution exists to the model. "Tuning" the model to match the nutrient loads and eutrophication condition in the waterbody of interest at the outset may not properly represent the conditions and response after nutrient load alteration. Thus, its ability to reliably meet the goal of management evaluation, i.e., predicting the benefit to be gained by management options, is limited.

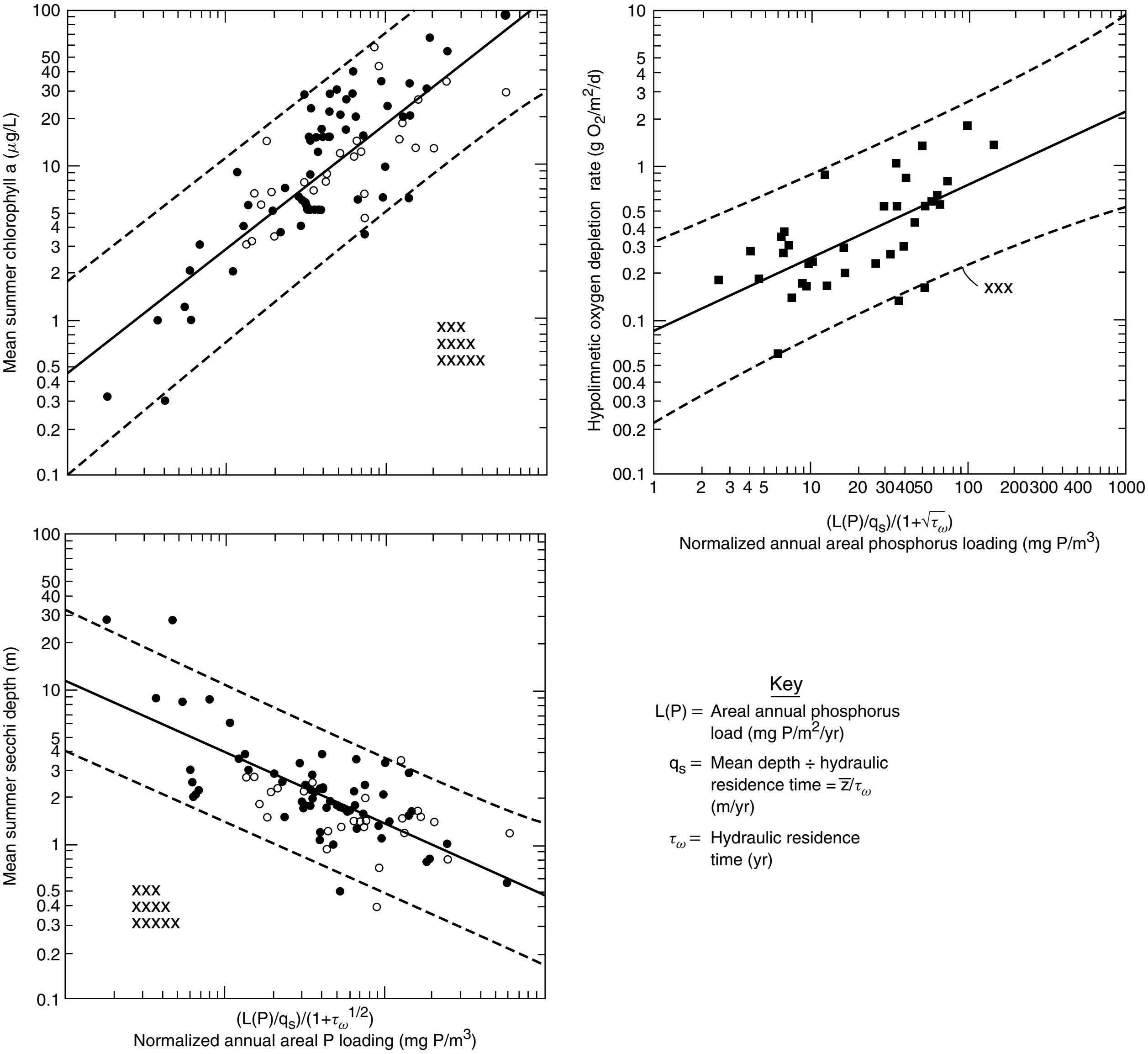

Figure 3. Relationships between normalized P load and eutrophication-related water quality response—U.S. OECD Eutrophication Study Results [after Rast and Lee (17)].

If the water quality problem is related to planktonic algae, the Vollenweider–OECD eutrophication modeling approach is the recommended approach for determining the reduction in nutrient loads/concentrations necessary to achieve the desired nutrient-related water quality in many lakes and reservoirs. Described by Rast and Lee (17) and amplified by Jones and Lee (18,19), this model empirically relates normalized phosphorus loading to eutrophication-related water quality parameters of chlorophyll, water clarity, and hypolimnetic oxygen depletion rate through relationships formulated by Vollenweider (20). These relationships take into account the influence of the key factors of the waterbody's mean depth, hydraulic residence time, and surface area on the utilization of phosphorus by algae within a waterbody. These models, based on the OECD (21) and post-OECD Eutrophication Study data, are shown in Fig. 3. Each point in each figure represents a lake, reservoir, or estuary for which the nutrient load and eutrophication response had been measured for at least a year to generate the model point. Jones and Lee (19) updated this model with data for more than 750 waterbodies in various parts of the world (Fig. 4). The use of this modeling approach and its reliability for predicting the changes in response parameters after a change in nutrient loading has been described by Rast et al. (22).

Rate of Recovery

One of the issues of particular concern in eutrophication management is the rate of recovery of a waterbody after reduction in the nutrient/phosphorus loads. Because large amounts of phosphorus are stored in lake sediments, some have incorrectly concluded that reducing the phosphorus load from the watershed would result in little improvement in water quality, especially in a waterbody with a long hydraulic residence time. However, Sonzogni et al. (23)

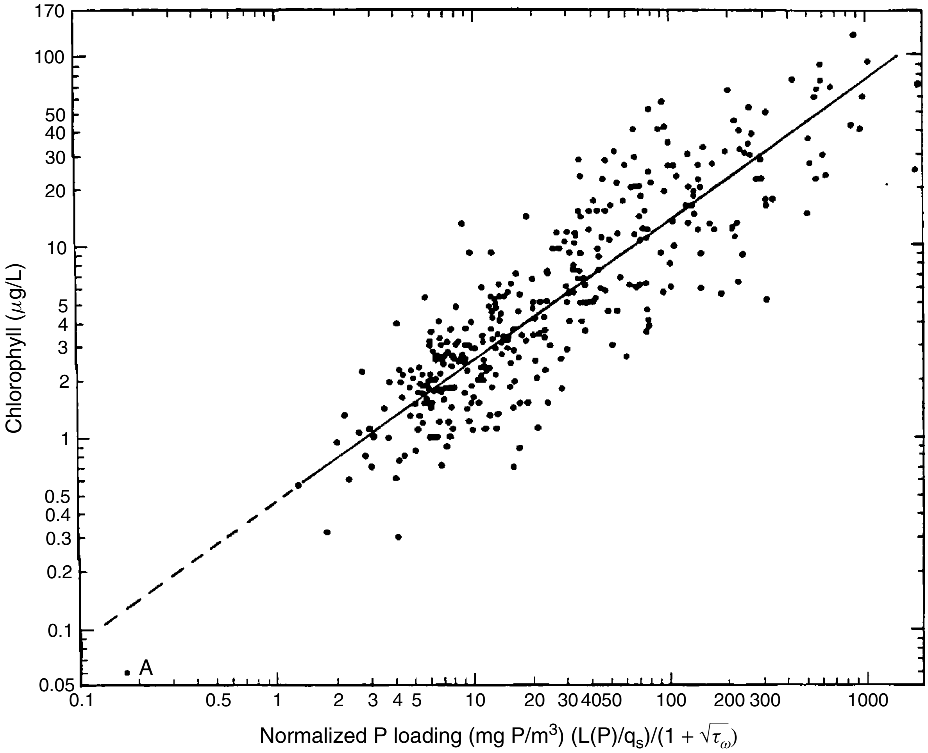

Figure 4. Updated relationship between normalized P load and planktonic algal chlorophyll response [after Jones and Lee (18)].

demonstrated that the rate of response in eutrophication-related water quality to reduction in phosphorus loading is governed by the phosphorus residence time in the waterbody. The P residence time in years is the total mass of phosphorus in the waterbody water column divided by the annual load, which is typically much shorter than the hydraulic residence time.

CONCLUSIONS AND RECOMMENDATIONS

Excessive fertilization, eutrophication, is a major cause of water quality impairment. Domestic wastewaters, urban stormwater runoff, and agricultural runoff/discharges are significant sources of nutrients that contribute to excessive fertilization of some waterbodies. Site-specific investigations are needed to determine the contribution of algal-available nutrients from these sources and the extent to which they can be controlled. Using the Vollenweider–OECD eutrophication modeling approach, the expected improvement in beneficial uses that could be achieved in many lakes or reservoirs by affecting a given load reduction and the expected recovery time can be estimated.

BIBLIOGRAPHY

1. Lee, G.F. (1973). Eutrophication. *Transactions of the Northeast Fish and Wildlife Conference*, pp. 39–60. Available: http://www.gfredlee.com/pexfert2.htm.
2. Richards, F.A. (1965). Anoxic basins and fjords. In: *Chemical Oceanography*. Riley and Skirrow (Eds.). Academic Press, New York.
3. U.S. EPA. (2000). *National Water Quality Inventory*. EPA 841-R-00-001, U.S. Environmental Protection Agency, Office of Water, Washington, DC.
4. Lee, G.F. and Jones, R.A. (1991). Effects of eutrophication on fisheries. *Reviews in Aquatic Sciences* **5**: 287–305. Available: http://www.gfredlee.com/fisheu.html.
5. Lee, G.F. and Jones-Lee, A. (1998). *Determination of Nutrient Limiting Maximum Algal Biomass in Waterbodies*. G. Fred Lee & Associates, El Macero, CA. Available: http://www.gfredlee.com/nut_limit.html.
6. U.S. EPA. (1998). *National Strategy for the Development of Regional Nutrient Criteria*. EPA 822-R-98-002, U.S. Environmental Protection Agency, Office of Water, Washington, DC.
7. U.S. EPA. (1999). *Protocol for Developing Nutrient TMDLs*. EPA 841-B-99-007, U.S. Environmental Protection Agency, Office of Water, Washington, DC.
8. Lee, G.F., Jones, R.A., and Rast, W. (1980). Availability of phosphorus to phytoplankton and its implication for phosphorus management strategies. In: *Phosphorus Management Strategies for Lakes*. Ann Arbor Press, Ann Arbor, MI, pp. 259–308. Available: http://www.members.aol.com/duklee2307/Avail-P.pdf.
9. Lee, G.F. and Jones-Lee, A. (2002). Developing nutrient criteria/TMDLs to manage excessive fertilization of waterbodies. *Proc. Water Environment Federation, TMDL 2002 Conference*, Phoenix, AZ, November. Available: http://www.gfredlee.com/WEFN.Criteria.pdf.
10. Lee, G.F. and Jones, R.A. (1988). The North American Experience in eutrophication control through phosphorus management. *Proc. Int. Conf. Phosphate, Water and Quality of Life*, Paris, France, February. Available: http://www.gfredlee.com/pexfert2.htm.
11. Rast, W. and Lee, G.F. (1983). Nutrient loading estimates for lakes. *J. Env. Eng.* **109**: 502–517.

12. Sharpley, A.N. (Ed.). (2000). *Agricultural and Phosphorus Management—The Chesapeake Bay*. CRC Press, Boca Raton, FL.
13. Logan, T. (2000). Nonpoint sources of pollutants to the Great Lakes: 20 years post PLUARG. In: *Nonpoint Sources of Pollution to the Great Lakes Basin*. Great Lakes Science Advisory Board, International Joint Commission Workshop Proceedings, February.
14. Sprague, L.A., Langland, M.J., Yochum, S.E., Edwards, R.E., Blomquist, J.D., Phillips, S.W., Shenk, G.W., and Preston, S.D. (2000). *Factors Affecting Nutrient Trends in Major Rivers of the Chesapeake Bay Watershed*. U.S. GS Water-Resources Investigations Report 00–4218, U.S. Geological Survey, Richmond, VA.
15. U.S. EPA. (2000). *National Management Measures to Control Nonpoint Source Pollution from Agriculture*. U.S. Environmental Protection Agency, Office of Water, Nonpoint Source Control Branch, Washington, DC.
16. Lee, G.F. and Jones-Lee, A. (2004). Assessing the water quality impacts of phosphorus in runoff from agricultural lands. In: *Environmental Impact of Fertilizer on Soil and Water*. W.L. Hall and W.P. Robarge (Eds.), *American Chemical Society Symposium* Series 872, Oxford University Press, Cary, NC, pp. 207–219. Available: http://www.gfredlee.com/ag_p-1_012002.pdf.
17. Rast, W. and Lee, G.F. (1978). *Summary Analysis of the North American (US Portion) OECD Eutrophication Project: Nutrient Loading-Lake Response Relationships and Trophic State Indices*. EPA 600/3-78-008, U.S. Environmental Protection Agency, Corvallis, OR.
18. Jones, R.A. and Lee, G.F. (1982). Recent advances in assessing the impact of phosphorus loads on eutrophication-related water quality. *J. Water Res.* **16**: 503–515.
19. Jones, R.A. and Lee, G.F. (1986). Eutrophication modeling for water quality management: An update of the Vollenweider-OECD model. *World Health Organization's Water Quality Bulletin* **11**(2): 67–74, 118. Available: http://www.gfredlee.com/voll_oecd.html.
20. Vollenweider, R.A. (1976). Advances in defining critical loading levels for phosphorus in Lake Eutrophication. *Mem. Ist. Ital. Idrobiol.* **33**: 53–83.
21. OECD. (1982). *Eutrophication of Waters, Monitoring, Assessment, and Control*. Organization for Economic Cooperation and Development, Paris.
22. Rast, W., Jones, R.A., and Lee, G.F. (1983). Predictive capability of US OECD phosphorus loading-eutrophication response models. *J. Water Pollut. Control Fed.* **55**: 990–1003. Available: http://www.gfredlee.com/pexfert2.htm.
23. Sonzogni, W.C., Uttormark, P.C., and Lee, G.F. (1976). A phosphorus residence time model: Theory and application. *Water Res.* **10**: 429–435.

CULTURAL EUTROPHICATION

Nitish Priyadarshi
Ranchi University
Ranchi, Jharkhand, India

INTRODUCTION

The phrase "cultural eutrophication" (= cultural enrichment) is becoming widely used to denote organic pollution resulting from human activities. Humans, through their various cultural activities, have greatly accelerated this process in thousands of lakes around the globe. Cultural or anthropogenic "eutrophication" is water pollution caused by excessive plant nutrients. Increased productivity in an aquatic system sometimes can be beneficial. Fish and other desirable species may grow faster, providing a welcome food source (1). Eutrophication produces "blooms" of algae or thick growths of aquatic plants stimulated by elevated phosphorus or nitrogen levels. There has been some uncertainty as to whether algal blooms result from increased concentrations of nitrate or phosphate or from some other cause. It is now commonly accepted that algal growth in fresh waters is generally restricted by phosphate concentrations, whereas in marine waters, it is restricted by nitrate concentration (2). In freshwaters, the nitrate concentration might, however, influence the kinds of algae that grow, some of which taint in drinking water or are toxic to animals.

Bacterial populations increase due to larger amounts of organic matter. The water becomes cloudy or turbid and has unpleasant tastes and odors. Cultural eutrophication can accelerate the "aging" of a waterbody enormously over natural rates. Lakes and reservoirs that normally might exist for hundreds or thousands of years can be filled in a matter of decades.

Cultural eutrophication also occurs in marine ecosystems, especially in near-shore waters and partially enclosed bays or estuaries. Partially enclosed seas, such as the Black, Baltic, and Mediterranean Seas, tend to be in especially critical condition (1). During the tourist season, the coastal population of the Mediterranean, for example, swells to 200 million people. Of the effluents from large cities, 85% go untreated into the sea.

SOURCES OF NUTRIENTS

Humans add excessive amounts of plant nutrients (primarily phosphorus, nitrogen, and carbon) to streams and lakes in various ways. Runoff from agricultural fields, field lots, urban lawns, and golf courses is one source of these nutrients. Untreated or partially treated domestic sewage is another major source. Sewage is a particular source of phosphorus to lakes when detergents contain large amounts of phosphates. The phosphates act as water softeners to improve cleaning action, but they also are powerful stimulants to algal growth when they are washed or flushed into lakes.

Agricultural Runoff

The enrichment material in agricultural runoff is derived from fertilizers applied to crops and from farm animal houses. Nitrogen used as fertilizer may be converted to nitric acid in soil and solubilize calcium, potassium and other ions, which become highly liable to leaching.

Domestic Sewage

Sewage is the most common source of nutrients and organic matter and undoubtedly, the greatest contributor to the eutrophication of lakes and ponds. Large quantities of nitrogen and phosphorus excreted by humans and animals enter into sewage. Phosphatic detergents in sewage (without tertiary treatment) may contain 15 to 35 mg/L of total nitrogen and from 6–12 mg/L of phosphorus (3). Untreated sewage, besides nutrients, also adds large quantities of nitrogenous organic matter.

Industrial Wastes

The nutrients in industrial effluents are variable in quality and quantity, depending on the process and type of industry. The wastes from certain industries, particularly fertilizers, chemicals, and food, are rich in nitrogen and phosphorus. Organically held phosphorus is more soluble, and there is concern that it will leach into surface waters, giving concentrations of 1 mg pL or more, when large amounts of cattle and pig slurries are applied to sandy soil, as in the Netherlands (2).

Urban Runoff

Urban runoff contains storm water drainage with organic and inorganic debris from various paved and grassed surfaces and fertilizers from gardens and lawns.

EFFECTS OF CULTURAL EUTROPHICATION

1. The excessive growth, or "blooms," of algae promoted by phosphates changed the water quality in Lake Erie and many other lakes. These algal blooms led to oxygen depletion and resultant fish kills. Many native fish species disappeared and were replaced by species more resistant to the new conditions. Beaches and shorelines were fouled by masses of rotting, stinking algae.
2. Decomposition of algal bloom leads to oxygen depletion in water. This with a high CO_2 level and poor oxygen supply, aquatic organisms begin to die, and the clean water turns into a stinking drain.
3. Algae and diatoms attain a high degree of dominance due to overfertilization. Algae and rooted weeds interfere with hydroelectric power, clog filters, retard water flow, and affect water quality and water works.
4. Macrophytes, particularly *Hydrilla*, *Potamogeton*, *Ceratophyllum*, and *Myriophyllum*, assume high population densities and make near-shore and shallow regions unsuitable for any purpose.
5. Filamentous green algae, such as *Spirogyra*, *Cladophora*, and *Zygnema*, form a dense floating mat or "blanket" on the surface when the density of the bloom becomes sufficient to reduce the intensity of solar light below the surface. These blankets often give shelter to several undesirable insects, including mosquitoes.
6. Eutrophication of a moderate level may be beneficial to fish production as it increases the food supply for fish in the form of algae. Fish ponds are often fertilized with nutrients to accelerate algal growth and increase fish productivity. But because the level of eutrophication increases due to human activities, the dominance of algal groups is taken over by blue-greens and the edible or game fish are replaced by hardy species of very little economic value.

BIBLIOGRAPHY

1. Cunningham, W.P. and Cunningham, M.A. (2002). *Principles of Environmental Science Inquiry and Applications*. Tata McGraw-Hill, New Delhi, pp. 234–235.
2. Wild, A. (1996). *Soils and the Environment: An Introduction*. Cambridge University Press, New York, p. 255.
3. Hume, N.B. and Gunnerson, C.E. (1962). Characteristics and effects of Hyperion effluent. *J. Water Pollut. Control Fed.* **34**: 15.

FISH CELLS IN THE TOXICOLOGICAL EVALUATION OF ENVIRONMENTAL CONTAMINANTS

LEONARD I. SWEET
Environmental Energy Group,
Engineering Labs Inc.
Canton, Michigan

INTRODUCTION

Aquatic life at risk to modulation by environmental contaminants include fish and their organs, tissues, cells, and subcellular processes. Through branchial, dermal, and oral absorption, as well as biomagnification, it is highly probable that fish are exposed to many chemical species, their metabolites, and their mixtures (e.g., aromatic hydrocarbons, carbamates, heterocyclic compounds, heavy metals, organophosphates, and halogenated compounds). Sources of toxicant exposure include primary anthropogenic emissions, municipal and hazardous waste landfills, incinerators, episodic and diffuse loadings, as well as global secondary sources that involve complex cycling across air–water (e.g., deposition, rain, snow), sediment–water, and biotic interfaces (e.g., vegetation, fish, birds).

Environmental contaminants of concern require research into their accumulation potential, toxicological potency, and heath effects, because they can adversely affect aspects of fish life histories through direct effects (e.g., on developing eggs and larvae) or by more indirect means (e.g., immunosuppression, and enhanced skin and liver disease). Chemical contaminant exposure can interfere with critical phases of the cellular response by destroying, sensitizing, or otherwise altering the functions of cells.

Fish models are increasingly recognized as useful for basic research, biomedical (e.g., carcinogenicity, pathology) and biotechnological applications (e.g., functional and comparative genomics, DNA isolation), and toxicity/environmental safety testing (e.g., lethality, endocrine and androgen modulation, developmental and immunotoxicity). Within these disciplines, fish cells are recognized as an important model because fish are among the oldest and most diverse vertebrates known, are an ecologically critical and pervasive species, are abundant and widely distributed, play a major role in the energetics of trophic levels, have feasible and practical considerations as test organisms (e.g., small size, economical maintenance and use, fecundity, smaller volume of test chemical use and disposal), and often as top predators are at additional risk of bioaccumulating chemical contaminants.

Fish species that are increasingly being used in toxicity testing include the following: rainbow trout, lake trout, fathead minnow, bluegill, zebrafish, guppy, carp, barbels, mummichog, yellow perch, and medaka. Trends in microbiotesting and high volume/throughput testing will lead to further developments in cost-effective microscale aquatic toxicity testing that use fish cells. Fish cells are also becoming recognized as more relevant toxicologically; recent evidence suggests that bloodborne environmental contaminants may be more bioavailable for transport inside the cell, versus transport to circulating binding proteins, than previously suspected. Surrogate tissues such as blood may also serve as an indication of target tissue exposure and cellular stress. The use of fish cells also enables investigations that seek to characterize the mechanisms of toxic effects.

Trends in Toxicity Testing

Prominent trends in toxicity testing include more *in vitro* tests and especially *in vitro* mechanistic assays, aimed at appreciating the importance of the molecular and cellular bases for the effects of chemical toxicants. For example, the assessment and role of cellular apoptosis in immunotoxicological methods is gaining recognition. Characterizing xenobiotic-induced or inhibited apoptosis provides the opportunity to detect subtle and reversible changes in the normal activity of cells and aids in explaining fish stress and time from exposure to toxicity. Other popular end points include heat shock proteins, signal transduction, endocrine modulation, oxidative stress, enzyme induction, influence of inflammatory mediators and endotoxin exposure, and cholinesterase activity. Further, cells harvested from hepatocytes and gills remain prominent choices for toxicity testing. These toxicological end points at biological and biochemical cellular levels attempt to evaluate the functional competence of specific organ systems and individuals.

In vitro response tests are often selected for the following reasons: primary cell exposures provide the necessary dose–response and mechanistic toxicity data; *in vitro* tests require minimal sacrifice of whole fish and minimal use of chemicals; tests are subject to less ethical scrutiny when using cells versus whole animals; and *in vivo* exposure response tests are often impractical because the fish are collected some distance from the laboratory and are only available infrequently and seasonally. The measured end points for assays with fish cells lines often include detection by flow cytometry and fluorescence microscopy. With regard to fish systems, *in vitro* immune cell toxicity studies have focused on the reduction of viability (e.g., assays such as trypan blue, neutral red differential uptake) or reactivity (e.g., phagocytic indexes).

Biomarkers are increasingly employed to evaluate chemically induced alterations in fish at the molecular, biochemical, cellular, or organismal level. These techniques that detect cellular departures from homeostasis may be applied in experimental effect studies or field monitoring and may be destructive or nondestructive to the fish test organisms. However, many of the classical biomarkers of exposure such as enzyme induction and tissue residue fail to identify toxicological hazards and clinical implications adequately that can impact exposed individuals and populations.

In all cases, good laboratory practice and appropriate methods for analgesia, anesthesia, and euthanasia must be employed. Standard operating procedures for tissue harvesting and cell isolation must be designed to minimize handling time prior to preservation or assay testing.

Importance and Policy Relevance

Fish, such as the lake trout in the Great Lakes, can be designated sentinel species of ecosystem health and are useful indicators for assessing chemical contaminant stress. Fish cells in toxicology research can delineate the magnitude of adverse effects of environmental contaminants and one can also extend the relevance of the data to understanding *in situ*, *in vivo*, and ecosystem end points.

With regard to assessing aquatic ecosystem health, toxicological end points at the cellular level can attempt to evaluate the structural integrity and functional competency of specific organ systems and individuals and attempt to address both quantitatively (e.g., internal and biologically effective dose) and qualitatively (e.g., early biological effect, altered structure/function) the more subtle toxic effects of low level contaminant exposure. The fish cell toxicity end point is important because it is rarely possible to predict the effects of urban, industrial, and agricultural pollutants on aquatic biota based solely on the composition and concentrations of contaminants. Thus, employing fish cells in the toxicological evaluation of environmental contaminants is important because it provides a potentially valuable risk assessment component, as well as tools for policy-makers and environmental, epidemiological, and fishery scientists.

Although the fish cells are important models of investigation independently, there is comparative evidence that fish cells and those of higher vertebrates are functionally similar. Further, ambient exposure routes for fish may approximate chronic inhalation tests in mammalian and vertebrate models. However, there are important differences between fish and human cells (e.g., the greater membrane rigidity, volume, surface area, and presence of a large nucleus in fish red blood cells).

In addition, there are potential challenges and obstacles in the interpretation and extrapolation of *in vitro* studies of fish, including (1) relevance of *in vitro* dose levels employed in the laboratory compared to *in vivo* environmental exposures; (2) ability to rule out concurrent exposure to other chemical, biological, or physical stressors; and (3) ascertaining clinical/environmental significance based on statistical significance.

Nonetheless, *in vitro* investigations of this type have advantages over *in vivo* and field studies for the following reasons: control over environmental exposure conditions, potentially reduced variability between experiments, the ability to evaluate thousands of individual cells readily, and the requirement of a small quantity of test chemicals to complete the exposure–response studies.

From an institutional viewpoint, some environmental policies state that contaminant-induced cellular changes are to be considered in decision making because they may directly influence ecologically important parameters, as well as human welfare. The science as applied to fish cells may not be sufficiently developed to allow rigorous hazard assessments, but there remains growing scientific and policy understanding of the subtle yet adverse effects of persistent contaminants on fish physiology.

Bioavailability and Mode/Mechanism of Action

In the aquatic environment, it is important to distinguish between the forms of toxicant exposure especially with regard to bioactivity, biodistribution, and potential toxicity. Speciation and bioavailability are strongly dependent on environmental factors (e.g., pH, redox state, dissolved oxygen, humic content, selenide and sulfide levels, mineral content, mercury content), physicochemical factors (e.g., solubility, partitioning, metal–ligand complexing, ionization), and biological factors (e.g., presence of methylating or demethylating microbes). In teleost species, some of the bioavailability is determined by the pH and chloride concentration of the stomach; at low pH and high chloride concentration, stomach conditions favor the formation of species more readily transported into the blood stream. There are also a number of intrinsic and extrinsic factors that may interact with toxic end-point parameters under investigation (e.g., reproductive status, age, stress, nutrition, toxicant mixtures, previous exposure, predation, food chain effects, habitat, density, environmental stochasticity, concurrent infections, stress, species, and genotype).

Knowledge of chemiometabolic enzyme and receptor systems in fish may prove useful in improving the interpretation of toxicological studies. Fish can clear select chemical toxicants via metabolic routes to a minor degree (e.g., feces, extraction from gill membranes, urine, bile, eggs, and mucus). It should be noted that in fish, as in mammals, the most important enzymes of chemical biotransformation include cyt-P450, UDP-, glucuron-, glutathione-, and sulfotransferases; and they are found most appreciably in the liver, although activity is found in the gills, intestines, and kidney. Interestingly, research evidence supports the presence of the aryl-hydrocarbon (Ah) receptor in teleost fish.

Characterizing the mechanisms of action of chemical toxicants is important for physiological, toxicological, and therapeutic reasons. For instance, once incorporated into biological tissues, the physiological and toxicological effects of metals are regulated by binding to specific ligands and excretion, involving competition for transport and cellular sites and inhibition of enzyme systems. Chemically induced toxicity may involve cell volume changes, alterations in cell permeability, and phospholipid structure. These chemically induced changes may lead to suppressed cell competence, induced or inhibited normal cell proliferation, altered hematology, altered ion balance and metabolism, and enhanced susceptibility to disease states. Certainly mechanistic investigations in ecotoxicology and etiopathogenesis have gained increasing importance especially with regard to prophylactic measures.

Uncertainty Analysis

By definition, uncertainties are inherent in ecological risk assessments and must be addressed in toxicological evaluations using fish cells. The sources of uncertainty include measurement or estimation of variables, natural variability/environmental stochastic over time and space (e.g., rainfall, wind velocity, temperature), and use of models that do not accurately reflect the environment or exposed populations of concern.

During toxicity testing, there is potential for notable individual variations in sensitivity and response. Certainly, it should be noted that overall *in vitro* toxic effects on fish do not necessarily also lead to *in vivo* toxicity. Many *in vitro* systems lack cell-to-cell contacts that may make them behave differently, as well as the influence of regulation by other physiological systems (e.g., nervous or immune). *In vitro* toxicology and mechanistic models are only as good as the level of *in vivo* understanding, and both must be considered in characterizing a chemical hazard. Further, subtle perturbations in cell function following *in vitro* exposure may not in every instance result in a relevant clinical effect, especially given the functional reserve, complexity, and adaptive responses. Other considerations for explaining differences in sensitivity can be attributed to genetics (e.g., expression of metallothionein), metabolism (although Phase I and II metabolism is similar to that of mammals), DNA repair, as well as the varying complexity of teleost fish tissue (e.g., for the immune system, the absence of lymph nodes, tonsils, bone marrow, and the presence of the pronephros).

Other sources of uncertainty are inherent and include experimental design and conduct, variability in chemical composition and purity, variability between sexes, variability across experiments, experimental uncertainty, animal to human extrapolation, higher dose to lower dose extrapolation, difference between commercial and environmental chemical species, persistence and exposure duration, and human variability in exposure and sensitivity. Finally, there is a fundamental assumption in toxicity testing that may lead to uncertainty—that lab tests of single individuals and relatively constant exposures are predictive of field-based intermittent exposures.

SUMMARY

In the end, the significance of studying fish cells in environmental toxicology and hazard assessment research is twofold: *first*, it provides comparative results of a battery of potential *in vitro* assays as an alternative system to *in vivo* toxicity tests and for screening and evaluating chemicals of concern; and *second*, it provides methods for diagnosing and predicting modulation and toxicity from low level exposure to environmental chemical agents that will aid in defining the mechanisms responsible for the observed effects. More investigations into the cellular responses of fish when given equally relevant and equitoxic doses are an important avenue for further study. A battery of available toxicity test methods on fish and fish cells will remain a backbone of environmental safety testing and chemical/biological risk analyses, especially given economic, logistical, and ethical concerns.

FISH CONSUMPTION ADVISORIES

Gary A. Buchanan
Alan Stern
Gloria Post
Division of Science, Research, and Technology
Trenton, New Jersey

INTRODUCTION

Fish consumption advisories can be defined as government health advice issued to the public concerning deleterious substances in fish and/or shellfish (e.g., crabs). These advisories typically provide advice on avoiding or limiting the consumption of certain types or sizes of fish, or on the number of fish meals to consume over a period of time in order to limit health risks. Fish consumption advisories were first issued in the United States in the mid-1970s (1).

HISTORY

Occasional contamination of shellfish and, to a lesser extent, finfish with pathogens and parasites has long been recognized as a potential public health hazard. Prior to fish consumption advisories addressing risks from environmental chemical contaminants, biological (e.g., bacterial) contamination resulted in the closure of shellfish beds and the removal of contaminated harvests from commerce. Such actions tend to be legally enforceable and function through the exercise of public health law. As such, they are not technically advisories, which, by their nature, are provided as guidance and rely on voluntary adherence. Nonetheless, experience with such actions paved the way for governments to consider fish consumption advisories as part of their larger public health responsibility. The recognition of environmental mercury poisoning through fish consumption in the Minimata Bay area of Japan in the 1950s raised the public consciousness about the potential for anthropogenic chemical contamination of dietary fish. However, at that time, and for several decades afterwards, concerns about mercury in fish focused on frank poisoning resulting from highly elevated levels. In addition, the impact of the Minimata contamination on an entire community consuming a common source of fish tended to focus attention on commercial fish. The indirect sources of mercury contamination of otherwise pristine waterbodies and its impact on recreational and subsistence anglers was not fully recognized until the early 1990s. Thus, mercury contamination in fish was treated similarly to biological contamination, and was addressed through regulatory control of permissible mercury levels in commercial fish. By the mid-1990s, recognition that mercury in fish posed health risks more subtle than frank poisoning led to the adoption of consumption advice to recreational and subsistence anglers based on maintaining exposure below a “virtual” safe level and, at the same time, advising fish consumption up to the maximum safe exposure level. The intent of this approach is to maximize the nutritional benefits to be derived from safe levels of fish consumption while avoiding the health risks associated with contaminant exposure. More recently, this approach has been applied to commercial fish, with the federal government (i.e., U.S. Food and Drug Administration, FDA) and various state governments issuing consumption advise for fish whose contaminant levels are nonetheless within regulatory limits for sale.

In general, the first major fish consumption advisories were issued for PCBs (polychlorinated biphenyls). In 1973, the FDA established a tolerance (i.e., a regulatory limit) of 5 parts-per-million (ppm) PCBs for fish in interstate commerce. The FDA subsequently lowered this tolerance to 2 ppm in 1984. Initially, many states used FDA action levels or tolerances in setting fish consumption advisories. However, unlike mercury, PCBs were recognized as carcinogens. Historically, noncarcinogens are considered to have a threshold level of exposure, below which no adverse effects are anticipated, whereas carcinogens have been considered to pose some risk at any level of exposure. Thus, rather than the all-or-nothing regulatory approach initially adopted for mercury, the presumption of some risk at all levels of consumption raised the need for advice that attempted to balance low risk against continued consumption. The FDA’s tolerance limit was not established with such a goal in mind, and thus, starting in the late 1980s and early 1990s, some states began issuing consumption advise to recreational and subsistence anglers that was designed to keep lifetime cancer risk from consuming PCBs in fish within an “acceptable” range. Based on historical concepts of “acceptable” cancer risk, this range generally extended from a risk of one-in-a-million (1×10^{-6}) to one-in-ten-thousand (1×10^{-4}) of developing a PCB-attributable cancer over the course of a 70-year lifetime. This approach was subsequently applied to known or suspected carcinogenic pesticides (e.g., chlordane).

SCOPE

Advisories are issued by local, tribal, state, and national (e.g., U.S. EPA) agencies. Typically, most advisories are issued by the state health department or environmental/natural resource agency. Based on data compiled by the U.S. EPA (2), 48 states had issued a total of 3,089 advisories in 2003, including species-specific, water body-specific, and statewide advisories. These advisories covered 35% of the nation’s total lake acres and 24% of the nation’s total river miles, and included coastal waters. Together, mercury and PCBs account for the great majority of advisories, with chlordane, dioxin, and DDT accounting for nearly all of the remaining advisories (2). A listing of all state advisories can be found on the U.S. EPA website (http://www.epa.gov/ost/fish/). In 2004, the U.S. FDA together with the U.S.EPA (3) issued a national mercury advisory for commercial and recreationally caught fish. This advisory included a "do not eat" advisory for shark, swordfish, king mackerel, or tilefish for the sensitive population (women of childbearing age, pregnant and breastfeeding women, and children). It also advised this sensitive population to check local fish advisories, or if no advice is available, to limit intake to one (6 ounce) meal per week of locally caught fish. Australia and New Zealand have offered similar advice for higher trophic level commercial species (e.g., marlin, shark, swordfish). The European Food Safety Authority (4) has also issued a mercury advisory to this sensitive subpopulation in the European Union concerning consumption of fish. Member countries in the European Union, other countries, and regional governments also issue fish advisories (e.g., British Food Standards, Health Canada, Ontario Ministry of the Environment) (5).

THEORY

Fish consumption advisories seek voluntary reduction of risk through modification of fish consumption behavior. Regulatory restrictions are generally applied when fish enter into commerce or when the risk posed by consumption is clear and immediate. In contrast, advisories tend to be predicated on the notion that government has, at most, limited authority to regulate individual behaviors like fish consumption. At the same time, government operates under a mandate to protect the public health, which is generally seen as a responsibility to advise, inform, and educate. Nonetheless, occasions exist where health risks are imminent or sufficiently elevated as to justify regulations closing specific waterbodies to fishing for some or all species. Often no clear *a priori* dividing line exists between conditions requiring advisory and regulatory actions, and decisions are often made on a case-by-case basis. Advisories are most appropriate when the risk is low to moderate, and are potentially balanced by associated benefits such as nutrition and the continuation of cultural practices. The most basic type of advisory is dichotomous—eat/don't eat. This approach is simple to communicate, but offers little opportunity to consider benefits as well as risks. More sophisticated advisories provide consumption frequency advice (e.g., no restriction, eat once per week, eat once per month... don't eat). Which frequency of consumption is recommended depends on the concentration of the contaminant in the particular species of fish such that eating a typical size serving with the specified frequency will not result in exceeding the maximum acceptable risk level. Thus, a fish for which an eat-once-per week advisory is issued would generally have about 25% of the concentration of the same contaminant as a fish for which an eat-once-per month advisory is issued. Consumption frequency approaches allow consumers to eat the maximum amount of fish that is consistent with remaining within a safe level of exposure or an "acceptable" level of risk. As the contaminant level in any given fish is not known, advisories are derived by extrapolating contaminant data from a limited sample of a given species of fish from a specific water body. Sometimes regional or statewide advisories are constructed by generalizing sample data from a group of representative waterbodies. Various statistical approaches are used for this extrapolation, including averaging concentration and selecting a concentration representing a given percentile of the distribution of sampled fish.

Advisories are generally constructed using common, default assumptions for factors such as body weight and portion size. For contaminants to which one subgroup in the population is more sensitive than the general population (such as methylmercury, which effects fetal neurological development at levels below those causing neurotoxicity in adults), advisories may provide separate advice to the sensitive group and the general population, which includes using different default body weights appropriate to each group

METHODS

Sampling

Fish tissue contaminant data are needed in order to develop fish advisories. Sampling methods vary by region and target species. Collection methods include traps, gill nets, trawls, electrofishing, hook and line fishing, as well as purchase from recreational or commercial fishermen. Fillet samples are typically analyzed, as these are often the target tissue for consumption, but some cultures do cook whole fish. Specialized tissues can also be targeted for certain species (e.g., hepatopancreas or "green gland" for crabs and lobster) in addition to the muscle. Individual fish samples or composite samples (i.e., multiple fish per sample) can be collected depending on needs and budget. Information on monitoring strategy, field procedures, target species, and target analytes are detailed by the U.S. EPA (6).

Analysis

A variety of methods are used to measure contaminants in fish species. Analytical methods and instrumentation have improved dramatically over the past few decades, allowing lower detection limits and compound discrimination (e.g., PCB congeners). For mercury, the cold vapor atomic absorption spectrophotometry is typically used (7). For organic contaminants (e.g., PCBs, pesticides, dioxins/furans, and polycyclic aromatic hydrocarbons), the methods vary, but gas chromatography/electron capture detection (GC/ECD) or GC/mass spectrometry (GC/MS) techniques are normally used (7). Additional information on sample handling, analytical methods, quality assurance, quality control, and data analysis are available (e.g., 6).

Risk Assessment

Risk-based advisories may be based on carcinogenic effects or noncarcinogenic effects. A toxicity factor (slope factor or reference dose, as described below) is based on toxicity data, usually from experimental animals and, less often, from humans. Advisories based on carcinogenic effects generally assume that no threshold exists below which there is no risk, so that any exposure to the contaminant poses some risk of cancer. Therefore, a target lifetime risk level (typically 10^{-4} to 10^{-6}, or 1 in 10,000 to 1 in 1,000,000) is chosen based on policy rather than scientific considerations. A carcinogenic potency factor (also called a slope factor) that relates daily dose to risk is developed from the toxicity data and is used to determine the daily dose of the contaminant (7).

$$\text{Daily Dose (mg/kg body weight/day)} = \frac{\begin{array}{c}\text{Lifetime Risk}\\ \text{(unitless)}\end{array}}{\begin{array}{c}\text{Potency factor}\\ \text{(mg/kg/day)}^{-1}\end{array}} \quad (1)$$

Advisories based on noncarcinogenic effects use the assumption that a threshold exists below which adverse effects other than cancer, such as organ toxicity, developmental effects, or reproductive effects are unlikely. A reference dose (in units of mg/kg body weight/day), below which no adverse effects are expected in the overall population, is developed by applying appropriate uncertainty factors to the dose at which effects occur in animals or humans, which is the daily dose that generally forms the basis for risk-based consumption advisories for noncarcinogenic effects.

Consumption advisories for fish with a given contaminant concentration can be derived as follows:

$$\begin{array}{c}\text{Consumption}\\ \text{Rate (meals/day)}\end{array} = \frac{\begin{array}{c}\text{Daily Dose}\\ \text{(mg/kg/day, see above)} \times\\ \text{Body Weight (kg)}\end{array}}{\begin{array}{c}\text{Contaminant}\\ \text{Concentration (mg/kg)} \times\\ \text{Meal Size (kg/meal)}\end{array}} \quad (2)$$

Assumptions must be made for meal size (typically 8 ounces or 227 grams) and body weight (typically 70 kg for an adult male and 62 kg for a pregnant female). Equation 2 provides the consumption rate in meals/day. If the consumption rate is calculated to be sufficiently large (e.g., 1 meal/day or greater), the advisory can be given as “unlimited consumption.” If the resulting consumption rate is less than the frequency corresponding to “unlimited consumption,” the advisory can be expressed in a convenient unit, such as meals/week, meals/month, meals/3 months, or meals/year. If the consumption rate is less frequent than a reasonable minimum (e.g., less frequent than once per year), the advisory may be given as “Do not eat.” The equations given above are a generalized form of the approach given by the EPA (6), which provides separate equations for carcinogens and noncarcinogens.

Separate fish consumption advisories may be developed for a contaminant to protect the general population and sensitive subpopulations. For example, for methyl mercury and polychlorinated biphenyls (PCBs), neurologic development is the endpoint of concern in the developing fetus, nursing infant, and young child. For the general population, paraesthesia (tingling of the extremities) is the endpoint of concern for methyl mercury and cancer is the endpoint of concern for PCBs.

Fish consumption advisories that do not rely directly on the risk-based approach described above have also been issued. For example, a general advisory has been issued by some states (e.g., Pennsylvania), advising consumption no more than one meal per week of freshwater sport fish. This approach is based on the precautionary principal that can be stated as: in the absence of specific data and given the tendency for bioaccumulative contaminants to be present in freshwater fish, it is prudent to limit consumption.

Another possible approach, for chemicals for which background dietary exposure from sources other than fish results in considerable risk, is to base the advisory on permitting a fractional increase above the background level. For example, background exposure to dioxin, primarily from dietary sources including meat and dairy product, have been estimated to result in a lifetime cancer risk of 1 in 1000 or 10^{-3} (8). This risk level is far above the risks typically used as the basis for risk-based consumption advisories. In such a case, the risk-based advisory approach would limit the consumption of fish, which is a beneficial part of a healthy diet, without actually providing any substantial reduction in risk from dioxin and related compounds.

ADVISORY EXAMPLE

Local or state advisories are typically more specific in terms of fish size, water body, and/or population as compared with national advisories. For example, in New Jersey, mercury advisories for largemouth bass list an advisory of “one meal per month” in Newton Lake for the general population and “do not eat” in Cooper River Lake for children and women of childbearing age (Table 1). Whereas, dioxin (2,3,7,8-tetrachlorodibenzo-*p*-dioxin) advisories in the Passaic River, NJ are “do not eat” for all fish species for all consumers (9).

COMMUNICATION

A very important aspect of fish consumption advisories is outreach to the public. Advisories result in public health protection only if their message is received, understood, and acted on. Outreach activities involve identifying populations at risk, developing an effective communication strategy, and implementation of that strategy using a variety of methods. For example, government agencies have used a variety of techniques, both traditional and unique, to get the advisory message to the public. Examples include publications (e.g., brochures), listing of advisories with fish regulations, posting on web pages, issuing press releases, public service announcements, videos, posting of warning signs at boat ramps and other public access points, direct outreach efforts using local community groups, and toll-free numbers. Risk communication guidance is available from government sources (e.g., 1).

Table 1. Example of Fish Consumption Advisory (8)

		Advisory/Prohibition		
		General Population[a,b] Range of Recommended Meal Frequency		High-Risk Individual[b,c]
		Lifetime Cancer Risk of 1 in 10,000	Lifetime Cancer Risk of 1 in 100,000	Recommended Meal Frequency
Location	Species	Do Not Eat More Than	Do Not Eat More Than	Do Not Eat More Than
Cooper River Lake (Camden Co.)	Largemouth Bass	Four meals per year	Do not eat	Do not eat
	Common Carp			
	Brown Bullhead	One meal per week	One meal per month	One meal per month
	Bluegill Sunfish			
Newton Lake (Camden Co.)	Bluegill Sunfish	One meal per week	One meal per month	One meal per month
	Brown Bullhead			
	Largemouth Bass	One meal per month	Four meals per year	Four meals per year
	Common Carp		One meal per year	Do not eat
Passaic River downstream of Dundee Dam and streams that feed into this section of the river.	All fish and shellfish*	Do not eat		Do not eat

[a] Range of Recommended Meal Frequency corresponds to a cancer risk of 1 in 10,000 to 1 in 100,000 over a lifetime.

[b] Eat only the fillet portions of the fish. Use proper trimming techniques to remove fat, and cooking methods that allow juices to drain from the fish (e.g., baking, broiling, frying, grilling, and steaming). One meal is defined as an eight-ounce serving.

[c] High-risk individuals include infants, children, pregnant women, nursing mothers, and women of childbearing age.

*Selling any of these species from designated waterbodies is prohibited in New Jersey (N.J.A.C. 7:25-18A.4).

BIBLIOGRAPHY

1. U.S. EPA. (1995). *Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories*. Risk Communication. Vol. 4. Office of Water, EPA 823-R-95-001.
2. U.S. EPA. (2004). National Listing of Fish Advisories, Fact Sheet. Office of Water, EPA-823-F-04-016, p. 6. Available: http://www.epa.gov/waterscience/fish/advisories/factsheet.pdf.
3. U.S. FDA and U.S. EPA. (2004). FDA and EPA Announce the Revised Consumer Advisory on Methylmercury in Fish. Press release P04-33, March 19, 2004. Available: http://www.fda.gov/bbs/topics/news/2004/NEW01038.html.
4. European Food Safety Authority. (2004). Press release issued March 18, 2004. Available: http://appl.efsa.eu.int/.
5. Ontario Ministry of the Environment. (2003). Guide to Eating Ontario Sport Fish 2003–2004, 22nd Edn. Queen's Printer for Ontario, Toronto, Ontario, p. 219. Available: http://www.ene.gov.on.ca/envision/guide/index.htm.
6. U.S. EPA. (2000). *Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories*. Fish Sampling and Analysis. Vol. 1. 3rd Edn. Office of Water, EPA 823-B-00-007.
7. U.S. EPA. (2000). Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories. Risk Assessment and Fish Consumption Limits. Vol. 2. 3rd Edn. Office of Water, EPA 823-B-00-008. Available: http://www.epa.gov/ost/fishadvice/volume2/index.html.
8. U.S. EPA. (2000). United States Environmental Protection Agency. Exposure and human health reassessment of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and related compounds. SAB Review Draft, EPA/600/P-00/001 Bg. Available: http://cfpub.epa.gov/ncea/cfm/part3.cfm.
9. New Jersey Department of Environmental Protection and New Jersey Department of Health and Senior Services. (2004). A Guide to Health Advisories for Eating Fish and Crabs Caught in New Jersey Waters. Trenton, NJ, p. 21. Available: http://www.state.nj.us/dep/dsr/njmainfish.htm.

READING LIST

U.S. EPA. (1996). *Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories*. Overview of Risk Management. Vol. 3.

FISHERIES: HISTORY, SCIENCE, AND MANAGEMENT

Robert T. Lackey
United States Environmental
Protection Agency
Corvallis, Oregon

The overall goal of fisheries management is to produce sustainable biological, social, and economic benefits from renewable aquatic resources. Fisheries are classified as *renewable* because the organisms of interest (fish, shellfish, reptiles, amphibians, and marine mammals) usually produce an annual biological surplus that, with judicious management, can be harvested without reducing future productivity. In contrast, *nonrenewable* resources (oil, coal, iron, and copper) are available in fixed quantities and are not replaced except over geologic time.

The benefits that humans gain from a fishery are diverse and may be enumerated in several ways. Most commonly, benefits are computed as *commodity* output—the weight or number of fish produced. Commodity output may be further split between the animals harvested by capture (fishing for wild animals) or culture (produced as captive animals)—commonly called the *capture* fisheries and the *culture* fisheries, respectively.

FISHERIES MANAGEMENT

The benefits are commonly measured as wholesale or retail economic *value* of the commodity output. Such benefits are easily calculated for *commercial* fisheries because the products are usually sold, but for sport or *recreational* fisheries, the quality of the fishing experience is very important, so measures of catch in weight, number, or value only partially measure the benefits provided to fishermen or to society. Measurements of the indirect economic value of recreational fishing that include the quality of the fishing *experience*, however, remain controversial. Even in commercial or subsistence fisheries, substantial benefits may be associated with *cultural* or *religious* aspects. Although such benefits are difficult to measure, they may be very important to the participants.

Beyond the direct benefits derived from harvested fish or the fishing experience, benefits are also derived by individuals and society from simply *knowing* that a particular natural resource exists (often called *existence* value). Society and individuals receive intangible benefits from preserving species and habitats, especially those in danger of extinction. Such benefits are often significant, but, like the benefits from recreational fishing, they are also exceedingly difficult to quantify in economic terms. The whale fishery is an example for which the value of leaving the animals unharvested currently is of greater benefit (primarily intangible) in most societies than the value (economic) of the harvested animals.

Whether or not measurable, fisheries management is increasingly being guided by ecological benefits mandated in treaties, laws, and government policies. For example, the *Convention on Biological Diversity* obligates signatory nations to preserve their biological diversity to the maximum possible extent. Many nations also have laws to protect species at risk of extinction, and these laws may be important constraints on the scope, type, and intensity of fishing that will be permitted.

In practice, the overarching management policy goal for managing a nation's fisheries is often stated in general terms such as:

> To ensure the attainment and continued satisfaction of human needs for present and future generations in an environmentally non-degrading, technically appropriate, economically viable, and socially acceptable manner, and such that land, water, plant, animal, and genetic resources are maintained.

The challenge for the fisheries manager is to translate such a general policy goal into a practical, effective program to maximize the benefits of specific fisheries to society.

FISH VERSUS FISHERIES

The words "fish" and "fisheries" have several meanings, and these terms often cause confusion. As traditionally used in fisheries management, *fish* typically includes the entire suite of *aquatic* organisms that are harvested (mackerel, tilapia, tuna, guppies, sea turtles, seals, whales, sea urchins, clams, squid, and frogs), or *could* be harvested if their numbers permitted. Thus, the term *fish* is not solely the *fin fish* (fish that have fins), so a fisheries manager may work with turtles, squid, or sponges, rather than fin fish. *Shellfish* (clams, crabs, lobsters) are also included under the broad definition of *fish*. In contrast to *fisheries* managers, *wildlife* managers generally deal with terrestrial mammals and birds (deer, wolves, bears, ducks, hawks, and whooping cranes).

A *fishery* is defined generically as a system composed of three interacting components: the aquatic biota, the aquatic habitat, and the human users of these renewable natural resources. Each of these components influences the fishery's performance. Understanding the entire system and its parts is often essential to successful management of a fishery.

There are many different types of fisheries and they may be classified in several ways:

- **Type of environment** (freshwater habitats—lakes, reservoirs, rivers, streams, and ponds; saltwater habitats—estuarine, coastal, and open ocean).
- **Method of harvest** (seining, trolling, trawling, fly casting, spearing, and dip netting).
- **Type of access permitted** (open access to fishing, open access with regulation, limited or purchased access, private property).
- **Organism of concern** (salmon, shrimp, bass, turtles, squid, cod, sharks, sea horses, whales, and swordfish).
- **Purpose of fishing** (commercial fishing for a product to sell, subsistence fishing for direct food consumption, or recreational fishing for sport and leisure).
- **Degree of wildness** of the target animals (totally wild and free-roaming animals, totally captive animals grown in ponds, or animals spawned in captivity, but released in the wild to be captured when they mature).

HISTORY OF THE HUMAN/FISH RELATIONSHIP

Fish have occupied an important place in human society for thousands of years. Early humans obtained fin fish, shellfish, and other aquatic life along the shores of lakes, rivers, and oceans. Archeological records document the use of fish spears 90,000 B.P. (90,000 years before present), nets 40,000 B.P. and fish hooks 35,000 B.P. The earliest documented human communities dependent on fishing were in the vicinity of Lake Mungo (Australia) 30,000 B.P. and Crete 8,000 B.P. The Egyptian aristocracy fished as a leisure activity at least 4,000 B.P. Fish have been raised in captivity for several thousand years.

As better preservation techniques developed (drying, smoking, and salting) and transportation improved, *commercial* fishing in the Middle Ages began shifting from local, small-scale activities to commercial, large-scale enterprises. Boat design and construction advanced, along with corresponding improvements in fishing gear and preservation techniques, especially the advent of canning. Canning represented a particularly important advancement because it permitted long-term storage and large-scale distribution of fishery products.

Cod fishing off eastern North America began in earnest in the early 1500s. By the 1600s, whaling was a prominent activity in many high seas locations. In the late 1800s, steam-powered ships, along with mechanized fishing techniques and refrigeration, enabled development of the large-scale industrial fisheries that still exist today. During the past 100 years, the global level of fishing has expanded continuously, a trend disrupted only briefly during the two world wars. After World War II, the intensity of commercial fishing especially increased.

Today, commercial fishing continues as a major economic sector in many countries. In addition to the large worldwide value of the catch, approximately 36 million people (15 million full-time, 13 million part-time, and 8 million occasional) are employed in the capture and culture fisheries.

Although documented for thousands of years, *recreational* fishing only relatively recently has become an activity enjoyed by large numbers of people. The sport surfaced during the Renaissance as a socially acceptable leisure activity and received broad visibility from the 1653 publication of *The Compleat Angler* by Izaak Walton. By the mid-1800s, recreational fishing was an important and common activity, particularly in North America and Europe. The scale of recreational fishing expanded greatly after World War II, especially in North America.

Millions of people fish recreationally, and they support a multibillion-dollar sport fishing industry worldwide (fishing gear and equipment, bait, boats, motors, outdoor clothing, lodging, and food). Recreational angling ranges in extremes from the serenity of fly-fishing in a remote alpine stream, to the noise and excitement of a bass fishing tournament held in a large reservoir, to the excitement of a child catching its first fish from the bank of a pond.

Beyond the widespread pursuit of wild fish for commercial or recreational purposes, many aquatic species are now successfully raised in aquaculture facilities. Raising fish in captivity (especially various carp species) for food has been practiced in China for at least 4,000 years. The Chinese also developed effective means to breed and raise fish for ornamental purposes, an early precursor to today's vast aquarium market. By the mid-1900s, carp, tilapia, catfish, trout, salmon, and shrimp were widely raised for the *food* market; several of these and other species such as pike, sunfish, bass, and walleye were raised for *stocking* to enhance recreational fishing opportunities.

CURRENT USES OF FISHERIES RESOURCES

More than 4,000 species of aquatic animals are harvested worldwide, totaling approximately 120 million metric tons annually. The total harvest includes *capture* fisheries (fish caught by nets, trawls, hooks) and *culture* fisheries (fish grown in ponds, cages, hatcheries), generally called *aquaculture*. The harvest tonnage from *capture* fisheries quadrupled between 1950 and 1990, but has leveled off or even declined since 1990. Harvest from *aquaculture* continues to increase. Fisheries harvest data are often of questionable accuracy, but the most recently available data on the combined capture and culture harvest suggest that China is the world's leading fish producer (32.5%), followed by Japan (5.1%), India (4.5%), United States (4.4%), and Russia (3.9%).

Aquacultural production continues to increase in importance worldwide and now accounts for approximately one-fifth of the total fish produced. Atlantic salmon, cultured in cages and pens, are raised in Norway, Scotland, Chile, Canada, United States, and elsewhere and provide fresh fish to the retail market year-round. Catfish, grown in ponds in the southern United States, and trout, grown in hatcheries in the northern United States, provide high quality, reliable, year-round products to the national retail market. Carp and tilapia are produced in large numbers, especially in Asia and Africa, and provide the primary source of animal protein for humans in those areas.

Worldwide annual per capita consumption of fish and shellfish is approximately 34 pounds (15 kg), but this varies considerably among regions and individual countries. Per capita consumption of fish has nearly doubled since the 1960s. Europeans and Asians tend to have the highest per capita consumption. Annual per capita consumption in the United States is approximately 15 pounds (7 kg); tuna, shrimp, pollock, salmon, and catfish are the top five.

In North America, the economic vitality of some rural communities depends on catering to the needs of recreational fishermen (fishing equipment, outdoor clothing, boats, motors, trailers, food, and lodging). In 1996, the estimated 35 million American adults (age 16 and older) who fished recreationally spent more than US$37 billion for goods and services related to the sport. The indirect economic impact of this direct expenditure totaled more than US$108 billion.

The market for aquarium and ornamental fish is important worldwide and continues to increase. Fish to supply this market are obtained from both wild and captive stocks. The live fish department is one of the most popular and profitable in many pet stores. Accompanying the sale of live fish is a market for aquaria, aquarium supplies, and, of course, fish food and medicine. Approximately 10 million ornamental *saltwater* fish are imported annually throughout the world. The number of *freshwater* fish imported is much higher and includes more than 5,000 species worldwide. The value of the U.S. retail ornamental fish market (live animals for aquaria and ornamental ponds only) is approximately US$3 billion.

CHARACTERISTICS OF AQUATIC ENVIRONMENTS

Lakes, streams, ponds, oceans, and estuaries are biologically productive, but their productivity has ecological constraints, and these constraints can be reduced by human actions. Many aquatic habitats have been altered by human actions (dredging, filling, damming, road building, pollution, and introduction of nonindigenous species), and their potential for producing sustained fish harvests has been reduced.

Variability is a pervasive characteristic of all ecosystems. Even in the total absence of human activity, aquatic ecosystems exhibit considerable fluctuations in the abundance of individual species. Fish species may vary in abundance several fold between years. In some years, there may be tremendous spawning success. In other years, the same species may have little or no reproduction. Variability in the aquatic environment makes it challenging to assess the likely biological consequences of fisheries management options such as adjusting harvest levels, changing gear regulations, or even placing a moratorium on fishing.

Climatic changes also alter the productive capacity of aquatic environments over the long term. Subtle shifts in ocean currents may cause some fish populations to collapse and others to thrive. Some changes in fish abundance caused by climatic or ocean shifts may happen over centuries and are not apparent without data sets of a century or more. Droughts, so apparent in their ecological effect on the terrestrial landscape, are also important to the aquatic environment, but their effects are usually much less visible. Major regional droughts, for example, are often correlated with increased upwelling of deeper, nutrient-rich ocean waters that stimulate increases in fish production. Even in the absence of all human activity, climate shifts cause the size of salmon runs to vary year-to-year and decade-to-decade.

EVOLUTION OF FISHERIES MANAGEMENT CONCEPTS

The history of fisheries management reflects the conventional wisdom of the day. During the past two centuries, the level of ecological understanding has greatly increased. Prior to the 1800s, most people presumed that biological resources from inland and marine waters were inexhaustible, given the level of harvest possible by the modest number of people who fished and the limited effectiveness of their fishing gear.

By the mid-1800s, the idea of unlimited natural riches from inland waters and the ocean was no longer credible. One popular approach to overcoming nature's constraints on sustainable harvest was to apply *animal husbandry* concepts, including the idea that "seeding nature" with fish produced in captivity would permit much greater levels of fishing. Conventional wisdom held that *aquaculture* could produce a nearly unlimited supply of fish of superior quality and according to a predictable schedule, just as farmers had long achieved for domestic livestock. In captivity, fish could be fed a high quality diet, protected from predation, and the quality of the product improved by selective breeding. If aquaculture performed as hoped, fisheries managers would no longer have to depend on the vagaries and limitations of nature for fisheries products.

For many species in many situations, aquaculture has worked well. Culture techniques greatly improved in the late 1800s. Selective breeding created animals better adapted to life in captivity. However, the expectation that aquaculture would be a solution to nature's limitations was not fully realized.

By the early 1900s, the limitations of aquaculture as a tool to supplement or replace wild fish were being recognized and *harvest regulation* was considered the more effective way to ensure sustained harvests. However, because the scientific underpinnings for many harvest regulations were poor and public pressure to continue heavy fishing was great, regulations were often modest, poorly enforced, and produced disappointing results in limiting harvests to sustainable levels.

As the recognition increased the need for regulations to control overfishing, an appreciation emerged that *habitat* was a limiting factor in fish yields, especially in inland and coastal waters. Thus, efforts to improve stream, lake, and estuarine habitats became more common throughout North America and continue today.

By the mid-1900s, *scientific* fisheries management was the dominant paradigm. The idea underlying this approach was that every fish population had the potential to produce a harvestable surplus and the largest surplus that could be harvested annually from that population (*maximum sustainable yield*) could be estimated by rigorous scientific analysis (stock assessment). The job of the fisheries manager was to control fishing pressure, using various regulations, at a level such that sustainable catch levels could be achieved in perpetuity. However, fishing pressure, as always, was very difficult to control; many fisheries ended up being overharvested, and yields eventually declined.

In the 1970s, the concept of *optimum sustainable yield* became popular, primarily in response to concerns that efforts to maximize the catch in recreational fisheries management overlooked, or at least undervalued, many important benefits that fishermen and society received. Less commonly, it was also used in managing commercial fisheries. Management goals using this approach tended to weigh more heavily the *quality* of the fishing *experience* or the *socioeconomic* aspects of fishing and to place less emphasis on the actual catch. Maximum benefits to society were usually achieved at catch levels *below* the maximum sustainable yield.

By the late 1900s, the trend in recreational fisheries management was toward *species* and *habitat* protection, especially in inland and coastal waters. The widespread recognition that some aquatic species were at risk of extinction led to public pressure to reverse such trends. The main causes of the decline of fish species was habitat alteration and introduction of nonnative fish species. Only rarely was overfishing the primary cause of precipitous declines in fish abundance. In fact, most endangered fish species have never been fished. "Endangered species" and "species at risk" legislation directed government agencies and fisheries managers to emphasize protecting species above catch. Under such legislation, fishing may be permitted only if it does not jeopardize legally protected species.

A recent trend in recreational and commercial fisheries management has been the emergence of the *stakeholder approach*. A stakeholder is any citizen or group potentially affected by, or having a vested interest in, an issue, program, action, or decision. The idea behind involving stakeholders in fisheries management is that society has a wide range of conflicting views on what fisheries management goals should be; therefore, it is desirable to include input from the full range of stakeholders in the process of defining goals and selecting management measures. Fisheries management plans developed with stakeholder involvement, it is assumed, therefore, have a higher likelihood of garnering widespread support.

Comanagement represents a further development of the stakeholder approach where some of the authority to manage the fishery is vested in the fishermen themselves or in organizations such as Indian or tribal governments. There are many current and proposed variants of comanagement, but they all transfer a degree of regulatory authority from government fisheries agencies to fishermen, associations, or other organizations. In some fisheries, a market is created for individual, transferable, fishing rights. In such situations, the right to fish may be purchased in an open market.

A widely accepted development in fisheries management has been the *precautionary principle*. The basic concept is that decision-makers (e.g., fisheries managers) ought to err on the side of caution in managing natural resources. Often the scientific basis for fisheries management decisions contains considerable uncertainty. Given the highly unpredictable future environmental and social conditions, it is wise for managers to use caution in managing fisheries.

LAKE FISHERIES MANAGEMENT

Lakes (including reservoirs formed by man-made dams) vary from large (Great Lakes, Caspian Sea) to small (farm ponds, minuscule alpine pools).

Maintaining at least reasonable good water and habitat quality is absolutely essential to nourishing healthy fish populations. Pollution control and abatement, although typically outside the direct purview of fisheries managers, are essential if management goals are to be achieved.

For large lakes, fisheries management involves primarily assessing, then selecting, fishing or harvest levels that are sustainable. Other management techniques include *intentional* introduction of nonnative species (e.g., Pacific salmon to the Great Lakes), or control of *unintentional* and undesirable introductions (e.g., sea lampreys in the Great Lakes).

There are more management options for small lakes and reservoirs. Small lakes may be manipulated by altering water levels or habitats (e.g., improving spawning areas, adding brush piles to provide hiding places for fish) or altering water quality (e.g., fertilizing low nutrient lakes, reducing the flow of nutrients into lakes that have excessive nutrients) to increase the productivity of desirable species. In some circumstances, improving access to the lake may promote greater use by fishermen. In extreme cases, small lakes may be *chemically rehabilitated* (e.g., all fish are removed by complete poisoning and a desirable mix of fish species reintroduced). Regulation of lake fishing usually involves limits on the *catch* (species, number, and size that may be kept), *time* (season and hours when fishing is permissible), and *gear* (type of fishing equipment and bait that may be used).

Reservoirs often present additional challenges to fisheries managers because they are built for other primary purposes (e.g., flood control, electricity generation, irrigation, water storage, transportation), and these uses often conflict with secondary purposes, such as achieving fisheries benefits for society. Water level fluctuations in many reservoirs, both daily and seasonally, which often result from electricity generation, flood control, and irrigation practices, have profound consequences for reservoir fish populations. In practice, fisheries managers must work collaboratively with many other groups to achieve a mix of societal benefits. Only a few of the benefits from most reservoirs are associated with fish, fishing, and environmental quality.

Management of many small lakes tends to be the responsibility of a single governmental or nongovernmental entity, greatly simplifying fisheries management. Large lakes, on the other hand, tend to have more complex managements, involving multiple agencies. Such inter-jurisdictional decision-making greatly compounds fisheries management problems. For example, because five nations surround the Caspian Sea, managing the sturgeon fishery (which produces highly valued caviar) sustainably has been difficult due to the lack of a single, enforceable management plan. As a result, Caspian Sea sturgeon populations have dropped 90% during the past several decades.

RIVERINE FISHERIES MANAGEMENT

Riverine systems describe a continuum of aquatic systems ranging from small creeks a few miles long to rivers the size of the Amazon, Mississippi, McKenzie, Yukon, and Columbia. One important characteristic is that many riverine systems pass through, or form the boundaries between, different political jurisdictions. Multiple political and management jurisdictions often result in unsatisfactory management of riverine fisheries when agencies fail to cooperate toward achieving common societal goals.

Rivers, especially larger ones, commonly undergo extensive habitat alteration resulting from human activities such as building dams, dikes, bridges, shipping channels, and waste treatment plants. Also, because of dikes and other structures, it is common to lose the connections between rivers and their productive flood plain channels and backwaters, habitats that often provide essential spawning and nursery areas for fish. In many cases, fisheries managers must work with highly altered ecosystems that are no longer suited for the fish species most valued by the public.

Harvest regulations are important in river fisheries, but habitat protection and improvement are especially important. Many rivers, both large and small, are severely polluted or altered by domestic, farm, and industrial waste, agricultural and urban runoff; water withdrawal for domestic, agricultural, and industrial use; siltation; and riparian (streamside) alterations. Poor water quality caused by pollution and lack of habitat diversity are often the limiting factors in developing successful fisheries management programs in rivers. The effects of pollution may be very subtle and indirect. Certain pollutants, for example, may make fish more vulnerable to predation by slightly reducing their ability to sense the presence of a predator.

Unlike lakes, streams and rivers are flowing systems that have a notable self-cleansing ability. Rapid turnover of water can lead to faster recovery of water quality than that in lakes or wetlands, where nutrients and pollutants may remain trapped in sediments for many years. However, the lower, coastal sections of rivers are often slow-moving and are often long-term repositories for contaminants.

Habitat alterations that adversely affect fish are common in riverine systems. The Colorado River of the southwestern United States, for example, is subject to numerous flow diversions and is reduced to a mere trickle by the time it reaches the Gulf of California. The Columbia River and its tributaries, arguably the most regulated river system in the world, contain 250 “large” dams and several thousand “small” ones. Such highly altered habitats no longer favor native fish species, many of which are migratory. Under these altered habitat conditions, certain nonnative fish species may prosper, and many native species are reduced in number or even extirpated.

Small riverine systems (streams, brooks, and creeks) are especially vulnerable to overfishing, habitat destruction, and the effects of land-use practices such as farming and urbanization. In extreme cases, management may include prohibition of fishing or, at least, highly restrictive fishing regulations. At the other extreme, some large riverine systems, especially in their lower reaches, may be managed similarly to large lakes and coastal fisheries. Commercial fishing may be the dominant type of fishing in many large rivers, especially in tropical areas.

COASTAL FISHERIES MANAGEMENT

The marine environment immediately adjacent to land supports *coastal fisheries* and includes nearshore marine, estuarine, and intertidal ecosystems. An *estuary* is a coastal waterbody that has a free connection to the ocean, and alternately with the tides, exhibits characteristics of both fresh- and saltwater environments. Because marine biota move into coastal rivers, the lower, intertidal reaches of rivers are often included as coastal fisheries. Intertidal environments are particularly important as nursery areas for the juveniles of many valuable saltwater fish species.

Coastal fisheries present an array of challenges to the fisheries manager. Human population density tends to be higher along the coasts, and this means that aquatic coastal habitat is likely to be substantially altered (e.g., sea walls, dredging, draining, and buildings) or polluted (e.g., from municipal and industrial waste, ship discharge, and runoff). One of the greatest challenges in managing coastal fisheries is the loss of coastal wetlands. These wetlands provide habitat for many adult fish and shellfish and are also essential breeding and rearing areas.

Coastal fisheries are also often heavily harvested, and there tend to be serious conflicts among user groups. Surf fishermen, crabbers, shrimp trawlers, oil extractors, shippers, boaters, swimmers, and sightseers all use the coastal environment in ways that often conflict with each other. Use can be intensive. There are more than 9 million salt water recreational anglers in the United States.

A major change in coastal fisheries management began in the 1970s as some nations extended their offshore management jurisdiction in an attempt to control fishing by foreign nations. In 1982, most nations adopted the United Nations *Law of the Sea Convention*, which recognized the 200-mile line separating the high seas from waters in the *exclusive economic zone* of the adjacent nation. If desired, individual nations could extend their exclusive economic zone to 200 miles under international law. The United States did so in 1983. In spite of the Law of the Sea Convention, overall fishing pressure has generally remained heavy, especially in exclusive economic zones, because domestic fleets soon replaced foreign fleets.

OPEN OCEAN OR HIGH SEAS FISHERIES MANAGEMENT

Open ocean fisheries are those that operate away from the coasts and often outside of any nation’s territorial waters. The two general categories of fish that are targeted in such fisheries are the *pelagic*, or open-water-dwelling, fish and the *demersa*l, or bottom-dwelling, fish. Pelagic fish species tend to feed and travel near the ocean surface. Demersal fish species tend to live on the continental shelves closer to shore. Tuna and swordfish are examples of commercially important pelagic species; cod, hake, flounder, and toothfish are important demersal species.

The inadvertent capture of nontarget species (called *bycatch*) is a serious management challenge in many fisheries but especially for open ocean and coastal fisheries. Perhaps a third (sometimes much more) of the catch is discarded by fishermen as not marketable. Bycatch may be the young of valued sport or commercial fish species or important food sources for sport or commercial fish species. Shrimp trawlers, for example, catch and discard large quantities of small fish while pursuing the much more valuable shrimp. Various types of fishing gear used by commercial fishermen can injure, and often kill, protected animals such as seabirds, marine mammals, and sea turtles. For example, the indirect catch of dolphins in tuna fishing has led to consumer boycotts and a demand for “dolphin-safe” tuna products. Likewise, devices that effectively exclude turtles from capture have been incorporated into the trawls used by commercial shrimp fishermen.

Habitat alteration caused by certain fishing gear is also a concern in some locations. Trawling (trawls are large, heavy nets dragged by fishing boats) may alter the physical and biological characteristics of the seabed, in particular sea grass beds and coral reefs, in ways detrimental to the well-being of target fish populations. Thus, fisheries managers may have to balance how to minimize the habitat alteration of sea floors caused by trawls while still permitting capture of the target species.

Serial depletion of fish populations is another challenge to managers of open ocean fisheries. Typically, this means that fishermen move to new fishing grounds as those closer to home are depleted. A related type of serial depletion is that caused by the development and use of improved fishing gear that allows fishermen to exploit new fishing grounds as the old ones are depleted. A recent trend has been for open ocean fishermen to move into fishing more deep water environments as more accessible nearshore stocks decline.

Open ocean fisheries management is currently in a state of flux; intense international efforts are being exerted to make them economically and ecologically viable within a framework of producing sustainable, but profitable catches. For example, one challenge for fisheries managers is the heavy subsidies many nations provide to commercial fishermen which creates *excess fishing capacity*. In many cases, fishing would not otherwise be profitable, but subsidies in the form of tax incentives or cash payments make it cost-effective for fishermen to continue fishing. Thus, in many cases, the laws of supply and demand that would tend to prevent overfishing do not come into play.

DIADROMOUS FISHERIES MANAGEMENT

Diadromous fish are those characterized by a life cycle of either spawning in freshwater environments and spending their adulthood in marine environments (*anadromous* species) or spawning in marine environments and spending adulthood in freshwater environments (*catadromous* species). Diadromous fisheries represent unique challenges to fisheries managers because target species often cross multiple jurisdictions. Unless management efforts are well coordinated, the combined effect of decisions by different jurisdictions can result in serious depletion of these resources.

Anadromous fish (e.g., salmon, American shad, striped bass, smelt, and sturgeon) are important species commercially and recreationally. Many rivers no longer support major spawning runs of anadromous fish.

Until the 1800s, large runs of Atlantic salmon were found in many coastal rivers of both western Europe and eastern North America. By the middle to late 1800s, salmon runs in the eastern United States and western Europe had been drastically reduced by the effects of overfishing, dams, and pollution. Overall, runs continue to be much reduced on both sides of the Atlantic. The largest remaining runs, although small by historic standards, occur in eastern Canada, Iceland, Ireland, Scotland, and the northern rivers of Norway. Aquaculture has largely replaced harvested wild fish as the source of Atlantic salmon for the retail market.

The seven species of Pacific salmon found on both sides of the North Pacific also have, overall, declined significantly from historic levels, but not as dramatically as Atlantic salmon. Hatchery production has been used to maintain some runs in the southern region of the range (e.g., Japan, Korea, California, Oregon, and Washington). In California, Oregon, Washington, Idaho, and southern British Columbia, runs have been depleted by past overfishing, dam construction, water withdrawal for irrigation, competition with hatchery-produced salmon, competition with various nonindigenous fish species, predation by marine mammals and birds, and climatic and oceanic shifts. Runs in the northern half of the range (e.g., Russian Far East, Alaska, Yukon, and northern British Columbia) are in much better condition. The northern runs have been abundant for the past several decades, but are likely to decline somewhat for several decades because ocean conditions in the North Pacific tend to shift on such a several-decade time cycle.

Striped bass, native to the East and Gulf coasts of North America, have been introduced to the Pacific Coast and are now found from Baja California to British Columbia. Overall, the species in its original range is less abundant than it was historically, but, nevertheless, catches are still substantial. Some runs on the west coast of North America do very well. The causes of the decline in its original range are similar to those that precipitated the drastic declines in salmon runs (i.e., dams, water diversions, pollution problems, and overfishing).

American shad are found from the Gulf of Mexico up the Atlantic coast and as far north as New Brunswick. Generally, in their native range, shad runs are much reduced from historic times due to dams, pollution, and overfishing. American shad have been introduced to the west coast of North America and have done well in some rivers, especially the Columbia and Sacramento—San Joaquin.

Several species of anadromous sturgeon are also of particular concern. Some are highly prized for their roe (eggs often sold as caviar) and must be carefully managed to avoid overfishing. Other sturgeon species are at risk of extinction, and drastic national and international measures may be required to protect these species.

Many other species have anadromous forms (e.g., smelt, alewife, blueback herring, and cutthroat, rainbow, brown, and brook trout) that support substantial fisheries in certain locations. In other locations, they are important in fisheries management because they are at risk of local extinction. Like salmon, all of these anadromous fishes are extremely vulnerable to dams and other impediments to migration; they are also sensitive to water diversions and pollutants.

A few species, such as American eels, are *catadromous*—they spawn in the ocean, but live their adult lives in freshwater. Eels are important commercial species in certain regions. They occur in rivers, lakes, estuaries, coastal areas, and open ocean. Their distribution ranges from the southern tip of Greenland, along the coast of North America, the Great Lakes, the Gulf of Mexico, the Caribbean, and as far south as northern South America. Most of the catch is exported to Europe and Asia.

AQUACULTURE

Aquaculture continues to expand its importance in both commercial and recreational fisheries management. The growth of aquaculture has been especially rapid during the past decade.

Food production is the most common objective in aquaculture. Commonly raised species are carp, Atlantic salmon, rainbow trout, catfish, tilapia, shrimp, oysters, mussels, and seaweed. Currently, one-third of the total world food fish supply is obtained from aquaculture, and this portion is increasing. The amount of farmed fish produced worldwide has more than doubled since 1989.

Producing fish in captivity for subsequent stocking to enhance, maintain, or initiate fishing is also important, especially in North America. Trout are commonly raised in the United States and Canada to support fisheries subject to intensive recreational fishing, the so-called "put-and-take" fisheries. On the Pacific coasts of North America and Asia, commercial enterprises operate hatcheries to maintain runs of salmon artificially to meet market demand, a practice called *ocean ranching*. After being hatched and reared from eggs, juvenile salmon are released to migrate to the ocean, spend several years growing to adult size, then return to the hatchery of origin to spawn. At the hatchery, they are captured, and some are then spawned artificially to obtain eggs for the next generation. The rest are processed for market.

A fairly recent development in fisheries management, *conservation aquaculture*, uses aquacultural techniques to produce fish threatened by extinction. These "captive breeding" efforts may be the only hope of preserving or recovering certain fish species or populations when natural reproduction is compromised.

SPECIES AND HABITAT PRESERVATION

By the 1990s, management objectives for many freshwater fisheries in North America had shifted from optimizing commodity output to protecting habitat or preserving imperiled species. Concerns about loss of *biological diversity* and *biological heritage* often eclipsed concerns for sustaining commercial or recreational catches. This was especially true for the Pacific coast salmon fisheries and fisheries in more remote, pristine areas such as national parks and wilderness areas.

Overall, alteration of aquatic habitat (e.g., dam construction, flood control structures, dredging to facilitate water transportation, filling to create useable land, sediment and pollution runoff, and acid rain) is one important cause of the tenuous status of many fish species. Thus, protecting fish *habitat* has become a prime focus of many fisheries management agencies.

Nonindigenous fish species (those not native to the area) include both *exotic* species (those from a foreign land), and *nonnative* species (those that have expanded beyond their native range), and they often adversely affect valued native fish populations. Nonindigenous species often compete with or prey upon commercially or recreationally important fish species. They may also hybridize with closely related native species and cause a distortion in the gene pool. Nonindigenous species have contributed to the decline of approximately two-thirds of the threatened or endangered fish in the United States. Many nonindigenous fish introductions have been the result of intentionally releasing bait fish after a day's fishing, unintended releases from international shipping activities, and releases and escapes from the aquaculture and aquarium trades.

Not all nonindigenous fish species are perceived as management problems. Many highly valued and heavily harvested fish species were intentionally introduced by fisheries managers and continue to enjoy widespread public support. Among the most widely introduced fish species in North America are Pacific salmon; rainbow, brown, and brook trout; striped bass; walleye; small- and largemouth bass; and bluegill. Recent trends in fisheries management have been away from introducing fish outside their native range.

Legislation to protect species from extinction also has affected the management priorities of fisheries agencies. For example, the U.S. Marine Mammal Protection Act (1972) and the U.S. Endangered Species Act (1973) are now the legal drivers for management of some fisheries (e.g., over much of the range of Pacific salmon in the United States, the primary management goal is to prevent extinction, not to increase catch).

Internationally, the *Convention on Biological Diversity* (1992) imposes legal obligations on all signatory countries to conserve their biodiversity, manage their fisheries resources in a sustainable manner, and promote fair and equitable distribution of the benefits of each nation's genetic and biological resources.

ECOSYSTEM MANAGEMENT

Beginning in the 1980s, a widespread view emerged that managing fisheries should be broadened in scope to include the entire ecosystem; hence, the rise of *ecosystem management*. A precise, universally accepted definition of ecosystem management has yet to emerge, but it is generally seen as the application of ecological, economic, and social information, options, and constraints to achieve desired social benefits within a defined geographic area and for a specified period.

Part of the appeal of ecosystem management is that it may better balance the suite of benefits (e.g., food fish, recreational fish, preserving endangered species, and preserving ecosystems) that society values. To date, ecosystem management has been most commonly implemented in public forests in North America. Efforts are now under way to apply the same concept to large lakes and open ocean ecosystems.

Adaptive management, the process of improving management effectiveness by learning from the results of carefully designed decisions or experiments, is often included in ecosystem management frameworks. The philosophy underlying adaptive management is the recognition that the ecological consequences of many fisheries management decisions are too uncertain to predict with confidence. Therefore, management decisions ought to be tentative and used to learn how the ecosystem responds. The information derived from such decisions allows the manager to adapt future decisions to reflect what has been learned from past decisions.

Another trend in fisheries management, now commonly practiced in Europe, is using the river basin as the management unit. River basins, or watersheds, have long been used in water management and pollution abatement but have only recently been adopted in fisheries management. River basins tend to be the preferred geographic level of fisheries management that international organizations use.

FUTURE OF FISHERIES MANAGEMENT

In the future, fisheries management will continue to reflect the overall values and preferences of the society within which it operates. Fisheries managers will strive to produce sustainable benefits from renewable biological resources, but society's needs will continue to evolve, resulting in different, and often conflicting, management goals. Efforts to maintain or increase the catch are likely to be tempered by society's growing interest in protecting the environment and preserving imperiled species.

Fish and fishing will remain important factors in the daily lives of many people but especially so in those areas where animal protein from agriculture is in relatively short supply. Overall, harvest pressure on most aquatic environments will continue to increase in concert with rising demand for animal protein for use as human and animal food and for recreation opportunities.

Harvest restrictions in many fisheries are likely to become more constraining as fisheries managers attempt to maintain sustainable yields and avoid fishery collapses. International trade in fisheries products is likely to be scrutinized to a greater degree in response to perceived environmental damage caused by excessive or inappropriate fishing. To counter past overfishing in the ocean, for example, there is likely to be increasing public pressure to create legally protected areas where fishing is forbidden (e.g., fish parks, marine reserves, and marine sanctuaries).

Protecting fish habitat will continue as a primary management goal in the future. There will be a continuing emphasis on protecting the environment in general and water quality in particular. The emphasis in enhancing water quality is likely to shift from controlling specific sources of pollution to reducing pollution from large-scale runoff. Fisheries managers will increasingly be focusing on water quality enhancement and pollution abatement. Introduction of nonindigenous species and genetically altered fish will be of increasing concern to fisheries managers.

In the future, aquaculture will likely provide a larger percentage of the total worldwide production of fisheries products. Domestication, genetic selection, genetic engineering, and other technological developments are likely to be employed to enhance aquacultural production.

READING LIST

Arrignon, J. (1999). *Management of Freshwater Fisheries*. Science Publishers, Inc., Enfield, NH, p. 582.

Avault, J.W. (1996). *Fundamentals of Aquaculture*. AVA Publishing Company, Inc., Baton Rouge, LA, p. 889.

Boon, P.J., Calow, P., and Petts, G.E. (1992). *River Conservation and Management*. John Wiley & Sons, New York, p. 470.

Boyd, C.E. and Tucker, C.S. (1998). *Pond Aquaculture Water Quality Management*. Kluwer Academic Publishers, Boston, MA, p. 700.

Clay, C.H. (1995). *Design of Fishways and Other Fish Facilities*. Lewis Publishers, Ann Arbor, MI, p. 248.

Cowx, I.G. (Ed.). (2000). *Management and Ecology of River Fisheries*. Fishing News Books, Ltd., Osney Mead, Oxford, UK, p. 444.

Gullucci, V.F., Saila, S.B., Gustafson, D.J., and Rothschild, B.J. (Eds.). (1996). *Stock Assessment: Quantitative Methods and Applications for Small-Scale Fisheries*. Lewis Publishers, Inc., Boca Raton, FL, p. 527.

King, M. (1995). *Fisheries Biology: Assessment and Management*. Fishing News Books, Blackwell Sciences, Ltd., Osney Mead, Oxford, UK, p. 341.

Knudsen, E. E., Steward, C.R., MacDonald, D.D., Williams, J.E., and Reiser, D. W. (2000). *Sustainable Fisheries Management: Pacific Salmon*. Lewis Publishers, Boca Raton, FL, p. 724.

Kohler, C.C. and Hubert, W.A. (Eds.). (1999). *Inland Fisheries Management in North America*, 2nd Edn. American Fisheries Society, Bethesda, MD, p. 718.

Lichatowich, J.A. (1999). *Salmon Without Rivers: A History of the Pacific Salmon Crisis*. Island Press, Washington, DC, p. 352.

Murphy, B.R. and Willis, D.W. (Eds.). (1996). *Fisheries Techniques*, 2nd Edn. American Fisheries Society, Bethesda, MD, p. 732.

Quinn, T.J. and Deriso, R.B. (1999). *Quantitative Fish Dynamics*. Oxford University Press, New York, p. 542.

Ross, M.R. (1997). *Fisheries Conservation and Management*. Prentice Hall, Upper Saddle River, NJ, p. 374.

Royce, W.F. (1996). *Introduction to the Practice of Fishery Science*, Revised Edn. Academic Press, San Diego, CA, p. 448.

Scalet, C.G., Flake, L.D., and Willis, D.W. (1996). *Introduction to Wildlife and Fisheries: An Integrated Approach*. W. H. Freeman and Company, New York, p. 512.

Sigler, W.F. and Sigler, J.W. (1990). *Recreational Fisheries: Management, Theory, and Application*. University of Nevada Press, Reno, NV, p. 418.

Templeton, R.G. (1995). *Freshwater Fisheries Management*. Fishing News Books, Blackwell Scientific, Ltd., Osney Mead, Oxford, UK, p. 241.

Welcomme, R.L. (2001). *Inland Fisheries: Ecology and Management*. Fishing News Books, Blackwell Scientific, Ltd., Osney Mead, Oxford, UK, p. 358.

FACTORS AFFECTING FISH GROWTH AND PRODUCTION

Roger C. Viadero
West Virginia University
Morgantown, West Virginia

The health and subsequent growth of fish are directly related to the quality of water in which the fish are raised. In general, factors affecting fish growth and production in freshwater aquatic systems can be classified as physical, chemical/biochemical, or a combination thereof. The physical properties of water that are important to fish production and growth include temperature and the concentrations of suspended and settleable solids; important chemical parameters include pH, alkalinity, hardness, and metals (e.g., iron, aluminum, calcium, etc.) The solubility of oxygen and ammonia gases, which vary as a function of other parameters such as temperature, are also key parameters in maintaining fish health. Recommended water quality criteria for salmonids are

Table 1. Summary of General Recommended Water Quality Criteria for Salmonids[a]

Analytic Parameter	Water Quality Criteria
pH	6.5–8.0 standard units
Total hardness and alkalinity	10–400 mg/L as $CaCO_3$
Calcium	<160 mg/L
Magnesium	<15 mg/L
Iron	<0.1 mg/L
Aluminum	<0.01 mg/L
Manganese	<1.0 mg/L

[a] Reference 1.

presented as typical chemical requirements of freshwater fish in Table 1 (1). However, note that specific water quality requirements are highly species dependent and further references should be consulted.

TEMPERATURE AND DISSOLVED OXYGEN (DO)

Water temperature is one of the most important physical factors affecting fish growth and production. Fish are cold-blooded animals which assume approximately the same temperature as their surroundings. Typically, fish are classified broadly as cold, cool, or warm water, depending on their tolerance for particular temperature ranges. Within each temperature classification, fish survival is bounded by an upper and lower temperature, between which an optimum temperature for growth exists. When temperatures vary outside the optimum range, decreased tolerance to changes in water quality constituents (particularly metabolites such as ammonia) and a decrease in immunological response can occur. Consequently, decreased growth/productivity and in some cases, mortalities, may result, depending on the magnitude of the deviation from the optimum temperature. For example, the survivable temperature range for rainbow trout (*Oncorhynchus mykiss*) is ~1 to 26 °C. However, depending on the reference source, the "optimum" temperature range for growth is 13 °C to 21 °C (1).

Representative optimum temperature ranges for common freshwater fish species are presented in Table 2. The survivable and optimum temperature ranges vary among species, so readers are referred to the comprehensive enumerations of temperature limitations for many freshwater fish species presented by Colt and Tomasso (2), Boyd and Tucker (3), Tomasso (4), Barton (5), Meade (1), and Andrews and Stickney (6).

Water temperature can also affect the solubility of gases which are important to fish health (e.g., dissolved oxygen and ammonia). In general, the solubility of gases decreases when water temperature increases because the chemical species have a greater affinity for the gas phase than the aqueous phase (7). For example, the dissolved oxygen concentration over a range of temperatures is presented in Table 3 (8).

Dissolved oxygen concentrations near saturation (≥95% saturation) are recommended for fish health and productivity (1). Further, the addition of oxygen to fish culture waters via aeration systems can aid in reducing the toxicity of un-ionized ammonia by driving the volatile gas off through mixing/aeration (9). Due to the toxicity of ammonia to fish, nitrogen species are presented and discussed later in a separate section.

Table 2. Representative Optimum Temperature Ranges for Common Freshwater Fish Species[a]

Name	Optimum Temperature Range, °C
Coho salmon (*Oncorhynchus kisutch*)	11–17
Rainbow trout (*Oncorhynchus mykiss*)	13–21
Smallmouth bass (*Micropteur dolomieui*)	18–24
Walleye (*Stizostedion vitreum vitreum*)	20–23
Channel catfish (*Ictalurus punctatus*)	25–30
Talapia (*T. mossambicia*)	28–30
Bluegill (*Leptomic machrochirus*)	29–32

[a] Reference 1.

Table 3. Dissolved Oxygen Concentration Over a Range of Temperatures[a,b]

Water Temperature, °C	Dissolved Oxygen Concentration, mg/L
0	14.62
5	12.80
10	11.33
15	10.15
20	9.17
25	8.38
30	7.63

[a] Reference 8.
[b] Under standard conditions in fresh water (Cl^- concentration = 0 mg/L).

SOLIDS

Solids in water can be in either suspended or settleable form; each can have both direct and secondary effects on the ability of water to support fish. High suspended solids concentrations, measured in the field as turbidity, can irritate fish gills and thus, lead to respiratory deficiencies in fish. Further, suspended solids can decrease water clarity and thus photosynthesis, which can lead to adverse affects on aquatic vegetation and consequently, on herbivorous fish species. High sedimentation in natural waterways can smother fish eggs. Similarly, sedimentation can degrade benthic macroinvertebrate habitats and thus have substantial adverse impacts on insectivorous fish species (10).

Organic matter is also often associated with the solid phase; thus, high suspended and/or settleable solids concentrations can indicate sources of increased oxygen demand in the water which may then lead to oxygen deficiencies for fish. Further, a direct linkage between organic solids loading and bacterial gill disease, as well as amoebae gill infestation, has been reported (11); solid-phase organic matter serves as a host site for opportunistic pathogens.

To control the effects of solids adequately on fish growth and productivity, it is recommended that total solids concentration be maintained at $\leq$80 mg/L. A comprehensive overview of solids removal technologies is presented in Metcalf and Eddy (12) and Wedemeyer (13). Additional representative works related to specific solids management technologies in fish culture systems include Viadero and Noblet (14), Kristiansen and Cripps (15), Summerfelt (16), and Chen et al. (17).

pH

In general, a pH of 6–9 is needed to support aquatic flora and fauna (7), though specific requirements are species-dependent. For instance, rainbow trout (*Oncorhynchus mykiss*) are relatively intolerant to acidic conditions, whereas green sunfish (*Lepomis cyanellus*) and *Rhinichthys atratulus* (black nose dace) are often found in waters recovering from impairment related to acid drainage (10). Other concerns related to fish growth as a function of pH include increases in toxicity of aluminum and ammonia at elevated pH (18,19).

ALKALINITY AND HARDNESS

Carbonate alkalinity is a measure of water's ability to neutralize strong acids due to the presence of carbonate species and is thus an important factor in buffering waters against abrupt changes in pH (7). In natural waters, carbonate alkalinity is derived mainly from the dissociation of $CaCO_3(s)$ and is consequently, directly related to water hardness. Further, elevated water hardness decreases the toxicity of metals such as copper and zinc through the formation of soluble metal–inorganic ligand complexes which are typically not as readily bioavailable to fish as free metal ions (7,18,20–23). A definitive recommended limit for total hardness in fish culture waters is not clear (22), though Heath (24) and Meade (1) noted positive effects on fish health in waters with hardness in excess of $\sim$400 mg/L as $CaCO_3$. In contrast, low alkalinities generally only contributed to fish stress when concentrations of CO_2, a weak acid often found in groundwater, were elevated.

DISSOLVED METALS

The metals of greatest concern for fish growth and production include copper, zinc, tin, cadmium, mercury, chromium, lead, nickel, arsenic, and aluminum (24). The USEPA (25) has listed several of these elements as priority water quality pollutants for human consumption of fish, including mercury, cadmium, and selenium. Heinen(18) synthesized the USEPA "priority elements" along with other metals of concern for effects on fish growth and production, as presented in Table 4.

Table 4. Prioritized List of USEPA "Priority Elements" and other Metals of Concern for Effects on Fish Growth and Production[a]

Analytic Parameter	Water Quality Criteria, Listed in Order of Importance
High	Mercury, cadmium, selenium
Medium	Arsenic, lead
Low	Chromium, copper, zinc, antimony, nickel, silver, beryllium, thallium

[a] Reference 18.

In addition to the priority pollutants listed in Table 4, other metals, including iron and aluminum, can have adverse effects on fish productivity and growth. For instance, both ferrous and ferric iron (Fe^{2+} and Fe^{3+}, respectively) can be toxic to fish. Ferrous iron was reportedly more toxic to fish than ferric iron, though in well-aerated waters, Fe^{3+} is the dominant form of dissolved iron. Ferrous iron, however, may be detrimental to fish health if precipitated in large quantities as $Fe(OH)_2(s)$, as it may "clog" fish gills (18).

Aluminum is toxic to rainbow trout (*Oncorhynchus mykiss*) at lower concentrations than those of other metals because it can accumulate on the gills as aluminum hydroxide and cause respiratory impairment (18). In addition, aluminum is highly toxic to trout in both dissolved and suspended forms; however, the symptoms of poisoning are different, depending on the form/phase of aluminum. Dissolved aluminum tends to cause acute mortalities, and suspended aluminum results in more chronic effects such as respiratory difficulties indicated by gill hyperplasia (26). It has been recommended that concentrations of total aluminum (both dissolved and suspended) be maintained at less than 0.01 mg/L for normal growth and survival of trout. However, the effects of complexation between aluminum and inorganic ligands may abate the effect of high dissolved aluminum concentrations by rendering a fraction of aluminum less bioavailable than free Al^{3+} ions (27).

Criteria for calcium and magnesium concentrations in trout waters are based on limited data (18); common recommendations include calcium concentrations $\leq$160 mg/L (22) and magnesium concentrations $\leq$15 mg/L, though an alternative limit of 28 mg/L Mg has also been suggested by (28).

NITROGEN SPECIES

Total ammonia nitrogen (TAN) is comprised of two species: (1) un-ionized ammonia (NH_3), which is highly toxic to fish and (2) ionized ammonia ($NH_4{}^+$), which is not believed to be toxic to fish (29). The widely accepted TAN limit for "no effect" is 0.0125 mg/L NH_3-N (29–31).

Long-term exposure to subacute un-ionized ammonia concentrations can decrease productivity (29,32) and change gill structure which may result in death from oxygen deficiency (11,33,34). Reduced growth of rainbow trout was observed at a concentration of 0.0166 mg NH_3-N/L. However, fish may be able to maintain sufficient oxygen uptake by changing cardiovascular and respiratory behaviors, though activity would decrease (35,36).

Water chemistry can also impact ammonia toxicity. For example, it has been widely reported that increased salinity and divalent cation concentrations may reduce the toxicity of NH_3-N by altering the osmoregulatory systems of fish and decreasing membrane permeability (29,37,38). Soderberg and Meade (29) studied acute and chronic toxicity to determine the effect of calcium and sodium on the toxicity of un-ionized ammonia to Atlantic salmon (*Salmo salar*) and lake trout (*Salvelinus namaycush*) and found that ionic strength was important in determining the effect of NH_3-N exposure on different species and sizes of fish. Temperature and pH were also significant water quality parameters because they affect the solubility and speciation of ammonia in solution. Methodologies to account for the effects of pH, temperature, and ionic strength are presented in Soderberg (9) and Soderberg and Meade (39).

Nitrite (NO_3^-) is produced by *Nitrosomonas* sp. through the biochemically mediated oxidation of ammonia. Nitrate concentrations in excess of 0.55 mg/L can result in methemoglobinemia, a condition in which the iron in blood cannot transport oxygen. As a consequence, oxygen uptake by fish becomes limited.

BIBLIOGRAPHY

1. Meade, J. (1989). *Aquaculture Management*. Van Nostrand Reinhold, New York.
2. Colt, J. and Tomasso, J. (2001). Hatchery water supply and treatment. In: *Fish Hatchery Management*, 2nd Edn. G. Wedemeyer (Ed.). American Fisheries Society, Bethesda, MD, pp. 91–186.
3. Boyd, C. and Tucker, C. (1998). *Pond Aquaculture Water Quality Management*. Kluwer Academic, Boston, MA.
4. Tomasso, J. (1997). Environmental requirements and noninfectious diseases. In: *Striped Bass and Other Morone Culture*. R. Harrell (Ed.). Elsevier, New York, pp. 253–270.
5. Barton, B. (1996). General biology of salmonids. In: *Principles of Salmonid Culture*. W. Pennell and B. Barton (Eds.). Elsevier, Amsterdam, pp. 29–95.
6. Andrews, J. and Stickney, R. (1972). Interactions of feeding rates and environmental temperature on growth, food conversion, and body composition of channel catfish. *Trans. Am. Fish. Soc.* **101**: 94–99.
7. Stumm, W. and Morgan, J. (1996). *Aquatic Chemistry*, 3rd Edn. John Wiley & Sons, New York.
8. Reynolds, T. and Richards, P. (1996). *Unit Operations and Processes in Environmental Engineering*, 2nd Edn. PWS, Boston, MA.
9. Soderberg, R. (1995). *Flowing Water Fish Culture*. Lewis, Boca Raton, FL.
10. Barbour, M., Gerritsen, J., Snyder, B., and Stribling, J. (1999). *Rapid Bioassessment Protocols for Use in Streams and Rivers; Benthic Macroinvertebrates and Fish*, 2nd Edn. USEPA 841-B-99-002, U.S. Environmental Protection Agency, Office of Water, Washington, DC.
11. Bullock, G. (1972). *Studies on Selected Myxobacteria Pathogenic for Fishes and on Bacterial Gill Disease of Hatchery-Reared Salmonids*. Technical Paper 60, U.S. Bureau of Sport Fisheries and Wildlife, Washington, DC.
12. Metcalf & Eddy. (2003). *Wastewater Treatment, Disposal, and Reuse*, 4th Edn. McGraw-Hill, Boston, MA.
13. Wedemeyer, G. (2001). *Fish Hatchery Management*, 2nd Edn. American Fisheries Society, Bethesda, MD.
14. Viadero, R. and Noblet, J. (2002). Membrane filtration for removal of fine solids from aquaculture process water. *Aquacultural Eng.* **26**(3): 151–169.
15. Kristiansen, R. and Cripps, S. (1996). Treatment of fish farm wastewater using sand filtration. *J. Environ. Qual.* **25**: 545–551.
16. Summerfelt, S. (1996). Engineering design of a water reuse system. In: *Walleye Culture Manual*. S. Summerfelt (Ed.). *NCRAC Culture Series* 101, Ames, IA.
17. Chen, S., Timmons, M., Aneshansley, D., and Bisogni, J. (1993). Suspended solids characteristics from recirculating aquaculture systems and design implications. *Aquaculture* **112**: 143–155.
18. Heinen, J. (1996). *Water Quality Criteria, Uptake, Bioaccumulation, and Public Health Considerations for Chemicals of Possible Concern in West Virginia Mine Waters Used for Culture of Rainbow Trout*. The Conservation Fund's Freshwater Institute, Shepherdstown, WV.
19. Alabaster, J. and Lloyd, R. (1980). *Water Quality Criteria for Freshwater Fish*. Butterworths, London.
20. Klontz, G. (1993). Environmental requirements and environmental diseases of salmonids. In: *Fish Medicine*. M. Stoskopf (Ed.). Saunders, Philadelphia, PA, pp. 333–342.
21. Sorensen, E. (1991). *Metal Poisoning in Fish*. CRC Press, Boca Raton, FL.
22. Piper, R., Mcelwain, I., Orme, L., McCraren, J., Fowler, L., and Leonard, J. (1982). *Fish Hatchery Management*. U.S. Fish and Wildlife Service, Washington, DC.
23. Snoeyink, V. and Jenkins, D. (1980). *Water Chemistry*. John Wiley & Sons, New York.
24. Heath, A. (1987). *Water Pollution and Fish Physiology*. CRC Press, Boca Raton, FL.
25. USEPA. (1993). *Guidance for Assessing Chemical Containment Data for Use in Fish Advisories*. vol. I. Fish Sampling and Analysis, EPA 823-R-93-002, U.S. Environmental Protection Agency, Washington, DC.
26. Freeman, R. and Everhart, W. (1971). Toxicity of aluminum hydroxide complexes in neutral and basic media to rainbow trout. *Trans. Am. Fish. Soc.* **100**: 644–658.
27. Viadero, R. and Tierney, A. (2003). Use Of treated mine water for rainbow trout (*Oncorhynchus mykiss*) culture—a preliminary assessment. *Aquacultural Eng.* **29**: 1–2, 43–56.
28. Wetzel, R. (1993). Microcommunities and microgradients: Linking nutrient regeneration, microbial mutualism, and high sustained aquatic primary production. *Neth. J. Aquatic Ecol.* **27**: 3–9.
29. Soderberg, R. and Meade, J. (1992). Effects of sodium and calcium on acute toxicity of un-ionized ammonia to Atlantic salmon and lake trout. *J. Appl. Aquaculture* **1**: 83–92.
30. Westers, H. and Pratt, K. (1977). Rational design of hatcheries for intensive salmonid culture, based on metabolic characteristics. *Progressive Fish-Culturist* **39**: 157–165.
31. Soderberg, R., Flynn, J., and Schmittou, H. (1983). Effects of ammonia on growth and survival of rainbow trout in intensive static-water culture. *Trans. Am. Fish. Soc.* **112**: 448–451.
32. Smith, C. and Piper, R. (1975). Lesions associated with chronic exposure to ammonia. In: *The Pathology of Fishes*. W.E. Ribelin and G. Migaki (Eds.). University of Wisconsin Press, Madison, WI, pp. 497–514.
33. Flis, J. (1968). Anatomicohistopathological changes induced in carp, *Cyprinus carpio*, by ammonia water. II. Effects of subtoxic concentrations. *Acta Hydrobiologica* **10**: 225–233.

34. Larmoyeux, J. and Piper, R. (1973). Effects of water reuse on rainbow trout to sub-lethal concentrations of ammonia. *Water Res.* **35**: 2–8.
35. Hughes, G. (1972). Morphometrics of fish gills. *Respir. Physiol.* **14**: 1–26.
36. Smart, G. (1976). The effect of ammonia exposure on gill structure of the rainbow trout (*Salmo gairdneri*). *J. Fish Biol.* **8**: 471–475.
37. Potts, W. and Fleming, W. (1970). The effects of prolactin and divalent ions on the permeability to water of *Fundulus kansae*. *J. Exp. Biol.* **53**: 317–327.
38. Lloyd, R. and Orr, L. (1969). The diuretic response by rainbow trout to sub-lethal concentrations of ammonia. *Water Res.* **3**: 335–344.
39. Soderberg, R. and Meade, J. (1991). The effects of ionic strength on un-ionized ammonia concentration (technical note). *Progressive Fish-Culturist* **53**: 118–120.

READING LIST

Bullock, G., Herman, R., Heinen, J., Nobel, A., Weber, A., and Hankins, J. (1994). Observations on the occurrence of bacterial gill disease and amoebae gill infestation in rainbow trout culture in a water recirculation system. *J. Aquatic Anim. Health* **6**: 310–317.

WATER NEEDS FOR FRESHWATER FISHERIES MANAGEMENT

Harold L. Schramm, Jr
U.S. Geological Survey
Mississippi State, Mississippi

THE SCOPE OF FISHERIES MANAGEMENT

"Fishery management" is variously defined, but common to all definitions is manipulating fishery resources to produce sustainable benefits for people. Fishery management relies on principles of aquatic ecology but focuses on the use of fishery resources.

A "fishery" has three components: the fish, the habitat, and the users (people). Effective fishery management embraces all three components. It is important to recognize that "the users" often are people who catch fish for sale or recreation, but the users are not necessarily consumptive. The thrust of this article, the dependency of fishery management on water, necessitates a focus on the habitat component of fisheries. However, no component is completely independent of the others, and the linkage between people and habitat is especially influential on fishery management.

WATER QUALITY

Good water quality is essential to fish survival, growth, and reproduction. Most fish require dissolved oxygen concentrations above 5 mg/L (parts per million), and 3 mg/L is considered stressful for many warm-water fish. Brief periods of hypoxia can be lethal. Fish survive over a pH range of 4.5–10, and acidic inputs from mine drainage or precipitation can be lethal. Freshwater fish vary in their tolerance of salinity; some species can survive in seawater, but others can tolerate less than 10% seawater. Many heavy metals and pesticides and their derivatives are toxic to fish, but some chemicals, such as mercury and PCBs, are accumulated by fish to levels toxic to other animals, particularly mammals and birds. Ammonia and nitrite can be toxic at less than 2 mg/L. Some evidence is now available that hormones in discharge water can disrupt endocrine function and interfere with fish reproduction. Tolerances of fish to water quality variables may vary among life stages; for example, eggs or larvae may tolerate much lower salinity or concentrations of a toxic chemical than adults.

Point-source pollution, although not abated, is relatively well controlled in most developed countries compared with nonpoint pollution—in essence, runoff from the watershed. Inputs of otherwise nontoxic and biologically essential chemicals, such as nitrates and phosphates, can have beneficial or benign effects on aquatic systems. However, excessive levels of these chemicals that result from agriculture, confined animal rearing facilities, and domestic and municipal development have dramatic effects on receiving waters and cause undesirable shifts in aquatic ecosystem structure and function. In this regard, it is important to recognize that many chemical inputs to aquatic systems are not directly injurious to fish but, by altering the composition and abundance of other biota, indirectly influence the health and survival of fish populations. A common example is oxygen depletion that results from planktonic algae blooms or excessive growth of aquatic plants (macrophytes). Although the plants usually produce the oxygen needed to sustain fish and other aquatic life, dense macrophytes or algae blooms fueled by high phosphorus concentrations can consume oxygen at night and during reduced-light conditions and cause fish kills. Oxygen depletion can also occur at the end of the growing season when excessively abundant plant material dies and decomposes.

WATER QUANTITY

It is obvious that fish need water, but there are aspects about water quantity that are far from obvious and are often overlooked by managers. In standing water systems, the amount of water impacts fish by affecting the types and amounts of habitats available and may alter thermal conditions. Declines in water level generally result in reduced areas of desirable aquatic habitat, and reduced water volume, in turn, impacts physicochemical conditions. Most freshwater fish spawn in relatively shallow water, and drawdowns at that time can cause spawning failures. The young of many fish find both shelter and food in shallow water, littoral zones (the nearshore area where rooted aquatic plants grow); dewatering these habitats reduces fish growth and survival and may result in year class failures. However, seasonal declines in water level can be desirable in shallow lakes where periodic drying of littoral zones is necessary to oxidize accumulated organic matter and consolidate soft, detritus-rich bottom sediments.

Rivers and streams are dynamic systems in terms of amount (discharge) and level (stage) of water, but the fish endemic to these systems are adapted to the annual cycle of fluctuations. Although peak and low flows and stages may vary from year to year, the annual cycle of rise and fall (the hydrograph) is rather predictable. Alteration of the annual hydrograph, such as may occur when a river is dammed for flood control or water supply, can interfere with spawning and recruitment by decoupling river elevation and thermal cycles, preventing access to essential habitats for floodplain-spawning fishes, or failing to provide necessary current velocity for fish that require flowing water for successful reproduction. River fluctuations are also important to adult stages of many fishes. Seasonal inundation of the floodplain provides access to rich food supplies for many fish, and receding waters flush nutrients, organic material, and biota, including small fish, into the river. As such, this “flood pulse” strongly influences fish production in river–floodplain ecosystems. High river stages also connect the river to floodplain lakes that were former river channels. Connection with the river can supply nutrients to the lake and allow in and out movements of fish. This connectivity, even if brief, unites the numerous different habitats in rivers necessary to sustain the diverse fish assemblage.

ALTERATION OF AQUATIC HABITATS

Water is essential to fish, but it is also essential to society and economic development. Natural river flows are altered for flood control, navigation, hydropower generation, water supply, and other anthropocentric purposes. Most of these purposes are accomplished by constructing dams to impound water and regulate flows. Dams are barriers to fish migrating to upstream spawning habitats and often curtail or even eliminate spawning. Further, the lake conditions created by dams eliminate habitats and flowing water essential to part or all of the life cycle of many native riverine fish. Reservoirs, by serving as sediment and nutrient traps and biological filters, can reduce the nutrients and sediments available to downstream systems. The resulting oligotrophication results in reduced fish production in downstream waters, a rather ironic reversal of escalating nutrient accumulation (eutrophication) common to many lakes and reservoirs.

Sedimentation is probably the greatest threat to river and stream fisheries resources. Although sedimentation is a natural process, excessive sediment input into waterways results from poor land management on the watershed. Sediment covers hard and rough bottoms essential for spawning and juvenile or adult stages of many fishes. The sediment smothers bottom-dwelling invertebrates; many are important food sources of fish; others are important to energy cycling in the stream. Sediment also provides an unstable substrate; thus, aquatic vegetation and bottom-dwelling animals that live in and on these soft substrates are often washed away during floods.

Natural thermal cycles are important to fish production, and “thermal enrichment” can be a form of pollution. Thermal enrichment, usually resulting from discharge of cooling water from electric generating plants or industry, can benefit fish production by lengthening the growing season and may benefit users by lengthening the fishing season. Unfortunately, adverse impacts usually outweigh benefits. Elevated temperatures can stress kill fish, alter their metabolic requirements, alter predator–prey dynamics, cause out-of-season spawning, and result in replacement of cool- and cold-water fish with species tolerant of warmer water. Water temperatures lower than natural thermal conditions can also be problematic. Coldwater, hypolimnetic discharges from reservoirs can reduce aquatic production downstream, impact fish growth and recruitment, and alter the fish assemblage in favor of cold-tolerant species.

BIOTIC POLLUTION

Introduced aquatic species are an expanding and insidious form of pollution. Introduced species arrive in waterways by various means, including intentional but ill-advised introductions. The greatest source of aquatic nuisance species is intercontinental shipping. More than 160 nonnative species have been documented in the Great Lakes. Most, if not all, arrived in the ballast water of large ships. These alien species can alter water quality, change the pathways of energy and nutrient flow, alter habitats, prey on or compete with native species, hybridize with native species, and introduce diseases and parasites. Rarely are the effects beneficial; and most introduced species, once established, cannot be eliminated.

COORDINATED MANAGEMENT

Most public-access waters are multiple-use resources. As such, fishing is just one use of a waterbody, which is especially true for the more than 10 million acres of large impoundments in the United States. These systems were built to accomplish one or more purposes—hydropower, flood control, water supply, and navigation—invariably directed at achieving social or economic benefits. Fishing, although typically included in benefit analysis, was not a primary purpose for constructing large reservoirs. Fulfilling the primary purposes for which the reservoir was built dominates reservoir operation, often diminishing fishery values. In the last two decades, resource managers have become more sensitive to fishery needs.

SHIFTING TO MORE HOLISTIC MANAGEMENT

The study of inland waters (limnology) is only a little more than a century old, and modern fishery management has been developing for only half that long. As knowledge grows, perspectives and paradigms shift. Fishery managers now recognize the value of managing fish assemblages rather than single species. Furthermore, managers are realizing that lakes and rivers are not isolated microcosms but rather part of larger watersheds or ecosystems, and what happens on the watershed or in a land-water ecosystem affects what happens in the water. Future management of fishery resources requires coordination and cooperation among diverse interests.

Acknowledgments
Helpful reviews of earlier versions of this manuscript were provided by Wayne A. Hubert, Robert T. Lackey, L. Esteban Miranda, and David W. Willis.

READING LIST

Dodds, W.K. (2002). *Freshwater Ecology: Concepts and Environmental Applications*. Academic Press, San Diego, CA.

Hall, G.E. and Van Den Avyle, M.J. (Eds.). (1986). *Reservoir Fisheries Management: Strategies for the 80s*. Reservoir Committee, Southern Division. American Fisheries Society, Bethesda, MD.

Kohler, C.C. and Hubert, W.A. (Eds.). (1999). *Inland Fisheries Management in North America*, 2nd Edn. American Fisheries Society, Bethesda, MD.

Lackey, R.T. and Nielsen, L.A. (Eds.). (1980). *Fisheries Management*. Halsted Press, John Wiley & Sons, New York.

Miranda, L.E. and DeVries, D.R. (Eds.). (1996). *Multidimensional Approaches to Reservoir Fisheries Management. American Fisheries Society Symposium* 16. American Fisheries Society, Bethesda, MD.

Murphy, B.R. and Willis, D.W. (Eds.). (1996). *Fisheries Techniques*, 2nd Edn. American Fisheries Society, Bethesda, MD.

Welcomme, R.L. (2001). *Inland Fisheries: Ecology and Management*. Fishing News Books, Blackwell Scientific, Oxford, UK.

AN OUTLINE OF THE HISTORY OF FISHPOND CULTURE IN SILESIA, THE WESTERN PART OF POLAND

Maria A. Szumiec
Polish Academy of Sciences
Chybie, Poland

The need for pond construction in Silesia arose from the lack of large natural waterbodies and the long distance to the sea. The abundance of water, benign climate, fertile soils, the growing demand for fish in the developing towns; and the possibility of transporting fish along the rivers favored the development of fish farming.

The Silesian territory (Fig. 1) situated on the junction of Slavonic and German nations was affected by numerous wars, but at the same time, it was a center where various cultures existed (2). The political position of Silesia changed several times. From 990 to 1327–1329, it belonged to Poland. The foundation of the Wratislawia bishoprics by Bolesław the Brave, the Polish king of the Piast dynasty, in 1000 was an important moment in the history of fishpond culture because monasteries were the first to play an important role in fish culture in the Middle Ages. In the twelfth century, division among the sons of a Polish prince from the royal dynasty of Piast gave rise to many Silesian principalities.

In 1327–1329, the different political orientation of the Silesian princes caused it to affiliate with the Czech Kingdom (Fig. 2). From the fifteenth century the southern part (Teschenensis and Oppoliensis-Ratiboriensis) was called Upper Silesia and the northern (Wratislawiensis), Lower Silesia. In 1526, Silesia became a part of the Habsburg Empire; in 1763 the entire Silesian territory was incorporated into Prussia; in 1918, it was partly affiliated with Poland, and in 1945, Silesia became part of Poland. Political, social, economic, and climatic variations distinctly affected the history of Silesian fishponds, which for ages played a manifold role in alleviating freshets and floods of rivers during wet periods and served as water resources during dry years.

This outline is based on secondary sources, chiefly the Polish literature including Czech, German, and Moravian citations. The historical data are derived from Encyclopedia PWN (1983–1987) (3).

THE FIRST PERIOD

The first development and breakdown of fishpond culture occurred between the twelfth and middle of the fifteenth century. Probably, Cistercians began to breed fish in ponds in the eleventh or twelfth century. Fishpond farming developed after the rivers and natural waterbodies had been wastefully exploited owing to the increasing population and strict observance of fasting. The highest progress in fish farming occurred in the fourteenth century and resulted from a rapid economic progress accompanying the growing mining industry and developing trade. Numerous fishponds as well as the introduction of the carp, probably from Czech and Moravia, allowed people to meet the religious demands. Devastating and depopulating wars, chiefly the incursion of the Tatars, obstructed fishpond development. The Hussite and Polish-Czech-Hungarian wars for the Czech throne ended the first period of development of fishpond culture in Silesia. Unfavorable climatic events were also responsible for this situation (4–6), although the outstanding development of fishpond culture might point to the fourteenth century as the warm climate phase.

THE SECOND PERIOD

In the middle of the fifteenth century, pond culture revived and developed till the end of the sixteenth century although some authors point to the sixteenth century as the beginning of the Little Ice Age (7). It resulted from a fairly peaceful time, warmed climate, popularization of carp culture and also due to privileges bestowed, mostly upon the oligarchy, by Ladislau of the Polish royal dynasty of Jagiellon, the king of Czech and Hungary.

Numerous ponds belonged to principalities and nobility, also to burghers, monasteries, and peasants. At the time of the highest development of pond fish culture, some places in the Silesian territory had the character of a lake district. Imagining the landscape at that time gives a map of a part of Upper Silesia elaborated in 1725 after a significant reduction in pond area in this territory (Fig. 3).

The fish from Upper Silesian ponds were transported by the River Vistula to Cracovia the capital of Poland

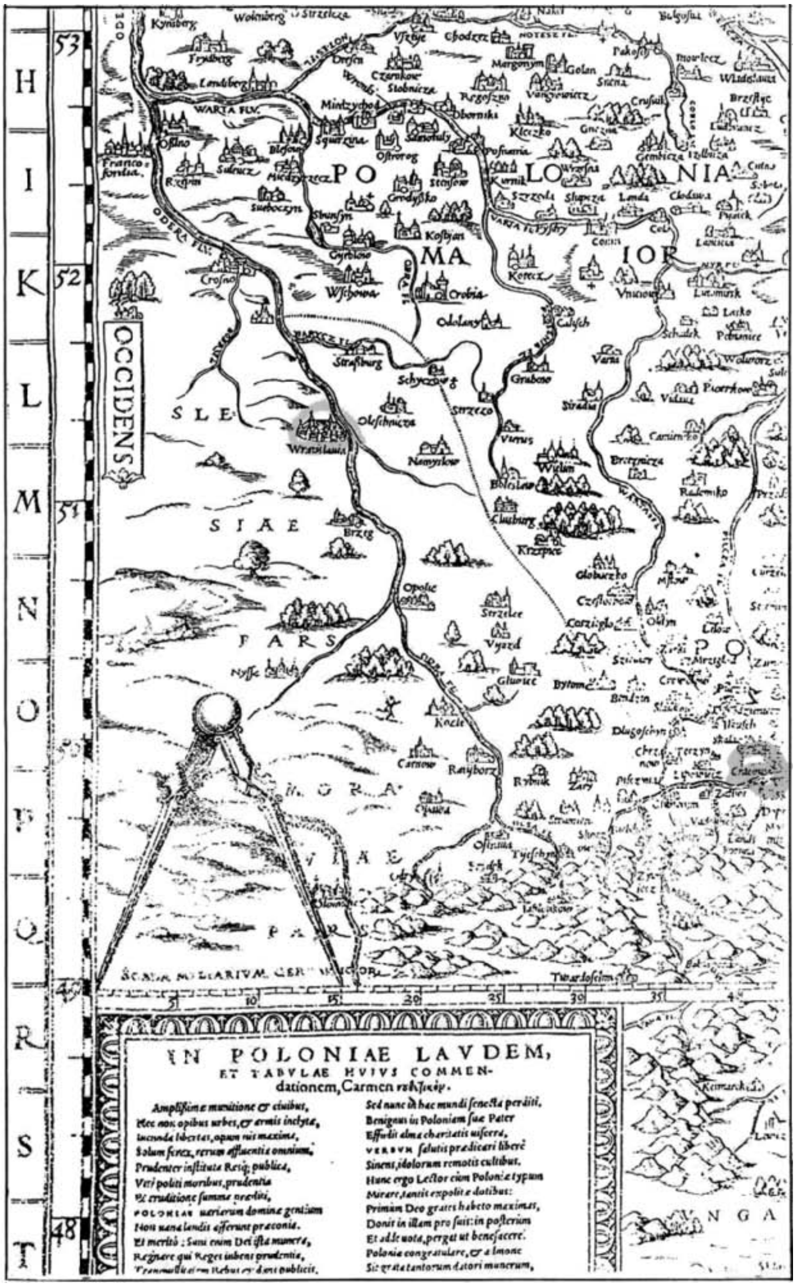

Figure 1. Silesia from a part of Venceslaus Grodeccius' map 1558–1562 (1).

where the king's court and an increasing number of citizens created high demand for fish. In some estates, the income from pond farming was estimated at 50–60% of total income.

The fish caught were distinguished as white and black fish; pike and carp were mentioned separately (Fig. 4). The separate rearing of fish fry in small ponds, chiefly by peasants, points to the high level of fish culture at that time. The need for the development of technologies for pond construction and water leveling found response in one of the first technical books in Polish written by Olbrycht Strumieński and printed in Cracow in 1573 (Fig. 5).

In towns, fishers' guilds were organized to preserve the acquired rules of fishing in rivers and of fish trading. A Czech annalist, J. Dubravius, in 1600 highly appreciated pond farming in Silesia, highlighting the abundance of inexpensive fish in comparison with Czech, Moravian, or Austrian (8,10).

Pond complexes also alleviated and prevented frequent freshets and floods in mountainous regions where forests disappeared owing to developing ferrous metallurgy. In periods of drought, pond waters were used to irrigate fields. Ponds were also built to reclaim marshy sites and to increase the area of cropland and meadows.

THE THIRD PERIOD

At the beginning of the seventeenth century, the fish farming economy collapsed, and this lasted to the middle of the nineteenth century. This was above all due to

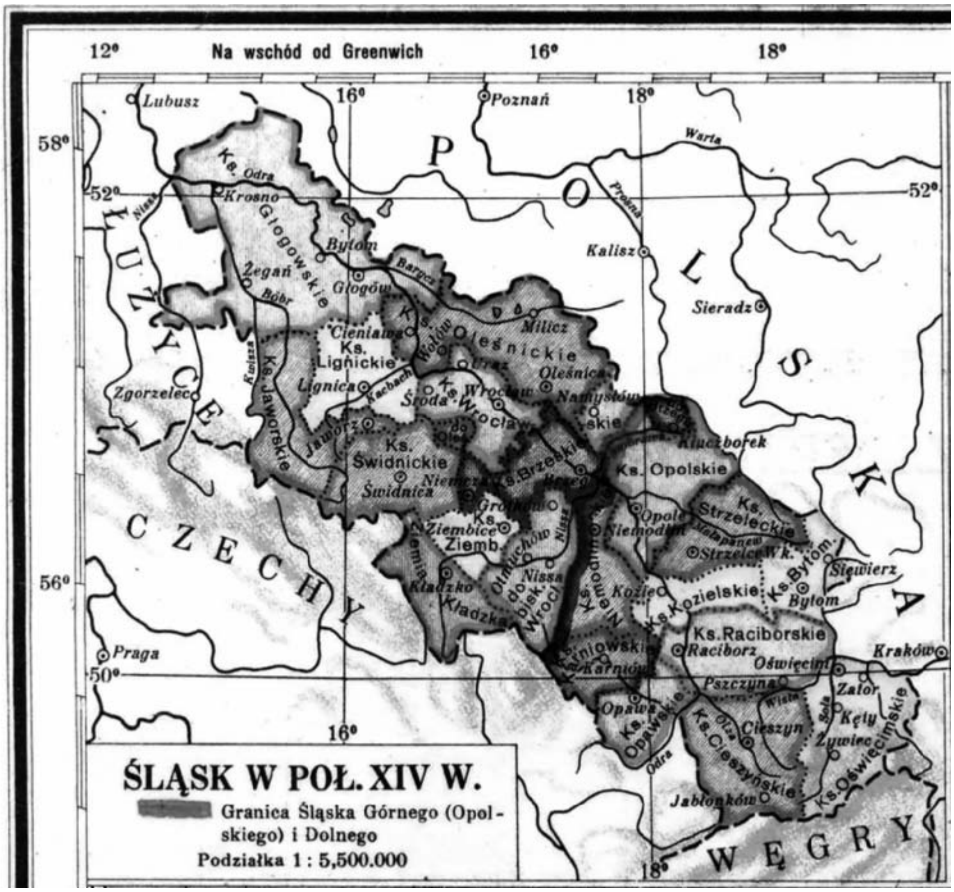

Figure 2. Partition of Silesia in the middle of fourteenth century. Semkowicz, W., Szkolny Atlas Historyczny. 1932, 2nd ed., 8, Książnica—Atlas, Lwów—Warszawa.

the Thirty Years' War (1618–1648) waged on a religious background. The war severely affected Upper Silesia. The behavior of foreign armies passing through the territory, for whose protection they had been called, had the most devastating effects. The majority of soldiers treated this province as enemy terrain causing enormous violence and the devastation made all inhabitants of numerous villages flee. The breakdown of agriculture due to depopulation created hunger and pestilence. Together with a very high contribution paid by the regent on behalf of the Austrian Empire, these events accounted for the poverty of the region (11). The religious repression was a further unfavorable stress on this country. The reconstruction of devastated ponds was impeded by the depopulation; however, the centers of large pond farms were renovated with the help of the villein service. A decrease in the level of groundwater due to the development of coal mining in Upper Silesia dried up shallow peasant ponds (Fig. 6).

The strengthening of the feudal social system and the excessive burden of villein service also affected peasant fish farming unfavorably. The primitive ineffective method of rearing all generations of fish in one pond returned (12,13). In Upper Silesia, the area of ponds decreased by 60% after the Thirty Years' War. A further reduction of about 20% occurred in the eighteenth century owing to the devastating Austrian–Prussian wars (1740–1742 and 1744–1745) (8). The high profit from cereal cultivation contributed to the decline of the fishpond economy and further reduced the area of ponds in the eighteenth century (12). However the area of ponds still reached 30,000 ha in Silesia, about 30% of the whole water surface area in the region at the turn of the nineteenth century (Fig. 7).

In the nineteenth century, the abolition of villein service made the renewal of pond farming very difficult in Upper Silesia because the work was based chiefly on the serf system. In Lower Silesia, the area and number of ponds was reduced to a smaller degree; they were reconstructed by hired specialists, although the regulation and canalization of the River Odra by Prussian authorities decreased the groundwater level and contributed to the disappearance of numerous irrigated territories, also fishponds.

A probable reason for the breakdown of the fishpond culture might also have been the cold weather that prevailed from the midseventeenth to the midnineteenth centuries whose maximum was in the midseventeenth (4,14–16). The cooling, called the Little Ice Age, was also confirmed by decreasing carp sales and an increasing price of this fish species in the Plesnensis Principality from 1690–1810 (Fig. 8).

THE FOURTH PERIOD

A period of continuous development in pond carp culture began from the middle of the nineteenth century and was maintained till the present. It resulted from a fairly

Figure 3. Numerous fishponds in a part of the Upper Silesia. (from a Jonas Nigrini map 1725, copyright by Macierz Ziemi Cieszyńskiej, 2000).

Two groups of freshwater fish distinguished in the sixtennth century

"whitefish"	"blackfish"
Bream - *Abramis brama* (probably also *A. sapa, A. ballerus, Blicca bjoerkna)*	Tench - *Tinca tinca* Crucian carp - *Carassius carassius* Burbot - *Lota lota*
Roach - *Rutilus rutilus* (probably also *Scardinius erythrophthalmus*)	Loach - *Misgurnus fossilis* (probably also spined loach - *Cobitis taenia*)
Perch - *Perca fluviatilis* (probably also *Stizostedion lucioperca*)	Wels - *Silurus glanis* Carp - *Cyprinus carpio*
Gudgeon - *Gobio gobio*	
Ruffe - *Gymnocephalus cernua*	
Pike - *Esox lucius*	

Figure 4. Two groups of freshwater fish distinguished in the sixteenth century (elaborated by M. Kuczyński).

peaceful time at the end of the nineteenth and the beginning of the twentieth centuries. The expansion of numerous studies on carp physiology and diet played an important role, but the main impulse for the culture development was observations by Tomas Dubisz of an accidental transfer of larvae during a freshet from spawning ponds to nursery ponds, previously plowed and sown with clover (Fig. 9) (18).

The observation caused him to introduce a new method of carp rearing, which ensured the abundance of food for

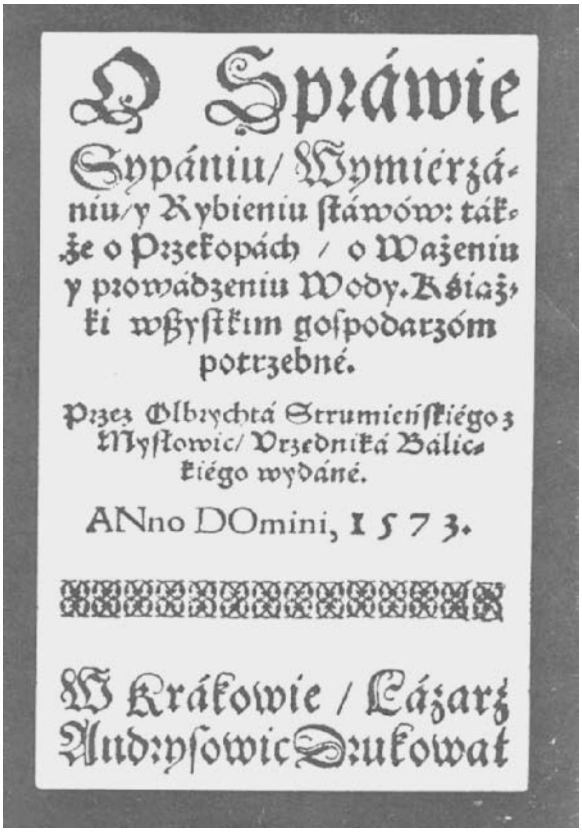

O Spráwie
Sypániu/ Wymierzániu/ y Rybieniu ſtáwów: także o Przekopách / o Ważeniu
y prowádzeniu Wody. Kśiążki wſzyſtkim goſpodarzóm
potrzebné.
Przez Olbrychtá Strumieńſkiego z
Myſłowic/ Vrzędniká Bálickiego wydáne.
ANno DOmini, 1573.
W Krákowie / Łázarz
Andrzyſowic Drukował

Figure 5. Cover of the book "About Managing, Pond Building and Stocking, Also About Digging, Water Levelling and Conducting, Book Useful for Every Farmer" by Olbrycht Strumieński, 1573 (9).

carp fry and shortened the rearing period of the market carp from 6–7 to 3 years. The effects of the fishpond culture also increased due to the warm period from the middle of nineteenth century (4,13). In Upper Silesia, most large fishponds were reconstructed by making them smaller and shallower; this created beneficial conditions for an increase in food resources for the carp and facilitated fish culture technology. However, during this time, large areas of ponds that vanished were turned into land ecosystems.

In the twentieth century, further steps in the intensification of fish culture consisted of the introduction of nitrogen–phosphorus fertilization in ponds. The application of this method in the 1950s increased yields from several kilograms to several hundred kilograms from 1 hectare of pond. A new intensified method of carp rearing introduced in fishery practice in the 1970s was based on greater stock densities and feeding fish pellets which covered the requirements of fish. By this method, yields rose to several thousand kilograms from 1 hectare. The polyculture of carp and herbivorous fish and the aeration of pond waters maintain proper environmental conditions (19).

In recent years, great attention was paid to the fishponds as ecosystems integrated with a catchment. Methods of the semi-intensive carp production, introduction of the polyculture with herbivorous fish species, and limited pond fertilization should ensure the sustainable functioning of carp ponds in river drainage. Moreover, these ecosystems enrich the river drainage in many ways (20).

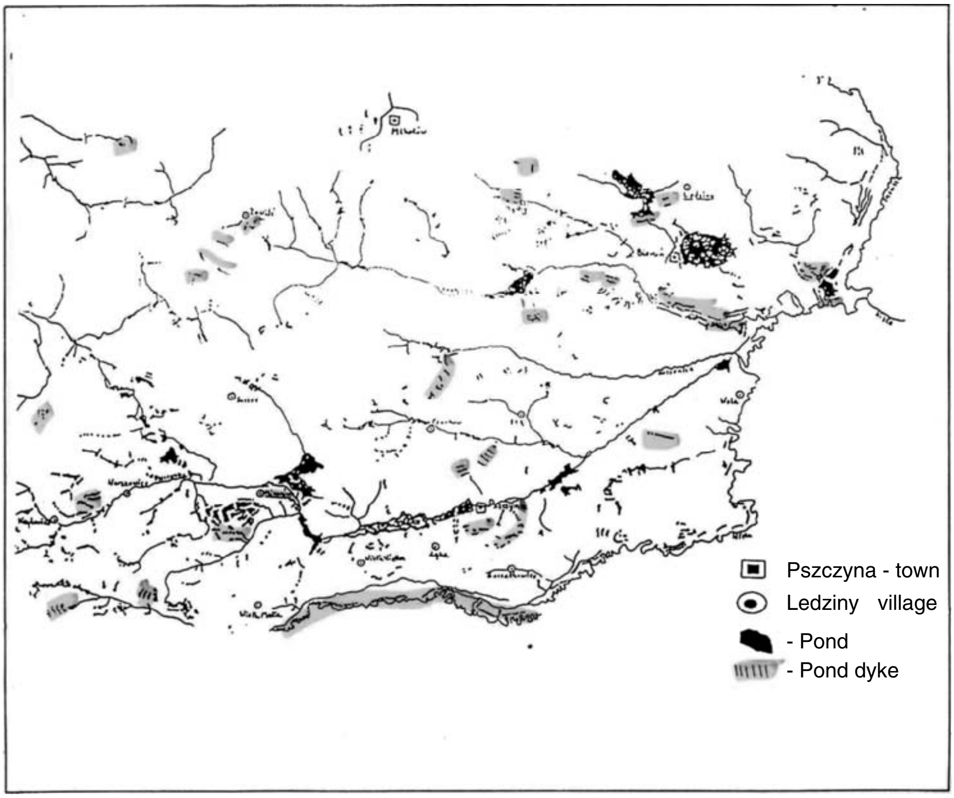

Figure 6. Numerous pond dikes in the Principatus Plessenemsis in the second half of the eighteenth century (from Schlenger's map elaborated by Janczak, Nyrek, and Wiatrowski 1961) (14).

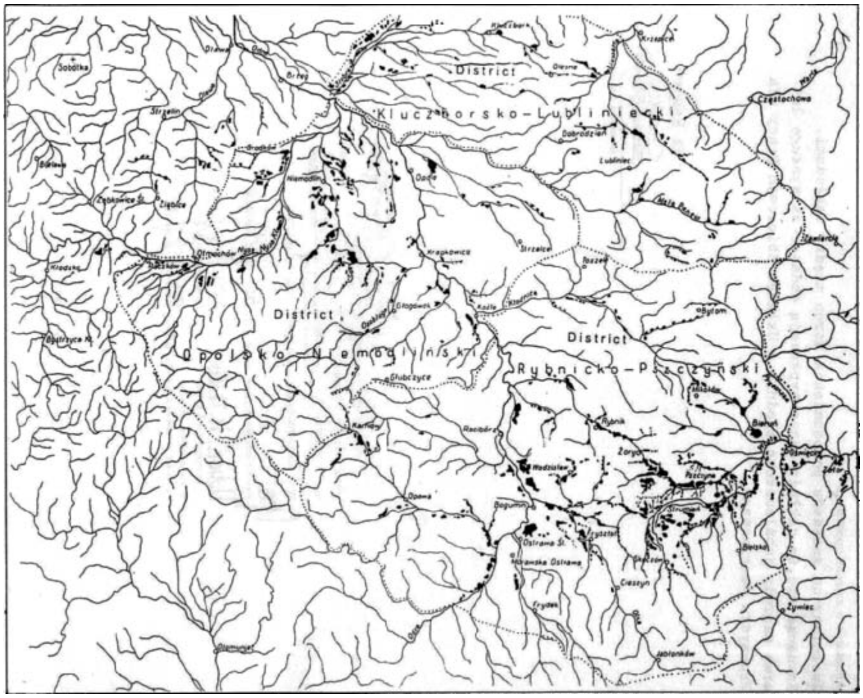

Figure 7. Water net and fishponds in Silesia in the first half of the eighteenth century (from Wieland-Schubbart's map 1736 elaborated by Nyrek (8).

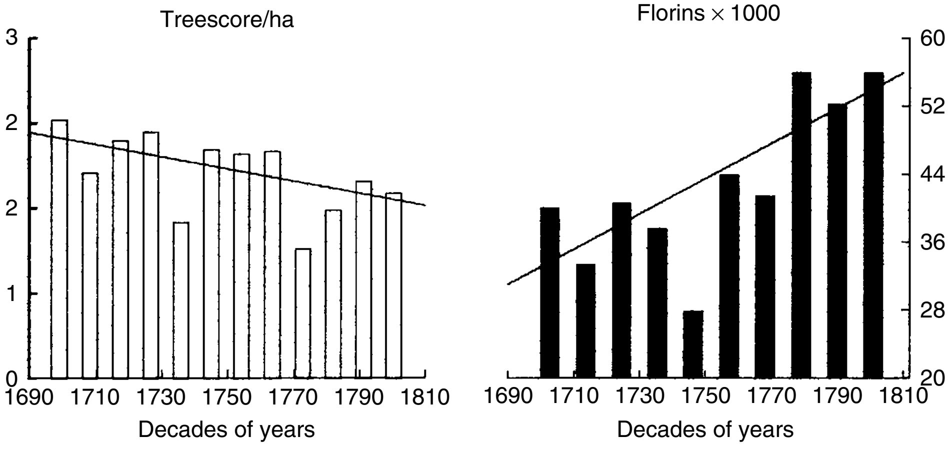

Figure 8. Average 10-year production and income from carp culture in ponds of Principatus Plessenemsis between 1690 and 1810.

BIBLIOGRAPHY

1. Buczek, C. (1939). Venceslaus Grodeccius: *Poloniae tabula, 1558–1562* (original form). *Monumenta Poloniae Cartographica (Atlas)*, sygn. 1423B IV. PAN, Kraków.
2. Olszewicz, B. (1936). *Najdawniejsze opisy geograficzne Śląska*. (The most ancient geographical description of Silesia). Wyd. Inst. Śląskiego, Katowice, Nasza Księgarnia, Warszawa.
3. *Encyklopedia Powszechna PWN*. (1983–1987). The General Encyclopedia, Polish Scientific Editions, PWN, Warszawa.
4. Semkowicz, W. (1922). Zagadnienia klimatu w czasach historycznych. (Problems of climate in historical times). *Przegl. Geogr.* **3**: 18–42.
5. Zubek, A. (1979). Janusowe oblicze klimatu (Janus' climate face). *Przegląd Techniczny* **13**: 42–43.
6. Aleksandrowicz, S.W. (1996). Holoceńskie fazy intensyfikacji procesów osuwiskowych w Karpatach (Holocen phases of intensification in the landslide processes). *Kwartalnik AGH. Geologia* **22**: 223–262.
7. Starkel, L. (2000). Heavy rains and floods in Europe during the last millennium. *Instytut Geografii UJ, Prace Geograficzne* **107**: 55–62.
8. Nyrek, A. (1966). *Gospodarka rybna na Górnym Śląsku od połowy XVI do połowy XIX wieku (Fish culture in the Upper Silesia from the middle of sixteenth to the middle of nineteenth century)*. Prace Wrocł. Tow. Nauk., A **111**.
9. Kwaśniewska-Mżyk, K. (1987). Strumieński, O. (1573). *O sprawie, sypaniu, wymierzaniu i rybieniu stawów, także o przekopach, o ważeniu i prowadzeniu wody. Książki wszystkim gospodarzom potrzebne*. (About managing, pond building and stocking, also about digging and water levelling and conducting. Books useful for every farmer). Inst. Śląski, Opole.
10. Inglot, S. (1968). *Zjawiska klimatyczno-meteorologiczne na Śląsku od XVI do połowy XIX wieku. Z badań nad wpływem posuch na rolnictwo Dolnego Śląska* (Climatic and

An den Erzherzoglichen Fiſchmeiſter, uſw. Herrn
Thomas Dubiſch,
Sketſchau. Oeſterr. Schleſien.

Wer da beiträgt, ſeinen Mitbürgern geſunde und wohlſchmeckende Nahrung zu beſchaffen, den darf man preiſen. Seit einer Reihe von Jahren haben in vielen Landen der Erde viel edle Männer hierfür auch dadurch geſorgt, daß ſie die Teiche, die Seen, die Bäche, die Ströme reicher an edlen Fiſcharten zu machen ſtrebten, nachdem allzulange von den Fiſchern ſorglos nur geerntet war, ohne zu ſäen, ohne rückſichtsvoll zu ſchonen – ohne, ſagen wir es: das Waſſer zu beſtellen, wie man die Aecker beſtellt, wenn ſie reiche Frucht geben ſollen. Schöne Erfolge ſolcher Arbeiten beginnen ſich zu zeigen, und das Herz des Menſchenfreundes erbaut ſich hieran um ſo mehr, da weithin über die Erde ein brüderlicher Austauſch nicht bloß der Erfahrungen, ſondern auch edlen Fiſcharten von Land zu Land, ja von Welttheil zu Welttheil ſtattfindet. Wir ſtehen zur Zeit noch mitten in dieſer Bewegung, aber wenn einſt ein ruhiger Forſcher zurückblickt, die Reſultate feſtſtellt und die Männer zählt, welche beſondere Verdienſte in dieſer Sache hatten – dann, geehrter Herr Dubiſch, wird ſehr warm auch Ihrer gedacht werden! Von frühester Jugend an ſind Sie mit vollem Eifer Fiſcher geweſen, haben aber nicht allein Fiſche fangen gelernt, ſondern ſich auch deren Pflege im Zoologiſchen Garten zu Wien gewidmet. Schon vor einem Vierteljahrhundert ward Ihnen eine große Auszeichnung für Salmonidenzucht verliehen. Aber Ihr größtes Verdienſt iſt ſicherlich die verbeſſerte Karpfenzucht – was Sie hierin leiſteten, kann Ihnen niemals vergeſſen werden. Das Dubiſchverfahren iſt ein weltberühmtes Wort geworden. In dieſer Beziehung hat auch der Deutſche Fiſcherei-Verein Ihnen ſehr viel Dank zu ſagen. Nicht in Heimlichkeit hüllten Sie Ihre ſchönen Erfolge bei dem erlauchten Erzherzog, dem Sie faſt zwanzig Jahre lang zu dienen die Ehre hatten. Sie zogen Schüler heran, welche Ihre Methode verbreiteten, und wenn der Deutſche Fiſcherei-Verein ſeine Mitglieder zum Beſten des Vereins zu Ihnen entſendete, ſo belehrten Sie dieſelben ohne Rückhalt, ſo daß wir großen Vortheil davon genoſſen haben und ſicherlich noch viel mehr genießen werden, wenn Ihre der Natur abgelauſchten Maßnahmen immer mehr in ihrer Zweckmäßigkeit, weil auf Naturgeſetzen begründet, erkannt werden. Im Stillen haben Sie alſo gewirkt, von hochgeſinnten Patrioten auf das Wärmſte anerkannt – wir nennen den Profeſſor Nowicki und den Baron Waltmann – es drängt uns aber, Ihnen auch Namens unſeres Vereins und aller Fiſchereifreunde Deutſchlands unſern innigſten Dank auszuſprechen, Ihnen jede Lebensfreude zu wünſchen. Möge es Gottes Wille ſein, daß Sie noch lange Jahre in ungetrübter Geſundheit verbringen.

Der Ausſchuß
des Deutſchen Fiſcherei-Vereins
von Behr
Königlicher Kammerherr
Präſident

Berlin, den

Figure 9. The Thomas Dubisz brevet (17).

meteorological phenomena in Silesia from the sixteenth to the middle of the nineteenth century. From investigations on impact of drought on agriculture in Lower Silesia). *Prace Wrocł. Tow Nauk.* B **139**: 9–47.

11. Chlebowczyk, J. (1966). Gospodarka Komory Cieszyńskiej na przełomie XVII-XVIII oraz w pierwszej połowie XVIII wieku (Economy of the Cieszyn Estate at the turn of the seventeenth and in the first half the eighteenth century). *Zakład Narodowy Imienia Ossolińskich*. Wydawnictwo Polskiej Akademii Nauk. Wrocław-Warszawa-Kraków.
12. Szczygielski, W. (1967). *Zarys dziejów rybactwa śródlądowego w Polsce* (An outline of the history of inland fisheries in Poland). PWRiL, Warszawa.
13. Nyrek, A. (1992). *Kultura użytkowania gruntów uprawnych, lasów i wód na Śląsku od XV do XX wieku*. (Land use culture of the arable land, forests and waters in Silesia from the fifteenth to twentieth century), Wyd. Uniw. Wrocł., *Seria Historia*, Vol. 47.
14. Nyrek, A. and Wiatrowski, L. (1961). Gospodarka rybna w księstwie pszczyńskim od końca XVII do początku XIX w. (Fish culture in Principatus Plessenensiş from end of the seventeenth to beginning of nineteenth century). *Zeszyty Nauk. Uniw. Wrocł.* A **3**: 6–128.
15. Stachlewski, W. (1978). *Klimat: przeszłość, teraźniejszość i przyszłość.* (Climate: past, present and future). PWN, Warszawa.
16. Bednarz, Z. (1996). June–July temperature variation for the Babia Góra National Park, Southern Poland, for the period 1650–1910. *Zeszyty Naukowe UJ, Prace Geograficzne* **102**: 523–529.
17. Łysak, A. (1996). Historia rybactwa rzecznego i stawowego w Polsce Południowej. (History of the river and pond fishery in southern Poland). *Materiały Konferencji: Aktualne zagadnienia gospodarki stawowej w Polsce*, Ustroń-Gołysz 1966, marzec 11–13, 19–30. Papers of the Conference: Current problems of pond culture in Poland. Ustroń-Gołysz March 11–13, 19–30.
18. Brzozowski, S. and Tobiasz, M. (1964). Z dziejów rybactwa małopolskiego. (From the history of Małopolska fishery). *Studia i materiały z dziejów nauki polskiej. PWN* B **9**: 1–99.
19. Szumiec, J. (1995). Effects of the diversified pond carp culture. Impact of different feeding and stock density on fish production. *Acta Hydrobiol.* **37**(Suppl. 1): 131–138.
20. Szumiec, M.A. (1996). The role of carp pond ecosystem management in wetlands conservation. *Acta Hydrobiol.* **37** (Suppl. 1): 13–20.

FLOODS

Steven Jennings
Eve Gruntfest

(from *The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts*, Wiley 2003)

INTRODUCTION

The interface between humans and hydrologic features across Earth's surface has helped shape human culture. From the earliest agricultural, complex societies established along some of the great rivers of the world to the bustling seaports of today, humans have gained from the myriad advantages of living in proximity to water. Fertile soil, ease of transportation, and availability of resources (both materials and energy) have allowed for the development of complex material and intellectual cultures. The relationship between water and humans also brings a great deal of risk. Flooding is one of these risks. The impact of floods on humans has been evident from *Genesis* to tonight's evening news. Early Mesopotamian maps may have been drawn to facilitate the reestablishment of property lines after flooding. While the impacts of flooding on humans have been positive in the case of fertile floodplains that support much of the world's agricultural productivity, there is the potential for a great deal of negative impact (Brown, 1984; Clark et al., 1985). Losses of life and property have focused the efforts of scientists, engineers, and government agencies on the prediction, control, and mitigation of floods and flood damage.

In spite of efforts to deal with flooding problems, monetary losses continue to rise at an alarming rate. In Venezuela in December 1999, two weeks of heavy rain resulted on December 15th in flash floods laden with soil, vegetation, and debris. Damages of US$3.2 billion, or 3.3% of the country's gross domestic product were reported. At least 20,000 people were killed. Generally, the number of lives lost due to flooding remains high. However, improvements in flood warnings particularly for major large-scale storms such as cyclones and typhoons have had dramatic effects. The severe 1991 cyclone in Bangladesh resulted in 140,000 dead and property losses of US$2.0 billion. A cyclone of similar intensity in 1993 resulted only in the loss of 126 lives. The early warnings and cyclone shelters accounted for the major improvement (www.ndndr.org). China has also witnessed a reduction in the number of lives lost to floods. While 3000 people died in 1998 floods, the 1998 floods were as great as those of 1931 and 1954 where the loss of lives was 145,000 and 33,000, respectively (www.ndndr.org). In October 1998 hurricane Mitch was the worst in the eastern Caribbean since 1780 when a hurricane killed 22,000 people. The death toll from Mitch is reported as 11,000. More than 3 million people were left homeless or were severely affected (www.ncdc.noaa.gov/ol/reports/mitch/mitch.html).

Floods take a variety of forms with the interplay of several factors leading to the inundation of normally dry land. Wohl (2000) identifies four primary challenges in reducing escalating flood damages. These are (1) estimating flood magnitude for a given recurrence interval, (2) accurately forecasting floods based on rapidly evolving weather conditions, (3) effectively operating flood-warning and evacuation procedures, and (4) establishing and enforcing land-zoning regulations. This chapter first discusses the contexts and causes of flooding, the first two points addressed by Wohl (2000). The second topic is the complex human responses to floods, points 3 and 4 of Wohl (2000). Many of these topics are illustrated with examples of floods.

DEFINITION OF FLOODS

Streams are linear water features that flow under the impetus of gravity. The amount of water contained in a stream is usually regulated by contributions of groundwater and surface runoff to the stream channel (Zaslavsky and Sinai, 1981; Knighton, 1998). Much of the time water in a stream flows within the confines of its channel. When inputs of water increase sufficiently, stream discharge leaves the stream channel and covers all or parts of the adjacent floodplain. Since the floodplain surface is usually a virtually flat surface and near the elevation of the stream channel, water can easily spread over the floodplain once water exceeds the elevation of the stream's banks. Most floods develop over a period of days or months as discharge increases gradually (Hirschboeck, 1987, 1988). Flash floods by contrast occur suddenly with little warning and are of short duration. Semiarid and arid areas are likely to experience flash floods (Reid and Frostick, 1987; Hassan, 1990). Flooding is not always associated directly with stream channels. Flooding occurs any time when water covers a surface that is normally not under water. Flooding can occur in coastal areas, low lying areas with poor drainage, or locations with inadequate urban drainage systems.

FACTORS THAT LEAD TO FLOODING

Floods have a multitude of causes. Some causes are related to what would be considered natural processes that would occur whether humans are present or not. Many causes have been affected by human activities. In some cases the severity of floods and the types of damage are a direct result of agriculture, urbanization, and the areas selected for development. In all cases flooding is related to increased discharge in stream channels.

Saturated Soil

Much of Earth's surface is covered by a weathered cover of regolith. Whether forming a true soil with well-developed horizons or a weakly developed detrital cover, the regolith is composed of a mix of mineral particles, organic fragments, and pore space. Commonly, much of the pore space is filled with air and to a lesser extent water. When large amounts of precipitation are received in a region, the pore space fills with water as the input of water from precipitation exceeds the output of water from the soil column to the water table. Decreases in

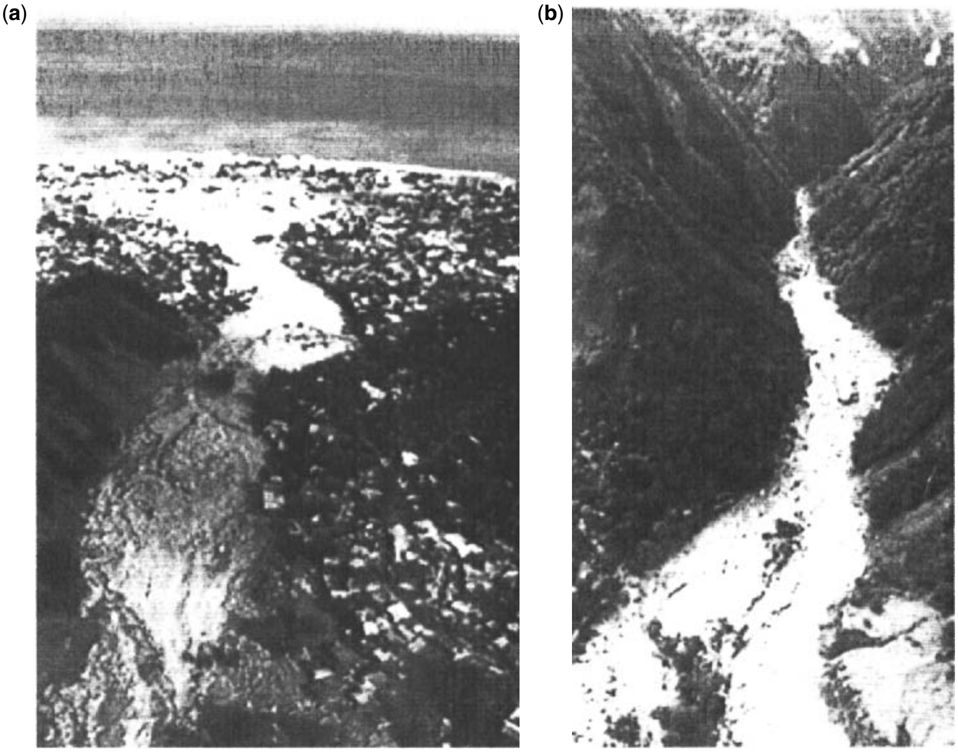
(a) (b)

Figure 1. Quebrada San Julián upstream of Caraballeda showing evidence of recent debris flows and flash floods. Note the high slope angles, large numbers of debris flow scars, and abundance of new alluvium and colluvium in the channel bed and fan surface.

infiltration lead to increases in runoff. The lag time between the precipitation event and the arrival of water to stream channels decreases significantly when soil saturation occurs. As a result, peak discharge increases significantly and the likelihood of overbank flow is high (Smith and Ward, 1998). Spatially, soil saturation may occur over large-scale basins, which leads to flooding in large areas. The peak discharge flows downstream and becomes concentrated in higher order streams causing flooding. In many cases saturation follows a period of high amounts of precipitation over a prolonged time period, possibly weeks or months (Wolman and Gerson, 1978; Ward and Robinson, 1990).

Basin Characteristics

Surface characteristics influence infiltration and runoff rates (Roberts, 1989; Kuhnle et al., 1996). Impervious surfaces such as exposed bedrock or a paved road accelerate surface runoff, thus decreasing lag time between the precipitation event and entrance of water into a nearby channel. Urbanized areas, therefore, with large percentages of impervious surface such as roofs, streets, and parking lots coupled with an engineered drainage system designed to move water quickly to stream channels greatly increase the chances that some flooding will occur after a significant precipitation event (Wolman, 1967; Hammer, 1972; Roberts, 1989; Newson, 1992). Conversely, rural areas with large areas of soil, natural vegetation, and the potential for a faster infiltration rate are less likely to have significant flooding resulting from a single precipitation event. Removal of as much as half the forest cover and a decrease of marsh land along the Yangtze River in China has led to increased flooding. Half a billion people, or 45% of China's total population, reside on the banks or floodplains of the Yangtze and the area produces about 42% of China's gross domestic product. In 1998, 79.6 million people in three Chinese provinces were affected by repeat flooding on the Yangtze. The floods killed more than 3000 people. Fourteen million people were evacuated and 21 million were made homeless (Weather.ou.edu/spark/AMON/v2_n3/News/DR_980819-China12.html).

Topography

Topography will influence the rate at which precipitation will be incorporated as stream discharge (Patton, 1988). Steep, rocky canyon walls have low infiltration rates as well as a great deal of potential gravitational energy that leads to the concentration of discharge during a short period of time (Strahler, 1964). Alluvial plains usually have a much longer lag time between a precipitation event and the introduction of runoff water into a stream channel. When land cover on steeper slopes is affected by perturbations such as wild fire or building-related oversteepening of slopes, the likelihood of mass movement events is greatly increased. These events are usually related to unstable regolith on steep slopes, which is susceptible to failure when sufficient precipitation is received. For example, see Fig. 1.

High Amounts of Precipitation

Flooding is created by the delivery of larger than normal amounts of runoff into stream channels (Smith and Ward, 1998, p. 67). Periods of above-average precipitation lead to floods. In some cases seasonal variability leads to great fluctuations in stream discharge. Wet–dry subtropical or monsoonal climates with distinctive seasons of precipitation lead to fluctuations from dry stream channels to potential flooding events. These cyclical events are related to large-scale atmospheric circulation patterns that operate through an annual or longer period. In the midlatitudes, the annual migration of subtropical high pressures and the polar front lead to distinct precipitation patterns. In the tropics, monsoonal flow can lead to large precipitation events (Milne, 1986). On longer time scales El Niño and La Niña events are persistent over several years and can lead to wet or dry conditions over large areas of Earth's surface from the Equator to the midlatitudes (Waylen and Caviedes, 1987; Pearce, 1988; Ely et al., 1994).

Extended Wet Periods

In many cases flooding is caused by the reception of precipitation over an extended time period, on the order of weeks to months, that leads to the saturation of soils in a large-scale region (Rodda, 1970b; Smith and Ward, 1998). This saturation leads to increased runoff at a time when streams are at capacity (Ward and Robinson, 1990). Additional water introduced to stream channels cannot be conveyed in the channel but is spread across the floodplain. Wet periods are related to synoptic conditions such as the position of the polar front that delivers cyclonic storms in quick succession. Poleward migration of subtropical air masses over continental areas such as the Mississippi River Basin help to supply large amounts of water to be precipitated by frontal activity. For example, see Fig. 2. In some locations rainfall may fall on snow-covered or frozen ground (Thomas and Lamke, 1962). These waters are unavailable to the hydrologic cycle as long as they remain in a solid form. In the case of the former, rainfall may accelerate the introduction of water into the stream network as snowmelt augments the precipitation already being received (Kattelman, 1990; Naef and Bezzola, 1990; Caine, 1995). The latter will greatly decrease the infiltration capacity of the soil causing most of the precipitation to quickly enter the stream network (Horton, 1933).

Decaying Tropical Cyclones

Some of the largest precipitation amounts received as the result of a single meteorological event have been associated with the movement of tropical cyclones (e.g., hurricanes, cyclones, and typhoons) poleward and over continents. These powerful cyclonic storms carry large

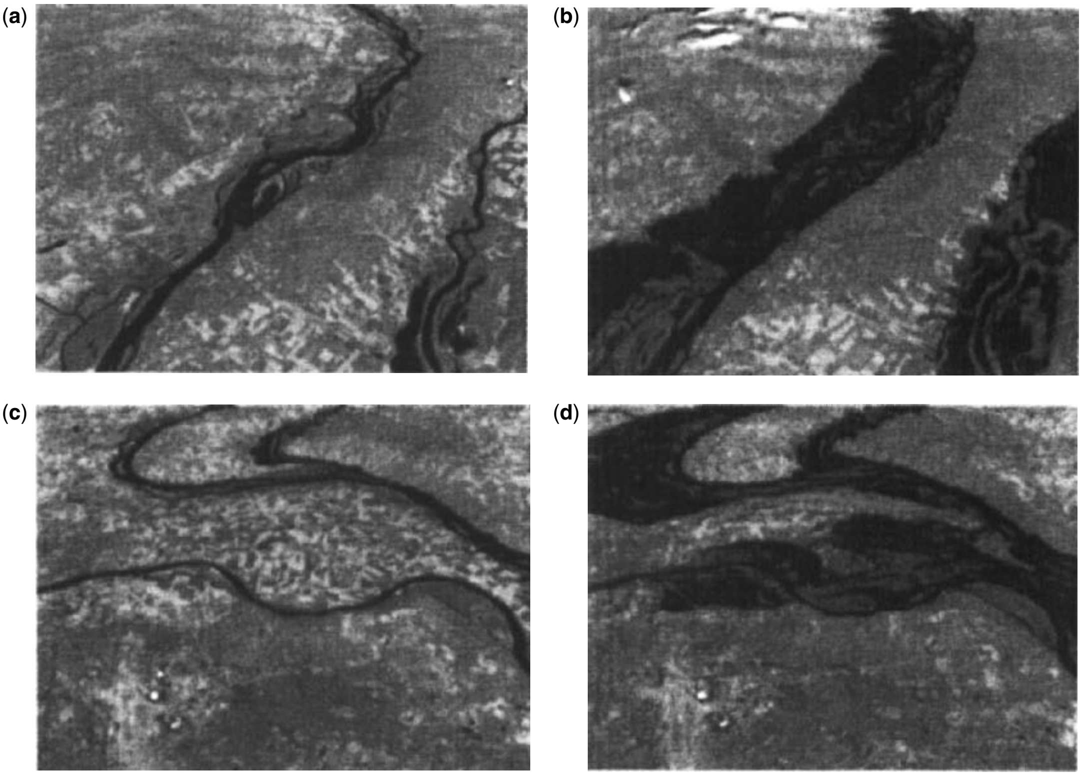

Figure 2. These scenes show various sections of the Mississippi River near St. Louis before and just after the 1993 floods, which peaked in late July/early August. The images show the area as seen by the LandSat Thematic Mapper (TM) instrument. The short-wave infrared (TM band 5), infrared (TM band 4), and visible green (TM band 2) channels are displayed in the images as red, green, and blue, respectively. In this combination, barren and/or recently cultivated land appears red to pink, vegetation appears green, water is dark blue, and artificial structures of concrete and asphalt appear dark gray or black. Reddish areas in the scenes during the flood show where water had started to recede, leaving barren land. See online version for color.

Figure 3. Water and sand washed inland to make travel difficult in North Topsail Island, North Carolina, after hurricane Fran.

amounts of warm moist air over land surfaces. While wind speeds associated with these storms decrease quickly after landfall, these decaying storms are capable of delivering precipitation over wide areas during a relatively short period of time, on the order of days to weeks. In some cases cyclonic storms associated with the polar front may exacerbate conditions by introducing a lifting mechanism that leads to increased condensation and precipitation. The relatively low-lying coastal plain of eastern North America is especially susceptible to damage from these types of storms (Bailey and Patterson, 1975; Hirschboeck, 1988). For example, see Fig. 3. In 1998 hurricane Mitch produced as much as 50 to 75 inches of precipitation in some areas of Central America. At least 11,000 deaths were associated with hurricane Mitch and more than 3 million people were left homeless or were severely affected (www.ncdc.noaa.gov/ol/reports/mitch/mitch.html).

Intense Thunderstorms

Thunderstorms are usually intense, short-lived storms that produce high winds, hail, and heavy rainfall. These storms can be caused by convection in moist tropical air masses over continental surfaces or fast-moving cold fronts that displace those moist air masses (Hirschboeck, 1987). When these storms develop over mountainous areas where the precipitation is concentrated by the topography the potential for large, catastrophic floods is great (Hall, 1981). For example, see Fig. 4. The eastern slope of the Rocky Mountains and the southwestern deserts of North America are common locations for the development of thunderstorms. As moist air encounters higher elevations in these locations, it is forced to rise. Unstable atmospheric conditions are created as mountain slopes heat and in turn heat the atmosphere. Adiabatic cooling causes condensation and the development of large cumulonimbus clouds that can reach the upper altitudes of the troposphere. Sometimes there is little movement associated with a thunderstorm or thunderstorm complex, with respect to the ground; heavy precipitation concentrated in a small geographical area can have catastrophic results.

Figure 4. Arizona flash flood, Wenden, Arizona. This community was flooded twice in late October 2000 when waters from Centennial Wash swept into the town. (Photo courtesy of U.S. Small Business Administration).

Quick Snowmelt

The storage of water in the form of snow temporarily removes that water from the hydrologic cycle. In many cases this sequestration of water is short term. Snow accumulates during winter especially at higher elevations and latitudes. With the onset of warmer spring and summer conditions, snowmelt supplies water to streams. A typical early warming may mean that snowmelt may be accelerated with large amounts of runoff entering stream channels. Mountain ranges in mid-latitude coastal regions such as the Coast Ranges and Sierra Nevada of California receive a significant portion of their annual precipitation in the form of snow. It is possible for warm early spring rains to fall on the snowpack, causing much faster runoff than normal (Bolt et al., 1975; Church, 1988). Another source of snowmelt is the subsurface introduction of heat from volcanic activity. Large volcanoes can be high enough to support permanent snow and ice cover. High temperatures associated with volcanic activity lead to the instantaneous melting of snow and ice. The melt water is commonly mixed with pyroclastic debris to form lahars (Smith, 1996).

Failure of Flood Control Structures

A variety of humanly constructed structures are used in an effort to limit the extent and severity of flooding (Gregory, 1995). Dams and levees are common flood control structures designed to contain water within designated areas (Brookes, 1985, 1988). These structures can fail because of construction errors, poor design, and overtopping by water (Biswas and Chatterjee, 1971; Costa, 1988). Flood control structures can fail because of the failure of a key component. For example, a spillway that erodes away has the potential to lead to the catastrophic failure of the entire dam as the water cuts downward. Sound structures may fail when the water retained by the structure exceeds the height of the structure. Large precipitation events or the displacement of water in a reservoir have the potential to send water flowing over the flood control structure (Kiersch, 1964). This may lead to the failure of the structure through erosion. Flooding may be exacerbated by these structures since a feature such as a levee tends to raise the stream level well above the floodplain. When a levee fails, a large amount of fluvial energy is concentrated through that break and a great deal of damage can occur near the break.

Cyclonic Storm Created Surges

In low-lying coastal locations a temporary increase in sea level associated with the approach and landfall of storms with significantly high winds and low central pressure can cause significant damage. Sea level rises in response to low pressure as it passes over ocean surfaces. Additionally, the upper portion of the water column is pushed into waves by the high winds. Storm surges can be more than 5 m above the normal high tide (Rappaport, 1994). In some areas such as bays coupled with low-lying deltas, like the Bay of Bengal and the mouth of the Ganges, where storm energy is concentrated, storm surges can reach high levels causing significant flooding (Frank and Husain, 1971; Murty and Neralla, 1992). Barrier islands are also susceptible to flooding by storm surges. Development on barrier islands along the southeastern coast of North America has led to rising property damage related to storms.

Mass Movement Events

A variety of mass movement events, while strictly not fluvial events, behave in a similar way to floods (Carson, 1976). The gravitationally fueled downhill movement of poorly consolidated regolith results from the introduction of meteoric water that adds weight and decreases hillslope cohesion. These events can do significant damage. Several types of mass movement events are composed of a larger percentage of sediments than a typical stream. Events such as mudflows, or lahars, commonly may approach the viscosity and velocity of streams. Valleys can be filled with fine-grained sediments as the deposits dewater following the initial surge of water and sediment. A variety of factors lead to mass movement events. The removal of plant cover by fire may expose soil surfaces so that infiltration rates may increase and lead to the accumulation of water along failure planes in the regolith. In areas with a subtropical wet–dry climate, such as the Mediterranean climate type, the burning of plant cover during the dry season and a subsequent wet season before the reestablishment of plant cover leads to mass wasting events (Rice et al., 1969; Campbell, 1975).

Human Responses to Flooding

There are no accurate estimates of the population in the world's floodplains. Even in the United States, only broad estimates are available, but the trends to increased vulnerability are clear. In 1955 U.S. floodplains had 10 million occupants. Thirty years later the number doubled to 20 million and by the mid-1990s about 12% of the national population lived in areas of periodic inundation. One sixth of the nation's floodplains are urbanized, and they contain more than 20,000 communities susceptible to flooding. Half of these communities have been developed since the early 1970s (Burby, 1985; Montz and Gruntfest, 1986; Alexander, 1993).

Many of the people at risk do not understand the potential consequences of the hazards they face. In the United States, flood damages exceed $2 billion annually. Only 20 to 30% of eligible structures are insured against flooding. Federal and state disaster assistance accounts for most of the difference. In the United States, almost two-thirds of the residential flood losses result from events that occur once every 1 to 10 years, even though the 100-year floodplain regulation is standard (Alexander, 1993).

In the United States, floods tend to be repetitive phenomena. From 1972 to 1979, 1900 communities were declared disaster areas by the federal government more than once, 351 were inundated at least three times, 46 at least four times and 4 at least five times. As of 1993, the United States was said to spend $9 billion a year on flood control and $300 million on flood forecasting (Alexander, 1993; Conrad, 1998).

Definitions of Structural and Nonstructural Measures

Adjustments to floods can be broadly classified into structural and nonstructural measures. Nonstructural approaches involve adjustment to human activity to accommodate the flood hazard (White, 1964; James, 1975; White, 1974) whereas structural methods are based on flood abatement or the protection of human settlement and activities against the ravages of inundation.

Structural change involves modification to the built environment to minimize or eliminate flood damage directly or flood channel construction changes. For example, see Fig. 5. Structural measures are expensive. They may give the illusion of security but the record shows otherwise (Alexander, 1993). The security can be temporary. A flood can occur that is bigger than the design of the channel or levee, and changing priorities in flood control projects that require higher reservoir levels for recreation or water supply can diminish the efficacy of structural measures (Williams, 1998).

The failure of structural flood control works poses a significant threat to the lives of the people who live

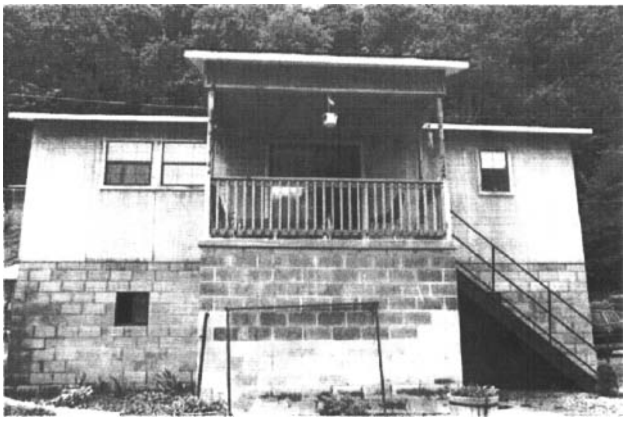

Figure 5. Elevated home in West Virginia is a mitigation success story. Risk is greatly reduced to homes elevated before a flood.

downstream from a massive structural project such as a dam. More than 2000 people died in 1969 in Italy when the Vaiont Dam collapsed (Blaikie et al., 1994). Because of stringent engineering standards and a system of inspections, the United States has seen few major failures. However, many structures are at the end of their design lives of 50, 75, or 100 years.

Structural flood control is still the dominant idea in many parts of the world. Following the 1927 Mississippi River floods, when river levees collapsed and 200 people died, 700,000 were displaced, and more than 135,000 buildings were damaged (Moore and Moore, 1989), the Army Corps of Engineers did not abandon its dream of controlling all floods. Rather, it proposed building large dams upstream to reduce flood peaks to the capacity of the floodway between the levees (Williams, 1998).

Until the 1970s, most flood loss reduction efforts involved structural solutions. Although nonstructural measures were discussed as alternatives, they were rarely implemented. The shift from mostly structural to mixed structural/nonstructural measures began in the 1970s and continues today. The mix of adjustments varies for each situation. In Europe almost all measures that are taken have elements of combined structural and nonstructural measures. There has also been a move to be antistructural. Some dikes are being removed in favor of nonstructural or more environmentally sensitive techniques (Smith and Ward, 1998).

Nonstructural measures include floodproofing, land-use planning, soil bioengineering, warning systems, preflood mitigation efforts, and insurance. The simplest nonstructural measure is to accept the loss. Another nonstructural measure is to provide postflood relief. Protection of floodplain residents and users, and the supply of relief when they suffer damage, are forms of hidden subsidy (Alexander, 1993). This category includes aid provided by the Red Cross, voluntary organizations, and governmental agencies.

Nonstructural measures include flood insurance and land-use management, acquisition and relocation, floodproofing, preflood mitigation preparedness, outdoor warning systems, and soil bioengineering.

DISCUSSION OF NONSTRUCTURAL MEASURES

Flood Insurance, Floodplain Mapping, and Land-Use Ordinances

In 1968 the U.S. National Flood Insurance Program (NFIP) was launched. It made affordable insurance available to residents in flood-prone areas. In 1999 more than 18,000 communities belonged to the program. Participating local governments require developers to meet minimum standards designed to avoid damages that might be inflicted by a catastrophic 100-year flood. The program also requires property owners to purchase flood insurance to receive a federally insured mortgage (Myers, 1996). Flood insurance is a means for placing some of the burden of losses onto the people who take (or make) the risk, namely the floodplain users and residents (Alexander, 1993). Communities can participate in a Community Rating System, established by the Federal Emergency Management Agency (FEMA), that allows them to show innovative strategies to reduce flood losses in return for lower insurance premiums for floodplain residents.

Before a community can participate in the flood insurance program, the flood hazard must be recognized, assessed, and mapped. These assessments include flood history, cost and types of past flood damages, maps of the limits of the 100-year flood (or other designated flood) on a topographic map, compilations of profiles and cross sections of the river to show the levels of past floods, and compilations of flood frequency curves and locally representative hydrographs.

FEMA works with the state and community governments to identify their flood hazard areas and publishes a Flood Hazard Boundary Map of those areas. When a community joins the NFIP, it must require permits for all construction or other development in these areas and ensure that the construction materials and methods used will minimize flood damage. However, there is not careful monitoring to be sure that reducing flood hazard in a particular area does not increase flood potential elsewhere. Often, the problems are just shifted to different locales. In return the federal government makes subsidized flood insurance available to those whose structures were in the flood hazard area prior to issuance of the flood maps. All others are eligible for flood insurance at actuarial rates. FEMA issues a Flood Insurance Rate Map after the Flood Insurance Study of risk zones and elevations has been prepared (http://floodplain.org/Jan32.htm).

Acquisition and Relocation

The most effective measure to reduce losses is to keep the floodplains free of development. However, in many river valleys in the world, it is too late for that option. One of the most promising strategies for reducing flood losses is the public acquisition of developed land susceptible to flooding (Conrad, 1998; www.fema.gov/mit/homsups.htm). The authorization for U.S. federal cost sharing for relocation is more than 30 years old. However, only recently have communities, tired by chronic flooding, taken advantage of funding packages and relocated. In one case, the entire town of Valmeyer, Illinois, was relocated. The town had a long history of floods. In 1943, 1944, and 1947 unusually high levels of the Mississippi caused flooding in the nearby bottomlands affecting Valmeyer. After the 1947 floods, the U.S. Army Corps of Engineers raised the levees protecting the reach of the flood-plain to 47 ft. On August 1, 1993, the flood overtopped the levees inundating Valmeyer, prompting its ultimate relocation. Since 1993 nearly 20,000 properties in 36 states and one territory have been bought out and over 25,000 families have moved from floodplains (http://www.nwf.org/nwf/pubs/higherground/intro.html).

Floodproofing

Floodproofing is a range of adjustments aimed at reducing flood damages to a structure or to the contents of buildings. There are three categories: (1) raising or moving the structure; (2) constructing barriers to stop floodwater from entering a building; and (3) wet flood proofing (U.S. Army Corps of Engineers, 1997).

Detection and Response Warning Systems

New technological advances in stream and rain gage networks and the increased regional floodplain management efforts have led to the adoption of thousands of local flood-warning systems. Many are simple detection systems and do not provide any mechanism for alerting the population at risk. In the United States until the 1990s warning or detection systems were planned and administered primarily at the local level.

Since then, the federal government including the Bureau of Reclamation, the U.S. Army Corps of Engineers, the National Oceanic and Atmospheric Administration, and the Federal Emergency Management Agency have actively participated in the installation and maintenance of detection and warning systems. Many systems are still managed by regional or local entities, but the percentage of federal dollars has increased substantially. Standards have also been established to help make the systems more compatible across regions (U.S. Department of Commerce, 1997).

An automated integrated network of stream and rain gages is being used in more than 1000 communities in the United States to help provide lead time for floods. Most of the systems are developed through collaborative efforts of many agencies. These ALERT systems (automated local evaluation in real time) have performed many functions other than flood warning, including helping in water supply decision making, fire weather forecasting, pollution monitoring, and providing data for river recreationists (Gruntfest and Huber, 1991). The availability of real-time data on the Internet also has increased interest in these monitoring systems (Gruntfest and Weber, 1998). The State of Arizona is developing a network for flood warning throughout the state. More than 30 agencies and communities are working together on the comprehensive ALERT system (http://www.alertsystems.org/saas/).

Warning systems may be nothing more than “cheap payoffs of the raingods.” Too often communities install rain gage/stream gage monitoring systems without a plan for getting the warning message disseminated. A warning system is only necessary once poor land-use decisions have been made, allowing people to settle in harm’s way. Many of the systems being built are not being adequately maintained to be reliable (Gruntfest and Huber, 1991; Parker and Fordham, 1996). Public education encouraging people to heed environmental cues is also being used. It is particularly difficult to provide adequate lead times for flash floods. Some communities do have drills to test the reliability and completeness of their systems to be sure the systems will operate when the conditions warrant.

As of 2001 a combination of factors increase the likelihood that automated detection systems may become more popular and more valuable. More powerful, less expensive computers, and World Wide Web access provide opportunities for inexpensive real-time weather data. While real-time stream and rain gage networks may be originally installed for flash flood forecasting, many agencies and users find the data useful for alternative purposes.

Soil Bioengineering

Anchored plantings along stream banks serve as the basis for this technique. Soil bioengineering and biotechnical engineering are cost-effective and environmentally compatible ways to protect slopes against surficial erosion and shallow mass movement. These approaches provide alternatives to structural channel “improvements.” They raise questions about the notion of why engineers ever considered that concrete-lined channels should be considered “improved” (Gray and Sotir, 1996). Generally, bioengineering solutions must also include a strategy to carry floodwaters away.

The bioengineering technique is gaining support throughout the United States and Europe. It is less expensive to install and less expensive to maintain as well. The broader adoption of soil bioengineering may radically alter floodplain management.

Combined Structural and Nonstructural Measures to Reduce Flood Losses

From the first attempts to reduce flood losses in the United States, structural measures were preferred for three main reasons: (1) their benefits appeared to be relatively easy to measure, (2) they did not require extensive and politically controversial land-use planning, and, (3) the federal cost-sharing agreements encouraged communities to select the most expensive engineering projects. These reasons were supported by a faith in the technology of structural measures to protect people and property from floods.

The record now shows that in spite of massive expenditures, flood losses have continued to rise. Since the 1960s, especially in the United States, there has been a call for a shift from primarily structural measures to control floods to nonstructural measures (Galloway, 1994; Larson, 1996; Williams, 1998). Land-use control is one of the most effective ways of reducing flood hazards. Statutes, ordinances, regulations, and compulsory purchases can be employed and relocation can be subsidized. A floodway left undeveloped through the city can become beautiful public open space.

CONCLUSION

Floods are generally caused by the combination of large amounts of precipitation and basin topography. For example, saturation of soils caused by large amounts of precipitation can lead to flooding. Urbanization of a drainage basin increases the amount of runoff reaching a channel and decreases the lag time between a precipitation event and peak flow. A variety of weather events lead to flooding, including extended wet periods, decaying tropical cyclones, intense thunderstorms, and quick snowmelt. In some cases humanly constructed structures designed to prevent flooding collapse, causing flooding or accentuating flooding. In low-lying coastal areas storm surges may cause significant flooding. Mass movement events are similar to flooding, although the proportion of sediments to water is larger than an alluvial flood with the outcome just as disastrous.

Humans respond to flooding in a variety of ways. Broadly defined these fall into two categories, structural and nonstructural measures. Structural measures include dams and dikes. Through time the efficacy of structural features has been questioned, and there has been a shift from purely structural approaches to controlling floods to a mix of structural and nonstructural flood mitigation strategies. Nonstructural measures include flood insurance, floodplain mapping, and land-use ordinances, acquisition and relocation, floodproofing, detection and response warning systems, soil bioengineering, and combined structural and nonstructural measures to reduce flood losses. Some progress is being made in addressing the hazards associated with flooding. The reduction of flood impacts continues at great expense, but vulnerability will continue to rise as long as more people build in floodplains, increasing the risk of catastrophic floods. Even the best warnings will not eliminate the risks increasingly being taken around the globe.

READING LIST

Alexander, D. (1993). *Natural Disasters*, Chapman and Hall, New York, p. 632.

Bailey, J.F. and Patterson, J.L. *Hurricane Agnes Rainfall and Floods, June–July 1972*. USGS Professional Paper 924.

Biswas, A.K. and Chatterjee, S. (1971). Dam disasters: an assessment. *Engineering Journal* **54**: 3–8.

Blaikie P., Cannon, T., Davis, I., and Wisner, B. (1994). *At Risk Natural Hazards, Peoples Vulnerability and Disasters*. Routledge, New York, p. 282.

Bolt, B.A., Horn, W.L., Macdonald, G.A., and Scott, R.F. (1975). *Geological Hazards*. Verlag, Berlin.

Brookes, A. (1985). River channelization: traditional engineering methods, physical consequences and alternative practices. *Progress in Physical Geography* **9**: 44–73.

Brookes, A. (1988). *Channelized Rivers: Perspectives for Environmental Management*. John Wiley, Chichester.

Brown, L.R. (1984). Conserving soils. In: L.R. Brown (Ed.). *State of the World*. Norton, New York, pp. 53–75.

Burby, R.J. (1985). *Flood Plain Land Use Management: A National Assessment*. Westview Press, Boulder, Colorado.

Caine, N. (1995). Snowpack influences on geomorphic processes in Green Lakes valley, Colorado Front Range. *Geographical Journal* **161**: 55–68.

Campbell, R.H. (1975). *Soil Slips, Debris Flows and Rainstorms in the Santa Monica Mountains and Vicinity, Southern California. US Geological Survey, Professional Paper 851.*

Carson, M.A. (1976). Mass-wasting, slope development and climate. In: *Geomorphology and Climate*. E. Derbyshire (Ed.). John Wiley, Chichester, pp. 101–136.

Church, M. (1988). Floods in cold climates. In: *Flood Geomorphology*. V.R. Baker, R.C. Kochel, and P.C. Patton (Eds.). John Wiley & Sons, New York, pp. 205–229.

Clark, E.H., Haverkamp, J.A., and Chapman, W. (1985). *Eroding Soils: The Off-Farm Impacts*. The Conservation Foundation, Washington, DC.

Conrad, D. (1998). *Higher Ground: Voluntary Property Buyouts in the Nation's Floodplains, A Common Ground Solution Serving People at Risk, Taxpayers and the Environment*. National Wildlife Federation, Washington, DC.

Costa, John E. (1988). Floods from Dam Failures. In: *Flood Geomorphology*. V.R. Baker, R.C. Kochel, and P.C. Patton (Eds.). John Wiley & Sons, New York, pp. 439–463.

Ely, L.L., Enzel, Y., and Cayan, D.R. (1994). Anomalous North Pacific atmospheric circulation and large winter floods in the southwestern United States. *Journal of Climate* **7**: 977–987.

Frank, N.L. and Husain, S.A. (1971). The deadliest tropical cyclone in history? *Bulletin of American Meteorological Society* **52**(6): 438–444.

Galloway, G. *A Blueprint for Change, Sharing the Challenge: Floodplain Management into the 21st Century*, report of the Interagency Floodplain Management Review Committee to the Administration Floodplain Management Task Force, Washington, DC, June 1994.

Gray, D.H. and Sotir, R.B. (1996). *Biotechnical and Soil Bioengineering Slope Stabilization: A Practical Guide for Erosion Control*. John Wiley & Sons, New York, p. 378.

Gregory, K.J. (1995). Human activity and paleohydrology. In: *Global Continental Palaeohydrology*. K.H. Gregory, L. Starkel, and V.R. Baker (Eds.). Wiley-Interscience, Chichester, pp. 151–172.

Gruntfest, E. and Huber, C.J. (1991). Toward a comprehensive national assessment of flash flooding in the United States. *Episodes* **14**(1): 26–34.

Gruntfest, E. and Weber, M. (1998). Internet and emergency management prospects for the future. *International Journal of Mass Emergencies and Disasters* **16**(1).

Hall, A.J. (1981). *Flash Flood Forecasting*, Operational Hydrology Report No. 18, WMO, Geneva.

Hammer, T.R. (1972). Stream channel enlargement due to urbanization. *Water Resources Research* **8**: 1530–1540.

Hassan, M.A. (1990). Observations of desert flood bores. *Earth Surface Processes and Landforms* **15**: 481–485.

Hirschboeck, K.K. (1987). Catastrophic flooding and atmospheric circulation anomalies. In: *Catastrophic Flooding*. L. Mayer and D. Nash (Eds.). Unwin Hyman, London, pp. 23–56.

Hirschboeck, K.K. (1988). Flood hydroclimatology. In: *Flood Geomorphology*. V.R. Baker, R.C. Kochel, and P.C. Patton (Eds.). John Wiley & Sons, New York, pp. 27–49.

Horton, R.E. (1933). The role of infiltration in the hydrologic cycle. *Transactions of the American Geophysical Union* **14**: 446–460.

Kattelman, R. (1990). Floods in the high Sierra Nevada, California, USA. In: *Hydrology in Mountainous Regions II. Artificial Reservoirs, Water and Slopes*. R.O. Sinniger and M. Monbaron (Eds.). IAHS Publ. No. 205, pp. 311–317.

Kiersch, G.A. (1964). The Vaiont River disaster. *Civil Engineering* **34**: 32–39.

Knighton, David (1998). *Fluvial Forms and Processes: A New Perspective*. Arnold, London.

Kuhnle, R.A., Binger, R.L., Foster, G.R., and Grissinger, E.H. (1996). Effect of land use changes on sediment transport in Goodwin Creek. *Water Resources Research* **32**: 3189–3196.

Larson, L. (1996). Lessons drawn from the 1993 flood. *Forum for Applied Research and Public Policy. Fall*, No. 3, pp. 102–104.

Milne, A. (1986). *Flood Shock*. Sutton, Gloucester.

Montz, B.E. and Gruntfest, E. (1986). Changes in American urban floodplain occupancy since 1958: the experiences of nine cities. *Applied Geography* **6**: 325–338.

Moore, J.W. and D.P. Moore, (1989). *The Army Corps of Engineers and the Evolution of Federal Floodplain Management*

Policy. Institute of Behavioral Science, Boulder, CO, *special publication* No. 30.

Murty, T.S. and Neralla, V.R. (1992). On recurvature of tropical cyclones and the storm surge problem in Bangladesh. *Natural Hazards* **6**(3): 275–279.

Myers, M.F. (1996). Midwest floods channel reforms. *Forum for Applied Research and Public Policy, 11, Fall*, No. 3, pp. 88–97.

Naef, F. and Bezzola, G.R. (1990). Hydrology and morphological consequences of the 1987 flood event in the upper Reuss valley. In: *Hydrology in Mountainous Regions II. Artificial Reservoirs, Water and Slopes, IAHS Publ.* R.O. Sinniger and M. Monbaron (Eds.). No. 205, pp. 339–346.

Newson, M. (1992). *Land, Water and Development: River Basin Systems and their Sustainable Management*. Routledge, London.

Parker, D. and Fordham, M. (1996). Evaluation of flood forecasting, warning and response systems in the European Union. *Water Resources Management* **10**(4): 279–302.

Patton, Peter C. (1988). Drainage basin morphometry and floods. In: *Flood Geomorphology*. V.R. Baker, R.C. Kochel, and P.C. Patton (Eds.). John Wiley & Sons, New York, pp. 51–64.

Pearce, F. (1988). Cool oceans caused floods in Bangladesh and Sudan. *New Scientist* **8**: 31.

Rappaport, E.N. (1994). Hurricane Andrew. *Weather* **49**(2): 51–60.

Reid, I. and Frostick, L.E. (1987). Flow dynamics and suspended sediment properties in arid zone flash floods. *Hydrologic Processes* **1**: 239–253.

Rice, R.M., Corbett, E.S. and Bailey, R.G. (1969). Soil slips related vegetation, topography and soil in southern California. *Water Resources Research* **5**: 647–659.

Roberts, C.R. (1989). Flood frequency and urban-induced channel change: Some British examples. In: *Floods: Hydrological, Sedimentological and Geomorphological Implications*. K. Beven and P. Carling (Eds.). John Wiley, Chichester, pp. 57–82.

Rodda, J.C. (1970). Rainfall excesses in the United Kingdom. *Transactions of the Institute of British Geography* **49**: 49–60.

Smith, K. and Ward, R.C. (1998). *Floods: Physical Processes and Human Impacts*. John Wiley & Sons, New York.

Smith, K. (1996). *Environmental Hazards: Assessing Risk and Reducing Disaster*, 2nd Edn. Routledge, London.

Strahler, Alan N. (1964). Quantitative geomorphology of drainage basins and channel networks. In: *Handbook of Applied Hydrology*. V.T. Chow (Ed.). McGraw-Hill, New York, pp. 440–474.

Thomas, C.A. and Lamke, R.D. (1962). Floods of February 1962 in southern Idaho and northeastern Nevada. *US Geological Survey Circular* **467**: 30.

U.S. Army Corps of Engineers. (1997). *Flood Proofing Techniques, Programs and References*. Available from Corps at CECW-PF, 20 Massachusetts Avenue, NW, Washington, DC, p. 26.

U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Office of Hydrology. (1997). *Automated Local Flood Warning Systems Handbook*, Weather Service Hydrology Handbook No. 2, February, Silver Spring, MD, unpaged.

Ward, R.C. and Robinson, M. (1990). *Principles of Hydrology*, 3rd Edn. McGraw-Hill, Maidenhead.

Waylen, P.R. and Caviedes, C.N. (1987). El Nino and annual floods in coastal Peru. In: *Catastrophic Flooding*. L. Mayer and D. Nash (Eds.). Unwin Hyman, pp. 57–77.

Williams, P. (February, 1998). Inviting trouble downstream. *Civil Engineering*, pp. 50–53.

Wohl, Ellen (Ed.). (2000). *Inland Floods*. Cambridge University Press.

Wolman, M.G. (1967). A cycle of sedimentation and erosion in urban river channels. *Geografiska Annaler* **49A**: 385–395.

Wolman M.G. and Gerson, R. (1978). Relative scales of time and effectiveness of climate in watershed geomorphology. *Earth Surf. Proc.* **3**: 189–208.

Zaslavsky, D. and Sinai, G. (1981). Surface hydrology: I-Explanation of phenomena; II-Distribution of raindrops; III-Causes of lateral flow; IV-Flow in sloping layered soil; V-In-surface transient flow. *Journal of Hydraulics Division American Society of Civil Engineers* **107**(HY1): 1–93.

SELECTED WEB PAGES RELATED TO NONSTRUCTURAL MEASURES

http://www.alertsystems.org/saas/ALERT User Group

http://FEMA.gov U.S. Federal Emergency Management Agency

http://ceres.ca.ca.gov/—State of California water resources agency

http://www.usace.army.mil/inet/functions/cw/cwfpms/fpms.htm U.S. Army Corps of Engineers with emphasis on floodplain management activities

http://www.nwf.org/nwf/pubs/higherground/intro.html National Wildlife Federation site for manuscript Higher Ground

http://member.aol.com/damsafety/homepage.htm Association of State Dam Safety Officials

http://www.ci.fort-collins.co.us/csafety/oem/index.htm Comprehensive emergency preparedness homepage from Ft Collins, Colorado an excellent reference.

http://web.uccs.edu/geogenvs/work/Eve/Beyond%20Flood%20Detection%20Final.html.

www.ncdc.noaa.gov/ol/reports/mitch/mitch.html—NOAA Website about Hurricane Mitch

FLOOD CONTROL STRUCTURES

Stefano Pagliara
Pietro Chiavaccini
Universitá di Pisa
Pisa, Italy

INTRODUCTION

Flood control structures are necessary to diminish the effects of excess water in streams. They can be distinguished in structures that provoke a reduction of peak flows and in structures that divert floodwaters through flood bypasses. Sometimes levees can also be used to control flood events.

STRUCTURES ABLE TO REDUCE PEAK FLOWS

They can be divided into retarding basins and reservoir flood storage.

Retarding Basins

The purpose of a retarding basin is to prevent flooding by controlling peak flows from storms to a level that could be safely carried by the downstream system (1).

It consists of a dam placed immediately upstream from the reach to be protected, preferably in an area where a great water volume is obtained with a small dam (2). The outlet discharge is generally regulated by valves or gates. The effect of a storage reservoir can be described with the equation of continuity:

$$Q_{in}(t) - Q_{ou}(t) = \frac{dV}{dt} \tag{1}$$

where Q_{in} is the inlet discharge, Q_{ou} the outlet discharge, and V the stored volume of water. Generally, the outlet discharge is a function of the water depth in the reservoir. A schematic draw of a storage reservoir is shown in Fig. 1, whereas the effect of water discharge is illustrated in Fig. 2. It is evident by a reduction of the peak flow and a modification of the duration of the flood event.

The type of the outlet should be chosen with respect for the storage characteristics of the reservoir and the nature of the flood problems. An ungated sluiceway is more preferable than a spillway operating as a weir, even if the latter is necessary for emergency discharge of a flood exceeding the design magnitude of the outlets.

Reservoir Flood Storage

Where flood damage at a number of locations on a river can be significantly reduced by construction of one or more reservoirs, or where a reservoir site immediately upstream from one damage center provides more economical protection than local protection works, reservoir flood storages should be considered. These structures permit storage of a portion of the flood flow in a selected area, generally placed laterally to the river. The reservoir is fed with a lateral weir (Fig. 3). The release of the volume of the stored water is ensured by gates or valves and regulated in a way that the emptying happens after the flood has passed.

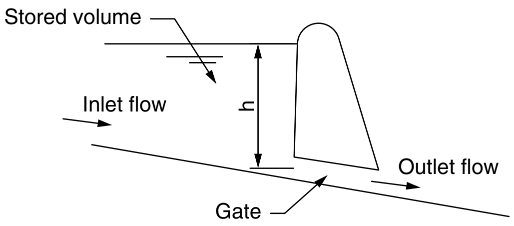

Figure 1. Schematic representation of a retarding basin.

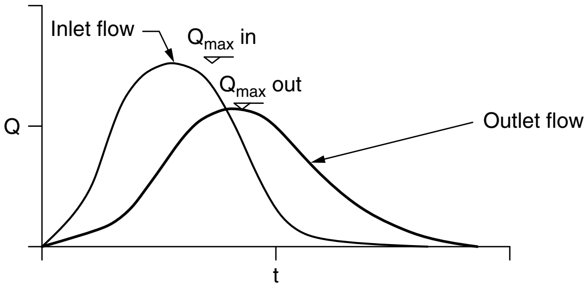

Figure 2. Effect of a retarding basin on flood discharge.

The effect of a storage reservoir is a diminution of peak flow without a modification of the duration of the flood (Fig. 4). The weir is dimensioned for the difference between the peak inflow (Q_{max} upstream) and the maximum discharge allowable downstream (Q_{max} downstream). Its height should be determined by evaluating the cost of the structure, because the taller it is, the longer it is. Besides, the height influences the point in which storage begins (point A of Fig. 4).

DIVERSION STRUCTURES

Flood bypasses ensure the diversion of floodwaters, which permits reduction of the flow passing in the original stream

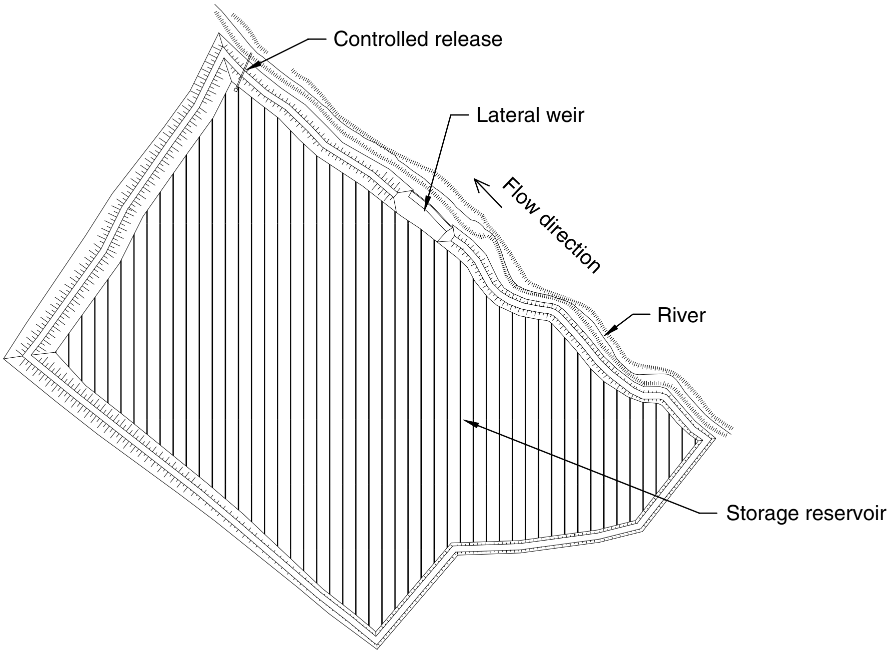

Figure 3. Storage reservoirs.

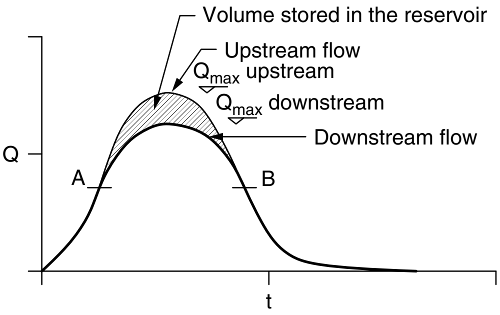

Figure 4. Effect of storage reservoir on flood.

to the allowable value (3). The diversion may return the water to the same channel at a point downstream or deliver it to another channel or different watershed. Besides, it not only conveys water in another reach, but also stores a substantial volume temporarily, thus serving as a large shallow reservoirs.

The diversion can be a second river working permanently (Fig. 5a) or if floodwater leaves the river at a controlled point, it works only during flood events (Fig. 5b). In the first case, modification to river morphology happens because there is a variation of the water tractive force in the original stream because of the diminution of the discharges. In the second case, the diversion remains dry for a long time and maintenance is necessary.

Admission of water to the bypass is achieved in different ways. In many cases, a low spot in the natural bank or a gap in the levee line is used. In some cases, a low section of the levee is created in a way that, when once overtopped, is washed out, and full discharge capacity is developed into the bypass. In other cases, a sill or a weir is used so that overflow occurs at a definite river stage. This last solution is suggested when overflow occurs frequently and the need for replacement of the levee section exists.

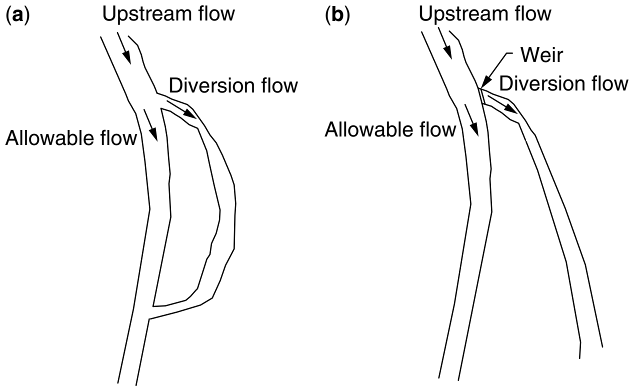

Figure 5. Diversion of flow.

LEVEES

Levees represent a barrier preventing overtopping (4). Levees are longitudinal dams erected parallel to a river and can be earth dikes or masonry construction. In the last case, they are called flood walls.

Levees are more used because their cost is low and material is easily available. They are generally built of material excavated near the river and placed in layers and compacted (Fig. 6). To increase the impermeability of the levee, a core of clay or sheetpile septa can be used. A drain ditch or tile line along the back toe of the line is advisable and the back slope should be flat enough to contain the seepage line.

The width of the top should be enough to permit the movement of maintenance equipment with a minimum of about 3 m. Bank slopes are usually very flat because the low cost of the material, and they should be protected against erosion by sodding, planting of shrubs and trees, or use of riprap. Global stability should be verified during the design of the levee.

When dimension of leaves are not compatible with available area flood walls can be used to reduce the space. They are designed to withstand the hydrostatic pressure (including uplift) exerted by the water when at flood level. Sometimes flood wall is backed by an earth fill, and in case of low discharges, it works as a retaining wall.

CONCLUSION

Flood control structures ensure a reduction of risk and damage in case of a flood event. The commonly used structures for mitigation of flood damage have been investigated, pointing out the difference and the significant characteristics to proceed in a correct design.

BIBLIOGRAPHY

1. Mays, L.W. (1999). *Hydraulic Design Handbook*. McGraw-Hill, New York.
2. Novak, P. Moffat, A.I., Nalluri, C. and Narayanan, R. (1990). *Hydraulic Structures*. Spoon Press.

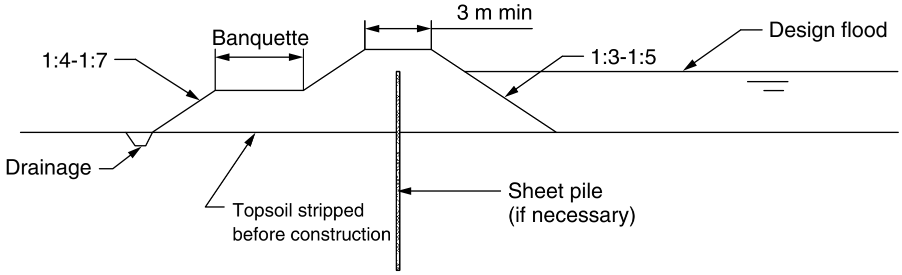

Figure 6. Typical levee cross section.

3. Linsley, R.K. Franzini, J.B., Freyberg, D.L., and Tchobanouglus, G. (1992). *Water Resources Engineering*. McGraw-Hill, New York.

4. CECW-EG. (2000). *Design and Construction of Levees*. EM-1110-2-1913.

FLOODS AS A NATURAL HAZARD

Luis Berga
ETSIn Caminos
Barcelona, Spain

The natural hazards suppose an important impact for human life and produce serious social effects and grave economic losses. Natural disasters constitute a constraint on the sustainable development, affecting its three basic mainstays: economics, social, and environmental.

In spite of the efforts made by the International Decade for Natural Hazard Reduction of the United Nations (1–3), the natural disasters in the world have experienced an increasing evolution during the last decades of the twentieth century, producing, at present, a mean of some 40,000 victims per year and mean economic losses of more than $60 billion per year. The number of major natural disasters during the period between 1950 and 1999 has been multiplied by 4.4, and by more than 15 times with relation to the economic damages (Fig. 1 and Table 1) (4,5).

The economic losses because of natural disasters are increasing with an exponential trend, having tripled during the last decade of the 1990s, with a current evaluation of more than $60 billion per year, as is shown in Fig. 2 (5).

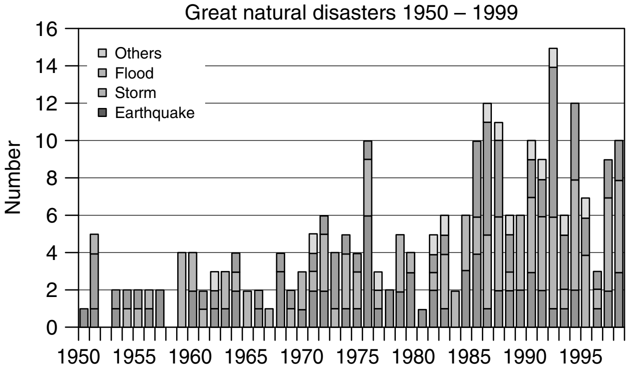

Figure 1. Number of natural disasters 1950–1999 (5).

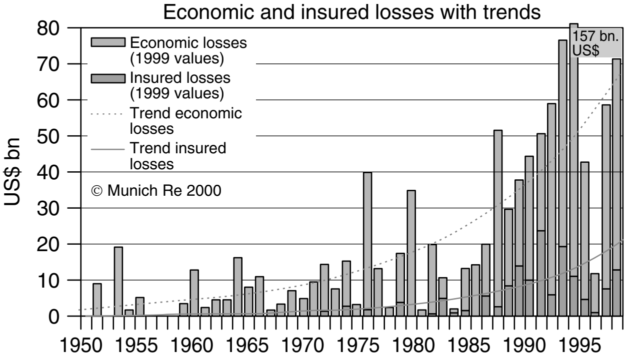

Figure 2. Natural disasters. Economic and insured losses with trends.

The greatest natural disasters of the last decades that have produced economic damages greater than $10 billion are shown in Table 2. It can be observed that of the 19 disasters, 8 are big floods, such as those of 1993 in the United States; 2002 in Central Europe; 1991, 1996, and 1998 in China (Fig. 3); those of North Korea in 1995; and those of the year 1999 in Venezuela, with more than 30,000 victims.

The floods constitute the most important disaster among the natural hazards. Floods represent about 30% of the total number of natural disasters and economic damages, and almost 25% of the fatalities produced by the natural disasters (Fig. 4). The extensive and enhanced statistical data show that, in the period 1975–2001, some 95 significant floods per year have been produced, which have caused a mean of some 11,000 fatalities per year, and have affected some 150 M people per year, which signifies that in the last decade of the twentieth century, 25% of the world population has been affected by the floods (6).

In recent decades, the impact caused by floods has been very important, and Table 3 refers to the most catastrophic floods that have occurred in the last fifteen years.

So then, experience shows that the impact caused by natural disasters and floods continues to increase progressively and, in many countries, constitutes a veritable restraint to sustainable economic development. For this, the UN decided in 1987 to create the International Decade for Natural Disaster Reduction (IDNDR) for the period 1990–2000 with the objective of reducing, by way of concerted international action, especially in the developing countries, the loss of lives, material damages, and the social and economic disorders caused by the natural

Table 1. Evolution of Number and Economic Losses of the Natural Disasters 1950–1999 (5)

Great Natural Disasters 1950–1999 Decade Comparison Losses in US$ Billion—1999 Values							
	Decade 1950–1959	Decade 1960–1969	Decade 1970–1979	Decade 1980–1989	Decade 1990–1999	Factor 80s:60S	Factor 90s:60S
Number	20	27	47	63	87	2.3	3.2
Economic Losses	39.6	71.1	127.8	198.6	608.5	2.8	8.6

Table 2. Major Natural Disasters in Relation with the Economic Damages

Country	Date	Disaster	Victims	Economic Damage (M $)
Japan	17.1.1995	Earthquake	6,348	100,000
United States	17.1.1994	Earthquake	61	44,000
China	May–Sep.1998	* Floods	3,656	30,700
United States	23–27.8.1992	Hurricane Andrew	62	26,500
China	27.6–13.8.1996	* Floods	3,048	24,000
Europe	Aug. 2002	* Floods	230	18,500
United States	27.6–15.8.1993	* Flood	50	16,000
Venezuela	13–16.12.1999	* Flash floods, landslides	30,000	15,000
Korea (North)	24.7–18.8.1995	* Floods	68	15,000
China	May–Sep.1991	* Floods	3,074	15,000
EUROPE	25.1–1.3.1990	Winter storms	230	14,800
Taiwan	20.9.1999	Earthquake	2,474	14,000
Armenia	7.12.1988	Earthquake	25,000	14,000
United States	1.4–27.6.1988	Drought		13,000
Turkey	17.8.1999	Earthquake	17,200	12,000
Italy	23.11.1980	Earthquake	2,914	11,800
China	21.6–20.9.1993	* Floods	3,300	11,000
America	20–30.9.1998	Hurricane Georges	4,000	10,000
Japan	26–28.9.1991	Typhoon Mireille (N° 19)	62	10,000

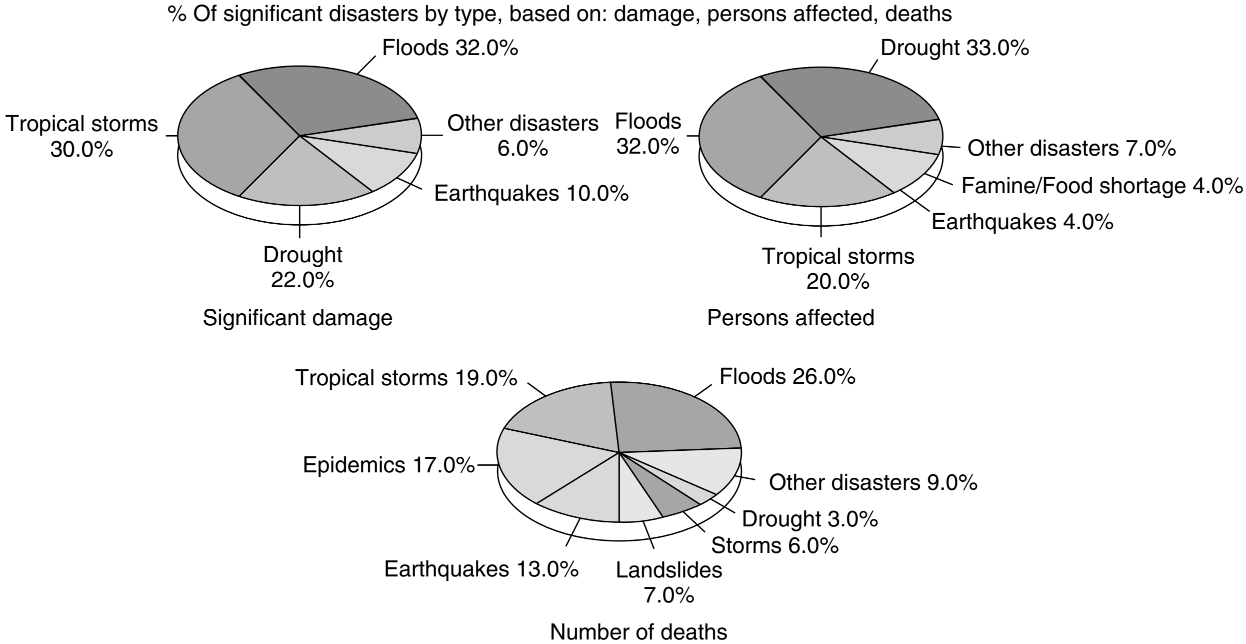

Figure 3. Major disasters around the world 1963–1992. Percentage of significant disasters by type, based on damage, persons affected, deaths.

hazards. Among the essential elements of the activities of the IDNDR, the following points stand out (3,7):

1. Greater emphasis in planification and preventive measures.
2. Adoption of integrated actions (structural and nonstructural) for the reduction of the disasters.
3. Establishment of forecasting and alarm systems compatible with the technology and culture of the countries.
4. Development of a social conscience of the necessity of the reduction of the impacts.

In order to continue the activities developed by the IDNDR and promote international cooperation, the UN has implanted the International Strategy for Disaster Reduction (ISDR), resolution of the General Assembly 54/219 of the year 2000 (8). Among the principal objectives of the ISDR, the urgent need to develop further and make use of the existing scientific and technical knowledge to reduce the vulnerability to natural disasters,

Table 3. Most Important Catastrophic Floods 1990–2002

Country	Year	Victims	Economic Losses (US M$)
China	Jul.–Aug. 1991	3,074	15,000
China	Jun.–Sep. 1993	3,300	11,000
USA	Jul.–Aug. 1993	38	15,600
Netherlands	Jan.–Feb. 1995	5	1,650
Norway	May.–Jun.1995	1	240
Korea (North)	Aug.–Sep. 1996	68	15,000
Korea (South)	July 1996	99	600
China	Jun.–Aug. 1996	3,048	24,000
China	Jul.–Aug. 1998	4,150	30,000
Bangladesh	Aug.–Sep. 1998	1,655	13,000
Central America (Hurricane Mitch)	Oct. 1998	20,000	4,000
America (Hurricane Georges)	Sep. 1998	4,000	10,000
Venezuela	Dec. 1999	20,000	15,000
Mozambique	Feb–Mar 2000	929	1,000
Central Europe	Aug. 2002	230	19,000
Korea	Sep.2002	150	800

Figure 4. China's 1998 flood.

stands out bearing in mind the particular needs of developing countries.

The ISDR has recently published a global review of disaster reduction initiatives "Living with Risk," which is "the first comprehensive effort by the United Nations system to take stock of disaster reduction initiatives throughout the world. This report discusses current disaster trends, assesses policies at mitigating the impact of disasters, and offers examples of successful initiatives. It also recommends that flood risk reduction be integrated into sustainable development at all levels—global, national and local" (6).

BIBLIOGRAPHY

1. IDNHR. (1987). *Confronting Natural Disasters*. National Academy Press, Washington, DC.
2. Housner, G.W. (1989). An international decade of natural disaster reduction: 1900–2000. *Natural Hazards*. **2**: 45–75.
3. IDNHR. (1994). *Yokohama Strategy and Plan of Action for a Safer World, World Conference on Natural Disaster Reduction*. Yokohama, Japan.
4. Zupka, D. (1988). *Economic Impact of Disasters*. Undro News.
5. Munich, R.E. (2002). *Topics. Annual Review of Natural Disasters*. Munich Re Group, Munich, Germany.
6. ISDR. (2004). *Living With Risk*. United Nations, Geneva.
7. UN. (1987). *General Assembly*. International Decade for Natural Disaster Reduction. A/Res/42/169.
8. UN. (1987). *General Assembly*. International Strategy for Disaster Reduction. A/Res/56/195.

FLOOD SOURCE MAPPING IN WATERSHEDS

BAHRAM SAGHAFIAN
Soil Conservation and
Watershed Management
Research Institute
Tehran, Iran

INTRODUCTION

Flood control projects are designed to protect people and properties against floods' adverse effects. No unique methodology exists, however, to identify flood source areas and to evaluate their contribution to the flood characteristics at any downstream point of interest. The number of nested hydrometric stations and the amount of available flood data measured simultaneously at the stations are usually insufficient for ranking flood source areas. This constraint leads us to take advantage of hydrologic simulation models.

Lumped modeling approach in flood studies precludes spatial prioritization with respect to downstream flood response. Thus, semi- or fully distributed modeling approaches are appropriate for flood control studies in large watersheds. However, the main issue is how to identify and prioritize watershed flood source areas based on their contribution to the flood response at the main outlet. The proposed methodology is also useful to determine source erosion and pollution areas.

UNIT FLOOD RESPONSE APPROACH

In this section, the approach proposed by Saghafian and Khosroshahi (1) for flood source mapping is briefly described. This approach enables identification and ranking of flood source areas within a watershed. Although the approach can be extended to a fully distributed discretization scheme, we consider delineated subwatersheds to represent homogeneous unit areas. Unit areas that produce the largest contribution to the flood peak at a given point subject to flooding (usually the outlet of the watershed draining the area of interest) are given highest rank among all source area units. One must note that ranking of subwatershed units based on their contribution to the discharge at the main outlet is not necessarily equivalent to the ranking based solely on the magnitude of peak discharge produced at the outlet of each unit. Another issue is that the nearest or largest subwatershed unit may not cause the largest impact on the outlet flood peak. These issues are studied later when a case study is presented. Particular emphasis will be put on the routing in the stream network that connects subwatershed areas because storage is a key factor in how flood source areas influence the outlet flood response. It is expected that the effect of stream routing is more pronounced for more frequent floods.

The flood source mapping procedure starts with collection of watershed data and historic rainfall-runoff data recorded at rain and stream gauges located inside or outside the watershed. A semidistributed rainfall-runoff model with stream routing capability is selected, and the model is properly calibrated and validated against the flood data of at least one hydrometric station located near the outlet of the watershed. Then the model is setup for the watershed of interest, if different from the watersheds bounded at hydrometric stations for which the model was calibrated.

Subwatersheds' flood hydrographs are simulated for a selected design storm, and flood routing is performed in the channel network connecting the outlets of subwatersheds to the watershed's main outlet. The routing results in a base outlet hydrograph. Then, in successive runs, each subwatershed unit is singly removed, and the resultant hydrographs at the main outlet are simulated (Fig. 1). The changes in flood peak of the generated hydrographs relative to the peak of the base hydrograph are computed. The procedure may be repeated for other design storms, and the sensitivity of the results could be examined.

We may now define a flood index to measure unit area contribution to the flood peak at the main watershed outlet:

$$FI_k = \frac{Q_{p,all} - Q_{p,all-k}}{Q_{p,all}} \times 100 \qquad (1)$$

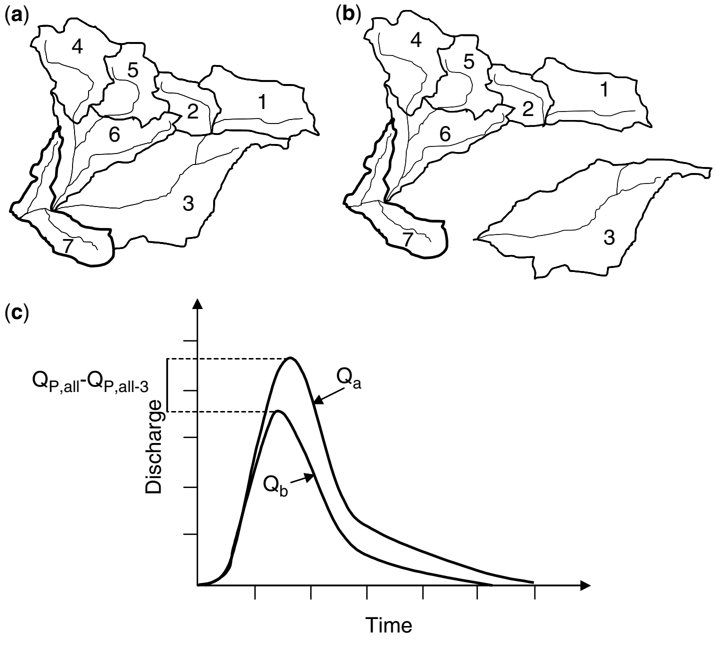

Figure 1. Unit flood response application on a watershed: (**a**) subwatershed units, (**b**) unit 3 removed, and (**c**) simulated flood hydrographs corresponding to cases (**a**) and (**b**).

where *FI* is the gross flood index of the *k*th subwatershed (in percent), $Q_{p,all}$ is the peak discharge of the base hydrograph (in m^3/s), $Q_{p,all-k}$ is the peak outlet discharge when the *k*th subwatershed is removed (in m^3/s), and A_k is area of the *k*th subwatershed (in km^2).

A different flood index can also be expressed in terms of contribution per unit area of subwatersheds as follows:

$$fi_k = \frac{Q_{p,all} - Q_{p,all-k}}{A_k} \tag{2}$$

where *fi* is the flood index of the *k*th subwatershed based on unit subwatershed area (in $m^3/s/km^2$).

The approach outlined above is entitled "unit flood response," abbreviated by UFR, based on its similarity to the unit response (UR) approach in groundwater management. In the UR approach, a groundwater model simulates the influence of a unit change in sink/source rates at selected well locations on design variables, such as hydraulic head or velocity, at specified locations (2). Suitable groundwater quantity/quality management scenarios may be studied based on the UR application. Although nonlinearity can be preserved by the model used in the unit flood response approach, the UR in groundwater studies assumes linearity by superimposing the individual unit responses at given locations.

A CASE STUDY

Damavand watershed with an area of 758 km^2 lies at 51°,46′,40″ to 52°,12′,05″ E longitude and 35°,32′,48″ to 35°,51′,39″ N latitude. High and low elevations in this relatively steep watershed are 4003 m and 1250 m above sea level, respectively. Rangeland, mostly in poor condition because of overgrazing, is the dominant vegetation cover. Agricultural areas are concentrated in the middle and outlet parts of the watershed. Average annual precipitation is reported as 443 mm. Although most precipitation occurs in the winter and early spring, severe floods in the region are often caused by high-intensity summer storms. The region is generally categorized as a flood prone area. As a result of limitations imposed by gradual availability of funds, one of the primary aims of flood control studies in this watershed is to identify and rank areas with higher impact on the outlet flood peak.

Boundary of Damavand watershed and its seven subwatersheds along with the contour map were digitized in a geographic information system (GIS). Maps of digital elevation model (DEM) and land slope were generated by the GIS. Figure 2 shows subwatershed boundaries on the DEM background. Table 1 summarizes some of the physiographic characteristics of the subwatersheds. A land-use map was produced through processing a Landsat satellite image of the area and field inspection. SCS curve number layer was generated by overlaying maps of land use and hydrologic soil groups (Fig. 3). Then, average subwatersheds' CN values were determined (see Table 1).

HEC-HMS hydrologic model (3) is applied in this case study. Specifically, SCS unit hydrograph method and Muskingum technique were applied to simulate rainfall-runoff transformation at subwatershed scale and stream routing, respectively. Rainfall intensity corresponding to the design storm of 50-year return period with duration equal to the watershed concentration time was drawn from the intensity-duration-frequency (IDF) curve of a local station. Average rainfall temporal pattern was developed at the same station, but no particular spatial pattern was considered in this case study.

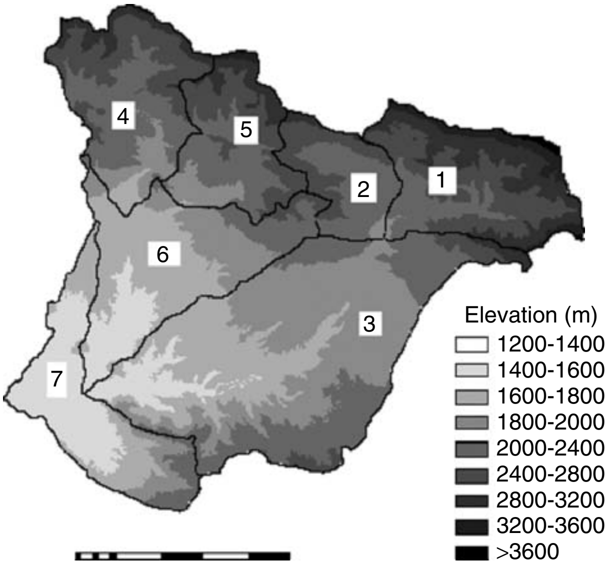

Figure 2. Subwatersheds of Damavand watershed on digital elevation model background.

Table 1. Damavand Subwatersheds Characteristics

Subwatershed Unit [1]	Area (km^2) [2]	Land Slope (%) [3]	Max Stream Length (km) [4]	Distance to Outlet (km) [5]	Stream Slope (%) [6]	Curve Number II [7]
1	97	47	16.8	32.1	10.5	81
2	46	27	12.3	32.1	5.3	71
3	253	19	36.2	4.8	5.1	78
4	96	37	19.3	18.6	8.2	80
5	70	35	14.8	20.7	6.7	80
6	112	12	24.9	4.8	6.7	78
7	84	13	16.4	0	6.0	86

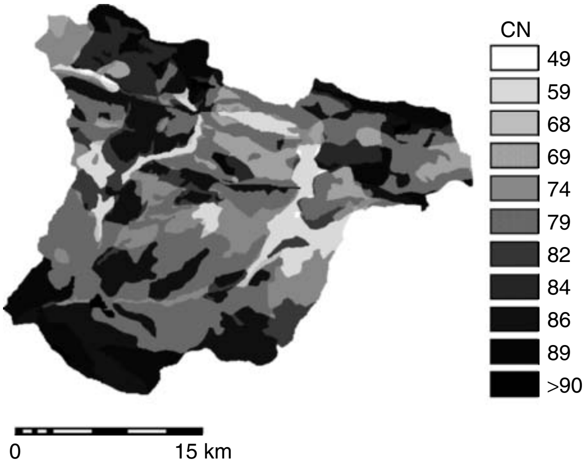

Figure 3. Map of curve number.

FLOOD INDEX ANALYSIS

HEC-HMS model was calibrated and validated for a number of rainfall-runoff events recorded at the outlet of the watershed. Calibration parameters were SCS initial loss and lag time. The 50-yr, 12-hr storm was simulated with peak discharge at the outlet of each subwatershed shown in Column 2 of Table 2. A peak discharge of 371.9 m^3/s was produced at the watershed outlet. The simulated outlet peak discharges after removing each subwatershed are reported in Column 4 of Table 2. For example, the peaks corresponding to cases with and without subwatershed 3 are 371.9 m^3/s and 270.0 m^3/s, respectively. Therefore, the peak is reduced by 101.9 m^3/s when subwatershed 3 is completely removed. Subwatershed 3's *FI* and *fi* flood indices are equal to $101.9/371.9 = 27.4\%$ (Column 5, Table 2) and $101.9/253 = 0.40$ $m^3/s/km^2$ (Column 7, Table 2), respectively. The priorities based on the absolute peak discharge contributions (*FI*) and per unit subwatershed area peak discharge contributions (*fi*) are respectively listed in Columns 6 and 8 of Table 2.

The overall mixed ranking of subwatersheds in Columns 3, 6, and 8 in Table 2 highlights the complex interdependent effect of subwatershed area, subwatershed characteristics, channel routing, and the location of flood source areas within a relatively large watershed. It is noted that the nearest (to the outlet) or largest and the most distant or smallest subwatersheds do not necessarily generate the highest and lowest contribution to the flood peak at the outlet, respectively. Similarly, subwatersheds producing the highest and lowest discharges at their own outlet may not rank first and last in *FI* or *fi*. Subwatersheds' absolute and specific discharges are compared in (Fig. 4), whereas *FI*'s and *fi*'s are shown in (Fig. 5).

One can observe, from (Fig. 5), that the rankings of absolute discharge contribution (*FI*) generally differ from those of per unit area discharge contribution (*fi*). Subwatersheds 5 and 4 rank first and second in the unit area discharge contribution, whereas subwatershed 2 attains the lowest rank. Subwatershed 3, with the highest ranking in the absolute discharge contribution (*FI*), ranks

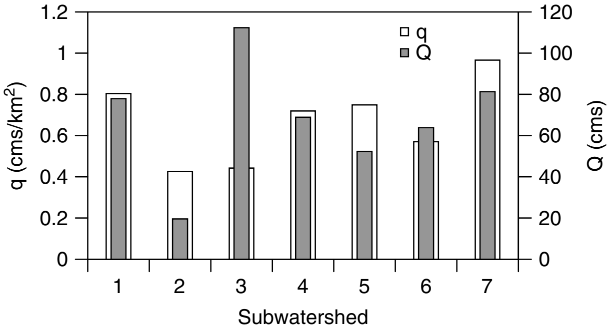

Figure 4. Comparison of absolute and specific subwatershed peak discharges.

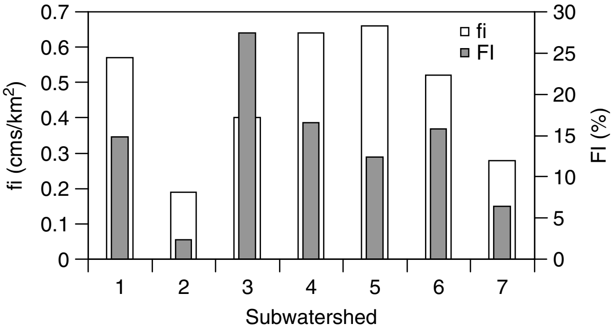

Figure 5. Comparison of subwatersheds' *FI* and *fi* flood indices.

Table 2. Simulated 50-yr Peak Discharges and the Results of UFR Application

Subwatershed [1]	Subwatershed Peak Discharge (m^3/s) [2]	Priority Based on Subwatershed Peak Discharge [3]	Outlet Peak Discharge Without Subwatershed (m^3/s) [4]	Flood Index *FI* (%) [5]	Priority Based on *FI* [6]	Flood Index *fi* (cms/km^2) [7]	Priority Based on *fi* [8]
1	77.9	3	316.8	14.8	4	0.57	3
2	19.6	7	363.0	2.4	7	0.19	7
3	112.2	1	270.0	27.4	1	0.40	5
4	69.1	4	310.0	16.6	2	0.64	2
5	52.3	6	325.8	12.4	5	0.66	1
6	63.7	5	313.3	15.8	3	0.52	4
7	81.3	2	348.1	6.4	6	0.28	6

Table 3. Flood Index (*fi*) Matrix for Pairs of Subwatershed.

Subwatershed	1	2	3	4	5	6	7
1	0.568	0.439	0.426	0.593	0.596	0.558	0.472
2	0.439	0.193	0.366	0.499	0.474	0.436	0.273
3	0.426	0.366	0.403	0.456	0.448	0.450	0.396
4	0.593	0.499	0.456	0.645	0.649	0.585	0.481
5	0.596	0.474	0.448	0.649	0.659	0.580	0.456
6	0.558	0.436	0.450	0.585	0.580	0.523	0.399
7	0.472	0.273	0.396	0.481	0.456	0.399	0.283

fifth in the unit area contribution (*fi*). Use of flood index (*fi*) is recommended in the analysis of spatial flood source prioritization required for optimum planning flood control measures at subwatershed scale.

The effect of spatial aggregation of unit source areas could also be examined by removing a pair of subwatersheds at each simulation run. The flood index matrix is shown in Table 3. The *fi* values in nondiagonal cells, say jth row and kth column, correspond to removing the jth and kth subwatershed simultaneously. When $j = k$ in diagonal cells, the results correspond to the case when a single subwatershed j has been removed (as already reported in Table 2). Clearly, the combination of subwatersheds 4 and 5 accumulated the highest impact on outlet flood peak. Conversely, the combination of subwatersheds 7 and 2 indicates the lowest impact.

SUMMARY

A technique, abbreviated by UFR, was described in this chapter for mapping flood source areas within a watershed. Subsequent elimination of subwatershed units in the rainfall-runoff modeling process allows separation of the effect of individual watershed area. Thus, contribution of each subwatershed unit may be quantified and its priority determined. The technique can be effectively applied in flood source area mapping at the desired spatial scale.

The UFR technique was applied in a case study, and the map of a 50-yr flood index was prepared (Fig. 6). Based on the results of this case study, it may be concluded that because of the complex nature of flood hydrograph propagation within the stream network, use of the UFR approach can be considered as the optimal way for prioritization of flood source areas.

Another issue is that implementing flood abatement measures in subwatersheds with low flood index may actually result in delaying peak arrival of such areas and could cause a more synchronized response with other subwatersheds, which is the case with subwatershed 7 in the presented case study.

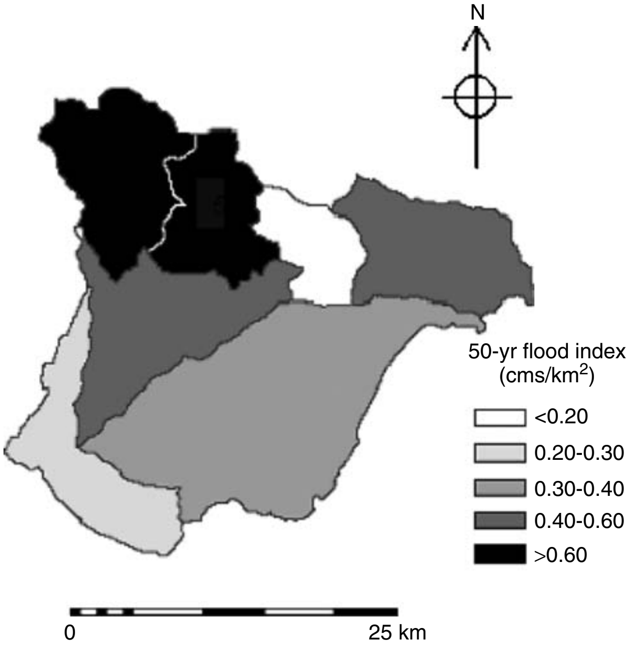

Figure 6. Map of 50-yr flood source units.

BIBLIOGRAPHY

1. Saghafian, B. and Khosroshahi, M. (in press). Unit response approach for priority determination of flood active areas. *J Hydrologic Engineer ASCE*.
2. Gorelick, S.M. (1983). A review of distributed parameter groundwater management modeling methods. *Water Resources Res.* **19**(2): 305–319.
3. Hydrologic Engineering Center. (2000). *Hydrologic Modeling System HEC-HMS: Technical Reference Manual*. US Army Corps of Engineers, Davis, CA.

URBAN FLOODING

Ole Mark
Asian Institute of Technology
Hørsholm, Denmark

Flooding in urban areas is an inevitable problem in many cities, and it causes huge costs to the society in structural and nonstructural damage. The problems that arize from urban flooding range from minor ones such as water getting into the basement of a few houses, to major incidents, where large parts of cities are inundated for several days. Most modern cities in the industrialized part of the world usually experience smaller scale local problems due mainly to insufficient capacity in their sewer systems during heavy rainstorms. Cities in other regions, including those in South/Southeast Asia, often have more severe problems because of much heavier local rainfall and lower drainage standards. This situation is continuously getting worse because many cities in the developing countries are growing rapidly these days but do not have the necessary funds to extend and rehabilitate their existing drainage systems.

CAUSES OF URBAN FLOODING

Urban flooding may be caused by a river or by local rainfall over a city. *River-induced flooding* is flooding of a city due to high water levels in rivers adjacent to the city. River water flows to lower lying areas, and floods them. *Rainfall-induced flooding* is flooding of a city due to local rainfall in the built-up areas of the city. Examples of rainfall-induced flooding in Bangkok can be seen in Figs. 2 and 3. This may occur several times a year on various scales. Mainly, inadequate existing drainage paths and their improper operation and maintenance cause these floods.

An interaction often exists between the drainage capacity of an urban drainage system and the water level in a recipient. For example, Bangkok was flooded from local rainfall during October 2002 due to reduced drainage capacity from the city to the river. The condition in the river was caused by a combination of very high runoff from the upstream river basin and a spring tide in the Gulf of Thailand, resulting in a very high water level in the river and hence drastically reduced drainage capacity from Bangkok to the river. No drainage by gravity could take place, and only drainage to the river by pumping was feasible.

Figure 1. Flooding in a suburb of Dhaka City, 1996. (Photo by Terry Van Kalken).

Figure 2. Flooding in Bangkok during a minor premonsoon rain in April 2002. (Photo by Dr. Ole Mark).

Figure 3. Flooding in Bangkok, 2002. (Photo by Mr. Ebbe Worm).

Individual streets in a city may flood due to numerous causes (1). Frequently, rainfall runoff starts as an overland flow on the street before entering the underground pipe system through catch pits. The water in the pipe system is then conveyed downstream but may return to the surface if the capacity of the pipes is exceeded. The duration of flooding on the street depends on the ground elevation, the intake capacity of the catch pits, and the capacity of the underground pipes. If the intake capacity of the catch pits is limited, large volumes of runoff will be transported along the surface during large rainstorms, even if there is sufficient capacity in the underground sewer network. Insufficient drainage capacity locally in the sewer system may then, in turn, cause surface flooding, even if the catch pit intake capacity is adequate.

THE EXTENT OF URBAN FLOODING

In Asian cities subject to regular flooding, people seem to have adjusted their daily lives to smaller flooding. However, major incidents have grave consequences. Jakarta, Indonesia, was flooded from an overnight rain during February 2002. During that flood, five people were killed and approximate 200,000 people were forced from their homes. Another urban flood occurred in Bangkok, Thailand, in 1983, which inundated the city for 6 months. The flood caused damage of approximately US$149 million. In 2000, an urban flood in Mumbai, India, claimed at least 18 lives, made hundreds of slum dwellers homeless, and plunged large areas of the city into darkness. In Dhaka City, Bangladesh, moderate rain can inundate parts of the city for several days. The situation was highlighted in

September 1996, when the residents of Dhaka experienced ankle to knee-deep water on the street (Fig. 1). Heavy traffic jams occurred because of stagnant water on the streets, and daily activities in parts of the city were almost paralyzed. At present, other major cities in Southeast Asia facing problems from regular urban flooding are Ho Chi Minh City, Hanoi, Vietnam, and Phnom Penh, Cambodia.

Examples of places outside South/Southeast Asia that suffer from urban flooding are Fukuoka and Tokyo, Japan (2); Playa del Gandia, Spain (3); and Houston, Texas, U.S.—where the storm "Allison" in June 2001 caused urban flood damage of the order of $2 billion at the Texas Medical Center in the Harris Gully watershed (4).

THE IMPACT OF URBAN FLOODING ON SOCIETY

The costs of urban flooding are often very difficult to estimate precisely. Urban flood damage may be divided into three groups (5):

- direct damage—typically material damage caused by water or flowing water
- indirect damage—traffic disruptions, administrative and labor costs, production losses, spreading of diseases
- social consequences—negative long-term psychological effects, due to decrease of property values in frequently flooded areas and slowed down economic development.

One way to quantify the cost of urban flooding from structural damage is to collect information about damage costs from insurance companies, as has been done in Norway (5). This process is not likely to be applicable to developing countries, where few families in flood-prone areas have insurance.

During urban floods in developing countries, it is often the poorer people who suffer the most, as they often live in areas of cheap or free land, which are more exposed to natural disasters and have no or only primitive drainage and a drinking water infrastructure. Open drains of insufficient capacity in such areas are often flooded causing material damage and danger of epidemics. The water depth in some inundated city areas may typically be up to the order of 40 to 70 cm, which creates considerable traffic congestion, infrastructure problems, and huge economic loss in production, as well as significant damage to property and goods.

Diseases spread when sewage is mixed with surface flood water which, in turn, may reach and contaminate local drinking water sources. This unfortunate interaction between drinking water, groundwater, and floodwater causes severe health problems and calls for an integrated water management policy. Examples of diseases are diarrhea or leptospirosis, which are spread by bacteria in the urine of rats. In the September 2000 flood in northeast Thailand, 6,921 cases of leptospirosis were reported; 244 of these resulted in loss of human life (6). Last, but not least, parasites seems to thrive when urban flooding occurs regularly. Moist soil provides a good environment for worm eggs to flourish, and flooded open drains spread eggs to new victims (1). Today, anthelmintica is available to kill parasites, but the parasites may gradually develop resistance to the drug and create new and more severe problems. The best way to manage parasite problems is to interrupt the life cycle of parasites, that is, to remove their natural environment by reducing the frequency and duration of flooding. For example, Moraes (7) found that reduced flooding decreased the prevalence of roundworm and hookworm by a factor of 2 and hookworm alone by a factor of 3.

WHAT CAN BE DONE TO UNDERSTAND AND REDUCE URBAN FLOODING?

The water-related infrastructure in cities consists of drinking water treatment facilities, water supply networks, sewer systems, and wastewater treatment plants. Many cities are old, and they have developed according to varying historical needs and visions. The sewer and drainage systems play the key role in efficient removal of storm water after rainstorms and prevention of urban flooding and its consequences. These systems, therefore, represent an important part of the urban infrastructure which was put in place through more or less continuous investments during the period of modern city development. This means that the layout and design of the infrastructure have gradually developed into rather complex systems. Inadequate planning and design as well as bad maintenance of these systems may cause severe economic and environmental damage in the cities, as many urban flooding cases prove. The problem is often aggravated if rainstorms coincide with high water levels in the sea or adjacent rivers during the monsoon season.

Many cities in the Western world manage local and minor flooding problems by using computer-based solutions. This involves building computer models of the drainage/sewer system, for instance, by using software such as MOUSE (8) and SWMM (9). Both of these software packages have been applied successfully to modeling major urban flooding events. The models are used to understand the often rather complex interaction between rainfall and flooding. Once the existing conditions have been analyzed and understood, alleviation schemes can be evaluated and the optimal scheme implemented.

At present, there are few studies on urban flooding that deal with both the conditions in the surcharged pipe network and extensive flooding on the catchment surface. However, it is feasible to model urban flooding based on the interaction between the pipe system and surface flooding, and this raizes new possibilities for managing urban flooding problems. Examples of this kind of modeling are Bangkok, Thailand (10); Dhaka City, Bangladesh (11); Fukuoka and Tokyo, Japan (2); Harris Gully, Texas, U.S. (4) and Playa del Gandia, Spain (3).

To model urban flooding, a hydrodynamic model must be applied, and the flow and water levels in the major system (the street network and surface storage between the houses) and the minor system (the pipe network) must be described accurately. The major and minor systems

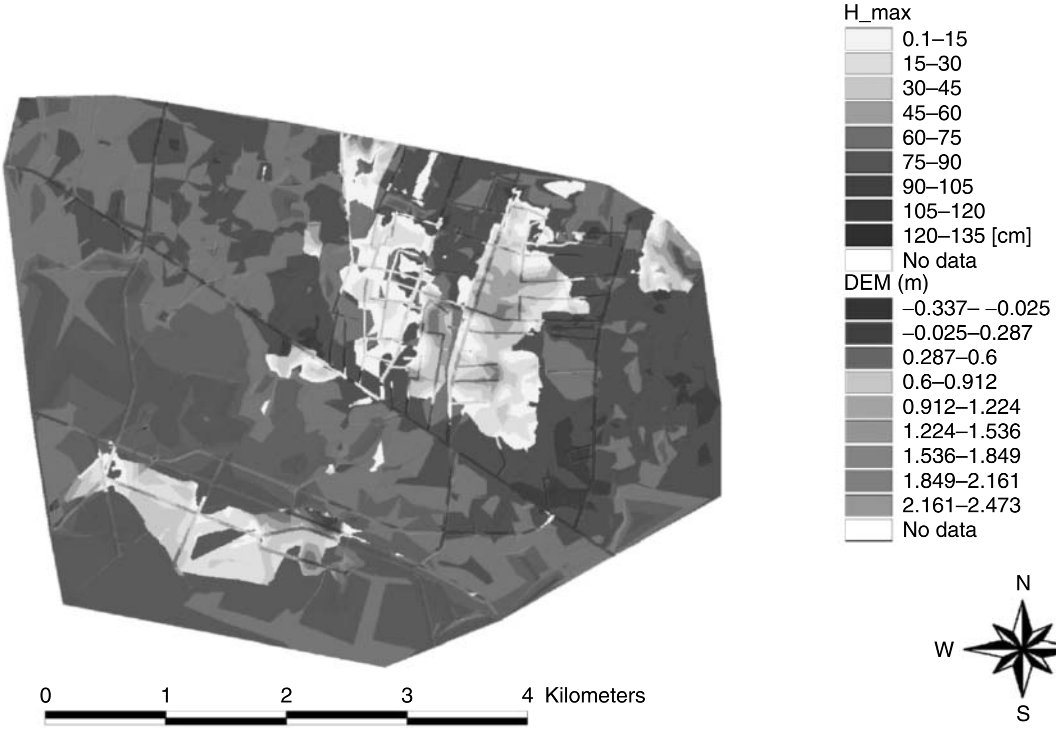

Figure 4. A flood inundation map for Bangkok (10).

route the rainfall runoff simultaneously in the pipes and on the streets. An exchange of water from the pipes to the streets (and vice versa) takes place, depending on the local hydraulic conditions in the catchment.

GIS plays a key role in data management for modeling urban flooding. A digital elevation model (DEM) is a powerful GIS tool for defining the catchment delineation and the storage of water on the surface. The DEM serves the dual purpose of input generation and result presentation of model simulations in terms of water depths in streets and between houses. An example of a flood inundation map for Bangkok (10) is shown in Fig. 4.

The handful of successful urban flood modeling applications carried out by computer models, which have traditionally been used to address local flooding problems in the industrialized part of the world, is very encouraging. The methodology is generic and can easily be transferred to other cities, if they have data available on the sewers and streets, as well as topographical maps.

FUTURE PERSPECTIVES AND CHALLENGES IN RELATION TO URBAN FLOODING

In 1997, the European Union (EU) passed a new and common set of directives, EN 752, for all EU member countries. This new standard calls for a new approach to the design of sewer systems in Europe. The old design criterion—design based on the return period of the rainfall—is replaced by a criterion based on the return interval of flooding. This new directive has boosted interest in understanding, research, and subsequent alleviation of urban flooding. Hence, it is foreseen that many urban flood applications will appear in the near future.

In the future, problems related to urban flooding will expand, as the cities in the developing countries are growing rapidly. So urban flooding must be seen in a broader perspective. An example is Dhaka City, which relies heavily (up to 97%) on groundwater for its water supply. During the last 25 years, the groundwater table has dropped by about 25 m (12). If this drop in the water table continues, it may generate problems for the city's water supply, and surface waters may need to be considered as additional resources. This is complicated by the growth of Dhaka City—areas that used to have permeable soils are being transformed into hard, impermeable surfaces. Such impermeable surfaces prevent replenishing groundwater storage and further aggravate groundwater problems. In turn, the runoff from the new impermeable surfaces generates additional surface runoff, which again increases the flooding in Dhaka City. Water supply pipes may be under low pressure during a period of flooding (e.g., due to power cuts), and polluted floodwater may enter the water supply network.

This poses an additional health risk to the population on top of the diseases spread by floodwater. In addition, there are considerable losses from the water supply network, so quite a large amount of drinking water is lost. This means that when dealing with urban flooding, one must remember that it involves not one problem, but a group of strongly interrelated problems, which have major impacts on the quality of life of people who live in flood-prone cities.

BIBLIOGRAPHY

1. Kolsky, P. (1998). *Storm Drainage*. Intermediate Technology Publications, London.

2. Ishikawa, Y., Morikawa, H., and Ilda, M. (2002). *Flood Forecasting and Management of Underground Spaces*. 9th International Conference on Urban Drainage, Portland, OR, September 8–13.
3. Tomicic, B., Mark, O., and Kronborg, P. (1999). *Urban Flooding Modelling Study at Playa de Gandia*. 3rd DHI Software Conference, June 7–9.
4. Holder, H.W., Stewart, E.J., and Bedient, P.B. (2002). *Modelling an Urban Drainage System with Large Tailwater Effects under Extreme Rainfall Conditions*. 9th International Conference on Urban Drainage, Portland, OR, September 8–13.
5. König, A., Sægrov, S., and Schilling, W. (2002). *Damage Assessment for Urban Flooding*. 9th International Conference on Urban Drainage, Portland, OR, September 8–13.
6. *Bangkok Post*. (2000) September 20, 2000.
7. Moraes, L.R.S. (1996). *Health Impacts of Drainage and Sewerage in Poor Urban Areas in Salvador, Brazil*. PhD Thesis, London School of Hygiene and Tropical Medicine, London.
8. Lindberg, S., Nielsen, J.B., and Carr, R. (1989). *An Integrated PC-Modelling System for Hydraulic Analysis of Drainage Systems*. The first Australian Conference on technical computing in the water industry: Watercomp '89, Melbourne, Australia.
9. Huber, W.C. and Dickinson, R.E. (1988). *Storm Water Management Model—SWMM, Version 4 User's Manual*. US Environmental Protection Agency, Athens, GA.
10. Boonya-Aroonnet, S., Weesakul, S., and Mark, O. (2002). *Modelling of Urban Flooding in Bangkok*. 9th International Conference on Urban Drainage, Portland, OR, September 8–13.
11. Mark, O., Apirumanekul, C., Kamal, M., and Praydal, P. (2001). *Modelling of Urban Flooding in Dhaka City*. UDM May 20–24, Orlando, FL.
12. Khan, H.R. and Siddique, Q.I. (1999). Urban water management problems in developing countries with particular reference to Bangladesh. *Int. J. Water Resour. Dev.* Special issue. ISSN 0790-0627.

FLOODWATER SPREADING

Bahram Saghafian
Soil Conservation and
Watershed Management
Research Institute
Tehran, Iran

INTRODUCTION

Providing usable water resources, especially in arid and semiarid regions, is going to be one of the major challenges of the twenty-first century. Therefore, alternatives in water resources development must be studied thoroughly as population growth increases the pressure on renewable and nonrenewable water resources. Most of the precipitation and surface runoff in areas subjected to water scarcity usually occur in seasons that do not coincide with periods of high water demand. On the other hand, growth in drinking and irrigation water demand has caused overuse of more reliable resources, such as groundwater, and has degraded their quality.

Flood spreading is a practical water and soil conservation technique that, if designed and maintained properly, can address one or a combination of water-related issues on local scales. Such issues may involve reduction of flood damage downstream by diverting a portion of the flood volume, controlling sheet and gully erosion over the land by maintaining vegetation cover, storing water in the soil profile or in underground aquifers for future use, improving soil fertility through sediment deposition, leaching saline soils, and irrigating cultivated and cropped lands where the rainy period and the growing period nearly coincide. The suitability and success of a flood spreading system depends on the project objective(s), nature of the soil, amount and temporal distribution of precipitation, land topography, and water demand.

Although flood spreading has been practiced widely with diverse objectives, the aim of this article is to present an overview of flood spreading schemes designed mainly for groundwater recharge and/or flood irrigation. "Water spreading" and "flood spreading" are used interchangeably in most literature, and we focus mainly on systems designed to operate on spate floodwater.

BACKGROUND

According to the former Soil Conservation Service (USDA, undated): "Water spreading is a specialized form of surface irrigation accomplished by diverting flood runoff from natural channels or watercourses and spreading the flow over relatively level areas. The diversion and spreading is controlled by a system of dams, structures, dikes or ditches, or a combination of these, designed to accommodate a calculated rate and volume of flow." Bennett (1) believes that water spreading has been limited largely to arid and semiarid areas used for range rehabilitation and crop production. He considers flood spreading a simpler and less expensive irrigation method than indirect irrigation with the aid of water stored in reservoirs.

Stoddart et al. (2) reported that spreading water schemes helped to increase forage production in the U.S. Great Plains and Southwest. The schemes consist of dams and dikes to divert and convey water to the land where it is absorbed. Heady (3) considers water spreading systems for spreading floodwaters to irrigate land, reduce and store sediment, or store water underground. For irrigation, a 1:20–30 ratio of spreader land to watershed area was suggested based on the work in Israel on the basis of 15% runoff, an average annual rainfall of 100 mm, and a 400-mm crop water demand. The rainfall threshold to generate runoff is increased for larger watersheds, thus, favoring small spreader systems. But the sediment per square area of watershed tends to increase for smaller watersheds. Heady (3) indicates that many such systems have failed due to inadequate design and the lack of maintenance.

Gupta et al. (1995) studied different water harvesting and conservation techniques. Their findings indicate that the benefit of water harvesting is high and shows a considerable increase in total biomass, tree height, root mass, and water use efficiency compared with the control plot.

HISTORY

The origin of flood spreading systems is not transparent in the literature. Bennett (1) states that floodwater use for crop production probably originated in arid regions where river overflow was used to water crops planted in adjacent lands. Writings and remains of irrigation dikes show that western South America, Mexico, and southwestern United States were some of the places where floodwater was used for farming.

Hudson (5) provides an overview of the traditional use of runoff in areas where rainfall amount does not meet crop demand. In parts of Arizona where the annual rainfall is 300 to 400 mm, Indian tribes used runoff from sandstone outcrops to water crops such as maize, squash, and melons planted at the base of hills (6). The Hopi tribe also applied flood farming in southwest North America and still continues this practice in smaller areas (7).

Kovda (8) reported the use of natural runoff in arid areas of the former Soviet Union to crop on flood terraces of large rivers. The technique is called "kair" farming. "Khaki" farming practiced in Turkmenistan allows runoff from mountain slopes to inundate the plains. After the plains have dried out, the land is cultivated and planted. Carr (9) reported similar evidence in Ethiopia. In a mainly pastoral community in Kenya where annual rainfall is less than 200 mm, small patches of sorghum are grown using a variety of runoff farming (10). Farming on valley bottoms is an ancient practice in Tunisia and the Negev desert (5), where annual rainfall is low, and in the level pans of one to three hectares formed in broad valleys of gentle slopes in Colorado, United States, where rainfall is 400 mm (11).

Runoff farming on level fields surrounded by earth banks has been carried out in the Khost plain, Paktia province in Afghanistan, where the runoff from stony hills is collected and led down to the fields. In two experimental small-scale projects in Kenya, water from a catchment area is diverted into the cultivated area, partitioned by earth bunds, into gently sloped terraces to prevent erosion (5). Stone spillways are built into the bunds to pass water to lower terraces. Construction of large-scale inundation spreading schemes, called "ahars," have been reported in semiarid regions of India, namely, in Bihar and Uttar Pradesh. Runoff during monsoon season is retained on very gentle slopes by low earth bunds such that the moisture stored can support a 5-month crop grown following the end of monsoon. Weirs are built to allow excess storm water to pass. An advantage of this system is regular leaching of saline soil in semiarid areas. Other forms of inundation farming are also reported in the Jaisalmer district of Rajasthan (12), where an earth bund is constructed across the valley plain to store water from surrounding hills. The catchment ratio is at least 15:1 where annual rainfall is 165 mm.

In India and Pakistan, a canal that has an elevated bed compared to the river, may be excavated in the river bank with a flatter slope than that of the river (5). The canal is filled by the flood overflow and runs dry during low flow periods. Maintenance of the canal entry is necessary for removal of deposited sediment. The water carried by the canal is then distributed over irrigation fields.

Implementation of flood farming practices, both in large- and small-scale experiments, have been also reported in some parts of Sudan, Morocco, Syria, Yemen, Australia, Jordan, Brazil, and Iran. The FAO (13) reports that the irrigated area under floodwater harvesting (spate irrigation) totals about 2 million hectares in North Africa and the Middle East.

SUITABILITY CRITERIA FOR FLOOD IRRIGATION

Bennett (1) numbers a few criteria for the suitability of land in flood irrigation projects. The criteria include rich fertile soil that has favorable permeability, gentle and smooth slopes, and sufficient rainfall. Annual rainfall of less than 8 inches (200 mm), or rainfall less than 4 to 5 inches during the growing season, may preclude installation of flood spreading systems. Range forage or supplemental feed may be produced by using flood spreading schemes where rainfall ranges from 8 to 14 inches (200 to about 350 mm).

The moisture stored in the soil must be sufficient to supply the crops until the next rainy season. Therefore, soils of high storage capacity, ideally medium to moderately fine texture of sufficient depth are preferable. Coarsely texture soil is not suitable because of its high drainage rate and low water holding capacity. Prinz (14) suggests a range of annual rainfall between 150 and 600 mm and a catchment to spreading area ratio between 100:1 and 10,000:1 and more for flood irrigation. Land slopes between 1 and 2 percent are suitable, and the water should not carry a high sediment load.

FLOOD SPREADING FOR ARTIFICIAL RECHARGE

Storing water underground has the benefits of no or slight evaporation loss, large storage capacity and no need to construct structures, and less vulnerability to drought and contamination. Besides augmenting groundwater storage, one of the primary objectives of artificial recharge may be to conserve and dispose of floodwaters. As a major source of water, storm runoff and floods overflowing the river banks or diverted by canals may be spread and stored underground for future use. Water (flood) spreading is perhaps the most widely used man-made system of artificial groundwater recharge. Water is spread over the ground, so it infiltrates into the soil and recharges the aquifer. The efficiency of the spreading system depends on several factors such as the spreading area, duration of recharge, soil hydraulic properties, volume of spread water, and topography.

Todd (15) classifies water spreading methods into stream, ditch and furrow, pit, basin, and flooding, out of which the last two methods are briefly mentioned here. Basin, the most favored and widely practiced method, is formed by dikes or levees or by excavation. Water is normally diverted from a stream and led to a single or series of basins, where water spills to lower basins. Large lands with gentle slopes are required for basin spreading. Sedimentation and sealing is likely to occur when turbid water is used, so the upper basins act as settling areas.

Periodic scraping, disking, or scarifying may be performed to restore recharge rates.

In the flooding method, suitable for a low 1–3% gradient topography, water (flood) is diverted from a stream to the spreading area. Erosion must be avoided as the shallow water flows over the area. The preferable texture of the soil is as coarse as possible to allow a high recharge rate. The cost of construction and maintenance is usually low for this method, but sediment, evaporation, and high floods are major problems. The former problem decreases the system efficiency, and the latter causes damage to canals, embankments, and spillways. The flooding method applied for groundwater recharge is generally similar to the flood spreading systems designed for irrigation.

FLOOD SPREADING SCHEMES

Many different flood spreading systems have been constructed in different countries. Flood spreading systems, mainly for irrigation, may be divided into two types: flow type and detention type (16). Free drainage and flow of water occurs in the first category thanks to a gentle land slope, whereas water is retained in the second type systems for infiltration. The flow type system is further divided into syrup-pan, spreader ditch, and dike and bleeder. The detention type is also classified into manual inlet control and automatic inlet control. The following describes different flow type systems, as stated by Jensen (16).

In the syrup-pan flow system, a single spreader ditch at the upper end of the field picks up the water from the main watercourse. Water spills over the sides of the spreader into the field divided into some sections by contour dikes. While flowing down the field, it partly infiltrates and partly runs off to the next section downstream. Contour dikes are broken at one end to allow water to proceed to the next section. At the end of the last section, there is a waste way or a channel leading excess water, if any, back to the main channel. Figure 1 presents a typical sketch of the system. In Fig. 2, a spillway along a contour dike is shown during operation.

In a spreader ditch flow system, a number of ditches divert water from a stream. The slope of the ditches decreases toward the lower end. Dikes and bleeder flow systems are a modified type of a syrup-pan system. Water flows through dikes to lower sections of the fields via tubes or weirs installed at intervals along the dikes. Emergency waterways connect each section to the main stream in case the tubes are blocked. Detention type systems are designed to divert and spread long duration flows with control over the depth of water.

DESIGN FACTORS

Several factors must be considered in the design of flood spreading systems. The meteorology and hydrology of the watershed, which produces the runoff for spreading, is of particular importance. Flow duration curves as well as flood values and the sediment load are to be studied. The major factors considered in the design of efficient spreading systems are diversion sites, slope of the spreading land, soil infiltration and water holding capacity, the area available for spreading, and economic considerations such as maintenance costs. Although a standard and widely practiced procedure for design of spreading systems does not exist, some irrigation and soil and water conservation books may include a chapter for designing them (16).

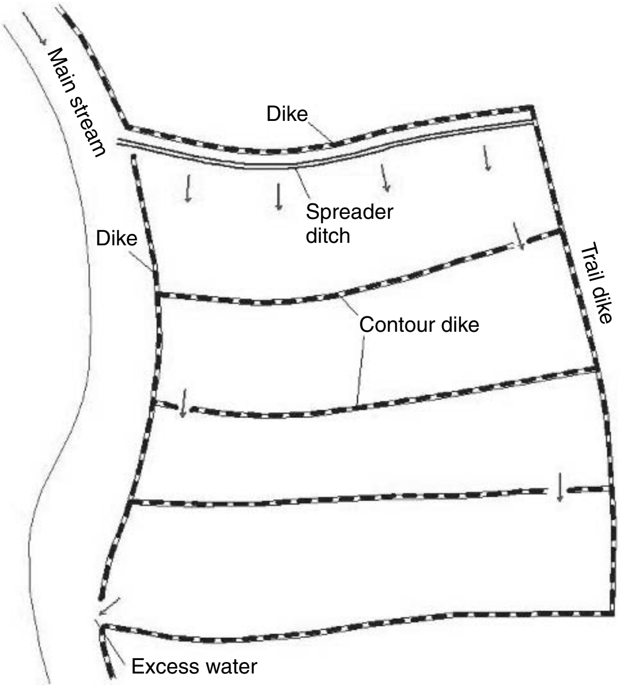

Figure 1. A typical syrup-pan flow system (modified from Reference 16).

Figure 2. A spillway along a contour dike during operation (Ghareh Baygon Flood Spreading Station, Fars Province, Iran).

ADVANTAGES AND PROBLEMS

Flood spreading systems have the advantages of generally low construction and maintenance costs and adaptability to rural areas because of their simple technology. By storing floodwater that would have been otherwise lost, local water demands can be partially met and flood damage downstream may be reduced. Other indirect benefits include environmental diversity, people migration control, and desert rehabilitation in arid areas.

The major problem of flood spreading projects is sedimentation. This decreases soil infiltration, which is a serious consequence in artificial recharge and flood irrigation systems. Efforts to maintain the efficiency of the system include growing vegetation, adding chemical soil conditioners, scarification, and leaving upper fields for siltation.

BIBLIOGRAPHY

1. Bennett, H.H. (2001). *Soil Conservation*. Agrobios, Jodhpur, India.
2. Stoddart, L.A., Smith, A.D., and Box, T.W. (1975). *Range Management*. McGraw-Hill, New York.
3. Heady, H.F. (1975). *Rangeland Management*. McGraw-Hill, New York.
4. Gupta, G.N., Bala, N., and Choudhary, K.K. (1995). Effect of run-off harvesting and conservation techniques on growth and biomass production of Prosopic cineraria. *Indian Forester* **121**: 702–710.
5. Hudson, N.W. (1987). *Soil and Water Conservation in Semi-Arid Areas*. FAO, Rome.
6. Billy, B. (1981). Water harvesting for dryland and floodwater farming on Navajo Indian Reservation 3–7. In: *Rainfall Collection for Agriculture in Arid and Semi-Arid Regions*. G.R. Dutt, C.F. Hutchinson, and A. Garduno (Eds.). Workshop at Tucson, AZ. Commonwealth Agric. Bureau, Fanham Royal, Bucks., UK.
7. UNEP. (1983). *Rain and Storm Water Harvesting in Rural Areas*. Tycooly International, Dublin.
8. Kovda, V.A. (1961). Land use development in the arid regions of the Russian plain, the Caucasus, and Central Asia. In: *A History of Land Use in Arid Regions*. Arid Zone Research, Vol. XVII, UNESCO, Paris.
9. Carr, M.K.V. (1979). *Land Use and Irrigated Agriculture in Woito and Lower Omo River Valleys*. UNDP/FAO Project ETH/78/013, Development of Irrigated Agriculture.
10. Morgan W.T.W. (1974). Sorghum gardens in south Turkana. *Geogr. J.* **140**: 80–93.
11. Mickelson, R.H., Cox, M.B., and Musick, J. (1965). Runoff water spreading on levelled cropland. *J. Soil and Water Conserv.* **20**(2): 57–60.
12. Kolarkar, A.S., Murthy, K.N.K., and Singh, N. (1983). Khadin—a method of harvesting water. *J. Arid Environ.* **6**: 59–66.
13. FAO. (1997). *Irrigation in the Near East Regions in Figures*. Water Report No. 9, FAO, Rome.
14. Prinz, D. (1996). Water harvesting: Past and future. In: *Sustainability of Irrigated Agriculture*. L.S. Pereria (Ed.). Proc. NATO Advanced Research Workshop, Vimeiro, March 21–26, 1994, Rotterdam, pp. 135–144.
15. Todd, D.K. (1980). *Groundwater Hydrology*. John Wiley & Sons, New York.
16. Jensen, M.E. (1981). *Design and Operation of Farm Irrigation Systems*. Monograph No. 3, ASAE, St. Joseph, MI.

MINIMUM ENVIRONMENTAL FLOW REGIMES

Peyman Daneshkar Arasteh
Soil Conservation and
Watershed Management
Research Institute (SCWMRI)
Tehran, Iran

INTRODUCTION

A growing demand exists for water by humans that not only reduced the amount of water available for future industrial and agricultural development but also has a profound effect on aquatic ecosystems and their dependent species (1).

Every inland, coastal, fresh, or brackish water ecosystem has specific water requirements for the maintenance of ecosystem structure, function, and the dependent species. For river-dependent aquatic ecosystems, normally referred to as environmental water requirements (EWR), environmental flow, or a new branch of hydrology, ecohydrology, (1) there is no universal definition of environmental flows. Flow is the governing influence in rivers, controlling many aspects of the physical environment and significantly affecting the biota of a riverine system. It is essential that adequate water be provided for the purpose of protecting the physical and ecological processes and features. Such flows are defined as EWR (2).

EWR are also defined as: "The water regimes needed to sustain the ecological values of aquatic ecosystems at a low level of risk" (3). Given that some systems are overallocated and unable to provide the full EWR, negotiation between stakeholders leads to the establishment of flows for the purpose of environmental protection. Such flows are defined as environmental water provisions, meaning: "The part of EWR that can be met" (3).

An environmental flow is the water regime provided within a river, wetland, or coastal zone to maintain ecosystems, and their benefits, where competing water uses exist and where flows are regulated. Environmental flows provide critical contributions to river health, economic development, and poverty alleviation. They ensure the continued availability of the many benefits that healthy river and groundwater systems bring to society (4).

Balancing the needs of the aquatic environment and other uses are becoming critical in many of the world's river basins as population and associated water demands increase (1).

In order to sustain the ability of freshwater-dependent ecosystems to support food production and biodiversity, environmental flows must be established scientifically, made legitimate, and maintained.

In recent years, a number of local and global assessments of water resources, current and future water use, water and food security projections, and water poverty analyses have been completed (1,2,5–10). Assessment of water availability, water use, and water stress at the global scale has also been the subject of increasingly intensive research over the course of the past decades. However, the requirements of aquatic ecosystems for water have not been considered explicitly in such assessments (1).

OBJECTIVES OF ENVIRONMENTAL FLOW

The following environmental flow objectives apply specifically to the management of flow volumes in a river basin (5):

a. Maintain appropriate minimum environmental flows (MEF) during low flow periods.
b. Provide appropriate 'flushing flow' and 'high flow' regimes.
c. Maintain, enhance, or restore species diversity, population structure, and community assemblages of aquatic biota and associated wildlife dependent on the instream environment.
d. Ensure identified populations of significant biota are adequately protected through provision of appropriate environmental flows.
e. Maintain and improve aquatic habitat, river structure, riparian vegetation, and water quality through provision of adequate environmental flows.
f. Provide a flow regime that preserves and, where possible, reinstates important riffle and pool habitats.
g. Minimize sedimentation and smothering of aquatic habitats through protection, maintenance, and restoration of indigenous riparian vegetation.

METHODS OF DETERMINING EWR

A range of environmental flow determination techniques exist and have been variously applied all around the world (4,11,12). Possible methodologies for calculating environmental flow requirements include:

Hydrological Methods

These methods are based on historical records and are simplest and the least data intense for estimating instream flows. They include the look up tables and desktop analysis.

Habitats Methods

They require some fieldwork to quantify the relationship between the selected variables describing the physical habitat behaviors and the river discharge. These methods are based on hydraulic simulations and habitat rating or modeling.

Holistic Methods

They are discussion-based approaches. The approaches offer a structured framework to the setting of instream flow requirements, based on expert opinion and stakeholder requirements. Holistic approaches make use of teams of experts and may involve participation of stakeholders, so that the procedure is holistic in terms of interested parties as well as scientific issues. Where methods have the characteristic of being holistic, they clearly have the advantage of covering the whole hydrological-ecological-stakeholder system. The disadvantage is that it is expensive to collect the relevant data.

For a review of these methods, reading of Kinhill (13), Arthington and Pusey (14), Arthington and Zalucki (15), and Arthington et al. (16) articles are suggested. The determination of environmental flow requirements requires consideration of the entire flow regime. However, a key distinction in how environmental flow requirements are determined exists between low flows and high flows including floods.

When determining environmental flow requirements, it is important to understand the tolerance of stream biota to certain volumes, timing, and duration of flows, which can be achieved by comparing the current flow regime with the natural flow regime.

REQUIRED DATA

a. Physical description of basin
b. Water quality conditions in basin
c. Description of key biota or environmental values (fauna and flora)
    i. Fish
    ii. Other vertebrates
    iii. Aquatic macroinvertebrates
    iv. Vegetation
d. Water resource development and hydrology
    i. Licensed water use
    ii. System operation
    iii. Stream flows

SOME RECOMMENDATIONS

Harby et al. (17) used a river system simulator analysis to optimize the environmental flow requirement of regulated river Maana in central southern Norway. The analysis recommended an environmental flow release of around 345 MLD (million liters per day). Zampatti and Lieshke (18) recommended 1.5 MLD as MEF for Diamond Creek Catchment, Australia. Zampatti and Close (6) recommended a MEF of 130, 8, and 9 MLD for Kiewa River, Running Creek, and Yackandandah Creek, Australia, respectively. Lieschke et al. (5) recommended maximum MEF of Plenty River system about 1.5 MLD. MDNR (19) recommended a MEF of 380 MLD for Potomac River system, Maryland. Smakhtin et al. (1) estimated EWRs needed to maintain a fair condition of freshwater ecosystems range globally from 20–50% of the mean annual river flow in a basin. These MEFs have been determined on the bases of the habitat and hydrologic-hydraulic methods. Arasteh (20) recommended a volume of 3400 million cubic meters per year as the annual minimum volume of EWR to survive Hamoon, Iran-Afghanistan trans-boundary wetlands on the base of a simple water balance method.

SUITABLE SITES

http://www.g-mwater.com.au
http://www.cgiar.org/iwmi/assessment/index.htm

BIBLIOGRAPHY

1. Smakhtim, V., Revenga, C., and Doll, P. (2004). *Taking Into Account Environmental Water Requirement in Global-Scale Water Resources Assessments*. Comprehensive Assessment Research Report 2, Comprehensive Assessment Secretariat, Colombo, Sri Lanka.
2. SKM. (2000). *Environmental Flow Study of the Yea River System*. Final Report Sinclair Knight Merz Pty. Ltd. and ID&A (SKM), Australia.
3. ARMCANZ/ANZECC. (1996). *National Principles for the Provision of Water for Ecosystems*. Agriculture and Resource Management Council of Australia and New Zealand (ARMCANZ) and Australian and New Zealand Environment and Conservation Council (ANZECC), Sustainable Land and Water Resources Management Committee, Subcommittee on Water Resources, Sydney, Australia.
4. Dyson, M., Bergkamp, G., and Scanlon, J. (2003). *Flow: The Essentials of Environmental Flows*. International Union for Conservation of Nature and Natural Resources (IUCN), Gland, Switzerland and Cambridge, UK.
5. Lieschke, J., Grgat, L., and Zampatti, B. (2000). *An Assessment of Environmental Flow Requirements for the Plenty River Catchment*. Department of Natural Resources and Environment- Freshwater Ecology, Australia.
6. Zampatti, B.P. and Close, P.G. (2000). *An Assessment of Environmental Flow Requirements for the Kiewa River*. Department of Natural Resources and Environment- Freshwater Ecology, Australia.
7. FAO. (2000). *Agriculture: Towards 2015 and 2030*. Technical Interim Report, Food and Agriculture Organization, Rome, Italy.
8. IWMI. (2000). World water supply and demand in 2025. In: *World Water Scenarios Analyses*. F.R. Rijsberman (Ed.). Earthscan Publications, London, UK.
9. UNESCO. (2003). *World Water Development Report*. United Nations Educational Scientific and Cultural Organization, Paris, France.
10. STORM. (2003). *Feasibility of water sensitive urban design in Alpine areas*. Draft Discussion Paper STORM Consulting Pty. Ltd., NSW, Australia.
11. Hass, L. (2002). Mediterranean water resources planning and climate change adaptation. *Proc. of Water, Wetlands and Climate Change*. Mediterranean Regional Roundtable, Athens, Greece, International Union for Conservation of Nature and Natural Resources (IUCN).
12. Symphorian, G.R., Madamombe, E., and van der Zaag, P. (2002). Dam operation for environmental water release: the case of Osborne Dam Save Catchment Zimbabwe. *Proc. of the 3rd Water Net Symposium*. Dar es Salaam, Tanzania. Available: http://www.waternetonline.ihe.nl/aboutWN/3rdProceedings.htm.
13. Kinhill. (1988). *Techniques for Determining Environmental Flow Requirements—A Review*. Kinhill Engineers, Department of Water Resources, Melbourne, Australia.
14. Arthington, A.H. and Pusey, B.J. (1993). In-stream flow management in Australia: methods, deficiencies and future directions. *Australian Biologist*. **6**: 52–60.
15. Arthington, A.H. and Zalucki, J.M. (1998). *Comparative evaluation of environmental flow assessment techniques: review of methods*. LWRDC Occasional Paper 27/98, Land and Water Resources Research and Development Corporation, Canberra, Australia.
16. Arthington, A.H., Brizga, S.O., and Kennard, M.J. (1998). *Comparative Evaluation of Environmental Flow Assessment Techniques: Best Practice Framework*. LWRRDC Occasional Paper 25/98, Land and Water Resources Research and Development Corporation, Canberra, Australia.
17. Harby, A. et al. (1999). Application of the river system simulator for optimizing environmental flow in a Norwegian regulated river. *Proc. of the 28th IAHR Congress*. CD-ROM.
18. Zampatti, B. and Leishke, J. (1999). *An Assessment of Environmental Flow Requirements for the Diamond River Catchment*. Department of Natural Resources and Environment-Freshwater Ecology, Australia.
19. MDNR. (1981). *The Potomac River low flow allocation agreement*. Article 2.C Maryland Department of Natural Resources (MDNR), Annapolis, MD.
20. Arasteh, P.D. (2004). *Development of a Distributed Model for Estimation of Regional Evaporation Using Remote Sensing Techniques*. Case Study: Hamoon Lakes of Iran. Ph.D. Thesis, Tarbiat Modarres University, Tehran, Iran.

FORENSIC HYDROGEOLOGY

LEO S. LEONHART
DAVID R. HARGIS
Hargis + Associates, Inc.
Tucson, Arizona

Forensic hydrogeology is a branch of environmental forensics, a relatively new discipline involving the use of environmental science to support litigation.

Forensics is derived from the Latin root *forensis*, meaning "of a forum," and is the study of formal debate. Accordingly, forensics involves arguing a position within any adversarial context. A legal context is not necessarily a requirement, but environmental forensics often support the interpretation or establishment of technical facts related to litigious matters. Generally, the product of a forensic hydrogeologic investigation is an opinion rendered in writing or orally via various legal vehicles such as affidavits, depositions, or trial testimony. The opinion may be direct or in rebuttal to an opinion proffered by an opposing expert. Due to the technical complexities of many legal matters pertaining to environmental issues, hydrogeologists often find themselves as members of interdisciplinary litigation support teams. These teams may comprise other experts in fields such as toxicology, chemistry, microbiology, engineering, industrial hygiene and safety, regulatory compliance, and industrial history and standards. Typical tasks for hydrogeologists include delineation of the source(s) and extent of contamination in subsurface soils and groundwater, interpretation of subsurface conditions controlling such flow and transport, analysis of groundwater flow and contaminant transport, determination of risk pathways toward potential receptors, and feasibility assessment of various remedial alternatives. The opinions rendered by the hydrogeology expert on these matters may be corroborated or supplemented by the opinion of the other experts involved.

There are also forensic applications in hydrogeology that involve litigation, but which do not involve contamination or toxic tort matters. Examples include settlement of damages between parties involving water supply issues, such as well interference or interruption of baseflow; flooding, such as watershed or hydrographic modifications resulting in increased runoff; or simply well drilling and design, involving poor yields or failures.

FORENSIC ISSUES IN HYDROGEOLOGY

There are a variety of legal issues whose resolution have required the support of forensic hydrogeologists. Recent cases have involved supporting determinations of allocation for environmental remediation costs among multiple responsible parties, circumstances "triggering" insurance policy coverage issues, environmental liabilities accompanying property transfers, and personal damages or tort claims from alleged exposure to contamination released from a site. Other claims have arisen over performance of contractors working on environmental cleanup or who were engaged in the construction or operations and maintenance of facilities that failed and allegedly resulted in contaminant releases.

Determination of fair and defensible cost allocation among parties potentially responsible for groundwater contamination is a complex undertaking. Various approaches involving allocation on the basis of area-, volume-, and mass-of-contamination, as well as weightings based on such factors, have been proposed (1,2). Additional approaches based on toxicity weighting have also been proposed for sites where multiple-groundwater contaminants are commingled. Such commingled plumes may also give rise to issues with regard to selecting alternative remedial technologies.

Depending on the period and type of coverage, exclusions, and various interpretations of what constitutes an occurrence, insurance claims involving liability for environmental damage may generate issues requiring input from hydrogeologists. For example, there may be matters related to the timing of releases relative to the timing of the policy coverage or to the enactment of certain "pollution exclusion" clauses within the policy. In such instances, the hydrogeologist may be able to opine on the likelihood that contamination occurred during the policy period and whether the contamination is persistent, requiring present-day remedial efforts.

Property transfer issues may be similar to the allocation and timing considerations considered earlier. Such matters require careful research of historical activities performed on the property. Hydrogeologists might use record searches, interviews, and review of historical photographic archives to reconstruct chronologies and thereby form opinions on the likelihood of certain operations as causes for contamination.

Tort claims are often the most complex issues and require extensive examination of risk factors from the source of the contamination, the transport pathway, and the point of exposure. Tort damages may not be limited to human exposure but may also extend to resource damage. In such matters, hydrogeologists might be able to help clarify the timing and fate of contamination from the source to the site of exposure using analytical or numerical models based on conceptualizations of the site hydrogeology.

TOOLS AND APPROACHES

In theory, the approach of hydrogeologists serving as experts in litigation is not fundamentally different from their serving in any other professional capacity. In other words, the tools of the trade are the traditional ones, and the approach involves traditional scientific methods of investigation. There are, however, issues that arise in contemporary forensic matters that are more interdisciplinary. In matters involving contaminant liabilities, environmental forensic specialists have been referred to as "industrial paleontologists," signifying their proficiency in unraveling the history of events leading to present-day conditions at a contamination site (3). Having been trained extensively in geology, hydrogeologists are particularly well suited for such "after-the-fact sleuthing."

As in any traditional investigation, forensic study is staged. The first stage is the acquisition and organization of pertinent data and information. This may involve simply compiling records, photoarchives, and searching databases, or it may include actual field data acquisition such as drilling, sampling, and analysis. At this stage, the hydrogeologist has various traditional and specialized tools available. For example, there are methods traditionally employed such as water level measurements and water and soil sample collection and analysis. Contaminant investigations may also incorporate isotopic analysis or chemical fingerprinting to assess timing or source issues to the extent materials released have a unique, detectable chemical signature. The basis for chemical fingerprinting might be weathering characteristics and composition of petroleum hydrocarbons, proprietary additives, alkyl-lead compounds, oxygenates, or isotopic ratios. Hydrogeologists might also consider the changing state of industrial technology. For example, much information has been gathered on the historical use of organic solvents (4). Some solvents were commercially available and in use at different times, as indicated in Table 1. Certain solvents were more prevalent in industrial use than others at different times. Other chemicals were banned for certain usages by regulatory agencies. A hydrogeologist who is knowledgeable about historical chemical usage/availability and can reconcile such information with chemical use and disposal history of the site can present informed opinions on the timing of releases. Using chemical associations in this manner and context is analogous to using index fossils in geochronology.

The next stage involves assessing and analyzing the data per the study objectives. This analysis is based on both the hydrogeologist's conceptualization or alternative conceptualizations of the hydrogeologic system and the impacts of site operations on the soil and groundwater. The hydrogeologist assembles information gathered to determine such things as the occurrence, flow direction,

Table 1. Year of First Commercial Availability of Selected Organic Chemicals[a]

Chemical	Year of First Commercial Availability
Carbon tetrachloride	1907
Trichloroethylene (TCE)	1908
1,2-Dichloroethane (DCA)	1922
Tetrachloroethene (PCE)	1925
DDT	1942
Chlordane	1947
Toxaphene	1947
Aldrin	1948
Dibromochloropropane (DBCP)	1955
Bromacil	1955
Gasoline additives	
Tetraethyl lead	1923
Methyl cyclopentadienyl manganese tricarbonyl (MMT)	1957
Tetramethyl lead	1966
Methyl tert-butyl ether	1980s

[a] Reference 3.

and use of groundwater; interaction with surface water; geologic cross sections; distribution of contaminants in the soil and groundwater; and other factors related to the overall hydrogeologic conditions at and in the vicinity of the site. Current and past operational information can then be used to interpret the contaminant history of the site. The approach for this step can vary widely according to the skills and experience of the expert as well as the history of site operations and the type and availability of site records. In some cases, the hydrogeologist may want to apply various models to simulate or project groundwater flow and solute transport conditions. Other types of modeling that are less familiar to traditional hydrogeology, but which can support interpretations on timing of release, might include various geochemical models, degradation models of various fuels and components, and corrosion models (for tanks and pipeline leaks).

ETHICS AND STANDARDS

In the process of performing these services, extreme care must be taken to organize and document the information used and generated, considering the likely degree of scrutiny by both technical and legal adversaries. In general, this requires strict adherence to professional ethical standards and methods of analysis, systematic maintenance of records, traceability, and disclosure.

With specific regard to the admissibility of expert opinion in the courtroom, some courts use criteria or standards to determine the acceptability of expert testimony. One such precedent was set in the opinion of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, which held that judges must function as gatekeepers in assuring that expert opinion proffered in the courtroom is valid and reliable (5). To this goal, the court suggested criteria based on the following questions:

1. Is the opinion testable and has it been tested?
2. Is the error rate associated with the technique or opinion acceptable?
3. Has the basis for the opinion survived peer review and has it been published?
4. Is it generally accepted among scientists in the pertinent field?

Other courts have simply applied a "general acceptance" criterion, which originated in the 1923 case of *Frye v. United States*, which dealt with admissibility of a polygraph-type examination. This court held that the scientific opinion must have achieved "general acceptance" in its particular field or context.

BIBLIOGRAPHY

1. Graves, B.J. et al. (2000). Allocating responsibility for groundwater remediation costs. *The Trial Lawyer—J. Strategy, Technique, Case Manage.* **23**: 159–171.
2. Marryott, R.A. et al. (2000). Allocating remedial costs at Superfund sites with commingled groundwater contaminant plumes. *J. Environ. Forensics* **1**: 47–54.
3. Clarke, J.H., Clarke, A.N., and Smith, J.S. (1999). Environmental forensics: Using science to resolve adversarial situations. *Environ. Prot.* **September**: 49–54.
4. Doherty, R.E. (2000). A history of the production and use of carbon tetrachloride, tetrachloroethylene, trichloroethylene, and 1,1,1-trichloroethane in the United States: Part 1—historical background; carbon tetrachloride and tetrachloroethylene and Part 2—trichloroethylene and 1,1,1-trichloroethane. *J. Environ. Forensics* **1**: 69–81, 83–93.
5. Faigman, D.L. (2002). Is science different for lawyers? *Science* **297**: 339–340.

FORESTS AND WETLANDS

Theodore A. Endreny
Emera Bridger
SUNY-ESF
Syracuse, New York

Forests and wetlands are landscape features that play integral roles in the hydrological cycle. Within the landscape, forests and wetlands are inextricably linked. Wetlands may exist within forest ecosystems, or wetland ecosystems can be characterized as "forested." This article will outline the characteristics of both forests and wetlands and some implications for water resource management.

FORESTS

Forests cover an estimated area of 4.17×10^9 hectares, which is roughly 32% of Earth's surface (1). A forest can be defined as "a biotic community predominated by trees and woody vegetation that covers a large area" (1), whereas others define forests by its vegetative character irregardless of area. Forests serve social, economic, and ecological functions, which impact water resources in

Table 1. Year of First Commercial Availability of Selected Organic Chemicals[a]

Chemical	Year of First Commercial Availability
Carbon tetrachloride	1907
Trichloroethylene (TCE)	1908
1,2-Dichloroethane (DCA)	1922
Tetrachloroethene (PCE)	1925
DDT	1942
Chlordane	1947
Toxaphene	1947
Aldrin	1948
Dibromochloropropane (DBCP)	1955
Bromacil	1955
Gasoline additives	
Tetraethyl lead	1923
Methyl cyclopentadienyl manganese tricarbonyl (MMT)	1957
Tetramethyl lead	1966
Methyl tert-butyl ether	1980s

[a] Reference 3.

and use of groundwater; interaction with surface water; geologic cross sections; distribution of contaminants in the soil and groundwater; and other factors related to the overall hydrogeologic conditions at and in the vicinity of the site. Current and past operational information can then be used to interpret the contaminant history of the site. The approach for this step can vary widely according to the skills and experience of the expert as well as the history of site operations and the type and availability of site records. In some cases, the hydrogeologist may want to apply various models to simulate or project groundwater flow and solute transport conditions. Other types of modeling that are less familiar to traditional hydrogeology, but which can support interpretations on timing of release, might include various geochemical models, degradation models of various fuels and components, and corrosion models (for tanks and pipeline leaks).

ETHICS AND STANDARDS

In the process of performing these services, extreme care must be taken to organize and document the information used and generated, considering the likely degree of scrutiny by both technical and legal adversaries. In general, this requires strict adherence to professional ethical standards and methods of analysis, systematic maintenance of records, traceability, and disclosure.

With specific regard to the admissibility of expert opinion in the courtroom, some courts use criteria or standards to determine the acceptability of expert testimony. One such precedent was set in the opinion of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, which held that judges must function as gatekeepers in assuring that expert opinion proffered in the courtroom is valid and reliable (5). To this goal, the court suggested criteria based on the following questions:

1. Is the opinion testable and has it been tested?
2. Is the error rate associated with the technique or opinion acceptable?
3. Has the basis for the opinion survived peer review and has it been published?
4. Is it generally accepted among scientists in the pertinent field?

Other courts have simply applied a “general acceptance” criterion, which originated in the 1923 case of *Frye v. United States*, which dealt with admissibility of a polygraph-type examination. This court held that the scientific opinion must have achieved “general acceptance” in its particular field or context.

BIBLIOGRAPHY

1. Graves, B.J. et al. (2000). Allocating responsibility for groundwater remediation costs. *The Trial Lawyer—J. Strategy, Technique, Case Manage.* **23**: 159–171.
2. Marryott, R.A. et al. (2000). Allocating remedial costs at Superfund sites with commingled groundwater contaminant plumes. *J. Environ. Forensics* **1**: 47–54.
3. Clarke, J.H., Clarke, A.N., and Smith, J.S. (1999). Environmental forensics: Using science to resolve adversarial situations. *Environ. Prot.* **September**: 49–54.
4. Doherty, R.E. (2000). A history of the production and use of carbon tetrachloride, tetrachloroethylene, trichloroethylene, and 1,1,1-trichloroethane in the United States: Part 1—historical background; carbon tetrachloride and tetrachloroethylene and Part 2—trichloroethylene and 1,1,1-trichloroethane. *J. Environ. Forensics* **1**: 69–81, 83–93.
5. Faigman, D.L. (2002). Is science different for lawyers? *Science* **297**: 339–340.

FORESTS AND WETLANDS

Theodore A. Endreny
Emera Bridger
SUNY-ESF
Syracuse, New York

Forests and wetlands are landscape features that play integral roles in the hydrological cycle. Within the landscape, forests and wetlands are inextricably linked. Wetlands may exist within forest ecosystems, or wetland ecosystems can be characterized as “forested.” This article will outline the characteristics of both forests and wetlands and some implications for water resource management.

FORESTS

Forests cover an estimated area of 4.17×10^9 hectares, which is roughly 32% of Earth’s surface (1). A forest can be defined as “a biotic community predominated by trees and woody vegetation that covers a large area” (1), whereas others define forests by its vegetative character irregardless of area. Forests serve social, economic, and ecological functions, which impact water resources in myriad ways. A salient economic use of forested land throughout the world is as a source of fuel and timber. The various hydrological functions that forests perform have substantial economic benefits as well, which ecologists and economists have tried to quantify in recent decades. These economic factors may be a driving force in forest and water resource management decisions.

DISTRIBUTION OF FORESTS

Forest types vary along temperature and moisture gradients. Forests generally occur in regions where the annual precipitation exceeds 38 cm, the frost-free period is at least 14 to 16 weeks in duration, and solar radiation is greater than 27 W/m^{-2} (1). Regional conditions, such as ocean currents and air mass circulation patterns, may also determine the range of forest types in a particular location.

Forested landscapes impact global hydrologic systems by vapor and gas exchange, and conversely, these hydrologic regimes impact the distribution and composition of forests. For instance, overharvesting of timber in semiarid regions may result in lower precipitation, thus further constraining the growth of vegetation (2). On a local scale, the composition of a forest will impact processes such as throughfall, residence time, chemical enrichment, litter characteristics, and soil characteristics (3).

FOREST STRUCTURE AND THE HYDROLOGIC CYCLE

Interception

Precipitation, or the deposition of liquid or solid water onto the Earth's surface, can fall directly onto the soil surface, be intercepted by vegetation or other surfaces, or flow to the ground via trunks and stems (4). The forest canopy plays a significant role in precipitation interception. It has been found that 10–20% of annual precipitation is removed by canopy interception, depending on climatic and vegetation characteristics (1).

Litter on the forest floor may also intercept and store water temporarily before it infiltrates into the soil. The leaching that occurs as water travels through the litter layer can have a substantial impact on the biogeochemistry of the site. The enrichment of precipitation intercepted by litter will vary among forest types because of differences in litter composition. On average, the litter layer in deciduous forests is thinner than in coniferous forests (3), therefore constituents (i.e., nitrogen, calcium) have been found to leach out in higher quantities at sites dominated by deciduous species as compared with conifer-dominated locations (3)

Evapotranspiration

A portion of the intercepted water may evaporate and return to the atmosphere. The rate of evaporation is determined by the amount of water present and the amount of solar energy available (5). Transpiration can also result in the loss of water vapor into the atmosphere. Transpiration is the process whereby water is released from plants' stomata on leaves and lenticels on stems. Transpiration rates can be a considerable output of water in a watershed depending on the type of vegetation and the age of the stand. On average, 95% of the water taken up by a plant's roots will be transpired, whereas the remaining 5% will be converted into biomass via photosynthesis (5). Chang (1) states "a mature forest can transpire as much as 1000 tons of water to produce 1 ton of wood," while noting that this figure can be even higher in some species. As a result of the difficulty of empirically differentiating the water lost by evaporation and that which is lost by transpiration, these two processes are discussed as single process called evapotranspiration.

Throughfall and Stemflow

Precipitation is also routed toward the forest floor by the processes of throughfall or stemflow. Throughfall is defined as the portion of the water that falls without being intercepted or that which was stored in the canopy that drops from the canopy to the forest floor (5). Stemflow is the water the travels along the stems of plants. Throughfall and stemflow are important processes because they influence the distribution of and quality of the water that reaches the soil surface, and thereby they have a "[significant] impact [on] forest biogeochemical cycles" (6). Leaching of nutrients from above ground biomass in deciduous hardwood forests can account for up to 20% of the annual nitrogen and potassium fluxes (6).

Despite its small volumetric contribution to the forest ecosystem, stemflow, in particular, has been found to have a significant impact on runoff generation, soil moisture patterns, and soil water chemistry (6). The rate and magnitude of stemflow has been found to be correlated with vegetation type and structure, storm event characteristics, and seasonality. In general, stemflow yields increase with the magnitude of a storm event (6). The chemical inputs from stemflow also vary based on seasonal variables. Concentrations of target constituents were found to be greater during the winter months than during the summer. Changes in canopy density, greater residence time in the bark, and the lower bark pH during the winter are some of the seasonal factors (6).

Infiltration and Runoff Mechanisms

The forest type, and its subsequent root system, plays an important part in soil genesis and structure, as well as the soil's infiltration capacity. In forest ecosystems, 40–70% of the total biomass is found below the soil surface. Therefore, root systems introduce a significant amount of organic material into the system, which facilitates the formation of aggregates, the organization of which determines soil structure The roots also remove soil moisture from the adjacent soil, which serves to increase soil aggregation (7).

In addition to the organic matter that is contributed by the forest root system, roots also move soil and form macropores that play an important role in infiltration and percolation. Woody plants can send their roots more than 5 m deep depending on the substrate (7). These roots often follow existing channels and voids, but as they grow, they expand these spaces. When roots die, these voids are left empty and can influence the hydraulic conductivity of the soil under certain conditions.

The type and structure of forests also affects the process of infiltration of water into the soil (1). The infiltration capacity is the "maximum rate at which precipitation can infiltrate into the soil" (5). In most forest systems, the infiltration capacity is generally high and infiltration excess overland flow does not generally occur and cannot be considered a significant runoff mechanism (5). That being said, surface conditions such as vegetation cover, roughness, crusting, soil temperature, and slope can determine the quantity of overland flow and its velocity (1).

The hydraulic conductivity is the ability of a soil to carry water or another liquid (5). The conductivity of a soil depends both on the properties of the soil as well as the properties of the liquid. The soil structure and the system of macropores will impact the conductivity. The water that firsts enters the soil after a storm event will follow preferential pathways. Typically, water infiltrating at the soil surface will follow the path of least resistance filling large pores, cracks, and animal burrows before the rest of the soil has been saturated. Forests, like was mentioned above, have a greater number of macropores because of root propagation and the decomposition of dead roots.

In addition, a healthy root system can reduce the amount of soil lost in overland flow events. Thus, soil stabilization and retention can prevent erosion and retard surface water pollution.

Management Implications

As a result of the integral relationship between forests and the hydrological cycle, the extent and quality of forested areas will have an impact on the water balance of the watershed. On average, peak flows of streams are generally lower in watersheds that have a high percentage of forested area (1). Activities that reduce forest cover, such as residential and industrial development, clearing for agricultural operations, and logging, may change the hydrological balance within a given watershed. According to Huber and Iroumé (8), intensive forest operations can have implications for water yield and quality. Removal of forest vegetation can impact soil moisture, infiltration capacity, and total water yield. A decrease in interception capacity, by removing or altering the forest canopy, can result in greater soil moisture.

For example, reduction of forest canopy area can cause an increase in snow accumulation on the ground. As the snow melts, an overall increase in water yield will occur because the storage capacity of the watershed has decreased. Also, an increase of impervious surfaces can decrease infiltration, contributing to a greater quantity of infiltration excess overland flow. Therefore, a greater quantity of water is released into streams.

WETLANDS

Wetlands are ecosystems inundated by water during the growing season, which produces soils dominated by anaerobic processes and in which the biota exhibits flood-tolerant adaptations (9). Wetlands cover approximately 4–5% of the Earth's surface. They exist in across a broad spectrum of climates from boreal peatlands to tropical mangrove forests (9). Although they make up a relatively small percentage of the total land area, wetlands can have disproportional affects on hydrological and biogeochemical cycles in local, regional, and continental contexts.

FORESTED WETLANDS

Forested wetlands are the dominant wetland type in the continental United States (10). These wetlands are of particular interest because they lie at the interface of upland and aquatic systems; these wetlands possess characteristics of both upland and aquatic systems while having distinctive properties found only in wetlands. This unique landscape position makes forested wetlands important to surface and groundwater quality as well as biodiversity. The forested wetlands of most interest to this discussion are swamps and riparian wetlands.

Swamps are wetlands that are dominated by trees rooted in hydric, mineral soils (9,11). Swamps have standing water through most or all of the growing season (11). Freshwater swamps, or deepwater swamps, have a broad geographic distribution ranging from the bald cypress-tupelo (*Taxodium distichum-Nyssa aquatica*) swamps that are found from the Midwest to the southern United States to the Atlantic white cedar (*Chamaecyparis thyoides*) swamps and red maple (*Acer rubrum*) swamps of the northeastern United States. The hydrogeology and landscape position of swamps will, in large part, determine the hydrologic and ecological functions that the wetland performs as well as the magnitude of these functions (11)

The riparian zone is the corridor of land on the banks of streams, rivers, and lakes. Riparian wetlands are found within the riparian zone, in areas that receive seasonal inputs of water through flooding of the adjacent body of water, but that do not have standing water throughout the growing season. Although riparian wetlands are not necessarily forested (i.e., wet meadows), they often are. Willows (*Salix* spp.), red alder (*Alnus rubra*), and green ash (*Fraxinus pennsylvanica*) are common bottomland species found in riparian wetlands. Riparian vegetation, particularly riparian forests, has a significant affect on stream geomorphology and ecology. It has been found that stream water temperatures, for example, are correlated with riparian soil temperatures. Riparian vegetation also provides allochthonous inputs that are used by aquatic biota as a source of carbon.

ROLE OF WETLANDS IN WATERSHED HYDROLOGY

Wetlands, especially forested wetlands, perform similar hydrologic functions to those performed by upland forests. Interception, evapotranspiration, throughfall, and stemflow are all processes that occur in wetlands. However, unlike upland forests, wetlands are more directly impacted by the hydrologic environment. The character of wetlands is directly affected by specific effects of hydrology (11). Hydrology leads to specialized vegetative communities; also, flow properties can determine productivity, and nutrient cycling is impacted by the hydrological characteristics of a wetland (11). Wetlands also provide important hydrological functions, including hydrological storage, biogeochemical cycling, and pollutant retention (9).

Hydrological storage includes flood attenuation, groundwater recharge, and in some cases, groundwater discharge. Wetlands attenuate flood waters by intercepting storm runoff and flood waters, storing the water for some period of time, and then releasing the water at a slower rate than if it had not been intercepted. Thus, wetlands reduce the peak flows and spread them out over time (11). The abatement of floods by wetlands has both economic and ecological implications. The reduction of peak flows also has water quality benefits. Wetlands reduce the velocity of the peak flows, which allows sediments and other pollutants to settle out from the water column. Flood reduction can also decrease scour of stream bed and banks, reducing the amount of material eroded from stream channels.

Wetlands contribution to groundwater is largely dependent on the geomorphology of the wetland. Groundwater dynamics can play an important role in some wetlands, but be negligible in others. Groundwater recharge is common in depression wetlands such as prairie potholes in the mid-western United States. These wetlands fill with water during the spring, and as the growing season progresses, water from the wetland infiltrates the soil and percolates into the aquifer. In other wetlands, such as fens, groundwater is discharged into the wetland and is the dominate source of water for the wetland. In other wetlands, neither significant groundwater recharge nor discharge occurs (11).

Biogeochemical cycling is another crucial function of wetlands. Wetlands specialized role in biogeochemical cycling is largely because of the alternative aerobic and anaerobic soil conditions throughout the year (11). When the wetland soils are saturated, they became anaerobic and a series of reduction reactions occur. The reduction of compounds can alter their bioavailability and prevent their introduction into surface and groundwaters. When the soil moisture decreases and aerobic activity again commences, these constituents can become available again. A large body of research has examined the removal of nitrates (NO_3^-) from subsurface water and phosphorous (P) from overland flow in riparian systems (12–15). The efficiency of these wetlands to remove nutrients and sediments is largely because of vegetation characteristics, preferential flow paths, and seasonality of inputs.

Wetlands are sometimes referred to as the "kidneys of the landscape" because they can remove nutrients, particularly nitrogen and phosphorus, pesticides, bacteria, heavy metals, and other potential pollutants from the surface or groundwater that they intercept. Similarly, researchers often conceptualize the landscape as a series of sources and sinks. Within this framework, wetlands are often thought of as sinks for pollutants. In reality, the relationship between the wetland and its watershed is often much more complex. At a certain point, wetlands may become saturated with a certain constituent and become a source rather than a sink (16). For example, when a high sediment yield exists in a watershed, wetlands may become a source of sediment. In large basins that have high upland yield, wetlands can actually account for 52–91% of the sediment discharged into streams. The role of wetlands as sources of sediment may be more variable in smaller watersheds, depending on the geomorphology and land use in the watershed (17).

The benefits of having wetlands throughout the watershed are now well known. Yet, anthropogenic disturbances as well as natural variation have contributed to the loss of wetlands throughout the world. It is estimated that, in the United States, the loss of 50% of the wetlands has occurred since European settlement (11). Wetlands are typically drained so that the land can be used for other uses, such as agricultural or residential development. The loss of wetlands from the watershed can translate to increased stream discharge, increased sediment yields, and eutrophication of surface waters.

Another anthropogenic disturbance that impacts riverine swamps and riparian wetlands is the reduction of flow in the stream channel or channel alteration. These disturbances impact the seasonal inputs of water from flooding and may cause changes in the composition and distribution of plant species. These changes may also impact the rate and magnitude of biogeochemical as well as hydrologic functions.

CONCLUSION

Although it may be helpful to look at forests and wetlands as two distinct ecosystems, it is important to understand the interconnections between them. Forest management regimes may determine the types of constituents as well as the amount of water that travels into bottomland areas by overland flow as well as subsurface flow paths. Subsequently, the condition of wetlands can have an effect on the storage and cycling of nutrients, sediments, and organic matter. The consideration of both of these ecosystems when making management decisions will enable managers to accentuate the ecosystem functions of one ecosystem without compromising the health of the other.

BIBLIOGRAPHY

1. Chang, M. (2003). *Forest Hydrology: An Introduction to Water and Forests*. Lewis Publishers, Boca Raton, FL.
2. Sahin, V. and Hall, M. (1999). The effects of afforestation and deforestation on water yields. *J. Hydrol.* **178**: 293–309.
3. Kavvadias, V.A. et al. (2001). Litterfall, litter accumulation and litter decomposition rates in for forest ecosystems in northern Greece. *Forest Ecol. Manag.* **144**: 113–127.
4. Crockford, R.H. and Richardson, D.P. (2000). Partitioning of rainfall into throughfall, stemflow, and interception: effect of forest type, ground cover and climate. *Hydrologic. Processes.* **14**: 2903–2920.
5. Hornberger, G.M., Raffensberger, J.P., Wiberg, P.L., and Eshleman, K.N. (1998). *Elements of Physical Hydrology*. John Hopkins University Press, Baltimore, MD.
6. Levia, D.F. and Frost, E.E. (2003). A review and evaluation of stemflow literature in the hydrologic and biogeochemical cycles of forested and agricultural ecosystems. *J. Hydrol.* **274**: 1–29.

7. Brady, N.C. and Weil, R.R. (2002). *The Nature and Properties of Soils*, 13th Edn. Prentice Hall, Upper Saddle River, NJ.
8. Huber, A. and Iroumé, A. (2001). Variability of annual rainfall partitioning for different sites and forest covers in Chile. *J. Hydrol.* **248**: 78–92.
9. Keddy, P.A. (2002). *Wetland Ecology: Principles and Conservation*. Cambridge University Press, Cambridge, UK.
10. Lockaby, B.G. and Connor, W.H. (1999). N:P balance in wetland forests: productivity across biogeochemical continuum. *Botan. Rev.* **65**(2): 171–185.
11. Mitsch, W.J. and Gosselink, J.G. (2000). *Wetlands*, 3rd Edn. John Wiley & Sons, New York.
12. Lowrance et al. (1985). *Risk Assessment*. McGraw Hill, New York.
13. Jordan, T.E., Correll, D.L., and Weller, D.E. (1993). Nutrient interception by a riparian forest receiving inputs from adjacent cropland. *J. Environ. Qual.* **22**: 467–473.
14. Ashby, J.A., Bowdon, W.B., and Murdoch, P.S. (1998). Controls on denitrification in riparian soils in headwater catchments of a hardwood forest in the Catskill mountains, USA. *Soil Biol. Biochem.* **30**(7): 853–864.
15. Cey, E.E., Rudolph, D.L., Aravena, R., and Parkin, G. (1999). Role of the riparian zone in controlling the distribution and fate of agricultural nitrogen near a small stream in southern Ontario. *J. Contaminant Hydrol.* **37**: 45–67.
16. Cooper, J.R., Gilliam, J.W., and Jacobs, T.C. (1987). Riparian areas as a control of nonpoint pollutants. In: *Watershed Research Perspectives*. D.L. Correll (Ed.). Smithsonian Institution Press, Washington, DC.
17. Phillips, J.D. (1989). Fluvial sediment storage in wetlands. *Water Res. Bull.* **25**: 867–873.

ROCK GLACIER

Alexander Brenning
Humboldt–Universität zu Berlin
Berlin, Germany

Active rock glaciers are the geomorphological expression of creeping mountain permafrost (Fig. 1). They consist of a debris–ice mixture that mechanically deforms and slowly moves downslope or downvalley and develop therefore a tongue-shaped or lobate landform with a front scarp and characteristic surface structures, such as furrows and ridges, as rheological expressions. Unlike "true" glaciers, rock glaciers do not present surface ice or exposed ice in general but contain only supersaturated ground ice.

Rock glaciers that do not move any longer but still contain ice are called inactive, and the sediment body that remains after a rock glacier has melted is called a relict (or fossil) rock glacier. Rock glaciers may become inactive for climatic or dynamic reasons. In the former case, the seasonally unfrozen block mantle has grown as a consequence of permafrost degradation. A dynamically inactive rock glacier, in contrast, may develop even in areas of continuous permafrost if it enters into flat terrain or if its thickness falls below a critical value. The front scarp of active rock glaciers generally reaches inclinations between 35° and up to 45°, but inactive ones show front slopes at or below the angle of repose, and the upper part of inactive front slopes is gentler than for activity. Relict rock glaciers can be recognized from their subdued appearance and irregular surface structure. Further classification criteria are presented in Table 1.

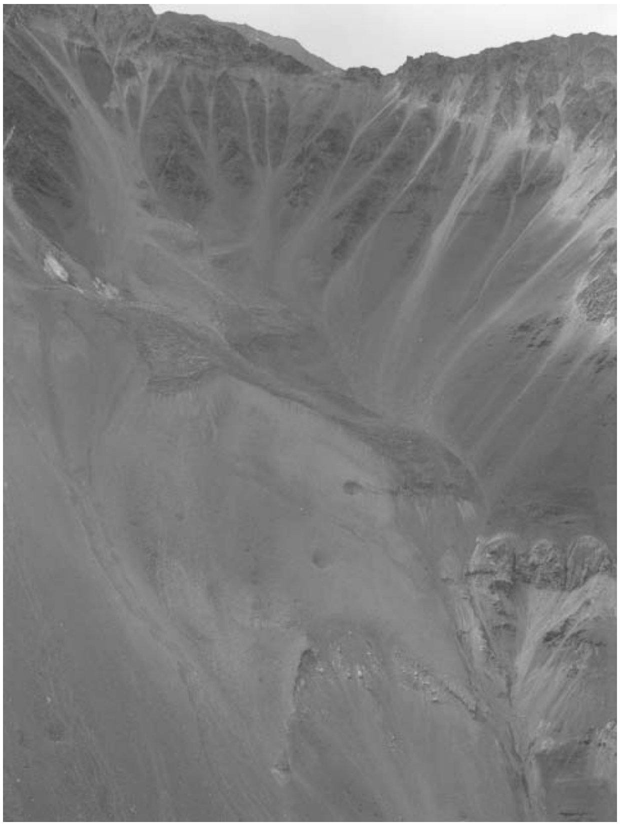

Figure 1. An active talus rock glacier in the Andes of Central Chile (Estero de las Yeguas Muertas, 33.5°S). (The crater-like depressions are produced by subrosion of gypsum).

In the 19th century, several authors already described rock glaciers, but only with the work of Wahrhaftig and Cox (1959) did modern and more comprehensive rock glacier research begin. However, until today, some confusion persists, mainly of the terminological kind concerning the precise differentiation of rock glaciers and glaciers (Table 1).

The development of rock glaciers is a consequence of the enrichment of debris with percolating snow meltwater, which freezes under permafrost conditions and forms interstitial and segregation ice. Therefore, rock glaciers are of periglacial origin. They may form out of talus accumulations (talus rock glaciers) or morainic debris (debris rock glaciers).

Rock glaciers contain between 40% and 60% of ice by volume and therefore constitute unsuspicious stores of water. In the Swiss Alps, rock glaciers contain only 0.03 km^3 of water equivalent per 1000 km^2 of mountain area, compared with 2.5 km^3 per 1000 km^2 stored within glaciers. In semiarid mountain areas, such as the Andes at 33°S, in contrast, the amount of water stored within rock glaciers is by one order of magnitude higher, while at the same time glacier sizes are smaller.

Table 1. Rock Glacier Taxonomy as Suggested by Barsch (1987)

Type	Talus rock glacier (material: talus) Debris rock glacier (morainic debris) Special rock glacier (other material)
Location	For talus rock glaciers: –Valley head, including cirques –Valley side walls –Footslope below cliff For debris rock glaciers: –Glacier terminus –Side of glacier
Connection to source area	Direct Not direct
Surface relief	Very well developed Subdued to well developed No furrows and ridges
Surface grain sizes	Earthy/blocky
Form	Singular/complex
Complexity	Multipart/multilobe/multiunit/multiroot
Shape	Tongue-shaped/lobate/transitional
Size (area)	Small ($<10^4$ m^2), medium, huge ($>10^5$ m^2)
Activity	Active/inactive/relict

Therefore, rock glaciers are important water resources in high mountain areas of continental semiarid and arid climates. In spite of this, little is known about their quantitative importance for river discharge.

The lower limit of rock glacier distribution is associated with the lower limit of discontinuous mountain permafrost, which is often attributed to the $-1\,^{\circ}$C to $-2\,^{\circ}$C isotherm of mean annual air temperature. Therefore, rock glaciers are considered good permafrost and climate indicators on a regional scale.

Concerning rock glacier movement, horizontal displacement rates range between a few centimeters and 1–2 m per year. Velocities are highest in the central lower part of the rock glacier and decrease laterally and toward the rock glacier's rooting zone. Furthermore, rock glaciers seem to move faster in summer than in winter.

The age of active rock glaciers may generally range between several thousand years and up to 10,000 years. Although several attempts have been made to obtain more precise data, absolute age determination has turned out to be extremely difficult. As climatically inactive rock glaciers are often situated at the lower limit of the distribution of active ones and up to 200 m below this limit, their inactivity probably developed since the Little Ice Age.

Recent rock glacier investigations focus increasingly on their physical properties. Geophysical methods, such as geoelectric soundings and refraction seismics, are being used in order to differentiate between permafrost and unfrozen ground along cross sections through the rock glacier. Furthermore, boreholes have been established in order to directly observe the internal structure of rock glaciers and monitor the change of deformation rates with depth. As the coarse boulder mantle of rock glaciers may strongly influence their energy balance, this issue is also being addressed within the more general framework of ground thermal regime in permafrost areas. At a regional scale, statistical methods are being applied in order to determine morphological and climatic controls on rock glacier formation.

READING LIST

Barsch, D. (1987). *Rock Glaciers: An Approach to their Systematics*. Giardino, Shroder, & Vitek (Eds.).

Barsch, D. (1996). *Rockglaciers*. Springer, Berlin.

Clark, M. (Ed.). (1988). *Advances in Periglacial Geomorphology*. John Wiley & Sons, New York.

Giardino, J.R., Shroder, J.F., and Vitek, J.J.D. (Eds.). (1987). *Rock Glaciers*. Allen & Unwin, Boston, MA.

Gorbunov, A.P. (1983). Rock glaciers of the mountains of middle Asia. *Proc. 4th International Conference on Permafrost, Fairbanks, Alaska*. National Academic Press, Washington, DC, pp. 359–362.

Haeberli, W. (1985). *Creep of Mountain Permafrost: Internal Structure and Flow of Alpine Rock Glaciers*. Mitteilungen der Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie 77, p. 142.

Hauck, C. (2001). *Geophysical Methods for Detecting permafrost in High Mountains*. Mitteilungen der Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie der ETH Zürich 171. PhD thesis, p. 194.

Wahrhaftig, W.B. and Cox, A. (1959). Rock Glaciers in the Alaska Range. *Geol. Soc. Am. Bull.* **70**: 383–436.

GREAT LAKES

Theodore A. Endreny
SUNY ESF
Syracuse, New York

GEOLOGIC ORIGINS

Lakes Superior, Michigan, Huron, Erie, and Ontario comprise the five interior freshwater lakes that extend across a total area of 244,000 km^2 (0.16% of world land area), occupy parts of six states and two provinces in two nations, and contain nearly 20% of the world's freshwater supply (see Fig. 1). Glaciers, although frozen, are the only contiguous surface freshwater volume that contains more freshwater; they were the geologic forces that carved these lakes and the surrounding drainage area from the Precambrian rocks and overlying marine deposits from earlier inland seas that are now rocks classified by their Cambrian, Ordovician, Silurian, Devonian, Mississippian, and Pennsylvanian periods. The current Pleistocene Epoch of glacier activity, whose vertical heights were 2 km, compressed and scoured the area during numerous advances and retreats. During each icesheet retreat, the ice-excavated lake basins were filled by meltwater, and the absence of ice overload burden triggered areas of earth uplift and the creation of drainage outlets. Prior to the St. Lawrence drainage that followed the most recent

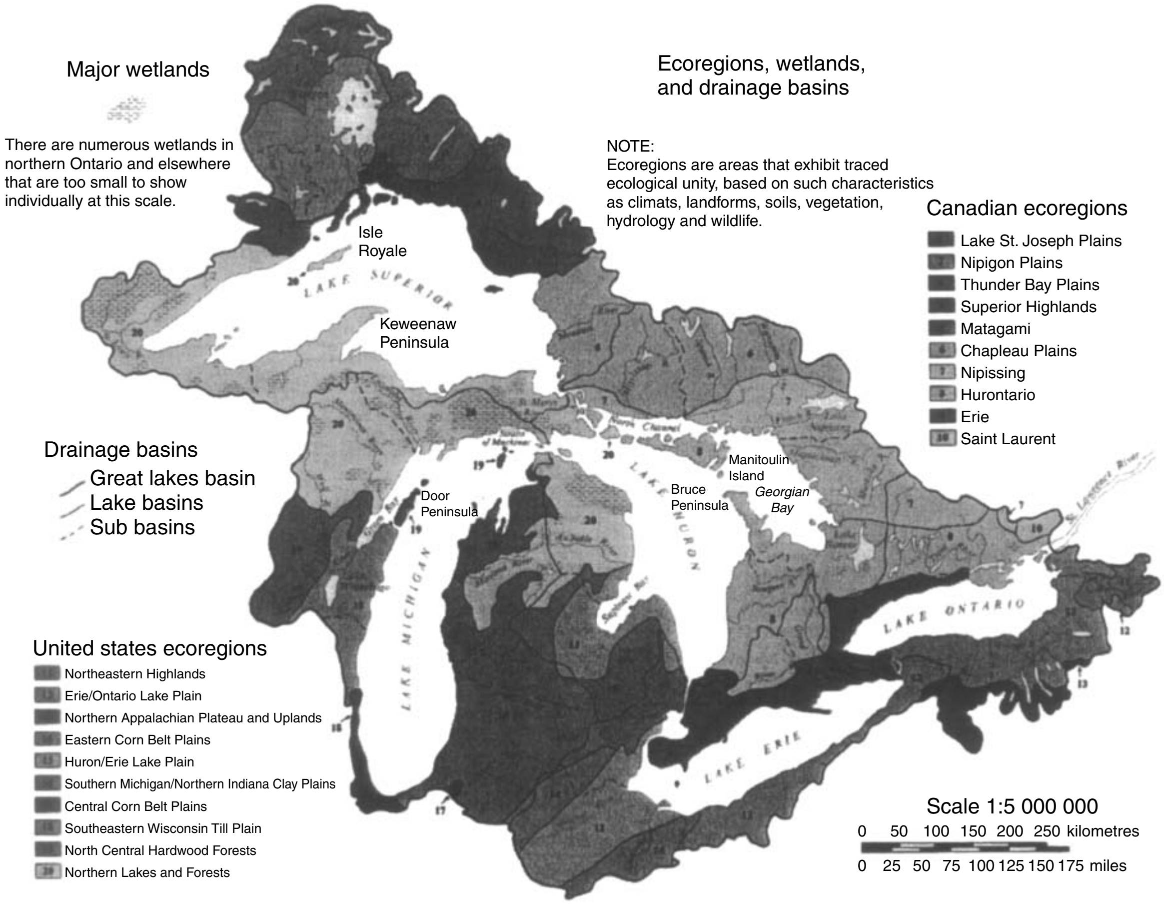

Figure 1. Great Lakes ecoregions, indicating unique features of the basin formed by geological forces, such as low-lying wetland areas.

Wisconsin icesheet retreat, lake drainage had spilled through Illinois River Valley toward the Mississippi, the Hudson River Valley, the Trent River, and the Ottawa River Valley. Uplift continues today, although more slowly.

Deposits from the glacial scour and retreat are largely poorly sorted mixtures of boulders, gravel, sand, silt, and clay, referred to as glacial drift and till. Exposed consolidated rocks of limestone, shale, sandstone, halite, and gypsum formed during the preglacial period of inland seas from sedimentary deposits are found in many river and glacial carved valleys; outcrops of igneous and other morph rocks are also revealed. Landforms include sand dunes; flat till plains; linear moraines; ramp shaped drumlins; and meltwater deposited eskers, kames, and outwash. Relatively low relief defines the drainage basin, but the lakes feature remarkable depth gradients, widths, and lengths, which are recorded in Table 1. It is apparent that extremes exist among the lakes; Erie and Superior have nearly a 350 m difference in maximum depth. The lakes are stunning, but the wetlands occupying the low gradient lands, as well as the isolated gorges and ravines carved by active rivers, also provide both ecological and geologic complexity.

Table 1. Physical Features of System

Category	Superior	Michigan	Huron	Erie	Ontario
Elevation at low level, m	183	176	176	173	74
Length, km	563	494	332	388	311
Breadth, km	257	190	245	92	85
Average depth, m	147	85	59	19	86
Maximum depth, m	406	282	229	64	244
Volume, km^3	12,100	4,920	3,540	484	1,640
Water area, km^2	82,100	57,800	59,600	25,700	18,960
Land drainage area, km^2	127,700	118,000	134,100	78,000	64,030
Total basin area, km^2	209,800	175,800	193,700	103,700	82,990
Shoreline length, km	4,385	2,633	6,157	1,402	1,146

HYDROLOGIC FLOWS AND CLIMATE

Great Lake flows are from Lake Superior and Michigan into Huron, Huron into Erie, and Erie into Ontario, which flows into the St. Lawrence Seaway through Canada and into the Atlantic. Numerous smaller inland lakes, rivers, and wetlands drain into the Great Lakes system, including some famous waterbodies such as the Finger Lakes, Lake Nipigon, Lake St. Clair, the Niagara River, and the Montezuma Wetlands of New York. Alterations to flow have occurred, such as diversions of Long Lac and Lake Ogoki to capture water from the Hudson Bay watershed, dams to drown the Lachine Rapids, and reversing the Chicago River to keep sewage from mixing with drinking water. Interesting dynamics of flow within the Great Lakes system include fall and spring turnover that breaks the Lake's thermocline strata and mixes hypolimnion and epilimnion waters, seiche, or sloshing of lake water height from shore to shore, an oscillation due to the large fetch of lake surface and strong winds piling water against the coast, and devastating surface waves caused by gale force winds. The Welland Canal, a unique man-made feature of the system, forms a navigable channel between Lakes Erie and Ontario, which are vertically separated by nearly 100 m, a drop that is dramatically illustrated by the Niagara Falls.

Surface water resources are enormous in volume and area; the precipitation, evaporation, and runoff components of the hydrologic cycle maintain this system that has a large flux rate, as shown in Table 2. Runoff into the lakes is smallest for Erie and Ontario, almost half that for Superior and Huron, but is augmented by the inflows from those same upper Great Lakes. The retention times follow a logarithmic decrease of 100-year intervals from the headwater Lakes of Superior and Michigan, which have large volumes and small inflows, to the yearly intervals for Erie and Ontario, which have enormous inflows and relatively small volumes. Fluctuation in lake levels has dramatic impacts on the internal and surrounding ecosystems and despite management efforts, is largely regulated by seasonal and annual (e.g., persistent) weather patterns. Groundwater reserves in the Great Lakes are present in multiple locations and at many different depths and flow rates. Besides their role in providing drinking water for numerous communities, they provide important base-flow release to keep the Lakes' feeder rivers perennial and flowing year-round.

Wetlands in the Great Lakes region are extensive and are represented by four distinct types classified as swamps, marshes, bogs, and fens. Swamps are distinguished by the presence of trees and shrubs and significant annual water cover. Marshes are known for their thick stands of aquatic plants rather than trees and like swamps, are most likely located in the southern Lake region. Bogs are most well known for their stagnant acidic water and slowly decaying, thick mats of sphagnum mosses; fens are filled by slowly moving, less acidic groundwater, dominated by sedges and grasses, but may have stunted shrubs and trees. Many of these wetlands are located along the coastal area, including the UNESCO Biosphere Reserve at Long Point on the north shore of Lake Erie.

Weather patterns and their long-term average condition known as climate have dynamic annual fluctuations in temperature and precipitation due to the Lakes' middle latitude location, which is a global meeting area for continental polar and maritime tropical air masses pushed by the jet stream. January mean daily air temperature ranges from −20 to 0 °C, and July mean daily air temperature ranges from 7.5 to 25 °C. Mean annual precipitation ranges from 600 to 1300 mm, and snowfall ranges from 150 to 350 mm. The mean annual frost-free period ranges from 220 days to a period less than 2 months north of Lake Superior. Rossby long waves perturbed by shortwave upper airflow patterns trigger the deepening of low-pressure systems, termed cyclogenesis, which results in open wave cyclones defined by the overrunning of warm fronts by cold fronts. Several such open wave cyclones can cover the Great Lakes region, and depending on the season, this can result in synoptic snowfall and rainfall as the cold, dry air mixes with the saturated Gulf Coast air.

Several other weather patterns are also familiar to the region. When the semipermanent Bermuda high-pressure system moves off the southeastern coast after autumn's first frost, the area may experience an Indian Summer as warmer Gulf Coast air is pushed north. Winter weather patterns common to this region include the severe lake effect snows, caused by cold air warming and absorbing lake water, and then cooling on the leeward shore and precipitating, and the northeastern (Nor'easter) winds caused by a strong high-pressure system in Quebec that bring Atlantic Ocean moisture into the region, with ice storms and wet snows. Spring weather is famous for its oscillation between extreme heat and cold, again as a result of the dynamic middle latitude looping of the Gulf Stream that can drag very warm spring Gulf Coast air northward or extremely cold, spring Hudson Bay air southward. Summer weather in the Great Lakes is moderated from extreme heat by the lake temperatures, which may set up lake–land breezes,

Table 2. Water Budgets for Great Lakes Water System ($m^3 s^{-1}$)

Category	Superior	Michigan	Huron	Erie	Ontario
Runoff to lake	14,000	11,000	14,000	7,000	9,000
Precipitation to lake	21,000	15,000	15,000	7,000	5,000
Evaporation from lake	15,000	10,000	13,000	7,000	3,000
Flow in by diversions	2,000	−1,000	0	−2,000	2,000
Flow in from upper lake	0	0	37,000	53,000	58,000
Flow out to lower lake/river	22,000	15,000	53,000	58,000	71,000
Retention time (yr)	191	99	22	2.6	6

as well as provide water vapor for convective afternoon thundershowers.

LAND AND WATER USE

Settlement in the Great Lakes extends back to Native American occupation following the Wisconsin icesheet retreat 10,000 years ago; the estimated population during the sixteenth century was between 60,000 and 117,000. The famous Iroquois confederacy of six tribes (Cayuga, Mohawk, Onondaga, Oneida, Seneca, and Tuscarora) occupied much of the New York lands draining into Lake Ontario. European settlers arrived in the early seventeenth century; boats were initially limited to the St. Lawrence Valley due to severe rapids. By 1680, French fur trading in Canada was well established and protected by army forts; by 1730, the British Fort Oswego protected colonies along the southern shore of Lake Ontario. The British captured many French forts by 1759, and the British retained control of the Great Lakes during the American Revolution, granting lands to nearly 40,000 loyalists who fled New England. The 2-year War of 1812 between America and Britain for the Great Lakes was claimed a victory by both sides; America secured the southern coast, and Canada retained its northern control.

Approximately 1 million people in the western Great Lakes region, who sowed crops, worked dairy, and built canals to ship exports and bring new products, firmly established an agricultural heritage for the region by the mid-1800s. Wetlands were drained, soils eroded by denuded lands, rivers were clogged by sediment, and many fish spawning habitats were destroyed. Soil conservation measures have helped stabilize soils, and chemical fertilizers, pesticides, herbicides and irrigation have been recommended for increasing production. Chemicals such as excess phosphorous and historic use of DDT, along with modern row crop monoculture and hormone disrupting chemicals, have significantly impacted the ecosystem. Agricultural production brings vital income and needed food to the region and occupies between 3 and 67% of basin land and upward of 33% of shoreline land, as reported in Table 3.

Logging and forestry occupy even greater amounts of land, from 21 to 91% in the basin (see Table 3), and commercially began in the 1830s by targeting the light and strong white pine, desired for ship and building construction. Forest industrial uses, primarily pulp and papermaking, but also tannins, created additional revenue and development of the region. Again, however, several negative environmental consequences of this industry, such as water pollution and extensive area, high grade, clear cutting have since been identified. Now it operates at lower volumes and uses best management practices to target sustainable operations. Commercial fisheries were also of major importance; in the late 1890s, harvests had grown by 20% per year since 1820, and annual harvests of 65,000 tonnes were recorded. Today, lake trout, sturgeon, and lake herring have largely been replaced by introduced species of smelt, alewife, splake, and Pacific salmon; average harvests have been around 50,000 tonnes. Recreation boating, swimming, and ecotourism along with sport fishing also represent significant uses and income on the Lakes.

Table 3. Basin and Shoreline Land Use Type (%)

Category	Superior	Michigan	Huron	Erie	Ontario
Basin Land Use					
Agricultural	3	44	27	67	39
Residential	1	9	2	10	7
Forest	91	41	68	21	49
Other	5	6	3	1	5
Shoreline Land Use					
Agricultural	n/a	20	15	14	33
Residential	n/a	39	42	45	40
Recreational	n/a	24	4	13	12
Commercial	n/a	12	32	12	8
Other	n/a	5	7	16	7

Industrial and commercial development cultivated along the shores of the Great Lakes provided waterpower and waterways for shipping, through which in the 1820s, the Erie Canal and Welland Canal were built. By 1960, water level controls allowed operation of the St. Lawrence Seaway, and the region's principal commodities, iron ore, coal, and grain, were heavily shipped between lakes and to the Atlantic Ocean. Industrial manufacturing, aggressively developed in this region, fostered the growth of many important cities such as Duluth, Marquette, Chicago, Detroit, Toledo, Hamilton, Cleveland, and some towns that had less waterborne commerce such as Rochester, Buffalo, and Toronto in Lake Ontario. Polluted discharges from these industrial centers caused a significant negative impact on the Great Lakes, further discussed in the next section. Episodes of cholera in 1854 and typhoid fever in 1891 were the impetus to reverse the flow of the sewage-spoiled Chicago River away from Lake Michigan, a drinking water source; Hamilton, faced with contaminated lakeshore water, installed deeper intakes into Lake Ontario.

Aquatic ecosystem resources and dynamics are extensive and complex. Driving the food web is sunlight, supplying energy for autotrophs such as phytoplankton, algae, and plants, which in turn feed the heterotrophs, from zooplankton to herbivore fish, which then feed the larger ecosystem carnivore and omnivore predators. Fish species for this region include alewife*, Atlantic salmon, bloater, bowfin, brook trout, brown trout*, burbot, carp*, Chinook salmon*, Coho salmon*, freshwater drum, lake herring, lake sturgeon, lake trout, lake whitefish, longnose sucker, muskellunge, northern pike, pink salmon*, pumpkinseed, rainbow smelt*, rainbow trout*, rock bass, round goby*, round whitefish, ruffe*, sea lamprey*, smallmouth bass, walleye, white bass, white perch*, white sucker, and yellow perch, where * indicates invasive species. This process continues upward to the amphibian, avian, and terrestrial herbivores and predators and is repeated in many adjoining wetlands and rivers. Popular avian and terrestrial animals in this region include cormorants, gulls, waterfowl, eagles, snapping turtles, whitetailed deer, moose, beaver, muskrat, fox, wolf, and black bear. Decomposition returns the organic nutrients to the mineral soil and inorganic system, an important process carried out by bacteria and fungal organisms. Natural and anthropogenic supplies of essential micro- and macronutrients are critical for the proper biological functioning of this food web, but the introduction of metals, industrial chemicals, and excess nutrients has caused problems, including bioaccumulation and biomagnification, some of them are discussed in the next section.

WATER RESOURCE MANAGEMENT

Pioneering international water resources management has been forged along the shores of the Great Lakes, possibly a challenge best met by two countries that have relatively stable political and economic systems. In 1909, the Boundary Waters Treaty established binational guidance on regulating water levels and flows, which have been the responsibility of the International Joint Commission (IJC), an agency of and for Canada and the United States interests. Separate legislation in 1972, called the Great Lakes Water Quality Agreement (GLWQA), established guidance on regulating water quality, again overseen by the IJC and implemented by numerous national, provincial, state, and small government agencies. Dams are operated on the St. Marys River, Niagara Falls, and the St. Lawrence and only directly affect Lake Superior water levels. These few control points prevent the IJC from dampening all fluctuation; municipal, manufacturing, and power production withdrawls (see Table 4) are affected by and affect levels. Weather exerts a larger control on water level variability, but the IJC has concluded that costs to engineer controls outweigh benefits. Beginning in 1973, the IJC has included erosion control within its water level management objectives of hydropower and navigation, and today the IJC has established new water level guidance to allow greater fluctuation, principally in Lake Ontario, for ecosystem health.

Water quality management under the GLWQA, signed in 1978, began with an analysis of the mass balance loading of phosphorous and other pollutants, leading the way for major subsequent advances in setting target loading limits and current total maximum daily load analysis. Nutrients and carcinogens have received the greatest research and management focus, but the initial analysis of hormone disrupting chemicals in the Great Lakes food chain has identified new concerns and represents a growing issue. An ecosystem approach was adopted by the GLWQA, which was implemented through more expansive Lakewide Management Plans (LaMPs), but areas of concern that contained extreme pollution were also targeted by Remedial Action Plans (RAPs). Innovative water resource management methods developed by the IJC and national agencies such as the Environmental Protection Agency and Environment Canada that encouraged citizen involvement in seeking and implementing solutions have since been studied and shared within the UNESCO Hydrology for the Environment Life and Policy (HELP) Initiative, a needs driven hydrology policy-management-science program to which Lake Ontario belongs. Advanced hydrologic research, such as remote sensing driven modeling and monitoring, coupled with consensus based watershed management, will ideally answer pressing water resource issues.

Table 4. Municipal, Manufacturing, and Power Production Water Withdrawals (10^6 m^3 yr^{-1})

Category	Superior	Michigan	Huron	Erie	Ontario
		Withdrawals			
Municipal	98	2,622	384	2,685	927
Manufacturing	1,133	8,608	2,158	9,820	2,935
Power production	740	12,131	4,852	12,791	13,282
		Consumptive Use			
Municipal	18	169	170	189	152
Manufacturing	71	785	89	1,409	125
Power production	9	214	62	178	174

New management issues have replaced the traditional water chemistry concerns and include the response to several invasive species. Of the many invasives, including crustaceans, fish, mollusks, and plants, some new introductions such as the spiny water flea and zebra muscle are wreaking havoc on industrial to recreational uses of the Great Lakes. Ballast dewatering operations of commercial shipping vessels have been identified as the major pathway for invasives, and U.S. Coast Guard regulations prohibiting this operation are in place. Other challenges include the restoration of coastal wetlands; many of them were destroyed by the removal of extreme peak and low water levels that are responsible for the killing back of monocrop grasses that outcompete the more complex ecosystems needed for proper fish spawning. Restoration of urban streams, including the separation of storm and sanitary sewers, will also be required to regain healthy tributary and headwater spawning areas for Lake species. Additionally greater management control is needed for agricultural, suburban, and urban management of nonpoint source runoff, currently the biggest source of sediment, nutrient, metal, organic, and chemical pollutant loading, now that point source discharges have largely been controlled. A final note must address the recent climate change analysis performed as part of the U.S. Global Change Research Program, which predicts a most likely scenario of lower lake levels due to the establishment of greater evaporation rates, which would potentially have devastating impacts on many recreational, commercial, industrial, navigational, and other Great Lakes activities.

READING LIST

Government of Canada and US Environmental Protection Agency. (1995). *The Great Lakes: An Environmental Atlas and Resource Book*, 3rd Edn.

U.S. Global Climate Change Research Program. (2000). *Preparing for a Changing Climate: Great Lakes Overview*. Great Lakes Regional Assessment Group.

GREENHOUSE GAS EMISSIONS FROM HYDROELECTRIC RESERVOIRS

Nathalie Barrette
Laval University
Québec, Canada

The presence of certain trace gases (e.g., CO_2 and CH_4) in the atmosphere is related to human activity (fossil fuel combustion, deforestation, and intensive agriculture), and to the biogeochemical processes that occur in natural environments (tropical and boreal forests and tundra). For many years now, scientists have been trying to determine the importance of natural environments as sinks for or sources of these trace gases (1–3). The terrestrial biosphere constitutes a major carbon sink (4); more than a third of the anthropic CO_2 emissions are fixed there (5). However, certain changes in land use (e.g., deforestation and the draining of marshes for agriculture) can lead to changes in the way greenhouse gases (GHGs) are produced or fixed by modifying the physicochemical characteristics of these soils. In the medium and long term, these changes are likely to invert the carbon sink capacity that has been attributed until now to certain natural environments (e.g., boreal regions). This article examines how one of these land-use changes, the creation of hydroelectric reservoirs, significantly modifies the carbon cycle in natural environments.

STATE OF RESEARCH

The scientific community's interest in GHG production by hydroelectric reservoirs originated, in part, from the controversy launched by the publication of a paper by Rudd et al. (6). In this publication, the authors concluded that "*. . .per unit of energy produced, greenhouse-gas flux to the atmosphere from some hydroelectric reservoirs may be significant compared to greenhouse-gas emission by fossil-fuelled electricity generation. . .*". This conclusion, based on a weak sampling of fluxes and a method of calculation described as inadequate by certain scientists, created keen interest among researchers (7,8) and the scientists affiliated with hydroelectric companies (9–11). Above and beyond the exercise of comparing energy production methods, one fact remains: that hydroelectricity can no longer be described "as clean energy" with respect to GHG emissions (12–17).

Boreal Reservoirs

One of the studies that has contributed greatly to furthering our knowledge focused on boreal reservoirs in central northern Québec (18).

It was initially assumed that CO_2 and CH_4 fluxes measured on the surface of the reservoirs came mainly from the degradation of organic material in the flooded soil (12). The first research endeavors to understand the processes involved were devoted to surface sampling of GHG emission. Parallel to sampling these surface fluxes, certain morphological parameters (depth of the water column, type of flooded soil), physicochemical (temperature of water), and meteorological data (wind speed, temperature) were also measured to determine their effects on the variability of these fluxes. Over the years, a significant amount of data was collected concerning GHG emissions from hydroelectric reservoirs in northern Québec. These data, as well as complementary work in microbiology, geochemistry, and other scientific fields, have made it possible to understand better the complex series of processes that control the annual carbon balance in the reservoirs. These processes include the production of GHGs through the degradation of flooded organic matter (soil and vegetation), the production/fixation of GHGs by bacteria and plankton in the water column, and the loss of GHGs to the atmosphere by diffusion, convection, and effervescence (production of gas bubbles) (19).

The processes responsible for producing GHGs in reservoirs and, in particular, production from the degradation of organic material in flooded soil are still being studied. Moreover, we can make no definitive qualitative or quantitative statements about the nature and importance of the various processes involved in producing CO_2 and CH_4. The original assumption about the origin of these gases supposed that the flooding that occurs when a hydroelectric reservoir is created starts significant bacterial activity that degrades the flooded terrestrial organic matter. This activity leads to mineralization of the organic matter and the liberation of CO_2 in the water column. There is likewise the creation of anoxic zones in the flooded soils or in the hypolimnion that supports methanogenesis (formation of CH_4). The carbon source at the origin of the reservoirs' GHG emissions is thus located mainly in the flooded soils (12).

A subsequent study by Weissenberger et al. (20) concerning the total quantity of submerged carbon versus the issue rates of CO_2 and CH_4 at the surface of the reservoirs anticipated that after 100 years, the submerged organic matter would be completely degraded. In practice, however, we are far from observing such a phenomenon. The Cabonga reservoir (central northern Québec), which is more than 70 years old, shows few signs of advanced degradation. A scuba diver observed that certain submerged conifers were still intact after being submerged all these years (9). Moreover, one statistical study by Duchemin et al. (17) showed that there was little difference in CH_4 fluxes for reservoirs of different ages and that there was only a very slight decrease (less than 1% per year) of CO_2 fluxes according to the age of the reservoirs. According to Weissenberger and collaborators (20), the degradation of organic material in the flooded soil cannot solely explain CO_2 and CH_4 fluxes observed at the surface of the reservoirs.

Migrating carbon (from the drainage basin) in transition in the reservoir seems to constitute another considerable source of the carbon involved in GHG production. In fact, new avenues of research suggest that GHG production initially comes from the degradation of organic material in the flooded soil. After a few years, GHG production seems to remain stable due to a major change in the trophic state of the reservoir. Reservoirs go from an autotrophic state (self-sufficient in terms of energy; photosynthetic production exceeds the loss due to respiration) to a heterotrophic or allotrophic state (dependency on an external source of energy, where photosynthetic production is lower than respiration). This change leads to the assimilation of migrating, allochthonous carbon by living organisms in the reservoir. The bacteria are probably responsible for transforming this carbon into GHGs, and the process seems to occur in the entire water column. All of this is still hypothetical and constitutes a new research project in itself (21).

With regard to the bacterial activity in the water column, experiments in microbiology by Duchemin and his collaborators allowed them to observe that a significant portion of the CH_4 was oxidized before it even reached the reservoir surface (12). The CH_4 is oxidized in the first 25–50 centimeters above the flooded soil. The rate of oxidation is significant in flooded peat bogs and very low in flooded podzols. According to Duchemin et al. (19), two conditions are necessary for methane oxidation: a dissolved oxygen concentration higher than $1\ \mathrm{mg\,L^{-1}}$ and a dissolved inorganic nitrogen concentration lower than $20\ \mu\mathrm{M}$. This process is at work mainly during the ice-free season because, the weak dissolved oxygen concentrations generated an anoxic zone in the winter which largely reduced methane oxidation. However, at the present time, very few observations are available to quantify the process of methane oxidation adequately.

Planktonic activity did not appear to be a determining factor in the GHG concentration of the hydroelectric reservoirs of central northern Québec. Though biological activity (fixation-respiration) had suggested that the maximum dissolved CO_2 concentration occurred at night, a study by Canuel et al. (1997) has shown that the maximum concentration actually occurs during the day. The authors thus deduced that the evolution of the GHG concentrations in the reservoir is controlled by processes other than planktonic activity.

Conversely, the processes responsible for transporting dissolved gases in the water column are relatively well understood. It is assumed that most of the transport occurs vertically, that is, there is little horizontal current and homothermy. The vertical transport of dissolved gases results from turbulent diffusion and convection. The turbulent diffusion is induced by mechanical forcing (e.g., wind) which acts in a stable density gradient. For its part, the convection occurs following density anomalies that cause an unstable profile. This can ensue from the deep absorption of solar radiation or from the loss of sensible heat or latent heat through evaporation or through ice melting. These two transport mechanisms (turbulent diffusion and convection) vary in intensity according to the period of the year and weather surface conditions. For example, in autumn, major convective movements are responsible for transporting a significant quantity of dissolved gas toward the reservoir surface due to the static instability generated by surface cooling (23). According to Duchemin and Levesque (24), effervescence (i.e., release by bubbles) constitutes another transfer mechanism of GHGs in a reservoir. However, the contribution of this mechanism is only significant for CH_4 in the very shallow parts of reservoirs (24). For CO_2, it is estimated that this transport mechanism contributes less than 1% of the total annual emissions (25).

Kelly et al. (26) conducted a study in northwest Ontario in which GHG emissions were measured before and after the controlled flooding of a peat bog and pond. It was observed that the flooding of the peat bog greatly modified the GHG emission balance. Before flooding, the peat bog constituted a CO_2 sink and a weak source of CH_4. After flooding, it became a significant source of CO_2 and CH_4. According to Kelly et al. (26), this CO_2 status inversion is partly caused by the death of almost all the vegetation in the natural environment, making it incapable of fixing CO_2, and by the decomposition of the submerged organic matter. Furthermore, the reinforcement of its status as a CH_4 source is due to an increase in the anoxic decomposition of the vegetation and to a reduction in the consumption of CH_4 (oxidation) by oxidizing bacteria. It would seem that the first mechanism is quantitatively more significant than the second.

Tropical Reservoirs

Reservoirs in tropical regions have also been studied. The first studies indicated that tropical reservoirs constitute even more significant GHG sources than northern reservoirs (15,16,24,27). According to Fearnside (27), the results obtained until now even call into question the idea that hydroelectricity in the Amazon region will produce less GHGs than the production of the same quantity of energy by other fossil fuels. According to Pearce (15), the Amazon basin brings together ideal conditions for producing methane; little dissolved oxygen, relatively high temperatures, and high concentrations of nutrients. This author also states that all the hydroelectric reservoirs of the world produce the equivalent of 7% of the total production of CO_2. Duchemin et al. (24) point out that Amazonian reservoirs produce ten times more CO_2 and three times more CH_4 than the hydroelectric reservoirs of northern Québec.

CONCLUSION

The biosphere and, in particular, the soils contain substantial carbon reserves. Flooding following the creation of a hydroelectric reservoir can therefore significantly modify this carbon reserve. The boreal regions and their soils have acted as a carbon sink since the last glaciation. This capacity can, however, be reversed. As for the tropical regions, their GHG emissions can be so significant as to call into question hydroelectric development.

From the point of view of the international effort to reduce GHGs, it would appear essential to compare the annual total GHG emissions arising from different energy forms; an exercise that has been carried out by several scientists (28–30). Duchemin (29) has shown that the boreal reservoirs have annual equivalent CO_2 emissions lower than those emitted by power stations fueled by coal or natural gas when producing an equivalent quantity of energy (see Fig. 1). On the other hand, he has also shown that certain tropical reservoirs can emit quantities

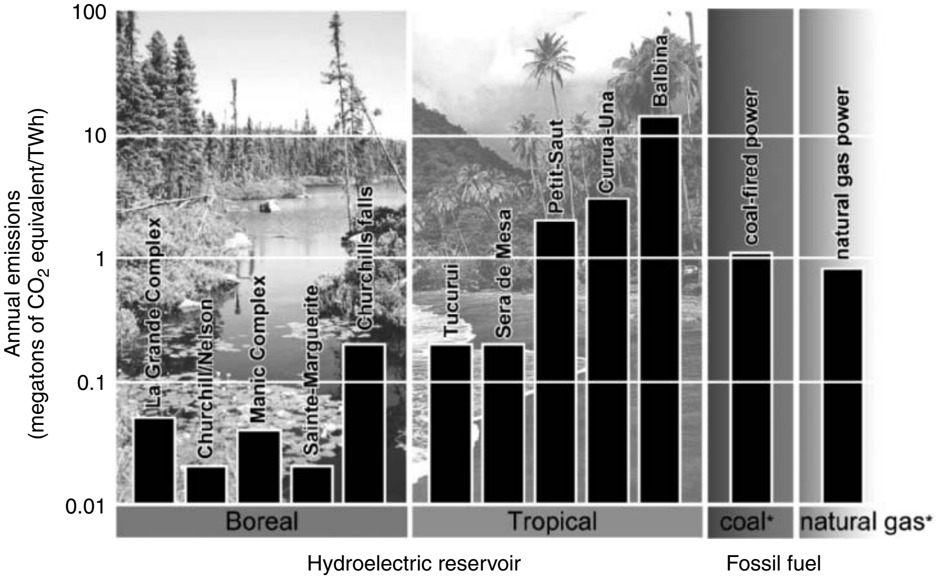

Figure 1. Greenhouse gas emissions from hydropower and fossil fuel power (29).

*average value

of equivalent CO_2 up to 14 times higher than the emissions from a coal-fired power station producing a comparable quantity of energy (see Fig. 1). It is particularly worth noting that shallow tropical reservoirs (6 m and less) emit up to two to three times more CO_2 per square meter than deep reservoirs (25 m and more). Thus, the site for constructing a hydroelectric reservoir in a tropical area can be a determining factor in its emission balance. The creation of hydroelectric reservoirs in zones of fairly flat topography should be avoided because the flooded surface areas will be substantial and will provide ideal conditions for developing the bacteria responsible for these emissions.

Duchemin insists that the current measurements of emissions in tropical reservoirs are still too limited to offer a clear answer to the question of energy choices. Moreover, he adds that future studies should introduce the concept of the net before-and-after-flooding balance to evaluate truly the changes in the carbon sink or source capacity of a given territory.

BIBLIOGRAPHY

1. Moore, T.R. and Knowles, R.K. (1987). Methane and carbon dioxide evolution from subarctic fens. *Can. J. Soil Sci.* **67**: 77–81.
2. Oechel, W.C. et al. (1993). Recent change of Arctic tundra ecosystems from a net carbon dioxide sink to a source. *Nature* **361**: 520–523.
3. Jarvis, P.G. (1995). The role of temperate trees and forests in CO_2 fixation. *Vegetation* **121**: 157–174.
4. Fan S. et al. (1998). A large terrestrial carbon sink implied by atmospheric and oceanic carbon dioxide data and models. *Science* **282**: 456–457.
5. Keeling, R.F., Piper, S., and Heimann, M. (1996). Global and hemispheric CO_2 sinks deduced from recent atmospheric oxygen measurements. *Nature* **381**: 218–221.
6. Rudd, J.W.M., Harris, R., Kelly, C.A., and Hecky, R.E. (1993). Are hydroelectric reservoirs significant sources of greenhouse gases? *Ambio* **22**(4): 246–248.
7. Svensson, B.S. and Ericson, S-O. (1993). Does hydroelectric power increase global warming? *Ambio* **22**(8): 569–570.
8. Rosa, L.P. and Schaffer, R. (1994). Greenhouse gas emissions from hydroelectric reservoirs. *Ambio* **23**(2): 164–165.
9. Gagnon, L. and Chamberland, A. (1993). Emissions from hydroelectric reservoirs and comparison of hydroelectric, natural gas, and oil. *Ambio* **22**(8): 568–569.
10. Chamberland, A., Bélanger, C., and Gagnon, L. (1996). L'hydroélectricité face aux autres options. *Ecodecision* 56–60.
11. Gagnon, L. (1997). Emissions from hydroelectric reservoirs and comparison of hydroelectricity, natural gas, and oil. *Inter. J. on Hydropower and Dams* **4**(4): 20–24.
12. Duchemin, E., Lucotte, M., Canuel, R., and Chamberland, A. (1995). Production of greenhouse gases CH_4 and CO_2 by Hydroelectric reservoirs of the boreal region. *Global Biogeochem. Cycles* **9**(4): 529–540.
13. Fearnside, P.M. (1995). GHG emissions from hydroelectric reservoirs and the energy policy implications. *Water, Air and Soil Pollution* **00**: 1–27.
14. Yundt, P.E. (March, 1996). *Estimation of Greenhouse gas emissions from Canadian hydroelectric dams and reservoirs*. Environment Canada, Oil, Gas & Energy Division. Research report.
15. Pearce, F. (1996). Trouble bubbles for hydropower-from China to Norway, new hydroelectric schemes are supposed to help cut emission of greenhouse gases. But will they? *New Scientist* **150**: 28–31.
16. Galy-Lacaux, C. et al. (1999). Long-term greenhouse gas emissions from hydroelectric reservoirs in tropical forest regions, *Global Biogeochem. Cycles* **13**(2): 503–517.
17. Duchemin, E., Lucotte, M., and Canuel, R. (1999). Comparison of static chamber and thin boundary layer equation methods for measuring greenhouse gas emission from large water bodies. *Environ. Sci. Technol.* **33**: 350–357.
18. Duchemin, E. (2000). *Hydroélectricité et gaz à effet de serre: évaluation des émissions des différents gaz et identification des processus biogéochimiques de leur production*, Ph.D. Thesis Sciences de l'Environnement, Univ. of Québec at Montréal, Montréal.

19. Duchemin, E., Lucotte, M., and Canuel, R. (1997). *Rapport annuel des activités de recherche 1996: Importance de la saison hivernale dans le bilan des émissions de gaz à effet de serre depuis les réservoirs hydroélectriques en milieu boréal*. *Annual Report*, pp. 58–71, Chaire de Recherche en Environnement Hydro-Québec-CRSNG-UQAM, UQAM, Montréal.
20. Weissenberger, S., Duchemin, E., Houel, S., Canuel, R., and Lucotte, M. (1999). *Rapport annuel des activités de recherche 1998*: Greenhouse gas emissions and carbon cycle in boreal reservoirs, *Annual Report*, pp. 136–151, Chaire de Recherche en Environnement Hydro-Québec-CRSNG-UQAM, UQAM, Montréal.
21. Weissenberger, S. (1999). *Rapport annuel des activités de recherche 1998: Émissions de gaz à effet de serre par les réservoirs boréaux: origine et cycle du carbone*, *Annual Report*, pp. 154–161 Chaire de Recherche en Environnement Hydro-Québec-CRSNG-UQAM, UQAM, Montréal.
22. Canuel, R., Duchemin, E., and Lucotte, M. (January 1997). *Handbook on greenhouse gases sampling and analytical techniques, paper presented at the International workshop on greenhouse gases from hydroelectric reservoirs: production, fate and means of study, Santarém, Bresil*.
23. Barrette, N. (2000). *Réservoirs hydroélectriques du moyen nord québécois: mesures et modélisation numérique des flux et des concentrations de gaz à effet de serre*, Ph.D. Thesis Sciences de l'Environnement, Univ. of Québec at Montréal, Montréal.
24. Duchemin, E. and Levesque, K. (1999). *Rapport annuel des activités de recherche 1998: Analyse statistique des flux d'émission atmosphériques de gaz à effet de serre depuis de grands réservoirs boréaux en période libre de glace*, *Annual Report*, pp. 20–96, Chaire de Recherche en Environnement Hydro-Québec-CRSNG-UQAM, UQAM, Montréal.
25. Duchemin, E., Lucotte, M., Canuel, R., and Chamberland, A. (1997). *Rapport annuel des activités de recherche 1996: Hydroélectricité et gaz à effet de serre*. *Annual report*, pp. 14–26, Chaire de Recherche en Environnement Hydro-Québec-CRSNG-UQAM, UQAM, Montréal.
26. Kelly, C.A. et al. (1997). Increase in fluxes of Greenhouse gases and methyl mercury following flooding of an experimental reservoir. *Environ. Sci. Technol.* **31**: 1334–1344.
27. Fearnside, P.M. (1995). Hydroelectric dams in the Brazilian amazon as sources of greenhouse gases. *Environ. Conserv.* **22**: 7–19.
28. St-Louis, V. et al. (2000). Reservoir surfaces as sources of GHG: a global estimate. *Bioscience* **50**(9): 766–775.
29. Duchemin, E. (2001). *Hydroélectricité et gaz à effet de serre*, Vertigo, 2(1). Available: http://www.vertigo.uqam.ca/.
30. Fearnside, P.M. (2001). Greenhouse gas emissions from a hydroelectric reservoir (Brazil's Tucurui dam) and the energy policy implications. *Water, Air, and Soil Pollution* 1–27.

GULLY EROSION

Wayne D. Erskine
State Forests of New South Wales
Beecroft, New South Wales
Australia

Gullies have formed rapidly in many parts of the world during the last 300 years following the change from traditional agriculture or nomadic indigenous lifestyles to more intensive European-based agriculture. They are deeper erosion channels than rills, but their differentiation is often unclear and arbitrary because of a continuum between the two features. Rills are shallow, ephemeral erosion channels that can be obliterated by standard tillage, alternating cycles of freeze-thaw, or rain-impacted shallow overland flows.

DEFINITION

Brice (1) proposed that gullies are recently extended drainage channels that transmit ephemeral flow, have steep sides, a width > 0.3 m, a depth > 0.6 m, and terminate at a steeply sloping or vertical, upstream primary nickpoint (see next section for definition of within-gully features). Gregory and Walling (2) concluded that gullies are characterized by

- ephemeral streamflow,
- steep sidewalls and nickpoints,
- a depth greater than rills so that they cannot be eliminated by plowing, crossed by a wheeled vehicle, or restored in the short term,
- a V-shaped cross section in fine-grained, resistant subsoils,
- a U-shaped cross section where the soil and subsoil are equally erodible, and
- unconsolidated sediments in the sidewalls.

Attempts to differentiate between gullies and streams on the basis of flow frequency are difficult for two reasons. First, gullies are not restricted to specific climatic regions but have been described in arid, semiarid, subhumid, and humid areas (11). Second, flow regimes vary systematically based on drainage basin area and climate, (31) and gullying is not restricted to specifically sized basins nor to areas of ephemeral flow.

Brice (1) adopted a depth > 0.6 m as representative of a gully, but a depth of 0.3 m has been proposed by soil conservationists (3) and a depth of 0.5 m by Prosser and Winchester (4). Gullies need not be defined by the exceedance of a critical depth because the depth of incision, cross-sectional shape and longitudinal profile of gullies are all dependent on the materials into which the gully is eroded (5–7).

To overcome the above problems, the definition proposed by Schumm et al. (8) is adopted. A gully is a relatively deep, recently formed eroding channel that forms on valley sides and on valley floors where no well-defined channel previously existed. In contradistinction, incised streams form by the erosion of a preexisting river channel (8). Valley-side gullies result in an expansion of the drainage network due to erosion of soil, colluvium, and/or bedrock in hillside hollows. Valley-floor gullies may be either continuous or discontinuous and erode valley-fills of alluvium, swamp deposits, and/or bedrock. Arroyos, which have been the subject of much research in the American southwest, are valley-floor gullies that have steeply sloping or vertical walls of cohesive, fine sediments and flat and generally sandy floors (9).

GULLY EROSION PROCESSES

Gully erosion occurs by a number of erosional processes involving nickpoint or nickzone initiation and upstream retreat (Fig. 1), downstream progressing degradation, and sidewall erosion. Nickpoints are an abrupt change in bed profile that involve a substantial, local increase in slope and retreat by migrating upstream at a rate that declines exponentially over time (5,8). Primary nickpoints mark the upslope limit of gully erosion, whereas secondary nickpoints rework sediment temporarily stored in the gully bed (Fig. 1). Nickzones are steep sections at the headward limit of erosion cycles and also migrate upstream over time without maintaining a vertical face (8). The rotation of nickpoints into nickzones usually occurs when the eroded valley fill is not stratified (6) or when it is uniform, noncohesive material (5). During downstream progressing degradation, bed slope is reduced, the bed often becomes armoured or veneered with gravel, if present, and bed material discharge declines, approaching a minimum value consistent with upstream sediment supply (10). During the evolution of gullies, initial upstream nickpoint migration is often replaced by subsequent downstream progressing degradation, as upstream sediment supply declines due to the exhaustion of most of the stored sediment within the gully network. These erosion cycles can be repeated many times. Sidewall or bank erosion following incision occurs by many processes that operate at different rates (7,11) and can, in some cases, erode more sediment than the initial nickpoint erosion (12).

Gully Sediment Yields

Actively developing gullies generate very high sediment yields that have caused concern to land and river managers throughout the world. However, rates of gully erosion are not included in equations for standard erosion prediction (13). The author's data on sediment yields from gullies of different age in southeastern Australia have been combined with that of Prosser and Winchester (4) to produce Fig. 2. Erskine and Saynor's (14) detailed review of the available sediment yield data for Australian rivers showed that most rivers have yields less than 50 $\mathrm{t km^{-2} y^{-1}}$, and many are also less than 10 $\mathrm{t km^{-2} y^{-1}}$. High yields in excess of 120 $\mathrm{t km^{-2} y^{-1}}$ have been found only for basins that have been extensively disturbed by agriculture or significantly gullied. Figure 2 demonstrates that developing gullies less than 5 years old produce sediment yields greater than 2000 $\mathrm{t km^{-2} y^{-1}}$, but that these high yields rapidly decline during the next 20 years. Williams (15) reported the very high sediment yield of 93,750 $\mathrm{t km^{-2} y^{-1}}$ for a small developing gully in the seasonally wet tropics of northern Australia, where sediment yields are usually very low by world standards (8). Neil and Fogarty (17) used farm dams to determine sediment yields from continuous and discontinuous gullies in comparison to native forest in southeastern Australia. They found that discontinuous gullies, on average,

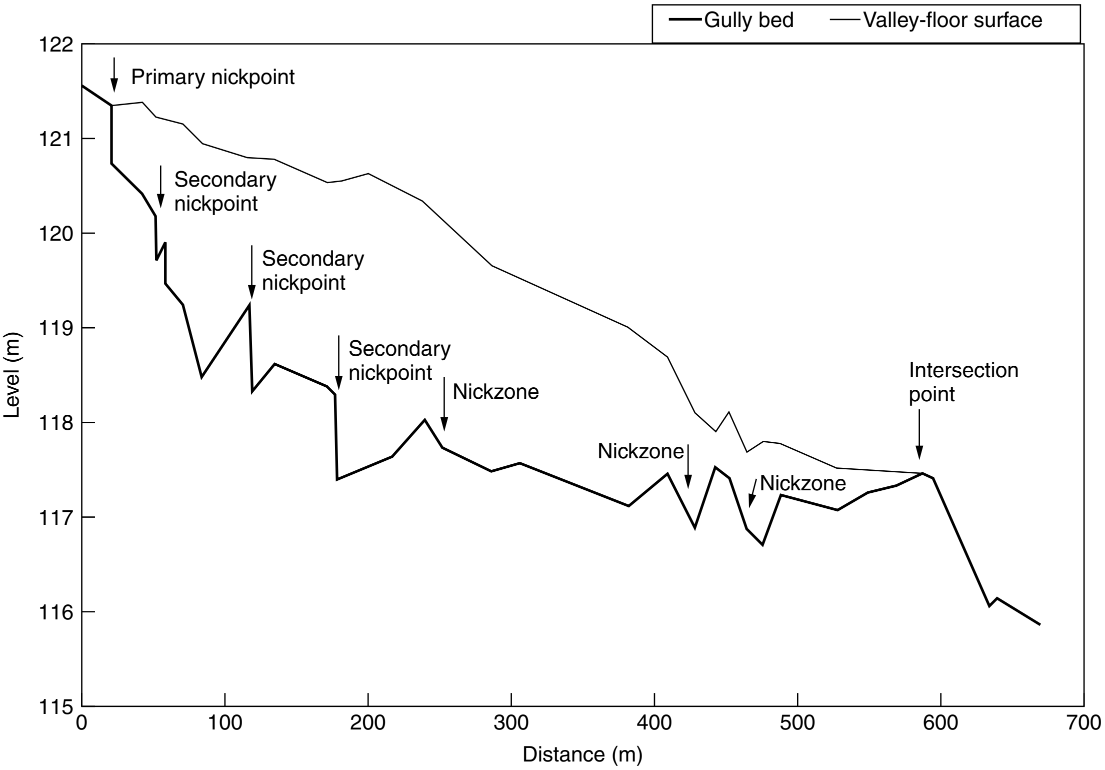

Figure 1. Longitudinal profile of the Dry Arm discontinuous valley-floor gully in the upper Wollombi Brook drainage basin in southeastern Australia, showing some of the main features.

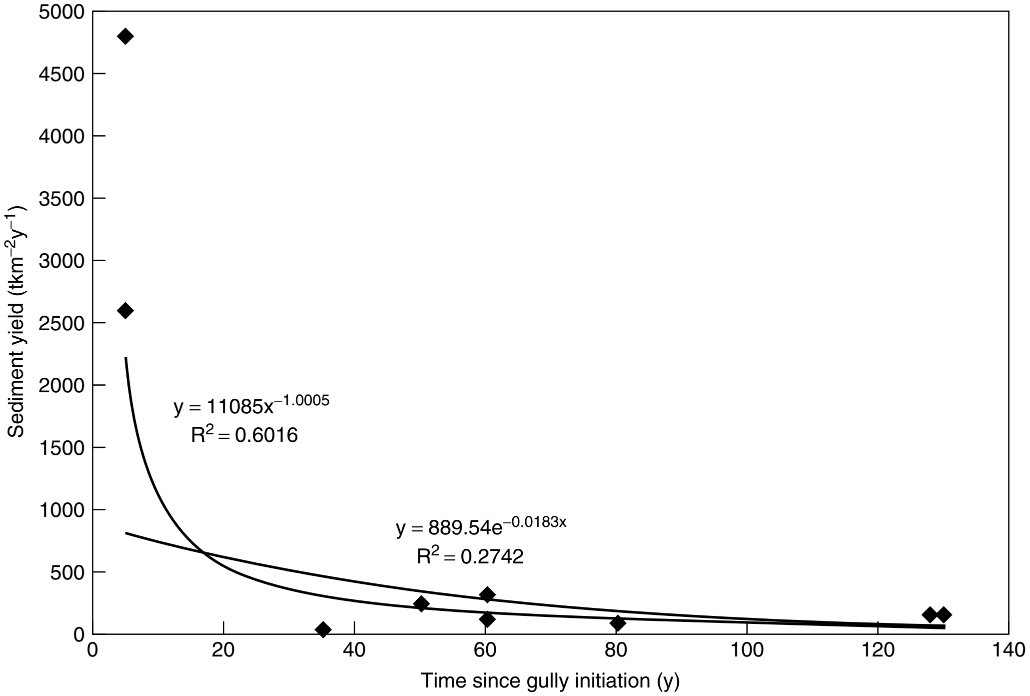

Figure 2. Variations in gully-derived sediment yields with time since gully initiation in southeastern Australia.

produced more than 11 times higher yields than native forests and that continuous gullies generated the highest yields at more than 73 times greater than native forests. Graf (18) demonstrated that an exponential decline in rates of gully erosion occurs in Colorado, consistent with relaxation following initial gully erosion. However, for the data in Fig. 2, a power function is a much better fit than an exponential function. Nevertheless, there is a rapid decline in sediment yield following the first 20 years after gully initiation.

Not all of the sediment generated by gully erosion is transported directly into higher order streams because significant amounts are often stored within the gully and immediately downstream. Melville and Erskine (19) found from detailed field measurements on a compound discontinuous gully system draining a 13.8 km^2 basin in southeastern Australia that 90% of the 190,000 m^3 eroded since European settlement was trapped in the gully bed and small floodplain within the gully as well as in the fan immediately downstream, without ever being exported from the basin. More recent research by the author on other discontinuous gullies in southeastern Australia accounted for all of the gully-eroded sediment which was entirely stored within the gully and the downstream fan. Similarly, Neil and Fogarty (17) estimated that as much as 60% of the material eroded from discontinuous gullies in southeastern Australia was deposited solely in the downstream fan and hence was not transported out of the basin.

Sediment movement in drainage basins is often restricted to clearly defined compartments that may be up to tens of kilometers long but are not linked to higher order streams (19,20). Therefore, gullying often results in reworking sediment from one temporary storage to another without a detectable impact on the sediment yields of large drainage basins (21).

CAUSES OF GULLY EROSION

Cooke and Reeves (9) proposed a deductive model of arroyo formation which maintains that a gully is initiated when the erosion potential of flows (erosivity) locally exceeds the resistance of the surficial materials (erodibility). Therefore, there are two fundamental changes, increased erosivity and increased erodibility, which either together or individually, can lead to gully initiation (9). Erosivity refers to the propensity for flows to detach and remove materials from a given area more rapidly than they are replaced from upstream and is a function of many interrelated hydraulic, channel form, and sediment load variables that are difficult to evaluate precisely (9). Nevertheless, recent research has demonstrated that grass is very important in increasing critical boundary shear stress for soil erosion (22,23). Erodibility is a measure of surficial sediment's resistance to erosion forces and is determined by soil texture, soil aggregation, soil coherence to wetting, soil organic matter content, and vegetative cover. However, there are many trigger mechanisms that can potentially lead to gully formation; they have been classified (Fig. 3) as natural geomorphic processes, gradational geomorphic thresholds, random climatic events, secular rainfall changes, and land use changes (9,24).

Figure 3 shows an expanded version of Cooke and Reeves's deductive model without the explicit links between specific trigger mechanisms and the induced changes in erosivity and/or erodibility. The reason for the omission of these explicit links is to show an uncluttered

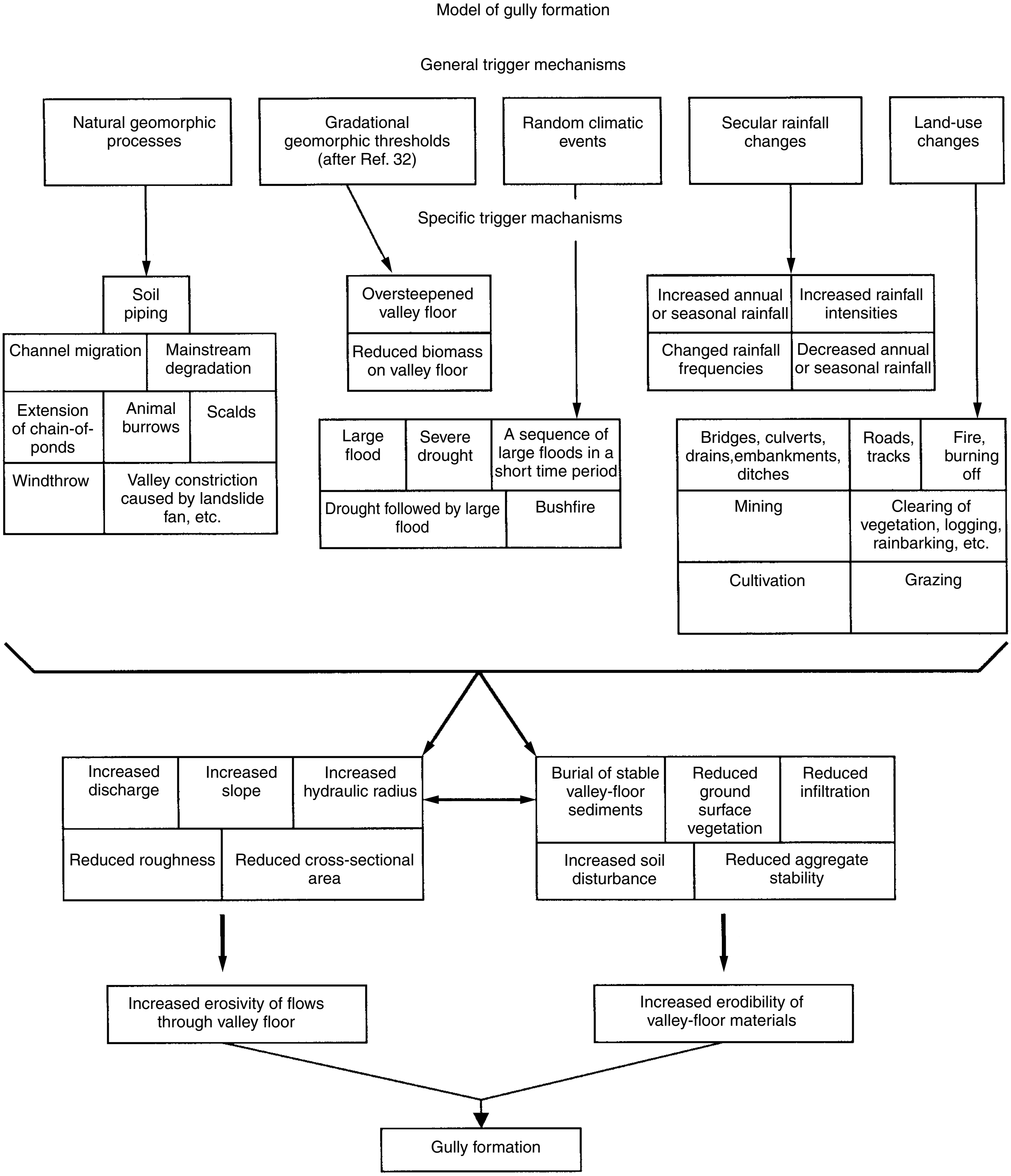

Figure 3. Deductive model of gully formation (modified from Refs. 9 and 24).

hierarchical representation of the interrelationships. Examples of the possible specific trigger mechanisms are outlined in the second tier, and the way they effect erosivity and erodibility is listed in the third tier of Fig. 3.

It is important to emphasize that the results of detailed investigations of gully initiation have found that apparently similar gullies can be formed in different areas as a result of different combinations of initial conditions and specific trigger mechanisms (9). For example, the author found that in upper Wollombi Brook, a 341 km^2 basin in humid southeastern Australia, for the period since first European settlement in the 1820s,

- the timing of valley-floor gully initiation was highly variable, ranging over at least 100 years and only rarely coincided with the period of peak agricultural productivity in the middle of the nineteenth century,
- most gullies were initiated at various times following long periods of sediment storage in valley segments without inflowing tributaries that eventually increased valley-floor slope until it exceeded a stability threshold (gradational geomorphic threshold in Fig. 3),
- some gullies were initiated by a catastrophic flood when up to 508 mm of rainfall were recorded in 12 hours in June 1949 (25) (random climatic event in Fig. 3),
- some gullies were initiated at various times by culvert construction, drains, and ditches in localized parts of valley floors (19) (land use changes in Fig. 3), and
- many valley floors remained ungullied despite extensive valley-floor forest clearing, high intensity agricultural activities involving repeated cultivation, extensive valley-floor roading, repeated high intensity wildfires, and intensive rural subdivision.

Figure 4 depicts the least-squares linear regression relationship between field-surveyed, mean valley-floor slope and basin area for 19 sites (nine ungullied and 10 gullied) in the upper Wollombi Brook basin. For the gullied sites, the valley-floor slope of the presumed initiation point of the gully was surveyed. As this is frequently buried by sand eroded from the upstream gully, multiple excavations were conducted to reveal the original valley-floor surface which was readily identified using the techniques outlined by Happ et al. (26). A locally steeper slope segment was found at the downstream end of each gully, as first reported elsewhere (27). The slope and constant of the regression equation are consistent with reported values elsewhere (13). The regression line effects a clear separation between gullied and ungullied sites (only one gullied site plots below and no ungullied sites plot above the line) and represents a gradational threshold slope for the initiation of gullying in this basin. Similar results have been reported elsewhere (8,13,27,28).

EFFECTS OF GULLY EROSION

The environmental, human, and economic effects of gully erosion documented by many authors (for examples, see Refs. 7–9, 13, 24, 26, 28–30) include all of the following:

1. high soil erosion rates and sediment yields
2. substantial loss of productive farmland
3. deterioration of downstream water quality
4. rapid and massive downstream sedimentation
5. replacement of a relatively stable, muddy substrate by very mobile sand
6. more peaked flood hydrographs
7. higher flow velocities due to decreased resistance to flow and greater flow depths
8. reduced baseflow persistence
9. reduced frequency of overbank flow
10. lower floodplain water tables
11. loss of valley-floor wetlands and seepage zones and their associated vegetation
12. reduced macroinvertebrate, reptile, amphibian and fish biodiversity and some local extirpations by the wholesale loss of aquatic habitats
13. severe farm access and management problems

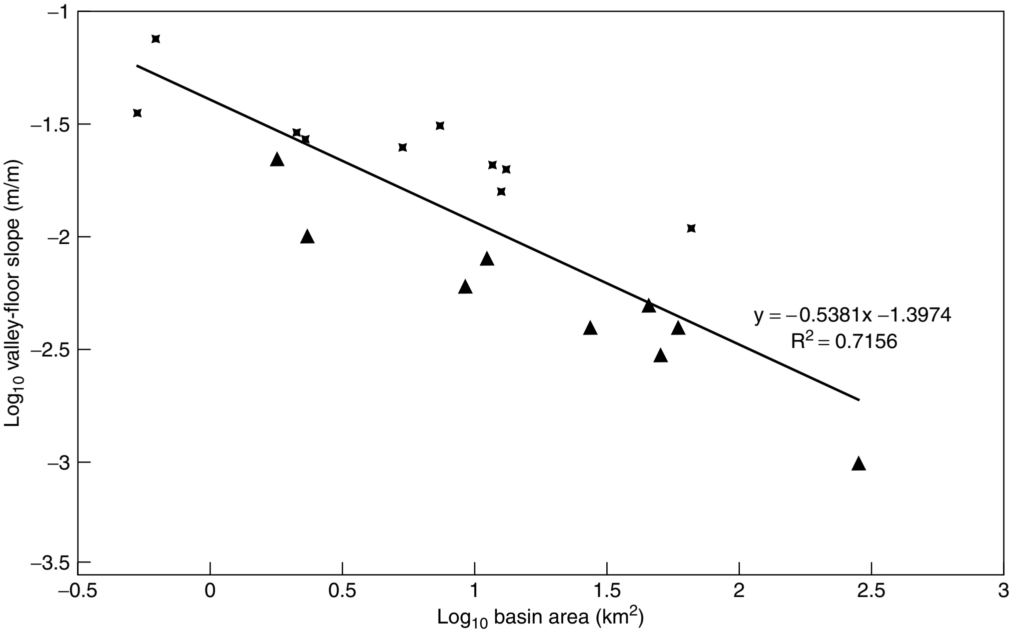

Figure 4. Bivariate plot of mean valley-floor slope versus drainage basin area for various gullied and ungullied sites in the upper Wollombi Brook drainage basin, southeastern Australia. The gullied sites are represented by crosses and the ungullied sites by triangles.

14. flooding of houses by soil-laden water
15. reduced farm incomes due to a decline in land productivity
16. disruption and dislocation of transport routes
17. damage to, and loss of, human structures, such as bridges, pipelines, pump sites and water intakes.

These effects have been recorded at the site of initial gully erosion as well as upstream and downstream (8,9).

BIBLIOGRAPHY

1. Brice, J.C. (1966). *Erosion and Deposition in the Loess-Mantled Great Plains, Medicine Creek Drainage Basin*. U.S. Geol. Surv. Prof. Paper **325-H**, Washington, DC.
2. Gregory, K.J. and Walling, D.E. (1973). *Drainage Basin Form and Process*. Edward Arnold, London, p. 456.
3. Rosewell, C.J., Crouch, R.J., Morse, R.J., Leys, J.F., Hicks, R.W., and Stanley, R.J. (1991). Forms of erosion. In: *Soils Their Properties and Management. A Soil Conservation Handbook for New South Wales*. P.E.V. Charman and B.W. Murphy (Eds.). Sydney University Press, South Melbourne, pp. 12–35.
4. Prosser, I.P. and Winchester, S.J. (1996). History and processes of gully initiation and development in eastern Australia. *Zeits. Geomorph. Suppl. Bd.* **105**: 91–109.
5. Brush, L.M. Jr. and Wolman, M.G. (1960). Knickpoint behaviour in noncohesive material: a laboratory study. *Geol. Soc. Amer. Bull.* **71**: 59–74.
6. Holland, W.N. and Pickup, G. (1976). Flume study of knickpoint development in stratified sediment. *Geol. Soc. Am. Bull.* **87**: 76–82.
7. Ireland, H.A., Sharpe, C.F.S., and Eargle, D.H. (1939). *Principles of Gully Erosion in the Piedmont of South Carolina*. U.S. Dept. Agric. Tech. Bull. No. 633.
8. Schumm, S.A., Harvey, M.D., and Watson, C.C. (1984). Incised channels. *Morphology, Dynamics and Control*. Water Resources Publications, Littleton, CO, p. 200.
9. Cooke, R.U. and Reeves, R.W. (1976). *Arroyos and Environmental Change in the American South-West*. Clarendon Press, Oxford, p. 213.
10. Galay, V.J. (1983). Causes of river bed degradation. *Water Res. Res.* **19**: 1057–1090.
11. Crouch, R.J. (1987). The relationship of gully sidewall shape to sediment production. *Aust. J. Soil Res.* **25**: 531–539.
12. Blong, R.J., Graham, O.P., and Veness, J.A. (1982). The role of sidewall processes in gully development: some N.S.W. examples. *Earth Surf. Processes Landf.* **7**: 381–385.
13. Vandaele, K., Poesen, J., Govers, G., and van Wesemael, B. (1996). Geomorphic threshold conditions for ephemeral gully incision. *Geomorphology* **16**: 161–173.
14. Erskine, W.D. and Saynor, M.J. (1996). *The Influence of Waterway Management on Water Quality with Particular Reference to Suspended Solids, Phosphorus and Nitrogen*. Department of Conservation and Natural Resources, Victoria, East Melbourne, p. 82.
15. Williams, M.A.J. (1976). *Erosion in the Alligator Rivers Area*. CSIRO (Aust.) Land Res. Ser. No. 38, pp. 112–125.
16. Erskine, W.D. and Saynor, M.J. (2000). *Assessment of the Off-Site Geomorphic Impacts of Uranium Mining on Magela Creek, Northern Territory, Australia*. Supervising Scientist Report 156, Supervising Scientist, Darwin, p. 59.
17. Neil, D. and Fogarty, P. (1991). Land use and sediment yield on the Southern Tablelands of New South Wales. *Aust. J. Soil Water Conserv.* **2**: 33–39.
18. Graf, W.L. (1977). The rate law in fluvial geomorphology. *Amer. J. Sci.* **277**: 178–191.
19. Melville, M.D. and Erskine, W.D. (1986). Sediment remobilization and storage by discontinuous gullying in humid southeastern Australia. *Int. Assoc. Hydrol. Sci.* Publ. No. 159, pp. 277–286.
20. Pickup, G. (1988). Modelling arid zone soil erosion at the regional scale. In: *Fluvial Geomorphology of Australia*. R.F. Warner (Ed.). Academic Press, Sydney, pp. 105–127.
21. Trimble, S.W. (1981). Changes in sediment storage in the Coon Creek Basin, Driftless area, Wisconsin, 1853 to 1975. *Science* **214**: 181–183.
22. Prosser, I.P., Dietrich, W.E., and Stevenson, J. (1995). Flow resistance and sediment transport by concentrated overland flow in a grassland valley. *Geomorphology* **13**: 71–86.
23. Prosser, I.P. and Slade, C.J. (1994). Gully formation and the role of valley-floor vegetation, southeastern Australia. *Geology* **22**: 1127–1130.
24. Hendrickson, D.A. and Minckley, W.L. (1984). Cienegas-vanishing climax communities of the American southwest. *Desert Plants* **6**: 131–175.
25. Erskine, W.D. (1996). Response and recovery of a sand-bed stream to a catastrophic flood. *Zeits. Geomorph.* **40**(3): 359–383.
26. Happ, S.C., Rittenhouse, G., and Dobson, G.C. (1940). *Some Principles of Accelerated Stream and Valley Sedimentation*. U.S. Dept. Agric. Tech. Bull. No. 695.
27. Schumm, S.A. and Hadley, R.F. (1957). Arroyos and the semi-arid cycle of erosion. *Am. J. Sci.* **255**: 161–174.
28. Patton, P.C. and Schumm, S.A. (1975). Gully erosion, Northwestern Colorado: a threshold phenomenon. *Geology* **3**: 88–90.
29. Melton, M.A. (1965). The geomorphic and palaeoclimatic significance of alluvial deposits in southern Arizona. *J. Geol.* **73**: 1–38.
30. O'Connor, N.A. and Lake, P.S. (1994). Long-term and seasonal large scale disturbances of a small lowland stream. *Aust. J. Mar. Freshwater Res.* **45**: 243–255.
31. Wolman, M.G. and Gerson, R. (1978). Relative scales of time and effectiveness of climate in watershed geomorphology. *Earth Surf. Processes* **3**: 189–208.
32. Begin, Z.B. and Schuhuam, S.A. (1984). Gradational thresholds and landform singularity: significance for Quaternary studies. *Quat. Res.* **21**: 267–274.

POTENTIAL HEALTH ISSUES ASSOCIATED WITH BLUE-GREEN ALGAE BLOOMS IN IMPOUNDMENTS, PONDS AND LAKES

Barry H. Rosen
U.S. Fish and Wildlife Service
Vero Beach, Florida

INTRODUCTION

Cyanobacteria, also known as blue-green algae, are the most primitive group of algae. Although they are most closely related to other bacteria, they have the same pigment for photosynthesis that other algae and plants have. They are simple but remarkably successful organisms. Individually, they are microscopic; however, large clusters of cells are easily visible as a surface scum (a type of algal bloom) on any stable body of water. Some kinds of blue-green algae produce natural toxins. Unfortunately, no visual technique can distinguish between toxin-producing blue-greens and those that are harmless. Ingestion of these toxins has caused the poisoning of domestic animals, sometimes resulting in death. These include cattle, horses, sheep, hogs, ducks, dogs, and wildlife. Human exposure and toxicity, ranging from intestinal problems to liver damage and fatalities, has only recently been documented.

TOXINS TYPES

The toxins are generally of two types: hepatotoxins and neurotoxins. Hepatotoxins affect the liver, disrupting the important proteins that keep the liver functioning. They generally act more slowly, and a higher dose is needed before death occurs; however, they also may be tumor promoters at low doses. Microcystins, the most common and important family of hepatotoxins, have over 60 variations of a basic cyclic peptide structure. Most of these are hepatotoxins, however, their toxicity depends on the specific amino acids that are part of the basic structure (Fig. 1). Related compounds include nodularin, nodulapeptins, anabaenopeptins, and aeruginopeptin. Another increasingly important hepatotoxin is cylindrospermopsin.

Neurotoxins are normally fast-acting; ingestion of a large dose causes paralysis of skeletal and respiratory muscles which results in death. These compounds are potent alkaloids rather than cyclic peptides. The most common forms are anatoxin, anatoxin-a(S), saxitoxin, neosaxitoxin, and related compounds (Fig. 2).

All of these toxins are difficult to identify and quantify and require detailed laboratory analysis using techniques such as liquid chromatography and mass spectroscopy. Canada, Australia, and Great Britain have developed a guideline level of 1 microgram microcystin per liter of water, or 1 part per billion (1 ppb) for drinking water, and this standard has been adopted by the World Health Organization. During algal blooms, toxin levels can greatly exceed 1 ppb.

SOURCES OF TOXINS

Hundreds of kinds of blue-green algae are known from aquatic habitats; however, only a handful of these are of concern. Some common ones that are known toxin producers include species of *Aphanizomenon, Microcystis*, and *Anabaena* (Fig. 3). Others include species of *Planktothrix (Oscillatoria), Nostoc, Anabaenopsis, Nodularia, Cylindrospermum, Cylindrospermopsis, and Lyngbya.*

Blue-green algae blooms may appear like thick pea soup or grass clippings on the water and are most common in late summer, although blooms can happen in spring or year-round in the warmest latitudes. Pigments and toxins are produced inside the cells and stay there as long as they are alive. When cells break down, usually when a bloom begins to die, their pigments are released into the water and may look like green or blue paint. Bloom die-off also leads to oxygen depletion and associated fish kills in a waterbody. This results from bacterial decomposition of these dead algal cells, which consumes the oxygen. Generally, cooler weather, rainfall, and reduced sunshine will lead to the collapse of an algal bloom. Some blooms last a few weeks; others persist for a few months, depending on environmental conditions. Nutrient enrichment often increases the amount of blue-green algae and may also enhance the dominance of noxious forms.

Control of Blue-Green Algal Blooms

Reducing the load of nutrients from the watershed and/or preventing the release of nutrients from bottom sediments are the primary control measures. In most systems, phosphorus is the critical nutrient. Because in-lake controls treat mostly symptoms, reducing the input of nutrients from the watershed is preferable.

In-lake controls include physical methods such as aeration, circulation, dilution, flushing, and light-limiting

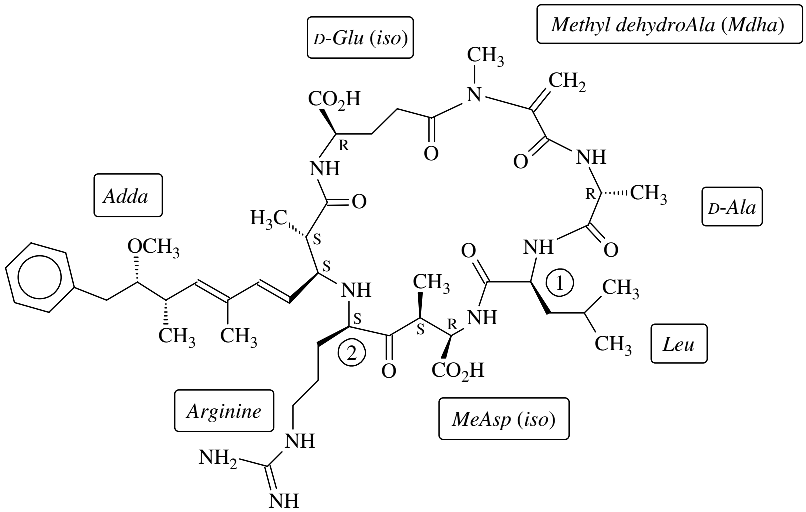

Figure 1. The generic structure of a microcystin. Amino acids substituted at positions 1 and 2 account for most of the variations in this family of compounds (e.g., microcystin-LR contains the amino acids leucine (L) and arginine (R) at positions 1 and 2, respectively) (courtesy of Dr. Gregory Boyer, SUNY-Syracuse).

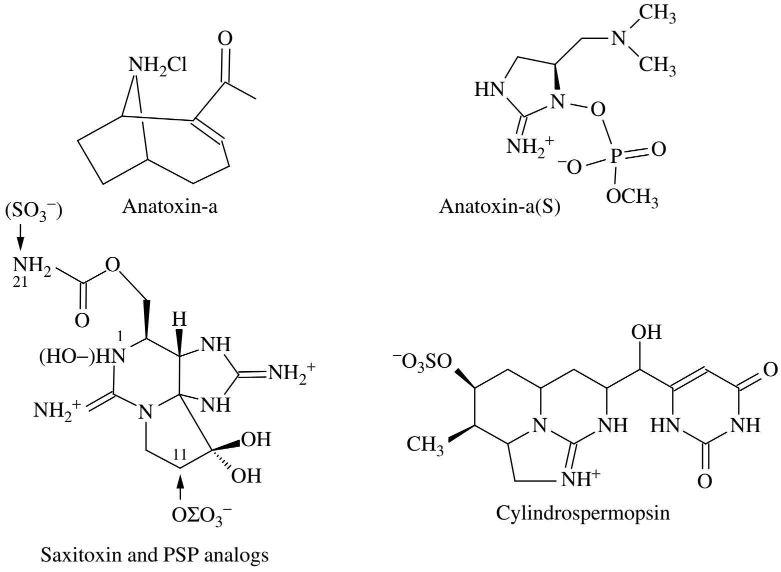

Figure 2. Structures of the alkaloid cyanobacterial toxins: anatoxin-a, anatoxin-a(S), saxitoxin, and cylindrospermopsin. Anatoxin-a, -a(S), and saxitoxin are neurotoxins. Cylindrospermopsin is a hepatotoxin. Saxitoxin is a representative of a large toxin family referred to as the paralytic shellfish poisons (courtesy of Dr. Gregory Boyer, SUNY-Syracuse).

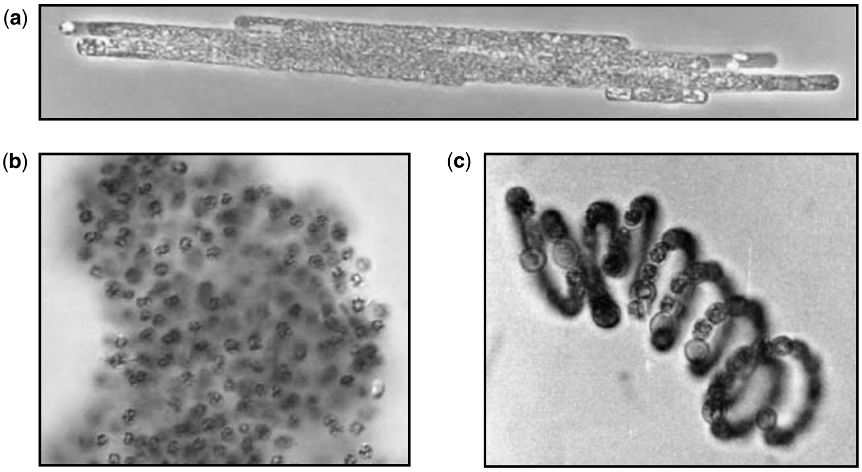

Figure 3. Known toxin producing blue-green algae as seen through a microscope. (**a**) *Aphanizomenon flos-aquae* (**b**) *Microcystis aeruginosa* (**c**) *Anabaena circinalis*.

dyes. Chemical controls include algicides and phosphorus inactivation. Biological controls include in-lake plantings and barley straw. All of these techniques have advantages and disadvantages that need to be addressed for controlling an algal bloom.

HEAT BALANCE OF OPEN WATERBODIES

Klaus Jöhnk
University of Amsterdam
Amsterdam, The Netherlands

INTRODUCTION

From a continuum–mechanical point of view, the heat balance or the thermodynamics of a lake, reservoir, or ocean is governed by three different types of processes: surface fluxes, those through the water surface and the sediment boundary of a water body; internal heat production; and external heat supply. Whereas surface fluxes are accounted for in boundary conditions, supply and production take place in the water volume and are described as additional production terms in the equation of heat (1). Turbulent diffusion and, to a much lesser degree, molecular diffusion lead to a certain temperature stratification, usually dividing a lake into a well-mixed, upper warm layer, the epilimnion, separated from it by a transition zone a colder hypolimnion. The temperature–depth distribution can be determined from the heat balance equation and boundary conditions describing the surface fluxes into the water. In contrast to this, in a well-mixed waterbody with a homogeneous temperature distribution, for example, a polymictic lake, all heat production and flux terms are lumped together in a net thermal energy flux. In this case, the heat production terms are treated as fluxes across the surface. Using such a simple system, one can determine the equilibrium temperature of the waterbody for a certain set of meteorologic conditions, that temperature to which the mixed waterbody is driven under the same prevailing meteorologic conditions. For a stratified lake or ocean, such an approach is only of limited use because it does not account for vertical transport of heat in the waterbody. But it can be a good approximation for the surface temperature when the stratification is strong, thus separating the hypolimnion from the well-mixed epilimnion.

The heat budget of a waterbody is determined mainly by the heat exchange across the water surface through long-wave (infrared) radiation, evaporation, and conduction, plus heat production through the absorption of short-wave radiation in the water column. To a far less extent, inflows (rivers, discharges) and outflows, precipitation, heat conduction across the bottom, and biogeochemical reactions alter the heat balance. These effects become relevant only in certain cases: in/outflow in reservoirs and lakes with short residence time (2); heavy precipitation by thunderstorms (3); sedimentary (or geothermal) heat flow in transparent shallow lakes (4); or in amictic or meromictic lakes, where the water body is decoupled from the atmosphere (5).

Special conditions occur in ice-covered lakes. Here, heat conduction and short-wave absorption through the snow and ice cover limit the heat entering the waterbody. The parameters describing these processes depend strongly on factors such as the type of ice, ice thickness, water inclusions, snow cover, and snow type (6). The description of the heat transfer across the air–ice interface down into the lake water thus needs to include additional formulations for snow and ice (7).

HEAT BALANCE

Meteorologic Driving Factors

Heat exchange across a water surface is related to the meteorologic conditions above the water surface and the water surface temperature itself. The main driving factors are air temperature, water vapor pressure, and wind speed above the water surface. They control conductive and evaporative heat transfer. Radiative heat fluxes are additionally influenced by clouds and atmospheric conditions. Usually only a limited set of meteorologic variables are known for calculating heat fluxes using empirical relations. These variables comprise air temperature T_a and relative humidity H, usually measured 2 m above ground level, wind speed U, cloud cover C, and global irradiance I. Instead of relative humidity, dew point temperature or water vapor pressure might be given. Wind speed drops significantly when measured nearer to the ground, so it has to be scaled to a fixed level, for example, 10 m above ground, using a logarithmic velocity profile depending on the roughness height of the terrain (8).

General Heat Balance Formulation

The main components of the net heat exchange Q_{net} between atmosphere and water can be divided into radiative and nonradiative terms. Radiative heat fluxes are distinguished according to their wavelengths. Short-wave radiation Q_s entering the water column has its major spectral parts in the range from 400–700 nm, whereas long-wave, infrared radiation is in a range of approximately 10 μm. In contrast to short-wave radiation, which penetrates the water column, infrared radiation from clouds and the atmosphere, Q_{lin}, is absorbed directly at the water surface. Depending on the surface temperature T_w of the waterbody, it behaves as a so-called blackbody and also emits infrared radiation, Q_{lout}. Nonradiative heat flux terms are due to evaporation or condensation Q_e and the sensible heat flux Q_c generated by conduction.

Omitting other less important heat fluxes such as geothermal heat flux or those by in/outflows, the general heat balance is given by

$$Q_{\mathrm{net}} = Q_s + Q_{\mathrm{lin}} + Q_{\mathrm{lout}} + Q_e + Q_c \tag{1}$$

The magnitudes of the different heat fluxes are shown in Fig. 1 for Lake Nieuwe Meer in The Netherlands.

Empirical Relations

The relations between meteorologic fields and heat fluxes are evaluated by empirical formulas. Usually they are derived from measurements at lakes or the ocean in certain climatic and/or geographic regions. One can find a multitude of parametric values for a certain empirical relation for different waterbodies. Thus, those relations cannot be generalized, and care must be taken in applying them to other settings. Collections of different formulations and parametric sets are described in the literature (e.g., 9–11). In the following, one set of these empirical relations is shown that exemplifies their structure. The heat flow direction is assumed here to be positive from the atmosphere into the waterbody.

Short-Wave Radiation. Solar irradiance at the water surface is partly reflected, depending on the solar zenith angle. The zenith angle Θ itself depends on latitude λ, the day of the year d, and daytime τ. In addition, the reflectivity r of the water depends on wave action, which

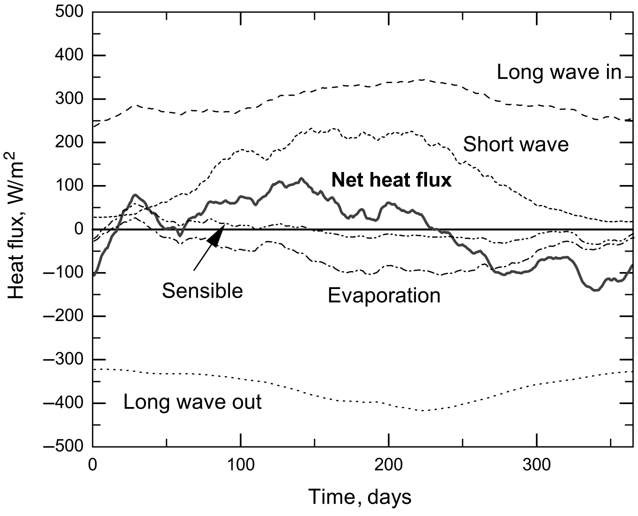

Figure 1. Heat fluxes for Lake Nieuwe Meer, The Netherlands, during the year 2002 (30-day running mean).

can be parameterized by the wind speed U across the water surface (12). If the global irradiance incident at the surface Q_{irr} is measured, the irradiance entering the water column is given by $Q_s = Q_{\text{irr}}[1 - r(\Theta, U)]$. If the irradiance is unknown, it can be calculated from the solar constant S and the angle of incident light at the location (13,14): $Q_{\text{irr}} = S \cos \Theta$. Obviously, this is an upper bound for a clear sky and must be modified for cloud cover by multiplying by an attenuation function $g_s(C)$ (9).

Short-wave radiation that penetrates into a waterbody decays with depth. In its simplest form, this process is described by the Lambert–Beer law: exponential decay $\exp(-kz)$ with depth z and vertical extinction coefficient k. More sophisticated models describing light extinction with depth that depend on water constituents, phytoplankton biomass, and spectral composition of the light can be found in Reference 14. In a well-mixed water column, the depth distribution of heat absorption does not affect the heat balance, as long as there is no significant heat flux into the sediment.

A simple formulation for heat flux into the water column generated by short-wave radiation is given below. Some effects such as the dependence of solar radiation on the Sun–earth distance, atmospheric attenuation, or local effects, for example, shading, are not included here.

Formulation. Earth's declination [day of year: d (days)]:

$$\delta = 23.45 \cdot \frac{2\pi}{360} \cdot \sin\left[\frac{2\pi}{365.25}(d + 283)\right] (\text{deg}) \tag{2}$$

Zenith angle [latitude: λ (deg); time: τ (hours)]:

$$\Theta = \arccos\left[\sin\delta \sin\lambda - \cos\delta \cos\lambda \cos\left(\frac{2\pi}{24}\tau\right)\right] (\text{deg}) \tag{3}$$

Refractive index air–water:

$$n_w = 1.33 \tag{4}$$

Snell's law [incidence angle: Θ; refraction angle: Θ_w (deg)]:

$$\sin\Theta_w = \frac{\sin\Theta}{n_w} \tag{5}$$

Reflection coefficient:

$$r = \frac{1}{2}\left\{\left[\frac{\tan(\Theta - \Theta_w)}{\tan(\Theta + \Theta_w)}\right]^2 + \left[\frac{\sin(\Theta + \Theta_w)}{\sin(\Theta + \Theta_w)}\right]^2\right\} (-) \tag{6}$$

Reflection coefficient depending on wind speed for large zenith angles (wind speed: U [m/s]):

$$r = 0.09[4.444 \exp(-0.17U) + 2.977]^{\left(\frac{180}{\pi}\Theta - 65.5\right)0.040816} (-),\quad \Theta > 65.5^\circ \tag{7}$$

Effect of cloudiness [Cloudiness: $C(-)$ range [0,1]]:

$$g_s = 1 - 0.65\, C^2 (-) \tag{8}$$

Solar constant (for the mean distance between the Sun and the earth)

$$S = 1353\ \text{W/m}^2 \tag{9}$$

Incident irradiance:

$$Q_{\text{irr}} = S \cos\Theta\ (\text{W/m}^2) \tag{10}$$

Heat flux of short-wave radiation entering the water column:

$$Q_s = (1 - r) g_s Q_{\text{irr}}\ (\text{W/m}^2) \tag{11}$$

Long-Wave Radiation from the Atmosphere and Clouds. Long-wave radiation from the atmosphere, clouds, and the water surface can be treated as blackbody radiation; thus, according to the Stefan–Boltzmann law, it is proportional to the fourth power of the body's surface temperature. For long-wave radiation into a lake, cloud cover increases the heat flux, which is taken into account by a quadratic expression for the cloudiness $g_l(C)$. The proportionality factor, emissivity ε_a, depends on a number of factors, such as air temperature, vapor pressure, and cloud height. A common expression for it is given (see e.g., Ref. 9) as a quadratic function of air temperature.

Formulation. Emissivity of the atmosphere [air temperature: T_a (°C)]:

$$\varepsilon_a = 0.919 \cdot 10^{-5}(T_a + 273)^2 (-) \tag{12}$$

Effect of cloudiness [cloudiness: $C(-)$ range (0, 1)]:

$$g_l = 1 + 0.17\, C^2 (-) \tag{13}$$

Stefan–Boltzmann constant:

$$\sigma = 5.6697 \cdot 10^{-8}\ \text{W m}^{-2}\ \text{K}^{-4} \tag{14}$$

Heat flux of long-wave radiation from the atmosphere and clouds:

$$Q_{\text{lin}} = \varepsilon_a \sigma g_l (T_a + 273)^4\ (\text{W/m}^2) \tag{15}$$

Long-Wave Radiation from a Water Surface. The emission of infrared radiation from a water surface takes the usual form for blackbody radiation at the temperature of the water surface, where the emissivity ε_w of water is constant. It yields the Stefan–Boltzmann law for long-wave radiation from a water surface, given below.

Formulation. Emissivity of water:

$$\varepsilon_w = 0.97 (-) \tag{16}$$

Heat flux of long-wave radiation from a water surface [water temperature: T_w (°C)]:

$$Q_{\text{lout}} = -\varepsilon_w \sigma (T_w + 273)^4\ (\text{W/m}^2) \tag{17}$$

Evaporation. Evaporation from a water surface, a heat loss term in evaporation and a source term for condensation, depends on the water temperature and the water vapor pressure above the water surface e_a (see, e.g., Reference 15). The latent heat of evaporation/condensation multiplied by the water's density $\rho_w L(T_w)$ describes the energy needed to evaporate a given volume of water. The amount of water per unit surface, the evaporation height, which evaporates per time, E, depends on the atmospheric conditions above the water column. It is proportional to the gradient of water vapor pressure, which can be approximated by the difference in saturated vapor pressure at ambient water temperature $e^{(\mathrm{sat})}(T_w)$ and the vapor pressure above the water column. The latter can be computed as the saturated vapor pressure at the dew point temperature T_{dew} or via the relative humidity H defined as the ratio of actual to saturated vapor pressure, depending on the meteorologic data available. The effect of wind speed on the evaporation $f_{\mathrm{wind}}(U)$ mixing the air above the water is often assumed to have a linear form (9,16). As humidity is increasing when moving downwind across a water surface, evaporation decreases slightly with the size of a lake (or the wind fetch). To correct for that effect, an additional factor $R_{\mathrm{size}}(A)$ dependent on lake size is applied (17,18).

Formulation. Latent heat of evaporation/condensation [air temperature: T_w (°C)]:

$$L = 2.5 \cdot 10^6 + 2365 T_w \ (\mathrm{J/kg}) \tag{18}$$

Saturated vapor pressure from the Clausius–Clapeyron equation [temperature: T (°C)]:

$$e^{(\mathrm{sat})}(T) = 1230 \exp\left[5362.97\left(\frac{1}{283} - \frac{1}{T+273}\right)\right] \ (\mathrm{Pa}) \tag{19}$$

Water vapor pressure in the air [dew point temperature: T_{dew}(°C); rel. humidity: $H(-)$]:

$$e_a = e^{(\mathrm{sat})}(T_{\mathrm{dew}}) = H e^{(\mathrm{sat})}(T_a) \tag{20}$$

Wind function [wind speed: U (m/s)]:

$$f_{\mathrm{wind}} = 1 + 0.59U(-) \tag{21}$$

Lake size scaling [surface area: A (km^2)]:

$$R_{\mathrm{size}} = \left(\frac{5}{A}\right)^{0.05} \ (-) \tag{22}$$

Evaporation rate:

$$E = 1.36 \cdot 10^{-11} f_{\mathrm{wind}} R_{\mathrm{size}} [e^{(\mathrm{sat})}(T_w) - e_a] \ (\mathrm{m/s}) \tag{23}$$

Density of water:

$$\rho_w = 1000 \ \mathrm{kg/m^3} \tag{24}$$

Heat flux by evaporation/condensation:

$$Q_e = -\rho_w L E \ (\mathrm{W/m^2}) \tag{25}$$

Sensible Heat. The sensible heat flux across a water surface depends on the temperature difference between the water and the overlying air. Assuming that the transport of heat in the atmosphere is similar to that of water vapor, the sensible heat flux due to conduction and convection can be modeled in parallel to the latent heat flux. The Bowen ratio B (19,20) relating these two fluxes can be approximated by the ratio of temperature differences and water vapor pressure differences across the surface and a linear dependence on air pressure p_a. Air pressure is either measured or approximated with a barometric height formulation to adapt for different elevations of the water surface.

Formulation. Bowen ratio [atmospheric pressure: p_a (Pa); air and water temperatures: T_a, T_w(°C)]

$$B = 0.61 \cdot 10^{-3} p_a \frac{T_w - T_a}{e^{(\mathrm{sat})}(T_w) - e_a}(-) \tag{26}$$

Sensible heat flux:

$$Q_c = B Q_e \ (\mathrm{W/m^2}) \tag{27}$$

BIBLIOGRAPHY

1. Hutter, K. and Jöhnk, K.D. (2004). *Continuum Methods of Physical Modeling—Continuum Mechanics, Dimensional Analysis, Turbulence*. Springer, Heidelberg.
2. Straskraba, M. and Hocking, G. (2002). The effect of theoretical retention time on the hydrodynamics of deep river valley reservoirs. *Int. Rev. Hydrobiol.* **87**: 61–83.
3. Patterson, G., Wooster, M.J., and Sear, C.B. (1997). Satellite-derived surface temperatures and the interpretation of the three-dimensional structure of Lake Malawi, Africa: The presence of a profile-bound density current and the persistence of thermal stratification. *Verh. Internat. Verein. Limnol.* **26**: 252–255.
4. Tsay, T.K., Ruggaber, G.J., Effler, S.W., and Driscoll, C.T. (1992). Thermal stratification modeling of lakes with sediment heat flux. *J. Hydraul. Eng.* **118**: 407–419.
5. Coulter, G.W. and Spigel, R.H. (1991). Hydrodynamics. In: *Lake Tanganyika and Its Life*. G.W. Coulter (Ed.). Oxford University Press, Oxford, pp. 49–75.
6. Ashton, G.D. (Ed.). (1986). *River and Lake Ice Engineering*. Water Resources Publications, Littleton, CO.
7. Launiainen, J. and Cheng, B. (1998). Modelling of ice thermodynamics in natural water bodies. *Cold Regions Sci. Technol.* **27**: 153–178.
8. Johnson, H.K. (1999). Simple expressions for correcting wind speed data for elevation. *Coastal Eng.* **36**: 263–269.
9. Henderson-Sellers, B. (1984). *Engineering Limnology*. Pitman, London.
10. Henderson-Sellers, B. (1986). Calculating the surface energy balance for lake and reservoir modeling: A review. *Rev. Geophys.* **24**: 625–649.
11. Martin, J.L. and McCutcheon, S.C. (1999). *Hydrodynamics and Transport for Water Quality Modeling*. CRC Press, Boca Raton, FL.
12. Preisendorfer, R.W. and Mobley, C.D. (1986). Albedos and glitter patterns of a wind-roughened sea surface. *J. Phys. Oceanogr.* **16**: 1293–1316.
13. Forsythe, W.C., Rykiel, E.J., Jr., Stahl, R.S., Wu, H., and Schoolfield, R.M. (1995). A model comparison for day length as a function of latitude and day of year. *Ecol. Model.* **80**: 87–95.

14. Kirk, J. (1983). *Light and Photosynthesis in Aquatic Ecosystems*. Cambridge University Press, Cambridge.
15. Brutsaert, W. (1982). *Evaporation into the Atmosphere*. D. Reidel, Dordrecht, The Netherlands.
16. Winter, T.C., Rosenberry, D.O., and Sturrock, A.M. (1995). Evaluation of 11 equations for determining evaporation for a small lake in the north central United States. *Water Resour. Res.* **31**: 983–993.
17. Marti, D.E. and Imboden, D. (1986). Thermische Energieflüsse an der Wasseroberfäche: Beispiel Sempachersee. *Schweiz. Z. Hydrol.* **48**: 196–228.
18. Condie, S.A. and Webster, I.T. (1997). The influence of wind stress, temperature, and humidity gradients on evaporation from reservoirs. *Water Resour. Res.* **33**: 2813–2822.
19. Bowen, I.S. (1926). The ratio of heat losses by conduction and by evaporation from any water surface. *Phys. Rev.* **27**: 779–787.
20. Lewis, J.M. (1995). The story behind the Bowen ratio. *Bull. Am. Meteorol. Soc.* **76**: 2433–2443.

HYDRAULICS

Stefano Pagliara
Pietro Chiavaccini
Universitá di Pisa
Pisa, Italy

INTRODUCTION

Hydraulics is the branch of physics that studies the equilibrium (hydrostatic) and the motion (hydrodynamics) of fluids. Fluids can be divided into gases and liquids. Hydraulics treats the liquid that generally are practically incompressible and occupies definite volumes.

THE PROPERTIES OF LIQUIDS

The density r is the mass per unit volume, and it is influenced both by the temperature and the pressure. For water at 4 °C, r = 1000 kg/m^3. The specific weight g is the weight of a unit volume of substance. For water, g = 9.813 kN/m^3. The viscosity m is the properties of a fluid that determine the resistance to a shearing force. For liquids, it decreases as the temperature increases, and for water at 20 °C, it is 1.14 10^{-3} kg/(m s). The surface tension is because of the molecular attraction and is responsible for the curvature of the fluid surface near the wall of vessels.

HYDROSTATIC FORCES

A liquid at rest in a vessel exerts a pressure on the base and on the wall. This pressure is called hydrostatic. If A is the area of the base, the force per unit area exerted is $p = W/A$, where W is the weight of the liquid in the vessel. Further, the pressure p applied to any horizontal plane passing through the liquid is $p = gh$, where h is the vertical distance of the plane below the free surface (Stevin's law). Besides, any external pressure applied to a fluid is transmitted undiminished throughout the liquid and onto the walls of the containing vessel (Pascal's principle), and the pressure at a given depth does not depend on the shape of the vessel containing the liquid or the amount of liquid in the vessel.

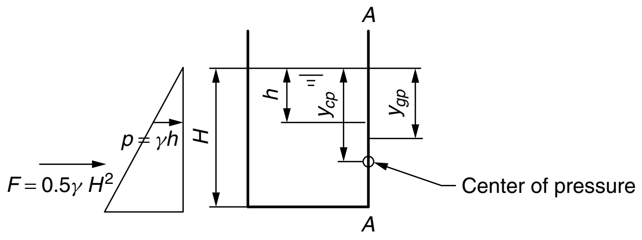

Figure 1. Pressure and resulting forces on the wall A–A of a vessel containing a liquid.

In a liquid at rest, viscous and shearing forces do not exist. The vertical distance h between any selected points in a liquid at rest and the free surface exposed to the atmosphere is called the head of the liquid.

The line of action of the resultant force passes through the center of pressure. This center of pressure lies below the centroid, which is an intuitive result, because pressure increases with depth.

It can be determined graphically or analytically using the following formula:

$$y_{cp} = \frac{I_{cg}}{y_{cg}A} + y_{cg} \tag{1}$$

where I_{cg} is the moment of inertia of the area A about its center of gravity axis and y_{cg} is the depth of the center of gravity below free surface (Fig. 1).

The horizontal component of the resulting force acting on any surface is equal to the normal force on the vertical projection of the surface and acts on the center of pressure for the vertical projection. Instead, the vertical component is equal to the weight of the liquid above the area, real or imaginary.

FLUID FLOWS

Fluid flow can be steady or unsteady, uniform or nonuniform. Steady flows occur if the velocity, pressure, and other fluid variables (density, viscosity) are constant with time. Uniform flow instead occurs when no variation in space exists. The two fundamental equations of fluid flows are the equation of continuity and the Bernoulli equation (or the momentum equation). The first expresses the principle of conservation of mass, and in steady conditions and for incompressible fluids becomes

$$A_1V_1 = A_2V_2 \tag{2}$$

where A_i and V_i are, respectively, the cross-sectional area and the average velocity at the generic section i.

The Bernoulli equation results from the application of the principle of conservation of energy, and in the direction of the flow is expressed as follows:

$$\frac{p_1}{\gamma} + z_1 + \frac{V_1^2}{2g} = \frac{p_2}{\gamma} + z_2 + \frac{V_2^2}{2g} + \Delta h \tag{3}$$

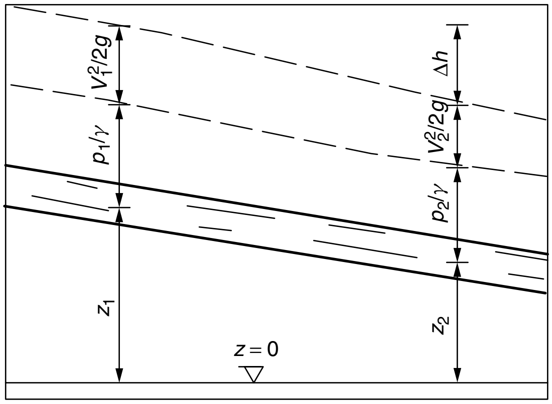

Figure 2. Graphical representation of Bernoulli equation.

where z_i is the height of the fluid above a reference level, p_i the pressure, V_i the velocity, g the acceleration because of gravity, and Dh the energy lost between the two sections. The term p/g represents the pressure head, z the geometric head, and $V^2/2g$ the velocity head, and their sum is the total energy. The graphical representation of the Bernoulli equation is illustrated in (Fig. 2).

FLOWS THROUGH ORIFICES AND WEIRS

For a tank with a hole with sharp edges (Fig. 3a) flowing in the atmosphere, the discharge Q through the orifice of area A can be obtained applying Bernoulli's equation:

$$Q = \mu A\sqrt{2gh} \tag{4}$$

where h is the water depth above the center of the orifice and m is a coefficient that takes into account the effect of friction and the contraction of the flow and is equal to 0.61–0.62.

A weir is a wall built in a channel cross section, over which the water flows from one level to another in a controlled way (Fig. 3b). The relationship that links the characteristics of a rectangular weir to the discharge, in the hypothesis that the velocity of the fluid approaching is small so that kinetic energy can be neglected, is:

$$Q = \mu' Bh\sqrt{2gh} \tag{5}$$

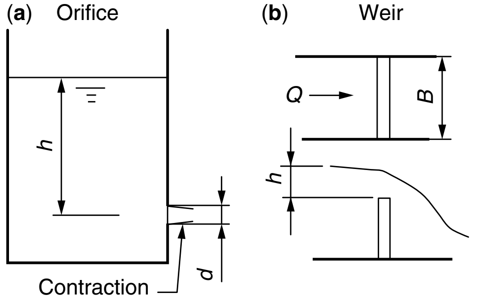

Figure 3. Flow through orifices (**a**); and weir (**b**).

where B is the width of the weir, h is the hydraulic depth over it, and m' is a coefficient (0.42 in case of sharp edge and 0.385 for a wide-crested weir).

FLOWS IN PIPES

Flow in pipes can be laminar or turbulent (1,2). In the first case, the fluids move in parallel layers. In the second one, the particles move in all directions, causing a complete mixing of the fluid. If the flow is laminar or turbulent, it can be evaluated with the Reynolds number, which expresses the ratio between inertial and viscosity forces:

$$\mathrm{Re} = \frac{\rho V d}{\mu} \tag{6}$$

where d is the diameter of the pipe. When Re<2000, flow is in laminar condition, whereas for values greater than 2000, it starts to be turbulent.

If a flow is laminar, the head losses Dh can be determined using the Hagen–Poiselle formula:

$$\Delta h = \frac{64}{\mathrm{Re}}\frac{L}{d}\frac{V^2}{2g} \tag{7}$$

where L is the distance between two sections.

For turbulent flow, the lost head can be given by means of the Darcy–Weisbach formula:

$$\Delta h = f\frac{L}{d}\frac{V^2}{2g} \tag{8}$$

where f is the friction factor, which depends on the values of Re and on the roughness of the wall of the pipe. According to Colebrook and White, f is expressed as:

$$\frac{1}{\sqrt{f}} = -2\log\left[\frac{(\varepsilon/d)}{3.71} + 2.51\frac{\mathrm{Re}}{\sqrt{f}}\right] \tag{9}$$

where e is the dimension of the roughness.

The head losses can also be because of singularity inserted in the pipes (curves, valves, entrance and exiting connections, etc.), and they are expressed as

$$\Delta h = k\frac{V^2}{2g} \tag{10}$$

where k is a coefficient depending on the characteristics of the singularity.

OPEN CHANNEL FLOWS

An open channel is a conduit in which the liquid flows with a free surface subjected to atmospheric pressure (3). For steady and uniform flow, the energy grade line is parallel to the liquid surface (which represents the hydraulic grade line). Define with S the slope of the channel and with R the hydraulic radius, the velocity of the flow can be determined with the Chezy formula:

$$V = C\sqrt{RS} \tag{11}$$

where C is a coefficient depending on the wall roughness determined with the following expression (Bazin):

$$C = \frac{87}{1 + m\sqrt{R}} \tag{12}$$

The coefficient m is given in Table 1.

The hydraulic radius represents the ratio between the wetted area and the wetted perimeter. For a rectangular channel, $R = Bh/(B + 2\,h)$ where B is the width and h the water depth.

Another expression to determine the hydraulic characteristic of open channel flow is the Manning equation:

$$V = \frac{1}{n} R^{2/3} \sqrt{S} \tag{13}$$

where n is the Manning coefficient (Table 2).

In open channel flow, the specific energy can be introduced as

$$E = y + \frac{V^2}{2\,g} \tag{14}$$

It represents the sum of water depth and velocity head, and in terms of unit discharge $q = Q/B$, it is

$$E = y + \frac{q^2}{2\,gy^2} \tag{15}$$

For a constant value of q and for rectangular channel, the specific energy has a minimum for a water depth

$$y_c = \sqrt[3]{\frac{q^2}{g}} \tag{16}$$

This water depth is called critical depth. For a given value of specific energy, it corresponds to the maximum unit discharge.

Table 1. Values for Bazin's Roughness Coefficient

Type of Channel	Bazin's m
Cement: planed wood	0.109
Planks: brick, cut stones	0.290
Roubble masonry	0.833
Earthen channel: very regular	1.54
Earthen channel: in ordinary condition	2.35
Earthen channel: very rough or weed grown	3.17

Table 2. Values for Manning Coefficient

Type of Channel	n
Natural stream: Clean and Straight	0.030
Natural stream: Major Rivers	0.035
Natural stream: Sluggish with Deep Pools	0.040
Excavated Earth Channels: Clean	0.022
Excavated Earth Channels: Gravelly	0.025
Excavated Earth Channels: Weddy	0.030
Floodplains: Pasture, Farmland	0.035
Smooth Steel	0.012
Smooth Steel	0.022

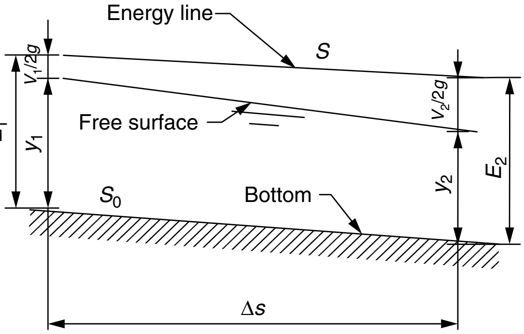

Figure 4. Nonuniform flow condition.

Flows occurring at depth below and above the critical depth are classified as supercritical and subcritical, respectively.

In steady, nonuniform flow conditions, the water depth varies with the distance. These conditions can usually be studied using Bernoulli's theorem and dividing the problem into lengths of approximately uniform condition. For each reach of length DS, the energy equation becomes

$$\Delta S = \frac{E_2 - E_1}{S_0 - S} \tag{17}$$

where E_2 and E_1 are the specific energies downstream and upstream, respectively, S_0 is the bottom slope, and S is the slope of the energy line (Fig. 4).

CONCLUSION

A compacted summary of the fundamental principles of hydraulics is presented in this paper. Aspects of hydrostatic and of the motion of fluids have been described from an engineering point of view.

BIBLIOGRAPHY

1. White, F.H. (1994). *Fluid Mechanics*. McGraw-Hill, New York.
2. Rouse, H. (1979). *Elementary Mechanics of Fluids*. Dover Publications, New York.
3. Chow, V.T. (1959). *Open Channel Hydraulics*. McGraw Hill, Kogakusha.

HYDRAULICS OF PRESSURIZED FLOW

Xing Fang
Lamar University
Beaumont, Texas

Hydraulics is a branch of fluid mechanics specifically dealing with water as a major fluid in many engineering systems. It is safe to consider water as an incompressible fluid for typical engineering applications. Pressurized flow usually refers to the flow in closed conduits. Hydraulics of pressurized flow involves three basic equations: continuity, energy, and momentum equations. Head loss is important in hydraulics of pressurized flow regardless of system design or flow analysis. Darcy-Weisbach and Hazen and Williams equations typically estimate friction head loss. The friction factor depends on Reynolds number and relative roughness of the pipe. Total head loss includes friction head loss caused by fluid viscosity and local head losses at changes in the cross section, bends, and various pipe fittings. A pump provides energy head to, and a water turbine takes energy head from, the pipe system.

Fluid mechanics is the study of behavior of liquids and gasses at rest or in motion subjected to various forces. Hydraulics is a branch of the fluid mechanics specifically dealing with water as a major fluid in many engineering systems. The way water behaves under various conditions depends primarily on its fundamental chemical and physical properties. The bulk modulus of water, which is a measure of its elasticity (or volume reduction under pressure), is about 2.2×10^9 N/m^2 under normal pressure. Therefore, it is safe to consider water as an incompressible fluid for practical purposes. An incompressible fluid does not change its density (ρ) with respect to pressure; therefore, it is also independent of location if no changes of temperature occur in the study area. Pressurized flow usually refers to the flow in closed conduits, e.g., in a water supply pipe system, where the flow is under pressure and full-pipe flow (Fig. 1), whereas pressure could be above or under the local atmospheric pressure.

Hydraulics of pressurized flow involves three basic equations: continuity, energy, and momentum equations. The equation of continuity is a mathematical statement of the Law of Conservation of Mass (water). Along a conduit, inflow rate at the upstream cross section 1 (Q_1) equals the outflow rate at the downstream cross section 2 (Q_2):

$$Q_1 = Q_2 \quad \text{or} \quad V_1 A_a = V_2 A_2 \tag{1}$$

where V_1 and V_2 are the average velocities at the cross sections 1 and 2 and A_1 and A_2 are the cross-sectional areas. When the branch pipe system is involved, the equation of continuity states that the sum of inflow equals the sum of outflow in the control volume system. The energy equation for incompressible fluid under no heat exchange can be expressed as

$$Z_1 + \frac{p_1}{\gamma} + \frac{V_1^2}{2g} = Z_2 + \frac{p_2}{\gamma} + \frac{V_2^2}{2g} + HL_{12} + H_p - H_t \tag{2}$$

Each term in Eq. 2 has units of length (e.g., meter or feet). The first three terms ($z, p/\gamma, V^2/2g$) are the energy per unit weight of water caused by its elevation, pressure, and velocity, respectively, and they are called elevation head, pressure head, and velocity head. Subscripts 1 and 2 stand for upstream and downstream cross sections 1 and 2, and g is the acceleration of gravity (9.81 m/s^2 or 32.2 ft/s^2). HL_{12} is energy loss per unit weight of water from the cross section 1 to 2 and is called total head loss. H_p and H_t are energy heads provided by a water pump or taken by a water turbine to/from the pressurized flow system:

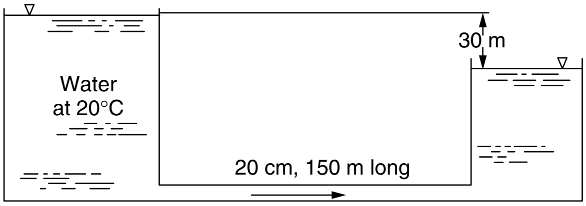

Figure 1. Pressurized pipe flow between two reservoirs.

$$H_p = \frac{P_E \eta}{\gamma Q} \quad \text{or} \quad H_t = \frac{P_T}{\eta \gamma Q} \tag{3}$$

where P_E is the electric power provided to a pump, P_T is the electric power generated by a turbine, η is overall efficiency, γ is specific weight of water, and Q is discharge through a pump or turbine. Some applications may not have pumps or turbine in the system (e.g., in Fig. 1). Head loss is an important term in hydraulics of pressurized flow regardless of system design or flow analysis. Total head loss includes friction head loss caused by fluid viscosity and local head losses at changes in the cross section, bends, and various pipe fittings. Friction head loss (HL_f) is usually determined from the Darcy–Weisbach equation:

$$HL_f = f \frac{L}{D} \frac{V^2}{2g} \tag{4}$$

where f is the friction factor, L is the pipe length between the cross section 1 and 2, D is the pipe diameter, and V is the mean velocity (Q/A). For a noncircular conduit, D is replaced by four times the hydraulic radius that is the flow area divided by wetted perimeter. The friction factor is a function of the Reynolds number (Re) and the relative roughness of a conduit (absolute roughness ε over diameter D). Absolute roughness for concrete, welded-steel and cast iron pipe are 0.3 to 3.0 and 0.12 and 0.25 mm, respectively, and it is a function of pipe material. For laminar flow, f is equal to 64/Re. For turbulent flow, the Moody diagram presented in many textbooks (1–3) can be used, and Swamee and Jain (4) also suggested an explicit formula in replace of the Moody diagram with restrictions:

$$f = \frac{1.325}{\left[\ln\left(\frac{\varepsilon}{3.7D} + \frac{5.74}{\mathrm{Re}^{0.9}}\right)\right]^2} \tag{5}$$

for $10^{-6} < \varepsilon/D < 10^{-2}$ and $5000 < \mathrm{Re} < 10^8$. For water supply engineering applications, the Hazen and Williams equation is typically used to estimate friction head loss as

$$HL_f = L \frac{4.727 Q^{1.852}}{C^{1.852} * D^{4.8704}} \quad \text{in English units} \tag{6a}$$

$$HL_f = L \frac{10.675 Q^{1.852}}{C^{1.852} * D^{4.8704}} \quad \text{in the international system of units} \tag{6b}$$

where C is a coefficient given in Table 1 for various types of conduit surface. Local head loss (HL_L) is typically expressed as function of velocity head:

$$HL_L = K \frac{V^2}{2g} \tag{7}$$

Table 1. Values of C for the Hazen–Williams formula [from Khan (1)]

Type of Pipe	C	Type of Pipe	C
Asbestos cement	140	Brass	130–140
Cast Iron		Concrete	
New, unlined	130	Steel forms	140
Old, unlined	40–120	Wooden forms	120
Cement lined	130–150	Centrifugally spun	135
Bitumastic, enamel-lined	140–150		
Tar-coated	115–135		
Copper	130–140	Galvanized iron	120
Lead	130–140	Plastic	140–150
Steel (riveted)	110	Steel (new, unlined)	140–150

Velocity V is typically taken from just downstream of the local fitting. K is the local head loss coefficient and typically determined by experiments. For example, perpendicular square entrance, $K = 0.5$; exit to a reservoir, $K = 1.0$; standard tee entrance, $K = 1.8$; globe valve when fully open, $K = 10$; and close return bend, $K = 2.2$. Many other local head loss coefficients are given in textbooks (1–3). Local head losses from various fittings are also represented as equivalent pipe lengths in many application fields. Therefore, the total head loss can be calculated by Darcy–Weisbach and Hazen–Williams equations alone when length (L) as the sum of equivalent pipe lengths from fittings and actual pipe length is used.

The fluid momentum equation can analyze forces required to keep a pipe bend and a vertical tee in place (Fig. 2). Fluid momentum equation is derived from the Newton's second law of motion as

$$\sum F = \dot{M}_o - \dot{M}_I \tag{8}$$

It means that the sum of all external forces applied on the control volume equals the difference of the rate change of momentum leaving (outflow) and entering (inflow) the control volume. It is easier to arrange the momentum equations in the direction of three rectangular axes x, y, and z.

Energy Eq. 2 can determine what flow rate the pressurized cast iron pipe in Fig. 1 can carry:

$$HL_{12} = f\frac{L}{D}\frac{V^2}{2g} = Z_1 - Z_2 = 30\ m = f\frac{150}{0.2}\frac{V^2}{2 * 9.81} \tag{9}$$

The cross sections 1 and 2 are located at the surface of upstream and downstream reservoirs, so pressure head and velocity head at the cross sections 1 and 2 are all zero. H_p and H_t are zero as no pump and turbine are in the system. The relative roughness (ε/D) of the cast iron pipe is 0.25 mm/20 cm = 0.00125; and by assuming that the Reynolds number is very large first, the friction factor can be estimated from Eq. 5 as 0.02. Therefore, flow velocity V can be calculated from Eq. 9 as 6.26 m/s. Now one needs to recheck the Reynolds number and recalculate friction factor and velocity:

$$\mathrm{Re} = \frac{VD}{\nu} = \frac{6.26\ m/s * 0.2\ m}{1.007 \times 10^{-6}} = 1.24 \times 10^6 \tag{10a}$$

$$f = \frac{1.325}{\left[\ln\left(\dfrac{0.25}{3.7 * 200} + \dfrac{5.74}{(1.24 \times 10^6)^{0.9}}\right)\right]^2} = 0.02 \tag{10b}$$

where ν is the kinematic viscosity of water at 20 °C. Because the friction factor is still 0.02, the final velocity is the same as the first trial, and the flow rate in the pipe is

$$Q = VA = 6.26\ \mathrm{m/s} * (\pi/4 * 0.2^2) = 0.196\ \mathrm{m^3/s} = 6.95\ \mathrm{ft^3/s} \tag{11}$$

where local head losses (entrance and exit) are relatively small and ignored in the above calculation. For fully turbulent conditions at high Reynolds numbers, discharge for a pressurized pipe flow can be calculated as

$$Q = \left[\frac{\pi}{4}\sqrt{2g}\left(1.14 - 0.86\ln\frac{\varepsilon}{D}\right)D^{2.5}\right]\left(\frac{HL_f}{L}\right)^{1/2} \tag{12}$$

BIBLIOGRAPHY

1. Khan, A. (1987). *Fluid Mechanics*. Holt, Rinehart and Winston, New York, pp. 320–326.

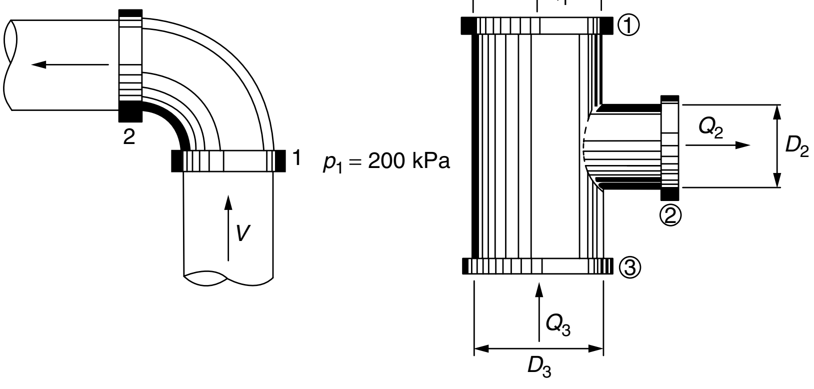

Figure 2. Examples for momentum applications: a pipe bend and a vertical tee.

2. Roberson, J.A., Cassidy, J.J., and Chaudhry, M.H. (1997). *Hydraulic Engineering*, 2nd Edn. John Wiley & Sons, New York.
3. Simon, L. and Korom, S.F. (1997). *Hydraulics*, 4th Edn. Prentice-Hall, Englewood Cliffs, NJ, pp. 142–143.
4. Swamee, P.K. and Jain, A.K. (1976). Explicit equations for pipe-flow problems. *J. Hydr. Div., Proc. ASCE*, pp. 657–664.

HYDROELECTRIC POWER

Sharad K. Jain
National Institute of Hydrology
Roorkee, Uttranchal, India

INTRODUCTION

Hydropower generation follows a simple concept: water falling from the force of gravity turns the blades of a turbine, which is connected to a generator. The rotating generator produces electricity. Prehistoric man was aware of the energy contained in falling water. One of the earliest devices to use this energy was the water wheel. The Romans used the energy of falling water to do many useful things. They constructed paddle wheels that turned with river flow and lifted water to troughs built higher than river level. The Egyptians and Greeks harnessed the power of river currents to turn wheels and grind grain before 2000 B.C. In the Middle Ages, more efficient water wheels were built for milling grain.

In the nineteenth century, the water turbine gradually replaced the water wheel, and dams were built to control the flow of water. Since then, the hydroelectric potential of rivers continued to be developed. The first modern turbine design was developed in 1849 by James Francis. According to estimates, hydroelectric power production has now risen to 2,000 billion kilowatt hours worldwide.

DAMS AND HYDROPOWER

As a dam holds back water in its reservoir, the water level rises creating a reservoir of potential power. This water is allowed to pass through the turbines of a hydroelectric power plant in conduits called penstocks. The water turns the turbines which are coupled with generators to produce electricity. This electricity passes through a transformer and reaches factories, shops, and homes on transmission lines. The principle and the technique to generate electricity from water remain the same regardless of the size of the project. A plant may serve a small community or a country. For example, many communities in remote areas of Nepal are not connected to the national grid and get electricity from mini/micro hydropower (size <2 MW) plants. The largest hydroelectric complex in the world now is on the Parana River between Paraguay and Brazil. It is known as the Itaipu Dam, and its 18 turbines produce 12,600 megawatts (MW) of electricity. The Itaipu Dam supplies about 80% of the entire electricity demand in Paraguay and 25% of the demand in Brazil. Even this will be surpassed by the hydropower plant at the Three Gorges Dam (under construction) across the Changjiang River (China) whose installed capacity will be about 18,000 MW!

QUANTITATIVE HYDROPOWER FACTS

Hydropower is the world's most important and cleanest source of renewable energy. It can be seen from Table 1 that nearly 17.5% of the world's electricity is produced by hydropower.

Canada is the world's biggest producer of hydropower, generating 350 TWh/year (nearly 62% of the country's total electricity production). Canada produces more than 13% of the global output of hydropower and is the world's second largest exporter of electricity after France.

Europe and North America have developed more than 60% of their hydropower (see Table 2), most of it during the twentieth century. In contrast, many countries in Asia, South America, and Africa currently use only a small portion of their potential hydropower, and a large hydropower potential still remains unexploited.

The electricity supplied by hydropower far exceeds the capacity of any other renewable energy resource. Norway meets virtually its entire (99.6%) electricity demand by hydropower. Twenty-five countries worldwide depend on hydropower for more than 90% of their electricity needs. Though fossil fuels are dominant in electricity generation, worldwide, more than 60 countries currently use hydropower for half or more of their electricity needs. Most of the installed hydroelectric capacity is in North America, Brazil, Russia, China, and Europe. The amount of hydropower generation by major producers is shown in Table 3.

Table 1. World Electricity Generation by Source, 1999[a]

Source	Percentage
Hydropower	17.5
Coal	38.1
Nuclear	17.2
Gas	17.1
Oil	8.5
Other	1.6

[a] Reference 1.

Table 2. Continentwise Percentage of Hydropower Potential that has been Developed

Continent	Percentage of Hydropower Potential that has been Developed	Percentage of Electricity Generated by Hydropower
Africa	7	2
Asia	20	39
Australia	40	2
Europe	65	13
N. America	61	26
S. America	19	18

Table 3. Hydropower Generation by Major Producers of the World[a]

Country	Hydropower Produced TWh	% of World Total
Canada	346	13.0
United States	319	12.0
Brazil	293	11.0
China	204	7.7
Russia	161	6.1
Norway	122	4.6
Japan	96	3.6
India	81	3.0
France	77	2.9
Sweden	72	2.7
Rest of the world	888	33.4
Total World	2659	100.0

[a] Reference 1.

ADVANTAGES OF HYDROPOWER

Hydroelectric power is a clean, renewable energy source. In contrast to other renewable sources of electricity, hydropower can supply a significant portion of the world's electricity needs, year after year. Because it can respond quickly to changes in electricity demand by controlling the release of water, hydropower can augment other energy sources, such as solar and wind energy. Besides, hydropower generation is better suited to meet demands for peak loads compared to thermal units.

These are some other advantages of hydropower generation

- Hydropower has furnished electricity to the world for more than a century, making it a time-tested, reliable technology of known costs and benefits.
- This is the most important renewable source of energy in the world. The Sun is the prime mover of the water cycle, and power can be produced as long as the rain/snow falls and rivers flow.
- As the raw material is virtually free, generating cost is almost free from inflation.
- Hydropower plants do not require much outlay on account of operation and maintenance and have long lives.
- Hydropower generation does not pollute the environment; no heat is produced, and no harmful gases are released.
- Hydropower power plants work at very high efficiency compared to the other major source (thermal power).
- Reservoirs can store water during times of low demand and can quickly start power generation. Thermal power plants take much longer to start up than hydropower plants.
- In multipurpose projects, generation of hydroelectric energy is combined with other uses such as irrigation, water supply, and flood control, resulting in tremendous savings of money and other resources.
- Small, mini-, and microscale hydropower projects can be tailor-made to provide power and minimize adverse environmental and social impacts.
- As a result of developments in turbine technology, efficient turbines whose capacities vary from several hundreds of MW to a few MW have been developed.

If waterpower is used to generate electricity in place of fossil fuels, emission of greenhouse gas can be substantially reduced. However, the creation of large reservoirs and accompanying infrastructures is being resisted by many local communities. Often, small hydropower projects built with the involvement of local communities provide an acceptable solution to electricity problems and the least harm to the environment. Many regions, particularly Asia, Europe, and North America, have developed substantial hydroelectric resources using small plants of <10 MW capacity.

DISADVANTAGES AND ISSUES IN HYDROPOWER GENERATION

The development of hydropower in many countries is not at desired rates for several reasons. Major hydropower projects have large initial costs and long gestation periods. Of late, these projects are being opposed because of harm that they can do to the environment. The reservoirs inundate forests, disturb wild life, and the local population may be displaced causing social tensions.

Careful planning and operation of hydropower facilities can minimize environmental damage, but environmental costs may hinder the development of hydropower in some areas. Another problem is the adverse effect of dams on river ecosystems. After construction of a dam, the river flow downstream is considerably reduced leading to significant changes in river ecology. The construction of a dam also restricts fish movement.

Proponents of these projects argue that many of the adverse impacts are temporary: they arise mainly during the construction phase. Problems such as forest submergence can be tackled by compensatory afforestation. The acceleration in regional development as a result of these projects can more than offset the harmful effects. Most problems are not insurmountable and can be largely overcome by careful planning and involving all stakeholders in decision-making.

COMPONENTS OF HYDROPOWER PROJECTS

In a storage project, the reservoir behind the dam stores water that generates electric power. The portion of the reservoir that is immediately upstream of the intake structure is known as the forebay. A penstock is a conduit to carry water from the forebay to turbines and gates, and valves are installed to control the flow of water. A surge tank is constructed to handle the problems of water hammer. The turbines and generators are installed in a powerhouse. The water comes out of the turbines through the draft tube and joins the tail water.

The term "static head" (see Fig. 1) denotes the difference between the water surface elevation in the forebay and the tail water level (TWL). The 'net head'

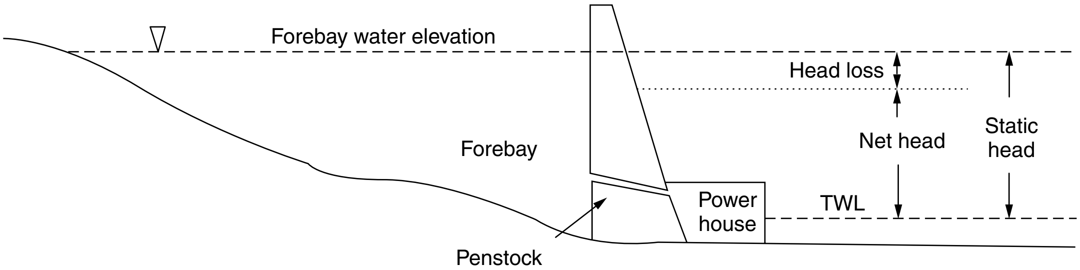

Figure 1. Terms related to head in a hydropower plant.

is the static head less losses in the penstock:

$$\begin{aligned}\text{Net head} &= \text{Static head} - \text{Losses}\\ &= \text{Water surface elevation in forebay}\\ &\quad - \text{TWL-Penstock losses} \qquad (1)\end{aligned}$$

Although the head is related to the height of the dam, a low dam can yield a high head if the powerhouse is located some distance downstream from the dam. The amount of hydropower generated is a function of the discharge and the hydraulic head:

$$P = 9.817QH\eta \qquad (2)$$

where P is electric power in kW, Q is the discharge through the power plant in m^3/s, and H is the net head in m. Further, η is the overall efficiency of the power plant expressed as a ratio (usually about 0.85) obtained by multiplying the turbine efficiency by the generator efficiency. Note that hydroelectric power generation depends on the rate of flow through the turbines and the effective head. Thus, power output can be controlled by releasing more water at a low head or less water at a high head. Most hydropower plants are constructed in hilly areas where steep slopes provide high heads.

TYPES OF HYDROPOWER PROJECTS

There are four major classifications of hydroelectric projects with respect to types of site development: storage, barrages, run-of-river, and pumped storage. Storage plants usually have heads in the medium to high range (greater than 25 m) and can store relatively large volumes of water during periods of high stream flow to provide supply for power generation during periods of low stream flow. The power plant is commonly located at the toe of the dam, though it might also be away from the dam. Peaking operation is frequently associated with storage projects, and this requires large and sometimes rapid fluctuations in releases of water. It is often necessary to provide facilities to even out the fluctuations in the discharge, if rapid changes in water levels below the project are not tolerable.

A barrage, also known as pondage, has a very small storage capacity. It can't effectively regulate flow and generates power according to the variation of load. Run-of-river plants have little or no storage and, therefore, must generate power from stream flow as it occurs with little or no benefit from at-site regulation. Run-of-river projects generally have productive heads in the low to medium range (5 to 30 m) and are frequently associated with navigation or other multipurpose developments. For a base-load run-of-river project to be feasible, the stream must have a relatively high baseflow. Falls in irrigation canals are sometimes used to generate energy.

PUMPED STORAGE SCHEMES

A significant and growing portion of the hydroelectric capacity worldwide is devoted to pumped storage facilities that are designed solely to provide power during peak loads. These projects depend on pumped water as a partial or total source of water. Pumped storage projects consist of a high level forebay where inflow or pumped water is stored until needed and a low level afterbay where the power releases are stored. Pumping and generation are done by units composed of reversible pump turbines and generator motors connecting the forebay and afterbay. During off-peak hours, excess electricity produced by conventional power plants is used to pump water from lower to higher level reservoirs. During periods of high demand, water is released from the upper reservoir to generate electricity. Such projects derive their usefulness because the demand for power is generally low at night and on weekends, and therefore, pumping energy at a very low cost will be available from idle generating facilities. The feasibility of pumped storage developments arises from the need for relatively large amounts of peaking capacity, the availability of pumping energy at a cheap rate, and a sufficiently long off-peak period. It results in large cost savings through more efficient use of base-load plants.

There are three types of pumped storage development: diversion, off-channel, and in-channel. In the diversion type of development, water is released through generating units into an afterbay in an adjacent basin. The off-channel type of pumped storage development is most suitable when a forebay site exists on a hill above a stream where an afterbay can be constructed. The head differential should be large, and the forebay site should be close to the afterbay to avoid head loss and reduce construction costs. The water requirement in this scheme is not large after the initial supply has been provided.

In the in-channel type of scheme, the reservoir of a conventional power project is used as a forebay. The afterbay could be a reservoir from a downstream project or a reservoir provided solely to serve as an afterbay. This scheme is more attractive if the cost of the afterbay is shared with other purposes.

LOAD OF HYDROPOWER PROJECTS

The demand for electrical energy is known as load. The ratio of the average power demand to peak power demand for the time period under consideration is known as the load factor, and it is computed daily, weekly, monthly, or annually. Thus,

$$\text{Load factor} = \frac{\text{Average power demand}}{\text{Peak power demand}} \tag{3}$$

With respect to the type of load served, hydropower projects can be classified in two categories: base-load plants and peaking plants. Base-load plants are projects that generate hydroelectric power to meet the base-load demand (the demand that exists 100% of the time).

Peaking plants generate power to supplement base-load generation during periods of peak demands. These plants must have sufficient capacity to satisfy peak demands, and enough water should be available. In general, a peaking hydroelectric plant is desirable in a system that has thermal generation units to meet base-load demands.

ESTIMATING HYDROPOWER POTENTIAL

Hydroelectric power potential is determined on the basis of the critical period, as indicated by the historical stream flow record. If a project serves more than one purpose and if, while serving another purpose, some of the storage or stream flow is not available for power production, the stream flow data should be adjusted to reflect the “loss.” Losses such as evaporation and leakage must also be deducted from the available flow. The amount of power generated over a time, or energy, is expressed in kilowatt-hour (kWh):

$$\text{kWh} = 9.817QHT\eta \tag{4}$$

in which kWh is the hydropower generated during the period (kWh) and T is the number of hours in the period. Two methods are used to estimate the hydropower potential at a given site: the flow duration curve method and the sequential stream flow routing (SSR) method. In the first method, the flow duration curve developed at the site using the stream flow data is the basic input. The net head for various discharges is estimated, and the hydropower equation is used to estimate the power generated at many points on the flow duration curve. In this way, a power duration curve is developed. The average annual energy and dependable capacity can now be calculated. This method is simple and fast, but it cannot take into account the installed capacity and key project features, such as power plant characteristics.

In the SSR method, the time step size and period of analysis are chosen, and the operation of the reservoir is simulated. For each time period, the reservoir outflow is computed, and the amount of energy generated is calculated using Equation 4. The process is repeated for all the time steps. Now the average annual, monthly, etc., generation as well as firm power can be computed. This method can take into account the reservoir and power plant characteristics. The results of this method are more realistic compared to the flow duration curve method.

OPERATION OF HYDROPOWER RESERVOIRS

Hydropower reservoirs store water to ensure supply as well as hydraulic head. During the filling season, the aim usually is to end the season with as much energy (as water) as possible in store in the system. The water stored in an upper reservoir (higher elevation) has higher potential energy, so upper reservoirs in a series should be filled first. After generating energy at an upper reservoir, the water can be captured in a downstream (lower) reservoir where energy can be generated again. The same logic also holds for any spill from an upper reservoir. Of course while storing water, one also has to examine the compatibility of hydropower use with the other uses of water.

During the drawdown, the objective is to maximize hydropower production for a given total storage amount vis-à-vis the demands. Recall that a system can generate the maximum amount of power when all reservoirs are full (the hydraulic heads will be the highest). If the available water is limited, it should be allocated among the reservoirs to maximize hydropower production. The governing variables are the storage and power generating capacities of reservoirs, the inflows, and the efficiencies of power plants. Note that in smaller reservoirs, the rate of increase of head per unit volume of additional water is higher compared to a large reservoir (that has more surface area). Referring to Fig. 2, the volume of water needed to increase the head by **x** units in a smaller reservoir (V_1) is less than the volume needed (V_2) to increase the head in a larger reservoir by the same increment.

The combination of the important variables in hydropower production, reservoir capacity, volume of inflows, and efficiency of power plant determines the ranking of the reservoirs to produce power. When reductions in storage are necessary, they are made from reservoirs that are least able to produce power. Conversely, an increase

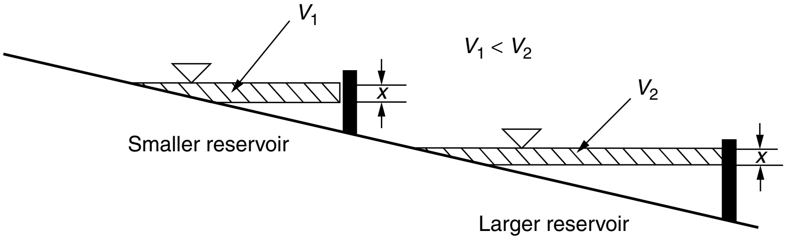

Figure 2. Change in head as capacities vary (after Reference 2).

in storage should be in reservoirs that have the greatest ability to produce power. The following reservoir ranking ratio can be used (2):

$$V_i = a_i e_i \left(\sum_{j=1}^{i} I_j \right) \quad (5)$$

where V_i is the increased power production per unit increase in storage, a_i is unit change in hydropower head per unit change in storage, and e_i is the power generation efficiency, all for reservoir i, and I_j is direct inflows and releases into reservoir j, for all reservoirs upstream of reservoir i. Note that reservoir 1 is the uppermost reservoir in the series. The reservoirs are ordered according to V_i values, and filling begins at the highest value of V_i and proceeds in descending order.

A number of software packages are available to analyze reservoir operation for hydropower generation. See HEC (3), McMahon and Mein (4), and Wurbs (5) for further details.

BIBLIOGRAPHY

1. IEA. (2001). *Key World Energy Statistics*. International Energy Agency, Paris. www.iea.org
2. Lund, J.R. and Guzman, J. (1999). Derived operating rules for reservoirs in series or in parallel. *J. Water Resour. Plann. Manage., ASCE* **125**(3): 143–153.
3. HEC. (1982). *HEC—5 Simulation of Flood Control and Conservation System, Program Users Manual, The Hydrologic Engineering Center*. U.S. Army Corps of Engineers, Davis, CA.
4. McMahon, T.A. and Mein, R.G. (1986). *River and Reservoir Yield*. Water Resources Publications, Colorado.
5. Wurbs, R.A. (1996). *Modelling and Analysis of Reservoir System Operations*. Prentice-Hall, New York.

HYDROELECTRIC RESERVOIRS AS ANTHROPOGENIC SOURCES OF GREENHOUSE GASES

Nicolas Soumis
Marc Lucotte
René Canuel
Sebastian Weissenberger
Stéphane Houel
Catherine Larose
University of Québec in Montréal
Montréal, Canada

Éric Duchemin
DREX Environment
Montréal, Canada

OVERVIEW

The ever-increasing demand for energy over the recent development of societies has spurred the construction of hydroelectric facilities. Since dams were first used to generate hydropower around 1890, their construction rate increased tremendously to peak during the 1950s and the 1980s (1). Today, about 25% of the 33,105 large dams (≥ 15 m height) listed by the International Commission on Large Dams (ICOLD) are used for hydropower generation (2) and currently provide 19% of the world's electricity supply (1). Although over 150 countries operate hydroelectric plants, Brazil, China, Canada, Russia, and the United States produce more than 50% of the world's hydropower (1). According to data from 1996, hydroelectric reservoirs worldwide cover an estimated 600,000 km^2 (3).

Apart from the benefits they provide, hydroelectric reservoirs entail several social and environmental drawbacks such as the loss of lands and displacement of peoples (1,4), downstream hydrological alterations (5–8), elevated methylmercury burdens in fish and human consumers (8–11), as well as the loss of biodiversity (1,8). Recently, they also proved to be nonnegligible sources of greenhouse gases (GHG) like carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O) (12–31). Owing to their implication in climate change (32–35), hydroelectric reservoirs have raised a sustained interest among both the scientific and political communities. Research during the last decade has brought valuable insights into the biogeochemical impacts stemming from their creation. Although it is now recognized that impoundment leads to both the disappearance of a GHG sink and the creation of a new anthropogenic GHG source (8,34), the duration of anthropogenic GHG emissions from these water bodies remains a matter of current debate (36).

ESTIMATES OF GHG EMISSIONS FROM HYDROELECTRIC RESERVOIRS WORLDWIDE

Table 1 provides statistical parameters for GHG fluxes at the surface of hydroelectric reservoirs. Spatial variability is an important issue that may interfere with regional estimates of GHG fluxes (31,33,35,36). Nevertheless, several data sets shown in Table 1 are normally distributed and mean values can therefore be used to estimate GHG fluxes on a latitudinal basis.

Exceptions are tropical CH_4 diffusive fluxes, which are not normally distributed, and N_2O diffusive fluxes for which data are too scarce. Likewise, bubbling fluxes are marred by either failure to comply with normality or insufficiency of data. In all of these cases, mean values must be used with caution as they poorly reflect the actual situation.

Generally speaking, tropical hydroelectric reservoirs possess the highest mean GHG fluxes, whereas the lowest mean fluxes are observed in temperate ones. A striking characteristic is the presence of CO_2 influxes in some temperate hydroelectric reservoirs from arid and semiarid regions (31). These distinctive GHG emission regimes are caused mainly by particular environmental conditions, such as temperature, which can influence GHG production processes (22,33,36). To a lesser extent, important differences in the mean age of hydroelectric reservoir sets in Table 1 may also intervene (33,35,36).

Table 1. GHG Fluxes at the Air/Water Interface of Hydroelectric Reservoirs Worldwide During the Open Water Period (12,14–29,31,37,38)

		Latitude (Climate) (Number of Reservoirs)[a]	Mean Age (y)	Range ($\mathrm{mg{\cdot}m^{-2}{\cdot}d^{-1}}$)	Median ($\mathrm{mg{\cdot}m^{-2}{\cdot}d^{-1}}$)	Mean ($\mathrm{mg{\cdot}m^{-2}{\cdot}d^{-1}}$)	Std Dev. ($\pm$ $\mathrm{mg{\cdot}m^{-2}{\cdot}d^{-1}}$)	Normality
Diffusive Fluxes	fCO_2	Boreal (12)	24	653 to 2500	1346	1459	600	yes
		Temperate (16)	44	−1195 to 2200	685	525	938	yes
		Tropical (21)	13	−142 to 13737	4789	5467	3746	yes
	fCH_4	Boreal (7)	29	3.5 to 22.8	10.0	10.8	6.5	yes
		Temperate (13)	49	1.3 to 15.0	7.0	6.7	3.7	yes
		Tropical (22)	16	5.7 to 233.3	21.7	50.5	66.0	no
	fN_2O	Boreal (3)	18	0.02 to 0.5	0.2	0.2	0.2	yes
		Temperate (0)	ND	ND	ND	ND	ND	ND
		Tropical (4)	8	5.6 to 800	34.8	218.8	388.3	no
Bubbling Fluxes	fCO_2	Boreal (2)	10	1.0	1.0	1.0	0	ND
		Temperate (1)	70	1.0	ND	ND	ND	ND
		Tropical (16)	14	0.02 to 26	0.3	2.5	6.5	no
	fCH_4	Boreal (5)	20	0.04 to 184.2	0.8	46.4	79.6	no
		Temperate (1)	70	14	ND	ND	ND	ND
		Tropical (21)	15	0 to 800	22.9	85.6	179.1	no
	fN_2O	Boreal (2)	27	0 to 0.03	0.02	0.02	0.02	ND
		Temperate (0)	ND	ND	ND	ND	ND	ND
		Tropical (0)	ND	ND	ND	ND	ND	ND

[a] Hydroelectric reservoirs that have been sampled multiple times are considered as seperate water bodies.

GHG DYNAMICS IN HYDROELECTRIC RESERVOIRS

GHG emissions are common in natural freshwater ecosystems (39–42). However, hydroelectric reservoirs differ from these natural ecosystems in several fundamental aspects (43–45) and therefore cannot be directly compared. Typically, these man-made water bodies are created by damming rivers or lakes. Consequently, the first outcome of hydroelectric reservoir creation is the flooding of terrestrial ecosystems and the subsequent death of the associated vegetation (10,13,33). Overall, this event likely corresponds to the permanent loss of a potential carbon sink (10,33,46), as exemplified by data in Table 2. The second outcome attributable to hydroelectric reservoirs is the creation of a net source of GHG through perturbations that are not observed in steady, natural aquatic ecosystems. These perturbations are highlighted by the GHG emission regimes of hydroelectric reservoirs, which generally overwhelm, at least temporarily, those of natural water bodies (10,13,15,16,21–23,33,35,56).

Emissions that would not have existed without the perturbations brought about by hydroelectric reservoirs are referred to as “anthropogenic GHG emissions.” Consequently, when evaluating these emissions, the following issues must be taken into account. First, the establishment of a valid reference state must be based on a comparison with GHG fluxes prior to flooding because of the major impacts of impoundment on terrestrial ecosystems (13,36). Second, the altered GHG dynamics following the impoundment phase as well as its temporal evolution must be considered (36).

Figure 1 summarizes the current knowledge on GHG dynamics in hydroelectric reservoirs. GHG emissions result from the superposition of short- and long-term

Table 2. Examples of Preimpoundment Carbon Fluxes from Terrestrial Ecosystems Commonly Flooded for the Creation of Hydroelectric Reservoirs[a]

Terrestrial Ecosystem	Study Site	Flux (g C·$\mathrm{m^{-2}{\cdot}y^{-1}}$)[b]	Reference
Boreal mesotrophic lake	Canada	+73	10
Boreal peatland (*Sphagnum* spp., *Ledum groenlandicum*)	Canada	−20 to −50	10, 47
Boreal aspen (*Populus tremuloides*) forest, mature	Canada	−80 to −290	48
Boreal pine (*Pinus sylvestris*) forest, mature	Finland	−214 to −252	49
Boreal spruce (*Picea mariana*) forest, mature	Canada	−60 to −158	48, 50
Temperate beech (*Fagus grandifolia)* forest, mature, CO_2 only	Germany	−490 to −494	51
Temperate pine (*Pinus tæda*) forest, CO_2 only	United States	−69	52
Temperate Mediterranean alpine grassland, CO_2 only	United States	+29 to −132	53
Tropical mesic savanna, considering/not considering bush fires	Australia	−100/−300	54
Tropical rain forest, mature	Brazil	+80 to −390	55
Natural lakes, worldwide average, CO_2 only	many countries	+70	40

[a] Carbon storage capacity of terrestrial ecosystems depends on several factors, such as their type, maturity level, and disturbance history, as well as climatic conditions (9,14,36,49–51).

[b] Positive sign indicates a carbon source; negative sign indicates a carbon sink.

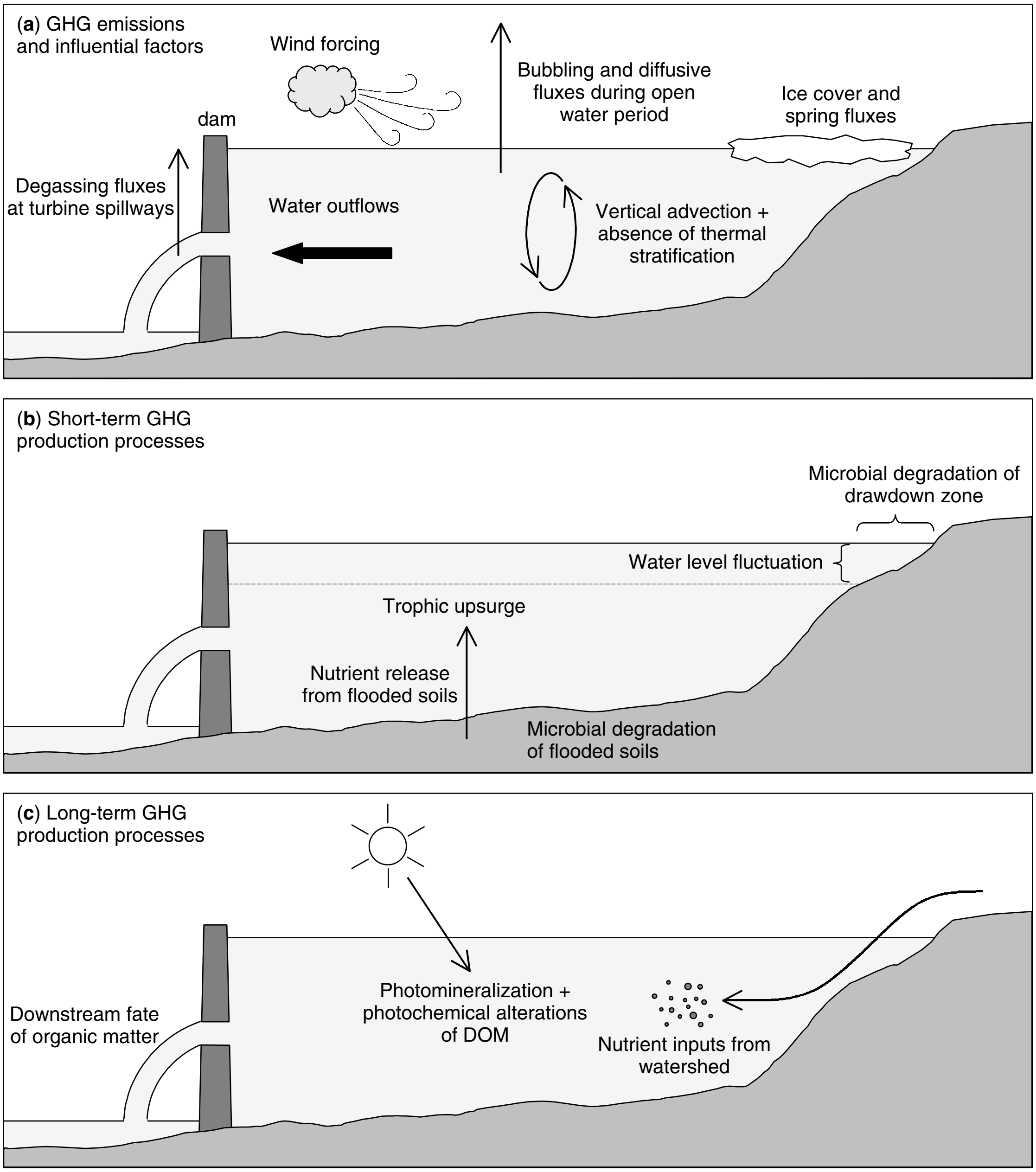

Figure 1. Dynamics of GHG in hydroelectric reservoirs. GHG emissions (**a**) are the result of short-term (**b**) and long-term (**c**) production processes.

production processes. Fundamental mechanisms involved in the production and emission of GHG are exhaustively described in limnology handbooks (57,58). Therefore, only the characteristics of hydroelectric reservoirs that lead to anthropogenic GHG emissions are hereafter described.

GHG Emission Processes and Influential Factors in Hydroelectric Reservoirs

GHG Fluxes at the Air/Water Interface During the Open Water Period. During the open water period, a fraction of GHG escapes from hydroelectric reservoirs through vertical advection (bubbling fluxes) (59) and molecular diffusion (diffusive fluxes) (60,61). GHG emission levels at the air/water interface may vary widely both within and between hydroelectric reservoirs, depending on the factors presented in Table 3.

Hydrometeorological forcings induce a particular hydrodynamic within hydroelectric reservoirs that enhance air/water interface GHG fluxes (36,43). For instance, wind generates surface waves and the vertical advection (Langmuir circulation) of water masses (43), thereby promoting gas transfer velocity across the air/water interface (61,62). For hydroelectric reservoirs with relatively extended surface areas and large fetches, wind leverage can be substantial (14,22,63,64). The operation regime of hydroelectric reservoirs also contributes to hydrodynamic features, which augment

Table 3. Main Factors Identified to Be Responsible for Inter- and Intra-Reservoir Variations in GHG Emission Levels at the Air/Water Interface (16,21,22,27,31,36)

Inter-Reservoir Variability	
Internal factors(pertaining to hydroelectric reservoir)	age of impoundment surface area morphometry and bathymetry physicochemistry operation regime
Regional factors(pertaining to watershed)	type of ecosystems flooded climatic feature pedology and geological basement ecological features
Intra-Reservoir Variability	
Spatial heterogeneities	water column depth water column temperature presence or absence of macrophytes hydrodynamics (main channel vs. protected coves) exposure to meteorological forcings (wind and sun)
Temporal heterogeneities	interannual (age, variations in climate and operation regime) season cycle (water level, biology, ice cover if applicable) day-night cycle (biology, especially photosynthesis)

GHG emissions. Intermittent huge water outflows for hydropower generation enable vertical advection and impede thermal stratification of the water column (43). This frequent situation in hydroelectric reservoirs (23,26,31,45) allows hypolimnetic water to mix with the upper layers, thereby promoting GHG evasion to the atmosphere.

GHG Fluxes at the Air/Water Interface During Spring. During wintertime, ice covers boreal and some temperate freshwater bodies. As GHG are still being produced throughout this period (26), gas bubbles accumulate under the ice. When the ice melts in the spring, GHG evade to the atmosphere at greater rates than during the open water period (22,26). However, as most studies on hydroelectric reservoirs are conducted in summer, spring GHG fluxes are often overlooked.

Degassing Fluxes at Turbine Spillways. GHG also escape from the water column at turbine spillways when hydropower is produced. Such degassing fluxes are caused by the sudden change in pressure as well as the increased air/water exchange surface (17,19) and strongly depend on the volume of outflows, the GHG concentrations within the water column, and the depth of water outlets on the dam. Although an important term of the GHG budget of hydroelectric reservoirs (36), degassing fluxes have seldom been assessed (19,31,46).

GHG Production Processes in Hydroelectric Reservoirs

Flooded Soils and Drawdown Zones. Hydroelectric reservoirs create a new interface between water and soils, thereby triggering the onset of major biogeochemical perturbations. Following impoundment, the physicochemical features of flooded soils promote the production of GHG through microbial degradation of organic matter (10,13). Biomass degradation is generally associated with oxygen depletion at the bottom of tropical hydroelectric reservoirs (19,23,28,65), a phenomenon that seems to be transient or inexistent in boreal and temperate ones (16,26,31,66).

Drawdown zones refer to the extended portion of hydroelectric reservoir banks that are periodically flooded by fluctuating water levels that occur during the operation phase. These zones act as intermittent water/soil interfaces and also constitute favorable sites for the production of GHG (67).

Typically, flooded organic matter undergoes an initial phase of rapid decomposition of labile compounds, followed by the slower decomposition of more recalcitrant compounds (13,28). Microbial degradation of flooded soils thus decreases with time and, according to studies on both carbon loss and isotopic tracers (δ^{13}C), appears to be limited to approximately ten years (68,69).

Trophic Upsurge. Following impoundment, a fraction of compounds from flooded soils and drawdown zones are released into the water column of hydroelectric reservoirs (28,68,70). This nutrient (C, N, P) pulse increases water turbidity, alters feeding habits as well as trophic levels of the planktonic community, and ultimately disrupts its initial photosynthesis:respiration ratio (22,69–73). Overall, these drastic shifts within the planktonic community structure provoke the transition of the aquatic system to a more allotrophic state (trophic upsurge) (74). Consequently, respiration overwhelms photosynthesis, thus enhancing the production of GHG in the water column.

Figure 2 illustrates the temporal progression of bacterioplankton in hydroelectric reservoirs. After the impoundment phase, bacterial biomass increases sharply (Fig. 2a). Then, as nutrient inputs from flooded soils gradually decrease, physicochemical parameters within the

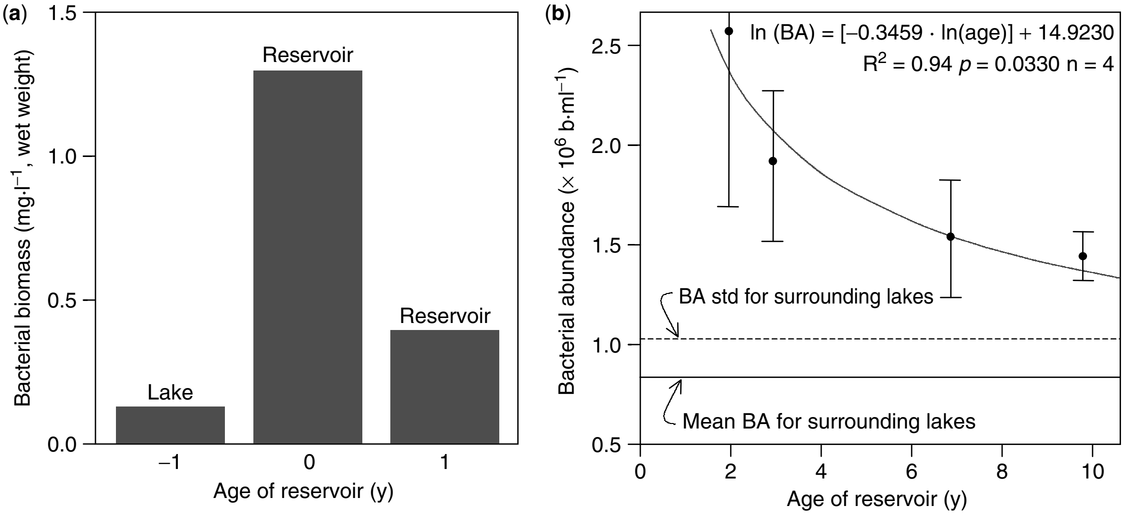

Figure 2. Temporal progression of bacterioplankton in reservoirs. Year "0" corresponds to the impoundment event. (**a**) Bacterioplankton biomass before and after the impoundment of a experimental reservoir created from a small boreal lake (70). (**b**) Evolution of bacterioplankton abundance (BA) following the impoundment of the boreal hydroelectric reservoir Laforge-1 (22,38,75). Error bars are standard deviation.

water column tend to evolve toward natural conditions (66,76,77). In accordance with the bottom-up theory, bacterial abundance also declines (65,70), but, as suggested in (Fig. 2b), a substantial time lag exists between the depletion of flooded nutrients and the anticipated onset of an equilibrium state within the bacterioplanktonic population.

Allochthonous Inputs from the Watershed. The actual contribution of allochthonous organic matter to anthropogenic GHG emissions from hydroelectric reservoirs is unclear (36). Nevertheless, the fact that long-term GHG emissions from certain hydroelectric reservoirs still overwhelm those of natural water bodies (33,35) and do not stem from the degradation of flooded soils (68,69) supports the hypothesis that these anthropogenic emissions are fueled by inputs from the watershed (22,36,56,78). If so, allochthonous inputs would only contribute to anthropogenic GHG emissions from hydroelectric reservoirs if they are mineralized to a greater extent than in natural ecosystems.

Two phenomena may account for such a case. On the one hand, photochemical reactions play a key role in the carbon cycle of aquatic systems (79–84) through both direct photomineralization of dissolved organic matter (DOM) into CO_2 (85–89) and molecular alterations that modify the bioavailability of DOM to bacteria (90,91). Empirical data suggest that these mechanisms play a greater role in hydroelectric reservoirs than in natural water bodies. For instance, a study conducted on a boreal hydroelectric reservoir has indicated that DOC photoreactivity exceeded that of boreal and temperate natural lakes significantly (94). This situation is not due to qualitative aspects of DOM, but rather to differences in age of water, as well as iron and manganese concentrations (92).

On the other hand, hydroelectric reservoirs disrupt natural streamflows and increase the water residence time (7,8,36,93), thereby retaining organic matter and hypothetically allowing for increased degradation. It is thus likely that hydroelectric reservoirs modify the mineralization pattern of organic matter at the scale of the watershed. However, this important issue has yet to be addressed.

DURATION OF ANTHROPOGENIC GHG EMISSIONS FROM HYDROELECTRIC RESERVOIRS

As shown in (Fig. 3) and elsewhere (14,33,35), GHG emissions significantly decline with the age of hydroelectric reservoirs. The duration of anthropogenic GHG emissions is closely related to the perturbations caused by their creation and operation. As these perturbations occur at different time scales (Fig. 1), anthropogenic GHG emissions from hydroelectric reservoirs may last for decades (33,35).

The duration of anthropogenic GHG emissions from hydroelectric reservoirs can be grossly estimated through a heuristic comparison with natural lakes. Higher GHG emission levels from the former likely indicate that certain biogeochemical perturbations still intervene. (Fig. 3) suggests that CO_2 diffusive fluxes exceeding the mean of natural lakes worldwide (704 mg $CO_2 \cdot m^{-2} \cdot d^{-1}$) (40) can last up to 70 years. However, weak regression coefficients (R^2) between the magnitude of GHG fluxes and the age of hydroelectric reservoirs, in (Fig. 3) and in previous literature reviews (33,35), indicate that the latter parameter poorly explains the decline of GHG fluxes. Ultimately, the accurate determination of the duration of anthropogenic GHG emissions will depend on an exhaustive appraisal of the mechanisms involved in the production of GHG.

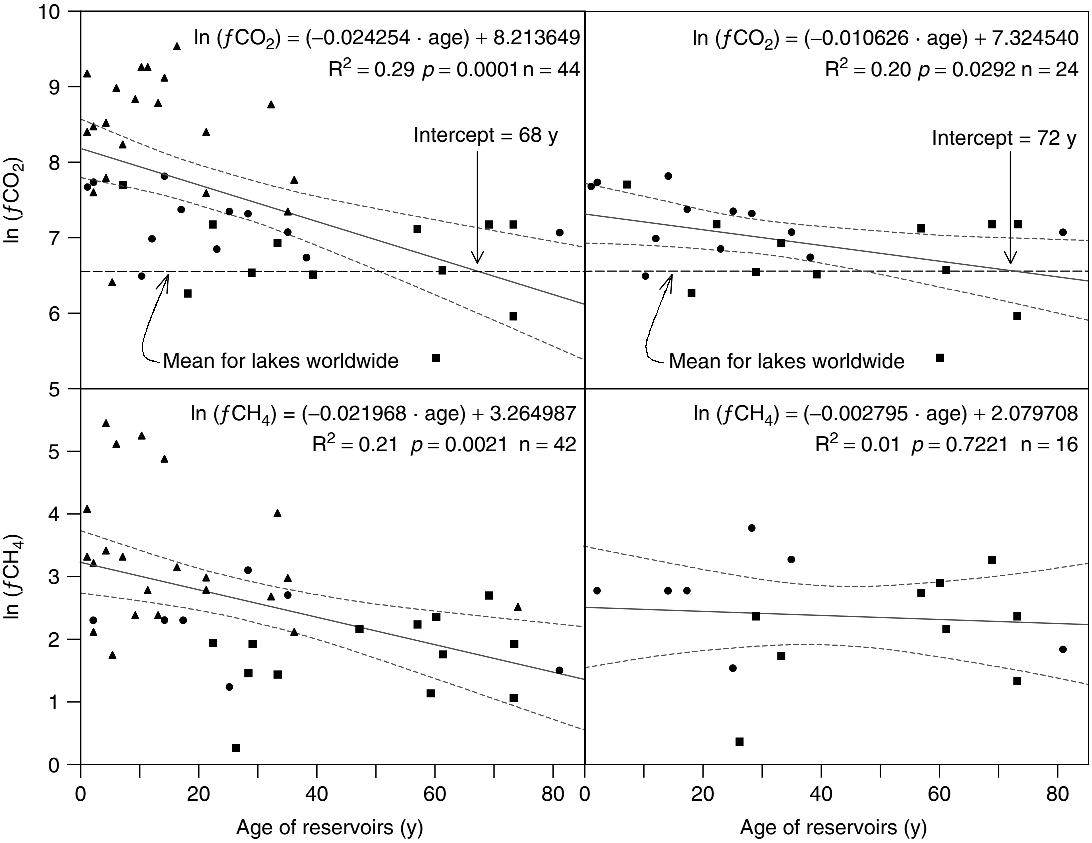

Figure 3. Temporal evolution of CO_2 and CH_4 diffusive fluxes from hydroelectric reservoirs during the open water period (12,14–16,19–29,31,37,92). Left panel: Includes hydroelectric reservoirs from all climatic zones (boreal, temperate, and tropical). Right panel: Boreal and temperate hydroelectric reservoirs only. Hydroelectric reservoirs with CO_2 influxes have been discarded and those that have been sampled during different years are considered as seperate water bodies. Circles = boreal; squares = temperate; triangles = tropical. Dashed curved lines represent the 95% confidence intervals.

POLITICAL ASPECTS

As a result of the threat of climate change, anthropogenic GHG emissions have become of major concern for political entities. In 1990, the Intergovernmental Panel on Climate Change's (IPCC) First Assessment Report set the basis for the United Nations Framework Convention on Climate Change (UNFCCC), which was further implemented by the Kyoto Protocol in 1997 (34). The UNFCCC establishes an overall international structure to deal with climate change. Its mandate is to require signatory countries to gather relevant information on GHG emissions (national GHG inventories) as well as to develop strategies to prevent or mitigate these emissions (94).

From this perspective, the energy sector represents an area where GHG abatements are possible through strategies such as the "clean development mechanism" (CDM). As defined by the Marrakesh Accords, a project is considered a CDM provided that it substitutes a higher GHG-emitting option (95) [for instance, building hydropower plants instead of thermal ones, the latter being commonly recognized as emitting more GHG (28,33,35)]. Certificates of GHG emission reduction can then be acquired through the World Bank's Prototype Carbon Fund (PCF) (96).

Within the framework of the CDM, anthropogenic GHG emissions from hydroelectric reservoirs cannot be ignored. However, their inclusion into national GHG inventories is a prerogative that has yet to be achieved. To date, only GHG emissions produced by the degradation of flooded soils are considered as anthropogenic by IPCC (97). Nevertheless, the eventual part of GHG emissions stemming from allochthonous inputs should also be considered in the next IPCC guidelines, provided that enough evidence exists to substantiate such a situation. In any case, the inclusion of hydroelectric reservoirs into national GHG inventories is planned for the second commitment under the Kyoto Protocol, effective as of 2012. Countries that are characterized by large surface areas covered by hydroelectric reservoirs should then observe a significant increase in their national GHG inventories.

BIBLIOGRAPHY

1. World Commission on Dams (WCD). (2000). *Dams and Development. A New Framework for Decision-Making*. Earthscans Publications, London, p. 448.
2. International Commission on Large Dams (ICOLD). (2003). *World Register of Dams*. ICOLD, Paris, p. 340.

3. Pearce, F. (1996). *New Sci.* **150**: 28–31.
4. Cernea, M.M. (1988). *Involuntary Resettlement in Development Projects: Policy Guidelines in World Bank-Financed Projects*. Technical paper no. 80, World Bank, Washington, DC, p. 92.
5. Dynesius, M. and Nilsson, C. (1994). *Science* **266**: 753–762.
6. McCully, P. (1996). *Silenced Rivers: The Ecology and Politics of Large Dams*. Zed Books, London, p. 368.
7. Gillilan, D.M. and Brown, T.C. (1997). *Instream Flow Protection*. Island Press, Washington, DC, p. 427.
8. Rosenberg, D.M. et al. (1997). *Environ. Rev.* **5**: 27–54.
9. Hecky, R.E. et al. (1991). *Advances in Mercury Toxicology*. T. Suzuki (Ed.). Plenum Press, New York, pp. 33–52.
10. Kelly, C.A. et al. (1997). *Environ. Sci. Technol.* **31**: 1334–1344.
11. Lucotte, M. et al. (Eds.). (1999). *Mercury in the Biogeochemical Cycle*. Springer, Berlin, p. 352.
12. Smith, L.K. and Lewis, W.M. (1992). *Global Biogeochem. Cycles*. **6**: 323–338.
13. Rudd, J.W.M. et al. (1993). *Ambio*. **22**: 246–248.
14. Keller, M. and Stallard, R.F. (1994). *J. Geophys. Res.* **99**: 8307–8319.
15. Kelly, C.A. et al. (1994). *EOS Trans. AGU.* **75**: 332–333.
16. Duchemin, É. et al. (1995). *Global Biogeochem. Cycles*. **9**: 529–540.
17. Galy-Lacaux, C. (1996). *Modifications des Échanges de Constituants Mineurs Atmosphériques Liées à la Création D'une Retenue Hydroélectrique*. Ph.D. Thesis, Paul Sabatier University, Toulouse, France, p. 200.
18. Hellsten, S.K. et al. (1996). *Measured Greenhouse Gas Emissions from Two Hydropower Reservoirs in Northern Finland*. IAEA advisory group meeting on assessment of greenhouse gas emissions from the full energy chain for hydropower, nuclear power, and other energy sources, Hydro-Québec, Montréal, pp. 1–12.
19. Galy-Lacaux, C. et al. (1997). *Global Biogeochem. Cycles*. **11**: 471–483.
20. Schellhase, H.U. et al. (1997). *Environ. Prof.* **19**: 48–57.
21. Duchemin, É. et al. (1999). *Étude sur la Production et L'émission de Gaz à Effet de Serre par les Réservoirs Hydroélectriques de L'entreprise et des Lacs Naturels Environnants*. Report no 21046-99027c, Hydro-Québec, Montréal, p. 47.
22. Duchemin, É. (2000). *Hydroélectricité et Gaz à Effet de Serre: Évaluation des Émissions des Différents Gaz et Identification des Processus Biogéochimiques de leur Production*. Ph.D. Thesis, Institute of Environmental Sciences, University of Québec in Montréal, Montréal, Canada, p. 308.
23. Duchemin, É. et al. (2000). *Verh. Internat. Verein. Limnol.* **27**: 1391–1395.
24. Lima, I.B.T. et al. (2000). *Verh. Internat. Verein. Limnol.* **27**: 1437–1440.
25. Matvienko, B. et al. (2000). *Verh. Internat. Verein. Limnol.* **27**: 1415–1419.
26. Huttunen, J.T. et al. (2002). *Global Biogeochem. Cycles* **16**: 3.1–3.17.
27. Lima, I.B.T. et al. (2002). *Verh. Internat. Verein. Limnol.* **28**: 438–442.
28. Ministério da Ciência e Tecnologia (MCT). (2002). *Emissões de Dióxido de Carbono e de Metano Pelos Reservatórios Hidrelétricos Brasileiros*. COPPE, Brasília, p. 119.
29. Åberg, J. et al. (2004). *Water Res.* **38**: 531–538.
30. Bergström, A.K. et al. (2004). *Arch. Hydrobiol.* **159**: 25–42.
31. Soumis, N. et al. (2004). *Global Biogeochem. Cycles* **18**: GB3022, doi: 10.1029/2003 GB002197.
32. Intergovernmental Panel on Climate Change (IPCC). (1996). *The Science of Climate Change*. Press Syndicate of the University of Cambridge, New York, p. 572.
33. St-Louis, V.L. et al. (2000). *BioScience* **50**: 766–775.
34. Intergovernmental Panel on Climate Change (IPCC). (2001). *Climate Change 2001: The Scientific Basis*. Cambridge University Press, New York, p. 944.
35. Duchemin, E. et al. (2002). *World Res. Rev.* **14**: 334–353.
36. World Commission on Dams (WCD). (2000). *Workshop on Dam Reservoirs and Greenhouse Gases (Part III)*. Secretariat of the World Commission on Dams, Cape Town, South Africa, p. 17.
37. Kelly, C.A. and Rudd, J.W.M. (2000). *BioScience* **50**: 766–775.
38. Soumis, N. unpublished data.
39. Kling, G.W. et al. (1992). *Hydrobiologia.* **240**: 23–36.
40. Cole, J.J. et al. (1994). *Science* **265**: 1568–1570.
41. del Giorgio, P.A. et al. (1997). *Nature* **385**: 148–151.
42. del Giorgio, P.A. et al. (1999). **80**: 1422–1431.
43. Ford, D.E. (1990). *Reservoir Limnology: Ecological Perspectives*. K.W. Thornton et al. (Eds.). John Wiley & Sons, New York, pp. 15–41.
44. Thornton, K.W. (1990). In: *Reservoir Limnology: Ecological Perspectives*. K.W. Thornton et al. (Eds.). John Wiley & Sons, New York, pp. 1–13.
45. Straškraba, M. (1998). *Internat. Rev. Hydrobiol.* **83**: 1–12.
46. Fearnside, P.M. (2002). *Water Air Soil Pollut.* **133**: 69–96.
47. Wieder, R.K. (2001). *Ecol. Appl.* **11**: 327–342.
48. Arain, M.A. et al. (2002). *Can. J. For. Res.* **32**: 878–891.
49. Kolari, P. et al. (2004). *Global Change Biol.* **10**: 1106–1119.
50. Potter, C.S. et al. (2001). *Can. J. For. Res.* **31**: 208–223.
51. Knohl, A. et al. (2003). *Agr. Forest Meteorol.* **118**: 151–167.
52. Luo, Y.Q. et al. (2003). *Global Biogeochem. Cycles* **17**: 1021.
53. Xu, L.K. and Baldocchi, D.D. (2004). *Agr. Forest Meteorol.* **123**: 79–96.
54. Williams, R.J. et al. (2004). *Func. Plant Biol.* **31**: 415–422.
55. Miller, S.D. et al. (2004). *Ecol. Appl.* **14**: S114–S126.
56. Weissenberger, S. et al. (1999). *Greenhouse gas emissions and carbon cycle in boreal reservoirs, Proceeding of the International Conference on Greenhouse Gas Emissions From Dams and Lakes*. Rio de Janeiro, Brasil, Dec. 4–5, 1998, COPPE, Rio de Janeiro.
57. Wetzel, R.G. (2001). *Limnology. Lake and River Ecosystems*, 3rd Edn. Academic Press, New York, p. 1006.
58. Kalff, J. (2002). *Limnology*. Prentice-Hall, Upper Saddle River, NJ, p. 592.
59. Chanton, J.P. and Whiting, G.J. (1995). In: *Biogenic Trace Gases: Measuring Emissions from Soil and Water*. P.A. Matson and R.C. Harriss (Eds.). Blackwell Science, Boston, pp. 98–125.
60. MacIntyre, S. et al. (1995). In: *Biogenic Trace Gases: Measuring Emissions from Soil and Water*. P.A. Matson and R.C. Harriss (Eds.). Blackwell Science, Boston, pp. 52–97.
61. Cole, J.J. and Caraco, N.F. (1998). *Limnol. Oceangr.* **43**: 647–656.
62. Crusius, J. and Wanninkhof, R. (2003). *Limnol. Oceanogr.* **48**: 1010–1017.
63. Duchemin, E. et al. (1999). *Environ. Sci. Technol.* **33**: 350–357.

64. Barrette, N. (2000). *Réservoirs Hydroélectriques du Moyen Nord Québécois: Mesures et Modélisation Numérique des Flux et des Concentrations de Gaz à Effet de Serre*. Ph.D. Thesis, Institute of Environmental Sciences, Unversity of Québec in Montréal, Montréal, Canada, p. 167.
65. Dumestre, J.F. et al. (1999). *Hydrobiologia.* **400**: 75–83.
66. Shetagne, R. (1994). *Arch. Hydrobiol. Beih. Ergebn. Limnol.* **40**: 223–229.
67. Jugnia, L.B. et al. (2004). *Can. J. Microbiol.* submitted.
68. Houel, S. (2003). *Dynamique de la Matière Organique Terrigène dans les Réservoirs Boréaux*. Ph.D. Thesis, Institute of Environmental Sciences, University of Québec in Montréal, Montréal, Canada, p. 121.
69. Hélie, J.F. (2004). *Géochimie et Flux de Carbone Organique et Inorganique dans les Milieux Aquatiques de l'est du Canada*. Ph.D. Thesis, Dept. of Earth Sciences, University of Québec in Montréal, Montréal, Canada, p. 204.
70. Paterson, M.J. et al. (1997). *Can. J. Fish Aquat. Sci.* **54**: 1088–1102.
71. Pinel-Alloul, B. and Méthot, G. (1984). *Int. Revue. Ges. Hydrobiol.* **69**: 57–78.
72. Montgomery, S. (2000). *Cycle du Méthylmercure dans les Réservoirs Hydroélectriques et les Lacs Naturels Boréaux tel Que Précisé par la Caractérisation Biogéochimique de la Matière Particulaire en Suspension et du Plancton*. Ph.D. Thesis, Institute of Environmental Sciences, University of Québec in Montréal, Montréal, Canada, p. 260.
73. Weissenberger, S. in preparation.
74. Roué-Le Gall, A. (2001). *Modélisation du Fonctionnement Écologique d'un Réservoir Nouvellement mis en Eau: Estimation des Flux de Carbone Entre les Différents Compartiments Biologiques*. Ph.D. Thesis, Geological Informatic Center, École Nationale Supérieure des Mines, Paris, p. 210.
75. Planas, D. et al. (2002). *Comparison of Planktonic Metabolism Between Reservoirs and Lakes in Northern Québec (Canada)*. Phytoplankton Productivity Conference, Bangor, UK, Mar. 18–22.
76. Kenttämies, K. (1980). *Characteristics of the Water of Finnish Man-Made Lakes*. Report no 39, National Board of Waters, Finland, pp. 13–25.
77. Grimard, Y. and Jones, H.G. (1982). *Can. J. Fish Aquat. Sci.* **39**: 1473–1483.
78. Lavandier, P. (1990). *Hydrobiologia* **207**: 79–86.
79. Tranvik, L.J. (1992). *Hydrobiologia* **229**: 107–114.
80. Wetzel, R.G. et al. (1995). *Limnol. Oceanogr.* **40**: 1369–1380.
81. Bertilsson, S. and Allard, B. (1996). *Arch. Hydrobiol. Beih. Ergebn. Limnol.* **48**: 133–141.
82. Cole, J.J. (1999). *Ecosystems* **2**: 215–225.
83. Vähätalo, A.V. et al. (1999). *Biodegradation* **10**: 415–420.
84. Pace, M.L. et al. (2004). *Nature* **427**: 240–243.
85. Salonen, K. and Vähätalo, A. (1994). *Environ. Int.* **20**: 307–312.
86. Granéli, W. et al. (1996). *Limnol. Oceanogr.* **41**: 698–706.
87. Bertilsson, S. and Tranvik, L.J. (1998). *Limnol. Oceanogr.* **43**: 885–895.
88. Granéli, W. et al. (1998). *Biogeochemistry* **43**: 175–195.
89. Zepp, R.G. et al. (1998). *J. Photochem. Photobiol.* **46**: 69–82.
90. Moran, M.A. and Zepp, R.G. (1997). *Limnol. Oceanogr.* **42**: 1307–1316.
91. Moran, M.A. and Covert, J.S. (2003). In: *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter*. S. Findlay and R. Sinsabaugh (Eds.). Academic Press, New York, pp. 243–262.
92. Soumis, N. (2005). *Photochimie et Réactivité du Carbone Organique Dissous dans les Milieux Lentiques Boréaux et Autres Travaux sur les Émissions de Gaz à Effet de Serre*. Ph.D. Thesis, Institute of Environmental Sciences, University of Québec in Montréal, Montréal, Canada.
93. Vörösmarty, C.J. et al. (1997). *Ambio* **26**: 210–219.
94. United Nations (UN). (1992). Text of the United Nation Framework Convention on Climate Change. Available: http://unfccc.int/files/essential_background/background_publications_htmlpdf/application/pdf/conveng.pdf.
95. United Nations Framework Convention on Climate Change (UNFCCC). (2001). Marrakesh Accords and the Marrakesh Declaration (draft version). Available: http://unfccc.int/cop7/accords_draft.pdf (as on Oct. 2004).
96. Prototype Carbon Fund projects. (2004). Available: http://carbonfinance.org/pcf/router.cfm?Page=Projects (as on Oct. 2004).
97. Penman, J. et al. (Eds.). (2003). *Good Practice Guidance for Land-Use, Land-Use Change and Forestry*. Institute for Global Environmental Strategies (IGES), Hayama, Japan, p. 588.

READING LIST

Rosa, L.P. and Schaeffer, R. (1994). *Ambio.* **23**: 164–165.

Goulden, M.L. et al. (1998). *Science* **279**: 214–217.

HYDROLOGIC PERSISTENCE AND THE HURST PHENOMENON

DEMETRIS KOUTSOYIANNIS
National Technical University
Athens, Greece

INTRODUCTION

Unlike common random series like those observed, for example, in games of chance (dice, roulette, etc.), hydrologic (and other geophysical) time series have some structure, that is, consecutive values of hydrologic time series depend on each other. It is easy to understand that, for instance, in a monthly river flow series, a month of very high flow is likely to be followed by a month of high flow, too. River flow is interrelated to groundwater storage, so high flow indicates that groundwater storage will be high, too, and given that the groundwater flow is a slow process, it is expected that its contribution to river flow in the next month can be significant. This explains the dependence of consecutive values of hydrologic time series, which have been known as *short-range* (or *short-term*) *dependence, persistence, or memory*.

Interestingly, however, there is another kind of dependence observed on larger timescales, known as *long-range* (or *long-term*) *dependence, persistence, or memory*. This was discovered by Hurst (1), while investigating the discharge time series of the Nile River in the framework of the design of the Aswan High Dam, and was found in many other hydrologic and geophysical time series. This behavior is the tendency of wet years to cluster into multiyear wet periods or of dry years to cluster into multiyear drought periods. The terms “Hurst phenomenon” and “Joseph effect”, due to Mandelbrot (2) from the biblical story of the “seven years of great abundance” and the “seven years of famine,” have been used as alternative names for the same behavior. Since its original discovery, the Hurst phenomenon has been verified in several environmental quantities such as (to mention a few of the more recent studies) in wind power (3); global or point mean temperatures (4–6); flows of several rivers such as the Nile (7,8); the Warta, Poland (9), Boeoticos Kephisos, Greece (5), and the Nemunas, Lithuania (10); inflows of Lake Maggiore, Italy (11); indexes of North Atlantic oscillation (12); and tree-ring widths, which are indicators of past climate (8). In addition, the Hurst phenomenon has gained new interest today due to its relation to climate changes (5,13–15).

The possible explanation of long-term persistence must be different from that of short-term persistence discussed above. This will be discussed later. However, its existence is easy to observe even in a time series plot, provided that the time series is long enough. For example, in Fig. 1 (up) we have plotted one of the most well-studied time series, that of the annual minimum water level of the Nile River for the years 622–1284 A.D. (663 observations), measured at the Roda Nilometer near Cairo (16, pp. 366–385; 17). In addition to the plot of the annual data values versus time, the 5-year and 25-year averages are also plotted versus time. For comparison, in the lower panel of Fig. 1, we have also plotted a series of white noise (consecutive independent identically distributed random variates) with statistics the same as those of the Nilometer data series. We can observe that the fluctuations in the aggregated processes, especially of the 25-year average, are much greater in the real-world time series than in the white noise series. Thus, fluctuations in a time series on a large scale distinguish it from random noise.

STOCHASTIC REPRESENTATION OF THE HURST PHENOMENON

Quantification of long-term persistence is better expressed mathematically using the theory of stochastic processes. Let X_i denote a stochastic representation of a hydrometeorologic process where $i = 1, 2, \ldots,$ denotes discrete time with a time step or scale which for the purposes of this article is annual or multiannual. It is assumed that the process is stationary, a property that does not hinder exhibiting multiple scale variability. The stationarity assumption implies that its statistics are not functions of time. Therefore, we can denote its statistics without reference to time, that is, its mean as $\mu := E[X_i]$, its autocovariance as $\gamma_j := \mathrm{Cov}[X_i, X_{i+j}] (j = 0, \pm1, \pm2, \ldots)$, its autocorrelation $\rho_j := \mathrm{Corr}[X_i, X_{i+j}] = \gamma_j/\gamma_0$, and its standard deviation $\sigma := \sqrt{\gamma_0}$. Further, we assume ergodicity, so that these statistics can be estimated from a unique time series substituting time averages for expected values.

Let k be a positive integer that represents a timescale larger than the basic timescale of the process X_i. The aggregated stochastic process on that timescale is denoted as

$$Z_i^{(k)} := \sum_{l=(i-1)k+1}^{i\,k} X_l \tag{1}$$

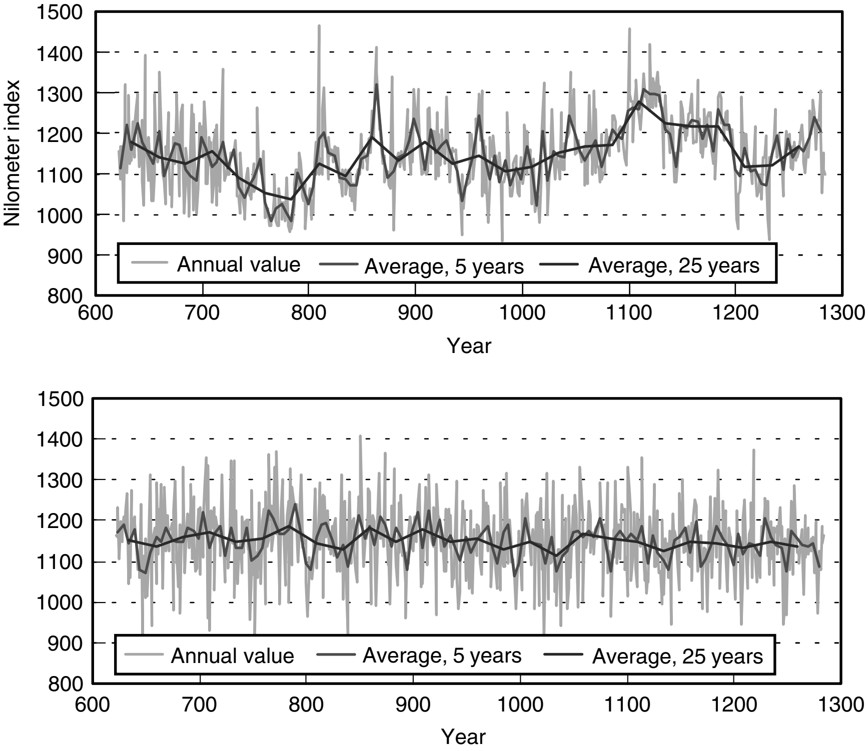

Figure 1. (Up) Plot of the Nilometer series indicating the annual minimum water level of the Nile River for the years 622–1284 A.D. (663 years); (down) a white noise series with same mean and standard deviation, for comparison.

The statistical characteristics of $Z_i^{(k)}$ for any timescale k can be derived from those of X_i. For example, the mean is

$$E[Z_i^{(k)}] = k\ \mu \tag{2}$$

and the variance and autocovariance (or autocorrelation) depend on the specific structure of γ_j (or ρ_j). In a process that exhibits the Hurst phenomenon, the variance $\gamma_0^{(k)}$ for timescale k is related to that of the basic scale γ_0 by

$$\gamma_0^{(k)} := \mathrm{Var}[Z_i^{(k)}] = k^{2H}\gamma_0 \tag{3}$$

where H is a constant known as the Hurst coefficient whose values are in the interval [0.5, 1]. The value $H = 0.5$ corresponds to random noise, whereas values in the interval [0, 0.5] are mathematically possible but without interest in hydrology. Consequently, the standard deviation is a power law of the scale or level of aggregation k with exponent H, that is,

$$\sigma^{(k)} := (\gamma_0^{(k)})^{1/2} = k^H\sigma \tag{4}$$

This simple power law can be easily used for detecting whether a time series exhibits the Hurst phenomenon and for determining the coefficient H, which is a measure of long-term persistence. Equation 4 calls for a double logarithmic plot of standard deviation $\sigma^{(k)}$ of the aggregated process $Z_i^{(k)}$ versus timescale k. In such a plot, called an aggregated standard deviation plot, Hurst behavior is manifested as a straight line arrangement of points corresponding to different timescales, whose slope is the Hurst coefficient. An example is depicted in Fig. 2 for the Nilometer series of Fig. 1. Clearly, the plot of the empirical estimates of standard deviation is almost a straight line on the logarithmic diagram of slope 0.85. For comparison, we have also plotted the theoretical curve for the white noise of slope equal to 0.5, significantly departing from historical data.

By virtue of Eq. 3, it can be shown that the autocorrelation function, for any aggregated timescale k, is independent of k and is given by

$$\rho_j^{(k)} = \rho_j = (1/2)(|j+1|^{2H} + |j-1|^{2H}) - |j|^{2H}$$
$$\approx H(2H-1)|j|^{2H-2} \tag{5}$$

which shows that autocorrelation is a power function of the lag. Consequently, the autocovariance $\gamma_j^{(k)} = \gamma_0^{(k)}\rho_j^{(k)}$ is a power law of both the scale k (with exponent $2H$) and the lag j (with exponent $2H - 2$).

The power spectrum of the process,

$$s_\gamma^{(k)}(\omega) := 2\sum_{j=-\infty}^{\infty} \gamma_j^{(k)}\cos(2\pi j\omega) \tag{6}$$

is given approximately by

$$s_\gamma^{(k)}(\omega) \approx 4\ (1\text{—}H)\gamma_0^{(k)}(2\ \omega)^{1-2H} \tag{7}$$

which is a power law of both the scale k (with exponent $2H$) and the frequency ω (with exponent $1 - 2H$).

The power law Eqs. 5 and 7 can be used, in addition or alternatively to Eqs. 3 and 4, to detect the Hurst behavior of a time series. Note that Hurst's (1) original formulation to detect this behavior was based on another quantity, the so-called rescaled range, which corresponds to the cumulated process of inflow minus outflow of a hypothetical infinite reservoir.

Equations 3–7 describe the second-order properties of the process $Z_i^{(k)}$. A generalization is possible, if we assume that the process of interest exhibits scale invariant properties in its (finite-dimensional joint) distribution function:

$$(Z_i^{(k)} - k\ \mu) \overset{\mathrm{d}}{=} \left(\frac{k}{l}\right)^H (Z_j^{(l)} - l\ \mu) \tag{8}$$

where the symbol $\overset{\mathrm{d}}{=}$ stands for equality in distribution. In this case, Eq. 3 can be obtained from Eq. 8 by setting $i = j = l = 1$ and taking the variance of both sides. Equation 8 defines X_i and $Z_i^{(k)}$ as stationary increments of a self-similar process. If, in addition, X_i (and hence $Z_i^{(k)}$) follows the normal distribution, then X_i (and $Z_i^{(k)}$) is called fractional Gaussian noise (FGN; (18)). Our interest here includes processes that may be not Gaussian, so we will limit the scaling property of Eq. 8 to second-order properties only and call the related process a *simple scaling signal* (SSS).

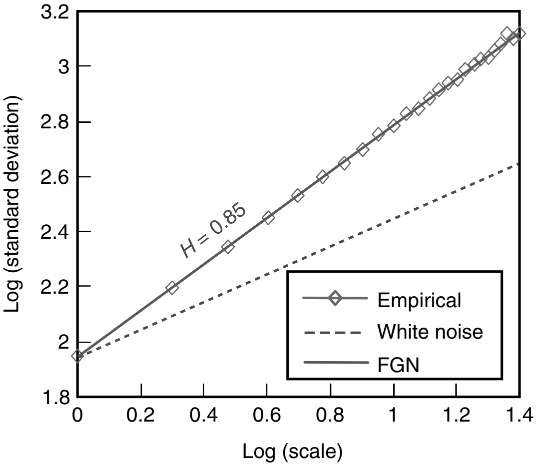

Figure 2. Aggregated standard deviation plot of the Nilometer series.

PHYSICAL EXPLANATIONS OF THE HURST PHENOMENON

As described in the Introduction, the concept of short-term persistence in hydrologic processes is easy to explain, whereas long-term persistence and the Hurst phenomenon are more difficult to understand. Mesa and Poveda (19) classify the Hurst phenomenon as one of the most important unsolved problems in hydrology and state that "something quite dramatic must be happening from a physical point of view." However, several explanations have been proposed. These can be classified in two categories, physically based and conceptual.

Klemeš (20) proposed an explanation that may be classified in the first category. According to this, the Hurst behavior of hydrologic records can be explained by representing the hydrologic cycle by a "circular cascade of semi-infinite storage reservoirs" where the output from one reservoir constitutes an essential part of input into the next. He showed that, even with an originally uncorrelated Gaussian forcing, the outputs grew progressively more Hurst-like as the complexity of the system increased, for example, the number of reservoirs in the cascade. Another example of such a hydrologic system was suggested in Klemeš (21).

Beran (17, pp. 16–20) describes two physically based model types that lead to system evolution (in time or space) with long-range dependence. The first model type applies to critical phenomena in nature such as phase transition (transition from a liquid to a gaseous phase or spontaneous magnetization of ferromagnetic substances). For some critical system temperature, the correlation of the system state at any two points decays slowly to zero, so the correlation in space can be represented by Eq. 5. The second type is related to models based on stochastic partial differential equations, which, under certain conditions, result in solutions with long-range dependence. These models provide sound links of long-range dependence with physics but are very complex.

A simple model of this category was studied by Koutsoyiannis (6). This model assumes a system with purely deterministic dynamics in discrete time, which, however, results in a time series with an irregular appearance exhibiting the Hurst phenomenon. The system dynamics is based on the simple map

$$x_i = g(x_{i-1}; \alpha) := \frac{(2-\alpha)\min(x_{i-1}, 1-x_{i-1})}{1-\alpha\min(x_{i-1}, 1-x_{i-1})} \tag{9}$$

where x_i is the system state, assumed to be scalar, at time i and $\alpha < 2$ is a parameter. This map, known as a generalized tent map, has been used in the study of dynamic systems. For example, the map approximates the relation between successive maxima in the variable $x(t)$ from the Lorenz equations that describe climatic dynamics (22, p. 150). Koutsoyiannis (6) demonstrated that this model can describe a system subject to the combined action of a positive and a negative feedback. If the parameter α is assumed to vary in time following the same map, $\alpha_i = g(\alpha_{i-1}; \lambda)$, then we obtain the double tent map,

$$u_i = G(u_{i-1}, \alpha_{i-1}; \kappa, \lambda) := g(u_{i-1}; \kappa\ \alpha_{i-1}) \tag{10}$$

Both parameters κ and λ should be < 2, whereas the domain of u_t is the interval [0, 1]. If the system domain is the entire line of real numbers, we can apply an additional transformation to shift from [0, 1] to $[-\infty, \infty]$:

$$x_t = \ln[u_t/(1-u_t)] \tag{11}$$

The model behavior with respect to parameters κ and λ is depicted in Fig. 3. We observe that for small values of the parameter κ, the time series synthesized by the model exhibits short-range dependence with a Hurst coefficient around 0.5. For large values of κ, the model yields runaway behavior. However, in between the two noninteresting areas, there is an area of parameter values, shaded in Fig. 3, in which the resulting time series exhibits long-range dependence. All three types of behavior are observed for negative values of λ (not shown in Fig. 3), whereas for $\lambda > 1$, only the runaway behavior is observed, regardless of the value of κ.

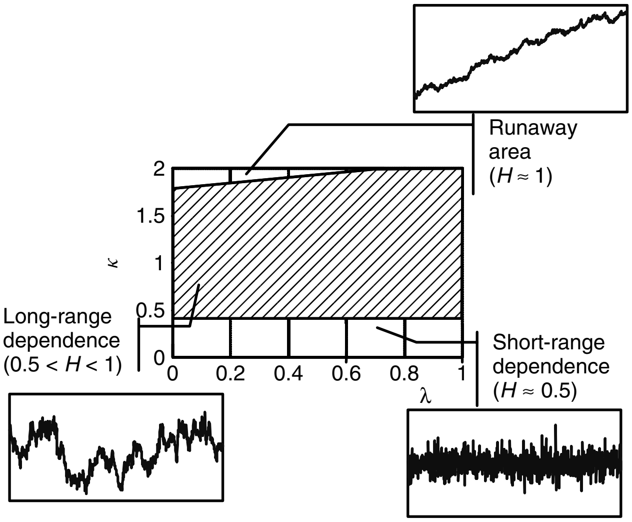

Figure 3. Schematic of the general behavior of the double tent map in terms of the ranges of its parameters κ and λ.

The second category, labeled here "conceptual," does not aim at explaining the physical mechanism leading to Hurst behavior of historical records of some natural or other processes but examines different stochastic mechanisms that might produce realizations resembling the patterns of the observed empirical time series. For example, Klemeš (20) analyzed several variants of the "changing mean" mechanism which assumes that the mean of the process is not a constant determined by the arithmetic mean of the record but varies through time. Specifically, he performed numerical experiments with the following Gaussian random processes:

1. a process with the mean alternating periodically between two values after constant time intervals called "epochs";
2. a process with a monotonic linear trend in the mean, throughout the entire series length;
3. a process similar to 1 but with epoch lengths taking two different values with probabilities p and $1 - p$; and
4. a process with a Gaussian-distributed mean randomly varying from epoch to epoch, the epoch length also varying randomly and following either a uniform, exponential, or (single parameter) Pareto distribution.

The processes behaved increasingly Hurst-like as their structure changed from 1 to 4. This behavior was most influenced by the distribution of epoch lengths, whereas the distribution of the mean itself had little effect.

The effect of periodical patterns, which are extensions of those of model 1, have been thoroughly studied by Montanari et al. (23), who, however, noted that such patterns are unusual in real data. The effect of monotonic deterministic trends, which are extensions of model 2, was studied by Bhattacharya et al. (24), who showed mathematically that a trend of the form $f(t) = c\,(m + t)^{H-1}$, where t denotes time, c a nonzero constant, m a positive constant, and H a constant in the interval [0.5, 1], results in time series exhibiting the Hurst phenomenon with a Hurst coefficient precisely equal to H. We may note, however, that this kind of nonstationarity with a monotonic deterministic trend spanning the whole length of a time series can hardly represent a long time series of real data, even though in short time series it seems to be realistic. For example, to refer to the Nilometer series of Fig. 1, if one had available only the data of the period 700–800, one would detect a "deterministic" falling trend of the Nile level; similarly, one would detect a regular rising trend of the Nile level between the years 1000–1100. However, the complete picture of the series suggests that these trends are parts of large-scale random fluctuations rather than deterministic trends.

Based on this observation, Koutsoyiannis (8) proposed a conceptual explanation, which can be regarded as an extension of Klemeš's model 4 and is also similar to other proposed conceptual models, as will be discussed later. More specifically, Koutsoyiannis (8) demonstrated that superimposition of three processes with short-term persistence results in a composite process that is practically indistinguishable from an SSS process.

This demonstration is reproduced here in Figs. 4–5. It starts by assuming a Markovian process U_i, like that graphically demonstrated in Fig. 4a, where mean $\mu := E[U_i]$, variance γ_0, and lag one autocorrelation coefficient $\rho = 0.20$. The specific form of this process is an AR(1) one, $U_i := \rho\, U_{i-1} + E_i$, where E_i is white noise, and its autocorrelation is

$$\mathrm{Corr}[U_i, U_{i+j}] = \rho^j \tag{12}$$

The autocorrelation function is shown in Fig. 5a along with the autocorrelation function of SSS with the same lag one autocorrelation coefficient (0.20). We observe the large difference of the two autocorrelation functions: that of the Markovian process practically vanishes at lag 4, whereas that of SSS has positive values for lags as high as 100.

In the next step, a second process V_i is constructed by superimposing another Markovian process M_i:

$$V_i = U_i + M_i - \mu \tag{13}$$

Here the process M_i is constructed in a different way, rather than using the AR(1) model, yet without loosing its Markovian behavior, so that its autocorrelation is

$$\mathrm{Corr}[M_i, M_{i+j}] = \varphi^j \tag{14}$$

for $\varphi > \rho$. More specifically, a continuous time process M (see explanatory sketch in Fig. 4b) with the following properties was assumed: (1) it has mean μ and some variance Var[M], (2) any realization m of M lasts N years and is independent of previous realizations, and (3) N is a random variable exponentially distributed with mean $\lambda = -1/\ln\,\varphi$. (This means that N can take nonintegral values). In other words, M takes a value $m_{(1)}$ that lasts n_1 years, then it changes to a value $m_{(2)}$ that lasts n_2 years, etc. (where the values $m_{(1)}$, $m_{(2)}$, ... can be generated from any distribution). The exponential distribution of N indicates that the points of change are random points in time. If we denote M_i the instance of the M process at discrete time i, it can be shown that M_i is Markovian with lag one autocorrelation φ. This way of constructing M_i allows us to interpret V_i as a process similar to U_i but with mean M_i that varies randomly in time (rather than being constant, μ) shifting among randomly determined values $m_{(1)}, m_{(2)}, \ldots$, each lasting a random time period with average λ. It can be easily shown from Eq. 13 that the autocorrelation of V_i for lag j is

$$\mathrm{Corr}[V_i, V_{i+j}] = (1 - c)\rho^j + c\,\varphi^j \tag{15}$$

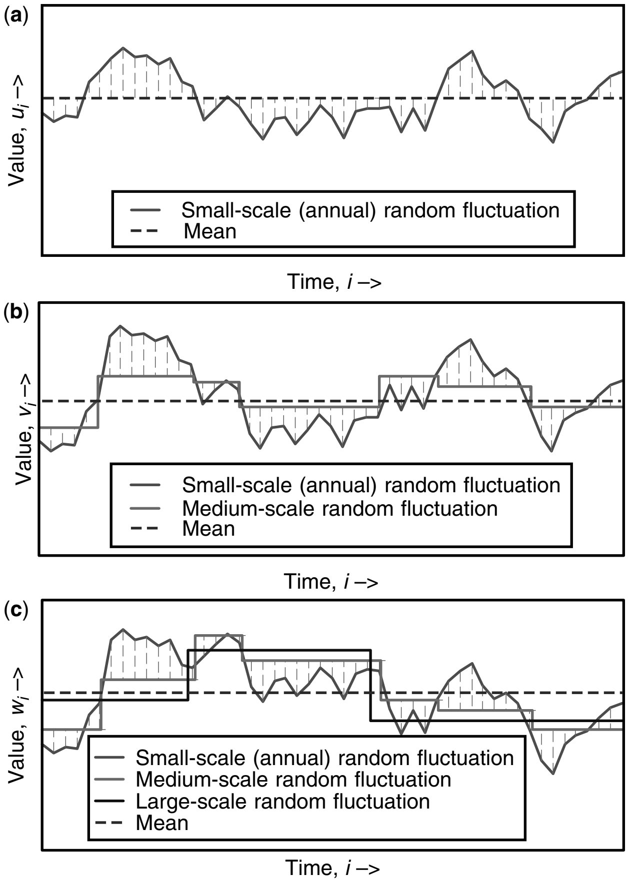

Figure 4. Illustrative sketch of multiple timescale random fluctuations of a process that can explain the Hurst phenomenon: (**a**) a time series from a Markovian process with constant mean; (**b**) the same time series superimposed on a randomly fluctuating mean on a medium timescale; (**c**) the same time series further superimposed on a randomly fluctuating mean on a large timescale (from Reference 8).

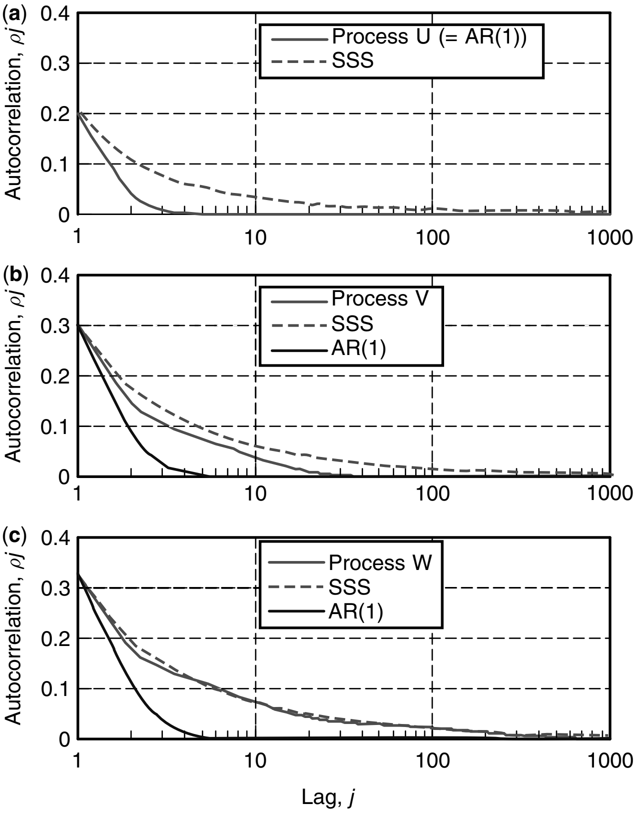

Figure 5. Plots of the example autocorrelation functions of (**a**) the Markovian process U with constant mean; (**b**) the process U superimposed on a randomly fluctuating mean on a medium timescale (process V); (**c**) the process V further superimposed on a randomly fluctuating mean on a large timescale (process W). The superimposition of fluctuating means increases the lag one autocorrelation (from $\rho_1 = 0.20$ for U to $\rho_1 = 0.30$ and 0.33 for V and W, respectively) and also shifts the autocorrelation function from the AR(1) shape (also plotted in all three panels) toward the SSS shape (also shown in all three panels) (from Reference 8).

where $c := \mathrm{Var}[M_i]/(\mathrm{Var}[M_i] + \mathrm{Var}[U_i])$. Setting, for instance, $\lambda = 7.5$ years ($\varphi = 0.875$) and $c = 0.146$, we get the autocorrelation function shown in Fig. 5b, which has departed from the AR(1) autocorrelation and approached the SSS autocorrelation.

In a third step, another process W_i is constructed by superimposing on V_i a third Markovian process P_i:

$$W_i = V_i + P_i - \mu = U_i + M_i + P_i - 2\,\mu \tag{16}$$

P_i is constructed in a way identical to that of M_i, but with lag one autocorrelation $\xi > \varphi$, so that the mean time between changes of the value of P is $\nu = -1/\ln \xi$. Working as in the previous step, we find

$$\mathrm{Corr}[W_i, W_{i+j}] = (1 - c_1 - c_2)\rho^j + c_1\varphi^j + c_2\xi^j \tag{17}$$

where c_1 and c_2 are positive constants (with $c_1 + c_2 < 1$). Setting, for instance, $\lambda = 7.5$ years ($\varphi = 0.875$), $\nu = 200$ years ($\xi = 0.995$), $c_1 = 0.146$, and $c_2 = 0.036$, we get the autocorrelation function shown in Fig. 5c, which has now become almost indistinguishable from the SSS autocorrelation for time lags from 1 to 1000.

In conclusion, a Markovian underlying process can result in a nearly SSS process if there occur random fluctuations of the mean of the process on two different scales (e.g., 7.5 and 200 years), yet the resulting composite process is stationary. If we consider that fluctuations occur on a greater number of timescales, the degree of approximation of the composite process to the SSS process will be even better and can cover time lags greater than 1000 (although the extension to lags beyond 1000 may not have any practical interest in hydrology). In conclusion, the *irregular* changes of climate that, according to the National Research Council (25, p. 21), occur on all timescales can be responsible for and explain the Hurst phenomenon.

This demonstration bridges several ideas that had been proposed to explain the Hurst phenomenon, rather than being a novel explanation. As already discussed, it is similar to Klemeš's model 4, except for the setting of multiple timescales for the fluctuation of the mean and the emphasis on the stationarity of the composite process. Here, note that Klemeš referred to all his "changing mean" models as models with nonstationarity in their mean, even though this is strictly true only for models 1 and 2. He did point out that his final models in group 4 were in fact stationary and that he kept the term "nonstationary" for all changes in the mean to communicate the fact (elaborated in more detail in Ref. 26) that one cannot tell the difference from the pattern of a single "nonstationary-looking" time series (which even a stationary model is designed to mimic), but his explanation has sometimes been missed and led to a misconception about his work by some authors (including this one, who expresses his apology).

The idea of irregular sporadic changes in the mean of the process appeared also in Salas and Boes (27) but not in connection with SSS and not in the setting of multiple timescales. The idea of composite random processes with two timescales of fluctuation appeared in Vanmarcke (28, p. 225). The idea of an explanation of the Hurst phenomenon as a mixture of scales appears in Mesa and Poveda (19). The idea of representing SSS as an aggregation of short-memory processes is the principle of the well known fast fractional Gaussian noise algorithm (FFGN, 29) and is also studied as a possible physical explanation of the Hurst phenomenon by Beran (17, p. 14). The difference in the above described explanation is the aggregation of only three short-memory processes.

IMPORTANCE OF THE HURST PHENOMENON

The presence of the Hurst phenomenon increases dramatically the uncertainty of climatic and hydrologic processes. If such a process were random and our information on this were based on a sample of size n, then the uncertainty in the long term, which can be expressed in terms of the variance of the estimator of the mean, $\overline{X}$, would be

$$\mathrm{var}[\overline{X}] = \frac{\sigma^2}{n} \tag{18}$$

This offers good approximation for a process with short-term persistence, as well, but it is not valid for a process with long-term persistence. Instead, the following relation holds (5;17, p. 54; 30):

$$\mathrm{var}[\overline{X}] = \frac{\sigma^2}{n^{2-2H}} \tag{19}$$

The difference between Eqs. 18 and 19 becomes very significant for large values of H. For example, in a time series of $n = 100$ years of observations and standard deviation σ, according to the classical statistics (Eq. 18), the standard estimation error, the square root of $\mathrm{var}[\overline{X}]$, is $\sigma/10$. However, for $H = 0.8$, the correct standard error, given by Eq. 19, is $\sigma/2.5$, four times larger. To have an estimation error equal to $\sigma/10$, the required length of the time series would be 100,000 years! Obviously, this dramatic difference induces substantial differences in other common statistics as well (5).

A demonstration of the difference in estimates related to climate is given in Fig. 6. Here, a long climatic time series (992 years) was used that represents the Northern Hemisphere temperature anomalies with reference to the 1961–1990 mean (Fig. 8, up). This series was constructed using temperature sensitive palaeoclimatic multiproxy data from 10 sites worldwide that include tree rings, ice cores, corals, and historical documents (31,32). The time series was studied in relation to the Hurst phenomenon by Koutsoyiannis (5), and it was found that the estimate of the Hurst coefficient is 0.88. In the upper panel of Fig. 6, the point estimates and the 99% confidence limits of the quantiles of the temperature anomalies have been plotted for the probability of nonexceedance, u, ranging from 1 to 99%, assuming a normal distribution, as verified from the time series, and using the classical statistical estimators. This is done for two timescales, the basic one ($k = 1$) that represents the annual variation of temperature anomaly and the 30-year timescale, which typically is assumed to be sufficient to smooth out the annual variations and provide values representative of the climate. (For the latter, the averaged rather than aggregated time series, i.e., $z_i^{(30)}/30$, has been used.)

If classical statistics is used (Fig. 6, upper panel), then it is observed that, due to the large length of the series, the confidence band is very narrow and the point estimates for the basic and the aggregated timescale differ significantly. The variability of climate, as expressed by the distribution of the average at the 30-year timescale, is very low, despite much higher variability on the annual scale. This justifies the saying "Climate is what you expect, weather is what you get." Things change dramatically, if the statistics based on the hypothesis of long-term persistence (5) are used with $H = 0.88$. This is depicted in the lower panel of Fig. 6, where it is observed that the variation in the 30-year average is only slightly lower than that of the annual values and the confidence band has dramatically widened for both timescales. This could be expressed by paraphrasing the above proverb to read "Weather is what you get, climate is what you get—if you keep expecting for many years." The consequences of these differences in estimating climatic uncertainty due to natural variability are obviously very significant.

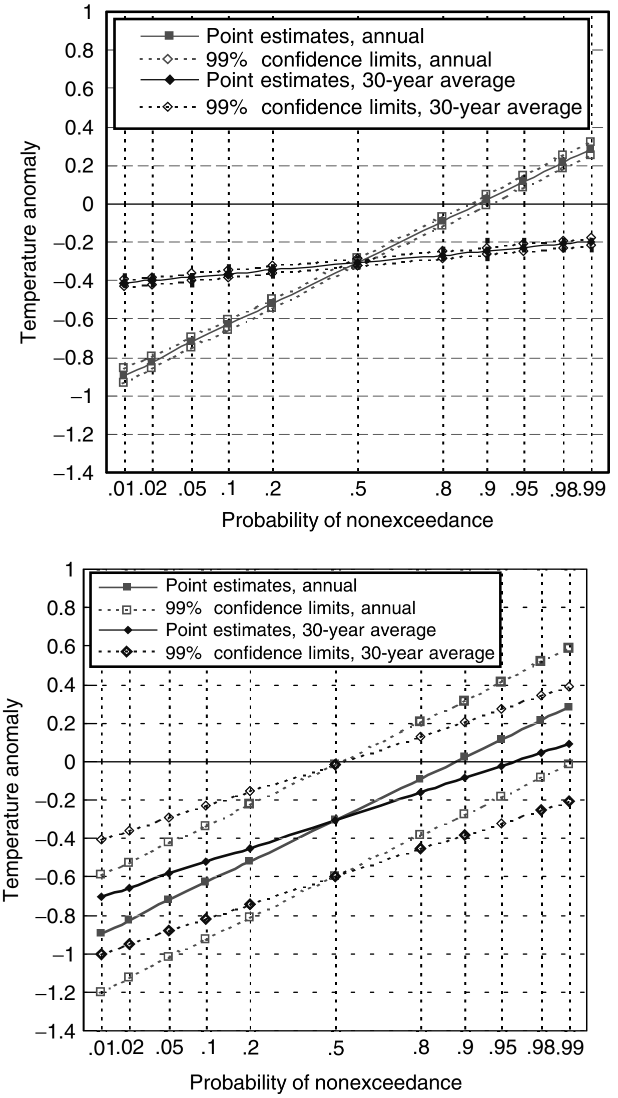

Figure 6. Point estimates of quantiles and 99% confidence limits thereof on the basic timescale (annual values, $k = 1$) and the 30-year timescale (30-year averages, $k = 30$), for the Jones's time series of the Northern Hemisphere temperature anomalies: (up) using classical statistics; (down) using adapted statistics.

The consequences in water resources engineering and management are even more significant. Particularly, because the notion of climate implies long timescales, it is to be expected that the practical importance of the Hurst phenomenon increases in projects whose operating cycles span long periods of time. A typical example may be large reservoirs with multiyear flow regulation (33; see also article SW-776, Reliability Concepts in Reservoir Design). For small-to-middle range reservoirs, it is generally regarded that the effect of the Hurst phenomenon appears to be within the margin of error of hydrologic data used for their design and operation. However, even in hydrosystems with small reservoirs or no reservoirs at all, as becomes obvious from the previous discussion, the effect on the Hurst phenomenon is significant if the uncertainty (not only the expected value) of water availability is to be assessed.

SIMPLE ALGORITHMS TO GENERATE TIME SERIES REPRODUCING THE HURST PHENOMENON

Several algorithms have been developed to generate time series that reproduce the Hurst phenomenon. Among these, we discuss here the simplest ones that can be applied, even in a spreadsheet. These are based on the previously discussed properties of SSS and can be used to provide good approximations of SSS for practical hydrologic purposes.

A first, rather "quick and dirty" algorithm can be very easily formulated based on the deterministic double tent map (Eqs. 10 and 11). The problem with the resulting time series is that consecutive generated values are too regularly and smoothly related. This can be avoided by discarding some of the generated values x_i and holding only the values $x_{\nu j}$, for some $\nu > 1$ and for $j = 1, 2, \ldots, n$, where n is the required series length. Figure 7 depicts the attained Hurst coefficient in a time series generated from the double tent map (Eq. 10) either untransformed or transformed (Eq. 11) for $\nu = 1$ and 4. This figure can serve as a tool to estimate the parameter κ required to achieve a certain Hurst coefficient H (assuming $\lambda = 0.001$). A time series so generated can then be transformed linearly to acquire the required mean and standard deviation.

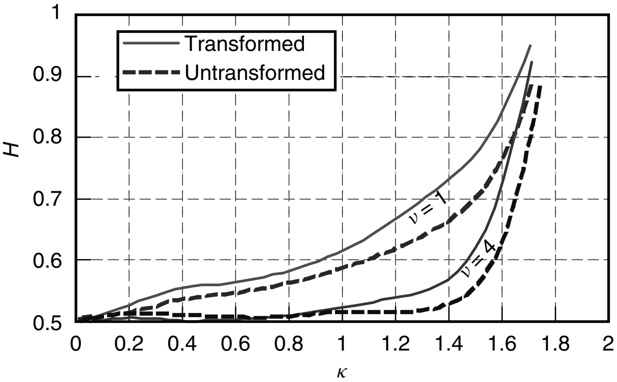

Figure 7. Hurst coefficient of a time series generated from the double tent map (Eqs. 10 and 11) for parameter values $\lambda = 0.001$, κ ranging from 0 to 2, and $\nu = 1$ and 4. For κ approaching 2, the double tent map has runaway behavior.

By appropriately choosing the initial values α_0 and u_0, one can obtain a time series that can have a presumed general shape; this requires applying a random search optimization technique. An example of the application of this algorithm to the Jones data set already discussed is depicted in Fig. 8.

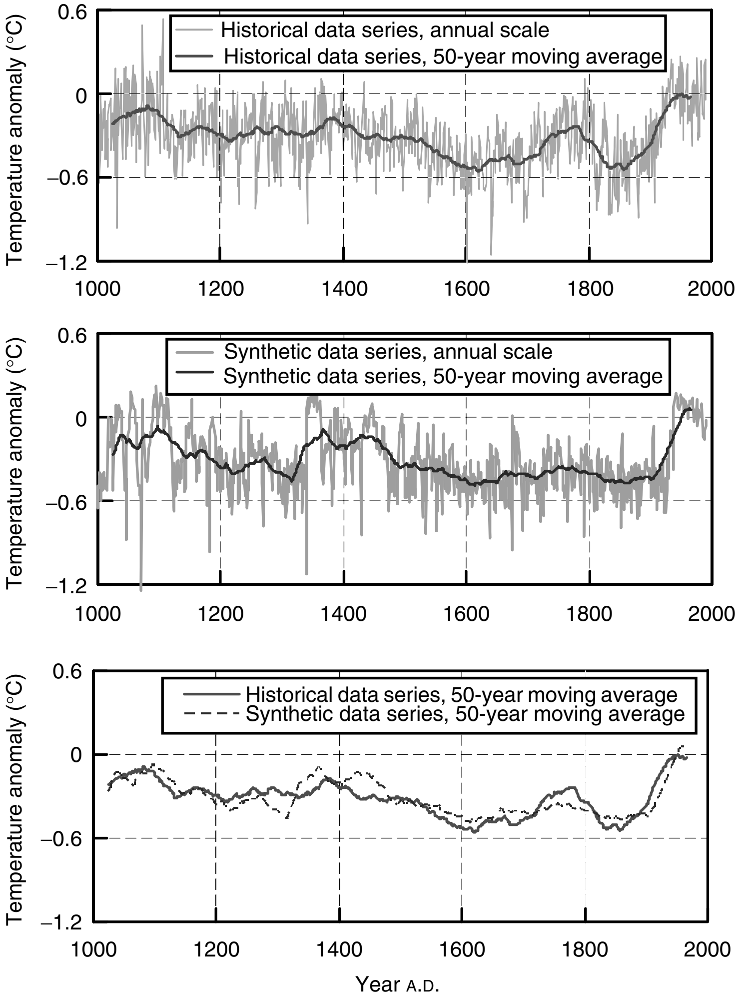

Figure 8. (Up) Plot of the Jones data series indicating Northern Hemisphere temperature anomalies with reference to the 1961–1990 mean; (middle) a synthetic time series generated by the double tent map fitted to the Jones data set and assuming $\nu = 4$; (down) comparison of the synthetic and original time series in terms of their 50-year moving averages.

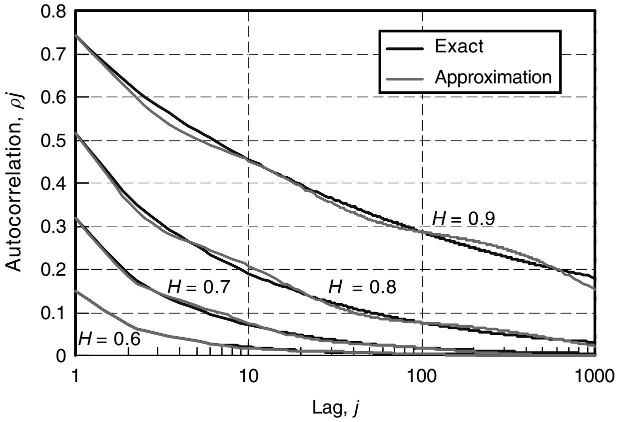

Figure 9. Approximate autocorrelation functions based on Equations 17 and 20–21 vs. the exact SSS autocorrelation functions (Eq. 5) for various values of the Hurst exponent H (from Reference 8).

As we saw earlier, the weighted sum of three exponential functions of the time lag (Eq. 17) can give an acceptable approximation of the SSS autocorrelation function on a basic timescale. This observation leads to an easy algorithm to generate SSS. The following equations (from Reference 8) can be used to estimate the parameters ρ, φ and ξ:

$$\rho = 1.52\ (H - 0.5)^{1.32}$$
$$\varphi = 0.953 - 7.69\ (1 - H)^{3.85} \tag{20}$$

$$\xi = \begin{cases} 0.932 + 0.087\,H & H \leq 0.76 \\ 0.993 + 0.007\,H & H > 0.76 \end{cases} \tag{21}$$

The remaining parameters c_1 and c_2 can then be estimated so that the approximate autocorrelation function (Eq. 17) matches the exact function (Eq. 5) for two lags, for example, lags 1 and 100. (Their values are obtained by solving two linear equations). Comparison plots of approximate autocorrelation functions based on Equations 17 and 20–21 versus the exact SSS autocorrelation functions (Eq. 5) for various values of the Hurst exponent H are shown in Fig. 9. Equations 16 and 17 may be interpreted as representing the sum of three independent AR(1) processes, with lag one correlation coefficients ρ, ϕ, and ξ, and variances $(1 - c_1 - c_2)\ \gamma_0$, $c_1\gamma_0$, and $c_2\gamma_0$, respectively. Thus, the generation algorithm is as simple as the generation of three AR(1) series and their addition.

The simple expressions of the statistics of the aggregated SSS process make possible a disaggregation approach for generating SSS (8). Specifically, let us assume that the desired length n of the synthetic series to be generated is 2^m where m is an integer (e.g., $n = 2, 4, 8, \ldots,$); if not, we can increase n to the next power of 2 and then discard the redundant generated items. We first generate the single value of $Z_1^{(n)}$ knowing its variance $n^{2H}\gamma_0$ (from Eq. 3). Then we disaggregate $Z_1^{(n)}$ into two variables on the timescale $n/2$, $Z_1^{(n/2)}$ and $Z_2^{(n/2)}$, and we proceed this way until the series $Z_1^{(1)} \equiv X_1, \ldots, Z_n^{(1)} \equiv X_n$ is generated (see explanatory sketch in Fig. 10).

We consider the generation step in which we disaggregate the higher level amount $Z_i^{(k)} (1 < i < n/k)$ into two lower level amounts $Z_{2\,i-1}^{(k/2)}$ and $Z_{2\,i}^{(k/2)}$ so that

$$Z_{2\,i-1}^{(k/2)} + Z_{2\,i}^{(k/2)} = Z_i^{(k)} \tag{22}$$

Thus, it suffices to generate $Z_{2\,i-1}^{(k/2)}$ and then obtain $Z_{2\,i}^{(k/2)}$ from Eq. 22. At this generation step, we have available the already generated values of previous lower level time steps, $Z_1^{(k/2)}, \ldots, Z_{2\,i-2}^{(k/2)}$, and of next higher-level time steps, $Z_{i+1}^{(k)}, \ldots, Z_{n/k}^{(k)}$ (see Fig. 10). Theoretically, it is necessary to preserve the correlations of $Z_{2\,i-1}^{(k/2)}$ with all previous lower level variables and all next higher level variables. However, we can get a very good approximation if we consider correlations with only one higher level time step behind and one ahead. Under this simplification, $Z_{2\,i-1}^{(k/2)}$ can be generated from the linear relationship,

$$Z_{2\,i-1}^{(k/2)} = a_2\, Z_{2\,i-3}^{(k/2)} + a_1\, Z_{2\,i-2}^{(k/2)} + b_0\, Z_i^{(k)} + b_1\, Z_{i+1}^{(k)} + V \tag{23}$$

where a_2, a_1, b_0 and b_1 are parameters given by

$$\begin{bmatrix} a_2 \\ a_1 \\ b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 + \rho_3 & \rho_4 + \rho_5 \\ \rho_1 & 1 & \rho_1 + \rho_2 & \rho_3 + \rho 4 \\ \rho_2 + \rho_3 & \rho_1 + \rho_2 & 2(1 + \rho_1) & \rho_1 + 2\rho_2 + \rho_3 \\ \rho_4 + \rho_5 & \rho_3 + \rho_4 & \rho_1 + 2\rho_2 + \rho_3 & 2(1 + \rho_1) \end{bmatrix}^{-1} \bullet \begin{bmatrix} \rho_2 \\ \rho_1 \\ 1 + \rho_1 \\ \rho_2 + \rho_3 \end{bmatrix} \tag{24}$$

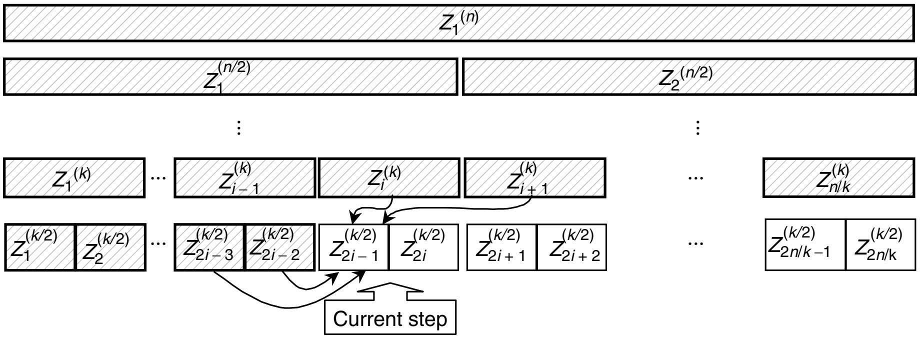

Figure 10. Explanatory sketch of the disaggregation approach for generating SSS. Gray boxes indicate random variables whose values have been already generated prior to the current step, and arrows indicate the links to those of the generated variables that are considered in the current generation step (from Reference 8).

with ρ_j given by Eq. 5 and V is an innovation with variance,

$$\mathrm{Var}[V] = \gamma_0^{(k/2)}(1 - [\rho_2, \rho_1, 1+\rho_1, \rho_2+\rho_3][a_2, a_1, b_0, b_1]^T) \tag{25}$$

where the superscript T denotes the transpose of a vector.

All parameters are independent of i and k, and therefore they can be used in all steps. When $i = 1$, there are no previous time steps, and thus the first two rows and columns of the above matrix and vectors are eliminated. Similarly, when $i = n/k$, there is no next time step, and thus the last row and column of the above matrix and vectors are eliminated.

The power law of the power spectrum of SSS allows generating an SSS time series X_i by filtering a series of white noise V_i using the symmetrical moving average (SMA) scheme (34):

$$X_i = \sum_{j=-q}^{q} a_{|j|}\, V_{i+j} = a_q\, V_{i-q} + \cdots + a_1\, V_{i-1} + a_0\, V_i + a_1\, V_{i+1} + \cdots + a_q\, V_{i+q} \tag{26}$$

where q theoretically is infinity but in practice can be restricted to a finite number, as the sequence of weights a_j tends to zero for increasing j. Koutsoyiannis (8) showed that the appropriate sequence of a_j is

$$a_j \approx \frac{\sqrt{(2-2H)\gamma_0}}{3-2H}(|j+1|^{H+0.5} + |j-1|^{H+0.5} - 2|j|^{H+0.5}) \tag{27}$$

The sequence length q must be chosen at least equal to the desired number of autocorrelation coefficients m that are to be preserved. In addition, the ignored terms a_j beyond a_q must not exceed an acceptable tolerance $\beta\sigma$. These two conditions result in

$$q \geq \max\left[m, \left(\frac{2\beta}{H^2 - 0.25}\right)^{1/(H-1.5)}\right] \tag{28}$$

Thus, q can be very large (on the order of thousands to hundreds of thousands) if H is large (e.g., > 0.9) and β is small (e.g., < 0.001). Approximate autocorrelation functions based on Eqs. 26 and 27 versus the exact SSS autocorrelation functions (Eq. 5) for various values of H and q are shown in Fig. 11.

This method can also generate non-Gaussian series with skewness ξ_X by appropriately choosing the skewness of the white noise ξ_V. The relevant equations for the statistics of V_i, which are direct consequences of Eq. 26, are

$$\left(a_0 + 2\sum_{j=1}^{s} a_j\right) E[V_i] = \mu,$$

$$\mathrm{Var}[V_i] = 1,$$

$$\left(a_0^3 + 2\sum_{j=1}^{q} a_j^3\right) \xi_V = \xi_X \gamma_0^{3/2} \tag{29}$$

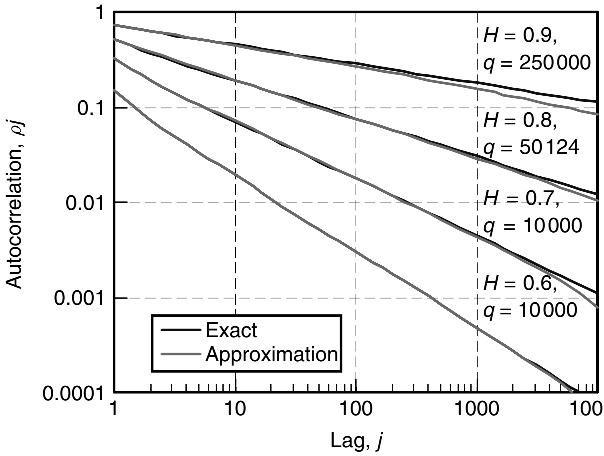

Figure 11. Approximate autocorrelation functions based on Equations 26 and 27 vs. the exact SSS autocorrelation functions (Eq. 5) for various values of the Hurst exponent H and of the number of weights q (from Reference 8).

CONCLUDING REMARKS

More than half a century after its discovery, the Hurst phenomenon has been verified as almost omnipresent in several processes in nature (e.g., hydrology), technology (e.g., computer networks), and society (e.g., economics). However, its consequences are still not widely understood or are ignored; to quote Klemeš (20), it is still regarded by many as “a ghost to be conjured away.”

For example, in stochastic hydrologic simulations that are used in hydrosystem modeling, the Hurst phenomenon is generally not reproduced. The most widespread stochastic hydrology packages have not implemented any types of models that reproduce the Hurst phenomenon. However, today methodologies exist, implemented in software packages, which can reproduce the Hurst phenomenon even in complicated situations, such as a multivariate setting with multiple timescales and asymmetric probability distributions (15,34,35). In addition, as described before, reproducing the Hurst phenomenon in univariate problems is quite simple.

In hydrologic analysis, it has been a common practice to detect falling or rising monotonic “trends” in available records, assume that these are deterministic components, and then “subtract” them from the time series to obtain a “detrended” time series, which is finally used in subsequent analyses. This common technique, which is described in several hydrologic texts, obviously contradicts the Hurst phenomenon. The “trends” are large scale fluctuations, the basis of the Hurst phenomenon. They could be regarded as deterministic components if a sound, physically based model could capture them and also predict their evolution in the future. This, however, is not the case. The a posteriori fitting of a regression curve (e.g., a linear equation) on historical data series has no relation to deterministic modeling. The subtraction of the “trends” from the time series results in a reduction of the standard deviation, an artificial decrease of uncertainty. This is exactly opposite to the real meaning of the

Hurst phenomenon, which, as analyzed before, increases uncertainty substantially.

Even without adopting this “detrending” technique, hydrologic statistics, the branch of hydrology that deals with uncertainty, in its current state is not consistent with the Hurst phenomenon. Typical statistics used in hydrology such as means, variances, cross- and autocorrelations and Hurst coefficients, and the variability thereof, are based on classical statistical theory, which describe only a portion of natural variability, and thus its results may underestimate dramatically the natural uncertainty and the implied risk.

The situation is even worse in climatology, which again uses the classical statistical framework but on longer timescales (e.g., 30 years). As demonstrated before, the consequences of the Hurst phenomenon for natural variability increase as the timescale increases. Recently, many researchers are involved in detecting anthropogenic climatic changes mostly using classical statistical tests, without taking into account the Hurst phenomenon. If statistical estimators consistent to the Hurst phenomenon are used, a choice more consistent with nature, it is more unlikely that such tests will result in statistically significant changes.

Acknowledgments
The enlightening general discussions of Vit Klemeš about the nature of the Hurst phenomenon and related issues, as well his detailed suggestions and comments on this particular article are gratefully appreciated.

BIBLIOGRAPHY

1. Hurst, H.E. (1951). Long term storage capacities of reservoirs. *Trans. ASCE* **116**: 776–808.
2. Mandelbrot, B.B. (1977). *The Fractal Geometry of Nature*. Freeman, New York.
3. Haslett, J. and Raftery, A.E. (1989). Space–time modelling with long-memory dependence: Assessing Ireland’s wind power resource. *Appl. Statist.* **38**(1): 1–50.
4. Bloomfield, P. (1992). Trends in global temperature. *Climatic Change* **21**: 1–16.
5. Koutsoyiannis, D. (2003). Climate change, the Hurst phenomenon, and hydrological statistics. *Hydrol. Sci. J.* **48**(1): 3–24.
6. Koutsoyiannis, D. (2003). *A toy model of climatic variability with scaling behaviour*. *Hydrofractals ’03, Int. Conf. Fractals Hydrosciences*, August 2003, Monte Verita, Ascona, Switzerland, ETH Zurich, MIT, Université Pierre et Marie Curie. Available: http://www.itia.ntua.gr/g/docinfo/585/.
7. Eltahir, E.A.B. (1996). El Niño and the natural variability in the flow of the Nile River. *Water Resour. Res.* **32**(1): 131–137.
8. Koutsoyiannis, D. (2002). The Hurst phenomenon and fractional Gaussian noise made easy. *Hydrol. Sci. J.* **47**(4): 573–596.
9. Radziejewski, M. and Kundzewicz, Z.W. (1997). Fractal analysis of flow of the river Warta. *J. Hydrol.* **200**: 280–294.
10. Sakalauskienė, G. (2003). The Hurst phenomenon in hydrology. *Environ. Res. Eng. Manage.* **3**(25): 16–20.
11. Montanari, A., Rosso, R., and Taqqu, M.S. (1997). Fractionally differenced ARIMA models applied to hydrologic time series. *Water Resour. Res.* **33**(5): 1035–1044.
12. Stephenson, D.B., Pavan, V., and Bojariu, R. (2000). Is the North Atlantic oscillation a random walk? *Int. J. Climate* **20**: 1–18.
13. Evans, T.E. (1996). *The effects of changes in the world hydrological cycle on availability of water resources*. In: *Global Climate Change and Agricultural Production: Direct and Indirect Effects of Changing Hydrological, Pedological and Plant Physiological Processes*. F. Bazzaz and W. Sombroek (Eds.). FAO and John Wiley, Chichester, West Sussex, UK, Chap. 2.
14. Koutsoyiannis, D. (2003). *Hydrological statistics for engineering design in a varying climate*. *EGS-AGU-EUG Joint Assembly, Geophysical Research Abstracts, Vol. 5*, Nice, April 2003, European Geophysical Society, American Geophysical Union. Available: http://www.itia.ntua.gr/g/docinfo/565/.
15. Koutsoyiannis, D. and Efstratiadis, A. (2004). *Climate change certainty versus climate uncertainty and inferences in hydrological studies and water resources management*. *1st General Assembly of the European Geosciences Union, Geophysical Research Abstracts, Vol. 6*, Nice. Available: http://www.itia.ntua.gr/g/docinfo/606/.
16. Toussoun, O. (1925). Mémoire sur l’histoire du Nil. In: *Mémoires a l’Institut d’Egypte*. vol. 18, pp. 366–404.
17. Beran, J. (1994). *Statistics for Long-Memory Processes*, *vol. 61 of Monographs on Statistics and Applied Probability*. Chapman and Hall, New York.
18. Mandelbrot, B.B. (1965). Une classe de processus stochastiques homothétiques a soi: Application à la loi climatologique de H. E. Hurst. *C. R. Acad. Sci. Paris* **260**: 3284–3277.
19. Mesa, O.J. and Poveda, G. (1993). The Hurst effect: the scale of fluctuation approach. *Water Resour. Res.* **29**(12): 3995–4002.
20. Klemeš, V. (1974). The Hurst phenomenon: A puzzle? *Water Resour. Res.* **10**(4): 675–688.
21. Klemeš, V. (1978). Physically based stochastic hydrologic analysis. *Adv. Hydroscience* **11**: 285–356.
22. Lasota, A. and Mackey, M.C. (1994). *Chaos, Fractals and Noise, Stochastic Aspects of Dynamics*. Springer-Verlag, Berlin.
23. Montanari, A., Taqqu, M.S., and Teverovsky, V. (1999). Estimating long-range dependence in the presence of periodicity: an empirical study. *Math. Comput. Modeling* **29**: 217–228.
24. Bhattacharya, R.N., Gupta, V.K. and Waymire, E. (1983). The Hurst effect under trends. *J. Appl. Probab.* **20**: 649–662.
25. National Research Council (Committee on Opportunities in the Hydrologic Sciences). (1991). *Opportunities in the Hydrologic Sciences*. National Academy Press, Washington, DC.
26. Klemeš, V. (1976). Geophysical time series and catastrophism. *Catastrophist Geol.* **1**(1): 43–48.
27. Salas, J.D. and Boes, D.C. (1980). Shifting level modelling of hydrologic time series. *Adv. Water Resour.* **3**: 59–63.
28. Vanmarcke, E. (1983). *Random Fields*. MIT Press, Cambridge, MA.
29. Mandelbrot, B.B. (1971). A fast fractional Gaussian noise generator. *Water Resour. Res.* **7**(3): 543–553.
30. Adenstedt, R.K. (1974). On large sample estimation for the mean of a stationary random sequence. *Ann. Statist.* **2**: 1095–1107.
31. Jones, P.D., Briffa, K.R., Barnett, T.P., and Tett, S.F.B. (1998). *Millennial temperature reconstructions. IGBP PAGES / World Data Center-A for Paleoclimatology Data Contribution Series #1998-039*, NOAA/NGDC Paleoclimatology

Program, Boulder, CO. Available: ftp.ngdc.noaa.gov/paleo/contributions_by_author/jones1998/.

32. Jones, P.D., Briffa, K.R., Barnett, T.P., and Tett, S.F.B. (1998). High-resolution paleoclimatic records for the last millennium: Interpretation, integration and comparison with General Circulation Model control-run temperatures. *Holocene* **8**(4): 455–471.
33. Klemeš, V., Sricanthan, R., and McMahon, T.A. (1981). Long-memory flow models in reservoir analysis: What is their practical value? *Water Resour. Res.* **17**(3): 737–751.
34. Koutsoyiannis, D. (2000). A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. *Water Resour. Res.* **36**(6): 1519–1533.
35. Langousis, A. and Koutsoyiannis, D. (2003). *A stochastic methodology for generation of seasonal time series reproducing overyear scaling. Hydrofractals '03, Int. Conf. Fractals Hydrosciences*, August 2003, Monte Verita, Ascona, Switzerland, ETH Zurich, MIT, Université Pierre et Marie Curie. Available: http://www.itia.ntua.gr/g/docinfo/586/.

UNIT HYDROGRAPH

Christophe Cudennec
Ecole Nationale Supérieure
Agronomique
Rennes, France

On the scale of a hydrologic event, a river basin schematically has two influences. The first is the yield of rainfall transformed into effective rainfall (in terms of runoff production), and the second is the transfer of runoff flows generated all over the basin territory to the unique outlet. The concept assumes that all transfer processes of the whole basin can be synthesized in a single linear function, called the unit hydrograph of the basin considered (1). It is at the root of the idea of the transfer function and is a kind of functional signature of the basin (2). This means that a characteristic function can be identified for a given basin. It is assumed that this function underlies any rainfall–runoff event and synthesizes all relevant hydrologic processes and geographic influences on the temporal scale of the event and the spatial scale fixed by the basin territory. Furthermore this means that the unit hydrograph should really be the hydrograph at the outlet, following an average impulse unit effective rainfall input, and that the response of the basin to any actual effective rainfall chronicle can be simulated through the sum of successive elementary affinities of the unit hydrograph. Such a holistic purpose necessarily faces problems of simplifying the natural complexity and of identifying the unit hydrograph itself.

The mathematical translation of the unit hydrograph concept is based on the following assumptions and analytical proposals:

- One considers a 1-unit height of effective rainfall generated uniformly over the basin territory during a time-unit duration.
- The response at the outlet of the basin is the characteristic unit hydrograph $u(t)$ during time t.
- This analytical function is used as a linear operator, which means that it respects the affinity and additivity properties. The response to any height of uniform unit effective rainfall is proportional to the unit hydrograph. The response to two following unit effective rainfalls is the sum of two following unit hydrographs. These two properties allow calculating the discharge chronicle $Q(t)$ generated by a chronicle of effective rainfall $\overline{R_{\mathrm{eff}}}(t)$:

$$Q(t) = \frac{S}{\Delta t} \cdot \sum_{\tau=1}^{\min(t,\theta)} \overline{R_{\mathrm{eff}}}(t-\tau+1)u(\tau) \qquad (1)$$

where Δt: the discrete time pace;
$Q(t)$: the mean discharge during the time step $[t-\Delta t; t]$;
$\overline{R_{\mathrm{eff}}}(t)$: the height of effective rainfall, average in space, during the time step $[t-\Delta t; t]$;
$u(t)$: the value of the unit hydrograph during the time step $[t-\Delta t; t]$;
S: the basin surface area;
θ: the length of the unit hydrograph;
τ: the time abscissa of the unit hydrograph.

If one considers Δt close to zero, the effective rainfall input is similar to a Dirac function, and the consequent unit hydrograph is then called an instantaneous unit hydrograph. Equation 1 can than be expressed continuously as a convolution:

$$Q(t) = S\overline{i_{\mathrm{eff}}}(t) * u(t) \qquad (2)$$

$$\Leftrightarrow Q(t) = S\int_0^t \overline{i_{\mathrm{eff}}}(t-\tau)u(\tau)d\tau \qquad (3)$$

where $\overline{i_{\mathrm{eff}}}(t)$ is the intensity of the effective rainfall, average in space, at time t. $u(t)$ is then the kernel function of the convolution.

The initial unit hydrograph concept strongly simplifies the complexity of basin hydrology:

- Its globality does not easily allow accounting for space heterogeneities and variabilities.
- It applies only to rapid components of runoff.
- Its linearity and stationarity can be criticized because actual hydrologic events are nonlinear.
- The claim of describing all transfer processes within a basin is strong regarding the differences between hillslope and channel processes.

But these problems have been studied through different proposals for identifying it.

The analytical expression for the unit hydrograph can be deduced from a comparison of actual rainfall and discharge gauging (3,4). It can also be deduced from conceptual assumptions, based mostly on the concepts of linear reservoir and linear channel (3). The most famous conceptual proposal is Nash's model (5,6), based on an analogy between a basin and n identical linear reservoirs in series. The unit hydrograph obtained is the following gamma law:

$$u(t) = \frac{1}{k(n-1)!}\left(\frac{t}{k}\right)^{n-1} e^{-\frac{t}{k}} \quad (4)$$

where the parameters n and k have to be calibrated, but where it is a priori assumed that the gamma law is the shape of the unit hydrograph.

To show the influence of basin geomorphology on unit hydrograph identification, a theory has been proposed by Rodriguez-Iturbe and Valdès (7). They define the unit hydrograph as the probability density function of the water travel time to the outlet. They estimate it by combining the probabilities of paths through the Strahler states and the probabilities of residence times in the different Strahler states of the river network. The result is called the geomorphological unit hydrograph.

Parallel to this, some models are based on the isochrone notion. It is already the case for the rational method used since the nineteenth century (8) and is coming back in models based on the geomorphological area and width functions because they are easily available from geographic databases and GIS processing (9,10). This geomorphological basement of the unit hydrograph identification opens, today, new prospects because it allows splitting transfers through hillslopes and through a river channel (11), considering nonlinear transfer processes through a stable geometric structure (8), and accounting for variability of effective rainfall input along the isochrone areas (12).

BIBLIOGRAPHY

1. Sherman, L.K. (1932). Streamflow from Rainfall by the Unit Hydrograph Method. *English News Record* **108**: 501–505.
2. Beven, K. (2001). *Rainfall-Runoff Modelling, The Primer*. John Wiley, Chichester, UK, p. 360.
3. Chow, V.T. (Ed.). (1964). *Handbook of Applied Hydrology*. McGraw Hill, New York.
4. Duband, D., Obled, Ch., and Rodriguez, J.Y. (1993). Unit hydrograph revisited: an alternate iterative approach to UH and effective precipitation identification. *J. Hydrol.* **150**: 115–149.
5. Nash, J.E. (1957). *The Form of the Instantaneous Unit Hydrograph*. C.R. et rapports, AIHS., IUGG., Toronto, Sept., 3–14 pp. 114–118.
6. Nash, J.E. (1959). Systematic determination of unit hydrograph parameters. *J. Geophys. Res.* **64**: 111–114.
7. Rodriguez-Iturbe, I. and Valdès, J.B. (1979). The geomorphologic structure of hydrologic response. *Water Resour. Res.* **15**: 1409–1420.
8. Dooge, J.C.I. (1959). A general theory of the unit hydrograph. *J. Geophys. Res.* **64**: 241–256.
9. Beven, K. and Kirkby, M.J. (Eds.). (1993). *Channel Network Hydrology*. John Wiley, Chichester, UK, p. 319.
10. Rodriguez-Iturbe, I. and Rinaldo, A. (1997). *Fractal River Basins; Chance and Self-Organization*. Cambridge University Press, Cambridge, p. 547.
11. Robinson, J.S., Sivapalan, M., and Snell, J.D. (1995). On the relative roles of hillslope processes, channel routing, and network geomorphology in the hydrologic response of natural catchments. *Water Resour. Res.* **31**: 3089–3101.
12. Cudennec, C., Gogien, F., Bourges, J., Duchesne, J., and Kallel, R. (2002). *Relative Roles of Geomorphology and Water Input Distribution in an Extreme Flood Structure*. IAHS Red books, to be published (March).

HYDROLOGICAL PROCESSES AND MEASURED POLLUTANT LOADS

Ramakar Jha
K.D. Sharma
National Institute of Hydrology
Roorkee, Uttranchal, India

Vijay P. Singh
Louisiana State University
Baton Rouge, Louisiana

HYDROLOGICAL PROCESSES

Hydrological processes involved in the circulation of water between the atmosphere, the land surface, and underground and its return back to the atmosphere is know as the hydrologic cycle (Fig. 1). In the circulation process, water vapors are transported by winds and air currents through the atmosphere (1). When the air mass cools sufficiently, the water vapor condenses into clouds, and a portion falls to the ground as precipitation in the form of snow, rain, sleet, or hail. Water that falls to the ground as precipitation follows many paths on its way back to the atmosphere. The water may be intercepted and taken up by the plants; it may be stored in small depressions or lakes; it can infiltrate the soil; or it can flow over the surface to a nearby stream channel. The sun may cause the water to evaporate directly back into the atmosphere, or the force of gravity may pull it down through the pores of the soil to be stored for years as slowly moving groundwater. Some of the water flowing through the ground returns to the surface to supply water to springs, lakes, and rivers. Water on the ground surface, in streams, or in lakes can return to the atmosphere as vapor through the process of evaporation. Water used by plants may return to the atmosphere as vapor through transpiration, which occurs when water passes through the leaves of plants. Collectively known as evapotranspiration, both evaporation and transpiration occur in greatest amounts during periods of high temperatures and wind, dry air, and sunshine.

This movement of water supports life on the earth and is mainly governed by the energy of the sun, the force of gravity, land use/land cover pattern, soil type, and geological characteristics. The quantities of water in the atmosphere, soils, groundwater, surface water, and other components are constantly changing because of the dynamic nature of the hydrological processes. The magnitude of various storage components including soil water, snow packs, lakes, reservoirs, and rivers can be altered by human activities. With the water

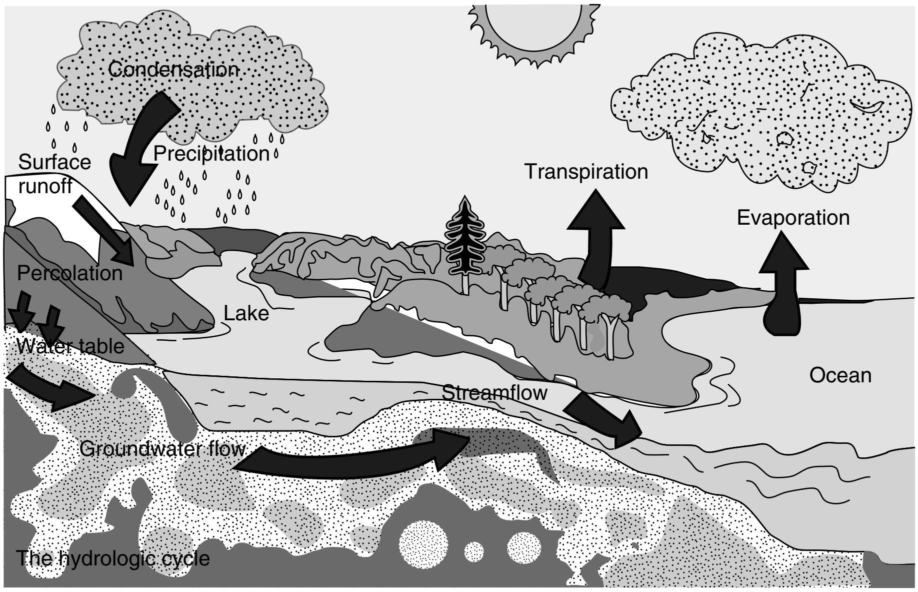

Figure 1. Circulation of water—the hydrologic cycle.

budget approach, experts can examine existing watershed systems, quantify the effects of management impacts on the hydrologic cycle, and, in some cases, predict or estimate the hydrologic consequences of proposed activities. Further, the budget can be estimated in a simplified manner, if components are categorized into input, output, and storage (2). Based on the principle of conservation of mass, inputs such as rainfall, snowmelt, and condensation must balance with changes in storage and outputs, which include stream flow, groundwater seepage, and evapotranspiration (outflow subtracted from the inflow is equal to change in storage). This hydrologic balance, or water budget, is an application of the conservation of mass law expressed by the equation of continuity:

$$I - O = \Delta S \tag{1}$$

where $I = \text{inflow}$, $O = \text{outflow}$, and $\Delta S = \text{change in storage}$.

It is to be noted that Eq. 1 does not consider the quality aspects in water budgeting using conservation of mass theory. As a result of increasing population, growth steady rise in irrigation activities, rapid expansion of industries, urbanization, and modern ways of living, tremendous pressure on the available water has been exerted, which is highly uneven in its spatial and temporal distribution both in quantity and quality. Therefore, the water pollution is a major environmental concern all over the world. In the following section, pollutant loads in hydrological variables generated and altered because of human activities have been presented.

MEASURED POLLUTANT LOADS

It is essential for water resources planners and managers to look for water pollution generated in different variables of hydrologic cycle because of human activities. It has been found that among various hydrological variables, rainfall, runoff (including snowmelt runoff), seepage, and groundwater flow are affected, and the water quality has been degraded beyond permissible limits in many instances.

Pollution Measurement in Rainfall

Acids are the main source of pollution in the rainfall. They occur naturally or because of human activities. Both the phenomena are discussed in the following section.

Pollution Because of Natural Phenomena. Natural sources of acids exist, such as volcanoes, natural geysers, and hot springs. Nature has developed ways of recycling these acids by absorbing and breaking them down. These natural acids contribute to only a small portion of the acidic rainfall in the world today. In small amounts, these acids actually help dissolve nutrients and minerals from the soil so that trees and other plants can use them for food. The large amounts of acids produced by human activities overload this natural acidity.

Pollution Because of Human Activities. On a daily basis, human activities (industrial, agricultural, and residential) cause vast quantities of natural and synthetic chemicals to be emitted into the atmosphere. Once released, the substances are dispersed throughout the globe by air currents.

Scientists have discovered that air pollution from the burning of fossil fuels is the major cause of acid rain (Fig. 2). Power plants use coal and oil to produce the electricity we need to heat and light our homes and to run our electric appliances. Natural gas, coal, and oil are burnt to heat homes. Cars, trucks, and airplanes use gasoline, another fossil fuel. The smoke and fumes from burning fossil fuels rise into the atmosphere and combine with the moisture in the air to form acid rain (3).

When water droplets form and fall to the earth, they pick up particles and chemicals that float in the air. Even

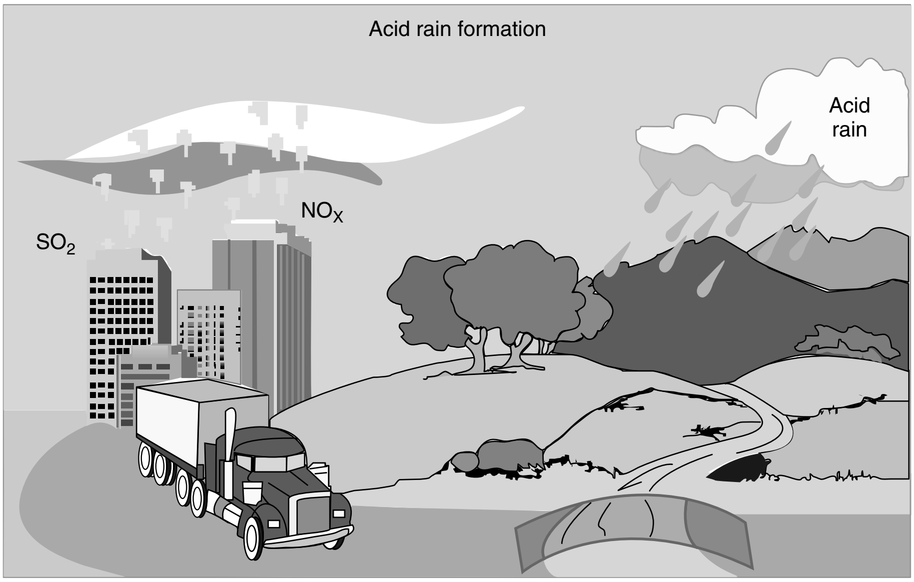

Figure 2. Acid rain formation.

clean, unpolluted air has some particles such as dust or pollen. Clean air also contains naturally occurring gases such as carbon dioxide. The interaction between the water droplets and the carbon dioxide in the atmosphere gives rain a pH of 5.6, making even clean rain slightly acidic. Other natural sources of acids and bases in the atmosphere may lower or raise the pH of unpolluted rain. However, when rain contains pollutants, especially sulfur dioxide and nitrogen oxides, the rain water can become very acidic.

The main chemicals in air pollution that create acid rain are sulfur dioxide and nitrogen oxides. Acid rain usually forms high in the clouds where sulfur dioxide and nitrogen oxides react with water, oxygen, and oxidants, which forms a mild solution of sulfuric acid and nitric acid. Sunlight increases the rate of most of these reactions. Rainwater, snow, fog, and other forms of precipitation containing those mild solutions of sulfuric and nitric acids fall to the earth as acid rain with a consequent decrease of water pH (pH below 5.5–5.6). Figure 3 depicts the acid rain formation in the atmosphere.

Effects of Acid Rain. Acid rain does not account for all of the acidity that falls back to earth from pollutants. About half the acidity in the atmosphere falls back to the earth through dry deposition as gases and dry particles. The wind blows these acidic particles and gases onto buildings, cars, homes, and trees. In some instances, these gases and particles can eat away the things on which they settle. Dry deposited gases and particles are sometimes washed from trees and other surfaces by rainstorms. When that happens, the runoff water adds those acids to the acid rain, making the combination more acidic than the falling rain alone. The combination of acid rain plus dry deposited acid is called acid deposition.

Acid rain can contaminate drinking water, damage vegetation and aquatic life, and erode buildings and

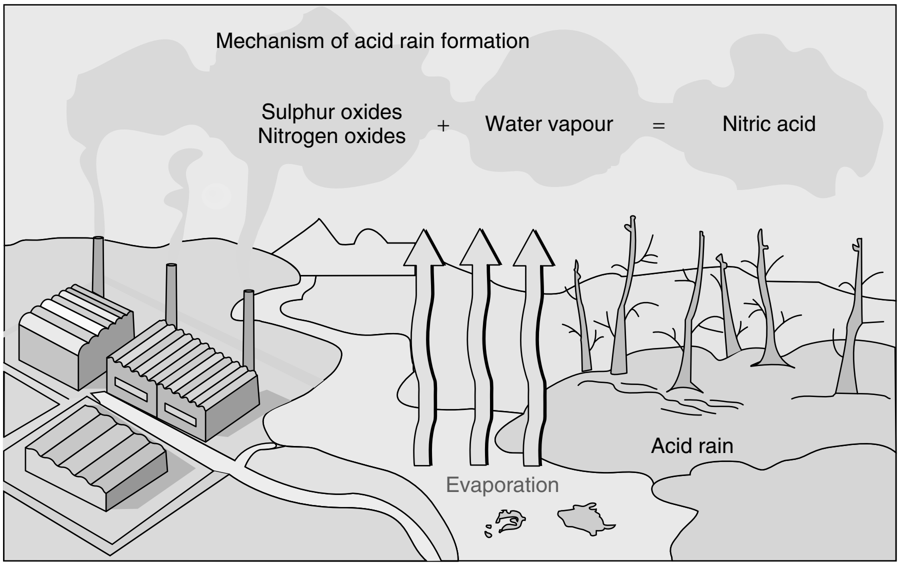

Figure 3. Acid rain formation in the atmosphere.

monuments. The water bodies most susceptible to change because of acid precipitation are those whose catchments have shallow soil cover and poorly weathering bedrock, e.g., granite and quartzite. These soil types are characterized by the absence of carbonates that could neutralize acidity. The runoff water from such areas is less buffered than from areas such as limestone catchments, with an adequate level of carbonate. Such catchments and waters are termed acid-sensitive (poorly buffered), and can suffer serious ecological damage because of artificially acidified precipitation from air masses downwind of major emissions.

Numerous studies show that acid precipitation damages forests and may cause significant decreases in productivity. Acid precipitation causes foliar damage to birch and pines, impairs seed germination of spruce seeds, erodes protective waxes from oak leaves, and leaches nutrients from plant leaves. Scientists believe that damage to trees may increase the likelihood of avalanches and landslides because trees help retain snow cover and soil cover on steep mountainsides.

Concern for agriculture has also been raised by numerous researchers. Acid precipitation is particularly harmful to buds; therefore, acids falling on plants in the spring might impair growth. Acids may also damage plants by altering the soil. For example, acid rain may leach important elements from the soil, resulting in lower yield and reduced agricultural output. Acidification of soils may also impair soil bacteria and fungi that play an important role in nutrient cycling and nitrogen fixation, both essential to normal plant growth.

The toxic action of acidity is most pronounced in its effect on the fish and macroinvertebrate fauna of streams and lakes. Dissolved aluminium concentrations as low as 0.05–1 mg can kill such organisms, whereas lower levels may have nonlethal effects such as respiratory difficulty, impaired growth, and reproductive ability, which, in the long term, may lead to their elimination.

Pollution Measurement in Runoff

The portion of water that does not infiltrate the soil but flows over the surface of the ground to a stream channel is called surface runoff. Water always takes the path of least resistance, flowing downhill from higher to lower elevations, eventually reaching a river or its tributaries. Runoff is the greatest source of water pollution, contributing from 50–60% of the pollutant load. As this water flows over agricultural lands, urban areas, hills, and forest land, it picks up different types of pollutants such as silt particles, microbial contaminants, pesticides, fertilizer, pet waste, litter, industrial wastes, urban wastes, and oil.

Natural Sources of Runoff Pollution. Heavy rain falling on exposed soil can cause substantial leaching of nitrate, some of which goes directly into rivers, but most of which percolates into the groundwater from where it may eventually reach the rivers if no natural denitrification occurs. Assessment of trends in nitrate concentrations should be undertaken on a long-term basis (i.e., frequent sampling for more than 10 years' duration).

In regions of high erosion (e.g., steep slopes, heavy rainfall, highly erodible rocks) where natural total suspended solids (TSS) already exceed 1 gm per litre, intensive agriculture aggravates the natural erosion rates.

Increased mineral salts in rivers may develop from potash mines, salt mines, iron mines, and coal mines. Mining wastes result in increases in specific ions only, such as Cl^- and Na^+ from potash and salt mines, and SO_4^{2-} from iron and coal mine wastes. The changes in ionic contents and the ionic ratio of waters are very often linked to pH changes. Mine wastewaters are generally very acidic ($pH \leq 3$), whereas industrial wastes may be basic or acidic.

Increased evaporation and evapotranspiration in the river basin (mainly in arid and subarid regions) also increase minerals. Evaporation affects all ions and as calcium carbonate reaches saturation levels, calcium sulphate-rich waters or sodium chloride-rich waters are produced. Increased evaporation and evapotranspiration in the river basin (mainly in arid and subarid regions) also increase minerals. Salinization resulting from evaporation usually leads to more basic pH levels.

In colder regions where snowmelt has a significant hydrological influence, the accumulated acidic deposition in the snow may be released when it melts. Melting snow during spring releases its acid in a sudden torrent, quickly elevating the acidity of lakes and streams. This surge of acids coincides with the sensitive reproductive period for many species of fish. The first 30% of the melt water contains virtually all of the acid and typically has a pH of 3 to 3.5, which is toxic to eggs, fry, and adult fish as well, which can cause a sudden acid pulse, which may be more than one pH unit lower than normal (4).

Runoff Pollution Because of Human Interference. There are many causes for water pollution, but two general categories exist: direct and indirect contaminant sources (Fig. 4). Direct sources include effluent outfalls from factories, refineries, waste treatment plants, etc. that emit fluids of varying quality directly into water supplies.

Indirect sources include contaminants that enter from soils/groundwater systems and from the atmosphere via rain water. Soils and groundwater contain the residue of human agricultural practices (fertilizers, pesticides, etc.) and improperly disposed of industrial wastes (5). Atmospheric contaminants are also derived from human practices (such as gaseous emissions from automobiles,

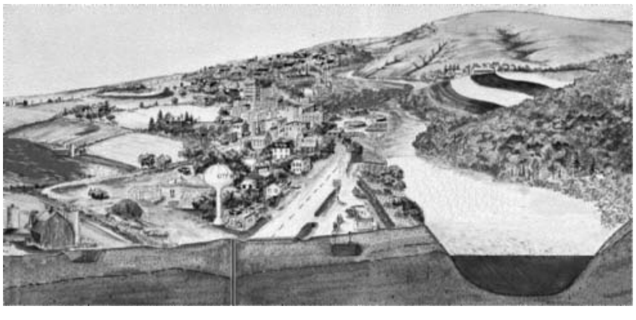

Figure 4. Representation of direct and indirect sources of pollution.

Table 1. Nonpoint Sources Pollution and Pollutants

Category of Nonpoint Source	Pollution Causing Activities	Principal Pollutants
Agriculture	Crop production, animal production, farm production	Nutrients, salts, fertilizer, pesticides, organic matters
Forestry	Access road construction and operation, harvesting systems, logging, crop regeneration, other silvicultural processes	Nutrients, pesticides, organic pollutants
Mining	Exploration, construction of facilities, mine operation, mine abandonment	Nutrients, waters, dissolved minerals, salinity
Construction	Land development, transportation and communication networks, water resources facilities	Chemicals, biological materials
Urban runoff	Precipitation discharges containing pollutants, accelerated and concentrated by urban surfaces and collection system	Organic materials, coliform bacteria, pesticides, nutrients, heavy metals
Hydrologic modifications	Channel modifications, farm drainage, dams, resource recovery, and related activities	Nutrients, pesticides, thermal chemicals, micro-organisms
Residual wastes	Foregoing categories create residual waste not discharged to water but conveyed by runoff and infiltration	Both hazardous and nonhazardous

factories, and even bakeries as discussed earlier). Table 1 reflects the pollution causing activities and their principal pollutants of nonpoint sources of water pollution.

Many physical, chemical, and biological parameters are measured to determine the overall quality of water. However, the quality of water is based not only on the concentration of substances but also on the intended use of the water. A person using water for drinking would have a different set of criteria for determining the water quality compared with a person using the water for swimming. In Tables 2, 3, and 4, the water quality criteria based on the use of water (such as drinking, bathing, and irrigation) is presented (8).

Pollutant Loads in Groundwater

Where water infiltrates the ground, gravity pulls the water down through the pores until it reaches a depth in the ground where all of the spaces are filled with water (Fig. 5). At this point, the soil or rock becomes saturated, and the water level that results is called the water table. The water table is not always at the same depth below the land surface. During periods of high precipitation, the water table can rise. Conversely, during periods of low precipitation and high evapotranspiration, the water table falls. The area below the water table is called the saturated zone, and the water in the saturated zone is called groundwater. The area above the water table is the unsaturated zone.

Groundwater is found in aquifers, which consist of soil or rock in the saturated zone that can yield significant amounts of water. In an unconfined aquifer, the top of the aquifer is defined by the water table. Confined aquifers are bound on the top by impermeable material, such as clay. Water in a confined aquifer is normally under pressure and can cause the water level in a well to rise above the water table. If the water rises above the ground surface, it is designated a flowing artesian well. A perched water table occurs when water is held up by a low permeability material and is separated from a second water table below by an unsaturated zone. In the saturated zone, groundwater flows through the pores of the soil or rock both laterally and vertically.

Water moving from an aquifer and entering a stream or lake is called groundwater discharge, whereas any water entering an aquifer is called recharge. An aquifer may receive recharge from these sources, an overlying aquifer, or more commonly from precipitation followed by infiltration. The pollutants present in soil also move along the water entering the aquifer and pollute the groundwater.

How We Contaminate Groundwater. Any addition of undesirable substances to groundwater caused by human activities is considered to be *contamination* (9). It has often been assumed that contaminants left on or under the ground will stay there. This assumption has been shown to be wishful thinking. Groundwater often spreads the effects of dumps and spills far beyond the site of the original contamination. Groundwater contamination is extremely difficult, and sometimes impossible, to clean up.

Groundwater contaminants come from two categories of sources: point source and distributed, or nonpoint source. They are as follows:

Point sources

The main sources are:

- On-site septic systems
- Leaky tanks or pipelines containing petroleum products
- Leaks or spills of industrial chemicals at manufacturing facilities
- Underground injection wells (industrial waste)
- Municipal landfills
- Livestock wastes
- Leaky sewer lines
- Chemicals used at wood preservation facilities
- Mill tailings in mining areas
- Fly ash from coal-fired power plants
- Sludge disposal areas at petroleum refineries
- Land spreading of sewage or sewage sludge
- Graveyards
- Road salt storage areas

Table 2. General Quality Criteria for Raw Waters Used for Organized Community Water Supplies (Surface and Groundwater)

(a) Primary parameters (frequency of monitoring may be daily and/or even continuous using automatic instruments for few parameters like pH, DO, and Conductivity)

Parameter	Range/limiting value		Note
	Use with Only Disinfection	Use After Conventional Treatment	
pH	6.5–8.5	6–9	To ensure prevention of corrosion in treatment plant and distribution system and interference in coagulation and chlorination
Color, Pt scale, Hz units	<10	10–50	Color may not get totally removed during treatment (not applicable in monsoon period)
Total suspended solids, mg/l	<10	10–50	High SS may increase the cost of treatment
Odor, dilution factor	<3	3–10	May not be easily tackled during treatment to render water acceptable
Dissolved oxygen, (% saturation)	90–110	80–120	
Biochemical oxygen demand, mg/l	<3	3–5	Could cause problems in treatment, larger chlorine demands, and residual taste and odor problems.
Total Kjeldahl' Nitrogen, mg/l	<1	1–3	Same as above
Ammonia, mg/l	<0.05	0.05–1.5	Same as above
Fecal coliform, MPN/100 ml	<200	200–2000	The criteria world be satisfied if, during a period of one week, not more than 5% samples show greater than 20,000 MPN/100 ml, and not more than 20% of samples show greater than prescribed limit
Conductivity, umhos/cm	<1000	1000–2000	High conductivity may cause unpalatable mineral taste and physiological disorders
Chloride, mg/l	<200	200–300	May cause physiological impact and unpalatable mineral taste
Sulphates, mg/l	<100	100–250	May cause digestive abnormality on prolonged consumption
Phosphates, mg/l	<0.1	0.1–0.3	May interfere with coagulation
Nitrate, mg/l	<30	30–50	High nitrate/nitrite may cause methamoglobinemia
Fluoride, mg/l	<1.0	1.0–1.5	Prolonged consumption of water containing high F may cause fluorosis
Surfactants, mg/l	<0.1	0.1–0.2	May impair treatability and cause foaming

Note: There should not be any significant wastewater discharge in the upstream (up to 5 km) of the water intake point.

(b) Additional Parameters for Periodic (Say Monthly/seasonal) Monitoring

Parameters	Desirable	Acceptable	Note
Dissolved iron, mg/l	<0.3	3–5	Higher iron affects the taste of beverages and causes stains
Copper, mg/l	<0.5	0.5–1.0	May result in damage of liver
Zinc, mg/l	<0.1	<5.0	May cause bitter stringent taste
Arsenic, mg/l	<0.002	0.002–0.05	Can cause hyperkertosis and skin cancer in human beings
Cadmium, mg/l	<0.001	0.001–0.005	Toxic to man
Total-Cr, mg/l	<0.05	<0.05	Toxic at high doses
Lead, mg/l	<0.05	0.04–0.05	Irreversible damage to the brain in children, anaemia, neurological disfunction, and renal impairment
Selenium, mg/l	<0.01	<0.01	Toxic symptoms similar to arsenic
Mercury, mg/l	<0.0002	0.0002–0.0005	Deadly, poisonous, and carcinogenic
Phenols, mg/l	<0.001	<0.001	Toxic and carcmogenic; may also cause major problem of taste and odor
Cyanides mg/l	<0.05	<0.05	Larger consumption may lead to physiological abnormality
Any Polycyclic Aromatic hydrocarbons, mg/l	<0.0002	<0.0002	Carcinogenic
Total Pesticides, mg/l	<0.001	<0.0025	Tend to bioaccumulate and biomagnify in the environment; toxic and carcinogenic

Note: The parameters mentioned need to be examined only when (i) there are known natural sources in the upstream basin region likely to contribute, (ii) there are known problems in respect to the parameter, (iii) there are other well-founded apprehensions.

- Wells for disposal of liquid wastes
- Runoff of salt and other chemicals from roads and highways
- Spills related to highway or railway accidents
- Coal tar at old coal gasification sites
- Asphalt production and equipment cleaning sites

Among the more significant point sources are municipal landfills and industrial waste disposal sites. When either

Table 3. Quality Criteria for Waters of Mass Bathing Reaches

Parameter	Desirable	Acceptable	Note
Fecal coliform, MPN/100 ml	<100	100–1000	If MPN of fecal coliform is noticed to be more than 100/100 ml, then regular tests should be carried out. The criteria would be satisfied if, during a period, not more than 5% samples show greater than 5000 MPN/100 ml, and not more than 20% of samples show greater than 1000/100 ml
pH	6–9	6–9	
Color Pt scale Hz units	No abnormal color	No abnormal color	No color of anthropogenic origin
Mineral oil, mg/l	No film visible	No film visible	
Methylene blue active substances, mg/l	<0.3	0.3–1.0	Skin problem likely
Phenols, mg/l	<0.005	0.005–0.01	Skin problems and odor problem
Transparency (Sechhi depth)	>2 m	0.5–2 m	
Biochemical oxygen demand, mg/1	<5	5–8	High organic matter may be associated with coliform pathogens
Dissolved oxygen, % saturation	80–120	60–140	–
Floating matter of any type	Absent	Absent	

Note: No direct or indirect discharge of untreated domestic/industrial wastewater within 5 km upstream.

Table 4. General Water Quality Criteria for Irrigation Waters (6,7)

Parameter	Generally Acceptable	Relaxation for Special Planned (Exceptional Notified) Cases	Note
Conductivity, umhos/cm	<2250	2250–4000	The irrigation water having conductivity more than 2250 µs/cm at 25 °C may reduce vegetative growth and yield of the crops. It may also increase soil salinity, which may affect its fertility
Fecal coliform, MPN/100 ml	<5000	No limit	No limit for irrigating crops not eaten raw
pH	6–9	5–9.5	Soil characteristics important
Biochemical oxygen demand, mg/l	<100	No limit	Land can absorb organic matter faster than water; however, the conditions of water logging should be avoided. Stagnant water should not persist for more than 24 hours
Floating materials such as wood, plastic, rubber, etc.	Absent	No limit	May inhibit water percolation
Boron, mg/l	<2	<2	Boron is an essential nutrient for plant growth; however, it becomes toxic beyond 2 mg/l
SAR	<26	No limit	SAR beyond 26 may cause salinity and sodicity in the soil. When it exceeds the limit, method of irrigation and salt tolerance of crops should be kept in mind
Total heavy metals mg/l	< 5.0	5.0–15.0	

of these occur in or near sand and gravel aquifers, the potential for widespread contamination is the greatest. Other point sources are individually less significant, but they occur in large numbers all across the country. Some of these dangerous and widespread sources of contamination are septic tanks and leaks and spills of petroleum products and of dense industrial organic liquids.

Septic systems are designed so that some of the sewage is degraded in the tank and some is degraded and absorbed by the surrounding sand and subsoil. Contaminants that may enter groundwater from septic systems include bacteria, viruses, detergents, and household cleaners, which can create serious contamination problems. Despite the fact that septic tanks and cesspools are known sources of contaminants, they are poorly monitored and studied very little.

Nonpoint (Distributed) Sources

Nonpoint sources comprise the other type-generalized discharges of waste whose location cannot be identified. The main sources are agriculture, forestry, mining, construction, urban runoff, hydrologic modifications, and residual wastes. Table 1 reflects the pollution-causing activities and their principal pollutants of nonpoint sources of water pollution.

Contamination can render groundwater unsuitable for use. Scientists also predict that, in the next few decades, more contaminated aquifers will be discovered, new contaminants will be identified, and more contaminated groundwater will be discharged into wetlands, streams, and lakes.

Once an aquifer is contaminated, it may be unusable for decades. The residence time, as noted earlier, can be

Figure 5. Infiltration of surface water.

anywhere from two weeks to 10,000 years. Furthermore, the effects of groundwater contamination do not end with the loss of well water supplies. Several studies have documented the migration of contaminants from disposal or spill sites to nearby lakes and rivers as this groundwater passes through the hydrologic cycle, but the processes are not as yet well understood. In Canada, pollution of surface water by groundwater is probably at least as serious as the contamination of groundwater supplies. Preventing contamination in the first place is, by far, the most practical solution to the problem. Prevention can be accomplished by the adoption of effective groundwater management practices by governments, industries, and all individuals. Although progress is being made in this direction, efforts are hampered by a serious shortage of groundwater experts and a general lack of knowledge about how groundwater behaves.

BIBLIOGRAPHY

1. Chow, V.T. (1959). *Open Channel Hydraulics*. McGraw-Hill, NewYork.
2. Brooks, K.N., Ffolliott, P.F., Gregersen, H.M., and Thames, J.L. (1993). *Hydrology and the Management of Watersheds*. Iowa State University Press, Ames, IA.
3. USEPA. (2003). *Mid Atlantic Integrated Assessment*. Office of Research & Development, USEPA Region II, Philadelphia, PA.
4. Jones and Bisson. (1984).
5. Jha, R., Ojha, C.S.P., and Bhatia, K.K.S. (2005). Assessment of pollution outflow from large agricultural areas. In: *The Encyclopedia of Water*. John Wiley & Sons, New York.
6. Meybeck, M. (2001). River Basin under Anthropocene conditions, In: *Science and Integrated Coastal Management*. B. von Bodungen and R.K. Turner (Eds.). Dahlem Univ. Press, Berlin, Germany.
7. Trivedi, R.C. (2003). *Water quality standards, 2nd International Conference on Water Quality Management*. CBIP, India, pp. 18–33.
8. Bartram, J. and Ballance, R. (Eds.). (1996). *Water Quality Monitoring: A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes*. Chapman Hall, London.
9. Cherry, J.A. (1987). Groundwater Occurrence and Contamination in Canada. In: *Canadian Aquatic Resources*. M.C. Healey and R.R. Wallace (Eds.). *Canadian Bulletin of Fisheries and Aquatic Sciences*. **215**: 395. Department of Fisheries and Oceans, Ottawa, Canada.

HYDROLOGIC THRESHOLDS

Bellie Sivakumar
University of California
Davis, California

INTRODUCTION

Hydrologic science has witnessed an exponential growth in the last century, thanks to sheer necessity (or even mere curiosity) and technological advances. Population growth and lifestyle changes have contributed (both positively and negatively) to hydrologic science, in terms of water resources availability, planning, and management. Invention of powerful computers, remote sensing, geographic information systems, worldwide networking facilities, and many other technologies have, in one way or another, allowed better hydrologic data collection, analysis, and sharing, which, in turn, facilitated more effective and efficient water resources planning and management. Despite these developments, our lack of ability to adequately plan and manage our water resources continues to be a matter of grave concern.

Setting aside, for the present discussion, the population growth and the socio-politico-economic scenarios, our inability to adequately plan and manage our water resources may largely be attributed to two factors:

1. our inability to collect all the relevant data for the system of interest; and
2. our inability to understand the past, model the present, and forecast the future, even with collected data; in other words, our inability to accurately “replicate and extrapolate” the data.

In regard to 1, we may take some solace by our experience, albeit limited, that often only a few processes dominate responses in a given hydrologic system and, therefore, may focus on collecting more reliable data relevant only to these dominant processes (1). This concept, fittingly termed by Grayson and Blöschl (1) as the "Dominant Processes Concept" (DPC), has been receiving increasing attention among hydrologists in one way or another (2–4). More recently, Sivakumar (5) presented a new proposal for moving forward with the DPC, emphasizing on integration of theoretical concepts and practical knowledge for more effective and efficient data collection and analysis procedures.

As for 2, notwithstanding data quantity and quality issues, our inability to accurately replicate and extrapolate the data may largely be because of the inherent limitations of our modeling and forecasting procedures and the inappropriate selection of data sets used therein for "training" (learning or calibration) and "testing" (verification or validation), respectively. Sivakumar (6) points out this "data handling" problem in a rather different and multifaceted manner, by reviewing sample studies on recent data modeling and forecasting techniques in hydrology, including artificial neural networks and chaos theory. More specifically, Sivakumar (6) argues that any bias resulting in our extrapolations is in line with the procedural or data bias in our learning, which is chosen to capture (the statistics of) our desired "extreme" or "average" extrapolations. To overcome this problem, he suggests the use of different models for different ranges of data, rather than a single model for the entire data range, as is commonly used. Deriving an analogy between humans and catchments, and emphasizing their internal characteristics and responses to external events, Sivakumar (6) discusses the usefulness of "thresholds" in identifying these data ranges. Reiterating that the "threshold" concept is already in wide use in different implicit forms in hydrology, he stresses the need for bringing these implicit forms under one common umbrella, and explicitly termed as "thresholds," for everyone to understand.

In the spirit of the proposal by Sivakumar (6), this article presents further ideas as to the potential role of thresholds in hydrology. Specifically, it aims to launch the threshold concept at a much wider spectrum than the narrow time-series/parameter-optimization spectrum presented by Sivakumar (6), which is attempted by discussing three important issues: (a) relevance of thresholds in hydrology; (b) challenges in the identification and quantification of thresholds; and (c) importance of interdisciplinary efforts on threshold research.

THRESHOLDS IN HYDROLOGY

Simply put, thresholds are "critical" points separating different states (or phases) that exist or are possible in a given system. These points are critical because they are the "enablers" or "barriers" for transitions between states that may have enormous differences in their properties (e.g., physical, chemical, and biological). As these transitions are often "catastrophic", thresholds are generally defined as points separating "average" and "extreme" events.

The relevance of thresholds in hydrology should be evident even from our simple observations of "average", "high", and "low" rainfall periods, which often lead to "normal", "flood", and "drought" periods, respectively. Whether it is rainfall, runoff, or any other hydrologic phenomenon, transitions between different states are essentially the result of the underlying physical mechanisms. These physical mechanisms are, in turn, governed by the characteristics of the land, ocean, and atmosphere. As natural and man-made changes to land, ocean, and atmosphere occur, even within their individual realms, so do changes to the underlying physical mechanisms and transitions from one state to another. When such changes occur across these domains, the "probability" of changes to physical mechanisms and transitions between states increases, because a much more complex coupled land-ocean-atmospheric system is in effect.

The land-ocean-atmospheric system as a whole is a "network" composed of interconnected parts, which in themselves are "networks" (e.g., hydrologic system) composed of their own interconnected parts, and so on. In this network(s), the relationships between the parts are circular (in an abstract, but not necessarily spatial, sense), which causes "feedback" because the actions of each component "feeds back" on itself (e.g., rainfall-evaporation-rainfall or rainfall-runoff-evaporation-rainfall). This feedback mechanism can be either positive or negative, but both types are relevant and important; a negative feedback stabilizes the system, whereas a positive feedback often destabilizes the system.

Feedbacks are not only inherent in hydrologic systems but are also inherently nonlinear in their nature. The nonlinear nature of the feedbacks, when combined with the complex nature of the hydrologic systems, may cause bifurcations, i.e., rapid, unpredictable, and often catastrophic transitions to new states, occurring at critical points or thresholds. One state that can develop at a bifurcation is chaos, which demonstrates sensitive dependence to initial conditions or "butterfly effect" [e.g., (7,8)]. This effect, i.e., infinitesimal changes in inputs leading to enormous differences in outcomes, amplifies uncertainty and limits predictability. In fact, the most complex behavior in hydrologic systems (and other networks) often occur at the edge of chaos, i.e., at a transition from chaos to order or vice-versa. A particular signature of this state is the presence of fractals, i.e., patterns in space and/or time that are self-similar on many scales [e.g., (9)].

A related property of a feedback, nonlinear, chaotic, and fractal system is self-organization, which refers to the spontaneous development (organization) of the system and its evolution to different and more complex states. Self-organizing systems demonstrate the "emergence" of properties. The emergent properties are not inherent in any of the parts of the system examined in isolation, but emerge only as a result of the interactions between the parts that characterize the system as a whole. Self-organization is particularly relevant for hydrologic systems. For example, a river exists as an interaction between the flow of water, the materials making up the riverbank and channel, the elevation and terrain of the land, the climate, temperature and rainfall, the law of gravity, the forces of turbulence, and more. Although it is possible to understand each component of the river by reduction, this is not adequate for understanding the river as a whole, because the emergent properties and evolution of the river are governed by the interactions among the components, which stretch beyond the domain of any one single component.

IDENTIFICATION AND QUANTIFICATION OF THRESHOLDS

With the relevance and importance of thresholds in understanding hydrologic systems evident, the immediate question is how to identify and quantify them. In this regard, knowledge of general system characteristics (e.g., climatic condition, geography, land use, wet or dry state) and system evolution (i.e., history of changes to the system and the associated changes to hydrologic events, such as rainfall amounts, frequency of floods and droughts) may provide some basic information on the types of thresholds (i.e., identification) and their levels (i.e., quantification). With this information, additional threshold criteria may be established using methods and tools, available at our disposal, based on traditional linear and reductionistic approaches.

We may be able to achieve some success with the above; however, it will only be partial, because linear and reductionistic approaches only study the system components in parts and attempt to estimate important statistics, whereas hydrologic systems are governed by nonlinear feedbacks, bifurcations, chaos, fractals, and emergent properties that run across the individual component parts and require a nonlinear and holistic approach(es).

As much as the concept of a nonlinear and holistic approach is exciting, the following point is noteworthy. Although a nonlinear and holistic approach may be *the* approach for the system as a whole, it may *not* be the one for the individual component parts after all, for which a linear and reductionistic approach may perform better. As a “complete” explanation of a system requires accurate analysis of the parts (reductionism) AND the whole system complexity (holism), a coupled reductionistic-holistic approach may work, in all probability, better than either of the two approaches when adopted independently. The impending challenge, therefore, is to devise such a coupled approach that would provide equal importance to the whole system as well as to its parts.

New theoretical developments and methodological advancements may indeed be needed to address this challenge. To what extent will we give priority to this challenge and how quickly will we be successful in formulating a reliable coupled reductionistic-holistic approach remain to be seen. Although we must start laying the foundation for this task, sooner rather than later, we must also find ways to understand, use, interpret, and even integrate the existing theories and methods better than we have thus far, because the latter may provide some important information that may, subsequently, become significant in implementing the former.

An examination of hydrologic literature clearly reveals the opportunities and possibilities in this regard, as evident from the following two examples. Sivakumar (12), through a discussion of three fundamental ideas of chaos theory (maybe less popular among hydrologists), presents an interpretation of “chaos theory as a bridge between our popular views of determinism and stochasticity”. Addressing the Dominant Processes Concept, Sivakumar (5) proposes an integration of ideas from nonlinear dynamic theory, expert system, and optimization technique, whose usefulness may be viewed at a much broader spectrum of model simplification.

The fact that interpretation and integration of existing theories and development of new ones often require knowledge of not only hydrology but also other fields, including atmospheric and ocean sciences, ecology, geomorphology, mathematics, and physics, our ability to succeed in identifying and quantifying thresholds mainly lies in our ability to perform collaborative work with researchers in these fields. This topic is discussed next.

INTERDISCIPLINARY RESEARCH ON THRESHOLDS

Thresholds are problem-dependent. Although a connection between thresholds of inputs (e.g., rainfall) and outputs (e.g., runoff), often exists a threshold for an input does not automatically lead to a corresponding threshold in an output. For example, in a rainfall-runoff-sediment transport scenario, the threshold for erosion (a result of runoff) may be different from that for runoff (a result of rainfall), which, in turn, may be different from that for rainfall (a result of atmospheric mechanisms). Similarly, thresholds are also scale-dependent. Although scale-invariance in space and time exists in hydrologic systems and phenomena, different thresholds may exist at different scales, depending on the dominant catchment parameter at the given scale (e.g., slope, land use) and the dominant catchment mechanism (e.g., erosion, interception, evapotranspiration).

These observations indicate the complications involved in even arriving at a definition for thresholds for a given hydrologic system. Also, hydrologic systems play different roles with respect to other related systems (e.g., atmospheric, ecologic, environmental, and geomorphic systems) and vice-versa, which makes the problem even more difficult. Further complications develop when consideration of a one-to-one system scenario or a multiple-system scenario is needed. In view of these, the only hope we can have is to approach the problem from an interdisciplinary view, and multidisciplinary to be more specific, involving researchers from as many related fields as possible.

Establishing interdisciplinary collaborations is not an easy task, particularly at a time we, hydrologists, ourselves seem to be moving away from each other, with our emphasis on specific data analysis techniques rather than common hydrologic problems (6). There is encouraging news, however, as efforts are already being made to address the threshold problem from an interdisciplinary perspective. For instance, the Environmental Hydroscience Discipline Group of the University of Western Australia is working toward the organization of a multidisciplinary workshop entitled “Thresholds and Pattern Dynamics—A New Paradigm for Predicting Climatic Driven Processes,” which is planned to be held July 4–7, 2005 (13). The purpose is to bring together researchers from different disciplines, including seismology, hydrology, ecology, geomorphology, and economics. A session on introduction of threshold concepts in these disciplines is also planned, an important focus of which is "learning each other's problems and language," a point also highlighted by Sivakumar (6) in his proposal of "thresholds" for its simplicity and generality. We will need to make continued efforts in this direction to be successful in solving the threshold "conundrum."

CONCLUDING REMARKS

Following up on a recent initiation for "thresholds" in hydrology (6), this article discussed their relevance, identification and quantification, and the importance of interdisciplinary research. The discussion points to the need for a new paradigm, or at least a major paradigm shift, for studying thresholds. As thresholds play crucial roles in future predictions and risk assessment, for both the whole system and its component parts, only a (new) coupled reductionistic-holistic approach seems appropriate. Formulation of such an approach, however, may benefit from existing theories and methods, particularly their integration. Interdisciplinary efforts would significantly help in these directions. Hope certainly exists with the availability and advancement of technology, communication facilities, and human resources.

BIBLIOGRAPHY

1. Grayson, R.B. and Blöschl, G. (2000). Summary of pattern comparison and concluding remarks. In: *Spatial Patterns in Catchment Hydrology: Observations and Modeling*. R.B. Grayson and G. Blöschl (Eds.). Cambridge University Press, Cambridge, UK, pp. 355–367.
2. Woods, R. (2002). Seeing catchments with new eyes. *Hydrol. Process* **16**: 1111–1113.
3. Young, P.C. and Parkinson, S. (2002). Simplicity out of complexity. In: *Environmental Foresight and Models: A Manifesto.* M.B. Beck (Ed.). Elsevier Science, The Netherlands, pp. 251–294.
4. Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R. (2003). Downward approach to hydrological prediction. *Hydrol. Process* **17**: 2101–2111.
5. Sivakumar, B. (2004). Dominant processes concept in hydrology: moving forward. *Hydrol. Process* **18**(12): 2349–2353.
6. Sivakumar, B. (2005). Hydrologic modeling and forecasting: role of thresholds. *Environ. Model Softw.* In press.
7. Lorenz. (1963).
8. Gleick, J. (1987). *Chaos, Making a New Science*. Viking, New York, pp. 1–352.
9. Mandelbrot, B.B. (1982). *The Fractal Geometry of Nature*. W. H. Freeman, New York, pp. 1–468.
10. Rigon, R., Rinaldo, A., and Rodriguez-Iturbe, I. (1994). On landscape self-organization. *J. Geophys. Res.* **99**(B6): 11,971–11,993.
11. Rodriguez-Iturbe, I. and Rinaldo, A. (1997). *Fractal River Basins: Chance and Self-Organization*. Cambridge University Press, Cambridge, UK, pp. 1–547.
12. Sivakumar, B. (2004). Chaos theory in geophysics: past, present and future. *Chaos Soliton Fract.* **19**(2): 441–462.
13. Sivapalan, M. (2004). Personal Communication.

GANGA RIVER, INDIA

Sharad K. Jain
National Institute of Hydrology
Roorkee, Uttranchal, India

The Himalayas, the great mountain chain of Asia are the source of three major river systems of the world, the Indus, the Ganga (or the Ganges), and the Brahmputra. The large and fertile plains of the Indus and Ganga in the subcontinent of Northern India have been the cradles of one of the greatest ancient civilization, the Indus Valley Civilization. The River Ganga which occupies nearly one-third of the geographical area of India is the most important and sacred river of India.

So legendary has been the socioeconomic, cultural, and religious saga of this great river that the Indian mythology and history are full of stories and incidents woven around the river and numerous pilgrim spots along its course like Rishikesh, Hardwar, Allahabad, Varanasi, and Patna.

DESCRIPTION OF GANGA BASIN

The catchment area of the Ganga falls in four countries, India, Nepal, Tibet (China), and Bangladesh. Many important tributaries of the Ganga originate in the Himalayas in India and Nepal; Bangladesh lies in the deltaic region. The total length of the river is 2525 km which makes it the fifteenth longest river in Asia and the thirtyninth longest in the world. The combined basin of the Ganga, Brahmaputra, and Meghna rivers in India, Nepal, and Bangladesh is also known by the name *Greater Ganga Basin*. The major part of the geographical area of Ganga Basin lies in India.

Though the headwaters region of the Ganga in the Himalayas is dotted by a number of mighty tributaries, the Bhagirathi river that rises from the Gangotri glacier near Gomukh at an elevation of about 7010 m above mean sea level is traditionally considered the source of River Ganga.

The other main stream that originates in the Uttranchal state of India is the Alakhnanda. Flowing downhill, Bhagirathi and Alakhnanda are joined by a number of streams such as the Mandakini, the Dhuli Ganga, and the Pindar. These two rivers (Bhagirathi and Alakhnanda) meet at a place called Devprayag and thereafter, the combined flow is known by the name Ganga. The River Ganga enters into plains at the pilgrimage center at Haridwar, and from there it flows in a south/southeasterly direction. Yamuna is the most important tributary of the Ganga that joins it on the right bank at Allahabad. The total length of the Yamuna from its origin to Allahabad is 1,376 km, and the drainage area is 366,223 sq. km. The Yamuna is a mighty river in itself and has a number of tributaries such as the Chambal, the Sind, the Betwa, the Ken, and the Tons. After confluence with Yamuna, the Ganga River flows east and is joined by a number of tributaries such as the Ramganga, the Gomti, the Ghaghra, the Gandak, the Bagmati, the Kosi, the Sone and the Damodar.

The Ganga delta is said to begin at a place known as Farakka where a barrage has been constructed to control river flow. At about 40 km downstream from Farakka, the river splits into two arms. The right arm called River Bhagirathi flows south and enters the Bay of Bengal about 150 km downstream from Calcutta. The left arm, known as Padma, turns east and enters Bangladesh. While flowing in Bangladesh, the Padma meets the River Brahmputra at a place known as Goalundo. The combined flow, still known as Padma, is joined by another mighty river Meghna at Chandpur, 105 km downstream from Goalundo. Further down, the river ultimately flows into the Bay of Bengal. The delta of the Greater Ganga Basin, one of the largest in the world, is known by the name Sunderbans.

For hydrologic studies, the entire stretch of the Ganga River in India can be divided into three stretches or reaches. The upper reach extends from the origin to Narora, the middle reach from Narora to Ballia, and the lower reach from Ballia to its delta.

An index map of the basin is given in Fig. 1.

HYDROLOGY OF THE GANGA BASIN

The Ganga Basin extends across an area of 1,086,000 sq. km. The drainage area of the basin lying in India is 861,452 sq. km, nearly 26.2% of the geographical area of the country and 79.32% of the Ganga Basin. Some tributaries such as the Ghagra, the Gandak, and the Kosi drain areas in Nepal amounting to 190,000 sq. km. The delta of the Greater Ganga Basin covers an area of 56,700 sq. km. The headwater reaches of the river receive a considerable part of the precipitation as snow, and some peaks are permanently snow covered. The average annual rainfall in the basin varies from 35 cm at the western end to nearly 200 cm near the delta.

The average annual discharge of the Ganga, the Brahmaputra, and the Meghna rivers is 16.65, 19.82, and 5.1 thousand m^3/s, respectively. The average annual flow of the Ganga at Farraka is about 525×10^9 m^3. There are large variations in the flow of the Ganga with time. Snow and glacier melt during the hot months (March to June) give large summer flows in Himalayan rivers such as the Ganga and its tributaries. The maximum discharge in these rivers is observed during monsoon months (June to September). At Goalundo, the average annual flow of the river is 11,470 m^3/s. The maximum and minimum flow at this site is 70,934 and 1161 m^3/s (1). The peak flow at Farakka in 1971 was estimated at 70,500 m^3/s.

The surface water resource potential of the Ganga and its tributaries in India has been assessed at 525×10^9 m^3 out of which 250×10^9 m^3 is considered useable NCIWRD (1). Based on the 1991 census, the per capita water availability was near 1471 m^3 per year. Though the Ganga Basin is bestowed with an abundant water resource, its occurrence/availability both in quantity and quality is not uniformly distributed either spatially or temporally. More than 75% of the annual rainfall occurs in the monsoon months of June to September. As a result, large areas are subjected to floods on one hand and droughts on the other.

The storage potential of the Ganga Basin in India has been identified at 84.46×10^9 m^3. However, till 1995, a total of only 36.8×10^9 m^3 of storage space was created. Water resources development schemes to create storage of 17.06×10^9 m^3 are under construction, and projects to provide another 29.56×10^9 m^3 of storage are in the pipeline. The total replenishable ground water resource of the Ganga Basin is estimated at 171.0×10^9 m^3, of which about 48.6×10^9 m^3 was being used by 1999. The River Ganga carries one of the world's highest sediment loads, equal to nearly 1451 million metric tons per annum.

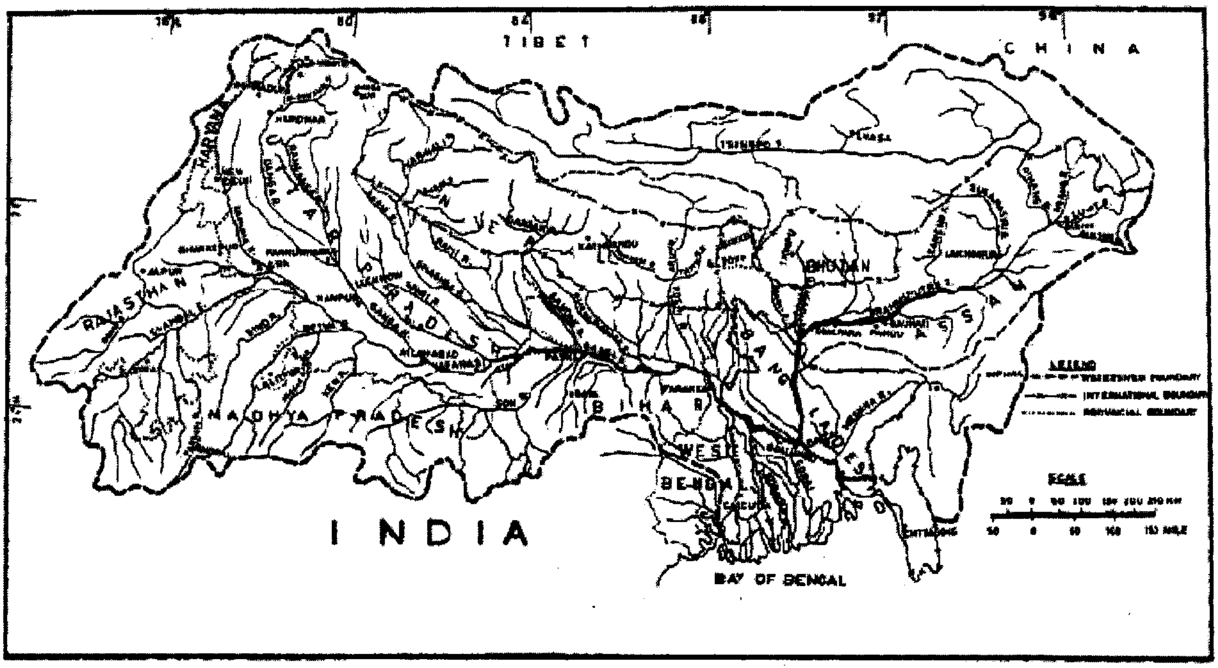

Figure 1. Index map of the Greater Ganga Basin (2).

The Ganga Basin is the largest basin in India, both catchmentwise and populationwise. The total population in the Basin per the 1991 census has been estimated at 356.8 million, which is 42% of India as a whole. The average population density in the Basin was 414 persons per sq. km as against 267 for all of India. Per the 1991 census, the Basin had 111 urban centers that had populations of more than 0.1 million, and of these, seven cities had populations exceeding a million.

The Ganga and its tributaries have formed a large flat and fertile plain in North India. The availability of abundant water resources, fertile soil, and a suitable climate have given rise to a highly developed agriculture-based civilization and one of the most densely populated regions of the world. The net sown area in the Ganga Basin in India is around 44 million hectares (M-ha), and the net irrigated area is 23.41 M-ha.

The hydroelectric potential of the Ganga Basin has been assessed as 10,715 MW at a 60% load factor. Of the 142 identified schemes in the Basin, 22 schemes that have a total installed capacity of 2437 MW are operating and 12 schemes whose installed capacity is about 2716 MW are in various stages of construction.

WATER RESOURCES DEVELOPMENT PROJECTS IN GANGA BASIN

The development of water resource projects in the Ganga Basin has a long history. One of the most remarkable of these works which includes many great civil engineering structures is the Upper Ganga Canal system whose construction was commenced by the then British government in the 1840s to ward off drought and famine in the western part of Uttar Pradesh state. Completed in 1854, the Upper Ganga Canal was to irrigate 0.7 million ha area. The canal takes off from the Ganga at Bhimgoda weir near Hardwar. The main canal is 230 km long and is one of the most exquisitely constructed structures. To expand the command area, the discharge capacity of the canal was recently augmented to nearly 297 m^3/s. The command area of the canal is about 0.924 million hectares.

The Madhya Ganga Canal takes off from Ganga at Raoli barrage near Bijnor. The capacity at the head of this 115 km long canal is 234 m^3/s. It is operated to provide irrigation to paddy crops in 114,000 hectares as well as to augment the supply to the Upper Ganga Canal system. The Lower Ganga Canal takes off from the Narora weir across the Ganga and was constructed in 1879. The main canal is 100 km long and irrigates 0.5 million ha. Near the Indo–Bangladesh border, a barrage has been constructed at Farakka to divert water to the river Hooghly to increase and maintain adequate depths of flow for navigation and operation of the Calcutta port.

The Yamuna river emerges from the hills near Tajewala where the water is taken off by the western and eastern Yamuna canals. It flows a further 280 km down to Okhla near Delhi from where the Agra Canal takes off. Chambal river is an important tributary of the Yamuna whose catchment area is 139,468 sq. km. Three important storages constructed in the basin include Gandhisagar, Jawahar sagar, and Rana Pratap Sagar cascade, which provide live storage of 8,500 million cubic m. The barrage at Kota diverts water into canals on either side that irrigate a 0.57 million ha agricultural area.

The Ramganga dam on the tributary of the same name has created a reservoir at Kalagarh that has a live storage capacity of 2,190 million m^3. A feeder channel takes off from the dam to provide extra supplies of water to the upper and lower Ganga canal and the Agra canal, besides direct irrigation. A 666,000 ha cropped area will be irrigated by this project. A system of reservoirs in the Damodar valley, Konar, Tilaiya, Tenughat, Maithon, and Panchet dam and Dugapur barrage have largely controlled floods in this subbasin. The Damodar Valley Corporation is responsible for the integrated regulation of most of these dams.

Some other important projects on the tributaries of the Ganga include Matatila on Betwa, Sarda sagar on the Sarda, Obra and Rihand on the Rihand river, and Mayurakshi on the Mayurakshi River. The Sarda canal provides protective irrigation to a nearly 0.6 million ha area. The Tehri dam on Bhagirathi (under construction) will provide live storage capacity of 2,613 million cubic m to be used for power generation and irrigation. Some other major projects under construction are Rajghat on the Betwa, Bansagar on the Sone, and Lakhwar-Vyasi on the Yamuna. A few projects are also being planned in collaboration with the government of Nepal. To boost the use of waterways, some segments of the Ganga River have been declared national waterways.

PROBLEMS IN WATER RESOURCE DEVELOPMENT OF THE GANGA BASIN

The basic problem in using water resources in the Ganga Basin is that in relation to the relatively large annual flow in the Basin, the storage capacity of existing and foreseeable reservoirs in India is not large enough to permit conservation of flows during the high flow season. The live storage capacity of all reservoirs in the Ganga Basin is less than one-sixth of the annual flow, which does not permit the desired degree of flow regulation. The majority of the good storage sites of the Basin are in Nepal, and this is leading to a delay in constructing dams. Lean season flows in the basin without adequate storage backup are not sufficient to meet the requirements for various demands. Besides, monsoon flows in the basin are so high that the Ganga and its tributaries remain in spate almost every year.

In the Ganga Basin, the flooding problem is confined mainly to the middle and terminal reaches. In general, the severity of the problem increases from west to east and from south to north. The worst flood-affected states in the Ganga Basin are Uttar Pradesh, Bihar, and West Bengal. In Uttar Pradesh, flooding is largely confined to the eastern districts; the rivers that cause flooding include the Sarada, the Ghagra, the Rapti, and the Gandak. The major causes of flooding here are drainage congestion and bank erosion. North Bihar is in the grip of floods almost every year due to spillage of rivers. In West Bengal, floods are caused by drainage problems as well as tidal effects. The Ganga

Flood Control Commission was set up by the government of India for flood management in the Ganga Basin.

WATER QUALITY CONSIDERATIONS

The numerous cities located in the Ganga Basin generate and discharge huge quantities of wastewater; a large portion eventually reaches the river through the natural drainage system. Industrial complexes which have sprung up along the rivers also dump considerable pollution loads into the river. Over the years, the Ganga and its tributaries have become the channels of transport of industrial effluents and the drains for the wastewater of cities. It is estimated that some 900 million liters of sewage is dumped into the Ganga every day; three-fourths of the pollution in the Ganga is from untreated municipal sewage. In particular, the middle reach of the basin between Kanpur and Buxar is the most urbanized and industrialized and also the most polluted segment of the Basin. Municipal and industrial wastes in dangerous concentrations enter the watercourse in this segment and pose a grave threat. Murti et al. (3) repeat results of a detailed study on Ganga River.

The pollution load dumped in the river by humans is a serious health hazard to the dense population of the Basin. Recognizing the magnitude of this problem and realizing the importance of water quality as a cardinal element of river management, the government of India started the planning and execution of several programs to check the pollution of the River Ganga and its tributaries. An ambitious program, known as the Ganga Action Plan (GAP), was initiated in 1985. Pollution abatement works for the River Ganga had been taken up in 25 class I towns (population above 1,00,000) along the main River Ganga under the three basin states of Uttar Pradesh, Bihar, and West Bengal. The main objectives of the GAP were

1. reduction of the pollution load on the river and improving the water quality as a result thereof, and
2. establishment of domestic/municipal wastewater treatment systems emphasizing resource recovery to make such systems self-sustainable as far as possible.

In addition, GAP was to serve as a model to demonstrate the methodology of improving the water quality of the other polluted rivers and waterbodies of the country to their designated best use class. A multipronged approach was adopted to achieve the objectives of the GAP. Similar other plans are in various stages of implementation in some tributaries of the Ganga.

The task of restoring the Ganga to its pristine glory will not be easy. Nevertheless, it will be worth the trouble because the sustainable development of a large geographical area and population crucially depends on it.

BIBLIOGRAPHY

1. NCIWRD. (1999). *Integrated Water Resources Development—A plan for action. Report of the National Commission for Integrated Water Resources Development*. Govt. of India, Ministry of Water Resources, New Delhi.
2. Chatruvedi, M.C. and Rogers, P. (1985). *Water Resources Planning—Some Case Studies for India*. Indian Academy of Sciences, Bangalore.
3. Krishna Murti, C.R., Bilgrami, K.S., Das, T.M., and Mathur, R.P. (1991). *The Ganga—A Scientific Study*. Northern Book Centre, New Delhi.

INTERCEPTION

Hubert H.G. Savenije
Delft University of Technology
The Netherlands

Interception is a key process in hydrology that is receiving relatively little attention. It is the first process of the hydrological cycle that a raindrop experiences after falling from the sky, and hence, all subsequent processes depend on it. Neglect of the interception process can lead to serious modeling mistakes. Interception is defined as the evaporative process that feeds moisture back to the atmosphere shortly after the rainfall has ceased (within approximately 1 day). It is the evaporation from all surfaces that have been moistened by precipitation, including leaves, soil, mulch, rock, paved, and build-up surfaces. Evaporation from interception can amount to a considerable part of all terrestrial evaporation, depending on the climate and region. The interception process behaves differently at different time scales. Formulas are presented for the computation of interception at short, daily, and monthly time scales.

INTRODUCTION

To hydrologists who focus on the terrestrial part of the hydrological cycle, precipitation is generally seen as the starting point of the water cycle. Although one could start with any other point in the cycle as well, e.g., advection of moisture to the land or evaporation from the oceans, it is common to begin the terrestrial water cycle with the rainfall, probably because rainfall is more easily measurable than are other incoming fluxes. This cycle involves many interconnected processes (fluxes), separation points, feedback loops, and stocks where the water resides over longer or shorter periods of time. Starting from the rainfall, the first process that comes into view is interception. It retains the water before it can continue its path in the water cycle, and it allows for a direct feedback loop to the atmosphere. Depending on the precise definition of this process, interception can be understood to mean different things.

First, it is important to decide whether we consider interception as a stock or a flux, or as the combination of both. If we consider it a stock, then it is the amount of rainfall that can be temporarily stored on the land (with its natural and manmade cover) to be evaporated shortly after (or during) the rainfall event. Here, we actually mean the interception capacity. If we consider interception as a flux, then it is the evaporation from intercepted water, which we express in millimeters per unit of time. If we consider interception as an integrated process, then it is the sum of the rate of change of intercepted water and its evaporation:

$$I = \frac{dS_I}{dt} + E_I \qquad (1)$$

where I is the interception process (mm/d), S_I is the interception storage (mm), and E_I is the evaporation from interception (mm/d). The interception I is the part of the rainfall that is intercepted and, after a short period of time, turns into the evaporative flux E_I. For a time scale in the order of 1 day, it is safe to assume that $I = E_I$.

Next, it is important to define the location of the interception process in the hydrological cycle. The most logical place is between the atmosphere and the first separation point where the rainfall splits into interception, surface runoff, and infiltration (see Fig. 1). The interception stock (S_I) is located at the first separation point. In this definition, the interception process includes evaporation from wet leaves, wet land cover (included manmade structures and roads), wet mulch, wet forest floor, and even wet soil. In short, it is the fast evaporation mechanism that dries moist land cover during and directly after the rain.

A more narrow definition of interception, often used, is the difference between rainfall and "throughfall," the rain that falls through the leaves of, e.g., a tree. In this definition, interception is merely the amount of water retained by the leaves of a tree. This definition is not workable for hydrologists because it leaves evaporation from land surface not catered for. Hydrologists who do use this definition, however, combine wet surface evaporation with all other evaporative fluxes and call it evapotranspiration, which is not really a process, but rather a combination of different evaporation processes with different time scales and characteristics (1).

The total evaporation E in a catchment is the sum of several different processes: interception I, transpiration T, surface evaporation E_s, and open water evaporation E_o (2). In the broader definition of interception, we combine evaporation from leaf interception with wet surface evaporation occurring on the same day as the rainfall. The amount of interception thus defined is larger than the amounts quoted in the literature based on the difference between rainfall and throughfall.

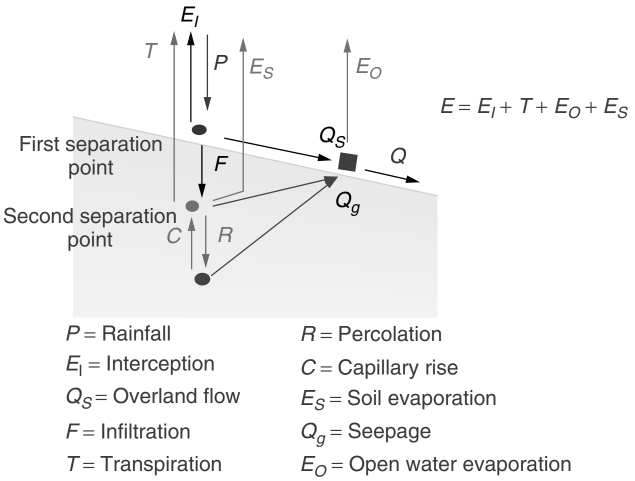

Figure 1. Separation points at the landscape interface.

THE IMPORTANCE OF INTERCEPTION

Interception is one of the most underrated, and underestimated, processes in rainfall–runoff analysis. One could call interception the "Cinderella of Hydrology." She may be beautiful and rather straightforward but apparently not interesting enough to be adequately represented in hydrological models. Some models even disregard interception completely [particularly event-based models, based on the simplified catchment model of Dooge (3, 4)], the argument being that it is generally a small proportion of the total evaporation. This disregard is a mistake, for several reasons.

First, the amount of interception is not small. Beven (5) states that evaporation from intercepted water on leaf surfaces in rough canopies can be efficient and a significant component of the total water balance in some environments. Calder (6) shows that in the upland forest catchments of Britain, evaporation from interception amounts to 35% in areas with an annual rainfall of more than 1000 mm/a, but that it is higher in areas with lower rainfall, amounting to about 40–50% in areas with 500–600 mm/a. In Calder's definition, however, interception is merely the difference between rainfall and throughfall. But if we use the broader definition presented above, interception is considerably larger. After a rainfall event, not only the leaves of vegetation are wet, but also the surface underneath (rock, soil, mulch layer, roads, build-up area, etc.), which becomes dry within the same day, particularly in warm climates. Moreover, many stagnant pools continue to evaporate until a day after the event. The additional wet surface evaporation can be as much as the interception by leaves, particularly in dry climates. These interception processes are fast, having an average residence time in the order of 1 day.

Note that the wetted soil surface should not be considered part of the soil moisture that feeds the transpiration process. The wet surface (extending to several millimetres of soil depth) feeds back the intercepted water through direct evaporation and not via a delayed transpiration process. Even a stretch of dry sand, without vegetation, can intercept water. After a rainfall event, a wet "crust" of soil is formed, underlain by dry sand, which dries out again within a day. This soil can intercept several millimeters of rainfall.

Transpiration is different. First, it is a physiological process intimately tied to CO_2 assimilation. The time scale of transpiration is determined by the soil moisture stock, which makes the time scale of the process much longer (average residence times varying between weeks and months depending on the soil depth). Finally, the process of transpiration does not change the isotope composition of the evaporated moisture, whereas evaporation of a water surface does. One can distinguish fast transpiration and delayed transpiration. Fast transpiration is from shallow rooted plants (typically grass and annual crops) with a time scale of less than a month; delayed transpiration is from deeply rooted plants (trees, shrubs, perennial crops), which have a time scale longer than a month. Fast transpiration only draws on the upper soil layer (until 50 cm depth), whereas delayed transpiration draws on deeper soil layers. Open water evaporation can be considered separately, when necessary, and is identifiable relatively simply.

DETERMINING EVAPORATION FROM INTERCEPTION

Computation of interception on an instantaneous basis is done by a Rutter model (7), based on Equation 1. For practical purposes, however, and in view of the typical time scale of the process where wetness caused by a rainfall event evaporates within a time span of a day, interception is well captured on a daily time scale. The daily process of interception is conveniently described as a threshold process, with a daily threshold D (mm/d). Daily interception (mm/d) is then computed as

$$I = \min(P,D) \tag{2}$$

which implies that if the rainfall is less than the threshold, the interception is equal to the rainfall and otherwise it is equal to the threshold value.

Monthly evaporation from interception can be determined efficiently, even for poorly gauged catchments through De Groen's method. De Groen (8) developed an analytical model for the computation of monthly interception based on the Markov property of daily rainfall[1] and using a daily interception capacity D.

$$I_m = P_m\left(1 - \exp\left(-\frac{n_r D}{P_m}\right)\right) \tag{3}$$

where I_m is the monthly evaporation from interception (mm/month), P_m is the monthly rainfall (mm/month), and n_r is the number of raindays per month (d/month). She demonstrated that, as a result of the Markov property of daily rainfall, the number of raindays in a month is a function of the monthly rainfall.

$$n_r = \frac{30p_{01}}{1 - p_{11} + p_{01}} \tag{4}$$

where p_{01} is the probability of a rainday after a dry day and p_{11} is the probability of a rainday after a rainday. These probabilities appear to be power functions of monthly rainfall. These power functions can be derived from time series analysis of selected raingauges with daily observations. De Groen (9) demonstrated that the coefficients of the power functions are representative for larger areas and can be interpolated between rainfall stations, similar to the regional distribution of monthly rainfall, which is surprising, because daily rainfall is known to have a very small spatial correlation (a few kilometer), but apparently the Markov property has a high one, of the same order of magnitude as monthly rainfall (100–1000 km).

Figure 2 shows the relative contributions of monthly interception (I) and transpiration (T) to total evaporation (E) in the Mupfure catchment. The Mupfure river, at Beatrice in Zimbabwe, drains a small catchment of 1215 km^2. It has an annual rainfall of about 800 mm/year. There is no substantial open water. The plots are based on monthly values over the period 1970–1979. The monthly interception has been computed by De Groen's formula with a threshold value of 4 mm/day. This threshold, may seem large but it caters to more than the difference between rainfall and throughfall (including wet surface evaporation) and there is generally more than one rainfall event per day, each of which contributes to evaporation from interception. Pitman (10) argues that interception in Southern Africa can be as much as 8 mm/day in forests. De Groen (8) considers 2–5 mm/day appropriate, depending on the land use, but the true value is not exactly known. The use of a higher or lower value of D changes the partitioning between I and T. The transpiration is computed as the difference between the interception and the total evaporation, which was determined on the basis of a monthly water balance. As a result, Fig. 2 gives an impression of how interception and transpiration relate and their order of magnitude, for different amounts of monthly rainfall.

As an aside, in Fig. 2, one can distinguish scatter dots and dots that appear to follow a pattern (the drawn lines). The scatter dots that occur in the low rainfall months are the result of transpiration from soil moisture stocks built-up during the wet season, which are accessed by deeply rooted vegetation. The scatter dots in the months with high rainfall are caused by limited solar radiation constraining evaporation. The drawn curves represent the contributions of interception and transpiration to the fast evaporation, occurring within 1 month, i.e., interception and fast transpiration.

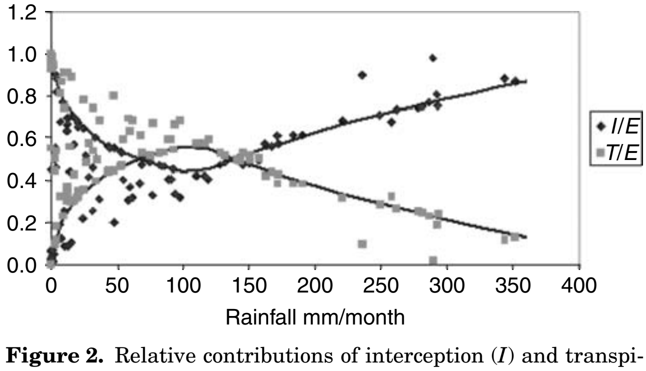

Figure 2. Relative contributions of interception (I) and transpiration (T) to monthly evaporation (E) in the Mupfure catchment in Zimbabwe, with $D_d = 4$ mm/day.

---

[1] The Markov property implies that the probability that a certain day is a rainday depends purely on whether the previous day was a wet or a dry day.

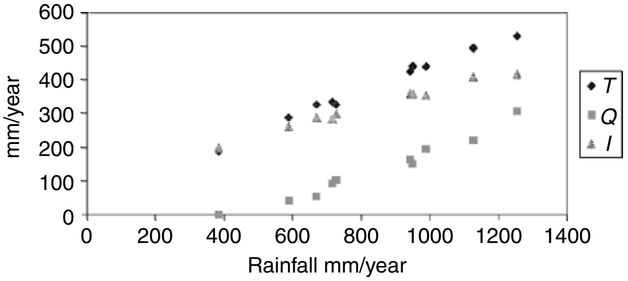

Figure 3. Annual rainfall partitioning into runoff (Q), interception (I), and transpiration (T), Mupfure basin at Beatrice, with $D_d = 4$ mm/day.

It can be seen that, overall in the Mupfure catchment, evaporation from interception is as significant as transpiration, and that in the wet months, transpiration is dominated by fast transpiration (the drawn curve), as there is no moisture deficit and the soil moisture stock is being replenished. More importantly, interception is dominant in wet months. Fig. 3 shows the annual partitioning of rainfall into transpiration (T), runoff (Q), and interception (I), indicating that interception comprises a considerable amount of the annual rainfall, with a tendency for interception to dominate in very dry years.

Besides interception forming a considerable part of monthly evaporation, it is crucial for determining antecedent conditions in rainfall–runoff modeling. It is often said that during storm events, evaporation from interception is negligible as compared with other fluxes. This may be true during the event, which is of short duration, but it is untrue in relation to the build-up of the antecedent conditions. It is widely recognized that the success of event-based modeling depends to a large extent on the antecedent conditions, particularly the distribution of the soil moisture in the unsaturated soil. If we want to get that right, it is important to split the rainfall into the effective part that contributes to soil moisture (and hence the rainfall–runoff process) and the ineffective part that does not, i.e., interception.

Finally, evaporation from interception is the most important process in moisture recycling to support continental rainfall. Figure 2 illustrates this. It shows that in months with high rainfall (i.e., the wet season), interception is an important mechanism. In dry months, transpiration is often dominant, but in these months (typically after the wet season), there are hardly any rainfall generating mechanisms to benefit from this feedback. Shuttleworth (2) observed that half the evaporation from interception occurs during the storm, which provides instant moisture feedback. Hence the moisture feedback to the atmosphere, which is such an important mechanism to support continental rainfall in the Sahel, and the Amazon (11,12), relies primarily on interception. The reason why transpiration is relatively small in wet months is because part of the energy available for evaporation is consumed by the interception process, which precedes transpiration, and because solar radiation is inhibited by clouds in wet months.

CONSEQUENCES OF UNDERESTIMATING INTERCEPTION

The picture of the Mupfure basin may not be representative for other climatic regions, but it serves to make a point. Interception is an important mechanism that cannot be simply neglected or lumped with other evaporation mechanisms. Disregarding or underestimating interception can lead to serious modeling mistakes, particularly when one uses automated calibration techniques. If interception is modeled incorrectly, the error will be compensated by other parameters, to satisfy the goodness-of-fit criterion. If a range of interception thresholds is tested, then one may find a correlation between interception and other soil and groundwater parameters. Clearly such relations are spurious.

The most common mistake of lumping interception with transpiration leads to an overdimensioning of the soil moisture stock, which can be seen easily. If the interception is forced through the transpiration process, a correct representation of the total flux and time scale in the model can only be achieved if the soil moisture stock is overdimensioned. If transpiration and interception are of the same order of magnitude, then the modeled soil moisture stock should be double its "real" value.

BIBLIOGRAPHY

1. Savenije, H.H.G. (2004). The importance of interception and why we should delete the term evapotranspiration from our vocabulary. *Hydrol. Processes* **18**(8): 1507–1511.
2. Shuttleworth, W.J. (1993). Evaporation. In: *Handbook of Hydrology*. D.R. Maidment (Ed.). McGraw-Hill, New York, Chapter 4.
3. Dooge, J.C.I. (1973). *Linear Theory of Hydrologic Systems*. Agriculture Research Service Technical Bulletin No. 1468, U.S. Department of Agriculture.
4. Dooge, J.C.I. (2003). Linear theory of hydrologic systems. *EGU Reprint Series*, 1, European Geosciences Union, Katlenburg-Lindau, Germany.
5. Beven, K.J. (2001). *Rainfall-runoff Modelling, the Primer*. John Wiley & Sons, New York.
6. Calder, I.R. (1990). *Evaporation in the Uplands*. John Wiley & Sons, New York.
7. Rutter, A.J., Kershaw, K.A., Robins, P.C., and Morton, A.J. (1971). A predictive model of rainfall interception in forests. I. Derivation of the model from observations in a plantation of Corsican pine. *Agricultural Meteorol.* **9**: 267–384.
8. De Groen, M.M. (2002). *Modelling Interception and Transpiration at Monthly Time Steps; Introducing Daily Variability Through Markov Chains*. Ph.D. dissertation, IHE-Delft. Swets & Zeitlinger, Lisse, The Netherlands.
9. De Groen. (1992).
10. Pitman, W.V. (1973). *A Mathematical Model for Generating Monthly River Flows from Meteorological Data in Southern Africa*. University of Witwatersrand, Department of Civil Engineering, Hydrological Unit, South Africa.
11. Eltahir, E.A.B. and Bras, R.L. (1994). Precipitation recycling in the Amazon basin. *Q. J. Roy. Meteorol. Soc.* **120**: 861–880.
12. Savenije, H.H.G. (1995). New definitions for moisture recycling and the relation with land-use changes in the Sahel. *J. Hydrol.* **167**: 57–78.

KINEMATIC SHOCK

Vijay P. Singh
Louisiana State University
Baton Rouge, Louisiana

Kinematic shocks form due to a variety of factors, including spatial and temporal variability of inflow, systems geometry, and initial and boundary conditions. This study provides a short overview of the formation, propagation, tracing, and fitting of shocks in hydrologic modeling, in particular, referring to rainfall–runoff modeling.

INTRODUCTION

The kinematic shock is a discontinuity representing a sudden rise or surge in flow depth. For example, flood waves have an intrinsic tendency to steepen as they propagate downstream due to lateral inflow or geometric constriction and eventually form a shock. The discontinuity can also represent a sudden decrease in flow depth. For example, when flow advances over a porous bed, there comes a time when the flow rate is not sufficient to satisfy the infiltrating demand of the bed and is, therefore, suddenly absorbed by the bed. This phenomenon is observed in ephemeral streams and irrigation borders and furrows. Discontinuities in flow are also observed in geologic formations due to fracture, fissures, and sudden changes in porous media properties. Shocks are also observed in landscape evolution and formation of deltas, river bed profiles, and river junctions.

Documented evidence of the occurrence of shocks is hard to find, but shocks have been observed in numerical solutions (1). Shocks lead to physically unrealistic multiple-valued and nonconvergent solutions (2). Therefore, proper consideration of shocks in hydrologic and hydraulic modeling is needed. There are four aspects of shocks that need consideration: (1) formation, (2) propagation, (3) tracing of the shock path, and (4) shock fitting. From a numerical standpoint, two main issues need to be addressed: (1) the physical relevance of kinematic shocks and (2) the adequacy of numerical schemes in the presence of shocks. In many kinematic cascade models, numerical schemes, such as the Lax–Wendroff scheme, do not exhibit multiple-valued solutions, and therefore, the presence of shocks goes unnoticed. This may partly be due to smoothing out of shocks by the numerical schemes. For many engineering problems, this may be acceptable if shocks are small and of short duration. However, these schemes may yield unrealistic or undesirable multiple hydrographs. Then, kinematic shocks warrant proper consideration.

SHOCK FORMATION

A range of factors leads to shock formation. These factors can be classified into three groups. In the first group are initial and boundary conditions. Lateral inflow and outflow form the second group. System geometric characteristics, such as surface roughness, geometric shape, and surface slope, are in the third group. It is, however, difficult to single out the contribution of each individual factor. Ponce and Windingland (3) attempted to determine flow and channel characteristics that either tend to promote or inhibit development of a kinematic shock. The flow and channel characteristics considered by them were the inflow hydrograph peak, the Froude number, the time to peak, the base-to-peak-flow ratio, the kinematic roughness parameter, and the kinematic exponent. The inflow hydrograph was represented by a gamma distribution. Of all the parameters, the size of the wave was the most important in the occurrence or nonoccurrence of the shock. They concluded that a kinematic shock is most likely to occur when there are a low base-to-peak-flow ratio, a high Froude number, and a wide and sufficiently long channel. A flood wave traveling in a steep, initially dry channel would be a potential candidate for shock formation.

ORDER OF DISCONTINUITY

Two kinds of discontinuities occur in hydrology. A discontinuity of the first kind (or first order) represents a sudden, finite change in flow depth. In that case, the differential equation of continuity no longer applies. The partial derivatives cease to be defined so that some additional condition is needed. This condition may be derived using the theory of weak solutions replacing the differential equation by an integral equation or by multiplying the conservation of mass to a frame comprising the shock (4). Both methods yield the same result in the form of a jump condition. According to the second method, the differential continuity equation is replaced by an equation stating that the flow into one side of the shock equals the flow out on the other side. Mathematically, this can be expressed as

$$U = \frac{Q_2 - Q_1}{h_2 - h_1} = \frac{\Delta Q}{\Delta h} \tag{1}$$

where subscripts 1 and 2 denote the variables calculated in front of and behind the moving shock, U is the shock velocity which must be between the characteristic velocities on either side of the shock (5), h is the flow depth, and Q is the flow discharge.

A discontinuity of the second kind (or second order) is a small change in flow depth h. If $h_2 - h_1 = dh$ is small, the expression for U reduces to

$$U = \frac{dQ}{dh} = c(h) \tag{2}$$

where c is the velocity of the discontinuity between flow depths h and $h + dh$. A small change in flow depth, dh, if maintained, is propagated through a flow depth h at velocity U. A line of constant depth in the $x - t$ diagram, therefore, describes the motion of the boundary between flows of depth h and $h + dh$, such that its slope is necessarily equal to the flow wave celerity, c.

For flows where the depth (h) increases downstream, the condition for forming a discontinuity of the first kind can be expressed in the following equivalent terms: (1) The lines of constant depth in the x–t diagram, if continued away from the t axis or x axis, would intersect; (2) the wave celerity, c, increases with depth of flow; and (3) the $Q - h$ curve is concave to the h axis. If these conditions are not satisfied, a first-order discontinuity is not formed. This can be illustrated in terms of second-order discontinuity. If c increases with h, small depth changes in higher flow regions upstream will move faster downstream than those in lower order flow regions and overtake them. This means that the depth gradient increases until a first-order discontinuity is formed. If c decreases with h, the reverse takes place; the depth gradient decreases, and any discontinuity is dispersed.

Three types of first-order shocks can be identified. To that end, we consider an x–t diagram which has two types of characteristics: A and B, as shown in Fig. 1. The A characteristics originate from the x axis and therefore depend on the initial condition. The B characteristics originate from the t axis and therefore depend on the boundary condition. These two types of characteristics are separated by a bounding or limiting characteristic that originates from the origin. The first kind of shock (or discontinuity) forms when two A characteristics intersect. This kind of shock does not occur in overland flow on an initially dry plane where effective rainfall depends only on time. The second kind of shock is designated as mixed shock and is formed by the intersection of one A characteristic and one B characteristic. Clearly, the limiting characteristic plays a major role in this case. The third kind of shock is boundary dependent and forms when two characteristics intersect.

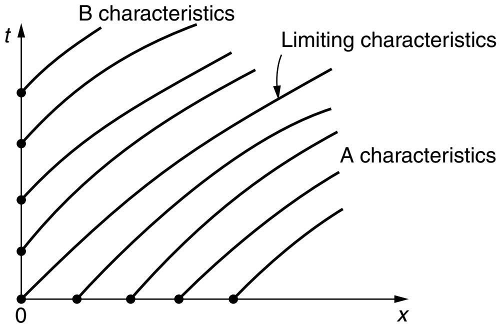

Figure 1. Characteristics of a plane in the (x, t) space.

It may be interesting to consider some possible cases of boundary-dependent shock formation. When rainfall increases with time and c increases with h, a shock will be formed by the intersection of the B characteristics, as shown in Fig. 2a. This can be prevented only by stopping it at the discontinuity. The shock is small and is fed by the characteristics running into it. Another example is when the initial region of varying h is small and there is effectively an initial discontinuity that is propagated along the line of discontinuity, as shown in Fig. 2b. Again, the discontinuity is small and is fed by lines running into it.

Another case is a sudden rise in rainfall at a given time within a small region of rapid change. A shock is

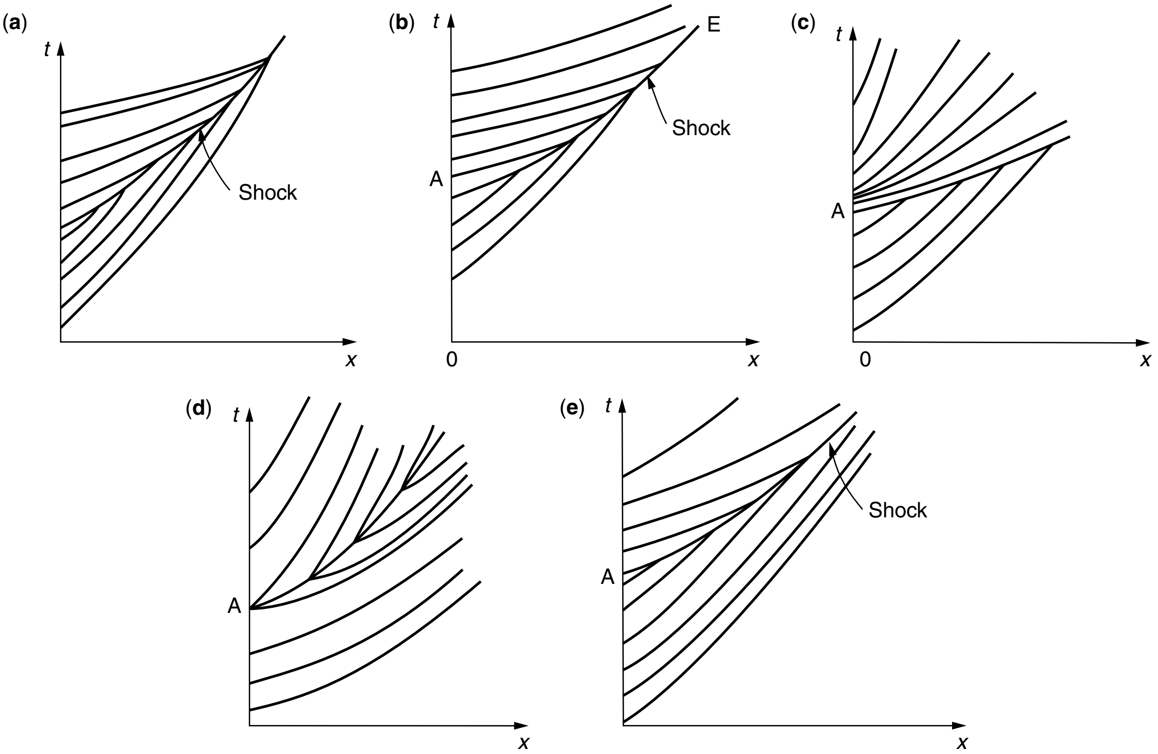

Figure 2. (**a**) Rainfall increases continuously with t; (**b**) rainfall increases over a short time; (**c**) rainfall suddenly increases at A in the limit; (**d**) rainfall increases, but without limit; and (**e**) rainfall increases to a limit and then decreases.

formed and occupies a region. It is fed by characteristics running into its lower boundary, as shown in Fig. 2c. This assumption is waived in Fig. 2d. Still another case occurs when rainfall increases to a limit and then decreases. In this case also, the shock occupies a region, as shown in Fig. 2e.

SHOCKS IN PLANAR FLOWS

For constant effective rainfall, r, the B characteristics are given by (6)

$$h(t, t_0) = h_1(t_0) + r(t - t_0) \tag{3}$$

$$x(t, t_0) = \frac{\alpha}{r}[h_1(t_0) + r(t - t_0)]^n - \frac{\alpha}{r}h_1^n(t_0) \tag{4}$$

where $t_0 (0 \leq t_0 \leq \infty)$ is the time when a B characteristic intersects the t axis, h_1 is the flow depth at the boundary $x = 0, n$ is the kinematic flow exponent, and α is the kinematic roughness coefficient. A shock may occur if $\partial x(t, t_0)/\partial t_0 > 0$, and the derivative can be obtained by differentiating Eq. 4. The region of intersecting characteristics can be obtained by taking the derivative of Eq. 4 with respect to t_0 and equating them to zero. Then, these two derivative equations can be solved to yield the coordinates of points on the characteristic envelope. The smallest values of these coordinates define the first point in the region of B characteristics where the formation of the kinematic shock is initiated, as shown in Fig. 3.

SHOCKS IN A KINEMATIC CASCADE

The mathematical properties of a kinematic cascade were investigated by Kibler and Woolhiser (1,7). A cascade is a representation of a watershed or a portion thereof by a series of planes discharging into a channel. Thus, the outflow of one plane forms the upper boundary condition of flow for the next. They developed a criterion, based on the properties of adjacent flow-plane pairs, to predict when a shock occurs in a cascade. They modified their Lax–Wendroff solution by placing instantaneous jumps in the output hydrograph corresponding to the arrival of shocks. They found that the Lax–Wendroff scheme effectively smoothed out shocks but was adequate when corrected for known shocks for solving the kinematic cascade. Schmid (8) derived a generalized shock formation criterion for infiltrating cascades that also encompasses the cascades with time-dependent rates of effective rainfall. On an infiltrating surface, the time to ponding marks the onset of overland flow. Ponding time can be computed by using the hydraulic properties of soils, and it strongly influences shock formation. For an infiltrating cascade, the ratio of respective ponding times of two adjacent, infiltrating planes is of major importance.

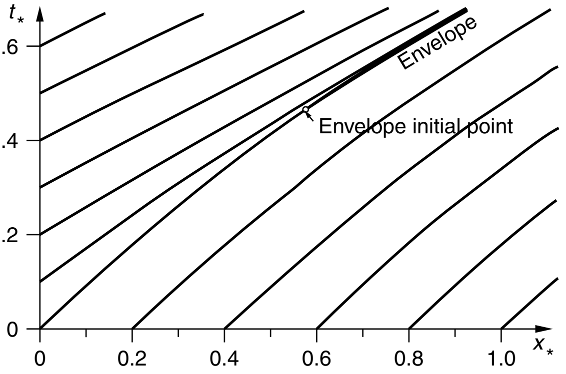

Figure 3. Intersection of characteristics in the (x, t) plane.

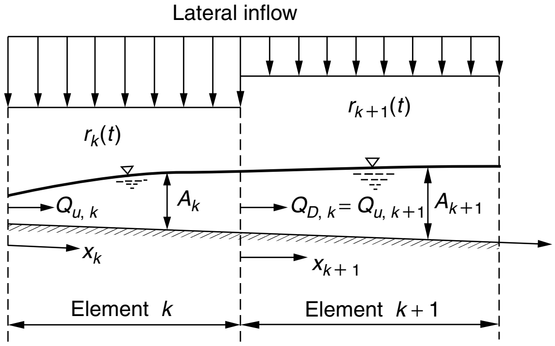

Figure 4. Two geometric elements in series, each has spatially uniform lateral inflow and topographic characteristics.

In a kinematic cascade, the criterion for shock formation is when the time to ponding of the upper plane is less than the time to ponding of the lower plane. If a cascade has two planes and overland flow starts on the upper plane rather than on the lower plane, a shock will form as soon as the rate of effective rainfall on the upper plane becomes nonzero, as shown in Fig. 4. Shocks can also occur when the time to ponding on the upper plane is greater than that on the lower plane (6).

SHOCKS ON CURVED SURFACES

A curved surface, where the slope continuously varies, can be considered a generalization of the kinematic cascade. Harsine and Parlange (9) dealt with kinematic shocks on curved surfaces. If the kinematic roughness coefficient varies in space or is a function of x, then one condition for the shock to occur is when $d\alpha_0/dx_0$ is a maximum. When $x = x_0$, the points where the B characteristics originate from the x axis, at that point α is α_0. Similarly, the effect of spatial variations on other surface geometric features can be derived.

SHOCK FITTING

Due to the spatial and temporal variability of rainfall and watershed characteristics, innumerable shocks may develop when the kinematic wave theory is applied to rainfall–runoff modeling, flow through unsaturated media, flow over porous beds, and dam break modeling. Because shock formation is intrinsic to the theory, approximate methods of shock fitting have been developed and are employed. In general, shock fitting entails (1) determining the characteristics intersecting each other on the shock path, (2) locating the points of intersection or determining the shock path, and (3) shock fitting.

It is assumed that a shock or discontinuity in h occurs first at the initial point on the characteristic envelope, and then the location of this moving shock is traced in the $x-t$ plane as the solution of the kinematic wave equations is generated. Across this discontinuity, however, the differential form of the continuity equation does not hold, but a control-volume form of this equation does, which leads to Eq. 1. In this equation, discharge is expressed by the kinematic depth–discharge relation. Thus,

$$\frac{dx_s}{dt} = \alpha \frac{h_2^n - h_1^n}{h_2 - h_1} \tag{5}$$

where x is the x coordinate of the discontinuity. The path of this moving shock in the x–t plane results from the solution of Eq. 5:

$$x_s(t) = x_0 + \alpha \int_{t_0}^{t} \frac{h_2^n - h_1^n}{h_2 - h_1} \tag{6}$$

where (x_0, t_0) are the coordinates of the initial point on the envelope. An iterative numerical solution of Eq. 6 is straightforward.

Borah et al. (10) developed an approximate procedure for shock fitting in unsteady-state flow routing. This method avoids finding the exact location of the shock origin by discretizing the upstream hydrograph, and it preserves the effect of shocks by routing them as they appear. It permits the use of the same numerical scheme to route both characteristic and shock waves. The procedure is particularly efficient when the initial condition represents a dry or uniform flow condition and only the outflow hydrograph is desired.

BIBLIOGRAPHY

1. Kibler, D.F. and Woolhiser, D.A. (1972). Mathematical properties of the kinematic cascade. *J. Hydrology* **15**: 131–147.
2. Croley, T.E. and Hunt, B. (1981). Multiple-valued and nonconvergent solutions in kinematic cascade models. *J. Hydrology* **49**: 121–138.
3. Ponce, V.M. and Windingland, B. (1985). Kinematic shock: sensitivity analysis. *J. Hydraul. Eng.* **111**(4): 600–611.
4. Whitham, G.B. (1974). *Linear and Nonlinear Waves*. Wiley Interscience, New York.
5. Lax, P.D. and Wendroff, B. (1960). Systems of conservation laws. *Commun. Pure Appl. Math.* **13**: 217–237.
6. Singh, V.P. (1996). *Kinematic Wave Modeling in Water Resources: Surface Water Hydrology*. John Wiley & Sons, New York.
7. Kibler, D.F. and Woolhiser, D.A. (1970). *The Kinematic Cascade as a Hydrologic Model*. Hydrology Paper No. 30, Colorado State University, Fort Collins, CO.
8. Schmid, B.H. (1990). On kinematic cascades: derivation of generalized shock criterion. *J. Hydraul. Res.* **28**(3): 331–340.
9. Hairsine, S.Y. and Parlange, J.Y. (1987). Kinematic shock waves on curved surfaces and application to the cascade approximation. *J. Hydrology* **87**: 187–200.
10. Borah, D.K., Prasad, S.N., and Alonso, C.V. (1980). Kinematic wave routing incorporating shock fitting. *Water Resour. Res.* **16**(3): 529–541.

KINEMATIC WAVE METHOD FOR STORM DRAINAGE DESIGN

TOMMY S.W. WONG
Nanyang Technological University
Singapore

INTRODUCTION

Ever since Lighthill and Whitham showed that the main body of a natural flood wave moves as a kinematic wave, there has been continual interest in applying the kinematic wave theory to drainage design. A drainage design method based on the kinematic wave theory has several inherent advantages: physically based parameters, facility for hydrograph prediction, and feasibility of obtaining analytical solutions for certain idealized basin and channel conditions. In this article, based on the theory, a drainage design method is developed for estimating the design discharge for a series of overland planes and for a V-shaped basin.

APPLICABILITY OF KINEMATIC WAVE THEORY

The applicability of the kinematic wave theory to overland and open channel flow of with sufficient accuracy compared to the solution from the Saint-Venant equations has been investigated by several researchers (1–3). For overland flow, the applicability of the theory can be defined by the Morris and Woolhiser (13) criterion,

$$kF_o \geq 5 \tag{1}$$

where k is the kinematic flow number and F_o is the Froude number at the end of the plane at equilibrium. The parameter kF_o can be related to the physical characteristics of an overland plane and the rainfall intensity as follows (4, p. 66):

$$kF_o = 8586 \frac{S_o^{1.3} L_o^{0.4}}{n_o^{0.6} i_n^{0.6}} \tag{2}$$

where S_o is the slope of the overland plane, L_o is the length of the overland plane, n_o is Manning's roughness coefficient of a plane surface, and i_n is the net rainfall intensity. The units are m m^{-1} for S_o, m for L_o, and mm h^{-1} for i_n.

For open channel flow, the applicability of the theory can be defined by the Ponce et al. (2) criterion,

$$\tau = \frac{TS_c u_c}{y_c} > 83 \tag{3}$$

where τ is the dimensionless wave period, T is the wave period that can be taken as twice the time-of-rise of the flood wave (5), S_c is the channel bed slope, u_c is the steady-state, uniform, channel flow mean velocity, and y_c is the steady-state uniform flow depth in the channel. The units are s for T, m m^{-1} for S_c, m s^{-1} for u_c and m for y_c. As a rule of thumb, the American Society of Civil Engineers (6, p. 58; 7, p. 87) simplified the criterion to

$$S_c > 0.002 \tag{4}$$

In general, the theory is applicable to overland flow and most open channel flow where the backwater effect is not significant (8, p. 42).

RAINFALL INTENSITY–DURATION RELATIONSHIP

To estimate the design discharge of a desired recurrence interval, the rainfall intensity–duration curve of the same recurrence interval is used. Analyses of the total rainfall curves show that, for a given recurrence interval, the rainfall intensity varies inversely with the rainfall duration and can be mathematically described by (9; 10, p. 71)

$$i = a/(c + t_r)^b \tag{5}$$

where i is the rainfall intensity, t_r is the rainfall duration; and a, b, and c are constants. To facilitate the derivation of an explicit expression for design discharge, Eq. 5 is reduced to (11, p. 20)

$$i = a{t_r}^{-b} \tag{6}$$

Using Eq. 6 with a single set of a and b values cannot fit the entire rainfall intensity-duration curve, but Chen and Evans (12) showed that by dividing the rainfall curve into segments, it is possible to fit the entire rainfall curve using different values of a and b for each segment.

Further, the effect of abstraction losses such as initial and infiltration losses can be incorporated into Equation 6 by substituting the net intensity, i_n, for rainfall intensity, i, and the net duration, t_n, for rainfall duration, t_r:

$$i_n = a_n {t_n}^{-b_n} \tag{7}$$

where a_n and b_n are constant within one segment of the rainfall curve but vary from segment to segment. The relationship between the net intensity, i_n, the rainfall intensity, i, and the rate of infiltration, f, is

$$i_n = i - f \tag{8}$$

and the relationship between the net duration, t_n, the duration, t_r, and the initial loss, I, is

$$t_n = t_r - \left(\frac{60I}{i}\right) \tag{9}$$

The units for Eqs. 6–11 are mm h^{-1} for i, i_n, and f; min for t_r and t_n; and mm for I. Typical values of initial loss and uniform infiltration rates for various surface types are reported by Stephenson and Meadows (4, p. 57).

TIME OF CONCENTRATION FOR A SERIES OF PLANES

Based on the kinematic wave theory, Wong (13) showed that the time of concentration, t_o, for a series of planes, is

$$t_o = \sum_{j=1}^{N} 7L_{oj} \times \left(\frac{n_{oj}}{\sqrt{S_{oj}}}\right)^{3/5} \left\{ \frac{\left[\sum_{r=1}^{j}(i_{nr}L_{or})\right]^{3/5} - \left[\sum_{r=1}^{j-1}(i_{nr}L_{or})\right]^{3/5}}{\sum_{r=1}^{j}(i_{nr}L_{or}) - \sum_{r=1}^{j-1}(i_{nr}L_{or})} \right\} \tag{10}$$

where N is the number of planes, j is the jth plane under consideration in the direction of flow for calculating the time of concentration, and r is the rth plane under consideration in the direction of flow for calculating the equilibrium inflow and outflow of the jth plane. The units are min for t_o, m for L_o, m m^{-1} for S_o, and mm h^{-1} for i_n. Equation 10 is applicable to a series of planes in which each plane has uniform slope and roughness, is subject to uniform net rainfall intensity, and has a turbulent or nearly turbulent flow regime. Within this condition, it can be applied to a cascade of planes, to planes of different roughness, to planes of different soil types resulting in different net intensities, to planes subject to different rainfall intensities, and to planes that combine all these variables. In the application of Eq. 10, the values of n_o may be selected from the recommended values of Engman (14), the Institution of Engineers, Australia (15, p. 300), or the American Society of Civil Engineers and Water Environment Federation (10, p. 88).

If the net intensity is the same for all planes, Eq. 10 reduces to

$$t_o = \sum_{j=1}^{N} \frac{7}{{i_n}^{2/5}} \left(\frac{n_{oj}}{\sqrt{S_{oj}}}\right)^{3/5} \left[\left(\sum_{r=1}^{j} L_{or}\right)^{3/5} - \left(\sum_{r=1}^{j-1} L_{or}\right)^{3/5} \right] \tag{11}$$

If there is only one plane, both Eqs. 10 and 11 reduce to (1)

$$t_o = \frac{7}{{i_n}^{2/5}} \left(\frac{n_o L_o}{\sqrt{S_o}}\right)^{3/5} \tag{12}$$

DESIGN DISCHARGE FOR A SERIES OF PLANES

For a homogenous, rectangular plane, Stephenson and Meadows (4, pp. 49–51) and Chen and Wong (16) showed that the maximum peak discharge for the plane may occur under the condition of full-area or partial-area contribution. Chen and Wong (17) and Wong (18) further showed if the maximum discharge occurs under the condition of partial-area contribution, the maximum discharge is related to the rainfall by $b_n = 1$ in Equation 7. However, this condition may not apply to a series of planes that are heterogeneous. A trial and error process is required to determine the maximum discharge for the series of planes. The maximum discharge is the design discharge for the planes.

For planes subject to uniform net rainfall intensity and for full-area contribution, an explicit expression for the peak discharge can be derived by equating the time of concentration of the planes to the net rainfall duration. This expression can also be used to estimate the peak discharge from the downstream portion of the series of planes. For full-area contribution, the peak discharge per unit width of plane, q_o, is related to the design rainfall intensity, i_o, as follows:

$$q_o = \frac{i_o \sum_{j=1}^{N} L_{oj}}{3.6 \times 10^6} \tag{13}$$

The units are m^2 s^{-1} for q_o, mm h^{-1} for i_o, and m for L_o. Equating t_n Eq. 7 to t_o Eq. 11 to obtain the design rainfall intensity, i_o, and substituting it into Eq. 13 results in the following explicit expression for the peak discharge q_o:

$$q_o = \frac{\sum_{j=1}^{N} L_{oj}}{3.6 \times 10^6} \left\{ \frac{{a_o}^{1/b_o}}{\sum_{j=1}^{N} 7\left(\frac{n_{oj}}{\sqrt{S_{oj}}}\right)^{3/5} \times \left[\left(\sum_{r=1}^{j} L_{or}\right)^{3/5} - \left(\sum_{r=1}^{j-1} L_{or}\right)^{3/5}\right]} \right\}^{b_o/[1-(2b_o/5)]} \tag{14}$$

where a_o and b_o are the respective values of a_n and b_n in Eq. 7 for $i_n = i_o$ and $t_n = t_o$. The units are $\mathrm{m^2\,s^{-1}}$ for q_o, m for L_o, and $\mathrm{m\,m^{-1}}$ for S_o.

If there is only one plane, Eq. 14 reduces to

$$q_o = \frac{1}{3.6\times10^6} \left\{ \frac{{a_o}^{1/b_o} {L_o}^{(1-b_o)/b_o}}{7\left(\frac{n_o}{\sqrt{S_o}}\right)^{3/5}} \right\}^{b_o/[1-(2b_o/5)]} \tag{15}$$

TIME OF TRAVEL IN A CHANNEL

For a drainage basin that comprises overland planes and drainage channels, the time of concentration for the basin is the summation of the overland time of concentration and the time of travel in the channel (8, p. 90). Yen (19) asserted that these two flow components should be evaluated separately, as they are really two separate, sequential systems. Further, Kibler and Aron (20) showed that time of concentration methods that consider the two flow components separately give better estimates.

Based on the kinematic wave theory, Wong (21) and Wong and Zhou (22) showed that for channels that have a negligible backwater effect subject to uniform lateral inflow and constant upstream inflow, the times of travel in these channels for seven different cross sections are as follows:

1. Wide rectangular channel

$$t_t = 0.0167 \left(\frac{n_c}{\sqrt{S_c}}\right)^{3/5} L_c B^{2/5} \left(\frac{{Q_d}^{3/5} - {Q_u}^{3/5}}{Q_d - Q_u}\right) \tag{16}$$

2. Square channel

$$t_t = 0.0289 \left(\frac{n_c}{\sqrt{S_c}}\right)^{3/4} L_c \left(\frac{{Q_d}^{3/4} - Q_u^{3/4}}{Q_d - Q_u}\right) \tag{17}$$

3. Deep rectangular channel

$$t_t = 0.0265 \left(\frac{n_c}{\sqrt{S_c}}\right) \frac{L_c}{B^{2/3}} \tag{18}$$

4. Triangular channel

$$t_t = 0.0236 \left(\frac{n_c}{\sqrt{S_c}}\right)^{3/4} L_c \left(\frac{z^2+1}{z}\right)^{1/4} \times \left(\frac{{Q_d}^{3/4} - {Q_u}^{3/4}}{Q_d - Q_u}\right) \tag{19}$$

5. Vertical curb channel

$$t_t = 0.0198 \left(\frac{n_c}{\sqrt{S_c}}\right)^{3/4} L_c \left(\frac{1+\sqrt{z^2+1}}{\sqrt{z}}\right)^{1/2} \times \left(\frac{{Q_d}^{3/4} - {Q_u}^{3/4}}{Q_d - Q_u}\right) \tag{20}$$

6. Parabolic channel

$$t_t = 0.0272 \left(\frac{n_c}{\sqrt{S_c}}\right)^{9/13} L_c B^{2/13} \left(\frac{{Q_d}^{9/13} - {Q_u}^{9/13}}{Q_d - Q_u}\right) \tag{21}$$

7. Circular channel

$$t_t = 0.0290 \left(\frac{n_c}{\sqrt{S_c}}\right)^{4/5} \frac{L_c}{B^{2/15}} \left(\frac{{Q_d}^{4/5} - {Q_u}^{4/5}}{Q_d - Q_u}\right) \tag{22}$$

where t_t is the time of travel in the channel, n_c is Manning's roughness coefficient of the channel surface, S_c is the channel bed slope, L_c is the channel length, B is the dimension defined in Fig. 1, Q_d is the constant downstream

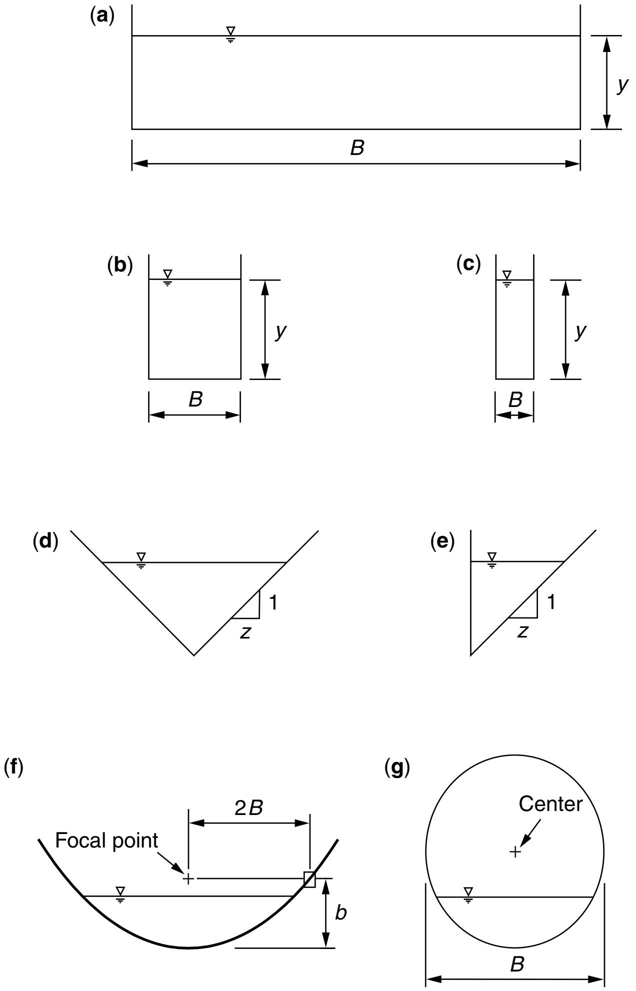

Figure 1. Channels of different cross section: (**a**) wide rectangular ($B \gg y$); (**b**) square ($B = y$); (**c**) deep rectangular ($B \ll y$); (**d**) triangular; (**e**) vertical curb; (**f**) parabolic; (**g**) circular.

outflow from the channel, and Q_u is the constant upstream inflow to the channel.

Further, Wong (23) showed that the time of travel in a general rectangular channel (i.e., a rectangular channel that has any flow depth) is

$$t_t = 0.0167\left(\frac{n_c}{\sqrt{S_c}}\right)^{3/4} L_c\left[\frac{(1+2\mu)^2}{\mu}\right]^{1/4}\left(\frac{{Q_d}^{3/4}-{Q_u}^{3/4}}{Q_d-Q_u}\right) \tag{23}$$

where μ is a parameter relating flow depth y to channel breadth B as follows:

$$\mu = \frac{y}{B} \tag{24}$$

Based on Manning's equation, the value of μ can be determined from

$$\left[\frac{\mu^5}{(1+2\mu)^2}\right]^{1/3} = \frac{n_c Q}{B^{8/3}\sqrt{S_c}} \tag{25}$$

where Q is the discharge in the channel. Equation 8 can be applied only to channels where the variation in μ is small. For channels that have a large variation in μ, the channel can be divided into longitudinal segments, and the time of travel is determined for each segment.

The units for Eqs. 16–25 are min for t_t; m m^{-1} for S_c; m for L_c, B, and y; and m^3 s^{-1} for Q_d, Q_u, and Q. In the application of Eqs. 16–23, the values of n_c may be selected from those recommended by Chow (24, pp. 110–113) or Acrement and Schneider (25).

DESIGN DISCHARGE FOR A V-SHAPED BASIN

Consider a V-shaped basin comprised of two identical overland planes and a drainage channel, as shown in Fig. 2. The two planes are subject to a uniform intensity i_n, and the runoff from the planes becomes the lateral inflow to the channel. The time of concentration for the basin, t_b, is the summation of the overland time of concentration, t_o, and the time of travel in channel, t_t, as follows (8, p. 90)

$$t_b = t_o + t_t \tag{26}$$

The overland time of concentration, t_o, is given by Eq. 12. Based on Eqs. 16–23, the times of travel in the channel, t_t, are as follows:

1. Wide rectangular channel

$$t_t = 7\left(\frac{B}{2i_n L_o}\right)^{2/5}\left(\frac{n_c L_c}{\sqrt{S_c}}\right)^{3/5} \tag{27}$$

2. Square channel

$$t_t = \frac{1.26}{(2i_n L_o)^{1/4}}\left(\frac{n_c L_c}{\sqrt{S_c}}\right)^{3/4} \tag{28}$$

3. Deep rectangular channel

$$t_t = \frac{0.0265}{B^{2/3}}\left(\frac{n_c L_c}{\sqrt{S_c}}\right) \tag{29}$$

4. Triangular channel

$$t_t = \frac{1.03}{(2i_n L_o)^{1/4}}\left(\frac{n_c L_c}{\sqrt{S_c}}\right)^{3/4}\left(\frac{z^2+1}{z}\right)^{1/4} \tag{30}$$

5. Vertical curb channel

$$t_t = \frac{0.863}{(2i_n L_o)^{1/4}}\left(\frac{n_c L_c}{\sqrt{S_c}}\right)^{3/4}\left(\frac{1+\sqrt{z^2+1}}{\sqrt{z}}\right)^{1/2} \tag{31}$$

6. Parabolic channel

$$t_t = \frac{2.84B^{2/13}}{(2i_n L_o)^{4/13}}\left(\frac{n_c L_c}{\sqrt{S_c}}\right)^{9/13} \tag{32}$$

7. Circular channel

$$t_t = \frac{0.594}{B^{2/15}(2i_n L_o)^{1/5}}\left(\frac{n_c L_c}{\sqrt{S_c}}\right)^{4/5} \tag{33}$$

8. General rectangular channel

$$t_t = \frac{0.726}{(2i_n L_o)^{1/4}}\left(\frac{n_c L_c}{\sqrt{S_c}}\right)^{3/4}\left[\frac{(1+2\mu)^2}{\mu}\right]^{1/4} \tag{34}$$

The units are min for t_t; mm h^{-1} for i_n; m for L_o, L_c, B, and y; and m m^{-1} for S_c.

For the V-shaped basin, the maximum peak discharge may occur under the condition of full-area or partial-area contribution. The maximum discharge is the design discharge for the basin. For planes subject to uniform net rainfall intensity and for full-area contribution, the design rainfall intensity, i_b, can be derived by equating the time of concentration of the basin, t_b, Eq. 26 to the net rainfall

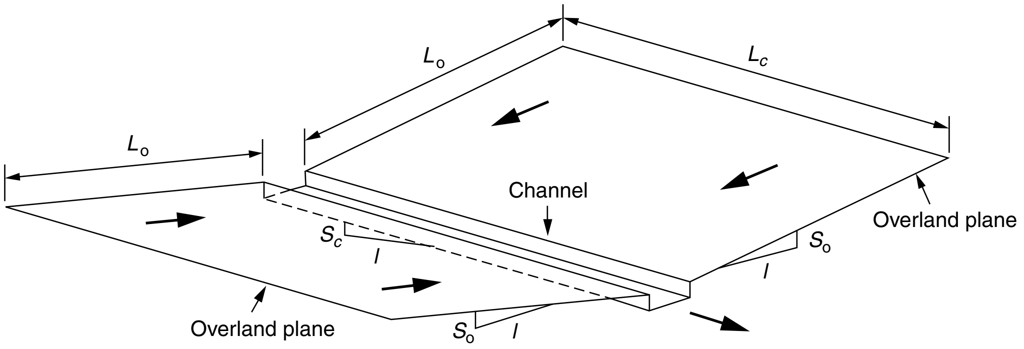

Figure 2. A V-shaped drainage basin.

duration, t_n Eq. 7. The peak discharge for the basin, Q_b, is then derived from i_b as follows:

$$Q_b = \frac{2 i_b L_o L_c}{3.6 \times 10^6} \tag{35}$$

The units are $\mathrm{m^3\ s^{-1}}$ for Q_b, $\mathrm{mm\ h^{-1}}$ for i_b, and m for L_o and L_c. Equation 35 can also be used to estimate the peak discharge under the partial-area contribution (i.e., only from the downstream portion of the basin).

Further, for a V-shaped basin that has a wide rectangular channel, an explicit expression for the peak discharge can be derived. The design rainfall intensity, i_b, can be obtained by substituting Eqs. 12 and 27 in Eq. 26 to obtain t_b and equating it to t_n Eq. 7. Substituting i_b in Eq. 35 results in the following explicit expression for the peak discharge Q_b:

$$Q_b = \frac{2L_oL_c}{3.6 \times 10^6} \left\{ \frac{{a_b}^{1/b_b}}{7\left[\begin{array}{c} \left(\dfrac{n_oL_o}{\sqrt{S_o}}\right)^{3/5} \\ + \left(\dfrac{B}{2L_o}\right)^{2/5} \left(\dfrac{n_cL_c}{\sqrt{S_c}}\right)^{3/5} \end{array} \right]} \right\}^{b_b/[1-(2b_b/5)]} \tag{36}$$

where a_b and b_b are the respective values of a_n and b_n in Eq. 7 for $i_n = i_b$ and $t_n = t_b$. The units are $\mathrm{m^3\ s^{-1}}$ for Q_b; m for L_o, L_c, and B; and $\mathrm{m\ m^{-1}}$ for S_o and S_c.

BIBLIOGRAPHY

1. Woolhiser, D.A. and Liggett, J.A. (1967). Unsteady one-dimensional flow over a plane—the rising hydrograph. *Water Resour. Res.* **3**(3): 753–771.
2. Ponce, V.P., Li, R.M., and Simons, D.B. (1978). Applicability of kinematic and diffusion models. *J. Hydraul. Div., ASCE* **104**(HY3): 353–360.
3. Morris, E.M. and Woolhiser, D.A. (1980). Unsteady one-dimensional flow over a plane: partial equilibrium and recession hydrographs. *Water Resour. Res.* **16**(2): 355–360.
4. Stephenson, D. and Meadows, M.E. (1986). *Kinematic Hydrology and Modelling*. Elsevier, New York.
5. Ponce, V.P. (1991). The kinematic wave controversy. *J. Hydraul. Eng., ASCE* **117**(4): 511–525.
6. American Society of Civil Engineers. (1996). River hydraulics. *Technical Engineering and Design Guides as Adapted from the U.S. Army Corps of Engineers No. 18*. American Society of Civil Engineers Press, New York.
7. American Society of Civil Engineers. (1997). Flood-runoff analysis. *Technical Engineering and Design Guides as Adapted from the U.S. Army Corps of Engineers No. 19*. American Society of Civil Engineers Press, New York.
8. Overton, D.E. and Meadows, M.E. (1976). *Stormwater Modeling*. Academic Press, New York.
9. Wenzel, H.G. (1982). Rainfall for urban stormwater design. In: *Urban Stormwater Hydrology*. D.F. Kibler (Ed.). Water Resources Monograph 7, American Geophysical Union, Washington DC, pp. 35–67.
10. American Society of Civil Engineers and Water Environment Federation. (1992). Design and construction of urban stormwater management systems. *ASCE Manuals and Reports of Engineering Practice No. 77 or WEF Manual of Practice FD-20*, American Society of Civil Engineers and Water Environment Federation, New York and Alexandria, VA.
11. Wong, T.S.W. (1992). *An Introduction to Kinematic Wave Method for Storm Drainage Design*. Hillview Publications, Singapore.
12. Chen, C.N. and Evans, R.R. (1977). Application of Kinematic Wave Method to Predict Overland Peak Flows. *Proceedings of International Symposium on Urban Hydrology, Hydraulics and Sediment Control*, University of Kentucky, Lexington, KY, pp. 113–118.
13. Wong, T.S.W. (1996). Time of concentration and peak discharge formulas for planes in series. *J. Irrig. Drainage Eng., ASCE* **122**(4): 256–258.
14. Engman, E.T. (1986). Roughness coefficients for routing surface runoff. *J. Irrig. Drainage Eng., ASCE* **112**(1): 39–53.
15. Institution of Engineers, Australia. (1987). *Australian Rainfall and Runoff*. D.H. Pilgrim (Ed.). Canberra, Australia.
16. Chen, C.N. and Wong, T.S.W. (1992). *Comparison of Peak Discharges from Full and Partial Areas of An Overland Plane*. International Symposium on Urban Stormwater Management, Sydney, Australia, pp. 94–97.
17. Chen, C.N. and Wong, T.S.W. (1993). Critical rainfall duration for maximum discharge from overland plane. *J. Hydraul. Eng., ASCE* **119**(11): 1040–1045.
18. Wong, T.S.W. (1996). Influence of partial-area effect on design discharge of homogeneous plane. *Hydrological Sci. Technol.* **12**(1–4): 83–98.
19. Yen, B.C. (1982). Some measures for evaluation and comparison of simulated models. Urban Stormwater Hydraulics and Hydrology. *Proceedings of 2nd International Conference on Urban Storm Drainage*, Urbana, Illinois, B.C. Yen (Ed.). Water Resources Publications, Littleton, CO, pp. 341–349.
20. Kibler, D.F. and Aron, G. (1983). *Evaluation of Tc methods for urban watersheds*. *Frontiers in Hydraulic Engineering: Proceedings of Cambridge Conference*, Massachusetts, H.T. Shen (Ed.). ASCE, New York, pp. 553–558.
21. Wong, T.S.W. (2001). Formulas for time of travel in channel with upstream inflow. *J. Hydrologic Eng., ASCE* **6**(5): 416–422.
22. Wong, T.S.W. and Zhou, M.C. (2003). Kinematic wave parameters and time of travel in circular channel revisited. *Adv. Water Resour.* **26**(4): 417–425.
23. Wong, T.S.W. (2002). Generalized formula for time of travel in rectangular channel. *J. Hydrologic Eng., ASCE* **7**(6): 445–448.
24. Chow, V.T. (1959) *Open Channel Hydraulics*. McGraw-Hill, New York.
25. Acrement, G.J. and Schneider, V.R. (1989). *Guide for Selecting Manning's Roughness Coefficients for Natural Channels and Flood Plains*. U.S. Geological Survey Water-Supply Paper 2339.

KINEMATIC WAVE AND DIFFUSION WAVE THEORIES

Vijay P. Singh
Louisiana State University
Baton Rouge, Louisiana

A wide range of phenomena, natural as well as man-made, in physical, chemical, and biological hydrology exhibit characteristics similar to those of kinematic or diffusion waves. The wide range of phenomena suggests that these waves are very pervasive. This study presents theories that are considered fundamental to advancing the state of the art of hydrology.

INTRODUCTION

The term "wave" implies a disturbance traveling upstream, downstream, or remaining stationary. We can visualize a water wave propagating where the water itself stays very much where it was before the wave was produced. We witness other waves that travel as well, such as heat waves, pressure waves, and sound waves. There is obviously the motion of matter, but there can also be the motion of form and other properties of matter.

Flow phenomena, according to the nature of particles composing them, can be classified into two categories: (1) flows of discrete noncoherent particles and (2) flows of continuous coherent particles. Examples of flows in the first category include traffic flow; transport of sand and gravel in pipes, flumes, and rivers; evolution of ripples, dunes, and bars on a river bed; settlement of sediment particles in a tank; and movement of microbial organisms. Exemplifying flows in the second category are overland flow, flood movement, baseflow, snowmelt, movement of glaciers, infiltration, evaporation, solute transport, ion exchange, and chromatographic transport. Because of their wave-like behavior, these flow phenomena can be described by the hydrodynamic wave theory or one of its variants, including the kinematic wave theory, diffusion wave theory, gravity wave theory, and linear wave theory.

A survey of geophysical literature reveals that the kinematic and diffusion wave theories have been applied to a wide spectrum of problems, including watershed runoff modeling, flood routing in rivers and channels, movement of soil moisture, macropore flow, subsurface storm flow, erosion and sediment transport, debris flow, solute transport, ion exchange, chromatography, sedimentation, glacial motion, movement of snowmelt water, flow over porous beds in irrigation borders and furrows, river-ice motion, vertical mixing of coarse particles in gravel-bed rivers, to name but a few. Other fields outside the realm of geophysics that apply the kinematic and diffusion wave theories include traffic flow, movement of agricultural grains in a bin, and blood circulation. It can be argued that within certain limitations a wide range of hydrologic phenomena can be approximated quite closely by kinematic and diffusion wave theories. Outside of these limitations, the theories provide only a crude approximation and the full hydrodynamic wave theory or one of its higher order variants may be needed.

Before presenting the theories, one may ask, What is meant by kinematic and diffusive? Although papers and textbooks in hydrology define kinematic and diffusive waves mostly from an approximation of the momentum equation and occasionally from defining flux laws, a comprehensive account of these terms is lacking in literature. Therefore, a short discussion to that end is in order.

WHAT IS KINEMATICS?

The term "kinematic" involves consideration of such macroscopic properties of a flow phenomenon as velocity, discharge, concentration, and time of travel. It does not involve consideration of forces and mass. Thus, the connotation of "kinematic" is that it is phenomenological and less physical, but its representation may be derived from or related to dynamic considerations involving forces, mass, momentum, and energy.

Although studies based on kinematics existed long before the development of Greek science, Ampere (1) is credited with introducing the term "kinematics" for the study of motion in his famous essay on the classification of sciences. His intent was to create a discipline for the study of motion without regard to the forces involved. Thus, kinematics is defined as the science that deals with the study of motion in the most general way. The motion can be translatory, dilatational, or rotational (angular) (2).

It may be instructive to view kinematics in the context of geometry and dynamics. From the standpoint of dimensions, geometry implies only length, whereas kinematics involves length and time, and dynamics considers length, mass, and time. Thus, kinematics is somewhere between geometry and dynamics, perhaps closer to the latter. Although kinematics is related to geometry, it can also be studied without geometry by means of numerical description of the functions as the Babylonians seem to have done it—a precursor to curve fitting. In kinematics, we are concerned with the position of a point at a certain time and also with the displacement accomplished and the distance traversed in a certain time interval. Other functions of time such as velocity and acceleration can be introduced. Even higher derivatives such as the rate of change of acceleration, called jerk, and even higher order derivatives can be included.

It should be noted that the introduction of mass does not necessarily make some motions strictly dynamic; they may remain strongly and functionally kinematic. For instance, the formulation of equations for conservation of volume in fluid mechanics is very much in the realm of kinematics. When the equations for conservation of mass are formulated, they are still in the same domain; the two treatments are very similar. Even the discussion of momentum flux and kinetic energy flux are fundamentally kinematic. In general, there are terms in the equations of dynamics that are essentially kinematic. The difference between kinematics and dynamics thus requires close scrutiny.

KINEMATIC WAVE THEORY

In a seminal contribution, in 1955, Lighthill and Whitham developed the kinematic wave theory, and they gave a full account of the theory for describing flood movement in long rivers and traffic flow on long crowded roads (3,4). The theory is comprised of three components: (1) the theory of kinematic continuous waves, (2) the theory of shock waves, and (3) the theory of the formation of shock waves from continuous waves. Iwagaki (5) developed a approximate method of the characteristics for routing steady-state flow in open channels of any cross-sectional shape that have nearly uniform lateral flow and proposed that the method would be applicable to the hydraulic analysis of runoff estimates in river basins. Implicit in his method was the kinematic wave assumption. Thus, Iwagaki (1955) can also be credited with independently conceiving of the kinematic wave concept and developing the method of kinematic wave routing.

Following Lighthill and Whitham (3,4), kinematic waves exist if it can be assumed with sufficient accuracy that there is a functional relationship between flux, concentration, and position. Thus, the kinematic wave theory is mathematically expressed by the law of conservation of mass through the continuity equation and a flux–concentration relation. The coupling of these two equations leads to a first-order partial differential equation that has only one system of wave characteristics. The essence of this is that the wave property is derived from the continuity equation alone. The governing equation (or continuity equation) is only of the first order, so kinematic waves possess only one system of characteristics. This would also imply that kinematic waves travel only in a downstream direction and that the kinematic wave theory cannot accommodate waves that travel in the upstream direction as in backwater flow. In the absence of any lateral inflow and/or outflow, the flux along the characteristics is constant, and the kinematic waves are nondispersive and nondiffusive. However, the waves are dispersive and diffusive when there is lateral inflow and/or outflow, as, for example, in rainfall runoff.

Continuous Waves

If the volume of a quantity passing a given point x in unit time is denoted by flux F and the concentration C is defined as the volume of the quantity per unit distance, then for a one-dimensional flow system without any lateral sources and sinks, the continuity equation or the mass conservation law is

$$\frac{\partial C}{\partial t} + \frac{\partial F}{\partial x} = 0 \tag{1}$$

Equation 1 is first order and states that the volume of the quantity in a small element of length changes at a rate equal to the difference between the rates of inflow and outflow.

Fundamental to the development of the kinematic wave theory is the development of a flux–concentration relation. A general flux–concentration relation can be assumed as

$$F = F(C, x) \tag{2}$$

Equations 1 and 2 have two unknowns, F and C, and can therefore be combined into one equation having one unknown. Multiplying Eq. 1 by

$$c = \left(\frac{\partial F}{\partial C}\right)_{x\,\text{constant}} = c(C, x) \tag{3}$$

one obtains

$$\frac{\partial F}{\partial t} + c\frac{\partial F}{\partial x} = 0 \tag{4}$$

Equation 4 is the kinematic wave equation. The quantity c has various names. In hydraulics, it is called wave velocity or celerity, and it is referred to as mobility in vadose zone hydrology. It is the slope of the flux–concentration relation at a fixed x. This is not the same as the mean velocity u which is expressed as

$$u = F/C \tag{5}$$

The relation between c and u is thus obvious:

$$c = \frac{d}{dC}(uC) = u + C\frac{du}{dC} \tag{6}$$

Equation 6 shows that c is greater then u, if u increases with C as in river flow; it is less than u if u decreases with C as in sedimentation; and it is equal to u if u does not change with C. Equation 4 states that F is constant in waves traveling past a point whose celerity c is given by Eq. 3. The kinematic wave equation 4 has one system of characteristics traveling only in the downstream direction, given by $dx = c\,dt$, and along each of these characteristics, F is fixed.

Although Eq. 3 constitutes the basic building block of the kinematic wave theory, a more popular, but more restrictive, derivation of the theory in hydrology is obtained by assuming that the local acceleration, convective acceleration, and pressure gradient (or concentration gradient) in the momentum equation are negligible. This assumption states an equivalence between frictional and gravitational forces. When only the local and convective acceleration terms are assumed negligible, the remainder of the momentum equation (i.e., the pressure gradient equaling the difference between gravity and frictional slopes) leads, in conjunction with the continuity equation, to diffusion wave theory. When the full momentum equation is used jointly with the continuity equation, the result is the dynamic wave theory. In general, the local acceleration and the convective acceleration are of the same order of magnitude but are opposite in sign, thus counteracting each other. In a wide range of problems in hydrology, the frictional and gravitational terms are dominant, and this is one of the reasons for the popularity of the kinematic wave theory.

Formation of Kinematic Shocks

A kinematic shock is a discontinuity representing a sudden rise or surge in flow depth. For example, during wave movement, faster moving waves overtake slower moving waves, and at a fixed position, there will be an increase in flux and concentration as functions of time, leading to shock formation. Thus, flood waves have an intrinsic, nonlinear tendency to steepen as they propagate downstream, eventually forming a shock. In the characteristic plane, shock formation results from the intersection of characteristics. After some time, the shock weakens and dissolves into a region of uniform flow. Shocks arise in a variety of hydrologic processes. A short but general discussion of shocks is relevant, however, and is given here. Shocks or discontinuities are of first or second order. A discontinuity of the first kind in the concentration is defined by a sudden change of concentration at a certain level. In such a case, the differential equation of continuity (Eq. 4) no longer applies;

instead, it is replaced by an equation stating that the flow of mass into one side of the shock is equal to that on the other side of the shock. Denoting the velocity of discontinuity by U, mass conservation leads to

$$U = \frac{F_2 - F_2}{C_2 - C_1} \tag{7}$$

where subscripts 1 and 2 denote the quantities ahead and behind the shock. Equation 7 shows that, in general, the discontinuity is not at rest but moves with velocity U which is the slope of the chord joining the points (F_1, C_1) and (F_2, C_2) on the F–C diagram.

A discontinuity of the second kind is represented by a very small change in concentration. If $C_2 - C_1 = dC$ is small, the expression for U becomes $dF/dC = c$, the celerity. In this case the velocity of a discontinuity between concentrations C and $C + dC$ is the same as the celerity c. A small change, if maintained, is propagated through the medium of concentration C at velocity c similarly to the propagation of sound through air at a definite velocity. A line of constant concentration, therefore, describes the motion of a boundary between media of concentrations C and $C + dC$ so that its slope is necessarily equal to c. The complete modification of concentration at a position as well as its profile can be characterized as a series of small discontinuities propagated through the medium.

Determination of Kinematic Shocks

Many factors affect the formation of shocks, and these factors can broadly be distinguished as of three types: (1) initial and boundary conditions, (2) lateral inflow and outflow, and (3) watershed geometric characteristics. A full account of the theory of shock formation was given by Lighthill and Whitham (3,4), and a mathematical treatment of the formation and decay of shocks was given by Lax (6).

Carrier and Pearson (7) developed a method for determining the enveloping curve for any region of intersecting characteristics. Croley and Hunt (8) analytically determined boundary-dependent shocks in planar flow. Hairsine and Parlange (9) dealt with kinematic shocks on curved surfaces. Kibler and Woolhiser (10) investigated geometry-dependent shocks using a kinematic cascade. They derived the mathematical properties of a kinematic cascade and developed a criterion, based on the properties of adjacent flow-plane pairs, to predict when a shock would occur in a cascade. They also developed a numerical procedure for shock fitting. General properties of shock waves along with continuous kinematic waves were discussed. Full equations of motion were employed to investigate the structure of the kinematic shock.

Building on the work of Kibler and Woolhiser (10), Schmid (11) investigated the effect of lateral inflow and outflow as well as geometry on the development of shocks. He derived a generalized criterion for shock formation in an infiltrating cascade that also encompasses cascades that have time-dependent rates of effective rainfall. Using simulated conditions, Ponce and Windingland (12) determined flow and channel characteristics that either tend to promote or inhibit development of kinematic shocks. Borah (13) and Borah et al. (14) developed approximate but efficient numerical methods for determining the shock path and shock fitting.

DIFFUSION WAVE THEORY

The diffusion wave theory is comprised of (1) the formation of continuous waves, (2) the formation of shocks, and (3) the propagation of shock waves. The diffusion wave formulation is obtained from the continuity equation and a diffusive flux law. In surface water hydrology, the diffusive flux law is obtained by neglecting the inertial terms in the momentum equation, whereas in groundwater and subsurface water hydrology, it is expressed by Darcy's law. When the two equations are combined, the result is a second-order partial differential equation. The second-order differential term is associated with a coefficient expressed in terms of what is referred to as diffusivity. This term introduces diffusion. The first-order spatial derivative and the coefficient associated with it together lead to steepening or breaking of waves. Clearly, theses two terms have opposite tendencies of nonlinear wave steepening and diffusion. The diffusivity term is always positive.

When a shock is formed, it can be discontinuous or continuous. If the diffusivity is small, the shock is a rapid but continuous increase in flow concentration over a narrow range. Breaking due to nonlinearity is balanced by diffusion in the narrow range to yield a steady-state profile. The sign of the derivative of the concentration plays an important role in determining whether the shock will break forward or backward.

FLUX LAWS

The term "flux" is defined in two ways in hydraulics. First, flux denotes a volume of any quantity per unit area per unit time. Thus, its dimensions are L/T if the quantity is, say, runoff. Second, flux is defined as any quantity per unit time. For example, volume flux is the volume of a quantity per unit time as exemplified by discharge, mass flux is mass per unit time, momentum flux (it has the same dimensions as force) is momentum per unit time, energy flux (kinetic energy flux has the same dimensions as power) is energy per unit time, and so on. Both definitions are used in water and environmental engineering.

Flux laws are fundamental to the development of transport theories. The flux laws most common in environmental and water sciences are of two types: (1) power laws and (2) gradient laws. Kinematic wave flux laws are the most popular flux laws of the power type, and Darcy's law and the Darcy–Buckingham law are the most popular gradient type flux laws. Generalized flux laws that specialize into power and gradient laws are also used. An example of a generalized flux law is Burgers' law or a generalized version thereof (15). Power type flux laws lead to kinematic wave equations, whereas the gradient type flux laws lead to diffusion waves governed by elliptical or parabolic partial differential equations.

Algebraic Laws

A general expression for the flux–concentration relation is given by Equation 2. One of its special forms is the popular kinematic wave flux law expressed as

$$F = \alpha C^n \tag{8}$$

where α is a parameter and n is an exponent. The meaning and interpretation of α and n may vary with the problem to which Eq. 8 is applied. Equation 8 leads to the famous Chezy and Manning equations which are popular in studies of overland flow and flood routing in open channels and rivers. At a fixed location x, $(\partial F/\partial C)$ defines the wave celerity $c : c = (\partial F/\partial C)|_{x\,\text{fixed}}$ which is not the same as the average velocity of flow. When $n = 1$, Eq. 8 becomes

$$F = \alpha C \tag{9}$$

This is a linear flux–concentration law used to describe the movement of meltwater runoff from snowpack.

Beven (16) employed two somewhat uncommon three-parameter forms of the flux–concentration relation for his channel network routing model:

$$F = \frac{C(a + bk) - k}{1 - bC} \tag{10}$$

and

$$\frac{F}{a(1 - \exp(-kF)) + bF} = C \tag{11}$$

where a, b, and k are parameters.

Another flux law is defined as

$$F = \alpha C^2(C_o - C) \tag{12}$$

where C_o is the maximum value of C and α is a parameter. Equation 12 is used to describe the settlement of particles in a dispersion. Still another flux law used to describe the movement of sediment in flumes and pipes is

$$F = u_0 C\left(1 - \frac{C}{C_0}\right) \tag{13}$$

where u_0 is the velocity of a single particle when there are no other particles in the flow. All of these flux laws lead to kinematic waves.

Gradient Laws

The gradient-type flux law is expressed as

$$F = G(\partial C/\partial x) \tag{14}$$

where G is some function. As an example,

$$F = \beta(\partial C/\partial x)^m \tag{15}$$

where β is a parameter and m is an exponent. Equation 15 is a nonlinear version of Darcy's law. When $m = 1$, Eq. 15 becomes

$$F = \beta(\partial C/\partial x) \tag{16}$$

Equation 16 is Darcy's law, where β describes the negative of the saturated hydraulic conductivity and C the hydraulic head.

A more general form of Eq. 16 arises when β depends on C:

$$F = \beta(C)\partial C/\partial x \tag{17}$$

Equation 17 is the Darcy–Buckingham law where C describes the hydraulic head and β the negative of the unsaturated hydraulic conductivity which varies with moisture content. When a gradient flux law is coupled with the continuity equation, the resulting partial differential equation turns out to be of parabolic type under transient flow conditions and of elliptical type under steady-state conditions.

Generalized Flux Laws

A generalized flux law is obtained by combining the power and gradient-type flux laws. In general,

$$F = G(C, \partial C/\partial x, x) \tag{18}$$

where G is some function and its argument is defined by C, $\partial C/\partial x$, and x. One form of Eq. 18 is

$$F = \alpha C^n + \beta(\partial C/\partial x)^m \tag{19}$$

where m and n are exponents. Equation 19 specializes into (15)

$$F = \alpha C^n + \beta(\partial C/\partial x) \tag{20}$$

and

$$F = \alpha C + \beta \partial C/\partial x \tag{21}$$

Equation 21 is Burgers' flux law (17) used in turbulence modeling; when used in the continuity equation, the resulting partial differential equation is Burgers' equation. Equations 19 and 20 have been employed to rout flows in open channels (15).

VALIDITY OF KINEMATIC AND DIFFUSION WAVE THEORIES

Lighthill and Whitham (3) have shown that, for Froude numbers below 1 (appropriate to flood waves), dynamic waves are rapidly attenuated and kinematic waves become dominant. Using a dimensionless form of the Saint-Venant (SV) equations, Woolhiser and Liggett (18) obtained the kinematic wave number, K, as a criterion for evaluating the adequacy of the kinematic wave (KW) approximation. For K greater than 20, the kinematic wave approximation was considered an accurate representation of the SV equations in modeling overland flow. Morris and Woolhiser (19) modified this criterion by explicitly including a Froude number corresponding to normal flow, F_O, and showed, based on numerical experimentation, that $F_O^2 K \geq 5$ is a better indicator of the adequacy of the kinematic wave approximation.

Using a linear perturbation analysis, Ponce and Simons (20) derived properties of the kinematic wave, diffusion wave, and dynamic wave representations in modeling open channel flow. Menendez and Norscini (21) extended the work of Ponce and Simons (20) by including the phase lag between the depth and velocity of flow. Based on the propagative characteristics of a sinusoidal perturbation, they derived criteria to evaluate the adequacy of kinematic wave and dynamic wave approximations. Daluz Vieira (22) compared solutions of the SV equations with those of the kinematic wave and dynamic wave approximations for a range of F_O and K and defined regions of validity of these approximations in the $K - F_O$ space. Fread (23) developed criteria for defining the application range of the kinematic wave and dynamic wave approximations. Ferrick (24) defined a group of dimensionless-scale parameters to establish the spectrum of river waves, using continuous transitions between wave types and subtypes.

It is possible to make inferences about the diffusion wave theory using the Vedernilkov number, N, which is the ratio of relative celerity of kinematic waves to the relative celerity of diffusive waves. If N is less than one, flow is stable, and if N is greater than one, it is unstable. Thus, $N = 1$ gives the condition of neutral stability. Clearly, a surface disturbance attenuates in stable flow, amplifies in unstable flow, and undergoes no change in neutral stable flow.

The comparative studies cited before, show that many cases satisfy the conditions for the validity of kinematic wave theory and that in surface water hydrology, kinematic waves dominate in such cases. Other wave types may exist but they are either short-lived or play a minor role. This is further elaborated by considering two special cases: (1) uniform, unsteady-state flow and (2) steady-state, nonuniform flow.

Unsteady-State Uniform Flow

Using simplified conditions, Singh (2,25–30) derived error equations specifying error as a function of time for the kinematic wave approximation of space-independent flows. He considered four different types of scenarios depending upon the presence of lateral inflow or rainfall and infiltration: (1) Lateral inflow is constant, and there is no infiltration; if there is, it is included in lateral inflow. (2) Both lateral inflow and infiltration are considered constant. (3) Both lateral inflow and infiltration are included, but their difference is zero. (4) There is no lateral inflow, but infiltration is included. These scenarios were analyzed under two types of initial conditions: (1) the plane or channel is initially wet, and (2) the plane or channel is initially dry. In all, 18 cases were analyzed.

For space-independent flow, the continuity equation takes the form,

$$\frac{dh}{dt} = i - f \tag{22}$$

and the momentum equation takes the form,

$$\frac{du}{dt} = g(S_0 - S_f) - \frac{iu}{h} \tag{23}$$

where g is the acceleration of gravity, i is the lateral inflow (or rainfall intensity), f is the infiltration rate, S_0 is the bed slope, and S_f is the slope of the energy line. For a kinematic wave approximation, Eq. 23 becomes

$$S_0 = S_f \tag{24}$$

which can be expressed as

$$u = \left(\frac{S_0}{\beta}\right)^{0.5} h^{0.5} \tag{25}$$

where β is a parameter. Singh (2,26,27,29,30) derived error equations for all cases which turned out to be Riccati equations of the form,

$$\frac{dE}{d\tau} = C_0(\tau) + C_1(\gamma,\tau)E + C_2(\gamma,\tau)E^2 \tag{26}$$

where E is the error defined as (kinematic wave solution-dynamic wave solution)/dynamic wave solution, τ is dimensionless time, γ is a dimensionless parameter, and $C_i, i = 1$, 2, and 3, are Riccati coefficients. The parameter γ is analogous to a kinematic wave parameter and is defined as

$$\gamma = \frac{4g^2\beta S_0 h_0}{{i_0}^2} \tag{27}$$

or

$$\gamma = 4\beta g S_0 \tag{28}$$

depending upon whether or not lateral inflow (i_0) is included. In general, the kinematic wave approximation is very good if γ is equal to or greater than 10. As time progresses, that is, $\tau \geqslant 10$, the kinematic wave approximation converges to the dynamic wave solution.

Steady-State Nonuniform Flow

Steady-state nonuniform flows are encountered in a variety of natural situations. In overland flow, the steady state is attained for constant rainfall after the flow depth at the outlet has reached equilibrium. This same is true for channel flow subject to constant lateral inflow. For a channel receiving a constant inflow of long duration at its upstream boundary, the flow at the downstream end would reach equilibrium. The steady-state solution aids in understanding the nature of the surface water profile. It may help define the condition for use of zero depth in place of zero influx at the upstream boundary. When the rainfall duration is much longer than the time of equilibrium, steady-state surface water profiles are very useful. Pearson (31) examined the criteria for using the kinematic wave approximation of the SV equations for shallow water flow. For steady-state one-dimensional flow over a plane, he derived a new criterion as $K \geqslant 3+5/F_O^2$, where K is the kinematic wave number and F_O is the Froude number corresponding to normal flow.

Govindaraju et al. (32,33) provided a comprehensive discussion of the DW theory for steady-state nonuniform flow. They presented both numerical and analytical results for flux-type downstream boundary conditions and using zero inflow at the upstream boundary. Parlange et al. (34) investigated errors in the KW and diffusion wave (DW) approximations by comparing their predictions with the numerical solution of the SV equations under steady-state conditions. They suggested splitting the solution into two regions, one near the downstream end of the plane and the other covering most of the plane. Singh and Aravamuthan (35–38) derived errors in the kinematic and diffusion wave approximations under four conditions: (1) zero flow at the upstream boundary, (2) finite depth at the upstream boundary, (3) critical flow depth at the downstream end, and (4) zero-depth gradient at the downstream boundary. For economy of space, only the KW case will be discussed. Depending on the inclusion of lateral inflow and infiltration, 21 cases were analyzed for the four different scenarios mentioned before. By comparing the kinematic wave solution with the dynamic wave solution, error equations were derived for all cases. For time-independent flow, the governing equations are

$$\frac{dh}{dx} = i - f \tag{29}$$

and

$$\frac{d}{dt}\left(\frac{1}{2}u^2 + gh\right) = g(S_o - S_f) - \frac{iu}{h} \tag{30}$$

The error equations were generalized Riccati type equations:

$$\frac{dE}{dx} = C_0E + C_1E^2 + C_2E^3 + C_3E^4 + C_4E^5 \tag{31}$$

where the parameters, $C_i, i = 1, 2, 3$, and 4, are nonlinear functions of x, h, K, and F_O which are as defined before. In most cases, it was found that the downstream boundary exercised significant influence on the adequacy of the kinematic wave approximation. Away from the boundaries, that is, $0.1 \leqslant x/L \leqslant 0.9$, where L is the length of the plane, the kinematic wave approximation was a good approximation (39).

BIBLIOGRAPHY

1. Ampere, A.M. (1834). *Essai Sur La Philosophie De Sciences, or Exposition Analytique D Une Classification Naturelle De Toutes Les Connaissances Humaines*. Bachelier, Paris, France.
2. Singh, V.P. (1996). *Kinematic Wave Modeling in Water Resources: Surface Water Hydrology*. John Wiley & Sons, New York.
3. Lighthill, M.J. and Whitham, G.B. (1955). On kinematic waves: 1. Flood movement in long rivers. *Proc. R. Soc., London, Series A* **229**: 281.
4. Lighthill, M.J. and Whitham, G.B. (1955). On kinematic waves: 2. A theory of traffic flow on long crowded roads. *Proc. R. Soc., London, Series A* **229**: 317–345.
5. Iwagaki, Y. (1955). *Fundamental studies on the runoff analysis by characteristics*. Bulletin 10, Disaster Prevention Research Institute, Kyoto University, Kyoto, Japan, p. 25.
6. Lax, P.D. (1972). The formation and decay of shock waves. *Am. Math. Mon.* **79**: 227–241.
7. Carrier, G.F. and Pearson, C.E. (1976). *Partial Differential Equations: Theory and Practice*. Academic Press, New York.
8. Croley, T.E. and Hunt, B. (1981). Multiple-valued and non-convergent solutions in kinematic cascade models. *J. Hydrology* **49**: 121–138.
9. Hairsine, S.Y. and Parlange, J.Y. (1987). Kinematic shock waves on curved surfaces and application to the cascade approximation. *Journal of Hydrology* **67**: 187–200.
10. Kibler, D.F. and Woolhiser, D.A. (1970). *The Kinematic Cascade as a Hydrologic Model*. Hydrology Paper 39, Colorado State University, Fort Collins, CO.
11. Schmid, B.H. (1990). On kinematic caseades: derivation of a generalized shock criterion. *Journal of Hydraulic Research* **28**(3): 331–340.
12. Ponce, V.M. and Windingland, B. (1985). Kinematic shock: sensitivity analysis. *J. Hydraul. Eng.* **111**(4): 600–611.
13. Borah, D.K. (1989). Runoff simulation model for small watersheds. *Trans ASAE* **32**(3): 881–886.
14. Borah, D.K., Prasad, S.N., and Alonso, C.V. (1980). Kinematic wave routing incorporating shock fitting. *Water Resour. Res.* **16**(3): 529–541.
15. Singh, V.P. and Prasana, M. (1999). A generalized flux law, with an application. *Hydrological Processes* **13**: 73–87.
16. Beven, K. (1979). On the generalized kinematic routing method. *Water Resour. Res.* **15**(5): 1238–1242.
17. Burgers, J.M. (1948). A mathematical model illustrating the theory of turbulence. In: *Advances in Applied Mechanics*. R. Von Mises and T. Von Korman (Eds.). Academic Press, New York, pp. 171–199.
18. Woolhiser, D.A. and Liggett, J.A. (1967). Unsteady one-dimensional flow over a plane: the rising hydrograph. *Water Resour. Res.* **3**(3): 753–771.
19. Morris, E.M. and Woolhiser, D.A. (1980). Unsteady one dimensional flow over a plane: partial equilibrium hydrograph. *Water Resour. Res.* **16**(2): 355–360.
20. Ponce, V.M. and Simons, D.B. (1977). Shallow wave propagation in open channel flow. *J. Hydraul. Div., ASCE* **103**(HY12): 1461–1475.
21. Menendez, A.N. and Norscini, R. (1982). Spectrum of shallow water waves: an analysis. *J. Hydraul. Div., ASCE* **108**(HY1): 75–93.
22. Daluz Vieira, J.A. (1983). Conditions governing the use of approximations for the Saint-Venant equations for shallow surface water flow. *J. Hydrology* **60**: 43–58.
23. Fread, D.L. (1985). *Applicability Criteria for Kinematic and Dynamic Routing Models*. Laboratory of Hydrology, National Weather Service, NOAA, U.S. Department of Commerce, Silver Spring, MD.
24. Ferrick, M.J. (1985). Analysis of river wave types. *Water Resour. Res.* **21**(2): 209–220.
25. Singh, V.P. (1993). Quantifying the accuracy of hydrodynamic approximations for determination of flood discharges. *J. IWRS* **13**(3 & 4): 172–185.
26. Singh, V.P. (1994). Accuracy of kinematic-wave and diffusion-wave approximations for space-independent flows. *Hydrological Processes* **8**: 45–62.
27. Singh, V.P. (1994). Accuracy of kinematic-wave and diffusion-wave approximations for space-independent flows with lateral inflow neglected in the momentum equation. *Hydrological Processes* **8**: 311–326.
28. Singh, V.P. (1994c). Derivation of errors of kinematic-wave and diffusion-wave approximations for space-independent flows. *Water Resour Manage* **8**: 57–82.
29. Singh, V.P. (1995). Accuracy of kinematic-wave and diffusion-wave approximations for space-independent flows on infiltrating surfaces with lateral inflow neglected in the momentum equation. *Hydrological Processes* **9**: 783–796.

30. Singh, V.P. (1996). Errors of kinematic wave and diffusion wave approximations for space-independent flows on infiltrating surfaces. *Hydrological Processes* **10**: 955–969.
31. Pearson, C.P. (1989). One-dimensional flow over a plane: criteria for kinematic wave modeling. *J. Hydrology* **111**: 39–48.
32. Govindaraju, R.S., Jones, S.E., and Kavvas, M.L. (1988). On the diffusion wave model for overland flow: 1. Solution for steep slopes. *Water Resour. Res.* **24**(5): 734–744.
33. Govindaraju, R.S., Jones, S.E., and Kavvas, M.L. (1988). On the diffusion wave model for overland flow: 2. Steady state analysis. *Water Resour. Res.* **24**(5): 745–754.
34. Parlange, J.Y., Hogarth, W., Sander, G., Rose, C., Haverkamp, R., Surin, A., and Brutsaert, W. (1990). Asymptotic expansion for steady state overland flow. *Water Resour. Res.* **26**: 579–583.
35. Singh, V.P. and Aravamuthan, V. (1995). Accuracy of kinematic wave and diffusion wave approximations for time-independent flows. *Hydrological Processes* **9**: 755–782.
36. Singh, V.P. and Aravamuthan, V. (1995). Errors of kinematic wave and diffusion wave approximations for time-independent flows. *Water Resour. Manage.* **9**: 175–202.
37. Singh, V.P. and Aravamuthan, V. (1996). Errors of kinematic wave and diffusion wave approximations for steady state overland flows. *Catena* **27**: 209–227.
38. Singh, V.P. and Aravamuthan, V. (1997). Accuracy of kinematic wave and diffusion wave approximations for time-independent flow with momentum exchange included. *Hydrological Processes* **11**: 511–532.
39. Singh, V.P., Aravamuthan, V., and Joseph, E.S. (1994). Errors of kinematic wave and diffusion-wave approximations for time-independent flows in infiltrating channels. *Irrig. Sci* **15**: 137–145.

KINEMATIC WAVE FLOW ROUTING

VIJAY P. SINGH
Louisiana State University
Baton Rouge, Louisiana

Flow routing on planes, in channels, and in storm sewers using the kinematic wave theory is quite popular, especially for urban drainage design and planning. The small watershed sizes and the often fairly steeper prismatic conduits in an urban environment are well suited for application of the theory. Little or no streamflow data are required to develop the kinematic wave routing method. In nonurban environments, the kinematic routing wave method, corrected for dynamic effects, is also becoming increasingly popular.

KINEMATIC WAVE FORMULATION FOR CHANNEL FLOW ROUTING

The kinematic wave equation is derived from the continuity equation,

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0 \tag{1}$$

and can be expressed in terms of the discharge, Q, as the dependent variable for a channel:

$$\frac{1}{c}\frac{\partial Q}{\partial t} + \frac{\partial Q}{\partial x} = 0, \quad 0 \le x \le L \tag{2}$$

where A is the area of flow cross section, c is the travel speed of the flood wave called the kinematic wave speed and may, at a particular cross section and for a given Q, be evaluated from the Kleitz–Seddon law,

$$c = \frac{dx}{dt} = \left(\frac{dQ}{dA}\right)\bigg|_{x\,\text{fixed}} = \frac{1}{B}\left(\frac{dQ}{dh}\right)\bigg|_{x\,\text{fixed}} \tag{3}$$

where B is the channel flow width, h is the depth of flow, and A is the flow cross-sectional area. When c is constant, Eq. 2 becomes linear and will have a translatory solution. For a variable c, it is a nonlinear equation. In practice, c is usually estimated from a rating curve relating discharge to depth of flow, expressed in general by the Chezy or Manning formula:

$$Q = CAR^m S_f{}^{0.5} \tag{4}$$

where R is the hydraulic radius, C is an empirical resistance coefficient, S_f is the frictional slope = bed slope, S_o, and m is an exponent. Because R depends on A, Eq. 4 is usually written as

$$Q = \alpha A^n, \alpha = C(S_f)^{0.5}, n \in [1, 3] \tag{5}$$

where α is the resistance parameter and n is an exponent. Therefore,

$$c = n\alpha A^{n-1} = n\frac{Q}{A} = nu\left(\frac{Q}{\alpha}\right)^{(n-1)/n} \tag{6}$$

where u is the mean flow velocity.

To solve Eq. 2, the initial and boundary conditions are usually assumed as

$$Q(x, 0) = Q_0(x) \tag{7}$$

$$Q(0, t) = Q_u(x), Q_u(0) = Q_0(0) \tag{8}$$

If $Q_0(x) = 0, 0\langle x\langle L$, then the channel is initially dry. $Q_u(t)$ is such that an analytical solution is not tractable except when $Q_u(t) = Q_u =$ constant. Therefore, numerical solutions with attendant discretization of the solution domains are employed. A number of numerical schemes have been proposed to solve the kinematic wave equations (1). Figure 1 shows a finite-difference grid for a computational cell in (x, t) space. The discretization, however, has its own flaws. It is known that numerical solutions of Eq. 2 introduce variable amounts of numerical diffusion, and the solution resembles a diffusion wave rather than a kinematic wave (2). In first-order numerical schemes, the numerical diffusion is uncontrolled, making the solution depend on the grid size. In second-order schemes, the numerical diffusion vanishes, but a certain amount of numerical dispersion (of the third order) exists (3). Some workers (4) argue that a small amount of numerical diffusion should be welcome, for the kinematic wave solution

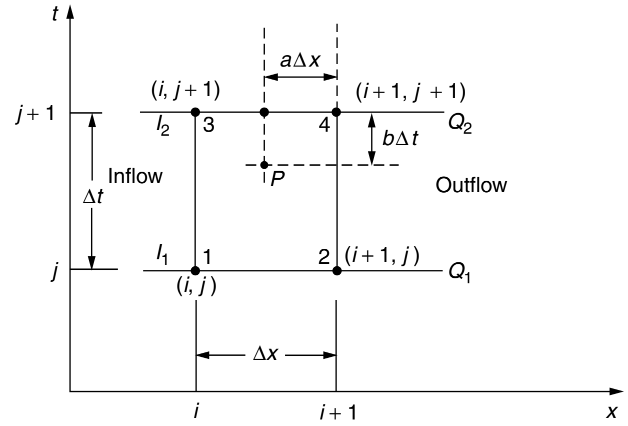

Figure 1. Finite-difference grid for a computational cell in the (x, t) space; a and b are weighting factors in space and time, respectively.

stands to benefit from it and is then applicable to a wider class of problems. Therefore, complete elimination of diffusion may not be desirable from a practical standpoint, for natural diffusion is inherent in physically realistic flows. Thus, one way to accomplish the twin objective of achieving a kinematic wave solution and an acceptable amount of diffusion is to match physical and numerical diffusion judiciously. This procedure has been successfully used for stream-flow routing (5–7). Three possible fully off-centered schemes, which are frequently used, are shown in Fig. 2. The stability depends on the scheme used.

GENERALIZED METHOD FOR A NUMERICAL SOLUTION

A fairly general and flexible finite-difference formulation of the kinematic wave model given by Eq. 2 can be developed, and it can be shown that several existing models are special cases of this formulation. To that end, we employ a rectangular $x - t$ grid, as shown in Fig. 1, where a $(0 \leq a \leq 1)$ and b $(0 \leq b \leq 1)$ are the weighting factors in space and time directions, respectively. Let p be a point in the x–t grid around which the finite-difference approximation is attained. If $W(x, t)$ is any variable denoting either $A(x, t)$ or $Q(x, t)$, then,

$$\frac{\partial W}{\partial x} = \frac{\Delta W}{\Delta x} = \frac{1}{\Delta x}[(1-b)(W_{i+1}^{j} - W_{i}^{j}) + b(W_{i+1}^{j+1} - W_{i}^{j+1})] \tag{9}$$

and

$$\frac{\partial W}{\partial t} = \frac{\Delta W}{\Delta t} = \frac{1}{\Delta t}[(1-a)(W_{i+1}^{j+1} - W_{i+1}^{j}) + b(W_{i}^{j+1} - W_{i}^{j})] \tag{10}$$

where subscript i denotes the space coordinate x and j denotes the time coordinate t.

Similar to Eqs. 9 and 10, Eq. 1 can be written in finite difference around a point p as,

$$\frac{1}{\Delta t}[(1-a)(A_{i+1}^{j+1} - A_{i+1}^{j}) + a(A_{i}^{j+1} - A_{i}^{j}) + \frac{1}{\Delta x}[(1-b)(Q_{i+1}^{j} - Q_{i}^{j}) + b(Q_{i+1}^{j+1} - Q_{i}^{j+1})] = 0 \tag{11}$$

Depending on the values of a and b, point p may lie anywhere within or on the boundaries of the rectangular grid element. Usually Q_{i+1}^{j+1} and A_{i+1}^{j+1} are unknowns, and these at previous points (i, j), $(i + 1, j)$, and $(i, j + 1)$ are known. Arranging the unknowns on the left side and knowns on the right side, Eq. 11 can be written as

$$bQ_{i+1}^{j+1} + (1-a)A_{i+1}^{j+1}\frac{\Delta x}{\Delta t} = bQ_{i+1}^{j} + (1-a)A_{i+1}^{j}\frac{\Delta x}{\Delta t} - Q_{i+1}^{j} - bQ_{i}^{j} + aA_{i}^{j}\frac{\Delta x}{\Delta t} + bQ_{i}^{j+1} - a\frac{\Delta x}{\Delta t}A_{i}^{j+1} \tag{12}$$

The kinematic wave assumption requires a unique relation between Q and A as in Eq. 5. Hence, there remains actually only one unknown on the left side of Eq. 12. Let

$$f(Q_{i}^{j}) = bQ_{i}^{j} - aA_{i}^{j}\frac{\Delta x}{\Delta t} \tag{13}$$

and

$$g(Q_{i}^{j}) = bQ_{i}^{j} + (1-a)A_{i}^{j}\frac{\Delta x}{\Delta t} \tag{14}$$

Taking advantage of Eqs. 13 and 14, Eq. 12 can be expressed as

$$g(Q_{i+1}^{j+1}) = g(Q_{i+1}^{j}) + f(Q_{i}^{j+1}) - f(Q_{i}^{j}) + Q_{i}^{j} - Q_{i+1}^{j} \tag{15}$$

Using the knowledge of the Q–A relationship, functions $f(Q)$ and $g(Q)$ can be computed. First, the right side of Eq. 15 is evaluated. Then, Q_{i+1}^{j+1} is evaluated by using the Q–A relationship. Thus, calculations can proceed either in the x direction or the t direction. The generalized scheme is stable for $a/b \leq c_n \leq (1-a)/(1-b)$, where c_n is the Courant number $= c\Delta t/\Delta x$. When $a = b = 0.5$, the scheme is second-order accurate. This means that point p is centrally located in the rectangular grid. As p departs from the center or as a and b depart from the value of 0.5, truncation errors of the order $O(\Delta x)$ and $O(\Delta t)$ increase,

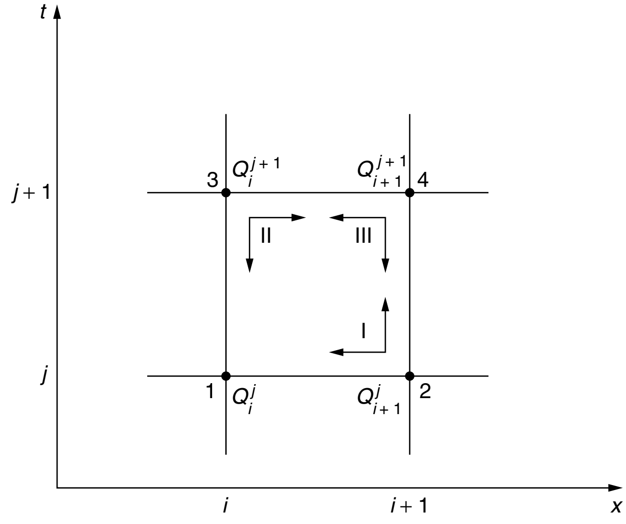

Figure 2. Three fully off-centered numerical schemes for solving kinematic wave equations.

respectively and independently. This truncation is the cause of the numerical diffusion encountered in numerical kinematic channel flow routing. Thus, the desired level of numerical diffusion can be achieved by an appropriate choice of a and b values. A more complete account of the stability and convergence properties of the generalized scheme is given by Smith (8), who showed that several flood routing models are special cases of Eq. 15.

Reservoir Flow Routing

The routing of flow through a reservoir involves a volume balance and the assumption that reservoir storage is controlled by outflow. The volume balance can be expressed as

Rate of change of storage = inflow − outflow,

$$(A^{j+1}_{i+1} - A^{j}_{i+1})\frac{\Delta x}{\Delta t} = \frac{1}{2}(Q^{j}_{i} + Q^{j+1}_{i}) - \frac{1}{2}(Q^{j}_{i+1} - Q^{j+1}_{i+1}) \quad (16)$$

By rearranging Equation 16 and comparing Eq. 12, one gets

$$0.5Q^{j+1}_{i+1} + A^{j+1}_{i+1}\frac{\Delta x}{\Delta t} = 0.5Q^{j}_{i+1} + A^{j}_{i+1}\frac{\Delta x}{\Delta t} - Q^{j}_{i+1} - 0.5Q^{j}_{i} + Q^{j}_{i} + 0.5Q^{j+1}_{i} \quad (17)$$

Equation 17 is a special case of Eq. 12 where $a = 0$ and $b = 0.5$.

Muskingum Method

Flow routing in channels, using the Muskingum method, is based on Eq. 16 and a relation between storage (S), inflow (I), and outflow (Q) (9):

$$S = K[x_m I + (1 - x_m)Q] \quad (18)$$

where x_m ($0 \leq x_m \leq 1$) is a weighting factor and K is travel time. The right side of Eq. 18 defines prism storage and wedge storage; the latter depends on the amount by which inflow exceeds outflow. Using Eq. 3,

$$K = \frac{\Delta x}{c} = \Delta x\left(\frac{dA}{dQ}\right) = \frac{\Delta x \Delta A}{\Delta Q} \quad (19)$$

In the Muskingum method, c is considered constant for a channel reach. Then,

$$KQ = \Delta x Q\frac{\Delta A}{\Delta Q} = \Delta x\frac{Q}{u} = \Delta x\frac{Q}{c} = \Delta x A \quad (20)$$

Equations 18–20 can be combined with the volume balance equation and then written in finite-difference form, similar to Eq. 12, as

$$0.5Q^{j+1}_{i+1} + (1 - x_m)A^{j+1}_{i+1}\frac{\Delta x}{\Delta t} = 0.5Q^{j}_{i+1} + (1 - x_m)A^{j}_{i+1}\frac{\Delta x}{\Delta t} - Q^{j}_{i+1} - 0.5Q^{j}_{i} + x_m A^{j}_{i}\frac{\Delta x}{\Delta t} + Q^{j}_{i} + 0.5Q^{j+1}_{i} - x_m A^{j+1}_{i}\frac{\Delta x}{\Delta t} \quad (21)$$

If $a = x_m$ and $b = 0.5$, Eq. 21 is a special case of Eq. 12. A nonlinear Muskingum method is discussed by Koussis and Osborne (10) and Singh and Scarlatos (11).

Brakensiek's Models

Brakensiek (12) employed the kinematic wave method for flood routing in prismatic channels. He used three numerical schemes which are special cases of Eq. 12 for

$a = 0.5$ and $b = 0.5$ Model I (centered)
$a = 0.5$ and $b = 1.0$ Model II (implicit)
$a = 0.5$ and $b = 0.0$ Model III (explicit)

His flood routing results were sensitive to model type. The first two schemes were more stable than the explicit scheme. The use of an implicit scheme did not require simultaneous solution of Q and A at a new point.

HYMO Model

Williams (13) developed a hydrologic modeling method, called HYMO, using a variable travel time routing procedure. The routing employs the volume balance using arithmetic averages to define discharges over a time increment and end areas over a distance increment. This amounts to using $a = b = 0.5$. A distinguishing feature of HYMO is said to be the variable travel time as a function of the unknown discharge, and therefore it involves an iterative calculation. Smith (8), however, has shown that the travel time has to be a constant. The travel time, T, is calculated in HYMO as

$$T = \frac{\text{reach length}}{\text{average velocity}} = \frac{2S}{I + O} \quad (22)$$

where S is the storage based on end areas, I is the inflow, and Q is the outflow. The wave velocity c, therefore, is

$$c = \frac{L}{T} = \frac{1}{2}\frac{I + O}{A_{\text{av}}} = \frac{Q^{j}_{i} + Q^{j}_{i+1}}{A^{j}_{i} + A^{j}_{i+1}} \quad (23)$$

This implies that

$$c = \frac{dQ}{dA} = \frac{Q}{A} \quad (24)$$

This means that the wave velocity in HYMO is defined by the $Q - A$ chord rather than by the tangent. This is true only for a linear channel where

$$u(x) = \frac{Q}{A} = f(x) \quad (25)$$

This implies that the velocity, u, will be independent of discharge or storage and travel time must necessarily be a constant.

SSARR Model

The U.S. Army Corps of Engineers (14) uses a continuous-time, stream-flow simulation and reservoir regulation model, called SSARR. The model represents a channel by a number of reaches through each of which storage is essentially routed. This amounts to using Eq. 12 where $a =$

0.0 and $b = 0.5$ for each reach. Flood wave modification, therefore, depends on the number of reaches employed.

SWMM Model

The storm water management model (SWMM), developed by the U.S. Environmental Protection Agency (15), is widely used in the United States for urban storm water analysis. In the channel routing phase, a finite-difference form of the continuity equation is employed using weighting coefficients, w_x and w_t, respectively, in space and time, which essentially are $w_x = 1 - a$, and $w_t = b$. When convective acceleration terms are included in the model, then $a = 0.45$ and $b = 0.55$ are used.

MUSKINGUM–CUNGE METHOD

Cunge (2) proposed an explicit finite-difference scheme for solving Eq. 2, which may also be a basis for a generalized treatment of kinematic wave models. The scheme centers the time derivative by taking $b = 0.5$ and retains the weighting coefficient a in space. The celerity, c, is taken as an average constant value for the reach or computational cell. Equation 2 can be written in finite-difference form as

$$\frac{1}{c\Delta t}[(1-a)(Q_{i+1}^{j+1} - Q_{i+1}^{j}) + a(Q_i^{j+1} - Q_i^j)] + \frac{1}{2\Delta x}(Q_{i+1}^j - Q_i^j + Q_{i+1}^{j+1} - Q_i^j) = 0 \tag{26}$$

By solving for Q_{i+1}, Eq. 26 can be written as the classical Muskingum equation:

$$Q_{i+1}^{j+1} = C_1 Q_i^j + C_2 Q_i^{j+1} + C_3 Q_{i+1}^j \tag{27}$$

or

$$Q_4 = C_1Q_1 + C_2Q_2 + C_3Q_3$$

where

$$C_1 = \frac{\Delta t + 2aK}{2K(1-a) + \Delta t} \tag{28}$$

$$C_2 = \frac{2K(1-a) - \Delta t}{2K(1-a) + \Delta t} \tag{29}$$

$$C_3 = \frac{\Delta t - 2aK}{2K(1-a) + \Delta t} \tag{30}$$

and

$$K = \frac{\Delta x}{c} \tag{31}$$

$$C_1 + C_2 + C_3 = 1 \tag{32}$$

The unit of K is time, and it has the connotation of storage-delay time, travel time, translation time, or lag time. It can be estimated from either the Kleitz–Seddon law, hydrographic analysis, or observations.

Cunge (2) showed that Eqs. 26–32 constitute a second-order approximation of the diffusion wave equation if the weighting coefficient a is evaluated as

$$a = \frac{1}{2}\left(1 - \frac{Q}{BS_o\Delta x}\right) \tag{33}$$

where S_0 is the channel bed slope. Cunge derived Eq. 23a from a Taylor series expression of $Q(x + \Delta x, t + \Delta t)$ in the finite-difference form of Eq. 2 and a comparison with the coefficients of the diffusion wave equation for regular channels.

Equation 33 specializes to the equation of Kalinin and Milyukov (16) for $a = 0$:

$$\Delta x = \frac{Q}{S_oBc} = \frac{Q}{S_o(dQ/dh)} \tag{34}$$

where h is the flow depth. Equation 34 can be used to compute the length of subreaches for flow routing in a channel reach.

In the traditional Muskingum method, a single value of a or Δx is used for the entire duration of routing. That can be accomplished by evaluating Eq. 33 or 34 for a representative value of Q and corresponding values of B and c, or dQ/dh. In irregular channels, the average values of these parameters for the subreach should be used. The value of a, it is found, depends on Δx.

Ponce (17) proposed a simplified Muskingum routing equation by taking $\Delta t/K$ and a in Eqs. 26–31 as

$$\frac{\Delta t}{K} = 1, a = 0 \tag{35}$$

Then, it follows that

$$C_1 + C_2 + C_3 = \tfrac{1}{3} \tag{36}$$

leading to the simplified Muskingum routing equation,

$$Q_{i+1}^{j+1} = \tfrac{1}{3}[Q_i^j + Q_i^{j+1} + Q_{i+1}^j] \tag{37}$$

Physically, Eqs. 35 and 36 fix Δt and Δx, and K is the time taken by a flood wave traveling a distance Δx at celerity c given by Eq. 31. Therefore, the Courant number c_N is

$$c_N = \frac{c\Delta t}{\Delta x} = 1 \tag{38}$$

implying equality between flood-wave celerity c and grid celerity $\Delta x/\Delta t$.

Ponce (17) defined a grid Reynolds number R_g (the ratio of physical or hydraulic diffusivity to numerical diffusivity) from Eqs. 33 and 36 or Eq. 34 as

$$R_g = \frac{Q}{BS_oc\Delta x} = 1 \tag{39}$$

which states that the channel diffusivity $Q/(2BS_o)$ is equal to the grid diffusivity $c\Delta x/2$. In Ponce's method,

Δx and Δt can be computed from flow variables by using Eqs. 35 and 39. For example, Δx is given by Eq. 34 or

$$\Delta x = \frac{q}{cS_o} \tag{40}$$

and Δt by Eq. 39,

$$\Delta t = \frac{\Delta x}{c} = \frac{q}{S_o c^2} \tag{41}$$

where q is the unit-width discharge. For natural (nonprismatic) channels, Eq. 37 relates the channel friction to cross-sectional shape and expresses the steady-state discharge relation. Using Eq. 38 in Eq. 40 yields

$$\Delta x = \frac{A}{nBS_o} \tag{42}$$

and

$$\Delta t = \frac{\Delta x}{c} = \frac{A^{2-n}}{\alpha n^2 BS_o} \tag{43}$$

The routing coefficients C_1, C_2, and C_3 can be expressed in terms of c_N and R_g by using Eqs. 31, 33, 38, and 39:

$$C_0 = 1 + c_N + R_g \tag{44}$$

$$C_1 = \frac{1 + c_N - R_g}{C_o} \tag{45}$$

$$C_2 = \frac{-1 + c_N + R_g}{C_o} \tag{46}$$

$$C_3 = \frac{1 - c_N + R_g}{C_o} \tag{47}$$

where c_N and R_g are expressed in terms of $B, c, \Delta x, \Delta t$, and Q.

If there is lateral inflow q_L, the Muskingum equation 27 can be modified as

$$Q_{i+1}^{j+1} = C_1 Q_i^j + C_2 Q_i^{j+1} + C_3 Q_{i+1}^j + Q_L \tag{48}$$

in which

$$Q_L = \frac{2cq_L \Delta t}{(\Delta t/K) + 2(1-a)} \tag{49}$$

where q_L is the unit-width lateral inflow. Because $c = \Delta t/K$ and if $K = \Delta t$ and $a = 0$, Eq. 50 reduces to

$$Q_L = \frac{2q_L \Delta x}{3} \tag{50}$$

It is pertinent at this point to comment on the role of the bed slope S_o and the frictional slope S_f in the Muskingum–Cunge method. In prismatic channels, S_o is used as one of the main channel parameters. Under steady-state, uniform conditions, it is known that $S_o = S_f = S_e$, where S_e is the energy slope. Under steady-state but nonuniform conditions, $S_o \neq S_e = S_f$. In the Muskingum–Cunge method, linearization (or the Taylor series approximation) of the flow is performed around the steady-state uniform or equilibrium flow condition represented by Q_o which permits using S_o in Eq. 33 or 34. In actual practice, however, S_o is not often readily determined, owing largely to local channel bed irregularities. For such channels, S_f is more readily obtained from a given value of Manning's roughness coefficient. Thus, it is more convenient to use S_f as an estimator of S_o. Ponce (18) used S_f in lieu of S_o in his work on linearized diffusion wave modeling.

MUSKINGUM–CUNGE METHOD WITH VARIABLE PARAMETERS

Ponce and Yevjevich (6) considered the space–time variation of K and a as the flow varied. In this case, Δt is usually fixed, and Δx and S_o are specified for each computational cell comprising four grid points. This requires determining flood wave celerity c and the unit-width discharge q for each computational cell. The Muskingum–Cunge method requires calculating the average values of Q and c for each computational cell to compute K and a. Ponce and Yevjevich (6) state three ways of determining c and q for use in computing C_N and R_g: (1) using a two-point average of values at grid point (i,j) and $(i+1,j)$; (2) using a three-point average of values at grid points (i,j), $(i+1)$, and $(i,j+1)$; and (3) using a four-point average calculation by iteration. The average values for a four-point scheme can be expressed bt taking the average at four points (19). Holden and Stephenson (19) showed through numerical experiments that parameter a is virtually independent of grid spacing. Furthermore, the lower the value of a, the greater the attenuation in flood peak. This means that the closer a is to zero, the greater attenuation occurs as numerical diffusion. Although the values of a found are close to 0.5 for many flow conditions, the value of a drops significantly lower for river channels because more gentle bed slopes lead to higher attenuation of a flood hydrograph. The implicit formation of the Muskingum–Cunge method has no particular advantage over the explicit one, and hence the latter may be preferable.

KOUSISS MODEL

Koussis (20) proposed a generalized numerical solution of Eq. 2 in which discretization was carried out only in space and continuous functions for time derivatives were retained. Considering c as an average value for the computational cell, Eq. 2 was written as an ordinary differential equation in Q_{i+1}. Subject to the initial condition $Q_{i+1}(0) = Q_i(0) = Q(x, t = 0)$, the resulting equation was solved by assuming a linear variation of $Q_j(t)$ over Δt and then expressing it in the form of the Muskingum equation.

MODEL PARAMETERS

Four parameters are involved in the previous kinematic wave models: Δx, Δt, a, and c. These parameters are used to compute the routing coefficients. The parameters can be constant (linear case) or variable (nonlinear case).

Spatial Grid Spacing

A single-valued storage function cannot be prescribed for long river reaches, and floods should be routed by dividing the reach length L into N subreaches of length Δx. Laurenson (21,22) has shown that Δx and the weighting parameter a are not independent of each other in that a decreases with Δx or as storage becomes less distributed. Because a cannot be negative, Δx must have a lower limit, or N must have a finite upper limit. Nash (23) argued that unrealistic negative outflows at the start of the outflow hydrograph were the result of applying the Muskingum method to distributed storages ($a \neq 0$). Weinmann and Laurenson (1) showed that these negative outflows can be ignored under certain conditions.

Temporal Grid Spacing

If Δt is much smaller than $\Delta x/c$, then the disturbance might not have traveled distance Δx. For $a \neq 0$, computation of variables at the current section may lead to unrealistic results such as a dip in the outflow hydrograph or negative outflows for an initially dry channel. The best time step for the kinematic wave model can be taken as $\Delta t = \Delta x/c$.

Weighting Factor, *a*

Strictly speaking, a can take on any value from zero to one. The value $a > 0.5$ results in amplifying flood waves and is therefore not important in flood routing. The value $a = 0.5$ for all values of $\Delta x/\Delta t$ results in pure translation of a flood wave (2), corresponding to fully distributed storage. The values $a < 0.5$ introduce numerical diffusion resulting in wave attenuation. For the extreme case, $a = 0$ corresponds to fully concentrated storage or a reservoir that has an instantaneous response and attenuating inflow peak. This discussion then shows that when Δx and Δt are chosen within prescribed limits, a flood wave can be both attenuated and translated by an appropriate choice of a. This is similar to what is done by using lag and route models that treat translation and attenuation separately. Translation can also be achieved by choosing $a = 0$ and an appropriate value of N, when concentrated storages are arranged in series, as in the Kalinin–Milyukov model. Frequently, a constant value of a will be sufficient, unless it varies greatly.

Wave Celerity, *c*

In kinematic wave models, c for a given reach is a function only of flow depth, for the rating curve is a single-valued function. In the coefficient representation of these models, a flood wave is attenuated by taking $a < 0.5$. Similar to kinematic wave models, the storage routing models commonly employed in hydrology use a single-valued storage–discharge curve. Koussis (5,20,24,25), Williams (13), and Ponce (4), amongst others, included a looped rating curve in kinematic wave models and thereby corrected them for dynamic effects.

To determine the average reciprocal wave celerity, the channel reaches should be selected as morphologically similar. For natural rivers, the variations in channel cross sections, bed slope, and roughness may be so great that it may not suffice to use only two $A(Q)$ curves at i and $(i + 1)$ cross sections, unless Δx is quite small. Nevertheless, it is reasonable to compute the average reciprocal wave celerity from steady flow $A(Q)$ curves at several cross sections. In prismatic channels, this may be computed from the slope of the $A(Q)$ curve for unsteady-state flow at i and $(i + 1)$ cross sections.

BIBLIOGRAPHY

1. Weinmann, P.E. and Laurenson, E.M. (1979). Approximate flood routing methods: a review. *J. Hydraul. Div., ASCE* **105**(4): 638–639.
2. Cunge, J. (1969). On the subject of a flood wave computation method (Muskingum method). *J. Hydraul. Res.* **7**(2): 205–230.
3. Ponce, V.P. and Theurer, F.B. (1982). Accuracy criteria in diffusion routing. *J. Hydraul. Div., ASCE* **108**(HY6): 747–757.
4. Ponce, V.M. (1986). Diffusion wave modeling of catchment dynamics. *J. Hydraul. Eng.* **112**(8): 716–727.
5. Koussis, A.D. (1983). Unified theory of flood and pollution routing. *J. Hydraul. Eng.* **109**(12): 1652–1664.
6. Ponce, V.M. and Yevjevich, V. (1978). Muskingum-Cunge method with variable parameters. *J. Hydraul. Div., ASCE* **104**(HY12): 1663–1667.
7. Price, R.K. (1978). A river catchment flood model. *Proc. Inst. Civ. Eng., London, Part 2* **65**: 655–668.
8. Smith, A.A. (1980). A generalized approach to kinematic flood routing. *J. Hydrology* **45**: 71–89.
9. Singh, V.P. (1988). *Hydrologic Systems*. Vol. 1. Rainfall-Runoff Modeling. Prentice-Hall, Englewood Cliffs, NJ.
10. Koussis, A.D. and Osborne, B.J. (1986). A note on nonlinear storage routing. *Water Resour. Res.* **22**(13): 2111–2113.
11. Singh, V.P. and Scalatos, P. (1987). Analysis of the nonlinear Muskingum flood routing. *J. Hydraul. Eng.* **113**(1): 61–79.
12. Brakensiek, D. (1967). Kinematic flood routing. *Trans. ASAE* **10**(3): 340–343.
13. Williams, J.R. (1975). HYMO flood routing. *J. Hydrology* **26**: 17–27.
14. U.S. Army Corps of Engineers. (1972). *SSARR Model: Streamflow Synthesis and Reservoir Regulation*. U.S. Army Corps of Engineers Portland District, Portland, OR.
15. Huber, W.C., Heany, J.P., Medina, M.A., Peltz, W.A., Sheikh, H., and Smith, G.F. (1975). *Stormwater Management Model User's Manual. Document EPA-670/2-75-017*. National Environment Research Center, U.S. Environmental Protection Agency, Cincinnati, OH.
16. Kalinin, G.P. and Milukov, P.I. (1958). On the computation of unsteady water flow along channels by the use of reach travel curves (in Russian). *Meteorology and Hydrology* **7**: 18–25.
17. Ponce, V.M. (1979). Simplified Muskingum routing equation. *J. Hydraul. Div., ASCE* **105**(HY1): 85–91.
18. Ponce, V.M. (1981). *Development of Algorithm for the Linearized Diffusion Method of Flood Routing*. SDSU Civil Engineering Series 8114, San Diego State University, San Diego, CA, p. 84.
19. Holden, A.P. and Stephenson, D. (1995). Finite difference formulations of kinematic equations. *J. Hydraul. Eng.* **121**(5): 423–426.
20. Koussis, A.D. (1976). *An Approximate Dynamic Flood Routing Method*. Paper L1, Procedings of the International Symposium on Unsteady Flow in Open Channels, New Castle-upon-Tyne, Northumberland, UK.

21. Laurenson, E.M. (1962). *Hydrograph Synthesis by Runoff Routing*. Report 66, Water research Laboratory, The University of New South Wales, Manly Vale, Kennsington, New South Wales, Australia.
22. Laurenson, E.M. (1964). A catchment storage model for runoff routing. *J. Hydrology* **2**: 141–163.
23. Nash, J.E. (1959). A note on the Muskingum flood routing method. *J. Geophys Res.* **64**: 1053–1056.
24. Koussis, A.D. (1978). Theoretical estimation of flood routing parameters. *J. Hydraul. Div., ASCE* **104**(HY1): 109–115.
25. Koussis, A.D. (1980). Comparison of Muskingum difference schemes. *J. Hydraul. Div., ASCE* **106**(HY6): 925–929.

RELIABILITY CONCEPTS IN RESERVOIR DESIGN

Demetris Koutsoyiannis
National Technical University
Athens, Greece

INTRODUCTION

The reliability of a system is defined as the probability that the system will perform the required function for a specified period of time under stated conditions. Reliability is the complement of the probability of failure (or risk), the probability that the "loading" will exceed the "capacity." Denoting α as the reliability, β the probability of failure, and $P[\omega]$ the probability of an event ω, the mathematical expression of this definition is

$$\alpha := P[L(t) < C(t); t \in \Pi] =: 1 - \beta \tag{1}$$

where $L(t)$ and $C(t)$ represent the loading and capacity, respectively, at time t, within a certain time period Π (e.g., a year). The failure of a system can be classified as *structural failure* or *performance failure*. Structural failure involves damage of the structure or facility, hindering its ability to function as desired in the future, whereas performance failure does not necessarily involve structural damage but rather inability of the system to perform as desired at some time within the period of interest, which results in temporary unfavorable consequences.

RESERVOIR DYNAMICS

A reservoir's function is to regulate natural inflows, which vary irregularly, to provide outflows at a more regular rate that is determined by water demand for one or more uses (water supply, irrigation, hydropower), temporarily storing the surplus, when inflows exceed outflows. Reservoir dynamics are more conveniently expressed in discrete rather than continuous time.

The quantities that are necessary to describe dynamics are the following:

Storage S_t. More precisely known as *active storage,* it is the volume of water stored, at time t, above the minimum level, which is determined either technically (i.e., as the level of the lowest valve of off-take) or legally by a decree imposing rules for a reservoir's operation. Active storage S_t ranges between zero and a maximum value c imposed by the reservoir size, which corresponds to the level of the spillway crest (or some specified level above it when the sluice gates are constructed over the spillway). During floods, excess water is routed through the spillway, which causes temporary storage above the normal limit c. This is known as flood control storage. Water storage below the minimum level, known as *dead* or *inactive storage,* serves two main purposes: It provides volume for sediment accumulation and environmental protection, as it protects the habitat of the reservoir during dry periods by hindering complete emptying. Associated with the last function is also conservation of the quality of the landscape. This article is focused on the design of the active storage of a reservoir; some notes on additional storage zones are contained in the last section of the article.

Net Inflow X_t. It is the algebraic sum of cumulative inflows to the reservoir from time $t-1$ to time t, minus the losses during the same time period. Inflows include runoff from the catchment upstream of the reservoir (typically, the main component of inflows), rainfall on the surface area of the reservoir, and, possibly, water artificially conveyed from other sources (e.g., interbasin transfers through tunnels or pipelines). Losses include evaporation from the surface area of the reservoir, possibly seepage to groundwater, and leakage under or through the dam.

Water Demand, δ_t. It is the sum of all water requirements for the different water uses served by the reservoir for the time period $(t-1, t)$. The demand may vary with time (e.g., due to seasonal agricultural demand or due to some rule, usually based on the quantity of water in the reservoir).

Release, R_t. Also known as draft, withdrawal, or outflow, it is the actual amount of water taken from the reservoir to satisfy water demand during the time period $(t-1, t)$. When there is a sufficient amount of water in the reservoir, R_t equals demand δ_t; otherwise $R_t < \delta_t$.

Spill, W_t. It is the excess water that, during times of floods and simultaneously high reservoir storage, cannot be stored in the reservoir due to the upper reservoir storage limit c.

Reservoir dynamics are easily expressed by the law of mass conservation, or equivalently, the water balance equation. Considering that S_t is limited between 0 and c, the water balance equation is easily formulated as

$$S_t = \max(0, \min(S_{t-1} + X_t - \delta_t, c)) \tag{2}$$

In addition, the release is determined as

$$R_t = \min(S_{t-1} + X_t, \delta_t) \tag{3}$$

and the spill as

$$W_t = S_{t-1} - S_t + X_t - R_t = \max(0, S_{t-1} + X_t - \delta_t - c) \tag{4}$$

Equations 2–4 apply when the inflow and withdrawal occur at constant rates throughout the period $(t, t-1)$—this could be called the "steady" model. A simple

21. Laurenson, E.M. (1962). *Hydrograph Synthesis by Runoff Routing*. Report 66, Water research Laboratory, The University of New South Wales, Manly Vale, Kennsington, New South Wales, Australia.
22. Laurenson, E.M. (1964). A catchment storage model for runoff routing. *J. Hydrology* **2**: 141–163.
23. Nash, J.E. (1959). A note on the Muskingum flood routing method. *J. Geophys Res.* **64**: 1053–1056.
24. Koussis, A.D. (1978). Theoretical estimation of flood routing parameters. *J. Hydraul. Div., ASCE* **104**(HY1): 109–115.
25. Koussis, A.D. (1980). Comparison of Muskingum difference schemes. *J. Hydraul. Div., ASCE* **106**(HY6): 925–929.

RELIABILITY CONCEPTS IN RESERVOIR DESIGN

DEMETRIS KOUTSOYIANNIS
National Technical University
Athens, Greece

INTRODUCTION

The reliability of a system is defined as the probability that the system will perform the required function for a specified period of time under stated conditions. Reliability is the complement of the probability of failure (or risk), the probability that the “loading” will exceed the “capacity.” Denoting α as the reliability, β the probability of failure, and $P[\omega]$ the probability of an event ω, the mathematical expression of this definition is

$$\alpha := P[L(t) < C(t); t \in \Pi] =: 1 - \beta \qquad (1)$$

where $L(t)$ and $C(t)$ represent the loading and capacity, respectively, at time t, within a certain time period Π (e.g., a year). The failure of a system can be classified as *structural failure* or *performance failure*. Structural failure involves damage of the structure or facility, hindering its ability to function as desired in the future, whereas performance failure does not necessarily involve structural damage but rather inability of the system to perform as desired at some time within the period of interest, which results in temporary unfavorable consequences.

RESERVOIR DYNAMICS

A reservoir’s function is to regulate natural inflows, which vary irregularly, to provide outflows at a more regular rate that is determined by water demand for one or more uses (water supply, irrigation, hydropower), temporarily storing the surplus, when inflows exceed outflows. Reservoir dynamics are more conveniently expressed in discrete rather than continuous time.

The quantities that are necessary to describe dynamics are the following:

Storage S_t. More precisely known as *active storage,* it is the volume of water stored, at time t, above the minimum level, which is determined either technically (i.e., as the level of the lowest valve of off-take) or legally by a decree imposing rules for a reservoir’s operation. Active storage S_t ranges between zero and a maximum value c imposed by the reservoir size, which corresponds to the level of the spillway crest (or some specified level above it when the sluice gates are constructed over the spillway). During floods, excess water is routed through the spillway, which causes temporary storage above the normal limit c. This is known as flood control storage. Water storage below the minimum level, known as *dead* or *inactive storage,* serves two main purposes: It provides volume for sediment accumulation and environmental protection, as it protects the habitat of the reservoir during dry periods by hindering complete emptying. Associated with the last function is also conservation of the quality of the landscape. This article is focused on the design of the active storage of a reservoir; some notes on additional storage zones are contained in the last section of the article.

Net Inflow X_t. It is the algebraic sum of cumulative inflows to the reservoir from time $t-1$ to time t, minus the losses during the same time period. Inflows include runoff from the catchment upstream of the reservoir (typically, the main component of inflows), rainfall on the surface area of the reservoir, and, possibly, water artificially conveyed from other sources (e.g., interbasin transfers through tunnels or pipelines). Losses include evaporation from the surface area of the reservoir, possibly seepage to groundwater, and leakage under or through the dam.

Water Demand, δ_t. It is the sum of all water requirements for the different water uses served by the reservoir for the time period $(t-1, t)$. The demand may vary with time (e.g., due to seasonal agricultural demand or due to some rule, usually based on the quantity of water in the reservoir).

Release, R_t. Also known as draft, withdrawal, or outflow, it is the actual amount of water taken from the reservoir to satisfy water demand during the time period $(t-1, t)$. When there is a sufficient amount of water in the reservoir, R_t equals demand δ_t; otherwise $R_t < \delta_t$.

Spill, W_t. It is the excess water that, during times of floods and simultaneously high reservoir storage, cannot be stored in the reservoir due to the upper reservoir storage limit c.

Reservoir dynamics are easily expressed by the law of mass conservation, or equivalently, the water balance equation. Considering that S_t is limited between 0 and c, the water balance equation is easily formulated as

$$S_t = \max(0, \min(S_{t-1} + X_t - \delta_t, c)) \qquad (2)$$

In addition, the release is determined as

$$R_t = \min(S_{t-1} + X_t, \delta_t) \qquad (3)$$

and the spill as

$$W_t = S_{t-1} - S_t + X_t - R_t = \max(0, S_{t-1} + X_t - \delta_t - c) \qquad (4)$$

Equations 2–4 apply when the inflow and withdrawal occur at constant rates throughout the period $(t, t-1)$—this could be called the “steady” model. A simple modification of the equations allows for the case where the inflow (or withdrawal) is highly seasonal and can (in the limit) be modeled as a sudden occurrence; this could be called the “sudden” model. These two models, called “simultaneous” and “staggered” by Pegram (1), bound all the behavior observed in real reservoirs.

DEFINITION OF RELIABILITY APPLIED TO A RESERVOIR

Now, the previously stated general definition of reliability [also known as dependability; e.g., (2) p. 312] can be applied to a reservoir. It is observed that the failure of a reservoir's function is a performance failure, a failure to meet the water demand. At time t, the loading is the water demand δ_t, and the capacity is the sum of S_{t-1} (storage at time $t-1$) and X_t (inflow from time $t-1$ to t). Thus, application of Equation 1 yields

$$\alpha = 1 - \beta = P[\delta_t < S_{t-1} + X_t] \tag{5}$$

Considering Equations 2 and 3, the following equivalent and more convenient expressions are found:

$$\alpha = P[S_t > 0], \beta = P[S_t = 0] \tag{6}$$

and

$$\alpha = P[R_t = \delta_t], \beta = P[R_t < \delta_t] \tag{7}$$

In this context, the demand δ_t is regarded as a known quantity at any time instant t. All other quantities involved, namely, S_t, X_t, R_t, and W_t, are regarded as random variables. Given the storage capacity c, the demand δ_t, and the probability distribution and autocorrelation functions of the input X_t, in theory, the probability distributions of output variables S_t, R_t, and W_t can be determined in terms of that of X_t; this, however, is not an easy task due to the nonlinearity of the dynamics expressed in Equations 2–4. Theoretically, once the distribution function of S_t or R_t has been determined, the reliability α is obtained directly from Equation 6 or 7, respectively. However, although the theoretically based calculations unavoidably involve complete knowledge of the distribution function of S_t and R_t, the reliability α can be determined in an alternative, much simpler, manner. Under the assumptions of stationarity and ergodicity, α can be estimated from a historical (for an existing reservoir) or synthesized (via simulation) time series of storage s_t or release r_t of adequate length n. (Here lower case symbols were used for values of the random variables S_t and R_t.) Specifically, the estimate of α based on Equation 6 is

$$\alpha = \frac{1}{n}\sum_{t=1}^{n}[1 - U(-s_t)] \tag{8}$$

where $U(x)$ is Heaviside's unit step function, where $U(x) = 1$ for $x \geq 0$ and $U(x) = 0$ for $x < 0$. Correspondingly, the estimate of α based on Equation 7 is

$$\alpha = \frac{1}{n}\sum_{t=1}^{n} U(r_t - \delta_t) \tag{9}$$

The purpose of the sum in Equation 8 (or 9) is to count the periods when storage is not zero (or release r_t equals the demand δ_t). Thus, reliability is expressed as the proportion of time steps when the system performs as desired. For this reason, α has also been termed *time-based* reliability. Here, it must be observed that although Equations 5–7 are mathematically equivalent to each other, as are Equations 8 and 9, when applied to historical time series, they may result in different estimates. For example, a city's water supply may not be allowed to empty, as restrictions on releases are applied before this situation is reached. In such a case, Equation 8 may result in the erroneous estimate $\alpha = 1$, whereas Equation 9 will estimate the reliability correctly if the desired demand, before restrictions, is entered into the calculations. Thus, Equation 9 is preferable when dealing with historical time series, but both are equivalent in synthetic simulations; moreover, application of Equation 8 is faster as it does not require simulating releases at all (only Equation 2 needs to be applied).

The stationarity assumption that was inherent in the previous analysis is satisfactory when the time step is a year (either calendar or hydrologic). However, the annual time step is usually too large and hides the variation of both inflows and demand within a year, which may result in a failure some time within the year that is recovered at the end of the year. Therefore, a smaller time step (e.g., monthly) is usually chosen, so that one year corresponds to $k > 1$ (e.g., 12) time steps. On this finer timescale, all processes depend on the time step in a periodic manner, that is, they are cyclostationary. Yet, the reliability and failure probability are usually expressed on the annual scale, in which stationarity is redeemed. To shift from the finer scale to the annual scale, the rule adopted is that a failure occurring in one or more finer scale time steps is regarded as a failure for the year. Using this rule, Equation 6 becomes

$$\alpha' = P\left[\bigcap_{i=1}^{k}(S_{t-i} > 0)\right], \beta' = P\left[\bigcup_{i=1}^{k} S_{t-i} = 0\right] \tag{10}$$

where the symbols α' and β' were used instead of α and β to distinguish from time-based reliability, whereas the symbol ‘$\bigcap$’ indicates that all of the following events should occur simultaneously and ‘$\bigcup$’ indicates that any of the following events should occur. Similarly, Equation 7 becomes

$$\alpha' = P\left[\bigcap_{i=1}^{k}(R_{t-i} = \delta_{t-i})\right], \beta' = P\left[\bigcup_{i=1}^{k} R_{t-i} < \delta_{t-i}\right] \tag{11}$$

or alternatively,

$$\alpha' = P\left[\sum_{i=1}^{k} R_{t-i} = \sum_{i=1}^{k}\delta_{t-i}\right], \beta' = P\left[\sum_{i=1}^{k} R_{t-i} < \sum_{i=1}^{k}\delta_{t-i}\right] \tag{12}$$

Likewise, Equations 8 and 9 become

$$\alpha' = \frac{k}{n}\sum_{p=1}^{n/k}\min\{[1 - U(-s_t)]; t = k(p-1)+1, \ldots, kp\} \tag{13}$$

$$\alpha' = \frac{k}{n}\sum_{p=1}^{n/k}\min\{U(r_t - \delta_t); t = k(p-1)+1, \ldots, kp\} \qquad (14)$$

respectively. The sums in Equations 13 and 14 count the number of years in which no failure has occurred. Apparently, α' and β' provide information on the occurrence of a failure within a year and not in the time period during which the failure lasted. Therefore, they have been known as *occurrence-based* reliability and failure probability, respectively. An overall indication of the duration of failures within an average year can be obtained by applying Equation 8 or 9 and estimating β, the time-based probability of failure.

Apart from occurrence-based and time-based reliability, an additional reliability measure has been often used, which is not expressed in terms of probability (and thus, literally does not comply with the general definition of reliability). This is the so-called *volumetric* or *quantity-based* reliability, expressed as the ratio of the average release to demand:

$$\alpha_V = 1 - \beta_V = E\left[\sum_{i=1}^{k} R_t\right] \bigg/ \sum_{i=1}^{k} \delta_t \qquad (15)$$

Given that a failure that occurs in a year does not extend over the whole year, and, in addition, the release during the failure is not necessarily zero but some positive quantity smaller than demand, it is easily concluded that

$$\alpha' \leq \alpha \leq \alpha_V \qquad (16)$$

Among the three measures of reliability, the most important and most frequently used is the severest, the occurrence-based reliability α'. Another means for expressing virtually the same concept is the return period or recurrence interval of emptiness, T. This is the mean time between two consecutive empty states of the reservoir, and it is none other than the reciprocal of the probability of failure β' (1):

$$T := 1/\beta' = 1\ (1 - \alpha') \qquad (17)$$

which is expressed in years (given that α' and β' are expressed on an annual timescale). The concept of the return period of emptiness of a reservoir is similar to that typically used for design floods. The difference is that in design floods, failure is the exceedance of the magnitude of the design flood, whereas failure in a reservoir is emptying of the reservoir. Typical design values of reliability and return period for reservoir design are $\alpha' = 99\%$ ($T =$ 100 years) for municipal water supply reservoirs, $\alpha' = 70 - 85\%$ ($T = 3.3 - 6.7$ years) for irrigation reservoirs in subhumid climates, and $\alpha' = 80 - 95\%$ ($T = 5 - 20$ years) for irrigation reservoirs in arid climates (2, p. 313).

TRADITIONAL RESERVOIR DESIGN PROCEDURES

Most hydraulic structures, such as flood protection works and drainage networks, whose load varies randomly, have been designed on a probabilistic basis, adopting a certain reliability level or, equivalently, a certain return period for the design flood. Traditionally, however, this has not been the case in reservoir design, which has rarely been based on sound probability. This is obvious even from the terminology traditionally used. For example, the use of the term *firm yield* implies a nonprobabilistic, or failure-free concept. Specifically, the firm yield of a reservoir has been defined as the draft or withdrawal that lowers the water content in a reservoir from a full condition to its minimum allowable level just once during a critical historical drought. (3, p. 27.8; 4, p. 534). It has been characterized as essentially the no-failure yield (3, p. 27.8). In a probabilistic context, however, any draft has a nonzero probability of failure (unless the demand is less than the hypothetical lower bound of the inflow distribution, which can be plausible only for perennial streams; this is unusual).

Several procedures have been widely used in reservoir design, which are rather deterministic and not consistent with the reliability concept. The most common has been *mass curve analysis* and its variations. A mass curve is a plot of cumulative inflow volumes (typically based on historical discharge records) as a function of time. Using this plot, the firm yield, as well as the required reservoir storage to attain this firm yield, can be determined graphically. In addition, the method can determine the required storage for a smaller target release. This graphical method was developed 120 years ago (5) and has been widely used until now, although criticized (6) for not providing information on the probability of failure and for the fact that the reservoir capacity determined by this method increases as the arbitrary length of available observed inflow data increases. As shown by Feller (7), this increase is asymptotically proportional to the square root of the length of the record.

A first variation of the method is its application using synthetic, rather than observed, data (6). This eliminates the drawback of the arbitrary length of the record and also provides some measure of uncertainty by applying the same procedure using different generated synthetic series. However, this kind of description of uncertainty is not consistent with a rational definition of reliability (e.g., that of the previous section).

A more theoretical flavor for the method has been given by the so-called *range analysis*, commenced by the work of Hurst (8; see also 9, p. 184). Range is essentially the algebraic difference of the maximum and minimum departures of the mass curve from the straight line that joins its starting and ending points. The range concept has greatly contributed to the understanding and description of the so-called Hurst phenomenon in hydrology, climatology, and other geophysical sciences. Applied to a reservoir, the range represents the required storage of a reservoir operating without any spill or other loss and providing a constant outflow equal to the mean flow. Obviously, this is an oversimplification of a real reservoir. On the other hand, the range concept involves complexity in estimation, and simpler and more efficient methods have been proposed that can be used instead of range analysis (10).

An additional design method is the so-called *sequent-peak analysis* (11, p. 274; 12, p. 400). Essentially, it is a tabulated version of mass-curve analysis and can also incorporate in the calculations, apart from runoff, the effects of precipitation, evaporation, and leakage. The method does not involve the reliability concept, nor does it consider spills from the reservoir.

SIMPLIFIED RELIABILITY-BASED PROCEDURES FOR RESERVOIR DESIGN

As already mentioned, the analytical determination of reliability in the general case of a reservoir fed by inflows with seasonality, that have arbitrary probability distribution, and autocorrelation functions is a very difficult, if not impossible, task (1). Therefore, existing analyses have been based on several simplifications. However, the results of such analyses are very useful, at least for the initial stage of reservoir design. The typical simplifying assumptions are to

- neglect secondary inflows (precipitation) and losses (evaporation, leakage);
- neglect seasonality by the adoption of an annual time step;
- neglect autocorrelation and assume that inflows are independent in time;
- use a specific distribution function for inflows, typically two-parameter such as normal, lognormal, or gamma.

The objective of such probability-based theoretical analyses is to determine the relation of the following three quantities:

- reservoir size c; it is usually standardized as $\kappa := c/\sigma$, where σ is the standard deviation of annual net inflow X_t;
- demand δ, which is assumed constant for all years; it is usually standardized as $\varepsilon := (\mu - \delta)/\sigma$, where μ is the mean of annual net inflow X_t; ε has been termed the standardized inflow (8; 3, p. 27.7) or the *drift* (1);
- probability, expressed either as reliability α, probability of failure β, or return period T; because of the annual time step used, time-based reliability (Equations 6, 7) is identical to occurrence-based reliability (Equations 10–12).

The first among the probabilistic approaches used in such analyses is discretization of reservoir storage into several zones, each representing a certain state, and the use of a Markov chain model to represent transitions from state to state (13,14; see also 9, p. 264).

A second method is stochastic (Monte Carlo) simulation, in which a long synthetic series of inflows is generated from the appropriate distribution function and then transformed into a series of storage values using Equation 2; reliability is then easily determined from the storage time series using Equation 8. Gould (15) used this method to propose a reservoir size–yield–reliability formula, fitted to 240 sets of Monte Carlo simulations for various combinations of demand, reservoir size, and skewness of gamma distributed inflows:

$$(\varepsilon + 0.15)[\kappa + d_1(\alpha, \gamma)] = d_2(\alpha, \gamma) \quad (18)$$

where d_1 and d_2 are coefficients depending on reliability α and skewness γ and are given by nomographs (see also 2, p. 323). McMahon and Mein (16) adapted this formula to indicate reliability more explicitly; this can be estimated from the standardized normal variate z_α corresponding to α, using the equation

$$z_\alpha = 2\sqrt{\varepsilon(\kappa + d(\alpha)\sigma/\mu)} \quad (19)$$

where d is a coefficient depending on reliability α and is given by a table (see also 3, p. 27.14).

A more accurate and rigorous theoretical methodology to estimate the reservoir size–yield–reliability relationship was developed by Pegram (1). This was based on finite-difference and integral equations, which employ reservoir dynamics (Equation 2) in a probabilistic context to determine the return period of emptiness. Pegram applied his methodology for normal, lognormal, and discrete inputs both independent and serially correlated. His results, when compared to those of the Gould method (Equations 18–19) indicate that the latter underestimates the reservoir size required to attain a certain reliability level. Using Pegram's results for normally distributed inflows, which were verified and expanded here with extended simulations, the following approximate relationship has been established:

$$\ln(T - 1) = 2(\varepsilon + 0.25)(\kappa + 0.5)^{0.8} \quad (20)$$

This is valid for $T > 2 (\alpha > 0.5)$ and can be alternatively written as

$$\begin{aligned}\ln(T - 1) &= -\ln(1/\alpha - 1) \\ &= (2/\sigma^{1.8})(\mu + 0.25\sigma - \delta)(c + 0.5\sigma)^{0.8} \quad (21)\end{aligned}$$

For known mean μ and standard deviation σ of inflows, Equation 21 can directly yield either the reliability α for a known reservoir size and demand, the reservoir size c for a given demand and reliability, or the demand δ that can be met at a given reliability for a known reservoir size. Equation 20 is graphically depicted in Fig. 1 in comparison with Pegram's exact results. This equation is suggested for preliminary estimates, but it should be applied with caution for the reasons explained in the next section.

Effects of Inflow Characteristics on Reservoir Size

As explained earlier, simplified design procedures such as that using Equation 20 are based on a number of abridging assumptions about inflows. Significant differences may appear if these assumptions are not valid. More specifically, what may cause significant departures from Equation 20 are hydrologic persistence, especially long-term, and the seasonal distribution of inflow and demand. Less significant differences are caused by the skewness of inflows and secondary inflows and losses.

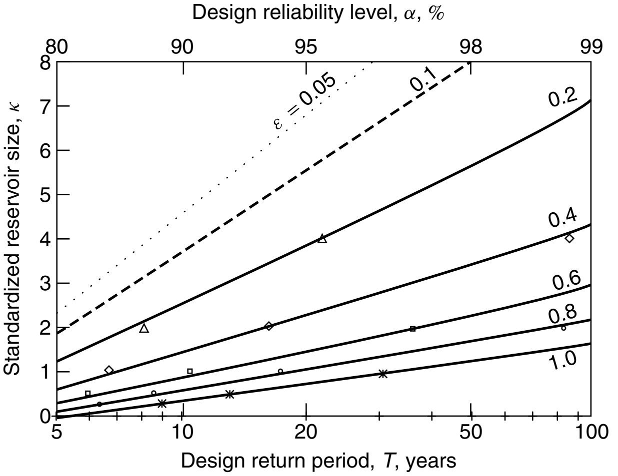

Figure 1. A simple representation of reservoir size–yield–reliability: Standardized reservoir size ($\kappa = K/\sigma$) required to achieve a certain drift ($\varepsilon = (\mu - \delta)/\sigma$) at a certain reliability level for independent inputs normally distributed. Lines are constructed from Equation 20, whereas plotted points are theoretical results by Pegram (1) for $\varepsilon = 0.2$ (triangles), 0.4 (diamonds), 0.6 (squares), 0.8 (circles), and 1.0 (stars).

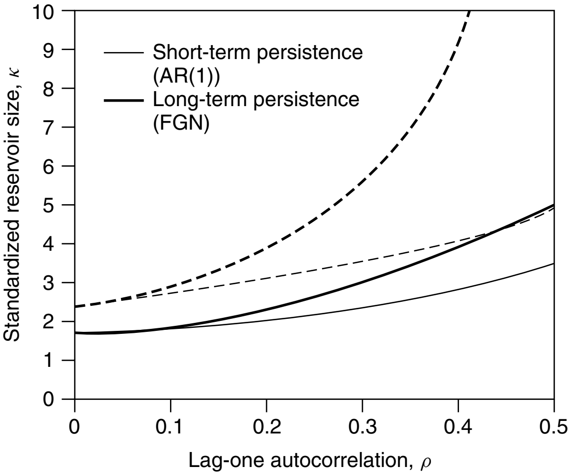

Figure 2. Effects of short-term and long-term hydrologic persistence of inflows on the standardized reservoir size κ required to achieve a drift $\varepsilon = 0.8$ at a reliability level $\alpha = 98\%$ (continuous lines) and a drift $\varepsilon = 0.2$ at a reliability level $\alpha = 90\%$ (dashed lines). Results are obtained by simulation.

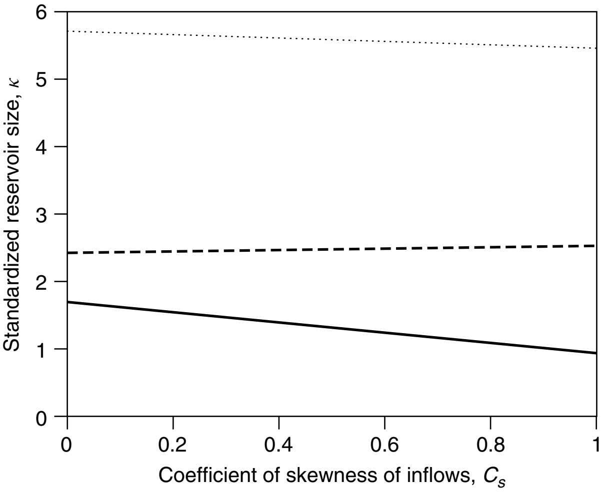

Figure 3. Effect of the skewness of inflows on the standardized reservoir size κ required to achieve a drift $\varepsilon = 0.8$ at a reliability level $\alpha = 98\%$ (continuous line), a drift $\varepsilon = 0.2$ at a reliability level $\alpha = 90\%$ (dashed line), and a drift $\varepsilon = 0.2$ at a reliability level $\alpha = 98\%$ (dotted line). Results are obtained by simulation using independent two-parameter gamma distributed inflows.

The effect of hydrologic persistence is demonstrated in Fig. 2, which depicts the standardized reservoir size κ required to achieve two combinations of drift and reliability ($\varepsilon = 0.8, \alpha = 98\%$ and $\varepsilon = 0.2, \alpha = 90\%$) versus the lag-one autocorrelation coefficient, ρ. Two cases of hydrologic persistence have been examined, short-term and long-term. In short-term persistence, it was assumed that the inflows follow the autoregressive process of order 1 [AR(1) or Markov], whereas in long-term persistence, it was assumed that they follow the fractional Gaussian noise (FGN) process with Hurst exponent $H = \ln(2 + 2\rho)/\ln 4$. Obviously, the effect of persistence is very significant, especially for long-term persistence and high demand (low drift ε). Thus, the required reservoir size for $\varepsilon = 0.2$ and $\alpha = 90\%$ is $c = 2.4\sigma$ when $\rho = 0$ ($H = 0.5$) and becomes 4 times larger ($c = 9.6\sigma$) when $\rho = 0.4$ ($H = 0.74$).

To quantify the effect of seasonal variation of inflow and demand, it is observed that in the worst case, the total annual inflow comes before the beginning of withdrawal (the "sudden" model). The reservoir dynamics in Equation 2 assumes that inflow and withdrawal are distributed evenly during a year (the "steady" model). If inflow precedes withdrawal, Equation 2 should be modified to read $S_t = \max[0, \min(S_{t-1} + X_t, c) - \delta_t]$. It is then easily determined that extra storage equal to δ is required in addition to that estimated from Equation 21. This worst case, however, is not very realistic, except in arid regions. In the real world, the extra storage capacity required (in addition to that estimated from Equation 21) is a percentage of δ. This can be as high as 50% for water supply reservoirs and 80% for irrigation reservoirs in semiarid regions.

The effect of the skewness of inflows is demonstrated in Fig. 3, which depicts the standardized reservoir size κ required to achieve three combinations of drift and reliability ($\varepsilon = 0.8, \alpha = 98\%$; $\varepsilon = 0.2, \alpha = 90\%$; and $\varepsilon = 0.2, \alpha = 98\%$) versus the coefficient of skewness, C_s. It can be observed that the effect of skewness is not significant; for low draft (high drift ε), this effect can be beneficial (lowering of required storage), but for high draft, the required storage is practically insensitive to skewness.

GENERALIZED SIMULATION PROCEDURE FOR RELIABILITY-BASED RESERVOIR DESIGN

The simplest general procedure for estimating the reservoir size–yield–reliability relationship in an accurate and detailed manner and for any arbitrary inflow characteristics is stochastic (Monte Carlo) simulation. A simplified simulation procedure is outlined as follows:

1. Generate a series of inflows X_t at an appropriate timescale (e.g., monthly) using an appropriate stochastic model (e.g., 17, 18; see also the entry Stochastic Simulation of Hydrosystems).
2. Assume a reservoir size c.
3. Calculate a series of reservoir storages using Equation 2.
4. Estimate the reliability using Equation 13.
5. Repeat steps 3–4 for different reservoir sizes.

This procedure is very simple to execute even in a tabulated form on a spreadsheet.

The drawback of stochastic simulation is that it requires a vast simulation length to find accurate results. It can be shown that the required number of simulated time steps (e.g., months) to estimate the occurrence-based failure probability β' at an acceptable error $\pm\varepsilon\beta'$ and confidence γ is

$$n = k(z_{(1+\gamma)/2}/\varepsilon)^2(1/\beta' - 1) \tag{22}$$

where k is the number of time steps per year and z_p is the p-quantile of the standard normal distribution. For instance, for $k = 12, \gamma = 95\%$ $(z_{(1+\gamma)/2} = 1.96), \varepsilon = 10\%$, and $\beta' = 0.01$, this yields $n = 456,000$ months (38,000 years). Today, this is not a major problem as the required computer time for such a simulation length can be less than one second in a common PC.

The simplified procedure described can be extended to a detailed simulation procedure, which includes, in addition to runoff, the precipitation, evaporation, and leakage of the reservoir. In this case, level–area–volume and level–leakage relationships are required to establish the functions $a(S)$ and $l(S)$ which yield reservoir area a and leakage l for any storage S (usually using interpolation from arrays of tabulated values). In this case, the runoff Q, has to be expressed in equivalent depth units, as are precipitation P and evaporation E. If f is the catchment area, the net inflow becomes

$$X_t = Q_t[f - a(S_{t-1})] + (P_t - E_t)a(S_{t-1}) - l(S_{t-1}) \tag{23}$$

where it was implicitly assumed that variations in the reservoir area and leakage within a time step are not large, so that $a(S_{t-1})$ and $l(S_{t-1})$ can be assumed representative of the entire period within a time step. When the detailed simulation procedure is employed, three time series, instead of one, have to be synthesized. Obviously, Q_t and P_t are cross-correlated, and therefore they cannot be generated independently of each other; a bivariate stochastic model is needed in this case. On the other hand, E_t can be generated independently of the other two series.

Obviously, the simulation procedure, either simplified or detailed, can be applied directly using historical, rather than synthesized inputs. However, due to short record length, the accuracy of results will not be satisfactory.

Extension of the simulation method, combined with optimization, for a multiple reservoir system can be found in References 19 and 20.

DESIGN OF ADDITIONAL STORAGE ZONES OF A RESERVOIR

The design of flood control storage is typically reliability-based but in a very different context from that described for active storage. A failure of the flood routing function of a reservoir is not a performance failure but a structural one: an overtopping of the dam due to a severe flood can result in collapse of the dam. Therefore much lower levels of probability of failure are adopted, of the order of $10^{-3} - 10^{-6}$. The typical steps here are (1) estimation of a design storm, based on statistical analysis of rainfall, for an appropriate return period ($10^3 - 10^6$ years); (2) estimation of the inflow hydrograph using an appropriate rainfall–runoff model; and (3) routing this hydrograph through the spillway and estimation of the outflow hydrograph and the maximum water level. Implicit assumptions in the entire procedure, like the assumption that the reservoir is full at the beginning of the flood, decrease the risk further. It must be noted that several procedures have been proposed that are supposedly risk-free. These are based on the so-called probable maximum precipitation concept. However, it has been argued that a risk-free procedure is an illusion and the value of probable maximum precipitation can be exceeded by a certain probability (e.g., of the order of 10^{-5}; 21).

The sizing of dead storage has been typically based on the expected sediment accumulation in the reservoir for a certain design period (e.g., of the order of 10^2 years). Because of the large design period and the accumulative character of this process, the approach to this problem is very different. Expected values, rather than probabilities, are involved in the calculations. The additional objectives that the dead storage serves, environmental protection (protection of the habitat of the reservoir during dry periods) and conservation of the quality of the landscape, have not been given special attention, until now, and have not been considered in the design procedure. One would expect that, some years after construction, these additional objectives would not be served adequately because the water in the dead volume would be reduced due to sediment accumulation. Fortunately, however, this has not been the case: the implicit assumption that sediments will reach the bottom of reservoir near the dam is not verified. Thus, in large reservoirs, sediment accumulation occurs mainly in the active zone of reservoir (near the entrance of the river to the reservoir) and much less in dead storage. The unfavorable consequence is the reduction of active storage. These problems need to be further investigated in future reservoir designs.

Acknowledgment
The review by the editor and the comments by G.G.S. Pegram are gratefully appreciated.

BIBLIOGRAPHY

1. Pegram, G.G.S. (1980). On reservoir reliability. *J. Hydrol.* **47**: 269–296.
2. Raudkivi, A.J. (1979). Hydrology. *An Advanced Introduction to Hydrological Processes and Modelling*. Pergamon Press, Oxford.

3. McMahon, T.A. (1993). Hydrologic design for water use, Ch. 27. In: *Handbook of Hydrology*. D. Maidment (Ed.). McGraw-Hill, New York.
4. Chow, V.T., Maidement, D.R., and Mays, L.W. (1988). *Applied Hydrology*. McGraw-Hill, New York.
5. Rippl, W. (1883). The capacity of storage reservoirs for water supply. *Proc. Inst. Civ. Eng.* **71**: 270–278.
6. Schultz, G.A. (1976). *Determination of deficiencies of the Rippl-diagram method for reservoir sizing by use of synthetically generated runoff data*, Proceedings XIIth Congress of ICOLD (International Commission on Large Dams). March/April 1976, Mexico City.
7. Feller, W. (1951). The asymptotic distribution of the range of sums of independent random variables. *Ann. Math. Stat.* **22**: 427.
8. Hurst, H.E. (1951). Long term storage capacities of reservoirs. *Trans. ASCE* **116**: 776–808.
9. Kottegoda, N.T. (1980). *Stochastic Water Resources Technology*. Macmillan Press, London.
10. Koutsoyiannis, D. (2002). The Hurst phenomenon and fractional Gaussian noise made easy. *Hydrol. Sci. J.* **47**(4): 573–596.
11. Mays, L.W. and Tung, Y.-K. (1992). *Hydrosystems Engineering and Management*. McGraw-Hill, New York.
12. Mays, L.W. (2001). *Water Resources Engineering*. John Wiley & Sons, New York, NY.
13. Moran, P.A.P. (1959). *The Theory of Storage*. Methuen, London.
14. Zsuffa, L. and Gálai, A. (1987). *Reservoir Sizing by Transition Probabilities: Theory, Methodology, Application*. Water Resources Publications, Littleton, CO, p. 186.
15. Gould, B.W. (1960). Water supply headworks storage estimation. *Aust. Civ. Eng. Constr.* **1**(9): 46–52.
16. McMahon, T.A. and Mein, R.G. (1986). *River and Reservoir Yield*. Water Resources Publications, Littleton, CO.
17. Koutsoyiannis, D. (2000). A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. *Water Resour. Res.* **36**(6): 1519–1533.
18. Koutsoyiannis, D. (2001). Coupling stochastic models of different time scales. *Water Resour. Res.* **37**(2): 379–392.
19. Nalbantis, I. and Koutsoyiannis, D. (1997). A parametric rule for planning and management of multiple reservoir systems. *Water Resour. Res.* **33**(9): 2165–2177.
21. Koutsoyiannis, D. and Economou, A. (2003). Evaluation of the parameterization-simulation-optimization approach for the control of reservoir systems. *Water Resour. Res.* **39**(6): 1170, 1–17.
21. Koutsoyiannis, D. (1999). A probabilistic view of Hershfield's method for estimating probable maximum precipitation. *Water Resour. Res.* **35**(4): 1313–1322.

LAKES

WALTER GELLER
UFZ-Centre for Environmental Research
Magdeburg, Germany

Lakes are waterbodies of considerable size that are larger than ponds. *Shallow lakes* are mixed by the wind from the surface to the ground, and the daylight reaches the bottom. *Stratified lakes* are deep enough to be vertically divided into a lighted surface-near *euphotic* zone, where photosynthesis of green plants is possible, and a deep-water *aphotic* zone, which is too dark for photosynthetic activities. In many lakes, thermal vertical stratification is observed as a warm layer near the surface, the epilimnion, and a colder layer below, the hypolimnion. The epilimnion is wind-exposed and well mixed, whereas the deep water layer of the hypolimnion is not included in the mixis process. Introductions to the ecology, hydrology, and limnology of lakes are available in textbooks, handbooks (1,2), and reviews (3–9).

ORIGIN OF LAKE BASINS

Lake basins containing standing waterbodies were formed by different processes; tectonic forces; volcanic, glacial, fluviatile activities; or by solution of easily dissolvable rock such as limestone. Hutchinson (1) ordered the lakes by their mode of origin into about 80 categories. Typical examples are listed by Meybeck (8,10), who also gives numbers of the types of origin, of size, and of age.

Tectonic lakes are found in geological rift valleys; the deepest are Lake Baikal (11) and Tanganyika (12). The Caspian Sea and the Aral Sea (13) are rest areas of the ancient Thetis Sea which was closed by tectonic movements, and they are the largest inland waters in area (374,000 km^2, 64,500 km^2). The oldest inland waters are found among tectonic lakes, some of these, it is estimated, are older than 20 to 30 million years.

Volcanic processes led to the origin of different lake types. *Maar lakes* emerged after singular volcanic events (steam explosions in contact zones of magma and groundwater). Lava dams or emerging volcanoes resulted in reservoir-like lakes, such as L. Kivu and Laguna del Laja (Chile). Large lakes were formed on lava plateaus (L. Myvatn, Iceland) and in *caldera* basins, whereas *crater lakes* are usually small. Lakes at active volcanoes are often heavily acidified by acidic brines rinsing into the lake water (14).

The main types of glacially formed lakes are (1) small *kar lakes* at the slopes of mountains at different altitudes, (2) *piedmont lakes* where glaciers from valleys reached the foreland at the foot of the mountains, and (3) extended *lake districts from low altitude glaciations* as on the Canadian Shield and the circum-Baltic area. Postglacial lakes of temperate climatic zones are only 8 to 12 thousand years old.

In the lowland area of large rivers, the flow of the running rivers often changes, leaving standing waters, oxbows or side waters, as *fluvial lakes*, which were parts of the rivers in the past or are permanently in open connection with the river system.

Lakes are more numerous as size decreases. Meybeck (8,10) gives the estimated worldwide number of lakes, ordered by size classes and by the origin of lake basins. A list of the deepest and largest lakes, by area and by water volume, is given by Herdendorf (15,16), including many uncertainties on the basic data of even the 50 most important inland waters of the world.

WATER BUDGET AND SALINITY OF LAKES

The oceans contain 83.5% of the world's water, the inland waters only 0.015%. Lakes contain a total water volume of 205,000 km^3; about 50% of this total volume is freshwater, on the one side, and saline water on the other side. The deepest and oldest lake, Baikal, has a volume of about 20% of the total freshwater volume of lakes and is the worldwide largest singular freshwater resource. The annual fluxes through the pools of the atmosphere, rivers, and groundwater are 496, 38, and 12 thousand km^3; exchange times are 10 days, 16 days, and 700 years, respectively. The volume of freshwater lakes is 100 thousand km^3; exchange times are 1 month to 500 years [data after (17,18)].

Lakes are standing waters that are usually connected with and included in the flow system of river catchments. The lakes within the catchment systems can be ordered by their degree of connection to the overall water flows (19). Lakes with intensive through-flow have high rates of flushing and water exchange and short retention times. The lake volume compared with the volume of all inflows per year gives the virtual *filling time*. The relation of the outflow and the lake volume gives the net exchange rate of the lake water. Including groundwater exchange, precipitation on the lake surface, and evaporation, the relation of all inputs to all losses of water gives the gross exchange rate; its reciprocal value is the *retention time* or average age of the water.

Depending on the given regional water balance, lake basins are permanently flushed, have an outflow, are in a steady state of the flow of water and of dissolved salts, and, thereby, contain freshwater. In areas with a negative water balance, the river catchments are *endorheic* systems, and terminal lake basins occur without outflows; here, the water finally evaporates, and the inflowing salts accumulate, leaving *saline lakes*. Saline lakes are typical of (semi)arid areas and contain nearly half of the world's lake water volume (20). Saline lakes (with salinities >3%) (21,22) are found in all climatic zones, warm, temperate, and cold regions, where negative water balances are given by large-scale climatic zones or by regional orographic conditions (rain-lee east of the Rocky Mountains in North America, east of the Andes in South America, or in the Iranian basin). Such endorheic river catchments, in which the water flow ends in terminal lakes, are the Jordan River system with the Dead Sea, the system of Lake Titicaca with Lake Poopó as the terminal saline lake, and the system of Lake Tahoe and Pyramid Lake in North America. In Central Australia, Lake Eyre is a large, endorheic saline basin. The highest possible salt concentrations are found in the Dead Sea (~300 g/L) and in Deep Lake (Antarctica) which never freezes; the water reaches −14.5 °C during the winter (22).

STANDARD COMPOSITION OF FRESHWATERS

Freshwaters, soft waters as well as hard waters, are chemically dominated by the carbonate system, $CO_3{}^{2-}$ and $HCO_3{}^{-}$ as anions, and Ca^{2+} and Mg^{2+} as cations; together they are the main constituents and the buffering system of neutral freshwaters. This is caused by rainwater that is weakly acidic from CO_2 and, thereby, dissolves the carbonatic minerals of soils and rock (23,24). In areas poor in carbonates, the water contains low concentrations of dissolved salts, and in other areas of dominating limestone, the water is rich in dissolved carbonates. The qualitatively similar composition of freshwaters is described as a worldwide "standard water composition" (25,26).

The gases in air, 78.09% N_2, 20.95% O_2, and 0.03% CO_2, are in physical equilibrium between air and water. At low temperatures, higher concentrations of gases dissolve in the water than at high temperature. The maximum content of dissolved oxygen and CO_2 in water is 14.5 (8.9) mg O_2/L and 1.005 (0.51) mg CO_2/L at 0 °C (20 °C).

STRATIFICATION AND MIXING

Lakes are usually stratified into an upper layer, the epilimnion, which is well mixed by the wind, and a deep waterbody, the hypolimnion, which is not included in the mixing process. Stratification occurs seasonally in the temperate climatic zones, or in a diurnal–nocturnal rhythm in the tropical zones (Fig. 1). Ice covers block energy inputs by wind into the water column during the winter. The regime of stratification and mixis of the water column, as given by the climate, defines the types of lakes as (1) *monomictic*, (2) *dimictic*, or (3) *polymictic*, according to the number of periods of total mixing (*holomixis*) per year. Tropical lakes at low altitude are polymictic; they mix during the night and stratify during the day. Dimictic lakes have a stable stratification with a warm epilimnion during the summer and an ice cover during the winter; they show holomixis during the two seasonal periods of homothermy. One holomixis occurs in spring after ice melting when the surface water has the same temperature as the hypolimnion, usually 4 °C, and a second during late autumn before freezing. Monomictic lakes mix only once per year. The warm-monomictic lakes of oceanic-temperate regions never freeze and show holomixis during the winter. Cold-monomictic lakes are found in polar climatic regions; they are covered by ice during large parts of the year and melt only in summer when the water column reaches 4 °C and mixes from the surface to the ground.

The stability of stratification of a warm epilimnion and a cold hypolimnion determines whether wind forces mix only the uppermost layers, or the entire water column from the surface to the bottom. Stratification stability depends on the density gradient between the epilimnetic and the hypolimnetic partial waterbodies. Indicative numbers are the *Wedderburn Number* or the *Lake Number*, which, at given threshold values, predict optional holomixis with the next wind-forcing (27).

Climatic Zones and Lake Temperatures

Lakes are integrators of the regional climate and weather development (28). The surface temperatures follow the actual air temperature with a time delay, and lakes of the same region show time courses of similar surface temperatures. The deep water of lakes in polar and

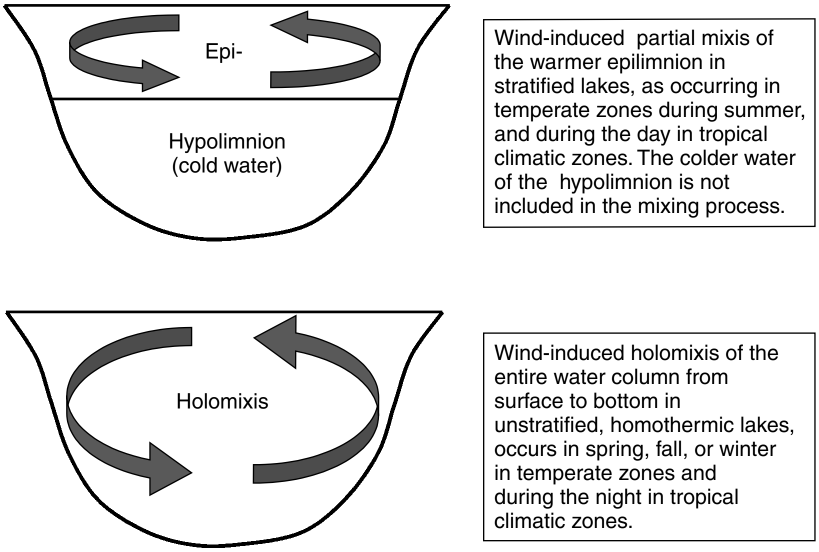

Wind-induced partial mixis of the warmer epilimnion in stratified lakes, as occurring in temperate zones during summer, and during the day in tropical climatic zones. The colder water of the hypolimnion is not included in the mixing process.

Wind-induced holomixis of the entire water column from surface to bottom in unstratified, homothermic lakes, occurs in spring, fall, or winter in temperate zones and during the night in tropical climatic zones.

Figure 1. Seasonal or diurnal succession of stratification into warm epilimnion and cold hypolimnion and unstratified water column when the water is homothermic with complete overturn of the water from surface to ground (holomixis).

temperate continental climatic zones is at 4 °C, the temperature of water at its maximum density. The deep water of warm-monomictic lakes of temperate oceanic climatic zones in the Southern Hemisphere, in New Zealand and South America and also in western and southern Europe, has a temperature above 4 °C. The hypolimnetic temperatures remain from the last holomixis at the end of winter, and changes reflect long-term climatic conditions (29).

Waves and Oscillations

Driven by wind, waves emerge on the surfaces of lakes; their heights can be approximated by $h = 0.105\ (\text{fetch})^{0.5}$ (wave height and fetch given in cm; the *fetch* is the maximum length of the wind blowing across an open lake surface) (9). Less peculiar are oscillations along the lake axis, which have amplitudes of few centimeters at the lake surface (*seiches*). Their frequencies depend on the length of the lake axis, between minutes in small lakes and one day in large lakes. At the boundary layer between the epi- and hypolimnion, however, internal oscillations can reach wave amplitudes of 10 m or more. The periods of these internal oscillations are slower and reach 1 month in the largest lakes, such as Baikal or Tanganyika (1,27).

ECOSYSTEMS OF LAKES

The ecosystems of lakes belong to the best described and understood ecosystems because the composition of the community is limited in number of species and the populations are limited by the lake size. The seasonal development of the populations, their change in species composition, and the growth and decrease of the populations are well investigated and analyzed in terms of numbers of individuals as well as in terms of energetics and species composition (30).

Lakes have characteristic compartments as partial habitats of their ecosystems. The ecosystem of the pelagic zone of lakes is dominated by planktonic microorganisms, larger zooplankton, and fish. The phytoplankton is autotrophic and limited to the euphotic layers near the surface where light intensities are sufficient for photosynthesis and oxygen production. The limit between the *euphotic* and the *aphotic zone* is usually at depths where the light intensity decreases to <1% of that at the lake surface.

The *pelagic zone* is the body of free water of the lake, the *littoral zone* is the euphotic part at the lakeshore, and the *profundal* is the aphotic dark layer of the *benthic zone* below the littoral. The benthic community outside the littoral zone, the profundal, consists of the organisms living in and on the sediments, microorganisms, macrozoobenthos, and benthic fish species. Below the light limit for photosynthesis, the species of the benthic community, macro- and microorganisms, are heterotrophic and consumers of oxygen. All particles in the water eventually sink to the lake bottom, form annual layers of growing sediments (varves), and leave an archive of residuals. Therefore, the history and the development of the lake can be reconstructed from the microlayers of a sediment core, with respect to eutrophication and nutrient levels and impacts of climatic changes.

The littoral zone is the euphotic part of the benthic zone. The living community consists of autotrophic macrophytes and algae and of heterotrophic animals and microorganisms. The belt of macrophytes appears as an interface between land and water; wetland plants are on the land side of the belt, followed by plants rooting in the water and emerging in the air, swimming plants, and lastly submerged plants that cover the light bottom of the deeper littoral zone. The partial ecosystem of the littoral zone connects the land and water, and simultaneously, as a biologically reactive zone, retains and changes the flow of nutrients and other allochthonous inputs of matter into the lake's pelagic zone.

Trophic State and Productivity

Trophy is a measure of the level of nutrients and of the biological productivity resulting from nutrients supplies.

The organismic productivity of aquatic ecosystems, measurable as the formation of biomass, is limited by the availability of the rarest nutrient element (C,N,P,S,Fe,...). In temperate zones, phosphorus is usually the least available element because there is no gaseous transport form through the air and P is tightly sorbed on soil particles and retained on the land. The production of organic substance is limited by the availability of phosphorus, so the trophy (productivity) of lakes can be related to the P-content of the water at the start of the growing season (31,32). The process of growth needs a proportion of 106C:16N:1P as the atomic ratio of biomass (33,34). The primary production of phytoplankton and macrophytes—depending on P—reaches annual rates of 10 to 50 g C/m^2 of organic carbon in nutrient-poor lakes and more than 500 g C/m^2 in very nutrient-rich systems.

The biomass produced sediments to the lake bottom, where this dead organic substance is decayed by oxygen-consuming bacteria. Depending on the trophic state, the oxygen conditions in the hypolimnic water body reach different end points in late summer and fall. In oligotrophic lakes, the nutrients and productivity are low, as well as the following decomposition and oxygen losses in deep water. In eutrophic lakes, however, the nutrient supply and the resulting productivity are high. After decomposition in deep water layers, the oxygen content in the hypolimnion may reach critically low limits in late summer. The consequences of high nutrient supplies are mitigated in deep lakes with high hypolimnic volumes but are more severe in shallower lakes with small hypolimnic volumes and small oxygen reserves therein (Fig. 2).

Food Chains, Filtrators, and Size Spectra in plankton

Within the pelagic zone of lakes, the species of the plankton community can be structured by the food chain from the primary producers, the green algae, to herbivores, zooplankton species that feed on these algae, and carnivores that feed on herbivorous zooplankton. The efficiency of energy transfer from the lower to the higher trophic levels is about 10% from plants to animals and up to 30% from animal prey to their predators. Because there are losses of unused particles and dissolved organic matter (DOM) on all levels, recycling by bacteria as decomposers of DOM is important. Based on the bacteria as food, leading via bacteria-feeding protozoans to larger predators, the “lost” DOM is recycled again into the “normal” food chain by the “microbial loop.” This secondary food chain is important for the overall efficiency of the ecosystem as a processor of energy and its ability to recycle nutrient elements.

In oligotrophic systems, the organisms, prey and predators, are present in low concentrations, and a predator has to find and to grasp each single prey individually. The prey in such food chains is usually smaller than its predator by a factor of 10 to 20 by body length. Therefore, in oligotrophic planktonic systems, the food chain from green algae, via zooplankton to fish needs four to five subsequent trophic levels. The overall transfer efficiency with its losses at any transfer is small. In the more productive systems of meso- to eutrophic lakes, however, the concentrations of organisms from algae and bacteria to zooplankton and fish are high, and additional, new types of filtering predators appear. Filtrators use a filtering system, comb-like sieves in cladocerans (*Daphnia*), fine gill structures in filter-feeding fish (Coregonids), to collect large numbers of very small prey particles. The linear size relation between prey and filter-feeding predators is 100 × to 1000×, and, thereby, the length of the food chain becomes short, three levels from algae to fish. In temperate lakes, we find *Daphnia* as the keystone filtrator on the level of herbivores all over the world. The filtering, zooplankton-feeding fish of the genus

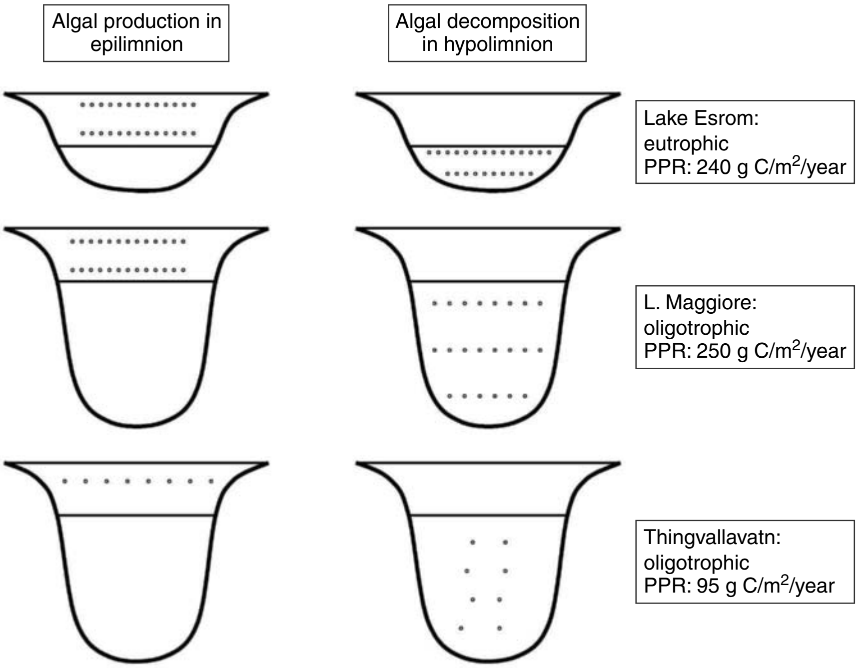

Figure 2. The relationship between the supply and demand of oxygen in deep water is determined by the intensity of the respective decomposition rate and the volume of the hypolimnetic waterbody of a given lake. The oxygen demand with algal decomposition consumes a large percentage of the supply in the small hypolimnic waterbody of L. Esrom, it is medium in L. Maggiore with its big hypolimnion, and it is unimportant in Thingvallavatn because the mass of algae is small and the big volume of the hypolimnion has a big oxygen reserve. L. Esrom is eutrophic, L. Maggiore has the same annual primary production (PPR) of algae, but is oligo-/mesotrophic because of its morphology. Thingvallavatn is a typical oligotrophic lake because of low primary production and its depth (redrawn from Reference 35).

Coregonus are found only in the Northern Hemisphere. In the ancient Lake Tanganyika, an evolution of a “new” filtrator occurred, *Limnotrissa*, where a zooplankton-feeding sardine, led to very high fish production. In other ancient tropical lakes, herbivorous *Tilapia* species appeared that filter-feed on algae.

The food web of any ecosystem might be described on the level of species composition and populations, according to its taxonomic structure. The system, however, might be alternatively viewed with an ataxonomic approach and explained functionally as a system of trophic levels linked by energy transfers: the “trophic-dynamic system” (34,36).

Another way to describe the planktonic food web and to explain some of its functions and properties is to view its size structure and the continuity of sizes as a spectrum. Investigations of the entire planktonic system revealed the plankton as a spectrum of size classes, in which the biomass of all size classes is equal. Many small organisms are present with the same sum of biomass, as are the fewer organisms of larger sizes. The regular pattern of body sizes over all organisms of plankton (37,38) corresponds to growth and physiological rates in a reciprocal way, fast overturn in the pools of small organisms and slow overturn in the big ones.

REGIONAL LAKE LIMNOLOGY

Glacial Lakes

The glaciated areas of the Canadian Shield, the Scandinavian mountains, and the adjacent sedimentary platforms of the circum-Baltic countries have been eroded by glaciation that left ten thousands of large and small lake basins. The large basins of the Hudson Bay and the Baltic Sea are corresponding central basins. The Baltic Sea was intermittently an inland lake and then became a brackish shelf sea. The surrounding areas from Finland to northern Germany have extended lake districts that cover large parts of the land. The “limnic ratio” is a measure of the percentage of the recent lake area of the deglaciated land area (8): world total deglaciated area: 6.8%; Scandinavian Shield: 12.2%; Finland: 9.4%; Norway: 13.9%; Sweden: 8.55%; South Baltic Platform: 2.2%; Canadian Shield:10.3%; and Patagonia: 3.0%. The lakes of this origin are the largest in Europe, Lakes Ladoga and Onega in western Russia and the large southern Swedish lakes, and in North America, the Laurentian Great Lakes and the large lakes of Canada (Table 1).

Formerly glaciated mountain districts are found on all continents (Table 1). The fjord lakes and foothill lakes in the deglaciated mountain areas are found in the Northern and the Southern Hemispheres of the temperate climatic regions in Scandinavia, the loghs/lochs of Ireland and Scotland, the English Lake District, the Alps, New Zealand, the Patagonian mountain areas of Chile and Argentina, and the Central and Northern Asian mountains. *Fjord lakes* are the deepest glacial lakes. Their mountain valley erosion basins have peculiarly steep borders and flat bottoms with depths of 500 m or more. Where the glacial erosions from valleys extend into the

Table 1. Large Glacial Lakes on Shields, Sedimentary Platforms, and in Mountain Areas

Glacial Lakes	Surface Area, km^2	Max. Depth, m	Water Volume, km^3	Retention Time Years	Age Origin, 10^3 y	Data Source
Canadian Shield lakes						(8,10,49)
Superior	82,100	405	12,200	180	8, glacial	
Huron	59,500	223	3,500	48	8, glacial	
Michigan	57,750	281	4,800	31	8, glacial	
Great Bear	31,326	446	2,200		glacial	
Great Slave	28,568	614	2,088		tectonic, glacial	
Andean lakes						(39,40)
Argentino	1,410	500				
Gral. Carrera, Chile/Buenos Aires, Arg.	1,892	586	>500		tectonic, glacial	
Nahuel Huapi, Arg.		464				
Llanquihue, Chile	871	317	159	74		
New Zealand						(74)
Taupo	616	165	59	10.6	volcanic, glacial	
Waikitipu	289	380				
Wanaka	180	311				
Baltic lakes						(8,10,72)
Ladoga, Russia	18,130		908		>8, baltic	
Onega, Russia	9,700		292		>8, baltic	
Vänern, Sweden	5,648		152	8.8	6.5, baltic	
Vättern, Sweden	1,912	128	74	58	6.5, baltic	
Mälaren, Sweden	1,096		14	2.3	6.5, baltic	
Alpine piedmont lakes						(28,54)
Geneva, W.-Alps	580	310	89.9		piedmont	
Constance, N.-Alps	472	253	47.7	4.15	piedmont	
Garda, S.-Alps	370	346	49	26.6	piedmont	
Maggiore, S.-Alps	213	370	37.5	4.1	piedmont	
Como, S.-Alps	146	410	22.5	4.5	piedmont	

foothill areas of the foreland, the characteristic *piedmont lakes* cover large areas, but they are shallow compared with fjord lakes or fjord-like parts of the same lake (e.g., L. Garda, L. Argentino, L. Buenos Aires, show both, deep fjord-like parts and large, shallow foothill areas).

Lakes in the (sub-)polar regions are covered by ice permanently or during the largest part of the year. Lake Vanda (Antarctica) has a permanent, clear, ice cover. The deepest water layers are stabilized by saline inputs and trap the down-welling light energy, leading to a temperature of 25 °C above the bottom (63 m) (22). Some large lakes in the Antarctic were found by satellite radar imaging under several thousand meters of glacial ice cover. The largest of these lakes, L. Vostok, is the size of Lake Ontario (39).

TROPICAL AND ANCIENT LAKES

The lakes of the (sub-)tropical climatic zone are described in textbooks by Beadle (42) and Serruya (43). The most important difference from temperate lakes is the diurnal variability of temperatures that are as large as the seasonal variability between summer and winter in temperate lakes. The restricted time of nocturnal holomixis presumably limits the maximum depth of mixis in deep tropical lakes. Therefore, the very deep lakes Tanganyika and Malawi are meromictic, and, consequently, have anoxic, deep waterbodies. The deep African Rift Valley Lakes are ancient lakes and, therefore, have peculiarities beyond their tropical sites (12).

The largest and deepest lakes of the world are also the oldest lakes: Baikal, Tanganyika, Malawi, and Issik-Kul. The most important ancient lakes are listed in Table 2. The old lakes originate from the Tertiary and have a long development with changing water levels that could be reconstructed by paleolimnological investigations. The best investigated ancient lake is L. Biwa, whose history is based on sediment cores that are 200 and 1000 m long, covering 0.5 to 1 million years (44). The oldest and deepest of the ancient lakes gave time to their biota for evolutionary processes which led to adaptive radiations of—mostly endemic—species, found 10 to 100× more numerously than in young lakes. There are >400 species of Cichlid fish species in Tanganyika and Malawi and >200 species of Gammarid crustacean species in Baikal. The ecosystem structure is severely changed by new keystone species in the food web and new top carnivores: in Lake Baikal the pelagic Gammarid *Macrohectopus* and Cottoidei fish and the freshwater seal *Phoca sibirica*. In Lake Tanganyika, the species *Limnotrissa* directs the food chain to high production of fish, whereas in the other large African lakes the insect larva *Chaoborus* plays the role of the main zooplankton feeder. This species leaves the lake when it becomes an adult insect, so the branch of the food chain is then lost for the lake ecosystem.

MAN-MADE LAKES: RESERVOIRS AND PIT LAKES

Man-made lakes are built in great numbers, ten thousands of reservoirs and pit lakes in the voids of surface mining. They show limnological peculiarities that come primarily from the management regime for reservoirs and from geogenic impacts on water chemistry for pit lakes in coal and ore mining districts. The state of reservoir building and the problems are presented by the World Commission on Dams (http://www.dams.org). Reviews and other comprehensive scientific publications on mining lakes are in an initial phase (50).

REGIONAL PROGRAMS FOR PROTECTING LAKES AND COLLECTING DATA

The protection of lakes and other surface waters is different regionally. The Environmental Protection Agency (EPA) has set standards in North America, and the

Table 2. Basic Data on Ancient Lakes[a]

	Surface Area, km^2	Maximum Depth m	Water Volume, km^3	Age & Origin, Years
African Rift Valley				
Tanganyika	32,900	1471	18,900	20 mio, tectonic
Malawi	22,490	709	6,140	2 mio, tectonic
Victoria	68,460	92	2,700	20,000, tectonic
Asia				
Dead Sea	1,020	433	188	tectonic
Caspian	374,000	1025	78,200	>5 mio, tectonic
Baikal	31,500	1637	23,000	20 mio, tectonic
Aral	64,100		1,020	>5 mio, tectonic
Issyk-Kul	6,240	702	1,730	25 mio, tectonic
Biwa	681	104	27.6	2 mio, tectonic
N. America				
Tahoe	500	501	156	2 mio, tectonic
S. America				
Maracaibo	13,010		280	36 mio, tectonic
Titicaca	8,562	284	903	2.8 mio, tectonic

[a] References (8,44–49).

European Union in the EU countries. The regulations of the EU-Water Framework Directive have initiated many activities to register, describe, and improve surface waters and their state in Europe (see documents of the European Environmental Agency (EEA) and EU-Commission). There are data collections on world lakes for some regions and for selected countries: ILEC, EEA and EU-documents, EPA-documents (8,12,13,15,16,21,42,51–71) (31,32,72–75). For a series of important single lakes, there are monographic descriptions (11,76–80).

BIBLIOGRAPHY

1. Hutchinson, G.E. (1957). *A Treatise on Limnology. 1: Geography, Physics, and Chemistry*. John Wiley & Sons, New York.
2. Hutchinson, G.E. (1967). *A Treatise on Limnology. 2: Introduction to Lake Biology and the Limnoplankton*. John Wiley & Sons, New York.
3. Hutchinson, G.E. (1975). *A Treatise on Limnology. 3: Limnological Botany*. John Wiley & Sons, New York.
4. Hutchinson, G.E. (1993). *A Treatise on Limnology. 4: The Zoobenthos*. John Wiley & Sons, New York.
5. Le Cren, E.D. and Lowe-McConnell, R.H. (1980). *The Functioning of Freshwater Ecosystems. International Biological Programme*; 22. Cambridge University Press, Cambridge, UK.
6. Margalef, R. (1983). *Limnología*. Omega (Ed.). Barcelona, Spain.
7. Taub, F.B. (1984). *Lakes and Reservoirs. Ecosystems of the World*. Elsevier Science, Amsterdam.
8. Meybeck, M. (1995). Les lacs et leur bassins. In: *Limnologie générale*. R. Pourriot and M. Meybeck (Eds.). Masson, Paris, pp. 6–59.
9. Wetzel, R.G. (2001). *Limnology—Lake and River Ecosystems*, 3rd Edn. Academic Press, San Diego.
10. Meybeck, M. (1995). *Physics and Chemistry of Lakes*. A. Lerman, D.M. Imboden, and J.R. Gat (Eds.). Springer, Berlin, New York, pp. 1–36.
11. Kozhova, O.M. and Izmest'eva, L.R. (1998). *Lake Baikal*, 2nd Edn. Biology of Inland Waters, Backhuys, Leiden, Germany.
12. Johnson, T.C. and Odada, E.O. (1996). *The Limnology, Climatology and Paleoclimatology of the East African Lakes*. OPA, Amsterdam.
13. Mandych, A.F. (1995). *Enclosed Seas and Large Lakes of Eastern Europe and Middle Asia*. SPB, Amsterdam.
14. Varekamp, J.C., Pasternack, G.B., and Rowe, G.L., Jr. (2000). *J. Volcanology Geothermal Res.* **97**: 161–179.
15. Herdendorf, C.E. (1982). *J. Great Lakes Res.* **8**: 379–412.
16. Herdendorf, C.E. (1990). *Large Lakes, Ecological Structure and Function*. M.M. Tilzer, and C. Serruya (Eds.). Springer, Berlin, New York, pp. 3–38.
17. Schwarz, H.E., Emel, J., Dickens, W.J., Rogers, P., and Thompson, J. (1990). *The Earth as Transformed by Human Action*. B.L. Turner, W.C. Clark, R.W. Kates, J.F. Richards, J.T. Mathews, and W.B. Meyer (Eds.). Cambridge University Press, Cambridge, UK, pp. 253–270.
18. Lvovitch, M.I. (1973). *Trans. Am. Geophys. Union* **54**: 28–42.
19. Magnuson, J.J. and Kratz, T.K. (2000). *Verh. Int. Verein. Limnol.* **27**: 74–87.
20. Williams, W.D. (1996). *Verh. Int. Verein. Limnol.* **26**: 61–79.
21. Hammer, U.T. (1986). *Saline Lake Ecosystems of the World*. Junk, Dordrecht, The Netherlands.
22. Hobbie, J.E. (1984). *Lakes and Reservoirs*. F.B. Taub (Ed.). Ecosystems of the World. Elsevier, Amsterdam, pp. 63–105.
23. Stumm, W. and Schnoor, J.L. (1995). *Physics and Chemistry of Lakes*. A. Lerman, D.M. Imboden, and J.R. Gat (Eds.). Springer, Berlin, New York. pp. 185–215.
24. Stumm, W. and Morgan, J.J. (1995). *Aquatic Chemistry 3*. John Wiley & Sons, New York.
25. Rodhe, W. (1949). *Verh. Int. Verein. Limnol.* **10**: 377–386.
26. Livingstone, D.A. (1973). *Data of Geochemistry 6*. Prof. Pap. U.S. Geological Survey, Denver, CO.
27. Imboden, D.M. and Wüest, A. (1995). *Physics and chemistry of lakes*. A. Lerman, D.M. Imboden, and J.R. Gat (Eds.). Springer, Berlin, New York, pp. 83–138.
28. Livingstone, D.M. (1993). *Verh. Int. Verein. Limnol.* **25**: 75–81.
29. Ambrosetti, W. and Barbanti, L. (1999). *J. Limnol.* **58**: 1–9.
30. Reynolds, C.S. (1997). *Vegetation Processes in the Pelagic: A Model for Ecosystem Theory. Excellence in Ecology*, No. 9. Ecology Institute, Oldendorf, Germany.
31. Vollenweider, R.A. (1980). *OECD Eutrophication Programme. The Canadian Project*. OECD Eutrophication Programme, Burlington, VT.
32. Vollenweider, R.A. and Kerekes, J. (1980). *OECD Eutrophication Programme. Synthesis Report*. OECD Eutrophication Programme, Burlington, VT.
33. Parsons, T.R., Takahasi, M., and Hargrave, B. (1977). *Biological Oceanographic Processes 2*. Pergamon Press, New York.
34. Odum, E.P. (1971). *Fundamentals of Ecology* 3. W.B. Saunders, Philadephia.
35. Jónasson, P. (1996). *Verh. Internat. Verein. Limnol.* **26**: 1–33.
36. Lindeman, R.L. (1942). *Ecology* **23**: 339–418.
37. Sheldon, R.W., Prakash, A., and Sutcliffe, W.H. (1972). *Limnol. Oceanogr.* **17**: 327–340.
38. Gaedke, U. (1992). *Limnol. Oceanogr.* **37**: 1202–1220.
39. Bell, R.E. and Karl, D.M. (1998). *Lake Vostok Workshop—Final Report*. NSF, Washington, DC.
40. Geller, W. (1992). *Aquatic Sci.* **54**: 37–57.
41. Murdie, R.E., Pugh, D.T., and Styles, P. (1999). *Geo-Marine Lett.* **18**: 315–320.
42. Beadle, L.C. (1981) *The Inland Waters of Tropical Africa*. Longman, London.
43. Serruya, C. and Pollingher, U. (1983). *Lakes of the Warm Belt*. Cambridge University Press, New York.
44. Horie, S. (1984). *Lake Biwa*. Monographiae Biologicae. W. Junk, Dordrecht, The Netherlands.
45. Kawanabe, H., Coulter, G.W., and Roosevelt, A.C. (1999). *Ancient Lakes*. Kenobi Productions, Ghent, UK.
46. Kira, T. (1993). *Data Book of World Lake Environments. 1: Asia and Oceania*. International Lake Environment Committee; UNEP, Kusatsu.
47. Kira, T. (1994). *Data Book of World Lake Environments. 2: Africa and Europe*. International Lake Environment Committee; UNEP, Kusatsu.
48. Kira, T. (1995). *Data Book of World Lake Environments. 3: The Americas*. International Lake Environment Committee; UNEP, Kusatsu.
49. Kozhova, O.M. and Izmest'eva, L.R. (1998). *Lake Baikal*, 2nd Edn. *Biology of Inland Waters*. Backhuys, Leiden.
50. Castro, J.M. and Moore, J.N. (2000). *Environ. Geol.* **39**: 1254–1260.

51. Allan, R.J., Dickman, M., Gray, C.B., and Cromie, V. (1994). *The Book of Canadian Lakes. Monograph Series* No. 3. The Canadian Association on Water Quality, Gloucester, Ontario.
52. Boulton, A.J. and Brock, M.A. (1999). *Australian Freshwater Ecology*. Gleneagles, Glen Osmond, Australia.
53. Evans, M.S. (2000). *Aquatic Ecosystem Health Manage.* **3**: 65–79.
54. Fricker, Hj. (1980). *OECD Eutrophication Programme. Regional Project Alpine Lakes*. Swiss Federal Board for Environmental Protection, Bern.
55. IARH (2003). *National data book of Argentine lakes and lagoons and their environment*. (preliminary report). Istituto Argentino de Recursos Hidrocos, Buenos Aires.
56. ILEC (1988). *Data Book of World Lake Environments. A Survey of the State of World Lakes*. Otsu, Japan.
57. ILEC (1994). *The Pollution of Lakes and Reservoirs. UNEP Environment Library* 12. UNEP, Nairobi.
58. Jin, X. (1995). *Lakes in China*. China Open Press, Beijing.
59. Jolly, V.H. and Brown, J.M. (1975). *New Zealand Lakes*. Auckland University Press, Oxford University Press, Auckland, Oxford.
60. Kawanabe, H., Coulter, G.W., and Roosevelt, A.C. (1999). *Ancient Lakes*. Kenobi Productions, Ghent, UK.
61. Kira, T. (1993). *Data Book of World Lake Environments. 1: Asia and Oceania*. International Lake Environment Committee; UNEP, Kusatsu.
62. Kira, T. (1994). *Data Book of World Lake Environments. 2: Africa and Europe*. International Lake Environment Committee; UNEP, Kusatsu.
63. Kira, T. (1995). *Data Book of World Lake Environments. 3: The Americas*. International Lake Environment Committee; UNEP, Kusatsu.
64. Kristensen, P. and Hansen, H.O. (1994). *European Rivers and Lakes. EEA Environmental Monographs*; 1. European Environmental Agency, Copenhagen.
65. Maitland, P.S., Boon, P.J., and McLusky, D.S. (2002). *The Freshwaters of Scotland*. John Wiley, Chichester, UK.
66. Moriarty, Ch. (2002). *Studies of Irish Rivers and Lakes*. Marine Institute, Dublin.
67. Munawar, M. and Hecky, R.E. (2001). *The Great Lakes of the World (GLOW) Food-Web, Health and Integrity*. Backhuys, Leiden.
68. Rast, W. and Lee, G.F. (1978). *OECD Eutrophication Programme. The North American Project*. OECD Eutrophication Programme. EPA, Corvallis, OR.
69. Robertson, A. and Scavia, D. (1984). *Lakes and Reservoirs. Ecosystems of the World*. F.B. Taub (Ed.). Elsevier, Amsterdam, pp. 135–176.
70. Tilzer, M.M. and Serruya, C. (1990). *Large Lakes—Ecological Structure and Function*. Springer, Berlin, New York.
71. Taub, F.B. (1984). *Lakes and Reservoirs. Ecosystems of the World*. Elsevier Science, Amsterdam.
72. Willén, E. (1984). *Lakes and Reservoirs. Ecosystems of the World*. F.B. Taub (Ed.). Elsevier, Amsterdam, pp. 107–134.
73. Williams, W.D. (1984). *Lakes and Reservoirs Ecosystems of the World*. F.B. Taub (Ed.). Elsevier, Amsterdam, pp. 499–520.
74. Viner, A.B. (1987). *Inland Waters of New Zealand*. DSIR, Wellington.
75. Zhadin, V.I. and Gerd, S.V. (1963). *Fauna and Flora of the Rivers, Lakes and Reservoirs of the U.S.S.R.* Translation 1963. Israel Progr. Sci. Transl. Ltd., Jerusalem.
76. Dejoux, C. and Iltis, A. (1992). *Lake Titicaca*. Monographiae Biologicae. Kluwer Academic, Dordrecht, The Netherlands.
77. Jónasson, P. (1992). *Thingvallavatn*. OIKOS, Odense.
78. Létolle, R. and Mainguet, M. (1996). *Der Aralsee*. Springer, Berlin, New York.
79. Saijo, Y. and Hayashi, H. (2001). *Lake Kizaki*. Backhuys, Leiden.
80. Wood, R.B. and Smith, R.V. (2002). *Lough Neagh*. Kluwer Academic, Dordrecht, The Netherlands.

THE THEORY OF ALTERNATIVE STABLE STATES IN SHALLOW LAKE ECOSYSTEMS

ADRIAN E. WILLIAMS
APEM Ltd.
Manchester, United Kingdom

OVERVIEW

The idea of alternative stable states was first proposed in the late 1960s (1), and within ecological communities, it was described mathematically in the 1970s (2,3). The contemporary view of shallow lakes is that two alternative stable states can exist over a range of nutrient concentrations: a clear water state dominated by submerged macrophytes and a turbid phytoplankton-dominated state (4–8). Unique states may be found at either end of the nutrient continuum, but over a wide range of nutrient concentrations (<50 to several thousand $\mu g\ l^{-1}$ TP), either of these two states can exist, stabilized by a number of buffer mechanisms (5,6,9–11).

A “ball and cup” diagram is a simple way to visualize this idea (Fig. 1). The position of the ecological system (the ball) can change between the two states and settles at the bottom of the cup (stable equilibrium). The slope of the cup determines both the speed and the direction of the system change to a stable state. At either low or high

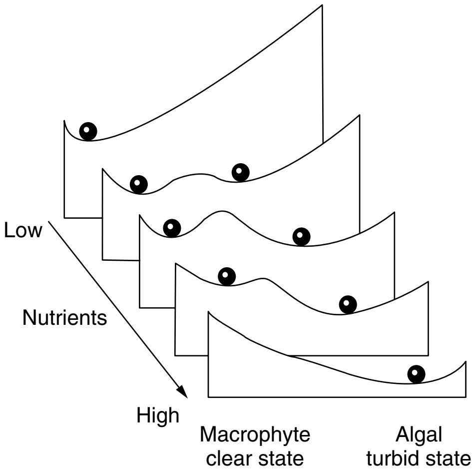

Figure 1. “Ball and cup” schematic of the potential for two stable states (clear water macrophyte-dominated versus turbid algal-dominated) to exist within a shallow lake under a wide range of nutrient conditions.

nutrient levels, one of the two stable states exists—the ball can only be in one location because of the slope of the horizon. However, between these points the system can rest in either cup (state). The ease with which one stable state can be switched to the other is indicated by the position of the two balls with respect to one another and the slopes that separate them. Therefore, if the starting position is a pristine, low-nutrient (oligotrophic), macrophyte-dominated, clear water lake, the nutrient levels can increase substantially before the only possible scenario is a turbid algal-dominated lake. However, to reverse the situation takes more than just reducing the nutrients as the ball can remain in the turbid state even at relatively low nutrient levels. To switch the state prematurely requires that a force be applied to the ball to push it over the edge of the algal cup and into the macrophyte state.

Therefore, each stable state has a number of buffer systems that help maintain that state, and through top-down and bottom-up processes, fish, plankton, invertebrates, macrophytes, and nutrients are all entwined in reciprocal feedback mechanisms that not only determine the state that is present (12,13) but can also be manipulated to force a shift.

Macrophyte buffer mechanisms that help maintain and stabilize a clear water state in shallow lake systems are numerous (4,14–16). Macrophytes act as habitats and refugia for macroinvertebrates and cladocerans that reduce the epiphyton and phytoplankton communities by grazing (6,17). Davis (18) suggested that the plant bed environment, often deoxygenated, favors grazers by discouraging their predators. Macrophytes may also release allelopathic chemicals (19); partake in the "luxury uptake" of nutrients, removing available nutrients for phytoplankton (20–22); and reduce available nitrate by anaerobic decomposition processes such as denitrification (23). Moreover, certain species of macrophytes have been shown to oxidize the sediment and reduce the release of phosphorus. With less available phosphorus in the water column, phytoplankton populations are reduced (24,25). In addition, stands of macrophytes reduce water movement within them. In consequence, suspended sediment resettles and turbidity falls, aiding the growth of macrophytes (26). Moss (10) suggests that even if light penetration is falling, submerged macrophytes may switch from low growing forms to taller species and that plants heavily encrusted with epiphyton may be able to shed their leaves and produce new growth. These processes competitively disadvantage phytoplankton and can buffer the macrophyte-dominated state.

Once established, on the other hand, phytoplankton can buffer the switch back to a macrophyte state by growing much earlier in the season than macrophytes in temperate regions. In doing so, they can curtail the development of turions, rhizomes, or seeds by shading in spring or reduce the formation of propagules by shading and competition for CO_2 in late summer (10). Submerged plants require carbon for growth, but algae have shorter CO_2 and HCO_3 diffusion pathways owing to their small size, thus removing the carbon available to the bulkier macrophytes (27). Some researchers have suggested that certain blue-green algae may even release chemicals toxic to macrophytes (28). In addition, phytoplankton-dominated open water has few refugia for grazing Cladocera; thus, any Cladocera venturing forth are typically removed by zooplanktivorous fish, which may be compounded by the fact that, with reduced macrophytes, few large macroinvertebrates would be available to large fish, favoring smaller sized fish feeding largely on zooplankton (6). Finally, in poorly flushed systems, phytoplankton species can consist of large filamentous or colonial blue-green algal species that are inedible to zooplankton, which along with the above mechanisms, all help to maintain phytoplankton-dominated water states over a wide range of nutrient concentrations.

By understanding the mechanisms that buffer and thus maintain the alternate stable states, it is possible to attempt to force a shift or switch from one state to another by manipulating top-down and bottom-up processes. Such attempts are often made during lake restoration projects when a desire exists to shift an algal-dominated turbid lake into a clear water macrophyte-dominated state. However, the long-term success of lake restoration projects, although they use knowledge of top-down and bottom-up processes, is not guaranteed and unexpected results are common (29–31). The uncertainty that surrounds these outcomes often stems from a lack of understanding about the buffer systems that stabilize the alternate states within the specific lake under restoration and an oversimplification of the top-down and bottom-up processes involved (13).

BIBLIOGRAPHY

1. Lewontin, R.C. (1969). The meaning of stability. *Brookhaven Symp. Biol.* **22**: 13–23.
2. Noy-Meir, I. (1975). Stability of grazing systems: An application of predator-prey graphs. *J. Ecol.* **63**: 459–481.
3. May, R.M. (1977). Thresholds and breakpoints in ecosystems with a multiplicity of states. *Nature* **269**: 471–477.
4. Blindow, I., Andersson, G., Hargeby, A., and Johansson, S. (1993). Long term pattern of alternative stable states in two shallow eutrophic lakes. *Freshwater Biol.* **30**: 159–167.
5. Scheffer, M. et al. (1993). Alternative equilibria in shallow lakes. *T.R.E.E.* **8**(8): 275–279.
6. Timms, R.M. and Moss, B. (1984). Prevention of growth of potentially dense phytoplankton populations by zooplankton grazing, in the presence of zooplanktivorous fish, in a shallow wetland ecosystem. *Limnol. Oceanogr.* **29**(3): 472–486.
7. Hosper, S.H. (1989). Biomanipulation, new perspective for restoring shallow, eutrophic lakes in The Netherlands. *Hydrolog. Bull.* **23**: 5–11.
8. Irvine, K., Moss, B., and Balls, H. (1989). The loss of submerged plants with eutrophication II. Relationships between fish and zooplankton in a set of experimental ponds, and conclusions. *Freshwater Biol.* **22**: 89–107.
9. Moss, B. and Leah, R.T. (1982). Changes in the ecosystem of a guanotrophic and brackish shallow lake in eastern England; potential problems in its restoration. *Internationale Revue der Gesamten Hydrobiologie* **67**: 625–639.
10. Moss, B. (1990). Engineering and biological approaches to the restoration from eutrophication of shallow lakes in

which aquatic plant communities are important components. *Hydrobiologia* **200/201**: 367–377.

11. Scheffer, M. (1990). Multiplicity of stable states in freshwater systems. *Hydrobiologia* **200/201**: 475–486.
12. Williams, A.E., Moss, B., and Eaton, J. (2002). Fish induced macrophyte loss in shallow lakes: Top-down and bottom-up processes in mesocosm experiments. *Freshwater Biol.* **47**: 2216–2232.
13. Williams, A.E. and Moss, B. (2003) Effects of different fish species and biomass on plankton interactions in a shallow lake. *Hydrobiologia* **491**: 331–346.
14. Canfield, Jr., D.E. et al. (1984). Prediction of chlorophyll *a* concentrations in Florida lakes: Importance of aquatic macrophytes. *Canad. J. Fisheries Aquat. Sci.* **41**: 497–501.
15. Coops, H. and Doef, R.W. (1996). Submerged vegetation development in two shallow, eutrophic lakes. *Hydrobiologia.* **340**: 115–120.
16. Romo, S., Van Donk, E., Gylstra, R., and Gulatis, R. (1996). A multivariate analysis of phytoplankton and food web changes in a shallow biomanipulated lake. *Freshwater Biol.* **36**: 683–696.
17. Carpenter, S.R. and Lodge, D.M. (1986). Effects of submerged macrophytes on ecosystem processes. *Aquatic Botany* **26**: 341–370.
18. Davis, J.C. (1975). Minimal dissolved oxygen requirements of aquatic life with emphasis on Canadian species: a review. *J. Fisheries Res. Board Canada* **32**: 2295–2332.
19. Wium-Andersen, S., Cristophersen, A.C., and Houen, G. (1982). Allelopathic effects on phytoplankton by substances isolated from aquatic macrophytes (Charales). *Oikos* **39**: 187–190.
20. Phillips, G.L., Eminson, D., and Moss, B. (1978). A mechanism to account for macrophyte decline in progressively eutrophicated freshwaters. *Aquatic Botany* **4**: 103–126.
21. Van Donk, E., Gulati, R.D., and Grimm, M.P. (1989). Food web manipulation in Lake Zwemlust: positive and negative effects during the first two years. *Hydrolog. Bull.* **23**: 19–34.
22. Ozimek, T., Van Donk, E., and Gulati, R.D. (1993). Growth and nutrient uptake by two species of *Elodea* in experimental conditions and their role in nutrient accumulation in a macrophyte dominated lake. *Hydrobiologia.* **251**: 13–18.
23. Takii, S. and Fukui, M. (1996). Comparison of anaerobic mineralization processes in sediments between littoral reed and offshore sites in a shallow hypertrophic lake. *Hydrobiologia.* **319**: 37–45.
24. Carpenter, S.R., Elser, J.J., and Olson, K.M. (1983). Effects of roots of *Myriophyllum verticillatum* L. on sediment redox conditions. *Aquatic Botany* **17**: 243–250.
25. Van Donk, E., Grimm, M.P., Gulati, R.D., and Klien Breteler, J.P.G. (1990). Whole lake food web manipulation as a means to study community interactions in a small ecosystem. *Hydrobiologia* **200/201**: 275–289.
26. Schiemer, F. and Prosser, M. (1976). Distribution and biomass of submerged macrophytes in Neusiedlersee. *Aquatic Botany* **2**: 289–307.
27. Simpson, P.S. and Eaton, J.W. (1986). Comparative studies of the photosynthesis of the submerged macrophyte *Elodea canadensis* and the filamentous algae *Cladophora glomerata* and *Spirogyra* sp. *Aquatic Botany* **24**: 51–12.
28. Van Vierssen, J. and Prins, Th.C. (1985). On the relationship between the growth of algae and aquatic macrophytes in brackish water. *Aquatic Botany* **21**: 165–179.
29. Marsden, M.W. (1989). Lake restoration by reducing external phosphorus loading: The influence of sediment phosphorus release. *Freshwater Biol.* **21**: 139–162.
30. De Melo, R., France, R., and McQueen, D.J. (1992). Biomanipulation: Hit or myth? *Limnol. Oceanogr.* **37**(1): 192–207.
31. Moss, B. et al. (1997). Vertically challenged limnology; contrasts between deep and shallow lakes. *Hydrobiologia* **342/343**: 257–267.

NOAA LAKE LEVEL FORECAST FOR LAKE MICHIGAN RIGHT ON TARGET

NOAA Great Lakes
Environmental Research Lab

July 17, 2003—How high or low the water level is in the five Great Lakes has long been of interest to those who use the lakes for commercial and recreational boating and fishing, energy, and as a source of drinking water. The levels have been measured and forecast monthly for many years, and NOAA scientists hit the bull's-eye with their July forecast of the water level of Lake Michigan.

"We were within inches of the actual measurement," said Cynthia Sellinger, deputy director of the NOAA Great Lakes Environmental Research Laboratory in Ann Arbor, Mich.

GLERL uses meteorological forecasts from the NOAA National Weather Service to forecast the lake levels six months out. Back in January, GLERL's Lake Michigan forecast was 176.61 meters—about 579 feet; in July the actual level was 176.65 meters—approximately 2 inches higher than the forecast. Sellinger meets monthly with her

---

colleagues from Environment Canada and the U.S. Army Corps of Engineers to forecast the levels.

Since 1997, Lake Michigan has dropped 4.1 feet, the largest drop in that short of time span since records have been kept. Lakes Michigan and Huron are often considered one body of water as they are connected by the Straits of Mackinac, located between the upper and lower peninsulas of the state of Michigan.

The other lakes are more difficult to forecast, because the water flow of Superior and Ontario are regulated by power companies and Erie gets 80 percent of its water from Michigan and Huron, Sellinger explained.

The Great Lakes—Superior, Michigan, Huron, Ontario, and Erie—are the major bodies of water within the Great Lakes Basin. The basin is located along the international boundary between Canada and includes portions of eight states and the provinces of Ontario and Quebec. Combined, these lakes make up the largest surface fresh water body in the world—they contain 90 percent of the U.S.'s fresh water and 20 percent of the world's fresh water.

The Great Lakes carry thousands of recreational and commercial ships annually. Lake levels are important to shippers, who must load the lake carriers with fewer goods when the levels are low and to recreational boaters, who often have to find dock space in deeper water. The Lakes provide drinking water to the 40 million residents as well as water to power and cool electric generating plants.

Lake levels have been dropping for the past six years and lake evaporation has increased significantly.

"The only good thing that comes out of lower lake levels is more beach for beach-goers," according to Sellinger. However, that also exposes and destroys some vital breeding habitat for lake wildlife.

Lake levels have been forecast before GLERL was formed in 1974. Part of the laboratory's mandate is to provide the lake level forecasts, a task that was given to GLERL when the research arm of the U.S. Army Corps of Engineers was transferred to the laboratory.

NOAA is dedicated to enhancing economic security and national safety through the prediction and research of weather and climate-related events and providing environmental stewardship of the nation's coastal and marine resources. NOAA is part of the U.S. Department of Commerce.

RELEVANT WEB SITES

NOAA Great Lakes Environmental Research Lab
NOAA Great Lakes Water Levels
NOAA Great Lakes Satellite Information

MEDIA CONTACT

Jana Goldman, NOAA Research, (301) 713–2483.

(Photo courtesy Greg Lang of the NOAA Great Lakes Environmental Research Laboratory.)

SUBMERGED AQUATIC PLANTS AFFECT WATER QUALITY IN LAKES

Karl E. Havens
South Florida Water Mgmt. District
West Palm Beach, Florida

BACKGROUND INFORMATION

A principal cause of water quality impairment in lakes and reservoirs is enrichment by excessive levels of nutrients, resulting in undesirable conditions that include noxious and/or toxic algae blooms (1), and taste and odor problems (2). The nutrient content of lake water is determined by a myriad of factors, including external loading rate, water residence time and depth, geochemical interactions in the water body, and biological processes. Biological processes are particularly important in shallow lakes. Fish and benthic animals recycle nutrients from bottom sediments, buoyant algae carry nutrients upward in the water column, settling rates of algae determine loss rates for water column nutrients, and plants facilitate various nutrient-scavenging processes (3,4). Submerged plants, in particular, can have major influences on water quality in terms of clarity, biomass of noxious algae, taste and odor, organics, and suspended solids—factors of major relevance to society's use of the water.

This chapter addresses four questions:

1. What types of submerged plants occur in lakes, and how does type of plant relate to water quality benefits?
2. What factors control the biomass and spatial extent of submerged plants?
3. What are the processes whereby plants affect lake water quality, in particular, the concentrations of nutrients, turbidity, and the occurrence of noxious algae?
4. How can lakes be managed to promote the development of submerged plants that will help to improve water quality?

For a detailed overview of the ecological role of submerged plants in shallow lakes, the reader may consult Jeppesen et al. (5) or Scheffer (3). For identification of different types of plants in lakes and for general information about their biology and distribution, there are a number of web-based resources provided by major universities; for example: http://aquat1.ifas.ufl.edu (University of Florida, Center for Aquatic and Invasive Plants).

TYPES OF SUBMERGED AQUATIC VEGETATION RELATIVE TO WATER QUALITY BENEFITS

Lakes can support rooted vascular plants (true plants that have tissues for gas and materials transport); plant-like macroalgae, such as *Chara* (shrimp grass) that do not have vascular tissues or roots; and filamentous algae that can form dense mats on the lake bottom, sometimes covering rocks and even other plants. Filamentous algae are least desirable from a water quality standpoint, because at high biomass, they can kill other plants by shading and they form floating noxious mats at the lakeshore when they decompose (3). *Chara* generally is a pioneer species that rapidly invades lake habitats after a disturbance (6,7), and because it does not have roots, all of its accumulated nutrient content may be recycled back into the water column if unfavorable conditions (e.g., reduced light, high wind, or cold water) cause the plants to be lost. Certain vascular plants, such as *Ceratophyllum* (coontail), also do not have roots and would have this same problem. In regard to *Chara*, note that clear water and low nutrient levels are most often associated with dense beds of this plant in temperate (8,9) and subtropical (7) lakes and that processes other than direct incorporation of nutrients into biomass may be most important for water quality improvement by *Chara* beds. Rooted vascular plants—common examples in North America are *Vallisneria* (eelgrass) and *Potamogeton* (pondweed)—are more permanent than *Chara*, and because they can transfer nutrients to their underground roots prior to seasonal “die-off,” they have potential to provide a better long-term nutrient sink.

FACTORS CONTROLLING THE BIOMASS AND EXTENT OF SUBMERGED VEGETATION

Scientists and lake managers have given considerable attention to the factors that control the biomass and spatial extent of submerged plants in lakes. Obviously, water depth plays a key role because ultimately it determines the amount of light that reaches the lake bottom where plants germinate from seeds (or oospores, in the case of algae) and begin to grow. However, light availability at any given depth is heavily influenced by the turbidity of the water, which is determined by the amount of suspended solids, algal cells, and dissolved organic matter. Hence, a number of authors have identified significant relationships between the maximal depth for submerged plant occurrence in lakes and simple indicators of turbidity, such as Secchi disk transparency (10–13). The relationships vary from lake to lake. For example, in Lake Okeechobee, Florida, the published models predict that submerged plants can occur at depths between 0.7 and 1.8 m, whereas plants actually occur at depths as great as 2.0 m. This may reflect the occurrence of plants that are adapted to low light levels in this turbid shallow lake (14). In general, *Chara* can occur at greater depths than vascular plants, for any given level of turbidity (15).

Experimental studies have confirmed the causal effect of underwater light availability on submerged plant growth (14,16). However, when evaluating field data, it can be difficult to tell whether plants are responding to variations in underwater irradiance or whether they are the cause of high irradiance in certain areas of a lake because they have various functions that reduce turbidity. It is a classic “chicken versus egg” paradox. Some shallow lakes display zones of clear water immediately adjacent to turbid water of the same depth (8), the only difference is that the clear areas have dense beds of submerged aquatic vegetation (SAV). In such instances, one might conclude that the plants have caused the clear water. In Lake Okeechobee, Florida, clear water expanded over an ever-increasing area during summer 2000, coincident with the expansion of large beds of *Chara*. Areas adjacent to the plant beds had more turbid water, but they also had slightly greater depths (6). It was not possible to establish a cause–effect relationship. Scheffer (3) provides a good overview of the issue of correlation versus causation as it pertains to SAV and turbidity in lakes, for those interested in more information on this subject.

In addition to depth and turbidity, sediment type may influence the distribution and biomass of submerged plants. Soft sediments that are readily suspended by wind, waves, and fish may not support plants, especially types like *Chara* that do not have roots. In Lake Okeechobee, which has areas of sand, peat, and mud sediments, *Chara* does not occur at locations that have soft organic mud, even when depth and light conditions appear to be favorable (7). A simple empirical model, based on water depth and sediment type, was able to predict at 75% accuracy the spatial distribution of *Chara* over a 170 km^2 landscape in that lake. The sediment type found also was the major factor controlling where plants could occur in Lake Kinneret, Israel (17), whereas water depth was the factor controlling the timing of plant occurrence.

Other factors that can influence the distribution and biomass of submerged plants include the degree of exposure to wind and waves (13,18), the slope of the lake bottom (19), and as mentioned before, overgrowth by filamentous algae (20).

PROCESSES WHEREBY SUBMERGED VEGETATION AFFECTS WATER QUALITY

When one compares lakes with and without dense submerged plants or follows the conditions in a lake over time as plants increase, striking changes in water quality are apparent. Water can change from murky green (algae blooms) to crystal clear (Fig. 1), and at the same time, nutrient levels in the water can dramatically decline. Even a relatively low biomass of submerged plants can attenuate peak concentrations of suspended algae (phytoplankton) in a lake’s water column (21; Fig. 2). The benefits of clear water and plants for the fish, wildlife, and humans who use the resource are obvious.

Submerged plants mediate changes in water quality by a variety of mechanisms. The focus here is on processes that control nutrient and light availability, rather than food web interactions. However, it should be recognized that a potentially important function of submerged plants is to provide refuge for large zooplankton (22). During the daytime, large zooplankton can escape fish predation by

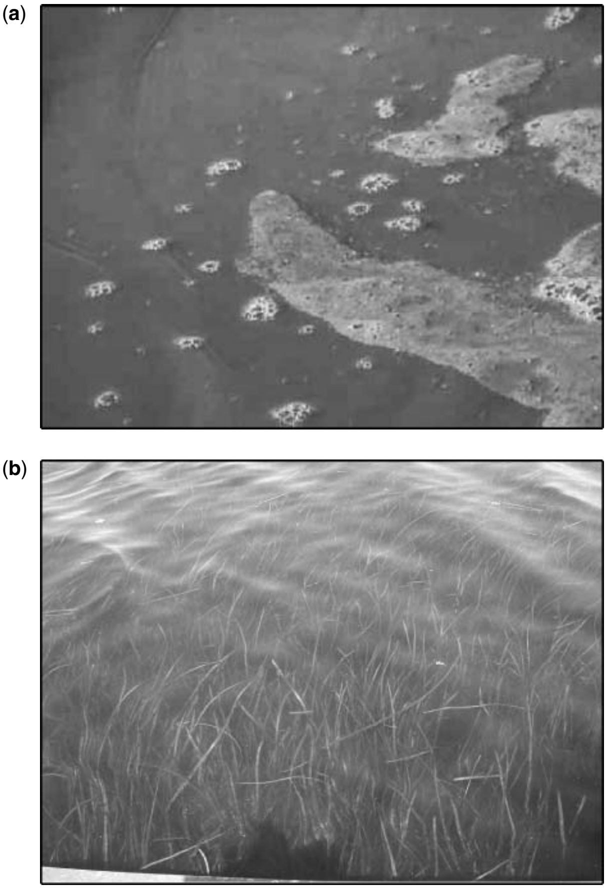

Figure 1. Photographs of typical conditions in turbid, phytoplankton-dominated lakes that have high levels of nutrients in the water column and blooms of algae (**a**), versus clear, plant-dominated lakes that have low nutrient levels in the water and no algal blooms. These photographs were taken by the author at Lake Okeechobee, Florida.

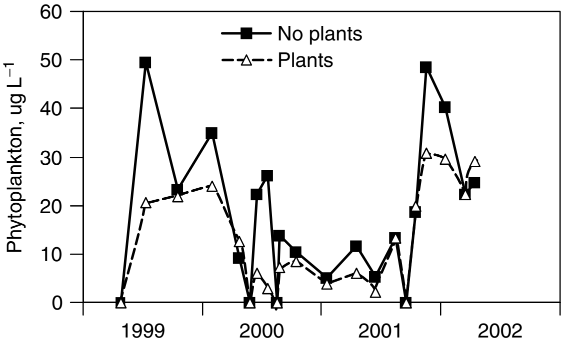

Figure 2. Data from the shoreline region of Lake Okeechobee, Florida, showing the concentrations of phytoplankton chlorophyll (μgL^{-1}) at sampling stations with plants versus stations without plants. The presence of plants dampens the peaks in phytoplankton, as a result of various processes described in the text. Modified from a figure in Havens (20).

migrating into the plant zone, and at night, when they are not susceptible to those visual predators, they can migrate out into the open water to graze on phytoplankton. The result is theoretically a reduction in algal biomass and reduced risk of nuisance algal blooms, although uncertainty remains about the magnitude of this effect. One of the complicating issues is that plants may also provide refuge for the small fish that eat zooplankton (22).

In a conceptual framework (Fig. 3), one mechanism whereby submerged plants can benefit water quality is the uptake (**U**) of dissolved nutrients that otherwise could fuel algal blooms and the array of adverse impacts shown in the diagram. Benthic algal mats can also take up nutrients from the water, further reducing availability for the bloom-forming phytoplankton. Much of the nutrient uptake by plants actually is associated with attached algae (periphyton) that grow on plant surfaces (23,24). Vascular plants and benthic algae also take up nutrients from lake sediments; it has been documented that benthic algal mats can substantially reduce the diffusive flux (**D**) of nutrients from sediments to the overlying water in both shallow freshwater and estuarine systems (25–27). All of the plants and algae mentioned here can at times excrete (**E**) dissolved nutrients back into the sediment or water column. Conditions that favor net growth of plants and/or benthic algal mats favor a net uptake of nutrients, which translates into better water quality.

Nutrients present in the water column or sediments in a dissolved form can also be transformed (**T**) into particulates that ultimately settle (**S**) to the lake bottom at a rate that depends on water column mixing (or lack thereof), the density of the water, and the characteristics of the particles. Plants can stimulate both the transformation of dissolved nutrients to particulates in the water and the rate of settling. In regard to transformation (**T**), when plants carry out intense photosynthesis, the pH of the surrounding water can be substantially elevated. At high pH (e.g., >10), much of the dissolved inorganic carbon (C) occurs as carbonate ions $(CO_3)^{-2}$ (28), which can be precipitated to the sediments, or onto the surface of plants, as calcium carbonate ($CaCO_3$). More important, from a water quality standpoint, is that this process can be accompanied by coprecipitation of phosphorus (29), as insoluble hydroxyapatite [$Ca_{10}(PO_4)_6(OH)_2$]. Phosphorus is a nutrient that often is limiting to phytoplankton in freshwater lakes (30). Therefore, coprecipitation could help reduce phytoplankton growth and suppress formation of blooms. However, it must be noted that at high pH, the release of dissolved phosphorus from lake sediments (**D**) is sometimes enhanced, even under oxidized conditions (31). Depending on circumstances, this could counteract the effects on particle transformation.

Plants play an important role in settling of particles (**S**) by slowing the rate of water movement (32), and acting as sediment "traps" by intercepting particles using their leaves and stems; the particles subsequently sink to the lake bottom (33). The uptake of nutrients by periphyton, which ultimately provide particulate material (dead cells) to the sediments is another loss process mediated by plants. Once the particles settle within a plant bed, they are less likely to be re-suspended (**R**) because dense beds of submerged plants significantly reduce the amount of wind and wave energy that reaches the sediment surface (34).

Finally, it has been suggested that submerged plants might inhibit the growth of suspended algae by releasing

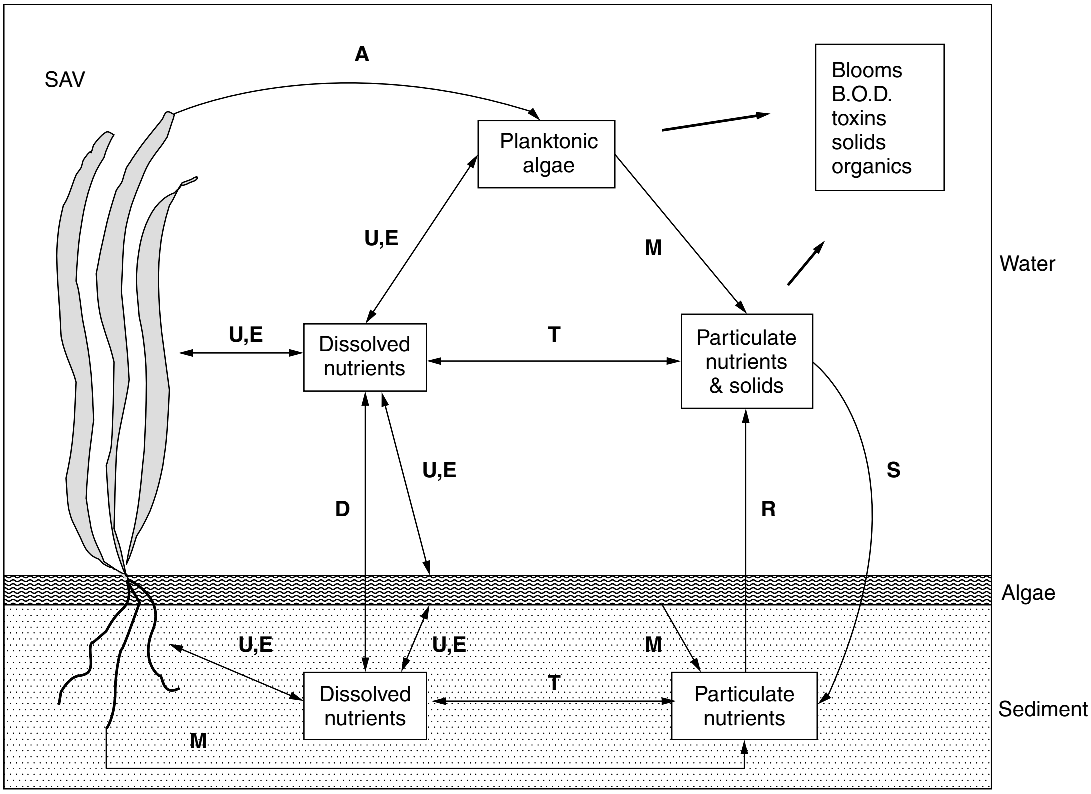

Figure 3. A simple conceptual model, illustrating how submerged aquatic plants and benthic algal mats influence the dynamics of phosphorus in a typical eutrophic lake. The arrows correspond to various processes: **A** = release of allelopathic chemicals; **U** = uptake of dissolved nutrients; **E** = excretion of dissolved nutrients; **T** = transformation of nutrients from dissolved to particulate form, or vice versa; **D** = diffusion of dissolved nutrients from sediments to water, or vice versa; **M** = mortality; **S** = settling of particles; and **R** = resuspension of particles by wind, waves, or bioturbation. B.O.D. = Biological Oxygen Demand. The two large arrows point to impacts of algal blooms and solids on the water resource. This figure is modified from a similar model presented in Havens et al. (23).

allelopathic chemicals (**A**) into the water column (35). This effect has been documented in controlled tank studies with a high biomass of plants and restricted water volume, but uncertainty remains regarding its importance in lake ecosystems.

MANAGING LAKES TO PROMOTE SUBMERGED VEGETATION

The total biomass of primary producers in a lake is controlled by nutrient input to the ecosystem (36), whereas partitioning of that biomass into phytoplankton and plants is determined by a complex set of factors. Understanding how particular environmental conditions influence submerged plants is critical to facilitating plant survival and growth, and in turn, maintaining good water quality (see below). Lakes dominated by soft organic sediments or heavily influenced by wind-driven waves are not well suited for submerged plants. However, even under these conditions, options may exist, such as dredging to remove soft mud or construction of wave barriers to provide protected areas for plants to grow (37). In lakes that have favorable sediments and low to moderate wave energy along the shoreline, the key factor controlling plants is likely to be light availability, which as noted before, varies with the depth and turbidity of the water. One of the factors controlling turbidity is the nutrient content of the water, which determines the abundance of phytoplankton cells that absorb light. Scheffer (38) noted that there is a complex relationship between nutrients and turbidity in lakes that varies with the presence or absence of plants. This relationship (Fig. 4) is highly relevant when considering how to establish plants in a nutrient-rich, turbid lake.

As a lake becomes enriched with nutrients, the turbidity of the water increases. In lakes with plants, the turbidity at any given nutrient level is lower than that in lakes without plants due to the mechanisms described before (Fig.3). Once a “critical turbidity” is reached, where light does not allow net growth to occur, plants are lost, and turbidity quickly increases (upward arrow), as the nutrients formerly sequestered by plants go into the biomass of phytoplankton in the water. To restore plants to a lake that has reached this state, nutrients must be reduced to a lower level than where the switch previously occurred (downward arrow). Both the clear-water, plant-dominated state and the turbid-water, phytoplankton-dominated state are highly stable—it takes a large

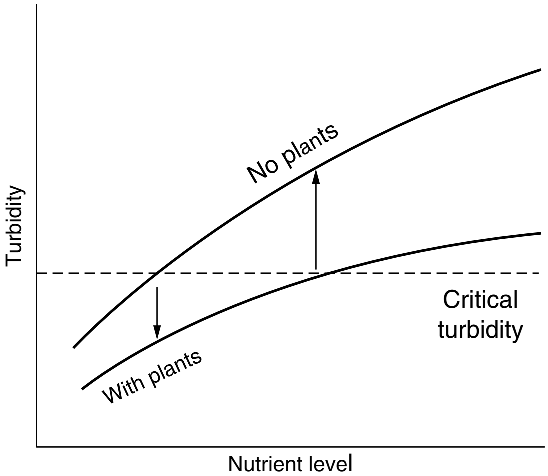

Figure 4. Relationship between turbidity and nutrient level in lake water, in lakes with versus without plants. Lakes with plants have a lower turbidity for any given nutrient level than lakes without plants. When nutrient levels increase in lakes with plants, the plants can be lost once a critical turbidity is reached, at which point the net growth of plants is zero. When this occurs, turbidity can rapidly increase (upward arrow). To restore the plants in such a lake (downward arrow), the nutrient level must be substantially reduced. This figure is modified from a diagram in Scheffer (3).

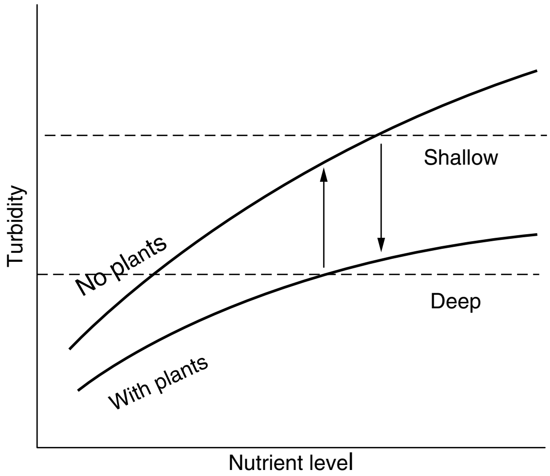

Figure 5. Effects of water depth on the occurrence of plants in turbid lakes. By substantially lowering the water level of a lake, the critical turbidity is increased (upper dashed line), and this can allow plants to recover (downward arrow), even though the nutrient level remains high. If the water level increases substantially (lower dashed line), the plants may be lost, even at a lower level of nutrients than that where they were recovered (upward arrow). Unless there is capability for strong control of water levels and no other adverse impacts of keeping the lake low (e.g., impacts on water supply), the ultimate solution to keeping plants and clear water in a lake is to reduce external nutrient inputs. This figure is based on information in Scheffer (3).

input of nutrients to switch a plant-dominated lake to a phytoplankton lake and an even larger reduction of nutrients to switch the lake back (3,38). This stability of states reflects the biological feedback loops that occur in the system. Plants cause clear water, which favors more plants, which cause even clearer water…etc. Phytoplankton cause turbid water, which reduces plants, which favors more phytoplankton, which further increases turbidity and so on.

Note that at extremes along the nutrient axis (extremely low or high nutrient levels), plants can exist in only one state—clear or turbid, respectively. However, for lakes that have moderate to high levels of nutrients and the possibility of alternative stable states (38), one method for switching a lake from phytoplankton to plant dominance is to reduce inputs of limiting nutrients, such as phosphorus and nitrogen (3,39). Ultimately, this is the path to lake restoration. However, an interim management strategy that can sometimes be used in tandem with nutrient reduction is to lower the lake's water level. Critical turbidity is a function of depth, so that in a deep lake, relatively low turbidity might be sufficient to prevent plant growth. However, in shallow water, enough light may reach the lake bottom to allow plants to germinate and grow. By lowering the water level, it is possible to raise the critical turbidity (Fig. 5), so that plants can become reestablished at a higher level of nutrients than that when the lake had deeper water. At this point, the lake has essentially been reset onto the lower curve.

A switch from turbid to clear water happened in the shoreline region of Lake Okeechobee, Florida, in summer 2000, after an intentional lowering of the water level was followed by a major drought (40). The water level in the lake dropped by more than 2 m, and the spatial extent of submerged plants increased from less than 10 to over 170 km^2 (16). Water in shoreline areas, which was highly turbid from algae blooms in the late 1990s, became crystal clear in 2000–2002. Recently, it has somewhat deteriorated, as higher water levels led to reduced biomass of plants. Ultimately, reduced nutrient input is the solution to the water quality problem in this and other nutrient-enriched lakes (39), because as long as alternative states exist, uncontrollable events, such as marked increases in water level or wind storms, could potentially shift a plant-dominated lake back to a turbid lake with phytoplankton blooms.

SUMMARY

Water quality in shallow lakes and reservoirs can be strongly influenced by biological processes, including those associated with fish, invertebrates, algae, and plants. The most striking influences often are those associated with submerged aquatic plants, which include rooted vascular plants, loosely attached macroalgae, such as *Chara*, and filamentous benthic algal mats. Where conditions favor high biomass and spatial extent of these plants, water turbidity generally is low, biomass of phytoplankton is low, and quality of water for drinking and other uses is good. The plants maintain these conditions by removing nutrients from the water; supporting attached algae that also sequester nutrients; stabilizing bottom sediments; and a variety of other physical, chemical, and biological processes. Actions to favor development of submerged plants in shallow, nutrient-enriched lakes can help to maintain good water quality, although the long-term solution still is to reduce external nutrient inputs.

BIBLIOGRAPHY

1. Paerl, H.W. (1988). Nuisance phytoplankton blooms in coastal, estuarine, and inland waters. *Limnol. Oceanogr.* **33**: 823–847.
2. Smith, V.H. et al. (2002). Managing taste and odor problems in a eutrophic drinking water reservoir. *Lake Reservoir Manage* **18**: 319–323.
3. Scheffer, M. (1998). *The Ecology of Shallow Lakes*. Chapman and Hall, London, UK, p. 357.
4. Havens, K.E. and Schelske, C.L. (2001). The importance of considering biological processes when setting total maximum daily loads (TMDL) for phosphorus in shallow lakes and reservoirs. *Environ. Pollut.* **113**: 1–9.
5. Jeppesen, E., Sondergaard, M., Sondergaard, M., and Christoffersen, K. (Eds.). (1998). *The Structuring Role of Submerged Macrophytes in Lakes*. Springer, Berlin, p. 423.
6. Havens, K.E. et al. (2001). Hurricane effects on a shallow lake ecosystem and its response to a controlled manipulation of water level. *Sci. World J.* **1**: 44–70.
7. Havens, K.E. et al. (2002). Large-scale mapping and predictive modeling of submerged aquatic vegetation in a shallow eutrophic lake. *Sci. World J.* **2**: 949–965.
8. Scheffer, M. et al. (1994). Vegetated areas with clear water in turbid shallow lakes. *Aquatic Bot.* **49**: 193–196.
9. Van den Berg, M.S., Coops, H., Meijer, M.L., Scheffer, M., and Simons, J. (1998). Clear water associated with a dense *Chara* vegetation in the shallow and turbid Lake Veluwemeer, The Netherlands. In: *The Structuring Role of Submerged Macrophytes in Lakes*. E. Jeppesen, M. Sondergaard, M. Sondergaard, and K. Christoffersen (Eds.). Springer, Berlin, pp. 339–352.
10. Canfield, D.E., Langeland, K.A., Linda, S.B., and Haller, T.T. (1985). Relations between water transparency and maximum depth of macrophyte colonization in lakes. *J. Aquatic Plant Manage.* **23**: 25–28.
11. Chambers, P.A. and Kalff, J. (1985). Depth distribution and biomass of submerged aquatic macrophyte communities in relation to Secchi depth. *Can. J. Fish Aquatic Sci.* **42**: 701–709.
12. Middleboe, A.L. and Markager, S. (1997). Depth limits and minimum light requirements of freshwater macrophytes. *Freshwater Biol.* **37**: 553–568.
13. Hudon, C., Lalonde, S., and Gagnon, P. (2000). Ranking the effects of site exposure, plant growth form, water depth, and transparency on aquatic plant biomass. *Can. J. Fish Aquatic Sci.* **57**(Suppl. 1): 31–42.
14. Grimshaw, H.J. et al. (2002). The effects of shading on morphometric and meristic characteristics of wild celery, *Vallisneria americana*, transplants from Lake Okeechobee, Florida. *Archiv fur Hydrobiologia* **155**: 65–81.
15. Blindow, I. (1992). Long and short term dynamics of submerged macrophytes in two shallow lakes. *Freshwater Biol.* **28**: 15–27.
16. Van Dijk, G.M. and Van Vierssen, W. (1991). Survival of a *Potamogeton pectinatus* L. population under various light conditions in a Shallow Eutrophic Lake Lake Veluwe in The Netherlands. *Aquatic Bot.* **39**: 121–130.
17. Gafny, S. and Gasith, A. (1999). Spatially and temporally sporadic appearance of macrophytes in the littoral zone of Lake Kinneret, Israel: Taking advantage of a window of opportunity. *Aquatic Bot.* **62**: 249–267.
18. Chambers, P.A. (1987). Nearshore occurrence of submersed aquatic macrophytes in relation to wave action. *Can. J. Fish Aquatic Sci.* **44**: 1666–1669.
19. Duarte, C.M. and Kalff, J. (1986). Littoral slope as a predictor of maximum biomass of submerged macrophyte communities. *Limnol Oceanogr.* **31**: 1072–1080.
20. Havens, K.E. (2003). Submerged aquatic vegetation correlations with depth and light attenuating materials in a shallow subtropical lake. *Hydrobiologia* **493**: 173–186.
21. Weisner, S.E.B., Strand, J.A., and Sandsten, H. (1997). Mechanisms regulating abundance of submerged vegetation in shallow eutrophic lakes. *Oecologia* **109**: 592–599.
22. Jeppesen, E., Lauridsen, T.L. Kairesalo, T., and Perrow, M.R. (1998). Impact of submerged macrophytes on fish-zooplankton interactions in lakes. In: *The Structuring Role of Submerged Macrophytes in Lakes*. E. Jeppesen, M. Sondergaard, M. Sondergaard, and K. Christoffersen (Eds.). Springer, Berlin, pp. 353–368.
23. Havens, K.E. et al. (2001). Complex interactions between autotrophs in shallow marine and freshwater ecosystems: Implications for community responses to nutrient stress. *Environ. Pollut.* **113**: 95–107.
24. Vadeboncoeur, Y. and Steinman, A.D. (2002). Periphyton function in lake ecosystems. *Sci. World J.* **2**: 1449–1468.
25. Hansson, L.A. (1988). Effects of competitive interactions on the biomass development of planktonic and periphytic algae in lakes. *Limnol Oceanogr.* **33**: 121–128.
26. Hansson, L.A. (1989). The influence of a periphytic biolayer on phosphorus exchange between substrate and water. *Archiv fur Hydrobiologie* **115**: 21–26.
27. McGlathery, K.J., Krause-Jensen, D., Rysgaard, S., and Christensen, P.B. (1997). Patterns of ammonium uptake within dense mats of the filamentous macro-alga *Chaetomorphalinum*. *Aquatic Bot.* **59**: 99–115.
28. Golterman, H.L. (1969). *Methods for Physical and Chemical Analysis of Fresh Waters*. IBP Handbook No. 8. Blackwell Scientific, Oxford, UK, p. 213.
29. Murphy, T., Hall, K., and Yesaki, I. (1983). Co-precipitation of phosphate and calcite in a naturally eutrophic lake. *Limnology Oceanogr.* **28**: 58–67.
30. Schindler, D.W. (1977). Evolution of phosphorus limitation in lakes. *Science* **195**: 260–262.
31. James, W.F., Barko, J.W., and Field, S.J. (1996). Phosphorus mobilization from littoral sediments of an inlet region in Lake Delavan, Wisconsin. *Archiv fur Hydrobiologie* **128**: 245–257.
32. Fonseca, M.S., Fisher, J.S., Zieman, J.C., and Thayer, G.W. (1982). Influence of the sea grass, *Zostera marina* L., on current flow. *Estuarine Coastal Shelf Sci.* **15**: 351–364.
33. Carpenter, S.R. (1981). Submersed vegetation: an internal factor in lake ecosystem succession. *Am. Naturalist* **118**: 372–383.
34. Vermaat, J.E., Santamaria, L., and Roos, P.J. (2000). Water flow across and sediment trapping in submerged macrophyte beds of contrasting growth form. *Archiv fur Hydrobiologia* **148**: 549–562.
35. van Donk, E. and van de Bund, W.J. (1998). Impact of submerged macrophytes including charophytes on phyto- and zooplankton communities: Allelopathic versus other mechanisms. *Aquatic Bot.* **72**: 261–274.
36. Liboriussen, L. and Jeppesen, E. (2003). Temporal dynamics in epipelic, pelagic, and epiphytic algal production in a clear and a turbid shallow lake. *Freshwater Biol.* **48**: 418–431.

37. Canfield, D.E. Bachmann, R.W., and Hoyer, M.V. (2000). A management alternative for Lake Apopka. *Lake Reservoir Manage* **16**: 205–221.
38. Scheffer, M. (1989). Alternative stable states in eutrophic shallow freshwater systems: A minimal model. *Hydrobiological Bull.* **23**: 73–85.
39. Moss, B., Madgwick, J., and Phillips, G. (1997). *A Guide to the Restoration of Nutrient-Enriched Shallow Lakes*. W.W. Hawes, UK, p. 180.
40. Steinman, A.D., Havens, K.E., Carrick, H.J., and Van Zee, R. (2002). The past, present, and future hydrology and ecology of Lake Okeechobee and its watersheds. In: *The Everglades, Florida Bay, and Coral Reefs of the Florida Keys: An Ecosystem Sourcebook*. K.G. Porter and J. Porter (Eds.). CRC Press, FL, pp. 19–37.

LAKES—DISCHARGES TO

Laurel Phoenix
Green Bay, Wisconsin

Lakes have suffered from a long history of receiving pollutants from various sources. These pollutants degrade water quality, with implications for aquatic and terrestrial ecosystems as well as for a variety of issues for human populations dependent on that water and those ecosystems. The degree to which these pollutants degrade the lake depends on the amount and types of pollutants as well as the physical characteristics of the lake. For an introductory overview of how discharges to lakes are harmful, the pollutant sources and how pollutants are transported into the lake are described. The physical characteristics that determine the assimilative capacity of a lake are then covered. Finally, the effects that different pollutants have on lakes and their dependent ecosystems are explained as well as the ramifications for aquatic species, birds, wildlife, and humans.

POLLUTANT SOURCES AND TRANSPORT

Polluted discharges into lakes can come from watercraft, terrestrial sources, and atmospheric deposition. Watercraft are categorized into ships, boats, and recreational equipment. Terrestrial pollutant sources can be categorized into point and nonpoint sources. Point sources of pollution emanate from a single location, like a factory's pipe dumping effluent into a stream or lake. Nonpoint sources of pollution, often called "polluted runoff," can come from all over a landscape and from many locations at once, such as fertilizer runoff from suburban lawns or oil washing off city streets. Atmospheric deposition has characteristics of both point and nonpoint source pollution and so is treated separately here. Atmospheric deposition of pollutants originating from smokestacks travels through the air and then falls into the water.

Watercraft

Ships discharge a variety of pollutants into lakes. The ballast water that ships discharge into the water contains oils and invasive species. The entry of the zebra mussel into the Great Lakes is a good example of how an aquatic ecosystem can be significantly changed by the introduction of just one new species. Now, regulations require ships to dump their ballast water in the open ocean before entering the Great Lakes. Used oil from ship engines has also been dumped into lakes for years. Sewage from ships has also been dumped untreated into lakes, although cruise ships have recently been pressured to install onboard sewage treatment systems so that only treated effluent is dumped. Atmospheric deposition from smokestacks is becoming a more serious problem with increased global shipping. Ships burn low-grade diesel fuel and these carcinogenic compounds then enter lakes.

Boats used for recreation also pollute lakes. Boat owners may dump their sewage straight into the lake rather than pumping it out dockside at sewage receptacle facilities. Gas spills often happen at docks when gas pumping is not done with care, and some gas and oil are emitted into the water from outboard motors.

Smaller recreational equipment like jet skis pollute water with oil and gasoline at a far higher rate than outboard motors. Jet skis also damage littoral habitat, kill aquatic animals, or frighten aquatic species and disturb mating or nesting behaviors. Chromosomal damage in fish has been linked to the fossil fuel discharges from jet skis.

Terrestrial Point Sources

The most common point sources polluting lakes are sewage systems and industrial effluent pipes. Sewage can enter lakes either through intentional discharge of raw or treated sewage, or through occasional discharges from combined sewer overflows (CSOs). Depending on the city and on the country, one can find cities discharging raw sewage into lakes rather than spending the money to treat it properly. Sewage treatment plants in other cities may have primary, secondary, or tertiary sewage treatment, resulting in various types and concentrations of pollutant discharge into lakes or their tributaries. Combined sewer overflows are found in cities where stormwater and sewage flow down the same pipes, rather than flowing through separate stormwater and sanitary sewers. In rain events, the additional stormwater entering sewage pipes may be more than a sewage treatment plant's designed capacity. When this happens, sewer pipe outlets are designed to open and discharge raw sewage into a local stream or lake to prevent too much sewage from reaching the treatment plant; passing through the plant without sufficient treatment; and/or backing up into the city. Consequently, the raw or partially treated sewage from some cities contributes a daily material loading to a lake, whereas combined sewer overflows discharge sporadically as a result of heavy rain or rapid snowmelt. Some of the constituents of sewage are human waste, bacteria, viruses, chemicals, detergents, heavy metals, and pharmaceutical and personal care products (PPCPs) (1) such as antibiotics, steroids, birth control pills, shampoo, hair dye, and perfume.

Industrial pollutants enter a tributary or lake by a pipe, either intentionally through a permitting program or by accidental release. In the United States, any industrial facility, including sewage treatment plants, must have a National Pollutant Discharge Elimination System (NPDES) permit to discharge anything into a water body. The type, quantity, and concentration of pollutants from various industrial processes are written into the permit, thus limiting the effluent discharge to the permit limits. Pollutants may be chemical, biological (e.g., bacteria, viruses, or organic matter that deoxygenates the water), thermal, radioactive, metals, and so on. Because there is a permitting system, and since the name implies a reduction or elimination of pollutant discharges, it is widely assumed that permits issued are still within the assimilative capacity of the water body, and that the appropriate government agencies monitor and enforce permits to ensure that the water is not becoming degraded. Unfortunately, this is not the case, and rivers and lakes are receiving rising material loads.

Terrestrial Nonpoint Sources

Nonpoint source pollutants move into a lake from its entire drainage area. Nonpoint sources are grouped into urban runoff, mining, agriculture, silviculture, groundwater, and human transport. Urban runoff includes fertilizers, pesticides, and pet waste from lawns; oil, rubber, deicing chemicals, and metals from roads; and sediment from construction.

Mining contributes acid mine drainage to tributaries or lakes or can put cyanide into the water when it is used to leach gold out of gold ore. Strip mining transports sediment into streams. Agriculture contributes a significant portion of the overall nonpoint source pollutant load to lakes. These pollutants include sediment from erosion, pesticides mobilized from the soil by rain, and overapplication of commercial fertilizer or manure.

Silviculture primarily contributes pesticides and sediment to lakes. Pesticides used on tree farms or areas sprayed for infestations (e.g., gypsy moth) are transported through rain and snowmelt into water bodies. Sediment from road building for timber trucks or from clear-cutting or harvesting on steep slopes is also discharged into streams and lakes.

There are two primary types of pollutants moving through groundwater. Leaking septics contribute organic sewage material, *Escherichia coli*, viruses, and toxic household cleaners. Acid rain filters into soils, and the resulting chemical reaction then mobilizes mercury, manganese, aluminum, and zinc into the groundwater and then into lakes through subsurface flow. Humans can transport invasive species unintentionally or with purpose. For example, zebra mussels can “hitchhike” on a pleasure boat being hauled from one water body to another, or canals have been built that connect previously distinct ecosystems and allow migration of species into a new lake. Intentional transfer of invasive species happens when people get rid of pet fish by flushing them down the toilet or dumping them directly into streams or lakes, and this fish will now be an exotic species for that lake. Other exotic species can be placed in a lake with the intent of using them to control other aquatic animals or plants, but several of these experiments (e.g., carp) have backfired; new unexpected problems arise from the exotic introduction.

Atmospheric Deposition

Although some scientists may regard atmospheric deposition as a type of nonpoint source pollution, it is categorized here as distinct from terrestrial point or nonpoint source pollution because it comes from discrete sources (smokestacks), but from numerous locations. The type of airborne pollutant discharged from smokestacks depends on the type of industrial facility, the type of energy they use, and their chemical processes. For example, electrical generation plants burning coal, or other industries that use coal for their energy source, can emit lead, mercury, zinc, arsenic, uranium, sulfur dioxide, and nitrogen oxides. All of these can enter the water surface either through wet (rain or snow) or dry deposition. The heavy metals, radioactive uranium, and acid rain constituents have different effects on lakes. Downwind lakes close to these smokestacks will receive a heavier loading of heavy metals, whereas the SO_2 and NO_x that contribute to acid rain can travel much farther and degrade more distant lakes. Other industries emitting other kinds of chemicals or metals into the air will affect lakes according to how those metals or chemicals interact with lake chemistry or lake biota.

Atmospheric deposition of mercury is bioconverted by bacteria into methylmercury, a fat-soluble neurotoxin. It then moves up through and beyond the aquatic food chain. Regions receiving deposition of both mercury vapor and acid rain will result in an even higher conversion of elemental mercury in methylmercury, moving first into the water and then into the fish, wiping out entire fish species. Acid rain also mobilizes aluminum, manganese, and zinc from soils into the lake.

PHYSICAL CHARACTERISTICS OF LAKES

What is the assimilative capacity for a lake for certain types or amounts of pollutants? A lake’s ability to absorb pollutants without degrading its biological and chemical processes is a function of its physical characteristics. Some limnological characteristics are described here. (For more information into lake processes and response to pollutants, see Reference 2.) What is the lake’s volume and depth? The greater the lake volume, the more it can dilute pollutants down to less harmful concentrations. Vertical temperature differences found in deep lakes can stratify water into layers. This stratification, and whether or not the lake “turns over” (a mixing of the layers by turbulence in the water) in spring and fall, controls the amount of oxygen, nutrients, and other chemical cycling occurring in the upper or lower waters of the lake.

Therefore, stratification influences which types of pollutants may stay or have an effect only in the upper strata, only in the lower strata, or whether pollutants have access and impact throughout the entire water column and the sediment below. Because stratification controls temperature, it controls the rates of biochemical reactions, and these rates may also provide a buffering effect against some pollutants. Similarly, because stratification controls oxygen levels, pollutant loadings of organic materials (e.g., paper pulp, sewage effluent, manure) will have more negative impact during the times of year that oxygen levels are lowest. Consequently, the same daily amounts of material loadings to a lake can have different effects on the lake depending on the season.

The flushing rate of a lake describes how long it takes for new water inputs to the lake to entirely replace existing lake water with new water. How often the lake theoretically refills with new water affects how fast some of the pollutants entering a lake may be flushed out before they can harm lake processes very much. The higher the flushing rate, the more diluted some pollutants can be because some pollutants will not be able to stay in the lake for long.

The typology of a lake, whether it is naturally more oligotrophic or eutrophic, affects its ability to absorb some pollutants without changing lake characteristics significantly. Over hundreds or thousands of years, a lake naturally receives more nutrient inputs and naturally undergoes eutrophication, or nutrient enrichment, and slowly changes from oligotrophic to eutrophic. When rapid nutrient inputs are due to human action, this is called cultural eutrophication. A more oligotrophic lake is typically deep and has very clear green or blue water, high oxygen, few nutrients or calcium, and limited phytoplankton or other plant species. At the other end of the spectrum is a eutrophic lake, which is shallow and has opaque green or brownish green water, very low oxygen in summer, high nutrients and calcium, and frequent algal blooms in reaction to the high nutrient concentration. So, for example, a naturally oligotrophic lake would become more eutrophic in response to high sewage effluent or farm manure inputs, and a more eutrophic lake would have less assimilative capacity for high loadings of sewage effluent, since the lake is already heavily loaded with nutrients.

POLLUTANT FATE AND CONSEQUENCES

The various sources and transportation mechanisms of pollutants entering lakes of varying characteristics affect lakes in different ways and to different degrees. This section gives a brief sampling of some of the negative effects discharges to lakes can have on the water, sediment, and biota. (For a more comprehensive explanation of the various effects of different pollutant discharges into lakes, see Reference 3.)

Water

Lake water can be degraded in numerous ways. For example, heavy nutrient loading from sewage, manure, and runoff from agricultural soils high in phosphorus reduce the amount of oxygen in the water. Moreover, various chemical compounds and metals entering the lake change the natural rates of chemical cycling in the lake by changing the ion concentrations in the lake. Acid rain forces calcium out of solution as manganese, zinc, and aluminum are brought in.

Sediments

Pollutant deposition into lake sediments can seriously alter existing water chemistry. Chemical pollutants deposited into the lake sediment alter the chemical balance and exchange of chemicals between the sediment and water. Sediment becomes a long-term storage site for pollutants. If sediment gets stirred up later on because of turnover, floodflows from tributaries, or dredging, even more pollutants are released into the water and the food chain and once again chemical exchanges between the lake bottom and the water are altered. For example, lead in sediments deposited up through the 1970s from vehicle exhaust is still found in lake sediments today. Altering water chemistry of a lake affects all biota in the lake.

Plants

Plants are vulnerable in several ways to pollutant discharge. If the discharge into the lake is sediment from agricultural, silvicultural, or construction activities, this new sediment layer can smother existing plants. Chemical pollutants can kill off some aquatic plants. Thermal pollution from some industrial plants or their cooling towers can kill plants by altering water temperature. Invasive species introduced by ballast water or by "hitchhiking" on pleasure boats hauled from one lake to another can wipe out plants by increased depredation. Acid rain reduces phytoplankton as well as zooplankton populations, undermining the aquatic food chain.

Fish

Fish are also very vulnerable to discharges into lakes. Some fish can lose their food source if aquatic plants are harmed or eliminated by pollutants. Thermal pollution from industrial cooling water affects fish metabolism and reproduction. It also stresses fish by increasing oxygen demand in the water and makes fish more vulnerable to disease. Reproduction can also be harmed by contact or ingestion of some chemicals. There might be fewer eggs, damaged eggs, or no eggs. Recently, estrogen in animal manure, human hormones, caffeine, and wood-pulp plant effluents are among some of the suspected causes of male fish producing eggs in their sex organs (4). Chemicals like chlorine are added in sewage treatment or flushed through cooling tower pipes to kill bacterial slimes and then are discharged into the lake and kill fish. Biomagnification in the food chain of substances like mercury, polychlorinated biphenyls (PCBs), or dioxins can pass from the sediments through phytoplankton and zooplankton into smaller and then bigger fish. High acid rain amounts not only wipe out entire fish species, but also cause lakes and streams to become fishless.

Fish can even serve as a transport mechanism for these chemicals from one region to another. For example, salmon have been found to contaminate lakes when they go upstream for spawning and then die in streams and lakes and release the PCBs or dioxins into the lake water and sediments. Beluga whales in the St. Lawrence River downstream of the Great Lakes have suffered impaired fertility and death from the DDT, PCBs, lead, mercury, and insecticides they have been exposed to from discharges into the river as well as the Great Lakes. Once they have been dissected, their bodies must be disposed of according to hazardous waste regulations. Invasive species can harm fish. A northern snakehead, perhaps a former aquarium pet dumped by its owners, was recently caught in Chicago's Burnham Harbor on Lake Michigan (5). Snakeheads are considered "voracious eaters" and can consume many of the smaller fish in a lake. Invasive species may also start competing with existing fish by preying on the same food sources. Another exotic, the zebra mussel, affects fish by clearing the water and thus making fish more visible to prey.

Shellfish

Like fish, shellfish can also experience reduced reproduction from chemicals and hormones. For example, spawning behavior is altered by exposure to some antidepressants. Shellfish may face increased competition for food from invasive species. Shellfish also absorb and concentrate pollutants in their flesh, making them dangerous to eat.

Wildlife

Wildlife such as bear or cormorants can be harmed by biomagnification of PCBs, mercury, and so on when they eat the polluted fish.

Humans

Human populations are also harmed in numerous ways. Drinking water is degraded when a lake has industrial and agricultural chemicals, PPCPs, heavy metals, and so on in its source water. We know from painful experience how mercury causes Minimata disease or how cadmium, a byproduct of metal refinery, causes a disease that softens the bones. Algal blooms from cultural eutrophication give drinking water a bad taste and smell.

Shoreline use and value can also be diminished. Some lake pollution will force beaches to close, reducing recreational resources for locals and possibly keeping tourists away and out of the water because of *E. coli*, the smell and slime of algal blooms, and the toxins some blue-green algae produce. Heavy nutrient inputs into a lake can cause eutrophication that results in water that looks like pea soup and has a horrific smell. These can also cause a reduction in recreational use, when boats get caught up in algal mats or weedy macrophytes and tourists avoid an area, and lower the value of lakeshore properties. If fish or bear are unsafe to eat, then recreational opportunities are diminished. Tourists wanting to hunt and fish may go elsewhere, and those that continue to eat unsafe fish can harm their own health or have babies with birth defects.

PRESENT AND FUTURE CHOICES

Pollutant sources, transport, and fate have serious consequences for lakes. Cities will either have more degraded drinking water, with potential health consequences, or they will have to spend enormous amounts of money to purify the water enough to make it safe for human consumption. The effects of pollutants on aquatic life can not only affect food chains within the lake but can also harm birds, wildlife, and humans that depend on the aquatic species for food. The externalities of industrial production, sewage or stormwater treatment, and other land uses are then spread to new populations as well as rebounding on polluter populations themselves. Their quality of life degrades along with their drinking water, recreational options, and fish and wildlife populations.

BIBLIOGRAPHY

1. For more information, see: http://www.epa.gov/esd/chemistry/pharma/index.htm.
2. Wetzel, R. (2001). *Limnology: Lake and River Ecosystems*, 3rd Edn. Academic Publishers, Burlington, MA, 2001; or S.W. Effler (Ed.) (1996). *Limnological and Engineering Analysis of a Polluted Urban Lake*. Springer, New York.
3. Chiras, X., Reganold, X., and Owen, X. (2002). *Natural Resource Conservation*, 8th Edn. Prentice-Hall, Upper Saddle River, NJ, Chap. 11.
4. Fahrenthold, D.A. (2004). "Male bass in Potomac producing eggs," *Washington Post*. October 15, p. A01.
5. "Snakehead found in Lake Michigan," *Chicago Tribune*. October 14, 2004.

LASERS SCAN LEVEES FROM THE AIR

John Hall
New Orleans District
U.S. Army Corps of Engineers

New Orleans District is taking to the air to learn more about the ground.

For good reason—for large jobs flying surveys are faster and cheaper than ground surveys. And the eye-in-the-sky is particularly handy in New Orleans District's 30,000 square miles of coastal Louisiana, where "ground" is a relative term.

The district is using laser technology called lidar, for LIght Detection And Ranging. It is analogous to radar, which stand for RAdio Detection And Ranging, except that lidar uses light beams instead of radio waves. A scanning laser rangefinder emits a beam of coherent light from a helicopter flying 45 miles per hour at an altitude of 220 feet. (The laser beam is invisible to the naked eye, and not intense enough to cause harm.) The beam bounces back and the lidar system's instruments gather vast amounts of data which can be sliced, diced, and served up in many forms. The process is almost effortless, compared to the labors of ground surveys in wetlands and bayous.

New Orleans District's first use of lidar was on the Lake Pontchartrain and Vicinity Hurricane Protection Project. Total cost was $124,000, with most of the dollars going to the contractor, John E. Chance and Associated, Inc., which specializes in corridor mapping (electric lines, highways, and levees).

A laser scanning rangefinder make a hefty instrument package for a helicopter (Photo courtesy of New Orleans District)

"We can rapidly determine levee heights without doing a ground survey," said Al Naomi, senior project manager of the hurricane project. "You can get a lot of data rapidly."

Indeed. The system produces 8,000 to 20,000 data points per second. But what does all this data really mean?

"It provides all three dimensions, the X, Y, and Z coordinates," said Terral Broussard, a district retiree who works for the contractor. This is trigonometry in action—three coordinates used to locate a point in space. The X and Y axis are length and width; the Z axis is the elevation.

Lidar surveys have been completed for all 125 miles of the Lake Pontchartrain and Vicinity project (www.mvn.usace.army.mil, click on hurricane box). These levees and floodwalls are in St. Bernard, Orleans, Jefferson, and St. Charles parishes. But lidar isn't limited to looking at levees. In New Orleans District, it is being considered for the Comite River Diversion Project, a flood control project to be built near Baton Rouge.

For land surveys, the laser is not the only sensor. Chance's helicopter also carries a video camera, and location is provided by the satellite-based Global Positioning System. The video provides high-resolution color images. On the computer screen, these pictures may be viewed as an overlay precisely placed amid the lidar data. Or, you can just turn the video loose and watch the levee roll by as if you were on the helicopter.

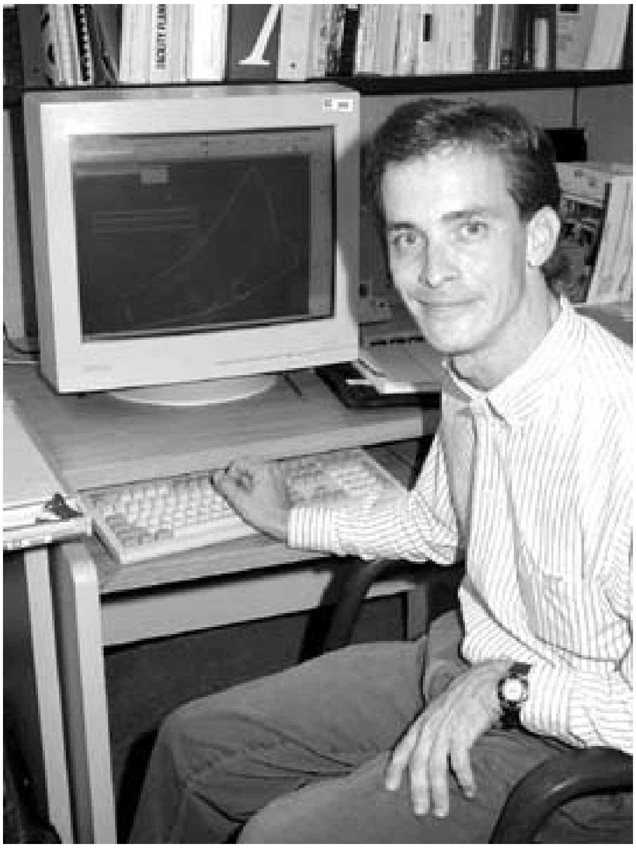

Mike Brennan, civil engineer, projects lidar information on his computer. (Photo courtesy of New Orleans District)

To illustrate, Mark Huber of Survey Section pulled up on his computer a lidar aerial view of the Causeway toll plaza in Metairie, La. Bright colors marked the elevations. In mid-screen, he overlaid a video picture with the tollbooths, parked cars, roadways, and even the bike path that circles underneath, all clearly visible.

"Lidar may not be quite as accurate as ground surveys," Naomi said. "But it's faster, cheaper and requires no landowner interface."

"Other (Corps) districts have used lidar for land mapping," Huber said. "We're the first to pull the final product from the raw data, creating levee cross sections and centerline profiles."

How extensively New Orleans District will use lidar is unclear. While lidar cannot retrieve data from water surfaces, vegetation does reflect its signals. That makes it a candidate for wetland surveys where ground surveying is inherently difficult.

Lidar apparently has a lot of use as a quick look, and for studies. The question is whether it will prove accurate enough for project design. But, according to Huber, no errors exceed six inches compared with ground surveys.

LEVEES FOR FLOOD PROTECTION

Stefano Pagliara
Simone Pozzolini
Università di Pisa
Pisa, Italy

Levees were among the first river training structures built to protect human settlements and activities from floods. A levee is an embankment whose primary purpose is to furnish protection from high water due to river floods.

Levees are similar to small earth dams, but there are important differences:

- Levees are subject to high water levels only for a limited period of time (of the order of hours or rarely days), so that a levee embankment is generally not subject to saturation conditions.
- Levee alignment is dictated primarily by flood protection requirements; this often results in construction on poor foundations.
- Borrow is generally obtained from areas adjacent to the levee site, which produce fill material that is often far from ideal.

Embankments that are subject to water loading for prolonged time periods or permanently should be designed using earth dam criteria.

DESIGN OF LEVEES

Levee design involves numerous factors, dealing with both hydraulic and geotechnical considerations (Fig. 1). Design factors may vary, and no general procedure can be established. However, it is possible to present a general stepwise design procedure (1) based on successful experience that can be used as a starting point in levee design. Such a procedure may be summarized by the following steps:

1. Conduct a geologic study based on a thorough review of available data; initiate preliminary subsurface explorations.
2. Analyze preliminary exploration data and establish preliminary soil profiles, borrow locations, and embankment sections.
3. Initiate final exploration to provide
   - additional information on soil profiles
   - undisturbed strengths of foundation materials
   - more detailed information on borrow areas and other required excavations.
4. Using the information obtained in step 3,
   - determine both embankment and foundation soil parameters, and refine preliminary sections where needed.
   - compute rough quantities of suitable material, and refine borrow area locations.
5. Divide the entire levee into reaches of similar foundation conditions, embankment height, and fill material, and assign a typical trial section to each reach.
6. Analyze each trial section for
   - foundation and embankment seepage;
   - slope stability;
   - settlement;
   - trafficability of the levee surface;
   - maintenance requirements.
7. Design special treatment to preclude any problems determined from step 6.
8. Based on the results of step 7, establish final sections for each reach, and define construction methods.
9. Design embankment slope protection.

Field Investigation and Laboratory Testing

Levee construction requires both field and laboratory investigation.

Field Investigation. The extent of field investigation depends on several factors (1). Conditions requiring extensive field investigation include poor experience in the area with respect to levee performance, significant levee height, bad foundation conditions, long high water duration, low quality of levee materials, high risk conditions in case of levee failure, and the presence of a concrete structure in the levee alignment.

Field investigation and testing are generally performed in three successive steps:

1. preliminary office geologic study;
2. field geologic survey;
3. field testing (preliminary phase and final phase).

Each step is described in Table 1.

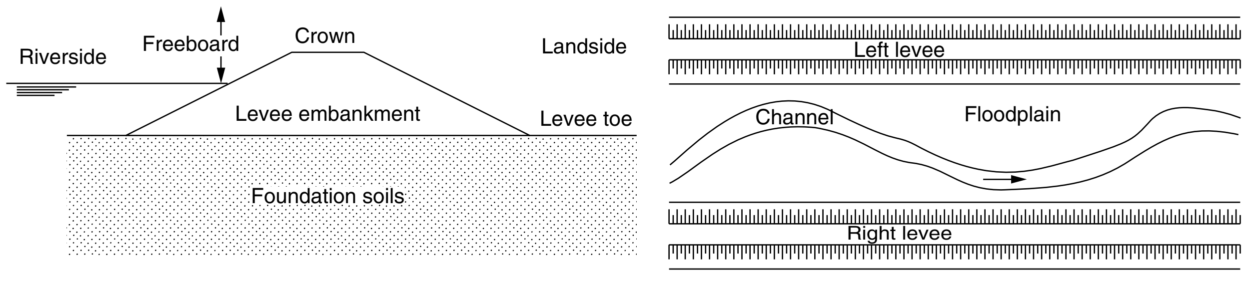

Figure 1. Sketch of a levee (cross section and plan view).

Table 1. Office Study and Field Investigation

Office study	Collection and study of — topographic, soil, and geologic maps — aerial photographs — existing information on soil and ground characteristics
Field survey	Observations and geologic analysis of the area, including such features as — riverbank slopes, rock outcrops, earth and rock cuts or fills — surface materials — poorly drained areas — evidence of instability of foundations and slopes — emerging seepage. — Physiographic features
Field testing	Preliminary phase — Widely spaced disturbed sample borings and test pits — Geophysical surveys or cone penetrometer test, used to interpolate between widely spaced borings — Borehole geophysical tests — Preliminary estimates of foundation strength Final phase — Additional disturbed sample borings — Undisturbed sample borings in critical reaches — Field vane shear tests, useful if undisturbed samples are unavailable. — Field pumping tests, providing information on large scale permeability of *in situ* deposits — Water table observations by piezometers in foundations and borrow areas

Laboratory Testing. Laboratory testing programs for levees vary from minimal to extensive, depending on the importance of the project and on the site conditions. Soil tests to be performed will vary according to the soil classification (Table 2).

Embankment Design

A levee cross section is generally trapezoidal berms may be added to ensure stability or seepage control. Berms are always used for high levee design.

Embankment design requires detailed analysis for levees of significant height, bad foundation conditions, or unsuitable embankment materials.

Embankment Geometry

Slope. Mechanical stability is largely dependent on the side slopes of the levee. The slope steepness is generally selected by considering the types of compaction, fill material characteristics, and foundation strength. For high levees, a soil mechanics stability analysis is required. For lower levees laid on proven foundations, recommended riverside slopes are about 1V:3H÷1V:3.5H; landside slopes are about 1V:2H÷1V:2.5H (V = Vertical; H = Horizontal).

Freeboard. Freeboard (the difference between the levee crown elevation and the high water stage) is used to account for uncertainties in levee performance. Many levees have been designed that have a 1-meter freeboard. Nowadays, the freeboard concept is less used in levee design. Risk-based analysis directly accounts for hydraulic uncertainties, and the geotechnical uncertainties are considered in the deterministic analysis of foundations and embankment behavior.

Crown Width. The width of the levee crown depends primarily on roadway requirement; it is in general in the range of 2.5–5 meters. To provide access for normal maintenance operations and flood fighting operations, minimum widths of 3.0–3.5 m are commonly used.

Stability Analysis. The principal mechanisms of levee failure are

Table 2. Laboratory Testing

Soil tests of fine-grained cohesive materials

- — Visual classification and water content determination
- — Atterberg limits
- — Permeability
- — Consolidation
- — Compaction
- — Shear strength

Soil tests of pervious materials

- — Visual classification
- — *In situ* density and relative density determination
- — Granulometry
- — Permeability
- — Consolidation
- — Shear strength

- — overtopping
- — piping
- — embankment or foundation slides
- — surface erosion.

Common methods used to analyze levee embankment stability are (2)

- — methods assuming definite boundary stress values;
- — methods based on the various soil characteristics that determine the failure state;
- — methods considering the degree of safety for force and momentum equilibrium;
- — methods comparing deformation work with the potential energy of the structure.

Primary attention should be devoted to the problems of defining the shear strengths, unit weights, geometry, and limits of possible sliding surfaces.

The more severe loading conditions for a levee are (1)

- — End of construction. Levee embankment and foundation stability must be analyzed using drained conditions for free draining soils and undrained conditions for low permeability soils.
- — Sudden drawdown. Excess pore water pressure can develop if a prolonged flood stage saturates the upstream embankment portion and then falls faster than the soil can drain. This can negatively affect riverside slope stability.
- — Steady seepage from full flood stage. Flood water levels can sometimes fully saturate levee embankments. This is an uncommon condition for ordinary levees but may be critical for landside slope stability.
- — Earthquake. Earthquake loadings are not normally considered in analyzing the stability of levees because of the low probability of coincidence with floods. Depending on the seismic characteristics of the study area, determination of liquefaction susceptibility may instead be required.

Shear strength selection under various loading conditions must consider soil type (free draining or low permeability) and pore water pressure conditions. Minimum factors of safety for the various stability analyses should be selected according to local regulations and to sound engineering judgement, considering failure consequences.

Measures to Increase Stability

Methods of Improving Embankment Stability. Levee stability can be enhanced by changes in embankment section, such as

Flattening Embankment Slopes. Flattening embankment slopes will usually increase the stability of an embankment against a shallow foundation type failure.

Stability Berms. Berms concentrate additional weight where it is most needed and force a substantial increase in the failure path. Berms can be an effective means of stabilization for shallow foundation and embankment type failures and also for more deeply seated foundation failures.

Methods of Improving Foundation Stability. Levees located on foundation soils of inadequate shear strength require some type of foundation treatment or the use of alternative design features in the levee section, such as a wider levee crest and flatter slopes. Foundation deposits that are prone to cause problems include very soft clays, sensitive clays, loose sands, natural organic deposits, and debris deposited by humans.

Means of dealing with inadequate foundation soils include

Excavation and Replacement. This procedure, though extremely effective, is economically feasible only where deposits of unsuitable material are not excessively deep.

Stage Construction. A levee is built in successive intervals of time where the strength of the foundation material is inadequate to support embankment weight, if built continuously at a pace faster than the foundation material can drain. Rest periods during levee construction permit dissipation of pore water pressures.

Stage construction is appropriate when pore water pressure dissipation is reasonably rapid. The rate of consolidation can be increased by using vertical drains. The disadvantages of this method are the delays in construction operations and uncertainty as to its scheduling and efficiency.

Densification of Loose Sands. Sands is generally not densified due to economic constrictions; possible exceptions are the presence of important structures in a levee system.

Levee Settlement. Evaluation of postconstruction settlement is important when settlement can result in loss of freeboard of the levee or damage to structures in the embankment. Detailed settlement analyses by theoretical analysis should be made when significant consolidation is expected, as under high embankment loads, embankments of highly compressible soil or built on compressible foundations, and near structures in levee systems founded on compressible soils.

Where foundation and embankment soils are pervious or semipervious, most of the settlement will occur during construction. For impervious soils, it is usually conservatively assumed that all the calculated settlement of a levee built in a normal sequence of construction operations will occur after construction.

Sometimes a levee is overbuilt by a given percentage of its height (varying from 0–10%) to take into account anticipated settlement of the foundation and within the levee fill itself, but overbuilding increases the severity of stability problems and may be impracticable for some foundations.

Seepage Control

Seepage flows develop when river waters flow through the embankment body or through foundation soils. Seepage is a very slow process, and the establishment of a seepage line generally requires a long time (2). As levees are exposed to high water levels for relatively short periods, the danger from seepage is more limited than that for earth dams.

Foundation Underseepage. Without control, underseepage flows in pervious foundations beneath levees may result in excessive pore water pressure, rupturing phenomena, and piping beneath the levee embankment. The principal seepage control measures are described in the following paragraphs (1). The effectiveness of seepage control measures can be evaluated by determination of flow-net, empirical methods, or mathematical solutions of seepage flow equations (Fig. 2).

Cutoffs. A cutoff beneath a levee stratum is one of the most positive means for blocking seepage through pervious soils. Cutoffs may consist of excavated trenches backfilled with compacted earth or slurry trenches usually located near the riverside toe. Steel sheet piling is generally not entirely watertight due to interlock leakage, but its use can significantly reduce the possibility of foundation piping. A cutoff must penetrate approximately 95% or more of the thickness of pervious strata to be effective, so constructing it is not economically feasible in the presence of thick pervious strata.

Riverside Blankets. Levees are frequently situated in alluvial areas on foundations that have natural surface covers of fine-grained impervious to semipervious soils overlying pervious sand and gravel. If these surface strata (or blankets) are continuous and extend riverward for a considerable distance, they can effectively reduce seepage flow and pore pressures landside of the levee.

The effectiveness of the blanket depends on its thickness, length, distance to the levee riverside toe, and permeability. Protection of the riverside blanket against erosion is an important feature.

Landside Seepage Berms. Landside berms provide additional weight useful for counteracting uplift seepage forces. Berms also cause an increase in the hydraulic length of seepage flows, providing a reduction in pressure forces at the landside toe.

Seepage berms can improve the performance of existing impervious or semipervious top stratam but may also be placed directly on pervious deposits. Moreover, seepage berms increase the levee's landside slope stability.

Pervious Toe Trench. Partially penetrating toe trenches can improve seepage conditions close to the levee toe in a levee situated on deposits of pervious material overlain by little or no impervious material.

Pervious toe trenches are useful for controlling shallow underseepage. To collect deeper seepage flows, relief well systems may be used in conjunction with pervious toe trenches. Trench dimensions will vary according to the volume of expected underseepage, desired reduction in uplift pressure, and construction practicality. The trench backfill must be designed as a filter material. To improve drainage efficiency, trenches are often provided with perforated pipes to collect seepage.

Pressure Relief Wells. Pressure relief wells are used to reduce uplift seepage forces by providing controlled outlets for seepage that would otherwise emerge uncontrolled landward of the levee. They are installed where pervious strata underlying a levee are too deep to be penetrated by cutoff or toe drains and when the construction of landside seepage berms is not feasible due to space limitations.

The design of pressure relief well systems involves the determination of well spacing, size, and penetration in the underlying strata. Factors to be considered include

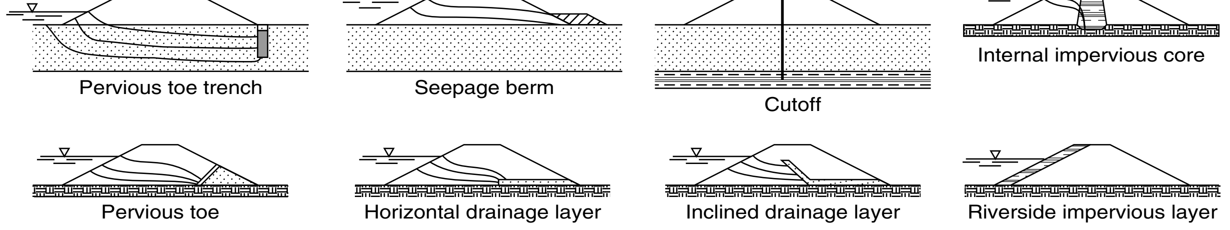
Pervious toe trench
Seepage berm
Cutoff
Internal impervious core
Pervious toe
Horizontal drainage layer
Inclined drainage layer
Riverside impervious layer

Figure 2. Seepage control devices.

depth, stratification and permeability of foundation soils, distance to the effective source of seepage, characteristics of the landside top stratum, and the degree of pressure relief desired.

Seepage Through Embankments. Embankment seepage emerging on the landside slope results in high seepage forces that decrease slope stability. Moreover seepage can soften fine-grained fill in the vicinity of the landside toe or lead to fine materials piping. Generally flood water stages do not act against a levee long enough to make seepage flows exit on the landside slope, but the combination of the effects of flood stages with a period of heavy precipitation may lead to consistent seepage flows. Various means to manage seepage trough embankments are described hereafter.

Pervious Toe Drain. Pervious toe drains provide a ready exit for seepage through an embankment and can lower the phreatic surface sufficiently so that no seepage will emerge on the landside slope. They can also be combined with partially penetrating toe trenches as a method for controlling shallow underseepage.

Horizontal Drainage Layers. Horizontal drainage layers essentially serve the same purpose as pervious toes but can extend further under the embankment requiring a relatively small amount of additional material. They can also protect the base of the embankment against high uplift pressures where shallow foundation underseepage is occurring. Sometimes horizontal drainage layers also carry away seepage from shallow foundation drainage trenches.

Inclined Drainage Layers. An inclined drainage layer is very effective for controlling internal seepage and is used extensively in earth dams, but it is rarely used in levee construction because of its cost. It might find use for short levee reaches in urban areas. An inclined drainage layer completely intercepts embankment seepage regardless of the degree of stratification in the embankment or the material type riverward or landward of the drain. Inclined drains must be tied into horizontal drainage layers to provide an exit for the collected seepage.

The design of pervious toe drains and horizontal and inclined drainage layers must ensure that such drains have adequate thickness and permeability to transmit seepage without appreciable head loss and, at the same time, prevent migration of finer soil particles. Moreover, the design of drainage layers must satisfy the criteria for filter design.

Landside Berms. Landside berms can also be very useful in embankment seepage control. For a brief description of their characteristics, refer to the previous paragraphs.

Embankment Zoning. To counteract seepage flows, a central impervious core or covering a riverside slope with an impervious layer can be useful. The latter is generally more economical than a central impervious core and, in most cases, is entirely adequate. As a matter of fact, levee embankments are often constructed as homogeneous sections as zoning is usually used in earth dams or in major levees.

PROTECTION OF RIVERSIDE SLOPES

The protection needed on a riverside slope to withstand the erosional forces of waves and stream currents depends on a number of factors such as floodwater stage duration, bank slope, erosion susceptibility of the embankment materials, presence of abrupt transitions, short-radius bends, or structures riverside of the levee that increase flow turbulence.

Several types of slope protection have been used. In general, high-class slope protection, such as riprap, articulated mat, soil cement, or paving should be provided on riverside slopes beneath bridges and adjacent to structures passing through levee embankments.

LEVEE CONSTRUCTION

Almost any soil is suitable for constructing levees, except very wet, fine-grained soils or highly organic soils. Levees of homogeneous cross section are generally made of silt, silty sand, or sandy clay. Fine-grained soils act as sealing materials due to their low permeability, but coarse components increase levee stability. The choice of material is often governed by the extent of local soils and by the technological experience of the contractor.

Levee embankments can be built with various degree of compaction. The central portion of the embankment may be compacted or semicompacted, but riverside and landside berms (for seepage or stability purposes) may be constructed of uncompacted or compacted fill. When foundations have adequate strength and where space is limited in urban areas both with respect to the quantity of borrow and levee geometry, compacted levee fill construction by earth dam procedures is frequently selected. This involves using select material, water content control, and compaction procedures. Where foundations are weak and compressible, high-quality fill construction is not justified because these foundations can support only levees with flat slopes. In such cases, uncompacted or semicompacted fill is appropriate. Uncompacted fill is generally used where the only available borrow is very wet and frequently has high organic content and where rainfall is very high during the construction season.

Except in seismically active areas or other areas requiring a high degree of compaction, compaction by vibration other than that afforded by tracked bulldozers is not generally necessary.

LEVEE REHABILITATION

As many levee systems are very old, the rehabilitation of existing levees is gaining greater prominence. Many techniques have been studied to rehabilitate existing levees. Recently, new working techniques that use geosynthetic materials extensively, have been introduced. Table 3 shows conventional and innovative levee rehabilitation techniques (3).

Table 3. Levee Rehabilitation Techniques

Problem	Conventional Rehabilitative Method	Innovative Rehabilitative Method
Overtopping protection	Rebuild	Cellular confinement system
	Vegetation	Reinforced grass
	Concrete slabs	Soil cement
Current and wave attack	Vegetation	Reinforced grass
	Revetment	Concrete block system
	Gabions	Soil cement
Rainfall surface erosion	Vegetation	Turf reinforcement mat
	Chemical stabilization	
Embankment seepage	Toe drain	Biopolymer chimney drain
	Conventional chimney drain	Biopolymer toe trench
Foundation seepage	Conventional toe trench	Jet grouted cutoff
	Conventional cutoff	
	Riverside blanket	
	Landside seepage berm	
	Pressure relief well	
Slope instability	Drainage	Reinforced soil slope
	Slope flattening and benching	Soil nailing
	Conventional restraint structure	Pin pile
	Chemical treatment	Stone-filled trench
		Geosynthetic drainage system
		Anchored geosynthetic system

BIBLIOGRAPHY

1. U.S. Army Corps of Engineers. (2000). *Design and Construction of Levees*. Engineering Manual No. 1110-2-1913 30.
2. Peter, P. (1982). *Canal and River Levees*. Elsevier Scientific Publishing Company, Amsterdam.
3. Perry, E.B. (1998). *Innovative Methods for Levee Rehabilitation*. Prepared for U.S. Army Corps of Engineers, Technical report REMR-GT-26.

LIMNOLOGY

Walter Geller
Annette Geller
UFZ Center for Environmental Research
Magdeburg, Germany

Limnology is the science of inland waters, also called epicontinental waters (1), including fresh and saline, standing (lenitic) and running (lotic) surface waters, estuaries, wetlands (2,3), and hyporheic (subterranean) waters (4) as ecosystems. Limnology integrates hydrological, physical, and chemical factors as framework constraints for ecological structures and processes. The first comprehensive publication coining the term "Limnology" was written by F. A. Forel (5) on Lake Geneva (Lac Léman) with chapters on geography, hydrography, geology, climatology, hydrology, hydraulics, thermics, optics, acoustics, chemistry, biology, history, navigation, and fisheries (see Ref. 6).

Limnology was traditionally focused on lake systems, as seen by the classic textbook of Hutchinson, *A Treatise on Limnology*, in Refs. 7–12. Other textbooks on limnology that present lake systems in detail, running waters, however, are treated only in a few chapters (13–16). In 1922, the International Association of Theoretical and Applied Limnology (SIL) was founded, and limnology became the science of aquatic ecosystems. The volumes of the book series *Die Binnengewässer* (edited by A. Thienemann since 1925) present the basics of limnological knowledge. The first concepts to compare lakes (17–20) were directed to lake types, referring to size and morphology, climate, the behavior to stratify and mix seasonally or diurnally (21), and different nutrient contents resulting in different productivity and oxygen regimes (temperate, tropical and polar lakes, mono-, dimictic, and polymictic lakes, oligotrophic and eutrophic lakes, clear-water and dystrophic brown-water lakes). Regional limnology was included by Hutchinson (7) and (22) with comprehensive surveys over the lakes of the world, their mode of origin, size, and number.

Studies of hydrobiologists first focused on an inventory of plants and animals, and then on the microorganisms in standing and running waters. A more general concept of productivity and energy flow in food chains was developed to describe the input and fate of matter and energy in ecosystems (23). Lakes all over the world were investigated according to Lindeman's "Trophic-Dynamic Concept" on a large-scale comparison of the functioning of lake ecosystems (24), being a part of the International Biological Programme (IBP). The nonbiological fields of limnology were treated by physicists (25–27) and chemists (28,29) in basic reviews and the textbooks *Aquatic Chemistry* and *Physics and Chemistry of Lakes*.

Institutions for limnological research, all focusing on lake limnology, were founded during the early decades of the twentieth century in Austria, England, France, Germany (including a first station on tropical limnology in Manaos, Brazil), Italy, the Netherlands, Sweden, Switzerland, Russia at Lake Baikal, the United States, and Canada. The only early stations for river ecology were at the Amazon and Volga rivers.

The limnology of creeks and rivers as part of aquatic ecology developed later than lake limnology (30,31). As running waters are more difficult to characterize, the development of concepts considered to suitably cover the relevant processes still proceeds. Typologies of rivers can be based on abiotic features as hydrological characteristics, water flow and its variability, the amount of dissolved and suspended matter, salt content and turbidity, and the course of temperature along the river and through time (32). *The Rivers Handbook* of Calow and Petts (33) has chapters on hydrology, physical and chemical characteristics, the biota, the ecosystem, and perturbations and biological impacts, and it addresses monitoring programs; modeling of hydrological development, of water quality, of sediment transport, and of biological responses; and the evolution of habitats.

First, in the early twentieth century fisheries, biologists divided European rivers along their stretch into distinct zones according to slopes, flow velocities, and the communities of fish (zones of salmonids, of cyprinids) (34). Benthos biologists amended this approach by including benthic invertebrates (rhitron, potamon) (35). In 1980, Vannote et al. (36) established the "River Continuum Concept" using the continuous shift along river stretches (1) of the relation between autotrophy and heterotrophy, and (2) of the relation between autochthonous nutrients, that are produced in the river, and allochthonous, terrestrial-borne nutrients. Rare or seasonal events as floods temporarily include the floodplains into the riverine system. The pulsed changes of dry and high-water conditions are essential for these systems adapted to regular floods, which affect the floodplains and offer new ecological niches (e.g., fruit-eating fish species in flooded Amazon forests) (37). The hereby derived "Flood Pulse Concept" sees the river as part of a larger ecosystem including the floodplains as system components being temporarily dry land or shallow flooded side-parts of the stream. The ecological continuity from water to land is approached by the "Concept of Ecotones" (38). The scale of integrated characteristics and management of the landscape results from enclosing the whole catchment area of the river system, based on the aquatic ecosystems, both rivers and lakes as partial systems. To characterize the structure and the hydrological processes within landscapes with respect to water flows, the integrated approach of the concept of "Ecohydrology" (39) was developed. The scale of catchment areas is the basis for the European Union (EU) "Water-Framework-Directive" that was established to keep or to reach a good/maximum chemical state and a good/maximum ecological state in all running waters and lakes of the EU-member states (40,41). Furthermore, special commissions have been established to deal with trans-boundary problems of water protection of rivers and lakes (Table 1).

Human use of land and waters changed the quality of inland waters with increasing population densities and by increasing loads of waste and nutrients impacting lakes and rivers. The limnology of creeks and rivers was predominantly developed within applied sciences, aiming at fisheries, water pollution by wastewater, the limits of self-purification of running waters, and solution of problems caused by pollution. Rivers have received urban wastewater for a long time. The pollution effects were first scientifically described at the beginning of the twentieth century. Kolkwitz and Marsson (42,43) characterized rivers from the source of pollution along stretches of self-purification by the benthic community's changes of species composition, being the base of the "Saprobic System." The saprobic system is still used as a biological tool to discriminate polysaprobic, mesosaprobic, and oligosaprobic conditions (44).

Wastewater pollution is a more severe threat for lakes than for rivers, because the organic loads, being nutrients for the aquatic organisms, remain in the lakes for years according to their water renewal time. The cumulative enrichment of nutrients, especially of phosphorus and nitrogen, with its biological consequences in lake waters (e.g., algal blooms) and the increasing of oxygen deficiencies in deep water layers, called "eutrophication," was observed in most lakes in densely settled areas of Europe and in North America. The increase of total phosphorus in lakes depends on the population density in the lake's catchment area. The limits for acceptable P-loadings of morphologically different lakes have been treated in comparative studies of the OECD in Europe and in Canada (45–48).

Systematic measures to reduce the inputs of P and N into surface waters have been taken for several decades (e.g., purification of wastewater in treatment plants, banning of detergents containing phosphate, abstraction of wastewater from lakes by ring-pipes). The results in many lakes clearly show that these efforts have been successful. The present state of nutrient loads of rivers and lakes in Europe and the development during the past is documented in reports, which can be downloaded from the European Environmental Agency (http://www.eea.eu.int/). Eutrophication of European lakes, in terms of phosphorus concentrations, generally is decreasing. The actual phosphorus concentrations are highest in the Eastern European countries and lowest in the Nordic countries. The state of surface waters in North America is documented in reports accessible from the U.S. Environmental Protection Agency (http://www.epa.gov).

Geological regions poor in carbonates are sensitive to acid rain as it changes the chemical composition of seeping runoff and ground- and surface-waters (49). Rain-acidified inland waters are carbonate-free, enriched in sulfate and aluminum, and acidic within a range between 4.5 and 5.5 pH units. Reports on long-term development and large-scale geographic distribution in EU-countries (50) showed that in Norway 6000 lakes (51) and in Sweden 17,000 lakes (52) were acidified. In southern Norway, 9630 fish populations were lost (53). Measures against acid rain and acidification of soils, lakes, and rivers started by reducing the emissions of acidic smoke (54) and direct treatment

Table 1. Commissions Established to Manage Inland Waters, River Basins, and Lake Catchments Including Trans-boundary Problems Within Their Catchment Areas. The Table Exemplarily Shows National and International Commissions and the Respective Lakes and Rivers

River, Lake, Basin	Commission	Web Page
Danube	International Commission for the Protection of the Danube River (ICPDR)	http://www.icpdr.org/pls/danubis/DANUBIS.navigator
Rhine	International Commission for the Protection of the Rhine River (IKSR)	http://www.iksr.org/
Elbe	International Commission for the Protection of the River Elbe (IKSE, SMOL)	http://www.ikse-mkol.de/html/ikse/ikse/deutsch/indexd.htm
Oder	International Commission for the Protection of the River Oder (signed by Germany, Czech Republic and Poland on 11 April 1996)	
Dnieper	International DNIPRO Fund (IDF)—National Program of Environmental Sanitation of River Dnipro Basin and Drinking Water Quality Improvement	http://greenfield.fortunecity.com/hunters/228/toppage1.htm
Bodensee, Lake Constance	International Commission for the Protection of the Lake Constance	http://www.igkb.de
Lake Geneva	Commission Internationale pour la Protection des Eaux du Lac Léman contre la pollution (CIPEL)	http://www.cipel.org
Great Lakes Basin	The International Joint Commission prevents and resolves disputes between the United States of America and Canada under the 1909 Boundary Waters Treaty. It assists the two countries in the protection of the trans-boundary environment, including the implementation of the Great Lakes Water Quality Agreement.	http://www.ijc.org http://www.epa.gov/glnpo
Lake Tahoe	Lahontan Regional Water Quality Control Board	http://www.epa.gov
Lake Washington	The Cedar River Basin Plan combines a traditional King County Basin Plan with a Non-point Source Pollution Action Plan. The Plan was prepared under the policy direction of the Watershed Management Committee composed of key tribal, state and local government agencies and non-govern-mental organizations, and a Citizens Advisory Committee of area residents.	http://dnr.metrokc.gov/wlr/watersheds/cedar-lkwa.htm/
Mississippi River Basin	The Mississippi River Initiative began 1997 and has developed into a coordinated federal effort to keep illegal pollution, ranging from raw sewage to industrial waste, out of the river and to restore the river and surrounding communities. The Initiative employs cooperative efforts of the Dept. of Justice, EPA, Customs Service, Coast Guard, Fish and Wildlife Service, FBI, and institutions of 26 U.S. states of the Mississippi River Basin.	http://www.epa.gov
Aral Sea basin	Inter-State Commission for Water Coordination (ICWC): Water ministers of the five states in the basin, Kazakhstan, Kyrgyzstan, Turkmenistan, Tajikistan and Uzbekistan; the ICWC was established 1992 with joint responsibility for water management with the two river basin agencies (Amu-Darya and Syr-Darya).	

Note: See IWAC (www.iwac-riza.org) for complete list of trans-boundary co-operations in Europe.

of acidic waters. In Sweden, 200,000 tons of lime were spread every year over 20 years, and 6000 lakes were treated. About half of the acidified area was restored until the early 1990s (55–57).

Manmade lakes worldwide amount to tens of thousands of reservoirs and pit lakes from surface mining on coal, gravel, and ores. Reservoirs show limnological peculiarities (58,59) that primarily result from the management regime. The state of reservoir building and its problems are presented by the World Dam Commission (http://www.wdc.org). The water chemistry of pit lakes from coal or ore mining is determined by geogenic impacts, mostly acidification and contamination by heavy metals. Reviews and other comprehensive scientific publications on mining lakes are in an initial state (60).

Some traditional institutions for limnology investigated "their" lakes over a long period. Eutrophication and acidification of surface waters by acid rain were recognized because of these long-term observations. The effects of countermeasures, stepwise reoligotrophication, and neutralization were also monitored hereby. Only long datasets by traditional long-term investigations (e.g., "Long Term Ecological Research") render possible the analyses of the responses of lakes and rivers (61). Presently, these datasets are reevaluated to track the effects of long-term changes of climate on the physical, chemical, and ecological behavior of lake systems and rivers (62,63). In addition, paleolimnologists get information for previous changes of climate and of nutrient inputs altering in the course of historical and geological times, by investigating the fossil rests of organisms and other suitable proximate-indicators in laminated lake sediments (64–66). In most lakes of temperate regions, the sediments show data reaching back to the end of the last glaciation, and to some million years in the oldest, tertiary lakes like Lake Biwa, Japan (67).

BIBLIOGRAPHY

1. Margalef, R. (1994). *Limnology Now.* Elsevier, Amsterdam.
2. Mitsch, W.J. (1994). *Global Wetlands—Old World and New*. Elsevier, Amsterdam.
3. Roggeri, H. (1995). *Tropical Freshwater Wetlands*. Kluwer Academic Publishers, Dordrecht.
4. Schwoerbel, J. (1961). *Arch. "Hydrobiol." Suppl.* **25**: 182–214.
5. Forel, F.A. (1892). *Le Léman: Monographie Limnologique*. F. Rouge, Lausanne, pp. 543, 651, and 715.
6. Edmondson, W.T. (1994). *Limnology Now: A Paradigm of Planetary Problems*. R. Margalef (Ed.). Elsevier, Amsterdam.
7. Hutchinson, G.E. (1957). *A Treatise on Limnology. 1: Geography, Physics, and Chemistry*. John Wiley & Sons, New York.
8. Hutchinson, G.E. (1967). *A Treatise on Limnology. 2: Introduction to Lake Biology and the Limnoplankton.* John Wiley & Sons, New York.
9. Hutchinson, G.E. (1975). *A Treatise on Limnology. 3: Limnological Botany.* John Wiley & Sons, New York.
10. Hutchinson, G.E. (1993). *A Treatise on Limnology. 4: The Zoobenthos.* John Wiley & Sons, New York.
11. Ruttner, F. (1962). *Grundriß der Limnologie. 3. Aufl.* Walter de Gruyter & Co., Berlin.
12. Pourriot, R. and Meybeck, M. (1995). *Limnologie générale*. Masson, Paris.
13. Margalef, R. (1983). *Limnología.* Omega, Barcelona.
14. Horne, A.J. and Goldman, C.R. (1994). *Limnology*. McGraw-Hill, New York.
15. Wetzel, R.G. (2001). *Limnology—Lake and River Ecosystems*, 3rd Edn. Academic Press, San Diego.
16. Schönborn, W. (2003). *Lehrbuch der Limnologie*. E. Schweitzerbart'sche Verlagsbuchhandlung, Stuttgart.
17. Birge, E.A. and Juday, C. (1911). *Bull. Wisc. Geol. Nat. Hist. Surv.* **22**: 259.
18. Naumann, E. (1919). *Lunds Univ. Arsskrift N.F.II* **17**: 1–22.
19. Thienemann, A. (1915). *Verh. Naturh. Ver. Preuss. Rheinl.* **71**: 281–389.
20. Thienemann, A. (1918). *Arch. Hydrobiol.* **12**: 1–65.
21. Hutchinson, G.E. and Löffler, H. (1956). *Proc. Nat. Acad. Sci. Wash.* **42**: 84–86.
22. Meybeck, M. (1995). *Physics and Chemistry of Lakes.* A. Lerman, D.M. Imboden, and J.R. Gat (Eds.). Springer, Berlin.
23. Lindeman, R.L. (1942). *Ecology.* **23**: 339–418.
24. Le Cren, E.D. and Lowe-McConnell, R.H. (1980). *The Functioning of Freshwater Ecosystems. International Biological Programme*, 22. Cambridge Univ. Press, Cambridge.
25. Mortimer, C.H. (1974). *Mitt. Internat. Verein. Limnol.* **20**: 124–197.
26. Imberger, J. and Patterson, J.C. (1990). *Advances in Applied Mechanics 27*. J.W. Hutchinson and T.Y. Wu (Eds.). Academic Press, Cambridge.
27. Imboden, D.M. and Wüest, A. (1995). *Physics and Chemistry of Lakes.* A. Lerman, D.M. Imboden, and J.R. Gat (Eds.). Springer, Berlin.
28. Stumm, W. and Morgan, J.J. (1995). *Aquatic Chemistry*, 3rd Edn. John Wiley & Sons, New York.
29. Lerman, A., Imboden, D., and Gat, J. (1995). *Physics and Chemistry of Lakes*. Springer-Verlag, Berlin.
30. Behning, A. (1928). *Das Leben der Wolga.* Schweizerbart'sche Verlagsbuchhandlung, Stuttgart.
31. Hynes, H.B.N. (1970). *The Ecology of Running Waters.* University of Toronto Press, Toronto.
32. Meybeck, M. (1996). *Verh. Internat. Verein. Limnol.* **26**: 81–96.
33. Calow, P. and Petts, G.E. (1992). *The Rivers Handbook*. Blackwell, Oxford, pp. 523, 526.
34. Huet, M. (1949). *Schweiz. Z. Hydrol.* **11**: 332–351.
35. Illies, J. and Botosaneanu, L. (1963). *Mitt. Internat. Verein. Limnol.* **12**: 1–57.
36. Vannote, R.L. et al. (1980). *Can. J. Fish. Aquat. Sci.* **37**: 130–137.
37. Junk, W.J., Bayley, P.B., and Sparks, R.E. (1989). *Canadian Special Publication of Fisheries and Aquatic Sciences* **106**: 110–127.
38. Naiman, R.J., Decamps, H., and Fournier, F. (1989). *The Role of Land/Inland Water Ecotones in Landscape Management and Restoration—a Proposal for Collaborative Research.* UNESCO, Paris.
39. Zalewski, M., Janauer, G.A., and Jolankaj, G. (1997). *Ecohydrology: A New Paradigm for the Sustainable Use of Aquatic Resources. Technical Documents in Hydrology* No. 7. UNESCO, Paris.
40. EU. (2000). *EU-Water-Framework-Directive (WFD): Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy.* Official Journal L 327, 22/12/2000.
41. WATECO. (2000). *Economics and the Environment—The Implementation Challenge of the Water Framework Directive, a Guidance Document. Brüssel.* Available: http://www.forum.europa.eu.int
42. Kolkwitz, R. and Marsson, M. (1908). *Ber. Bot Ges.* **26**: 505–519.
43. Kolkwitz, R. and Marsson, M. (1909). *Int. Rev. Ges. Hydrobiol.* **2**: 126–159.
44. Liebmann, H. (1951). *Handbuch der Frischwasser- und Abwasser-Biologie.* Oldenbourg, München.
45. Vollenweider, R.A. (1968). *Die wissenschaftlichen Grundlagen der Seen- und Fließgewässereutrophierung unter besonderer Berücksichtigung des Phosphors und des Stickstoffs als Eutrophierungsfaktoren.* 69. Report DAS/CSJ. OECD.
46. Vollenweider, R.A. (1980). *OECD Eutrophication Programme. The Canadian Project.* OECD Eutrophication Programme.
47. Vollenweider, R.A. and Kerekes, J. (1980). *OECD Eutrophication Programme. Synthesis Report.* OECD Eutrophication Programme. OECD, Burlington.

48. Fricker, H.J. (1980). *OECD Eutrophication Programme. Regional Project Alpine Lakes*. Swiss Federal Board for Environmental Protection, Bern.
49. Stumm, W. and Schnoor, J.L. (1995). *Physics and Chemistry of Lakes*, 2nd Edn. A. Lerman, D. Imboden, and J. Gat (Eds.). Springer, Berlin.
50. Luckewille, A., Jeffries, D.S., Johannessen, M., Raddum, G., Stoddard, J.L., and Traaen, T. (1997) *The nine year report: acidification of surface water in Europe and North America—long term developments (1980s and 1990s)*. NIVA Report serial no. 3637-97.
51. Henriksen, A., Skjelkvale, B.L., Mannio, J., Wilander, A., Jensen, J.P., Moiseenko, T., Harriman, R., Traaen, T., Fjeld, E., Vuarenmaa, J., Kortelainen, P., and Forsius, M. (1997) *Regional Lake Surveys 1995 in Finland, Norway, Sweden, Denmark, Russian Kola, Russian Karelia, Scotland and Wales—Results*. NIVA Report serial no. 3645-97.
52. Leonard, J. and Crouzet, P. (1999) *Acidification*. Topic Report. Copenhagen. Available: http://europa.eu.int. p. 85
53. Hesthagen, T., Sevaldrud, I.H., and Berger, H.M. (1999). *Ambio* **28**: 112–117.
54. Likens, G.E., Driscoll, C.T., and Buso, D.C. (1996). *Science* **272**: 244–246.
55. Olem, H. (1991). *Liming Acidic Surface Waters*. Levis Publ., Chelsea.
56. Henrikson, L. and Brodin, Y.-W. (1995). *Liming of Acidified Surface Waters*. (Eds.). Springer, Berlin.
57. Erikson, F., Hörnström, E., Mossberg, P., and Nyberg, P. (1983). *Hydrobiologia* **101**: 145–164.
58. Straskraba, M., Tundisi, J.G., and Duncan, A. (1993). *Comparative Reservoir Limnology and Water Quality Management. Developments in Hydrobiology*. Kluwer Academic Publishers, Dordrecht.
59. Tundisi, J.G. and Straskraba, M. (1999). *Theoretical Reservoir Ecology and its Applications*. Backhuys Publishers, Leiden.
60. Castro, J.M. and Moore, J.N. (2000). *Environmental Geology* **39**: 1254–1260.
61. Goldman, C.R. (1988). *Limnol. Oceanogr.* **33**: 1321–1333.
62. George, D.G. and Harris, G.P. (1985). *Nature* **319**: 536–539.
63. Ambrosetti, W. and Barbanti, L. (1999). *J. Limnol.* **58**: 1–9.
64. Anderson, N.J. (1993). *Trends Ecol. Evol.* **8**: 356–361.
65. Frey, D.G. (1969). *Mitt. Internat. Verein. Limnol.* **17**: 7–18.
66. Smol, J.P. (1992). *J. Aquat. Ecosystem Health* **1**: 49–58.
67. Horie, S. (1984). *Lake Biwa*. Monographiae Biologicae. W. Junk, Dordrecht.

ADSORPTION OF METAL IONS ON BED SEDIMENTS

Chakresh K. Jain
National Institute of Hydrology
Roorkee, India

The adsorption characteristics of cadmium on bed sediments of the River Ganga (India) have been studied to determine the tolerance of the river system for the heavy metal load. The effect of various controlling parameters, initial concentration, solution pH, sediment dose, contact time and particle size, have been evaluated. The optimum contact time needed to reach equilibrium was approximately 60 minutes and was independent of the initial concentration of cadmium ions. The adsorption curves are smooth and continuous leading to saturation, suggesting a possible monolayer coverage of cadmium ions on the surface of the adsorbent. The extent of adsorption increases with pH. Furthermore, the adsorption of cadmium increases as adsorbent doses increase and as adsorbent particle size decreases. The two geochemical phases, iron and manganese oxides, probably support the adsorption of cadmium ions. The adsorption data have been analyzed by using the Langmuir and Freundlich adsorption models to determine the mechanistic parameters of the adsorption process. Isotherms have been used to obtain thermodynamic parameters, free energy change (ΔG^o), enthalpy change (ΔH^o) and entropy change (ΔS^o). The negative values of free energy change (ΔG^o) indicate the spontaneous nature of the adsorption of cadmium on bed sediments, and the positive values of enthalpy change (ΔH^o) suggest the endothermic nature of the adsorption process. The uptake of cadmium is controlled by bulk as well as intraparticle diffusion mechanisms.

INTRODUCTION

Heavy metals added to a river system by natural and man-made sources during their transport are partitioned among different compartments of aquatic ecosystems, such as water, sediment, and biota (1). The added heavy metals then undergo many different chemical changes, whereby a high degree of variation in metal concentrations occurs (1,2). Metals that do not remain soluble in water are adsorbed and accumulate on bottom sediments that act as a sink (3). The distribution of heavy metals in the bottom sediments is affected by the mineralogical and chemical composition of the suspended material, anthropogenic influences, and *in situ* processes such as deposition, sorption, and enrichment in organisms (4). Thus, in the natural conditions of the river water, suspended and riverbed sediments play an important role in buffering higher metal concentrations in water, particularly by adsorption and/or precipitation. Therefore, the study of sediments and their sorptive properties can provide valuable information relating to the tolerance of the system to the added heavy metal load and may determine the fate and transport of pollutants in the aquatic environment.

The tremendous increase in the use of heavy metals during the past few decades has resulted in an increased concentration of metals in waterbodies and has great environmental significance due to their toxicity and nonbiodegradable nature (5). Cadmium, one of the most toxic metals, finds its way to watercourses through wastewater discharges from metal plating and cadmium–nickel battery industries, phosphate fertilizers, mining, pigments, stabilizers, and alloys (6). The toxic effects of cadmium include high blood pressure, kidney damage, and destruction of testicular tissue and red blood cells (7). Even low concentrations of cadmium cause hypertensive diseases. The International Agency for Research on Cancer classified cadmium as one of the chemicals that is carcinogenic to humans (8). Therefore, understanding the environmental conditions and mechanisms that regulate the mobilization and distribution of cadmium in the environment is essential.

Gardiner (9) studied the effect of various parameters on the adsorption of cadmium on river mud and other natural solids. Koelmans and Lijklema (10) studied the adsorption of cadmium onto sediment and suspended solids in Lake Volkerak in The Netherlands and reported that cadmium is bound almost completely to geochemical iron, manganese, and organic phases. Wiley and Nelson (11) examined the influence of various factors on the adsorption of cadmium onto the sediments of Sturgeon Lake in Oregon and reported that pH is the most critical parameter affecting cadmium adsorption. Christensen (12) and Palheiros et al. (13) also reported almost similar findings. Bajracharya et al. (14) studied the effect of zinc and ammonium ions on the adsorption of cadmium on sand and soil and reported that both ions suppress adsorption capacity significantly. Fu and Allen (15) studied the adsorption of cadmium by oxic sediments using a multisite-binding model. The model has been used satisfactorily to predict the extent of adsorption across the pH range 4.5–7.0.

Despite the apparent wealth of information on adsorption processes, little is known about quantitatively describing adsorption by coarser sediment. The important components of the suspended load for geochemical transport are silt, clay, hydrous iron and manganese oxides, and organic matter. Generally, adsorption studies for sediment less than 50 μm in size predominate (13,16); those for sediment more than 75 μm in size are lacking because sorption decreases as particle size increase. Although clay and silt adsorb metal ions much better than coarser fractions of sediment, one should take into account that most river sediments contain 90–95% sand and only 0–10% clay and silt. Therefore, in river systems that have a high sand percentage and low clay and silt content, the overall contribution of the sand to adsorption of metal ions could be comparable to or even higher than that of the clay and silt fraction. Our earlier efforts to establish this hypothesis have been quite successful (17–19).

Recently, we studied the adsorption characteristics of zinc on bed sediments of the River Ganga (20). In this article, we report the adsorption of cadmium on bed sediments of the River Ganga to assess the tolerance of the river system to the added cadmium load. The importance of geochemical phases has been investigated. Adsorption data have been analyzed by using adsorption models to determine the mechanistic parameters of the adsorption process, and isotherms have been used to obtain thermodynamic parameters.

THE RIVER SYSTEM

The River Ganga is a perennial river formed by the confluence of two smaller rivers at Deoprayag. Bhagirathi is one of them originating at Gaumukh in the Gangotri Glacier, 3129 m above mean sea level; the other is Alaknanda that originates in Sapta Tal Glacier. After covering a distance of about 220 km in the Himalayas, it enters the plains at Hardwar, and after meandering across a distance of about 2290 km in the plains of Uttar Pradesh, Bihar, and West Bengal, it joins the Bay of Bengal through a large number of branches flowing in India and Bangladesh (Fig.1).

Physiographically, the area is generally flat, except for the Siwalik Hills in the north and northeast of the catchment. The area is devoid of any relief features of provinces except from deep gorges cut by drains and rivers flowing through the area. The drainage pattern is dominated by the River Ganga, which is the only major river flowing through the area. The meltwater of the glaciers in the upper Himalayas maintains a perennial supply of water in the River Ganga.

The region is characterized by a moderate type of subtropical monsoonal climate. It has a cool, dry, winter season from October to March, a hot, dry, summer season from April to June, and a warm rainy season from July to September. The average annual rainfall across the Ganga Basin varies from 780 mm in the upper part to 1040 mm in the middle course and 1820 mm in the lower delta of Bangladesh (21). It has been observed that the rainfall is heaviest in the northern region of Hardwar district, close to the foothills of the Himalayas, and becomes less southward. The major land use is agriculture, and there is no effective forest cover.

Geologically, the area is a part of the west Indogangetic plain, which is composed mainly of Pleistocene and subrecent alluvium brought down by the river action from the Himalayan region. At Hardwar, the sequence of sandstone and shales along with gravel beds and clays are grouped within the upper and middle Siwaliks. The soils of this region do not form a compact block. They differ from valley to valley and slope to slope according to different ecological conditions. The bed of the river is rocky down to Hardwar. The soils of the area are predominantly loam to silty loam and are normally free from carbonates. The pollution status of the River Ganga has been described earlier (22,23).

EXPERIMENTAL METHODOLOGY

Freshly deposited sediments from shallow water near the bank of the River Ganga at Hardwar in the state of Uttaranchal were collected in polyethylene bags and brought to the laboratory. Samples were taken from the upper 5 cm of the sediments where flow rates were low and sedimentation was assumed (24,25).

The size distribution of the sediment samples was analyzed by using nylon sieves to obtain various fractions. The textural features of the sediments were observed, and a preliminary classification made according to grain size and distinctive geochemical features. The important geochemical phases for the adsorption process are organic matter, manganese oxides, iron oxides, and clays. The contents of manganese oxide and iron oxide were measured as total manganese and total iron, respectively, and were extracted from the sediment samples using acid

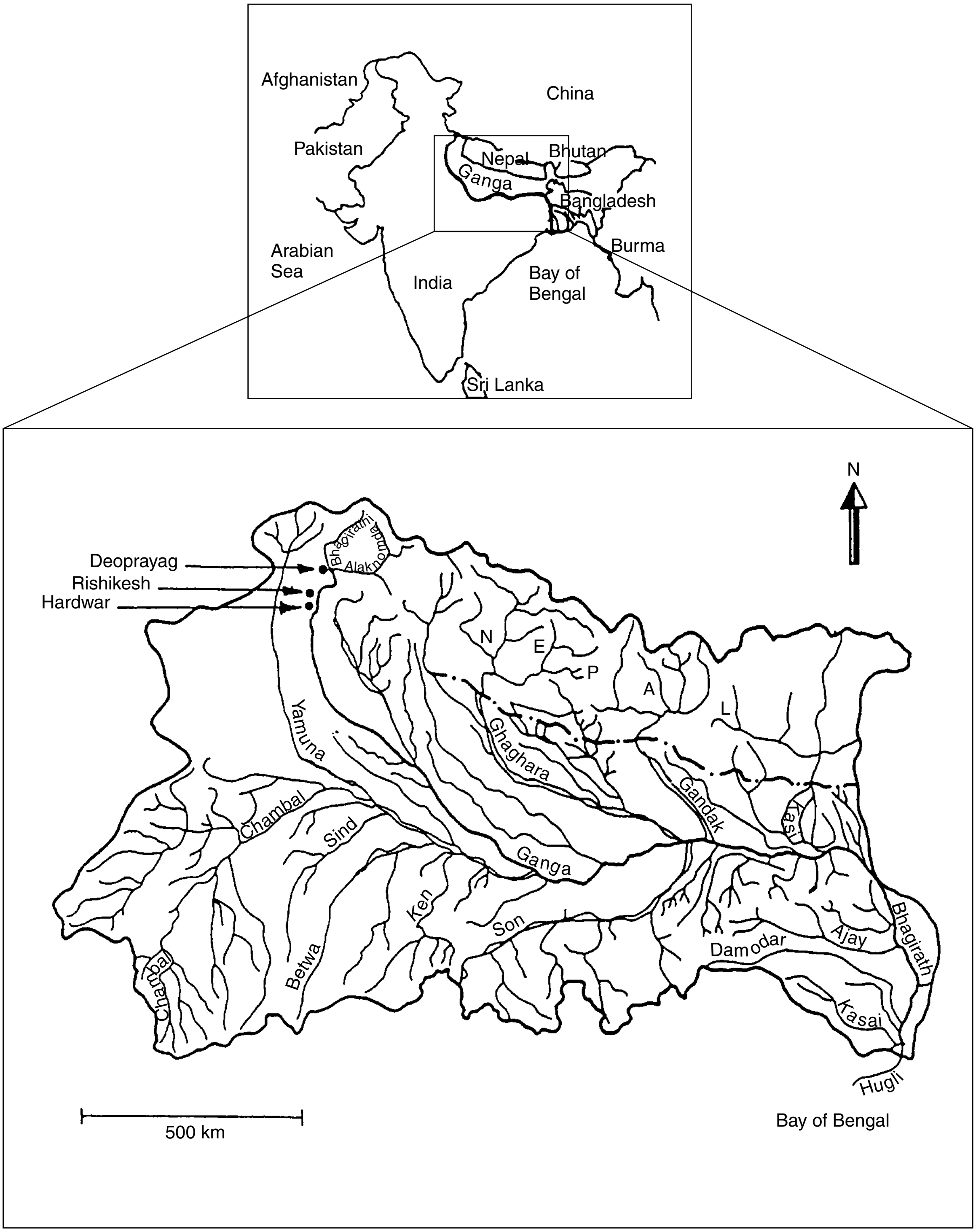

Figure 1. The Ganga Basin showing the locations of the sediment sampling sites.

digestion ($HF + HClO_3 + HNO_3$) in an open system. Organic matter was determined by oxidation with hydrogen peroxide.

All chemicals used in the study were obtained from Merck, India, and were of analytical grade. Aqueous solutions of cadmium were prepared from cadmium nitrate. Deionized water was used throughout the study. All glassware and other containers were thoroughly cleaned by soaking in detergent followed by soaking in 10% nitric acid for 48 h and finally rinsed with deionized water several times prior to use.

Adsorption experiments were conducted in a series of Erlenmeyer flasks of 100 mL capacity covered with Teflon sheet to prevent introduction of any foreign particles. Fifty mL of cadmium ion solution (200–2000 μg/L) was transferred in the flasks together with desired adsorbent doses (W_s in g/L), and placed in a water bath shaker maintained at 30 °C. A pH of 6.5 ± 0.1 was maintained throughout the experiment by using dilute HNO_3 and NaOH solutions. Aliquots were retrieved periodically and filtered through 0.45 μm cellulose nitrate membrane filters. The filters were soaked in 1% v/v HNO_3 for 1 h and thoroughly rinsed with deionized water prior to use.

The concentration of cadmium ions was determined by flame atomic absorption spectrometry, using a Perkin-Elmer Atomic Absorption Spectrometer (Model 3110) with air–acetylene flame. The detection limit for the cadmium ion was 0.0005 mg/L. Operational conditions were adjusted to yield optimal determinations. Quantification of metals was based on calibration curves of standard solutions of cadmium ion. These calibration curves were determined several times during the period of analysis.

RESULTS AND DISCUSSION

The content of important geochemical phases (iron and manganese) in different fractions of the sediment along with weight percentages are given in Table 1. The sediment has a coarse texture and is composed of more than 90% sediment of size >75 μm and <10% silt and clay. The organic content of the sediment was of the order of 0–1%. The background cadmium level in the various fractions of the sediments was negligible in the unpolluted zone, compared to the amount of adsorbate added for the adsorption tests. This confirms the absence of any cadmium particulate attached to the sediment particles.

It is evident from the data that the amount of manganese and iron in the various fractions of the sediment decreases as particle size increases. This indicates the possibility that the two geochemical phases act as active support material for the adsorption of cadmium ions. However, the relative contribution of the individual components could not be obtained from the present studies because the type and composition of a sediment's mineral and organic fractions vary simultaneously in natural systems, and the effect of individual constituents cannot be isolated. The content of iron in the sediment fractions is relatively higher and indicates the possible presence of iron minerals other than hydroxides. However, this result should be confirmed by further investigations. It is further evident from Table 1 that the sediment fraction of 150–210 μm particle size constitutes 68.6% of the total sediment load. Therefore, it was considered appropriate to study the adsorption of cadmium ions on this fraction (150–210 μm) and to compare it to the clay and silt fraction (<75 μm particle size) to demonstrate the importance of the coarser fraction in controlling metal pollution.

Table 1. Characteristics of Sediments

Sediment Fraction, μm	Weight, %	Total Mn, mg/g	Total Fe, mg/g
<75	9.8	1.570	68.2
75–150	13.5	1.420	62.0
150–210	68.6	1.199	58.0
210–250	6.6	0.844	55.1
250–300	0.7	0.830	52.1
300–425	0.6	0.823	52.0
425–600	0.2	0.824	52.0

OPERATING VARIABLES

Equilibrium Time (*t*)

To determine the equilibrium time for the adsorption process, adsorption experiments were performed for the uptake of cadmium ions for different contact times and for a fixed adsorbent dose of 0.5 g/L and an initial cadmium concentration of 1000 μg/L for the two particle sizes of adsorbent (0–75 and 150–210 μm) at pH 6.5 (Fig. 2). The solution pH for the experiments was close to that encountered in the river water. These plots indicate that the concentration of cadmium ions in solution becomes asymptotic to the time axis such that there is no appreciable change in the remaining concentration after 60 minutes in both fractions. This time is presumed to represent the time at which an equilibrium concentration is attained. All additional experiments were conducted for 60 minutes.

According to Weber and Morris (26), the uptake for most adsorption processes varies almost proportionately as $t^{1/2}$ rather than as contact time. Therefore, plots of

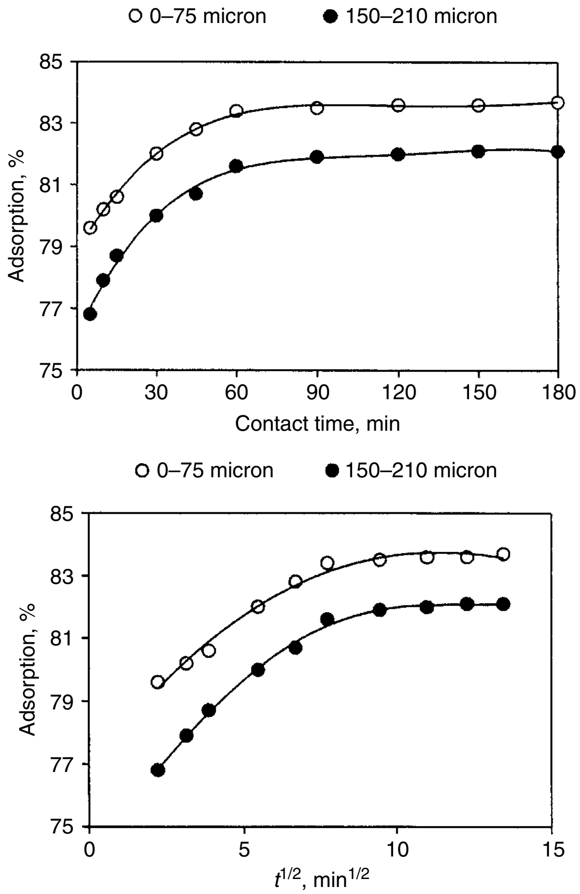

Figure 2. Effect of contact time (t and $t^{1/2}$) on percent adsorption of cadmium ions.

cadmium adsorbed, C_t versus $t^{1/2}$, are also presented for the two particle sizes of adsorbent in Fig. 2. It is clearly evident that adsorption of cadmium ions follow three phases: (1) an early extremely fast uptake (Phase I), (2) a transition phase (Phase II), and (3) an almost flat plateau section (Phase III). Phase I is attributed to the instantaneous use of the most readily available adsorbing sites on the adsorbent surface (bulk diffusion). Phase II, exhibiting additional removal, is attributed to the diffusion of the adsorbate from the surface film into the macropores of the adsorbent (pore diffusion or intraparticle diffusion), stimulating further migration of adsorbate from the liquid phase onto the adsorbent surface. Phase III, a plateau section, represents the equilibrium state.

Adsorption Isotherm

The adsorption isotherms for cadmium adsorption on the bed sediments, shown in Fig. 3, have a fixed adsorbent dose of 0.5 g/L at a pH of 6.5 ± 0.1. The adsorption data indicate linear distribution in the range 0 to 1000 μg/L. It is evident that for the same equilibration time, the amount of cadmium adsorbed is greater for greater initial concentrations of cadmium ions. In addition, the percentage adsorbed is greater for lower initial concentrations of cadmium and decreases as the initial concentration increases (Fig. 3). This is obvious because more efficient use of the adsorptive capacities of the adsorbent is expected from a greater driving force.

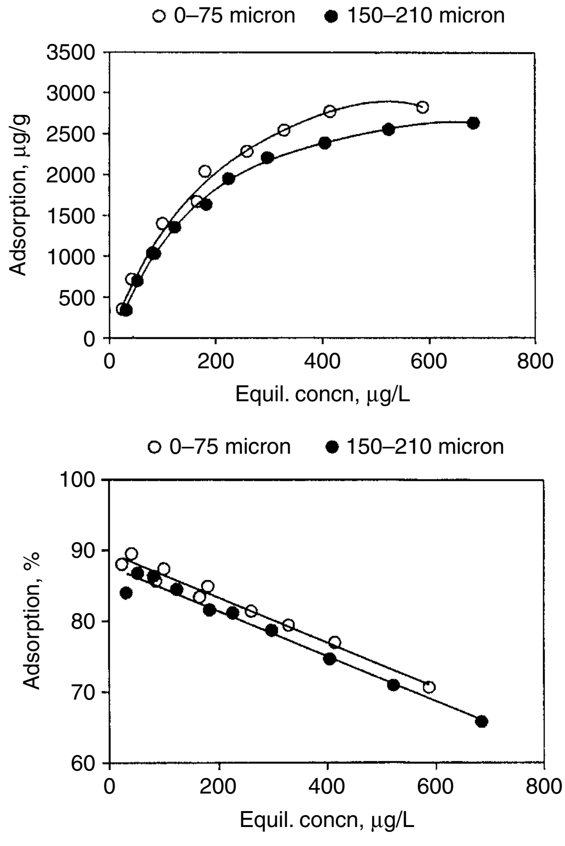

Figure 3. Adsorption of cadmium ions at different concentrations.

Comparing the two plots, it is clearly evident that the affinity of cadmium is greater for the <75-μm fraction, clay and silt, compared to the coarser fraction. It is also evident from Table 1 that the <75 μm fraction contains more iron and manganese than the 150–210 μm fraction, indicating the possibility of an association of these substrate (iron and manganese) with clay and silt particles. These findings illustrate the possible role of clay and silt as sites for cadmium adsorption. However, due to the paucity of data, correlation statistics could not be attempted to confirm this statement. The clay and silt constitute <10% of the total sediment load; comparing the weight percentages of the <75 μm to the 150–210 μm and their corresponding adsorption capacities for cadmium ions, it is clear that the coarser sediment contributes more than the clay and silt fraction in controlling cadmium pollution.

Effect of pH

The adsorption of cadmium on the riverbed sediments was studied across the pH range 2–7 for a fixed initial concentration of cadmium ($C_i = 1000$ μg/L) and an adsorbent dose of 0.5 g/L at particle sizes of 0–75 and 150–210 μm (Fig. 4). The pH of the solution was adjusted by using dilute hydrochloric acid and sodium hydroxide solutions. The pH was measured before and after the solution had been in contact with the sediment; the difference between the two values was generally less than 0.1 pH unit. A general increase in adsorption as the pH of solution increased was observed up to pH 6.0 for both fractions of the sediment. From the results, it is evident that the pH for maximum uptake of cadmium ion is 6.0. Further, it is apparent that the adsorption of cadmium rises from 6.5% at pH 2.0 to 83.3% at pH 6.0 for the clay and silt fraction (0–75 μm) and from 3.9% at pH 2.0 to 81.5% at pH 6.0 for the coarser sediment fraction (150–210 μm). A similar trend for pH was reported by Palheiros et al. (13) for the adsorption of cadmium on riverbed sediment.

Adsorbent Dose (W_s)

The effect of adsorbent dose on the adsorption properties of the bed sediments of the River Ganga was studied at

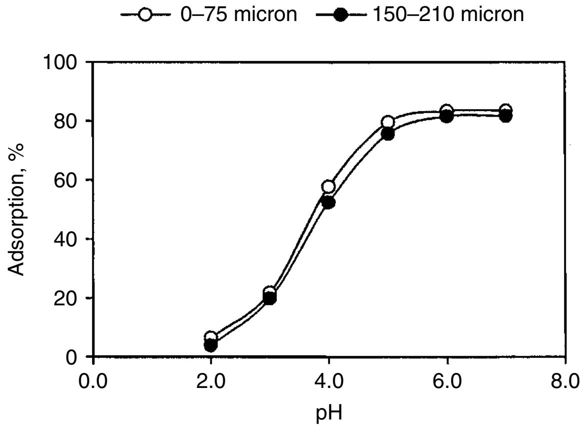

Figure 4. Effect of pH on adsorption of cadmium ions.

pH 6.5 using different adsorbent doses varying from 0.5 to 2.5 g/L and at a fixed initial cadmium concentration of 1000 μg/L (Fig. 5). The experiments were conducted in a water bath shaker to disperse the sediment particles in the aqueous medium. It was observed that for a fixed initial concentration of cadmium ($C_0 = 1000$ μg/L), the adsorption of cadmium per unit weight of adsorbent decreases as adsorbent load increases. On the other hand, the percent adsorption increases from 83.4 to 93.7% for the 0–75 μm fraction as the adsorbent load increases from 0.5 to 2.5 g/L. The adsorption of cadmium was higher for the 0–75 μm fraction compared to the 150–210 μm fraction. Because of the higher surface area as well as higher content of iron and manganese in the 0–75 μm fraction; these are the main driving forces for the adsorption of cadmium ions.

Particle Size (d_p)

The effect of adsorbent particle size on cadmium adsorption is shown in Fig. 5 for a fixed initial concentration of cadmium ($C_i = 1000$ μg/L), an adsorbent dose of 0.5 g/L, and at pH 6.5. These plots reveal that for a fixed adsorbent dose, smaller particles adsorb more cadmium. Further, it is observed that the percentage of cadmium adsorbed decreases from 83.4% on the 0–75 μm fraction to 77.4% on the 425–600 μm fraction. This result occurs because adsorption is a surface phenomenon; the smaller particle sizes offered a comparatively larger surface area and hence higher adsorption occurred at equilibrium. The higher content of iron and manganese in the 0–75 μm sediment fraction also accounts for greater adsorption of cadmium in this fraction compared to larger size fractions of the sediment.

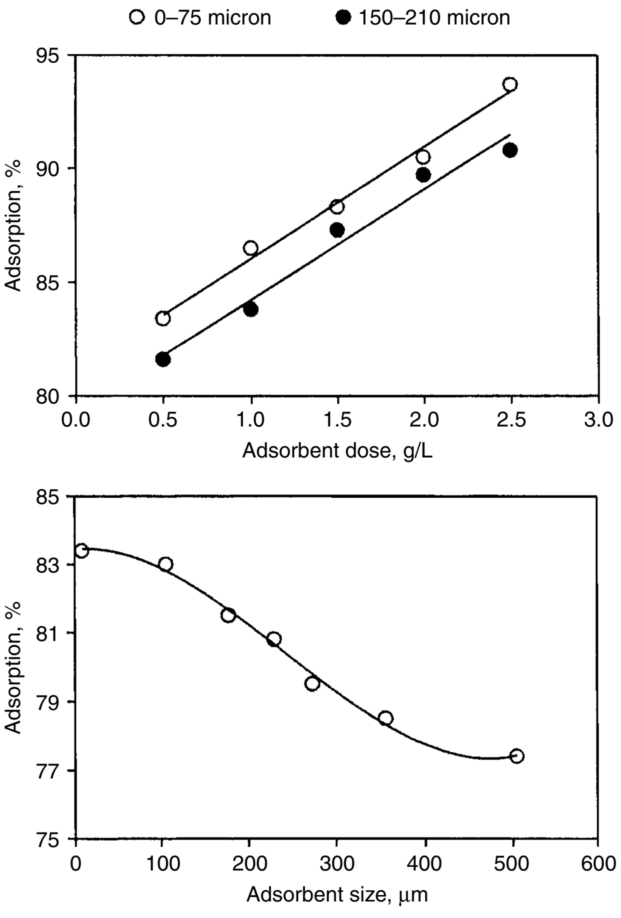

Figure 5. Effect of adsorbent dose and size on adsorption of cadmium ions.

ADSORPTION MODELS

Adsorption data for a wide range of adsorbate concentrations are most conveniently described by adsorption models, such as the Langmuir or Freundlich isotherm, which relate adsorption density q_e (metal uptake per unit weight of adsorbent) to equilibrium adsorbate concentration in the bulk fluid phase, C_e. The adsorption data for cadmium on the bed sediments of the River Ganga was analyzed using Langmuir and Freundlich models to evaluate the mechanistic parameters of the adsorption process.

Langmuir Model

Langmuir's isotherm model is valid for monolayer adsorption onto a surface containing a finite number of identical sites. The Langmuir treatment is based on the assumption that maximum adsorption corresponds to a saturated monolayer of solute molecules on the adsorbent surface, that the energy of adsorption is constant, and that there is no transmigration of adsorbate in the plane of the surface.

The linear form of Langmuir isotherm equation is represented by the following equation:

$$\frac{1}{q_e} = \frac{1}{Q^o} + \frac{1}{bQ^oC_e} \tag{1}$$

where q_e is the amount adsorbed at equilibrium time (μg/g), C_e is the equilibrium concentration of the adsorbate ions (μg/L), and Q^o and b are Langmuir constants related to maximum adsorptive capacity (monolayer capacity) and energy of adsorption, respectively. When $1/q_e$ is plotted against $1/C_e$, a straight line of slope $1/bQ^o$ and intercept $1/Q^o$ is obtained (Fig. 6), which shows that the adsorption of cadmium ions follows the Langmuir isotherm model. The Langmuir parameters, Q^o and b, are calculated from the slope and intercept of the graphs and are given in Table 2. These values may be used for comparison and correlation of the sorptive properties of the sediments.

Freundlich Model

The Freundlich equation, widely used for many years, is applicable to isothermal adsorption. This is a special case of heterogeneous surface energies in which the energy term, b, in the Langmuir equation varies as a function of surface coverage, q_e, strictly due to variations in the heat of adsorption (27). The Freundlich equation has the general form,

$$q_e = K_F C^{1/n} \tag{2}$$

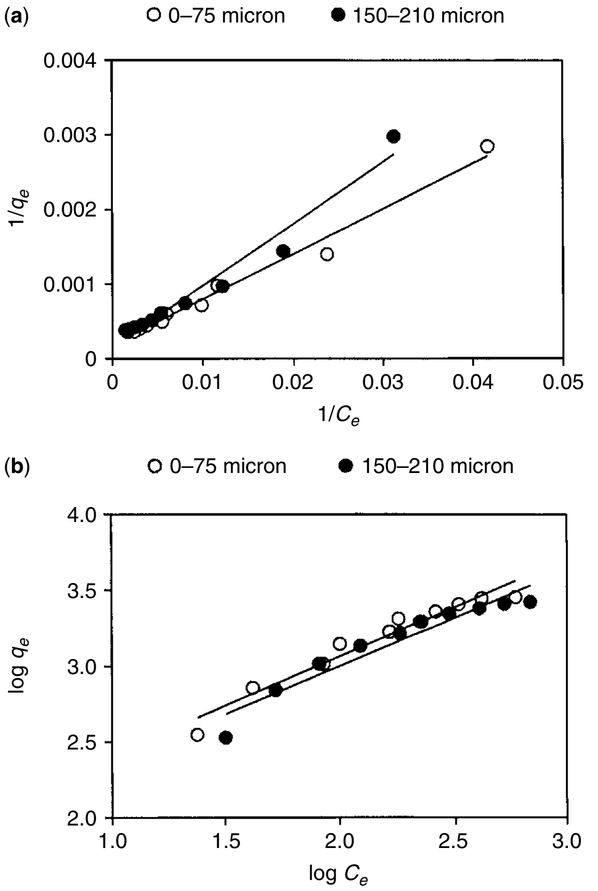

Figure 6. Graphical representation of adsorption isotherms: (**a**) Langmuir isotherm and (**b**) Freundlich isotherm.

The Freundlich equation is basically empirical but is often useful to describe data. Data are usually fitted to the logarithmic form of the equation,

$$\log q_e = \log K_F + \frac{1}{n}\log C_e \tag{3}$$

where q_e is the amount adsorbed (μg/g), C_e is the equilibrium concentration of the adsorbate ions (μg/L), and K_F and n are Freundlich constants related to adsorptive capacity and adsorptive intensity, respectively. When log q_e is plotted against log C_e, a straight line of slope $1/n$ and intercept log K_F is obtained (Fig. 6). This plot reflects satisfaction of the Freundlich isotherm model for the adsorption of cadmium ions on Ganga River sediment. The intercept of the line, log K_F, roughly indicates the adsorptive capacity, and the slope, $1/n$, indicates adsorptive intensity (28). The Freundlich parameters for the adsorption of two metal ions are given in Table 3.

Table 2. Langmuir Parameters for the Adsorption of Cadmium

Sediment Fraction, μm	Adsorption Maxima, Q^o, mg/g	Binding Energy Constant, b, mg/L
0–75	5.0	3.306
150–210	10.0	1.206

Table 3. Freundlich Parameters for the Adsorption of Cadmium

Sediment Fraction, μm	Adsorptive Capacity, K_F mg/g	Adsorptive Intensity, $1/n$
0–75	0.0045	1.765
150–210	0.0043	1.726

The Freundlich type adsorption isotherm is an indication of the surface heterogeneity of the adsorbent, and the Langmuir type isotherm corresponds to the surface homogeneity of the adsorbent. This leads to the conclusion that the surface of the bed sediments of the River Ganga is made up of small heterogeneous adsorptive patches that are similar to each other in respect to adsorptive phenomena (29).

THERMODYNAMICS

The effect of temperature on the adsorption of cadmium on bed sediments of the River Ganga is shown in Fig. 7 for a fixed adsorbent dose of 0.5 g/L and an initial cadmium concentration of 1000 μg/L at pH 6.5 for the two sediment fractions (0–75 μm and 150–210 μm). The temperature range used in the study was from 20 to 40 °C. It is evident from the plots that the adsorption of cadmium increases as temperature increases for both fractions, which may be mainly due to an increase in the number of active sites caused by the breaking of some bonds.

Thermodynamic parameters such as free energy change (ΔG^o), enthalpy change (ΔH^o), and entropy change (ΔS^o) were determined by using the following equations (30–32).

$$\Delta G^o = -RT\ln K_c \tag{4}$$

$$\Delta G^o = \Delta H^o - T\Delta S^o \tag{5}$$

where

ΔG^o = change in free energy, kJ/mol;
ΔH^o = change in enthalpy, kJ/mol;
ΔS^o = change in entropy, J/mol/K;
T = absolute temperature, K;
R = gas constant = 8.314×10^{-3};
K_c = equilibrium constant, defined as

$$K_c = \frac{C_{Ae}}{C_e} \tag{6}$$

where C_{Ae} and C_e are the equilibrium concentrations (μg/L) of the metal ions on the adsorbent and in the solution, respectively.

Equations 4 and 5 can be combined and rewritten as

$$\log K_c = \frac{\Delta S^o}{2.303R} - \frac{\Delta H^o}{2.303RT} \tag{7}$$

When log K_c is plotted against $1/T$, a straight line of slope $\frac{\Delta H^o}{2.303R}$ and intercept $\frac{\Delta S^o}{2.303R}$ is obtained (Fig. 8).

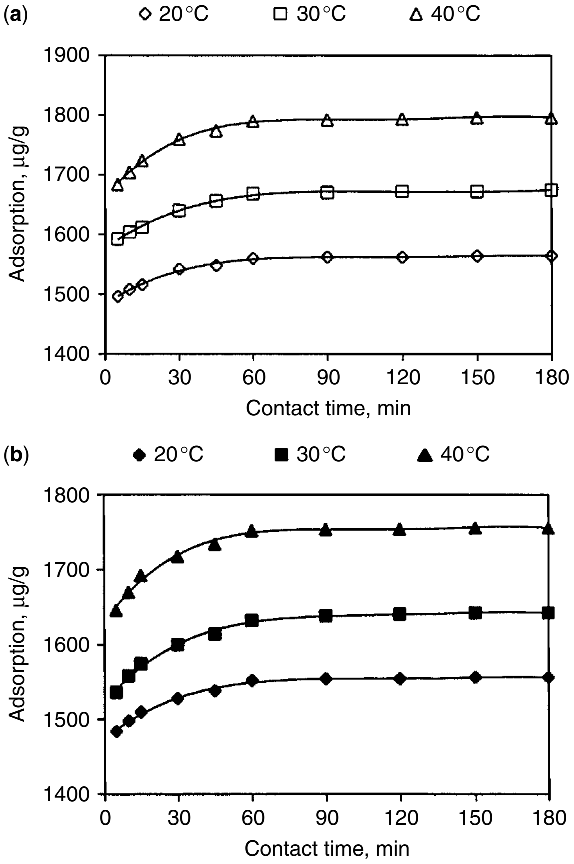

Figure 7. Effect of contact time and temperature on adsorption of cadmium ions: (**a**) sediment fraction = 0–75 µm; (**b**) sediment fraction = 150–210 µm.

The values of ΔH^o and ΔS^o were obtained from the slope and intercept of the van't Hoff plots of $\log K_c$ vs $1/T$ (Fig. 8). The thermodynamic parameters for the adsorption process are given in Table 4.

Positive values of ΔH^o suggest the endothermic nature of the adsorption. Negative values of ΔG^o indicate the spontaneous nature of the adsorption process. However, the negative value of ΔG^o decreased as temperature increased, indicating that the spontaneous nature of adsorption is inversely proportional to the temperature. The positive values of ΔS^o show increased randomness at the solid/solution interface during the adsorption process. The adsorbed water molecules, which are displaced by the adsorbate species, gain more translational energy than is lost by the adsorbate ions, thus allowing the prevalence of randomness in the system. The enhancement of adsorptive capacity at higher temperatures may be attributed to the enlargement of pore size and/or activation of the adsorbent surface (33).

Adsorption Dynamics

The rate constant of adsorption is determined by using the Lagergren first-order rate expression, (34–36):

$$\log(q_e - q) = \log q_e - \frac{k_{\mathrm{ad}}}{2.303} \times t \tag{8}$$

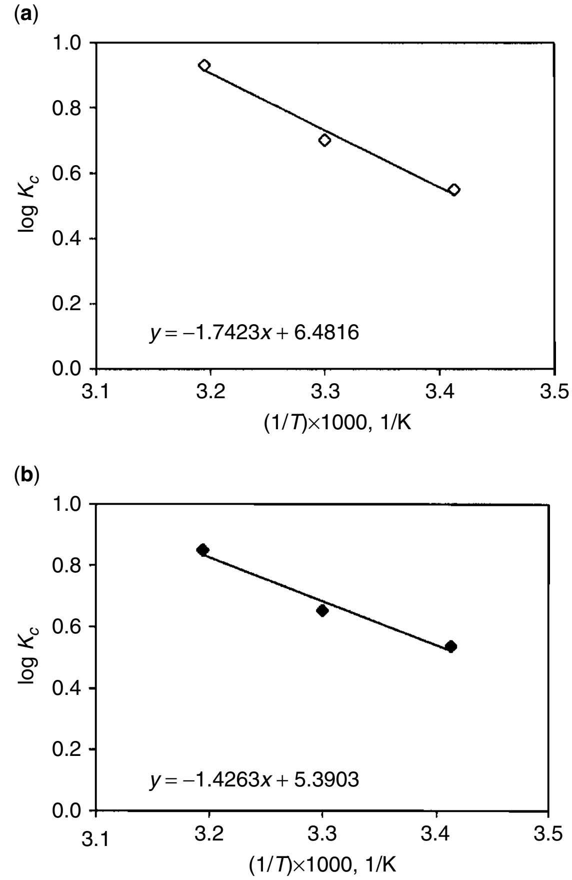

Figure 8. Van't Hoff plot for adsorption of cadmium ions: (**a**) sediment fraction = 0–75 µm; (**b**) Sediment fraction = 150–210 µm.

Table 4. Thermodynamic Parameters for the Adsorption of Cadmium

Sediment Fraction, µm	Temperature, °C	K_c	ΔG^o, kJ/mol	ΔH^o, kJ/mol	ΔS^o, J/mol/K
0–75	20	3.545	−3.083	33.36	124.1
—	30	5.024	−4.066	—	—
—	40	8.524	−5.576	—	—
150–210	20	3.444	−3.012	27.31	103.2
—	30	4.494	−3.786	—	—
—	40	7.064	−5.087	—	—

where q and q_e are the amounts of metal adsorbed (µg/g) at time t(min) and at equilibrium, respectively, and k_{ad} is the Lagergren rate constant for adsorption (1/ min). The straight line plots of $\log(q_e - q)$ versus t for different concentrations (Fig. 9) and temperatures (Fig. 10) indicate the applicability of this equation. Values of k_{ad} were calculated from the slope of the linear plots and are presented in Tables 5 and 6 for different concentrations and temperatures, respectively. The values of Lagergren rate constants (k_{ad}) were almost the same at both concentrations, whereas the value increases as temperature increases.

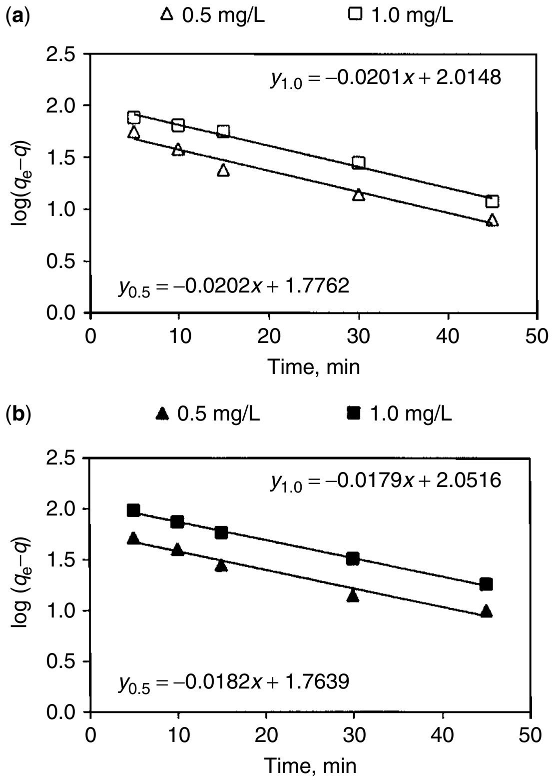

Figure 9. Lagergren plots at different initial concentrations: (**a**) sediment fraction = 0–75 μm; (**b**) sediment fraction = 150–210 μm.

Intraparticle Diffusion

The rate constant for intraparticle diffusion (k_{id}) is given by Weber and Morris (37):

$$q = k_{\mathrm{id}} t^{1/2} \tag{9}$$

where q is the amount adsorbed (μg/g) at time t (min). Plots of q versus $t^{1/2}$ are shown in Figs. 11 and 12 for different initial concentrations and temperatures, respectively. All plots have the same general features, an initial curved portion followed by a linear portion and a plateau. The initial curved portion is attributed to bulk diffusion, the linear portion to intraparticle diffusion, and the plateau to the equilibrium. This indicates that transport of cadmium ions from the solution through the particle solution interface, into the pores of the particle as well as the adsorption on the available surface of sediment, are both responsible for the uptake of cadmium ions. The deviation of the curves from the origin also indicates that intraparticle transport is not the only rate limiting step. The values of the rate constants (k_{id}) were obtained from the slope of the linear portion of the curves for each concentration of metal ions (Table 5) and temperature (Table 6). The values of intraparticle rate constants (k_{id}) increase as concentration and temperature increase.

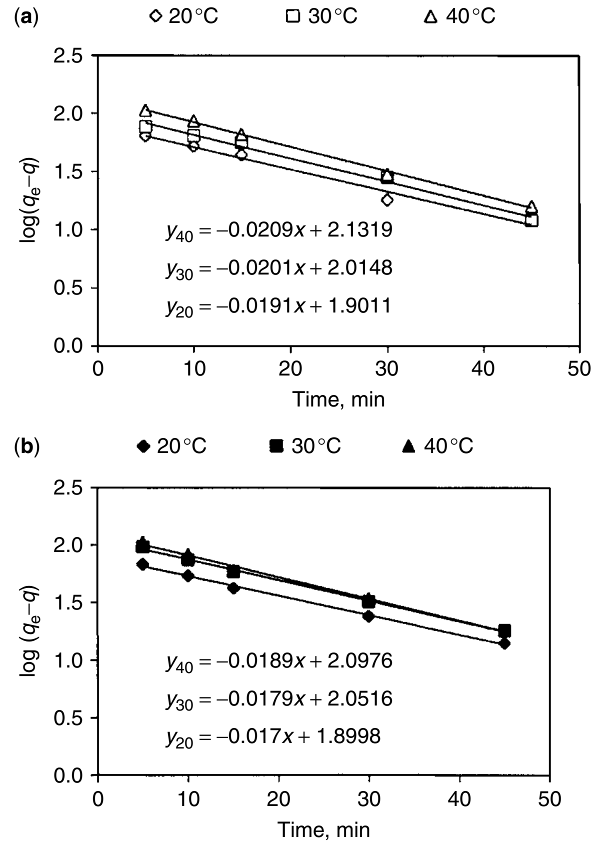

Figure 10. Lagergren plots at different temperatures: (**a**) sediment fraction = 0–75 μm; (**b**) sediment fraction = 150–210 μm.

Table 5. Rate Constants at Different Concentrations

Sediment Fraction, μm	Concentration, μg/L	Lagergren Rate Constant k_{ad} (1/ min)	Intraparticle Rate Constant k_{id}, μg/g/min
0–75	500	4.65×10^{-2}	3.84
—	1000	4.63×10^{-2}	9.88
150–210	500	4.19×10^{-2}	5.85
—	1000	4.12×10^{-2}	8.86

Table 6. Rate Constants at Different Temperatures

Sediment Fraction, μm	Temperature °C	Lagergren Rate Constant k_{ad} (1/ min)	Intraparticle Rate Constant k_{id}, μg/g/min
0–75	20	4.40×10^{-2}	8.11
–	30	4.63×10^{-2}	11.80
–	40	4.81×10^{-2}	14.29
150–210	20	3.92×10^{-2}	7.50
–	30	4.12×10^{-2}	10.00
–	40	4.35×10^{-2}	12.35

CONCLUSION

The study has demonstrated that river sediments provide good buffering capacity for metal loads. Although cadmium ions have more affinity for the clay and silt fraction of the sediment, the overall contribution of the coarser fraction to adsorption is greater compared to the clay and silt fraction. The pH of the solution is the most important parameter in the control of cadmium pollution.

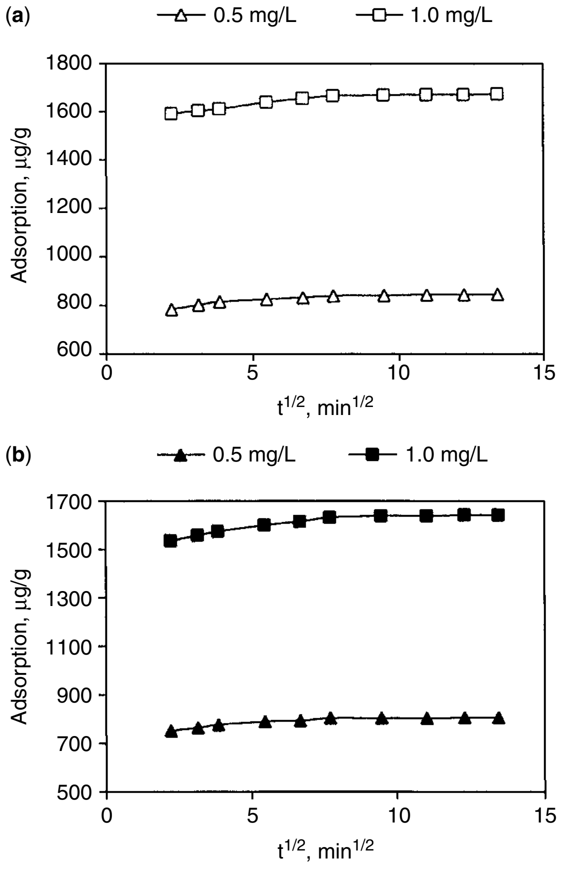

Figure 11. Intraparticle diffusion plots at different initial concentrations: (**a**) sediment fraction = 0–75 μm; (**b**) sediment fraction = 150–210 μm.

Iron and manganese oxides support the adsorption of cadmium. The adsorption data follow both the Langmuir and Freundlich isotherm models. The kinetic data suggest that the adsorption of cadmium is an endothermic process, which is spontaneous at low temperature. The uptake of cadmium is controlled by both bulk as well as intraparticle diffusion mechanisms. It is stated that in natural river systems, salinity and pH greatly affect the speciation of the metal ions and thus greatly interfere with their fixation on clay minerals. Adsorption/desorption equilibria and complexation with fulvic and humic acids also play an important role in speciation. Further studies are being planned to understand better the processes that affect the sorption of cadmium in natural systems.

BIBLIOGRAPHY

1. Moore, J.W. and Ramamoorthy, S. (1984). *Heavy Metals in Natural Waters, Applied Monitoring and Impact Assessment*. Springer-Verlag, New York.
2. Forstner, W. and Wittman, G.T.W. (1979). *Metal Pollution in the Aquatic Environment*. Springer-Verlag, New York.
3. Hakanson, L. (1980). An ecological risk index for aquatic pollution control-A sedimentological approach. *Water Res.* **14**: 975–1001.
4. Forstner, W. and Muller, G. (1975). *Factors controlling the distribution of minor and trace elements* (heavy metals V, Li, Sr) in recent lacustrine sediments, Resumes des Publications IX Congress Intl. De Sedimentologie, Nice, Theme II, p. 6.

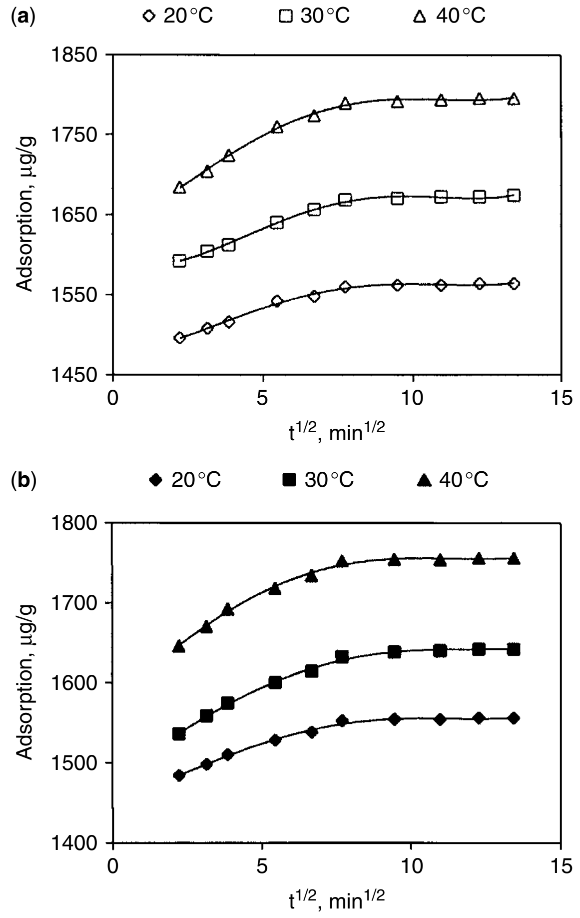

Figure 12. Intraparticle diffusion plots at different temperatures: (**a**) sediment fraction = 0–75 μm; (**b**) sediment fraction = 150–210 μm.

5. Zuane, J.D. (1990). *Hand Book of Drinking Water Quality*, Van Nostrand Reinhold, New York.
6. Low, K.S. and Lee, C.K. (1991). Cadmium uptake by the moss calympers delesertii, besch. *Bioresour. Technol.* **38**: 1–6.
7. (1980). *Encyclopedia of Environmental Sciences*, 2nd Edn. McGraw-Hill, New York, p. 354.
8. IARC (International Agency for Research on Cancer). (1976). Mongraphs on the evaluation of carcinogenic risk of chemicals to man, In: *Cadmium and Cadmium Compounds*, 11, p. 3974.
9. Gardiner, J. (1974). The chemistry of cadmium in natural water—II. Adsorption of cadmium on river muds and naturally occurring solids. *Water Res.* **8**: 157–164.
10. Koelmans, A.A. and Lijklema, L. (1992). Sorption of tetrachlorobenzene and cadmium to sediments and suspended solids in lake Volkerak/Zoom. *Water Res.* **26**: 327–337.
11. Wiley, J.O. and Nelson, P.O. (1984). Cadmium adsorption by aerobic lake sediments. *J. Environ. Eng., ASCE*, **110**(1): 226–243.
12. Christensen, T.H. (1984). Cadmium soil sorption at low concentrations: I. Effect of time, cadmium load, pH, and calcium, *Water, Air, and Soil Poll.* **21**: 105–114.
13. Palheiros, B., Duarte, A.C., Oliveira, J.P. and Hall, A. (1989). The influence of pH, ionic strength and chloride concentration

of the adsorption of cadmium by a sediment. *Water Sci. Tech.* **21**: 1873–1876.

14. Bajracharya, K., Barry, D.A., Vigneswaran, S., and Gupta, A.D. (1996). Heavy metal adsorption in soil: Comparison of bisolute adsorption models and laboratory experiments. In: *Water Quality Hydrology*. V.P. Singh and B. Kumar (Eds.). Vol. 3, Kluwer Academic Publishers, The Netherlands, pp. 19–26.
15. Fu, G. and Allen, H.E. (1992). Cadmium adsorption by oxic sediments. *Wat. Res.* **26**: 225–233.
16. Gagnon, C., Arnac, M., and Brindle, J.R. (1992). Sorption interactions between trace metals (Cd and Ni) and phenolic substances on suspended clay minerals. *Water Res.* **26**: 1067–1072.
17. Jain, C.K. and Ram, D. (1997). Adsorption of lead and zinc on bed sediments of the river Kali. *Water Res.* **31**(1): 154–162.
18. Jain, C.K. and Ram, D. (1997). Adsorption of metal ions on bed sediments. *Hydrological Sciences Journal* **42**(5): 713–723.
19. Jain, C.K. and Ali, I. (2000). Adsorption of cadmium on riverine sediments: quantitative treatment of the large particles. *Hydrological Processes* **14**: 261–270.
20. Jain, C.K. (2001). Adsorption of zinc onto bed sediments of the river Ganga: Adsorption models and kinetics. *Hydrological Sciences Journal* **46**(3): 419–434.
21. Krishna Murti, C.R. (1991). Some physical features of the Ganga basin. In: *The Ganga: A Scientific Study*. Krishna Murti, C.R., Bilgrami, K.S., Das, T.M., and Mathur, R.P. (Eds.). The Ganga Project Directorate, New Delhi, India.
22. Jain, C.K. (1998). *Adsorption of Zinc on Bed Sediments of River Ganga*, Technical Report, CS(AR)9/98-99, National Institute of Hydrology, Roorkee, India.
23. Jain, C.K. (2002). A hydro-chemical study of a mountainous watershed: the Ganga, India. *Water Res.* **36**(5), 1262–1274.
24. Sakai, H., Kojima, Y., and Saito, K. (1986). Distribution of metals in water and sieved sediments in the Toyohira river. *Water Res.* **20**: 559–567.
25. Subramanian, V., Grieken, R.V., and Vant, D.L. (1987). Heavy metal distribution in the sediments of Ganges and Brahmaputra rivers. *Environ. Geol. Water Sci.* **9**(2): 93–103.
26. Weber, W.J. Jr. and Moris, J.C. (1963). J. Sanit. Eng. Div., ASCE, SA2, 31, cited in Faust S.D. and Aly O.M., 1987. *Adsorption Processes for Water Treatment*, 65–122, Butterworth, London, pp. 65–122.
27. Adamson, A.W. (1967). *Physical Chemistry of Surfaces*, 2nd Edn. Interscience Publishers, Inc., New York.
28. Weber, W.J. Jr. (1972). *Physico-chemical Processes for Water Quality Control*. John Wiley & Sons, New York, NY, pp. 208–210.
29. Ajmal, M., Khan, A.H., Ahmad, S., and Ahmad, A. (1998). Role of sawdust in the removal of copper (II) from industrial wastes. *Water Res.* **32**(10): 3085–3091.
30. Singh, A.K., Singh, D.P., Pandey, K.K., and Singh, V.N. (1988). Wollastonote as adsorbent for removal of Fe(II) from water. *J. Chem. Technol. Biotechnol.* **42**: 39–49.
31. Catena, G.C. and Bright, F.V. (1989). Thermodynamic study on the effects of β-cyclodextrin inclusion with anilinonaphthalene sulfonates. *Anal. Chem.* **61**: 905–909.
32. Frajii, L.K., Hayer, D.M., and Werner, T.C. (1992). Static and dynamic fluorescence quenching experiments for the physical chemistry laboratory. *J. Chem. Edu.* **69**: 205–215.
33. Vishwakarma, P.P., Yadava, K.P., and Singh, V.N. (1989). Nickel(II) removal from aqueous solutions by adsorption on fly ash. *Pertanika* **12**: 357–366.
34. Singh, D.B., Prasad, G., Rupainwar, D.C., and Singh, V.N. (1989). As(III) removal from aqueous solution by adsorption. *Water Air and Soil Pollution* **42**: 373–386.
35. Namasivayam, C. and Ranganathan, K. (1995). Removal of Cd(II) from waste water by adsorption on waste Fe(III)/Cr(III) hydroxide. *Water Res.* **29**: 1737–1744.
36. Zhang, L., Zhao, L., Yu, Y., and Chen, C. (1988). Removal of lead from aqueous solution by non-living rhizopus nigricans. *Water Res.* **32**(5): 1437–1444.
37. Weber, W.J. Jr. and Moris, J.C. (1962). *Advances in Water Pollution Research*. Proc. 1st International Conference on Water Pollution Research. Vol. 2, Pergamon Press, Oxford, UK, p. 231.

MICROBIOLOGY OF LOTIC AGGREGATES AND BIOFILMS

Eric A. Strauss
Upper Midwest Environmental
Sciences Center
La Crosse, Wisconsin

INTRODUCTION

The microbial community is a vital component of lotic (flowing water) ecosystems and is important for nutrient cycling, organic matter decomposition, downstream transport of materials, and transfer of energy to higher trophic levels. Some of this microbial community is transported downstream as unattached cells; however, much of it is associated with suspended organic particles (i.e., aggregates) or attached to available substrate in biofilms. Research on microbial activity within organic aggregates has mostly been limited to marine and lake systems. As a result, more is known about the microbial communities and microbial processes in marine and lake aggregates compared with aggregates in lotic ecosystems. Conversely, biofilms have been studied in great detail in many systems, including lotic ecosystems. Many excellent and detailed reviews are available on the topics of aggregates and biofilms (1–11). Included here is a general overview of the microbiology and ecology of these topics and how they are relevant to lotic ecosystems.

LOTIC AGGREGATES

Origins and Terminology

Microorganisms readily colonize substrates in lotic ecosystems to form diverse and active communities. Suspended organic matter is often rich in nutrients and easily metabolized carbon, providing an especially good substrate for microbial colonization and growth and for the subsequent formation of lotic aggregates. Lotic aggregates can be defined as the interactive association between a microbial community and suspended organic matter in a flowing water system. The organic matter that is a central component of lotic aggregates may have a variety of origins including allochthonous and autochthonous particles and byproducts of animal and microbial activity. Allochthonous organic matter (e.g., wood, leaves, and terrestrial debris) readily enters lotic systems through various pathways. Larger particles tend to quickly settle onto the channel bed or get trapped in channel structures, whereas the smallest particles often become entrained in the flowing water. Over time, physical and biological processes fragment the large particles into smaller detrital particles, which can then become resuspended into the water column. Organic matter particles of autochthonous origin (e.g., biofilm slough and aquatic vegetation) also may become fragmented and entrained in the current in the same manner as those particles of allochthonous origin. Organic byproducts of animal activity such as fecal pellets and discarded invertebrate exoskeletons are common in highly productive systems and can be nutrient-rich substrates for microbial colonization (12). Aggregates also may form in conjunction with dissolved organic matter (13) or colloidal organic polymers that are small (1.5 nm to 0.45 μm), high-molecular-weight byproducts of organic matter decomposition that remain suspended in the water (14).

Aggregates have the potential to enlarge as they travel downstream by combining with other aggregates (aggregation) and to gather additional materials primarily through impaction (Fig. 1) (15). However, the physical forces exerted by flowing and turbulent water usually fragment larger aggregates (disaggregation), maintaining the aggregates at a smaller size compared with those found in nonflowing systems (16). Although relatively small, the size of lotic aggregates are still highly variable and can be further classified into microaggregates (<5 to 500 μm) and macroaggregates (>500 μm) (7). Larger aggregates are also sometimes called "river snow" in keeping with terminology used to describe similar aggregates present in marine (marine snow) and lake (lake snow) ecosystems (17,18). In the field of stream ecology, lotic aggregates may also be called seston, which can be defined as particles entrained in flowing water. A list of common terms and abbreviations used to describe suspended matter in lotic systems is provided in Table 1.

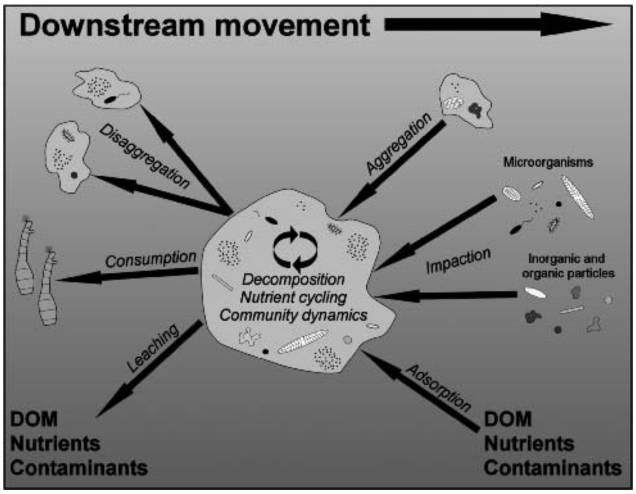

Figure 1. Conceptual diagram illustrating processes affecting lotic aggregate size and composition. DOM = dissolved organic matter.

Table 1. Common Terms, Abbreviations (in parentheses), and Sizes Used to Classify Suspended Matter in Flowing Water Systems

Term	Size
Dissolved organic matter (DOM)	<0.45 μm
Particulate organic matter (POM)	>0.45 μm
Course particulate organic matter (CPOM)	>1 mm
Fine particulate organic matter (FPOM)	0.45 μm to 1 mm
Aggregate	>0.45 μm
Macroaggregate	>500 μm
Microaggregate	<500 μm
Colloid	1.5 nm to 0.45 μm
Floc	>0.45 μm
Seston	>0.45 μm

The size ranges listed here are generally accepted definitions, but they can vary somewhat depending on application.

Microbial Community

Bacteria and other microorganisms secrete a viscous, slimy mixture of exopolymers (i.e., chains of polysaccharides) called glycocalyx. In general, the properties of exopolymers can contribute many beneficial functions to organisms including assisting in locomotion and adhesion to surfaces, and protection from abrasion, dehydration, and nutrient loss. More specifically, in lotic aggregates, the exopolymers bind the contents of aggregates together and assist in the capture of additional nutrients, organic matter, and other microorganisms. Much of the exopolymers associated with lotic aggregates are likely produced by microorganisms within the aggregate, but they also may accumulate via impaction with suspended exopolymers.

Microbial communities thriving within the fragile matrix of exopolymers and particles consist of a diverse heterogeneous assemblage of bacteria, algae, fungi, protozoa, and sometimes zooplankton. Compared with marine and lake ecosystems, relatively few studies have examined the microbial community structure of aggregates in lotic systems. However, many studies on marine aggregates have demonstrated that the concentration of algae, zooplankton, protozoa, and bacteria within aggregates is much greater, often 2–3 orders of magnitude, than in the surrounding water (7). Larger aggregates often support larger microbial communities, but smaller particles may be relatively more densely populated per unit surface area (19). Detailed microscopic analyses have shown that bacteria are not uniformly distributed within or on the surface of aggregates but form microcolonies (7). These microbial communities have temporally dynamic compositions that change quickly in response to conditions both inside and outside of the aggregate and can also exhibit seasonal patterns (17,20).

Inorganic (nitrogen, phosphorus, and silica) and organic (carbohydrates, amino acids, and dissolved organic carbon) nutrients are often 1–2 orders of magnitude more enriched in aggregates compared with surrounding water (7). In addition, many heavy metal and organic contaminants also readily adsorb to the aggregate matrix. As a result of these enriched conditions, downstream movement of aggregates is an important mechanism of transport for nutrients, organic materials, and contaminants. Moreover, these nutrient-rich substrates result in high growth efficiencies of heterotrophic bacteria associated with the lotic aggregates, which suggests that these aggregates are hot spots for microbial processes and the transfer of organic matter into bacterial biomass (20). Enzymes produced by aggregate-bound bacteria solubilize organic matter 1–2 orders of magnitude faster than the bacterial cells can assimilate the dissolved matter (21). This decomposition results in aggregate mass loss and leaching of dissolved organic matter back into the flowing water. Finally, consumption of lotic aggregates by many species of detritus- and filter-feeding organisms is a vital link of energy transfer between microbial and higher trophic levels (Fig. 1).

BIOFILMS

Biofilm Structure

Microbial biofilms are the thin slime layers that are observed on surfaces and substrates (e.g., rocks, sediment, vegetation, leaves, wood) in aquatic systems. Biofilms (sometimes called microbial mats) are similar to aggregates and possess many of the same general structural and functional attributes. For example, biofilms (like aggregates) contain microbial communities that are vital to lotic ecosystem function and are responsible for many key production, decomposition, and nutrient cycling processes (22). Highly developed biofilms are generally a few millimeters thick and can develop on virtually any available surfaces in natural lotic systems, provided the system has sufficient nutrients. Biofilm development is especially rapid on surfaces that contain or leach nutrients (e.g., particulate organic matter, POM) (2). Available surfaces are still colonized in nutrient-poor or oligotrophic systems, but biofilms are generally less developed. Under oligotrophic conditions, biofilm bacteria are often able to obtain limited nutrients by concentrating organics on the exopolymer matrix, using waste products of nearby microorganisms, and by pooling their biochemical resources (i.e., enzymes) to break down organic particles (11).

Historically, biofilms were thought to be a relatively homogeneous distribution of bacteria in an exopolymer matrix and nutrient and material delivery throughout the biofilm would be dominated by diffusion (2). However, recent advances in the analytical tools used to study biofilms have allowed scientists to more accurately characterize the structure or architecture of biofilms. For example, the application of confocal scanning laser microscopy has allowed nondestructive inspection of intact biofilm samples. This technique has revealed the physical structure of established biofilms to consist of a diverse microbial assemblage arranged in microcolonies and interspersed within a heterogeneous exopolymer matrix containing protrusions and a network of pores and channels (2,4). Microsensors, another relatively recent research tool, have been used to examine chemical properties in spatially explicit areas within biofilms and have found distinct differences in dissolved oxygen concentrations between the pores and channels in the matrix and the matrix itself. These differences suggest that such structures may facilitate water movement and nutrient delivery within the biofilm. Microsensors have also been used to quantify vertical chemical gradients in biofilms that have important implications for microbial community structure and biogeochemical processing (Fig. 2). Chemical gradients will be discussed further in the next section.

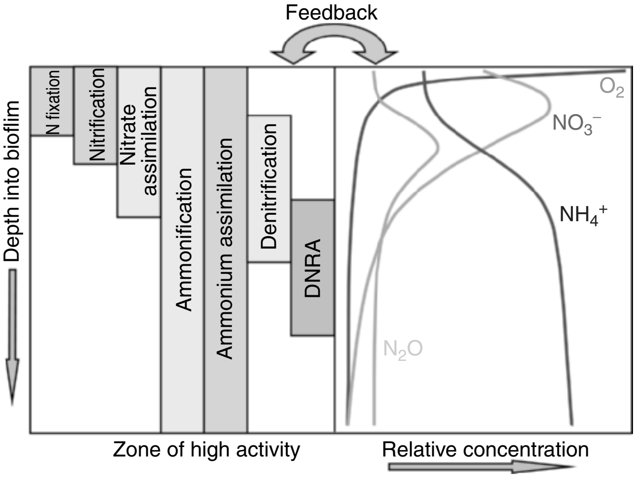

Figure 2. General patterns of nitrogen cycling (left) and relative concentrations of nitrogen constituents (right) in a lotic biofilm [redrawn and modified from Paerl and Pinckney (3)]. Feedback between the chemical and biological components facilitates the formation and equilibrium of the gradient (i.e., the chemical environment within the biofilm affects microbial processes and vice versa). DNRA = dissimilatory nitrate reduction to ammonium.

The physical stress caused by the turbulent flow of water in lotic systems can have significant effects on biofilm structure and organization. Many studies have observed that biofilms subjected to fast flowing water tend to be thinner, smoother, denser, more stable, and have lower surface sinuosity than biofilms in slower or nonflowing water (23,24). Cells occupying the surface of lotic biofilms have been observed to be arranged in formations that can withstand the higher shear stress; examples of these formations include hexagon-shaped structures and ripples positioned parallel to flow direction (24). Freely oscillating streamer-like structures are also sometimes observed adhered to the surface of lotic biofilms (8). Streamers are predominately single species colonies that elongate over time in the absence of disturbance.

Microbial Processes

The rates and types of microbial processes at work in any particular biofilm are largely regulated by the physical and chemical characteristics of the lotic environment. Several types of algae and cyanobacteria flourish in biofilms if sufficient nutrients are available and the overlying water is shallow and clear enough to allow adequate light penetration to the substrate. Oxygen (O_2) generated by these primary producers can reach very high levels and is a key factor supporting aerobic biofilm processes. In fact, under high-light conditions, the surficial area of well-lit biofilms can produce O_2 concentrations 2–3 times greater than atmospheric concentrations (25). Much, if not all, of the oxygen is respired as it is transported and diffused deeper into the biofilm creating a vertical gradient of oxygen concentrations throughout the biofilm (Fig. 2). Chemical gradients are common in biofilms and control which microbial processes are able to occur at certain depths. The opposite is also true; microbial processes control the availability of certain ions. The feedback relationship between these two factors quickly forms a relatively stable gradient equilibrium that is particularly important for the transformations and cycling of certain nutrients, primarily nitrogen, carbon, and sulfur.

The nitrogen cycle is a complex group of microbial processes that transform, oxidize, or reduce nitrogen ions among the various oxidation states. These processes occur under different environmental conditions, all of which can be located in distinct locations along the chemical gradients found in biofilms (Fig. 2). Nitrogen fixation, which is the conversion of atmospheric dinitrogen gas (N_2) to biologically available ammonium (NH_4^+), is predominately done by cyanobacteria; thus, it occurs in the upper, high-light areas of the biofilm. Nitrification, which is the oxidation of NH_4^+ to nitrate (NO_3^-), also occurs in the upper areas because it is an aerobic process. Bacteria that have the enzymes to reduce NO_3^- for assimilation likely occur throughout the biofilm profile, but they can only use NO_3^- where it is present. Organic matter is ubiquitous throughout the biofilm; therefore, ammonification (the release of NH_4^+ from organic matter) and the resulting NH_4^+ are also widespread. Ammonium assimilation into microbial biomass is not oxygen-dependent and can take place wherever NH_4^+ is present. Denitrification, which is the anaerobic reduction of NO_3^- to nitrite (NO_2^-), nitrous oxide (N_2O), and N_2, occurs deeper in the biofilm where oxygen is low and NO_3^- is high. Dissimilatory NO_3^- reduction to NH_4^+ is strictly an anaerobic process that occurs in the lowest areas of the biofilm where NO_3^- is still available.

It is generally accepted that bacterial activity and production of biofilm biomass is resource limited in lotic systems. Thus, nutrient availability and the quantity and quality of the dissolved organic carbon in the water have strong positive effects on bacterial activity in the biofilm (26,27). Total bacterial activity also often correlates well with primary production in biofilms because of the influx of oxygen and high-quality organic carbon provided by algae (27,28). Finally, vertical hydrodynamics of stream channels (i.e., upwelling and downwelling) can be an important means of nutrient and dissolved organic matter (DOM) delivery through the biofilm, which can subsequently increase microbial activity (29).

Under certain chemical or physical conditions, specialized biofilms that are dominated by unique groups of microorganisms may develop. For example, orange-colored iron-oxidizing biofilms are prevalent in areas where groundwater with high concentrations of reduced iron (Fe^{2+}) is discharged into oxygenated waters (30). Another example is the colorful biofilms of thermophilic bacteria and cyanobacteria that are commonly associated with water flowing from hot springs (Fig. 3). The colors of these biofilms depend primarily on the chlorophyll content in the cells, but also on the ratio of chlorophyll to carotenoids (yellow to red pigments) (31).

BIBLIOGRAPHY

1. Johnson, B.D., Kranck, K., and Muschenheim, D.K. (1990). Physico-chemical factors in particle aggregation. In: *The Biology of Particles in Aquatic Systems*. R.S. Wotton (Ed.). CRC Press, Boca Raton, FL, pp. 57–81.
2. Costerton, J.W. et al. (1995). Microbial biofilms. *Ann. Rev. Microbiol.* **49**: 711–745.
3. Paerl, H.W. and Pinckney, J.L. (1996). A mini-review of microbial consortia: their roles in aquatic production and biogeochemical cycling. *Microbial Ecol.* **31**: 225–247.

Figure 3. Orange-colored biofilm associated with hot spring outflow draining into the Firehole River, Yellowstone National Park, Wyoming. The bright orange color is because of pigments in thermophilic cyanobacteria and bacteria in the biofilm.

4. Davey, M.E. and O'Toole, G.A. (2000). Microbial biofilms: From ecology to molecular genetics. *Microbiol. Molecular Biol. Rev.* **64**: 847–867.
5. Kiørboe, T. (2001). Formation and fate of marine snow: Small-scale processes with large-scale implications. *Scientia Marina* **65**: 57–71.
6. Paerl, H.W. and Kuparinen, J. (2002). Microbial aggregates and consortia. In: *Encyclopedia of Environmental Microbiology*. G. Bitton (Ed.). John Wiley & Sons, New York, pp. 160–181.
7. Simon, M., Grossart, H-P., Schweitzer, B., and Ploug, H. (2002). Microbial ecology of organic aggregates in aquatic ecosystems. *Aquat. Microbial Ecol.* **28**: 175–211.
8. Stoodley, P., Sauer, K., Davies, D.G., and Costerton, J.W. (2002). Biofilms as complex differentiated communities. *Ann. Rev. Microbiol.* **56**: 187–209.
9. Turley, C.M. (2002). The importance of 'marine snow'. *Microbiol. Today* **29**: 177–179.
10. Zimmermann-Timm, H. (2002). Characteristics, dynamics and importance of aggregates in rivers—an invited review. *Int. Rev. Hydrobiol.* **87**: 197–240.
11. Prakash, B., Veeregowda, B.M., and Krishnappa, G. (2003). Biofilms: A survival strategy of bacteria. *Curr. Sci.* **85**: 1299–1307.
12. Wotton, R.S. and Malmqvist, B. (2001). Feces in aquatic ecosystems. *BioScience* **51**: 537–544.
13. Kerner, M. et al. (2003). Self-organization of dissolved organic matter to micelle-like microparticles in river water. *Nature* **422**: 150–154.
14. Thurman, E.M. (1985). *Organic Geochemistry of Natural Waters*. Martinus Nijhoff/Dr. W. Junk Publishers, Dordrecht, The Netherlands, p. 497.
15. Wotton, R.S. (2004). The ubiquity and many roles of exopolymers (EPS) in aquatic systems. *Scientia Marina* **68**: 13–21.
16. Wotton, R.S. (1996). Colloids, bubbles, and aggregates—a perspective on their role in suspension feeding. *J. N. Amer. Bentholog. Soc.* **15**: 127–135.
17. Böckelmann, U., Manz, W., Neu, T.R., and Szewzyk, U. (2000). Characterization of the microbial community of lotic organic aggregates ('river snow') in the Elbe River of Germany by cultivation and molecular methods. *FEMS Microbiol. Ecol.* **33**: 157–170.
18. Neu, T.R. (2000). *In situ* cell and glycoconjugate distribution in river snow studied by confocal laser scanning microscopy. *Aquat. Microbial Ecol.* **21**: 85–95.
19. Berger, B., Hoch, B., Kavka, G., and Herndl, G.J. (1996). Bacterial colonization of suspended solids in the River Danube. *Aquat. Microbial Ecol.* **10**: 37–44.
20. Grossart, H.P. and Ploug, H. (2000). Bacterial production and growth efficiencies: Direct measurements on riverine aggregates. *Limnol. Oceanog.* **45**: 436–445.
21. Smith, D.C., Simon, M., Alldredge, A.L., and Azam, F. (1992). Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. *Nature* **359**: 139–142.
22. Battin, T.J., Kaplan, L.A., Newbold, J.D., and Hansen, C.M. (2003). Contributions of microbial biofilms to ecosystem processes in stream mesocosms. *Nature* **426**: 439–442.
23. Liu, Y. and Tay, J.-H. (2001). Metabolic response of biofilm to shear stress in fixed-film culture. *J. Appl. Microbiol.* **90**: 337–342.
24. Battin, T.J. et al. (2003). Effects of current velocity on the nascent architecture of stream microbial biofilms. *Appl. Environ. Microbiol.* **69**: 5443–5452.
25. Fenchel, T., King, G.M., and Blackburn, T.H. (1998). *Bacterial Biogeochemistry: The Ecophysiology of Mineral Cycling*. Academic Press, San Diego, CA, p. 307.
26. Tank, J.L. and Webster, J.R. (1998). Interaction of substrate and nutrient availability on wood biofilm processes in streams. *Ecology* **79**: 2168–2179.
27. Fischer, H. (2003). The role of biofilms in the uptake and transformation of dissolved organic matter. In: *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter*. S.E.G. Findlay and R.L. Sinsabaugh (Eds.). Academic Press, San Diego, CA, pp. 285–313.
28. Espeland, E.M., Francoeur, S.N., and Wetzel, R.G. (2001). Influence of algal photosynthesis on biofilm bacterial production and associated glucosidase and xylosidase activities. *Microbial Ecol.* **42**: 524–530.
29. Battin, T.J. (2000). Hydrodynamics is a major determinant of streambed biofilm activity: From the sediment to the reach scale. *Limnol. Oceanog.* **45**: 1308–1319.
30. Emerson, D. and Revsbech, N.P. (1994). Investigation of an iron-oxidizing microbial mat community located near Aarhus, Denmark: Field studies. *Appl. Environ. Microbiol.* **60**: 4022–4031.
31. Brock, T.D. (1994). *Life at High Temperatures*. Yellowstone Association for Natural Science, Yellowstone National Park, Wyoming, p. 31.

MICROORGANISMS IN THEIR NATURAL ENVIRONMENT

Matthias Koschorreck
Center of Environmental
Research (UFZ)
Magdeburg, Germany

GENERAL REMARKS

In their natural aquatic environment, microbes face conditions that principally differ from the picture of a mixture of different microbes swimming around in a kind of liquid culture. The study of microorganisms in laboratory cultures often gives limited information about the role and behavior of microorganisms in their natural habitat. Our understanding of microbes in their natural aquatic habitat was considerably increased by the development of *in situ* methods in the last decades. The introduction of microsensors allowed characterizing the microbial habitat on the size scale of microorganisms. The recent advances in microscopic techniques (confocal laser scanning microscopy, atomic force microscopy) offer the possibility of real 3-D examination of microbial aggregates or biofilms. Molecular techniques such as *in situ* hybridization to study the phylogeny and physiology of microorganisms in their natural environment are rapidly, improving.

A characteristic of the natural habitat is its heterogeneity. Hot spots of high activity may be surrounded by a "desert." Natural habitats are characterized by **gradients**. The co-occurrence of various, sometimes opposing, gradients creates a high diversity of ecological niches.

Depending on transport velocities and reaction rates, these gradients can occur on a meter scale (in open water) down to a micrometer (in sediments) or even submicrometer scale (in microbial aggregates) (Fig. 1). The natural environment is not static but subject to fast **temporal fluctuations**. Temporal fluctuations occur on different timescales from seasonal effects down to very short-term changes within seconds or minutes. A phototrophic microbial mat, for example, can shift from oxygen oversaturation to anoxia within minutes when it is darkened (1). Microorganisms have developed different strategies to cope with changing conditions. The largest part of the bacteria in aquatic systems is in a metabolically inactive dormant state, waiting for favorable conditions. Motility, often in combination with chemotaxis, is an important strategy for following changing gradients. Microorganisms can also alter their phenotypic properties in response to changing conditions. On the scale of microorganisms, surface effects are important and **diffusion** becomes the dominant transport process. Solid surfaces are covered by a diffusive boundary layer (DBL) (2) (Fig. 1).

The term microbes means bacteria, algae, fungi, and protozoa. This article focuses on the ecology of bacteria in aquatic environments. A general overview of aquatic microbiology can be found in Ford (3), Hurst et al. (4), or Sorokin (5).

HABITATS

Aquatic systems offer different habitats for microbes to live in: free water, sediment, and hyporheic interstitial (6). Of special importance are interfaces such as the surface of water, sediment, rocks (epilithon), and aquatic plants (epiphyton). Microbes that live in **free water** (pelagic organisms) can be dispersed as single cells or concentrated in aggregates (Fig. 2). Organic aggregates, which are larger than 500 μm, are called lake, river, or marine "snow" (7). They are hot spots of microbial life in aquatic systems and are enriched in nutrients compared to the surrounding water. The formation of aggregates allows direct syntrophic interactions such as the exchange of metabolic intermediates.

Natural waters typically contain between 0.5 and 5×10^6 bacteria per mL (5). Microbial numbers, diversity, and activities are usually higher in the **sediment** and at surfaces than in free water. Sediments typically contain about 10^9 bacteria per mL. The fraction of metabolically active cells is also higher in sediment compared to free water. In a Swedish lake, only 4% of the pelagic bacteria, 37% of epiphytic bacteria, and 46% of sediment bacteria were metabolically active (8). This can have several causes such as higher nutrient availability, higher habitat diversity, or protection against predation. Microbial numbers and activities usually decrease exponentially as sediment depth increases. The absence of turbulent mixing allows the patchiness and spatial (mostly vertical) gradients to form in the sediment. The situation is especially complex in the rhizosphere of aquatic plants.

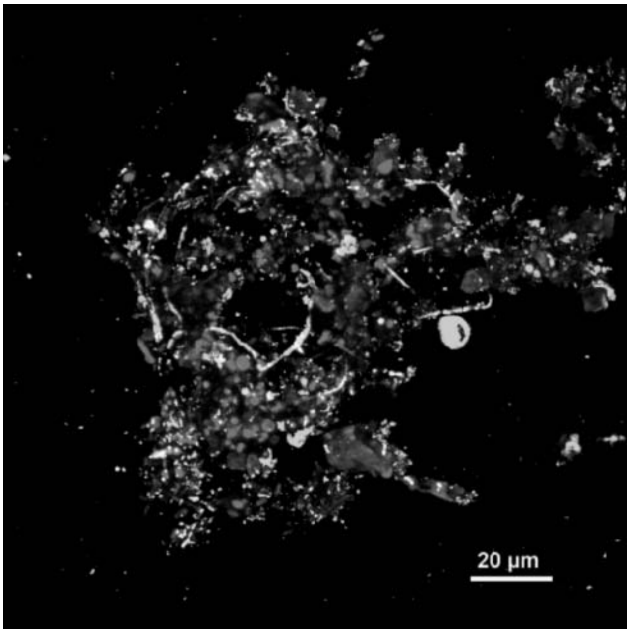

Figure 2. Confocal laser scanning micrograph of a lotic aggregate from the River Saale, Germany. Maximum intensity projection of 31 sections at 1 μm distance from four channels. Color allocation: reflection = white, autofluorescence of algae = blue, nucleic acid staining of bacteria = green, lectin stained glycoconjugates = red (courtesy of Thomas R. Neu) This figure is available in full color online.

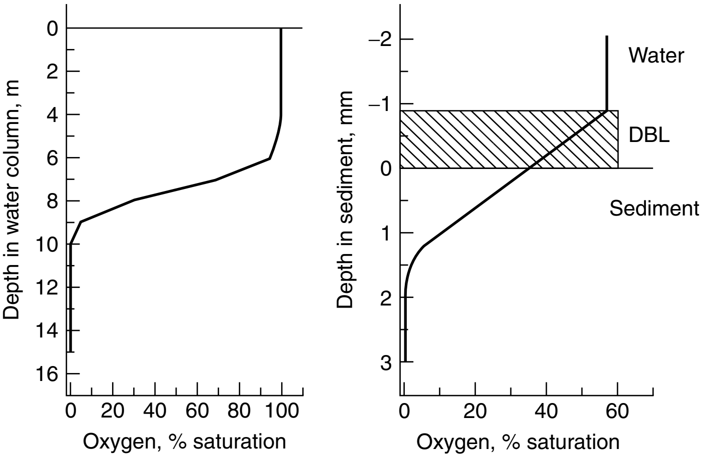

Figure 1. Vertical gradient of oxygen (**a**) in a lake with anoxic hypolimnion and (**b**) at the sediment water interface. DBL = diffusive boundary layer.

Surfaces are often covered by biofilms (9,10) or microbial mats (11–13). Living attached to a surface allows microbes to stay in favorable places under conditions of moving water. A biofilm is a "surface accumulation, which is not necessarily uniform in time or space, that comprises cells immobilized at a substratum and frequently embedded in an organic polymer matrix of microbial origin" (10). The solid surface on which a biofilm develops is called the "substratum." Living conditions in biofilms may differ substantially from conditions in free water. Biofilms are enriched in nutrients, are a place of intensive biotic interactions, and may offer protection against unfavorable conditions. Biofilms are not stable but undergo different successional stages. Biofilm growth starts due to adsorption of cells to a substratum. As the biofilm growth thickers, its composition changes, transport processes influence metabolic activities in the biofilm, and it becomes more and more affected by shear stress which finally leads to detachment and washout of parts of the biofilm.

The liquid–gas interface is characterized by high concentrations of oxygen and hydrophobic substances and high light intensities. It is colonized by a special microflora called neuston (14).

Microbes play an important role in **artificial aquatic systems** such as wastewater treatment plants, biotechnological reactors, and drinking water supplies. Living conditions in such systems can differ substantially from natural habitats and are often characterized by high nutrient content and homogeneity (in reactors) or very low nutrient content (in drinking water). The general habitat characteristics and the adaptations of microorganisms, however, are similar in natural and artificial aquatic systems.

ABIOTIC FACTORS

Water

Per definition, aquatic bacteria live in habitats where water is available. Many natural waterbodies, however, temporarily dry out due to water level fluctuations. Organisms in such habitats develop strategies to survive periods of dessication. They may form spores or other resting stages.

Even if microbes are swimming in liquid water, it might not be available to them because of low water activity. Water activity expresses the ratio of water in the vapor phase to the amount in vapor saturated air (ranging from 0 to 1.0). Low water activity can be caused by high salt concentration. Most bacteria need a water activity higher than 0.98 which is the activity of sea water at 25 °C (14). Especially resistant halobacteria can live at water activities as low as 0.6. They are found in salt lakes or salt saturated evaporation ponds.

Substrates

Microbes need different types of substrates. These include electron donors, electron acceptors, macronutrients (carbon, nitrogen, and sulfur), micronutrients (e.g., different metals and trace elements). Many organisms also depend on organic substrates. The availability of electron acceptors (O_2, $NO_3{}^{2-}$, Fe^{III}, $SO_4{}^{2-}$, CO_2) is a major determinant of which physiological groups of microorganisms are predominant. Surface waters are mostly oxic, and aerobic microbes dominate. In sediments and under certain circumstances also in free water (groundwater, hypolimnion), conditions are anoxic, and anaerobic organisms dominate microbial populations (Fig. 1). Anoxic bacteria are also frequently found in oxic environments (15). These can be facultative anaerobes such as denitrifiers and also oxygen resistant anaerobes or strict anaerobes that live in anoxic microsites (e.g., inside microbial aggregates). The oxic–anoxic interface is a place of high microbial activity and special adaptations (16). At the oxic–anoxic interface, redox cycling of mobile electron carriers such as Fe^{II}/Fe^{III} or $H_2S/SO_4{}^{2-}$ can occur on a small spatial scale.

Heterotrophic microbes depend on organic substrates as a carbon source; autotrophs such as photosynthetic organisms or nitrifying bacteria can use CO_2 that is available in most aquatic systems which are buffered by a carbon buffer system. Exceptions are acidic waters where inorganic carbon is available only as physically dissolved CO_2 and potentially limits autotrophic growth.

Concentrations of nutrients are usually low in natural environments. Consequently, cells in natural water are usually much smaller than cells grown in culture media, and generation times are longer. Microbes have developed several strategies to survive starvation. Most obvious is the formation of metabolically inactive resting stages like spores or the increased affinity of nutrient uptake systems. All nutrients have to cross the diffusive boundary layer before they can be taken up by an organism. The diffusion flux to a cell depends on its surface to volume ratio, so smaller organisms are more efficient in diffusive substrate uptake (17). Microbial growth in natural environments is often very slow (18).

Temperature

Temperature is the second most important parameter besides nutrient availability that controls microbial growth in natural aquatic systems. Aquatic microbes can live in liquid water at all temperatures between freezing and boiling. The world record for hyperthermophilic bacteria is currently 113 °C, and many extreme thermophiles cannot growth at "normal" temperatures (19). Such extreme temperatures occur at submarine hydrothermal vents where water remains liquid at temperatures above 100 °C due to the hydrostatic pressure. Most natural aquatic habitats, however, are characterized by low temperatures. In deep lakes and in the oceans, great parts are colder than 5 °C, and microbial processes are slow. The temperature in the natural environment is usually below the optimum growth temperature of bacteria. Even polar or alpine ice is colonized by microorganisms, and the boundary between surface ice and water is a place of microbial life (20).

Light

Microbes are affected by light in different ways, both directly or indirectly. Light is the energy source of photosynthesis which can take place under aerobic and anaerobic conditions. Anoxic photosynthesis typically takes place at the chemocline where anoxic conditions meet maximum light intensities. High abundances of purple sulfur bacteria can be concentrated in only a few centimeter thick layer of water (21). It is especially important for phototrophic organisms to maintain their vertical position in the water column. To counteract sedimentation, microbes may use active movement or regulate buoyancy by intracellular gas vesicles (22). In moving water, pelagic organisms are continuously transported from regions of high light near the surface to darker conditions further down and back (23).

At the water surface or on mud surfaces, solar radiation can be inhibiting for microbial life. One strategy to prevent photodamage is migration of motile organisms into the sediment during periods of high illumination (24). Microorganisms have pigments to protect them from photodamage. UV radiation can alter nutrient conditions by mobilizing dissolved organic matter (25) or by reducing iron (26).

pH

Natural aquatic systems cover the whole pH range from 0 to 14, and all of them are inhabited by microbes (27). Acidophilic bacteria (28) live in extremely acidic waters. Basophiles and alkaliphiles live in hypersaline waters. The activity of microorganisms can modify the pH of their habitat. The oxidation of iron or sulfur compounds can lower the pH to extreme values below pH 2. Photosynthetic activity can create local pH values higher than pH 10 in phototrophic biofilms (29).

Shear Stress

Water movement affects microbes mechanically in aquatic habitats and is an important regulator of biofilm growth. High current velocities cause detachment and washout of attached microorganisms (9). On the other hand, drift is an import mechanism for colonizing new habitats.

BIOTIC INTERACTIONS

Food Web

Microbes play an important role in the food web in aquatic environments. Bacteria, flagellates, and microzooplankton consume dissolved organic matter (DOM) which is released by phytoplankton. By this **microbial loop** (30), energy and nutrients are returned to the main food chain. In unproductive aquatic systems, bacterial production can exceed phytoplankton production and make these systems net sources of CO_2 (31).

Bacteria **compete** with eukaryotic algae and phagotrophs for resources. Usually, bacteria play a larger role in low productivity (oligotrophic) ecosystems compared to highly productive (eutrophic) systems. Oligotrophic systems that have low nutrient concentrations and higher proportions of soluble organic C favor heterotrophic bacteria to phytoplankton and phagotrophic heterotrophs (32). The microbial mineralization of organic matter is a complex network of different reactions carried out by different organisms.

Syntrophy

In anoxic environments, the degradation of organic matter resulting in the accumulation of uncompletely oxidized intermediates such as acetate or hydrogen gas is energetically unfavorable. Methanogens and sulfate reducers can, for example, consume hydrogen gas and lower its partial pressure so that the hydrogen producing process becomes thermodynamically favorable. This “interspecies hydrogen transfer” allows the degradation of compounds which are refractory to degradation by the pure culture of a single species (33). A prerequisite of such syntrophic interaction is that the organisms are close spatially.

Grazing

Bacterial numbers in pelagic systems and in surface sediment are often controlled by grazing. Microorganisms have developed different strategies against grazing. They can change cell size, change cell shape to longer forms, or form multicellular aggregates (34). Some bacteria are not digested by higher organisms but live as endosymbionts inside bigger aquatic organisms.

Bioturbation

The activity of macroorganisms can greatly modify the sediment environment (35). Organisms and particles can be moved by bioturbation, and porewater chemistry might be changed. Bioturbation can increase both homogeneity or patchiness of the sediment and introduce short-term changes of conditions.

Bacteriophages

Viruses are the most abundant biological entities in aquatic systems. Typical abundances of these 20 to 200 nm particles are 10^{10} per liter in surface water. They probably infect all organisms and influence biogeochemical and ecological processes as well as gene transfer among microorganisms (36).

CONCLUDING REMARK

Due to the development of new methods for studying microbes in their natural environment, we now have a first impression of the complexity of the natural habitat. Many issues such as biotic interactions in aggregates and biofilms, the reaction of microbes to quickly changing environmental conditions, and their distribution in very heterogeneous environments remain important topics for further research. The study of the distribution, role, and function of microbes in their natural aquatic environment is one of the most exciting areas in aquatic ecology.

BIBLIOGRAPHY

1. Revsbech, N.P. and Jørgensen, B.B. (1986). Microelectrodes: Their use in microbial ecology. *Adv. Microb. Ecol.* **9**: 293–351.

2. Boudreau, B.P. and Jørgensen, B.B. (Eds.). (2001). *The Benthic Boundary Layer—Transport Processes and Biogeochemistry*. Oxford University Press, Oxford.
3. Ford, T.E. (1993). *Aquatic Microbiology—An Ecological Approach*. Blackwell, Boston.
4. Hurst, C.J., Knudsen, G.R., McInerney, M.J., Stetzenbach, L.D., and Walter, M.V. (Eds.). (1997). *Manual of Environmental Microbiology*. ASM Press, Washington, D.C.
5. Sorokin, Y.I. (1999). *Aquatic Microbial Ecology*. Backhuys Publishers, Leiden.
6. Pusch, M. et al. (1998). The role of micro-organisms in the ecological connectivity of running waters. *Freshwater Biol.* **40**: 453–495.
7. Simon, M., Grossart, H.P., Schweitzer, B., and Ploug, H. (2002). Microbial ecology of organic aggregates in aquatic ecosystems. *Aquatic Microb. Ecol.* **28**: 175–211.
8. Haglund, A.L., Törnblom, E., Boström, B., and Tranvik, L. (2002). Large differences in the fraction of active bacteria in plankton, sediments and biofilm. *Microb. Ecol.* **43**: 232–241.
9. Characklis, W.G. and Marshall, K.C. (Eds.). (1990). *Biofilms, Ecological and Applied Microbiology*. John Wiley & Sons, New York, NY.
10. Bryers, J.D. (Ed.). (2000). *Biofilms II—Process Analysis and Application*. John Wiley & Sons, New York, NY.
11. Cohen, Y. and Rosenberg, E. (1989). *Microbial Mats—Physiological Ecology of Benthic Microbial Communities*. American Society for Microbiology, Washington.
12. Stal, L.J. and Caumette, P. (Eds.). (1994). *Microbial Mats—Structure, Development and Environmental Significance. NATO ASI Series*. Springer, Berlin.
13. Krumbein, W.E., Paterson, D., and Stal, L.J. (Eds.). (1994). *Biostabilization of Sediments*. BIS Verlag, Oldenburg.
14. Schlegel, H.G. and Jannasch, H.W. (1991). Prokaryotes and Their Habitats. In: A. Balows, H.G. Trüper, M. Dworkin, W. Harder, and K.H. Schleifer (Eds.). *The Prokaryotes*. Springer, Berlin, pp. 75–125.
15. Peters, V. and Conrad, R. (1995). Methanogenic and other strictly anaerobic bacteria in desert soil and other oxic soils. *Appl. Environ. Microb.* **61**: 1673–1676.
16. Brune, A., Frenzel, P., and Cypionka, H. (2000). Life at the oxic-anoxic interface, microbial activities and adaptations. *FEMS Microb. Rev.* **24**: 699–710.
17. Koch, A. (1996). What size should a bacterium be? A question of scale. *Ann. Rev. Microbiol.* **50**: 317–348.
18. Koch, A.L. (1997). Microbial physiology and ecology of slow growth. *Microbiol. Molec. Biol. Rev.* **61**: 305–318.
19. Rothschild, L.J. and Mancinelli, R.L. (2001). Life in extreme environments. *Nature* **409**: 1092–1101.
20. Thomas, D.N. and Diekmann, G.S. (2002). Antarctic sea ice—a habitat for extremophiles. *Science* **295**: 641–644.
21. Overmann, J., Beatty, J.T., Krouse, H.R., and Hall, K.J. (1996). The sulfur cycle in the chemocline of a meromictic salt lake. *Limnol. Oceanogr.* **41**: 147–156.
22. Konopka, A. (1982). Buoyancy regulation and vertical migration by *Oscillatoria rubescens* in Crooked Lake, Indiana. *Br. Phycol. J.* **17**: 427–442.
23. Reynolds, C.S. (1997). Vegetation processes in the pelagic, a model for ecosystem theory. In: *Excellence in Ecology*. O. Kinne (Ed.). Ecology Institute, Oldendorf.
24. Kingston, M.B. (1999). Effect of light on vertical migration and photosynthesis of *Euglena proxima* (Euglenophyta). *J. Phycol.* **35**: 245–253.
25. Karentz, D. et al. (1994). Impact of UV-B radiation on pelagic freshwater ecosystems, report of working group on bacteria and phytoplankton. *Arch. Hydrobiol.* **43**: 31–69.
26. Herzsprung, P. et al. (1998). Vertical and annual distribution of ferrric and ferrous iron in acidic mining lakes. *Acta Hydrochim. Hydrobiol.* **26**: 253–262.
27. Seckbach, J. (Ed.). (1999). *Enigmatic Microorganisms and Life in Extreme Environments*. Kluwer, Dordrecht.
28. Johnson, D.B. (1998). Biodiversity and ecology of aciophilic microorganisms. *FEMS Microb. Ecol.* **27**: 307–317.
29. Revsbech, N.P., Nielsen, J., and Hansen, P.K. (1988). Benthic primary production and oxygen profiles. In: *Nitrogen Cycling in Coastal Marine Environments*. T.H. Blackburn and J. Sörensen (Eds.). John Wiley & Sons, New York, NY, pp. 69–83.
30. Azam, F. et al. (1983). The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* **10**: 257–263.
31. del Giorgio, P.A., Cole, J.J., and Cimbleris, A. (1997). Respiration rates in bacteria exceed phytoplankton production in unproductive aquatic systems. *Nature* **385**: 148–151.
32. Cotner, J.B. and Biddanda, B.A. (2002). Small players, large role, microbial influence on biogeochemical processes in pelagic aquatic ecosystems. *Ecosystems* **5**: 105–121.
33. Schink, B. (1991). Syntrophism among Prokaryotes. In: *The Procaryotes*. A. Balows, H.G. Trüper, M. Dworkin, W. Harder, and K.H. Schleifer (Eds.). Springer, Berlin, pp. 276–299.
34. Jürgens, K. and Sala, M. (2000). Predation-mediated shifts in size distribution of microbial biomass and activity during detritus decomposition. *Oikos* **91**: 29–40.
35. Aller, Y., Woodin, S.A., and Aller, R.C. (Eds.). (2001). *Organism-Sediment Interactions*. University of South Carolina, Columbia.
36. Fuhrman, J.A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature* **399**: 541–548.

CALIBRATION OF HYDRAULIC NETWORK MODELS

SRINIVASA LINGIREDDY
LINDELL E. ORMSBEE
DON J. WOOD
University of Kentucky
Lexington, Kentucky

INTRODUCTION

Availability of reliable network modeling software coupled with affordable computing hardware technology has led to rapid growth in the use of hydraulic network models for design, planning, and operational studies of water distribution systems. Friendly graphical user interfaces (GUIs) and integrated computer-aided design (CAD) and geographic information systems (GIS) modeling environments helped drastically reduce model development time and the associated costs. The validity of these models, however, depends largely on the accuracy of input data and the assumptions made in developing the model. Although carefully developed models tend to have greater control on much of the data associated with the model, certain model parameters exist that are either not readily available or difficult to obtain. Such parameters typically include pipe roughness factors and nodal demands and their spatial and temporal distribution. As a result of the difficulty of obtaining economic and reliable measurements of both of these parameters, final model values are normally determined through the process of model calibration. Model calibration involves adjustment of these and other uncertain network model parameters until the model results closely approximate actual observed conditions as measured from field data. In general, a network model calibration effort should encompass seven basic steps:

1. Identification of intended use of the model
2. Identification of calibration model parameters and their initial estimates
3. Model studies to determine calibration data sources
4. Data collection
5. Macrocalibration
6. Sensitivity analysis
7. Microcalibration

These steps have been successfully tested by the authors on several model calibration studies. Each of these steps is discussed in the following sections.

IDENTIFICATION OF INTENDED USE OF THE MODEL

Before calibrating a hydraulic network model, it is important to first identify its intended use (e.g., pipe sizing for master planning, operational studies, design projects, rehabilitation studies, water quality studies) and the associated type of hydraulic analysis (steady-state versus extended-period). Usually, the type of analysis is directly related to the intended use. For example, water quality and operational studies require an extended-period analysis, whereas some planning or design studies may be performed using a steady-state analysis. In the latter, the model predicts system pressures and flows at an instant in time under a specific set of operating conditions and demands (e.g., average or maximum daily demands). In extended-period analysis, the model predicts system pressures and flows over an extended period (typically 24 hours).

Both the intended use of the model and the associated type of analysis provide some guidance about the type and quality of collected field data and the desired level of agreement between observed and predicted flows and pressures (1). Models for steady-state applications can be calibrated using multiple static flow and pressure observations collected at different times of day under varying operating conditions (e.g., fire hydrant flow test data). On the other hand, models for extended-period applications require field data collected over an extended period (e.g., one to seven days).

In general, a higher level of model calibration is required for water quality analysis or an operational study than for a general planning study. For example, determining ground elevations using a topographic map may be adequate for the latter, whereas the former may require an actual field survey. Such considerations obviously influence the methods used to collect the necessary model data and the subsequent calibration steps.

IDENTIFICATION OF CALIBRATION MODEL PARAMETERS AND THEIR INITIAL ESTIMATES

The second step in calibrating a hydraulic network model is to determine initial estimates of the primary model parameters. Although most models will have some degree of uncertainty associated with several model parameters, the two model parameters that normally have the greatest degree of uncertainty are the pipe roughness coefficients and the demands to be assigned to each junction node.

Pipe Roughness Values

Initial estimates of pipe roughness values may be obtained using average literature values or directly from field measurements. Various researchers and pipe manufacturers have developed tables that provide estimates of pipe roughness as a function of various pipe characteristics, such as pipe material, pipe diameter, and pipe age. Table 1 shows typical Hazen–William roughness factors (2). Although such tables may be useful for new pipes, their specific applicability to older pipes decreases significantly as the pipes age. In addition, the variation in roughness with pipe age depends on other factors, such as water chemistry, and will vary from location to location. As a result, initial estimates of pipe roughness for all pipes other than relatively new pipes should normally come directly from field testing. Even when new pipes are being used, it is helpful to verify the roughness values in the field because the roughness coefficient used in the model may actually represent a composite of several secondary factors such as fitting losses and system skeletonization.

Adjusting Pipe Roughness vs. Pipe Diameter

The approach of adjusting pipe roughness values assumes that the pressure drop is entirely because of frictional and other minor losses within the pipeline. If the intended use for the calibrated model does not include water quality simulations, this approach of adjusting pipe roughness value will have little or no impact on model results. On the other hand, this approach might seriously affect results from water quality simulations, which is particularly true when a substantial portion of the distribution system comprises fairly old pipelines. Loss of carrying capacity in such pipes is not only because of increased roughness but also because of the reduced flow area resulting from tuberculation, scaling, sedimentation, and biofilm growth. Flow area reductions as high as 50% are not uncommon in certain distribution systems. Although adjusting pipe roughness to reflect the reduction in carrying capacity will not affect hydraulic gradeline and pressure calculations, significant deviations in simulated water quality parameters are possible as the erroneous velocities are used in water quality calculations. Therefore, for more accurate water quality predictions, it might be prudent to adjust the pipe diameters by limiting roughness coefficients to certain predetermined values. For example,

Table 1. Typical Hazen–William Pipe Roughness Factors

Pipe Material	Pipe Age	Pipe Size (Inches)	C Factor
Cast Iron	New	All Sizes	130
	5 years old	12 and Over	120
		8	119
		4	118
	10 years old	24 and Over	113
		12	111
		4	107
	20 years old	24 and Over	100
		12	96
		4	89
	30 years old	30 and Over	90
		16	87
		4	75
	40 years old	30 and Over	83
		16	80
		4	64
	50 years old	40 and Over	77
		24	74
		4	55
Welded Steel	Values of C the same as for cast-iron pipes, 5 years older		
Riveted Steel	Values of C the same as for cast-iron pipes, 10 years older		
Wood Stave	Average value, regardless of age		120
Concrete or Concrete Lined	Large sizes, good workmanship, steel forms		140
	Large sizes, good workmanship, wooden forms		120
	Centrifugally spun		135
Vitrified	In good condition		110
Plastic or Drawn Tubing			150

1 inch = 25.4 mm.

if the computed Hazen–William roughness coefficient for a pipeline is 40, one could limit the coefficient to about 90 and compute the pipe diameter that would produce the measured pressure drop.

Spatial and Temporal Distribution of Nodal Demands

The second major parameter determined in calibration analysis is the average (steady-state analysis) or temporally varying (extended-period analysis) demand to be assigned to each junction node. Initial average estimates of nodal demands can be obtained by identifying a region of influence associated with each junction node, identifying the types of demand units in the service area, and multiplying the number of each type by an associated demand factor. Alternatively, the estimate can be obtained by first identifying the area associated with each type of land use in the service area and then multiplying the area of each type by an associated demand factor. In either case, the sum of these products is an estimate of the demand at the junction node. Although in theory the first approach should be more accurate, the latter approach is more expedient. Estimates of unit demand factors are available from water resource handbooks and textbooks (3).

MODEL STUDIES TO DETERMINE CALIBRATION DATA SOURCES

After model parameters have been estimated, the accuracy of the model parameters can be assessed, which is done by executing the computer model using the estimated parameter values and observed boundary conditions and then comparing the model results with the results from actual field observations. Collecting accurate and appropriate field data is crucial for proper calibration of a hydraulic network model. One should use at least 10 to 20 pressure readings measured at locations evenly spread over the water distribution system for steady-state calibration. These pressure readings should ideally be collected when the network is under “stressed” conditions, i.e., during a peak-hour demand condition, during filling cycle of a ground-level pumped storage tank, or during hydrant flow tests. Under these conditions, the increased flowrates make the pressure readings more sensitive to small changes in pipe roughness values. The hydrant flows that produce more than 15 psi drop from static pressures are ideal for calibration studies. In most distribution systems, the peak-hour demands may not provide 15 psi drop at sufficient number of locations. Ground-level pumped storage tanks are not common in most distribution systems. Therefore, hydrant flow tests provide a logical alternative for pressure reading in most systems. However, not every hydrant flow will result in an adequate pressure drop. Hydrants located at or near large pressure mains, supply sources, and elevated storage tanks tend to provide less than the desired pressure drop. On the other hand, certain hydrants might yield very little flow while showing large pressure drops (e.g., 80 psi of pressure drop while yielding only 200 gpm) because of localized losses such as constricted flow areas. Although a network model that is yet to be calibrated may not be meaningful for planning and operational studies, it can provide some insight into the locations for hydrant flow tests that provide reasonable flows and adequate pressure drops. The following procedure is recommended to locate hydrants for calibration field studies.

- Perform hydrant flow analysis on all hydrants in the distribution system. Hydrant flow analysis appears to be a standard feature in most commercial modeling packages and provides information on flowrate that a hydrant can deliver while maintaining a minimum pressure of 20 psi throughout the distribution system. One could use the nearest junction nodes in place of hydrants if the model does not contain the hydrant information.
- Compute the flowrate per unit pressure drop (difference between static and residual pressure) for each hydrant.
- Rank order the hydrants from lowest to highest of flowrate per unit pressure drop.
- Select about 15 to 25 hydrants from the top of the list that are evenly scattered throughout the distribution system. Care must be taken to avoid hydrants with unusually low flowrates to pressure drop ratios. A closer look at these hydrants and related flows might reveal problems with the network model. The authors have also found that unusually high flowrate to pressure drop ratios could be because of network anomalies that can be corrected by a close observation of network model parameters.

DATA COLLECTION

The next step is to acquire calibration data from field studies. This section describes the general procedures for collecting and recording field data.

Fire Flow Tests

Fire flow tests are useful for collecting both discharge and pressure data for use in calibrating hydraulic network models. Such tests are normally conducted using both a normal pressure gauge (for measuring both static and dynamic heads) and a Pitot gauge (for use in calculating discharge). In performing a fire flow test, at least two separate hydrants are first selected for use in the data collection effort. One hydrant is identified as the pressure or residual hydrant, whereas the remaining hydrant is identified as the flow hydrant. The general steps for performing a fire flow test are illustrated by McEnroe et al. (4).

In order to obtain sufficient data for an adequate model calibration, it is important that data from several fire flow tests be collected. Before conducting each test, it is also important that the associated system boundary condition data be collected, which includes information on tank levels, pump status, etc. It is a common practice for the local fire departments to conduct hydrant flow tests and record the time of day and corresponding flows and pressures. However, in most cases, such records do not include the boundary conditions associated with each hydrant flow test, as the main purpose for their tests is to rate the fire hydrant and not necessarily for hydraulic calibration. Therefore, care must be taken to avoid hydrant flow data that does not include the associated boundary conditions data. Table 2 may be used as a template when collecting calibration data using hydrant flow tests.

Telemetry Data

In addition to static test data, data collected over an extended period of time (typically 24 hours) can be very useful for use in calibrating network models. The most common type of data will include flowrate data, tank water level data, and pressure data. Depending on the level of instrumentation and telemetry associated with the system, much of the data may be already collected as part of the normal operations. For example, most systems collect and record tank levels and average pump station discharges on an hourly basis. These data are especially useful in verifying the distribution of demands among the various junction nodes. If such data are available, the data should first be checked for accuracy before use in the calibration effort. If such data are not readily available, the modeler may have to install temporary pressure gauges or flowmeters in order to obtain the data. In the absence of flowmeters in lines to tanks, inflow or discharge flowrates can be inferred from incremental readings of the tank level.

In systems with ground-level pumped storage tanks (Fig. 1), the tanks are filled during off-peak demand periods using the high system pressures. Pressures in the distribution system drop dramatically during tank-filling operation, thereby providing an opportunity to use the associated data for model calibration. Tank telemetry data, along with a few static pressure readings within the distribution system where a significant drop occurs, could be used for calibrating the network model in lieu of hydrant flow test data.

EVALUATE MODEL RESULTS

Using fire flow data, the model is used to simulate the discharge from one or more fire hydrants by assigning the observed hydrant flows as nodal demands within the model. The flows and pressures predicted by the model are then compared with the corresponding observed values in an attempt to assess model accuracy. In using telemetry data, the model is used to simulate the variation of tank water levels and system pressures by simulating

Table 2. Sample Hydrant Flow Data Collection Table for Calibration Studies

			Flowing Hydrant					Residual Hydrant				Tank Levels			Pump Status			Remarks
Test	Date	Time	Name	Elevation	Static Pressure	Discharge Pressure	Flowrate	Name	Elevation	Static Pressure	Residual Pressure	A	B	….	A	B	….	
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
13																		
14																		
15																		
16																		
17																		
18																		

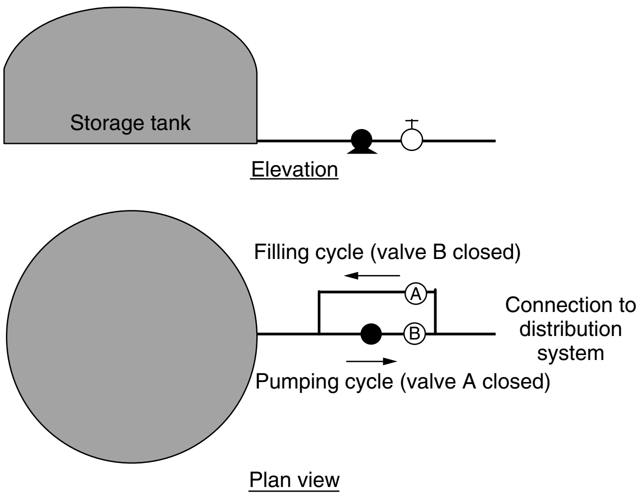

Figure 1. Schematic of a ground-level pumped storage facility.

the operating conditions for the day over which the field data was collected. The predicted tank water levels are then compared with the observed values in an attempt to assess model accuracy.

Model accuracy may be evaluated using various criteria. The most common criteria are absolute pressure difference (normally measured in psi) or relative pressure difference (measured as the ratio of the absolute pressure difference to the average pressure difference across the system). In most cases, a relative pressure difference criteria is normally preferred. For extended period simulations, comparisons are normally made between the predicted and observed tank water levels. To a certain extent, the desired level of model calibration will be related to the intended use of the model. For example, a higher level of model calibration will normally be required for water quality analysis or an operational study as opposed to use of the model in a general planning study. For most applications, a maximum state variable (i.e., pressure grade, water level, flowrate) deviation of less than 5–10% is generally regarded as satisfactory (1).

Deviations between results of the model application and the field observations may be caused by several factors, including: (a) erroneous model parameters (i.e., pipe roughness values and nodal demand distribution), (b) erroneous network data (i.e., pipe diameters, lengths, etc.), (c) incorrect network geometry (i.e., pipes connected to the wrong nodes, etc.), (d) incorrect pressure zone boundary definitions, (e) errors in boundary conditions (i.e., incorrect PRV value settings, tank water levels, pump curves, etc.), (f) errors in historical operating records (i.e., pumps starting and stopping at incorrect times), (g) measurement equipment errors (i.e., pressure gauges not properly calibrated, etc.), and (h) measurement error (i.e., reading the wrong values from measurement instruments). The last two sources of errors can hopefully be eliminated or at least minimized by developing and implementing a careful data collection effort. Elimination of the remaining errors will frequently require the iterative application of the last three steps of the model calibration process—macrocalibration, sensitivity, and microcalibration. Each of these steps is described in the following sections.

MACROCALIBRATION

In the event that one or more of the measured variable values are different from the modeled values by an amount that is deemed to be excessive (i.e., greater than 30%), it is likely that the cause for the difference may extend beyond errors in the estimates for either the pipe roughness values or the nodal demands. Possible causes for such differences are many, but may include

- valves that are open in the system but modeled as closed or partly closed
- valves that are closed or partly closed in the system but not modeled correctly
- inaccurate pump curves or pumps defined by useful power rather than head/flow data
- incorrect pump (on/off) status data or pump operating speeds for variable speed pumps
- inaccurate tank telemetry data
- inaccurate flow or pressure gauges
- incorrect pipe sizes and lengths (typographical errors)
- inaccurate pressure regulating valve settings
- incorrect network geometry
- inaccurate junction or hydrant elevations
- hydrants located on wrong pipes near multiple pipe junctions (Fig. 2)
- incorrect pressure zone boundaries
- missing or incorrect demand data—especially pertaining to large users (e.g., hospitals, schools, etc.)

The best way to adequately address such errors is to systematically review the data associated with the model in order to ensure its accuracy (5). In most cases, some data will be less reliable than other data. This observation provides a logical place to start in an attempt to identify the problem. Model sensitivity analysis provides another means of identifying the source of discrepancy. For example, if it is suspected that a valve is closed, this assumption can be modeled by simply closing the line in the model and evaluate the resulting pressures. Potential errors in pump curves can sometimes be minimized by simulating the pumps with negative inflows set equal to observed pumps discharges (6).

Plotting pressure contours pertaining to the simulation that produced undue low pressure (compared with the measured value) might provide an insight into certain problem areas. For example, a significant flow constriction (wrong pipe diameter or partially closed valve) in the model would produce rapid pressure drops that could be easily traced using the pressure contour plots. Probing into headloss/1000 values in the vicinity of the hydrant under question might help trace the problem pipes. In general, it would be relatively easier to trace the problems if the simulated pressure reading is significantly lower than the measured value. Only after the model results

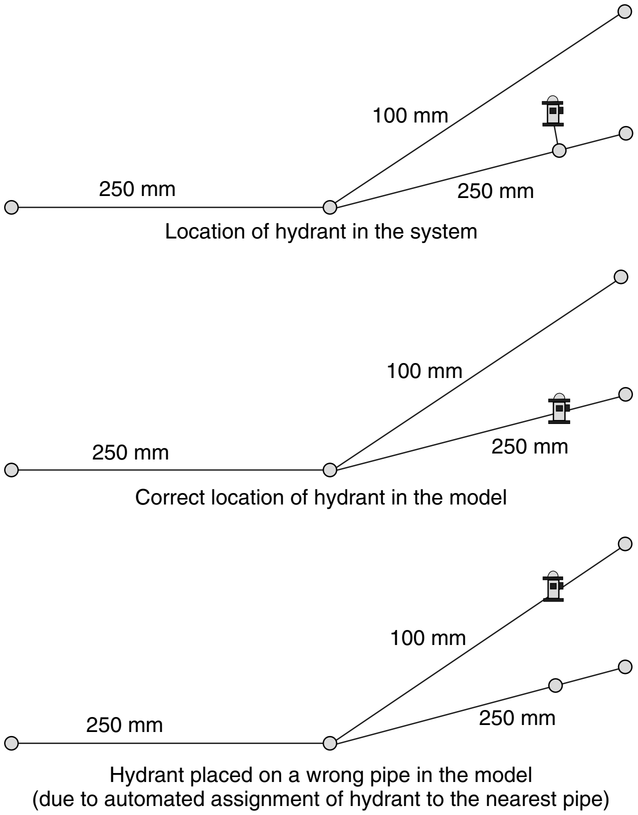

Figure 2. Example depicting misplaced hydrant.

and the observed conditions are within some reasonable degree of correlation (usually less than 20% error) should the final step of micro-level calibration be attempted.

SENSITIVITY ANALYSIS

Before attempting a micro-level calibration, it is helpful to perform a sensitivity analysis of the model in order to help identify the most likely source of model error. This analysis can be accomplished by varying the different model parameters by different amounts and then calculating the associated effect. For example, many current network models have as an analysis option the capability to make multiple simulations in which global adjustment factors can be applied to pipe roughness values or nodal demand values. By examining such results, the user can begin to identify which parameters have the most significant impact on the model results and thereby identify potential parameters for subsequent fine tuning through micro-level calibration.

MICROCALIBRATION

After the model results and the field observations are in reasonable agreement, a micro-level model calibration should be performed. As discussed previously, the two parameters adjusted during this final calibration phase will normally include pipe roughness and nodal demands. In many cases, it may be useful to break the microcalibration into two separate steps: (a) steady-state calibration, and (b) extended-period calibration. In performing a steady-state calibration, the model parameters are adjusted to match pressures and flowrates associated with static observations. The normal source for such data is from fire flow tests. In an extended-period calibration, the model parameters are adjusted to match time-varying pressures and flows as well as tank water level trajectories. In most cases, the steady-state calibration will be more sensitive to changes in pipe roughness, whereas the extended-period calibration will be more sensitive to changes in the distribution of demands. As a result, one potential calibration strategy would be to first fine tune the pipe roughness parameter values using the results from fire flow tests and then try to fine tune the distribution of demands using the flow/pressure/water level telemetry data.

Historically, most attempts at microcalibration have typically employed an empirical or trial and error approach. Such an approach can prove to be extremely time-consuming and frustrating when dealing with most typical water systems. The level of frustration will, of course, depend somewhat on the expertise of the modeler, the size of the system, and the quantity and quality of the field data. Several commercial software packages offer an optimized calibration feature that would ease some of the frustration associated with model calibration. Genetic algorithm (GA)-based techniques for model calibration appear to be popular with the commercial software packages. The optimized calibration techniques obtain an optimal set of decision variables (e.g., pipe roughness values and nodal demands) that minimize the average deviation between measured and calculated pressures and flows while satisfying all system and bound constraints (7). Although one would expect this to be a one-time process, several factors force repeated use of this approach until a more satisfactory set of decision variables are obtained. One of the reasons for repeated use of optimized calibration runs is the way the pipelines are grouped. GA-based approaches solve for optimal group roughness values instead of roughness values for individual pipelines. Alternatively, they can solve for an optimal multiplication factor for the preassigned roughness values for all the pipes in a group. Pipes of similar roughness characteristics (estimated based on material, age, and diameters of the pipes) are treated as one group. However, this is a highly subjective process and therefore offers opportunities for regrouping and reapplying the optimized calibration technique. Similarly, the explicit bound constraints on the pipe group roughness values offer opportunities for relaxing/tightening of the bounds and reapplying the optimized calibration technique. The iterative procedure for model calibration using optimized calibration approach is summarized in the following.

- Setup the model to include calibration data (observed pressures and flow readings).
- Group the pipes based on similar roughness characteristics.
- Estimate the roughness values for each group of pipes.
- Set up an upper and lower bound for each group of pipes based on estimated group roughness values. As

an initial guess, these bounds should be within 20% to 30% of the estimated values.

- Perform optimized calibration.
- Compare measured and calibrated pressure readings.
- If not satisfactory, or if any of the calibrated group roughness values reach the bounds, relax the bounds by a small percentage and rerun optimized calibration.
- Repeat the trials until satisfactory results are obtained.
- If the bounds had to be relaxed beyond a reasonable value for a group of pipes, it would imply a problem with network data and not necessarily the roughness values. Go back to macrocalibration probing the network in and around the area of concern, and repeat macrocalibration, sensitivity analysis, and microcalibration steps.

Although this procedure outlines the iterative method for obtaining optimal set of group roughness values, it can be easily modified when calibrating the model for other parameters such as nodal demands and their spatial and temporal distribution (7)

SUMMARY AND CONCLUSION

Network model calibration should always be performed before any network analysis planning and design study. A seven-step methodology for network model calibration has been proposed. Historically, one of the most difficult steps in the process has been the final adjustment of pipe roughness values and nodal demands through the process of micro-level calibration. With the advent of recent computer technology, it is now possible to achieve good model calibration with a reasonable level of success. As a result, little justification remains for failing to develop good calibrated network models before conducting network analysis. It is expected that future developments and applications of both GIS and SCADA technology will lead to even more efficient tools.

BIBLIOGRAPHY

1. Walski, T.M. (1995). Standards for model calibration. *Proc. AWWA Computer Conf.*, Norfolk, VA, pp. 55–64.
2. Pipe2000. (2004). Users Guide, Pipe2000 Series Models GUI—KYPIPE Hydraulic Model. KYPipe LLC, Lexington, KY.
3. Cesario, L. (1995). *Modeling, Analysis and Design of Water Distribution Systems*. American Water Works Association, Denver, CO.
4. McEnroe, B., Chase, D., and Sharp, W. (1989). Field testing water mains to determine carrying capacity. *Technical Paper EL-89*, Environmental Laboratory of the Army Corps of Engineers Waterways Experiment Station, Vicksburg, MS.
5. Walski, T.M. (1990). Sherlock Holmes meets Hardy Cross, or model calibration in Austin, Texas. *J. AWWA*. **82**(3): 34.
6. Cruickshank, J.R. and Long, S.J. (1992). Calibrating Computer Model of Distribution Systems. *Proc. AWWA Computer Conf.*, Nashville, TN.
7. Lingireddy, S. and Ormsbee, L.E. (2002). Hydraulic network calibration using genetic optimization. *Civil Environ. Engineer. Syst.* **19**(1): 13–39.

NUMERICAL MODELING OF CURRENTS

Rudi Rajar
University of Ljubljana
Ljubljana, Slovenia

INTRODUCTION

Description of Currents

Currents—water movement—appear in all surface waters: rivers, lakes, coastal seas, and oceans. This flow of water is almost always turbulent, which means with many irregular eddies of different size. Currents are sometimes near to steady movement (not changing with time), but mostly the movement is unsteady, significantly changing with time.

In rivers the currents are more aligned nearly in one dimension (1D), although in larger rivers they can be quite complex and three-dimensional (3D). In shallower rivers or coastal seas the modeling of currents is sometimes two-dimensional (2D), but in reality *all* currents are three-dimensional (3D); that is, the water velocities are in all three directions, although often they are an order of magnitude greater in the two horizontal directions than in the vertical direction.

In surface waters the current circulation has different orders of magnitude.

Global currents are large-scale currents, typical for every world ocean. They are often quasisteady in long-term average, and with their enormous transport of heat from the equator to the poles, and vice versa, they significantly affect climate and weather on the continents. The main driving forces are the temperature gradients between the poles and the equator and the wind. Due to the earth's rotation, the main pattern of global currents is typically from the poles toward the equator on the west side of continents (cold currents), and from the equator to the poles on the east side of continents (warm currents). Some of these global currents are the cold Peru (or Humboldt) current along the west coast of South America (toward the equator), the warm Kuroshio current from the equator along the east coast of Asia, and the warm Gulf Stream along the east coast of North America. The last is deflected by the topography of the coastline toward northern Europe, causing a relatively mild climate along the Scandinavian coastline.

Local currents (mesoscale and microscale currents) are currents in smaller regions, as in coastal (or shelf) seas, which are defined as shallower than about 200 m, to distinguish them from the oceans. Of course, local currents at a smaller scale appear also in lakes and rivers.

Importance of Current Circulation

Numerical modeling of currents usually includes modeling of additional parameters that are important in water ecology. When solving practical problems, the following parameters must be determined: (1) water current (circulation), (2) transport of sediments with currents, (3) transport of any contaminants, and (4) biochemical processes of contaminants.

Water current circulation is always the basis for calculation of the transport of any contaminant. But in many cases, knowledge of the current velocities is used directly, for example, in designing river or marine coastal protection, where the computation of current patterns helps determine bank and riverbed erosion and the optimum form of the protection structures. Transport of sediment is also important from two points of view. In river reservoirs, lakes, or coastal bays, sedimentation is often a big problem. Over the years, river reservoirs upstream of dams are filled with sediments and thus lose their useful storage volume. Even more important is the fact that most contaminants are bound to suspended sediments and often over 90% of the contaminant transport is the result of sediment transport with less than 10% in dissolved form. Determining the transport by currents of any contaminants in surface waters is very important. These can be nutrients, affecting the growth of algae and possibly causing eutrophication, or any toxic chemical. Coastal seas often receive large amounts of contaminants from rivers. Many coastal settlements discharge their wastewater (sometimes untreated) into coastal seas. Determination of currents and contaminant transport—most often by numerical modeling—is essential to determine consequences of pollution and to plan the optimum location of wastewater outlets. A special problem is oil spill, where numerical modeling of the spreading oil slick can help in clean up or can help to find the culprit of the pollution. Heat can also be considered as a kind of contaminant. Coastal thermal or nuclear power plants dispose of their waste heat into the sea. As the water temperature should not increase over a certain limit (usually 3 °C over the "natural" temperature), these power plants should be located at sites where strong currents transport away and mix the waste heat. Because contaminants are not only transported and spread (dispersed) over the water body but are also prone to different transformations due to biochemical processes, another quite difficult task of numerical models is to simulate these processes.

The importance of global currents is crucial. As already stated, global currents significantly affect climate and weather on the continents. The El Niño phenomenon is well known to cause catastrophic changes of weather over large parts of continents. Global currents are also important carriers of nutrients and contaminants of all types, and they influence the fish migration. Many oceanographic studies are carried out to determine details of global currents, their changes (e.g., prediction of the effect of future global warming), and their important influence on the environment.

Methods of Determination of Currents

To determine currents in surface waters, there are three methods: numerical modeling, physical modeling, and field measurements. Numerical modeling is the most appropriate methodology. However, most often it should be combined with field measurements and sometimes with physical modeling.

Physical models were widely used before the era of fast computers. Studied regions were built at a reduced scale in large hydraulic laboratories, and water was pumped to the models. Current velocities and other important parameters were then measured and used to model nature, using physical laws of similitude. However, there are several difficulties in applying this methodology: physical models are expensive and time consuming to build and operate; it is practically impossible to simulate very large regions; and it is almost impossible to reliably simulate biochemical processes.

Field measurements are becoming very important. As numerical models almost always demand calibration and verification of the simulated results, reliable data for comparison can only be obtained by field measurements. This technology is developing very fast. Remote sensing (e.g., satellite images) can provide the pattern of sea surface temperature, or chlorophyll concentration over very large regions. As these patterns depend on transport by currents, this methodology can be used for verification and calibration of numerical models. *In situ* measurements (directly by ships, buoys) are also necessary in combination with the remote sensing technique.

In fact, reliable simulation of currents and contaminant transport can only be attained by the combined application of numerical models and field measurements.

NUMERICAL MODELS

What Is a Numerical Model?

A simplified definition of numerical model is the following: a computer code that can describe or "simulate" a certain physical phenomenon, in this case current circulation and contaminant transport in surface water bodies. The simulation is never perfect, as the real physical phenomenon is too complex, and too many parameters and processes are involved. However, most often, the simulations are accurate enough to essentially help at solving different scientific and engineering problems.

Every numerical model basically consists of three parts: (1) basic equations, describing the physical phenomena; (2) all the necessary input data (e.g., description of topography, data on physical properties of water and contaminants); (3) "forcing factors"—the data about forces that cause water and sediment movement and sources of contamination. The modelers call these "boundary conditions," because these data determine the behavior of the model and essentially influence the results.

A numerical model can be compared to a physical model: in both cases the form (topography) of the water body must be given (for numerical model in digital form), as well as the physical properties of the water and boundary conditions. The final results are the same in both cases. The result of a numerical model can even be presented as an animation, showing, for example, water circulation—similar to a movie taken on a physical model.

Dimensions of Numerical Models of Currents

One-dimensional models simulate the flow in one dimension only—for example flow along a river. Topography is given with cross sections; the distance between them depends on how big the river is and on the required accuracy. For unsteady flow, the result is time-dependent depth and velocity in each cross section. Contaminant concentration can also be calculated in every cross-section and at every time step.

Two-dimensional models are usually used in shallow coastal seas or lakes. The computational region (in the horizontal plane) is divided into control volumes of rectangular or other polygonal form (possibly also curvilinear) and bottom and bank topography is given as input data. There is only one control volume from surface to the bottom. The result of simulation is depth and two components of horizontal velocity and possibly contaminant concentration in every control volume and at every time step.

Three-dimensional models are similar to 2D models: the difference is that the computational region is additionally divided into layers along the depth. In this way, the whole computational domain is divided into "bricks" in all three dimensions. Besides the parameters mentioned with 2D models, the vertical velocity component is also computed.

Three-dimensional models are the most sophisticated and basically the most accurate. However, the real accuracy and reliability of the models depend on the accuracy of the input data and on the calibration, verification, and validation of the models.

Forcing Factors

The water movement or, current is caused and influenced by four "forcing factors."

Wind has an important effect on the water surface. With its shear stress it causes the surface layer of water to move basically in the direction of the blowing wind. Due to the law of continuity, this movement also causes movement of other water layers. If the phenomenon is taking place in a closed basin, layers near the bottom must move in the opposite direction to the wind. In general, the movement is complex and depends on the bottom and bank topography. In larger seas and oceans, the earth's rotation essentially influences the current pattern (1). A good numerical model must take into account all these influences.

Tide influences the current pattern. When simulating currents inside a semienclosed bay or coastal sea, the boundary condition for the model is represented by the time variation of tidal level at the open boundaries.

Inflow of rivers also affects the currents, with its inflow of mass and momentum.

Density gradients—variations of density with space and/or time—also cause water circulation. The density changes are most often caused by gradients of temperature and salinity; therefore, this type of forcing is also called "thermohaline forcing."

The relative importance of the four forcing factors can vary with different cases. Wind and river inflow are usually the most important forcing factors in lakes and smaller coastal seas, while tide is the most important factor in small and middle-sized coastal areas. Density gradients and winds are very important for ocean circulation. A good numerical model should include all four forcing factors, although in many cases some less important ones can be neglected entirely.

Numerical Methods

The water movement or current is described by a system of differential equations, basically a scalar continuity equation (mass conservation) and a vector momentum equation, which has as many scalar components as the dimension of the model. Because these equations are nonlinear, the solution is very complex. Basically two groups of numerical methods are used for their solution: finite difference methods and finite element methods. With the first group of methods, the computational domain is divided into control volumes of polygonal forms (possibly with curvilinear sides), while with finite element methods, the domain is divided into different unequal geometric bodies. The finite difference methods are a bit simpler at transforming the differential equations into finite difference forms, while the finite element method can better adjust the control volumes to irregularly shaped boundaries. In the simulation of currents, the finite difference models are more often used.

The number of control volumes into which the computational domain is divided depends on the accuracy of the input data and on the required accuracy of the results. Typically, the domain is divided into 100–150 cells in each horizontal direction, and into 15–30 layers along the depth. Computing time is long and fast computers are needed for the computation of currents.

Verification, Calibration, and Validation of Models

For each simulated case, the applied model has to be checked for reliability.

Verification of a model is the first test, which shows whether the phenomenon is correctly simulated. Then the model should be ***calibrated***. Most often there are a few parameters in the model simulations for which the proper values are not accurately known, so a series of simulations are carried out with variable values of these parameters. The simulation results are compared with results of measurements (or possibly results of another model). In this way the proper values of the parameters are determined. (Example parameters are diffusion coefficients for turbulence or dispersion of contaminants.) When the simulated results are in acceptable agreement with measurements, the model is "verified" and "calibrated." It is then used to simulate other similar cases, for which the model parameters were not calibrated. If the results are in agreement with measurements, the model is "validated."

CASE STUDIES

A: Transport of Radioactive Contaminants in Japan Sea

The Japan Sea is a large, almost enclosed sea between the Japanese islands, Sakhalin Island, the Siberian coast, and Korea. Its length is about 1800 km, its width 900 km (see Fig. 1), and the maximum depth is about 3900 m. The relatively weak exchange of water with the Pacific Ocean is achieved through four straits (Fig. 1), the widest being Tsushima Strait between Korea and the Japanese islands.

A large amount of radioactive (RA) waste has been deposited in the Japan Sea, off Vladivostok, at a depth

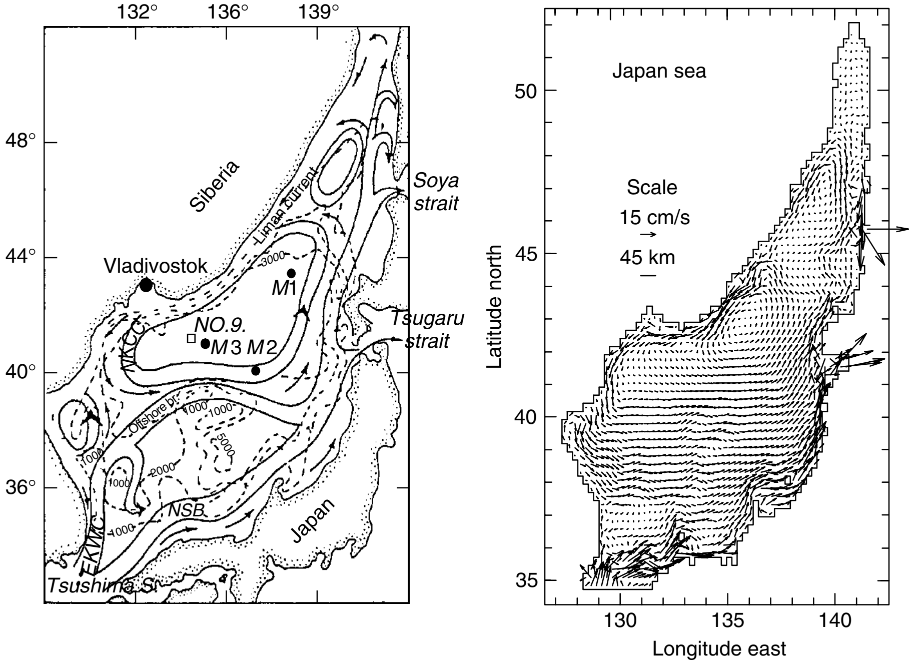

Figure 1. Measured surface currents in the Japan Sea (left) and surface currents simulated by the PCFLOW3D model (right). (From Ref. 3.).

of about 3500 m. Concern existed in Japan that leaking containers with RA waste could cause contamination of fish and other marine organisms in the Japan Sea. A research model should answer the questions: (1) In what concentrations and after what time would the RA contaminants reach the surface layers? (2) Where might they be transported by currents and dispersion? In the simulations, the possible leaking of the radionuclides was supposed to be at location No. 9 (Fig.1).

The first task was to determine the current circulation (4). A 3D numerical model PCFLOW3D was used (2,3). The 3D circulation was driven by three forcing factors: thermohaline forcing, wind, and the inflow/outflow surface currents. The computations were carried out taking into account the measured seasonally averaged temperature and salinity distribution and seasonal winds. A uniform wind field was accounted for (9.5 m/s in winter, 0 m/s in summer). The inflow current through the Tsushima Strait (a branch of the global Kuroshio current) was taken into account.

After some calibration, the results of the model were compared with the observed pattern of the surface currents (Fig.1). It can be seen that agreement is good; the simulated spirals have the same direction and strength as the measured ones. Some measurements also confirmed proper simulation of bottom currents.

On the basis of the computed winter and summer hydrodynamic circulation, transport and dispersion of radioactive contaminants were further simulated. The calculated amount of radioactive contaminants released in 90 d was 1 TBq (10^{12} becquerels), from dumping location No. 9 (Fig. 1).

The final results of the simulations have shown that the RA contaminants would reach the surface mainly in the northern part of the Japan Sea. The first contaminants would reach the surface layers (above 180 m depth, where fishing is possible) after 3 years; the amount would increase in time and would attain a maximum after about 30 years. But the maximum concentrations would be on the order 1E-2 Bq/m^3, which is far below the natural background values and hence would not represent any danger for the environment.

B: Simulation of Mercury Cycling in the Mediterranean Sea

Mercury is a highly toxic metal, especially in its organic form, monomethylmercury. An international project carried out in 2002–2004 had the goal to determine the natural and anthropogenic sources of mercury in the Mediterranean region and to determine the distribution of the most important mercury forms by the use of a three-dimensional numerical model (5).

Again, the current circulation had to be determined first. It was calculated for four annual seasons. Forcing factors were the following: (1) thermohaline forcing, that is, density gradients due to temperature and salinity distribution (obtained by measurements) in the whole computational domain and also along the depth; (2) wind, and (3) inflow from rivers and exchange of water through the Gibraltar and Bosporus. Verification and calibration of the model were carried out on the basis of numerous measurements. Figure 2 shows a comparison of simulated and measured currents for the summer season.

Transport and dispersion of mercury were then simulated, sources arriving from rivers, through straits from adjacent seas, and from some known point sources of contamination. Verification and calibration of the transport–dispersion part of the model were carried out on the basis of numerous measurements of relevant parameters. Three cruises by a large research ship *Urania* over the whole Mediterranean Sea provided measurements. Figure 3 shows the distribution of mercury, simulated over one year of inflow from the sources of contamination.

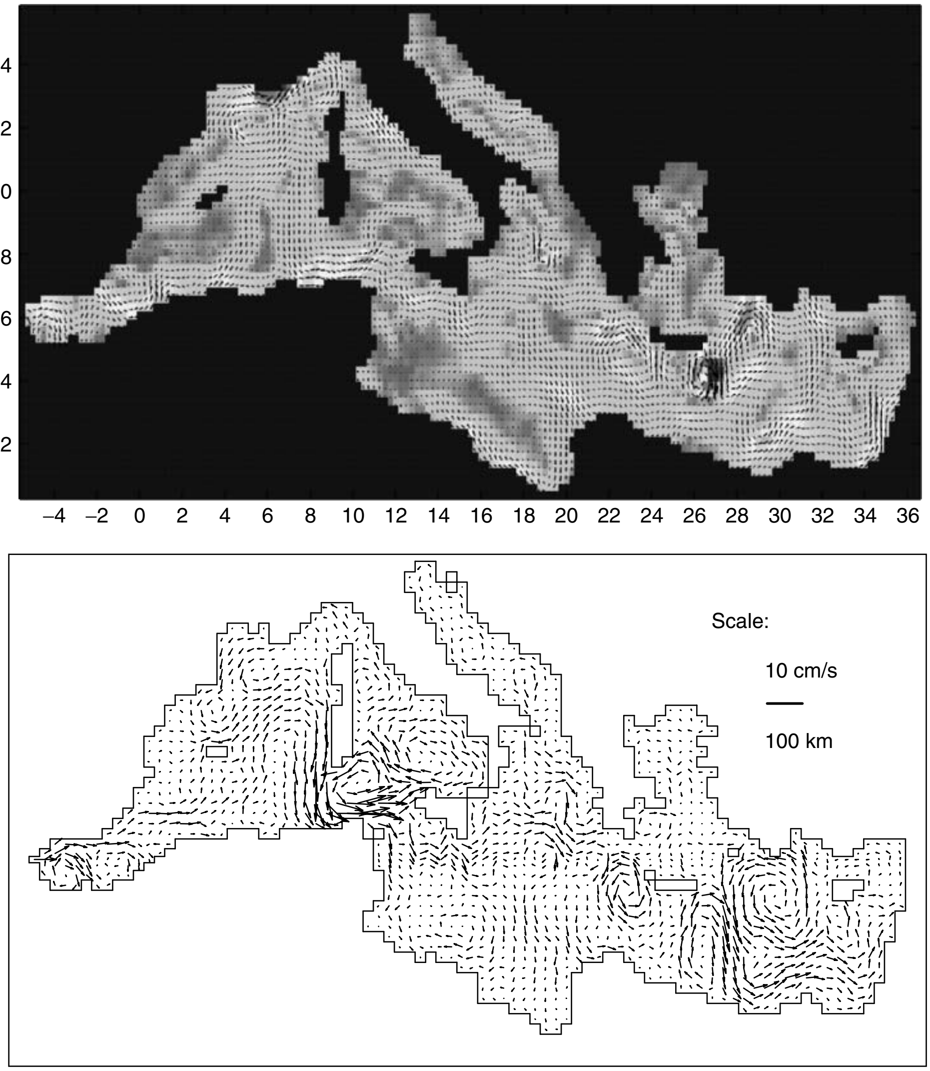

Figure 2. (Top) Measured average currents in the Mediterranean in the summer season. (Bottom). Simulated currents for the same season.

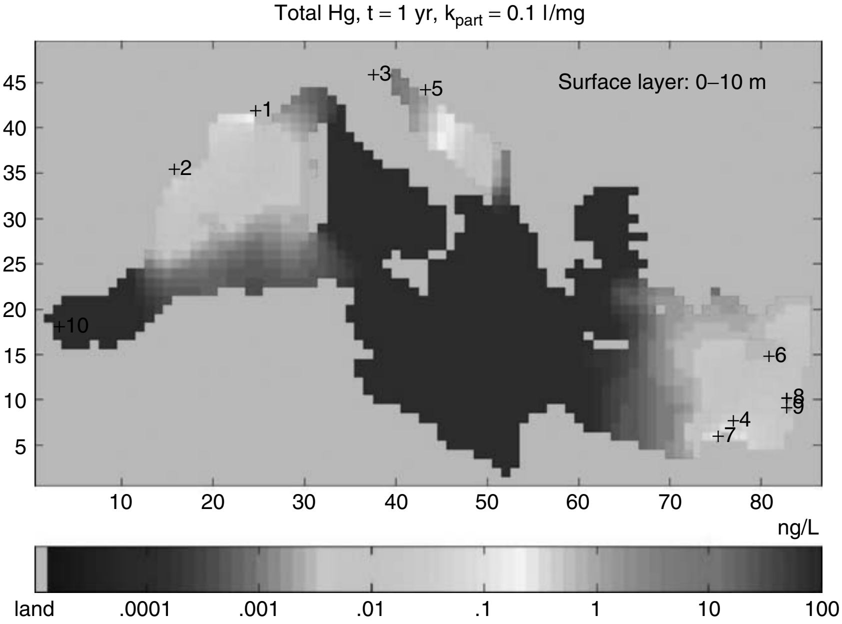

Figure 3. Concentration of total mercury in the surface layer of the Mediterranean Sea, after 1 year of simulation by PCFLOW3D model. Sources of contamination are from rivers and point sources, marked by numbers.

One of the goals of the project was also to couple the water model with the atmospheric model (which also simulates circulation and mercury transport) into a general integrated model. As there is permanent exchange of mercury between the water surface and the atmosphere, the integrated model gives a more realistic result about global mercury cycling and contamination in the region. The general model was finally used as a tool to predict possible measures to diminish concentration of toxic mercury in the Mediterranean region.

BIBLIOGRAPHY

1. Duxbury, A.C. and Duxbury A.B. (1989). *An Introduction to the World's Oceans*. Wm. C. Brown Publishers, Dubuque, IA.
2. Rajar, R. and Četina, M. (1997). Hydrodynamic and water quality modelling: an experience. *Ecol. Model.* **101**: 195–207.
3. Četina, M., Rajar, R., and Povinec, P. (2000). Modelling of circulation and dispersion of radioactive pollutants in the Japan Sea. *Ocean. Acta* **23**(7): 819–836.
4. Yoon, J.H. (1995). Numerical modelling of the ocean miniature "Japan Sea." In: *Proceedings of the International Workshop on Numerical Prediction of Ocean Variations*, Tokyo, March.
5. Rajar, R., Četina, M., Žagar, D., and Petkovšek, G. (2004). Three-dimensional simulation of mercury cycling in the Mediterranean Sea. In: *International Conference on Mercury as a Global Pollutant,* Ljubljana, June, Part 2, pp. 1324–1327.

UNCERTAINTY ANALYSIS IN WATERSHED MODELING

Thorsten Wagener
The Pennsylvania State University
University Park, Pennsylvania

INTRODUCTION

All watershed modeling studies unavoidably contain uncertain elements. This uncertainty is propagated into the model predictions and impacts subsequent decision-making. If not taken into account properly, this can lead to model predictions that are assumed to be more reliable than they actually are. The usefulness of these predictions in any decision-making context therefore becomes questionable.

A watershed model is the numerical representation of our perception of the hydrologic system at hand. It consists of a model structure (i.e., underlying equations), model parameters (i.e., system invariants), and model states. The model is driven by measurements of variables entering the watershed (e.g., precipitation and energy) and compared to observations of the system response (e.g., streamflow or groundwater heads).

SOURCES OF UNCERTAINTY

Sources of uncertainty in this context are manifold and basically include every element that is needed for the modeling of a watershed as described above. The main uncertainties are listed and described here. The first uncertainty present when modeling watershed systems originates from the modelers' limited understanding of the real world (i.e., uncertainty in the perceptual model used).

Perceptual Model (PM) Uncertainty. A PM (1) is a conceptual representation of the watershed that is subsequently translated into mathematical (numerical) form in the model. The PM is therefore based on our understanding of the real-world watershed system. This understanding might be low, particularly with respect to subsurface system characteristics and water flowpaths, and therefore our PM might be uncertain.

The next group of uncertainties is related to the mathematical/numerical model and its components. These components and the related uncertainties are:

Model Structure (Equation) Uncertainty. The ambiguity present in the PM usually means that a variety of model structures seem to be feasible representations of the watershed at hand prior to any modeling being undertaken. Unfortunately, analysis of the performance of the selected model structures in reproducing observed watershed behavior is often insufficient to distinguish between them and some ambiguity about an appropriate structural system representation remains.

Model Parameter Uncertainty. Models of watershed hydrology necessarily aggregate the natural heterogeneity of watershed characteristics and processes into individual model components. These components are characterized by model parameters that describe, for example, storage size, water residence time, or flowpath. The scale at which these parameters can be measured is commonly different from the scale at which they are applied within the model. At least some of the model parameters therefore have to be estimated through a process of parameter adjustment (manually or automatically), where the parameters are varied until observed and modeled system behavior agree as much as possible. Most watershed models contain a large number of nonlinearly interacting parameters, which makes this a very difficult task, and some uncertainty commonly remains with respect to the question of whether a unique best parameter set has been found.

Model State Uncertainty. With respect to state uncertainty one has to distinguish between initial state uncertainty and state uncertainty after a certain time series of data has been processed by the model. The first uncertainty can usually be taken care of by including a warm-up period that is excluded from the performance analysis. Alternatively, the initial state can be optimized, similarly to the parameter adjustment process described above. State uncertainty during model simulation stems from biases in the model predictions. It can be reduced using data assimilation techniques, that is, approaches that recursively assimilate measurements of the model output variable from previous time steps to adjust the model states to reduce predictive uncertainty.

Variable Uncertainty. This is another, less often considered uncertainty, originating from the problem of simulated and observed variables not being exactly the same quantity (2). For example, predictions and measurements of soil moisture might be at different scales and therefore a perfect match should not be expected or is not even something the modeler should strive for (3).

Input (Forcing) and System Response Data Uncertainty. There are also uncertainties that are related to the measurements of forcing and response variables. These stem, for example, from an inability to capture the spatial distribution of rainfall or from problems in capturing the true streamflow discharge. These uncertainties can only be reduced outside the modeling process, for example, through better measurement techniques.

Generally, no satisfactory means of considering all these uncertainties simultaneously in a watershed modeling exercise have yet been developed (3). Additionally, the modeler is faced with the problem that, in many cases, the distance between the (erroneous) observation and the uncertain prediction of streamflow is the only measure of deviation (error) at each time step, providing very little information for model evaluation.

APPROACHES TO UNCERTAINTY ANALYSIS

A wide variety of techniques for uncertainty analysis in watershed hydrology are available and often applied (3,4). These methods differ, for example, in underlying philosophy and assumptions and in sampling strategies applied. Little guidance is available to help modelers decide what approach might be the most suitable one for their situation (1). While the differences are large, there is some overlap between all the approaches. In principle, all techniques for uncertainty analysis can be reduced to three basic steps. The main potential approaches that are applied for uncertainty analysis in watershed hydrology are described based on this three-step approach.

1. *Define What Constitutes a Behavioral Model*. The definition of what constitutes a behavioral model can roughly be divided into two types of definitions, those that follow a strict statistical approach, and those using set-theoretic measures not consistent with traditional statistics. The former are based on assumptions about the structure of the model residuals, that is, the differences between observed and simulated system response. Typical assumptions are that the error is Gaussian and has a constant variance in time (3). The second type is less bound by a formal statistical framework, but this is commonly defended with the statement that the structure of the uncertainty is unknown in most modeling studies and that restricting assumptions are therefore not justifiable. Examples of the second type include the generalized likelihood uncertainty estimation (GLUE) (2) framework and fuzzy measures.
2. *Find All Models that Comply With This Definition in the Feasible Parameter Space*. The search for all models that are consistent with the definition consists of two parts. First is the search for all parameter sets within a certain model structure that provide behavioral predictions. This search is commonly performed using various types of random sampling (e.g., uniform random sampling or types of stratified sampling) or using population evolution based techniques. The use of Monte Carlo Markov chains can be very efficient in this context if the assumptions that have to be made about the shape of the response surface are justified.

   The second part of the search is related to the problem that potentially different model structures can provide behavioral predictions for a certain case. This is an even more difficult task since the model space is not just infinite, but also unbounded. No systematic approach exists to explore this search space. Approaches suggested include the use of ensembles of model structure (e.g., in a Bayesian framework) (5) or the inclusion of a more complex error model that takes the imperfection of the selected model structure into account (6).
3. *Propagate the Predictions of All Identified Models into the Output Space While Considering Other Uncertainties (e.g., in the Forcing Data)*. Three basic approaches are applied for the propagation stage, first-order, statistical, and set-theoretic. The first approach only propagates the mean and the variance of the predicted variable, based on means and variances of inputs and parameters. An overview of the most popular first-order approaches is provided by Melching (7). Statistical approaches sample from the posterior parameter distributions derived in step 2 and usually make assumptions about the distributions of input uncertainties to derive ensembles of predictions. Ensembles of inputs and/or parameter sets, not necessarily sampled from strict statistical distributions, are also propagated in the set-theoretic approach.

CONCLUSION

Uncertainty is an unavoidable element in any watershed modeling exercise. It has to be considered if the model simulations should be used in a decision-making context. Main uncertainties include those stemming from a lack of understanding of the hydrologic system, the problem of identifying appropriate model structure and parameters, and uncertainties in the measured data. While many approaches exist that can be used for uncertainty analysis, little guidance is available to choose the best approach for a certain study. A three-step structure is common to all approaches and potential implementations are discussed above. Future developments in this area will focus on a more realistic representation of the uncertainties present, techniques that allow for the explicit and simultaneous consideration of all uncertainties, and better sampling strategies.

BIBLIOGRAPHY

1. Beven, K. (2004). Does an interagency meeting in Washington imply uncertainty? *Hydrol. Proc.* **18**: 1747–1750.

2. Beven, K.J. (2000). *Rainfall–Runoff Modeling: The Primer*. John Wiley, Chichester, UK.
3. Gupta, H.V., Beven, K.J., and Wagener, T. (2004). Model calibration and uncertainty estimation. In: *Encyclopedia of Hydrological Sciences*. M. Anderson (Ed.). John Wiley, Chichester, UK.
4. Duan, Q., Sorooshian, S., Gupta, H.V., Rousseau, A., and Turcotte, R. (Eds.). (2003). *Advances in Calibration of Watershed Models*. AGU Monograph, pp. 29–47.
5. Newman, S.P. (2002). Accounting for conceptual model uncertainty via maximum likelihood Bayesian model averaging. *Acta Univ. Carolinae Geol.* **46**(2/3): 529–534.
6. Kennedy, M.C. and O'Hagan, A. (2001). Bayesian calibration of computer models. *J. R. Statist. Soc.* **63**(3): 425–464.
7. Melching, C.S. (1995). Reliability estimation. In: *Computer Models of Watershed Hydrology*. V.P. Singh (Ed.). Water Resources Publishers, Washington, DC, pp. 69–118.

WATERSHED MODELING

Thorsten Wagener
The Pennsylvania State University
University Park, Pennsylvania

INTRODUCTION

Models of watershed hydrology are important tools in hydrological investigations for operational and research purposes. Areas of application are the extension of streamflow time-series in space (1–3) and time (4,5), evaluation of a watershed response to climate (6,9) and/or land-use variability and change (8,9), calculation of design floods (10,11), as load models linked to water quality investigations (12), real-time flood forecasting (13–15), and the provision of boundary conditions for atmospheric circulation models (16,17). This multitude of, often very specific, tasks and circumstances has contributed to the development of a vast number of watershed models starting in the early 1960s (see Todini (18), for a historical review). These models are usually a mixture of linear and nonlinear functions, combined to represent those processes occurring in the specific watershed and important for the study objectives at hand. Watershed models are also used in water quality studies, but the focus here is simulation of the quantitative response and general features of these models.

WATERSHED PROCESSES

The hydrologic cycle is a closed system in the sense that no water is lost. Only part of the processes that constitute the hydrologic cycle are relevant for watershed modeling. A watershed is therefore an open system with respect to inputs (mainly precipitation and energy) and outputs (mainly streamflow, but also surface and subsurface losses). A schematic description of the main processes in a typical watershed system is shown in Fig. 1. Models of watershed hydrology represent these processes using different mathematical approaches and different levels of spatial detail. It is important to consider that the part of the watershed system reproduced in the model usually depends on the prevailing processes and also on the modeling purpose and available data (3). For example,

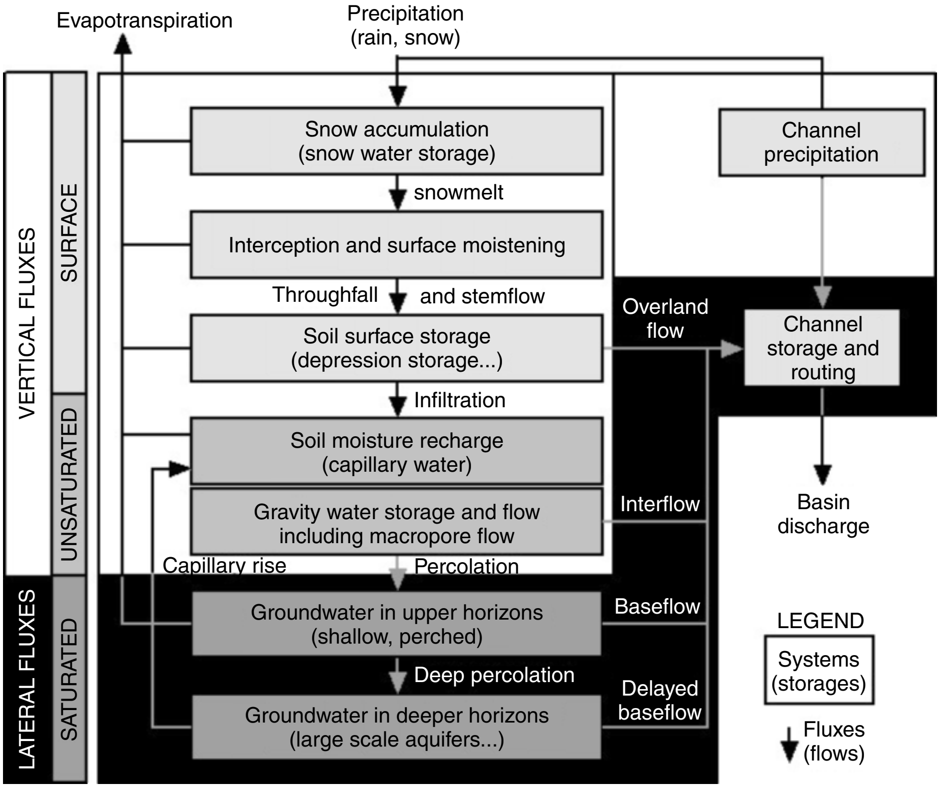

Figure 1. Schematic of watershed processes (after Reference 19).

routing processes are usually not required if the modeling purpose is simulation of the watershed outflow at a monthly time-step.

WATERSHED MODEL CLASSIFICATION

It is common to classify watershed models into three distinct types (20; Fig. 2): (1) *empirical* (also called data-based, metric, or black-box), (2) *conceptual* (also called parametric, explicit soil moisture accounting, or gray-box) and (3) *physically based* (also called mechanistic or white-box) model structures. There are, however, models that have characteristics of more than one group.

Empirical models use available time series of input and output variables (precipitation, streamflow, temperature, etc.) to derive both the model structure and the corresponding parameter values. They are therefore purely based on the information retrieved from the data and generally do not include prior knowledge about catchment behavior and flow processes, hence, the name black box. Popular examples of empirical models are artificial neural networks (ANN; 21,22) and transfer functions (TF; 23). Empirical models are usually spatially lumped; they treat the catchment as a single unit. Variants on purely empirical TF models are *data-based mechanistic models*. They constrain the degrees of freedom to those structures that are physically interpretable, thereby using the hydrologist's understanding of the natural system (24–26).

Conceptual models (3,27) are built from storage elements that represent parts of the watershed where water is temporarily stored, such as soil, aquifers, or streams. These elements are filled through fluxes such as rainfall, infiltration, or percolation, and empty through evapotranspiration, runoff, drainage, etc. Based on a conceptualization of the real-world watershed, the modeler, specifies the structure of these models a priori, for example, the number of and connections between storage elements. These models are also usually spatially lumped. This means that large parts of the (heterogeneous) watershed are integrated in a single (homogeneous) element. The consequence of this process is that the model parameters lose some of their physical meaning and cannot be measured on the required scale. Rather, the modeler has to use observations of the watershed response to find appropriate values for the model parameters (see next section). Some models are based on a semidistributed approach, in which different models represent individual subwatersheds. This approach is finding increased attention in the literature (28). Conceptual models form the large majority of models used in operational hydrology. Their dependence on measurements of the watershed response (mainly streamflow) limits their use in ungauged watersheds. However, research is ongoing to resolve this problem (3,29).

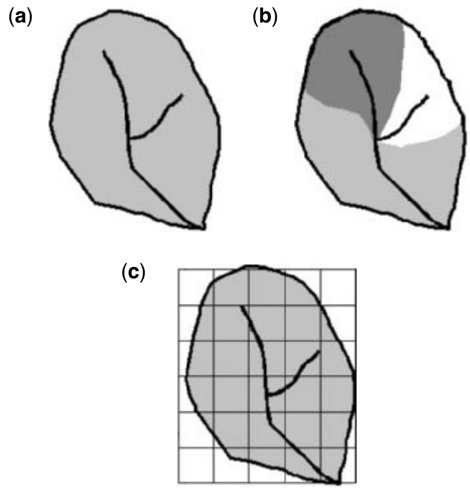

Figure 2. Classification of watershed models according to spatial discretization used: (**a**) lumped, (**b**) semidistributed, and (**c**) distributed.

Physical models (30,31) are based on conservation of mass, momentum, and energy. The spatial discretization applied is usually based on grids, but sometimes also on some type of hydrologic response unit or triangular irregular networks (32). Physical models are therefore particularly attractive for studies that require a high level of spatial detail, such as studies on soil erosion or diffuse-source pollution (33). The initial idea underlying these models was that the degree of physical realism on which these models are based would be sufficient to relate their parameters, such as soil moisture characteristics and unsaturated zone hydraulic conductivity functions for subsurface flow or friction coefficients for surface flow, to physical characteristics of the catchment (18), thus eliminating the need for observed system response to condition the parameters of the model. However, the currently available physical models do not fulfill this ideal. They suffer from extreme data demand, scale-related problems (e.g., the measurement scales are different from the process and model (parameter) scales), and overparameterization (34). A number of key parameters—applied to a large number of elements—still has to be estimated from measurements to capture the uniqueness of a specific watershed (35). The expectation that these models could be applied to ungauged catchments has therefore not been fulfilled (36). Beven (34) argues that this type of model is applied in a way similar to lumped conceptual models, though on a different scale.

A general characteristic of all current models is that they do not provide reliable predictions in ungauged basins (3). The general behavior of a watershed is reasonably constrained by the model structure (i.e., the constituent equations), but some adjustment of model parameters is required to capture the unique characteristics of the system under investigation. A large number of popular models of watershed hydrology can be found in Singh (37) and Singh and Frevert (38,39).

WATERSHED MODELING PROCEDURES

Many authors have suggested procedures for using watershed models. Examples are given in Beven (32) and in Anderson and Burt (40). They vary slightly in the steps included and definitions used. A procedure using commonly found elements is shown in Fig. 3 (after

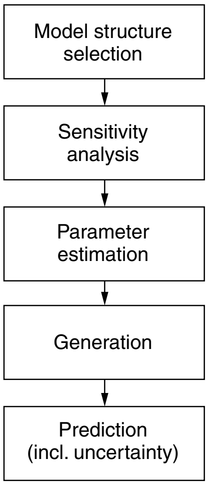

Figure 3. Typical elements of a watershed modeling procedure.

Reference 3). The elements considered are model structure selection, sensitivity analysis, parameter estimation, generalization, and prediction (including uncertainty estimation).

The first step for the modeler is to *select* (or develop) *a model structure*, a combination of mathematical equations, that describes the watershed system as a function of the system characteristics (e.g., saturation or infiltration excess dominated runoff production), the modeling objective (e.g., do the effects of routing have to be considered if only the monthly response is wanted?), and the available data (e.g., a simple evapotranspiration function has to be used if measurements of variables such as humidity or wind speed are not available). Personal preferences and experience in using a particular model structure might also play an important role in the selection process. Woolhiser and Brakensiek (41, p. 15) conclude that "*objective methods of choosing the best model* (structure) *have not yet been developed, so this choice remains part of the art of hydrological modelling*." However, future modeling tools might include knowledge-based approaches to support the decision-making process of the hydrologist for selecting an appropriate model structure.

Sensitivity analysis tests the sensitivity of the model output to changes in the model parameters (or even to uncertainty in the model input or in other model elements). This can be done before parameter estimation to find parameters that are particularly important for model performance—the estimation process can then be focused on these, or it can be done after the calibration as part of the uncertainty analysis—to identify parameters that are particularly uncertain. The simplest approach is a perturbation technique where one parameter is varied over its feasible range while the remaining parameters are kept at reasonable values, although more sophisticated approaches exist (32).

As stated before, the current generation of watershed models generally requires some sort of *parameter estimation* to enable the model to match observed and calculated system behavior, for example, to give reliable predictions. This process is sometimes referred to as calibration. Within this step, the performance of selected parameter sets in reproducing the observed system behavior is evaluated (Fig. 4). either by visual inspection or by calculating some function F that summarizes how well simulated and observed system behavior match:

$$F = f(e) \tag{1}$$

where f is some aggregation function applied to e, the vector of residuals calculated as

$$e_t = z_t - y_t \tag{2}$$

where z is the observed system response, for example, a time series of streamflow; y is the calculated system response; and t is the time step. The modeler attempts to make the function F as small (or large depending on the chosen formulation) as possible, either by manually adjusting the parameters or through some automatic procedure. See Duan et al. (42) for a state-of-the-art review of current techniques.

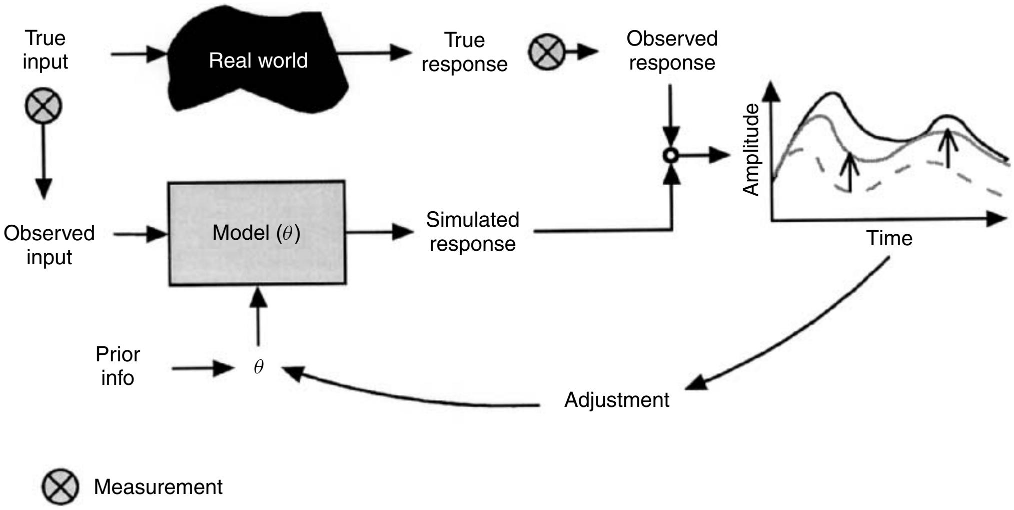

Figure 4. Calibration of a watershed model. The symbol θ represents a parameter vector, y is the system output (usually streamflow), and t denotes time.

The problem with this step of conditioning the parameters on observations is that it can result in parameter values that depend on the specific time series used (e.g., a particular climate period). Some modelers include a *generalization* stage to avoid this. During this stage, the model (including the parameter set found most representative of the system at hand) is applied to another time series of measured system response, not used during parameter estimation, but usually of the same variable. A model is assumed generally applicable if its performance does not degrade significantly during this process (43).

The model can then be used for *prediction*. However, there is a wide variety of factors that introduce *uncertainty* into the modeling process. These are for example, the inability of the modeler to find a "best" parameter set, uncertainties in measuring input and output data, and problems in finding an appropriate model structure to represent the system under investigation. These uncertainties have to be considered and incorporated in the model predictions. It is therefore sensible to state these uncertainties explicitly and include confidence limits that denote the ranges in which the predictions of the model are most likely to fall (44,45).

BIBLIOGRAPHY

1. Manley, R.E. (1978). Simulation of flows in ungauged basins. *Hydrological Sci. Bull.* **23**(1): 85–101.
2. Sefton, C.E.M. and Howarth, S.M. (1998). Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. *J. Hydrol.* **211**: 1–16.
3. Wagener, T., Wheater, H.S., and Gupta, H.V. (in press). *Rainfall-Runoff Modelling in Gauged and Ungauged Catchments*. Imperial College Press/World Scientific, London/River Edge, NJ, UK/USA.
4. Lees, M.J. and Wagener, T. (2000). *Review of HYCOM rainfall-runoff modeling undertaken for the review of Southern Water's source yields (1997)*. Unpublished Report to Southern Water, Imperial College Consultants, London, UK.
5. Lees, M.J. and Wagener, T. (2000). *HYCOM Rainfall-Runoff Modelling of Bewl-Darwell Reservoir Source Rivers*. Unpublished Report to Southern Water, Imperial College Consultants, London, UK.
6. Jakeman, A.J., Chen, T.H., Post, D.A., Hornberger, G.M., Littlewood, I.G., and Whitehead, P.G. (1993). Assessing uncertainties in hydrological response to climate at large scale. Macroscale modelling of the hydrosphere. *Proc. Yokohama Symp.*, IAHS Publ. No. 214, pp. 37–47.
7. Cameron, D., Beven, K.J., and Naden, P. (2001). Flood frequency estimation under climate change (with uncertainty). *Hydrol. Earth Syst. Sci.* **4**(3): 393–405.
8. Binley, A.M., Beven, K.J., Calver, A., and Watts, L.G. (1991). Changing responses in hydrology: Assessing the uncertainty in physically based model predictions. *Water Resour. Res.* **27**(6): 1253–1261.
9. Parkin, G. , O'Donnell, D., Ewen, S., Bathurst, S.C., O'Connell, P.E., and Lavabre, S. (1996). Validation of Catchment Models for Predicting Land-use and Climate Change Impacts 2. Case Study for a Mediterranean Catchment. *J. Hydrol.* **175**, pp. 595–613.
10. Calver, A., Lamb, R., and Morris, S.E. (1999). River flood frequency estimation using continuous runoff modelling. *Proc. Inst. Civ. Eng. Water Maritime Energy* **136**: 225–234.
11. Lamb, R. (2000). An approach to the calibration of a conceptual rainfall-runoff model for flood frequency estimation by continuous simulation. *Water Resour. Res.* **35**: 3103–3114.
12. Mroczkowski, M., Raper, G.P., and Kuczera, G. (1997). The quest for more powerful validation of conceptual catchment models. *Water Resour. Res.* **33**(10): 2325–2335.
13. Kitanidis, P.K. and Bras, R.L. (1980). Real-time forecasting with a conceptual hydrologic model 1. Analysis of uncertainty. *Water Resour. Res.* **16**(6): 1025–1033.
14. Kitanidis, P.K. and Bras, R.L. (1980). Real-time forecasting with a conceptual hydrologic model 2. Applications and results. *Water Resour. Res.* **16**(6): 1034–1044.
15. Moore, R.J. (1999). Real-time flood forecasting systems: Perspectives and prospects. In: R. Casale, and C. Margottini (Eds.), *Floods and Landslides: Integrated Risk Assessment*, Springer-Verlag, Berlin, pp. 147–189.
16. Dümenil, L. and Todini, E. (1992). A rainfall-runoff scheme for use in the Hamburg climate model. In: J.P. O'Kane (Ed.), *Advances in Theoretical Hydrology*, Elsevier, Amsterdam, pp. 129–157.
17. Wood, E.F., Lettenmaier, D.P. and Zatarin, V.G. (1992). A land-surface hydrology parameterisation with sub-grid variability for general circulation models. *J. Geophys. Res.* **97**(D3): 2717–2728.
18. Todini, E. (1988). Rainfall-runoff modelling—past, present and future. *J. Hydrol.* **100**: 341–352.
19. Becker, A. (1992). Criteria for a hydrologically sound structuring of large scale land surface process models. In: J.P. O'Kane, (Ed.) *Advances in Theoretical Hydrology*, Elsevier, Amsterdam, pp. 97–111.
20. Wheater, H.S., Jakeman, A.J., and Beven, K.J. (1993). Progress and directions in rainfall-runoff modelling. In: A.J. Jakeman, M.B. Beck, and M.J. McAleer (Eds.), *Modelling Change in Environmental Systems*, John Wiley, Chichester pp. 101–132.
21. Hsu, K.-L., Gupta, H.V., and Sorooshian, S. (1995). Artificial neural network modelling of the rainfall-runoff process. *Water Resour. Res.* **31**(10): 2517–2530.
22. Dawson, C. and Wilby, R. (1998). An artificial neural network approach to rainfall-runoff modelling. *Hydrological Sci. J.* **43**: 47–66.
23. Young, P.C. (1984). *Recursive Estimation and Time-Series Analysis: An Introduction.* Springer-Verlag, Berlin, Germany.
24. Young, P.C. (1992). Parallel processes in hydrology and water quality: A unified time-series approach. *J. IWEM* **6**: 598–612.
25. Jakeman, A.J. and Hornberger, G.M. (1993). How much complexity is warranted in a rainfall-runoff model? *Water Resour. Res.* **29**(8): 2637–2649.
26. Young, P.C. and Beven, K.J. (1994). Data-based mechanistic modelling and the rainfall-flow nonlinearity. *Environmetrics* **5**: 335–363.
27. O'Connell, P.E. (1991). A historical perspective. In: D.S. Bowles, and P.E. O'Connell (Eds.) *Recent Advances in the Modelling of Hydrologic Systems*, Kluwer Academic, Dordrecht, pp. 3–30.
28. Boyle, D.P., Gupta, H.V., Sorooshian, S., Koren, V., Zhang, Z., and Smith, M. (2001). To wards improved streamflow forecasts: The value of semi-distributed modelling. *Water Resour. Res.* **37**(11): 2739–2759.
29. Abdullah, F.A. and Lettenmaier, D.P. (1997). Development of regional parameter estimation equations for a macroscale hydrologic model, *J. Hydrol.* **197**: 230–257.

30. Freeze, R.A. and Harlan, R.L. (1969). Blueprint for a physically-based digitally simulated hydrologic response model. *J. Hydrol.* **9**: 237–258.
31. Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., and Rasmussen, J. (1986). An introduction to the European Hydrological System—Système Hydrologique Europeén, SHE. 1. History and philosophy of a physically-based, distributed modelling system. *J. Hydrol.* **87**: 45–59.
32. Beven, K.J. (2000). *Rainfall-Runoff Modeling: The Primer*. John Wiley, Chichester, U.K.
33. Refsgaard, J.C. and Abbott, M.B. (1996). The role of distributed hydrological modelling in water resource management. In: M.B. Abbott, and J.C. Refsgaard (Eds.), *Distributed Hydrological Modelling*, Kluwer Academic, Dordrecht, pp. 1–16.
34. Beven, K.J. (1989). Changing ideas in hydrology—The case of physically-based models. *J. Hydrol.* **105**: 157–172.
35. Calver, A. (1988). Calibration, sensitivity and validation of a physically based rainfall-runoff model. *J. Hydrol.* **103**: 102–115.
36. Refsgaard, J.C. and Knudsen, J. (1996). Operational validation and intercomparison of different types of hydrological models. *Water Resour. Res.* **32**(7): 2189–2202.
37. Singh, V.P. (Ed.). (1995). *Computer Models of Watershed Hydrology*. Water Resources Publishers, Highlands Ranch.
38. Singh, V.P. and Frevert, D. (Eds.). (2002). *Mathematical Models of Large Watershed Hydrology*, Vol. 1. Water Resources Publishers, Highlands Ranch.
39. Singh, V.P. and Frevert, D. (Eds.). (2002). *Mathematical Models of Small Watershed Hydrology*, Vol. 2. Water Resources Publishers, Highlands Ranch.
40. Anderson, M.G. and Burt, T.P. (1985). Modelling strategies. In: M.G. Anderson, and T.P. Burt (Eds.), *Hydrological Forecasting*, John Wiley & Sons, New York, pp. 1–13.
41. Woolhiser, D.A. and Brakensiek, D.L. (1982). Hydrologic system synthesis. In: C.T. Haan, H.P. Johnson, and D.L. Brakensiek (Eds.), *Hydrologic Modelling of Small Watersheds*. ASAE Monograph, Washington, DC, pp. 3–16.
42. Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N., and Turcotte, R. (Eds.). (2003). *Calibration of watershed models*. AGU Water Science and Application, 6, Washington, DC.
43. Klemeš, V. (1986). Operational testing of hydrological simulation models. *Hydrological Sci. J.* **31**(1): 13–24.
44. Beven, K.J. and Binley, A.M. (1992). The future of distributed Models: model calibration and uncertainty prediction. *Hydrol. Process.* **6**: 279–298.
45. Gupta, H.V., Beven, K., and Wagener, T. (2004). Model calibration and uncertainty estimation. In: M.G. Anderson (Ed.), *Encyclopedia of Hydrological Sciences*, John Wiley, Chichester, UK.

MODELING OF WATER QUALITY IN SEWERS

Ole Mark
Asian Institute of Technology
Hørsholm, Denmark

Modeling of water quality in sewers emerged during the late 1990s. The water quality and amount of water in sewers are dominant factors for designing wastewater treatment plants and for analyzing water quality impacts from sewer overflows. In addition, sewage under anaerobic conditions may cause corrosion of pipes, malodors, and pose health risks for sewage workers. Hence, it is of utmost importance to analyze and understand the transport and transformation of wastewater in sewers. Ideally, modeling of sewer systems and wastewater treatment plants should be based on the same general principles and done in integrated models to pursue an integrated understanding and management of sewers, wastewater treatment plants, and receiving waters. However, it must be stressed that the transformation of wastewater in sewers is not identical to the processes in activated sludge and biofilm systems—so even though several similarities exist between sewers and wastewater treatment plants, modeling of water quality in sewers is a discipline of it own. This article focuses on how to model water quality in sewers.

THE SEWER AS A PHYSICAL, BIOLOGICAL AND CHEMICAL REACTOR[1]

To model water quality in sewers, it is necessary to know the various types of sewers and their water quality characteristics. For a detailed description of the processes in sewer, refer to specialized literature (1). Around the world, all sewers can basically be divided into three categories:

1. storm sewers
2. wastewater sewers
3. combined sewers.

Storm Sewers. Storm sewers dominantly transport rain water, mainly runoff, during rain from city areas. Typically, storm water runoff has low concentrations of pollution and the concentration time for the storm sewer system is usually rather short—of the order of minutes to a few hours. If the concentration time is short and the concentrations are low, then aerobic conditions will dominate, and there will be only a low level of transformation and biological activity in the storm water.

Sanitary Sewers. Sanitary sewers transport wastewater from residences, commercial areas, and industries. Depending on the local layout of the sewer system, the wastewater may be transported by gravity or in pumped lines (in areas with insufficient ground slope available). The concentration time for sanitary sewers may be up to 1–2 days. Wastewater typically has high concentrations of pollution and biodegradable material, so wastewater sewers are potentially a place of high biological activity.

Combined Sewers. Combined sewers transport just wastewater in dry weather and both wastewater and

---

[1] The terminology "The sewer as a physical, chemical and biological reactor" has been largely promoted through the work of the Sewer Systems & Processes Working Group under IWA and in particular promoted by Prof. Thorkild-Hvitved-Jacobsen, Aalborg University, Denmark, and Prof. Richard Ashley, University of Bradford, UK.

runoff from city areas during rain. The concentration time in a combined sewer system, may be rather, short or it may rise to the order of 1–2 days. During dry weather flow, combined sewers operate like sanitary sewers, but during rain, rainwater that has a lower concentration of pollutants is added, and the transport time of the water to the wastewater treatment plant is significantly reduced. When a combined sewer system is overloaded, the excess water is discharged into receiving waters (a river, lake, or a coastal area), and this combined sewer overflow (CSO) may seriously deteriorate the water quality in the receiving waters.

The impact on recipients from overflows from sewer systems can be divided into three typical problems:

1. toxic impact due to un-ionized ammonia discharged mainly from CSOs
2. oxygen depletion
3. hygienic problems caused by fecal coliform.

Hence, the selection of an appropriate modeling approach for water quality in sewers depends both on the type of sewer system and the nature of the problem. Un-ionized ammonia in a combined sewer system may be accurately modeled by applying an advective dispersion model, considering that ammonia is conservative. However, it is a must to include biological and chemical processes to model hydrogen sulfide in a pumped wastewater pipe whose retention time is more than 3–4 hours. The following sections focus on modeling various types of water quality problems in sewers.

DETERMINISTIC MODELING OF WATER QUALITY

The strength of deterministic models for water quality in sewers is increased understanding of the individual processes and their interaction. Deterministic modeling of water quality in sewers and urban drainage can be divided into two groups:

1. quasi-conservative pollution
2. nonconservative pollution.

Modeling water quality in sewers is a challenging task, which assumes that the model is built and verified step by step. Note that a small error in the hydrodynamic model is carried 100% into the water quality model. A volume error of 10% more water in the hydrodynamic model will directly imply that the concentration of pollution in the water quality model will be too small. At present, modeling of pollutant transport may be calculated by using an advective–dispersion model such as MOUSE TRAP (3). Models transporting the pollution by pure advection (no dispersion) are InfoWorks (4) and the SWMM models (EPA SWMM, MIKE SWMM, and XP SWMM) (5). As software for modeling sewers develops rapidly, it is suggested that the reader always seek the latest information from the software suppliers of the models.

Modeling of Quasi-Conservative Pollution

The term *Quasi-Conservative Pollution* covers pollution that is not significantly transformed while it is in the sewer system. Ammonia and phosphorous are water quality parameters, which certainly are not 100% conservative, but which in most cases can be modeled with sufficient accuracy assuming that they are conservative (6,7).

Dissolved conservative pollution can typically be modeled by assigning a measured concentration to the wastewater and to the rain. The resulting concentration in the sewers is then effectively a dilution of the wastewater based on the land use and local industrial load to the sewer system. At this point, it is important that the model applied for pollution modeling represents the transport of pollution inside each pipe. The model must have computational grid points within each pipe otherwise there will be too much dispersion in long pipes, resulting in lowered peaks of concentration.

Modeling of Nonconservative Pollution

The term *Nonconservative Pollution* covers situations where biological and chemical processes play a significant role during the time the water passes through the sewer/drainage system. In this case, modeling the pollution by an advective–dispersion model will be insufficient and will give misleading results. Some examples of *Nonconservative Pollution* are BOD, COD, and hydrogen sulfide. If the sewer system has a very short concentration time, even BOD and COD may be modeled by ignoring the water quality processes, but a contribution from erosion of sediment deposits in the sewers must still be included.

The advective–dispersion model basically has only one calibration parameter, the dispersion coefficient, because advection of the pollution depends fully on the hydrodynamics. To determine the dispersion coefficient, a tracer measurement can be done with salt, rhodamine, uranine, or a similar substance, or the model can be calibrated based on measurements during rainfall of a quasi-conservative pollutant such as ammonia. A tracer study is highly recommended because it verifies that the model accurately describes the transport of a conservative substance under known flow conditions. In addition to finding the dispersion coefficient, the tracer measurement program serves three purposes:

- verification of the assumptions about infiltration to the sewer system
- the possibility of fine-tuning the Manning numbers along the sewers in the tracer measurement
- validation of the capability of the modeling system to describe the physical transport of the dissolved pollutants.

In conclusion, a tracer study serves well as a verification process of a hydrodynamic model. If the model fails to reproduce the tracer results, then there is no reason to continue the study to model anything related to water quality. The final set of input data to the advection-dispersion model consists of

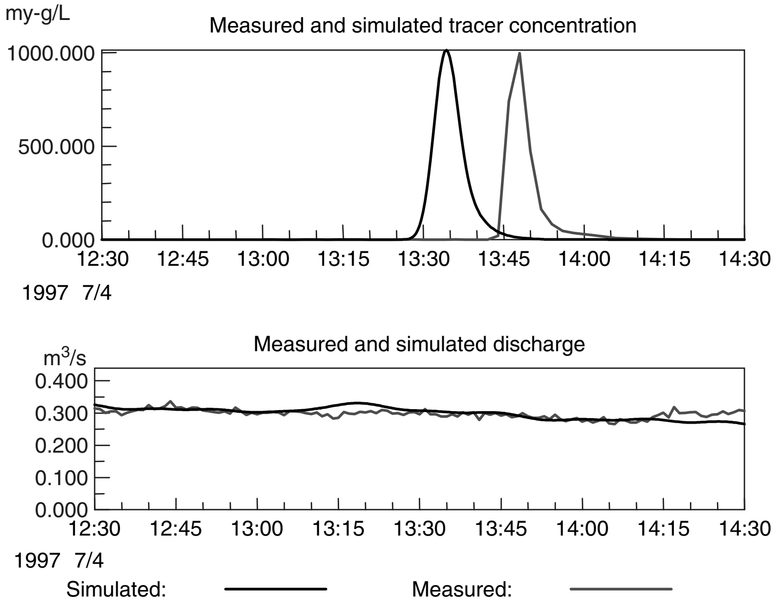

Figure 1. Example of tracer study results, showing that the hydrodynamic model needs to be recalibrated.

1. diurnal variation of dry weather flow
2. diurnal variation of concentration
3. infiltration
4. rainfall runoff
5. major industrial loads in terms of flow and concentration.

The discrepancies between a tracer study and advective dispersion, shown in Fig. 1, are for the city of Helsingborg, Sweden (7).

During the measurements, the tracer dose was added as a half-minute long pulse. From the figure, it can be seen that there was a very good match between the measured and simulated flow, but the timing of the measured and the simulated peaks was out of phase for the tracers. It was found that this occurred because the Manning numbers were never calibrated properly and only a flow gauge and not a water level gauge was available at this location in the sewers. Simulation results by applying an advective dispersion model for modeling various water quality parameters can be seen in Fig. 2 for the Helsingborg sewer system in Sweden.

Sediment Attached Pollution

Pollution may be attached to sediments, especially BOD, COD and phosphorous (2,8). When suspended sediment or sediment deposits exist in sewers, it is of utmost importance to include a model that can also simulate sediment transport processes in the sewers—including erosion and deposition of sediments. Principles for modeling sediment transport in sewers can be found in Mark et al. (9) and IWA (2).

Modeling Nonconservative Pollution by Application of a Water Quality Model

Nonconservative pollution in sewers must be modeled by describing the water quality processes along with the advective dispersion mode, preferably based on a hydrodynamic model. At this time, only the MOUSE TRAP model has such capabilities (3). The SWMM models (EPA SWMM, MIKE SWMM and XP SWMM) and InfoWorks (4) have only an advection model.

The transformation of pollution may either be modeled by applying a model based on decay rates or by using a model based on the physical, biological, and chemical processes, such as the WATS model (1). The WATS model has thus been successfully applied to wastewater modeling in the Estoril sewer system in Portugal (10). Good case studies demonstrating the applicability of water quality process models during rain are still lacking at this time.

To calibrate a water quality process model, it is of utmost importance to have a minimum of two stations measuring hydraulic and water quality parameters in the important sewers in the catchment because the water quality parameters will determine a process rate, which can be found only by comparing the measurements from two or more measurement stations. Some typical water quality processes under aerobic conditions can be seen in Fig. 3.

MODELING HYDROGEN SULFIDE

Understanding and modeling hydrogen sulfide is important because it introduces problems such as corrosion of pipes and other metal constructions in sewers. It is toxic, it poses a serious health risk to the sewer maintenance staff, and it releases noxious odors to the city environment. Further, it may inhibit sewage treatment processes (11). A high sulfide content may be toxic to fish in streams affected by sewer overflow (12). Hydrogen sulfide develops under anaerobic conditions in sewers, and it interacts with anaerobic water quality processes (1,13). Hydrogen sulfide has traditionally been modeled by using empirical equations such as: Thistlethwayte (14), Boon and Lister (15), Pommery & Parkhurst (16), and Nielsen

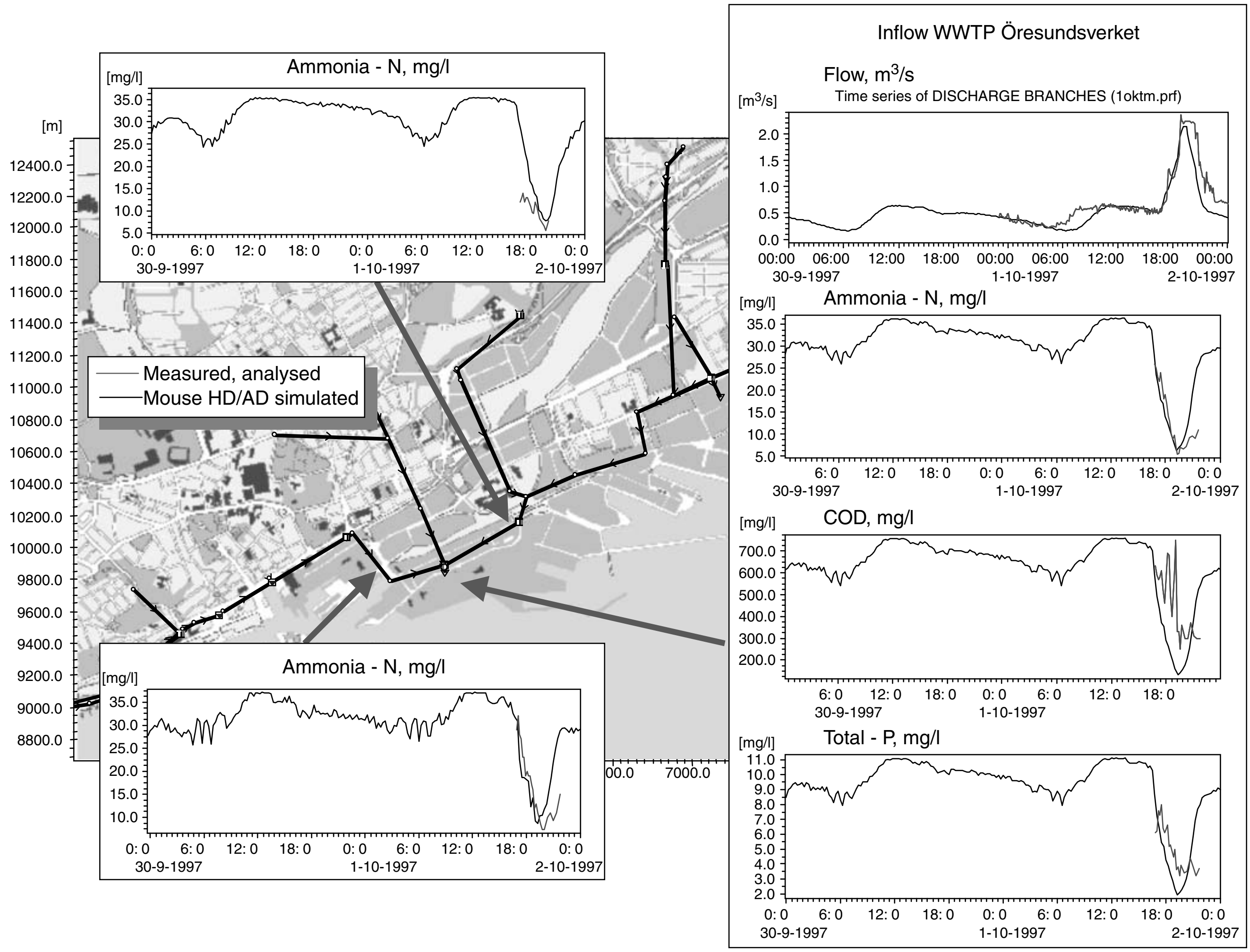

Figure 2. Wet weather sampling. Examples of measured flow and wastewater analyses compared to preliminary MOUSE hydrodynamic and advective dispersion simulation results.

and Hvitved-Jacobsen (17). Note that the four empirical models in these references were developed by using data from pressure mains, so they are not generally applicable to all sewer systems. At present, the only hydrogen sulfide model in commercial software is the MOUSE TRAP model (18). This model includes an advective dispersion model and four empirical equations for sulfide buildup. In addition, oxygen demand and extra Michaelis–Menten expressions are included in the model, so it approaches a more deterministic description of sulfide production.

A more recent and generally applicable model for hydrogen sulfide in sewers is the WATS model (1). The WATS model applies an integrated aerobic and anaerobic concept to the transformation of organic matter and sulfur in wastewater (1).

MODELING FECAL COLIFORM IN SEWERS

Experience modeling fecal coliforms has shown that coliform cannot be considered conservative or without natural decay in the receiving waters (20), for example, in the sea, where the main controlling parameters for bacterial decay are salinity, temperature, and light intensity. At present, limited knowledge is available concerning the transport of coliform in sewers; it is modeling fecal coliform in sewers by using a first-order decay rate is suggested (21); it is conservative when the concentration time in the sewer system is short—up to a few hours (22).

A PROCEDURE FOR MODELING WATER QUALITY IN SEWERS

In most cases, sediment affects pollution significantly, and when the concentration time of the sewer system is long (as a rule of thumb, 3–4 hours), the interaction of BOD/COD and oxygen must be taken into account in the model to produce reliable results. When the oxygen level plays a significant role, for instance, in modeling hydrogen sulfide, a water quality process model is a must. A consistent procedure for water quality modeling is outlined here. Each step in the modeling procedure must be completed and verified before continuing to the next step. If the model results at a step are unsatisfactory, then it is not feasible to continue because the outcome of each step depends on

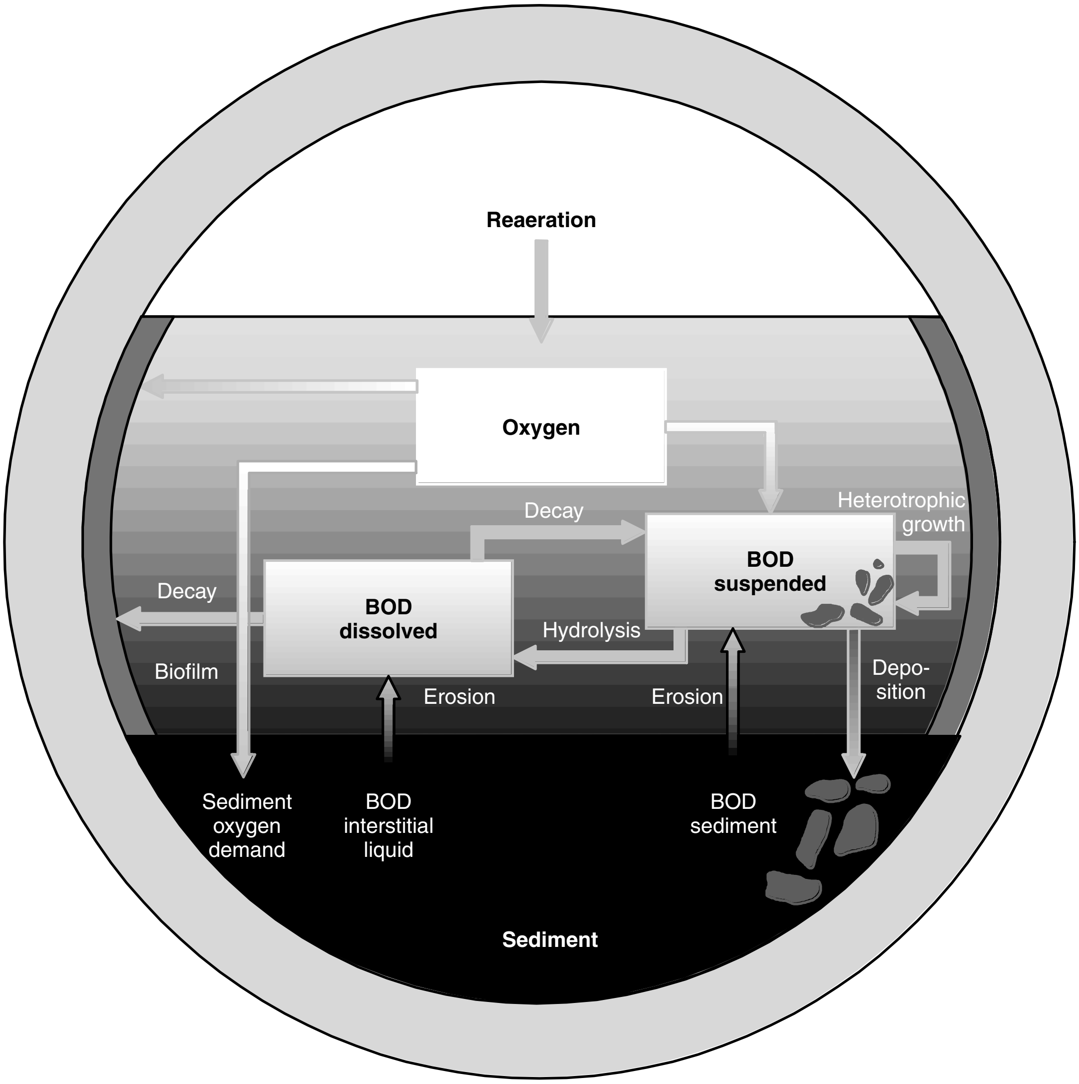

Figure 3. Example of water quality processes in wastewater under aerobic conditions.

the success of the previous steps. The modeling procedure is as follows:

1. Make a tracer verification of the hydrodynamic model, and apply a pure advection and/or advective–dispersion model.
2. Model a real quasi-conservative pollutant, for example, ammonia for both dry weather flow and rainstorms.
3. Analyze the ratio of dissolved and particulate COD and BOD, and check if a “first flush” exists.
4. If the water quality parameters in the study are attached to sediment, then model the sediment transport of the suspended solids.
5. To determine the water quality process rates, obtain measurements of flow and pollution for relevant periods at upstream and a downstream locations in the sewer system.
6. Model dissolved COD and BOD using the process rates from step 5.
7. Model total COD and BOD as the outcome of the modeling of the dissolved COD/BOD and the sediment transport of particulate COD and BOD.

LONG-TERM SIMULATION OF POLLUTION

High computational speed has made it feasible to carry out hydrodynamic and water quality simulations applying deterministic models. These models allow computing annual loads of pollutants to a WWTP or annual loads discharged from CSOs. In addition, extreme events can be modeled, such as those giving rise to high CSO discharges of water and pollutants that can cause fish kills. Through long term simulation, extreme events can be assigned return periods, providing information concerning how often various events occur. In Fig.4, the result from a 3-month simulation is seen. Long-term simulations generally require data spanning three to five times the (return) period of interest.

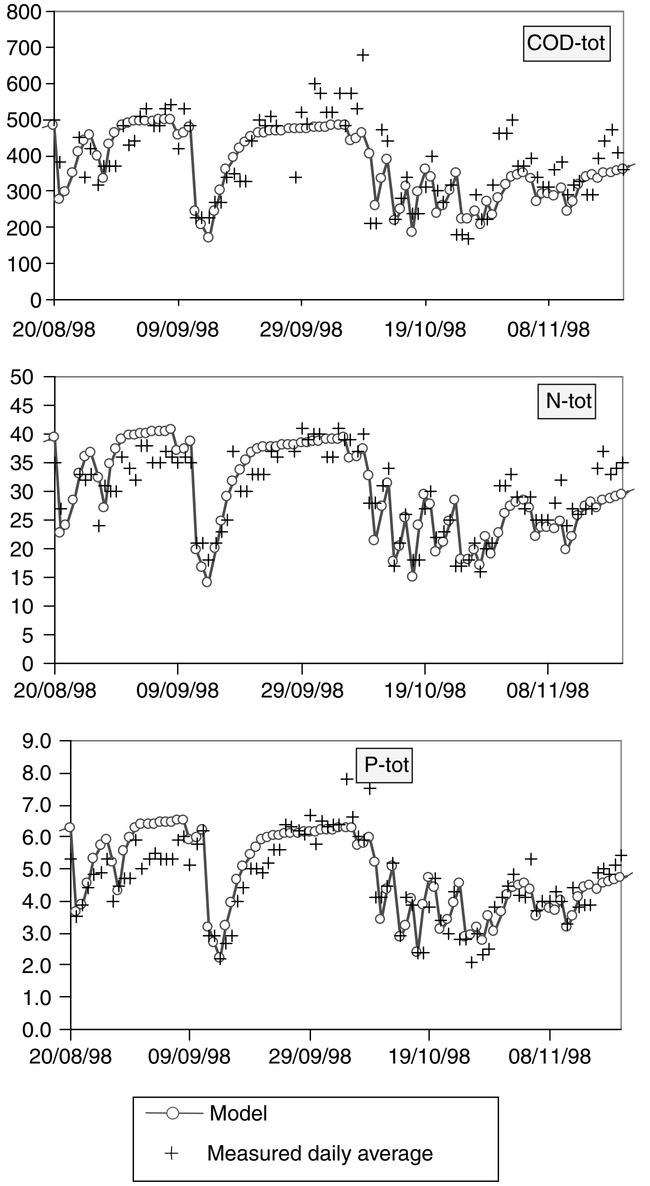

Figure 4. Simulated ammonium concentration in the inflow to the Øresundsverket, WWTP, Helsingborg, compared with analyzed 24 hourly values (each workday).

FUTURE PERSPECTIVES AND CHALLENGES IN RELATION TO MODELING OF WATER QUALITY IN SEWERS

At present, it is feasible to model water quality and pollution in sewers. However, for successful modeling, it is of great importance to understand the hydrodynamics, the transport of pollution, and the potential associated biological and chemical processes. Further, a model describing the dominant processes is essential. And, finally, a systematic procedure must be applied whereby the performance of the water quality model is verified step by step.

Holistic analysis of the sewer system, the wastewater treatment plant, and receiving waters is still an emerging area that has potential for integrated analyses of overall performance. The sewer system and the water quality processes in the sewers play a key role in this analysis, and, as more information becomes available concerning modeling of water quality in sewers, more solutions, and hopefully, better and sustainable solutions may be developed for handling the water in sewer systems.

BIBLIOGRAPHY

1. Hvitved-Jacobsen, T. (2002). *Sewer Processes—Microbial and Chemical Process Engineering of Sewer Networks*. CRC Press, Boca Raton, FL.
2. IWA/IAHR Joint Committee on Urban Drainage. (2003). *Solids in Sewers*. IWA Publishing.
3. Garsdal, H., Mark, O., Dørge, J., and Jepsen, S.-E. (1994). *MOUSE TRAP: Modelling of Water Quality Processes and the Interaction of Sediments and Pollutants in Sewers. Specialized International Conference: The Sewer as a Physical, Chemical & Biological Reactor*. Aalborg, Denmark.
4. Bouteligier, R., Vaes, G., and Berlamont, J. (2001). *Dissolved Pollutant Modelling in Urban Drainage Systems—A Case Study of the Wezembeek Oppem Catchment on the Transport of Ammonium using INFO WORKS CS*. *MTERM Workshop on Modelling of Water Quality in Urban Drainage Systems*, Asian Institute of Technology, December.
5. Huber, W.C. and Dickinson, R.E. (1988). *Storm Water Management Model—SWMM, Version 4 User's Manual*. US Environmental Protection Agency, Athens, GA.
6. Mark, O., Appelgren, C., and Kosir, M. (1996). *Water Quality Modelling for the Ljubljana Master Plan*. *7th International Conference on Urban Storm Drainage*, Hanover, Germany.
7. Mark, O., Hernebring, C., and Magnusson, P. (1998). Optimisation and control of the inflow to a wastewater treatment plant using integrated modelling tools. *Water Sci. Technol.* **37**(1): 347–354.
8. Crabtree, R.W. (1989). Sediments in Sewers. *J. Inst. Water Environ Manage.* **3**.
9. Mark, O., Appelgren, C., and Larsen, T. (1995). Principles and approaches for numerical modelling of sediment transport in sewers. *Water Sci. Technol*. **31**(7): 107–115.
10. Mourato, S., Almeida, M., Matos, J., and Hvitved-Jacobsen, T. (2002). Modelling in-sewer pollutant degradation processes in the Costa do Estoril sewer system. *Sewer Processes and Networks*. France—Accepted for publication in *Water Science and Technology* in 2003.
11. Bentzen, G. et al. (1995). Controlled dosing of nitrate for prevention of H2S in a sewer network and the effects on the subsequent treatment process. *Water Sci. Technol.* **31**(7): 293–302.
12. Nielsen, P.H., Raunkjær, K., Norsker, N.H., Jensen, N.A., and Hvitved-Jacobsen, T. (1992). Transformation of wastewater in sewer systems—a review paper. *Water Sci. Technol.* **25**(6): 17–31.
13. Tanaka, N., Hvitved-Jacobsen, T., and Horie, T. (1998). Transformation of wastewater organic matter in sewers under changing aerobic/anaerobic conditions. *Water Sci. Technol.* **72**(6): 651–666.
14. Thistlethwayte, D.K.B. (Ed.). (1972). *The Control of Sulfides in Sewage Systems*. Butterworth, Sydney, Australia, p. 173.
15. Boon, A.G. and Lister, A.R. (1975). Formation of sulphide in rising main sewers and its prevention by injection of oxygen. *Prog. Water Technol.* **7**: 289–300.
16. Pommery, R.D. and Parkhurst, J.D. (1976). *The Forecasting of Sulphide Build-up Rates in Sewers*. *Proceeding of the*

8th International Conference on Water Pollution Research, Sydney.

17. Nielsen, P.H. and Hvitved-Jacobsen, T. (1988). Effect of sulfate and organic matter on the hydrogen sulfide formation in biofilms of filled sanitary sewers. *J. Water Pollut. Control Fed.* **60**: 627–634.
18. Erichsen, A.C., Andersen, K.K., and Mark, O. (1999). *Sulphide Occurrence in Sewer Networks—a New Topic in MOUSE TRAP. 3rd DHI Software Conference*, Helsingør, Denmark.
19. Van Assel, J. and Carrette, R. (2002). A practical application of Urban Pollution Modelling in Flanders (Belgium): The catchment of Tielt. *Water Sci. Technol.* **45**(3): 125–129.
20. Evison, L.M. (1986). Comparative studies on the survival of indicator organisms and pathogens in fresh and sea water. *Water Sci. Technol.* **20**: 1051–1054.
21. Rauch, W. et al. (2002). Mathematical modeling of integrated urban drainage systems. *Water Sci. Technol.* **45**(3): 81–94.
22. Tomicic, B., Lutzen, A., and Mark, O. (2001). *Integrated Modeling of the Sewer System and the Receiving Waters for the Island of Ischia. Conference: UDM*. May, Orlando, FL.

MODELING OF URBAN DRAINAGE AND STORMWATER

Ole Mark
Asian Institute of Technology
Hørsholm, Denmark

Sutat Weesakul
DHI Water & Environment
Bangkok, Thailand

INTRODUCTION

The construction and operation of urban drainage and stormwater systems have historically been driven by the objectives of maintaining public hygiene and preventing flooding. All over the world, cities have grown rapidly during the last century, and city authorities have built and continuously expanded storm and wastewater drainage facilities. Many old cities have developed according to their varying historical needs and visions. Hence, the layout and design of the infrastructure have gradually developed into rather complex systems, which require tools to handle such situations.

An understanding of the physical system and its interaction with the environment is a prerequisite for effective planning and management of urban drainage and stormwater systems. In this respect, computer models provide the opportunity for well-structured analyses of the current performance of a system and the effect of potential changes in the system; they offer a sound scientific framework for coordinated management and planning. Apart from assessing different scenarios, the models also help to improve process understanding. Urban drainage models are used to understand the often rather complex interaction between rainfall and overflow/flooding. Once the existing conditions have been analyzed and understood, alleviating schemes can be evaluated, and the optimal scheme implemented.

Today's advances in computer technology allow many cities all over the world to report overflows from their sewers yearly using a combination of measurements and modeling, and in many places, they manage local and minor flooding problems using computer-based solutions. This involves developing computer models of the drainage/sewer system, for instance, by using software such as InfoWorks (1), MIKE URBAN (2), MOUSE (3), PCSWMM (4), and SWMM (4). All are hydrodynamic modeling packages that solve the one-dimensional equations of St. Venant.

TYPES OF MODELS USED FOR URBAN DRAINAGE AND STORMWATER SYSTEMS

Any model is a simplified mathematical representation of a physical system. This representation may be based on a ***deterministic*** approach (i.e., a fixed relationship between physical disturbance and its effect, such as model results) or it could be ***stochastic***, involving terms of probability in the model inputs and the interpretation of model results. Models for urban drainage and stormwater are predominantly deterministic models, and the text below refers only to this class.

Furthermore, deterministic models applied to urban drainage can be roughly classified as ***physical models*** and ***conceptual models***. Placing a model in one of these classes relies on the level of mathematical sophistication of the treatment of underlying physical processes in the model. Practically, models are classified according to the importance of empirical parameters for the models' ability to describe these processes accurately. The reliance on empirical parameters classifies the model as conceptual and emphasizes the need for validation against field measurements. In this context, various hydrologic models would belong to the class of conceptual models; a hydrodynamic network model is an example of a physical deterministic model.

BRIEF HISTORY OF URBAN DRAINAGE MODELING

Modeling of urban sewer and drainage systems dates back to the late 1960s when the U.S. EPA funded the development of the first version of the EPA Stormwater Management Model (SWMM). The first version 1.0 of SWMM was released in 1971. Around this time, Wallingford (UK) and DHI Water and Environment (DK) developed their first one-dimensional flow models, but it was not until 1982–86, with the release of the commercial software packages, MicroWASSP (1982), MOUSE (3), and PCSWMM (4), that they were no longer 'labeled' as 'in-house software.' These modeling packages came very soon after the first PCs hit the market. Since the release of personal computers (PC), modeling software for urban drainage has closely followed and fully used the full powers of the PC. As PCs have become more powerful, the modeling software has increased in complexity, and the size of the urban drainage model has gone from a maximum of around 250–500 pipes in the mid-1980s to modeling the drainage systems of mega cities, (e.g., models with thousands of pipes). Computer modeling of urban drainage has not added much new within the basic theory of urban hydrology and hydraulics, but it brought fast and user-friendly computations of the St. Venant equations to the broad engineering community. During the 1990s, modeling of urban drainage broke new ground by developing both theoretical and modeling approaches for sediment transport and water quality in sewers, processes which today are still not fully understood. Lately, major modeling packages have been fully embedded in Geographical Information Systems (GIS) to ease model building based on the registrations in the databases of the local authorities and to present model output in conjunction with the cities' infrastructural data.

CALIBRATION AND VERIFICATION OF MODELS: FIELD MEASUREMENTS

Urban drainage models are expected to reproduce the behavior of the modeled system with a high level of accuracy. This is usually ensured through parameter calibrations and the verification of model results against the measured performance of the real system. Calibration involves adjusting the model's key parameters, to minimize the differences between the model results and the field measurements (e.g., water levels). A continuous period or a set of intermittent events used for calibration should include a full range of expected operational conditions in the system. Verification provides proof that the model generates results within the acceptable error range. Reliable verification should be carried out for a simulation period or intermittent event(s), which are different and independent of those used for the model calibration.

All models must, if possible, be calibrated before application. Conceptual models require more attention in this respect compared to fully deterministic, physical models because the assumptions of the conceptual models must also be proven valid. Thus, in urban drainage modeling, the focus of model calibration is usually on hydrologic models (conceptual), whereas the deterministic hydrodynamic models of the drainage network typically require only minor adjustments for accurate performance. However, independently of the applied model type, good modeling practice requires thorough model verification before application.

Reliable field measurements in sewers and drains and the results of laboratory analysis (in the case of water quality modeling) are essential for successful model verification (i.e., for the model application as a whole). The harsh environment in sewers and urban drains (poisonous gases and pathogenic microorganisms in confined conduits, aggressive fluids, floating debris), as well as particular hydraulic conditions (alternating free-surface and pressurized flow, rapid flow surges, etc.) requires special monitoring equipment and highly trained specialists for good results. Routinely measured variables include water level, flow velocity, conductivity, temperature, pH, etc. The chemical and biological properties of the water, as well as sediment characteristics, are obtained by manual or automated sampling, followed by laboratory analyses.

APPLICATIONS OF MODELS TO URBAN DRAINAGE AND STORMWATER SYSTEMS

Urban drainage and stormwater models are usually set up and applied for specific analyses, and a model of the same part of a city may be differently schematized and calibrated for different objectives, based on the application type. The next sections briefly describe common applications of urban drainage and stormwater modeling:

- modeling of water quality
- modeling of urban flooding
- modeling the interaction between urban drainage systems, receiving waters, and wastewater treatment plants.

MODELING OF WATER QUALITY

Modeling of wastewater is important for predicting the inflow of both water quantity and quality to wastewater treatment plants and for precise descriptions of overflow from sewer systems (overflow of a mix of wastewater and rainwater from combined sewer systems)—combined sewer overflows (CSO) and overflow from sanitary sewer systems (SSO). Traditionally, overflows are modeled by using a rainfall runoff model as input to a hydrodynamic pipe flow model in conjunction with a diurnal description of the wastewater produced by the people living in the city.

In the past, the performance of a sewer system was characterized by most Environmental Protection Agencies (EPAs) by the number of overflow events per year. If a sewer system had less than, e.g., five overflows/year, it was considered satisfactory performance. This was a performance criterion which was feasible to check and implement, but it could not account for the actual volumes spilled. It was realized that five overflows/year could be spilling one thimble of water five times, or it could be huge volumes of water spilled five times. When modeling was introduced, it was possible to compute volumes from all overflows in a drainage system. In the past, they all had to be measured to get a picture of the complete overflow from the system. Relatively cheap and reliable application of models for calculating overflow volumes and pollution emissions for the entire system changed the EPA focus from a frequency/year to a consideration of the yearly overflow volumes, whenever that was more appropriate. For instance, for lakes, the accumulated overflow of nitrogen and phosphorus from the drainage system is most important, whereas the number of spills is less important. For smaller streams, the loads from extreme events causing high ammonia or low oxygen concentrations may be the dominant processes. The approach to sewer rehabilitation/optimization was changed from a reduction of the number of spills/year to a reduction of the yearly overflow volumes together with consideration of the load to receiving waters. In other words, "The focus went downstream."

The load to streams and rivers would typically be computed based on a conceptual hydraulic model such as the SAMBA model (5). This kind of model is simplified so that it includes only computation of a time–area runoff hydrograph together with the most significant elements in the pipe/drainage system, weirs, pumps, locations with divergent flow, and outlets. Due to the simplifications, the calculations are very fast and suitable for scanning a large number of alternative layouts of drainage systems. In the early days of computer modeling, this kind of modeling was the only way to calculate annual loads and extreme statistics of discharges of water and pollutants based on historical time series of rain events. The model focused on the requirements from the EPA.

The introduction of water quality modeling of urban drainage such as MOSQITO (6) and MOUSE TRAP (7) in the middle of the 1990s, extended the possibility of hydrodynamic modeling of "clean" water to include modeling of the transport of pollution and sediments in the sewer and drainage systems. This further enhanced the description of the modeling of wastewater and the pollution loads to the receiving water because this kind of water quality modeling provides an accurate description of the dilution process when rainwater mixes with sewage, together with an estimate of the water quality processes in the system. The outcome was more accurate computations of concentrations of conservative pollutants in the CSO and, for example, ammonia can be modeled accurately in most cases. For more information concerning modeling water quality, refer to the section in the Encyclopedia of Water, **Modeling of Water Quality in Sewers**.

During the 1990s, long-term simulations of drainage network hydraulics and water quality gradually became standard practice, facilitating simulation of long historical time series of rain and their impact in terms of overflow or load on the wastewater treatment plant. This combination of the simulation of wet weather conditions by using the dynamic pipe flow model and a simple hydrologic model to simulate the dry weather conditions makes it possible to simulate many years (>20 years, but feasible length of simulation period depends on model size, computer speed, and rainfall data duration) continuously within a reasonable computational time on a PC. The results can be presented based on the statistical occurrence of impacts, for example, in terms of overflow volume. Understanding the behavior of the system and its interaction with the environment, gained from long-term simulation, is superior to simulation of selected single rains because the uncertainty from selecting a few representative rains is removed. Furthermore, based on a long-term simulation, the statistics are made on the computed output (the impact of the rain) in terms of CSO, SSO, or flooding. In the past, the statistics were made during the selection of the rains because as they would be input for the simulation.

MODELING OF URBAN FLOODING

The problems arising from urban flooding range from minor ones, such as water getting into the basements of a few houses, to major incidents where large parts of cities are inundated for several days (Fig. 1). Most modern cities in developed countries typically have smaller scale problems, due to locally insufficient sewer capacity. Other cities, for instance, in Asia-Pacific, can have more severe problems because there is insufficient drainage and much heavier local rainfall. The situation is further aggravated because cities in developing countries grow rapidly these days, but do not have the funds necessary to extend and rehabilitate their existing drainage systems. For more information concerning urban flooding, refer to the section in the Encyclopedia of Water, **Urban Flooding**.

Figure 1.

Modeling urban flooding differs significantly from modeling pipe flow, which has a confined geometry can be measured accurately and easily. The flow is basically one-dimensional and accurately described by the St. Venant equations. When the water leaves the drainage/stormwater system, it flows onto a three-dimensional ground surface where the flow routes and paths are not well defined. The surface must be described by a digital elevation model containing roads and houses in the flooded areas. Modeling surface flooding takes on additional data in an environment where the local geometry is difficult to describe accurately and the flow pattern may be one- or two-dimensional. Recently, a model was developed that combined a one-dimensional pipe flow model and a two-dimensional surface flood model (8). This approach allows exchanging surcharge water and flow in the pipe network The results can be used to identify hazardous flood zones, where high velocity occurs (9). Another two-dimensional model using a shallow water equation has been developed to study the flow characteristic of street flooding when water escapes from manholes (10). Modeling urban flooding is an emerging area where applications have started to appear during the last decade. For further information, refer to Mark et al. (11).

Examples of cities that experience urban flooding problems on a larger scale are Bangkok, Thailand; Dhaka, Bangladesh; Ho Chi Min City, Vietnam; Mumbai, India; and Phnom Penh, Cambodia. In the year 2000, nearly 17,000 telephone lines in Mumbai went out of service after flooding occurred, and the electricity supply was switched off for a long time in most areas for safety reasons. The floods reached 1.5 m deep in the locations that were most severely inundated. Fifteen lives were lost in the Mumbai flooding (Times of India, July 25, 2000). In Dhaka, even a little rain can inundate parts of the city for several days. The situation was highlighted in September 1996, when the residents of Dhaka experienced ankle to knee-deep water in the streets. Heavy traffic jams occurred because of the stagnant water on the streets, and daily activities in parts of the city were almost paralyzed. In 1983, Bangkok was flood inundated for six months, which caused direct damage worth approximately 6,600 million baht ($149 million) (12).

MODELING THE INTERACTIONS AMONG URBAN DRAINAGE SYSTEMS, RECEIVING WATERS, AND WASTEWATER TREATMENT PLANTS

"Integrated modeling" in urban drainage is a term describing "*Modeling of the interaction between urban drainage systems, the receiving waters, and wastewater treatment plants*." The term, "receiving waters" is used here for groundwater, streams, rivers, lakes, the sea, etc. Most urban areas have these three components. Few who are involved in managing these systems would disagree with the statement: "It is logical and beneficial to plan, design and operate these components in an integrated manner" (13).

Integrated modeling is a complex exercise due to the sheer size of the model problem and also due to the different modeling approaches for each subsystem that has developed during modeling history. The hydrodynamics in flow and water levels can be directly transferred between subsystems, but presently water quality modeling is handled completely differently in channels/rivers and at the wastewater treatment plant. This causes problems in the interface and data transfer between subsystems, which again impacts accurate modeling of WQ processes.

The logic behind integrated management is old, and a number of separate tools exist. The advent of computer models has been the biggest single factor in the development of integrated analyses. Because of technology limitations in the past, historical reality has been removed from integration, and independent pragmatic criteria for each component have been the norm (13). These are some examples.

- Sewers are typically designed to contain storms of a specific return frequency without surcharging, rather than protecting against flooding.
- Combined sewer overflows are engineered to pass forward flow dictated by downstream sewer or WWTP capacity rather than the ability of the receiving water to accept the overflows.
- Treatment plants themselves are often designed to cope with an arbitrary multiple of dry weather flow, rather than treating all flow from the sewer system.

In most parts of the world, there have been fragmented planning and management of the sewer system, the wastewater treatment plant, and receiving waters. Furthermore, there is a general lack of methodology and technology available for integrated planning and management. Finally, there is simply a lack of a regulatory framework for "integrated thinking." An exception to this "rule" is the UPM procedure (14) in the United Kingdom, which is highly recommended reading. Today's challenge is to move the legislation from individual consideration of each subsystem's performance to integrated management of the urban wastewater system.

The quality of an integrated model for an urban drainage system depends strongly on the quality of the WQ sewer model, which today is the weakest link in integrated modeling. Despite these uncertainties, integrated modeling is still very beneficial, and the future will surely be dominated by integrated modeling.

REAL-TIME MODELING

At present, hydrologists working in urban areas are facing many new challenges imposed by the ever-changing hydrologic environment in cities. In such conditions, emphasis is focused on managing the urban systems as efficiently as possible within the existing infrastructure by applying currently available information and technology, where real-time modeling plays an increasingly important role. It occurs in two types of applications:

- modeling real-time control and active real-time control of urban drainage systems
- modeling as a real-time decision support tool.

MODELING REAL-TIME CONTROL AND ACTIVE REAL-TIME CONTROL OF URBAN DRAINAGE SYSTEMS

In the early 1990s real-time control of sewers (RTC) was recognized as a big leap forward to achieving superior operation of many sewer systems. Since then, RTC has certainly been proven an eminent solution for some sewer systems to reduce the risk of flooding and the amount of CSO at the same time. But RTC has not been implemented to the extent it was foreseen in the early 1990s, and it seems as if RTC never became "the big thing" that was envisaged. Some of the reasons behind the slow implementation of RTC in the real world are due to the risks of applying RTC, and the safety precautions which have to be taken to ensure smooth and secure operation of the RTC system in case of a power failure or other abnormal incidents.

The aim of RTC modeling of an urban drainage/stormwater system is to develop the strategies applied in real life for pumps, weirs, and gates and to evaluate the impact and risk from malfunctions, in a system controlled by RTC. For example, what is the flood risk if a weir or a gate is stuck during a rain?

Today RTC is applicable for

- reducing the flooding risk by using the storage capacity within the sewer system,
- reducing pollution spills from sewer overflows by retaining more stormwater and sewage within the

sewer system as a result of attenuated flows and storage,

- reducing capital costs by minimizing the storage and flow carrying capacity requirements within the system,
- reducing operating costs by optimizing pumping costs and providing the information necessary to implement effective maintenance procedures, and
- enhancing wastewater treatment plant performance by balancing inflow loads and allowing the plant to operate closer to its design capacity, thereby reducing the variability of the final effluent.

When the feasibility of RTC is evaluated for a sewer system, it is important to be aware of the changes in flow pattern introduced by RTC. If RTC is applied during dry weather flow, for example, to level out the flow to a wastewater treatment plant, it must be investigated if the new flow conditions introduced by using RTC cause sediment deposits to develop in new and undesired places. If this is the case, then sediment deposits may increase the risk of flooding, or the entire saving gained by applying RTC may be spent on increased cleaning and maintenance. Neither would be a good outcome.

MODELING AS A REAL-TIME DECISION SUPPORT TOOL

Applying a model in conjunction with real-time data (e.g., rain data) provides means for improving the information about the current status of the urban drainage system. In general, rain gauges provide information only at specific locations. The application of a model provides the possibility of filling information gaps between the rain gauges and computing the future impact of the rain. Models can be automatically executed in real time, for example, when the rainfall at specified locations exceeds preset threshold values. Based on the forecast by a hydrologic model or an urban drainage model, the level of information about the hydrologic system can be extended into the future. Real-time hydrologic information makes it possible to detect potential hazards shortly before or just after they happen, for example, heavy rainfall recorded in certain parts of the catchment that causes flooding, or overflows, and actions can be taken to mitigate the problems in the drainage system. Another example is that local traffic control can redirect traffic to roads not flooded or provide decision support to the operators of pumps and weirs in the system (15).

The application of real-time hydrologic information is not the full solution to overflow and flooding problems, but it may mitigate flood problems and in this way reduce the economic losses and the stress on people and the environment. Finally, the information generated by a real-time hydrologic information system can be applied by using the historical data for design and maintenance analyses to achieve better functionality of the urban hydrologic system before the next heavy rain arrives. In this fashion, the system can also be used to train operators.

FUTURE PERSPECTIVES AND CHALLENGES IN MODELING URBAN DRAINAGE

Calibration Support for Urban Drainage and Stormwater Systems

Today, urban drainage and stormwater models are calibrated to reproduce measurements of flow, water levels, and velocity through a manual calibration procedure. The reliability of the modeling is highly dependent on the adequacy of the calibration procedure. Traditionally, the model is calibrated manually using a trial-and-error parameter adjustment procedure. At present, however, automatic calibration support is being widely discussed (16) as a replacement for the manual calibration procedure, which is subjective and can be time-consuming. Automatic methods for model calibration take advantage of the speed and power of computers, and are objective and relatively easy to implement. Automatic calibration supports are still emerging. So they need to be checked, and the calibration must be thoroughly evaluated by an experienced engineer because the automatic calibration procedure may stop at a local minimum of the objective function, or it may not converge sufficiently toward the chosen objective due to deficiencies in the measured data.

WATER QUALITY MODELING

Modeling water quality in sewers is still an emerging area because water quality and sediment transport processes are difficult to describe precisely and the theory is still developing. At present, a number of practical operational problems in sewers push the research toward sustainable urban drainage systems and practical solutions to handle significant problems such as hydrogen sulfide production. Hydrogen sulfide causes a noxious smell in the public environment; it corrodes pipes and pumps in sewer networks; and poses a health risk to people working inside sewers, for example, during maintenance. When a sewer pipe collapses due to corrosion from hydrogen sulfide, it imposes huge costs on society to restore the pipe. Therefore, there is a need to understand and manage the phenomenon of hydrogen sulfide formation, and the emerging development of modeling tools for this problem makes it possible to assess causes and alleviating measures.

MODELING THE INTERACTION BETWEEN URBAN DRAINAGE SYSTEMS, RECEIVING WATERS, AND WASTEWATER TREATMENT PLANTS

The future challenge within integrated modeling is to develop a consistent modeling formulation for water quality in sewers, the wastewater treatment plant, and receiving waters, which can be carried across the three subsystems. Furthermore, water quality processes in sewers are at present still an emerging area of research, and need more research. Finally, legislation must be revised nearly all over the world to deal with permits to systems which are optimized from a holistic point of view. Today, around the world, most legislation focuses

Next Page

only on one single sub-system, independently of the rest. This will open optimization of the total urban drainage system rather than optimizing each subsystem individually.

BIBLIOGRAPHY

1. Andrews, T. (2002). Building better integrated network data. *World Water Environ. Eng.* July: 26.
2. Andersen, H.S., Larsen, L.C., Lindberg, S., Tamasauskas, H., and Mark, O. (2004). *MIKE URBAN—Integration of Modelling and GIS*. *6th Int. Conf. Urban Drainage Modelling—UDM'04*, Dresden, Germany.
3. Lindberg, S, Nielsen, J.B., and Carr, R. (1989). An integrated PC-modelling system for hydraulic analysis of drainage systems. *1st Aust. Conf. Tech. Comput. Water Ind.: Watercomp '89*, Melbourne, Australia.
4. James, W. and Robinson, M.A. (1984). PCSWMM3, Version 3 of the runoff, executive and extended transport blocks adapted for the IBM-PC, *Proc. Stormwater Water Qual. Modelling Conf.*, sponsored by USEPA and Ontario Ministry of Environment, Burlington, Ontario, Sept. 6–7, published by CHI Hamilton, pp. 39–52.
5. Johansen, N.B., Linde-Jensen, J.J., and Harremoës, P. (1984). Computing combined system overflow based on historical rain series. Chalmers University of Technology, *3rd Int. Conf. Urban Storm Drainage*, Göteborg, Sweden.
6. Blanc, D., Kellagher, R., Phan, L., and Price, R. (1995). FLUPOL-MOSQITO, models, simulations, critical analysis and development. *Water Sci. Technol.* **32**(1): 185–192.
7. Garsdal, H., Mark, O., Dørge, J., and Jepsen, S.-E. (1994). MOUSE TRAP: Modelling of water quality processes and the interaction of sediments and pollutants in sewers. *Specialized Int. Conf.: Sewer Phys. Chem. Biol. Reactor*, Aalborg, Denmark.
8. Alam, J.M. (2003). Two-dimensional Urban Flood Modeling for Real Time Flood Forecasting for Dhaka city, Bangladesh. AIT Thesis no. WM-02-06. Asian Institute of Technology, Thailand.
9. Nania L., Gomez M., and Dolz, J. (2002). Analysis of risk associated with urban runoff, case study: City of Mendoza, Argentina. *9th Int. Conf. Urban Drainage*, Sept. 8–13, Portland, OR.
10. Schmitt T.G., Schilling, W., Saegrov, S., and Nieschulz, K.P. (2002). Flood risk management for urban drainage system by simulation and optimisation. *Global Solutions for Urban Drainage, 9th Int. Conf. Urban Drainage*, Sept. 8–13, Portland, OR.
11. Mark, O., Weesakul, S., Apirumanekul, C., Boonya Aroonnet, S., and Djordjević, S. (2004). Potential and limitations of 1-D modelling of urban flooding. Accepted for *J. Hydrol.*
12. Boonya-aroonnet, S., Weesakul, S., and Mark, O. (2002). Modelling of urban flooding in Bangkok. 9*th Int. Conf. Urban Drainage*, Portland, OR.
13. Rauch, W. et al. (2002). Mathematical modeling of integrated urban drainage systems. *Water Sci. Technol.* **45**(3): 81–94.
14. Foundation for Water Research. (1998). *Urban Pollution Management (UPM) Manual*, 2nd Edn. FR/CL 0009.
15. Mark, O. et al. (2002). A real-time hydrological information system for Bangkok. *Int. Conf. Urban Hydrol. 21st Century*, Oct. 14–18, Kuala Lumpur, Malaysia.
16. Wangwongwiroj, N., Schlütter, F., and Mark, O. (2003). Principles and practical aspects of an automatic calibration procedure for urban rainfall-runoff models. Accepted for the *J. Urban Water.*

MODELING UNGAUGED WATERSHEDS

THORSTEN WAGENER
The Pennsylvania State University
University Park, Pennsylvania

INTRODUCTION

Models of watershed hydrology are irreplaceable tools in today's research and practice. Their areas of application are wide reaching, from water management and flood forecasting, to their use as load models for water quality studies. A vast number of these models have been developed since the 1960s and they differ in temporal and spatial discretization, processes described, and in the constituent equations used. However, there are also similarities that have consequences for the way these models are commonly applied. All models of watershed hydrology aggregate the real hydrologic system on a particular element scale in space and time. The spatial scale might vary from grid cells of tens of meters to models that treat the whole catchment as a single unit, and the temporal scale might vary, for example, from 15-minute intervals to monthly time steps. It is common to assume homogeneity of processes or watershed characteristics on scales smaller than the one applied, that is, parameters are assumed to be effective on a certain scale (Fig. 1), although the heterogeneity of the real world on smaller scales is sometimes described by distribution functions (2). The characteristics of each model element, for example, storage or infiltration capacity, are described by parameters within the model. A common problem is that the scale on which these characteristics can be measured are usually different, mostly smaller, than the scale of the model element. The effect of this difference between model and measurement scale is that one has to revert to alternative methods to estimate model parameters. The usual approach is to observe the responses of the real hydrologic system, for example, streamflow, and compare them to predictions of the model. The modeler then adjusts the model parameters, in a process usually referred to as calibration, until model predictions and observations are as close as possible. Calibration can be performed using manual or automatic techniques and the available literature on this subject is immense. Between 3 to 10 years of observations are required for calibration, depending on the model complexity and the informational content of the data (3). Shorter periods can suffice when the data sufficiently trigger the response modes of the model (4).

Alternatives to model calibration have to be found when no or insufficient time series of the variable under investigation are available for this process. This is a common problem, even in countries that have extensive measuring networks such as the United Kingdom that has more than 1400 gauging stations (5). It is also possible that sufficiently long streamflow time series are available but that the modeling objective is the prediction of a different variable, for example, groundwater levels. Then, a calibration with respect to the variable under study

Previous Page

only on one single sub-system, independently of the rest. This will open optimization of the total urban drainage system rather than optimizing each subsystem individually.

BIBLIOGRAPHY

1. Andrews, T. (2002). Building better integrated network data. *World Water Environ. Eng.* July: 26.
2. Andersen, H.S., Larsen, L.C., Lindberg, S., Tamasauskas, H., and Mark, O. (2004). *MIKE URBAN—Integration of Modelling and GIS*. *6th Int. Conf. Urban Drainage Modelling—UDM'04*, Dresden, Germany.
3. Lindberg, S, Nielsen, J.B., and Carr, R. (1989). An integrated PC-modelling system for hydraulic analysis of drainage systems. *1st Aust. Conf. Tech. Comput. Water Ind.: Watercomp '89*, Melbourne, Australia.
4. James, W. and Robinson, M.A. (1984). PCSWMM3, Version 3 of the runoff, executive and extended transport blocks adapted for the IBM-PC, *Proc. Stormwater Water Qual. Modelling Conf.*, sponsored by USEPA and Ontario Ministry of Environment, Burlington, Ontario, Sept. 6–7, published by CHI Hamilton, pp. 39–52.
5. Johansen, N.B., Linde-Jensen, J.J., and Harremoës, P. (1984). Computing combined system overflow based on historical rain series. Chalmers University of Technology, *3rd Int. Conf. Urban Storm Drainage*, Göteborg, Sweden.
6. Blanc, D., Kellagher, R., Phan, L., and Price, R. (1995). FLUPOL-MOSQITO, models, simulations, critical analysis and development. *Water Sci. Technol.* **32**(1): 185–192.
7. Garsdal, H., Mark, O., Dørge, J., and Jepsen, S.-E. (1994). MOUSE TRAP: Modelling of water quality processes and the interaction of sediments and pollutants in sewers. *Specialized Int. Conf.: Sewer Phys. Chem. Biol. Reactor*, Aalborg, Denmark.
8. Alam, J.M. (2003). Two-dimensional Urban Flood Modeling for Real Time Flood Forecasting for Dhaka city, Bangladesh. AIT Thesis no. WM-02-06. Asian Institute of Technology, Thailand.
9. Nania L., Gomez M., and Dolz, J. (2002). Analysis of risk associated with urban runoff, case study: City of Mendoza, Argentina. *9th Int. Conf. Urban Drainage*, Sept. 8–13, Portland, OR.
10. Schmitt T.G., Schilling, W., Saegrov, S., and Nieschulz, K.P. (2002). Flood risk management for urban drainage system by simulation and optimisation. *Global Solutions for Urban Drainage, 9th Int. Conf. Urban Drainage*, Sept. 8–13, Portland, OR.
11. Mark, O., Weesakul, S., Apirumanekul, C., Boonya Aroonnet, S., and Djordjević, S. (2004). Potential and limitations of 1-D modelling of urban flooding. Accepted for *J. Hydrol.*
12. Boonya-aroonnet, S., Weesakul, S., and Mark, O. (2002). Modelling of urban flooding in Bangkok. 9^{th} *Int. Conf. Urban Drainage*, Portland, OR.
13. Rauch, W. et al. (2002). Mathematical modeling of integrated urban drainage systems. *Water Sci. Technol.* **45**(3): 81–94.
14. Foundation for Water Research. (1998). *Urban Pollution Management (UPM) Manual*, 2nd Edn. FR/CL 0009.
15. Mark, O. et al. (2002). A real-time hydrological information system for Bangkok. *Int. Conf. Urban Hydrol. 21st Century*, Oct. 14–18, Kuala Lumpur, Malaysia.
16. Wangwongwiroj, N., Schlütter, F., and Mark, O. (2003). Principles and practical aspects of an automatic calibration procedure for urban rainfall-runoff models. Accepted for the *J. Urban Water.*

MODELING UNGAUGED WATERSHEDS

Thorsten Wagener
The Pennsylvania State University
University Park, Pennsylvania

INTRODUCTION

Models of watershed hydrology are irreplaceable tools in today's research and practice. Their areas of application are wide reaching, from water management and flood forecasting, to their use as load models for water quality studies. A vast number of these models have been developed since the 1960s and they differ in temporal and spatial discretization, processes described, and in the constituent equations used. However, there are also similarities that have consequences for the way these models are commonly applied. All models of watershed hydrology aggregate the real hydrologic system on a particular element scale in space and time. The spatial scale might vary from grid cells of tens of meters to models that treat the whole catchment as a single unit, and the temporal scale might vary, for example, from 15-minute intervals to monthly time steps. It is common to assume homogeneity of processes or watershed characteristics on scales smaller than the one applied, that is, parameters are assumed to be effective on a certain scale (Fig. 1), although the heterogeneity of the real world on smaller scales is sometimes described by distribution functions (2). The characteristics of each model element, for example, storage or infiltration capacity, are described by parameters within the model. A common problem is that the scale on which these characteristics can be measured are usually different, mostly smaller, than the scale of the model element. The effect of this difference between model and measurement scale is that one has to revert to alternative methods to estimate model parameters. The usual approach is to observe the responses of the real hydrologic system, for example, streamflow, and compare them to predictions of the model. The modeler then adjusts the model parameters, in a process usually referred to as calibration, until model predictions and observations are as close as possible. Calibration can be performed using manual or automatic techniques and the available literature on this subject is immense. Between 3 to 10 years of observations are required for calibration, depending on the model complexity and the informational content of the data (3). Shorter periods can suffice when the data sufficiently trigger the response modes of the model (4).

Alternatives to model calibration have to be found when no or insufficient time series of the variable under investigation are available for this process. This is a common problem, even in countries that have extensive measuring networks such as the United Kingdom that has more than 1400 gauging stations (5). It is also possible that sufficiently long streamflow time series are available but that the modeling objective is the prediction of a different variable, for example, groundwater levels. Then, a calibration with respect to the variable under study

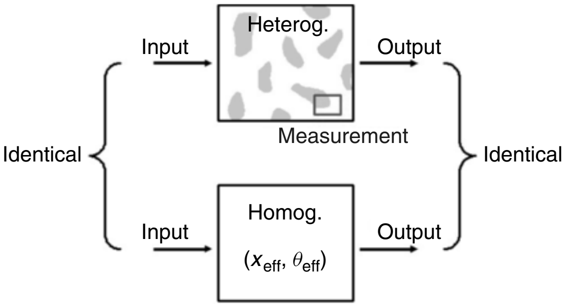

Figure 1. Definition of effective parameters and differences in scale between measurements and hydrologic modeling. (Modified from Reference 1).

would not be possible. With respect to the modeling of ungauged watersheds, there is also a difference whether the required predictions are limited to particular events or whether continuous simulation is required. Some reasonable estimates for event-based models can be derived, but this is not the case for continuous models where the predictions are highly uncertain (6). These two cases are therefore treated separately here.

MODELING THE EVENT-BASED WATERSHED RESPONSE

Empirical models are usually used for modeling individual events. In these models, the response characteristics of watersheds (e.g., mean annual flood or the percentage runoff) are related to watershed descriptors (e.g., area, drainage density, or dominant soil types) using regression equations (7;8, p. 301).

Another empirical approach that is very popular for event-based modeling in ungauged watersheds in many parts of the world is the curve number (CN) method. This approach was originally developed by the Soil Conservation Service for watersheds in the United States (9;10, pp. 147 et ff.). The basis of this technique is the assumption that the ratio of direct runoff to total precipitation is equal to the ratio of retained water to the potential maximum retention. The value of the potential maximum retention S can be calculated using values for CN:

$$S = \frac{1000}{\text{CN}} - 10 \tag{1}$$

where S is calculated in inches. Actual values for CN (0–100) given in tables or graphs are a function of soil type, land use, and antecedent moisture conditions (10, pp. 148–149), which makes this approach so attractive for modeling ungauged watersheds. These graphs and tables were originally derived from measured rainfall-runoff data on a small watershed or hillslope scale (8, p. 184).

These parametrically simple event-based models have been applied with some success (7,11). However, there is a trend to move from event-based models to those that provide continuous simulation because the initial conditions for event-based models are a major source of uncertainty. A recent workshop report on challenges in hydrologic predictability noted "in watershed rainfall-runoff transformation ... initial and boundary conditions are the critical issues" (12, p. 17). This problem has led to a more holistic approach to flood management in some countries. In the United Kingdom, for example, there is a move to replace event-based modeling with continuous approaches (13,14).

MODELING THE CONTINUOUS WATERSHED RESPONSE

Several approaches to estimating parameters in ungauged watersheds are available for continuous simulation models. The two most common approaches are (1) the derivation of regression relationships between model parameters and watershed characteristics (5,6,15–17) and (2) estimation of parameter values from measurable watershed (mainly soil) properties (18–23).

In the first approach, a chosen model structure is calibrated to a large number of watersheds for which sufficiently long and informative observations are available. An attempt is then made to derive regression equations that predict its value using a combination of several watershed characteristics. A separate equation is commonly derived for each model parameter. The parameter values in the ungauged watershed can then be estimated using the derived equations and a prediction can be made. Figure 2 shows a typical procedure for extrapolating parameters from gauged watersheds using regression analysis (6). The steps are as follows (Fig. 2): (1) Select catchments and their characteristics. (2) Select and calibrate the local model structure. (3) Select and calibrate the regional model structure. (4) Predict flow at the ungauged site. Currently, this is probably the most often applied technique; however, Wagener et al. (6) found considerable uncertainty in typical predictions using this approach.

Sometimes, it might be possible to derive at least some of the model parameters directly from measurable watershed characteristics. The scale difference between model parameters and measurements might be relatively small for some parameters if the model uses a very fine spatial distribution, or simple equations can be used to derive these parameters from a combination of watershed characteristics. Koren et al. (22), for example, show how storage capacities can be estimated from soil properties such as field capacity and wilting point. These properties are usually derived from point samples analyzed on the laboratory scale. This makes using these values for lumped parameter estimation questionable because *there is generally no theory that allows the estimation of the effective values within different parts of a heterogeneous flow domain from a limited number of small scale or laboratory measurements* (8). On the other hand, this approach does not assume that all the model parameters are independent as in the earlier mentioned regression technique. The idea of Koren et al. (22) is therefore rather to derive good initial estimates for a subsequent calibration procedure in gauged watersheds, that is, to reduce the calibration effort, and also for ungauged watershed and distributed modeling approaches. Very few examples can be found in the literature where models, using only measured parameters, have been applied without further calibration. It is unlikely that reliable

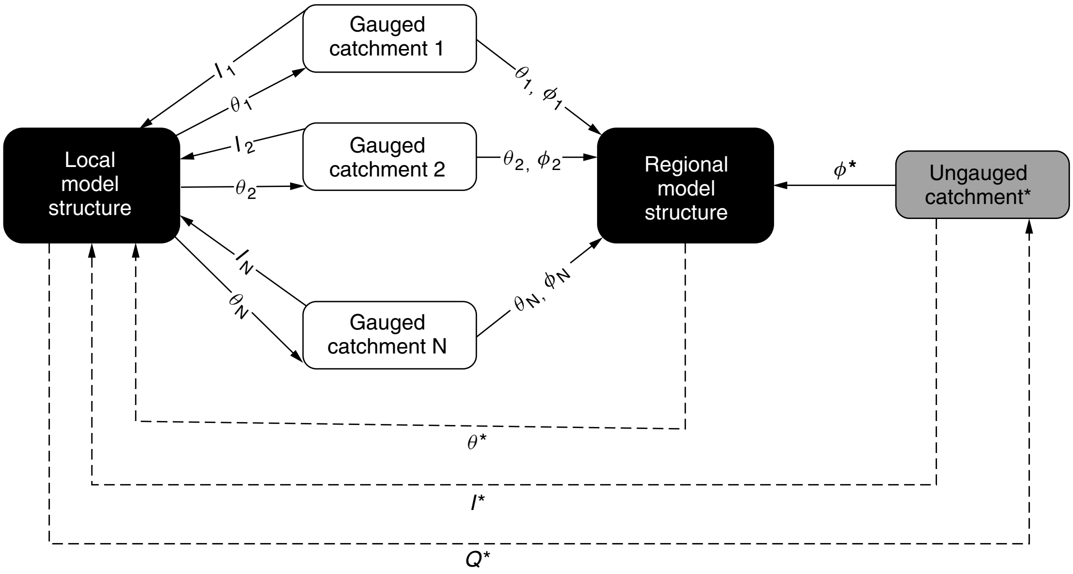

Figure 2. Schematic description of the statistical regionalization approach. (Modified from Reference 6).

predictions can be obtained by this approach from the current generation of model structures.

CONCLUSION

The ungauged problem is currently an area of extensive research and it can be expected that considerable progress will be made during the coming years (6,24,25). The complexity of the problem requires a holistic approach that can be provided only by a wide variety of hydrologists working on different topics. However, the potential value of the scientific outcome is very high. Current predictions in ungauged watersheds have to be considered as very uncertain though and must be used carefully in decision-making.

BIBLIOGRAPHY

1. Grayson, R. and Blöschl, G. (Eds.). (2000). *Spatial Patterns in Catchment Hydrology*. Cambridge University Press, Cambridge, UK.
2. Moore, R.J. and Clarke, R.T. (1981). A distribution approach to rainfall runoff modeling. *Water Resour. Res.* **17**(5): 1367–1382.
3. Yapo, P.O., Gupta, H.V., and Sorooshian, S. (1996). Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data. *J. Hydrol.* **181**: 23–48.
4. Jakeman, A.J. and Hornberger, G.M. (1993). How much complexity is warranted in a rainfall-runoff model? *Water Resour. Res.* **29**(8): 2637–2649.
5. Sefton, C.E.M. and Howarth, S.M. (1998). Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. *J. Hydrol.* **211**: 1–16.
6. Wagener, T., Wheater, H.S., and Gupta, H.V. (in press). *Rainfall-Runoff Modeling in Gauged and Ungauged Catchments*. Imperial College Press, London.
7. IH. (1999). *Flood Estimation Handbook* (5 volumes). Centre for Ecology and Hydrology, Wallingford, UK.
8. Beven, K.J. (2000). *Rainfall-Runoff Modeling: The Primer*. John Wiley, Chichester, U.K.
9. SCS (1972). *National Engineering Handbook*, Section 4, Hydrology. U.S. Department of Agriculture, Washington, DC.
10. Chow, Ven Te, Maidment, D.R., and Mays, L.W. (1998). *Applied Hydrology*. McGraw-Hill, New York.
11. NERC (1975). *The flood studies report* (5 volumes). Natural Environment Research Council, Wallingford, UK.
12. (National Research Council). (2003). *Predictability & Limits-to-Prediction in Hydrologic Systems*. National Academy Press, Washington, DC.
13. Lamb, R. (2000). An approach to the calibration of a conceptual rainfall-runoff model for flood frequency estimation by continuous simulation. *Water Resour. Res.* **35**: 3103–3114.
14. Wheater, H.S. (2002). Progress and prospects for fluvial flood modelling. *Philos. Trans. R. Soc.* **360** (1796): 1409–1431.
15. Jakeman, A.J. et al. (1992). A systematic approach to modelling the dynamic linkage of climate, physical catchment descriptors and hydrological response components. *Math. Comput. Simul.* **33**: 359–366.
16. Seibert, J. (1999). Regionalisation of parameters for a conceptual rainfall-runoff model. *Agric. Forest Meteorol.* **98–99**: 279–293.
17. Abdullah, F.A. and Lettenmaier, D.P. (1997). Application of regional parameter estimation schemes to simulate the water balance of a larger continental river. *J. Hydrol.* **197**: 258–285.
18. Beven, K.J., Kirkby, M.J., Schoffield, N., and Tagg, A. (1984). Testing a physically-based flood forecasting model (TOPMODEL) for three UK catchments. *J. Hydrol.* **69**: 119–143.
19. Parkin, G. et al. (1996). Validation of catchment models for predicting land-use and climate change impacts. 2. Case study for a Mediterranean catchment. *J. Hydrol.* **175**: 595–613.
20. Refsgaard, J.C. and Knudsen, J. (1996). Operational validation and intercomparison of different types of hydrological models. *Water Resour. Res.* **32**(7): 2189–2202.
21. Loague, K.M. and Kyriakidis, P.C. (1997). Spatial and temporal variability in the R-5 infiltration data set: Déjà

vu and rainfall-runoff simulations. *Water Resour. Res.* **33**: 2883–2896.

22. Koren, V., Smith, M., and Duan, Q. (2003). Use of a priori parameter estimates in the derivation of spatially consistent parameter sets of rainfall-runoff models. In: Q. Duan, H.V. Gupta, S. Sorooshian, A.N. Rousseau, and R. Turcotte (Eds.), *Calibration of Watershed Models*. American Geophysical Union Water Science and Application Series, 6, pp. 239–254.
23. Atkinson, S.E., Woods, R.A., and Sivapalan, M. (2002). Climate and landscape controls on water balance model complexity over changing time scales. *Water Resour. Res.* **38**(12): 1–17.
24. Sivapalan, M. (2003). Prediction in ungauged basins: A grand challenge for theoretical hydrology. *Hydrological Processes(invited commentary)*. **17**(15): 3163–3170.
25. Sivapalan, M. et al. (2003). IAHS decade on predictions in ungauged basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sci. J.* **48**(6): 867–880.

CORPS TURNED NIAGARA FALLS OFF, ON AGAIN

Frank Balon
Buffalo District,
U.S. Army Corps of Engineers

Niagara Falls, one of the Seven Natural Wonders of the World, attracts millions of tourists each year. But 30 years ago tourists flocked to see an unprecedented sight the mighty Falls reduced to a mere trickle by the engineering skill of Buffalo District.

Niagara Falls is 19 miles downstream from Lake Erie. Goat Island divides the Niagara River into two channels, separating Niagara Falls into American Falls on the American side and Horseshoe Falls on the Canadian side.

Public concern arose in the mid-1960s regarding the talus (rock fragments) accumulation at the base of the American Falls, which many felt detracted from its natural beauty. Three major rockfalls occurred in 1931, 1954 and 1959, depositing about 130,000 cubic yards of rock at the Falls' base.

In 1965, Congress authorized the U.S. Army Corps of Engineers to study the measures needed to preserve and enhance the beauty of the American Falls. The results of the Corps' study were submitted to the International Joint Commission (IJC) which, at the request of the American and Canadian governments, began an independent inquiry in 1967. The Commission established the American Falls Study Board, which included representatives from the Corps, Environment Canada, and eminent landscape architects to consider possible alternatives for preserving American Falls' future.

The IJC recommended to both countries that American Falls be temporarily dewatered to facilitate a timely and thorough study. The governments agreed, and on June 12, 1969, a rock cofferdam stopped the flow over American Falls. The Falls remained dry until November 25, 1969, when a backhoe removed the cofferdam.

The stark contrast between Niagara Falls with water and without actually increased tourist volume to the area

Once stripped of its cascading white waters, the 1,100-foot-wide, 200-foot-high precipice revealed sights never before seen. One rock outcrop closely resembled a human profile. Another giant block of rock at the fall's base was larger than most houses in the city of Niagara Falls.

Special measures were taken to prevent damage to aquatic life in the American Falls channel. Terrestrial vegetation in the channel's small islands were protected and irrigated. Sprinklers kept Rochester shale wet on the face of the American Falls. In addition, railings at viewing areas were relocated and mass rock was stabilized to protect workers and the viewing public from injury.

While American Falls was dewatered, the board conducted a detailed geologic exploration. Forty-six core holes, totaling 4,882 feet, were bored. Dye and water testing on the completed holes identified weak points in the rock mass. Face mapping included topographic, stratigraphic and structural studies. Additional testing included terrestrial photogrammetry of the Falls' face, mapping rock fractures and joints, and measuring both

---

Workers were lowered by cranes, in safety cages, to inspect the face of the falls

water pressure in rock joints and horizontal movement in the adjacent rock mass.

"It was thrilling to walk on the American Falls riverbed shortly after dewatering and to view closely the huge exposed boulders," said Andrew Piacente, a civil engineering technician in the Water Control Branch. Piacente served as the district engineer's field representative to the American Falls Working Committee and the American Falls Board of Control during the project.

Talus studies involved examining the cobbles and boulders to determine their size, rock type and condition. The talus blocks were photographed and mapped. Talus depth ranged from 25 to 50 feet.

The board considered three alternative methods for removing all or part of the talus, and estimated the time and costs associated with each. The rock removal methods involved using either a cableway system between the U.S. and Canada, large cranes on and below the Falls' crest, or a large rock crusher together with a portable conveyor. The board determined a cableway with land disposal to be the most practical method.

Data gathered from the geologic exploration was used to build a realistic model of the American Falls 1/50th its actual size. The model, built by Ontario Hydro, included removable talus blocks, which allowed accurate simulation of various talus arrangements. The turbulence, mist, illumination, and volume of water were all closely duplicated.

Far from ruining the Falls' tourist appeal, the dewatering significantly increased tourist volume. People traveled thousands of miles to see the Corps unparalleled engineering feat.

"I've seen the Falls five or six times before, in winter and in summer," said a tourist quoted in the June 13, 1969 issue of the *Buffalo Courier Express*. "But this is the best time to see it. It's so unique."

Public opinion was considered very important. Public displays describing the Falls and the Board's undertakings were exhibited in both countries. The dewatering program received intensive national and international media coverage. Public opinion was strongly in favor of maintaining the Falls' natural appearance, leaving the accumulated talus fully intact.

In the end, the board agreed. After exhaustive deliberation, the IJC recommended that the U.S. and Canadian governments leaving the talus totally intact, considering the cost of removal, the irreversible nature of the project, and the resulting accelerated erosion. In addition, the IJC concluded that artificial means should not be employed to prevent further erosion of the Falls' crest.

The timeless beauty of majestic Niagara Falls continues to inspire and serve as a symbol of international cooperation between the U.S. and Canada. A 1975 IJC report to both governments stated that "consideration of the preservation and enhancement of the beauty of the American Falls cannot be limited to their physical aspects. The appeal and fascination of the Falls mean different things to different people. Their beauty is in the eye, the mind, and the heart of the beholder."

OPEN CHANNEL DESIGN

Xing Fang
Lamar University
Beaumont, Texas

An open channel flow is a flow having a free surface, which is subject to atmospheric pressure. This type of flow is typically found in rivers, creeks, irrigation canals, drainage conduits or ditches, culverts, spillways, and sanitary sewers. When water flows partially in a closed conduit (e.g., sewer pipe), this flow should be treated as an open channel flow. Open channel design is to design a channel to carry a certain amount of flow from one location to another under any provided channel geometry and topographic conditions. Design discharge that a channel needs to convey/carry is predetermined by a hydrological analysis. The analysis could be a simple application of the rational method ($Q = CIA$, where Q is the peak discharge, C is the watershed runoff coefficient, I is the average design rainfall intensity, and A is the upstream watershed area) or a complex watershed hydrological modeling, which may include rainfall design hyetograph, rainfall loss estimation, transformation of rainfall excess into runoff, and hydrograph routing at rivers detention/retention ponds, and reservoirs.

Parameters to define channel geometry include channel shape, cross-sectional area (A), wetted perimeter (P), hydraulic radius (R), top width (T), and hydraulic depth (D). Table 1 gives proprieties for common geometric channel shapes. Side slope of a channel (z, z_1, z_2 in Table 1)

Table 1. Geometric Properties of Different Channel Cross Sections

Channel Definition	Area (A)	Wetted Perimeter (P)	Hydraulic Radius (R)	Top Width (T)	Hydraulic Depth (D)
Rectangle	By	$b + 2y$	$\frac{by}{b+2y}$	b	y
Trapezoid with equal side slopes	$(b + zy)y$	$b + 2y\sqrt{1+z^2}$	$\frac{(b+zy)y}{b+2y\sqrt{1+z^2}}$	$b + 2zy$	$\frac{(b+zy)y}{b+2zy}$
Triangle with equal side slopes	zy^2	$2y\sqrt{1+z^2}$	$\frac{zy}{2\sqrt{1+z^2}}$	$2zy$	$0.5y$
Trapezoid with unequal side slopes	$by + 0.5y^2(z_1 + z_2)$	$b + y(\sqrt{1+z_1^2} + \sqrt{1+z_2^2})$	$\frac{by + 0.5y^2(z_1+z_2)}{b + y(\sqrt{1+z_1^2} + \sqrt{1+z_2^2})}$	$b + y(z_1 + z_2)$	$\frac{by + 0.5y^2(z_1+z_2)}{b + y(z_1+z_2)}$
Triangle with unequal side slopes	$0.5y^2(z_1 + z_2)$	$y(\sqrt{1+z_1^2} + \sqrt{1+z_2^2})$	$\frac{0.5y^2(z_1+z_2)}{y(\sqrt{1+z_1^2} + \sqrt{1+z_2^2})}$	$2(z_1 + z_2)$	$0.5y$
Circle	$(\theta - \sin\theta)d_0^2/8$	$0.5\theta d_0$	$\left(1 - \frac{\sin\theta}{\theta}\right)d_0/4$	$d_0 \sin(0.5\theta)$ or $2\sqrt{y(d_0 - y)}$	$\left(\frac{\theta - \sin\theta}{\sin\theta/2}\right)d_0/8$
Parabola	$2/3\ Ty$	$T + \frac{8y^2}{3T}$	$\frac{2T^2y}{3T^2 + 8y^2}$	$\frac{3A}{2y}$	$2/3y$
Round cornered rectangle $Y > r$	$\left(\frac{\pi}{2} - 2\right)r^2 + (b + 2r)y$	$(\pi - 2)r + b + 2y$	$\frac{\left(\frac{\pi}{2} - 2\right)r^2 + (b+2r)y}{(\pi - 2)r + b + 2y}$	$b + 2r$	$\frac{(\theta/2 - 2)r^2}{b + 2r} + y$
Round-bottomed triangle	$\frac{T^2}{4z} - \frac{r^2}{z}(1 - z\cot^{-1}z)$	$\frac{T}{z}\sqrt{1+z^2} - \frac{2r}{z}(1 - z\cot^{-1}z)$	$\frac{A}{P}$	$2[z(y - r) + r\sqrt{1+z^2}]$	$\frac{A}{P}$

Symbol definition: y is depth; b is bottom width; z, z_1, and z_2 are side slopes of horizontal versus vertical; d_o is diameter; and r is radius for rounded corners of rectangle and triangle. Angle $\theta = 2\cos^{-1}[(d_o - 2y)/d_o]$, which ranges from 0 to 360 degrees.

is given as the horizontal increase with a unit vertical increase. Except for rectangular channels, hydraulic depth is not equal to the water depth y (Table 1) because the depth of water is variable in each cross section along flow lines. Hydraulic depth or mean depth (A/B) is therefore used and computed by dividing the cross-sectional area (A) by the width of the free surface (i.e., top width B). Wetted perimeter (P) is the total length of the channel boundary at a section wetted by the flowing liquid, and hydraulic radius (R) is determined by dividing channel cross-sectional area (A) by wetted perimeter (P). For very wide natural rivers (top width B is much larger than water depth D), hydraulic radius is approximately equal to mean depth D.

The flow in an open channel is uniform if the depth of flow does not vary along the length of the channel. The volumetric flow rate (discharge in ft^3/sec or m^3/sec) in a uniform open channel flow can be determined from the Continuity (conservation of mass) Equation 1 and Manning's Equation 2:

$$Q = VA \tag{1}$$

$$V = \frac{m}{n}R^{2/3}{S_0}^{1/2} \tag{2}$$

where V is the mean velocity of flow at a cross section, $m = 1.0$ (m$^{1/3}$/s) in SI units or $m = 1.49$ (ft$^{1/3}$/s) in English (British Gravitational) units, n is the Manning's roughness coefficient (dimensionless), R is the hydraulic radius, S_0 is the channel bed slope, and A is the cross-sectional area. Typical values of n for various channel conditions are given in Table 2. For channels with different roughness

Table 2. Typical Manning's n Coefficients (1)

Type of channel	Minimum	Normal	Maximum
Riveted and spiral steel	0.013	0.016	0.017
Coated cast iron	0.010	0.013	0.014
Uncoated cast iron	0.011	0.014	0.016
Galvanized wrought iron	0.013	0.016	0.017
Black wrought iron	0.012	0.014	0.015
Corrugated metal	0.021	0.024	0.030
Glass	0.009	0.010	0.013
Cement mortar	0.011	0.013	0.015
Finished concrete	0.010	0.012	0.014
Concrete culvert, straight	0.010	0.011	0.013
Concrete culvert with bends, connections	0.011	0.013	0.014
Concrete sewer with manholes, inlets, etc.	0.013	0.015	0.017
Wood stave	0.010	0.012	0.014
Clay drainage tile	0.011	0.013	0.017
Brick work	0.012	0.015	0.017
Earthen channel, straight, clean	0.017	0.020	0.025
Channel, straight with short grass, few weeds	0.022	0.027	0.033
Natural creeks and small streams (clean, straight, full stage, no rifts or deep pools)	0.025	0.030	0.033
Major streams (top width at flood stage greater than 100 ft)	0.025	—	0.060
Flood plains (pasture with high grass)	0.030	0.035	0.050

coefficients (n_i, $i = 1, \ldots, N$) over the cross section (e.g., main channel and floodplain), a composite roughness (n_e) can be calculated by Equation 3 (2,3):

$$n_e = \left(\frac{\sum_{i=1}^{N} P_i n_i^{3/2}}{\sum_{i=1}^{N} P_i} \right)^{2/3} \qquad (3)$$

For a composite channel (not simple geometry in Table 1), one should apply Manning's equation to compute discharges for each subchannel and summarize them to obtain the discharge in the whole channel.

For open channel design, channels should be classified as nonerodible channels and erodible channels. Most nonerodible channels are constructed and lined by using concrete, rip-rap, interlocking blocks, geotextiles, and vegetation, which can withstand erosion satisfactorily under all operational velocities. Unlined channels formed in natural materials (e.g., sand, gravel, sandy loam, firm soil, stiff clay) are generally erodible, especially under high velocity as encountered, e.g., during floods. To prevent the overtopping of channel lines by surface waves and surges, a *freeboard* is recommended; it is defined as the vertical distance from the top of the channel to the water surface at the design discharge. An average freeboard from 1.4 ft to 3.6 ft for discharges from 10 to 3000 ft^3/sec, respectively, is recommended by the U.S. Bureau of Reclamation (4). Number of economical factors, for example, cost for land requisition controlled by channel top width, cost for lining channel controlled by wetted perimeter and materials, and cost of excavation controlled by channel area, can affect final selection of channel geometry.

In open channel design, the dimensions of a channel are computed for uniform flow; and Manning's equation (Equation 2) is applied after channel cross section (e.g., rectangle or trapezoid) and design discharge are specified. For lined channel design, one typically designs the channel with the best hydraulic section (Table 3), where the channel section has the least wetted perimeter for a given cross-sectional area and the channel is most hydraulic efficient. The best hydraulic section may not give the most economical channel because it could result as a very wide channel with a higher cost for land requisition. The semicircle has the least perimeter among all sections with the same area, but it may not be the most practical to construct with conventional materials. If one wants to design a concrete trapezoidal channel capable of carrying 500 cfs with a channel slope of 0.0002 ft/ft. If one uses the best hydraulic section for trapezoidal channel given in Table 3 and applies Manning's equation Equation (2) as given below, water depth (y) is determined as 8.8 ft and channel width is 10.1 ft:

$$500\,cfs = \frac{1.49}{0.015}\left(\sqrt{3}y^2\right)\left(\frac{y}{2}\right)^{2/3}(0.0002)^{1/2} \qquad (4)$$

If a channel is placed in erodible material, the permissible velocity method can assure channel stability. For a trapezoidal unlined channel design, approximate permissible side slopes for various materials are given in Table 4 and maximum permissible velocities and roughness coefficients n are given in Table 5. One may use the best hydraulic section concept as the first estimate of the geometry for an erodible channel by using permissible side slope. If the velocity of the best hydraulic section at the design discharge is greater than the maximum permissible velocity, one needs to redo the design by using the maximum permissible velocity as the design velocity.

Table 4. Channel Side Slopes for Various Kinds of Materials (1)

Materials	Side Slopes (Horizontal:Vertical)
Rock	Nearly vertical
Muck and peat soils	1/4:1
Stiff clay or earth with concrete lining	1/2:1 to 1:1
Earth with stone lining	1:1
Firm Clay	1 1/2:1
Loose Sandy soil	2:1
Sandy Loam	3:1

Table 5. Maximum Permissible Velocities and Manning's *n* Values for Common Channel Materials (1,5)

Materials	V (ft/s)	Manning's n
Fine sand	1.50	0.020
Sandy loam	1.75	0.020
Silt loam	2.00	0.020
Firm loam	2.50	0.020
Stiff clay	3.75	0.025
Fine gravel	2.50	0.020
Coarse gravel	4.00	0.025

Table 3. Best Hydraulic Sections (1)

Cross Section	Area A	Wetted Perimeter P	Hydraulic Radius R	Top Width T	Mean Hydraulic Depth, D
Trapezoid (half of hexagon)	$\sqrt{3}y^2$	$2\sqrt{3}y$	$1/2\,y$	$4/3\,\sqrt{3}y$	$3/4\,y$
Rectangle (half of a square)	$2y^2$	$4y$	$1/2\,y$	$2y$	y
Triangle (half of a square)	y^2	$2\sqrt{2}y$	$1/4\,\sqrt{2}y$	$2y$	$1/2\,y$
Semicircle	$0.5\pi y^2$	πy	$1/2\,y$	$2y$	$0.25\pi y$
Parabola $T = 2\sqrt{2}y$	$4/3\sqrt{2}y^2$	$8/3\sqrt{2}y$	$1/2\,y$	$2\sqrt{2}y$	$2/3\,y$

Note: y = Maximum water depth in cross section.

If one wants to design an earth trapezoidal channel using sandy loam and capable of carrying 500 cfs with channel slope of 0.0002 ft/ft, first design the channel for a best hydraulic cross section, which means that the hydraulic radius is half of the water depth (y). For a sandy loam channel, the recommended side slope $m = 3$ and application of Manning's equation results in a water depth of 8.2 ft and a flow velocity of 2.55 ft/s, which is great the maximum permissible velocity ($V_{\max}$) for sandy loam (Table 5). One has to use the maximum permissible velocity method to redesign the channel cross section:

$$V = 1.75 = \frac{1.49}{n} R_H^{\frac{2}{3}} S^{\frac{1}{2}} = \frac{1.49}{0.02} R_H^{\frac{2}{3}} (0.0002)^{\frac{1}{2}}$$

$$\therefore \quad R_H = 2.14 \text{ ft}$$

$$\therefore \quad A = by + 3y^2 = Q/V = 500/1.75 = 285.7 \qquad (5)$$

$$\therefore \quad P = b + 2y\sqrt{m^2 + 1} = b + 2y\sqrt{10} = A/R_H = 285.7/2.14 = 133.5 \qquad (6)$$

Using a trial-and-error method to solve Equations 5 and 6, one can get $y = 2.27$ ft and $b = 119.0$ ft. The final channel cross section with a 1-ft freeboard gives:

$$P_{tot} = 119.0 + 2(3.27)\sqrt{10} = 139.7 \text{ ft } (42.58 \text{ m})$$

$$T = 119.0 + 2 \times 3 \times 3.27 = 138.6 \text{ ft } (42.25 \text{ m})$$

$$A_{tot} = 119.0(3.27) + 3(3.27)^2 = 421.2 \text{ ft}^2 \ (39.0 \text{ m}^2)$$

BIBLIOGRAPHY

1. Chow, V.T. (1959). *Open-Channel Hydraulics*. McGraw Hill, Boston, MA.
2. Einstein, H.A. (1934). Der Hydrauliche oder Profil-radius. *Schweizerishe Bauzeitung* **103**(8): 89–91.
3. Horton, R.A. (1933). Separate roughness coefficients for channel bottom and side. *Eng. News Record* **111**(2): 652–653.
4. U.S. Bureau of Reclamation (1952). Linings for Irrigation Canals.
5. Fortier, S. and Scobey, F.C. (1926). Permissible canal velocities. *Tran. ASCE* **89**: 940–954.

ORGANIC COMPOUNDS AND TRACE ELEMENTS IN FRESHWATER STREAMBED SEDIMENT AND FISH FROM THE PUGET SOUND BASIN

Dorene E. MacCoy
Robert W. Black
National Water Quality
Assessment Program
U.S. Geological Survey

As part of the National Water-Quality Assessment (NAWQA) Program, the USGS is investigating contaminants in streambed sediment and aquatic organisms and their relation to land use. One such study is being done in the Puget Sound Basin, which is located in northwestern Washington State and includes streams and rivers that drain to the Puget Sound, the Strait of Georgia, and the Strait of Juan de Fuca, but does not include marine waters. The basin encompasses 13,700 square miles; forest, urban, and agriculture are the principal land uses.

Urban stream

Agricultural stream

SUMMARY OF FINDINGS

Organochlorine Compounds—Highest concentrations of organochlorine compounds in streambed sediment were at an urban site on Thornton Creek near Seattle, where total chlordane, DDT, DDD, and DDE were found to exceed Canadian sediment quality guidelines. Concentrations are compared to Canadian guidelines because there are no sediment quality guidelines in the State of Washington.

Highest concentrations of organochlorine compounds in sculpin (a bottom-feeding fish) were found at the Thornton Creek site. Polychlorinated biphenyls (PCBs) and total DDT (DDT + DDD + DDE) exceeded New York State criteria for protection of fish-eating wildlife at Thornton Creek, and total PCB's exceeded these criteria at the West Branch Kelsey Creek at Bellevue, another urban site.

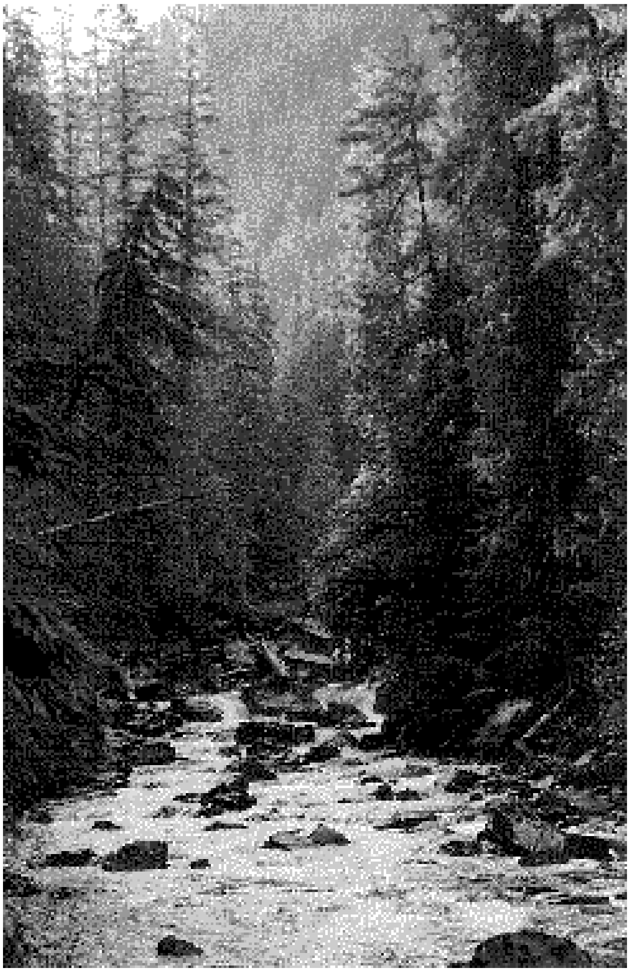

Forest stream

Concentrations are compared to New York State criteria because there are no criteria for protection of fish-eating wildlife in the State of Washington.

Effects—Elevated levels of organochlorine compounds such as DDT and PCBs are toxic to all animals and can bioconcentrate in tissue, cause tumors, and cause hormonal and behavioral problems. They can also suppress the immune and respiratory systems and cause abnormal development in aquatic species. The primary effect on aquatic communities is to reduce numbers of sensitive species, allowing species that are more resistant to contaminants to become dominant (Harte and others, 1991).

PAHs—Polycyclic aromatic hydrocarbon concentrations in streambed sediment exceeded Canadian guidelines in most urban streams and were highest at the West Branch Kelsey Creek at Bellevue.

Effects—Many PAHs, such as benzo(a)anthracene, benzo(a)pyrene, and chrysene, are carcinogenic, causing tumors in fish and other animals, and are acutely toxic to some organisms. Noncarcinogenic PAHs, such as fluoranthene, phenathrene, and pyrene, are also toxic to some organisms. The effects on aquatic organisms of the PAHs found in sediment at Kelsey Creek are unknown, but concentrations of benzo(a)pyrene as high as those observed in this study can cause precancerous tumors in fish (Eisler, 1987).

Trace elements—Concentrations of arsenic, cadmium, lead, mercury, and zinc frequently exceeded forest and reference conditions in streambed sediment and sculpin in urban streams.

Effects—Elevated levels of arsenic, cadmium, lead, mercury, and zinc may not be of concern in a naturally metal-rich region such as Puget Sound because the aquatic system has adapted to this type of environment, but excessive amounts of these elements can affect the nervous, respiratory, circulatory, and reproductive systems of aquatic organisms, as well as affect their development and feeding habits (Rand and Petrocelli, 1985).

DATA COLLECTION AND ANALYSIS

Streambed sediment and whole sculpin tissue were analyzed to assess the occurrence and distribution of contaminants and to better understand the fate of contaminants in the environment (Table 1). We collected samples in September 1995 from 18 sites, which were characterized on the basis of the predominant land use in the stream's basin—4 agricultural sites, 9 urban sites, 2 forest sites, and 3 reference sites, which are mostly forested and receive minimal impact from humans. At each site we collected the top 2–3 centimeters of streambed sediment in depositional areas; predatory bottom-feeding fish (sculpin) were collected from 17 of these sites. (See Crawford and Luoma, 1994, and Shelton and Capel, 1994, for a more complete description of the methods used.)

Fine-grained sediment and tissue accumulate trace elements and organic compounds associated with anthropogenic (human-related) activities. Sculpin are bottom-feeding fish that are not usually consumed by humans but are eaten by other fish and fish-eating wildlife. Organic compounds analyzed for were organochlorine pesticides, total PCBs, and other organic compounds (of the other organic compounds, only PAH values that exceeded Canadian guidelines are reported because they may have the most potential to harm aquatic and related organisms).

EVALUATION OF DATA

We compared our data to guidelines and criteria for organic compounds to show possible adverse effects to aquatic

Table 1. Contaminants Analyzed for in Streambed Sediment and Whole Sculpin Tissue from the Puget Sound Basin

Contaminant	Sediment[1]	Tissue
Organochlorine compounds	31 pesticides; total PCBs	26 pesticides; total PCBs
Other organic compounds	64	not analyzed[2]
Trace elements	44	22

[1] Finer than 2.0 millimeters for organic compounds, finer than 63.0 micrometers for trace elements.

[2] Tissue analysis too costly for this study.

organisms and fish-eating wildlife, and to local forest and reference conditions for selected trace elements to show possible effects of land use (Fig. 1).

Organic compounds—We compared levels of organochlorine compounds and PAHs detected in sediment to draft interim freshwater sediment quality guidelines developed by the Canadian Council of Ministers of the Environment (CCME). These guidelines were developed from toxicity and species abundance data for benthic organisms from studies throughout North America and represent total concentrations in sieved and unsieved sediment samples (CCME, 1995). The guidelines used are the threshold effects level (TEL), below which adverse effects to aquatic organisms are expected to occur **rarely**, and the probable effects level (PEL), above which adverse effects are predicted to occur **frequently**. Concentrations that exceed these guidelines may or may not have adverse effects on aquatic organisms; the comparisons should be used to indicate **potential** sediment quality problems that may warrant further study.

We compared concentrations of organochlorine compounds in sculpin to New York State Department of Environmental Conservation (NYSDEC) criteria (Newell and others, 1987). These criteria were determined from laboratory experiments using fish-eating wildlife and are considered one of the best sets of criteria for evaluating the effects of contaminated fish tissue on wildlife.

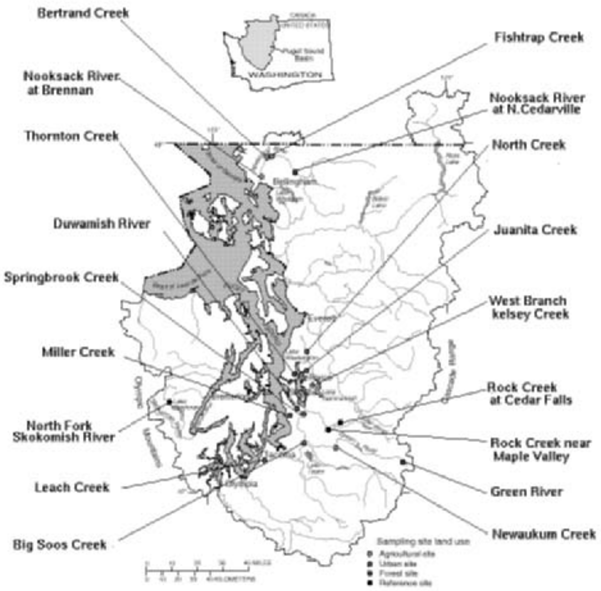

Figure 1. Puget sound basin.

Trace elements—We compared concentrations of selected trace elements in streambed sediment and sculpin

Figure 1 table. Organic compounds that exceed Canadian probable effects levels (PEL) and New York State Department of Environmental Conservation (NYSDEC)criteria, and trace elements that exceed median of forest and reference conditions at sampling sites in the Puget Sound Basin

Site	Sediment	Tissue
Bertrand Creek near Lynden	arsenic, cadmium, chromium, nickel, zinc	zinc
Nooksack River at Brennan	chromium, nickel, zinc	arsenic, chromium, lead, mercury, nickel, zinc
Thornton Creek near Seattle	arsenic, cadmium, chromium, lead, mercury, nickel, zinc, DDT	arsenic, lead, mercury, PCBs
Duwamish River at golf course at Tukwila	arsenic, cadmium, lead, zinc	arsenic, mercury
Springbrook Creek at Tukwila	arsenic, cadmium, chromium, lead, mercury, nickel, zinc	no tissue sampled
Miller Creek near Des Moines	arsenic, cadmium, chromium, lead, mercury, nickel, zinc	arsenic, lead, mercury
North Fork Skokomish River at Staircase Rapids	arsenic, chromium, nickel, zinc	arsenic
Leach Creek near Steilacoom	arsenic, cadmium, chromium, lead, nickel, zinc	arsenic, zinc
Big Soos Creek above hatchery near Auburn	arsenic, cadmium, chromium, lead, nickel	—
Fishtrap Creek at Flynn Road	arsenic, cadmium, chromium, lead, nickel, zinc	arsenic, mercury
Nooksack River at North Cedarville	nickel	—
North Creek below Penny Creek near Bothell	arsenic, cadmium, chromium, lead mercury, nickel, zinc	arsenic, mercury
Juanita Creek at La Juanita	arsenic, cadmium, chromium, lead, nickel, zinc	arsenic, cadmium, lead, mercury
West Branch Kelsey Creek at Bellevue	arsenic, cadmium, chromium, lead, mercury, nickel, zinc, PAHs	arsenic, cadmium, lead, mercury, zinc, PCBs
Rock Creek at Cedar Falls near Landsburg	arsenic, cadmium	chromium, nickel, zinc
Rock Creek near Maple Valley	cadmium, lead, zinc	mercury
Green River above Twim Camp	mercury	arsenic, nickel, zinc
Newaukum Creek near Black Diamond	cadmium, lead, zinc	—

tissue to median concentrations from the forest and reference sites. Land-use impacts may cause concentrations from the agricultural and urban sites to exceed these medians.

ORGANIC COMPOUNDS DETECTED IN STREAMBED SEDIMENT

Organochlorine pesticides were detected at 3 of the 18 sites sampled for streambed sediment: an agricultural site on Fishtrap Creek in the northern part of the basin, an urban site on Thornton Creek near Seattle, and a reference site on Rock Creek near Maple Valley. The highest concentrations were found at the urban site on Thornton Creek (Fig. 2).

PAHs were most frequently detected in streambed sediment samples from urban streams. The highest concentrations were found in the sample taken fromWest Branch Kelsey Creek at Bellevue (Table 2). (See also Tables 3–7.)

Definitions of organic compounds found in the Puget Sound Basin

DDT (dichlorodiphenyltrichloroethane) is an organochlorine insecticide banned from use in the U.S. in 1972. Total DDT refers to the sum of DDT and its breakdown products DDE and DDD.

Chlordane is an organochlorine insecticide banned from use in the 1980's. Total chlordane refers to the sum of *cis*-chlordane, *trans*-chlordane, *cis*-nonachlor and *trans*-nonachlor.

HCB (hexachlorobenzene) is a fungicide used as a seed and soil treatment, restricted from use in the 1980's.

Dieldrin is an organochlorine insecticide with restricted use in the U.S. since the1970's.

Heptachlor epoxide is a breakdown product of the organochlorine insecticide heptachlor. It was used in the U.S. until the1970's.

PCBs (polychlorinated biphenyls) are by-products of a variety of industrial products. Manufacture was stopped in the 1970's. There are over 209 breakdown products of PCBs, and total PCB refers to the sum of all forms detected.

PAHs (polycyclic aromatic hydrocarbons) are natural by-products of forest fires. Other sources include the steel and petroleum industry, the manufacture of coal tar and asphalt, power generation, burning trash, and vehicle emissions. Tons are emitted to the atmosphere and introduced to aquatic environments through oil spills and sewage discharge.

ORGANOCHLORINE COMPOUNDS DETECTED IN TISSUE

Total PCBs and/or at least 1 of 26 organochlorine pesticides were detected in tissue at 2 agricultural and 6 urban sites. The highest concentrations and greatest ranges of organochlorine compounds were detected at the Thornton Creek site (Fig. 3).

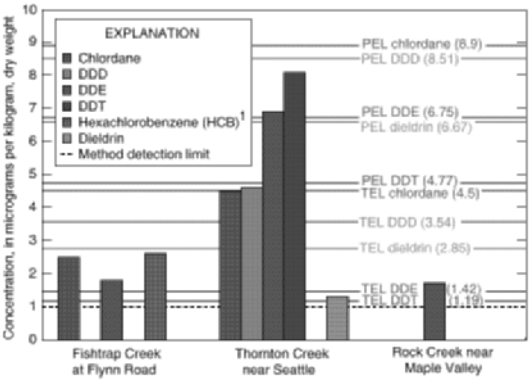

Figure 2. Concentrations of organochlorine pesticides in streambed sediment from selected sites in the Puget Sound Basin compared to Canadian criteria.(TEL, threshold effects level; PEL, probable effects level). [1]No published Canadian data for HCB.

Table 2. Concentrations of Polycyclic Aromatic Hydrocarbons (PAHs) in Streambed Sediment that Exceed Canadian Guidelines [values in micrograms per kilogram, dry weight; nd, not detected or below 50 micrograms per kilogram method detection limit; bold, above TEL; *italics*, above PEL]

Site Name	Benzo(a)anthracene[1]	Benzo(a)pyrene	Chrysene[1]	Fluoranthene[2]	Phenanthrene[2]	Pyrene[2]
Fishtrap Creek	**54**	**53**	50	91	nd	**87**
Duwamish River	**52**	Nd	56	91	nd	79
Springbrook Creek	**370**	**450**	**520**	**890**	**370**	**770**
Juanita Creek	**76**	**73**	**83**	**150**	**78**	**120**
West Branch Kelsey Creek	*680*	*1700*	*950*	*2800*	*850*	*2300*
Leach Creek	**57**	**62**	**65**	100	**51**	**94**
Miller Creek	**100**	**120**	**130**	**230**	**120**	**200**
North Creek	Nd	Nd	nd	61	nd	**56**
Thornton Creek	**220**	**310**	**270**	**470**	**200**	**410**
Rock Creek near Maple Valley	**270**	Nd	**200**	**320**	**150**	**240**
Rock Creek at Cedar Falls	50	54	63	nd	160	nd

[1] Weakly carcinogenic (Eisler, 1987).

[2] Noncarcinogenic.

Table 3. Organochlorine Compounds in Sediment from the Puget Sound Basin, September 1995 (micrograms per kilogram dry weight in less than 2 millimeter size fraction bottom material unless otherwise specified; MDL, method detection limit; nd, nondetect or below MDL; (), detection below MDL; E, detection level different from MDL, —, no published Canadian guidelines)

				Agricultural	Agricultural	Agricultural	Agricultural	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Forest	Reference	Reference	Reference
Site Type	MDL	Interim Canadian Sediment Quality Guidelines, Threshold Effects Level[1]	Interim Canadian Sediment Quality Guidelines, Probable Effects Level[1]	Bertrand Creek Near Lynden, WA	Fishtrap Creek at Flynn Road at Lynden, WA	Newaukum Creek Near Black Diamond, WA	Nooksack River at Brennan, WA	Big Soos Above Hatchery Near Auburn, WA	Duwamish River at Golf Course at Tukwila, WA	Springbrook Creek at Tukwila, WA	Juanita Creek at La Juanita, WA	West Branch Kelsey Creek at Bellevue, WA	Leach Creek near Steilacoom, WA	Miller Creek near Des Moines, WA	North Creek Below Penny Creek Near Bothell, WA	Thornton Creek Near Seattle, WA	Green River Above Twin Camp Creek near Lester, WA	Nooksack River at North Cedarville, WA	Rock Creek Near Maple Valley, WA	Rock Creek at Cedar Falls Near Landsburg, WA	North Fork Skokomish River at Staircase Rapids Near Hoodsport, WA
USGS site number				12212500	12212100	12108500	12213140	12112600	12113390	12113375	12120490	12119850	12091300	12103326	12125900	12128000	12103380	12210700	12118500	12113375	12056495
Aldrin	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Chloroneb	5.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(20)	nd	nd
Dacthal (DCPA)	5.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Dieldrin	1.0	2.85	6.67	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	1.3	nd	nd	nd	nd	nd
Endrin	2.0	2.67	62.4	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(10)	nd	nd
Heptachlor	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Heptachlor epoxide	1.0	0.6	2.74	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Hexachlorobenzene	1.0	—	—	nd	2.6	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Isodrin	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Lindane (gamma-BHC, gamma-HCH)	1.0	0.94	1.38	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mirex	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Oxychlordane	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Pentachloroanisole	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
PCB	50	34.1	277	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Toxaphene	200	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Endosulfan	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
alpha-BCH	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
beta-HCH (beta-BHC)	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(5)	nd	nd
cis-Chlordane	1.0	4.5	8.9	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	1.6	nd	nd	nd	nd	nd
cis-Nonachlor	1.0	—	—	nd	2.5E	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
cis-Permethrin	5.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(20)	nd	nd
o,p′-DDD	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(5)	nd	nd
o,p′-DDE	1.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
o,p′-DDT	2.0	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

(*continued overleaf*)

Table 3. (*Continued*)

				Agricultural	Agricultural	Agricultural	Agricultural	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Forest	Reference	Reference	Reference
Site Type	MDL	Interim Canadian Sediment Quality Guidelines, Threshold Effects Level[1]	Interim Canadian Sediment Quality Guidelines, Probable Effects Level[1]	Bertrand Creek Near Lynden, WA	Fishtrap Creek at Flynn Road at Lynden, WA	Newaukum Creek Near Black Diamond, WA	Nooksack River at Brennan, WA	Big Soos Above Hatchery Near Auburn, WA	Duwamish River at Golf Course at Tukwila, WA	Springbrook Creek at Tukwila, WA	Juanita Creek at La Juanita, WA	West Branch Kelsey Creek at Bellevue, WA	Leach Creek near Steilacoom, WA	Miller Creek near Des Moines, WA	North Creek Below Penny Creek Near Bothell, WA	Thornton Creek Near Seattle, WA	Green River Above Twin Camp Creek near Lester, WA	Nooksack River at North Cedarville, WA	Rock Creek Near Maple Valley, WA	Rock Creek at Cedar Falls Near Landsburg, WA	North Fork Skokomish River at Staircase Rapids Near Hoodsport, WA
o,p′-Methoxychlor	5.0	—	—	nd	nd	Nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
p,p′-DDD	1.0	3.54	8.51	nd	nd	Nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	4.6	nd	nd	nd	nd	nd
p,p′-DDE	1.0	1.42	6.75	nd	1.8	Nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	6.9	nd	nd	1.7E	nd	nd
p,p′-DDT	2.0	—	—	nd	nd	Nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	8.1	nd	nd	nd	nd	nd
p,p′-Methoxychlor	5.0	—	—	nd	nd	Nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(20)	nd	nd
trans-Chlordane	1.0	—	—	nd	nd	Nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	1.2	nd	nd	nd	nd	nd
trans-Nonachlor	1.0	—	—	nd	nd	Nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	1.7	nd	nd	nd(5)	nd	nd
trans-Permethrin	5.0	—	—	nd	nd	Nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(8)	nd	nd	nd(20)	nd	nd
organic carbon (grams per kilogram)				8.1	20	8.5	3.2	19	3.9	20	4.5	13	3.1	4.8	11	9.1	5.9	12	250[2]	14	11
inorganic carbon (grams per kilogram)				nd	nd	Nd	nd	nd	nd	nd	nd	nd	0.2	nd	nd	0.9	nd	nd	nd	nd	nd
organic + inorganic carbon (grams per kilogram)				8.1	20	8.5	3.2	19	3.9	20	4.5	13	3.3	4.8	11	10	5.9	12	250	14	11
particle size (percent < 62 micron)				5	9	7	8	19	5	26	5	10	5	9	4	7	4	33	17	4	6
duplicate size fraction (percent <62 micron)					8											6					

---

[1] CCME(Canadian Council of Ministers of the Environment), 1995, Protocol for the derivation of Canadian sediment quality guidelines for the protection of aquatic life: Winnipeg, Manitoba, report CCME EPC-98E, Task Group on Water Quality Guidelines, 38 p.

[2] The value for organic carbon is very high for this sample. The value has been verified by the U.S. Geological Survey National Laboratory, but may not represent typical conditions in this stream.

Table 4. Metals in Streambed Sediment from Sites in the Puget Sound Basin, September 1995 (micrograms per gram unless otherwise specified, in less than 63 micron size fraction dry weight; MDL, method detection limit; nd, not detected or below MDL; (), detection below MDL)

		Agricultural	Agricultural	Agricultural	Agricultural	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Forest	Reference	Reference	Reference
Site Type	MDL	Bertrand Creek Near Lynden, WA	Fishtrap Creek at Flynn Road at Lynden, WA	Newaukum Creek Near Black Diamond, WA	Nooksack River at Brennan, WA	Big Soos Creek Above Hatchery Near Auburn, WA	Duwamish River At Golf Course at Tukwila, WA	Juanita Creek at La Juanita, WA	West Branch Kelsey Creek at Bellevue, WA	Leach Creek Near Steilacoom, WA	Miller Creek Near Des Moines, WA	North Creek Below Penny Creek Near Bothell, WA	Thornton Creek Near Seattle, WA	Green River Above Twin Camp Creek Near Lester, WA	Nooksack River at North Cedarville, WA	Rock Creek Near Maple Valley, WA	Rock Creek at Cedar Falls near Landsburg, WA	North Fork Skokomish River at Staircase Rapids Near Hoodsport, WA
USGS site number		12212500	12212100	12108500	12213140	12112600	12113390	12120490	12119850	12091300	12103326	12125900	12128000	12103380	12210700	12118500	12117695	12056495
Aluminum (%)	0.005	4.5	5.4	6.2	8.4	6	7.1	7.3	6.9	6.3	7.2	5.8	6.7	7.5	8.7	3.9	7.2	7.9
Antimony	0.1	1	1	0.5	0.7	1	1	1	1	1	2	2	3	0.5	0.6	1	0.6	0.6
Arsenic	0.1	15	21	6.2	nd	23	12	12	16	10	15	33	19	2.5	7	5.3	7.9	7.7
Barium	1	400	600	370	630	410	430	520	530	420	500	490	510	390	580	230	370	670
Beryllium	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	2
Bismuth	10	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Cadmium	0.1	0.6	0.9	0.4	0.2	0.3	0.6	0.4	1.5	0.7	0.8	0.7	1.4	0.2	0.1	0.7	0.3	0.2
Calcium (%)	0.005	1.6	1.9	1.8	2.8	2	2.2	2.2	2.1	2.5	2.3	2.1	2.5	1.9	3.4	2.2	1.9	1.1
Cerium	4	30	32	35	44	34	36	39	32	35	36	27	35	54	40	15	36	88
Chromium	1	78	95	50	140	99	57	110	110	93	110	120	120	26	80	54	65	110
Cobalt	1	23	26	20	24	16	17	17	18	15	16	15	18	17	21	8	17	22
Copper	1	26	38	45	50	40	44	29	82	29	44	28	58	31	40	42	44	48
Europium	2	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Gallium	4	12	10	13	17	12	15	14	14	11	14	12	13	19	18	8	14	19
Gold	8	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Holmium	4	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Iron (%)	0.005	21	8.6	4	4.8	4.1	4.9	3.8	3.9	3.2	3.6	4.1	4.2	4.4	4.4	1.7	3.7	4.4
Lanthanum	2	14	15	15	21	15	17	17	15	16	17	12	16	23	19	10	18	40
Lead	4	12	30	21	8	18	33	43	110	71	73	46	190	7	nd	23	13	13
Lithium	2	10	20	20	30	20	20	20	30	10	20	20	20	30	20	20	20	50
Magnesium (%)	0.005	0.75	0.98	0.87	2.4	1.2	1.1	1.2	1.1	0.9	1.1	1	1.3	0.78	1.9	0.57	0.87	1.4
Manganese	4	1300	9200	3200	800	1500	1700	1500	2300	3800	930	2200	2100	850	750	570	1000	940
Molybdenum	2	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	0.07
Mercury	0.02	0.06	0.09	0.1	0.06	0.08	0.1	0.08	0.34	0.08	0.15	0.18	0.53	0.34	0.06	0.11	0.1	nd
Neodymium	4	15	15	17	20	16	18	18	15	15	16	12	17	26	20	9	17	40

(*continued overleaf*)

Table 4. (*Continued*)

		Agricultural	Agricultural	Agricultural	Agricultural	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Forest	Reference	Reference	Reference
Site Type	MDL	Bertrand Creek Near Lynden, WA	Fishtrap Creek at Flynn Road at Lynden, WA	Newaukum Creek Near Black Diamond, WA	Nooksack River at Brennan, WA	Big Soos Creek Above Hatchery Near Auburn, WA	Duwamish River At Golf Course at Tukwila, WA	Juanita Creek at La Juanita, WA	West Branch Kelsey Creek at Bellevue, WA	Leach Creek Near Steilcoom, WA	Miller Creek Near Des Moines, WA	North Creek Below Penny Creek Near Bothell, WA	Thornton Creek Near Seattle, WA	Green River Above Twin Camp Creek Near Lester, WA	Nooksack River at North Cedarville, WA	Rock Creek Near Maple Valley, WA	Rock Creek at Cedar Falls near Landsburg, WA	North Fork Skokomish River at Staircase Rapids Near Hoodsport, WA
Nickel	2	58	70	30	120	60	31	63	57	37	57	52	65	19	61	23	36	54
Niobium	4	nd	4	8	11	9	9	9	9	8	9	9	10	14	10	4	10	15
Phosphorus (%)	0.005	0.21	0.26	0.21	0.13	0.18	0.19	0.12	0.16	0.09	0.09	0.28	0.18	0.12	0.12	0.16	0.13	0.14
Potassium (%)	0.05	0.5	0.67	0.61	1.3	0.7	0.89	0.81	0.82	0.71	0.82	0.62	0.79	1	1.2	0.41	0.59	1.6
Scandium	2	11	13	14	19	13	15	16	15	13	15	12	15	20	17	7	13	17
Selenium	0.1	0.9	1.4	1.2	0.7	1	0.4	0.4	0.6	0.3	0.4	1.6	0.6	0.9	0.7	4.6	1.1	1.4
Silver	0.1	0.3	0.3	0.3	0.1	0.3	0.3	0.2	0.5	0.2	0.3	0.4	2.2	0.2	0.1	0.4	0.3	0.2
Sodium (%)	0.005	1.1	1.3	1	2.1	1.5	1.6	2	1.6	2	2	1.5	1.9	1.3	2.2	0.85	1.3	1.6
Strontium	2	230	240	210	480	240	300	310	260	310	330	250	310	180	630	190	220	190
Sulfur (%)	0.05	0.14	0.17	0.16	0.2	0.13	0.1	0.07	0.13	0.05	0.05	0.17	0.14	0.07	0.32	0.28	0.08	0.1
Tantalum	40	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Thorium	4	4.1	nd(3.1)	7.3	3.9	nd(4.5)	8.8	4.2	4.8	5.2	3.5		3.8	4.9	4.1		4.6	11
Tin	10	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Titanium (%)	0.005	0.32	0.37	0.38	0.53	0.37	0.45	0.44	0.42	0.42	0.44	0.33	0.44	0.56	0.5	0.18	0.44	0.49
Uranium	0.05	1.47	1.99	1.35	2.4	1.85	1.68	1.85	1.91	1.78	1.79		1.94	1.86	1.81		1.61	3.14
Vanadium	2	65	120	100	160	110	110	120	120	110	120	120	120	100	140	54	110	130
Yttrium	2	18	17	20	20	18	19	18	16	17	17	15	17	35	19	14	15	26
Ytterbium	1	2	2	2	2	2	2	2	2	2	2	1	2	3	2	1	1	2
Zinc	4	130	250	150	99	93	140	140	470	180	190	180	330	94	85	95	88	110
Organic + inorganic carbon (grams per kilogram)	0.01	5.68	6.62	8.21	1.19	6.66	3.68	2.9	6.37	2.44	3.07	10.7	5.11	5.43	0.84	22.7[1]	8.89	3.91
Inorganic carbon (grams per kilogram)	0.01	0.17	0.17	0.04	0.01	0.03	0.02	0.02	0.03	0.03	0.01	0.06	0.03	0.03	0.02	0.05	0.05	0.01
Organic carbon (grams per kilogram)	0.01	5.51	6.45	8.17	1.18	6.63	3.66	2.88	6.34	2.41	3.06	10.1	5.08	5.4	0.82	22.2	8.84	3.9

[1] The value for organic carbon is very high for this sample. The value has been verified by the U.S. Geological Survey National Laboratory, but may not represent typical conditions in this stream.

Table 5. Semivolatile Organic and Polychlorinated Biphenyl (PCB) Compounds in Streambed Sediment from the Puget Sound Basin, September 1995 (micrograms per kilogram unless otherwise specified in less than 2 millimeter size fraction, dry weight; MDL, method detection limit; nd, not detected or below MDL; (), detection below MDL; E, detection level different from MDL,—, no published Canadian guidelines)

Site Type				Agricultural	Agricultural	Agricultural	Agricultural	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Forest	Reference	Reference	Reference
Compound	MDL	Canadian Interim Sediment Quality Guidelines for Polycyclic Aromatic Hydrocarbons (Threshold Effect Levels)^1	Canadian Interim Sediment Quality Guidelines (Probable Effect Levels)^1	Bertrand Creek Near Lynden, WA	Fishtrap Creek at Flynn Road at Lynden, WA	Newaukum Creek Near Black Diamond, WA	Nooksack River at Brennan, WA	Big Soos Creek Above Hatchery Near Auburn, WA	Duwamish River At Golf Course At Tukwila, WA	Juanita Creek at La Juanita, WA	West Branch Kelsey Creek at Bellevue, WA	Leach Creek Near Steilcoom, WA	Miller Creek Near Des Moines, WA	North Creek Below Penny Creek Near Bothell, WA	Thornton Creek Near Seattle, WA	Green River Above Twin Camp Creek Near Lester, WA	Nooksack River at North Cedarville, WA	Rock Creek Near Maple Valley, WA	Rock Creek At Cedar Falls Near Landsburg, WA	North Fork Skokomish River at Staircase Rapids Near Hoodsport, WA
USGS site number				12212500	12212100	12108500	12213140	12112600	12113390	12120490	12119850	12091300	12103326	12125900	12128000	12103380	12210700	12118500	12117695	12056500
1,2,4-Trichlorobenzene	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
Dichlorobenzene	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
1,2-Dimethylnaphthalene	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	13E	nd(250)	28E	nd
1,3-Dichlorobenzene	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
1,4-Dichlorobenzene	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
1,6-Dimethylnaphthalene	50	—	—	nd	17E	nd	19E	nd	23E	nd	15E	nd	nd	nd	15E	nd	23E	93	120	nd
1-Methyl-9H-fluorene	50	—	—	nd	16E	nd	26E	nd	nd	nd	36E	nd	nd	nd	35E	nd	19E	nd(250)	25E	nd
1-Methylphenanthrene	50	—	—	nd	13E	nd	16E	nd	28E	27E	70	21E	25E	16E	31E	nd	12E	nd(250)	100	nd
1-Methylpyrene	50	—	—	nd	30E	nd	nd	nd	35E	37E	100	37E	33E	nd	71	nd	nd	nd(250)	48	nd
2,2′-Biquinoline	50	—	—	38E	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	29
2,3,6-Trimethylnaphthalene	50	—	—	nd	11E	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	13E	nd(250)	42E	nd
2,4-Dinitrotoluene	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
2,6-Dimethylnaphthalene	50	—	—	23E	33E	26E	23E	41E	29E	nd	25E	nd	nd	nd	25E	nd	43E	110	79	nd
2,6-Dinitrotoluene	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
2-Chloronaphthalene	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
2-Chlorophenol	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
Pentachlorophenol	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	21E	nd
2-Methylanthracene	50	—	—	nd	22E	nd	nd	nd	35E	36E	58	27E	30E	nd	50E	nd	30E	140	25E	nd
3,5-Xylenol	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	22E	nd(250)	nd	nd
4-HCY Penphenanthrene	50	—	—	nd	13E	nd	nd	nd	nd	28E	150	23E	33E	27E	55	nd	nd	95	20E	nd

(*continued overleaf*)

Table 5. (*Continued*)

Site Type				Agricultural	Agricultural	Agricultural	Agricultural	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Forest	Reference	Reference	Reference
Compound	MDL	Canadian Interim Sediment Quality Guidelines for Polycyclic Aromatic Hydrocarbons (Threshold Effect Levels)[1]	Canadian Interim Sediment Quality Guidelines (Probable Effect Levels)[1]	Bertrand Creek Near Lynden, WA	Fishtrap Creek at Flynn Road at Lynden, WA	Newaukum Creek Near Black Diamond, WA	Nooksack River at Brennan, WA	Big Soos Creek Above Hatchery Near Auburn, WA	Duwamish River At Golf Course At Tukwila, WA	Juanita Creek at La Juanita, WA	West Branch Kelsey Creek at Bellevue, WA	Leach Creek Near Steilcoom, WA	Miller Creek Near Des Moines, WA	North Creek Below Penny Creek Near Bothell, WA	Thornton Creek Near Seattle, WA	Green River Above Twin Camp Creek Near Lester, WA	Nooksack River at North Cedarville, WA	Rock Creek Near Maple Valley, WA	Rock Creek At Cedar Falls Near Landsburg, WA	North Fork Skokomish River at Staircase Rapids Near Hoodsport, WA
4-Bromophenyl-phenylether	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
4-Chloro M-Cresol	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
4-Chlorophenyl-phenylether	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
Acenaphthene	50	—	—	nd	17E	nd	nd	nd	nd	nd	41E	nd	21E	nd	19E	nd	nd	nd(250)	18E	nd
Acenaphthylene	50	—	—	nd	22E	nd	nd	nd	nd	nd	38E	nd	nd	nd	39E	nd	nd	nd(250)	20E	nd
Acridine	50	—	—	nd	nd	nd	nd	nd	32E	36E	130	28E	35E	31E	54	nd	nd	nd(250)	nd	nd
Anthracene	50	—	—	nd	28E	22E	nd	23E	28E	40E	160	25E	36E	29E	71	nd	nd	84	9E	nd
9,10-Anthraquinone	50	—	—	nd	30E	nd	nd	nd	38E	45E	310	33E	56	29E	89	nd	nd	nd(250)	48E	nd
Azobenzene	50	—	—	nd	21E	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	24E	nd(250)	36E	nd
Benzo(a)anthracene	50	31.7	385	nd	54	26E	35E	31E	52	76	680	57	100	35E	220	nd	37E	270	50E	nd
Benzo(a)pyrene	50	31.9	782	nd	53	nd	nd	27E	45E	73	1700	62	120	36E	310	nd	nd	nd(300)	54	nd
Benzo(b)fluoranthene	50	—	—	nd	66	nd	nd	41E	65	92	970	84	120	49E	260	nd	nd	nd(340)	55	nd
Benzo(c)quinoline	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
Benzo(g,h,i)perylene	50	—	—	nd	34E	nd	nd	nd	37E	58E	440	43E	62E	35E	170	nd	nd	nd(250)	nd	nd
Benzo(k)fluoranthene	50	—	—	nd	41E	22E	nd	25E	50	74	790	57	110	34E	230	nd	nd	nd(230)	34E	nd
bis(2-Ethylhexyl) Phthalate	50	—	—	53	180E	67	61E	55	110	120	2400E	270	220	93E	990E	55	74E	1100E	51E	84
Butylbenzyl Phthalate	50	—	—	44E	44E	43E	45E	39E	44E	46E	87	42E	90	45E	110	46E	39E	nd(250)	35E	50
C8-Alkylphenols	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
Carbazole	50	—	—	nd	15E	nd	nd	20E	29E	38E	180	25E	38E	27E	63	nd	nd	nd(250)	12E	nd
Chrysene	50	57.1	862	24E	50	15E	12E	25E	56	83	950	65	130	31E	270	26E	11E	200	63	nd
Di-n-butyl Phthalate	50	—	—	73E	41E	44E	50E	45E	60	62	59	45E	52	48E	58	72E	42E	300	38E	53
Di-n-octyl Phthalate	50	—	—	nd	nd	nd	nd	nd	nd	nd	140E	55E	nd	nd	nd	nd	55E	nd(250)	53E	nd
Dibenzo(a,h)anthracene	50	—	—	nd	nd	nd	nd	nd	nd	nd	280	nd	36E	nd	93	nd	nd	nd(250)	nd	nd
Dibenzothiophene	50	—	—	nd	17E	nd	20E	nd	nd	20E	63	21E	25E	29E	38E	nd	nd	nd(250)	25E	nd

Table 5. (*Continued*)

Site Type				Agricultural	Agricultural	Agricultural	Agricultural	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Forest	Reference	Reference	Reference
Compound	MDL	Canadian Interim Sediment Quality Guidelines for Polycyclic Aromatic Hydrocarbons (Threshold Effect Levels)[1]	Canadian Interim Sediment Quality Guidelines (Probable Effect Levels)[1]	Bertrand Creek Near Lynden, WA	Fishtrap Creek at Flynn Road at Lynden, WA	Newaukum Creek Near Black Diamond, WA	Nooksack River at Brennan, WA	Big Soos Creek Above Hatchery Near Auburn, WA	Duwamish River At Golf Course At Tukwila, WA	Juanita Creek at La Juanita, WA	West Branch Kelsey Creek at Bellevue, WA	Leach Creek Near Steilcoom, WA	Miller Creek Near Des Moines, WA	North Creek Below Penny Creek Near Bothell, WA	Thornton Creek Near Seattle, WA	Green River Above Twin Camp Creek Near Lester, WA	Nooksack River at North Cedarville, WA	Rock Creek Near Maple Valley, WA	Rock Creek At Cedar Falls Near Landsburg, WA	North Fork Skokomish River at Staircase Rapids Near Hoodsport, WA
Diethyl Phthalate	50	—	—	27E	18E	24E	29E	25E	20E	20E	28E	24E	25E	27E	29E	28E	24E	150	16E	23
Dimethyl Phthalate	50	—	—	nd	9E	nd	8E	nd	19E	nd	nd	nd	nd	nd	10E	nd	9E	nd(250)	50	nd
Fluoranthene	50	111	2355	34E	91	28E	16E	42E	91	150	2800	100	230	61	470	nd	16E	320	36E	nd
9H-Fluorene	50	—	—	50	20E	50	25E	23E	17E	21E	69	23E	28E	27E	39E	nd	17E	130	18E	nd
Hexachlorobenzene	50	—	—	nd(1)	2.6	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd
Indeno[1,2,3-cd]pyrene	50	—	—	nd	nd	nd	nd	33E	52	79	1600E	58	92	37E	300E	nd	nd	nd(250)	nd	nd
Isophorone	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
Isoquinoline	50	—	—	nd	nd	nd	19E	nd	nd	26E	nd	nd	nd	nd	nd	nd	nd	97	nd	nd
N-Nitroso-Diphenylamine	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
N-Nitroso-Di-n-Propyl Amine	50	—	—	nd	20E	nd	nd	nd	nd	nd	36E	nd	nd	nd	36E	nd	nd	nd(250)	29E	nd
Naphthalene	50	—	—	nd	nd(5)	nd	14E	nd	nd	nd	11E	nd	nd	nd	nd(5)	nd	10E	71	30E	nd
Nitrobenzene	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
Pentachloroanisole	50	—	—	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd(1)	nd
Pentachloronitrobenzene	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
Phenanthrene	50	41.9	515	13E	47E	21E	16E	28E	46E	78	850	51	120	28E	200	14E	20E	150	160	nd
Phenanthridine	50	—	—	nd	nd	nd	nd	nd	nd	30E	61	nd	26E	nd	41E	nd	nd	nd(250)	21E	nd
Phenol	50	—	—	31E	14E	23E	9E	25E	45E	31E	30E	17E	16E	26E	29E	31E	8E	120	11E	18
Pyrene	50	53.0	875	33E	87	26E	15E	37E	79	120	2300	94	200	56	410	nd	17E	240	39E	nd
Quinoline	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
bis(2-Chloroethoxy)methane	50	—	—	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(250)	nd	nd
p-Cresol	50	—	—	nd	nd	150	82	330	64	62	34E	28E	28E	31E	35E	nd	360	130	23E	nd
Organic Carbon (grams per kilogram)				8.1	20	8.5	3.2	19	3.9	4.5	13	3.1	4.8	11	9.1	5.9	12	250[2]	14	11
Inorganic Carbon (grams per kilogram)				nd	nd	nd	nd	nd	nd	nd	nd	0.2	nd	nd	0.9	nd	nd	nd	nd	nd
Organic + inorganic Carbon (grams per kilogram)				8.1	20	8.5	3.2	19	3.9	4.5	13	3.3	4.8	11	10	5.9	12	250	14	11

---

[1] CCME (Canadian Council of Ministers of the Environment), 1995, Protocol for the derivation of Canadian sediment quality guidelines for the protection of aquatic life: Winnipeg, Manitoba, report CCME EPC-98E, Task Group on Water Quality Guidelines, 38 p.

[2] The value for organic carbon is very high for this sample. The value has been verified by the U.S. Geological Survey National Laboratory, but may not represent typical conditions in this stream.

Table 6. Organochlorine Compounds in Whole Sculpin Tissue from Sites in the Puget Sound Basin, September 1995 (micrograms per kilogram wet weight unless otherwise specified; MDL, method detection limit; nd, nondetect or below MDL; E, estimated value MDL; (), detection different from MDL)

Site Type	MDL	Agricultural	Agricultural	Agricultural	Agricultural	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Forest	Reference	Reference	Reference
		Bertrand Creek Near Lynden, WA	Fishtrap Creek at Flynn Road at Lynden, WA	Newaukum Creek Near Black Diamond, WA	Nooksack River at Brennan, WA	Big Soos Creek Above Hatchery Near Auburn, WA	Duwamish River At Golf Course At Tukwila, WA	Juanita Creek At La Juanita, WA	West Branch Kelsey Creek At Bellevue, WA	Leach Creek Near Steilacoom, WA	Miller Creek Near Des Moines, WA	North Creek Below Penny Creek Near Bothell, WA	Thornton Creek Near Seattle, WA	Green River Above Twin Camp Creek Near Lester, WA	Nooksack River At North Cedarville, WA	Rock Creek Near Maple Valley, WA	Rock Creek At Cedar Falls Road Near Landsburg, WA	North Fork Skokomish River At Staircase Rapids Near Hoodsport, WA
USGS site number		12212500	12212100	12108500	12213140	12112600	12113390	12120490	12119850	12091300	12103326	12125900	12128000	12103380	12210700	12118500	12117695	12056500
Aldrin	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
DCPA	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Dieldrin	5	nd	7.2	nd	nd	nd	nd	10	9.9	nd	7.9	nd	27	nd	nd	nd	nd	nd
Endrin	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Heptachlor	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Heptachlor epoxide	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	6.9	nd	nd	nd	nd	nd
Hexachlorobenzene	5	nd	nd	nd	nd	nd	nd	7.1	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Lindane (gamma HCH)	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mirex	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Oxychlordane	5	nd	nd	nd	nd	nd	nd	9	nd	nd	nd	nd	12	nd	nd	nd	nd	nd
Pentachloroanisole	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Total PCB	50	nd	nd	nd	nd	nd	nd	110	120	nd	nd	nd	310	nd	nd	nd	nd	nd
Toxaphene	200	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alpha-BHC	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Beta-BHC	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
cis-Chlordane	5	nd	11	nd	nd	nd	nd	14	12	6.2	9.1	nd	30	nd	nd	nd	nd	nd
cis-Nonachlor	5	nd	6.4	nd	nd	nd	nd	5.8	nd	nd	nd	nd	12	nd	nd	nd	nd	nd
Delta-BCH	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
o,p′-DDD	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
o,p′-DDE	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
o,p-Methoxychlor	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
p,p′-DDD	5	nd	8.1	nd	nd	nd	nd	nd	17E	nd	nd	nd	55E	nd	nd	nd	nd	nd
p,p′-DDE	5	30	27	nd	nd	nd	nd	18	18	5	7.5	5.1	97	nd	nd	nd	nd	nd
p,p′-DDT	5	nd(6.4)	21	nd	nd	nd	nd	nd(5.8)	nd	nd	7.5	nd	64	nd	nd	nd	nd	nd
p,p-Methoxychlor	5	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
trans-Chlordane	5	nd	5.2	nd	nd	nd	nd	nd	7.2	nd	nd	nd	14	nd	nd	nd	nd	nd
trans-Nonachlor	5	nd	14	nd	nd	nd	nd	15	14	6	13	nd	32	nd	nd	nd	nd	nd
Sample weight		10	10	10	10	10	10	10.1	10	10	10	9.9	10.2	10	10	10	10	10
Lipids %		4.2	3	2.1	2	2.3	4.2	9	4.1	4.2	3.6	2.4	5.3	3.1	7.6	2.1	2.4	4.1

Table 7. Metals in Whole Sculpin Tissue from Sites in the Puget Sound Basin, September 1995 (microgram per gram or ppm dry weight; MDL, method detection limit; nd, nondetect or below MDL; (), detection different from MDL)

		Agricultural	Agricultural	Agricultural	Agricultural	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Urban	Forest	Reference	Reference	Reference
Site Type	MDL	Bertrand Creek Near Lynden, WA	Fishtrap Creek At Flynn Road At Lynden, WA	Newaukum Creek Near Black Diamond, WA	Nooksack River at Brennan, WA	Big Soos Creek Above Hatchery Near Auburn, WA	Dwamish River at Golf Course at Tukwila, WA	Juanita Creek at La Juanita, WA	West Branch Kelsey Creek at Bellevue, WA	Leach Creek Near Steilacoom, WA	Miller Creek Near Des Moines, WA	North Creek Below Penny Creek Near Bothell, WA	Thornton Creek Near Seattle, WA	Green River Above Twin Camp Creek Near Lester, WA	Nooksack River at North Cedarville, WA	Rock Creek Near Maple Valley, WA	Rock Creek at Cedar Falls Road Near Lands-burg, WA	North Fork Skokomish River at Staircase Rapids Near Hoodsport, WA
USGS site number		12212500	12212100	12108500	12213140	12112600	12113390	12120490	12119850	12091300	12103326	12125900	12128000	12103380	12210700	12118500	12117695	12056500
Arsenic	0.2	nd	0.4	nd	0.4	0.3	1.1	0.5	0.7	0.4	0.4	0.5	0.6	0.5	0.3	0.2	0.3	0.5
Barium	0.1	2.3	3.7	2.6	8.4	2.4	2.4	2.4	1.3	1.3	1.9	5.1	4.1	2.9	6.5	3.2	9.2	3.2
Beryllium	0.2	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(0.3)	nd	nd	nd	nd	nd	nd
Boron	0.2	0.7	1	0.5	0.9	0.4	1.1	2.1	1.2	0.9	1.6	1.2	1.6	1	2.7	0.6	0.4	0.4
Cadmium	0.2	nd	nd	nd	nd	nd	nd	0.3	0.2	nd	nd	nd(0.3)	nd	nd	nd	nd	nd	nd
Chromium	0.5	1.3	1.4	1.7	2.3	1.4	1.5	1.1	1	1	0.9	1.4	1.3	1.7	1.4	1.7	2.1	1.4
Cobalt	0.2	nd	0.2	0.3	0.7	nd	0.2	0.2	0.2	nd	nd	nd(0.3)	nd	0.3	0.8	0.2	0.4	0.3
Copper	0.5	2.9	3.5	2.2	3.4	2	2	1.9	6.8	1.8	2.3	1.6	2.1	2.3	2.1	2.2	2.1	1.2
Iron	1	167	160	79.7	250	84.4	153	45.5	46.7	54.9	73.2	40.3	71.3	71.7	113	29.6	51.7	156
Lead	0.2	nd	nd	nd	1.1	nd	nd	0.2	0.3	nd	0.2	nd(0.3)	1	nd	nd	nd	nd	nd
Manganese	0.1	11.2	32.6	21.6	18.6	16.8	13.6	28.5	21.1	57.6	7.3	32.1	15.5	6.1	17	9.6	10.1	7.8
Mercury	0.1	0.1	0.2	0.1	0.33	0.14	0.66	0.76	0.23	0.16	0.31	0.31	0.25	nd	0.18	0.4	0.1	nd(0.1)
Molybdenum	0.2	nd	nd	nd	0.3	nd	0.2	nd	0.2	nd	nd	nd(0.3)	nd	0.2	0.2	nd	0.2	nd
Nickel	0.2	1	1.1	1.3	3.7	0.9	1.3	1	0.8	0.7	0.8	1.3	1.1	1.8	1.7	1.3	2.1	0.8
Selenium	0.1	1.2	2.1	2.6	3.6	1	1.3	2.4	1.2	2.7	0.7	2.7	1.5	2.7	4.8	3.6	2.4	4.8
Silver	0.2	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(0.3)	nd	nd	nd	nd	nd	nd
Strontium	0.1	31.3	49.8	65.2	98.9	57.5	44.4	38.2	27.4	32.4	32.4	75	43.7	28.1	66.6	79.2	123	105
Uranium	0.2	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd(0.3)	nd	nd	nd	nd	nd	nd
Vanadium	0.1	nd	0.4	0.5	1.7	0.3	1	0.5	0.2	0.5	0.3	1	0.7	0.6	0.5	0.2	0.5	0.4
Zinc	0.5	129	41	61.5	71	59.1	51	66.5	128	71.4	49.6	58.6	58.1	68.1	58	66.7	72.4	55.9
% water		78.3	76.9	79.9	80.1	78.6	77.4	72.8	74.9	77.6	75.8	77.3	75.7	78.5	74.2	76.5	78.2	75

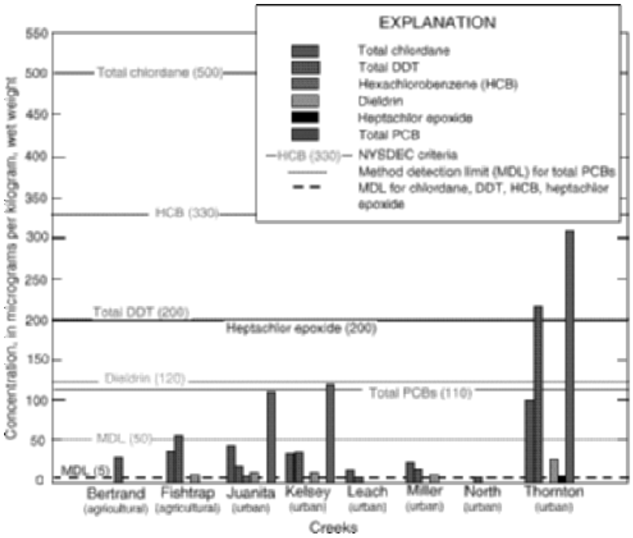

Figure 3. Organochlorine compounds in whole sculpin tissue from selected creeks in the Puget Sound Basin compared to New York State Department of Environmental Conservation (NYSDEC) criteria.

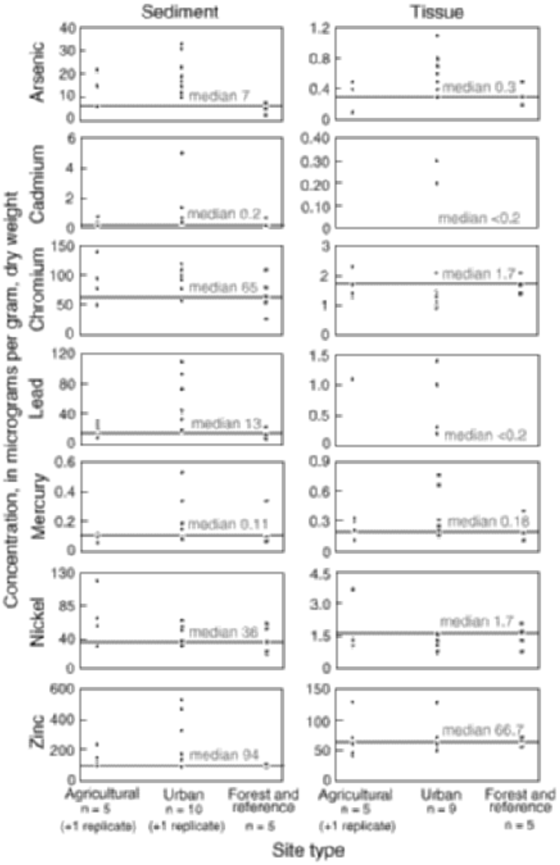

Figure 4. Concentrations of selected trace elements detected in streambed sediment and whole sculpin tissue collected in the Puget Sound Basin compared to the median of forest and reference site data. (n, number of samples; +1 replicate, one replicate sample data included; <, less than).

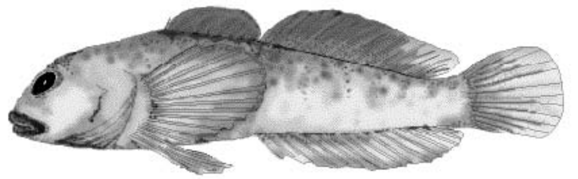

TRACE ELEMENTS DETECTED IN STREAMBED SEDIMENT AND WHOLE SCULPIN TISSUE

Concentrations of arsenic, cadmium, chromium, lead, mercury, nickel, and zinc were elevated in streambed sediment and sculpin from agricultural and urban sites compared to concentrations from the forest and reference sites. Arsenic and the heavy metals cadmium, lead, mercury, and zinc had the highest concentrations and the greatest range of concentrations at urban sites compared to those from the agricultural and the combined forest and reference sites, indicating possible enrichment of these elements in the urban areas (Fig. 4). The concentrations detected do not necessarily have negative impacts on the environment, but do suggest that land use may have led to increased levels of these elements.

READING LIST

CCME (Canadian Council of Ministers of the Environment), 1995, *Protocol for the derivation of Canadian sediment quality guidelines for the protection of aquatic life: Winnipeg*, Manitoba, report CCME EPC-98E, Task Group on Water Quality Guidelines, p. 38.

Crawford, J.Kent, and Luoma, Samuel N., 1994, *Guidelines for studies of contaminants in biological tissues for the National Water-Quality Assessment Program: U.S.* Geological Survey Open-File Report 92-494, p. 69.

Eisler, R., 1987, *Polycyclic aromatic hydrocarbon hazards to fish, wildlife, and invertebratesa synoptic review: Laurel, Md., Patuxent Wildlife Research Center, U.S. Fish and Wildlife Service*, Biological Report 85 (1.11), p. 81.

Harte, John, Holdren, Cheryl, Schneider, Richard, and Shirley, Christine, 1991, *Toxics A to Z—a guide to everyday pollution hazards*: Berkeley, Calif., University of California Press, p. 479.

Newell, Arthur J., Johnson, David W. and Allen, Laurie K. 1987, *Niagara River biota contamination project-Fish flesh criteria for piscivorous wildlife: New York State Department of Environmental Conservation, Division of Fish and Wildlife, Bureau of Environmental Protection*, Technical Report 87-3, p. 182.

Rand, G.M., and Petrocelli, S.R., 1985, *Fundamentals of aquatic toxicology methods and applications*: New York, Hemisphere Publishing Corp., p. 666.

Shelton, L.R., and Capel, P.D., 1994, *Guidelines for collecting and processing samples of stream bed sediment for analysis of trace elements and organic contaminants for the National Water-Quality Assessment Program: U.S.* Geological Survey Open-File Report 94-458, p. 20.

IMPERVIOUS COVER—PAVING PARADISE

Sandra Bird
U.S. Environmental
Protection Agency

THE IMPORTANCE OF IMPERVIOUS COVER

Nonpoint source pollution is pollution from diffuse sources such as urban/suburban areas and farmlands; this is now recognized as the primary threat to water quality in the United States. As urban and suburban development increases, the amount of land covered with impervious surfaces—areas where infiltration of water into the underlying soil is prevented—also increases. Roadways and rooftops account for the majority of this impervious area. Research in recent years has consistently shown a strong relationship between the percentage of impervious cover in a watershed and the health of the receiving stream. Scientists generally agree that stream degradation consistently occurs at even relatively low levels of imperviousness (10 to 20%). Increased impervious surfaces alter stream hydrology resulting in lower flows during droughts and higher peak flows during floods. Roadways and other impervious areas channel pollutants directly into streams without their being processed during transport through the soil. With advance planning and identification of at-risk watersheds, total impervious cover can be reduced during development within a watershed, and steps can be taken to mitigate the impacts of added impervious cover.

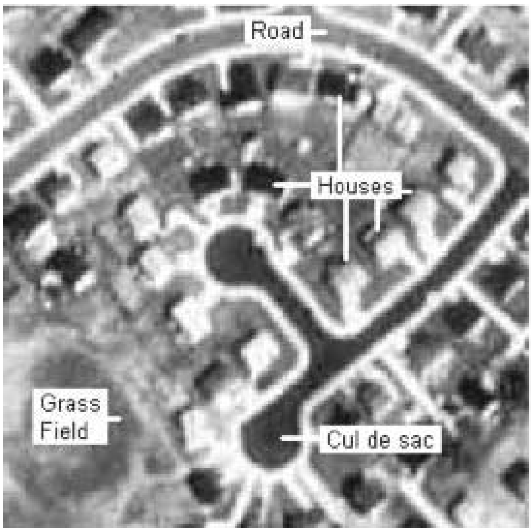

Examples of impervious cover

GIS AND SPATIAL DATA

A Geographic Information System (GIS) is a computer system capable of assembling, storing, manipulating, and displaying geographically referenced information (i.e., data identified according to set locations). The way maps and other data have been stored or filed as layers of information in a GIS makes it possible to perform complex analyses.

With the aid of GIS software, a human analyst can efficiently and accurately identify and categorize ground features visible in computer-generated versions of aerial photographs that represent true map distances available throughout the nation. High-resolution satellite imagery is rapidly expanding use of remote sensing techniques for impervious cover estimation. Classified land cover data are now available throughout the country. Detailed road networks and block level census data are other GIS data sources available to aid in estimating impervious cover.

RESEARCH OBJECTIVE

Impervious cover is proposed as an indicator of aquatic conditions for sub-watersheds throughout the country. Researchers are focusing on methods that would be useful in doing region-wide environmental assessments. The usefulness of impervious cover as an indicator is a function of the ease and accuracy for estimating it. Testing is underway to determine with what degree of accuracy impervious cover can be estimated for sub-watershed areas from data available throughout a region.

RESEARCH SUMMARY

An impervious cover test data set for 56 sub-watersheds in Frederick County, MD was developed and used to evaluate different estimation techniques suitable for application to a regional-scale characterization. Using a combination of data sources, researchers were able to estimate the percentage of impervious cover in a watershed to within +/− 1%.

Additional impervious cover data sets are being developed from aerial photographs for approximately 209 watersheds in the Atlanta, GA area for two time periods. These data sets will be used to confirm estimation methods and to determine the accuracy of methods to project future changes in impervious surface area. The estimation technique that was developed in this research project can be used by local and national agencies to target watersheds for monitoring and mitigation efforts.

PHYTOREMEDIATION BY CONSTRUCTED WETLANDS

Rupali Datta
Dibyendu Sarkar
Alpana Khairom
Chacharee Therapong
University of Texas at San Antonio
San Antonio, Texas

A constructed wetland is a wastewater treatment facility that duplicates the processes in natural wetlands. Constructed wetlands are complex, integrated systems in which water, plants, animals, microorganisms, and the environment interact to improve water quality. Phytoremediation, defined as cleansing of the environment using vegetation, is the primary water treatment process in constructed wetlands. The advantages of phytoremediation using constructed wetlands are that it is environmentally friendly, a sustainable technology, and cost-effective; the primary disadvantages are potentially incomplete removal of pollutants and pest attraction. Successful implementation of phytoremediation technology in constructed wetlands is a function of a number of interrelated biophysicochemical factors, including the types and fate of toxicants, characteristics of soil and sediment, plant species and plant uptake, hydrology and water flow, and microbial population. Hence, phytoremediation by constructed wetlands generally reflects an integrated multidisciplinary effort that combines plant biology, microbiology, surface chemistry, hydrology, and environmental engineering. The initial construction of wetlands, taking into account all these factors to design the most appropriate and effective method for removing target pollutants is labor-intensive and expensive, but low maintenance and operating costs make constructed wetlands a much more economically viable alternative to conventional physical and chemical treatments. Above all, the use of plants in environmental cleanup may guarantee a greener and cleaner planet for us and for future generations.

INTRODUCTION

The rising demands of modern day society for economic development and improved standards of living are leading to increased use of raw materials (1). The use of chemicals has played a key role in the development of agriculture, industry, housing, transport, textiles, and health. Their use has contributed significantly to the rise in the standard of living among populations around the world, but a major side effect has been the continuous release of both natural and man-made substances, including gases, heavy metals, volatile organic compounds, soluble organic compounds, suspended solids, and nitrogen and phosphorus compounds into the air, water, and terrestrial environments. Recent research is increasingly focused on the ecological and human health effects of contaminants, and guidelines or standards on soil, drinking water, irrigation water, crop tissues, foodstuff, biosolids, and fertilizers are now becoming available (2). In addition to the human impact, some contamination problems result from geologic activity. For example, reports on reduced iron in Denmark, fluoride in Bavaria and Moldova, and arsenic in Bangladesh and India illustrate the fact that natural geochemical conditions may cause locally elevated concentrations of metals (1).

Environmental remediation technologies use physical, chemical, or biological processes that attempt to eliminate, reduce, isolate, or stabilize a contaminant or a multitude of contaminants (3). Depending on the technology used, the process may take place either at the location of the contamination (*in situ*), or the contaminated soil or water may be removed for *ex situ* treatment. The cost of environmental remediation using conventional techniques can be very high, a fact that has turned out to be the driving force in the development of new remediation technologies. Phytoremediation, defined as the cleansing of the environment using vegetation, has emerged as an effective, inexpensive, and environmentally friendly alternative for remediating contaminated soil and ground water (4). Commercial use of phytoremediation is currently rather limited; most of the technologies being developed are primarily at the experimental stage. The most developed and widely used phytoremediation technique is the use of constructed wetlands for treating wastewater from municipal sewage treatment facilities and industrial processing operations (5).

Wetlands prevent floods, store water for times of drought, and most importantly, clean silt and pollutants from runoff by trapping sediment and contaminants. By doing so, they prevent chemical and sedimentary pollution of rivers, lakes, and coastal waters; hence, environmentalists refer to wetlands as nature's kidneys (6). The physicochemical properties of wetlands provide many positive attributes for remediating contaminants (7). The rhizosphere of wetland vegetation provides an enriched culture zone for microbes involved in degradation (8). Furthermore, sediments or soils in wetlands display reducing conditions, which promotes certain remediation reactions (9). Chemical pollutants are also taken up by wetland plants, where they are accumulated and biotransformed to less toxic and/or immobile states (10) or are volatilized into the atmosphere (11). Plants also supply fixed carbon to sediments, which is an energy source for bacterial transformation of chemical contaminants (12). Constructed wetland ecosystems are typically characterized by carefully selected and induced biophysicochemical properties favorable for chemical, bio- and phytoremediation of contaminants, such as enhanced surface retention, complimentary redox conditions, hyperaccumulating vegetation, and degrading microbial population, and hence, provide conditions that help transform harmful pollutants to less harmful products or essential nutrients that can be used by the biota (13).

Although natural wetlands and terrestrial phytoremediation are traditionally used for removing wastes, phytoremediation using constructed wetlands is becoming more popular due to their efficiency in treating pollutants and cost-effectiveness. This article provides an overview of the types of constructed wetlands, their advantages and disadvantages, the pollution removal mechanisms (with particular emphasis on phytoremediation attributes), and their effectiveness in removing organic pollutants, nutrients, pathogens, and heavy metals.

TYPES OF CONSTRUCTED WETLANDS

A constructed wetland is a designed, man-made complex of saturated substrates, emergent and submerged vegetation, animal life, and water that simulates natural wetlands for human use and benefits (14). Constructed wetlands offer an economical and largely self-maintaining alternative to conventional treatment of water contaminated by a variety of pollutants such as organic chemicals, toxic metals, and nutrients (14).

Wetland systems are typically described in terms of the position of the water surface (15). Accordingly, there are two basic types of constructed wetlands:

- subsurface flow systems
- free-water surface flow systems

Subsurface Flow Systems

A subsurface flow (SF) wetland is specifically designed for treating certain types of wastewater and is typically constructed as a bed or channel containing appropriate media (Fig. 1). Coarse rock, gravel, sand, and other soils have been used, but a gravel medium is most common in the United States and Europe (15). The medium is typically planted with the same types of vegetation present in marshes, and the water surface is designed to remain below the top surface of the medium. The sizes of these systems range from small on-site units designed to treat septic tank effluents from individual homes, schools, apartment complexes, and parks that treat a few hundred gallons of wastewater per day to 1.5×10^7 liters per day systems that treat municipal wastewater (15). There are approximately 100 SF systems in the United States that treat municipal wastewater, the majority of these treat less than 3.8×10^3 m^3/day (15). The most commonly used vegetation in SF wetlands includes cattail (*Typha* sp.), bulrush (*Scirpus* sp.), reed (*Phragmites* sp.), and sedge (*Carex* sp.). The submerged plant roots provide a substrate for microbial processes. Because most emergent macrophytes can transmit oxygen from leaves to roots, there are aerobic areas on the root surfaces. The remainder of the submerged environment in SF wetlands tends to remain devoid of oxygen (17). These systems are very useful in removing biochemical oxygen demand (BOD), total suspended solids (TSS), metals, and organic pollutants.

Free-Water Surface Flow Systems

Free-water systems (FWS) are designed to simulate natural wetlands, such as bogs, swamps, and marshes where water flows over the soil surface at a shallow depth (Fig.1). In FWS wetlands, water flows over a vegetated soil surface from an inlet point to an outlet point. The sizes of FWS wetlands systems range from small on-site units designed to treat septic tank effluents to large units that occupy more than 40,000 acres (15). Similar to SF systems, the most commonly used emergent vegetation includes cattail, bulrush, and reeds (17). The plant canopy formed by the emergent vegetation shades the water

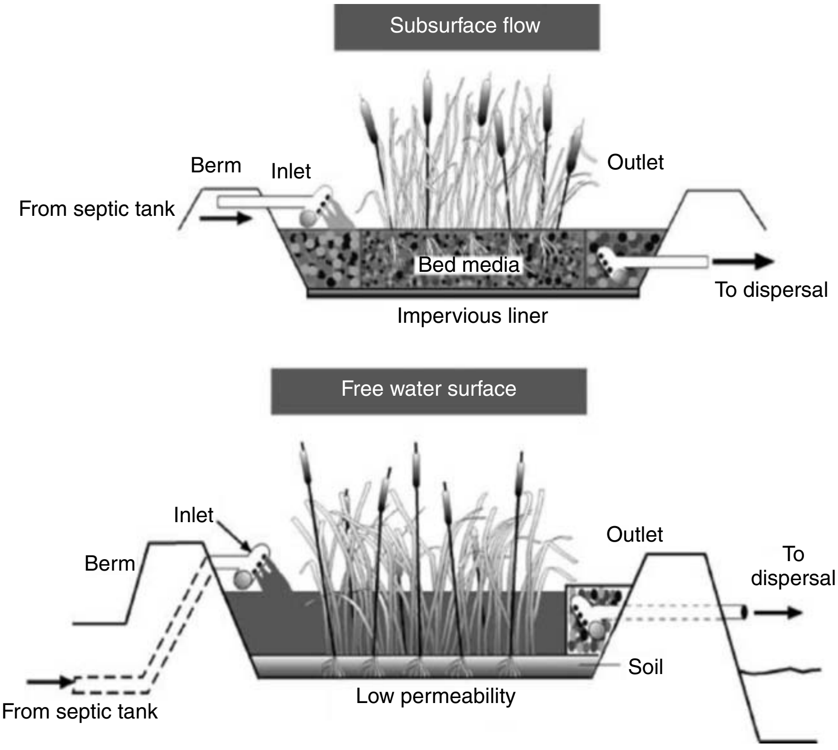

Figure 1. Subsurface flow and free-water systems (16).

surface, prevents the persistence and growth of algae, and reduces wind-induced turbulence in the water flowing through the system (15). Submerged portions of the living plants, the litter accumulated from the previous growth, and the standing dead plants provide a substrate for the microorganisms responsible for biological treatment in the system. Oxygen is available at the water surface and on living surfaces, allowing for aerobic activities. However, the bulk of the activities in the wetland is anaerobic (15). The low velocity and laminar flow of the influent in FWS systems result in very effective removal of TSS that contain BOD components, fixed forms of total nitrogen (TN) and total phosphorus (TP), trace levels of metals, and complex organics. The particulates are either oxidized or reduced and release soluble forms of BOD, TN, and TP, which are either adsorbed by soil/sediment or removed by microbial or plant populations in the wetland (15).

The use of FWS constructed wetlands has increased significantly since the late 1980s. SF wetlands are not now typically selected for larger flow municipal systems. The higher cost of the rock or gravel media makes a large SF wetland uneconomical compared to a FWS wetland despite the smaller area required for a SF wetland (15). Cost comparisons have shown that at flow rates above 60,000 gallons per day, it is usually cheaper to construct a FWS wetland system. However, there are exceptions where public access, mosquito, or wildlife issues justify selecting a SF wetland. Currently, FWS systems are widely distributed in 32 states in the United States (15).

ADVANTAGES AND DISADVANTAGES OF CONSTRUCTED WETLANDS

Constructed wetlands provide various advantages. This widens their appeal among professionals, from engineers and those involved in the workings of wastewater treatment facilities to the environmentalists and people concerned with recreation. Unlike some water issues in which the advantages to one group are disadvantages to another, the effective operation of constructed wetlands can provide broad benefits to a wide range of interests (18). Some of the more obvious benefits are listed here:

- Constructed wetlands are less expensive to build than other treatment options. There is no need for high cost, electromechanical equipment or for high consumption of electrical energy. If properly constructed and maintained, the wetlands provide efficient treatment of wastewaters (6).
- Operating and maintenance expenses (energy and supplies) are low. They require only periodic (rather than continuous) on-site labor (6).
- They are tolerant of fluctuating hydraulic and contaminant loading rates (19).
- They facilitate water recycling and reuse (15).
- Constructed wetlands have great tolerance for varying amounts of wastewater loading. The loading rates can be altered by varying livestock production levels, changing climatic conditions, and changing management (6).
- They can be aesthetically pleasing and may also provide additional benefits such as wildlife habitat, public recreation, environmental education, and groundwater recharge (6).

However, there are certain obvious disadvantages and negative side effects from the operation, maintenance, and effectiveness of constructed wetlands:

- Constructed wetlands require a continuous supply of water and also require water level management to sustain vegetation (6).
- They require more land area than the conventional treatment plants, if advanced treatment is desired (14).
- Surface wetlands attract mosquitoes and may cause other pest problems (14).
- Treatment efficiency may vary based on the effects of season and weather (6).
- High levels of ammonia can destroy vegetation in wetlands (6).
- The long-term effectiveness of constructed wetlands is not well documented. Wetland aging may contribute to a decrease in contaminant removal rates over time.

POLLUTANT REMOVAL MECHANISMS

Pollutant removal processes occur by interaction with wetland vegetation, the water column, and the wetland substrate. The processes may be physical, chemical, or biological (20).

Physical pollutant removal processes in wetland systems occur due to the presence of plant biomass and substrate media. Plants physically retard the pathways of sedimentation; the soil or gravel substrate acts as a filter bed, thereby aiding physical removal of suspended solids by straining (20).

Chemical reactions between inorganic pollutants can lead to their precipitation from the water column as insoluble compounds. Exposure to atmospheric gases and sunlight can lead to the breakdown of certain organic pollutants and destruction of pathogens. Antibiotic chemicals can also play a role in removing pathogens from wastewater (20).

Biological processes play a major role in removing pollutants in constructed wetland systems. Six major biological reactions have been identified as aiding the pollutant removal performance of wetlands, photosynthesis, respiration, fermentation, nitrification, denitrification, and biological phosphorus removal (7). A more detailed discussion of biological mechanisms, particularly in regard to bio- and phytoremediation, is presented in a later section.

As in any other biological wastewater treatment system, wetland process rates depend on various environmental factors, such as temperature, pH, oxygen availability, and hydraulic and pollutant loads (21,22). Under some environmental conditions, process rates may slow down or cease altogether, or even reverse, releasing pollutants. In general, the effectiveness of pollutant removal processes that rely on biological activities may be reduced due to a decrease in metabolic activities caused by low temperature; many metabolic activities are less effective if the pH is either too high or too low (20). Hydraulic overloading occurs when the flow exceeds the design capacity, thus reducing the actual hydraulic retention time. Pollutant overload occurs when the influent pollutant loads exceed the process removal rates; for example, nutrient overloads can cause excess algal bloom leading to eutrophication, and decomposition of excessive organic matter can consume dissolved oxygen resulting in anaerobic conditions (20–22). Nutrient and organic matter overloading may even reverse many pollutant removal processes (20).

PHYTOREMEDIATION ATTRIBUTES OF CONSTRUCTED WETLANDS

Phytoremediation using wetlands is a potentially economical remediation alternative that plays an important role in sustaining and restoring the environment. It has been an important aspect of constructed wetlands, which have been used successfully to detoxify large volumes of wastewater that contain dilute concentrations of contaminants. Pollutants in wastewater are generally removed by plants in a symbiotic association with beneficial soil bacteria (23). For many years, scientists and engineers have observed improvement in the quality of water flowing through natural wetlands. This knowledge is of invaluable importance in designing the most effective constructed wetland systems. However, assessing the phytoremediation potential of constructed wetlands is still a rather complex issue due to variable conditions of hydrology, soil/sediment types, plant species diversity, and seasonal effects on plant growth and water chemistry (24). Williams (7) provides an analytical assessment of the phytoremediation attributes of constructed wetlands, as discussed following.

The physicochemical properties of wetlands provide many positive attributes for remediating contaminants (7). The expansive rhizosphere of wetland herbaceous shrub and tree species provides an enriched culture zone for microbes involved in degrading pollutants (8,25,26). Most wetland soil/sediment zones have strongly reducing conditions (9). Remediation reactions that require reducing conditions are enhanced under the low redox potential of wetland systems (27).

However, anaerobic reducing conditions also decrease the depth to which plant roots penetrate, restricting plant uptake and rhizosphere actions to shallower aquifer levels (7). Zones of wetlands interfacing with flowing water systems can provide conditions for aerobic oxidation-dependent degradation reactions. Redox conditions within rhizospheres can also be altered by oxygen delivery by plant roots (28).

The oxidation states of wetlands are also strongly related to hydrology. The rise and fall of the water table can greatly alter redox zones (7). Soil porosity, percolation rates, and horizontal hydrologic conductivity rates control the rate of oxygen delivery to subsurface layers. Slower vertical and horizontal flow rates with high microbial activity quickly produce reducing conditions. Conversely, more rapid subsurface flows increase oxygen delivery, thereby slowing the decline in redox (7). Plant roots can transport oxygen to the surrounding rhizosphere (28–30) along with plant tissue air spaces connected to belowground biomass (31). Complex interactions among the microbial community, plant species composition, soil/sediment characteristics, and hydrology create a complex set of remediation reactions in wetland ecosystems (7).

Phytoremediation reactions are influenced by seasonal variability of physical and chemical characteristics in response to terrestrial and climatic events (7,32). Plant photosynthesis and standing crop biomass exhibit dramatic seasonal changes in nearly all tidal freshwater vegetation communities (9). Long-term ecological succession can shift plant species composition in the wetlands thereby altering phytoremediation characteristics because phytoremediation by one wetland community may not be the same as that which the new successional plant community develops (7). Hence, successful long-term phytoremediation must carefully consider the management implications of secondary successions to avoid ecological shifts away from optimal plant community structure (33).

Given all the factors that can influence the effectiveness of a wetland system and keeping in mind that specific requirements vary depending on the functional role of wetland plants in the treatment systems, broad generalizations can still be made about the characteristics of plant species that are most suitable for constructed wetland wastewater treatment systems (34):

- ecological acceptability, that is, no significant weed or disease risks or danger to the ecological or genetic integrity of surrounding natural ecosystems;
- tolerance to local climatic conditions, pests, and diseases;
- tolerance of pollutants and hypertrophic waterlogged conditions;
- ready propagation and rapid establishment, spread, and growth;
- high pollutant removal capacity, either through direct assimilation or storage, or indirectly by enhancing microbial transformations.

Cattails (*Typha* sp.), bulrushes (*Scirpus* sp.), reeds (*Phragmites* sp.), and sedges (*Carex* sp.) generally meet these requirements and hence, are the typical macrophytes (rooted plants that anchor to the substrate) of choice in the majority of constructed wetland systems, including both subsurface flow and freshwater surface flow types (20).

EFFECTIVENESS OF PHYTOREMEDIATION IN WETLANDS

Constructed wetlands can remove a range of pollutants, including biodegradable organic substances, metals, and nutrients. Natural wetlands have lower efficiency in removing pollutants than constructed wetlands (10). The growth and adaptation of plants to anoxic conditions in wetland sediments drives many of the biological and chemical processes to transform the pollutants to a variety of forms that differ in mobility and toxicity (12). For example, the activity of plant roots alters the surrounding conditions in the sediment, enhancing the rate of transformation and fixation of metals (35).

The following section, derived primarily from Williams (7) and Sundaravadivel and Vigneswaran (20), presents a review of the performance of both constructed and natural wetlands in regard to phytoremediation of the aforementioned three major classes of chemical pollutants.

Removal of Organics

Organic contaminants, such as trichloroethylene (TCE) and pesticides, have been remediated effectively by a variety of phytoremediation actions of terrestrial and wetlands species (7). Cunningham et al. (36) identified the following phytoremediation actions for successfully removing organic contaminants: accumulation into biomass, phytovolatilization, cellular degradation, and rhizosphere degradation. Hybrid poplars have had increasing success in phytoremediating TCE (37) and other volatile organic compounds (VOCs). Several VOCs, including TCE, were removed from surficial groundwater and subsequently degraded by poplar (*Populus deltoids* x *trichlocarpa*) test plots near toxic disposal sites in Maryland (38). Nietch et al. (39) experimented with bald cypress (*Taxodium distichum*) to remediate TCE and found that the TCE flux through bald cypress into surrounding air space decreased from day to night and from August to December. Much of this flux was controlled by evapotranspirational seasonality.

Nzengung et al. (40) summarized phytoremediation research on halogenated organic chemicals (HOC) by *Elodea candensis, Myriophyllum aquaticum, Salix nigra*, and *Populus deltoides* and identified four major phytoremediation processes: (1) rapid sequestration by partitioning to lipophilic plant cuticles; (2) phytoreduction to less halogenated metabolites; (3) phytooxidation to haloethanols, haloacetic acids, and unidentified metabolites; and (4) assimilation into plant tissues as nonphytotoxic products. Additional types of VOC, semivolatile, and nonvolatile compounds were evaluated for their phytoremediation potential by hybrid poplar (41). Burken and Schnoor (41), developed the transpiration stream concentration factor (TSCF) and the root concentration factor (RCF) for 12 such compounds, including aniline, phenol, nitrobenzene, benzene, atrazine, and toluene. Newman et al. (42) reported phytoremediation via reductive dechlorination by *Populous trichocarpa* and emphasized the ecological concern of selecting a phytoremediating species that grows well in the local habitat, while not ecologically destabilizing existing communities. In another study, Ensley (43) reported that duckweed (*Lemna gibba*) reductively dechlorinates chlorinated phenols.

Removal of Metals

Wetland sediments are generally considered a sink for metals and, in the anoxic zone, may contain very high concentrations of metals in a reduced state (44). As such, the bioavailability of metals in wetlands is low compared to that in terrestrial systems with oxidized soils. Water-soluble and exchangeable metals are most bioavailable, metals precipitated as inorganic compounds or metals complexed by humic substances or metals adsorbed onto hydrous oxides are potentially bioavailable, and metals precipitated as insoluble sulfides and metals bound to the crystal lattice of minerals are essentially unavailable (45). Because of reducing conditions, the depth of wetlands to which plants can penetrate is limited, and this restricts the uptake of contaminants and rhizosphere actions to shallower levels (7).

Understanding the role of phytoremediation in metal removal from wetlands has progressed over the years from basic uptake studies to quantification of chemical speciation and bioavailability and the role of genetic engineering in metal hyperaccumulation (46–50). In New Zealand, submerged, rooted aquatic species, such as *Ceratophyllum demersum*, were hyperaccumulators of arsenic (CF or concentration factor of 20,000) compared to cattail, *Typha orientalis* (CF of 100), a popular wetland plant species (46). Water hyacinths (*Eichhornia crassipes*), well known for their use in wastewater treatment, have also been found effective in accumulating heavy metals and in the uptake of pesticide residues, PCB, DDE, DDD, and DDT (47,51). Apparently, water hyacinths carried out phytoremediation by root sorption, concentration, and/or metabolic degradation (47). Duckweed (*Lemna minor*) and water velvet (*Azolla pinnata*) both effectively removed iron and copper at low concentrations; however, water velvet does not grow from May to October, and such production seasonality translates to a limited phytoremediation interval (52).

Removal of heavy metals by wetland hydrophytes presents the problem of plant biomass collection to prevent decomposition, thus recycling accumulated metals, because many of the wetland plant species typically undergo much faster decomposition compared to terrestrial species (7). Phytomining, the recovery of accumulated trace metals, might be side benefits of phytoremediation; however, concerns about plant matter getting into the food chain by direct consumption or by decomposition pathways (49) include considerable ecological and human health problems. Careful harvesting techniques must be used to prevent loss of metal-enriched plant biomass. Biomass recycling is a major concern because even wave action can remove fresh biomass prior to harvesting (7).

Removal of Nutrients

Removal of macronutrients, nitrogen and phosphorus, by stabilized constructed wetlands is a complex cyclic process believed to involve a number of "conceptual compartments," such as soils, sediments, and litter/peat (80% of the total nutrient load), water column (15–20%), and plants/other biota (5%) (20,53). The primary mechanisms of nutrient removal from wastewater in constructed wetlands are microbial processes such as nitrification and denitrification as well as physicochemical processes such as the fixation of phosphate by iron and aluminum in the soil filter (17). The hydraulic retention time, including the length of time the water is in contact with the plant roots, affects the extent to which the plant plays a significant role in the removal or breakdown of pollutants (17).

A survey of literature on the nutrient removal efficiencies of constructed wetlands indicates extreme variations (20). Phosphorus removal efficiency is strongly dependent on loading rate, 65–95% removal is achieved at loading rates less than 5 g/m^2/y. However, removal efficiency decreases to 30–40% or even less when phosphorus loadings are greater than 10–15 g/m^2/y (54). Nitrogen removal efficiencies reportedly decline at mass loading rates above 20 kg/ha/d (20). Total nitrogen removals up to 79% are reported at higher loading rates of up to 44 kg/ha/d from constructed wetland systems in the United States (55).

Nutrient removal by wetland systems is also a function of climate; tropical climatic conditions generally tend to increase nutrient removal efficiencies (20). The nitrogen and phosphorus removal efficiencies of a full-scale plant treating domestic wastewater in Bhubaneshwar, India, are over 70% and 43%, respectively (56). A pilot scale study in Central India also indicated higher removal efficiencies for both nitrogen and phosphorus in the range of 58 to 65% (57). A full-scale, gravel bed, horizontal flow system build in Kathmandu, Nepal, reports nitrogen removal in the range of 80–99% and phosphorus removal from as little as 5% to 69% (58).

SUMMARY AND FUTURE DIRECTIONS

Constructed wetlands can be designed to remove a wide range of pollutants from wastewater. Due to their low operating, maintenance, and energy costs, constructed wetlands could rapidly become the ideal method to achieve sustainable wastewater management goals. Several laboratory-scale mesocosmic studies have established the versatile nature of wetlands to remove pollutants, but only a few large-scale field studies have been conducted to gauge the remediation potential of constructed wetlands. The use of constructed wetlands to treat wastewater is a relatively new and emerging technology. The initial results are encouraging, but the full range of capabilities and limitations of this technology have not yet been uncovered.

Degradation of pollutants by wetlands involves complex interactions among wetland species, sediment/water geochemistry, and climate. Therefore, it is crucial to determine the effectiveness of remediation by implementing a detailed monitoring program that accounts for the interactions of numerous variables ranging from the composition of the plant species to the seasonality of hydrologic flow. Carefully planned monitoring programs are essential to determine the effectiveness of the wetland remediation strategy and are also instrumental in achieving regulatory acceptance.

Public concern during the past 20 years has strengthened regulations on wastewater discharges and resulted in substantial progress in treating point sources of water pollution, especially for large cities and major industries. Widespread implementation of the constructed wetlands treatment technology may accomplish the same for small communities, small industries, farmers, and developers, where low cost technologies are important. The available knowledge on constructed wetlands has grown extensively over the last decade, but much more research is needed to determine the capabilities of plants and microorganisms to remove pollutants from wastewater. As more wetland projects are undertaken and further field-based research is conducted, the treatment possibilities of constructed wetlands will be better understood.

BIBLIOGRAPHY

1. Holt, M.S. (2000). Sources of chemical contaminants and routes into the freshwater environment. Review article. *Food Chem. Toxicol.* **38**: S21–S27
2. Adriano, D.C. (1986). *Trace Elements in the Terrestrial Environment*, 2nd Edn. Springer-Verlag, New York.
3. USDA. (1996). *More moderate economic growth expected in the rest of 1996 and 1997*. Economic Research Service. Industrial Uses/IUS-6/September 1996.
4. Ensley, B.D. (2000). Rationale for use of phytoremediation. In: *Phytoremediation of Toxic Metals Using Plants to Clean Up the Environment*. I. Raskin and B.D. Ensley (Eds.). John Wiley & Sons, New York, pp. 3–11.
5. Kadlec, R.H. (1999). Constructed wetlands for treating landfill leachate. In: *Constructed Wetlands for the Treatment of Landfill Leachates*. G. Melamoottil, E.A. McBean, and F. Rovers (Eds.). Lewis Publishers, Boca Raton, FL, pp. 7–31.
6. Simeral, K.D. (2001). *Using constructed wetlands for removing contaminants from livestock wastewater*. Ohio State University Fact Sheet, School of Natural Resources, Columbus, OH. TDD No. 800-589-8292 (Ohio only) or 614-292-1868. http://ohioline.osu.edu/a-fact/0005.html.
7. Williams, J.B. (2002). Phytoremediation in wetland ecosystems: Progress, problems, and potential. *Crit. Rev. Plant Sci.* **21**(6): 607–635.
8. Meharg, A.A. and Cairney, J.W.G. (2000). Ectomycorrhizas—extending the capabilities of rhizosphere remediation? Review. *Soil Biol. Biochem.* **32**: 1475–1484.
9. Odum, W.E., Smith, T.J., Hoover, J.K., and Mclvor, C.C. (1984). *The ecology of tidal freshwater marshes of the United States East Coast: A community profile*. U.S. Dept. of the Interior, Fish and Wildlife Service, FWS/OBS-83/17.
10. Terry, N. (2000). *The San Francisco Bay constructed wetland treatment system: The use of constructed wetland treatment systems for the removal of toxic trace elements from electric utility wastewater: Role of vegetation*. EPRI, Palo Alto, CA, pp. 1–48.
11. Hansen, D., Duda, P., Zayed, A.M., and Terry, N. (1998). Selenium removal by constructed wetlands: Role of biological volatilization. *Environ. Sci. Technol.* **32**: 591–597.
12. Horne, A.J. (2000). Phytoremediation by constructed wetlands. In: *Phytoremediation of Contaminated Soil and Water*. N. Terry and G. Bañuelos (Eds.). Lewis, Boca Raton, FL, pp. 13–39.
13. Kadlec, R.H. and Knight, R.L. *Treatment Wetlands*. Lewis, Publishers, Boca Raton, FL.
14. Hammer, D.A. and Bastian, R.K. (1989). Wetlands ecosystems: Natural water purifiers? In: *Constructed Wetlands for Wastewater Treatment Municipal, Industrial and Agricultural*. D.A. Hammer (Ed.). Lewis, Chelsea, MI, pp. 5–19.
15. USDA-NRC and USEPA. (2000). *A Handbook of Constructed Wetlands: General Considerations*. Vol. 1. EPA- 843B00005. http://www.epa.gov/owow/wetlands/pdf/hand.pdf

16. National Environment Services and Center. (2001). Common onsite components. Constructed wetlands. http://www.nesc.wvu.edu/nsfc/frd10/constructedwetlands.htm
17. Stottmeister, U., Wiessner, A., Kuschk, P., Kappelmeyer, U., Kastner, M., Bederski, O., Muller, R.A., and Moormann, H. (2003). Effects of plants and microorganisms in constructed wetlands for wastewater treatment. *Biotechnol. Adv.* **22**(1–2): 93–117.
18. Gelt, J. (1997). Constructed wetlands: Using human ingenuity, natural processes to treat water, build habitat. *Arroyo* **9**(4). http://ag.arizona.edu/AZWATER/arroyo/094wet.html
19. Costa, C. (2003). *Applications of constructed wetlands treatment technology: An overview of the technology. Air Waste Manage. Assoc., South. Sect. Annu. Conf.*, Savannah, GA, July 30–Aug. 1, 2003.
20. Sundaravadivel, M. and Vigneswaran, S. (2001). Constructed wetlands for wastewater treatment. *Crit. Rev. Environ. Sci. Technol.* **31**(4): 351–409.
21. Young, R., White, R., Brown, M., Burton, J., and Atkins, B. (1998). *The Constructed Wetland Manual*. Vol. 1. NSW Department of Infrastructure, Planning and Natural Resources (Natural Resource Management), New South Wales, Australia.
22. Young, R., White, R., Brown, M., Burton, J., and Atkins, B. (1998). *The Constructed Wetland Manual*. Vol. 2. NSW Department of Infrastructure, Planning and Natural Resources (Natural Resource Management), New South Wales, Australia.
23. Krotz, R.M., McCaskey, T.A., Zhou, S.D., Lino, S., Dawkins, R.A., and Ruf, M.E. (1996). *Constructed wetlands transform raw sewage into pond water*. Highlights of Agricultural Research 43(3), Fall 1996, Alabama Agricultural Experiment Station, Auburn, AL. http://www.ag.auburn.edu/aaes/communications/highlights/fall96/wetlands.htm
24. Angier, J.T., McCarty, G.W., Rice, C.P., and Bialek, K. (2002). Influence of a riparian wetland on nitrate and herbicides exported from an agriculture field. *J. Agric. Food Chem.* **50**(15): 4424–4429.
25. Olsen, L.D. and Lorah, M.M. (1998). Natural attenuation of chlorinated NOCs in wetlands. *Ground Water Currents* (EPA) **29**: 1–3.
26. Macek, T., Mackova, M., and Kas, J. (2000). Exploitation of plants for the removal of organics in environmental remediation. Review article. *Biotechnol. Adv.* **18**: 23–34.
27. Chapelle, F.H. (1996). *Identifying redox conditions that favor the natural attenuation of chlorinated ethenes in contaminated ground-water systems. Symp. Nat. Attenuation Chlorinated Organics Ground Water*. EPA/540/R-96/509. U.S. Environmental Protection Agency, Washington, DC. pp. 17–20.
28. Weissner, A., Kuschk, P., Kastner, M., and Stottmeister, U. (2002). Abilities of helophyte species to release oxygen into rhizospheres with varying redox conditions in laboratory-scale hydroponic systems. *Int. J. Phytoremediation* **4**(1): 1–15.
29. Brix, H. (1993). Wastewater treatment in constructed wetlands: System design, removal processes, and treatment performance. In: *Constructed Wetlands for Water Quality Improvement*. G.A. Moshiri (Ed.). CRC Press, Boca Raton, FL, pp. 9–22.
30. Wetzel, R.G. (1993). Constructed wetlands: scientific foundations are critical. In: *Constructed Wetlands for Water Quality Improvement*. G.A. Moshiri (Ed.). Lewis, Boca Raton, FL, pp. 3–8.
31. Valiela, I., Teal, J.M., and Persson, N.Y. (1976). Production and dynamics of experimentally enriched salt marsh vegetation: below-ground biomass. *Limnol. Oceanogr.* **2l**: 245–252.
32. Kalff, J. (2002). *Limnology: Inland Water Ecosystems*. Prentice-Hall, Upper Saddle River, NJ.
33. Williams, J.B., Coleman, H.V., and Pearson, J. (1999). *Implications of pH effects and succession for phytoremediation in wetlands*. In: G.A. Uzochukwu and G.B. Reddy (Eds.). *Proc. 1998 Natl. Conf. Environ. Remediation Sci. Technol.*, Battelle Press, Columbus, OH, pp. 243–248.
34. Tanner, C.C. (1996). Plants for constructed wetland treatment systems—a comparison of the growth and nutrient uptake of eight emergent species. *Ecological Eng.* **7**: 59–83.
35. Wang, T. and Peverly, J.H. (1999). Iron oxidation states on root surfaces of a wetland plant (*Phragmites australis*). *Soil Sci. Soc. Am. J.* **63**: 247–252.
36. Cunningham, S.D., Shann, J.R., Crowley, D.E., and Anderson, T.A. (1997). Phytoremediation of contaminated water and soil. In: *Phytoremediation of Soil and Water Contaminants*. E.L. Kruger, T.A. Anderson, and J.R. Coats (Eds.). ACS Symposium Series 664. American Chemical Society, Washington, DC, pp. 2–19.
37. Gordon, M., Choe, N., Duffy, J., Ekuan, G., Heiman, P., Nuiznieks, I., Newman, L., Ruszaj, M., Shurtleff, B.B., Strand, S., and Wilmoth, J. (1997). Phytoremediation of trichloroethylene with hybrid poplars. In: *Phytoremediation of Soil and Water Contaminants*. E.L. Kruger, T.A. Anderson, and J.R. Coats (Eds.). ACS Symposium Series No. 664, American Chemical Society, Washington, DC, pp. 177–185.
38. Compton, H.R., Haroski, D.M., Hirsh, S.R., and Wrobel, J.G. (1998). Pilot scale use of trees to address VOC contamination. In: *Bioremediation and Phytoremediation, Chlorinated and Recalcitrant Compounds*. G.B. Wickramanayake and R.E. Hinchee (Eds.). Battelle Press, Columbus, OH, pp. 245–250.
39. Nietch, C.T., Morris, J.T., and Vroblesky, D.A. (1999). Biophysical mechanisms of trichloroethene uptake and loss in bald cypress growing in shallow contaminated groundwater. *Environ. Sci. Technol.* **33**(17): 2899–2904.
40. Nzengung, V.A., Castillo, R.M., Gates, W.P., and Mills, G.L. (2001). Abiotic transformation of perchloroethylene in homogeneous dithionite solution and in suspensions of dithionite-treated clay mineral. *Environ. Sci. Technol.* **35**: 2244–2251.
41. Burken, J.G. and Schnoor, J.L. (1998). Predictive relationships for uptake of organic contaminants by hybrid poplar trees. *Environ. Sci. Technol.* **32**(21): 3379–3385.
42. Newman, L.A., Doty, S.L., Gery, K.L., Heilman, P.E., Muizneiks, I.E., Shang, T.Q., Siemieniec, S.T., Strand, S.E., Wang, X.E., Wilson, A.M., and Gordon, M.P. (1998). Phytoremediation of organic contaminants: A review of phytoremediation research at the University of Washington. *J. Soil Contam.* **7**(4): 531–542.
43. Ensley, B.D. (1997). *The Use of Plants for the Remediation of Environmental Contamination*. In: P.T. Kostecki and E.J. Calabrese (eds.), *12th Annu. Conf. Contam. Soils—Anal., Site Assessment, Fate, Environ. Hum. Risk Assessment, Remediation Regulat.*, Amherst, MA, October 20–23, 1997. Environmental Health Sciences Program, School of Public Health, University of Massachusetts, Amherst, MA.
44. Weis, J.S. and Weis, P. (2004). Metal uptake, transport and release by wetland plants: Implications for phytoremediation and restoration. *Environ. Int.* **30**: 685–700.
45. Gambrell, R.P. (1994). Trace and toxic metals in wetlands—a review. *J. Environ. Qual.* **23**: 883–891.
46. Reay, P. (1972). The accumulation of arsenic-rich natural water by aquatic plants. *J. Appl. Ecol.* **9**: 557–565.

47. Gupta, G.C. (1980). Use of water hyacinths in wastewater treatment. *J. Environ. Health* **43**(2): 8082.
48. Debusk, T.A., Laughlin, R.B., Jr., and Schwartz, L.N. (1996). Retention and compartmentalization of lead and cadmium in wetland microcosms. *Water Resour.* **30**(11): 2707–2716.
49. Khan, A.G., Kuek, C., Chaudhry, T.M., Khoo, C.S., and Hayes, W.J. (2000). Role of plants, mycorrhizae and phytochelators in heavy metal contaminated land remediation. *Chemosphere* **41**: 197–207.
50. Datta, R. and Sarkar, D. (2004). Effective integrations of soil chemistry and plant molecule biology in phytoremediation of metals: An overview. *Environ. Geoscience* **11**(2): 53–63.
51. Wang, Q., Cui, Y., and Dong, Y. (2002). Phytoremediation of polluted waters: Potentials and prospects of wetland plants. *Acta Biotechnol.* **22** (1–2): 199–208.
52. Jain S.K., Vasudevan, P., and Jha, N.K. (1989). Removal of some heavy metals from polluted water by aquatic plants: Studies on duckweed and water velvet. *Biological Wastes* **28**: 115–126.
53. Faulkner, S.P. and Richardson, C.J. (1990). Physical and chemical characteristics of freshwater wetland soils. In: *Constructed Wetlands for Wastewater Treatment*, 2nd Edn. D.A. Hammer (Ed.). Lewis, Chelsea, MI, pp. 41–72.
54. Faulkner, S.P. and Richardson, C.J. (1993). Physical and chemical characteristics of freshwater wetland soils. In: *Constructed Wetlands for Water Quality Treatment*, CRC Press, Boca Raton, FL, p. 41.
55. Watson, J.T., Reed, S.C., Kadlec, R.H., Knight, R.L., and Whitehouse, A.E. (1990). Performance expectations and loading rates for constructed wetlands. In: *Constructed Wetlands for Wastewater Treatment*, 2nd Edn. D.A. Hammer (Ed.). Lewis, Chelsea, MI, pp. 319–351.
56. Juwarkar A.S., Oke, B., Juwarkar, A., and Patnaik, S.M. (1995). Domestic wastewater treatment through constructed wetland in India. *Water Sci. Technol.* **32**(3): 291–294.
57. Billore, S.K., Singh, N., Sharma, J.K., Dass, P., and Nelson, R.M. (1999). Horizontal subsurface flow gravel bed constructed wetland with phragmites karka in central India. *Water Sci. Technol.* **40**(3): 163–171.
58. Shreshtha, R.R., Haberl, R., and Laber, J. (2000). *Approach of constructed wetland technology transfer to Nepal*. *Proc. 1st World Water Congr. Int. Water Assoc.*, Paris, July 2000, IWA, London, UK.

UNRECOGNIZED POLLUTANTS

G. Fred Lee
Anne Jones-Lee
G. Fred Lee & Associates
El Macero, California

INTRODUCTION

The U.S. water pollution control program focuses on the control of a limited group of chemicals, the 126 current "priority pollutants." The chemicals included on the list of priority pollutants were originally selected in the mid-1970s through a litigation settlement. The list did not receive appropriate peer review for its representation of the chemical contaminants of most pressing importance to public health and environmental quality. That notwithstanding, the priority pollutants remain the focal point of water quality investigation and management. Limited attention is given to the evaluation and regulation of the many thousands of other potential pollutants that are in municipal and industrial liquid and solid wastes, even in light of the myriad of other chemicals currently in commercial and personal household use, and those that come into use annually.

Daughton (1) highlighted the growing concern about unrecognized, unregulated pollutants, indicating that over 22 million organic and inorganic substances exist, with nearly 6 million commercially available. The current water quality regulatory approach addresses fewer than 200 of these chemicals. He also noted,

> *"Since the 1970s, the impact of chemical pollution has focused almost exclusively on conventional "priority pollutants," especially on those collectively referred to as "persistent, bioaccumulative, toxic" (PBT) pollutants, "persistent organic pollutants" (POPs), or "bioaccumulative chemicals of concern" (BCCs).*
>
> *The "dirty dozen" is a ubiquitous, notorious subset of these, comprising highly halogenated organics (e.g., DDT, PCBs).*
>
> *The conventional priority pollutants, however, are only one piece of the larger risk puzzle.*
>
> *Regulated pollutants compose but a very small piece of the universe of chemical stressors to which organisms can be exposed on a continual basis.*

Similarly, Dr. K. Hooper of the Hazardous Materials Laboratory, California EPA Department of Toxic Substances Control, recently noted (2),

> *"Over the past 25 years, tens of thousands of new chemicals (7 chemicals per day) are introduced into commerce after evaluation by USEPA. Few (100–200) of the 85,000 chemicals presently in commerce are regulated. We have reasons to believe that a much larger number than 200 adversely affect human health and the environment."*

Periodically, unregulated chemicals that are in use are discovered to be widespread pollutants, posing a significant threat to public health and the environment. Examples of such findings and illustration of the need to address the broader issue of previously unrecognized pollutants are presented below.

EXAMPLES OF PREVIOUSLY UNRECOGNIZED POLLUTANTS

Pharmaceuticals and Personal Care Products (PPCPs)

Increasing attention is being given to pharmaceuticals and personal care products (PPCPs) as environmental pollutants. At the California Bay Delta Authority (CBDA) Contaminant Stressors Workshop, Dr. Christian Daughton, Chief, Environmental Chemistry Branch, US EPA National Exposure Research Laboratory, addressed this issue in his presentation entitled, "*Ubiquitous Pollution from Health and Cosmetic Care: Significance, Concern, Solutions, Stewardship—Pollution from Personal Actions*" (1).

Daughton (1) pointed out that a wide variety of chemicals that are introduced into domestic wastewaters are being found in the environment. Various chemicals (pharmaceuticals) used by individuals and to treat pets, outdated medications disposed of into sewage systems, and treated and untreated hospital wastes discharged to domestic sewage systems end up in receiving waters for domestic wastewater treatment plant effluents. Further, transfer of sewage solids ("biosolids") to land, industrial waste streams, landfill leachate, releases from aquaculture of medicated feeds, etc. also introduce these chemicals into the environment. Many of these chemicals are not new and have been present in wastewaters for some time. However, they are only now beginning to be recognized as potentially significant water pollutants and are largely unregulated as water pollutants.

According to Daughton (1),

> *"PPCPs* [Pharmaceuticals and Personal Care Products] *are a diverse group of chemicals comprising all human and veterinary drugs (available by prescription or over-the-counter; including the new genre of "biologics"), diagnostic agents (e.g., X-ray contrast media), "nutraceuticals" (bioactive food supplements such as huperzine A), and other consumer chemicals, such as fragrances (e.g., musks) and sun-screen agents (e.g., mehylbenzylidene camphor); also included are "excipients" (so-called "inert" ingredients used in PPCP manufacturing and formulation)."*

Although the full range of impacts of PPCPs is just beginning to be investigated, PPCPs are being found to have adverse impacts on aquatic ecosystems. For example, Daughton (1) discussed the relationship between PPCPs and endocrine disrupters, which are believed to be responsible for causing sex changes in fish. In addition, in a feature article in *Environmental Science and Technology*, Eggen et al. (3) reviewed a number of the issues pertinent to understanding the impacts of PPCPs and other chemicals that can cause endocrine disruption, DNA damage/mutagenesis, deficiencies in immune system, and neurological effects in fish and other aquatic life. (Additional information on PPCPs is available at www.epa.gov/nerlesd1/chemistry/pharma/ index.htm.)

Perchlorate

Perchlorate (ClO_4^-), used in highway safety flares and other applications, is another example of a chemical that is now being found in surface and groundwaters in sufficient concentrations to pose a threat to human health. Perchlorate is derived from several sources. Silva (4) of the Santa Clara Valley, CA Water District, noted the potential for highway safety flares to be a significant source of perchlorate contamination to water, even when the flares are 100% burned. According to him, one fully burned flare can leach up to almost 2,000 μg of perchlorate, and

> *"A single unburned 20-minute flare can potentially contaminate up to 2.2 acre-feet* [726,000 gallons] *of drinking water to just above the California Department of Health Services' current Action Level of 4 μg/L* [for perchlorate]" (4).

Silva also pointed out that more than 40 metric tons of flares were used/burned in 2002 alone in Santa Clara County. California's Office of Environmental Health Hazard Assessment recently conducted an evaluation of the hazards of perchlorate in drinking water. The 4 μg/L action level for perchlorate in drinking water was based on the detection limit; it has been revised to 6 μg/L based on the recent OEHHA evaluation.

Polybrominated Diphenyl Ethers (PBDE)

Another unrecognized, unregulated pollutant is the polybrominated diphenyl ethers (PBDEs) used as a flame retardant on furniture and other materials. Hooper (2) recently discussed finding PBDE in human breast milk and in San Francisco Bay seals, and the fact that archived human breast milk shows that this contamination has been occurring for over 20 years. According to McDonald (5) of the California Environmental Protection Agency, Office of Environmental Health Hazard Assessment,

> *"Approximately 75 million pounds of PBDEs are used each year in the U.S. as flame retardant additives for plastics in computers, televisions, appliances, building materials and vehicle parts; and foams for furniture. PBDEs migrate out of these products and into the environment, where they bioaccumulate. PBDEs are now ubiquitous in the environment and have been measured in indoor and outdoor air, house dust, food, streams and lakes, terrestrial and aquatic biota, and human tissues. Concentrations of PBDE measured in fish, marine mammals and people from the San Francisco Bay region are among the highest in the world, and these levels appear to be increasing with each passing year."*

PBDEs are similar to PCBs and are considered carcinogens. Some of the PBDEs are being banned in the United States and in other countries.

Pesticides

Another example of unidentified pollutants was given by Kuivila (6). She discussed the fact that approximately 150 pesticides used in California's Central Valley exist that are a threat to cause water quality problems in the Delta and its tributaries. The current pesticide water quality regulation program considers only about half a dozen of those.

CONCLUSIONS AND RECOMMENDATIONS

The presence of untold, unregulated pollutants in environmental systems, as illustrated above with the examples of PPCPs and others, is not unexpected based on the approach that is normally used to define constituents of concern in water pollution control programs. Based on the vast array of chemicals that are used in commerce, many of which are or could be introduced into aquatic systems from wastewater and stormwater runoff, it is likely that many other chemicals will be discovered in the future that are a threat to public health or aquatic ecosystems. A pressing need exists to significantly expand water quality monitoring programs to specifically search for new, previously unrecognized water pollutants. As demonstrated by the perchlorate and PBDE situations, monitoring programs that focus on priority pollutants stand to be significantly deficient in properly defining constituents of concern with respect to impairing the beneficial uses of waters.

BIBLIOGRAPHY

1. Daughton, C.G. (2004). Pharmaceuticals and personal care products (PPCPs) as environmental pollutants: Pollution from personal actions. *Presented at the California Bay-Delta Authority Contaminant Stressor Workshop*, Sacramento, CA. Contact Daughton at daughton.christian@epa.gov. Also see http://search.epa.gov/s97is.vts and http://www.epa.gov/esd/chemistry/pharma/tracers.htm.
2. Hooper, K. (2003). *Lessons from the PBDEs—the value of monitoring community body burdens using breast milk*. In: 6th Bienniel State of the Estuary Conference, *The San Francisco Bay-Delta Estuary: Changes and Challenges*, San Francisco Estuary Project, Oakland, CA.
3. Eggen, R.I.L., Behra, R., Burkhardt-Holm, P., and Escher, B.I. (2004). Challenges in ecotoxicology: Mechanistic understanding will help overcome the newest challenges. *Environ. Sci. Technol.* **38**(3): 58A–64A.
4. Silva, M.A. (2003). Safety flares threaten water quality with perchlorate. Report of Santa Clara Valley Water District. http://www.valleywater.org/Water/Water_Quality/Protecting_your_water/_Lustop/Perchlorate.shtm.
5. McDonald, T.A. (2003). *The risk posed by the PBDEs, a class of flame retardants*. In: 6th Bienniel State of the Estuary Conference, *The San Francisco Bay-Delta Estuary: Changes and Challenges*, San Francisco Estuary Project, Oakland, CA.
6. Kuivila, K. (2000). *Pesticides in the Sacramento-San Joaquin Delta: State of our Knowledge*. Abstracts from the CALFED Bay-Delta Program Science Conference 2000, Sacramento, CA.

READING LIST

Daughton, C.G. (2004) "Non-Regulated Water Contaminants: Emerging Research," *Environ. Impact Assess. Rev.* **24** (7–9): 711–732.

Ternes, T.A., Joss, A., and Siegrist, H. (2004). "Scrutinizing Pharmaceuticals and Personal Care Products in Wastewater Treatment," *Environ. Sci. Technol.* **38**(20): 393A–399A.

POLLUTION OF SURFACE WATERS

Nitish Priyadarshi
Ranchi University
Jharkhand, Ranchi, India

INTRODUCTION

According to the American College Dictionary, pollution is defined as "to make foul or unclean; dirty." Water pollution occurs when a body of water is adversely affected by the addition of large amounts of materials to the water. When it is unfit for its intended use, water is considered polluted.

Comprising over 70% of the earth's surface, water is undoubtedly the most precious natural resource that exists on our planet. Without the seemingly invaluable compound of hydrogen and oxygen, life on the earth would be nonexistent. It is essential for everything on our planet to grow and prosper. Although we as humans recognize this fact, we disregard it by polluting our rivers, lakes, and oceans. We are slowly but surely harming our planet to the point where organisms are dying at a very alarming rate. In addition to innocent organisms dying off, our drinking water has become greatly affected, as is our ability to use water for recreation. To combat water pollution, we must understand the problem and become part of the solution.

Surface water pollutants are generally classified into these eight major categories:

1. sewage and waste
2. infectious agents
3. particulate soil and mineral matter
4. dissolved toxic pollutants.
5. mineral and chemical compounds
6. radioactive nuclides
7. thermal pollutants
8. organic chemical exotics

Surface water should be free from the following contaminants (1):

1. Compounds that impart color, turbidity, and odor such as greases, oils, toxic metals, phenols, and organics
2. Toxic radionuclides that acutely affect the physiology of humans, animals, and aquatic organisms
3. Substances that may precipitate to form objectionable deposits or float on the surface as oil, scum, or debris
4. Heavy materials that check the growth of aquatic flora and fauna
5. Thermal effluents should not increase the water temperature by 3 to 5 °F.
6. Substances that are likely to enhance the growth of undesirable aquatic life

Components of surface water are listed in Table 1.

Table 1. Typical Analysis of Some Surface Waters[a]

Component (ppm)	X[b]	Y[c]
Silica	9.5	1.2
Calcium	4.0	36.0
Magnesium	1.1	8.1
Iron (III)	0.07	0.02
Total hardness as $CaCO_3$	14.6	123
Sodium	2.6	6.5
Potassium	0.6	1.2
Sulphate	1.6	22
Chloride	2.0	13
Nitrate	0.41	0.1
Bicarbonate	18.3	119
Total dissolved solids	34	165

[a] Reference 2.
[b] X: Pardee Reservoir, Oakland, California, U.S.A., 1976 (water supply source for East Bay Area of San Francisco).
[c] Y: Niagara River, New York, U.S.A. (water supply source for city of Niagara Falls).

SOURCES OF SURFACE WATER POLLUTION

Surface water comes in direct contact with the atmosphere, streams, rivulets, and surface drains. So a continuous exchange of dissolved and atmospheric gases occurs and wastes are added through water conveyance.

There are many causes of surface water pollution but two general categories exist: direct and indirect contaminant sources. Direct sources include effluent outfalls from factories, refineries, waste treatment plants, decomposed plant animal matter, and radioactive materials and wastes. Indirect sources include effluent contaminants that enter the surface water supply from surface soils and from the atmosphere via rainwater. Soils contain the residue of human agricultural practices (fertilizers, pesticides, etc.) and improperly disposed of industrial wastes. Atmospheric contaminants are also derived from human practices (such as gaseous emissions from automobiles, factories, and even bakeries).

Impact of Livestock Production Practices on Surface Water Quality

Livestock practices that can impact surface water quality include both intensive and nonintensive operations. Intensive operations include feedlots (>500 head of cattle), dairies, and wintering sites; nonintensive operations include pasture, cow-calf operations, and watering sites for cattle. Waste management and disposal can also impact water quality. Livestock density is not the only factor affecting water quality; siting and management are also important considerations. Water quality parameters related to livestock production include nutrients (nitrogen and phosphorus), microorganisms (e.g., bacteria, faecal coliforms, *Cryptosporidium, Giardia*), and organic materials such as livestock wastes.

Impact of Pesticides on Surface Water Quality

Pesticides can move into surface water by both point and nonpoint source pollution. Point source pollution occurs when pollutants originate from a single event or fixed site. Point source events include chemical runoff during improper storage, mixing/loading, disposal or misapplication to waterbodies. Nonpoint source pollution is the movement of pesticides from areas across watersheds over time into the surface water. Runoff (water flow) and erosion (soil particles) are common forms of nonpoint source pollution of surface water from agriculture. Pesticides are most susceptible to runoff immediately after application when they are in the “mixing zone.” The mixing zone is the thin layer (1/10 to 1/3 inch) at the soil surface.

Impact of Fertilizers on Surface Water Quality

Fertilizers given to crops are not always fully consumed; part remains in the soil absorbed by soil colloids and influences the quality of surface water when it is dissolved. Chemical fertilizers consist mainly of relatively simple compounds of nitrogen, phosphorus, and potassium, which are nutritive elements for plants. Nitrogen when that enters surface water tends to cause eutrophication of lakes and other surface water reservoirs.

Impact of Mining

In many countries, water pollution related to the mining industry can be traced back to older times compared with that caused by other industries. The characteristics of waste water from mining industries vary greatly depending on the kind of mine (3).

Metal Mining Industry. Many kinds of wastewater are discharged. For example, mine water, ore dressing water, wastewater from smelters, in general, contain metal ions, sulfuric acid and suspended solids at high concentrations.

Coal Mining Industry. Wastewater from coal mines consists mainly of mine water, coal dressing water, trace and major elements and contains a large quantity of suspended solids.

Sulfur Mining. In general, wastewater from sulfur mines contains large quantities of sulfuric acid and iron and has a low pH.

Stone-Quarrying and Clay-Supply Industries. These industries supply limestone, building stones, porcelain clay, and so on. Wastewater from the sites of these industries generally contains large amounts of suspended solids. Mining activities generate large quantities of solid wastes in most cases from large heaps. Such wastes not only pollute surface water by exuding toxic substances and also by rapid generation of suspended solids in surface water when the heap is washed away by a heavy rainfall or a flood.

Food Processing Industry

A common characteristic of wastewater from all types of food processing industries is the high organic content. Materials contained in wastewater can be divided roughly into carbohydrate, protein, fat, and mixtures of them. In many instances, wastewater contains oil, nitrogen, and phosphorus. In starch production, sugar refining, and brewing, the BOD level of waste water is especially high.

Paper and Pulp Industries

These are typical water consuming industries, and the levels of COD and suspended solids in wastewater are extremely high. A large quantity of waste paper fiber is also discharged as suspended solids together with wastewater.

Iron and Steel Industry

Wastewater from the cooling and cleaning processes for coke furnace gas contains ammonia, cyanide, and phenols. There is also wastewater from the dust collecting process of each furnace that contains suspended solids (coke dust and ore) and that from the pickling process, which contains acid, iron, and oil.

Electroplating Industry

Toxic substances released through wastewater are various heavy metals (cadmium, zinc, copper, etc.), cyanide, hexavalent chromium, acids, and alkalies.

Additional Forms of Water Pollution

Additional forms of water pollution exist in the forms of petroleum, radioactive substances, and heat. Petroleum often pollutes surface waterbodies in the form of oil from oil spills. Besides supertankers, offshore drilling operations contribute a large share of pollution. One estimate is that one ton of oil is spilled for every million tons of oil transported. This equals about 0.0001%. Radioactive pollution can be human-made or natural. Radioactive substances are produced as waste from nuclear power plants and from the industrial, medical, and scientific use of radioactive materials, or they can come from naturally occurring radioactive isotopes in water like radon. Specific forms of waste are from uranium and thorium mining and refining.

The last form of water pollution is heat. Heat is a pollutant because increased temperatures result in the deaths of many aquatic organisms. These increases in temperature are caused by discharges of cooling water from factories and power plants.

BIBLIOGRAPHY

1. Sharma, B.K. and Kaur, H. (1994). *Water Pollution*. Goel, Meerut, India, p. 56.
2. Snoeyink, V.L. and Jenkins, D. (1980). *Water Chemistry*. John Wiley & Sons, New York, p. 6.
3. Trivedi, P.R. and Raj, G. (1997). *Water Pollution*. Akashdeep, New Delhi, India, p. 11.

POND AQUACULTURE—MODELING AND DECISION SUPPORT SYSTEMS

Maria A. Szumiec
Polish Academy of Sciences
Chybie, Poland

SOLAR RADIATION AND TEMPERATURE IN SHALLOW STAGNANT WATERS

Physical environment plays a controlling role in all processes and events that occur in waters. Knowledge of water's dynamic characteristics and model elaboration is basic. Shallow stagnant waters are characterized by the individuality resulting from their small capacity. This is displayed in their receptivity to external factors, such as weather variability, that differentiates them from lakes. In a temperate climate, quick and distinct changes in meteorological events are displayed in variations in the solar radiation that penetrates into shallow waters, thus, in variations in light and thermal conditions. These changes are then transferred directly and/or via chemical processes to different levels of the biological chain. In fishponds, the final link in the chain is the stenothermic organism which is affected by the ambient temperature in a complex way. Penetrating light plays the decisive role in fish survival, limiting the concentration of oxygen dissolved in the water. The thermal conditions inhibit development and growth of the warm water fish, such as carp, throughout the considerable part of the year.

An attempt to approach the time-vertical distribution of penetrating solar radiation and temperature is presented in descriptive form and in simple mathematical language. The majority of the material is obtained from constructed carp ponds situated in a temperate climate ($\phi = 49°52'\mathrm{N}, \lambda = 18°48'\mathrm{E}$). Both the differentiated pond morphometry and the various periods of their use allow following their effects on the physical environment. Their average surface area is equal to some thousand square meters, and the average depth varies from 120 cm in the warm season to 200 cm in winter. The simultaneous time-vertical distribution of temperature is presented in ponds whose surface areas are from some hundred to some thousand square meters and from 20–120 cm deep (Fig. 1). The gravitational water flow through the ponds requires inclination of their bottoms from the water inlet to its outlet. In deeper ponds, the difference in depth between shallow and deep places can exceed 1 m. Dikes one to two meters high above the water surface limit the wind stress, especially in smaller ponds.

Two periods limited by water temperature are distinguished; warm—from the beginning of April to the end of October, when the water temperature is above 4 °C, and cold—from November till March, when the temperature drops below 4 °C. Most of the examples presented deal with period from May till September, when the thermal conditions stimulate growth and development of the carp organism.

In the warm part of the year the water feeding ponds compensates only for evaporation and infiltration losses. During the winter, the scanty water flow through ponds covered with ice protects them from oxygen deficiency. The solar radiation flux density was monitored in the wavelength range 0.3–3.0 μm.

The Warm Season

Absorbed solar radiation is the main source of heat in shallow waters; the greatest heat losses take place in the evaporation process (1). In early spring, solar radiation

Figure 1. View from the air over the ponds of the Institute of Ichthyobiology and Aquaculture of the Polish Academy of Sciences at Gołysz.

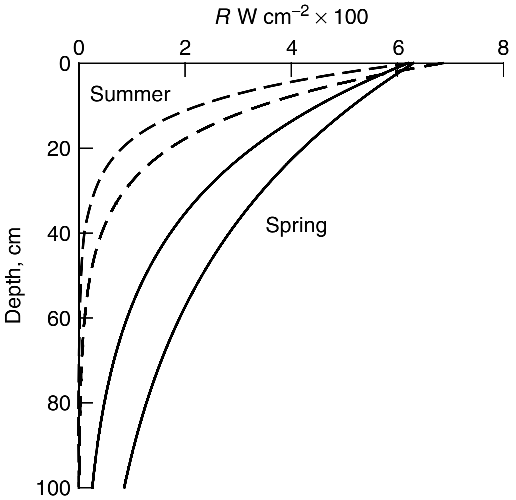

Figure 2. Density of the vertical solar radiation flux R under the free water surface in ponds during sunny and calm days; higher extinction presents ponds with a denser carp stock.

penetrates to the pond bottom. At night, cooling of the surface layers makes the whole water stratum circulate; a breeze enhances this process. Favorable light conditions during the day and nutrients released from the bottom during water circulation stimulate the development of flora and fauna and distinctly diminish the radiative flux penetrating to the bottom. Intensive absorption of the long waves and intensive scattering of the short waves of the radiative spectrum lead to high extinction in the surface layers and to an exponential decrease of R with the pond depth z (Fig. 2). On sunny days under the free water surface, $R(z)$ takes the form of

$$R(z) = R_0 e^{-\alpha z}$$

The coefficient α of vertical extinction is equal to 0.02 cm^{-1} in spring and increases in late summer; it also increases in ponds densely stocked with carp (Fig. 2) (2,3).

Values of α in ponds are approximately ten times higher than in most lakes, where they usually are computed per 1 m (4–6). A wind whose speed exceeds 3 m s^{-1} stimulates the circulation of the entire water stratum and also limits the penetration of vertical radiation to a certain degree. This distinct stratification of vertical radiation manifests in a distinct vertical decrease in temperature and oxygen content in ponds (Fig. 3).

In late summer in the deeper parts of ponds fertilized or densely stocked with carp, the oxygen concentration decreases below the saturation point (7). Macrophytes limit the radiation penetrating into ponds; however, the largest decrease occurs below dense floating plants. During sunny, calm days, the regular diel and seasonal solar radiative cycles follow the Sun's elevation in the whole water stratum (Fig. 4).

They can be approximately represented at all the depths by the geometric function

$$R(z,t)_d = \{A_R(z,t)\sin[\pi/T_R(z,t)]\}_d$$

Amplitude $A_{R,d}$ and period $T_{R,d}$ decrease approximately exponentially with depth. A decrease in $T_{R,d}$ shortens the day at the bottom (8). The seasonal cycle of $R(z,t)_s$ follows the Sun's elevation only in the surface water layer (Fig. 5). In the deeper parts, the average monthly radiative flux R decreases from May to September owing to decreasing water transparency.

An equation for the distribution $R(z,t)_s$ of seasonal–vertical solar radiation in ponds during sunny day is not a simple geometric function, though it can be satisfactorily computed using the first terms of the Fourier series

$$R(z,t)_s = \{\Sigma_n A(z)\sin[n\pi/T(z)t + \varphi(z)]\}_s$$

where $(A, T, \varphi)_s$ are the seasonal amplitude, period, and phase of $R(z,t)_s$ in the period from May–September.

The greater the surface area of ponds, the greater the role of wind in water circulation and in radiation extinction (9). Momentary cloudiness deforms the time–vertical

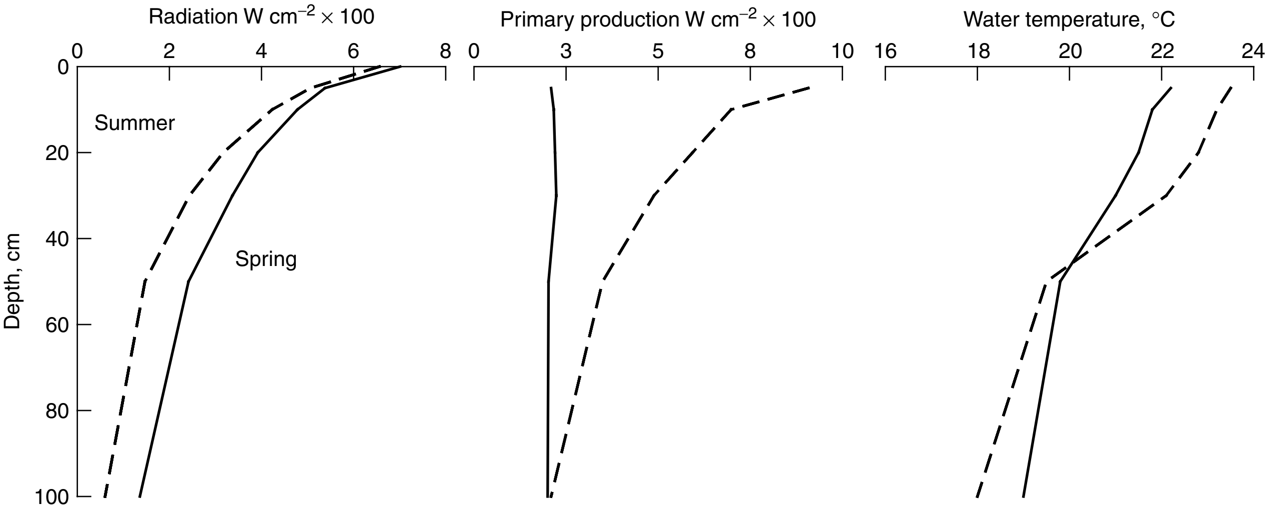

Figure 3. Example of the simultaneous vertical distribution of radiative flux, primary production, and water temperature in ponds on a sunny, calm summer day.

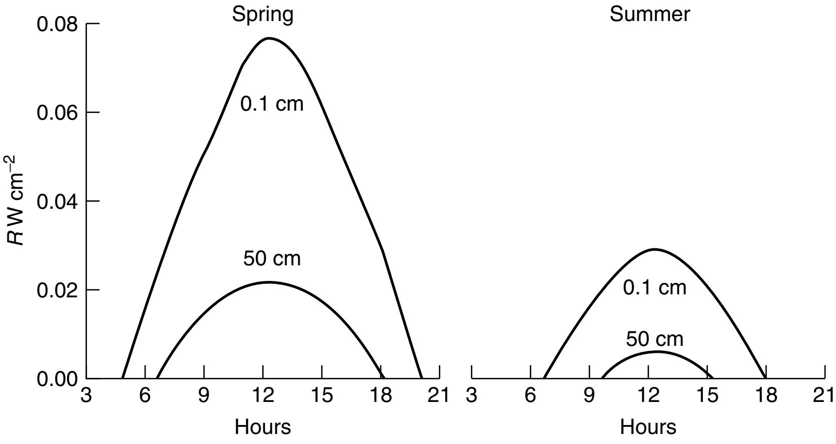

Figure 4. Diel cycles of radiative flux in ponds on sunny days.

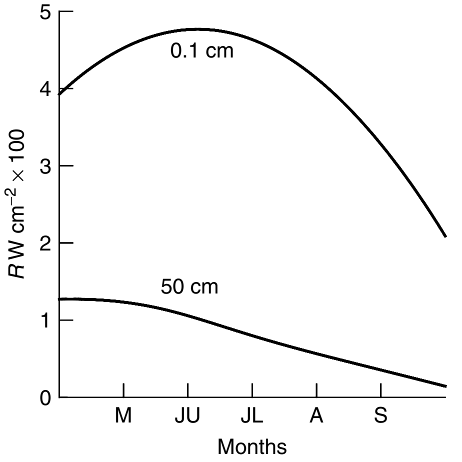

Figure 5. Seasonal cycles of radiative flux in ponds on sunny days.

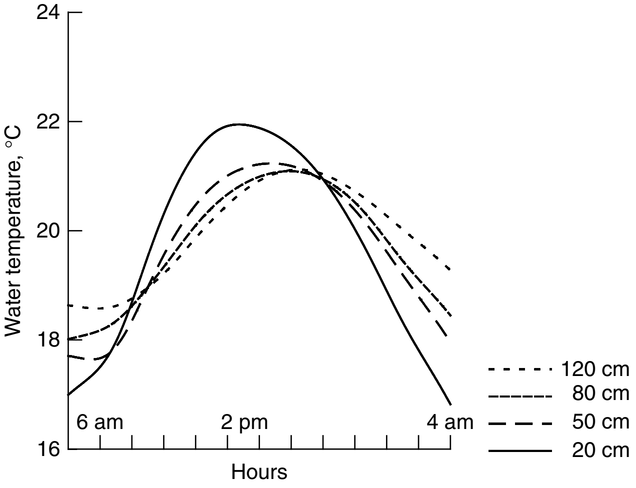

Figure 6. Simultaneous smoothed diel temperature cycles in ponds of different average depths on a sunny, calm summer day.

radiative flux and diminishes its density; cloudiness maintained for a whole day causes the daily radiative amplitude to approximate zero.

In shallow waters, the time–vertical distribution of the temperature ϑ is chiefly a resultant value of the absorbed radiation and heat losses in the evaporation process. During sunny, calm days, temperature cycles are semiregular geometric functions; that have with a slightly quicker increase in temperature during the morning hours and a distinctly slower decrease in the afternoon (Figs. 6 and 7). During the night, the temperature difference between surface and bottom water layers approaches zero. The diel amplitude and vertical gradient decrease exponentially as the average depth z of the pond increases (10,11).

In May, when sunny but cold northern air masses advect, the temperature in very shallow waters, for example, about 20 cm deep, can decrease to a few degrees in the early morning and increase to over 25 °C at noon. These great time variations may be lethal for some organisms, such as e.g., carp fry (12).

The Cold Part of the Year

In late autumn and in winter, the radiative flux reaches only a few percent of that recorded in summer and is additionally limited in ponds covered by ice. Only transparent ice allows radiation to penetrate to the bottom and to create favorable conditions for primary production. Under nontransparent ice, water flow through ponds is required to improve oxygen conditions. In non frozen ponds, the action of wind in stirring water ensures sufficient oxygen concentration for the fish.

The first autumnal deep cooling equalizes and lowers the temperature in the whole water stratum to 4 °C. A further temperature decrease is manifested in inverse stratification. Strong wind accompanied by cooling stirs the water, obstructs freezing, and lowers the temperatures in the whole stratum, sometimes even below 0.1 °C. In ponds covered by ice, the temperature in the lowest water layers is heated from the bottom and usually increases to 2–2.5 °C. Such vertical temperature distribution is normally maintained from the end of December to early March. The warm, windy weather in the second half of this month removes the ice cover and raises the water temperature to 4 °C, equalizing differences between the surface and the bottom (10).

In a temperate climate, the ice cover usually disappears one or more times from ponds (13), at first leading to momentary water heating and, under the consecutive

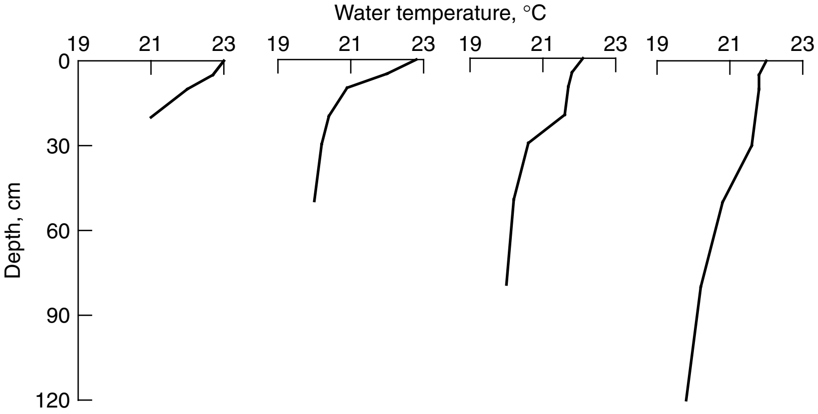

Figure 7. Simultaneous vertical temperature distribution in ponds of different average depths on a sunny, calm summer day.

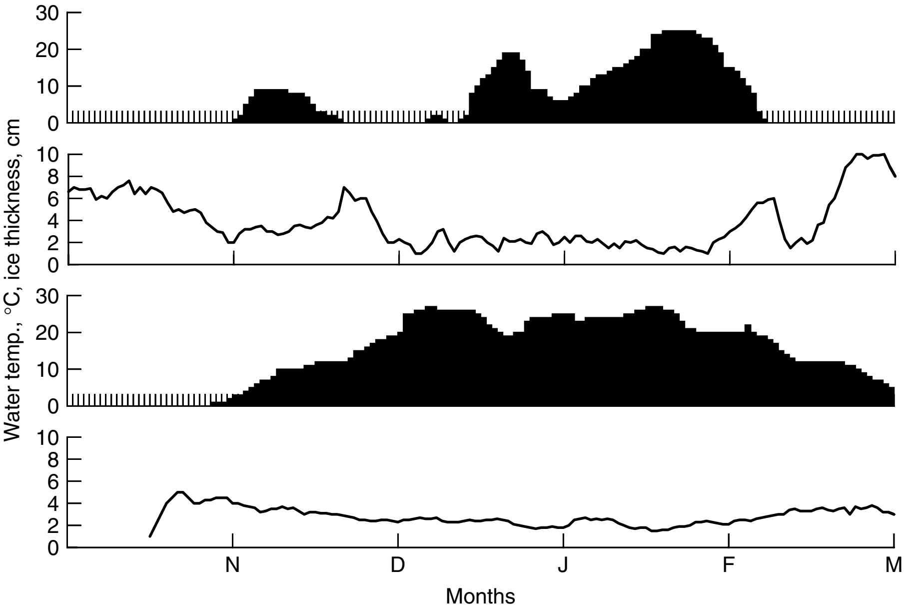

Figure 8. Examples of the average diel temperature patterns in the pond's bottom water layer during winter for momentary disappearances and permanent ice cover.

cooling, to a deep temperature decrease (Fig. 8). This creates unfavorable conditions for warm water organisms, for example, it increases the mortality of carp juveniles.

SUMMARY

The individuality of shallow stagnant waters, such as carp ponds, results from their small capacity and differentiates of their physical environment from lake environments. The density of the solar radiative flux that penetrates into water and the vertical temperature and oxygen stratification in ponds situated in a temperate climate are characterized by high extinction and vertical gradients, approximately ten times higher than in most lakes. Vertical extinction and attenuation increase from spring to late summer. Despite its high variability, average diel radiative courses can be represented by a simple geometric function, the seasonal distribution by the first several terms of the Fourier series. The diel cycle of water temperature in ponds corresponds to the annual cycle in lakes. The effect of increasing pond depths is displayed in the exponential decrease of the diel temperature amplitude and vertical gradient. During winter in ponds covered by ice, the temperature in the bottom water layer is maintained between 2 and 2.5 °C. Short warm weather periods lead to the momentary disappearance of ice and to a temperature increase, the consecutive cooling—to a considerable temperature decrease.

BIBLIOGRAPHY

1. Szumiec, M. (1969). Bilans cieplny stawów karpiowych na przykładzie Gospodarstwa Gołysz. (Thermal balance of carp ponds by example of Gołysz Experimental Farm). *Acta Hydrobiol.* **11**: 137–178.
2. Szumiec, M.A. (1982). The influence of meteorological conditions and eutrophication on some physical elements of the water bodies environment. *Acta Hydrophysica* **27**: 217–221.
3. Szumiec, M.A. and Augustyn, D. (1985). Studies on intensification of carp farming 3. Meteorological conditions, solar radiation, and water temperature in ponds. *Acta Hydrobiol.* **27**: 159–172.
4. Stenz, E. (1953). Strahlungsdurchlässigkeit einiger Seen Polens. *Acta Geoph. Pol.* **1**: 44–55.

5. Hutchinson G.E. (1957). *A Treatise on Limnology*. Vol. I. John Wiley & Sons, New York, p. 1015.
6. Szumiec, M. (1975). Wpływ kierowanej eutrofizacji na energię słoneczną przenikającą w głąb stawów. (The effect of controlled eutrophication on solar radiation penetrating into ponds). *Acta Hydrobiol.* **17**: 149–182.
7. Szumiec, M. (1971). Solar radiation, water temperature and primary production in carp ponds. *Pol. Arch. Hydrobiol.* **18**: 175–187.
8. Szumiec, M.A. (1986). Weather impacts on carp farming. *Aquaculture Hungarica (Szarvas)* **5**: 179–189.
9. Szumiec, M. (1981). Wpływ cech morfometrycznych zbiorników typu stawowego na bilans ciepła i temperaturę wody. (The influence of the morphometry of pond—type reservoirs on the heat balance and water temperature). *Przegl. Geofiz.* **26**: 143–160.
10. Szumiec, M.A. (1984). *Termika stawów karpiowych. (Thermal properties of carp ponds)*. PWN, Warszawa, p. 136.
11. Goubier-Martin, M. (1991). *Climat et aquaculture. Application aux etangs de la Dombes. Doctorat de Geographie sous la direction de Mr. Marcel Leroux, Professeur. Universite Jean Moulin—Lyon III, Faculte des Letters et Civilizations*. Laboratoire de Geographie Physique, p. 483.
12. Szumiec, M.A. and Białowąs, H. (2002). *Effect of temperature and genetics on carp juvenile survival in ponds. EAS Conference "Pond Aquaculture in Central and Eastern Europe in the 21st Century*. Special Publication No. 31, 71–74.
13. Szumiec, M.A. and Augustyn, D. (1997). Warunki hydrometeorologiczne w sezonie hodowlanym 1996 i w zimie 1995/1996. (Hydro-meteorological conditions in the farming season 1996 and in winter 1995/1996). *Komunikaty Rybackie* **2**: 12–16.

PUMPING STATIONS

Joseph D. Bankston
Louisiana State University
Baton Rouge, Louisiana

Surface water pumping stations are employed either to move water from an available surface location to where it is wanted or to remove it from a location where it is not wanted. Examples of the former include municipal or industrial water supply and irrigation, whereas municipal drainage pumps are examples of the latter.

Although differing in purpose, all have similarities of function and components. A pumping station must provide a water inlet for water to flow to the pump, a pump suitable for the water and delivery requirements, and a means of powering the pump. Although these requirements are interdependent to some degree, a reasonable approach is to consider them in order.

WATER INLET

Surface water frequently carries with it flotsam, jetsam, and aquatic species, both plant and animal, which must be prevented from entering the pump without severely restricting the flow of water. In addition, surface water levels often fluctuate. Settling basins, reservoirs, grids, screens, and multiple intakes or floating intakes have been used to address these problems.

Reservoirs and settling basins even out the water supply reducing fluctuation in the water level and because of reduced velocity of water flow, allow debris either to settle out or float to the surface. The reduced water flow also reduces the impact from objects in the water, which reduces the required strength of the intake structure.

It is preferable to obtain the water from a high-quality spot. For example, in many reservoirs, a preferred area is about 1 meter below the surface. As the level of the reservoirs change, this entails for a fixed structure, multiple ports located at different levels and valved so that the desired level may be chosen. Low entering velocities are desirable to avoid heavy loading of floating matter, sediment, and fish. Velocities of less than 0.5 ft/sec are preferable. Higher velocities result in heavier loading and greater maintenance requirements of the intakes. Current and foreseeable environmental conditions should be considered when siting the intakes. For example, winds that stir up sediments near shore may necessitate siting the intake away from the shoreline, as may pollution caused by shoreline development. Submerged intakes are possible if they do not interfere with navigation. Some small stations may employ floating intakes in which the intake is mounted on a floating structure such as a raft and connected to the pump with flexible pipe.

Pumps are guarded by physical barriers such as inlet grids or screens located upstream of the pump. These are usually arranged in series, beginning with those that have the largest openings and, by necessity, are the largest and strongest and proceeding to more restrictive barriers. These barriers must be cleaned periodically to prevent the accumulation of debris from severely restricting the water flow. Cleaning may be manual or automatic.

PUMPS

A pump (or pumps) is the second major component of a pumping plant. A pump must be capable of producing the required head while delivering the desired flow rate. It must also be able to handle the suction lift and the water quality or lack thereof. The latter is dependent on the water—its corrosiveness (salty water or chemical laden) and how much and what kind of trash or debris it may contain. An appropriate pump is constructed of materials that resist the corrosive characteristics of the water and of a design that is capable of handling the anticipated debris.

The head, suction lift, and flow rate are dependent on the piping system and the pump's characteristics. The piping system and pump interact to determine the operating point of the pump—flow rate and pressure. The pump cannot independently control these parameters. As the flow rate increases, the head (pressure) required to force the water through the pumping system also increases. A pump will typically have reduced capacity (flow) as the head it is pumping against increases. Thus, as the flow rate increases, the pressure a pump is capable of producing decreases while the pressure required to move the water increases. The operating point occurs at that flow rate under which the pressure produced by the pump is equal to the pressure required to move the water. The efficiency of a pump varies greatly with its operating point, with maximum efficiency occurring at a particular flow rate and falling off at higher or lower flow rates.

Manufacturers typically list an operating range in which the pumps performance is satisfactory; outside this range, the efficiency of the pump may be low and damage to the pump could occur. The required output of a pumping plant often varies over time. In order to avoid operating outside of the pumps designed range, a pumping station may employ multiple pumps or, when feasible, storage. Multiple pumps, although more costly initially, allow the flexibility of operating the pumps at an efficient operating point and also supply redundancy in case of pump failure.

Water is stored to equalize pumping rates over time. One example of storage is a system of drainage canals that themselves are drained by a pump (as in New Orleans, where much of the city lies below sea level and water must be pumped to provide drainage). The drainage canals provide storage volume as well as a means of flowing the water to the pumping station. Economies result from operating smaller pumps at their rated loads rather than large pumps at conditions less than optimum for the pump.

Two types of pumps are typically used in pumping stations—centrifugal and axial flow. Centrifugal pumps use centrifugal force to move water from one point to another and to develop pressure. Water enters at the center of an impeller, where it is spun and forced at a high velocity to the outer edge by the rotating impeller. The water is discharged into a casing where some of the velocity is converted into pressure. In an axial flow pump, both the water flow and the force exerted on the water by the pump is in a straight line along the axis of the pump. Axial flow pumps use a propeller or screw to apply force to the water. Axial flow pumps are particularly suited for high-volume low-head (pressure) applications. Their efficiency is high, especially when the head is in the range of 8 to 20 feet. An advantage of an axial flow pump is its ability to handle some debris. Centrifugal pumps, particularly high efficiency ones, often have close tolerances and may not handle trash as well. Centrifugal pumps are capable of operating efficiently at higher heads than axial flow pumps.

Suction lift is suction that a pump must apply to suck water into its inlet. It includes the height the pump is located above the water source, the friction loss in the pipe leading from the water source to the pump, and the energy imparted to the water as velocity. The amount of suction lift that a pump can produce is limited and varies with the individual pump. Excessive suction lift can cause inefficient operation, damage to the pump, or a failure to operate. For these reasons, it is preferable for the pump to be immersed in the water source or for the water to flow to the pump under gravity.

POWER

Pumps are powered by electric motors or internal combustion engines. Internal combustion engines include diesel, gasoline, natural gas, and liquefied petroleum gas (LPG) engines, with diesel being the most common. Each power source has advantages and disadvantages, many of which are site- and application-dependent. Deciding factors may include ability to do the job, reliability and availability of the power source and fuel supply, cost of equipment and installation, expected useful life, cost and ease of maintenance, convenience of operation, and the cost of energy to operate.

With the large variety of motors and engines available, a suitable example of either should be available and capable of powering the chosen pump. In general, internal combustion engines become more applicable as the power requirements and hours of use increase. Fuel efficiency of higher horsepower engines is usually better than for low horsepower engines (when each is properly matched to the load), and the higher fixed costs of engines as compared with motors can be spread over more operating hours as the hours of use increase. In addition, engines may be operated at varying speeds that provide more flexibility than most electric motors.

Advantages to electric motors exist if the cost of providing service and the cost of electricity, including energy, demand, and standby charges, are not prohibitive. The electric motor provides ease of operation, long life, requires minimal maintenance, and maintains its performance level year after year. In addition, initial costs are usually less than the cost of internal combustion engines, and the reliability is higher. They can, however, be shut down by the loss of electrical power. If the ability to operate during times when power is frequently lost (for example, during periods of flooding after a hurricane) is a necessity, or the pumping station is located in an area that suffers from frequent power outages, the internal combustion engine may be a better choice.

Natural gas engines and electric motors are dependent on the availability of natural gas service and electric service, respectively. If this service is not available, or the cost of obtaining it is too great, than diesel, gasoline, or LPG engines, which can be supplied with fuel from storage tanks, become the feasible power sources. Storage tanks allow considerable freedom in siting the power unit, but they also present additional concerns. The tanks should be designed to prevent pollution caused by leaks, spills, or evaporation, which includes guarding against leaks or spills and provides a means of containing the spill and of cleaning after a spill. For this reason, underground tanks are usually avoided. Above-ground tanks have their own sets of problems. They are more vulnerable to attack or vandalism and tend to have a greater degree of fuel evaporation than underground tanks. Not only does evaporation result in fuel loss, it may also lead to higher gum content of the fuel and poor engine performance and clogged filters, which, if severe enough, can prevent an engine from operating. Adulteration of the fuel can occur by the condensation of water inside the tank or, in the case of diesel fuel, by bacteria that feed on the fuel.

Energy costs are estimated under the conditions at which the pumping station operates. A thorough analysis requires the expected operating schedule including loading, time of occurrence, and duration of operation. Loading is particularly important in determining internal combustion engine fuel usage as the efficiency of delivered power is greatly impacted by the degree to which the engine is loaded. For example, Dr. Claude Boyd of Auburn

University obtained comparative efficiency data of an engine under different loading conditions. The engine, running at 1800 rpm, required 1.6 gal/hr of fuel to provide 4.9 horsepower. Increasing the load to 16.9 horsepower, a greater than three-fold increase, required 2.0 gal/hr, or only 25% greater fuel consumption. Similarly, the same engine running at 950 rpm required 0.7 gal/hr to supply 4.8 horsepower and 1.2 gal/hr to supply 16.7 horsepower. Under these different loading conditions, the number of horsepower hours produced by the same engine for each gallon of fuel consumed ranged from 3.1 to 13.9. Although not generally as sensitive to loading as internal combustion engines, electric motors also fall off in efficiency when underloaded, particularly at 50% or less of capacity. Energy costs for electricity is impacted by operating schedule as electric charges for commercial customers include demand and service charges in addition to energy costs. Demand and energy charges are charges for having electric power available, even if it is not used.

Maintenance costs can significantly affect the viability of a particular power option. Maintenance costs are considerably higher for internal combustion engines than for electric motors. Exact costs are difficult to pin down as they depend on duty cycle, cost of labor, degree of maintenance, and the individual engine or motor. A rough rule of thumb is to access internal combustion engines a 1 to 1.5 cent per horsepower-hour maintenance cost penalty in comparison with electric motors.

READING LIST

Auley, R.H. and Knutson, G.D. (1993). Cost comparison: engine vs. electric, *Irrigation J.* May-June.

Bankston, J.D. and Baker, F.E. (1995). *Powering aquaculture equipment*, Southern Regional Aquaculture Center Publication No. 375.

New, L.L. (1992). *Pumping Plant Efficiency and Irrigation Costs*. Texas Agricultural Extension Service, The Texas University System.

Electric Power Research Institute. (1988). *Small Cogeneration System Cost and Performances*, Project EM5954.

REGULATED RIVERS

Geoffrey Petts
University of Birmingham
Birmingham, United Kingdom

The regulation of river flows by dams and reservoirs has been fundamental to the advancement of civilization, and as dryland development continues, river regulation for water supply and power production continues to be central to economic growth. However, river ecosystems have evolved to become adapted to the natural pattern of flow variability from season to season and from year to year. Flow regulation has had dramatic impacts on both running-water and floodplain ecosystems, resulting in major changes in biodiversity and to fisheries.

River regulation involving the redistribution of water in space and time has been fundamental to the advancement of human society (see Water in History). Dams have been built to impound rivers and store runoff in large reservoirs for use in dry seasons, in drought years, for long-distance transfer to centers of high water demand—for irrigation agriculture or urban use—and for hydroelectric power production. It has been estimated that 800,000 dams of all sizes have been built on the world's rivers. The largest have dams over 150 m high and reservoirs with storage capacities greater than 25,000 million cubic meters or have a hydroelectric generation capacity of more than 1000 MW. Today, large dams contribute directly to about 15% of global food production and nearly 20% of the world's electricity supply (1). However, these dams, reservoirs, and diversions have had widespread ecological impacts. Worldwide, reservoirs have drowned about 1 million km^2 of land and in the United States 25,000 km of fast-flowing main rivers have been converted into stillwater lakes and on the Columbia-Snake system alone, 822 km of river habitats have been replaced by reservoirs. Here, the focus is on the regulated rivers downstream from dams.

By 1994, it was estimated that nearly 80% of the large river systems in Europe, the former Soviet Union, the United States, and Canada were at least moderately regulated by dams, reservoirs, and diversions (2). In the United States, the nation's major rivers, the Mississippi, Colorado, and Columbia, are completely regulated by a combination of large dams and weirs. Impoundments trap a river's sediment load and about 25% of the global flux of river sediments is trapped in reservoirs. Reservoirs can also alter the water quality of the river downstream of the dam, especially the annual pattern of temperature variations. However, the major impact of dams and reservoirs is to regulate the flow regime along the river below the dam and in many cases also to significantly reduce the total runoff (3).

In the 1990s, worldwide annual water withdrawals of 4430 km^3/yr represented between 35% and 50% of the average annual renewable resource, and irrigation agriculture consumed some 2500 km^3 of water. Obvious indicators of unsustainable water use include the failure of the Hwang He and Colorado Rivers to reach the sea in dry seasons and the shrinkage of the Aral Sea by 25,000 km^2. In the case of the Nile River, the influence of the Aswan Dam on the flow regime extends for 1000 km, changing the hydrographic conditions over the continental shelf of the southeast Mediterranean. Ecological indicators include the loss of productive floodplain wetlands, dramatic changes of channel size and shape, the degradation of river and floodplain fisheries, and the incursion of saltwater into coastal floodplain and delta areas.

Flow regulation includes the reduction of average annual runoff; reduced seasonal flow variability or reversal of the seasonal pattern to meet irrigation demands; altered timing of annual extremes, reduced flood magnitudes, and the imposition of unnatural, short-duration pulses associated with hydropower generation. Until the 1960s, most dams and reservoirs were designed and operated with little regard for the river and its flow-dependent ecosystems. In most cases, the only discharges from the dam were compensation flows—often a constant minimum flow—to meet the needs of industry along the river, especially for water mills, and other riparian uses, with higher flows during the rare periods of dam spill.

During low flow periods, many regulated rivers became dry until unregulated tributaries restored some flow. A survey in 1987 by the Institute of Hydrology of supply, regulating, and hydropower reservoirs in the United Kingdom showed mean compensation flows to the river below the dam of 15–17% of average daily flow, but for reservoirs operated only for water supply the compensation flow was zero. In dryland areas, often minimum *flows* were increased during the natural low-flow season to meet the needs of irrigation and other abstractors downstream, and along many rivers minimum water *levels* have been raised by weirs to meet navigation requirements.

All reservoirs regulate floods. The basic concept of flood regulation is empty space and reservoirs having a large flood-storage capacity in relation to the annual runoff can exert complete control upon the annual hydrograph of the river downstream. However, even when a reservoir is full to spillweir level, temporary storage of floodwaters within the lake plays an important role in reducing the maximum rate of outflow from the reservoir when the water area of a reservoir is 2% or upward of the catchment area. Along many regulated rivers, the mean annual flood has been reduced by more than 25% and there are examples where the 50-year flood has been reduced by over 20%.

Flood regulation and the maintenance or increase of minimum flows has led to major changes in the size and shape of river channels (4). Degradation has occurred immediately below some dams because of the trapping of the sediment load behind the dam, but the reduction of channel size in response to the reduced flood magnitudes has been dramatic—channel capacities being reduced by 50% or more in some cases. Rates of channel narrowing have been particularly rapid along some semiarid rivers where the maintenance of perennial flows encouraged the rapid establishment and growth of woody vegetation on exposed sediments. Elsewhere, rapid aggradation has occurred at confluences between regulated rivers and tributaries carrying high sediment loads forming deltas and bars that can extend for several kilometers below the confluence.

In the past, the most disturbing feature of regulated rivers was that the environmental changes resulting from regulation were not anticipated. A scientific journal, *Regulated Rivers*, was founded in 1988 to provide a focus for synthesizing studies of the environmental impacts of river regulation. In 2002 the journal changed its name and emphasis to *River Research and Applications*, reflecting major changes in both research on, and management of, regulated rivers. For 25 years attention had been moving to the determination of flows required to protect water-dependent ecosystems. By the start of this millennium there had developed a vanguard movement to restore regulated rivers and to reconnect rivers with their floodplains; to advance decision-making processes to balance consumptive water use with that needed to maintain a healthy river; and to allocating our limited water resources to meet the needs of both humans and riverine ecosystems (5). This issue is addressed in this volume under Environmental Flows.

BIBLIOGRAPHY

1. WCD (2000). *Dams and Development: A Framework for Decision-making*. World Commission on Dams, Earthscan, London.
2. Dynesius, M. and Nilsson, C. (1994). Fragmentation and flow regulation of rivers in the northern third of the World. *Science* **266**: 753–762.
3. Petts, G.E. (1984). *Impounded Rivers*. John Wiley, Chichester, UK.
4. Petts, G.E. and Gurnell, A.M. (2005). Dams and geomorphology. *Geomorphology*, Washington, DC.
5. Postel, S. and Richter, B. (2003). *Rivers for Life*. Island Press, Washington, DC.

RESERVOIRS-MULTIPURPOSE

Sharad K. Jain
National Institute of Hydrology
Roorkee, Uttranchal, India

NEED FOR RESERVOIRS

Reservoirs are developed for two main functions. The first is to store water to even out the fluctuations in river flow and match the availability to demand. Generally, the major part of annual stream flow is available during the few months of the rainy season, but the demand for water arises all year round. Therefore, it is necessary to store the excess water in the months of high flows and use it when the natural stream flow is inadequate to meet the demand. The water stored in a reservoir may also be diverted to faraway places where it is needed.

The second function is to raise the level of the water upstream of the dam to provide a hydraulic head. The creation of storage and head allows dams to generate electricity. The other important purpose of a reservoir is to control flooding in downstream areas.

The early Mesopotamians were the first dam builders. The Sumerians built networks of irrigation canals in the plains along the lower Tigris and Euphrates Rivers about 6500 years ago. Earth and rock-filled dams had been built around the Mediterranean, in the Middle East, China, and Central America by the late first millennium B.C. The remains of impressive dams and aqueducts in Spain are fine examples of the ingenuity of Roman engineers. South Asia has a long history of dam building. Long earthen embankments were built to store water for irrigation. The remains of the Indus Valley civilization which flourished 4000 to 5000 years ago show that these people had built well-planned networks of water supply and drainage works. The oldest reservoir in operation today is the Aftentang reservoir (capacity 100 million m^3), constructed west of Shanghai during 589 to 581 B.C. (1). The oldest continuously operating dam still in use is the Kofini flood control diversion dam and channel constructed in 1260

B.C. on the Lakissa River upstream of the town of Tiryns, Greece, which it continues to protect.

Worldwide, there were more than 36,000 large dams by 1986 (2); many more have been constructed since then. The Asian continent accounts for more than 64% of all dams, and China has built most of them. Engineers of the former USSR have built most of the large reservoirs; they are followed by engineers from Canada. More than 80% of the dams in the world are earth and rock-filled types.

A reservoir contains a number of structural features other than the main dam. Spillways are used to discharge water when the reservoir level threatens to become dangerously high. Dams built across broad plains may include long lengths of ancillary dams and dikes. A schematic diagram of a reservoir is given in Fig. 1.

A reservoir project requires huge amounts of money, manpower, land, and other resources and significantly affects the environment, population, and economy of the region. Once these projects are in place, it is not easy to undo their impacts. Therefore, they should be carefully planned to impart the maximum possible benefit to the national economy.

Multipurpose Reservoirs

The typical purposes of a reservoir are water supply for domestic and industrial uses, irrigation, navigation, generation of hydroelectric power, recreation, and flood control. Depending on the purposes served by the reservoir, it may be classified as either single purpose or multipurpose. A single purpose reservoir serves only one purpose. A multipurpose reservoir serves a combination of purposes. As a rule of thumb, the bigger a reservoir, the more purposes it can serve.

While planning a reservoir, all available data should be analyzed, and, if necessary, further information should be gathered so that the best decisions are made with respect to location, size and type of structure, and auxiliary facilities. After the best site is selected from amongst the potential sites, additional investigations are carried out to finalize the project details. These include engineering, geologic and hydrologic investigations, and economic analysis. The water yield at a site is commonly estimated using flow duration curves.

Storage Zones in a Reservoir

For ease of analysis and operation, the entire reservoir storage space is conceptually divided into a number of zones by drawing imaginary horizontal planes at various elevations. The lowest zone is the dead (or inactive) zone. The bulk of the storage capacity for conservation is provided in the conservation (or active) storage zone. The top level of the conservation zone is termed the full reservoir level (FRL) or normal pool level. If the storage space above the FRL is exclusively reserved for flood control, the maximum storage capacity is dead plus active storage. The maximum water level (MWL) is the highest elevation to which the water is allowed to rise in a reservoir.

Dead storage is provided in a reservoir for accumulating sediments or to provide the minimum head for the turbines of a hydropower dam.

Reservoir Storage–Yield Analysis

The procedures estimating the storage capacity needed to meet given demands or the possible yield from a given project constitute the storage–yield analysis. The reservoir storage–reliability–yield relationship relates inflow characteristics, reservoir capacity, release, and reliability. The inflow process, the reservoir storage, and the outflow process constitute the stream-flow regulation system. Its components are the reservoir storage capacity (S), yield (q), and a measure of reliability (R). The relationship among these variables can be symbolically designated by a function (3):

$$\phi = \phi(S, q, R) \tag{1}$$

where $S > 0, q > 0, 0 < R \leqslant 100\%$. This function is called the regulation regime function, or storage–yield function.

Critical period techniques are the earliest methods of storage–yield analysis. The critical period is the period in which an initially full reservoir, passing through various states (without spilling), empties. The *mass curve method* proposed by Rippl was the first rational method for computing the required storage capacity of a reservoir. This simple method is commonly used in the planning stage. Define a function $X(t)$ as

$$X(t) = \Sigma_t x(t)\,dt \tag{2}$$

where $x(t)$ represents inflows. The graph of $X(t)$ versus time is known as the mass curve. The reservoir capacity is obtained by finding the maximum difference between cumulative inflows and cumulative releases for the most critical period of recorded flow. This method, although

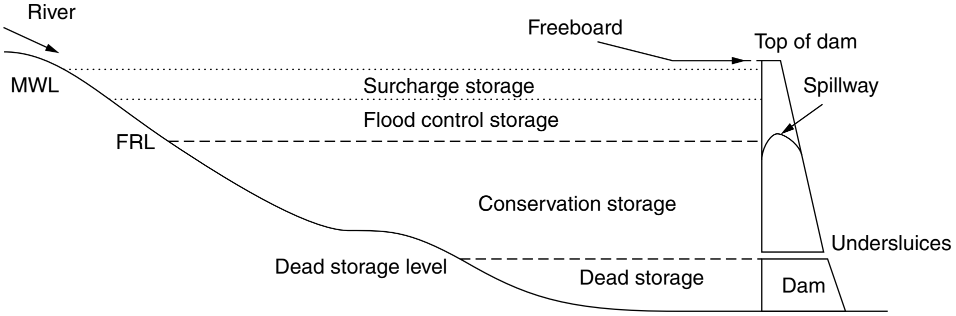

Figure 1. Schematic diagram of a reservoir.

simple and straightforward, has a few shortcomings. It is suitable when the draft is constant, and it is not possible to consider evaporation losses properly. The sequent peak algorithm overcomes some of the limitations of the mass curve method.

Systems analysis techniques such as optimization and simulation are frequently used for storage–yield analysis. Among the optimization techniques, linear programming and dynamic programming have been particularly suitable. Simulation is a widely used technique for solving a variety of problems in designing and operating a water resources system. The reason is that this approach can be realistically and conveniently used to examine and evaluate the performance of a set of alternative options. The simulation approach can be used as a stand-alone method or can be used to modify further and test the results of critical period or optimization methods. McMahon and Mein (4) describe storage–yield analysis techniques in detail.

Flood Control Storage Capacity

The requirement of storage for flood control conflicts with the requirements for conservation. Conservation requirements, such as water supply and hydropower generation, require the storage space to be full, whereas flood control requires empty storage space.

The demands for water supply and hydroelectric power are relatively deterministic, whereas the demand for flood control storage is highly stochastic. Furthermore, the usual time period for analysis of conservation is 1 month, whereas for flood control, it is of the order of a few hours. The storage requirement can be estimated by using the design flood hydrograph. An initial reservoir level at which this flood hydrograph impinges on the reservoir is assumed. The maximum level attained by the reservoir is computed by routing the hydrograph through the reservoir. The maximum height of a dam is obtained after adding the freeboard to this level.

RESERVOIR OPERATION

The efficient use of water resources requires judicious design and also proper operation. After a reservoir comes into being, the benefits that could be reaped depend largely on how well it is operated. Reservoir operation forms an important part of planning and managing a water resources system. Postconstruction, detailed guidelines are given to the operator which enable him to make appropriate management decisions. A reservoir operating policy specifies the amount of water to be released from storage at any time depending on the state of the reservoir, demands, and any information about the likely inflow to the reservoir. The operating problem for a single-purpose reservoir is to decide about the releases to be made, so that the benefits for that purpose are maximized. For a multipurpose reservoir, additionally, it is required to allocate the release optimally among various purposes. The complexity of the problem of reservoir operation depends on the extent to which the various intended purposes are compatible. If the purposes are compatible, comparatively less effort is needed for coordination.

Conflicts in a Multipurpose Reservoir

While operating a multipurpose reservoir, a number of conflicts can arise in satisfying various purposes. These are as follows:

Conflicts in Reservoir Space. These conflicts occur when a reservoir (of limited storage) is required to satisfy incompatible purposes, for example, water conservation and flood control. The satisfaction of conservation purposes requires the reservoir to be filled to the maximum extent, whereas the objective of flood control is best met when the reservoir has sufficient vacant space. The critical decision in regulation is whether to fill the reservoir or keep it vacant.

Conflicts among Purposes. Conflicts can also arise among conservation purposes when the pattern of water use is different and the requirement of storage space for one purpose is not in conformity with the other. The water required for uses such as irrigation, municipal, and industrial water is consumed and cannot be shared with any other use.

Conflicts within the Same Purpose. A deficit of water in a reservoir can be distributed over time in different ways. A typical decision is whether to cut the supply now so that there is a small deficit for a longer period or to postpone the cut for the future and risk a bigger shortage for a shorter time. The impact of these two decisions will be different.

Simple Approaches to Reservoir Operation

A reservoir is operated according to a set of rules to store and release water, depending on the purposes it is required to serve. The release decisions are made in accordance with the available water, current and forecasted inflows, demands, weather outlook, and time of the year. The drawdown–refill cycle of a reservoir is usually 12 months long, except when the reservoir capacity is large in relation to stream flows. In many regions of the world, the filling and drawdown periods are distinctly separated.

Standard Linear Operating Policy

The simplest of reservoir operating policies is the standard linear operating policy (SLOP), graphically represented in Fig. 2. According to this policy, if the amount of water available in a particular period, is less than the target demand (T), all available water is released. If the available water is more than the target demand but less than target demand plus available storage capacity, a release equal to the target demand is made, and the excess water is stored in the reservoir. If there is no space to store the excess water even after meeting the target demands, all the water in excess of the maximum storage capacity is released.

Note that SLOP is a one-time operating policy; future implications of the current decision are ignored. This type of time-isolated release of water is neither beneficial nor desirable. Though this policy is not used in day-to-day operation, it is frequently used in planning studies.

Rule Curves

A rule curve or level specifies the desired storage to be maintained in a reservoir during different times of

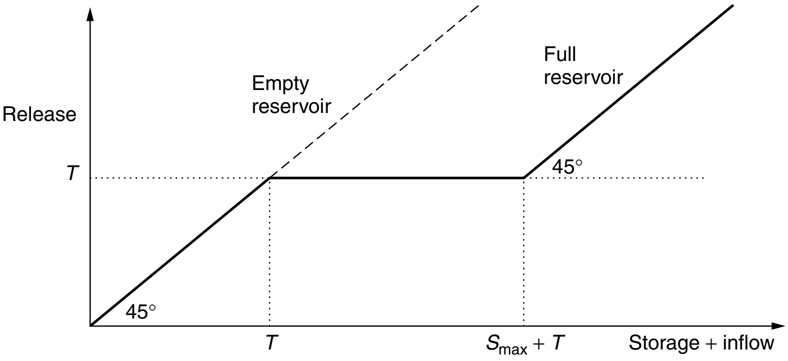

Figure 2. Graphical representation of standard linear operating policy.

the year while trying to meet various demands. The implicit assumption is that a reservoir can best satisfy its purposes if the storage levels specified by the rule curve are maintained at different times. Different rule curves may be developed for different purposes, such as water supply for domestic and industrial use, irrigation, hydropower generation, and flood control. The rule curve as such does not give the releases from the reservoir; these depend on actual inflows and demands.

The derivation of rule curves depends on the type and purposes of the reservoir. The data from the critical period are used for this purpose. Clearly, if the reservoir can meet the demands during the critical year, it will be able to serve its purpose for all other years. These days, many software packages are available for preparing rule curves.

While operating a reservoir following rule curves, if the reservoir level at any time is above the rule curve, releases are made to meet all conservation demands. If the reservoir water level is in the vicinity of the rule curve level, the release is restricted so that the reservoir level does not fall appreciably below the rule curve. If for some reason, the reservoir level drops below the rule curve, the release should be curtailed with the objective of returning to the rule curve at the earliest.

Rule curves implicitly reflect the established trade-off among various project objectives in the long run. These curves specify only the ideal levels to be maintained; the operators can use their experience and judgment to distribute the excess or deficit to maximize benefits. The operation of a reservoir may become quite rigid by strictly following the rule curves. To provide flexibility, different rule curves may be specified for different circumstances, viz., normal, dry, and wet. An operating schedule that gives some leverage to the operator to use his experience and hydrologic forecasts is termed a flexible schedule. Of course, this requires using detailed input data, good models, trained manpower, and careful planning. But at the same time, the benefits can be very large.

Concept of Storage Zoning

The entire reservoir storage space can be conceptually divided into a number of zones by drawing imaginary horizontal planes at various elevations. The sizes of these zones can vary with time as shown in Fig. 3.

Reservoir operators are expected to maintain the reservoir level within the specified zones. This conceptual division of a reservoir into a number of zones and the rules governing the maintenance of reservoir levels in a specified zone are based on the assumption that, at a specified time, an ideal storage zone exists for the reservoir and benefits can be maximized by keeping the storage in this zone. This concept is in some way akin to a rule curve and has the added advantage that it gives more flexibility to the decision-maker.

Operation of a Multireservoir System

The benefits from the joint operation of a system of reservoirs can be substantially larger than the sum of benefits from the operation of individual reservoirs. A system may consist of reservoirs in series, in parallel, or a combination. Various approaches for developing operating policies for a system of reservoirs have been discussed by Wurbs (5).

While managing multiple reservoirs, all reservoirs are maintained in the same zone at any time, to the

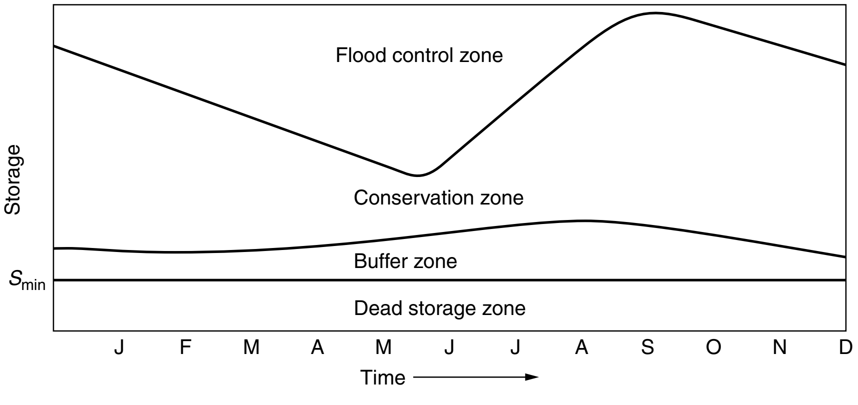

Figure 3. Variation of reservoir zones with time.

extent possible. This is necessary to have balance among reservoirs after an unexpected or extreme hydrologic event. There are three approaches for balancing reservoir contents. The first, known as the “equal function” policy, is to keep all reservoirs at their same zonal position. The second is based on a reservoir ranking or priority concept—the entire zone of the lowest ranking reservoir is used fully before drawing water from the next lowest ranking reservoir, and so on. The third concept is based on a “storage lag” policy. Withdrawals from the zones of some reservoirs are made before withdrawals are begun from the same zones of other reservoirs. After a certain volume has been released from the initial group of reservoirs, releases are made from all reservoirs, maintaining the percentage difference of the available zone volumes.

Conditional rules are also used to regulate multiple-reservoir systems. These policies define reservoir releases as a function of the existing storage volumes and the time of the year and also as a function of the expected natural inflows into the reservoirs for some prespecified future time period.

Reservoir Operation for Flood Control

A storage reservoir with gates is one of the most effective means of flood control. The moderation of a flood through storage is achieved by storing part of the flood volume in the rising phase of the hydrograph and releasing it gradually in the receding phase of the flood. The degree of moderation or flood attenuation depends on the empty storage space available in the reservoir when the flood impinges it. The flood control pool must be emptied as quickly as downstream flooding conditions allow; this will reduce the risk of highly damaging future releases, should major floods occur in quick succession.

In multipurpose reservoirs located in the regions where floods can be experienced at any time and flood control is one of the main purposes, space is permanently allocated for flood control at the top of the conservation pool. The amount of space depends upon the magnitude of likely floods. If the floods are experienced only in a particular season, space is allocated only during that season.

Reservoir regulation consists of storing the peak flows over and above the safe (nondamaging) carrying capacity of the channel at the damage point in the reservoir. The reservoir is emptied after the passage of the flood to make space for controlling subsequent floods. In Fig. 4, ABCDE represents the inflow hydrograph. The line ZZ represents the nondamaging carrying capacity of the river channel downstream of the reservoir. From point B to point D, the natural flow in the river exceeds its safe carrying capacity. If there were no reservoir, the flood water would have spilled over the channel banks and caused damage during the time B to D. The moderated release from the reservoir under ideal operation is given by the dotted curve AGDF. As soon as the inflow begins to increase, the release is gradually increased, till some point G where the release equals the safe carrying capacity. While the inflows from point B to D exceed the safe carrying capacity of the downstream channel, the release is maintained within a safe range by storing the volume in the segment BCD in the flood control zone of the reservoir. After point D, the inflows continue to fall rapidly but the release, while still in a safe zone, exceeds the inflow, so that the reservoir is quickly emptied.

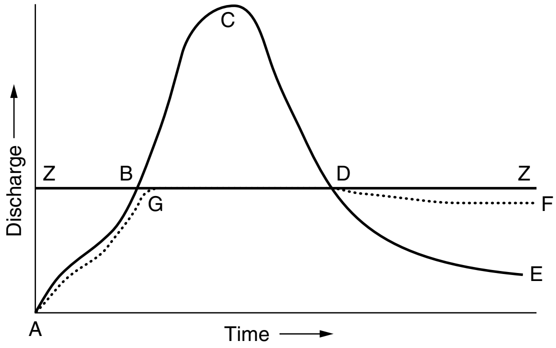

Figure 4. Ideal operation of a reservoir for flood control.

This is an ideal situation which is possible only if the perfect foreknowledge of the hydrograph is available. In actual operation, the release curve will deviate from the ideal shape. For example, if the operator makes smaller releases in the early part of the hydrograph, it is likely that the reservoir will completely fill before point D. In that eventuality, the operator will be forced to make releases at rates exceeding the safe carrying capacity of the downstream channel thereby causing flood damage. Conversely, if at the beginning of a flood, the operator starts making higher releases in the expectation of a major flood and such a flood does not occur, the reservoir may not fill to the desired level by the end of the filling season.

SYSTEMS ENGINEERING FOR RESERVOIR MANAGEMENT

Determination of the reservoir operating policy for efficient management of available water is a complex problem. Many attempts have been made to solve this problem using optimization and simulation models.

Optimization is the science of choosing the best solution from a number of possible alternatives. Optimization methods find a set of decision variables such that the objective function is optimized. Linear programming (LP) and dynamic programming (DP) techniques have been extensively used in water resources. Loucks et al. (6) illustrated applications of LP, nonlinear programming (NLP), and DP to water resources. Yeh (7) presented a state-of-the-art review and discussed in detail the various optimization models for multipurpose, multireservoir operating problems. Simonovic (8) provided a state-of-the-art review and applications of the systems approach to reservoir management and operation. Wurbs (9) reviewed the application of optimization, simulation, and network-flow models.

Simulation is the process of designing a model of a system and conducting experiments with it to understand the behavior of the system and to evaluate various strategies for operating it. The essence of simulation is to reproduce the behavior of the system. It allows controlled experimentation without disturbing the real system. However, simulation analysis does not yield an immediate optimal answer which can be arrived at iteratively. The simulation technique is possibly the most powerful tool for studying complex systems.

The simulation models associated with reservoir operation include the mass-balance computation of reservoir inflows, outflows, and changes in storage. They may also provide an economic evaluation of flood damage, hydropower benefits, irrigation benefits, and other similar characteristics. A large number of reservoir planning and operating studies have used simulation models because such models can provide a more realistic representation of reservoir systems and their operations. They also allow added flexibility to derive responses which cannot be readily defined in economic terms (recreational benefits, preservation of fish and wildlife). The results of simulation models are easier to explain to decision-makers and system operators because the ideas inherent in simulation modeling can be understood easily. Thus, these models are effective tools for a dialogue with operators and decision-makers.

A number of general-purpose computer software programs are available nowadays that can be used for analysis related to planning, design, and operation of multipurpose reservoirs. Some well-known simulation models are the HEC-5 model developed by the Hydrologic Engineering Center, (10), the Acres model (11); the Streamflow Synthesis and Reservoir Regulation (SSARR) Model (12), and the Interactive River System Simulation (IRIS) model (13). Jain and Goel (14) presented a generalized simulation model for conservation operation of a multipurpose reservoir system. Lund and Guzman (15) discussed derivation of operating rules for reservoir systems.

Real-Time Reservoir Operation

A multipurpose reservoir can be efficiently operated if the time interval between the occurrence of an event and the execution of the control adapted for that event is short. In real-time operation, release decisions for a finite future time horizon are made based on the condition of the reservoir at the instant when these decisions are made and the forecast about the likely inflows/demands over this time horizon. After a certain time interval, the new information about the system becomes available, the forecasts are updated, and the decisions are modified.

Real-time operation is especially suitable during floods where the catchment response changes rapidly and decisions have to be made quickly and adapted frequently. Successful application of the real-time operating procedure requires a good telemetry system through which data can be observed on-line. Real-time flood forecasting involves estimating discharge in a river at some period prior to its occurrence. It is very useful in real-time operation because the forecast leadtime proves useful in mitigating some of the adverse effects of flooding. Real-time hydrologic forecasting depends, to a large extent, on timely availability of good quality hydro-meteorological data at a forecasting station.

BIBLIOGRAPHY

1. Morris, G.L. and Fan, J. (1998). *Reservoir Sedimentation Handbook*. McGraw-Hill, New York.
2. ICOLD (1988). *World Register of Large Dams*, International Commission on Large Dams, Paris.
3. Klemes, V. (1981). *Applied Stochastic Theory of Storage in Evolution, in Advances Hydroscience*. V.T. Chow (Ed.). Vol. 12. Academic Press, New York.
4. McMahon, T.A. and Mein, R.G. (1986). *River and Reservoir Yield*. Water Resources Publications, Colorado.
5. Wurbs, R.A. (1996). *Modelling and Analysis of Reservoir System Operations*. Prentice-Hall, New York.
6. Loucks, D.P., Stedinger, J.R., and Haith, D.A. (1981). *Water Resources Systems Planning and Analysis*. Prentice-Hall, Englewood Cliffs, NJ.
7. Yeh, W.W.-G. (1985). Reservoir management and optimization models: a state-of-the-art review. *Water Resour. Res.* **21**(12): 1797–1818.
8. Simonovic, S.P. (1992). Reservoir system analysis: closing gap between theory and practice. *Journal of Water Resources Planning and Management, ASCE* **118**(3): 262–280.
9. Wurbs, R.A. (1993). Reservoir system simulation and optimization models. *J. Water Resour. Plann. Manage, ASCE* **119**(4): 455–472.
10. HEC (1982). *HEC-5 Simulation of Flood Control and Conservation System, Program Users Manual*. The Hydrologic Engineering Center, U.S. Army Corps of Engineers, Davis, USA.
11. Sigvaldason, O.T. (1976). A simulation model for operating a multi-purpose multi-reservoir system. *Water Resour. Res.* **12**(2): 263–278.
12. USACE (1987). *SSARR Users Manual: Streamflow Synthesis and Reservoir Regulation*. US Army Corps of Engineers, Portland, Oregon, USA.
13. Loucks, D.P., Salewicz, K.A., and Taylor, M.R. (1989). *IRIS: An Interactive River System Simulation Model, General Introduction and Description*. International Institute for Applied Systems Analysis, Laxenburg, Austria.
14. Jain, S.K. and Goel, M.K. (1997). *Software for Reservoir Analysis*. Report No. UM-1/96–97, National Institute of Hydrology, Roorkee.
15. Lund, J.R. and Guzman, J. (1999). Derived operating rules for reservoirs in series or in parallel. *J. Water Resour. Plann. Manage., ASCE* **125**(3): 143–153.

DAM REMOVAL AS RIVER RESTORATION

Chih Ted Yang
Colorado State University
Fort Collins, Colorado

River is a dynamic system. It makes spatial and temporal adjustments of its channel geometry, profile, and pattern to maintain a dynamic equilibrium condition compatible with changing independent variables or constraints applied to the river system. The dynamic adjustments of a river are accomplished through the processes of scour, sediment transport, and deposition. Human activities, especially the construction and operation of dams, have profoundly affected most rivers in the world. As a result of these activities, many rivers have lost their ability to maintain or recover their dynamic equilibrium and functions at a self-sustained level. The main purpose of river restoration is to reestablish the basic functions of transporting water and sediment to sustain the ecosystem of rivers and their corridors.

Dams are built and operated for flood control, irrigation, water supply, hydropower generation, navigation, recreation, and environmental purposes. Each dam was designed to last for a given useful lifetime with certain economic and other benefits in mind. A dam can also impose certain risk to downstream residents and properties. Most of the dams, especially small dams, were designed with an estimated useful life of 50 to 100 years. A dam may be removed because of one or more than one of the following factors:

- Economics
- Dam safety and security
- Legal and financial liability
- Ecosystem restoration
- Fish passage improvement
- Site restoration
- Recreation

Sediments are trapped in reservoirs after the completion of dams. The amount of sediments trapped in a reservoir depends on the trap efficiency of the reservoir, which is a function of sediment inflow, ratio between average annual inflow and reservoir volume, reservoir shape, and the dam's operational rules and facilities. A detailed sediment management plan must be developed before a dam's removal to prevent:

- Excessive downstream river channel aggradation, channel widening, bank erosion, increase of flood stage, plug water intake structure, turbidity, and descrupt aquatic habitats
- Adverse impact on water quality
- Excess sediment removal cost and difficulties of finding appropriate sediment storage site
- Head cutting of tributaries

Engineering considerations of dam removal include complete or partial removal; materials used in the dam construction; rate of dam removal and reservoir drawdown; and ability to draw reservoir pool and construction of bypass channel or low-level outlets. Table 1 summarizes some of the factors that need to be considered in dam removal.

Dam removal can be an engineering method for river restoration. Because of the dynamic nature of river adjustments in response to natural and manmade events, comprehensive studies of possible responses of a river must be carried out before a dam's removal. Lane (1) proposed the following qualitative equation to maintain a dynamic equilibrium among sediment load Q_s, sediment size d, water discharge Q, and river slope S.

$$Q_s d \propto QS \tag{1}$$

Table 1. Relationship Between Dam Decommissioning and Sediment Management Alternatives [modified from ASCE (4) by Randle and Greimann (5)]

Sediment Management Alternative	Dam Decommissioning Alternatives		
	Continued Operation	Partial Dam Removal	Full Dam Removal
No Action	• Reservoir sedimentation continues at existing rates, • Inflowing sediment loads are reduced through watershed conservation practices, or • Reservoir operations are modified to reduce sediment trap efficiency.	• Only applicable if most of the dam is left in place. • The reservoir sediment trap efficiency would be reduced. • Some sediment may be eroded from the reservoir.	• Not applicable.
River Erosion	• Sluice gates are installed or modified to flush sediment from the reservoir. • Reservoir drawdown to help flush sediment.	• Partial erosion of sediment from the reservoir into the downstream river channel. • Potential erosion of the remaining sediment by sluicing and reservoir drawdown.	• Erosion of sediment from the reservoir into the downstream river channel. Erosion rates depend on the rate of dam removal and reservoir inflow. The amount of erosion depends on the ratio of reservoir width to river width.
Mechanical Removal	• Sediment removed from shallow depths by dredging or by conventional excavation after reservoir drawdown.	• Sediment removed from shallow depths before reservoir drawdown. • Sediment removed from deeper depths during reservoir drawdown.	• Sediment removed from shallow depths before reservoir drawdown. • Sediment removed from deeper depths during reservoir drawdown.
Stabilization	• The sediments are already stable because of the presence of the dam and reservoir.	• Retain the lower portion of the dam to prevent the release of coarse sediments or retain most of the dam's length across the valley to help stabilize sediments along the reservoir margins. • Construction of a river channel through or around the reservoir sediments.	• Construction of a river channel through or around the existing reservoir sediments. • Relocate a portion of the sediments to areas within the reservoir area that will not be subject to high-velocity river flow.

where Q and Q_s = water and sediment discharge, respectively; d = sediment particle diameter, and S = channel slope.

Yang (2) introduced the following quantitative equation for the prediction of dynamic adjustment of a river based on his unit stream power equations (3,4)

$$C_t = I\left(\frac{VS}{\omega}\right)^J \tag{2}$$

or

$$\frac{Q_t}{Q} = I\left(\frac{QS}{WD\omega}\right)^J \tag{3}$$

where W = channel width, D = channel depth, ω = sediment fall velocity, C_t = total bed-material concentration, Q_t = total bed-material load, and I, J = coefficients.

Because sediment fall velocity is directly proportional to the square root of sediment diameter and J has an average value of 1.0 for rivers, (3) can be simplified to

$$\frac{Q_t d^{0.5}}{K} = \frac{Q^2 S}{WD} = \frac{Q^2 S}{A} \tag{4}$$

where A = channel cross-sectional area, d = median sediment partial diameter, and K = a site-specific parameter.

Equation (4) was used by Yang (7) for the prediction of river morphologic changes because of the construction and operation of a dam. Equation (4) can also be used for the prediction of dynamic adjustments of a river after the removal of a dam.

The rate and timing of a staged reservoir drawdown during the dam removal process should follow the following criteria, especially for the purpose of river restoration (6):

- The drawdown rate is slow enough so downstream flood wave does not cause damages.
- The release of coarse sediment is slow enough not to cause excessive downstream aggradation.
- The concentration of fine sediment is not too great or duration too long to cause an unacceptable impact on water quality and aquatic environment.

During the dam removal process, a comprehensive monitoring program should be carried out to determine

- Reservoir sediment erosion and redistribution
- Hill slope stability along reservoir and downstream river channel adjustment
- Water quality and suspended sediment concentration
- River bed aggradation and flood stage
- Aquifer characteristics
- River channel plan form and channel geometry adjustments
- Effect of large woody debris
- Effects on coastal process

Sediments in a reservoir can be removed by dry excavation, mechanical dredging, hydraulic dredging, sediment conveyance by river flow, and partial sediment removal and reservoir stabilization. Sediment removal by river flow is the least lost and most suitable method for river restoration purposes. This method also requires detailed studies of the hydrologic, hydraulic, sediment transport, scour, deposition, and river morphology. Because of the complexity of these studies, computer models are often used by experienced engineers to simulate and predict the dynamic adjustment of a river after the dam removal.

The U.S. Army Corps of Engineers' HEC-6 (8) and Bureau of Reclamation's GSTARS3 (9) and GSTAR-1D (10) are some of the commonly used computer models in public domain. Bureau of Reclamation's models and user's manuals can be downloaded by users by going to http://www.usbr.gov/pmts/sediment/ and following the links therein. In addition to computer models, physical models are also used for short-term simulation and prediction of erosion, sediment transport, and deposition during the dam removal process.

Dam removal and river restoration often require a multidisciplinary team approach with team members consist of engineers, environmentalists, river morphologists, fishery and aquatic experts, and representatives from local community. Useful references include, but are not limited to, books on *Stream Corridor Restoration: Principles, Process, and Practice* (11), *Erosion and Sedimentation Manual* (12), and *Sediment Transport: Theory and Practice* (7).

BIBLIOGRAPHY

1. Lane. (1955). *Sediment Transport*. McGraw-Hill, New York.
2. Yang, C.T. (1986). *Dynamic adjustment of rivers*. *Proceedings of the 3rd International Symposium on River Sedimentation*, Jackson, MS, pp. 118–132.
3. Yang, C.T. (1973). Incipient motion and sediment transport. *J. Hydraulics Division, ASCE* **99**(HY10): 1679–1704.
4. Yang, C.T. (1979). Unit stream power equations for total load. *J. Hydrol.* **40**: 123–138.
5. American Society of Civil Engineers. (1997). *Guidelines for Retirement of Dams and Hydroelectric Facilities*, New York.
6. Randle, T.J. and Greimann, B. (1992). Chapter 8 Dam Decommissioning and Sediment Management. In: *Erosion and Sedimentation Manual*. U.S. Bureau of Reclamation Technical Service Center, Denver, CO.
7. Yang, C.T. (1996). *Sedimentation Transport: Theory and Practice*. McGraw-Hill, New York (reprint by Krieger Publishing Company, 2003).
8. U.S. Army Corps of Engineers. (1977). *Generalized Computer Program, HEC-6, Scour and Deposition in Rivers and Reservoirs, User's Manual*. The Hydrologic Engineering Center, Davis, CA (revised 1993).
9. Yang, C.T. and Simões, F.J.M. (2002). *User's Manual for Generalized Sediment Transport model for Alluvial River Simulation version 3.0 (GSTARS3)*. U.S. Bureau of Reclamation, Technical Service Center, Denver, CO.
10. Yang, C.T., Huang, J.C., and Greimann, B.P. (2004). *User's Manual for Generalized Sediment Transport for Alluvial Rivers—One Dimension (GSTAR-1D)*. U.S. Bureau of Reclamation, Technical Service Center, Denver, CO.
11. Federal Interagency Stream Restoration Work Group. (1998). *Stream Corridor Restoration—Principles, Processes, and Practice*.
12. Yang, C.T. (2004). *Erosion and Sedimentation Manual*. U.S. Bureau of Reclamation Technical Service Center, Denver, CO.

RIPARIAN SYSTEMS

Theodore A. Endreny
SUNY-ESF
Syracuse, New York

Riparian areas are those along the bank of a watercourse, which is a relatively narrow transition zone between a downslope aquatic waterbody and an upslope terrestrial watershed. Riparian areas are predominantly associated with a river or stream, but also define the bank of a lake, wetland, or other standing water. The riparian system refers to the atmospheric, vegetation, and soil components; their physical, chemical, and biological properties; and their interaction with each other and the upslope terrestrial and downslope aquatic hydrologic cycle. It is important to recognize that riparian areas may border both influent and effluent streams, known as those that lose or gain water to the groundwater.

DEFINITION

Riparian systems may be wetlands, but there are nonwetland riparian areas, and there is no jurisdictional definition of a riparian system as there is for wetlands, defined under Section 404 of the Clean Water Act (33 CFR 328.3(b)). For a discussion of the wetland function of riparian areas, see the entry for Wetlands. As for wetland areas, various federal agencies have attempted to define riparian systems to help in identify them. According to the U.S. Fish and Wildlife Service (FWS), either the presence of a set of obligate water-loving vegetative species (e.g., phreatophytes that have water-table roots) different from that of adjacent areas or more rigorously growing vegetative species similar to upslope areas distinguish the riparian area. The U.S. Department of Agriculture (USDA) considers that riparian areas are further distinguished from the surrounding lands by saturated soils resulting from a shallower depth to the water table, which implies a distinct hydrologic regime.

SYSTEM INPUTS, OUTPUTS, AND TRANSFORMATIONS

A riparian system, in hydroecological terms, can be viewed from its inputs, outputs, and transformations. Inputs arrive via the atmosphere, surface soil, and subsurface soil media, and outputs may travel within these same media after hydrologic and biochemical fluxes between and within the atmospheric, vegetative, and soil components. Atmospheric inputs include short-wave, or solar, and long-wave radiation, wet and dry precipitation, and wind-advected constituents. Surface inputs include upslope storm runoff water, overbank water from a flooded watercourse, eroded soil, as well as sorbed, entrained, and dissolved chemical species. Subsurface inputs include local and regional discharge from upgradient groundwater flow paths, bank storage from the watercourse, and dissolved and entrained chemical species.

Riparian systems are open systems that provide complex and useful chemical sinks, sources, and internal transformations of constituents important in aquatic chemistry and watercourse health. Energy flow into the riparian system, which provides riparian biota growth, reproduction, and maintenance, is derived principally from high-energy sunlight and the use of its short-wave radiation in autotrophic photosynthesis to form organic material from water and carbon dioxide. In the presence of an electron donor, such as oxygen, the organic material is harvested for its energy by heterotrophic organisms in respiration, the reverse process of photosynthesis. Because this energy is associated with the organic material, the food web energy flow has an associated mass cycle. In this cycle, the carbon and associated nutrients (e.g., phosphorous and nitrogen) pass through numerous organisms up the food web, and after excretion or organism death, the carbon and nutrients stop moving up the food web and are either transformed or decomposed back to their inorganic forms as part of mineralization. Microbes, such as bacteria and fungi, and rotifers (e.g., worms) are the key agents in transformation and decomposition.

Riparian input constituents include nutrient species (e.g., nitrogen and phosphorous), chemical pollutants from agrochemicals (e.g., herbicides and pesticides), metals from urban and industrial activities (e.g., lead, zinc, and copper), organic materials from watershed activities (e.g., waste from livestock or domestic pets and vegetative debris), suspended solids (e.g., sand, silt, and clays from eroded soils), pathogens from wildlife (e.g., *Giardia* and *Cryptosporidium*), and in residential areas, thermally heated water and human trash.

In time, the riparian system may accumulate these constituents from surface water that slows and deposits entrained constituents, from soil adsorbing (at the surface) and absorbing (internal to the structure) dissolved or entrained constituents, from vegetation taking the constituents into its root or stem system, or from water stores that fill in the surface or subsurface zone. Biochemical processes operating in the riparian system may transform material inputs that are retained, and another component of inputs that exceed storage may then discharge as outputs with subsequent wind advection, seasonal and storm based surface runoff, seasonal watertable fluctuations, and subsurface flow.

Riparian atmospheric and surface component transformations include mainly photolysis, saprophytic decay, animal digestion, and weathering. Microbial (e.g., fungi and bacteria) subsurface transformations are likely to fluctuate between those of aerobic and anaerobic conditions as the water table fluctuates. Fungi are solely aerobic, but bacteria can operate in both aerobic conditions and anaerobic and have a greater range of transformative options. Two governing principles for microbial transformations of chemicals require that they be energetically favorable and that enzymes are available as catalysts.

The biodegradation potential for a chemical is determined by the free energy change in its reaction with other environmentally available chemicals. This is determined from knowledge of the reduction–oxidation, or redox, potential of the riparian area, which has an ecological sequence from aerobic (O_2) to denitrification (NO_3^-), to manganese reduction (MnO_2), to iron [$Fe(OH)_3$] reduction, to sulfate reduction (SO_4^{2-}), to methanogenesis (CO_2). Given that many enzymes are extremely specific and that mutations occur in bacteria, some individuals in a given strain of bacteria may accomplish transformations not observed in other individuals.

VEGETATED RIPARIAN BUFFERS

Riparian zones supporting forest or grass cover are described structurally as vegetative filter strips (VFS) or functionally as riparian buffer zones (RBZs). Three defined zones for the riparian buffer have been identified by the USDA (1) for agricultural nonpoint source (NPS) runoff and water quality management. The most upslope area is zone 3 and as proposed, contains grasses that slow NPS runoff velocities and allow settling out of entrained constituents such as sediment and sorbed pollutants. The middle area, zone 2, dedicated to managing forestland that captures overland flow leaving zone 1, provides a root-based carbon substrate for microbes (fungi and bacteria) transforming nutrients in subsurface flow, as well as income for the landowner. The near-stream area, is zone 1, which is an undisturbed forest, provides function and habitat for the aquatic system. There are no fixed recommendations for the width of these zones, and many landowners are cautious about dedicating otherwise revenue generating land for forested buffers. In some states, recommended widths for the entire buffer have averaged between 10 and 30 m.

In zone 1, direct biotic and abiotic controls of the vegetative buffer on the aquatic system are many. The vegetative canopy provides cover from incoming solar radiation and blocks outgoing evening long-wave radiation for stable temperature control, as well as protection from predators. The vegetative zone also drops large wood debris into the watercourse for in-stream habitat; creates turbulent eddies (e.g., hydraulic circular flows) and conditions favorable for mobile species seeking to feed or rest by creating areas of refuge from strong downstream currents; and provides substrate for aquatic plants and sessile organisms. Further, the vegetative root wads and downed stems along the bank provide soil stability and erosion control, and the small woody and leaf debris enter the aquatic food web and feed shredders, nymphs; algae, and trout.

Zone 3 and 2 overland flow interception, slowing, settling, and filtering of runoff-entrained constituents requires that the runoff enter in a thin, dispersed sheet and that concentrated channels of flow be removed and flow respread. Methods for spreading the flow include constructed and natural methods (e.g., limbs from a tree). In addition to surface channels of short-circuited flow discharging directly to the watercourse and thereby avoiding riparian storage and transformation, flow can also preferentially short-circuit in the subsurface soil. Such short-circuiting might include preferential flow through macropores and agricultural tile drains that pass too quickly for effective microbial transformation, as well as deep groundwater flows that upwell into the stream at its base and bypass the microbial populations (see Fig. 1). Additional management interventions include removing accumulated sediment berms from the upslope edge of zone 3, as needed, and replacing vegetative materials that are not fixing or assimilating nutrients at target efficiency.

Vegetative riparian structure and function for the adjacent watercourse, it has been widely documented, provide critical resources for aquatic ecosystem health. Inside the United States, a great deal of this work was performed in the Coastal and Piedmont Regions in Maryland, Virginia, and North Carolina; studies in southern Finland and Sweden are also widely cited (2). Most studies have focused on nutrient transformations of nitrogen and phosphorous, and results indicate that nutrient retention varies as a function of nutrient load, riparian upslope steepness, and riparian buffer width, as well as vegetative types, health, subsurface flow paths, and microtopography (e.g., surface short-circuiting). Total nitrogen reduction estimates given for subsurface inputs are between 67 and 89% (3). Phosphorous reduction in U.S. riparian buffers ranges from 41 to 93% (4), and removal efficiency, it has been observed, decreases with time.

RIPARIAN BUFFER BEST MANAGEMENT PRACTICES

In U.S. Environmental Protection Agency (USEPA) research in the mid-Atlantic highland areas of Pennsylvania, Maryland, and Virginia, random sampling networks of benthic macroinvertebrate health, recorded as a biotic index (BI), trend strongly with riparian habitat health. For watersheds with land use stress, such as agricultural or residential occupancy, stream reaches that have low BI values (e.g., impaired) were also areas where riparian habitat was typically absent or degraded, whereas reaches that have high BI values had attendant healthy riparian habitat. Recommendations from these studies include using forested riparian areas as watershed restoration and best management practices (BMPs) to reduce upslope pollutant loads and improve aquatic health.

Hydrologic investigations in urban environments, such as the National Science Foundation (NSF) funded Long Term Ecological Research (LTER) research site in Baltimore, Maryland, have found differences between urban and rural riparian systems. Although the riparian

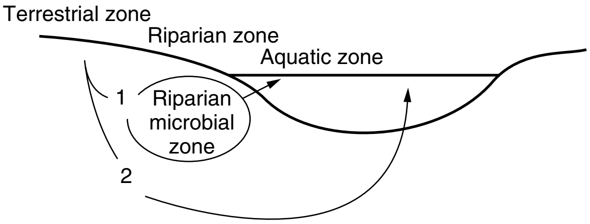

Figure 1. Flow paths from terrestrial to the aquatic zone showing (1) passage through the riparian microbial zone and (2) bypass beneath the riparian microbial zone where nutrient transformations occur.

system in both urban and rural cases is the transition between terrestrial and aquatic processes, the gradual blending of these zones found in rural areas is more abrupt in urban areas. In fact, in areas where streams have been channelized for flood control, leveed from their natural floodplain, and upslope ground water recharge has been reduced by impervious cover, urban riparian systems are drier than their counterpart rural systems. In such cases, vegetative species that are obligatory water loving will not survive in the urban riparian system that has greater depth to the water table and more pronounced mesic periods. Instead, the more common species found in urban riparian systems are obligatory highland, or dry tolerant, trees and shrubs.

For new storm water management under the Clean Water Act, known as the Municipal Separate Storm Sewer System (MS4), the USEPA has required permitting for most municipalities to document pollutant discharge reductions. These reductions may be obtained by BMPs and for retrofitting to existing urban structure, will require smaller footprint design and engineering than a traditional 10 to 30 m wide riparian system. One BMP design that borrows water treatment functional components from riparian buffers is referred to as a vegetated bioretention device, or rain garden. In a USEPA Factsheet, this device is suggested for redirecting runoff from parking lots, driveways, and urban streets, when collocated with storm drain inlets, vertically through the BMP and into the groundwater or a subsurface drainpipe. The mixture of vegetation, soils, and microbes will transform and reduce total Kjeldahl nitrogen (TKN, e.g., ammonia and organic nitrogen) by 68 to 80%; total phosphorous by 70 to 83%; lead, copper, and zinc metals by 93 to 98%; and total suspended solids, organics, and bacteria by 90% (5). Given that this system is intended for urban areas where water tables are, deeper plant species should be drought tolerant, and in colder regions, salt tolerant in areas where roads are deiced.

BIBLIOGRAPHY

1. Welsch, D. (1991). *Riparian Forest Buffers—Function and Design for Protection and Enhancement of Water Resources USDA-FS Pub. No. NA-PR-07-91*. U.S. Department of Agriculture, Radnor, PA.
2. Haycock, N., Burt, T., Goulding, K., and Pinay, G. (Eds.). (1997). *Buffer Zones: Their Processes and Potential in Water Protection*. Quest Environmental, Hertfordshire, UK, pp. 1–326.
3. Correll, D.L. (1997). Buffer zones and water quality protection: general principles. In: *Buffer Zones: Their Processes and Potential in Water Protection*. N. Haycock et al. (Eds.). Quest Environmental, Hertfordshire, UK, pp. 7–20.
4. Uusi-Kamppa, J., Turtola, E., Hartikainen, H., and Ylaranta, T. (1997). The interactions of buffer zones and phosphorous runoff. In: *Buffer Zones: Their Processes and Potential in Water Protection*. N. Haycock et al. (Eds.). Quest Environmental, Hertfordshire, UK, pp. 43–53.
5. USEPA. (1999). *Storm Water Technology Fact Sheet: Bioretention*. EPA-832-F-99-012. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

RIVERS

Pierre Y. Julien
Chih Ted Yang
Colorado State University
Fort Collins, Colorado

Rivers are natural streams typically larger than a creek. The distinction may be somewhat subjective, but creeks cannot be traveled by boat and rivers can. Rivers are the result of precipitation and runoff over land surfaces, followed by surface erosion and sediment transport. As water runs off to lower elevations, the erosion of the underlying earth and soil carves channels into the landscape. These natural channels convey the floodwaters from rainfall and snowmelt and also drain the aquifers during periods of low flow. Rivers are perennial when there is flowing water all year round. Rivers that dry out during part of the year are ephemeral streams, also called arroyos. Related nouns include riverbank or riverside to describe the banks of a river, as well as riverbed to designate the stream bottom. Other adjectives related to rivers include riverine, riparian, fluvial, and fluviatile.

The discharge, or rate of outflow, of a river depends on the channel width, flow depth, and flow velocity. The flow velocity depends on the flow depth, or hydraulic radius, the slope of the river, and the roughness of the riverbed. The volume of water that flows through a river is affected by the duration and intensity of rainfall precipitation or snowmelt in the drainage basin of the river. Over geologic times, it is possible for rivers to be deflected into a different drainage system. This phenomenon is called a capture and results in an increase in drainage area.

Traditionally river systems have been classified as young, mature, or old depending on their stage of development over geologic times. A young river typically flows in a steep valley and has a steep bed slope and numerous riverbed irregularities. Cliffs or escarpments and differences in the resistance of rocks create irregularities in a riverbed and can thus cause rapids and waterfalls. Young systems are characterized by discontinuous longitudinal profiles, lakes, and variable hydraulic geometry. At a mature stage, the river valley has a wide floor and flaring sides and more pronounced erosion by its tributaries. An old river system tends to have regular longitudinal profiles and a gradual decrease in riverbed slope in the downstream direction. Geologic discontinuities such as faults and active tectonics can cause sudden changes in bed elevation that effectively can rejuvenate old fluvial systems. The age classification of rivers is diminishing in popularity now that quantitative studies of river behavior are more common.

The topography of natural rivers depends on the balance between the quantity of sediment in the stream and its transport capacity. A river is alluvial when it can transport its own riverbed sediment and can form its own channel geometry. The weathering process of the land of a given watershed produces large masses of alluvium that can be carried by rivers over geologic periods. An alluvial river tends to eliminate irregularities and form a smooth

Table 1. Water and Sediment Loads of Selected Rivers[a]

		Discharge			
	Catchment Area	Water		Sediment	
River	10^6 km^2	m^3/s	mm/yr.	10^6 ton/yr	mm/yr
Amazon	7.0	100,000	450	900	90
Mississippi	3.9	18,000	150	300	55
Congo	3.7	44,000	370	70	15
La Plata/Parana	3.0	19,000	200	90	20
Ob	3.0	12,000	130	16	4
Nile	2.9	3,000	30	80	15
Yenissei	2.6	17,000	210	11	3
Lena	2.4	16,000	210	12	4
Amur	2.1	11,000	160	52	15
Yangtse Kiang	1.8	22,000	390	500	200
Wolga	1.5	8,400	180	25	10
Missouri	1.4	2,000	50	200	100
Zambesi	1.3	16,000	390	100	50
St. Lawrence	1.3	14,000	340	10	6
Niger	1.1	5,700	160	40	25
Murray-Darling	1.1	400	10	30	20
Ganges	1.0	14,000	440	1,500	1,000
Indus	0.96	6,400	210	400	300
Orinoco	0.95	25,000	830	90	65
Orange River	0.83	2,900	110	150	130
Danube	0.82	6,400	250	67	60
Mekong	0.80	15,000	590	80	70
Hwang Ho	0.77	4,000	160	1,900	1,750
Brahmaputra	0.64	19,000	940	730	800
Dnjepr	0.46	1,600	110	1.2	2
Irrawaddi	0.41	13,000	1,000	300	500
Rhine	0.36	2,200	190	0.72	1
Magdelena (Colombia)	0.28	7,000	790	220	550
Vistula (Poland)	0.19	1,000	160	1.5	5
Kura (USSR)	0.18	580	100	37	150
Chao Phya (Thailand)	0.16	960	190	11	50
Oder (Germany/Poland)	0.11	530	150	0.13	1
Rhone (France)	0.096	1,700	560	10	75
Po (Italy)	0.070	1,500	670	15	150
Ishikari (Japan)	0.016	230	450	6	270
Tiber (Italy)	0.013	420	1,000	1.8	100
Tone (Japan)	0.012	480	1,250	3	180
Waipapa (New Zealand)	0.0016	46	900	11	5,000

[a] After Reference 1.

gradient from its source to its base level. The lowest level to which a river flows is called the base level. Sea level is the ultimate base level, but a lake or reservoir may serve as a local and temporary base level.

Slight changes in sediment transport capacity or in sediment supply to a stream can result in a significant change in bed elevation and planform geometry. For instance, if a stream receives more sediment than it can transport the result is aggradation, or a raise in riverbed elevation through time. Conversely, when the sediment transport capacity exceeds the upstream supply, the water will erode the riverbed and riverbanks and degrades the riverbed.

Similar changes normally take place laterally in river bends. Typically, the velocity increases on the concave riverbank and the increased sediment transport capacity results in erosion and lateral migration of a meandering channel. Simultaneously, the decreased flow velocity near the convex bank of the river decreases transport capacity and results in sediment accumulation and aggradation through time to form point bars. As it approaches base level, downward cutting is replaced by lateral cutting, and the river widens its bed and valley and develops a sinuous course that forms exaggerated loops and bends called meanders. A river may open up a new channel across the arc of meander, called a neck cutoff, thereby cutting off a river segment and creating an oxbow lake as shown in Fig. 1.

River velocity is one of the primary factors that control the quantity and size of rock fragments and sediment carried by the river. When rivers reach oceans, lakes, or reservoirs, the decrease in flow velocity results in net deposition of the sediment load. Part of the load carried by the stream is deposited in the riverbed or on the floodplain beyond the channel. The flow of water during floods exceeds the conveyance of the stream channel, and

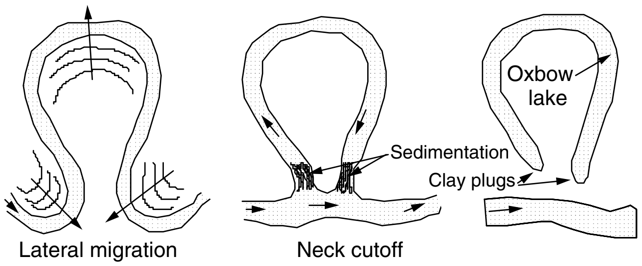

Figure 1. Oxbow lake formation process.

the water level increases to overtop the riverbank and flood adjacent floodplains. As the velocity reduces further on the floodplain, the deposition of sediment forms natural levees. Over long periods of time, landforms produced by deposition include floodplains, deltas, and alluvial fans.

Some typical characteristics including the watershed size, the water and sediment discharges of large rivers are presented in Table 1. The flow and drainage area of the Amazon far exceed that of any other river. On the other hand, the Yellow River (Huang Ho) and the Ganges have tremendous sediment concentrations for very large streams. Several textbooks describe the mechanics of river formation. Some of the recommended texts for future reading and reference in river morphology and river mechanics include Julien (1), Yalin and da Silva (2), Wohl (3), Yang (4) and the recent text of Bridge (5).

BIBLIOGRAPHY

1. Julien, P.Y. (2002). *River Mechanics*. Cambridge University Press, Cambridge, UK.
2. Yalin, M.S. and da Silva, A.M.F. (2001). *Fluvial Processes*. IAHR Monograph, Bern, Switzerland.
3. Wohl, E. (2000). *Mountain Rivers*. Water Resources Monograph, AGU, Washington, DC.
4. Yang, C.T. (1996). *Sediment Transport: Theory and Practice*. McGraw-Hill, New York.
5. Bridge, J.S. (2003). *Rivers and Floodplains, Forms Processes and Sedimentary Record*. Blackwell, New York.

RIVER AND WATER FACTS

National Wild and Scenic Rivers
System - National Park Service

WATER FACTS

- River Lengths
- The National Wild and Scenic Rivers System has only 11, 303 river miles in it—just over one-quarter of one percent of our rivers are protected through this designation, and this protection is often contended.
- Currently, 600,000 miles of our rivers lie behind an estimated 60,000 to 80,000 dams.
- The United States has 3,500,000 miles of rivers. The 600,000 miles of rivers lying behind dams amounts to fully 17% of our river mileage.
- The Missouri River is about 2,540 miles long, making it the longest river in North America. The Nile is the longest river in the world at 4,132 miles as it travels northward from its remote headwaters in Burundi to the Mediterranean Sea.
- The 8 longest rivers in the U.S. are (in descending order) Missouri, Mississippi, Yukon, St. Lawrence (if you count the Great Lakes and its headwaters as one system), Rio Grande, Arkansas, Colorado, Ohio.
- The 8 largest rivers in the U.S., based on volume, are (in descending order) Mississippi, St. Lawrence, Ohio, Columbia, Yukon, Missouri, Tennessee, Mobile.
- Water covers nearly three-fourths of the earth's surface.
- Most of the earth's surface water is permanently frozen or salty.
- Over 90% of the world's supply of fresh water is located in Antarctica.
- The earth's total allotment of water has a volume of about 344 million cubic miles. Of this:
  - 315 million cubic miles (93%) is sea water!
  - 9 million cubic miles (2.5%) is in aquifers deep below the earth's surface.
  - 7 million cubic miles (2%) is frozen in polar ice caps.
  - 53,000 cubic miles of water pass through the planet's lakes and streams.
  - 4,000 cubic miles of water is atmospheric moisture.
  - 3,400 cubic miles of water are locked within the bodies of living things.
- If all the world's water were fit into a gallon jug, the fresh water available for us to use would equal only about one tablespoon.
- It doesn't take much salt to make water "salty." If one-thousandth (or more) of the weight of water is from salt, then the water is "saline."
- Saline water can be desalinated for use as drinking water by going through a process to remove the salt from the water. The process costs so much that it isn't done on a very large scale. The cost of desalting sea water in the U.S. ranges from $1 to $16 per 1000 gallons.
- The overall amount of water on our planet has remained the same for two billion years.
- The United States consumes water at twice the rate of other industrialized nations.
- 1.2 Billion—Number of people worldwide who do not have access to clean water. 6.8 Billion—Gallons of water Americans flush down their toilets every day.
- Each day almost 10,000 children under the age of 5 in Third World countries die as a result of illnesses contracted by use of impure water.
- Most of the world's people must walk at least 3 hours to fetch water.

- By 2025, 52 countries—with two-thirds of the world's population—will likely have water shortages.
- The average single-family home uses 80 gallons of water *per person* each day in the winter and 120 gallons in the summer. Showering, bathing and using the toilet account for about two-thirds of the average family's water usage.
- The average person needs 2 quarts of water a day.
- During the 20th century, water use increased at double the rate of population growth; while the global population tripled, water use per capita increased by six times.
- Water use in the United States alone leaped from 330 million gallons per day in 1980 to 408 million gallons per day in 1990, despite a decade of improvements in water-saving technology.
- On a global average, most freshwater withdrawals—69%—are used for agriculture, while industry accounts for 23% and municipal use (drinking water, bathing and cleaning, and watering plants and grass) just 8%.
- Water used around the house for such things as drinking, cooking, bathing, toilet flushing, washing clothes and dishes, watering lawns and gardens, maintaining swimming pools, and washing cars accounts for only 1% of all the water used in the U.S. each year.
- Eighty percent of the fresh water we use in the U.S. is for irrigating crops and generating thermoelectric-power.
- More than 87% of the water consumed in Utah is used for agriculture and irrigation.
- Per capita water use in the western U.S. is much higher than in any other region, because of agricultural needs in this arid region. In 1985, daily per capita consumption in Idaho was 22,200 gallons versus 152 gallons in Rhode Island.
- A corn field of one acre gives off 4,000 gallons of water per day in evaporation.
- It takes about 6 gallons of water to grow a single serving of lettuce. More than 2,600 gallons is required to produce a single serving of steak.
- It takes almost 49 gallons of water to produce just one eight-ounce glass of milk. That includes water consumed by the cow and to grow the food she eats, plus water used to process the milk.
- About 6,800 gallons of water is required to grow a day's food for a family of four.
- The average American consumes 1,500 pounds of food each year; 1,000 gallons of water are required

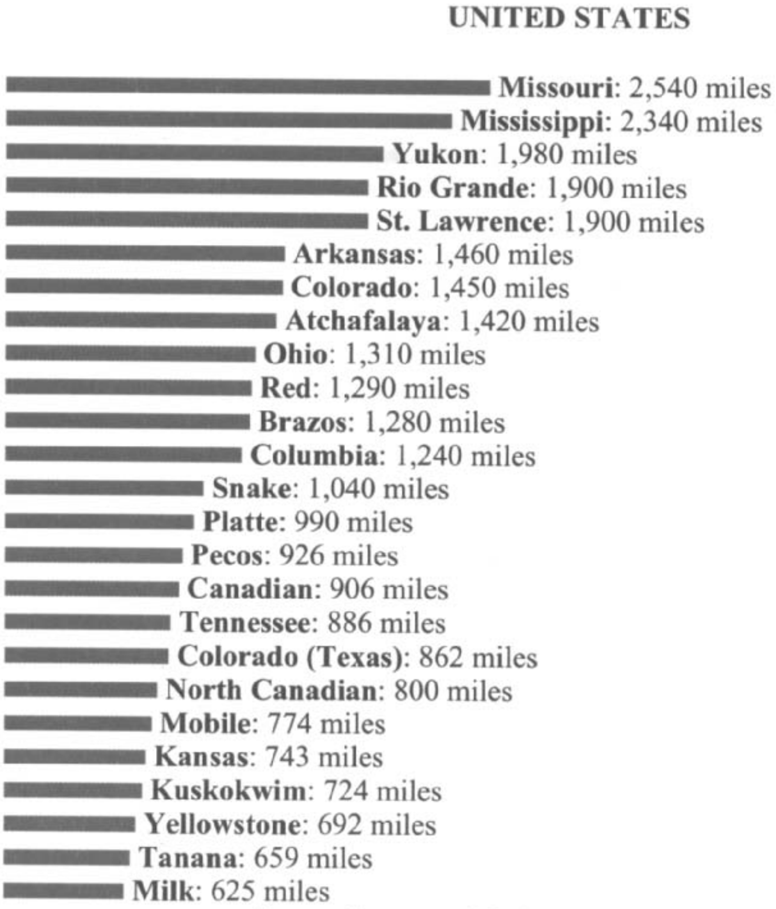

(Source: Kammerer, J.C., Largest Rivers in the United States, U.S. Geological Survey Fact Sheet OFR 87-242 rev. 1990)

to grow and process each pound of that food.—1.5 million gallons of water is invested in the food eaten by just one person! This 200,000-cubic-feet-plus of water-per-person would be enough to cover a football field four feet deep.

- About 39,090 gallons of water is needed to make an automobile, tires included.
- Only 7% of the country's landscape is in a riparian zone—only 2% of which still supports riparian vegetation.
- The U.S. Fish and Wildlife Service estimate that 70% of the riparian habitat nationwide has been lost or altered.
- More than 247 million acres of United States' wetlands have been filled, dredged or channelized—an area greater than the size of California, Nevada and Oregon combined.
- Over 90% of the nearly 900,000 acres of riparian areas on Bureau of Land Management land are in degraded condition due to livestock grazing.
- Riparian areas in the West provide habitat for more species of birds than all other western vegetation combined—80% of neotropical migrant species (mostly songbirds) depend on riparian areas for nesting or migration.
- Fully 80% of all vertebrate wildlife in the Southwest depend on riparian areas for at least half of their life.
- Of the 1200 species listed as threatened or endangered, 50% depend on rivers and streams.
- One fifth of the world's freshwater fish—2,000 of 10,000 species identified—are endangered, vulnerable, or extinct. In North America, the continent most studied, 67% of all mussels, 51% of crayfish, 40% of amphibians, 37% of fish, and 75% of freshwater mollusks are rare, imperiled, or already gone.
- At least 123 freshwater species became extinct during the 20th century. These include 79 invertebrates, 40 fishes, and 4 amphibians. (There may well have been other species that were never identified.)
- Freshwater animals are disappearing five times faster than land animals.
- In the Pacific Northwest, over 100 stocks and subspecies of salmon and trout have gone extinct and another 200 are at risk due to a host of factors, dams and the loss of riparian habitat being prime factors.

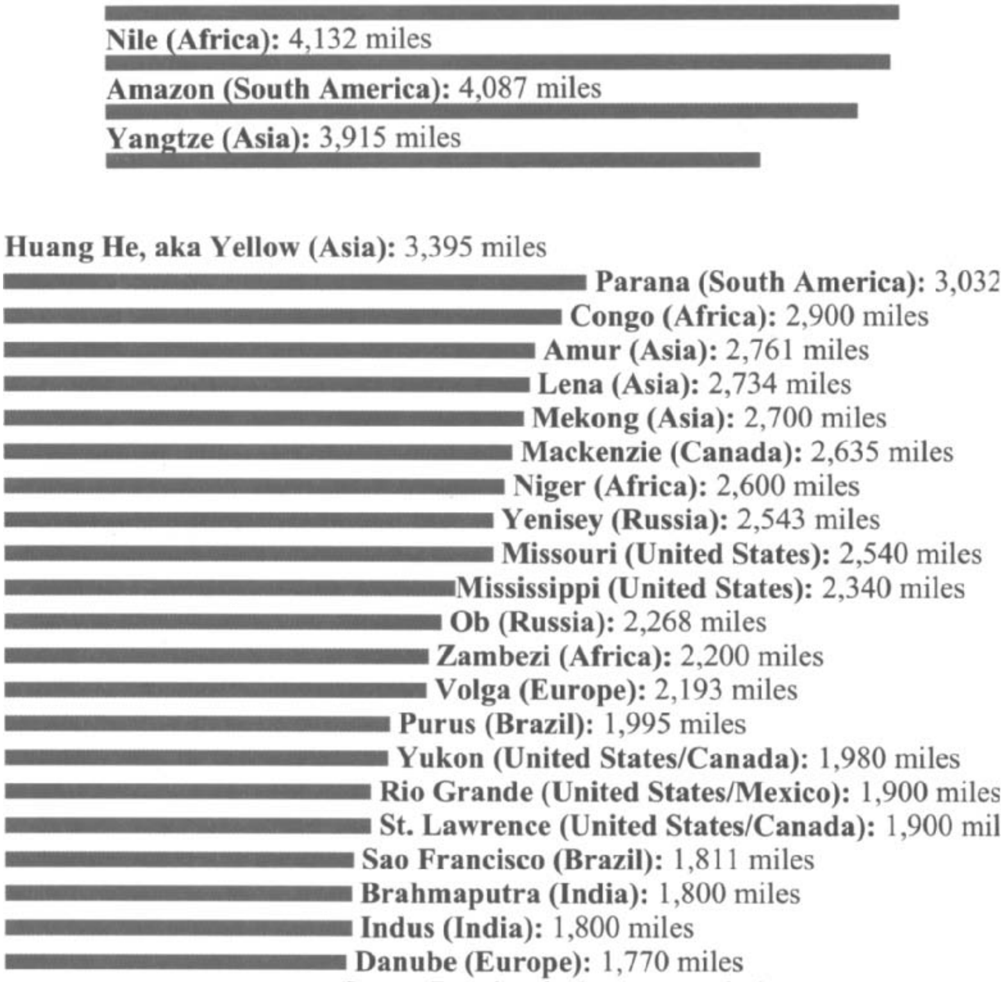

(Source: Encyclopedia Britannica and others)

- A 1982 study showed that areas cleared of riparian vegetation in the Midwest had erosion rates of 15 to 60 tons per year.
- One mature tree in a riparian area can filter as much as 200 pounds of nitrates runoff per year.
- At least 9.6 million households and $390 billion in property lie in flood prone areas in the United States. The rate of urban growth in floodplains is approximately twice that of the rest of the country.
- If all the water in the Great Lakes was spread evenly across the continental U.S., the ground would be covered with almost 10 feet of water.
- One gallon of water weighs 8.34 pounds.

RIVER LENGTHS

It's not so easy to define how long a river is. If a number of tributaries merge to form a larger river, how would you define where the river actually begins? Here, we define river length as the distance to the outflow point from the original headwaters where the name defines the complete length. (Source: Statistical Abstract of the U.S., 1986).

WORLD

Estimates for the length of the world's rivers vary wildly depending on season of the year, who is doing the measuring, the capabilities of the cartographer and his equipment and sources. However, the biggest cause of disagreeing measurements is the inclusion or exclusion of tributaries. For example, many sources lump the Mississippi and Missouri Rivers into one river system, making it one of the longest in the world. The same is true of rivers such as the Ob-Irtysh system in Asia. Considered as a whole, it is one of the ten longest rivers in the world. Removing the Irytish drops the Ob down to 15th position—assuming the rivers ahead of it also weren't measured with massive tributaries included. Here, we have tried to separate the major tributaries. You can easily find other sources that disagree with these numbers; please do not send us further questions on this.

These numbers were taken from the Encyclopedia Britannica and tributaries were separated out with help from sources like Comptons Encyclopedia and others.

SEDIMENT LOAD MEASUREMENTS

Wayne D. Erskine
State Forests of New South Wales
Beecroft, New South Wales, Australia

"No matter how sophisticated, analysis and interpretation can never substitute for well collected data" (1). These words are particularly appropriate for stream sediment loads because most sediment moves during rare, large floods that collectively account for a very small portion of time when access and safety problems are acute (1).

MODES OF SEDIMENT TRANSPORT

Two modes of sediment transport by rivers, suspended load and bed load, are usually recognized on behavioral grounds. *Suspended load* refers to the relatively fine part of the total sediment load that is transported in continuous or intermittent suspension. For sediment to remain in suspension, flow turbulence and velocity must be sufficiently great to counter the tendency of suspended sediment to settle to the riverbed. The two main factors determining sediment settling velocity are grain size and density. *Bed or traction load* refers to the relatively coarse part of the total sediment load that is transported along or close to the riverbed by rolling, sliding, or saltation. The latter term refers to sand movement in a series of hops across a riverbed in a distinct, concentrated layer.

The two modes of sediment transport are not supplementary concepts because a certain grain size can be transported in suspension by a given discharge but as bed load by another smaller discharge. Einstein et al. (2) proposed two alternative modes on the basis of experimental work on the Enoree River in South Carolina, United States. *Wash load* is the finer part of the total sediment load that is not present in bed material in significant quantities and is supplied to and transported through the measurement reach. *Bed-material load* is the coarser part of the total sediment load that is composed of grains found in bed material in appreciable quantities and is transported at a rate dependent on stream discharge. For Einstein et al.'s (2) measurement reach on the Enoree River, 96% of the measured suspended sediment during a flood was finer than 0.351 mm, and not more than 7% of the bed material was finer than 0.351 mm. The rate of transport of sediment coarser than 0.351 mm was a function of discharge, whereas the transport rate for sediment finer than 0.351 mm was related to upstream sediment supply, irrespective of whether it was transported in suspension or by traction. This limiting grain size varies among rivers depending on watershed geology, relief, rainfall, vegetation, land use, channel form, and flood peak discharge and its variability.

Vertical, Lateral and Temporal Distribution of Suspended Sediment at a River Cross Section

If suspended sediment exhibits a uniform vertical and lateral distribution and changes little during short periods of time, then the suspended load can be measured at any point in a river cross section to obtain a mean concentration. Hyperconcentrations of suspended sediment ($400 - 1000$ kg/m^3) in the Yellow River, China, were uniformly distributed in the vertical (3). However, a large body of similar work on other rivers without hyperconcentrated suspended sediment has usually shown that the minimum concentration is found at the water surface and the maximum concentration is found near the riverbed (4,5). Furthermore, the coarsest fractions, which are usually sand, exhibit the greatest variation (5).

Nevertheless, Curtis et al. (4) reported that concentrations of particles as small as 1 or 2 μm (clay) in the Amazon River still showed slight but discernible differences between the riverbed and surface. They concluded that suspended sediment fluxes based only on surface samples are likely to contain significantly large errors.

A number of studies reviewed on the lateral variation of suspended sediment concentration at a cross section found that variations of 20% were common and that some were as large as 70% (5). This is caused by lateral variations in flow velocity and depth (and the consequent development of mobile bedforms) and by incomplete mixing of sediment-laden inflows from small tributaries into large rivers (Fig. 1). The latter effect needs to be evaluated carefully before selecting a field measurement site that should exhibit well-mixed conditions (Fig. 1).

Measured instantaneous fluctuations in depth-integrated (see below) suspended sediment concentration can vary by up to 24% (5,6). The problems of vertical, lateral, and temporal variations in suspended sediment concentration at a cross section are considerable and indicate that suspended sediment samples must be collected during a meaningful period of time from the full flow depth at multiple verticals on a river cross section.

Suspended Load. The suspended sediment load (Q_s in tonnes/day) is calculated from:

$$Q_s = 0.0864QC_s \tag{1}$$

where Q is the discharge (m^3/s) and C_s is the suspended sediment concentration (mg/L).

It is, therefore, necessary to know the discharge at the measurement site and also the mean suspended sediment concentration to calculate reliable suspended loads. For this reason, sediment load measurement sites should be located at well-mixed sections at river gauging stations.

Three requirements for an efficient suspended sediment sampler are

Figure 1. Upstream view of the Sacramento River at Knights Landing, California, showing a muddy plume originating from the tributary in the middle left of the photograph which follows the right bank for some distance before completely mixing with the less turbid mainstream flow (W.D. Erskine photograph).

1. the flow velocity at the sampler intake should be the same as stream flow velocity,
2. the sampler should cause minimum flow disturbance at the sampling point, and
3. the sampler intake should be oriented into the flow in both vertical and horizontal planes (7).

Six main types of suspended sediment samplers have been identified (7–10):

1. instantaneous samplers
2. point-integrating samplers
3. depth-integrating samplers
4. single stage samplers
5. pumping samplers
6. continuous monitoring techniques.

Instantaneous samplers collect a water–sediment mixture at a single time at a single point in a vertical in the stream. Clearly, they cannot account for any temporal, vertical, or lateral variations in concentration and hence are not recommended for use in sediment load programs. However, they are appropriate for other water quality or limnological programs. Point-integrating samplers collect a water–sediment mixture at a predetermined depth in a single vertical of a river cross section during a short time period and, therefore, are time-integrated samplers. A valve is used to open the nozzle, and a separate air exhaust expels the air displaced by the sample. A series of depth settings acceptable for point-integrating samplers was found (11). The U.S. Federal Inter-Agency Sedimentation Project (6,8,10) developed a series of such samplers (USP61, USP 63, USP 72); USP 61 is illustrated in Fig. 2.

Depth-integrating samplers continuously collect a water–sediment mixture at a rate proportional to the flow velocity from the water surface to near the bed as they are lowered and raised at a uniform rate at a series of verticals at a river cross section (6,9,10). The resultant suspended sediment concentration is a time- and velocity-weighted mean. Gregory and Walling (7) recommended that the vertical transit rate (the rate of lowering and raising) should not exceed 40% of the maximum flow velocity in the vertical. The intake nozzle diameter can also be changed to ensure that the sample bottle capacity is not exceeded for deep and/or fast streams. The U.S. Federal Inter-Agency Sedimentation Project (6,8,10) developed a series of such samplers (USD 43, USDH 48, USDH 59, USDH 75, USDH76, USDH 81, USD 49, USD 74, USD 77); USDH 48 and 59 are illustrated in Fig. 2. There is an unsampled zone about 0.1m high immediately above the bed for depth-integrated samplers because of the inclined sample bottle (Fig. 2) to ensure that bed material is not scooped into the nozzle when the sampler reaches the bed. Depth-integrated samplers have been the accepted standard for hand field measurements of suspended load for the last 50 years (6,7,9,10).

Depth-integrated suspended sediment sampling is conducted according to either the equal transit rate (or equal width increment) method or the equal discharge

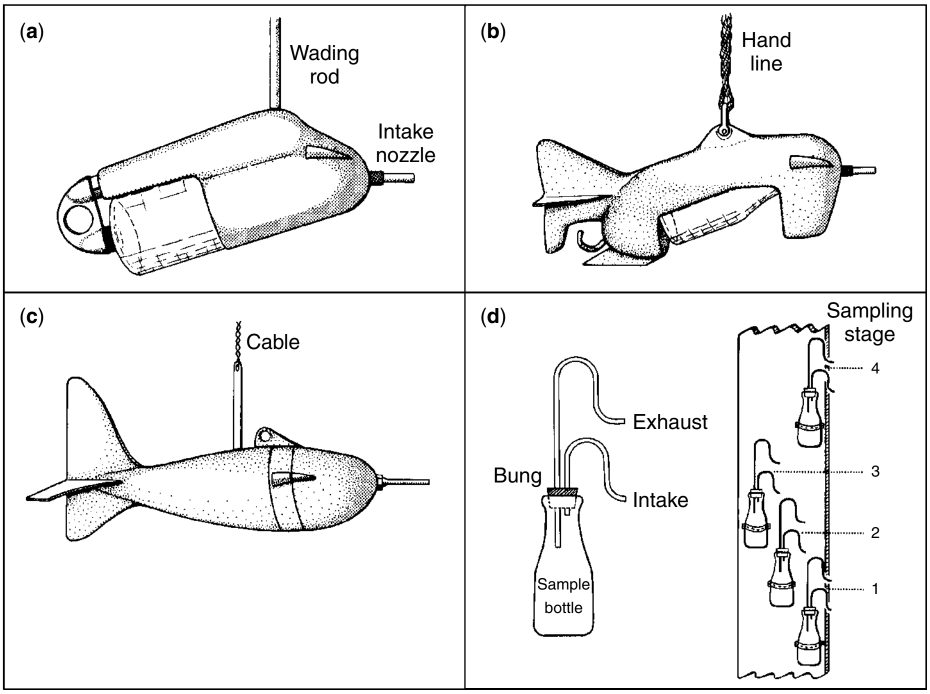

Figure 2. Suspended sediment samplers from Reference 7. (**a**) USDH 48 depth-integrating sampler; (**b**) USDH 59 depth-integrating sampler; (**c**) USP 61 point-integrating sampler which can also be used as a depth-integrating sampler if the valve is left open; and (**d**) single-stage (rising stage) samplers.

increment method (8,9). The equal transit rate method involves moving the sampler at a constant rate between the water surface and the bed at each of a series of verticals equally spaced across the channel. No prior knowledge of the velocity pattern is required at the sampling site to obtain a mean suspended sediment concentration for the whole section. Edwards and Glysson (8) developed guidelines for determining the number of verticals needed to achieve a given relative standard error based on the percentage of sand in the sample and an index of variability computed from the flow velocity and depth. The equal discharge increment method involves dividing the cross section into a series of segments of equal discharge and collecting depth-integrated samples at the centroid of each segment. Clearly, the latter method can be used only when complete velocity–area gauging has also been completed and calculated. This is not always the case.

Single-stage samplers are a type of automatic point-integrating device (10). They consist of a sample bottle equipped with intake and exhaust tubes which fills automatically when the stream first reaches the stage that submerges the sampler (Fig. 2). Such samplers usually do not collect a sample representative of the mean suspended sediment concentration at the cross section and can be contaminated by postinundation flows which can pass through the sampler.

Pump samplers remotely obtain detailed time series variations in suspended sediment concentrations at a fixed point in the stream and are required for accurate load determinations (9,10,12). Sampler intakes should be set either at a fixed position characterized by sufficient flow turbulence to ensure minimum variation in concentration throughout the cross section (13) or on a depth proportional intake boom (12). The latter consists of a hinged boom attached to the riverbed at a pivot and suspended by a float; the intake is set at an appropriate proportion of the flow depth (12). The boom should be at least twice as long as the maximum expected stage, and the intake should be set at an acceptable depth, 0.6 was adopted by Eads and Thomas (12). The boom rises and falls with the flood stage and can be depressed by transported large woody debris to avoid damage because it is oriented downstream. Nevertheless, a single pump sampler intake still equates to a point-integrating approach to sampling (12), and the representativeness of the sample intake point must always be determined.

Pump samplers can be programmed to collect water–sediment samples at specific turbidity thresholds (14) or by various stratified or variable probability plans that reduce the variance in the sediment load. Thomas and Lewis (15) evaluated flow-stratified, time-stratified, and selection-at-list time (SALT) sampling plans and found that

- time-stratified sampling generally gives the smallest variance in storm sediment loads;
- flow-stratified and SALT sampling have difficulties only with small storms because of lack of data; and
- flow-stratified sampling produces the lowest variance for longer, more complex hydrographs that have numerous peaks.

A number of simultaneous pump and depth-integrated samples should be obtained to determine if the pump sample intake position is appropriate to produce suspended sediment concentrations that are equivalent to, or can be converted to equivalence with, depth- and width-integrated samples (12,13). Pumped-sample suspended sediment concentrations obtained from a single, fixed-position intake in a zone of turbulent flow generally varied within 10% of depth-integrated samples without a consistent trend to over- or underestimate concentrations (13). Pump-sample suspended sediment concentrations collected by a constant depth proportion intake were closely related to depth- and width-integrated samples ($r^2 = 0.98$) but pumped concentrations were higher for concentrations above approximately 100 mg/L (12). These results highlight that pump samplers should be used for automatic suspended sediment sampling, providing that the resultant concentrations are equivalent to, or can be converted to equivalence with, depth- and width-integrated samples.

The most common continuous monitoring technique uses turbidity and measures the attenuation or scattering of an incident beam of radiation or light. Turbidity can be measured with an *in situ* probe at frequent time intervals, and the resultant values can often be converted to depth- and width-integrated suspended sediment concentrations (13,14,16–18). Reported r^2 values for relationships between depth-integrated suspended sediment concentration and turbidity can be as high as 0.90 to 0.99 (13,15). However, such relationships are usually site-specific and more reliable where suspended sediment is dominated by silt and clay (14,16).

Bed Load. A range of different portable and permanent bed load samplers has been developed for use in sand- and gravel-bed rivers, but their performance has often been poor (19). Portable samplers include baskets (7,19), trays (7,19), pressure-difference samplers (19–23), acoustic meters (19), and pumps (24). Permanent samplers are associated with major in-stream structures and include vortex tubes (25), pits (19), slot with a conveyor belt emptying into a continuous weighing hopper (26,27), and the Birkbeck bed load sampler (28). The latter consists of slots in the riverbed fitted with boxes resting on water-filled pressure pillows and combined with pressure transducers for stream height measurement (28). Most early samplers were unsatisfactory because of inaccurate measurements, variable and poor sampling efficiency (19), inconsistent sampler placement on the riverbed, and the oscillatory nature of bed load transport (19,29). The slot with a conveyor belt (26,27) is the most innovative and reliable technique devised to date but is also very expensive.

The Helley–Smith pressure-difference bed load sampler was designed and developed by the U.S. Geological Survey, Water Resources Division to be compatible with depth-integrated suspended sediment samplers (22) and is one of the most accurate (20,21,23). For sediments whose diameters are between 0.5 and 16 mm, the sampler has a near-perfect sampling efficiency that is higher for finer sediment and lower for coarser sediment (20,21). Sample bags can be filled to about 40% capacity with sediment larger than the mesh size without reducing hydraulic efficiency (20,21). Larger than standard sampling bags and increased bag mesh size improved sampling efficiencies in streams that have high sand and organic matter fluxes because of reduced bag plugging (23,30,31). Short sampling times also increase sampling efficiency in the same streams (31). Thin-walled samplers (1.5 mm) also measure larger sand fluxes than thick-walled (6.3 mm) samplers due to local hydraulic effects (32). To operate efficiently, the sampler must have a reasonable fit between the sampler base and the riverbed and will not sample satisfactorily when the bed is highly irregular (21). The inlet orifice diameter of the original Helley–Smith sampler (76.2 mm) closely approximates the unsampled zone between the intake nozzle and the bed for a depth-integrated suspended sediment sampler (22). Therefore, depth-integrated suspended load samplers collect intermittently suspended bed material and wash load whereas pressure-difference bed load samplers collect sediment moving by rolling and sliding, bedform migration, and saltation.

There is considerable temporal and spatial variability inherent in bed load transport rates (20,21,29,32,33) which makes field measurements difficult. Bed load fluxes measured at a fixed sampling point during constant discharge on a sand-bed stream that has dune bedforms range from near zero to approximately four times the mean rate, and about 60% are less than the mean (29,32,33). Field measurements indicate that highest transport rates occur in the center of the channel and diminish near the margins (32). The U.S. Geological Survey provisional method for bed load gaugings entails two complete traverses of a cross section and at least 20 equally spaced, point measurements for 30–60s at least 0.5m apart on each pass (21). Although instantaneous bed load fluxes are highly variable, it is usually possible to define a mean flux for a given discharge, provided thorough field measurements are made (21,32). Marked differences in bed load fluxes and their relationship to discharge occur between pools and riffles (21).

BIBLIOGRAPHY

1. Thomas, R.B. (1985). *Measuring Suspended Sediment in Small Mountain Streams*. U.S. Dept. Agr. For. Serv. Pacific Southwest For. Range Expt. Station Gen. Tech. Report **PSW-83**, Berkeley, CA.
2. Einstein, H.A., Anderson, A.G., and Johnson, J.W. (1940). A distinction between bed-load and suspended load in natural streams. *Trans. Am. Geophys. Union* **21**: 628–633.
3. Boliang, Z. and Zhan, N. (1982). Analysis of the vertical distribution of high sediment concentrations in the Yellow River and studies of methods of observation. *Int. Assoc. Hydrol. Sci. Publ. No.* **137**: 421–430.
4. Curtis, W.F., Meade, R.H., Nordin, C.F., Jr., Price, N.B., and Sholkovitz, E.R. (1979). Non-uniform vertical distribution of fine sediment in the Amazon River. *Nature* **280**: 381–383.
5. Task Committee on Preparation of Sedimentation Manual Committee on Sedimentation of the Hydraulics Division. (1969). Sediment measurement techniques: A Fluvial sediment. *J. Hydraul Div. Proc. Am. Soc. Civ. Eng* **95**: 1477–1514.
6. Benedict, P.C. (1948). Determination of the suspended sediment discharge of streams. *Proc. 1st Fed. Inter-Agency Sedimentation Conf.*, Denver, 1947, Bureau of Reclamation, U.S. Dept. of the Interior, Washington, DC, 55–65.
7. Gregory, K.J. and Walling, D.E. (1973). *Drainage Basin Form and Process*. Edward Arnold, London, p. 456.
8. Edwards, T.K. and Glysson, G.D. (1999). *Field methods for measurement of fluvial sediment*. Techniques of water-resource investigations of the United States Geological Survey C2(3) Applications of Hydraulics, Washington, DC.

9. Guy, H.P. and Norman, V.W. (1970). *Field methods for measurement of fluvial sediment*. Techniques of water-resource investigations of the United States Geological Survey C2(3), Washington.
10. Witzigman, F.S. (1965). A summary of the work of the Inter-Agency Sedimentation Project. *Proc. 2nd Fed. Inter-Agency Sedimentation Confer.*, 1963, U.S. Dept of Agric. Misc. Publ. **970**: 166–177.
11. *Federal Inter-Agency Sedimentation Project, Measurement of Sediment Discharge of Streams*. (1948). Federal Interagency Sedimentation Project Report No. 8, Minneapolis, MN.
12. Eads, R.E. and Thomas, R.B. (1983). Evaluation of a depth proportional intake device for automatic pumping samplers. *Water Res. Bull.* **19**: 289–292.
13. Walling, D.E. and Teed, A. (1971). A simple pumping sampler for research into suspended sediment transport in small catchments. *J. Hydrol.* **13**: 325–337.
14. Lewis, J. (1996). Turbidity-controlled suspended sediment sampling for runoff-event load estimation. *Water Resour. Res.* **32**: 2299–2310.
15. Thomas, R.B. and Lewis, J. (1995). An evaluation of flow-stratified sampling for estimating suspended sediment loads. *J. Hydrol.* **170**: 27–45.
16. Kunkle, S.H. and Comer, G.H. (1971). Estimating suspended sediment concentrations in streams by turbidity measurements. *J. Soil Water Conserv.* **26**: 18–20.
17. Walling, D.E. (1977). Assessing the accuracy of suspended sediment rating curves for a small basin. *Water Resour. Res.* **13**: 531–538.
18. Walling, D.E. (1977). Limitations of the rating curve technique for estimating suspended sediment loads, with particular reference to British rivers. *Int. Assoc. Hydrol. Sci. Publ. No.* **122**: 34–48.
19. Hubbell, D.W. (1964). *Apparatus and techniques for measuring bedload*, U.S. Geol. Surv. Water Supply Paper 1748, Washington, DC.
20. Emmett, W.W. (1980). *A field calibration of the sediment trapping characteristics of the Helley-Smith bedload sampler*. U.S. Geol. Surv. Prof. Paper **1139**, Washington, DC.
21. Emmett, W.W. (1981). Measurement of bed load in rivers. *Int. Assoc. Hydrol. Sci. Publ. No.* **133**: 3–15.
22. Helley, E.J. and Smith, W. (1971). *Development and calibration of a pressure difference bedload sampler*. U.S. Geol. Surv. Open File Report.
23. Johnson, C.W., Engleman, R.L., Smith, J.P., and Hanson, C.L. (1977). Helley-Smith bed load samplers. *J. Hydraul Div. Proc. Am. Soc. Civ. Eng.* **103**: 1217–1221.
24. Langley, H.I. (1976). *Bedload sampling by pumping. Symposium on Inland Waterways for Navigation, Flood Control and Water Diversions*. Vol. II, Colorado State University, Fort Collins, 1976, Am. Soc. Civ. Eng., N.Y.: 945–959.
25. Hayward, J.A. and Sutherland, A.J. (1974). The Torlesse stream vortex-tube sediment trap. *J. Hydrol. (N.Z.)* **13**: 41–53.
26. Leopold, L.B. and Emmett, W.W. (1976). Bedload measurements, East Fork River, Wyoming. *Proc. Nat. Acad. Sci. USA* **73**: 1000–1004.
27. Leopold, L.B. and Emmett, W.W. (1977). 1976 bedload measurements, East Fork River, Wyoming. *Proc. Nat. Acad. Sci. USA* **74**: 2644–2648.
28. Reid, I., Layman, J.T., and Frostick, L.E. (1980). The continuous measurement of bedload discharge. *J. Hydrol. Res.* **18**: 243–249.
29. Carey, W.P. (1985). Variability in measured bedload-transport rates. *Water Res. Bull.* **21**: 39–48.
30. Beschta, R.L. (1981). Increased bag size improves Helley-Smith bed load sampler for use in streams with high sand and organic matter transport. *Int. Assoc. Hydrol. Sci. Publ. No.* **133**: 17–25.
31. O'Leary, S.J. and Beschta, R.L. (1981). Bedload transport in an Oregon coast range stream. *Water Res. Bull.* **17**: 886–894.
32. Pitlick, J. (1988). Variability of bed load measurement. *Water Resour. Res.* **24**: 173–177.
33. Hubbell, D.W., Stevens, H.H., Jr, Skinner, J.V., and Beverage, J.P. (1985). New approach to calibrating bed load samplers. *J. Hydraul. Eng.* **111**: 677–694.

SEDIMENTATION

CHIH TED YANG
PIERRE Y. JULIEN
Colorado State University
Fort Collins, Colorado

Sedimentation is a science of interest to geologists, river morphologists, environmentalists, and engineers. Geologists are interested in the formation, mineral content, and age of sedimentary rocks. River morphologists are interested in the formation of rivers, their patterns, and profiles. Environmentalists are interested in the impacts of sediments on water quality, aquatic habitat, and ecosystems. Engineers are interested in the movement of sediment and its impact on the engineering design and operation of hydraulic structures. Due to these differences of interests and emphases, sedimentation studies can vary from qualitative to quantitative or a combination of both.

This article emphasizes the hydraulics of sediment transport. Due to the complexity of sediment transport, only basic assumptions and approaches used by engineers are summarized here. More detailed descriptions and analyses of sediment transport, scour, and deposition can be found in text books such as those by Julien (1), Yang (2), Simons and Sentürk (3), and Yalin (4).

Sediment transport can be classified as bed load, suspended load, and total load. Two basic approaches are used to study sediment transport, the probabilistic and deterministic approaches. Einstein (5) developed the most well-known probabilistic theory for sediment transport. Engineering applications of Einstein's transport functions are limited due to laborious computation procedures and extensive field data requirements. However, the modified Einstein method (6) has been used by engineers to estimate total bed-material load where direct bed load measurements are difficult to obtain.

The deterministic approach assumes that there is a one-to-one functional relationship between the dependent variable of sediment transport rate, or concentration, and independent variables. Commonly used independent variables are water discharge, average flow velocity, water surface or energy slope, shear stress, stream power, and unit stream power.

Figure 1 shows that unit stream power has the strongest one-to-one correlation with sediment discharge.

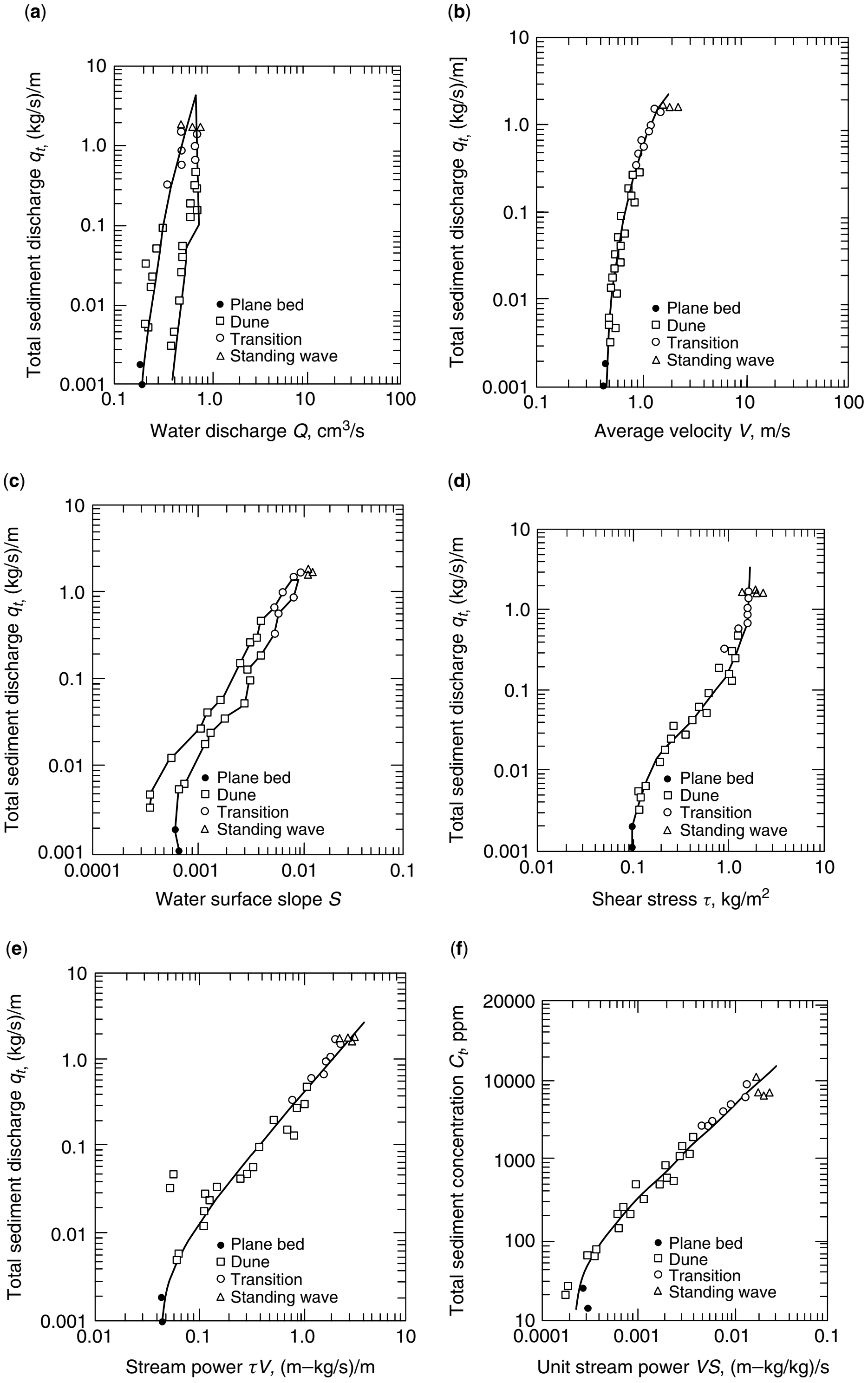

Figure 1. Relationships between total sediment discharge and (**a**) water discharge, (**b**) velocity, (**c**) slope, (**d**) shear stress, (**e**) stream power, and (**f**) unit stream power, for 0.93-mm sand in an 8-ft wide flume (7).

Yang's 1973 (8) dimensionless unit stream equation is

$$\log C_{\mathrm{ts}} = 5.435 - 0.286 \log \frac{\omega d}{\upsilon} - 0.457 \log \frac{U_*}{\omega} + \left(1.799 - 0.409 \log \frac{\omega d}{\upsilon} - 0.314 \log \frac{U_*}{\omega}\right) \log\left(\frac{VS}{\omega} - \frac{V_{\mathrm{cr}} S}{\omega}\right) \tag{1}$$

where C_{ts} = total sediment concentration in ppm by weight, VS = unit stream power, U_* = shear velocity, υ = kinematic viscosity of water, ω = fall velocity of sediment, d = median sediment particle diameter, V = average flow velocity, S = slope, and V_{cr} = critical flow velocity at incipient motion.

The independent variables in Eq. 1 can be measured directly in the field except the fall velocity of sediment.

The fall velocity can be obtained from figures based on experimental results recommended by the U.S. Inter-Agency Committee on Water Resources, Subcommittee on Sedimentation (9). The fall velocity can also be computed by Rubey's 1933 formula (10).

The sediment concentration or sediment load computed from different formulas can vary drastically from each other and from measurements. Figure 2 shows comparison among the computed total sediment discharges using different formulas and the measured results from the Niobrara River near Cody, Nebraska. Table 1 summarizes the ratings of some of the commonly used sediment transport formulas by the American Society of Civil Engineers Task Committee for users to consider.

Riverbed sediment particles are not uniform in size. Computations of sediment transport by size fraction are necessary to obtain realistic results for engineering purposes. Computer models are often used to simulate and predict the process of scour, transport, and deposition in rivers and reservoirs. The U.S. Army Corps of Engineers' HEC-6 (13) model and the Bureau of Reclamation's GSTARS 2.1 (14) and GSTARS3 (15) are some of the commonly used models in the public domain for solving river engineering and river morphology problems where the movement of sediment, scour, and deposition are subjects of concern. The Bureau of Reclamation's models and user's manuals can be obtained by going to http://www.usbr.gov/pmts/sediment/ and following the links therein.

Table 1. Summary of Rating of Selected Sediment Transport Formulas[a]

Formula Number	Reference	Type	Comments
1	Ackers and White	Total load	Rank = 3[b]
2	Engelund and Hansen	Total load	Rank = 4
3	Laursen	Total load	Rank = 2
4	MPME[c]	Total load	Rank = 6
5	Yang	Total load	Rank = 1, best overall predictions
6	Bagnold	Bed load	Rank = 5
7	Meyer-Peter and Müller	Bed load	Rank = 7
8	Yalin	Bed load	Rank = 8

[a] Reference 12.

[b] Based on a mean discrepancy ratio (calculated over observed transport rate) from 40 tests using field data and 165 tests using flume data.

[c] Meyer-Peter and Müller's formula for bed-load and Einstein's formula for suspended load.

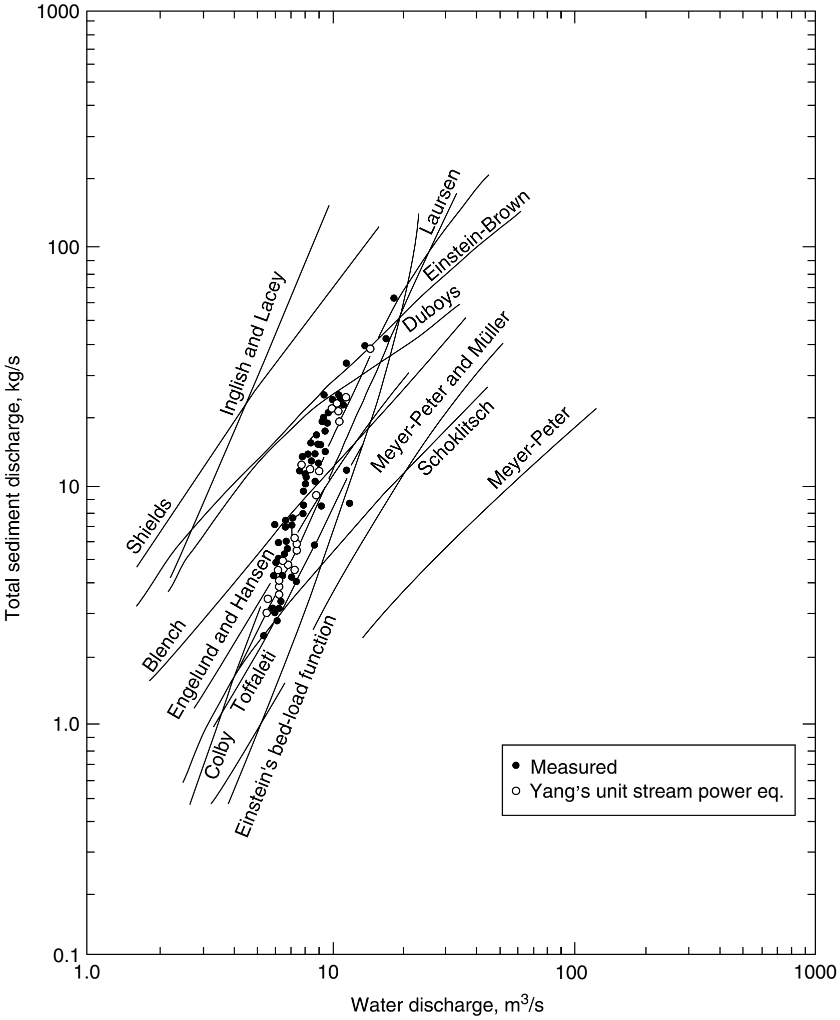

Figure 2. Comparison among measured total sediment discharge of the Niobrara River near Cody, Nebraska, and computed results of various equations (11).

BIBLIOGRAPHY

1. Julien, P.Y. (1998). *Erosion and Sedimentation*. Cambridge University Press, Cambridge, UK, p. 434.
2. Yang, C.T. (1996). *Sediment Transport: Theory and Practice*. McGraw-Hill, New York (reprint by Krieger, 2003), p. 396.
3. Simons, D.B. and Sentürk, F. (1977). *Sediment Transport Technology*. Water Resources Publication, Washington, DC, p. 807.
4. Yalin, M.S. (1972). *Mechanics of Sediment Transport*. Pergamon Press, Los Angeles, CA.
5. Einstein, H.A. (1950). *The Bedload Function for Sediment Transport in Open Channel Flows*. U.S. Department of Agriculture Soil Conservation Service Technical Bulletin No. 1026.
6. Colby, B.R. and Hembree, C.H. (1955). *Computation of Total Sediment Discharge, Niobrara River near Cody, Nebraska*. U.S. Geological Survey Water Supply Paper 1357.
7. Yang, C.T. (1972). Unit stream power and sediment transport. *J. Hydraul. Div. ASCE* **98**(HY10): 1805–1826.
8. Yang, C.T. (1973). Incipient motion and sediment transport. *J. Hydraul. Div. ASCE* **99**(HY10): 1679–1704.
9. U.S. Interagency Committee on Water Resources. (1957). *Subcommittee on Sedimentation*. Report No. 12, Some Fundamentals of Particle Size Analysis.
10. Rubey, W.W. (1933). Settling velocity of gravel, sand, and silt particles. *Am. J. Sci.* **25**: 325–338.
11. Yang, C.T. (1977). *The Movement of Sediment in Rivers. Geophysical Survey* 3. D. Reidel, Dordrecht, The Netherlands, pp. 39–68.
12. ASCE. (1982). ASCE Task Committee on Relations between Morphology of Small Streams and Sediment Yield of the Committee on Sedimentation of the Hydraulics Division. *J. Hydraul. Div. ASCE* **108** (HY11): 1328–1365.
13. U.S. Army Corps of Engineers. (1977). *Generalized computer program, HEC-6, Scour and Deposition in Rivers and Reservoirs, User's Manual*. The Hydrologic Center, Davis, CA (revised 1993).
14. Yang, C.T. and Simões, F.J.M. (2000). *User's Manual for GSTARS 2.1 (Generalized Stream Tube model for Alluvial River Simulation version 2.1)*. U.S. Bureau of Reclamation, Technical Service Center, Denver, CO.
15. Yang, C.T. and Simões, F.J.M. (2002). *User's Manual for GSTARS3 (Generalized Sediment Transport model for Alluvial River Simulation version 3.0)*. U. S. Bureau of Reclamation, Technical Service Center, Denver, CO.

SEDIMENTATION AND FLOTATION

Roger C. Viadero
West Virginia University
Morgantown, West Virginia

SEDIMENTATION

Overview

Waters and wastewaters typically contain particulate matter ranging from a few microns to centimeters in diameter, although the actual size distribution is often highly variable. One of the most common methods for the removal of particulate matter from waters and wastewaters is simple gravity settling, often referred to as "sedimentation."

In water and wastewater treatment, four general classes of sedimentation can occur: Type 1, discrete settling; Type 2, flocculent settling; Type 3, hindered or "zone" settling; and Type 4, compression settling. In Type 1 settling, particles settle individually with no interparticle interactions. Consequently, no change in particle size occurs (i.e., floc formation). This system is often analyzed using *Stokes' Law*, based on a force balance on a single particle settling in a viscous medium under laminar flow conditions. The development and application of this method of analysis is demonstrated in a subsequent subsection.

In flocculent settling, small particles aggregate to form larger particles. As the particles increase in size, the particle settling velocity increases (for particles of equal density). In order to make this process occur, chemicals are added to induce particle coalescence/coagulation and to promote the formation of "floc." In water treatment, a coagulant such as [hydrated] aluminum sulfate ("alum"), ferric sulfate, ferric chloride, silica, clays, and polymers can all be used to promote the formation of larger particles. In order to determine the optimum coagulant dose for a particular water, bench-scale tests must be performed. Further, it should be noted that the effectiveness of individual coagulants can be highly pH-dependent. For example, the formation of $Al(OH)_3(\downarrow)$ through the addition of alum requires $5 \leq pH \leq 8$ and sufficient alkalinity for the reaction to initiate, as presented in Eq. 1. (Note that alum can also be hydrated with 18 H_2O.)

$$\mathrm{Al_2(SO_4)_3 \cdot 14H_2O} + \underbrace{\mathrm{3Ca(HCO_3)_2}}_{\mathrm{ALK}} \rightarrow \mathrm{2Al(OH)_3}(\downarrow) + \mathrm{3CaSO_4} + \mathrm{14H_2O} + \mathrm{6CO_2} \quad (1)$$

Hindered or "zone" settling occurs when interparticle interactions inhibit the discrete settling of adjacent particles. In this case, particles tend to remain in a fixed position with respect to one another. Consequently, the agglomerated particles settle as a zone and a distinct solid-liquid interface develops at the top of the settling mass. This process occurs without the addition of chemical coagulants. In Type 4 or "compression settling", a structure is formed in which settling occurs via gravity compression of the structure, because of the weight of particles. Compression settling typically occurs in samples with high solids concentration, such as the secondary clarifier in wastewater treatment or in sludge thickening facilities (1,2).

Analysis and Modeling

To develop an understanding of the physics involved in sedimentation, it is useful to examine the simplest case of discrete settling. Based on the general force balance presented in Fig. 1, the velocity of a solid spherical particle of a given diameter and density settling in a viscous medium can be determined according to Eq. 2a.

$$F_{\mathrm{gravity}} - F_{\mathrm{buoyant}} - F_{\mathrm{drag}} = m_{\mathrm{p}} a_{\mathrm{s}} \quad (2a)$$

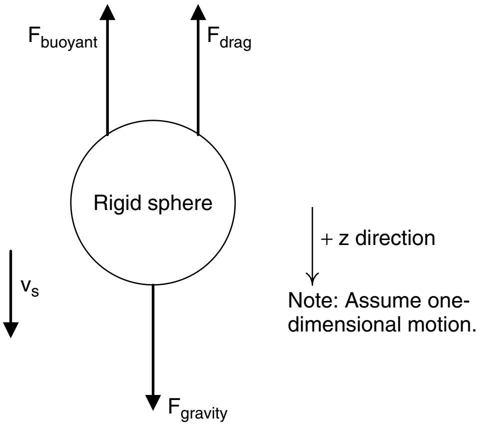

Figure 1. General force balance for a solid spherical particle settling in a viscous medium.

where $F_{\mathrm{gravity}} =$ gravitational force, $F_{\mathrm{buoyant}} =$ buoyant force, $F_{\mathrm{drag}} =$ drag force, $m_{\mathrm{p}} =$ particle mass, and $a_{\mathrm{s}} =$ acceleration of settling particle. If it is presumed that a terminal settling velocity has been reached (i.e., $a_{\mathrm{s}} = 0$), the force balance can be rewritten as:

$$F_{\mathrm{gravity}} - F_{\mathrm{buoyant}} - F_{\mathrm{drag}} = m_{\mathrm{p}} a_{\mathrm{s}} = m_{\mathrm{p}} \underbrace{\frac{dv_{\mathrm{s}}}{dt}}_{0} = 0 \qquad (2b)$$

The terms in the force balance can be written explicitly as presented in Eqs. 3a, 3b, and 3c.

$$F_{\mathrm{gravity}} = V_{\mathrm{p}} g \rho_{\mathrm{p}} \qquad (3a)$$

$$F_{\mathrm{bouyant}} = V_{\mathrm{p}} g \rho_{\mathrm{w}} \qquad (3b)$$

$$F_{\mathrm{drag}} = C_{\mathrm{D}} A_{\mathrm{p}} \rho_{\mathrm{w}} \frac{v_{\mathrm{s}}^2}{2} \qquad (3c)$$

where $V_{\mathrm{p}} =$ particle volume, $g =$ gravitational constant (9.8 m/s^2), $\rho_{\mathrm{p}} =$ density of particle (typical value = 1,400 kg/m^3), $\rho_{\mathrm{w}} =$ density of water (1,000 kg/m^3 under standard conditions), $C_{\mathrm{D}} =$ drag coefficient, $A_{\mathrm{p}} =$ cross-sectional area of the particle, and $v_{\mathrm{s}} =$ particle settling velocity.

Based on the general force balance presented in Fig. 1, and using Eqs. 2 and 3, the settling velocity for spherical particles of a given diameter and density can be determined as:

$$v_{\mathrm{s}} = \sqrt{\frac{4}{3} \frac{g(\rho_{\mathrm{p}} - \rho_{\mathrm{w}}) d_{\mathrm{p}}}{C_{\mathrm{D}} \rho_{\mathrm{w}}}}, \ \mathrm{Re} \geq 1 \qquad (4)$$

As only one-dimensional motion is considered, vector notation has been dropped and the sign convention presented in Fig. 1 is used throughout.

The drag coefficient for rigid spheres can be calculated according to Eq. 5, where C_{D} is a function of the hydraulic flow regime, established using the Reynolds number, Re, given in Eq. 6.

$$C_{\mathrm{D}} = \frac{24}{\mathrm{Re}} + \frac{3}{\sqrt{\mathrm{Re}}} + 0.34 \qquad (5)$$

$$\mathrm{Re} = \frac{v_{\mathrm{s}} d_{\mathrm{p}} \rho_{\mathrm{w}}}{\mu_{\mathrm{w}}} = \frac{v_{\mathrm{s}} d_{\mathrm{p}}}{\nu_{\mathrm{w}}} \qquad (6)$$

When spherical particles settle in a viscous medium under laminar flow conditions (Re < 1), the first term in Eq. 5 predominates (i.e., $C_{\mathrm{D}} \simeq 24/\mathrm{Re}$) and the expression for settling velocity becomes more simple:

$$v_{\mathrm{si}}(\text{Stokes'}) = \frac{g(\rho_{\mathrm{p}} - \rho_{\mathrm{w}}) {d_{\mathrm{pi}}}^2}{18 \mu_{\mathrm{w}}}, \ \mathrm{Re} < 1 \qquad (7)$$

This form of the settling velocity is known as "Stokes' Equation." As most waters and wastewaters contain particles of many different sizes, an additional subscript "i" is typically added to denote particles of the same size in an otherwise heterogeneous distribution. Consequently, a different settling velocity would have to be calculated for different size particles (1,2).

Application

In order to adequately analyze particle removal through sedimentation, it is necessary to establish a framework for analysis. In this case, consider the longitudinal cross section of a settling basin shown in Fig. 2. Assume a uniform distribution of particles at the entrance, along which particles can enter at any height. Further, any particle that settles into the sludge zone is permanently removed from the system. Now, define the "critical settling velocity" (v_{sc}) as the settling velocity of the smallest particle that will be 100% removed. To further ease the analysis of this system, the critical settling velocity vector will be broken into its horizontal (v_{h}) and vertical (v_{sc}) components. Note that in sedimentation, we are primarily concerned with motion in the vertical direction, so the vertical component of the settling velocity is often labeled "v_{sc}." The trajectory of particles with a velocity, v_{si}, equal to v_{sc} is also presented in Fig. 2. In this case, 100% of the particles are removed from the system, regardless of their entry position at the inlet.

Similarly, the trajectory of particles with a settling velocity greater than v_{sc} is presented in Fig. 3. [The trajectory of particles with a critical settling velocity (dashed lines) has been added to Fig. 3 for reference.] In this case, 100% of the particles are removed from the system, regardless of their entry position at the inlet.

The trajectory of particles with $v_{\mathrm{si}} < v_{\mathrm{sc}}$ is presented in Fig. 4. By examining the trajectories of these particles,

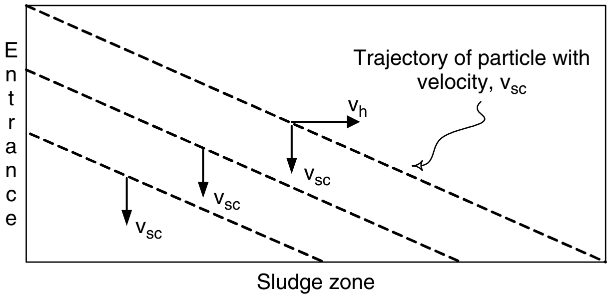

Figure 2. Trajectory of particles with a settling velocity equal to v_{sc}.

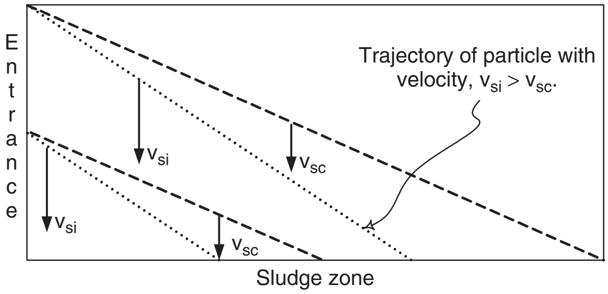

Figure 3. The trajectory of particles with a settling velocity greater than v_{sc}.

we can see that it is possible to have some particles "escape" through the outlet. Likewise, some particles may become trapped in the sludge zone, depending on their respective position at the inlet and the magnitude of v_{si} relative to v_{sc}. Consequently, when the settling velocity of an individual particle is less than the critical settling velocity, fractional particle removal will occur; the fraction of particles removed when $v_{si} < v_{sc}$ is given by Eq. 8 (1,2).

$$F_r = \frac{v_{si}}{v_{sc}} \tag{8}$$

Process Design and Implementation

In a typical water treatment process, sedimentation follows coagulation and flocculation, as presented schematically in Fig. 5. In activated sludge wastewater treatment, two sedimentation basins are used, as presented in Fig. 6. With regard to sedimentation basin design, it is common to determine the hydraulic detention time, θ_H, and hydraulic surface loading rate, S, presented in Eqs. 9 and 10, respectively. Typical design parameters for sedimentation tanks used in primary and secondary wastewater treatment are presented in Tables 1 and 2, respectively (1,3).

$$V_{basin} = \theta_H \cdot Q \tag{9}$$

$$S = \frac{Q}{A} \tag{10}$$

where V_{basin} = volume of sedimentation basin, θ_H = hydraulic detention time, Q = water/wastewater flow rate, and A = basin surface area.

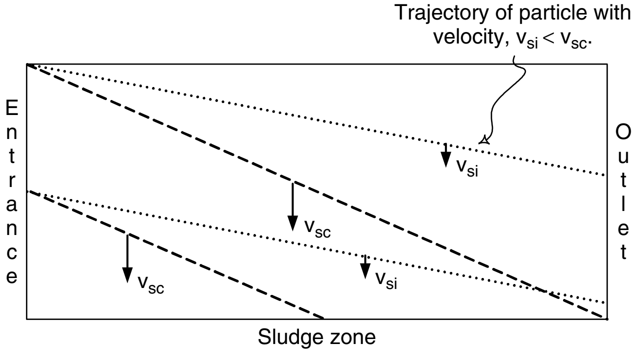

Figure 4. The trajectory of particles with a settling velocity less than v_{sc}.

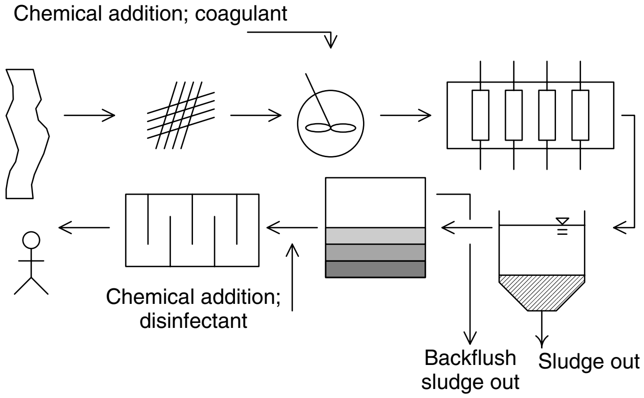

Figure 5. Schematic showing location of sedimentation in typical water treatment process.

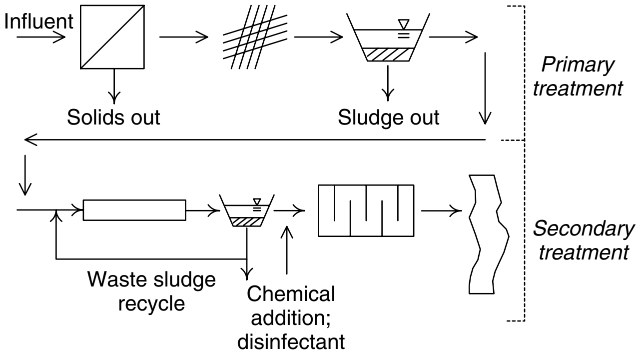

Figure 6. Schematic showing location of two sedimentation basins used in activated sludge wastewater treatment.

Table 1. Typical Design Parameters for Sedimentation Tanks Used in Primary Wastewater Treatment (1)

Primary Treatment Followed by Secondary Treatment	Range	Typical
Hydraulic detention time, hr	1.5–2.5	2.0
Average overflow rate (average flow), gal/ft^2·d	800–1,200	
Overflow rate (peak flow), gal/ft^2·d	2,000–3,000	2,500
Primary Treatment Followed by Waste Activated Sludge		
Hydraulic detention time, hr	1.5–2.5	2.0
Average overflow rate (average flow), gal/ft^2·d	600–800	
Overflow rate (peak flow), gal/ft^2·d	1,200–1,700	1,500

The data presented in Table 1 are for general reference purposes. In actual system design, factors such as water flow rate, basin depth, and desired solids removal must be considered (4).

FLOTATION

Flotation is a process for the separation of particulate matter from liquids in which bubbles are introduced into

Table 2. Typical Design Parameters for Sedimentation Tanks Used in Secondary Wastewater Treatment (3)

Secondary Treatment	Range
Average overflow rate (average flow), gal/ft$^2\cdot$ d	400–800
Overflow rate (peak flow), gal/ft$^2\cdot$ d	1,000–1,200
Average solids loading rate, lb/ft$^2\cdot$ d	0.8–1.2
Peak solids loading rate, lb/ft$^2\cdot$ d	2.0
Depth, ft	12–20

a wastewater and the solid particles rise with the bubbles. The particulate matter is then skimmed off the water surface and clarified water is removed from the bottom of the basin (1). A comparative schematic of sedimentation and flotation processes is presented in Fig. 7. Flotation is typically applied in cases where particles are difficult to remove via sedimentation; examples include synthetic fibers in the textile industry, wood fibers in the pulp and paper industry, and oily wastewaters (5). Flotation is also used in activated sludge thickening (1).

In general, air flotation and dissolved air flotation are the two types of flotation processes used most widely in industry. In air flotation, air is introduced via fine bubble diffusers. In this process, problems with diffuser clogging are common. Additionally, a large fraction of the operating cost is associated with the need to supply large amounts of power to force air bubbles through the diffusers.

Dissolved air flotation (DAF) is different from air flotation mainly in the conditions under which air enters the system. A schematic of a typical DAF process used to thicken activated sludge is presented in Fig. 8 (1). In dissolved air flotation, air is injected into wastewater under pressure. In typical wastewater applications, air is pressurized to 345–483 kPa (3.4–4.8 atm; 50–70 psi) (5), which results in an increase in the solubility of air in water. When the pressure is released to the atmosphere, the excess dissolved air comes out of the aqueous phase and rises. In order to enhance the particle removal efficiency of DAF processes, it is common to add chemicals that promote flocculation to the influent prior to pressurization. Reagents used for this purpose are similar to those used in coagulation/flocculation processes (e.g., alum, polymers, etc.) (1,5).

Vacuum flotation is a third technique in which wastewater is saturated with air under atmospheric pressure. A vacuum is applied to the wastewater in an enclosed tank, where the air is pulled from solution and particles rise to the top. The application of vacuum flotation is generally limited to small operations, because higher capital and operating costs associated with maintaining a closed reactor system.

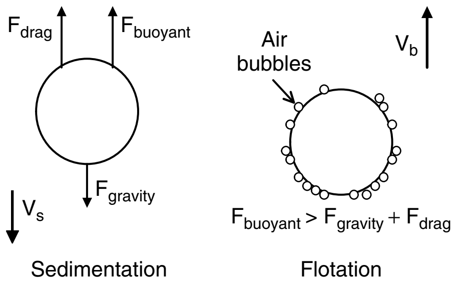

Figure 7. Comparative schematic of sedimentation and flotation.

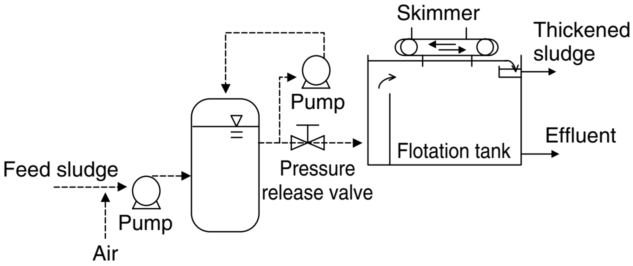

Figure 8. Schematic of a typical dissolved air flotation process used to thicken activated sludge. (Note: no recycle).

FLOTATION VERSUS SEDIMENTATION?

The selection of either sedimentation or flotation as a solid/liquid separation process depends on a large number of factors, including particle density, overflow rate, aesthetic requirements, sludge volume production rate, ease of operation, and availability of reliable design data. However, the guidelines presented in Table 3 can be used to compare and screen these processes for potential application.

BIBLIOGRAPHY

1. Metcalf and Eddy. (2003). *Wastewater Treatment, Disposal, and Reuse*, 4th Edn. McGraw-Hill, Boston, MA.
2. Tchobanoglous, G. and Schroeder, E. (1987). *Water Quality*. Addison Wesley, Reading, MA.
3. Water Pollution Control Federation. (1985). *Clarifier Design*. Manual of Practice FD-8.

Table 3. General Guidelines/Comparison for Selection of Sedimentation and Flotation Processes

Flotation	Sedimentation
Generally favored when particles with low density (low specific gravity) are to be removed.	Substantially less complicated to operate (e.g., fewer maintenance and operating problems, requires less skilled staff than flotation processes).
Can be used when space and/or capital are limited because higher overflow rates can be applied, which results in smaller tank requirements.	Requires less power for operation.
Odor problems can be minimized because of the lack of septic conditions.	Can easily be designed from the wealth of reliable performance data available.
A thicker sludge and, consequently, a lower sludge volume is produced.	

4. Reynolds, T. and Richards, P. (1995). *Unit Operations and Processes in Environmental Engineering*, 2nd Edn. PSW Publishing, Boston, MA.
5. Eckenfelder, W. (2000). *Industrial Water Pollution Control*, 3rd Edn. McGraw-Hill, Boston, MA.

RESERVOIR SEDIMENTATION

Sharad K. Jain
National Institute of Hydrology
Roorkee, Uttranchal, India

The detachment and transportation of soil is termed soil erosion. Every stream carries some sediment in suspension and moves larger particles along its bed. The deposition of sediment in channels or reservoirs creates a variety of problems, such as raising stream beds, meandering and overflow along the banks, and, of course, depletion of storage capacity in reservoirs.

The sediment content in river waters varies with flow conditions. It is small in lean flow months, but it attains a maximum value during floods. The total quantity of sediment transported annually to the sea by rivers of the world is about 2×10^{10} tons or about 13.5 km^3 in volume. Worldwide, reservoirs are annually losing about 1% of their storage capacity or about 65 km^3/year. The rate of loss of reservoir storage in the United States is about 0.22%/year which is equivalent to 2020 million m^3/year. Based on weighted average data from 144 reservoirs in India, the annual loss of gross storage was estimated at 0.44%.

Sediment source and reservoir locations are not uniformly distributed; the problem is really severe in some countries. China has more than 80,000 reservoirs that are annually losing, on average, 2.3% of their storage capacity due to sedimentation.

INTRODUCTION

Soil erosion is considerably high in arid climates. In India, about 5333 million tonnes (16.35 t/ha) of soil is detached annually due to agriculture and associated activities. Of this, about 29% is carried away to oceans by rivers. Nearly 10% of it is deposited in reservoirs resulting in loss of 1–2% of storage capacity (1). The sediment transport of selected rivers is given in Table 1.

Each sediment particle transported by flow is affected by two dynamic forces: a horizontal component acting in the direction of flow and a vertical component due to gravity; there is also the force of water turbulence. The specific gravity of soil materials is about 2.65, so particles of suspended sediment tend to settle to the channel bottom, but upward currents in the turbulent flow counteract gravitational settling. The sediment inflow and outflow in natural river reaches are mostly in balance. A reservoir changes the flow characteristics and its sediment transport capacity. Because the reservoir width is much bigger than the river width, the velocity of flow entering into it decreases tremendously and there is a dampening of turbulence. Consequently, the flow is unable to transport all the sediments, and the particles begin to deposit. First, the larger suspended particles and most of the bed load are deposited at the mouth of the reservoir. The smaller particles remain in suspension for a long time and some may leave the reservoir with water.

Table 1. Sediment Transported by Selected Rivers[a]

River	Drainage Area, km^2	Average Sediment Concentration, kg/m^3	Erosion Modulus, $ton/km^2/yr$
Nile	2,978,000	1.25	37
Missouri	1,370,000	3.54	159
Colorado	637,000	27.5	212
Indus	969,000	2.49	449
Irrawaddy	430,000	0.70	695
Brahmaputra	666,000	1.89	1090
Red	119,000	1.06	1092
Ganges	955,000	3.92	1519
Liaohe	166,300	6.86	240
Yangtze	1,807,200	0.54	280
Yellow	752,400	37.6	2,480

[a] Adapted from Reference 2.

RESERVOIR SEDIMENTATION

The accumulation of sediments is one of the principal factors that threaten the longevity of reservoirs. Sometimes a project is not constructed just because the silting rate is so high that the reservoir will fill up before the investment is fully recovered.

The ultimate destiny of all reservoirs is to be filled by sediment. There are instances of reservoirs filling up within a few years of their operation. The Sanmexia dam was the first major dam on the middle reaches of the Yellow River. In the first 18 months after dam closure, 1.8 billion metric tons of sediment accumulated in the reservoir, representing a trap efficiency of 93% (3). The Xinghe Reservoir in Shaanxi Province took 2 years to construct but only 1 year to fill with sediment.

The right approach to solving a reservoir sedimentation problem has three segments: (1) collect and analyze field data, (2) set up appropriate models, and (3) develop an operational policy for the reservoir. Deposition and scouring may differ considerably when different operating policies are adopted.

The sediment deposits in a reservoir can be divided in three groups: topset beds, foreset beds, and bottomset beds (see Fig. 1). Topset beds are composed of large size sediment deposits but may also have fine particles. These extend up to the point where the backwater curve ends. The downstream limit of the topset bed corresponds to the downstream limit of bed material transport in the reservoir. These deposits cause a minor reduction in reservoir storage capacity. The foreset deposit is the face of the delta deposit advancing toward the dam. It is a transition zone that has steeper slopes and decreasing grain size. The bottomset beds consist of fine sediments that are deposited beyond the delta by turbidity currents or nonstratified flow. This pattern of deposits may change due to reservoir drawdown, slope failures, and extreme floods.

For proper reservoir management, knowledge about the sediment deposition pattern in various zones is

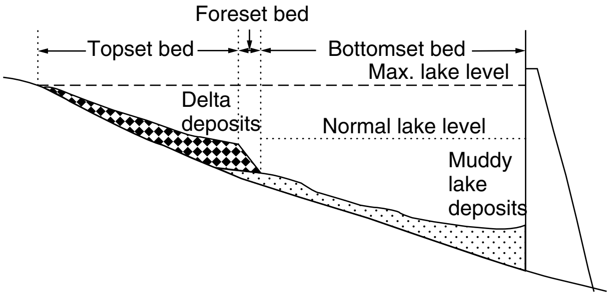

Figure 1. Sediment deposit zones in a reservoir.

essential. Timely remedial measures can be undertaken with correct knowledge of the sedimentation processes in a reservoir. A common procedure for tackling a sediment problem is to designate a portion of reservoir storage for sediment accumulation. But this approach does not solve the problem; it merely postpones the date when it becomes serious.

Factors Influencing Reservoir Sedimentation

Two dominant factors that influence the rate of silting in a reservoir are (1) capacity to inflow ratio (CIR) and (2) sediment content in the water flowing in. The other factors are the texture and size of the sediment, the trap efficiency of the reservoir, and the method of reservoir operation. The CIR is the ratio of reservoir storage capacity to mean annual inflow. A reservoir whose CIR is more than 50% is considered hydrologically large, may have significant carryover, and its trap efficiency will also be large. Note that the sediment inflow depends on the catchment area, too. All other things remaining the same, a dam of the same capacity in an upper catchment will have a higher rate of silting compared to a dam of the same design but constructed lower down in the valley.

The two principal factors mentioned above have a complete range of interplay. A reservoir that has a small CIR and small sediment inflow and one that has a large CIR and large sediment inflow may have more or less the same rate of sedimentation. A high rate of silting can also be expected from a high CIR and a high sediment content in the inflow. On the other hand, a high CIR and low sediment content in inflows will result in a low rate of silting.

Trap Efficiency

The trap efficiency of a reservoir is the ratio of sediment retained in the reservoir to the sediment inflow. It depends primarily on the sediment characteristics, the detention time of the inflow, the method of reservoir operation, and the age of the reservoir. A small reservoir on a large stream passes most of its inflow so quickly that the finer sediments are discharged downstream. A larger reservoir, on the other hand, may retain water for several years and the outflow from it may be completely devoid of suspended sediments. The trap efficiency of a reservoir decreases with age as the reservoir capacity is reduced by sediment accumulation. Thus complete filling of the reservoir may require a very long time.

The trap efficiency can be computed from the sediment inflow and outflow data. Brune (4) analyzed data from 44 reservoirs in the States; 40 were normal ponded reservoirs whose catchment areas vary from 0.098 sq. km to 478,110 sq. km and CI ratios range from 0.0016 to 2.05. The database also included two desilting basins and two semidry reservoirs. This analysis revealed that the laws of sediment deposition are the same for all types of reservoirs and the factors influencing the trap efficiency are independent of the size of the reservoir. Brune (4) presented a set of envelope curves between CIR and trap efficiency (see Fig. 2) that shows a semilogarithmic curvilinear relation between trap efficiency and CIR.

Sedimentation and Life of a Reservoir

The term *life of a reservoir* appears to be a misnomer because reservoirs do not have a single well-defined life which denotes two functional states: *ON* and *OFF*. Rather, they show a gradual degradation of performance. The 'life' of a reservoir for different purposes will be affected at different times. Murthy (5) defined the following terms connected to the life of a reservoir.

Useful Life. The period during which the capacity occupied by sediment does not prevent the reservoir from serving its intended primary purpose.

Economic Life. This is determined by the time after which the effect of various factors, such as physical deterioration by sedimentation and changing requirements for project services, cause the operating costs of the reservoir to exceed the additional benefits from its continuation.

Design Life. The period that is adopted for economic analysis. This is either the useful life or the shorter of the expected economic life or fixed span of life (say 50/100 years).

Full Life. The number of years required for full depletion of reservoir capacity by sedimentation.

LOSS OF STORAGE CAPACITY

Observations show that in reservoirs that have small sluicing capacity with respect to normal floods and have no upstream reservoirs, the siltation rate is comparatively high in the first 15–20 years and thereafter it falls off and may ultimately become negligible. From the data

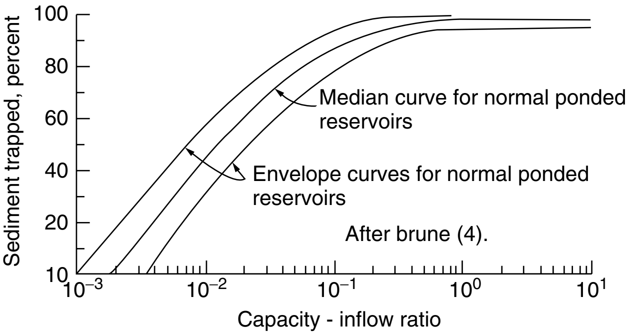

Figure 2. Brune's curves (4).

of reservoir capacity surveys, Shangle (6) found that the sedimentation rates in major reservoirs (storage > 100 million m^3) in India that have completed more than 50 years of their useful lives varied from 0.30 to 4.89 Ha-m/100 sq. km/year. The rate for those major reservoirs that have completed less than 50 years of their useful lives varied from 0.34 to 27.85 Ha-m/100 sq. km/year. Deyi and Fan (7) derived an empirical formula to determine the average annual loss rate of reservoir capacity.

The rate of siltation of a reservoir normally shows a falling trend with time. A plausible explanation is that the obstruction by the dam causes the dips and flanks of the storage basin to fill up with silt in early years. A stage comes when the river section adjusts itself to carry the normal discharge and disposal of the suspended load in the area of the reservoir is harmonized with the condition of the flow. Besides, the progressive development of deltas above the reservoir helps in trapping some of the silt load. Shrinkage and settlement of deposited silt also takes place with time due to superimposed loads of additional silt. This results in reduction in silt volume. However, a complete explanation of this behavior is not available.

Unit Weight of Sediments Deposited in a Reservoir

The unitweight of sediments is the dry weight (kg) per unit volume (cubic meter) of the material. It has been observed that the density of deposited sediment varies from 500–2000 kg/m^3. The important factors that influence the unit weight of deposited sediment are the manner of reservoir operation, the texture and size of sediment particles, and the consolidation rate. The reservoir operation is the most influential of these factors. If the water level of a reservoir is lowered from time to time, the deposited sediment is exposed to the Sun and air and gets dense. There may be considerable time for consolidation before floods. The degree of consolidation depends on the weight of overlying material, its exposure, sediment size, and time. The reservoir is always filled that has a low density of deposit. Power and irrigation reservoirs belong to the intermediate class. Generally, lower densities are observed in the vicinity of a dam under submerged conditions whereas higher densities are noticed in the upstream portions of reservoirs.

Lara and Pemberton of U.S.B.R. developed a method for estimating the initial unitweight of sediment deposits when the size analysis of the incoming sediment and the proposed reservoir operating schemes are known. Reservoir operations were classified in different types as follows:

Reservoir Type	Reservoir Operation	W_c	W_m	W_s
I	Sediment always submerged or nearly submerged	26	70	97
II	Normally moderate to considerable reservoir drawdown	35	71	97
III	Reservoir normally empty	40	73	97
IV	Riverbed sediments	60	73	97

The unitweight of the sediment deposits can be estimated from:

$$\gamma = 16.05(W_cP_c + W_mP_m + W_sP_s)$$

where γ is unitweight in kg/m^3; P_c, P_m, P_s are percentages of clay, silt, and sand, respectively, of the incoming sediment; and W_c, W_m, W_s are the coefficients of clay, silt, and sand, respectively, which may be obtained from the table above (15).

Distribution of Sediments in Reservoirs

Sediment distribution in a reservoir is important, and this requires careful consideration in planning and design stages. The pattern of sedimentation helps in predicting the extent to which services will be affected at various times and the remedial actions to be taken. The designer is interested to know the height of sediment accumulation to fix the sill elevation of the outlets, the penstock gate elevation, and to estimate the region where a delta would be formed and the consequent increase in backwater levels. Finally, the pattern is necessary to plan for recreational facilities. Four types of distribution patterns of deposits have been identified (3):

Delta Deposits. These deposits contain the coarsest fraction of the sediment load that is rapidly deposited in the zone of inflow.

Wedge-shaped Deposits. These are thickest at the dam and become thinner moving upstream.

Tapering Deposit. These occur when deposits become progressively thinner moving toward the dam.

Uniform Deposits. The thickness of these deposits is more or less uniform in the reservoir. Such deposits are not very common.

The consequences of reservoir sedimentation are gradual reduction in benefits. The extent of loss depends on the type and nature of the purposes being served and the rate of loss of storage capacity. The loss turns out to be really high when the replacement cost of the lost storage is considered.

A commonly used empirical method to estimate the new reservoir profile is discussed next.

Empirical Area Reduction Method

Based on the analysis of the sediment survey data of many reservoirs in the United States, it was found that the sediment distribution pattern depends on reservoir geometry, operation, and sediment characteristics. Depending upon the elevation–capacity characteristics, reservoirs are classified into four types: (a) gorge, (b) hill, (c) flood plain–foot hill, and (d) lake. Empirically derived sediment distribution curves are used to distribute the sediment throughout the reservoir section. The *Empirical Area Reduction* method, as revised by Lara (8), is the most popular for predicting a new reservoir bed profile. It is necessary first to estimate the amount of sediment deposited in the reservoir. The computational steps have been given in many books, such as Morris and Fan (3).

RESERVOIR SURVEYS

Sediments accumulated in an existing reservoir can be determined by periodically running sediment surveys. It is a direct measurement procedure for assessing the volume and pattern of deposits. Sediment data collected during the surveys are analyzed to determine the specific weights of the deposits, their grain size distribution, sediment accumulation rates, and reservoir efficiencies. Recent advances in technology have considerably reduced the effort to carry out reservoir surveys and analyze data.

The advantages of reservoir surveys are as follows:

1. The reservoir survey can be less costly than continuous sediment measurement at several locations in the catchment.
2. The accuracy of these surveys is usually very high, particularly if advanced equipment is used.
3. It is possible to estimate the total sediment (bed and suspended) load carried by a river.

There are some limitations of the reservoir sedimentation survey:

1. Such surveys do not provide any information about the variation of sediment yield with time and give only the total sediment accumulated since the last survey.
2. This method does not provide subcatchmentwise sediment yield.
3. This approach is not very effective where sedimentation is small as the errors of measurement may mask the true sedimentation rates.
4. To find the total sediment inflow, information about sediment outflow is also needed.

It is essential to have an accurate map of the reservoir on an appropriate scale before commencing a survey. Important reservoir features, such as the full reservoir level (FRL) along the periphery, the position of the dam, outlets, inflowing streams, etc., along with other details such as the position of nearby bridges, roads, and villages should be marked on the map. Horizontal and vertical control points are fixed at a suitable interval on the reservoir circumference. Next, cross-sections are planned at a suitable spacing, depending on the reservoir size.

The contour and range methods are two basic techniques for surveying reservoirs. The selection of a method depends on the quantity and distribution of sediment indicated by field inspections, the shape of the reservoir, the purpose of the survey, and the degree of accuracy desired. The volume of accumulated sediment is computed by subtracting the revised capacity from the original capacity at a given reservoir elevation (usually the FRL).

The frequency of surveying reservoirs depends on the sediment accumulation rate and the cost of a survey. Reservoirs that have high accumulation rates are surveyed more often than those with lower rates. Generally, the reservoirs are surveyed every 3 to 10 years. Special circumstances may necessitate a change in the established schedule. A reservoir might be surveyed after a major flood that has brought in a heavy sediment load. A survey may also be run following the closure of a major dam constructed upstream in the same catchment. An upstream dam reduces the free drainage area and hence reduces the sediment inflow.

Remote sensing techniques, enabling acquisition and analysis of synoptic data over a broad spectral range, are an alternative to the conventional way of data acquisition and processing. This technique is highly cost- and time-effective compared to conventional methods. The methodology for assessing reservoir sedimentation is described by Goel et al. (9).

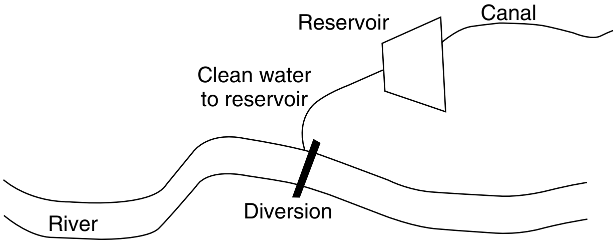

Figure 3. An offstream reservoir.

CONTROL OF SEDIMENT INFLOW INTO A RESERVOIR

The problem of reservoir sedimentation is complex but not unmanageable. It can be largely controlled by judicious design, construction, and management. The Yellow River in China is notorious for very high volumes of sediments in its water. The Sanmenxia Dam was the first dam built in the middle reach of the Yellow River in China. After the impounding commenced, about 1.8 billion metric tonnes of sediments accumulated during the first 18 months, representing a trap efficiency of 93%. The balance between sediment inflow and outflow was restored by providing high capacity bottom outlets, and reservoir operation was substantially changed.

There are various ways to manage the sedimentation problem, and the effectiveness of an approach depends on the site conditions. A single technique will not work everywhere. Broadly, the methods are

- Control the sediment inflow into the reservoir.
- Do not allow the entering sediment to settle in the reservoir.
- Remove the settled sediment from the reservoir.

Regarding the first approach, broadly, there are three ways to prevent sediment from entering a reservoir. These are (1) constructing reservoirs away from streams, (2) physical barriers in the way of sediment movement, and (3) better watershed management to check soil erosion.

Offstream Reservoirs

If the site conditions permit, a reservoir can be constructed away from the main stream (Fig. 3). This reservoir is filled with water of low sediment concentration that is diverted from the main river. The flows that contain high amounts of sediment load are excluded from diversion. An additional advantage of an offstream reservoir is that only a desired amount of flow is diverted to it, and therefore, its capacity can be limited.

Check Dams

A check dam is a small dam a few meters high that is constructed across a stream in the headwater region to control channel erosion. Most of the incoming sediment is deposited behind the check dam and relatively clear water comes out downstream. These dams can be highly effective in controlling reservoir sedimentation. The cost of a small check dam is not very high. These dams are more efficient for sediment trapping if they are spaced farther apart. These dams are not a long-term solution to the problem because they do not control the sediment erosion, and once the reservoir behind the check dam is filled with sediments, it no longer serves the purpose.

SEDIMENT ROUTING

This is an effective method for controlling reservoir sedimentation. It includes methods to manage the hydraulic behavior of the reservoir to allow the maximum sediments to pass through. The concepts of sediment routing were developed in China. The guiding principle is *discharge the muddy water, impound the clear water*. It is an environmental friendly approach because the characteristics of sediment transport are not significantly changed.

Reservoir Drawdown

Typically, the rising limb of the hydrograph carries larger amount of sediments than the falling limb. In this approach, the reservoir water level is bought down in the beginning of a flood to pass turbid flow without deposition.

Density Currents

A density current is the gravity flow of fluid under, over, or through another fluid of approximately equal density or density whose differs by a small amount from that of the primary current. Density currents are generated when sediment-laden water enters a relatively still waterbody. The density current venting approach (Fig. 4) provides a clear unhindered path to sediment-laden flows that are released through low-level outlets.

Sediment Bypass

This arrangement, shown in Fig. 5, prohibits sediment-laden flow from entering the reservoir. It consists of a diversion structure upstream of the dam by which the flows with heavy sediment concentration are diverted to a bypass channel or conduit which joins the main river downstream of the dam. As a result, relatively 'clean' flows enter the reservoir.

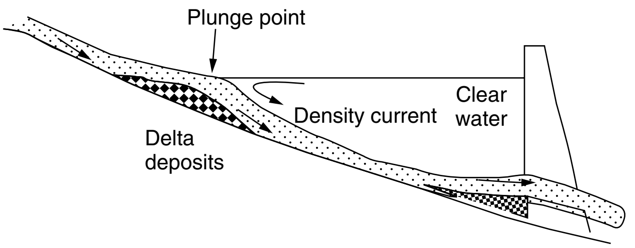

Figure 4. Venting of density current through bottom outlet in a reservoir.

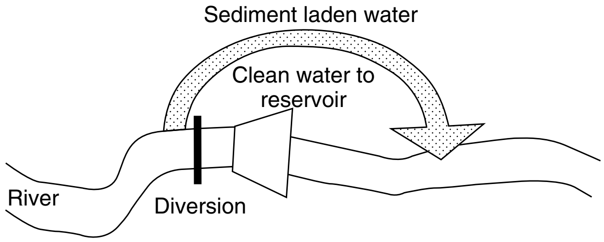

Figure 5. Sediment bypass.

BIBLIOGRAPHY

1. Dhruva Narayana, V.V. (1995). *Research in Soil and Water Conservation in India with Special Emphasis on Watershed Management*. Scientific Contribution No. INCOH/SAR-5/95, National Institute of Hydrology, Roorkee, India.
2. Qiang, N. and Dai, D. (1980). The problems of river sedimentation and the present status of its research in China. In: *Proceedings of the International Symposium on River Sedimentation, Beijing. The Chinese Society of Hydraulic Engineering* (Vol. 1). Guanghua Press, UNESCO, pp. 19–39.
3. Morris, G.L. and Fan, J. (1998). *Reservoir Sedimentation Handbook*. McGraw-Hill, New York.
4. Brune, G.M. (1953). Trap efficiency of reservoirs. *Trans. Am. Geophys. Union* **34**(3): 407–418.
5. Murthy, B.N. (1977). *Life of Reservoirs, Technical Report No. 19*. Central Board of Irrigation and Power, New Delhi.
6. Shangle, A.K. (1991). *Reservoir Sedimentation—Status in India*. Jalvigyan Sameeksha, Vol. V. INCOH, National Institute of Hydrology, Roorkee, India, pp. 63–70.
7. Deyi, W. and Fan, J. (1991). Method of preserving reservoir capacity. In: *Lecture Notes of Regional Training Course on Reservoir Sedimentation and Control*. Central Water Commission, New Delhi.
8. Lara, J.M. (1962). *Revision of Procedures to Compute Sediment Distribution in Large Reservoirs*. US Bureau of Reclamation, Denver, CO.
9. Goel, M.K., Jain, S.K., and Agrawal, P.K. (2002). Assessment of sediment deposition rate in Bargi reservoir using digital image processing. *Hydrological Sci. J.* **47**(S): S81–S92.

WATER FROM SATURATED RIVER SEDIMENT—SAND ABSTRACTION

Stephen W. Hussey
Dabane Trust
Burnside, Bulawayo, Zimbabwe

Water is available in the saturated sediment of sand-filled rivers. The removal of water from this sediment is known as sand abstraction. Despite the terminology, it is not the removal of sand but of water from a surface dry, seasonal river where the water is retained in the sediment, not abstraction of water from a flowing river.

Wherever there is erosion of land within the catchment area of a river, soil, sand, or gravel is carried into the river channel. The river becomes silted with sediment retained in depressions in the riverbed, behind rocks, and where the slope of the river is sufficiently low for sediment to be deposited rather than washed away.

In arid and semiarid areas, the sediment that accrues is often coarse grained, and, as the river basin drains, water is retained in the sediment in useable quantities. In spite of the intermittent flow of seasonal rivers, as the rivers drain and dry out, water is held in the sediment.

BACKGROUND INFORMATION

History and Traditional Use

There have been sand-filled rivers for centuries in arid and semiarid areas where intermittent rainfall and long dry periods exist (Fig. 1). These rivers have become increasingly filled with sediment since the land has become populated and built up. As surface water dropped into the sediment, people had to dig down to find their daily water requirements. This practice continues today in many arid areas.

Traditionally, people lived in river valleys, and as the seasonal rivers dried out, they excavated inverted cone-shaped holes to expose the water level in the sediment. As the loose, unstable sand continually filled the hole, people were forced to re-excavate their holes. A natural development was to insert an opened ended drum into the river sediment as a liner. This made a ‘sand well’ that was more stable and did not fill with sand after every use (Fig. 2). The water abstracted was then used for domestic, livestock, and small-scale gardening purposes.

Typical Uses

Domestic Use. A properly managed sand-abstraction system provides a suitable source of water for domestic use. A variety of schemes have been developed ranging from small-scale, individual families using basic hand pumps to reticulated systems for small towns.

Irrigation Schemes. These range from small family gardens to large-scale commercial schemes. Where gardens can be established on low river banks (above the flood line), sand-abstraction systems provide an excellent, low-cost option for conveying water to nutrition gardens. One small hand pump (for instance, a Rower) is sufficient to irrigate adequately more than 200 m^2 of garden—an average size brushwood-fenced garden (Fig. 3).

Figure 1. Typical sand river.

Figure 2. Traditional sand well.

Figure 3. A small-scale community garden where water is drawn by hand pumps from the sand river in the background.

Livestock Watering. Sand-abstraction systems provide a very suitable source of water for livestock, even in less than optimum conditions. A hand pump, either driven directly into the sand, discharging into a portable trough on the river sand, or a more complete system drawing water onto the river bank are both appropriate livestock watering systems.

Livelihood. Sand abstraction can provide water for small-scale, rural industry such as brickmaking, household construction and repair, and small livestock projects.

ADVANTAGES AND DISADVANTAGES

Advantages

1. Clean, naturally filtered water free of the contamination associated with open surface water such as dams.
2. The water has a natural taste free from mineral salt contamination often associated with groundwater.
3. Easy to access when compared with the labor involved in building a dam or digging a well.
4. On the traditional and small-scale level, the technology does not require pumps and expensive drilling equipment. It is a low-level technology that people can operate for themselves without outside intervention.
5. Lower cost per cubic meter yield of water compared with ground- or surface water developments.
6. Where extensive sediment deposits exist, large volumes of water can be held.
7. Reduced evaporation from water retained in the sediment as opposed to open surface water.
8. Low environmental impact from installation of sand-abstraction systems.
9. The technology can be used in various other medias such as tube wells and gravel beds.

Disadvantages

1. The system is as vulnerable as any other supply to inadequate recharge from poor rainfall and to seasonal water losses.
2. As there is no confined aquifer at the point of abstraction, if sites are not carefully selected, water can drain away from the site.
3. If not adequately secured, equipment can be washed away in flash floods.
4. The system is based on environmental degradation, it works only where there has been severe soil erosion.
5. At some sites, silt layers form and move through the river sand to clog installations.
6. In certain conditions, infiltration galleries and well points may be subject to bio-fouling.

EPHEMERAL RIVERS

In ephemeral river systems, people have, for generations, been drawing water from river pools and from the river sand. The system works best in large, slow-flowing rivers that cross extensive plains away from the watershed. Conditions for sand abstraction are ideal where rivers cut through igneous soils. Soils in these areas tend to contain large proportions of coarse-grained sands that are carried into and remain in the river systems and allow large volumes of water to be retained. The system is dependent to some degree on continuing soil erosion and is inappropriate in small fast-flowing rivers below the watershed that contain small amounts of unstable sand.

Table 1. Comparison of Community Water Supply Quality—Groundwater and Sand-Abstraction Systems—Zimbabwe

	Recommended Limit	Max Allowed Limit	Source Borehole 1	Source Borehole 2	Source Borehole 3	Source Open Sand Abst	Source 'Safe' Sand Abst
Conductivity (mS/m)	70	300	460	400	940	50	40
pH	6–9	5.5–9.5	7.3	7.5	8.1	7.9	7.3
Turbidity (NTU)	1.0	5.0	30	120	55	20	6.5
Total hardness	20–230	650	1100	110	370	130	120
Calcium ($CaCO_3$)	NS	NS	780	20	100	120	68
Magnesium ($CaCO_3$)	100	150	320	90	270	10	52
Sodium (Na)	100	400	390	600	910	20	9.7
Potassium (K)	NS	NS	6.5	4.0	5.8	1.5	1.5
Iron (Fe)	0.1	1.0	6.6	11.0	0.9	0.5	1.1
Manganese	0.05	1.0	2.6	0.4	0.2	0.1	2.8
Alkalinity ($CaCO_3$)	NS	NS	610	190	500	360	200
Chloride	250	600	22	110	58	2	8
Sulfate	200	600	1600	1700	290	1.7	<0.01
Phosphate	NS	NS	0.1	<0.01	<0.01	<0.01	<0.01
Ammonia total nitrate nitrogen	6.0	10.0	0.3	0.1	0.5	<0.01	<0.01
	–	–	<0.01	<0.01	0.5	<0.01	<0.01
Fluoride	1.0	1.5	0.9	<0.01	<0.01	<0.01	<0.01
Approx dissolved salts	500 (WHO)	1500 (WHO)	2300	2000	4700	250	200
Oxygen absorbed 4 h, 27 °C	NS	NS	2.0	<0.01	8.0	3.3	<0.01

[a] NS: not significant.

[b] WHO

The system is conditional on water freely percolating through sediment to the point of abstraction. Silt and clay create an impenetrable barrier to an abstraction system, and thus sand abstraction in such conditions is not an appropriate option. The ideal situation is a slow-flowing, wide river that has deep, coarse sediment. The ideal site is above a natural rock barrier or in a depression of the riverbed (typically, a former pool), which is continually recharged by water percolating from an expanse of sand above.

The system can be made to work on smaller rivers that contain coarse-grained sand by constructing a subsurface dam to retain sediment. This consists of a barrier to retain the water in the sediment. It is constructed across the river from the riverbed to the surface of the river sediment. A further possibility is a sand dam that is an impoundment, generally a weir, constructed to raise the level of river sediment. The impoundment is designed to silt up over several years; the entire length of the wall is raised by only some 0.3 to 0.5 meters a year. Coarse-grained sand is impounded and the finer silt is washed out of the weir each year. Water can be abstracted from these dams as through any other sand-abstraction system.

GEOGRAPHICAL CONTEXT

Suitable ephemeral rivers are extensive in much of southern Africa, parts of east Africa, the Sahel, the Middle East, the Indian subcontinent, the Southeast United States, parts of Mexico, and parts of Latin America and Australia. The technology of sand abstraction is particularly used in these arid and semiarid areas:

southern Africa—South Africa, Namibia, Swaziland, Botswana, Zimbabwe, and Zambia;
east Africa—Kenya, Ethiopia, Sudan, and Somalia.

WATER QUALITY

Unlike groundwater supplies, where water may be contaminated with mineral salts, water obtained from sand-abstraction sources is generally not tainted. Surface water supplies are very quickly contaminated by livestock that wade, urinate, and defecate in the water. Surface water should, therefore, not be considered fit for human consumption without treatment. Where extensive sands exist, the entire waterway acts as an enormous sand filter bed for a sand-abstraction system. Water drawn from a properly constructed scheme may thus be considered potable and require no treatment. However, unprotected sites can become clogged with deposits of clay or fouled with livestock droppings.

SUSTAINABILITY—VILLAGE-LEVEL OPERATION AND MAINTENANCE

1. Sand-abstraction systems draw water from river sands, and thus they are typically shallow water sources requiring only simple hand pumps, which communities that have few tools and little expertise can operate and maintain. In their simplest form, while still ensuring a protected water supply, buckets are quite adequate for drawing water.
2. Sand-abstraction sites can be identified by villagers, whereas groundwater sources generally require either a geophysical survey or dowsing. A greater sense of ownership and responsibility is thus promoted by sand-abstraction schemes.
3. Many groundwater sources are deep and beyond the capacity of village people to operate and maintain. Boreholes are expensive to drill and can seldom be undertaken by rural communities on their own. In many areas, the aquifers are small and are quickly overabstracted, resulting in blocked fissures and ultimately dry boreholes or wells. Boreholes and wells in sandstone localities are prone to collapse or become completely blocked by fine sediment that slowly percolates into the borehole and granite areas.
4. Small-scale units that can be managed by a Village Water Committee are a very practical option, especially in the Southern Hemisphere, considering the large number of possible sand-filled river sites.
5. In some rivers, there is a possibility that flooding will wash the well point (abstraction point) away. A further complication is that silt layers can move through the river sand and clog installations. Only practical experience can rule out these possibilities.

METHODS OF ABSTRACTING WATER FROM SAND

There are several systems of sand abstraction that have been developed over the years. Each depends on equipment that can be installed into the water-bearing river sand and at all times remain in free moving water. Schemes range in size from small-scale, hand-operated supplies right through to large-scale mechanized (diesel or electric-powered) irrigation schemes on large rivers (Figs. 4 and 5).

Well Screens

Well points or well screens, whose slots are generally 0.5, 1.0, or 1.5 mm wide, prevent entry of river sand into the

Figure 4. Small-scale, hand pump sand-abstraction scheme.

Figure 5. Submersible pumps drawing water for a commercial irrigation scheme from the sand of the Limpopo River in the background.

pump system. Slot size is selected in accordance with the coarseness of the river sand so that sand whose grain size is a diameter greater than the slot is prevented from entering. Smaller sand is initially drawn through the slots but coarser sand quickly collects around the immediate screen and then blocks the entry of finer sand; thus a natural graded filter is developed around the well point.

One or more well screens are connected to a suction pump on the river bank. Installations are done either when the river water is at its lowest level so that the well screen can be dug into the sand, or when the river sand is saturated with water to full depth. Well screens can then be pushed, or 'jetted,' into the lower levels of sand in an artificial 'quicksand' condition, caused by a water jet from a motorized centrifugal pump. This system of installation is relatively technically complex and depends on materials and equipment not readily available (Fig. 6).

Manifold and Well Points

This system depends on low velocity of water at the point of abstraction, sufficient to draw off water without drawing sand into the system. It is best installed toward the end of the dry season when the river water level is at its lowest. River sand is removed to the water-bearing level and a 'manifold' (a large diameter pipe) is laid on the water-bearing sand. Well points are installed at an angle into the sand and connected alternately at each side of the manifold. The manifold is connected to a suction pump on the riverbank. The number of well points, manifold size, and piping to the riverbank are calculated according to the velocity of water required in different parts of the system and on the quantity of water required. This is probably the most common system presently used, and, in general, all necessary materials are readily available.

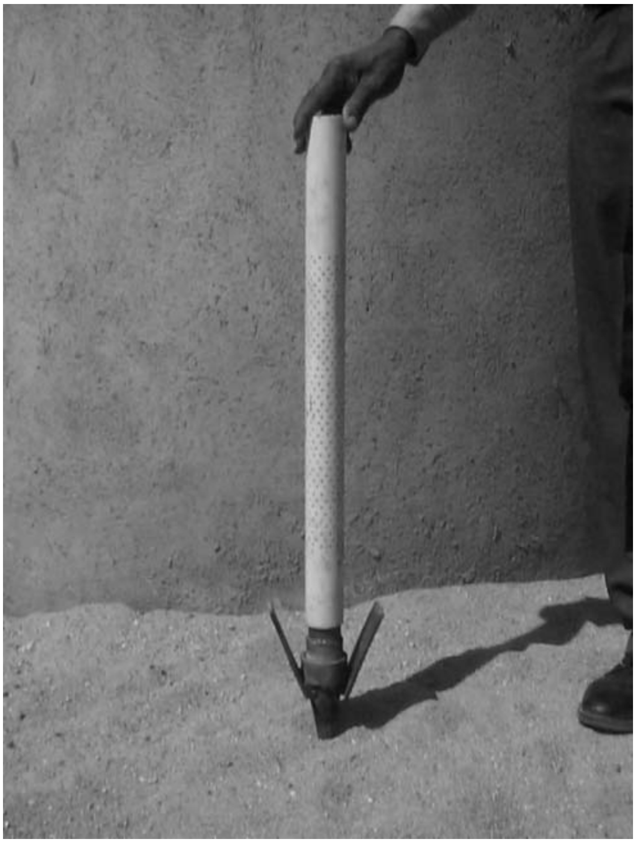

Figure 6. Well points suitable for small-scale sand-abstraction system.

Infiltration Galleries

Many original sand-abstraction installations were gravity supply systems and consisted of a number of interconnected well points. Sand had to be kept out of the pipes so that water could run in and flow to a sump or false well on the riverbank from where it could be pumped out. Modern synthetic materials have recently improved the potential of this system; before they were introduced, it was difficult to keep sand out. However, it is still difficult to ensure that the pipe system is sufficiently deep in the sand to keep it in water year-round. Further, it necessitates a lot of digging into the riverbank. It is relatively inexpensive and does not require complex equipment or expertise to operate or maintain. A windlass or basic hand pump can be used to draw water.

Caisson

Water is abstracted through the slotted base of a large (1.0 to 1.5 m diameter) flat cylindrical or slightly conical ‘caisson’ connected by pipes to a suction pump on the riverbank. The ‘caisson’ is dug into river sand as deeply as possible, and as the water level drops during the season, it is lowered to keep it in water continually. As in any of the systems, the installation is complete once a level has been reached that remains in water year-round. This system is more awkward due to the continual redigging and lowering, but once at its lowest point, can be used where silt tends to accumulate because of the large surface area for abstraction. However, slow recharge through the silt to the abstraction zone can significantly impede abstraction.

COST-EFFECTIVENESS

Small-Scale Schemes

Small sand-abstraction installations do not require any complex machinery such as tractors, graders, drilling rigs, or welders. People at the village level can undertake the installations themselves using primarily locally available equipment. Well points can be easily fabricated in PVC or steel and either jetted, driven, or dug into the saturated sediment. Water can then be abstracted with basic low-tech pumps that do not require expensive operation and maintenance.

Commercial Schemes

Some heavy machinery is needed for initial preparation of the river. However, overall, there is less material to be removed when installing sand-abstraction well points or infiltration galleries than in the constructing an earth embankment dam. Pumps that would typically be used to draw water from a dam or aquifer-based irrigation scheme may also be used in systems based on sand abstraction.

SEDIMENT TRANSPORT

Matjaz Mikos
University of Ljubljana
Ljubljana, Slovenia

SEDIMENT SOURCE AREAS

Sediment transport is a natural process of transporting solid particles called sediments from erosion sources to deposition or sedimentation areas. This process has been shaping landscapes throughout geological time. Sediments, sedimentary rocks, and the process by which they were formed are scientifically studied by sedimentology.

One important source of sediments is soil erosion. Geologic soil erosion occurs as surface removal of farm, forest, and other erodible soils in the form of sheet erosion, interrill and rill erosion, gully erosion or deflation, mainly caused by impacts of raindrops (splash erosion, rain erosion) and overland flow on bare soils or land without dense vegetative cover to protect soils. It may be checked on farm soils by such farm management practices as contour plowing, crop rotation, and mulching. Accelerated erosion is the speeding up of erosive processes, such as deforestation, overgrazing, and construction, either directly or indirectly, by the intervention of humans. Its most obvious effects can be seen in the alteration of a river’s regime by improper land use or by human intervention into the catchment area. Other important sediment sources are deposits of mass wasting phenomena, driven by gravity, such as rockfalls, rockslides, landslides, debris flows, and mudflows, exposed to fluvial processes of erosion, transportation, and sedimentation. The subjects of these processes are fluvial sediments—solid particles and particles of rock material that have been transported from their sources and deposited elsewhere by flowing water. A sediment particle is a solid particle, which settles after being suspended in a fluid, but the term usually includes all detritus produced by erosion processes and deposited by fluvial, lacustrine, marine, glacial, and aeolian agencies in the process of sedimentation. The term fluvial sediment is usually reserved for unconsolidated material found in a fluvial system.

PHYSICAL PROPERTIES

Sediment transport involves the two-phase flow of a water–sediment mixture. The most important physical properties of water as a nearly incompressible fluid are the mass density of the fluid (mass of fluid per unit volume), the specific weight of the fluid (fluid weight per unit volume of fluid), the dynamic viscosity (proportional factor between the shear stress causing fluid deformation and the deformation rate), and the kinematic viscosity (the dynamic viscosity of a fluid divided by the mass density of the same fluid). Both the density and the viscosity of water decrease as the temperature increases.

The most important physical properties of sediments as a solid phase can be divided into

— the properties of a single sediment particle: the mass density of solid particles (the mass per unit volume), the specific weight of solid particles (the solid weight per unit volume of solid), the submerged specific weight of a particle (the difference between the specific weights of solid particles and water), the specific gravity (the dimensionless ratio of the specific weight of a solid particle to the specific weight of a fluid at a standard reference temperature), the sediment size (e.g., sieve diameter, fall diameter, nominal diameter), the sediment shape (different shape factors and indexes, e.g., sphericity), and the sediment roundness (different indexes using the sharpness of the edges or corners of a particle to classify particles from angular to rounded);

— the bulk properties of sediments: the particle size frequency distribution (measured by wet or dry sieving equipment such as mechanical shakers or sedimentation tubes and represented as a histogram or a cumulative curve) and corresponding particle size parameters (mean, median, mode, sorting, skewness, kurtosis), the angle of repose (usually given as the slope angle of a cone of submerged loose material under incipient sliding conditions), the porosity (the volume of voids per total volume), the void ratio (the volume of voids per volume of solid), the unit weight (the weight of solid and water in voids per unit volume), and the dry unit weight (the weight of solid per unit volume).

The most important physical properties of a water-sediment mixture are the specific weight of the mixture (the total weight of solid and water in voids per unit volume), the specific mass of a mixture (the total mass of solid and water in voids per unit volume), the volumetric sediment concentration (the volume of solids divided by the total volume), the dynamic viscosity of a mixture (a function of the dynamic viscosity of water and the volumetric sediment concentration), and the kinematic viscosity of a mixture (the dynamic viscosity of a mixture divided by the mass density of the same mixture).

INCIPIENT MOTION

Due to the stochastic nature of sediment transport, it is difficult to define precisely the flow condition at which a sediment particle will begin to move. Fluid flow around sediment particles resting on the bottom of the flow exerts forces that tend to initiate motion. The resisting force of noncohesive sediment relates to submerged particle weight. Threshold conditions occur when the hydrodynamic overturning moment of forces acting on a single particle exceeds the resisting moment. Then the particle is in incipient motion, thus defining the threshold condition for sediment transport. The hydrodynamic forces acting on a noncohesive sediment particle are the submerged particle weight, the lift force, the drag force, and the resistance force. Most incipient motion criteria are derived from either the shear stress or the velocity approach. Because of the stochastic nature of sediment transport, probabilistic approaches have also been introduced.

The well-known Shields shear stress approach determines the threshold condition from the shear stress, the submerged specific weight of the particle, the particle diameter, the kinematic viscosity, and gravitational acceleration. These factors are grouped into two dimensionless quantities: the dimensionless shear stress called the Shields parameter (relating the shear stress and the submerged weight of a sediment particle), and the shear velocity Reynolds number (expressing the ratio of sediment size to the laminar sublayer thickness). The threshold of motion for granular material of uniform size depends on whether laminar or turbulent flow conditions prevail around the sediment particle; it is given in the form of a Shields diagram showing an experimentally determined relationship between both dimensionless quantities at the threshold condition. The value of the critical Shields parameter depends on whether the boundary is in a hydraulically smooth regime, a transition regime, or a completely rough regime. Typically, gravel-bed rivers fall within the completely rough regime, and sand-bed rivers fall within the transition and partially hydraulically smooth regime. Shear velocity appears in both the Shields parameter and the shear velocity Reynolds number, so it is possible to define the Shields diagram using the dimensionless particle diameter as an abscissa. For hydraulically rough turbulent flows, the critical shear stress becomes linearly proportional to sediment size.

A velocity approach equivalent to the Shields approach defines the relationship between critical flow velocities and the mean sediment size in the form of a diagram (e.g., Hjulstrom diagram) or defines the relationship between dimensionless critical average flow velocity (expressing the ratio between the average critical flow velocity at incipient motion and the sediment fall velocity in clear water) and the shear velocity Reynolds number (Yang criterion).

The probabilistic approach to incipient motion (proposed by Gessler), gives the mean condition that there is a 50% chance for a given sediment to move under specified flow conditions (turbulence fluctuations) and sediment conditions (particle position, orientation, and protrusion into the flow with respect to particles of different sizes and to bed forms).

MOVEMENT OF FLUVIAL SEDIMENTS

When flow conditions exceed the criteria for incipient motion, sediment particles on a streambed start to move. Once a sediment particle in a stream is in motion, the applied hydraulic load governs the mode of its motion. Generally, there are two modes of motion: near or on the streambed and in suspension.

Sediment that moves near or on the streambed is called bed load transport. Bed load can sometimes be divided into contact load, composed of particles rolling, sliding, or pushing, and saltating load, composed of particles bouncing, hopping, or jumping. The size of the particles that can be transported by saltation is usually correlated with the flow velocity and water density. The bed load, the amount of solid material carried on or near a streambed, usually amounts to less than 10% of the total sediment transported in large alluvial rivers. In mountain rivers and steep high-gradient streams, the major part of the total load may actually be bed load due to numerous mass-movement processes in the watershed and in-channel erosion of deposited fluvial sediments, such as in pools. Under fully developed sediment transport conditions, there is a steady exchange between bed load and suspended load.

Suspended load refers to sediment that stays in suspension for an appreciable length of time. In most natural streams, sediments are mainly transported as suspended load. Discharge-weighted total suspended solids concentrations range from practically 0 mg L^{-1} (clear water) during low flows in mountain environments to several kg L^{-1} during high flows in the form of hyperconcentrated flows, which typically occur in tropical regions and in loess or volcanic ashes.

Based on the modes of sediment transport, the total load is the sum of the bed load and the suspended load.

Based on the source of material transported, the total load can also be classified as the sum of the bed-material load and the wash load. The wash load consists of fine materials that are finer than those found in the streambed. The amount of wash load depends mainly on the supply from the watershed, not on the hydraulics of a stream. Consequently, it is difficult to predict the wash load based on the hydraulic characteristics of a stream.

An important part of a lowland stream's load may also be a dissolved load carried in solution. The proportion varies according to the climate, the chemical nature of the rocks, and the proportion of runoff contributing to the stream flow in relation to the amount of groundwater flow. The quantity of dissolved solids is given by the total concentration of dissolved material in water, measured by complete evaporation of a given quantity of water. The parameter is of use in any examination of the rate of chemical denudation and of the dissolved load of streams. Discharge-weighted total dissolved solids concentrations range from 5000 mg L^{-1} in arid saline environments to 5 mg L^{-1} in tropical rainforest regions.

TRANSPORT CAPACITY OF STREAMS

The ability of a stream to transport sediment, as measured by the maximum quantity (mass or volume) of sediment that can be carried past a specific point in the stream in a given unit of time is called transport capacity. The transport capacity increases as the water discharge becomes greater or the stream gradient becomes steeper, and decreases, as the particle size of the sediment becomes larger. Stream transport capacity is a function of bed width; for a given water discharge and gradient, the flow velocity near the streambed is lower in a wide, shallow stream than in a narrow, deep one.

The sediment discharge ratio in a stream is the ratio between the sediment discharge and the water discharge. When defining sediment discharge, the total solid transport in a stream through its cross section is taken regardless of the mode of sediment motion. The sediment-rating curve is then an empirical expression of the relationship between the stream water discharge and the stream sediment discharge at a given point. It is shown as an equation where the sediment discharge is defined as the water discharge multiplied by the mean sediment concentration and can then be written as a potential function of water discharge.

Flow competence is a term that is used in fluvial geomorphology and hydrology to indicate the ability of a stream to move particles of a particular size as bed load. It refers to the largest particle size that can be carried by a particular stream velocity. The largest particle that can be transported increases generally as a high (e.g., sixth) power of the stream velocity. The stream flow velocity at the bed is controlled by such factors as stream gradient and hydraulic radius. At equal water depth and gradient, near-bed flow velocity is higher in a wide shallow stream, where the hydraulic radius is practically the same as the average water depth, than in a narrow deep one, where, on the contrary, the hydraulic radius is much less than the average water depth. But a large, slowly moving stream may carry a large quantity of small particles in suspension, and although its (suspended) transport capacity is high, its flow competence is small. Conversely, a small but rapidly flowing stream can move relatively large particles (i.e., it has a high flow competence) although its transport (bed load and suspended load) capacity is small because of its limited water discharge.

Sediment yield is defined as the mean sediment load (total sediment discharge) carried by a stream per unit area of the catchment area; it gives some measure of the rate of erosion in a drainage basin in addition to the transport capacity of the stream itself. To establish the sediment budget of a catchment area, knowledge of sediment production in sediment source areas as well as sediment yield and sediment transport capacity of streams is important.

SEDIMENT TRANSPORT FUNCTIONS

Sediment transport is the mass transfer of solids by flowing water. Its mathematical representation in the form of a transport function needs the determination of sediment transport rates, given as the mass or weight of solids (sediment particles) transported through a cross section of a stream per unit time. Specific sediment transport rates are often given only per unit bed width to be then multiplied by the active streambed width.

Some of the classical formulas for estimating sediment transport rates were derived mainly for bed load, neglecting suspended load. In the way they approach the relation between water discharge and bed load discharge, these bed load formulas can be divided into groups that use shear stresses, energy slope, water discharge, bed forms, probability—turbulent flow fluctuations, stochastic modeling of single particle step lengths and rest periods, regression of laboratory data, and the equal mobility hypothesis as a basis for estimating bed load rates.

The starting point for suspended load equations is the vertical distribution of time-averaged velocities and sediment concentrations at given hydraulic conditions. It can be obtained from the exchange theory that states that under steady-state equilibrium conditions, the downward flux of sediment due to the fall velocity must be balanced by the net upward flux of sediment due to turbulent fluctuations. To obtain the suspended load in many such transport functions, one should know the suspended sediment concentration at some distance above the streambed.

Most total load equations are actually total bed-material load equations. In the comparison between computed and measured total bed-material load, wash load should be subtracted from the measurement before the comparison in most cases. There are two general approaches to determining total load. The first is to compute bed load and suspended load separately and then add them together to obtain total load. The second is to determine the total load function directly without dividing it into bed load and suspended load. A sediment particle may be transported as bed load at one time and as suspended load at another time or location. Except for coarse materials, which are mainly transported as bed load, total bed-material load equations should be used for determining sediment transport capacity in natural streams.

STREAM AGGRADATION AND DEGRADATION

Sediment transport is a process that interrelates erosion and sedimentation. When the rate of sediment supply from upstream is higher than a stream's sediment transport capacity, the streambed will start to aggrade at a rate defined by the difference between the rate of sediment supply and the sediment transport rate of the stream. The flow in the stream will be saturated with sediment, its transport capacity fully used. But if a stream's sediment transport capacity exceeds the rate of sediment supply from upstream, the balance of sediment load has to come from the channel itself. In this case, the channel starts to degrade.

The particle size distribution of the bed material on a streambed generally depends on the magnitude of the applied hydraulic loads in relation to the mobility of sediment particles of different sizes. In natural gravel-bed and sand-bed rivers, size nonuniformity of fluvial sediments is rather a general rule. The applied hydraulic load expressed as the applied shear stress can move none of the particles (zero sediment transport, no motion), finer and more exposed particles only (weak sediment transport, partial motion), or all particles (fully developed sediment transport, full motion). It is normally observed that without motion of finer particles, the streambed surface is composed of the original bed material. Due to motion of finer particles at a hydraulic load not sufficiently high to displace coarser particles and because of the nonuniformity of the bed-material size, finer material is transported at a faster rate than coarser material, and the remaining bed material on the bed surface becomes coarser. The bed load fractions are finer than those of the bed surface. This stage of the sediment transport is called the selective transport phase. The transition between the zero sediment transport stage and the weak sediment transport stage is smooth. This coarsening process (vertical sorting) on the streambed stops once a layer of coarse material completely covers the streambed and protects the finer materials beneath it from being transported. After this process is completed, the streambed is armored; the coarser layer is called an armor layer. Due to the variation in the flow condition of a natural stream (flood waves), usually more than one layer of armoring material is required to protect finer material beneath it from being eroded. When the applied hydraulic load is large enough to break the coarse armor layer, the finer particles are no longer shielded by stable coarser particles. All available particles of all sizes enter into motion and are found as bed load, as soon as the applied hydraulic load exceeds the threshold of motion of coarser particles. This concept is referred to as the equal-mobility concept for which all fractions of sediment enter into motion at the same value of applied shear stress. When the armor layer breaks, the transition between the weak sediment transport stage and the fully developed sediment transport stage is rather sharp.

The armoring process can be combined with incipient motion criteria to compute the extent of stream degradation. When there is not enough coarse material to develop an armor layer, degradation can be computed using the stable slope approach. This method computes the volume of eroded material in a given stream reach due to limited sediment supply. The limiting slope or final stable slope of the stream can be computed using a sediment transport equation. When there is downstream bedrock or other type of fixed control in the stream, the limiting slope will start at that point and extend in the upstream direction.

DOWNSTREAM FINING IN STREAMS

Sediment transport is definitely selective in many ways and thus is interrelated to sorting processes. Generally, sorting is defined as a process by which materials are graded according to one of their particular attributes, such as shape, size, and density. In sedimentology, the natural sorting of sediments by particle size is a basis for classification into well sorted or poorly sorted deposits. Sediments that have been repeatedly reworked by marine waves or have traveled a long distance downstream in a fluvial system are usually well sorted into different-sized particles, because different particle sizes have different settling velocities. In general, sediments transported by glaciers and mass movement are poorly sorted. Fluvial sediments tend to be well sorted.

Sorting processes in a stream can be observed in three directions: laterally (e.g., differences between coarser sediments in the thalweg and finer sediments on a bar), vertically (e.g., differences between coarser armor layer and finer sublayers), and longitudinally (downstream decrease of mean sediment size due to decreasing flow competence).

The abrasion of fluvial sediments covers attrition processes during passive and active phases of sediment transport, such as breakage, chipping, wearing, and grinding. Consequently, fluvial sediments change their particle size distribution as sediment particles are size-dependently reduced in size and their roundness generally increases. Only breakage, more common in steep high-gradient streams or bedrock reaches, will increase the angularity of sediment particles.

Downstream fining is a process in natural streams that describes an approximately exponential decrease in mean sediment size under ideal conditions. It covers sorting processes and fluvial abrasion, as well as *in situ* processes such as chemical or frost weathering of temporarily stored but exposed fluvial sediments, for example, on a bar. Generally, sorting is the prevailing process responsible for fining; especially in specific situations as when a gravel-bed river suddenly turns into a sand-bed river. Abrasion generally prevails in steep high-gradient streams, where flow competence is high and sorting is thus less pronounced. The exponential decrease is often disrupted in sedimentary links, where fresh sediments from tributaries flow laterally into a stream. When sediment transport over longer reaches of a stream is to be modeled, sorting processes and fluvial abrasion of sediment particles should therefore be accounted for.

MODELING OF SEDIMENT TRANSPORT

On one hand, the problem can be studied practically in the field by observing and measuring most relevant process parameters or in a laboratory using physical (hydraulic) modeling. On the other hand, it can be studied theoretically using numerical modeling and also using appropriate field and laboratory data for model validation.

Sediment transport can be thus defined as a natural process manifesting in sediment-laden flows, and it is usually studied as part of applied mechanics (e.g., river mechanics) or river (loose boundary) hydraulics. The main problem to be solved is to establish functional relationships that relate hydrodynamics (water movement) to sediment movement or vice versa.

For sediment-laden flow, three types of movement can be differentiated:

- the mixture may be considered a Newtonian fluid (the volumetric concentration of solids is very small, e.g., less than 1%)—typically sediment transport in streams and rivers falls into this category;
- the mixture behaves as a quasi-Newtonian fluid (the volumetric concentration of solids remains small, e.g., less than 8–10%)—typically turbidity currents and concentrated suspension fall into this category;
- the mixture behaves as a non-Newtonian fluid (the volumetric concentration of solids becomes important, e.g., more than 8–10%)—typically hyperconcentrated flows (turbidity currents and suspensions) and debris flow fall into this category.

At low sediment concentration and low deformation rates, sediment-laden flow obeys Newton's law of deformation. The governing equations of motion are the continuity and the dynamic equations. These equations can be given in different coordinate systems, such as Cartesian, cylindrical, or spherical, in one, two, or three dimensions.

Modern mathematical models are two- or three-dimensional models, which can be divided into two groups: coupled and uncoupled models. Coupled models solve the governing equations for continuity (mass-balance) and dynamics (momentum-balance) of two-phase flow simultaneously; uncoupled models solve hydrodynamic equations separately from the sedimentologic equations. An important advantage of coupled models is that they can handle intensive unsteady-state processes, such as river changes during a dam break or flash flood waves. Sediment transport numerical models can predict morphological changes and help river engineers to take appropriate measures. Contemporary numerical models use computer power extensively to account for nonuniformity of fluvial sediments when developing fractional models rather than using different mean diameters of sediment.

READING LIST

Chien, N. and Wan, Z. (1999). *Mechanics of Sediment Transport*. ASCE Press, Reston, VA.

Graf, W. (1998). *Fluvial Hydraulics—Flow and Transport Processes in Channels of Simple Geometry*. John Wiley, Chichester, UK.

Raudkivi, A.J. (1998). *Loose Boundary Hydraulics*. A. A. Balkema, Rotterdam, CO.

Simons, D.B. and Sentürk, F. (1992). *Sediment Transport Technology—Water and Sediment Dynamics*. Water Resources Publications, Littleton, CO.

Van Rijn, L.C. (1993). *Principles of Sediment Transport in Rivers, Estuaries and Coastal Seas*. Aqua Publications, Amsterdam.

Wohl, E. (2000). *Mountain Rivers, Water Resources Monograph 14*. American Geophysical Union, Washington, DC.

Yalin, M.S. (1992). *River Mechanics*. Pergamon Press, Oxford, UK.

Yang, C.T. (1996). *Sediment Transport—Theory and Practice*. McGraw-Hill, New York.

STOCHASTIC SIMULATION OF HYDROSYSTEMS

Demetris Koutsoyiannis
National Technical University
Athens, Greece

INTRODUCTION

Simulation is defined as a technique for imitating the evolution of a real system by studying a model of the system. The model is an abstraction, a simplified and convenient mathematical representation of the actual system, typically coded and run as a computer program. If the model has a stochastic[1] element, then we have *stochastic simulation*. The term stochastic simulation sometimes is used synonymously with the *Monte-Carlo method*.

Stochastic simulation is regarded as mathematical experimentation and is appropriate for complex systems, whose study based on analytical methods is laborious or even impossible. For such systems, stochastic simulation provides an easy means to explore their behavior by answering specific 'what if' questions. Moreover, stochastic simulation can be viewed as a numerical method for solving mathematical problems in several fields such as statistical inference, optimization, integration, and even equation solving. Under certain conditions, stochastic simulation is more powerful than other more common numerical methods (e.g., numerical integration of high-dimensional differential equations).

[1] The term *stochastic* is etymologized from the Greek verb *στοχάζομαι* initially meaning to aim, point, or shoot (an arrow) at a target (*στόχος* is a target). Metaphorically, the verb meant to guess or to conjecture (the target). The modern Greek meaning is to imagine, to think, to meditate. It appears that the word *stochastic* is found in English since the seventeenth century, and the obsolete meaning pertains to conjecture. Its use as a scientific term is attributed to the Swiss mathematician Jakob Bernoulli (1654–1705), who in his famous masterpiece *Ars Conjectandi* realized that randomness and uncertainty are important aspects of our world and should be objects of scientific analysis. In its modern sense, *stochastic* can be regarded as synonymous with random or probabilistic, but it is mostly used for processes that indicate a mixture of structure and randomness; the term *stochastic process* was used in 1932 by A. N. Kolmogorov.

Due to their complexity, hydrosystems, including water resource systems, flood management systems, and hydropower systems, are frequently studied using stochastic simulation. A generalized solution procedure for hydrosystems problems, including systems identification, modeling and forecasting, hydrologic design, water resources management, and flood management, is discussed. Emphasis is given on the stochastic representation of hydrologic processes, which have a dominant role in hydrosystems. Peculiarities of hydrologic and other geophysical processes (seasonality, long-term persistence, intermittency, skewness, spatial variability) gave rise to substantial research that resulted in numerous stochastic tools appropriate for applications in hydrosystems. Four examples of such tools are discussed: (1) the multivariate periodic autoregressive model of order 1 [PAR(1)], which reproduces seasonality and skewness but not long-term persistence; (2) a generalized multivariate stationary model that reproduces all kinds of persistence and simultaneously skewness but not seasonality; (3) a combination of the previous two cases in a multivariate disaggregation framework that can respect almost all peculiarities except intermittency; and (4) the Bartlett-Lewis process that is appropriate for modeling rainfall and emphasizes its intermittent character on a fine time scale.

A BRIEF HISTORY

Synthetic stream-flow records were first used early in the twentieth century by Hazen (1) in studies of water supply reliability. Their construction, however, was not based on the theory of stochastic processes, then not developed, but on merging and rescaling observed records of several streams. This early work emphasizes the need for long synthetic records and the importance of simulation in water resources technology. The foundation of stochastic hydrology followed the significant developments in mathematics and physics in the 1940s, as well as the development of computers. Specifically, it followed the establishment of the Monte Carlo method, which was invented by Stanislaw Ulam in 1946. Notably, Ulam conceived the method while playing solitaire during convalescence from an illness, in an attempt to estimate the probabilities of success of the plays. As Ulam describes the story in some remarks later published by Eckhardt (2), "After spending a lot of time to estimate them by pure combinatorial calculations, I wondered whether a more practical method than 'abstract thinking' might not be to lay it out say one hundred times and simply observe and count the number of successful plays." Soon after, the method grew to solve neutron diffusion problems by himself and other great mathematicians and physicists in Los Alamos (John von Neumann, Nicholas Metropolis, and Enrico Fermi), and was first implemented on the ENIAC computer (2,3). The 'official' history of the method began in 1949 when a paper was published by Metropolis and Ulam (4).

In the field of water resources, the most significant initial steps were the works by Barnes (5) for generating uncorrelated annual flows at a site from normal distribution, Maass et al. (6) and Thomas and Fiering (7) for generating flows correlated in time, and Beard (8) and Matalas (9) for generating concurrent flows at several sites.

The classic book on time series analysis by Box and Jenkins (10) also originated from different, more fundamental scientific fields. However, it has subsequently become very popular in stochastic hydrology. Box and Jenkins developed a classification scheme for a large family of time series models. Their classification distinguishes among autoregressive models of order p [AR(p)]; moving average models of order q [MA(q)]; combinations of the two, called autoregressive-moving average [ARMA(p, q)] models; and autoregressive integrated moving average [ARIMA(p, d, q)] models. However, despite a large family, Box-Jenkins models do not fully cover the needs of hydrologic modeling, as they do not comply with some peculiarities of hydrologic and other geophysical processes. This gave rise to substantial research that resulted in numerous stochastic tools appropriate for application to water resources.

UTILITY OF STOCHASTIC SIMULATION IN HYDROSYSTEMS

Due to the significant uncertainties inherent in hydrosystems, among which the major uncertainty is hydrologic uncertainty (related to the unknown future of inflows to hydrosystems), an estimate of a system's reliability is important for its design and operation. The reliability of a system is defined as the probability that a system will perform the required function for a specified period of time under stated conditions (11, p. 434). Reliability is the complement of the probability of failure or risk, the probability that "loading" will exceed "capacity." In many instances, the risk can be estimated by analytical means, so stochastic simulation is not required. For example, in the design of dikes that confine a river's flow, the risk of overtopping dikes can be estimated in a typical probabilistic manner, provided that a long enough record of floods of the river exists (some decades). The estimating procedure includes selecting a probabilistic model (e.g., an extreme value distribution function), fitting the model based on the available record, and estimating the probability that a flood will exceed the discharge capacity of the designed river cross section (the estimate of the latter is a matter of hydraulics). Behind this procedure, there are two implicit assumptions that make the methodology appropriate for this problem:

1. The project under study (the dikes) does not modify the natural flow regime, so that if the project had been constructed many years before, the observed flow record would not be altered. Thus, the assumed probabilistic model, although fitted to past data, is still valid after construction of the project.
2. The quantity (flood discharge) whose exceedance was assumed to be the risk is the same quantity, for which we have observed data. Thus, the probabilistic model that was constructed for this quantity can yield the risk directly.

In many cases, however, these assumptions are not valid. Let us first examine the case where assumption 2 is untrue. For example, we may have available rainfall data, from which we can construct a probabilistic model for extreme rainfall intensity, and wish to estimate the probability of exceeding the flood discharge. In this case, we can use a simple one-to-one mapping (transformation) of rainfall to discharge values (e.g., to adopt the relation known as the rational formula), so that the risk of exceeding a certain discharge level equals the risk of exceeding the corresponding rainfall level. This methodology usually incorporates serious oversimplifications and ignorance of certain factors that affect the actual hydrologic process (e.g., retention and infiltration). A more realistic methodology is to use a more detailed model that transforms a rainfall series (not each isolated value) into a discharge series, also considering all processes involved in this transformation. In this case, we can use simulation to obtain a discharge series.

Assumption 1 can be untrue in many cases as well. For example, constructing a dam will alter the flood regime at the dam and downstream, as the spillway outflow does not equal the natural inflow (attenuation occurs due to temporal flood storage). The construction of a storm sewer will also modify the contributing areas and flow times in the area (in addition, it is impossible to have observed data for the sewer discharge in its design phase). Another typical example is a reservoir (see the entry RELIABILITY CONCEPTS IN RESERVOIR DESIGN), whose storage (a quantity that determines the risk, which is the probability of emptying the reservoir) did not exist before the construction of the reservoir. Obviously, in all these cases where assumption 1 is not valid, assumption 2 is also not valid. Thus, we will proceed as in the previous paragraph where simulation is the most appropriate procedure for obtaining a series of data values for the quantity of interest.

Even in an existing project (e.g., an existing reservoir), where the quantity of interest can be measured directly to obtain a historical record, simulation may be necessary again to assess the impacts of several possible changes in the future that were not experienced in the past. For example, a change in water use (e.g., an increase in water demand) and a change in land use or climate, which alters water availability, calls for simulation to estimate a series for the quantity of interest in the scenario examined.

The previous discussion explains why in most studies of hydrosystems (except cases where both assumptions listed previously are valid), it is necessary to simulate to transform some input time series of initial quantities into some output time series of the final quantities of interest. By grace of the power of computers, the simulation methodology has greatly replaced older methodologies that used simplified one-to-one transformations. But why should simulation be stochastic?

In stochastic simulation, the input time series are no longer the observed records but synthetic time series constructed by an appropriate stochastic model. An observed time series is unique and has a limited length equal to the period of observations. On the contrary, a stochastic model can produce as many time series as required and of any arbitrary length. The utility of a long time series becomes obvious in steady-state simulations (12, p. 1220), when estimating a low value of the probability of failure (risk). For example, in a problem where the accepted probability of failure is 1% per year, apparently several hundreds of simulated years are needed to detect a few failures. The utility of ensemble time series (as opposed to the unique observed record) becomes obvious in nonsteady-state problems (i.e., in terminating simulations) and in forecast problems in which the initial conditions (present and past values of the processes of interest) are known. In these cases, stochastic simulation offers the possibility of different sample paths of the quantity of interest, instead of having a single value at a time, so that we can estimate expected values and confidence zones.

COMPONENTS AND SOLUTION PROCEDURE OF STOCHASTIC SIMULATION

The components and the steps followed in the stochastic simulation of a hydrosystem are shown in Fig. 1. The entire procedure includes two main model components (marked 1 and 2 in Fig. 1) and two simpler procedures (marked 3 and 4 in Fig. 1). The first model component is the stochastic model of inputs, which produces a vector $\mathbf{X}(\boldsymbol{\mu}, \omega)$ of hydrological inputs (e.g., a time series of rainfall, evaporation, and river flow, depending on the problem studied) to the hydrosystem, where $\boldsymbol{\mu}$ is a vector that contains the parameters of hydrologic inputs (all estimated from the available records of observations) and ω denotes a sample path realization of the random variables (i.e., ω can be thought of as representing the randomness in the system, e.g., all random numbers in a simulation run). At a minimal configuration, the vector of parameters $\boldsymbol{\mu}$ includes mean values, standard deviations, autocorrelations (at least for lag one), and cross-correlations (for multiple-site models).

The second component is the transformation model which takes the inputs $\mathbf{X}(\boldsymbol{\mu}, \omega)$ and produces the outputs $\mathbf{Z}(\mathbf{X}(\boldsymbol{\mu}, \omega), \boldsymbol{\lambda})$ (e.g., river flow, if $\mathbf{X}$ is rainfall and evaporation, or reservoir release and storage, if $\mathbf{X}$ is river flow); here the vector $\boldsymbol{\lambda}$ contains parameters of the transformation model (e.g., parameters that determine the hydrologic cycle in a basin and/or parameters that determine the operation of a specific project such as a reservoir).

The third component is a procedure that takes the outputs $\mathbf{Z}(\mathbf{X}(\boldsymbol{\mu}, \omega), \boldsymbol{\lambda})$ and determines a sample performance measure $L((\mathbf{Z}(\mathbf{X}(\boldsymbol{\mu}, \omega), \boldsymbol{\lambda}))$ of the system that corresponds to the sample realization represented by ω. This performance measure depends on the problem examined; for instance, in a flood design problem, it can be the risk of exceeding a specified flood level; in a reservoir design problem, it can be either the risk of emptying a reservoir or the attained release for a stated reliability.

By repeating the execution of these three components using different simulation runs, represented by different ω, we can obtain an ensemble of simulations and a sample of performance measures, from which we can estimate the true (independent of ω) performance measure of the

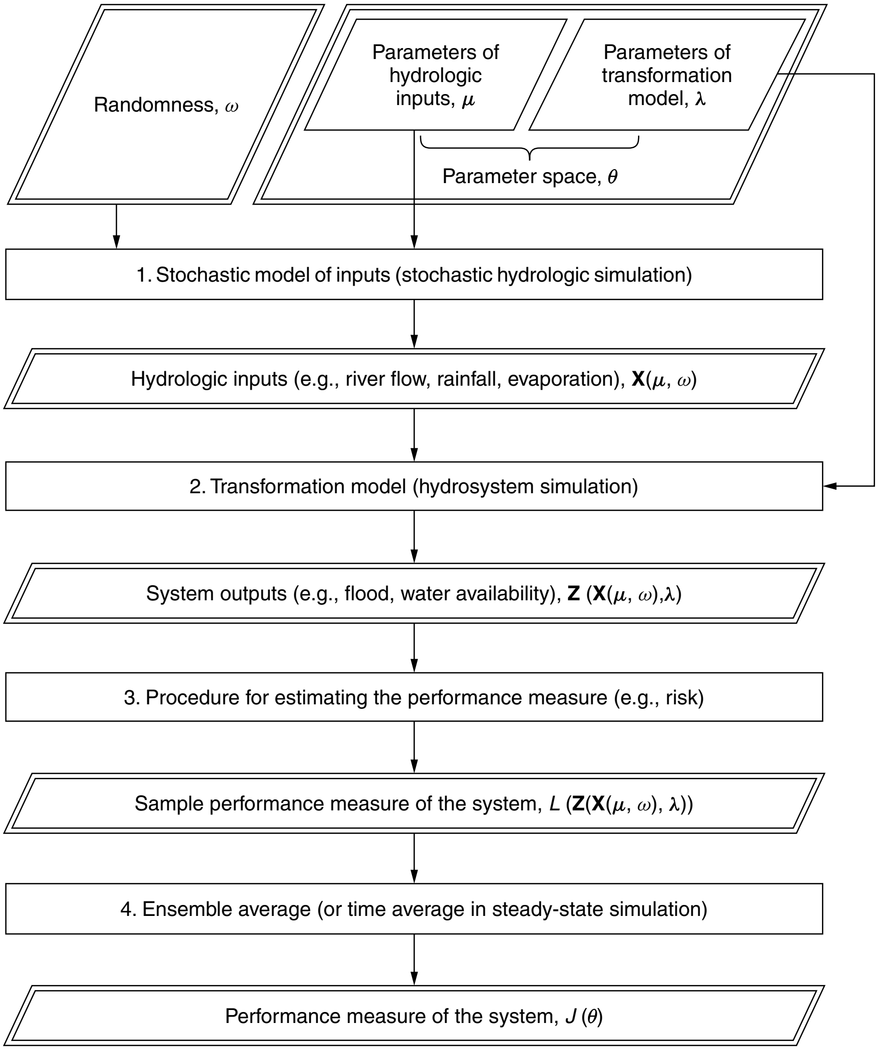

Figure 1. Schematic representation of the components and the solution procedure in hydrosystem simulation. Rectangles represent the components of the solution procedure, and parallelograms represent inputs and outputs of the different components.

system $J(\boldsymbol{\theta}) := E[L(\mathbf{Z}(\mathbf{X}(\boldsymbol{\mu}, \omega), \boldsymbol{\lambda}))]$, where $E[\ ]$ denotes the expected value and $\boldsymbol{\theta} := (\boldsymbol{\mu}, \boldsymbol{\lambda})$. However, if the system is stationary and ergodic (in other words, if we have a steady-state simulation), L will tend to J, as the simulation length tends to infinity. Therefore, a single instance of the sample performance measure, estimated from a simulation of large length, is an adequate estimate of the true performance measure. This is the case, for instance, in a reservoir simulation of constant water demand. Conversely, if the water demand is growing in time (a common situation in practice), the simulation is no longer steady-state, and numerous runs, typically of a short length, must be performed to estimate the true performance measure.

In the following part of the article, we will focus on the model component 1, the stochastic model of inputs.

TYPICAL BOX–JENKINS MODELS

Let X_i denote the process of interest (e.g., rainfall or streamflow at a site) where i denotes discrete time. To generate a time series of X_i, we start generating a sequence of independent identically distributed variables (iid, also known as white noise) V_i that have a specified distribution function (e.g., Gaussian). This is known as generation of random numbers; a concise introduction to this topic can be found in Papoulis (13) and a more detailed presentation can be found in Ripley (14). If it can be assumed that X_i is stationary (i.e., it has a probability distribution function that does not vary in time, which is typically the case if the time step is a year), the unstructured sequence of V_i can be converted to a structured sequence X_i by means of a recursive relationship, whose general form was studied by Box and Jenkins (10). From the large family of Box-Jenkins processes, those that have been widely used in stochastic hydrology are special cases of the equation,

$$X_i = aX_{i-1} + a'X_{i-2} + bV_i + b'V_{i-1} \tag{1}$$

where a, a', b, and b' are parameters that are estimated from the autocovariance properties of the process X_i. The special cases are

1. the iid or white noise or AR(0) process in which $a = a' = b' = 0; b \neq 0$;
2. the Markovian or AR(1) process in which $a' = b' = 0; a, b \neq 0$;

3. the AR(2) process in which $b' = 0; a, b, a' \neq 0$;
4. the ARMA(1, 1) process in which $a' = 0; a, b, b' \neq 0$.

The complete form of (1), known as the ARMA(2, 1) process, has been not used so frequently in hydrology. The special cases listed preserve autocovariance properties for lag 0 (case 1) to lag 2 (cases 2 and 3); beyond these lags, the model autocovariance is zero (case 1) or tends to zero exponentially (cases 2–4).

PECULIARITIES IN STOCHASTIC REPRESENTATION OF HYDROLOGIC PROCESSES

Simple stationary models, such as previously described, are often not the best choice in hydrologic modeling because of several peculiarities of hydrologic processes; the most important are discussed here.

Seasonality. When the timescale of interest is finer than annual, hydrologic processes cannot be regarded as stationary because of the effect of the season of the year on the properties of the process. A simple method often used to take into account seasonality is to standardize the process X_i using seasonal values of mean and standard deviation by setting $Y_i := (X_i - \mu_i)/\sigma_i$ and assuming that Y_i is a stationary process that can be modeled, for instance, by (1); here μ_i and σ_i denote the mean and standard deviation, respectively, of X_i, which, it is assumed, vary with i in a periodic manner. This standardization approach, however, is flawed; the stationarity assumption for Y_i implies that, apart from the mean and standard deviation, other statistical properties of X_i like autocorrelation and skewness do not vary in season, which is not true. A more precise way of modeling seasonality is to assume a cyclostationary (also known as periodic) process, that is expressed as in (1), but has parameters $a, b\ a', b', \ldots$, and statistics of the noise variables V_i that vary with i in a periodic manner.

Long-Term Persistence. Box–Jenkins models such as (1) are essentially of the short memory type; their autocorrelation structure decreases rapidly with lag time. However, the study of long historical time series of hydrologic and other geophysical processes has revealed that autocorrelations may be significant for large lags of 50 or 100 years. This property is related to the tendency of stream flows to stay above or below their mean for long periods, observed for the first time by Hurst (15), or, equivalently, to multiple timescale fluctuations of hydrologic processes (16). Therefore, models such as (1) are proven inadequate in stochastic hydrology, because it is important that the long-term persistence of hydrologic processes is reproduced (see the entry RELIABILITY CONCEPTS IN RESERVOIR DESIGN).

Intermittency. On fine timescales, some hydrologic processes such as rainfall and in some cases stream flow appear as intermittent processes; thus, rainfall alternates between two states, dry (zero rainfall) and wet (positive rainfall). This is manifested in the marginal probability distribution of rainfall depth by a discontinuity at zero. Box–Jenkins processes such as that in (1) need to be truncated to represent this discontinuity; this is not so easy, nor common. To model intermittency, alternative two-state processes, such as two-state Markov chains (17, p. 302) and point process models (18) have been proposed.

Skewness. Another peculiarity of hydrologic processes is the skewed distribution functions observed mostly on fine and intermediate timescales. This is not so common in other scientific fields whose processes are typically Gaussian. Therefore, attempts have been made to adapt standard models to enable treatment of skewness (19–22). Skewness is caused mainly by the fact that hydrologic variables are nonnegative and sometimes intermittent. Therefore, successful modeling of skewness indirectly contributes to avoiding negative values of simulated variables; however, it does not eliminate the problem, and some ad hoc techniques (such as truncating negative values) are often used in addition to modeling skewness.

Spatial Variation. Hydrologic processes evolve in both time and space. Typically, time series models consider only temporal evolution. The most precise mathematical representation of hydrologic processes can be achieved by extending the indexing set of the process from one dimension (representing time) to three dimensions (one for time and two for space). However, multidimensional modeling is not easy and has been implemented only in a few cases (for example in continuous time and space modeling of rainfall; 23). A midway solution, which is the most common in stochastic hydrology, is to use multivariate models, that describe the temporal evolution of the process simultaneously at a number of points. The same method can be used directly to model more than one cross-correlated hydrologic process (e.g., rainfall and runoff) at the same location simultaneously.

In the following sections, we give some characteristic examples of models that respect these peculiarities and together can deal with a large spectrum of problems in stochastic hydrologic simulation:

1. The multivariate periodic autoregressive model of order 1 [PAR(1)], which reproduces seasonality and skewness but not long-term persistence;
2. A generalized multivariate stationary model that reproduces all kinds of persistence and simultaneously skewness but not seasonality;
3. A combination of the previous two cases in a multivariate disaggregation framework that can respect almost all the previously listed peculiarities, except intermittency which such types of models may not easily handle;
4. The Bartlett–Lewis process that is appropriate for modeling rainfall and emphasizes its intermittent character on a fine timescale, but only on a single point basis.

The Multivariate PAR(1) Model

Let $\mathbf{X}_s := [X_s^1, X_s^2, \ldots, X_s^n]^T$ represent a hydrologic process on a subannual (e.g., monthly) timescale (δ) and at n locations (the subscript T denotes the transpose of a vector or matrix). The PAR(1) model is similar to the AR(1) model but has periodically varying parameters. In a multivariate setting, it is expressed by

$$\mathbf{X}_s = \mathbf{a}_s\mathbf{X}_{s-1} + \mathbf{b}_s\mathbf{V}_s \tag{2}$$

where $\mathbf{a}_s$ and $\mathbf{b}_s$ are $(n \times n)$ matrices of parameters and $\mathbf{V}_s$ is a vector of innovations (independent, random variables in both time and location) of size n. The time index s can take any integral value, but the parameters $\mathbf{a}_s$ and $\mathbf{b}_s$ are periodic functions of s whose period $k := 1$ year$/\delta$ (e.g., 12 if δ is one month). This model can reproduce the following set of statistics:

1. the mean values, i.e., the k vectors $\boldsymbol{\mu}_s := E[\mathbf{X}_s]$ each of size n;
2. the variances and lag-zero cross-covariances among different locations, i.e., the k matrices $\boldsymbol{\sigma}_{ss} := \mathrm{Cov}[\mathbf{X}_s, \mathbf{X}_s]$ $= E[(\mathbf{X}_s - \boldsymbol{\mu}_s)(\mathbf{X}_s - \boldsymbol{\mu}_s)^T]$ (where Cov[ ] denotes covariance), each of size $(n \times n)$;
3. the lag-1 autocovariances at each location, i.e., the k vectors $\boldsymbol{\gamma}_{s,1} := [\gamma_{s,1}^1, \ldots, \gamma_{s,1}^n]^T$, where $\gamma_{s,\tau}^l := \mathrm{Cov}[X_s^l, X_{s-\tau}^l] = E[(X_s^l - \mu_s^l)(X_{s-\tau}^l - \mu_{s-\tau}^l)]$, each of size n (notice the notational identity $\gamma_{s,0}^l \equiv \sigma_{ss}^{ll}$);
4. the third moments, the k vectors $\boldsymbol{\xi}_s = \mu_3[\mathbf{X}_s] = E[(X_s^l - \mu_s^l)^3, l = 1, \ldots, n]^T$ each of size n (where $\mu_3[\,]$ denotes the third central moment of a random variable or random vector).

The model parameters $\mathbf{a}_s$ and $\mathbf{b}_s$ are typically determined by the moment estimators that are

$$\mathbf{a}_s = \mathrm{diag}(\gamma_{s,1}^l/\gamma_{s-1,0}^l, l = 1, \ldots, n) \tag{3}$$

$$\mathbf{b}_s\mathbf{b}_s^T = \boldsymbol{\sigma}_{ss} - \mathbf{a}_s\boldsymbol{\sigma}_{s-1,s-1}\mathbf{a}_s \tag{4}$$

These equations are extensions of the seasonal model of those for the stationary Markov model given by Matalas and Wallis (19, p. 63). In an alternative estimate, a full (rather than a diagonal) matrix $\mathbf{a}_s$ can be derived, which enables preserving the lag-1 cross-covariances among different locations. However, the more parsimonious formulation in Eq. 3 is sufficient for most cases. The calculation of $\mathbf{b}_s$, given the product $\mathbf{b}_s\mathbf{b}_s^T$ from Eq. 4, is not a trivial issue. A generalized methodology to do this operation, also known as the extraction of the square root of a matrix, was proposed by Koutsoyiannis (21). Another group of model parameters are the moments of the auxiliary variables $\mathbf{V}_s$. The first moments (means) are obtained by

$$E[\mathbf{V}_s] = \mathbf{b}_s^{-1}(\boldsymbol{\mu}_s - \mathbf{a}_s\boldsymbol{\mu}_{s-1}) \tag{5}$$

By definition 1, the variances are $\mathrm{Var}[\mathbf{V}_s] = [1, \ldots, 1]^T$, and the third moments are obtained by

$$\mu_3[\mathbf{V}_s] = \left(\mathbf{b}_s^{(3)}\right)^{-1}(\xi_s - \mathbf{a}_s^{(3)}\xi_{s-1}) \tag{6}$$

where $\mathbf{a}_s^{(3)}$ and $\mathbf{b}_s^{(3)}$ denote the matrices whose elements are the cubes of $\mathbf{a}_s$ and $\mathbf{b}_s$, respectively.

A GENERALIZED MULTIVARIATE STATIONARY MODEL RESPECTING LONG-TERM PERSISTENCE

The most difficult and often the most important task in simulating hydrologic processes on an annual scale is to reproduce long-term persistence. The Box–Jenkins processes are inappropriate for this purpose. Other types of models such as fractional Gaussian noise (FGN) models and broken line models (whose comprehensive discussion can be found in Ref. 24) have several weak points such as parameter estimation problems, narrow type of autocorrelation functions that they can preserve, and their inability to reproduce skewness and simultaneously to perform in multivariate problems. In a recent paper (22), all these problems were remedied; the proposed generalized methodology can perform in multivariate problems for all categories of statistics listed in points 1–4 of the previous section and, in addition, for the autocovariances at all locations for any lag r.

The setting of the method is stationary, rather than cyclostationary, so all statistics and parameters are not functions of time, which is reflected in the notation used. For example, the autocovariance for lag τ is denoted as $\boldsymbol{\gamma}_\tau := [\gamma_\tau^1, \ldots, \gamma_\tau^n]^T$ where $\gamma_\tau^l := \mathrm{Cov}[X_i^l, X_{i-\tau}^l]$. Remember that long-term persistence implies nonignorable autocovariances for high lags (e.g., of the order $10^2 - 10^3$). Such autocovariances can be described by a power-type (as opposed to the exponential type of ARMA processes) function such as

$$\gamma_\tau^l = \gamma_0^l(1 + \kappa^l\beta^l\tau)^{-1/\beta^l} \tag{7}$$

where κ^l and β^l are constants. This generalized autocovariance structure (GAS) incorporates as special cases the exponential ARMA type structure (for $\beta = 0$) and the FGN structure (for a special combination of κ^l and β^l; see Ref. 22). The constants κ^l and β^l can be estimated by fitting Eq. 7 to the sample autocovariance estimates; note that Eq. 7 can be used for lags beyond a certain lag τ_0, thus allowing the possibility of specifying different values (i.e., the historical values of the sample autocovariance estimates) for smaller lags.

In each of the locations, the process X_i^l can be expressed in terms of some auxiliary variables V_i^l, uncorrelated in time i (i.e., $\mathrm{Cov}[V_i^l, V_m^k] = 0$ if $i \neq m$), but correlated in different locations l for the same time i by using

$$X_i^l = \sum_{r=-q}^{q} a_{|r|}^l V_{i+r}^l \tag{8}$$

This equation defines the so-called symmetrical moving average (SMA) scheme. Like the conventional (backward) moving average (MA) process, the SMA scheme transforms a sequence of temporally uncorrelated variables V_i^l into a process with autocorrelation by taking the weighted average of a number of V_i^l. In the SMA process, the weights a_r^l are symmetrical about a center (a_0^l) that corresponds to the variable V_i^l. The number of variables V_i^l that define X_i^l is $2q + 1$, where q is theoretically infinity but in practice can be restricted to a finite number, as the sequence of weights a_r^l tends to zero for increasing r. Koutsoyiannis (22) showed that the discrete Fourier transform $s_a^l(\omega)$ of the a_r^l sequence is related to the power spectrum $s_\gamma^l(\omega)$ of the process (i.e., the discrete Fourier transform of the sequence of γ_τ^l) by

$$s_a^l(\omega) = \sqrt{2 s_\gamma^l(\omega)} \tag{9}$$

This enables easy and fast (using the fast Fourier transform) computation of the sequence of a_r^l, even if the terms of the sequence are thousands. The computation includes transforming the sequence of γ_τ^l to $s_\gamma^l(\omega)$, calculating $s_a^l(\omega)$ from Eq. 9, and inversely transforming $s_a^l(\omega)$ to the sequence of a_r^l.

The auxiliary variables V_i^l, by definition, have unit variances, means $E[V_i^l]$, and third central moments $\mu_3[V_i^l]$ given by

$$\left(a_0 + 2\sum_{j=1}^{s} a_j\right) E[V_i^l] = \mu^l, \quad \left(a_0^3 + 2\sum_{j=1}^{q} a_j^3\right) \mu_3[V_i^l] = \xi^l \tag{10}$$

Their variance–covariance matrix $\mathbf{c} := \mathrm{Cov}[\mathbf{V}_i, \mathbf{V}_i]$ has elements c^{lk} that can be expressed in terms of σ^{lk} (the elements of the variance–covariance matrix $\boldsymbol{\sigma}$ of $\mathbf{X}_i$), and the sequences a_i^l and a_i^k by

$$c^{lk} = \sigma^{lk} / \sum_{r=-q}^{q} a_{|r|}^l a_{|r|}^k \tag{11}$$

Given the matrix $\mathbf{c}$, the vector of variables $\mathbf{V}_i = [V_i^1, V_i^2, \ldots, V_i^n]^T$ can be generated using the simple multivariate model,

$$\mathbf{V}_i = \mathbf{b}\mathbf{W}_i \tag{12}$$

where $\mathbf{W}_i = [W_i^1, W_i^2, \ldots, W_i^n]^T$ is a vector of innovations with unit variance, independent both in time i and in location $l = 1, \ldots, n$, and $\mathbf{b}$ is a matrix of size $n \times n$ such that

$$\mathbf{b}\,\mathbf{b}^T = \mathbf{c} \tag{13}$$

The other parameters needed to define model Eq. 12 completely are the vector of mean values $E[\mathbf{W}]$ and third moments $\mu_3[\mathbf{W}]$ of W_i^1. These can be calculated in terms of the corresponding vectors of V_i^1, already known from Eq. 10, by

$$E[\mathbf{W}] = \mathbf{b}^{-1} E[\mathbf{V}], \quad \mu_3[\mathbf{W}] = (\mathbf{b}^{(3)})^{-1} \mu_3[\mathbf{V}] \tag{14}$$

STOCHASTIC DISAGGREGATION TECHNIQUES

Seasonal models that can reproduce the long-term persistence of hydrologic processes do not exist at present. If the timescale of interest is finer than annual and, simultaneously, respecting of long-term persistence is important, a two-scale approach is followed. A stationary stochastic model like that described in the previous section is used to generate the annual time series. These are then disaggregated into a finer timescale so that the periodicity and short-term memory of the process of interest are respected. Traditionally, the latter task has been tackled by the so-called disaggregation models, which were initially proposed by Valencia and Schaake (25) and improved since then by the contribution of several researchers (for an outline of such contributions, see Refs. 26 and 27). These are purposely designed models to generate a process at the finer timescale given than at the coarser one. Specifically, they do not model the process of interest in the lower level timescale itself, but rather they are hybrid schemes simultaneously using both timescales. Sometimes (owing to nonlinear transformations of variables) these models cannot ensure consistency with the higher level process. Then, adjusting procedures are necessary to restore consistency (26,28).

A different approach was recently proposed by Koutsoyiannis (29), which is a generalized framework for coupling stochastic models on different timescales. This approach couples two independent stochastic models appropriate, respectively, for the coarser (annual) and finer (e.g., monthly) scales using a transformation that modifies the output of the latter to become consistent with the series produced by the former model. To demonstrate this approach, we will assume that the coarse scale model is the multivariate SMA model described in the previous section, which produces annual series $\mathbf{Z}_i$, and the finer scale model is the multivariate PAR(1) model described two sections before, which produces monthly series $\mathbf{X}_s$. Consistency of the two series requires that they obey

$$\sum_{s=(i-1)k+1}^{ik} \mathbf{X}_s = \mathbf{Z}_i, \tag{15}$$

where k is the number of fine-scale time steps within each coarse-scale time step ($k = 12$ in our example). The annual series $\mathbf{Z}_i$ is generated first. The finer scale model is run independently of the coarser scale one, without any reference to the known $\mathbf{Z}_i$, and produces monthly series $\tilde{\mathbf{X}}_s$. If we aggregate the latter on the annual scale (by means of Eq. 15), we will obtain some annual series $\tilde{\mathbf{Z}}_i$, which will apparently differ from $\mathbf{Z}_i$. In a subsequent step, we modify $\tilde{\mathbf{X}}_s$ thus producing $\mathbf{X}_s$ consistent with $\mathbf{Z}_i$ (in the sense that they obey Eq. 15) without affecting the stochastic structure that characterizes $\tilde{\mathbf{X}}_s$. For this modification, we use a linear transformation $\mathbf{X}_s = \mathbf{f}(\tilde{\mathbf{X}}_s, \tilde{\mathbf{Z}}_i, \tilde{\mathbf{Z}}_i)$ which has been termed the coupling transformation. This is given by (29)

$$\mathbf{X}_i^* = \tilde{\mathbf{X}}_i^* + \mathbf{h}(\mathbf{Z}_i^* - \tilde{\mathbf{Z}}_i^*) \tag{16}$$

where

$$\mathbf{X}_i^* := [\mathbf{X}_{(i-1)k+1,\ldots,}^T \mathbf{X}_{ik}^T]^T \tag{17}$$

$$\mathbf{Z}_i^* := [\mathbf{Z}_i^T, \mathbf{Z}_{i+1}^T, \mathbf{X}_{(i-1)k}^T]^T \tag{18}$$

$$\mathbf{h} = \mathrm{Cov}[\mathbf{X}_i^*, \mathbf{Z}_i^*]\{\mathrm{Cov}[\mathbf{Z}_i^* \mathbf{Z}_i^*]\}^{-1} \tag{19}$$

and $\tilde{\mathbf{X}}_i^*$ and $\tilde{\mathbf{Z}}_i^*$ are defined in terms of $\tilde{\mathbf{X}}_s$ and $\tilde{\mathbf{Z}}_i$ in a manner identical to that of the definition of $\mathbf{X}_i^*$ and $\mathbf{Z}_i^*$.

It is clarified that the vector $\mathbf{X}_i^*$ contains the monthly values of all 12 months of year i for all examined locations (e.g., for five locations, $\mathbf{X}_i^*$ contains $12 \times 5 = 60$ variables) and the vector $\mathbf{Z}_i^*$ contains (1) the annual values of the current year; (2) the annual values of the next year; and (3) the monthly values of the last month of the previous year (e.g., for five locations, $\mathbf{Z}_i^*$ contains $3 \times 5 = 15$ variables). Items (2) and (3) of $\mathbf{Z}_i^*$ are included to ensure that the transformation will preserve the covariance properties among the monthly values of each year and also the covariances for the previous and next years as well. Note that at the stage of the generation at year i, the monthly values of year $i-1$ are known (therefore, in $\mathbf{Z}_i^*$, we enter monthly values of the year $i-1$), but the monthly values of year $i+1$ are not known (therefore, in $\mathbf{Z}_i^*$, we enter annual values of the year $i+1$, which are known).

The quantity $\mathbf{h}(\mathbf{Z}_i^* - \tilde{\mathbf{Z}}_i^*)$ in Eq. 16 represents the correction applied to $\tilde{\mathbf{X}}$ to obtain $\mathbf{X}$. Whatever the value of this correction, the coupling transformation will ensure preservation of first- and second-order properties of variables (means and variance-covariance matrix) and linear relationships among them (in our case the additive property, Eq. 15). However, it is desirable to have this correction as small as possible so that the transformation does not seriously affect other properties of the simulated processes (e.g., the skewness). It is possible to make the correction small enough, if we keep repeating the generation process for the variables of each period (rather than performing a only single generation) until a measure of the correction becomes lower than an accepted limit. This measure can be defined as

$$\Delta = (1/m)||\mathbf{Z}_i^{\prime *} - \tilde{\mathbf{Z}}_i^{*}|| \qquad (20)$$

where $\mathbf{Z}_i^{\prime *}$ and $\tilde{\mathbf{Z}}_i^{\prime *}$ are respectively, $\mathbf{Z}_i^*$ and $\tilde{\mathbf{Z}}_i^*$ standardized by the standard deviation (i.e., $Z_i^{\prime *l} := Z_i^{*l}/\{\mathrm{Var}[Z_i^{*l}]\}^{1/2}$), m is the common size of $\mathbf{Z}_i^{\prime *}$, and $\tilde{\mathbf{Z}}_i^{\prime *}$, and $|| \cdot ||$ denotes the Euclidian norm.

POINT PROCESS MODELS

On even finer timescales such as daily or hourly, the intermittency of hydrologic processes dominates, and stochastic models like those described earlier can hardly describe it. Point process models have been the most widespread approach to representing intermittent hydrologic processes, particularly, rainfall. As a representative example, we summarize here the rainfall model based on the Bartlett–Lewis process; this was chosen due to its wide applicability and experience in calibrating and applying it to several climates. Accumulated evidence of its ability to reproduce important features of the rainfall field from the hourly to the daily scale and above can be found in the literature (30–32). This type of model has the important feature of representing rainfall in continuous time; the statistical properties on any discrete timescale are directly obtained from those in continuous time, and this enables model fitting combining statistics of different timescales.

The Bartlett–Lewis rectangular pulse model assumes that rainfall occurs in the form of storms of certain durations and each storm is a cluster of random cells; each storm has constant intensity during the time period it lasts. These are the general assumptions of the model (Fig. 2):

1. Storm origins t_i occur following a Poisson process at rate λ (this means that durations between consecutive storm origins, $t_i - t_{i-1}$, are independent identically distributed following an exponential distribution whose parameter is λ).
2. Origins t_{ij} of cells of each storm i arrive following a Poisson process at rate β.
3. Arrivals of each storm i terminate after a time v_i exponentially distributed with parameter γ.
4. Each cell has a duration w_{ij} exponentially distributed with parameter η.
5. Each cell has a uniform intensity X_{ij} with a specified distribution.

In the original version of the model, it is assumed that all model parameters are constant. In a modified version, parameter η is randomly varied from storm to storm using a gamma distribution of shape parameter α and scale parameter ν. Subsequently, parameters β and γ also vary so that the ratios $\kappa := \beta/\eta$ and $\phi := \gamma/\eta$ are constant.

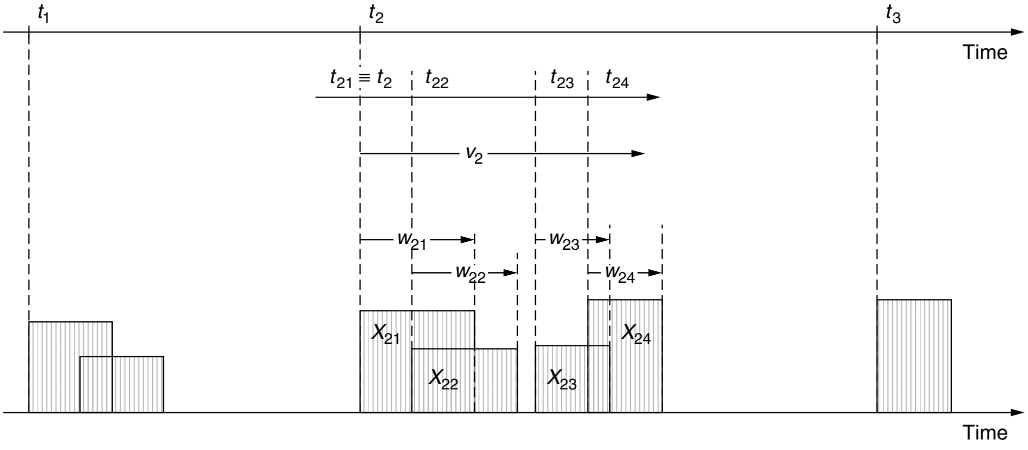

Figure 2. Explanatory sketch for the Bartlett–Lewis rectangular pulse model.

The distribution of the uniform intensity X_{ij} is typically assumed exponential with parameter $1/\mu_X$. Alternatively, it can be chosen as a two-parameter gamma with mean μ_X and standard deviation σ_X. Thus, in its most simplified version, the model uses five parameters, $\lambda, \beta, \gamma, \eta$, and $\mu_{X,}$ and in its most enriched version, seven parameters, $\lambda, \kappa, \phi, \alpha, \nu, \mu_X$, and σ_X.

The equations of the Bartlett–Lewis model, relating the statistical properties of the rainfall process in discrete time to the model parameters, may be found in the references mentioned before. These equations serve as the basis for model fitting.

CONCLUDING REMARKS

Stochastic simulation is a powerful method, easily applicable, and extremely flexible. Its main advantage is its ability to perform in complex systems describing them faithfully, without simplifying assumptions. However, it is an approximate procedure, and the accuracy of its results depends on the sample size. In addition, it is a slow procedure, as the estimation error decreases inversely proportional to the square root of the simulation length (i.e., for half the error, we need four times greater simulation length). Today, this is not a major problem as the progress in computer technology makes attainable even a vast simulation length in reasonable computer time.

In addition to the estimation error due to a finite simulation length, another significant source of uncertainty is always the limited historical records (based on hydrologic measurements), which are used to fit probabilistic or stochastic models. This source of uncertainty, which concerns the simulation method and also any method, including an analytical one, is forgotten sometimes, so the following points of caution should be stressed:

- The choice of a particular stochastic model and the estimation of its parameters are always based on available historical records, which are the only authentic source of information.
- Simulated (synthetic) hydrologic records do not replace historical records.
- The generation of a synthetic record (whose length is usually a multiple of that of the historical record) does not add any information nor does it extend the historical record length.

In conclusion, the following points should be added:

- In problems that can be solved analytically (as in the example of the design of dikes discussed earlier), stochastic simulation is not the preferable method.
- Stochastic simulation becomes a powerful numerical method when studying a complex system and analytical (or other numerical) methods are not applicable, are very difficult, or require oversimplifying assumptions for the system.

Acknowledgment
The review by the editor and the comments by A. Efstratiadis are gratefully appreciated.

BIBLIOGRAPHY

1. Hazen, A. (1914). Storage to be provided in impounding reservoirs for municipal water systems. *Trans. Am. Soc. Civ. Eng.* **77**: 1539.
2. Eckhardt, R. (1989). Stan Ulam, John von Neumann and the Monte Carlo method. In: *From Cardinals to Chaos*. N.G. Cooper (Ed.). Cambridge University Press, New York.
3. Metropolis, N. (1989). The beginning of the Monte Carlo method. In: *From Cardinals to Chaos*. N.G. Cooper (Ed.). Cambridge University, New York.
4. Metropolis, N. and Ulam, S.M. (1949). The Monte Carlo method. *J. Amer. Stat. Assoc.* **44**: 335–341.
5. Barnes, F.B. (1954). Storage required for a city water supply. *J. Inst. Eng. Australia* **26**(9): 198–203.
6. Maass, A., Hufschmidt, M.M., Dorfman, R., Thomas, H.A., Jr., Marglin, S.A., and Fair, G.M. (1962). *Design of Water Resource Systems*. Harvard University Press, Cambridge, MA.
7. Thomas, H.A. and Fiering, M.B. (1962). Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation. In: *Design of Water Resource Systems*. A. Maass, M.M. Hufschmidt, R. Dorfman, H.A. Thomas, Jr., S.A. Marglin, and G.M. Fair (Eds.). Harvard University Press, Cambridge, MA.
8. Beard, L.R. (1965). Use of interrelated records to simulate streamflow. *Proc. ASCE, J. Hydraul. Div.* **91**(HY5): 13–22.
9. Matalas, N.C. (1967). Mathematical assessment of synthetic hydrology. *Water Resour. Res.* **3**(4): 937–945.
10. Box, G.E.P. and Jenkins, G.M. (1970). *Time Series Analysis, Forecasting and Control*. Holden Day.
11. Chow, V.T., Maidement, D.R, and Mays, L.W. (1988). *Applied Hydrology*. McGraw-Hill, New York.
12. Winston, W.L. (1994). *Operations Research, Applications and Algorithms*, 3rd Edn. Duxbury, Belmont, CA.
13. Papoulis, A. (1990). *Probability and Statistics*. Prentice-Hall, London.
14. Ripley, B.D. (1987). *Stochastic Simulation*. John Wiley & Sons, New York.
15. Hurst, H.E. (1951). Long term storage capacities of reservoirs. *Trans. ASCE* **116**: 776–808.
16. Koutsoyiannis, D. (2002). The Hurst phenomenon and fractional Gaussian noise made easy. *Hydrol. Sci. J.* **47**(4): 573–596.
17. Haan, C.T. (1977). *Statistical Methods in Hydrology*. The Iowa State University Press, Ames, IA.
18. Waymire, E. and Gupta, V.K. (1981). The mathematical structure of rainfall representations, 2, A review of theory of point processes. *Water Resour. Res.* **20**(10): 1453–1465.
19. Matalas, N.C. and Wallis, J.R. (1976). Generation of synthetic flow sequences. In: *Systems Approach to Water Management*. A.K. Biswas (Ed.). McGraw-Hill, New York.
20. Todini, E. (1980). The preservation of skewness in linear disaggregation schemes. *J. Hydrol.* **47**: 199–214.
21. Koutsoyiannis, D. (1999). Optimal decomposition of covariance matrices for multivariate stochastic models in hydrology. *Water Resour. Res.* **35**(4): 1219–1229.
22. Koutsoyiannis, D. (2000). A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. *Water Resour. Res.* **36**(6): 1519–1533.
23. Waymire, E., Gupta, V.K., and Rodriguez-Iturbe, I. (1984). A spectral theory or rainfall intensity at the meso-β scale. *Water Resour. Res.* **17**(5): 1273–1985.

24. Bras, R.L. and Rodriguez-Iturbe, I. (1985). *Random Functions in Hydrology*. Addison-Wesley, Reading, MA.
25. Valencia, D. and Schaake, J.C. (1973). Disaggregation processes in Stochastic Hydrology. *Water Resour. Res.* **9**(3): 211–219.
26. Grygier, J.C. and Stedinger, J.R. (1988). Condensed disaggregation procedures and conservation corrections for stochastic hydrology. *Water Resour. Res.* **24**(10): 1574–1584.
27. Koutsoyiannis, D. (1992). A nonlinear disaggregation method with a reduced parameter set for simulation of hydrologic series. *Water Resour. Res.* **28**(12): 3175–3191.
28. Koutsoyiannis, D. and Manetas, A. (1996). Simple disaggregation by accurate adjusting procedures. *Water Resour. Res.* **32**(7): 2105–2117.
29. Koutsoyiannis, D. (2001). Coupling stochastic models of different time scales. *Water Resour. Res.* **37**(2): 379–392.
30. Rodriguez-Iturbe, D., Cox, R., and Isham, V. (1987). Some models for rainfall based on stochastic point processes, *Proc. R. Soc. London, A* **410**: 269–298.
31. Rodriguez-Iturbe, D., Cox, R., and Isham, V. (1988). A point process model for rainfall: further developments. *Proc. R. Soc. London, A* **417**: 283–298.
32. Onof, C. and Wheater, H.S. (1993). Modelling of British rainfall using a random parameter Bartlett-Lewis rectangular pulse model. *J. Hydrol.* **149**: 67–95.

STORAGE AND DETENTION FACILITIES

XING FANG
Lamar University
Beaumont, Texas

The increase of population in the cities results in urbanization of undeveloped areas, which increases impervious areas, e.g., paved roads, parking facilities, residential houses, and commercial property development. The effects of urbanization include a pronounced impact on the characteristics of an area's hydrology; i.e., urbanization leads to an increase of peak flow rate and a reduction of the time to peak of runoff hydrographs. A common practice in storm water management is that the increase in peak flow rate caused by urbanization must be mitigated by reducing the peak flow rate to attain the predevelopment value by using storage and detention facilities. A storage and detention facility (pond or basin) is a low-lying area (natural or constructed), which is designed to hold a specific amount of runoff temporarily and used to attenuate peak flow rate. Once runoff has been collected in a detention facility, it is released to downstream conveyances (e.g., channels, bayous, or lakes) in a controlled manner such that downstream flooding and other adverse impacts are prevented or alleviated. Detention facilities can be surface and underground storage basins. Underground detention facilities are often used for small commercial development and consist of a series of connected large pipes or prefabricated custom chambers manufactured specifically for underground detention; e.g., a supermarket can install an underground detention pipe system under its parking lot to satisfy the zero runoff increase requirement specified by a regulatory agency. A surface detention pond is typically much larger in area and used for regional stormwater management, and it is a depressed or excavated area with earth embankments. Figure 1 shows a schematic of inflow hydrograph (after urban development) and outflow hydrograph from the outlet of a detention pond, and the peak flow of outflow hydrograph is equal to or smaller than the peak flow before development, which is much lower than the peak flow in the inflow hydrograph. From Fig. 1, one can see that the required volume of a detention pond is equal to the difference in area under inflow and outflow hydrographs.

Detention ponds can be classified as wet or dry ponds, depending on whether the principle outlet is constructed above the lowest point in the depression. The wet pond, also known as the retention pond, is designed to retain some of the inflow inside the pond. About 80% or more of detention facilities in the United States and Canada are dry ponds during no rain season, and they provide the maximum capacity for runoff detention during the next rainfall event. Sometimes pipes or channels may be necessary to divert inflow into detention facilities if the detention pond is outside of the main drainage pathway (off-line pond). An on-line pond is positioned along the drainage pathway, and all runoff naturally pass through it without any inflow diversion or collection system. Detention storage designs must consider the size of the pond, site selection, outflow structure (weir or orifice) design, construction of pond (cutting and filling), maintenance cost, and estimates of service life. Multiple ponds may be necessary to achieve the detention requirement under certain site conditions, particularly where the topography is flatter.

Design criteria for detention ponds vary widely. In the United States, detention regulations are typically adopted at the municipal or county level. Design of a detention pond like any other hydrologic design is always associated with one or more design storm events, e.g., typically specified as a certain return period. For example, a drainage district is to design a detention pond to hold all runoff for a 300-acre developed area during a 100-year, 24-hour rainfall event. The 100-year, 24-hour rainfall for the study area is 12 inches. If available land for the detention pond is 40 acres, what will be the depth for the pond? The depth will be 7.5 ft (300 acre × 12 inches/40 acre/12). This

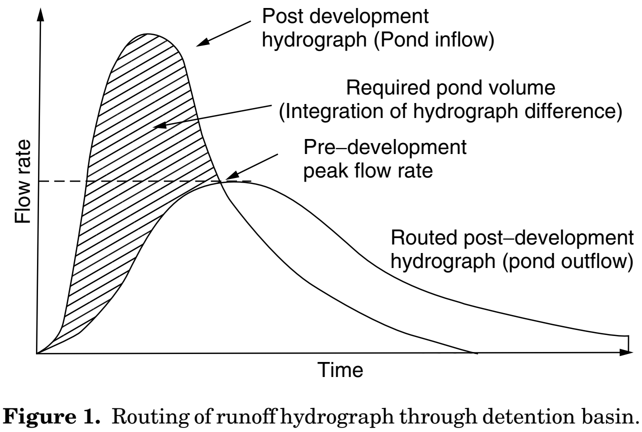

Figure 1. Routing of runoff hydrograph through detention basin.

detention pond design is simple because the pond holds all runoff during the rainfall event and the district plans to manually control outflow-gate to release the flow after the rainfall event. This design does not require any hydrologic routing analysis, but it requires a large amount of land requisition (40 acres).

The design of the storage volume (size) and outlet structure often requires knowledge of the predevelopment peak flow and an estimate of the inflow hydrograph under urbanized conditions. It needs to perform hydrologic routing analysis for the inflow hydrograph into the pond. Hydrograph routing (1) can predict the temporal and spatial variation of an inflow hydrograph as it traverses a river reach or reservoir (pond). Routing through a detention pond is meant to solve the continuity equation by using the storage–outflow relation. The storage-outflow relation is typically determined from two relationships: outflow (discharge) versus water surface elevation and storage (volume) versus water surface elevation. The first relationship is determined by the outflow structure design, and the second one is determined by the detention pond geometry (change horizontal area with respect to elevation) from contour or topographic maps.

Hydraulic analysis for detention pond design determines outflow through the outlet under different water surface elevations inside the pond, and outflow rate depends on type and geometry of the outlet. The most common outlet structures for detention ponds can be of orifice-type, weir-type (broad-crested, sharp-crested, spillways, and v-notch weir), or a combination of the two. The exit flow can be classified as submerged or free and depends on the effective head of water surface on the upstream side of the outlet structure. The outflow from each type of outlet structure is a function of several variables. For the flow through the orifice-type (e.g., a culvert), it is function of orifice area and the difference between the head water elevation and the centroid of the orifice. For the weir-type outlet (emergency spillway), it is function of effective crest length and the head over the weir crest. For the V-notch weir, it is a function of the notch angle and the head over the notch bottom. The riser pipe acts as weir at low heads and like an orifice at higher heads. The equations required to calculate the outflow from the detention pond for different outlet structures can be found in many textbooks and design handbooks (2,3).

A detention pond can be formed in a low-lying area by only constructing the earth embankment with the outflow structure and it could be fully constructed, i.e., digging a pond for certain depth (construction cost is high but sometimes that is the only choice). The side slopes of a constructed pond should be stabilized where needed to prevent the erosion. The side slopes should be 3H: 1V or flatter (4). Outlets should have trash racks, and course gravel packing should be provided if a perforated riser outlet prevents clogging of the riser caused by accumulation of sediment particles. Criteria for selecting the site for install of a pond should include the site's ability to support the environment as well as the cost effectiveness of locating a pond at that specific site. The pond should be located where the topography of the site allows for maximum storage at minimum construction cost (5). A site-specific constraint for pond construction includes underlying bedrock that would require expensive cutting operations to excavate. In designing the wet detention pond, the site must have adequate base-flow from the groundwater or from the drainage area to maintain the permanent pool. The storage capacity of the detention pond may be maintained properly by designing the outlet structure properly (appropriate opening size), so that the velocity of the water through the outlet as well as through the pond does not cause erosion of the bed material, interns which may cause clogging of the outlet opening.

The detention pond design process can be summarized as the following steps:

1. Select appropriate site and shape for detention pond design, and select material for detention structure construction.
2. Estimate inflow hydrograph and the peak inflow from watersheds to the detention pond.
3. Estimate the volume of water that can be stored in the detention pond and establish an elevation-storage relationship from the geometry of the pond.
4. Select the proper type of outlet structure and design the outlet opening.
5. Perform hydrologic routing analysis to determine the peak flow rate of the outflow hydrograph, to check weather it is smaller than the allowable one (peak rate before development) or it is safe, which does not cause flooding at downstream. If it does not satisfy the criteria, redesign the outlet or increase the height of the retaining structure to increase the storage.

Estimation of the size of a detention pond can also be accomplished by some simplified methods (1), e.g., natural storage loss method, modified Rational triangular hydrograph method, and NRCS TR-55 procedure, and by using computer software packages. Under normal operation of a detention pond, it is assumed that sufficient storage is available to regulate a flood without overtopping of the pond. Under special circumstances, the emergency operation scheme is implemented when the storage capacity is limited and inflows caused by heavy rainfall events (with a return period longer than the design return period) are expected to exceed the storage capacity (overtopping). Therefore, an emergency spillway should be provided to control overflow relief for large storms. The detention pond can also be designed to improve the water quality by removing some pollutants through sedimentation and other mechanisms as runoff is temporarily stored in the pond before the water is draining to another natural water body.

BIBLIOGRAPHY

1. Viessman, Jr., W. and Lewis, G.L. (2003). *Introduction to Hydrology*, 5th Edn. Prentice Hall, Englewood Cliffs, NJ.
2. Roberson, J.A., Cassidy, J.J., and Chaudhry, M.H. (1997). *Hydraulic Engineering*, 2nd Edn. John Wiley & Sons, New York.

3. McCuen, R.H. (1998). *Hydrologic Analysis and Design*, 2nd Edn. Prentice-Hall, Englewood Cliffs, NJ.
4. Mays, L.W. (1990). *Hydraulic Design Handbook*. McGraw Hill, New York.
5. NVPDC (Northern Virginia Planning District Commission). (1990). *Evaluation of Regional BMPs in the Occoquan Watershed*. Annandale, VA.

URBAN STORMWATER RUNOFF WATER QUALITY ISSUES

G. Fred Lee
Anne Jones-Lee
G. Fred Lee & Associates
El Macero, California

INTRODUCTION

Urban creeks and lakes can provide important habitats for aquatic life, as well as aesthetic assets to communities. A key component of this resource is the quality of water in these waterbodies. This chapter is devoted to a review of water quality problems in urban creeks and lakes caused by stormwater runoff-associated pollutants.

The primary function of many urban creeks is the conveyance of stormwater to prevent flooding. Often they have been channelized to assist in achieving rapid removal of stormwater from an urban area. This channelization, coupled with the development (paving) in the urban creek watershed, is at odds with providing high-quality aquatic life habitat. Urban creek flows can vary from a few cubic feet per second of groundwater-based flow to a thousand or more cubic feet per second during flood flow conditions associated with major runoff events. The high flows are detrimental to developing and maintaining desirable aquatic life habitat. Urban creeks also are frequently receptacles for waste materials, litter and debris, including shopping carts, yard waste, and so on. At the same time, urban creeks and lakes can provide important aesthetic amenities and, in some cases, recreational fisheries and nursery areas for aquatic life. The fisheries in urban creeks can range from a sustainable trout fishery to carp- or minnow-dominated waters. Some urban lakes provide good warmwater sport fisheries for bass, bluegill, and so on.

The Center for Watershed Protection report entitled, “Impacts of Impervious Cover on Aquatic Systems” (1), contains information on some aspects of the impact of urban stormwater runoff on water quality. This report is an expansion/update of earlier work by Schueler (2) on the impact of urbanization (paving) of an area on the waterbodies receiving the runoff from the area. Burton and Pitt (3) developed a *Stormwater Effects Handbook*, which provides background information on the water quality problems associated with stormwater runoff from urban and, to a lesser extent, rural areas. They discuss impacts on receiving water uses and sources of stormwater pollutants. Most of this over-900-page handbook is devoted to a discussion of approaches for assessing the characteristics of stormwater runoff and its impacts on receiving water quality. Another source of information on the impact of urban stormwater runoff-associated chemical constituents on water quality is the *Stormwater Runoff Water Quality Science/Engineering Newsletter* available from http://www.gfredlee.com.

REGULATING URBAN STORMWATER RUNOFF WATER QUALITY IMPACTS

The U.S. EPA (4) promulgated the national regulations that require cities with populations greater than 100,000 to develop stormwater runoff water pollution control programs to control pollution to the maximum extent practicable (MEP) using best management practices (BMPs). These requirements are also applicable to highway stormwater runoff. The EPA did not define MEP or the BMPs that are to be used. Urban stormwater runoff pollution is to be regulated by National Pollution Discharge Elimination System (NPDES) permits that ultimately require compliance with water quality standards. Although NPDES permits for domestic and industrial wastewater discharges are to prevent violations of water quality standards at the point of discharge or at the edge of a mixing zone if provided for, the NPDES permits issued to municipalities for controlling urban stormwater pollution, while requiring ultimately compliance with water quality standards, did not specify a date by which compliance with water quality standards is to be achieved. The current regulations require the use of a BMP ratcheting-down process to ultimately, but as of yet at an undefined date, control violations of water quality standards in the runoff. Justification for this difference in approach for regulating urban stormwater runoff, compared with urban wastewater discharges, develops from the significantly different characteristics of urban stormwater runoff.

ASCE (5) and CASQA (6) list a variety of “BMPs” that are advocated for the “treatment” of urban stormwater runoff to control water pollution, including ponds that allow settling of some chemical constituents, grassy areas that allow settling and removal of some chemicals, and infiltration basins that allow infiltration of the stormwater into groundwaters. These BMPs were largely based on hydraulic factors without evaluation of their effectiveness in treating stormwater runoff to achieve water quality standards. To the contrary, it has been found that ponds (detention basins), grassy swales, and so on cannot be considered adequate for treating urban stormwater runoff to achieve compliance with water quality standards.

The cost of retrofitting conventional BMPs to developed urban areas has been estimated to be on the order of $1 to $3 per person per day for the population served by the storm sewer collection system. These costs are primarily for the acquisition of land and effectively restrict implementation of these BMPs to new developments where their cost can be incorporated into the cost of the development. Although conventional BMPs are being installed especially in new developments, the most popular conventional BMPs will not be adequate to treat urban or highway stormwater runoff to achieve compliance with water quality standards. To achieve compliance with water quality standards for urban area and highway stormwater runoff, it will be necessary to construct, operate, and maintain treatment works of the type used in advanced wastewater treatment. The costs of such retrofitted treatment works in developed areas is projected to be on the order of $5 to $10 per person per day. These very high costs will require that a different approach be developed for regulating urban area stormwater runoff.

EVALUATION OF WATER QUALITY IMPACTS OF STORMWATER RUNOFF

Although there has been considerable work on the chemical characteristics of urban stormwater runoff, little work has been devoted to evaluating the water quality/beneficial use impacts of this runoff. Especially in light of the tremendous costs associated with providing for treatment/control of stormwater runoff, it is important to properly assess whether a chemical constituent derived from stormwater runoff that is present in an urban stream or lake is in a chemical form that is toxic or bioavailable, i.e., can cause pollution. Failure to make this evaluation can lead to expenditure of large amounts of public funds for the development and installation of so-called best managements practices that effect little or no improvement of the beneficial uses of an urban stream or lake or other waterbody receiving the stormwater runoff.

The quality of water in urban creeks, at times, is dominated by urban stormwater runoff-associated constituents. In the late 1970s/early 1980s, the U.S. EPA conducted a Nationwide Urban Runoff Program (NURP) in 28 communities across the United States. The NURP studies provided information on concentrations and loads of a variety of potential pollutants in urban stormwater runoff. Pitt and Field (7) summarized the results of the NURP studies, as did WEF/ASCE (8). Although the U.S. EPA NURP studies provided data on the concentrations and loads of a variety of potential pollutants in urban stormwater runoff, they failed to address true water quality issues, i.e., the impacts of the potential pollutants on the beneficial uses of the receiving waters for the runoff (9).

Heavy Metals

In the fall of 1998, the California Storm Water Quality Task Force conducted a review of constituents that are present in urban area and highway stormwater runoff in sufficient concentrations to cause violations of U.S. EPA water quality criteria and/or California Toxics Rule standards, which were developed by the U.S. EPA (10) for the state. Copper, lead, and zinc were found in almost all urban street and highway stormwater runoff in concentrations that would violate U.S. EPA worst-case-based water quality criteria and state standards based on those criteria. Sometimes cadmium and mercury were also present above those criteria/standards. These findings indicate that there is a potential for certain heavy metals in urban stormwater runoff to be toxic to aquatic life in urban creeks.

Lee and Taylor (11,12) presented the results of a study of heavy metal concentrations and aquatic life toxicity in ten different Upper Newport Bay (Orange County, California) watersheds during 1999–2000. Several watersheds had predominantly urban land use. Lee and Taylor found several heavy metals, including copper, zinc, and lead, in concentrations above water quality criteria/standards. Through toxicity identification evaluation (TIE) studies, Lee and Taylor (12) found, as have others, that heavy metals in urban residential area and highway stormwater runoff are in nontoxic forms. However, this is not necessarily the case for heavy metals in industrial stormwater runoff. Several examples exist in which heavy metals such as zinc from galvanized roofs or copper from copper roofs can be present in industrial stormwater runoff in sufficient concentrations and available forms to be toxic to aquatic life.

Aquatic Life Toxicity

Several constituents are normally present in urban-area stormwater runoff that could cause aquatic life toxicity. The constituents of greatest concern are the heavy metals, including copper, zinc, lead, and occasionally cadmium. Toxicity measurements of urban stormwater runoff from several areas (12–14) have shown that although runoff from urban residential and commercial areas may be toxic to *Ceriodaphnia* (a U.S. EPA standard freshwater zooplankton test organism), that toxicity has not been because of heavy metals. Toxicity identification evaluations have shown that the toxicity measured was caused by the organophosphorus (OP) pesticides diazinon and/or chlorpyrifos. Although the OP pesticides are of concern because of their toxicity to a few types of zooplankton, they are not toxic to fish or algae at the concentrations typically found in urban runoff.

Pesticide-Caused Toxicity

Lee et al. (13,14) reviewed the topic of OP-pesticide-caused toxicity. Diazinon and chlorpyrifos have been, or will soon be, phased out of urban use by the U.S. EPA because of their potential toxicity to children. Chlorpyrifos can no longer be sold for use as a pesticide in urban areas. The U.S. EPA and the registrants have agreed that it will no longer be legal to sell diazinon for urban use after December 2004. These OP pesticides are being replaced by others, especially the pyrethroid pesticides, in urban areas. However, the replacement pesticides have not been evaluated by the U.S. EPA Office of Pesticide Programs for their potential to cause aquatic life toxicity in stormwater runoff from their point of application. Several pesticides are more toxic to fish and zooplankton than the OP pesticides. Further, many of the pyrethroid pesticides tend to sorb strongly to soil particles and, therefore, will be transported in particulate form and accumulate in sediments. Weston (15) and Weston et al. (16) have reported finding that some sediment-sorbed pyrethroid-based pesticides are bioavailable to some benthic organisms. It is unclear whether this bioavailability leads to toxicity. It could, however, cause toxicity in urban streams and lakes and their sediments that is adverse to aquatic life-related beneficial uses of the waterbody.

Dissolved Oxygen

Stormwater runoff events can cause significant dissolved oxygen (DO) depletion in urban streams and other nearby waterbodies. DO measurements made by the DeltaKeeper (17) in waterbodies just before, during, and after a runoff event showed that the DO before the event was adequate for maintenance of aquatic life, i.e., above about 5 mg/L. However, shortly after the event began, the DO in some of the waterbodies dropped to less than 1 mg/L and stayed depressed for several days. A stormwater runoff event in November 2002, and another in August 2003, which were the first major runoff events of the summer/fall, led to large fish kills in half a dozen or so of those waterbodies.

Nutrients

Urban stormwater runoff contains elevated concentrations of various nutrients (nitrogen and phosphorus compounds) that can lead to excessive fertilization of urban creeks, lakes, and downstream waterbodies. Kluesener and Lee (18) and Rast and Lee (19) determined the nutrient loads associated with urban stormwater runoff. In addition to being derived from stormwater runoff, nutrients, especially nitrate, can also be present in groundwater flow to urban creeks and lakes, which can be an important source of nitrate.

Cowen and Lee (20) reported that part of the algal available P in urban stormwater runoff was derived from the leaching of tree leaves and flowers. Cowen and Lee (21) conducted studies of the algal available phosphorus in urban stormwater runoff in several urban areas. Lee et al. (22) summarized the results of those studies and those of others on algal available P in urban and agricultural runoff. In general, it has been found that the algal available P in stormwater runoff from urban and agricultural areas is equal to the sum of the soluble orthophosphate plus about 20% of the particulate phosphorus. Therefore, about 80% of the particulate phosphorus (which can be most of the phosphorus load in such runoff) does not support algal growth.

pH

There can be sufficient primary production in urban creeks and lakes to cause significant diel (over a 24-hr day) changes in pH and dissolved oxygen, which is especially true for those urban streams that have only limited areas where extensive canopy from trees along the bank shades the water. The U.S. EPA (23) Gold Book water quality criterion limits the pH of waters to 9. It is not unusual for the pH of waterbodies to exceed that value in the late afternoon, at the height of photosynthetic activity, and to be several units lower in early morning.

Ammonia

It is possible for the ammonia concentrations in urban creeks to be sufficiently high to violate ammonia water quality criteria based on potential toxicity to aquatic life in an urban stream or lake, which is especially likely if a significant storm sewer discharge contains storm sewer-accumulated sludge/sediments scoured during a runoff event, or scour of stream sediments, which would tend to have high ammonia concentrations. The impact of that ammonia would be exacerbated in those urban streams and lakes that are highly productive as they would tend to have an elevated pH in mid-afternoon because of photosynthetic activity.

Sanitary Quality

Especially during dry weather flow, urban stormwater runoff and, in some situations, drainage ways such as creeks in urban areas, often have greatly elevated concentrations of total coliforms, fecal coliforms, and *E. coli*. The U.S. EPA (24) announced that it was going to require that states adopt a revised contact-recreation criterion for fresh water based on the measurement of *E. coli*. *E. coli* has become the standard recommended organism for assessing the sanitary quality of a freshwater with respect to contact recreation. It is also a useful indicator of potential pathogens in domestic water supplies. Enterococci have become the standard fecal indicator organism for marine waters. The U.S. EPA (4) announced that it was implementing its 1986 criteria for those bacteria in states bordering Great Lakes and in ocean waters that had not adopted those criteria by April 2004. In 2005, the U.S. EPA will develop revised contact recreation-based water quality criteria for the inland waters of the United States.

In many communities, the design of the sanitary sewerage (collection) system is such that discharges of raw sewage to urban waterways can be associated with pump station power failure, blockage of the sewer, and other factors. Further, sanitary sewerage systems are sometimes poorly maintained, with the result that there can be discharges of raw sewage to nearby watercourses on an ongoing basis through leaks in the sewerage system. In addition, animals, including birds, can contribute significant amounts of fecal coliforms and *E. coli* to stormwater runoff, which, in turn, can cause urban creeks to have poor sanitary quality.

With increased emphasis on managing the water quality impacts of urban stormwater runoff in some parts of the country, such as Southern California (especially in the Santa Monica Bay watershed, because of the adverse impacts on sanitary quality of Santa Monica Bay beaches), efforts are being made to control *E. coli* and other pathogen indicators in stormwater runoff, as well as in separate storm sewers during dry weather flow. Ultimately, through comprehensive studies that are now being developed in the Los Angeles Basin and elsewhere, information will be gained on the specific sources of *E. coli* and the potential for their control. Information on the current understanding and control of the sanitary quality of urban stormwater runoff is available in the proceedings of the US EPA 2004 national Beaches conference: http://www.epa.gov/beaches/.

Total Organic Carbon (TOC)

Based on U.S. EPA regulations, domestic water supplies that have a total organic carbon (TOC) concentration above about 2 mg/L may be required to treat the water to remove the total organic carbon to that level, in order to reduce the potential for formation of trihalomethanes (THMs) and other disinfection byproducts during the disinfection of the water supply. This situation raises the question as to whether urban stormwater runoff could be a significant contributor of TOC to urban creeks and ultimately to downstream waterbodies that are used for domestic water supply purposes. Site-specific investigations need to be conducted to evaluate this situation for a particular waterbody.

Excessive Bioaccumulation of Hazardous Chemicals in Edible Aquatic Organisms

Fish and other edible aquatic organisms taken from some urban streams have been found to contain excessive concentrations of legacy pesticides such as DDT, dieldrin, and chlordane, derived from their former use in urban areas as well as from current runoff from urban areas that had been agricultural. In addition, fish and other aquatic life in urban streams can contain excessive concentrations of PCBs and dioxins/furans. As discussed by Lee and Jones-Lee (25), dioxins are known to be present in stormwater runoff from urban areas and highways and can, therefore, be present in urban streams and lakes, especially in the sediments. PCBs are sometimes found in urban stream fish because of spills of electrical transformer PCBs that have occurred in the urban stream watershed or illegal discharges of PCBs from industrial sources to the storm sewer system.

An example of this type of situation occurred in the Smith Canal in the city of Stockton, California. Some edible fish taken from that canal in 1998 contained concentrations of PCBs at levels that are considered hazardous for consumption because of the increased risk of cancer. Lee et al. (26) conducted a study on Smith Canal sediments to determine the total concentrations of PCBs in the sediments and their bioavailability using the U.S. EPA standard sediment bioavailability test procedure with *Lumbriculus variegatus*. It was found that although the sediments had high TOC, which would tend to make the PCBs less bioavailable, there still was significant uptake of the PCBs from the sediments by *Lumbriculus*, which indicates that those organisms would be a food-web source of the excessive PCBs that are found in higher trophic-level edible fish taken from parts of the Smith Canal.

Some measurements of mercury in urban stormwater runoff have shown that the concentrations are sufficient to potentially lead to excessive bioaccumulation of mercury in edible fish tissue. In urban streams or lakes where bioaccumulation of mercury is a potential concern, fish should be examined to determine if they have excessive bioaccumulation of mercury. Lee and Jones-Lee (25) and Lee (27) have provided guidance on approaches that should be followed to evaluate excessive mercury bioaccumulation by examination of edible fish tissue.

PAHs, Oil and Grease, and Unrecognized Hazardous/Deleterious Organic Chemicals

Numerous organic compounds are not pesticides or organochlorine bioaccumulatable chemicals but are of potential concern in urban stormwater runoff. These compounds include oil and grease, PAHs, and others included in the group of "total organic carbon." Within the oil and grease and TOC fractions in urban stormwater runoff can be thousands of unregulated organic chemicals that pose a threat of toxicity to aquatic life and/or bioaccumulate in edible aquatic life where they pose a threat to higher trophic-level organisms, including humans. Many chemicals have been in use and entering the environment for many years but have not been regulated. For example, Silva (28) of the Santa Clara Valley Water District, California has reported that sufficient perchlorate leaches from a flare used at a highway accident to contaminate 726,000 gallons of drinking water with perchlorate above the California Department of Health Services action level of 4 μg/L.

Daughton (29) indicated that although there are more than 22 million organic and inorganic substances, with nearly 6 million commercially available, fewer than 200 are addressed by the current water quality regulations. He noted special concern that in general pharmaceuticals and personal care products (PPCPs) are not regulated but can pose significant water quality concerns. Daughton stated, "Regulated pollutants compose but a very small piece of the universe of chemical stressors to which organisms can be exposed on a continual basis." Additional information on PPCPs is available at http://www.epa.gov/nerlesd1/chemistry/pharma/index.htm.

Suspended Sediment/Turbidity

If an urban creek watershed contains areas of new construction and/or if the urban creek watershed and the creek have soils that readily erode, there can be significant increases in suspended solids/turbidity in the creek during runoff events. The increased turbidity makes the water turbid (muddy), which can affect aquatic life habitat.

Trash

Urban creeks are notorious for accumulating materials that people discard, including grocery carts, tires, paper, Christmas trees and shrubbery, and lawn trimmings. Although some of these items can inhibit flow and thus lead to flooding, some of this material also provides habitat for aquatic organisms in the creek. The primary adverse impact of trash is on the aesthetic quality of the waterbody. Some creeks receive large amounts of trash, which is evidenced by the "creek days" that environmental/public groups conduct, when debris of various types is removed from the creek. With increased emphasis being placed on controlling trash in stormwater runoff in the Los Angeles area pursuant to a TMDL issued to control trash in urban stormwater runoff (30), there could be a reduction in the total amount of trash that is dumped into Los Angeles area urban creeks.

Aquatic Life Habitat

As part of its Water Quality Criteria and Standards Plan (24), the U.S. EPA specifically delineated urban stormwater runoff as a cause of deteriorated aquatic life habitat. The habitat degradation is a result of a variety of factors including channelization and increased urban stream flow caused by paved development in the watershed.

The CWP (1) report contains an extensive discussion of the impact of urbanization with the associated increase in impervious cover (e.g., paving) in urban stream watersheds, on the hydrological and morphological characteristics of urban streams. It reported that when the percentage of impervious cover in an urban stream's watershed exceeds about 10%, the stream's characteristics are typically impacted. When the impervious cover exceeds about 25%, severe impacts on the waterbody's characteristics tend to occur.

As part of the implementation of its Water Quality Criteria and Standards Plan, the U.S. EPA plans to pursue the use of bioassessment methodology to determine the degree of degradation caused by urban stormwater runoff that would need to be corrected to develop desirable aquatic life habitat in urban streams and other waterbodies that receive urban stormwater runoff (24). Thus far the U.S. EPA and state water pollution control agencies seem to have made little progress toward achieving this goal. Information on the U.S. EPA's current program in this area is presented at http://www.epa.gov/ebtpages/watewaterbioassessment.html.

OVERALL

It has become evident that a need exists for comprehensive water quality monitoring/evaluation programs to determine, for representative locations, the real, significant water quality-use-impairments that are occurring in urban lakes and streams (and, for that matter, downstream waters) receiving urban area and highway stormwater runoff. This monitoring/evaluation program should include defining the specific sources of the constituents that lead to the water quality/use impairments. Once the water quality problems have been defined and the sources of the responsible pollutants have been identified, then a reliable evaluation can be made of the management practices that can be implemented to control the pollution of urban streams and lakes by urban area stormwater runoff-associated constituents. In general, because of the high cost of treatment, it is likely that the management practices will focus on source control, as opposed to treatment of the stormwater runoff.

The U.S. EPA's announced "Strategy for Water Quality Standards and Criteria" (31) includes development of wet weather water quality standards. These standards would more appropriately consider how chemical constituents in stormwater runoff impact the beneficial uses of receiving waters. They would likely include a weight-of-evidence evaluation of the relationship between the concentrations of toxic/available forms of constituents in stormwater runoff and their impacts on aquatic-life-related resources in the waterbodies receiving the runoff. Lee and Jones-Lee (32) have reviewed this approach for managing urban area stormwater runoff water quality (33).

BIBLIOGRAPHY

1. CWP. (2003). Impacts of impervious cover on aquatic systems. Watershed Protection Research Monograph No. 1, Center for Watershed Protection, Ellicott City, MD. Available: http://centerforwatershedprotection.goemerchant7.com/index.cgi.
2. Schueler, T. (1994). The importance of imperviousness. *Watershed Protection Techniques* **2**(4): 100–111.
3. Burton, G.A., Jr. and Pitt, R.E. (2002). *Stormwater Effects Handbook: A Toolbox for Watershed Managers, Scientists, and Engineers*. Lewis Publishers, Boca Raton, FL.
4. U.S. EPA. (2004). EPA Proposes Federal Water Quality Standards for Beaches. U.S. Environmental Protection Agency, Washington, DC. Available: http://www.epa.gov/beaches/.
5. ASCE. (2000). "National Stormwater Best Management Practices (BMP) Database." American Society of Civil Engineers/US Environmental Protection Agency. Available: http://www.asce.org/peta/tech/nsbd01.html.
6. CASQA. (2003). "Stormwater Best Management Practice (BMP) Handbooks." California Stormwater Quality Association, Livermore, CA. Available: http://www.cabmphandbooks.com.
7. Pitt, R.E. and Field, R. (1990). Hazardous and Toxic Wastes Associated with Urban Stormwater Runoff. *Proc. of the Sixteenth Annual BREL Hazardous Waste Research Symposium*, US EPA Office of Research and Development, EPA/600/9-90 037, pp. 274–289.
8. WEF/ASCE. (1998). Urban Runoff Quality Management, Water Environment Federation Manual of Practice No. 23 and American Society of Civil Engineers Manual and Report on Engineering Practice No. 87, Water Environment Federation, Alexandria, VA, and American Society of Civil Engineers, Reston, VA.
9. Lee, G.F. and Jones, R.A. (1981). Will EPA's nationwide urban runoff study achieve useful results? *Civil Engi.* **51**: 86–87.
10. U.S. EPA. (2000). Water Quality Standards; Establishment of Numeric Criteria for Priority Toxic Pollutants for the State of California; Rule. U.S. Environmental Protection Agency, Region 9, *Federal Register* 40 CFR Part 131, Vol. 65, No. 97, [FRL-6587-9], RIN 2040-AC44, San Francisco, CA.
11. Lee, G.F. and Taylor, S. (2001). Results of Heavy Metal Analysis Conducted During 2000 in the Upper Newport Bay Orange County, CA Watershed. Report of G. Fred Lee & Associates, El Macero, CA. Available: http://www.members.aol.com/apple27298/Heavy-metals-319h.pdf.
12. Lee, G.F. and Taylor, S. (2001). Results of Aquatic Toxicity Testing Conducted During 1997–2000 within the Upper Newport Bay Orange County, CA Watershed. Report of G. Fred Lee & Associates, El Macero, CA. Available: http://www.members.aol.com/apple27298/295-319-tox-paper.pdf.
13. Lee, G.F., Taylor, S., and County of Orange Public Facilities and Resources Department. (2001). Upper Newport Bay Water Quality Enhancement Project, Final Report. Agreement Nos. 8-023-258-0 and 8-174-250-0, submitted to State

Water Resources Control Board, Santa Ana Regional Water Quality Control Board and Orange County Public Facilities and Resources Department to meet the requirements of the U.S. EPA 319(h) Project, G. Fred Lee & Associates, El Macero, CA and RBF Consulting, Irvine, CA, May.

14. Lee, G.F., Taylor, S., and County of Orange Public Facilities and Resources Department. (2001). Upper Newport Bay/San Diego Creek Watershed 205(j) Water Quality Planning Grant, Final Report. Agreement No. 7-037-250-0, U.S. EPA 205(j) Project, May.
15. Weston, D.P. (2002). Toxicological implications of changing pesticide use in the Central Valley. *Norcal SETAC News* **13**(1): 15–16.
16. Weston, D.P., You, J., and Lydy, M.J. (2004). Distribution and toxicity of sediment-associated pesticides in agriculture-dominated water bodies of California's Central Valley. *Environ. Sci. Technol.* **38**(10): 2752–2759.
17. DeltaKeeper. (2002). Low Dissolved Oxygen Problems in the City of Stockton Waterways during November 2002. Presentation to the Central Valley Regional Water Quality Control Board, Sacramento, CA, December.
18. Kluesener, J.W. and Lee, G.F. (1974). Nutrient loading from a separate storm sewer in Madison, Wisconsin. *J. Water Pollut. Control Fed.* **46**: 890–936.
19. Rast, W. and Lee, G.F. (1983). Nutrient loading estimates for lakes. *J. Environ. Eng. Div. ISCHIA* **109**: 502–517. See also closure discussion, (1984). "Nutrient estimates for lakes," *J. Environ. Eng.* **110**: 722–724.
20. Cowen, W.F. and Lee, G.F. (1973). Leaves as a source of phosphorus. *Environ. Sci. Technol.* **7**: 853–854.
21. Cowen, W.F. and Lee, G.F. (1976). Phosphorus availability in particulate materials transported by urban runoff. *J. Water Pollut. Control Fed.* **48**: 580–591.
22. Lee, G.F., Jones, R.A., and Rast, W. (1980). Availability of phosphorus to phytoplankton and its implication for phosphorus management strategies. In: *Phosphorus Management Strategies for Lakes*. Ann Arbor Press, Ann Arbor, MI, pp. 259–308. Available: http://www.members.aol.com/duklee2307/Avail-P.pdf.
23. U.S. EPA. (1987). Quality Criteria for Water 1986. EPA 440/5-86-001, U.S. Environmental Protection Agency, Washington, DC.
24. U.S. EPA. (1998). Water Quality Criteria and Standards Plan—Priorities for the Future. EPA 822-R-98-003, U.S. Environmental Protection Agency, Office of Water, Washington, DC, June.
25. Lee, G.F. and Jones-Lee, A. (2002). Organochlorine Pesticide, PCB and Dioxin/Furan Excessive Bioaccumulation Management Guidance. California Water Institute Report TP 02–06 to the California Water Resources Control Board/Central Valley Regional Water Quality Control Board, California State University Fresno, Fresno, CA, December. Available: http://www.gfredlee.com/OClTMDLRpt12-11-02.pdf.
26. Lee, G.F., Jones-Lee, A., and Ogle, R.S. (2002). Preliminary Assessment of the Bioaccumulation of PCBs and Organochlorine Pesticides in *Lumbriculus variegatus* from City of Stockton Smith Canal Sediments, and Toxicity of City of Stockton Smith Canal Sediments to *Hyalella azteca*. Report to the DeltaKeeper and the Central Valley Regional Water Quality Control Board, G. Fred Lee & Associates, El Macero, CA, July. Available: http://www.gfredlee.com/SmithCanalReport.pdf.
27. Lee, G.F. (2003). Regulating Mercury in the Water Column and Sediments. Report to Dredge Tailings Workgroup, by G. Fred Lee & Associates, El Macero, CA. Available: http://www.gfredlee.com/TotalMercuryandDissolvedMercuryStandards-rev.pdf.
28. Silva, M.A. (2003). Safety Flares Threaten Water Quality with Perchlorate. Report of Santa Clara Valley Water District. Available: http://www.valleywater.org/Water/Water_Quality/Protecting_your_water/_Lustop/Perchlorate.shtm.
29. Daughton, C.G. (2004). Pharmaceuticals and Personal Care Products (PPCPs) as Environmental Pollutants: Pollution from Personal Actions. Presentation at California Bay-Delta Authority Contaminant Stressors Workshop, U.S. Environmental Protection Agency, Las Vegas, NV. Available: http://daughton.christian@epa.gov.
30. LARWQCB. (2003). Trash TMDL for Los Angeles River Watershed. Los Angeles Regional Water Quality Control Board, Los Angeles, CA. Available: http://www.swrcb.ca.gov/rwqcb4/html/meetings/tmdl/tmdl_pollutant_trash.html.
31. U.S. EPA. (2003). Strategy for Water Quality Standards and Criteria: Setting Priorities to Strengthen the Foundation for Protecting and Restoring the Nation's Waters. U.S. Environmental Protection Agency, Washington, DC. Available: http://www.epa.gov/waterscience/standards/strategy/final.pdf.
32. Lee, G.F. and Jones-Lee, A. (2003). Regulating Water Quality Impacts of Port and Harbor Stormwater Runoff. Proc. International Symposium on Prevention of Pollution from Ships, Shipyards, Drydocks, Ports, and Harbors, New Orleans, LA, November. Available on CD ROM from http://www.ATRP.com. Available: http://www.members.aol.com/duklee2307/PHStormwater-papfinal.pdf.
33. U.S. EPA. (1990). National pollutant discharge elimination system permit application regulations for storm water discharges; final rule. *Federal Register* **55**(222): 47990–48091.

RIVERS AND STREAMS: ONE-WAY FLOW SYSTEM

CHRISTOPHE CUDENNEC
Ecole Nationale Supérieure
Agronomique
Rennes, France

Broadly speaking, rivers and streams are inscribed in the structural relief lines of the lowest local altitudes, the thalwegs. But they do not necessarily strictly follow these lines when they are considered on a very precise scale, because their exact drawing is constrained by local particularities such as the substratum, the vegetation, and anthropic artifacts. Inversely, there is not necessarily a river or a stream in a thalweg because the permanent flow and the landscape marking depend on many hydrologic and geomorphological factors.

Because of this particular position of rivers and streams, they have a major role in hillslope drainage, whatever the kinds of waterways through the hillslopes are. Then, any river or stream receives water from the sides and from upward, and makes it flow downward through a slope-oriented sense. Thus, a point of a stream or a river has a topological and functional relationship with all of its up and downward points. This relationship leads to defining the so-called 'one-way flow system' and makes structured objects emerge on a larger scale. Rivers and streams meet and are organized in a hierarchical branched network.

A drainage network is made up of particular points: upstream extremities called sources, junctions called confluences, and a unique downstream extremity called the outlet—and the identification of this outlet is directly related to the definition of a watershed. The channel joining two such points is a link (1–3). To characterize the position of any link within the system, one can identify its magnitude (4), which is the number of sources it drains. One can also attribute an order to it, according to one of the proposed ordering schemes (1,5). Particularly, according to Strahler's scheme, the notion of (network)-stream is defined as a set of successive links of the same order, from up to downstream (6).

From this identification of individual and collective characteristics of river and stream systems, many works have been performed to describe, explain, quantify, and simulate the organizational schemes and rules of the constituents and furthermore, to base the simulation of water flows on these schemes and rules.

Planar observation of individual channels led to observing different patterns and shapes, mostly based on sinuosity, braiding, and anabranching. The cross section of a given site in a channel can present very different shapes and sizes. Moreover, planar observation of drainage networks also points out different patterns. Such individual and network patterns result from complex interactions between the substratum, the vegetation, the anthropic constraints and influences, and the water fluxes. Moreover, these complex interactions are scale-dependent through a wide range of spatial, temporal, and frequency scales. Nevertheless, these patterns can be described by typologies and by quantifying some relevant geometric criteria (7,8). Among them are the sinuosity characteristics such as the meandering radius and the wavelength (9); the bankfull width, depth, and slope; and the drainage density (1,10).

Within the range of scale-dependencies, the relationship with the corresponding watershed is a major issue. First, even if the development of a drainage network follows up and down dynamics (1,6,11), the initiation of a channel has been related to planar geometric characteristics of the upward territory (1,6), called the source area. Furthermore this dependence can be linked to hydrologic processes and to slopes (12,13). Thus, the drainage density appears to depend partly on geometric and partly on hydrologic considerations (1,13,14), which generates great geographic and temporal variability (3,15).

Second, the local geometric and hydraulic characteristics of a channel are related to characteristics of the watershed at whose outlet it is located (8,16). The first such relation, found by Hack (17), is the expression of the length of the longest waterway as a power law of the area of the watershed. Then, other power law relationships have been identified and studied for a range of regimes and substratum between the discharge of a given frequency of occurrence and the area and between stream channel geometric parameters (depth, width, slope, and velocity) and the discharge.

The channel slope at any point is thus related to the upstream conditions. Along a waterway from a source to the outlet, the upstream characteristics vary, particularly the drainage area, and the channel slope presents a covariation. The longitudinal profile of a waterway demonstrates this covariation that generally presents a typical decreasing shape (18), whose characteristics and variability depend mostly on geology and climate and sometimes, on anthropic effects.

These interpretations of local characteristics within systemic trends are elements of an organization of the drainage network. Other watershed-scale elements tend to show that there is a kind of organization of the hydrogeomorphological one-way flow system. These elements come from the analysis of quantitative characteristics which are identified from geometric, topological, and hydrologic considerations.

"The pioneer work" in this domain is the identification of Horton laws of network planar composition, first proposed with Horton's ordering scheme (1) and then verified with Strahler's (6,19) by the "network definition" of a stream. These laws show that the number of streams of order i, the mean length of streams of order i, the mean total area contributing to streams of order i, and the mean slope of streams of order i broadly satisfy geometric series, according to characteristic ratios, respectively, the bifurcation, the length, the area, and the slope ratios. These laws were recognized as geometric-scaling relationships and were integrated in the fractal analysis of the river network organization (20,21), which was then applied to a wide range of parameters and functions (16,22). The width function, particularly, has both geomorphological and hydrologic significance and is deeply studied (7,16).

All these observations and quantifications of the one-way flow system of rivers and streams are characteristic of a given state of a system and can furthermore help follow the evolution of the system. The channels themselves and their relationships with upstream channels and hillslopes, and the network structure evolve according to natural trends and eventually to anthropic constraints. It appears then that the characteristic observations and quantifications tend to particular values or shapes which show a natural underlying tendency of energy optimality or of self-organized criticality (16,23).

Furthermore, on the basis of these observations and quantifications, many modelings can be proposed. First, organization models try to produce realistic networks, based on deterministic and/or statistical reasoning using topological, percolation, optimality, or self-organization rules (16). Second, hydraulic and hydrologic models can base the quantitative description of water fluxes on such actual and theoretical geomorphological considerations (7,16).

BIBLIOGRAPHY

1. Horton, R.E. (1945). Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Bull. Geol. Soc. Am.* **56**: 275–370.
2. Strahler, A.N. (1952). Hypsometric (area altitude) analysis of erosional topography. *Bull. Geol. Soc. Am.* **63**: 117–142.

3. Strahler, A.N. (1964). Quantitative geomorphology of drainage basins and channel networks. In: *Handbook of Applied Hydrology*. V.T. Chow (Ed.). McGraw-Hill, New York, Section 4-II, p. 38
4. Shreve, R.L. (1966). Statistical law of stream numbers. *J. Geol.* **74**: 17–37.
5. Strahler, A.N. (1957). Quantitative analysis of watershed geomorphology. *EOS Trans AGU* **38**: 912–920.
6. Schumm, S.A. (1956). Evolution of drainage systems and slopes in badlands at Perth Amboy, N.J. *Bull. Geol. Soc. Am.* **87**: 597–646.
7. Beven, K. and Kirkby, M.J. (Eds.). (1993). *Channel Network Hydrology*. John Wiley, Chichester, UK, p. 319.
8. Thorne, C.R., Hey, R.D., and Newson, M.D. (1997). *Applied Fluvial Geomorphology for River Engineering and Management*. John Wiley, Chichester, UK, p. 376.
9. Leopold, L.B. and Langbein, W.B. (1966). River meanders. *Sci. Am.* **214**: 60–69.
10. Horton, R.E. (1932). Drainage basin characteristics. *EOS Trans AGU* **13**: 350–361.
11. Glock, W.S. (1931). The development of drainage systems: a synoptic view. *Geogr. Rev.* **21**: 475–482.
12. Montgomery, D.R. and Dietrich, W.E. (1988). Where do channels begin? *Nature* **336**: 232–234.
13. Montgomery, D.R. and Dietrich, W.E. (1989). Source areas, drainage density, and channel initiation. *Water Resour. Res.* **25**: 1907–1918.
14. Tarboton, D.G., Bras, R.L., and Rodriguez-Iturbe, I. (1992). A physical basis for drainage density. *Geomorphology* **5**: 59–76.
15. Abrahams, A.D. (1984). Channel networks: a geomorphological perspective. *Water Resour. Res.* **20**: 161–168.
16. Rodriguez-Iturbe, I. and Rinaldo, A. (1997). *Fractal River Basins; Chance and Self-Organization*. Cambridge University Press, Cambridge, p. 547.
17. Hack, J.T. (1957). *Studies of Longitudinal Profiles in Virginia and Maryland*. U.S. Geological Survey Professional Paper, 294-B.
18. Scheidegger, A.E. (1967). A stochastic model for drainage patterns into an intramontane trench. *Bull. Assoc. Sci. Hydrol.* **12**: 15–20.
19. Shreve, R.L. (1969). Stream lengths and basin areas in topologically random channel networks. *J. Geol.* **77**: 397–414.
20. La Barbera, P. and Rosso, R. (1987). Fractal geometry of river networks. *EOS Trans AGU* **68**: 1276.
21. Tarboton, D.G., Bras, R.L., and Rodriguez-Iturbe, I. (1988). The fractal nature of river networks. *Water Resour. Res.* **24**: 1317–1322.
22. Dodds, P.S. and Rothman, D.H. (1999). A unified view of scaling laws for river networks. *Phys. Rev. E.* **59**: 4865–4877.
23. Bak, P. (1996). *How Nature Works*. Springer Verlag, New York, p. 212.

STREAMFLOW

XING FANG
Lamar University
Beaumont, Texas

A knowledge of hydrology is fundamental to decision-making processes where water is a component of the engineering system. Streamflow is an important component of the hydrologic cycle. The primary input in a hydrologic cycle is precipitation (rainfall or snowfall). Many forms of rainfall losses occur on Earth's surface, e.g., interception by surface vegetation, depression storage, and infiltration to soil. Rainfall excess is the net rainfall after subtracting rainfall losses from gross rainfall and turns into surface runoff. Surface runoff can eventually gather into natural streams or rivers, and flows in streams/rivers are referred to as streamflow. Streamflow typically quantifies as volumetric discharge in cubic meter or feet per second (cms or cfs). Streamflow is not directly recorded, even though streamflow is one of the most important parameters in hydrologic studies. Typically, the water level at a specific stream cross section is recorded and streamflow is deduced by means of a rating curve. The rating curve is developed by using a set of discharges and gauge heights measured over a period of months or years. Discharges used for rating curve development can be determined by numerical integration after flow velocities are measured at several locations along the cross section and where velocity changes. The principal sources of streamflow data for the United States are the U.S. Geological Survey, U.S. Natural Resources Conservation Service (NRCS), U.S. Forest Service, and U.S. Agricultural Research Service.

Streamflow, at a given location on a stream, is usually represented by a hydrograph. A streamflow hydrograph is a graphical representation of instantaneous discharge at a given location with respect to time. Two types of hydrographs are particularly important: the annual hydrograph and the storm hydrograph. The annual hydrograph is a plot of streamflow versus time over a year, and it shows long-term water balance of precipitation, evaporation, and streamflow in a watershed. From the annual hydrograph, a stream may be classified as a perennial river, ephemeral river, or snow-fed river. A perennial river typical under a humid climate has continuous flow conditions year round and even during no-rainfall periods. An ephemeral river in an arid climate has long periods when the river is dry (no flow condition). A snow-fed river has the bulk of the runoff yield occurring in the spring and early summer from snowmelt, and variation of discharge is smaller than that for other types of rivers. Annual hydrographs typically contain some spikes caused by rain storms. A storm hydrograph is a streamflow hydrograph during and after a storm.

An example storm hydrograph with typical timing parameters and rainfall hyetograph is shown in Fig. 1. A hydrograph is the response of a watershed under certain rainfall input, and it is affected by the watershed's physical, geological, vegetation, and climatic features. The integration of area under a hydrograph between any two points in time gives the total volume of runoff passing the stream cross section during the time interval. A typical hydrograph consists of rising limb, crest segment, receding limb, and base flow. The rising limb depends on duration and intensity of rainfall, antecedent moisture condition, and drainage characteristics of basin. The crest segment contains a peak that represents the highest concentration of runoff that usually occurs soon after the rainfall has ended. The receding limb represents the withdrawal of the

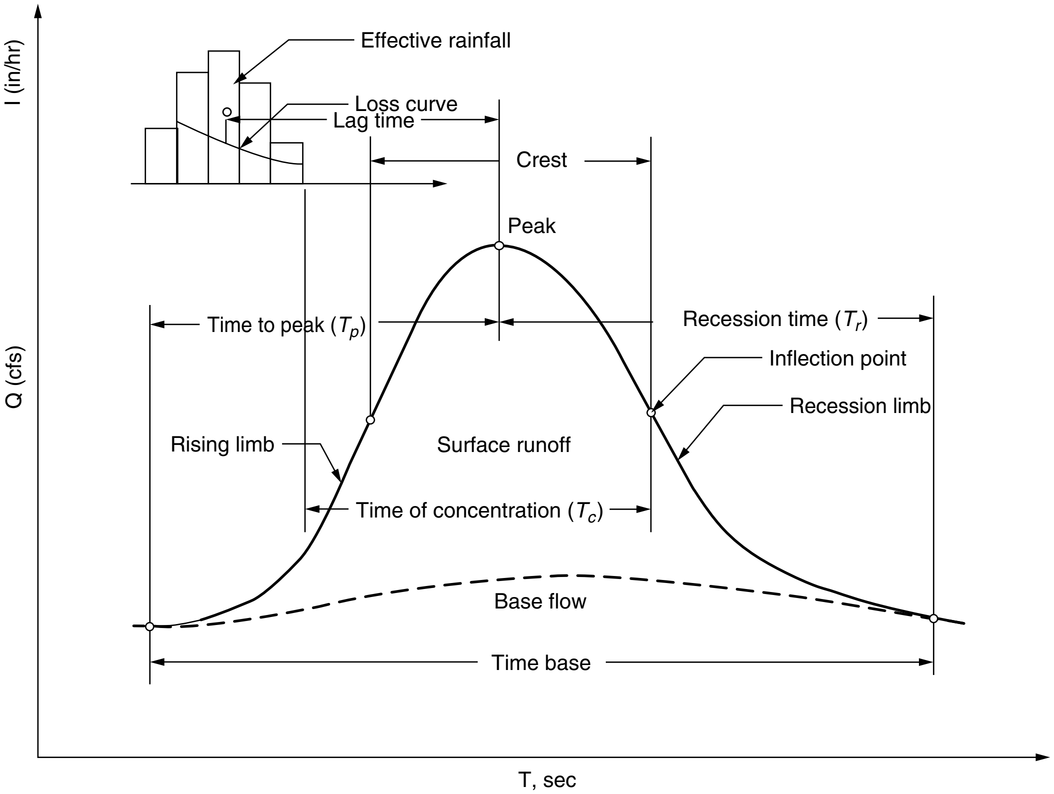

Figure 1. Schematic streamflow hydrograph including typical timing parameters and a rainfall hyetograph.

water from the storage after surface inflow to the channel has ceased. The base flow is that part of runoff that exists even before the occurrence of rainfall (e.g., in the perennial river during the no rainfall period). The total streamflow hydrograph has two components: direct runoff hydrograph (DRH) and base flow. The direct runoff hydrograph is the transformation of effective rainfall passing through a watershed. Base flow is the part of precipitation that percolates downward until it reaches ground water table and eventually discharges into the stream.

Study of a streamflow hydrograph is important for water management and environmental studies as it provides the peak flow rates and for which hydraulic structures can be designed. For the design problems, runoff volume alone is not adequate because the given volume of water may or may not present a flood hazard, but it depends on the time distribution of flood runoff. If the total volume is distributed over a long period of time, flood damages will be minimal, and if the volume is distributed for a short time, even though the volume is small, the damage may be hazardous. Several time parameters have been developed that reflect the timing of runoff, for example, lag time, time to peak, time of concentration, and recession time in Fig. 1. Although direct runoff begins with the commencement of effective rainfall (Fig. 1), the largest portion of runoff generally lags the rainfall because it takes time for runoff to travel from any location within the watershed to the outlet. The basin lag time could be defined as the time from the center of mass of the rainfall excess to the peak discharge rate on the hydrograph (Fig. 1). The time to peak is the time from the beginning to the peak discharge in a simple (single peak) direct runoff hydrograph (Fig. 1). Both the lag time and the time to peak of a hydrograph correlate to the time of concentration of a watershed. The time of concentration (T_c) is the time it takes a water parcel to travel from the hydraulically most distal part of watershed to the outlet or reference point downstream (1–5). In hydrograph analysis, the time of concentration is the time difference between the end of rainfall excess and the inflection point of a hydrograph where the recession curve begins as shown in Fig. 1.

Generally, for a hydrological design, a hydrologist is required to provide peak rates of discharge, a stage height of a hydraulic structure, or a complete discharge hydrograph for a design frequency. The peak discharge of streamflow can be obtained from long-term historical records or from regression equations developed from recorded streamflow. A compete hydrograph, which is needed for design of reservoirs and detention and retention ponds, can be estimated from various rainfall-runoff models or regional dimensional or dimensionless (unit) hydrographs (4). The unit hydrograph is the hydrograph that results from unit rainfall excess distributed uniformly over the drainage area at a constant rate for an effective duration (6). For example, NRCS has a dimensionless unit hydrograph (DUH) procedure for design applications. This DUH developed by Victor Mockus (3) was derived from a large number of natural unit hydrographs from watersheds varying widely in size and geographical locations. Time to peak (T_p) is equal to the watershed

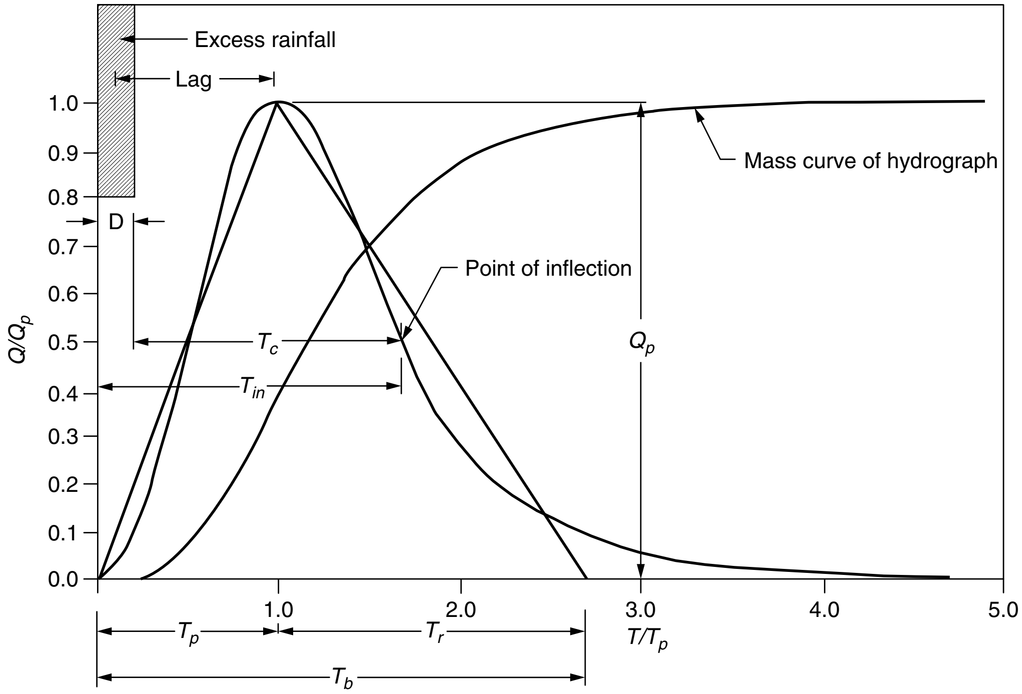

Figure 2. NRCS synthetic unit hydrographs including dimensionless hydrograph and triangular unit hydrograph.

lag time (T_L) plus half of the rainfall excess duration or the duration of unit hydrograph (D in Fig. 2).

$$T_p = T_L + D/2 \qquad (1)$$

This dimensionless unit hydrograph also can be represented by an equivalent triangular hydrograph (Fig. 2). These characteristics of the NRCS unit hydrograph represent values that have been adopted for an *average* watershed. For NRCS DUH, the peak discharge (Q_p) is given as

$$Q_p = KA/T_p \qquad (2)$$

where Q_p is in cfs (cubic feet per second), A (area) is in mile^2, and T_p is in hours. K is the peak rate factor (PRF) and is considered equal to 484, assuming a triangular hydrograph with a time base being 8/3 T_p (Fig. 2). K is related to the internal storage characteristic of a basin and can vary considerably depending on watershed characteristics and scale (size) of a basin. For example, K can range from a value of nearly 600 for steep mountainous conditions to a value nearly to 300 in the flat coastal plains (swampy country) of the state (3). For a very flat, high-water-table watershed, the NRCS peak rate factor of 484 or even 300 likely is too large. NRCS DUH also has an empirical relationship for average lag time that is assumed to be 0.6 T_c (the time of concentration).

$$T_L = 0.6\ T_c \qquad (3)$$

BIBLIOGRAPHY

1. Kirpich, Z.P. (1940). Time of concentration of small agricultural watersheds. *Civil Eng.* **10**(6): 362.
2. USACE. (1966). *Generalized Computer Programs - Unit Hydrograph and Loss Rate Optimization*. Hydrologic Engineering Center, U.S. Army Corps of Engineers (USACE), Davis, CA.
3. NRCS (National Resource Conservation Service). (1972). *National Engineering Handbook*. U.S. Dept. of Agriculture, Washington, DC.
4. McCuen, R.H. (1998). *Hydrologic Analysis and Design*, 2nd Edn. Prentice-Hall, Englewood Cliffs, NJ.
5. Garg, S.K. (2001) *Irrigation Engineering and Hydraulic Structures*. Khanna Publishers, Baltimore, MD.
6. Sherman, L.K. (1932). Stream-flow from rainfall by the unit-graph method. *Eng. News-Record* **108**: 501–505.

WATER QUALITY IN SUBURBAN WATERSHEDS

Charles W. Kirby
Paciulli, Simmons & Assoc.
Fairfax, Virginia

The story of water quality in suburban watersheds is primarily that of nonpoint source pollution (NPS). There may be instances where point sources contribute significantly to degraded water quality in suburban areas, such as privately owned or municipal sewage treatment plants, but the most salient characteristic of suburban basins is the conversion of land from a previous forested or agricultural land cover to highly impervious surfaces such as rooftops, roads, and sidewalks.

When rain or snow strikes the earth's surface, the water may infiltrate into the soil and become part of the groundwater, or it may flow over the surface and enter a stream system as runoff. Under natural conditions, a high proportion of precipitation soaks into the ground and does not directly enter the surface water network, although groundwater may reappear as surface water at the interface between streams and the water table. Forested watersheds can absorb huge amounts of precipitation without experiencing runoff; often one inch or more of precipitation may be absorbed into the soil (1–7). Alteration of the land surface within a watershed reduces infiltration and increases the proportion of precipitation directly entering the waterways as runoff. The increase in the percentage of impervious surface from urbanization in a stream basin is the primary threat to water quality in suburban areas. Zoppou (8) presents an up-to-date review of the numerical models currently used to simulate urban runoff.

Construction activities in land development pose a direct threat to stream quality by sedimentation from the newly cleared land surface. The removal of vegetation destabilizes the soil and allows rainfall to move large amounts of sediment rapidly into the waterways, burying the habitat for macroinvertebrates and juvenile fish and directly killing aquatic vegetation by interfering with photosynthesis (9,10) (Fig. 1). Many local governments require sediment control measures during construction to reduce damage to downstream waters, but many investigators have found that current erosion and sediment control methods are often ineffective (6,11–13).

Streams in suburban watersheds exhibit radically altered flow regimes compared to those in unaltered basins (14). The hydrographs of urbanized streams typically have higher peaks during storms and show more frequent flows above base level because precipitation that formerly soaked into the soil and joined the groundwater is now shunted into the stream network. Urban streams also have steeper rising and falling limbs of the hydrograph, as shown in Fig. 2. More frequent and rapid flooding is known as “flashiness.” These effects destabilize the streambed, which had reached a state of dynamic equilibrium in the predevelopment condition (9,15). The result is erosion of the bed and banks and accompanying increases in sedimentation and direct destruction of habitat for aquatic flora and fauna. Other physical alterations include increased temperature and increase in sediment load. The effects of velocity and other hydraulic parameters on stream fauna are presented cogently and entertainingly by Vogel (16).

Figure 1. Suburban streams typically have entrenched, widened channels with unstable banks.

Suburban development affects surface water chemistry as well, and suburban watersheds are second only to agricultural areas in degraded water quality from chemical contributions. Nutrients such as phosphorus and nitrogen are often increased by higher population density in a watershed. The use of lawn fertilizers may allow runoff of high concentrations of both elements during storms, and home users of these chemicals often apply amounts far in excess of the capacity of the turf to absorb and use them (3). Golf courses are often sources for these nutrients as well

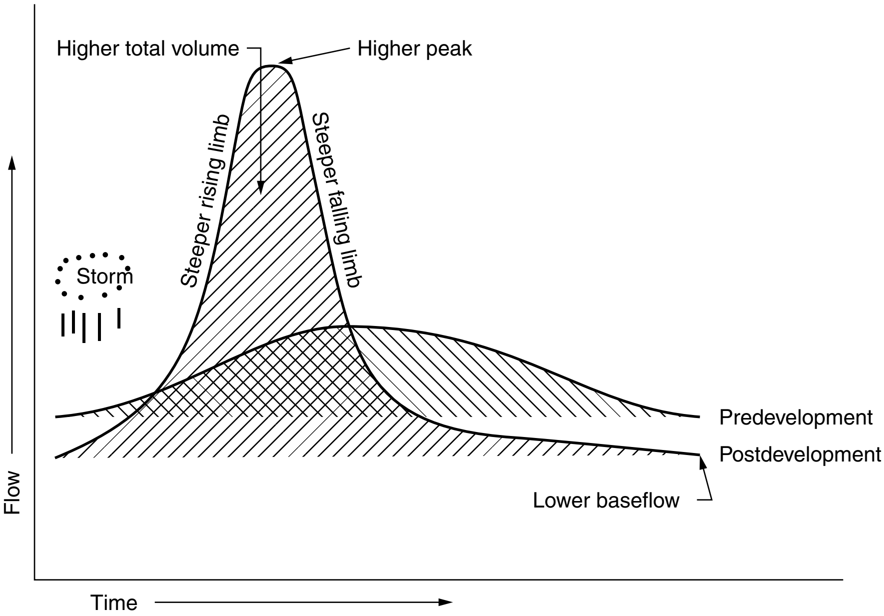

Figure 2. Increased impervious surface in suburban watersheds causes higher volume of runoff, higher peak flow, and lower base flow levels than in undeveloped watersheds.

as other pollutants in suburban areas. Nitrogen has no federally mandated maximum concentration, but such a limit has been set for phosphorus at 1 part per million by the Clean Water Act.

A number of metals are found in elevated concentrations in suburban surface waters. These include nickel, zinc, lead, mercury, cadmium, and copper. The sources of these pollutants range from dry deposition of lead and mercury to particles from automobile brakes and other systems in other cases. Metal concentrations rarely reach the acutely toxic levels found in surface waters drained from industrial sites, but suburban metals pollution nevertheless contributes to degradation of the macroinvertebrate fish, and periphyton populations. In lentic communities and higher order lotic habitats, adsorbed metals may remain permanently in bed sediments (17).

Current stormwater management practices codified in zoning ordinances and municipal facilities standards manuals concentrate almost exclusively on control of erosion and sedimentation (E&S). This emphasis can be traced to the early awareness of civil engineers and water resource managers of the loss of topsoil caused by poor agricultural and construction management practices in the first half of the twentieth century (18). The measures implemented in the period following World War II in most parts of the country are a significant improvement over the prior lack of attention to stream quality issues, but they have proven quite inadequate to preserve aquatic habitat, biotic diversity, and water quality. Engineering solutions to stormwater management problems go under the rubric of best management practices (BMPs). Erosion and sediment control policies for Fairfax County Virginia, an urban county near Washington, DC, are typical of jurisdictions that have recognized water and stream quality problems and the political will to address those problems through engineering standards. Fairfax County BMP requirements for E&S control during construction include

- Design and construction of detention ponds to limit postdevelopment runoff volume from the entire site to levels no higher than the predevelopment condition. The predevelopment and postdevelopment conditions must be demonstrated by hydrologic and hydraulic modeling, according to accepted standards.
- Design of outfall from any stormwater management structures must meet maximum velocity standards to prevent scouring and erosion of the natural stream channel receiving discharge from the constructed structures. The maximum permissible velocity must be hydraulically modeled prior to plan approval and must take into account the soil type in the receiving natural stream channel.
- Temporary E&S controls during construction such as temporary stormwater detention ponds and installation of silt fences intended to prevent sediment-laden runoff from denuded areas from entering off-site streams.
- Prohibitions or limitations on development within floodplains, wetlands, or within 100 feet of a tributary stream.

Environmental regulators in Fairfax County have sent letters to engineering firms and land development companies encouraging the use of innovative BMPs such as bioretention, dry swales, filtration structures, and extended detention ponds, which offer advantages in emulating the predevelopment surface water and groundwater hydrology, but does not require them. However, it is often perceived that innovative BMPS entail higher installation and maintenance costs and are rarely used in many localities. This condition may be expected to change as programs such as the Stage II National Pollutant Discharge Elimination System and the Total Maximum Daily Loads efforts advance to the implementation stage during the next several years. Both programs will place increasing responsibility on local regulatory agencies to decrease negative impacts on receiving waters from the full range of urban hydrology characteristics.

BIBLIOGRAPHY

1. Natural Resources Conservation Service. (1986). *Urban Hydrology for Small Watersheds*. Technical Release 55. US Department of Agriculture, Natural Resources Conservation Service, Washington, DC.
2. Watson, I. and Burnett, A.D. (1995). *Hydrology, An Environmental Approach*. CRC Lewis Publishers, Boca Raton, FL.
3. Northern Virginia Planning District Commission. (1992). *Northern Virginia BMP Handbook*. Northern Virginia Planning District Commission, Engineers and Surveyors Institute, Annandale, VA.
4. Allan, J.D. (1995). *Stream Ecology: Structure and Function of Running Waters*. Chapman and Hall, London.
5. Barbour, M.T., Gerritsen, J., Snyder, B.D., and Stribling, J.B. (1999). *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish*, 2nd Edn. EPA 841-B-99-002. U.S. Environmental Protection Agency, Office of Water, Washington, DC.
6. FitzHugh, T. (2001). *Watershed Characteristics and Aquatic Ecological Integrity: A Literature Review*. The Nature Conservancy Freshwater Initiative, Arlington, VA.
7. Hynes, H.B.N. (1970). *The Ecology of Running Waters*. University of Liverpool Press, Liverpool, UK.
8. Zoppou, C. (2001). Review of urban storm water models. *Environ. Modelling Software* **16**: 195–231.
9. Leopold, L.B., Gordon Wolman, M., and Miller, J.P. (1964). *Fluvial Processes in Geomorphology*. W.H. Freeman, San Francisco, CA.
10. Leopold, L.B. (1994). *A View of the River*. Harvard University Press, Cambridge, MA.
11. Schueler, T. (2000). The Importance of Imperviousness. *Watershed Protection Tech.* **1**: 100–111.
12. McRae, C.R. (1996). *Experience from Morphological Research on Canadian Streams: Is Control of the Two-Year Frequency Runoff Event the Best Basis for Stream Channel Protection? Effects of Watershed Development and Management on Aquatic Ecosystems*. L.A. Roesner (Ed.). American Society of Civil Engineers, New York.
13. Jones, C.R., Via-Norton, A., and Morgan, D.R. (1996). Bioassessment of BMP effectiveness in mitigating stormwater impacts on aquatic biota. *Effects of Watershed Development*

and Management on Aquatic Ecosystems. L.A. Roesner (Ed.). American Society of Civil Engineers, New York.

14. Bhaduri, B., Minner, M., Tatalovich, S., and Harbor, J. (2001). Long-term hydrologic impact of urbanization: a tale of two models. *J. Water Resour. Plann. Manage.* **127**(1): 13–19.
15. Sovern, D.T. and Washington, P.M. (1996). Effects of urban growth on stream habitat. *Effects of Watershed Development and Management on Aquatic Ecosystems*. L.A. Roesner (Ed.). American Society of Civil Engineers, New York.
16. Vogel, S. (1994). *Life in Moving Fluids*. Princeton University Press, Princeton, NJ.
17. Mitsch, W.J. and Gosselink, J.G. (1993). *Wetlands*. Van Nostrand Reinhold, New York.
18. Leopold, L.B. and Maddock, T., Jr. (1954). *The Flood Control Controversy, Big Dams, Little Dams, and Land Management*. Ronald Press, New York.

SURFACE WATER POLLUTION

RAMAKAR JHA
K.K.S. BHATIA
National Institute of Hydrology
Roorkee, Uttaranchal, India

VIJAY P. SINGH
Louisiana State University
Baton Rouge, Louisiana

C.S.P. OJHA
Indian Institute of Technology
Roorkee, Uttaranchal, India

INTRODUCTION

Water is a basic necessity for human survival and sustenance of civilization. The ever-increasing population, steadily rising irrigation activities, rapidly expanding industries, growing urbanization, and rising standard of living exert tremendous pressure on the available water resources (1). The temporal and spatial distribution of the quantity and quality of these resources are highly uneven. Water is used for irrigation, industry, power generation, drinking, bathing, recreation, fisheries, wildlife propagation, pollution abatement, etc. Water in appropriate quantities of specified quality is required for each use (2). Therefore, the quantity as well as quality of river flow should be monitored in river management program.

Surface water pollution is a major environmental concern all over the world (3). The simplest definition of surface water pollution is “the loss of any of the actual or potential beneficial uses of water caused by any change in its composition due to human activity.” The beneficial uses of water are varied and if water is rendered unsuitable for any of the purposes as indicated earlier, then it is polluted to a greater or lesser degree, depending on the extent of the damage caused. Pollution eventually diminishes the aesthetic quality of lakes and rivers. Even more seriously, when contaminated water destroys aquatic life and reduces its reproductive abilities, then it eventually menaces human health.

Natural waters cannot be expected to be immediately suitable for the whole range of beneficial uses. For instance, nobody should expect to drink the water of any river without treatment. Depending on the local usage, weather, and other factors, the water, in its natural state, may be turbid or very highly colored and, hence, unattractive to the user. More importantly, river water is most unlikely to be microbiologically pure, that is, free of those minute organisms (bacteria and the like) that can cause disease in humans. In these cases, water pollution means that, because of some human activity, the water is no longer suitable for some potential use, such as drinking, even after treatment. To illustrate, if a river from which water is taken for treatment for drinking receives a discharge of chemical or possibly toxic waste, the water may be rendered completely unsuitable for eventual distribution. The local treatment works are usually designed to remove natural contaminants—coloring matter, particles causing turbidity, micro-organisms, etc—and therefore may be unable to cope with a heavy load of completely foreign matter, such as chemical waste.

SOURCES OF POLLUTION

The surface water pollution in rivers, lakes, reservoirs, coastal areas, and other surface water resources may occur mainly because of point-source pollution and nonpoint-source pollution. The term point-source pollution refers to pollutants discharged from one discrete location or point, such as an industry or municipal wastewater treatment plant, to surface water resources (Fig. 1). The effluent outfall from factories, refineries, waste treatment plants, etc. emits fluids of varying quality directly into surface waters.

The term nonpoint-source pollution refers to the discharge of pollutants that cannot be identified as coming from one discrete location or point. The pollution outflow from nonpoint source (NPS) pollution, or diffuse pollution, include contaminants that enter the water supply from soils/groundwater systems and from the atmosphere via rain water. Soils and groundwaters contain the residue of

Figure 1. Polluted river in United Kingdom.

human agricultural practices (fertilizers, pesticides, etc.) and improperly disposed of industrial wastes.

A majority of pollution instances are caused by sudden or continuing, accidental or deliberate, discharge of polluting material, which on first consideration might not seem harmful or offensive at all. Such pollution events are caused by the discharge of nontoxic organic matter waste from creameries, sewage (treated or untreated), manure slurry, food production waste, chemicals, and silage effluent to water bodies. When an uncontrolled discharge of organic material occurs, the constraining factor on the growth of the bacterial pollution is removed at a stroke. An immediate abundance of food exists and a corresponding plenitude of dissolved oxygen initially. Bacterial growth is promptly stimulated and the population increases rapidly, consuming the available oxygen as it does so.

The growth of bacteria, therefore, tends to reduce the amount of oxygen dissolved in the water. The extent of oxygen depletion that occurs depends on the rapidity with which the stream takes up oxygen from the atmosphere, i.e., its re-aeration capacity. This capacity is greatest in fast-flowing, turbulent streams and least in deep, slow-flowing rivers. In addition, the loss of oxygen may be counteracted by the photosynthesis of green plants, which produce oxygen during daylight. Where the degree of pollution is severe, these compensating factors may be insufficient to prevent the oxygen content of the water decreasing to very low levels, or, in the worst case, to anaerobic conditions, where a complete absence of free oxygen exists. In such conditions, foul-smelling compounds such as hydrogen sulphide may be formed.

Deoxygenation is the most important potential effect of organic waste discharge. However, toxic compounds, such as ammonia, may be present in such a waste, particularly where they have been stored for some time and have become septic, which can happen with farm waste, in which the presence of compounds, such as ammonia and hydrogen sulphide, probably contribute as much as the deoxygenating capacity of these wastes to the devastating effect on fish stocks in pollution incidents. When toxic pollution occurs, the effects are often direct and immediately apparent—fish are killed, the flora and fauna of the water receiving the pollution may be wiped out, different visible effects will be noticeable, and there may be noxious smells. The principal effects are those of direct poisoning by the offensive pollutants discharged.

EFFECTS OF POLLUTION ON DRINKING WATER

The surface water pollution occurring because of point and/or nonpoint sources of pollution creates a health hazard in many ways. Table 1 shows the effect of various water quality variables for drinking purposes.

TREATMENT OF POLLUTANTS

For many years, the main goal of treating municipal wastewater was simply to reduce its content of suspended solids, oxygen-demanding materials, dissolved inorganic compounds, and harmful bacteria. In recent years, however, more stress has been placed on improving means of disposal of the solid residues from the municipal treatment processes. The basic methods of treating municipal wastewater fall into three stages: primary treatment, including grit removal, screening, grinding, and sedimentation; secondary treatment, which entails oxidation of dissolved organic matter by means of using biologically active sludge, which is then filtered off; and tertiary treatment, in which advanced biological methods of nitrogen removal and chemical and physical methods such as granular filtration and activated carbon absorption are employed. The handling and disposal of solid residues can account for 25–50% of the capital and operational costs of a treatment plant. The characteristics of industrial wastewater can differ considerably both within and among industries. The impact of industrial discharges depends not only on their collective characteristics, such as biochemical oxygen demand and the amount of suspended solids, but also on their content of specific inorganic and organic substances. Three options are available in controlling industrial wastewater. Control can take place at the point of generation in the plant; wastewater can be pretreated for discharge to municipal treatment sources; or wastewater can be treated completely at the plant and either reused or discharged directly into receiving waters. Figure 2 illustrates the treatment procedure of polluted water.

Raw sewage includes waste from sinks, toilets, and industrial processes. Treatment of the sewage is required before it can be safely buried, used, or released back into local water systems. In a treatment plant, the waste is passed through a series of screens, chambers, and chemical processes to reduce its bulk and toxicity. The three general phases of treatment are primary, secondary, and tertiary. During primary treatment, a large percentage of the suspended solids and inorganic material is removed from the sewage. The focus of secondary treatment is reducing organic material by accelerating natural biological processes. Tertiary treatment is necessary when the water will be reused; 99% of solids are removed and various chemical processes are used to ensure the water is as free from impurity as possible.

INDIAN SCENARIO OF SURFACE WATER POLLUTION

Water Stress and Water Scarcity in India

In India, because of uneven distribution of rainfall, the available per capita water resources have been varying in different river basins. As a result, despite good annual rainfall, some river basins fall in the category of water-scarce and water-stressed regions, whereas many others suffer from absolute scarcity. The water availability in any region or country is reflected by “water stress index” (4). This index is based on the minimum per capita water required for basic household needs and to maintain good health. A region whose renewable fresh water availability is below 1700 cubic meters/capita/annum is a “water stress” region, whereas the one whose availability falls below 1000 cubic meters/capita/annum experiences “water scarcity.” The current national average per capita water

Table 1. Effects of Water Quality Parameters for Drinking Purposes

S. No.	Parameter	Prescribed Limits IS:10500, 1991 Desirable	Permissible Limit	Probable Effects
1.	Color (hazen unit)	5	25	Makes water aesthetically undesirable
2.	Odor	Essentially free from objectionable odor		Makes water aesthetically undesirable
3.	Taste	Agreeable		Makes water aesthetically undesirable
4.	Turbidity (NTU)	5	10	High turbidity indicates contamination/pollution
5.	PH	6.5	8.5	Indicative of acidic or alkaline waters, affects taste, corrosivity, and the water supply system
6.	Hardness as $CaCO_3$(mg/l)	300	600	Affects water supply system (scaling). Excessive soap consumption, calcification of arteries. There is no conclusive proof, but it may cause urinary concretins, diseases of kidney or bladder, and stomach disorders
7.	Iron (mg/l)	0.03	1.00	Gives bitter sweet astringent taste. Causes staining of laundry and porcelain. In traces, it is essential for nutrition
8.	Chloride (mg/l)	250	1000	May be injurious to some people suffering from diseases of heart or kidneys. Taste, indigestion, corrosion, and palatability are affected
9.	Residual Chlorine (mg/l) only when water is chlorinated	0.20	—	Excessive chlorination of drinking water may cause asthma, colitis, and exzema
10.	Total Dissolved Solid (mg/l) (TDS)	500	2000	Palatability decreases and may cause gastro-intestinal irritation in humans; may have laxative effect particularly on transits and corrosion; may damage water system
11.	Calcium (Ca) (mg/l)	75	200	Causes encrustation in water supply system. Although insufficiency causes a severe type of rickets, excess causes concretions in the body such as kidney or bladder stones and irritation in urinary passages. (Essential for nervous and muscular system, cardiac functions, and in coagulation of blood)
12.	Magnesium (Mg) mg/l	30	100	Its salts are cathartics and diuretic. High conc. may have laxative effect particularly on new users. Magnesium deficiency is associated with structural and functional changes. It is essential an activator of many enzyme systems
13.	Copper (Cu) mg/l	0.05	1.50	Astringent taste but essential and beneficial element in human metabolism. Deficiency results in nutritional anemia in infants. Large amount may result in liver damage, cause central nervous system irritation, and depression. In water supply, it enhances corrosion of aluminium in particular
14.	Sulphate (SO_4) mg/l	200	400	Causes gastro-intestinal irritation along with Mg or Na, can have a cathartic effect on users, concentration of more than 750 mg/l may have laxative effect along with Magnesium
15.	Nitrate (NO_3) mg/l	45	100	Cause infant methaemoglobinaemia (blue babies) at very high concentration, causes gastric cancer and adversely affects central nervous system and cardiovascular system
16.	Fluoride (F) mg/l	1.0	1.50	Reduces dental carries, very high concentration may cause crippling skeletal fluorosis

Table 1. (*Continued*)

S. No.	Parameter	Prescribed Limits IS:10500, 1991 Desirable	Prescribed Limits IS:10500, 1991 Permissible Limit	Probable Effects
17.	Cadmium (Cd) mg/l	0.01	—	Acute toxicity may be associated with renal, arterial hypertension, itai-itai disease (a bone disease). Cadmium salt causes cramps, nausea, vomiting, and diarrhea
18.	Lead (pb) mg/l	0.05	—	Toxic in both acute and chronic exposures. Burning in the mouth, severe inflammation of the gastrointestinal tract with vomiting and diarrhea; chronic toxicity produces nausea, several abdominal pain, paralysis, mental confusion, visual disturbances, anemia, etc.
19.	Zinc (Zn) mg/l	5	15	An essential and beneficial element in human metabolism. Taste threshold for Zn occurs at about 5 mg/l, imparts astringent taste to water
20.	Chromium (Cr + 6) mg/l	0.05	—	Hexavelant state of chromium produces lung tumors, can produce cutaneous and nasal mucous membrane ulcers and dermatitis
21.	Boron (B) mg/l	1.00	5.00	Affects central nervous system, its salt may cause nausea, cramps, convulsions, coma, etc.
22.	Alkalinity mg/l $CaCO_3$	200	600	Imparts distinctly unpleasant taste, may be deleterious to humans in presence of high pH, hardness, and total dissolved solids
23.	Pesticides mg/l	Absent	0.001	Imparts toxicity and accumulates in different organs of human body affecting immune and nervous systems, may be carcinogenic
24.	Phosphate (PO_4) mg/l	No guideline		High concentration may cause vomiting and diarrhea; stimulate secondary hyperthyroidism and bone loss
25.	Sodium (Na)	No guideline		Harmful to persons suffering from cardiac, renal, and circulatory diseases
26.	Potassium (K) mg/l	No guideline		An essential nutritional element, but in excess is cathartic
27.	Silica (SiO_2) mg/l	No guideline		—
28.	Nickel (Ni) mg/l	No guideline		Nontoxic element but may be carcinogenic in animals; can react with DNA resulting in DNA damage in animals
29.	Pathogens (a) total coliform (per 100 ml) (b) Fecal coliform (per 100 ml)	1	10	Causes water borne diseases like coliform, jaundice, typhoid, cholera, etc.; produces infections involving skin mucous membrane of eyes, ears, and throat

availability figure per annum is 2464 cubic meters, implying that India is not yet in the "water stress" range. However, this is only the national average figure. There are several parts of India that are water stressed (Fig. 3)—the regions in the Indus, Krishna, and Ganga subbasins. Regions under east-flowing rivers between Mahanadi and Pennar and west-flowing rivers of Kachchh and Kathiawar are experiencing water scarcity, whereas the regions under east-flowing rivers between Pennar and Kanyakumari are suffering with absolute water scarcity (situation where the per capita availability falls below 500 cubic meters/annum). The per capita water availability here falls as low as 411 cubic meters (5), as shown in Fig. 3. The annual rainfall received by India is unevenly distributed across its different parts and across different times of the year. As a result despite good annual rainfall, some river basins fall in the category of water-scarce and water-stressed regions, whereas many others suffer from absolute scarcity.

Surface Water Pollution in India

River Water Pollution. In addition to water scarcity and water stress, many rivers are being used indiscriminately for disposal of municipal, industrial, and agricultural wastes, thereby polluting the river water beyond the permissible limits; Fig. 4 shows major rivers.

As can be seen from Fig. 2, the river water is unfit even for irrigation purposes at some places. Thus, it has

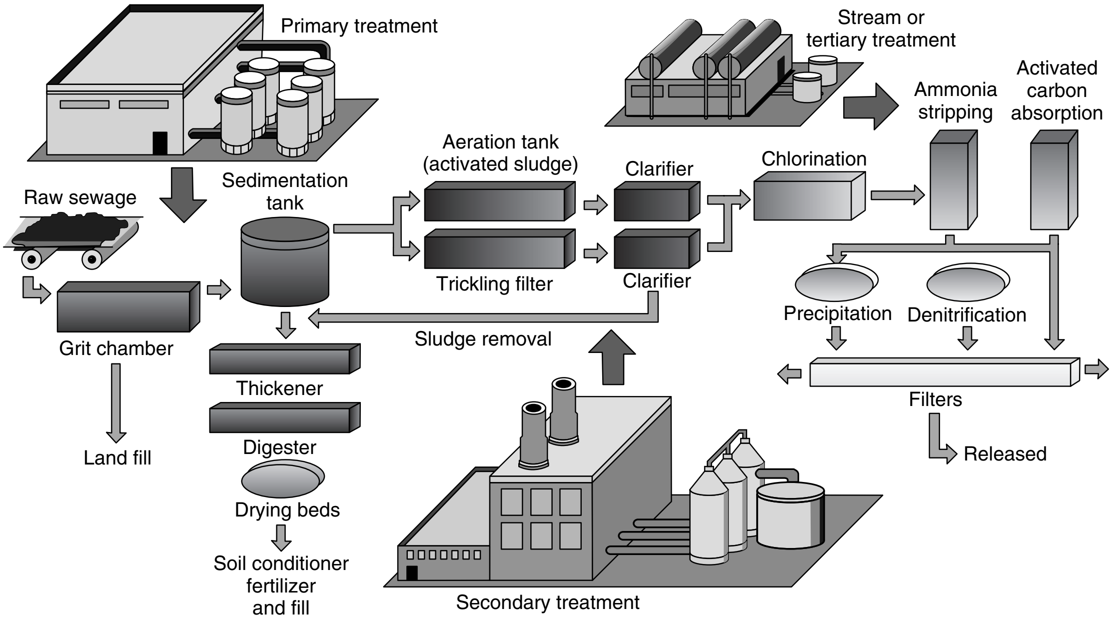

Figure 2. Wastewater treatment.

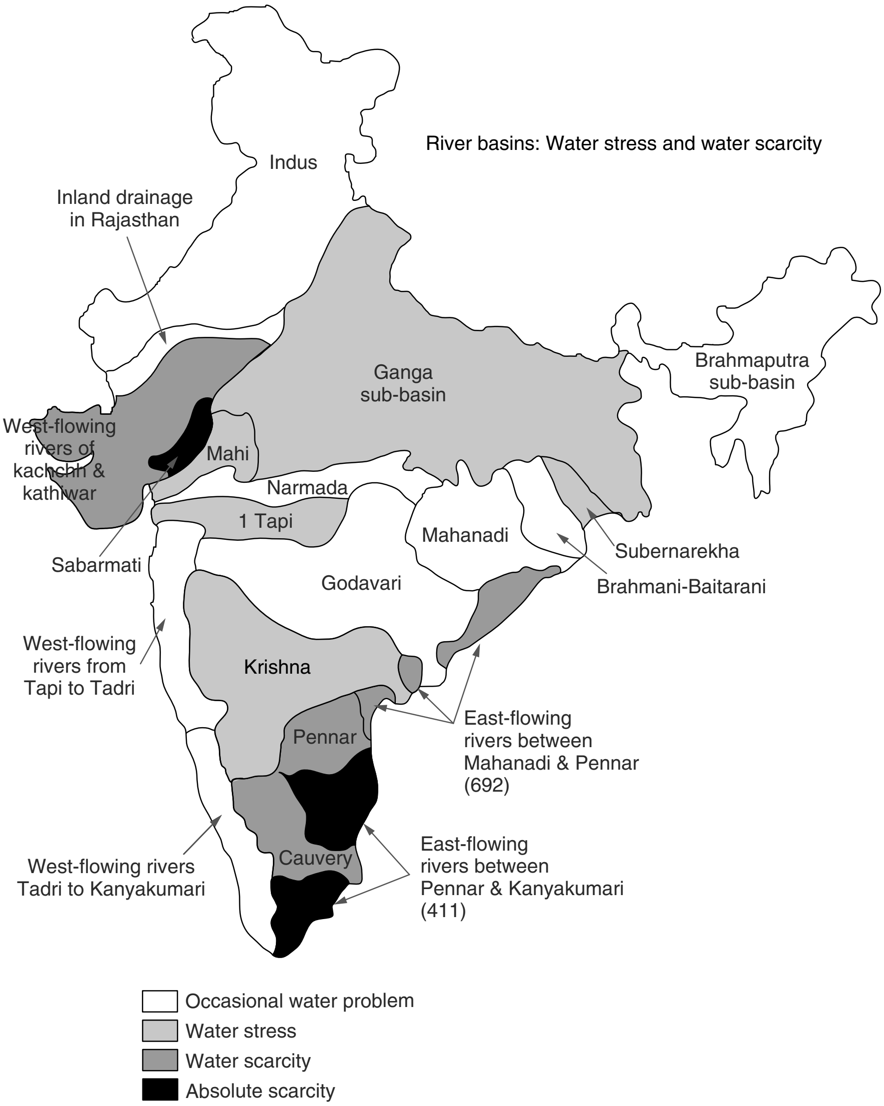

Figure 3. Water stress and water scarcity in India.

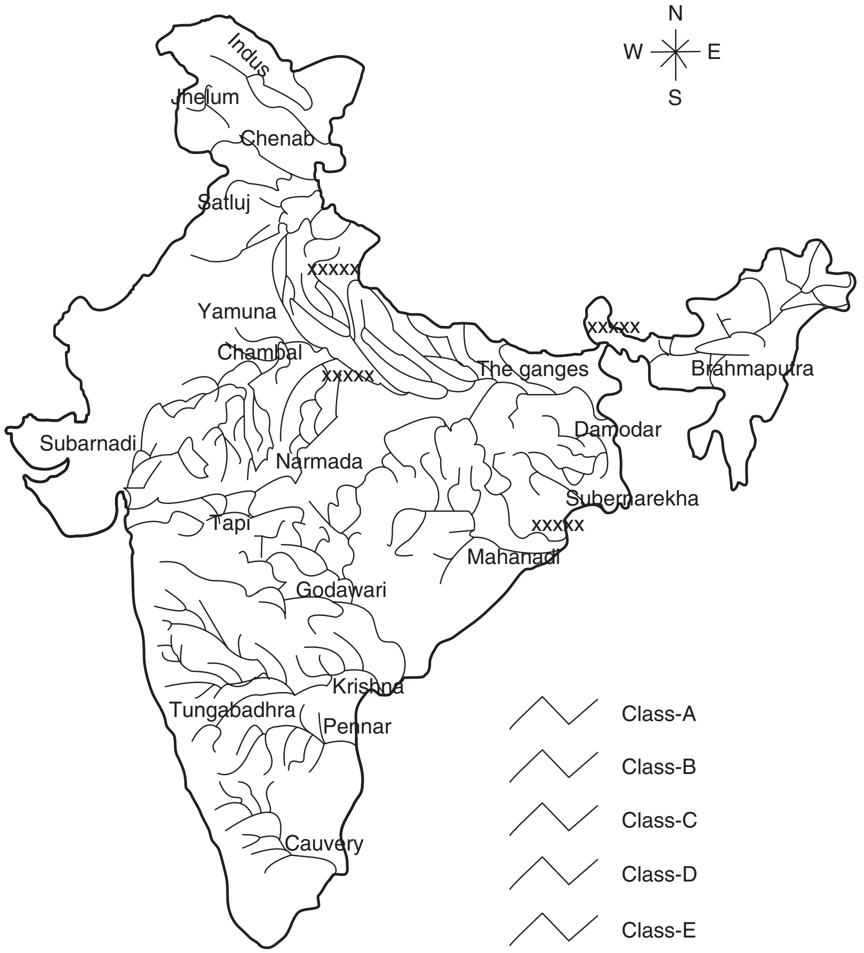

Figure 4. Generalized pattern of major rivers in India. Rivers will all be classified A to E.
Class A: Fit for drinking after proper disinfection with the addition of chlorine or bleaching powder.
Class B: Fit for bathing.
Class C: Fit for drinking only after proper treatment (screening to remove physical matters or particulate such as paper, plastic etc.
Class D: This is fit for fish and wildlife.
Class E: Suitable only for industrial cooling, irrigation, etc.

become absolutely essential to evaluate the environmental impacts of surface water resources to minimize the progressive deterioration in water quality. Water quality issues needing to be addressed with respect to rivers are discussed below.

Change in Physical Characteristics. Temperature, turbidity, and total suspended solids (TSS) in rivers can be greatly affected by human activities such as agriculture, deforestation, and the use of water for cooling. For example, the upward trend in soil erosion and the related increase in TSS in rivers can be seen in most of the mountainous regions in India (6).

Contamination by Fecal and Organic Matter. In India, fecal contamination is still the primary water quality issue in rivers, especially where human and animal wastes are not adequately collected and treated. Although this applies to both rural and urban areas, the situation is probably more critical in fast-growing cities.

The release of untreated domestic or industrial wastes high in organic matter into rivers results in a marked decline in oxygen concentration (sometimes resulting in anaerobic conditions) and an increase in ammonia and nitrogen concentrations downstream of the effluent input. The most obvious effect of the release of organic matter along the length of the river is the depletion of oxygen downstream of the discharge as shown by the so-called “oxygen-sag curve,” which plots dissolved oxygen concentration against distance. Industrial activities that discharge large organic loads include pulp and paper production and food processing. Fecal matter affects the use of water for drinking water source or bathing water, as well as ecological health of river.

Toxic Pollutants: Organics and Heavy Metals. Organic pollutants (mostly chemicals manufactured artificially by man) are also becoming an important water quality issue. Uncontrolled discharge of industrial wastewater often causes pollution because of toxic metals. Other sources of metal pollution are leachates from urban solid waste landfills and mining waste dumps.

Rivers such as the Yamuna, which pass through large towns and cities, are often badly affected with organic pollutants. Another example is that of Damodar River, which is polluted with heavy metals developing mostly from electroplating, tanning, and metal-based industries (6).

River Eutrophication. During the 1950s and 1960s, eutrophication (nutrient enrichment leading to increased plant arid algae growth) was observed mostly in lakes and reservoirs. Since the 1970s, the increasing levels of phosphates and nitrates entering rivers, particularly

Table 2. Grossly Polluted and Less Polluted Stretches of Some Major Rivers

Basin	River/ Tributary	Polluted Stretch Class	Existing Class	Desired Parameters	Critical
		I. Grossly Polluted Stretches			
Ganga	Yamuna	(i) Delhi to confluence with Chambal	Partly D	C	DO, BOD, coliforms
		(ii) In the city limits of Delhi, Agra, and Mathura	Partly E -do-	B	-do-
	Chambal	D/s of Nagda and D/s Kota (approx 15 km for both places)	Partly D Partly E	C	BOD,DO
	Damodar	D/s of Dhanbad to Haldia	Partly D Partly E	C	BOD, Toxic
	Gomti	Lucknow to confluence with Ganga	Partly D Partly E	C	DO, BOD, coliforms
	Kali	D/s of Modinagar to confluence with Ganga	Partly D Partly E	C	-do-
	Khan	(i) In the city limits of Indore	E	B	-do-
		(ii) D/s of Indore	E	D	-do-
	Kshipra	(i) In the city limits of Ujjain	E	B	-do-
		(ii) D/s of Ujjain	E	D	-do-
	Hindon	Saharanpur to confluence with Yamuna	E	D	DO, BOD, Toxic
Godavari	Godavari	(i) D/s ofNasik to Nanded	Partly D	C	BOD
		(ii) City limits of Nasik and Nanded	Partly E -do-	B	BOD
Krishna	Krishna	Karad to Sangli	Partly D Partly E	C	BOD
Subemarekha	Subemarekha	Hathi dam to Baharagora	Partly D Partly E	C	DO, BOD, coliforms
Sabarmati	Sabarmati	Immediate upstream of Ahmedabad city up to Sabarmati Ashram	E	B	DO,BOD, coliforms
		II. Less Grossly Polluted Stretches			
Ganga	Betwa	Between Vidisha and Mandideep and Bhopal	D	C	BOD, Total coliforms
Krishna	Krishna	Dhom dam to Narso Babri (Mah)	D	C	BOD and coliforms
		Tributary streams	D	C	-do-
		Up to Nagarjunasagar dam	D	C	-do-
		From Nagarjunasagar dam to upstream of Repella (AP)	D	C	-do-
	Bhadra	Origin to downstream of KICCL of Bhadra dam (Karnataka)	D	C	Total coliforms
	Tunga	Thirthahalli to confluence with Bhadra	C	B	Total coliforms
Cauvery	Cauvery	i) From Talakaverito 5 km of Mysore District BorderYagni (Karnataka)	C	A	Total coliforms DO, BOD, coliforms BOD, Total coliforms Total coliforms
		ii) From KR Sagar Dam to Hogenekkal (Karnataka)	E	C	
		iii) From Pugalur to Grand Anicut (Tamil Nadu)	E	C	
		iv) Grand Anicut to Kumbhakonam (Tamil Nadu)	E	C	
Brahmani	Baitami	Upstream of Chandbali	D	B	BOD and coliforms
Baitami	Brahmani	Upstream of Dharmshalla	D	B	BOD and coliforms
Tapi	Tapi	From city limits of Nepanagar to the city limits of Burhanpur (MP)	E	A	DO,BOD

Source: Central Pollution Control Board, New Delhi.

in developed countries, were largely responsible for eutrophication occurring in running waters. In India, isolated reports have appeared for some river reaches, especially in plains around agriculture tracts of land.

In small rivers, eutrophication is said to promote macrophyte (large plants) development, whereas in large rivers, phytoplankton (algae) are usually more dominant than macrophyte. In such situations, the chlorophyll concentration of the water may reach extremely high values because of the fact that this pigment is present in all plants. Eutrophication can result in marked variations in dissolved oxygen arid pH throughout the day. The changes in water quality caused by eutrophication can be a major cause of stress to fish because of the release, at high pH, of highly toxic gaseous ammonia and depletion of oxygen after sunshine hours.

Salinization. Industrial and mining waste pollution results in increase in specifications. Evaporation, however, increases the concentration of all ions.

Table 2 shows the water quality of grossly polluted and less polluted major rivers of India. It is found that most of the rivers lie under the category C, which is suitable only for irrigation purposes.

Lake Water Pollution. Lakes serve as traps for pollutants carried by rivers and groundwater draining the watershed. The pollutant concentration in the lake usually builds up because of evaporation of water from the lake's surface unless a natural flushing with good quality water occurs.

Eutrophication. Simply speaking, eutrophication is the biological response to excess nutrient input to a lake. The production of biomass and its death and decay results in a number of effects, which individually and collectively result in impaired water use. The most important of these effects are decreased dissolved oxygen levels, release of odorous compounds (e.g., H_2S), and siltation.

Many important lakes in India [e.g., Hussein Sagar (Hyderabad), Nainital (Uttar Pradesh) and Dal (Jammu and Kashmir)] have reportedly progressed to advanced eutrophication levels (6).

Lake Acidification. One of the major issues related to lakes in particular, and to freshwaters in general, is the progressive acidification associated with deposition of rain and particulates (wet and dry deposition) enriched in mineral acids. The problem is characteristic of lakes in specific regions of the world that satisfy two major critical conditions: the lakes must have soft water (i.e., low hardness, conductivity, and dissolved salts) and be subjected to "acid rain."

To date, lake acidification has not been reported as a problem in India.

Bioaccumulation and Biomagnification. The processes of bioaccumulation and biomagnification are extremely important in the distribution of toxic substances (discharged in waste effluents) in fresh water ecosystems. The concentration of pollutants within the organism because of bioaccumulation and biomagnification depends on the duration of exposure of the organism to the contaminated environment and its trophic level in the food chain. Several fold increases in trace contaminant concentrations have been commonly observed in lakes and estuarine environments.

BIBLIOGRAPHY

1. Gleick, P.H. (1993). *Water in Crisis: A Guide to the World's Freshwater Resources*. Oxford University Press, New York.
2. Chapman, D. (2001). *Water Quality Assessment- A Guide to the Use of Biota, Sediments and Water in Environmental Monitoring*, 2nd Edn. E & FN SOPN.
3. World Bank. (1992). *World Development Report 1992*. Oxford University Press, New York.
4. Falkenmark, M. and Widstrand, C. (1992). *Population and Water Resources: A Delicate Balance*. Population Bulletin, Population Reference Bureau.
5. Chitale, M.A. (1992). Population and water resources of India. In: *Science, Population and Development*. V. Gowariker (Ed.). Unmesh Publications, Pune, p. 448.
6. Ministry of Water Resources, Government of India. (2004). Available: http://wrmin.nic.in/.

READING LIST

Center for Science and Environment. (1985). *The State of India's Environment 1984–85: The Second Citizens' Report*. Centre for Science and Environment, New Delhi, p. 394.

Jha, R., Ojha, C.S.P., and Bhatia, K.K.S.B. (2005). Dilution-mixing zones and design flows. *The Encyclopedia of Water*. John Wiley & Sons, New York.

Pawlow, J.R. (1983). *Technical Guidelines Manual for Performing Waste Load Allocations*. U.S. Environmental Protection Agency, Washington, DC.

SURFACE RUNOFF AND SUBSURFACE DRAINAGE

Joseph Holden
University of Leeds
Leeds, United Kingdom

Runoff is the movement of water across or through the land surface. Sometimes, the word runoff is used simply to refer to river discharge. However, there are many different types of runoff, including overland flow (surface runoff) or throughflow (subsurface runoff, sometimes also called interflow). Figure 1 illustrates the main hillslope runoff pathways for water. Precipitation can either hit the surface of the hillslope directly or be intercepted by vegetation. This intercepted water can be stored on leaves and tree trunks which shelter the ground beneath, or it can trickle down to reach the surface via stem flow. There are then two possibilities for direct precipitation or stem flow, once it reaches the hillslope, either to infiltrate into the soil or to pond up and flow over the surface as overland flow. Water that enters the soil may percolate down or move laterally

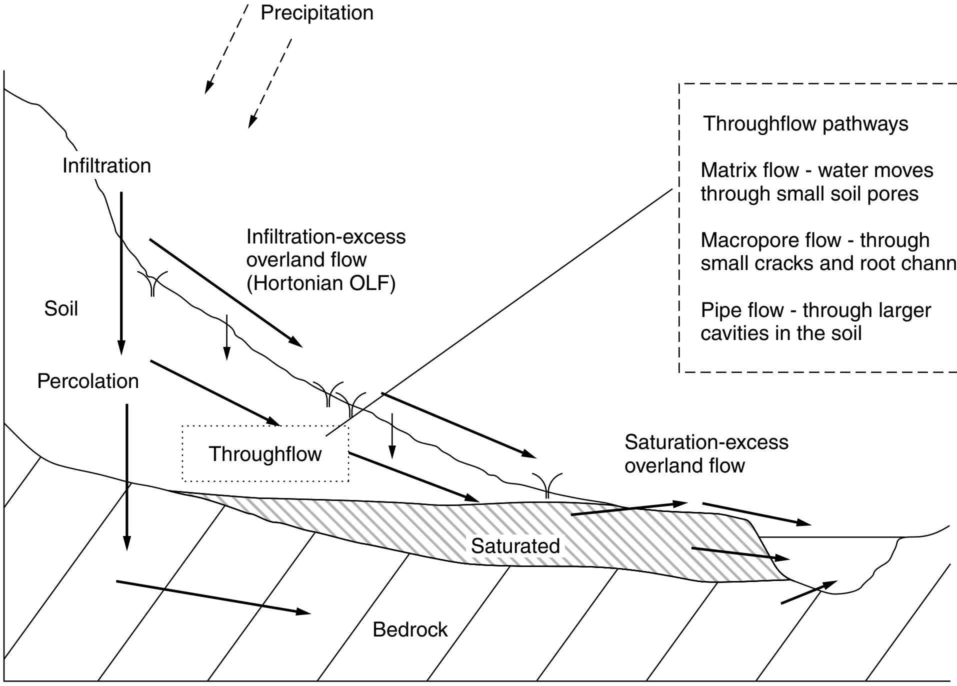

Figure 1. Main hillslope runoff pathways.

through the soil as throughflow, or it may return to the surface where the soil becomes saturated. The following sections provide details on each of the main interlinked overland flow and subsurface drainage mechanisms.

INFILTRATION-EXCESS OVERLAND FLOW

If the surface water supply is greater than the rate of infiltration into the soil, then surface storage will occur (even in urban catchments in small surface depressions). When the surface depressions are filled, they will start to overflow; this is called Hortonian overland flow (OLF) or infiltration-excess OLF. In Horton's theory, infiltration divides rainfall into two parts. One part goes via OLF to the stream channel; the other goes initially into the soil and then through groundwater flow to the stream or into groundwater storage or is lost by evapotranspiration. The dominance of Horton's (1) theory meant that research into subsurface flow mechanisms was neglected until the 1970s as it was assumed that infiltration-excess OLF was the major contributor to fast responding flow in rivers (2).

In many temperate areas of the world, infiltration-excess OLF is a rare occurrence except in urban locations. The infiltration capacity of many soils is too high to produce infiltration-excess overland flow. Infiltration-excess OLF is more likely in semiarid areas where soil surface crusts have developed and rainfall can be particularly intense. Often infiltration-excess OLF will occur only on spatially localized parts of a hillslope such as in tractor wheelings on arable land. This spatially localized infiltration-excess OLF is known as the partial contributing area concept (3). This suggests that only parts of the catchment or hillslope contribute infiltration-excess OLF, rather than the whole catchment, as Horton had originally suggested.

SATURATION-EXCESS OVERLAND FLOW

When water infiltrates into a soil, it fills up the available pore spaces. When all the pore spaces are full, the soil is saturated, and the water table is at the surface. Therefore, any extra water has difficulty entering the soil because it is saturated. Hence, OLF occurs. This type of OLF is known as saturation-excess overland flow. Saturation-excess OLF can occur at much lower rainfall intensities than required to generate infiltration-excess OLF. Saturation-excess OLF can occur even when it is not raining. This might happen, for example, at the foot of a hillslope. Water draining through the soil is known as throughflow. Throughflow from an upslope can fill up the soil pores at the bottom of the slope, and so the soil becomes saturated (Fig. 1). Any extra water is then forced out onto the surface to become OLF. This water, known as 'return flow,' is a component of saturation-excess OLF. Therefore saturation-excess OLF is more likely to occur at the bottom of a hillslope or on shallow soils where there is restricted pore space for water storage.

The area of a catchment or hillslope that produces saturation-excess OLF varies through time. During wet seasons, for example, more of a catchment or hillslope will be saturated and therefore can generate saturation-excess OLF than during dry seasons. During a rainfall, if the catchment starts off relatively dry, then not much of the area will generate saturation-excess OLF; but as rainfall continues, then, more of the catchment becomes saturated, especially in the valley bottoms, and therefore a larger area of the catchment will produce saturation-excess OLF. The fact that the area of a catchment in which saturation-excess OLF occurs tends to vary is known as the 'variable source area concept' (4). The variable source area model has become the dominant concept in catchment hydrology (2).

Saturation-excess OLF is a much more important process than infiltration-excess OLF across temperate zones. The main differences between the two OLF types are related to the water flow paths. For infiltration-excess OLF, all of the flow is fresh rainwater that has not been able to infiltrate into the soil. However, saturation-excess OLF is often a mixture of water that has been inside the soil (return flow) and fresh rainwater that reaches the hillslope surface.

THROUGHFLOW

If water infiltrates into soil, several things can happen; it can be taken up by plants and transpired (or be lost from the soil by evaporation); it can continue to percolate down into the bedrock; or it can travel laterally through the soil or rock—this is called throughflow.

Most water reaches rivers worldwide either by throughflow, through the soil layers or through bedrock. Throughflow can maintain low flows (baseflow) in rivers by slow subsurface drainage of water and can also contribute to peak flows (storm flow) by generating saturation-excess OLF and as an important process in its own right (5). There are different ways water can move through soil as throughflow, and this impacts the timing of water delivery to the river channel. Soils are not uniform deposits; they have cracks and fissures within them. Water can move through the very fine pores of soil as matrix flow, or it can move through larger pores called macropores (macropore flow), or even larger cavities called soil pipes (pipe flow). Water that moves through the soil matrix occurs in a laminar fashion whereas flow within macropores and pipes is turbulent.

Matrix Throughflow

It is possible to measure throughflow and overland flow on hillslopes by digging a trench across the slope and intercepting and collecting the flow from different soil layers. However, this is not always feasible, and disturbance of hillslope soils is not desirable. Flow through the matrix of a porous substance such as soil that occurs in a laminar fashion should behave according to Darcy's Law. This allows us to calculate the likely rate of water movement through a porous medium when it is saturated (saturated hydraulic conductivity). Thus, it is possible to estimate the amount of flow taking place as matrix throughflow. Often, the saturated hydraulic conductivity will vary with depth and soil type. Sandy soils typically have high hydraulic conductivity compared to clay soils that have low hydraulic conductivity. Therefore, water will drain through sandy soils more quickly than through clay soils. Lateral throughflow through the soil matrix will occur in any soil in which the hydraulic conductivity declines with depth. If both soil and bedrock remain permeable at depth, however, then percolation remains vertical and little lateral flow can occur; infiltrating water will only recharge groundwater storage (5).

Water also moves through soils on hillslopes that are unsaturated. Even after a long drought, most soils contain some water. This suggests that gravitational drainage and evapotranspiration are not the only forces at work in moving water within soils and that other forces involved must be very strong (2). It is possible to calculate how the interacting forces of gravity, matric potential (soil water tension), positive pore water pressure, and other forces that operate balance out. Once this soil water energy balance has been determined (by measuring pore water pressure and relative altitude, for example), it is possible to map spatially the directions of subsurface water movement. Water in different parts of the soil can move in different directions due to capillary, gravitational, and other forces. Often, it is possible to show that throughflow water preferentially flows into hillslope hollows, and thus these areas become more important contributors to subsurface flow and saturation-excess OLF.

Macropore and Pipe Throughflow

Macropores are pores larger than 0.1 mm in diameter, and soil pipes are those larger than 1 mm in diameter that transport water through the soil. Macropores and pipes can promote rapid, preferential transport of water and chemicals through the soil due to their size and also because they are connected and continuous across sufficient distances to bypass agriculturally and environmentally important soil layers (6). Water may drain more quickly through macropore and soil pipe networks than through the soil matrix. Darcy's Law does not provide a good estimate of subsurface drainage rates if a soil contains many hydrologically active soil pipes or macropores because it is valid only for nonturbulent flows. Macropores may not take up much space in the soil, and if they are open at the soil surface, they will often take up only a tiny proportion of the soil surface. Despite their small spatial role, macropores can still have a high impact on runoff and play a large role in throughflow drainage as water can preferentially flow through them. A study in Niger on a crusted sandy soil showed that 50% of infiltrated water moved through macropores (7). Some studies in upland peat catchments have indicated that 30% of throughflow moves through macropores (8) and 10% through soil pipes (9).

Groundwater Flow

Groundwater is water held below the water table in soil and rock. Therefore, groundwater flow has, to some extent, already been discussed before. However, further treatment of groundwater flow as a separate component is required because of its worldwide importance. In many catchments, water is supplied to the stream from groundwater in the bedrock. This is water that has percolated down through the overlying soil and entered the bedrock. Rock has small pores, fractures, and fissures. Therefore, it is possible to use Darcy's Law to investigate flow rates through bedrock where fissures are at a minimum. Where there are large fractures such as in cavernous limestone areas, then it may not be so useful. Groundwater may be a large store of water, but for it to be available to supply river flow, the holding material (rock or soil) needs to be not just porous but permeable. That is to say that a rock (or soil) may be porous but relatively impermeable either because the pores are not connected or because they are so small that water can only be forced through them with difficulty. Conversely a rock that has no voids except one or two large cracks will have low porosity and therefore a poor store of water, but because water can pass easily through the cracks, the permeability will be high (5).

RUNOFF PROCESS CONTROLS OF RIVER FLOW

Surface and subsurface flow processes are important controls on river flow and the response of a catchment to precipitation. The dominance of different types of runoff processes is controlled by local climate and catchment features such as geology, topography, soils, and vegetation. Land management can influence runoff processes by changing the infiltration capacity of the soil surface, altering the internal structure of a soil, or by changing the catchment water balance (e.g., through deforestation). Depending on the nature of the aquifer, baseflow discharge may be uniform throughout the year, or peak discharge may lag significantly behind precipitation inputs (2). Some catchments lack any significant groundwater storage and therefore have limited baseflow.

The occurrence of hillslope flow processes in the catchment and their relative dominance affect the speed at which water is delivered to a stream. Overland flow is typically much faster than subsurface drainage. This is why urbanization can lead to increased flood risk downstream. Surfaces created by urbanization reduce infiltration capacity and promote the formation of infiltration-excess OLF. Matrix flow is slower than macropore flow which, in turn, is a form of drainage slower than pipe flow. Many soil pipe flow waters can have velocities equal to that of overland flow.

Where infiltration-excess overland flow dominates the hillslope runoff response, then a river hydrograph is likely to have a short lag time and high peak flow. If throughflow in the small soil pores (matrix flow) dominates the runoff response on the hillslopes, then the hydrograph may have a much lower peak and longer lag times. However, throughflow contributes to saturation-excess OLF so throughflow can still lead to rapid and large flood peaks. In some soils, only a small amount of infiltration may be needed to cause the water table to rise to the surface. There may even be two river discharge peaks caused by one rainfall, where the first peak is saturation-excess overland flow and some precipitation directly in the channel and the second peak is much longer and larger caused by subsurface throughflow accumulating at the bottom of hillslopes and valley bottoms before entering the stream channel. Throughflow may also contribute directly to storm hydrographs by a mechanism called piston or displacement flow. This is where soil water at the bottom of a slope (old water) is rapidly pushed out of the soil by new fresh infiltrating water entering at the top of a slope. Determination of whether the water is old water being pushed out of a slope or new water can be done hydrochemically (by comparing the water to precipitation chemistry) or by using dyes. Where surface saturation occurs to any great extent such as on wide valley bottoms, saturation-excess OLF will dominate the flow response with higher peak discharges and lower lag times than are characteristic of throughflow storm contributions. Where soils are fairly impermeable through much of their profile, surface saturation may be extensive, and much more rainfall is translated into runoff. Where permeable soils overlie impermeable bedrock, however, throughflow can sometimes account for most storm discharge (5).

Runoff processes are by no means independent of one another, and water travelling across the surface at one point may later take the form of subsurface flow through the matrix and then flow through macropores before being returned to the soil matrix, for example. Hillslope runoff is dynamic, and the dominance of runoff production processes varies across both time and space.

BIBLIOGRAPHY

1. Horton, R.E. (1933). The role of infiltration in the hydrological cycle. *Trans. Am. Geophys. Union* **14**: 446–460.
2. Holden, J. (Ed.). (2004). *Introduction to Physical Geography; Processes, Landscapes and Environmental Change*. Pearson, Harlow.
3. Betson, R.P. (1964). What is watershed runoff? *J. Geophys. Res.* **69**: 1541–1552.
4. Hewlett, J.D. (1961). Watershed management. In: *Report for 1961 Southeastern Forest Experiment Station*. US Forest Service, Ashville, NC, pp. 62–66.
5. Burt, T.P. (1996). The hydrology of headwater catchments. In: *River Flows and Channel Forms*. G.E. Petts and P. Calow (Eds.). Blackwell, pp. 6–31.
6. Beven, K.J. and Germann, P. (1982). Macropores and water flow in soils. *Water Resour. Res.* **18**: 1311–1325.
7. Leonard, J., Perrier, E., and de Marsily, G. (2001). A model for simulating the influence of a spatial distribution of large circular macropores on surface runoff. *Water Resour. Res.* **37**: 3217–3225.
8. Holden, J., Burt, T.P., and Cox, N.J. (2001). Macroporosity and infiltration in blanket peat: the implications of tension disc infiltrometer measurements. *Hydrological Processes* **15**: 289–303.
9. Holden, J. and Burt, T.P. (2002). Piping and pipeflow in a deep peat catchment. *Catena* **48**: 163–199.

TRACE ELEMENTS IN WATER, SEDIMENT, AND AQUATIC BIOTA—EFFECTS OF BIOLOGY AND HUMAN ACTIVITY

NITISH PRIYADARSHI
Ranchi University
Ranchi, Jharkhand, India

INTRODUCTION

Trace elements are defined here as those elements that generally occur in water at concentrations of less than 1 mg/L.

In recent years, the role of trace element studies in environmental chemistry is gaining attention. The may be porous but relatively impermeable either because the pores are not connected or because they are so small that water can only be forced through them with difficulty. Conversely a rock that has no voids except one or two large cracks will have low porosity and therefore a poor store of water, but because water can pass easily through the cracks, the permeability will be high (5).

RUNOFF PROCESS CONTROLS OF RIVER FLOW

Surface and subsurface flow processes are important controls on river flow and the response of a catchment to precipitation. The dominance of different types of runoff processes is controlled by local climate and catchment features such as geology, topography, soils, and vegetation. Land management can influence runoff processes by changing the infiltration capacity of the soil surface, altering the internal structure of a soil, or by changing the catchment water balance (e.g., through deforestation). Depending on the nature of the aquifer, baseflow discharge may be uniform throughout the year, or peak discharge may lag significantly behind precipitation inputs (2). Some catchments lack any significant groundwater storage and therefore have limited baseflow.

The occurrence of hillslope flow processes in the catchment and their relative dominance affect the speed at which water is delivered to a stream. Overland flow is typically much faster than subsurface drainage. This is why urbanization can lead to increased flood risk downstream. Surfaces created by urbanization reduce infiltration capacity and promote the formation of infiltration-excess OLF. Matrix flow is slower than macropore flow which, in turn, is a form of drainage slower than pipe flow. Many soil pipe flow waters can have velocities equal to that of overland flow.

Where infiltration-excess overland flow dominates the hillslope runoff response, then a river hydrograph is likely to have a short lag time and high peak flow. If throughflow in the small soil pores (matrix flow) dominates the runoff response on the hillslopes, then the hydrograph may have a much lower peak and longer lag times. However, throughflow contributes to saturation-excess OLF so throughflow can still lead to rapid and large flood peaks. In some soils, only a small amount of infiltration may be needed to cause the water table to rise to the surface. There may even be two river discharge peaks caused by one rainfall, where the first peak is saturation-excess overland flow and some precipitation directly in the channel and the second peak is much longer and larger caused by subsurface throughflow accumulating at the bottom of hillslopes and valley bottoms before entering the stream channel. Throughflow may also contribute directly to storm hydrographs by a mechanism called piston or displacement flow. This is where soil water at the bottom of a slope (old water) is rapidly pushed out of the soil by new fresh infiltrating water entering at the top of a slope. Determination of whether the water is old water being pushed out of a slope or new water can be done hydrochemically (by comparing the water to precipitation chemistry) or by using dyes. Where surface saturation occurs to any great extent such as on wide valley bottoms, saturation-excess OLF will dominate the flow response with higher peak discharges and lower lag times than are characteristic of throughflow storm contributions. Where soils are fairly impermeable through much of their profile, surface saturation may be extensive, and much more rainfall is translated into runoff. Where permeable soils overlie impermeable bedrock, however, throughflow can sometimes account for most storm discharge (5).

Runoff processes are by no means independent of one another, and water travelling across the surface at one point may later take the form of subsurface flow through the matrix and then flow through macropores before being returned to the soil matrix, for example. Hillslope runoff is dynamic, and the dominance of runoff production processes varies across both time and space.

BIBLIOGRAPHY

1. Horton, R.E. (1933). The role of infiltration in the hydrological cycle. *Trans. Am. Geophys. Union* **14**: 446–460.
2. Holden, J. (Ed.). (2004). *Introduction to Physical Geography; Processes, Landscapes and Environmental Change*. Pearson, Harlow.
3. Betson, R.P. (1964). What is watershed runoff? *J. Geophys. Res.* **69**: 1541–1552.
4. Hewlett, J.D. (1961). Watershed management. In: *Report for 1961 Southeastern Forest Experiment Station*. US Forest Service, Ashville, NC, pp. 62–66.
5. Burt, T.P. (1996). The hydrology of headwater catchments. In: *River Flows and Channel Forms*. G.E. Petts and P. Calow (Eds.). Blackwell, pp. 6–31.
6. Beven, K.J. and Germann, P. (1982). Macropores and water flow in soils. *Water Resour. Res.* **18**: 1311–1325.
7. Leonard, J., Perrier, E., and de Marsily, G. (2001). A model for simulating the influence of a spatial distribution of large circular macropores on surface runoff. *Water Resour. Res.* **37**: 3217–3225.
8. Holden, J., Burt, T.P., and Cox, N.J. (2001). Macroporosity and infiltration in blanket peat: the implications of tension disc infiltrometer measurements. *Hydrological Processes* **15**: 289–303.
9. Holden, J. and Burt, T.P. (2002). Piping and pipeflow in a deep peat catchment. *Catena* **48**: 163–199.

TRACE ELEMENTS IN WATER, SEDIMENT, AND AQUATIC BIOTA—EFFECTS OF BIOLOGY AND HUMAN ACTIVITY

NITISH PRIYADARSHI
Ranchi University
Ranchi, Jharkhand, India

INTRODUCTION

Trace elements are defined here as those elements that generally occur in water at concentrations of less than 1 mg/L.

In recent years, the role of trace element studies in environmental chemistry is gaining attention. The deficiencies or excess distribution of these in water and foodstuffs may pose serious regional health hazards and cause diseases in human beings and animals. Proper evaluation of the environmental impact of trace elements on humans, animals, plants, and aquatic biota requires a multidisciplinary approach involving physicians and geoenvironmental scientists.

TRACE ELEMENTS IN WATER

Trace elements enter the aquatic environment in small soil particles that are redistributed during runoff as suspended particles in streams, lakes, and coastal marine locations. Trace elements in water are important; they are toxic to humans and aquatic life in excess, and they accumulate in food chains. Trace element activity in aquatic systems depends on the element species formed by natural hydrodynamic, chemical, and biological forces.

The concentrations of trace elements in water are often much lower than would be expected from either equilibrium solubility calculations or supply to water from various sources. The most common reason for the low concentrations is adsorption of elements onto a solid phase.

To understand the behavior of any trace elements in waters, it is essential to know the chemical forms of the element in the water. Anomalously high concentrations are often related to the presence of stable complexes in solution (1).

Adsorption of manganese and iron oxides, in particular, is probably the most important process in maintaining the concentrations of trace elements at levels far below those predicted by equilibrium solubility calculations. In general, elements in solution as cationic species are more likely to be adsorbed than anionic species. An important exception is the strong adsorption of phosphate anions by ferric oxyhydroxides (1).

TRACE ELEMENTS IN SEDIMENTS

The significance of studies of sediments in the context of water pollution was realized in the 1960s. Publications show that trace element studies in sediments began in the United States in 1963, the Netherlands in 1966, the Soviet Union in 1967, the Federal Republic of Germany and Sweden in 1968, Poland in 1969, Hungary in 1970, Australia in 1972, Israel and Japan in 1973, Switzerland in 1974, the United Kingdom in 1975, Yugoslavia in 1976, and so on.

The study of sediments is all the more important because, on one hand sediments are potential carriers of elements from a watershed to the receiving body of water, and, on the other, they play a significant role in the attenuation of toxic trace elements in polluted aquatic environments.

Sediments commonly form major sinks for trace element contaminants released during both past and present geological, industrial, and other human activities. Commonly, trace elements released from industrial activity are adsorbed onto other mineral surfaces (2).

Most trace elements in sediments (mercury, lead, cadmium, zinc, copper, nickel, cobalt, manganese, chromium, and silver) are chemically bound and require heat and low pH to convert them to soluble form. Various physical, chemical, and biological factors (such as the scouring action of water currents, dredging, changes in chemical parameters and in microbial activity) can bring about the redistribution and partial solution of trace elements from sediments.

Clay minerals that comprise a significant part of sediments together with freshly precipitated iron hydroxides, amorphous silicic acids, and organic substances can sorb cations from solution and release equivalent amounts of other cations into solution. However, the behavior of different clay minerals toward the various trace elements is not yet fully understood. It is reported that the exchange capacity increases in the order:

$$\text{Kalonite} < \text{chlorite} < \text{illite} < \text{montmorillonite}$$

The affinity of trace elements toward clay minerals has been established as follows (3):

$$\text{Pb} > \text{Ni} > \text{Cu} > \text{Zn}$$

Further, it has been established that Cr and Cu favor bonding to illite whereas V, Co, and Ni favor bonding to montmorillonite.

A general increase of trace element concentration from coarse to fine-grained fractions has generally been noted. Trace element content in particulate matter is 1000 to 100,000 times higher than the corresponding trace element content in the associated aquatic phase. The presence of a large proportion of heavy minerals results in characteristic enrichment of trace elements, particularly in fine sand-silt fractions. On the other hand, the high proportions of light minerals such as quartz and feldspar tend to have the opposite effect on the sedimentary trace element budget (4).

Dahlberg Model

The variations in the trace element content of stream sediments can be expressed as a function of potential controlling factors by the following model of Dahlberg (5):

$$T = f(LHGCVMe)$$

where T = resulting trace element concentration, f = function of potential controlling factors, L = influence of lithologic units, H = hydraulic effects, G = geologic features, C = cultural (man-made) influences, V = type of vegetative cover, M = the effect of mineralized zones, and e = error plus effects of additional factors not explicitly defined in the model.

In this model, factors other than C create natural pollution that can be recorded as the background value. However, it is not always not possible to distinguish between contamination of industrial or domestic origin and natural pollution.

The trace elements immobilized in bottom sediments do not necessarily remain in that condition but may be released as a result of changes in the aquatic milieu. These are the main causes of remobilization:

1. changes in redox conditions
2. elevated salt concentrations
3. lowering of pH
4. increased use of natural and synthetic complexing agents that can form soluble metal complexes
5. microbial activity

TRACE ELEMENTS IN AQUATIC BIOTA

The role of trace elements in aquatic biota has been treated in earlier literature reviews with emphasis on the toxicity of the individual elements. Only since mercury and cadmium poisoning in Japan has the accent been shifted toward investigations dealing with the influence of trace elements on the metabolism of aquatic biota and the ability of the latter to accumulate both essential and nonessential elements. Essential trace elements are used by aquatic biota in only trace amounts, and a slight increase in the concentration in water could lead to a significant element increase in aquatic biota. Experiments on the impact of trace elements on aquatic biota have been carried out for many years (6–9).

Phytoplankton

The relative short life of phytoplankton (a few days to a few months) makes it very susceptible to changes in abiotic parameters. The element concentrations, therefore, can vary considerably within a short period of time. Knauer and Martin (10) established that the respective element concentrations for cadmium, copper, manganese, zinc, and lead underwent comparable changes in phytoplankton in Monterey Bay, California. The cadmium concentrations in the water distinctly decreased whereas at the same time, the cadmium content in phytoplankton increased. A similar observation was made for manganese.

Macroalgae

Macroalgae can also be used as indicators of elemental pollution. Their elemental content is directly related to the elemental concentrations in water. A detailed analysis of bladder wrack (*Fucus vesiculosus*) populations in coastal waters around Great Britain revealed significantly higher contamination in the eastern Irish Sea and North Sea by zinc, iron, manganese, copper, nickel, lead, silver, and cadmium than in other coastal waters surrounding Great Britain. The maximum values in the entire investigation area were 962 ppm zinc, 1517 ppm iron, 190 ppm manganese, 28.4 ppm copper, 18.0 ppm nickel, 9.0 ppm lead, 0.79 ppm silver, and 20.8 ppm cadmium (dry weight).

In the compiled data of macroalgae investigations, cadmium values are generally between 1 and 2 ppm. Copper concentrations vary widely. Minimum content is near 6 ppm. The lower concentrations of lead are between 2 and 3 ppm. Chromium appears in an average concentration of 2–3 ppm. The minimum concentrations of manganese lie between 6 and 10 ppm. Nickel has a concentration range similar to that of lead. A maximum of 72 ppb has been reported.

Mosses

It has generally been observed that mosses have a particular storage capacity for lead. Lead content of 277.4 ppm was found in *Fontinalis antipyretica* in the Elsenz River (Federal Republic of Germany). Samples taken from the heavily polluted Ruhr River had cadmium concentrations between 0.17 and 13.2 ppm, and the record amount of 13.2 ppm occurred in *Lagarosiphon major* from Lago Maggiore in Italy.

Higher plants

The roots of higher aquatic plants enable them to incorporate additional elements from sediment (interstitial water). Trace element enrichment differs from species to species. High arsenic enrichment was identified in different families of higher aquatic plants from the Waikito River, New Zealand. Maximum amounts up to 0.5 g arsenic per kg of dry mass were recorded. Different amounts of Cd, Zn, Pb, and Cu in the roots, stems, and leaves of different species were also identified. The highest concentrations were found in the leaves.

Zooplankton

High mercury concentration has been reported in zooplankton from different parts of the world. The maximum concentration recorded was 25.21 ppm Hg. The mercury concentrations in plankton samples gradually diminished with distance from the source. Zooplankton samples taken off the eastern coast of United States between Cape Hatteras and Cap Cod had an average mercury content of 0.5 ppm. From South African coastal waters, a maximum of 132 ppm of zinc was recorded. Cobalt varied from 0.004–0.26 ppm.

Trace Elements in Fish

The study of fish muscle tissue is one means for investigating the amount of elements entering humans by food chain enrichment and has therefore been investigated more than other organs. The absolute increase in trace elements in muscle tissue of contaminated fish is often much lower than that in other organs (11). A generally higher species-specific variation in element content cannot, therefore, always be determined by analyzing muscle tissue. Furthermore, it appears that the musculature becomes enriched by trace elements only when contamination is extremely high. In the musculature of the roach (*Rutilus rutilus*) from a section of the Neckar River, a 25-fold relative increase between minimum and maximum cadmium contamination was registered. In the liver, there was a 156-fold increase and in the kidney, a 196-fold increase (12). When contamination is severe, primarily the gills show elevated cadmium concentrations. The scales of Atlantic salmon (*Salmon salar*) and brown trout (*Salmon trutta*) were analyzed by Abdullah et al. (13) to determine element contamination. In part of the sample, a distinct correlation was established between concentrations in the scales and in the environment.

Of all the elements, mercury is known to be the most toxic. Several cases have been reported of mercury contamination in fish and shellfish. One of the most striking investigations is that of Miller et al. (14), who detected similar mercury concentrations in museum specimens of tuna and swordfish species and freshly caught fish. In the field study, it was found that, on average, 88.9% of the total mercury in fish musculature was methylmercury, irrespective of the weight, length, and species of the fish. The assumption can be made that most of the mercury in the ecosystem is available to fish as organic mercury compounds.

Trace Elements in Aquatic Biota in Sediments

The distribution of elements in the aquatic milieu shows that the highest element concentrations are generally found in bottom sediment. The amount of elements concentrated there present, however, a particular danger to organisms, especially to those that live in the sediment and enter the food chain. Kushner (15) found bacteria in the Ottawa River of which 50% and more were resistant to 1 ppm Hg^{2+} and some were even resistant to 10 ppm Hg^{2+}. These bacteria could transform Hg^{2+} into a volatile mercury compound. The accumulation of radioactive cadmium from sediments and sea water by polychaetes was observed by Ueda et al. (16). Worms in contact with sediments accumulated six times more cadmium than worms not in contact with sediments.

Zn, Cu, Pb, and Cd elemental values in goldfish were much higher than those in catfish. This discrepancy can probably be explained by the different feeding habits of the species; the goldfish burrow when searching for food, whereas the catfish is a predatory fish.

EFFECTS OF GEOLOGY

This is the source of baseline or background levels. It is to be expected that in areas characterized by element-bearing formations, these elements will also occur at elevated levels in the water and bottom sediments of the particular area. Trace elements may be derived from the weathering of rocks. Average concentrations of some trace elements in rocks are shown in Table 1.

Whether a particular trace element goes into solution during weathering depends on the mineral in which the element occurs and on the intensity of chemical weathering. Many trace elements do not substitute readily in feldspars or common ferromagnesian minerals. They may be present in chemically resistant accessory minerals such as zircon, apatite, or monazite, or as sulfides, which weather rapidly in oxygenated water. The resistant minerals generally remain unaltered unless weathering is very intense (gibbsite formation), and so alteration of them never causes high trace element concentrations in water. Where many elements occur in high concentrations and hence in many economic deposits, the elements occur as sulfides (e.g., copper, zinc, molybdenum, silver, mercury, lead), and these sulfides often contain selenium, arsenic, and cadmium. Sulfides weather rapidly, so such ore deposits can give rise to local high concentrations of dissolved trace elements.

EFFECTS OF HUMAN ACTIVITY

Human activities introduce trace elements into the hydrosphere in many ways. Burning of fossil fuels and smelting of ores put elements into the hydrosphere, where they are washed out by rain into surface water.

Table 1. Concentration of Selected Elements in Rocks (ppm)[a]

	Granite	Basalt	Shale	Sandstone	Limestone
Aluminum	Major	Major	Major	Major	4200
Arsenic	2	2	13	1	1
Beryllium	3	1	3	–	–
Boron	10	5	100	35	20
Cadmium	0.13	0.2	0.3	–	0.03
Chromium	10	170	90	35	11
Cobalt	4	48	19	0.3	0.1
Fluorine	800	400	740	270	330
Iodine	0.5	0.5	2	1	1
Lead	17	6	20	7	9
Lithium	30	17	66	15	5
Manganese	450	1500	850	50	1100
Mercury	0.03	0.001	0.4	0.03	0.04
Molybdenum	1	1.5	2.6	0.2	0.4
Nickel	10	130	68	2	20
Rare earths	0.5–70	1–80	1–80	0.05–15	0.05–8
Scandium	10	30	13	1	1
Selenium	0.05	0.05	0.6	0.05	0.9
Thallium	1.5	0.2	1.4	0.8	–
Thorium	14	2.7	12	5.5	2
Titanium	Major	Major	Major	Major	400
Uranium	3	1	4	2	2
Vanadium	50	250	130	20	20

[a] Reference 17.

A wide range of trace elements in fossil fuels are either emitted into the environment as particles during combustion or accumulate in ash which may itself be transported and contaminate water. Some of the elements arising as pollutants from fossil fuel combustion are Pb, Cd, Zn, As, Sb, Se, Ba, Cu, Mn, and V.

Municipal sewage and industrial effluent introduce trace elements directly. Mining activities can result in release of elements because previously impermeable rocks are broken up and exposed to water and because sulfide-containing rocks are exposed to oxygen, resulting in rapid alteration and dissolution. In the mining industry, there are two major contaminant waste streams: (1) elements discharged in solution via mine drainage and (2) particulate grains of ore-forming or related minerals released after ore processing (18). Underground disposal of toxic wastes, including radioactive wastes, has the potential to release a variety of substances to groundwater and hence to surface waters.

Agriculture constitutes one of the very important nonpoint sources of trace elements. The main sources are

1. impurities in fertilizers: Cd, Cr, Mo, Pb, U, V, Zn, (e.g., Cd and U in phosphatic fertilizers)
2. pesticides: Cu, As, Hg, Pb, Mn, Zn
3. wood preservatives: As, Cu
4. wastes from intensive pig and poultry production: Cu, As
5. composts and manures: Cd, Cu, Ni, Pb, Zn, As

The human input of trace elements in many rivers and lakes is many times greater than the natural input.

BIBLIOGRAPHY

1. Drever, J.I. (1982). *The Geochemistry of Natural Waters*. Prentice-Hall, Englewood Cliffs, NJ, pp. 297–317.
2. Rubio, B., Pye, K., Rae, J.E., and Rey, D. (2001). Sedimentological characteristics, heavy metal distribution and magnetic properties in subtidal sediments, Ria de Pontevedra NW Spain. *Sedimentology* **48**: 1277–1296.
3. Mitchell, R.L. (1964). Trace elements in soil. In: *Chemistry of the Soil*. F.E. Bear (Ed.). Reinhold, pp. 320–368.
4. Mehrotra, M.N. (1990). The role of sediments in environmental pollution: A case study of the Ganga at Varanasi. *J. Indian Assoc. Sedimentologists* **9**: 1–14.
5. Dahlberg, E.C. (1968). Applications of a selective simulation and sampling technique to the interpretation of stream sediment, copper anomalies near South-Mountain. *Econ. Geol.* **63**: 409–417.
6. Carpenter, K.E. (1924). A study of the fauna of rivers polluted by lead mining in the Aberystwyth district of Cardiganshire. *Ann. Appl. Biol.* **11**: 1–23.
7. Weber, H. (1934). Vergiftungsversuchemit Kupfersulfat, *Z. Hydrol.* **4**: 105–126.
8. Dawson, A.B. (1935). The hemopoietic response in the catfish, *Ameiurus nebulosus* to chronic lead poisoning. *Biol. Bull.* **23**: 298–311.
9. Jones, J.R.E. (1946) The oxygen consumption of *Gasterostens aculeatus* L. in toxic solutions. *J. Exp. Biol.* **23**: 298–311.
10. Knauer, G.A. and Martin, J.H. (1973). Seasonal variations of cadmium, copper, manganese, lead and zinc in water and phytoplankton in Monterey Bay, California. *Limnol. Oceanogr.* **18**: 597–604.
11. Forstner, U. and Wittman, G.T.W. (1979). *Metal Pollution in the Aquatic Environment*. Springer-Verlag, New York, pp. 271–318.
12. Muller, G. and Prosi, F. (1978). Verteilung Von Zink, Kupfer und Cadmium in verschiedenen Organen von Plotzen (Rutilus rutilus L.) aus Neckar und Elsenz. *Z. Naturforsch.* **33c**: 7–14.
13. Abdullah, M.I., Banks, J.W., Miles, D.L., and O'Grady, K.T. (1976). Environmental dependence of manganese and zinc in the scales of Atlantic salmon, *Salmo salar* (L.) and brown trout, *Salmo trutto* (L.). *Freshwater Biol.* **6**: 161–166.
14. Miller, G.E. et al. (1972). Mercury concentrations in museum specimens of tuna and swordfish. *Science* **175**: 1121–1122.
15. Kushner, D.J. (1974). Microbial dealing with heavy metals. *Proc. Int. Conf. Transp. Persistent Chem. Aquatic Ecosyst.*, Ottawa, Canada.
16. Ueda, T., Nakamura, R., and Suzuki, Y. (1976). Comparison of Cd accumulation from sediments and sea water by polychaete worms. *Bull. Jpn. Soc. Sci. Fish.* **42**: 299–306.
17. Martin, J.M. and Meybeck, M. (1979). Elemental mass balance of material carried by major world rivers. *Mar. Chem.* **7**: 173–206.
18. Pirrie, D. et al. (2003). The spatial distribution and source of arsenic, copper, tin and zinc within the surface sediments of the Fal Estuary, Cornwall, U.K. *Sedimentology* **50**: 579–595.

INNOVATIVE PENS HATCH THOUSANDS OF TROUT

Richard Dowling
Pittsburgh District
U.S. Army Corps of Engineers

For a fellow who doesn't fish, Pittsburgh District's Mike Fowles has sure helped to put a lot of trout into the creels of Pennsylvania anglers.

US Army Corps
of Engineers
Natural Resource
Management

A fish and wildlife specialist for the U.S. Army Corps of Engineers for the last 26 years, Fowles has worked out a landmark agreement with a local chapter of Trout Unlimited to establish a cooperative fish hatchery in the tailrace of the district's Youghiogheny Lake. The effort has put an estimated 20,000 trout ranging from 12 to 16 inches into various area waters, and earned Pittsburgh District recognition across the state as a friend of fish.

The Chestnut Ridge Chapter of Trout Unlimited (TU) this spring signed a five-year extension to its memorandum of understanding (MOU) to run the hatchery. Other signers along with the Corps were the Pennsylvania Fish and Boat Commission and D/R Hydro, operators of an electric power generating plant at the dam. The initial three-year MOU was due to expire in June, and has been hailed as a success beyond anyone's expectations.

The key to this success, according to Fowles, is the use of "cage culture" nursery pens actually submerged in the flowing water of the Youghiogheny River, rather than the more common hatchery practice of long concrete tanks with artificial flows and aeration.

The fish arrive in early summer as three-inch fingerlings, and are fed a high-quality protein meal by TU members, who are also responsible for trash pick-up and maintenance of the submerged pens. Corps and D/R Hydro monitor water quality in the tailrace to insure that the young trout have the best possible living conditions. The result a year later is that thousands of the former fingerlings are more than a foot long, and ready to be stocked. "This is a really good growth rate," said Fowles.

Some of the fish go right into the Youghiogheny River (one of Western Pennsylvania's premiere trout streams) while others are trucked to other nearby waters. The in-river hatchery offers much better growth and survival rates for the trout, which include rainbows, brown trout, brook trout, and an exotic strain of rainbows known as "palomino" trout.

Sections of the river downstream from the hatchery are designated "trophy trout" sections by the State Fish and Boat Commission, and draw anglers from throughout the eastern U.S.

The state provides the trout fingerlings each spring. For its part, the hydroelectric company provides free power from its nearby generators to support the hatchery.

"This is truly a win-win situation for all the partners," said Fowles on the day of that the partners extended the MOU. "This is an outstanding example of how Pittsburgh District is developing new partnerships to accomplish our natural resource management and environmental stewardship objectives."

Tom Shetterly, president of the Chestnut Ridge Chapter of TU, said about a half dozen of his group's members rotate the daily feeding duty for the young trout through the year. "When it's time to stock, we get 12 to 15 guys there," he said. The annual feeding bill runs about $2,000.

He said that the Corps of Engineers "is just a great partner with us; you guys have just bent over backwards to make this program better than anybody could have ever expected." He also praised the efforts of other Pittsburgh District personnel, including Mike Koryak, district limnologist (one who studies fresh waters), and Pat Docherty, area operations manager.

Working with the Corps and the state, TU has also received permission to place gravel for enhanced natural spawning beds in the river and the stilling basin below the Yough Dam.

"We're seeing some of the first fish we stocked come back to the same area to spawn," Shetterly said. Raising the fish in submerged cages apparently has imprinted them with the characteristics of their immediate surroundings, something that Shetterly says has never been accomplished elsewhere. "This thing is huge," he marveled. "It's going to be world-famous."

Another bonus of the in-river cage culture is the way the trout benefit from natural food, adding up to 12 ounces in weight gain for each pound of artificial fish food added to the cages. Waste from the trout and any food that goes through the bottom of the cages also combine to increase the alkalinity of the water, thus improving the environment, according to Shetterly.

The trout hatchery is one of many programs supporting fishing in Pittsburgh District. For example, working with the Yough Walleye Association, the district is supporting efforts to put crushed limestone into Tub Run, a tributary of Youghiogheny Lake, to neutralize the highly-acidic waters, a legacy of abandoned coal mining.

Youghiogheny Lake is one of 16 multi-purpose flood control reservoirs built and operated by Pittsburgh District in Western Pennsylvania, New York, Ohio, West Virginia, and northern Maryland. Pittsburgh District also works with various state fish and wildlife agencies to conduct annual fish census sampling at its navigation locks on the Ohio River.

All in all, Fowles has worked hard to help improve fish habitat and fishing as a sport, but don't ask him to tell you his own fishing stories. "I don't fish," he said simply. "And I don't hunt either, though of course I recognize both as legitimate outdoor activities."

He is passionate about bird watching, a hobby he and his wife picked up in 1973 while an undergraduate at Ohio University. His logbook shows he has sighted thousands of bird species all over the world.

Thousands of brook, rainbow and brown trout are raised each year in these submerged pens in the tailwaters if the Youhiogheny River. (Photo courtesy of Pittsburgh District)

"My list is rapidly approaching the 5,000 mark, a goal I set for myself when I first started birding." Comparing the two sports, he says of birders, "Oh, we get our prey. We just don't kill 'em."

WATERSHED

Karen D. Brettschneider
Houston-Galveston Area
Council
Houston, Texas

DEFINITION

Everyone lives in a watershed. A watershed is the hydrologic or geographic boundary of a river system. Watersheds can be delineated for any waterbody, whether it is the creek flowing through your neighborhood or the Mississippi River (Fig. 1).

Watersheds are defined by first selecting a point along a stream or river. The stream and its tributaries, if any, are then traced backward from there, and the topography of the terrain is analyzed. The watershed boundary is defined by the high point, or ridge, between streams flowing to separate major waterbodies or "stems." All waters falling within the watershed boundary, in theory, eventually flow to the original point designated to define the watershed.

WATER QUALITY

All factors and activities within a watershed directly or indirectly affect the water quality of river systems. Although local climate, geology, soil type, and vegetation play an important role in determining of the water quality conditions within a watershed, one of the most important factors affecting water quality is land use. Runoff from different land use types, such as agricultural, industrial, urban, or even barren or deforested land may contribute a number of pollutants to a waterbody. Some of these pollutants include excessively applied fertilizers or pesticides, animal waste, toxic chemicals, sewage from broken sewer lines or septic tanks, oil, and sediment.

Best management practices on all types of land use help reduce the amount of pollution entering waterways. Preserving, restoring, or creating wetlands, which "filter out" many pollutants is another way to help ensure clean water and a clean environment within a watershed.

FLOOD POTENTIAL

Flooding is also influenced by factors such as the percentage of impervious cover within the watershed boundary. In addition, concrete drainage control ditches and canals have been built over the years to help control flooding. When a channel is lined with concrete, natural aquatic habitats are disturbed or destroyed, and pollutants

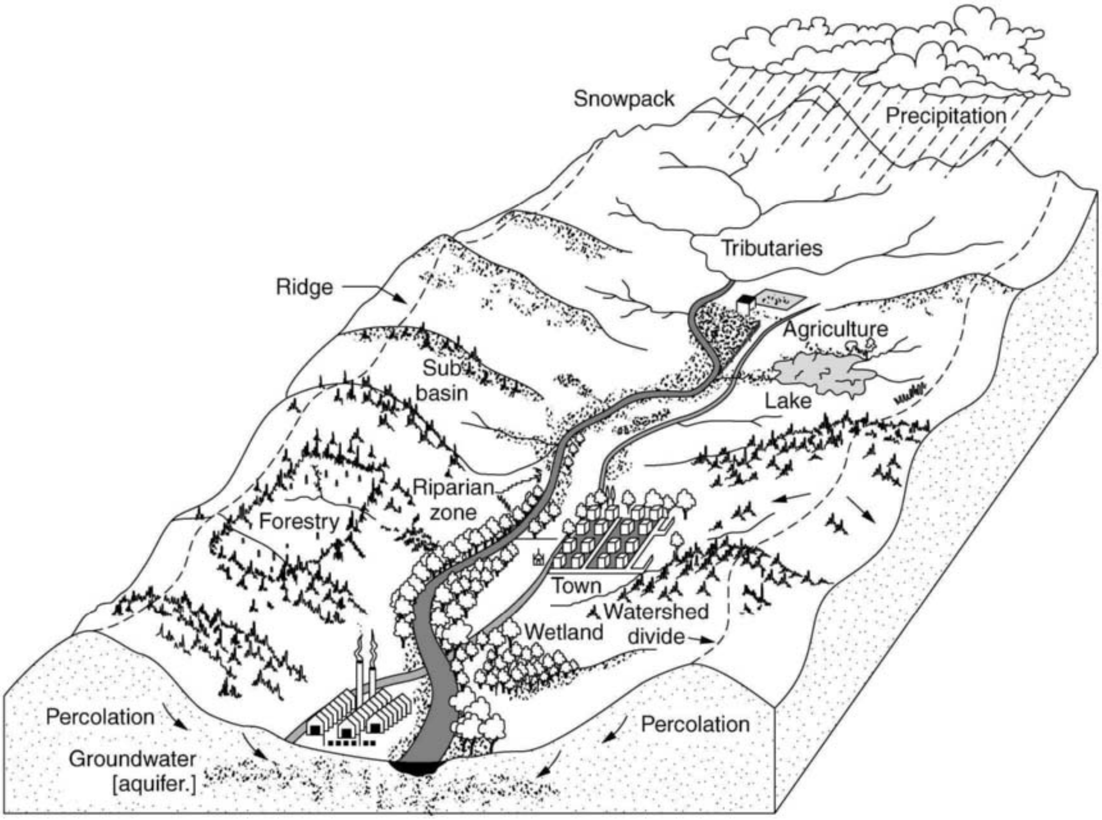

Produced by lane council of governments

Figure 1. Sample drawing of a watershed (Used with permission of the Lane Council of Governments, Eugene, Oregon).

can more easily enter a waterway due to the lack of vegetation to control and filter runoff.

Recently, engineers have used more “environmentally friendly” types of flood control. Some these include constructed wetlands and detention ponds. In these cases, habitat is not destroyed, but created, and pollutants are kept from entering waterways.

Although a person or entity may not be located near a waterbody, the results of various actions eventually flow downhill or through stormwater infrastructure to creeks and streams.

COMBUSTIBLE WATERSHEDS

V. Kulik
CYPUM PTY LTD
Canberra City, Australia

Existing hydrological networks can be used for conflagration prediction. Runoff data allows predicting high fire danger earlier, highlights the most dangerous watersheds, and makes forecasting more foolproof. In some cases, a hydroforecast improves the likelihood of preventing a major disaster. This technique has certain differences from standard meteorological methods: (1) It predicts conflagrations not for an administrative or geographical region, but for a watershed—so, a conflagration can be forecast for the area around water supply reservoirs. (2) It allows forecasting the fire danger earlier and estimates of the amount of time available for active response. (3) It excludes most false alarms. (4) It is more accurate, simply because it uses hydrographic data, which are more representative than meteorological data. (5) It is more foolproof and less dependent on errors in observations. (6) It depends less on the stochastic nature of summer rains and other meteorological elements. (7) It is simpler than any other method.

THE SIMPLEST TECHNIQUE

1969

Watershed = 671 sq.km, 30 km SSW from Canberra City (capital of Australia), adjacent to the water supply watersheds of Cotter River (mainly eucalyptus forest overlying granitic soils).

1970

45° = baseflow recession during cool periods (semilog scale for cumecs).

1971

Vertical scales have to be the same for all years.

1972

Every summer there are many small fires. A fire > 1 sq. km starts, when the discharge < 0.1 cumecs and the recession > 45°.

1973

For this and 4 previous years: cumulative evaporation = two cumulative rainfalls.

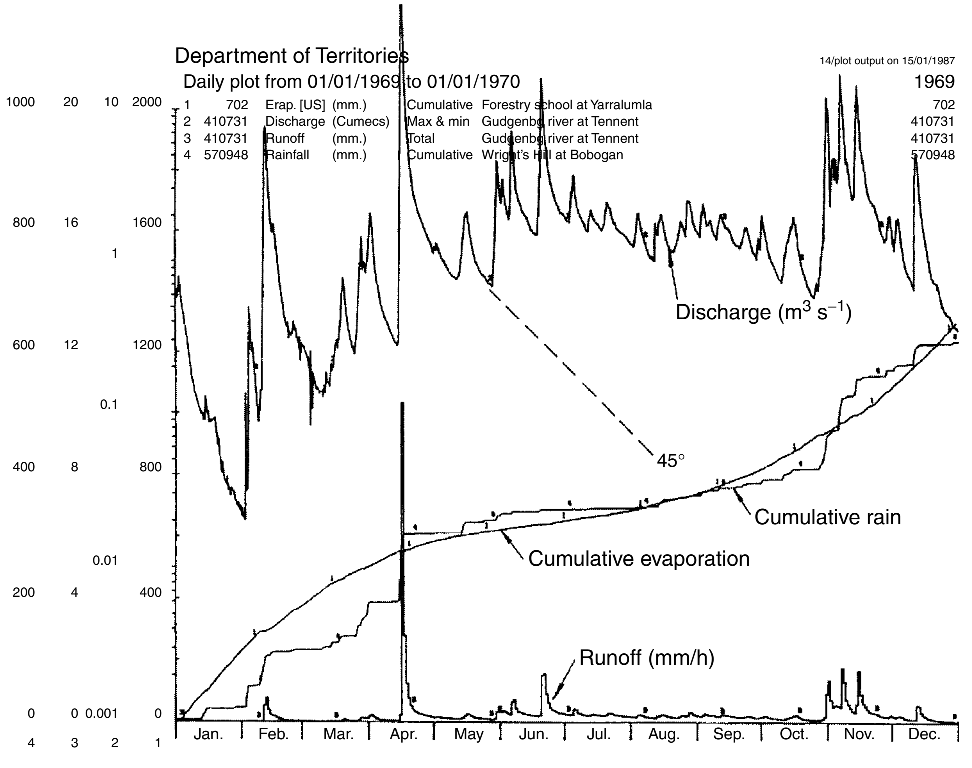

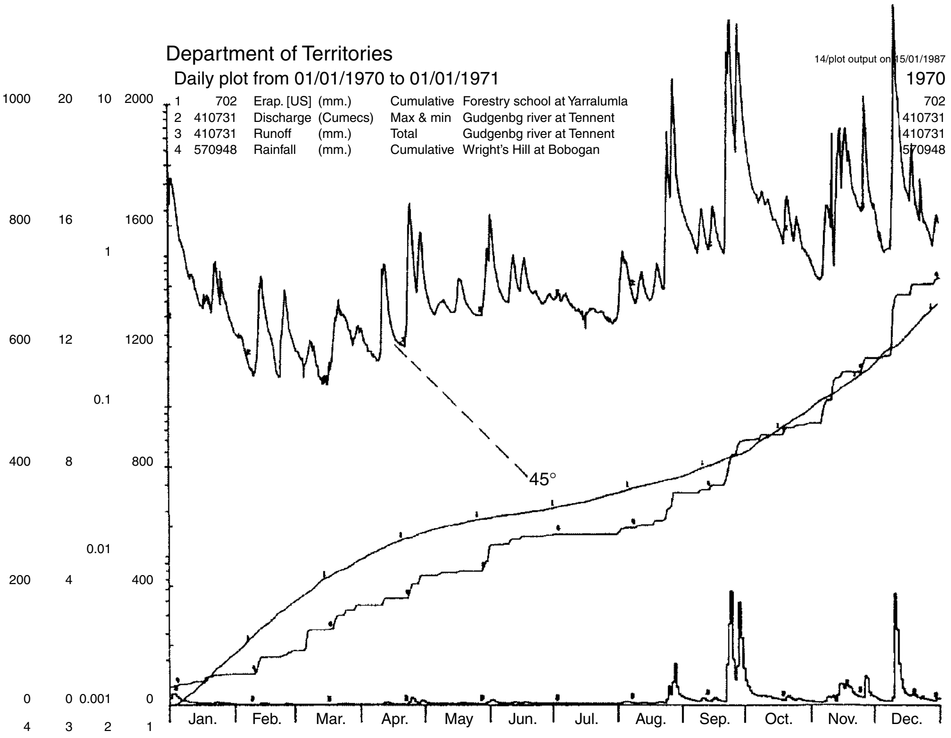

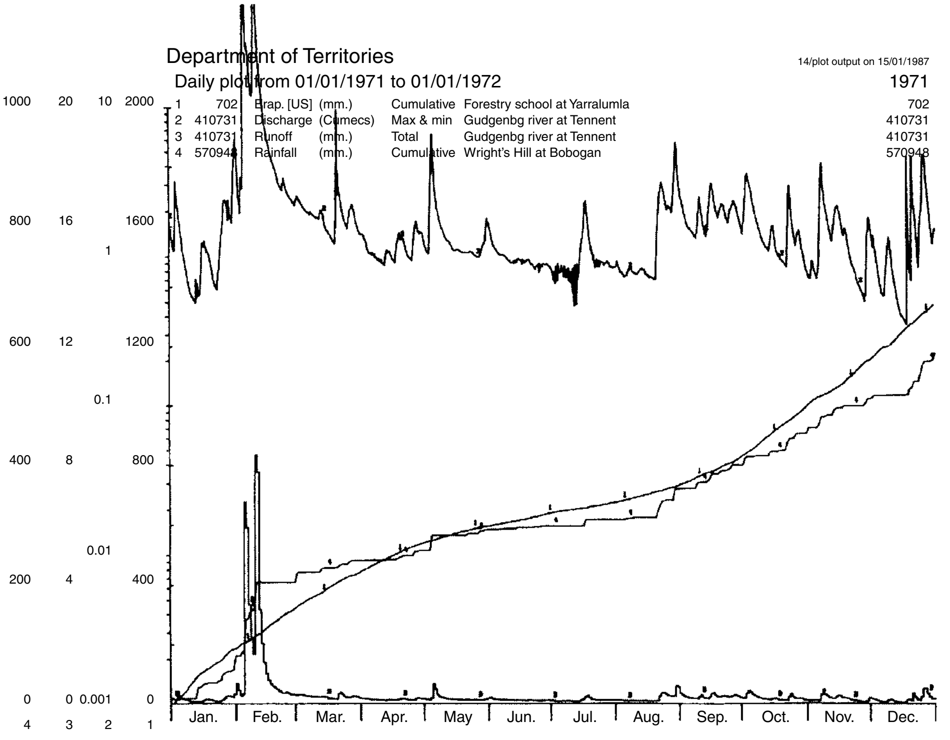

1974

Floods: Cumulative evaporation < two cumulative rainfalls. Meteo database problem: "negative rain" for October 1974. Dryness of fuel affects the intensity of a fire. Several special indexes are used to forecast it. There are no approved international standards for calculating these indexes, and even in one country, several different drought and fire danger indexes may be used (1–2).

The idea behind any antecedent moisture content index is to "sum" the rainfall and evaporation across a certain time interval. Hence, a mistake in a daily rain record could be fatal, not only for this daily index, but also for subsequent days.

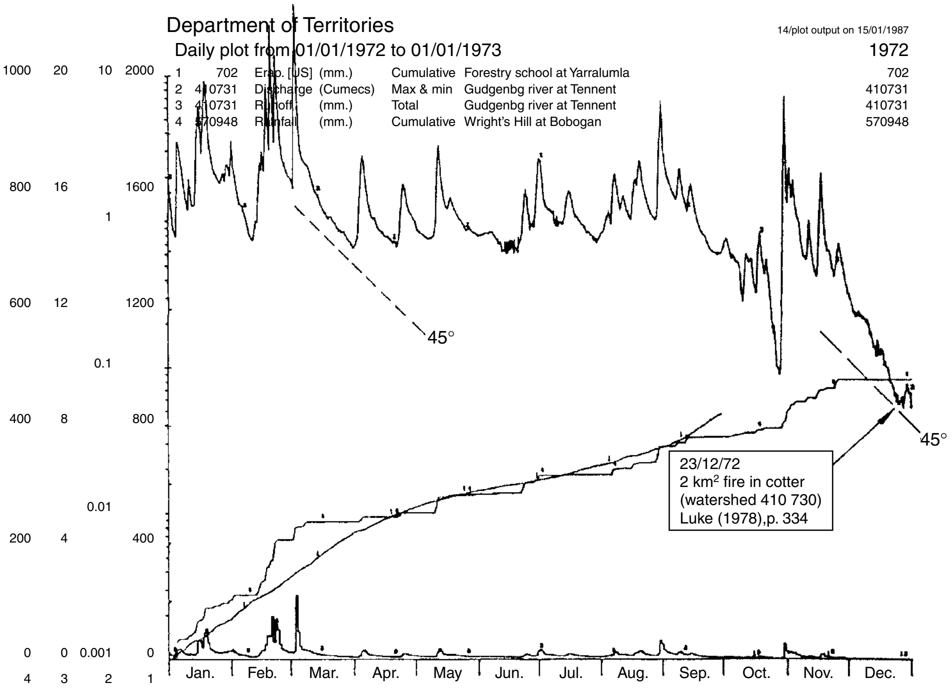

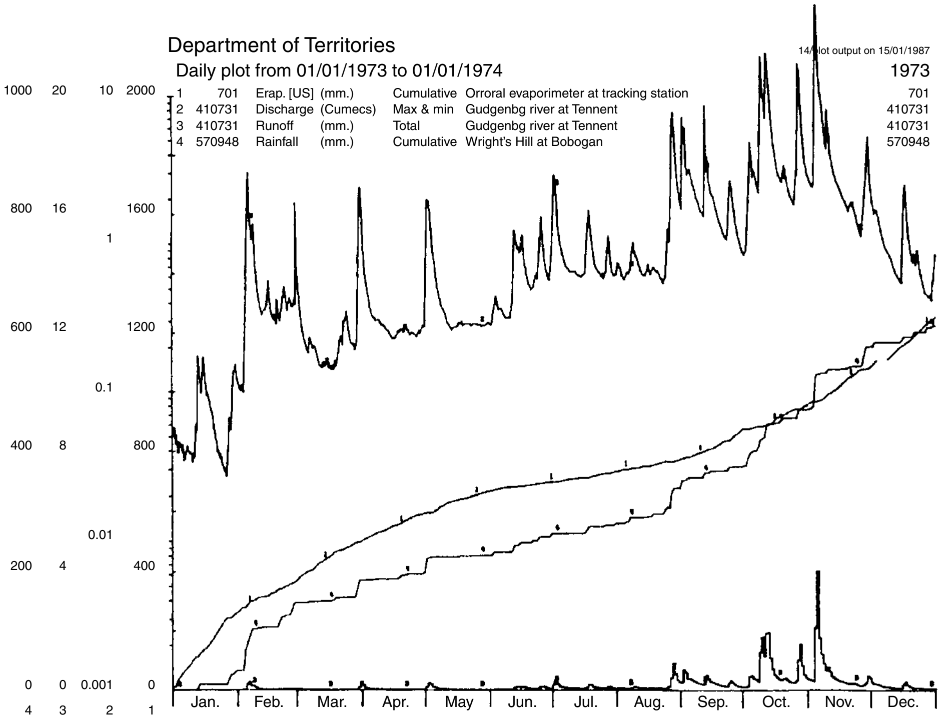

1975

Small fires only. Defect of meteo data: evaporation. Turner (2, p. 12) stated that "a single day of careless weather observations could strongly affect fire weather indices."

1976

Similar problem with meteo data (as in 1972, December 1973, 1975).

1977

Low flow does not indicate a fire, if the recession = 45° (compare with 1972).

1978

February: discharge = 0.1 cumecs, but the recession is concave. Small fires only.

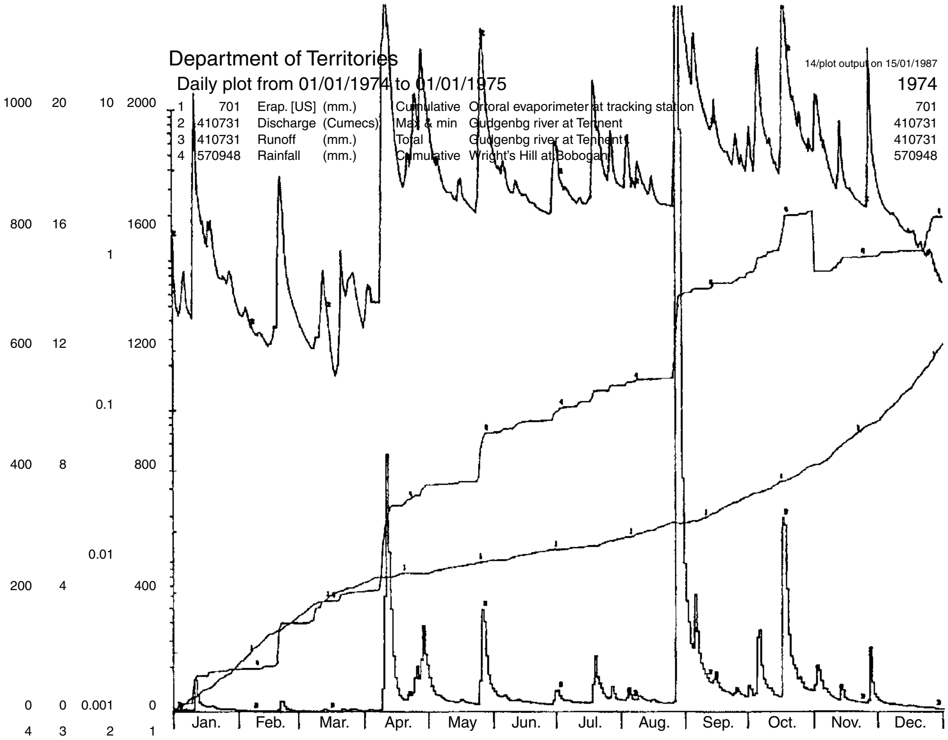

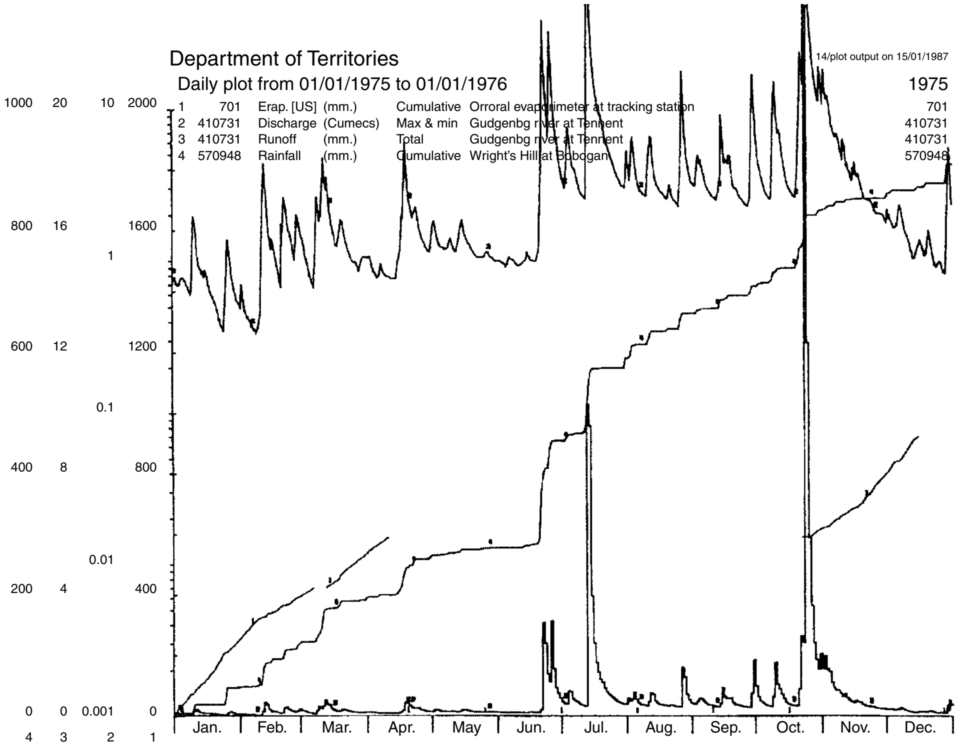

1979

Small fires only. Cumulative evaporation > two cumulative rainfalls, meteo fire danger is "extreme." It does not mean, however, that the meteo indices and other techniques of bushfire forecasting are to be abolished and forgotten. The hydrologic technique only supplements them.

1980

Cumulative evaporation > two cumulative rainfalls, but the recession is concave. Small fires only. Bushfires can start almost any time. Even during autumn rains, lightning could start a bushfire. A wet forest will not burn as well as a dry one, but it may burn. However,

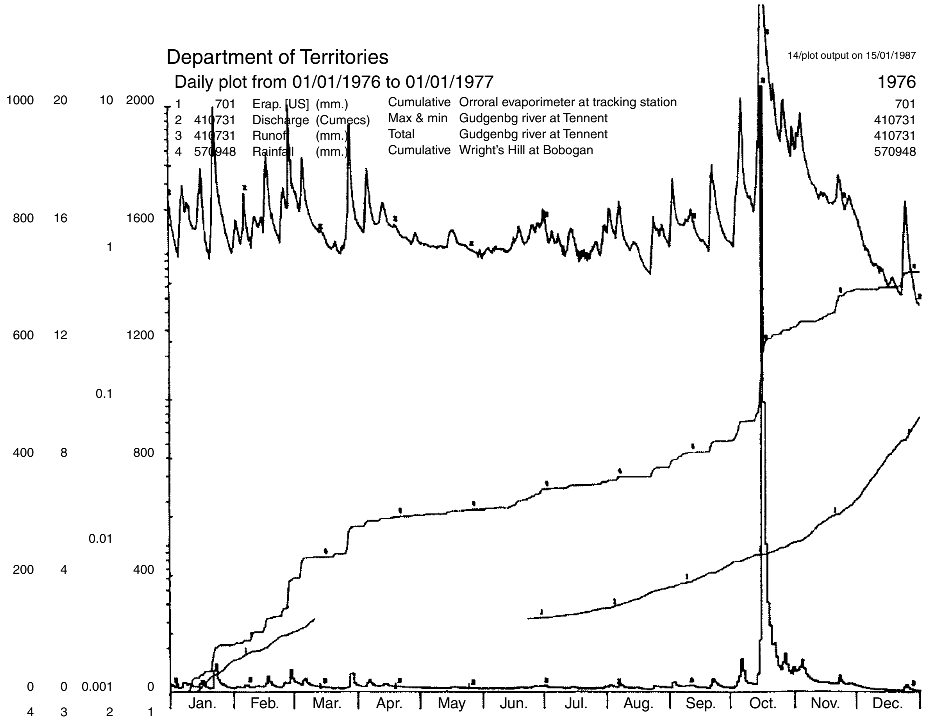

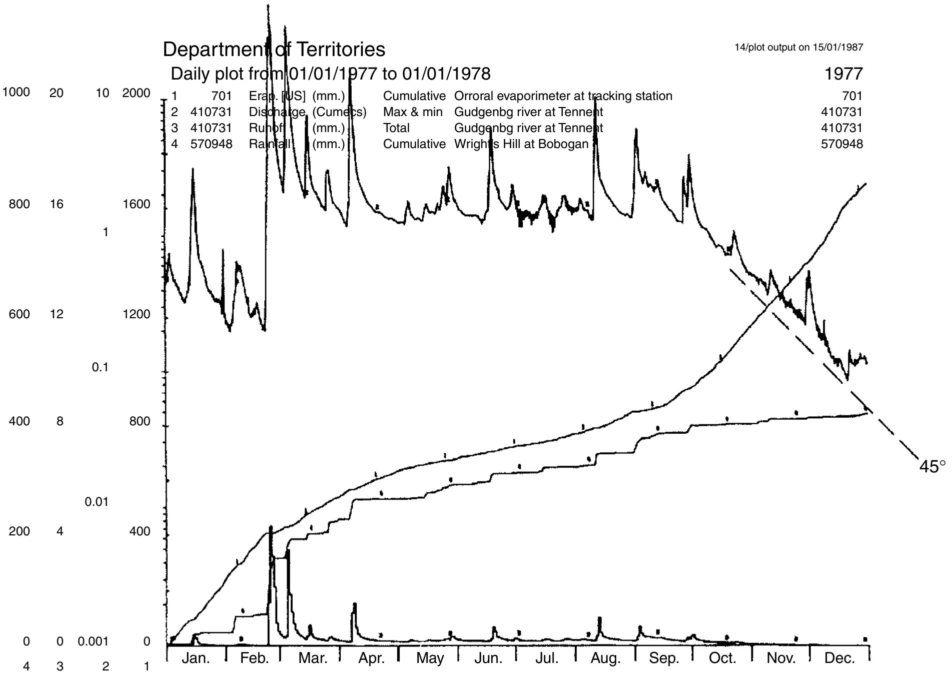

the intensity of such bushfires is usually very low, and it is generally thought that such bushfires are not a major hazard. They decrease the amount of the accumulated fuel and hence to some extent, prevent the more dangerous, intense conflagrations.

1981

Cumulative evaporation > two cumulative rainfalls, but the recession is concave. Small fires only. Small fires are used for prescribed burnings and are considered generally a useful tool when controlled effectively and sufficiently, little damage results in the long-term components of the environment. These artificial bushfires are called Fuel Reduction Burnings (3–5). The 1981 hydrograph shows that it was a good year for prescribed burnings.

1982

There are two groups of predictors: (a) Early predictors: (a1) low flow; (a2) change of the curvature of the recession curve from concave to convex (semilog paper); (a3)

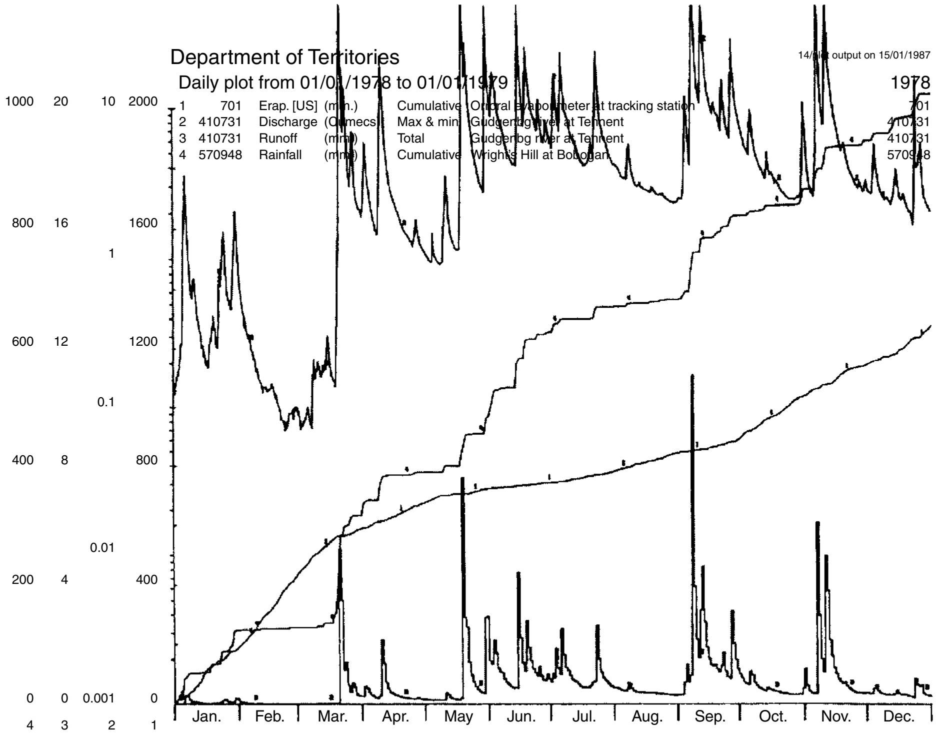

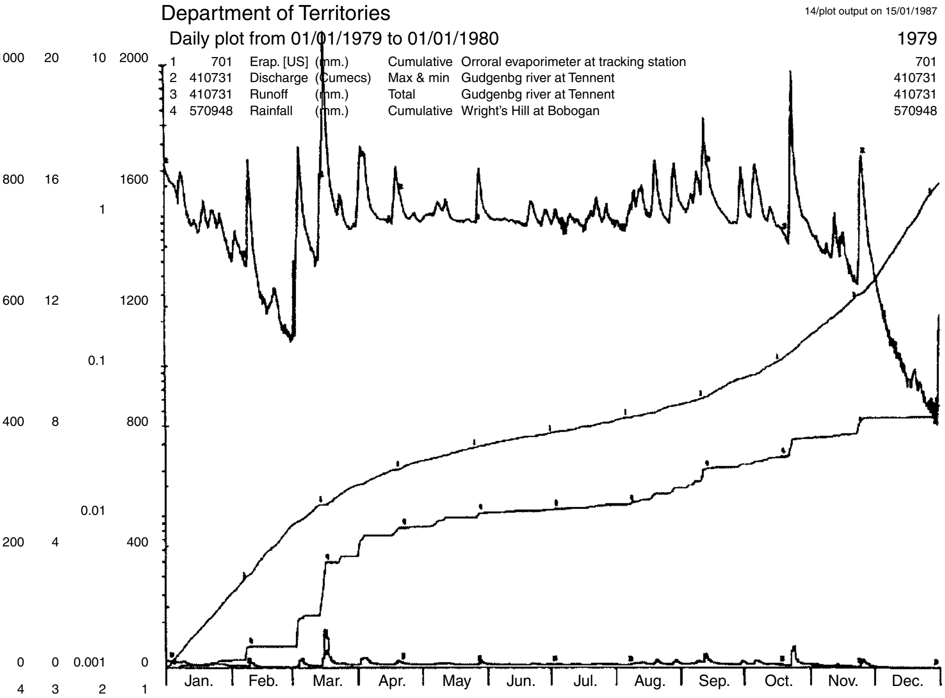

accelerated recession. These early predictors can be used to estimate how many days there are before the critical fire danger starts (by extrapolating the recession curve). (b) Later predictors: (b1) stabilization of the runoff to a "zero flow" value; (b2) Steep rate of recession after intermediate minor rains; (b3) Convex recession after intermediate rains. (b4) Abrupt stabilization of runoff to "zero flow" several days after a minor rain. The manifestation of these late predictors means that the probability of a conflagration has increased dramatically.

1983. Combustion January 9, 1983

In January 1983, a wild fire burnt out approximately 58 km^2, representing 30% of the Corin Dam catchment on the Cotter River and 300 km^2 of the adjacent Gudgenby River catchment. It was a conflagration of

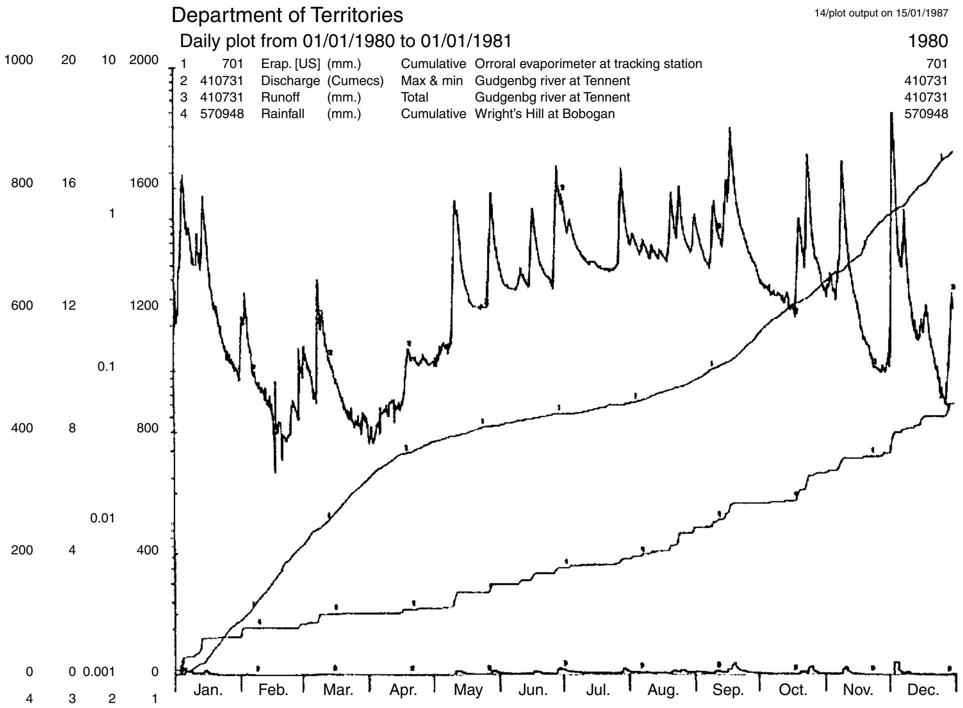

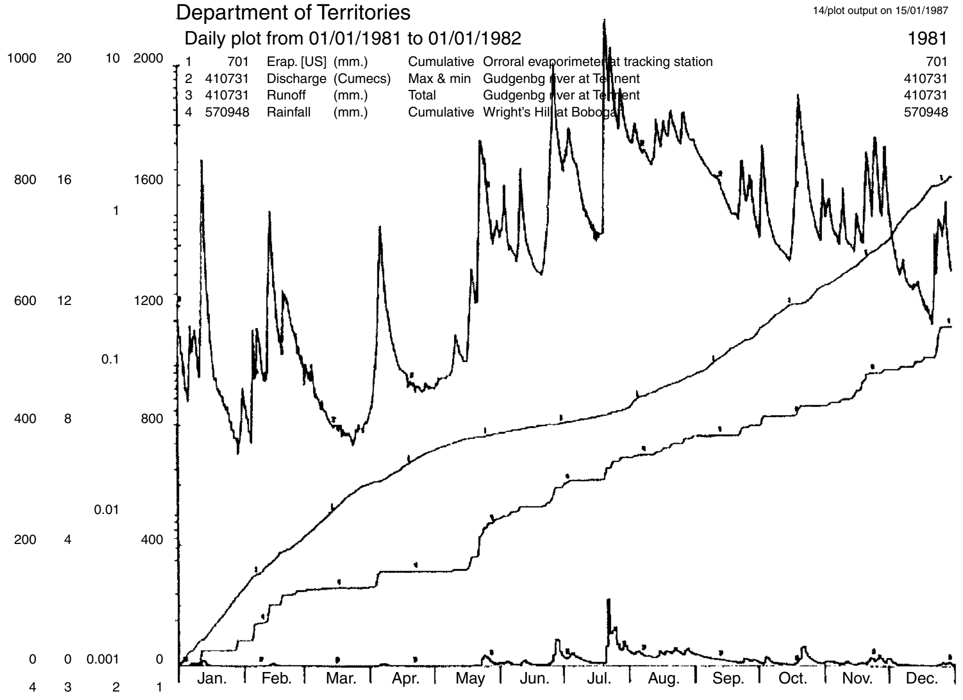

very high intensity. The conflagration completely burned the Boboyan pine tree plantation (300 ha), despite all attempts to save it (by concentrating firefighting resources on protecting the plantation). Later on, attempts to control the conflagration were only partly successful, as the conflagration continued to spread for about 3 weeks. (Note recession after intermediate rains). The approximate cost of conflagration suppression operations was $477,000 (6).

SUMMARY

A major forest conflagration is a very rare phenomenon; to collect a sufficient amount of information for development of the forecasting technique, it is necessary simply to analyze more cases. We ask everybody who has at hand data on watersheds, affected by a conflagration, to try this simple technique, to develop it, and to communicate their results.

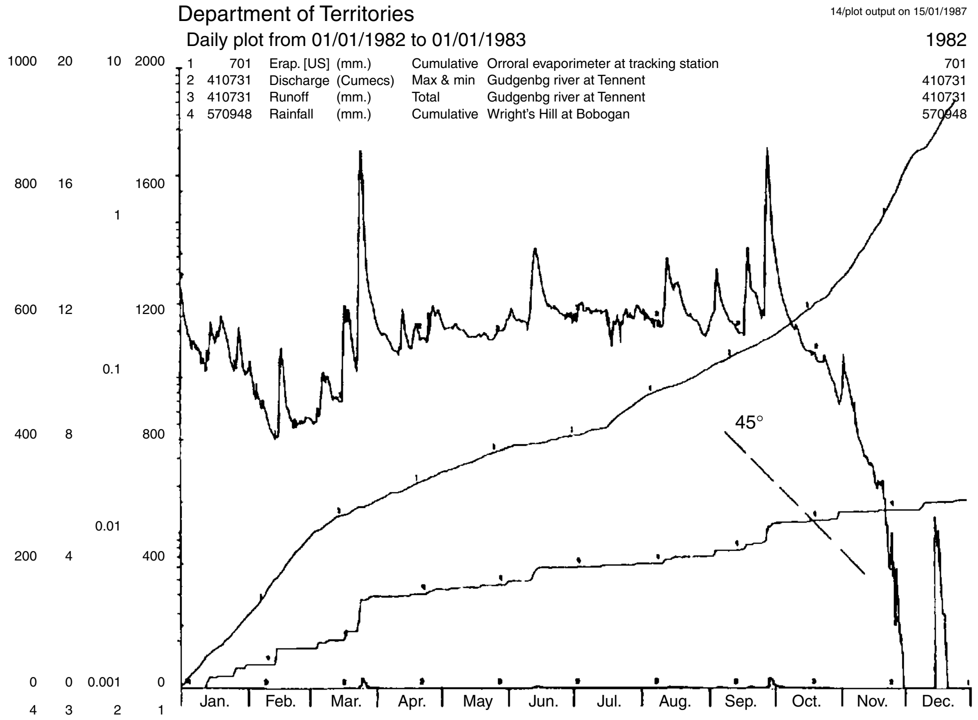

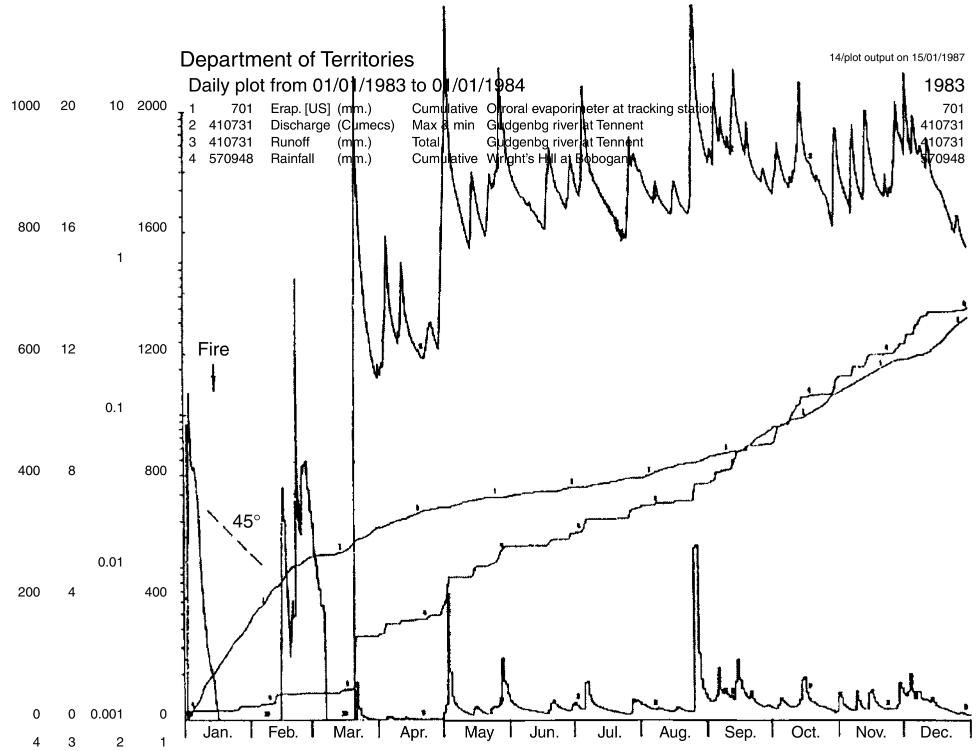

BIBLIOGRAPHY

1. Van, W. (1975). *A Comparison of The Canadian and American Forest Fire Danger Rating System.* Information Report PS-X-59, Patawana Forest Experimental Station, Chalk River, Ontario, Canada, p. 19. Available: http://www.nps.gov/seki/fire/links.htm.
2. Turner, J.A. (1978). *Weather in Canada Forest Fire Danger System: A User Guide to National Standards and Practice.* Canada p. 39. Available: http://www.nofc.forestry.ca/fire/.
3. Luke, R.H. and McArthur, A.G. (1978). *Bushfires in Australia.* Australian Government Publishing Service, Canberra. Available: http://life.csu.edu.au/hazards/3WildFires.html.

4. Cheney, N.P. (1976). Bushfire disasters in Australia 1945–1975. *Australian Forestry* **39**: N4, 245–268. Available: http://www.bbm.csiro.au/vesta/.
5. O'Loughin, E.M., Cheney, N.P., and Burns, J. (1982). The Bushranger Experiment: hydrological response of eucalypt catchment to fire. *First Nat. Symp. for Hydrol.*, Melbourne, pp. 132–138. Available: http://www.nps.gov/gis/metadata/sagu/sagu_rxburns.html http://www.prescribed-fire.org/state.html.
6. *The Gudgenby Fire—9 January 1983.* (1983). Operative report of ACT fire brigade unit. Canberra 1983, p. 37. Available: http://www.esb.act.gov.au/esb.htm.

TIME OF CONCENTRATION AND TRAVEL TIME IN WATERSHEDS

Bahram Saghafian
Soil Conservation and
Watershed Management
Research Institute
Tehran, Iran

INTRODUCTION

Time of concentration, denoted by T_c throughout this article, is one of the oldest and most fundamental concepts in hydrologic science and engineering. Time of concentration and travel time (T_t) are two measures of watershed runoff response time. Travel time represents a broader hydrologic term, and time of concentration may be considered a special case of travel time. Several hydrologic analysis and design procedures require direct knowledge of T_c or T_t. In the rational method, T_c must be known to determine the duration of design storm and thus, the flood peak. The Muskingum stream routing method requires wave travel time through stream reaches as a measure of the storage time constant. The application of time–area or Clark rainfall–runoff transformation is optimum when the spatial distribution of travel time is known.

DEFINITIONS

Time of concentration is usually defined as one of the following categories:

1. The time that it takes for water to travel from the most distant part of a watershed to the outlet. Some literature appropriately emphasizes that the distance has a hydraulic base and not necessarily a geometric base.
2. The wave travel time from the most hydraulically remote point to the outlet of a watershed.
3. The time elapsed for which all parts of a watershed contribute to direct runoff at the outlet.
4. In hydrograph analysis, T_c is defined as the time between the end of excess rainfall and the inflection point on the falling limb of the hydrograph. It is assumed that the inflection point indicates the end of direct runoff. The centroid of the rainfall hyetograph is sometimes defined as the beginning of T_c.

Definitions 2, 3, and 4 represent a similar interpretation of watershed time response, but definition 1 is based on water velocity which is different from wave celerity. As we will see in the next section, some methods explicitly rely on definition 1 for estimating T_c. Definitions 2 and 3 are similar but depend on the duration of rainfall excess (T_r). If, in theory, excess rainfall intensity is constant and T_r is sufficiently long or at least equal to the time to equilibrium (T_e) at which the watershed reaches a steady state, then $T_c = T_e$. On the other hand, when $T_r < T_e$, definitions 2 and 3 are still valid, but the time in neither is equal to T_e. The case of $T_r < T_e$ represents a dominant case in the hydrology of watersheds, so time of concentration becomes event-dependent, not just intensity-dependent, and this is not convenient in practice. Even almost all hydraulic T_c equations avoid dealing with the issue of variable rainfall duration.

Travel time follows definition 2 and may be determined from any point in the watershed to another point of interest downstream. Thus, time of concentration is specifically a travel time whose starting and ending points are the hydraulically farthest and the outlet of the watershed, respectively. By restricting travel time to channel systems, Linsley et al. (1) describes T_t as the time between a crest at one station and the crest at a downstream station. Again, some definitions of travel time refer to definition 1, in which the water travel time is considered rather than wave travel time. T_t is also a subject in underground flow, but in this article only surface water is discussed.

THEORETICAL CONSIDERATIONS

Uncertainty in computing T_c and T_t representing watershed time responses is due mainly to simplifications of complex rainfall–runoff processes. Watershed time response depends strongly on watershed static characteristics, such as shape, area, slope, length, roughness (neglecting long-term changes), and drainage pattern. On the other hand, time response is a function of nonstationary factors such as spatial and temporal rainfall variations. Therefore, watershed time response is generally event-dependent and not a stationary character of a watershed.

Following definition 2, we first assume:

1. Excess rainfall intensity remains constant through time.
2. Rainfall duration is sufficiently long to reach a state of equilibrium.

The most comprehensive analytical derivation of travel time in watersheds with spatially variable characteristics is presented by Saghafian and Julien (2). The general relation for the travel time from location X_1 to location X_2 downstream is as follows:

$$T_t = \int_{X_1}^{X_2} \frac{dx}{c} \qquad (1)$$

where x is the distance measured along the hydraulically longest flow path from X_1 to X_2 and c is the wave celerity.

Skipping details of substitutions and using the Manning resistance equation, one can finally obtain

$$T_t = \int_{X_1}^{X_2} (1-\gamma)\left(\frac{a_1}{Q_e}\right)^{\gamma}\left(\frac{n}{a_2^{2/3}S_0^{1/2}}\right)^{1-\gamma} dx \qquad (2)$$

where Q_e is the equilibrium discharge, a_1, a_2, and γ are flow cross-sectional parameters, S_0 is the bed slope corresponding to the *kinematic travel time*, and n is the Manning roughness coefficient.

The equilibrium discharge Q_e passing through a given cross section at location x is equal to the excess rainfall rate spatially integrated over the drainage area of that cross section. This is mathematically expressed by

$$Q_e(x) = \int_0^{A(x)} i dA \qquad (3)$$

where $A(x)$ is the drainage area or flow accumulation at distance x and i is the excess rainfall intensity at any point draining to x. The substitution of Q_e in Equation 2 requires knowledge of the spatial variability of excess rainfall and of the drainage area along the hydraulically longest flow path.

Total travel time (T_t) can be separated into the travel time for overland flow, T_{to}, and for channel flow, T_{tc}, such that $T_t = T_{to} + T_{tc}$. Travel time may be computed for all locations in the watershed using Eq. 2. Note that Eqs. 2 and 3 hold for spatially variable excess rainfall. For $X_1 = 0$ at the most hydraulically distant point and $X_2 = L$ at the outlet of the watershed of interest, $T_t = T_c$. Note that the hydraulically longest path may be unknown at first, so one must determine T_t by trial and error.

Certain relations must be obtained for flow cross-sectional area $A_x(h,x)$, hydraulic radius $R(h,x)$, bed slope $S_0(x)$, and equilibrium discharge $Q_e(x)$, or alternatively i and $A(x)$, so that the travel time integral for channel flow can be calculated. Both A_x and R may be expressed as functions of flow depth h, either precisely through geometric relationships or approximately through regression curves. In a general form, $A_x = a_1 h^{b_1}$, and $R = a_2 h^{b_2}$, where a_1, a_2, b_1, and b_2 are constants for a given cross section. Also in Eq. 2, $\gamma = 2b_2/(2b_2 + 3b_1)$ when using Manning's equation. For overland flow and flow in wide channels, a_2, b_1, and b_2 equal one and $\gamma = 0.4$. Details for solving Eq. 2 in a discretized raster model are discussed in Saghafian and Julien (2).

If the excess rainfall intensity i is uniformly distributed in space, then $Q_e = iA$, and the term $1/i$ can be taken out of the integral. Then, the travel time is proportional to $i^{-\gamma}$:

$$T_t = Ci^{-\gamma} \qquad (4)$$

where C is a watershed hydrogeomorphologic index conglomerating spatially distributed characteristics such as surface roughness, slope, flow length, flow accumulation, drainage pattern, and channel geometry. For the special case of watersheds with wide channels, γ equals 0.4.

Consider a simple case of an overland plane of length L and slope S_0. Equation 2 reduces to the following equation for travel time from the upstream of the plane ($X_1 = 0$) to its lower edge ($X_2 = L$):

$$T_t = \frac{n^{0.6}L^{0.6}}{i^{0.4}S_0^{0.3}} \qquad (5)$$

The travel time obtained from Eq. 5 is essentially the time of concentration of the overland plane. For travel time in channels, Eq. 2 can be applied similarly with knowledge of the required parameters.

In practice, however, the application of these derivations is time-consuming and requires data that may not be readily available. In the following section, we describe the most commonly used T_c methods. Comparison of the such methods with the theoretical derivation can help to evaluate practice results versus theory.

METHODS AND FORMULAS

Numerous references have discussed different methods and formulas to compute T_c. However, some hydrology literature is not clear in the definition of appropriate terms in formulas, applicability, limitations, condition of original data used in developing the formulas, and even the units. In some methods, variables such as length, slope, and roughness are used with slightly different meanings. It is not possible to review all methods here, hence the discussion is limited to a few more widely practiced methods described in well-known hydrology references.

There are two general approaches for estimating T_c: hydraulic-based and empirical. The hydraulic laws of free surface flow are applied in the former approach in which the travel times for sequential reaches are added up. In the latter approach, however, empirical formulas are derived from field data, and the watershed is treated mostly as a lumped unit. One example of hydraulic-based methods is the method outlined in the previous section. This section proceeds with description of other hydraulic methods.

Hydraulic Methods

The kinematic wave model is the preferred method for the overland flow portion:

$$T_c = 0.94 n^{0.6} L^{0.6} i^{-0.4} S^{-0.3} \qquad (6)$$

where T_c is in minutes, n is the Manning roughness coefficient, L is the flow length (in feet), i is the rainfall intensity (in inches per hour), and S is the slope (ft/ft). The coefficient of Eq. 6 becomes 0.88 in SI units with T_c in seconds, L in meters, and i in meters per second. In theory, as in Eq. 5, the coefficient must be one. The resulting difference may be partly attributed to flow convergence. Typical values of overland Manning's n are given by Engman (3), Ponce (4), and McCuen (5). It is recommended that L not exceed 300 feet. Iteration is required in applying Equation 6 for the rational method because both T_c and i are unknown but related through intensity-duration frequency (IDF) curves.

The Soil Conservation Service (6) proposed the average-velocity method for T_c, which can also be applied more broadly to compute T_t, as follows:

$$T_c = \sum_{j=1}^{N} \frac{L_j}{60V_j} \tag{7}$$

where T_c is in minutes, L_j is the length of the jth flow segment (in feet), V_j is the velocity through the jth segment (in feet per second), and N is the number of segments. This equation and a similar one proposed by Chow et al. (7), clearly follow definition 1 that is based on flow velocity rather than wave velocity. The hydraulically longest flow path can be divided into N segments according to different flow types: sheet, concentrated, or channel. The segments should be approximately uniform in slope, roughness, and cross section (for channels). Charts of overland flow velocity are provided as a function of slope and cover by the SCS (6). The charts are part of a method known as the "upland method." A curve that represents concentrated flow velocity in small upland gullies is also provided (6). The flow velocities generally vary from 0.2 ft/s for mildly sloped forest lands to 15 ft/s for steep paved areas.

Equation 7 may be used to estimate the travel time through a channel. Manning's formula corresponding to a 2-year or bankful discharge may be used to compute the average velocity. For greater design return periods and/or higher than bankful discharges, however, average flow velocity must be computed considering the floodplain condition. In using Eq. 7, note that wave celerity may be 1.4 to 2.0 times the water velocity (1). So there are relatively large errors in computing channel travel time based on average water velocity. Therefore, using the wave celerity in computing travel time is recommended. Theoretically, the kinematic wave celerity is equal to βV in wide channels, where $\beta = 1.67$ for Manning's equation and $\beta = 1.5$ for Chezy's formula (4). If the Muskingum storage time constant (K) is already determined for a reach, this value can be used as a reasonable measure of wave travel time through that reach.

Linsley et al. (1) suggested station hydrograph analysis to obtain travel time in a stream. If the local inflow between the stations is negligible, travel time can be determined from the lag time between the hydrograph peaks measured at the two stations. Travel time is not a constant characteristic for a reach, so a range of stage-discharges is necessary to study the change in travel time as a function of stage. Travel time is expected to decrease as stage increases to bankful, and then it increases at stages higher than bankful.

Empirical Methods

Most empirical methods developed to estimate time of concentration assume the watershed as a simple lumped unit. Kirpich (8) presented one of the most commonly used formulas for estimating T_c based on data from some small watersheds in the United States. For Tennessee watersheds ranging in size from 1 to 112 acres, slopes from 3% to 10%, and well-defined channels, the following formula was proposed by Kirpich (8):

$$T_c = 0.0078L^{0.77}S^{-0.385} \tag{8}$$

where T_c is in minutes, L is the length from the farthest point along the channels to the outlet (in feet), and S is the slope along L or approximately the average watershed slope (in feet per feet). T_c is multiplied by 0.4 for concrete and asphalt overland surfaces, and by 0.2 for concrete channels. With units of L in km and S in m/m, the coefficient in Eq. 8 becomes 3.97.

The Bransby–Williams formula (9), which is used for watersheds up to 50 sq. miles in size, is written as

$$T_c = LA^{0.4}D^{-1}S^{-0.2} \tag{9}$$

where T_c is in hours, L is the longest distance from the watershed boundary to the outlet (in miles), A is the watershed area (in sq. miles), D (in miles) is the diameter of a circle with area A, and S is the average elevation drop along L (in feet per 100 feet). With units of T_c in minutes, L in km, A in km^2, and S in m/m, and by converting D to A:

$$T_c = 14.6LA^{-0.1}S^{-0.2} \tag{10}$$

Equation 10 is one of the most widely used formulas.

A variant of the Kirpich formula was developed by the USBR (10) for small mountainous watersheds in California:

$$T_c = 60(11.9L^3/H)^{0.385} \tag{11}$$

where T_c is in minutes, L is the length of the longest stream (in miles), and H is the elevation difference between the watershed divide and the outlet (in feet).

Johnston and Cross (11) presented a relationship for T_c derived from the data of several watersheds ranging in size from 25 to 1624 sq. miles:

$$T_c = 4.7r^{-2}(L/S)^{0.5} \tag{12}$$

where T_c is in hours, r is the stream branching factor, L is the length of the main stream (in miles), and S is the average stream slope (in feet per mile). One may alternatively omit r and use 5.0 as the coefficient of the equation.

The SCS lag equation, developed from data on agricultural watersheds, can be modified by assuming that T_c is equal to 1.67 times the watershed lag:

$$T_c = \frac{1.67L^{0.8}[(1000/\mathrm{CN}) - 9]^{0.7}}{1900S^{0.5}} \tag{13}$$

where T_c is in minutes, L is the length of the longest flow path (in feet), CN is the SCS curve number, and S is the average watershed slope (in %). This equation has been found acceptable for urban watersheds smaller than 2000 acres in size with completely paved areas.

GENERAL RECOMMENDATIONS

Selection of the equation for estimating T_c and T_t is subjective. Some general recommendations, however, can be proposed. First try to make up your mind about the proper definition of time of concentration and travel time. Definition 2 is based on wave travel time and is preferred, but water velocity in definition 1 can be also transformed into wave celerity. Use definition 4 when hyetograph-hydrograph data are available, although rainfall spatial variability in large watersheds and nonreproducibility of response time are two major issues that are not addressed in this definition. Another important consideration is the preference to estimate T_c and T_t as functions of design rainfall intensity in an iterative procedure.

Next is the choice between hydraulic and empirical methods. Choose a hydraulic method when proper data can be collected to apply it and the estimate of T_c is considered a critical step in the project. If repetitive use of the selected method for various watersheds and for many outlet points inside the watershed of interest is required, the development of a GIS-based computer model based on Eq. 2 may be justified. Curves of T_c can be constructed for any watershed by running the model for various rainfall intensities. This leads to calibration of γ and C in Eq. 4.

The next option in hydraulic-based methods may be the combination of the kinematic wave model (Eq. 6) for overland flow, the upland method for concentrated flow, and Manning's equation for channel flow. Travel time through channel reaches must be adjusted based on wave celerity, as discussed in "Hydraulic methods."

In using empirical methods, one must be aware of the conditions and assumptions on which the development of these methods was originally based. It is important to compare the characteristics of the watershed considered with the range of watershed area, land use/cover, slope, and rainfall intensity of the original data set used in developing that empirical method. Simplicity and practicality of empirical methods comes at the expense of lost generality and accuracy.

BIBLIOGRAPHY

1. Linsley, R.K., Kohler, M.A., and Paulhus, J.L.H. (1949). *Applied Hydrology*. McGraw-Hill, New York.
2. Saghafian, B. and Julien, P.Y. (1995). Time to equilibrium for spatially variable watersheds. *J. Hydrol.* **172**: 231–245.
3. Engman, E.T. (1986). Roughness coefficients for routing surface runoff. *J. Irrig. Drainage Eng. ASCE* **112**(1): 39–59.
4. Ponce, V.M. (1989). *Engineering Hydrology*. Prentice-Hall, Englewood Cliffs, NJ.
5. McCuen, R.H. (1989). *Hydrologic Analysis and Design*. Prentice-Hall, Englewood Cliffs, NJ.
6. Soil Conservation Service. (1972). *Travel Time, Time of Concentration, and Lag, National Engineering Handbook*. U.S. Department of Agriculture, Washington, DC.
7. Chow, V.T., Maidment, D.R., and Mays, L.W. (1988). *Applied Hydrology*. McGraw-Hill, New York, Chapter 5, pp. 164–165.
8. Kirpich, Z.P. (1940). Time of concentration of small agricultural watersheds. *Civ. Eng.* **10**(June): 362.
9. Williams, G.B. (1922). Flood discharges and the dimensions of spillways in India. *Engineering (London)* **134**: 321.
10. U.S. Bureau of Reclamation. (1973). *Design of Small Dams*. Washington, DC.
11. Johnston, D., and Cross, W.P. (1949). *Elements of Applied Hydrology*. Ronald Press, New York.

WATERSHED HYDROLOGY[1]

Peter E. Black
State University of New York
College of Environmental
Science and Forestry
Syracuse, New York

INTRODUCTION

Consideration of a drainage basin as the basic unit of supply and management goes beyond merely calling attention to the need for watershed management based on what we typically visualize of the natural environment. This article seeks to establish a broad base for anyone who wishes to base policy decisions on sound science, in this case, on some readily identified and easily comprehended fundamentals about natural and disturbed hydrologic environments. This article organizes, defines, and explains the importance of seven basic functions of watersheds, augmented by some essential topics aimed at a more thorough understanding of the functions.

It is essential to consider a watershed as a basic unit of landscape for managing the water resource. The reasoning underlying this practical concept is actually rooted in the fundamental role of water on this planet more than on the obvious practical management needs. The watershed is, first, a *system* whose primary focus is on interactions between the topography and runoff-producing events. Topography defines a watershed's limits and is a useful framework within which to examine, describe, and manipulate its *functions*, or how the watershed works. Runoff-producing events, rain, snow, or ice storms, and snowmelt induce stream flow that provides sites for chemical reactions and habitat for wildlife as well as pathways for energy, nutrients, and waste products between the land and the oceans. Second, the watershed is a dynamic *hydrologic unit*, whose behavior interacts with biological systems to construct a range of *aquatic ecosystems* that have unique properties reflecting the *quantity, regimen,* and *quality* of the water resource. These watershed functions blend on larger geographical and temporal scales with physical, chemical, and energy components of the universe, thereby affording the opportunity for humans to establish working landscapes that are environmentally sensitive, sensible, and sustainable.

Terminology: Three Important Terms

A watershed (also *catchment* or *drainage basin*) is the natural unit of land on which water collects, is stored, and

---

[1] The material in this article is summarized from *Watershed Hydrology*, Second Edition, 1996, CRC Press/Lewis Publishers, Boca Raton, FL, and, in part, from an animated and narrated nutshell autotutorial short course on compact disk accompanied by a manual under the title *Concepts of Watershed Hydrology*, published by the author in 2002 (www.waterbudget.com).

is redistributed to the atmosphere, streams, and larger water bodies. The simplest definition is "an area of land on which precipitation collects and runs off through a common outlet." The "outlet" is the lowest place on the watershed; it is selected as the best location for monitoring stream flow, its quality, and temporal distribution. The term also applies to the high ground between two drainage basins, which is also known as a *topographic divide*.

Hydrology is *the study of the movement and storage of water in the natural and disturbed environment*. The term also refers to the *condition of the hydrologic environment at some specified time or place*, as in "the hydrology of Eagle River just prior to the June flood...."

Antecedent moisture conditions refers to *a description of meteorologic, vegetative cover and/or land use circumstances, soil characteristics, and the amount of water in storage on a watershed that collectively affect runoff behavior at the start of a hydrologic event of interest*.

Units

Both English and Metric System units are used throughout aquatic disciplines. Most of the well-established stream flow and water quality records made during the twentieth century in the United States and Great Britain were in units of inches, feet, and acres. Other regions of the developed world used the cgs (or metric) system of millimeters, meters, and hectares. Hydrologists need to be familiar with both sets of units and should be able to convert readily from one to the other. Each has benefits. For example, conversion from water quality units of *concentration* in milligrams per liter (mg/L) to *load* in kilograms per hectare is easier in the metric system. The large volume of water represented by the *acre-foot* (43,560 square feet covered by water one foot deep) can be readily derived from the basic unit of stream flow in the English system, one cubic foot per second (cfs, or second-foot), which is a reasonable minimum measurable flow in the field. Thus, a constant flow of 1 cfs for 1 day is very close to 2 *acre-feet/day*, or about a third of a million gallons.

WATERSHED FUNCTIONS

The seven watershed functions are arranged in three groups: three *hydrologic* functions, two *ecological* functions, and two *response* functions. Large watersheds are made up of many small watersheds, so the discussion here focuses on *small* watersheds, where stream flow is dominated by storm flow or discharge from runoff-causing events.[2] How runoff events from small watersheds combine to produce discharges on large watersheds is the subject of extensive research and complex models and research techniques. Runoff from large watersheds is dominated by baseflow (discharge from a groundwater reservoir). The behavior of these two types of flow is very different, but once the water is in the stream, visual separation is often difficult. In contrast, however, the chemical characteristics of storm flow and baseflow are also quite different: they have different storage histories that are reflected chemically. That provides a developing field of research into using water chemistry to understand, model, and therefore predict hydrologic behavior.

Hydrologic Functions

The names of the hydrologic functions derive simply from what they do. The British term for the drainage basin is *catchment;* the U. S. term is *watershed*. Both are self-descriptive. Between them, the water is *stored*.

Collection. The collection function is clearly the first step in a process-oriented description of the hydrologic cycle on land.[3] An essential first consideration in this function is where, how, for how long, and in what form the runoff-causing event impacts the watershed. For example, the same rainstorm at different locations on a watershed will result in different distributions of stream flow (Fig. 1). Thus, a small-area but intense[4] summer thunderstorm near the outlet of a watershed will produce a short-duration, high-peak runoff flow in contrast to the same storm occurring at the upper end of the watershed where the length of stream provides time for the runoff to spread out in time and space, thereby producing a lower peak but longer duration of flow. The hydrograph for the storm near the outlet is said to be "flashy," a common characteristic of small watersheds as well as impervious parking lots and urban watersheds.

It is clear that the location of the runoff-producing event is of considerable importance in determining hydrologic behavior. Size and type of storm, therefore, also become important. Thus, in a region where summer thunderstorms are prevalent, short-duration, flashy stream flow may be expected, especially if the watershed under consideration is at a high elevation where slopes are steeper, soils are thinner, and intense storms can readily cover the entire small watershed. These are *weather* events. In contrast, snowmelt is a *climatic* event that embraces a longer time and greater areal extent than thunderstorms. In between, cyclonic storms and hurricanes may be the dominating determinant of runoff behavior on intermediate size and midelevation watersheds. Climate and weather are characteristic of different areas, so hydrologic behavior is similarly geographically determined.

Summarizing, the collection function reflects how and where the water that falls on a watershed influences the storage and discharge functions. Further, the different types of runoff-causing events interact with and define

---

[2] A small watershed has been defined as (1) from 200 acres to 1 square mile in size for applying the Rational Formula to determine peak flows (1); (2) one in which runoff behavior is dominated by *overland flow* as opposed to *channel flow* (2), and, for legislative/programmatic purposes; and (3) one that is up to 250,000 acres (PL 566, 1954). It is suggested here that the small watershed be defined *functionally* as one on which the water resource issue of concern be under the primary influence of *storm flow* as opposed to *baseflow*.

[3] There is, of course, a hydrologic cycle on waterbodies, too. It consists largely of the same functions as those on terrestrial (land) environments, but in the oceans, without the discharge, attenuation, or flushing functions.

[4] Expressed in inches or millimeters per hour.

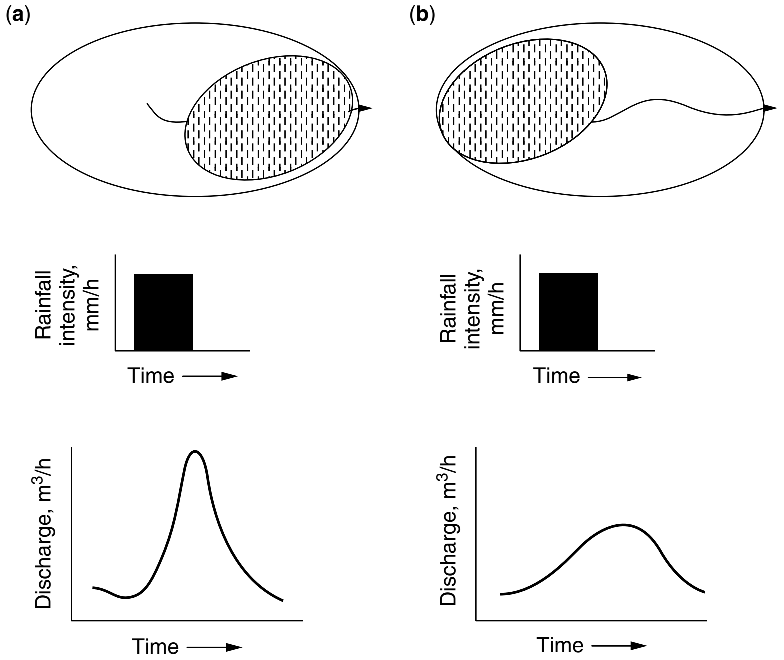

Figure 1. Storm hydrographs from the same rainstorm (hatched area) at different locations on the same watershed. (**a**) Storm is near the outlet. (**b**) Storm is near the headwaters. The volume of water running off might be the same (the area under the hydrographs), but the time pattern of runoff is different.

typical *hydrographic regions* in which homogeneous areas exhibit similar hydrologic behavior.

Storage. The storage function is the focus of a small watershed (Fig. 2). There are six types of storage in and on a watershed: *channel, depression, retention, detention, groundwater*, and *vegetation* storage.

When water arrives at the land surface, three things can happen to it. First, water may fall directly into channels, including small lakes and ponds, as well as streams and rivers. From these, of course, it is simply a continuing process for this water to flow out of the watershed by joining other waters in the stream and be reclassified as channel runoff.

Second, water may be held on the surface as *depression* storage in a thin film or as puddles. Wetted surfaces of all types—on blades of grass or leaves, rooftops, parking lots, even in some cases wetlands—may be considered depression storage. Snowpack is also a form of depression storage. The water thus detained on the surface may enter the soil if the pores in the soil are not already full and are large enough to allow water in; often, water in depression storage evaporates from the surface under the influence of wind, a warm surface, or sunlight following the storm. Runoff from depression storage is minimal.

Third, water may infiltrate (move into) the soil. How much and how fast this occurs depends on the openings or pores in the soil, whether there are small soil grains that can be washed into and clog the pores, how much of the soil pore space is already filled with water, and how much and how fast the water already in the soil can drain downward (*percolation*). Some larger scale factors that affect the infiltration rate[5] include surface roughness, slope of the land, and extent of soil particle size variety, often dependent on the geologic history of the area and its erosion (known collectively as *geomorphology*). If the rate at which the precipitation falls—*rainfall intensity*—is greater than the capacity of the soil to absorb it—*infiltration capacity*—then excess water runs off the top of the soil as *surface runoff*.

Once in the soil, water may be held tightly (close to soil surfaces) and for long periods in the capillary (or *micropores*) against the force of gravity, or it may be held temporarily (not as close to soil particle surfaces) in the larger, noncapillary or *macropores*. The water *retained* for a long time is called *retention storage,* whereas the water temporarily *detained* is *detention storage.*[6] Upon infiltrating the soil, water will move first into retention storage if it is not filled (where the water is pulled in and held at higher tensions) and into detention storage only if the retention storage is already filled, or recharged.

Water in retention storage is held tightly against the force of gravity, so water cannot flow out of retention storage. In contrast, water can flow out of detention storage and is referred to as subsurface runoff.[7] Subsurface flow may go either to a nearby stream or to a *groundwater* reservoir, the body of water beneath the *water table*. A groundwater reservoir is the saturated zone from which water tends to flow off rather slowly as *baseflow* in contrast to the greater degree of flashiness exhibited by *storm flow*. The latter is made up of channel, surface, and subsurface runoff, sometimes referred to as "wet weather flow."

Topography (and geological history) plays a significant role in the development of soil storage reservoirs. Elevation, slope, aspect (of slopes), orientation (of the

---

[5] The rate at which water moves across the air–soil interface, in inches or millimeters per hour.

[6] The difference may be readily committed to memory by recalling the term used to describe the punishment for improper behavior in school when a student would be sent to a temporary hour-long study hall (*detention*), not commitment to life in prison (*retention*).

[7] The terms "runoff" and "flow" are often used interchangeably.

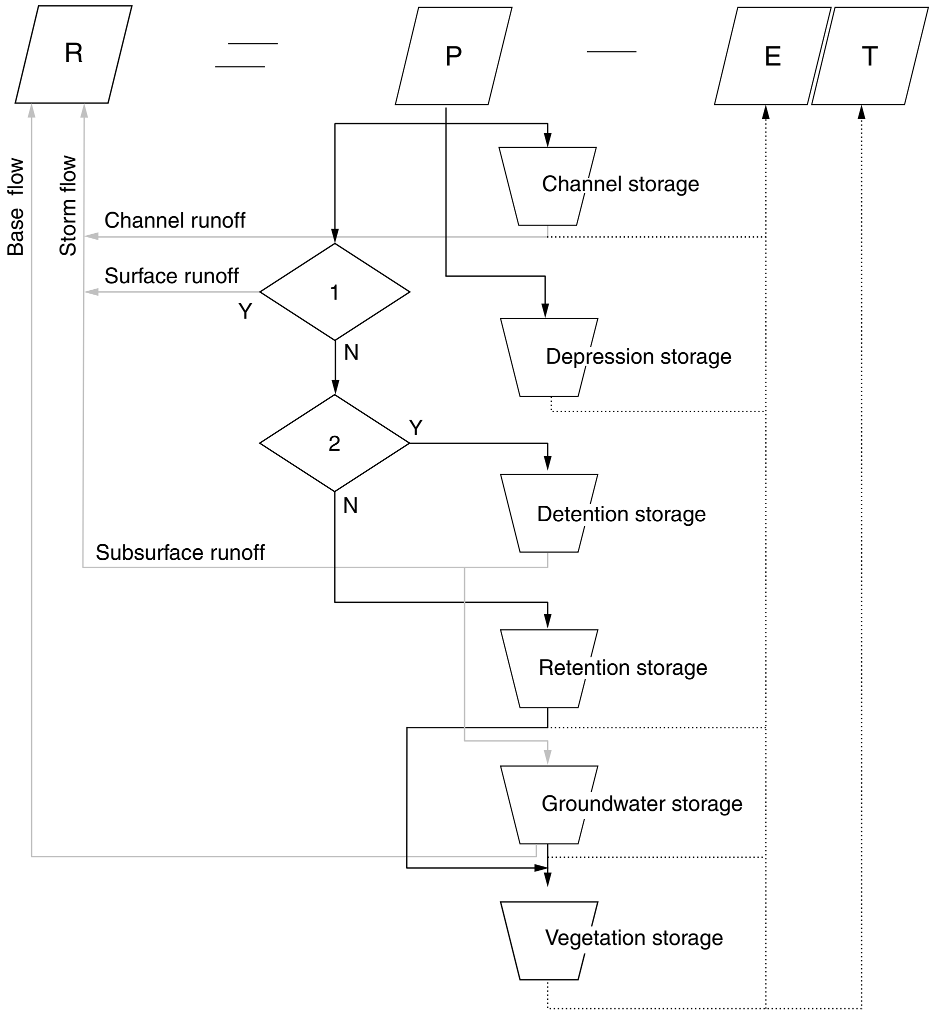

Figure 2. The hydrologic cycle depicted by focusing on the six types of watershed storage and the basic water balance equation (R = Runoff; P = Precipitation; E = Evaporation; and T = Transpiration) at the top. Choice box #1 asks "Does precipitation intensity exceed infiltration capacity?" and choice box #2 asks "Is retention storage full?" Some features, storm flow and baseflow, are kept separate to emphasize that they behave differently; water can leave retention storage only through evapotranspiration (E_t) processes; and water in detention storage, which flows out in response to gravity within about 24 hours, generally is not around long enough to supply E_t at all. See text for full description.

stream), and the distribution of those features throughout the watershed affect the antecedent moisture conditions as well as the storage characteristics of soil reservoirs. For example, south-facing slopes in the Northern Hemisphere are drier than their northern counterparts because they receive more incident energy that is used to evaporate moisture. Soils on south-facing slopes are less well developed and thinner, have lower water holding capacities, and a greater percentage of the soil pores in detention storage where less water is available for vegetation, therefore fewer and/or different species. In short, south-facing slopes tend to be drier than north-facing slopes in the Northern Hemisphere and, of course, the opposite south of the equator.

Finally, water may be stored in *vegetation*. Water enters vegetation when vapor condenses on plant tissues, when it is intercepted as rain or snow, or when absorbed directly through leaves and bark. More water enters through the roots, however, and, because it ascends in response to atmospheric evaporative tension working through tiny openings in leaves called stomates or stomata,[8] it is often combined with the water that evaporates from depression (surface) storage on the vegetative parts as *evapotranspiration* (E_t). E_t refers to all of the evaporative processes and may be modeled as an object of research that focuses on the amount of energy available for evaporation (3). Evapotranspiration is strong enough to pull water away from high-tension films close to soil particles (retention storage). In contrast, water in detention storage is held loosely in the macropores and flows out usually within 24 hours; during that time there usually are waning storm clouds, high humidity, and often cooler temperatures all of which diminish the energy (and water) available for evapotranspiration. The bottom line is that water held in detention storage is not generally available to plants or E_t.

From this examination of the storage function, it is clear that storage type and distribution in the soil profile and topographically over the watershed modify the broad hydrographic patterns imprinted by the climate and weather factors that characterize the collection function. The infiltration capacity concept and the distribution of retention and detention storage are two of the most important concepts in watershed hydrology. And, in practice, the relationship between infiltration capacity and precipitation intensity combines to define the limits to surface runoff and attendant flooding that occurs at the

---

[8] The process is known as *transpiration*.

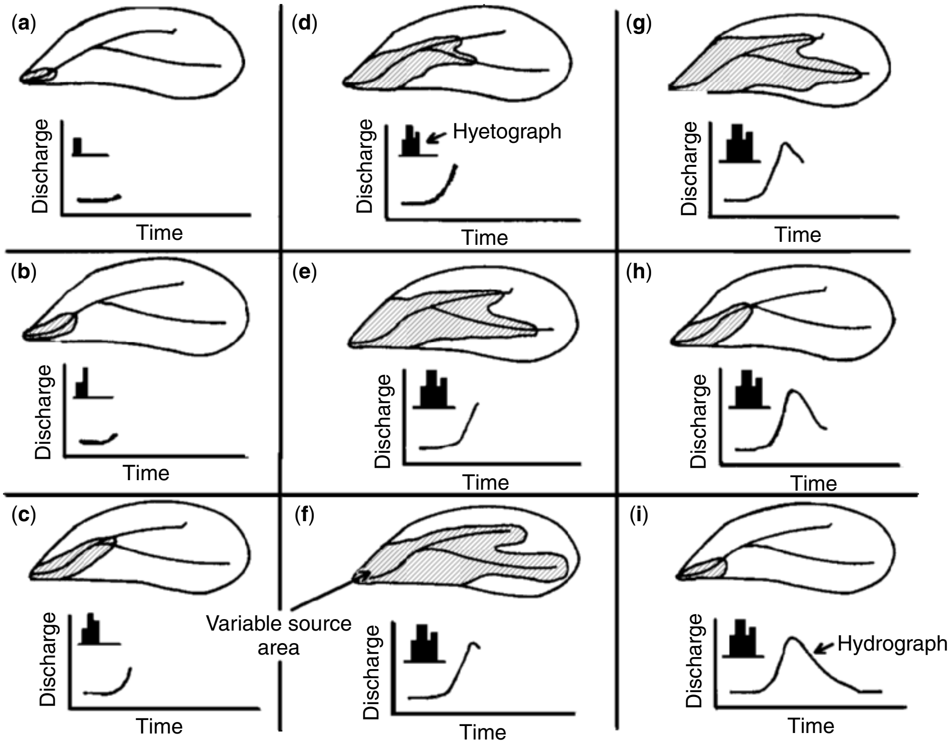

Figure 3. The variable source area. During a storm that uniformly, instantaneously, and completely covers a watershed, the zone contributing to runoff changes.

air–soil interface where humans have the greatest impact on the hydrologic cycle.

Summarizing, the storage function focuses on how water is retained and detained on a watershed and what its characteristics are as it moves between storage reservoirs. Storm flow dominates small watershed runoff behavior. The fact that water can readily absorb elements and compounds gives rise to the chemical function (below) and makes it possible to use water chemistry to monitor and model water storage.

Discharge. As we have just seen, there are two categories of runoff: *baseflow* and *storm flow*, which include channel runoff, surface runoff,[9] and subsurface runoff. Storm runoff is ramified in the *storm hydrograph*, a plot of runoff over time that translates unique watershed characteristics into a typical discharge (often, a flood). Changing land use or vegetative cover can profoundly affect the storm hydrograph on small watersheds and, if done over a wide enough area, even on large ones. The baseflow hydrograph also exhibits patterns that are unique to the watershed from which it flows. However, baseflow integrates the runoff over longer times and broader landscapes, so it is less sensitive to land use and vegetative cover changes. Both types of flow display typical patterns for hydrographic regions, allowing extrapolation of hydrologic information from one watershed to another within the same region. Thus, regional flood behavior curves may be estimated for ungauged streams. The storm hydrograph is influenced primarily by *weather*; the baseflow hydrograph is influenced primarily by *climate*.

An important concept of discharge is that the portion of the watershed that contributes to storm runoff changes during the runoff-causing event, even if uniform rainfall that instantaneously and completely covers the watershed is assumed. The variable source area model identified by Hewlett and Hibbert (4) promotes understanding how and when water gets to the stream (Fig. 3). It also provides a sound scientific basis for protecting and restoring the near-stream *riparian zone* areas, also referred to as *buffer zones* or simply as *buffers*. Water flowing off or through the soil, or in saturated stream bottom zones, appears rather quickly as storm flow, so it follows that water from these zones is likely to be a major source of nonpoint sources of pollution.[10] *Isotope hydrology,* which monitors specified elements or ions in the environment, is one method of tracking water in time and space as it moves through natural and disturbed environments enabling identification of how “old” and “new” water contributes to stream flow. Other isotopic elements and dyes may also be used.

Summarizing, the discharge function integrates antecedent moisture conditions and characteristics of the collection and storage functions. Storm and base flow portions of runoff need to be analyzed separately because they arrive at the stream from different watershed storage reservoirs, are influenced by different characteristics, and consequently behave differently. Both the *storm* and *baseflow* hydrographs are unique to a particular watershed. Hydrographs may be used to analyze, describe, and

---

[9] Under undisturbed and most natural conditions, surface runoff (also referred to as *overland flow*) does not occur. Exceptions are evident on steep slopes that have shallow soils and intense precipitation.

[10] Pollutants that are diffused across the surface of the land and have no apparent end-of-the-pipe outlet as do point sources.

predict stream behavior on ungauged watersheds. A great deal may be learned about the watershed by careful analyzes of the storm and baseflow hydrographs and their relationships to other hydrologic parameters (e.g., precipitation, infiltration, and soil storage) of the drainage basin.

Ecological Functions

Ecological functions refer to *chemical reactions* and *habitat*. Except for nuclear reactions, virtually all chemical reactions on the earth, in our solar system, and throughout the universe, take place in the presence of water, which therefore plays an essential role in environmental chemistry and life itself. The aquatic chemistry of our body cells, organs, and blood, in soil water, streams, wetlands, ponds, and lakes, and in the oceans define life. Life is considered as existing only in the presence of water and, consequently, the ongoing search for extraterrestrial life focuses on a search for water.

Chemical Reactions. As drops of water form in the atmosphere following cooling and condensation, carbon dioxide readily dissolves in them. This shifts the *reaction*[11] (or pH) of natural rainfall from a neutral value of 7.0 to about 5.7 (slightly acid), which makes precipitation rather corrosive: it dissolves many of the elements and/or compounds found in rocks and soils that derive from them. Carbon dioxide, dihydrogen oxide (water), and several forms of carbonate all disassociate (to hydrogen and hydroxyl ions) in aquatic environments and result in various concentrations of hydrogen ions; hence the use of pH to monitor one of the major water quality parameters. Others include *conductivity*, an index to dissolved salts based on how much electricity the water conducts; *temperature*, a measure of the energy content of the water; *turbidity*, a measure of the amount of suspended solids in the water; *dissolved gases* (oxygen, carbon dioxide, methane, ozone, etc.); *dissolved nutrients* (nitrogen, phosphorous, potassium, and sulfur); *heavy metals* (iron, lead, mercury, etc.); *color*, an indication of dissolved substances such as tannin; and *contaminants*, including toxins and disease-causing bacteria and/or viruses, and spores.

Some of the chemical exchanges that are evident in the presence of water may, in turn, be used to assist in evaluating the hydrology of a waterbody: where it came from, with what it has come in contact, and so on. For example, silica dissolves readily when water moves through sandy soil. If that water flows into a wetland, abundant plant life will remove silica from the water column to produce cells, which differ from animal cells in that they have silica-based cell walls. And, because the great growth of plants in wetlands ensures a plentiful supply of decayed plant and animal remains, water flowing out of a wetland will tend to be low in oxygen and high in carbon dioxide (as well as low in silica and pH).

Water is considered a universal solvent (although it won't dissolve the noble metals gold, silver, platinum, etc.) and naturally conveys waste materials from our body cells, organs, individuals, communities, and nations. It even assists in carrying nutrients *upstream* in the bodies of migratory salmon and related species that feed in the oceans and die when they have completed their reproductive journey to the headwaters of the stream of their birth, ironically, succeeding because of the unique stream chemistry that they can trace to the stream in which they were spawned. It is also believed that it plays an essential role in the carbon cycle (5).

The hydrologic cycle also interacts with the energy balance of the earth in two ways. First, water vapor is one of the chief greenhouse gases along with carbon dioxide, methane, ozone, and other two-atom molecule gases. Second, water vapor in the form of cloud cover plays a direct role in the energy balance of the planet by reflecting solar radiation to space, thereby affecting the temperature of the atmosphere and the earth's surface. We need to explore ways in which our land and water management strategies and policies might adversely or beneficially impact the role of water in these critical interactions.

Summarizing, the *chemical reaction* function brings a whole new and exciting dimension to the field of watershed hydrology and land management, not only water chemistry, but environmental and astrophysical chemistry. These are some very significant and fundamental characteristics of the earth and the universe.

Habitat. The habitat function embraces the observation that wildlife (terrestrial, aquatic, and avian) evolved under a variety of conditions and require that they be maintained to survive. Ecosystem support for all species needs to be preserved and/or restored to ensure continued biodiversity in our environment, especially aquatic environments. Human beings themselves may have evolved at the beginnings of civilization at the seashore, in the presence of water and all that it has to offer as sustenance, food, protection, and a myriad of services (6).

This function is not dealt with in depth in this paper, but it is important to point out that a great deal of time and money have been invested in restoring (and preserving) wetlands, and recently increased activity in watershed restoration as well. These preservation and restoration efforts are often lackadaisically directed toward restoring habitat rather than hydrologic functions. But, when habitat—particularly aquatic habitat—is restored, the chances are very good that hydrologic functions have been restored as well.

Summarizing, the habitat function might be viewed as a surrogate for evaluation of restoration of watershed functions that are harder to visualize, measure, or model. The importance of viable ecosystems and associated biodiversity, however, is a value of aquatic habitats that transcends even semiserious personal desires.

Response Functions

The response functions are so called because they represent practical functions that play important roles

---

[11] A measure of the relative acidity and alkalinity of a compound in terms of the hydrogen ion concentration. The range is from 1 to 14 (the numbers are logarithms to condense an otherwise very broad scale), where 7.0 is neutral.

in our aquatic environments. The responses really involve how the storm hydrograph reacts with other considerations: in the *attenuation function*, the peak discharge, Q_p, is seen to diminish as the area of a watershed increases. In the *flushing function*, the concentration of dissolved and/or suspended substances appears to relate to the stages of the storm hydrograph in a predictable manner.

Attenuation. It is the nature of storms to have the most (depth) and most intense (depth per unit time, e.g., inches per hour) precipitation at or near the center, and precipitation diminishes until both become zero. This observation flies in the face of the earlier assumptions of uniform, instantaneous, and complete storm coverage: in fact, the larger the watershed, the longer it takes for water falling at any point to reach the outlet, in addition to the fact that there will be a place where the rainfall is most intense and greatest in depth. Given, too, that slopes and stream gradients generally diminish downstream, it follows that the length of time between the start (or centroid) of the runoff-causing event and the time of occurrence of the consequent peak flow Q_p (in this case, Q_p is reported in units of cubic measure per unit time *per unit area,* e.g., cubic feet per second per square mile, csm, or cubic meters per second per square kilometer) also diminishes (Fig. 4).

Summarizing the attenuation function, it is easy to remember that the three "*L*s" go together: thus the storm peak in runoff per unit area is *L*ower and *L*ater on *L*arger watersheds.

Flushing. Flushing is a natural characteristic of all aquatic ecosystems except the ocean. As relatively pure water arrives at the land surface (or channels), it dilutes whatever substances are already in the stream (Fig. 5). Then, as water flows off the near-stream zones within the variable source area, it carries with it soil particles and dissolved materials and starts a cleansing or flushing action in the stream/watershed system. As the velocity and volume of stream flow increases, more and more particles deposited earlier in the streambed may also be picked up and transported downstream. A plot (*pollutograph*) of the concentration (e.g., mg/L) of dissolved and suspended materials increases until peak flow occurs and then quickly diminishes. Put another way, the plot of the concentration over time, the pollutograph, peaks during the rising limb of the storm hydrograph (or during the rising stage of the annual hydrograph's spring peak, for example), then rapidly decreases until the prerunoff-causing event stream flow level is reached.

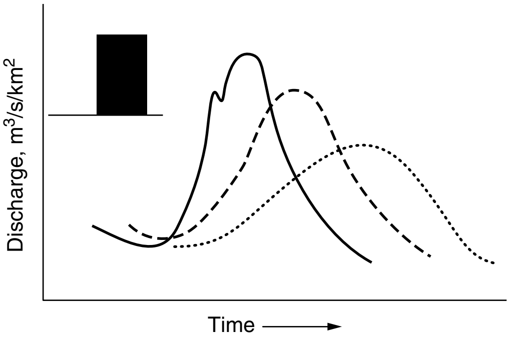

Figure 4. The initial upstream storm hydrograph diminishes and spreads out as the peak moves downstream. Note: the vertical axis is in volume *per unit area,* and the initial channel interception prepeak on the rising limb disappears.

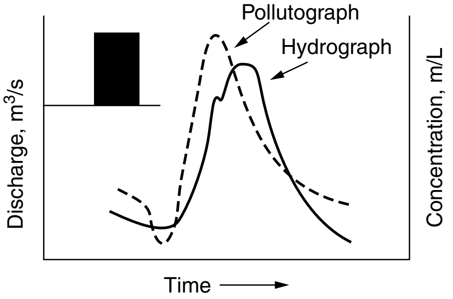

Figure 5. Flushing of some suspended solid (read on right-hand axis) superimposed on storm hydrograph (read on left-hand axis).

Some of the potential pollutographs are well known. For example, a great deal of field data and physical modeling have produced reliable suspended-sediment pollutographs. However, there is a paucity of data and studies on pollutographs for dissolved gases, pH, temperature, and dissolved solids. Intuitively, pollutographs generally follow this pattern. But there are not enough studies on this topic to back up that assertion.

Summarizing the flushing function, nearly pure water from runoff-causing events dilutes water in the stream and initiates dilution and then a flushing action, within and near the stream, that peaks during the rising limb and returns to the preevent level commensurate with stream flow.

SUMMARY

Essential to life as we know it and expect that it evolved on other planets, water pervades our local, regional, national, and worldly environments. We have the greatest access to it locally, but our attention to careful science-based management of water on all scales is demanded to ensure our sustainability. Understanding water in the natural and disturbed environments starts with honoring basic hydrology on a watershed basis.

BIBLIOGRAPHY

1. Dunne, T. and Leopold, L.B. (1978). *Water in Environmental Planning*. W. H. Freeman and Company, San Francisco, CA, p. 818.
2. Chow, V.T. (1970). *Handbook of Applied Hydrology*. McGraw-Hill Book Co., Inc., New York, NY.
3. Reifsnyder W.E., and Lull, H.W. (1965). *Radiant Energy in Relation to Forests*. Technical Bulletin No. 1344, Forest Service, USDA, United State Government Printing Office, Washington, DC.
4. Hewlett, J.D. and Hibbert, A.R. (1967). *Factors Affecting the Response of Small Watershed to Precipitation in Humid*

Areas. Proceedings of the International Symposium on Forest Hydrology, edited by Slupper and Lull, p. 275.

5. Lovett, R.A. (2002). Rain might be leading carbon sink factor. *Science* **296**: 1787.
6. Morgan, E. (1972). *The Descent of Woman*. Stein and Day, New York, NY, p. 284.

WATERSHED MANAGEMENT FOR ENVIRONMENTAL QUALITY AND FOOD SECURITY

G. S. Shrivastava
University of the West Indies
St. Augustine, Trinidad,
West Indies

Watershed management is a broad term. It comprises the planning, implementation, and operation of programs and practices for the optimal utilization and preservation of natural resources such as water, forest, and wildlife. The term is rooted in forestry, and the adverse impacts of deforestation relating to soil erosion, sedimentation, and on water yield are well known (1), and they need not be elaborated here. Suffice it to say that the frequency and severity of floods and droughts, mudslides, and water quality are closely linked to the state of watershed management. In this article, however, watershed management is considered with specific reference to environmental quality and food security.

Environmental quality is an illusive term that embraces physical, psychological, and emotional factors, and biodiversity is an important aspect of the same. Furthermore, an increasing awareness of the critical interdependence of ecology and hydrology has led to the emergence of a new study field of ecohydrology (2). It is a frontier full of challenging and unexplored questions, embracing problems, which are crucially related to biodiversity and ecosystem functioning; e.g., how do hydrologic systems control (or how they are controlled by) ecological systems?

In a given landscape, its flora and fauna dynamically interact with the hydrologic cycle over varying spatial and temporal scales, and the ecological structure and function of ecosystems are intimately tied to the hydrological regime of floods and droughts. Accordingly, ecological diversity is a function of varying flow. Moreover, it is essential to not only identify these interactions, but also to ascertain the limits to perturbation before the ecosystems begin to break down. In this context, estimation of instream flows plays a vital role for the preservation of environmental quality (3). Furthermore, two other aspects are important as well. First, within an overall watershed management strategy, ecotourism development should be mild to moderate lest it endangers the very environment that the tourists want to experience. Second, it should be recognized that the places where biodiversity is greatest are often the same places where conservation measures are most difficult to enforce because of poverty, difficulty in controlling poaching, and habitat destruction.

Food security can be simply defined as the ability of a nation to feed itself by means of its agricultural resources, and such ability is largely dependent on irrigation and drainage infrastructure. Inadequacy of such infrastructure leads to poor crop yields and loss of fertile land because of salinization. However, central to irrigation is the availability of water resources, and it is the watersheds that generate this vital resource. Unfortunately, in some parts of the world, such as in the Middle East and the Caribbean (4,5), watershed management does not get the attention it deserves, because, virtual water, embedded in grain imports, creates an illusory sense of food security, which in turn slows the pace of socioeconomic reform needed to improve watershed management.

In conclusion, although watershed management policies are shaped in the political arena, it is hydrology that can steer decisions in the right direction. Therefore, political and economic pressures should not prevail over hydrologic evidence and reasoning. Ultimately, there is either science and sustainability or crisis and conflict in the sphere of watershed management.

BIBLIOGRAPHY

1. Southwick, C.H. (1996). *Global Ecology in Human Perspective*. Oxford University Press, Oxford, UK.
2. Kundzewicz, Z.W. (Ed.). (2002). Special section on ecohydrology. *Hydrol. Sci. J.* **47**(5): 797–832.
3. Shrivastava, G.S. (2003). Estimation of sustainable yield of some rivers in Trinidad. *J. Hydrol. Eng.* **8**(1): 35–40.
4. Shrivastava, G.S. (2003). Water resources and food security: A Caribbean case study. *J. Water Maritime Eng.* **156**(4): 351–353.
5. Allan, J.A. (2002). Hydro-peace in the Middle-East: Why no water wars? A case study of the Jordan River Basin. *SAIS Rev.* **12**(2): 35–43.

WATER HYACINTH—THE WORLD'S MOST PROBLEMATIC WEED

Adrian E. Williams
APEM Ltd.
Manchester, United Kingdom

INTRODUCTION

Water hyacinth (*Eichhornia crassipes*), a perennial, herbaceous monocotyledon member of the pickerelweed family (Pontederiaceae), is native to tropical America. It is usually found as a floating plant with mature specimens having long feathery roots that can be over 2 m in length (Fig. 1). However, in shallow waters, individuals can root themselves into sediment, and individual plants can even be found fully rooted on banksides. Water hyacinth is found in a number of differing morphological growth forms, and the morphology and architecture of leaves, petioles, and clonal groups can be highly plastic (1–3). However, two forms dominate (4,5). One, typically found in open

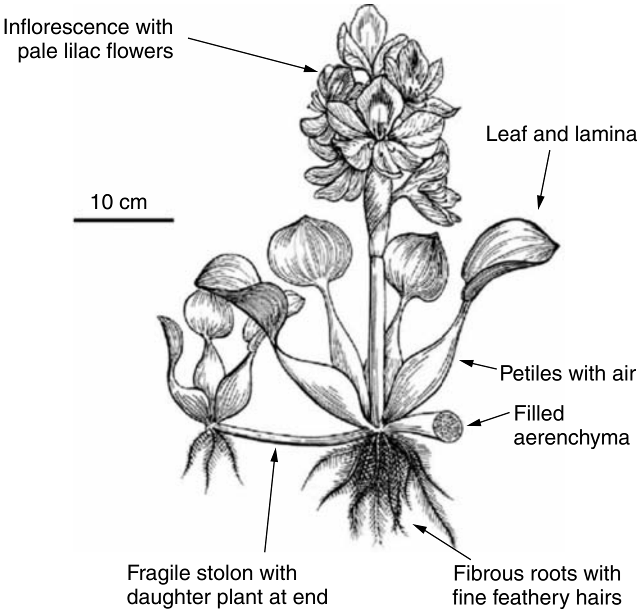

Figure 1. Water hyacinth (*Eichhornia crassipes*).

water or at the edges of plant mats, is characterized by short bulbous and very buoyant petioles containing air-filled aerenchyma, whereas another form occurs in crowded areas, such as in the middle of plant mats, with slender, tall, nonbulbous petioles. The short bulbous form can develop into the tall nonbulbous form if crowding occurs and sufficient nutrients are available (1). Petioles can therefore grow to over a meter in length, although 50 cm is usually more typical. Center and Spencer (2) found that in mature plants, between 6 and 8 leaves or lamina are present on each plant and that an equilibrium establishes between new leaves appearing and old leaves senescing.

Plants can double their biomass in 6 days (6) and although water hyacinth produces seeds, which form part of the waterbodies seed bank, plants generally reproduce through clonal propagation with daughter plants being produced at the ends of fragile runners or stolons. These stolons are easily broken by wind, current, or animal and boat movements, thus allowing the detached clones to disperse and colonize new areas.

In the center of each mature rosette of leaves rises a broad stem, on top of which is an inflorescence that usually bears between 6 and 14 lily-like pale lilac or violet flowers. Each flower consists of 6 petals, the upper one marked with blue and yellow, 6 stamens, and a 3 segmented ovary. The fruit that develops comprises of a simple capsule containing in the region of 400 to 500 seeds. After flowering, the inflorescence bends down into the water and seeds sink to the bottom of the waterbody where they can remain dormant for 20+ years (7) before germinating if the correct conditions prevail. For instance, a waterbody may become temporarily dry, resulting in all the parent plants dying; however, following the replenishment of water to the system, the seeds can germinate, thus providing a new stock of plants.

The fine feathery hairs that surround the main fibrous roots have a purple tinge and capture small particles as well as removing nutrients from the water column. Ideal growth conditions include a high nutrient concentration, especially of nitrogen and phosphorus, with water hyacinth thriving especially well where anthropogenic influences have artificially raised these levels. In addition, tropical temperatures, plenty of sunlight, and a high relative humidity all benefit the growth and expansion of water hyacinth.

With different morphological forms and an ability to adapt the growth form to the prevailing conditions, water hyacinth is able to colonize new areas very quickly. Although mats often develop out from the shoreline of lakes, rivers, and reservoirs, detached floating mats can also exist that are many kilometres in diameter and several meters thick. However, in its native habitat, the plant is kept in check through both natural pests and, more importantly, annual floods. During these periods, rivers may rise by over 10 meters, resulting in a torrent of water that sweeps the plants out to sea where they perish because of the effects of salinity. A few individuals remain, however, snagged on the banks and protected in quieter backwaters, which act as the new source of the following years stock. Overall, therefore, water hyacinth only becomes a significant problem when it finds its way into areas outside of its natural range.

THE PROBLEM

Water hyacinth was seen as a botanical speciality and, because of its beautiful and delicate inflorescence, used to decorate ornamental ponds (6,8). Therefore, from its native range, it was spread around the world during the nineteenth century. It was recorded in Egypt in the 1870s, the United States in the 1880s, Australia and southern Asia by the 1890s, China and the Pacific by the early 1900s, South Africa in 1910, East Africa by the 1930s, and West Africa by the 1970s. It is now established throughout tropical and warm temperate regions of the world (5,9). Moreover, as demonstrated in the Democratic Republic of Congo, where two plants were seen to produce 1200 daughter plants over four months (10), water hyacinth has massive growth potential. In terms of biomass, a m^2 of the weed can increase by up to 500 g each day so within a year, a hectare of plant can increase by over 1800 tons fresh weight, which, of course, does not include the subsequent growth of the new material. This exceptional growth rate is often partly attributable to the presence of large quantities of available nitrogen and phosphorus, and unfortunately, because anthropogenic eutrophication has affected the great majority of waterbodies throughout the world, the growth of water hyacinth and the problems associated with it have been greatly exacerbated. As such, and further aided by its ability to colonize new habitats and a lack of natural controls outside of its native range, water hyacinth is now commonly described as the world's most problematic aquatic weed (Fig. 2).

Invasive species such as water hyacinth have been referred to as a form of biological pollution because they can upset equilibria between native species that have

Figure 2. Lake Victoria in 1995. A mat of water hyacinth stretches out into the lake resulting in enormous ecological, sociological, and economic impacts.

formed slowly over millennia. When an invasive plant such as water hyacinth enters a new habitat it will out-compete and thus displace native vegetation, thereby becoming established. Some of the ways such displacements affect the native biota include a reduction in biodiversity, a decrease in habitat and food, and an overall change to ecological processes (11).

In the case of water hyacinth, native plants along the fringes of rivers and lakes are eliminated through successional shifts and, generally, a loss of biodiversity occurs (12). Furthermore, as native flora and fauna have coevolved over long periods of time, replacement of native vegetative systems with ones dominated by water hyacinth typically alters these relationships with accompanying reductions in available habitat and food. For instance, under the dense mats that can form, a reduced light and oxygen climate can lead to a decrease in plankton primary production, resulting ultimately in the death of native fish.

Ecosystem processes may change as a result of the spread of water hyacinth. Native ecosystems have developed under and adapted to particular abiotic factors and processes, such as rates of nutrient cycling, rainfall patterns, and fire regimes. The presence of alien species can alter these processes, which has a knock on effect to the ecosystem as a whole (13). For instance, where water hyacinth has been introduced, it can disrupt the natural hydrologic cycle by transpiring greater quantities of water, up to three times more, than native vegetation in the same habitat (11). Water tables and some surface water habitats are thus reduced and native species impacted.

Water hyacinth, with its characteristic thick fringes and floating mats, also has more direct impacts on anthropogenic activities. The weed can impede navigation, even in motorized boats, and transportation costs may increase because of the necessity of taking alternative indirect routes or as a result of motoring through the mats, which, if physically possible, increases fuel consumption (14,15). The weed has affected fishermen by fouling their nets and traps, reducing catches, and preventing access to landing sites. In West Africa, the inland waters of Ghana, Burkina Faso, and Togo have been particularly hard hit. Here, not only has the weed slowed down the boats of fishermen, it has also made landing difficult to the point that more of their catch goes off (14). Fish are often an important source of protein to lakeside communities, and without it, health suffers. Moreover, the mats are breeding grounds for carriers of human and animal disease such as malaria, schistosomiasis, encephalitis, and river blindness (11,16). Irrigation schemes have been put under pressure and, indeed, irrigation channels blocked. In many cases, the weed has reduced the supply of clean potable water and caused difficulties in water extraction. Unsurprisingly, its presence has increased disputes between local communities and overall has been responsible for the translocation of some communities away from affected areas when local watering points and lake access have become unavailable. Furthermore, in Lake Victoria, apart from the reduction in national revenue from the reduced export of fish produce, tourism has declined and intermittent closure of the hydroelectric plant in Jinga, because of weed buildup, has caused disruption within Uganda's capital, Kampala (6,16,17).

Unsurprisingly, therefore, the World Bank, together with the United Nations, the International Monetary Fund, and a plethora of other international and national organizations, has pumped hundreds of billions of dollars into tackling the problem of water hyacinth around the world.

METHODS OF CONTROL

Once established, water hyacinth is extremely difficult, if not impossible, to eliminate, and attempts to do so often adversely impact remnant native vegetation (18). More achievable, however, is the control of the weed, but the elimination or even control of waterweeds is a costly process both in fiscal and temporal terms. Three basic methods exist to remove water hyacinth: physical, chemical, and biological controls.

Physical control involves employing manual labor and mechanical harvesters to remove the weed. The use of physical barriers, such as floating booms, to corral water hyacinth and prevent its spread as well as aiding ease of physical removal can also be implemented (14). Often, manual removal at landing stages can be undertaken extensively by locals if they are provided with equipment such as rakes and wheelbarrows. Aquatic vegetation cutters or "Swamp devils" have been used widely throughout the world, and booms are often used in conjunction to reduce weed expansion (19). These techniques have often worked in that weed can be contained and removed. However, often as a result of the sheer scale of the problem, benefits are typically short lived as a cleared site can be inundated with new weed in a matter of days, either through growth of any remaining weed or immigration of plants from outside of the site. Furthermore, booms can become overburdened from the weight of the weed (20) and machinery can breakdown. Unfortunately, with a lack of expertise, tools, and spare parts, the expensive harvesters have, on occasion, been left to rust on the shore. However, physical control has been used successfully throughout the world to manage water hyacinth (21), and in Nigeria, such methods are considered relatively effective (22). Nonetheless, physical control has to be intensive and regular and is thus ultimately expensive. The cost of clearing a hectare of weed is estimated to be between US $2400–30,000 (23,24). Therefore, for a waterbody of more than a few tens of hectares, physical control is not a feasible long-term control option.

Chemical control is, worldwide, one of the most commonly used methods of macrophyte repression. Chemicals can be applied from the air, water, or land, and some degree of accuracy as to where the herbicide lands can be achieved. Chemicals such as diquat, glyphsophate, amitrole, and the amine and acid formulations of 2,4-D are the most effective and commonly used chemicals against water hyacinth. Application needs to be undertaken by trained individuals using correct dosages and applied at appropriate stages of growth. When they are applied incorrectly, problems can be encountered, not least of which is killing nontargeted plant species such as aquatic macrophytes or agricultural crops (18). Herbicides themselves typically only cost in the region of US $ 50 ha^{-1} of treated area; however, application costs whether by plane, boat, or hand are in addition to this (25). Therefore, herbicide use is only viable over small areas; on the larger scale application, it can become prohibitively expensive. However, whatever the operational costs, if weeds grow faster than chemicals can be applied, then the control of water hyacinth is clearly not accomplished (11). Moreover, if the dead or dying weed is not removed, the rotting biomass can lead to localized deoxygenation, which can, in turn, lead to other detrimental impacts on the waterbody (19,26). Furthermore, although plants may appear to have been controlled, the problem may not have been solved if, as in the case of water hyacinth, seeds remain buried in the sediment and propagate some years later or, as is more typical, plants remaining in the marginal vegetation recolonize the waterbody. Repeated applications are therefore required (5).

The main problem associated with herbicide use, and indeed the reason they are not used as widely as they might be, is the potential effect, real or perceived, of chemicals on humans either directly through ingestion or indirectly through reduced income. Many affected rivers and lakes are major sources of drinking water, and therefore the problem of ingestion is clear. Also, the sites most affected by water hyacinth, and therefore most in need of herbicide application, are typically adjacent to lakeshore communities and farmlands, which increases the risk of destroying nontargeted plants and crops. Furthermore, fear concerning the impact on native biota, not just the loss of nontargeted plants, but also the potential for bioaccumulation of contaminant dioxins through the food chain, is also an area of concern for governments considering their own populations health and the sale of fish to international exporters (27). Lake Victoria, for instance, supports an extremely valuable export fishery that earns critical foreign exchange for the riparian countries. The European Union, which is a major importer of fish from the lake, made it very clear that widespread application of chemicals could threaten the marketability of the fish. Finally, if biological control is used (see below), the application of herbicides may have impacts on the biological agents. Although experiments have found that diquat, glyphsophate, and 2,4-D do not increase mortality rates of *Neochetina eichhorniae* (Fig. 3), they did move away from treated plants to untreated ones (28). In addition, researchers found that *N. bruchi* mortality was increased by several of the commonly used herbicides (29).

On the whole, chemical control, if undertaken correctly, can be very effective in the short term, but as with physical control, the use of herbicides is too expensive for repeated use on large expanses of water.

Biological control involves the use of host-specific agents that are naturally occurring enemies of invasive weeds in their native bioregion. The United States Department of Agriculture initiated biological control in Argentina in 1961, and since then, weevils, moths, and fungi have been tested for host specificity (30). Host-specific agents have now been successfully used to control water hyacinth in a variety of localities around the globe (5,11,31). Indeed, by 1998 a total of 7 agents had been released in 33 countries (32). Biological control of water hyacinth has therefore become very popular, and two weevils in particular, *N. eichhorniae* and *N. bruchi*, have been used extensively and successfully in North America, Asia, Australia, and Africa.

In the long term, once the weed is under control, a natural flux between agent and weed should exist, thus

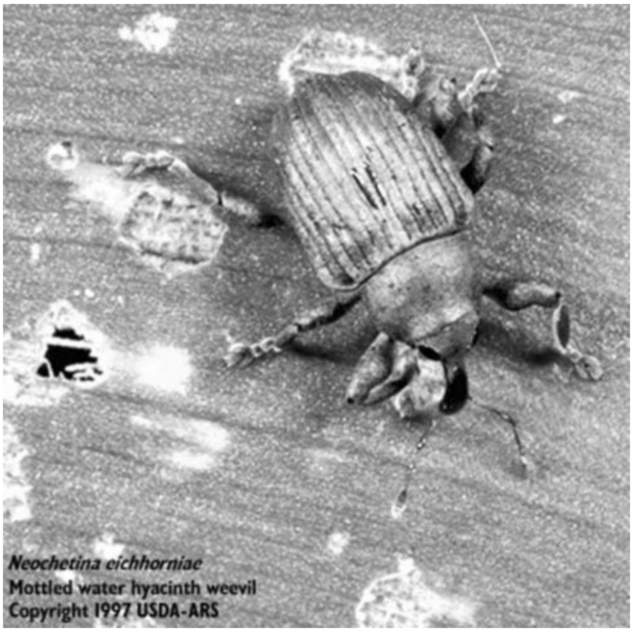

Figure 3. Biological control using the mottled water hyacinth beetle—Neochetina eichhorniae.

keeping in balance the density of weed and numbers of weevils present (33), although, as with all dynamic systems, fluctuations in both populations will occur (34). Biological control can provide an environmentally sensitive, cost-effective, and permanent solution to the problem of water hyacinth. However, the two drawbacks, assuming that host specificity is guaranteed, is that biological control usually takes between 3 to 5 years for the weevil populations to reach a level from which control can be achieved (5,25) and, moreover, it does not always work (35–37).

Overall, of most use in the control of water hyacinth is an integrated method of control. Integrated control employs all three of the above control methods in an overall management strategy. Each of the three is tailored to suit a specific problem, and the extent to which each is used varies with location. In general, however, mechanical and chemical controls are used initially to bring short-term benefits, whereas the biological approach, taking longer to implement and develop, follows after the first two control strategies (11,25). In this regard, maintaining weevil-rearing centers is an essential feature of active management to ensure that weevil populations remain at an adequate level for control.

One other type of control exists that is of key importance but yet is not mentioned as often as the above control methods—the control of nutrient inputs. The levels of available nitrogen and phosphorus have often been cited as the most important factors in limiting water hyacinth growth (38–43). Indeed, a Commonwealth Science Council survey of exotic floating weeds throughout Africa (34) recommended "that action be taken at a national and sub-regional level to reduce nutrient input (particularly from urban and industrial waste) into water bodies, and to develop land-use schemes for long-term utilisation and sustainable development of watersheds." However, in 1993, at a symposium on water hyacinth control, only three papers from 13 country and regional reports mentioned the importance of managing nutrient inputs through improved watershed management. In 1998, during a global working group on biological and integrated control of water hyacinth, in a session dedicated to integrated management, only two out of the 13 papers presented identified the importance of nutrients in controlling water hyacinth. A similar scenario was seen in 2000, at the second meeting of the global working group.

It is understandable and necessary that from within affected countries scientists and practitioners of weed control turn to mechanical, chemical, and biological means of short-term control and management. Moreover, in the short term, all biological possibilities should be vigorously explored, as appropriate insects are available and have been used elsewhere often to good effect. Yet the only real long-term solution to the water hyacinth problem is not biological control but rather the control of anthropogenic eutrophication through sustainable management of watersheds and resources. The reliance of developing nations on biological control as a long-term solution rather than sustainable management and development of their treasured resources is of major concern.

BIBLIOGRAPHY

1. Pieterse, A.H. (1978). The water hyacinth (*Eichhornia crassipes*) - a review. *Abstract of Tropical Agriculture* **4**: 9–42.
2. Center, T.D. and Spencer, N.R. (1981). The phenology and growth of water hyacinth (*Eichhornia crassipes* (Mart.) Solms) in a eutrophic north-central Florida lake. *Aquatic Botany* **10**: 1–32.
3. Watson, M.A., Carrier, J.C., and Cook, G.S. (1982). Effect of exogenously supplied gibberelic acid (GA_3) on pattern of water hyacinth development. *Aquatic Botany* **13**: 57–68.
4. Wright, A.D. and Purcell, M.F. (1995). *Eichhornia crassipes* (Mart.) Solms-Laubach. In: *The Biology of Australian Weeds*. Vol. 1. R.H. Groves, R.C.H. Shepherd, and R.G. Richardson (Eds.). R.G & F.J. Richardson Publications, Melbourne, Australia, p. 314.
5. Julien, M.H., Giffiths, M.W., and Wright, A.D. (1999). *Biological Control of Water Hyacinth. The Weevils Neochetina Bruchi and N. Eichhorniae: Biologies, Host Ranges and Rearing, Releasing and Monitoring Techniques for Biological Control of Eichhornia Crassipes*. ACIAR Monograph No. 60, p. 87.
6. Lindsey, K. and Hirt, H.M. (1999). *Use Water Hyacinth—A Practical Handbook of Uses for the Water Hyacinth from Across the World*. Marianum Press, Kisubi, Uganda, p. 114.
7. Das, R.R. (1969). A study of reproduction in *Eichhornia crassipes* (Mart.) Solms. *Tropical Ecol.* **10**: 195–198.
8. Penfound, W.T. and Earle, T.T. (1948). The biology of water hyacinth. *Ecological Monographs* **18**: 447–472.
9. Gopal, B. (1987). *Water Hyacinth: Aquatic Plant Studies*. Elsevier Science Publishers, Amsterdam, p. 471.
10. Ivens, G.W. (1967). *East African Weeds and their Control*. Oxford University Press, Nairobi, p. 244.
11. De Groot, P.J. (1993). Introduction and summary. In: *Control of Africa's Floating Water Weeds*. A. Greathead and P. J. De Groot (Eds.). Commonwealth Science Council, **CSC 93** (AGR-18) Proc. 295, p. 187.
12. Willoughby, N.G., Watson, I.G., and Twongo, T. (1996). Preliminary studies on the effects of water hyacinth on the diversity, abundance and ecology of the littoral fish fauna in Lake Victoria, Uganda. In: *The Limnology, Climatology and Paleoclimatology of the East African Lakes*. T.C. Johnson and E. Odada (Eds.). Gordon & Breach, Toronto, pp. 643–655.
13. Flack, S.R. and Benton, N.B. (1998). Invasive Species and Wetland Biodiversity. *National Wetlands Newsletter* **20**: 7–11.
14. Akinyemiju, O.A. (1987). Invasion of Nigerian waters by water hyacinth. *J. Aquat. Plant Manag.* **25**: 24–26.
15. Food and Agriculture Organization. (2000). Fighting water weeds in West Africa. *News and Highlights*, no. 705.
16. Twongo, T., Bugenyi, F.W.B., and Wanda, F. (1995). The potential for further proliferation of water hyacinth in Lakes Victoria, Kyoga and Kwania and some urgent aspects for research. *African J. Trop. Hydrobiol. Fisheries* **6**: 1–10.
17. Twongo, T. and Balirwa, J.S. (1995). The water hyacinth problem and the biological control option in the highland lake region of the upper Nile basin - Uganda's experience. *The Nile 2002 Conference: Comprehensive Water Resources Development of the Nile Basin - Taking Off*. Arusha, Tanzania.

18. Njuguna, S.G. and Thitai, G.N.W. (1993). The control of floating water-weed infestations in Kenya. In: *Control of Africa's Floating Water Weeds*. A. Greathead and P. J. De Groot (Eds.). Commonwealth Science Council, **CSC 93** (AGR-18) Proc. 295, p. 187.
19. LVEMP. (2000). *Water Hyacinth and other Invasive Weeds in Lake Victoria: A Status Report*. Lake Victoria Environmental Management Project, Fisheries Resources and Research Institute, Jinja, Uganda.
20. Goodland, R. (1995). *Uganda: Owen Falls Hydroproject: An Environmental Reconnaissance of the Water Hyacinth Problem*. Environment Department Report, World Bank, Washington, DC, p. 23.
21. NAS. (1976). *Making Aquatic Weeds Useful: Some Perspectives for Developing Countries*. National Academy of Sciences, Washington, DC, p. 174.
22. Juweto, E. (1993). Mechanical control of water hyacinth: The Nigerian experience. In: *Control of Africa's Floating Water Weeds*. A. Greathead and P. J. De Groot (Eds.). Commonwealth Science Council, **CSC 93** (AGR-18) Proc. 295, p. 187.
23. Haller, W.T. (1996). Evaluation of the Kelpin 800 aquatic weed harvester. *Aquatic* **18**: 10–15.
24. Farri, T.A. and Boroffice, R.A. (1999). An overview on the status and control of water hyacinth in Nigeria. In: *Proceedings of the First IOBC Global Working Group Meeting for the Biological and Integrated Control of Water Hyacinth*. M.P. Hill, M.H. Julien, and T.D. Center (Eds.). Weeds Research Division, ARC, South Africa, p. 182.
25. Howard, G.W. and Harley, K.L.S. (1998). How do floating aquatic weeds affect wetland conservation and development? How can these effects be minimised? *Wetlands Ecol. Manag.* **5**: 215–225.
26. Mallya, G.A. (1999). Water hyacinth in Tanzania. In: *Proceedings of the First IOBC Global Working Group Meeting for the Biological and Integrated Control of Water Hyacinth*. M.P. Hill, M.H. Julien, and T.D. Center (Eds.). Weeds Research Division, ARC, South Africa, p. 182.
27. Murty, A.S. and Mulendema, B.C. (1999). Growth of the water hyacinth on Lake Kariba and planning for integrated control. In: *Proceedings of the First IOBC Global Working Group Meeting for the Biological and Integrated Control of Water Hyacinth*. M.P. Hill, M.H. Julien, and T.D. Center (Eds.). Weeds Research Division, ARC, South Africa, p. 182.
28. Haag, K.H. (1986). Effects of herbicide application on mortality and dispersive behaviour of the water hyacinth weevils, *Neochetina eichhorniae* and *Neochetina bruchi*. *Environ. Entomol.* **15**: 1192–1198.
29. Patnaik, N.C., Mohapatra, L.N., Ghode, M.K., and Misra, B.K. (1987). Effects of some weedicides on the water hyacinth weevil (*Neochetina bruchi* Hustach) and its weed host. *Plant Protection Bulletin - India* **39**: 23–26.
30. Center, T.D. (1994). Biological control of weeds: Water hyacinth and water lettuce. In: *Pest Management in the Subtropics. Biological Control—A Florida Perspective*. D. Rosen, F.D. Bennet, and J.L. Capinera (Eds.). Intercept Ltd., Andover, UK, p. 737.
31. Harley, K.L.S. (1990). The role of biological control in management of water hyacinth, *Eichhornia crassipes*. *Biocontrol News and Information* **11**: 11–22.
32. Julien, M.H. and Griffiths, M.W. (1998). *Biological Control of Weeds. A World Catalogue of Agents and their Target Weeds*. CAB International, Wallingford, UK, p. 240.
33. Howard, G.W. (2000). Control options: Freshwater invasives. In: *Best Management Practices for Preventing and Controlling Invasive Alien Species*. G. Preston, A.G. Brown, and E. van Wyk (Eds.). Symposium Proceedings, Cape Town, South Africa, p. 316.
34. Harley, K.L.S. (1993). Floating aquatic weeds: What are they? Where are they? How should we manage them? In: *Control of Africa's Floating Water Weeds*. A. Greathead and P. J. De Groot (Eds.). Commonwealth Science Council, **CSC 93** (AGR-18), Proc. 295, p. 187.
35. Salmah, M.R., Mansor, M., and Ahmad, A.B. (1991). A preliminary study of the distribution of *Neochetina eichhorniae*; a possible biological agent for water hyacinth in Kerian District, Perak. In: *Proceeding Asas dan Gunaan dalam Biologi*. A.A. Bidin (Ed.). UKM, New Delhi.
36. McFadyen, R.E.C. (2000). Successes in biological control of weeds. In: *Proceedings of the X International Symposium on Biological Control of Weeds*. N.R. Spencer (Ed.). Montana State University, Bozeman, MT, p. 149.
37. Julien, M.H., Giffiths, M.W., and Stanley, J.N. (2001). *Biological Control of Water Hyacinth 2. The Moths Niphograpta Albiguttalis and Xubida Infusellus: Biologies Host Ranges and Rearing, Releasing and Monitoring Techniques for Biological Control of Eichhornia Crassipes*. ACIAR Monograph No. 79, p. 90.
38. Carignan, R. and Neiff, J.J. (1994). Limitation of water hyacinth by nitrogen in subtropical lakes of the Paraná floodplain (Argentina). *Limnol. Oceanog.* **39**: 439–443.
39. Heard, T.A. and Winterton, S.L. (2000). Interactions between nutrient status and weevil herbivory in the biological control of water hyacinth. *J. Appl. Ecol.* **37**: 117–127.
40. Musil, C.F. and Breen, C.M. (1985). The development from kinetic coefficients of a predictive model for the growth of *Eichhornia crassipes* in the field. I. Generating kinetic coefficients for the model in greenhouse culture. *Bothalia* **15**: 689–703.
41. Reddy, K.R., Agami, M., and Tucker, J.C. (1989). Influence of nitrogen supply rates on growth and nutrient storage by water hyacinth (*Eichhornia crassipes* (Mart.) Solms) plants. *Aquat. Bot.* **36**: 33–43.
42. Reddy, K.R., Agami, M., and Tucker, J.C. (1990). Influence of phosphorus on growth and nutrient storage by water hyacinth (*Eichhornia crassipes* (Mart.) Solms) plants. *Aquat. Bot.* **37**: 355–365.
43. Reddy, K.R., Agami, M., D'Angelo, E.M., and Tucker, J.C. (1991). Influence of potassium supply on growth and nutrient storage by water hyacinth. *Bioresource Technol.* **37**: 79–84.

WATER QUALITY IN PONDS

Nitish Priyadarshi
Ranchi University
Ranchi, Jharkhand, India

INTRODUCTION

A pond is a body of water usually smaller than a lake, encircled by vegetation, and generally shallow enough for sunlight to reach the bottom. Rotted plants can grow in any spot within the pond. Widespread plant growth, such as pickerelweed, spatterdock, and cattails, is quite common in and around ponds and aids in creating a habitat for various forms of animal life, including insects (e.g., damselflies and dragonflies), bullfrogs, greenfrogs, red winged blackbirds, marsh wrens, and bats.

Temporary ponds, that is, ponds that are dry for part of the year, are especially interesting and support a unique community. Organisms in such ponds must be able to survive in a dormant stage during dry periods or be able to move in and out of ponds, as can amphibians and adult aquatic insects. The temporary pond is a favorable place for those organisms adapted to it because interspecific competition and predation are reduced (1).

The chemical composition of any pond water ultimately depends on those elements and compounds in solution, in suspension, and those accumulating along the pond bottom, most of which come secondarily to the ponds from runoff. Other substances are produced by chemical reactions between the different elements and/or compounds by vegetation growing in or around the ponds.

POND WATER QUALITY

Salts

Water in closed pond basins is unlike the waters of other ponds in two respects: (1) it continually experiences a change in chemical composition because (2) it continually increases its concentration of various salts, mainly due to a high evaporation rate (2). However, the increase in salts in a closed basin is not infinite. Certainly, as water flows into a dry pond basin, perhaps from the first rains of late summer or early fall, it carries the salts from rocks surrounding the lake basin and dissolves the surface salts of the playa, providing they have not been removed by deflation. As precipitation declines, generally after the early spring rains, evaporation becomes greater than inflow which then starts the inevitable process of concentration of the salts in solution. Salts concentrated in the waters of closed basins depend mainly on the rock type(s) of the surrounding basins and/or the chemical characters of any influents and, to a minor degree, on several inorganic ions in rainwater.

Pollutants

This catchall category of pollutants means anything added to the pond that is not wanted in the pond. Pollutants can consist of items that may or may not be harmful to human beings as well as fish, and they may or may not be visible. They may float on the surface, sink to the bottom, or dissolve in the water. They may come from outside the pond or from within the pond itself, for example, oil leaking from a submerged pump. Some pollutants are easy to identify and control and/or remove, such as leaves, pollen, and dead rats; some just add to the filter load if not removed but cause few other problems if not in excessive amounts, for example, bird droppings. Most harmful pollutants that dissolve in water are hard to identify or quantify. Surface water runoff that can enter the pond is often a major source of pollutants. That is why all ponds should be designed with a raised edge or at least some type of channel around them so that the surface water does not enter.

Dissolved Oxygen

Whenever air is in contact with water, whether through natural or artificial means, a transfer of oxygen from the air to the water takes place until the water becomes saturated. Plants under light convert carbon dioxide to oxygen in water. Fish, plants at night, and aerobic bacterial action consume the oxygen.

As the air components dissolve in the water, a point is reached where no more can be added. This point is called saturation. The saturation points are different for each of the gases and depend on several factors, but temperature is the most important. As the temperature increases, the water simply cannot hold as much of each type of gas. For oxygen, the approximate saturation level is 11.5 mg/L at 50 °F, 9 mg/L at 70 °F, and 7.5 mg/L at 90 °F. Impurities added to the water (i.e., salt) decrease these saturation levels.

Water whose oxygen concentration is less than 3 mg/L will not generally support fish. When the concentration falls to about 3–4 mg/L, fish start gasping for air at the surface. Levels from 3–5 mg/L can normally be tolerated for short periods. Above 5 mg/L, almost all aquatic organisms can survive indefinitely, provided other environmental parameters are within allowable limits.

Surface oxygen from 11–15 mg/L creates a highly eutrophic condition. Depletion of dissolved oxygen in pond water can encourage the microbial reduction of nitrate to nitrite and sulfate to sulfide, giving rise to odor problems. It can also cause an increase in the concentration of iron (II) in solution (3).

Color

The color of pond water depends on the depth, the color of the surrounding area and sky, the scattered light rays coming from the water, and in most cases, absorption of scattered light rays by dissolved or suspended materials.

Pond water is typically blue for pure, clear water, and various shades of blue–green to a definite green, then into the yellows and browns; the various colors are produced by the innumerable contaminants (Table 1). In general, black pond water results from high organic content, and blue pond water results from an absence of contaminating materials. The color of most pond water exists because of light scattering by vibrating water molecules blue predominates because molecular scattering is greater for shorter wavelengths.

Ammonia

Ammonia (NH_3), in parts per million (ppm), is the first measurement to determine the “health” of a biological converter. Ammonia should not be detectable in a pond that has a “healthy” bioconverter. The ideal and normal measurement of ammonia is zero. When ammonia is dissolved in water, it is partially ionized depending upon the pH and temperature. Ionized ammonia is called ammonium and is not toxic to fish. As the pH drops and the temperature decreases, ionization and ammonium increase, which decreases the toxicity. The presence of ammonia in ponds at higher than geogenic levels is an important indicator of fecal pollution (5).

Table 1. Variation in Water Color[a]

Color	Contaminants
Clear blue (pure water)	No contaminants of suspended materials; color due to molecular scattering
Bluish green	Suspension of blue-green algae, phytoplankton
Green	Colloidal $CaCO_3$ (hard water)
Yellowish green	Suspended sulfur
Greenish yellow	High plankton content, excluding humic debris
Yellow	Diatoms
Yellow brown	Diatoms-fluorescent compounds, humic acids, carboxylic acids, organic matter
Gray	*Spirostomum*
Black	Organic debris, oxidation of ferrous iron, *Stentor*
Red to orange to purple	Suspended clays, ferrous hydroxide, *Oscillataria rubescens, Euglena sanquinea, Haematococcus pluviolis, Dunaliella salina, Bacterium halobium*, purple sulfur bacteria, zooplankton.

[a] Reference 4.

Ammonia tends to block oxygen transfer from fish gills to the blood and can cause both immediate and long-term gill damage. The mucous-producing membranes can be destroyed, reducing both the external slime coat and damaging the internal intestinal surfaces. Fish suffering from ammonia poisoning usually appear sluggish, often at the surface as if gasping for air.

Ammonia is a gas primarily released from fish gills as a metabolic waste from protein breakdown and from some lesser secondary sources such as bacterial action on solid wastes and urea.

When ammonia is detected (assuming a pH of about 7.5),

1. Increase aeration to maximum. Add supplemental air if possible.
2. Stop feeding the fish if detected in an established pond.
3. For an ammonia level of 0.1 ppm, conduct a 10% water change out. For a level of 1.0 ppm, conduct a 25% change out.
4. Consider transferring fish if the ammonia level reaches 2.5 ppm.
5. Retest in 12 to 24 hours.

Nitrite

Nitrite is produced by the autotrophic *Nitrosomonas* bacteria that combine oxygen and ammonia in the bioconverter and to a lesser degree on the walls of the pond.

Nitrite has been termed the invisible killer. The pond water may look great, but nitrite cannot be seen. It can be deadly, particularly to the smaller fish, in concentrations as low as 0.25 ppm. Nitrite damages the nervous system, liver, spleen, and kidneys of fish.

Whenever 0.25 ppm of nitrite or more is detected in a pond,

1. Increase aeration to maximum. For a nitrite level of 1 ppm or greater, add supplemental air, if possible.
2. Discontinue use of any UV sterilizers and ozone generators.
3. For a nitrite level less than 1 ppm, conduct a 10% water change out, and add 1 pound of salt per hundred gallons of changed water.
4. For nitrite levels of 4.0 or greater, consider transferring fish.

Nitrate

Nitrate is produced by the autotrophic *Nitrobacter* that combine oxygen and nitrite in the bioconverter and to a lesser degree on the walls of the pond. The nitrate concentration in surface water is normally low but can reach high levels as a result of agricultural runoff, refuse dump runoff, or animal wastes. Concentrations from zero to 200 ppm are acceptable. Ammonia and nitrite are toxic to fish, but nitrate is essentially harmless.

The nitrate concentration is controlled naturally through routine water change outs and to a lesser degree through plant/algae consumption.

Temperature

The temperature of the water in a pond is seldom constant; variations depend on depth, season, and geography. Pond water is heated by absorption of solar radiation. Events generally happen faster at higher temperature and in smaller ponds. Over normal temperature ranges, biological activity doubles for each 10 °C rise in temperature. The toxicity of ammonia increases as the temperature rises, and the amount of dissolved oxygen that the water can hold decreases.

Direct sunlight during the day can cause the temperature to rise higher, and heat loss on clear nights can cause the temperature to drop lower than that of shaded ponds. A clear night sky can absorb a large amount of heat from a small pond and actually drive the pond temperature below air temperature.

Plant Nutrients and Cultural Eutrophication

Pond water clarity (transparency) is affected by sediments, chemicals, and the abundance of plankton organisms; it is a useful measure of water quality and water pollution. Ponds that have clear water and low biological productivity are said to be oligotrophic. By contrast, eutrophic waters are rich in organisms and organic materials. Eutrophication, an increase in nutrient levels and biological productivity, often accompanies successional changes in ponds. Surrounding runoffs bring in sediments and nutrients that stimulate plant growth. Over time, ponds often fill in, becoming marshes or even terrestrial biomes. The rate of eutrophication depends on water chemistry and depth, volume of inflow, mineral content of the surrounding watershed, and the biota of the pond itself (6).

Human activities can greatly accelerate eutrophication, an effect called cultural eutrophication. Cultural eutrophication is caused mainly by increased nutrient input into a waterbody. Eutrophication produces "blooms" of algae or thick growth of aquatic plants stimulated by elevated phosphorus or nitrogen levels (Fig. 1). Bacterial populations then increase, fed by larger amounts of organic matter. The water often becomes cloudy or turbid and has unpleasant tastes and odors. Cultural eutrophication can accelerate the "aging" of a waterbody enormously over natural rates. Ponds that normally might exist for hundreds of years can be filled in a matter of decades.

pH

The pH of water is a measure of the acid–base equilibrium and, in most natural waters, is controlled by the carbon dioxide–bicarbonate–carbonate equilibrium. An increased carbon dioxide concentration in pond water will therefore lower pH, whereas a decrease causes it to rise. Temperature also affects the equilibria and the pH. In pure water, a decrease in pH of about 0.45 occurs as the temperature is raised by 25 °C.

A pH measurement helps to determine if water is suitable for fish. The pH should normally be between 7.0 and 8.5; it is probably acceptable if it is anywhere between 6.0 and 9.0. Although most fish could tolerate a pH as low as 5.0, bioconverter bacteria are subject to damage. Long-term conditions above pH 9.0 can cause kidney damage to some fish.

A sudden change of a half or more pH unit in an established pond is an indication that something happened, and the cause should be determined. An increasing pH trend in a pond is normally caused by lime leaching out of concrete and to a lesser degree by concentration due to evaporation and decomposing organic matter. Decreasing pH is due primarily to bacterial action that releases acidic compounds. Established ponds normally maintain their equilibrium pH if sludge and decaying organic material are routinely removed. Scheduled water change outs (10% per week for small pond, less for larger ponds) are also helpful. At pH extremes approaching 4 or 11, remove any remaining fish.

Figure 1. Eutrophic pond. Nutrients from agriculture and domestic sources have stimulated the growth of aquatic plants. This reduces water quality, alters species composition, and lowers the pond's aesthetic values.

Saprotrophic Organisms

Aquatic bacteria, flagellates, and fungi are distributed throughout a pond, but they are especially abundant in the mud–water interface along the bottom where bodies of plants and animals accumulate. A few of the bacteria and fungi are pathogenetic, attack living organisms, and cause disease, but the great majority begins attack only after an organism dies. When temperature conditions are favorable, decomposition occurs rapidly in a body of water; dead organisms do not retain their identification for very long but are soon broken up into pieces, consumed by the combined action of detritus-feeding animals and microorganisms, and their nutrients are released for reuse (1).

BIBLIOGRAPHY

1. Odum, E.P. (1971). *Fundamentals of Ecology*. W.B. Saunders, Philadelphia, PA, p. 12.
2. Reeves, C.C, Jr. (1968). *Introduction to Paleolimnology*. Elsevier, New York, pp. 42–50.
3. Guidelines for Drinking Water Quality. (1999). *Health Criteria and Other Supporting Information*. Vol. 2, 2nd Edn. World Health Organization, Geneva, Ph.D., A.I.T.B.S., New Delhi, p. 324.
4. Shapiro, J. (1956). *The Coloring Matter of Natural Waters*. Ph.D. Thesis. Yale University, New Haven, CT. (unpublished).
5. International Organization for Standardization. (1984). *Water Quality-Determination of Ammonium*. Geneva (ISO5664:1984; ISO6778:1984; ISO7150-1:1984).
6. Cunningham, W.P. and Cunningham, M.A. (2002). *Principles of Environmental Science Inquiry and Applications*. Tata McGraw-Hill, New Delhi, p. 234.

WATER TURBINE

Xing Fang
Lamar University
Beaumont, Texas

Water turbine, also called a hydraulic turbine, is a rotary machine that converts the power of flowing water into usable electronic or mechanical power. The word *turbine* is derived from the Latin word for "whirling" or a "vortex". Water turbines were developed in the nineteenth century during the Industrial Revolution, using scientific principals and methods, for industrial power. Now they are mostly used for hydropower generation, which typically requires a construction of a dam to form a reservoir. They are one of the cleanest producers of power, replacing the burning of fossil fuels and eliminating nuclear waste. They use a renewable energy source and are designed to operate for decades. They produce significant amounts of the world's electrical supply. On the other hand, there are some negative consequences; by their nature, water

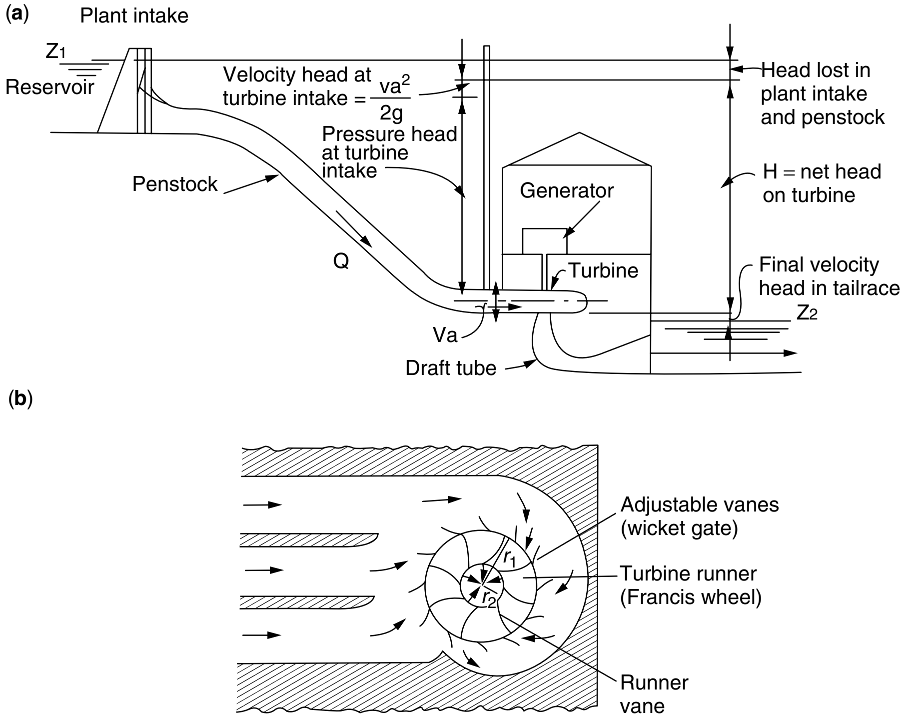

Figure 1. Hydropower arrangement for reaction turbine. (**a**) Hydropower arrangement from upstream reservoir to downstream river. (**b**) Plane view of reaction turbine installation.

turbines interrupt the natural ecology of rivers, stopping migrations of fishing.

A water turbine is usually connected to an elevated reservoir (Fig. 1) through large pipe (penstock), which transfers potential energy of water into kinetic energy. Total hydraulic head (H) available for a turbine is equal to the elevation difference of water surfaces between upstream reservoir (Z_1) and downstream river (Z_2) minus any friction and local energy losses (HL) though penstock, inlet, and turbine system. The electric power generated by a water turbine is

$$P = \frac{\gamma Q H \eta}{1000} = \frac{\gamma Q (Z_1 - Z_2 - HL)\eta}{1000} \text{ (SI units)}$$
$$P = \frac{\gamma Q H \eta}{550} \text{ (English units)} \tag{1}$$

where η is the turbine efficiency and P is the electric power in kilowatts for SI units and in horsepower for English units. Large industrial water turbines typically operate at efficiencies greater than 90%, which is much higher than overall efficiency of thermal power generation (on the order of 30%). From Equation 1, it indicates that decrease of energy losses in penstock and inlets results in an increase of power generation by a turbine. Power generation depends on elevation differences of water surface in upstream reservoir and downstream river, which fluctuate over season and over year. Water turbines can run continuously to provide electricity generation but typically vary power output over a season. Water turbine system can be relatively easy to turn on and turn off in comparison a with a thermal or nuclear power system; therefore, some of water turbines only provide electricity during high-power demand periods (e.g., during the day). When the power network has extra power during night, e.g., developed by thermal and nuclear power plants, extra power can pump water from a downstream reservoir to an upstream reservoir and can be used for future power generation. A pump turbine is used for the above purpose and is a hydraulic machine that operates as a pump in one mode of operation and as a turbine in a different mode. A turbine and generator can supply power under varying load demands at a constant rotational speed by adjusting control wicket gates (Fig. 1), but overall efficiency and performance of turbine will vary somewhat with the gate opening.

Water turbines are divided into two groups: impulse turbines and reaction turbines, according to their hydraulics action. Impulse turbines are the modern forms of water wheels that have been used for thousands of years for industrial power. The impulse turbine is typically used for high-head hydropower installation (e.g., $H > 1500$ ft) but for relative small discharge. In the impulse turbine, jets of water issuing from one or several nozzles impinge on vanes (buckets) of the turbine wheel (also called runner), which reverses the flow of water, and the resulting change in momentum (impulse) causes the turbine to spin. A nozzle in reducing flow area is to create a high speed of water jet (V_j). The fluid momentum equation can determine what forces can be developed on buckets because of a water jet. When the speed of turbine runner is half of the speed of water jet, the optimum power is developed by an impulse turbine:

$$P = \gamma Q \left(\frac{V_j^2}{2g}\right) \tag{2}$$

Reaction turbines are typically used for low-head (1–100 ft) and medium-head (90–1500 ft) hydropower plants with moderate-to-high discharges. For reaction turbines, only part of the available head (H) is converted into kinetic energy, and a substantial part remains as a pressure head that varies throughout the turbine water passage. Reaction turbines include radial-flow Francis turbines for medium-head and axial-flow Kaplan turbines with or without adjustable runner blades for low-head hydropower plants. In 1849, James B. Francis conducted sophisticated tests, developed engineering methods for water turbine design, and improved the efficiency of reaction turbine to over 90%. The Francis turbine, named for him, was the first modern water turbine. It is still the most widely used water turbine in the world today. The Kaplan turbine, a propeller type machine, was invented around 1913 by Victor Kaplan. It was an evolution of the Francis turbine, but it revolutionized the ability to develop low-head hydropower plants. S-type and bulb turbines are other axial flow reaction turbines used for tidal rivers or estuaries that have only several feet of water surface elevation differences. They can generate electric power for tidal flow in two different flow directions.

As the swirling mass of water spins into a tighter rotation of a reaction turbine, it tries to speed up to conserve energy. This property acts on the runner; in addition to the water's falling weight, water pressure decreases to zero as it passes through turbine blades and gives up its energy to make the turbine runner to rotate continuously. Francis turbines have fixed runner blades and use a series of adjustable wicket gates (Fig. 1) to control flow into turbine runner to achieve high efficiency under various flow rate and available head conditions. Flow for the Francis turbine enters the runner in a radial direction, whereas flow of the Kaplan turbine enters the runner in a direction parallel to the turbine's rotational axis. A Kaplan turbine can have adjustable blades for its runner to improve turbine efficiency.

The specific speed (n_s or N_s) of a turbine characterizes the turbine's shape and performance in a way that is not related to its size, which allows a new turbine design to be scaled from an existing design of known performance. The specific speed is the main criteria to select the correct turbine type for matching a specific hydropower development.

$$n_s = \frac{nQ^{1/2}}{g^{3/4}H^{3/4}} \quad \text{or} \quad N_s = \frac{NP^{1/2}}{H^{5/4}} \tag{3}$$

The specific speed n_s is a dimensionless parameter; n is the rotational speed of turbine runner in rps (revolution per second), Q is the discharge passing through turbine in ft^3/s or m^3/s, H is the available head in ft or m, and g is the acceleration of gravity as 32.2 ft/s^2 or 9.81 m/s^2. The specific speed N_s requires N in rpm (revolution per minute), P in horsepower, and H in feet. The specific speed of a turbine is given by the manufacturer (along with other ratings) and will always refer to the point of maximum efficiency. These allow accurate calculations to be made of the turbine's performance for a range of heads and flows. Given a flow and a head for a specific hydropower development site, and the rpm requirement of the generator, one can calculate the specific speed. The result is the main criterion for turbine selection; e.g., impulse turbines have N_s less than 20; Francis turbines have N_s about from15 to 85; and Kaplan turbines have N_s about from 65 to 195 (1).

What should be the diameter of a turbine runner (D)? It is dependent on the turbine's speed ratio ϕ:

$$D = \frac{60 \times \phi\sqrt{2gH}}{\pi N} \quad \text{or} \quad \phi = \frac{u}{\sqrt{2gH}} = \frac{\pi DN/60}{\sqrt{2gH}}$$

where u is the linear rotational speed of a turbine runner and N is the rotational speed in rpm. For a 60-cycle frequency, N is equal to 7200 divided by the number of poles (n) in the generator, which must be an even number. In Europe and South America, where the frequency is 50 cycles/second, the speed N is given as 6000/n. The ranges of speed ratios for maximum efficiency for the different types of turbines are 0.43–0.47 for impulse turbine, 0.5–1.0 for Francis turbine, and 1.5–3.0 for Kaplan propeller turbine.

Turbines and pumps are susceptible to cavitation. Cavitation can damage the turbine runner and affect turbine performance. The region of a turbine where cavitations are most likely to occur is on the downstream side of the turbine impeller blades. The cavitation index (σ) can evaluate the possibility of turbine cavitation:

$$\sigma = \frac{(P_{atm} - P_v)/\gamma - (Z_t - Z_2)}{H}$$

where P_{atm} and P_v are the local absolute atmospheric pressure and vapor pressure of water and Z_t and Z_2 (Fig. 1) are elevations of the downstream side of the turbine runner and tailwater. Lower values of σ indicate a greater tendency for cavitation. The critical sigma (σ_c) for different types of turbines are obtained experimentally and typically provided by turbine manufacturers. When σ is greater than σ_c, there is no cavitation; otherwise cavitation is highly possible.

BIBLIOGRAPHY

1. Roberson, J.A., Cassidy, J. J., and Chaudhry, M. H. (1997). *Hydraulic Engineering*, 2nd Edn. John Wiley & Sons, New York.

WETLANDS: USES, FUNCTIONS, AND VALUES

Matthew P. McCartney
International Water
Management Institute
Pretoria, South Africa

Throughout history, wetlands have played an important role in human development. They have brought benefits, but also caused difficulties for people. Their perceived value to society has varied from place to place and has changed over time.

In many arid and semiarid regions of the world, societies have flourished because of the water and other resources provided by wetlands. Many of the great civilizations of history have benefited from wetlands. For example, the Maya, Inca, and Aztec civilizations in Latin America; the Khmer civilization of Angkor in the lower basin of the Mekong; and the civilizations of the Nile, Niger, and Mesopotamia were all dependent on wetland resources. In contrast, in wetter parts of the world, wetlands were for many centuries widely thought of as unproductive wastelands often associated with disease, difficulty of access, and danger. In East Anglia in the United Kingdom, poor drainage within an area of approximately 3,870 km^2 of fens meant serious delays in sowing crops and loss of harvests because of flash floods. Furthermore, illnesses and epidemics were believed to originate in the bad air ("miasmas") of swamps and bogs. A common perception was that wetlands were "awful reservoirs of stagnated water, which poisons the circumambient air for many miles around about and sickens and frequently destroys many of the inhabitants" (1).

In Europe, land drainage schemes were initiated as early as the thirteenth century to dry out wetlands and improve their agricultural productivity. Major drainage schemes in Britain in the seventeenth century converted thousands of square kilometers of fens and marsh into, what is today, some of the most productive farmland in the country (2). Similar schemes were implemented across Europe and elsewhere in the world. During the twentieth century, technological advances coupled with greater intensity of agriculture, as well as urban and industrial development, increased the pace of change. Alteration in the water balance and the quality of water entering wetlands can result in desiccation and degradation of wetland habitat. Although very difficult to quantify, it is estimated that globally more than half of all wetlands have been destroyed as a consequence of human activities (3).

In recent years, it has been realized that wetlands are valuable ecosystems. It is now understood that they maintain environmental quality, sustain livelihoods, support immense biodiversity, and have important recreational and aesthetic qualities.

Wetlands continue to play an important role in sustaining the livelihoods of many millions of people, particularly in Africa and Asia. For example, the valley bottom wetlands in Africa are of major importance in agricultural systems. Saturated during the wet season, they remain wet into the dry season effectively extending the growing season. As a result, these wetlands act as a key resource for cultivators and pastoralists by providing a source of arable land or grazing throughout the dry season and during droughts. Similarly, in Bangladesh, many thousands of cattle that graze on floodplains during the dry season are watered from perennially flooded areas known locally as *beels*.

Insight into their value has led to a radical change in attitude toward wetlands and has been the primary reason for a range of national and international conservation policies. Throughout Europe, the United States, and many other countries, considerable efforts are now being made to protect and even recreate wetlands (4).

WETLAND FUNCTIONS

Wetlands are among the most productive ecosystems in the world, characterized by high primary productivity and a multiplicity of habitats that enables them to support a large diversity of life. The wide range of natural phenomena that occur within wetlands makes them valuable in socioeconomic as well as ecological terms.

Wetland values arise from the ecological functions they perform. Ecological functions are the result of the physical, chemical, and biological processes that occur through the interaction of the biotic (i.e., flora and fauna) and abiotic (i.e., soil and water) components that comprise an ecosystem. The exact nature of the functions is unique to any particular wetland and depends on site-specific factors, including climate, geology, topography, and the wetland's location within a catchment. However, it is now recognized that biodiversity per se, plays an important role in maintaining wetland processes (e.g., condensation and evapotranspiration) across a range of temporal and spatial scales (5).

Not all wetland functions are useful to people. Some are harmful. For example, provision of habitat for mosquitoes and other insects that transmit illnesses is a function of many wetlands that throughout history has resulted in incalculable human suffering. However, many wetland functions do enhance human well-being, and it is now widely accepted that, overall, the advantages of wetlands outweigh their disadvantages.

The *physical* benefits human populations derive from wetlands comprise "goods" and "services." Goods are the natural resources that occur within a wetland and can be exploited by people. Examples include plants and wildlife used for food, trees and reeds used for construction materials, and vegetation used for medicine. Services are the nonresource benefits that people gain from the interaction of wetland processes. Examples include hydrologic, water quality, and geomorphological functions, such as flood reduction, nutrient transformation, and sediment retention. A summary of some of the physical benefits that human societies gain from wetland functions is presented in Table 1.

The high productivity of wetland ecosystems is reflected in the high human populations that they can support. In Africa and Asia, hundreds of millions of people depend on wetlands for water, rice, fish, shrimp, and other food production. In many tropical wetlands, fisheries production is extremely high. The lower Mekong Basin supports an inland fishery whose annual yield is approximately 500,000 tons, and in Indonesia, half the commercial fish catch comprises species dependent on estuarine wetland systems (6). The freshwater and estuarine wetlands of the Nile, Congo, Rifiji, Tana, Niger, and Zambezi river catchments in Africa each sustain the livelihoods of many hundreds of thousands of people by providing fisheries; dry season grazing for cattle; water for domestic use; and agriculture and plants for traditional medicines, mat production, and firewood. Similarly the great riverine and floodplain systems of South and Central America support large numbers of subsistence and commercial fisheries. It is estimated that 75% of the commercial fish catch of the Amazon River depends on food chains originating in seasonally flooded forest wetlands (6).

Table 1. Examples of Wetland Functions that Benefit Human Populations

Wetland Function	Use by Society	Examples
• Storage of precipitation and runoff	• Water supply	• In Florida, U.S., approximately 5 million people, including many in the city of Miami are supplied with water from the wetlands of the Florida Everglades. • 10,000 people living in the vicinity of Laguna El Jocotal, a shallow floodplain lake in El Salvador, depend on the lake for their dry season water supply.
	• Flood protection	• A flood prevention value of about US$13,500 per hectare per year has been attributed to wetlands in the Charles River catchment in Massachusetts, U.S.
	• Agriculture	• In West Africa, the shallow groundwater and elevated soil moisture of about 5 million hectares of inland valley wetlands are used extensively in crop cultivation.
• Groundwater discharge	• Water supply	• 8% of the discharge of the Mzima springs in Tsavo National Park, Kenya, is diverted to supply the coastal city of Mombasa.
• Groundwater recharge	• Water supply	• The Hadejia–Nguru wetlands in Nigeria contribute to the recharge of an aquifer used for water supply by approximately 1 million people who live in the region. • Gravels underlying riverine forest wetlands adjacent to the Danube supply a large proportion of the water to the 1.6 million people of Vienna, Austria.
• Sediment retention	• Shoreline protection	• The mangroves of the Indus delta protect the coastline from wind and currents and shelter Port Qasim, Pakistan's second largest port.
• Nutrient transformation	• Wastewater treatment	• 12,000 hectares of wetland to the east of Calcutta, India, are used to treat approximately 68 million liters of sewage each day. In 10 days, it is transformed into "sweet water" suitable for irrigation.
	• Protection of water resources	• A riparian wetland system on the River Lambourn, UK, removed up to 85% of the nutrients in agricultural runoff before it reached the river.
• Biomass production (e.g., wood, grass, fish)	• Food	• Local people use the fruits, seeds, tubers, roots, and leaves of around 200 plants from the wetlands surrounding Lake Chilwa in Malawi. • The wetlands of the Kafue floodplain in Zambia support fisheries averaging 6,500 tons per year.
	• Building and construction materials	• In Matang Forest Reserve, Malaysia, 40,000 hectares of mangroves annually yield timber worth US$9 million.
• Maintenance of biodiversity	• Medicine	• The St. Regis Mohawk tribe in the United States uses some 85 wetland plants for medicine, including *Acorus calamus* used for chills, fevers, and coughs and *Lobelia cardinalis* used as a blood purifier. • In Peninsular Malaysia, the sale of medicinal plants originating in wetlands can earn households up to US$ 80 per month.
• Chemical cycling	• Gas regulation	• Wetlands are important in the biogeochemical cycling of carbon dioxide, methane, and hydrogen sulfide. • Globally, wetland peat deposits occupy just 3% of the world's land area but store 16–24% of the planet's soil carbon pool.

Some wetlands play an important role in maintaining or improving water quality. The large number of chemical and biological processes that take place within wetlands (including nitrification, denitrification, ammonification, and volatilization) may convert heavy metals to insoluble forms and remove nutrients from the water. In Uganda, the Nakivubo papyrus swamp receives semitreated effluent from the sewage works of the capital city, Kampala. During the passage of the effluent through the wetland, the papyrus vegetation absorbs nutrients and the concentration of pollutants is reduced before the water enters Lake Victoria, the principal water source for the city (7). In Malaysia, the North Selangor peat swamp protects the quality of water used for water supply and irrigation by promoting sedimentation, nutrient uptake, toxic metal removal, and pH buffering (8).

People also gain *nonphysical* benefits from wetland functions. These are associated with spiritual enrichment, cognitive development, and aesthetic experience. For example, the culture of the Lozi people in western Zambia has developed as a consequence of the hydrologic dynamics of the Barotse floodplain. Each year, cultural ceremonies are conducted when the people move out of the floodplain during periods of high water and back again when water levels fall. Such ceremonies are important in enhancing the social cohesiveness of communities. Elsewhere, wetlands such as the Florida Everglades in the United States, the Okavango Delta in Botswana and the Caroni Swamp in Trinidad are major tourist attractions, not only enriching the lives of visitors, but also generating income for others.

Wetlands support a rich diversity of flora and fauna. Some animals live permanently in wetlands (e.g., the manatee, swamp tortoise, and lechwe antelope); others use them as refuges and migrate to them when conditions of drought or food scarcity exist elsewhere or at particular stages of their life cycle. Wetlands are especially important for migratory birds. Of the 300 species of birds that migrate between Europe and Africa each year, a large number depend on wetland habitats to support them enroute. The Patanal wetland in South America provides habitat for some 865 species of birds, and the Okavango Delta and Kafue Flats in Africa each support 400 species. It is estimated that the Inner Delta of the Niger Delta hosts up to 1.5 million birds each year (9). The beauty and naturalness of what are perceived to be pristine and increasingly rare ecosystems, and the variety of wildlife that they support, are important and of value to increasing numbers of people worldwide.

Hence, wetlands bring a wide variety of both tangible and intangible benefits to a huge number of people. The way in which they do so is complex, multifunctional, and very often not obvious. Many groups of people benefit from wetlands, and very often the benefits are experienced far beyond the boundaries of the wetland itself. Many wetland contributions to human welfare accrue without people being aware of them. As a consequence, wetland functions have in the past often gone unrecognized in development and resource planning and management.

WETLAND VALUES

The values attributed to wetlands depend largely on social perceptions of the uses and the benefits to be gained from them. The view that wetlands are wastelands arose in countries where agricultural productivity is not usually limited by water scarcity, at a time when rapidly rising populations required increased food production and there was very little understanding of the benefits provided by wetlands.

Today, in the developed countries of the world, the situation is very different. The ecosystems that people most depend on are those in which the functions, composition, and structure are far removed from natural systems. High levels of social and economic security have been obtained through application of complex technologies and considerable alteration of natural ecosystems, including wetlands.

These societies are now sufficiently prosperous that they can contemplate a large number of options in the way that they manage ecosystems, and paradoxically, many of them are now placing very high value on the aesthetic quality and biodiversity functions of their remaining wetlands.

In contrast, many millions of people in much of the developing world remain directly dependent on ecosystems that are in a much more natural state. For many individuals, wetland functions provide the most important contributions to their livelihoods and welfare by providing food supplies, medicines, income, employment, and cultural integrity. Such communities often have limited alternative livelihood options, and this makes them particularly vulnerable to changes in the condition of the natural environment on which they depend.

As a consequence of these different perspectives, wetland values were, until recently, very different. In the developed world, the primary focus was on conservation of wetland “wilderness” for recreation and to maintain wildlife habitat, especially for waterfowl. In much of the developing world, the emphasis was on the utility of wetlands, in particular their role in food production and livelihood support. However, there is now a widespread convergence of views brought about by the realization of two factors. First, that although the natural resources they provide may not be as important in the developed world, wetlands, nevertheless, support or protect a wide range of economic activities: for example, through their role in flood reduction and water purification; second, that social and economic factors are the main reasons for wetland loss; poverty and environmental degradation are inextricably intertwined. Consequently, sustainable development is a prerequisite for successful conservation.

This convergence of views is reflected in the widely promulgated “wise use” policy of the Convention on Wetlands of International Importance especially as Waterfowl Habitat—commonly referred to as the Ramsar Convention. This states that “the wise use of wetlands is their sustainable utilisation for the benefit of humankind in a way compatible with the maintenance of the natural properties of the ecosystem” (10). Explicit in the guidelines of the Ramsar Convention is the need to identify the values that local people, who use wetlands directly, can bring to all aspects of wetland wise use.

The site-specific nature and diversity of wetland functions makes quantification of wetland values extremely difficult. However, the wise use of wetlands requires a way of comparing the various benefits and consequences of development options on wetland functions. Traditionally, valuation techniques were strictly financial. However, it is impossible to assess all the issues relevant to wetlands using money as the sole determinant in the assessment process. Consequently, in recent years, economists have developed a number of environmental valuation techniques that enable assessment of nonmonetary impacts (11). Such methods, for example, multicriteria analysis, enable assessment of nonmonetary values and are being increasingly used to evaluate alternatives in wetland management.

Management of environmental resources, including wetlands, will be a key human endeavor in the twenty-first century. The biotic impoverishment and disruption of wetland functions caused by past decisions and continuing mismanagement severely constrain the options for future management. If the benefits wetlands bring are to be used more wisely in the future, there is need for much greater insight into the dynamics of wetland functions and the links between wetlands and human society. Only through such understanding will the true value of wetlands be determined.

BIBLIOGRAPHY

1. Parkinson, R. (1811). *General View of the Agriculture of the County of Huntingdon*. Phillips, London. p. 351. Cited in: Sheail, J. The land-use history of the Huntingdonshire Fenland, with special reference to Holme Fen and Woodwalton Fen National Nature Reserves, Institute of Terrestrial Ecology, Natural Environment Research Council, 1976, p. 2.
2. Cook, H. and Williamson, T. (1999). *Water Management in the English Landscape: Field Marsh and Meadow*. Edinburgh University Press, Edinburgh.
3. Cosgrove, W.J. and Rijsberman, F.R. (2000). *World Water Vision: Making Water Everybody's Business*. Earthscan Publications Ltd, London.
4. Klötzli, F. and Grootjans, A.P. (2001). Restoration of natural and semi-natural wetland systems in Central Europe: progress and predictability of developments. *Restoration Ecol.* **9**: 209–219.
5. Lawton, J.H. (1999). What to conserve—species or ecosystems? In: *Ecosystem Management: Questions for Science and Society*. Maltby et al. (Eds.). University of London, London, pp. 55–62.
6. Finlayson, M. and Moser, M. (Eds.). (1991). *Wetlands*. International Waterfowl and Wetlands Research Bureau, Facts on File, Oxford, UK.
7. Kansiime, F. and Nalubega, M. (1999). *Wastewater Treatment by a Natural Wetland: the Nakivubo Swamp, Uganda*. A.A. Balkema, Rotterdam.
8. Khan, N. (1995). Protection of North Selangor Peat Swamp Forest, Malaysia. *Parks* **5**: 24–31.
9. Roggeri, H. (1995). *Tropical Freshwater Wetlands: A Guide to Current Knowledge and Sustainable Management*. Kluwer Academic, Dordrecht, The Netherlands.
10. Ramsar, Convention Bureau. (2000). *Ramsar Handbooks for the Wise Use of Wetlands*. Ramsar Convention Bureau, Gland.
11. Barbier, E.B., Acreman, M.C., and Knowler, D. (1997). *Economic Valuation of Wetlands: A Guide for Policy Makers and Planners*. Ramsar Convention Bureau, Gland.

WETLANDS OVERVIEW

U.S. Environmental Protection
Agency—Office of Water, Office
of Wetlands, Oceans
and Watersheds

WHAT IS A WETLAND?

Although wetlands are often wet, a wetland might not be wet year-round. In fact, some of the most important wetlands are only seasonally wet. Wetlands are the link between the land and the water. They are transition zones where the flow of water, the cycling of nutrients, and the energy of the sun meet to produce a unique ecosystem characterized by hydrology, soils, and vegetation—making these areas very important features of a watershed. Using a watershed-based approach to wetland protection ensures that the whole system, including land, air, and water resources, is protected.

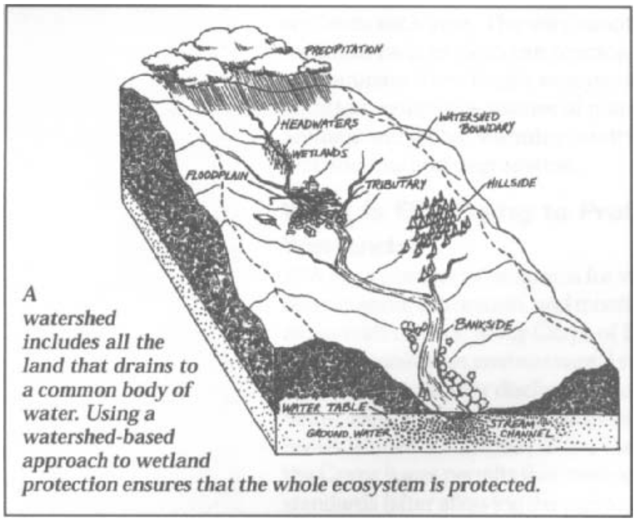

A watershed includes all the land that drains to a common body of water. Using a watershed-based approach to wetland protection ensures that the whole ecosystem is protected.

Wetlands found in the United States fall into four general categories—marshes, swamps, bogs, and fens. Marshes are wetlands dominated by soft-stemmed vegetation, while swamps have mostly woody plants. Bogs are freshwater wetlands, often formed in old glacial lakes, characterized by spongy peat deposits, evergreen trees and shrubs, and a floor covered by a thick carpet of sphagnum moss. Fens are freshwater peat-forming wetlands covered mostly by grasses, sedges, reeds, and wildflowers.

GOOD NEWS

Often called "nurseries of life," wetlands provide habitat for thousands of species of both aquatic and terrestrial plants and animals. These nurseries support the critical

---

Two-thirds of the 10 million to 12 million waterfowl of the continental United States reproduce in the prairie pothole wetlands of the Midwest, and in the winter millions of ducks like these can be found in the wetlands of the south-central United States

developmental stages for many plants and animals. Although wetlands are best known for being home to water lilies, turtles, frogs, snakes, alligators, and crocodiles, they also provide important habitat for waterfowl, fish, and mammals. Migrating birds use wetlands to rest and feed during their cross-continental journeys and as nesting sites when they are at home. As a result, wetland loss has a serious impact on these species. Habitat degradation since the 1970s has been a leading cause of species extinction.

> Living systems cleanse water and make it fit, among other things, for human consumption.
>
> *Elliot A. Norse*,
> in R.J. Hoage, ed.,
> Animal Extinctions, *1985*

> The nation behaves well if it treats the natural resources as assets which it must turn over to the next generation increased, and not impaired, in value.
>
> Theodore Roosevelt, 1907

Wetlands do more than provide habitat for plants and animals in the watershed. When rivers overflow, wetlands help to absorb and slow floodwaters. This ability to control floods can significantly prevent property damage and loss and can even save lives. Wetlands also absorb excess nutrients, sediment, and other pollutants before they reach rivers, lakes, and other waterbodies. They are great spots for fishing, canoeing, hiking, and bird-watching, and they make wonderful outdoor classrooms for people of all ages.

BAD NEWS

Despite all the benefits provided by wetlands, the United States loses about 60,000 acres of wetlands each year. The very runoff that wetlands help to clean can overload and contaminate these fragile ecosystems. In addition, nonnative species of plants and animals and global warming contribute to wetland loss and degradation.

This forested wetland on the Chincoteague National Wildlife Refuge on Virginia's Eastern Shore is part of the Atlantic flyway, where shorebirds and waterfowl rest before they migrate south for the winter

A freshwater pool at Assateague National Seashore in Virginia

WHAT IS EPA DOING TO PROTECT WETLANDS?

EPA has a number of programs for wetland conservation, restoration, and monitoring. EPA, along with the U.S. Army Corps of Engineers (Corps), establishes environmental standards for reviewing permits for discharges that affect wetlands, such as residential development, roads, and levees. Under Section 404 of the Clean Water Act, the Corps issues permits that meet environmental standards (after allowing the public to comment).

WORKING TOGETHER TO PROTECT AND RESTORE WETLANDS

In addition to providing regulatory protection for wetlands, EPA works in partnership with states, tribes, and local governments, the private sector, and citizen organizations to monitor, protect, and restore these valuable habitats. EPA is helping states and tribes incorporate wetland monitoring, protection, and restoration into their watershed plans. EPA is also working to develop national guidance on wetland restoration, as well as constructed wetlands used to treat storm water and sewage. Nationally, EPA's Five-Star Restoration Program provides grants and promotes information exchange through community-based education and restoration projects.

Wetland habitat along this Idaho riparian corridor provides food and shelter for many different wildlife species

EPA works with a variety of other federal agencies to protect and restore wetlands, including the U.S. Fish and Wildlife Service, the U.S. Department of Agriculture, and the National Marine Fisheries Service. EPA is working with these agencies and others to achieve a net increase of 100,000 acres of wetlands each year by 2005. EPA also partners with private interests and public organizations like the Association of State Wetland Managers, the National Association of Counties, the Izaak Walton League, local watershed associations, schools, and universities to advance conservation and restoration programs.

If bottomland hardwood swamps are protected, Bald Cypress trees like these can grow for more than 2000 years.

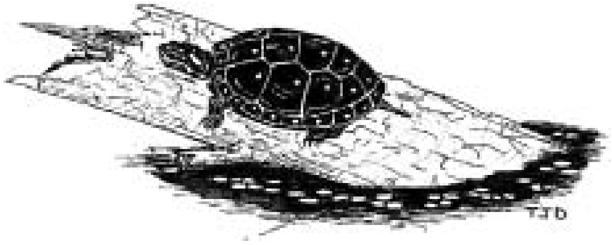

Spotted turtle

HOW CAN I HELP?

You can do many things to help protect wetlands in your area. First, identify your watershed and find the wetlands in your neighborhood. Learn more about them and share what you learn with someone you know! Encourage neighbors, developers, and state and local governments to protect the functions and values of wetlands in your watershed.

To prevent wetland loss or degradation, follow these simple guidelines:

- Invest in wetlands by buying duck stamps. Proceeds from these $15 migratory bird hunting stamps support wetland acquisition and restoration. The stamps are available on-line at the U.S. Fish and Wildlife Service's web site (www.fws.gov) or at your local post office.
- Rather than draining or filling wetlands, find more compatible uses that would not damage the wetlands, such as waterfowl and wildlife habitat.
- When developing your landscaping plan, keep wetlands in mind. Plant native grasses or forested buffer strips along wetlands on your property to protect water quality.

- "Get into" your wetland by participating in a volunteer monitoring program.
- Plan to avoid wetlands when developing or improving a site. Get technical assistance from your state environmental agency before you alter a wetland.
- Maintain wetlands and adjacent buffer strips as open space.
- Support your local watershed association.
- Plan a wetland program or invite a wetland expert to speak at your school, club, youth group, or professional organization.
- Build a wetland in your backyard. Learn how by visiting the U.S. Department of Agriculture's web site at www.nrcs.usda.gov/feature/backyard.

Wetlands can be found in every county and climatic zone in the United States.

CLASSIFICATION OF WETLANDS AND DEEPWATER HABITATS OF THE UNITED STATES

Lewis M. Cowardin
Virginia Carter
Francis C. Golet
Edward T. Laroe
U.S. Department of the Interior

USE OF THE CLASSIFICATION SYSTEM

This System was designed for use over an extremely wide geographic area and for use by individuals and organizations with varied interests and objectives. The classification employs 5 System names, 8 Subsystem names, 11 Class names, 28 Subclass names, and an unspecified number of Dominance Types. It is, of necessity, a complex System when viewed in its entirety, but use of the System for a specific purpose at a local site should be simple and straightforward. Artificial keys to the Systems and Classes (Appendix E) are furnished to aid the user of the classification, but reference to detailed definitions in the text is also required. The purpose of this section is to illustrate how the System should be used and some of the potential pitfalls that could lead to its misuse.

Before attempting to apply the System, the user should consider four important points:

(1) Information about the area to be classified must be available before the System can be applied. This information may be in the form of historical data, aerial photographs, brief on-site inspection, or detailed and intensive studies. The System is designed for use at varying degrees of detail. There are few areas for which sufficient information is available to allow the most detailed application of the System. If the level of detail provided by the data is not sufficient for the needs of the user, additional data gathering is mandatory.

(2) Below the level of Class, the System is open-ended and incomplete. We give only examples of the vast number of Dominance Types that occur. The user may identify additional Dominance Types and determine where these fit into the classification hierarchy. It is also probable that as the System is used the need for additional Subclasses will become apparent.

(3) One of the main purposes of the new classification is to ensure uniformity throughout the United States. It is important that the user pay particular attention to the definitions in the classification. Any attempt at modification of these definitions will lead to lack of uniformity in application.

(4) One of the principal uses of the classification system will be the inventory and mapping of wetlands and deepwater habitats. A classification used in the mapping is scale-specific, both for the minimum size of units mapped and for the degree of detail attainable. It is necessary for the user to develop a specific set of mapping conventions for each application and to demonstrate their relationship to the generalized classification described here. For example, there are a number of possible mapping conventions for a small wetland basin 50 m (164 feet) in diameter with concentric rings of vegetation about the deepest zone. At a scale of 1:500 each zone may be classified and mapped; at 1:20,000 it might be necessary to map the entire basin as one zone and ignore the peripheral bands; and at 1:100,000 the entire wetland basin may be smaller than the smallest mappable unit, and such a small-scale map is seldom adequate for a detailed inventory and must be supplemented by information gathered by sampling. In other areas, it may be necessary to develop mapping conventions for taxa that cannot be easily recognized; for instance, Aquatic Beds in turbid waters may have to be mapped simply as Unconsolidated Bottom.

HIERARCHICAL LEVELS AND MODIFIERS

We have designed the various levels of the system for specific purposes, and the relative importance of each will vary among users. The Systems and Subsystems are most important in applications involving large regions or the entire country. They serve to organize the Classes into meaningful assemblages of information for data storage and retrieval.

The Classes and Subclasses are the most important part of the system for many users and are basic to wetland mapping. Most Classes should be easily recognizable by users in a wide variety of disciplines. However, the Class designations apply to average conditions over a period of years, and since many wetlands are dynamic and subject to rapid changes in appearance, the proper classification of a wetland will frequently require data that span a period of years and several seasons in each of those years.

The Dominance Type is most important to users interested in detailed regional studies. It may be necessary to identify Dominance Types in order to determine which modifying terms are appropriate, because plants and animals present in an area tend to reflect environmental conditions over a period of time. Water regime can be determined from long-term hydrologic studies where these are available. The more common procedure will be to estimate this characteristic from the Dominance Types. Several studies have related water regimes to the presence and distribution of plants or animals (e.g., Stephenson and Stephenson 1972; Stewart and Kantrud 1972; Chapman 1974).

Similarly, we do not intend that salinity measurements be made for all wetlands except where these data are required; often plant species or associations can be used to indicate broad salinity classes. Lists of halophytes have been prepared for both coastal and inland areas (e.g., Duncan 1974; MacDonald and Barbour 1974; Ungar 1974), and a number of floristic and ecological studies have described plants that are indicators of salinity (e.g., Penfound and Hathaway 1938; Moyle 1945; Kurz and Wagner 1957; Dillon 1966; Anderson et al. 1968; Chabreck 1972; Stewart and Kantrud 1972; Ungar 1974).

In areas where the Dominance Types to be expected under different water regimes and types of water chemistry conditions have not been identified, detailed regional studies will be required before the classification can be applied in detail. In areas where detailed soil maps are available, it is also possible to infer water regime and water chemistry from soil series (U.S. Soil Conservation Service, Soil Survey Staff 1975).

Some of the Modifiers are an integral part of this system and their use is essential; others are used only for detailed applications or for special cases. Modifiers are never used with Systems and Subsystems; however, at least one Water Regime Modifier, one Water Chemistry Modifier, and one Soil Modifier must be used at all lower levels in the hierarchy. Use of the Modifiers listed under mixosaline and mixohaline (Table 1) is optional but these finer categories should be used whenever supporting data are available. The user is urged not to rely on single observations of water regime or water chemistry. Such measurements give misleading results in all but the most stable wetlands. If a more detailed Soil Modifier, such as soil order or suborder (U.S. Soil Conservation Service, Soil Survey Staff 1975) can be obtained, it should be used in place of the Modifiers, mineral and organic. Special Modifiers are used where appropriate.

RELATIONSHIP TO OTHER WETLAND CLASSIFICATIONS

There are numerous wetland classifications in use in the United States. Here we relate this system to three published classifications that have gained widespread acceptance. It is not possible to equate these systems directly for several reasons: (1) the criteria selected for establishing categories differ; (2) some of the classifications are not applied consistently in different parts of the country; and (3) the elements classified are not the same in various classifications.

The most widely used classification system in the United States is that of Martin et al. (1953) which was republished in U.S. Fish and Wildlife Service *Circular 39* (Shaw and Fredine 1956). The wetland types are based on criteria such as water depth and permanence, water chemistry, life form of vegetation, and dominant plant species. In Table 4 we compare some of the major components of our system with the type descriptions listed in *Circular 39*.

In response to the need for more detailed wetland classification in the glaciated Northeast, Golet and Larson (1974) refined the freshwater wetland types of *Circular 39* by writing more detailed descriptions and subdividing classes on the basis of finer differences in plant life forms. Golet and Larson's classes are roughly equivalent to Types 1–8 of *Circular 39*, except that they restrict Type 1 to river floodplains. The Golet and Larson system does not recognize the coastal (tidal) fresh wetlands of *Circular 39* (Types 12–14) as a separate category, but classifies these areas in the same manner as nontidal wetlands. In addition to devising 24 subclasses, they also created 5 size categories, 6 site types giving a wetland's hydrologic and topographic location; 8 cover types (modified from Stewart and Kantrud 1971) expressing the distribution and relative proportions of cover and water; 3 vegetative interspersion types; and 6 surrounding habitat types. Since this system is based on the classes of Martin et al. (1953), Table 4 may also be used to compare the Golet and Larson system with the one described here. Although our system does not include size categories and site types, this information will be available from the results of the new inventory of wetlands and deepwater habitats of the United States.

Stewart and Kantrud (1971) devised a new classification system to better serve the needs of researchers and wetland managers in the glaciated prairies. Their system recognizes seven classes of wetlands which are distinguished by the vegetational zone occupying the central or deepest part and covering 5% or more of the wetland basin. The classes thus reflect the wetland's water regime; for example, temporary ponds (Class II) are those where the wet-meadow zone occupies the deepest part of the wetland. Six possible subclasses were created, based on differences in plant species composition that are correlated with variations in average salinity of surface water. The third component of classification in their system is the cover type, which represents differences in the spatial relation of emergent cover to open water or exposed bottom soil. The zones of Stewart and Kantrud's system are readily related to our water regime modifiers (Table 1), and the subclasses are roughly equivalent to our Water Chemistry Modifiers (Fig. 8).

Wetlands represent only one type of land and the classification of this part separate from the rest is done for practical rather than for ecological reasons (Cowardin 1978). Recently there has been a flurry of interest in a holistic approach to land classification (in Land Classification Series, *Journal of Forestry*, vol. 46, no. 10). A number of classifications have been developed (e.g., Radford 1978) or are under development (e.g., Driscoll et al. 1978). Parts

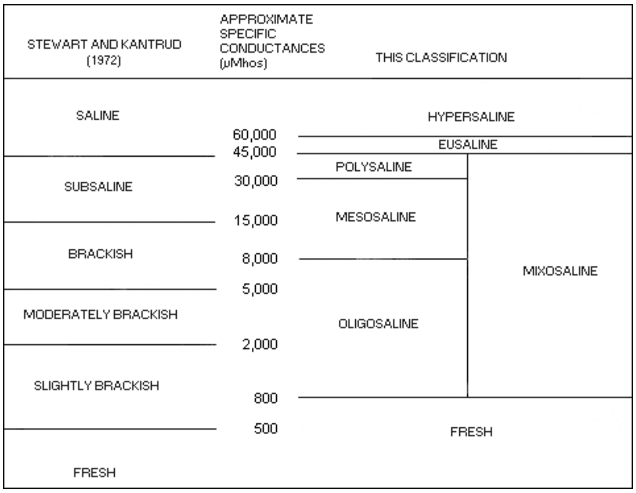

Figure 1. Comparison of the water chemistry subclasses of Stewart and Kantrud (1972) with Water Chemistry Modifiers used in the present classification system.

Table 1. Comparison of the Zones of Stewart and Kantrud's (1971) Classification with the Water Regime Modifiers used in the Present Classification System

Zone	Water Regime Modifier
Wetland-low-prairie	Non-wetland by our definition
Wet meadow	Temporarily flooded
Shallow marsh	Seasonally flooded
Deep marsh	Semipermanently flooded Intermittently exposed
Intermittent-alkali	Intermittently flooded (with eusaline or hypersaline water)
Permanent-open-water	Permanently flooded (with mixohaline water)
Fen (alkaline bog)	Saturated

of this wetland classification can be incorporated into broader hierarchical land classifications.

A classification system is most easily learned through use. To illustrate the application of this system, we have classified a representative group of wetlands and deepwater habitats of the United States (Plates 1–86).

URBAN RUNOFF

Miklas Scholz
The University of Edinburgh
Edinburgh, United Kingdom

INTRODUCTION

Urban runoff comprises the stormwater runoff from urban areas such as roads and roofs. Road and car park runoff is often more contaminated with heavy metals and organic matter than roof runoff. Runoff is usually collected in gully pots that require regular cleaning. Stormwater pipes transfer urban runoff either directly to watercourses or to sustainable urban drainage systems (also called best management practice in the United States) such as ponds or wetlands. However, urban runoff is frequently treated only by silt traps that need to be cleaned regularly. This chapter highlights the main areas in urban runoff research and contain general background information supported by case studies.

ROAD RUNOFF DRAINAGE

Background to Road Runoff

Conventional stormwater systems are designed to dispose of rainfall runoff as quickly as possible. This results in 'end of pipe' solutions that often involve large interceptor and relief sewers, huge storage tanks at downstream locations, and centralized wastewater treatment (1–3).

In contrast, sustainable urban drainage systems such as combined attenuation pond and infiltration basin systems (4) can be applied as cost-effective local 'source control' drainage solutions, for example, for collecting road runoff. It is often possible to divert all road runoff for infiltration or storage and subsequent recycling. Runoff from roads is a major contributor to the quantity of surface water requiring disposal, so this is a particularly beneficial approach when suitable ground conditions prevail (1). Furthermore, infiltration of road runoff can reduce the concentration of diffuse pollutants such as oil, leaves, metals, and feces, thereby improving the water quality of surface water runoff (3,4).

Case Study: Design and Operation of Urban Stormwater Ponds Treating Road Runoff

Combined wetlands and infiltration basins are cost-effective 'end of pipe' drainage solutions that can be used for local source control. The aims of this case study were to assess the constraints of the design and operation of these systems, the influence of wetland plants on infiltration rates, and the water treatment potential. Road runoff was first stored and treated in a constructed wetland before it overflowed into parallel infiltration basins; one was planted and the other was unplanted (Fig. 1). British Building Research Establishment and Construction Industry Research and Information Association and German Association for Water, Wastewater and Waste design guidelines were unacceptable because they did not take vertical soil property variations fully into account. Wetland plants in one infiltration basin had no significant influence on drainage properties. The water quality of both infiltration basins was not acceptable for recycling directly after the system setup (5).

ROOF RUNOFF DRAINAGE

Background to Roof Runoff

Combined wet and dry ponds as cost-effective 'end of pipe' drainage solutions can be used for local source control, for example, diversion or collection of roof runoff. It is often possible to divert all roof runoff for infiltration or storage and subsequent recycling. Runoff from roofs is a major contributor to the quantity of surface water requiring

Figure 1. Experimental ponds treating road runoff in Scotland.

Figure 2. Experimental infiltration pond for roof runoff in England.

disposal, so this is a particularly beneficial approach when suitable ground conditions prevail (1). Furthermore, roof runoff is usually cleaner than road runoff. However, diffuse pollution from the atmosphere and degrading building materials can have a significant impact on the water quality of any surface water runoff (3).

Case Study: Design and Operation of Stormwater Ponds for Roof Runoff Treatment

The purpose of this case study was to optimize design, operation and maintenance guidelines and to assess the water treatment potential of a stormwater pond system after 15 months of operation. The system was based on a combined silt trap, attenuation pond (wet pond) and dry pond construction used for drainage of roof water runoff from a single domestic property (Fig. 2). British Building Research Establishment and Construction Industry Research and Information Association, and German Association for Water, Wastewater and Waste design guidelines were tested. These design guidelines failed because they do not consider local hydrologic and soil conditions. The infiltration function for the dry pond is logarithmic and depends on the season. Furthermore, algae control techniques were successfully applied, and water treatment of rainwater runoff from roofs was largely unnecessary. However, seasonal and diurnal variations in biochemical oxygen demand, dissolved oxygen and pH were recorded (4).

URBAN RUNOFF AND GULLY POT EFFLUENT TREATMENT

Background to Gully Pot Effluent

Gully pots can be viewed as simple physical, chemical and biological reactors. They are particularly effective in retaining suspended solids and settleable solids (6). Currently, gully pot effluent is extracted once or twice per annum from road drains and transported (often over long distances) for treatment at sewage works (1). A more sustainable solution would be to treat gully pot effluent locally in potentially sustainable constructed wetlands, reducing transport and treatment costs. Furthermore, gully pot effluent treated by constructed wetlands can be recycled for cleansing gully pots and washing wastewater tankers. Heavy metals within urban runoff originate from fuel additives, car body corrosion, and tire and brake wear. Common metal pollutants from road vehicles include lead, zinc, copper, chromium, nickel and cadmium. Typical concentration ranges for most heavy metals are between 0.05 and 2 mg/L (7).

Case Study: Design and Operation of Constructed Wetlands Applied to Gully Pot Effluent Contaminated with Heavy Metals

The aim was to assess the treatment efficiencies for gully pot effluent of experimental vertical-flow constructed wetland filters containing *Phragmites australis* (Cav.) Trin. ex Steud. (common reed) and filter media of different adsorption capacities. Six out of 12 filters received inflow water spiked with heavy metals. For one year, hydrated copper nitrate and hydrated nickel nitrate were added to sieved gully pot effluent to simulate contaminated primary treated stormwater runoff. The inflow concentrations for dissolved copper, nickel and nitrate were approximately 1.0, 1.0 and 3.7 mg/L, respectively. For those six filters receiving metals, an obvious breakthrough of dissolved nickel was recorded after road gritting and salting during winter. Sodium chloride was responsible for leaching of dissolved nickel (Fig. 3). Reductions of copper, nickel, biochemical oxygen demand, and suspended solids were frequently insufficient compared to international secondary wastewater treatment standards. An analysis of variance indicated that all filters were similar in their treatment performance (8).

URBAN RUNOFF DISCHARGING INTO URBAN WATERCOURSES

Background to Urban Runoff Discharge to Waters

Stormwater runoff is usually transferred within a combined sewer (rainwater and sewage) or separate stormwater pipeline. Depending on the degree of pollution,

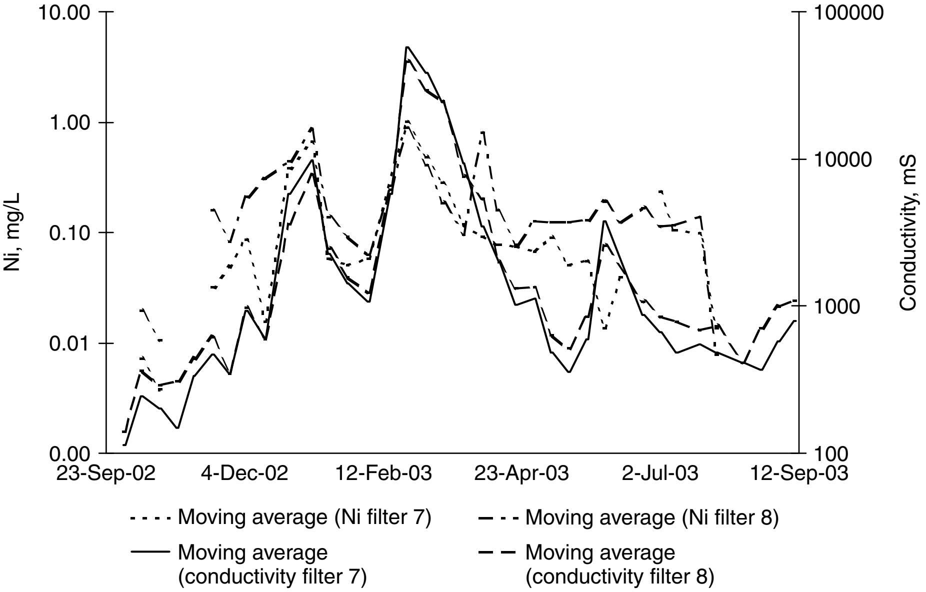

Figure 3. Relationship between the breakthrough of nickel (Ni) and conductivity within the outflows of two experimental constructed wetlands.

it may require further conventional treatment at the wastewater treatment plant. Alternatively, stormwater runoff may be treated by more sustainable technologies, including stormwater ponds, or simply by disposing it via ground infiltration or drainage into a local watercourse (1).

Common preliminary treatment steps for stormwater that is to be discharged into a local watercourse are extended storage within the stormwater pipeline and transfer through a silt trap (also referred to as a sediment trap). The silt trap is often the only active preliminary treatment of surface water runoff despite the fact that stormwater is frequently contaminated with heavy metals and hydrocarbons (9). Heavy metals within road runoff are from fuel additives, car body corrosion, and tire and brake wear (1).

The dimensioning of silt traps is based predominantly on the settlement properties of the particles and the maximum flow-through velocity. As a general rule, the outflow should be located on the same axis as the inflow, so that continuous flow ensures particle settlement over the full length of the silt trap (10). The structure should be small and simple to keep capital costs low.

Various international and British guidelines exist for designing outflow structures. In practice, designs vary considerably. However, silt traps are usually longer than wider by a factor of up to five. Silt trap depths vary considerably but are usually about 40 cm. However, often the lack of maintenance of outflow structures such as silt traps leads to problems years after the structure has been commissioned. Accumulated sediment has to be removed from silt traps to retain their original design performance and to avoid resuspension of accumulated pollutants during storms, causing pollution in receiving waterbodies (11). Watercourses in urban (built-up) areas that receive surface water runoff may be subject to pollution, particularly, if silt traps, for example, are insufficiently maintained (e.g., low frequency of sediment removal). Furthermore, the additional water load may contribute to local flooding (11).

Case Study: Urban Runoff Quality Associated with a Full Silt Trap Discharging into an Urban Watercourse

The aim was to assess the influence of a full silt trap at the end of a stormwater drainage pipe on the water quality of stormwater discharged into a semi-natural urban watercourse. For approximately eleven weeks, the water qualities of the preliminarily treated stormwater and of the receiving watercourse (Braid Burn) were studied. The mean outflow concentrations of suspended solids were 2.0 and 34.1 mg/L during dry and wet weather conditions, respectively. Suspended solids concentrations of up to 141.6 mg/L were recorded during storms. Suspended solids values for treated stormwater were often too high compared to international secondary wastewater treatment standards of approximately 30 mg/L. Pollutants, including heavy metals (e.g., zinc, copper, and nickel) accumulated in the silt trap (Fig. 4). However, high outflow velocities during heavy rainfalls did not result in clearly defined sediment layers due to sediment resuspension. Metals did not accumulate in the receiving watercourse (12).

BIBLIOGRAPHY

1. Butler, D. and Davies, J.W. (2000). *Urban Drainage*. E & FN Spon, London, UK.
2. Abbott, C.L. and Comino-Mateos, L. (2001) *In situ* performance monitoring of an infiltration drainage system and field testing of current design procedures. *J. Chartered Inst. Water Environ. Manage.* **15**(3): 198–202.
3. Ellis, J.B., D'Arcy, B.J., and Chatfield, P.R. (2002). Sustainable urban drainage systems and catchment planning. *J. Chartered Inst. Water Environ. Manage.* **16**(4): 286–291.
4. Scholz, M. (2003). Case study: design, operation, maintenance and water quality management of sustainable stormwater ponds for roof runoff. *Bioresource Technol.* **95**(3): 269–279.
5. Haf M., Götting, S., Scholz, M. and Heal, K. (2004). Case study: design and operation of sustainable urban stormwater ponds treating road runoff during the system set-up phase. *Proc. Br. Hydrological Soc. Int. Conf. Hydrol.: Sci. Pract. 21st*

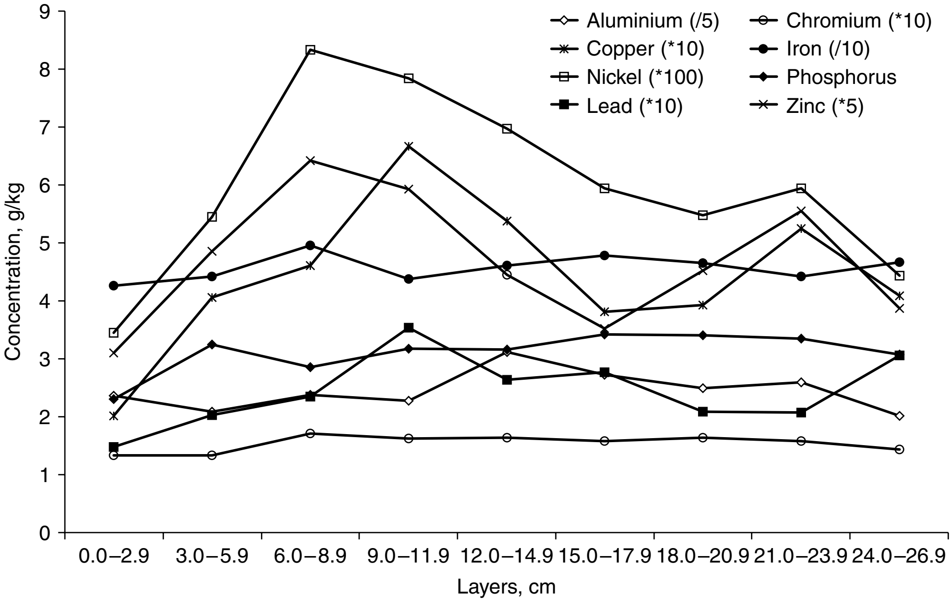

Figure 4. Metal concentrations in different sediment layers of a silt trap in Scotland.

Century, Imperial College, London, UK, July 12–16 2004, Vol. 2, pp. 280–286.

6. Bulc, T. and Slak, A.S. (2003). Performance of constructed wetland for highway runoff treatment. *Water Sci. Technol.* **48**(2): 315–322.
7. Cooper, P.F., Job, G.D., Green, M.B., and Shutes, R.B.E. (1996). *Reed Beds and Constructed Wetlands for Wastewater Treatment*. WRC, Swindon, UK.
8. Scholz, M. (2004). Treatment of gully pot effluent containing nickel and copper with constructed wetlands in a cold climate. *J. Chem. Technol. Biotechnol.* **79**(2): 153–162.
9. Ciszewski, D. (1998). Channel processes as a factor controlling accumulation of heavy metals in river bottom sediments: Consequences for pollution monitoring (Upper Silesia, Poland). *Environ. Geol.* **36**(1–2): 45.
10. Schmitt, F., Milisic, V., Bertrand-Krajewski, J.L., Laplace, D., and Chebbo, G. (1999). Numerical modelling of bed load sediment traps in sewer systems by density currents. *Water Sci. Technol.* **39**(9): 153.
11. Pontier, H., Williams, J.B., and May, E. (2001). Metals in combined conventional and vegetated road runoff control systems. *Water Sci. Technol.* **44**(11–12): 607.
12. Scholz, M. and Zettel, S. (in press). Stormwater quality associated with a full silt trap discharging into an urban watercourse. *J. Chartered Inst. Water Environ. Manage.*

URBAN WATER STUDIES

Karen D. Brettschneider
Houston-Galveston Area Council
Houston, Texas

INTRODUCTION

The urban landscape contains many elements that can alter the water quality of a waterbody. Before development occurs, grass and other natural vegetation hold sediment in place. Streams flowing through undeveloped terrain have no major sources of contaminants, such as bacteria (except for wildlife), that affect contact recreation. There are also no major sources of toxins that could negatively impact humans and wildlife. Development (residential, commercial, industrial, and agricultural, for example) brings issues that have the potential to compromise the character of streams through point source pollution or nonpoint source pollution (Fig. 1). The best way to prevent or deal with these issues is to have cooperation among a broad group of stakeholders in each watershed, the area of land (and tributaries) that drains to a particular waterbody. These individuals must work together to develop sound watershed management to protect or to restore the health of the ecosystem of that watershed.

WHAT HAPPENS TO THE QUALITY OF WATER WHEN DEVELOPMENT BEGINS?

One of the first concerns for water quality during development involves construction sites. When the natural vegetation is removed and bare soil is exposed, rainwater easily collects sediment, along with any pollutant attached to that sediment, and drains eventually to a stream. Additional sediment added to the water column may pose impacts to humans, such as complications with sources of public water supply, as well as to aquatic organisms. Sediment reduces sunlight for aquatic photosynthesis, clogs fish gills, disturbs bottom-dwelling organisms, and overall reduces the aesthetic value of the stream.

DURING AND AFTER DEVELOPMENT—POINT SOURCE AND NONPOINT SOURCE POLLUTION

Development includes the establishment of industrial plants, wastewater treatment plants, and other entities that discharge treated water to an outfall that connects to a stream. State and/or federal laws and standards regulate the effluent (point source) of these plants, but there still lies the potential for overflows, illegal discharges, and accidental spills.

Although there have been major strides in controlling pollutants from point sources, nonpoint source pollution (NPS) remains one of the most complex and elusive problems to deal with in watershed management. NPS pollution stems from those constituents that flow overground from diffuse sources directly into area waterways or indirectly through stormwater infrastructure to outfalls that connect to rivers and streams. One important factor to remember is that water that flows from homes to wastewater treatment facilities and then to a specified waterbody is clean as it leaves the plant. Stormwater receives no treatment. Left uncontrolled, these pollutants often have a significant impact on the quality of the water.

Some potential NPS pollutants carried in rainwater runoff include the following:

- excessive fertilizer applied to lawns or agricultural fields
- excessive pesticides and herbicides applied to lawns or agricultural fields
- grass clippings left on streets or blown down storm drains
- motor oil or other vehicle fluids from leaking automobiles or from illegal dumping
- trash
- animal waste
- sewage from broken sewer lines
- sewage from failing septic systems

Some of the consequences of these components entering a waterway include the following:

- excessive nutrient levels that may lead to eutrophic conditions
- depressed dissolved oxygen levels that can be harmful to aquatic species
- harmful elements exposed to wildlife
- increased bacteria counts
- impaired aesthetics
- impaired recreational value

APPROACHING THE ISSUES—WATERSHED MANAGEMENT

When analyzing water quality for an urban area, a watershed approach is the most effective means of defining problem areas and addressing concerns. Studying the various aspects within the watershed and how they interact is a scientifically sound approach to preventing and to solving water quality problems. Some of the various layers of information easily available to use in conjunction with water quality data include topographic information, detailed drainage information, soil types and characteristics, groundwater/aquifer information, vegetation types, climate information, and geology. Land cover or land use data can be obtained from aerial photography or other digital imagery.

Figure 1. An urban stream with many potential sources of pollution.

After gathering this type of information, participants in the watershed management plan may locate individual potential pollution sources, such as treatment plant or stormwater outfall locations. They may also locate potential areas of concern, such as areas of concentrated industrial activity, areas with large numbers of septic tanks, or areas with concentrated animal feeding operations. The latter activities are commonly found in rural areas, but these operations in the watershed influence the character of the waterbody as well.

Gathering of individuals, or stakeholders, from agriculture, industry, local business, government, local neighborhoods, the environmental profession, as well as other interested parties is the next step in addressing issues on a watershed scale. Each of these individuals can lend support in various aspects of the activities within the watershed.

Monitoring

Due to the size of urban areas, there are typically a number of entities that measure water quality for specific purposes. Gathering this information and coordinating among different groups, including city offices, county offices, river authorities, and other environmental groups can lead to a wealth of information about water quality throughout a watershed. Many agencies may be willing to modify their programs to help collect data in certain areas that are not monitored by any other entity. This action is commonly referred to as coordinated monitoring.

Common parameters used to assess the overall health of a waterbody include the following:

Field Measurements

- dissolved oxygen
- pH
- water temperature
- specific conductance

- salinity in marine waters
- Secchi depth
- flow
- days since last significant rainfall
- observational information, including water appearance and weather (and as applicable, biological activity; pertinent observations related to water quality or stream uses, e.g., boating, swimming, fishing, irrigation pumps, etc.; watershed or in-stream activities, e.g., bridge construction; unusual odors; specific sample information; and missing parameters)

Conventional Chemical Parameters

- ammonia –N
- nitrate + nitrite –N
- total phosphorus
- orthophosphorus
- chlorophyll *a*
- pheophytin
- total dissolved solids
- total suspended solids
- sulfate
- chloride
- total hardness

Bacteria

- fecal coliform
- *E. Coli* or *Enterococcus*

Flow Measurement

- direct measurement or USGS gage data
- flow estimate

Additional parameters for total maximum daily load (TMDL) purposes, load calculations, modeling, trend analysis, etc. or other specific purposes not related to the 305(b) screening:

- alkalinity
- total Kjeldahl nitrogen
- total organic carbon
- BOD5
- COD
- Toxics (metals or organics in water) for aquatic life/habitat monitoring

Routine monitoring should not include ambient toxicity, toxics in sediment or fish tissue, 24-hour dissolved oxygen measurements, and monitoring to determine the extent of impairment.

In addition to collecting data on water chemistry, biological and habitat data can assist in more comprehensively understanding the health of a watershed.

DEFINING ISSUES

Approaching concerns at a watershed level is appropriate in most situations; however, there may be issues confined to only one part of the watershed. Based on available data, researchers may be able to pinpoint a problem to a certain area. For example, the main stem of a channel may be contaminated with bacteria from a certain point downstream of a confluence with a tributary. Detailed study of that tributary's subwatershed may lead to sources of the bacteria that may help resolve the overall issue. So, while taking a watershed approach to water quality, it sometimes may be beneficial to go a step further to the subwatershed level, especially when trying to mitigate a particular pollutant. Eliminating or simply reducing input from small tributaries is the first step in cleaning up the entire watershed.

PREVENTING THE PROBLEMS—BEST MANAGEMENT PRACTICES AND PUBLIC EDUCATION

An important component of addressing water quality problems associated with NPS is through a preventive approach with using best management practices (BMP) at construction sites and other facilities to prevent contaminated runoff. Another strong preventive is public education. Once people know they can have a positive impact on the waterbodies in their community and watershed, then communities are one step closer to establishing healthy ecosystems in urban areas.

CONCLUSION

Urban areas often contain an immense network of flowing waterbodies containing quality aquatic habitat and recreational opportunities such as boating, canoeing, fishing, and swimming. At the same time, the many waterbodies that run through highly urbanized or impacted landscapes face challenges that require the cooperation of various entities to produce sound watershed management. The maintenance of a network of stakeholders and monitoring entities that works toward this goal is essential. Preventive steps such as BMPs, public education, and outreach opportunities are also very important in creating and maintaining quality places in which to live.

SUBGLACIAL LAKE VOSTOK

Brent C. Christopher
John C. Priscu
Montana State University
Bozeman, Montana

When Captain Robert F. Scott first observed the McMurdo Dry Valleys in December 1903, he wrote, "We have seen no living thing, not even a moss or a lichen... it is certainly a valley of the dead." Eight years later, Scott's team reached the South Pole and his entry referencing conditions on the polar plateau read "Great God! This is an awful place." Scott was unaware that life surrounded him in the dry valleys, and he could not have realized that microbial life could exist miles beneath his feet in an environment sealed from the atmosphere by Antarctica's expansive continental ice sheet. The realization that there was life on the Antarctic continent, other than that associated with the marine system, did not come to light until the seminal investigations initiated by the International Geophysical Year in the late 1950s and early 1960s. Biological studies during this period were exploratory in nature, describing life forms and their habitat in terrestrial and surface lake environments. These important pioneering studies reshaped our understanding of the potential for life in the coldest, driest, and windiest place on Earth. However, it was not until about 5 years ago that microbiological evidence suggested that the Antarctic interior could serve as an environment for life, which resulted from investigations conducted on a deep ice core from Vostok Station. These studies suggested that the cryosphere of Antarctica could support some of the most unusual and extreme microbial ecosystems on the planet. This information, coupled with the recent discovery of an extensive network of subglacial lakes beneath the East Antarctic Ice Sheet (EAIS), has changed our view of life in Antarctica. Although it is well documented and widely known that life exists in the permanently cold margins and marine environments of Antarctica, the cryosphere has traditionally been viewed as being devoid of life.

Subglacial Lake Vostok is by far the largest of the more than 140 subglacial lakes that have been identified thus far (Fig. 1), with a surface area of more than 14,000 km^2 and a depth in excess of 800 m, which makes it one of the largest lakes on Earth. The base of the central EAIS is relatively warm as a result of the combined effect of geothermal heating and the insulation caused by the overlying ~4 km of ice, despite surface air temperatures commonly below −60 °C. Lake Vostok has only recently become a focus of research, even though its presence was first noted in the early 1960s when Soviet Antarctic Expedition (SAE) pilots observed an extremely flat area near Vostok Station that could represent ice floating over water. These airborne

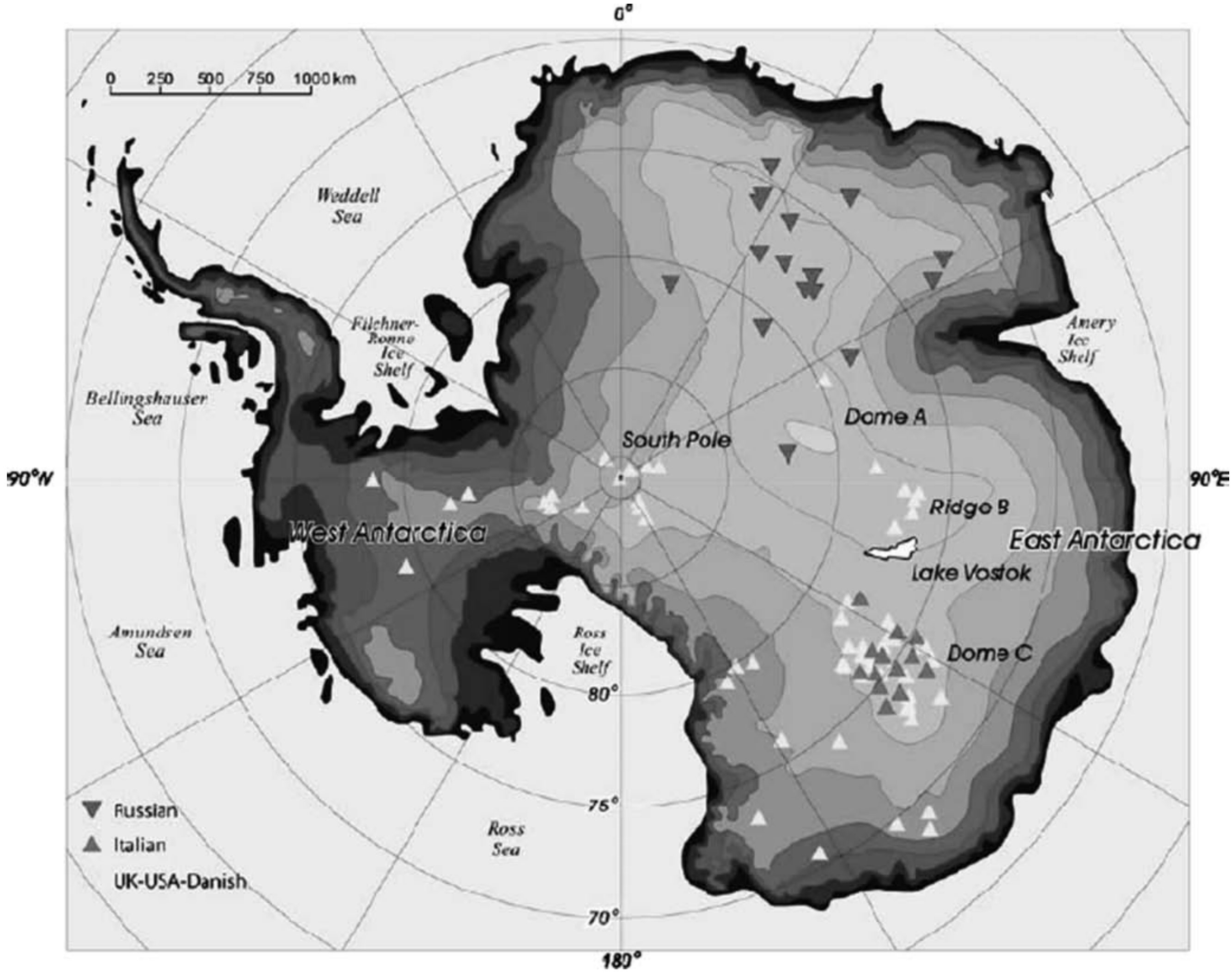

Figure 1. Locations of Antarctic subglacial lakes. Lakes discovered by Italian (blue triangles), Russian (red down white triangles), and U.K.–U.S.–Danish (yellow triangles) teams are included (Lake Vostok is shown in outline). The ice-sheet surface is contoured at 500-m intervals. Reproduced from Priscu et al., '*Subglacial Antarctic Lake Environments: International Planning for Exploration and Research*.' EOS, in press.

observations were soon confirmed by seismic data, radar profiles, and satellite images made near Vostok Station. The discovery of Lake Vostok and several additional subglacial lakes in central Antarctica during the early 1970s went relatively unnoticed by the biological scientific community until the late 1990s. However, curiosity about the nature of this environment and the potential existence of microbial life in subglacial lakes has recently intensified.

The site selection for Vostok Station by the SAE was truly fortuitous for subglacial lake research. Several deep ice cores have been extracted from the ice sheet at Vostok Station since drilling began in the mid-1960s (the first 500-m-deep borehole was made in 1965). The most recent and deepest (3623 m below the surface) ice core (borehole 5G) stopped ~120 m above the lake-ice interface, because of concerns of contaminating the lake. The upper 3310 m of the ice core has provided a detailed paleoclimate record spanning the past 420,000 years. Ice at depths between 3310 and 3538 m is deformed by contact with bedrock and does not contain useful paleoclimate data. The basal portion of the ice core from 3539 to 3623 m has a chemistry and crystallography distinctly different from the "normal"

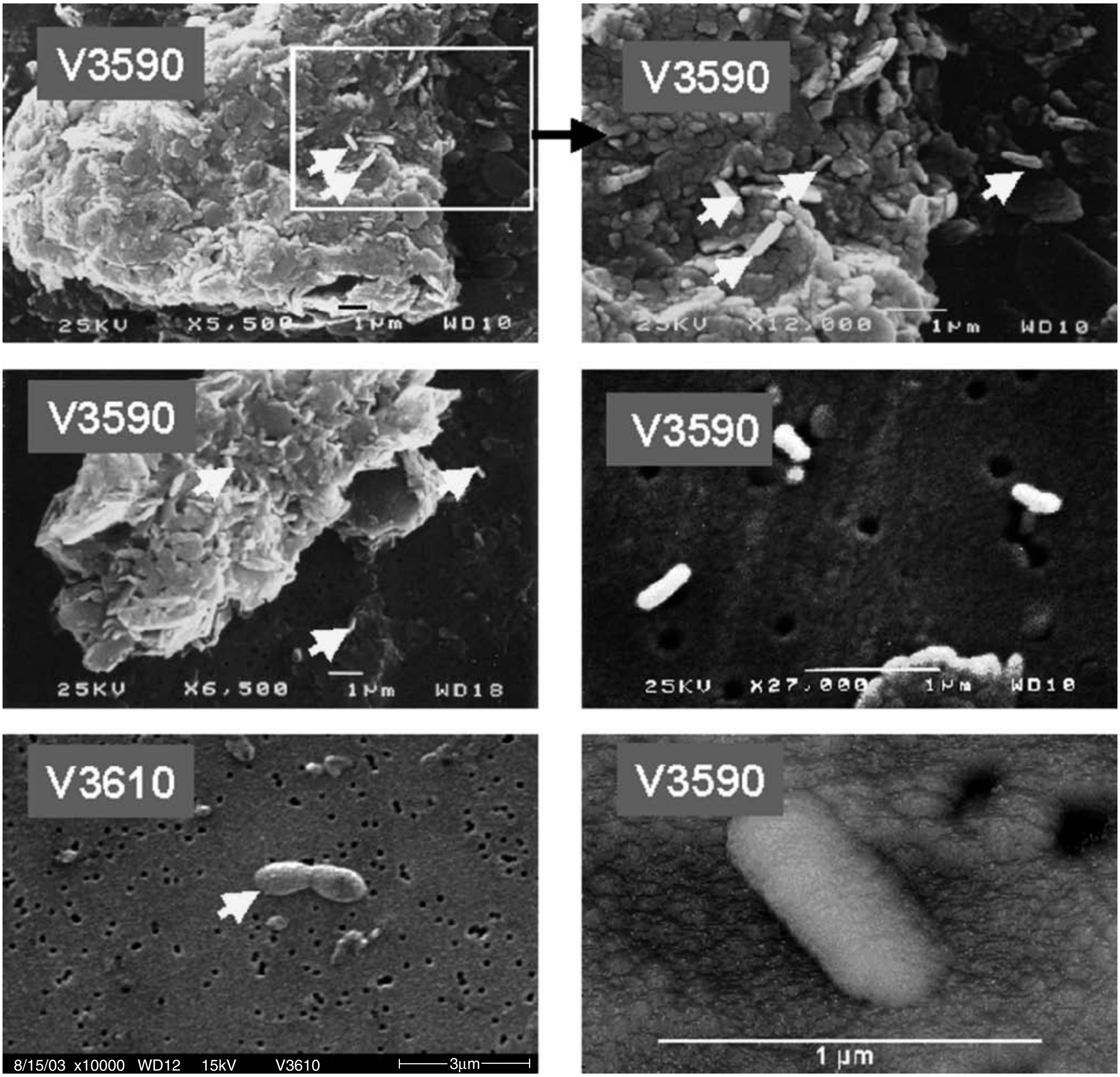

Figure 2. Scanning electron micrographs of microbes in Lake Vostok accretion ice. Samples are designated by their depth of recovery from below the surface in meters, and yellow arrows indicate prokaryotic cells. Melt water was filtered onto 0.2-µm polycarbonate filters sputter coated with 10-nm Au-Pd and imaged with a cryogenic scanning electron microscope. Adapted from Priscu et al., *Science* **286**: 2141–2144 (1999).

glacial ice above it. This basal ice has extremely low conductivity, large (up to 1 m) ice crystals, and numerous sediment inclusions. The mineral composition of ice-bound sediments below 3539 m is dominated by micas and is clearly different than typical crustal composition and particles within the overlying glacial ice. The isotopic composition of this basal ice in concert with other data indicates that it formed by the refreezing (accretion) of Lake Vostok lake water to the underside of the ice sheet. Thus, there is ~210 m of Lake Vostok water frozen on the underside of the glacial ice beneath Vostok Station. Recent airborne radar surveys have shown that the accretion ice is dynamic and actually flows out of the lake on the downstream side.

The freshwater in Lake Vostok originates from the overlying ice sheet, which melts near the shoreline of the lake and at the ice–water interface in the north. Accretion to the base of the ice sheet occurs in the central and southern regions, removing water from the lake. As the accreted ice is essentially gas-free, the lake water is thought to have a relatively high dissolved gas content supplied from air bubbles released from the overlying glacial ice. Estimates on the gas content of Lake Vostok suggest that it could have an oxygen concentration 50 times higher than that in the open ocean. Although it seems inevitable that viable microorganisms from the overlying glacial ice and in sediment scoured from bedrock adjacent to the lake are regularly seeded into the lake, the question remains whether these or preexisting microorganisms have established a flourishing community within Lake Vostok. Evidence for the existence of geothermal activity is supported by the recent interpretation by French scientists of He^3/He^4 data from accretion ice. These data imply that there may be extensive faulting beneath Lake Vostok, which results in geothermal plumes entering the bottom of the lake in the southern part of Lake Vostok. If this emerging picture is correct, Lake Vostok could harbor an ecosystem fueled by geochemical energy much like that observed in deep-sea vent systems.

Because of concerns of contaminating these pristine subglacial environments and the current lack of an adequate sampling strategy to meet these needs, no direct samples have yet been obtained from subglacial lakes. Therefore, the accreted ice retrieved from Vostok Station has offered the only opportunity for microbial ecologists and geochemists to investigate material derived from a subglacial lake. Microbiological and molecular-based studies of accreted ice by four independent laboratories have indicated that these cores contain bacteria (Fig. 2). Metabolic and culturing experiments demonstrate that a portion of the assemblage becomes metabolically active upon melting and exposure to nutrients which suggests Lake Vostok contains a viable microbial community. Molecular profiling of microbes within the accreted ice (by both culturing and direct 16S rDNA amplification) shows close agreement with present-day surface microbiota, consisting of phylotypes within the bacterial divisions *Proteobacteria* (alpha, beta, gamma, and delta), Low G + C Gram Positives, *Actinobacteria*, and the *Cytophaga–Flavobacterium–Bacteroides* line of descent. If microbes within the accreted ice are representative of the lake microbiota, this would imply that bacteria within Lake Vostok do not represent an evolutionarily distinct subglacial biota. The time scale of isolation within Lake Vostok (~20–25 million years) is not long in terms of prokaryotic evolution compared with their 3.7×10^9 year history on Earth. An alternative scenario is that glacial melt water entering the lake forms a lens overlying the Lake Vostok water column. If so, the microbes entrapped within accretion ice would likely have spent little time within the actual lake water (few, if any cell divisions occurring) before being frozen within the accretion ice. The microbes within the main body of the lake below such a freshwater lens may have originated primarily from basal sediments, rocks, and/or the preexisting environment before Lake Vostok became ice-covered. If so, their period of isolation may be adequate for significant evolutionary divergence, particularly given the potential selection pressures that may exist within this unusual subglacial environment.

Epifluorescence (of DNA-stained cells) and scanning electron microscopy have measured bacterial cell densities of 2.8×10^3 and 3.6×10^4 cells mL^{-1}, respectively, in Vostok accretion ice (3590 m below the surface). Additionally, the concentration of dissolved organic carbon (DOC) in this core was measured at 0.51 mg L^{-1}. Based on these values and partitioning coefficients for the water to ice phase change, it is estimated that the water in Lake Vostok had bacterial cell concentrations of 10^5 to 10^6 mL^{-1} and a DOC concentration of 1.2 mg L^{-1}. Using data on the volume of the Antarctic ice sheet and subglacial lakes in concert with published bacterial volume to carbon conversion factors, the cell number and carbon content within the Antarctic ice sheet and subglacial lakes can be estimated, assuming that concentrations of bacterial cells and DOC in other subglacial lakes are similar to those estimated for Lake Vostok. Based on these calculations, subglacial lakes contain about 12% of the total cell number and cell carbon with respect to the pools associated with the Antarctic continent (e.g., subglacial lakes + ice sheet). The number of cells and prokaryotic cell carbon content for subglacial lakes plus the ice sheet (1.00×10^{26} cells; 2.77×10^{-3} Pg) approaches that reported for Earth's freshwater lakes and rivers (1.3×10^{26}; 2.99×10^{-3} Pg), which suggests these environments contain nontrivial amounts of cells and carbon. These estimates of the number of prokaryotes and organic carbon associated with Antarctic subglacial lakes and glacial ice are clearly tentative and should be refined once additional data become available; however, they do imply that Antarctica contains an organic carbon reservoir that should be considered when addressing issues concerning global total carbon storage reservoirs and dynamics.

The seminal reports of life in subglacial Antarctic lakes have spawned many new ideas about how life evolved on our planet and others, how life adapts to extremely cold conditions, and how ecosystems function beneath 4 km of ice, and it has changed our view of the extent of Earth's biosphere. One new idea theorizes that microbes might thrive throughout most of the Antarctic ice sheet within intercrystalline veins. The veins, formed at junctions where three ice crystals meet, may offer a unique habitat for microbes in "solid" ice that has not yet been considered. These veins contain a relatively solute-rich solution that both lowers the freezing point and provides nutrition for microbial growth. This idea is supported by new data on the isotopic composition of gases within cores of Vostok glacial ice, which suggests biological modification. If it is eventually shown that such a microbial habitat does indeed exist in solid ice, then much of the Antarctic ice sheet may harbor active microbes, a notion previously unheard of. Important new information based on airborne radar surveys has shown that at least 140 more subglacial lakes exist beneath the EAIS. Most of these lakes exist near Dome C. These lakes, although not nearly as large as Vostok, are still substantial with at least one exceeding 500 km^2 in surface area. The density of lakes near Dome C, together with maps of bottom topography, implies that a hydrologically interconnected lake district may exist in this region of Antarctica. These lakes provide many new environments that presumably contain life in what was once thought to be an inhospitable environment.

Future research of microbial ecosystems in subglacial lakes depends on a plan in which water samples are collected and returned to the surface. The overriding and limiting issues of this entire strategy are environmental concerns and the control of contamination in both forward and return excursions into the lakes. As such, it is of prime importance that environmental stewardship precedes all scientific endeavors, and that such an undertaking not be attempted until adequate sampling techniques are established. To this end, the Scientific Committee on Antarctic Research (SCAR) has established an international body of specialists to outline a detailed plan for eventual lake entry and sample return. This plan calls for the establishment of a network of instruments that gather limnological data continuously, collection of water samples for return to the surface, and recovery of deep sediment cores that can reconstruct paleoclimate and geological records for Antarctica. The next 10 years should prove to be an interesting time of discovery for Antarctic science, one that follows the Antarctic tradition of melding interdisciplinary and international science. We can expect subglacial lakes to be at the forefront of such discovery because they remain one of the last unexplored frontiers on our planet.

READING LIST

Christner, B.C., Mosley-Thompson, E., Thompson, L.G., and Reeve, J.N. (2001). Isolation of bacteria and 16S rDNAs from Lake Vostok accretion ice. *Environmental Microbiol.* **3**: 570–577.

Dowdeswell, J.A. and Siegert, M.J. (2002). The physiography of modern Antarctic subglacial lakes. *Global and Planetary Change* **35**: 221–236.

Karl, D.M. et al. (1999). Microorganisms in the accreted ice of Lake Vostok, Antarctica. *Science* **286**: 2144–2147.

Priscu, J.C. et al. (1999). Geomicrobiology of subglacial ice above Lake Vostok, Antarctica. *Science* **286**: 2141–2144.

Priscu, J.C. and Christner, B.C. (2004). Earth's icy biosphere. In: *Microbial Diversity and Bioprospecting*. A.T. Bull (Ed.). ASM Press, Washington, DC.

Siegert, M.J., Dowdeswell, J.A., Gorman, M.R., and McIntyre, N.F. (1996). An inventory of Antarctic subglacial lakes. *Antarctic Science* **8**: 281–286.

Studinger, M. et al. (2003). Rystrom, Ice Cover, Landscape Setting, and Geological Framework of Lake Vostok, East Antarctica, Earth Planet. *Sci. Lett.* **205**(3–4): 195–210.

Tabacco, I.E. et al. (2003). Evidence of 14 new subglacial lakes in DomeC-Vostok area. *Terra Antarctica Reports* **8**: 175–179.

WATER—THE CANADIAN TRANSPORTER

Environment Canada

Water plays an important role in the transformation of the Canadian landscape by moving large amounts of soil, in the form of *sediment*. Sediment is eroded from the landscape, transported by river systems, and eventually deposited in a lake or the sea. For example, the Fraser River carries an average of 20 million tonnes of sediment a year into the marine environment.

The sediment cycle starts with the process of *erosion*, whereby particles or fragments are weathered from rock

---

This article is produced by Environment Canada and is reproduced with permission.

material. Action by water, wind, glaciers, and plant and animal activities all contribute to the erosion of the earth's surface. Fluvial sediment is the term used to describe the case where water is the key agent for erosion. Natural, or geologic, erosion takes place slowly, over centuries or millennia. Erosion that occurs as a result of human activity may take place much faster. It is important to understand the role of each when studying sediment transport.

Any material that can be dislodged is ready to be transported. The *transportation* process is initiated on the land surface when raindrops result in sheet erosion. Rills, gullies, streams, and rivers then act as conduits for sediment movement. The greater the discharge, or rate of flow, the higher the capacity there is for sediment transport.

The final process in the cycle is *deposition*. When there is not enough energy to transport the sediment, it comes to rest. Sinks, or depositional areas, can be visible as newly deposited material on a flood plain, bars and islands in a channel, and deltas. Considerable deposition occurs that may not be apparent, as on lake and river beds. A knowledge of sediment dynamics is an integral part of understanding the aquatic ecosystem.

Canada's waterways move many millions of tonnes of sediment annually in this never-ending cycle of erosion, transportation, and deposition. Sediment is measured and classified according to its dynamic characteristics:

- suspended load (suspended in the water)
- bed load (rolling or bouncing along the bottom)
- bed material (stationary on the bed)

WHY IS SEDIMENT IMPORTANT?

Sediment carried in water has a variety of effects: what are they and why are they important?

Toxic Chemicals

Sediment plays a major role in the transport and fate of pollutants and so is clearly a concern in water quality management. Toxic chemicals can become attached, or adsorbed, to sediment particles and then transported to and deposited in other areas. These pollutants may later be released into the environment. By studying the quantity, quality, and characteristics of sediment in the stream, scientists and engineers can determine the sources and evaluate the impact of the pollutants on the aquatic environment. Once the sources and impact are known, action can be taken to reduce the pollutants. The association of toxic chemicals with sediment is an issue of national importance.

Navigation

Deposition of sediment in rivers or lakes can decrease water depth, making navigation difficult or impossible. To ensure access, some of the sediment may be dredged from the stream or harbour, but this may release toxic chemicals into the environment. To determine how much dredging needs to be done and how often, water levels must be monitored, and the rates of sediment transport and deposition estimated. Sedimentation of navigation channels is a concern in the Fraser River (British Columbia), the Mackenzie River (Northwest Territories), and the Great Lakes-St. Lawrence system (Ontario and Quebec).

Fisheries/Aquatic Habitat

Streamborne sediment directly affects fish populations in several ways:

- Suspended sediment decreases the penetration of light into the water. This affects fish feeding and schooling practices, and can lead to reduced survival.
- Suspended sediment in high concentrations irritates the gills of fish, and can cause death.
- Sediment can destroy the protective mucous covering the eyes and scales of fish, making them more susceptible to infection and disease.
- Sediment particles absorb warmth from the sun and thus increase water temperature. This can stress some species of fish.
- Suspended sediment in high concentrations can dislodge plants, invertebrates, and insects in the stream bed. This affects the food source of fish, and can result in smaller and fewer fish.
- Settling sediments can bury and suffocate fish eggs.
- Sediment particles can carry toxic agricultural and industrial compounds. If these are released in the habitat they can cause abnormalities or death in the fish.

Forestry

Some forestry practices have negative impacts on the environment. Extensive tree cutting in an area may not only destroy habitat but increase natural water runoff and accelerate soil erosion. These can lead to increased flow and sediment loads in nearby streams. They can also release chemical substances occurring naturally in forest soils, and allow them to contaminate rivers or lakes. Both the chemicals and the additional sediment can harm fish and other organisms. Sediment problems resulting from forestry practices are prevalent in British Columbia, Ontario, Quebec, New Brunswick, and Newfoundland.

Water Supply

Sediment can affect the delivery of water. When water is taken from streams and lakes for domestic, industrial, and agricultural uses, the presence of sediment in the water can wear out the pumps and turbines. As this increases maintenance costs, it is important to determine the amount of sediment in the stream so that the appropriate equipment can be chosen when designing a water supply plant.

Energy Production

The amount of sediment transported affects both the size and the life expectancy of reservoirs created for power generation. A dam traps sediment that would normally be carried downstream, and that sediment decreases the size of the reservoir and thus its use for power generation. Therefore, it is necessary to know the amount of sediment to ensure the effective design of reservoirs for the long term.

Agriculture

Some farming practices increase soil erosion and add toxic chemicals to the environment. Thus, productive soil is lost to farms, sediment and pollutants are added to streams, and maintenance costs of irrigation systems are increased. Sediment data and information are necessary in the evaluation of cropping practices and their environmental effects. Sediment-related problems associated with agriculture occur across the country.

Case Study: Aquatic Habitat and Construction In St. John's, Newfoundland

A considerable amount of sediment is lost from building sites during construction. Depending on its location, this sediment finds its way to sewer or stream systems, increasing the costs of water treatment or impacting on aquatic habitat. In 1982, construction was started on a building to house the Institute for Marine Dynamics in St. John's, Newfoundland. The site, near the Rennie's River, required massive excavation. To minimize the impact of the sediment on the aquatic environment, a de-sedimentation facility was built. A settling pond was used, and alum was added to help settle the sediment, alum having no significant impact on the ecosystem. Without the de-sedimentation facility, toxic chemicals and sediment could have contaminated the fish, buried fish eggs, dislodged aquatic plants, and generally overwhelmed the aquatic habitat. Over three years, 1250 tonnes of sediment were kept from entering the river waters. The cost of the de-sedimentation was less that one tenth of one percent of the building costs, and a highly productive trout habitat was protected from contamination.

HOW IS SEDIMENT SAMPLED?

Sediment quantity and quality are sampled in a variety of ways:

- Specially designed *suspended-sediment* samplers are used to collect water/sediment samples that are analyzed for sediment quantity and sometimes quality.
- *Bed-load* samples are usually taken by lowering a specially made sampler to the stream bed. Resting there, the sampler traps the material moving along the bottom.
- *Bed-material* samples may be taken simply by hand from exposed bars or stream banks, or by samplers from the stream bed. Some samplers scoop sediment by simply digging into the bed, while other kinds extract a core from the bed.

Once collected, suspended-sediment samples are analyzed for concentration and particle size. This is usually done in a laboratory. The *concentration* is the ratio of sediment (dry weight) to the total water-sediment mixture, expressed as milligrams per litre (mg/L). The *particle size* is simply the size of the sediment particles. Depending on their size, they are classified as sand, silt, or clay.

To find out how much material is transported by a river, one can combine the concentration with the stream discharge, or flow. This gives the sediment load, which indicates the total amount of sediment transported over a certain time period, whether an hour, a day, month, or year. In such a way, it has been estimated that at Montreal the St. Lawrence River transports 2.3 million tonnes of sediment in suspension each year, or the equivalent of 230,000 truck-loads of soil.

Water, and more specifically the hydrologic cycle, plays a major role in driving the sediment life cycle. The amount of water and its distribution over time influence how and when sediment is sampled.

In Canada, sampling is usually for suspended-sediment data. Most of this sampling is done during high-flow conditions (spring, summer, and fall rainstorms), when most of the sediment is transported through the river system. However, a few samples may be taken at other times throughout the year, to better define the sediment regime. Bed-load sampling is typically undertaken in the spring, when high discharge mobilizes the stream bed. Bed-material may be sampled during the summer, when low-flow conditions may expose parts of the stream bed, making sampling easier.

Once the samples have been analyzed, the data on concentrations, particle sizes, or loads can be applied to engineering and environmental questions.

Average Annual Suspended Sediment Load (tonnes) for Selected Rivers in Canada

Glaciers retreated 10,000 years ago, leaving large amounts of easily erodible material across much of western Canada. In mountainous areas (e.g., the Fraser, Peace, and upper Mackenzie rivers), steepness and abundant water supply enable large amounts of sediment to be carried away. In contrast, the flat and dry conditions of the Prairies result in much lower sediment loads. In eastern Canada, where much of the land is bedrock, there is a limited sediment supply and therefore smaller loads.

SEDIMENT DATA AND INFORMATION

The measurement of sediments in streams in Canada dates back to 1948, in Saskatchewan. The federal government has conducted a national sediment program since 1961 in cooperation with the provinces, territories, and other interested agencies, such as hydroelectric companies. Data-collection techniques are standardized across the country to maintain data quality and comparability. Provincial governments also collect sediment data either as part of a regular sampling program or for specific

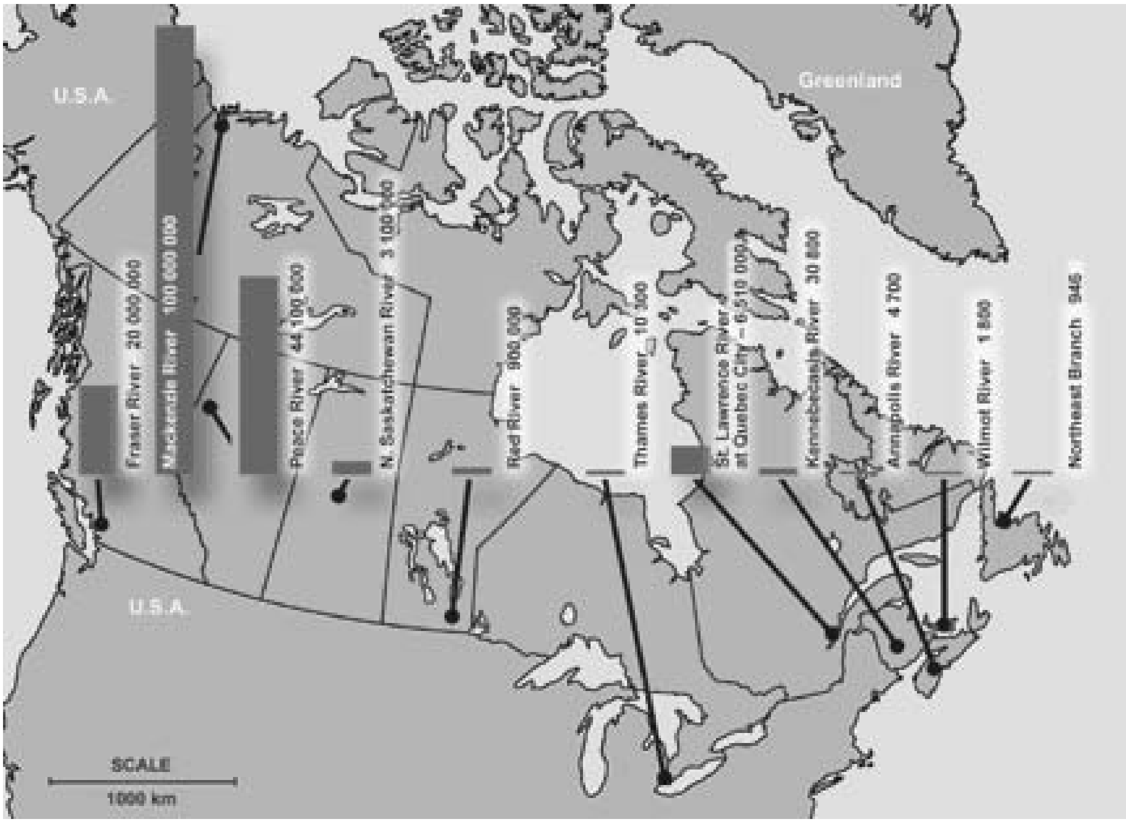

Glaciers retreated 10000 years ago, leaving large amounts of easily erodible material across much of western Canada. In mountainous areas (e.g., the Fraser, Peace, and upper Mackenzie rivers), steepness and abundant water supply enable large amounts of sediments to be carried away. In contrast, the flat and dry conditions of the Prairies result in much lower sediment loads. I eastern Canada, where much of the land is bedrock, there is a limited sediment supply and therefore smaller loads.

studies. Consulting engineers and planners, as well as university researchers, also carry out sampling for site-specific projects.

These data have been used extensively to address reservoir sedimentation, environmental impact assessment, sediment-associated contaminant transport, and other concerns.

Sediment data and information are available from a variety of sources. A large amount of data is contained in a national computer data base operated by Environment Canada. The data base contains historical and current data for about 750 stations throughout the country, about 300 of which are currently monitored.

The types of data that are stored in the national computer data base are as follows:

- suspended-sediment concentrations
- suspended-sediment loads
- suspended-sediment particle size
- bed load
- bed-load particle size
- bed-material particle size
- sediment quality

The data are published annually on CD-Rom and are available from Water Survey of Canada, Meteorological Service of Canada, Environment Canada, Ottawa, Ontario K1A 0H3, Tel.: (613) 992-7121, Fax: (613) 992-4288.

FRESHWATER SERIES A-8

Note: A resource guide, entitled Let's Not Take Water For Granted, is available to help classroom teachers of grades 5–7 use the information from the Water Fact Sheets.

FLOOD PREVENTION

Slobodan P. Simonovic
The University of Western Ontario
London, Ontario, Canada

INTRODUCTION

Although people cannot do much about floods, they can influence the nature and extent of flood losses. The latter can be achieved through flood management policies and programs. The first section of this contribution describes some of the causes of flooding in order to provide insight into why some flood protection and response polices might succeed in one area but fail in another. Since natural flood mechanisms exist and socioeconomic conditions vary throughout the world, a mix of responses is necessary to effectively deal with the flood problem. The second section of this contribution describes the nature of some specific flood protection activities. In this manner, the need for effective flood management is demonstrated.

CAUSES OF FLOODING

Hydrometeorological Mechanisms

Flooding is caused primarily by hydrometeorological mechanisms, acting either individually or in combination (1,2). The regional variability in the intensity of these mechanisms reflects the diversity of climate across the region. The relative importance of the hydrometeorological flood mechanisms varies considerably throughout the year because of their close link to climate. In particular, severe flooding may result when several of these mechanisms occur coincidentally, as can happen in the spring.

Snowmelt Runoff Floods. The melting of a snowpack that has accumulated during the winter months is a common flood type. This occurs in watersheds of all sizes, often in combination with storm-rainfall runoff and/or ice jams. The amount of snowmelt runoff is controlled directly by the thickness, ripeness, and extent of the snowpack and by the rate of melting. The greater the amount of snow on the ground during a melt period, the more water is potentially available for snowmelt runoff. Ultimately, however, the rate of snowmelt is controlled primarily by radiant energy (3). Relatively cool weather causes slow melting of the snowpack and the gradual release of meltwater, while warm weather speeds melting and releases meltwater more rapidly, thereby increasing the possibility of flooding. Severe flooding from snowmelt can happen when there is a rapid shift from cold to warm temperatures in the spring following a winter of high snow accumulation. It can also be triggered during the winter months by a sudden and rapid thaw. Since the climatic factors influencing both the accumulation of snow and the rate of melt are regional, snowmelt flooding commonly occurs over large areas and affects numerous watersheds.

Storm-Rainfall Floods. Heavy or torrential rainfall associated with convective storms and midlatitude and tropical cyclones (including hurricanes) can cause flooding, even after a period of drought. This type of flood can develop rapidly, particularly when the ground is saturated and the rate of rainfall far exceeds the capacity of the ground to absorb the water, or when a significant portion of the ground surface is covered with impermeable materials, such as concrete or pavement in urban areas. The magnitude of storm-rainfall flooding depends on the intensity and duration of the rainfall, the areal extent of the storm, prestorm ground moisture conditions, and drainage basin characteristics (e.g., topography, overburden thickness, vegetative cover, drainage density). Storm-rainfall floods can significantly affect watersheds of up to 100,000 km^2 in size throughout the world. Rivers with larger watersheds generally are unaffected because the zone of intense rainfall within a storm system commonly is concentrated within a smaller portion of the drainage basin. However, prolonged, heavy rainfall over a large area from a succession of storms may produce storm-rainfall flooding along a river draining a very large watershed.

Rain-on-Snow Floods. Rain-on-snow floods are a combination of snowmelt runoff and storm-rainfall floods. They occur when rainfall runoff is augmented by snowmelt, which increases the amount of water flowing into a stream system. The rate of snowmelt during a rainstorm is affected directly by air and rainfall temperatures, amounts of rainfall, and wind speeds. However, the amount of surface runoff released from a snowpack is controlled by snowpack ripeness, since an unripe snowpack will store more rainwater than a ripe snowpack. When heavy or sustained rainfall occurring in combination with other meteorological conditions leads to ripening and significant melting of a moderately deep snowpack, rain-on-snow can generate substantial runoff causing very severe flooding (3).

Ice Jam Floods. Ice jam floods result from the temporary obstruction of river flow by the buildup of ice fragments within the channel and can occur during both the freeze-up and breakup periods. However, ice jams during breakup are more likely to cause flooding. Once formed, an ice jam causes the river to rise immediately upstream and may overtop its banks, depending on the height of the obstruction relative to the sides of the channel. The failure of an ice jam can release a surge of water and ice downstream that causes a sudden rise in water levels and flow velocities downstream. Ice jams that form during the breakup period commonly coincide with the freshet flow arising from snowmelt runoff and can accentuate the level of flooding. Flooding from ice jams tends to be localized since it is dependent directly on the formation of an ice jam, and the tendency for an ice jam to form at any given location along a river is variable. Ice jam formation along a river is promoted by the presence of local sections of intact ice cover during breakup and/or local channel characteristics (e.g., channel shoaling, variation in channel width, channel splitting by islands or bars, and sharp bends). However, ice jams can also form behind bridges and other artificial structures that constrict a channel. The severity of ice jamming varies from year-to-year and depends on factors such as the harshness of the winter, the amount of ice decay and melting prior to breakup, and the amount of rise in river level immediately prior to and at the time of breakup. Ice jams are commonly associated with larger, north-flowing rivers.

Natural Dams

Floods can also be caused by the formation and failure of natural dams, although these events are far more localized and less frequent (at the national scale) than hydrometeorological flooding. Floods from natural dams occur due to the blockage of drainage by landslides, glaciers, and moraines. Flooding occurs upstream of the natural dam as a result of ponding, but also downstream if there is a failure of the dam (or in some instances the development of a tunnel under a glacier) that allows the rapid drainage of the impounded water. These "outburst floods" produce peak discharges that are proportional to the volume of the impoundment, rather than the area of the contributing watershed, and the resulting flood can be larger by an order of magnitude or more than the maximum expected hydrological flood for the stream. Such large floods consequently may cause enormous erosion and channel change along the flood paths for many kilometers downstream of a dam and represent a much greater potential risk than the flooding behind the dam. With some specific glacier dams, outburst floods have happened nearly annually over periods of up to several decades because the dam has reformed repeatedly after successive drainages of the impoundment.

FLOOD MANAGEMENT

Preventing flood damages is one important element of flood management. There are a large number of strategies and methods available today to address flood hazards and disasters. Floods are natural environmental events (2). They present a hazard and may cause a disaster only after humans construct an environment on flood-prone lands (1,4–6). Building structures within floodplains is a development strategy chosen frequently by many societies all over the world. As a result, floods and floodplains need to be "managed" (7). Today, the common framework for dealing with floods comprises four general types of activities: (1) modifying flooding, (2) modifying susceptibility to flooding, (3) modifying the impacts of flooding, and (4) preserving the natural and beneficial functions of floodplains. Modifying flooding refers to structural flood control measures that are designed and constructed to keep the flood hazard away from people and buildings. Modifying susceptibility to flooding refers to activities designed to keep people and buildings away from the flood hazard. These are generally called nonstructural mitigation measures and include a broad set of tools (8,9). Modification of impacts of flooding on people includes floodplain management activities. These activities do not reduce the amount of damage that floods might cause, but they help "spread" the costs associated with those damages. The fourth set of management activities seeks to preserve the natural and beneficial functions of floodplains from an ecological perspective.

Flood Management Process

The flood management process can be divided into three major stages: planning, flood emergency management, and postflood recovery (10). During the *planning* stage, different alternative measures (structural and nonstructural) are analyzed and compared for possible implementation in order to reduce flood damages in the region. The analysis of alternative measures involves project formulation for each measure, understanding advantages and disadvantages of alternative project arrangements, evaluation of positive and negative project impacts, and finally relative comparison of alternative measures. Flood *emergency management* includes regular appraisal of the current flood situation and daily operation of flood control works. A very important aspect of the appraisal process is identification of potential events that could affect the current flood situation (such as dike breaches, wind setup, and heavy rainfall). At this stage decisions are made on urgent major capital works and flood upgrading measures for existing structures. From the appraisal of the current situation information is initiated regarding evacuation and repopulation of different areas. *Postflood recovery* involves numerous hard decisions regarding return to the "normal life." Issues of main concern during this stage of the flood management process include evaluation of damages, rehabilitation of damaged properties, and provision of flood assistance to flood victims. At this stage also, all environmental impacts are evaluated and mitigation strategies selected.

Structural Flood Management Measure

Flood management has the objective to evaluate utilization of possible methods for reducing flood damages to existing buildings and other facilities and reducing flood risk to permit additional growth. Structural measures can be further divided into local protection measures and upstream flood protection measures (usually storage reservoirs). Structural measures emphasize construction of levees or walls to prevent inundation from floods below some specific design flood flow. Additional works may include drainage and pumping facilities for areas that are sealed off from precipitation runoff to the river by the levees; diversion structures to divert flow during the peak from the protected region; channel modifications to increase the hydraulic capacity or stability of the river; and one or more reservoirs upstream from the protected area to capture the volume of a designed flood and release it at nondamaging rates.

Nonstructural Flood Management Measures

Nonstructural measures include zoning (to limit the types of land uses permitted to those that may not be severely damaged by floods), protection of individual properties (e.g., waterproofing of the lower floors of existing buildings), a flood warning system (to evacuate residents and to move valuables), and flood insurance (to recognize the risks of floods and to provide compensation when damages are not avoidable at acceptable cost) (11,12).

Responding to Floods and Sharing the Risks

As no flood risk management program can provide absolute protection, it is also appropriate to examine the institutional arrangements for responding to and recovering from floods. In most countries there are important federal flood management agencies. They usually coordinate and encourage emergency preparedness activities within the federal government, and between federal and lower level governments. In the context of flood response, as for all disaster response, they place initial responsibility upon individuals. Based on the extent of the flood and on an individual's capacity to respond, responsibility can move from municipal, to regional, and finally to federal levels. Each level of government must request the support of the next one. Emergency preparedness and response is clearly a shared activity among individuals, the private sector, and all levels of government.

BIBLIOGRAPHY

1. Parker, D.J. (Ed.). (2000). *Floods.*, Vols. I and II, Rouledge, London.
2. Smith, K. and Ward, R. (1998). *Floods: Physical Processes and Human Impacts*. John Wiley & Sons, Hoboken, NJ.
3. Matthai, H.F. (1990). Floods. In: *Surface Water Hydrology. The Geology of North America*. Wolman, M.G. and Riggs, H.C. (Eds.). Volume O-1. Geological Society of America, Boulder, CO, pp. 97–120.
4. White, G.F. (1945). *Human Adjustments to Floods*. University of Chicago Press, Chicago, IL.
5. Alexander, D. (1993). *Natural Disasters*. UCL Press, London.
6. Mileti, D.S. (1999). *Disasters by Design: A Reassessment of Natural Hazards in the United States*. Joseph Henry Press, Washington, DC.
7. Hunt, C. (1999). A twenty-first century approach to managing floods. *Environments* **27**(1): 97–114.
8. Krimm, R.W. (1998). Making mitigation a reality. *Australian J. Emerg. Manage.* **13**(1): 60–64.
9. Bruce, J.P., Burton, I., and Egener, I.D.M. (1999). *Disaster Mitigation and Preparedness in a Changing Climate*. Institute for Catastrophic Loss Reduction, Toronto, Canada.
10. Simonovic, S.P. (1999). Social criteria for evaluation of flood control measures—Winnipeg case study. *Urban Water* **1**(2): 167–175.
11. Simonovic, S.P. (1999). Decision support system for flood management in the Red River Basin. *Can. Water Resour. J.* **24**(3): 203–223.
12. Simonovic, S.P. (2002). Two new non-structural measures for sustainable management of floods. *Water Int.* **27**(1): 38–46.

EFFECTS OF DDT IN SURFACE WATER ON BIRD ABUNDANCE AND REPRODUCTION—A HISTORY

J. Gordon Edwards
San Jose, California

Early in the DDT controversy so many allegations were made that it was difficult to respond to them all in the available time. Hundreds of scientists patiently refuted those allegations, however, and their contributions are summarized in many books and articles. The news media were notified that they had been misinformed by some environmentalists, and a few carried stories that agreed more with the facts. They also became aware that most scientists disputed the allegations that DDT caused cancer, mutations, hepatitis, poliomyelitis or other human illnesses. It was made clear that DDT residues did NOT "build up" in animal food chains, because they were metabolized or excreted by fish, birds and mammals.

BIRD POPULATIONS THREATENED?

Some persons attempted to link bird declines with the heavy usage of DDT between 1946 and 1960 and to imply that birds later increased after DDT was banned. When those claims were investigated with objectively, it became evident that most declines had occurred either before DDT was present, OR after levels of DDT had waned. The Environmental Defense Fund commented on the abundance of birds during the DDT years, with "increasing numbers of pheasants, quail, doves, turkeys and other game species." To the list of increasing bird populations, the Audubon Society added eagles, gulls, ravens, herons, egrets, swallows, robins, grackles, red-winged blackbirds, cowbirds, and starlings. The annual Audubon Christmas Bird Counts between 1941 (pre-DDT) and 1960 (after DDT use had waned) showed that at least 26 different kinds of birds became more numerous during those decades of greatest DDT usage. They documented an overall increase from 1,379 birds seen per observer in 1941 to 5,742 birds seen per observer in 1960. Statistical analyses of the Audubon data confirmed the perceived increases.

At Hawk Mountain, Pennsylvania, counts of migrating raptors were made daily by teams of ornithologists for more than 40 years. The results were published annually by the Hawk Mountain Sanctuary Association and revealed great increases (not decreases) of most kinds of hawks during the DDT years. Total numbers of hawks counted there increased from 9,291 in 1946 to 13,616 in 1956 and 20,196 by 1967. Sparrow hawks seen were as follows: 1946–98; 1956–192; 1965–408; and 1967–666.

The activists could not explain the population explosions of white-tailed kites in heavily-sprayed California, or of robins, starlings, cowbirds and red-winged blackbirds throughout North America. They tried, saying that not all birds were affected by DDT, but only predators. Because it was evident that predators were also multiplying rapidly, they decided that not all predators were affected, but only those that preyed on fish. But the fish-eating murres, ospreys, and gulls were rapidly increasing, too. The murres on Funk Island (along the North Atlantic Coast) increased during the DDT decades, from 15,000 pairs in 1945 to a million and a half pairs by 1971, and gannets there went from 200 pairs in 1942 to 3,000 pairs in 1970 (1).

Herring gull numbers in Massachusetts exploded during the DDT decades, from 2,000 pairs (1941) to 35,000 pairs (by 1971). Although they were on the "protected" list, Massachusetts gave the Audubon bird-lovers permission to poison 30,000 of them on Tern Island. The local Audubon president, William Drury, said "It's kind of like weeding a garden" (2).

Many other kinds of birds multiplied so well in the presence of DDT and DDE that they became destructive pests. In North Carolina, 6 million blackbirds ruined the town of Scotland Neck in 1970, polluting streams, depositing 9 inches of droppings on the ground, and fouling the forests where they roosted at night. Near Fort Campbell, Kentucky, 77 million blackbirds were roosting, and the resulting threat of human disease (histoplasmosis) was of great concern (3). When marshes were sprayed in the Midwest to control mosquitoes, a common result was reported: "Today in a small area of northern Ohio, 10 million redwings mill about in the cornfields after the nesting season." They destroyed the crops, while farmers begged for avicidal relief. In Virginia, the Department of Agriculture stated, "We can no longer tolerate the damage caused by redwings... 15 million tons of grain are destroyed annually—enough to feed 90 million people" (4).

Rachel Carson started environmental activists down the path to extremism when she wrote *Silent Spring* in 1962 (5). Later, the groups were joined by extremists opposed to technology and dedicated to reducing human populations in Third World Countries in every way possible, including insect-transmitted diseases.

In the front of her book, Rachel Carson wrote that it was "Dedicated to Dr. Albert Schweitzer, who said 'Man has lost the capacity to foresee and to forestall. He will end by destroying the earth.'" Miss Carson certainly knew he was referring to atomic warfare when she quoted that, but she implied that he meant there were hazards from pesticides such as DDT. Because I had found so many other errors in her book, I got a copy of Schweitzer's autobiography to see if he really mentioned DDT. He did, writing on page 253: "How much labor and waste of time these wicked insects do cause us... but a ray of hope, in the use of DDT, is now held out to us" (6).

She also wrote (5, p. 118): "Like the robin, another American bird [the bald eagle] seems to be on the verge of extinction." Yet the robin was never endangered at all! Roger Tory Peterson, America's leading ornithologist, also writing in 1962 (7), stated that the robin was "the most abundant bird in North America". There is no doubt as to who was correct! The eagles had declined before DDT, but they increased during the DDT years. In addition, the Audubon Society's Christmas Bird Counts revealed that twelve times more robins were counted, per person, in the 1960 count than were seen during the 1941 count (before DDT). In addition, there were 21 times more cowbirds, 38 times more blackbirds, 131 times more grackles, and so forth. No matter how deceitful her prose, the influence of Carson's *Silent Spring* was great, and it continues to shape environmentalist propaganda as well as U.S. policy.

I taught medical entomology for years and was concerned about malaria, yellow fever, typhus, encephalitis, river blindness, leishmaniasis, and other tropical diseases that were transmitted by blood-sucking insects and could be controlled only by insecticides. The National Academy of Sciences wrote in its 1970 book, *The Life Sciences*, that "In little more than two decades DDT has prevented 500 million deaths due to malaria and other diseases, that would otherwise have been inevitable" (8). I found no references to those diseases in *Silent Spring*, however she wrote (5, p. 187): "Only yesterday mankind lived in fear of scourges of smallpox, cholera and plague. Now our major concern is no longer with the disease organisms that once were omnipresent; sanitation, better living conditions and new drugs have given us control over infectious disease. Today we are concerned with a different kind of hazard that lurks in our environment." She then launched into many pages of unsupported attacks on insecticides, with never a mention of those millions of human lives that had already been saved by DDT!

Charles Wurster was the chief Environmental Defense Fund biologist. He got some tanks of sea water, put some marine algae in them, and poured large amounts of DDT into the tanks, to see if it would harm the algae. He described how ethanol was added to each flask of DDT, "in order to yield the desired concentrations [500 parts per billion]." The solubility of DDT in water is only 1.2 parts per billion, so we wondered about the validity of his research, exposing the algae to hundreds of times more DDT in solution than would be possible without the alcohol. In previous tests with the same kinds of algae, Ravenna Ukeles and others had found that 600 ppb of DDT in the water did not affect the algae.

Wurster published his non-scientific results in *Science* magazine (9). In 1969 Dr. Paul Ehrlich expanded on Wurster's hoax and wrote what he called a "scenario" based upon it. He titled his article "*Ecocatastrophe*," and published it in *Ramparts* magazine (10). His article began with:

> "The end of the ocean came late in the summer of 1979, and it came even more rapidly than biologists had expected. There had been signs for more than a decade, commencing with the discovery in 1968 that DDT slows down photosynthesis in marine plant life. It was announced in a short paper in *Science*, but to ecologists it smacked of doomsday. They knew that DDT and similar chlorinated hydrocarbons had polluted the entire surface of the earth, including the sea."

Thousands of school children were required to read Ehrlich's scenario. It was a surprise to see that same exact article published by Ehrlich in England a year later, in a journal titled "*The Year's Best Science Fiction*."

BIRD DATA

Although at first many anti-DDT activists said that DDT was killing off adult birds, they soon had to admit it was not. They were not "dropping from the sky dead," as alleged by some environmentalists (5, p. 124), (11), nor were they "falling out of the trees by the thousands," as other extremists wrote (12).

Wurster (13) referred to "increasing numbers of pheasants, quail, turkeys, and other game species," and Audubon naturalists added at least 26 other species that increased in numbers during the decades of greatest DDT use (14,15).

The propagandists changed their story, and said DDT was killing nestling birds. That was shown by tests to be untrue, when all of the food given to experimental nestlings contained DDT, but no adverse effects resulted (16). They next said perhaps DDT makes the eggs thin-shelled or kills the developing embryos. Legitimate tests proved that they were wrong again.

In their effort to make people believe that DDT in normal bird diets caused thinning of eggshells, the environmental industry fostered tendentious feeding experiments during which their birds were exposed to treatments that every poultryman knew would cause severe eggshell thinning. This included removing calcium from the food, starving the birds, depriving them of water, and cutting the light from 16 hours a day to 8 hours a day.

Obviously, heavy DDT usage was not eradicating birds of any kind but what was causing the great *increase* of birds in areas that had been sprayed with DDT? Causes included the reduction in numbers of blood-sucking insect parasites, which provided birds protection from avian malaria, rickettsial-pox, bronchitis, Newcastle disease and encephalitis. Also, more bird-feed seeds and fruits were available for birds after phytophagous insects had been killed by insecticide sprays. Perhaps even more important was the medicinal effect of DDT. When ingested, DDT triggered the induction of hepatic enzymes that detoxified potent carcinogens such as aflatoxins produced by *Aspergillus* fungi, which abound in the natural foods eaten by birds. Aflatoxins are carcinogenic at levels of 0.03 to 0.08 ppm in the diets of birds, fish and mammals. Many humans in third world countries also die because of aflatoxins in their food.

The anti-DDT extremists relied on just three or four studies in their attempt to convince the Environmental Protection Agency hearing examiner that DDT causes serious thinning of the shells of birds. The studies included Bitman et al. (17), and Porter and Wiemeyer (18). Of course they could not explain the thin shells that occurred before DDT was used, nor could they explain the lack of adverse effects on eggs of captive birds that were experimentally dosed with DDT.

Research by Bitman was designed in such away that thin eggshells would have resulted regardless of whether DDT was added to the diet. His Japanese quail were fed a diet containing only 0.56% calcium “to provide calcium stress during egg-laying.” The DDT concentration added to the diet was 4,000 times the concentration in normal human diets in the United States. They blamed the DDT, instead of the calcium deficiency, for the production of thinner shells.

Compare that with the 1970 tests by FWS researchers Tucker and Haegele (19). They fed different amounts of calcium to each of three different cages of quails, *without* DDT: 3% calcium to the “controls”, 1.73% calcium to another group, and 1% calcium to the third group. None got any DDT or DDE. The shells formed by the 1% calcium birds were 9.3% thinner than the controls. Remember, Bitman and his colleagues at the FWS laboratory fed their quail food with only 0.56% calcium! That would be expected to result in much thinner eggshells than did Tucker’s 1% calcium group, even if Bitman had not added any DDT or DDE to the food.

During Congressional testimony, I criticized Bitman’s calcium-deficient quail diets. He evidently understood my criticisms and subsequently performed additional research with his quail. This time he fed the quail 2.7% calcium in the diet, instead of 0.56%. The result was that both the DDT-fed and the DDE-fed birds produced heavier eggshells, and there was no difference in shell thickness between the DDT-fed birds and the ‘controls.’ He sought to publish those results in *Science* magazine, but this time they were rejected. He had to publish them, instead, in *Poultry Science* (20).

Why did *Science* refuse Bitman’s second article? The editor of *Science* had informed Dr. Thomas Jukes earlier that *Science* would never publish any article about DDT that was not antagonistic to that insecticide, and through the years he enforced that attitude. Anti-DDT writers simply kept expanding their anti-DDT propaganda in *Science* magazine. It was written by the same coterie, and for references in *Science*, they simply kept citing each other. No other studies were accepted by America’s leading science journal. Without that sheltered bias, the case against DDT would have quickly folded and millions more human lives would have been saved.

BENEFICIAL RESULTS OF DDT

There was no scientific basis for the allegations that DDT caused bird populations to experience declines. On the contrary, it is obvious that proper use of DDT actually increased the populations of fish, birds and mammals, because of the following actions: (*1*) DDT applications increased the yield of plant products in fields, woods, marshes, fencerows, and roadside ditches; (*2*) they increased protective vegetative cover for birds and mammals and provided better nesting and denning sites; (*3*) they killed mosquitoes, flies, fleas, lice, and mites in sprayed areas, not only reducing stress and annoyance, but also preventing bird diseases transmitted by those arthropods (avian malaria, bronchitis, rickettsia-pox, Newcastle disease, and so forth); and (*4*) DDT stimulated the liver to produce more hepatic microsomal enzymes that broke down toxic substances in the body, including “aflatoxins” that occur commonly in natural foods such as moldy grain, nuts, cotton seed. Those aflatoxins, resulting from *Aspergillus* fungi, are among the most potent carcinogens known in foods. DDT-induced hepatic enzymes prevent the development of tumors in birds and mammals, according to research by McLean and McLean (21), Edward Laws (22), and Salinskas and Okey (23).

The anti-DDT propaganda was refuted by scientific research and also by the multiplying hoards of birds counted during the Christmas Bird Counts and the Hawk Mountain Sanctuary migration counts.

The anti-DDTers were apparently not impressed, even when birds became so abundant that they destroyed fields of corn or broke down the tree branches upon which they roosted. They didn’t seem to realize that DDT had failed to harm the destructive birds. They apparently felt that DDT caused good birds to decline and bad birds to have population explosions!

PERSISTENCE OF DDT RESIDUES

During the seven months of EPA hearings on DDT, Dr. George Woodwell testified regarding an article written by him and Charles Wurster (24). The hearing transcript contained a statement that: “DDT residues in an extensive salt marsh on the south shore of Long Island average more than 13 pounds per acre” (This was discussed on p. 7232 of the hearing transcript, on January 13, 1972), as follows:

The USDA attorney asked Woodwell:

Q “Isn’t it a fact that after you initially studied this marsh you continued your samplings, and found as a result that you were getting an average of one pound per acre of DDT?"

A "No, I wouldn't agree with that."

Q "Didn't you also find out later that one of the areas where you took your samples was an area that was convenient for the mosquito Commission's spray truck, too?"

A "I can't say that I discovered that, either. I don't believe there's any evidence to that effect."

Q "Dr. Wurster, perhaps?" [We had told the attorneys the details of Wurster's Seattle testimony, and Wurster was in the audience, this day.]

A "I don't believe he knows that, either. I don't believe there's any evidence to that effect."

Q "Dr. Wurster, your coauthor, made the following statement at the Washington state hearings, and I'm quoting him verbatim: He testified: 'We have since sampled that marsh more extensively, and we found that the average in the marsh was closer to one pound per acre. The discrepancy was because our initial sampling was in a convenient place, and this turned out to be a convenient place for the mosquito Commission's spray truck, too.' Did you learn that after the fact, Doctor?"

A "That is a true statement in my experience. I did not know that Dr. Wurster had said that, but that is a true statement."

Q "Doctor, have you ever published a retraction of this 13 pounds per acre, or a further article which discloses the results of your further sampling which brings the average down to around one pound per acre?"

A "I never felt that this was necessary."

(Woodwell also admitted they had only taken six samples of the soil in that extensive marsh!)

BREAKDOWN AND DISAPPEARANCE OF DDT

Earlier, George Woodwell et al. wrote (25), that "6 billion pounds of DDT had been used, but only 12 million pounds could be accounted for in all of the earth's biota", and that was "less than a thirtieth of one year's production of DDT in the 1960's." He wrote that "most of the DDT has either been degraded to innocuousness or sequestered in places where it is not freely available." That contrasted so sharply with his testimony during the EPA hearings that a reporter asked him why he had completely omitted it from his testimony. Woodwell replied that the EPA lawyers told him not to mention the article, "lest my testimony be disallowed" (26).

A Department of Interior study at Gulf Breeze, Florida was reported in *Chemical Assays* (27). Huge glass containers were filled with sea water, a large amount of DDT was added, and the closed containers were suspended in the Gulf. Nearly every day, a sample was taken out for analysis. After 38 days, 92% of the DDT (and its metabolites, DDD and DDE), had disappeared.

In Washington state estuaries, the Bureau of Commercial Fisheries monitored pesticide residues in shellfish at 19 stations during 3 years of heavy DDT use (1966–1969). Ninety-three percent of the samples contained less than 10 parts per billion of DDT (and the highest level found was only 0.1 part per million). Only 3% of the samples contained chlorinated hydrocarbon insecticides. Shellfish are known to "concentrate" chlorinated hydrocarbons in their systems at levels 40,000 to 70,000 times as great as that in the surrounding water, so it is evident that the coastal waters were very free of DDT residues.

DDT is broken down rather rapidly by heat, cold, moisture, sunlight, alkalinity, salinity, soil micro-organisms, and a great many other things. Studies abound regarding DDT disappearance from soil, sand, leaf litter, plant tissues, animal tissues, and water.

Croker and Wilson applied DDT to a tidal marsh (28). In less than 24 hours, only traces remained, and even those traces disappeared in five days.

Finally, the environmental extremists had to resort simply to endlessly repeating all of the old allegations and hoping that most people had not heard about the refutations. The failure of the news media to report the truth helped them get away with that ruse, and pseudo-environmental magazines continued to frighten readers with the same old unsubstantiated myths.

BROWN PELICANS

In 1918, T. Gilbert Pearson and Robert Allen investigated habitats of brown pelicans along the 1,500-mile coastline of the Gulf of Mexico from Key West to the mouth of the Rio Grande. Belatedly reporting his results in *National Geographic* magazine in March 1934, (29), Pearson estimated a total of 65,000 pelicans during that survey. Robert Allen repeated that Gulf survey in 1934, after becoming president of the National Audubon Society. He discovered there had been an 82% decrease in pelican numbers in the 16 years since their earlier survey. He saw only 900 pelicans in Texas and practically none in Louisiana (30). In 1937, the results were published *in Adventures in Bird Protection* (31). According to *Conservation In The United States* (32, p. 414), 17 colonies existed along the Texas coast in 1918, but only one colony of 500 pelicans remained in 1939, and it declined to 200 by 1941.

I reported on these and dozens of other findings during my testimony in Congressman W.R. Poage's hearings for the House Agriculture Committee, on March 18 1971. Finley and Keith (of the U. S. Fish & Wildlife Service), however, were issuing bulletins stating that "A catastrophic decline has occurred on the Gulf Coasts of Texas and Louisiana where a population of over 50,000 brown pelicans has all but disappeared since 1961." The cause, they insisted, was DDT!

In 1961, van Tets saw the last active nesting colony in Louisiana, on North Island, containing only 200 pairs (33). Those, plus the 900 reported in Texas, fell far short of the alleged 50,000 conjured up by Finley, so I called him to ask for details. He answered on March 29, 1971, saying "In response to your inquiry, I have investigated the source of the reported 50,000 brown pelicans on the Texas and Louisiana coasts. The only comprehensive estimate was made by T.G. Pearson: 65,000 birds, in his 1934 book." Was it possible that Finley didn't know that the total in 1918 was for the entire 1,500-mile Gulf Coast and was in *National Geographic*, instead of a book? Was it also possible that he was unaware of Robert Allen's 1934 survey? He cited only five references, dated 1879 to 1955, and concluded "Although the reports are sketchy, Jim Keith and I both feel that the estimate of 50,000 is not unreasonably high." Five months later, however, (August 2, 1971), he wrote to Congressman Poage, stating: "The year 1961 was merely a hasty approximation of an unknown time. After reviewing the evidence, I think now that I should have said the 50,000 pelicans disappeared **by** 1961 [instead of **since** 1961]." (emphasis added). The truth was that the total number did not exceed 300 pelicans along the entire Texas coast in 1942 (before DDT), according to Audubon researchers. If the pelicans had somehow multiplied from less than 500 in 1959 to 50,000 before 1961, as Finley and Keith claimed, they would have had to do it in a habitat that was rich in DDT!

The significance of Finley's belated retraction of the untruthful "50,000 pelican disappearance" was great. It removed the possibility that DDT could have caused a pelican disappearance that never occurred! In fact, the declines obviously occurred long before any DDT was present, and resulted from human persecution, oil spills, gas blowouts, bombing, machine gun practice by the military, and hurricanes.

In 1968, Schreiber and DeLong surveyed California's Anacapa Island pelican colony for the Smithsonian Institution and wrote: "While fluctuations in numbers do occur, there has been no apparent decline in recent years along the California coast" (34). (There were no adverse effects from DDT).

In 1968, Jehl and Keith saw dozens of sick pelicans along the shore of the Gulf of California (35). Soon thereafter, coastal southern California experienced a severe epidemic of Newcastle Disease. A major symptom of that disease is the production of misshapen, soft-shelled eggs.

On January 28, 1969, the Santa Barbara Oil Spill occurred near the southern California coast. Oil surrounded Anacapa Island, and two biologists (Jehl and Risebrough) went out in March to see how much damage the oil was causing to sea lions. They also entered a brown pelican colony and it was reported in *Cry California* (36) that there were many soft or broken eggshells in and near the nests there. Jehl recalled that "As soon as the [poor] condition of the eggs was appreciated, Risebrough was able to postulate that chlorinated hydrocarbon pesticides [DDT] were probably responsible" (37, p. 18).

In California, the pelicans had experienced no difficulties during the 20 years of heavy DDT usage but suddenly suffered a reproductive failure just after the great oil spill surrounded their Anacapa Island colony, according to Jehl and Kennedy. Scientists should have been impressed by that, but the oil spill was not discussed in any reports about the pelican nesting failure.

The oil film, through which the pelicans were diving for fish, contained 21 ppm of mercury (38), and the anchovies upon which the pelicans fed contained 17 ppm of mercury. Mercury is known to have in serious adverse effects on birds, their eggs, and their young.

Environmentalists concealed the fact that California Fish & Game biologists had found that those anchovies contained high levels of lead, which is known to cause severe shell-thinning. They collected hundreds of pelican eggs from that colony during the next two summers, and their shells were measured with screw micrometers. In April 1972, I obtained all of those measurements and found that they clearly revealed inverse correlations between DDT residues and shell thicknesses. Some of the thinnest shells were those of eggs with low DDT levels, and the higher DDT concentrations were often in the thicker-shelled eggs (39). That was also reported by Switzer at the 1972 EPA hearings on transcript pages 1812–1836. I had discussed all of this during my testimony before Congressman Poage, and it was published as a document by the *Committee on Agriculture*, March 18, 1971 (40). Robert Finley, who was in charge of Fish & Wildlife Service pelican studies throughout the United States, responded to my testimony before Congressman Poage's Committee by writing Poage (August 2, 1971) that "there is not a shred of evidence that spilled oil is capable of causing thin-shelled eggs or otherwise affecting bird reproduction." On the contrary, Hartung (41) said that spilled oil reduces hatchability by 68%. Anonymous reports in *Conservation News* (NWF) (42) discussed "Embryo mortality from oil on feathers of adult birds." Dieter wrote that "oil on eggs kills 76 to 98% of embryos inside, and birds that have ingested oil produce 70% to 100% fewer eggs than normal" and "Liver and spleen atrophy may result from oil exposure." (43). King (44) said, "oil is a cause of pelican mortality for six weeks after a spill." Szaro (45) reported that "oil causes kidney nephrosis, pancreatic degeneration and lipid pneumonia." Birds may lose their ability to compensate for osmotic water loss when oil inhibits water uptake by the intestine, which in turn fails to trigger the necessary desalination of water by the nasal glands. He concluded that "It is possible that dehydration may cause the death of many oil-contaminated birds."

In 1923, Dawson described, in *Birds of California* (46), how "merely the brief presence of a human on Anacapa caused adult pelicans to spring from their nests, breaking their eggs with their feet as they did so, and the exposed eggs and nestlings were quickly eaten by marauding gulls." All such warnings were ignored by the Fish & Wildlife specialists! The FWS biologist in charge (Jehl) even admitted that **they had shotgunned pelicans sitting on their nests of eggs** (47), so they could compare the DDT residues in the birds with those in their eggs. They also admitted that they had **collected 72% of all intact eggs in the Anacapa pelican colonies for analysis!** Those analyses revealed an inverse statistical correlation between DDT residues and shell thickness! When the government biologists again invaded the Anacapa colony in 1969, the results must certainly have been anticipated, yet they sought to blame the nesting difficulties entirely on "traces of DDT in the fish" eaten by the birds.

My student and collaborator, Richard Main, was employed by the National Park Service for 2 years. He trapped many large Norway rats beside pelican nests in the Anacapa colony (in "safe traps"). The National Park Service wanted to poison the rats, to protect the young pelicans, but that was prohibited by the California Fish & Game Department because, they said, they feared that native rodents there might also be killed. Several feral cats also roamed about in the colony because a man named Frenchy left them there when he moved from Frenchy's Cove a few years earlier.

Franklin Gress, a graduate student at the University of California, studied the Anacapa pelicans in 1970. The desertion of the colony occurred following his flights to the island by helicopter and his hours sitting in nests in the colonies. His extensive notes were released in 1970 by the Resources Agency of the California Department of Fish & Game. His first visit was March 29, when he counted 127 nests, but he wrote that "The number of nesting birds dwindled, as nests were deserted, until by April 1st the colony was virtually abandoned." He returned on April 20th, and soon wrote that "the nesting attempts seemed to be irresolute." Two days later "the colony was suddenly deserted." A similar result followed his third visit, on June 1st, when he noted that "It was evident that the process of abandonment had begun." By June 20, "the colony was nearly abandoned, and there was no other nesting on the island." On June 29, he returned and reported that "only one nest contained a nestling... this was the only pelican to hatch in either colony on Anacapa Island in 1970." (It should be noted that the Park Rangers who patrolled the island by boat laughed and said they saw numerous young pelicans standing on nests during the intervals between Gress' disruptive visits.)

In 1971, at the request of the National Park Service and Secretary of the Interior William Pecora, Richard Main and I drew up protective regulations for the Anacapa pelican colonies, keeping the destructive government biologists out of the colony during the following breeding season, from March 15 to July 31. The oil was mostly gone, and there had not been time enough for DDT levels to decrease in the water or in the fish, yet the pelican productivity was great! The pelicans quickly recovered their breeding potential (48,49).

The Federal Fish & Wildlife Service and the California Fish & Game Department continued to issue pelican reports that never mentioned the oil spill, the rats, the disruptions by humans and helicopters, the egg predation by biologists, the lead in the pelican bones, or the epidemic of Newcastle Disease in coastal Southern California. They always attributed the alleged declines, however, to DDT!

Those allegations played a very important part in the EPA hearings that led to the banning of DDT in December 1972. We, and many National Park personnel, were convinced that they wanted to totally destroy the colony which would help the EPA to ban DDT. (Malaria would then continue to destroy millions of humans in the Third World and halt the population explosions there.) As it turned out, they succeeded in the broader goal, but failed to actually destroy the pelican colonies.

The Fish and Wildlife Service never publicly corrected its erroneous "since 1961" propaganda, so the news media continued to say that DDT nearly caused the marvelous pelicans to become extinct during the 1960s. Perhaps many people still think that the pelicans were saved from extinction "because attorney William Ruckelshaus banned DDT in 1972!"

That was enough to shake our confidence in the Fish & Wildlife Service, the Environmental Protection Agency, their pelican propaganda, and their integrity!

THE ENDANGERED OSPREY

Long before DDT was used in this country, Hickey (50) attributed a 70% decline of eastern ospreys to "pole trapping" around fish hatcheries. In 1971, Chesapeake Bay ospreys had no problem with DDT or DDE, but "lipids of osprey eggs had PCB levels from 545 to 2270 parts per million" (51). In 1972 Hickey avoided mentioning that ospreys were abundant throughout the west, despite massive applications of DDT there (52). Studies in Wisconsin and Michigan revealed an increase in nesting osprey productivity, from 11 young in 1965 to 74 young in 1970, before the DDT ban. Because data from other regions were sparse or unavailable, many incorrectly assumed that the population crashes were uniform throughout North America, but actually Ornithologist Charles Henny said, "there may be more of the birds in the West than ever before. Their future appears good" (53). Henny estimated that there were 5,000 or more breeding pairs in North America, plus non-breeding adults and juveniles. He wrote that "The population declines of the 1960s, largely attributed to DDT, were not as widespread as originally thought."

After DDT was developed, Hickey ignored the pole trapping and shooting of ospreys and blamed only DDT for all bird declines. I was a member of the Hawk Mountain Sanctuary Association and was impressed by its published data on ospreys. At Hawk Mountain, Pennsylvania, the survey of migrating ospreys during the DDT years revealed the following results: 191 ospreys in 1946, 254 in 1951, 352 in 1961, 457 in 1967, 529 in 1969, and 630 in 1972 (54). It was difficult to believe the charges that DDT threatened osprey extinction, when more were migrating every year after the introduction of DDT, setting new all-time records in several of those years.

There were many reports of high levels of mercury in the fish upon which ospreys depended for food. (Mercury causes drastic thinning of shells of eggs.) It was alleged, however, that DDT and DDE had contaminated the fish world-wide, and that, as a result, fish-eating birds could not reproduce successfully! Chesapeake Bay ospreys had no trouble with DDT or DDE, but their eggs contained high PCB levels (51). That was shown to cause drastic thinning of eggshells, as well as many other adverse effects, yet environmentalists continued to place the blame solely on DDT, despite the fact that feeding birds high levels of DDT alone did not cause birds to produce such thin eggshells.

After DDT was banned, the Hawk Mountain Sanctuary Association *Newsletter* reported in April 1976 that "For reasons we do not understand at all, the number of osprey going south is returning to something like normal, just 318 in 1974 and 279 in 1975." (Notice how many fewer there were in the years after DDT disappeared!)

BALD EAGLES

These birds have suffered much at the hand of humans, but not because of DDT. In 1917, Alaska enacted legislation providing a bounty of 50¢ per eagle and increased the bounty, later. From that time until 1952, the state paid bounties on 128,000 dead eagles (more than $100,000). An article in a 1921 issue of *Ecology* magazine (55) was titled "Threatened Extinction of the Bald Eagle."

In the 1930s, there were no records of bald eagles in most of New England, not more than 10 in Pennsylvania, and perhaps 15 or 20 along Chesapeake Bay. *Bird Lore* magazine (56) stated, "this will give some idea of the rarity of the eagle in eastern United States." That was 14 years before DDT was present, and the eagles surely could not have disappeared in anticipation of DDT!

In 1937, Bent wrote that "the bald eagle probably nested at one time over much of New England, but there are no recent authentic records of them nesting in the three southern states there." (57).

Each year, thousands of bird watchers participate in the annual Christmas Bird Count for the National Audubon Society. In 1941 (before any DDT was used), they saw 197 bald eagles (14), but in 1960 (after 15 years of heavy DDT use), the count was 891 bald eagles seen (15). If those results had been the reverse, I might have also thought that DDT was killing them!

The Hawk Mountain Sanctuary censuses, tabulated almost daily throughout the years, showed that the number of bald eagles migrating over Pennsylvania more than doubled during the first 6 years of heavy DDT use (1946–1952). (Their summaries also revealed great increases in most other kinds of raptors during the "DDT years.")

In 1943, in *Science News Letter* (58), it was stated "When the timber was cleared it was inevitable that the eagles had to go. The cities grew and fouled the rivers with sewage and industrial waste. The once-teeming fish population vanished. With their main source of food thus taken away, it was only natural that the eagles might vanish also." Notice that was written years before DDT appeared in North America!

In America's Eagle Heritage (59), the study of bald eagles in the Everglades is discussed. "In 1959 there were 24 active nests, which produced 18 young; in 1964 there were 51 nests, which produced 41 eaglets." Notice the sizeable increase! The chief Everglades biologist reported that "The population is stable and is reproducing at a rate more than adequate to maintain its numbers (and) I know of no evidence that the region ever supported a larger number of nesting eagles."

At the Patuxent Wildlife Research Center in Maryland (U.S. Department of Interior), autopsies were performed on all bald eagles found dead in eastern U.S. between 1960 and 1965. Coon et al. (60) stated that none of those deaths was blamed on DDT. Forty-six had been shot or trapped, and 7 died of injuries from flying into buildings or towers. Between 1965 and 1980, they performed autopsies on 652 more bald eagles from across the nation. The major cause of death was always shooting and trapping, but there were also many deaths from electrocution or impact with solid structures. The scientists concluded that "the role of pesticides has been greatly exaggerated." In 1972, the carcasses contained twice as much PCB as DDT, DDD, and DDE combined, and the brains had four times as much PCB" (61).

In 1983, New York State contained only three active eagle nests. Later, 150 bald eagles were imported from Alaska. Peter Nye (62) wrote, "My colleagues and I identified 80 locations in New York believed to have been bald eagle nesting sites between 1880 and 1960. By 1940 only a handful of pairs could be identified; yet [he said] the oft-mentioned culprit, DDT, wasn't put into use until the last few nesting eagles were already struggling for their survival" Nye and his colleagues imported eaglets from Alaska beginning in 1976, including 49 in the first 3 years. Now a healthy population is there.

Postupalski (63) found no significant correlation between DDT or DDE residues and shell thickness in a large series of Bald Eagle eggs. However, Stickel reported that bald eagles produced eggs whose shells were 14% to 21% thinner after lead, mercury, or a small dose of parathion was added to their food (64).

Krantz et al. (65) analyzed bald eagle eggs from three states and found that eggshells from Florida were thinnest (0.50 mm), followed by those from Maine (0.53) and Wisconsin (0.55). Analyses of their DDT residues revealed an inverse correlation: the eggs from Maine contained 21.76 ppm, twice as much as those from Florida and five times as much as those from Wisconsin. Obviously, DDT and DDE were not correlated with that shell thinning.

In Wisconsin and Michigan, an increase in nesting bald eagle productivity was noted, from 51 young produced in 1964 to 107 in 1970. In 1966, Fish & Wildlife biologists fed large amounts of DDT to captive bald eagles for 112 days, then concluded that "DDT residues encountered by eagles in the environment would not adversely affect eagles or their eggs" (64). Other wildlife authorities attributed bald eagle reductions to "widespread loss of suitable habitat," but said "illegal shooting continues to be the leading cause of direct mortality in both adult and immature bald eagles."

To summarize, the bald eagle in the lower 48 states was on the verge of extinction in the 1920s and 1930s, long before DDT was discovered. They were shot on sight for fun, bounty, or feathers, trapped accidentally, killed by impact with buildings and towers, or electrocuted by power lines. Many more may have died because of the high levels of PCBs in their lipids and brains. Thousands have now been reared in captivity and released into the wild. The increased numbers of bald eagles are admired and protected.

The most surprising thing is that the environmental industry and the news media continued erroneously to attribute the increase in eagle numbers from 1946 to 1970 to just one thing... the 1972 ban on DDT!

EGGSHELL THINNING

Shell-less eggs or eggs with thin or distorted shells are the result of any of a great number of factors. Older birds produce thin-shelled eggs, as do some young birds, and perfectly normal eggshells become 5% to 10% thinner as the developing embryo withdraws calcium for bone formation. Even the same bird produces varying eggshell thicknesses as time passes. Higher temperatures and dehydration cause thinner shells to be formed, as does higher relative humidity. A reduction in illumination causes drastic changes in shell formation, and the duration of the light has physiological effects as well as regulating the amount of food the birds have time to eat. Stress from excitement, fear, and noise is very destructive to egg-laying success. Simple restraint interferes with the transport of calcium throughout the bird's body, preventing it from reaching the shell gland and forming good eggshells. Lower than normal temperatures reduce the carbonic anhydrase activity, which causes thinner shells.

A great many environmental substances inhibit a bird's ability to produce normal eggshells (oil, lead, mercury, PCBs, 2,4-D, cadmium, lithium, chromium, selenium, and sulfur compounds in the diet). Sulfa drugs, especially sulfanilamide, may result in the complete absence of a shell around the yolk.

Tucker (66) reported that after water was withheld from quail for 36 hours, they produced eggshells 29.6% thinner when drinking was resumed. He also found that parathion ingestion caused 4.8% thinner shells, Sevin caused 8.7% thinning, mercury caused 8.6% thinning, and lead caused 14.5% thinner shells, but *p,p′*-DDT and DDE caused NO shell thinning and *o,p′*-DDT caused shells 0.5% thicker to be formed.

Many bird diseases, such as avian bronchitis, avian malaria, rickettsia pox, or Newcastle disease, may cause great thinning of eggshells. Such information has been well covered in the literature, including a great book, *The Avian Egg*, by the Romanoffs (67). The environmental extremists, who sought to blame DDT, never mentioned that book or Romanoff's *The Avian Embryo* (68).

A deficiency of phosphorus or Vitamin D also results in thin eggshells, but the most notorious cause of thinning is the lack of sufficient calcium. When the diet contains less than 1% calcium, the bird's leg bones may soften to such an extent that even walking is difficult. Wasn't it curious that the FWS researchers at Patuxent deliberately fed their DDT-dosed birds only food that was deficient in calcium, and then attributed the shell thinning to the massive amounts of DDT and DDE they added to the diets?

Dozens of researchers (and thousands of farmers) found that DDT did *not* cause shell-thinning. Jefferies fed his finches 50 to 300 micrograms per day of DDT, which "caused calcium metabolism to tend toward the production of *heavier* shells." (emphasis added) After 2 years they still formed shells that were 7% thicker (69). Frank Cherms fed quail 200 mg/kg of DDT for months, with no effect on shell thickness through four generations. DeWitt (70) fed his pheasants diets containing 50 ppm DDT throughout the year, and reported that the birds produced 23% more hatchable eggs than the controls (and 2 weeks later 100% of the DDT chicks were alive, but 6% of the "controls" had died).

It was demonstrated repeatedly in caged experiments that DDT, DDD, and DDE did not cause shell-thinning, even at levels hundreds of times greater than wild birds would ever accumulate. During many years of carefully controlled feeding experiments, Scott et al. (71) found "no tremors, no mortality, no thinning of egg shells and no interference with reproduction caused by levels of DDT which were as high as those reported to be present in most of the wild birds where 'catastrophic' decreases in shell quality and reproduction have been claimed." Scott et al. also found that, "DDT did not have any deleterious effect upon the sex hormones involved in egg production and indeed may have had a beneficial effect upon egg shell quality."

Early egg collectors routinely saved only thick-shelled eggs. Comparisons between the shell thickness of museum eggs and those collected more recently are misleading and have led to much of the eggshell propaganda during the DDT years. Hickey and Anderson (72) measured the shells of 1729 museum eggs, and reported differences in shell weights of the eggs of five bird species, but said "changes did not occur in the five other species we studied, or in dozens of others, not studied." Four of the five species they reported in their big table had eggs that were *heavier* after DDT use began than those produced during pre-DDT years. Golden eagle eggshells during the "DDT years" were 5% thicker than those produced before any DDT was used. Red-tailed hawk eggs just before DDT was present had much thinner shells than did the pre-1937 eggs, but during the years of heavy DDT use, the birds produced shells that were 6% thicker.

Claus and Bolander in *Ecological Sanity* (73) wrote,

> "We have nowhere seen reports of changes in the shell thickness of any species which exceed 50 to 60 microns (0.05 to 0.06 mm) and in some cases differences of 5 to 7 microns have been reported as 'significant decreases' (0.005 to 0.007 mm differences). Since only the 2nd and in some cases the 3rd decimal point differ, one wonders how significant those differences really were!"

Morris et al. (74) reported that "the most critical area of light intensity change is in the range between 8 and 16 hours and decreasing photoperiods depress rate of lay." It was interesting that Peakall (75) deliberately used those precise hours in his efforts to incriminate DDT. Peakall kept his ring-doves in a cage on a 16-hour light and 8-hour dark schedule. When ready for the experiment, he put the female of each pair into isolation and abruptly reduced the light to only 8 hours per day. At the same time, he began feeding the doves 10 ppm of DDT in all of their food! Poultry would have stopped laying entirely under those conditions, but the doves continued to produce good eggs. Apparently determined to stop their productivity, Peakall then began to inject intraperitoneally large amounts (150 mg/kg) of DDE (not DDT!) into each bird! (They still continued to lay eggs.)

Hauser (76) warned that "Providing adequate light is just as important as vaccination, parasite control, and feeding programs. When lightbulbs in the hen-house are dirty and the hen receives less light than needed, the stress may be severe enough to cause a premature molt." (The effect of Peakall's 50% reduction in illumination would obviously be severe!)

Details of such "experiments" were exposed and discussed in detail by many qualified biologists, however, the media, the pseudoenvironmentalists, and the financially-rewarded anti-DDT "researchers" failed to publicize frauds. Consequently, the general public remained misinformed about the causes of eggshell thinning, and anti-DDT extremists continued to accumulate money and power as a result.

Measuring Eggshell Thickness

The great book by Claus and Bolander (73) includes nearly 300 pages dealing with DDT frauds. They revealed many examples of the untruthfulness of anti-DDT propagandists. A section titled "The Turning of the Screw" (73, pp. 400–405) exposes the reasons that screw micrometers failed to measure bird eggshells accurately. The differences between thicknesses of shells were measured "to the nearest 0.01 mm" (10 microns) and were, in some cases, reported as "significant" thinning if the difference was 5 to 7 microns! Claus and Bolander reminded readers that a human red blood cell is about 8 microns across and a human hair may be 50 to 80 microns in diameter. Also, an eggshell that has a soft texture and an easily-compressible inner membrane might not be measured very accurately with a screw micrometer by a person seeking to prove that DDT caused eggshell thinning.

POLYCHLORINATED BIPHENYLS

There are known to be about 200 different PCBs in the environment. They came into usage in the 1930s as plasticizers on wires and cables and for protective coatings on metals, plastics, and wood. They are also paint extenders and heat exchange fluids in transformers. Every fluorescent light ballast contained liquid PCB, so gas chromatograph analyses for DDT performed in rooms with those lights distorted all results. The plastic tubing in the apparatus, and other plastics that had been in contact with the samples also resulted in false analyses of DDT. Samples to be analyzed were frequently stored in plastic or cellophane containers, which also produced PCBs. In 1985, Wolff and Peel (77) confirmed that the famous allegations of high DDT levels in Antarctic samples "were overblown by more than 100-fold" because the plastic bags in which they were stored contained so much PCB (78). In 1969, Peterle (79) attempted to measure DDT in Antarctic snow and melt-water. Only one of his five filters indicated any DDT presence, but based on that single filter, he wrote that "there should be five million pounds of DDT in Antarctic snow." (No other analyses confirmed his guess.)

PCBs occurred in most fish and wildlife samples at higher concentrations than did DDT residues. In 1970, Risebrough et al. pointed out that "PCBs are more important enzyme inducers than DDE, causing delayed breeding" (80). When ingested by birds, they were known to cause infertility, eggshell thinning, non-hatchability, and severe physiological damage. Curley et al. (81) stated that "PCBs throw off hormonal levels in birds, affecting calcium reserves, which in turn result in thinner eggshells."

In 1971, Risebrough and Spitzer stated that Chesapeake Bay ospreys had no trouble with DDT or DDE, but that their eggs contained PCB levels from 545 to 2,270 parts per million (82). Boyle (83) reported that Herring gulls from the Great Lakes contained an average of 2,224 ppm of PCB" and "signs of eggshell thinning were found in nine of thirteen species in the region."

Scott et al. (84) observed that "Many reports relating reproductive declines of wild birds to DDT and DDE were based on analytical procedures that did not distinguish between DDT and PCBs." Because PCBs caused such severe effects on fish, birds, and wildlife, it was noteworthy that so great a proportion of the difficulties were later admitted to have really been PCBs, instead of DDT. Feeding birds high levels of DDT alone did not cause birds to produce those thin eggshells.

By 1971, 5,000 tons of PCBs were being produced annually in the United States. After 20 years of faulty analyses of DDT, it was finally admitted that we had been misinformed and that many individuals who *knew* the truth had deliberately concealed it to foster banning DDT!

Lichtenstein et al. (85) prepared a table revealing GLC analyses of 11 Aroclor™ plasticizers and listing various chlorinated hydrocarbon insecticides whose peaks were identical to some peaks of PCBs. Those insecticides included DDT, DDD, Lindane, heptachlor, heptachlor epoxide, aldrin, and dieldrin.

Aquatic habitats have often been seriously threatened by PCBs. Harvey (86) and (87) stated that there was "twice as much PCB as DDT in Atlantic sea water," and Giam et al. (88) found similar levels in plankton from the Gulf of Mexico and the Caribbean. Hom et al. (89) reported that there were large amounts of PCB in the Santa Barbara Basin of southern California (near Anacapa Island), where a high "apparent DDE" concentration was reported in sediment that was deposited 12 years before any DDT or DDE existed. They explained that away by saying "we attribute the higher DDE concentration in the 1930 sediment to spurious contamination during collection, storage or analysis."

In 1965, Anderson et al. analyzed five samples of egg-pools, which they then claimed contained high levels of DDT. Years later, after DDT had been banned, they reanalyzed those five samples and belatedly reported that three of the five samples actually contained no DDT at all, and the other two had only a fourth as much as touted in their earlier article (90). Scott et al. (84, p. 364) reminded us that those authors had led the campaign to ban DDT because of its supposedly "proven effects on eggshells," but they later sought to implicate PCBs as having been the *real* hazard.

Sherman (91) called attention to these developments and observed that "This negated the putative data that were the basis for previous alleged sensational charges against DDT. For thirty years DDT was a scapegoat for artifacts and mimics of that pesticide."

If the errors that were so frequently repeated by anti-DDT activists were actually unintentional, they should have corrected them. That practice of not correcting misstatements also seemed to appeal to journalists, politicians, and environmentalists. As a result, DDT was eventually placed on trial by William Ruckelshaus, the head of the Environmental Protection Agency.

In April 1972, after 7 months presiding over those EPA hearings, Judge Edmund Sweeney concluded that "DDT is not a carcinogenic hazard to man... DDT is not a mutagenic or teratogenic hazard to man... The uses of DDT under the regulations involved here do not have a deleterious effect on freshwater fish, estuarine organisms, wild birds or other wildlife. The evidence in this proceeding supports the conclusion that there is a present need for the essential uses of DDT" (92).

Incredibly, EPA Administrator Ruckelshaus (an attorney, with no knowledge of the scientific facts) never attended a single day of the seven months of hearings, and (according to his special assistant, Marshall Miller, did not even read the transcript of those vital hearings) (93). However, he overruled the judge (and ignored the hundreds of truthful scientists who had traveled to testify during the hearings). He was a strong supporter of the Environmental Defense Fund, and deliberately ignored the scientists' 9,000 pages of testimony and personally banned DDT. Several of the serious errors and untruths in Ruckelshaus' "Opinion and Decision" concerning DDT were exposed and were published in the *Congressional Record* by Senator Barry Goldwater in 1972 (94). The information concerning his decision was never adequately reported by the news media or by the pseudoenvironmentalists. Ruckelshaus mistakenly applied the wrong name to DDT, erroneously stated that DDE was also an insecticide, and proposed that applications of DDT be replaced by parathion (which had already killed or injured hundreds of people).

John Quarles' June 3, 1982 affidavit to the U.S. District Court for Northern Alabama is of interest because it discussed the irresponsible Ruckelshaus ban of DDT. Quarles had served as General Council for Mr. Ruckelshaus in 1971 and 1972. He testified, "After seven months of hearings, the EPA Hearing Examiner made findings generally supportive of the position that DDT did not cause undue harm and that an adequate basis did not exist for cancelling the uses of DDT." Quarles added that "There were no findings that DDT had caused harm or would cause harm under a specific set of circumstances or at any particular time or place." This was of course not reported by the news media and was buried by every environmental group.

On April 26, 1979, in a letter to Allan Grant (the president of the American Farm Bureau Federation), Mr. Ruckelshaus wrote, "Decisions by the government involving the use of toxic substances are political, with a small 'p'. Science has a role to play, but the ultimate judgement remains political, (and) the power to make this judgement has been delegated to the Administrator of the EPA." By the time Ruckelshaus wrote that letter, he had become senior vice president of the (non-environmental) Weyerhauser Lumber Company! Later he took charge of the Browning-Ferris garbage disposal company and sought to dump trash and garbage into a pleasant little stream in Apanolia Canyon, near Half Moon Bay, California. That would have destroyed irreplaceable riparian habitats and an ecological complex of great diversity." Fortunately, local outrage against Ruckelshaus's scheme halted his destructive actions, *that* time!

MISIDENTIFICATIONS OF DDT

Glotfelty and Caro (95) commented that "misidentifications of chlorinated hydrocarbon insecticides resulted from interference by pigment-related natural products found in photosynthetic tissues."

Sims, plant pathologist and biochemist at University of California, discovered that certain red algae produce halogen compounds that had been misidentified as DDT metabolites. He said, "the natural halogens were unknown previously but are now known to be common in natural marine habitats. Compounds containing bromine or iodine rather than chlorine may also have registered as DDT on the gas chromatograph" (96).

Frazier et al. analyzed 34 soil samples that had been sealed in glass jars since they were collected in 1911 (97). The gas chromatograph indicated that five kinds of chlorinated hydrocarbon insecticides were in that soil, even though none of them had been in existence until 30 years after those samples were sealed. They concluded that "the apparent insecticides were actually misidentifications caused by the presence of co-extracted naturally-occurring soil components."

Coon confirmed that material he and his colleagues at the WARF Institute earlier believed to contain DDT residues actually had none at all (98). He also reported on a gibbon that was collected in Burma in 1935 and sealed in a tight container for 30 years before being analyzed. The analysis indicated DDE in the kidneys, testes, liver and fat tissues, even though DDE did not exist anywhere on earth until 6 years after the gibbon was preserved.

Bowman et al. (99) analyzed soil that was sealed in 1940 (years before DDT) and reported "a naturally-occurring extraneous substance appearing in pre-DDT soil that gave the same chromatographic retention times on GLC as DDT (62 ppm) and DDE (35 ppb)."

Hickey sent a robin that had been collected and stored in formalin in 1938 to the WARF Institute for analysis (99). Although it was collected 7 years before DDT was present, the analysis indicated that the robin contained DDT and DDD! Dozens of other robins were similarly analyzed in Michigan and Wisconsin before it became evident that mercury in the earthworms eaten by those birds actually caused the symptoms that had been attributed to DDT poisoning, including tremors, ataxia, and death. Massive amounts of DDT fed to caged robins failed to make them ill. Even nestling robins were not sickened after they were fed only food containing DDT.

DDT SAVING FORESTS

Despite the absence of any confirmation that wild birds suffered illnesses or mortality caused by DDT, activists succeeded in halting the use of DDT on American elm trees, following which thousands of the stately trees died in five midwestern states because of the Dutch Elm Disease that was transmitted by uncontrolled Bark Beetles. Later, millions of eastern oak trees were killed by gypsy moth larvae, which had previously been quickly controlled by a single spray of DDT. Their caterpillars defoliated more than 9 million acres of forest annually before DDT was available, but 1 pound of DDT per acre eradicated the gypsy moths from New England and all adjacent states. The National Audubon Society wrote "No damage was done to bird life, including nestlings." USDA biologists wrote (100), "without this insecticide, there is little hope of preventing the spread throughout their potential range." Environmentalists fought desperately to halt the use of DDT in the infested forests! Forest destruction accelerated after DDT was banned.

In 1974, a campaign was launched to spray DDT on forests in three northwestern states to halt a great Douglas-fir tussock moth outbreak. Permission to spray DDT was granted by the EPA, after 3 years of ridiculous obstinancy. The DDT spraying was then quickly completed, the pests were eradicated, and thousands of acres of Douglas-fir forests were saved. Analyses of vegetation, soil, birds, mammals, and stream inhabitants revealed no harmful effects caused by DDT. Ignoring all science, however, two local newspapers editorialized regarding the persistence of DDT. The *Vancouver Sun* and the *Lewiston Tribune* both wrote in editorials (101) that "DDT has a half-life of several thousand years." (It had only been in existence for 30 years!) I sent each editor copies of literature citing truthful scientific data, but neither of them responded.

BIOLOGICAL MAGNIFICATION UP FOOD CHAINS

A great effort was made by environmental extremists to convince us that there was "biological magnification of DDT up food chains." The theory was as follows: "Small crustaceans ingest DDT, and when hundreds of them are eaten by a fish, that fish might contain a higher DDT concentration than any one crustacean. If many of those fish are eaten by a duck, it might then harbor a higher concentration of DDT than any one fish. After hawks eat the ducks, they could theoretically contain tremendous amounts of DDT! The major difficulty with this theory is that it simply doesn't work that way, where DDT is involved, because DDT is eliminated in faeces, metabolized in blood and cells, and excreted in the urine.

A typical propagandist diagram shows 0.04 ppm DDT in the zooplankton or tiny crustaceans, compared with 25 ppm in the brain of a hawk at the top of the food chain (an increase of 600 times). However, propagandists said that "DDT levels increased 10 million times." They untruthfully sought to start the food-chain with 3 parts per trillion (ppt) of DDT in the water (*before* it entered any food chain!) They then concluded that "DDT is magnified a million or more times as it passes up the food chains."

DDT adheres to particles in the lake bed, so the mud or sand contains a much higher DDT concentration than the water itself. As Hickey pointed out (102), "biological magnification only begins *after* the DDT gets into some living creature." Transfers before that time are *not* "biological", and are not any part of any food chain!

Another source of misinformation is also obvious. If *wet*-weight samples of crustaceans are compared with *dry*-weight samples of fish muscle, the dry tissues will appear to have higher concentrations of DDT because the samples are not diluted by water. Duck muscles do not contain higher concentrations of DDT than do fish muscles, if the samples are prepared and analyzed similarly (103, p. 2160).

Freed (104) proved that about 85% of the DDT in a fish enters through the gills, and very little through the diet. Davis found that "trout pass all of the blood in their body through the gills every 64 seconds, causing rapid transfer of chemicals from the water into the body tissues (105). After the fish is eaten by a duck, most of the DDT passes through the duck's digestive tract and is eliminated.

Comparison of DDT concentrations at each step up a food chain are difficult, and the approximate levels of DDT at each step in the food chain are difficult to determine!

Animals usually store *lower* levels of DDT than were in their food. Anchovies in Monterey Bay contained no more DDT residues than the algae they fed upon (750 to 1,000 parts per billion). Sea lions were found to contain much lower concentrations of DDT than the fish upon which they fed.

Jefferies and Davis (69) reported that earthworms kept in soil saturated with dieldrin contained a lower concentration of dieldrin than the soil, and thrushes that were fed *only* on those earthworms never accumulated any higher concentration of dieldrin than the individual earthworms had contained.

POSTSCRIPT

On June 10, 1971, Philip Butler's report for the National Academy of Sciences claimed (106), "As much as 25% of the DDT produced to date may have been transferred to the sea." Two other panels of the NAS, headed by Kanwisher (107) and by Harvey (108), later refuted such reports. The *Washington Post* (109), however, expanded upon that untruthful claim, saying "Nearly 25% of all DDT manufactured to date is now in the world's oceans, where it is killing baby fish." Jacques Cousteau must have been confused by it all, for in his testimony before the Senate Commerce Committee on October 18, 1971, he stated, "For example, we now know that 25% of all DDT compounds so far produced are already in the sea, and finally they all will end up in the sea." That false statement was repeated in *U.S. News & World Report*, (110).

So, a false estimate of how much DDT *may* have gotten into the sea (made by the man whose own EPA laboratory at Gulf Breeze, Florida, had shown that DDT and its metabolites break down very rapidly in sea water), was escalated to a solid statement by the *Washington Post* (and alleged to "kill baby fish"). It was then blown completely out of context by Cousteau and the anti-DDT press, warning that **ALL of the DDT and its metabolites ever produced will "finally . . . end up in the sea."** The same sort of misstatements and deliberately false allegations are still being directed against many other chemicals that are of great potential benefit to mankind and the environment.

With scientists who behave like so many of those in this report, how can industry, education, conservation, and public health programs survive in America?

BIBLIOGRAPHY

1. Bruemmer, F. (1971). Massacre on Funk Island. *Animals Magazine*, April, pp. 554–555.
2. Associated Press. (1971). Audubon Society to Poison Gulls. *New York Times and San Francisco Chronicle*, Apr. 13.
3. Graham, F., Jr. (1971). Bye-bye blackbirds? *Audubon* **73**(5): 29–35.
4. Anon. (1967). The birds take over. *Bull. Virginia Department of Agriculture*, May, p. 6.
5. Carson, R. (1962). *Silent Spring*. Houghton Mifflin, Boston, MA.
6. Schweitzer, A. (1949). *Out of My Life and Thought—an Autobiography*. Henry Holt, New York, p. 203.
7. Peterson, R.T. (1963). *Life Nature Library: The Birds*. Time, Inc., New York.
8. National Academy of Sciences. (1970). *The Life Sciences: Recent Progress and Application to Human Affairs The World of Biological Research Requirements for the Future*. Washington, DC, p. 432.
9. Wurster, C.F. (1968). DDT reduces photosynthesis by marine phytoplankton. *Science* March **159**: 1474–1475.
10. Ehrlich, P.R. (1969). Ecocatastrophe. *Ramparts magazine* Sept **8**: 24–28.
11. Dustman, E.H. and Stickel, L.F. (1969). The occurrence and significance of pesticide residues in wild animals. *Ann. NY. Acad Sci.* **160**(1): 162–172.
12. Wurster, C.F. As long as it doesn't kill anybody. *Innovation Magazine*, January **17**: 1–9.
13. Wurster, C. (1969). *Congressional Record—Senate*. May 5, p. S4599.
14. Audubon Magazine. (1942). Audubon magazine's forty-second Christmas bird count. *Audubon*, January **44**: 1–75.
15. Cruikshank, A.D. (Ed.). (1961). Sixty-first Christmas bird count. *Audubon Field Notes* **15**(2): 84–300.
16. Mitchell, R.T., Blagbrough, H.P., and Van Elten, R.C. (1953). The effects of DDT upon the survival and growth of nestling songbirds. *J Wildl. Manage.* **17**: 45–54.
17. Bitman, J., Cecil, H.C., Harris, S.J., and Fries, G.F. (1969). DDT induces a decrease in eggshell calcium. *Nature* **224**: 44–46.
18. Porter, R.D. and Wiemeyer, S.N. (1969). Dieldrin and DDT: Effects on sparrow hawk eggshells and reproduction. *Science* **165**: 199–200.
19. Tucker, R.K. and Haegele, H.A. (1970). Eggshell thinning as influenced by method of DDT exposure. *Bull. Environ. Contam. & Toxicol.* **5**(3): 191–194.
20. Bitman, J., Lillie, R.J., Cecil, H.C., and Fries, G.F. (1971). Effect of DDT on reproductive performance of caged leghorns. *Poultry Sci.* **150**: 657–659.
21. McLean, A.E. and McLean, E.K. (1969). Diet and toxicity. *Br. Medical Bull.* **25**: 278–281.
22. Laws, E.R. (1971). Evidence of antitumorigenic effects of DDT. *Arch. Environ. Health*, Sept. **23**: 181–184.
23. Silinskas, K.C. and Okey, A.B. (1975). Protection by 1,1,1-trichloro-2,2- bis(p-chlorophenyl)ethane (DDT) against mammary tumors and leukemia during prolonged feeding of 7,12-dimethylbenz(a)anthracene to female rats. *J. Natl. Cancer Inst.* **55**(3): p. 653–657.
24. Woodwell, G.M., Wurster, C.F., and Isaacson, P.A. (1967). DDT Residues in an East Coast Estuary: A Case of Biological Concentration of a Persistent Insecticide. *Science* **156**: 821–824.
25. Woodwell, G.M., Craig, P.P., and Johnson, H.A. (1971). DDT in the biosphere: where does it go? *Science* **174**: 1101–1107.
26. Anon. (1972). A defeat for scientific integrity: there is evidence that scientists testifying in DDT hearing failed to report all of the facts. *Business Week*, July 8, p. 60.
27. Wilson, A.J., Forester, J., and Knight, J. (1970). Chemical Assays. *USDI Circular* **335**, August, p. 20.
28. Croker, R.A. and Wilson, A.J. (1965). Kinetics and effects of DDT in a tidal marsh ditch. *Trans. Am. Fish. Soc.* **94**: 152–159.
29. Pearson, T.G. (1934). Birds that cruise the coast and inland waters. *Natl. Geog.*, March, pp. 299–302.
30. Allen, R.P. (1935). Pearson's pelican survey in 1918. *Auk* **52**: 199.
31. Pearson, T.G. (1937). *Adventures in Bird Protection*. Appleton-Century, New York.
32. Gustafson, A.F., Guise, C.H., Hamilton, W.J., and Ries, H. (1949). *Conservation in the United States*. Comstock Publishing Company, Inc. (first published in 1939), Ithaca, NY.
33. van Tets, G.F. (1965) A review of the forms of social communication patterns in the Pelecaniformes. *Ornith. Monogr. No. 2*, Ameri. Ornith. Union, p. 88.
34. Schreiber, R.W. and DeLong, R.L. (1969). Brown pelican status in California. *Audubon Field Notes*. February **23**(1): 57–58.
35. Keith, J.0., Woods, L.A., and Hunt, E.G. (1970). Reproductive failure in brown pelicans on the Pacific Coast. *Transactions of the 35th North American Wildl. Conf.*, March, pp. 56–64.
36. Kennedy, D. and Hessel, J. (1969). *Cry California (summer)* **4**(3): 2–10.
37. Jehl, J.R. (1970). Is 30 million years long enough? *Pacific Discovery*, Jan/Feb, pp. 16–24.
38. D'Itri, F.M. and Trost, P.B. (1970). A paper presented to the *International Conf of Environmental Mercury Contamination*. Ann Arbor, MI, Sept 30.
39. Hazeltine, W. (1972). Disagreements on why brown pelican eggs are thin. *Nature* **239**: 410–411.
40. Edwards, J.G. (1971). Testimony. U.S. Congressional Committee on Agriculture, Washington, DC, March 18. *S.N. 92-A*, pp. 575–594.
41. Hartung, R. (1965). Some effects of oiling on reproduction of ducks. *J. Wildl. Manage.* **29**: 872–874
42. Anon. (1979). Oil on parent's feathers causes embryo mortality, *NWF Conservation News*, Oct. 15, 1979.
43. Dieter, M.P. (1977). Oil on eggshells kills 76 to 98% of the embryos. *Report of Interagency Energy Environment Research & Development*. pp. 35–42.
44. King, K.A. (1979). Oil causes pelican mortality for six weeks after spill. *Bull. Environ. Contam. Toxicol.* **23**: 800–805.
45. Szaro, R.C. (1977). Oil causes kidney nephrosis, pancreatic degeneration, and lipid pneumonia. *Proc. 42nd North American Wildlife and Natural Resources Conference* **375**: 375–376.
46. Dawson, W.L. (1923). *Birds of California*. South Moulton Co, Los Angeles, CA.
47. Jehl, J.R. (1970). *Proc. 35th North American Wildl Conf.*, pp. 56–64.

48. Main, R.E. (1970–1973). *Annual nesting surveys for Channel Island National Monument*, U.S. National Park Service.
49. Edwards, J.G. (1982). Brown pelicans revisited. *Nature Society News*, March 15, p. 24.
50. Hickey, J.J. (1943). *A Guide to Bird Watching*. Oxford University Press, London.
51. Risebrough, R.W. and Spitzer, R.R. (1971). Controversy continues over PCBs. *Chem. & Eng. News*, December 12, pp. 32–33.
52. Anderson, D.W. and Hickey, J.J. (1972). Eggshell changes in certain North American birds. Proc. of the 15th International Ornithological Congress. E.J. Brill, Leiden, pp. 515–540. See also, Hawk Mountain Sanctuary Newsletter, 1971.
53. Henny, C. (1976). Update: Ospreys. *Audubon*. March **78**: 130–131.
54. *Hawk Mountain Sanctuary Newsletters*. Hawk Mountain, PA, 1946 through 1972.
55. van Name, W.G. (1921). Threatened extinction of the bald eagle. *Ecology* **2**: 76–78.
56. Anon. (1930). The bald eagle bill. *Bird Lore Magazine* **32**: 164–167.
57. Bent, A.C. (1937). Raptorial birds of America. *U.S. Natl. Mus. Bull.* **167**: 321–349.
58. Thone, F. (1943). Ornithology, nature ramblings: vanishing American. *Sci. News Lett.*, July 3.
59. Anon. (1965). *America's Eagle Heritage*. USDI, Bureau of Sport Fisheries and Wildlife.
60. Coon, N.C., Locke, L.N., Cromartie, E., and Reichel, W.L. (1970). Causes of bald eagle mortality. *J Wildl. Dis.* **6**: 72–76.
61. Belisle, A.A. et al. (1972). Residues in fish, wilflife, and estuaries: residues of organochlorine pesticides, polychlorinated biphenyls, and mercury and autopsy data for bald eagles, 1969 and 1970. *Pest. Monitoring J.* **6**: 133–138.
62. Nye, P. (1992). The Adirondacks: Return of a native. *Nat. Hist.*, May, pp. 54–55.
63. Postupalski, S. (1971). Canadian Wildlife Service manuscript, April 8.
64. Stickel, L. (1966). Bald eagle pesticides relations. *Trans. 31st North American Wildl. Conf.*, pp. 190–200.
65. Krantz, W.C. et al. (1970). Organochlorine and heavy metal residues in bald eagle eggs. *Pest. Monitoring J.* **4**(3): 136–148.
66. Tucker, R.K. (1971). Other pollutants may cause thin egg shells. *Utah Sci.*, June, pp. 47–49.
67. Romanoff, A.L. and Romanoff, A.J. (1949). *The Avian Egg*. John Wiley & Sons, New York.
68. Romanoff, A.L. (1967). *The Avian Embryo*. John Wiley & Sons, New York.
69. Jefferies, D.J. and Davis, B.N.K. (1968). Dynamics of Dieldrin in soil, earthworms, and song thrushes. *J. Wildl. Manage.* **32**: 441–456.
70. DeWitt, J.B. (1956). Chronic toxicity to quail and pheasants of some chlorinated insecticides. *J Agric. Food Chem.* **4**: 863–866.
71. Scott, M.L. et al. (1975). Effects of PCBs, DDT, and mercury compounds upon egg production, hatchability and shell quality in chickens and Japanese quail. *Poultry Sci.* **54**(2): 350–368.
72. Hickey, J.J. and Anderson, D.W. (1968). Chlorinated hydrocarbons and eggshell changes in raptorial and fish-eating birds. *Science* **162**: 271–273.
73. Claus, G. and Bolander, K. (1977). *Ecological Sanity*. David McKay Co, New York.
74. Morris, T.R., Fox, S., and Jennings, R.C. (1964). The response of laying pullets to abrupt changes in daylength. *Br. Poultry. Sci.* **5**: 133–147.
75. Peakall, D.B. (1970). p,p'-DDT: effect on calcium metabolism and concentration of estradiol in the blood. *Science* **168**(931): 592–594.
76. Hauser, E.J. (1963). Planned light bulb maintenance. *Pac. Poultryman*, August. **69**(8): 14–16.
77. Wolff, E.W. and Peel, D.A. (1985). The record of global pollution in polar snow and ice. *Nature* **313**: 535–540.
78. George, J.L. and Frear, D.E.H. (1966). Pesticides in the Antarctic. *J. Appl. Ecol.* **3**(Suppl): 155–167.
79. Peterle, T.J. (1969). DDT in Antarctic snow. *Nature* **224**: 620.
80. Risebrough, R.W., Florant, G.L., and Berger, D.D. (1970). Organochlorine Pollutants in Peregrines and Merlins Migrating through Wisconsin. *Can. Field. Naturalist* **84**: 247–253.
81. Curley, A. et al. (1971). Polychlorinated biphenyls: Distribution and storage in body fluids and tissues of Sherman rats. *Environ. Res.* **4**: 481–495.
82. Risebrough, R.W. and Spitzer, R.R. (1971). Controversy continues over PCBs. *Chem. & Eng. News*, December 12, pp. 32–33.
83. Boyle, R.H. (1978). Update: getting rid of PCBs. *Audubon* **80**(6): 150–151.
84. Scott, M.L. et al. (1975). Effects of PCBs, DDT, and mercury compounds upon egg production, hatchability and shell quality in chickens and Japanese quail. *Poultry Sci.* **54**(2): 350–368.
85. Lichtenstein, E.P., Schulz, K.R., Fuhremann, T.W., and Liang, T.T. (1969). Biological interaction between plasticizers and insecticides. *J. Econ. Ent.* **62**(4): 761–765.
86. Harvey, G.R., Steinhauer, W.G., and Teal, J.M. (1973). Polychlorobiphenyls in North Atlantic ocean water. *Science* **180**: 643–644.
87. Harvey, G.R. (1974). DDT and PCB in the Atlantic. *Oceanus* **18**(1): 19–23.
88. Giam, C.S. et al. (1973). Chlorinated hydrocarbons in plankton from the Gulf of Mexico and Northern Caribbean. *Bull. Environ. Contam. Toxicol*. June **9**(6): 376–382.
89. Hom, W., Risebrough, R.W., Soutar, A., and Young, D.R. (1974). Deposition of DDE and polychlorinated biphenyls in dated sediments of the Santa Barbara basin. *Science* **184**: 1197–1199.
90. Anderson, D.W. et al. (1969). Significance of chlorinated hydrocarbon residues to breeding pelicans and cormorants. *Can. Field. Naturalist* **83**: 91–112.
91. Sherman, R.W. (1973). "Artifacts" and "mimics" of DDT and other organochlorine insecticides. *J. N. Y. Entomol. Soc.* **81**: 152–163.
92. Sweeney, E.M. (1972). Decision of the EPA hearing examiner, Judge Edmund Sweeney, April 26, *Congressional Record*, 40 CFR 164.32, 113 pp. Summarized in: Thomas, D.L. (1972). War on pesticides: DDT was number one on the casualty list. *Barron's*, May 1.
93. Anon. (1972). EPA chief did not read all evidence. *Santa Ana Register*, July 23, p. 1.
94. Edwards, J.G. (1972). The infamous Ruckelshaus DDT decision. *Congressional Record*. July 24, 1972, pp. S11545–47 (via Barry Goldwater).

95. Glotfelty, D.E. and Caro, J.H. (1970). Artifacts of dieldrin in gas chromatographic analysis of green plant material. *Anal. Chem.* **42**(2): 282–284.
96. Sims, J.J. (1977). Press Release June 15.
97. Frazier, B.E., Chesters, G., and Lee, G.B. (1970). “Apparent” organochlorine insecticide contents of soils sampled in 1910. *Pest. Monitoring J.* **4**(2): 67–70.
98. Coon, F.B. (1966). Electron capture gas chromatographic analysis of selected samples of authentic pre-DDT origin. Unpublished paper presented at New York City American Chemical Society Meeting, September, New York.
99. Bowman, M.C., Young, H.C., and Barthel, W.F. (1965). Minimal concentrations of aldrin, dieldrin, and heptachlor in soil for control of white-fringed beetles determined by parallel gas chromatographic and biological assays. *J. Econ. Entomol.* **58**: 896–902.
100. Beroza, M. and Knipling, E.F. (1972). Gypsy moth control with the sex attractant pheromone. *Science* **177**(July): 19–27.
101. *Vancouver Sun and the Lewiston Tribune*. (1974). December 12.
102. Hickey, J.J., Keith, J.A., and Coon, F.B. (1966). An exploration of pesticides in a Lake Michigan ecosystem. *J. Appl. Ecol.* **3**(Suppl): 141–154.
103. Abou-Donia, M.B. and Menzel, D.B. (1968) The metabolism *in vivo* of 1,1,1-trichloro-2,2-bis(p-chlorophenyl)ethane (DDT), 1,1-dichloro-2,2-bis(p-chlorophenyl)ethane (DDD) and 1,1-dichloro-2,2-bis(p-chlorophenyl)ethylene (DDE) in the chick by embryonic injection and dietary ingestion. *Biochem. Pharmacol.* **17**(10): 2143–2161.
104. Freed, V.H. (1968). *Environmental Health Science Center Progress Report*. Oregon State University, Corvallis, OR.
105. Davis, J.C. (1970). *J. Fish Res. Bd. Can.* **27**: 207–214.
106. Butler, P. (1971). *NAS Report*. June 10.
107. Kanwisher, J.W. (1973). DDT in the marine environment. *National Academy of Sciences Committee Report*. January 30, 1973, p. 9.
108. Harvey, G.R. (1973). DDT in the marine environment. *National Academy of Sciences Committee Report*. August 9, 1973, p. 8.
109. Cohn, V. (1971). DDT perils ocean life, study finds. *The Washington Post*. June 13, 1971.
110. Anon. (1971). Growing worry—dirty oceans and waterways. *U.S. News & World Report*. November 1, p. 61.

INSTREAM FLOW METHODS

G.V. Loganathan
Virginia Polytechnic Institute
and State University
Blacksburg, Virginia

Water use is typically classified into offstream use (withdrawals) and instream use. Offstream demands include industrial, commercial, agricultural, and domestic uses. Instream demands include hydropower, recreation, navigation, fish and wildlife habitat management, and assimilation of waste. Unlike offstream demands, which should be satisfied separately, instream demands in general are satisfied simultaneously. That is, the same quantity of water is utilized in instream demands. Therefore, the largest of these instream flow demands must be satisfied. This demand may be in terms of discharge, depth of flow, and velocity. Jowett (1) classifies the methods as *historic flow regime*, *hydraulic*, and *habitat* but points out that the hydraulic methods are not usually used to assess seasonal flow requirements. A key determinant in managing instream flows is the presence of an upstream reservoir. Low flow augmentation is possible with the availability of stored water and efficient management of the reservoir is an important aspect. When a reservoir is not available, the low flows in the stream form the basis for minimum instream flow assessment. The competing offstream demands, public preference, local economy, aesthetics, regulations, water rights, and water quality form the constraints.

When there is an abundance of water, the competing needs can be met easily. When there is only a limited supply, problems emerge. Because river flows are random, it is likely that low flows will occur and proper planning should address the issue. The determination of minimum instream flow may be based on an exceedance criterion such as 90% to guarantee a relatively uninterrupted supply; it may also be based on the nature of the demand such as maintaining a minimum depth or velocity. Hydropower depends on both available flow and hydraulic head. For runoff without storage, the amount of energy that can be guaranteed over 90% (to 97%) of the time is called the firm energy. The required flow is determined as the flow that will be equaled or exceeded 90% of the time. The plot of discharge against percent time equaled or exceeded is called the *flow-duration curve*. If a reservoir is available, a set of *rule curves* is typically followed in making the releases to optimize the available storage subject to the inflow behavior (2). Navigation and recreational boating require maintenance of flow depth based on the barge and boat size. The required flow is to be determined based on the depth requirement.

The procedure given here is in general followed to determine flows that are equaled or exceeded by a certain percentage. A list of moving average values for a fixed averaging duration such as 7 d is generated. Using these moving average values, flows corresponding to various return periods such as 10 yr (90% exceedance) are determined by *frequency analysis* (3,4). The resulting flow is called 7-d 10-yr flow or 7Q10. Virginia Department of Health recommends 1-d 30-yr flow as the characteristic low flow for drought-related studies. For fish habitat, the Tennant or Montana method recommends 30% of annual mean flow as the flow providing required channel flow conditions.

Dissolved oxygen, temperature, water quality, substrate, stream toe width, velocity, flow, and depth are considered key elements in preserving fish habitat. The instream flow incremental methodology (IFIM) utilizes detailed physical habitat simulation (PHABSIM) models. In PHABSIM, for each physical variable of depth, velocity, and substrate a habitat preference curve is defined for a specified species at a specified life stage. For a flow value, the pattern of distribution of depth, velocity, cover, and substrate is analyzed over a stream reach with the aid of a hydraulic model. Using the physical variables, the values from the habitat preference curves are combined into an index called weighted usable area (WUA). For a range of flow values, the WUA values can be obtained. Based on the weighted usable area values, a suitable set of flows can be determined for different levels of species protection and enhancement (4,5). Hardy (6) provides a comprehensive discussion of habitat modeling.

BIBLIOGRAPHY

1. Jowett, I.G. (1997). Instream flow methods: a comparison of approaches. *Regul. Rivers Res. Manage.* **13**(2): 115–127.
2. Mays, L.W. and Tung, Y.K. (1992). *Hydrosystems Engineering and Management*. McGraw Hill, New York.
3. Loganathan, G.V., Kuo, C.Y., and McCormick, T.C. (1985). *Methods of Analyzing Instream Flows, Bulletin 148*. Virginia Water Resources Research Center, Blacksburg, VA. Available: http://www.vwrrc.vt.edu/publications/Bulletin%20148.pdf.
4. McMahon, T.A. (1993). Hydrologic design for water use. In: *Handbook of Hydrology*. D.R. Maidment (Ed.). McGraw Hill, New York.
5. Gordon, N.D., McMahon, T.A., and Finlayson, B.L. (1992). *Stream Hydrology: An Introduction for Ecologists*. John Wiley & Sons, Hoboken, NJ.
6. Hardy, T.B. (1998). The future of habitat modeling and instream flow assessment techniques. *Regul. Rivers Res. Manage.* **14**(5): 405–420.

FLOODPLAIN

Mark R. Tompkins
University of California
Berkeley, California

The term *floodplain* has a range of physically and ecologically based definitions, in addition to a regulatory (i.e., flood control) definition. From the physical perspective, floodplains are a component of the dimensions and geometry of river channels, which as a whole reflect the runoff and sediment load they carry (1). Floodplains have been defined as the relatively flat level floor formed of sediment deposits, which extends from the edge of the channel to the valley walls (2). Nanson and Croke (3) define floodplains more precisely as “the largely horizontally-bedded alluvial landform adjacent to a channel, separated from the channel by banks, and built of sediment transported by the present flow-regime.” They identify lateral point bar accretion, overbank vertical accretion, and braid-channel accretion as the dominant floodplain forming processes in their genetic floodplain classification based on stream power and sediment texture. Figure 1 is a conceptual diagram of a natural river basin that identifies the valley floor, floodplain, river corridor, and channel areas of a basin. Floodplains and riparian areas immediately adjacent to active channels have been documented as very important in sediment retention during high flows (4). Floodplains also act to dissipate bed shear stress and attenuate flood peaks contained in the primary channel. In addition, flood peaks travel downstream faster when a river corridor becomes confined laterally from its floodplain. Floodplains provide important aquatic and terrestrial habitat in addition to significant hydrologic, water quality, and aesthetic benefits (5).

The biological and ecological importance of floodplains has been well documented. Many species of aquatic organisms require floodplains to complete stages of their life history (6,7), including reproduction, rearing, and hibernation. Floodplains also provide important refugia for aquatic organisms during episodes of high flow or vulnerability to predators (8). Aquatic invertebrates and fish seek refuge from high bed shear stresses and

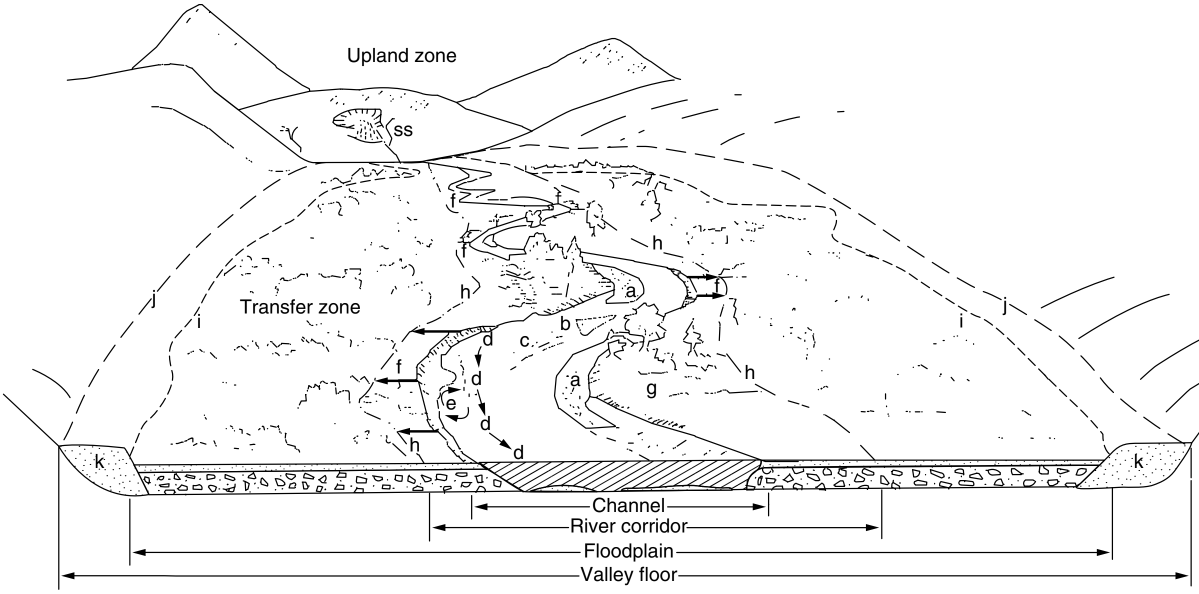

Figure 1. Generalized valley floor, floodplain, river corridor, and channel orientation (after Ref. 5). Letters refer to channel corridor characteristics: (a) point bar, (b) midchannel bar, (c) riffle, (d) pool, (e) secondary current, (f) bank erosion, (g) riparian vegetation, (h) corridor boundary, (i) floodplain boundary, (j) valley floor boundary, (k) terrace, (ss) sediment supply and storage zone.

flow velocities during floods in the hydraulically rough, slow-moving areas beyond the margins of the active channel and on the floodplain. The importance of these floodplain functions has been shown in population growth simulation models that predict increases in biomass of fish and aquatic invertebrates with increases in floodplain width (9). Floodplains have been recognized as critical elements of healthy river ecosystems (10). Without a connected and functioning floodplain, a river corridor ecosystem is incomplete. In their work on the flood pulse concept, Junk et al. (11) describe an aquatic–terrestrial transition zone that moves across the floodplain as flood waters rise. This moving transitional zone is essential to the successful recruitment, establishment, and growth of many species of vegetation, fish, and other organisms in river corridor ecosystems. Historically, floods were often viewed as destructive disturbances to floodplain ecosystems. More recently, however, the prevention of floods has been viewed as the more destructive disturbance to floodplain ecosystems.

The regulatory definition of floodplains originated because valley floors and floodplains have historically been attractive sites for human settlement (12). For thousands of years, humans have built civilizations along river corridors. In fact, the very “cradle of civilization” was built near the confluence of the Tigris and Euphrates Rivers in what is now southern Iraq. Most human development along river corridors has been established on floodplains. Floodplains are naturally flat, close to water, and have fertile soils, making them ideal for agriculture, building sites, and industry location (13). Therefore, a regulatory floodplain is a surface next to a channel that is inundated once during a given return period (14). Regulatory floodplains may or may not be related to physical floodplains. In most cases, some form of flood protection is provided on regulatory floodplains. The administration of regulatory floodplains differs from country to country. In the United States, Congress established the National Flood Insurance Program (NFIP) with the passage of the National Flood Insurance Act of 1968 to identify regulatory floodplains and administer flood protection. The Federal Emergency Management Agency (FEMA) currently administers the NFIP. In return for federally subsidized insurance, communities with access to the NFIP must adopt and enforce a floodplain management ordinance to reduce future flood risks to new construction in Special Flood Hazard Areas (SFHAs). A SFHA is an area within the 100-year floodplain (note this is not a physical reality, but rather a regulatory designation that is discussed in more detail below). Development may take place within a SFHA, provided that development complies with local floodplain management ordinances, which must meet the minimum federal requirements. Communities that participate in the NFIP must implement floodplain management measures in the development of their floodplain lands. FEMA expects these measures to constitute an overall community program of corrective and preventive measures for reducing future flood damage. In general, floodplain management measures address zoning, subdivision, or building requirements for structures in the floodplain and special-purpose floodplain ordinances.

The 100-year flood (and floodplain) was selected by an expert panel as the standard for protection within the NFIP. The 100-year flood (or the 1% annual-chance flood) was chosen on the basis that it provides a higher level of protection (FEMA does not specify what “higher” refers to) while not imposing overly stringent requirements or the burden of excessive costs on property owners. Flood Insurance Studies (FIS), which apply detailed hydrologic and hydraulic analyses, have been performed for many of the watersheds in the United States that are currently included in the NFIP. The results of these analyses include flood insurance rate maps (FIRMs) that identify areas of flood risk and, perhaps most importantly, the boundaries of the regulatory 100-year floodway. The regulatory floodway is defined as the channel of a stream plus any adjacent floodplain areas that must be kept free of encroachment so that the entire base flood (100-year flood) discharge can be conveyed with no greater than a 1.0-foot increase in the base flood elevation.

The references listed in the Bibliography provide a thorough treatment of floodplains for those wishing to develop a deeper understanding of this broad and diverse topic.

BIBLIOGRAPHY

1. Leopold, L.B., Wolman, M.G., and Miller, J.P. (1964). *Fluvial Processes in Geomorphology.* W.H Freeman, San Francisco, CA.
2. Anderson, M., Walling, D.E., and Bates, P. (Eds.). (1996). *Floodplain Processes*. John Wiley, Chichester, UK.
3. Nanson, G.C. and Croke, J.C. (1992). A genetic classification of floodplains. *Geomorphology* **4**: 459–486.
4. Schlosser, I.J. and Karr, J.R. (1981). Riparian vegetation and channel morphology impact on spatial patterns on water quality in agricultural watersheds. *Environ. Manage.* **5**: 233–243.
5. Brookes, A. and Shields, F.D. (1996). Perspectives on river channel restoration. In: *River Channel Restoration—Guiding Principles for Sustainable Projects*. A. Brooks and F.D. Shields (Eds.). John Wiley & Sons, Hoboken, NJ, pp. 1–19.
6. Smock, L.A. et al. (1992). Lotic macroinvertebrate production in three dimensions: channel surface, hyporheic, and floodplain environments. *Ecology* **73**(3): 876–886.
7. Merritt, R.W. and Lawson, D.L. (1992). The role of leaf litter macroinvertebrates in stream–floodplain dynamics. *Hydrobiologia* **248**: 65–78.
8. Townsend, C.R. and Scarsbrook, M.R. (1997). The intermediate disturbance hypothesis, refugia, and biodiversity in streams. *Limnol. Oceanogr.* **42**(5): 938–949.
9. Power, M.E., Parker, G., Dietrich, W.E., and Sun, A. (1995). How does floodplain width affect floodplain river ecology? A preliminary exploration using simulations. *Geomorphology* **13**: 301–317.
10. Bayley, P.B. (1995). Understanding large river-floodplain ecosystems. *Bioscience* **45**(3): 153–158.
11. Junk, W.J., Bayley, P.B., and Sparks, R.E. (1989). The flood pulse concept in river–floodplain systems. *Can. Spec. Publ. Fish. Aquat. Sci.* **106**: 110–127.

12. Dunne, T. and Leopold, L.B. (1978). *Water in Environmental Planning*. W.H. Freeman, New York.
13. Mount, J.F. (1995). *California Rivers and Streams*. University of California Press, Berkeley, CA.
14. Ward, R. (1978). *Floods: A Geographical Perspective*. John Wiley & Sons, Hoboken, NJ.

FISH PASSAGE FACILITIES

Mark R. Tompkins
University of California
Berkeley, California

A *fish passage facility* provides fish with a watered route around or through a barrier to migration in a fluvial system. Fish passage facilities have been designed for both natural and artificial barriers. Some passage facilities allow fish to propel themselves past the barrier, while others physically transport fish past the barrier. Fish passage facilities must provide passage during periods dictated by migratory requirements without injuring or unduly stressing fish. Fish passage facilities have been designed and constructed in North American, Europe, Russia, Latin American, Africa, Australia, New Zealand, China, and Japan. Depending on the region and type of application, fish passage facilities are also referred to as fishways, fish ladders, fish passes, fish bypasses, fish locks, and fish elevators.

Many fish species (e.g., Chinook salmon, American shad, striped bass) must travel upstream and downstream in river systems to complete their life cycles. The development of infrastructure (e.g., dams, water diversions, culverts at road crossings) on rivers has blocked fish passage on rivers around the world. In many cases, viable populations of some fish species have been threatened by these passage barriers. To mitigate this threat, fisheries professionals have designed and constructed fish passage facilities. Many fish passage facilities have been constructed as retrofits to existing barriers that were designed and built before fish passage was required, or before the technology was fully developed. More recently, some infrastructure on rivers has been designed in a "fish friendly" manner that integrates requirements for safe and effective fish migration in the original design (e.g., "fish-safe" turbines, pumps, spillways). As multispecies fish conservation and restoration increases in importance, fish passage facilities will be required to provide a range of hydraulic conditions (i.e., depths and velocities) throughout the year to provide passage for a range of fish species and life stages.

A variety of fish passage facility designs have been developed (see Table 1) to address a wide range of barriers to fish migration. However, asOdeh and Haro (1) note, all fish passage facilities include five common features: (1) the immediate reach of river or waterway served by the fish passage facility; (2) the entrance to the fish passage facility; (3) the fish passage facility itself; (4) the exit of the fish passage facility; and (5) the reach of river where fish exit the passage facility. An effective fish passage facility must allow fish to navigate each of these components of the facility safely and in a manner that meets the temporal requirements of their migratory movements.

Table 1. Fish Passage Facility Designs

Fish Passage Facility Type	Primary Passage Direction	Mode of Fish Movement	Energy Dissipation
Pool and weir	Upstream	Active	Turbulence in pools
Weir and orifice	Upstream	Active	Turbulence in pools
Denil	Upstream	Active	Partial reversal of flow direction
Vertical slot	Upstream	Active	Partial reversal of flow direction
Fish lock	Upstream	Passive	N/A
Fish lift	Upstream	Passive	N/A
Screen and bypass	Downstream	Active	N/A
"Nature-like"	Upstream/ downstream	Active	Turbulence from "natural bed elements"
Road crossing culverts	Upstream/ downstream	Active	Turbulence from roughness elements added to culvert

Fish passage facilities can facilitate both upstream and downstream passage around barriers in fluvial systems. However, most facilities are designed primarily for passage in one direction or the other. Table 1 lists the most common types of fish passage facilities and summarizes for each the primary passage direction (upstream or downstream), mode of fish movement (active or passive), and mechanism of energy dissipation (e.g., turbulence). Fish passage facilities must be designed to align hydraulics and physical characteristics of the facility with the swimming capacity and behavior of the species of concern. Designers must carefully consider flow velocities, depths, turbulence, dissolved oxygen levels, temperature, noise, light, and even odor in the design of a fish passage facility.

Upstream fish passage facilities must attract migrating fish to a specified location in the river downstream of the barrier with an appropriate flow, and then move the fish actively (i.e., by encouraging them to swim) or passively (i.e., by trapping and transporting them) around the barrier. Upstream fish passage facilities for adult anadromous salmonids (e.g., salmon and trout) and clupeids (e.g., shad and alewives) have a long history in North America and Europe. Upstream fish passage facilities have also been designed for juvenile catadromous fish (e.g., American eel). The most common upstream fish passage facilities are described briefly below.

Pool and Weir. The pool and weir facility consists of a series of pools separated by weirs (Fig. 1) that allow fish to ascend by jumping or swimming from the downstream end at the river to the upstream end at the top of the barrier. Flow energy is dissipated as turbulence in pools.

Weir and Orifice. The weir and orifice facility is similar to the pool and weir except that each dividing wall between

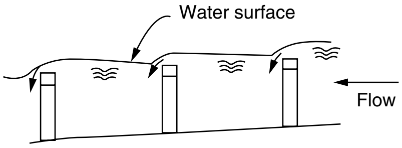

Figure 1. Pool and weir.

pools has a submerged orifice (Fig. 2). Some facilities of this type allow passage from pool to pool through an orifice or over a weir. Again, flow energy is dissipated as turbulence in pools.

Denil. The denil facility consists of a steep flume with vanes along the sides and bottom (Fig. 3). Fish ascend from the downstream end of the barrier to the upstream end by swimming through the center of the flume. The vanes in the denil partially reverse flow direction to dissipate the energy.

Vertical Slot. The vertical slot facility is similar to the pool and orifice type except that the orifice extends for the full height of the baffle between pools (Fig. 4), allowing migrating fish to pass upstream from pool to pool at any depth in the water column. Similar to the denil, the vertical slot dissipates energy by partially reversing the direction of the flow in the facility.

Fish Lock. A fish lock operates in a similar manner as a navigation lock (Fig. 5). Fish enter a chamber at the downstream end of the barrier that is flooded to raise the water surface elevation in the chamber. Fish are then passed upstream through chambers until the water surface elevation difference between the chamber and the upstream end of the barrier is small enough for the fish to pass.

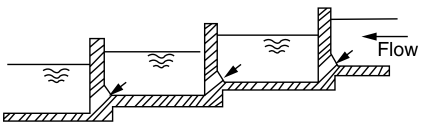

Figure 2. Weir and orifice.

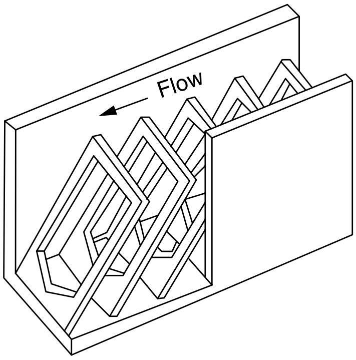

Figure 3. Denil.

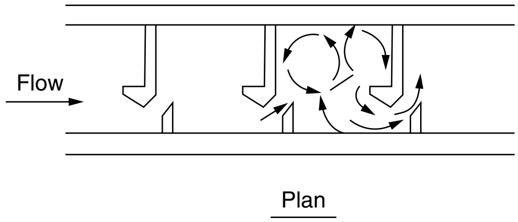

Figure 4. Vertical slot.

Fish Lift. A fish lift can include any mechanical means of transporting fish around a barrier such as tank trucks, large buckets, or other flooded containers (Fig. 6). Migrating fish must be lured into the flooded container and then provided with suitable conditions to maintain their health for the duration of the lift.

Downstream fish passage has received significantly less attention than upstream passage, and therefore the facilities designed to provide downstream passage are less developed. To date, no country has found a satisfactory solution to downstream migration problems, especially at large passage barriers (2). Downstream fish passage facilities have been developed primarily for juvenile anadromous fish. Most downstream fish passage facilities consist of some kind of screen and a bypass. The screen diverts fish away from potential harm associated with the passage barrier (e.g., turbines or agricultural diversions) and the bypass returns fish to the river downstream of the passage barrier. Bypasses must be designed to accommodate the behavior and swimming abilities of the fish species of interest. The most common screens are described briefly below.

Physical Screens. Physical screens (Fig. 7) typically divert downstream migrating fish away from potential harm using screens made of a wide variety of materials (e.g., perforated plates, metal bars, wedgewire, plastic or metal mesh). Screens are normally placed diagonally to flow and must meet maximum approach velocity (i.e., the velocity of flow passing through the screen) and mesh size criteria suitable to prevent impingement of the species of interest. This category of barriers includes drum screens, inclined plane screens, self-cleaning screens, and screens installed perpendicular to flow.

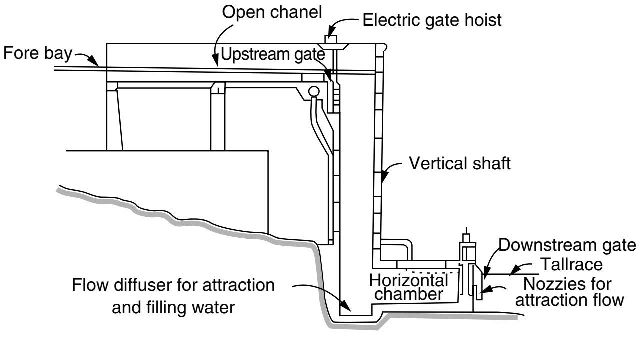

Figure 5. Fish lock.

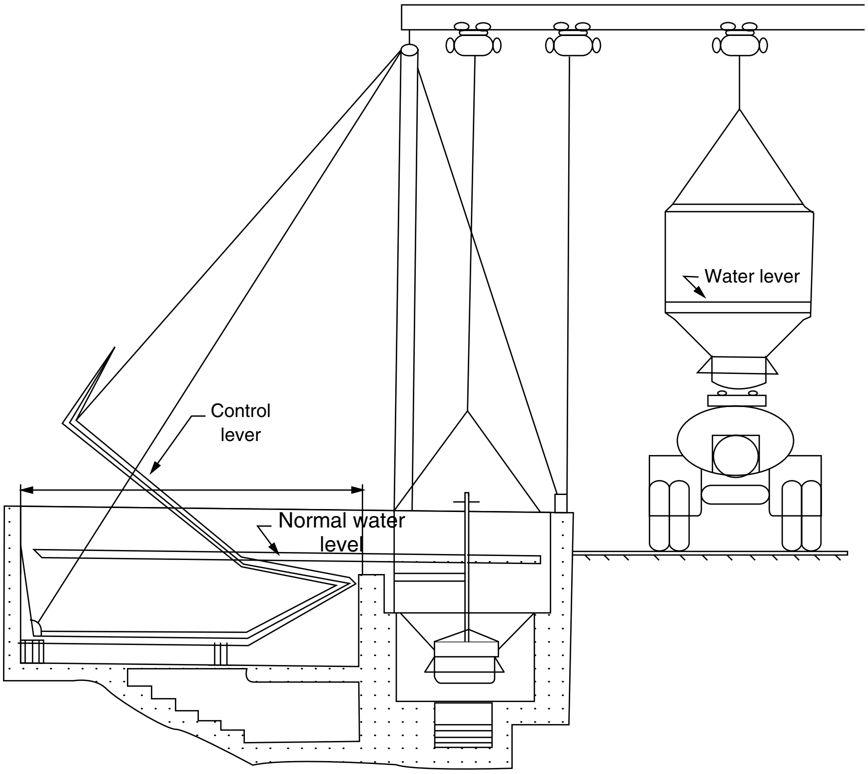

Figure 6. Fish lift.

Behavioral Screens. Behavioral screen facilities for downstream passage have been designed to take advantage of behavioral responses of the fish species of interest to attract or repel fish from a specified location. These approaches have relied on visual (e.g., air bubbles or light arrays), auditory (e.g., audio signal arrays), electrical (e.g., electric currents), and hydrodynamic (e.g., altered velocities and flow patterns) signals to elicit the desired response from the fish species of concern.

Some fish passage facilities provide both upstream and downstream passage to a variety of fish species. These facilities are typically designed for barriers with relatively small water surface elevation differences. Two of these types of facilities are described briefly below.

"Nature Like". Nature-like fish passage facilities incorporate features of natural channels to dissipate energy and provide suitable passage conditions for the fish species of interest (Fig. 8). Specific designs of this type include rock-ramp fishways, ramp fishways with boulders and pools, and boulder block fishways. Each of these approaches achieves upstream and downstream passage in a manner that is more aesthetically compatible with the natural stream environment.

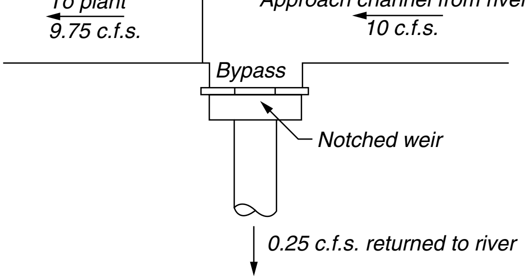

Figure 7. Physical screen.

Road Crossing Culverts. When designed to provide passage for fish under a range of flow conditions, culverts at road crossings can be considered fish passage facilities. Inappropriately designed culverts are often barriers to fish passage. New culvert designs and culvert retrofit designs for fish passage must consider the upstream and downstream migration behavior and the swimming ability of the fish species and life stages of interest. Culverts with baffles, slotted weirs, spoilers, and more natural roughness elements such as cemented boulders have been installed to provide adequate passage conditions.

A significant amount of research and development has occurred in the field of fish passage facilities. References 1–7 provide a good starting point for further investigation of this topic.

BIBLIOGRAPHY

1. Odeh, M. and Haro, A. (2000). Developments in fish passage engineering and technology. In: *Advances in Fish Passage Technology: Engineering Design and Biological Evaluation*. M. Odeh (Ed.). American Fisheries Society, Bethesda, MD, pp. 1–12.
2. EPRI. (1994). *Research Update on Fish Protection Technologies for Water Intakes*. Stone and Webster Eng. Corp. Report.

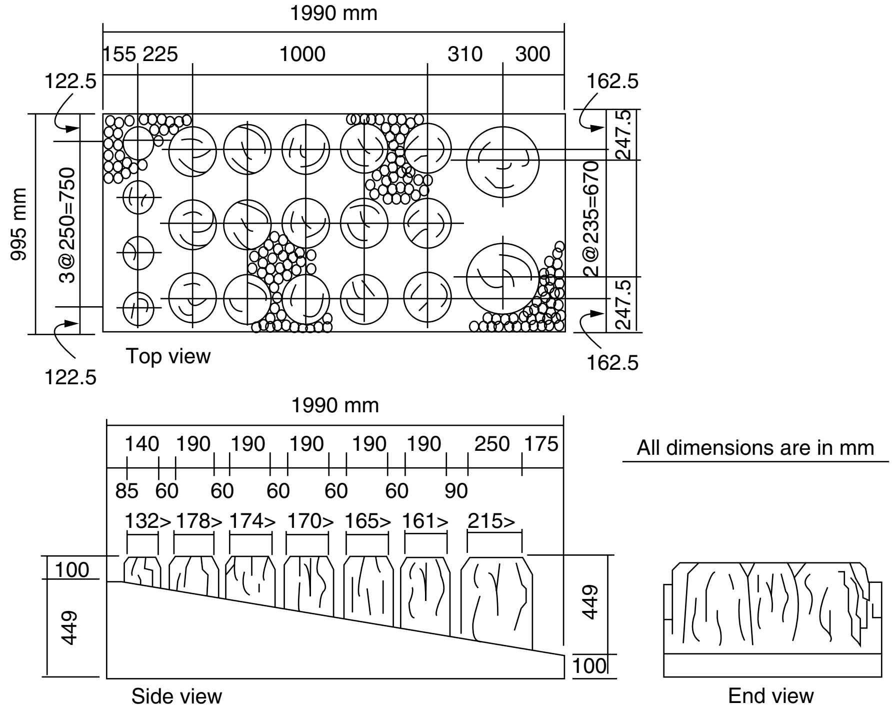

Figure 8. Nature-like fish passage facilities.

3. Bates, K. (1993). *Fish Passage Policy and Technology*. Proceedings of a symposium, K. Bates (compiler). Bioengineering Section of the American Fisheries Society, Portland, OR.
4. Bell, M.C. (1991). *Fisheries Handbook of Engineering Requirements and Biological Criteria*. U.S. Army Corps of Engineers, North Pacific Division, Portland, OR.
5. Clay, C.H. (1995). *Design of Fishways and Other Fish Facilities*, 2nd Edn. Lewis Publishers, Ann Arbor, MI.
6. Jungwirth, M., Schmutz, S., and Weiss, S. (Eds.). (1998). *Fish Migration and Fish Bypasses*. Fishing News Books, Oxford, UK.
7. Odeh, M. (1999). *Innovations in Fish Passage Technology*. American Fisheries Society, Bethesda, MD.

FISHING WATERS

Mark R. Tompkins
University of California
Berkeley, California

The term *fishing waters* refers to all bodies of water where fish are legally captured by humans. Fishing waters can be defined and regulated internationally, nationally, and locally. The four primary categories of fishing that occur in fishing waters are commercial, recreational, subsistence, and indigenous. More than one category of fishing can occur in a given fishing water, and extractive fishing does not have to occur in all fishing waters.

International fishing waters include vast areas of open ocean outside of the zones managed under national fishing regulations. Fishing in international waters is governed by conventions enacted by the United Nations. The most recent conventions dealing with international fishing waters are the Convention on Fishing and Conservation of Living Resources of the High Seas, and the Convention on the Law of the Sea. Nearly all of the fishing in international waters is commercial in nature.

National fishing waters include oceanic zones nearer to coasts and managed by individual nations. Each country manages fishing in these fishing waters differently. In the United States, national fishing waters are governed by The Magnuson–Stevens Fishery Conservation and Management Act. This Act was signed into law in 1976 and covers all fishing activities that occur in federal waters within the United States' 200 nautical mile limit, referred to as the Exclusive Economic Zone (EEZ). The Act also sets forth a system for the management and monitoring of the fish stocks in waters within the United States' EEZ. This system includes eight regional Fishery Management Councils to provide a forum for state, industry, and public participation in the management of American ocean fisheries. The National Marine Fisheries Service (NMFS) is the agency responsible for administering the fishery management requirements of the Act. The Fishery Management Councils develop fishery management plans for each zone that are administered by the NMFS and enforced by the U.S. Coast Guard. Most of the fishing in national waters is commercial in nature.

Local fishing waters can include near shore ocean areas, lakes, reservoirs, rivers, and streams. Again, each country manages local fishing waters differently. In the United States, the U.S. Fish and Wildlife Service together with state Departments of Fish and Game are primarily responsible for the management of fishing in local waters. These agencies are also involved in stocking programs that are used to maintain, enhance, and create fishing waters. In addition, these agencies assist in the management of subsistence and indigenous fisheries. State-issued licenses are typically required to fish in local fishing waters, and most of the fishing in these waters is recreational in nature.

As introduced above, the primary requirement for a body of water to be defined as a fishing water is for some type of fishing to occur in that body of water. Commercial fishing is the most extensive practice that defines a body of water as a fishing water. Commercial fishing occurs in every coastal region in the United States and many different techniques and types of equipment are employed in this endeavor. Any fishing that results in the sale of the captured fish is considered commercial fishing. Recreational fishing is probably the second most extensive practice that defines a body of water as a fishing water. Fishing is considered recreational when fish are captured for pleasure rather than profit. Recreational fishing includes both “catch-and-keep” fishing and “catch-and-release” fishing. Subsistence and indigenous fishing are significantly less extensive practices that define bodies of water as fishing waters. Subsistence fishing provides food or other products for the people who catch the fish, and indigenous fishing is provided for and managed by treaties between the federal government and native peoples (in the United States).

The topic of fishing waters is large and diverse. References 1–4 provide a useful introduction to the topic.

BIBLIOGRAPHY

1. Cochrane, K.L. (Ed.). (2002). *A Fishery Manager’s Guidebook: Management Measures and Their Application*. Food and Agriculture Organization of the United Nations, Rome.
2. Hanna, S., Blough, H., Allen, R., Iudicello, S., Matlock, G., and McCay, B. (2000). *Fishing Grounds: Defining a New Era for American Fisheries Management*. Island Press, Washington, DC.
3. Symes, D. and Phillipson, J. (2001). *Inshore Fisheries Management*. Kluwer Academic Publishers, Boston, MA.
4. Walters, C.J. and Martell, S.J.D. (2004). *Fisheries Ecology and Management*. Princeton University Press, Princeton, NJ.

LAND SURFACE MODELING

XU LIANG
University of California
Berkeley, California

INTRODUCTION

Land surface modeling consists of developing and applying computational models that integrate complicated hydrologic and physical processes that occur at the land surface, between the land surface and the atmosphere, and within the soil column. A land surface model or macroscale hydrologic model, as these types of models are called, has two important characteristics: (a) it considers processes related to both the energy and water budgets and their interactions, and (b) it operates over relatively large spatial domains with short temporal scales. Depending on their complexity, these models generally include representations for some or all of the following major physical processes:

- Snow and snowmelt
- Frozen soil
- Infiltration
- Surface runoff
- Subsurface runoff (e.g., flows from unsaturated and/or saturated zones)
- Evaporation and transpiration
- Vegetation dynamics
- Processes related to the energy budget (e.g., radiation, sensible heat flux, ground heat flux, etc.)

The above processes are interconnected and constrained by the available water and energy. Evapotranspiration links the energy and water aspects of a land surface model together. Soil moisture and surface temperature relate various different physical, hydrological, and biogeochemical processes to one another.

MOTIVATION

One of the main purposes of developing a land surface model is to couple it with (i.e., connect it to) atmospheric or climate models to study the influence of the atmosphere on the land surface, and vice versa, the feedbacks or impacts of the land surface (e.g., the processes listed above) on the atmosphere. Adequate representations of such two-way feedbacks (called land-atmosphere interactions) can improve the accuracy of numerical weather forecasts and climate predictions that play a dominant role in hydrologic forecasting. A land surface model (or a macroscale hydrologic model) is also a physically based tool to assess water resources and hydrologic impacts caused by climatic changes over large river basins or continents. Hydrologic applications at such large spatial scales are impossible without the development of the concept and approach of land surface modeling.

BACKGROUND

We often hear the terms “land surface model” and “hydrologic model.” Do they refer to the same thing? Strictly speaking, traditional hydrologic models do not belong to the category of land surface models, which is because hydrologic models do not describe energy budget-related processes, such as radiation and sensible and ground heat fluxes, although some may include limited energy related calculations such as that related to snow and snowmelt processes. In addition, land surface models have more detailed representations of vegetation-related processes than traditional hydrologic models. Furthermore, traditional hydrologic models mainly focus on the movement of water and its distribution at the land surface and within the vadose zone over small spatial domains with relatively large temporal scales. As atmospheric models deal with physical processes happening over much larger spatial domains at much shorter time steps or temporal scales (e.g., weather phenomena can occur on time scales of minutes), land surface models must also have the capability to deal with hydrologic and physical processes at short temporal scales over large spatial domains for representation of heat and water exchange between the land surface and the atmosphere.

Although land surface models and traditional hydrologic models are developed for different applications, they share some common features as evident by the evolution of these models. The work by Manabe (1) was the first to include land surface impacts in a climate model. Although his land surface model (called the "Bucket" model) includes overly simplified hydrologic processes, a simple energy balance equation, and no vegetation, Manabe's pioneer work ignited many innovative and significant development in the later generation of land surface models.

The evolution of land surface models since Manabe's Bucket model can be characterized by three major phases/stages. The first phase was led by the work of Deardorff (2) (e.g., the force-restore scheme), Dickinson et al. (3,4) (e.g., the BATS model), and Sellers et al. (5,6) (e.g., the SiB model) in which much better vegetation representations and more complex energy budget calculation are considered to realistically represent the energy exchange between the land surface and the atmosphere. To date, a large number of novel models were developed with their roots linked to these three schemes [e.g., (7–15)].

The second phase was led by the work of Liang et al. (16,17) (e.g., the VIC model), Famiglietti and Wood (18), and Schaake et al. (19) in which the effects of subgrid spatial variability on water and energy budgets (e.g., surface runoff, evapotranspiration, etc.) because of heterogeneity of soil properties, topography, vegetation, and precipitation are considered using statistical-dynamical approaches. Also, this phase is marked by improved parameterizations of hydrologic processes (e.g., infiltration, subsurface runoff, etc.) that were overly simplified in the past to provide realistic representations of water and energy exchanges between the land surface and the atmosphere. The oversimplifications of hydrologic processes [e.g., (1)] have been found to lead to significant errors in water and energy budget-related calculations (e.g., soil moisture, runoff, and evaporation) [e.g., (20)]. Crossley et al. (21) and Gedney and Cox (22) also noted that inadequate representations of major hydrologic processes can limit our ability of projecting future climate change. With the advancement over the last decade, major hydrologic processes over the land surface and within the vadose zone are now more adequately represented [e.g., (23–26)] in many land surface models.

Motivated by the need of a more comprehensive representation of the carbon cycle to address climate change issues, the third phase was led by the work of Bonan (27), Sellers et al. (6), Cox et al. (28), and Dai et al. (29) in which vegetation dynamics are included to account for carbon uptake by plants and feedbacks between climate and vegetation. Pitman (30) provides a comprehensive discussion of these models designed for coupling to climate models.

As the land surface impacts hydrologic, energy, and carbon cycles through the interactions among hydrological, physical, and biogeochemical processes, it is essential that these processes be considered or treated in an ***integrated*** way through a unified representation of the atmosphereland system. Land surface models have evolved substantially toward this goal. With the new generation of land surface models, it is now possible to address complex issues and quantify the impacts of surface water hydrology and vegetation on the physical processes that govern atmospheric circulation, and vice versa. For example, Xue (31) showed that degradation of the land surface can have a significant impact on the Sahelian regional climate by increasing surface air temperature and reducing precipitation, runoff, and soil moisture during the summer season. Xue et al. (32) also demonstrated the important role of the land surface for simulating more realistic monthly mean precipitation and flood areas over the United States. Entekhabi et al. (33) further showed the significant impact of soil moisture on numerical forecasting of extreme events (i.e., extreme events of drought and flood). Their study also established the role of land memory as an important source of climate predictability for the United States. Pitman et al. (34) showed the important impact of land cover change on rainfall trends in Australia that are consistent with observations.

LAND SURFACE (MACROSCALE HYDROLOGIC) MODELING FRAMEWORK

In a land surface model, processes associated with the water budget that were identified in the Introduction are connected by the Richards equation. The Richards equation can be written in one-dimensional, two-dimensional, or three-dimensional form. The one-dimensional (1-D) form is more typically used in land surface models. The change of soil moisture content throughout the soil column (i.e., from the soil surface to bedrock) with time can be described by the 1-D Richards equation as

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left(D(\theta)\frac{\partial \theta}{\partial z}\right) - \frac{\partial K(\theta)}{\partial z} - \sin k(z, t) \qquad (1)$$

where θ is the volumetric soil moisture content [L^3/L^3], $D(\theta)$ is the hydraulic diffusivity [L^2/T], $K(\theta)$ is the hydraulic conductivity [L/T], $\text{sink}(z, t)$ is a sink term (e.g., to represent transpiration, E_t, from the root zone), and z represents the vertical direction and is assumed positive downward in Eq. (1). If $q(z, t)$ represents the flux [L/T] across a horizontal plane at depth z and time t, then the boundary condition at the land surface can be expressed as

$$q(z, t)|_{z=0} = P - R - E_o \qquad (2)$$

where P is the precipitation rate, throughfall rate, or snowmelt rate [L/T]; R is the total surface runoff rate (i.e., overland flow) [L/T]; and E_o is the evaporation from bare soil [L/T]. If no flux across the lower boundary exists (i.e., the interface between soil column and bedrock at depth

$z = D$), but the drainage Q_b [L/T], then the lower boundary condition can be expressed as

$$q(z,t)|_{z=D} = Q_b \tag{3}$$

Processes associated with the energy budget as identified in the Introduction are connected to the water budget processes through evapotranspiration. Applying the energy balance equation at the land surface, we have

$$R_n = LE + SH + G \tag{4}$$

where R_n is the net radiation [EL^{-2}T^{-1}], LE is the evapotranspiration [EL^{-2}T^{-1}], SH is the sensible heat flux [EL^{-2}T^{-1}], and G is the ground heat flux [EL^{-2}T^{-1}], all of which can be expressed as a function of the surface temperature, T_s. The net radiation includes net shortwave radiation and net longwave radiation, which are expressed as

$$R_n = (1-\alpha)S_{\text{down}} + L_{\text{down}} - \varepsilon\sigma T_s^4 \tag{5}$$

where α is the surface albedo, S_{down} and L_{down} are the downward shortwave and longwave radiation [EL^{-2}T^{-1}], respectively, ε is the surface emissivity, and σ is the Stefan–Boltzmann constant ($\sigma = 5.67\times10^{-8}$ W K^{-4}m^{-2}). Furthermore, if E represents evapotranspiration [L/T], then E and LE are related as follows:

$$LE = L_e\rho_w E \tag{6}$$

where L_e is the latent heat of vaporization [EM^{-1}] and ρ_w is the density of water [M/L^3]. Hence, the evapotranspiration links the water and energy budgets.

Theoretically speaking, one can obtain the soil moisture θ at different soil depth and the surface temperature T_s at any time t based on Equations (1)– (6), provided the initial soil moisture condition is known, because the processes governing P, R, Q_b, E, SH, and G may be related to either θ or T_s, or both, among other things. Practically, however, it is not feasible to solve for θ and T_s analytically based on Equations (1)– (6) and the initial soil moisture condition unless extreme simplifications are made/introduced.

As a result of the complexity and nonlinearity of the physical processes, the soil moisture θ and surface temperature T_s are usually solved ***numerically*** in a land surface model. The degree of complexity with which the processes are represented physically and/or numerically determines the degree of complexity of a land surface model. For instance, a land surface model can consider either or both infiltration excess runoff and saturation excess runoff in the surface runoff (R) generation mechanisms [e.g., (24)]. Different approaches can also be used to describe evapotranspiration (e.g., Penman–Monteith, supply-demand, mass transfer, α-formulation, etc.), with different degrees of complexity and accuracy.

To numerically solve Equations (1)– (6) (or their variations), land surface models are applied to a study domain (i.e., a study region) that is divided by grids in the horizontal plane (e.g., at the land surface), called model grids, and the vertical soil column is divided into a number of discrete layers. Therefore, each modeling unit has specified lengths in the x- and y-horizontal directions associated to the model grid, and a depth in the vertical or along the soil column. The size of a model grid (i.e., its lengths in the x- and y-directions) determines the spatial resolution at which the land surface model is applied. Finer (coarser) spatial resolutions refer to smaller (larger) sizes of the horizontal modeling grids. The vertical resolution is normally defined by the depth of each vertical layer in the soil column. Higher vertical resolution implies smaller soil depth for each soil layer. Within a horizontal modeling grid, multiple land cover types are allowed. A study area may contain several modeling grids depending on its size. The vertical soil layers include both the saturated and unsaturated zones, separated by the groundwater table.

Most of the current generation land surface models lump the unsaturated and saturated zones together and treat the entire soil column as an unsaturated zone. As an exception, Liang et al. (35) presented a new approach to dynamically represent the movement of the groundwater table within the soil column in a land surface model to consider the impacts of surface water and groundwater interactions on energy and water exchanges. Their results showed the important impacts of such interactions on the partitioning of water budget components, and suggested more attention on issues in this area. Although the idea of investigating the impacts of surface and groundwater interactions on the land-atmosphere system has not yet received much attention, the dynamical representation [e.g., (35)] of groundwater table in land surface models may provide new insights about the water cycle as a whole. For example, is groundwater discharge a significant source of fresh water recycling to the estuaries and the oceans? Can groundwater recharge and discharge processes provide feedback mechanisms that affect climate?

Typical land surface models may, for example, have one vegetation layer, and three vertical soil layers in the soil column (i.e., from the land surface to the bedrock). The three soil layers are normally divided into a top thin soil layer (e.g., 5–10 cm deep), an upper soil layer (e.g., 20–200 cm deep), and a lower soil layer (e.g., from 50 cm to a few meters). The top thin layer is used to capture the quick bare soil evaporation following small summer rainfall events and to facilitate applications using remote sensing information on near-surface soil moisture. The upper soil layer is used to represent the dynamic response of the soil to rainfall events, and the lower layer is used to characterize the seasonal soil moisture behavior.

As a result of the characteristics and intended applications, land surface models typically have large grid sizes (e.g., from a few km^2 to hundreds of km^2 or larger). Therefore, it is very important to consider subgrid spatial variability (i.e., variations at spatial scales finer than the modeling grid) in soil properties, vegetation properties, topography, precipitation, etc., as these variables relate nonlinearly to the physical and hydrologic processes, such as the infiltration, runoff generation, evapotranspiration, etc., that are represented and solved only at the scale of the model grid. This problem is very challenging, and it is not fully addressed yet, although significant progress has been made in recent years.

After discretizing the study domain into modeling grids and vertical soil layers, Equations (1)–(6), together with the initial soil moisture conditions and the representations (or parameterizations) of each relevant physical and hydrologic process (e.g., surface runoff, subsurface runoff, evapotranspiration, etc.), T_s and θ are numerically solved at each model grid and for each soil layer to obtain R, Q_b, E, SH, and G, for example, at each time step t within a study period.

INPUTS AND OUTPUTS

Model Inputs

A typical land surface model requires four types of input information: (a) meteorological data, (b) vegetation data, (c) soil property data, and (d) model parameters.

Meteorological data may include precipitation, air temperature, humidity, surface pressure, wind speed, and downward shortwave and longwave radiation, or a subset of them, when the land surface model is executed offline, i.e., when the land surface model is not coupled with an atmospheric model. In the offline mode, some relationships (both empirical and semitheoretical) can be used to calculate humidity, surface pressure, and radiation, so that only the time series of precipitation, air temperature, and wind speed are needed. If the land surface model is coupled with an atmospheric model, the meteorological information listed above is provided by the atmospheric model instead.

Vegetation data generally include vegetation type, vegetation fractional coverage, leaf area index (LAI) or normalized difference vegetation index (NDVI), root distribution in the soil column (usually difficult to obtain), minimum stomatal resistance, displacement height, and roughness length, and their seasonal dependence. Depending on the complexity of the parameterizations, some models may require more information, such as green leaf fraction, stem and dead-matter area index, and light sensitivity factor, whereas others require less.

Soil property data usually include soil type, soil porosity, wilting point, field capacity, saturated hydraulic conductivity, soil water potential at saturation, soil heat capacity, and soil thermal conductivity. Similar to vegetation data, some models may require more information (e.g., the Clapp–Hornberger B parameter, soil temperature at the model lower boundary, etc.) than others, depending on the level of model complexity.

Most land surface models have some model parameters that cannot be precisely estimated because of observational limits or the absence of direct relationships with available observations. These model parameters are usually estimated based on model calibration (see the discussion on calibration in this encyclopedia) using observed time series, such as streamflow, evapotranspiration, etc. Methods for estimating model parameters through calibration can be found in the literature [e.g., (36–39)].

Model Outputs

A typical land surface model provides the following time series as output:

- snow water equivalent, depth of snowpack, temperature of snowpack;
- surface runoff, subsurface runoff, soil moisture for each soil layer;
- latent, sensible, and ground heat fluxes, and net radiation;
- surface temperature, and soil temperature of each soil layer.

Land surface models that consider surface and groundwater interactions can also provide additional information, such as time series of groundwater table depth, recharge, and discharge. Land surface models that are coupled with a river network routing scheme can also yield time series of streamflow as outputs.

REMARKS

In the past decade, the scientific community has made significant progress in the development of land surface models over three major phases/stages. Land surface models will continue to evolve to include more complete and important processes, such as surface and groundwater interactions, sediment transport, and biogeochemical processes (e.g., for studying carbon and nitrogen exchange), and to improve the representations (or parameterizations) of subgrid spatial variability associated with hydrologic and other physical processes of the integrated atmosphere-vegetation-land-soil system. It is anticipated that with further advancement of land surface models, processes such as floods, droughts, desertification, and the biogeochemical cycles (e.g., carbon, nitrogen, etc.) can be quantitatively and more realistically investigated, and through new discovery, improved understanding of the fundamental laws that control these phenomena may be revealed. Moreover, it is anticipated that new interdisciplinary fields will emerge that integrate atmospheric science, hydrology, public health, ecology, biology, etc. to investigate and predict potential epidemic outbreaks of environmental and/or waterborne diseases that result from integrated effects of natural extremes (e.g., flood and drought), human activity, land cover, land use change, etc.

BIBLIOGRAPHY

1. Manabe, S. (1969). Climate and the ocean circulation: 1, the atmospheric circulation and the hydrology of the Earth's surface. *Monthly Weather Rev.* **97**: 739–805.
2. Deardorff, J.W. (1978). Efficient prediction of ground surface temperature and moisture with inclusion of a layer of vegetation. *J. Geophys. Res.* **83**: 1889–1903.
3. Dickinson, R.E., Henderson-Sellers, A., Kennedy, P.J., and Wilson, M.F. (1986). *Biosphere–Atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model*. NCAR Tech. Note, TN-275+STR, p. 72.
4. Dickinson, R.E., Henderson-Sellers, A., Kennedy, P.J., and Wilson, M.F. (1993). *Biosphere–Atmosphere Transfer Scheme (BATS) version le as coupled to Community Climate Model*. NCAR Tech. Note, TN-387+STR.

5. Sellers, P.J., Mintz, Y., Sud, Y.C., and Dalcher, A. (1986). A simple biosphere model (SiB) for use within general circulation models. *J. Atmosp. Sci.* **43**: 505–531.
6. Sellers, P.J. et al. (1996). A revised land-surface parameterization (SiB2) for atmospheric GCMs. Part 1: model formulation. *J. Clim.* **9**: 676–705.
7. Noilhan, J. and Planton, S. (1989). A simple parameterization of land surface processes for meteorological models. *Mon. Weather Rev.* **117**: 536–549.
8. Xue, Y., Sellers, P.J., Kinter, J.L., and Shukla, J. (1991). A simplified biosphere model for global climate studies. *J. Clim.* **4**: 345–364.
9. Koster, R.D. and Suarez, M.J. (1992). Modeling the landsurface boundary in climate models as a composite of independent vegetation stands. *J. Geophys. Res.* **97**(D3): 2697–2715.
10. Ducoudre, N.I., Laval, K., and Perrier, A. (1993). SECHIBA: a new set of parameterizations of the hydrologic exchanges at the land-atmosphere interface within the LMD atmospheric general circulation model. *J. Clim.* **6**: 248–273.
11. Verseghy, D.L., McFarlane, N.A., and Lazare, M. (1993). CLASS—a Canadian Land Surface Scheme for GCMs II: vegetation model and coupled runs. *Int. J. Climatol.* **13**: 347–370.
12. Viterbo, P. and Beljaars, C.M. (1995). An improved land surface parameterization scheme in the ECMWF model and its validation. *J. Clim.* **8**: 2716–2746.
13. Wetzel, P.J. and Boone, A. (1995). A parameterization for land-atmosphere-cloud exchange (place)—documentation and testing of a detailed process model of the partly cloudy boundary-layer over heterogeneous land. *J. Clim.* **8**(7): 1810–1837.
14. Chen, F. et al. (1996). Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res.* **101**(D3): 7251–7268.
15. Desborough, C.E. and Pitman, A.J. (1998). The BASE land surface model. *Global Planetary Change* **19**: 3–18.
16. Liang, X., Lettenmaier, D.P., Wood, E.F., and Burges, S.J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.* **99**(D7): 14,415–14,428.
17. Liang, X., Lettenmaier, D.P., and Wood, E.F. (1996). A one-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model. *J. Geophys. Res.* **101**(D16): 21,403–21,422.
18. Famiglietti, J.S. and Wood, E.F. (1994). Multiscale modeling of spatially-variable water and energy-balance processes. *Water Resour. Res.* **30**(11): 3061–3078.
19. Schaake, J.C., Koren, V.I., Duan, Q.Y., Mitchell, K., and Chen, F. (1996). Simple water balance model for estimating runoff at different spatial and temporal scales. *J. Geophys. Res.* **101**(D3): 7461–7475.
20. Chen, T.H. et al. (1997). Cabauw experimental results from the project for intercomparison of land-surface parameterization schemes. *J. Clim.* **10**(6): 1194–1215.
21. Crossley, J.F., Polcher, J., Cox, P.M., Gedney, N., and Planton, S. (2000). Uncertainties linked to land-surface processes in climate change simulations. *Climate Dynamics* **16**(12): 949–961.
22. Gedney, N. and Cox, P.M. (2003). The sensitivity of global climate model simulations to the representation of soil moisture heterogeneity. *J. Hydrometeorol.* **4**(6): 1265–1275.
23. Koster, R.D., Suarez, M.J., Ducharne, A., Stieglitz, M., and Kumar, P. (2000). A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *J. Geophys. Res.* **105**(D20): 24,809–24,822.
24. Liang, X. and Xie, Z. (2001). A new surface runoff parameterization with subgrid-scale soil heterogeneity for land surface models. *Advanc. Water Res.* **24**(9-10): 1173–1193.
25. Milly, P.C.D. and Shmakin, A.B. (2002). Global modeling of land water and energy balances. Part I: the land dynamics (LaD) model. *J. Hydrometeorol.* **3**(3): 283–299.
26. Cherkauer, K.A. and Lettenmaier, D.P. (2003). Simulation of spatial variability in snow and frozen soil. *J. Geophys. Res.* **108**(D22): Art. No. 8858.
27. Bonan, G.B. (1995). Land-atmosphere CO_2 exchange simulated by a land surface process model coupled to an atmospheric general circulation model. *J. Geophys. Res.* **100**: 2817–2831.
28. Cox, P.M., Huntingford, C., and Harding, R.J. (1998). A canopy conductance and photosynthesis model for use in a GCM land surface scheme. *J. Hydrol.* **213**(1–4): 79–94.
29. Dai, Y.J. et al. (2003). The common land model. *Bull Amer. Meteorologic. Soc.* **84**(8): 1013–1023.
30. Pitman, A.J. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. *Int. J. Climatol.* **23**(5): 479–510.
31. Xue, Y. (1997). Biosphere feedback on regional climate in tropical north Africa. *Quarterly J. Royal Meteorologic. Soc.* **123**(542): 1483–1515.
32. Xue, Y., Zeng, F.J., Mitchell, K.E., Janjic, Z., and Rogers, E. (2001). The impact of land surface processes on simulations of the US hydrological cycle: A case study of the 1993 flood using the SSiB land surface model in the NCEP Eta regional model. *Monthly Weather Rev.* **129**(12): 2833–2860.
33. Entekhabi, D. et al. (1999). An agenda for land surface hydrology research and a call for the second international hydrological decade. *Bull Am. Meteorol. Soc.* **80**(10): 2043–2058.
34. Pitman, A.J., Narisma, G.T., Pielke, R.A., Sr., and Holbrook, N.J. (2004). Impact of land cover change on the climate of southwest Western Australis. *J. Geophys. Res.* **109**: D18109, doi:10.1029/2003JD004347.
35. Liang, X., Xie, Z., and Huang, M. (2003). A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model. *J. Geophys. Res.* **108**(D16): 8613, doi: 10.1029/2002JD003090.
36. Duan, Q., Sorooshian, S., and Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* **28**(4): 1015–1031.
37. Yapo, P.O., Gupta, H.V., and Sorooshian, S. (1998). Multi-objective global optimization for hydrologic models. *J. Hydrol.* **204**: 83–97.
38. Vrugt, J.A., Gupta, H.V., Bouten, W., and Sorooshian, S. (2003). A shuttled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model. *Water Resour. Res.* **39**(8): 1201, doi:10.1029/2002WR001642.
39. Vrugt, J.A., Gupta, H.V., Bastidas, L.A., Bouten, W., and Sorooshian, S. (2003). Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resour. Res.* **39**(8): 1214, doi:10.1029/2002WR001746.

ANIMAL FARMING OPERATIONS: GROUNDWATER QUALITY ISSUES

Thomas Harter
University of California
Davis, California

Animal farming operations (AFOs) include feedlots, dairies, cattle farms, hog farms, chicken farms, and other facilities that raise or maintain a significant number of animals either seasonally or year-round. With respect to groundwater quality, the fate, reuse, and disposal of animal manure produced at AFOs is of critical concern.

Groundwater benefits a wide variety of uses including production as drinking water for domestic/municipal use, production as irrigation water for agricultural/municipal uses, ecosystem support, and providing base flow in rivers and streams that is vital to the ecological health of these surface waters. AFO related groundwater quality issues result from potentially detrimental impacts during the collection, transport, storage, and land application of animal manure and animal wastes.

Potential groundwater pollutants originating from AFOs are (1) salts and nutrients, (2) bacteria that indicate the presence of pathogens, and (3) organic or inorganic chemicals specifically associated with the management of AFOs. Table 1 lists water quality standards and goals for some of these pollutants, given various groundwater uses. As in other nonanimal farms, any crop production areas that are part of the AFO are also subject to application of pesticides and commercial fertilizers, which may affect groundwater quality.

Animal manure is physically characterized by its weight, volume, moisture content, and solids content. Key chemical properties of manure include nitrogen, phosphorus, and potassium content, 5-day biological oxygen demand (BOD_5), and chemical oxygen demand (COD). Nitrogen content is differentiated by organic nitrogen, ammonium-nitrogen, and nitrate-nitrogen content. Typical values of the chemical composition of animal manure can be found in USDA-NRCS (1) or in the American Society of Agricultural Engineers (ASAE) Standards (X384.2: Manure Production and Characteristics) (2). The specific composition of freshly excreted animal manure varies over a wide range depending on the animal species, the age/production stage of the animal, the diet of the animal, and the effects of weather, season, and animal care. The specific operation of the manure collection, transport, and storage facilities also greatly affect animal manure and waste water composition due to mixing (including mixing with nonmanure wastewaters collected within the AFO facility), separation of solids and liquids, waste digestion, aging, and any incidental or constructed waste treatment processes.

Animal manure is subject to leaching into the subsurface anywhere within the AFO and its land application area, except in concrete-lined or otherwise isolated areas of the operation. This includes unlined feedlots, open corrals, manure solids storage areas, manure and wastewater storage lagoons, and the fields that receive manure from land application in either liquid or solid form or as a sludge amendment.

The leaching of chemicals and pathogens into groundwater is greatly affected by the timing and amount of recharge, which may originate from precipitation, irrigation, application of manure slurry, or from ponding of wastewater (e.g., in storage lagoons). Soil texture, soil permeability, organic matter content, cation exchange capacity, and soil temperature further influence the fate of some of the salts, nutrients, chemicals, and pathogens released from manure excretion and manure land applications.

NITROGEN

Animal manure contains predominantly organic nitrogen, N_{org}, and ammonium nitrogen, NH_4-N. Organic nitrogen is the dominant form of nitrogen in manure solids (particularly if they are separated from the liquid and dried). The liquid portion of animal manure is rich in both organic and ammonia nitrogen.

Organic nitrogen, if applied to land, is highly sorptive and has very limited mobility: significant amounts of organic nitrogen are usually found only in the top 5–30 cm, depending on land management practices. Over time, organic nitrogen will be converted to ammonium ion (NH_4^+), a process referred to as mineralization. The mineralization rate at which organic nitrogen converts to ammonium depends on the manure composition, moisture conditions, and temperature. After application onto soils or incorporation into soils, higher temperature under sufficiently moist conditions leads to rapidly increasing mineralization rates.

Ammonium is subject to two major environmental pathways: volatilization as ammonia gas (NH_3) and microbially facilitated oxidation to nitrate nitrogen, NO_3-N. Ammonia volatilization may account for a significant amount of the nitrogen loss from an AFO and is a major environmental pathway of the total nitrogen excreted in AFOs. High temperature and exposure of large surface areas of manure to wind or air movement (e.g., during the land application process) facilitate volatilization. Once incorporated into the soil, ammonium volatilization is limited due to strong sorption of the positively charged ammonium ion to soil particles. Ammonium mobility in soils is higher than that of organic nitrogen but still sufficiently limited to prevent significant ammonium movement to depths of more than 1–2 m. In sandy soils with limited cation exchange capacity, ammonium mobility can be enhanced if soil/groundwater redox conditions are anoxic.

Under aerated, oxic conditions, microbial oxidation of ammonium in soils is responsible for the conversion of ammonium to nitrate (NO_3^-). This process is referred to as "nitrification." The time frame needed for complete nitrification of soil ammonium varies from hours or days under warm, moist conditions to months under near freezing conditions.

Table 1. Selected Water Quality Standards and Guidelines for Various Uses

Contaminant	Drinking Water	Livestock Water Supply[c]	Industrial Uses[c]
Nitrate (as N)	10 mg/L[a]	100 mg/L	Petroleum: 2 mg/L
Selenium	0.045 mg/L[a]		
Total coliform bacteria	No more than one coliform-positive sample/month (or no more than 5% coliform-positive for systems that analyze more than 40 samples/month)[a]	<200 to <1,000,000	
Fecal coliform bacteria	0[a]	<1 to <10	
Iron	0.3 mg/L[b]	—	
Manganese	0.05 mg/L[b]	—	
Chloride	250 mg/L[b]	—	
Sulfate	250 mg/L[b]	<250 mg/L to <2,000 mg/L	
Total dissolved solids	500 mg/L[b]	<500 mg/L to <3,000 mg/L	Textile: 150 mg/L Cooling water: 1000 mg/L Paper: 1000 mg/L Chemical: 2500 mg/L

[a] Primary maximum contaminant levels, U.S. Environmental Protection Agency.
[b] Secondary maximum contaminant levels, U.S. Environmental Protection Agency.
[c] Reference 1 (available at http://www.nrcs.usda.gov).

Nitrate, a negatively charged ion, has no sorptive capacity and is highly mobile. It can be readily leached from shallow soils and contaminates groundwater. Under anoxic conditions, however, nitrate is reduced to harmless nitrogen gas ("denitrification"). Denitrification is significant in heavy soils and in fine-grained unsaturated sediments and rocks that are subject to anoxic conditions. Reducing conditions in groundwater also lead to significant denitrification.

Of concern to groundwater is the concentration of nitrate in the recharge to groundwater. The potential nitrate concentration is difficult to estimate because the amount of volatilization and denitrification that occurs after land application of manure is typically not known with a high degree of accuracy. Where denitrification and volatilization in the soil and unsaturated zone above the water table are not significant processes, the potential concentration of nitrate in recharge water can be obtained from estimated annual nitrogen leaching losses and from estimated amounts of annual recharge:

$$\mathrm{NO_3(mg/L)} = \frac{\text{annual N leaching loss (kg/ha)}}{\text{annual recharge (mm)}} \times 100 \times 4.42$$

For example, leaching losses of 50 kg/ha nitrate-nitrogen per year in a location with 250 mm annual recharge potentially leads to a nitrate concentration of 89 mg/L in recharge water. The limit in drinking water set by the U.S. Environmental Protection Agency is 45 mg/L.

PHOSPHORUS

Phosphorus is not a concern with respect to the use of groundwater as drinking or irrigation water. It is only a significant concern, where groundwater discharges into streams or lakes. Total phosphorus levels above 0.05–0.1 mg/L potentially lead to significant eutrophication of these surface waterbodies.

Manure contains significant amounts of phosphorus. However, phosphorus is strongly sorbed to soil particles and subject to plant uptake, when manure is applied to crops. On sandy soils with limited cation exchange capacity and on cracked clay soils, repeated applications of manure may result in some phosphorus leaching to shallow groundwater. If shallow groundwater is collected in tile drains, drain ditches, or if it discharges into nearby streams, elevated phosphorus concentration may become a significant concern.

SALINITY

Solid and liquid manure are rich in total salinity (determined by the amount of total dissolved solids), which is readily leachable through soil and into groundwater. Crops grown on fields that are fertilized with manure may consume a limited amount of that salinity. Salinity leaching out of the root zone into groundwater is not subject to any degradation or sorption.

PATHOGENS

Pathogens are ubiquitous in animal waste, but soils provide a significant buffer against groundwater pollution. In the subsurface, pathogens become entrained in dead-end pores, are collected via filtration onto grain surfaces, or chemically interact with the grain surfaces of the soil/aquifer material. Over time, pathogens die off ("inactivation"). However, where soils and shallow groundwater sediments have large pores (sandy soils, sandy and gravelly aquifer sediments, highly fractured rocks, and finely textured soils with macropores), pathogens travel significant distances in relatively short periods of time. Frequent detections of total coliform concentrations in shallow groundwater underneath AFOs are the result of limited pathogen filtration capacity in the soil or unsaturated zone that overlies the groundwater. Shallow wells, often small domestic supply wells, in the immediate vicinity of the AFO are most likely to be affected.

PHARMACEUTICALS (ANTIBIOTICS AND HORMONES)

Only recently, analytical methods have been established that are sensitive enough to detect these pharmaceutical compounds in water resources. They occur in significant concentrations in the inflow to and discharge from wastewater treatment plants, in surface water, and they are also present in animal manure. Research on the fate of these organic compounds in the subsurface is still in its infancy.

BIBLIOGRAPHY

1. U.S. Department of Agriculture Natural Resource Conservation Service. (1996). *Agricultural Waste Management Field Handbook*. U.S.D.A., Washington, DC.
2. American Society of Agricultural Engineers. (2004). Standard X384.2: Manure Production and Characteristics. ASAE, St. Joseph, MI (also available on the web: http://www.asae.org).

AQUACULTURE TECHNOLOGY FOR PRODUCERS

Shulin Chen
Bill Johnson
Brit Johnson
Qiang Pan
Washington State University
Pullman, Washington

INTRODUCTION

Aquaculture is the fastest growing sector of agriculture in many parts of the world because of the healthy nature of seafood, the increasing demands on protein worldwide, and the reduction of ocean catch. According to FAO statistics (1), aquaculture's contribution to global supplies of fish, crustaceans, and molluscs continues to grow, increasing from 3.9% of total production by weight in 1970 to 27.3% in 2000. Worldwide, the aquaculture sector has increased at a rate that exceeds that of any other agriculture sector, with an average compounded rate of 9.2% per year since 1970 compared with only 1.4% for capture fisheries and 2.8% for terrestrial farmed meat production systems. In 2000, the reported total aquaculture production was 45.7 million tons by weight and $56.5 billion by value with 10.1 million tons and $5.6 billion of the value coming from aquatic plants. More than half of the total world aquaculture production in 2000 was finfish, although growth in the other major species groups continues to be rapid with no apparent slowdown in production to date.

In the United States, the National Aquaculture Act of 1980 established a national policy "to encourage the development of aquaculture," which was then followed by the development of the *Aquaculture Research and Development Strategic Plan* by the Joint Subcommittee on Aquaculture. To develop a globally competitive U.S. aquaculture industry, the plan calls for (1) improving the efficiency of U.S. aquaculture production, (2) improving aquaculture production systems, and (3) improving the sustainability and environmental compatibility of aquaculture production.

The major limitations to the development of aquaculture in many regions are availability of water resources and environmental impact. Water is critical to aquatic animals, for it both serves as the medium that supplies the required oxygen and creates the physical environment that allows for movement of the animals. As world population has grown and industries have increased their production to meet the demands of the growing population, water has become a scarce resource that has had many conflicting demands placed upon it. Consequently, the major strategy for increasing aquaculture production with this limited water resource is production intensification. As fish biomass density increases, however, the amount of oxygen dissolved naturally in water becomes limited, thus necessitating supplemental aeration and oxygenation. In addition, the increased use of artificial feed as a means to developing a high density culture system leads to water pollution that stems from the facts that aquatic animals cannot completely utilize the feed and that the animals will excrete other metabolic products such as ammonia that are toxic to the aquatic animals at elevated levels. Internally, the waste produced deteriorates the water quality and makes the water unsuitable for fish to live in, while externally, discharge of the polluted water, if not properly managed, may lead to adverse effects on the environment.

Water uses and environmental impacts of different aquaculture operations can vary significantly because of the extremely diverse natures of the aquaculture systems and their associated technologies. In contrast to terrestrial farming systems, where the bulk of global production is based on a limited number of animal and plant species, more than 210 different farmed aquatic animal and plant species were reported in 2000 (1). This great diversity reflects the large number of aquatic species that are readily adaptable to the wide range of production systems and conditions present in the different countries and regions of the world. Here, we limit our scope to major freshwater systems for finfish production, namely, ponds, raceways, recirculating systems, and cage culture, as these system are predominant in production aquaculture.

POND SYSTEMS

Description

Ponds account for the majority of aquaculture production worldwide (2). In the simplest terms, a pond is an earthen or concrete impoundment that holds water. Water is generally not passed through the system and therefore water is normally only required to fill the pond and replace that which is lost to evaporation and seepage (2). Pond culture is practiced worldwide with it being the predominant system employed in Asia (3), where it totals over 287,400 tons of fish processed. Pond production of channel catfish is the largest single aquaculture enterprise in the United States (4).

Regardless of the species cultured, ponds can be classified by construction method, by location, and by use (2). Typical construction classifications are levee ponds, watershed ponds, hybrid ponds, marsh ponds, beach ponds, and intertidal ponds. Levee ponds are constructed by removing the soil from the area that will become the pond bottom and require a source of pumped water to fill and maintain the pond (5). Watershed ponds are built in hilly areas through the damming of a watercourse and obtain their water from the surrounding watershed. Hybrid ponds usually consist of two or three levees across a shallow drainage basin, which provides the majority of water for the pond, but they also require a source of pumped water to meet their needs (5). Marsh ponds in coastal areas are constructed with levees above ground level. Water salinity ranges from fresh to saline and filling and draining are usually done by pumping, although some marsh ponds have weirs to allow tidal waters to move in and out (2). Beach ponds are more prominent in Asia and are constructed on the shore in sandy locations where seepage is often a problem. Concrete or plastic may be used to avoid seepage problems and water ranging in salinity from brackish to saline is usually pumped with some ponds having weirs to allow for tidal flushing (2). Intertidal ponds are located just off shore in a location where a cove or bay may be blocked off with either a levee or net. Levees are often constructed out of coral rubble or rocks. The levee is designed so as to confine the cultured species while allowing for water exchange through tidal fluctuation (2). Examples of use classifications are broodstock pond, fingerling pond, and food-fish production pond. Each of the use classifications may be employed with any of the aforementioned construction classifications.

Pond Effluent and Its Constituents

Water is discharged from ponds for several reasons. First, unintentional discharge may occur due to rain if inadequate storage capacity is provided. Second, ponds may be drained either completely or partially during harvesting activities. Lastly, water may be discharged during flushing of the pond in an attempt to improve water quality or as a disease treatment (5). Effluent volume is a factor of pond use as pond drainings vary with use. In the U.S. channel catfish industry, broodstock ponds are drained every 1–5 years while fingerling ponds are drained annually and food-fish production ponds vary between 1 and 20 years (5).

Constituents of aquaculture ponds that may affect the receiving water body include solids, organic matter, nutrients, pH, dissolved oxygen, and toxic chemicals (5). Suspended solids (mainly phytoplankton cells) may be deposited in the receiving water and possibly lead to changes in benthic invertebrate and fish populations due to alterations in the biological oxygen demand (BOD). However, the BOD in aquaculture effluents is usually relatively low and is expressed over a longer time period than most wastewaters (6).

Properly managed pond cultures operate within the limits imposed by natural processes within the pond and make it possible to operate them with minimal impact on the environment (5). The fate of nutrients within a pond is of extreme importance to fish culture. In fact, simple mass balances show that with the above noted limited water exchange in ponds, concentrations of waste products would build to levels that could kill fish.

Waste products such as nitrogen and phosphorus generated by fish can be estimated by calculating the quantity of nutrients added to the pond through feed and so on, then subtracting the amount recovered when the fish are harvested. Unfortunately, organic matter is more difficult to calculate. Carbon may be retained by the fish, respired as carbon dioxide, or lost in waste products. In 1988, empirical values for organic matter generation in fecal solids were determined by Henderson and Bromage (7) to be between 0.3 and 0.7 kg/kg of feed. Two to four times more organic matter is produced within a pond due to phytoplankton growth from inorganic nutrients in fish waste than would be calculated solely as fish waste (8).

Luckily, natural biological, chemical, and physical processes known in limnology and waste water engineering take place within the pond, which remove much of the waste products from the water (5). In fact, after several years of production, catfish ponds average 6 mg/L total nitrogen, 0.5 mg/L total phosphorus, and 75 mg/L organic matter, which is 25%, 10%, and 10%, respectively of the expected values calculated by a mass balance over just a single year (5). Nutrient and organic levels never reach the levels predicted by simple mass balance calculations due to the following processes: nitrogen is lost from the water as a gas through denitrification and ammonia volatilization and as particulate organic matter settles to the bottom and decomposes; phosphorus is lost to pond soils as precipitates of calcium phosphate and particulate organic phosphorus; and additionally, microbial processes are continuously oxidizing organic matter within the pond (5).

Levels of dissolved oxygen, pH, or toxic chemicals within aquaculture pond water have been shown to have little effect on the receiving water body (5). In a stream study, dissolved oxygen concentrations were found to be higher downstream from catfish ponds compared to upstream levels (9). Pond water pH fluctuations beyond equilibrium levels of 6.5–8.5 are due to biological activity of phytoplankton, which usually does not coincide with discharge (10). Equilibrium pH will be quickly reestablished due to dilution, reactions with the receiving waters buffer system, and/or reaeration that will equalize dissolved carbon dioxide concentrations and moderate the pH (5). Pesticides and therapeutants used in pond culture must be EPA-approved chemicals and are used infrequently due to the expense of treating large volumes of water. Additionally, long residence times and intense microbial activity within the pond allow chemicals to be dissipated prior to discharge (5).

Pond Operation and Management

It is expected that regulatory programs in several nations will be implemented based on best management practices (BMPs) (11). Effluent management within the context of pond aquaculture poses a difficult challenge for monitoring and treatment due to infrequent discharges and high volumes of a dilute nature. Recommendations on BMPs that are being made to the pond-raised catfish industry (5) include reducing water exchange rates, providing storage within the ponds, using high-quality feeds and efficient feeding practices, managing assimilative capacity within the pond, providing adequate aeration and circulation, positioning mechanical aerators to minimize erosion, allowing solids to settle before discharging water, reusing water that is drained from ponds, and using effluents to irrigate terrestrial crops. In terms of pond construction and renovation practices, recommended practices include optimizing the ratio of watershed to pond area, protecting against erosion, and providing sedimentation basins.

RACEWAY SYSTEM

Description

Flow-through (FT) aquaculture systems can be defined as those production systems where water enters and leaves the system without being recycled for further use. These systems are characterized based on the type of culture unit employed and the pattern of water flow through the system. Culture units that have been used include ponds, tanks, and/or flumes called raceways. The water flow, in turn, is classified as either a single pass system or a serial reuse system in which water flows, respectively, through a single culture unit or several culture units before being discharged.

Flow-through systems are typically energetically passive systems, where water flows from the farm inlet to the outlet via gravity and aeration/reaeration occurs as a result of hydraulic water cascades before and after the culture units. Much information is available concerning FT production system design, construction, and use (12–16). Currently, the most prominent flow-through system is a serial reuse raceway system constructed of concrete.

Flow-through production systems offer several production advantages resulting from the design. Typical raceway culture units possess length and width ratios of 7:1 with depths generally less than 1 m. These dimensions facilitate grading and harvesting operations as well as the easy observation of fish (17). The continuous flow of clean water transports waste metabolites out of the culture units and provides dissolved oxygen needed for life support. In most U.S. raceway operations, there is typically a quiescent zone at the end of the raceway for solids capture. The quiescent zone facilitates the settling of fecal particles as no fish is present to create disturbance. The solids accumulated in the quiescent zone are removed periodically and stored in an offline settling basin.

Aquaculture production using FT raceway systems is common worldwide for many species of fish. In the United States, FT systems are the second most commonly used aquaculture production system (18). Warm, cool, and cold water species such as catfish *Ictalurus* sp., yellow perch *Perca flavescens*, and rainbow trout *Oncorhynchus mykiss*, respectively, are cultured in FT systems. However, coldwater fish production from government hatcheries and the commercial production of salmonids, particularly rainbow trout, are by far the most common uses of flow-through aquaculture systems.

Water Use

Because no water recycling occurs in FT systems, substantial quantities of water are utilized—more than any other type of aquaculture production system. In particular, the culture of coldwater fish, which are naturally more susceptible to poor water quality conditions and require higher dissolved oxygen concentrations than warm and cool water species, requires substantial water resources. However, for all practical purposes, FT system water use is nonconsumptive and does not reduce water availability for other uses.

Depending on farm location and the available water resources, there is a wide range of FT total farm flow rates, which are beneficial in understanding system water use. For example, in the United States, rainbow trout farm total flow rates can range from less than 1699 L/min to 5097×10^2 L/min for the largest farm. On a per unit production basis, typical FT systems are able to produce 6 kg/yr of fish per 1 L/min water flow (19). Based on these values, total water use for FT aquaculture trout production in the United States can be estimated. In the United States approximately 32 million kg of trout were produced in 2000 (20). This corresponds to a nonconsumptive cumulative water flow rate of 5.3×10^6 L/min, which is 15.5% of the consumptive water flow rate required to supply New York City's 9 million people (34.1×10^6 L/min). Clearly, the magnitude of the FT flow rate per production is directly related to the quantity of effluent and has important implications on subsequent FT system environmental impacts and effluent treatments.

Environmental Impact

Environmental regulations controlling FT effluent discharge have been established in many regions (21–23). Potential flow-through effluent pollutants can be grouped into three categories: (1) pathogenic bacteria or parasites, (2) therapeutic chemicals and drugs, and (3) metabolic products and food wastes (17). Regulations primarily target metabolic and food wastes, although recent regulations have begun to address the pathogenic and therapeutant components even though evidence of the environmental impact of the first two is limited (17).

Extensive work has been done to characterize FT effluents and the associated environmental impacts due to metabolic and food wastes. Flow-through effluents contain solid and dissolved organic and nutrient pollutants. Organic pollutants are derived from feces and uneaten feed and are described by biological and chemical oxygen demands, while the nutrient pollutants, nitrogen and phosphorus, are excreted from fish gills and urine. Flow-through effluent characterizations for locations across the world have been summarized in several reviews (17,24). These studies showed that effects from FT effluent discharge are site specific and range from insignificant to heavily impacted, although heavily impacted water bodies are few in the United States (25). The general insignificant impact of FT effluents on receiving waters has been linked to the low pollutant concentrations resulting from high system flow rates. A European study compared FT effluents to domestic wastewater and showed that total suspended solids (TSS), total phosphorus (TP), total nitrogen (TN), and biological oxygen demand (BOD) concentrations were far less, 3% or less, than the same components in medium strength domestic wastewater (26).

Effluent Treatment

The treatment of FT effluents is difficult due to the dilute concentrations, high flow rates, and the high costs of traditional treatment technologies (26). Currently, strategies to treat FT effluent discharge include implementing best management practices for solid waste control (23); optimizing dietary P requirements to reduce excess P excretions (27); and/or utilizing mechanical screen filters, submerged biofilters, or other effluent treatment processes (22,24). These strategies have effectively lowered pollutant discharges from many FT systems. However, the scale-up of the effluent treatment options for applications in large facilities is still problematic with applications of many of the best available technologies to FT effluents for solids and nutrient control being limited by treatment efficiencies and/or inhibitory treatment costs resulting from treatment process operations and/or sizing considerations (28).

NET PEN SYSTEM

Description

Net pens are commonly located in coastal waters although they may also be found in open ocean or freshwater ponds and lakes. Coastal net pens are usually a net bag suspended below a floating rigid top frame. Typical net pens of this type are 10–20 m per side at the surface and 10 or more meters deep. Open ocean net pens are designed with a more rigid structure and can be as large as thousands of cubic meters in volume. Freshwater net pens are generally relatively small rigid structures covered by netting or wire mesh.

Until recently, freshwater net pens were the most common form of aquaculture being utilized throughout the world. However, this culturing technique is no longer the dominant form for freshwater aquaculture, although freshwater net pen culturing is becoming increasingly popular in China. Today, its most dominant application is coastal aquaculture with a trend toward moving further offshore. Marine net pen aquaculture has become progressively more intensive due to new technologies, expansion of suitable sites, improved feeds, increased understanding of the cultured species biology, and increased water quality within farming systems. In some countries, marine net pen culture is the dominant culture method. Atlantic salmon are currently being cultured in net pens and cages in Canada, Chili, Ireland, Japan, Scotland, Norway, and the United States. Other species currently being cultured in net pens throughout the world include sea bream and sea bass in the Mediterranean Sea, yellowtail and red sea bream in Japan, and tuna and cobia culture is growing.

Objections to net pen culture include degradation of the environment, disease transmission, antibiotic use, and interaction of escapees with wild populations. Law suits have been filed for odors, excessive noise, and even visual pollution. Environmental impacts from intensive aquaculture include hypernutrification (eutrophication), benthic enrichment, increased biological oxygen demand, and changes in the bacterial flora. Serious impacts from net pen culture occurred in Japan in the 1990s, resulting in changes in culture practices, and Taiwan is planning for sustainable marine aquaculture as it moves toward cage culture.

Water Quality

Water quality may be affected by net pen culture due to waste metabolites released by the fish and dissolution of nutrients from the feed and feces. Impacts are affected by bottom topography, geography, and hydrographic conditions and tend to be more severe if the location chosen has poor water exchange. Net pen culture has been practiced worldwide in both marine and freshwater with little impact to substantial impact on water quality.

Benthic Environment

The best-studied aspect of marine net pen culture is the deposition of organic matter below the cages or pens. The culturing of animals in pens or cages results in some loss of feed and the production of feces. Uneaten feed and feces may accumulate under or near (usually no more than 30 m from the edge) the net pens or cages. Tidal flow, depth of site, composition, temperature, and salinity of the seawater are among the factors that influence the distribution of solids under each site. Accumulation of uneaten feed and feces can lead to anoxic sediments concurrent with changes in the microbial community and possibly leading to the release of methane, carbon dioxide, and hydrogen sulfide. As mentioned previously, there may or may not be links between water quality and the benthic environment. Therefore, those with more interest in this aspect are directed to Mahnken, who cited a number of studies covering nutrient dynamics and deposition with regard to the benthic community structure.

Methods of Reducing Impacts

As in other forms of aquaculture, methods to reduce the impact of aquaculture have been studied and proposed. Utilization of specifically formulated feeds and improving feeding efficiency have been recommended to reduce the amount of waste produced in both the water and benthic environment. Another approach, which has been successful in Switzerland, is to diaper the cages or net pens. The most common approach to dealing with waste accumulation is the proper placement of farms. This is leading to more farms being sited in open water with better flushing rates and ability to naturally assimilate nutrients.

RECIRCULATING SYSTEM

Description

Recirculating systems are aquaculture systems in which the majority of water is treated or conditioned continuously for repeated use. One of the major advantages of recirculating systems is that a high density of fish biomass can be kept at optimum temperatures. The recirculating system is a relatively new technology that has been developed in the last two decades. There are various designs for recirculating aquaculture systems. An effective recirculating system will usually include a culture tank, an aeration or oxygenation system, a biofilter, and a suspended solids removal unit. The biofilter provides a medium for aerobic bacteria to break down organic fish waste and to convert excreted toxic ammonia into nitrate. Some systems will include additional treatment units such as disinfection.

Water Use

In a recirculating system, the water is processed continuously to guarantee optimum growing conditions. Water is pumped into the tanks through mechanical and biological filtration systems and disinfection systems and then returned into the tanks. Two terms that are often used in relation to recirculating systems are recirculation rate and water exchange rate. Recirculation rate (turnover time) is the amount of water per unit of time pumped out of the tank, through the biofilter, and back into the tank. This can easily be determined by dividing the volume of water in the tank by the capacity of the pump (in gallons per minute). An increase in recirculation rate will increase the number of turnovers per day in the culture tank and thereby provide reduced ammonia levels. Most production recirculation systems are designed to provide at least one complete turnover per hour (24 cycles/day). Exchange rate refers to the amount of water replaced per day. Most recirculating systems recommend 5–10% water exchange rate per day depending on the sensitivity of the species to water quality, stocking density, and feeding rates. The major consumption of water in recirculating systems is associated with the removal of solids such as backwashing the particulate filter or biofilters, while another major consumption derives from the need to replace evaporative loss.

Water Treatment

An excellent culture environment for fish is crucial. In the recirculating systems, the major unit operations to consider for the maintenance of a good environment are ammonia removal/conversion and solids removal.

In respect to solids removal, waste solids, including uneaten feed and by-products of fish metabolism, have to be removed as quickly as possible. Left in the system, the solids could generate additional oxygen demand and produce carbon dioxide and ammonia nitrogen when undergoing bacterial decomposition. Solids removal is generally accomplished by the use of settlement, screen filters, or granular media filter such as bead filters or sand filters.

The most critical nonsolids water quality parameter is ammonia nitrogen. Ammonia is a principal excretory product of fish metabolism and in its un-ionized form is highly toxic to fish. Fish will excrete about 30 g of ammonia for each kilogram of food consumed (at 35% protein content in the feed). The most common technique used for ammonia removal is biofiltration that works on the principle of biological nitrification in which ammonia is oxidized successively to nitrite and finally to nitrate. Nitrate is much less toxic to fish, especially in freshwater. Another process called denitrification is an anaerobic process where nitrate is converted to nitrogen gas. Although not normally employed in commercial aquaculture facilities today, the denitrification process is becoming increasingly important as stocking densities increase and water exchange rates are reduced; resulting in excessive levels of nitrate in the culture system (29). Necessary conditions for nitrification are availability of dissolved oxygen, proper pH, and adequate alkalinity in the water. The rate of the biofiltration process is dependent on temperature, pH, salinity, surface area of the biofilter, and flow rate.

The types of biofilters commonly used in recirculating aquaculture systems include submerged biofilters, trickling biofilters, rotating biological contactors (RBCs), floating bead biofilters, and fluidized-bed biofilters. A common feature to all these biofilters is that they rely on a “fixed film” process in which nitrification bacteria develop on the surface of the support material of the filter and form an active bacterial biomass layer called biofilm. As fish culture water flows through the surface of the film, nutrients such as ammonia and oxygen will diffuse into the biofilm to be utilized by the nitrifying bacteria. The biofilm’s importance also derives from the fact that, as a result of the bacterial attachment to a media, they are free from the risk of being washed out from the system when significant water exchange rate occurs.

Environmental Impact

Recirculation technology, predominantly used in land-based operations, enables farmers to achieve high efficiencies, providing them with greater control over the rearing environment. With the advent of recirculation technology for aquaculture, the incoming source water is reused to varying degrees and undergoes various treatment processes to ensure its purity. The farmer thereby gains considerable control over the rearing environment, being able to manipulate temperature, dissolved oxygen, photoperiod, and water clarity. With more control over water parameters, the stocking densities can be increased, thereby allowing for greater economies of scale. Also, greater control is gained in the separation of uneaten feed and fecal material from the water, thereby preventing these nutrients from entering the downstream environment. Consequently, a recirculating system can significantly reduce effluent volume. However, a recirculating system cannot eliminate the production of fish waste. Thus, as with other animal production systems, a waste management plan is necessary for operation.

BIBLIOGRAPHY

1. Food and Agriculture Organization of the United Nations. (2002). *The World State of Fisheries and Aquaculture*. Available: http://www.fao.org/fi/default_all.asp.
2. Avault, J.W. (1996). *Fundamentals of Aquaculture, A Step-by-Step Guide to Commercial Aquaculture*. AVA Publishing, Baton Rouge, LA.
3. Lin, C.K. and Yi, Y. (2003). Minimizing environmental impacts of freshwater aquaculture and reuse of pond effluents and mud. *Aquaculture* **226**: 57–68.
4. USDA/NASS (United States Department of Agriculture/National Agricultural Statistics Service). (2001). *Catfish and Trout Production*. United States Department of Agriculture, National Agricultural Statistics Service, Washington, DC.
5. Tucker, C.S., Boyd, C.E., and Hargreaves, J.A. (2002). Characterization and management of effluents from warmwater aquaculture ponds. In: *Aquaculture and the Environment in the United States*. J.R. Tomasso (Ed.). U.S. Aquaculture Society, Baton Rouge, LA, pp. 35–75.
6. Boyd, C.E. and Gross, A. (1999). Biochemical oxygen demand in channel catfish *Ictalurus punctatus* pond waters. *J. World Aquaculture Soc.* **30**: 349–356.
7. Henderson, J.P. and Bromage, N.R. (1988). Optimizing the removal of suspended solids in aquaculture effluents in settlement lakes. *Aquacultural Eng.* **7**: 167–181.
8. Boyd, C.E. (1985). Chemical budgets for channel catfish ponds. *Trans. Am. Fisheries Soc.* **114**: 291–298.
9. Boyd, C.E. et al. (2000). Environmental assessment of channel catfish *Ictalurus puntatus* farming in Alabama. *J. World Aquaculture Soc.* **31**: 511–544.
10. Boyd, C.E. and Tucker, C.S. (1998). *Pond Aquaculture Water Quality Management*. Kluwer Academic Publishers, Boston.
11. Boyd, C.E. (2003). Guidelines for aquaculture effluent management at the farm-level. *Aquaculture* **226**: 101–112.
12. Burroughs, R. and Chenoweth, H. (1955). *Evaluation of Three Types of Fish Rearing Ponds*. Research Report 39, U.S. Department of the Interior, Fish and Wildlife Service, Washington, DC.
13. Wheaton, F.W. (1977). *Aquacultural Engineering*. John Wiley & Sons, Hoboken, NJ.
14. Klapsis, A. and Burley, R. (1984). Flow distribution studies in fish rearing tanks. Part 1. *Aquacultural Eng.* **3**: 108–118.
15. Timmons, M.B. and Youngs, W.D. (1991). Considerations on the design of raceways. In: *Aquaculture Systems Engineering*. P. Gioviannini (Ed.). American Society of Agricultural Engineers, St. Joseph, MI.
16. Wedemeyer, G.A. (Ed.) (2001). *Fish Hatchery Management*, 2nd Edn. American Fisheries Society, Bethesda, MD.
17. Hinshaw, J.M. and Fornshell, G. (2002). Effluents from raceways. In: *Aquaculture and the Environment in the United States*. J.R. Tomasso (Ed.). U.S. Aquaculture Society, Baton Rouge, LA.
18. USDA (United States Department of Agriculture). (1998). Census of aquaculture.
19. Timmons, M.B., Ebeling, J.M., Wheaton, F.W., Summerfelt, S.T., and Vinci, B.J. (2002). *Recirculating Aquaculture Systems*. Northeastern Aquaculture Center Publication No. 01–002.
20. United States Trout Farmers Association. (2000). Available: www.USTFA.org.
21. Rosenthal, H. (1994). Fish farm effluents and their control in EC countries: summary of a workshop. *J. Appl. Ichthyol.* **10**: 215–224.
22. Bergheim, A. and Brinker, A. (2003). Effluent treatment for flow through systems and European environmental regulations. *Aquacultural Eng.* **27**: 61–77.
23. MacMillan, J.R., Huddleston, T., Woolley, M., and Fothergill, K. (2003). Best management practice development to minimize environmental impact from large flow-through trout farms. *Aquaculture* **226**: 91–99.
24. Cripps, S.J. and Bergheim, A. (2000). Review: solids management and removal for intensive land-based aquaculture production systems. *Aquacultural Eng.* **22**: 33–56.
25. Sloan, D. (2001). *Verification of the State Listings for the 1998 EPA 303d and 305b: Impaired Waterbodies Due to Aquaculture*. Report to the Joint Subcommittee on Aquaculture, Aquaculture Effluents Task Force. National Association of State Aquaculture Coordinators, Raleigh, NC.
26. Cripps, S.J. (1994). Minimising outputs: treatment. *J. Appl. Ichthyo.* **10**(4): 284–294.
27. Barrows, F.T. and Hardy, R.W. (2001). Nutrition and feeding. In: *Fish Hatchery Management*, 2nd Edn. G.A. Wedemeyer (Ed.). American Fisheries Society, Bethesda, MD.
28. True, B.J. and Chen, S. (2004). Reducing phosphorous discharge from flow-through aquaculture I: facility and effluent characterization. *Aquacultural Eng*.
29. Timmons, M.B. Ebeling, J.M., and Wheaton, F.W. (2002). *Recirculating Aquaculture Systems*, 2nd Edn. Cayuga Aqua Systems, Ithaca, NY.

BIOFUEL ALTERNATIVES TO FOSSIL FUELS

Lucas Reijnders
Universiteit van Amsterdam
Amsterdam, The Netherlands

INTRODUCTION

Our species has been exclusively dependent on biomass for fuel during most of its existence. The use of coal and mineral oil emerged due to the scarcity of biofuels (1). By now, the world energy market, which supplies a yearly energy consumption of about 400×10^{18} joule (J), is dominated by fossil fuels. Still, about 14% of yearly world fuel consumption is based on biofuels (2,3). Most of this consumption is similar to the use before the Industrial Revolution and is located in the countryside of developing countries. It is often associated with poor indoor and outdoor air quality. Negative impacts on soils, on the availability of groundwater, and on water for generating hydroelectricity have also been linked to biofuels in developing countries (4–6).

Against this background, it may be somewhat surprising to see increased advocacy for biofuels. This concerns mainly "modern biomass": biofuels applied while using advanced conversion technology to generate electricity and heat and for automotive power (7).

There are several main drivers behind the increased advocacy for biofuels. First, there are efforts to become less dependent on mineral oil, driven by concerns about energy scarcity, political instability in the Middle East, and oil price shocks. A second driver is environmental concern. Some applications of biofuels result in lower emissions of such pollutants as hydrocarbons, carbon monoxide (CO), sulfur oxides, and nitrogen oxides. It has also been argued that in combusting biofuels, one simply puts back into the atmosphere carbon dioxide (CO_2) that has been recently fixed by photosynthesis, thus ensuring climate neutrality, in contrast to the combustion of fossil fuels which puts CO_2 in the atmosphere that was fixed many millions of years ago. The need to find an outlet for agricultural surpluses has also been conducive to biofuel advocacy.

Advocacy and also actual combustion of "modern biomass" are increasing. A major application is the use of ethanol and to a lesser extent fatty acid methylester as fuel in motorcars. Ethanol is used mainly in Brazil, North America, and the European Union (EU). In Brazil, sugarcane is the major feedstock for ethanol production, and in the EU and North America, the main feedstock for ethanol is starch from grains. The use of biomass for modern electricity and heat production is mainly on the increase in the European Union. Expansion of biomass-based waste incineration, cocombustion of biomass in coal-fired power plants, and energetic use of residual biomass by industry contribute significantly to this trend.

Here the implications of modern biomass use that are relevant for water are discussed. Some of these implications are direct. By clear felling forests for the supply of biomass, there can be a major direct impact on surface water, such as changes in peak flows and runoff. Other implications are indirect. The extent to which modern biomass influences atmospheric levels of greenhouse gases such as CO_2 and CH_4 (methane) is a determinant of climate change and this in turn influences the biogeochemical water cycle. Current climate change leads among other things to sea level rise (that in turn may lead to increasing intrusion of saltwater) and an increase in land surface precipitation of 0.5–1%/decade in mid- and high latitudes of the Northern Hemisphere, whereas there is a decrease of such precipitation in subtropical land areas of 0.3%/decade (8). Due to climate change, there may be changes in water erosion, seasonal river flows, and groundwater tables; impacts on biological activity; and the presence of species in water. Increasing temperatures also increase human freshwater use (9).

In this article, we deal with the following questions:

- — Is the use of biofuels "climate neutral"?
- — What is the impact of biofuels on (fresh) water use?
- — What is the impact of biofuels on substance flows that may impact water?

ARE BIOFUELS "CLIMATE NEUTRAL"?

To answer the question whether biofuels are "climate neutral," a life cycle perspective on biofuel combustion is necessary. Such a perspective includes all stages of the biofuel life cycle. For instance, if trees from forests are used as biofuel, the cycle starts with the seed and ends with the postharvesting fate of the forest. The main matters to consider in establishing the impact of the biofuel life cycle on climate are the input of fossil fuels during the life cycle and the impact of the biofuel life cycle on biogenic carbon stocks. Both influence atmospheric concentrations of greenhouse gases such as CO_2 and CH_4. One should also consider the fate of compounds containing nitrogen (N) in the biofuel life cycle. Such compounds may be partly converted into the potent greenhouse gas N_2O (dinitrogen oxide). This aspect tends to be less important than the impact on atmospheric gases containing carbon, but may still be significant in agricultural production of biofuels.

Because the use of fossil fuels is so dominant in worldwide energy consumption, it is not surprising that usually there is an input of fossil fuels in the biofuel life cycle. Quantitatively, such inputs vary widely depending on the kind of biofuel and its actual production. For instance, in forest-derived biofuels, the input of fossil fuels is usually 5–10% of the energy value represented by the biofuel (output). This gives an "output to input energy ratio" of 10–20, meaning that the biofuel can supply 10–20 times the equivalent of the input of fossil fuels in the biofuel life cycle. In animal meal that is cocombusted with coal in a number of European power plants, the output to input energy ratio is about 0.24 (10). Table 1 gives output to input energy ratios for several biofuels.

This table shows, among other things, that for ethanol from U.S. corn, there is no agreement on the energy ratio. The underlying disagreements result in part from different estimates of the actual fossil fuel inputs in the biofuel life cycle. When corn is used to produce the biofuel ethanol, there are also coproducts, such as cattle feed. In this case, the allocation of fossil fuel input to the different outputs (products) arises. One possibility is to allocate all fossil inputs to the biofuel. Another possibility is to divide the fossil fuel inputs by the products involved. Such an allocation may be done in different ways, for instance, based on a monetary or on an energy basis, and these different ways lead to different outcomes. For ethanol produced from corn in the United States, the U.S. Department of Agriculture (12) has argued that the biofuel energy output to fossil fuel energy input ratio is currently 1.3, whereas Pimentel (13) calculates an output to input energy ratio of 0.58–0.74, depending on the allocation to ethanol and its coproducts.

However, notwithstanding this disagreement, one may conclude that when ethanol is produced from U.S. corn starch, the input of fossil fuels is at least a large share of the energy output from the biofuel. A similar conclusion can, for instance, be drawn regarding oil-methylesters

Table 1. Ratios of Current Energy Output to Fossil Fuel Input for the Life Cycles of a Number of Biofuels[a]

Type of Biofuel and Application	Ratio of Energy Output to Fossil Fuel Input
Forest derived (short transport distances), for electricity production	10–20
Ethanol from Brazilian sugarcane (bagasse as production fuel), transportation fuel	~7
Ethanol enzymatically produced from woody crops, transportation fuel	~4
Chicken manure (energetic allocation) produced in the European Union, for electricity production	~1
Ethanol from U.S. corn, transportation fuel	0.58–1.3
Animal meal produced in the European Union for electricity production	0.24

[a] References 3, 10–13.

produced from oil crops such as rapeseed, sunflower, or soybean (14). More in general, one can conclude that burning biofuels is not only a matter of giving back to the atmosphere carbon that was recently fixed. Usually, there is the added emission of fossil fuel derived carbon compounds, such as CO_2. Especially when high-input agricultural production is involved and when further processing is fueled by fossil fuels, this emission tends to be important.

A second important matter to be considered is the effect of the biofuel life cycle on biogenic carbon stocks: biomass and soil organic carbon (C). The last stock is of great quantitative importance. It is currently about three times the atmospheric stock and about four times the stock of C in biomass. Since the beginning of the Industrial Revolution, the combustion of fossil fuels has added an estimated 270×10^{15} g C to the atmosphere, whereas biogenic sources (deforestation, biomass burning, and loss of soil organic carbon) added $\sim 136 \times 10^{15}$ g C. Depletion of the stock of soil organic C since the start of the Industrial Revolution is an estimated 78×10^{15} g C (15).

The biofuel life cycle can substantially reduce and add to biogenic carbon stocks. If biofuel is produced from deforestation and conversion into agricultural land, large emissions of carbon containing gases from soil to the atmosphere can occur. On the other hand, when biofuel is produced from reforestation or conservation tillage and the use of compost and manure, there can be net sequestration of carbon in the soil (15). In Table 2, this is exemplified by the life cycle emission of 68×10^2 g CO_2/kWh, when the production of forest derived fuel is accompanied by deforestation. The net sequestration of 19×10^2 g CO_2/kWh is from reforestation of agricultural land.

If we survey the relation between biofuel use and the emission of greenhouse gases that affect climate, it will be clear that the relation is more complicated than "giving back" recently fixed carbon dioxide to the atmosphere. As Tables 1 and 2 show, net life cycle additions of greenhouse gases to the atmosphere can be larger than that from combusting an equivalent amount of fossil fuel.

Table 2. Life Cycle CO_2 Emission Factors for Several Current Types of Power Generation in 10^2 g CO_2/kWh[a]

Type of Fuel	Life Cycle Emission Factor in 10^2 g CO_2/kWh
Hard coal	+10
Natural gas	+ 4
Forest derived biofuel	−19–+68

[a] Reference 3.

THE IMPACT OF BIOFUELS ON (FRESH) WATER USE AND WATER AVAILABILITY

The replacement of forests by plantations or annual biofuel crop production may lead to a decrease in available water due to increased drainage and increases in peak flows and storm flows. Soil degradation from clear felling of forests may lead to insufficient replenishment of groundwater reserves and increased peak and storm flows (16,17).

Water is also an important resource in biofuel production. Fast growing woody crops such as poplar and eucalyptus use relatively much water. For instance, Tuskan (18) has calculated that in the United States hybrid poplar requires roughly 1 meter/hectare yearly of (ground) water during midrotation. In southern China, it has been found that under comparable conditions, water tables under eucalyptus plantations were 50 cm lower than under mixed forests (19). Water use for biofuel crops may be expected to increase when temperatures rise (20).

Although there are some waterlogged areas where large uptakes of water by energy crops can be considered a benefit, it will be more common that water can be a limiting factor in biofuel production. Currently, in nearly 80 countries that have 40% of the world population, water demands already exceed additions to stock (21), and such countries are unlikely to embark on major energy crop production programs. In countries that currently have surplus water, if compared with demand, limitations may also emerge. For instance, in Ethiopia where eucalyptus is grown for biofuel production, competition between food crops and eucalyptus for scarce water resources has led to restrictions on the latter (4). Berndes (22) has pointed out that large-scale biofuel production countries such as Poland and South Africa may face absolute water scarcity, resulting in lowering of water tables.

Judicious choice of biofuel crop may limit water use. C_4 biofuel crops generally use water more efficiently than C_3 crops (22). Water efficiency may also be improved by some types of agroforestry, reducing runoff, direct evaporation of water, and drainage (17,23).

BIOFUEL SUBSTANCE FLOWS THAT MAY IMPACT WATER

The biofuel life cycle mobilizes large amounts of substances. When biofuels are produced by agriculture (e.g., ethanol made from starch or sugar), higher inputs of sediments, nutrients, and pesticides in water accompany increased agricultural land use. Such increases may also be large because land requirements may be large. If, for instance, the United States were to decide to cover current U.S. gasoline consumption by corn-based ethanol fuels, more land would be needed than current total U.S. cropland (11).

When biofuels are derived from tree plantations, erosion may be much increased. Studies of plantation forestry on slopes has shown that, especially in the planting and timber harvesting phase, erosion can be extensive, giving rise to silting of lakes and reservoirs and channel instability downstream (24). The sequestration of minerals may be reduced by clear felling, leading to increased levels in surface water (25).

Another matter for concern in tree plantations is the use of plant nutrients to stimulate biomass production. High productivity is strongly linked to large inputs of plant nutrients containing K (potassium), N, and P (phosphorus), and to poor efficiencies in output (nutrients in biofuel) to input ratios (16,26). This may lead to major nutrient losses to water.

The actual impact of biomass production, of course, strongly depends on the measures taken to limit such impacts. Woodlands generating yearly, on average, 3 metric tons of woody biomass per hectare will lose much less nutrients and soil than high productivity plantations (27). In the agricultural production of biofuels, conservation tillage, applying crop residues, terracing, and other provisions to trap sediment, reduce fertilizer application, and minimize hydrologic changes may limit impact.

The combined life cycle emissions of acidifying substances such as NO_x and SO_2 are for biofuels of the same order of magnitude as the corresponding emissions from fossil fuel life cycles (28–30). The generation of persistent organic pollutants such as chlorinated dioxins and benzofurans and polycyclic aromatic hydrocarbons is not less than that from fossil fuels (31,32). The mobilization of heavy metals and other hazardous elements may be substantial. In Denmark, for instance, the application of fly ash originating in the combustion of wood and straw to soil is banned because of its cadmium content (33). Klee and Graedel (34) have compiled a survey of elements in coal and the average concentration in dry plant material. Significant differences emerge from this survey. For a number of hazardous substances, including Hg, As, Pb, Cr, and Ni, concentrations in biomass are lower. On the other hand, there are also elements that may have an environmental impact that have higher concentrations in biomass, such as P, K, Co, and Cd.

The load of hazardous elements generated by biomass combustion may be much increased when such elements are added to biomass during its stay in the economy. Examples are the treatments of wood with chromated copper and arsenic salts or metaborates for conservation. When wastes of such treated woods are burned as biomass especially the level of hazardous elements in ash is elevated. Sewage sludge that is combusted as biomass is another case in point.

Again actual impacts will depend on the measures taken. For instance, state-of-the-art use of desulfurization and de-NO_x technology will strongly limit acidification risk for poorly buffered surface waters. And indefinite containment of hazardous elements in ashes will better protect ground- and surface water than application of ashes to soils or the use of settling ponds for ashes.

BIBLIOGRAPHY

1. Pomeranz, K. (2002). Political economy and ecology on the eve of industrialisation: Europe, China and the global conjuncture. *Am. Hist. Rev.* **107**: 425–446.
2. Minnen, J.G., van Ierland, E.C., and Nabuurs, J. (2003). Terrestrial carbon sinks and biomass. In: *Issues in International Climate Policy*. E.C. van Ierland, J. Gupta, and M.F.J. Kok (Eds.). Edward Elgar, Cheltenham, UK.
3. Reijnders, L. and Huijbregts, M.A.J. (2003). Choices in calculating life cycle emissions of carbon containing gases associated with forest derived biomass. *J. Cleaner Prod.* **11**: 527–532.
4. Jagger, P. and Pender, J. (2003). The role of trees for the sustainable management of less-favored lands: The case of eucalyptus in Ethiopia. *Forest Policy Econ.* **5**: 83–93.
5. Akinbami, J., Salami, A.T., and Siyanbola, W.O. (2003). An integrated strategy for sustainable forest-energy-environment interactions in Nigeria. *J. Environ. Manage.* **69**: 115–128.
6. Ludwig, J. et al. (2003). Domestic combustion of biomass fuels in developing countries: A major source of atmospheric pollutants. *J. Atmos. Chem.* **44**: 23–37.
7. Goldemberg, J. and Coelho, S.T. (2004). Renewable energy-traditional biomass vs modern biomass. *Energy Policy* **32**: 711–714.
8. IPCC (Intergovernmental Panel on Climate Change). (2001). *Climate Change: The Scientific Basis*. Cambridge University Press, Cambridge, UK.
9. Levin, R.B. et al. (2002). U.S. drinking water challenges in the twenty-first century. *Environ. Health Perspect.* **110**: 43–52.
10. Reijnders, L. and Huijbregts, M.A.J. (2005). Life cycle emissions of greenhouse gases associated with burning animal wastes in countries of the European Union. *J. Cleaner Prod.* **13**: 51–56.
11. Kheshgi, H.S., Prince, R.C., and Marland, G. (2000). The potential of biomass fuels in the context of global climate change: Focus on transportation fuel. *Annu. Rev. Energy Environ.* **25**: 199–244.
12. U.S. Department of Agriculture. (2002). Economic Report 814. Washington, DC.
13. Pimentel, D. (2004). Ethanol fuels: Energy balance, economics and environmental impacts are negative. *Nat. Resour. Res.* **12**: 127–134.
14. Venturi, P. and Venturi, G. (2003). Analysis of energy comparison for crops in European agricultural systems. *Biomass and Bioenergy* **25**: 235–255.
15. Lal, R. (2004). Soil carbon sequestration to mitigate climate change. *Geoderma* (in press).
16. Perry, D.A. (1998). The scientific base of forestry. *Annu. Rev. Ecol. Syst.* **29**: 435–466.
17. Ong, C.K. et al. (2002). Tree-crop interactions: Manipulation of water use and root function. *Agric. Water Manage.* **53**: 171–186.
18. Tuskan, G.A. (1998). Short rotation woody crop supply systems in the United States: What do we know and what do we need to know? *Biomass and Bioenergy* **14**: 307–315.

19. Zhou, G.Y. et al. (2003). Hydrological impacts of reafforestation with eucalyptus and indigenous species: A case study in southern China. *Forest Ecol. Manage.* **167**: 209–222.
20. Brown, R.A. et al. (2000). Potential production and environmental effects of switchgrass and traditional crops under current and greenhouse-altered climate in the central United States: A simulation study. *Agric. Ecosystems Environ.* **78**: 31–47.
21. Quadir, M. et al. (2003). Agricultural water management in water-starved countries. Challenges and opportunities. *Agric. Water Manage.* **62**: 165–185.
22. Berndes, G. (2002). Bioenergy and water—the implications of large scale bioenergy production for water use and supply. *Global Environ. Change* **12**: 253–271.
23. Droppelmann, K.J., Lehmann, J., Ephrath, J.E., and Berliner, P.R. (2000). Water use efficiency and uptake patterns in a runoff agroforestry system in an arid environment. *Agroforestry Syst.* **49**: 223–243.
24. Stott, T. and Mount, H. (2004). Plantation forestry impacts on sediment yields and downstream channel dynamics in the UK: A review. *Prog. Phys. Geogr.* **28**: 197–240.
26. Felker, P. and Guevarra, J.C. (2003). Potential commercial hardwood forestry plantations in arid lands—an economic analysis of Prosopis lumber production in Argentine and the United States. *Forest Ecol. Manage.* **186**: 271–286.
27. Pimentel, D. et al. (2002). Renewable energy: Current and potential issues. *Bioscience* **52**: 1111–1120.
28. Faaij, A. et al. (1998). Externalities of biomass based electricity production compared with power generation from coal in the Netherlands. *Biomass and Bioenergy* **14**: 125–147.
29. Annalamai, K. et al. (2003). Cofiring of coal and cattle feedlot biomass fuels. *Fuel* **82**: 1195–2000.
30. Makareviciene, V. and Janulis, P. (2003). Environmental effect of rapeseed oil methyl ester. *Renewable Energy* **28**: 2395–2403.
31. Wunderli, S. et al. (1996). Levels and congener pattern of PCDD/PCDF in fly ash and bottom ash from waste wood and natural wood burned in small to medium sized wood firing facilities in Switzerland. *Organohalogen Comp.* **27**: 231–236.
32. Johansson, I. and van Bavel, B. (2003). Levels and patterns of polycyclic aromatic hydrocarbons in incineration ashes. *Sci. Total Environ.* **311**: 221–231.
33. Pedersen, A.J. (2003). Characterisation and electrodialytic treatment of wood combustion fly ash for removal of cadmium. *Biomass and Bioenergy* **23**: 447–458.
34. Klee, R.J. and Graedel, T.E. (2004). Elemental cycles: A status report on human and natural dominance. *Annu. Rev. Environ. Resour.* (in press).

SOIL CONSERVATION

R.J. Rickson
M. McHugh
R.P.C. Morgan
Cranfield University
Silsoe, United Kingdom

INTRODUCTION

Soil is a precious natural resource, which performs many vital functions. As a growing medium, soil provides nutrients, air, and water to sustain life on earth, supporting global agricultural and manufacturing industries with renewable raw materials, enabling sustainable economic growth and prosperity. As a habitat, soil is an essential gene reservoir, maintaining a healthy biodiversity that enables micro- and macroorganisms to adapt successfully to environmental changes. Soil influences our natural and cultural landscapes by providing the foundation and raw materials for infrastructure, thus affecting the aesthetics of the world around us. Soils have enormous storage capacity, absorbing surface water, which helps to reduce flooding risk, and providing water supplies for agricultural, domestic, and industrial demands. Soil filters the water passing through it, and chemical transformations help neutralize potentially harmful substances. Soils also store nutrients, trace elements, and organic carbon. Soils are essential in the sequestration of carbon and in the absorption, storage, and reflection of solar energy—important processes that influence global climate and climate change.

Soil formation rates are very low, roughly equivalent to $0.1\ \mathrm{mm} \cdot \mathrm{y}^{-1}$ in temperate regions (1). With such limited rates of renewal, soil should be treated as a nonrenewable resource, which must be protected and conserved so that the functions of the soil, as listed above, can be maintained for future generations.

THE NEED FOR SOIL CONSERVATION

Soil conservation aims to prevent degradation of the soil resource. Such degradation occurs when one or more of the functions described above are impaired, either through *in situ* changes in soil properties, such as loss of soil organic carbon or depletion of nutrients, or through physical loss of the soil through erosion. As the latter process is irreversible, emphasis on erosion control is given here. Erosion is initiated when the kinetic energy from running water or wind is sufficient to detach and transport soil particles and aggregates from the soil mass. The eroded soil is transported and some or all of it may be deposited downslope. Often, however, the finer particles, mainly clays and organic matter, are transported into adjacent watercourses, where the nutrients associated with them can cause pollution, and deposition of the soil particles (sedimentation) can restrict water flow.

Soil conservation aims at conserving the soil properties associated with fertility (Table 1) by preventing *in situ* degradation and preventing physical loss of soil by controlling erosion. Ideally, soil conservation should also improve the levels of these fertility indicators to ensure continuing soil health and functionality and to minimize the detrimental impact of any future soil degradation process.

SOIL CONSERVATION APPROACHES

Soil conservation techniques can be classified into three approaches:

1. *Mechanical Methods*—the use of engineering structures to conserve fertility and control soil erosion.

Table 1. Main Components of Soil Fertility

Physical components	Soil depth Bulk density Soil water/water holding capacity Aggregation/soil structure Porosity/soil air
Biological components	Microorganisms Macroinvertebrates Soil respiration Organic matter/carbon
Chemical components	Extractable nutrients Essential elements Cation exchange capacity pH Electrical conductivity C/N ratio

2. *Agronomic Methods*—the use of vegetation or simulated vegetation to conserve fertility and control soil erosion.
3. *Soil Management Methods*—the manipulation of soil conditions to conserve fertility and control soil erosion.

These methods are interrelated and multipurpose in that they aim to conserve both soil fertility and control soil erosion simultaneously. In many situations, a single approach is insufficient on its own and several approaches need to be combined in an integrated soil conservation system.

Mechanical Methods of Soil Conservation

Introduction. Mechanical methods of soil conservation are applied to control soil erosion by water. Some techniques date back to ancient times such as terraces, which were adopted by Phoenician, Roman, Chinese, and Incan civilizations but are seldom used in modern times, due to high capital and labor costs incurred in construction and maintenance.

Types of Mechanical Soil Conservation Measures. Terraces are constructed on hill slopes where soil conservation is needed because of erosion risk. Terraces control soil erosion by reducing slope length and local slope steepness, thus reducing overland flow energy to erode and maintaining inherent soil fertility.

There are many different types of soil conservation terrace, often developed for local environmental conditions (e.g., the "fanya juu" terrace of eastern Africa). Generic types of terrace include the channel and bench terrace. Channel terraces are suited to slopes no steeper than 7° (12%) (2,3) (Fig. 1). They consist of a cut channel and a filled ridge (constructed from the cut material), excavated across the slope.

The distance between consecutive terraces downslope can be determined by equations based on site conditions (4,5). However, a general rule of thumb is to use a vertical interval between terraces of 1 m. This results in varying distances between terrace channels as the slope gradient changes, with shorter terrace intervals on steeper slopes. This may cause practical problems for mechanized farming.

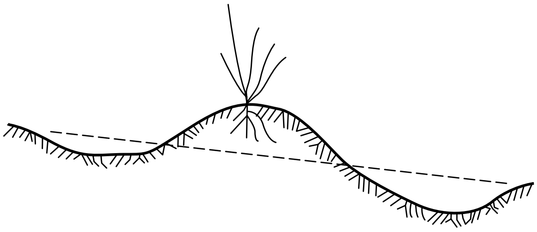

Figure 1. Cross section of a channel terrace (http//www.fao.org/documents).

The terrace channel reduces slope length, and overland flow generated between terrace channels discharges into the channel. The size (width × depth) and gradient of the channel across slope (usually at 1: 250 = 0.4% slope) determine its capacity in conveying overland flow to a safe outlet (see lined waterways below). The gentle gradient of the channel ensures that erosion within the channel itself does not occur.

Bench terraces are steplike structures, constructed across slope, which prevent soil erosion by dissecting the slope into smaller, shorter components, each with a level or near-level bench (Fig. 2). The dimensions of the benches depend on slope steepness, the desirable width of cultivable land (i.e., bench width), and the acceptable height of the steep wall of the terrace, or "riser" (4,5). The construction of the level bench requires the oversteepening of the terrace riser. To prevent the collapse of the riser, it must be supported with stones, rocks, masonry, timber, or simply a dense network of vegetation roots.

By concentrating and redirecting overland flow, terraces will change the natural hydrology, which may incur increased risk of erosion. To avoid this, flow in the terrace channels is directed to lined waterways, which are designed to carry increased volume and velocity of overland flow without causing erosion. These waterways, ideally located in natural depressions to minimize earth moving, are lined with dense, short vegetation, stones, or geotextiles. These linings permit high velocity flows without causing erosion. The width and depth of the waterway are determined by the volume of overland flow and the waterway slope and lining.

In conclusion, mechanical measures of soil conservation control the physical loss of soil by erosion. As the soil is kept *in situ*, soil fertility is also maintained.

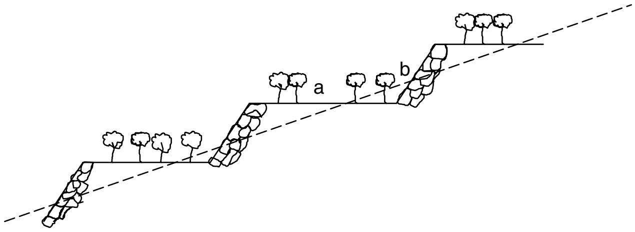

Figure 2. Cross section of bench terraces: a—level bench; b—riser.

Agronomic Methods of Soil Conservation

Introduction. Vegetation is highly effective at controlling soil erosion (6). Agronomic methods of soil conservation involve the use of vegetation to control soil erosion and to maintain or enhance soil fertility.

Water Erosion Control. The canopy, stem, and root components of plants interact with soil erosion processes. A plant canopy intercepts rainfall, changing raindrop sizes and reducing fall velocities, so that rainfall kinetic energy available for erosion is reduced. Plant stems distort overland flow paths and impart roughness to the flow, reducing flow velocity and thus kinetic energy to erode. Plant roots encourage infiltration, reducing overland flow generation, and also increase soil shear resistance against the shearing forces of overland flow. These beneficial effects of vegetation have been quantified by many workers (7–10). A fully mature, low growing, dense vegetation cover is able to control soil loss by two or three orders of magnitude compared to bare soil.

Agronomic methods of soil conservation apply these principles in the field. Maximum cover on the soil surface will protect the soil from erosive raindrops and overland flow, absorbing their energy and thus reducing the ability to erode. High stem and root densities reduce overland flow volumes, velocities, and energy to erode. These effects are achieved through cover cropping, high density planting, crop rotations, mulching, and erosion control mats. All of these control soil erosion and enhance soil fertility by contributing nutrients, organic matter, and carbon to the soil, thus encouraging microbial biomass and activity and improving soil structure and aggregate stability. A fertile, noneroding soil will be able to support a dense vegetative sward with high percentage cover, stem, and root densities, which in turn are effective at controlling erosion and enhancing soil fertility.

Vegetation is used to reduce slope length and steepness when grown in strips parallel to the contour. The strips may be of different crops (strip cropping) or as a single field strip across the slope. The strips restrict the generation of overland flow and encourage deposition of eroded sediment, so that local slope gradients might be reduced. The deposited sediments are also rich in nutrients, which are retained within the field, thus maintaining fertility and avoiding diffuse pollution off-site.

Agroforestry systems combine erosion control and maintenance of soil fertility (11). One common practice is to grow perennial trees/shrubs alongside annual crops. The trees form across-slope barriers to overland flow (also known as “alley cropping”), thus reducing erosion. While the annual crops may deplete soil nutrients in the topsoil, these can be replaced by release of nutrients assimilated in the leaves that fall from the trees. These nutrients come from reserves deeper in the soil by the tree roots, so competition for nutrients between the two components is minimal.

Wind Erosion Control. By exerting a drag on air flow, vegetation can reduce the velocity of wind below that required to initiate erosion and can enhance the deposition of soil already in the air. The most common use of vegetation is in the form of shelterbelts spaced at regular intervals at right angles to the direction of the erosive winds. Traditionally, trees have been used to form shelterbelts but they need to be supplemented by shrubs (bushes) and grasses to provide cover close to the ground surface, as 70–75% of the soil eroded by wind is carried within the lower 1 m of the atmosphere. Modern shelterbelts are based on species of willow and alder, both of which can be coppiced to encourage growth close to the ground. The most important characteristic of a shelterbelt is its density or porosity. The belt needs to be porous so that some air passes through the belt but at a low velocity. Clearly, if the belt is too sparse, the velocity of the wind is not reduced. A porosity of 40–50% is generally considered optimum. Belts are usually designed to maintain wind speeds below 75% of the speeds recorded over open ground. To achieve this, they need to be spaced at intervals of between 12 and 17 times the belt height.

The principles of shelterbelts can also be used to design in-field shelter systems in which crops are alternated on a strip system, aligned at right angles to the prevailing wind. Thus, strips of barley, spaced at 5–10 m intervals, can be alternated with strips of carrots, onions, or sugar beet to control erosion.

Since shelterbelts require a certain amount of time to become effective, they are not appropriate where immediate control of wind erosion is required. In these situations, windbreaks can be made from brushwood or plastic netting. Brushwood fences are commonly constructed on coastal dunes to reduce wind velocities and stabilize the mobile sand, which can then be seeded or planted with shrubs and trees to provide long-term stability and reclamation.

Soil Management Methods of Soil Conservation

Farming operations carried out up and down the slope will increase the risk of overland flow generation and loss of soil and nutrients. Changing to contour cultivations will reduce these risks significantly, but this is not always practical, especially on complex or steep slopes, where the arrangement of fields on the contour makes harvesting difficult and incurs high energy consumption. Contour ridging creates raised ridges across slope, which will reduce slope length an effective catchment area for overland flow generation, and intercept any overland flow that is generated. The across-slope ridges can be joined together at regular intervals by “ties” running up/downslope, thus creating small basins that trap any overland flow and eroded soil, conserving both soil and water within the ”tied ridges.”

The factors affecting the susceptibility of soil to erosion can be affected by soil management methods, which can both degrade or conserve soil. For example, intensive arable production can lead to severe soil degradation, as soil is broken down by numerous tillage practices into a loose, fine tilth seedbed. The resulting small soil aggregates are highly susceptible to erosion. Studdard (12) and Pochard (13) showed that under conventional arable production, there is little time between cultivations for aggregates to reform into larger, more stable, less erodible units. Soil structural stability is further reduced by loss of carbon by oxidation as tillage exposes deeper soil layers to the air. Soil faunal biomass, including fungal mycelia and earthworms, is important in maintaining cohesive bonds within and between aggregates, but it can be destroyed during soil disturbance.

Reducing the number and/or intensity of cultivations preserves soil structural stability and aggregation. Less disturbed aggregates are able to resist erosive forces, as cohesive bonds within the aggregates are maintained. Fewer, less invasive cultivations mean more of the vegetative biomass remains on or within the soil. This biomass is able to absorb the energy of erosive agents, so erosion is controlled. For example, a rough, trashy soil surface imparts friction to overland or air flow, causing turbulence and eddying, resulting in a reduced flow velocity and kinetic energy to erode.

The relative reduction in soil disturbance and retention of vegetative material on and within the soil are the principles behind the use of "conservation tillage." This term incorporates a continuum of soil management methods such as minimum till, zero till, mulch tillage, direct drilling, noninversion tillage, subsurface mulch tillage, and plough/plant cultivation (14). Conservation tillage aims at maintaining soil fertility and controlling erosion. It can also reduce input costs to agriculture, as fewer, less invasive cultivations expend less energy and fuel, resulting in dramatic adoption rates in the 1970s in the United States during the oil crisis. It should be noted that the success of conservation tillage is related to soil type and timeliness of operations. For example, there is evidence that reducing the number of cultivations can increase bulk density on structureless soils, leading to overland flow generation and flooding risk, which can only be alleviated by deep subsoiling periodically. Supporters of conservation tillage claim that these are short-term effects, which are offset in the medium to long term by the improvement of aggregation and inherent soil structure resulting from increased levels of biotic activity and organic matter levels.

Manipulation of the soil surface condition by tillage is used to control wind erosion. On light soils such as sands and silts, the soil is ploughed and then pressed with a heavy roller. Drilling is then carried out at 90° to the plough/press operation. The resultant tilth is rough and cloddy, to impart maximum friction to air flow, thus reducing wind speed and ability to detach and transport soil.

Soil conditioners are applied to soils susceptible to degradation, especially erosion. These products can perform the dual role of enhancing soil fertility and controlling erosion. Products with high organic matter (such as seaweed extract and tree bark) improve inherent soil fertility in terms of nutrients and organic matter, thus improving aggregation, structure, porosity, infiltration capacity, and water holding capacity.

Hydrophilic soil conditioners including bark, limestone, palm oil effluent, and seaweed extract increase aggregation and infiltration capacity, so overland flow is less likely to occur. The increase in organic matter from these products helps control erosion by increasing cohesive bonds and the stability of aggregates. Hydrophobic soil conditioners including latex, emulsions, and molasses form an impermeable film on the soil surface, which repels water and increases the soil's resistance to erosion. While these latter products control soil erosion, they have no effect on inherent soil fertility.

Table 2. Objectives of Soil Conservation Measures in Relation to Economic and Social Factors

1. The need for soil conservation should be recognized by all stakeholders—farmers, local communities, landowners, advisory services, government departments, NGOs, and politicians.
2. Soil conservation measures must be compatible with present and future management plans: they should not be imposed by external bodies such as researchers, scientists, advisory and extension services, or politicians.
3. Methods of construction will be based on available and appropriate technology and level of education. Technology transfer will be "bottom-up" not "top-down."
4. The easiest and minimal maintenance of soil conservation scheme will be adopted.
5. Present infrastructure should not be radically changed by soil conservation measures.
6. Mechanical measures should be laid out economically.
7. Crop selection will be based on economic as well as soil conserving properties.
8. Available grants and subsidies will be used. This has been demonstrated in the United States for many years. Farmers are eligible for USDA farm subsidies, but only if they demonstrate soil conservation practices have been implemented. A similar scheme will operate in England and Wales under the DEFRA Single Payment Scheme. To be compliant with the scheme, farmers have to draw up a simple risk-based soil management plan that protects against soil erosion and maintains soil organic matter and soil structure.
9. Conservation measures will ensure long-term productivity. Farmers are considered "land stewards," taking responsibility for conserving the soil resource for future generations. This concept is exemplified by the "Landcare" movement in Australia and New Zealand.

IMPLEMENTATION AND ADOPTION OF SOIL CONSERVATION MEASURES

While much is known of soil conservation techniques, they will fail unless the right social, economic, and political conditions prevail (15). Table 2 highlights the main considerations for the adoption and long-term success of soil conservation measures.

CONCLUSIONS

Soil is a sustainable resource, if treated carefully and with respect. Failing to protect and conserve soil can lead to rapid soil degradation, which in turn can be an irreversible process. Soil conservation can be defined as the maintenance of soil fertility and the control of soil erosion. These two objectives can be realized by appropriate management of the soil resource, using diverse soil conservation approaches. Application of these techniques will ensure that the diverse forms and functions of soil will be conserved for future generations.

BIBLIOGRAPHY

1. Young, A. (1998). *Land Resources: Now and for the Future*. Cambridge University Press, Cambridge, UK.
2. Sheng, T.C. (1972). A treatment-oriented land capability classification scheme. In: *Report on the Latin American Watershed Management Seminar*. FAO No. TA3112.
3. Sheng, T.C. (1972). Bench terracing. *J. Sci. Res. Council Jamaica* **3**: 113–127.
4. Morgan, R.P.C. (1995). *Soil Erosion and Conservation*, 2nd Edn. Longman, London.
5. Hudson, N.W. (1981). *Soil Conservation*, 2nd Edn. Batsford.
6. Coppin, N.J. and Richards, I.G. (1990). *Use of Vegetation in Civil Engineering*. Butterworths, London.
7. Laflen, J.M. and Colvin, T.S. (1981). Effect of crop residue on soil loss from continuous row cropping. *Trans. ASAE* **24**: 605–609.
8. Hayes, J.C., Barfield, B.J., and Barnhisel, R.I. (1984). Performance of grass filters under laboratory and field conditions. *Trans. ASAE* **27**(5): 1321–1331.
9. Babaji, G.A. (1987). *Some plant stem properties and overland flow hydraulics: a laboratory simulation.* Ph.D. thesis. Cranfield University at Silsoe.
10. Ree, W.O. (1949). *Flow of Water in Channels Protected by Vegetative Linings*. USDA Technical Bulletin 967.
11. Young, A. (1997). *Agroforestry for Soil Management*, 2nd Edn. CAB International in association with the International Centre for Research in Agroforestry.
12. Stuttard, M.J. (1985). *The Water Drop Test of Soil Erodibility*. Silsoe College Report, Cranfield University.
13. Pochard, L. (1986). *Effects of tilth on water conservation,* M.Sc. thesis. Cranfield University at Silsoe.
14. Cogo, N.P., Moldenhauer, W.C., and Foster, G.R. (1984). Soil loss reductions from conservation tillage practices. *Soil Sci. Soc. Am. J.* **48**: 368–373.
15. Stocking, M.J. and Tengberg, A. (1999). Soil conservation as incentive enough—experiences from Brazil and Argentina. In: *Incentives in Soil Conservation From Theory to Practice*. D.W. Sanders, P.C. Huszar, S. Sombatpanit, and T. Enters (Eds.). Science Publishers, London.

LANDSCAPE WATER-CONSERVATION TECHNIQUES

Douglas H. Fender
International Turf Producers
Foundation
Meadows, Illinois

When given information and technology, people will make the effort to conserve. Plant selection will have less impact on water use than either irrigation or soil preparation. Water budgets give consumers more options and personal choice, while providing an effective way to conserve water.

Outdoor water-conservation measures typically focus on reducing or eliminating landscape water use. But implementing new scientific findings and advanced technology and general education can go a long way toward achieving the same end, just as these methods have proved successful in conserving water indoors. People will act to conserve water and improve the environment when properly informed of and motivated by the best scientific knowledge and technology.

AN INDIVIDUAL'S RIGHT TO CHOICE

Numerous water-use studies have documented that, depending on an area's climate, residential outdoor water use can account for between 22 to 67% of total annual water use. Clearly, this represents a vast opportunity for conservation. But to maintain an individual's right to personal choice and to maximize the positive environmental benefits of landscaping, a variety of factors need to be addressed that take location into account. When dealing with living plants, a one-size-fits-all solution will not be effective. But proven advanced scientific landscape water-conservation principles and practices do exist, and these can be modified and refined according to area-specific needs (Fig. 1).

THE NEED FOR CLEAR AND CAREFUL DEFINITIONS

First, however, potential targets for widespread outdoor water conservation should be clearly and carefully defined. Too often, a narrow definition focuses exclusively on landscape water use. Narrow definitions often overlook potentially high water-use elements such as swimming pools or other water features, whose evaporative losses are as great as or greater than those from landscape application.

Other nonlandscape outdoor water uses include washing cars, driveways, sidewalks and siding—and even some types of children's water toys. In addition to conserving in these areas, many techniques can be applied landscape water conservation as well.

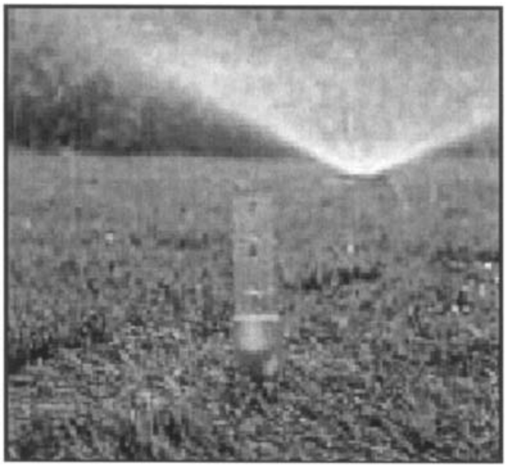

Figure 1. Rather than attempting to regulate or ban a specific water use, the water budget technique leaves the determination in the hands of the rate-paying water user.

THE WATER BUDGET PROGRAM

One of the most effective techniques is based on the practice of advising people how much water they can use, rather than telling them how they must use it. Termed a "water budget" or "water allocation" method, water providers establish a series of escalating allocation/pricing tiers so that every unit of water (i.e., 1000 gallons) in excess of a base quantity costs more than the previous unit.

Because outdoor water use can be measured and priced higher, people adjust their end uses according to their personal desires and financial concerns. This approach eliminates the need for contentious public hearings on landscape ordinances and the development of debatable plant lists, as well as the potential for draconian enforcement practices and so-called "water police."

Water budgets, for both indoor and outdoor water use, encourage individual freedom of choice and allow artistic expression by homeowners and landscape designers. Rather than attempting to regulate or ban a specific water use, this technique leaves the determination in the hands of the rate-paying water user. As we've seen, when gasoline prices rise, individuals can quickly adjust their use patterns. The same holds true for water.

Once purveyors decide how to allocate and price water, they then have the very important role of assisting in the development and distribution of scientifically based educational materials on water conservation. Again, successful indoor conservation practices can be easily converted to outdoor water conservation, particularly as it relates to landscape water use.

The Two-Track Strategy

Moving from general to more specific landscape water-conservation recommendations, a two-track strategy that emphasizes different approaches for existing landscapes and newly planned and installed landscapes may be most effective, as follows.

Existing Landscape Areas

Predrought/Premaximum Heat-Day Practices

1. **Increase water infiltration** with dethatching or hollow-core aerification of all lawn areas, as well as under the drip line of trees. Till garden areas to break up surface crusting, adding mulch where appropriate.
2. **Trim or prune** trees, shrubs, and bushes to remove low-hanging, broken or diseased parts and allow greater sunlight penetration throughout and beneath the plant. Generally speaking, plant water use is proportional to total leaf surface; thus, properly pruned plants should require less water.
3. **Fertilize all plants** (when soil temperatures reach at least 50 °F or 10 °C) with a balanced plant food that contains nitrogen (N), phosphorus (P), and potassium (K) according to the results of soil testing or as experience has shown is appropriate.
4. **Sharpen pruning shears and mower blades**. Dull blades encourage plant water losses and the introduction of disease.
5. **Establish or confirm soil type(s)** to match water-infiltration rates with future water-application rates and to determine if soil pH adjustments are recommended.
6. **Perform irrigation-system maintenance**, regardless of type (hose-end, drip, in-ground, etc.) to ensure maximum uniformity of coverage and overall operation. Repair or replace broken or damaged nozzles or heads. Flush drip system emitters to ensure proper flow. Ensure that rainfall shutoffs and other devices are working properly.
   a. Acquire and/or install hose-end water timers for all hose bibs.
   b. Adjust in-ground system controllers according to plant's seasonal needs.
7. **Upgrade in-ground irrigation systems** by adding soil-moisture meters, rain shutoff devices, or evapotranspiration (ET)-based controllers.
8. **Relocate drip emitters**, particularly around trees, to the outer edge of their drip lines. This will result in higher water use, but it will also encourage a better root system that will anchor the tree in high winds and provide the water and nutrients that are needed by large trees, shrubs, and other plants (Fig. 2).

Figure 2. Locate drip emitters to provide water and nutrients needed by large trees, shrubs, and other plants.

Figure 3. Water in the early morning.

9. **Confirm water-application rates** for hose-end or automatic systems to know the actual running times required to distribute a specified amount of water within a given amount of time.
10. **Water in early morning** (Fig. 3) when wind and heat are lowest, to maximize the availability of the water to the selected plants.
11. **Irrigate all plants infrequently and deeply**, according to local ET or soil-moisture requirements, to establish a deep, healthy root system. A core-extracting soil probe or even a simple screwdriver can help determine when to water if the more sophisticated ET rates are not available (Fig. 4). Professional turf managers gradually lengthen the interval between irrigations to create gradual water stress for deeper rooting.
12. **Cycle irrigation applications** (on-off-on-off) to allow penetration and avoid runoff. Depending on soil types, the running times may be from 5 to 15 minutes and off times from 1 to 3 hours. Repeat this cycle until necessary amounts of water are applied and maximum penetration is achieved.
13. **Begin regular mowing** when grass blades are one-third higher than desired postmowing length, and keep clippings on the lawn.
14. **Raise the mowing height** as the summer progresses to the highest acceptable level to encourage deep rooting. (Note: This has been a traditional recommendation, but further study is required to refine this approach and maximize effective water use and/or conservation.)

Drought or Maximum Heat-Day Practices—to Maximize Landscape Appearance

1. **Withhold fertilizers**, particularly nitrogen, on turfgrass; however, small amounts of potassium will aid in developing more efficient roots.
2. **Reduce mowing frequency** to minimize shock to turf areas.
3. **Reduce traffic on turf areas**, this will minimize wear and possible soil compaction.
4. **Adjust automatic timers** of in-ground irrigation systems according to plant's seasonally changing water needs.

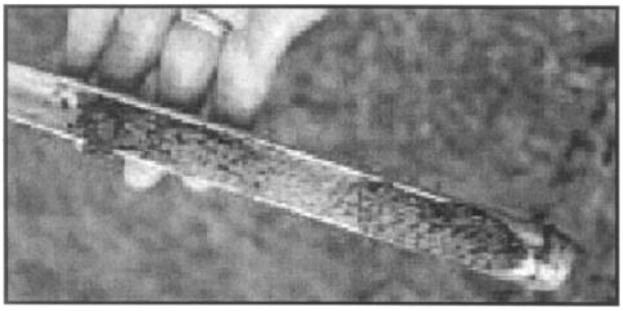

Figure 4. An example of a core-extracting soil probe.

Drought or Maximum Heat-Day Practices—If Dormant Turf Appearance is Acceptable

1. **Eliminate all traffic on turf areas**, including mowing, which will probably not be necessary because of the extremely slow growth rate.
2. **Adjust automatic timers to manual** or use hose-end sprinklers to apply approximately one-quarter inch of water a week. The dormant lawn will have a tan, golden, or light brown appearance; however, light/infrequent watering will be sufficient to maintain life in the crowns of the grass plant during this period.
3. **Minimize water applications** for all plant materials to the essential amounts needed to maintain plant vitality.

Postdrought or Maximum Heat-Day Practices Irrigate All Plants to Reestablish Soil-Moisture Levels, beginning with staged increases in watering and progressing toward deep and infrequent watering. By gradually lengthening the running times but adding greater spacing between watering applications, the initially shallower roots will extend to reach deeper soil moisture. Previously dormant turf will recover rather quickly, and other plants will regain their vigor.

Newly Planned or Installed Landscape Areas. Though fewer in number than existing landscape areas, newly planned and installed landscape areas can generally achieve greater water savings if all of the best design, plant selection, installation, and management practices now available are closely observed and fully implemented.

The basic principles of xeriscape landscaping provide an excellent starting point, providing they are fully understood and properly applied. However, note that it is incorrect to assume that a xeriscape is supposed to be a totally grassless landscape or one that uses only rocks, cactus, or driftwood.

Dr. Doug Welsh, former director of the National Xeriscape Council, wrote in *Xeriscape Gardening: Water Conservation for the American Landscape:* "In xeriscape landscaping we try to plan the amount of turf so that the investment in water will be repaid in use and beauty. In many instances grass *is* the best choice. For play areas, playing fields and areas for small pets, grass is often the only ground cover that will stand up to the wear. Turf also provides unity and simplicity when used as a design component."

Rather than duplicate information contained in several xeriscape manuals, this publication will focus on a limited number of very critical elements that can maximize landscape water conservation.

Efficient irrigation is without question an important water-conservation activity. People waste water; plants don't. Overwatering not only wastes water, but it also weakens or kills plants more than underwatering. Another wasteful practice seen all too often is misapplication of water, resulting in rotted fences and house siding, flooded

Figure 5. A wasteful practice seen all too often is misapplication of water, resulting in flooded sidewalks or driveways and rivers of water wastefully flowing down gutters.

sidewalks or driveways and rivers of water wastefully flowing down gutters (Fig. 5).

Though less so today, many new in-ground landscape-irrigation systems have been sold in the recent past on the basis of simplicity, for example, "set it and forget it." Homeowners, intimidated by the sophisticated appearance of the system's control box, would not modify the settings for seasonal plant water-use changes. Even worse, to be cost-competitive, many systems did not include readily available and relatively inexpensive soil-moisture meters, rain shutoff devices, or multistation programs. These deficiencies have resulted in overwatered landscapes, where water runs down the streets and systems continue to operate during torrential downpours.

Because of its high visibility, turf watering can be seen as the antithesis of water conservation and is often an out-and-out target for elimination. In some locales, "cash-for-grass" programs are used to pay homeowners handsomely to remove grass from their landscapes. One highly respected West Coast water official noted at a conservation convention: "It isn't the grass that causes a problem, it's the poorly designed and poorly operated irrigation system. I can't control the irrigation systems, but I can reduce the amount of grass in a landscape, and that will control the water-use problem created by bad irrigation." The fault was not the grass, but the fact that it was being improperly watered.

Soil analysis and improvements are another very important aspect of water conservation. The soil on most new residential and commercial landscape sites has literally been turned upside-down during the construction process, and the topsoil is placed beneath a layer of clay. Then it is compacted as hard as cement by equipment, piles of building materials, and construction-worker foot traffic. The soil's texture, chemistry, and natural flora and fauna are destroyed.

More water could be conserved and healthy landscape plants more easily grown by improving the soil before planting than by any other process or technique. The initial cost of adding topsoil and soil test-guided amendments to improve the soil may seem high, but the return on that investment will be even higher. Failing to improve the soil prior to planting, when it is most practical and efficient, will result in roots not being able to penetrate as deeply as possible and runoff occurring almost instantaneously. If the soil pH is not correct, plants will not be healthy, nutrients will not be used, and chemical leaching can take place.

Appropriate plant selection can be a source of frustration or misunderstanding and not produce the hoped-for water savings, particularly when it is combined with "practical turf areas." Lists of low water-use landscape plants can cause confusion because they are based on the incorrect assumption that those plants capable of surviving in arid regions are low water users, when the plants typically are only drought-resistant (Fig. 6).

The majority of turfgrass species and cultivars have been scientifically assessed for their evapotranspiration (water-use) rates and can be selected according to the needs of a specific climate. On the other hand, very few tree and shrub species and cultivars have undergone comparable quantitative water-use assessments. One stunning exception comes from research at the University of Nevada, where Dr. Dale Devitt found that "one oak tree will require the same amount of irrigation as 1800 square feet of low-nitrogen fertilized turfgrass!"

It is also important to understand that though "low water-use" plant lists were developed with the best intentions and purposes in mind, practically all of these lists have been based on anecdotal evidence or consensus judgments, not scientific measurement under controlled procedures. Quite simply, a given group decides that based on their experiences and suppositions, a certain

Figure 6. The minilysimeter gauges water use under actual turf conditions. This photo at Washington State University's research site in Puyallup, shows minilysimeters being weighed. The pots are weighed for daily water evaporation, rewetted, and returned to their "holes" in the turf. These scientific measurements under controlled procedures help researchers better understand low-water use.

plant should or should not be placed on a "low water-use" list. As has been seen time after time, it typically is not the plant that wastes water, but the person who is in charge of its care.

When establishing new lawns, turfgrass sod has been shown to require less water than seeding; the evaporative losses from bare soil are greater for seeded areas than they are for turf-covered soil beneath sod (Fig. 7).

Water "harvesting" and reuse is another water-conserving practice gaining greater use. It can be employed to conserve public water supplies and recharge groundwater supplies.

Historically, planners and designers have focused their efforts on moving rainwater and snowmelt away from a property as quickly as possible, giving little thought to the possible advantages of using that water for landscape purposes. Consider the fact that during a 1-inch rainfall, a 35-foot-by-60-foot roof (approximately 2000 square feet) will collect nearly 1250 gallons of water. Rushing this water to gutters and then sewers makes little sense when it could conceivably go into a system that could capture or distribute it across a landscaped area.

Another increasingly feasible source of additional landscape water is recycled or gray water. Some communities are installing dual water-delivery systems where one carries potable water for drinking, cooking, cleaning, and other general household uses and a second system delivers less thoroughly treated (but very safe) water for use on landscapes. On a small scale, some locales encourage collecting a home's gray water (from clothes washers, etc., but not toilets) for use on landscapes.

After being applied to a landscape area, harvested, recycled, or gray water is either transpired by plants and evaporates into the atmosphere or finds its way into groundwater supplies after it has been cleansed by the plant's root structure.

Figure 7. When establishing new lawns, turfgrass sod has been shown to require less water than seeding.

As science and technology continue to advance, new and better information, tools, and systems will become available to help people establish and maintain water-conserving and environmentally beneficial landscapes. An ongoing challenge will be keeping pace with these developments, sharing that information, and continually improving best management practices.

Just as we improve our health by using water to brush our teeth and to wash our bodies and clothes, applying water judiciously to a properly designed and installed landscape can improve our health and the general environment.

Efforts to eliminate landscape water use not only take away freedom of personal choice; they also bring many environmental, economic, and emotional drawbacks that could be more costly in the long run.

CROP WATER REQUIREMENTS

Mladen Todorovic
Mediterranean Agronomic
Institute of Bari
Valenzano, Bari, Italy

Crop water requirements (CWR) refer to the amount of water required to compensate for evapotranspiration losses from a cropped field during a specified period of time. Crop water requirements are expressed usually in mm/day, mm/month, or mm/season, and they are used for management purposes in estimating irrigation water requirements, irrigation scheduling, and water delivery scheduling.

The concept of crop water requirements is intimately connected with crop evapotranspiration because both refer to the same amount of water. Nevertheless, there is some difference between them. Crop evapotranspiration represents the water losses that occur (i.e., a hydrological term), but crop water requirement indicates the amount of water that should be supplied to account for these losses (i.e., an irrigation management term). In fact, this amount of water corresponds to the effective irrigation water supply necessary to reach the maximum yield. Therefore, the estimation of crop evapotranspiration precedes the estimation of crop water requirements where the latter usually represents the values of crop evapotranspiration aggregated over some period of time. For management purposes, crop evapotranspiration should be transformed into crop water requirements to be used for estimating irrigation water requirements and later, management of the irrigation water supply. Then, crop water requirements are satisfied by the amount of irrigation water effectively supplied to the root zone and the effective precipitation.

The concept of crop water requirements has become important due to the development of large engineering works when it was necessary to estimate the water volumes to be supplied to newly irrigated areas. In general, it is possible to distinguish between crop water requirements for long-term planning, where an average climate or a climate with a certain probability of occurrence can be used for a CWR estimate, and crop water requirements for real-time management, where climatic data of the ongoing season are applied.

The procedures for estimating crop water requirements correspond to those of crop evapotranspiration because both terms refer to the same amount of water. More details of measuring and estimating crop water requirements are given under the article on CROP EVAPOTRANSPIRATION of this Encyclopedia.

Example of an estimate of seasonal crop water requirements for tomato growing in the Mediterranean region based on a planting date of April 15th and where the average monthly reference evapotranspiration values are as follows: $\mathrm{ET}_{o_{\mathrm{April}}} = 3.5$ mm/day; $\mathrm{ET}_{o_{\mathrm{May}}} = 4.4$ mm/day; $\mathrm{ET}_{o_{\mathrm{June}}} = 5.1$ mm/day; $\mathrm{ET}_{o_{\mathrm{July}}} = 5.8$ mm/day; $\mathrm{ET}_{o_{\mathrm{August}}} = 5.6$ mm/day, and $\mathrm{ET}_{o_{\mathrm{September}}} = 4.2$ mm/day.

Calculation Procedure. Some data necessary for calculation are taken from the article on CROP EVAPOTRANSPIRATION of this Encyclopedia: the lengths of the growing stages are taken from Table 3 and K_c values from Table 4 of that article. The growing cycle of tomato can be presented graphically as in Fig. 1: K_c is fixed at 0.5 from April 15th to May 15th, then it increases linearly until June 24th when it reaches the value 1.15; K_c is fixed at 1.15 from June 24th to August 7th; then it decreases linearly to 0.8 on September 6th, the last day of growing cycle. Hence, the average monthly values of K_c are determined as follows: $K_{c_{\mathrm{April}}} = 0.5$; $K_{c_{\mathrm{May}}} = 0.57$; $K_{c_{\mathrm{June}}} = 0.94$; $K_{c_{\mathrm{July}}} = 1.15$; $K_{c_{\mathrm{August}}} = 1.04$; $K_{c_{\mathrm{September}}} = 0.835$.

Then, the crop water requirements are calculated for each month as

$$\begin{aligned}
\mathrm{CWR}_{\mathrm{April}} &= K_{c_{\mathrm{April}}} * \mathrm{ET}_{o_{\mathrm{April}}} * 15\\
&= 0.5 * 3.5 * 15 = 26.2 \text{ mm/month}\\
\mathrm{CWR}_{\mathrm{May}} &= K_{c_{\mathrm{May}}} * \mathrm{ET}_{o_{\mathrm{May}}} * 31\\
&= 0.57 * 4.4 * 31 = 77.8 \text{ mm/month}
\end{aligned}$$

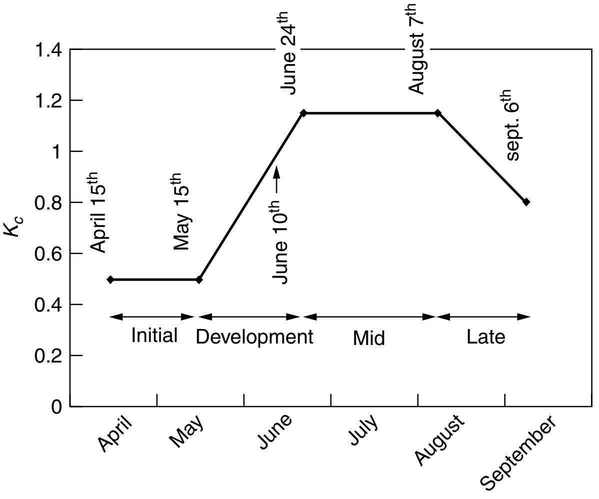

Figure 1. The variation of K_c for a tomato crop during the growing season.

$$\begin{aligned}
\mathrm{CWR}_{\mathrm{June}} &= K_{c_{\mathrm{June}}} * \mathrm{ET}_{o_{\mathrm{June}}} * 30\\
&= 0.94 * 5.1 * 30 = 143.8 \text{ mm/month}\\
\mathrm{CWR}_{\mathrm{July}} &= K_{c_{\mathrm{July}}} * \mathrm{ET}_{o_{\mathrm{July}}} * 31\\
&= 1.15 * 5.8 * 31 = 206.8 \text{ mm/month}\\
\mathrm{CWR}_{\mathrm{August}} &= K_{c_{\mathrm{August}}} * \mathrm{ET}_{o_{\mathrm{August}}} * 31\\
&= 1.04 * 5.6 * 31 = 180.5 \text{ mm/month}\\
\mathrm{CWR}_{\mathrm{September}} &= K_{c_{\mathrm{September}}} * \mathrm{ET}_{o_{\mathrm{September}}} * 6\\
&= 0.835 * 4.2 * 6 = 21.0 \text{ mm/month}
\end{aligned}$$

The seasonal crop water requirements represent the sum of the monthly values, 656.1 mm/season.

Crop water requirements can also be estimated daily, provided that daily values of reference evapotranspiration are available. There, K_c values can be interpolated similarly, as in the previous example, or they can be estimated using more complicated approaches (e.g., the dual K_c approach).

Crop water requirements of one crop variety vary from one location to another due to variations in climatic parameters and, therefore, to different reference evapotranspiration, different planting dates and lengths of growing season, different K_c values, etc.

AGRICULTURAL WATER USE EFFICIENCY (WUE) AND PRODUCTIVITY (WP)

PASQUALE STEDUTO
Mediterranean Agronomic
Institute
Valenzano, Italy

Among the different sectors of society, agriculture represents the major water consumer. Rain-fed and irrigated agriculture uses water at various degrees of *efficiency*, resulting in different levels of *productivity*.

The general term *efficiency (e)* indicates the ratio of *output* to *input* of variables (Eq. 1) associated with *transformation* (e.g., of energy) or *transport* (e.g., of mass and energy) *processes*:

$$e = \frac{\text{output}}{\text{input}} \tag{1}$$

The *efficiency* term (e) has two peculiarities: (1) the units are nondimensional and (2) the values are always between the theoretical limits of 0 and 1. In water use, the efficiency (e_w) has water amounts (e.g., volumes, quantities, or depth) in both *input* and *output* variables.

The general term *productivity* (p) still indicates a ratio of output to *input* of variables, as in Equation 1, but these *output* and *input* variables are not associated with *transport processes* and might be associated, but not necessarily, with a *transformation process*. The consequent peculiarities of the *productivity* term (p) do not match those of the *efficiency* term (e), that is, the units are dimensional, and the theoretical limits of 0 and 1 are lost. In water use, the productivity term (p_w) has water amounts (e.g., volumes, quantities, or depth) only in the *input*. The *output*, instead, generally takes the form of quantities produced: physical (e.g., dry matter or final yield of crops, or meat or dairy products from animal husbandry, etc.), economic (e.g., gross income, net revenue, etc.), or any other form.

The efficiency term e_w has to do more with the way water is used by the system, whereas p_w has to do more with the outcome from the system per unit of water it consumes. Nevertheless, it turns out that a change in e_w is reflected in a change in p_w, implying that there is a link between e_w and p_w, though the reverse is not the case. Moreover, though p_w cannot be strictly defined by Equation 1, it still resembles a sort of efficiency of the system to produce a given *output* for a given *input*. The undefined degree of commonality and distinction between e_w and p_w has led to some confusion in investigating agricultural water use.

A GENERAL FRAMEWORK FOR AGRICULTURAL WATER USE EFFICIENCY AND WATER PRODUCTIVITY

Figure 1 illustrates a suitable framework for identifying the *efficiency* (e_w) and the *productivity* (p_w) components within the domain of agricultural water use. The framework distinguishes irrigated from rain-fed agriculture and splits the overall path that water "travels" to produce a given yield into single segments of particular relevance and specific boundary conditions.

In irrigated agriculture, the source of water can be a reservoir (e.g., a dam) from which water will be conveyed and delivered to the farm. Applying Equation 1 to the whole path from the reservoir to a cropped field at farm level, will express the *efficiency* (e_w) in water transport where the *input* is the outflow from the reservoir (*Res_w_out*) and the *output* is the water effectively stored in the soil root zone (*Root_zone_w*). The whole path can

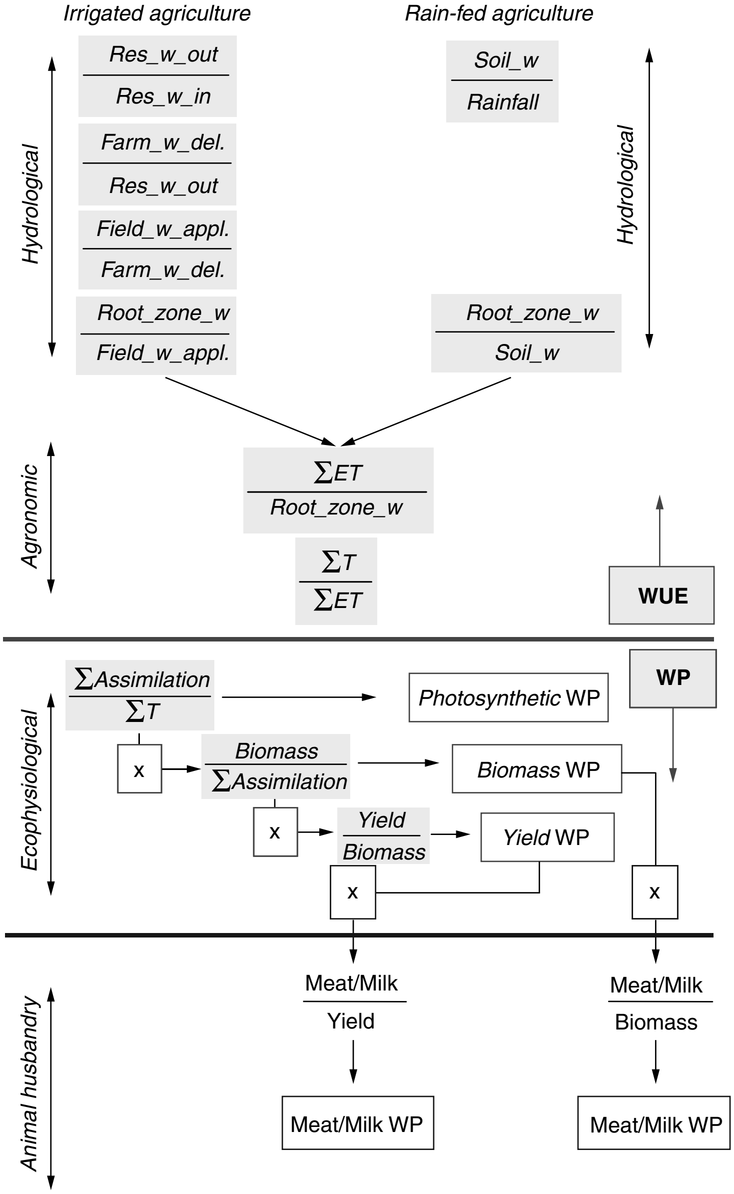

Figure 1. Suggested framework for analyzing components of *water use efficiency* and *water productivity* in agricultural crop production and animal husbandry. The overall path "traveled" by water is split into different segments where appropriate analysis can be performed. See text for explanation.

be split into different segments (e.g., from the dam to the collective irrigation system; from the collective system to the hydrant of a single farm, from the hydrant to the cropped field where an irrigation system is present, and from the outlets of the irrigation system into the root zone of the crop being irrigated) to investigate the single component efficiencies and the overall efficiency obtained from the product of the single-segment efficiencies.

In rain-fed agriculture, rainfall represents the source of water, and the ratio of the water stored in the soil (*Soil_w*) to the water from the rain (*Rainfall*) represents the rain storage *efficiency*. Out of all the water stored in the soil (*Soil_w*), only a portion may be available in the soil explored by crop roots (*Root_zone_w*), and this represents the available water *efficiency*. All efficiencies mentioned so far can be qualified as *hydrological*.

After the final allocation of water in the *root zone* (whether irrigated or rain-fed), water follows a common pathway for the crop production process: it enters the plants and evapotranspires into the atmosphere. We can identify two additional efficiencies that may be qualified as *agronomic*: one is expressed by the ratio of the amount of water used by evapotranspiration (ΣET) to the water stored in the soil root zone (*Root_zon_w*); the other is expressed by the ratio of the amount of effectively transpired (ΣT) and the total amount of evapotranspired water (ΣET). This last segment of the water path is essential because nontranspired water is not considered productive to the crop. Thus, a distinction has to be made between soil evaporation (E), which is a loss, and crop transpiration (T).

At this point, the water has finished its trip from the source, all the way through the system, passing through the soil, entering the plant, and ending up into the atmosphere, beginning as a liquid phase and ending up as a vapor phase. So far, all the expressions of *efficiency* for the different segments (either those pertaining to the *hydrological* pathway or those pertaining to the *agronomic* pathway) properly match the definition of Equation 1. because they are nondimensional and have theoretical limits between 0 and 1.

Examining the remaining segments of the framework reported in Fig. 1, we next enter the *physiological* (or better, *ecophysiological*) domain of the agricultural system where the term *efficiency* loses its rigorous meaning and it is more appropriate to talk in terms of *water productivity* (p_w). Three major expressions for p_w can be identified: (1) *photosynthetic water productivity* (*photosynthetic* p_w); (2) *biomass water productivity* (*biomass* p_w); and (3) *yield water productivity* (*yield* p_w).

Photosynthetic p_w is defined as the ratio of leaf net assimilation to leaf transpiration (*A*/*T*) and varies among species according to their photosynthetic pathways. These species can be grouped into three major types: the C_3, C_4, and CAM plants. All three types ultimately rely on the Calvin photosynthetic carbon reduction cycle for their CO_2 assimilation, but some additional features characterize those groups that make the *photosynthetic* p_w of CAM plants higher than that of C_4 plants, which in turn is higher than that of C_3 plants, in comparable environmental conditions. Values of *photosynthetic* p_w may also vary according to the nutritional status of the leaves and their age. Variable environmental conditions (e.g., radiation regime, dryness of the atmosphere, etc.) may also affect the values of *photosynthetic* p_w. Under equal ambient conditions, all species belonging to the same group type have similar *photosynthetic* p_w values.

Biomass p_w is defined as the ratio of cumulative aboveground biomass (biomass) of a crop canopy to cumulative transpiration (ΣT) by the crop. The relationship between cumulative aboveground biomass and ΣT is generally quite strict. The slope of the relationship may vary from year to year and from location to location, depending on the different environmental conditions and the biomass composition of the various species. When ΣT is normalized for a reference evaporative demand of a given climatic environment and the nutritional conditions are similar, many crop species follow the same slope and group into the same types as indicated for the *photosynthetic pathways* p_w (C_3, C_4, and CAM) and the main production compounds (e.g., carbohydrates, proteins, or lipids).

Yield p_w is defined as the product of *biomass* p_w times the harvest index (HI, i.e., the ratio of the marketable yield to aboveground biomass). *Yield* p_w values are variable among species and cultivars, depending on the values of HI, which in turn depend on the partitioning of assimilates in the various plant organs (e.g., leaves, stem, grains). Of course, HI can also vary greatly as a consequence of the environmental conditions where the plant is growing (e.g., water stress, nutritional deficit, etc.).

The water productivity concept can be extended to animal production systems when the crop's biomass and/or yield are used as feed. Then, the product of biomass or yield water productivity is multiplied by the factor of conversion of the feed into animal production (dairy product or meat). The ratio approach can be extended throughout the system up to the final step of interest (Fig. 1).

Furthermore, when the amount of product (the nominator of the productivity ratio) is expressed as gross income (e.g., multiplying it by the price) or as net income (e.g., deducting production costs), the ratio can represent the *economic water productivity* (*economic* p_w), better indicated as *water return* (r_w) having the unit of "currency" (dollars, euro, etc.) per unit amount of water (volume, quantity, or depth).

LARGE AREA SURFACE ENERGY BALANCE ESTIMATION USING SATELLITE IMAGERY

Peyman Daneshkar Arasteh
Soil Conservation and
Watershed Management
Research Institute (SCWMRI)
Tehran, Iran

INTRODUCTION

Many activities such as irrigation and agricultural scheduling, water resources planning, weather forecasting, atmospheric circulation modeling, energy resources management and development, and even health related issues involve studying the surface energy balance (EB) of the earth. Photosynthesis and other metabolic processes in green plants, evaporation from free water and soil surfaces, wind effects, atmospheric and oceanic circulation, and many other phenomena are mainly functions of solar radiation. In other words, they are functions of surface EB, in which solar radiation is the most important driving force.

Due to nonuniform properties of the earth's surface (albedo, heat capacity, soil cover, etc.) and also due to temporal variation of the earth's surface properties, the earth's surface phenomena must be monitored on spatial and temporal scales. Due to spatial variations, ground-based monitoring of the earth surface can cover only small areas because of economic restrictions.

Generally, installing a dense monitoring network is not feasible or it is technically impossible. For example, first-order climatology stations may be 100 km apart. But variations in soil texture, soil cover, topography and microclimate may occur on a much smaller spatial scale. On the other hand, interpolation/extrapolation of point observations may be insufficient due to weak ground coverage. And, the need for distributed models increases.

Satellite imagery provides a large view of the earth's surface. Such a spatially continuous, but time discrete, data acquisition system provides a suitable framework for distributed modeling, provided that the phenomenon to be modeled has spectral behavior.

All features on or near the earth's surface, including soil surface, water bodies, vegetation cover, and near surface atmosphere, exhibit spectral characteristics. Therefore, remote sensing can monitor the earth's surface directly or indirectly. During the past two decades, several satellites were launched with a large variety of sensor capabilities. Many articles in the literature report using remote sensing to study the energy partitioning of the earth's surface. In this respect, Bastiaanssen (1) developed a model for energy partitioning using a remote sensing algorithm named SEBAL (surface energy balance algorithm for land). Bastiaanssen et al. (2) used SEBAL to model evapotranspiration on a catchment scale in Spain, Niger, China, and Egypt. Bastiaanssen (3) also tested the model in Turkey. Silberstein et al. (4) correlated the surface effective temperature with the satellite-derived surface temperature to form the energy balance on a catchment scale.

Remote sensing is being applied in a large number of investigations, such as Xin and Shih (5), Laymon et al. (6), and Biftu and Gan (7) to model surface evaporation and canopy evapotranspiration. Bastiaanssen et al. (8) introduced the possibilities of applications in agriculture, and Arasteh (9) developed a model to estimate evaporation from the Hamoon Wetlands called HRSE (Hamoons remotely sensed evaporation) within an area of 4000 km^2 on the Iran–Afghanistan border using NOAA-AVHRR images.

The advantages of using remote sensing in hydrologic studies have been stated by Prince (10) and Dubayah et al. (11) as follows:

1. large aerial coverage
2. higher spatial resolution than ground-point observations
3. availability of long-term records from area without ground observations
4. possibility of change detection

In this article, the EB is studied using remotely sensed data and some common methods are introduced. Then, the EB components of the Hamoon Wetlands are described as a case study.

ENERGY BALANCE

All climatic and hydrologic phenomena in the atmosphere and on the earth's surface are affected by the EB of the surface. The EB is one of the main subjects in climatology, hydrology, civil and traffic engineering, biology, and health. EB is the balance between the inward and outgoing energy components of the surface.

Understanding the behavior of the terrestrial hydrosphere, biosphere, and climatic systems and their possible changes requires knowledge of the nature of the EB (11). The EB of a surface is the radiative transfer of energy through the top of the earth's surface and the molecular energy transfer into the soil matrix or free waterbody (12). Fig. 1 shows the components of the EB.

As shown in Fig. 1, the EB has several components. The net radiation flux of the surface, R_n, is the algebraic sum of the following components:

$$R_n + A - S - G - \lambda E - \Delta Q = 0, \qquad (1)$$

where A, S, G, λE, and ΔQ are advective and convective sensible heat, soil heat, latent heat fluxes and change in heat storage, respectively. All of the fluxes have the dimension of MT^{-3} with units such as MJ/m^2-day. In Eq. 1, all sources of energy (R_n and A) are positive values, and sink terms (S, G, λE, and ΔQ) are negative.

There is also a radiation balance at the surface, in which the net radiation is calculated from the algebraic sum of the inward and outward radiations (Eq. 2).

$$R_n = R_s \downarrow - R_s \uparrow + R_l \downarrow - R_l \uparrow, \qquad (2)$$

where R_s and R_l are short- and long-wave radiation, respectively; ↓ and ↑, represent inward and outward radiation to and from the surface.

Using Stephan's law, Eq. 2 changes to (6,13):

$$R_n = (1-\alpha)R_s + C\sigma(\varepsilon_{sky}{T_a}^4 - \varepsilon_{surface}{T_S}^4) \qquad (3)$$

where α is surface albedo; ε_{sky} and $\varepsilon_{surface}$ are the emissivities of the atmosphere and the surface, respectively; C is the cloudiness factor; σ is the Stephan–Boltzman constant (4.903×10^{-9} MJ/K^4-m^2-day); and T_s and T_a, are surface and near surface atmosphere temperatures (K).

Other components of EB, S and λE, may be determined by aerodynamic resistance relations (6,13,14) and G can be modeled on the basis of heat conduction theory (12):

$$S = \frac{\rho C_p}{r_a}(T_s - T_a) \qquad (4)$$

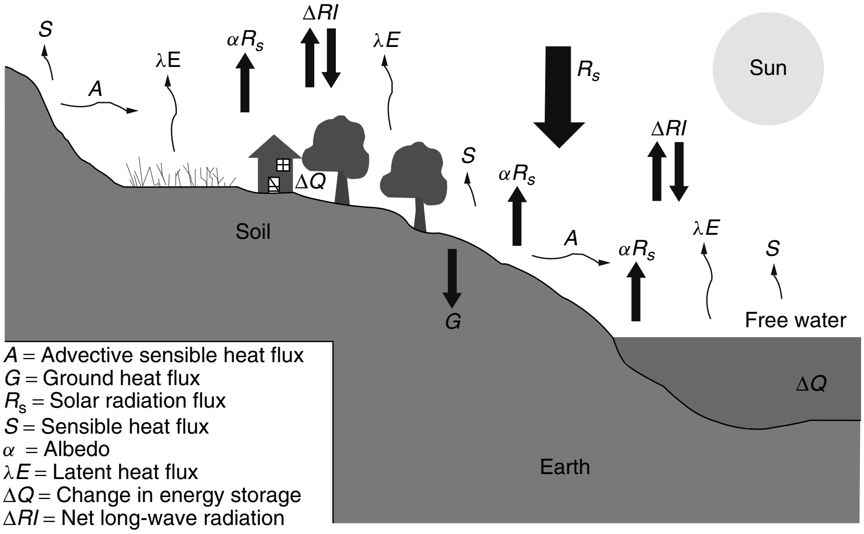

Figure 1. Schematic components of the EB.

$$\lambda E = \frac{\rho C_\mathrm{p}}{\gamma} \frac{e_\mathrm{S}(T_\mathrm{s}) - e_\mathrm{a}}{r_\mathrm{a} + r_\mathrm{s}} \tag{5}$$

$$G = dC_\mathrm{s} \frac{dT}{dt} \tag{6}$$

where ρ is air density (kg/m^3); C_p is the specific heat of air (MJ/kg-K); C_s is the soil matrix or waterbody heat capacity (MJ/m^3-K); γ is the psychometric constant; $\mathrm{r_a}$ and $\mathrm{r_s}$ are aerodynamic and surface resistances, respectively (s/m); e_s and e_a are saturated vapor pressure at T_s and actual air vapor pressure, respectively (kPa); d is the depth of soil or water (m); and dT/dt is the change in temperature per unit time.

Equations 3–6 have some variables that can be determined by remote sensing and others must be obtained from ground observations. The received solar radiation, R_s, is introduced in several texts on the basis of the earth–Sun geometric relationships (15). Determination of surface temperature and near surface air temperature, dew point, precipitable water in atmosphere column, vapor pressure deficit, and albedo are reviewed by Cracknell, (16), Prince et al. (10), Dubayah et al. (11), Menenti (17), Granger (18), and Schmugge et al. (19).

REMOTELY SENSED EB

Satellite remote sensing is a collection of sciences, technologies, and methods of inferring the earth's surface parameters by measuring the electromagnetic radiation of the surface (18). Visible (VIS), near infrared (NIR) and thermal infrared (TIR) channels are used for EB processing (16,17,19).

Some primary surface parameters such as surface albedo, surface temperature, air temperature, surface roughness, land cover, and land use may be provided by remote sensing. In the following, satellite data acquisitions are introduced.

Surface Albedo

Reflected directional radiance in VIS and NIR channels (0.4 to 2.5 μm) are being used to determine surface albedo (16–18). To determine surface albedo, Gutman (20) and Csiszar and Gutman (21) introduced the following methodology:

1. Normalizing VIS and NIR channels reflectances with respect to solar zenith angle (SZA):

$$\begin{aligned}(\rho_1)_\mathrm{TOA} &= R_1 \cos(\mathrm{SZA}) \\ (\rho_2)_\mathrm{TOA} &= R_2 \cos(\mathrm{SZA})\end{aligned} \tag{7}$$

   where R is reflectance; ρ is normalized reflectance; SZA is solar zenith angle (radians); 1 and 2 indexes represent VIS and NIR channels, respectively; and TOA is top of atmosphere.

2. Converting VIS and NIR narrowband reflectance to a broadband reflectance by a linear relation to determine top of atmosphere (TOA) reflectance:

$$(\rho)_\mathrm{TOA} = a + b(\rho_1)_\mathrm{TOA} + c(\rho_2)_\mathrm{TOA} \tag{8}$$

   where a, b, c are regression coefficients.

3. Correcting TOA reflectance for anisotropy to determine TOA albedo:

$$(\alpha)_\mathrm{TOA} = (\rho)_\mathrm{TOA}(\mathrm{ARF})^{-1} \tag{9}$$

   where ARF is the anisotropic reflectance factor.

4. Modifying TOA albedo for atmospheric effects to determine surface albedo:

$$\alpha = f[(\alpha)_\mathrm{TOA}, \mathrm{SZA,AWV,AA}] \tag{10}$$

   where α is albedo (%), AWV is atmospheric water vapor, and AA are atmospheric aerosols. Song and Gao (22) gave calibrated nonlinear relations between coefficients b and c and NDVI (normalized difference vegetation index) or SAVI.

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - R}{\mathrm{NIR} + R} \tag{11}$$

$$\mathrm{SAVI} = 1.5\left(\frac{\mathrm{NIR} - R}{\mathrm{NIR} + R + 0.5}\right) \tag{12}$$

where R is the red channel of the spectrum.

Instead of the linear relation of Equation 8 Granger (18) emphasized that the AVHRR-NIR channel could be used individually to determine the out of atmosphere reflectance and the corresponding albedo.

Surface Temperature

Surface temperature is the result of the thermodynamic equilibrium between the energy condition in the near surface atmosphere, soil surface, and subsurface soil. It is also a function of surface emissivity (19). The first step in inferring surface temperature by remote sensing is to determine the brightness temperature, T_B, of TIR channels. The T_B at satellite temperature is the temperature of a blackbody and can be determined by solving Planck's spectral distribution function for blackbody temperature (16):

$$R(T_\mathrm{B}) = \frac{\displaystyle\int_{\nu_1}^{\nu_2} E(\nu, T_\mathrm{B})\Phi(\nu)d\nu}{\displaystyle\int_{\nu_1}^{\nu_2} \Phi(\nu)d\nu} \tag{13}$$

where R is radiance (MJ/m^2-Sr-cm); T_B is brightness temperature (K); E is Planck's radiation energy of a blackbody (MJ/m^2-Sr-cm); Φ is spectral response function; and ν is wave number equal to the inverse of wavelength (cm^{-1}).

Spectral radiance, R, is determined from the sensor calibration equation on the basis of the image digital number. Equation 13 must be solved numerically for energy E. Then the brightness temperature, T_B, is determined by solving Planck's law (16):

$$T_\mathrm{B} = \frac{C_2\nu}{\ln\left(1 + \dfrac{C_1\nu^3}{E}\right)} \tag{14}$$

in which $C_1 = 1.1910659 \times 10^{-5}$ MJ/s-m^2-Sr-cm^{-4} and $C_2 = 1.438833$ [cm-K].

To determine surface temperature, there are several methods classified in three major categories:

1. Ground observation based calibration
2. Atmospheric effects back-modeling
3. Atmospheric effects omission (two-look method and multichannel method or split window)

Cracknell (18) has described these methods in detail. The split window method for determining surface temperature is the most common (10,18,23,24).

Air Temperature

Ground observation calibration methods can be used to determine air temperature. Air temperature can also, be derived from the thermal vegetation index (10,11) where the surface temperature of a closed canopy is assumed equal to its adjacent air temperature.

Sky and Surface Emissivities

Surface and near surface atmosphere emissions are related to surface indexes such as NDVI, SAVI, etc. (6,25):

$$\varepsilon_\mathrm{surface} = a_1 \cdot \mathrm{NDVI} + a_2 \tag{15}$$

$$\varepsilon_\mathrm{surface} = b_1 \cdot \ln(\mathrm{NDVI}) + b_2 \tag{16}$$

where a_1, a_2, b_1, and b_2, are regression coefficients.

Net Radiation, Sensible and Latent Heat Flux Densities

According to Eq. 3, from the earth–Sun geometry and satellite data acquisition, using previously mentioned methods, one can determine R_n. Using satellite derived T_s and T_a and ground based data, other energy components of the surface of Eqs. 4–6 can be determined. In this respect, surface and aerodynamic resistances and surface roughness length are related to LAI (leaf area index) and consequently to remote sensing by indexes such as NDVI or SAVI (3,25):

$$Z_{0\mathrm{m}} = \exp(a' \cdot \mathrm{NDVI} + b') \tag{17}$$

$$\mathrm{LAI} = a'' \cdot \ln(b'' + \mathrm{NDVI}) \tag{18}$$

$$\mathrm{LAI} = a''' \cdot \exp(b''' \cdot \mathrm{NDVI}) \tag{19}$$

$$\begin{aligned} r_\mathrm{s} &= f(\mathrm{LAI}) \\ r_\mathrm{a} &= f(Z_{0\mathrm{h}}) \end{aligned} \tag{20}$$

where $Z_{0\mathrm{m}}$ and $Z_{0\mathrm{h}}$ are the surface roughness length due to turbulence and heat, respectively; r_a and r_s, are aerodynamic and surface resistances; and a', a'', a''', b', b'', and b''' are regression coefficients.

Case Study

To show an example of a remotely sensed EB application, the study by Arasteh (9) on Hamoon Wetlands is summarized as follows:

The Hamoon Wetlands located between Iran and Afghanistan form a complex hydraulic-hydrologic system of three neighboring wetlands with an average water area of 4000 km^2 in wet years. The wetlands are transboundary and affect all human activities in the region (Fig. 2).

Determining an estimate of evaporation from the free water surface of the Hamoon Wetlands was the main objective of the study by Arasteh (9). For this, it was required to study remotely sensed EB. Arasteh (9) used energy partitioning to fulfill this objective. Using his algorithm with the NOAA-AVHRR-14 image of May 10, 1998, the net radiation, sensible heat, and latent heat of the Hamoon surface are shown in Fig. 3.

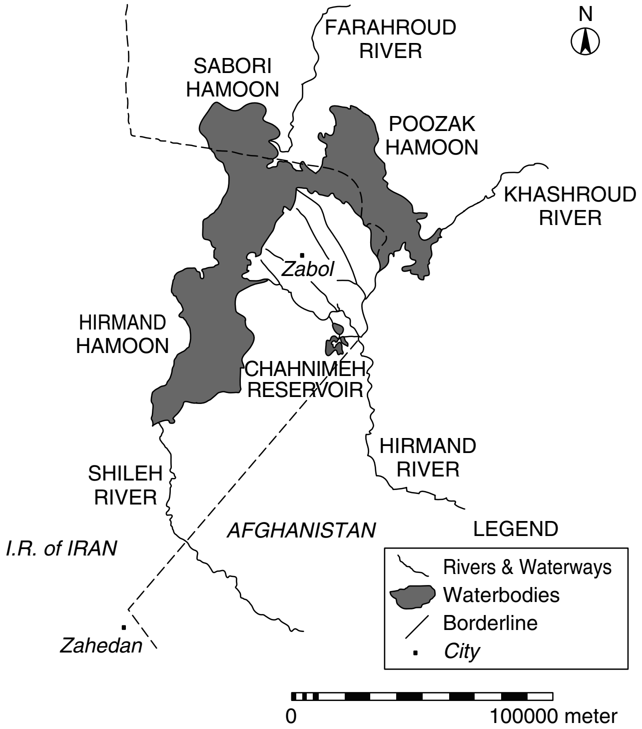

Figure 2. Hamoon Wetlands region between Iran and Afghanistan.

DATA ACCURACY

To use each method, it is necessary to consider its accuracy. Therefore, it is very important to consider the question of the accuracy of remotely sensed EB.

Considering Eqs. 1 and 3–6, there are various variables with their own sources of error. Cracknell (16) reported a variation of 0.5 to 1.4 K in the sea surface temperature for only a 2% change in surface emissivity. Menenti (17) reviewed the literature and reported errors of less than 3% in latent heat and obtained an average error of 15% for this component. Dubayah et al. (11) stated that it is not possible to estimate downwelling long-wave radiation with accuracy less than 20 W/m^2 using AVHRR data.

Finally, it can be concluded that the accuracy of a remotely sensed EB is related mainly to the spatial resolution of images, the spectral resolution (number of channels and bandwidth for each channel), the accuracy of ground-based observations, and the range of applicability of calibrated equations.

SOURCES OF SATELLITE EB

Carleton (26) summarized some sources of satellite based EB (Table 1). He emphasized METEOSAT, GMS, and ERBS satellite systems as the most important sources of global coverage of EB. The ERBS is a professional system. The Earth Radiation Budget Experiment (ERBE) included ERBS and NOAA 9–10 platforms.

Table 1. Some Satellite Systems Related to EB Generation[a]

Variable	Spectral Region	Sensor	Satellite	Type of Satellite
Surface albedo	VIS	SR	NOAA	Polar orbiter
	NIR	AVHRR	NOAA	Polar orbiter
		VISSR	GOES	Geostationary
Surface solar irradiance	VIS	SR	NOAA	Polar orbiter
		VISSR	GOES	Geostationary
Outgoing long-wave Radiation	TIR	SR	NOAA	Polar orbiter
		AVHRR	NOAA	

[a] Reference 26.

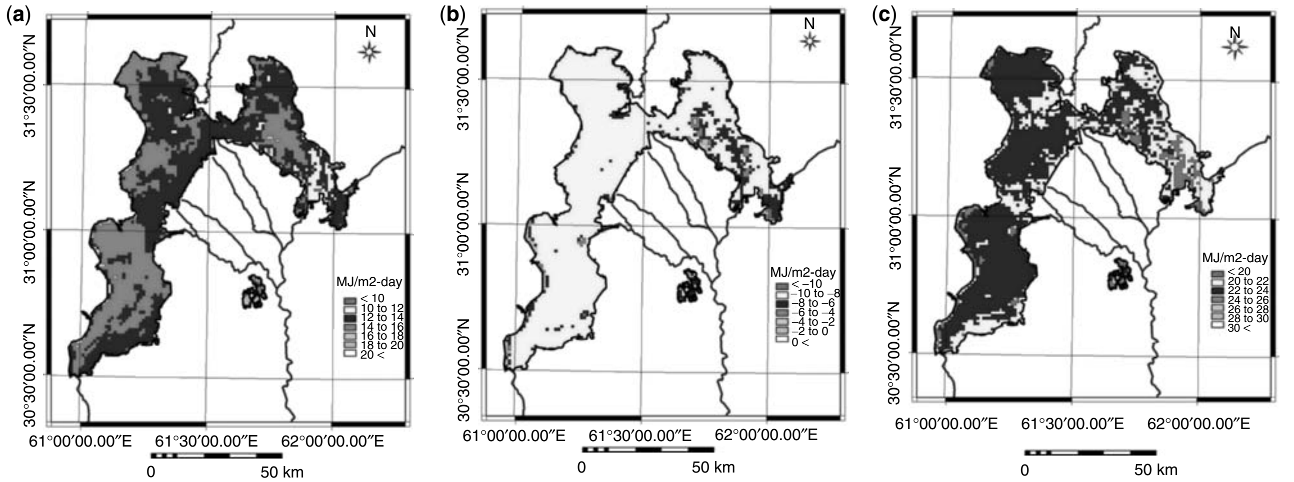

Figure 3. Energy partitioning for Hamoon Wetlands using the Arasteh (9) algorithm; (**a**) net radiation, (**b**) sensible heat, and (**c**) latent heat.

RELATED INTERNET SITES

1. http://www.saa.noaa.gov/nsaa/products/.
2. http://www.osdpd.noaa.gov/PSB/EPS/RB/.
3. http://orbit-net.nesdis.noaa.gov/goes/gcip/.
4. http://www.cdc.noaa.gov/~map/maproom/text/Climate_Pages/olr.shtml.
5. http://poes.nesdis.noaa.gov/climate/.
6. http://www.ncdc.noaa.gov/.
7. http://www.eurometeo.com/.

BIBLIOGRAPHY

1. Bastiaanssen, W.G.M. (1995). *Regionalization of Surface Flux Densities and Moisture Indicators in Composite Terrain*, Ph.D. Thesis, Agricultural University, Wageningen, the Netherlands.
2. Bastiaanssen, W.G.M. et al. (1998). A remote sensing surface energy balance algorithm for land (SEBAL): 2. Validation. *J. Hydrol.* **212–213**: 213–229.
3. Bastiaanssen, W.G.M. (2000). SEBAL-based sensible and latent heat fluxes in the irrigated Gediz Basin, Turkey. *J. Hydrol.* **229**: 87–100.
4. Silberstein, R.P., Sivapalan, M., and Wyllie, A. (1999). On the validation of a coupled water and energy balance model at small catchment scales. *J. Hydrol.* **220**: 149–168.
5. Xin, J.N. and Shih, S.F. (1991). NOAA polar-orbiting satellite APT data in lake evaporation estimation. *J. Irrig. Drainage Eng.* **117**(4): 547–557.
6. Laymon, C. et al. (1998). Remotely-sensed regional-scale evapotranspiration of a semi-arid Great Basin desert and its relationship to geomorphology, soils and vegetation. *Geomorphology* **21**: 329–349.
7. Biftu, G.F. and Gan, T.Y. (2000). Assessment of evapotranspiration models applied to a watershed of Canadian Prairies with mixed land-uses. *Hydrological Processes* **14**: 1305–1325.
8. Bastiaanssen, W.G.M., Molden, D.J., and Makin, I.W. (2000). Remote sensing for irrigated agriculture: Examples from research and possible applications. *Agric. Water Manage.* **46**: 137–155.
9. Arasteh, P.D. (2004). *Development of a Distributed Model for Estimation of Regional Evaporation using Remote Sensing Techniques, Case Study: Hamoon Lakes of Iran*, Ph.D. Thesis, Tarbiat Modarres University, Tehran, Iran.
10. Prince, S.D. et al. (1998). Inference of surface and air temperature, atmospheric precipitable water and vapor pressure deficit using advanced very high- resolution radiometer satellite observations: Comparison with field observations. *J. Hydrol.* **212–213**: 230–249.
11. Dubayah, R.O., Wood, E.F., Engman, E.T., Czajkowski, K.P., Zion, M., and Rhoades, J. (2000). Remote sensing in hydrological modeling, In: *Remote Sensing in Hydrology and Water Management*. G.A. Schultz and E.T. Engman (Eds.). Springer-Verlag, Berlin, Germany, pp. 85–102.
12. Stull, R.B. (1988). *An Introduction to Boundary Layer Meteorology*. Kluwer Academic, Dordrecht, the Netherlands, pp. 251–294.
13. Engman, E.T. and Gurney, R.J. (1991). *Remote Sensing in Hydrology*. Chapman and Hall, London, UK, pp. 85–102.
14. Albertson, J.D., Kiely, G., and Parlange, M.B. (1996). Surface fluxes of momentum, heat and water vapor. In: *Radiation and Water in Climate System*. E. Raschke (Ed.). Springer-Verlag, Berlin, Germany, pp. 59–82.
15. Burman, R. and Pochop, L.O. (1994). *Evaporation, Evapotranspiration and Climate Data*. Elsevier, Amsterdam, the Netherlands, pp. 28–54.
16. Cracknell, A.P. (1997). *The Advanced Very High Resolution Radiometer (AVHRR)*. Taylor and Francis, London, UK.
17. Menenti, M. (2000). Evaporation. In: *Remote Sensing in Hydrology and Water Management*. G.A. Schultz and E.T. Engman (Eds.). Springer-Verlag, Berlin, Germany, pp. 157–188.
18. Granger, R.J. (2000). Satellite derived estimates of evapotranspiration in Gediz Basin. *J. Hydrol.* **229**: 70–76.
19. Schmugge, T.J. et al. (2002). Remote sensing in hydrology. *Adv. Water Resour.* **25**: 1367–1385.
20. Gutman, G. (1988). A simple method for estimating monthly mean albedo of land surface from AVHRR data. *J. Appl. Meteorol.* **27**: 973–984.
21. Csiszar, I. and Gutman, G. (1999). Mapping global land surface albedo from NOAA-AVHRR. *J. Geophys. Res.* **104**: 6215–6228.
22. Song, J. and Gao, W. (1999). An improved method to derive surface albedo from narrowband AVHRR satellite data: Narrowband to broadband conversion. *J. Appl. Meteorol.* **38**: 239–249.
23. Andersen, H.S. (1997). Land surface temperature estimation based on NOAA-AVHRR data during the HAPEX-Sahel experiment. *J. Hydrol.* **188–189**: 788–814.
24. Watts, C.J. et al. (2000). Comparison of sensible heat flux estimates using AVHRR with scintillometer measurements over semi-arid grassland in northeast Mexico. *Agric. Forest Meteorol.* **105**: 81–89.
25. Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., and Holtslag, A.A.M. (1998). A remote sensing surface energy balance algorithm for land (SEBAL): 1. Formulation. *J. Hydrol.* **212–213**: 198–212.
26. Carleton, A.M. (1991). *Satellite Remote Sensing in Climatology*. Belhaven Press, London, UK, pp. 205–244.

SOIL EROSION AND CONTROL PRACTICES

MARIANNE MCHUGH
MICHELLE CLARKE
JULIA DUZANT
ROY MORGAN
JANE RICKSON
Cranfield University
North Wyke, Devon, United Kingdom

INTRODUCTION

Soil erosion is the removal of soil or other slope-forming materials by erosive agents, namely, water and wind. Accelerated soil loss, in excess of the rate of natural soil erosion, is perpetuated when surface vegetation is compromised or removed. It is therefore more likely to occur wherever land is used by humans for agricultural production, infrastructural construction, mineral abstraction, or recreation, and where measures for soil protection are not in place.

Soil erosion involves a two-phase process of soil detachment and transport, both of which require an energy input. Much of the energy of erosion derives from water or wind flow over the soil surface or is provided by gravity. Although reducing the energy of erosive forces is the most effective way of controlling soil erosion, one consequence of reducing energy may be the deposition of any material already eroded. As eroded sediment is often enriched with nutrients, organic matter, and contaminants preferentially adsorbed onto the sediment surface, sediment deposition may allow contaminant release into the environment. It is necessary, therefore, to reduce erosive energy at the source areas of sediment rather than after the sediment has been detached and transported.

SOIL EROSION

Water Erosion

Rainsplash (detachment and removal of soil particles and aggregates at the soil surface by raindrop impact) is an important soil-detaching agent. The energy available for erosion is related to raindrop size and velocity. Raindrops have greater kinetic energy than the same volume of runoff and are thus potentially more erosive, although most raindrop energy is expended in soil detachment, leaving little energy available for sediment transport. Rainsplash plays a further role in slope erosion through the detachment, dispersion, and redeposition of fine soil particles on the soil surface. Soil pores are blocked by the redeposited particles, resulting in a surface seal, a reduction in infiltration, and an increase in erosive runoff.

Rainsplash constitutes only a negligible proportion of total soil loss except where runoff is not generated, where it can be the dominant process of sediment removal (1). Where runoff does occur, particle suspension is greatly enhanced by raindrop impact (2) and rainsplash operates as a “feeder” mechanism to rill and runoff erosion. Runoff may also present a buffer between falling raindrops and the soil surface, so that detachment by rainsplash can be limited by runoff. This depends on water depth and raindrop size (3).

Runoff in unchanneled form has a relatively low kinetic energy and, while able to transport soil particles thrown into the flow by raindrop impact, is not able to detach new material. Hortonian runoff (when rainfall intensity exceeds the infiltration capacity of the soil) depends critically on rainfall intensity and amount and soil infiltration capacity. Saturation runoff (when rainfall amount exceeds soil storage capacity) depends on rainfall intensity, subsoil permeability, and the porosity and permeability of the topsoil.

Runoff channeled into rills or gullies has higher kinetic energy and is a much more powerful agent, able to detach and transport soil particles of all sizes, including gravels and small stones. Its considerable erosive power can account for up to 78% of the total material transported within an erosion event.

Wind Erosion

The main factor affecting wind erosion is the velocity of moving air. Because of the roughness imparted by vegetation, soil, and other obstacles, wind speeds are lowest nearest the ground surface. Above this, wind speed increases exponentially with height. The movement of soil particles can be related to critical shear velocity, which is the velocity of wind required to dislodge soil particles from the ground surface and is dependent on soil particle size and soil cohesiveness. Very high shear velocities are required to initiate soil particle movement by wind, and this energy may be translated to adjacent soil particles at the same time, thus causing a chain reaction of motion. Sediment-laden air also acts to detach and mobilize other soil particles by bombarding or impacting on the soil surface. Once in motion, soil particles are transported by wind in suspension (in the air), as surface creep (rolling along the ground surface), and as saltation (particle movement in jumps) (4).

Rates of Soil Erosion

Rates of soil erosion are normally expressed in terms of the mass or volume of soil removed from a unit area of land per unit of time (t or $m^3{\cdot}ha^{-1}{\cdot}y^{-1}$). Typical annual rates of erosion on hillslopes under forest or natural grassland are 0.01–0.5 $t{\cdot}ha^{-1}$ in areas of moderate slope steepness and 0.1–5 $t{\cdot}ha^{-1}$ on steeper slopes. These rates are generally considered low and representative of erosion as a natural geomorphological process. On agricultural land, annual rates are much higher, ranging from 0.1 to 200 $t{\cdot}ha^{-1}$, reflecting the poorer protection afforded by various crops. Arable land is particularly vulnerable to erosion seasonally when the soil is bare between ploughing and crop emergence and also from emergence until the crop canopy closes. Grazing land is vulnerable if overgrazing removes the vegetation cover. Even higher rates of erosion have been recorded on construction sites, often reaching 1000 or even 5000 $t{\cdot}ha^{-1}{\cdot}yr^{-1}$.

Since erosion is a natural process, it is implicit that it cannot be prevented. Erosion can, however, be controlled at an acceptable rate, which should be no faster than the rate at which new soil forms. As little is known about rates of soil formation, except that they are extremely slow, a value of 1 $t{\cdot}ha^{-1}$ (equivalent to 0.1 $mm{\cdot}y^{-1}$) is sometimes accepted as an average annual rate for temperate areas. A more pragmatic approach is to consider the environmental damage that can occur if no control measures are in place (Table 1). Decisions can then be made about the level of damage tolerable relative to the costs of preventing that damage.

Impacts of Soil Erosion

Soil is a valuable resource *in situ* but when eroded by wind and water it can cause significant impacts both on and off site. The on-site implications of soil erosion include declines in cultivable soil depth, available soil moisture, soil organic matter, nutrient content, soil fertility, and soil structure as well as large-scale losses in productive area through gullying or mass movement (4). These on-site effects are particularly important on agricultural land:

Table 1. Rates of Erosion at a Field Scale and Associated Levels of Damage from Individual Storm Events

Erosion Rate, $t{\cdot}ha^{-1}$	Class	Indicators
<1	Natural	No surface evidence of erosion; land well-protected by vegetation cover (90–100% ground cover); very low sediment concentrations in runoff water, which appears clear.
1–2	Very slight	No surface evidence of erosion; land protected by vegetation cover (70–90% ground cover); low sediment concentrations in runoff water, which is clear to slightly cloudy.
2–5	Slight	Some crusting of the soil surface; localized wash, small channels (rills) every 50–100 m spacing; 30–70% ground cover; slight risk of pollution if runoff discharges directly into water courses.
5–10	Moderate	Evidence of surface wash; rills spaced every 20–50 m; exposed tree roots mark level of former soil surface; slight to moderate surface crusting; 30–70% ground cover; risk of pollution if runoff discharges directly into water courses.
10–50	High	Continuous network of rill channels every 5–10 m or deeper channels (gullies) every 50–100 m; surface crusting of soils over large area; <30% ground cover; danger of pollution and sedimentation downstream.
50–100	Severe	Land scarred by continuous network of rills every 2–5 m or gullies every 20 m; sediment splays of coarse material; bare soil; siltation of water bodies; damage to roads by erosion and sedimentation; risk of muddy floods.
100–500	Very severe	Land heavily scarred by continuous network of channels with gullies every 5–10 m; soil heavily crusted; bare soil; severe siltation, pollution, flooding, and eutrophication problems downstream.
>500	Catastrophic	Extensive network of rills and gullies; large gullies (>100 m^2) every 20 m; most of original soil surface removed; severe damage from erosion, pollution, flooding, and sedimentation downstream.

annual agricultural damage costs due to erosion have been estimated at £264 million in the United Kingdom alone (5). The net effect of these impacts is a loss in productivity, which both restricts what can be grown and results in increased expenditure on fertilizers and irrigation to maintain yields. Later this threatens food production and may ultimately lead to land abandonment, with a concurrent decline in the value of land as it changes from productive farmland to wasteland (4). A further consequence of this is the unsustainable exploitation of land resources elsewhere.

The deposition of eroded sediment reduces the capacity of rivers and drainage ditches, enhances flooding, blocks irrigation canals, and shortens the design life of reservoirs (4). Deposited sediment also represents a risk to infrastructure. Chemicals, including fertilizers and pesticides, and nutrients such as nitrogen (N) and phosphorus (P) deposited with eroding soil may have detrimental effects on water quality for supply, recreation, and fisheries. Further impacts on aquatic ecosystems include increases in turbidity reducing light penetration and O_2 availability, thereby reducing plant growth and key biological activities such as photosynthesis; submergence and physical damage to floodplain and in-stream plants; and clogging of fish spawning gravels.

CONTROLLING SOIL EROSION

Measures to control erosion rely on reduction of the energy of erosive forces. Reducing the rate at which soil is detached and transported involves increasing interception and infiltration of rainfall and runoff; reducing runoff volume, velocity, and energy; decreasing the area over which runoff occurs; increasing soil resistance; and increasing soil aggregate stability. The practical measures available to achieve these conditions include mechanical, agronomic, and soil management methods of erosion control.

Mechanical Methods of Erosion Control

Mechanical methods of soil conservation are useful in the control of soil erosion by water. They include terraces, constructed on hill slopes to control soil erosion by reducing slope length and slope steepness, thus limiting the energy of runoff to detach and transport soil material (6). There are many different types of terrace but generic types include the channel and bench terrace. Channel terraces are suited to slopes no steeper than 7° (12%) (6) and consist of a cut channel and filled ridge (constructed from the cut material) constructed across the slope. The distance between consecutive terraces downslope is influenced by site conditions, with greater distances between terraces on gentler slopes and shorter terrace intervals on steeper slopes.

Bench terraces are steplike structures constructed across a slope to prevent erosion by dissecting the hill slope into smaller, shorter slope components, each with a level or near-level bench. The dimensions of the benches depend on slope steepness, the desirable width of cultivable land, and the acceptable height of the steep wall (riser) of the terrace, which is often related to soil depth. As construction of the level bench requires oversteepening of the terrace riser, this near-vertical wall must be supported with stones, masonry, or timber to prevent collapse.

In channel terrace systems, runoff is discharged into a channel constructed at a gentle gradient across slope to ensure that erosion within the channel itself does not occur. The channels are usually kept free of vegetation, so that any accumulated sediment can be excavated each year. From there, flow is directed to lined waterways, which are designed to cope with increased volume and velocity of runoff without causing erosion. These waterways, ideally located in natural depressions to minimize the need for earth moving, are lined with vegetation, stones, or geotextiles with high shear resistance to minimize the risk of erosion.

Agronomic Methods of Erosion Control

Water Erosion Control. Agronomic methods involve the use of vegetation to control soil erosion and are highly effective (7). A plant canopy intercepts rainfall, changing raindrop size distribution and reducing fall velocities, so that kinetic energy available for erosion via raindrop impact is reduced. Plant stems distort and impart roughness to runoff paths, reducing flow velocity and thus kinetic energy to detach and transport soil material. Plant roots encourage infiltration through the soil profile and thus reduce the risk of runoff generation. Lateral plant roots also increase the resistance of soil against the shearing forces of runoff.

Agronomic methods of erosion control include ensuring maximum cover on the soil surface, especially during high rainfall periods, to protect the soil from erosive raindrops and maintaining high stem and root densities to reduce runoff volumes, velocities, and energy. These can be achieved through cover cropping, high density planting, crop rotations, mulching, and erosion control geotextiles. A geotextile is any permeable textile material that is an integral part of a constructed project, structure, or system. It may be made of synthetic or natural fibers. Geotextiles have many applications in land and water management, but those used for erosion control are usually in the form of two- or three-dimensional mats, webs, nets, grids, or sheets.

Natural fiber geotextiles are biodegradable and designed to be laid over the surface of the soil to provide protection against erosion until a vegetative cover is established: in this way, geotextiles act as simulated vegetation. Artificial fiber geotextiles are designed to give permanent protection to the soil surface and are usually buried to reinforce the soil both before and after vegetation establishment. Surface-laid, natural fiber geotextiles are more effective at controlling soil erosion by raindrop impact because they provide good surface cover, high water absorption, and a rough surface on which water can pond. They also reduce erosion by runoff because they impart roughness to the soil surface and therefore reduce runoff velocity. Buried geotextiles reduce erosion to a lesser extent as they cannot control rainsplash erosion and actually increase erosion if they are backfilled with loosely consolidated, and therefore highly erodible, material. Over time, natural fiber geotextiles biodegrade and become less effective as vegetation establishes on the slope while artificial fiber mats continue to perform.

These measures also control soil erosion by contributing nutrients, organic matter, and carbon to the soil, thus encouraging microbial activity and improving soil structure and aggregate stability, which in turn further reduces the potential for soil erosion. Vegetation can also be used to reduce slope length and steepness (and thus erosion risk) when grown in strips parallel to the contour. The strips may be of different crops (strip cropping) or as a single field strip across the slope. The strips restrict the potential for runoff generation and encourage deposition of eroded sediment within the strips so that local slope gradients may be reduced. Common in areas where field boundaries (e.g., hedges) have been removed to increase field size and allow agricultural intensification, field grass strips can be established along former field boundaries, reinstating shorter slope lengths and reducing erosion risk. As farm machinery can traverse the field strips, this simple but effective soil erosion control measure may be adopted willingly by farmers.

Agroforestry systems, in which perennial trees or shrubs grow alongside annual crops, also help to control soil erosion. The trees or shrubs form across-slope barriers to runoff, thus reducing slope lengths and soil losses on the slope.

Wind Erosion Control. Vegetation is also widely used to control wind erosion. By exerting a drag on air flow, vegetation can reduce the velocity of the wind below that required to initiate erosion and can enhance the deposition of soil already being carried in the air. The most common use of vegetation is in the form of shelter belts spaced at regular intervals at right angles to the direction of the erosive winds. Traditionally, trees have been used to form shelterbelts but they need to be supplemented by shrubs (bushes) and grasses to provide cover close to the ground surface since some 70–75% of the soil eroded by wind is carried within the lower 1 m of the atmosphere. The most important characteristic of a shelterbelt is its density or porosity. If too dense, the belt acts like a solid wall and air flowing over the top will create erosive eddies in the lee. The belt therefore needs to be porous so that some air passes through the belt but at a low velocity. Conversely, if the belt is too sparse, the velocity of the wind is not reduced. A porosity of 40–50% is generally considered optimum. Belts are usually designed to maintain wind speeds below 75% of the speeds recorded in open ground. To achieve this, they need to be spaced at intervals of between 12 and 17 times the belt height.

While many shelterbelts comprise single rows of trees, some consist of three rows made up of alternating nurse trees, durable trees, and bushes. The advantage of these mixed belts is they can withstand the loss of an individual species due to disease or pest since other species within the belt can fill any gaps. They also provide a habitat for wildlife.

In arable farming, shelterbelts are considered boundary features and are placed around field edges. However, the principles of shelterbelts can also be used to design in-field shelter systems in which crops are alternated on a strip system, aligned at right angles to the wind. Thus, strips of barley, spaced at 5–10 m intervals, can be alternated with strips of carrots, onions, or sugar beet to control erosion. The barley is sown in late autumn to allow sufficient time for it to emerge before the main crop is drilled in the spring. The barley strips reduce wind velocity until the cover of the main crop is sufficient to protect the soil, at which point the barley is killed off with a selective herbicide so that it does not compete with the crop.

Since shelterbelts require a certain amount of time to grow and become effective, they are not appropriate where immediate control of wind erosion is required. In these situations, windbreaks made from brushwood or plastic netting can be used. The design principles remain the same, aiming for 40–50% porosity and spacings of 12–17 times the windbreak height. Brushwood fences are commonly constructed on coastal dunes to reduce wind velocities and stabilize the mobile sand, which can then be seeded or planted with shrubs and trees to provide long-term stability.

Soil Management Methods of Soil Conservation

The inherent susceptibility of a soil to erosion depends on the soil type and this is difficult to adjust. However, soil management can affect other factors that can increase or decrease soil erosion. For example, intensive arable production can lead to severe soil degradation as soil is broken down by numerous, invasive tillage practices into a loose, fine tilth seedbed. The resulting small soil aggregates are highly susceptible to dispersion, detachment, and transport by rainfall, runoff, and wind. Soil structural stability is further reduced by loss of carbon through oxidation as tillage exposes deeper soil layers to the air, and by loss of soil faunal biomass destroyed during soil disturbance.

The way in which soil is managed in arable production systems therefore has important impacts on soil erosion control. Farming operations carried out up and down the slope will increase the risk of runoff generation while contour cultivations will reduce these risks significantly. Contour ridging enhances this principle by creating raised ridges across slope to reduce local slope length, reduce effective catchment area for runoff generation, and intercept any runoff that is generated. A development of contour ridging is the use of "tied" ridges, where the ridges are joined at regular intervals by "ties" running up and down slope, so creating small basins that trap runoff and eroded soil.

Reducing the number and intensity of cultivations may preserve soil structural stability and aggregation. Less disturbed aggregates are able to resist the dispersive forces of rainfall, runoff, and wind as cohesive bonds within the aggregates are maintained. Soil microorganisms and macroinvertebrates, which also help bind soil particles together and increase aggregate stability, are also less disturbed. Fewer, less invasive cultivations mean more of the vegetative biomass remains on or within the soil (in the form of standing stubble, cut stems, and roots) and is available to absorb the rainfall and runoff energy otherwise available for soil detachment and transport.

The relative reduction in soil disturbance and retention of vegetative material on and within the soil are the principles behind the use of conservation tillage for the control of soil erosion. This umbrella term incorporates a continuum of soil management methods such as minimum tillage, zero tillage, mulch tillage, direct drilling, noninversion tillage, subsurface mulch tillage, and plough/plant cultivation. Conservation tillage aims at maintaining soil fertility and controlling erosion but has an added aim of reducing input costs to agriculture, as fewer, less invasive cultivations expend less energy and fuel costs.

Manipulation of the soil surface condition by tillage is used to control local wind speeds at the ground surface in an attempt to control soil loss by wind erosion. On light, friable soils such as sands and silts, primary cultivations (ploughing and pressing) are performed in one direction, followed by drilling at 90° to this. The resultant tilth is rough and cloddy and imparts maximum friction to air flow, thus reducing wind speed and the ability to detach soil. Tillage techniques can also be used to minimize water erosion, for example, the Aqueel minimizes erosion by maintaining water within surface depressions on the soil surface (8).

Soil conditioners applied to soils susceptible to erosion can perform the dual role of enhancing soil fertility and controlling erosion. Products with high organic matter (such as seaweed extract and tree bark) add nutrients and organic matter to inherent soil fertility, so improving aggregation, structure, porosity, infiltration capacity, and water holding capacity. Soil conditioners reduce soil erosion by either increasing runoff infiltration or increasing resistance to soil detachment. Hydrophilic soil conditioners include bark, limestone, palm oil effluent, and seaweed extract, which increase aggregation and infiltration capacity, so runoff is less likely to occur. The increase in organic matter from these products also helps control erosion by increasing cohesive bonds in the soil, thus improving the stability of aggregates against soil erosion processes. Hydrophobic soil conditioners, including latex, emulsions, and molasses, form a water-repellent film on the soil surface and increase soil resistance to detachment and transport by rainsplash and runoff.

CONCLUSION

The accelerated loss of soil through water and wind erosion processes is a globally important phenomenon with significant and wide-ranging implications for land and water quality. Controlling the impacts of soil erosion is therefore imperative. A range of control measures are presented, illustrating that soil erosion control can be effectively addressed from a mechanical, agronomic, or soil conservation perspective, depending on the environment, skills, time, and resources available.

BIBLIOGRAPHY

1. Imeson, A.C. (1977). Splash erosion, animal activity and sediment supply in a small forested Luxembourg catchment. *Earth Surf. Proc. Landforms* **2**(2–3): 153–160.
2. Bryan, R.B. (1979). The influence of slope angle on soil entrainment by sheet wash and rainsplash. *Earth Surf. Proc.* **4**: 43–58.
3. Palmer, R.S. (1964). The influence of a thin water layer on water-drop impact forces. *Int. Assoc. Sci. Hydrol. Publ.* **65**: 141–148.
4. Morgan, R.P.C. (1995). *Soil Erosion and Conservation*, 2nd Edn. Longman, London.
5. Environment Agency. (2002). *Agriculture and Natural Resources: Benefits, Costs and Potential Solutions*. Peterborough, UK.

6. Sheng, T.C. (1972). A treatment-oriented land capability classification scheme. In: *Report on the Latin American Watershed Management Seminar*. FAO No. TA3112.
7. Coppin, N.J. and Richards, I.G. (1990). *Use of Vegetation in Civil Engineering*. Butterworths, London.
8. Simba, X. (2004). The Aqueel and soil erosion. Available: http://www.aqueel.com/.

WATER TABLE CONTRIBUTION TO CROP EVAPOTRANSPIRATION

Chandra Madramootoo
Macdonald Campus of McGill University
Ste-Anne de Bellevue, Quebec, Canada

Evapotranspiration (ET) is the transfer of water vapor from the soil root zone to the atmosphere via the process of photosynthesis. Water vapor can also move directly from the ground surface to the atmosphere. The rate of ET depends on several climatic, soil water, and crop factors, including rainfall, relative humidity, wind speed, air temperature, solar radiation, cloud cover, soil moisture, and stage of crop growth. The plant transpires at the potential rate if soil water is not limiting, which depends on the soil water content.

Water is held in the soil root zone in three distinct phases (see Fig. 1): hygroscopic water, capillary water, and gravitational water. Because of the high tension of the hydroscopic water, it is not available to plants. Both capillary water and gravitational water are available for crop uptake in the ET process.

As the soil water near the soil surface evaporates, the water content near the surface is reduced and the moisture tension increases. Consequently, the tension gradient exceeds the gravitational gradient, and moisture will move upward from the water table toward the soil surface. If root zone soil moisture is between field capacity and saturation, the crop will transpire at the potential rate. In this case, gravitational water contributes to the ET. However, once the soil moisture drops below field capacity, ET is met from the capillary water. In shallow water table soils, the water table contributes to part of the ET, which is known as capillary rise or upward flux. Depending on the position of the water table, the soil type, and the potential ET, part or all of the crop ET may be met through upward flux. If all ET is not met by upward flux, the remainder is taken from the capillary water. As the gravitational water is depleted, and if there is no supply of water to the soil profile either by precipitation or irrigation, the water table will continue to drop. The water table can be further lowered by drainage and seepage. In arid regions, if irrigation is limited, crops depend on the upward flux of groundwater to meet ET demands. However, in regions of deep groundwater, the upward flux is very small, lower than crop ET, and the plants begin to wilt. They thus exhibit signs of drought stress.

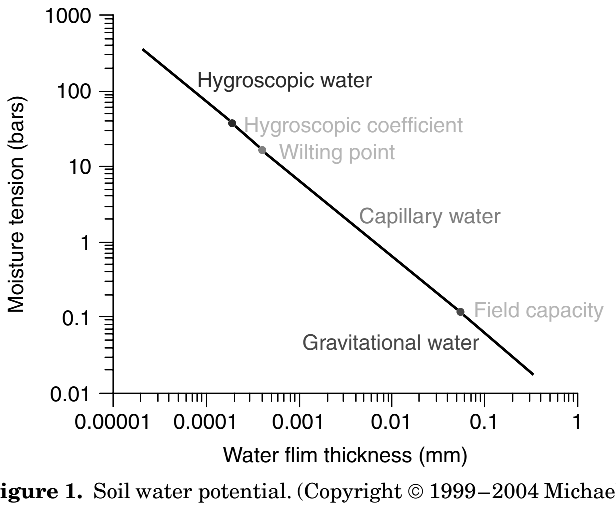

Figure 1. Soil water potential. (Copyright © 1999–2004 Michael Pidwirny).

The basic capillary rise equation ($h = 2\gamma \cos\alpha/\rho g R$) governs the height of the capillary rise (h) and is a function of the pore size distribution of the soil particles. Gardner (1) proposed the following equation to estimate the steady upward flux from a static water table:

$$q_{\max} \approx Aad^{-n} \qquad (1)$$

where $q_{\max}$ is the maximum evapotranspiration rate; A, a, and n are constants (n is a function of soil texture and is higher for coarser textured soils); and d is the depth to the water table.

Equation 1 shows that the ET rate decreases with water table height and will decrease more strongly for coarse textured soils. When the water table is close to the soil surface (low values of d), the maximum ET will be limited by the prevailing climatic conditions. At lower water table depths (high values of d), the ET will be limited by the transmissivity of the soil. Obreza and Pitts (2) developed relationships between upward flux and water table depth for four locations in Florida (Fig. 2).

The ability to meet crop ET from shallow water tables is used in a process known as subsurface irrigation (subirrigation). By making use of perforated plastic pipes installed below the soil surface (primarily for drainage), drainage outflows can either be restricted or water can be pumped back into the plastic pipes during the growing season, thus maintaining a shallow water table. Upward flux from the water table helps to meet the crop ET demands, which is a high energy and water efficient method of irrigation. Water application efficiency can be as high as 95%. It is also a low-cost irrigation system, because there is no need for additional investments in aboveground hardware. The only additional cost is a structure on the tile drain outlet.

The height at which the water table ought to be maintained depends on the crop. Shallow-rooted vegetable crops will need a water table of 30–50 cm below the soil surface. For deeper rooted crops, e.g., corn and soybeans, a water table of 60–75 cm will meet the crop ET requirement. Close monitoring of the water table depth is advantageous, especially in the wetter humid

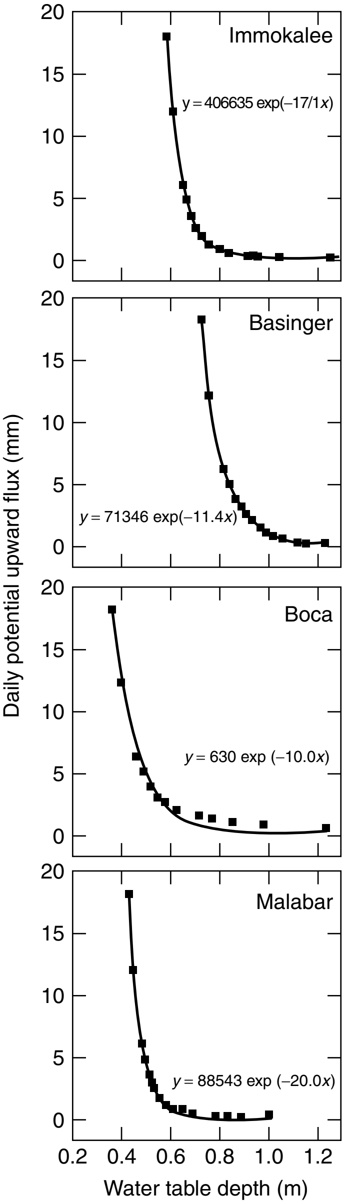

Figure 2. Upward flux vs. water table depth for 4 locations in Florida (2).

regions. Should heavy rainfalls occur, the water table may rise into the crop root zone and cause water logging. In this case, the tile outlet should be opened to lower the water table. The spacing between the tile lines also influences the height of the water table. Narrow-spaced pipe laterals will ensure a more uniform water table distribution. In just the same way that soil type influences upward flux, so too does it influence the lateral pipe spacing. Therefore, the ET rate and the soil type must be taken into account when designing a subirrigation system.

Soils affected by salinization and water logging pose a severe management problem. Salts tend to move upward by capillary action and enter the root zone, thus restricting crop growth and leading to an accumulation of salt in the root zone. Under these conditions, upward groundwater movement is harmful and subsurface drainage is required to lower the water table.

In forested regions, and in tree plantations, there could be large amounts of ET by the tree canopy, thus lowering the water table. This phenomenon is used in some parts of the world affected by salinity and water logging. Salt-tolerant trees and shrubs are planted to control water logging and salinity, which is known as biodrainage.

Modern irrigation practices are being promoted to make better use of the soil as a reservoir to store precipitation and irrigation water. By encouraging field practices that enhance infiltration and improving soil tilth, there is less surface runoff, and more water is stored at shallow depths that can be extracted by the plants' roots to meet the ET demands. This process is also known as rainwater harvesting in rainfed agriculture. The overall effect is that water is conserved in the soil rather than in large open reservoirs, thus reducing irrigation infrastructure costs, operation and maintenance costs, and water distribution losses. Higher irrigation efficiencies are thus obtained.

BIBLIOGRAPHY

1. Gardner, W.R. (1958). Some steady state solutions of unsaturated moisture flow equations with application to evaporation from a water table. *Soil Sci.* **85**: 228–232.
2. Obreza, T.A. and Pitts, D.J. (2002). Effective rainfall in poorly drained microirrigated citrus orchards. *Soil Sci. Soc. Am. J.* **66**: 212–221.

CROP EVAPOTRANSPIRATION

Mladen Todorovic
Mediterranean Agronomic
Institute of Bari
CIHEAM-IAMB,
Valenzano, Bari, Italy

Crop evapotranspiration is a physical process in which water passes from the liquid to the gaseous state while moving from the soil to the atmosphere. It refers to both evaporation from soil and vegetative surfaces and transpiration from plants. These two processes are considered together because they occur simultaneously in nature, and it is difficult to separate them when measured or estimated.

Crop evapotranspiration refers to the water losses from a cropped field, and this term is frequently used in irrigation scheduling and irrigation water management.

Crop evapotranspiration is affected by weather factors (radiation, air temperature, humidity, and wind speed), crop parameters (crop type, variety, and development stage), and local environmental and management conditions (soil salinity, application of fertilizers, tillage practices, planting density, presence of diseases and pests, windbreaks, irrigation practices, presence of groundwater and impermeable soil horizons, etc.). Accordingly, it is possible to distinguish between crop evapotranspiration under standard (optimal) conditions and crop evapotranspiration under nonstandard conditions. Standard conditions refer to disease-free, well-fertilized crops, grown in large fields, under optimal water supply, and achieving full production for the given climatic conditions. Nonstandard conditions are defined as those that differ from standard conditions. *Crop evapotranspiration under standard conditions* is usually denoted as $\mathrm{ET_c}$, and *crop evapotranspiration under nonstandard conditions* is marked as $\mathrm{ET_{c\_adj}}$ or $\mathrm{ET_{c\_corr}}$, where the suffixes "*adj*" and "*corr*" refer to "adjusted" and "corrected," respectively. Standard conditions concur with the optimal conditions of well-managed irrigation fields, so, crop evapotranspiration under nonstandard conditions is always lower than or equal to crop evapotranspiration under standard conditions. Therefore,

$$\mathrm{ET_{c\_adj}} = K_s\mathrm{ET_c} \qquad (1)$$

where K_s is the reduction coefficient ($K_s < 1$) that accounts for the difference between standard and nonstandard conditions.

MEASUREMENTS OF CROP EVAPOTRANSPIRATION

Crop evapotranspiration can be determined by measurements and estimates. In general, measurements are carried out at agrometeorological stations and experimental sites to evaluate estimating procedures and to calibrate and derive parameters of estimation methods. Different types of lysimeters and methods based on a water and energy balance are used to measure crop evapotranspiration.

Lysimeters are tanks located in the fields and filled with soil of the same type as the surroundings. In lysimeters and their surroundings of at least 100 m around (to allow the establishment of an equilibrium sublayer), crops are grown under homogeneous management practices (standard or nonstandard) to measure the amount of water lost by evapotranspiration. The lysimeters provide direct measurements of crop evapotranspiration and can be grouped in three categories: weighing, where evapotranspiration is measured by weighing the changes in soil water content in the tank; nonweighing of a constant water table where the water table level is maintained at the same level inside as outside the lysimeter, and nonweighing of the percolation type, where water stored in the soil is determined by sampling or neutron methods and rainfall and percolation are measured. Weighing lysimeters give the most accurate estimates for short time periods (e.g., 1 hour), and they are recommended for arid and semiarid areas; nonweighing lysimeters of the constant water table type are particularly suggested for areas where a shallow groundwater table exists, nonweighing lysimeters of the percolation type are frequently used in areas of high precipitation.

The *water balance method* is a direct measurement technique based on the principles of conservation of mass which state that the difference in flow into and out of a controlled volume of soil during a specified time interval corresponds to the changes in soil moisture. This method requires monitoring many parameters such as irrigation water supply, precipitation, surface and subsurface lateral flow, and deep percolation. The accuracy of crop evapotranspiration estimates by this method is usually low due to difficulties in measuring and controlling one or more terms of the water balance equation.

Energy balance methods are based on the energy balance equation which states that the net flux density of radiation is balanced by losses of latent heat, sensible heat, and by soil heat flux. The Bowen ratio approach is the commonly used energy balance method. It relies on the flux gradient theory and expresses the latent heat flux of evapotranspiration through the measured values of net radiation, ground heat flux, the difference in air temperature between two heights above the canopy, and the corresponding difference in vapor pressure between the same two heights. This method can be used for crop evapotranspiration estimates in hourly and shorter time intervals.

ESTIMATES OF CROP EVAPOTRANSPIRATION

The estimate of crop evapotranspiration relies on the so-called two-step approach where in the first step, a reference evapotranspiration is determined and then, in the second, the crop evapotranspiration ($\mathrm{ET_c}$)is calculated as a product of reference evapotranspiration ($\mathrm{ET_o}$) and crop coefficient K_c;

$$\mathrm{ET_c} = K_c\mathrm{ET_o} \qquad (2)$$

In Eq. 2, the reference evapotranspiration term refers primarily to the climatic demand, and the crop coefficient accounts mainly for the specific crop characteristics and partially for the management practices (e.g., frequency of soil wetness, etc.).

The concepts of reference surface and reference evapotranspiration ($\mathrm{ET_o}$) estimates have been evaluated significantly during the second half of the twentieth century when they have passed from free water surface and empirical and semi empirical formulas to a hypothetical crop of fixed characteristics and more mechanistic equations.

Numerous equations for estimating reference crop evapotranspiration have been developed and successively modified to account for specific climatic conditions and availability of input information. Some of them in widespread use are presented in Table 1 along with the required input weather data and recommended timescale of applications.

In general, the methods for $\mathrm{ET_o}$ estimates do not have global validity and require local validation and calibration when applied in different environments. The accuracy of an $\mathrm{ET_o}$ estimate has a strong relation with the number and quality of required input information, local climatic conditions, timescale of applications, etc. For example, the Hargreaves method (1) could be the most appropriate for monthly $\mathrm{ET_o}$ estimates, provided that only temperature data are available. Nevertheless, it should not be used for daily $\mathrm{ET_o}$ estimates, especially when days are hot and dry and wind speed is relevant. The same can be stated for the radiation method which shows good results in humid climates and tends to underestimate $\mathrm{ET_o}$ in

Table 1. Some Methods for Reference Evapotranspiration Estimates, Weather Variables Needed, and Suggested Timescale of Application[a]

Method	Temperature	Humidity	Wind Speed	Sunshine or Radiation	Evaporation	Timescale			
						Hour	Day	Week	Month
Blaney–Criddle	+	–	–	*	–				X
Hargreaves	+	–	–	*	–				X
Pan evaporation	–	–	–	–	+			X	X
Radiation	+	–	–	+	–			X	X
Penman	+	+	+	+	–		X	X	X
Penman–Monteith	+	+	+	+	–	X	X	X	X

[a] + must be measured; – is not necessary; * estimate required; X recommended timescale of application

arid climates. The pan evaporation method integrates the effects of different weather variables on evaporation from a specific free water surface and requires calibration to account for the local environment. None of these methods should be used for daily $\mathrm{ET_o}$ estimates due to limited input information and approximations introduced in the equations. For $\mathrm{ET_o}$ estimates on daily and shorter timescales, more mechanistic methods are needed that rely on the full set of input weather variables and combine both aerodynamic and energy balance principles.

The first physically based combination equation was derived by Howard L. Penman in 1948 to describe the evaporation from a free water surface, and until now, this equation has remained the basis of numerous successive modifications and approaches. The main drawback of the Penman approach (2,3) is in the assumption that the evaporation process originates on the outer surface of the leaves rather than within their substomatal cavities. This lack of generality was satisfactorily overcome by John L. Monteith (4,5), who combined energy balance, aerodynamic, and surface parameters in an evapotranspiration equation formally valid for uniform vegetation of any type, fully covering the ground, and in any state of water status.

During the last decades of the twentieth century, many tests around the world have proved the robustness and consistency of the Penman–Monteith approach (6–9) and suggested it as the standard method for reference evapotranspiration estimates. Consequently, it was recommended to introduce a new concept of reference evapotranspiration surface necessary to standardize computational procedures for estimating reference evapotranspiration. These procedures should be used in different regions and climates of the world and thus, avoid ambiguities when the values of reference evapotranspiration have to be estimated and compared.

New reference surfaces and corresponding methodologies for evapotranspiration estimates represent the results of investigation by two groups of scientists:

- the FAO (Food and Agricultural Organization) Expert Group on the Revision of FAO Methodologies for Crop Water Requirements, which published the FAO Irrigation and Drainage paper 56 and also made it available on the Internet at www.fao.org/docrep/X0490E/x0490e00.htm, and
- the ASCE-EWRI (American Society of Civil Engineers—Environmental Water Resources Institute) Task Committee on Standardization of Reference Evapotranspiration, which released the first draft of Technical Report in July 2002.

The FAO Expert Group defined the reference surface as “a hypothetical reference crop with an assumed crop height of 0.12 m, a fixed average surface resistance of 70 s/m (over 24 hours), and an albedo (reflectance of evaporating surface) of 0.23”.

The ASCE-EWRI Task Committee distinguishes between two reference surfaces: the first is a short crop with an approximate height of 0.12 m, (similar to clipped grass and corresponding to most of the reference locations in the world), and the second is a tall crop with an approximate height of 0.50 m (similar to full-cover alfalfa and resembling many reference locations across the United States).

Then, provided that a standard set of measured input data is available (including air temperature and humidity, wind speed and incoming solar radiation), the reference evapotranspiration ($\mathrm{ET_o}$) can be calculated by the standardized form of the Penman–Monteith equation:

$$\mathrm{ET_o} = \frac{0.408\Delta(R_n - G) + \gamma \dfrac{C_n}{T + 273} U_2(e_s - e_a)}{\Delta + \gamma(1 + C_d U_2)} \tag{3}$$

where $\mathrm{ET_o}$ is the reference evapotranspiration (mm $\mathrm{day^{-1}}$ for a daily time step or mm $\mathrm{hour^{-1}}$ for an hourly time step), R_n is the net radiation, (MJ $\mathrm{m^{-2}}$ $\mathrm{day^{-1}}$ for a daily time step or MJ $\mathrm{m^{-2}}$ $\mathrm{hour^{-1}}$ for an hourly time step), G is the soil heat flux density (MJ $\mathrm{m^{-2}}$ $\mathrm{day^{-1}}$ for a daily time step or MJ $\mathrm{m^{-2}}$ $\mathrm{hour^{-1}}$ for an hourly time step), T is the mean daily or hourly air temperature at 1.5 to 2.5 m height (°C), Δ is the slope of the saturation vapor pressure vs. temperature curve (kPa $^{\circ}\mathrm{C}^{-1}$), γ is the psychrometric constant (=0.066 kPa $^{\circ}\mathrm{C}^{-1}$), e_s is the saturated vapor pressure at air temperature (kPa), e_a is the prevailing actual vapor pressure (kPa), U_2 is the mean wind speed measured at 2 m height (m $\mathrm{s^{-1}}$), and C_n and C_d are numerator and denominator constants, respectively, that change with the type of reference surface and calculation time step.

Equation 3 represents the simplest way to estimate the reference evapotranspiration from both reference surfaces (short and tall vegetation) and for different (hourly and daily) time steps. The values for C_n and C_d are derived by simplifying several terms within the Penman–Monteith

Table 2. The Values for C_n and C_d to be Used in Equation 3 (after Ref. 11)

Calculation Time Step	Short Reference Surface (Grass-Like)		Tall Reference Surface (Alfalfa-Like)	
	C_n	C_d	C_n	C_d
Daily	900	0.34	1600	0.38
Hourly during daytime	37	0.24	66	0.25
Hourly during nighttime	37	0.96	66	1.7

equation. They are presented for different reference surfaces and various time steps in Table 2.

The calculation procedures for the reference evapotranspiration and intermediate meteorological parameters ($R_n, G, e_s, e_a, \Delta, \gamma$, etc.) are reported in many books and publications: ASCE Manual 70 (6), FAO 56 (10), the ASCE-EWRI Task Committee report (11), etc. Moreover, free software (REF-ET) for standardized calculations of reference evapotranspiration and intermediate parameters is available at www.kimberly.uidaho.edu/ref-et. An example of a reference evapotranspiration estimate using the standardized form of the Penman–Monteith equation is given here for both grass-like and alfalfa-like surfaces.

Example of a Calculation daily (for a location in southern Italy on June 10, 2002)

Given:

Net radiation $R_n = 14$ MJ m^{-2} day^{-1}; soil heat flux $G = 0$ (can be neglected especially for daily ET$_o$ estimates); minimum air temperature $T_{\min} = 14\,^\circ$C; maximum air temperature $T_{\max} = 28\,^\circ$C; wind speed at 2 m height $U_2 = 1.2$ m s^{-1}; minimum relative humidity (at $T_{\max}$) RH$_{\min} = 50\%$; maximum relative humidity (at $T_{\min}$) RH$_{\max} = 80\%$.

Calculation. The slope of the saturation vapor pressure curve Δ is calculated for mean air temperature $T = 21\,^\circ$C as

$$\Delta = \frac{4098\left[0.6108\exp\left(\dfrac{17.27T}{T+237.3}\right)\right]}{(T+237.3)^2}$$

$$= \frac{4098\left[0.6108\exp\left(\dfrac{17.27*21}{21+237.3}\right)\right]}{(21+237.3)^2}$$

$$= 0.1528\ \text{kPa}/^\circ\text{C}$$

The saturation vapor pressure is calculated as the average value of the saturated vapor pressures at $T_{\min}$ and $T_{\max}$:

$$e_{s\_T\min} = 0.6108\exp\left[\frac{17.28T_{\min}}{T_{\min}+237.3}\right]$$

$$= 0.6108\exp\left[\frac{17.28*14}{14+237.3}\right] = 1.60\ \text{kPa}$$

$$e_{s\_T\max} = 0.6108\exp\left[\frac{17.28T_{\max}}{T_{\max}+237.3}\right]$$

$$= 0.6108\exp\left[\frac{17.28*28}{28+237.3}\right] = 3.78\ \text{kPa}$$

$$e_s = \frac{e_{s\_T\min} + e_{s\_T\max}}{2}$$

$$= \frac{1.60+3.78}{2} = 2.69\ \text{kPa}$$

The actual vapor pressure is calculated as the average value of the vapor pressures at $T_{\min}$ and $T_{\max}$:

$$e_{a\_T\min} = \frac{\text{RH}_{\max}}{100}e_{s\_T\min} = \frac{80}{100}1.60 = 1.28\ \text{kPa}$$

$$e_{a\_T\max} = \frac{\text{RH}_{\min}}{100}e_{s\_T\max} = \frac{50}{100}3.78 = 1.89\ \text{kPa}$$

$$e_a = \frac{e_{a\_T\min} + e_{a\_T\max}}{2} = \frac{1.28+1.89}{2} = 1.58\ \text{kPa}$$

Then, for a short (grass-like) reference surface and daily ET$_o$ estimate, the constants C_n and C_d can be obtained from Table 2 ($C_n = 900$ and $C_d = 0.34$), and the reference evapotranspiration can be estimated from Equation 3 as

$$\text{ET}_o = \frac{0.408\Delta(R_n - G) + \gamma\dfrac{C_n}{T+273}U_2(e_S - e_a)}{\Delta + \gamma(1 + C_dU_2)}$$

$$= \frac{\begin{array}{c}0.408*0.1528*(14-0)\\ +0.066*\dfrac{900}{21+273}*1.2*(2.69-1.58)\end{array}}{0.1528+0.066*(1+0.34*1.2)}$$

$$= 4.65\ \text{mm/day}$$

Similarly, assuming that the input data remain the same, for a tall (alfalfa-like) reference surface, the constants C_n and C_d can be obtained from Table 2 ($C_n = 1600$ and $C_d = 0.38$), and ET$_o$ can be estimated as

$$\text{ET}_o = \frac{0.408\Delta(R_n - G) + \gamma\dfrac{C_n}{T+273}U_2(e_S - e_a)}{\Delta + \gamma(1 + C_dU_2)}$$

$$= \frac{\begin{array}{c}0.408*0.1528*(14-0)\\ +0.066*\dfrac{1600}{21+273}*1.2*(2.69-1.58)\end{array}}{0.1528+0.066*(1+0.38*1.2)}$$

$$= 5.43\ \text{mm/day}$$

The example shows that the reference evapotranspiration is greater from a tall alfalfa-like surface than from a short grass-like surface. This is due to the fact that both the aerodynamic and surface resistance are lower for the tall alfalfa-like surface than for the short grass-like surface.

The estimate of the reference evapotranspiration on an hourly basis can be done by using the same equation and appropriate values for C_n and C_d from Table 2. Moreover, the net radiation R_n should be expressed in MJ m^{-1} hour^{-1}, and the soil heat flux G should be taken into consideration as a portion of the energy term.

The Hargreaves equation represents an alternative solution for the estimate of reference evapotranspiration when input data are limited only to the values of minimum and maximum air temperature (the data widely available and published in hydrometeorological bulletins worldwide). This simple empirical equation has shown good performance when applied on a monthly scale and also on shorter timescales (e.g., weekly), when the wind speed is not strong and the climate is humid rather than dry.

Besides the minimum and maximum air temperatures ($T_{\min}$ and $T_{\max}$, respectively), the Hargreaves method (1) employs in the estimate the extraterrestrial radiation (R_a) which is a function of the latitude of the location and the day of the year (Julian day) under consideration that can be easily obtained from agrometeorological tables or calculated by using the following set of formulas:

$$R_a = 37.586\, d_r(\omega_s \sin\varphi \sin\delta + \cos\varphi \cos\delta \sin\omega_s) \quad (4)$$

where

d_r: relative distance Earth–Sun (rad)

$$d_r = 1 + 0.033\cos\left(\frac{2\pi}{365}J\right) \quad (5)$$

J: day of the year (Julian day) number (nondimensional),

$$J = \text{integer}\ (30.42M - 15.23) \quad (6)$$

M: month number (1–12) (nondimensional),

ω_s: sunset hour angle (rad),

$$\omega_s = \arccos(-\tan\varphi\tan\delta) \quad (7)$$

φ: latitude of location (rad),

δ: solar declination (rad),

$$\delta = 0.4093\sin\left(\frac{2\pi}{365}J - 1.405\right) \quad (8)$$

Then, the widely accepted form of the Hargreaves equation for an $\mathrm{ET_o}$ estimate (mm day^{-1}) for a reference grass surface is

$$\mathrm{ET} = 0.0023\frac{R_a}{\lambda}(T + 17.8)(T_{\max} - T_{\min})^{0.5} \quad (9)$$

where R_a is the extraterrestrial radiation (MJ m^{-2} day^{-1}), λ is the latent heat of vaporization (MJ kg^{-1})

$$\lambda = 2.501 - (2.361 * 10^{-3})T \quad (10)$$

T is the average temperature (°C)

$$T = \frac{T_{\min} + T_{\max}}{2} \quad (11)$$

and multiplier (0.0023), exponent (0.5), and the value of 17.8 in the second multiplier of the equation are the parameters which should be validated and eventually calibrated for different locations.

Example of a Calculation monthly (for a location in southern Italy in June 2002)

Given:

Minimum and maximum air temperature $T_{\min} = 14\,^\circ\mathrm{C}$ and $T_{\max} = 28\,^\circ\mathrm{C}$; latitude of location is $\varphi = 40^\circ = 40 * 3.14/180 = 0.697778$ rad; the month is June which means $M = 6$.

Calculation. Using the following sequence of Equations (6–5–8–7–4), the extraterrestrial radiation (MJ m^{-2} day^{-1}) is calculated as

$$\begin{aligned} J &= \text{integer}\ (30.42M - 15.23) \\ &= \text{integer}\ (30.42 * 6 - 15.23) = 167 \\ d_r &= 1 + 0.033\cos\left(\frac{2\pi}{365}J\right) \\ &= 1 + 0.033\cos\left(\frac{2 * 3.14}{365}167\right) = 0.96818 \\ \delta &= 0.4093\sin\left(\frac{2\pi}{365}J - 1.405\right) \\ &= 0.4093\sin\left(\frac{2 * 3.14}{365}167 - 1.405\right) = 0.40715256 \\ \omega_s &= \arccos(-\tan\varphi\tan\delta) \\ &= \arccos[(-\tan 0.697778)(\tan 0.40715256)] \\ &= 0.935331 \\ R_a &= 37.586 d_r(\omega_s \sin\varphi\sin\delta + \cos\varphi\cos\delta\sin\omega_s) \\ R_a &= 37.586 * 0.96818[0.935331(\sin 0.697778) \\ &\quad \times (\sin 0.40715256) + (\cos 0.697778)(\cos 0.40715256) \\ &\quad \times (\sin 0.935331)] = 29.267 \end{aligned}$$

Then, the average temperature and the latent heat of vaporization are calculated as

$$\begin{aligned} T &= \frac{T_{\min} + T_{\max}}{2} = \frac{14 + 28}{2} = 21\,^\circ\mathrm{C} \\ \lambda &= 2.501 - (2.361 * 10^{-3})T = 2.501 - 0.049581 \\ &= 2.45142\ \mathrm{MJkg}^{-1} \end{aligned}$$

Finally, the grass reference crop evapotranspiration is

$$\begin{aligned} \mathrm{ET} &= 0.0023\frac{R_a}{\lambda}(T + 17.8)(T_{\max} - T_{\min})^{0.5} \\ \mathrm{ET} &= 0.0023\frac{29.267}{2.45142}(21 + 17.8)(28 - 14)^{0.5} \\ &= 3.99\ \text{mm/day} \end{aligned}$$

This means that the grass reference evapotranspiration on a monthly basis for June 2002 was $3.99 * 30 =$ **119.7 mm month^{-1}**.

The average daily value of the $\mathrm{ET_o}$ estimate using the Hargreaves method (3.99 mm day^{-1}) is significantly lower than the value (4.65 mm day^{-1}) obtained from the estimate using the Penman–Monteith equation. This is due to the fact that the impact of wind speed and humidity are not taken into consideration in the Hargreaves equation. However, this equation can be used on a monthly basis to estimate an average reference evapotranspiration. Moreover, the estimates of $\mathrm{ET_o}$ using the Hargreaves equation can be improved by adapting the parameters to the local conditions.

The second step in the crop evapotranspiration estimate is the determination of crop coefficient K_c and the calculation of $\mathrm{ET_c}$ by Equation (2). The crop coefficient is a nondimensional factor that represents the ratio between the crop and reference evapotranspiration. The crop coefficient integrates the features of a typical field crop that change during its growing season and differ from the characteristics of a reference surface, which has a constant appearance and is a complete ground cover. These features comprise the crop height and crop aerodynamic properties (roughness), albedo (reflectance), canopy resistance, and soil evaporation from a cropped field.

K_c values vary from crop to crop during the growing season and depend also on weather conditions and irrigation practices (Fig. 1). It is possible to distinguish between the single crop coefficient K_c and the dual crop coefficient where the K_c is split into the basal crop coefficient (K_{cb}) accounting for crop transpiration and the evaporation coefficient (K_e) referring to soil evaporation. The use of a single K_c approach is recommended in regular applications for management purposes and the development of irrigation scheduling; the dual K_c approach should be applied only for specific purposes (e.g., research) and on a daily basis.

K_c values for different crops are given in many technical materials (FAO 24, FAO 33, FAO 56, etc.) referring to short grass as the reference surface. Nevertheless, it is important to emphasize that most K_c values reported in the literature are obtained with the modified Penman equation (FAO 24) and they should be revised and updated when used with the FAO Penman–Monteith method or some other equation for a reference evapotranspiration estimate. In fact, when crop evapotranspiration has to be calculated, the K_c values should be applied in conjunction with the equations with which they were obtained.

The planting months and the lengths of crop development stages for some widely distributed species are given in Table 3, and an indicative list of K_c values for different crop growth stages is presented in Table 4. The K_c values are given, assuming that the reference evapotranspiration is calculated for a grass-like surface. Both tables are prepared on the basis of the FAO data and some other investigations. However, these data should be used only in the absence of more reliable values.

Example of an estimate of evapotranspiration for a tomato crop growing in the Mediterranean region with the planting date of April 15th.

Calculation Procedure. The lengths of the growing stages are taken from Table 3 and K_c values from Table 4. Then, the growing cycle of tomato can be presented graphically as in Fig. 2.

The K_c values can be determined daily by using a simple linear interpolation. For example, on June 10th, the K_c value is 0.92. Then, assuming that the reference evapotranspiration is 4.65 mm day^{-1} (as calculated in the example here before for the grass-like reference surface using the Penman–Monteith equation), the tomato evapotranspiration for June 10th is estimated by using Eq. 2 as

$$\mathrm{ET_c} = K_c \mathrm{ET_o} = 0.92 * 4.65 = 4.28 \text{ mm/day}$$

Finally, if the tomato crop is not growing under standard (optimal) conditions, the crop evapotranspiration can be adjusted to the real conditions by applying Eq. 1. The values of reduction coefficient K_s are always lower than one and depend on the specific site conditions to which the crop is exposed during its growing cycle. In any case, in real management, checking the results of $\mathrm{ET_c}$

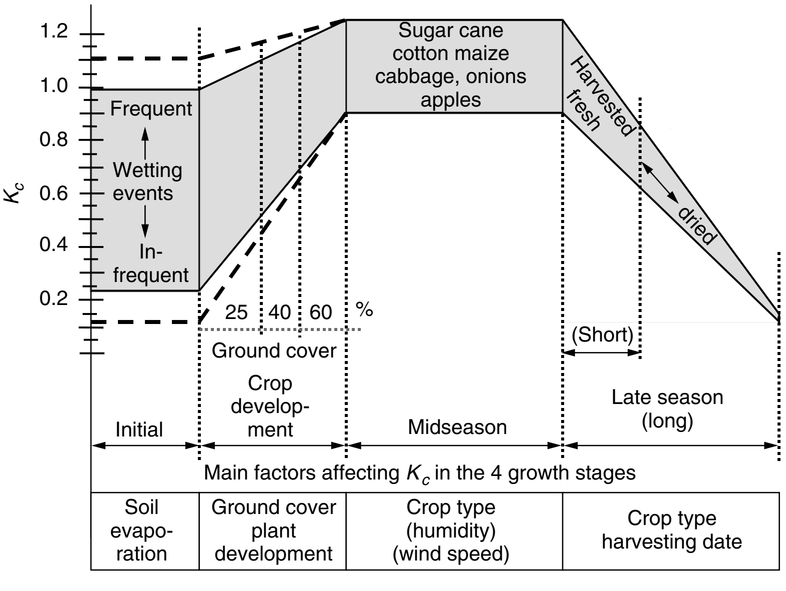

Figure 1. Main factors affecting the crop coefficient (K_c) during the four growing stages (adapted from FAO 56).

Table 3. Planting Months and Lengths of Crop Development Stages for Some Important Crops in Different Regions of the World (Adapted from FAO 56)

Crop	Planting Month	Length of Crop Development Stages					Region
		Initial	Develop.	Mid	Late	Total	
Tomato	January	30	40	40	25	135	Arid regions
	Apr/May	35	40	50	30	155	Calif., USA
	January	25	40	60	30	155	Calif. desert, USA
	Oct/Nov	35	45	70	30	180	Arid regions
	April/May	30	40	45	30	145	Mediterranean
Potato	Jan/Nov	25	30	30/45	30	115/130	Semiarid climate
	May	25	30	45	30	130	Continental climate
	April	30	35	50	30	145	Europe
	Apr/May	45	30	70	20	165	Idaho, USA
	December	30	35	50	25	140	Calif. desert, USA
Sugar beet	March	30	45	90	15	180	Calif., USA
	June	25	30	90	10	155	Calif., USA
	September	25	65	100	65	255	Calif. desert, USA
	April	50	40	50	40	180	Idaho, USA
	May	25	35	50	50	160	Mediterranean
	November	45	75	80	30	230	Mediterranean
	November	35	60	70	40	205	Arid regions
Soybeans	Dec	15	15	40	15	85	Tropics
	May	20	30/35	60	25	140	Central USA
	June	20	25	75	30	150	Japan
Cotton	March–May	30	50	60	55	195	Egypt, Pakistan, Calif.
	March	45	90	45	45	225	Calif. desert, USA
	September	30	50	60	55	195	Yemen
	April	30	50	55	45	180	Texas
Sunflower	April/May	25	35	45	25	130	Medit., Calif.
Barley/Oats/Wheat	November	15	25	50	30	120	Central India
	March/April	20	25	60	30	135	35–45° Lat.
	July	15	30	65	40	150	East Africa
	April	40	30	40	20	130	
	November	40	60	60	40	200	
	December	20	50	60	30	160	Calif. desert
Winter Wheat	December	20	60	70	30	180	Calif., USA
	November	30	140	40	30	240	Mediterranean
	October	160	75	75	25	335	Idaho, USA
Maize (grain)	April	30	50	60	40	180	East Africa (alt.)
	Dec/Jan	25	40	45	30	140	Arid climate
	June	20	35	40	30	125	Nigeria (humid)
	October	20	35	40	30	125	India (dry, cool)
	April	30	40	50	30	150	Spain, Calif.
	April	30	40	50	50	170	Idaho, USA
Sorghum	May/June	20	35	40	30	130	USA, Pakis., Medit.
	March/April	20	35	45	30	140	Arid regions
Rice	Dec; May	30	30	60	30	150	Tropics; Medit.
	May	30	30	80	40	180	Tropics
Grapes	April	20	40	120	60	240	Low latitudes
	March	20	50	75	60	205	Calif., USA
	May	20	50	90	20	180	High latitudes
	April	30	60	40	80	210	Midlatitudes (wine)
Citrus	January	60	90	120	95	365	Mediterranean
Apples/Pears	March	20	70	90	30	210	High latitudes
	March	20	70	120	60	270	Low latitudes
	March	30	50	130	30	240	Calif., USA
Olives	March	30	90	60	90	270	Mediterranean

Table 4. Indicative Crop Coefficient Values for Different Growing Stages of Some Important Crops under Optimal Management Conditions and Reference ET Calculated for a Grass-like Surface (Adapted mainly from the FAO 56)

Crop	K_{c_initial}[a]	K_{c_mid}	K_{c_end}	Maximum Crop Height (m)
Tomato	0.4–0.6	1.15–1.20	0.70–0.90	0.60
Potato	0.4–0.6	1.15	0.60–0.70	0.40
Sugar beet	0.4–0.5	1.20	0.70	0.50
Soybean	0.3–0.5	1.15	0.50	0.5–1.0
Cotton	0.4–0.5	1.15–1.20	0.60–0.70	1.2–1.5
Sunflower	0.3–0.5	1.0–1.15	0.40	2.0
Barley	0.3–0.5	1.15	0.25	1.0
Oats	0.3–0.5	1.15	0.25	1.0
Spring Wheat	0.3–0.5	1.15	0.25–0.40[b]	1.0
Winter Wheat				
• with frozen soils	0.4	1.15	0.25–0.40[b]	1
• with nonfrozen soils	0.7	1.15	0.25–0.40[b]	1
Maize, Field	0.3–0.5	1.20	0.35–0.60[c]	2.0
Sorghum				
• grain	0.3–0.5	1.0–1.10	0.55	1–2
• sweet	0.3–0.5	1.20	1.05	2–4
Rice	1.05	1.20	0.90	1.0
Grapes				
• table or raisin	0.30	0.85	0.45	2.0
• wine	0.30	0.70	0.45	1.5–2.0
Citrus, no ground cover				
• 70% canopy	0.70	0.65	0.70	4
• 50% canopy	0.65	0.60	0.65	3
• 20% canopy	0.50	0.45	0.55	2
Citrus, active ground cover or weeds				
• 70% canopy	0.75	0.70	0.75	4
• 50% canopy	0.80	0.80	0.80	3
• 20% canopy	0.85	0.85	0.85	2
Apples/Pears				
• no ground cover, killing frost	0.45	0.95	0.70	4
• no ground cover, no frost	0.60	0.95	0.75	4
• active ground cover, killing frost	0.50	1.20	0.95	4
• active ground cover, no frost	0.80	1.20	0.85	4
Olives (40 to 60% ground cover)	0.65	0.70	0.70	3–5

[a] The higher values should be adopted when the surface is frequently wetted.

[b] The higher value is for the hand-harvested crop.

[c] The higher value is for harvesting at high grain moisture.

estimates obtained from equations with the results of on-field monitoring of soil and/or plant water status is highly recommended.

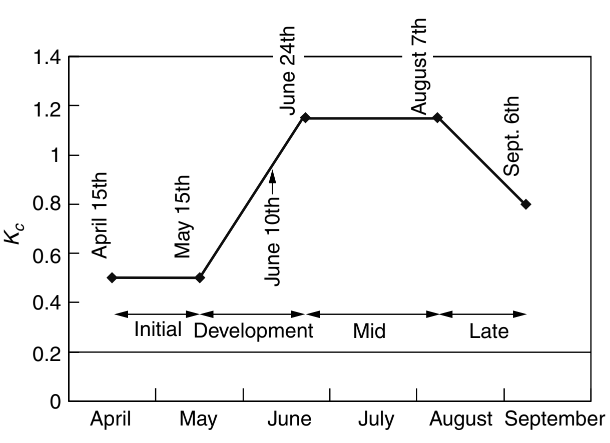

Figure 2. The evaluation of K_c values of a tomato crop during the growing season.

BIBLIOGRAPHY

1. Hargreaves, G.H. and Samani, Z.A. (1985). Reference crop evapotranspiration from temperature. *Appl. Eng. Agric.* **1**: 96–99.
2. Penman, H.L. (1948). Natural evaporation from open water, bare soil, and grass. *Proc. R. Soc. London, Ser. A* **193**: 120–146.
3. Penman, H.L. (1953). *The Physical Basis of Irrigation Control*. 13th Int. Hort. Congr., Vol. 2, pp. 913–922.
4. Monteith, J.L. (1965). Evaporation and atmosphere. In: *The State and Movement of Water in Living Organisms. Symp. Soc. Exp. Biol.*, XIX, pp. 205–234.
5. Monteith, J.L. (1973). *Principles of Environmental Physics*. Edvard Arnold, London.
6. Jensen, M.E., Burman, R.D., and Allen, R.G. (1990). *Evapotranspiration and Irrigation Water Requirements*. ASCE Manuals and Reports on Engineering Practices, N° 70. American Society Civil Engineers, New York.
7. Smith, M., Allen, R.G., Monteith, J.L., Perrier, A., Pereira, L.S., and Segeren, A. (1991). *Report of the Expert Consultation on Procedures for Revision of FAO Guidelines for Prediction of Crop Water Requirements*. UN-FAO, Rome, Italy.

8. Choisnel, E., de Villele, O., and Lacroze, F. (1992). *Une Approche Uniformise'e du Calcul de L'évapotranspiration Potentielle pour L'ensemble des Pays de la Communauté Européenee, Com. Commun, Européennes*. EUR 14223 FR, Luxembourg.
9. Allen, R.G., Smith, M., Perrier, A., and Pereira, L.S. (1994). An update for the definition of reference evapotranspiration. *ICID Bull.* **43**(2): 1–34.
10. Allen, R.G., Pereira, L.S., Raes, D., and Smith, M. (1998). *Crop Evapotranspiration. Guidelines for Computing Crop Water Requirements*. FAO Technical Paper 56, UN-FAO Rome, Italy.
11. ASCE Task Committee for Standardization of Reference Evapotranspiration. (2002). *The ASCE Standardized Reference Evapotranspiration Equation*. Environmental and Water Resources Institute of American Society of Civil Engineers, ASCE, Reston, VA.

WATER POLLUTION FROM FISH FARMS

Asbjørn Bergheim
RF-Rogaland Research
Stavanger, Norway

Alexander Brinker
Fischereiforschungsstelle des
Landes Baden-Württemberg
Langenargen, Germany

WASTE OUTPUT

Wastes from aquaculture plants include all materials used in the process that are not removed from the system during harvest. The principle wastes from aquaculture are uneaten feed, excreta, chemicals, and therapeutics. In addition, the term "waste" can also refer to pathogens and dead or escaped fish.

Generally, the quantity of waste is closely connected to the culture system used. *Intensive* farm systems, typically monoculture of carnivorous finfish in the temperate zone reliant on artificial feed, may cause serious local pollution. So-called semi-intensive farm systems are supplied natural feed sources, such as vegetation, oil cakes, cereal bran, and organic-chemical fertilizers. The latter systems dominate the tropical/subtropical production of herbivorous or omnivorous fish, e.g., the major production of carps and tilapia, and the waste output to the surrounding waters is much lower than from intensive fish farms. This is illustrated in Fig. 1, where the nutrient budget is estimated for intensive farms (cage or flow-through raceway systems) and for semi-intensive farms (Nile tilapia produced in static pond water). In highly intensive farms, as much as 74–84% of the nutrient inputs are released into surrounding waters, whereas only 2–6% of the nutrients are released from semi-intensive farms. Semi-intensive production of carps takes place mostly in static ponds only drained during harvest, where supplied nutrients not assimilated in fish are mainly trapped in the pond sediments. The Chinese production of carps and tilapia in freshwater constitutes more than one third of the global aquaculture production (China: 12 million MT in 1997, 1). Thus, the retained fraction of nutrients in the pond sediments is vital to avoid serious water pollution.

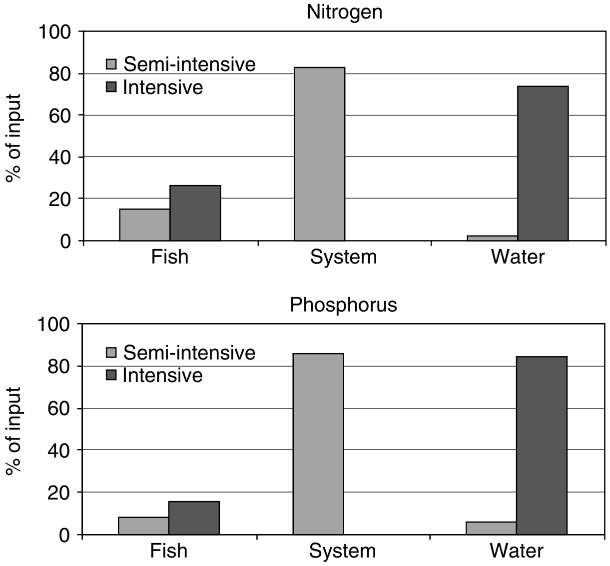

Figure 1. A comparison of the nutrient budget of a semi-intensive and an intensive aquaculture system (see text). After Edwards (2).

In distinct regions, fish farms represent a dominating water pollution source. Along the western and northern coast of Norway, cage farming of salmon and trout contributed about 55% of P and 17% of N of the total input in 1998 (see Fig. 2). Since then, the annual production has increased from c. 400,000 MT to c. 600,000 MT. The total nutrient input to this part of the coastline has strongly increased over the last 20 years because of the growing aquaculture industry. The average specific loading was 10.8 kg P and 52.0 kg N per ton of produced fish. For population equivalents (p.e.):6000 p.e. as Total P and 4000 p.e. as Total N per MT produced fish. Besides, uneaten feed and feces from cage farms load the recipient with biodegradable organic matter that may lead to significant oxygen depletion [corresponds to 300–500 kg as BOD_7 per MT produced fish, (4)].

A wide range of chemicals are used in the aquaculture industry, including disinfectants, chemotherapeutants (antibacterial, antifungal, and antiparasitic compounds), pigments incorporated into feeds, and compounds employed in and applied to construction materials (e.g., antifoulants) (5). Of particular interest is the use of chemotherapeutants because of the quantities discharged into the aquatic environment. In Norway, the only country that has kept records, around 50 MT of antibacterials were introduced into the coastal environment by the aquaculture industry in 1987. The use has, however, fallen to below 1 MT annually because of development of efficient vaccines against bacterial diseases. In semi-intensive, tropical freshwater farms, antibacterials are sparsely applied.

ENVIRONMENTAL IMPACTS

Most reported studies of environmental effects are related to intensive aquaculture. The quality of wastes from fish farms and its temporal variability are important

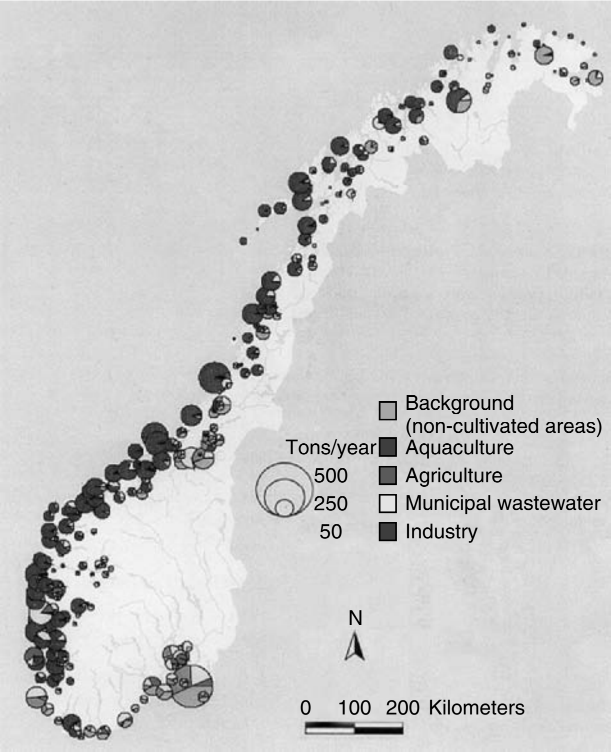

Figure 2. Estimated phosphorus inputs to the Norwegian coastal waters from the main sources in 1998 (3).

determinants of environmental impacts (6). Aquaculture wastes are rich in P with a TP/TN ratio of around 5. Thus, fish farms have been expected to be a potential source for eutrophication in freshwater sources as P is the limiting nutrient. Generally, the eutrophication rates in lakes and reservoirs influenced by fish farming seem to be lower than those predicted by lake models. Håkanson et al. (7) observed very little, if any, increased concentrations of TP in two Swedish lakes influenced by cage farm emissions. The main causes seemed to be linked to (1) direct uptake of feces and waste feed from farmed fish by wild fish, and (2) a significant amount of P from the cages that are eliminated from the lake water by different processes, e.g., sedimentation.

Nutrients emitted from cage farms are found to be highly biologically available for marine algae (8): The bioavailability of P in effluent from aquaculture, agriculture, and municipal sewage was 50/60%, 30%, and 65/70%, respectively, whereas more than 80% of N in all effluent types was available for algal growth. Despite that aquaculture is the dominating nutrient source along the western-northern coast of Norway, the cage farms are located in a zone with huge water exchange (The Gulf Stream) and large recipient capacity.

Deposition of solids from cage farms may strongly effect the benthic ecosystems beneath the cages (6). Such materials have shown to increase the extent of anaerobic sediment layers with dramatic consequences for the original biological communities. Under intensive cage farms, the bacterial activity in anaerobic sediments often causes out-gassing of methane and hydrogen sulphide. In addition, sedimentary phosphorus might be released at oxygen depletion in deeper water. Studies of benthic communities all indicate that detrimental impacts are limited to the immediate vicinity of the intensive cage (or mussel) operation.

Among other consequences of modern aquaculture are environmental effects caused by antibiotic use, e.g., drug residues in water, sediments and organisms, and introduction of exotic species for aquaculture purposes. Welcomme (9) reported that 98 species of fish had been introduced internationally involving inland waters. Also, accidentally or deliberately released farmed fish can affect their local wild forms by behavior and genetic impacts.

ENVIRONMENTAL REGULATIONS

A wide range of regulations and standards control fish farming and the discharge of effluent in parts of the world, such as North America and Europe. Aquaculture relies on precious water resources, and; as populations increase, access to good quality water and the ability to discharge effluents will become more restrictive. In many countries, aquaculture legislation is poorly developed. During the last few years, however, interest has grown to develop a comprehensive regulatory framework for aquaculture whose goals should be to protect the industry,

Table 1. Governmental Regulations for Fish Farm Effluents in Some European Countries per 1999 (11)

	Country									
Criteria	B	DK	F	GR	IRL	NL	N	S	SF	UK
Limitation of production		§			§		§	§	§	
Effluent treatment		§	§	§	§	§			§	§
Limitation of N and P load	§	§	§	§	§		§		§	§
Limitation of organic load	§	§	§	§	§	§	§			§
Feed composition		§	§					§	§	
Feed conversion ratio		§	§				§		§	
Environmental impact assessment		§	§		§		§	§		§
Rules for usage of chemicals					§		§			§
Monitoring of effluent water		§			§		§			§
Taxes on effluent water and/or feeds	§		§			§				§

B: Belgium, DK: Denmark, F: France, GR: Greece, IRL: Ireland, NL: Holland, N: Norway, S: Sweden, SF: Finland, UK: United Kingdom.

the environment, other resource users, and the consumer, and to clarify rules for resource use (10).

A brief review of regulations controlling discharge of effluent from fish farms in ten European countries is presented in Table 1.

The waste load from fish farms can also be reduced by effluent treatment attempts (see WASTE WATER TREATMENT IN FISH FARMS).

BIBLIOGRAPHY

1. FAO. (2000). FISHSTATE Plus—Version 2.3. Available: http://www.fao.org/fi/statist/fisoft/fishplus.asp.
2. Edwards, P. (1993) In: *Environment and Aquaculture in Developing Countries, ICLARM Conf. Proc.* R.S.V. Pullin, H. Rosenthal, and J.L. Maclean (Eds.). Vol. 31, pp. 139–170.
3. NIVA. (2000). Report from Norwegian Institute for Water Research, O-98140, Oslo, Norway.
4. Aasgaard, T. and Bergheim, A. (1996). In: *Aquaculture and Water Resource Management*. D.J. Baird, M.C.M Beveridge, L.A. Kelly, and J.F. Muir (Eds.). Blackwell Science, London, UK, pp. 50–80.
5. Beveridge, M.C.M. (1996). *Cage Aquaculture*, 2nd Edn. Fishing News Books, Blackwell Science Ltd, London, UK.
6. Costa-Pierce. (1996).
7. Håkanson, L., Carlsson, L., and Johansson, T. (1998). A new approach to calculate the phosphorus load to lakes from fish farm emissions. *Aquacultural Eng.* **17**: 149–166.
8. NIVA. (1992). Report from Norwegian Institute for Water Research, O-89156, Oslo, Norway.
9. Welcomme, R.L. (1988). *International Introductions of Inland Aquatic Species*. FAO Fish. Tech. Paper, 294. FAO, Rome, Italy.
10. Kongkeo, H. (2001). In: *Aquaculture in the Third Millennium*. R.P. Subasinghe, P.B. Bueno, M.J. Phillips, C. Hough, S.E. McGladdery, and J.R. Arthur (Eds.). NACA/FAO, Bangkok, Thailand, pp. 267–293.
11. Bergheim, A. and Brinker, A. (2003). Effluent treatment for flow through systems and European environmental regulations. *Aquacultural Eng.* **27**: 61–77.

WORLD'S MAJOR IRRIGATION AREAS

JACOB W. KIJNE
Herts, United Kingdom

EXTENT OF IRRIGATION

Irrigation is the artificial watering of land by technical means, such as canals, pumps, sprinklers, or drip systems, to supply moisture to crops. Surface irrigation, such as basin and furrow irrigation, in which water is led to the fields through canals and ditches, has been practiced for over two millennia.

About 20% of the world's agricultural land is irrigated, but some 40% of the food and fiber production worldwide takes place on irrigated land. For cereal production, it is close to 60% (1). FAO expects this share to increase further in the next three decades. Much of this expansion will take place in developing countries, increasing their irrigated land perhaps by as much as 20% before 2030. This is a net expansion, which implies that agricultural land lost to urban growth and soil salinization will be compensated.

At present, more than 60% of the world's irrigated land is in Asia. Table 1 shows the ten top countries and river basins in terms of irrigated land area. About two thirds of the irrigated land in Asia is devoted to rice and wheat production.

Worldwide, about 70% of water withdrawals is supplied to agriculture. Table 2 shows for the ten top irrigation countries the total water withdrawal (for agriculture, industrial, and domestic use) as a percentage of the renewable water resource in the country and the percentage of this total withdrawal used in agriculture.

Although food production to 2020 is expected to increase faster in developing countries than in developed countries, experts suggest that food production will not keep pace with demand. In 1999/2000, developing countries produced 55% of the world grain production and accounted for 61% of world grain consumption. They also imported 231 million tons of grain, equivalent to 72% of worldwide imports, which suggests that developing countries although playing a major role in irrigated agriculture production still depend on international agricultural trade for their food security.

DIFFERENT METHODS OF IRRIGATION

The three main types of irrigation are canal (or surface) irrigation, sprinkler irrigation, and trickle or drip irrigation. In canal irrigation, water is distributed through a network of canals, consisting of one or more main canals conveying water from the source (e.g., a reservoir or a barrage in a river), distributary (secondary) canals from the outlets in the main canal, tertiary canals (called watercourses in the Indian subcontinent) from the outlets in the distributary canals to the farm gate, and field canals. Canal irrigation systems have lower construction and operating costs, simpler maintenance, and lower skilled labor demands than do pressurized delivery systems. A disadvantage of canal irrigation over piped systems is that the soil must convey and infiltrate the water over the fields. Because soil properties are highly variable, both spatially and temporally, designing water-efficient surface irrigation at the farm level is complicated. Much research has been done on determining the optimum flow rate for a given length of furrow and vice versa. It is now claimed that in furrow irrigation design, it is possible to attain an application efficiency (i.e., the ratio of the average depth of water applied to the root zone, and the average depth of water applied to the field) of about 70% (2). However, the reality in the field is probably different. Application efficiency refers to the average depths of application, but the uniformity of distribution is also important. It expresses the often large differences in the depths of water stored in the root zone along the length of the fields resulting from local differences in soil structure and texture and variable infiltration amounts in low and high spots in the field. Precision land leveling and automatic devices (e.g., for the release of surge flows) have

Table 1. Top Ten Countries and River Basins for Irrigated Land

Country	Area (million ha)	% of World Total	River Basin	Area (million ha)
India	59	21	Yangtze	31
China	54	20	Mekong	23
United States	22	8	Xi jiang	19
Pakistan	18	6	Ganges	11
Iran	8	3	Irriwaddy	10
Mexico	6.5	2	Brahmaputra	9
Indonesia	4.8	2	Yellow river	9
Thailand	4.8	2	Indus	8
Turkey	4.5	1.5	Chao Praya	7.5
Bangladesh	4	1.5	Krishna	6

Source: FAOSTAT On-line Statistical Service (http://apps.fao.org) and IWMI, 2002. World Irrigation and Water Statistics 2002, International Water Management Institute, Colombo, Sri Lanka.

Table 2. Water Withdrawal as Percentage of Renewable Water Resource, and its Share Supplied to Irrigation

Country	Percent of Water Resource Withdrawn (2000)	Percent of Total Withdrawal Supplied to Agriculture (1995)
India	26	96
China	19	89
USA	17	41
Pakistan	70	96
Iran	54	95
Mexico	17	85
Indonesia	3	94
Thailand	8	85
Turkey	15	83
Bangladesh	1	96

Source: same as Table 1.

improved attainable application efficiency and uniformity of distribution, while reducing the labor requirements of surface irrigation (3).

Sprinkler irrigation and drip (trickle) irrigation eliminate the need for land leveling and allow more frequent water application without alternating saturation and desiccation of the upper soil layers. But most importantly, these newer methods have improved attainable application efficiencies and, hence, lose less water. Nowadays, commercial crops that were traditionally furrow irrigated are increasingly irrigated by drip irrigation methods. For example, a subsurface drip irrigation system was chosen to replace furrow and sprinkler irrigation systems on a commercial sugarcane estate in Swaziland. A postinvestment evaluation indicated that sucrose content increased by 15% and that 22% less water was used in the drip system than in the sprinkler irrigation system it replaced (4). Similar yield increases and water savings were noted in drip irrigated corn (maize) in a project in arid southeast Anatolia, Turkey (5).

CHANGES IN IRRIGATION OBJECTIVES OVER TIME

In the colonial period, the dominant reason for installing irrigation was to prevent famines during drought years. An additional concern was flood control in the deltas of the big rivers in Asia and Africa. Famine prevention necessitated an irrigation system that would provide some water to many farmers to assure the main harvest and revenue collection for the colonial power with a minimum of control and regulation. In the Asian subcontinent, this was achieved through a form of timed rotation of a continuous flow to different watercourses, in which a shortfall in supplies is shared equally by all users. The system was not expected to provide enough water for maximum yields.

From the 1950s, irrigation rapidly expanded, especially in Asia. The peak in funding by the international development banks for construction of irrigation occurred in the 1970s and 1980s. Internal rates of return on irrigation projects between 1965 and 1985 were in the range of 12–25% (6). This period coincided with the green revolution in Asia when cereal production increased many fold as a result of the introduction of high yielding varieties of rice and wheat, enhanced fertilizer use, and improved irrigation systems. During this period, irrigation's main objectives were to obtain maximum yields per unit land through high water and fertilizer inputs. Concurrently, extensive research was done to determine and predict water requirements and evapotranspiration rates, i.e., evaporation from the soil and transpiration from the growing crop. This topic was central to the efforts to match irrigation supplies as closely as possible with the evaporative demand of the environment to maximize crop production.

After about 1985, irrigation investments dropped sharply because of lower grain prices (from which the urban population benefited more than the farmers), rising construction costs, and hence lower returns on investments. Growing opposition to further large-scale irrigation projects, such as the Namada in India and the Three Gorges Dam in China, contributed to this reduced interest in irrigation investments. Although investments slowed, irrigation development in Asia nevertheless continued unabated. Experience and research have shown that supportive policies and laws have at least as much effect on irrigation development as does the availability of international investments. Specifically, national priorities, strength in the rule of law, a free market economy, local and user participation in planning and management, as well as encouragement of the private sector have been shown to stimulate irrigation development (7).

Pumped groundwater is increasingly used for irrigation. This increase has been at such an extent that many groundwater resources are now overexploited. Recycled drainage water and wastewater have also been added as possible sources of irrigation water. Generally, some treatment of wastewater has been deemed necessary before it can safely be used in crop production. But many developing countries do not have or enforce regulations covering the use of wastewater, or lack the resources to treat wastewater adequately. From the 1990s, the irrigation strategy focused on the control and integrated management of all water resources for agricultural and nonagricultural uses, which made irrigation managers realize that low irrigation efficiencies on the farm do not imply that water is wasted. If the drainage flow is fed back into the canal and reused downstream, and seepage to the groundwater is recycled by pumping, the relevant parameter is no longer application efficiency at the farm level but at the subsystem or system level, or even at the river basin level. For example, the Nile valley in Egypt with its notoriously inefficient irrigation at the farm level has a basin level application efficiency of about 95% because of its extensive reuse of drainage water (8). Hence very little water, and of no further use because of its low quality, is discharged into the Mediterranean Sea. These examples from Third World countries follow earlier developments of a similar nature in the United States and Australia (e.g., respectively, the Imperial Irrigation District with the Colorado River, and the Murray-Darling Basin).

IRRIGATION PERFORMANCE ASSESSMENT

Performance parameters express the degree to which the objectives of irrigation are met. The chosen parameters are usually limited to those pertaining to performance of the irrigation infrastructure and the system management. Because of its large data requirement, the variability in performance among farmers and for different crops is rarely studied. In one instance, where available information allowed quantification of the way farmers managed their irrigation supplies, it was found that around half the cotton farmers irrigated either excessively or deficiently although the average performance was close to the optimum for maximum yield per unit land (9). Excessive irrigation was also found in the Gezira irrigation system in Sudan. Here the variation was attributed to unattended furrow irrigation, aging infrastructure causing siltation, and weed growth in the canals (10). Large variabilities among farmers obviously indicate a significant potential for improvement.

In protective irrigation systems (implying proportional sharing of a limited resource among the widest number of people), performance is evaluated by measuring the flow rates that reach the tail ends of watercourses. This measurement assesses the equity of distribution of the available water to as many farmers as possible. Head-tail differences in water supplies continue to be the case in many large-scale irrigation systems. Failure to recognize and address equity considerations (i.e., fairness as understood by the farmers) in performance assessment has been a major cause of problems in operation and maintenance of irrigation system.

When the irrigation objective in the second half of the twentieth century shifted from protective to productive irrigation (involving the right to maximum economic return for a limited number of farmers), the performance parameter of choice became the yield per unit of land. Toward the end of the century, water for all uses became a major concern, and irrigation management came to be seen as one element of integrated water resource management. At the same time, performance assessment of irrigation systems was no longer based only on profitability and revenue collection, but it came to include the social objectives of food security, poverty alleviation, and environmental protection.

Because of increasing pressure on available water resources, irrigation managers worldwide are now compelled to improve the performance of their systems. The list of performance parameters has consequently grown and includes all aspects of maintenance and operation of the infrastructure, output per unit irrigation water and per unit of land, financial aspects (e.g., cost recovery), and environmental performance (e.g., quality and quantity of drainage flows). Benchmarking, using a standard against which performance can be assessed, has now been applied to irrigation management. By comparing with the norm, it attempts to identify deficiencies in a system's infrastructure and organizational structures. Benchmarking thus intends to make continuous improvements, especially in the quality and reliability of the services provided to the farmers. It also intends to stimulate an institutional culture that supports innovation and reliability in service providers (11).

IRRIGATION MANAGEMENT

Three main types of irrigation management can be distinguished: commercial estates producing usually one crop (e.g., sugarcane, cotton, tobacco, pineapples, etc.), canal irrigation systems with many small farmers, and individual farmers managing their source of irrigation water, usually a tubewell or small reservoir. Of these, large-scale canal irrigation systems cover the largest area, especially in Asia, and provide employment to millions of people. Improvements in the management of these systems could have a major beneficial impact on food production. This type of management, usually done by state organizations such as irrigation departments, has therefore received the most attention. Common institutional deficiencies in water resource management are described elsewhere (12).

Along field channels, and often significantly above the field channels, farmers control the actual operation of the system. In that part of the system that is controlled by management, including reservoirs and the main and secondary (or distributary) canals, considerable discrepancy often exists between the nominal rules and procedures incorporated in the design and the actual operation of the system.

Because of widespread dissatisfaction with state management of irrigation systems, financial institutions who invested in irrigation development have stimulated state agencies since the 1990s to hand over all or part of irrigation management to the private sector. Greater participation of the farmers in system management and user-operated and managed irrigation systems as found in developed countries are examples to be followed in the Third World. Privatization was expected to lower the cost of service provision to the farmers while increasing its quality. Conditions for success of privatization include:

- A legal framework either existing or tailor-made to provide the legal background for enforceable contracts
- Quality control and other regulations
- Active participation of the farmers in the implementation of the changes
- Institutions that control and enforce the rules and regulations
- Rights to water and use of the infrastructure
- Impartial adjudication of disputes
- Sound arrangements about the financial responsibilities regarding maintenance of the irrigation infrastructure (13)

An alternative to complete privatization was introduced in Pakistan where irrigation system management was decentralized and the participation of farmers in system management was promoted. Management at some secondary canal level was handed over to the farmers' organizations. A recent evaluation of one of these sites indicates that the water distribution and allocation has become more equitable as a result of the handover, which is particularly beneficial to tail-end farmers (14). Additional benefits were that illegal tampering of outlets had almost ceased, and that cost recovery had improved significantly. However, experience with irrigation management transfer elsewhere in developing countries has not been uniformly positive. Different performance parameters show improvement in some systems but decline in others. In some systems, an increase in farmers' fees for operation and maintenance and in fee collection has occurred, but nevertheless major maintenance is deferred. The effect of management transfer on agricultural production and farm income is equally uncertain.

In recent years, many countries have gained experience with institutional reform of their water sectors, including Australia, India, Philippines, Malaysia, Turkey, Mexico, Argentina, and Chili. Institutional changes will continue to shift toward user-management and water markets, with a more regulatory role for the state. Although the long-term impact of these various efforts has not yet been assessed, the Australian experience clearly demonstrates that if farmers are to pay for the full cost of irrigation services, the irrigation agencies have to become accountable to them.

SUSTAINABILITY AND FUTURE OF IRRIGATED AGRICULTURE

Irrigated agriculture competes for water with other uses and often contributes to degradation of land and water resources. The net social benefit of irrigated agriculture is the difference between returns to the farmer and society as a whole, and the cost to society associated with its negative aspects (e.g., the cost of disposal of saline and polluted drainage water). Because no general agreement exists on the real benefits and costs of irrigated agriculture, whether irrigated agriculture is sustainable in the long run has become controversial. Many definitions of sustainability exist. A strict definition is that sustainability is achieved when irrigation and drainage are conducted on-farm and within irrigation districts, in a manner that does not degrade the quality of land, water, and other resources (15). Others define sustainability of irrigated agriculture as the ability to continue extracting net positive returns from water and land for an indefinite period of time (16). In contrast to the first definition, the latter one is not inconsistent with some degree of environmental degradation, which implies that it is not necessarily true for all ecosystems that the optimal rate of degradation is zero.

Linked with this issue of sustainability is the question of whether there is enough irrigable land and water for future needs. FAO studies (17) suggest there is still scope for expanding irrigation to satisfy future needs. Their estimates suggest that in developing countries, only half of the potential irrigated land (about 400 million ha) is in use. Moreover, only 20% of this presently unused potential area will be in production by 2030. This assessment takes into account the limitations imposed by the availability of water. In the developing countries as a whole, only about 7% of renewable water resources were withdrawn (18) for irrigation. And the projection is a 14% increase in water withdrawals for irrigation by 2030. Water availability is considered critical when at least 40% of the renewable resource is used for irrigation. Of the countries listed in Table 2, the situation is critical only in Pakistan and Iran. FAO expects this to be the case for all of South Asia and West Asia/North Africa by 2030. Although a 14% increase in water withdrawn for irrigation over a period of 30 years may not appear to be much, the development of additional resources will be increasingly difficult. For example, monsoonal river discharges are part of the unutilized resource, but they are extremely expensive to capture.

In another study (19), the potential demand (i.e., without water supply constraints) and actual consumption (realized demand given the limitations in supply, and calculated for consumption rather than withdrawal) of irrigation water were assessed. This study expects the potential demand in developing countries to grow by 12%, but the actual consumption should grow less because of water scarcity and improvements in yields per unit water. The difference in growth of the potential and realized demands leads over time to lower values of the proportion of the potential demand realized in actual consumption of irrigation water, as shown in Table 3. The expected improvements in water productivity (kg/m^3) in irrigated cereal production range from about one third in sub-Saharan Africa to two thirds in South Asia.

With respect to the potential for irrigation expansion in the industrial countries, most studies agree that there

Table 3. Proportion of Potential Demand Realized in Actual Consumption

	1995	2025
Asia	0.81	0.76
Latin America	0.83	0.75
Sub-Saharan Africa	0.73	0.72
West Asia/North Africa	0.78	0.74

Source: Rosegrant et al. (19).

is less scope and need for expansion in America and Europe. At present some irrigated land in California goes out of production as the water rights are sold to cities for their domestic water supply. In both the United States and Europe, some rainfed agricultural land is now left idle as there is no need for more production. If the need developed, agricultural productivity could probably be enhanced on rainfed land at a lower cost than would be involved in expanding irrigated land. Moreover, future production increases in developed countries are expected to come from higher yields per unit of land and per unit of water.

It has been argued that water pricing might reduce the demand for irrigation water. However, water used for irrigation is not simply an economic resource but also a social good. A change in the price of water would affect the urban price of food, which in turn could have major political consequences. It would be more appropriate to charge for the water actually consumed than for what is supplied in irrigation, but that is practically impossible. Another practical difficulty with water pricing as an instrument for affecting its use is that for the price to have an effect on water demand, it must be significantly increased, most likely more than politically feasible. A case in point is China, where according to the World Bank, water for agriculture is heavily subsidized. A better option for improving water use efficiency in irrigation appears to be rationing water where the demand exceeds supply (20).

A contentious issue is how to improve water productivity (e.g., kg/m^3 or $\$/m^3$) when water resources are scarce and higher water rates are not appropriate. One option is through deficit irrigation: applying less water than the crop would need for maximum production per unit of land. Another option is to design a responsive irrigation system in which all farmers have access to water as and when needed, which is expected to lead to less wastage. The latter, however, is difficult and expensive to realize in large-scale irrigation systems with many fragmented small holdings. Providing farmers in deficit irrigation with less water than they would freely demand is expected to encourage farmers to grow more water-efficient and high-value crops. The downside of deficit irrigation, though, could be the possible salt buildup in the root zone in the absence of sufficient leaching of salts from the soil profile. Model studies have shown that on-demand irrigation is profitable where water is relatively abundant or high-value crops are grown that would benefit from extreme reliability and accuracy in irrigation supplies. Yet, deliberately restricting irrigation deliveries at no less than 80% of the potential full demand, crops can still produce well without salinization of the root zone. The productivity of water (kg/m^3) could then be almost 50% higher with deficit irrigation than when water is supplied to meet full crop demand (21).

The first challenge for the future, then, is to improve crop production on irrigated land by managing water better to achieve optimal production per unit of water, which is more easily realized in developed countries than in developing ones, but because of the population growth, the latter have a greater need for more food production. Present thinking is that more food export from industrial countries is not the solution. Imports have to be paid for in hard currency and staple crops should be grown where the people live. However, this could change as a result of globalization and reductions of tariffs in agricultural trade. Hence, for now and the foreseeable future, irrigated agriculture is essential for worldwide food security. The second challenge thus becomes to enhance irrigated agriculture while minimizing the adverse environmental impacts.

BIBLIOGRAPHY

1. FAO. (2002). *World Agriculture: Towards 2015/2030*. Food and Agricultural Organization of the United Nations, Rome, Italy.
2. Jurriens, J. and Lenselink, K.J. (2001). Straightforward furrow irrigation can be 70% efficient. *Irrigation and Drainage* **50**: 195–204.
3. Walker, W.R. and Skogerboe, G.V. (1987). *Surface Irrigation: Theory and Practice*. Prentice-Hall, Englewood Cliffs, NJ.
4. Merry, R.E. (2003). “Dripping with success”: The challenges of an irrigation redevelopment project. *Irrigation and Drainage* **52**: 71–83.
5. Yazar, A., Metin Sezen, S., and Gencel, B. (2002). Drip irrigation of corn in the southeast Anatolia Project (GAP) area in Turkey. *Irrigation and Drainage* **51**: 293–300.
6. Jones, W.I. (1995). *The World Bank and Irrigation*. Operations Evaluation Department, World Bank, Washington, DC.
7. Hargreaves, G.H. (2003). Lessons from success and/or failure of irrigation development. *Irrigation and Drainage* **52**: 31–38.
8. Keller, A. and Keller, J. (1995). *Effective Efficiency: A Water Use Concept for Allocating Freshwater Resources*. Water Resources and Irrigation Division Discussion Paper 22. Winrock International, Arlington, VA.
9. Lorite, I.J., Mateos, L., and Fereres, E. (2004). Evaluating irrigation performance in a Mediterranean environment. II. Variability among crops and farmers. *Irrigation Science* **23**: 85–92.
10. Ibrahim, A.A., Stigter, C.J., Adam, H.S., and Adeeb, A.M. (2002). Water-use efficiency of sorghum and groundnut under traditional and current irrigation in the Gezira scheme, Sudan. *Irrigation Science* **21**: 115–125.
11. (2004). *Irrigation and Drainage* **53**(2).
12. Frederiksen, H.D. (1996). Institutional principles for sound management of water and related environmental resources. In: *Water Resources: Environmental Planning, Management and Development*. A.K. Biswas (Ed.). McGraw-Hill, New York, pp. 529–577.
13. Vermillion, D. and Sagardoy, J.A. (1999). *Transfer of Irrigation Management Services: Guidelines*. FAO Irrigation and

Drainage Paper 58. Food and Agricultural Organization of the United Nations, Rome, Italy.

14. Latif, M. and Pomee, M.S. (2003). Irrigation management turnover: an option for improved utilization of limited water resources in Pakistan. *Irrigation and Drainage* **52**: 261–272.
15. Oster, J.D. and Wichelns, D. (2003). Economic and agronomic strategies to achieve sustainable irrigation. *Irrigation Science* **22**: 107–120.
16. Barker, R., Dawe, D., and Inocencio, A. (2003). Economics of water productivity in managing water for agriculture. In: *Water Productivity in Agriculture: Limits and Opportunities for Improvement*. J.W. Kijne, R. Barker, and D. Molden (Eds.). CABI Publishing, Wallingford, UK, pp. 19–35.
17. See endnote 1.
18. At field level, water withdrawn for irrigation includes, besides that actually transpired by the growing crop, evaporation from wet soil and water surfaces, leakage from channels, seepage below the water table and drainage water from the field.
19. Rosegrant, M.W., Cai, X., and Cline, S.A. (2002). *Global Water Outlook to 2025*. Averting an impending crisis. A 2020 Vision for Food, Agriculture, and the Environment Initiative Report. International Food Policy Research Institute, Washington, DC.
20. Perry, C. (2001). Water at any price? Issues and options in charging for irrigation water. *Irrigation and Drainage* **50**: 1–7.
21. Sarwar, A. and Perry, C. (2002). Increasing water productivity through deficit irrigation: evidence from the Indus Plains of Pakistan. *Irrigation and Drainage* **51**: 87–92.

IRRIGATION IN THE UNITED STATES

FREDERIC LASSERRE
Laval University
Quebec City, Canada

CONTRASTED TRENDS

The total amount of farmland, comprising both cropland and pastureland, is slowly decreasing in the United States at 71.2 million hectares (M ha) (176 million acres) in 1998, but irrigated lands kept increasing from 16.7 Mha in 1974, 19.8 Mha in 1982, 19.9 Mha in 1992, and 25 Mha in 2000. The geography of the increase in irrigated lands is contrasted. The national yearly increase for the period 1974–2000 is 1.57%, but east central and Atlantic states witnessed much stronger growth, well above 4%, on average, whereas Arizona, Montana, Nevada, Texas, Oklahoma, and Wyoming saw their irrigated lands shrink overall (Fig. 1) (1).

But this 26-year period masks a three-period evolution. From 1974 to 1982, irrigated surfaces grew on average by 2.18% yearly; sharp increases occurred in states such as South Carolina (+29%/year), Georgia (+22.7%/year), and Minnesota (+19.1%/year); other western states such as Texas and New Mexico already displayed a downward trend in their irrigated surfaces, and eastern States as well such as New Jersey and West Virginia. From 1982 to 1992, the growth is much more moderate: 0.08% per

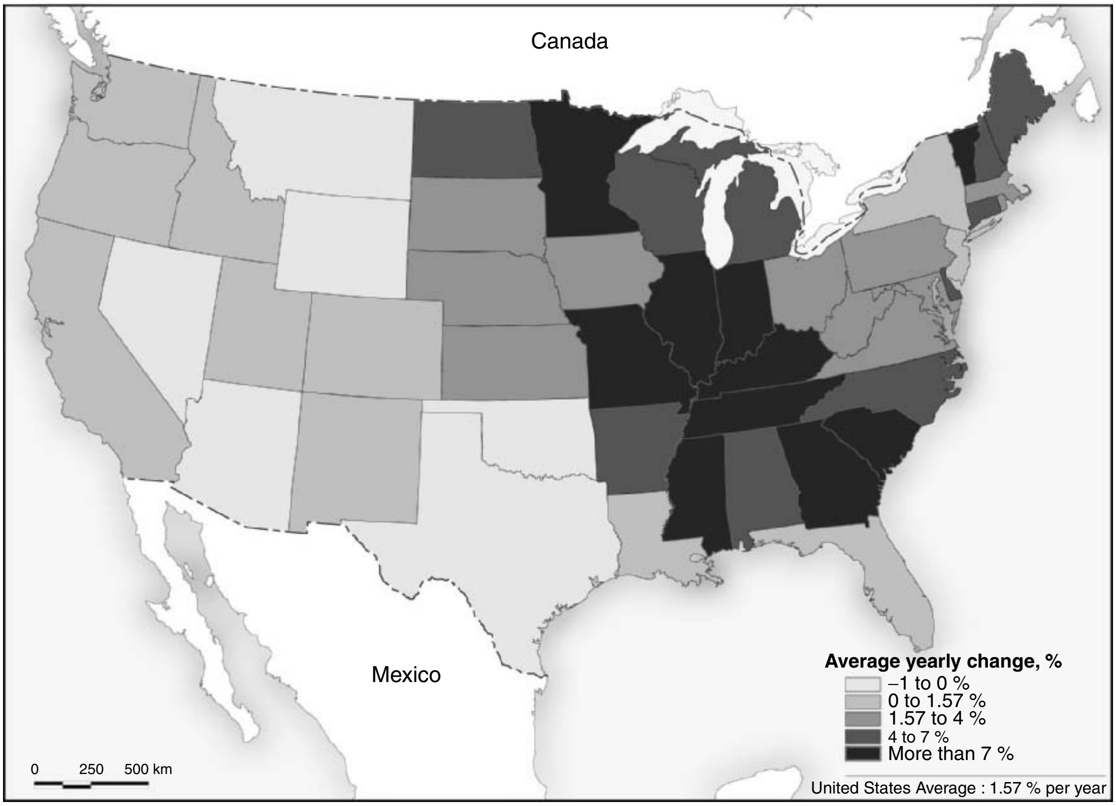

Figure 1. Irrigated surfaces evolution, 1974–2000.

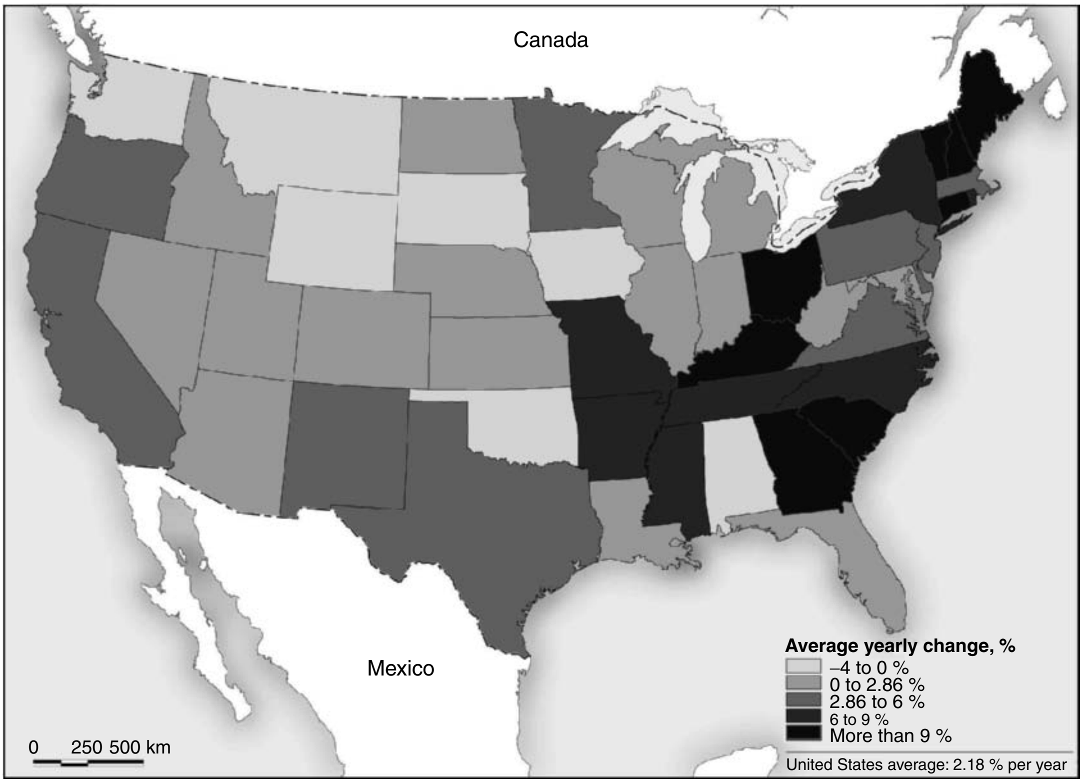

Figure 2. Irrigated surfaces evolution, 1992–2000.

year in the nation, with most states located in the western part of the country witnessed a decrease. Then growth picked up from 1992 to 2000 at a national yearly average of 2.86%. Some western states (Fig. 2) (Oregon, Montana, Wyoming, Oklahoma, South Dakota) still experienced a decline, but others grew appreciably (California: +3.7%/yr; Texas: +3.5%). The strongest growth was experienced in the Atlantic and east central regions, growth rates exceeded 9% in Georgia, South Carolina, Ohio, Kentucky, and several New England states (Tables 1 and 2).

This spectacular growth of irrigated surfaces in the eastern United States can be explained by two factors: a production objective, the desire of farmers to increase yields or switching to more water-demanding crops, such as corn, sugar, vegetables; a financial factor as well; many farmers in this wetter part of the country irrigate their cropland to avoid any risk of yield decrease. Tight markets, strong pressures from banks, and rising debt have made some farmers risk averse.

THE IRRIGATION COUNTRY: WATER IN THE DRY WEST

However strong the irrigation growth might have been in the eastern part of the country and even though a slight decline has begun to take place in a few western states, absolute irrigated surfaces still show that irrigation is concentrated mainly in the western and central United States. In 1974 as well as in 1992 (Figs. 3 and 4), the picture depicted by the maps underlines the dominant share in irrigated surfaces in states west of the 100th meridian: in 1974, irrigated land in California, Idaho, Colorado, Nebraska, and Texas accounted for 58.3% of all irrigated land. In 2000, irrigated land in the same five states still made up 51% of irrigated land countrywide.

Of course, natural conditions largely explain this spatial pattern: rainfall becomes much scarcer west of the 100th meridian and becomes more abundant again only in the northwestern part in the country, on the Pacific coast (Fig. 5). Most of the west gets less than 20 inches of rain a year; patches of California, Nevada, Arizona, Idaho, Utah, Wyoming, and even Washington receive less than 10 inches. Water is present, however, in the region: aquifers do contain large amounts of water (for instance, the Ogallala aquifer underneath Texas, Oklahoma, and Nebraska). And mountain ranges generate rivers; the Colorado drains a large basin in the area. Tapping this resource became a political objective during the conquest of the West during the second half of the nineteenth century to tame nature and embody the conquest of the land, even though it implied transferring agricultural techniques that belonged to the wetter eastern part of the country. Water schemes were designed to develop the resource, rivers were dammed, and canals built to divert rivers, paid for by the federal budget and delivering water that was and still is largely subsidized. Agriculture is flourishing in the arid west, but it is only thanks to the development of what Donald Worster (3) calls a "hydraulic society," where water is the main resource and cements the very fabric of the economy and society. This explains why, today, it is still a very political and emotional issue fiercely debated, although the age of large-scale dams and developments seems to be gone.

Table 1. Top 15 States with Increases in Irrigated Surfaces, 1974–2000 and 1992–2000

1974–2000	%/year
South Carolina	11.78
Georgia	10.60
Vermont	9.19
Missouri	8.74
Mississippi	8.72
Indiana	8.05
Minnesota	7.98
Illinois	7.64
Tennessee	7.23
Kentucky	7.20
Maine	6.99
Alabama	6.41
Arkansas	6.18
Michigan	5.76
Delaware	5.59
North Carolina	5.29
1992–2000	**%/year**
Connecticut	17.20
Maine	17.07
New Hampshire	17.07
South Carolina	11.91
Vermont	11.45
Kentucky	11.44
Georgia	9.87
Ohio	9.51
Missouri	8.18
North Carolina	7.13
New York	7.09
Arkansas	6.61
Rhode Island	6.59
Tennessee	6.34
Mississippi	6.12
New Jersey	5.98

Table 2. Fifteen States Experiencing the Lowest Increase in Irrigated Surfaces, 1974–2000 and 1992–2000

1974–2000	%/year
New Jersey	1.39
Oregon	1.27
Louisiana	1.13
Florida	1.08
Idaho	1.05
California	1.02
Washington	0.70
Colorado	0.65
New Mexico	0.54
Oklahoma	−0.06
Texas	−0.06
Montana	−0.09
Arizona	−0.64
Nevada	−0.70
Wyoming	−0.88
1992–2000	**%/year**
Illinois	1.33
Wisconsin	0.89
Colorado	0.88
Maryland	0.77
Louisiana	0.57
Indiana	0.46
Arizona	0.26
West Virginia	0.00
Oklahoma	−0.12
Washington	−0.55
South Dakota	−0.58
Montana	−1.73
Alabama	−1.96
Wyoming	−2.88
Iowa	−3.85

WATER WITHDRAWALS

The geography of water withdrawals for irrigation (Fig. 6) confirms the dominant role of irrigation in agriculture in the West: large withdrawals take place in this part of the country in every state but a few of them and reaches levels as high as 70.1 billion m^3/year in California (see Table 3).

But the geography of water withdrawals also shows that eastern states withdraw appreciable volumes of water for irrigation: Florida, of course (9.1 billion m^3/year), largely because of the fruit and sugar industries; but also Georgia (2.2 Bm^3/year); North and South Carolina (cotton, peanuts, corn, tobacco); New Jersey; Massachusetts (0.33 Bm^3/year); and Michigan. In the three latter states, large water withdrawals are largely accounted for by the redevelopment of fruit and vegetable cultures for nearby urban markets. Central States in the Missouri/Mississippi basin also use significant water for irrigation: Nebraska (14.4 Bm^3/year), Arkansas (13.3 Bm^3/year), Kansas (5.7 Bm^3/year), Mississippi (2.1 Bm^3/year), Missouri (2.1 Bm^3/year), Louisiana (1.8 Bm^3/year), and Oklahoma (1.2 Bm^3/year).

Some states rely heavily on groundwater for irrigation. In the lower 48 states, 19 states rely on groundwater for more than 40% of irrigation withdrawals; levels are as high as 49.7% for Montana and 49.1% for Kentucky. On the other hand, 17 states pump less than 30% of their irrigation water from aquifers. Reliance on groundwater is not connected to the size of the withdrawals; states along the lower Mississippi and Missouri basins and around the Great Lakes tend to pump a larger proportion of their irrigation water from surface water.

As a result of the huge withdrawals taking place in the West for irrigation, shares of irrigation water in total freshwater withdrawals (Fig. 7) are nearly always above 60%, except for Kansas (56.3%), Oklahoma (40.7%), Texas (34.8%), and North Dakota (12.7%) (Table 4).

In the eastern United States, the share of irrigation is consistently below 15%, except for Arkansas, Florida, Mississippi, Georgia (17.8%) and Missouri (17.4%), and is less than 1% in 12 states (Indiana, Vermont, Kentucky, Iowa, Virginia, New York, Alabama, Ohio, Tennessee, Pennsylvania, West Virginia, and the District of Columbia). Thus, although it is developing fast in the eastern part of the country and may at times in specific places, represent a problem for water supply in the summertime, irrigation remains by far the major demand sector only in the West. This explains the pressure exerted on ground or surface waterbodies by environmental problems such as the overexploitation (Colorado River)

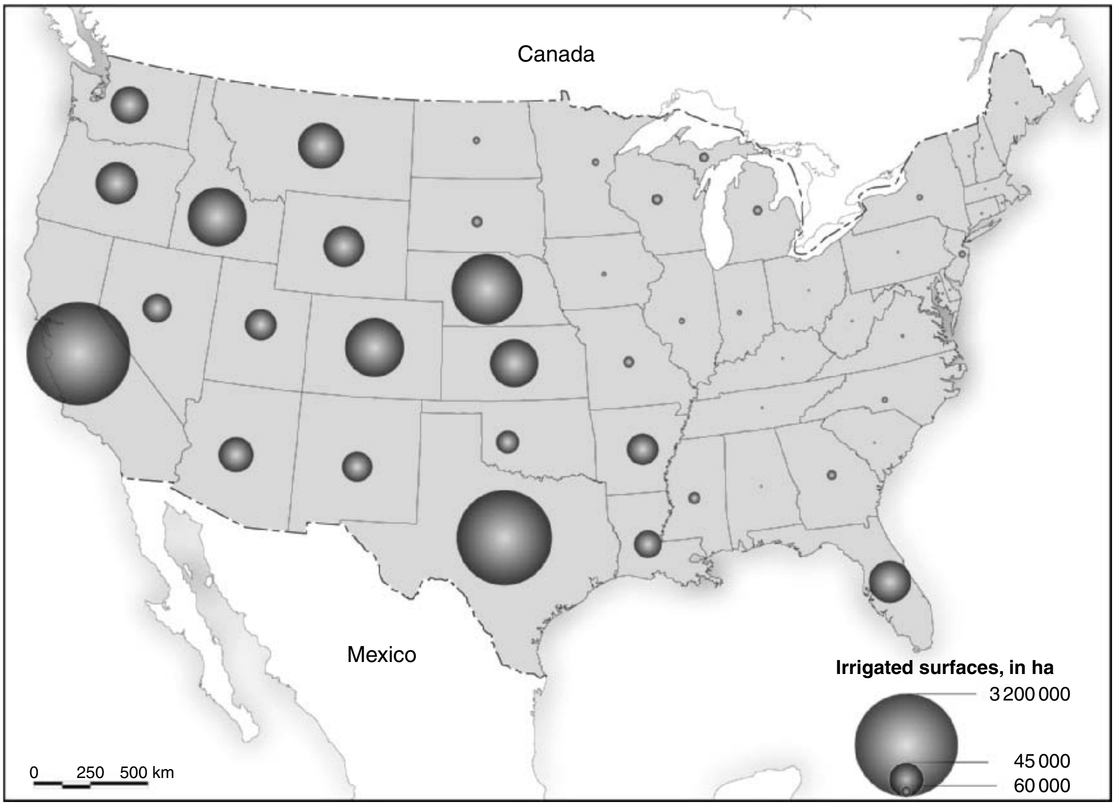

Figure 3. Irrigated surface per state, 1974.

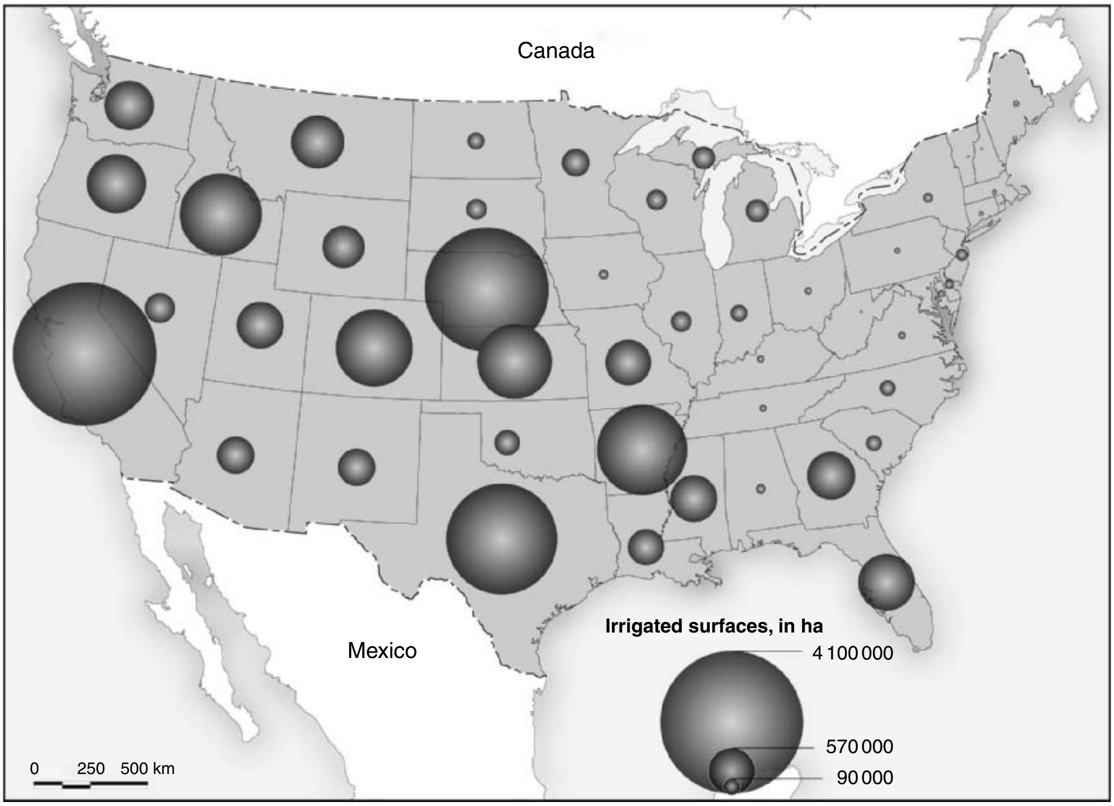

Figure 4. Irrigated surface per state, 2000.

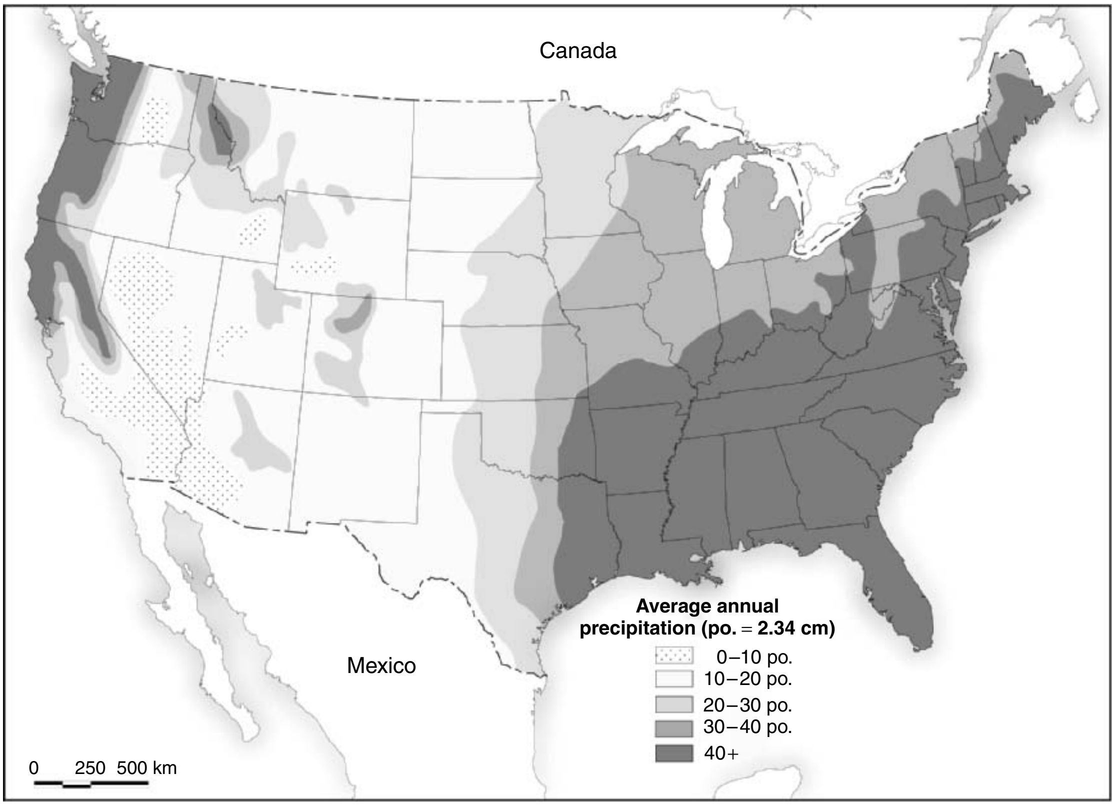

Figure 5. The dry western United States (1,2).

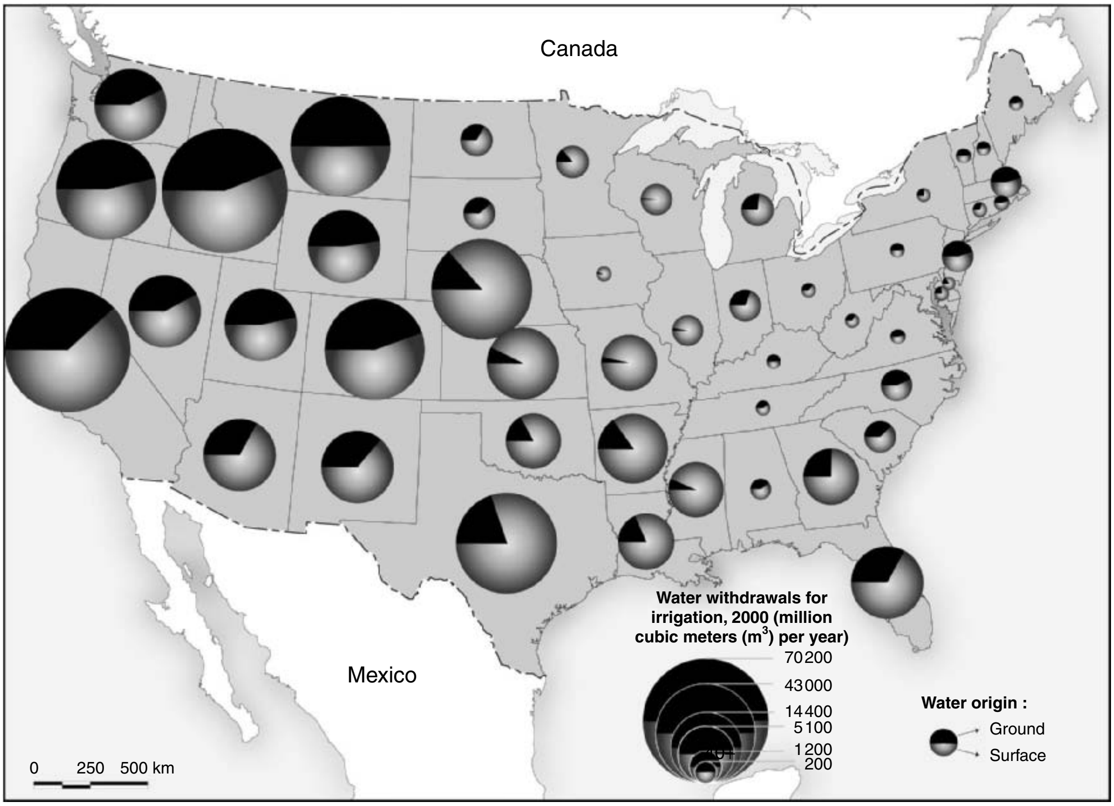

Figure 6. Water withdrawals for irrigation, 2000.

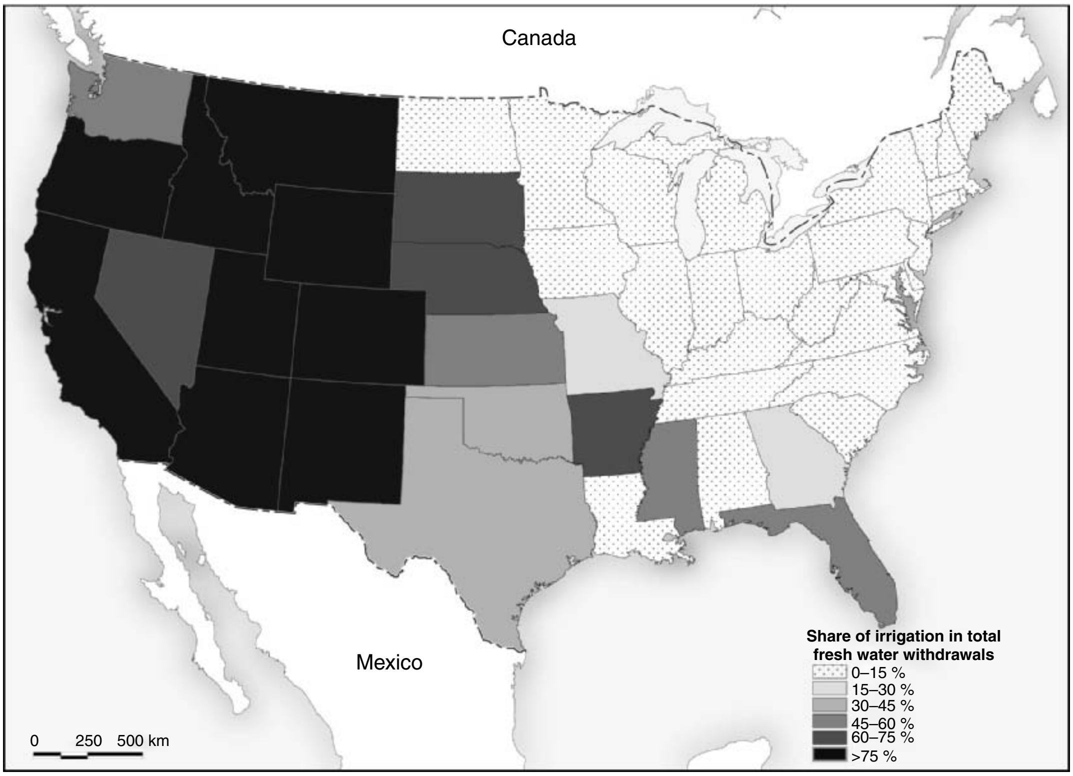

Figure 7. Share of irrigation in total fresh water withdrawals, 2000.

or the salinization of rivers (such as the Gila River). It also entails a developing political debate as to whether the sharing of the resource should leave so much subsidized water to the agriculture sector, as thirsty cities and industries have begun buying up expensive water rights from farmers to satisfy their fast growing needs.

FARMING AND IRRIGATION TECHNIQUES

The agricultural sector displays a contrasting picture, as can be seen in Fig. 8. If absolute water withdrawals for irrigation are very large in the West and if agriculture there remains by far the largest demand sector, the share of irrigated land in total farmland is not necessarily dominant (Table 5).

California, Arkansas, and Mississippi are the only states whose share of irrigated land represents more than half of the total exploited land. The extent of fruit and vegetables crops in California, as well as cotton growing in the Mississippi valley, accounts for such high shares. Strong fruit and vegetable cropping (Washington, New Jersey, Michigan, Florida, Delaware), sugar (Florida), potato (Idaho), tobacco and peanuts (Georgia) are, in their respective states, dominant cultures where irrigation is extensively practiced, to enable cropping (as for most cultures in the West) or to reduce, even eliminate, the financial risk that could stem from poor rains in an otherwise pluvial agriculture.

Table 3. Top Twenty States for Irrigation Water Withdrawals, 2000 (in Billion Cubic Meters per Year)

California	Idaho	Colorado	Montana	Oregon
70.1	43	29.4	22.5	16.1
Texas	Nebraska	Arkansas	Wyoming	Arizona
15.3	14.4	13.3	12.2	11.5
Utah	Florida	Washington	New Mexico	Kansas
10.3	9.1	7.6	6.4	5.7
Nevada	Georgia	Mississippi	Missouri	Louisiana
5.2	2.2	2.14	2.1	1.8

Table 4. States Whose Share of Irrigation in Total Fresh Water Withdrawals is Greater Than 50%, 2000

State	%
Montana	95.93
Wyoming	91.15
Colorado	90.63
Oregon	87.76
New Mexico	87.73
Idaho	87.28
Utah	81.07
Arizona	80.51
California	79.43
Nevada	74.98
Arkansas	72.66
Nebraska	72.05
South Dakota	70.64
Washington	57.63
Kansas	56.25
Florida	52.70
Mississippi	50.14

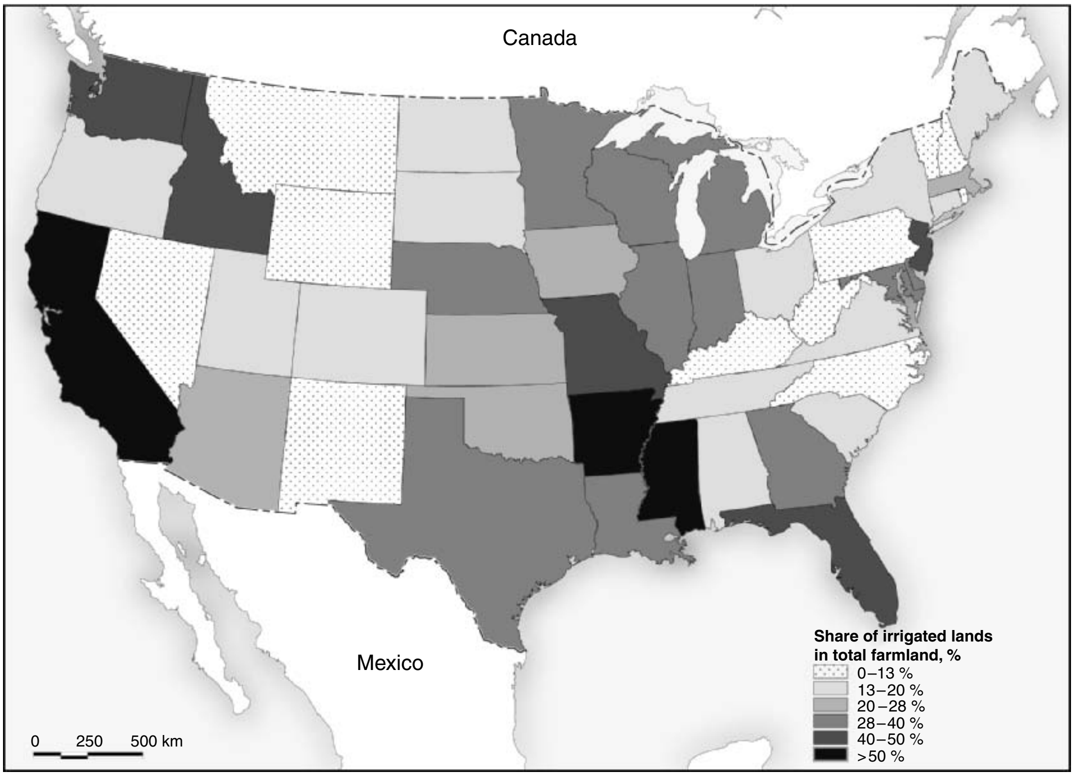

Figure 8. Share of irrigated land in farmland, 1998.

Table 5. States Whose Share of Irrigated Land in Farmland (1998) is Greater Than 30%

State	%
Maryland	31.07
Illinois	31.32
Nebraska	31.37
Georgia	31.89
Delaware	32.22
Minnesota	37.18
Texas	38.16
Louisiana	39.10
Wisconsin	39.26
Michigan	39.78
New Jersey	41.78
Florida	41.78
Missouri	43.02
Idaho	43.33
Washington	45.72
Mississippi	52.36
Arkansas	65.17
California	65.90

These risk-averse farming sectors partly explain why the water application rate per surface unit can be so high in the more humid East (Fig. 9). The low cost of water is certainly a factor, but it rather accounts for profligate rates in the West, as water remains heavily subsidized in this part of the country, whereas farmers in the East, who produce high-value crops such as fruits and vegetables, reason that the extra water spread in their fields is worth the guarantee that they will not suffer from a few percent loss in their crops because of poor rains (Table 6). This is a consequence of the increasing role of financial institutions in agribusiness, as banks have stepped up their pressure on farmers who, in turn, tend to be willing to spend more on water to avoid potential losses from natural factors.

The cost factor is certainly better reflected in the irrigation technique employed. Irrigation in the East is dominated by the sprinkler technique, an easy-to-implement technology with a higher water productivity than gravity or surface watering. Drip irrigation is also

Table 6. Top Fifteen States for Water Application Rate in Thousand m^3/ha, 2000

State	Rate, 000 m^3/ha
Montana	24.1
Arizona	21.7
Idaho	21.2
Massachusetts	20.97
Wyoming	19.4
Colorado	15.96
Nevada	14.87
Oregon	13.87
Utah	13.54
California	12.8
New Mexico	11.8
Washington	8.9
Florida	8.2
North Carolina	6.8
South Carolina	6
Arkansas	5.4

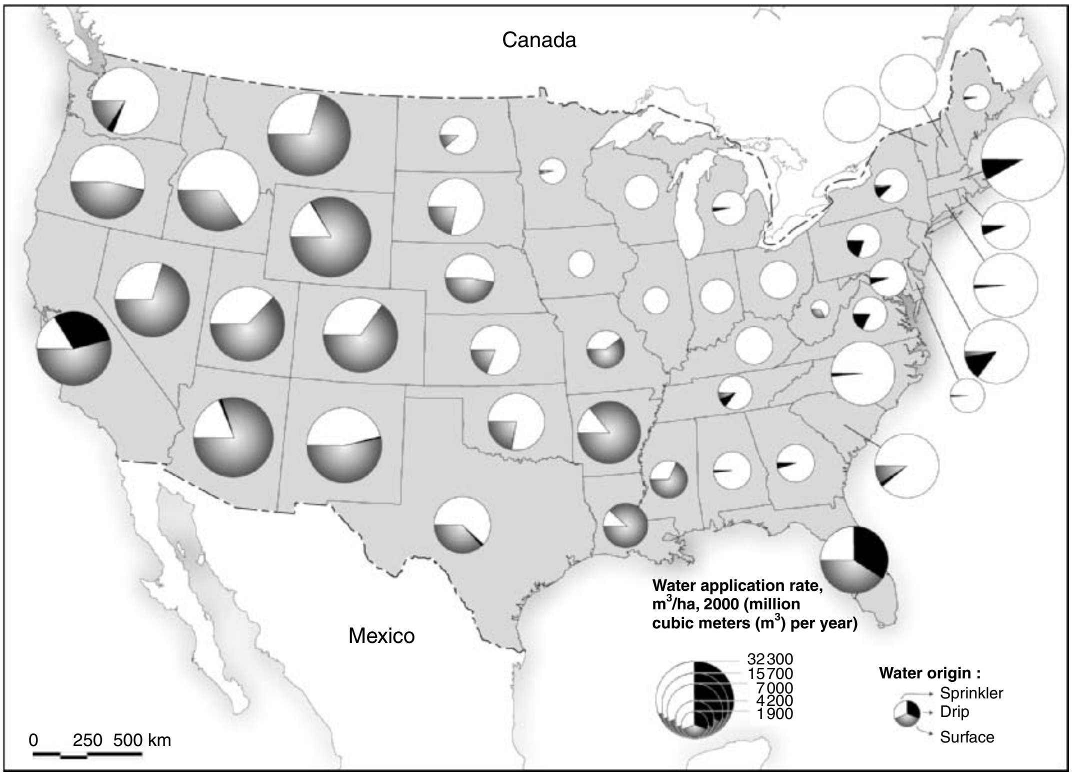

Figure 9. Water application rate, m^3/ha, 2000.

used in Florida (34.1% of volume of water spread), Pennsylvania (20%), Virginia (17.9%), New Jersey (12.3%), New York (10.8%), and Massachusetts (8.3%). But west of the Mississippi or in its lower course, surface irrigation is either dominant or holds a significant share of total water used in fields (see Table 7).

Whereas it is well-documented that its water productivity is low (<40%), in water-scarce areas, the extent of surface irrigation definitely implies improvement potential but points to a major political problem, the very low cost farmers still pay for water carried by expensive water-development projects paid for by all taxpayers in the United States. Fees collected by irrigation districts often do not even pay for the direct costs of moving the water from the dams to the fields, let alone refund the investment.

Table 7. States Where the Share of Surface Irrigation is Greater Than 30%, 2000

State	Share, % of Total Water Used for Irrigation
Louisiana	88.30
Arkansas	86.03
Wyoming	83.10
Arizona	79.82
Montana	70.93
Nevada	70.48
Mississippi	68.03
Colorado	65.29
Utah	62.41
Missouri	59.55
California	54.16
New Mexico	53.11
Nebraska	47.44
Oregon	46.08
Florida	40.73
Texas	36.83
Idaho	34.67
West Virginia	33.33

The low share of sprinkler and drip irrigation in water-scarce areas creates a large productivity improvement potential. Converting a fraction of surface-irrigated fields into sprinkler-irrigated fields could vastly reduce the water application rate. There is a significant relation in the West between a large share of surface irrigation and high application rates. Turning to a more sprinkler-intensive agriculture could reduce water withdrawals and consumption and help relieve the region from the ongoing, bitter, water-sharing debate. But this cannot be done by mere campaigns in favor of more effective watering techniques; farmers are well-aware of them. The issue again boils down to the pricing of water, which is definitely a question that most politicians would not want to tackle directly.

Broadening the topic, these observations underline the structure of water use that agriculture has developed in the United States as of 2000. In the East, a burgeoning and fast-growing irrigation sector relies on water-saving techniques but can use water generously. This water use does not stem from climate constraints, but from financial reasoning, avoiding risk and increasing surface productivity. In the West, water use in irrigation was the only way to develop intensive agriculture, with crops imported from Europe or the more humid East. But, because of the history of agricultural development, water prices still prevent the adoption of more efficient irrigation techniques. In the context of a slow erosion of irrigation in the West due to rising competition for water from cities and the industry and also due to unsustainable water withdrawal rates in several river basins and/or aquifers, it is important to take into account the improvement potential that agriculture displays in its water use. Massive water diversions are probably not necessary because agriculture, by far the main water user in the West, can reduce its consumption by turning to more efficient irrigation systems. But this comes at a cost.

Acknowledgments
This research was made possible thanks to a grant from the FQRSC.

BIBLIOGRAPHY

1. Riebsame, W. and Robb, J. (Eds.). All maps are designed with data from USGS and the Agricultural Census, except Fig. 5, adapted from (1997) Atlas of the New West. University of Colorado at Boulder Press, WW Norton.
2. Lasserre, F. and Descroix, L. (2003). *Eaux et Territoires*. PUQ, Québec.
3. Worster, D. (1985). *Rivers of Empire. Water, Aridity, and the Growth of the American West*. Oxford University Press, New York.

IRRIGATION WELLS

Muhammad Nadeem Asghar
International Water
Management Institute (IWMI)
Lahore, Pakistan

In arid and semiarid areas, rainfall is markedly variable in magnitude, time of occurrence, and spatial distribution. Irrigation is required to meet the demand of agriculture, but surface water supplies are also characterized as scarce and unreliable. Consequently, farmers consider it necessary to create a dependable water supply, which is adequate and in a desired location. This management option has spurred farmers to develop the technology of irrigation wells.

History started with the development of irrigation wells to be used for lifting water from rivers (surface water), surface storage (rainfall), and subsurface dug holes (underground water). The discharge rates of such irrigation wells depend on the capacity of the lifting devices.

As pumps were introduced, a boom in another kind of irrigation well occurred to exploit underground water for irrigated agriculture. Promising technologies of irrigation wells include tube wells, skimming wells, dug wells, and radial wells. The discharge rates of these irrigation wells depend on the capacity of the lifting devices and also on the hydrogeologic characteristics of the aquifer.

A tube well has a single strainer that penetrates vertically into the aquifer and operates under suction to extract groundwater at a discharge rate usually greater than 28 L/s. In unconfined aquifers that have fresh groundwater lenses underlain by saline groundwater layers, these wells are prone to extract groundwater of deteriorated quality due to upward movement of saline groundwater that enters the well screen. The thickness of these fresh groundwater lenses varies from a few meters to more than 30 meters, depending on the recharge source. In the surroundings of canals and rivers, this thickness can be more than 30 meters. The diameter of tube-well strainers is usually greater than 10 cm. The greater the thickness of the fresh groundwater lens, the higher the possibility of extracting greater discharge, which requires a larger diameter pipe for the strainer. This type of irrigation well, however, can extract groundwater without the problem of deteriorated quality provided that the well penetrates at less than 60% of the depth of the fresh groundwater lens and its operating time does not exceed 6 hours in any day (1).

For unconfined aquifers where the thickness of the fresh groundwater lens is less than 25 meters, skimming wells are introduced. This type of irrigation well is multistrainer, spread at some horizontal distance from the suction point, and these strainers are joined to a common pump. Like tube wells, the strainers in skimming wells penetrate vertically into the aquifer. However, designing a skimming well to include four to six strainers of 5 cm diameter that spread at a 1.5-m horizontal distance from the suction point makes the design cost-effective and minimizes operation and maintenance requirements. These strainers should be installed at less than 60% of the depth of the fresh groundwater lens of the aquifer. To reduce the risk of quality decline, its operating time should also not exceed 6 hours in any day. This type of irrigation well also works under suction and can usually extract groundwater at discharge rates less than 28 L/s (2).

Dug wells can be used to extract fresh groundwater from aquifers when thickness of fresh groundwater lenses is less than fifteen meters. Depending on the well diameter, there are two kinds of dug wells found in various parts of the world; small dug wells and large dug wells. In both dug wells, groundwater gathers under gravity, and pumps are used to extract this groundwater. Dug wells are mostly suitable for smallholder irrigated agriculture.

A large dug well has a diameter large enough for workers to enter it for construction and maintenance. This type of irrigation well can be dug as deep as 90% of the depth of the fresh groundwater lens of the aquifer and can still avoid the risk of entering the underlying saline groundwater at the bottom of the dug well (3).

A small dug well, on the other hand, has a diameter too small to enter and is constructed from the ground surface using special tools. The treadle pump, which is a human-powered pump, also uses this type of irrigation well as a source of groundwater. For small-scale irrigation, these pumps were first introduced into Bangladesh in the early 1980s, and they are now making their mark in Africa (4).

Radial wells are also multistrainer wells. Unlike tube wells and skimming wells that have vertically installed strainers, radial wells have horizontal strainers. These strainers facilitate skimming a very thin (even less than 3 meters) fresh groundwater lens as they provide large discharges per unit of drawdown—a desirable feature to avoid upward movement of saline groundwater (5). These strainers (usually of 5 cm diameter) can be joined directly to the sump—a collection dug hole, or they are first connected to collector pipes (usually of 10 cm diameter), and then these pipes are joined to the sump. Groundwater comes under gravity to these horizontal strainers, and pumps are used to extract it from the sump. This type of irrigation well, which is an innovative and modern type of dug well, can also be used to support smallholder irrigated agriculture.

BIBLIOGRAPHY

1. Asghar, M.N., Prathapar, S.A., and Shafique, M.S. (2002). Extracting relatively fresh groundwater from aquifers underlain by salty groundwater. *Agric. Water Manage.* **52**: 119–137.
2. Asghar, M.N., Saeed, M.M., Ashraf, M., and Qureshi, A.S. (2002). Potential of skimming wells for extracting relatively fresh groundwater from unconfined aquifers underlain by saline groundwater. *Proc. 2nd Int. Conf. New Trends Water Environ. Eng. Safety Life: Eco-compatible Solutions for Aquatic Environ.* June 24–28, Capri, Italy.
3. Asghar, M.N., Shafique, M.S., Ahmad, S., and Kahlown, M.A. (2002). *Root Zone Salinity Management Using Fractional Skimming Wells with Pressurized Irrigation: Annual Progress Report 1999–2000*. Working Paper No. 51, International Water Management Institute, Lahore, Pakistan.
4. Kay, M., and Brabben, T. (2000). *Treadle Pumps for Irrigation in Africa*. IPTRID. International Programme for Technology and Research in Irrigation and Drainage. Knowledge Synthesis Report No. 1.
5. Kelleners, T.J. (2001). *Effluent Salinity of Pipe Drains and Tube Wells: A case study from the Indus plain*. Ph.D. Thesis, Wageningen University, Alterra ILRI, Wageningen.

AGRICULTURE AND LAND USE PLANNING

Andy Seidl
Colorado State University
Fort Collins, Colorado

Julia Freedgood
American Farmland Trust
Washington, District of Columbia

Fen C. Hunt
U.S. Department of Agriculture
Washington, District of Columbia

INTRODUCTION

Many owners of agricultural lands, particularly those at the urban fringe, face strong pressures to convert their lands to residential or commercial uses. If these forces of change occurred in an environment of perfectly functioning markets, then there would be no economic problem or failure and no economic justification for public policy in the form of planning. We would simply observe the optimal rate of land conversion to reflect societal preferences and changes in income and population over time. However, the forces contributing to agricultural land conversion are rife with informational and policy- driven distortions in both agricultural and nonagricultural land markets, among them, zoning and other local policies that encourage development and inflate land values beyond the reach of agriculture, high-risk and low-economic returns to commodity production, higher incomes and the provision of health insurance in off-farm occupations, unintended consequences of state and federal farm and nonfarm policies, and improvements in telecommunications and in other infrastructure that mute the cost advantages of urban living. Moreover, agricultural lands provide many types of value to landowners, private citizens, communities, and society at large that are not reflected in the market value of agricultural products or in the capitalized value of those products in establishing land lease and sale prices. As a result of both supply- and demand-side dynamics, the market for agricultural land usually fails to reflect social values. Therefore, public policy in the form of some sort of land-use planning is justified.

The American public values agricultural land for many reasons (1,2), beyond its ability to produce food, fiber, and timber. These include curbing negative impacts of ill-managed growth; providing flood control; improving water quality; maintaining groundwater rechargeability; mitigating air pollution, ozone, and greenhouse gas emissions; providing community buffers or greenways; reducing traffic congestion; and easing the fiscal burden of public services. Such values can be reflected within a community by the relative intensity of alternative development; the opportunity for public goods and services, such as parks and trails; wildlife habitats, scenic views, and recreational opportunities; and proximity to urban dwellers (see, for a variety of sources, Ref. 3).

Public values for agricultural land preservation have translated into public support for establishing agricultural land preservation programs at the local, state, and federal levels. A mixture of both regulatory (4,5) and incentive-based approaches (4,6) have resulted in the permanent protection of nearly 1.5 million acres of agricultural land from conversion to nonagricultural uses (7) and the concentration of millions of more acres in agricultural production through zoning. Agricultural land protection through the various tools of land-use planning is a complex issue and process. This article provides only a glimpse of fundamental approaches and resources available for agricultural land-use planning and preservation.

AGRICULTURAL LAND USE PLANNING ISSUES

What land should be preserved and for what purposes? What level of private and public resources should be made available for fee simple land or development rights purchases? How should priorities be set to protect critical and sensitive areas with limited private and public funds to achieve the best results? Both private and public institutions must address these questions and concerns.

Agricultural land-use planning and preservation constitutes sensible resource management for sustainable agriculture and more broadly for sustainable development. It also stands on its own as a local, state, or national policy goal to mitigate land fragmentation and haphazard development. Various tools, resources, and model programs are available to facilitate agricultural land-use planning and preservation. Each person and community has unique goals, challenges, and collective knowledge that will determine the appropriate mixture of tools to guide its private and public actions in crafting effective local land-use management and policy (8).

Flexibility and knowledge in local conditions are the most pervasive features of successful programs to preserve lands for multiple functions meeting agricultural operations, managing land use, and achieving environmental, social, and economic vitality objectives (4). Scientists and educators at the land-grant universities provide research-based information, analyses, and educational programs that will help facilitate this process at the local and state levels.

Public programs for agricultural land-use planning and preservation may benefit both landowners and the general public. Agricultural land-use planning and preservation programs may be motivated to preserve production agriculture for national food security or for local food choice reasons. In this case, the following are many of the key characteristics typically considered to prioritize land parcels chosen for preservation:

1. Soil quality and productivity
2. Agricultural infrastructure
3. Farming methods, including conservation practices
4. Unique or critical land quality
5. Imminent threat of conversion
6. Critical size of a parcel for a viable agricultural operation
7. Proximity to other protected land for a critical mass to achieve effectiveness
8. Importance to local agricultural and economic vitality (9,2).

Alternatively, the public may be more concerned with issues related to urban growth pressure or preservation of ecological integrity and cultural resources (10). They may perceive preserving agricultural land as a way to sustain community values, lifestyle, or its natural and cultural heritage. In this case, land preservation programs may rank land parcels by the ability to provide direct environmental or economic benefits of open space and scenic beauty, economic opportunity, such as agro-tourism, wildlife habitats, community buffers or greenways, source water or watershed protection, clean air or carbon sequestration, historical or cultural significance, recreational opportunities, such as hunting, curbing current and anticipated development pressure, and/or complementary community practices, such as agricultural zoning and growth management (1,11).

NATIONAL, STATE, AND LOCAL AGRICULTURAL LAND PLANNING AND PRESERVATION PROGRAMS

U.S. Department of Agriculture (USDA)

Public interest in agricultural land protection, especially against fragmentation and haphazard development, has increased steadily since the 1950s. As this interest has grown, the preservation toolbox has expanded from local and state to the federal level. USDA programs designed to preserve working agricultural landscapes include the Farm and Ranch Lands Protection Program (FRPP), the Forest Legacy Program (FLP), and the Grassland Reserve Program.

The FRPP (www.nrcs.usda.gov/programs/FRPP) leverages federal funds with state and local funds to purchase conservation easements on prime and locally important or unique land by limiting conversion to nonagricultural uses. The FLP (www.fs.fed.us/spf/coop/programs/loa/flp.shtml) establishes partnerships between the Forest Service (USFS) and state forestry agencies to protect environmentally important private forestlands from conversion to nonforest uses. Through conservation easements or rental agreements, the GRP (www.nrcs.usda.gov/programs/GRP) protects and restores grasslands from conversion to cropland and other nonagricultural uses and enables viable ranching operations to restore plant and animal biodiversity.

Many state and local ranking systems are based on variations of the federal Land Evaluation and Site Assessment (LESA) system (www.nrcs.usda.gov/PROGRAMS/lesa/index.html). LESA encompasses a portfolio of land attributes to estimate the relative value of agricultural land. LESA provides a framework to measure the merits of keeping individual parcels in agricultural operation based on local preservation objectives. LESA recently has been tied to GIS technology to facilitate mapping and planning scenarios.

State and Local Programs

Over the past 50 years, public organizations at the state and local levels have used a variety of tools to preserve agricultural land. Examples of these tools include differential tax assessment programs, agricultural district programs, right-to-farm or nuisance protection laws, transfer of development rights (TDR) programs, purchase of agricultural conservation easements—also known as development rights (PACE or PDR) programs—growth management or urban boundaries and agricultural zoning, among a variety of other tools. Private organizations such as land trusts also have gotten involved, typically by accepting donated conservation easements or facilitating the process for landowners to sell easements to a public program.

Beginning in 1956 with the inception of Maryland's differential property tax assessment program, today every state except Michigan has a differential property tax assessment program that assesses agricultural land based on its use value, instead of its fair market value. The program either reduces property taxes on agricultural lands or defers taxes for as long as the land remains in agriculture. Michigan, Wisconsin, and New York also have programs to allow farmers to claim income tax credits to offset their local property tax bills.

California enacted the first agricultural districts law in 1965, commonly known as the Williamson Act. Landowners are allowed to create "agricultural preserves" areas. New York was the first state to create a comprehensive agricultural district program to protect farmland and support the farming business. Agricultural district programs are authorized in 16 states (7) to create legally recognized geographic entities where agricultural activities and their land bases are encouraged and protected. In addition, all 50 states have nuisance protection statutes (also known as "right-to-farm" laws) for agricultural operations.

In 1967, Boulder County in Colorado (www.co.boulder.co.us/lu) established a TDR program to protect agricultural lands and open space. This program enables landowners to transfer the development rights on one parcel of land to another parcel of land, such as from an agricultural zone to designated higher density development areas. Besides maintaining working agricultural landscapes, TDR programs may be designed for multiple purposes, such as to conserve environmentally sensitive areas or preserve historic landmarks. As of 2000, Montgomery County in Maryland had more than 40,000 acres, which accounted for 60% of the national total, enrolled in the TDR program (7).

In 1974, Suffolk County in New York enacted the first PACE (also known as purchase of development rights, PDR) program. King County in Washington and the states of Maryland, Massachusetts, and Connecticut quickly followed suit. PACE programs are voluntary on the landowner's part and permanently protect agricultural land by removing the development rights. As of 2004, PACE programs operate in 23 states, including 19 statewide and more than 45 local programs (7). Five additional statewide programs have been authorized to acquire easements but have not purchased any to date. In particular, local PACE programs are expanding probably because of incentives from the federal FRPP, which provides matching funds to support local, state, and private easement acquisitions, as well as from landowner and community interest in saving land for agriculture. From 2000 to 2004, the number of local PACE programs increased 32% (7).

Programs at the local and state levels take a variety of forms. Communities need to use a combination of tools and techniques as part of a comprehensive effort to plan for appropriate land use and protect agricultural land. Examples of other common tools that communities use include comprehensive planning; agricultural zoning; limitations on infrastructure, regional tax sharing; urban growth boundaries and growth management; and agricultural viability and economic development programs. Information on conservation easements, private conservation and donation, and local land trusts can be found at the Land Trust Alliance (LTA) (www.lta.org). Information on farmland protection and stewardship can be found on the Farmland Information Center (FIC) (www.farmlandinfo.org), which has a series of fact sheets on farmland protection tools and an extensive collection of local and state land-use and land preservation laws, statistics, and literature.

Examples of statewide PACE programs include California (www.consrv.ca.gov/DLRP/index.htm), Maryland (www.malpf.info), Michigan (www.michigan.gov/mda/0,1607,7-125-1567_1599—, 00.html), Pennsylvania (www.agriculture.state.pa.us/farmland/cwp/view.asp?a=3&q=122454&agricultureNav= |4451|), and Vermont (www.vhcb.org/conservation.html). Prominent local PACE programs can be found in Boulder County, Colorado (www.co.boulder.co.us/openspace); Montgomery County, Maryland (www.mc-mncppc.org/legacy_open_space/index.shtm); and Lancaster County, Pennsylvania (http://www.co.lancaster.pa.us/planning/site/default.asp).

Land-Grant Colleges and Universities (LGUs)

Many programs related to agricultural land preservation may be found in the Land-Grant University System, including general national level programs in applied research and outreach education, specific endowed chairs in land-use and planning, and regional or state level programs addressing land-use policy and planning.

Each U.S. state and territory has at least one land-grant university (LGU) (www.csrees.usda.gov/qlinks/partners/state_partners.html), including state agricultural experiment station(s) and a state cooperative extension service. Both the agricultural experiment stations and the nationwide Cooperative Extension Service (www.csrees.usda.gov/Extension/index.html) have a long history of engagement in research, teaching, education, extension, and outreach on issues surrounding land use and policy. Information on land-use management and policy (www.csrees.usda.gov/nea/nre/in_focus/ere_if_nercrd.html) can be found on the CSREES website. Researchers at the USDA's Economic Research Service (ERS) (www.ers.usda.gov/Topics/view.asp?T=103610) also conduct economic research on land-use policy issues.

The increasing importance and complexity of land-use issues have led to the establishment of two endowed chairs at LGUs: (1) the C. William Swank Rural-Urban Policy Chair at the Ohio State University (aede.osu.edu/programs/Swank); and (2) the John Hannah Distinguished Professor in Land Policy at Michigan State University (www.landpolicy.msu.edu). Both land-use endowed chairs provide substantial leadership at the state and national levels to address these complex issues through research, education, extension, and outreach (12).

Examples of land-use programs at LGUs include the University of California at Davis (aic.ucdavis.edu/research1/land.html), Colorado State University (dare.agsci.colostate.edu/csuagecon/extension/pubstools.htm#LandUse) (13), Michigan State University (ntweb11a.ais.msu.edu/luaoe/index.asp), the Ohio State University (landuse.osu.edu/), Pennsylvania State University (www.cas.nercrd.psu.edu/Land_use.htm), Purdue University (www.ces.purdue.edu/anr/landuse/), University of Wisconsin (www.uwsp.edu/cnr/landcenter), and University of Wyoming (www.uwyo.edu/openspace).

Nongovernmental Organizations

For landowners, one of the most common ways to permanently preserve their agricultural land is to donate a conservation easement to a qualified conservation organization or public entity. If donating the entire easement value is not feasible, landowners can make a partial donation. According to the LTA (www.lta.org/aboutlt/census.shtml), more than 1260 local and regional land trusts were operating in the United States in 2000 and had collectively preserved more than 6.2 million acres of land and other natural and cultural resources. Information about local and state growth management, land-use planning, and agricultural land preservation in general can be found at the American Farmland Trust (AFT) (www.farmland.org), the American Planning Association (www.planning.org), the Smart Growth Network (www.smartgrowth.org), the Sprawl Resource Guide (www.plannersweb.com/sprawl.html), and the Lincoln Institute of Land Policy (www.lincolninst.edu/indexhigh.asp).

Increasingly, in partnership with local, state, or federal programs, many nongovernmental organizations (NGOs), especially land trusts, work to preserve agricultural lands and open space. Until recently, except in California, most NGOs accepted donated conservation easements. As the 2002 Farm Bill made qualified NGOs eligible to receive federal funds, land trusts have since increased in purchasing agricultural conservation easements. Most land trust activities take place at the local and state levels, such as the Colorado Coalition of Land Trusts (www.cclt.org/) and Sonoran Institute (www.sonoran.org/index.html). However, prominent national organizations working on land preservation include the AFT (www.farmland.org/), the LTA (www.lta.org/), and the Trust for Public Land (www.tpl.org).

CONCLUDING REMARKS

Agricultural land policy and protection is a pressing issue of public concern. The failure of land markets to reflect the social value of extensive land uses relative to intensive uses and the irreversibility of land-use intensification decisions create a particularly fluid and pressing policy environment for lay people, land-use practitioners, and researchers alike. Federal, state, and local governmental agencies and nongovernmental organizations of various scales are engaged in land-use planning as are researchers and outreach educators at many Land Grant Universities throughout the United States. In this article, we have provided a brief overview of the issues and resources and ongoing programs directly influencing agricultural land protection and land-use policy. We have taken pains to point to many of the resources currently available on the Internet, recognizing that this is an area of inquiry characterized by rapid change and development. As such, traditional print resources, although valuable, cannot be expected to keep pace with the inchoate informational environment found within the agricultural land-use planning and protection arena. Interested readers are encouraged to consult the prominent individual practitioners, program managers, and researchers in the field to benefit from the most up-to-date information available.

BIBLIOGRAPHY

1. Hellerstein, D., Nickerson, C., Cooper, J., Feather, P., Gadsby, D., Mullarkey, D., Tegene, A., and Barnard, C. (2002). *Farmland Protection: The Role of Public Preferences for Rural Amenities*. USDA Economic Research Service. Agricultural Economic Report #815.
2. Abdalla, C. (2004). What the public values about farm and ranch land: Workshop summary. Northeast Regional Center for Rural Development, University Park, PA. Available: www.cas.nercrd.psu.edu/Land_use.htm.
3. Northeast Regional Center for Rural Development (NERCRD). (2004). Available: www.cas.nercrd.psu.edu/Land_use.htm, Accessed September 4, 2004.
4. Freedgood, J. (1997). *Saving America's Farmland: What Works*. The American Farmland Trust, Northampton, MA.
5. Libby, L.W. and Abdalla, C. (2002). Protecting farmland at the fringe: Do regulations work? Swank Program in Rural-Urban Policy, Department of Agricultural, Environmental, and Development Economics, The Ohio State University, Columbus, OH. Available: aede.osu.edu/programs/Swank.
6. De Cuir, N., Sokolow, A.D., and Woled, J. (Eds.). (2003). Compensating Landowners for Conserving Agricultural Land. Agricultural Issues Center, University of California, Davis, CA. Available: www.aic.ucdavis.edu.
7. Farmland Information Center (FIC). (2004). American Farmland Trust. Available: www.farmlandinfo.org, Accessed September 4, 2004.
8. McLeod, D., Coupal, R., Seidl, A., Inman, K., and Taylor, D. (2003). Opportunities and challenges for land use research and outreach in the Intermountain West. *J. Extension* **41**(5). Available: www.joe.org/joe/2003october/a5.shtml.
9. Farm Foundation. (2004). Issue Report: Farmland preservation. Farm Foundation Issue Report Issue 1, April 2004. Oakbrook, IL. Available: www.farmfoundation.org/documents/FINALIssueReport4-15-04.pdf.
10. Seidl, A. (2001). Colorado professionals' concerns, abilities, and needs for land use planning. *J. Extension.* **39**(4). Available: www.joe.org/joe/2001august/rb5.html.
11. Loomis, J.B., Rameker, V., and Seidl, A. (2004). A hedonic model of public market transactions for open space protection. *J. Environ. Planning Manage.* **47**(1): 83–96.
12. (1999). *The Performance of State Programs for Farmland Protection: Conference Proceedings*. Swank Program in Rural-Urban Policy, Department of Agricultural, Environmental, and Development Economics, The Ohio State University, Columbus, OH.
13. Seidl, A. (2004). Failing markets and fragile institutions in land use: Colorado's experience. In: *Land Use Problems and Conflicts: Causes, Consequences and Solutions*. S.J. Goetz, J.S. Shortle, and J.C. Bergstrom (Eds.). Routledge Publishers. In press.

WATERLOGGING

Suresh K. Gupta
CSSRI
Karnal, Haryana, India

According to Clayton, land is waterlogged if the water table is within ± 150 cm of the natural surface. A modification to account for the yield of the crop was made and decided to term a land waterlogged if the yield is reduced to 25% of the optimum value due to a rise in the subsoil water. A committee constituted by the Central Water and Power Commission (India) defined waterlogging on the basis of a definition proposed by Kuntze. According to this definition, an area is said to be waterlogged when water stagnates on the land surface or the water table rises to an extent such that soil pores in the root zone of crop become saturated, resulting in restrictions in the normal circulation of air, a decline in the level of oxygen, and an increase in the level of carbon dioxide. Due to its simplicity, the definition proposed by Clayton has been adopted in many countries, and apparently it accounts for both surface and subsurface waterlogging. Thus, the word waterlogging is used to designate stress due to water stagnation or a shallow water table because similar conditions could occur as a result of any of the two causative factors. However, to designate the problem of water stagnation on the land surface, besides waterlogging, surface stagnation and flooding have also been used synonymously. Deep submergence depths caused by floods are generally excluded from waterlogging because the causative factors as well as the solutions to such a problem could be different.

VISUAL OBSERVATIONS FOR PROBLEM IDENTIFICATION

The assessment of the degree and extent of waterlogging can usually be made through simple observations, interviews with local officers of the agricultural and irrigation departments, discussions with agricultural technicians and farmers, through communication, and possibly some field measurements. Field reconnaissance is necessary to find out if the problem exists. Look for the following to arrive at a reasonable conclusion.

- Standing water or wet spots in parts of the field where the crop stand is poor. Standing water in low spots after a prolonged dry period is a useful indicator of the problem.
- Observe wells, gravel pits, and deep channels, which show the depth to groundwater. In India, the water table in open wells is commonly measured and reported. If there are few wells, then install pits or auger small observation wells into the soil to depths of 1–2 m below the soil surface. If soil horizons are reached which are gley, wet, and may contain black or red mottles, one can assume that soils are poorly drained at this level (1).
- Vegetative cover is a good index to the depth to the water table in many areas. Certain grasses/crops and trees are more common in waterlogged lands. It is a useful practice to look for these features while surveying. Trees such as willows, cottonwood, and poplar often thrive in high water table areas. Reed grass and sedges are also common.
- A general yellowing of the crop is noticed in areas that are affected by surface stagnation or have a high water table.
- Salts in the form of a white crust on the soil surface reveal the problem of waterlogging. Since the problem could occur even when the crust is not present, so its absence does not mean that there is no problem of waterlogging.
- Absence of surface drains or their condition (full of vegetation or plugged up) could also indicate a problem of waterlogging/drainage.

FIELD MEASUREMENTS TO ASSESS THE PROBLEM

In the field, waterlogging is measured/assessed through physical measurement of the depth to the water table using open wells, tube wells, or observation wells. The depth to the water table in the range of 0–1.5 m would usually indicate a problem of waterlogging, although it would depend upon the soil and crop characteristics. A relatively shallow water table may not be a problem in a coarse-textured soil, but even a relatively deeper water table could cause a problem in a fine-textured soil.

In the field, the oxygen diffusion rate (ODR) is a good measure of oxygen deficiency. ODR is measured with an oxygen diffusion rate meter. Measurements are usually made at a 10-cm depth replicated at 5–10 locations. Stolzy and Letey (2), in a review on plant response to measured ODR, concluded that the roots of many plants do not grow in soils with ODR values of $20 \times 10^{-8}\ \mathrm{g\,cm^{-2}\,min^{-1}}$. For germination and emergence of seeds, minimum ODR values of the order of $40\text{–}80 \times 10^{-8}\ \mathrm{g\,cm^{-2}\,min^{-1}}$ are mentioned. The values of ODR that indicate adequate aeration are in the range of $20\text{–}35 \times 10^{-8}\ \mathrm{g\,cm^{-2}\,min^{-1}}$ (2).

The oxidation–reduction potential (redox potential, RP) of the soil is also used as an indicator of the problem of waterlogging. In practice, the RP of the soil is measured with an oxygen meter using an Ag–AgCl reference electrode. The readings can be converted relative to the standard hydrogen electrode by adding 222 mV. As long as O_2 is present in a significant amount, the RP remains high (800–900 mV). The oxidation–reduction potential for reduction of NO_3^- to NO_2^- is 420 mV at pH 7 and 530 mV at pH 5. Nitrate is readily reduced to N_2O or N_2 at these oxidation–reduction potentials.

CAUSES OF WATERLOGGING

A large number of factors combine to create the problem or affect the degree of the problem of waterlogging. Geomorphic features like palaeochannels, local depressions, lower elements of slopes, and flat flood plains are invariably subject to varying degrees of waterlogging. Unfavorable topography of the terrain has a profound effect on the magnitude of the problem. Climatic conditions such as heavy rainstorms, a characteristic of monsoon climatic regions, coupled with poor natural drainage and/or poor upkeep of the drainage system result in surface stagnation of water. Development activities such as construction of roads, bridges, railway lines, and buildings result in choking of natural drainage thus aiding surface stagnation and a rise in the water table. Soil texture and structure having direct bearing on water storage and transmission could cause surface stagnation of water. Formation of a hard layer at plow depth in rice–wheat rotation prolongs the surface stagnation of water, especially in areas of poor on-farm water management. Prolonged heavy precipitation and a high water table resulting in reduced absorption of rainwater is a sure prescription for surface water stagnation. Increasing encroachment of wetlands for developmental activities compounds the situation.

Many factors that result in surface stagnation aid the buildup of the water table in a region. However, a rise in the water table is an inevitable consequence of introducing surface irrigation through the interbasin transfer of water. Several factors together determine the rate at which the water table would rise in a given setting. Sources of excess soil water that result in high water tables include high precipitation in humid regions, surplus irrigation water and canal seepage in irrigated lands, and artesian pressure. The geomorphologic setting of an area and the presence of impervious substrata/hard pan such as clay pan, calcic layer, duripan, or fragipan could obstruct the downward movement of water and cause short-term waterlogging. If such conditions persist for long, the land might develop a perched water table.

EFFECTS OF WATERLOGGING

Both air and water in appropriate proportions are essential for plant growth. Disturbance of the delicate balance between these two components as a result of waterlogging results in poor aeration. Except for winter conditions, an air-filled porosity of less than 10% is considered to indicate inadequate aeration. Because gases in general, diffuse about 1000 times more slowly in water than in air (3), this causes an imbalance in the proportion of oxygen and carbon dioxide in the root zone. This imbalance manifests its effect on the physiological processes of the plants and on the nutrient status of the soils. The chemistry and microbiology of submerged lands is quite different from the chemistry and microbiology of nonsubmerged lands (4). This can affect the availability of many nutrients. For example, nitrogen can undergo denitrification more readily and be lost to the atmosphere as a gas. Leaching of plant nutrients could also result in their nonavailability to plants. The anaerobic (reducing) environment results in changes that can result in deficiencies (reduced uptake) or toxicities (excess). For example, the increased availability of silicon apparently decreases the uptake of iron, whereas sulfide, ferrous, and manganese ions accumulate in waterlogged soils. The important inorganic and organic compounds that are produced in submerged soils are iron, sulfides, and organic acids. Ponnamperuma (4) reported that the level of ferrous iron in submerged soils often exceeds 3000 ppm. Such levels could be toxic to rice plants, although the degree of toxicity may depend upon the pH and the presence of hydrogen and ferrous sulfide. Soil sulfate under highly reduced conditions as might occur in submerged land is reduced to H_2S. Some of the organic acids such as acetic, propionic, and butyric are found in submerged soils; acetic acid is the most abundantly available carboxylic acid. These acids are toxic to rice seedlings.

All biological processes are strongly influenced by soil temperature. Waterlogged soils have a large heat capacity and therefore, are relatively colder than dry soils. Crop growth in waterlogged soils, therefore starts later and is slower than in dry soils.

The unconfined compressive strength of a soil is exponentially related to the moisture content. Water content is the chief factor determining the strength and also the deformation characteristics of cohesive soils. As a result, dry soils can support heavy loads because of the rigidity of the individual soil particles and the high internal friction of the soil mass. At high moisture content, the soil is susceptible to plastic deformation and consolidation. With increasing soil moisture, it loses its load-bearing capacity altogether. As a result, heavy machinery cannot be employed on waterlogged lands. A direct consequence of this would be delays in the sowing/harvesting operations of crops. To manipulate the soil without compaction or puddling, the moisture content of the soil must be below some critical level, depending upon the soil type. Compared with sand, clays have a narrow range of water content in which tillage will readily break the soil. The timeliness of the operation significantly affects the yield of crops in humid regions and in soils that have high clay content. Even the quality of the crop is affected because of the delay in harvesting.

Besides what has been stated before, excess water and the resulting continuously wet root zone can lead to some serious and fatal diseases of the root and stem (5).

It is now apparent that the adverse effect of temporary waterlogging depends not only on plant species but also on the physiological stage of growth, the time and duration of waterlogging, the light intensity, the temperature, and the fertility of the soil. Injury to crops is particularly severe when crops are flooded on hot days. The ill effects are generally greater if the waterlogging conditions are accompanied by salinity/sodicity. It is now known that no single factor affects crop yield in waterlogged soils. Generally, it is a combination of poor aeration, reduced uptake of plant nutrients, imbalance between the uptake of various ions, and the toxic effects of various elements/cations.

The ill-effects of surface stagnation of water and a high water table extend beyond the agricultural sector. The anticipated adverse effects of waterlogging are listed in Table 1.

SURFACE STAGNATION AND CROP YIELDS

Excess water in the crop root zone is injurious to plant growth. Crop yields are drastically reduced on poorly drained soils, and in cases of prolonged waterlogging,

Table 1. Anticipated Losses due to Waterlogging

Agriculture	Landscape and Infrastructure	Socioeconomic
Decline in agricultural production	Damage to infrastructure	Increased socioeconomic disparity
Restriction on crops	Landscape degradation	Increased expenditure on health related services
Decline in product quality	Decline in ecosystem health	Migration from rural to urban areas
	Damage to soil health	Increased gender disparity

plants eventually die. Crops vary in their tolerance to surface stagnation. Considerable variation in waterlogging tolerance exists between and within species (6,7). Some crops thrive if water stagnates on cropped land. Lowland rice is one such fine example. However, most dry foot crops are adversely affected, and their yields are reduced. The late Arthur Hodgson, considered a world leader on the topic, extensively and intensively studied the effect of surface stagnation of water on cotton yield. Williamson and Kriz (8) reviewed experiments on flooding and waterlogging. Their conclusion is that when plants are actively growing, the severity of injury depends on the growth stage and the time of the year at which flooding takes place. The influence of the latter is closely related to soil and air temperatures. The Indian experience, however has been that in semireclaimed alkaline lands, surface stagnation at an early growth stage is relatively more harmful than at later growth stages. It might be due to the fact that most crops are sensitive to salts at early growth stages.

Considerable work has been reported on the flooding tolerance of rice (9–11). The extent of damage or yield reduction as a result of water stagnation depends on the crop tolerance, crop growth stage at which water stagnation occurs (12–15), duration of water stagnation/flooding (16–18), soil type, air and soil temperature, and other agroclimatic conditions. Studies have also been conducted on the imbalance between the uptake of various ions to identify the causes of yield reduction (19–22) and amelioration of adverse effects through additional doses of nutrients (21,23,24).

Because crop conditions vary over time, it would be essential and useful to integrate the degree of surface stagnation and duration to assess adverse effects. Some limited attempts in this direction have been made. In spite of their limitations, these concepts could serve as starting point for more elaborate investigations.

To integrate the effect of excess water and time on crop yield, tomato yields were related to the number of days that the soil is >10 mm wetter than field capacity. It has been shown that, on average, the yield declined by 0.26% for every day that the soil is >10 mm wetter than field capacity (5,25).

Gupta and Pandey (26), following Sieben (27), integrated the effect of depth of surface stagnation on the yield of rice. According to their hypothesis, rice yield might not suffer due to occasional high depths of water stagnation. However, at some time, the yield might start decreasing. According to this hypothesis, the stress due to excess water depth, Ewd, could be expressed as follows:

$$\mathrm{Ewd} = \sum_{i=1}^{n} \sum_{j=1}^{m} (D_{ij} - d) P_{ij} F_j \qquad \text{for } D_{ij} \geq d \qquad (1)$$

in which Ewd is the depth day index, cm days; D_{ij} is the depth of stored water at the ith depth interval and jth crop stage, cm; d is the optimum depth of ponding, cm; P_{ij} are the number of days for which submergence occurs in the jth growth stage, days; m is the number of growth stages; n is the number of days in a given growth stage; and F_j is a constant depending on the stage of growth of a given variety. Using the data reported by Gupta and Pandey (28), it was observed that the crop yield decreases only to the extent of 6.4% at a depth day index of 250 (26). The yield is not reduced until the depth day index is 150. This technique is a powerful tool to work out the severity of waterlogging as well as the actual time to carry out drainage.

Gupta et al. (29), while reviewing the Indian experience in crop yield influenced by the duration of water stagnation, attempted to develop a crop production function as affected by duration of water stagnation. It was found that a piecewise linear model of the same nature reported by Maas and Hoffman (30) fit the studied data sets well. Crops have a threshold duration up to which there is no yield decline till the threshold is reached; beyond that, there is a linear decrease in the crop yield till the duration at which the crop yield is likely to be zero. Gupta et al. (31) confirmed the validity of the model while comparing the production functions for three winter crops. The duration of water stagnation at which decline in yield would be 50%, is 7.2 days for wheat, and it is 7.9 and 12.2 for mustard and barley.

GROUNDWATER DEPTH AND TOLERANCE OF CROPS

Per se, a shallow depth of groundwater is not harmful to plants. Indirectly, it determines the prevailing moisture conditions and therefore has an influence on water supply, aeration, and the heat properties of soils. As such, the yields of most crops are affected when groundwater is shallow. As with surface stagnation, crops also vary in their tolerance to a high water table (Table 2). Lowland rice can thrive on soils that have a shallow water table.

Studies on the Ill-Effects of a High Water Table

Numerous laboratory and lysimeter/field experiments on the effect of water table depth on crop yields have been conducted at various locations. Lewis and Wor (32) reported that orchards in Oregon require, ideally, a water table depth of 180–250 cm, but a higher water table persisting for 3 to 4 days following rain or irrigation does no harm to the plants. Deep-rooting fruit trees tend to suffer most when the water table is close to the surface such that it helps establish a shallow rooting system. Roe

Table 2. Tolerance Levels of Crops to High Groundwater Table (Groundwater at 50 cm)

Tolerance Level	Crops
High tolerance	Sugarcane, potatoes, rice, willow, plum, broad beans, strawberries, some grasses
Medium tolerance	Sugar beet, wheat, oats, citrus, bananas, apple, barley, peas, cotton, pears, blackberries, onion
Sensitive	Maize, tobacco, peaches, cherries, olives, peas, beans, date palm

Source: Internet Search.

(33) reviewed the work on yields of vegetables related to water table depth in the eastern United States. The general finding is that a water table from 30–60 cm from the surface is required for most vegetable crops. Penman (34) observed that citrus trees in Australia, remained in healthy condition for the first 8–10 years where a water table was within 120 cm from the surface. Adverse effects appeared thereafter.

Williamson and Kriz (8) provide a good account of the tolerance of crops to a constant water table. The data given agree fairly well with values mentioned by Williamson and van Schilfgaarde (35), Williamson and Carrekar (36), and Feddes (37). In these papers, required water table depths of 60 to 90 cm for sandy soils and 100 to 150 cm for clay soils are mentioned; the depth within these ranges depends on soil type, crop, and climatologic conditions. The relative tolerance of crops to waterlogging has also been listed according to the cause of damage (8). It is now known that the responses of agricultural crops vary genetically and physiologically to the water regime and therefore, crops that are not adapted to a high water table receive a severe setback when roots encounter a high water table. The extent of damage depends upon the crop, the crop growth stage at which high water table conditions occur, the duration of the high water table, the soil type and agroclimatic conditions.

Visser (38) described an intensive survey of prevailing groundwater table depths in The Netherlands. A general crop production function was described as (1) an increasing yield as depth increases at shallow depths, (2) a certain optimum range, and (3) a yield decline at greater groundwater depths. At shallow depths, the crop suffers from lack of aeration, but at greater depths, water deficiency is the cause of yield depression. This kind of crop production function is expected when more or less the same amounts of irrigation are given to crops.

Though accepting this logic, it is now known that the contribution of the water table to meet a part of the crop water requirement should not be overlooked (39,40). By properly scheduling irrigation, the shape of the crop production function could be altered significantly. The actual experience and field results substantiate the observation that using careful management, crop production can be increased even when the water table is shallow. For an equivalent yield of a wheat crop under a different water table, the amount of irrigation needs to be decreased as water table depth decreases. For example, Pandey and Sinha (41) observed that an equivalent yield of wheat could be obtained when the number of irrigation were varied from 0 to 4 as the water table was lowered from <75 cm to 150 cm. No hard and fast rule can decide the optimum depth to the water table for a given crop. The information given in the literature should be treated only as a guideline. To take advantage of the shallow water table contribution to the crop water requirement, controlled drainage has been perfected. It has been shown that controlled drainage helps to reduce water stress at least in low rainfall years and more N use by plants (42,43).

Under field conditions, soil–plant systems vary continuously; evaluating the effect of water table on plant growth would require integration of water table depth over time during the entire growing season.

Sieben (27), when reviewing the results of experimental drainage fields in the Dutch Ysselmeer polders, drew special attention to the influence of high fluctuating winter water tables. He took 30 cm below surface as a critical level and computed so-called SEW_{30} values from

$$\mathrm{SEW}_{30} = \sum_{i=1}^{n}(30 - x_i) \tag{2}$$

where x_i are daily water table depths below the surface. For this computation only x_i values smaller than 30 are taken into account, so that the sum is a measure of the excess over the 30-cm level. Large SEW values generally indicate poor drainage conditions. It seems that at SEW values of 100 to 200, there is a decrease in yield. The analysis also revealed that the tolerance of four crops is in the order of winter wheat > peas > oats > winter barley. The yield of wheat is not adversely affected till this index passes the 200 mark, whereas barley is tolerant to waterlogging only up to an index of 120.

Questions of the validity of this concept have been raised mainly on the basis that identical SEW values do not necessarily imply identical conditions. It is also argued that the concept does not account for the stage of growth at which shallow depths to the water table occur (44). Although, it should be possible to modify the Sieben Eq. 2 to avoid such criticism, yet the nonavailability of data on crop growth stage dependent factors would usually make such modification an academic exercise. A recent attempt to determine crop susceptibility factors for five growth stages for corn and soybean could provide answers to such a criticism (45). These authors also concluded that normalized crop susceptibility factors for wet stress might be roughly approximated for some crops from existing data on yield response to dry stress.

Comparative Assessment of Stress Due to Surface Stagnation and a High Water Table

Several experimental results exist in the literature where adverse effects on crop performance at the same site due to surface stagnation as well as a high water table have been studied (19,46). However, in many studies, it has not been possible to compare the results, as the stress levels cannot be compared. Ahmad et al. (46) however, imposed equivalent stress levels based on Equation 2, as proposed by Sieben (24). At the same stress level due to submergence and depth to the water table of 15 cm, crop vegetative growth and shoot uptake of N, P, and K were significantly less for submergence than due to shallow depth to the water table.

MANAGEMENT AND RECLAMATION OF WATERLOGGED LANDS

Waterlogging is a widespread problem under different agroclimatic conditions, so several options have been field tested to minimize the adverse effects of waterlogging. Among them are land forming (bed plantation, raised and sunken beds); crop selection in favor of tolerant crops; skipping or delaying irrigation so that crops can draw a part of their water requirement from the shallow water table; applying less water per irrigation; cultural practices such as hoeing and weeding, including chemical control of weeds; application of additional doses of nutrients through soil or foliar application, and mulching to minimize secondary salinization. However, note that these are short-term measures and might or might not be useful in each and every case. Therefore, improvement in drainage seems to be essential, as it would be the only permanent solution to the problem.

The lack of drainage or the inadequacy of it is rapidly becoming a major constraint on agricultural production. The productivity of agricultural lands can be sustained only if drainage improvements are undertaken on cropland currently affected by submergence or high water tables. Very often, the natural drainage in an area along with good water management is sufficient to eliminate excess water and to preclude the need for drainage systems. However, there are many situations where surface drainage would be required. In humid regions as well as in monsoon climatic conditions where rain storms of 100 mm or more in one day can be anticipated, even in arid and semiarid regions, surface drainage is an essential prerequisite to avoid water stagnation. Farmers who have heavy textured soils, soils with a plow layer that develops in lowland rice-wheat systems, alkaline lands with poor water absorption characteristics, or those who rely mainly on surface irrigation should have adequate surface drainage facilities to remove excess water. A uniform slope of about 1:1500 is desirable to drain irrigation water or rainfall off a field.

If recharge to the groundwater is more than the discharge and is natural drainage is unable to take care of this recharge, such a situation calls for subsurface drainage. Subsurface drainage may be accomplished either through the construction of open trenches or through buried clay or concrete tiles or perforated pipes. It must, however, be realized that most crops have an optimum depth to the water table for optimum crop yields. If the water table were lowered beyond this depth, there would not be any significant effect on the yield of the crops, and additional investment in intensive drainage would not be beneficial.

BIBLIOGRAPHY

1. Robinson, G.W. (1949). *Soils, Their Origin, Constitution and Classification*, 3rd Edn. Murby, London.
2. Stolzy, L.H. and Letey, J. (1964). Measurement of oxygen diffusion with platinum micro-electrodes. Correlation of plant response to soil oxygen diffusion rates. *Hilgardia* **35**: 567–576.
3. Grable, A.R. (1966). Soil aeration and plant growth. *Adv. Agron.* **18**: 57–106.
4. Ponnamperuma, F.N. (1972). The chemistry of submerged soils. *Adv. Agron.* **24**: 29–96.
5. Poysa, V.W., Tan, C.S., and Stone, J.A. (1987). Flooding stress and the root development of several tomato genotypes. *Hort. Science* **22**: 24–26.
6. Finn, B.J., Bourget, S.J., Nielson, K.F., and Dow, B.K. (1973). Effect of different soil moisture tensions on grass and legume species. *Canadian J. Soil Sci.* **41**: 16–23.
7. Francis, C.M. and Poole, M.L. (1973). Effect of waterlogging on the growth of annual medicago species. *Aust. J. Exp. Agric. Anim. Husb.* **13**: 711–713.
8. Williamson, R.E. and Kriz, G.J. (1970). Response of agricultural crops to flooding, depth of water table and gaseous composition. *Trans. ASAE* **13**: 216–220.
9. Joshi, M.S. and Dastane, N.G. (1965). Excess water tolerance of summer cereals. *Indian J. Agron.* **10**: 289–298.
10. Halim, A., Ching, H.F., Standifer, L.C., Yaacob, O., and Yoong, S.C. (1973). Some factors in maize production. In Crop diversification in Malaysia. *The Proc. of the Conf. held at Kualalumpur, Malaysia*.
11. Purvis, A.S. and Williamson, R.E. (1972). Effect of flooding and gaseous composition of the root environment on growth of corn. *Agron. J.* **64**: 674–678.
12. Erickson, A.E. and van Doren, D.M. (1960). The relation of plant growth and yield to soil oxygen availability. *7th, Int. Cong. Soil Sci. Trans.* (Madison, WI) **3**: 428–434.
13. Orchard, P.W. and Jessop, R.S. (1985). The response of sorghum and sunflower to short-term water logging. III. Root growth effects. *Plants and Soil* **88**: 421–340.
14. Soomaro, A.B. and Waring, S.A. (1987). Effect of temporary flooding on cotton growth and nitrogen nutrients in soils with different organic matter levels. *Aust. J. Agric. Res.* **38**: 91–99.
15. Sorte, N.V. et al. (1997). Effect of water logging on performance of soybean cultivars at various growth stages. *Journal of Soils and Crops* **7**: 171–174.
16. Howell, T.A. and Hiler, E.A. (1974). Effect of inundation period on seedling growth. *Trans. ASAE* **17**: 276–284, 294.
17. Chauhan, Y.S. (1987). Screening for Tolerance to Salinity and Waterlogging: Case Studies with Pigeon pea and Chickpea. In: *Proceedings of the Consultants Workshop on Adaptation of Chickpea and Pigeon pea to Abiotic Stresses*. N.P. Saxena and C. Johansen (Eds.). ICRISAT, A.P., India, pp. 93–103.
18. Singh, Y.V., Swarup, A., and Gupta, S.K. (2004). Effect of short-term waterlogging on growth, yield and mineral composition of sorghum. *Agrochimica, XLVI* (in press).
19. Chaudhary, T.N., Bhatnagar, V.K., and Prihar, S.S. Corn yield and nutrients uptake as affected by water table depth and soil submergence. *Agron. J.* **67**: 745–749.
20. Sharma, D.P. and Swarup, A. (1988). Effect of short-term flooding on growth, yield and mineral composition of wheat in sodic soil under field conditions. *Plant and Soil* **107**: 137–143.
21. Swarup, A. and Sharma, D.P. (1993). Influence of top dressed nitrogen in alleviating adverse effects of flooding on growth and yield of wheat in a sodic soil. *Field Crop Res.* **35**: 93–100.
22. Hoang, B., Johnson, J.W., Nesmith, D.S., and Bridges, D.C. (1995). Nutrients accumulation and distribution of wheat genotype in response to waterlogging and nutrient supply. *Plant and Soil* **172**: 47–54.

23. Paul, H. and DeLong, W.A. (1949). Phosphorus studies (I) effect of flooding on soil phosphorus. *Sci. Agric.* **29**: 137–147.
24. Hodgson, A.S. (1982). The effects of duration, timing and chemical amelioration of short-term water logging during furrow irrigation of cotton in a cracking clay. *Aust. J. Agric. Res.* **33**: 1019–1028.
25. Gales, K.R. (1976). *Effects of Waterlogging on Plant Water Relationships*. Annual Report 1975, ARC Letcombe Laboratory, pp. 40–42.
26. Gupta, S.K. and Pandey, R.N. (1983). Effect of depth and time of rainwater storage in the field on the yield of rice during the monsoon. *Ind. J. Agric. Sci.* **53**: 804–806.
27. Sieben, W.H. (1964). Effect of drainage conditions on soil N supply and yield. *Landbouwk Tijdscher* **76**: 784–802.
28. Gupta, S.K. and Pandey, R.N. (1979). Crop and water yields as affected by rainwater storage in rice fields- a field evaluation. *Field Crops Res.* **3**: 365–371.
29. Gupta, S.K., Singh, R.K., and Pandey, R.S. (1992). Surface drainage requirement of crops: application of a piecewise linear model for evaluating submergence tolerance. *Irrigation & Drainage systems* **6**: 249–261.
30. Maas, E.V. and Hoffman, G.J. (1977). Crop salt tolerance: current assessment. *J. Irrig. Drainage Div.* **103**: 115–134.
31. Gupta, S.K. et al. (2004). A comparative assessment of tolerance of three *Rabi* crops to water stagnation on soil surface. *J. Indian Soc. Soil Sci.* submitted.
32. Lewis, M.R. and Wor, A. (1931). *Orchard Drainage in the Medford Area- Jackson County, Oregon*. Oregon State Agri. Coll. Exp. Sta. Circ. 100.
33. Roe, H.B. (1936). *Influence of Depth of Groundwater Level on Yields of Crops Grown on Peat Land*. Minnesota Agri. Exp. Sta. Bull. No. 330.
34. Penman, F. (1938). Soil conditions at bamawn and ballendella in relation to citrus growth. *Jour. Dept. Agric. Victoria*: 1–36.
35. Williamson, R.E. and van Schilfgaarde, J. (1965). Studies on crop response to drainage. II. Lysimetric studies. *Trans. ASAE* **8**: 98–102.
36. Williamson, R.E. and Carrekar, J.R. (1970). Effect of water table on evapotranspiration and crop yield. *Trans. ASAE* **13**: 168–170, 176.
37. Feddes, R.A. (1971). *Water, Heat and Crop Growth*. Med. Landbouwhogeschool, Wageningen, p. 184.
38. Visser, W.C. (1958). *De Landbouwaterhuishouding in Nederland*. Comm. Onderz. Landb. Waterhuish. Ned. TNO. Rapport-nr, **1**: p. 231.
39. Timsina, J., Garrity, D.P., and Pandey, R.K. (1993). Water table gradients effect on the performance of diverse cowpea cultivars. *Agronomy J.* **85**: 359–367.
40. Ibrahim, S.M. (1999). Wheat cultivation under limited irrigation and high water table conditions. *Egyptian J. Soil Sci.* **39**: 361–372.
41. Pandey, S.L. and Sinha, A.K. (1971). Studies on water table position, soil properties, crop growth and yield under drained and undrained conditions. *Indian J. Agronomy* **16**: 494–501.
42. Evans, R.O., Cummings, J.R., and Gilliams, J.W. (1989). *Controlled Drainage: A Best Management Practice in North Carolina*. ASAE Paper 892695, St. Joseph, MI.
43. Fisher, M.J. et al. (1999). Water table management, nitrogen dynamics and yield of corn and soybean. *Soil Sci. Soc. Am. J.* **63**: 1786–1795.
44. Cannel, R.Q. et al. (1980). Effect of water logging at different stages of development on the growth and yield of winter wheat. *J. Sci. Food and Agric.* **31**: 117–132.
45. Evans, R.O., Skaggs, R.W., and Sneed, R.E. (1990). Normalized crop susceptibility factors for corn and soybean to excess water stress. *Trans. ASAE* **33**: 1153–1161.
46. Ahmad, N., Kanwar, R.S., Kaspar, T.C., and Bailey, T.B. (1992). Effect of soil surface submergence and a water table in vegetative growth and nutrient uptake of corn. *Trans. ASAE* **35**: 1173–1177.
47. Tsai, S. and Chu, T.M. (1997). Effects of Waterlogging on Growth and Yield of Maize VII. Recovery Responses of Maize Plants after Drainage. *Bulletin of National Pingtung Polytechnic Institute* **6**: 85–93.

WATER QUALITY MANAGEMENT IN AN AGRICULTURAL LANDSCAPE

Marianne McHugh
Michelle Clarke
Julia Duzant
Cranfield University
North Wyke, Devon, United Kingdom

INTRODUCTION

Water quality is important for a wide range of reasons including human health, food production, recreation, and biodiversity. Globally, agriculture exerts big demands and impacts on water supplies. In arid areas where groundwater may be the significant or sole source of water, agriculture can have profound effects on the rate and composition of groundwater recharge (1), while in tropical areas, high annual rainfall levels and the demands of tropical crops mean that water quality concerns are dominated by agrochemicals, untreated wastewater, and eroded sediment (2). In temperate zones, moderate rainfall throughout the year means there are few periods when agricultural activities can take place without some risk to water quality and, consequently, temperate water resources are also at constant risk of contamination from agriculture.

The range of agriculturally derived pollutants that may compromise water quality is listed in Table 1. Delivery of these contaminants to water may be associated with surface water runoff or with the movement of eroded soil particles. Contamination of water in agricultural landscapes also occurs directly, such as when grazing animals enter watercourses for access, drinking water, or shade. The contamination of water has widespread implications for human and animal health, for plant and animal biodiversity, and for safe and economical water industries, including water abstraction, fish farming, and water sports. Because of these implications, the protection and maintenance of water quality is legislated both nationally and internationally. In the following sections, the sources, mechanisms, and mitigation of agriculturally derived threats to water quality are described.

NITROGEN (N)

Nitrogen, as nitrate, is soluble, is retained inefficiently in soils, and is only weakly adsorbed by eroded sediment.

Table 1. Pollutants of Concern Within Agricultural Landscapes (5)

Category	Main Pollutants	Example Sources	Example Environmental Impacts
Soil	Silt, phosphorus, pesticides	Soil cultivation and drainage, overgrazing, outdoor pig systems; diary and beef herds	Silting of river bed gravels damaging fish spawning, invertebrate and plant habitats; excessive turbidity affecting plant growth and fish behavior; excessive sedimentation affecting rooted plants; economic costs to water abstractors; acts as carrier of other pollutants such as P and pesticides
Slurry, manure, dirty water	Organic matter, nutrients, pathogens, ammonia, metals	Slurry and manure storage and application to land; excretion directly to land; runoff from farm yards and buildings; irrigation	Deoxygenation and eutrophication, affecting all aquatic life; blooms of toxic blue-green algae; carrier of pathogens and pollutants
Fertilisers	Nitrate, phosphate	Inorganic fertilizers; slurry and manure	Eutrophication causing ecological degradation; increased water treatment costs
Veterinary medicines	Antibiotics, hormones, growth regulators, parasite control	Slurry and manure application to land; sheep dip application and disposal	Toxicity to aquatic species; contamination of potable supplies
Disinfectants	Halogens, organic chemicals, phosphorus	Use in diaries, pig and poultry housing	Toxicity to aquatic species; contamination of potable supplies; eutrophication causing ecological degradation

Runoff from agricultural farms is therefore a major source of N entering rivers, lakes, and coastal waters, particularly in areas of intensive livestock farming where stocking rates are high and feed is imported. Nitrogen levels in waters draining arable land are also high because of fertilizer inputs (surplus N not absorbed by the crops is discharged into groundwater, fresh water, and coastal waters), and because ploughing encourages organic N mineralization [the process by which organic N is converted to plant-available inorganic N (ammonium) by microorganisms]. Subsurface movement and loss of N also increases with sward age and drainage.

Nitrogen pollution of waters has significant detrimental effects. Apart from acidification due to nitrification of ammonium and the toxicity of various forms of N, the most significant impact is that of eutrophication, described as the most important environmental problem facing aquatic ecosystems in Europe (3). The process of eutrophication in surface waters can lead to a decrease in species diversity, oxygen depletion, increased sedimentation, and increases in toxic blue-green algae. Nitrogen also represents a hazard to human, particularly baby, health when present in drinking water supplies.

Preventing N pollution would require shifting from intensive arable and horticultural land use to extensive unfertilized grassland, which is socially and economically untenable. Therefore, the most effective ways to minimize excessive N leaching from agriculture include switching to organic farming (which would reduce N inputs), using minimum cultivation (which would reduce runoff generation), and planting a cover crop (such as rye or mustard) after harvest and before the next crop (which would reduce runoff generation and increase N uptake). Additional beneficial management practices include choosing high yielding crops to maximize uptake of N, calculating fertilizer requirements making allowance for existing soil N and organic amendments, avoiding unnecessary fertilizer applications, applying fertilizers away from watercourses, maintaining disease-free crops (as these cannot use N efficiently), and practising careful irrigation to avoid runoff. Alternatively, a curative approach of blending water with low N supplies or biological or chemical treatment to encourage denitrification may be employed (4), although this approach is more costly and less effective than preventing N pollution in the first instance.

PHOSPHORUS (P)

In the United Kingdom more than 50% of P entering surface waters is now thought to be derived from agriculture (5). There are three sources of P available in the landscape for mobilization and transfer to water bodies:

1. Soils contain different background levels of P and, therefore, different levels of P sensitivity to natural transfers in both soluble and sediment-laden forms through runoff and subsurface water flows. Transfer rates depend on the parent material, rainfall, and vegetation cover.
2. Their high specific surface area means soil particles detached during soil erosion have high adsorption capacity for nutrients such as P, which results in an enrichment of P in eroded sediments. The rates of detachment and transfer are accelerated by human impacts such as arable farming and livestock grazing.
3. Transfers can occur where P is added to the land in manures and fertilizers, although best management

practices are available to farmers to minimize these effects.

The relative proportions of P loss in sediment-bound and dissolved forms are highly site specific. Sediment-associated P constitutes the majority of P transported in surface runoff from agricultural land (5). However, dissolved P has a greater impact on water quality than sediment-bound P because it is more readily biologically available (6).

Phosphorus can negatively affect the quality of soils, groundwater, and surface water through eutrophication. In fresh waters, where P is considered the main limiting plant nutrient, elevated P loads tend to favor a few species and promote excessive algal growth. Submerged plants are lost and there are consequential impacts on invertebrate and fish species. As well as the diminished conservation status of eutrophic waterbodies, elevated P concentrations have been associated with impaired human and animal health.

To prevent these impacts, mitigation approaches address P sources or transfer mechanisms of both sediment-associated and runoff-dissolved P. Source approaches try to limit soil P content or its susceptibility to transfer, and include applying P based purely on the needs of the crop and the existing amount of residual soil P, or using chemical amendments to alter soil chemistry and hence reduce the solubility of soil P. Adding alum, for example, encourages P to precipitate with Fe (iron) and has been successfully applied to reduce runoff P concentrations (7). The alternative approach to controlling P is to limit transfer by reducing the occurrence and magnitude of transfer agents, especially runoff and sediment production. Measures to do this include increasing ground cover, contour cultivation, and minimum tillage. Uncultivated buffer strips at field edges, constructed wetlands, and ponds are also efficient at retaining P in agricultural runoff, but edge of field P transport can still occur in soils experiencing significant amounts of lateral subsurface flow (such as sandy soils). Also, while buffer strips can contain P bound on sediment, over time P storage within the buffer strip may create an active source of P flushed out in runoff during storm events.

ORGANIC MATTER

Organic matter as a pollutant concerns animal feces and urine (1) derived directly from grazing (unhoused) livestock, (2) deposited within housing or yards and stored, often with soiled bedding, as slurry or manure, or (3) combined with dairy parlor washings, silage effluent, and runoff from livestock yards and housing as dirty water.

Runoff of slurry, manure, or dirty water after application to land can occur with excess application, when heavy rain follows spreading, or where the terrain is too steep or insufficiently vegetated to ensure plant uptake of such applications. Contamination is also more likely where applications of organic material are close to drains and watercourses, where livestock have uncontrolled access to watercourses, and from farms with inadequate or poorly managed dirty water management systems.

Field or farmyard runoff containing organic matter may result in chronic organic enrichment of streams, ponds, lakes, and ditches. The organic material is broken down by bacteria causing deoxygenation of the water, because the biochemical oxygen demand of these materials can be one to two orders of magnitude higher than that of untreated domestic sewage. This loss of oxygen leads, in turn, to loss of invertebrate and fish life. In addition, organic matter from diffuse agricultural sources, especially from slurry and manure, may contain fecal bacteria, including *Escherichia, Cryptosporidium*, and *Giardia* species, which, in sufficient quantities, may be harmful or even fatal to human health.

Measures to minimize the risk of organic matter contamination of watercourses can be applied at the source, or during application or transfer stages. Manure can be treated to reduce pollution potential before land spreading, using chemical, biological, physical, or thermal processes. These processes may have additional benefits such as the recovery of nutrients including P (8). Alternatively, diet manipulation can be used to reduce the nutrient content of excreta without compromising animal performance (9). Within the farm area, runoff from animal enclosures and yards requires collection or diversion to storage or treatment areas to prevent runoff to watercourses. This is easier if clean water from roofs is diverted before coming into contact with organic material. Stored dirty water may then be applied to land, so that nutrients can be taken up by the crop. In storage systems for organic material, overflow or seepage can be prevented through careful siting of stores to avoid permeable soils and rock, effective sealing of the storage unit, regular checking and emptying to minimize overflow, and the diversion of surface water or rainwater (10).

Reducing organic matter contamination of water can be achieved by maintaining adequate storage facilities and land for application, maximizing knowledge of manure nutrient content and crop nutrient requirements, and employing best management practices, such as avoiding spreading when the soil is saturated or frozen, and minimizing application to steeply sloping land near watercourses. Applying soil erosion and runoff control measures is also effective in reducing organic pollutant transfer. The longer manure is in or on the soil before crops can use its nutrients, the greater the risk and opportunity for loss of nutrients and contaminants to surface waters. Spreading organic matter in spring or summer ensures maximum use of nutrient content and avoids the risk of spreading on frozen ground or before rainfall. Spreading should be at a rate appropriate to the crop and variable within the field according to slope, morphology, and proximity to watercourses. The pollution potential of grazing animals can be minimized by restricting access to streams. Once applied to land, transport of organic pollutants in runoff and sediment may be hindered by contour cultivation and the use of buffer strips.

SEDIMENT

The detachment and transport of soil particles from the land is known as soil erosion. Eroded soil particles enter rivers during periods of high rainfall and consequently high surface runoff. Sediment loads depend on land use, cultivation techniques, stocking densities, drainage, soils, slopes, geology, and rainfall. Heavy sediment loads can occur as a result of soil compaction (which may be associated with loss of organic matter, heavy machinery, multiple cultivations, or excessive livestock densities), soil capping, cultivation of inappropriate land (e.g., steep land), concentrated flow pathways along farm tracks and drainage systems, and the removal of natural runoff interceptors (such as hedges and uncropped land).

Sediment is considered a pollutant in its own right and as a carrier of adsorbed pollutants. Sediment itself increases water turbidity and smothers the bed of rivers and lakes when deposited in watercourses. Fine sediment particles clog river gravels, reducing water flow and aeration of the river bed and hence reducing the survival of fish eggs, invertebrates, and plants. Also, through their adsorption onto soil particles and aggregates, the transfer of sediment carries P, pesticides, pathogenic fecal bacteria, heavy metals, and other contaminants into surface waters.

Mitigation of the impacts of sedimentation may be focused at the source, although interception of sediment during transfer is also effective. Full consideration of site soil conditions is vital before any access or agricultural work is undertaken, as is avoiding unnecessary or ill-timed cultivations and access when the soil is wet and most susceptible to degradation. Positive measures to prevent soil erosion include contour cultivation and drainage, minimum or no tillage cultivation, establishing terraces or strip-cropping, maintaining vegetation cover through restricted grazing and the planting of cover crops, and maintaining and promoting soil organic matter content through mulching and incorporation of crop residues. Measures to minimize transport and delivery of eroded sediments to watercourses include employment of uncultivated buffer strips, grassed waterways, and sediment traps to filter suspended sediment from runoff. Sedimentation ponds or constructed wetlands can reduce runoff velocity and encourage sedimentation outside the watercourse system (11).

PESTICIDES AND OTHER CHEMICALS

A pesticide is defined as any substance prepared and used for the destruction of any pest. Pesticides therefore include herbicides, insecticides, fungicides, antifoulants, sheep dips, masonry and timber preservatives, and veterinary medicines. Agriculture and horticulture account for over 80% of pesticide sales, and applications in crop production are the main source of pesticides that may leach (move in water through the soil profile) to groundwater or run off to surface water in dissolved or particulate form. Pesticide characteristics, soil type, cultivation, and drainage all affect pesticide transport to water. Heavy rain after spraying increases the chance of pesticide movement in runoff. Loss may also occur through spray drift and volatilization, and research has shown that a high proportion of pesticides in surface waters may come from very small farmyard spills and drips from handling and sprayers: these often fall onto hard surfaces, from which they are readily washed into drains, ditches, and watercourses. The use and disposal of sheep dip is a particular risk to fresh waters: water is contaminated by dips located near watercourses and by inappropriate disposal of spent dip. Veterinary medicines and disinfectants used on dairy herds also pose a risk to the aquatic environment.

Pesticides are likely to have toxic effects on aquatic organisms either directly through toxicity or indirectly through changes in the environment. The impacts of pesticides range from rapid destruction of the tissues in contact with the pesticide to slow deterioration in a physiological process. Exposure to sublethal concentrations may cause genetic, physiological, or behavioral changes, including reproduction impairment, inhibition of brain activities, and growth reductions (12).

Pesticides applied to cropland are degraded by microbial and chemical reaction in the soil, temporarily immobilized by adsorption onto soil particles and organic matter, taken up by plants or pests, removed when the crop is harvested, or lost to the environment through runoff in surface water, volatilization into the atmosphere, and leaching to groundwater. For weakly to moderately adsorbed pesticides, the major carrier is runoff, and therefore controlling the loss of pesticides and persistent organic pollutants to watercourses involves reducing runoff. The most important factor affecting the amount of pesticide transported is the amount of pesticide present at the soil surface at the time of precipitation or irrigation. Reducing pesticide pollution can be promoted through the use of natural predators, crop rotations, integrating pesticide control with soil conservation, and using mechanical procedures, such as conservation tillage and vegetative strips (buffers), to reduce runoff and soil erosion and hence the loss of organic chemicals. The way in which the pesticide is applied may also affect pesticide loss. For example, tillage may incorporate the pesticide into the soil, so reducing the risk of pesticide loss by runoff.

HEAVY METALS

Heavy metals (high density elements naturally present in very low abundance in nature) of concern in agricultural landscapes include arsenic, cadmium, boron, selenium, nickel, lead, and mercury. Agricultural use of mineral fertilizers, feed concentrates, and medicines all contribute to farm inputs of heavy metals, although atmospheric deposition may also play a role, as in the deposition of lead and cadmium (13). Heavy metal transfer to water occurs through leaching and through adsorption onto eroded soil particles (preferentially onto clay particles) subsequently deposited within waterbodies (14).

When leached to surface waters, heavy metals are toxic to aquatic organisms (of particular concern is the impact of mercury on fish) and can impact negatively on drinking water quality. Several heavy metals, including mercury and lead, are known to bioaccumulate in the aquatic food chain.

Mitigating heavy metal contamination of water requires amending the source, plant uptake, and transfer of heavy metals. The choice of soil amendments (e.g., manure, compost, sludge, or mineral fertilizers) impacts on the heavy metal content of agricultural soils and hence on availability for transfer to waterbodies. Reducing manure use will reduce inputs of copper and zinc, for example, but the corresponding increase in mineral fertilizer required to maintain productivity may increase soil cadmium. Alternatively, changing farm management, for example, from arable to mixed farming, may reduce heavy metals at the source because of the overall reduction in fertilizers and manures used.

Minimizing the availability of heavy metals for transfer from the field can be achieved by increasing the ratio of grass to arable crop species or choosing particular arable crops to maximize plant uptake (e.g., metal removal from soil is much higher for sugar beet than for potatoes or wheat) (13). Alternatively, soil can be managed to minimize heavy metal transfer, through liming to raise pH levels and hence minimize leaching, or the application of soil conservation measures (such as minimum tillage and contour cultivation) to minimize heavy metal losses associated with soil erosion.

CONCLUSION

Water for food production, drinking, recreation, and ecology is an important but increasingly threatened global resource. The impacts of agricultural land use on the quality of water resources are numerous and include serious economic, health, and environmental implications from a wide range of specific pollutant threats. There are, however, a wide variety of measures available to minimize the sources, movement, and impacts of these agriculturally derived pollutants, most of which help to control pollution from more than one source. Control of water pollution within agricultural landscapes is, therefore, both necessary and achievable.

BIBLIOGRAPHY

1. Oren, O., Yechieli, Y., Böhlke, J.K., and Dody, A. (2003). Contamination of groundwater under cultivated fields in an arid environment, Central Arava Valley. *Israel. J. Hydrol.* **290**(3–4): 312–328.
2. Hernández-Romero, A.H., Tovilla-Hernández, C., Malo, E.A., and Bello-Mendoza, R. (2004). Water quality and presence of pesticides in a tropical coastal wetland in southern Mexico. *Mar. Pollut. Bull.* **48**(11–12): 1130–1141.
3. McGarrigle, M.L. (1998). Impact of eutrophication on Irish river water quality. In: *Eutrophication in Irish Waters*. J.G. Wilson (Ed.). Royal Irish Academy, Dublin, Ireland.
4. Hatch, D., Goulding, K., and Murphy, D. (2002). Nitrogen. In: *Agriculture, Hydrology and Water Quality*. P.M. Haygarth and S.C. Jarvis (Eds.). CABI, Wallingford, UK.
5. DEFRA. (2002). *Agriculture and Water: A Diffuse Pollution Review. The Government's Strategic Review of Diffuse Water Pollution from Agriculture in England*. DEFRA, London.
6. Frink, C.R. (1991). Estimating nutrient exports to estuaries. *J. Environ. Qual.* **20**: 717–724.
7. Campbell, K.L. and Edwards, D.R. (2001). P and water quality impacts. In: *Agricultural Non-point Source Pollution*. W.F. Ritter and A. Shirmohammadi (Eds.). Lewis Publishers, Boca Raton, FL.
8. Greaves, J., Hobbs, P., Chadwick, D., and Haygarth, P. (1999). Prospects for the recovery of phosphorus from animal manures: a review. *Environ. Technol.* **20**(7): 697–708.
9. Chadwick, D.R. and Chen, S. (2002). Manures. In: *Agriculture, Hydrology and Water Quality*. P.M. Haygarth and S.C. Jarvis (Eds.). CABI, Wallingford, UK.
10. Ritter, W.F. (2001). P and water quality impacts. In: *Agricultural Non-point Source Pollution*. W.F. Ritter and A. Shirmohammadi (Eds.). Lewis Publishers, Boca Raton, FL.
11. Novotny, V. (2003). *Water Quality: Diffuse Pollution and Watershed Management*, 2nd Edn. John Wiley & Sons, Hoboken, NJ.
12. Gevao, B. and Jones, K.C. (2002). Pesticides and persistent organic pollutants. In: *Agriculture, Hydrology and Water Quality*. P.M. Haygarth and S.C. Jarvis (Eds.). CABI, Wallingford, UK.
13. de Vries, W., Romkens, P.F.A.M., van Leeuwen, T., and Bronswijk, J.J.B. (2002). Heavy metals. In: *Agriculture, Hydrology and Water Quality*. P.M. Haygarth and S.C. Jarvis (Eds.). CABI, Wallingford, UK.
14. Haygarth, P.M. and Jarvis, S.J. (1999). Transfer of phosphorus from agricultural soils. *Adv. Agron.* **66**: 195–249.

CLASSIFICATION AND MAPPING OF AGRICULTURAL LAND FOR NATIONAL WATER-QUALITY ASSESSMENT

Robert J. Gilliom
Gail P. Thelin
U.S. Geological Survey

INTRODUCTION

The U.S. Geological Survey's (USGS) National Water-Quality Assessment (NAWQA) Program is designed to describe the status and trends in the quality of the nation's ground- and surface-water resources and to link the status and trends with an understanding of the natural and human factors that affect the quality of water. The program integrates information about water quality at a wide range of spatial scales, from local to national, and focuses on water-quality conditions that affect large areas of the nation or occur frequently within small areas.

The building blocks of the NAWQA Program are Study-Unit investigations, which will be conducted in 60 major hydrologic basins (Study Units) of the nation (Fig. 1). The Study-Unit investigations consist of intensive assessment activity for 3 years, followed by 6 years of less intensive monitoring, with the cycle repeated perennially (Leahy and others, 1990). Twenty Study Units will be in an intensive data-collection and analysis phase during each particular year, and the first complete cycle of intensive investigations of all 60 Study Units will be completed in 2002. The 60 NAWQA Study Units cover about one-half of the conterminous United States, encompass 60–70 percent of national water use and of the population

---

89K

served by public water supplies, and include diverse hydrologic systems that differ widely in the natural and human factors that affect water quality.

The distribution of Study Units ensures that the most important regional and national water-quality issues can be addressed by comparative studies among Study Units. NAWQA National Synthesis projects combine results of Study Unit Investigations with existing information from other programs and studies of the USGS and other agencies and researchers to produce regional and national-scale assessments for priority water-quality issues. The first water-quality issues to be focused on for National Synthesis are nutrients and pesticides.

Comparative studies among study units and national or regional aggregations of data for National Synthesis require consistent data on factors that affect the sources, behavior, and effects of contaminants and other factors that determine water-quality conditions. Natural features, such as geology, and factors related to human activities, such as land-use distribution, provide an environmental framework for assessing influences on water quality in different hydrologic systems. The emphasis of NAWQA on large-scale water-quality issues affected by human activities makes land-use characterization one of the most important aspects in study design and in the analysis of causes of water-quality conditions. Land-use characterization based on a nationally consistent classification system provides a basis for comparing the influences of human activities among Study Units. It also serves as a framework upon which to add many other types of county-based data on human activities, such as waste discharge or chemical use, that are associated with land use.

The purpose of this paper is to describe a system for classifying agricultural land for national water-quality assessment. The system focuses on classification of agricultural land within the 48 contiguous states. Alaska and Hawaii were included in basic data analysis but omitted from classification results because of their unique characteristics. Agricultural land use is one of the most important influences on water quality at national and regional scales, particularly in relation to the first NAWQA National Synthesis topics, nutrients and pesticides. Considered nationwide, agricultural land has large areal extent, a high degree of land disturbance, and high use of agricultural chemicals and water. Individual areas of agricultural land, however, can vary widely in their characteristics. There is a great diversity of agricultural activities in the nation, which follow distinct regional patterns influenced by environmental setting and economics. Different mixes of agricultural activities characteristic of particular regions can have widely varying influences on water quality because of differences in management practices and natural environmental setting.

METAL TOLERANCE IN PLANTS: THE ROLES OF THIOL-CONTAINING PEPTIDES

Weihua Zhang
Yong Cai
Florida International University
Miami, Florida

Chelation and compartmentalization are important mechanisms for metal tolerance in plants. Thiol-containing peptides, i.e., *glutathione* (GSH), *phytochelatins* (PCs), and *metallothioneins* (MTs), are metal ligands and play important roles in *metal tolerance* in plants. This article summarizes the present knowledge about the functions of these thiol-containing *peptides* in plant heavy metal tolerance, especially for Cd, As, and Cu. GSH participates in the amelioration of metal-induced oxidative stress. Metal-GSH complexes are the substrate for PC synthesis. Thus, enhanced GSH synthesis can increase metal tolerance in plants. PCs, a set of thiol-rich peptides with the general structure $(\gamma\text{-GluCys})_n\text{-Gly}$, play an essential role in constitutive Cd and tolerance, whereas PCs are not involved in adaptive Cd tolerance. In contrast, PCs are required for both constitutive and adaptive As tolerance in plants. PCs may play a role in constitutive Cu tolerance in some plants, but they are not involved in Ni or Zn tolerance. MTs may play an important role in constitutive Cu tolerance, and may be involved, to some extent, in Cd tolerance in plants.

INTRODUCTION

Heavy metals are those elements that have densities in excess of 5 g/cm^3 (1). About 40 elements fit this definition (in this paper, we consider metalloid arsenic (As) to be a heavy metal, because its properties are similar to those of other heavy metals). Heavy metals are often toxic to plants, although some heavy metals, i.e., Cu and Zn, are essential nutrients for plants and other organisms at low concentrations. Prolonged exposure to elevated concentrations can cause toxicity symptoms for both essential and nonessential heavy metals (e.g., Cd, Pb, Hg, As, etc.). The toxicity may result from deactivation of proteins via the binding of metals to sulfhydryl groups, competition for the active site between the heavy metal and an essential element resulting in deficiency effects (2,3), or production of free radicals and reactive species induced by heavy metals (3,4). Terrestrial plants are immobile organisms, and metal ion uptake mechanisms are not very selective. As a consequence, plants will take up substantial amounts of toxic heavy metals from metalliferous soils. Plants must have developed strategies that enable them to tolerate the presence of toxic heavy metals. Possible mechanisms for metal tolerance in plant cells are: (a) metal binding to cell wall; (b) reduced transport across the cell membrane; (c) active efflux; (d) compartmentalization; and (e) chelation (5). Chelation and subsequent compartmentalization play important roles in metal tolerance in plants. A variety of compounds, such as organic acids, amino acids, glutathione (GSH), phytochelatins (PCs), and metallothioneins (MTs), have been identified as the chelators of heavy metals. Therefore, these ligands may be involved in metal tolerance in plants.

Metal tolerance mechanisms in plants are categorized as constitutive tolerance mechanisms that exist in normal plants and adaptive tolerance mechanisms that are present in metal-tolerant plants including hyperaccumulators (6). Constitutive metal tolerance mechanisms are complex and are controlled by multiple genes. Adaptive metal tolerance mechanisms, however, seem to involve relatively simple control mechanisms (7). Thiol-containing peptides glutathione (GSH), phytochelatins (PCs), and metallothioneins (MTs) are thought to play important roles in metal tolerance in plants (3,8–10). This article summarizes the present knowledge about the functions of these thiol-containing peptides in plant heavy metal tolerance, especially for Cd, As, and Cu.

GLUTATHIONE

Glutathione (GSH), the tripeptide γ-gultamylcysteinylglycine, is the major source of nonprotein thiols in most plant cells. GSH is synthesized from γ-glutamic acid, cystein (Cys), and glycine (Gly) by a two-step ATP-dependent reaction. The first reaction synthesizes γ-glutamyl cystein (γ-EC) from γ-glutamic acid and Cys by γ-glutamylcysteine synthetase (γ-EC synthase); the second step forms GSH from γ-EC and Gly catalyzing by GSH synthase. GSH plays an important role in protecting plants from oxidative stress induced by xenobiotics and some heavy metals, thus increasing metal resistance in plants (11). A substantial number of reports about oxidative stress induced by a variety of metals exist, e.g., Cd (12–14), Cu (4,15,16), Ni (17), Hg (18), and As (19). GSH is involved in the amelioration of metal-induced oxidative stress in plants. For example, a wild type of *Arabidopsis thaliana* exposed to Cu showed changes in GSH and glutathione disulfide (GSSG) levels, suggesting their participation in the amelioration of metal-induced oxidative stress (15). A decrease in GSH synthesis leads the hypersensitivity of plants to metals, which has been shown in *Arabidopsis* exposed to Cd (12), *Holcus lanatus* exposed to As (19), and *Silene cucubalus* exposed to Cu (4).

As metal-GSH complexes are the substrate of phytochelatin synthase (PCS) (9), synthesis of PCs may lead the depletion in the GSH pool. In the presence of metals, GSH-metal complex may limit the rate of PC synthesis (20). Therefore, an increase in GSH levels facilitates the increase in PC synthesis and the sequestration of metals (13). It has been demonstrated that increases in GSH synthesis can improve metal tolerance in plants. For example, engineered Indian mustard, which overexpresses GSH, shows an increase in Cd tolerance and accumulation (20,21); elevated GSH concentrations are also involved in conferring tolerance to Ni-induced oxidative stress in *Thlaspi* Ni hyperaccumulators (17).

Vacuolar compartmentalization is an important mechanism for heavy metal tolerance in bacteria and plants. Some metals, i.e., Cd and As, need PCs to mediate the vacuolar compartmentalization in bacteria and plants (22–25). Recent studies indicate that GSH can also mediate vacuolar compartmentalization of metals, e.g., Cd, As, and Hg, in bacteria (26–28). In yeast *Saccharomyces cerevisiae*, which lacks PCS, a Cd-GSH complex can be directly transported to vacuoles by YCF1, one of ATP-binding cassette (ABC) transporter proteins located in tonoplasm (26,29–31). YCF1 was also found to transport As-GSH and Hg-GSH complexes to vacuoles in bacteria (27,28). Recently, BPT1, another ABC transporter in vacuoles, was found in yeast. Similar to YCF1, BPT1 also transports GSH conjugates including the Cd-GSH complex (32). However, this mechanism of metal tolerance involving active sequestration of the metal-GSH complexes into vacuoles by ABC transporters has not yet been reported in plants.

PHYTOCHELATINS

Structures

PCs are a set of thiol-rich peptides with the general structure (γ-GluCys)$_n$-Gly ($n = 2$ to 11), which are synthesized from GSH catalyzed by PCS (33–36). They were first identified in the yeast *Schizosaccharomyces pombe* and termed Cadystins (37,38), and isolated and characterized by Grill et al. (39). PCs have been found in all plants studied to date and some microorganisms (40). They are also found in animals, i.e., nematode *Caenorhabditis elegans* (41). Interestingly, some iso-phytochelatins (iso-PCs) are identified in some plants under heavy metal exposure (1). Grill et al. found that a few members of the *Fabales* produced homoPCs under Cd exposure (42). HomoPCs are synthesized from homoglutathione, in which β-Ala replaces Gly as a terminal amino acid (42). Klapheck et al. found that hydroxymethyl-PCs [(γ-GluCys)$_n$-Ser] were induced by Cd in several species of the family Poaceae (43). They are thought to be formed from hydroxymethyl-glutathione [(γ-GluCys)$_n$-Ser]. Meuwly et al. discovered γ-Glu-Cys-Glu, another homologue of GSH. It is not surprising that another iso-PC, (γ-GluCys)$_n$-Glu ($n = 2 - 3$), was induced by Cd in maize seedlings (44). Recently, iso-PCs (γ-GluCys)$_n$-Gln ($n = 3 - 4$) were isolated from the hairy roots of horseradish exposed to Cd (45). In addition, some PC-related compounds, i.e., desglycine PCs [(γ-GluCys)$_n$] (46,47), Cys-(γ-GluCys)$_n$-Gly ($n = 1$–5), and Cys-(γ-GluCys)$_n$ (47), were identified in maize.

Biosynthesis

PCs are not primary gene products. They are synthesized from GSH by a specific γ-glutamylcysteine dipeptidyl transpeptidase (EC 2.3.2.15)—PCS in the presence of metal ions (48). This enzyme transfers γ-Glu-Cys from GSH or PC to an acceptor molecule (i.e., GSH and PC) (49). PC synthesis reaction can be rapidly induced *in vivo* by exposure to a variety of metal cations (Cd^{2+}, Pb^{2+}, Zn^{2+}, Sb^{3+}, Ag^{+}, Hg^{2+}, Cu^{+}, Sn^{2+}, Au^{3+}, Bi^{3+}) and metalloid anions [arsenate (As^{V}) and arsenite(As^{III})] (1,3,10,37,50,51). The reaction is self-regulated. Once the free metal ions are chelated by the reaction products PCs, the reaction terminates (48,49,52). PCS activity was identified for the first time in cell-free extracts from suspension cultures of *Silene cucubalis* by Grill et al. (48), and similar PCS activities have been detected in pea (52), tomato (49), and *Arabidopsis* (53). In 1999, genes encoding PCS were first cloned and characterized in *Arabidopsis* (*AtPCS1*), *S. pombe* (*SpPCS*), and wheat (*TaPCS1*) (54–56). Homologous genes encoding PCS have been identified in other species, i.e., *C. elegans* (*CePCS*) (57), *Arabidopsis* (*AtPCS2*) (55,58), soybean Glycine max (*GmhPCS1*) (59), and *Brassica juncea* (Indian mustard) (*BjPCS1*) (60).

PCS is expressed constitutively, and the levels of this enzyme in cell cultures or intact plants are generally unaffected by exposure to metal ions (61). Analysis of the expression of *AtPCS1* indicated that levels of mRNA were not affected by exposure of plants to Cd (54,55). However, RT-PCR analysis of *TaPCS1* expression in wheat roots indicated induction of mRNA under Cd exposure (56). In addition, a recent study showed *AtPCS1* of *Arabidopsis thaliana* was regulated at transcriptional level during the early stages of plant development (62). Therefore, in some species or specific stages of plant development, PCS activity may be regulated both at transcriptional and posttranslational levels.

Earlier studies have shown that PCS activity requires the binding of free metal ions to the enzyme (48,49,52,63,64). However, a recent study has shown that metal binding per se is not responsible for the activation of PCS, but glutathion-like peptides containing blocked thiol groups are required for the enzyme activity (9). A similar conclusion was reached by Oven et al. (59). Therefore, the real substrates of PCS are metal-thiolate complexes.

Roles of PCs in Metal Tolerance

PCs are suggested to play a role in metal detoxification and homeostasis (1,37). However, no direct evidence that PCs are involved in homeostasis currently exists (51). It has been demonstrated that PCs play an essential role in constitutive Cd and As tolerance in plants. In addition, PCs may be required for adaptive enhanced As tolerance. Some indications exist that PCs may play a role in Ag, Cu, and Hg tolerance (55,56). However, the role of PCs in Cu tolerance in plants has been debated (54–56,65). A paucity of information exists about the roles of PCs for Ag and Hg tolerance in plants. Zn and Pb are weak activators for PC synthesis (48,50), and genetic and biochemical evidence has shown that PCs are not involved in Zn tolerance in plants (53–55,65–68). No evidence exists that PCs are involved in Pb tolerance in plants. Ni cannot induce PC synthesis (48,50), thus PCs are not involved in Ni tolerance in plants.

Cadmium. Cd is a strong activator for PC synthesis, and PCs and Cd can form detectable complexes *in vivo* (48,50). PCs play an essential role in normal constitutive Cd tolerance in plants (3,10,37,51,64,69,70). This conclusion is supported by a variety of biochemical and genetic evidence. *In vitro*, PCs have been shown to substantially decrease the toxicity of Cd to metal-sensitive plant enzymes, such as alcohol dehydrogenase, glyceraldehyde 3-phosphate dehydrogenase, nitrate reductase, ribulose 1,5-diphosphate carboxylase, and urase (71). In normal *S. vulgaris*, treatment with buthionine sulfoximine (BSO), which inhibits the synthesis of the PC precursor γ-EC, resulted in hypersensitivity to Cd (72). The most convincing evidence comes from genetic studies. Cd-sensitive *cad1* mutants of *Arabidopsis* are deficient in PCS activity and do not form Cd-PC complexes. Consequently, these mutants are sensitive to Cd (67,68). Expression of cDNA clones of *AtPCS1* from *Arabidopsis* and a similar gene *TaPCS1* from wheat in *S. cerevisiae* resulted in a dramatic increase in Cd tolerance (55,56). The mutants of *Arabidopsis* and *S. pombe* that lack PCS genes exhibited hypersensitivity to Cd and were unable to synthesize PCs (54,56). Azuki bean cells lacking PCS activity were also hypersensitive to Cd (73). All of this evidence strongly suggests that PCs play an essential role in constitutive Cd tolerance.

In addition to formation of Cd-PC complexes in cytoplasm, sequestration of Cd-PC complexes to vacuoles is also required for Cd tolerance (22,74). In vacuoles, the complexes are stabilized by incorporation of acid-labile sulfide (75). In yeast *S. pombe*, vacuolar sequestration process is mediated by HMT1 (22). HMT1, a member of ABC-type transport proteins located in tonoplasm, can transport PCs and Cd-PC complexes to vacuoles (76). Vacuolar compartmentalization of Cd-PC complexes has also been observed in plants, such as tobacco (75,77), and oat roots (23). However, plant genes encoding Cd-complex transporter have not been characterized yet (40).

Although PCs are essential for normal constitutive Cd tolerance, PCs are not involved in adaptive enhanced Cd tolerance in Cd-tolerant plants. In studies that compared Cd-tolerant plants with normal *Silene vulgaris*, De Knecht et al. found that Cd-tolerant plants exhibited a much lower PC synthesis and a lower rate of longer PC synthesis under Cd exposure than normal plants, and their Cd tolerance was not reduced by BSO treatment (72,78). These observations suggest that the enhanced Cd tolerance in Cd-tolerant *S. vulgaris* is not because of differential production of PCs and synthesis of longer chain PCs, although PCs are essential for constitutive Cd tolerance in normal *S. vulgaris* (72,78). Similar results were obtained in other studies of *S. vulgaris* (65,79). In Zn/Cd hyperaccumulator *Thlaspi caerulescens*, PC synthesis under Cd exposure was lower than in nonhyperaccumulators *Thlaspi arvense* (80), and BSO treatment did not increase sensitivity to Cd (65). These results suggest that PCs are not involved in adaptive enhanced Cd tolerance in this plant. The possible mechanism for the adaptive enhanced Cd tolerance in these Cd-tolerant plants is PC-independent vacuolar sequestration (65). However, this process remains to be elucidated.

Arsenic. Similar to Cd, As^{III} can induce high levels of PC synthesis in plants (50,81). However, it is difficult to detect As-PC complexes *in vivo*, although formation of As^{III}-tris-thiolate complexes is predicted on a theoretical basis (50). With improved experimental conditions, Schmoger et al. demonstrated the formation of As^{III}-PC complexes in plants by size exclusion chromatography and electrospray ionization-mass spectrometry (24). Constitutive As tolerance in plants requires the involvement of PCs. This process may include reduction of As^{V} to As^{III}, formation of As^{III}-PC complexes, and sequestration of the complexes into vacuoles (65,81,82). Results from various studies support the role of PCs in constitutive As detoxification in normal plants. For example, BSO-treated cell cultures of tobacco (83,84) and *Rauvolfia serpentina* (83,84), *Holcus lanatus* (6,82,85), and *Cytisus Striatus* (6,82,85) were found to be hypersensitive to As. The strongest evidence for the role of PCs in As detoxification comes from genetic studies of *Arabidopsis* and *S. pombe* (54,55). PCS-deficient mutants of *Arabidopsis* and *S. pombe* are highly sensitive to Cd and As (54). Conversely, the expression of *AtPCS1* from *Arabidopsis* in yeast *S. cerevisiae* resulted in enhanced resistance to As (55). This evidence strongly supports the argument that PCs are essential for constitutive As tolerance in plants.

Unlike enhanced Cd tolerance in Cd-tolerant plants, enhanced As tolerance in As-tolerant nonhyperaccumulator plants may require PCs. In As-tolerant *H. lanatus* and *C. Striatus*, suppression of phosphate/As^{V} uptake via an altered high-affinity phosphate/As^{V} uptake system is the adapted As tolerance mechanism that is mainly responsible for the enhanced As tolerance (6,19,82,86). However, PCs are also involved in the enhanced As tolerance in these As-tolerant plants via synthesis of higher levels and/or longer chain of PCs, because BSO treatment can cause these As-tolerant plants to be more sensitive to As (6,82,85). These observations suggest that constitutive production of PCs may be involved in the adaptive As tolerance in these plants. In As hyperaccumulator *Pteris vittata*, the major mechanism responsible for As hypertolerance is not related to restriction of As uptake, but may be related to enhanced vacuolar compartmentation of As (87). PCs were suggested to play no role in As hypertolerance in *P. vittata* (88). However, considering that PCs are required for enhanced As tolerance in the As-tolerant nonhyperaccumulators, PCs may play, to some extent, a role in As hypertolerance in this hyperaccumulator, which is supplementary to the adaptive As tolerance mechanisms (87).

Copper. Cu is an essential micronutrient for plants. The mechanisms for plant Cu tolerance have not been elucidated. Cu ion can induce PC synthesis, and formation of Cu^{+}-PC complexes are observed *in vitro* and *in vivo* (89,90). The role of PC in Cu tolerance in plants has been debated. Some indications exist that PCs may, to some extent, be involved in constitutive Cu tolerance. For instance, *S. pombe* cells, for which the PCS gene expression has been disrupted, exhibit hypersensitivity to Cd and Cu (56); PCS-deficient *Arabidopsis* mutant *cad1-3* is slightly sensitive to Cu, Hg, and Ag (54,91). However, the evidence is not conclusive. Another genetic study showed that PCS-deficient *S. pombe* was insensitive to Cu, Zn, Hg, Se, or Ni ions (54). *Arabidopsis* exposure to concentrations of Cu ion did not produce Cu-PC complexes, but a change in glutathione and glutathione disulfide contents was observed (15). In normal *S. vulgaris*, BSO treatment had no effect on the plant sensitivity to Cu, indicating that PCs are not essential for constitutive Cu tolerance (65).

PCs are not involved in adaptive enhanced Cu tolerance in Cu-tolerant plants. Differential Cu tolerance in Cu-tolerant and normal *S. vulgaris* is not because of differential PC production (92). The sensitivity to Cu is unaffected by BSO treatment both in Cu-tolerant and normal *S. vulgaris*, suggesting that PCs are not involved in adaptive Cu tolerance (65). A similar conclusion was obtained in Cu-tolerant *S. cucubalus* (4).

METALLOTHIONEINS (MTs)

MTs are a class of Cys-rich polypeptides (Mr: 6000–7000) that can bind metals via thiol groups of their Cys residues (93). MTs were first reported as Cd binding protein in horse kidney (94). MTs have been found in animals, plants, eukaryotic microorganisms, and some prokaryotes (40). MTs contain characteristic motifs: Cys-Cys, Cys-X-Cys, and Cys-X-X-Cys (5). MTs are divided into two classes (95). Mammalian MTs containing 20 highly conserved Cys residues are categorized as Class I (96), whereas MTs from plants, fungi, and invertebrate animals are grouped in Class II (97). According to the arrangement of Cys residues, plant MTs are further classified into four types (40,95). MTs are gene-encoded polypeptides, and their syntheses are transcriptionally activated by several factors such as hormones, cytotoxic agents, and metals, including Cd, Zn, Hg, Cu, Au, Ag, Co, Ni, and Bi metal ions (5). A wide variety of regulatory factors and expression in many organisms suggest that MTs have important functions (40). Although MTs have been studied for more than 40 years, their function has not been completely understood in animals (3,98). They may play a role in protection against metal toxicity, function as antioxidants, and are thought to be involved in plasma membrane repair (3). The only confirmed function of MTs in animals is detoxification of Cd and Zn (96,98).

MTs have been less studied in plants than in animals. The functions of MTs in plants also remain to be elucidated. Some indications exist that MTs may play a role in Cu and probably Cd tolerance in plants. Cu tolerance in *Arabidopsis* depends on the degree of Cu-induced expression of type 2 MT genes (99,100). When MT1a and MT2a genes from *Arabidopsis* were expressed in a yeast strain lacking one of its endogenous MT genes (*cup1*), the MT genes conferred the yeast a high level of tolerance to Cu and a moderate level of tolerance to Cd (101,102). Transgenic tobacco plant with expression of yeast MT gene can accumulate significantly higher quantities of Cu than control plants (103). When an MT2b-like gene (*SvMT2b*) from a highly Cu-tolerant *S. vulgaris* was expressed in yeast, this gene restored cadmium and copper tolerance in different hypersensitive strains (104). When an MT2-like gene (*tyMT*) from *Typha latifolia* was expressed in Cu-sensitive MT-deficient strain of yeast, it restored the yeast's metal tolerance (105). Transgenic *Arabidopsis*, with expression of the *tyMT* gene, showed an increased tolerance to heavy metals Cu and Cd (105). Results from these studies suggest that MTs play an essential role in constitutive Cu tolerance in these plants.

However, other studies do not support this conclusion. For instance, In Indian mustard, Cu treatment caused a strong increase in γ-ECS mRNA, GSH, and PCs in roots and shoots, whereas MT2 mRNA was decreased in response to the $CuSO_4$ treatments (106,107); In *Vicia faba*, MT1 and MT2 transcript levels are not significantly affected by treatment with Cu, Zn, or Cd (108,109). The observed decrease or no change in MT expression under Cu exposure may be caused by increased levels of PC synthesis, suggesting that Cu quenched by PCs may be favored over sequestration by MTs in Indian mustard and *Vicia faba* (106,109). As mentioned above, genetic evidence has shown that PCs may play a role in constitutive Cu tolerance in some plants (55,56), and constitutive metal tolerance mechanisms are complicated and controlled by multiple genes (7). Therefore, both PCs and MTs may be involved in constitutive Cu tolerance in plants. For some reasons, the PC-dependent mechanism appears to dominate in some plants, whereas the MT-dependent mechanism is favored in others

MTs may also be involved in adaptive enhanced Cu tolerance in plants. Cu-tolerant *S. vulgaris* was found to have significantly higher transcript levels of type 2b MT gene than normal *S. vulgaris* (104). The increase in the MT gene expression suggests that constitutive production of MTs also contribute to the enhanced Cu tolerance, although an adaptive mechanism is mainly responsible for the enhanced Cu tolerance in *S. vulgaris* (104). Restriction of Cu uptake may be the adaptive mechanism for the enhanced Cu tolerance in plants. By comparison of normal to Cu-tolerant *Silene cucubalus*, De Vos et al. suggested that the adaptive Cu tolerance in this plant does not depend on the production of PCs but is related to the plant's ability to prevent GSH depletion resulting from Cu-induced PC synthesis, e.g., by restriction Cu uptake (4). A recent study has shown that efflux of Cu across the root plasma membrane plays a role in the Cu tolerance in Cu-tolerant *S. vulgaris* (110).

In animals, MTs play an essential role in Cd detoxification. In plants, this role is mainly occupied by PCs (40). However, MT expression can enhance Cd tolerance in plants. For instance, expression of mouse MT1 gene and *Nicotiana glutinosa* MT gene confers Cd resistance in transgenic tobacco plants (111,112), suggesting that MT may play a role, to some extent, for Cd tolerance in plants. An increase in Cd tolerance was also found in transgenic *Arabidopsis* with expression of an MT gene (105). The roles of MTs in Cd tolerance in plants need further investigation and confirmation.

CONCLUSIONS

GSH, PCs, and MTs play important roles in metal tolerance in plants. GSH participates in the amelioration of metal-induced oxidative stress, and metal-GSH complexes are the substrate for PC synthesis. Thus, enhanced GSH synthesis can increase metal tolerance in plants. PCs play an essential role in constitutive Cd tolerance, whereas PCs are not involved in adaptive Cd tolerance. In contrast, PCs are required for both constitutive and adaptive As tolerance in plants. PCs may play a role in constitutive Cu tolerance in some plants, but they are not involved in Ni or Zn tolerance. MTs may play an important role in constitutive Cu tolerance, and may be involved, to some extent, in Cd tolerance in plants.

BIBLIOGRAPHY

1. Zenk, M.H. (1996). *Gene.* **179**: 21–30.
2. Van Assche, F. and Clijsters, H. (1990). *Plant, Cell and Environment* **13**: 195–206.
3. Hall, J.L. (2002). *Journal of Experimental Botany* **53**: 1–11.
4. De Vos, C.H.R., Vonk, M.J., Vooijs, R., and Schat, H. (1992). *Plant Physiology* **98**: 853–858.
5. Mejare, M. and Bulow, L. (2001). *Trends in Biotechnology* **19**: 67–73.
6. Schat, H., Llugany, M., Vooijs, R., Hartley-Whitaker, J., and Bleeker, P.M. (2002). *J. Exp. Bot.* **53**: 2381–2392.
7. Schat, H., Llugany, M., and Bernhard, R. (2000). *Phytoremediation of Contaminated Soil and Water*. N. Terry and G. Banuelos (Eds.). CRC Press, Boca Raton, FL, pp. 171–188.
8. Noctor, G. et al. (1998). *Journal of Experimental Botany* **49**: 623–647.
9. Vatamaniuk, O.K., Mari, S., Lu, Y-P., and Rea, P.A. (2000). *Journal of Biological Chemistry* **275**: 31451–31459.
10. Cobbett, C.S. and Goldsbrough, P.B. (2000). *Phytoremediation of Toxic Metals* 247–269.
11. Xiang, C. and Oliver, D.J. (1998). *Plant Cell* **10**: 1539–1550.
12. Xiang, C., Werner, B.L., Christensen, E.L.M., and Oliver, D.J. (2001). *Plant Physiology* **126**: 564–574.
13. Siripornadulsil, S., Traina, S., Verma, D.P.S., and Sayre, R.T. (2002). *Plant Cell* **14**: 2837–2847.
14. Iannelli, M.A., Pietrini, F., Fiore, L., Petrilli, L., and Massacci, A. (2002). *Plant Physiology and Biochemistry* **40**: 977–982.
15. Wojcik, M. and Tukiendorf, A. (2003). *Biologia Plantarum* **46**: 79–84.
16. Boominathan, R. and Doran, P.M. (2003). *Biotechnol Bioeng.* **83**: 158–167.
17. Freeman, J.L. et al. (2004). *Plant Cell* **16**: 2176–2191.
18. Ali, M.B. et al. (2000). *Bulletin of Environmental Contamination and Toxicology* **65**: 573–582.
19. Hartley-Whitaker, J., Ainsworth, G., and Meharg, A.A. (2001). *Plant, Cell and Environment* **24**: 713–722.
20. Zhu, Y.L., Pilon-Smits, E.A.H., Jouanin, L., and Terry, N. (1999). *Plant Physiology* **119**: 73–79.
21. Zhu, Y.L. et al. (1999). *Plant Physiology* **121**: 1169–1177.
22. Ortiz, D.F. et al. (1992). *EMBO Journal* **11**: 3491–3499.
23. Salt, D.E. and Rauser, W.E. (1995). *Plant Physiology* **107**: 1293–1301.

24. Schmoger, M.E., Oven, M., and Grill, E. (2000). *Plant Physiology* **122**: 793–801.
25. Hartley-Whitaker, J. et al. (2001). *Plant Physiology* **126**: 299–306.
26. Li, Z.S. et al. (1997). *Proc. Natl. Acad. Sci. USA* **94**: 42–47.
27. Rosen, B.P. (1999). *Trends Microbiol.* **7**: 207–212.
28. Gueldry, O. et al. (2003). *Eur. J. Biochem.* **270**: 2486–2496.
29. Falcon-Perez, J.M., Mazon, M.J., Molano, J., and Eraso, P. (1999). *J. Biol. Chem.* **274**: 23584–23590.
30. Tommasini, R. et al. (1998). *Plant J.* **13**: 773–780.
31. Gomes, D.S., Fragoso, L.C., Riger, C.J., Panek, A.D., and Eleutherio, E.C. (2002). *Biochim Biophys Acta.* **1573**: 21–25.
32. Klein, M. et al. (2002). *FEBS Lett.* **520**: 63–67.
33. Grill, E., Winnacker, E.L., and Zenk, M.H. (1987). *Proc. Natl. Acad. Sci. USA* **84**: 439–443.
34. Rauser, W.E. (1990). *Annual Review of Biochemistry* **59**: 61–86.
35. Steffens, J.C. (1990). *Annual Review of Plant Physiology and Plant Molecular Biology* **41**: 553–575.
36. Ruegsegger, A., Schmutz, D., and Brunold, C. (1990). *Plant Physiology* **93**: 1579–1584.
37. Rauser, W.E. (1995). *Plant Physiology* **109**: 1141–1149.
38. Cobbett, C.S. (1999). *Trends Plant Sci.* **4**: 335–337.
39. Grill, E., Winnacker, E.L., and Zenk, M.H. (1985). *Science* **230**: 674–676.
40. Cobbett, C. and Goldsbrough, P. (2002). *Annual Review of Plant Biology* **53**: 159–182.
41. Vatamaniuk, O.K., Bucher, E.A., Ward, J.T., and Rea, P.A. (2002). *Trends in Biotechnology* **20**: 61–64.
42. Grill, E., Gekeler, W., Winnacker, E.L., and Zenk, M.H. (1986). *FEBS Lett.* **205**: 47–50.
43. Klapheck, S., Fliegner, W., and Zimmer, I. (1994). *Plant Physiology* **104**: 1325–1332.
44. Meuwly, P., Thibault, P., and Rauser, W.E. (1993). *FEBS Lett.* **336**: 472–476.
45. Kubota, H., Sato, K., Yamada, T., and Maitani, T. (2000). *Phytochemistry* **53**: 239–245.
46. Bernhard, W.R. and Kagi, J.H. (1987). *Experientia* **52**: 309–315.
47. Chassaigne, H., Vacchina, V., Kutchan, T.M., and Zenk, M.H. (2001). *Phytochemistry* **56**: 657–668.
48. Grill, E., Loeffler, S., Winnacker, E.L., and Zenk, M.H. (1989). *Proc. Natl. Acad. Sci. USA* **86**: 6838–6842.
49. Chen, J., Zhou, J., and Goldsbrough, P.B. (1997). *Physiologia Plantarum* **101**: 165–172.
50. Maitani, T., Kubota, H., Sato, K., and Yamada, T. (1996). *Plant Physiology* **110**: 1145–1150.
51. Cobbett, C.S. (2000). *Plant Physiology* **123**: 825–832.
52. Klapheck, S., Schlunz, S., and Bergmann, L. (1995). *Plant Physiology* **107**: 515–521.
53. Howden, R., Andersen, C.R., Goldsbrough, P.B., and Cobbett, C.S. (1995). *Plant Physiology* **107**: 1067–1073.
54. Ha, S.B. et al. (1999). *Plant Cell* **11**: 1153–1164.
55. Vatamaniuk, O.K., Mari, S., Lu, Y-P., and Rea, P.A. (1999). *Proc. Natl. Acad. Sci. USA* **96**: 7110–7115.
56. Clemens, S., Kim, E.J., Neumann, D., and Schroeder, J.I. (1999). *EMBO Journal* **18**: 3325–3333.
57. Clemens, S., Schroeder, J.I., and Degenkolb, T. (2001). *European Journal of Biochemistry* **268**: 3640–3643.
58. Cazale, A-C. and Clemens, S. (2001). *FEBS Letters* **507**: 215–219.
59. Oven, M., Page, J.E., Zenk, M.H., and Kutchan, T.M. (2002). *Journal of Biological Chemistry* **277**: 4747–4754.
60. Heiss, S., Wachter, A., Bogs, J., Cobbett, C., and Rausch, T. (2003). *J. Exp. Bot.* **54**: 1833–1839.
61. Cobbett, C.S. (2000). *Current Opinion in Plant Biology* **3**: 211–216.
62. Lee, S. and Korban, S.S. (2002). *Planta.* **215**: 689–693.
63. Loeffler, S., Hochberger, A., Grill, E., Winnacker, E.L., and Zenk, M.H. (1989). *FEBS Letters* **258**: 42–46.
64. Di Toppi, L.S., and Gabbrielli, R. (1999). *Environmental and Experimental Botany* **41**: 105–130.
65. Schat, H., Llugany, M., Vooijs, R., Hartley-Whitaker, J., and Bleeker, P.M. (2002). *Journal of Experimental Botany* **53**: 2381–2392.
66. Harmens, H., Den Hartog, P.R., Ten Bookum, W.M., and Verkleij, J.A.C. (1993). *Plant Physiology* **103**: 1305–1309.
67. Howden, R. and Cobbett, C.S. (1992). *Plant Physiol.* **100**: 100–107.
68. Howden, R., Goldsbrough, P.B., Andersen, C.R., and Cobbett, C.S. (1995). *Plant Physiology* **107**: 1059–1066.
69. Goldsbrough, P. (2000). *Phytoremediation of Contaminated Soil and Water*. N. Terry and G. Banuelos (Eds.). CRC Press, Boca Raton, FL, pp. 221–233.
70. Clemens, S., Simm, C., and Maier, T. (2003). *Biopolymers* **8**: 255–288.
71. Kneer, R. and Zenk, M.H. (1992). *Phytochemistry* **31**: 2663–2667.
72. De Knecht, J.A., Koevoets, P., Verkleij, J.A.C., and Ernst, W.H.O. (1992). *New Phytologist.* **122**: 681–688.
73. Inouhe, M., Ito, R., Ito, S., Sasada, N., Tohoyama, H., and Joho, M. (2000). *Plant Physiology* **123**: 1029–1036.
74. Speiser, D.M., Abrahamson, S.L., Banuelos, G., and Ow, D.W. (1992). *Plant Physiology* **99**: 817–821.
75. Reese, R.N., White, C.A., and Winge, D.R. (1992). *Plant Physiol.* **98**: 225–229.
76. Ortiz, D.F., Ruscitti, T., McCue, K.F., and Ow, D.W. (1995). *Journal of Biological Chemistry* **270**: 4721–4728.
77. Voegeli-Lange, R. and Wagner, G.J. (1990). *Plant Physiol.* **92**: 173–212.
78. De Knecht, J.A. et al. (1994). *Plant Physiology* **104**: 255–261.
79. Leopold, I., Gunther, D., Schmidt, J., and Neumann, D. (1999). *Phytochemistry* **50**: 1323–1328.
80. Ebbs, S., Lau, I., Ahner, B., and Kochian, L. (2002). *Planta.* **214**: 635–640.
81. Sneller, F.E.C. et al. (1999). *New Phytologist.* **144**: 223–232.
82. Bleeker, P.M., Schat, H., Vooijs, R., Verkleij, J.A.C., and Ernst, W.H.O. (2003). *New Phytologist.* **157**: 33–38.
83. Nakazawa, R., Ikawa, M., Yasuda, K., and Takenaga, H. (2000). *Soil Science and Plant Nutrition* **46**: 271–275.
84. Schmoger, M.E., Oven, M., and Grill, E. (2000). *Plant Physiol.* **122**: 793–801.
85. Hartley-Whitaker, J., Woods, C., and Meharg, A.A. (2002). *New Phytologist.* **155**: 219–225.
86. Meharg, A.A. and Macnair, M.R. (1992). *Journal of Experimental Botany* **43**: 519–524.
87. Zhang, W., Cai, Y., Downum, K.R., and Ma, L.Q. (2004). *Environ Pollut.* **131**: 337–345.
88. Zhao, F.J. et al. (2003). *New Phytologist.* **159**: 403–410.
89. Reese, R.N., Mehra, R.K., Tarbe, E.B., and Winge, D.R. (1988). *J. Biol. Chem.* **263**: 4186–4192.

90. Mehra, R.K. and Mulchandani, P. (1995). *Biochemical Journal* **307**: 697–705.
91. van Vliet, C., Anderson, C.R., and Cobbett, C.S. (1995). *Plant Physiol.* **109**: 871–878.
92. Schat, H. and Kalff, M.M.A. (1992). *Plant Physiology* **99**: 1475–1480.
93. Stillman, M.J. (1995). *Coordination Chemistry Reviews* **144**: 461–511.
94. Margoshes, M. and Vallee, B.L. (1957). *JACS* **79**: 4813.
95. Cobbett, C. and Goldsbrough, P. (2002). *Annu. Rev. Plant Biol.* **53**: 159–182.
96. Klaassen, C.D., Liu, J., and Choudhuri, S. (1999). *Annual Review of Pharmacology and Toxicology* **39**: 267–294.
97. Robinson, N.J., Tommey, A.M., Kuske, C., and Jackson, P.J. (1993). *Biochem J.* **295**: 1–10.
98. Palmiter, R.D. (1998). *Proc. Natl. Acad. Sci. USA* **95**: 8428–8430.
99. Murphy, A. and Taiz, L. (1995). *Plant Physiol.* **109**: 945–954.
100. Murphy, A., Zhou, J., Goldsbrough, P.B., and Taiz, L. (1997). *Plant Physiol.* **113**: 1293–1301.
101. Zhou, J. and Goldsbrough, P.B. (1994). *Plant Cell* **6**: 875–884.
102. Zhou, J. and Goldsbrough, P.B. (1995). *Mol. Gen. Genet.* **248**: 318–328.
103. Thomas, J.C. et al. (2003). *Biotechnol. Prog.* **19**: 273–280.
104. van Hoof, N.A. et al. (2001). *Plant Physiol.* **126**: 1519–1526.
105. Zhang, Y.W., Tam, N.F.Y., and Wong, Y.S. (2004). *Plant Science* **167**: 869–877.
106. Schafer, H.J., Greiner, S., Rausch, T., and Haag-Kerwer, A. (1997). *FEBS Lett.* **404**: 216–220.
107. Schafer, H.J., Haag-Kerwer, A., and Rausch, T. (1998). *Plant Mol. Biol.* **37**: 87–97.
108. Foley, R.C. and Singh, K.B. (1994). *Plant Mol. Biol.* **26**: 435–444.
109. Foley, R.C., Liang, Z.M., and Singh, K.B. (1997). *Plant Mol. Biol.* **33**: 583–591.
110. Van Hoof, N.A.L.M. et al. (2001). *Physiologia Plantarum.* **113**: 225–232.
111. Pan, A. et al. (1994). *Plant Mol. Biol.* **24**: 341–351.
112. Suh, M.C., Choi, D., and Liu, J.R. (1998). *Mol. Cells* **8**: 678–684.

MICROIRRIGATION

Dorota Z. Haman
University of Florida
Gainesville, Florida

Microirrigation has been defined in the Standards of the American Society of Agricultural Engineers (ASAE) as "frequent application of small quantities of water on or below the soil surface in form of drops, tiny streams or miniature spray." In microirrigation systems, water is applied through emitters or applicators placed along a water delivery line (lateral line) adjacent to the plant row (1). Microirrigation includes all methods of frequent water application, in small flow rates, on or below the soil surface. The terms drip, trickle, and spray irrigation, common since the 1960s, have been supplanted by the term "microirrigation," which is now commonly used throughout the industry. Microirrigation is sometimes called "localized irrigation" to emphasize that only a targeted part of the soil volume is wetted during water application. The other name that is frequently and incorrectly used is "low volume irrigation." In reality, a microirrigation system is characterized by low flow rates that are expressed in terms of water volume per time (L/h, m^3/h), but the total volume (liters or cubic meters) delivered can be large.

Microirrigation systems are designed to transport water from a source through a delivery network of pipes and various types of emitters to a crop. The general goal of a microirrigation system is to provide irrigation water uniformly and efficiently to a crop to help meet the water needs of plants and to maintain a favorable root zone water balance. Ideally, the volume of water is applied directly to the root zone in quantities that approach the consumptive use of the plants. Additional goals of microirrigation can be to increase fruit or fiber production, to maintain crop visual quality, to protect plants and/or fruits and flowers from extreme temperature conditions (hot or cold), to apply crop or soil chemicals (fertilizers, pesticides, etc.), or to dispose of wastewater, which can be an important water and nutrient resource.

Microirrigation is usually characterized by the following features: (1) low water application rates, (2) high frequency of water application, (3) water applied near or into the root zone, (4) low pressure delivery system, and (5) delivery of fertilizers and other agricultural chemicals with irrigation water (fertigation and chemigation).

In microirrigation systems, water is distributed using an extensive hydraulic pipe network (Fig. 1) that conveys water from its source to the plant. Outflow from the irrigation system occurs through emitters placed along the water delivery (lateral) pipes in the form of droplets, tiny streams, or miniature sprays. The emitters can be placed either on or below the soil surface. Microirrigation systems are classified by the type of emitters used in the system. These are drip, bubbler, spray jet (or microspray), and microsprinkler. The system can also be subdivided into the point source emitters and line source emitters (Fig. 1). Point source emitters are spaced sufficiently far apart to create individual discrete water patterns and they are used for trees or widely spaced shrubs. Line source emitters are used for row crop production. The emitters are placed along the line closely enough to create a continuous wetted band along the crop row. In addition, line source (drip) systems can apply water above the ground or below the ground. Subsurface drip irrigation (SDI) systems are becoming more popular due to potentially high efficiency.

A typical flow rate of a point source emitter is less than 12 L/h. The flow rate of a line source emitter is given per length of lateral and is also less than 12 L/h per meter. Bubblers emit much larger volumes of water, but usually less than 225 L/h, where the microsprinklers and jet sprays are limited to 175 L/h.

The emitter is designed to reduce the operating pressure from the supply line, and a small volume of water is discharged at the emission point at low pressure. Emitters can vary from sophisticated, constant

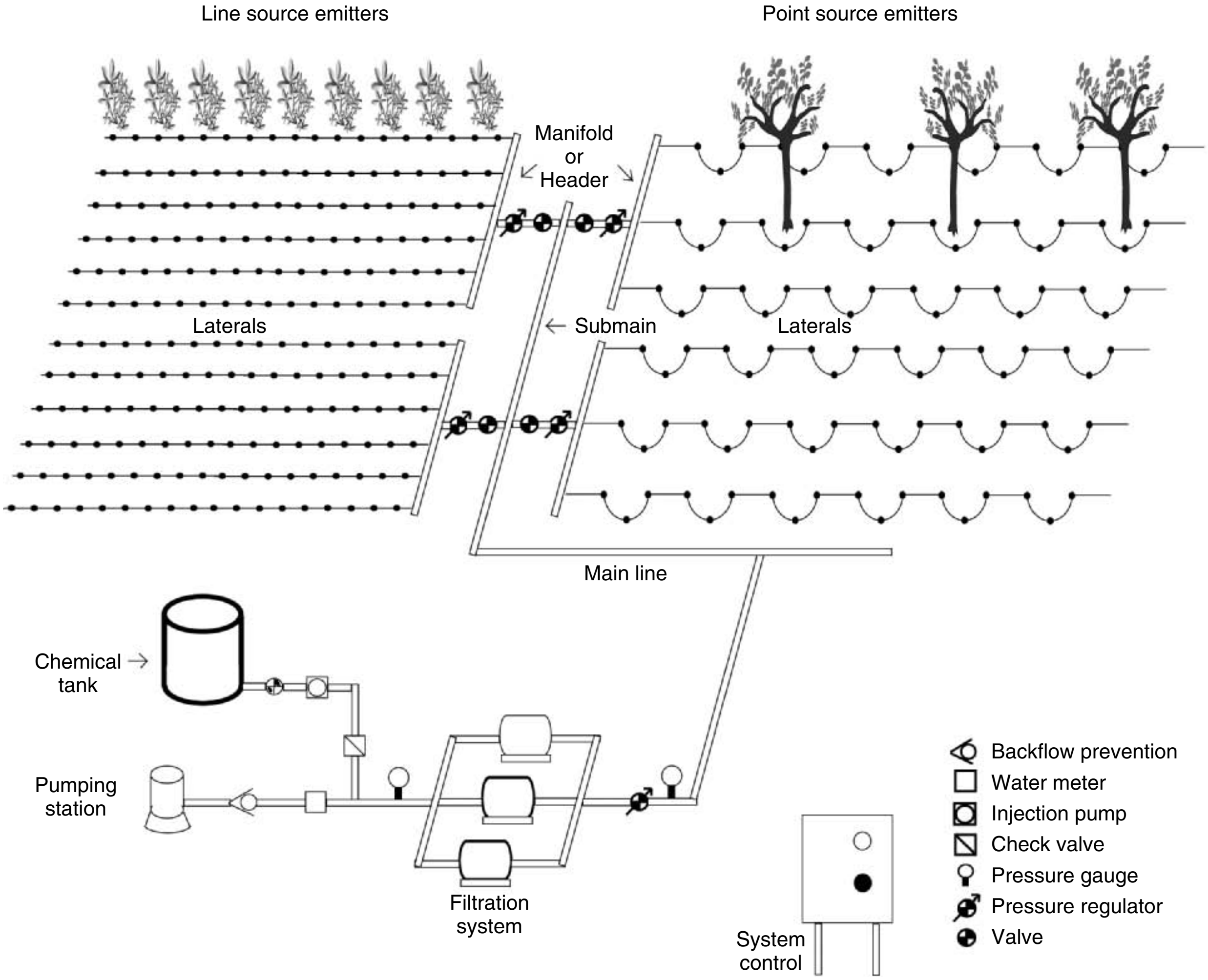

Figure 1. An example of a basic microirrigation system.

flow rate at variable pressure types of devices (pressure-compensating emitters) to very small, simple orifices. A large number of different types of emitters have been developed in attempts to find a perfect device. The main objective is to assure uniformity of water distribution. It is essential that the discharge from the emitter is uniform and that it not change significantly with small pressure variations in the system. At the same time, the emitter should be constructed in such a way that it does not clog very easily. The cost and the size are also important. The emitters presently available on the market can be classified into five distinct categories: (1) long path emitters, (2) short path orifice emitters, (3) vortex emitters, (4) pressure-compensating emitters, and (5) porous pipe or tube emitters (line source drippers).

Emitters are placed along laterals, the final water delivery lines, designed for uniform water distribution. The lateral line is generally constructed of flexible polyvinyl chloride (PVC) or polyethylene hose (PE). It is often placed above the ground, but it can be buried. For row crops, in the line source type of microirrigation system, a lateral line combines the function of the line and the emitter. These include laterals constructed from porous pipe, twin-bore PE pipe, or PE pipe provided with evenly spaced, built-in emitters. Each manufacturer of line source emitters supplies data with allowable length of laterals and recommended pressure for particular line source systems. Recommendations are based on the hydraulic characteristics of the product. The emitters, which may be connected to the lateral using a variety of connectors during installation of the irrigation system, are usually placed at predetermined positions, for example, at the base of the irrigated tree.

Lateral lines are attached to manifolds or submains. The manifolds and submains, which distribute the water to the specific parts of the field, are usually constructed from flexible smooth-walled, noncollapsible, black PE or flexible PVC pipe that can be left on the soil surface or buried. If rigid (usually white or purple) PVC pipe is used for the lateral line, it should be buried beneath the soil surface for protection against sunlight damage and prevention of algae growth in the line. The controls for adjustment of flow rate and pressure are usually located in the submain or manifold line along with valves and timing devices for the separate parts of the field. Water is delivered to the field by the main line. The main line is usually constructed from white PVC pipe that is buried beneath the soil surface for protection against harmful sunlight. Purple color is required for all PVC components in the system using reclaimed wastewater. The pipe should be properly rated for the particular application and able to withstand the design pressure in the system.

A main control station, often called the "control head," is usually located close to the water supply. A typical control station includes the pump, a backflow prevention system, a chemical injection system for fertilizers, chlorine, or other chemicals, and a combination of different filters. A main line valve and flowmeter are also included in the control head. Microirrigation systems can be controlled manually or automatically. Automatic control can be electromechanical (clock) or electronic (computer). The controller is often located next to other components of the control station. It can control the main valve, chemical injection, backflushing of filters, solenoid valves, and other controls located at remote locations in the irrigation system. Depending on the system, all or some of these components can be automated.

In microirrigation systems, only a portion of the total field area is wetted. The goal is to only wet the root zone of the crop of interest. Water flowing from the emitter is distributed in the soil by gravity and capillary forces creating the contour lines presented in Fig. 2, often referred to as "onion" patterns. The exact shape of the wetted volume and moisture distribution will depend on the soil texture, initial soil moisture, and, to some degree, the rate of water application. In the line source type of microirrigation tube that is frequently used for row crops such as vegetables, where the emitters are spaced very closely, individual "onion" patterns connect, creating a continuous moisture zone along the row.

Microirrigation systems have many potential advantages compared to other irrigation methods. Most of them are related to the low rates of the system. It can be argued that some of these benefits are not unique to a microirrigation system. However, certain combinations of these advantages are responsible for uniqueness of microirrigation in contrast to other systems.

WATER SAVINGS

Irrigation water requirements can be smaller with microirrigation when compared with other irrigation methods since a smaller portion of the soil volume is wetted, there is decreased evaporation from the soil surface, and there is reduction or elimination of the runoff. The losses due to evaporation from the soil are significantly reduced compared with other irrigation systems because only a small surface area under the plant is wetted and it is usually well shaded by the foliage. A microirrigation system allows for a high level of water control application and water can be applied only when needed in precisely controlled amounts. Deep percolation can be minimized or avoided.

SMALLER FLOW RATES

Since the rate of water application (L/S) in microirrigation systems is significantly lower than in other systems, smaller sources of water can be used for irrigation of the same acreage. The delivery pipes, the pump, and other components of the system can be smaller and therefore more economical. The systems operate under low pressure (35–200 kPa) and require less energy for pumping than high pressure irrigation systems.

APPLICATION OF CHEMICALS

Microirrigation systems allow for a high level of control of chemical applications. The plants can be supplied with the exact amount of fertilizer required at a given time. They are applied directly to the root zone and reduction in the total amount of fertilizer used is possible. There is also an advantage to the frequent application of fertilizers

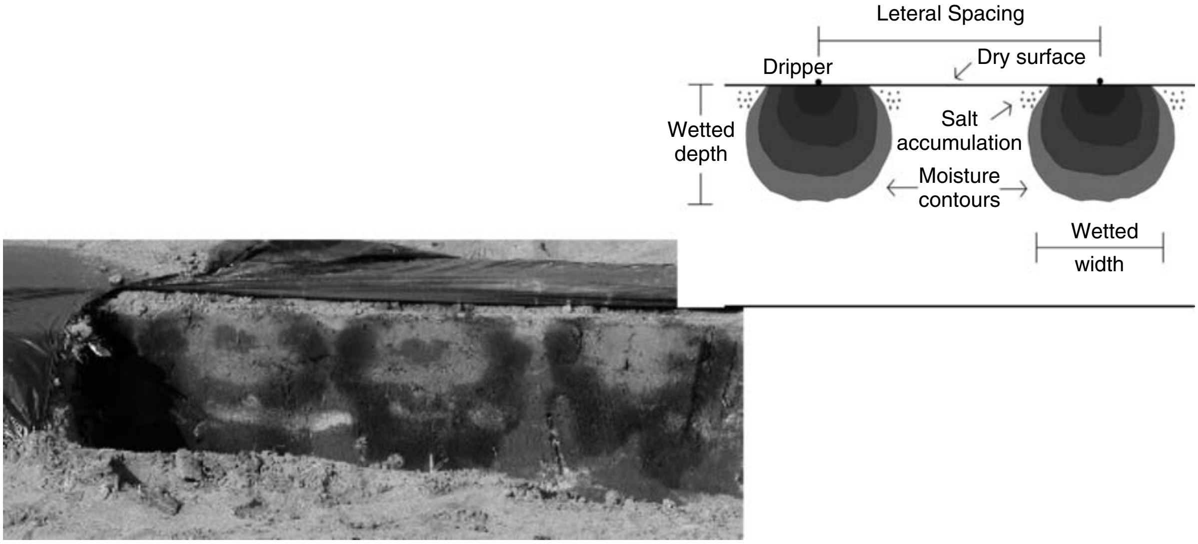

Figure 2. Wetted pattern under drip irrigation (line source type of emitter) in a sandy soil.

through the system (fertigation) in humid climates. In case of rain, only a small portion of recently applied fertilizer will be washed out and it can easily be replaced through the irrigation system, reducing the potential environmental hazard due to nutrient leaching from the production system. Fertigation is more economical, provides better distribution of nutrients throughout the season, and reduces groundwater pollution due to the high concentration of chemicals that could ordinarily move with deep percolated water. Other chemicals, such as herbicides, insecticides, fungicides, nematicides, growth regulators, and carbon dioxide can also be efficiently applied through microirrigation systems to improve crop production. This process of chemical application is called chemigation.

WATER SOURCES WITH HIGH SALT CONTENT

Water with relatively high salt content can be used in microirrigation systems. For optimum plant growth a certain range of total water potential in the root zone must be maintained. The potential defines how difficult it is for a plant to extract water from the soil. Large negative numbers are characteristic of very dry soils with low total water potentials, while potentials near zero reflect soils near saturation. The total water potential in the root zone is a sum of the matrix potential and osmotic potential. Since matrix potential is close to zero under well-managed microirrigation (high moisture content), the osmotic potential component can be a relatively large negative value, indicating high salt content, without a harmful effect on plant growth. This is not true for other irrigation systems.

IMPROVED QUALITY OF THE CROP

Microirrigated plants are supplied very frequently with small amounts of water and the stress due to the moisture fluctuation in the root zone is reduced to the minimum, often resulting in larger and better quality yield. In arid climates, or during dry seasons, proper water management can control the harvest timing.

ADAPTATION TO ANY TOPOGRAPHY

Microirrigation systems can operate efficiently on hilly terrain if appropriately designed and managed. A well-managed microirrigation system will minimize runoff even on hilly terrain.

ADDITIONAL ADVANTAGES OF MICROIRRIGATION SYSTEMS

During dry seasons or in arid climates, disease and insect damage can be reduced under the microirrigation system since the foliage of the plant is not wetted. With a small portion of soil surface being watered, field operations can be continued during irrigation. Water distribution is not affected by the wind in drip irrigation. However, wind can have some effect on spray jet patterns. Germination and water uptake by the weeds between the rows can also be significantly reduced since these areas remain dry. Microirrigation systems can also be extensively automated, decreasing labor and operating costs.

To operate satisfactorily, a microirrigation system has to be correctly designed and managed to account for the physical properties of soil, quality of irrigation water, and water requirements of the grown plants. Usually, microirrigation requires a higher management level than other irrigation systems. With all the advantages listed above, a microirrigation system is not a system without problems.

CLOGGING

One of the biggest problems encountered under microirrigation is the clogging of the emitters. The small openings can easily be clogged by soil particles, organic matter, bacterial slime, algae, or chemical precipitates. Microirrigation systems require very good filtration (most often recommended is 200 mesh screen) even with a high quality water supply.

MOISTURE DISTRIBUTION

Moisture distribution depends largely on the soil type being irrigated by the microirrigation system. In some soils, for example, deep sands, very little lateral water movement (low capillary forces) can create many problems. Under these conditions it is difficult to wet a significant portion of the root zone since most of the water moves downward due to gravity. It is also more difficult to manage the irrigation without deep percolation since only a small amount of water can be stored in the wetted volume desired. Increasing the number of emitters per plant may improve water distribution in the soil. As a result, coarse sands require much closer spacing of emitters than fine soils. In general, for any soil, the number of emitters and their spacing must be based on the geometry of wetted soil volume. It is important to realize that the microirrigation system wets only a limited portion of the potential soil root volume. Most of the plants can perform very well under these conditions. However, there is a minimum volume of roots that have to be wetted or a reduction in yield will result.

SALT BUILDUP

Microirrigation systems can use saline water. However, a problem may occur from salts accumulating at the edges of the wetted zone during prolonged dry periods. Light rain can wash these salts into the root zone and cause injury to the plants. In arid climates, where the rainfall is less than 10 in./yr, an additional irrigation system (sprinkler or surface) may be necessary to leach accumulated salts from the soil. In areas with heavy rainfall the salts will be washed out of the root zone before significant accumulation occurs.

Table 1. Extent of Microirrigation (in hectares) in Selected Countries and the World During the Period 1981–2000

Country	1981	1986	1991	2000
United States	185,300	392,000	606,000	1,050,000
India	20	0	55,000	260,000
Australia[a]	20,050	58,758	147,011	258,000
Spain	0	112,500	160,000	230,000
South Africa	44,000	102,250	144,000	220,000
Israel	81,700	126,810	104,302	161,000
France	22,000	0	50,953	140,000
Mexico[a]	2,000	12,684	60,000	105,000
Egypt	0	68,450	68,450	104,000
Japan[a]	0	1,400	57,098	100,000
Italy	10,300	21,700	78,600	80,000
Thailand[a]	0	3,660	45,150	72,000
Colombia[a]	0	0	29,500	52,000
Jordan	1,020	12,000	12,000	38,300
Brazil[a]	2,000	20,150	20,150	35,000
China	8,040	10,000	19,000	34,000
Cyprus	6,000	10,000	25,000	25,000
Portugal	0	23,565	23,565	25,000
Chinese Taipei[a]	0	10,005	10,005	18,000
Morocco[a]	3,600	5,825	9,766	17,000
Other[a]	50,560	38,821	100,737	177,000
Total world	436, 590	1, 030, 578	1, 826, 287	3, 201, 300

[a] Areas for these countries adjusted from 1991 figures according to the average percentage increase of other countries that had provided updated information to ICID.
Source: Adapted from Reference 2.

INITIAL COST

The initial investment and maintenance cost for a microirrigation system may be higher than for some other irrigation methods. Filters, chemical injectors, and possible automation components add to the cost of a microirrigation system. Actual costs will vary considerably depending on the selection of a particular microsystem, required filtration equipment, water quality, water treatment, and selection of automation equipment.

ADDITIONAL DRAWBACKS

Rodents and insects can create additional maintenance problems by chewing holes in the plastic. In addition, some components of the system can easily be damaged by persons unaware of their locations. A microirrigation system does not provide significant frost protection when compared to sprinkler irrigation; however, it is commonly used for this purpose in citrus production where it can provide some protection.

CURRENT IRRIGATED AREA

In 1977, the Food and Agriculture Organization of the United Nations (FAO) estimated that the total global irrigated area was 223 million ha. By 1996 this estimate increased to about 262 million ha. The 2000 irrigation survey conducted by the *Irrigation Journal* (2001) listed 25.5 million ha irrigated in the United States. Approximately 12.7 million ha (49.9%) were irrigated by sprinkler irrigation, 11.5 million ha (45.1%) by gravity irrigation, and 1.3 million ha (4.9%) by microirrigation.

Results of the surveys conducted by the International Commission on Irrigation and Drainage (ICID), summarized (Table 1) by Reinders (2), indicate that microirrigation has increased from 0.4 to 3.2 million ha between 1981 and 2000. In the United States, California, Florida, Washington, Texas, Hawaii, Georgia, and Michigan account for approximately 91% of the microirrigated land area. It is a small fraction of the total irrigated area; however, many of the high value crops that require intensive production practices are grown under microirrigation and the area of marginal lands that are microirrigated is increasing.

Almost all crops can be suitable for microirrigation; however, due to the high cost, microirrigation is primarily used on high value perennial crops, tree and vine crops, fruits, vegetables, and ornamentals. There is growing interest in applying microirrigation to lower valued field crops, such as cotton and corn or sugar cane, through the use of multiyear subsurface drip irrigation (SDI). Application of microirrigation for landscaping, greenhouses, and nurseries has also increased tremendously. Large containerized nursery plants are typically irrigated with microirrigation. In addition, an increase of microirrigation in residential properties and highway roadsides often prompted by water restriction and/or high cost can be observed in many areas of the United States.

BIBLIOGRAPHY

1. ASAE Standards. (2003). *Design and Installation of Microirrigation Systems*. ASAE EP405.1 FEB03, ASAE, St. Joseph, MI.

2. Reinders, F. (2000). Micro-irrigation: a world overview. In: *Proceedings of the Sixth International Micro-Irrigation Congress*. Capetown, South Africa, Oct. 22–27

MICROIRRIGATION: AN APPROACH TO EFFICIENT IRRIGATION

Harenda Singh Chauhan
G.B. Pant University of
Agriculture and Technology
Pantnagar, India

Microirrigation, earlier known as Drip or Trickle irrigation is a highly efficient method of water application. It developed and expanded globally at a fast rate in the last two decades in more than 35 countries with maximum coverage in the United States. This is most appropriately applicable to widely spaced crops such as fruits, vegetables and field crops. Its different variations consist of surface trickle, bubbler, microsprinkler, spray, mechanical move, pulse, subsurface drip, all of which are covered under a single term microirrigation. In the present article, its advantages, components, system design, fertigation, automation and other aspects has been described. Also the extent of application to different crops, and in different countries, limitations, scope, research trends and strategies of promotion has been discussed.

INTRODUCTION

The world is facing a serious shortage of fresh water due to dwindling resources. There is growing competition for clean water making less water available for agriculture, the largest consumer of water which accounts for more than 70% of total withdrawals for the world at large. The average irrigation efficiency of the world is fairly low. As per Gittinger (1), if this efficiency can be raised by 10%, the total conserved water will equal the water required in the world for all other uses. According to Smith (2), sustainable food production will depend on the judicious use of water resources to meet future food demands. Water productivity in terms of output per unit of food per cubic meter of water needs to be increased both in irrigated and rain-fed agriculture. In short, there is need for more crop per drop.

As urbanization, industrialization, and other uses increase, the availability of water for agriculture is likely to be reduced. By 2030, it is estimated (3) that irrigated areas will have expanded by 23%. However, decreasing water availability will allow an increase of only 12% more water for agriculture. This would necessitate an increase in water productivity by increasing water use efficiency and would require improving irrigation efficiency from the present 43% to 50% by 2030. This would in consequence require reducing irrigation water losses by introducing appropriate technologies and improved water management and cultural practices.

Water management involves various important aspects, in which there is the necessity and large scope for improvement. One of the important aspects is the field application of irrigation water. In surface application, common efficiencies range from 50–60%, whereas in drip irrigation it ranges around 90%. Although a microirrigation system appears initially costly and not necessarily applicable to similar land conditions as surface irrigation, it is definitely water saving and improves the yield and quality of crops. This article discusses microirrigation and its nature, extent, and future scope.

Historical Development

Earlier experiments on drip irrigation were done in Germany around 1860 using clay pipes having open joints, for both irrigation and drainage. In the United States, some work was done at Colorado State University in 1913. With the introduction of perforated plastic pipes in Germany, Canada, and the United States for subirrigation, this concept received quite an impetus. In these efforts, the work of Symcha Blass, an Israeli engineer is worth mentioning; he observed in 1940 that a big tree near a leaking pipe showed more vigorous growth than other trees in the area. From this observation, he arrived at the concept of irrigation, wherein water was applied in small quantity, virtually drop by drop. The earlier drip irrigation system consisted of plastic capillary tubes attached to large pipes. The experiments around 1960 at Arava in Israel showed very good results. With the development of new plastic components after World War II, great headway was made in manufacturing plastic pipes and other accessories. Drip system components were gradually properly established and sold in different countries. Per one survey, this system has been adopted in 35 countries and is installed in an area of 1,784,846 ha.

Considering the rapid growth of microirrigation, six International Congresses have been organized so far: the first in Israel in 1971, the second in California in 1974, the third in California in 1985, the fourth in Australia in 1988, the fifth in Florida in 1995, and the sixth in South Africa in 2000.

MICROIRRIGATION: DEFINITION AND TYPES

According to Nakayama and Bucks (4), trickle irrigation is the slow application of water on, above, or beneath the soil by surface trickle, subsurface trickle, bubbler, mechanical move spray, or pulse systems.

In this method, water is applied by low pressure pipes at a low rate near the root zone, in small quantities at frequent intervals. The water is applied as discrete or continuous drops, small streams or sprays, through emitters placed near the plants. It is also known as localized irrigation, drip irrigation, trickle irrigation, daily flow irrigation, and microirrigation.

Microirrigation is presently the most accepted name that includes all variations such as surface trickle, subsurface trickle, bubbler, mechanical move spray, and pulse, which are briefly described below.

Surface Trickle. In this system, the emitters and laterals are placed on the soil surface. It is the most common type and is also called drip irrigation. It is used for widely spaced as well as row crops. Discharge rates are lower than 12 L/h for single outlet, point source emitters, and less than 12 L/h/m for line source emitters.

Bubbler. In this method, water is applied to the soil surface in a small stream or fountain with a point source discharge much higher than surface or subsurface trickle but less than 225 L/h. A small basin is required for proper water distribution.

Spray. In this system, water is applied as a small spray, fog, or mist Discharge is less than 175 L/h, and it is used to irrigate trees or other widely spaced crops.

Mechanical Move. In this system, the concept of the bubbler is extended to large spaced row crops. including travelling trickle and drag or hose reel systems.

Pulse. It has high discharge rate emitters, and thus short application times. The discharge ranges are 4–10 times larger than surface trickle emitters. This system has less clogging problems but needs inexpensive pulse emitters and automatic controllers.

Subsurface Drip Irrigation. This is another variation of drip irrigation, also covered under the term microirrigation. Its only difference from surface drip irrigation is that it is installed below the soil surface.

In this system, water is applied slowly below the surface through emitters whose discharges have the same range as surface trickle emitters. This is useful for small fruit or vegetable crops. Historically, it is the oldest modern irrigation method used in the United States. House (5) successfully used it for irrigating apples, alfalfa, and cereal crops with porous subsurface laid pipes. He, however, found it extremely costly and thus not economically feasible for field crops. The advancement in plastic materials, however has made it economically feasible for several crops. Blass (6) applied drip irrigation using laterals with buried drippers, but the practice was discontinued because of emitter clogging and root penetration. However, the practice was revived for the following reasons: (1) The same system could be used for several consecutive crops with minimum tillage. (2) The labor requirement in removing and installing the tubing with each crop was reduced. (3) There was less interference with weeding, spraying, or harvesting of the crop with buried laterals.

CROPS SUITED AND ADVANTAGES

Crop Suited

There are many crops for which drip irrigation has been used. The major fruit crops for which it has been used are citrus, deciduous, strawberry, grape, mango, pomegranate, banana, peach, pear, apricot, papaya, and coconut. Some of the important vegetables, cash crops, and plantation crops, to which it has been applied are tomato, potato, sugarcane, cotton, olive, coconut, watermelons, cardamom, and ornamental trees.

Advantages

These are the main reasons that drip irrigation is used:

1. It provides a high degree of control for water application. Using such control, it is possible to attain much higher efficiencies of the order of 90% compared to sprinkler irrigation (60–80%) and surface irrigation (50–60%). Thus, it provides high water saving and high water use efficiency compared to other methods of water application.
2. It supplies water at a rate sufficient to satisfy evaporative demand by maintaining high matric and osmotic soil water potentials, which minimize water and osmotic stresses. In conventional systems of water application such as flood or sprinkler methods, the fluctuations in soil matric and solute potentials are comparatively large. The lesser fluctuation and thus less stresses result in better plant growth, an increase in yield, and better quality of fruits.
3. There is partial wetting of soil with drip irrigation. Because of this, less water is lost through evaporation. However, on a seasonal basis, there may not be more water saving for this reason than in conventional methods. But this helps in reducing weeds, resulting in a lower cost of weed control. It is helpful in carrying out further cultivation operations, such as spraying and harvesting, without damaging soil structure.
4. The foliage remains dry, so there is less growth and incidence of plants diseases, reducing the use of pesticides and fungicides.
5. It permits using saline water. The increase in salt concentration because of soil drying between two irrigations is less in drip than in other irrigation methods. Salts from the wetted portion are frequently leached from the active root zone. Leaf burn is avoided because there is no foliar application in this method.
6. It can be well adapted to marginal soils, such as soils with high permeability and low water holding capacity, sands, leached tropical soils, and also for steep land and slowly permeable soils.
7. It requires less agricultural chemicals such as fertilizers, herbicides, and pesticides. Because their application is according to precise plant needs, the quantity required is less, leaching losses are less, and use is more efficient. It is thus cost as well as labor saving, and causes less nitrogen pollution in groundwater.
8. It uses less energy compared to high head sprinklers.
9. It is very well adapted to greenhouses in water and nutrient application without foliage wetting, such as needed for flowers and potted plants. It is also useful for water and fertilizer application in mulched crops and with wind tunnels.

Table 1. Yield Increase, Water Saving, and Payback Period

Sl.No.	Crop	Investment, \$/(Rs.)[a]	Payback, Years	Water Saving, %	Yield Increase, %
1.	Pomegranate	625/(30,000)	1	45	45
2.	Grape	917/(44,000)	>1 year	48	23
3.	Sugarcane	989/(47,500)	1	56	33
4.	Banana	979/(47,000)	1	49	52
5.	Tomato	625/30,000	NA	39	50
6.	Chili pepper		NA	62	44
7	Sweet lime			61	50
8	Watermelon			40	88
9	Okra			40	16
10	Cabbage			60	2
11	Cotton			53	26

[a] Rs. refers to Indian currency; 1\$ = Rs48 (variable with time).

Advantages of Subsurface Drip

Although another variation of microirrigation, its advantages need special mention. According to Phene et al. (7), it has the following advantages compared to surface drip: (1) There is a substantial increase in water use efficiency(WUE). (2) It enables precise fertilization with high frequency irrigation. There is negligible evaporation and almost total elimination of deep percolation and nitrate–nitrogen leaching. (3) There is long-term sustainability of the system.

WATER SAVING AND YIELD INCREASE

The most important aspects of microirrigation are water saving and increased yield compared to conventional irrigation. The experience for a few crops grown in India, as reported by Uriel (8) (slightly modified), are given in Table 1.

Table 1 shows that there is a 40 to 60% saving of water with drip irrigation over conventional irrigation. This water can be diverted to increase the area under irrigation or to other uses. Similarly, the yield increases from 23 to 88%. For vegetables and closely spaced crops, the investment is about \$979 (Rs.47000)/ha, whereas for trees, it is around \$625 (Rs.30000)/ha. For bananas, the payback is less than 1 year, whereas for fruits like pomegranate, the payback is 1 year.

COMPONENTS AND SYSTEM DESIGN

Components

Emission Devices. This is the most important component, for which sometimes the method of irrigation itself is named. This component acts as energy dissipater, reducing the inlet pressure head (0.5–1 atmosphere) to zero atmospheres at the outlet. It is generally made of polypropylene. Some types may be called orifice emitters, orifice-vortex emitters, long path emitters, compensating long path emitters, long path multiple outlet emitters, grove and flap short path emitters, grove and disc short path emitters, and twin wall emitters. The discharge rate of an emitter is characterized by its mean at normal operating pressure and the coefficient of manufacturing variation. Its operation varies in the range of 1–2 atmosphere. In orifice type emitters, the pressure is dissipated through a small hole 0.4–0.6 mm. They are highly prone to clogging. In long path emitters, the pressure is dissipated by flow through a long narrow path. The simplest type is the microtube. The diameters of microtubes vary from 0.6–1 mm. The energy loss in long path emitters can be increased by creating a tortuous long path in the form of a spiral or labyrinth. Sometimes, the energy can be dissipated through perforations or small holes. Biwall or twin bore are such systems. In these systems, there is a small tube or a supply chamber operating at a high pressure and fitted with large, widely placed orifices. For each inner orifice, there are many outer orifices in the emission chamber. There are several in-line systems with different designs conceived by different manufacturers.

In-Line Drippers. In-line drippers are based on the long path principle. They are molded plastic accessories, in which the thread is a long narrow passage governing the slow discharge of water. Some emitters are attached to the lateral tubing and in others, the lateral tubing forms the outer wall of the tubing fixed along with the line. They may have discharges of 2–8 L/h at 1 atmosphere.

On-Line Drippers. These drippers are fixed on the lateral by punching suitable holes in the pipe. They may be classified as nonpressure compensating type or pressure compensating type.

Simple Type with Laminar Flow. In this type, the discharge is directly proportional to pressure. It has a simple thread, labyrinth, zigzag path, or other arrangement to dissipate energy.

Turbo Key Type. This is a blockage resistant and pressure compensating type, made of stabilized polymers, available in 2,4, and 8 L/h discharge.

Pressure Compensating Type. This type, provides uniform discharge between pressures of 0.3–3.5 atm. This type dripper is provided with a high quality rubber diaphragm to control pressure. It gives 2,3,4, and 8 L/h discharge at varying pressures and is suitable for sloping topography.

Built-in Dripper Tube. In this type, the drippers are welded to the inside of the tube during extrusion of polyethylene pipes. They are provided with an independent pressure compensating discharge mechanism and wide water passage to prevent clogging.

Other Components

Lateral Pipes. These are pipes with diameters of 12–33 mm and wall thicknesses sufficient to withstand pressures of 4–6 atmospheres, to which the emitters are connected.

These pipes should be flexible, noncorrosive, resistant to solar radiation, and able to withstand temperature fluctuation. Normally, pipes are made of diameters 10,12,16, and 20 mm, 1–3 mm thick, and generally laid above the ground. They are commonly manufactured from low density polyethylene (LDPE) or linear low density polyethylene (LLDPE). LLDPE provides better protection from the ultraviolet rays of the sun.

Main and Submain Pipes. They are placed below the ground and supply water to the laterals. They are made from rigid PVC or high density polyethylene. The pipes may be 65 mm or more in diameter and can withstand pressures of 6 kg/cm^2. These pipes are laid underground for long life. For submains, rigid PVC, HDPE, or LDPE pipes are generally used; diameters are 32–75 mm and can withstand pressures of 2.5 kg/cm^2. They are commonly laid above ground.

Filters. This is an important component used to reduce or minimize blocking or clogging of emitters. They are of three types.

Media Filter. Required to remove organic matter such as algae. It is made up of a circular tank filled with layers of coarse sand and different sizes of gravel and valves for flushing the assembly in case of clogging. It is available in different sizes and capacities as required by the system.

Hydrocyclone Filter. Also known as a vortex sand separator. When the irrigation water has more sand, it is used as a prefilter to remove the sand before it enters the drip system. It is generally followed by a screen filter.

Screen Filter. It is fitted in series with a gravel filter to further remove impurities like fine sands, dust, etc. from the water. It consists of a single or double perforated cylinder placed in a plastic or metal container. Generally, 100–200 mesh screens are used in these filters. Screens are specified by the diameter of the pipe, recommended flow rate, total surface area of the filter, or by the cleaning method. The head loss needs to be periodically measured, and if it exceeds the permissible specified limits, it needs to be cleaned.

Fertigation and Chemigation

The process of adding the fertilizer with irrigation water is also known as fertigation. When other agrochemicals for control of insects and pests are used along with fertilizers, instead of fertigation it may be termed chemigation. Fertilizer application through microirrigation not only helps in providing fertilizer economy, it also minimizes health concerns. However, the potential to apply other chemicals is considered a great risk. Microirrigation cannot be used for chemigation with chemicals such as fungicides, herbicides, and growth regulators, that need to be applied to plant foliage. Three methods commonly used for fertigation consist of a bypass pressure tank, a venturi, or direct injection system.

In a bypass system with the closing of the main valve, a certain fractional quantity of flow is made to pass through the fertilizer tank. This bypassed water, dissolving the fertilizer, goes to the drip system. In the venturi type, a suction head is created at the constriction which sucks the fertilizer through the system. In the direct injection type, pumps of piston or diaphragm type, operated by system pressure inject fixed quantities of fertilizers into the system.

Automation

This refers to operation of the system with minimum manual intervention and is useful in large areas. This may be done in three ways. The first method is time-based. In this method, the time of operation is estimated according to the volume of water required and the average flow rate of water. The duration of irrigation required has to be determined for each section. The duration for each valve is fed into the controller along with system start time, and the controller clock is set with the current day and time. The clock actuates the starting time of the first valve and simultaneously the pump. After the duration of the first valve is over, the first valve is closed, or the controller switches to the next valve. The second method is volume-based. In this system, the preset amount of water can be applied in the field by using automatic control metering valves. Automation using a volume-based system is of two types. In one type, the valve with pulse output provides one pulse after completing one dial of the valve. The controller accepts the pulse input and counts the volume per pulse. The volume required in a segment, can be programmed in the controller. In the second type, the valves are placed near each segment, and no controller is required. The valves are interconnected in series through a control tube. During sequential operation, only one valve remains open. The third system is real-time feedback based on the demand of the plant. The plant itself determines the degree of irrigation required. Sensors such as tensiometers, relative humidity sensors, rain sensors, and temperature sensors control irrigation scheduling. Conventional scheduling is based on refill and full point. A real-time system controls and holds soil moisture near a constant value.

System Design

Collection of General Information. This requires collection of basic data including information about water source; topography; crops to be grown; general nature of soil and its characteristics such as texture, depth, infiltration etc.; climate; and rainfall. Such information is necessary for irrigation scheduling.

Field Layout. Depending on the type and location of the source, size, geometric layout, and topographical features, the planning is done. A contour map is helpful, but if the highest point, lowest points, ridge line, and drainage lines can be located, the layout of the system and the lengths of mains and laterals are decided.

Crop Water Requirement. The peak irrigation requirement (PIR) expressed in L/day/plant, may be estimated from the following relationship:

$$\text{PIR} = E^{*}{}_{\text{pan}} \times A \times K_{\text{p}} \times K_{\text{c}} \times K_{\text{r}} \times E_{\text{a}} + L_{\text{r}} - R$$

where

$E^{*}{}_{\text{pan}}$ = (mm/day) is the average maximum pan evaporation;

A = total area allocated to each plant, m^2;

K_{p} = the pan factor (Eto = $K_{\text{p}}.E_{\text{pan}}$ is reference crop evapotranspiration; K_{p} is less than 1 and may be taken as 0.8);

K_{c} = crop coefficient, selected for a given crop, growth stage and climatic conditions and has to be experimentally determined or adopted from available information;

K_{r} = a reduction factor, based on crop ground coverage, also known as the canopy factor varies from 0–1, and for a canopy occupying full K_{r}, is equal to 1;

E_{a} = overall application efficiency ($E_{\text{a}} = k_{\text{s}}.k_{\text{u}}$, k_{s} is a coefficient, whose value is less than 1 and takes care of deep percolation and other losses, k_{u} takes care of application uniformity smaller than 1),

L_{r} = extra amount of water needed for leaching, taken as 0, where there is no salinity and no leaching problem;

R = rainfall taken as 0, where in an arid region, a crop is to be grown only with irrigation.

For many maturing crops with good water holding capacity, different factors may be taken as 1, and a rapid estimate may be made using the relationship,

$$\text{PIR} = 0.8E_{\text{pan}} \times A \times E_{\text{a}}$$

Monthly Water Requirement. As per INCID (9), a simplified estimate may be obtained as

$$V = E_{\text{p}} \times K_{\text{c}} \times K_{\text{p}} \times A \times N(\text{liters})$$

where E_{p} is the mean monthly pan evaporation (mm/day), A is the area to be irrigated m^2, and N is the number of days in a month.

Net volume to be applied may be written as

$$V_{\text{n}} = V - R_{\text{e}}$$

where R_{e} is the effective rainfall (mm).

The operating hours of a system for a month may be obtained as

$$T = \frac{V_{\text{n}} \times W_{\text{p}}}{\begin{array}{c}(\text{Drippers/plant}) \times \text{plant numbers}\\ \times\text{dripper discharge}\end{array}}$$

where W_{p} is the wetted area.

Operating hours per application may be written as

$$T_{\text{a}} = T/N_{\text{m}}(N_{\text{m}} \text{ is number of application/month})$$

System Design

Laterals may be designed by considering the flow of each lateral (de) obtained by multiplying the discharge of each dripper by the number of drippers per lateral. The diameter of submains may be designed by considering the flow of each submain line (ds), obtained by multiplying by the number of laterals.

The design of a main line may be done by considering the total flow of all the submains (dm), obtained by multiplying (ds) by the number of laterals.

The design of diameters of laterals, submains, and mains may be done by considering the pressure drop due to the friction of each lateral, submain, and main using the head loss relationship of Blassius or the William and Hazen equation.

The pump horsepower may be determined from the friction head losses of mains, submains, and laterals (H_{f}), the static head for the pump-well system (H_{s}), and the operating pressure head of the emitter (H_{e}):

$$\text{Total head } H = H_{\text{f}} + H_{\text{s}} + H_{\text{e}}$$

Thus,

$$\text{HP} = \frac{H \times \text{d}_{\text{m}}}{75 \times M_{\text{p}} \times M_{\text{m}}}$$

where M_{p} and M_{m} are the efficiencies of the motor and the pump, and d_{m} is the discharge of the main in liters per second.

EXTENT OF MICROIRRIGATION

There has been a very fast rate of adopting and extending microirrigation in different countries in the last 20 years, and several surveys have been carried out to find its coverage.

From the third international survey by ICID, reported by Bucks (10), it was found that from an area of 412,760 ha in 1981, it increased to 1,082,631 ha in 1986, and 1,768,987 ha in 1991. This was a 63% increase in 5 years and a 329% increase in 10 years. The major users of microirrigation in decreasing order, according to a survey in 1991, were the United States, Spain, Australia, South Africa, and Israel. Other countries having sizable microirrigation were Thailand, Columbia, Jordan, Brazil, China, Cyprus, Portugal, Chinese Taipei, and Morocco.

The fourth survey has been in process through the endeavor of Kulkarni of ICID. Preliminary feedback has been reported from 1981–2000, 19 years, by Reinders (11)

Table 2. Microirrigation Area in a Few Countries Over 19 Years[a]

Country	1981, ha	1986, ha	1991, ha	2000, ha
USA	185,000	392,000	606,000	1,050,000
India	20	0	55,000	260,000
Australia	20,050	58,738	147,011	258,000[b]
Spain	0	112,500	160,000	230,000
South Africa	44,000	102,250	144,000	220,000
Israel	81,700	126,810	104,302	161,000
France	22,000	0	50,953	140,000
Mexico	2,000	12,684	60,000	105,000*
Egypt	0	68,450	68,450	104,000
Japan	0	1,400	57,098	100,000[b]
Italy	10,300	21,700	78,600	80,000
Others	70,770	434,036	294,873	523,300
World	436,590	1,030,578	1,826,287	3,201,300

[a] Reference 11.

[b] Hectarage adjusted from 1991 figures according to the average percentage increase of other countries, as per data obtained by Dr. Kulkarni of ICID.

and is given in Table 2. From 1981–1986, a 136% increase took place and from 1986–1991, a 77% increase took place. From 1981–2000, 19 years, usage increased 633%.

Although there has been a tremendous increase in the use of microirrigation, the total area under microirrigation continues to remain only 0.8% of the total irrigated area of the world (10).

Besides the ICID surveys, according to Uriel (8), there have been several other assessments of the extent of microirrigation in different countries which are briefly given below.

According to a survey, it was estimated that in the Mediterranean coastal plains in Palestine, about 70.5% of vegetables was irrigated by drip irrigation, 28% by sprinklers, and a negligible area by traditional methods. It was found that 87% of farmers in the Jordan Rift Valley can grow two crops with modern methods in comparison with 83% in the coastal region. In Jordan, mainly in the Jordan Rift Valley, 64% of farmers use drip irrigation.

In the mountainous area around Lake Titicaca between Peru and Bolivia, a combination has been made between simple rural greenhouse and drip irrigation for growing potatoes and vegetables. Potatoes grown in greenhouses produced 350 kg/50 square meters an increase of 500% in yield in comparison to traditional methods. In Mauritius, about 33,000 small farmers cultivate about 25,000 ha of sugarcane. In several other countries like Nicaragua, southern Argentina, Thailand, Turkey, Peru, Mozambique, and China, drip irrigation is being tried for vegetables and orchard crops.

According to Smajstrla et al. (12), about 20% of the U.S. microirrigated fruit crops consist of Florida fruit crops which are primarily citrus. About 358,000 ha area in Florida is occupied by fruit crops most of which is citrus. Microirrigation provides better freeze protection and faster growth of citrus trees. Over 11% of 1.9 mha of commercial agricultural crops is microirrigated, of which 94% is in fruit crops, primarily citrus.

Among field crops, microirrigation has been extended to sugarcane and tobacco; 90% of Florida sugarcane is grown in muck soils, where subirrigation is practiced. Only 10% of sugarcane is in sandy soil which has the potential for microirrigation.

Vegetable Crops

Vegetable crops are grown in over 140000 ha in Florida, most of which is irrigated. About 2100 ha of strawberries grown in Florida are drip irrigated because of production benefits.

Per Chieng and Gulik (13), of 940,000 ha of land under irrigation in Canada, about 4700 ha were under microirrigation, and the trend is increasing. Most of the large scale microirrigation systems were in the province of British Columbia. In most of these areas, chemigation was being practiced, and the systems were used for tree fruit, field vegetables, greenhouse, flowers, and vines

Drip Irrigation in some indigenous form has been practiced in India for a long time. Perforated earthenware pipes have been used for quite some time in different parts of the country such as Maharashtra and Rajasthan. Similarly, perforated bamboo pipes and pitchers have been in use. In Meghalaya, they have used bamboo drip irrigation systems for betel pepper and areca nut crops by using low flows from hill streams. Drip irrigation in modern form was introduced in the early 1970s in India. Significant development has taken place in the 1980s. Per some estimates, from a small coverage of 1500 ha in 1985, the area increased to 6000 ha in 1988, 66,000 ha in 1993, and 260,000 in 2000.

The adoption of drip irrigation has been maximum in acute water scarce areas of Maharashtra state for commercial and horticultural crops such as coconut, grapes, banana, fruit trees, sugarcane, and plantation crops. About 85% of the installed drip systems in India has taken place in Maharashtra. The states of Karnataka, Tamil Nadu, Gujarat, and Andhra Pradesh have been gradually picking up.

According to Jain (14) there were 0.38 mha of land under microirrigation, making India the second largest in the world next only to the United States which has 1.055 mha.

Extent of Application to Different Crops and Trends

Per the survey, microirrigation has been applied differently to different crops, according to Bucks (10). It has been applied most to tree crops, covering 41.5%. Among fruits, deciduous types constitute the second largest coverage having a share of 16.7%. Citrus fruits constitute the third next highest, having a coverage of 13% of all the area covered. Olives/nuts are another crop having a share of 5.1%. Among vines, it has been applied most to grapes having a share of 11.1%.

The other tree crops under major application of microirrigation consist of avocado, mango, olive, nut tree, and coconut. There has been quite an increase of microirrigation in vines and vegetables which constitute 26% of the area under microirrigation. Application of microirrigation was done for cash crops like sugarcane in Hawaii on a large scale on about 42,000 ha in 1991. Coconut occupied an area of 37,000 ha. Among vegetables, it has a maximum coverage under field vegetables.

Greenhouses are expensive, so it is understandable that the area covered under such a system would be less than that of field crops. Microirrigation has been applied to a large number of unspecified crops. The area under this category in the 1991 survey was 21.5% of the total area of 1.8 mha.

In the trend of component use, it was observed that bubblers and microsprinklers were being installed more for tree crops than individual drippers. Line source applications were being used more than individual drippers. Pressure compensating drippers were increasing in use in some countries but were generally creating problems after 2 to 5 years. Fertigation was increasing in many countries such as the United States, Spain, Australia, Israel, South Africa, and Mexico. In other countries, it was still in the initial stages. Emitter clogging was generally caused by sediments, precipitated salts, or bacterial slimes. These problems were gradually solved with bubblers and microsprinklers as well as by better fabrication. To control clogging, generally, acids and chlorine were applied. The application efficiencies ranged from 60–90%.

The costs of installation in the third survey ranged from US $ 2000 to US $ 7000 per ha. It was generally found that line sources (drip tapes, etc.) used in vegetable and field crops were more expensive than drippers or microsprinklers used for tree crops. Systems with average costs from $ 2000 to $ 4000 were economically and efficiently operated for trees, vegetables, and field crops.

Extent of Subsurface Drip Irrigation

Although economical and highly efficient, the design and laying of subsurface drip is more sophisticated and complex, and its installation fairly soil and crop specific. Thus, in spite of its being promising, its application is still in its infancy. Per the ICID third survey, it had a coverage of 54,000 ha in the United States, 2500 ha in China, 150 ha in Israel, and 2184 ha in other countries, a global coverage of 58,834 ha. According to Phene et al. (15) reporting figures from the Irrigation Journal's 1992, irrigation survey out of a total of 690.2 kha of microirrigation coverage in the United States about 76.6 kha or 11.1% was under SDI. In California SDI had grown to 61.3 kha in 1992. This constituted about 80% of all SDI in the United States. The different crops to which it has been applied were potato, tomato, strawberry, cantaloupe, lettuce, cotton, grape, hops, apples, peach, almond, walnut, turf, and ornamentals.

LIMITATIONS, RESEARCH TRENDS, SCOPE, AND PROMOTION

Limitations

Emitter Clogging. This is one of the most important problems in the use of microirrigation. This may occur due to sand, rust, microorganisms or other impurities carried by irrigation water or because of chemicals precipitating in the flow system. Root penetration also reportedly causes clogging, which occurs mostly in buried systems. Proper filtration may remove suspended particles. Suspended organic matter and algae, etc. can be removed by media filters, whereas larger suspended materials can be removed by centrifugal separators. Appropriate chemicals can eliminate clogging due to precipitation of salts in the system.

Mechanical Damage. This may be caused by rodents, birds, animals, or careless use of implements. Chemicals are now available, which if mixed with plastics, discourage or prevent damage by rodents. With subsurface drip, there is minimum cultivation, and damage due to handling or implements can be avoided.

Salinity Hazard. With the use of brackish water for irrigation, salts tend to accumulate at the wetting front. The problems are caused in arid regions if irrigation methods other than drip are used and salts cannot be leached. The problem can be reduced by using sprinkler or flood methods of irrigation to leach the salts wherever feasible.

Lack of Microclimatic Control. This method cannot commonly provide frost protection as done by sprinklers, but a variation of microirrigation, known as microsprinklering, also provides quite effective frost protection.

Higher Costs. Compared to surface and sprinkler irrigation systems, installation costs are higher. Thus, it is feasible only for appropriate field crops and fruits, where returns are cost-effective. In some developing countries, to promote economy in using irrigation water, subsidies have been given for its installation; that has given good results in extending its adoption.

Lack of Awareness. This is a system requiring high skill in design, installation, operation, and management; it requires proper demonstration and training to extend its use to farmers in developing countries.

Research Trends

The issues and type of research in different countries vary according to need. However in the United States, which has the largest area in microirrigation, per Phene (16), research trends center around four areas: (1) determination of water requirements and irrigation scheduling, (2) nutrient management/fertigation methods and water quality, (3) system design and uniformity, and (4) expert systems and simulation models. The area of water requirement is considered important in view of the difficulty of accurate determination of water requirements and irrigation scheduling. The efforts related to this area were directed toward use of personal computers for data acquisition and control systems and electronic sensors for automation. Efforts were also directed to using plant measurements for scheduling irrigation and regional and local automatic scheduling, measuring, and calculating crop evapotranspiration as well as use of simulation models. Nutrient management was considered important in view of high nitrate-nitrogen pollution because of surface application and the scope of its reduction through fertigation. Because of improvements in materials hardware availability, quality, and manufacturing uniformity, there is a good potential for improvement in designing microirrigation systems and developing microcomputer programs. Expert systems have the ability of facilitating the transfer of knowledge from researchers to users. The soil water–plant relationship is a complex phenomenon, in which, expert systems can play a useful role in irrigation selection, scheduling, and management

Future Scope

It may be observed that the world food requirement is likely to increase as population increases. On the other hand, water availability for food production is gradually shrinking. Water scarcity has become the single biggest threat to food security, human health, and natural ecosystems. The only way to match higher food production with declining water resources seems to be optimization of available water resources through efficient water management Application of microirrigation has demonstrated great savings in water and fertilizer and also increased crop yield and improvement of the quality of produce. Due to this, there has been a fast growth of this technology, and the area has increased to more than double in 5 years since 1981 and to more than 4 times in the next 5 years. Such a trend shows the great possibility of further increase in the future. According to future projections, by the year 2000, the area will increase to more than eight times its coverage in 1981.

During its period of progress, its applicability has been tried for several other fruit and field crops with encouraging results. It has shown tremendous scope for flowers and vegetables and use in greenhouses and nurseries. Lots of good work has been done in landscaping, and there is a good prospect for further expansion in this area.

Success in the use of sewage as well as saline water and poor and degraded soils has further enlarged its scope. The experiments in Abu Dhabi of growing eucalyptus and casuarinas through drip irrigation using waters of salinity of 2000 mg/L and 12,000 mg/L, respectively, for stabilising sand dunes, offers great hope of greening the world's deserts.

Subsurface drip has emerged as a new promising technology, useful in further economizing water use and reducting labor, and is likely to become commercialized in the near future.

According to Uriel (8), good prospect has been observed for further expansion of microirrigation in several developing countries. In China, there is a plan for water conservation in five provinces. There is also a plan for improvement of on-farm irrigation in an area of 2,50,000 ha with sprinkler, drip, and flood irrigation. In the desert region of Nuberia in Egypt, drip irrigation is expected to cover an area of 100,000 ha of vegetables and fruits. In Nicaragua, a world relief organization has established 100 model farms to introduce drip irrigation for vegetables. In Argentina, 20 model farms for drip irrigation have been developed for vegetables, flowers, and orchards. In Peru, small areas were laid for drip irrigation to grow asparagus and vine yards on 900 hectares. In Turkey, a large scale project is in the process of improving community irrigation of which 1200 ha is for drip irrigation. In Mozambique, citrus orchards next to a dam are planned to irrigate 210 ha. In India, microirrigation has been applied in a number of states such as Maharashtra, Gujarat, Karnataka, and Tamil Nadu. But there is still a large scope for its application in other states. In the state of Andhra Pradesh, a large project of microirrigation is on the anvil.

In developing countries, however, there may be a necessity of providing subsidies in the initial stages and carrying out training and demonstrations to improve the awareness of technical staff and farmers in the interest of economizing on water use and environmental health. Besides increasing awareness, there is a great need to increase the technical competence and skill of technical people involved in installation and maintenance. Development of a strong database and providing optimum design, installation, and maintenance are great needs in developing countries and would boost microirrigation expansion a long way.

In the developed world, microirrigation offers hope to economize, reduce, and alleviate the use of excess fertilizer and the pesticide leaching hazard in surface irrigation and subirrigation and pollution of groundwater. This would help improve water quality and environmental health.

Strategies for Promotion

The future global shortage of water is very well recognized. The agricultural sector continues to be the major user of water. As the future share of water for other uses increases, the share available for the agricultural sector is likely to be further reduced. It is, therefore, essential to stress the economic and efficient use of water through microirrigation and other improved technologies of efficient water application. The developed countries have already adopted this technology to quite a large extent, though there is still scope for further adoption and extension of the more sophisticated systems such as use of automation, SDI, etc. But it is more important for the developing countries faced with increasing population and the need for more food production. Some possible approaches to meet such objectives are identified:

1. Coordinated efforts of scientists, engineers, research institutions, and industry are needed to develop appropriate technologies for field adoption of microirrigation.
2. There is a necessity to create a database on different aspects of microirrigation for specific crops/cropping pattern in different agroclimatic zones of a region. It is specially important to match the water requirement of a crop/cropping pattern to the economic viability of a system.
3. The manufacturing industry should establish a research and development base for continued improvement of the equipment to attain better performance and economy.
4. Besides working on larger systems, the industry and researchers should also develop technologies

for smaller systems suited to farmers. Small landholders have successfully used and managed drip systems for small plots.

5. It is necessary to carry out large scale demonstrations by industry, government, and nongovernment institutions on different crops in different regions to establish economic viability and promote awareness among farmers/users.
6. For the application of fertilizers and pesticides along with microirrigation, there is a necessity of generating a database on the feasibility and compatibility of such practices with a given soil water/crop condition which is duly approved by a recognized scientific body.
7. Before a microirrigation system is made available, it is necessary to ensure that the different components are manufactured according to established or adopted standards of a country, so that the materials and manufactured products give the desired performance and life for the system.
8. Microirrigation, though providing efficient and economic use of water and improving the yield of crops, is comparatively an expensive irrigation system specially suited to fruit plantation and vegetables. Besides the technical feasibility of production for a profitable return, it is also necessary to look into infrastructural aspects, such as postharvest technology, transport, and marketing as well as pricing policy. Unless the infrastructural aspect is taken care of and transport and marketing are streamlined, fruits and vegetables, being perishable products, cannot be profitably disposed of, and this expensive system cannot be sustained.
9. In many developing countries, liquid fertilizers are not indigenous and are imported; they are not affordable. Policy decisions may be necessary to enable in-country manufacture.
10. The present cost of installation may prove high for small and marginal farmers in developing countries. It needs to be examined as a policy that keeping in mind the global shortage of water, it may be worthwhile to partly subsidize the installation cost in the initial stages. Easy disbursement of subsidies that avoid delays may be desirable.

BIBLIOGRAPHY

1. Gittinger, J.P. (1985). *Economic Analysis of Agricultural Projects*. John Hopkins Press, Baltimore, MD, pp. 29–37.
2. Smith, M. (2000). Optimising Crop Production and Crop Water Management Under Reduced Water Supply. *Proc. 6th Int. Microirrigation Congr.*, Cape Town, South Africa, p. 2.
3. FAO. (2000). *Agriculture: Towards 2015/2030:Technical Interim Report*. Rome, Italy, p. 247.
4. Nakayama, F.S. and Bucks, D.A. (1986). *Trickle Irrigation for Crop Production*. Elsevier Science, the Netherlands, p. 383.
5. House, E.B. (1920). Irrigation by Means of Underground Porous Pipe. *Colorado State Experiment Station Bull*.
6. Blass, S. (1971). Drip irrigation. In: *Drip (Trickle) and Automated Irrigation in Israel*. Proceedings of Expert panel Vol. (1)10–28. Water Commissioners Office, Ministry of Agriculture, Tel Aviv, Israel.
7. Phene, C.J., Davis, K.R., Hutmacher, R.B., and McCormick, R.L. (1987). Advantages of subsurface irrigation for processing tomatoes. *Acta Horticulturae* **200**: 101–113.
8. Uriel, O.R. (2000). Technological Development of Micro Irrigation and its Applications in Developing Countries. *Proc. 6th Int. Microirrigation Congr.*, Cape Town, South Africa, pp. 2–4.
9. INCID. (1994). *Drip Irrigation in India: Status Report by Indian National Committee of Irrigation and Drainage*, p. 175.
10. Bucks, D.A. (1993). Microirrigation worldwide usage report. *Proc. 15th ICID Congr. Irrigation Drainage*, The Hague, The Netherlands, pp. 11–33.
11. Reinders, F.R. (2000). Micro irrigation: A world overview. *Proc. 6th Int. Micro irrigation Congr.*, Cape Town, South Africa, p. 2.
12. Smajstrla, A.J., Boggess, W.G., Boman, B.J., Clark, G.A., Haman, D.Z., Knox, G.W., Locascio, S.J., Obreza, T.Z., Parsons, L.R., Rhodes, F.M., Yeager, T., and Zazueta, F.S. (1985). Status and growth of microirrigation in Florida. *Proc. 5th Int. Microirrigation Congr.*, Orlando, FL, pp. 325–330.
13. Chieng, S. and van der Gulik, T. (1993). Microirrigation and chemigation: The British Columbia (Canada) experience. The workshop on micro irrigation worldwide. *Proc. 15th ICID Congr. Irrigation Drainage*, The Hague, the Netherlands Dec. 2, 1993, pp. 169–177.
14. Jain, R.B. (2000). Development of microirrigation in India. *Proc. Int. Conf. Micro and Sprinkler Irrigation*, Organized by Central Board of Irrigation and Power, Jalgaon, India, pp. 116–124.
15. Phene, C.J., Hutmacher, R.B., and Ayars, J.E. (1993). Subsurface drip irrigation in the USA. *Proc. Workshop Sprinkler and Drip Irrig. Syst.*, Organized by Central Board of Irrigation and Power, Dec 2–6, New Delhi, India, pp. 157–162.
16. Phene, C.J. (1995). Research trends in microirrigation. *Proc. 5th Int. Microirrigation Congr.*, April 2–6, Orlando, FL, pp. 6–24.

PLANT AND MICROORGANISM SELECTION FOR PHYTOREMEDIATION OF HYDROCARBONS AND METALS

T.C. McIntyre
(from *Phytoremediation: Transformation and Control of Contaminants*, Wiley 2003)

SUMMARY OF PRACTICAL IMPLICATIONS

Information on plants and other organisms that have the potential for phytoremediation of contaminants aids in the planning and establishment of removal and restoration projects for contaminated sites. Environment Canada is compiling databases on plants, bacteria, fungi, and other organisms that remediate petroleum hydrocarbons and accumulate metals and metalloids. The databases PHYTOPET (for PHYTOremediation of

PETroleum hydrocarbon contaminants) and PHYTOREM (for PHYTOREmediation of Metals) are user-friendly, searchable, and summarize all the known literature. Environment Canada has also drafted a Protocol, "Environmental Technology Verification (ETV) Protocol for Metal Contaminants." The program provides suppliers of specified environmental technologies and services that meet the required criteria, with the right to use the licensed ETV verification mark.

PHYTOREMEDIATION APPLICATIONS

Phytoremediation has good potential as a flexible and cost-effective means to clean up contaminated sites. Two of the most important applications are the use of plants and other organisms to degrade petroleum hydrocarbons and accumulate metals and metalloids (hereafter the term metals covers both). Plants and other organisms extract and degrade contaminants using a variety of processes such as:

1. Phytoextraction—the capability of terrestrial plants, algae, and plant tissues to sequester toxic elements, especially metals, by uptake or biosorption. Contaminated organisms can be harvested for disposal or extraction (biomining when at commercial levels) (1–6).
2. Rhizofiltration—the use of plant roots to absorb and sometimes precipitate contaminants from polluted waters (e.g., 7).
3. Phytostabilization—the use tolerant plants to stabilize contaminants by reducing bioavailability.
4. Phytodegradation—the use of plants and associated microorganisms to degrade organic pollutants.
5. Phytovolatilization—the use of plants to volatilize pollutants.

Information on useful phytoremediation species is limited and scattered in the literature. Environment Canada has organized information into two databases called PHYTOPET and PHYTOREM. The PHYTOPET database is a compilation and global inventory of organisms, with a focus on plants, for phytoremediation of petroleum hydrocarbons. PHYTOREM is a similar database for metals and metalloids.

Plants and other organisms have a remarkable capacity to accumulate contaminants, e.g., the common sunflower (*Helianthus annuus*) (8), Indian mustard (*Brassica juncea*) (9), alpine pennycress (*Thlaspi caerulescens*) (10,11), and brake fern (*Pteris vittata*) (12). Some of the highest accumulation of metals occurs in the New Caledonia tree (*Sebertia acuminata*) (13) and the Australian shrub (*Hybanthus floribundus*) [reported by Streit and Stumm in 1993 (14) as first reported by Harborne in 1988 (15)]. The purpose of this chapter is to describe the compilation of data for plants and other organisms that degrade hydrocarbons, or accumulate, hyperaccumulate, or tolerate metals.

PLANTS THAT TOLERATE AND SPUR DEGRADATION OF TOTAL PETROLEUM HYDROCARBONS

The database PHYTOPET specializes in plant species that spur the rhizodegradation of petroleum hydrocarbons in terrestrial and aquatic environments, including wetlands. The plants (both monocots and dicots), bacteria, protozoa, and fungi covered by the database show potential for phytoremediation of a wide range of hydrocarbons—from crude oil to polyaromatic hydrocarbons, such as chrysene and benzo(*a*)pyrene (Hutchinson et al. this book, Olson et al. this book). In the database, special attention has been paid to the Canadian plants from the oil-producing regions in the Western Prairie and Boreal ecozones for reclamation of oil-contaminated sites. The relationships between plants, microorganisms, and other species that can phytoremediate similar chemicals or mixtures of contaminants (metals, pesticides, and hydrocarbons) have also been considered.

The database uses many of the same fields as the PHYTOREM database for plants (exchanging information on the metal elements for information on the hydrocarbons). In addition, the database compiles information on plant-associated microorganisms. The information includes concentrations of the contaminant before and after treatment, length of treatment period, soil characteristics, age of the plant when exposed to the contaminant, and requirements for phytoremediation. Data in PHYTOPET also include tolerance to salinity and Western Canadian occurrence.

Many native grasses show promise to phytoremediate hydrocarbons [e.g., *Agropyron* spp., gramas (*Bouteloua* spp.), and buffalo grasses (*Buchloe* spp.)]. Poplar (*Populus* spp.) trees also are potentially useful. Cultivated plants with phytoremediation potential include carrot (*Daucus carota*), red fescue (*Festuca rubra*), alfalfa (*Medicago sativa*), and ryegrass (*Lolium* spp.).

PLANTS THAT ACCUMULATE METALS

Scope and Organization of the PHYTOREM Database

The global inventory of PHYTOREM covers both terrestrial and aquatic plants and other organisms (such as bryophytes, lichen, fungi, algae, and bacteria) that have potential value for phytoremediation of metals. Species included in the database have an ability or potential to tolerate, accumulate, or hyperaccumulate specific metals, or be useful as sorbents for metals.

An intensive literature search was conducted for the database, for accumulation of or tolerance to the 19 selected elements listed in Table 1. Reviews on metal accumulation provided global data (16–23). Although a large number of species are known to grow over serpentine substrates, only those specifically shown to be hyperaccumulators or to accumulate substantial levels of heavy metals are included in the database.

The database contains information on organisms that are both hyperaccumulators and accumulators (Table 2). The term hyperaccumulator was originally coined for

Table 1. Metals and Metalloids Scanned for the Development of the PHYTOREM Database

Aluminum (Al)	Cobalt (Co)	Mercury (Hg)	Radium (Ra)
Arsenic (As)	Copper (Cu)	Molybdenum (Mo)	Strontium (Sr)
Beryllium (Be)	Chromium (Cr)	Nickel (Ni)	Uranium (U)
Cadmium (Cd)	Lead (Pb)	Palladium (Pd)	Zinc (Zn)
Cesium (Cs)	Manganese (Mn)	Platinum (Pt)	

plants with greater than 1000 milligrams per kilogram (0.1 percent) nickel in dried aboveground tissues (34,35). This definition has been broadened to other elements and all parts of plants. Values for accumulation were based on either literature, or a value of 100 or 200 milligrams per kilogram of dry weight was assigned where none has been specified.

Information in the database includes geographical origin of the organisms and habitat characteristics; taxonomy; environmental effects, health effects, and uses of the species; cultivation practices; sources of material or species studied; weedy or other significant relatives; and mode of action (i.e., tolerance and accumulation). See Table 3 for details of criteria and fields. The database is worldwide in scope, but focuses on species from Canada, the U.S., and high-altitude species from warmer regions.

The database, compiled using *Microsoft Access 97®* software, allows for easy access by many database formats and programs and can be updated as needed. The database consists of three organism and two reference tables (Fig. 1a). The organism tables are (1) vascular plants; (2) algae, lichens, fungi, and mosses; and (3) bacteria. The main table of reference includes all citations that relate to concentration levels of elements accumulated or tolerated, and the second is a table of general botanical and reference works consulted. The database as designed can be searched by queries (Fig. 1b) and includes three data entry forms (botany, metal, and plant) and reporting formats for organisms and reference records.

Data sets can be quickly retrieved by applying appropriate search criteria, and viewed in tabular format. An example of a data screen for a species retrieved through the filtering process is given in Fig. 2.

The database has specific search criteria, some of which are as follows:

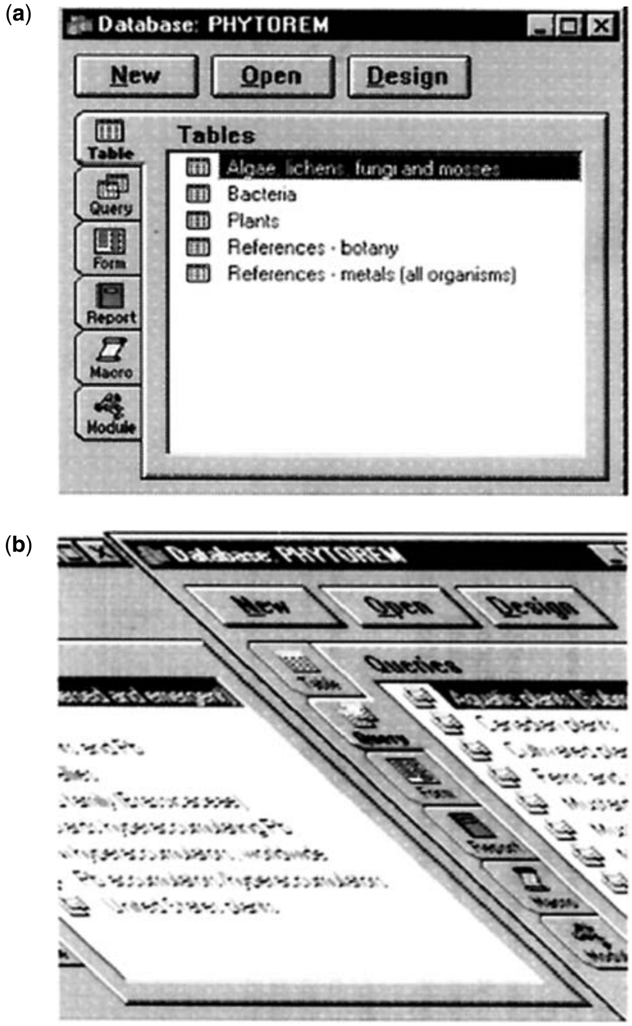

Figure 1. PHYTOREM database windows showing contents from the tables (**a**) and query (**b**) menus.

1. To access the elements radium and uranium, use the codes Rd and Ur respectively. Searches for accumulators or hyperaccumulators, however, use commonly used short forms for these elements. For example, the character code RaA will bring up accumulation data for radium.
2. To retrieve all records of species accumulating (A), hyperaccumulating (H), tolerating (T), or precipitating (P) an element, use the symbol for the element surrounded by asterisk in the elem_action field, e.g., "*Pb*" as the query or filter criterion. To retrieve species accumulating or hyperaccumulating a particular element, the search criterion should be a combination of the element and the desired accumulation code, e.g., "*PbH*" for lead hyperaccumulators.
3. There is no separate field to distinguish coniferous species; however, species can be retrieved by using the family Pinaceae or by retrieving all records with a growth form of TR for tree.
4. Species are reported as rhizofiltration agents for heavy metals. Studies dealing with nutrient culture experiments, although not aimed at identifying such agents, have been recorded with a Y for "yes" in the field biofiltr_use, because of the potential for such use.

Table 2. Values and Sources for Hyperaccumulation and Accumulation Criteria for Inclusion in the PHYTOREM Database

Element	Hyperaccumulation Criteria (Milligrams per Kilogram of Dry Weight)	Plants, Where Appropriate, Which Were the Basis of Criteria	References for Data	Accumulation criteria (Milligrams per Kilogram of Dry Weight)	Number of Records in the Database	Species with Highest Recorded Value {origin} [Value (Milligrams Per Kilogram of Dry Weight)]
Al	1000	Barley (*Hordeum vulgare*), horse bean (*Vicia faba*)	24	100	25	Hairy goldenrod (*Solidago hispida*) {Canada} [6820]
As	1000		25	100	4	Colonial bent (*Agrostis tenuis* = *Agrostis capillaris*) {cultivation} [2000]
Cd	100		2	10	37	Eel grass (*Vallisneria spiralis*) {India} [6242]
Co	1000		17	100	27	*Haumaniastrum robertii* {Africa} [10,200]
Cr	1000	General review	17	100	35	Alfalfa (*Medicago sativa*) {cultivation} [7700]
		Dicoma niccolifera, Sutera fodina	26,27			
		Alfalfa (*Medicago sativa*)	28			
		Smooth Water hyssop (*Bacopa monnieri*) (aquatic species)	29			
		Azolla spp.	30			
		Duckweed (*Spirodela polyrhiza*), kariba weed (*Salvinia molesta*)	31			
		Water lettuce (*Pistia stratiotes*)	32			
Cs	?			?	1	Sunflower (*Helianthus annuus*) {cultivation} [high absorption rate]
Cu	1000		17	100	67	Creosote bush (*Larrea tridentata*) {U.S.} [23,700 biosorption]
Hg	1000		16,17	100	35	Water lettuce (*Pistia stratiotes*) {pantropical} [1100]
Mn	10,000		33	200	28	*Macademia neurophylla* {New Caledonia} [51,800]
Mo	1500	Alpine pennycress (*Thlaspi caerulescens*)		?	1	Alpine pennycrest (*Thlaspi caerulescens*) {Europe} [1500–1800]
Ni	1000		34	100	372	*Psychotria douarrei* {New Caledonia} [47,500]
			35			Shrub violet (*Hybanthus floribundus*) {Australia} [2% total ash content]
Pb	1000		17	200	79	Indian mustard (*Brassica juncea*) {cultivation} [26,200]
Sr	?			?	1	Sunflower (*Helianthus annuus*) {cultivation} [high absorption Rate]
Ur	?			?	3	Sunflower (*Helianthus annuus*) {cultivation} [> 15,000]
Zn	10,000		17	500	48	Alpine pennycrest (*Thlaspi caerulescens*) {Europe} [52,000]

? indicates that criteria have not been defined for hyperaccumulation or accumulation of these elements.

Note: No reports were found for accumulation of beryllium, palladium, platinum, and radium.

Table 3. Fields, Description, and Access for PHYTOREM Database

Name	Description [and explanatory comments]	Plants	Bacteria	Algae	Access
Entry_seq	Automatic counter field to record sequence of data entry [allows multiple records taken from a single summary document to be grouped together for purposes of review and data management]	√	√	√	
Type_org	Vascular plant = VP, bacteria = BA, algae = AG, lichens = LI, fungi = FU, and bryophytes = BR [field allows tables of different organisms to be combined and provides the ability to distinguish records by type of organism]	√	√	√	
Growth_form	For plants only: fern = FE, graminoid = GR, herb = HE, shrub = SH, succulent = SU, vine = V, and tree = T	√			
Sci_name	Complete scientific name with authorities	√	√	√	√
Synonym	Common synonym	√		√	
Com_name	Common English name(s)	√	√	√	
Cv_strain	Cultivar or strain name, or code; also transgenic variants	√	√	√	
Family	Taxonomic category: family	√		√	√
Order	Taxonomic category: order	√		√	√
Subclass	Taxonomic category: subclass	√		√	√
Class	Taxonomic category: class	√		√	√
Duration	For plants only: A(nnual), B(iennial), A/B, P(erennial), A/P, and B/P	√			
Origin	Country or region of origin of plants on which report(s) is based [Canadian and U.S. species or those of other continents or countries can be sorted]	√	√	√	√
World_range	World range of the species	√			
Primary_habitat	T = terrestrial, A = aquatic, T/A = terrestrial and aquatic [provides the ability to distinguish between terrestrial, aquatic, marsh, and wetland species (T/A)]	√		√	√
Hab_descr	Habitat description [allows for sorting to distinguish plants from serpentine soils and to distinguish cultivated species (crop plants)]	√			
Namer_occ	North American occurrence: N(ative); E(xotic); X (not present); ? (status unknown) [allows North American species to be categorized]	√			
Namer_spp	Indication of whether other species in the genus are present in North America	√			
Sig_relatives	Other species of significance in the genus	√			
Cult_wild	For plants: crop plant = C, horticultural species = H, wild = W. For algae and other: cultured = C or wild = W	√		√	√
Cult_info	Propagation and/or test studies, or experimental conditions	√		√	
Cult_source	Source of cultivated material as indicated in publication	√	√	√	
Impact_attributes	UN = unknown, WD = weed, NX = noxious properties, HY = hybridizes, and PH = disease and insect pest host	√			
Impact_description	Documentation of potential impact [information on invasive potential of an exotic weed]	√			
Uses	General uses of plant and citations [medicinal or edible]	√			
Bio_notes	For plants: notes on such topics as pollination and dispersal mechanism. For bacteria: notes on important and interesting aspects of biology.	√	√		
Gen_notes	For plants and others: toxicity of metal to the plant or other organisms and other pertinent information and citations. For bacteria: general notes on procedures of the study or results and citations.	√	√	√	
Elem_conc_ref	Citation and concentration of elements with organ of storage	√	√	√	√
Elem_form_ref	Element form if specifically indicated [e.g., Cu(II)]	√	√	√	
Elem_chel_ref	Element, chelate used and citation	√	√	√	

Table 3. (*Continued*)

Name	Description [and explanatory comments]	Plants	Bacteria	Algae	Access
Elem_action	Hyperaccumulate = H, accumulate = A, tolerate = T, rhizosphere concentration = R, and precipitate = P [e.g., PbH is lead hyperaccumulation]	√	√	√	√
Storage_sites	For plants: sites where element concentrations were measured [e.g., root, shoot, or leaf]. For bacteria: sites where element measured [e.g., cell wall].	√	√		
Biofiltr_use	Use for biofiltration: yes (Y) or no (N) [used to distinguish species useful for such a function as indicated in the publication]	√	√	√	
Biofiltr_ref	References to the biofiltration publications [generally the same as the primary reference]	√	√	√	
Tolerance_info	Information on conditions under which tolerance occurs	√		√	
Summary_ref	Reference to summary papers where several to many species are listed	√	√	√	√
Primary_ref	Primary literature reference [obtained from summary reference if summary document was used]	√	√	√	√

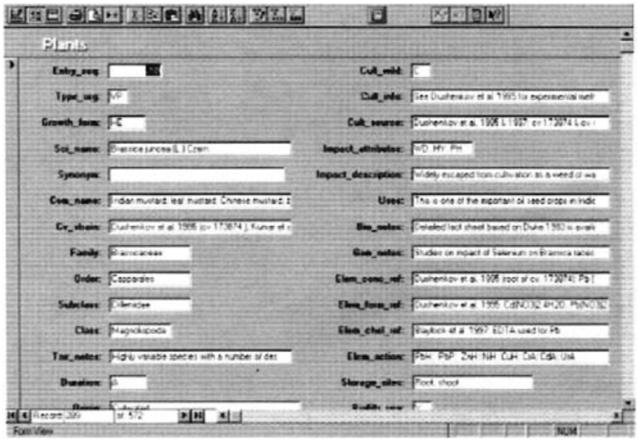

Entry_seq	[illegible]	Cult_wild	C
Type_sig	VP	Cult_info	See Dushenkov et al. 1995 for experimental meth
Growth_form	[illegible]	Cult_source	Dushenkov et al. 1995 & 1997; cv 173874 [illegible]
Sci_name	Brassica juncea (L.) Czern	Impact_attributes	WD; HY; PH
Synonyms		Impact_description	Widely escaped from cultivation as a weed of wa
Com_name	Indian mustard; leaf mustard; Chinese mustard;	Uses	This is one of the important oil seed crops in India
Cv_strain	Dushenkov et al. 1995 (cv 173874); Kumar et	Bio_notes	Detailed fact sheet based on Duke 1983 is avail
Family	Brassicaceae	Gen_notes	Studies on impact of Selenium on Brassica [illegible]
Order	Capparales	Elem_conc_ref	Dushenkov et al. 1995 (root of cv. 173874); Pb (
Subclass	Dilleniidae	Elem_form_ref	Dushenkov et al. 1995: Cd(NO3)2 4H2O; Pb(NO3)2
Class	Magnoliopsida	Elem_chel_ref	Blaylock et al. 1997 EDTA used for Pb
Tax_notes	Highly variable species with a number of dis	Elem_action	PbH; PbP; ZnH; NiH; CuH; CrA; CdA; UA
Duration	A	Storage_sites	Root; shoot

Record: 389 of 572

Figure 2. Example of a plant report form from a search of the PHYTOREM database.

Summary of the Database Contents

The database provides an extensive, though preliminary, collection of records on phytoremediation potential. At present, the database contains records of 775 species spread over 76 families, 39 orders, 9 subclasses, and 2 classes (dicots and monocots). Data come from 39 countries of origin (based on names used in the publications).

The greatest proportion (465) of species recorded in the plant table accumulate, hyperaccumulate, or tolerate a single metal of the 19 elements surveyed. Some species take up two elements (66) or three (25), with a balance of 15 species capable of accumulating four or more elements. Table 4 summarizes the species in the latter category with elements and degree of accumulation.

The highest number of records (260) was found for nickel followed by lead and copper (Fig. 3). No records were found of species accumulating beryllium, palladium, platinum, or radium. Species with the highest concentration of the elements recorded in the database (listed in the right-hand column of Table 2) provide some interesting possibilities for phytoremediation, including several commonly cultivated species such as sunflower (*Helianthus annuus*) and alfalfa (*Medicago sativa*).

Data Update Recommendations

Updates that could be considered include (1) information from a more extensive literature search, (2) fields for data on storage sites (e.g., roots), (3) mechanisms of tolerance or accumulation of metals by plants, (4) species that remediate radionuclides, (5) plants that remediate organic and other inorganic compounds and stabilize contaminated sites, and (6) expansion of data for algae, in particular those with a high absorptive capacity. In addition, the database could also provide pictures of plants and sites.

INORGANIC PHYTOREMEDIATION VERIFICATION PROTOCOL

Environment Canada is formulating a protocol to establish the limits of applicability for the phytoremediation of metals and metalloids in contaminated soil and water that will aid economic adoption and sustainable use. The pollutants (no organic contaminants to this point) are arsenic (As), cadmium (Cd), chromium (Cr), copper (Cu), lead (Pb), mercury (Hg), nickel (Ni), selenium (Se), silver (Ag), and zinc (Zn). Use of the verification protocol will allow the use of a Canadian Environmental Technology Verification symbol in remediation products. The applicant will be required to establish that a particular application has been tested (and under what conditions) (e.g., proof of principle at the bench, greenhouse, plot, or full-scale testing). Only applications that involve phytoextraction, rhizofiltration, phytovolatilization, phytostabilization, and combinations of these physiological processes will be certified initially. All phytoremediation applications (or technologies) will be evaluated based on the following criteria and information:

1. Total amount, concentration, and speciation of elements in contaminated soil, rhizosphere soil water,

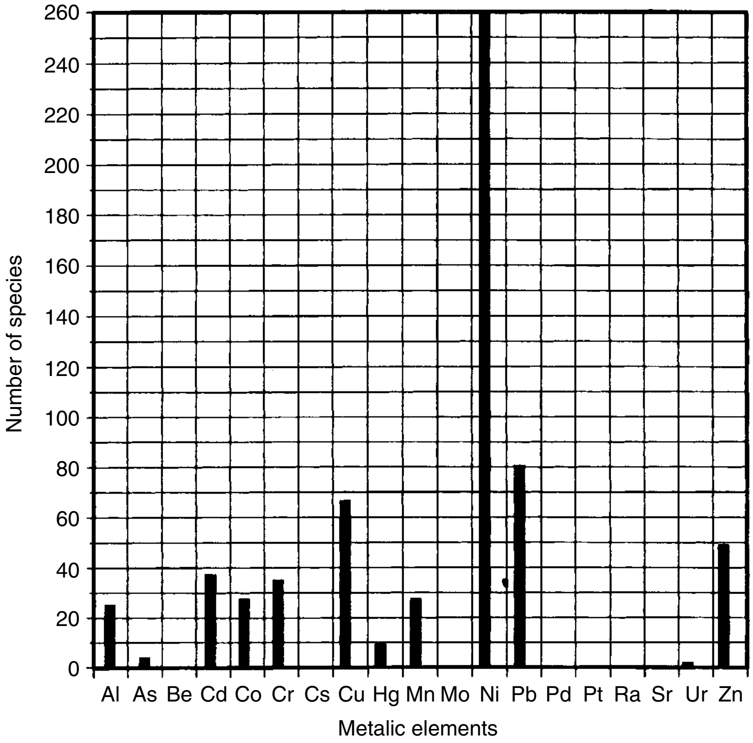

Figure 3. Number of species records for the 19 elements scanned.

Table 4. Plants Capable of Accumulating Four or More Metals

Scientific Name	Common Name	Origin and Characteristics	Elements and Degree of Accumulation[a]
Azolla filiculoides	Water fern	Africa, floating	CuA, NiA, PbA, and MnA
Bacopa monnieri	Water hyssop	India, emergent species	HgA, CuH, CrH, PbA, and CdH
Eichhornia crassipes	Water hyacinth	Pantropical/subtropical, troublesome weed	CdH, CrA, ZnA, HgH, PbH, and CuA
Hydrilla verticillata	Hydrilla	Southern Asia but introduced and spreading as a troublesome weed in the warmer states of the USA	CdH, CrA, HgH, and PbH
Lemna minor	Duckweed	Native to North America and widespread	PbH, CdH, CuH, and ZnA
Pistia stratiotes	Water lettuce	Pantropical and native to southern U.S. but an aquatic weed	CuT, CdT, HgH, and CrH
Salvinia molesta	Water fern	India	CrH, NiH, PbH, and ZnA
Spirodela polyrhiza	Giant duckweed	Native to North America	CdH, NiH, CrH, PbH, and ZnA
Vallisneria americana	Tape grass	Native to Europe and North Africa but widely cultivated in the aquarium trade	CuH, CdH, CrA, and PbH
Brassica juncea	Indian mustard	Cultivated	PbH, PbP, ZnH, NiH, CuH, CrA, CdA, and UrA
Helianthus annuus	Sunflower	Cultivated	PbH, UrH, SrH, CsH, CrA, CdA, CuA, MnA, NiA, and ZnA
Agrostis castellana	Bent grass	Portugal	AsH, PbA, ZnA, MnA, and AlA
Thlaspi caerulescens	Alpine pennycress	Europe	ZnH, CdH, CoH, CuH, NiH, PbH, and CrA
Athyrium yokoscense	Fern	Japan	CuH, CdA, ZnH, and PbH

[a] H: hyperaccumulator; A: accumulator; P: precipitator; and T: tolerant.

water below the root zone, groundwater, wastewater, roots, shoots, and air (where phytovolatilization is a factor) during and after treatment

2. Amount and timing of supplements added to treated soils and waters as related to mobilization or immobilization of contaminants, including co-contaminants, and expected effects
3. Seepage or infiltration volumes from treatment wetlands or soil plots and concentrations of contaminants and by-products for the range of expected moisture conditions
4. Range of meteorological conditions for effective phytoremediation application including, but not limited to, seasonal precipitation, air temperatures, humidity, wind speed and direction, and solar radiation
5. Time required to achieve cleanup or risk-reduction standards for various levels of initial contamination in soil or water, plants selected, soil or sediment conditions, and weather
6. Soil, sediment, or water quality during and after the treatment, including texture, mineralogy, moisture content, cation-exchange capacity for soils, organic carbon and nutrient content, pH, and redox potential for all media
7. List of plants and description of communities and any rooting depth, growing habits, and foliage requirements for effective cleanup
8. Unusual ecological conditions to avoid including endangered and threatened plants and wildlife, weeds, invasive species, bioaccumulation (in the absence of netting, fencing, motion detectors, or alarms), and wetland changes (e.g., drying up of seepage wetlands)
9. Risk-reduction or cleanup levels achieved in soil, groundwater, and air for different land uses
10. Residual management options (e.g., land filling in hazardous waste landfills, composting, and pyrolysis) and any risks during handling, transport, and disposal of residuals
11. Closure plans and options for vegetative cover

These criteria will be focused on specific issues for terrestrial and aquatic remediation. Criteria and information requirements will be weighted, depending on the application. This protocol is different from that used in the U.S. (36,37); however, the differences are expected to diminish over time with continued collaborative projects between Environment Canada and the U.S. Environmental Protection Agency.

CURRENT AND FUTURE EXPECTATIONS

Plant-based remediation and restoration applications are expected to receive increasing attention in Canada and elsewhere as options for removal of the more recalcitrant inorganic and organic constituents at contaminated sites. These Environment Canada databases and protocol could provide valuable assistance in the pre-selection of organisms, help clarify the regulatory responsibilities of proponents in securing necessary approvals prior to the use of specific organisms, and provide valuable insight into genetic materials to adapt or modify representative Canadian plant species to the climatic circumstances unique to specific sites. Future initiatives include (1) database expansion to include one for radioactive materials (PHYTORAD); (2) a Canadian network to identify, store, protect, propagate, and possibly distribute germplasm of candidate plant cultivars; (3) protocols that outline responsible provisions; (4) methods to recover inorganic contaminants from contaminated biomass; (5) a complete overview of the regulatory issues inherent in the use of plants as remediation and restoration agents; (6) definition of bioavailability and the role of plants in natural attenuation; and (7) establishment of mechanisms to limit phytoremediation site access by herbivores and omnivores.

Acknowledgments and Disclaimer. The author is extremely grateful for the opportunity to realize the conceptual development of PHYTOREM and PHYTOPET databases through the funding of the Canadian Biotechnology Strategy and the Panel on Energy Research and Development. Data were compiled and provided by Dr. Erich Haber of National Botanical Services in Canada for PHYTOREM and by Dr. J.J. Germida, Dr. R.E. Farrell, and Dr. C.M. Frick of the University of Saskatchewan for PHYTOPET. Environment Canada headquarter's library staff, Jean-François Belanger and Marie Jetten and the Canadian Museum of Nature librarian Mireille Boudreau were invaluable. Finally the author extends an incredible debt of gratitude to Dr. Marcos Alvarez of Environment Canada for his editorial support. Copies of the PHYTOREM and PHYTOPET databases are available on CD-ROM from the author at terry.mcintyre@ec.gc.ca.

Disclaimer. Environment Canada does not endorse the use of, nor does it attempt to determine the merits of any specific technology or technology provider through publication in this document, nor does Environment Canada assume any liabilities with respect to the use of any information, apparatus, method, or process discussed in this document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. Further, the views expressed in the document remain solely those of the authors.

BIBLIOGRAPHY

1. Wehrheim, B. and Wettern, M. (1994). Biosorption of cadmium, copper and lead by isolated mother cell walls and whole cells of *Chlorella fusca*. *Appl. Microbio. Biotechnol.* **41**: 725–728.
2. Brown, S.L., Chaney, R.L., Angle, J.S., and Baker, A.J.M. (1995). Zinc and cadmium uptake by hyperaccumulator *Thlaspi caerulescens* and metal tolerant *Silene vulgaris* grown on sludge-amended soils. *Environ. Sci. Technol.* **29**: 1581–1585.
3. Kim, Y.H., Yoo, Y.J., and Lee, H.Y. (1995). Characteristics of lead adsorption. *Biotechnol. Lett.* **17**: 345–350.
4. Wang, T.C., Weissman, J.C., Ramesh, G., Varadarajan, R., and Benemann, J.R. (1995). Bioremoval of toxic elements

with aquatic plants and algae. In: *Bioremediation of Inorganics*. R.E. Hinchee, J.L. Means, and D.R. Burris (Eds.). Third International *In Situ* and On-Site Bioreclamation Symposium. Battelle Press, Columbus, Ohio. Held in San Diego, California, April 24–27, pp. 65–69.

5. Wilson, M.W. and Edyvean, R.G. (1995). Biosorption for the removal of heavy metals from industrial wastewaters. In: *Biodeterioration and Biodegradation* 9. A. Bousher and M. Chandra (Eds.). Proceedings of the 9th International Biodeterioration and Biodegradation Symposium. Leeds, UK, Sept. 1993, pp. 47–51.
6. Gardea-Torresdey, J.L., Bibb, J., Tiemann, K.J., Gonzalez, J.H., and Arenas, J. (1996a). Adsorption of copper ions from solution by heavy metal stressed *Larrea tridentata* (creosote bush) biomass. In: *Proceedings of the HSRC/WERC Joint Conference on the Environment*. L.E. Erickson, D.L. Tillison, S.C. Grant, and J.P. McDonald (Eds.). Great Plains/Rocky Mountain Hazardous Substance Research Center, Kansas State University, Manhattan. Held in Albuquerque, New Mexico, May 21–23. (http://www.engg.ksu.edu/HSRC/96proceed/gardeal.html).
7. Dushenkov, V., Kumar, N.P.B.A., Motto, H., and Raskin, I. (1995). Rhizofiltration: the use of plants to remove heavy metals from aqueous streams. *Environ. Sci. Technol.* **29**: 1239–1245.
8. Salt, D.E., Blaylock, M., Kumar, N.P.B.A., Dushenkov, V., Chet, B.D. Ensley, I., and Raskin, I. (1995). Phytoremediation: a novel strategy for the removal of toxic metals from the environment using plants. *Biotechnology* **13**: 468–474.
9. Watanabe, M.E. (1997). Phytoremediation on the brink of commercialization. *Environ. Sci. Technol.* **31**: 182–186.
10. Baker, A.J.M., Reeves, R.D., and Hajar, A.S.M. (1994). *Heavy metal accumulation and tolerance in British populations of the metallophyte Thlaspi caerulescens J. & C. Presl (Brassicaceae) New Phytol*. **127**: 61–68.
11. Comis, D. (1995). Green remediation: using plants to clean the soil. *J. Soil Water Conserv.* May–June: 184–187.
12. Ma, L.Q., Komar, K.M., and Tu, C. (2001). A fern that accumulates arsenic. *Nature* **409**: p. 579.
13. Brown, K.S. (1995). The green clean. *BioScience* **45**: 579–582.
14. Streit, B. and Stumm, W. (1993). Chemical properties of metals and the process of bioaccumulation in terrestrial plants. In: *Plants as Biomonitors: Indicators for Heavy Metals in the Terrestrial Environment*. B. Markert (Ed.). VCH Publishing Group (Wiley-VCH Verlag), New York, pp. 31–62.
15. Harborne, J.B. (1988) *Introduction to Ecological Biochemistry*, 3rd Edn. Academic Press, London, 356 pp.
16. Atri, F.R. (1983). *Schwermetalle und Wasserpflanzen*. Gustav Fischer Verlag, Stuttgart, Germany.
17. Baker, A.J.M. and Brooks, R.R. (1989). Terrestrial higher plants which hyperaccumulate metallic elements—a review of their distribution, ecology and phytochemistry. *Biorecovery* **1**: 81–126.
18. Shaw, A.J. (Ed.). (1990). *Heavy Metal Tolerance in Plants: Evolutionary Aspects*. CRC Press, Boca Raton, Florida.
19. Baker, A.J.M., Proctor, J., and Reeves, R.D. (1991). The vegetation of ultramafic (serpentine) soils. In: *Proceedings of the First International Conference on Serpentine Ecology*. A.J.M. Baker, J. Proctor, and R.D. Reeves (Eds.). Intercept Ltd., Andover, United Kingdom. Held at University of California, Davis, June 19–22.
20. Roberts, B.A. and Proctor, J. (Eds.). (1992). *The Ecology of Areas with Serpentinized Rocks: a World View*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
21. Markert, B. (1993). *Plants as Biomonitors: Indicators for Heavy Metals in the Terrestrial Environment*. VCH Publishing Group (Wiley-VCH Verlag), New York.
22. Frankenberger, W.T., Jr. and Losi, M.E. (1995). Applications of bioremediation in the cleanup of heavy metals and metalloids. In: *Bioremediation: Science and Applications*. H.D. Skipper and R.F. Turco (Eds.). *Soil Science of America Special Publication* 43: 173–210.
23. Reeves, R.D., Baker, A.J.M., Borhidi, A., and Berazaín, R. (1996). Nickel-accumulating plants from the ancient serpentine soils of Cuba. *New Phytol.* **133**: 217–224.
24. Grauer, U.E. and Horst, W.J. (1990). Effect of pH and nitrogen source on aluminium tolerance of rye (*Secale cereale* L.) and yellow lupin (*Lupinus luteus* L.). *Plant Soil* **127**: 13–21.
25. Porter, E.K. and Peterson, P.J. (1975). Arsenic accumulation by plants on mine waste (United Kingdom). *Sci. Total Environ.* **4**: 365–371.
26. Wild, H. (1974). Indigenous plants and chromium in Rhodesia. *Kirkia* **9**: 233–241.
27. Brooks, R.R. and Yang, X.H. (1984). Elemental levels and relationships in the endemic serpentine flora of the Great Dyke, Zimbabwe, and their significance as controlling factors for this flora. *Taxon* **33**: 392–399.
28. Tiemann, K.J., Gardea-Torresdey, J.L., Gamez, G., Rodriquez, O., and Sias, S. (1997). Study of the ligands involved in metal binding to alfalfa biomass. Abstract 9. In: *Abstracts From The 1997 Conference on Hazardous Waste Research*. L.E. Erickson, M.M. Rankin, S.C. Grant, and J.P. McDonald (Eds.). Great Plains/Rocky Mountain Hazardous Substance Research Center, Kansas State University, Manhattan, Kansas. Held in Kansas City, Missouri, May 20–22, (http://www.engg.ksu.edu/HSRC/97abstracts/doc9.html).
29. Gupta, M., Sinha, S., and Chandra, P. (1994). Uptake and toxicity of metals in *Scirpus lacustris* L. and *Bacopa monnieri* L. *J. Environ. Sci. Health* **A29**: 2185–2202.
30. Priel, A. (1995). Purification of industrial wastewater with the Azolla fern. *World Water Environ. Eng.* **18**: p. 13.
31. Srivastav, R.K., Gupta, S.K., Nigam, K.D.P., and Vasudevan, P. (1994). Treatment of chromium and nickel in wastewater by using aquatic plants. *Water Res.* **28**: 1631–1638.
32. Sen, A.K., Mondal, N.G., and Mandal, S. (1987). Studies of uptake and toxic effects of Cr (VI) on *Pistia stratioides*. *Water Sci. Technol.* **19**: 119–127.
33. Baker, A.J.M. and Walker, P.L. (1990). Ecophysiology of metal uptake by tolerant plants. In: *Heavy Metal Tolerance in Plants: Evolutionary Aspects.* A.J. Shaw (Ed.). CRC Press, Boca Raton, FL, pp. 155–177.
34. Brooks, R.R., Lee, J., Reeves, R.D., and Jaffré, T. (1977). Detection of nickeliferous rocks by analysis of herbarium specimens of indicator plants. *J. Geochem. Explor.* **7**: 49–77.
35. Reeves, R.D. (1992). The hyperaccumulation of nickel by serpentine plants. In: *The Vegetation of Ultramafic (Serpentine) Soils*. A.J.M. Baker, J. Proctor, and R.D. Reeves (Eds.). Intercept Ltd., Andover, U.K., pp. 253–277.
36. Remediation Technology Demonstration Forum (RTDF) (2000). *A database of existing phytoremediation field sites* (http://www.rtdf.org/public/phyto/siteprof/usersearch/phyto_list.cfm).
37. Interstate Technology and Regulatory Council (ITRC) (2001). *Phytotechnology technical and regulatory guidance document. Phytotechnology Team*. Washington, DC. (http://www.itrcweb.org).

READING LIST

Azadpour, A. and Matthews, J.E. (1996). *Remediation of metal-contaminated sites using plants*. *Remediation* Summer: 1–18.

Black, H. (1995). Absorbing possibilities: phytoremediation. *Environ. Health Persp.* **103**: 1106–1108.

Coghlan, A. (1997). Plants gobble up waste explosives. *New Scientist* 1 February: p. 19.

Cunningham, S.D. and Berti, W.R. (1993). Remediation of contaminated soils with green plants: an overview. *In Vitro Cell. Dev. Biol.* **29P**: 207–212.

Cunningham, S.D. and Lee, C.R. (1995). Phytoremediation: plant-based remediation of contaminated soils and sediments. In: *Bioremediation: Science and Applications.* H.D. Skipper and R.F. Turco (Eds.). *Soil Science of America Special Publication* 43: 145–156.

Gardea-Torresdey, J.L., Tang, L., and Salvador, J.M. (1996). Copper adsorption by esterified and unesterified fractions of Sphagnum peat moss and its different humic substances. *J. Hazard. Mat.* **48**: 191–206.

Hartman, W.J., Jr. (1975). *An evaluation of land treatment of municipal wastewater and physical siting of facility installations*, U.S. Department of the Army, Washington, DC.

Kim, I. (1996). *Harnessing the green clean*. *Chem. Eng.* December: 39–41.

Moffat, A.S. (1995). Plants proving their worth in toxic metal cleanup. *Science* **269**: 302–303.

Parry, J. (1995). *Plants absorb heavy metals: nature provides a low-cost alternative to chemical extraction or immobilization*. *Pollut. Eng.* February: 40–41.

Rouhi, A.M. (1997). Plants to the rescue: crops with big appetites for certain elements could be key players in soil cleanup. *Chem. Eng. News* **75**(2): 21–23.

NITRATE POLLUTION PREVENTION

Brane Maticic
Ljubljana, Slovenia

Groundwater and springwater are sources used for water supply in many countries. Pollution of surface and groundwater is caused mostly by the chemicals used in industry, urban wastes, and agriculture. The pollution of **ground water** by **nitrates,** as well as pollution of **crops** that we consume, is one of the most serious environmental concerns in most countries. A primary characteristic of most agriculturally related groundwater pollution is that it is often widely distributed (nonpoint source), whereas localized animal farms, agrochemical and crop waste disposal and storage, industry, and urban wastes may lead to point pollution.

Using high amounts of nitrogen in fertilizers leads to an increase of nitrate in plants, where a part of the nitrate, not absorbed by plants, is washed away by rain and pollutes groundwater. Drinking water and eating vegetables with high amounts of nitrate and nitrite are toxice specially for children and babies because it may cause methemoglobinemia.

Acceptable daily intake (ADI) of nitrates is 0–5 mg/kg of body weight and of nitrites is 0–0.2 mg/kg of body weight. Children younger than 6 months are not allowed to consume food that contains nitrites. (Standards are set by common Committee FAO/SZO for nutrient additives—JECFA.) The limit values, that are supposed to ensure safe amounts of toxic elements for the majority of people have only relative value. The well-known statement by Paracelsus (1493–1541) is, “dosis sola facit venenum” (dose determines if a substance is toxic or not). All these concepts fail for substances with irreversible, especially carcinogenic or mutagenic effects. For carcinogenic and mutagenic substances, it is evident from the viewpoint of toxicology that prohibiting consumption is needed; any amount of such substances consumed are supposed to be a risk for human health.

According to the EU Nitrate Directive, the maximum standard for nitrate in drinking water (according to the value adopted by WHO) is 50 mg NO_3/L of water.

Bottled water is not supposed to contain any NO_2, whereas regular drinking water can contain up to 0.005 mg/L of NO_2-N under normal conditions and not more than 0.05 mg/L of NO_2-N in extraordinary circumstances.

When considering the possibility of pollution of groundwater by nitrates due to surpluses of nitrogen in the soil, it is essential to touch on the reasons for nitrate surpluses. The quantity of nitrate in vegetables (freshly edible crops) depends on the relative speed of two physiological processes: absorption of nitrates by the roots of plants and protein synthesis with consumption of nitrate and its reduction to nitrite and to ammonium. If the speed of absorption of nitrate is higher than the rate of protein synthesis, nitrate automatically accumulates in plants. On the other hand, if nitrate is transformed into protein in step with its absorption, nitrate is present only in small amounts. A disproportion between nitrogen absorption and protein synthesis may be caused for two reasons which sometimes are synergistic: (1) excessive quantities of nitrogen in the soil and (2) slowing down of protein synthesis.

The nitrate level in plants depends on the type of soil, climate conditions (water in the soil), soil cultivation practices, age of plants, and type of vegetable (cultivar and variety). The nitrate level is always higher in young plants. An excess of nitrate in the soil results above all from chemical nitrogenous fertilizers particularly in the form of nitrates, which are easily available to vegetables. Several factors affect the intensity of protein synthesis;the most important are the following: light, the state of the foliage, the lack of minerals, particularly of trace elements, and the use of certain pesticides. The feebler the light, the richer plants are in nitrate (crops under glass, particularly out of season, shady locations or overcrowding, short days, and weak sunshine). If leaves are insufficiently developed or damaged by parasitic attack, the nitrate content in the roots of plants is raised. Lack of minerals slows down protein synthesis which affects the accumulation of nitrates. All factors that obstruct normal biochemical processes in the soil or in plants can slow down protein synthesis and therefore accelerate the accumulation of nitrates for example, lack of water in the soil, excessive water in the soil due to inadequate drainage, inadequately cultivated soil, and use of herbicides.

It is not nitrate that is toxic but the compounds derived from it, nitrites and nitrosamines. There are two reasons that nitrate derived compounds are toxic:

1. Nitrates oxidize the ferrous ion of hemoglobin into ferric, methemoglobin is formed, and transport of oxygen is hindered (methemoglobinemia).
2. Nitrosamines are carcinogenic.

Nitrites and nitrosamines are formed during the microbiological process that takes place immediately after harvest whenever vegetables—medium rich in nitrate—are stored under anaerobic conditions at ambient temperature (sawerkrout, etc.). The very same processes take place in cooked vegetable stored at ambient temperature.

Many years of experiments have been conducted at the Agrohydrologic Research Station in Ljubljana, Slovenia (B. Maticic), on the impact of different water application levels in the soil and different levels of mineral nitrogen in the soil on the content of nitrate, nitrite, and ammonium on crop yield for many different vegetables. To prove the importance of sustainable nitrogen fertilization of crops and optimal treatment of crops with water (irrigation), we present some results of the interaction between nitrogen application and water application for three different vegetable crops: cabbage, red beet, and celery. The results of "water–nitrogen fertilizer–yield" production functions as well as "water–nitrogen fertilizer–nitrate" functions for these three crops are presented in Figs. 1–3.

The increase of water application levels caused a decrease in NO_3 content in vegetables until the water application reached a certain optimal level; a further increase of water application also increased NO_3 content. The influence of water application (irrigation) on NO_3 content is greater than the influence of fertilization with nitrogen. Extreme amounts of water (minimal, when plants are suffering from drought and maximal, when plants have too much water for normal growth) increase the level of NO_3 in vegetables.

It is obvious that fertilization with nitrogen increases the nitrate level in vegetables. For red beet, the maximal increase in nitrate level, caused by fertilization with nitrogen, was 450%; for cabbage only 60%. The increase

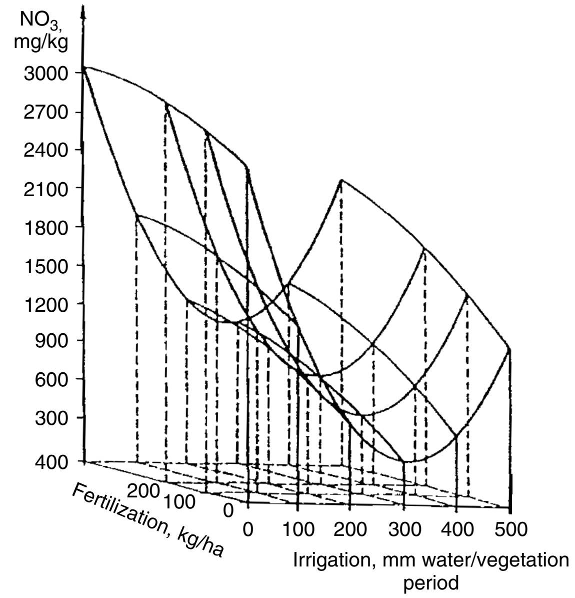

Figure 2. Nitrate content in red beet due to different water and different nitrogen treatments.

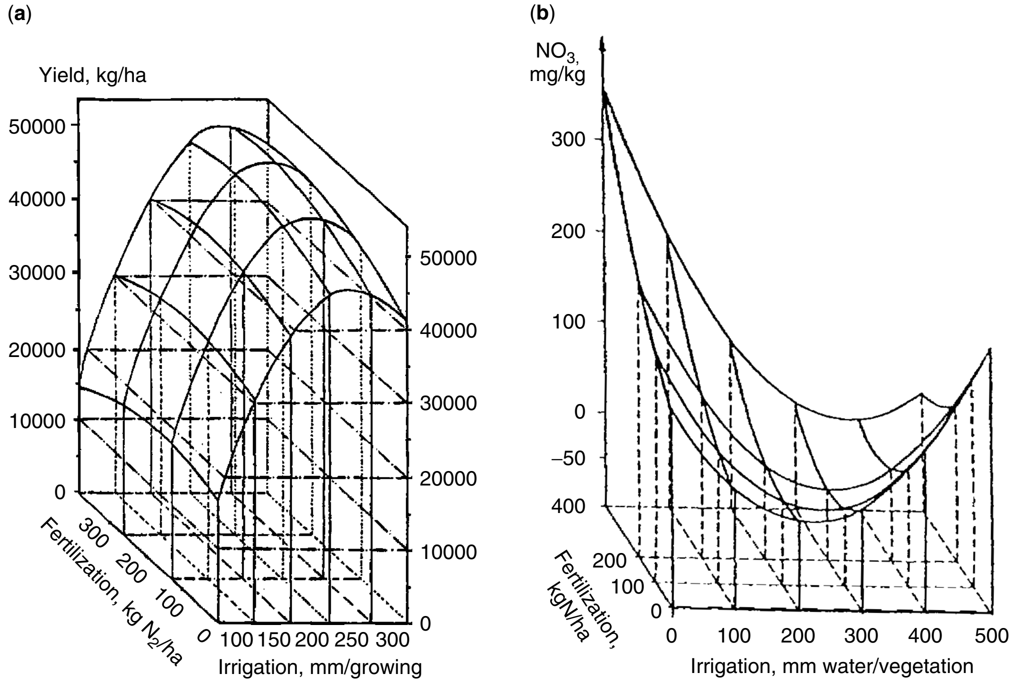

Figure 1. The impact of interaction between water application (irrigation) and nitrogen fertilization on crop yield (cabbage) and content of nitrate in the crop: (**a**) yield of cabbage due to different water and nitrogen treatments; (**b**) nitrate in cabbage due to different water and nitrogen treatments.

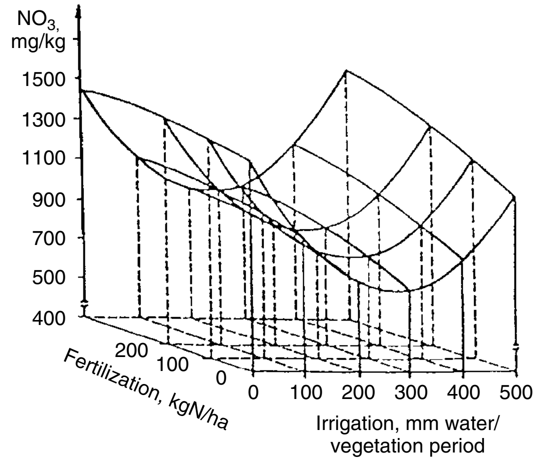

Figure 3. Nitrate content in celery due to different water and different nitrogen treatments.

in nitrogen doses from 0 to 400 kg N/ha caused an NO_3 content more than fifteen times higher in celery leaves and more than ten times higher in celery bulbs.

A significant reduction of nitrate and nitrite levels in vegetables was observed with optimal water supply and with small to medium doses of nitrogen fertilizers. According to the results, the recommended allowable water depletion from the soil for vegetable growers would be to maintain soil water conditions within the range of 20–30% of available water in the soil, and nitrogen fertilizer should be applied in several small doses.

To identify nitrogen surpluses (that may influence the pollution of groundwater and surface waters) due to agricultural technology used in different countries, it is advisable to demonstrate a "**mineral (nitrogen) balance**" on regional and farm levels using a **normative approach**. This method takes into account nitrogen input from mineral fertilizers, animal wastes, and the deposition from the atmosphere reduced by nitrogen uptake by harvested crops and ammonia losses to the atmosphere (30%).

It is also recommended to do thorough surface and groundwater quality **monitoring** according to the methodology recommended by international organizations with as many sampling points as possible for basic physical, chemical, and bacteriological analyses.

A mineral (nitrogen) balance provides information closer to the nitrate problem. Ammonia losses during storage do not directly affect groundwater. However, deposition from the atmosphere must be taken into account because it has an impact on nitrate concentration in groundwater. The average nitrogen deposition from the atmosphere is 20 kg nitrogen/ha (or 15.5 kg N/ha according to RIVM, The Netherlands). The nitrogen balance at the farm level is based on the so-called surface balance approach. The input factors include minerals provided by purchase of fertilizers and manure, minerals from production of animal manure, and deposition from the atmosphere. The output factor includes nitrogen from harvested crops. The assessments are based on elaborations by the Federal Agricultural Research Center, Braunschweig, Germany.

To reduce mineral (nitrogen) surpluses in agriculture and to meet the standards of nitrate and nitrite in drinking water and crops, it is necessary to control water and food quality in connection with nitrates and nitrites and to exact certain regulations regarding change in agriculture, especially animal excrement management. These regulations should give different norms for sustainable behavior in agricultural technology according to "Nitrate Directives" and "Code of Good Agricultural Practice." The most important norms should be as follows:

1. The highest quantity of manure allowed on agricultural land as well as limitations for the use of the manure in specific soil conditions:
   - The maximal allowed intensity of raising animals should be 3 LU/ha for cattle or 2 LU/ha for pigs and poultry.
   - Application of organic manure should not be allowed during winter on frozen soil.
   - Application of organic fertilizers should not be allowed on soil saturated with water.
   - Application of organic fertilizers should not be allowed in temporarily flooded areas.
   - Application of organic fertilizers should not be allowed near streams (10 m away from the stream) and in depressions where there is no runoff of water.
   - Applying liquid manure on bare soil should not be allowed in the period from Nov. 15 till Feb. 15.
   - Application of organic fertilizers on water aquifer protected areas should be very restricted and done only in agreement with local authorities and regulations valid for those areas.
   - In the vicinity of springwater and in underground water pumping areas, wastewater should not be drained to springwater or underground water in any case.
2. The highest quantity of N, P_2O_5, and K_2O allowed per hectare should be 170 kg N, 120 kg P_2O_5 and 300 kg K_2O.
3. Animal wastes should be stored in suitably arranged dung yards and cesspools so that there is no danger of leaking through and pouring over underground water.
4. The adjustment of farms to these regulations should be done as follows:
   - the adjustment of the number of animals (LU) according to the area of land available on the farm,
   - possible rent of additional land according to the contract, the construction of necessary dung yards and cesspools for hard and liquid manure, according to the restrictions.

Implementation of these regulations should be controlled by strict agricultural inspection in each country.

In countries where groundwater is relatively deep and drainage systems are supposed to drain "perched" water tables and to consider the hydraulics of water flow in the soil profile, drainage can prevent groundwater from pollution by minerals and pesticides that contaminate the upper layer of the soil. Polluted water flows to open canals (collectors) and further to rivers. In order not to pollute water in the rivers, it is an idea (B.Maticic) to install phytoremediation lagoons at the end of open canals (collectors) before water is discharged into surface water streams (rivers). This action would undoubtedly protect the quality of groundwater and water in rivers, and therefore act as an environmental measure.

When the world is going through many changes regarding exploitation of its natural resources and environmental protection, a new idea for efficient and sustainable draining of soil is of great importance. It provides the opportunity to obtain a common goal a “sustainable development.”

For many countries, where agriculture is a very important part of the economy, efficient and sustainable water management of agricultural land is very important. In the process of the transition of our society, there is strong emphasis on sustainable development, and it is believed that the field of agriculture is especially sensitive in this regard. That is why a big effort to overcome the mistakes of past development is being put forward to establish a whole new approach to agricultural planning and production with emphasis on the rational use of natural resources and modernized use of drainage and irrigation.

A determination exists to change old attitudes toward agricultural areas as underdeveloped areas, used only for mass food production, regardless of impacts on the environment. We believe that by careful planning and protection of the fertile soil in agricultural areas and forests as well as water resources, it is possible to achieve living conditions which would be attractive to people to remain on their land and continue farming traditions, as well as to the others to enjoy the environment of the countryside.

BIBLIOGRAPHY

1. Soren, R. and Boie, S.F. (1994). *National and EC Nitrate Policies*. Kobenhavn.
2. Matičič, B., Avbelj, L., Vrevc, S., and Jarc, A. (1995). *Water Pollution by Nitrate in Slovenia*. DELO, March 19.
3. Matičič, B. and Avbelj, L. (1996). *Water Pollution by Nitrate in Slovenia: Future Standards and Policy Instruments. ICID Proc.: 6th Drainage Workshop* - DRAINAGE AND THE ENVIRONMENT, TOPICS No. 4, Ljubljana.
4. Matičič, B. (1995). *Water Pollution by Nitrate in Slovenia*. Monograph.
5. Matičič, B. (1995). *Standards of Nitrate in Agriculture: The Case of Slovenia*. Monograph.
6. Matičič, B. (2004). *Drainage and Irrigation*. Contribution for Global Water Partnership Publication.
7. Brower, F.M. et al. (1994). *Mineral Balances of the European Union at Farm Level*. Agricultural Economics Research Institut, the Hague, the Netherlands.
8. *The Council Directive of 12 December 1991 Concerning the Protection of Waters Against Pollution Caused by Nitrates From Agricultural Sources*. (91/676/EEC). Official Journal of the European Communities, 31.
9. Water - Quality Survey, Unesco/WHO. (1978). *Manual on Water—Quality Monitoring*. WHO, 1988.
10. Schleef, K.H. and Kleinhanß, W. (1994). *Mineral Balances in Agriculture in the EU, Part I: The Regional Level*. Braunschweig, Germany.

NITRIFICATION

David D. Tarkalson
University of Nebraska-Lincoln
North Platte, Nebraska

Nitrification is the biologically facilitated oxidation of ammonium ($NH_4{}^+$) to nitrate ($NO_3{}^-$). The process of nitrification is performed by ammonium and nitrite ($NO_2{}^-$) oxidizing bacteria found in soil and waters (Table 1). In soil and water, ammonium is supplied through decomposition of organic matter by microbes and/or anthropogenic or natural additions of inorganic nitrogen sources containing ammonia (NH_3), urea, or ammonium. Nitrification consists of two independent microbially facilitated reactions. In the first reaction, *ammonium* oxidizers oxidize ammonium to nitrite, and in the second reaction, *nitrite* oxidizers oxidize nitrite to nitrate.

The purpose of nitrification (microbial facilitated conversion of ammonium to nitrate) is to provide nitrifying bacteria the energy they require for metabolism. The nitrification reactions are represented in Fig. 1, and the bacteria involved are listed in Table 1. There are two different microbially mediated nitrification reactions.

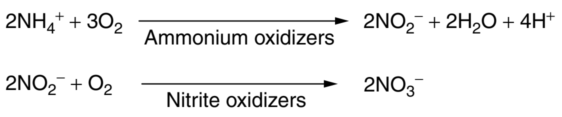

$$2NH_4{}^+ + 3O_2 \xrightarrow[\text{Ammonium oxidizers}]{} 2NO_2{}^- + 2H_2O + 4H^+$$

$$2NO_2{}^- + O_2 \xrightarrow[\text{Nitrite oxidizers}]{} 2NO_3{}^-$$

Figure 1. Nitrification reactions.

Table 1. Common Species of Nitrifiers and their Habitats[a]

Bacterial Species	Habitat
Ammonium oxidizers	
Nitrosomonas europea	Soil, water, sewage
Nitrosospira briensis	Soil
Nitrosococcus oceanus	Marine
Nitrosococcus mobilis	Marine
Nitrosococcus nitrosus	Marine
Nitrosolobus multiformis	Soil
Nitrosovibrio tenuis	Soil
Nitrite oxidizers	
Nitrococcus mobilis	Marine
Nitrobacter gracilis	Marine
Nitrobacter winogradski	Soil
Nitrobacter agilis	Soil, water

[a] Ref. 1.

Chemoautotrophic nitrification is the use of carbon dioxide and carbonates as a carbon source for cell construction and oxidation of inorganic molecules such as NO_3^- or NO_2^- as an energy source. Chemoautotrophic nitrification is carried out only by a select group of bacteria. Heterotrophic nitrification is the use of organic compounds for both a carbon and energy source. Heterotrophic nitrification is executed by a wide variety of heterotrophic bacteria and fungi (2). Chemoautotrophic nitrification is believed to be up to 1000 times faster and is better understood than heterotrophic nitrification. Chemoautotrophic nitrification dominates in neutral to alkaline agricultural soils, whereas heterotrophic nitrification dominates in acid forest soils (2).

RELATIONSHIP BETWEEN SOILS AND WATERS

Nitrification is a key component of the nitrogen cycle (Fig. 2). The nitrogen cycle represents the fundamental transformations of nitrogen in the environment. In agricultural soil systems, we strive to manage this cycle better to increase crop production and minimize the negative impacts of nitrogen on the environment. Implementing management practices that will help manage nitrogen in agricultural systems to protect surface and groundwaters is an example. Nitrate, the end product of the nitrification process, is highly soluble in water and is not readily adsorbed to soil particles; therefore, if nitrate is not used by plants or microbes as a nutrient, it can move vertically through soils to groundwater and in runoff to surface waters resulting in problems from nutrient enrichment. Because nitrification is a microbially mediated process, many management practices aimed at reducing nitrate leaching focus on the nitrification process.

FACTORS INFLUENCING NITRIFICATION

Assuming an adequate supply of ammonium and a sufficient population of nitrifying bacteria, there are several factors that affect nitrification in natural systems. These factors include pH, moisture, temperature, and oxygen concentration.

pH

Soil and water pH influence the availability of many essential nutrients such as calcium and phosphorus (3). Microbes require these nutrients for growth and development; therefore, pH ranges that promote adequate availability of these nutrients are required for nitrification. Nitrification in soils is optimized at pH 8.5 but can proceed in a pH range of 4.5 to 10 (3). Figure 3 illustrates the relationship between soil pH and nitrification. At low pH levels, many essential nutrients are less available for use by plants and microbes.

Moisture

Nitrifying bacteria require water to live and function. In soils, the optimum water content for nitrification is at field capacity (water remaining in soil after gravity has removed excess water after a period of saturation). Research shows

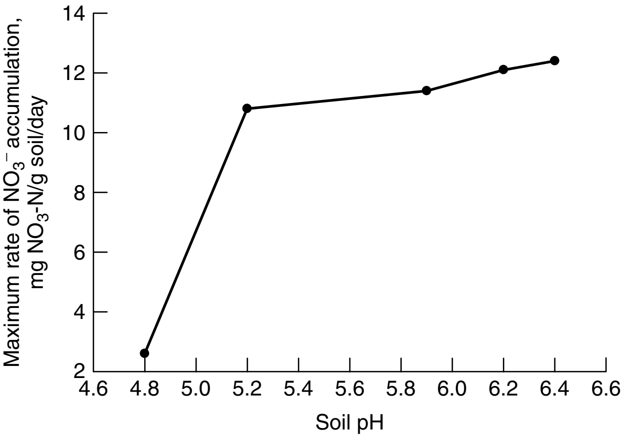

Figure 3. Effect of soil pH prior to incubation on nitrate accumulation in soil treated with 169 kg N as NH_4NO_3/ha and incubated at 23 °C for 30 days (4).

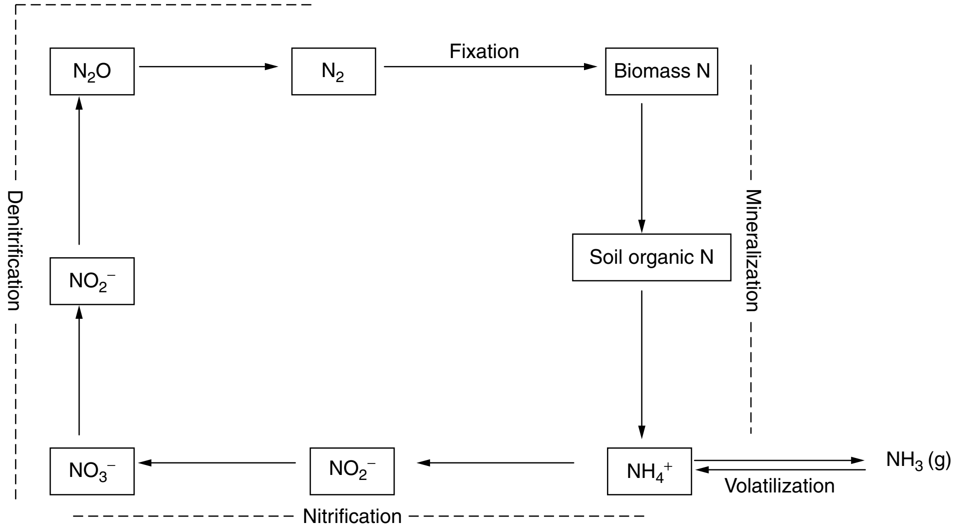

Figure 2. Nitrogen cycle.

Next Page

that, as water content increases in soil from dry to an optimum content, nitrification rates increase (Fig. 4). However, waterlogged conditions drastically decrease oxygen content and thus decrease the rate of nitrification. The water content at field capacity varies, depending on the soil physical properties such as texture and organic matter content.

Temperature

Generally, as temperature increases to a point so does the rate of nitrification (Fig. 5). Studies have shown that ammonium is converted to nitrite at temperatures from 0 °C to 65 °C and nitrite is converted to nitrate from 0 °C to 40 °C. Rates of nitrification dramatically decrease above and below these ranges. Optimum temperatures for nitrification are between 30 °C and 35 °C.

Oxygen Concentration

Figure 1 illustrates the need for oxygen in the nitrification process. Diffusion rates and concentrations of oxygen in soils and water will directly influence the rates of nitrification.

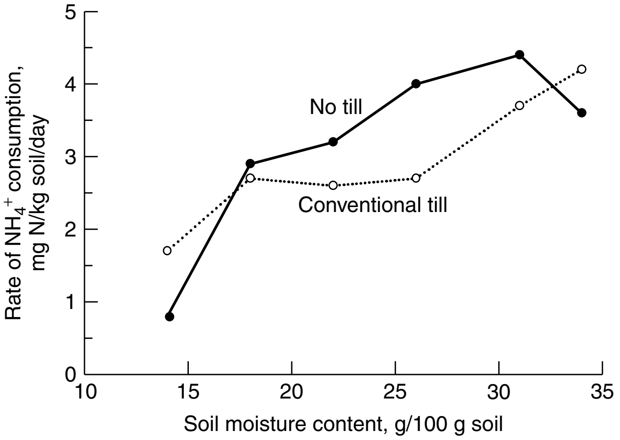

Figure 4. Nitrification rate of two tillage systems of a Maury silt loam soil column related to soil moisture content (5).

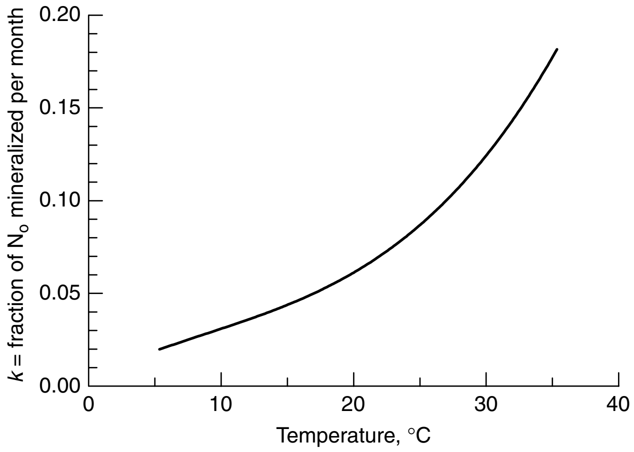

Figure 5. Fraction of N mineralized per month, k, in relation to temperature (k was estimated graphically for observed average monthly air temperatures) (6).

MANAGEMENT OF NITRIFICATION IN SOIL TO REDUCE NITRATE LOSSES TO SURFACE AND GROUNDWATER

Reducing nitrate losses in agricultural soils that receive large nitrogen (ammonia and ammonium) input from fertilizers is important to maintain high quality water resources. Chemicals called nitrification inhibitors have been developed that inhibit the nitrogen oxidative processes of nitrifying bacteria when added to agricultural soils concurrently with ammonium based fertilizers. In other words, nitrogen in the form of ammonium persists in the soil for longer periods of time that allow for increased plant use of the ammonium. Ammonium is adsorbed to soil particles; therefore, it does not readily move with water like nitrate. Nitrification inhibitors have a finite effective functionality in soil because soil microbes degrade them over time. Studies have shown that nitrification suppression by the inhibitor dicyandiamide can last 3 months, depending on the amount added with fertilizer (3).

Management practices that use knowledge of the factors that influence nitrifiers are common. For example, most farmers apply their ammonium based nitrogen fertilizers late in the fall when soil temperatures have declined to a point where nitrification is inhibited. If the farmer applies the fertilizer when soil temperatures are sufficiently high, then the ammonium is nitrified to nitrate and rainfall and/or snowmelt can leach the nitrate out of the root zone to the groundwater before the next crop can use the nitrogen.

Nitrification is a natural biological process involved in the transformation and fate of nitrogen in the environment. However, anthropogenic influences can either accentuate or reduce nitrification rates and can have a large impact on water quality.

BIBLIOGRAPHY

1. Tate, R.L. (1995). *Soil Microbiology*. John Wiley & Sons, New York.
2. Coyne, M.S. (1999). Soil microbiology. *An Exploratory Approach*. Delmar, Albany, NY.
3. Tisdale, S.L., Nelson, W.L., Beaton, J.D., and Havlin, J.L. (1985). *Soil Fertility and Fertilizers*. Macmillan, New York.
4. Dancer, W.S., Peterson, L.A., and Chesters, G. (1973). Ammonification and nitrification of N as influenced by soil pH and previous N treatments. *Soil Sci. Soc. Am. Proc.* **37**: 67–69.
5. Rice, C.W. and Smith, M.S. (1983). Nitrification of fertilizer and mineralized ammonium in no-till and plowed soil. *Soil Sci. Soc. Am. J.* **47**: 1125–1129.
6. Stanford, G., Carter, J.N., Westermann, D.T., and Meisinge, J.J. (1977). Residual nitrate and mineralizable soil nitrogen in relation to nitrogen uptake by irrigated sugarbeets. *Agron. J.* **69**: 303–308.

OCCURRENCE OF ORGANOCHLORINE PESTICIDES IN VEGETABLES GROWN ON UNTREATED SOILS FROM AN AGRICULTURAL WATERSHED

Mariana Gonzalez
Karina S.B. Miglioranza
Universidad Nacional de Mar del Plata
Mar del Plata, Argentina
and
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Buenos Aires, Argentina

Julia E. Aizpún de Moreno
Víctor J. Moreno
Universidad Nacional de Mar del Plata
Mar del Plata, Argentina

INTRODUCTION

The role of agrochemicals in modern agriculture is continuously evolving, and their contribution to crop protection continues to increase. At a global level, the use of pest control agents (i.e., pesticides) has proved to assist the solving of many problems facing human health and food production. Organochlorine pesticides (OCPs) are a class of compounds of particular concern in the environment because of their recalcitrance in natural solids, global transport, distribution, and toxicity, which leads to include them into the group of persistent organic pollutants (1). This group had been widely used since early 1940, with an intensified usage during the Second World War. At that time, crops were duplicated and malaria incidents were reduced in India from 75 millions to less than 5 millions persons affected in 10 years as the result of DDT usage. However, such usage has occasionally been accompanied with hazards to man and the environment. Target species became resistant, leading to the application of higher doses of pesticides. For these reasons, most developed countries banned or restricted their usage since 1970. However, in the poorer countries, OCP restrictions or prohibitions date from the later 1990s. In addition, many of the banned or withdrawn pesticides from the developed market are still produced and sold in developing country market (2). Pesticides such as DDT are still in use in India and Brazil with public sanity purposes (3–5).

When pesticide are agriculturally applied, roughly 85–90% of pesticides never reach the target organism, but disperse through the air, soil, and water (6,7). Pesticides that are bound in soils or taken up by plants and animals are drawn into rivers and lakes and move into the aquatic food chains inducing severe damage to aquatic life (8,9). The environmental behavior of OCPs is largely related to their physical and chemical characteristics. Some of them are volatile enough to evaporate, reaching the upper atmosphere and being carried hundreds or thousands of kilometers by a process known as “Global Distillation” (1,10,11). Thus, the capability of these pollutants for long-distance transfer, precipitation, and accumulation in environmental materials results in the formation of background concentrations in remote regions from the emission sources (12). Therefore, OCPs can be found in many places where they had never been applied. They have been linked to carcinogenicity and endocrine disruption in mammals, and concerns over toxicity are exacerbated by their hydrophobicity, which results in bioaccumulation in fatty tissues and biomagnification through food chains (13).

The general population can be exposed to low levels of pesticides in three general ways:

1. vector control for public health and other nonagricultural purposes
2. environmental residues
3. food residues

Regarding food residues, consumption of plants or plant products, such as vegetables, fruits, and grains, form a major part of the food consumption of human beings and cattle (14). Moreover, vegetables are an essential component of our diet because of their nutritional value (15).

The impact of soil contaminants on human health and accumulation of pesticides in food crops is a great concern because of increased dietary exposure of consumers (16).

Considering that vegetation is the link among the atmosphere, soil, and human food supply, contamination of plants will have a great influence on the total daily intake of a substance (14). Thus, it is important to understand those processes by which pollutants enter into this environmental compartment.

ABSORPTION AND ACCUMULATION OF OCPs IN PLANTS

Plants may accumulate OCPs via different pathways:

1. adsorption to the root surface
2. root uptake and transport to the upper plant tissues
3. uptake of airborne vapors by aerial plant parts (stem, leaves, and fruits)

The relative significance of each route by which OCPs can be taken up into fruits and vegetables is likely to be influenced by a wide range of factors (i.e., physicochemical characteristic of the pesticide, environmental conditions, and specie-specific characteristics) (17,18) (Fig. 1).

Air to Plant Route

OCPs are emitted to the atmosphere via various routes: by transfer to the airflow during the treatment of agricultural crops, evaporation from treated surfaces of plants and soils, and finally by wind erosion (12).

Simonich and Hites (19) presented a simplified mechanism of pollutants absorption by plants from air. Air to vegetation transfer of gaseous pollutants will be on the partition coefficient of the pollutant in the octanol

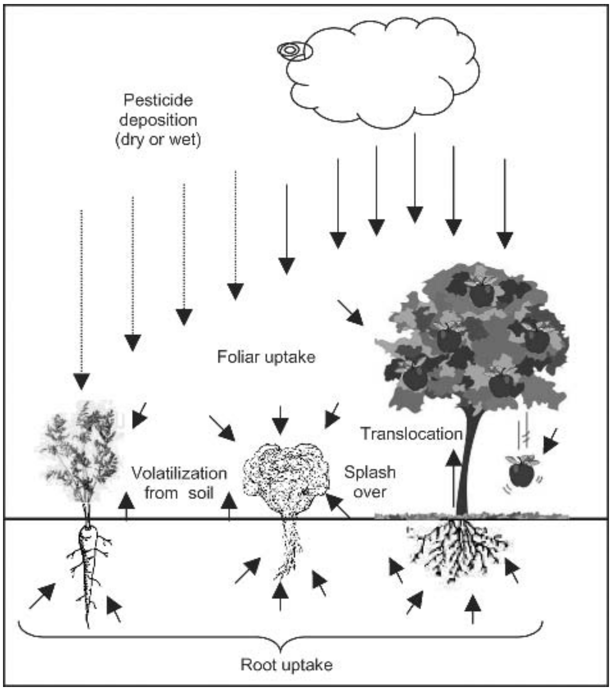

Figure 1. Pathways for accumulation of Organochlorine Pesticides by plants. Modified of Lovett et al. (18).

air system, the surface area, and lipid concentration in plant tissues.

Plant cuticles are the first barrier in the absorption of airborne and fall out OCPs in all aboveground organs and parts of higher terrestrial plants (fruit, flowers, stems, and leaves). Moreover, gaseous contaminants can directly reach the intercellular aerial medium through the open stomata (20–22). It was found that OCP residues in plant tissues decrease very slowly because of their low solubility in water and high solubility in oils. Thus, sorption of pesticides by plant surfaces depends significantly on the morphological features of the leaf, velocity, waxing, etc. (12,23). Moreover, vapor transfers of OCPs from soil to the aboveground parts of plants would be expected to occur to varying extents for all crops. For vegetables such as lettuce, courggetes, or zucchinis, which develop in contact with or in close proximity to the ground, it is possible that uptake of OCPs could also occur by direct adsorption from the soil and by splash-over of soil particles during heavy rain (12,23,24).

Soil to Plant Route

The transition of OCPs from soils into plants (35–70%) will be, in general, higher than into adjacent environments: water (2–18%) and atmosphere (18%) (12).

The uptake of OCPs by plant roots is affected by the bioavailability of the pesticide in the soil, leading to a competition between the organic fraction of the soil and the root material. Higher soil adsorption will result in a lower plant bioavailability (25).

After root uptake, plants can translocate pesticides via xylem to the aboveground plant parts. The extent of such mobilization will be on pesticide solubility and plant characteristics, such as root exudates (26). Thus, pesticides whose kow are higher than 4 will not be able to translocate from root to aerial tissues in most plant species (17). However, an exception has been reported for vegetables from the cucumis and cucurbitacea genera, which are able to translocate highly hydrophobic compounds, such as the insecticide DDT, PCBs, and dioxinas, from roots to fruits (26–29).

OCPs OCCURRENCE IN VEGETABLES

As was previously mentioned, vegetables are essential components of the human diets. As pesticide residues remain on fruits and vegetables, it is important to know the magnitude of such contamination from the point of view of safety to consumers. Thus, a growing concern for safer foods has led research into increased pesticides residues monitoring.

In order to regulate and control such residues, the concept of maximum residues levels (MRL) was established. This level represents the maximum concentration of the residue (expressed in mg/Kg) that is legally permitted in specific food items and animal feeds (31).

OCP residues in vegetables are included in the Food and Drug Administration (FDA) monitoring program because they are still present in the environment (32). A residual content of these compounds in foods today is still quantifiable, and their avoidance in farming is largely impossible (33).

Regular surveys and monitoring programs of OCP residues have been carried out for decades in developed countries, such as the United States (32,34), the United Kingdom (35), and Spain (36). Although in some developing countries, such as India (37), Tunisia (38), Nigeria (39), Rumania (40), Brazil, the Philippines (41), Egypt (42–44), Honduras (45), and Argentina (46), reports about OCP levels in vegetables were made, they were not made as a part of governmental monitoring programs. Most of the reports showed that OCPs, such as HCHs, DDTs, Aldrins, Chlordane, and Heptachlor, were present in all analyzed samples. In the United States, about 1% of the samples violated the established MRL (32), but in developing countries, such as Brazil, this value can increase up to 3.7% (41). The presence of OCPs in vegetable samples is the consequence of either recent and current use in developing countries and/or the result of their persistent residues 30 years after their prohibition.

One alternative to synthetic pesticide use is organic farming. Thus, two decades ago, growing environmental awareness and concerns about safer foods have lead people to question modern agricultural practices, which has been reflected in an increasing demand for organic production, which is perceived as less damaging to the environment and to be healthier than conventionally grown foods (47,48).

ORGANIC AGRICULTURE: WORLD WIDE SITUATION

Organic agriculture has developed rapidly during the last decade, and governments have encouraged the development of this tillage practice, mainly in developed countries. Chemophobia is the commonest reason for the public to choose organic food on the assumption that such food is free of synthetic pesticides (49).

Organic farms distinguish from all other forms of farming, namely conventional, by a rejection of soluble minerals as fertilizer and synthetic pesticides in favor of natural ones (49). However, organic food contains pesticide traces, although the amounts are lower than in conventional produce (50,51). Although it is likely that organically grown foods are lower in pesticide residues, very little documentation of residues levels exists (52). Moore et al. (33) studied the residue of eight OCPs in two matrices (carrots and winter squash) used to elaborate conventional and organic baby food brands. They found no detectable residues among all types of food. These results were the consequence of processing fruits and vegetables (washing, peeling, and cooking) that largely eliminates residues, as was largely demonstrated (53–55).

Today, developing countries represent a source of a wide range of organic produce, especially tea, coffee, cocoa, herbs, and spices, as well as tropical and subtropical fruits and vegetables for Europe and the United States (56).

Vegetable diversity produced in Argentina reaches 115 species with field crops covering a surface area of some 600,000 hectares (57), and about 160,000 hectares are destined to organic production. The European Union, EE.UU, and Switzerland are the main markets of destiny for these products (58). Therefore, the study of the factors that may impact in this kind of production is relevant. Most organic farms are supposed to remain small and self-sufficient, whereas conventional farms are increasing in size as the result of economic pressures.

The environmental fate of OCPs in agricultural environments of Argentina have been studied, including the impact of these pollutants in watersheds characterized by rural and small-case farming land use (59–62). OCPs were found in vegetables from organic and conventional farming (46,52,63,64). The presence of these contaminants in the organic farm was an interesting finding.

OCPs IN VEGETABLES ORGANICALLY CULTIVATED

Previous to the settlement of the organic farm, the area was characterized by natural vegetation. The farm is surrounded by important hedgerows and trees, and neighboring fields on both sides are conventionally farmed since its establishment about 30 years ago. Agrochemical use was a common practice in surroundings fields, and unsprayed field margins were not established.

From the organic farm, several vegetable species were studied in order to estimate the OCPs levels in their tissues. The produces include leafy vegetables (lettuce and chard), tomato, and leek. Most of the studied crops showed similar OCP levels between root and aboveground plan parts (stem or leaves). This observation was not consistent with previously published data where plants were cultivated in treated soils (28,30,65). Thus, in the organically treated soils, the air to plant route would be as important as the soil to plant route in all species.

The main OCPs found in vegetables were the same that were previously reported for conventional produce: HCHs, DDTs, and Chlordane being these pesticides forbidden for agricultural purposes. Moreover, high levels of endosulfans (α- and β-isomers and sulfate metabolite) were also found as a consequence of their current usage. All species accumulated OCPs efficiently from soil in their roots, leading to high root bioconcentration factors. Common practices in organic tillage constitute the enrichment of soil with vegetal rests. In this farm, pine needles collected form surrounding forests are usually added to the soil at the beginning of the growing periods. Pine needles, rich in waxes, are known to accumulate highly lipophilic pollutants from the atmosphere (66,67). Therefore, those soils enriched with this amendment showed the highest level of OCPs. Moreover, plants growing in such soils showed enrichment in the more lipophilic compounds in their roots.

Aerial tissues showed a great variation in accumulation regarding species and tissues. In lettuce, differences between inner and outer leaves were found, with the lowest concentrations in the outer leaves despite the higher lipid content, maybe by the growth dilution effect or weathering. Moreover, plant morphology could be an important factor affecting OCPs uptake and distribution (23,24). Thus, lettuce with outer leaves enlarging inner leaves had higher OCP levels in stem and inner leaves, probably as a consequence of the accumulation of pollutants between outer and inner leaves. On the other hand, chard plants with an inverse wide-high relation and longer stem showed higher OCP levels in outer leaves (51).

Tomato fruits contained OCPs not only in their peel but also in the inner flesh tissues, which indicate absorption rather than adsorption (63). Total OCPs in flesh were 3-fold greater than in peel, probably as a consequence of either continuous uptake by aerial tissues (fruit peel, leaves, and stem) of airborne OCPs, followed by translocation via phloem throughout the plant and accumulation in fruit flesh or removal of pesticides from peel by weather influences as photodegradation, volatilization, and so on (17).

Acaricide Dicofol is still in use in this country and has been reported to contain a significant amount of DDT as manufacturing impurity. Thus, it represents a potential continuous source of DDT in the studied area. DDT breaks down to DDE and DDD in soil and generally with time the parent/metabolite ratio decrease (68). DDT/DDE ratios found > 1 suggest that besides Dicofol use, new DDT can be continuously transported to the atmosphere of the organic farm. The presence of this new DDT can be attributed to illegal use or to atmospheric transport from those countries, mostly tropical, where DDT is still in use (61). Moreover, data suggest that long-range transport could be the predominant source of chlordane in the farm. The α-/γ-HCH ratio can be used as diagnostic of HCH sources, inputs of γ-HCH (lindane) decrease the ratio found in air and vegetation containing a background of technical HCH (α-/γ- = 4–6) (68). The low α-/γ-HCH ratio found in all samples reflect recent inputs of lindane to the atmosphere of the studied area.

Endosulfan is one of the few organochlorine insecticides still in use; it is primarily used on summer crops such as tomatoes in the agricultural belt on the northwest area of the studied watershed. Occurrence of endosulfan in aerial tissues of tomato agrees with their summer usage in the surrounding farms, showing the relevance of drift pesticide during application events. Research has reported that drift of endosulfans from aerial agricultural spraying can produce lethal concentration to fish in shallow exposed water bodies 200 m away from the target spray area (8). In this watershed, it is also common during summer months that a massive "fish kill" can occur. One possible explanation could be that lethal concentration inputs of endosulfans applied in the surrounding area that reach the waterbody. Moreover, surface runoff from the organic farm may constitute a source of endosulfans to the waterbody.

As a whole, it was demonstrated that plants bioaccumulated OCPs efficiently from both routes, air and soil. In organic tillage, it is common that crop residues are reincorporated into soil after harvest. Thus, soil OCP values may reflect contributions from decomposed plant material rather than changes either within soil or via the atmospheric deposition. Therefore, plant residues constitute a way of entry of atmospheric pollutants into of the organic farm, leading to the subsequent availability for root uptake. This fact will be relevant for plants whose edible parts grow below ground (i.e., carrots, potatoes) or those that are able to translocate them. Despite OCPs being detected in all edible tissue (leaves, stem, and fruit in the studied vegetables), the levels were below the maximum residue limits (MRL) considered by both Código Alimentario Argentino (CAA) (69) and the Codex Alimentarius (CA) (31).

Despite the practice of organic farming management, OCPs were present in the soil environment and vegetable tissues confirming the importance of contamination from adjacent conventional farms. Therefore, the occurrence of OCP residues in the organic samples may be attributed to ecosystem factors such as pesticide runoff, drift from pesticide used by neighboring farming, volatilization, leaching, and washout. Thus, when organically grown vegetables are subject to such potential sources of contamination, the meaning of the term "organic" becomes obscure and is difficult to interpret especially from the perspectives of the public consumers.

BIBLIOGRAPHY

1. Wania, F. and Mackay, D. (1996). *Environ. Sci. Technol.* **30**(9): 390–396.
2. Kannan, K., Tanabe, S., Ramesh, A., Subramanian, A., and Tatsukawa, R. (1992). *J. Agric. Food. Chem.* **40**: 518–525.
3. Minelli, E.V. and Ribeiro, M.L. (1996). *Bull. Environ. Contam. Toxicol.* **57**: 691–696.
4. Mathiessen, P. (1985). *Environ. Pollut.* **10**: 189–211.
5. Ramesh, A., Tanabe, S., Murse, H., Subramanian, A.N., and Tatsukawa, R. (1991). *Environ. Pollut.* **74**: 293–307.
6. Moses, M. et al. (1993). *Toxicol. Ind. Health.* **9**(5): 913–959.
7. Van der Werf, H.M.G. (1996). *Agric. Ecosys. Environ.* **60**: 81–96.
8. Berril, M., Coulson, D., McGillivray, L., and Pauli, B. (1998). *Environ. Toxicol. Chem.* **17**: 1738–1744.
9. Hapeman, C.J., Dionigi, C.P., Zimba, P.V., and McConnel, L.L. (2002). *J. Agric. Food Chem.* **50**: 4382–4384.
10. Iwata, H., Tanabe, S., Sakai, N., and Tatsukawa, R. (1993). *Environ. Sci. Technol.* **27**(6): 1080–1098.
11. Iwata, H., Tanabe, S., Ueda, K., and Tatsukawa, R. (1995). *Environ. Sci. Technol.* **29**(3): 792–801.
12. Galiulin, R.V., Bashkin, V.N., and Galiulina, R.A. (2002). *Water Air Soil Poll.* **137**: 179–191.
13. Patlak, M. (1996). *Environ. Sci. Technol.* **30**: 540–544.
14. McLachlan, M.S. (1996). *Environ. Sci. Technol.* **30**: 252–259.
15. Kumari, B., Madan, V.K., Kumar, R., and Kathpal, T.S. (2002). *Env. Monitor. Asses.* **74**: 263–270.
16. Torres, C.M., Picó, Y., and Mañes, J. (1996). *J. Cromatogr. A.* **754**: 301–331.
17. Trapp, S. and McFarlane, C. (1995). *Plant Contamination: Modeling and Simulation of Organic Chemical Processes*. Lewis Plublishers, Boca Raton, FL, p. 254.
18. Lovett, A.A., Foxall, C.D., Creaser, C.S., and Chewe, D. (1997). *Chemosphere.* **34**: 1421–1436.
19. Simonich, S.T. and Hites, R.A. (1995). *Env. Sci. Technol.* **29**(12): 2905–2914.
20. Schreiber, L. and Schonherr, J. (1992). *Environ. Sci. Technol.* **26**(1): 153–159.
21. Riederer, M. (1990). *Environ. Sci. Technol.* **24**(6): 829–836.
22. Bakker, M.I., Bass, W.J., Sijm, D.T.M.M., and Kolloffel, C. (1998). *Phytochemistry.* **47**(8): 1489–1993.
23. Ripley, B.D., Ritcey, G.M., Ronald Harris, C., Denomme, M.A., and Lissemore, L.I. (2003). *J. Agric. Food Chem.* **51**: 1328–1335.
24. Trapp, S. and Matthies, M. (1995). *Environ. Sci. Technol.* **29**(9): 2333–2338.
25. Ryan, J.A., Bell, R.M., Damidson, J.M., and O'Connor, G.A. (1988). *Chemosphere.* **17**(12): 2299–2323.
26. Campanella, B. and Paul, R. (2000). *Int. J. Phytoremediation.* **2**: 145–158.
27. Hülster, A., Müller, J.F., and Marschner, H. (1994). *Environ. Sci. Technol.* **28**: 1110–1115.
28. Müller, J.F., Hülster, A., Päpke, O., Ball, M., and Marschner, H. (1994). *Chemosphere*: 2175–2181.
29. White, J.C. (2000). *Inter. J. Phytoremediation.* **2**: 133–144.
30. Kumari, B. et al. (2003). *Env. Monit. Asses.* **87**: 311–318.
31. Codex Alimentarius. (2001). *FAO/WHO Codex Alimentarius Commission, Maximum Residue Limits for Pesticides*. FAO/WHO, Rome, Italy.
32. Food and Drug Administration. (1997). Pesticide program-residue monitoring. Available at http://www.cfsan.fda.gov.
33. Moore, V.K., Zakib, M.E., and Zakib, M.J. (2000). *Food Chem.* **71**: 443–447.
34. Luke, M.A., Masumoto, H.T., Cairns, T., and Hundley, H.K. (1988). *J. Assoc. Off. Anal. Chem.* **71**: 415–433.
35. MAFF. (1989). *Food Surveillance Paper 25*. HMSO, London, p. 71.
36. Chavarri, M.J., Herrera, A., and Ariño, A. (2004). *J. Sci. Food Agric.* **84**: 1253–1259.
37. Kaphalia, B.S., Takroo, R., Mehtrota, S., Nigam, U., and Seth, T.D. (1990). *J. Assoc. Off Anal. Chem.* **73**: 509–512.
38. Driss, M.R. and Boquera, M.L. (1987). *Acta Biol. Hungarica.* **38**: 101–106.

39. Adeyeye, A. and Osibanjo, O. (1999). *The Sci. Total Environ.* **23**: 227–233.
40. Hura, C., Leanca, M., Rusu, L., and Hura, B.A. (1999). *Toxicol. Lett.* **107**: 103–107.
41. Mansour, S.A. (2004). *Toxicol.* **198**: 91–115.
42. Abou-Arab, A.A.K., Solivan, K.M., El-Tantawy, M.E., Ismail, B.R., and Naguib, K.H. (1999). *Food Chem.* **67**: 357–363.
43. Abou-Arab, A.A.K., Solivan, K.M., and Naguib, K.H. (1998). *Bull. Nutr. Inst.* **18**(1): 117–137.
44. Abou-Arab, A.A.K. (1999). *Food Chem.* **65**: 509–514.
45. Kammerbauer, J. and Moncada, J. (1998). *Environ. Poll.*: 171–181.
46. Miglioranza, K.S.B., Aizpún de Moreno, J.E., Moreno, V.J., Osterrieth, M.L., and Escalante, A.H. (1999). *Environ. Poll.* **105**: 91–99.
47. Schifferstein, H.N.J. and Ouete Ophuis, P.A.M. (1998). *Food Qual. Pref.* **9**(3): 119–133.
48. Williams, P.R.D. and Hammit, J.K. (2001). *Risk Analysis.* **21**(2): 319–330.
49. Trewavas, A. (2004). *Crop Prot.* **23**: 757–781.
50. Baker, B.P., Benbrook, C.M., Groth, E., III, and Lutz Benbrook, K. (2002). *Food Add. Contam.* **19**(5): 427–446.
51. Gonzalez, M., Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2004). *J. Food Chem. Toxicol.* In press.
52. Bourn, D. and Prescott, J. (2002). *Cri. Rev. Food Sci. Nutr.* **42**(1): 1–34.
53. Schattenberg, H.J., III, Geno, P.W., Hsu, J.P., Fry, W.G., and Parker, R.P. (1996). *J. AOAC Int.* **79**: 1447–1453.
54. Krol, W.J., Arsenault, T.R., Pylpiw, H.M., Jr., and Incorvia Mattina, M.J. (2000). *J. Agric. Food Chem.* **48**: 4666–4670.
55. Ramesh, A. and Vijayalakshmi, A. (2002). *Pest Manag. Sci.* **58**: 1048–1054.
56. Barret, H.R., Browne, A.W., Harris, P.J.C., and Cadoret, K. (2002). *Food Policy.* **27**: 301–318.
57. Ministerio de la Producción. (2003). Secretaria de Agricultura Ganadería, Pesca y Alimentos de la República Argentina. Available at http://www.sagpya.mecon.gov.ar/0-0/index/agricultura/index_agricultura.htm.
58. SENASA. (2003). *Dir. Nac. de Fisc. Agroal. Dir. Cal. Agroal. Coord. Prod. Ecól.* "Situación de la Producción Orgánica en la Argentina durante el año 2003", Bs. As. Argentina.
59. Miglioranza, K.S.B. et al. (2002). *Environ. Sci. Poll. Res.* **9**: 250–256.
60. Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2003). *J. Soil Sed.* **4**: 264–265.
61. Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2003). *Environ. Toxicol. Chem.* **22**: 712–717.
62. Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2004). *Water Res.* **38**: 1765–1772.
63. Gonzalez, M., Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2003). *J. Agric. Food Chem.* **51**: 1353–1359.
64. Gonzalez, M., Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2003). *J. Agric. Food Chem.* **51**: 5024–5029.
65. Mattina, M.J.I., Ianucci-Berger, W., and Dykas, L. (2000). *J. Agric. Food. Chem.* **48**: 1909–1915.
66. Krauthacker, B., Romanic, S.M., and Reiser, E. (2001). *Bull. Environ. Contam. Toxicol.* **66**: 334–341.
67. Hank, M., Umlauf, G., and Melchlam, M.S. (1994). *Environ. Sci. Technol.* **32**: 1546–1543.
68. Bidleman, J.F. and Falconer, R.L. (1999). *Environ. Sci. Technol*: 206a–209a.
69. CAA. (1969). *Código Alimentario Argentino: Reglamento Alimentario, Normas Sobre Higiene*. Ley Nacional 18.284.

PESTICIDE CHEMISTRY IN THE ENVIRONMENT

Karina S.B. Miglioranza
Mariana Gonzalez
Universidad Nacional de Mar del Plata
Mar del Plata, Argentina
and
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Buenos Aires, Argentina

Julia E. Aizpún de Moreno
Víctor J. Moreno
Universidad Nacional de Mar del Plata
Mar del Plata, Argentina

INTRODUCTION

Pesticides have greatly enhanced the production and quality of food, feed, and fiber, as well as the control of disease vectors and pests adversely affecting the health and welfare of the world. Specifically, pesticides have been used in: (a) agriculture to increase productivity and quality and quantity of food, feed, and fiber; (b) forestry for pest control in silviculture; (c) the industrial, commercial, municipal, and military sector for rodent and weed/brush control around plant sites and in right-of-ways; (d) medical vector control; and (e) urban environment for termite control around structures and pest control in gardens (1). As a consequence, of all the environmental contaminants, pesticides have probably been the most widely criticized because of their direct use in natural systems. The nature of pesticide usage often requires broadcast distribution over large areas of crops. This wide treatment of the crop environment often results in treatment of adjacent areas as well. The use of pesticides is indispensable for conventional labor and extensive farming systems. Although the use of pesticides has resulted in increased crop production and other benefits, it has raised concerns about potential adverse effects on the environment and human health.

ENVIRONMENTAL BEHAVIOR OF PESTICIDES

The proportion of applied pesticide reaching the target pest has been found to be less than 0.3%, thus leaving over 99% elsewhere in the environment (2). Environmental entry and transport processes can occur either between or during rainfall events, including wet and dry deposition, foliar washoff, and spray drift. Moreover, pesticide dissolved in atmospheric water vapor may be deposited either on the soil surface directly or indirectly to the foliar surface, with subsequent removal during rainfall events via the process of foliar washoff. Once in the soil, pesticides may volatilize into the air, enter the surface and groundwater by runoff or leaching, be taken up by plants and soil organisms, or stay in the soil (3–13).

As the use of pesticides in agriculture inevitably leads to exposure of nontarget organisms (including humans), undesirable side effects may occur on some species, communities, or on ecosystems as a whole. Therefore, the knowledge of pesticide environmental behavior, like adsorption to soil, leaching to groundwater, and volatility in the atmosphere, is of primary concern for an accurate assessment of the risk to the environment and humans (14).

SORPTION TO SOIL COMPONENTS (ORGANIC AND INORGANIC)

The soil is a complex medium whose physical, chemical, and biological properties are highly heterogeneous. The horizontal and vertical distribution of textural and structural elements in soil, and the differences in physicochemical properties, lead to variations in sorption, desorption, transport, and permeability of soil. The soil serves as a temporary or permanent sink for any chemical substance that hits its surface (12).

The sorption of pesticide on soil reduces its mobility; it is dependent on the physical and chemical properties of the soil as well as the molecular characteristics of the pesticide. Soil organic matter is the main soil constituent responsible for sorption of nonionic pesticides. In order to assess pesticide mobility, a sorption constant, based on the soil organic carbon (K_{oc}, $dm^3\ kg^{-1}$), can be used. Thus, $K_{oc} = K_d/oc$, where K_d is a measure of the extent of pesticide sorption by the soil and oc is the fraction of the organic carbon presents in the soil (15).

A significant proportion, typically ranging from 20% to 70%, of a pesticide or its metabolites may remain in the soil as “persistent residue,” sequestered in soil colloids (16).

The sequestration results in a diminution in the rate and extent of biodegradation by microorganisms in the field. The decline in toxicity and bioavailability as a result of the accompanying sequestration may represent a natural remediation process (17,18).

BIOLOGICAL AND CHEMICAL DEGRADATION

Losses of pesticide in the soil via microbiological and chemical pathways are collectively termed degradation. The rates of degradation can be characterized by a half-life (DT50). In soil, this rate increases with temperature and water content (19). The products of pesticide degradation, named metabolites, may have environmentally undesirable characteristics. Therefore, when the evaluation of the environmental fate of a pesticide is carried out, it is recommended to take the fate of metabolites into consideration.

VOLATILIZATION

Volatilization is a major cause of pesticide loss from target areas, particularly when pesticides are applied to the surfaces of soils or plants. Atmospheric transport and deposition are the major processes for distributing many pesticides around the world (20). The most appropriate criterion of the volatilization rate of a pesticide is Henry's law constant (K_h, dimensionless), the ratio of the vapor pressure to the water solubility (21,22). Compounds with K_h much greater than 2.5×10^{-5} are volatile, with volatility decreasing with time, whereas compounds with K_h much smaller than 2.5×10^{-5} are much less volatile, with volatility increasing with time (21–23).

RUNOFF AND LEACHING

Water may disperse pesticides into the environment via foliar washoff, surface runoff, and leaching. Runoff may contribute to pollution of surface water and leaching to contamination of groundwater.

Runoff is defined as the water and any dissolved or suspended matter it contains that leaves a plot, field, or small single-cover watershed in surface drainage (15). Runoff may include dissolved, suspended particulate, and sediment-adsorbed pesticides. Pesticides that remain at the soil surface for longer periods of time because they are strongly adsorbed and resistant to degradation and volatilization will be more susceptible to runoff, whereas incorporation into the soil will reduce runoff risk (24). Soluble pesticides may be more readily leached into the soil during the initial rainfall.

Pesticide leaching may cause groundwater pollution. The extent to which groundwater contamination occurs will depend on pesticides properties, soil characteristics, drainage rate, and water table depth. A combination of mobility and persistence determines whether a compound will be degraded during its residence time in the zone above the groundwater (21,25). Gustafson et al. (25) propose a single numerical index for predicting the water contamination potential of pesticides, the groundwater ubiquity score (GUS). It is defined as $GUS = \log(DT_{50}) \times (4 - \log(K_{oc}))$.

Pesticides detected in groundwater generally have GUS values exceeding 2.8, whereas compounds with GUS values below 1.8 were not detected in groundwater.

Rate of pesticide leaching in soil decreases with increasing organic matter content and depth of the surface zone with high biological activity. Moreover, the presence of macropores (cracks, worm holes, root channels) in soils enhances the hazard of pesticide leaching to the groundwater. Via these macropores, water and solutes can be transported rapidly to the subsoil and groundwater, bypassing much of the soil matrix (26).

PESTICIDES GROUPS

Over the last 20–25 years, a drastic change in the general characteristics of pesticides has occurred. They have changed from those with extremely low water solubility, strongly sorbed with long half-life to those that are more water soluble, only slightly sorbed, and less persistent. According to the chemical structure, pesticides can be classified as organochlorines, extremely hydrophobic with very high partition coefficient, and as organophosphates, carbamates, and pyretroids, which are more hydrophilic with lower partition coefficient value (27). Moreover, pesticides can be conveniently divided into classes, depending on which particular pest they are directed toward. Thus, the main groups are the herbicides, insecticides, and fungicides.

Pesticide properties, combined with the amount applied, determine how many pesticides can be transported along each route into the environment.

As was mentioned above, the low water solubility and the high partition coefficient of the organochlorine pesticides (OCP) confer to these pollutants a high environmental persistence. For this reason, the monitoring of these compounds is highly recommended.

WHAT ARE "ORGANOCHLORINE PESTICIDES"?

Organochlorine pesticides are important classes of environmental contaminants. Owing to their environmental persistence, toxic properties, and potential for bioaccumulation, they can be classified as "persistent organic pollutants" (POPs). Of additional concern is the ability of many POP compounds to undergo long-range atmospheric transport, developing from their semivolatile nature, which has resulted in their distribution in nearly all environmental compartments on a global scale, often far removed from their places of production and application (28–30).

OCP have been widely used throughout the world since the 1950s, although the use of many of these chemicals has been subjected to restriction and/or banned in many countries in recent years. However, the use of OCP in tropical and subtropical countries has been increasing in recent years in agricultural and public health programs (31). In addition to current usage, significant environmental burdens remain for many of these chemicals, which can act as secondary sources from which releases to the environment can occur, even decades after initial use (32).

OCP enter the soil by deposition from air, drift, or by washing off from plant surfaces during rainfall or irrigation. As a result of their strong hydrophobic (lipophilic) character, these compounds are mainly found associated with organic matter in soil and lipid tissues of organisms.

Organochlorine Pesticide Distribution in Relation to Land Uses: A Study Case

Knowledge of pesticide concentration and spatial distribution is necessary in order to asses the environmental impact of these compounds. A field study was carried out as part of an integral research about the dynamics of OCPs in a shallow lake watershed, Argentina, in relation to surrounding land uses (12).

The primary objective of this study was to assess the effect of land uses, such as agricultural, recreational (used for hiking, camping, and campfires), and natural (forested area that have never been cultivated) on the occurrence and distribution of organochlorine pesticides in soils from Argentina.

The organochlorine compounds analyzed included α-, β-, γ-, and δ-HCH; p,p′-DDT and its degradation products; p,p′-DDE and p,p′-DDD; heptachlor and its metabolite (heptachlor epoxide); aldrin; dieldrin; α- and γ-chlordane; and α- and β-endosulfan. Three depths were selected from each land use (0–15 cm, 15–30 cm, and 45–55 cm).

The results showed that, as was expected, the soil total organic carbon (TOC) decreased with depth for all three land uses. For the natural soils, total OCP concentrations were well correlated with the levels of TOC through the soil profile, whereas in agricultural and recreational soils, this correlation was not observed. These results imply that although TOC is known to be the most important OCP sorbent in soil surface horizons, additional factors exist, such as particle size characteristics and organic matter composition, and physicochemical characteristics of OCPs involved in pesticide retention.

Natural soil accumulated OCP to a greater extent than the other soils, even though it has never received direct OCP application. It is located on a hill that is clothed with trees and grass cover, which would act as an OCP trap, avoiding the OCP to leach up through soil profile, volatilization, and minimizing the risk of soil erosion. The highest concentrations were found in the surface layer, decreasing gradually with depth as a result of an undisturbed soil profile. DDT (472.5 ng/g dry wt), DDE (99.98 ng/g dry wt), and HCHs (25.44 ng/g dry wt) were the main pesticides found. Moreover, the levels of OCP in soils from an organic farm near to these natural soils ranged between 4–12 ng/g dry wt depending on the enrichment with pine needles collected from the surrounding natural areas (33,34). It is known that the waxy surfaces of pine needles has the ability to bioaccumulate airborne lipophilic pollutants (35). The soils from the organic farm could be considered as "natural soil," because it has never received direct pesticides application, either the lack of permanent vegetal cover and the soil disturbation by tillage practice lead to lower OCP levels. However, in both soils (natural and organic farm), a similar OCP pattern was found, as a consequence of the exposure to the same agricultural environment.

The values of natural soils are similar to those found by Dimond and Owen (36), (in Maine) for forest soils in sprayed areas during 1958 to 1967.

DDT was heavily used in Argentina, but it was totally banned in 1990. Commercial acaricide (Dicofol), still in use in Argentina, has been reported to contain a significant amount of DDT as manufacturing impurity and might represent a continuous source of DDT in the studied area. However, the high values of DDT found in these soils suggest that "new" DDT continues to be atmospherically transported from other countries, mostly tropical, where DDT is still in use (37–40).

High levels of heptachlor epoxide, a metabolite of heptachlor, were also found possibly because of physicochemical processes, such as photolysis and photodechlorination, and because of the abundance of microorganisms and edaphic biota in the surface layer of these natural soils, biotransformation processes are also assigned.

Chlordane residues were found in all the samples with the highest values being those from the surface layer of natural soil. The ratios γ-/α-chlordane found in these soils suggest direct application of technical chlordane in this zone.

Regarding to HCH group, the surface layer of the three studied soils showed higher concentration of lindane than the deeper depth, although there were not marked differences among depths. Lindane is one of the few OCPs analyzed in this work that is currently being used in Argentina.

Those soils dedicated to recreational activities showed the highest OCP values in the middle depth. During flooding periods, weathering of the surface layer occurs; as a result, a clearance of OCPs and predominance of HCHs are observed. The contribution of these compounds would come from the lagoon water; significant evidence exists that HCHs are the main pesticide group found in the aquatic biota of this waterbody (41).

In the agricultural soil during the last 25 years, a traditional tillage has been carried out in which a moldboard plough was used at a depth of 25–30 cm and in which straw was burned. This kind of tillage practice enhances organic carbon and fine particles (clay and silt) loss (30% in relation to natural soils), jointly OCP loss by volatilization or surface runoff. In addition to this management practice, the topography enhanced by slopes, wind, and water erosion during rain and irrigation periods justify these losses. Despite this, high values of OCPs in the surface layer were detected, with DDT being the main pesticide found, supporting the assumption that, because these pesticides (except lindane and endosulfan) are forbidden in Argentina, they would reach these soils via atmospheric transport.

Finally, soil characteristics, topographic position, and land uses justify the differences in OCP accumulation among the studied soils. The accumulation pattern represent a dynamic balance between deposition and erosion in association with wind-induced circulation, tillage management practices, and effects by natural processes, such as rainfall and flooding periods, according to different soils.

The results of this study suggest that natural areas should not be overlooked when assessing the sources and monitoring the occurrence of OCPs.

RELATION TO SURFACE WATER

The use of pesticides on agricultural land may have a negative effect on nearby aquatic ecosystems (12,13,42,43). This effect is determined by the specific chemicals used, the place and methods of application of the chemicals, and the various mechanisms available for transporting compounds to the waterbodies, with soil erosion being a major pathway (44). The extent to which pesticides move from treated soils depends on the physicochemical properties of the different active ingredients, the soil management practices, and the combination of environmental variables, such as frequency, intensity, and duration of precipitation (45).

Hydrophobicity and persistence are the main properties of pesticides that control their accumulation in sediment and aquatic biota. Thus, this fact confirms the current necessity of OCP monitoring.

CONCLUDING REMARKS

It is known that pesticides are released into the environment as formulated products that are designed to deliver the active ingredient to the target organisms. During the last forty years, a transition in the type of used pesticides has taken place. Thus, the highly persistent and bioaccumulatives organochlorine pesticides were replaced by organophosphates, carbamates, and pyretroids, currently used, with faster degradation in the environment.

However, despite pesticides undoubtedly contributing to increase farm productivity and human health, they also create several problems including widespread accumulation of residues with damage to wildlife, fisheries, beneficial insects, and even humans as a result of their dispersion toward nontarget areas. Today, the concern is the impact of those “environmental healthy pesticides,” because they are being linked to a high fish, wildlife, and bees toxicity, and some of them, because of their high water solubility, are potential groundwater contaminants. For these reasons, the monitoring of pesticide applications is necessary because it can be helpful in modifying the application techniques or use of particular chemicals to avoid adverse effects. Concern over possible impacts of pesticides on the environment and humans is not confined just to the highly industrialized nations. Developing countries are also experiencing the same concern exacerbated by lack of infrastructure and resources to deal with them.

BIBLIOGRAPHY

1. Himel, C.M., Loats, H., and Bailey, G.W. (1990). *Pesticides in the Soil Environment: Processes, Impact and Modeling*. H.H. Cheng (Ed.). Soil Science Society of America, Inc., Madison, WI, pp. 7–52.
2. Pimentel, D. (1995). *J. Agric. Environ. Ethics.* **8**: 17–29.
3. Rapaport, R.A. et al. (1985). *Chemosphere.* **14**: 1167–1173.
4. Rapaport, R.A. and Eisenreich, S.J. (1988). *Environ. Sci. Technol.* **22**: 931–941.
5. Bidleman, T.F., Widequist, V., Jansson, B., and Soderlund, R. (1987). *Atmosph. Environ.* **21**: 641–654.
6. Bidleman, T.F. et al. (1989). *Arctic.* **42**: 307–313.
7. Swackhamer, D.C., McVeety, B.D., and Hites, L.A. (1988). *Environ. Sci. Technol.* **22**: 664–672.
8. Patton, G.W., Walla, M.D., Bidleman, T.F., and Barrie, L.A. (1991). *J. Geoph. Res.* **96**: 10867–10877.
9. Kolpin, D.W., Thurman, E.M., and Goolsby, D.A. (1996). *Environ. Sci. Technol.* **30**: 335–340.
10. Miglioranza, K.S.B. et al. (1999). *Environ. Poll.* **105**: 91–99.
11. Miglioranza, K.S.B. et al. (2002). *Environ. Sci. Poll. Res.* **9**: 250–256.
12. Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2003). *Environ. Toxicol. Chem.* **22**: 712–717.
13. Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2003). *J. Soil Sed.* **4**: 264–265.
14. Van der Werf, H.M.G. (1996). *Agric. Ecosys. Environ.* **60**: 81–96.
15. Leonard, R.A. (1990). *Pesticides in the Soil Environment: Processes, Impacts and Modeling.* H.H. Cheng (Ed.). Soil Society of America, Book Series 2, Madison, WI, pp. 303–349.

16. Calderbank, A. (1989). *Environ. Contam. Toxicol.* **108**: 71–103.
17. White, J.C., Alexander, M., and Pignatello, J.J. (1999). *Environ. Toxicol. Chem.* **18**: 182–187.
18. Alexander, M. (1997). *Environmentally Acceptable Endpoints in Soil*. D.G. Linz and D.V. Nakles (Eds.). American Academic of Environmental Engineers, Annapolis, MD, pp. 43–136.
19. Walker, A. (1976). *Pestic. Sci.* **7**: 41–49.
20. Atlas, E.L. and Schauffler, S. (1990). *Long-Range Transport of Pesticides*. D.A. Kurtz (Ed.). Lewis, Chelsea, MI, pp. 161–183.
21. Jury, W.A., Spencer, W.F., and Farmer, W.J. (1984). *J. Environ. Qual.* **13**: 573–579.
22. Spencer, W.F. and Cliath, M.M. (1990). *Long-Range Transport of Pesticides*. D.A. Kurtz (Ed.). Lewis, Chelsea, MI, pp. 1–16.
23. Clendening, L.D., Jury, W.A., and Ernst, F.F. (1990). *Long-Range Transport of Pesticides*. D.A. Kurtz (Ed.). Lewis, Chelsea, MI, pp. 47–60.
24. Larsson, S.J. et al. (1995). *Chemosphere.* **31**: 3305–3321.
25. Gustafson, D.I. (1989). *Environ. Toxicol. Chem.* **8**: 339–357.
26. Beven, K. and Germann, P. (1982). *Water Resourc. Res.* **18**: 1311–1325.
27. Carlsen, R.F. and Smith, S.N. (1987). *Silent Spring Revisited*. G.J. Marco et al. (Eds.). American Chemical Society, Washington, DC, pp. 71–83.
28. Larsson, P., Jarnmark, C., and Sodergren, A. (1992). *Mar. Poll. Bull.* **25**: 281–287.
29. Iwata, H. et al. (1994). *Environ. Poll.* **85**: 15–33.
30. Muir, D.C.G. et al. (1996). *Environ. Sci. Technol.* **30**: 3609–3617.
31. Forget, G. (1991). *J. Toxicol. Environ. Health.* **32**: 11–31.
32. Finizio, A., Bidleman, T.F., and Szeto, S.Y. (1998). *Chemosphere.* **36**: 345–355.
33. Gonzalez, M., Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2003). *J. Agric. Food Chem.* **51**: 1353–1359.
34. Gonzalez, M., Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2003). *J. Agric. Food Chem.* **51**: 5024–5029.
35. Krauthacker, B., Romanic, S.M., and Reiser, E. (2001). *Bull. Environ. Contam. Toxicol.* **66**: 334–341.
36. Dimond, J.B. and Owen, R.B. (1996). *Environ. Poll.* **92**: 227–230.
37. Ostromogil'skii, A.K., Kokorin, A.O., and Afanas'ev, M.L. (1987). *Soviet Meteorol. Hydrol.* 28–35.
38. Ramesh, A. et al. (1991). *Environ. Poll.* **74**: 293–307.
39. Matthiessen, P. (1985). *Environ. Poll.* **10**: 189–211.
40. Bidleman, T.F. and Falconer, R.L. (1999). *Environ. Sci. Technol.* **33**: 206A–209A.
41. González Sagrario, M.A., Aizpún de Moreno, J.E., Moreno, V.J., and Escalante, A.H. (1998). *Environ. Sci.* **6**: 153–170.
42. Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2004). *Environ. Sci. Poll. Res.* **11**: 227–232.
43. Miglioranza, K.S.B., Aizpún de Moreno, J.E., and Moreno, V.J. (2004). *Water Res.* **38**: 1765–1772.
44. Munn, M.D. and Gruber, S.J. (1997). *Environ. Toxicol. Chem.* **16**: 1877–1887.
45. Cheng, H.H. (1990). *Pesticides in the Soil Environment: Processes, Impacts and Modeling*. Soil Society of America, Book Series 2, Madison, WI, p. 539.

REMEDIATION OF PESTICIDE-CONTAMINATED SOIL AT AGRICHEMICAL FACILITIES

William R. Roy
Illinois State Geological Survey
Champaign, Illinois

AGRICHEMICAL FACILITIES

An agrichemical or crop production facility is a commercial agricultural application business. These facilities store, sell, mix, load, and apply pesticides and fertilizers to fields for farmers and livestock producers. There are about 9,250 of these relatively small businesses in the United States and about 850 registered agrichemical businesses in Illinois. Pesticide mixing, loading, and handling sites are diverse and vary in the types of products handled at the site, the size of the facility, and the construction and type of containment (1).

Sources of Water and Soil Contamination

Bulk liquid and dry pesticides and fertilizers are typically mixed, repackaged, or transferred from one container to another. Accidental and incidental chemical releases such as spills, tank leaks, hose breaks, and transport accidents can result in contaminated water and soil at agrichemical facilities. Moreover, catastrophic events such as fires and floods can also result in soil, sediment, and water contamination.

Extent of Water and Soil Contamination

About 25% of the groundwater samples collected from 52 agrichemical facilities in Illinois contained 11 pesticides (2). The four most frequently occurring pesticides were bentazon, atrazine, trifluralin, and metolachlor. Nitrate was detected in 16 samples at concentrations ranging from 0.17 to 59 mg N/L. In additional sampling at four agrichemical facilities, 17 contaminants (nitrate, pesticides, and pesticide degradates) were detected. Atrazine, for example, detected in 15 of 22 samples, ranged from 0.54 to 61 μg/L. Environmental property audits conducted at 49 randomly selected agrichemical facilities in Illinois revealed that pesticides were detected in at least one soil sample collected at every site and at all sampling depths (3). The most frequently detected pesticides were alachlor, atrazine, metolachlor, trifluralin, pendimethaline, cyanazine, metibuzin, metribuzin DA, butylate, and α-BHC. Pesticides were detected to depths of 4.5 m. About 50% of all pesticide detections, however, were found in the gravel layer that serves as the parking lot/road base for each facility. In an extreme case, the gravel layer at an abandoned agrichemcial facility in Illinois contained 4.4% trifluralin by weight (4). Pesticide-contaminated soil and groundwater at retail agrichemical facilities have also been documented in Minnesota, California, Florida, Michigan, Wisconsin (1), and Arkansas (5).

Movement of Pesticides and Fertilizers

A portion of the pesticides and fertilizers in soil may leach into groundwater. Many of the commonly used pesticides and fertilizers are fairly mobile and can be leached as surface water infiltrates into the ground. The contaminated gravel, in which the bulk of pesticides may occur, poses a risk to groundwater quality (4). Hence, contaminated soil and gravel can adversely impact groundwater and possibly drinking water in rural communities. Because of this potential, it may be necessary to remediate contaminated soil to minimize its impact on water quality.

CLEANUP OBJECTIVES

Groundwater Cleanup Objectives

The values for groundwater-cleanup objectives depend on the specific chemical and the type of groundwater at the site. For example, in Illinois, Class I groundwater is called Potable Resource Groundwater, and Class II groundwater is General Resource Groundwater (6). A major distinction between the two types of groundwater is the hydraulic conductivity of water-bearing geologic materials: groundwater in an aquifer that has hydraulic conductivity equal to or greater than 1×10^{-4} cm/s is classified as Class I Groundwater. Therefore, Class II groundwater would be expected in less permeable materials. The groundwater quality standards for Class II groundwater are less stringent than those for Class I groundwater. Class II groundwater cleanup objectives are five times greater than those given for Class I groundwater. For example, the objective for atrazine in Class I groundwater is 3 μg/L, whereas it is 15 μg/L for Class II groundwater (6).

Soil Cleanup Objectives

The remediation of contaminated soil also requires establishing some type of cleanup goals or objectives. Such objectives are used to determine whether the soil is "contaminated" in a regulatory context. Depending on the specific type, agrichemicals are subject to various physicochemical and biological processes that can reduce their concentrations in soil and groundwater. Unlike other common contaminants such as chlorinated hydrocarbons, agrichemicals are *intentionally* released into the environment to promote plant growth and to limit the spread of weeds. A mixture of application-rate equivalents and risk-based processes to derive soil cleanup objectives have been promulgated in the Midwest. For example, the Illinois Department of Agriculture provides a default soil-cleanup objective (SCO) for surface soil (0 to 12 inches) and subsurface soil (12 to 60 inches) for each major pesticide used in Illinois. A facility also has the option of developing site-specific SCOs. These SCOs were derived from the application of soil screening levels (SSLs) developed by the U.S. Environmental Protection Agency (7,8). The resulting SCOs take into account the movement and chemical fate of a given chemical in soil–water systems. In the application of the SSL concepts, three physicochemical-fate mechanisms are considered: dilution in groundwater, adsorption by soil organic matter, and volatilization. Degradation by biotic or abiotic mechanisms is not considered. The SSL model provides a dynamic link between groundwater concentrations and soil concentrations. The default subsurface (12 to 60 inches depth) SCOs were calculated using Class I groundwater quality standards. The default surface SCOs were also based on Class I groundwater standards with the exception of acetochlor, alachlor, aldicarb, atrazine, cyanazine, dimethoate, disulfoton, phorate, simazine, and terbufos. For these pesticides, application-rate equivalents are used as SCOs. As given in (9), each SCO was based on pesticide label rates to the upper foot of soil or of a soil/gravel mixture.

EXCAVATION OF CONTAMINATED SOIL

Once the SCOs have been established, the areas and depths of contamination by agrichemicals at an agrichemical facility can be estimated. The delineation of the area and depth of material to be removed is not an exact science. In practice, the boundaries for excavation are subjectively interpolated between the locations where soil cores were collected. The boundaries for excavation may also, in part, be determined by any physical constraints such as facility buildings, railroad tracks, or other permanent structures. Once the contaminated materials have been excavated from the dealership, they must be disposed in an environmentally acceptable manner. There are three options available: land application or spreading, landfilling, and above-ground treatment.

OPTIONS FOR REMEDIATION

Land Application

Land application for the disposal of soils contaminated with agrichemicals is a common approach in the Midwest and may be the least expensive and most readily available disposal technique (10). It is based on the premise that if contaminated soil is spread in a thin layer on an agricultural field, the pesticides and/or fertilizers will dissipate to smaller and harmless concentrations. In practice, the pesticides in the contaminated material are applied at the application-rate equivalent that would be appropriate for that field and would be expected to biodegrade with time. If more than one pesticide or fertilizer is present, the application rate is based on the most limiting rate for the chemical detected in the contaminated soil sediment, or fill.

When pesticide-containing soil is land applied, it may be necessary to minimize the amount of coarse material in the soil before application. Attempting to spread soil that contains a significant amount of coarse material such as gravel can damage most spreading equipment and be objectionable to the landowner or farmer. A practical technique to remove the coarse material is to use field-scale screeners. The excavated, unsorted material is fed into some type of mechanical screener to separate the coarse debris from the finer soil. Depending on the size of the equipment and the mesh size of the screen, a screener can process between 50 to 150 tons per hour. The coarse material may be used as a road base, field roads, or placed back in the original excavation. The fine material is then land applied using spreading equipment. There is a relatively narrow window of opportunity in which site remediation and landspreading can occur: postharvest to preplanting. During the planting season, agrichemical facilities are in their busiest periods and are reluctant to allow any site excavation because it can hinder the operation of the facility and service to their customers. During the winter, it is advisable to avoid freezing conditions because the soil cannot be handled effectively if frozen.

Often, there is not enough room at an agrichemical facility for soil screening. In actual practice, the excavated soil is transported from the facility to an area where the soil can be screened and land applied. Compensation to local farmers or land owners for the use of land to establish a staging area for the screening process and for the agricultural fields can add to the costs of remediation.

Landfilling

Historically, pesticide-contaminated soil excavated at agrichemical facilities has been placed in landfills. When land application is not possible, using contaminated soil as a refuse cover or backfill may be an option. However, it may be more expensive to place the soil in a landfill. For example, if transportation costs are ignored, the average cost of landfilling pesticide-contaminated soil in the Midwest is about $25 per cubic yard compared with an average of $10 per cubic yard for land spreading. Moreover, limitations in landfill capacity may be a barrier to the disposal of contaminated soil in some areas.

Soil Washing

Although land application and landfilling are commonly used approaches for pesticide-contaminated soil at agrichemical facilities, other options are under study. It has been suggested that soil washing could be applied for a relatively wide range of pesticides (11). Soil washing is a treatment process in which contaminants are extracted from the excavated material using a solution containing chemical surfactants. For example, under laboratory conditions, a 3% solution of the nonionic surfactant polyoxethylene-4-lauryl ether was efficacious for removing atrazine from contaminated fill samples that were collected at an agrichemical facility (12). After 1 hour of mixing the fill–liquid suspensions, about 85% of the atrazine was extracted, compared with 26% removal by deionized water.

Bioremediation

Bioremediation is another option for remediating contaminated soil. Biodegradation is a process in which microorganisms transform or alter the structure and properties of chemical compounds by enzymatic biochemical reactions. Biochemical transformations can break down toxic, anthropogenic compounds into less toxic metabolites, and/or into inorganic compounds. Bioremediation is the application of biodegradation to reduce the concentrations of contaminants in soil by enhancing the physicochemical conditions for degradation, the creation of optimum conditions for remediating contaminated sites. For example, the biodegradation of trifluralin, metolachlor, and pendimethaline in gravel fill samples was promoted by mixing them with yard-waste compost under laboratory conditions (13). The initial concentration of trifluralin (2200 μg/kg) was reduced to about 505 μg/kg in a 1:1 fill:compost mixture after 40 days.

The basis for enhanced biodegradation is creating an environment where microbes are more numerous thereby increasing the probability of significant degradation. There are several critical environmental conditions that affect microbial activity: moisture content, oxygen, redox potential, pH, nutrients, temperature, and especially the supply of nitrogen and phosphorous. If the contaminated soil contains a large amount of carbon but little nitrogen and phosphorous, biodegradation will cease when available N and P are depleted. Therefore, fertilization of the contaminated fill materials may be required as a management technique to enhance biodegradation. For example, the addition of cornmeal can enhance the extent of biodegradation of pesticide-contaminated soil from an agrichemical facility when the material is land applied (14).

The addition of a mixture of cornmeal, surfactant, fertilizer, and clean soil, under laboratory conditions, increased the rate of biodegradation of atrazine, alachlor, and metalochor in gravel fill samples that were collected at three different agrichemical facilities (15). The fill samples were then either moistened with sterilized deionized water to 25% moisture content, or amended with the surfactant-fertilizer solution containing 0.1% oleic acid, urea, polyoxyethylene-12-octyl phosphate (modified from the oleophilic fertilizer Inipol EAP 22). After 8 weeks of incubation, the amount of atrazine recovered decreased by 55 to 74% in fill–soil samples amended with cornmeal and the surfactant-fertilizer mixture. Similar reductions of alachlor and metolachlor were also observed in those samples (Fig. 1). No significant degradation of atrazine, alachlor, and metolachlor was observed in the unamended and sterilized fill samples. During the same incubation period, however, the rate of reduction of atrazine, alachlor, and metolachlor was slower in samples amended with cornmeal alone. The biodegradation half-lives of atrazine in the fill samples amended with corncob and the surfactant-fertilizer mixture ranged from 3.9 to 6.8 weeks.

Phytoremediation

Phytoremediation is a technique in which plants, primarily hybrid poplar trees, legumes, and grasses, are planted in contaminated soil. The plants have the potential to take up and incorporate contaminants into plant tissues, enhance biodegradation in the root zone, and reduce the volume of groundwater available to transport contaminants off-site. There are few, however, published examples of field-scale applications of phytoremediation at agrichemical facilities. Poplar trees were planted at an agrichemical facility in Illinois to remove nitrogen and herbicides from shallow surface groundwater (Eric Aitchison, Ecolotree, Inc., personal communication). A total of 440 trees planted grew to a height of 15 feet in 17 months, and a significant volume of groundwater was apparently used by the trees. Phytoremediation may have a future,

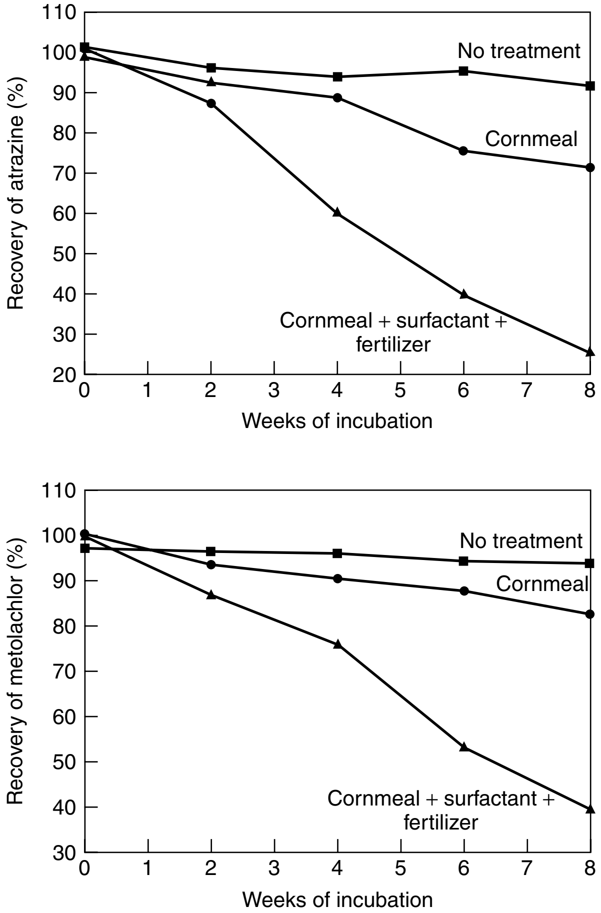

Figure 1. Amount of atrazine (top) and metolachlor (bottom) in soil–fill samples and the control sample (no treatment) in the presence of ground cornmeal and the cornmeal–soil–surfactant–fertilizer combination as a function of incubation time (15).

but additional research and development are needed. Phytoremediation, like bioremediation, may be too slow when rapid remediation is required. Phytoremediation may be suitable for inactive or abandoned agrichemical facilities. Moreover, relatively large concentrations of phytotoxic chemicals in soil and groundwater can limit plant growth (16).

OTHER APPROACHES

Other approaches for remediating pesticide-contaminated soil have been proposed such as incineration, solidification/stabilization, thermal desorption, radio-frequency heating, supercritical carbon dioxide extraction, and dechlorination by zero-valent iron (10,11,17,18). However, these approaches may not be cost-effective for a typical agrichemical facility. For example, incineration of soil may cost from $800 to $3,000 per cubic yard (17). However, if the contaminated soil also contains chemicals that are defined as hazardous under the Resource Conservation and Recovery Act, the soil may have to be incinerated (1).

THE SELECTION OF THE REMEDIAL APPROACH

Past handling practices have resulted in pesticide-contaminated soil and groundwater at retail agrichemical facilities. During about the last ten years, however, new containment regulations have been implemented in the Midwest which will hopefully reduce incidental releases of pesticides and fertilizers. For example, the Secondary Containment, Operational Area Containment, and Containment Management and Operations in Illinois (19) require the use of a containment system designed to intercept, retain, and recover operational and accidental spillage, leakage, wash water, and agrichemical residues. When contaminated soil, fill, or sediment is detected at an agrichemical facility during a site assessment, however, remediation may be required to minimize the adverse impact of the material on surface and groundwater. Techniques for remediating pesticide-contaminated soil is a current area of research, and new techniques and technologies are currently under study. However, land application and landfilling may be the most cost-effective, routine, and available technique to remediate contaminated soil at retail agrichemical facilities. Experience gained in the Midwest may be very useful for other domestic and international facilities.

BIBLIOGRAPHY

1. Buzicky, G., Liemandt, P., Grow, S., and Read, D. (1992). Agricultural chemical site remediation and regulations. *Pesticide Waste Management. Technology and Regulation*. J.B. Bourke (Ed.). American Chemical Society Symposium Series 510, Chapter 21.
2. Goetsch, W.D., Kirbach, G.C., and Black, W.F. (1993). Pesticides in well water and groundwater at agrichemical facilities in Illinois: an initial investigation. In: *Agrichemical Facility Site Contamination Study*. Illinois Department of Agriculture, Chap. 4.
3. Krapac, I.G., Roy, W.R., Smyth, C.A., and Barnhardt, M.L. (1995). Occurrence and distribution of pesticides in soil at agrichemical facilities in Illinois. *J. Soil Contamination* **4**: 206–226.
4. Roy, W.R., Krapac, I.G., and Chou, S.F.J. (1999). Chemical fate and transport of atrazine in soil gravel materials at agrichemical distribution facilities. *J. Soil Contamination* **8**: 365–387.
5. Senseman, S.A., Lavy, T.L., and Daniel, T.C. (1997). Monitoring groundwater for pesticides at selected mixing/loading sites in Arkansas. *Environ. Sci. Technol.* **31**: 283–288.
6. *Illinois Administrative Code. Title 35: Environmental Protection*. Chapter I: Pollution Control Board. Part 620. Groundwater Quality.
7. U.S. Environmental Agency. (1996). *Soil Screening Guidance: User's Guide*. EPA/540/R-963505.
8. U.S. Environmental Agency. (1996). *Soil Screening Guidance: Technical Background Document*. EPA/540/R-95/128.
9. *Illinois Administrative Code. Title 8: Agriculture and Animals*. Chapter I: Department of Agriculture. Part 259. Agrichemical Facility Response Action Program.
10. Felsot, A. (1996). Options for cleanup and disposal of pesticide wastes generated on a small scale. *J. Environ. Sci. Health* **B31**: 365–381.

11. Koustas, R.N. and Fischer, D. (1998). Review of separation technologies for treating pesticide-contaminated soil. *J. Air Waste Manage. Assoc.* **48**: 434–440.
12. Roy, W.R., Chou, S.F.J., and Krapac, I.G. (1995). Fate and transport of atrazine in fill materials at agrichemical facilities: A status report. In: *Proc. 5th Annu. Conf. Illinois Groundwater Consortium*, Makanda, IL, March 29–30, 1995, pp. 117–131.
13. Cole, M.A., Liu, X., and Zhang, L. (1994). Plant and microbial establishment in pesticide-contaminated soil amended with compost. In: *Enhanced Biodegradation of Pesticides in the Environment*. K.D. Racke and J.R. Coats (Eds.). American Chemical Society, Washington, DC, pp. 249–268.
14. Felsot, A.S., Mitchell, J.K., and Dzantor, E.K. (1994). *Use of Landfarming to Remediate Soil Contaminated by Pesticide Waste*. Report no. HWRIC RR-070, Hazardous Waste Research and Information Center, Champaign, IL.
15. Roy, W.R., Krapac, I.G., and Chou, S.F.J. (1996). Fate and Transport of Atrazine in Fill Materials at Agrichemical Facilities: An Update. In: *Proc. 6th Annu. Conf. Illinois Groundwater Consortium*, Makanda, IL, March 27–28, 1996, pp. 261–270.
16. Arthur, E.L., Perkovich, B.S., Anderson, T.A., and Coats, J.R. (1999). Degradation of an atrazine and metolachlor herbicide mixture in pesticide contaminated soils from two agrochemical dealerships in Iowa. *Water Air, Soil Pollut.* **113**: 1–17.
17. Risatti, J.B. and Zagula, S. (1993). Potential cost effective technologies for remediation of Illinois agrichemical facilities. In: *Agrichemical Facility Site Contamination Study*. Illinois Department of Agriculture, Chap. 7.
18. Comfort, S.D., Shea, P.J., Machacek, T.A., Gaber, H., and Oh, B.T. (2001). Field-scale remediation of a metolachlor-contaminated spill site using zerovalent iron. *J. Environ. Qual.* **30**: 1636–1643.
19. *Illinois Administrative Code. Title 8: Agriculture and Animals*. Chapter I: Department of Agriculture. Part 255. Agrichemical Facilities.

PESTICIDE OCCURRENCE AND DISTRIBUTION IN RELATION TO USE

Xuefeng Chu
Grand Valley State University
Muskegon, Michigan

GENERAL RELATIONSHIP BETWEEN PESTICIDE OCCURRENCE AND USE

In the United States, about 1 billion pounds of pesticides are used each year to control weeds, insects, and other organisms and about 80% of them are used in agriculture (1). Widespread use of pesticides over the past several decades has led to their occurrence in groundwater, surface water, aquatic biota and sediment, and the atmosphere (2). In 8200 sample analyses conducted by the National Water Quality Assessment (NAWQA) program during 1992–1996, 58 pesticides were detected at least once at or above 0.01 μg/L in both surface and groundwaters in 20 of the nation's major hydrologic basins (3). The NAWQA studies also revealed that more than 90% of water and fish samples from all streams contained one or, more often, several pesticides and about 50% of the sampled wells contained one or more pesticides; the highest detection frequencies were in shallow groundwater. In particular, the top 15 pesticide compounds found in water are among those with the highest current use, although some organochlorine compounds (e.g., DDT) heavily used in the 1960s were found most often in fish and sediments. The most frequently used pesticides occurred in geographic and seasonal patterns that mainly correspond to distributions of land use and associated pesticide use (4). Clearly, a significant relation can be expected between the use of pesticides and their occurrence in both surface and groundwaters. However, the extent of their occurrence and spatial and temporal distributions in surface and groundwater (how, when, and where the pesticides occur) may vary significantly.

ENVIRONMENTAL PATHWAYS OF PESTICIDES VERSES THEIR OCCURRENCE AND DISTRIBUTION

The environmental pathways of pesticides determine their occurrence and distribution. Following release into the environment, pesticides can be washed off the target application fields into adjacent surface waterbodies (e.g., streams, rivers, lakes) by runoff and erosion, leached by percolating water down through the vadose zone into the underlying aquifers, volatized into the atmosphere, taken up by plants, and accumulated in aquatic organisms through the food chain and even in human bodies. Thus, pesticides follow complex pathways and patterns of occurrence and distribution in the environment (Fig. 1).

TRANSPORT AND TRANSFORMATION PROCESSES OF PESTICIDES VERSES THEIR OCCURRENCE AND DISTRIBUTION

Pesticides, existing in three phases (dissolved, adsorbed, and vapor phases) in the environment, undergo a series of physical, chemical, and biological transport and transformation processes such as advection, dispersion, degradation, sorption, volatilization, and partitioning between dissolved and vapor phases in surface and groundwater (Fig. 1). These essential transport and transformation processes dominate pesticide occurrence and distribution in the environment. Due to the complexity of these processes, the occurrence and distribution of pesticides in surface and groundwater can be extremely variable and are affected by numerous factors such as pesticide use, pesticide properties, land use, soil types, topographic features, agricultural practices, climatic conditions, hydrology, and hydrogeologic settings. Whether or not pesticides occur in a certain part of the environment (a specific river reach, soil layer, or aquifer) depends on three essential elements: (1) availability of pesticide, (2) availability of flowing water, and (3) accessibility/pathways.

VARIABILITY IN THE RELATIONSHIP BETWEEN PESTICIDE OCCURRENCE AND USE

The relationship between pesticide occurrence and use varies for different pesticides and geographic regions.

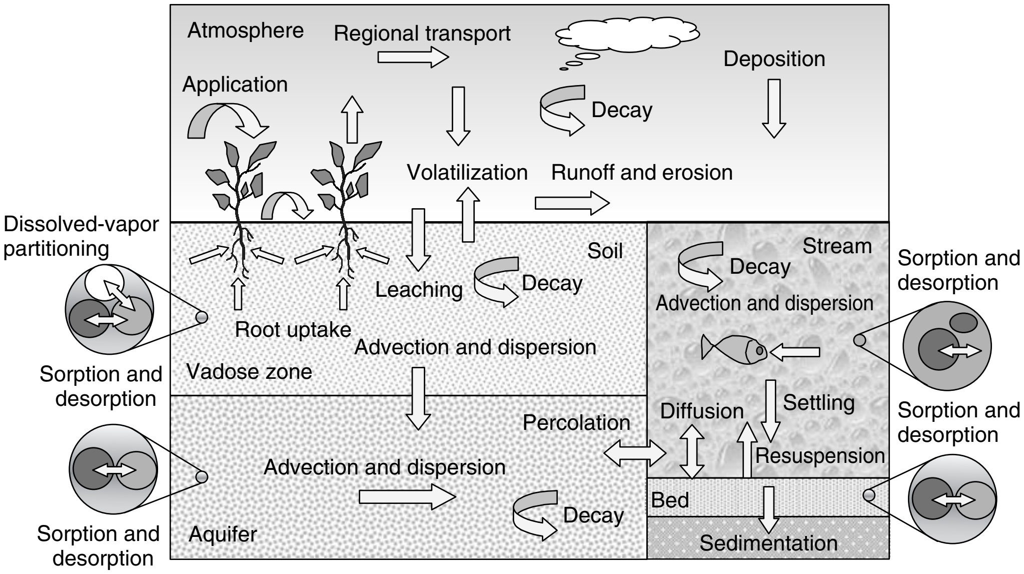

Figure 1. Fate and transport of pesticides and their occurrence in the environment.

Even for a specific pesticide at one site, the relationship changes with time due to the variability and complexity of hydrology. In addition, the relationship for surface water is different from that for groundwater due to their distinct hydrologic characteristics and pesticide fate and transport processes. Surface and groundwater systems have dissimilar timing of pesticide occurrence, temporal and spatial distributions, and magnitude of pesticide exposure levels. A pesticide application may result in a peak concentration in streams and shallow soils in hours, whereas it may take days, even months, to reach a peak exposure level in a groundwater system. The peak concentrations in groundwater can also be much smaller than those in surface water and shallow soils. A time shift of arrival of peaks can often be observed in the transport direction in both surface and subsurface water systems.

FIVE KEY PESTICIDE-USE ISSUES RELATED TO OCCURRENCE AND DISTRIBUTION

In terms of pesticide use, five key issues that affect pesticide occurrence and distribution need to be addressed: (1) which pesticide is applied (type and properties of pesticide), (2) how much and how often the pesticide is applied (amount and frequency of application), (3) when the pesticide is applied (timing of application), (4) how the pesticide is applied (method of application), and (5) where the pesticide is applied (domain of application). The types of pesticides that possess specific physical, chemical, and biological properties affect their occurrence and distribution directly by altering their environmental persistence, mobility, and other fundamental transport processes. The amount of applied pesticides determines the overall level of pesticides potentially available for occurrence in surface and groundwater, and the application frequency may determine the possibility and extent of pesticide buildup. The timing and method of application of pesticides affect occurrence primarily by changing their environmental pathways and thus spatial and temporal distributions and exposure levels. Additionally, the overall patterns of pesticide occurrence and distribution vary with pesticide-treated crops (application domain) that have different growing periods and canopy characteristics and require different soil–water–pesticide management practices.

EFFECTS OF USE OF PESTICIDES ON THEIR OCCURRENCE AND DISTRIBUTION

Effects of Pesticide Properties on its Occurrence and Distribution

The capabilities of sorption and volatilization of pesticides, characterized by distribution coefficient and Henry's law constant, determine their existing forms (dissolved, adsorbed, or vapor phases) and thus their mobility. These two processes and the degradation of pesticides are critical to their occurrence and distribution. Pesticides of lower sorptive capability tend to move quickly with flowing water. The spread of this type of pesticide throughout the environment is, to a great extent, determined by water movement. In contrast, long-lived pesticides of strong sorption, such as DDT, may last decades in the environment (4).

Effects of Application Quantity and Frequency on Pesticide Occurrence and Distribution

Frequent applications of pesticides increase the total amount of pesticides released into the environment and hence endanger the vulnerable hydrologic system. In particular, the accumulative effect of long-lived and strongly sorbed pesticides is often of great concern.

Effects of Application Timing on Pesticide Occurrence and Distribution

The combined timing of pesticide application and rainfall/irrigation can dominate exposure levels of some organophosphate insecticides in both surface and subsurface environments (5). A pesticide application, immediately followed by rainfall or overcanopy irrigation, provides an extremely high potential for the pesticide to transport from target fields to nearby surface waterbodies with runoff and erosion and to leach into soils and the underlying groundwater.

Effects of Application Methods on Pesticide Occurrence and Distribution

Pesticides can be applied by different methods, such as over-canopy spray, under-canopy ground surface spray, and soil-incorporated application. Application methods of pesticides affect their occurrence in the environment by changing their initial distribution and hence altering their following transport–transformation pathways and environmental fate. In addition, irrigation methods (e.g., over-canopy and under-canopy irrigation) and their combination with various pesticide application methods can be critical to pesticide occurrence and distribution. High variability of concentrations of some pesticides during irrigation periods can be attributed to differences in irrigation frequency and water management (6).

SUMMARY

The occurrence and distribution of pesticides in the environment are closely related to their use and are determined by a series of pesticide transport and transformation processes. The occurrence and use of pesticides follow extremely complex and dynamic patterns that are affected by numerous factors related to pesticide use, hydrology, and agricultural activities. Different ways of pesticide use (application quantity, frequency, timing, and method) can lead to distinct environmental fates of pesticides in terms of their occurrence, extent, as well as spatial and temporal distributions.

BIBLIOGRAPHY

1. U.S. Geological Survey. (2000). Pesticides in stream sediment and aquatic biota. U.S. Geological Survey Fact Sheet 092-00.
2. Barbash, J.E., Thelin, G.P., Kolpin, D.W., and Gilliom, R.J. (1999). Distribution of major herbicides in ground water of the United States, U.S. Geological Survey Water-Resources Investigations Report 98–4245.
3. U.S. Geological Survey. (1998). Pesticides in surface and ground water of the United States: Summary of results of the national water quality assessment program (NAWQA), Pesticides National Synthesis Project.
   Available at: http://ca.water.usgs.gov/pnsp/allsum/index.html
4. U.S. Geological Survey. (1999). The Quality of our nation's waters—nutrients and pesticides. U.S. Geological Survey Circular 1225.
5. Chu, X. and Mariño, M.A. (2004). Semidiscrete pesticide transport modeling and application. *J. Hydrol.* **285**(1–4): 19–40.
6. Domagalski, J.L. (1997). Results of a prototype surface water network design for pesticides developed for the San Joaquin River Basin, California. *J. Hydrol.* **192**: 33–50.

ASSESSMENT OF POLLUTION OUTFLOW FROM LARGE AGRICULTURAL AREAS

Ramakar Jha
K.K.S. Bhatia
National Institute of Hydrology
Roorkee, Uttaranchal, India

C.S.P. Ojha
Indian Institute of Technology
Uttaranchal, India

The pollution from large agricultural areas in general is discussed. The potential of remote sensing and geographical information system as input for NPS pollution assessment is emphasized. Mathematical models commonly used for assessing of nonpoint source (NPS) pollution are also described. Assessment of nutrient pollution from three agricultural watersheds of a typical Indian river basin ranging from 2–5 km^2 is considered as an illustrative example.

INTRODUCTION

The sources of surface water pollution typically fall into one of two categories: point-source pollution and nonpoint-source pollution. The term point-source pollution refers to pollutants discharged from one discrete location or point, such as an industry or municipal wastewater treatment plant, to surface water resources. The term nonpoint-source pollution refers to the discharge of pollutants, which cannot be identified as coming from one discrete location or point. The pollution outflow from agricultural areas is known as nonpoint source (NPS) pollution or diffuse pollution. NPS pollution occurs when rainfall, snowmelt, or irrigation water runs over land or through the ground, picks up pollutants, and deposits them in rivers, lakes, and coastal waters or introduces them into the groundwater. Sediment, nutrients, and pesticides are common agricultural pollutants, which cause surface water and groundwater quality problems (1–3).

Estimation of NPS pollution from cropland runoff is an important feature of regional water quality planning. The primary objective of NPS pollution estimation is to protect or restore the designated use of a water resource by reducing pollutant delivery to the water resource. All nonpoint sources of pollution are not equal. Many nonpoint sources of pollution are insignificant, whereas other sources contribute substantially to water resource impairment. Topographic, hydrologic, and agronomic factors often combine to make some nonpoint sources more detrimental to the beneficial use of water resources than others. Therefore, a method or strategy to identify and prioritize for treatment NPS areas, which are more detrimental than others, is desirable. A land treatment strategy should be developed to guide the selection and implementation of best management practices (BMPs). The BMP approach involves identifying and implementing land use practices in agricultural areas that prevent or reduce nonpoint pollution. In the case of erosion-sedimentation, many such practices are well known for agricultural activities. However, for some other types of pollutants, the BMP may not be known. In such instances, research that relates land use to water quality is needed.

ROLE OF GEOGRAPHICAL INFORMATION SYSTEM (GIS) AND REMOTE SENSING FOR NPS POLLUTION ASSESSMENT

Like any environmental phenomena, there is a spatial dimension to the management of water quality and the control of nonpoint-source pollution in agricultural watersheds. Understanding the spatial relationships among the various pollution sources in a watershed is critical to the successful implementation of the best agricultural management practices. A geographical information system (GIS) enables the effective integration, management, and analysis of disparate data sets related to chemicals, soils, climate, topography, land cover, and land use. These data sets are important driving variables for a nonpoint pollution model (4).

Several strategies for integrating distributed water quality models and GIS have emerged during the past few years (5–7). In most applications, the GIS was used primarily to generate model input data and to display output data from the model—an approach usually referred to as loose coupling of a model and the GIS (8–10). In loose coupling, there are two options: (1) loose coupling through interchange of data files in ASCII format between the model and the GIS and (2) loose coupling using a common binary file (10). An interface program (e.g., pre- and postprocessor) is normally used to convert and organize the GIS data in the form that the model requires. The loose coupling can often be time-consuming and problematic, particularly, for large watersheds. The loose coupling also, depends on the data structure of the GIS and the data file format specified by the water quality model (11,12).

The second strategy of coupling the GIS and water quality models involves close coupling in which the control programs in the GIS software are slightly modified to provide an enhanced environment for data transfer between the model and the GIS database (8–10). The options for loose and close coupling have significant overlap, depending on the characteristics of the water quality model. However, in close coupling, information is passed between the model and the GIS via memory-resident data models rather than external files. This improves the interactive capabilities and performance of closely coupled modeling environments.

The third strategy of coupling water quality models and the GIS is tight coupling or full integration (8–10). This strategy is based on incorporating the functional components of one system (e.g., the model) within the other, thereby eliminating the use of interface programs. The most widely used option of tight coupling involves developing tightly coupled seamless interfaces between the model and the GIS. The major benefits, of tight coupling strategy include improved performance, enhanced user interface, and increased problem-solving capabilities through efficient analysis of "what if?" scenarios. There is also little or no redundancy in the development process because the systems can leverage off each other.

Data acquired from remote sensors on aircraft and satellites offer great promise in providing numerous model inputs. Water quality models used to determine nonpoint source pollution vary in type and extent of land use for which these are best suited. Water bodies, urban areas, forests, cropland, and bare soils or rock can generally be differentiated very well. Individual species of vegetation often can be separated, surface soil properties can be delineated, and the percentage of vegetation or residue cover on the land can sometimes be estimated. The spectral, spatial, and temporal characteristics of remotely sensed data make those data useful for models that require knowledge of land use, cover, and conditions when calculating nonpoint-source pollutant loads. The relevant studies are available in many textbooks and the literature.

MATHEMATICAL MODELS FOR ESTIMATING POLLUTANTS FROM AGRICULTURAL AREAS

Mathematical models have been used extensively since the late 1960s to estimate pollutant outflows from agricultural areas. The term mathematical model may be new, but all of us have used mathematical models since we took algebra in school. Such models are mathematical representations of physical, chemical, biological, social, economic, and related processes. The advancement in computers, has made it possible now to evolve different mathematical approaches (from simple linear regression to complex differential equations) for obtaining favorable outcomes. Some of the mathematical models used for NPS pollution estimation are as follows: (1) SWMM (13), (2) HSPF (14), (3) ANSWERS (15), (4) CREAMS (16), (5) SWAM (17), (6) NTRM (18), (7) SWRRB (19), (8) SHE (20), (9) AGNPS (21), (10) SWAT (22), and (11) ANSWER-2000 (23).

Most of the mathematical models cited before simulate hydrologic, chemical, and physical processes involved in the transport of sediment, nutrients, and pesticides. In general, mathematical models for estimating NPS are divided into two classes: (1) a temporal class that includes event simulation and continuous simulation and (2) a spatial class that includes lumped parameters and distributed parameters.

In the temporal class, the continuous simulation models are preferred because they simulate all NPS pollution producing events and can be used to estimate total maximum daily loads and seasonal variations in NPS loading. However, in the spatial class, distributed parameter models are preferred. The principal advantage of distributed parameter models over lumped parameter models is that distributed parameter models represent the spatial variability of watershed features (soil, topography, land use etc.) more accurately. Moreover, the distributed parameter models can also estimate pollutant losses at different locations within the agricultural area. Some of the most commonly used distributed parameter

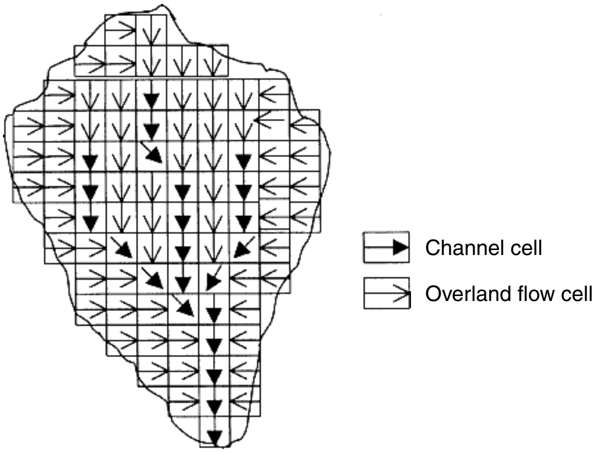

Figure 1. Mixed cell and routing process in overland and channel flow.

continuous simulation models are AGNPS and ANSWER-2000. The processes involved in these models are explained next.

In the beginning of the modeling process, a large agricultural area is divided into smaller grids or cells. A typical arrangement of such cells is shown in Fig. 1. The flow, sediments, and nutrients available at the beginning of the cell, coupled with those generated within the cell, are used to assess these values at the outlet of the cell. Subsequent routing of these values can lead to the estimation of flow as well as the NPS pollution load at the outlet of the agricultural watershed.

In a hydrological process, the rainfall added to an agricultural area in the beginning is intercepted by the canopy cover, evapotranspirated or infiltrated. Infiltration capacity is very high initially because of movement of water to the unsaturated zone, and water within this layer percolates downward. Water is discharged to the channel drainage system, when infiltration capacity is exceeded by the water supply and the depression storage has been satisfied. The total inflow to the channel drainage system is a combination of surface runoff, interflow, and base flow, which is calculated by using a water depletion function to gradually diminish it. A few specific relevant equations for the computation of different flow components are shown in Appendix 1. However, many more equations can be obtained from the literature.

The sediment transport capacity and potential sediment are important variables of nonpoint source pollution from a large agricultural area. As seen from Appendix 1, the sediment transport capacity is based on a sediment concentration ratio, c, estimated with a shear–stress relationship between the dominant flow shear on the soil and the critical stress based on the Shields criteria (24). The potential sediment supply due to rainfall can be calculated from an empirical relationship for rainfall energy rate adapted from the Universal Soil Loss Equation (USLE) method. The method accounts for the effects of canopy cover by introducing cover factors into the calculations of soil detachment due to rainfall and runoff. Similar models available in the literature can also be applied for estimating the sediment transport capacity and potential sediment.

The mathematical approaches used for the NPS water quality component in this chapter are widely used NPS models (Appendix 2). These were developed for the CREAMS model (25). Young et al. (20) modified the algorithms applied in the CREAMS model and created the AGNPS model, suitable on watershed scale. The nutrient simulation is normally divided into two parts to handle the soluble nutrient in the runoff and in the sediment separately (21). For the soluble part, the general assumption is that the rate of change in concentration of soluble nutrients at the surface is proportional to the difference between existing concentrations and concentration in rainfall. The available nutrient content at the surface is a result of combining the residual nutrient at the surface with the amount from fertilizer application. The available nutrient due to rainfall is estimated from the nutrient concentration in rainfall and given as input data. The movement rates are evaluated using nutrient leaching and runoff extraction coefficients given as input data for the model.

The mathematical equations are used separately for each grid or cell where the total runoff, sediment, and nutrient quantities are calculated by adding the surface runoff contributions from various grids or cells to the base flow. The routing through the channel system is achieved using a storage-routing technique based on continuity. Deposition of sediment and decay of nutrients are estimated using fractions of the transported mass. Mixing cell models, like other transport models, is subject to numerical dispersion affected in part by cell size and the assumption of complete mixing. Appendix 2 summarizes some of the commonly used equations.

ASSESSMENT OF NUTRIENTS FROM A TYPICAL INDIAN RIVER BASIN

To calibrate and test the mathematical equation shown in Appendixes 1 and 2, the nonpoint-source pollutants from three agricultural watersheds, ranging from 2–5 km^2, of Kali Basin, western Uttar Pradesh, India have been analyzed. River Kali in western Uttar Pradesh, India, has a significant socioeconomic value for nearby areas. It receives many point and nonpoint sources of pollution (26–29). The study area (Fig. 2) is part of the Yamuna basin in the Indo-gangetic Plains, composed of Pleistocene and subrecent alluvium, and lies between 29° 13′ to 30°N latitude and 77° 35′ to 77° 45′ E longitude. The mean rainfall across the basin is 1000 mm, which occurs mainly during the monsoon period. The basin area lies between elevations from 276 m above mean sea level to 221 m above mean sea level, and the major land use is agriculture. The soils of the area are loam to silty loam and are normally free from carbonates. Agricultural waste is also an important factor contributing to the pollution of the river water.

The basin boundary, drainage pattern, contour maps, spot heights and built-up area maps for River Kali, were digitized using ILWIS—GIS. For the analysis, the spatial data set available in grids or cells has been transformed

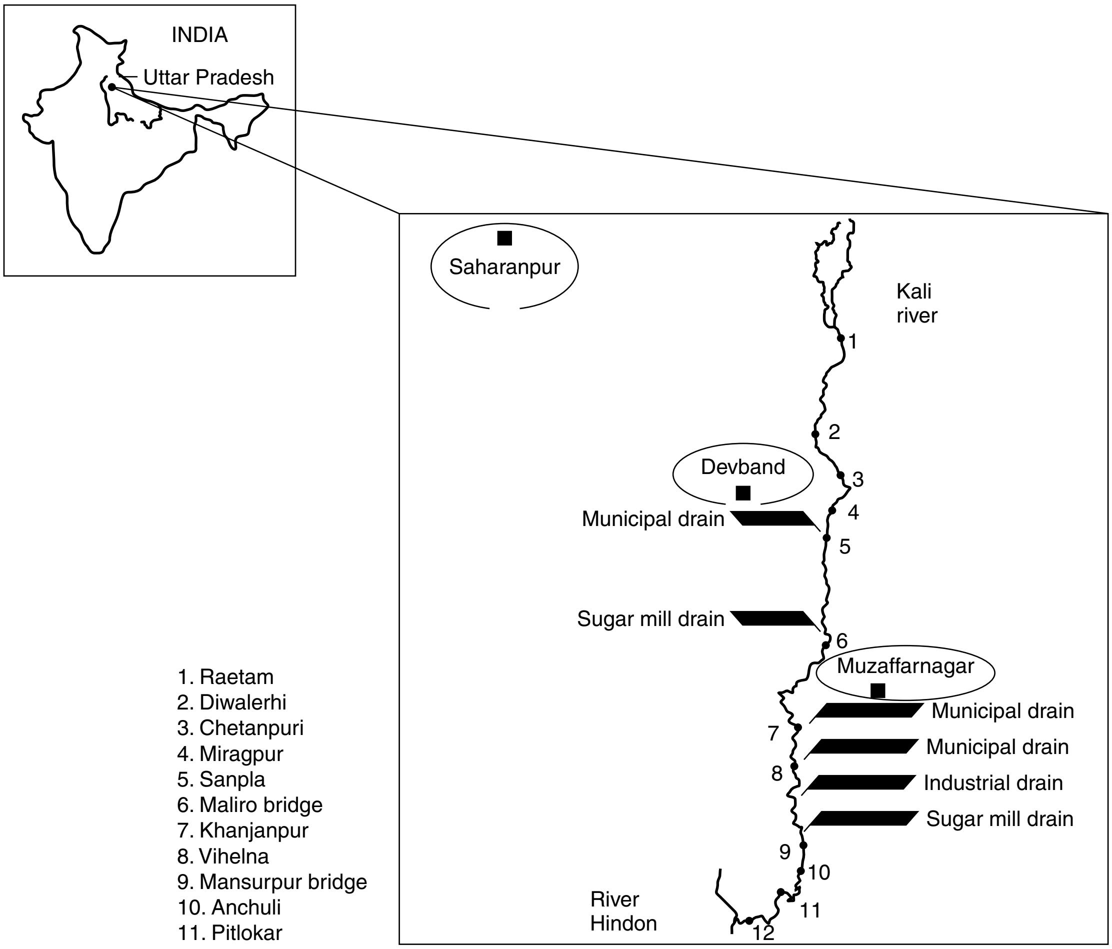

Figure 2. Location of River Kali in India.

to a digital elevation model (DEM), followed by filtering to arrive at a slope map, or a flow path map. For delineating land use/land cover maps, classification and quantification of different crops, and its temporal variation, IRS LISS III multispectral imageries and IRS-PAN data for the study period have been used. Ground truth verification was done to verify the results obtained using remote sensing data.

Further, extensive sampling from different sampling points giving a total of 576 data sets for nitrate (NO_3) and orthophosphate (o-PO_4) was done from March 1999 to February 2000. The nutrients were analyzed following a standard method (30). The NPS nutrient (NO_3 and o-PO_4) values obtained using mathematical equation based on the mixing cell concept, given in Appendixes 1 and 2, were compared with the observed nutrient values for all four agricultural watersheds (Fig. 3). It was found that the coefficient of determination, r^2, between observed and computed values of nutrients was greater than 0.85 in all four agricultural watersheds. It was also found that the error estimate viz., standard error, was less than 2 for all cases.

SUMMARY

To estimate pollution from a large agricultural area, distributed parameter and continuous simulation models are preferred because they simulate pollution producing events, total maximum daily pollution load, its seasonal variations, and represent the pollutant losses at different locations within the agricultural area. Remote sensing and GIS are important tools coupled with the mathematical models to estimate pollution from a large agricultural area. An illustrative case study from India demonstrates that NPS pollution can be computed well using the mixing cell concept with remote sensing and GIS support.

BIBLIOGRAPHY

1. US EPA. (1984). *Report to Congress: Non-Point Source Pollution in the United States*, U.S. Government Printing Office, Washington, DC.
2. US EPA. (1986). *Pesticides in ground water: Background document 8*. Office of ground water protection. Washington, DC.
3. Gould, G.A. (1990). Agriculture, non-point source pollution and federal law. *University of California at Davis Law Review* **23**(3): 461–471.
4. Joao, E.M. and Walsh, S.J. (1992). GIS implications for hydrologic modelling: simulation of non-point pollution generated as a consequence of watershed development scenarios. *Computers, Environment and Urban Systems* **16**: 43–63.

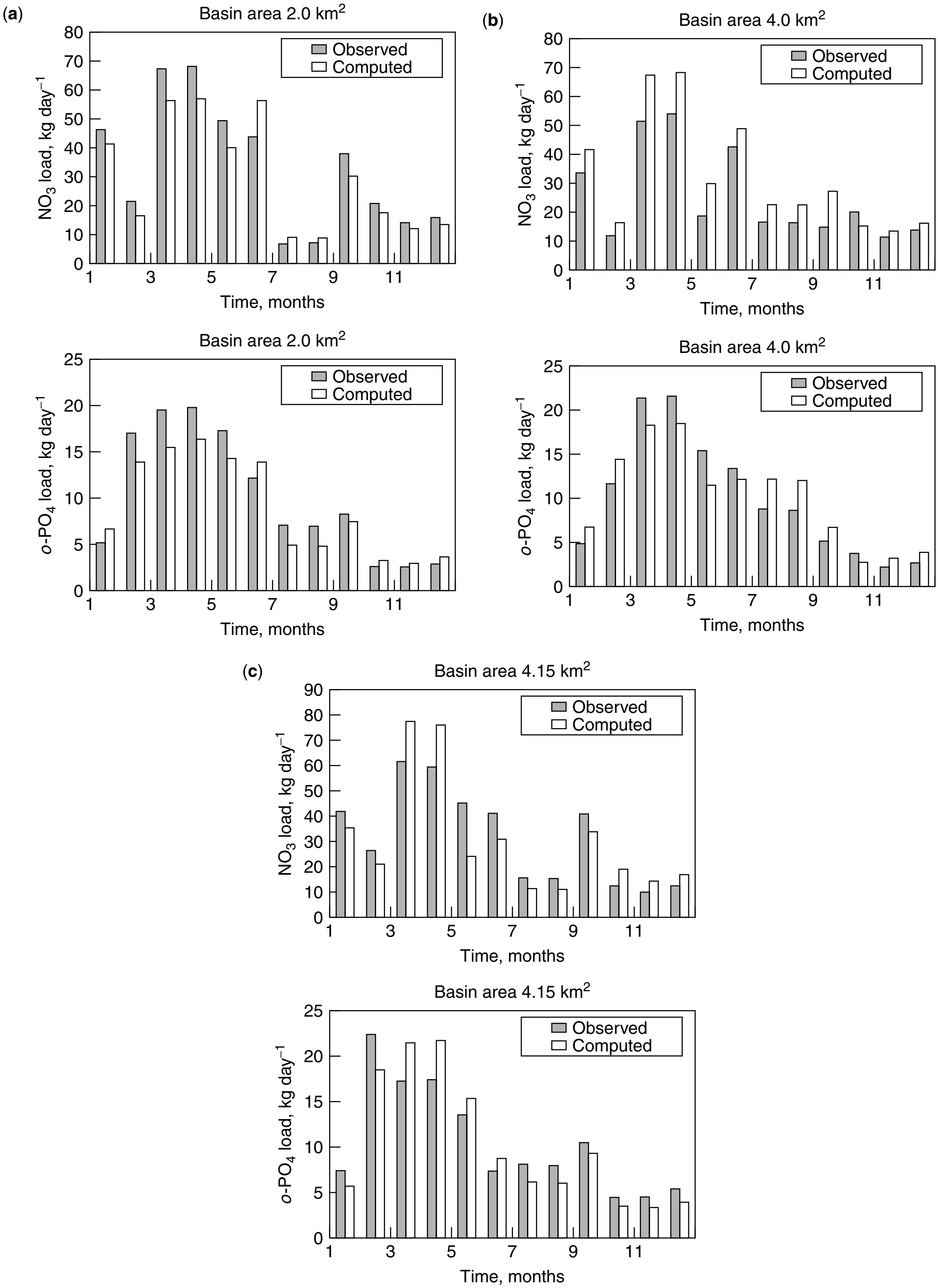

Figure 3. Observed and computed NPS nutrients from River Kali, India.

5. Maidment, D.R. (1993). GIS and hydrologic modelling. In: *Environmental Modelling with GIS*. M.F. Goodchild, B.O. Parks, and L.T. Steyaert (Eds.). Oxford University Press, New York, pp. 147–167.

6. Steyaert, L.T. and Goodchild, M.F. (1994). Integrating geographical information systems and environmental simulation models—A status review. In: *Environmental Information Management and Analysis*. W.K. Michener, J.W. Brunt, and

APPENDIX 1

Specific Equations for Different Variables Used in NPS Estimation

Variables	Author(s)	Governing Equation^a
Overland flow	Manning's formula	$Q_r = \frac{1}{R_3}(D_1 - D_S)^{5/3} S_1^{1/2} A$
Interflow	(31)	$Q_{\text{int}} = R_{\text{ec}}(W_{\text{ac}} - R_{\text{et}})$
Infiltration and surface detention	(32)	$\frac{dF}{df} = k\left[1 + \left(\frac{(m - m_0)(P_{\text{ot}} + D_1)}{F}\right)\right]$
	(33)	$P_{\text{et}} = \alpha \frac{s(T_a)}{s(T_a) + \zeta}(K_n + L_n)\frac{1}{\rho \lambda_y}$
Interception	(34)	$V = (S_i + C_p E_a t_R)(1 - e^{-kP})$
Surface storage	(31)	$D_S = S_d(1 - e^{-kP_e})$
Evapotranspiration	(35)	$P_{\text{et}} = 0.0075 R_a C_1 \delta_t^{1/2} T_{\text{avg}}$
Transport capacity	(36)	$Y_C = 2.65 pcr_f$
Shear-stress relationships	(36)	$\tau_d = \frac{\beta}{\beta + 1}\left(\frac{60}{K_f}\right)\gamma H_L S_o$
		$\tau_c = (\sigma - 1)\gamma \phi D_{50}$
Rainfall soil detachment	(36)	$G_{\text{rf}} = E_{\text{rf}}(1 - GC)CFD$
Runoff soil detachment	(36)	$G_{\text{ro}} = E_{\text{ro}} D$
Potential sediment supply	(36)	$Y_S = (G_{\text{rf}} + G_{\text{ro}})\Delta t$
Soluble nutrient (nitrogen)	(21)	$C_{\text{RON}} = \frac{(N_{\text{AVS}} - N_{\text{AVR}})}{F_{\text{POR}}} * [e^{(-N_{\text{DMV}} I_{\text{EFF}})} - e^{(-N_{\text{DMV}} I_{\text{EFF}} - N_{\text{RMV}} R_{\text{OFF}})}] + \frac{N_{\text{RNC}} R_{\text{OFF}}}{P_{\text{EFF}}}$
Sediment attached nutrients	(21)	$N_{\text{SED}} = N_{\text{SCN}} Y_{\text{SED}} ER$ $P_{\text{SED}} = P_{\text{SCN}} Y_{\text{SED}} ER$ $ER = aY_{\text{SED}}^b T_f$

V = interception depth (mm), S_i = storage capacity (mm), C_p = ratio of vegetated surface area, E_a = evaporation rate (mm h^{-1}), t_R = duration of the rainfall (h), k = constant (mm^{-1}), P = precipitation (mm), D_s = depression storage (mm), S_d = surface retention value (mm), P_e = cumulative rainfall excess (mm), F = total depth of infiltrated water (mm), t = time (s), K = saturated conductivity (mm s^{-1}), m = average moisture content of the soil, m_o = initial soil moisture content, P_{ot} = capillary potential (mm), D_1 = detention storage (mm), P_{et} = potential evapotranspiration rate (mm d^{-1}), α = equilibrium factor, $s(T_a)$ = slope of the saturation vapor pressure vs. temperature curve, ξ = psychrometric constant, K_n = short-wave radiation, L_n = long-wave radiation, ρ = mass density of water, λ_y = latent heat of vaporization, R_a = total incoming solar radiation (mm), C_t = temperature reduction coefficient, δ_t = temperature difference (°F $\Leftarrow$°C), T_{avg} = mean temperature (°F $\Leftarrow$°C), Q_{int} = interflow (m^3 s^{-1}), R$_{\text{ec}}$ = coefficient representing depletion, W_{ac} = water accumulation in the upper zone storage (mm), R_{et} = retained storage (mm), Q_r = channel inflow (m^3 s^{-1}), $D_1 - D_s$ = runoff depth above ponding, R_3 = combined roughness and channel-length parameter optimized for each land class, S_1 = average overland slope, A = area of the element (m^2), Y_C = sediment transport capacity (kg m^{-2}), c = volumetric sediment concentration ratio $= A\left(\frac{\tau_d}{\tau_c}\right)^B$, $A = 0.00066$, $B = 1.61$, r_f = runoff (mm), τ_d = dominant flow shear stress, τ_c = critical stress, β = discharge parameter (= 5/3), γ = water specific weight (kg m^{-2} s^2), H_L = average runoff depth (mm), S_o = average overland slope, K_f = overland flow friction = 60 + 3140 GC$^{1.65}$, GC = ground cover factor, D_{50} = median size of soil particles (mm), σ = specific weight of sediment, ϕ Shields entrainment function $= \frac{0.11}{R^*} + 0.0211 \log_{10} R^*$, R^* = Reynolds number $= \frac{\sqrt{\tau_d \gamma D_{50}}}{\nu}$, ν = kinematic viscosity of water (m^{3}s^{-1}), G_{rf} = rate of soil detachment due to rainfall (kg m^{-2} h), CF = canopy factor, D = soil erodibility factor (gJ^{-1}), E$_{\text{rf}}$ = rate of rainfall energy (Jm^{-2} h) $= i(11.9 + 8.7 \log_{10} i)$, i = rainfall intensity (mm h^{-1}), G_{ro} = rate of soil detachment due to runoff (kg m^{-2} h), E_{ro} = rate of energy input to the soil by the flow (Jm^{-2}h) $= \left(\frac{60}{K_f}\right)\gamma \frac{Q_L}{2} S_o$, Q_L = unit flow discharge (m^2 h^{-1}), S_o = element slope, Δt = time increment (h), C_{RON} = soluble nitrogen concentration in runoff (kg ha^{-1}), N_{AVS} = available nitrogen content in the surface (kg ha^{-1}), N_{AVR} = available nitrogen in rainfall (kg ha^{-1}), N_{DMV} = rate of downward movement of nitrogen into soil, N_{RMV} = rate of nitrogen movement into runoff, I_{EFF} = effective infiltration (mm), R_{OFF} = total runoff (mm), F_{POR} = porosity factor, N_{RNC} = nitrogen contribution due to rain (kg ha^{-1}), P_{EFF} = effective precipitation (mm), N_{SED} = overland nitrogen transported by the sediment (kg ha^{-1}), N_{SCN} = soil nitrogen concentration (.001 gNg^{-1}soil), ER = nutrient enrichment ratio ($a = 7.4$, $b = -0.2$), T_f = correction factor for soil texture (0.85 for sand, 1.0 for silt, 1.15 for clay, and 1.50 for peat).

S.G. Stafford (Eds). Taylor and Francis Inc., Bristol, PA, pp. 333–356.

7. Chen, C.L., Gomez, E., Chen, C.W., Wu, C.M., Lin, J.J., and Chen, I.L. (1995). An integrated watershed model with GIS and windows application. *Proceedings of International Symposium on Water Quality Modelling*. American Society of Agricultural Engineers, St. Joseph, MI, pp. 242–250.

APPENDIX 2

Specific Equations for Routing Component of NPS Assessment

Variables	Author(s)	Governing Equation[a]
Grid or cell for sediment	(25)	$Y_{\text{SEDout}} = 1000[(1 - S_{\text{Dep}})(Y_{\text{SEDab}} + Y_{\text{SEDin}})]$
Grid or cell for nutrients	(25)	$CC_{\text{RONout}} = \frac{100}{R_{\text{OFF}}}[(1 - N_{\text{Dec}})(C_{\text{RONab}} + C_{\text{RONin}})]$

[a] The subscripts 1 and 2 indicate the beginning and end of the time step; $I_{1,2}$ = inflow to the reach (m^3s^{-1}) consisting of overland flow, interflow base flow and channel flow from all contributing upstream basin elements; $O_{1,2}$ = outflow from the reach (m^3s^{-1}); $S_{1,2}$ = storage in the reach (m^3); Δt = time step of the routing (s); R_2 = channel roughness parameter; A_x = channel cross-section area (m^2); S_o = channel slope; Y_{SEDout} = sediment leaving the cell (ppm); S_{Dep} = deposition fraction; Y_{SEDab} = sum of all the sediment entering the cell (kg ha^{-1}); Y_{SEDin} = sediment generated within the element; $\text{CC}_{\text{RONout}}$ = soluble nutrients concentration in runoff leaving the cell (ppm); N_{Dec} = nutrients decay fraction; CC_{RONab} = sum of all the nutrient entering the cell (kg ha^{-1}); CC_{RONin} = nutrient generated within the element.

8. Livingston, D. and Raper, J. (1993). Modelling environmental systems with GIS: theoretical barriers to progress. In: *Innovations in GIS I*. M.F. Worboys (Ed.). Taylor and Francis, Bristol, PA, pp. 229–240.
9. Nyerges, T.L. (1993). Understanding the scope of GIS: its relationship to environmental modelling. In: *Environmental Modelling with GIS*. M.F. Goodchild, B.O. Parks, and L.T. Steyaert (Eds.). Oxford University Press, New York, pp. 75–93.
10. Fedra, K. (1993). GIS and environmental modelling. In: *Environmental Modelling with GIS*. M.F. Goodchild, B.O. Parks, and L.T. Steyaert (Eds.). Oxford University Press, New York, pp. 35–50.
11. Olivieri, L.J., Schaal, G.M., Logan, T.J., Elliot, W.J., and Motch, B. (1991). *Generating AGNPS input using remote sensing and GIS. ASAE paper no. 91–2662.* American Society of Agricultural Engineers, St. Joseph, MI.
12. Donovan, J.K. (1991). The U.S. EPA's Arizona rangeland project: integration of GIS and a non-point source pollution model. *Proceedings of the 11 Annual ESRI User Conference*, Environ. Systems Research Institute, Redlands, CA, pp. 261–272.
13. Metcalf and Eddy, Inc. Univ. of Florida and Water Resources Engineers, Inc. (1971). *Storm Water Management Model* (4 vols.). U.S.Environmental Protection Agency. Washington, DC.
14. Johanson, R.C., Imhoff, J.C., and Davis, H.H. (1980). *User's manual for hydrologic simulation program—Fortran (HSPF). EPA—600/9-80-015,* U.S. Environmental Protection Agency, Athens, GA, p. 684.
15. Beasley, D.B., Huggins, L.F., and Monke, E.J. (1980). ANSWERS: a model for watershed planning. *Trans. ASAE* **23**(4): 938–944.
16. Knisel, W.G., Jr. (Ed.). (1980). CREAMS: a field scale model for chemicals, runoff and erosion from agricultural management systems. *Cons. Res. Repsot. No. 26.* U.S. Dept. Agr., Washington, DC, p. 640.
17. DeCoursey, D.G. (1982). *ARS small watershed model*. Paper no. 82–2094–Am. Soc. Agr. Eng. St. Joseph, MI, p. 33.
18. Shafter, N.J., Gupta, S.C., Linden, D.R., Molina, J.A.E., Clapp, C.E., and Larson, W.E. (1983). Nitrogen-tillage-residue management model (NTRM). In: *Analysis of Ecological Systems: State of the Art in Ecological Models*. Elsevier, New York, pp. 525–544.
19. Williams, J.R. and Nicks, A.D. (1985). SWRRB- A simulator for water resources in rural basins: An overview. *Proc. Natural Resources Modelling Symp.*, Pingree Park, CO, October 16–21, 1983. pp. 17–24. USDA-ARS. ARS-30.
20. Abbott, M.B., Bathurst, J.C., Cunge, J.A., Connell, P.E., and Rasmussen, J. (1986). An introduction to the European hydrological system—systeme hydrologique european, 'SHE' 2: structure of physically based, distributed modelling system. *J. Hydrol.* **87**: 61–77.
21. Young, R.A., Onstad, C.A., Bosch, D.D., and Anderson, W.P. (1986). *AGNPS Agricultural Nonpoint Source Pollution Model*. A Large Watershed Analysis Tool. USDA. ARS Conserv. Res. Report No. 35, Springfield, VA.
22. Arnold, J.G., Williams, J.R., Srinivasan, R., King, K.W., and Griggs, R.H. (1995). *SWAT—Soil and Water Assessment Tool: Draft User's Manual*. U.S. Department of Agriculture, Agricultural Research Service, Grassland, Soil and Water Research Laboratory, Temple, TX.
23. Bouraoui, F. and Dillaha, T.A. (1996). ANSWERS-2000: Runoff and sediment transport model. *J. Environ. Eng. ASCE* **122**(6): 493–502.
24. Simons, D.B. and Senturk, F. (1976). *Sediment Transport Technology*. Water Resources Publications, Fort Collins, CO .
25. Frere, M.H., Ross, J.D., and Lane, L.J. (1980). The nutrient submodel. In *CREAMS: A Field Scale Model for Chemicals, Runoff and Erosion from Agricultural Management Systems*. Ed. W.G. Knisel, pp. 65–86, USDA, Cons. Research Report No. 26.
26. Ghosh, N.C. and McBean, E.A. (1998). Water quality modelling of the Kali River, India. *Water, Air, and Soil Pollution* **102**: 91–103.
27. Jain, C.K., Bhatia, K.K.S., and Seth, S.M. (1998). Assessment of point and non-point sources of pollution using a chemical mass balance approach. *Hydrol. Sci. J.* **43**(3): 379–390.
28. Jha, R., Ojha, C.S.P., and Bhatia, K.K.S. (2000). Development of deoxygenation and reaeration rate coefficients for a small tributary of river Hindon, U.P., India. *International Conference on Integrated Water Resources Management (ICIWRM-2000)*, New Delhi, India, pp. 464–474.
29. Jha, R., Ojha, C.S.P., and Bhatia, K.K.S. (2001). Refinement of predictive reaeration equations for a typical indian river. *Hydrol. Process.* **15**(6): 1047–1060.
30. APHA. (1985). Standard methods for the examination of water and wastewater (16th). American Public Health Association, Washington DC, p. 1268.
31. Kouwen, N. (1988). WATFLOOD: a micro-computer based flood forecasting system based on real-time weather radar. *Ca. Water Res. J.* **13**(1): 62–67.
32. Philip, J.R. (1954). An infiltration equation with physical significance. *Soil Sci.* **77**: Jan–Jun.

33. Priestley, C.H. and Taylor, R.J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Weather Rev.* **100**(2): 81–92.
34. Linsley, R.K., Kohler, M.A., and Paulhus, L. (1949). *Applied Hydrology*. McGraw Hill, New York.
35. Hargreaves, G.H. and Samani, Z.A. (1982). Estimating potential evapo-transpiration. *Tech. Note, J. Irrig. and Drain. Engrg., ASCE* **108**(3): 225–230.
36. Hartley, D.M. (1987). Simplified process model for water and sediment yield from single storms. *Trans. ASAE* **30**(3): 710–723.

DEEP-WELL TURBINE PUMPS

Blaine Hanson
LAWR
Davis, California

Deep-well turbine pumps are used to pump groundwater to the surface, and in some cases, they may be used to pump water out of sumps and supply pressure. Deep-well turbines consist of a housing or bowl, impellers, and a shaft, all of which are installed in the well [Figs. 1(a) and 1(b)]. The housing or bowl contains both the impeller and the diffuser vanes. The shaft connects the impeller to the electric motor or engine, whereas the impeller transfers the energy developed by the motor or engine to the water. Diffuser vanes guide the flow of water from the impeller to the discharge point of the bowl.

Impellers transfer energy from the motor or engine to the water either by centrifugal force or by lifting action. They may be enclosed or semi-open. Enclosed impellers have a disk or shroud on both sides of the vanes (Fig. 2). These impellers require a close clearance between a wearing ring, located at the inlet of the impeller, and the impeller skirt of the bottom extension of the impeller. In semi-open impellers, the shroud is only at the top of the impeller. Semi-open impellers have a shroud only on the top (Fig. 2). A close clearance (0.076 mm–0.177 mm) between the bottom of the vanes and the pump housing is necessary for good performance. Semi-open impellers are less susceptible to wear from sand because they can be raised slightly to reduce abrasion. This process, however, will decrease pump output and efficiency.

Mixed flow impellers are often used in deep-well turbine pumps. Mixed flow impellers use both centrifugal and lifting forces to pump water. When the impeller rotates, centrifugal forces develop in the water inside the impeller and cause water to flow toward the impeller's outer edge. This flow, in turn, results in water flowing into the center or eye of the impeller. Rotating a mixed flow impeller also transfers energy to the water by a lifting force. Impellers that pump water only through centrifugal force are classified as radial flow impellers. Impellers that pump water only through lifting forces are classified as axial or propeller impellers.

Some deep-well turbines use a line-shaft to connect the impellers to an electric motor or engine. Line shafts are supported by bearings—either oil-lubricated or water-lubricated—spaced at intervals along the shaft length. Oil-lubricated bearings are housed within a tubing inside the column pipe (the pipe that carries the water to the surface). An oil reservoir located on the pump head drips oil down the shaft.

Submersible pumps are deep-well turbines with a waterproof electric motor installed in the well below the pumping level. The motor is attached directly beneath the bowl assembly, and water intake is between the motor and the pump. A specially designed motor, slimmer and more compact than a line-shaft motor, is installed in the well. As both motor and pump are situated inside the well, submersible pumps do not require a line shaft.

In submersible pumps, it is crucial that enough water circulate around the motor to keep it cool. As water in the well should flow past the motor casing into the pump intake, submersible motors should be located above the level at which most of the flow enters the well. A shroud can be installed over the pump intake to force water to flow past the motor.

As the impellers and bowls must be submerged inside the well, the diameters of the impeller and bowl are limited by the diameter of the well, which in turn limits the output (capacity and pressure) developed by each impeller. This limitation can be overcome if the bowls are installed in stages, with each stage adding to the output developed by the preceding stages. A five-stage pump will, therefore, have five bowl assemblies, each containing an impeller, with the water guided from stage to stage by diffuser vanes built into the bowl assembly. Figure 1 shows a two-stage pump.

Because deep-well turbine pumps are installed in the well, pump size is limited by the diameter of the well casing. If the pump is too large for the casing, rocks from the gravel pack can become lodged between the casing, which makes it difficult to remove the pump and—particularly if the well is crooked—the pump may be difficult to install. The clearance between the bowls and the well casing should be at least 25 mm. Narrower clearances may complicate installation or, in crooked wells or where gravel becomes lodged between the pump and casing, may make it difficult for the bowls to be removed for repair.

The output of the pump must be sufficient to lift the water to the ground surface and provide the desired pressure at the desired flow rate or pump capacity. However, in some cases, deep-well turbines lift the water only and a centrifugal pump located at the discharge point of the well, commonly called a booster pump, is placed in line with the deep-well turbine pump to supply pressure.

The output of a pump is described by its performance characteristic curves, developed by pump manufacturers. These curves show the amount of total dynamic head developed by a single stage at a particular capacity, the pump efficiency at that capacity, and the horsepower requirements of the pump. Performance curves show which pump will provide the desired total head and capacity and determine what size electric motor or engine the pump will require.

Figure 3(a) shows the total head and capacity performance curve for a typical pump. In a no-flow condition (shut-off), the total head is maximum, or 22 m for this

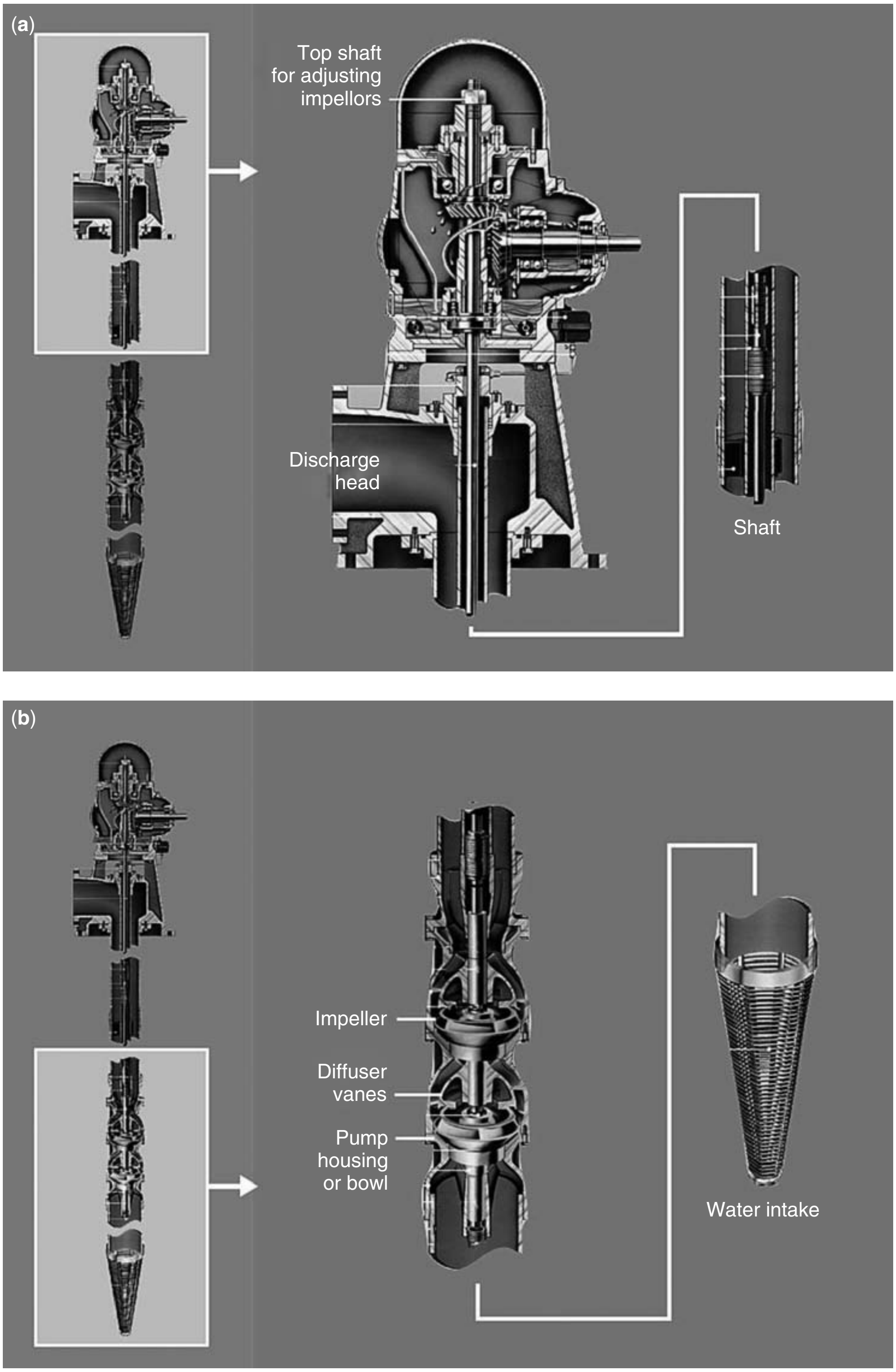

Figure 1. Cutaway of a deep-well turbine pump that shows (**a**) the pump head, gear head, and shaft; and (**b**) pump housing, impellers, and shaft. [*Source*: Hanson (1)].

pump. As the pump capacity increases, the total head decreases. At low capacities, most of the pump output consists of total head (normally measured as the sum of the pumping lift and the discharge pressure head), whereas at high capacities, most of the pump output is in capacity.

Pump efficiency increases as pump capacity increases until a maximum efficiency is reached [Fig. 3(b)]. Thereafter, pump efficiency decreases as capacity continues to increase.

Metric horsepower (mhp) increases slightly as capacity increases to a maximum, and then it may decrease slightly [Fig. 3(c)].

The performance curves are only for one-stage, but they can develop performance curves for multiple-stage pumps. For example, Fig. 3(a) shows that in a one-stage pump, 14.6 m of head is developed at 3028 L/min. The mhp is 14.2. In a three-stage pump, the head developed at 3,028 L/min will be three times that of a single stage, or 43.8 m. The mhp will also be tripled, to 42.6 mhp, and the pump efficiency may change slightly. Usually the manufacturer's curve will show the efficiency adjustments. The performance curve for a multiple-stage pump can therefore be estimated by multiplying the head and bhp of the performance curve by the number of stages for each capacity.

Pump performance curves have a variety of shapes. Some head/capacity curves are steep (large decreases in head occur as capacity increases), whereas others are flat (small decreases in head occur as capacity increases). The best curve will depend on changes in pumping conditions. Where pumping levels fluctuate significantly, a steep curve may allow for significant changes in pumping lift with only slight changes in pump capacity. A flat curve may help to maintain a relatively constant pressure over a wide range of flow rates.

An initially efficient pump with the necessary head and capacity can become inefficient. Inefficiency is frequently caused by wear because of sand in the well water, but it can also be caused by a mismatch between the pump and water system characteristics because of long-term changes in groundwater levels or changes to the water supply system. A mismatched pump is one that is operating properly but at an efficiency much less than the potential maximum efficiency. "Repairing" a mismatched pump will not improve efficiency.

The pump's performance can be evaluated by first testing the pump and then comparing that information with the manufacturer's performance curves. The pump test consists of measuring the pumping lift, discharge pressure, capacity, and input power to the pump and then calculating overall pumping plant efficiency—that is, the combined efficiency of the pump and motor or engine. The overall efficiency will be less than the pump efficiency reflected in the manufacturer's performance curve because of the inefficiency of the motor or engine.

Pump tests are conducted by utility companies, pump dealers, and consultants. In defining the current status of the pump, the test may provide any or all of the following information:

- Pumping lift or level—elevation difference between the discharge pipe at the pump head and the water level in the pumping well

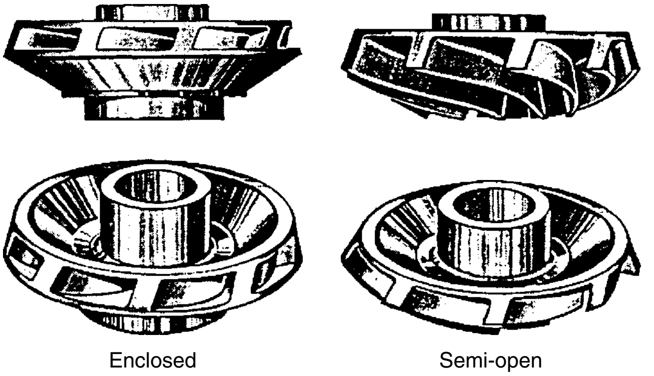

Enclosed Semi-open

Figure 2. Enclosed and semi-open impellers. [*Source*: Hanson (2)].

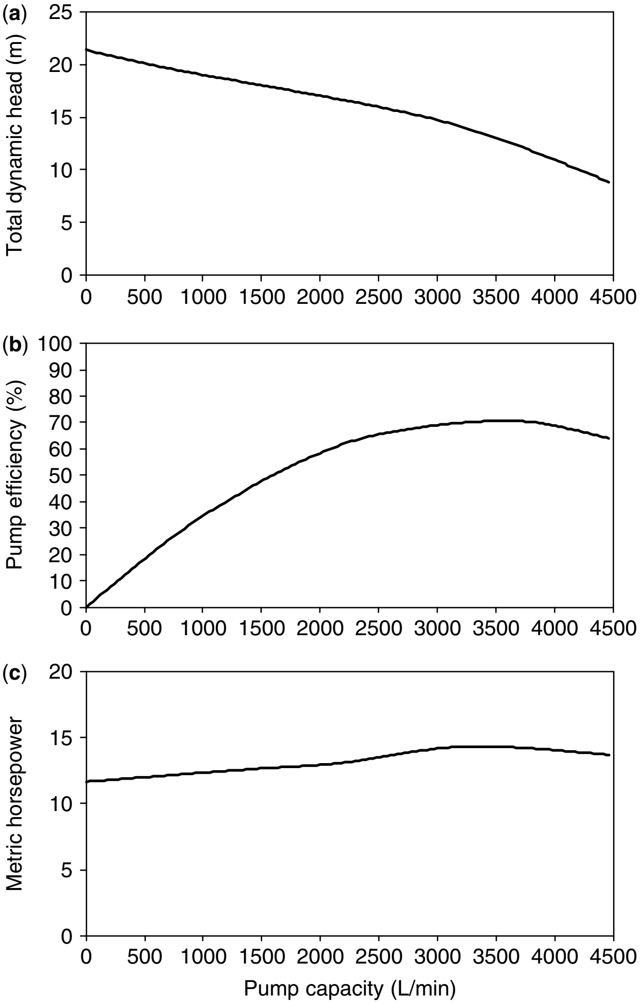

Figure 3. Pump performance curves that show the relationships between (**a**) total head and capacity, (**b**) pump efficiency and capacity, and (**c**) metric horsepower and capacity.

- Static or standing water level—elevation difference between discharge pipe and water level in the well before pumping
- Drawdown—difference between the pumping level and the static water level
- Discharge pressure—pressure at the discharge point of the pumping plant measured with a pressure gauge
- Discharge pressure head—the height of the column of water needed to provide the pressure at its base—calculated by multiplying the pressure by a conversion constant (0.704, 0.1021, or 10.21 for discharge pressure in psi, kPa, or bars, respectively)
- Total head—sum of the pumping lift and the discharge pressure head (note that if the pump is situated deep in the well, pressure losses caused by friction in the column pipe may have to be added to the total head)
- Pump capacity—water flow rate of the pumping plant.

- Well yield—flow rate or capacity divided by the drawdown (a well yield decreasing over time may indicate plugging in the well)
- Input horsepower to electric motor
- Percent motor overload
- Input kilowatts to motor
- Kilowatt-hours per acre-foot of pumped water
- Overall pumping plant efficiency

The pumping lift is measured by lowering an electric water level indicator (two electrodes attached to a wire connected to an ampere meter) or some other measuring device into the space between the well casing and the column pipe. In some cases, an air line is permanently installed to measure the pumping lift. This measurement requires access to the inside of the well casing; without such access, the pump test cannot be performed.

The pump capacity is measured with a flow meter installed in the discharge pipe. To operate properly, a flow meter requires a straight section of pipe eight to ten pipe diameters long immediately upstream from the meter to prevent or reduce turbulence in the water. Too much turbulence from valves, bends, or elbows will prevent accurate flow measurements. It is also recommended that a straight section of pipe two pipe diameters long be installed immediately downstream from the flow meter.

The input power of an electric motor or engine must be determined to calculate the overall pumping plant efficiency. Input horsepower of electric motors can be easily calculated. In some cases, the power meter will provide this information. Determining the input horsepower of engines is much more complicated and will require measuring the fuel flow rate to the engine.

The overall pumping plant efficiency can be calculated by the following equation:

$$E_o = (Q \times H)/(4634 \times \text{mhp})$$

where E_o is the overall pumping plant efficiency, Q is the capacity (L/min), H is the total head (m), and mhp is the metric input horsepower.

The pump test data should be compared with the manufacturer's performance curves to determine how the pump's actual performance compares with the manufacturer's recommended performance. This comparison step is important because it is possible to have a pumping plant measuring low efficiency but nonetheless operating properly. Repairing such a pump would be of no benefit. If the operating point is close to the head/capacity performance curve, the pump is operating properly, regardless of the overall efficiency calculated from the test data. If a significant difference exists between the performance curve and the operating point, the pump may be worn and should be repaired or replaced.

The pump test data shown in Fig. 4 illustrate the evaluation process. Two pump tests were conducted—the first when the pump was worn, with an overall efficiency of 39% and the second after the pump was repaired, with an overall efficiency of 50%. The worn pump showed a significant difference between the manufacturer's curve and the operating point, whereas the repaired pump showed only a slight difference. However, the efficiency of the repaired pump suggests that the pump is mismatched to its operating conditions and probably should be replaced by a matched pump.

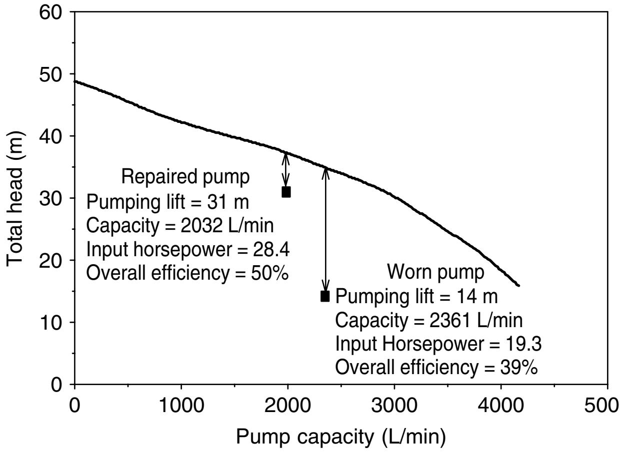

Figure 4. Comparison of pump test data with manufacturer's total head-capacity curve.

Adjusting the impellers can partially restore the capacity and total head of deep-well turbine pumps using semi-open impellers that are slightly or moderately worn. The adjustment consists of slightly lowering the pump shaft to which the impellers are connected by turning the nut at the top of the motor or gear head. The adjustment must also account for line shaft elongation that occurs when the pump is operating.

BIBLIOGRAPHY

1. Hanson, B.R. (2002). Improving pumping plant efficiency does not always save energy. *California Agriculture* **56**(4): 123–127.
2. Hanson, B.R. (2000). *Irrigation Pumping Plants*. University of California Division of Agriculture and Natural Resources Publication 3377.

MICROBIAL QUALITY OF RECLAIMED IRRIGATION: INTERNATIONAL PERSPECTIVE

Absar Alum
Arizona State University
Tempe, Arizona

Wastewater reuse has been practiced in different forms throughout the history of human civilization. Previously, sewage had been dumped in rivers and lakes or used locally to irrigate crops. Explosive population growth and diminishing fresh water resources around the globe have now made wastewater a valuable resource, especially in arid and semiarid areas of the world. The introduction of modern wastewater treatment processes has given new dimensions to water reuse. Reclaimed waters are commonly used for landscape or recreational irrigation (parks, golf courses, athletic fields, school yards, highway medians, and public and private lands) and agricultural crop irrigation (field crops, vegetables, orchids, and tree plantations).

Human feces, which may have microbial constituents as high as 10–30%, are one of the major sources of infectious and noninfectious agents found in municipal wastewater. Thus, wastewater can contain high concentrations of excreta-related microorganisms. The principal microorganisms (infectious agents) that may be present in raw sewage can be classified into three broad groups: bacteria, parasites (protozoa and helminths), and viruses. During an outbreak, pathogen numbers in local sewage go up. More than 30 excreta-related diseases of public health importance have been documented, and many of these are of specific importance in wastewater reuse schemes. Table 1 summarizes the major infectious agents potentially present in raw domestic wastewater and their respective diseases.

TYPES OF MICROORGANISMS

Bacteria

Bacteria are microscopic organisms, ranging from approximately 1 to 6 μm in length. Human and animal intestinal tracts are colonized by different types of bacteria, which are routinely shed in the feces. The number of bacteria in feces can be in the range of 10^8–10^{11} per gram of fecal material, and these bacteria end up in municipal wastewater (1). Coliforms, such as *Escherichia coli,* and *Enterobacter* constitute a small group of intestinal bacteria, which has historically been used as an indication that an environment is contaminated by human excreta. In addition to coliforms, some pathogenic bacteria, such as *Salmonella, Shigella, Vibrio, Mycobacterium, Clostridium, Leptospira*, and *Yersinia*, can also colonize part of a gastrointestinal tract, resulting in gastroenteritis.

A number of opportunistic bacterial pathogens can also be found in waste and reclaimed water. These opportunistic pathogens include *Pseudomonas, Streptococcus, Flavobacterium*, and *Aeromonas* species (2).

Parasites

Parasites are a very diverse group of microorganisms; they can be further categorized as protozoan, metazoan, or helminth. Parasites often have a complex life cycle sometime requiring a stage in intermediate hosts. The environmental stage of parasites found in wastewater is normally called cyst/oocyst (protozoans) or egg (helminth). In general, parasite cysts or eggs are larger than bacteria. Parasites are present in the feces of infected persons; however, healthy carriers may also excrete them. Cysts do not reproduce in the environment but can survive in soil for months or even years, depending on environmental conditions.

Entamoeba histolytica, the most dangerous protozoan parasite, occurs worldwide, although the incidence in the

Table 1. Pathogenic Microorganisms Associated with Wastewater

Group of microorganisms	Examples	Disease
Bacteria	*Campylobacter jejuni*	Gastroenteritis (campylobacteriosis)
	Clostridium	Gastroenteritis
	Enteropathogenic *Escherichia coli*	Gastroenteritis
	Leptospira	
	Mycobacterium	
	Salmonella typhi	Typhoid fever
	Salmonella sp.	Gastroenteritis (salmonellosis)
	Shigella	Bacillary dysentery
	Vibrio cholera	Cholera
	Yersinia sp.	Gastroenteritis (yersiniosis)
Viruses	Adenoviruses	Gastroenteritis (conjunctivitis, Respiratory)
	Caliciviruses	Gastroenteritis
	Coxackie viruses	Flu-like symptoms, gastroenteritis
	Echoviruses	Flu-like symptoms, aseptic meningitis
	Hepatitis A virus	Infectious hepatitis
	Hepatitis E virus	Infectious hepatitis
	Polioviruses	Poliomyelitis
	Reoviruses	Respiratory diseases
	Rotavirus	Acute gastroenteritis
Protozoan	*Balantidium*	Balantidiasis (gastroenteritis)
	Cryptosporidium	Cryptosporidiosis (diarrhea)
	Entamoeba coli	Diarrhea, ulceration
	Entamoeba histolytica	Amebiasis (amoebic dysentery)
	Giardia lamblia	Giardiasis (diarrhea)
Helminth	*Ascaris* sp.	Ascariasis (roundworm infection)
	Necator americanus	Ancylostomiasis (hookworm infection)
	Taenia sp.	Taeniasis (tapeworm infection)
	Trichuis trichuria	Trichuriasis (whipworm infection)

United States is not very well documented. *Cryptosporidium parvum* oocysts are the most important biological water contaminants in the United States (3).

Several helminthic parasites may be found in wastewater. *Ascaris lumbricoides* (roundworm) is one of the most important human parasites. It is estimated that approximately 25% of the world's population is infected by this nematode (4).

Viruses

Viruses are smaller than bacterial disease agents and range in size from approximately 20 to 200 nm. They are obligate intracellular parasites that can multiply only within a host cell. Viruses that can replicate in the gastrointestinal tract of humans or animals are referred to as enteric viruses. The presence of enteric viruses in sewage is a matter of concern because they can be transmitted through the fecal–oral route. Morethan 120 different human enteric viruses have been reported to date. An infected person can excrete as many as 10^6-10^{10} plaque-forming units (pfu) of enteric viruses per gram of fecal material.

The most important human enteric viruses are enteroviruses (polio, echo, and Coxsackie), caliciviruses, rotaviruses, reoviruses, parvoviruses, and adenoviruses. Caliciviruses are the most frequently reported and documented that are transmitted by contaminated water. Even though the etiologic agent cannot be successfully cultivated in the laboratory, there is epidemiological evidence that caliciviruses are the most common cause of viral gastroenteritis. Hepatitis A, the virus that causes infectious hepatitis, is one of the most resistant to environmental stresses (5). Rotaviruses are the most infectious of all enteric viruses (6) and thus can be considered a high risk to health if present in wastewater.

MECHANISM OF DISEASE TRANSMISSION

There are several ways in which an individual can acquire disease from wastewater use, and those can be broadly categorized into two modes; direct ingestion and indirect ingestion. Humans can acquire an etiologic agent either directly by contact, ingestion, or inhalation of reclaimed irrigation water or indirectly by contact with the objects previously contaminated by the reclaimed water. Following are some of the conditions for a successful infection or disease from exposure to reclaimed wastewater.

- *Etiologic Agent*: The disease-causing agent (bacteria, virus or a parasite) must be present in wastewater and hence, in the wastewater from that community.
- *Survival of Etiologic Agent*: The disease-causing agent must survive all wastewater treatment processes to which it is exposed.
- *Exposure*: The disease competent person must come in contact, directly or indirectly, with reclaimed water carrying the disease agent.
- *Infectivity and Infective Dose*: The etiologic agent must be infectious, moreover, the disease-causing agent must be sufficiently numerous at the time of contact to cause illness.

The exposure to an etiologic agent does not always result in disease. Disease results from a series of complex interactions between the host and the infectious agent, which can be defined in terms of the pathogenicity and virulence of the disease-causing agent and the relative susceptibility of the host. The susceptibility of the host depends on the general health of the subject and the specific pathogen. Age and the nutritional and immunological status of subjects are some of the other factors that determine the outcome of an infection.

MONITORING MICROORGANISMS

The detection, isolation, and identification of all microorganisms in a wastewater sample are very difficult, time-consuming, and cumbersome. The detection efficacy and sensitivity are influenced by various physicochemical parameters of the water sample, which further undermine the significance of such a strategy. To avoid the necessity of undertaking such a huge task, which has no equivalent value, indicator microorganisms are used to estimate the presence of fecal pathogens in water samples. An indicator should be present in equal or higher numbers and be more resistant than pathogens. Historically, fecal coliforms, in particular, *E. coli*, have been used as indicators of fecal contamination (7).

REMOVAL OF MICROORGANISMS BY WASTEWATER TREATMENT PROCESSES

The purpose of most advanced treatment processes is to remove either inorganic or organic constituents. Therefore, the removal of biological contaminants by these processes is only incidental in many cases and, generally is not too great. Microbial removal efficiencies of some treatment processes are given Table 2. Conventional biological treatment processes (trickling filters, activated sludge, and oxidation ponds) reduce the quantities of biological organisms in raw or settled sewage but do not eliminate them.

The most important treatment process for pathogen destruction is disinfection. In the United States, the most common disinfectant for both water and wastewater is chlorine. The efficiency of disinfection with chlorine depends on the water temperature, pH, length of contact, degree of mixing, presence of interfering substances, concentration and form of the chlorinated species, and nature and concentration of the organisms to be destroyed.

TYPES OF RECLAIMED IRRIGATION PRACTICES

Reclaimed irrigation can be practiced in the following five different ways:

- Flood irrigation, which wets almost all the land surface;

Table 2. Concentration of Various Microorganisms in Raw and Treated Wastewater[a]

Microorganisms	Number per 100 Raw Sewage	Reduction in Number of Microbes ($\log_{10}$) Primary Effluent	Secondary Effluent	Tertiary Effluent
Fecal coliform (MPN)	10^9	2	1	6
Salmonella (MPN)	8×10^3	1	2	1
Enteric viruses (PFU)	5×10^4	0.5	1	5
Helminth ova	8×10^2	1	2	1
Giardia cyst	10^4	0.5	0.5	3

[a] Reference 8.

- Furrow irrigation, which wets only part of the ground surface;
- Sprinkler irrigation, in which the soil is wetted in much the same way as by rain;
- Surface drip irrigation, in which water is applied at each individual plant at an adjustable rate;
- Subsurface drip irrigation, in which the topsoil remains relatively dry but the subsoil is saturated.

CONTROL OF HUMAN EXPOSURE

Four groups of people are at potential risk from reclaimed irrigation:

- agricultural field workers and their families;
- crop handlers;
- consumers (of crops, meat, and milk);
- population living near the affected fields; and
- recreational area visitors, players, etc.

GUIDELINES AND TREATMENT PRACTICES FOR WASTEWATER REUSE IN AGRICULTURE

All reuse guidelines or criteria have a single basic objective, to ensure the safety of public health. Wastewater is known to harbor a variety of different microorganisms, influenced by various environmental factors, so it is possible to address different aspects of this issue. Reuse guidelines can be based on the level of indicator microorganisms, epidemiological and/or risk models, and treatment practices.

Early guidelines for wastewater reuse in agriculture were concerned only with the pathogen content of wastewater. A World Health Organization (WHO) meeting of experts on the reuse of effluents in 1971 reviewed the information concerning health risks and concluded that only a limited health risk would result from unrestricted crop irrigation with wastewater containing no more than 100 coliforms per 100 mL. Following the recommendation of the WHO meeting in 1971, new epidemiological evidence was accumulated, and earlier studies were reevaluated. A meeting in Engelberg in 1985 of public health experts, environmental scientists, and epidemiologists reviewed the work that had been done and concluded that the risk of infection was lower than previously thought, particularly, from bacterial and viral pathogens (9).

The WHO guidelines for wastewater reuse, which were proposed by the Engelberg meeting in 1985 and reaffirmed in 1989, were aimed at providing a high degree of health safety to workers and consumers, but at the same time to be realistic in the light of resource constraints and limited technology in many countries.

In the United States the State of California has a long history of reclaimed water regulation. The relatively stringent California standards are not economically or technologically feasible for developing countries.

INTERNATIONAL STANDARDS

The WHO water reuse guidelines (Table 3) have served as a baseline for the development of reclaimed irrigation standards throughout the world generally and in developing countries particularly. On the other hand, California criteria, which represent the higher end, are extensively followed by developed countries. Different countries have adapted both standards to develop their regional microbial standards for reclaimed irrigation (Table 4).

Italy became the first European country to regulate reclaimed irrigation when it adopted wastewater reuse guidelines in 1977 under its 1976 national Water Law. Although the Italian guidelines were published 1 year earlier than California guidelines, both essentially follow the same approach. The region of Sicily has adopted different regulations, which are much closer to the WHO approach. As an added precaution, this regulation forbids irrigating vegetables that are eaten raw with any type of reclaimed wastewater.

Cyprus produces approximately 25 Mm^3 of urban effluent every year. Government has planned to expand irrigated agriculture by introducing reclaimed irrigation. Provisional standards for reclaimed water irrigation were introduced to regulate reclaimed irrigation, which take into account the special ecological and environmental conditions of Cyprus. These standards, which are stricter than the WHO guidelines, ensure the best possible application of the effluent for irrigation.

In France, reclaimed irrigation of crops is practiced mostly around big cities as the easiest way to treat and dispose of wastewater. In the early 1990s, the growing demand for irrigation water along with periodic droughts resulted in increased interest in water reuse. French health authorities [the Conseil Supérieur díHygiène Publique de France (CSHPF)] issued the “health guidelines for reuse, after treatment, of wastewater for crop

Table 3. Microbiological Quality Guidelines for Wastewater Use in Agriculture[a] by WHO

Category	Reuse Conditions	Exposed Group	Intestinal Nematodes (Mean Arithmetic Number of Viable Eggs per liter)	Faecal Coliforms (Geometric Mean, Number per 100 mL)	Wastewater Treatment Expected to Achieve the Required Microbiological Quality
A	Irrigation of crops likely to be eaten uncooked, sports fields, public parks	Workers, consumers, public	1	1000	A series of stabilization ponds designed to achieve the microbiological quality indicated, or equivalent treatment
B	Irrigation of cereal crops, industrial crops, fodder crops, pasture and trees	Workers	1	No standard recommended	Retention in stabilization ponds for 8–10 days or equivalent helminth and fecal coliform removal
C	Localized irrigation of crops in category B if exposure of workers and the public does not occur	None	Not applicable	Not applicable	Pretreatment as required by irrigation technology, but not less than primary sedimentation

[a] Reference 10.

Table 4. Guidelines and Criteria for Wastewater Reuse in Irrigation in Various Countries[a]

Country	Crop Type	Treatment Required	Microbial Criteria (max.)
Australia	Residential use, municipal irrigation, unrestricted crop irrigation	Secondary, tertiary, filtration and disinfection	<10 thermotolerant coliforms/100 mL
	Ornamental ponds with public access, restricted crop irrigation, irrigation of pasture and fodder crops for dairy animals, fire fighting	Secondary and disinfection	<100 thermotolerant coliforms/100 mL
	Municipal irrigation with restricted crop irrigation, irrigation of pasture and fodder crops for grazing animals	Primary sedimentation and lagooning or secondary plus disinfection (if required)	<1000 thermotolerant coliforms/100 mL
	Irrigation for turf production, silviculture, nonfood chain aquaculture	Primary sedimentation and lagooning or secondary	$<10,000$ thermotolerant coliforms/100 mL
Israel	Cotton, sugar beet, dry fodder seeds, forest irrigation	None	None
	Green fodder, nuts, citrus, olives, peanuts, bananas, almonds, etc.	None	None
	Deciduous fruits, conserved vegetables, cooked and peeled vegetables	Disinfection with chlorination (60 minutes contact time)	250 coliforms/100 mL
	Unrestricted crops, including vegetables eaten raw, parks, and lawns	Sand filtration and disinfection with chlorination (60 minutes contact time)	12 coliforms/100 mL (80%)
			2.2 coliforms/100 mL (50%)
Japan	Landscape irrigation	Disinfection (>0.4 mg/L residual chlorine)	0.0 *E. coli*/100 mL
Kuwait	Fodder, food crops not eaten raw, forest land	Advanced water treatment (1 mg/L residual chlorine after 12 h@20 °C)	10,000 coliforms/100 mL
	Food crops eaten raw	Advanced water treatment (1 mg/L residual chlorine after 12 h@20 °C)	100 coliforms/100 mL
Saudi Arabia	All irrigation purposes (unrestricted)	Advanced water treatment	2.2 coliforms/100 mL
	All irrigation purposes	Treatment plants and waste settling ponds	<1 intestinal nematode egg/100 mL
South Africa	Irrigation of dry fodder crops, seed crops, trees, non-recreational parks, nurseries (restricted access)	Primary and secondary	<1000 fecal coliforms/100 mL
	Food crops not eaten raw, cut flowers, orchards and vineyards, pastures, parks	Primary, secondary, and tertiary; oxidation pond system	<1000 fecal coliforms/100 mL
	Pasture for milking animals	Primary, secondary, and tertiary	0.0 fecal coliforms/100 mL
	Food crops eaten raw, lawns, nurseries (unrestricted access)	Advanced (general) drinking water standards	Drinking water standards

[a] Reference 11.

and green spaces irrigation"(12). The guidelines stipulate maintaining of extensive water quality data. Wastewater reuse authorizations are granted on a case-by-case basis after review of a highly detailed dossier of French reclaimed irrigation regulations based on WHO guidelines but complement them with strict rules of application such as a minimum distance required between irrigation sites and residential areas or roadways.

In Mexico, wastewater is extensively used to irrigate crops. Because a very significant part of Mexican farm produce is sold in U.S. markets, Mexico has developed very comprehensive reclaimed irrigation regulations that are stricter than WHO suggested standards (Tables 5 and 6).

Israel has developed very comprehensive reclaimed water usage regulations because reclaimed water constitutes a significant part of its national water budget. About 92% of the total wastewater is collected, and 72% is reclaimed for various purposes. Almost 42% of total wastewater is reused for irrigation (13). The Israel Ministry of Health has established criteria for reclaimed irrigation and water reuse that are stricter than WHO standards. According to Israeli standards, settled effluent can be used on industrial and fodder crops, pastures and hay, vegetables eaten cooked, fruit trees, ornamental plants, or seed plants. Irrigation should be stopped 1 month before the harvest of apples, pears, and plums; broccoli and cauliflower when furrows are irrigated, and tomatoes used for canning if furrows are irrigated.

WHO guidelines have been debated for their effectiveness to protect public health. A large part of the answer probably lies in the treatment requirements of limit values. In spite of their safety, the stringency of the California standards is acting as a barrier to their widespread adoption worldwide. The WHO point of view is often criticized as being too lax, but the California approach is too technologically intensive and expensive for developing countries. One must realize that where raw wastewater is directly reused, a widespread practice in many developing countries, the WHO guidelines are already a major step forward.

However, other requirements also seem to be necessary to complement WHO guidelines. Based on an extensive analysis of existing guidelines worldwide, the need for developing health-related chemical criteria for land application of reclaimed wastewater has also been reported by the WHO (14). This, added to the fact that the controversy between the 'WHO camp' and the 'California camp' has now been raging for a number of years means that the time has come to update the existing guidelines on an international basis. The door is therefore now open worldwide for a 'third approach' integrating the epidemiological knowledge base generated since 1989 and the latest technological developments in wastewater treatment.

Table 6. Crops Regulated by 1991 Mexican Reuse Guidelines

Salad Vegetables	Vegetables	Herbs	Fruits
Beet root	Beet	Coriander	Blackberries
Celery	Broccoli	Epazote	Melon
Cucumber	Cabbage	Mint	Strawberries
Lettuce	Carrot	Parsley	Sweet turnip
Radish	Cauliflower		Watermelon
Red tomato	Courgette		
Watercress	Garlic		
	Green tomato		
	Mushrooms		
	Onion		
	Papalo		
	Spinach		
	Wild greens		

Table 5. Conditions for Irrigation of Restricted Crops Under 1991 Mexican Guidelines

Type of Irrigation	Type of Wastewater[a]	Minimum Time Between the Last Irrigation and the Harvest (days)	Crops/Banned Crops
Flood	1	20	All crops listed in Table 6, except garlic, cucumber, sweet turnip, melon, and watermelon
	2	20	All crops listed in Table 6, except melon and watermelon
	3	20	All crops listed in Table 6
	4	20	All salad crops, vegetables, and fruits in general
Furrow	1	15	All crops listed in Table 6, except garlic, cucumber, sweet turnip, melon, watermelon, and green tomato
		20	No restrictions
	2	20	All crops listed in Table 6, except garlic, cucumber, sweet turnip, melon, watermelon, and green tomato
	3	20	All crops listed in Table 6, except melon and watermelon
	4	20	All salad crops, vegetables, and fruits in general
Spray	1	20	All crops listed in Table 6, except garlic, cucumber, sweet turnip, melon and watermelon
	2,3,4	20	All salad crops, vegetables and fruits in general

[a] Wastewater Type 1: <1000 total coliforms/100 mL and 0 helminth eggs/liter;
Wastewater Type 2: 1–1000 fecal coliforms/100 mL and ≤1 helminth eggs/liter;
Wastewater Type 3: 1001–100,000 fecal coliforms/100 mL, and no restriction on helminth eggs;
Wastewater Type 4: >100,000 fecal coliforms/100 mL and no restriction on helminth eggs.

BIBLIOGRAPHY

1. Talaro, K.P. and Talaro, A. (1999). *Foundations in Microbiology*. McGraw-Hill, New York, pp. 396–397.
2. Ashbolt, N.J. et al. (1995). The identification of human health significance of environmental aeromonads. *Water Sci. Technol.* **31**: 263–270.
3. Rose, J.B., Lisle, J.T., and LeChevallier, M. (1997). Waterborne cryptosporidiosis: Incidence, outbreaks, and treatment strategies. In: *Cryptosporidium and Cryptosporidiosis*. R. Fayer (Ed.). CRC Press, Boca Raton, FL, pp. 93–110.
4. Ellis, K.V., Rodrigues, P.C.C., and Gomez, C.L. (1993). Parasite ova and cysts in waste stabilization ponds. *Water Res.* **27**: 1455–1460.
5. Alum, A. (2001). *Control of Viral Contamination of Reclaimed Irrigated Vegetables by Drip Irrigation*. Ph.D. Dissertation, University of Arizona, Tucson, AZ.
6. Gerba, C.P., Rose, J.B., Haas, C.N., and Crabtree, K.D. (1996). Waterborne rotavirus: a risk assessment. *Water Res.* **30**: 2929–2940.
7. American Public Health Association. (1989). *Standard Methods for the Examination of Water and Wastewater*. APHA, AWWA, WEF, Washington, DC.
8. National Research Council. (1996). *Use of Reclaimed Water and Sludge in Food Crop Production*. National Academy Press, Washington, DC, p. 91.
9. The Engelberg Report. (1985). *Health Aspects of Wastewater and Excreta Use in Agriculture and Aquaculture*. *Int. Ref. Center Waste Disposal (ICRWD) News* No. 23, pp. 11–18.
10. World Health Organization. (1989). *Health Guidelines for the Use of Wastewater in Agriculture and Aquaculture*. Technical Report Series No. 776. World Health Organization, Geneva. p. 74.
11. SAEPA. (1997). *South Australian Reclaimed Water Guidelines*. Draft 10.1.97. South Australian Environmental Protection Authority.
12. CSHPF. (1991). *Recommandations Sanitaires Concernanat L'Utilisation, Apre's puration, des Eaux Residuaires Urbaines Pour L'irrigation des Cultures etdes Espaces Verts*. Circulaire DGS/SD1.D./91/N 51. Conseil Supérieur díHygiène Publique de France, Paris.
13. Friedler, E. (2001). Water reuse—an integral part of water resources management: Israel as a case study. *Water Policy* **3**: 29–39.
14. Chang, A.C., Page, A.L., and Asano, T. (1995). *Developing Human Health Related Chemical Guideline for Reclaimed Wastewater and Sewage Sludge Applications in Agriculture*. World Health Organization, Geneva.

READING LIST

Asano, T. and Levine, A.D. (1996). Wastewater reclamation, recycling and reuse: past present and future. *Water Sci. Technol.* **33**: 1–14.

US EPA. (1992). *In U.S. EPA, Office of Water and Wastewater and compliance (Ed.). Guidelines for water reuse*. US EPA, Washington, DC.

World Health Organization. (1973). *Reuse of Effluents: Methods of Wastewater Treatment and Health Safeguards*. Technical Report Series No. 517. A report of Meeting of Experts, WHO, Geneva.

SOIL SALINITY

Dennis L. Corwin
Salinity Laboratory
Riverside, California

DEFINITION, SOURCES, EFFECTS, AND GLOBAL IMPACTS

Soil salinity refers to the presence of major dissolved inorganic solutes in the soil solution (i.e., aqueous liquid phase of the soil and its solutes), which consist of soluble and readily dissolvable salts, including charged species (e.g., Na^+, K^+, Mg^{2+}, Ca^{2+}, Cl^-, $HCO_3{}^-$, $NO_3{}^-$, $SO_4{}^{2-}$, and $CO_3{}^{2-}$), nonionic solutes, and ions that combine to form ion pairs. The primary source of salts in soil and water is the geochemical weathering of rocks from the earth's upper strata, with atmospheric deposition and anthropogenic activities serving as secondary sources. The predominant mechanism causing the accumulation of salt in the root zone of agricultural soils is loss of water through evapotranspiration (i.e., combined processes of evaporation from the soil surface and plant transpiration), which selectively removes water, leaving salts behind.

The accumulation of soil salinity can result in reduced plant growth, reduced yields, and, in severe cases, crop failure. Salinity limits plant water uptake by reducing the osmotic potential, making it more difficult for the plant to extract water. Salinity may also cause specific ion toxicity or upset the nutritional balance of plants. In addition, the salt composition of the soil solution influences the composition of cations on the exchange complex of soil particles, which influences soil permeability and tilth.

Irrigated agriculture, which accounts for 35–40% of the world's total food and fiber, is adversely affected by soil salinity on roughly half of all irrigated soils (totaling about 250 million ha), with over 20 million ha severely effected by salinity worldwide (1). As a result of the potential detrimental impacts of soil salinity accumulation, it is a crucial soil chemical property that is routinely measured and monitored.

LABORATORY METHODS OF SOIL SALINITY MEASUREMENT

Although a variety of instrumentation has been developed for the measurement of soil salinity in the field, the most common technique is still laboratory analysis of aqueous extracts of soil samples. In the laboratory, soil salinity is commonly determined from the measurement of the electrical conductivity (EC) of soil solution extracts. The current-carrying capacity of soil solution is proportional to the concentration of ions in the solution. Soil salinity is quantified in terms of the total concentration of the soluble salts as measured by the EC of the soil solution in $dS\,m^{-1}$ (2).

To determine EC, measurements are made in a cell containing two electrodes of constant geometry and distance of separation (3). The soil solution is placed between the two electrodes. An electrical potential is imposed across the electrodes, and the resistance of the solution between the electrodes is measured. The measured conductance is a consequence of the solution's salt concentration and the electrode geometry whose effects are embodied in a cell constant. At constant potential, the current is inversely proportional to the solution's resistance:

$$\mathrm{EC}_T = k/R_T \tag{1}$$

where EC_T is the electrical conductivity of the solution in $\mathrm{dS\,m^{-1}}$ at temperature T (°C), k is the cell constant, and R_T is the measured resistance at temperature T. One $\mathrm{dS\,m^{-1}}$ is equivalent to one $\mathrm{mS\,cm^{-1}}$ and one $\mathrm{mmho\,cm^{-1}}$, where $\mathrm{mmho\,cm^{-1}}$ are the obsolete units of EC.

Partitioning of solutes over the three soil phases (i.e., gas, liquid, solid) is influenced by the soil/water ratio at which the extract is made, so the ratio must be standardized to obtain results that can be applied and interpreted universally. Customarily, soil salinity has been defined in terms of laboratory measurements of the EC of the saturation extract ($\mathrm{EC_e}$) because it is impractical for routine purposes to extract soil water from samples at typical field water contents. One widely used technique is to obtain an extract by vacuum filtration of a saturated soil paste made with distilled water (4). Commonly used extract ratios other than a saturated soil paste are 1:1, 1:2, and 1:5 soil/water mixtures. However, extracts at these ratios only provide relative salinity because the soils are adjusted to unnaturally high water contents not found in the field. Soil salinity can also be determined from the measurement of the EC of the soil solution at some defined field water content ($\mathrm{EC_w}$). An example would be the EC at field capacity, which represents the water content of soil 2–3 days after irrigation when free drainage is negligible. Theoretically, $\mathrm{EC_w}$ is the best index of soil salinity because it is the salinity actually experienced by the plant root. Nevertheless, $\mathrm{EC_w}$ has not been widely used for two reasons: (a) it varies over the irrigation cycle as the soil water content changes and (b) methods for obtaining soil solution samples are too labor and cost intensive to be practical for field-scale applications (5).

Soil solution extracts can be obtained by various means. For disturbed samples, soil solution can be obtained in the laboratory by displacement, compaction, centrifugation, molecular adsorption, and vacuum- or pressure-extraction methods. For undisturbed samples, $\mathrm{EC_w}$ can be determined either in the laboratory on a soil solution sample collected with a soil-solution extractor (see Fig. 1) or directly in the field using *in situ*, imbibing-type porous-matrix salinity sensors (see Fig. 2).

Electrolytic conductivity increases at a rate of approximately 1.9% per degree centigrade increase in temperature in the range of 15–35 °C. Customarily, EC is expressed at a reference temperature of 25 °C for purposes of comparison. The EC (i.e., $\mathrm{EC_e}$ or $\mathrm{EC_w}$) measured at a particular temperature T (°C), EC_T ($\mathrm{dS\,m^{-1}}$), can be adjusted to a reference EC at 25 °C, $\mathrm{EC_{25}}$ (8):

$$\mathrm{EC}_{25} = f_t \cdot \mathrm{EC}_T \tag{2}$$

$$f_T = 0.4470 + 1.4034\exp(-T/26.815) \tag{3}$$

where f_T is a temperature conversion factor.

Theoretical and empirical approaches are available to predict the EC of a solution from its solute composition (3).

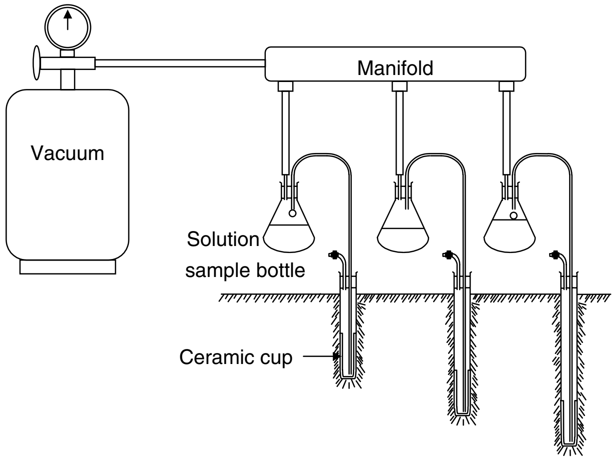

Figure 1. Diagram of a basic suction cup extractor setup for sampling the soil solution. Courtesy of Corwin (6).

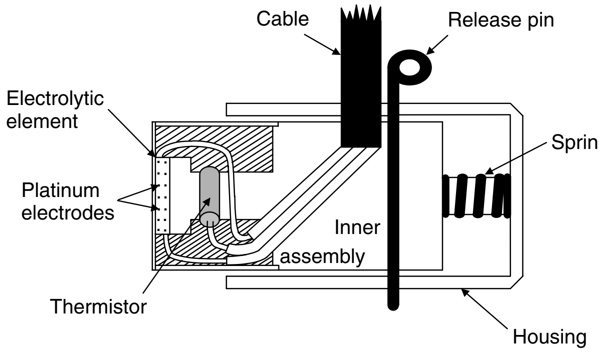

Figure 2. Schematic of a porous-matrix salinity sensor. Courtesy of Corwin (7).

Equation 4 is an example of a theoretical approach based on Kohlrauch's Law of independent migration of ions, where each ion contributes to the current-carrying ability of an electrolyte solution:

$$\mathrm{EC} = \sum_i \mathrm{EC}_i = \sum_i \frac{c_i(\lambda_0^i - \beta\sqrt{c_i}}{1000} \tag{4}$$

where EC is the specific conductance ($\mathrm{dS\,m^{-1}}$), EC_i is the ionic specific conductance ($\mathrm{dS\,m^{-1}}$), c_i is the concentration of the *ith* ion ($\mathrm{mmol_c\,L^{-1}}$), $\lambda^{i}{}_{0}$ is the ionic equivalent conductance at infinite dilution ($\mathrm{cm^2\,S\,mol_c{}^{-1}}$), and β is an empirical interactive parameter (9). Equation 5 shows an empirical equation developed by Marion and Babcock (10):

$$\log \mathrm{TSS} = 0.990 + 1.055 \log \mathrm{EC} \quad (\mathrm{r}^2 = 0.993) \tag{5}$$

where TSS is the total soluble salt concentration ($\mathrm{mmol_c\,L^{-1}}$).

FIELD METHODS OF SOIL SALINITY MEASUREMENT

Serious doubts exist about the ability of soil core samples, soil solution extractors, and porous-matrix salinity sensors (also known as soil salinity sensors) to provide representative soil water samples. Small soil core samples, extractors, and salt sensors do not adequately integrate spatial variability because of their small sphere of measurement; consequently, Biggar and Nielsen (11) suggested that these are "point samples" that can provide qualitative measurement of soil solutions, but not quantitative measurements unless their field-scale variability is established.

As salinity is a highly spatially variable soil property, the use of EC_e to measure salinity at field scales is generally impractical because of the need for hundreds and even thousands of soil samples. The use of EC_e to measure salinity at field scales is only practical when soil sampling is directed using correlated spatial information. Two potential sources of correlated spatial information used to direct where soil samples should be taken to measure EC_e are: (a) visual crop observations and (b) geospatial measurements of apparent soil electrical conductivity (EC_a) using electrical resistivity (ER), electromagnetic induction (EMI), or time-domain reflectometry (TDR). These sources of spatial information are used to direct soil sampling to characterize field-scale soil salinity variation.

Visual Crop Observation

Visual crop observation is a quick and economical method, but salinity development is detected after crop damage has occurred. For this reason, visual observations are unrealistic because crop yields are reduced to obtain soil salinity information. However, remote imagery potentially represents a quantitative approach to visual observation. Remote imagery may offer a potential for early detection of the onset of salinity damage to plants.

Apparent Soil Electrical Conductivity

As a result of the time and cost of obtaining soil solution extracts, developments in the measurement of soil EC have shifted to the measurement of electrical conductivity of the bulk soil, referred to as the apparent soil electrical conductance (EC_a). Apparent soil electrical conductivity measures the conductance through not only the soil solution but through the solid soil particles and via exchangeable cations that exist at the solid-liquid interface of clay minerals. The techniques of ER, EMI, and TDR measure EC_a.

The interpretation of EC_a measurements is not trivial because of the complexity of current flow in the bulk soil. Three pathways of current flow contribute to the EC_a measurement: (a) a liquid phase pathway via salts contained in the soil water occupying the large pores, (b) a solid-liquid phase pathway primarily via exchangeable cations associated with clay minerals, and (c) a solid pathway via soil particles that are in direct and continuous contact with one another (5). As a result of the three pathways of conductance, the EC_a measurement is influenced by several soil physical and chemical properties: (a) soil salinity, (b) saturation percentage, (c) water content, and (d) bulk density. The saturation percentage and bulk density are both closely associated with the clay content. Measurements of EC_a as a measure of soil salinity must be interpreted with these influencing factors in mind. Traditionally, EC_e has been the standard measure of salinity used in all salt-tolerance plant studies. As a result, a relation between EC_a and EC_e is needed to relate EC_a back to EC_e, which in turn is related to crop yield. In spite of these complicating factors, EC_a is the only viable means of spatially characterizing soil salinity at field scales.

As a result of the ease and quickness of measuring EC_a and the ability to mobilize EC_a-measurement equipment, geospatial measurements of EC_a are the most reliable means of characterizing the spatial variability of soil salinity at field scales. Geospatial measurements of EC_a are not used to directly measure soil salinity but rather to direct soil sampling from which EC_e measurements are subsequently made. From the geospatial EC_a measurements and a site-specific relation between EC_a and EC_e, a map of soil salinity distribution across the landscape and through the soil profile can be created. Detailed procedures and protocols for conducting an EC_a survey to characterize the spatial variability of soil salinity with EC_a-directed soil sampling are provided by Corwin and Lesch (12).

ER methods introduce an electrical current into the soil through current electrodes at the soil surface. The difference in current flow potential is measured at potential electrodes that are placed in the vicinity of the current flow (Fig. 3). The electrode configuration is referred to as a Wenner array when four electrodes are equidistantly spaced in a straight line at the soil surface. The two outer electrodes serve as the current or transmission electrodes, and the two inner electrodes serve as the potential or receiving electrodes. The depth of penetration of the electrical current and the volume of measurement increase as the interelectrode spacing, a, increases. For a homogeneous soil, the soil volume measured is roughly πa^3, and the depth of penetration is roughly equal to a. Additional electrode configurations are discussed by Burger (14).

By mounting the electrodes to "fix" their spacing, considerable time for a measurement is saved. Veris Technologies[1] has developed a commercial mobile system for measuring and mapping EC_a using the principles of ER coupled to a global positioning system (GPS). The Veris Model 3100 consists of a pair of transmission coulter electrodes and a pair of potential coulter electrodes to measure the voltage drop. Two sets of arrays permit the measurement of EC_a to depths of 0–30 cm and 0–90 cm.

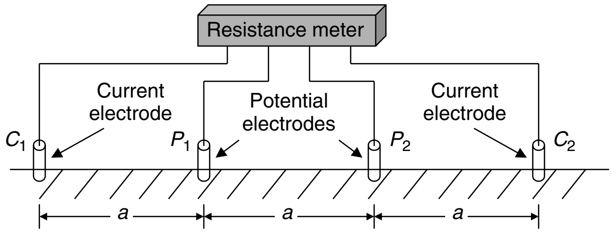

Figure 3. Schematic of electrical resistivity four electrodes, where C_1 and C_2 represent the current electrodes, P_1 and P_2 represent the potential electrodes, and a represents the interelectrode spacing. Courtesy of Corwin and Hendrickx (13).

---

[1] Veris Technologies, Salina, Kansas, USA (www.veristech.com). Product identification is provided solely for the benefit of the reader and does not imply the endorsement of the USDA.

EMI instrumentation consists of a transmitter coil and a receiver coil (Fig. 4). The transmitter coil located at one end of the instrument induces circular eddy-current loops in the soil. The magnitude of these loops is directly proportional to the EC in the vicinity of that loop. Each current loop generates a secondary electromagnetic field that is proportional to the value of the current flowing within the loop. A fraction of the secondary induced electromagnetic field from each loop is intercepted by the receiver coil at the opposite end of the instrument, and the sum of these signals is amplified and formed into an output voltage, which is related to a depth-weighted EC_a. The amplitude and phase of the secondary field will differ from those of the primary field as a result of soil properties (e.g., clay content, water content, salinity), spacing and orientation of the coils, frequency, and distance from the soil surface (15).

The two most commonly used EMI conductivity meters in soil science are the Geonics[2] EM-31 and EM-38. The EM-38 has had considerably greater application for agricultural purposes because the depth of measurement corresponds roughly to the root zone (i.e., 1.5 m), when the instrument is placed in the vertical coil configuration. In the horizontal coil configuration, the depth of the measurement is 0.75–1.0 m. The operation of the EM-38 equipment is discussed in Hendrickx and Kachanoski (15).

Mobile EM equipment developed at the USDA-ARS Salinity Laboratory (16) is available for appraisal of soil salinity and other soil properties (e.g., water content and clay content) using an EM-38 or a dual-dipole EM-38 unit. The dual-dipole EM-38 conductivity meter simultaneously records data in both dipole orientations (horizontal and vertical). The mobile EM equipment is suited for the detailed mapping of EC_a and correlated soil properties at specified depth intervals through the root zone. The advantage of the mobile EM equipment is that it is noninvasive, so it can be used in dry, frozen, or stony soils that are not easily measured with the invasive “fixed-array” approach, which requires good electrode-soil contact. The disadvantage of the EM approach would be that the EC_a is a depth-weighted value, which is nonlinear with depth (17).

ER and EM are both well suited for field-scale applications because their volumes of measurement are large, which reduces the influence of local-scale variability. However, ER is an invasive technique that requires good contact between the soil and four electrodes inserted into the soil; consequently, it produces less reliable measurements in dry, frozen, or stony soils than the noninvasive EM measurement. Nevertheless, ER has a flexibility that has proven advantageous for field application, i.e., the depth and volume of measurement can be easily changed by altering the spacing between the electrodes.

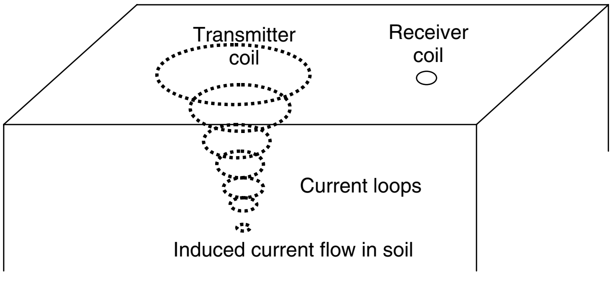

Figure 4. Transmitter coil, receiver coil, and induced current flow for EMI.

Time-domain reflectometry (TDR) was initially adapted for use in measuring water content. Later, Dalton et al. (18) demonstrated the utility of TDR to also measure EC_a, based on the attenuation of the applied signal voltage as it traverses the medium of interest. Although TDR has been demonstrated to compare closely with other accepted methods of EC_a measurement, it is still not sufficiently simple, robust, or fast enough for the general needs of field-scale soil salinity assessment (5). Only electrical resistivity and EM have been adapted for the geo-referenced measurement of EC_a at field scales and larger.

PLANT TOLERANCE TO SOIL SALINITY

The primary effect of soil salinity is on plant growth. Virtually all plant salt-tolerance studies conducted in the past have related plant yield response to EC_e, which is another reason why EC_e continues to be used as the primary measure of soil salinity. A salt-affected soil is one that contains enough soluble salts to hamper the growth of the crop. The proportions of dissolved ions (primarily Na^+, Ca^{2+}, Mg^{2+}, Cl^-, $SO_4{}^{2-}$, and HCO^{3-}) can vary, and the effects on plants can vary. Generally, however, plants respond similarly to salinity over a fairly wide range of combinations of salt. To quantify yield response to soil salinity, Maas and Hoffman (19) proposed a two-piece linear response function with a tolerance plateau of slope zero and a concentration-dependent line whose slope indicates the yield reduction per unit increase in salinity:

$$Y_r = 100 - b\,(EC_e - a) \tag{6}$$

where a is the salinity threshold ($dS\,m^{-1}$), b is the slope expressed in percent per $dS\,m^{-1}$, and EC_e is the mean EC_e taken from the root zone. Values for a and b can be found in Maas (20).

BIBLIOGRAPHY

1. Rhoades, J.D. and Loveday, J. (1990). Salinity in irrigated agriculture. In: *Irrigation of Agricultural Crops*. B.A. Stewart and D.R. Nielsen (Eds.). *Agron. Monogr.* No. 30, SSSA, Madison, WI, pp. 1089–1142.
2. U.S. Salinity Laboratory Staff. (1954). *Diagnosis and Improvement of Saline and Alkali Soils*. *USDA Handbook* 60, U.S. Government Printing Office, Washington, DC.
3. Jurinak, J.J. and Suarez, D.L. (1996). The chemistry of salt-affected soils and waters. In: *Agricultural Salinity Assessment and Management*. K.K. Tanji (Ed.). ASCE, New York, pp. 42–63.
4. Rhoades, J.D. (1996). Salinity: electrical conductivity and total dissolved solids. In: *Methods of Soil Analysis,*

---

[2] Geonics Limited, Mississauga, Ontario, Canada. Product identification is provided solely for the benefit of the reader and does not imply the endorsement of the USDA.

Part 3—Chemical Methods. D.L. Sparks (Ed.). *SSSA Book Series* No. 5, SSSA, Madison, WI, pp. 417–435.

5. Rhoades, J.D., Chanduvi, F., and Lesch, S. (1999). *Soil Salinity Assessment: Methods and Interpretation of Electrical Conductivity Measurements*. FAO Irrigation and Drainage Paper #57, Food and Agriculture Organization of the United Nations, Rome, Italy.
6. Corwin, D.L. (2002). Solute content and concentration—measurement of solute concentration using soil water extraction: suction cups. In: *Methods of Soil Analysis, Part 4—Physical Methods*. J.H. Dane and G.C. Topp (Eds.). *SSSA Book Series* No. 5, SSSA, Madison, WI, pp. 1261–1266.
7. Corwin, D.L. (2002). Solute content and concentration—measurement of solute concentration using soil water extraction: porous matrix sensors. In: *Methods of Soil Analysis, Part 4—Physical Methods*. J.H. Dane and G.C. Topp (Eds.). *SSSA Book Series* No. 5, SSSA, Madison, WI, pp. 1269–1273.
8. Sheets, K.R. and Hendrickx, J.M.H. (1995). Non-invasive soil water content measurement using electromagnetic induction. *Water Resour. Res.* **31**: 2401–2409.
9. Harned, H.S. and Owen, B.B. (1958). *The Physical Chemistry of Electrolyte Solutions*. Reinhold Publ. Corp., New York.
10. Marion, G.M. and Babcock, K.L. (1976). Predicting specific conductance and salt concentration of dilute aqueous solution. *Soil Sci.* **122**: 181–187.
11. Biggar, J.W. and Nielsen, D.R. (1976). Spatial variability of the leaching characteristics of a field soil. *Water Resour. Res.* **12**: 78–84.
12. Corwin, D.L. and Lesch, S.M. (In press). Characterizing soil spatial variability with apparent soil electrical conductivity: I. Survey protocols. *Comput. Electron. Agric. Agriculture.*
13. Corwin, D.L. and Hendrickx, J.M.H. (2002). Solute content and concentration—Indirect measurement of solute concentration: electrical resistivity—wenner array. In: *Methods of Soil Analysis, Part 4—Physical Methods*. J.H. Dane and G.C. Topp (Eds.). *SSSA Book Series* No. 5, SSSA, Madison, WI, 1282–1287.
14. Burger, H.H. (1992). *Exploration Geophysics of the Shallow Subsurface*. Prentice Hall PTR, Upper Saddle River, NJ.
15. Hendrickx, J.M.H. and Kachanoski, R.G. (2002). Solute content and concentration—indirect measurement of solute concentration: nonintrusive electromagnetic induction. In: *Methods of Soil Analysis, Part 4—Physical Methods*. J.H. Dane and G.C. Topp (Eds.). *SSSA Book Series* No. 5, SSSA, Madison, WI, 1297–1306.
16. Rhoades, J.D. (1993). Electrical conductivity methods for measuring and mapping soil salinity. In: *Advances in Agronomy*. Vol. 49. D.L. Sparks (Ed.). Academic Press, San Diego, CA, pp. 201–251.
17. McNeil, J.D. (1980). *Electromagnetic Terrain Conductivity Measurement at Low Induction Numbers*. Tech. Note TN-6, Geonics Limited, Ontario, Canada, pp. 1–15.
18. Dalton, F.N., Herkelrath, W.N., Rawlins, D.S., and Rhoades, J.D. (1984). Time-domain reflectometry: simultaneous measurement of soil water content and electrical conductivity with a single probe. *Science*. **224**: 989–990.
19. Maas, E.V. and Hoffman, G.J. (1977). Crop salt tolerance—current assessment. *J. Irrig. Drainage Div. ASCE.* **103**(IR2): 115–134.
20. Maas, E.V. (1996). Crop salt tolerance. In: *Agricultural Salinity Assessment and Management*. K.K. Tanji (Ed.). ASCE, New York, pp. 262–304.

MAINTAINING SALT BALANCE ON IRRIGATED LAND

Jacob W. Kijne
Herts, United Kingdom

DEFINITION

The salt balance of an area is described by an equation that compares the movement of salt into and out of an area. The salt fluxes are in balance when inflow and outflow of salt are equal. Salt moves in solution, and the water and salt balances are therefore linked as in Eq. 1, which describes the salt balance of an irrigated area:

$$IC_{\mathrm{i}} + RC_{\mathrm{r}} + GC_{\mathrm{g}} = PC_{\mathrm{p}} + \Delta S \tag{1}$$

where: I = irrigation water infiltrating the root zone
R = rainfall entering the root zone
G = capillary flow from the groundwater into the root zone
P = deep percolation from the root zone
C = salt concentration, and the subscripts referring to irrigation, rainfall, capillary flow, and percolation
ΔS = the change in salt content of the soil solution in the root zone

The water flows can be expressed in m^3/ha/day or any other convenient set of units, such as mm/month, in which the unit of time indicates the period over which the salt balance is measured. The unit of salt concentration, C, is usually mg/L or part per million (ppm). Apart from areas close to the sea, it is reasonable to assume that C_{r} is zero. It is also often assumed that C_{g} equals C_{p}, which is reasonable for annual periods but not necessarily true for shorter periods.

For steady-state conditions, the salt concentration in the soil solution remains the same, and hence, ΔS equals zero. The ratio of the amounts of drainage water (percolating from the root zone) and the amount of water entering the root zone as irrigation and rainfall is called the leaching fraction, LF. Hence:

$$\mathrm{LF} = P\Delta t/(I+R)\Delta t = P/(I+R) \tag{2}$$

where Δt equals the time period under consideration.

The ratio of the electrical conductivity (EC) of the drainage water to that of the applied water ($A = I + R$) can be written as $\mathrm{EC_p/EC_a}$. For steady-state conditions, this ratio is proportional to the inverse of the leaching fraction (1/LF). For solutions consisting mainly of chloride salts, the proportionality constant is 1. For concentrated solutions, the proportionality constant is less than 1. Often irrigation water also has a proportionality constant of less than 1, and its value depends on the LF. For example, the proportionality constant of Colorado River water dropped to below 1 when LF became less than about 20% (1). The lower the value of LF, the lower the proportionality constant. Leaching fractions of less than 20% are common in irrigation practice. The relationship between $\mathrm{EC_p/EC_a}$ and 1/L is also influenced by the solubility of soil lime, which is a function of the partial pressure of $\mathrm{CO_2}$.

PRESENCE OF SALTS

The salts carried in irrigation water come from salts that are continuously released from rocks and soils by mineral weathering and dissolution. Because of their low rainfall, most arid and semiarid lands are rich in these primary salts. The salts were deposited over time in sedimentary rock layers and are dissolved in the waters of aquifers. Unfavorable salt balances are common in arid lands: The global extent of salt-affected lands is considerable. It has been estimated that worldwide 100 million ha of irrigated land suffer from salinity, 20% of which is severely affected (2). Annually some 2 to 3 million ha of potentially productive agricultural land are taken out of production because of salinization. How much of this land is reclaimed (to various degrees) and then cultivated again is unknown. Wide divergence in the salinity figures is reported by different institutions, probably because of incomplete data and different definitions of salinity-affected lands; e.g., estimates for salt-affected land in India, according to different sources, range from 7 to 26 million ha, or between 17% and 60% of the irrigated land. For Pakistan, the most likely figure is some 40%, Israel 13%, Australia 20%, China 15%, Iraq 50%, and Egypt 35% of the irrigated land (3).

In many coastal regions, salt water intrusion has a strong negative effect on the regional salt balance. The problem is associated with groundwater extraction from deep fractured aquifers in the coastal region, often for drinking water in a coastal city. Sea water intrusion into the delta of rivers also upsets the regional salt balance. Here the cause is usually a much diminished outflow from the river because of upstream extractions for agricultural and industrial use. Examples include the Colorado River in the United States, the Murray in South Australia, and the Rhine in the Netherlands.

SALT BALANCE OF THE ROOT ZONE

Soluble salts added with the irrigation water must be removed by leaching to maintain the salt balance in the root zone of irrigated land. If not, crop yields will drop and—in case of predominantly sodium salts—soil crusting is likely to occur with an adverse effect on infiltration and hydraulic conductivity of the soil. The less soluble salts gradually precipitate when solutions become more concentrated and are again mobilized when the solutions become diluted by mixing with freshwater. Depending on whether precipitation or dissolution of salts (e.g., of lime or gypsum) occurs in the soil, the amount of salt leached from the root zone must be less than or more than that applied by irrigation. Generally, again depending on the occurrence of precipitation and dissolution reactions, when too much irrigation water is applied to the soil, more salts are added with the water, and in consequence, the drainage water also contains more salts than necessary. If the drainage water is returned to the river downstream, water quality becomes worse. Salt concentrations in the root zone must therefore be controlled to sustain irrigated agriculture over the long term.

When the quality of irrigation water deteriorates, for example, by extensive use of (slightly) saline groundwater when canal water became scarce, nonsteady conditions occur and cation reactions will change the salt load of drainage water. Because of exchange of sodium and magnesium for calcium on the clay particles, mineral precipitation could double in a soil profile that was initially high in exchangeable calcium. These nonsteady conditions could last for many years before a new steady state evolves. Irrigation thus interferes with the geohydrological and geochemical regimes of the area. When irrigation was introduced in areas with low groundwater tables, excessive irrigation often occurred and deep percolation was the agent of interference. The main consequences of this increased groundwater recharge are a high watertable and the mobilization of natural salt residues of the semiarid lands. The latter is referred to as salt pickup.

The relationship among the salinity of the applied water, the leaching fraction, and the resulting soil salinity is an important one. Crop yields and the salinity of drainage water could be more accurately predicted if one could unambiguously calculate the soil salinity resulting from irrigation applications of known salinity and at a specified leaching fraction. Table 1 presents various relationships among leaching fraction (LF), the dimensionless ratio of the average weighted root-zone salinity (Cs), and the average salinity of applied water (Ca).

The values in the table are based on steady-state conditions. Because of differences between the underlying assumptions, considerable variation exists in the calculated ratios of root zone salinity to irrigation water salinity. For the four methods of Table 1, the salinity-tolerance data are from threshold salinity-response functions, and the leaching equations ignore the effect of sodium salts on the soil structure. The relation between soil and water salinity as governed by leaching is a dynamic one and hard to predict, subject as it is to site-specific feedback mechanisms among growth of the crop (hence, their water use through evapotranspiration), weather conditions (5), and the leaching of salts (6). A contributing factor is the variability in measured values of the EC of saturated-soil extracts. The coefficient of variation of the EC of soil moisture at saturation is about 50% (7).

Below a leaching fraction of 0.1, a small change in the leaching amount can make a large difference in root zone salinity. The ratio of root zone salinity to irrigation water salinity is less sensitive to changes in the leaching amount at LF values between 0.1 and 0.4, which are most common. In this range of LF values, root-zone salinity increases about linearly with the salinity of the applied water. With the introduction of precision irrigation to reduce water losses and reduce drainage flows, leaching fractions tend

Table 1. Relationships Between Leaching Fraction and Ratio of Soil Salinity Over Applied Water Salinity

LF	Cs/Ca[1]	Cs/Ca[1]	Cs/Ca[1]	Cs/Ca[1]
0.05	3.00	7.00	4.0	10.50
0.1	2.00	5.00	2.6	5.50
0.2	1.25	3.00	1.4	3.00
0.3	1.00	2.50	1.3	2.15
0.4	0.83	2.35	1.0	1.75

Source: Kijne (4).

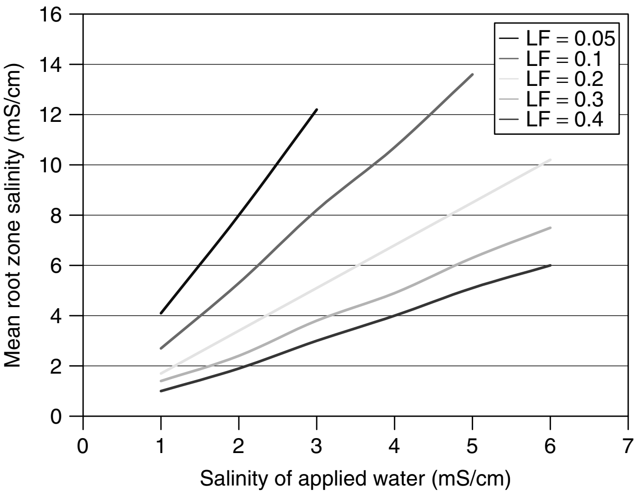

Figure 1. Root zone salinity as a function of the applied water and the leaching fraction. [*Source*: Hoffman and van Genuchten (8)].

to be lower with an enhanced risk of salt accumulation in the root zone.

The relationship of the fourth column in Table 1 uses a root water uptake function that is exponential with depth. This function is often used in modeling studies where a relationship between leaching and root-zone salinity is required. It is plotted in Fig. 1.

LEACHING OF SALT TO PREVENT YIELD REDUCTION

In irrigation practice, the leaching requirement is usually linked to the expected yield reduction. During leaching, soil salinity in the root zone increases with depth. In planning the desired leaching fraction for crop production purposes, it is commonly assumed that the EC at the lower root zone boundary corresponding to 25% to 50% yield reduction is still acceptable. The weighted average EC value for the entire root zone (weighted according to root distribution) would then be less than at the lower root zone boundary, and the resulting yield reduction is assumed to be less than 25% to 50%. This result gives some indication of the required leaching fraction based on the relation between yield and soil salinity.

The rate of downward flow is highest shortly after irrigation when the soil water content is still above or near field capacity. Thereafter leaching continues at a much-reduced rate. When the soil profile dries between irrigations, the soil solution becomes more concentrated. In many soils, the soil solution at field capacity is about twice as concentrated as when the soil is saturated. A further complication is that not all downward flow is equally effective in leaching salts from the root zone. Leaching is more effective when water moves through the soil mass, rather than through cracks between aggregates. Water moving through cracks and wormholes is called preferential flow. The efficiency of leaching depends on the fraction of irrigation water moving down the soil profile mixing with the soil solution. At the start of the leaching process as much as three-quarter of the applied water may flow through cracks in clay soils. Once the soil swells up with moisture, cracks close and the leaching process becomes more efficient.

LEACHING CONTROL

Leaching practices influence the quality and quantity of drainage water. Their control is important when rising watertables are a problem (waterlogging) or where the salt load in the drainage water needs to be brought down. In such cases, restrictions are placed on the amount of salt farmers are allowed to discharge in the drainage water. Experience in Australia and the United States indicates that strict adherence to a set of site-specific regulations on the amount and quality of drainage flows and recharge to the groundwater is necessary to maintain the salt balance in irrigated agriculture. Because of the presence of monitoring systems, farmers in those countries can make well-informed choices about the crops to be grown and the cropping intensities they adopt. Thus, agricultural choices are made within a strict set of rules and the farmers are made responsible for the consequences of their decisions. The successful introduction of regulations with respect to permissible pumping from and recharge to the groundwater, including salt loading, in developing countries will be conditional on the development of adequate monitoring systems and management institutions.

RIVER SALINITY

All arid zone rivers have natural salt profiles resulting from mobilization of salts in the catchment area and saline seeps. Also, irrigation-induced transport contributes to river salinity as drain water with fossil salts from pumped groundwater discharges into a river. Increases in the salinity of rivers and streams in many dry parts of the world pose an ecological hazard that has largely been overlooked. This lack of attention for the ecological impact of increased salinity in inland waters is unfortunate considering the vulnerability of aquatic ecosystems to increased salt levels (9).

Salinity profiles for four rivers are shown in Fig. 2, which illustrates the various degrees to which salts are returned to the river or remain in the land and the groundwater (10). The data used in Fig. 2 are averages that ignore the seasonal changes in salt load as affected by the flow regime of the rivers. Most drainage water from agricultural land in Pakistan's Punjab is being reused, either from surface drains or pumped up from shallow groundwater. In fact, in some irrigation systems, one half to two thirds of the irrigation water is pumped from the groundwater. According to one estimate, about 116 million tons of salt are added annually to the irrigated land in Pakistan (11). Because of the absence of extensive drainage, the leached salts are returned to the land rather than disposed of in the Indus River system or in evaporation ponds. The average annual salt load of the Indus River water measured at rim stations in the catchment is estimated at 33 million tons, whereas the outflow to the sea contains only 16.4 million tons. Hence about half of the annual salt influx remains in the land

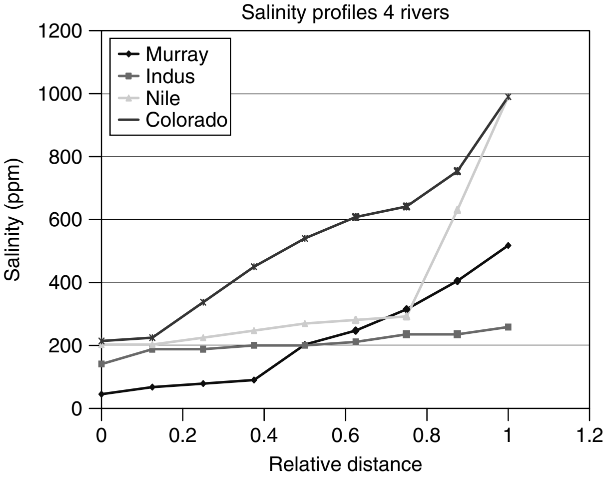

Figure 2. Salinity profiles of four rivers. [*Source*: Smedema (10)].

and in the groundwater (10). Most accumulation takes place in the Punjab, which is in sharp contrast with Egypt where a large portion of the irrigated land is underlain by subsurface drains that take the drainage water back to the Nile River. The salts do not stay in the Nile Basin but are discharged into the Mediterranean Sea. During part of the year, the salt content in the lower Indus is much lower than in the lower Nile and more salt disposal into the Indus could be accepted. However, during critically low flow periods, such disposals would not be possible as the salinity of the Indus water would exceed acceptable limits. The only option during those periods would be to store the drainage water temporarily for release during high flood periods. Extending the Left Bank Outfall Drain, now operating in Sindh, into Pakistan's Punjab may provide a more permanent, but also more expensive, solution than the present inadequate number of evaporation ponds.

The salt load of the Murray River shows considerable seasonal changes. At high flows when most of the river flow originates in the upper catchment, the salt influx is low, but at low flows of the river, when more of the river flow comes from groundwater base flow, the salt influx is higher. At low flows, salinity levels of the Murray are sometimes lowest downstream as the salinity of some tributaries is so much less than that of the main river that the water is diluted.

The salinity profile of the Colorado River is the result of large-scale irrigation development in the basin and concentration of the river water because of evaporative losses from storage reservoirs along its course. Under the salinity control projects of the Colorado River basin, several measures are undertaken to reduce further increases in river salinity. These measures include control of canal seepage by lining, restrictions on the quality of drainage flow back into the river, on-farm water savings by land leveling and changes in cropping patterns, retirement of irrigated land, and watershed management.

These examples of analyzing river salinity profiles clearly demonstrate their potential importance in understanding salt balances of irrigated lands at the scale of river basins.

MEASURING AND MODELING THE SALT BALANCE

Although the methodology for monitoring salinity at field level has been developed, using it in practice is problematic. The ability to diagnose and monitor field-scale salinity has improved considerably during the last 30 years or so. The apparent electrical conductivity of the soil can be measured by means of four-electrode sensors, either surface array or insertion probes, remote electromagnetic induction sensors (e.g., EM31 and EM38), and time-domain reflectometric sensors. Estimation of solute concentration from such measurements requires a calibration model relating soil solution EC to bulk EC. Many examples of the application and calibration of this technique illustrate the limited usefulness of manufacture-provided calibration curves and the different degrees of sophistication of the calibration models (12).

Various models have been developed to simulate transport through soils [e.g., LEACHM (13), SWAP (14)] that require soil solution EC data for their validation. These models are based on physical processes of water infiltration, redistribution, and evaporation. They enhance the understanding of water flow and chemical transport behavior and have been applied to simulate the terms of the water and salt balance (15). Widespread application of the models is restricted by the need for calibration with locally collected data and matching adjustment of some of the parameters. For example, in an Indian study with LEACHM-C, the root mean square errors (RSME) in the comparison of observed and predicted values of soil profile salinity ranged from 28% to 70%. After the water retention parameters were adjusted, the RMSE values ranged from 13% to 24% (16).

Similar problems of uncertainty in model application are associated with high watertables developing from excessive percolation in irrigated areas without adequate drainage. Plant roots take up water from shallow watertables as capillary rise meets part of the water demand of the crop. But if the shallow groundwater is saline, salt accumulation in the root zone affects crop production. Several models have been applied to this situation to ascertain the correct balance between the potentially positive and negative effects of shallow groundwater (17).

The spatial variability of soils and the temporal variability of the hydraulic parameters, especially in clay soils, as well as the largely unknown quantity of lateral groundwater flow constrain the applicability of models to larger areas. Yet, the focus of attention should be changed to the larger scale because of increased recycling of drainage water either through pumping from groundwater or by reusing water from surface drains. It is no longer sufficient to maintain a favorable salt balance at field scale; the salt balance should be maintained at system or river basin scale. As is apparent from the examples of river salinity mentioned earlier, the need for a holistic approach to the salt balance of irrigated areas is large and urgent.

BIBLIOGRAPHY

1. Oster, J.D. and Rhoades, J.D. (1990). Steady-state root zone salt balance. In: *Agricultural Salinity Assessment and*

Management. K.K. Tanji (Ed.). American Society of Civil Engineers, New York, pp. 469–503.

2. Ghassemi, F., Jakeman A.J., and Nix, H.A. (1995). Salinisation of land and water resources: Human causes, management and case studies. CABI/University of New South Wales Press Ltd., Melbourne, Australia.
3. Gleick, P.H. (1993). Water in crisis. A guide to the world's fresh water resources. Pacific Institute for Studies in Development, Environment, and Security, Stockholm Environmental Institute, Oxford University Press, Oxford, UK.
4. Kijne, J.W. Water productivity under saline conditions. In: *Water Productivity in Agriculture: Limits and Opportunities for Improvement*. J.W. Kijne, R. Barker, and D. Molden (Eds.). CABI Publishing, Wallingford, UK, pp. 89–102.
5. Shalhevet, J. (1994). Using water of marginal quality for crop production: major issues. *Agricultural Water Management* **25**: 233–269.
6. Dinar, A., Rhoades, J.P., Nash, P., and Waggoner, B.L. (1991). Production functions relating crop yield, water quality and quantity, soil salinity and drainage volume. *Agricultural Water Management* **19**: 51–66.
7. Tedeschi, A., Beltran, A., and Aragues, R. (2001). Irrigation management and hydrosalinity balance in a semi-arid area of the middle Ebro river basin (Spain). *Agric. Water Management* **49**: 31–50.
8. Hoffman, G.J. and van Genuchten, M.Th. (1983). Water management for salinity control. In: *Limitations to Efficient Water Use in Crop Production*. H. Taylor, W. Jordan, and T. Sinclair (Eds.). American Society of Agronomy Monograph. Madison, WI, pp. 73–85.
9. Williams, W.D. (1987). Salinization of rivers and streams: An important environmental hazard. *Ambio* **16**: 180–185.
10. Smedema, L. (2000). Irrigation-induced river salinization: Five major irrigated basins in the arid zone. International Water Management Institute, Colombo, Sri Lanka.
11. Chaudhry, M.R. and Bhutta, M.N. (2000). Soil and water contaminations from agro-industrial chemical in Pakistan. In: *Proceedings Agroviron 2000: 2nd International Symposium on New Technologies for Environmental Monitoring and Agro-Applications*. Trakya University, Tekirdag, Turkey, pp. 499–494.
12. Triantafallis, J., Laslett, G.M., and McBratney, A.B. (2000). Calibrating an electromagnetic induction instrument to measure salinity in soil under irrigated cotton. *Soil Science Society of America Journal* **64**: 10009–1017; G. Amente, J.M. Baker, C.F. Reece. (2000). Estimation of soil solution electrical conductivity from bulk soil electrical conductivity in sandy soils. *Soil Science Society of America Journal* **64**: 1931–1939.
13. Wagenet, R.J. and Hutson, J.L. (1987). LEACHM: leaching estimation and chemistry model. Version 2. Water Resources Institute Consortium, Center for Environmental Research, Cornell University, Ithaca, NY.
14. van Dam, J.C. et al. (1997). Theory of SWAP version 2.0. Technical Document 45. Wageningen Agricultural university and DLO Winand Staring Centre, Wageningen, the Netherlands.
15. Trantafallis, J., Huckel, A.I., and Odeh, I.O.A., (2003). Field-scale assessment of deep drainage. *Irrigation Science* **21**: 183–192.
16. Hagi-Bishow, M. and Bonnell, R.B. (2000). Assessment of LEACHM-C model for semi-arid saline irrigation. *International Commission on Irrigation and Drainage (ICID) Journal* **49**(1): 29–42.
17. Prathapar, S.A., Meyer, W.S., Bailey, M.A., and Poulton, D.C. (1996). A soil and groundwater simulation model: SWAGSIM. *Environmental Software* **11**: 151–158; X. Wang, P.H.J. Hollanders, S. Wang, S. Fang. (2004). Effect of field groundwater table control on water and salinity balance and crop yield in the Qingtongxia Irrigation District, China. *Irrigation and Drainage* **53**: 263–275.

READING LIST

Topp, G.C., Reynolds, W.D., and Green, R.E. (Eds.). (1992). *Advances in Measurement of Soil Physical Properties: Bringing Theory into Practice*. *SSSA Special Publication* 30, Soil Science Society of America, Madison, WI.

Skaggs, R.W. and van Schilfgaarde, J. (1999). *Agricultural Drainage*. Agronomy Monograph 38. American Society of Agronomy, Madison, WI.

SALT TOLERANCE

Suresh K. Gupta
CSSRI
Karnal, Haryana, India

Salinity has adversely affected irrigated agriculture for thousands of years, probably since the first human civilization in Mesopotamia. The fall of this ancient civilization is attributed to both waterlogging and soil salinity (4). David Shedidan in his book "Desertification of the United States" calls soil salinization of irrigated lands one of the deadliest forms of desertification (5). No other region, where irrigated agriculture is practiced, can be an exception to this kind of desertification. In recent years, soil salinity has assumed a worldwide dimension and has become a serious threat to at least one-third of irrigated lands in the arid and semiarid regions where lots of irrigation water is used without providing adequate drainage. Although a long-term solution to the problem is to carry out improvements in land drainage, differences in salt tolerance of crops provide opportunities to continue growing increasingly salt-tolerant crops on salty lands to buy time before investments are arranged for reclamation.

SALT TOLERANCE: DEFINITION

The literal meaning of tolerance is to endure, sustain, or put up with. Thus, the salt tolerance of a plant could be defined as the plant's capacity to endure the effects of excess salt in the medium of root growth (1). That plants can withstand a fair amount of salts without adverse effects is implicit in this definition. In fact, a fair proportion of salts in the growing medium is needed for good growth. There are instances where plant growth is stimulated at low concentrations of salts, and only thereafter, the yield starts declining (2). Apparently, the problem of soil salinity develops only when the salts accumulate in the root zone to an extent that it becomes harmful to the plants. The salt tolerance of a plant is also expressed in terms of the yield decrease expected for a given level of salinity compared to the yield under nonsaline conditions (3). Many times, salt tolerance and salt resistance have been used synonymously although, the two do not convey the same meaning. Salt resistance is a plant's ability to decrease or prevent the stress from penetrating into its tissues, but tolerance is the ability of the plant to eliminate, reduce, or repair the stress once it has already penetrated.

Salt or salinity in this article is used as a general term for salt stress, which could manifest itself from altogether different causes and compositions of salts under different situations. The stress could be due to bulk salinity of the soil water caused by excess soluble salts. Electrical conductivity (EC) of the saturation extract (EC_e) of a bulk soil solution has been widely used as a measure of this kind of stress (6). The adverse effects could also appear due to high exchangeable sodium, usually expressed in terms of exchangeable sodium percentage (ESP), due to excess ions such as chlorides and sodium expressed in meq L^{-1}, or due to excess toxic ions such as boron and fluorine, expressed in mg L^{-1}. A plant tolerant to one kind of stress need not necessarily be tolerant to other kinds of stresses. Therefore, a large number of relative tolerance tables are now available in the literature (1). These tables can provide a first approximation of the relative tolerance of plants to the anticipated kind and level of stress.

BIOLOGICAL RESPONSE TO SALINITY

The response of plants to salinity has been extensively studied and widely reported. For the purpose of discussions in this section, salinity effects are grouped under the heads (1) the effect of excess soluble salts, (2) the effect of high exchangeable sodium, and (3) a specific ion effect or toxicity of specific ions.

Effect of Excess Soluble Salts

Adverse effects appear to be related to the bulk soil salinity of the soil solution. As the salt concentration increases, the osmotic pressure of the soil water increases which affects the availability of water to the plants. When the osmotic pressure of the soil solution is high, the plants are unable to extract water as readily as they can from a relatively nonsaline soil. As such, even if water is present in the soil, it might not be available to the plants and results in physiological drought or osmotic desiccation. To survive, the plants must adjust osmotically and build up high internal solute concentrations, as is the case with salt accumulating halophytes. In salt sensitive plants, excess amounts of salt in the plant cells can be harmful because of damage to cell membrane, enzymes, and organelles.

The limited uptake of water hinders the uptake of ions by the plant root system. Thus, plants might encounter deficiencies of nutrients. Because the soil solution has high concentrations of certain ionic species, some of these ions, which are otherwise not harmful, might become toxic when absorbed in excess amounts. Salt-sensitive glycophytes have inadequate control over ion uptake, so such excess absorption of ions cannot be ruled out. Besides, preferential absorption of one ion may also retard the absorption of other essential plant nutrients necessary for the normal growth of plants. Evidence of potassium deficiencies and inhibition of absorption of nitrate from saline soils can be cited.

Even glycophytes, which are salt tolerant and possess the ability to adjust osmotically, exhibit growth reduction largely on account of energy expenditure for uptake and synthesis of solutes.

Exchangeable Sodium

High ESP of a soil is associated with structural deterioration of soils. As a result, the edaphic environment is adversely affected and results in restricted air and water movement in the root zone. The factors that limit the uptake of water in alkali/sodic soils are a shallow wetting zone, an impermeable layer (crust) on the surface, low unsaturated hydraulic conductivity, and low available moisture content. Besides water availability, sodium toxicity, as a result of high ESP, can also affect the yield. High ESP can also disturb the mineral nutrition of the plant because it results in nutritionally inadequate amounts of Ca, or it might retard the absorption of some ions (7).

Specific Ion Effects

Some constituents of salts, for example, chlorides, which are harmless to most crops at moderate concentrations, may be toxic to some crops at higher concentrations, for example, citrus and grapes. In addition, some elements like lithium, and boron are toxic to plants even though they may be present in only very small quantities. Under such situations, plants would be subjected to toxic effects, nutritional imbalance, and deficiencies of certain essential elements.

In all the cases discussed in this section, it is believed that the adverse effects are usually due to the cumulative effects of several factors, although one may be dominating others in many cases.

DIFFERENCES IN SALT TOLERANCE OF PLANTS

Plants differ greatly in tolerance to salts. A wide range of salt-tolerant plants that evolved over millions of years exists in the plant kingdom. Plants with natural salt tolerance can survive in water of varying degrees of saltiness, from slightly brackish in the case of barley to ocean water in some hardy species of mangrove trees. Plant species called 'halophytes' are highly salt tolerant; some of them withstand salt concentrations in excess of that of seawater. At least 1500 of them are currently known (8), many of them have potential for commercial exploitation (9). These plants contain genes that protect them from salts in a variety of ways, such as growing salt-blocking membranes, excreting salt through their leaves, or sequestering salt in microscopic compartments called vacuoles. The natural vegetation of arid regions, xerophytes, also has high salt tolerance. On the other hand, almost all crop species are glycophytes, plants that do not grow well in the presence of salts. Most fruit trees and vegetables are salt-sensitive plants. Generally, vegetable crops are more tolerant than fruit crops but are more sensitive to salt than field and forage crops. Within field and arable crops, beans, pulses, oil-seed crops, and cereal crops, in that order, show increasing sensitivity to salts. Many wild relatives of cultivated plants, however, are reportedly highly tolerant to soil salinity, alkalinity, and other associated conditions (10).

VISIBLE SYMPTOMS ON PLANTS GROWING UNDER HIGH SALINITY

Within a narrow range around a plant's salt tolerance level, no visible symptoms might appear on plants under field situations Yet, the crop yield might still be affected. At high salinity, the plants show signs of wilting, a similar symptom that appears when a plant is water stressed. The leaves of salt stressed plants may turn yellow, a condition called chlorosis. It happens because the excess sodium blocks iron uptake and the lack of iron inhibits production of chlorophyll. Stunted growth and deep bluish green foliage of plants followed by low production are other common signs in plants grown under high salinity. The toxic effects that are characteristic of the specific ion causing toxicity usually show on plant parts.

FACTORS INFLUENCING SALT TOLERANCE

Absolute tolerance to salinity cannot be expressed explicitly. It varies considerably with edaphic and environmental factors, which influence salt tolerance in plants. The most important among the edaphic factors are the physical condition of the soil, soil fertility, salt species present in the soil solution, salt distribution in the profile, salinity level, and irrigation methods. The environmental factors include temperature, humidity, radiation, and pollutants in the environment. It has been found that poor aeration as a result of inadequate drainage may adversely affect plant response to soil salinity. For several crops, sensitivity to salts is known to differ markedly according to the type of salts present and as a result, isosmotic solutions of different composition might affect the plants differently. Sorghum plants exhibited greater growth reduction with sulfate type than chloride type salinity (11). Therefore, the type of salinity might alter the order of tolerance in crops. Since salinity stress is expressed as the total sum of the matric potential and the osmotic potential, the salt tolerance of crops would be significantly different in frequently irrigated than infrequently irrigated conditions. As such, shallow and frequent irrigations are recommended for crops growing under a saline/sodic environment. Because of the same reasoning, a plant response at the same salinity is better in drip irrigation than with surface irrigation. Gupta and Dubey (12) reported that irrigation water of much higher salinity could be used for growing vegetable crops with pitcher irrigation than with conventional surface irrigation techniques, as reported by Minhas and Gupta (13). Apparently, the salt tolerance of most crops under pressurized modern irrigation techniques needs to be worked out.

Plants respond better to salt stress in cool and humid than in hot and dry weather. Studies on several crops such as alfalfa, bean, beet, carrot, cotton, onion, squash, strawberry, clover, salt grass, and tomato have shown that salinity decreased yields more when these crops were grown at high temperatures (14–16). Hoffman et al. (17) observed that alfalfa yields were highest under ozone concentrations found in many agricultural areas in California at moderate salinity levels which were otherwise found to reduce growth in the absence of ozone.

Besides the environmental and edaphic factors discussed in the foregoing section, biological factors such as the varieties and rootstock of plants and the stage of growth also cause significant differences in the response of plants to salt stress. Evidence is abounds revealing inter- and intragenic differences in plants' response to salt stress. Green alga (*Dunaliella vindis*) survives in the highly saline Dead Sea at a salinity level exceeding 35 $\mathrm{dS\,m^{-1}}$. A survey of reported data on the response of crop plants to increasing salinity levels reveals about eightfold differences among them in tolerance to salinity stress (18). An evaluation of more than 1000 wheat genotypes revealed a wide variation in the tolerance to soil sodicity (given by the pH of the soil solution). These wheat genotypes could be classified into tolerant, that grew well up to a pH of 9.6 (ESP and pH are well correlated for alluvial soils of the Indo-Gangetic plains) and sensitive, that did not grow well unless the pH was below 8.5. In between these two classes, varieties could be classified as medium tolerant and medium sensitive (19). Aswathappa et al. (20) recorded that some acacia species grew in NaCl concentrations of 1000 $\mathrm{mol\,m^{-3}}$, whereas others stopped growing at 200 $\mathrm{mol\,m^{-3}}$.

Some crops are known to be sensitive to salt stress at germination and others at emergence and early seedling growth. In cereal crops, the early productive stage has been found sensitive. Recent studies indicate that corn (21), sorghum (22), and cowpea (23) are most sensitive during the vegetative and early reproductive stages and least sensitive during the grain filling stage. It has been found that the tolerance of wheat increased with ontogeny. The soil saturation extract values at which yield would decline by 50% is 9.3 $\mathrm{dS\,m^{-1}}$ for seeding to crown root initiation stage and increases with age to 13.2 $\mathrm{dS\,m^{-1}}$ for dough stage to maturity (24). Though data sets are increasingly added to the literature to highlight such differences in response, some criticism is also appearing because of the limitations of the methods used in such evaluations. It is highlighted that presently used conventional methods of soil tests fail to reveal the true stress to which a plant is subjected at different growth stages. For example, it is difficult to say whether the young roots at the time of the emergence and initial establishment stage are really sensitive to salts or are exposed to a high salt stress particularly when the salt accumulate at the soil surface. Similarly, the salinity of the soil water that is approximately equal to that of the irrigation water at the top of the root zone is several times more than the salinity of irrigation water at the bottom. It is natural that at the later growth stages, plant roots moving deep into the profile might experience greater salt stress than that given by the average soil salinity. Thus, one would tend to agree to a general statement that in many crops and plants, increased tolerance with age has been observed, yet the data are not sufficient to draw conclusions in absolute terms about the differences in salt tolerance at different growth stages.

CHARACTERIZATION AND ASSESSMENT OF SALT TOLERANCE

A number of models have been used to appraise the relative tolerance of crop plants. In some models, survival of the plants has been used as a criterion whereas in others, vegetative growth, absolute yield, and relative yield have been used to characterize relative salt tolerance. When comparisons are made among the models, some drawbacks are apparent in a few of them. For example, survival of the plants used in many ecological works is very often at the expense of production. Thus, it seems inappropriate for agricultural situations. Similarly, vegetative growth has a limitation. It has been seen that, where vegetative growth has been affected, many plants pick up later and yield almost at the optimum level. Absolute yield is useful agriculturally, as it tells the farmer what to expect, but has little use for quantifying differences in tolerance between crops with different levels of yield, for example, sugar cane and wheat. The use of relative yield compared to a control (say, non saline environment) allows comparisons between crops and cultivars, where the yield is expressed in different units or where the components of economic yield differ between them. More and more reports are now appearing where crop production functions are plotted in terms of relative yield as a continuous function of increasing salinity or for that matter ESP or the concentration of a particular ion. At this stage, it seems to be the most appropriate criterion for evaluating the relative salt tolerance of plants, although some limitations are known.

Maas and Hoffman (3), probably for the first time, proposed a linear response model, containing three independent parameters, the non saline control yield ($Y_{\max}$), a salinity threshold ($\mathrm{EC_t}$), below which yield (Y) was unaffected, and a slope (S) representing the decrease in yield with increase in salinity above this level. $\mathrm{EC_0}$ represents the salinity at which there is no yield. Mathematically, the response function is written as

$$Y = Y_{\max} \qquad 0 \le \mathrm{EC} \le \mathrm{EC_t} \tag{1}$$

$$Y = Y_{\max} - S(\mathrm{EC} - \mathrm{EC_t}) \qquad \mathrm{EC_t} < \mathrm{EC} \le \mathrm{EC_0} \tag{2}$$

$$Y = 0 \qquad \mathrm{EC} > \mathrm{EC_0} \tag{3}$$

The relative yield of the crop as a ratio of the nonsaline control could also be used in this model. Then, Y would become RY, the relative yield and $Y_{\max}$ would be equal to 1 or 100 depending upon whether the relative yield is expressed in terms of a fraction or percentage.

The piecewise linear model has been used to produce tables of salinity response functions for a number of crops (1,3). Later studies have established that such a model can effectively describe the production function as affected by high ESP (25). A model proposed independently by Oosterbaan is also based on the same premise and uses segmented regressions defining the threshold as a breakpoint. The model proposed by Oosterbaan (26) seems to be quite flexible and could be used for many more situations besides studying the salt tolerance of crops.

Based on this model, threshold salinity, the salinity at which a 50% ($\mathrm{EC_{50}}$) reduction in yield is expected, and the salinity at which the yield is zero ($\mathrm{EC_0}$) have been used to assess the relative salt tolerance of crops. Both the threshold salinity and $\mathrm{EC_0}$ are not well defined (2), so $\mathrm{EC_{50}}$ seems to be the most appropriate parameter to assess the relative salt tolerance of crops (18,27), as it integrates both the slope and the threshold. The data derived from Maas (28) show that both $\mathrm{EC_t}$ and slope overlap between two classes (Table 1). Such overlap gets masked while calculating the $\mathrm{EC_{50}}$. van Genuchten (2) proposed a sigmoid model, defined in terms of $\mathrm{EC_{50}}$ and p where again $\mathrm{EC_{50}}$ is the salinity where the yield is reduced by 50%, and p is an empirical constant defining the shape of the curve. This model also fits well with almost all the data sets of the Maas and Hoffman model, so it attests to the fact that $\mathrm{EC_{50}}$ could be a good parameter to define relative salt tolerance.

The use of this model in field experimentation, besides other limitations, may be difficult where differences between genotypes are small. A limitation of this model is explained well in Fig. 1. Using $\mathrm{EC_{50}}$ as the criterion, it would be concluded that the tolerance of four crops is in the order crop a > crop b = crop c > crop d. On the other hand, within crops b and c, one would tend to conclude that crop b is more tolerant to salt at low and medium salinity levels and crop c is more tolerant at high levels of salinity. Another problem with this model is that it can be used to identify tolerance as distinct from yield potential in many crop cultivars. In wheat, for instance, Kharchia 65, a wheat variety from Rajasthan (India) showed the highest degree of salt tolerance though it was not one of the top yielders amongst the varieties tested at Karnal (India). Notwithstanding such minor difficulties, this approach has the potential of application in categorizing the relative salt tolerance of crops.

RELATIVE SALT TOLERANCE OF CROPS

Extensive data on this issue are part of the literature. The most comprehensive information on this aspect is compiled and reported by Maas (1). This includes information on the relative salt tolerance of herbaceous crops (fiber, grains, grasses, forage, and special crops), vegetable and fruit crops, woody crops, ornamental shrubs, trees, and ground cover crops to the bulk soil salinity of the growing media. It

Table 1. Range of Model Parameters in Describing the Relative Salt Tolerance of Agricultural Crops

Relative Salt Tolerance of Crops	$\mathrm{EC_t}$ (dS $\mathrm{m^{-1}}$)	Slope (%)
Sensitive	1–2	15–35
Moderately sensitive	1–3	6–20
Moderately tolerant	3–7	4–20
Tolerant	4–8	4–7

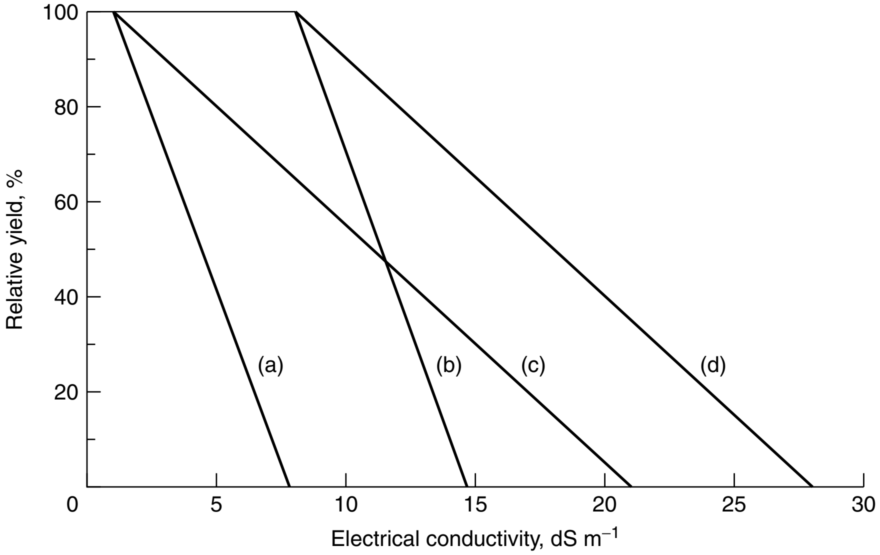

Figure 1. Crop salt tolerance as revealed by the piecewise linear model of Maas and Hoffman.

also lists the tolerance of agricultural crops and fruit-crop cultivars and rootstocks to chloride and the tolerance of agricultural crops and citrus and stone fruit rootstocks to boron. A number of tables listing the relative tolerance of crops have also appeared, but most are based on data reported by Maas and his co-workers. A table showing the relative tolerance of field, vegetable, forage, and fruit crops is given in Appendix 1. Gupta and Sharma (25) reported on the tolerance of a few agricultural crops and grasses to soil ESP. Porwal et al. (29) compiled the available data in India on tolerance of crops to saline water and presented the threshold and slope of the Maas and Hoffman function in terms of irrigation water quality. Per se, saline water might not directly affect crop yield, but once the leaching fractions under a given set of conditions are known, this data could be effectively used to assess crop tolerance to soil salinity. Later on, similar data were compiled and reported by Minhas and Gupta (13).

Although extensive literature is available on the relative tolerance to salts, considering the different agroclimatic conditions, different methods employed, varying farming conditions, and varieties used, the results obtained do not always agree. As an example, the alkali tolerance of rice in the literature has been rated as medium, but in the Indian context, it has been rated as highly tolerant. This difference arises because of the fact that the earlier rating was for direct seeded rice and the India rating is based upon transplanted rice where 35–40 day seedlings have been used for transplanting.

SALT TOLERANCE AND MANAGEMENT OF POOR QUALITY LAND AND WATER RESOURCES

Several treatments and management practices can reduce the salt level in soil. However, there are some situations where it is either not possible or not economically practical to attain desirably low soil salinity levels. In such cases, proper plant selection is one way to moderate yield reductions caused by excessive soil salinity. The ability of a plant to grow on salty lands or to remove salt from such soil makes salt-tolerant plants potentially useful for restoring salt-damaged soils. Salt-tolerant plants can also provide farmers with greater flexibility in crop selection besides reducing the leaching requirement. The difference in the salt tolerance of plants is now being exploited both on experimental and field scales to use salty water resources for crop production and to dispose of them in an ecofriendly manner with reduced investments. It is achieved through sequential concentration of water. In such a system, following each cycle of use, the volume of water decreases, but its salinity increases. Thus, after each cycle, a more salt tolerant crop is chosen before finally disposing of the water in an evaporation tank or by some other means (Fig. 2). This kind of practice has also been referred as a multiple water use system in the literature. The quantity of water to be disposed of at this stage is much less, so the investment in an evaporation tank is significantly reduced.

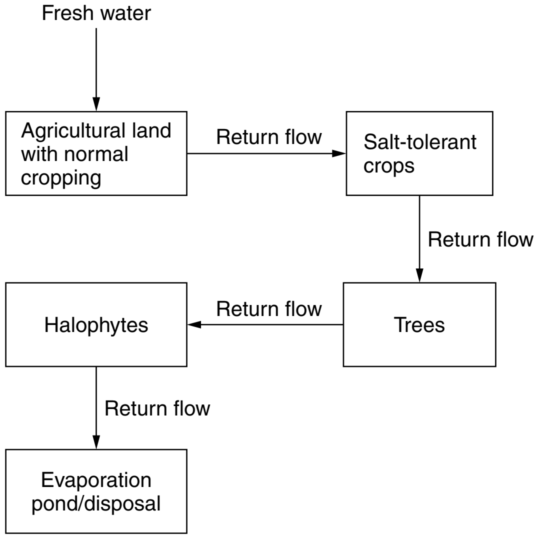

Figure 2. A schematic representation of sequential concentration for managing water resources.

RECENT DEVELOPMENTS IN SALT TOLERANT CROPS

Conventional breeding for developing salt-tolerant plants has been attempted for a long time, and the subject of breeding for salt tolerance has been reviewed by many workers (30). The lack of success in generating tolerant varieties [given the low number of varieties released and their limited salt tolerance (31)] would suggest that conventional breeding practices are not enough. Plant breeding from the existing gene pools could help to achieve small increases in salinity tolerance. Advances in molecular biology as well do not appear to have helped greatly in the development of salt-tolerant genotypes. There is good evidence that salt tolerance in truly tolerant species (halophytes) is a multifaceted trait. In halophytes, tolerance depends upon ion exclusion or compartmentation, synthesis of compatible solutes, regulation of transpiration, control of ion leakage through the apoplast, membrane characteristics, and the ability to tolerate low K:Na in the cytoplasm. It has been demonstrated that a range of traits is also important in crops. This suggests that the overall trait of salt tolerance is determined by a number of genes.

Twenty years ago, Epstein (32) argued for development of crops in which the consumable portion is botanically a fruit, such as grain, berries, or pomes, and that have a truly halophytic response to salinity. In these plants, Na^+ ions would accumulate mainly in leaves, and because water transport to fruits and seeds is primarily through the phloem pathway (i.e., the intercellular connections), much of the salt from these organs would be screened. Initially, this hypothesis looked beyond realization. But in the last few years, scientists have reported increasing success in using genetic engineering to unravel the secrets of the genes and proteins that govern salt tolerance. A genetically engineered tomato that thrives in brackish water was produced in mid-2001 by a team of plant biologists at UC Davis (33). In late 2001, Purdue University scientists discovered a protein that guides salt into plants (34). Biotechnology specialist Alan McHughen at UC Riverside plans to begin research into producing salt-tolerant avocado trees. We also have other transgenic plants, for example, canola. Up to 10% of the dry weight of the leaves in this plant is sodium chloride. Schroeder and his colleagues have also discovered a gene that helps protect plants from salt damage by keeping salt in the roots. These results clearly support Epstein's argument.

With some success coming our way, it must not be forgotten that much more basic research is yet to be conducted. To develop genetically modified plants that have the triple characteristics of high yield, good quality, and tolerance to salts will not be an easy task in the foreseeable future. Besides, the genetically modified plants must be stable, so their descendants can be reliably grown with the added traits, and the plants must be tested for safety. This might take years of testing. Moreover, recent criticism and experience with genetically modified plants might hamper the progress of work. For example, some environmental groups say that large agricultural companies are rushing too quickly to introduce genetically modified crops without regard for potential dangers. Genetically engineered corn to produce a natural insecticide has been opposed on the grounds that it could kill insects that are not pests, such as the monarch butterfly. Even more worrisome, early genetically modified plants weren't very useful to consumers. An earlier genetically engineered tomato called the Flav-Savr, also produced from research at UC Davis, flopped in the mid-1990s. To resolve such issues would require much more effort. Till then, an integrated approach based on conventional plant breeding, molecular biology, and genetic engineering should continue.

BIBLIOGRAPHY

1. Maas, E.V. (1990). Crop salt tolerance. In: *Agricultural Salinity Assessment and Management*. K.K. Tanji (Ed.). American Society of Civil Engineers, New York, pp. 262–304.
2. Van Genuchten, M.Th. (1983). *Analyzing Crop Salt Tolerance Data: Model Description and Users Manual*. Research Report No. 120. US Salinity Laboratory, Riverside, CA, p. 50.
3. Maas, E.V. and Hoffman, G.J. (1977). Crop salt tolerance: Current assessment. *J. Irrig. Drainage Div.* **103(IR 2)**: 115–134.
4. Rhoades, J.D. (1990). Soil salinity-causes and control. In: *Techniques for Desert Reclamation*. John Wiley & Sons, New York, pp. 109–134.
5. Shedidan, D. (1981). Desertification of the United States. *Council on Environment Quality*. USA.
6. U.S. Salinity Laboratory Staff. (1954). Diagnosis and Improvement of Saline and Alkali Soils. *USDA Handbook 60*. US Salinity Laboratory, Riverside, CA.
7. Maas, E.V. and Grieve, C.M. (1987). Sodium-induced calcium deficiency in salt-stressed corn. *Plant Cell Environ.* **10**: 559–564.
8. Aronson, J.A. (1989). *HALOPH: A Data Base of Salt tolerant Plants of the World. Tucson, Arizona*. Office of Arid Lands Studies, the University of Arizona, Tucson, AZ, p. 77.
9. Glenn, E., Brown, J.J., and Blumwald, E. (1999). Salt-tolerant mechanisms and crop potential of halophytes. *Crit. Rev. Plant Sci.* **18**: 227–255.
10. Stalker, H.T. (1980). Utilization of wild species for crop improvement. *Adv. Agron.* **33**: 111–147.
11. Lauchi, A. and Epstein, E. (1984). Mechanism of salt tolerance in plants. *Calif. Agric.* **38**: 18.
12. Gupta, S.K. and Dubey, S.K. (2001). Pitcher irrigation for water conservation and use of saline water in vegetable production. *Indian Farming*. **51**: 33–34,36.
13. Minhas, P.S. and Gupta, R.K. (1992). *Quality of Irrigation Water—Assessment and Management*. Indian Council of Agricultural Research, New Delhi, India.
14. Ahi, S.M. and Powers, W.L. (1938). Salt tolerance of plants at various temperatures. *Plant Physio.* **13**: 767–789.
15. Magistad, O. et al. (1943). Effect of salt concentration, kind of salt and climate on plant growth in sand culture. *Plant Physio.* **18**: 151–166.
16. Hoffman, G.J. and Rawlins, S.L. (1971). Growth and water potential of root crops as influenced by salinity and relative humidity. *Agron. J.* **63**: 877–880.

17. Hoffman, G.J., Maas, E.V., and Rawlins, S.L. (1975). Salinity ozone interactive effects on alfalfa yield and water relations. *J. Environ. Quality* **4**: 326–331.
18. Rana, R.S. (1986). Breeding crop varieties for salt affected soils. In: *Approaches for Incorporating Drought and Salinity Resistance in Crop Plants*. V.L. Chopra and R.S. Paroda (Eds.). Oxford and IBH, New Delhi, pp. 24–55.
19. Singh, K.N. (1990). *Wheat varieties in alkali soils. Lecture in Summer Institute on Innovative Methods and Emerging Trends in Plant Breeding Research*. CCS Haryana Agricultural University.
20. Aswathappa, N., Morcar N.E., and Thomson, L.A.J. (1986). Salt tolerance of Australian tropical and sub-tropical acacias. *Proc. Int. Workshop on Aust. Acacias in Developing Countries*. J.W. Tumbull (Ed.). Gympie, Australia, pp. 70–73.
21. Maas, E.V. et al. (1983). Salt sensitivity of corn at various growth stages. *Irrig. Sci*. **4**: 45–57.
22. Maas, E.V., Poss J.A., and Hoffman, G.J. (1986). Salinity sensitivity of sorghum at three growth stages. *Irrig. Sci*. **7**: 1–11.
23. Maas, E.V. and Poss, J.A. (1989). Sensitivity of cowpea to salt stress at three growth stages. *Irrig. Sci*. **10**: 313–320.
24. Minhas, P.S., Sharma O.P., and Patil, S.G. (1998). *25 Years of Research on Management of Salt-Affected Soils and Use of Saline Water in Agriculture*. Central Soil Salinity Research Institute, Karnal, p. 220.
25. Gupta, S.K. and Sharma, S.K. (1990). Response of crops to high exchangeable sodium percentage. *Irrig. Sci*. **11**: 173–179.
26. Oosterbaan, R.J. (1981). *Crop yield, soil salinity and water table depths in Pakistan*. Annual Report, International Institute for Land Reclamation and Improvement, Wageningen, the Netherlands.
27. Royo, A., Susin, V., and Aragüés, R. (1991). Salt tolerance of barley: Analysis of response functions at emergence and at adult stages. In: *Improvement and Management of Winter Cereals under Temperature, Drought and Salinity Stresses*. A. Acevedo et al. (Eds.). Proceedings of the International Symposium; October 1987; Cordoba, Spain, pp. 373–394.
28. Maas, E.V. (1984). Crop tolerance. *Calif. Agric*. **103**: 20–21.
29. Porwal, A.K., Misra, K.L., and Gupta, S.K. (1990). Assessment of crop tolerance to saline irrigation waters. *Bhu. Jal. News*. **5**(4): 35–40.
30. Shannon, M.C. (1990). Genetic approaches for developing economic salt tolerance crops. In: *Agricultural Salinity Assessment and Management*. K.K. Tanji (Ed.). American Society of Civil Engineers, New York, pp. 161–185.
31. Flowers, T.J., Koyama, M., and Yeo, A.R. (1998). Breeding for salt tolerance in crop plants in the age of molecular biology. In: *Salinity Management in Agriculture*. S.K. Gupta et al. (Eds.). Central Soil Salinity Research Institute, Karnal, India, pp. 303–313.
32. Epstein, E. (1983). Crops tolerant to salinity and other mineral stresses. In: *Better Crops for Food*. J. Nugent and M. O'Connor (Eds.). Ciba Foundation Symposium, London, pp. 61–82.
33. Zhang, H.X. and Blumwald, E. (2001). Transgenic salt tolerant tomato plants accumulate salt in the foliage but not in the fruits. *Nat. Biotechnol.* **19**: 765–768.
34. Shi, H.H., Ishitani, M., Kim, C., and Zhu, J.K. (2000). The *Arabidopsis thaliana* salt tolerance gene *SOS1* encodes a putative Na^+/H^+ antiporter. *Proc. Nat. Aca. Sci.* **97**: 6896–6901.

APPENDIX A

Relative Tolerance of Crop/Plants to Soil Salinity (EC_e)[a]

Tolerant	Moderately Tolerant	Sensitive
	Field Crops	
Barley	Rye	Black gram
Sugar beet	Wheat	Bengal gram
Rape	Oat	Green gram
Cotton	Rice	Lentil
	Sorghum	
	Corn	
	Sunflower	
	Castor	
	Safflower	
	Soybean	
	Pearl millet	
	Linseed	
	Cluster bean	
	Pigeon pea	
	Cowpea	
	Sesame	
	Groundnut	
	Vegetable Crops	
Garden beets	Bitter gourd	Radish
Asparagus	Bottle gourd	Carrot
Spinach	Brinjal	Celery
Amaranths	Tomato	Green beans
	Cabbage	Coriander
	Pea	Cumin
	Lady's finger	Mint
	Onion	
	Potato	
	Carrot	
	Turnip	
	Sweet potato	
	Dolichos	
	Sponge gourd	
	Watermelon	
	Muskmelon	
	Chilies	
	Fenugreek	
	Garlic	
	Forage Crops	
Salt grass	Perennial rye grass	Meadow foxtail
Bermuda grass	Dallis grass	Alsike clover
Rhodes grass	Sudan grass	Red clover
Bridsfoot	Alfalfa	Burnet
Barley (hay)	Rye (hay)	
	Oats (hay)	
	Wheat (hay)	
	Orchard grass	
	Blue grass	
	Fruit Crops	
Date palm	Pomegranate	Pear
	Guava	Apple
	Fig	Orange
	Olive	Plum
	Grape	Peach
	Ber	Mango
	Kagzi Lime	Avocado

---

[a] The values in the same column under a given category and for the type of crop/plant are in descending order of tolerance. The values are qualitative and are to be used as a first approximation.

GROUNDWATER ASSESSMENT USING SOIL SAMPLING TECHNIQUES

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Groundwater assessment for environmental projects involves significant soil sampling to evaluate hydrologic and lithologic conditions. Soil samples in environmental investigations are typically collected at a minimum of every one to two meters and at changes of lithology or obvious signs of contamination, based on soil staining or discoloration of the soil or in the case of volatile organic compounds, organic vapor readings. Continuous coring is recommended for a minimum of at least one soil boring per site to characterize the subsurface adequately.

VADOSE ZONE

Soil samples collected in the vadose zone above the water table are in the unsaturated zone. Groundwater is moving generally vertically, and some horizontal movement is caused by rootlets, faults, impermeable layers, or other conduits. Soil samples collected within the capillary fringe or below the groundwater table are in the saturated zone.

The capillary fringe moves up and down with seasonal variations of the groundwater table. Petroleum hydrocarbons move vertically from the source, such as an underground storage tank, through the vadose zone. As the hydrocarbons move vertically through the lithologic section, in general, the finer grained the soil, the greater the lateral spreading of the contaminant (Fig. 1).

K_p = Conc soil/conc gas Soil zone	Unsaturated zone
K_p = Conc soil/conc gas Intermediate zone	
K_h = Conc gas/conc water Capillary fringe Water table	Saturated zone
K_d = Conc soil/conc water Groundwater	

Figure 1. Equilibrium expressions for partioning volatile organic compounds among aqueous, vapor, and sorbed phases.

CAPILLARY FRINGE ZONE

The capillary fringe, within the saturated zone, is the zone immediately above the water table, where water is drawn upward by capillary attraction. Due to irregularities in pore sizes among the grains of sediment, the capillary water does not rise to an even height above the water table but rather forms an irregular capillary fringe. The capillary fringe is higher in fine-grained soils such as a fine silt than in coarse grained soil such as a gravel due to greater tensions created by the smaller pore openings. The capillary fringe can be the zone of direct evaporation of groundwater if the water table is close enough to the surface (1).

The vertical movement during a complete hydrologic cycle can be a few centimeters, in other cases, it can be tens of meters, depending on the recharge of the water, the proximity of pumping wells, and other factors. This movement greatly affects the zone above the water table, which includes the capillary fringe zone, especially when it is impacted by light nonaqueous phase liquids (LNAPL), also called floating product or free-product.

When the free-product of a hydrocarbon fuel, such as gasoline, moves up and down through the hydrologic cycle, a smear zone develops in the zone above the water table. Some of this highly impacted soil zone (HISZ) is above the current capillary fringe zone and may exist below the top of the current groundwater table. Therefore, a HISZ equals the cumulative thicknesses of the migrating capillary fringe zones and frequently exists in areas where free-product or high concentrations of dissolved hydrocarbons have existed for a long period (Fig. 2).

For hydrocarbon fuels, which are lighter than water, the soil sample just above the water table is typically collected and analyzed. This sample may contain the highest concentration of contaminants within a soil boring. Sometimes, buried or submerged contaminants are trapped beneath the water table when water levels rise, leaving a former free-product zone below the water table (2). In these cases, soil samples collected within the groundwater zone would be appropriate (Fig. 3).

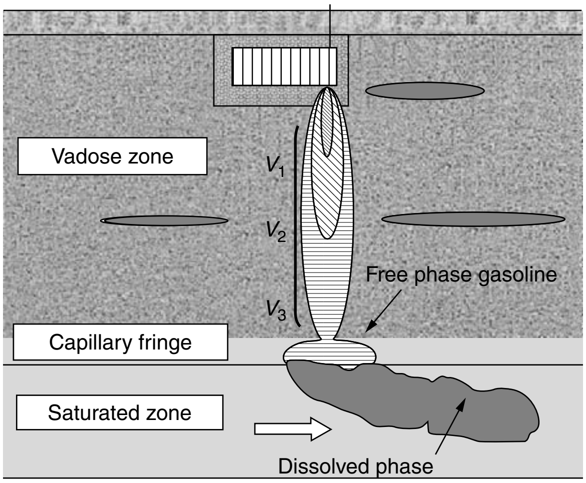

Figure 2. Gasoline movement from source (underground storage tank) to saturated zone (2).

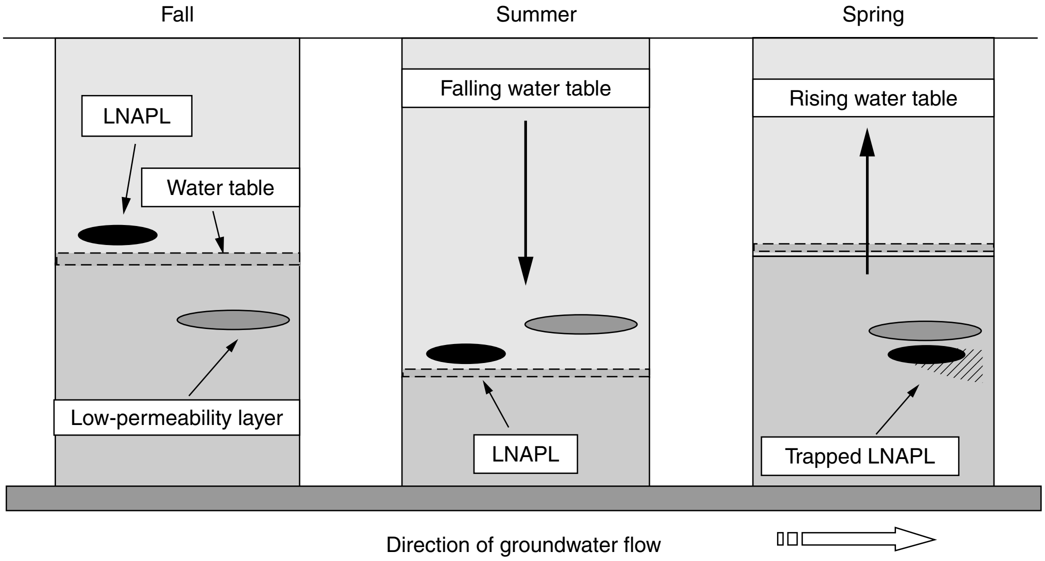

Figure 3. Trapping gasoline contaminants below the water table by fluctuating groundwater elevations (2).

For dense nonaqueous phase liquids (DNAPLs) such as chlorinated solvents, which have a specific gravity greater than water, soil samples are collected throughout the aquifer for hydrogeologic characterization. The highest concentrations of these dense liquids may be at the base of the aquifer, above the top of the next underlying aquitard.

SATURATED ZONE

Soil samples collected in the saturated or phreatic zone below the water table are used to evaluate the aquifer's characteristics. Soil samples collected below the groundwater table are generally not analyzed at a chemical laboratory, in favor of the more representative groundwater sample. Nonetheless, these saturated soil samples can provide a wealth of information about aquifer characteristics, total organic carbon, solid oxygen demand, porosity and permeability, as well as free-product distribution. Continuous soil coring of aquifer sediments allows complete lithologic analysis for potential groundwater remedial designs. These soil samples help scientists and engineers understand groundwater dynamics and the vertical and lateral movement of contaminants in the subsurface. Although groundwater sampling is imperative for groundwater assessments, data from soil samples used in conjunction with groundwater data provide a more complete understanding of subsurface conditions.

More detailed soil sampling information can be obtained from Testa (3), Jacobs (4), American Society for Testing Materials (5), California Regional Water Quality Control Board (6), and Morrison et al. (2). There are numerous minor differences in soil sampling and record keeping techniques and methods among field samplers, regional variations, and specific regulatory requirements that vary from state to state. A generalized summary of selected environmental soil sampling techniques follows.

SOIL SAMPLING METHODS

Undisturbed or minimally disturbed samples are obtained by driving a sampler into the soil, either manually, or using a drilling rig. The soil sampler is retrieved and opened, and the sample liner containing the soil sample is extracted. Samples are collected during subsurface investigations in vertical borings using rigs. Split-spoon samplers and direct push core samplers are used frequently for shallow soil sampling. The inner sample liners are typically composed of polyethylene, brass, or stainless steel. Samples may also be collected from tank pits and soil piles by driving a sampling tube into the soil by hand or using a drive-hammer.

Some analyses or guidelines allow collecting disturbed samples using a trowel, and then placing them in glass jars. Photoionization detectors (PID) are used in the field to measure organic vapors in soil samples. Soil samples should be collected in accordance with regulatory guidelines, which may vary among states and locales. The sample liners, containing soil, are typically sealed with Teflon tape and plastic caps at both ends. The samples are then labeled and stored in a refrigerated environment.

Small variations in soil lithology within several inches can lead to large differences in the results of chemical analysis. Duplicate soil samples, if taken, are collected as near as possible to the actual initial soil sample. Typically, these duplicate samples are collected in the same lithologic zone as the original sample. Duplicate samples should be a minimum of 5% of the samples. To maintain sample integrity, soil samples are delivered to certified laboratories using chain-of-custody procedures.

Lithologic Description

Soils and unconsolidated deposits for environmental projects are commonly described according to American Society for Testing Materials (ASTM) Method D 2488-84 (5) and the United Soil Classification System (USCS)

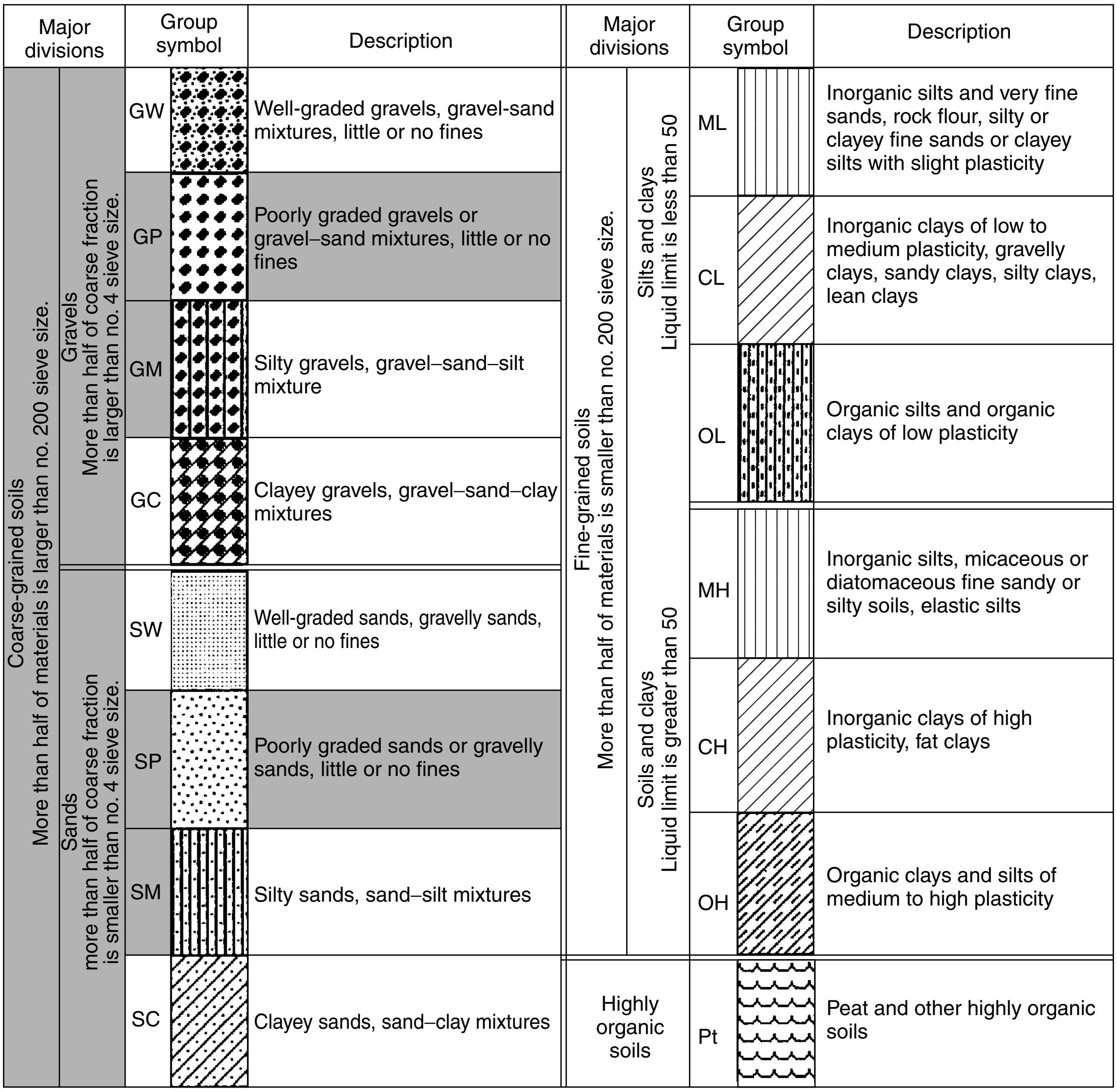

Major divisions		Group symbol	Description
Coarse-grained soils More than half of materials is larger than no. 200 sieve size.	Gravels More than half of coarse fraction is larger than no. 4 sieve size.	GW	Well-graded gravels, gravel-sand mixtures, little or no fines
		GP	Poorly graded gravels or gravel–sand mixtures, little or no fines
		GM	Silty gravels, gravel–sand–silt mixture
		GC	Clayey gravels, gravel–sand–clay mixtures
	Sands more than half of coarse fraction is smaller than no. 4 sieve size.	SW	Well-graded sands, gravelly sands, little or no fines
		SP	Poorly graded sands or gravelly sands, little or no fines
		SM	Silty sands, sand–silt mixtures
		SC	Clayey sands, sand–clay mixtures
Fine-grained soils More than half of materials is smaller than no. 200 sieve size.	Silts and clays Liquid limit is less than 50	ML	Inorganic silts and very fine sands, rock flour, silty or clayey fine sands or clayey silts with slight plasticity
		CL	Inorganic clays of low to medium plasticity, gravelly clays, sandy clays, silty clays, lean clays
		OL	Organic silts and organic clays of low plasticity
	Soils and clays Liquid limit is greater than 50	MH	Inorganic silts, micaceous or diatomaceous fine sandy or silty soils, elastic silts
		CH	Inorganic clays of high plasticity, fat clays
		OH	Organic clays and silts of medium to high plasticity
Highly organic soils		Pt	Peat and other highly organic soils

Figure 4. Unified soil classification system (USCS).

for physical description and identification of soils. The USCS was first used in the geotechnical engineering field and is now commonly used in the environmental field. Unfortunately, the USCS is not very useful for describing lithologic characteristics, geologic properties, or the origin or depositional environment of the soil or sediments (Fig. 4).

Other soil identification systems include the Burmister Soil Identification System (BS) for unconsolidated deposits, which is commonly used with the USCS. The other soil description system is the Comprehensive Soil Classification System (CSCS) developed by the U.S. Department of Agriculture. The CSCS describes soils as to agricultural productivity potential and best agricultural land use.

The ASTM Soil Classification Flow Chart and the USCS are generally accepted soil description methods used in the engineering and environmental fields. Descriptions of moisture, density, strength, plasticity, etc. are made using ASTM guidelines. Stratigraphic, genetic, and other data and interpretations are usually also recorded in the boring log. For consolidated deposits, including igneous, metamorphic, and sedimentary rocks, the USCS is combined with other geologic characteristics such as weathering, sorting, sphericity, and separation.

Color is commonly described by comparing the soil sample with a Munsell Rock Color Chart (7). Each layer of soil is identified by the geologist or engineer using the following items: color, soil type (USCS), classification symbol, color, consistency or relative density, moisture, structure (if any), and modifying information such as grain sizes, particle shape, cementation, plasticity, stratification, and other notes.

LITHOLOGIC INTERPRETATION

Once the lithology has been determined, geologists and engineers compare the data to develop a geologic model for the depositional environment. Knowing the type and origin of the soil, sediments and rocks provides a better understanding of groundwater and contaminant movement and allows predictive modeling of an aquifer. Once the geochemical data have been evaluated, a site conceptual model can be developed, and a corrective action plan and remedial design can begin if the site requires environmental cleanup.

BIBLIOGRAPHY

1. Fetter, C.W. (1993). *Contaminant Hydrogeology*. Macmillan, New York, p. 458.
2. Morrison, R., Beers, R., and Hartman, B. (1998). *Petroleum Hydrocarbon Contamination: Legal and Technical Considerations*. Argent Communications Group, Foresthill, CA.
3. Testa, S.M. (1994). *Geologic Aspects of Hazardous Waste Management*. CRC Press, Boca Raton, FL, pp. 145–187.
4. Jacobs, J. (2000). Monitoring well construction and sampling techniques. In: *Standard Handbook of Environmental Science, Health and Technology*. J. Lehr (Ed.). McGraw-Hill, New York, pp. 11.46–11.68.
5. American Society for Testing Materials. (1984). *Standard Practice for Description and Identification of Soils (Visual-Manual Procedure)*. Method D 2488-84. ASTM Philadelphia, PA.
6. California Regional Water Quality Control Board. (1989). *Leaking Underground Fuel Tank Field Manual: Guidelines for Site Assessment, Cleanup, and Underground Storage Tank Closure*. Sacramento, CA.
7. Munsell Color. (1988). *Munsell Soil Color Charts*. Munsell Color. Baltimore, MD.

SKIMMED GROUNDWATER

Muhammad Nadeem Asghar
International Water
Management Institute (IWMI)
Lahore, Pakistan

In several unconfined aquifers under irrigated agricultural areas, the groundwater system consists of a saturated porous medium containing miscible fluids (fresh groundwater and saline groundwater). The native groundwater is deep and saline because of the marine origin of the hydrogeologic formation. However, fresh groundwater lenses from deep percolation of the extensive water conveyance and distribution system, as well as from irrigation and rainfall, overlie this native saline groundwater.

According to Sufi et al. (1), over 20 BCM of fresh groundwater is being recharged annually in the Indus Basin of Pakistan.

The thickness of these fresh groundwater lenses varies from a few meters to more than 100 meters. For instance, in the Indus Basin of Pakistan, this is around 30 m thick in the lower or central parts of the *doabs* (area between two rivers), 60 m or more along the margins of the *doabs*, and approximately 150 m near the rivers and canals (2). When groundwater is extracted from these fresh groundwater lenses, it is called skimmed groundwater. Different types of irrigation wells are available for extracting this skimmed groundwater (3).

Under natural conditions, the saline groundwater tends to remain separate from the overlying fresh groundwater. However, when waterbodies of variable solute concentrations are in contact, molecular diffusion causes mixing across the line of contact. Due to groundwater and saline water movement, a zone of dispersion, which is also known as the interface, forms between these two fluids. This interface, which varies in thickness, is not static but responsive to recharge and discharge mechanisms. Thus, the processes of diffusion and dispersion result in a transition zone where the salinity gradually changes from completely fresh to fully saline.

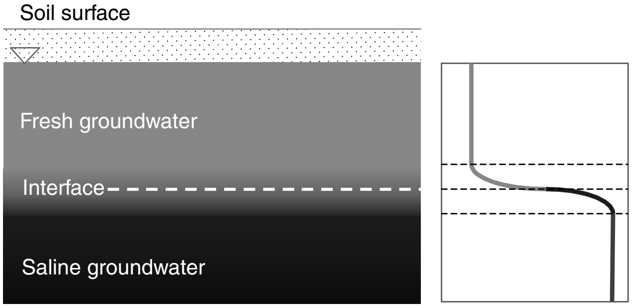

Therefore, the use of irrigation wells for skimming fresh groundwater may cause mixing of the ‘interface’ either within the well or in the aquifer. As a result, the irrigation well starts supplying skimmed groundwater of marginal quality. Therefore, the decision regarding the use of skimmed groundwater must consider the quality of groundwater with depth, the local hydrogeologic conditions of the aquifer, the socioeconomics of skimming well installation and operation, the discharge rate, the duration of pumping, and above all, the tolerable limits of skimmed groundwater quality for its intended uses (4).

Farmers are using this skimmed groundwater to fill the gap between the supply and demand of surface water in the conjunctive water use environment of the Indus Basin of Pakistan (5). This irrigation water resource has emerged as a formidable poverty reduction tool. Potentially, skimmed groundwater is a huge resource of irrigation water. In the Indus Basin of Pakistan alone, it has been estimated that nearly 200 billion cubic meter (BCM) of fresh groundwater is lying above salty groundwater (6), and covers about 30% of the irrigated area of Pakistan.

BIBLIOGRAPHY

1. Sufi, A.B., Latif, M., and Skogerboe, G.V. (1998). Simulating skimming well techniques for sustainable exploitation of groundwater. *Irrig. Drainage Syst.* **12**: 203–226.
2. Asghar, M.N., Prathapar, S.A., and Shafique, M.S. (2002). Extracting relatively fresh groundwater from aquifers underlain by salty groundwater. *Agric. Water Manage.* **52**: 119–137.

3. Saeed, M.M., Asghar, M.N., and Bruen, M. (2003). Options for skimming fresh groundwater in the Indus Basin of Pakistan: a review. *J. Groundwater Hydrol.* **45**(3).
4. Asghar, M.N., Saeed, M.M., Ashraf, M., and Qureshi, A.S. (2002). Potential of Skimming Wells for Extracting Relatively Fresh Groundwater from Unconfined Aquifers Underlain by Saline Groundwater. *Proc. 2nd Int. Conf. New Trends Water Environ. Eng. Safety Life: Eco-compatible Solutions Aquatic Environ.*, June 24–28, Capri, Italy.
5. Wahaj, R. and Asghar, M.N. (2002). Farmers' Management Responses to the gap between Supply and Demand of Canal Water in a Large-Scale Irrigation System. *Proc. 18th Congr. Int. Comm. Irrig. Drainage, Food Prod. Conditions Water Scarcity, Increasing Population Environ. Pressure*, July 21–28, Montreal, Canada.
6. NESPAK. (1983). *Drainage and Irrigation in Shallow Sweet Water Areas Overlying Saline Water Using Radial Collector Wells*. Technical Appraisal, NESPAK Ltd, Lahore, Pakistan.

SOIL MOISTURE MEASUREMENT—NEUTRON

Mukand Singh Babel
Asian Institute of Technology
Pathumthani, Thailand

The water (liquid and/or vapor phase) contained in the soil above the water table is called soil moisture. Soil moisture is a function of precipitation, soil texture, porosity, the rate of soil moisture uptake by plants, and the rate of evapotranspiration. Measurements of soil moisture are needed for several purposes, including irrigation planning and scheduling, agronomic studies, flow and transport processes in the unsaturated zone, and geotechnical investigations.

Soil moisture can be measured by direct and indirect methods. Direct methods are those wherein water is removed from a soil sample by evaporation, leaching, or chemical reaction, and the amount removed is determined. Indirect methods involve measuring some property of the soil that is affected by soil moisture content or measuring a property of some object placed in the soil, which comes to water equilibrium with the soil. The indirect methods generally require calibration. They are suitable for quick and repetitive measurements and do not require destructive sampling as needed in direct gravimetric methods (1).

Although a variety of indirect techniques are available for measuring soil moisture content, they are neither equivalent nor applicable in all situations. Thus, choosing a suitable method and following a well-defined procedure are important for obtaining reproducible and reliable information. Various *in situ* recording sensors, water budget models, and experimental approaches are used for soil moisture measurement, including neutron probe, time domain reflectometry, electromagnetic induction, electron resistivity tomography, tensiometers, heat dissipation sensors, and thermocouple psychrometers. Modern methods of soil moisture measurement such as airborne radar and satellite imagery are still at the research and development stage.

The neutron probe also called the "neutron moisture meter" is a nondestructive, indirect method commonly used for repetitive field measurement of soil water content volumetrically.

PRINCIPLE

The neutron probe is based on the principle of neutron thermalization. Hydrogen nuclei are nearly the same size and mass as neutrons and therefore, can scatter and slow neutrons. High-energy (0.1–10 Mev.), fast (1600 km/s) neutrons emitted from a radioactive source, such as americium–beryllium or radium–beryllium are slowed and changed in direction by elastic collisions with atomic nuclei of hydrogen atoms in soil. This process is called "thermalization," and the low-energy (about 0.03 eV) neutrons, after collision, are called "thermal/thermalized neutrons." The density of the thermal neutrons (thermal cloud of neutrons) depends on the soil–water content; the vast majority of hydrogen in soil is associated with water. Higher water content in soil leads to increased thermalization and, thus, a denser thermal cloud. A "slow neutron" detector is installed adjacent to the source (emitter) to measure the cloud density. The measurement is usually in the form of a "count ratio"; a higher count ratio denotes higher water content and vice versa. Thus, the density of thermal neutrons can be calibrated against the water content volumetric. (2).

THE EQUIPMENT

Figure 1 shows a typical neutron moisture meter. The equipment consists of

1. a neutron moisture depth probe containing a radioactive source that emits high-energy neutrons, and a detector of thermalized neutrons.
2. protective shielding (for gamma rays) of lead and polyethylene or paraffin for neutron absorption. The depth probe is stored within the shield when not in use.
3. a meter assembly that includes a pulse counter to register the counts generated by the thermal neutron detector, a built-in computer, a display screen, and a keyboard.
4. a thin-walled aluminum or plastic access tube. The tube diameter is recommended by manufacturers for optimized equipment performance.
5. a radioisotope license, radioisotope leak test kit, and radiation monitoring equipment, as required by the license.
6. calibration curves and parametric data for calibrating the computer of the equipment, if needed.

CALIBRATION OF EQUIPMENT

The sources of hydrogen other than water (the primary source of hydrogen) in a given soil are assumed constant and are accounted for during calibration. Similarly,

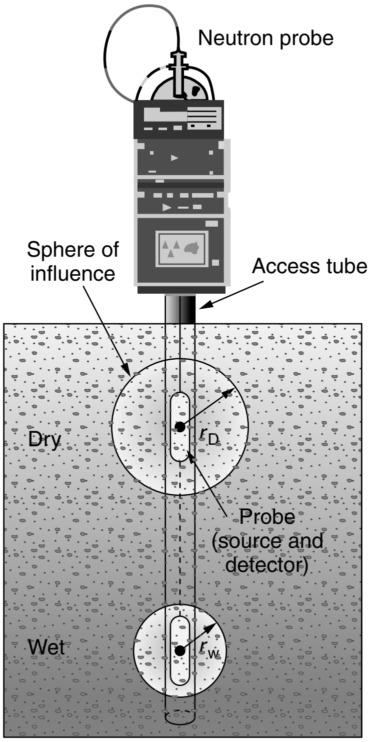

Figure 1. Neutron probe device. r_D and r_W are the radii of measurements in dry and wet soils (*Source*: Reference 3).

several nonhydrogen substances including C, Cd, B, Cl, and Li may be present in some soils and may also thermalize fast neutrons. However, their effect is compensated for through soil specific calibration. The neutron probe is calibrated by paired measurements of soil moisture content and neutron probe counts. Usually, the volumetric moisture content, θ_v is linearly related to the slow neutron count ratio (CR) as follows:

$$\theta_v = a + b(\mathrm{CR})$$

where CR is the ratio of slow neutron counts at a specific location in the soil and a standard count obtained with the probe in its shield. A calibration curve can be obtained with a sufficient number of measurements. This calibration curve is approximately the same for many soils. Figure 2 depicts a typical calibration curve.

The sphere of influence of neutrons around the radiation source varies with water content from about 0.15 m for wet soil (almost saturated) to more than 0.60 m for very dry soil, depending on how far fast neutrons must travel to collide with a requisite number of hydrogen nuclei. An approximate relationship for the radius of influence, r (in cm), as a function of soil water content is given by

$$r = 15(\theta_v)^{-1/3}$$

PROCEDURE

1. Use the soil auger to form a hole for installing the access tube. The access tube is left so as to protrude about 10 cm above the soil surface and is covered with an empty can or a stopper to keep water and debris from entering.
2. Place the probe unit over the access tube to make a measurement. Select an appropriate counting time, and make several standard counts while the probe is in the shield. This will correct for any electronic changes in the counting circuits.
3. Make one or more counts at each selected measurement depth.
4. Use the calibration curve to convert the count ratios to volumetric moisture content, or read the moisture content directly if the equipment has a built-in computer.

COMMENTS AND PRECAUTIONS

1. In regard to the radiation hazard of devices with radioactive sources, it is strongly recommended that the equipment be used according to the manufacturer's instructions, which must comply with safe procedures dictated by the radioactive source license.
2. Site-specific calibration may be required for a soil of unusual neutron-absorption properties or if the diameter or material of the access tube is different from the manufacturer's recommendations.
3. As the zone of measurement is spherical and at least greater than 0.15 m in radius, the probe depth increment in the access tube should not be less than

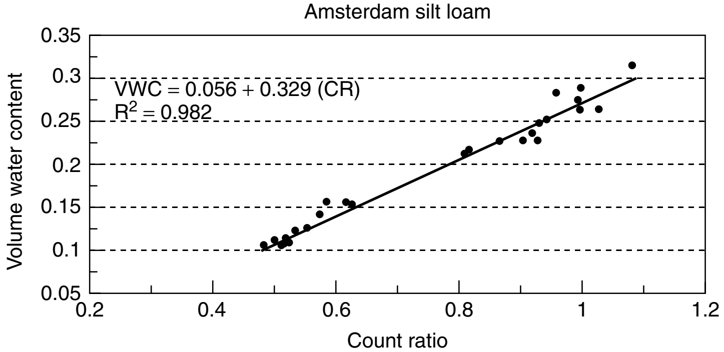

Figure 2. A neutron probe calibration relationship (*Source*: Reference 3).

0.15 m. The approach to the soil surface should be no closer than 0.15 m because a portion of the neutrons may escape from the soil.

ADVANTAGES AND DISADVANTAGES

The neutron probe moisture meter has several advantages:

1. It is a widely accepted method of soil moisture measurement with highly reproducible results.
2. After initial tube installation, no further monitoring is required.
3. Both horizontal and vertical installation are possible.
4. It can measure soil water content at multiple depths and locations.
5. It is possible to couple the equipment with other devices.
6. The equipment produces reliable results for a long period.

Following are the disadvantages of the equipment:

1. It contains a source of radioactivity and, therefore, cannot be automated. It may also pose a health risk if adequate precautions are not taken.
2. The equipment is expensive and requires calibration.
3. It has relatively limited spatial resolution of measurement.

BIBLIOGRAPHY

1. Gardner, W.H. (1986). Water content. In: *Methods of Soil Analysis*, *Part 1*. A. Klute (Ed.). *Agronomy Monograph* No. 9, American Society of Agronomy, Inc. and Soil Science Society of America, Inc., Madison, WI.
2. Carter, R.M. (1993). *Soil Sampling and Methods of Analysis*. Lewis, Boca Raton, FL.
3. Sumner, E.M. (2000). *Handbook of Soil Science.* CRC Press, Boca Raton, FL.

SOIL N MANAGEMENT IMPACT ON THE QUALITY OF SURFACE AND SUBSURFACE WATER

Dibyendu Sarkar
Vandana Vandanapu
Rupali Datta
University of Texas at San Antonio
San Antonio, Texas

Nitrogen (N) in the forms of nitrate or ammonium is a nutrient needed for plant growth. When the amount of soil-supplied N is considered deficient for satisfactory crop yields, N fertilizer is added to supplement the soil-supplied N. If N is supplied in excess of crop needs, a surplus is created that can cause significant water quality problems. Until recently, the main reason for investigations into the forms of N in soil systems below the rooting depth of plants was to evaluate the loss of nitrate to crops and to estimate the loss of production. In recent years, much attention has been focused on forms of N as potential pollutant sources of both groundwater and surface water. Nitrate concentrations above a certain level in drinking water may have critical health effects on animals and human infants. Alternately, N can also contribute to surface water contamination via surface water runoff and erosion. Excessive nitrate concentrations in surface water bodies, such as streams or lakes, can accelerate algae and plant growth, resulting in oxygen depletion. Nitrogen must be managed properly to protect water quality. Best management practices (BMPs) are methods that are the most effective, practical means of preventing or reducing nutrient pollution from nonpoint sources. Nitrogen BMPs primarily involve a change in fertilizer management practices, such as changing the timing, placement method, and/or the amount of fertilizer applied to reduce the possibilities of contaminating water resources. In addition, conservation tillage practices, cover crops, and crop rotation all help reduce surface runoff, erosion, and leaching of soil N. It is imperative to continue the development and implementation of new and effective management practices that increase N recovery and reduce potential N losses to the environment.

INTRODUCTION

Nitrogen is one of the most abundant elements on earth. It exhibits a number of oxidation states and combines with a number of other elements to form organic and inorganic compounds (1). Nitrogen accounts for 78% of the earth's atmosphere as elemental nitrogen (N_2) gas (2). However, the majority of the earth's N (98%) is in rock, sediment, and soils. The amount of N in rocks is about fifty times more than that in the atmosphere (3). For mineral soils in the United States, the approximate range of total N content in the surface six inches is 0.02% to 0.50% (4). Growing plants, animals, and microbial populations need a continual source of N. It is an essential component of the proteins that build cell material and plant tissue. In addition, it is necessary for the function of other biochemical agents, including chlorophyll (which makes photosynthesis possible), enzymes (which help organisms carry out biochemical processes and assimilate nutrients, and nucleic acids such as DNA and RNA which are involved in reproduction) (5). Of all the major plant nutrients, N is often the most important determinant of plant growth and crop yield. Without adequate N, plant growth is restricted. Characteristic symptoms of N deficiency in plants include yellowing and death of leaves and stunted growth. Just as too little N can cause problems, too much N can also cause problems. These problems can extend to plants, humans, animals, and the environment. The problems posed to the environment occur when excess N in soils is carried away with surface runoff to lakes, rivers, groundwater, and other aquatic ecosystems. Excess N in rivers, lakes, and groundwater can be toxic to humans and causes water quality problems in natural water systems (6). Nitrogen, particularly as nitrate, is the most common contaminant in aquifer systems (7). Hallberg (8) points to agriculture as the most substantial anthropogenic source of nitrate, and Keeney (9) suggests that this is caused by intensive and extensive land use for crop and animal production. The United States Environmental Protection Agency (USEPA) has identified eutrophication as the main problem in U.S. surface waters that have impaired water quality (10). Excessive amounts of nutrients, especially N and phosphorus (P), speed up eutrophication (11). Excess N in the estuaries of the oceans enhances the growth of aquatic organisms to the point that they affect water quality and lower dissolved oxygen levels (12–14). This affects the metabolism and growth of oxygen requiring species, causing hypoxia (12,13). Implementation of N-BMPs (best management practices) allows efficient, profitable production and minimizes N movement to surface water and groundwater. This article addresses the concept of N movement in soils and the impacts of N management on surface and subsurface water quality.

SOURCES OF NITROGEN IN THE ENVIRONMENT

Atmosphere

The earth's atmosphere is made up of 78% N as a colorless, odorless, nontoxic gas. The same N_2 gas found in the atmosphere can be found in spaces between soil particles. However, plants cannot use this form of N (15). A group of microbes (*Rhizobium* bacteria) that naturally exists in the soil converts N to ammonium. In *biological fixation*, nodule-forming *Rhizobium* bacteria inhabit the roots of leguminous plants and through a symbiotic relationship, convert atmospheric N_2 to a form the plant can use (16).

Atmospheric fixation occurs during lightning storms. The enormous energy of lightning breaks down N_2 molecules and enables their atoms to combine with oxygen in the air to form nitrogen oxides. These dissolve in rain forming nitrates that are carried to the earth. Atmospheric N fixation probably contributes some 5–8% of the total N fixed (5).

The large amount of fossil fuels used to produce electricity and in manufacturing industries oxidizes N because the intense heat given off by their combustion causes the molecules to react with each other, thereby adding N to the air as nitrates. These nitrates fall to the earth during precipitation, adding more N to the soil (17).

Animal Manure

Animal manures are important sources of N in the terrestrial environment. Manure usually contains both organic and inorganic N forms. The amount of N in manure depends on the animal's species, age, diet, and bedding materials. Ammonium N is the primary inorganic form in manure and is readily available to crops. Organic N becomes available to plants as manure decomposes; 20 to 50% of organic N is available to the first crop after application. Much of the remaining organic N becomes available in subsequent years as mineralization continues (18).

Legumes

Leguminous crops, such as alfalfa, beans, and peas, generally do not require N fertilization because the *Rhizobium* bacteria living on their roots take N from the air and convert it to forms that plants can use. Legumes provide some of the N for their own growth and often build up a supply that can be used by crops that follow (19).

Commercial Fertilizers

Commercial fertilizers are derived from the atmospheric N pool. Industrial processes used to manufacture commercial fertilizers usually combine atmospheric N with hydrogen and/or oxygen and/or carbon and then convert the products to various N fertilizers such as ammonium nitrate, anhydrous ammonia, and urea (4). These fertilizers are generally applied directly to soils.

Crop Residues

Crop residues returned to the soil are important sources of N for succeeding crops if they are properly managed (20). Nitrogen exists in crop residues in complex organic forms; the residues are decayed by soil microbes, thereby converting N and other nutrients to inorganic forms, which then become available for plant uptake (16).

Organic Soil Matter

Organic soil matter is a major source of N in terrestrial ecosystems. Organic matter is composed primarily of a relatively stable material called humus. Even though the soil has a large pool of N stored in soil humus, this N is mostly insoluble and is not available for plant uptake. Organic N must be mineralized to inorganic forms by soil microbes before plant uptake can occur.

FORMS OF NITROGEN IN SOILS

Nitrogen in soils is found in organic and inorganic forms, both of which are important sources of N for plant and microbial uptake. Organic N occurs in crop residues and organic matter as protein compounds, amino acids, amino sugars, and complex, largely unidentified substances (4). More than 90% of all soil N exists as organic N, and for the most part, is not available for plant uptake. Natural biological processes carried out by microorganisms in soil convert organic N to inorganic forms, which plants can use (15). Inorganic N exists in soil primarily in three forms: elemental N, ammonium N (NH_4^+) and nitrate N (NO_3^-). Ammonium N and nitrate N are two forms of soil N that can be readily used by plants, whereas elemental N cannot be used directly by plants. However, some N-fixing bacteria, particularly *Rhizobium* that coexists with leguminous plants, use N in the air (elemental N) and converts it in to a form of N that can be used by plants.

NITROGEN LOSSES FROM SOILS

Nitrogen is lost from soil in several ways: leaching, denitrification, volatilization, soil erosion and runoff, and crop removal.

Leaching

Leaching is the loss of soluble N. The NO_3^- ion is negatively charged and is not attracted to soil's clay and humus, which are also negatively charged. So it stays in the water between soil particles. As water moves through soil, the nitrate in the soil solution moves along with the water (21). This loss most frequently occurs in areas of heavy rainfall, under excessive irrigation, and in coarsely textured soils.

Denitrification

Denitrification is the conversion of NO_3^- molecules to gaseous forms that are lost to the environment. When there is a deficit of oxygen in the soil (i.e., under anaerobic conditions), certain bacteria meet their energy needs by reducing NO_3^- to dinitrogen gas or to nitrogen oxide (N_2O). These gases are released to the atmosphere, resulting in a loss of N from soil (22). Denitrification can be significant when soils are saturated with water.

Volatilization

Volatilization is the loss of N from the soil as ammonia (NH_3) gas. NH_3 volatilization commonly takes place when N is in an organic form called urea. Urea may originate from animal manure, fertilizers and, to a lesser degree, from the decay of plant materials. Urea is converted to ammonium carbonate by urease, an enzyme found in soils. Ammonium carbonate easily dissociates to ammonium and carbonate ions. The carbonate ions result in a relatively short-lived increase in soil pH. Ammonium ions convert at high pH to NH_3 gas, which may escape to the air (23). This results in a net loss of N from the soil system.

Surface Runoff and Soil Erosion:

Surface runoff and soil erosion can also cause significant losses of soil N. Runoff and erosion losses include both inorganic and organic N. Erosion leads to the removal of organic N in the upper soil profile (24) and NH_4^+ N, which is adsorbed to soil particles (25). Nitrate N is water soluble and is generally not adsorbed to soil particles; hence, it quickly moves off the soil surface as runoff with increased rainfall.

Crop Removal/Harvest

Substantial amounts of N are lost from soil through crop removal (16). When crops are harvested, the N source is removed. Crop removal accounts for a major portion of the N that leaves the soil system.

HUMAN HEALTH AND WATER QUALITY CONCERNS RELATED TO LOADINGS OF N INTO SURFACE AND GROUNDWATER

Nitrogen is the one element of most concern as it relates to water quality. Although N can be found in many forms in the soil, the nitrate form primarily affects water quality (15). The contamination of surface water or groundwater by NO_3^- N presents a potential health threat to human population. High NO_3^- N levels in water are a concern because the ingestion of such water can reduce the oxygen carrying capacity of the blood. Nitrate is reduced to NO_2^-, which oxidizes the Fe of the hemoglobin in the blood, and methemoglobin, which cannot carry oxygen, is formed (9). Infants have greater potential than adults to suffer these ill effects. In the United States, the maximum NO_3^- N level for potable water has been set at 10 mg/L (26). Secondary and related deleterious effects of exposure to NO_3^- include increased respiratory infections, inhibition of iodine uptake by the thyroid, and possible reproductive problems (27).

Another area of concern is the toxic effect of ammonia (NH_3) on freshwater aquatic life (28). It has been known since the early 1900s that NH_3 is toxic to fish and that this effect varies with water pH and temperature. A concentration of 0.02 mg/L as unionized NH_3 is the current limit in freshwater in the United States (29).

The most significant water quality concern with respect to N is the overenrichment or eutrophication of surface waters. Nitrates stimulate the growth of algae and other plankton, which provide food for higher organisms; however, excess NO_3^- N can cause over production of plankton. As these planktons die and decompose, they use up the oxygen dissolved in water, which causes an oxygen shortage for other aquatic organisms. Elevated N concentrations have altered natural aquatic floral and faunal population dynamics, exacerbated occurrences of hypoxia and anoxia, and sped eutrophication in the Gulf of Mexico (12,30).

NITROGEN BEST MANAGEMENT PRACTICES: IMPACT ON SURFACE AND GROUNDWATER QUALITY

Nitrogen from fertilizers as well as from other sources (e.g., biosolids, animal manures) has a direct impact on water quality. The primary goal of N best management practices (N-BMPs) is to manage N inputs to crop production to prevent degradation of water resources while maintaining profitability. Typically, BMPs are farming methods that assure optimum plant growth and minimize adverse environmental effects (31). By putting N-BMPs into practice, N losses from agricultural soils can be controlled. There are many BMPs currently in use; some of the more popular soil N management practices and their potential effects on surface and subsurface water quality have been categorized by Owens (32) as follows.

Crop Management

Cover Crops. The use of cover crops in cropping systems provides several benefits. However, only in recent years has much of the attention been focused on their potential to reduce NO_3^- leaching losses (32). The key to reducing NO_3^- N losses is to keep soil nitrate concentrations low during periods of high leaching potential. Nitrate can leach to groundwater when the soil contains significant NO_3^- N and when water percolates below the root zone. These conditions often exist during the winter months, when percolation of winter rainfall can be high and uptake of water and nitrate by plants is minimal (33).

Cover crops influence nitrate leaching in two ways: water budget effects and N uptake effects. Cover crops can influence the water budget by reducing the water content of the soil during their active growth stage (33,34). In humid climates, this water usage can reduce nitrate leaching significantly. Cover crops function by accumulating the inorganic soil N between main crop seasons and by holding it in an organic form, thus preventing it from leaching. Nonleguminous cover crops can immobile as much as 70% of the available NO_3^- N in the upper soil profile (35). Generally, nonlegumes are about three times more efficient than legumes in reducing nitrate leaching (36). Meisinger et al. (33) reported that cover crops reduced both the mass of N leached and the NO_3^--N concentration of leachate by 20–80% compared to no cover crop control. They also determined that grasses and *Brassicas* were two to three times more effective than legumes in reducing NO_3^--N leaching. Because, grasses and grains become quickly established in the fall and develop an extensive root system, they are more efficient than legumes in capturing soil nitrate and preventing late fall and winter leaching to groundwater (35,37). In a western Oregon study of nitrate movement in the soil profile following harvest of a broccoli crop fertilized with 250 lb N/acre, Hemphill and Hart (38) demonstrated the effectiveness of a fall-planted rye cover crop in reducing the movement of soil nitrate to the soil profile. Areas with cover crop contained 0.18 mg/kg nitrate at 30–40 inches of soil depth compared to 5.38 mg/kg nitrate at the same depth where no cover crop was grown.

Controlling erosion is an important component of surface water quality protection. Cover crops protect against soil erosion (39,40). Cover crops can be effective in reducing runoff and erosion, especially during the winter when there is little crop residue (32). Farmland is most susceptible to erosion when there is no vegetative ground cover or plant residue on the soil surface. Cover crops provide vegetative cover when a crop is not present to cushion the force of falling raindrops, which otherwise would detach soil particles and make them more prone to erosion. The rate of runoff is also slowed, thus improving moisture infiltration into the soil (41). When legumes are used as the winter cover crop, fixed N in the legumes can be supplied to the subsequent crop so that less fertilizer N has to be applied. This reduces the amount of inorganic N, which is readily available for removal in surface runoff (32).

Crop Rotations. Understanding the relationship between N and crop rotation is very important in N management decisions (42). Although there are several benefits to crop rotations (e.g., lower fertilizer N input because of residual N from legumes, erosion reduction, minimization of crop diseases, reduction in insect problems), relatively little attention has been given to impacts of NO_3^- N leaching by most crop sequences (32). Proper crop sequencing can sometimes be used to improve fertilizer-N efficiency and reduce nitrate leaching. Johnson et al. (43) showed that unfertilized soybeans scavenge large quantities of residual N fertilizer from a previous corn crop. Monocultures of crops, such as corn, winter wheat, and sorghum, require high N input. It is possible to reduce N use in a farming system by rotating these crops with small grains, which require less N, or with legumes, which may require no added N. This type of crop diversification should result in less N in the soil profile and, in general, less leaching.

The combination of crops grown in a rotation also affects the potential for groundwater contamination by N. For example, less N moves beyond the rooting zone with a corn–soybean rotation compared to continuous corn (44). Including perennial legumes or non leguminous crops in rotations also decreases NO_3^--N losses. Baker and Melvin (45) found much lower NO_3^--N concentrations beneath alfalfa than for corn or soybean.

The key to successful erosion control with crop rotations is soil cover. Fallow land has the highest erosion potential in any cropping system. Crop rotations that include small grains or other cover crops create a vegetative cover, which protects the soil surface from erosion (46). Grasses and legumes or small grains provide more protection against water erosion than row crops (47). Cropping systems with a higher frequency of sod reduce soil erosion. Growing cover crops with low residue crops and rotating high residue crops with low residue crops are also effective erosion control practices. Some crop rotations do not reduce erosion unless other practices, such as cover crops and residue management are followed (48). Crop rotations that use the land more intensively, such as corn, wheat, and soybeans grown in 2 years, produce larger amounts of biomass during the rotation and are more effective in reducing erosion than a continuous cropping sequence (49).

Tillage Practices. Soil erosion and runoff contribute to the degradation of surface water quality. Tillage management is an important factor to be considered when attempting to reduce soil erosion and subsequent sedimentation (50). Tillage practices that leave the soil surface unprotected greatly increase the potential for runoff, erosion, and nutrient losses from fields (44). Conventional tillage buries the protective crop residue cover and disturbs the soil structure; increased rainfall may then lead to amplified runoff and erosion (51). Conservation tillage practices, which reduce runoff and erosion, reduce N losses (52–54). Such practices leave greater amounts of crop residue and cause less soil disturbance. Both factors help reduce runoff and erosion, reduce transport of soluble and sediment-attached chemicals, and promote infiltration (55). Baker and Laflen (56) concluded that conservation tillage should reduce soil erosion by 75–90% (depending on the amount of surface residue) compared to conventional tillage. Conservation tillage also increases infiltration and reduces average annual runoff volumes by about 25% compared to conventional tillage (57). The majority of the N in surface runoff is associated with sediments, so conservation tillage usually results in a net decrease in N losses.

Several of the factors that reduce surface runoff promote an increase in subsurface flow (32). Conventional tillage disrupts surface soil structures, destroying macropores in the tilled layer. It also enhances the mineralization of organic N from crop residues and soils and increases the accumulation of residual NO_3^- N. In recent years, technological innovations have led to reduced levels of tillage in many areas. This trend seems to protect groundwater in some areas because of reduced N mineralization and increased N immobilization (44). The effect of tillage on macropores and the flow of water and chemicals through macropores are important (58–60). However, using conservation tillage, macropores remain intact from the soil surface to their full depth in the soil profile (61–63). Macropores potentially increase the amount of water entering the soil; this increase in infiltration leads to more leaching. Tillage has both advantages and disadvantages with respect to groundwater; information from studies on local soils should be used to determine if conservation tillage would have positive or negative impacts on subsurface water quality.

Fertilizer Management

Nitrogen Rate. The one most important practice affecting the N contribution to groundwater is the rate of N application (64). Excessive application of N fertilizer increases the likelihood that N is leached into surface water and groundwater (47). The amount of N from other sources should be given appropriate credit to avoid excess fertilizer N. The soil can supply N to plants as mineral N, usually NO_3^- N, and as N mineralized from organic matter. There can be substantial overwinter carryover of residual NO_3^- (65), which would allow reducing of subsequent N fertilizer applications.

Nitrogen rate recommendations also include credits for N fixed by legumes. Winter leguminous cover crops and legumes in rotation can provide a significant amount of N for a subsequent nonleguminous crop (66–68). Some legumes can provide as much as 120 kg N/ha, which shows strongly that credit needs to be given for legume N to reduce excessive N fertilizer applications (32). Similarly, N application rates can be significantly adjusted to account for manure application. By failing to account for these sources of N in determining the correct application rate, a surplus of N may be created; this surplus may leach to groundwater (69).

Similarly, N credits need to be given for manure applications. Recent reports of long-term manure applications indicate that high rates of application can pose a hazard for NO_3^--N contamination in groundwater (32).

Fertilizer Placement. Correct fertilizer placement is a BMP that can help protect both surface and groundwater quality from nutrient pollution. The two most often used fertilizer placement strategies are banding and broadcast placement (70).

The efficiency of N fertilizer is generally increased by band placement that provides many agronomic and environmental advantages over broadcast applications (71). Banding causes an increase in plant N use efficiency. High crop-use efficiency leaves less N in the soil to leach into groundwater (70). Banding of N below or with the seed is by far the best N-BMP to protect surface water quality because no N is left on the soil surface for erosional loss. A broadcast-incorporated application leaves a portion of the N fertilizer on or near the soil (70). A study in Montana found that spring wheat yields are approximately 6% higher when urea was banded 1 inch below the seed and between rows compared to broadcast (72).

Fertilizer Timing. The timing of N fertilizer application is an important factor that affects the efficiency of fertilizer N because the interval between application and crop uptake determines the length of exposure of fertilizer N to loss processes, such as leaching and denitrification. Timing N applications to reduce the chance of N losses through these processes can increase the efficiency of fertilizer N use (21).

Fertilizer applied in the fall causes groundwater degradation. Fertilizer-N applications should be managed very carefully on sandy or gravelly soils due to the high potential for leaching. Fall applications are not recommended on these types of soils (44). Fall applications of anhydrous NH_3 or urea on finer textured soils should be delayed until soil temperature reaches 45 °F or less when conversion to NO_3^- is slow, hence, the potential for NO_3^--N leaching losses is reduced. Split application of N fertilizers as opposed to a single application can reduce potential N losses by up to 30% and thereby reduce groundwater contamination (73).

Organic Wastes. Animal manure is a good source of N. It benefits plant growth, improves soil structure, and provides nutrients. However, manure can become an environmental problem when it pollutes groundwater and surface water as a result of improper handling, storage, and application on erosion-prone soils (74). This is primarily due to the particulate nature of the material compared to the more soluble forms of N in inorganic fertilizers (44).

The other major concern with manure is the movement of NO_3^- N to groundwater. Manure contains mostly organic and NH_4^+ N. These forms of N do not move through soils. However, soil microorganisms can convert these compounds to soluble NO_3^- N that may leach to groundwater. Nitrogen not mineralized in the first year after application of manure generally becomes part of the soil organic matter and releases N relatively slowly. Constant annual manure applications that supply N to meet the entire crop demand ultimately cause excessive fertilization. Potential pollution problems resulted when manure was applied on a long-term basis at recommended rates (475 kg/N ha) in Alberta, Canada (75).

To use manures efficiently as a nutrient source for crop production, the spread pattern must be uniform throughout the field (76). Nonuniform applications of manure result in improper nutrient crediting and can increase the possibility of overfertilization, which threatens groundwater and surface water quality (77). Manure should be applied when crops have the highest need for nutrients. The quicker the nutrients are used, the less the chances of pollution due to runoff into surface watercourses or leaching into groundwater sources.

Manure should be stored in properly designed facilities that are protected from excessive runoff, flooding, or overflow that would allow contamination of surface water. Injecting manure is especially important in preventing contamination of surface water supplies. Injection methods place liquid manure below the soil surface, thus eliminating surface runoff on sloping soils and volatilization of ammonia from the manure on all soils (76).

FUTURE DIRECTIONS

As Owens (32) points out, much is known about using N for efficient crop production and minimizing its adverse impacts on surface and subsurface water quality. Nevertheless, there are areas where further knowledge is needed to fill in the gaps and to improve upon what is already known. Following Owens (32), the areas mentioned here demand particular attention in research:

- *Subsurface water quality*: Much of the reported research in this area has been conducted with soil columns and lysimeters. There is need for more groundwater quality research on a field or watershed scale. A systems approach, in which existing knowledge is harnessed into quantitative, explanatory models, is more attractive and acceptable.
- *Urban use*: Fertilization of lawns and turf areas and the subsequent impacts on groundwater quality need more regulatory attention. In many states, there is no environmental law protecting water resources from overfertilization of privately owned lawn or gardens, although according to a Gallup poll, 22 million U.S. homeowners spent more than $14.6 billion on professional landscape/lawn care/tree care services in 1997. Research on this aspect of N management has been limited and needs to be revisited as a potential major source of N pollution.
- *Soil N Availability Tests*: Crops are probably more deficient in N than in any other element, and yet there are no widely accepted methods of testing soils other than testing for residual NO_3^- N and NH_4^+ N (78). Considerable attention needs to be directed toward developing soil tests that provide an accurate in-season assessment of N need for a crop, yet take into account variation in soil types and kinds of crop while recommending fertilizer N needs.
- *Education*: Knowledge about N fertilizer efficiency and techniques for minimizing nitrate leaching are available but are not consistently used. The real value of continued research will be the adaptation and implementation of these technologies, which require a major educational/outreach effort.

BIBLIOGRAPHY

1. Rowe, M.L. and Stinnett, S. (1975). *Nitrogen in the Subsurface Environment*. Publication # 600375030. Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC.
2. Follett, R.F. (2001). Nitrogen transformation and transport processes. In: *Nitrogen in the Environment: Sources, Problems and Management*. R.F. Follet and J.L. Hatfield (Eds.). Elsevier, Amsterdam, pp. 17–44.
3. Stevenson, F.J. (1982). Origin and distribution of nitrogen in soil. In: *Nitrogen in Agricultural Soils*. F.J. Stevenson (Ed.). Agronomy Monograph No. 22, American Society of Agronomy, Madison, WI, pp. 1–42.
4. Wells, K.L., Sims, J.L., and Smith, M.S. (1997). *Nitrogen in Kentucky Soils, AGR-43*. University of Kentucky Cooperative Extension Service, University of Kentucky, Lexington, KY.
5. Henry, C., Sullivan, D., Rynk, R., Dorsey, K., and Cogger, C. (1999). *Managing Nitrogen from Biosolids*. Publication # 99–508. Northwest Biosolids Management Association and Washington State Department of Ecology, Olympia, WA.
6. Hallberg, G.R. and Keeney, D.R. (1993). Nitrate. In: *Regional Ground-Water Quality*. W.M. Alley (Ed.). Van Nostrand Reinhold, New York, pp. 297–322.
7. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
8. Hallberg, G.R. (1989). Nitrate in groundwater in the United States. In: *Nitrogen Management and Ground Water Protection*. R.F. Follett (Ed.). Elsevier, Amsterdam, pp. 35–74.
9. Keeney, D.R. (1982). Nitrogen management for maximum efficiency and minimum pollution. In: *Nitrogen in Agricultural Soils*. F.J. Stevenson (Ed.). Agronomy Monograph 22, American Society of Agronomy, Madison, WI, pp. 605–649.
10. USEPA. (1996). *Environmental indicators of water quality in the United States:* EPA 841-R-96-002. EPA Office of Water, Washington, DC.
11. Baird, J.V. (1990). *Soil Facts: Nitrogen Management and Water Quality.* AG-439-2. North Carolina Cooperative Extension Service. North Carolina State University, Raleigh, NC. Available: http://www.soil.ncsu.edu/publications/Soilfacts/AG-439-02/index.htm.
12. Rabalais, N.N., Wiseman, W.J., Turner, R.E., Sengupta, B.K., and Dortch, Q. (1996). Nutrient changes in the Mississippi River and systems responses on the adjacent continental shelf. *Estuaries* **19**: 386–407.
13. Downing, J.A., Rabalais, N.N., Diaz, R., Zimmerman, R., Baker, J.L., and Duffy, M.D. (1999). *Gulf of Mexico Hypoxia: Land-Sea Interactions*. Council for Agricultural Science and Technology, Task Force Report No. 134.
14. Howarth, R.W. (2000). *Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution*. National Academy Press, Washington, DC.
15. Killpack, S.C. and Buchholz, D. (1993). *Nitrogen in the Environment: How Nitrogen Enters Groundwater*. Water Quality Initiative publication WQ256. University of Missouri Extension, University of Missouri-Columbia, Columbia, MO.
16. O'Leary, M., Rehm, G., and Schmitt, M. (1994). *Understanding Nitrogen in Soils*. FO-03770-GO. University of Minnesota Cooperative Extension Service, University of Minnesota, Minneapolis, MN. Available: http://www.extension.umn.edu/distribution/cropsystems/DC3770.html.
17. Nebel, B. and Wright, R. (2000). *Environmental Science*. Prentice-Hall, Upper Saddle River, NJ.
18. Shapiro, C., Wortmann, C.S., and Deloughery, R. (2002). *Manure Testing: What to Request?* Nebraska Cooperative Extension NF02-507. University of Nebraska-Lincoln, Lincoln, NE.
19. Peacock, B. (1996). *Balancing the Nitrogen Budget*. NG2-96. University of California Cooperative Extension, University of California, Tulare, CA.
20. Wells, K.L., Thomas, G.W., Sims, J.L., and Smith, M.S. (2000). *Managing Soil Nitrates for Agronomic Efficiency and Environmental Protection.* AGR-147. University of Kentucky Cooperative Extension Service, University of Kentucky, Lexington, KY.

21. Vitosh, M.L., Johnson, J.W., and Mengel, D.B. (1995). *Tri-state Fertilizer Recommendations for Corn, Soybeans, Wheat, and Alfalfa.* Ext. Bull. E-2567. Michigan State University, East Lansing, MI.
22. Trautmann, N.M., Porter, K.S., and Wagenet, R.J. (1998). *Nitrogen: The Essential Element*. Pesticide Fact Sheets and tutorials. Cornell Cooperative Extension, Cornell University, Ithaca, NY.
23. Stecker, J.A. (1993). *Nitrogen management for no-tillage systems in Missouri.* Agricultural publication G9175. University of Missouri Extension, University of Missouri-Columbia, Columbia, MO.
24. Travis, J.W., Halbrendt, N., Hed, B., Rytter, J.L., Anderson, E., Jarjour, B., and Griggs, J. (2003). A practical guide to the application of compost in vineyards. Available: http://fpath.cas.psu.edu/compostguide.pdf. Accessed 21 March 2004.
25. Khan, M. (1999). *Nitrogen Loss from Soils*. NDSU crop and pest report, Issue 3. North Dakota State University, Fargo, ND.
26. USEPA. (1973). *Nitrogenous Compounds in The Environment.* EPA-SAB-73-001. U.S. Government Printing Office, Washington, DC.
27. Follett, J.R. and Follett, R.F. (2001). Utilization and metabolism of Nitrogen by humans. In: *Nitrogen in the Environment: Sources, Problems and Management*. R.F. Follett and J.L. Hatfield (Eds.). Elsevier, Amsterdam. pp. 65–92.
28. USEPA. (1976). *Quality Criteria for Water*. U.S. Government Printing Office, Washington, DC, p. 256.
29. Sauer, T.J., Alexander, R.B., Brahana, J.V., and Smith, R.A. (2001). The importance and role of watersheds in the transport of Nitrogen. In: *Nitrogen in the Environment: Sources, Problems, and Management*. R.F. Follett (Ed.). Elsevier, Amsterdam, pp. 65–92.
30. Alexander, R.B., Smith, R.A., and Schwarz, G.E. (1995). The regional transport of point and nonpoint source N to the Gulf of Mexico. In: *Proc. Gulf of Mexico Hypoxia Conf.*, December 5–6, Kenner, LA.
31. Lilly, J.P. (1991). *Best Management Practices for Agricultural Nutrients*. Publication No. AG-439-20. North Carolina Cooperative Extension Service, North Carolina State University, Raleigh, NC.
32. Owens, L.B. (1993). Impact of soil N management on the quality of surface and subsurface water. In: *Soil Processes and Water Quality*. R. Lal and B.A. Stewart (Eds.). Lewis, Boca Raton, FL, pp. 137–163.
33. Meisinger, J.J., Hargrove, W.L., Mikkelsen, R.L., Williams, J.R., and Benson, V.W. (1991). Effects of cover crops on ground water quality. In: *Cover Crops for Clean Water.* W.L. Hargrove (Ed.). Soil and Water Conservation Society, Ankeny, IA, pp. 57–68.
34. Russelle, M.P. and Hargrove, W.L. (1989). Cropping systems: Ecology and management. In: *Nitrogen Management and Groundwater Protection*. R.F. Follett (Ed.). Elsevier, Amsterdam, pp. 277–317.
35. Muller, J.C., Denys, D., Morlet, G., and Mariotta, A. (1988). Influence of catch crops on mineral nitrogen leaching and its subsequent plant use. In: *Nitrogen Efficiency in Agricultural Soils*. D.S. Jenkinson and K.A. Smith (Eds.). Elsevier Applied Science, New York.
36. Roos, D. (1998). Cover crops for sustainable vegetable production. Available: http://www.geocities.com/RainForest/Canopy/1118/lecture.html.
37. Meisinger, J.J., Shipley, P.R., and Decker, A.M. (1990). Using winter cover crops to recycle nitrogen and reducing leaching. In: *Conservation Tillage for Agriculture in the 1990s*. J.P. Mueller and M.G. Wagger (Eds.). Spec. Bull. 90-1, North Carolina State University, Raleigh, NC, pp. 3–6.
38. Hemphill, D.D. and Hart, J. (1991). *Effect of Nitrogen Rate and Placement on Broccoli Yield and Nitrogen Uptake*. Report to the Oregon Processed Vegetable Commission.
39. Dabney, S.M. (1998). Cover crop impacts on watershed hydrology. *J. Soil Water Conserv.* **53**(3): 207–213.
40. Kaspar, T.C., Radke, J.K., and Laflen, J.M. (2001). Small grain cover crops and wheel traffic effects on infiltration, runoff, and erosion. *J. Soil Water Conserv.* **56**: 160–164.
41. Hartwig, N.L. (2003). Cover crops and conservation tillage for soil erosion control on cropland. Pennsylvania State University, University Park, PA. Available: http://www.agronomy.psu.edu/Extension/LivingMulch/sect9toc.htm.
42. Al-Kaisi, M. (2001). Value of crop rotation in nitrogen management. *Integrated Crop Management Newsletter* articles. Iowa State University, Ames, IA.
43. Johnson, J.W., Welch, L.F., and Kurtz, L.T. (1975). Environmental implications of N fixation by soybeans. *J. Environ. Qual.* **4**(3): 303–306.
44. Seelig, B.D. and Nowatzki, J. (2001). *Working to Avoid Nitrogen Contamination*. AE-1218. NDSU Extension Service, North Dakota State University, Fargo, ND.
45. Baker, J.L. and Melvin, S.W. (1994). Chemical management, status, and findings. In: *Agricultural Drainage Well Research and Demonstration Project—Annual Report and Project Summary*. Iowa Department of Agriculture and Land Stewardship and Iowa State University, Ames, IA, pp. 27–60.
46. Dickey, E., Harlan, P., Vokal, D., and Kisling Crouch, C.J. (1982). *Effects of Agricultural Runoff on Nebraska Water Quality*. Neb Guide G82-586. University of Nebraska Cooperative Extension Service, University of Nebraska-Lincoln, Lincoln, NE.
47. Hairston, J.E. (1995). *Soil Management to Protect Water Quality—Cropping and Nonstructural Agronomic Practices for Soil Management*. ANR-790-4.3.3. Alabama Cooperative Extension Services, Auburn University, Auburn, AL.
48. USDA. (1996). *Conservation Crop Rotation Effects on Soil Quality*. Natural Resources Conservation Service, Agronomy Technical Note No. 2, August 1996.
49. Heath, M.E., Metcalfe, D.S., and Barnes, R.F. (1976). *Forages—The Science of Grassland Agriculture*, 3rd Edn. Iowa State University Press, Ames, IA.
50. Weston, D. (1994). *Water Quality: The Tillage Component.* AE-1072. NDSU Extension Service, North Dakota State University, Fargo, ND.
51. Vanderwel, D. and Abday, S. (2001). An introduction to water erosion control. Alberta Agriculture, Food and Rural Development. Accessed 19 March 2004. Available: http://www.agric.gov.ab.ca/agdex/500/72000003.html.
52. Legg, J.O. and Meisinger, J.J. (1982). Soil nitrogen budgets. In: *Nitrogen in Agricultural Soils*. F.J. Stevenson (Ed.). Agronomy Monograph 22, ASA-CSSA-SSSA, Madison, WI, pp. 503–565.
53. Laflen, J.M. and Tabatabai, M.A. (1984). Nitrogen and phosphorus losses from corn-soybean rotations as affected by tillage practices. *Trans. Am. Soc. Agric. Eng.* **27**: 58–63.
54. Owens, L.B. and Edwards, W.M. (1993). Tillage studies with a corn/soybean rotation: Surface runoff chemistry. *Soil Sci. Soc. Am. J.* **57**(4): 1055–1060.

55. Thomas, G.W., Smith, M.S., and Philips, R.E. (1989). Impact of soil management practices on nitrogen leaching. In: *Nitrogen Management and Groundwater Protection*. R.F. Follett (Ed.). Elsevier, Amsterdam, pp. 247–276.
56. Baker, J.L. and Laflen, J.M. (1983). Water quality consequences of conservation tillage. *J. Soil Water Conserv.* **38**(3): 186–193.
57. Baker, J.L. and Johnson, H.P. (1983). Evaluating the effectiveness of BMPs from field studies. In: *Agricultural Management and Water Quality*. G.W. Bailey and F.W. Schaller (Eds.). Iowa State University Press, Ames, IA, pp. 281–304.
58. Thomas, G.W., Blevins, R.L., Phillips, R.E., and McMahon, M.A. (1973). Effect of killed sod mulch on nitrate movement and corn yield. *Agron. J.* **65**: 736–739.
59. Quisenberry, V.L. and Phillips, R.E. (1976). Percolation of surface-applied water in the field. *Soil Sci. Soc. Am. J.* **40**: 484–489.
60. Edwards, W.M., Shiptalo, M.J., and Norton, L.D. (1988). Contribution of macroporosity to infiltration into continuous corn no-tilled watershed: Implications for contaminant movement. *J. Contam. Hydrol.* **3**: 193–205.
61. Hauck, R.D. (1984). *Nitrogen in Crop Production*. ASA-CSSA-SSSA, Madison, WI.
62. Lal, R. and Stewart, B.A. (1994). *Soil Processes and Water Quality*. Lewis, Boca Raton, FL.
63. Pierzynski, G.M., Sims, J.T., and Vance, G.F. (1994). *Soils and Environmental Quality*. Lewis, Boca Raton, FL.
64. Smucker, D. (2001). Nitrogen management practices for potato production. *The Vegetable Growers News*. July 2001.
65. Bundy, L.G. and Malone, E.S. (1988). Effect of residual profile nitrate on corn response to applied nitrogen. *Soil Sci. Soc. Am. J.* **52**: 1377–1383.
66. El-Hout, N.M. and Blackmer, A.M. (1990). Nitrogen status of corn after alfalfa in 29 Iowa fields. *J. Soil Water Conserv.* **45**: 115–117.
67. McVay, K.A., Radcliffe, D.E., and Hargrove, W.L. (1989). Winter legume effects on soil properties and nitrogen fertilizer requirements. *Soil Sci. Soc. Am. J.* **53**: 1856–1862.
68. Smith, S.J., Schepers, J.S., and Porter, L.K. (1990). Assessing and managing agricultural nitrogen losses to the environment. In: *Advances in Soil Science*. B.A. Stewart (Ed.). Springer-Verlag, New York, pp. 1–43.
69. Randall, G.W. (1996). BMPs For nutrient management. Ill. *Fert. Conf. Proc.*, January 29–31, 1996.
70. Mahler, R.L. and Mahler, K.A. (1985). *Nitrogen BMPs for Small Farms*. WQ-20. University of Idaho Cooperative Extension system, University of Idaho, Moscow, ID.
71. Swdyky, R. (2003). Nitrogen—split application—A risk management tool for poor soil moisture conditions. *Saskatchewan Soil Conserv. Assoc.* (SSCA) *Newslett.*, Issue 37. Available: http://ssca.usask.ca/Issue37/Rich1.htm.
72. Jacobsen, J.S., Tanaka, D.L., and Bauder, J.W. (1993). Spring wheat response to fertilizer placement and nitrogen rate with limited moisture. *Commun. Soil Sci. Plant Anal.* **24**: 187–195.
73. Hairston, J.E. and Stribling, L. (1995). *Fertilizer Management to Protect Water Quality—Fertilizer Application Techniques*. ANR-790-4.4.3. Alabama Cooperative Extension Services, Auburn University, Auburn, AL.
74. McFadden, V.J. and Melvin, S. (1991). *Protecting our Water Quality with Livestock Waste Management.* PM 1428b. Iowa State University Cooperative Extension Service, Iowa State University, Ames, IA.
75. Chang, C., Sommerfeldt, T.G., and Entz, T. (1991). Soil chemistry after eleven annual applications of cattle feedlot manure. *J. Environ. Qual.* **20**: 475–480.
76. Randall, G.W. and Schmitt, M.A. (1993). *Best Management Practices for Nitrogen Use Statewide in Minnesota*. FO-06125-GO.University of Minnesota, Cooperative extension service, University of Minnesota, MN. Available: http://www.extension.umn.edu/distribution/cropsystems/DC6125.html.
77. Hendrickson, S. (2002). *BMPs (best management practices) to Protect Groundwater Quality—Nitrogen Management*. University of Wisconsin Extension, University of Wisconsin, Madison, WI.
78. Dahnke, W.C. and Johnson, G.V. (1990). Testing soils for available nitrogen. In: *Soil Testing and Plant Analysis*, 3rd Edn. R.L. Westerman (Ed.). Soil Science Society of America, Madison, WI, pp. 127–139.

SOIL PHOSPHORUS AVAILABILITY AND ITS IMPACT ON SURFACE WATER QUALITY

Dibyendu Sarkar
University of Texas at San Antonio
San Antonio, Texas

Phosphorus (P) is ubiquitous in soils and is an essential element for plant and animal growth. Without enough soil P, crop yields will be low. However, excess soil P may become a part of surface runoff and can accelerate eutrophication of receiving fresh waters. Eutrophication has been identified as the main problem in U.S. surface waters that has impaired water quality. Although P and nitrogen (N) both influence eutrophication, P is generally the limiting nutrient in freshwater bodies. The degree of P input in surface waters depends on the soil availability of P and the degree of transport of dissolved and particulate P, especially in agricultural lands that have been fertilized for a long period of time. The availability of P in soils is controlled by various factors, such as soil pH, clay content, soil organic matter, amorphous Fe/Al oxide and carbonate content, soil moisture content, and degree of soil compaction. The amount of available P is generally estimated by a variety of soil-test methods, including both chemical extractants and nondestructive techniques simulating plant root action. Phosphorus transport from the source to surface waters is controlled by the rate of erosion and surface runoff. Hence, reducing P loss in agricultural runoff may be brought about by implementing proper source and transport control strategies. Popular source control techniques include soil P management, estimating thresholds for environmental risk assessment, manure management, and refining feed rations, common transport management techniques include conservation practices such as reduced tillage, cover crops, and buffer strips in critical areas of P export from a watershed.

Phosphorus makes up about 0.12% of the earth's crust. It forms complex compounds with a wide variety of elements—about 150 minerals are known that contain at least 0.44% P (1). Phosphorus is ubiquitous in soils. It is found naturally in all soils (1) and ranges between less than 0.04% in the sandy soils of the Atlantic and Gulf coastal plains to more than 0.3% in the soils of the northwestern United States (2). Phosphorus is an essential element for plant and animal growth. Its input has long been recognized as necessary to maintain profitable crop and livestock production (3). Without enough soil phosphorus, crop yields will be low (4). However, many soils in the United States reportedly contain P well in excess of crop needs, which may enter lakes and streams in runoff from landscapes that drain to surface water bodies. Sources of P contamination of surface waters are numerous and include both nonpoint and point sources, such as agriculture, municipal sewage treatment plants, individual septic treatment systems, decaying plant material, runoff from urban areas and construction sites, stream bank erosion, and wildlife (5). Drainage of P-rich runoff to surface water bodies may increase their biological productivity by accelerating eutrophication, the natural aging of lakes or streams brought on by nutrient enrichment. The United States Environmental Protection Agency (USEPA) has identified eutrophication as the main problem in U.S. surface waters that have impaired water quality (6). Because it causes increased growth of undesirable algae and aquatic weeds, as well as oxygen shortages, eutrophication restricts water use for fisheries, recreation, industry, and drinking (7). In addition, associated periodic surface blooms of cyanobacteria (blue-green algae) occur in eutrophic drinking water supplies and may pose a serious health hazard to livestock and humans. Although eutrophication is a natural process, it can be hastened by changes in the land use of a watershed that increase the amount of nutrients added to an aquatic system (7). Recent outbreaks of the dinoflagellate *Pfiesteria piscicida* in the eastern United States, most notably in the Chesapeake Bay, have been linked to high nutrient levels in affected waters (3). Nitrogen (N) and P both influence eutrophication, but P is the critical element in most freshwaters (7). Due to low natural levels of P, biological productivity in surface waters is usually limited by P availability (5). The identification of P as a major cause of eutrophication in surface waters led to early emphasis on point sources delivering P to surface waters, for example, sewage systems delivering P-enriched wastewater. According to the National Research Council (8), overall trends show about equal numbers of U.S. rivers that have increasing and decreasing P loads. In general, the decreases are linked to point-source reductions, and the increases are linked to nonpoint source increases that result from increased agricultural land use (9). This review will address the concept of soil availability of P in agricultural lands and the relationship between P availability, P transport, and surface water eutrophication. Popular practices for P management will also be discussed.

PHOSPHORUS IN SOILS

The total P content of most surface soils is low, averaging only 0.6%. This compares to an average soil content of 0.14% N and 0.83% potassium (K) (2). In most soils, the P content of surface horizons is greater than that of the subsoil. Except in special situations, added P tends to be fixed by the soils where it is applied and allows little movement down through the soil profile (7). In addition, P is cycled from roots to aboveground parts of the plant and redeposited in crop residues on the soil surface. This builds up organic material and stimulates biological activity in surface layers. In reduced tillage systems, fertilizers and manures are incorporated only to shallow depths, thereby exacerbating P buildup in the top 5 inches of soil (7).

Many factors influence the P content of soil. Among these are (1) the type of parent material from which the soil is derived, (2) the degree of weathering, and (3) climatic conditions. Among other factors influencing soil P levels are the rates of crop removal, erosion, and P fertilization (2).

There are two primary forms of P in soil systems—organic and inorganic. The organic form of P in many soils represents more than half of the total P. Organic P has three different sink/sources (10), namely nucleic acids (e.g., guanine, adenine, thymine, uracil), phosphate esters of inositate (e.g., phytin), and phospholipids (e.g., lecithin). Biological processes in the soil, such as microbial activity, tend to control the mineralization of organic P. Mineralization is the breakdown or conversion of organic P to inorganic solution P (7). The inorganic form of P in soils (orthophosphates) exist in various forms: (1) oxide–phosphate complexes, such as clay mineral edge-P complexes, and (2) minerals of apatite (Ca phosphates), Mn phosphates, Fe phosphates, Al phosphates, and to a lesser degree, pyrophosphate (10).

AVAILABILITY OF SOIL PHOSPHORUS

Loosely defined, P availability refers to the amount of soil P that is readily available for uptake by plants. If the amount of available P exceeds plant needs, the excess P may become available for drainage to surface water bodies as part of the surface runoff, thereby degrading surface water quality. As a rule of thumb, soluble P is considered the most available. However, the solubility of many inorganic P compounds, as well as organic P, is generally low, and only small amounts of soil P are in solution at any one time. In fact, most soils contain less than a 1 lb acre^{-1} of soluble P (2). Soluble P in soils, either from fertilizer or natural weathering, reacts with clay, iron, and aluminum compounds in the soil and is converted to less mobile (hence, less “available”) forms by the process of P fixation. Because of the fixation processes (e.g., sorption and precipitation), P moves very little in most soils, stays close to its place of application, and crops seldom absorb more than 20% of fertilizer P during the first cropping system after application. This fixed, residual P remains in the rooting zone and becomes slowly available to succeeding crops (2).

Factors Controlling Availability of Soil Phosphorus

Soil P availability is influenced by the following factors (2):

- *Soil pH*: Soil pH affects both adsorption and precipitation of phosphorus. Precipitation of P as slightly soluble Ca phosphates occurs in calcareous soils whose pH values are around 8.0. Under acidic conditions, P is precipitated as Fe or Al phosphates that have low solubility. From a precipitation standpoint, the maximum availability of P generally occurs in the pH range of 6.0 to 7.0 (2). As a typical oxyanion, phosphate tends to be adsorbed more at lower pH than at higher pH. Hence, P generally tends to be more available in high pH soils, unless precipitated as low-solubility basic salts. Amorphous Fe and Al oxides that have high specific surface strongly adsorb P in soil systems (11).

- *Soil Organic Matter*: High amounts of organic matter in soils promote mineralization of organic P, thereby providing available P for plant growth. Organic matter also acts as a chelating agent, thereby preventing the formation of insoluble iron phosphates. Application of organic materials such as manure, plant residues, or green manure to soils that have high pH values increases the availability potential of mineral forms of P in the soil (2).
- *Clay Content*: Fine-textured soils such as clay and clay loams have greater P-fixing capacity than coarse-textured sandy soils. Among soils of similar clay mineralogy, P fixation increases as clay content increases and crystallinity decreases (12). Among clay minerals, 1:1 type clays (kaolinite) have greater P-fixing capacity than 2:1 type clays (e.g., montmorillonite, illite). Soils formed under high rainfall and high temperatures generally contain large amounts of kaolinitic clays and therefore have relatively lower soil P availability.
- *Soil Aging*: Fixation of soil P increases with time of contact between soluble P and soil particles. Consequently, more efficient use of fertilizer P is generally obtained by applying the fertilizer shortly before planting the crop (2). As soil ages, the possibility that P binds to soil particles increases, thereby decreasing its potential availability.
- *Soil Compaction and Soil Moisture*: Soil compaction reduces the soils oxygen supply and decreases the ability of plant roots to absorb soil P. Compaction reduces aeration and pore space in the root zone. This reduces P uptake and plant growth. Compaction also decreases the soil volume that plant roots penetrate and limits their access to soil P. Many studies have been conducted on the effect of soil moisture on P availability and response to applied P (13). Power et al. (14) reported that 53% of the variation in the fertilizer response of spring wheat on medium P soils resulted from variation in soil moisture. MacKay and Barber (15) showed that P availability decreased at soil moisture contents above and below field capacity because of the effects on both P diffusion and root growth rates. Uptake of P by corn dropped more than total plant weight when soil moisture was decreased below field capacity, indicating that the decrease in uptake was not likely due only to decreased demand.

In addition to the factors discussed before, several other factors define soil P availability (13). These factors include, but are not limited to soil solution P concentration, the amount of solid phase P that can enter the soil solution, the dissolution/desorption rate, the diffusion rate, Fe content, carbonate content, carbonate specific surface area, and organic P content (16–19).

Phosphorus Availability Indices

To evaluate P availability in soils, numerous soil tests have been developed that extract varying amounts of P, depending on the types of extractants used. The extractants can be generally classified into several categories (20):

- Water or unbuffered salt solutions (e.g., $CaCl_2$).
- Dilute concentrations of weak acids (e.g., lactate, acetate) with or without a complexing agent (e.g., F^-, EDTA).
- Dilute concentrations of strong acids (e.g., HCl, H_2SO_4) with or without a complexing agent (e.g., F^-, lactate, EDTA).
- Buffered alkaline solutions (e.g., $NaHCO_3$, NH_4HCO_3) with or without a complexing agent (e.g., EDTA).
- Anion exchange resin or iron oxide impregnated filter paper strips.
- Isotopic exchange with ^{32}P.

Water derives readily soluble P; unbuffered/neutral salt solutions derive easily exchangeable P (21). Dilute strong acid solutions solublilize Ca-P, Al-P, and to a lesser extent, Fe-P (22). Complexing agents are generally introduced to prevent readsorption of P by Fe oxides (20). Anion exchange resins and iron oxide impregnated filter papers (nondestructive to soil components in contrast to chemical extractants) function as sinks that simulate the action of plant roots by continuously removing dissolved P from the soil solution. The quantity of P that is isotopically exchangeable within a specified time interval gives an estimate of labile (hence, available) soil P (20).

Criteria for a suitable soil test have been detailed in several reports (13). Thomas and Peaslee (23) felt that the extract should (1) remove a reproducible and consistent proportion of the soil's available P and (2) reflect the extent and nature of reactions between soil and added P. Barrow (24) also felt that a soil test was an integration of several factors affecting soil P status (e.g., soil P buffering capacity, P quantity, soil P reaction time) into one numerical value. Kamprath and Watson (25) noted that (1) the extractant should extract a proportionate fraction of available P from soils that differ in other properties, (2) the procedure should work with reasonable accuracy and speed, and (3) the soil test P value determined should be correlatable with P uptake. On the basis of the major soil types and their properties, various states in the United States have adopted different soil test methods to predict available P. A field rating system, the "P index," has been developed to assess the potential for soil P to contribute to surface water pollution (26).

TRANSPORT OF SOIL PHOSPHORUS

Phosphorus enters lakes and streams in runoff from landscapes that drain to those surface water bodies (5).

The eutrophication process is clearly accelerated by the addition of P to surface waters. All forms of P (soluble, adsorbed, precipitated, and organic) are susceptible to transport from source soils to water bodies. Transport of soil P occurs primarily through surface flow, although the background levels of P entering streams and lakes via subsurface flow certainly reflect the impacts of land use. For example, in soils that have considerable slope, surface or overland flow can infiltrate into the soil during movement down the slope, move laterally through a conduit in the soil, and reappear as surface flow (7). However, in most cases, the concentration of P in subsurface flow, it has been found is quite low and well below the eutrophication threshold. In overland flow (surface runoff), P may be dissolved in runoff water (soluble/dissolved P or DP). The small quantity of soluble or readily desorbable P in most soil environments is due to the low solubility of phosphate minerals and the considerable adsorption capacities of clays and Fe/Al oxides for P. These factors result in the majority of total P transport that occurs as particulate P (PP). Particulate P enters streams and other surface waters, first undergoing a solubilization reaction (e.g., desorption) before becoming available for aquatic biota. The mechanisms involved in soluble P transport include initial desorption or dissolution of P bound by soil particles, followed by water movement from the source soil to a stream or a lake that later intercepts a sensitive water body (27).

Factors Controlling Surface Transport of Soil Phosphorus

Two major factors influence P losses from the landscape and movement to surface water bodies (5):

- *Erosion*: Particulate P is the primary form of P contributing to degradation of surface water quality. Erosion largely determines PP movement in the landscape. Sources of PP in surface waters include eroding surface soil, stream banks, channel beds, and plant materials. The contribution of eroded soil material to P enrichment of surface waters is rather complex (5). First, part of the sediment eroded from the landscape deposits at the toe slope in depressional areas, along field boundaries, or in grassy riparian zones. Thus, sediment yield from the field to a water body is only a portion of that eroded from the slope. Second, during the movement of sediment, the finer fractions are preferentially transported, whereas larger particles tend to settle out. Therefore, the P concentration and reactivity of eroded particulate material is usually greater than that of the source soil. The enrichment of P may increase as much as six-fold as the relative movement of fine to coarser particles increases. For this reason, fine-textured soils (clay loam, silty clay loam, silty clay, or clay) have a higher potential of supplying P to surface waters if erosion occurs (5).
- *Runoff*: As discussed earlier, runoff water across the soil surface can contain significant concentrations of DP. In addition to the readily available P loss from surface soils via desorption/dissolution during rainfall, DP can also be lost from standing vegetation via spring snowmelt because P contained in the tissue is released due to breakdown of plant cells by freezing and thawing. The removal of P from plant residue by overland drainage may account for differences among watersheds and seasonal fluctuations in P movement (5). Concentrations of DP in subsurface flow are low because the P-deficient subsoils sorb much of soluble P contained in the water percolating through the soil profile. Exceptions may occur in organic, highly permeable, peaty, or waterlogged soils, which tend to fix less P (7).

There is a well-defined relationship between the DP and PP fractions of total P in runoff as a function of soil erosion. As erosion increases, the PP fraction of total P increases, whereas the DP fraction decreases significantly (5). Irrigation, especially furrow irrigation, can significantly increase the potential for soil and water contact and therefore, can increase P loss by both surface runoff and erosion in return flows. Furrow irrigation exposes unprotected surface soil to the erosive effect of water movement. The process of irrigation also has the potential to increase greatly the land area that can serve as a possible source for P movement, a fact that is especially important in the western United States (3) where the surface water bodies suffer from an increased threat of eutrophication.

ENVIRONMENTAL MANAGEMENT OF SOIL PHOSPHORUS

When looking at management and/or remedial practices to minimize the environmental impact of P, several important factors must be considered. To cause an environmental problem, there must be a source of P (i.e., high soil P levels, manure or fertilizer applications, etc.), and it must be transported to a sensitive location (i.e., by runoff, erosion, or to a lesser extent, leaching). Hence, reducing P loss in surface runoff (thereby lowering the eutrophication potential of surface waters) may be brought about by source and transport control strategies (3).

Source Management

Source management attempts to minimize soil P buildup by controlling the amount of P applied to a given area. One of the most highly recommended avenues for source management is by measuring (and monitoring) soil test P in agricultural lands and estimating threshold levels for environmental risk assessment.

Decades of fertilization at rates exceeding those of crop removal have resulted in widespread increases in the P status of many agricultural soils, even to levels considered “excessive.” Environmental concern has forced many states to consider developing recommendations for P applications and watershed management based on the potential for P loss in agricultural runoff (3). As a result, well-defined “agronomic” and “environmental” thresholds for soil test P have been developed for many states. The environmental threshold levels range from two times in Michigan (40 ppm P-agronomic and 75 ppm P-environmental) to four times in Texas (44 ppm vs. 200 ppm). However, one still cannot accurately project that a soil test level above an expected crop response level exceeds crop needs and therefore is potentially polluting. What will be crucial in terms of managing source P based in part on soil test levels will be the interval between the critical soil P value for crop yield and runoff P (3).

Among the other popular source management techniques currently being implemented in farmlands are manure management (e.g., composting, lime or alum amendment, separation of solids from liquids); reduction of off-farm inputs of P in animal feeds (3); and controlling the rate, method and time of P application (5).

Transport Management

Phosphorus loss via erosion and runoff may be reduced by increasing residue cover on the soil surface through conservation tillage (5). Results from a flat, tiled, clay loam site in Michigan show significantly lower losses of sediment, total P, and soluble P for chisel tillage compared to moldboard plow tillage. However, sediment reductions were much greater than the reductions in total or soluble P. In other research, DP concentration in the runoff from no-till practices has often been greater than that from conventional practices (5).

Strategically placed and properly designed filter strips, it has been shown, effectively reduce erosion and P movement, especially PP (5). However, the strips have had little consistent effect on reducing DP concentrations.

Other common practices for reducing potential P transport include terracing, contour tillage, riparian zones, cover crops, and impoundments (settling basins). Basically, all these methods reduce the impact of rainfall on the soil surface, reduce runoff volume and velocity, and increase soil resistance to erosion. Most of these practices are more efficient in reducing PP than DP, and none of them should be relied on as the primary or sole practice to reduce P losses in surface runoff (3). Several researchers report little decrease in lake productivity (i.e., eutrophic plant growth) with reduced P inputs, following implementation of conservation measures (28,29). Thus, the effect of remedial measures in the contributing watershed will be slow for many cases of poor water quality (7).

BIBLIOGRAPHY

1. Cathcart, J.B. (1980). World phosphate reserves and resources. In: *The Role of Phosphorus in Agriculture*. F.R. Khasawneh, E.C. Sample, and D.J. Kamprath (Eds.). ASA-CSSA-SSSA, Madison, WI.
2. Griffith, B. (2002). Phosphorus. *Efficient Fertilizer Use Manual*. IMC Global Inc.
3. Sharpley, A.N. et al. (1999). *Agricultural Phosphorus and Eutrophication*. ARS-149. USDA.
4. Stout, W.L., Sharpley, A.N., and Pionke, H.B. (1998). Reducing soil phosphorus solubility with coal combustion by-products. *J. Environ. Qual.* **27**: 111–118.
5. Randall, G. et al. (1998). Phosphorus. *Transport to and Availability in Surface Waters*. University of Minnesota Extension Service Publication, Minneapolis, MN.
6. U.S. Environmental Protection Agency. (1996). *Environmental Indicators of Water Quality in the United States*. EPA 841-R-96-002.
7. Sharpley, A.N. (1999). *Managing Phosphorus for Agriculture and the Environment*. College of Agricultural Sciences Cooperative Extension Services Publication, The Pennsylvania State University, University Park, PA.
8. National Research Council. (1993). Soil and water quality. *An Agenda for Agriculture*. National Academy Press, Washington, DC.
9. Voss, R. and Griffith, B. (1998). *Phosphorus and Surface Water*. Illinois Fertilizer and Chemical Association Report, St. Anne, IL.
10. Evangelou, V.P. (1998). Environmental soil and water chemistry. *Principles and Applications*. John Wiley & Sons, New York.
11. Peters, J.M. and Basta, N.T. (1996). Reduction of excessive bioavailable phosphorus in soils by using municipal and industrial wastes. *J. Environ. Qual.* **25**: 1236–1241.
12. Sanchez, P.A. and Uehara, G. (1980). Management considerations for acid soils with high phosphorus fixation capacity. In: *The Role of Phosphorus in Agriculture*. F.R. Khasawneh, E.C. Sample, and D.J. Kamprath. (Eds.). ASA-CSSA-SSSA, Madison, WI.
13. Fixen, P.E. and Grove, J.N. (1990). Testing soils for phosphorus. In: *Soil Testing and Plant Analysis*, 3rd Edn. R.L. Westerman (Ed.). SSSA Book Series No. 3. SSSA, Madison, WI.
14. Power, J.F., Brown, P.L., Army, T.J., and Klages, M.G. (1961). Phosphorus responses by dryland spring wheat as influenced by moisture supplies. *Agron. J.* **53**: 106–108.
15. MacKay, A.D. and Barber, S.A. (1985). Soil moisture effects on root growth and phosphorus uptake by corn. *Agron. J.* **77**: 519–523.
16. Soper, R.J. and El Bagouri, I.H.M. (1964). The effect of soil carbonate level on the availability of added and native phosphorus in some calcareous soils. *Can. J. Soil Sci.* **44**: 337–344.
17. Fixen, P.E. and Ludwick, A.E. (1982). Residual available phosphorus in near-neutral and alkaline soils: II. Persistence and quantitative estimation. *Soil Sci. Soc. Am. J.* **46**: 335–338.
18. Holford, I.C.R. (1977). Soil properties related to phosphate buffering in calcareous soils. *Commun. Soil Sci. Plant Anal.* **8**: 125–137.
19. Cole, C.V., Innis, G.S., and Stewart, J.W.B. (1977). Simulation of phosphorus cycling in semi-arid grasslands. *Ecology* **58**: 1–15.
20. Kuo, S. (1996). Phosphorus. In: *Methods of Soil Analysis*. D.L. Sparks. (Ed.). Part III: Chemical Methods. SSSA Book Series No. 5. SSSA, Madison, WI.
21. O'Connor, G.A. and Sarkar, D. (1999). *Fate of Land Applied Residuals-Bound Phosphorus*. Final Report, DEP WM 661.
22. Nelson, W.L., Mehlich, A., and Winters, E. (1953). The development, evaluation, and use of soil tests for phosphorus availability. *Agronomy* **4**: 153–188.
23. Thomas, G.W. and Peaslee, D.W. (1973). Testing soils for phosphorus. In: *Soil Testing and Plant Analysis*. L.M. Walsh and J.D. Beaton. (Eds.). SSSA Book Series No. 3. SSSA, Madison, WI.
24. Barrow, N.J. (1980). Evaluation and utilization of residual phosphorus in soils. In: *The Role of Phosphorus in Agriculture*. F.R. Khasawneh, E.C. Sample, and D.J. Kamprath (Eds.). ASA-CSSA-SSSA, Madison, WI.
25. Kamprath, E.J. and Watson, M.E. (1980). Conventional soil and tissue tests for assessing the phosphorus status of soils. In: *The Role of Phosphorus in Agriculture*. F.R. Khasawneh,

E.C. Sample, and D.J. Kamprath. (Eds.). ASA-CSSA-SSSA, Madison, WI.

26. Sims, J.T. (1993). Environmental soil testing for phosphorus. *J. Prod. Agric.* **6**: 501–507.
27. Pierzynski, G.M., Sims, J.T., and Vance, G.F. (1994). *Soils and Environmental Quality*. Lewis, Boca Raton, FL.
28. Young, T.C. and DePinto, J.V. (1982). Algal-availability of particulate phosphorus from diffuse and point sources in the lower Great Lakes basin. *Hydrobiologia* **91**: 111–119.
29. Gray, C.B.J. and Kirkland, R.A. (1986). Suspended sediment phosphorus composition in tributaries of the Okanagan Lakes, BC. *Water Res.* **20**: 1193–1196.

SOIL WATER ISSUES

Eliot C. Roberts
Rosehall Associates
Sparta, Tennessee

Landscape plants and *turfgrass* can eliminate erosion and conserve water most effectively by considering many variables in the makeup of the landscape. These variables include the sizes of mineral components of soil, the components of the root zone, *soil moisture capacity*, soil texture, soil pH, the soil's *biodegradable capacity*, and its mulching capacity. All these factors are explained in this entry.

Because water and soil are the lifeblood of the planet, any policy that governs water use or proposes water conservation must incorporate the dynamic relationship that takes place between plants, the soil and water. The soil on which all plants grow is a highly valuable natural resource. We have a responsibility to protect it from erosion and to conserve and enrich it for future generations, just as we have a responsibility to use our supplies of water judiciously and maintain or improve its quality. *Landscape* plants generally and turf-grasses specifically can help achieve both of these goals.

To realize these benefits, it's helpful to look at what scientists have learned about the characteristics of soil such as texture, size of soil particles, the presence of organisms, capacity to hold moisture, acidity–alkalinity, and the presence of pollutants.

Soils are not homogenous, inert materials. They are composed of mineral particles that include sand, silt, and clay, as well as living and dead *micro/macroflora* and fauna, chemicals, air, and water in various percentages.

SIZES OF THE MINERAL COMPONENTS OF SOIL

From largest to smallest, mineral components of soil consist of the following:

- Stones = 10 to 100 millimeters in diameter (25 per inch).
- Gravel = 2 to 10 millimeters
- Coarse sand = 0.2 to 2 millimeters
- Fine sand = 0.02 to 0.2 millimeter
- Silt = 0.002 to 0.02 millimeters
- Clay = smaller than 0.002 millimeter. Clay particles are so small that they are measured in microns (0.002 millimeter is equivalent to 2 microns)

In terms of comparative size, if we enlarge a clay particle to the size of an apple, then a silt particle on the same scale would be the size of a limousine and a medium sand particle would be the size of the White House in Washington D.C. Because soil particles have relative sizes this small, there are many of them. For example, a pound of medium sand contains about 2.5 million particles, a pound of silt contains more than 2.5 billion particles, and a pound of clay contains more than 40 trillion particles. On the basis of total particle surface, one pound of sand would account for 20 square feet; silt would present 220 square feet of surface, and clay would have 5500 square feet of particle surface (Fig. 1).

THE MULTIPLE COMPONENTS IN A LANDSCAPE ROOT ZONE

Also present in most soils are high microbe counts within the root zone. Often there are more than 900 billion for each pound of soil. In each 1000-square-foot surface to a 6-inch root-zone depth, there will be a total of about 45 quadrillion organisms. As these organisms complete their life cycle and die, they deposit into the soil some 10 pounds of nitrogen, 5 pounds of phosphorus, 2 pounds of potassium, a half-pound of calcium, a half-pound of manganese, and one-third of a pound of sulfur for each 100 pounds of dead organisms on a dry-weight basis.

Soil microbiological processes also convert organic matter into humus. This is an ongoing reaction of great importance. Humus helps to form and stabilize soil aggregates that are essential for deep and extensive root growth. Humus also contributes to the process within the soil that holds and releases nutrients for plant growth.

In addition, many small animals known as soil fauna occupy the root zones of plants and contribute to the living nature of the soil. Depending upon soil conditions that are favorable for these macroorganisms, from 1 million to 2 million may be present for each 1000 square feet of root zone. The live weight of these organisms would range from 15 pounds to 30 pounds per 1000 square feet.

Of course, the water molecule associated with the soil is exceptionally small. One fluid ounce of water contains approximately 1,000,000,000,000,000,000,000,000 (24 zeros, or one trillion trillion) molecules.

GRASS PLANTS HAVE A TREMENDOUS POTENTIAL FOR ROOT GROWTH

Grasses fit right in with the sizes and numbers of soil particles found within this fascinating system. For example, there may be as many as 35 million individual grass plants per acre, or about 800,000 per 1000 square feet. No other type of plant culture involves such crowding. Roots grow down into the soil, and it is there that grass plants have a tremendous potential for root growth—up to 375 miles of roots from one plant and as many as 14 million

Figure 1. In terms of comparative size, if we enlarge a clay particle to the size of an apple, then a silt particle on the same scale would be the size of a limousine, and a medium sand particle would be the size of the White House in Washington, DC.

individual roots that may have a total surface area of 2500 square feet. Thus, root numbers and surfaces fit well within the very small spaces surrounding aggregated soil particles.

It is important to understand soil properties so we can appreciate how turfgrass needs moisture to grow and enhance the environment (Fig. 2).

SOIL'S CAPACITY TO HOLD MOISTURE

Soils differ in their capacity to hold moisture. Heavier clay and silt soils hold more moisture. Sandy soils can lose moisture through leaching as it runs through the root zone and down into the subsoil. Grasses with well-developed, deep root systems add sufficient organic matter to help hold moisture in the soil and thus prevent leaching.

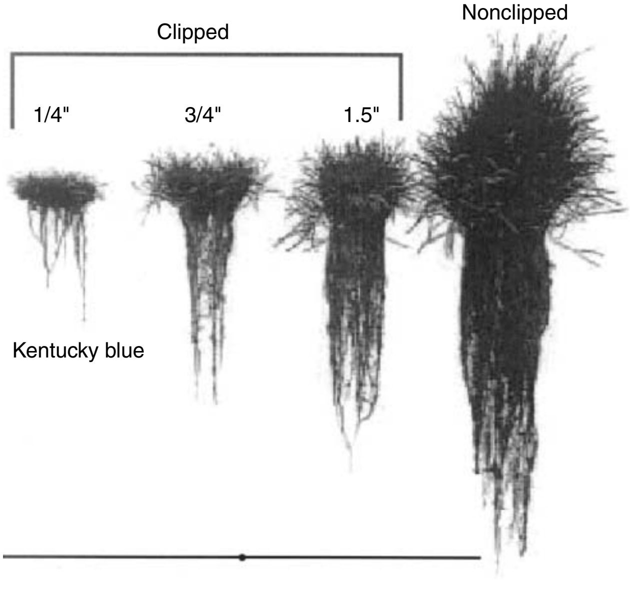

Figure 2. Higher mowing of turfgrass promotes a good root system. For Kentucky bluegrass, 1.5 inches is about right.

THE INFLUENCE OF SOIL TEXTURE ON WATER PENETRATION

The texture of the soil (as determined by the amount of sand, silt and clay) and the amount of thatch (organic deposit between green leaves and roots) influence the speed of water penetration into the soil. In general, heavy soils have many smaller pore spaces and take water in slowly. Sandy soils with fewer but larger pore spaces take moisture more rapidly, unless they are inherently hydrophobic or hard to wet. Soils and thatch that are hard to wet must be watered slowly with small amounts of water applied over longer periods of time to prevent runoff. Sandy soils require less water to penetrate to a given depth. Loam soils need intermediate amounts of water, and clay soils require more water to reach the same depth. As solid particles in the soil decrease on a percentage basis, moisture-holding capacity increases, and soil aeration decreases.

THE DEGREE OF ACIDITY AND ALKALINITY IN SOILS

Soil may either be acid, neutral, or alkaline. Soil pH (the measurement of the degree of acidity and alkalinity) is influenced by soil properties, biological influences, and climatic influences. Under acidic soil conditions, silt and clay particles tend to exist as individual units. Under more alkaline soil conditions, where calcium and magnesium are more plentiful, clay and silt particles group together to form granules. These provide improved soil structure, which results in more favorable balances of air and water in the soil. Where soils are acid and have poor structure, water penetration is much slower.

Soils become acid as carbon dioxide changes to carbonic acid in the soil, acid-reacting fertilizers are used on a continuing basis, or acid rainfall lowers soil pH values. Often a combination of all three of these causes occurs.

In addition to the effect of acidity on physical soil properties, nutrient fixation and availability also are modified, depending on the degree of acidity or alkalinity. For example, phosphates are most available from pH 5.5 to 7.5. Above and below these levels, phosphates are tied up with other minerals, and their availability is reduced. Regular soil tests can determine the need for lime or sulfur or for fertilizer mineral nutrients for specific plant types.

SOIL AS A BIODEGRADABLE AGENT

Biologically healthy soil is the best known medium for decomposing of all sorts of organic compounds, including pesticides and pollutants transported by air and water. These chemicals are known to be biodegradable. This is an ongoing process, which changes these substances into harmless compounds plus carbon dioxide and water. Limited prescribed use of pesticides is not harmful to beneficial soil organisms and should continue to be an important, well-accepted part of plant culture.

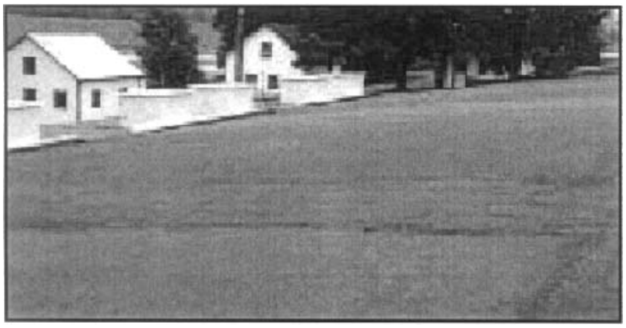

Figure 3. Dr. Thomas Watschke, Pennsylvania State University, created this highly controlled watershed site documenting that established turfgrass has a dramatic, positive effect on reducing nutrient and pesticide pollutants from water runoff.

Water and any pollutants associated with it infiltrate into the ground more quickly on grass-covered soils than on any other surface. Thus, runoff is diminished. Infiltration rates may be as high as 7 inches an hour on sandy soils and as low as 0.10 inch per hour on clay soils. Thus, recharge of purified groundwater is an important benefit. An acre left in open space provides an average of 600,000 gallons of recharge each year in humid regions. Grasses may use up to 10% of the water infiltrated, leaving 90% to recharge the local aquifer (Fig. 3).

GRASS GROUNDCOVER PROVIDES A LIVING MULCH

Good horticultural practice involves using mulches to conserve soil moisture and increase soil productivity. Unlike many landscape plants that are either widely spaced or simply annual in their growth habits, grass groundcover provides a living mulch over the soil surface. This is essentially perennial and provides long lasting soil and water conservation benefits.

Instead of viewing green-lawn groundcovers as static liabilities, these areas can be seen as dynamic, ever-changing populations of plants and animals living within and above the soil. All grasses are natural soil builders. Particularly in residential areas, lawns and landscapes help sustain the soil. Within the soil are large populations of micro- and macroorganisms that are highly competitive. These create a living, moist soil environment best suited to sustaining productive landscape soils, while at the same time purifying our water supply.

WATER SPREADING

Peyman Daneshkar Arasteh
Soil Conservation and
Watershed Management
Research Institute (SCWMRI)
Tehran, Iran

INTRODUCTION

In arid and semiarid regions, irrigation water shortage forces farmers to use floodwater for crop production. Water spreading is a form of runoff farming. It is a form of flood irrigation accomplished by diverting and spreading runoff from sloping areas over adjacent flood plains (1). Water spreading involves the construction of basins, pits, or barriers in or near natural stream channels to impound water and cause it to infiltrate the ground surface rather than leave the basin as surface runoff (2,3).

In other words, diverting or collecting runoff from natural channels, gullies, or streams with a system of dams, dikes, ditches, or other means, and spreading it over relatively flat areas is called water or floodwater spreading (1,4,5).

In a large part of the world, runoff water harvesting is best known and practiced in the semiarid areas (6,7). Runoff farming, which is identical with water harvesting for irrigation purposes, was already used for more than 2000 years. For example, runoff irrigation facilities in the Negev Desert in Israel were built 1300 to 2900 years ago (8). Flood irrigation also has been practiced in other parts of the world, such as Iran, for many years. Traditional flood irrigated farms, called band-sar, can be observed in 400,000 ha of Khorasan, the northeast province of Iran. Band-sar controls about one-fiftieth of 200–250 mm annual precipitation and one-fourteenth of 14,000 MCM annual surface runoff of Khorasan Province, Iran (9). Many sources exist addressing the floodwater spreading affects on soil, groundwater, wild habitats, etc. Baghernejad (10) studied physicochemical and micromorphological changes in topsoil because of flood spreading in Damghan Playa, Iran. Kowsar and Yazdian (11) investigated the fate of the dissolved nitrogen in floodwater that reached the watertable in Bisheh Zard Basin, Iran, because of flood spread over the land. The floodwater was from a geological formation containing NO_3 and NH_4. They found that the NO_3 exceeded the N required by plants of the region (rangelands and forest area), and purification of water must be in order to prevent the groundwater pollution. Nejabat (12) reported that a 2500 ha floodwater spreading system affected about 7500 ha of adjacent area in Gareh Bygon, Iran. He used a technique on the base satellite imagery to determine the increase in cultivated and rangeland areas and the decrease in wind erosion in water spread area.

Floodwater spreading has several advantages that are summarized by FAO (13) as follows:

a. Controlling floodwater damage
b. Reducing erosion and sedimentation
c. Increasing groundwater recharge
d. Improving plant cover and vegetation of the land
e. Increasing economic return for people
f. Increasing rangeland productivity
g. Increasing animal and agricultural production
h. Increasing soil fertility
i. Reclamation of renewable resources
j. Land use change by providing an optimized model for use of flow efficiency land and farming
k. Job creation by fruit trees plantation and increasing the income of local people

l. Improving the ecological balance of the environment
m. Reclamation of the pasture and vegetation cover
n. Providing ground for public participation
o. Providing ground for sustainable development

Floodwater spreading also has some disadvantages (14), listed as follows:

a. Affected by water shortage and climatic risks
b. Affected by farmers' experiences rather than on the scientifically well-established techniques
c. The possible conflicts between upstream and downstream users
d. Possible harm to fauna and flora adapted to running waters and wetlands
e. Requirement of relatively large labor inputs
f. Need to a relatively large natural watershed area
g. Need to appropriate investment and policy reforms to enhance the contribution of rainfed agriculture

In this article, water spreading as an irrigation technique will be introduced.

FLOODWATER SPREADING

Water harvesting refers to methods used to collect water from sources where the water is widely dispersed and quickly changes location or form and becomes unavailable or that is occurring in quantities and at a location where it is unusable unless some intervention is practiced to gather the water to locations where it can provide benefits (15). A number of water harvesting techniques exist that are practiced in many of the water scarce areas of the world, including (15,16):

a. Roof rainwater collection
b. Terracing
c. Small dams
d. Runoff enhancement
e. Runoff collection
f. Water holes and ponds
g. Tanks
h. Floodwater spreading

Floodwater spreading or spate irrigation is one of the common methods of runoff farming in arid and semiarid areas with rainfall intensity more than the infiltration rate of the catchment. USDA emphasized that a major difference between the water spreading system and other surface irrigation techniques, such as border and basin systems, is that the water spreading system is designed to meet precipitation and runoff conditions of the area and apply runoff to cropped fields, whereas other systems are designed to deliver water in accordance with plant needs (17).

Floodwater spreading has at least an inlet to divert flow from natural channels to the system, a delivery channel, and some contour spreader channels to spread water on the ground as overland flow. Figure 1 shows a schematic layout of a floodwater spreading system. Figures 2, 3, 4, 5 show a diversion dam and intake system with sediment control facility, a delivery channel, and the spreaders of flood spreading system, respectively.

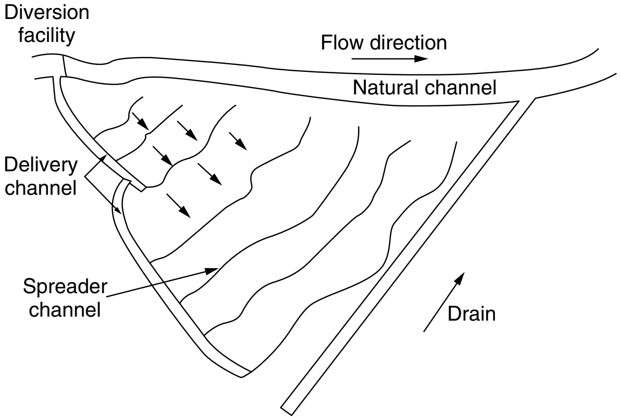

Figure 1. Schematic layout of a floodwater spreading system.

Photo by: R. Hasanzadeh

Figure 2. Diversion dam of Chikhab flood spreading system, Dehloran, Iran.

Photo by: R. Hasanzadeh

Figure 3. Intake and sediment control facilities of Poldasht flood spreading system, Makoo, Iran.

After a flood is diverted from the natural waterway, it reaches the spreading fields by the diversion or delivery channel and contour spreader. Sheet flow over

Photo by: R. Hasanzadeh

Figure 4. Diversion or delivery earthen canal of Sarchahan flood spreading system, Hajiabad, Iran.

Photo by: R. Hasanzadeh

Figure 5. Contour spreader and downstream cultivated field of Poldasht flood spreading system, Makoo, Iran.

the spreading area provides sufficient time to infiltrate more and more water to the soil and completes water holding capacity of the plant root zone. Excess water will be drained to a downstream contour spreader from one or more spillways. Figure 6 shows a schematic flood spreading system with some spillway, and Figs. 7, 8, 9 show three types of spillways, a cemented sandbag, a gabion type, and a cemented rockwork spillway, respectively.

Photo by: R. Hasanzadeh

Figure 7. A cemented sandbag-type spillway of Poldasht flood spreading, Makoo, Iran.

Photo by: R. Hasanzadeh

Figure 8. A gabion-type spillway of Chikhab flood spreading, Dehloran, Iran.

GOVERNING EQUATIONS

Overland flow because of floodwater spreading is on the base of shallow water hydraulics. The general characteristic of shallow water flow (SWF) is that the vertical scale is much smaller than typical horizontal scale, which is true in many situations. SWF surface is nearly horizontal, which allows a considerable simplification in the mathematical formulation and numerical solution by assuming the pressure distribution to be hydrostatic. However, they may not exactly be two-dimensional (2D). The flow exhibits a 3D structure, because of bed friction, just as in boundary layers. Moreover, density stratification, because of a difference in temperature or salinity, causes variations in the third direction. In many SWFs, these 3D effects are negligible, and it is sufficient to consider the depth averaged form, which is 2D in the horizontal plane (18). This restricted form is commonly indicated by the term shallow water equations (SWEs).

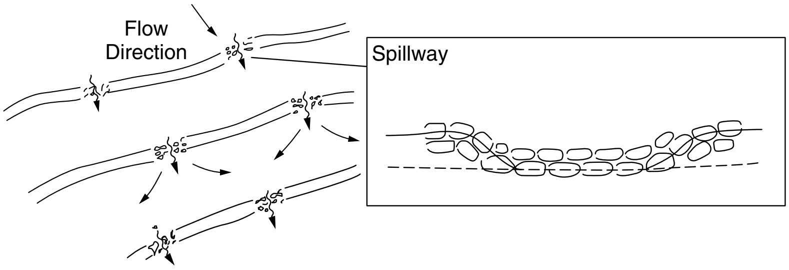

Figure 6. Schematic spillway (16).


Photo by: R. Hasanzadeh

Figure 9. A cemented rock work spillway of Chikhab flood spreading, Dehloran, Iran.

In this article, we limit our attention to incompressible fluid (water) and constant temperature and salinity. By assuming hydrostatic pressure, and with no Coriolis and wind effects and negligible effective stresses, the 2D depth averaged equations of motion are written as (19,20):

Continuity equation:

$$\frac{\partial h}{\partial t} + \frac{\partial (uh)}{\partial x} + \frac{\partial (vh)}{\partial y} = -I \tag{1}$$

x-momentum equation:

$$\frac{\partial (uh)}{\partial t} + \frac{\partial (u^2 h)}{\partial x} + \frac{\partial (uvh)}{\partial y} + g\frac{\partial}{\partial x}\left(\frac{h^2}{2}\right) = gh(S_{0x} - S_{fx}) + \frac{uI}{2} \tag{2}$$

y-momentum equation:

$$\frac{\partial (vh)}{\partial t} + \frac{\partial (v^2 h)}{\partial y} + \frac{\partial (uvh)}{\partial x} + g\frac{\partial}{\partial y}\left(\frac{h^2}{2}\right) = gh(S_{0y} - S_{fy}) + \frac{vI}{2} \tag{3}$$

where h is flow depth; u and v are flow velocities in x and y directions, respectively; I is the infiltration rate; t is time; g is gravitational acceleration; S_{0x} and S_{0y} are ground surface slopes because of x and y directions, respectively; and S_{fx} and S_{fy} are friction slopes in the direction of x and y axes, respectively.

Numerical solutions of SWEs with full or simple forms were described in several articles. Menendez (19), Akanbi and Katapodes (20), and Vreugdenhil (18) introduced numerical methods of solution of 2D SWEs. Arasteh and Vahhadj (21) developed a quasi2D finite difference semianalytical model of kinematic waveform of SWEs for Bisheh Zard flood spreading system, Iran. Arasteh (22) used an implicit finite difference scheme for the mentioned quasi2D model. Arasteh and Banihashemi (23–25) developed a solution method for implicit four points finite difference method to solve 1D kinematic wave SWE, called growing coefficient matrix, and used it for a typical floodwater spreading plot in advance phase of flow to the end of the spreading area.

DESIGN

NRCS (4) introduced the main conditions for applying the floodwater spreading as:

a. Soils have suitable intake rates and adequate water holding capacities for the crop to grow
b. Soils are suitable for production of feed, forage, or grain crops
c. The topography is suitable for the diversion or collection and spreading of water to achieve the desired results
d. Runoff or stream flow is available at the time of the year and in a volume sufficient to increase plant growth
e. Flows can be collected or delivered and spread and exceed water returned without causing excessive erosion
f. Fish and wildlife will not be significantly adversely affected
g. Grazing of the spreading area can be controlled

NRCS (4) also described the water quantity and quality requirement of flood spreading. FAO pointed out that water spreading can be used under 100 to 350 mm annually with a slope of less than 1% and even topography (16). USDA divided floodwater spreading systems into flow-type and detention-type systems (17). The flow-type systems incorporate free drainage from the irrigated area and are divided into spreader-ditch, syrup-pan, and dike-bleeder types. Detention-type systems retain the applied water on the irrigated area until it has infiltrated and are divided into manual and automatic inlet control subsystems. Apart from the type of system, two parameters must be determined as design criteria of the system, design application depth (DAD) and divertible volume of flow rate (DV or DQ). USDA stated that DAD could be determined by a relationship between intake rate, flow duration, or time of concentration, and DV or DQ could be determined as a ratio of volume of storm or peak discharge of upstream catchment.

SUITABLE SITES

http://www.fao.org/docrep/u3160E/33160e00.htm
http://www.fao.org/ag/agl/aglw/wharv.htm
http://www.eng.warwick.ac.uk/ircsa/index.htm

BIBLIOGRAPHY

1. Holtzworth, L. and Lacey, J. (1994). *Species Selection, Seeding Techniques and Management of Irrigated Pastures in Montana and Wyoming*. EB 99. Montana State University, Bozeman, MT.
2. OTA. (1983). *Water Related Technologies for Sustainable Agriculture in US Arid and Semi-Arid Lands*. United States Congress, Office of Technology Assessment, Washington, DC.

3. OTT. (2001). *Technologies Affecting Groundwater*. United States Office of Technology Transfer, Washington, DC.
4. NRCS. (1985). *Waterspreading*. Natural Resources Conservation Service Practice Standard No. 640.
5. Prinz, D., Tauer, W., and Vogtle, Th. (1994). Application of remote sensing and geographic information systems for determining potential sites for water harvesting. *Proc. of the XII CIGR Congress on Expert consultation on Water Harvesting for Improved Agricultural Production*, Washington, DC. 79–87.
6. Prinz, D. (1994). Water harvesting and sustainable agriculture in arid and semi-arid regions. *Proc. of the CIHEAM Conference on Land and Water Resources Management in Mediterranean Region*, Washington, DC. 745–762.
7. Alem, G. (1999). Rainwater harvesting in Ethiopia: an overview. *Proc. of the 25th WEDC Conference on Integrated Development for Water Supply and Sanitation*, Washington, DC. 387–390.
8. Evenari, M., Shanan, L., and Tadmor, N. (1971). *The Negev: The Challenge of a Desert*. Harvard University Press, Cambridge, MA.
9. Arabkhedri, M. and Partovi, A. (1997). Band-sar, a practice in optimum using of water and soil resources in an arid region. *Proc. of the 8th International Conference on Rainwater Catchment System*. 34–40.
10. Baghernejad, M. (1999). Artificial recharge system and the fate of dissolved and suspended particles in floodwater, a case study in Damghan Playa, Iran. *Proc. CD of the 9th International Conference on Rainwater Catchment System*.
11. Kowsar, S.A. and Yazdian, A.R. (1999). Geologic nitrogen in Agha Jari formation of Bisheh Zard Basin: a dilemma. *Proc. CD of the 9th International Conference on Rainwater Catchment System*.
12. Nejabat, M. (1999). Improving environmental characteristics in a wild area around a floodwater spreading system, a case study. *Proc. CD of the 9th International Conference on Rainwater Catchment System*.
13. FAO. (1999). Floodwater utilization: a successful approach towards sustainable development in arid environment. *Proc. of the Conference on the Multifunctional Character of Agriculture and Land*. Available at http://www.fao.org/mfcal/pdf/ft_iran.pdf.
14. Prinz, D. and Malik, A.M. (2002). Runoff farming. Available at http://www.wca-infonet.org/id/81755.
15. Pereira, L.S., Cordery, I., and Iacovides, I. (2002). Coping with water scarcity. IHP technical documents in hydrology. *UNESCO*. **58**: 57–63.
16. Critchley, W., Siegert, K., and Chapman, C. (1991). *Water Harvesting*. FAO, Rome, Italy.
17. Hart, W.E., Collins, H. G., Woodward, G., and Humpherys, A.S. (1983). Design and operation of gravity or surface systems. In: *Design and Operation of Farm Irrigation Systems*. M.E. Jensen (Ed.). pp. 550–559.
18. Vreugdenhil, C.B. (1994). *Numerical Methods for Shallow Water Flow*. Kluwer Academic Publisher, Amsterdam, the Netherlands, p. 261.
19. Menendez, A.N. (1987). On the accuracy of numerical methods for the two dimensional shallow water equation. *Proc. of the XXII Congress of IHAR*. 372–377.
20. Akanbi, A.A. and Katapodes, N.D. (1988). Model for flood propagation in initially dry land. *J. Hydraul. Engin.* **14**(7): 689–706.
21. Arasteh, P.D. and Vahhadj, S.R. (1997). Mathematical modeling of shallow water flow in a flood spreading system. *Proc. of the 8th International Conference on Rainwater Catchment System*. 173–181.
22. Arasteh, P.D. (2002). *Quasi-2D Simulation of Shallow Water Flow in a Flood Spreading System*. Soil Conservation and Watershed Management Research Institute, Research Report No. 77-0510300000-01, p. 80 (in Farsi).
23. Arasteh, P.D. and Banihashemi, M.A. (2003). Numerical solution of kinematic wave equation to simulate overland flow using four-point method with growing coefficient matrix. *Proc. of the 4th Iranian Hydraulics Conference*, Tehran, Iran, 753–760 (in Farsi).
24. Arasteh, P.D. and Banihashemi, M.A. (2003). Application of 1D kinematic wave model to simulate advance phase of overland flow in a flood spreading plot. *Proc. of the 3rd Aquifer Management Congress*, Tehran, Iran, 140–147 (in Farsi).
25. Arasteh, P.D. and Banihashemi, M.A. (2004). Possibility of application of kinematic wave model to simulate overland flow on flood spreading plots. *J. Pajouhesh va Sazandegi.* **61**: 14–23 (in Farsi).

SPRINKLER IRRIGATION

Muhammad Nadeem Asghar
International Water
Management Institute (IWMI)
Lahore, Pakistan

In arid and semiarid environments, supplemental irrigation has become a prerequisite to minimize risks of crop failures in both rain-fed and irrigated agricultural areas. Usually, surface irrigation, in which the entire or most crop area is flooded, is used for irrigating these croplands (Fig. 1). More often, however, water management skills on the farm are lacking, water supplies may be uncertain, and so irrigation efficiency tends to below.

Because of the present scarce global water resources, improving irrigation efficiency—getting more crop per drop—and making the best use of water for agriculture are prerequisites for the future (1). Shifting from surface irrigation to sprinkler irrigation, which is a method of irrigation under pressure in which water is sprinkled

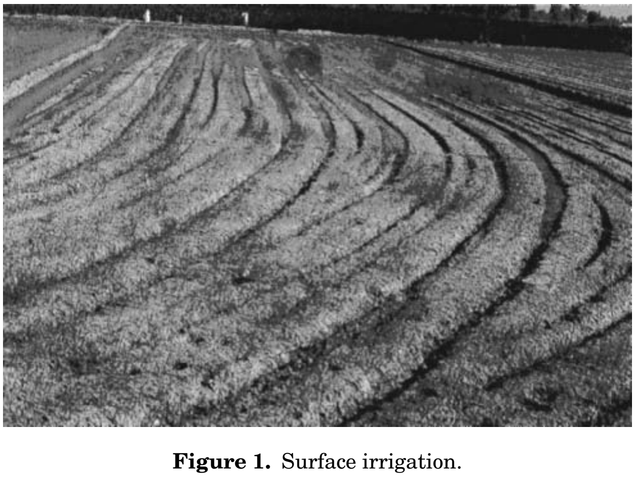

Figure 1. Surface irrigation.

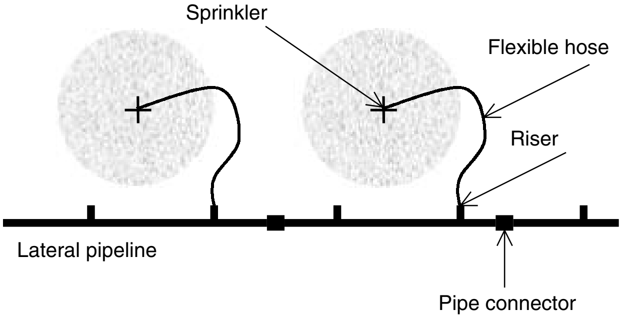

Figure 2. Lateral pipeline.

Figure 4. Hose-move sprinkler irrigation.

to imitate rainfall, will contribute substantially to these goals. In contrast to surface irrigation, sprinkler irrigation can help to irrigate frequently with just the right amount of water to avoid crop distress and to avoid overirrigation (Figs. 2 and 3).

In recent decades, various sprinkler irrigation methods and installations have been developed to meet the needs of farmers, particularly smallholders. Hand-move sprinkler irrigation is the most widely adopted and least expensive system operated with pressure ranges from 2.0–3.5 bars. The sprinklers are mounted on lateral pipelines temporarily laid across a field. To reduce labor requirements, this system was modified to hose-move sprinkler irrigation (Fig. 4). In this modified system, lateral pipelines are placed permanently, and the sprinklers, mounted on tripod stands, are not fitted directly to the lateral pipelines but are connected to them via flexible hoses. The hoses and sprinklers can be moved laterally on either side to cover a number of lateral positions (2).

Generally, in sprinkler irrigation, the uniformity of distribution is more sensitive to wind. However, this irrigation provides 'air-conditioning' effects on the crops if used in anti frost sprinkling and can also reduce pesticide applications if used as a spray. It is also fairly well established that a plan of heavy sprinkling at infrequent intervals will ordinarily produce a deeper root structure and thus better crop growth and yields than frequent, light sprinkling. However, light irrigation may be beneficial when there is a minimum crop water requirement. Under shallow water table conditions, when groundwater contributes significantly to meet the crop water requirement, light irrigation is also required to dilute the salinity in the effective root zone. Sprinkler irrigation may be used to germinate and establish salt-sensitive crops by keeping the seedbed adequately moist and salt-free by uniformly applying small depths of irrigation water.

Figure 3. Hand-move sprinkler irrigation.

When sprinkler irrigation is done with saline water, it brings water into contact with foliage, and salt deposits on leaves may adversely affect some crops. However, the susceptibility of plants to leaf injury from saline sprinkled water depends on leaf characteristics that affect the rate of absorption and is not generally correlated with tolerance to soil salinity. The degree of spray injury varies with weather conditions, especially the water deficit of the atmosphere. Visible symptoms may appear suddenly following irrigation when the weather is hot and dry. Increased frequency of sprinkling, in addition to increased temperature and evaporation, increases the salt concentration in the leaves due to adsorption and results in leaf damage. However, irrigation at night (or any other low evaporation period) minimizes the salt concentration in the leaves due to adsorption (3). Furthermore, evaporation losses during sprinkling can also be minimized by avoiding midday irrigation.

High capital costs, nonavailability of needed equipment in the local market, energy dependence, and lack of local expertise to operate and maintain sophisticated hydraulic equipment are the main constraints that hinder the acceptability of the system by the farmers, particularly smallholders. For the same reasons, surface irrigation is still likely to be dominant in the future, even though it wastes water and is a major cause of waterlogging and salinization.

However, the integration of surface irrigation and sprinkler irrigation could help in controlling salinity in the root zone with minimum application of irrigation water by improving leaching efficiency. In this practice, sprinkler irrigation helps in uniform distribution of groundwater, and surface wetness due to sprinkling that helps in minimizing the application of surface water still can meet the crop water requirement. This surface wetness also allows salts to dissolve in the root zone, which improves the leaching efficiency of surface irrigation. Therefore, the value of sprinkler irrigation can still be promised by promoting it together with surface irrigation (4).

BIBLIOGRAPHY

1. FAO. (2002). *Crops and Drops—Making the Best Use of Water for Agriculture*. On-line Publications, Food and Agriculture Organization of the United Nations, Rome.
2. Phocaides, A. (2000). *Technical Handbook on Pressurized Irrigation Techniques*. On-line Publications, Food and Agriculture Organization of the United Nations, Rome.
3. Kruse, E.G. (1995). On-farm irrigation and drainage practices. In: *Agricultural Salinity Assessment and Management*. K.K. Tanji (Ed.). ASCE, New York, pp. 349–371.
4. Asghar, M.N. et al. (2002). Helping Measures for Farmers to Mitigate Drought in Canal Water Scarce Areas of the Indus Basin, Pakistan. *Proc. Natl. Symp. Drought Water Resour. Pakistan*, Center of Excellence in Water Resources Engineering, UET, Lahore, pp. 63–73.

STOMATES

R. Andrés Ferreyra
Ag Connections, Inc.
Murray, Kentucky

Stomates or stomata (singular: stomate or stoma) are structures on the surfaces of plant leaves that allow gas exchange (transpiration, i.e., loss of water vapor, uptake of CO_2, and emission or uptake of O_2) between the interior of a leaf and the atmosphere.

Plant leaves are covered by a waxy layer, the cuticle, that prevents water loss from the cells on the leaf surface, or epidermis. The cuticle does not completely cover the leaves, however; it is interrupted by microscopic pores, surrounded by pairs of specialized guard cells (Fig. 1a). A stoma is a unit composed of a pore and its guard cells. Guard cells, unlike regular epidermal cells, usually contain chloroplasts (photosynthesizing organelles containing chlorophyll). In some plants, there are also additional, specialized subsidiary cells that differ in shape from regular epidermal cells and participate in the osmotic changes that drive guard cell movement.

Stomata can be found in all aboveground parts of plants but are more frequent in leaves. They usually appear on one (the lower) side of leaves but may appear on both. When stomata occur on the lower (abaxial) leaf surface, the plant is said to be hypostomatous; hyperstomatous plants have stomata only on the upper (adaxial) leaf surface, and plants with stomata on both surfaces are called amphistomatous. In this latter case, abaxial stomata are generally more numerous. The number of stomata per unit area is called the stomatal density. This figure may vary when plants are grown under different environmental conditions such as atmospheric CO_2 concentration.

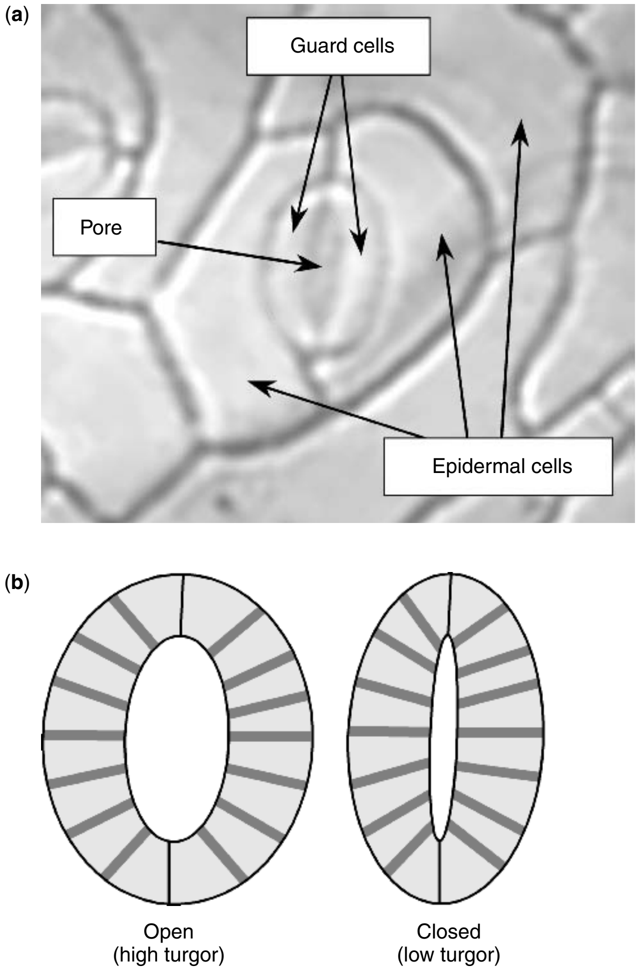

Figure 1. (**a**) Micrograph of a peanut (*Arachis hypogaea* L.) epidermis, showing the guard cells, the stomatal pore, and epidermal cells: (**b**) Diagram of open and closed stomata, showing cellulose microfibrils.

The size and density of stomata vary widely among species (Table 1). When fully open, stomatal pores are 3–12 μm wide and from 10 to more than 30 μm long. Species with smaller stomata tend to have a higher density that can range from 60–80 per mm^2 in corn to 150 in legumes such as alfalfa and clover, 300 in apple, and more than 1000 in scarlet oak (1). Dicotyledon leaves with a netted venation pattern generally have a scattered arrangement of stomata, whereas monocotyledons and conifers tend to have their stomata arranged parallel to the main axis of leaves.

Some plants locate their stomata at the bottom of recessed hollows and sometimes grow hair-like structures above them, apparently to limit water loss by creating a high-humidity environment above the pore. Most plant leaves have small open spaces under the pore, called substomatal cavities (Fig. 2).

STOMATAL CONDUCTANCE MODEL

Transpiration and its dependence on stomatal control have been traditionally described by using an electric analog model based on conductances or their reciprocal, resistances. The basic assumption of the conductance

Table 1. Occurrence and Dimensions of Stomata for Several Representative Species[a]

			Stomatal Density, Pores/mm^2		Guard Cell Dimensions, μm, for each Leaf Surface				Dimensions of Stomatal Pore, Abaxial, μm	
					Adaxial		Abaxial			
Plant Type	Representative Species	Common Name	Upper	Lower	Length	Width	Length	Width	Length	Depth
Moss	*Polytrichum commune*	Hair cap moss	16	16	46	15	46	15	15	12
Fern	*Osmunda regalis*	Royal fern	0	67			56	19	30	15
Conifer	*Pinus silvestris*	Scotch pine	120	120	28	7	28	7	20	6
Dicot tree	*Tilia europea*	European linden	0	370			25	9	10	8
Dicot herb	*Helianthus annuus*	Sunflower	120	175	35	13	32	14	17	
Dicot herb (xerophyte)	*Sedum spectabilis*	Autumn joy	35	56	32	10	33	10	20	18
Monocot herb	*Allium cepa*	Onion	175	175	42	19	42	19	24	18
Monocot C_3 grass	*Avena sativa*	Oats	50	45	52	15	56	13	20	10
Monocot C_4 grass	*Zea mays*	Corn	98	108	38	10	43	12	20	10

[a] Adapted from Reference 1.

model is that the diffusion of gases within plant tissues and in the atmospheric boundary layer is restricted by resistances and that the resulting system of gas fluxes into and out of leaves can be described by the equations developed for analyzing electrical circuits (Fig. 3).

Nobel (2) described the diffusion of gases along stomatal pores by applying Fick's law and a stomatal conductance (which we will denote as $g_{\mathrm{w}}^{\mathrm{st}}$). This is a widely used parameter in modeling applications; stomatal conductance, the rate of flow per unit area of the leaf, is considerably easier to measure than the flux density within a stomatal pore. Stomatal conductance is closely related to stomatal aperture; Jones (3) showed how to estimate conductance from observed stomatal dimensions. Given the dynamic behavior of stomatal aperture, however, conductance is highly variable, especially in natural conditions. Time-averaged values are typically used in calculations.

The cuticular conductance, $g_{\mathrm{w}}^{\mathrm{cut}}$, is probably the most uncertain of all conductances in this model. Traditionally it has been assumed to be negligible, but experimental data show that cuticular transpiration varies greatly among species. Larcher (4) generalized that it constitutes 10–33% of total transpiration, but some researchers presented even higher values (5).

Nobel proposed calculating $g_{\mathrm{w}}^{\mathrm{ias}}$, the conductance of intercellular air spaces, based on the average distance between mesophyll cell surfaces within a leaf and the inner surface of stomatal pores. Consistent application of the conductance model would thus require measurements on leaf cross-sectional images. Parkhurst (6) postulated that intercellular air space is a very important, usually ignored, component of the leaf gas exchange pathway, although it is more relevant for photosynthesis than for transpiration.

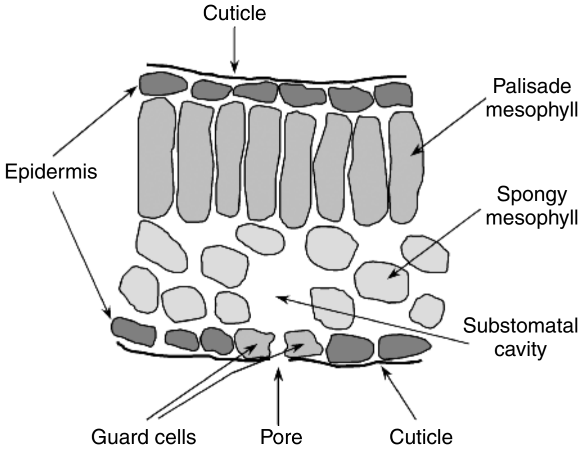

Figure 2. Diagram of a (hypostomatous) leaf cross section, showing different tissues and cell types.

Calculations with the electric analogy are simplified by replacing all conductances with resistances, taking $r_{\mathrm{w}}^{\mathrm{i}} = \left(g_{\mathrm{w}}^{\mathrm{i}}\right)^{-1}$, where the superscript "i" stands for "bl," "st," "cut," or "ias" (boundary layer, stomata, cuticle, and intercellular air space, respectively). Figure 3 presents the corresponding equivalent electrical circuit. The cuticular resistance acts as a resistance in parallel with the series sum of the stomatal and intercellular space resistances. According to this scheme, the total leaf resistance to water movement can be described as

$$r_{\mathrm{w}}^{\ell} = \frac{\left(r_{\mathrm{w}}^{\mathrm{ias}} + r_{\mathrm{w}}^{\mathrm{st}}\right) \cdot r_{\mathrm{w}}^{\mathrm{cut}}}{r_{\mathrm{w}}^{\mathrm{ias}} + r_{\mathrm{w}}^{\mathrm{st}} + r_{\mathrm{w}}^{\mathrm{cut}}} + r_{\mathrm{w}}^{\mathrm{bl}}$$

and the total leaf conductance is

$$g_{\mathrm{w}}^{\ell} = \left(r_{\mathrm{w}}^{\ell}\right)^{-1}$$

Nobel (2) provided some representative conductance values: for crop plants such as beet, spinach, tomato, and pea, $g_{\mathrm{w}}^{\mathrm{ias}}$ ranges from 800 to 4000 mmol/m^2/s (20 to 100 mm/s); for open stomata, $g_{\mathrm{w}}^{\mathrm{st}}$ ranges from 200 to 800 mmol/m^2/s (5 to 20 mm/s). These values are somewhat lower for other agricultural crops.

THE STUDY OF STOMATA

Malpighi saw pores on leaf surfaces in 1674. Grew showed them in his anatomical work of 1682. The term "stomata"

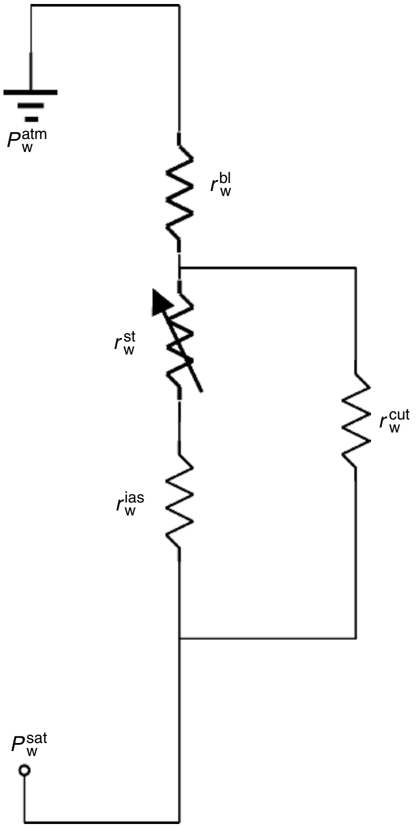

Figure 3. Conductance model of leaf water vapor movement, expressed in resistance form using electrical symbols. P_w^{atm} and P_w^{sat} are the water vapor pressure in the atmosphere and on the surface of mesophyll cells, respectively. The r_w^i are the resistances to water vapor movement in the boundary layer (i = "bl"), in stomata ("st"), through the cuticule ("cut"), and in the intercellular air spaces ("ias").

may have been first used by A. de Candolle in 1827, and research on stomatal behavior began earnestly in the mid-1900s (1). Gas exchange through small pores such as stomata was explained formally in physical terms by Brown and Escombe (7). Later, Stålfelt (8) and Bange (9) showed how, in moving air (when boundary resistance is low), transpiration and stomatal aperture are closely correlated. More recently, Zeiger (10) reviewed the multiple environmental influences on stomatal behavior; the review by Outlaw (11) shows the current state of the art.

The crucial role of stomata in controlling plant CO_2 uptake and water loss has motivated the search for methods of measuring stomatal aperture and conductance. Weyers and Meidner (12) extensively reviewed these techniques. A few salient methods are discussed briefly below.

Microscopy

In some plants, pieces of epidermis can be peeled off the leaf and observed on the microscope after fixing in alcohol. However, it is very difficult to peel the epidermis from some plants, and the process can also affect stomatal aperture. An alternative is to make impressions of the epidermis while the leaves are still on the plant, using substances such as collodion, silicone rubber, dental paste, or nail polish. These impressions are later peeled off and examined on the microscope (1).

The above mentioned methods involving epidermal manipulation may not be fully representative of *in vivo* conditions because the material used to make the impression may invade the pore or stomatal aperture may change during the process. There is evidence that ultrafast freezing can prevent aperture changes (13). This, coupled with scanning electron microscopy, can capture stomatal characteristics in great detail (e.g., 14).

Anatomically Explicit Models

Most implementations of conductance models assume that diffusion paths for both CO_2 and water vapor are identical, thus making transpiration and photosynthesis proportional to each other as stomatal aperture changes. This assumption ignores the intracellular, liquid-phase components of the CO_2 pathway, although Nobel (2) postulated that liquid-phase conductance values are significant.

Parkhurst (6) criticized conductance models "in part because they are usually one-dimensional representations, but also because they treat continuously interacting processes as if they were sequential." The lumped-parameter conductance approach, as opposed to using spatially distributed parameters and processes, may introduce error even in the apparently simple simulation of diffusion within intercellular spaces, especially when estimating photosynthesis: assimilation really occurs throughout the tortuous intercellular space concurrently with a gradual fall in CO_2 concentration, rather than lumped at the end of the flow path as in typical conductance models.

However, lumping the properties of the leaf interior was unavoidable until recent developments in numerical methods made it possible to solve diffusion equations in very complex domains. Parkhurst (6) proposed a complex distributed electrical analog model, and Pachepsky and Acock (15) created a very sophisticated finite-element model, 2DLEAF, that explicitly represents leaf tissues in a two-dimensional domain based on measurements made on micrographs of epidermal peels and leaf cross sections. Driven by environmental variables such as photon flux density, air and leaf temperature, gas concentrations in the atmosphere, and a novel form of parameterizing the leaf boundary layer (16), the model then simulates the fluxes of water vapor, CO_2, and O_2. Calculations with 2DLEAF have shown that distributions of water vapor and CO_2 concentrations in the leaf interior differ from each other, that photosynthesis and transpiration rates are not proportional to each other, and that stomatal aperture affects transpiration more than it affects photosynthesis (15); this confirms the statement by Parkhurst (6) that low stomatal conductance reduces water loss more than it reduces CO_2 uptake.

Infiltration

Another experimental technique for estimating stomatal aperture involves infiltrating graded solutions of different viscosity into leaves. The viscosity of the most viscous solution that can infiltrate the pores serves as a measure of aperture (17). However, this method does not allow calculating stomatal resistances, and, due to the differences in stomatal anatomy among species, it is better suited to relative measurements of stomatal aperture within the same species (3).

Mass or Viscous-Flow Porometers

The combined resistance of the stomata and the intercellular air spaces of the mesophyll can be estimated for an amphistomatous leaf by forcing pressurized air through the leaf from one surface to the other and then measuring the airflow rate (or the time needed for a specified drop in pressure). The flow rate is inversely proportional to the viscous-flow resistance.

Many such mass- or viscous-flow measuring devices, or porometers, have been described in the literature (18–20). However, absolute calibration is difficult because the flow path inside the leaf is complex and the series circuit measured by the instrument (air flows in through one leaf surface, the mesophyll, and out the other surface) is not representative of the actual water vapor diffusion paths, where the stomata of the two leaf surfaces are in parallel (3). Additionally, mass flow porometer designs have been cumbersome (albeit ingenious), and contemporary interest in separating the adaxial, abaxial, and mesophyll resistance components has decreased interest in mass flow instruments in favor of the diffusion porometer.

Diffusion Porometers

More advanced porometers generally include small cuvettes or chambers that can be attached to leaves and either measure the time necessary for a specified change in the humidity in the chamber (transit-time diffusion porometers; 21) or measure the water vapor concentration differential between the air leaving the chamber and the air entering it (continuous-flow diffusion porometers; 22). They can work with one or both leaf surfaces.

Diffusion porometers are powerful, albeit expensive. They are portable, operate rapidly (seconds to a few minutes), and store data in an onboard memory for subsequent downloading to a personal computer. Although they can work with individual leaf surfaces (unlike mass-flow porometers), they must still average over a relatively large leaf area and thus cannot describe the behavior of individual stomata. They are also affected by any factor that influences water loss (e.g., cuticle conductance), not just stomatal conductance.

Jones (3) suggested the following equation to estimate leaf conductance from a continuous-flow porometer:

$$g_{\mathrm{w}} = (r_{\ell\mathrm{w}} + r_{\mathrm{aw}})^{-1} = \frac{u_{\mathrm{e}}(x_{\mathrm{w0}} - x_{\mathrm{we}})[1 - (x_{\mathrm{ws}} + x_{\mathrm{w0}})/2]}{A(1 - x_{\mathrm{w0}})(x_{\mathrm{ws}} - x_{\mathrm{w0}})}$$

where g_{w} is the total conductance to water vapor (mol/m^2/s); $r_{\ell\mathrm{w}}$ is the leaf (combined stomata, cuticle, intercellular air space, and cell wall) resistance to water vapor (m^2 s/mol); r_{aw} is the boundary layer resistance to water vapor, typically specified by the porometer manufacturer; u_{e} is the flow rate entering the chamber (mol/s); x_{we} is the mole fraction of water vapor in the air entering the chamber (mol/mol); x_{w0} is the mole fraction of water vapor in the air leaving the chamber (mol/mol); x_{ws} is the saturation mole fraction of water vapor at the chamber temperature (mol/mol), where the relative humidity of the outlet air is equal to $x_{\mathrm{w0}}/x_{\mathrm{ws}}$; and A is the surface area of the leaf exposed in the chamber (m^2).

STOMATAL FUNCTION

Stomata open and close due to changes in the shape of guard cells. Changes in the concentration of ions inside guard cells make water move into or out of the cells osmotically. The corresponding changes in hydrostatic pressure (turgor pressure, the pressure exerted by the fluid in the cell that presses the plasma membrane against the cell wall) in turn make the cells swell or shrink. Two additional factors influence the change in shape; first, guard cell walls have a set of cellulose microfibrils arranged radially around the cell (radial micellation). When the cell volume tends to increase with turgor pressure, the microfibrils force a lengthening of the cells rather than an increase in diameter. Additionally, the inner walls of guard cells, those adjacent to the pore, are more rigid than the outer walls. Thus, a lengthening of the cell translates into an increase in its curvature, and consequently into an opening of the stoma (Fig. 1b).

The intake of ions for stomatal opening occurs through the guard cell membrane. A special form of enzyme called proton ATPase removes protons from the inside of the cell. This creates an electrical gradient across the membrane; potassium ions can move into the cell (through specialized ion channel proteins embedded in the membrane) driven by this gradient. Chlorine and calcium ions, as well as starch, play additional roles in this process. For further reading, Nicholls and Ferguson (23) presented an excellent introduction to ATPases and active transport through membranes; Mansfield et al. (24) discussed the role of calcium.

Stomatal aperture responds to numerous environmental factors such as light, temperature, relative humidity, CO$_2$ concentration, and pollutants. It also depends on internal variables such as water stress, periodic metabolic oscillations (circadian rhythms), and the concentration of signaling substances such as abscisic acid (ABA) and cytokinins. Outlaw (11) and Zeiger (10) explained all of this in detail; a brief outline is given below.

In general, stomata tend to open in the light, when there is energy available for photosynthesis. An interesting exception is that of desert succulent plants of the crassulacean acid metabolism (CAM) type. They open their stomata at night (when atmospheric water demand is low), store CO$_2$ as malic acid, and then convert that into sugars via photosynthesis during the day with their stomata closed and the plant safe from water loss.

Stomatal aperture is generally inversely proportional to CO$_2$ concentration. Some researchers believe that plants sense the concentration of CO$_2$ in the leaf interior and respond to that; others believe that the plant responds to the ratio of internal to external CO$_2$ concentrations.

Stomata tend to close at midday, suggesting that plants can either sense atmospheric humidity, or sense transpiration rate and react to limit water loss. Temperature also affects stomatal aperture differently for different species; in general, temperature extremes close the stomata. However, it is difficult to separate the high temperature effect from that of a high vapor pressure deficit (i.e., high water demand). Wind effects are complex. On one hand, wind reduces boundary layer resistance and thus tends to increase transpiration. On the other, there are reports of plants closing their stomata in response to wind or vibration.

ABA is a plant hormone believed to be the main signaling mechanism for stomatal control. Roots subjected to dry soil (or hypoxia due to flooded soil) produce ABA, which moves upward to the leaves through plant vascular tissues. ABA interacts with guard cells, allowing the membranes to leak potassium ions out of the cell, reducing its concentration gradient, which is accompanied by a loss of water and turgor pressure and subsequently, stomatal closure. Tal and Imber (25) presented an interesting experimental confirmation of the central role of ABA in stomatal control. They showed that mutants that cannot synthesize ABA show permanent wilting because they cannot close their stomata to stop water loss. When ABA is applied to these plants from an external source, however, the stomata close, turgor pressure is restored, and wilting stops.

STOMATAL REGULATION

Stomatal aperture provides plants with a control mechanism with which they can balance the entry of the CO_2 used for photosynthesis and the water lost by diffusion through the pore from evaporation from the surfaces of cells inside the leaf (2). Transpiration is usually considered an inevitable, undesirable side effect of CO_2 uptake. However, transpiration does have important roles: transpiration provides energy to transport water (and hence, nutrients and signaling substances) in the plant, and it also dissipates the heat of direct sunlight (by evaporative cooling).

Many ideas have been proposed to explain stomatal behavior in terms of optimization of cost (transpiration) and benefits (photosynthesis). Optimally, stomatal aperture would vary throughout the day such that there is maximum photosynthesis and minimum transpiration. This implies a complex control system that monitors several input variables (several of which were mentioned in the previous section). Jones (3) and Nobel (2) provided interesting conceptual discussions in terms of feedback and feedforward systems. The latter explains the popular hypothesis that stomata behave so that the sensitivity of photosynthesis and transpiration to stomatal aperture (i.e., the corresponding partial derivatives with respect to stomatal aperture) are a constant multiple of each other. Cowan (26) and Farquhar and Sharkey (27) discussed this and other optimization strategies in greater detail.

BIBLIOGRAPHY

1. Kramer, P.J. and Boyer, J.S. (1995). *Water Relations of Plants and Soils*. Academic Press, London.
2. Nobel, P.S. (1983). *Biophysical Plant Physiology and Ecology*. W.H. Freeman, New York.
3. Jones, H.G. (1992). *Plants and Microclimate*, 2nd Edn. Cambridge University Press, Cambridge, UK.
4. Larcher, W. (1995). *Physiological Plant Ecology*, 3rd Edn. Springer, New York.
5. Antipov, N.I. (1971). Correlation between cuticular and stomatal transpiration in woody and bushy plants. *Sov. Plant Physiol.* **18**(6):1026–1030.
6. Parkhurst, D.F. (1994). Diffusion of CO_2 and other gases inside leaves. Tansley Review No. 65. *New Phytol.* **120**: 449–479.
7. Brown, H.T. and Escombe, F. (1900). Static diffusion of gases and liquids in relation to the assimilation of carbon and translocation of plants. *Phil. Trans. R. Soc. London Ser. B* **193**: 223–291.
8. Stålfelt, M. (1932). Der stomatäre regulator in der pflanzlichen transpiration. *Planta* **17**: 22–85.
9. Bange, G.G.J. (1953). On the quantitative explanation of stomatal transpiration. *Acta Bot. Neerl.* **2**: 255–297.
10. Zeiger, E. (1983). The biology of stomatal guard cells. *Annu. Rev. Plant Physiol. Plant. Mol. Biol.* **34**: 441–475.
11. Outlaw, W.H. (2003). Integration of cellular and physiological functions of guard cells. *Crit. Rev. Plant Sci.* **22**(5):503–529.
12. Weyers, J.D.B. and Meidner, H. (1990). *Methods of Stomatal Research*. Longman, Scientific & Technical, Harlow, UK.
13. van Gardingen, P.R., Jeffree, C.E., and Grace, J. (1989). Variation in stomatal aperture in leaves of *Aven fatua* L. observed by low-temperature scanning electron microscopy. *Plant Cell Environ.* **12**: 887–898.
14. Shiraishi, M., Hashimoto, Y., and Kuraishi, S. (1978). Cyclic variations of stomatal aperture observed under the scanning electron microscope. *Plant Cell Physiol.* **19**: 637–645.
15. Pachepsky, L.B. and Acock, B. (1996). A model 2DLEAF of leaf gas exchange: Development, validation, and ecological application. *Ecol. Modelling* **93**: 1–18.
16. Pachepsky, L.B., Ferreyra, R.A., Collino, D., and Acock, B. (1999). Transpiration rates and leaf boundary layer parameters for peanut analyzed with the two-dimensional model 2DLEAF. *Biotronics* **28**: 1–12.
17. Hack, H.R.B. (1974). The selection of an infiltration technique for estimating the degree of stomatal opening. *Ann. Bot.* **38**: 93–114.
18. Darwin, F. and Pertz, D.F.M. (1911). On a new method of estimating the aperture of stomata. *Proc. R. Soc. London Ser. B* **84**: 136–154.
19. Meidner, H. (1992). Developments in mass flow porometry. *J. Exp. Bot.* **43**: 1309–1314.
20. Shimshi, D. (1977). Fast-reading viscous flow porometer. *New Phytol.* **78**(3):593–598.
21. Kanemasu, E.T., Thurtell, G.W., and Tanner, C.B. (1969). Design, calibration, and field use of a stomatal diffusion porometer. *Plant Physiol.* **44**: 881–885.
22. Beardsell, M.F., Jarvis, P.G., and Davidson, B. (1972). A null-balance diffusion porometer suitable for use with leaves of many shapes. *J. Appl. Ecol.* **9**: 677–690.
23. Nicholls, D.G. and Ferguson, S.J. (1992). *Bioenergetics 2*. Academic Press, San Diego, CA.
24. Mansfield, T.A, Hetherington, A.M., and Atkinson, C.J. (1990). Some current aspects of stomatal physiology. *Annu. Rev. Plant Phys. Plant. Mol. Biol.* **41**: 55–75.
25. Tal, M. and Imber, D. (1971). Abnormal stomatal behavior and hormonal imbalance in flacca, a wilty mutant of tomato.

III. Hormonal effects on the water status in the plant. *Plant Physiol.* **47**: 849–850.

26. Cowan, I.R. (1982). Regulation of water use in relation to carbon gain in higher plants. In: *Encyclopedia of Plant Physiology*. Vol. 12B. O.L. Lange and J.D. Bewdey (Eds.). Springer-Verlag, Berlin, pp. 535–562.
27. Farquhar, G.D. and Sharkey, T.D. (1982). Stomatal conductance and photosynthesis. *Annu. Rev. Plant Physiol.* **33**: 317–345.

CROP WATER STRESS DETECTION USING REMOTE SENSING

M. Susan Moran
USDA ARS Southwest
Watershed Research Center
Tucson, Arizona

Pablo J. Zarco-Tejada
Instituto de Agricultura
Sostenible (IAS-CSIC)
Córdoba, Spain

Thomas R. Clarke
USDA ARS U.S. Water
Conservation Laboratory
Phoenix, Arizona

Remote sensing in the optical spectrum has potential to detect plant manifestations of both transient and chronic crop water stress. Crop water status and evaporation rate have been directly related to thermal indexes, such as the crop water stress index (CWSI), the water deficit index (WDI), and the thermal kinetic window (TKW) index. Measurements of surface reflectance in narrow visible and near-infrared wavelengths have potential for monitoring crop photosynthesis and fluorescence. Surface reflectances in wide bands of the visible, near-IR, and short-wave IR spectrum have been used to detect plant manifestations of chronic crop water stress, such as phenological adaptations and leaf expansion and loss. To apply these approaches with satellite-based sensors for farm-scale irrigation scheduling and crop management, there are still needs for improved sensor design, timely image delivery, and reduced image cost.

INTRODUCTION

Numerous studies have shown that crop water stress has a direct effect on crop growth, development, yield, and ultimately, on farmers' profits. New management techniques in arid and semiarid regions are designed for controlled deficit irrigation to ensure low water loss without yield reduction. To achieve this delicate balance between water use and crop yield, farm managers need an operational means to quantify plant water deficit and evaluate the effects of stress on a given crop species at any stage of development. The following sections present the plant manifestations of water stress that can be detected using passive remote sensing methods with sensors that acquire data in the visible (VIS), near-infrared (NIR), short-wave infrared (SWIR), and thermal infrared (TIR).

CROP ADAPTATIONS TO WATER STRESS

Turner (1) provided a good synthesis of the effects of water deficit on crop plants and physiological adaptations to transient stress. He identified crop physiological adaptations associated with three types of drought resistance—drought escape, drought tolerance at high water potential, and drought tolerance at low water potential—where "drought" was defined as a period without significant rainfall. The following discussion focuses on the physiological adaptations that might have the greatest effect on the reflective and thermal spectral regions and on the optical properties of plants that would allow stress detection by remote sensing.

Drought escape is the ability of the plant to complete its life cycle before serious soil and plant water deficits develop. For example, studies have shown that wheat can hasten maturity in response to mild water deficits at the critical time between flowering and maturity. When stress occurred between floral initiation and wheat ear emergence, the number of wheat tillers that produced ears was less than that under well-watered conditions.

Drought tolerance at high tissue water potential is sometimes referred to as drought avoidance; it allows plants to endure drought periods while maintaining high plant water status. One such crop adaptation is the reduction of water loss through increased stomatal and cuticular resistance. This is expressed in increased epidermal waxes of leaves and a reduction in general plant productivity. Another adaptation is to reduce the radiation absorbed by the plant through leaf movement (e.g., leaf cupping, paraheliotropism, or wilting) or increased leaf pubescence and waxiness. Drought tolerance is also achieved by reduction of leaf area through decreased leaf expansion, reduced tillering and branching, and leaf shedding.

Water stress is often expressed by the inability of some crop species to maintain cell turgor at low water potentials. The capability of crops to control cell turgor varies widely by species (1). These variations in cell turgor under drought conditions can influence the closure of stomatal apertures and the rates of photosynthesis, evaporation, and leaf expansion.

Optical Properties of Plant Leaves

A great deal has been written on the optical properties of crop leaves in the VIS, NIR, SWIR, and TIR domains (2). In the visible domain, leaf reflectance is affected mainly by leaf pigments such as chlorophyll, xanthophyll, carotenoids, and anthocyanins, as well as by the leaf structure and the dry matter content. In the NIR, reflectance depends on the anatomic structure of the leaves and increases with the number of cell layers and the size of the cells. In the SWIR, reflectance is mainly affected by the leaf water content; strong water absorption bands are at 1.45, 1.95, and 2.5 μm. It is generally reported that leaves under water stress show a decrease in reflectance in the NIR spectrum, a reduced red absorption in the chlorophyll active band (0.68 μm), and a consequent blue shift of the red edge (3). In the TIR, there is a direct link between the process of plant water evaporation and the plant thermal response (i.e., water evaporates and cools the leaves).

The theory to explain these leaf optical properties and the effect of leaf constituents on leaf reflectance and transmittance was developed into a leaf model by Allen and Richardson (4). The PROSPECT model (5), based on a plate model for multiple layers, is widely used to simulate leaf optical properties as function of leaf chlorophyll concentration (C_{ab}), equivalent water thickness, dry matter content, and leaf internal structure; it enables simulation of leaf reflectance and transmittance as a function of different stresses.

Optical Properties of Plant Canopies

Though knowledge of the optical properties of individual leaves contributes to our understanding of the processes involved, field studies have shown that leaf optical properties alone are not sufficient to understand canopy-level reflectance. Canopy reflectance is a function of the leaf optical properties previously described, plus viewing and illumination geometry, canopy structure, and soil reflectance. For example, Guyot et al. (6) reported that the most important factor that influences the reflectance of a plant canopy is its geometric structure, not its leaf reflectance. Jackson and Ezra (7) concluded that water stress-induced changes in visible, NIR, and SWIR reflectance of a cotton canopy were due largely to canopy geometry changes rather than leaf physiological/anatomic changes in all but the red spectral band.

Furthermore, Myers et al. (8) identified seven parameters (in addition to leaf reflectance) that determined crop canopy reflectance, of which only the first three could be directly related to crop stress: (1) transmittance of leaves, (2) amount and arrangement of leaves, (3) characteristics of other components of the vegetation canopy, (4) characteristics of the background, (5) solar zenith angle, (6) look angle, and (7) solar azimuth angle. Canopy reflectance models enable the simulation of these canopy structural characteristics and viewing geometry on canopy-level reflectance. Closed agricultural canopies can easily be assumed to be horizontally homogeneous, plane-parallel turbid mediums, and therefore the plate model theory can be applied successfully. On the other hand, open crop canopies, showing heterogeneous structure with different shapes, need the development of models that specifically treat the effects of openings, shadows, and canopy geometry on the modeled reflectance. A detailed discussion of the different canopy reflectance model approaches by turbid-medium, geometric, hybrid, and ray-tracing methods is provided by Goel (9).

Similarly, the TIR emittance of a plant canopy is a function of the temperatures of both the plant components and the soil. T_c is the canopy temperature, defined by Norman et al. (10) as the TIR temperature in which the "vegetation dominates the [measurement] field of view minimizing the effect of soil." T_o is the temperature of the soil surface. T_s is the surface composite temperature, defined by Norman et al.(10) as the "aggregate temperature of all objects comprising the surface," which was shown by Kustas et al. (11) to be a linear function of T_c and T_o.

REMOTE SENSING FOR DETECTING CROP WATER STRESS

During the past 30 years, progress has been made in using remotely sensed data to retrieve information useful for irrigation scheduling and management. The basic approaches have focused on parameters directly related to crop water status (e.g., crop water loss, evaporation, metabolism, photosynthesis, and fluorescence) and plant manifestations of chronic crop water stress (e.g., phenology and leaf expansion and loss).

Crop Evaporation

An important breakthrough in using remote sensing for water stress detection was the development of the Idso–Jackson crop water stress index (CWSI) (12,13). Jackson et al. (12) derived the theoretical CWSI ($\mathrm{CWSI_t}$) by relating the canopy–air temperature difference ($T_c - T_a$) to net radiation (R_n), vapor pressure deficit (VPD), soil heat flux (G), volumetric heat capacity of air (ρC_p), psychrometric parameters (γ and Δ), and canopy and aerodynamic resistances (r_c and r_a), where

$$(T_c - T_a) = [r_a(R_n - G)/\rho C_p]\{\gamma(1 + r_c/r_a)/[\Delta + \gamma(1 + r_c/r_a)]\} - \{\mathrm{VPD}/[\Delta + \gamma(1 + r_c/r_a)]\} \quad (1)$$

Taking the ratio of actual (E for any r_c) to potential (E_p for $r_c = r_{cp}$) crop evaporation rates gives

$$E/E_p = [\Delta + \gamma^*]/[\Delta + \gamma(1 + r_c/r_a)] \quad (2)$$

where r_{cp} is the canopy resistance at potential evaporation. Jackson et al. (12) defined the $\mathrm{CWSI_t}$, ranging from 0 (ample water) to 1 (maximum stress), as

$$\mathrm{CWSI_t} = 1 - E/E_p = [\gamma(1 + r_c/r_a) - \gamma^*]/[\Delta + \gamma(1 + r_c/r_a)] \quad (3)$$

Though Jackson et al. (12) provided a thorough theoretical approach for computing CWSI, the concept is more universally applied using a semiempirical variation proposed by Idso et al. (13) based on the "non-water-stressed baseline." This baseline is defined by the relation between ($T_c - T_a$) and VPD under nonlimiting soil moisture conditions, example, when the plant water is evaporating at the potential rate (Fig. 1). Such non-water-stressed baselines have been determined for many different crops, including aquatic plants and both preheading and postheading growth periods for small grains (14).

Application of CWSI with satellite- or aircraft-based measurements of surface temperature is generally restricted to almost full-canopy conditions, so that the surface temperature sensed is equal to the canopy temperature (15). To deal with partial plant cover

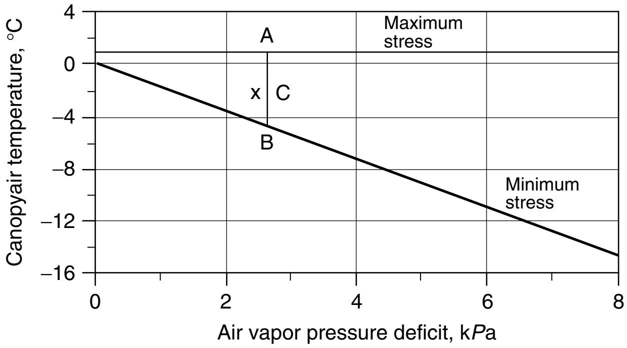

Figure 1. Canopy minus air temperature ($T_c - T_a$) versus vapor pressure deficit (VPD) for well-watered (line labeled minimum stress) and maximally stressed alfalfa based on measurements at various sites across the United States. The CWSI for a measured alfalfa canopy is computed as the ratio of the distances CB and AB (13).

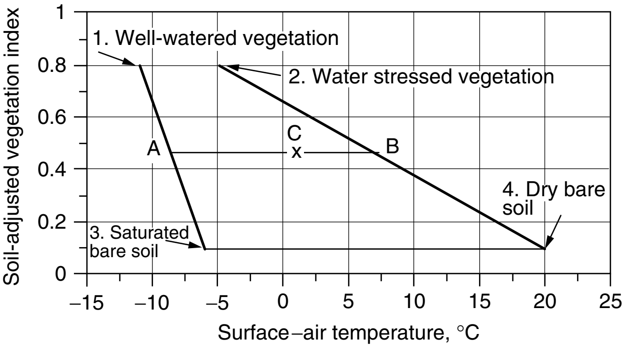

Figure 2. The trapezoidal shape that would result from the relation between surface temperature minus air temperature ($T_s - T_a$) and the soil-adjusted vegetation index (SAVI). Using a measurement of $T_s - T_a$ at point C, the ratio of actual to potential evaporation (which is equivalent to the Water Deficit Index or WDI) is equal to the ratio of the distances CB and AB (16).

conditions, Moran et al. (16) developed a Water Deficit Index (WDI) which combined measurements of reflectance with surface temperature measurements (a composite of both the soil and plant temperatures), as expressed by

$$\mathrm{WDI} = 1 - E/E_\mathrm{p} = [(T_\mathrm{s} - T_\mathrm{a})_\mathrm{m} - (T_\mathrm{s} - T_\mathrm{a})_\mathrm{r}]/[(T_\mathrm{s} - T_\mathrm{a})_\mathrm{m} - (T_\mathrm{s} - T_\mathrm{a})_\mathrm{x}] \qquad (4)$$

where subscripts m, x, and r refer to minimum, maximum, and actual, respectively. The WDI is operationally equivalent to the CWSI for full-cover canopies, where $T_s = T_c$. Graphically, WDI is equal to the ratio of distances AC/AB in the trapezoidal shape presented in Fig. 2, where WDI = 0.0 for well-watered conditions and WDI = 1.0 for maximum stress conditions. The left edge of the Vegetation Index/Temperature (VIT) trapezoid corresponds to $T_s - T_a$ values for surfaces evaporating at the potential rate; the right edge corresponds to $T_s - T_a$ values for surfaces in which no evaporation is occurring. In practice, WDI uses the Penman–Monteith energy balance equation to define the four vertices of the VIT Trapezoid that encompasses all possible combinations of a spectral vegetation index (e.g., Normalized Difference Vegetation Index or NDVI) and $T_s - T_a$ for one crop type on one day (Fig. 2).

Another promising approach for operational application is the use of remotely sensed crop coefficients (the ratio of actual crop evaporation to that of a reference crop) for estimating an actual, site-specific crop evaporation rate from readily available meteorologic information (17,18). This approach requires only a measure of NDVI and is simply an improvement of an approach already accepted and used by farmers to manage crops, where such improvements include increases in the accuracy of evaporation estimates and, by using images, the ability to map within-field and between-field variations.

Crop Metabolism

Quite distinct from the CWSI and WDI, Burke et al. (19) developed a concept of thermal stress in plants that linked the biochemical characteristics of a plant with its optimal leaf temperature range. The thermal kinetic window (TKW) is the range of temperatures within which the plant maintains optimal metabolism. For example, the TKW for cotton growth is 23.5 to 32 °C; the optimum temperature is 28 °C, and biomass production is directly related to the amount of time that canopy temperatures are within the TKW (Fig. 3), provided that insolation, soil moisture, and nutrients are nonlimiting. TKWs have been identified for several crop species, including cotton, wheat, soybean, potato, cucumber, and bell pepper. In practice, the TKW provides a biological indicator of plant health that is being used for irrigation management.

Crop Photosynthesis

Measurements of leaf chlorosis, which causes diminished leaf photosynthetic capacity, may be particularly suitable for early water stress detection. Differences in remote sensing reflectance due to changes in chlorophyll concentration levels have been detected in the green peak and along the red edge spectral region (690 to 750 nm) (20). The red edge occurs between the wavelengths of 0.69 and 0.76 μm due to the change in reflectance caused by chlorophyll absorption in the red spectrum and multiple scattering from leaves in the NIR spectrum (21) (Fig. 4). Several hyperspectral

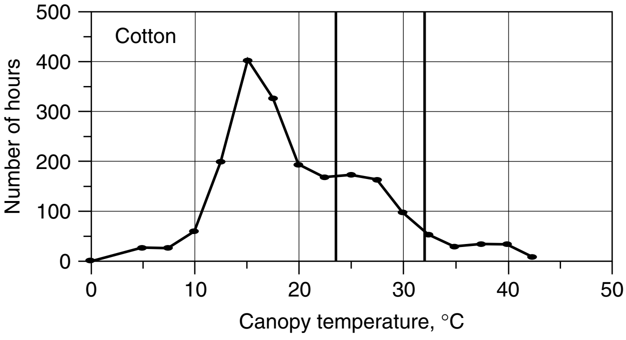

Figure 3. A frequency distribution of seasonal canopy temperatures of cotton, where the vertical lines represent the temperature range that comprises the species-specific thermal kinetic window (TKW) as determined from the changes in the apparent K_m with temperature (19).

indexes proposed in the literature track and quantify chlorophyll concentration for both leaf and canopy-level studies (22,23), allowing remote sensing detection methods to identify vegetation stress through the influence of chlorophyll content variation. These narrow-band vegetation indexes are, among others, the Simple Ratio Pigment Index, SRPI (ρ_{430}/ρ_{680}); Normalized Phaeophytinization Index, NPQI ($\rho_{415} - \rho_{435})/(\rho_{415} + \rho_{435}$); Photochemical Reflectance Index, PRI ($\rho_{531} - \rho_{570})/(\rho_{531} + \rho_{570}$); Normalized Pigment Chlorophyll Index, NPCI ($\rho_{680} - \rho_{430})/(\rho_{680} + \rho_{430}$); Greenness Index, G ($\rho_{554}/\rho_{677}$); Structure Intensive Pigment Index, SIPI ($\rho_{800} - \rho_{450})/(\rho_{800} + \rho_{650}$); and red edge reflectance-ratio indexes such as (ρ_{740}/ρ_{720}), ($\rho_{734} - \rho_{747})/(\rho_{715} + \rho_{726}$), ($\rho_{750}/\rho_{700}$), and ($\rho_{750}/\rho_{710}$). See a review by Zarco-Tejeda et al. (23) for more detail.

In agricultural canopies that have large spectral contributions from the soil background and LAI variation in different growth stages, combined indexes are proposed to minimize background soil effects while maximizing the sensitivity to chlorophyll concentration. CARI (chlorophyll absorption in reflectance index) (24) and MCARI (modified chlorophyll absorption in reflectance index) (25) have been proposed for combination with SAVI (soil-adjusted vegetation index) (26) and OSAVI (optimized soil-adjusted vegetation index) (27) to reduce background reflectance contributions. Successful chlorophyll estimation for corn canopies at different growth stages was achieved using the TCARI/OSAVI combined index (28) and with MCARI/OSAVI on open-tree crop canopies (29).

New developments related to vegetation stress, that have potential application to water stress detection, focus on monitoring chlorophyll fluorescence (CF). Red and far-red light is emitted from photosynthetic green plant tissues in response to photosynthetically active radiation. Several reviews offer details on theory, measurement methods, and interpretation, as well as the relationship to photosynthesis, plant physiological status, and photosynthetic functioning (30–32). In particular, relationships between steady-state and dark-adapted chlorophyll fluorescence features with water stress status have been investigated for both active and passive sensors (33–37). This suggests the potential for stress detection using far-distance remote sensing instruments.

Steady-state CF and photosynthetic rates are inversely related, such that CF is low when photosynthesis is high, although it can be reversed because of an intensified protective quenching action on CF production. The total amount of CF emitted by Photosystems I (PS-I) and II (PS-II) is typically less than 5% of total light absorbed. Thus, accurate remote detection of natural CF would enable the measurement of canopy-level CF without direct contact with the vegetation canopy. Zarco-Tejeda et al. (38) demonstrated that CF could be detected at both leaf and canopy levels using multiple approaches, including observing its effects on the red edge spectral region.

Crop Phenology

There is evidence that chronic crop water stress can either hasten (1) or delay (39) crop development, depending on the species involved and the timing of water stress. The timing and duration of stress is also of critical importance to ultimate yield; example, if intense water stress occurs during anthesis of a wheat crop, the impact on final grain yield is usually much greater than if this same stress occurs earlier or later in crop development. Tucker et al. (40) showed that crop phenological stage could be determined using a combination of spectral data and accumulated temperature units (growing degree day). Idso et al. (39) equated the slope of the spectral vegetation index (VI) over time with the rate of senescence and correlated this slope with final grain yield for wheat and barley under stressed and nonstressed conditions (Fig. 5). In a similar study, Fernández et al. (41) found that the hydric stress on wheat could be determined by the slope of the NDVI with time.

Chronic crop water stress can also manifest itself in reduced leaf expansion and leaf loss. There have been several studies that have monitored stress-induced reductions in biomass and GLAI using remote sensing techniques (42–44). Though these results are encouraging, this approach is limited by the fact that the VI–GLAI relation is exponential, leading to saturation of the NIR response at GLAI values of five to six (2). A better measure of the effects of water stress on leaf expansion and biomass production might be the amount of solar radiation intercepted, which is related directly to plant growth. Kumar and Monteith (45) showed that the fraction of absorbed photosynthetically active radiation (fAPAR) is related linearly to NIR/red reflectance. More recent work

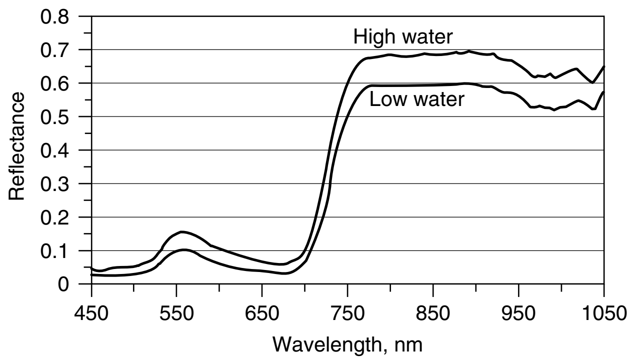

Figure 4. The expected trends of spectral reflectance of a crop canopy with high and low water treatments (21).

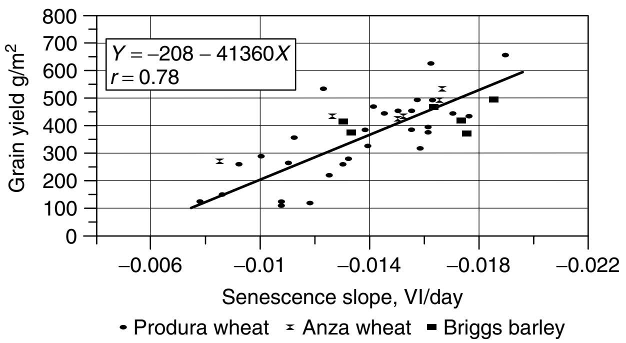

Figure 5. Final grain yield of Produra wheat, Anza wheat, and Briggs barley versus the slope of the transformed vegetation index (TVI6) during the senescence period (39).

by Pinter (46) has shown that this relation between VI and fAPAR is independent of variations in the solar zenith angle, thus increasing the usefulness of this remote sensing approach.

CONCLUDING REMARKS

The CWSI, WDI, and TKW are good examples of remote sensing indexes that link surface temperature measurements to crop and soil evaporation rates. Quantification of leaf chlorosis based on surface reflectance in narrow wavelength bands also has potential for detecting transient water stress. Season-long measurements of surface reflectance in wide wavelength bands have been used to detect the effects of chronic plant stress, such as delayed (or hastened) crop phenology and leaf expansion or loss. These are only some examples of the many approaches available for detecting crop water stress using remote sensors. However, to realize the full potential of satellite-based estimation of crop water stress, it will be necessary to continue sensor development, improve image availability and timely delivery, and reduce image cost.

BIBLIOGRAPHY

1. Turner, N.C. (1977). Drought resistance and adaptation to water deficits in crop plants. In: *Stress Physiology In Crop Plants*. H. Mussell and R.C. Staples (Eds.). Wiley-Interscience, New York, pp. 344–372.
2. Bauer, M.E. (1985). Spectral inputs to crop identification and condition assessment. *Proc. IEEE* **73**: 1071–1085.
3. Horler, D.N.H., Dockray, M., and Barber, J. (1983). The red edge of plant leaf reflectance. *Int. J. Remote Sensing* **4**: 273–288.
4. Allen, W.A. and Richardson, A.J. (1968). Interaction of light with a plant canopy. *J. Opt. Soc. Am.* **58**: 1023–1028.
5. Jacquemoud, S. and Baret, F. (1990). Prospect: A model of leaf optical properties spectra. *Remote Sensing Environ.* **34**: 75–91.
6. Guyot, G., Dupont, O., Joannes, H., Prieur, C., Huet, M., and Podaire, A. (1984). Investigation into the mid-IR spectral band best suited to monitoring vegetation water content. *Proc. 18th Int. Symp. Remote Sensing Environ.*, October 1–5, 1984, Paris, France, pp. 1049–1063.
7. Jackson, R.D. and Ezra, C.E. (1985). Spectral response of cotton to suddenly induced water stress. *Int. J. Remote Sensing* **6**: 177–185.
8. Myers, V.I., Bauer, M.E., Gausman, H.W., Hart, W.G., Heiman, J.L., MacDonald, R.B., Park, A.B., Ryerson, R.A., Schmugge, T.J., and Westin, F.C. (1983). Remote sensing application in agriculture. In: *Manual of Remote Sensing*, 2nd Edn. R.N. Colwell (Ed.). American Society of Photogrammetry and Remote Sensing, Falls Church, VA, pp. 2111–2228.
9. Goel, N.S. (1988). Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data. *Remote Sensing Rev.* **4**: 1–212.
10. Norman, J.M., Divakarla, M., and Goel, N.S. (1995). Algorithms for extracting information from remote thermal-IR observations of the Earth's surface. *Remote Sensing Environ.* **51**: 157–168.
11. Kustas, W.P., Choudhury, B.J., Inoue, Y., Pinter, P.J., Jr., Moran, M.S., Jackson, R.D., and Reginato, R.J. (1990). Ground and aircraft infrared observations over a partially-vegetated area. *Int. J. Remote Sensing* **11**: 409–427.
12. Jackson, R.D., Idso, S.B., Reginato, R.J., and Pinter, P.J., Jr. (1981). Canopy temperature as a crop water stress indicator. *Water Resour. Res.* **17**: 1133–1138.
13. Idso, S.B., Jackson, R.D., Pinter, P.J., Jr., Reginato, R.J., and Hatfield, J.L. (1981). Normalizing the stress-degree-day parameter for environmental variability. *Agric. Meteorol.* **24**: 45–55.
14. Idso, S.B. (1982). Non-water-stressed baselines: A key to measuring and interpreting plant water stress. *Agric. Meteorol.* **27**: 59–70.
15. Millard, J.P., Reginato, R.J., Idso, S.B., Jackson, R.D., Goettelman, R.C., and Leroy, M.J. (1980). Experimental relations between airborne and ground measured wheat canopy temperatures. *Photogrammetric Eng. Remote Sensing* **46**: 221–224.
16. Moran, M.S., Clarke, T.R., Inoue, Y., and Vidal, A. (1994). Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index. *Remote Sensing Environ.* **49**: 246–263.
17. Bausch, W.C. (1993). Soil background effects on reflectance-based crop coefficients for corn. *Remote Sensing Environ.* **46**: 213–222.
18. Hunsaker, D.J., Pinter, P.J., Jr., Barnes, E.M., and Kimball, B.A. Estimating cotton evapotranspiration crop coefficients with a multispectral vegetation index. *Irrig. Sci.* (in press).
19. Burke, J.J., Mahan, J.R., and Hatfield, J.L. (1988). Crop-specific thermal kinetic windows in relation to wheat and cotton biomass production. *Agron. J.* **80**: 553–556.
20. Rock, B.N., Hoshizaki, T., and Miller, J.R. (1988). Comparison of *in situ* and airborne spectral measurements of the blue shift associated with forest decline. *Remote Sensing Environ.* **24**: 109–127.
21. Filella, I., and Peñuelas, J. (1994). The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *Int. J. Remote Sensing* **15**: 1459–1470.
22. Carter, G.A. (1994). Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *Int. J. Remote Sensing* **15**: 697–704.
23. Zarco-Tejada, P.J., Miller, J.R., Mohammed, G.H., Noland, T.L., and Sampson, P.H. (2001). Scaling-up and model inversion methods with narrow-band optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Trans. Geoscience Remote Sensing* **39**: 1491–1507.
24. Kim, M.S., Daughtry, C.S.T., Chappelle, E.W., McMurtrey, J.E., III, and Walthall, C.L. (1994). The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (Apar). *6th Symp. Phys. Meas. Signatures Remote Sensing*, Jan. 17–21, 1994, Val D'Isere, France, pp. 299–306.
25. Daughtry, C.S.T., Walthall, C.L., Kim, M.S., Brown de Colstoun, E., and McMurtrey, J.E., III (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing Environ.* **74**: 229–239.
26. Huete, A.R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing Environ.* **25**: 295–309.
27. Rondeaux, G., Steven, M., and Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing Environ.* **55**: 95–107.
28. Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., and Dextraze, L. (2002). Integration of hyperspectral

vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing Environ.* **81**: 416–426.

29. Zarco-Tejada, P.J., Miller, J.R., Morales, A., Berjón, A., and Agüera, J. (2004). Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote Sensing Environ.* **90**: 463–476.
30. Papageorgiou, G. (1975). Chlorophyll fluorescence: An intrinsic probe of photosynthesis. In: *Bioenergetics of Photosynthesis*. Govindjee (Ed.). Academic, New York, pp. 319–371.
31. Krause, G.H. and Weis, E. (1984). Chlorophyll fluorescence as a tool in plant physiology. II. Interpretation of fluorescence signals. *Photosynthesis Res.* **5**: 139–157.
32. Schreiber, U., Bilger, W., and Neubauer, C. (1994). Chlorophyll fluorescence as a nondestructive indicator for rapid assessment of *in vivo* photosynthesis. *Ecological Stud.* **100**: 49–70.
33. Evain, S., Flexas, J., and Moya, I. (2004). A new instrument for passive remote sensing: 2. Measurement of leaf and canopy reflectance changes at 531 nm and their relationship with photosynthesis and chlorophyll fluorescence. *Remote Sensing Environ.* **91**: 175–185.
34. Souza, R.P., Machado, E.C., Silva, J.A.B., Lagôa, A.M.M.A., and Silveira, J.A.G. (2004). Photosynthetic gas exchange, chlorophyll fluorescence and some associated metabolic changes in cowpea (*Vigna unguiculata*) during water stress and recovery. *Environ. Exp. Bot.* **51**: 45–56.
35. Flexas, J., Briantais, J-M., Cerovic, Z., Medrano, H., and Moya, I. (2000). Steady-state and maximum chlorophyll fluorescence responses to water stress in grapevine leaves: A new remote sensing system. *Remote Sensing Environ.* **73**: 283–297.
36. Cerovic, Z.G., Goulas, Y., Gorbunov, M., Briantais, J-M., Camenen, L., and Moya, I. (1996). Fluorosensing of water stress in plants: Diurnal changes of the mean lifetime and yield of chlorophyll fluorescence, measured simultaneously and at distance with a [tau]-LIDAR and a modified PAM-fluorimeter, in maize, sugar beet, and kalanchoë. *Remote Sensing Environ.* **58**: 311–321.
37. Valentini, R., Cecchi, G., Mazzinghi, P., Scarascia Mugnozza, G., Agati, G., Bazzani, M., De Angelis, P., Fusi, F., Matteucci, G., and Raimondi, V. (1994). Remote sensing of chlorophyll a fluorescence of vegetation canopies: 2. Physiological significance of fluorescence signal in response to environmental stresses. *Remote Sensing Environ.* **47**: 29–35.
38. Zarco-Tejada, P.J., Pushnik, J., Dobrowski, S., and Ustin, S.L. (2003). Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. *Remote Sensing Environ.* **84**: 283–294.
39. Idso, S.B., Pinter, P.J., Jr., Jackson, R.D., and Reginato, R.J. (1980). Estimation of grain yields by remote sensing of crop senescence rates. *Remote Sensing Environ.* **9**: 87–91.
40. Tucker, C.J., Elgin, J.H., Jr., and McMurtrey, J.E., III. (1979). Temporal spectral measurements of corn and soybean crops. *Photogrammetric Eng. Remote Sensing* **45**: 643–653.
41. Fernández, S., Vidal, D., Simón, E., and Solé-Sugrañes, L. (1994). Radiometric characteristics of Triticum aestivum cv. Astral under water and nitrogen stress. *Int. J. Remote Sensing* **15**: 1867–1884.
42. Gardner, B.R. and Blad, B.L. (1986). Evaluation of spectral reflectance models to estimate corn leaf area while minimizing the influence of soil background effects. *Remote Sensing Environ.* **20**: 183–193.
43. Pinter, P.J., Jr., Jackson, R.D., Idso, S.B., and Reginato, R.J. (1981). Multidate spectral reflectances as predictors of yield in water stressed wheat and barley. *Int. J. Remote Sensing* **2**: 43–48.
44. Mogensen, V.O., Jensen, C.R., Mortensen, G., Thage, J.H., Koribidis, J., and Ahmed, A. (1996). Spectral reflectance index as an indicator of drought of field grown oilseed rape (*Brassica napus* L.), *Eur. J. Agron.* **5**: 125–135.
45. Kumar, M. and Monteith, J.L. (1981). Remote sensing of crop growth. In: *Plants and the Daylight Spectrum*. H. Smith (Ed.). Academic, London, pp. 133–144.
46. Pinter, P.J., Jr. (1993). Solar angle independence in the relationship between absorbed PAR and remotely sensed data for alfalfa. *Remote Sensing Environ.* **46**: 19–25.

VACUUM GAUGE TENSIOMETER

Mahbub Alam
Kansas State University
Garden City, Kansas

Soil water plays an important role in providing water for plant growth. The vacuum gauge tensiometer is a soil water measuring device that is sensitive to soil water change. The quantity of soil water that adheres around soil particles determines the energy level at which it is being retained or pulled. In other words, the plant root has to exert energy to pull water out of a soil matrix. This energy is also expressed as tension the roots experience in removing the water. Tensiometer is like a mechanical root but shows the tension in the vacuum gauge attached to it. It is an indirect method to determine soil water status. Direct measure of water content is time-consuming and labor-intensive, and the result lags behind the current time. Hence, irrigators depend on indirect methods. Vacuum gauge tensiometer is inexpensive, easy to use, and a preferred measurement method, especially for course-textured soils.

A vacuum gauge tensiometer with its components is shown in Fig. 1. Variations from this design exist on the market. A tensiometer is a sealed, water-filled tube equipped with a vacuum gauge on the upper end and a porous ceramic tip on the lower end. The basic components include: a cap, body tube, vacuum gauge with dial, and a porous ceramic tip (Fig. 1).

The Cap: The cap must seal the tensiometer body tube airtight or the device will not work. Some models have a reservoir, like the one shown in Fig. 1, which serves as a water supply source for the body tube. In these models, the cap assembly has an extended stem to hold a neoprene stopper that seals the body tube from outside atmosphere.

Body Tube: The body tube provides support and a liquid connection between a porous tip and the vacuum gauge. Tensiometers come in various lengths. Standard lengths are 6, 12, 18, 24, 36, 48, and 60 inches.

Vacuum Gauge: A hermetically sealed vacuum gauge (Bourdon[1] vacuum gauge) with a round dial is attached

---

[1] A thin-walled flattened tube used in a gauge to read pressure or vacuum developed by Bourdon, an inventor in France some 150

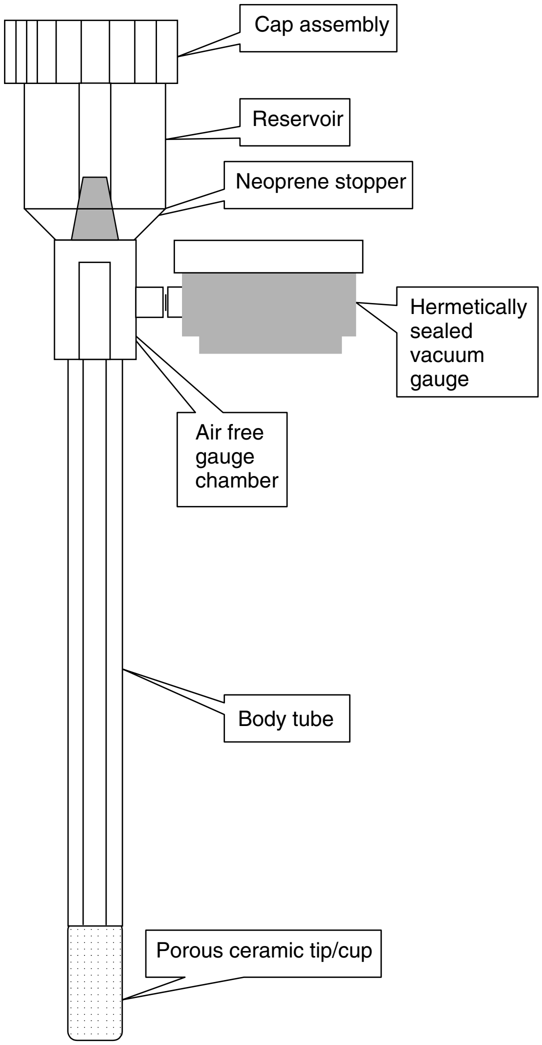

Figure 1. Sketched view of a vacuum gauge tensiometer showing its components.

at the air-free gauge chamber of the body tube. The round dial in the gauge is marked with a scale in centibars to indicate the gauge reading. The scale is calibrated to read in centibar, or hundredths of one "bar." A bar is the unit of pressure, either positive or negative, that has been adopted for the expression of soil water tension. The bar is an international unit of pressure in the metric system and equivalent to 14.5 lb/sq. inch, or 0.987 atmospheres. One centibar is equal to 1 kPa (kilopascal).

Ceramic Tip/Cup: The ceramic tip is built like a tubular cup, and it is porous. The openings of the pores are so small that when saturated with water, air cannot pass through within the range of soil water tensions to be measured. As soil water content outside the porous cup in the soil matrix declines from saturation, water inside the tube starts moving out through the porous tubular tip, which causes a partial vacuum inside the tensiometer that can be read on the vacuum gauge. The needle in the gauge starts moving above the face of the dial. The dial reading, indicated by the needle, shows the suction or tension at which the water is being pulled by the surrounding soil.

A reading of zero corresponds to a completely saturated condition, regardless of the type of soil. A reading of 80 kPa (centibar) indicates a dry condition for sandy soils and would hurt sensitive crops. This number is also the functional upper limit for tensiometer readings. A tension higher than 85 kPa (centibar) will cause the water inside the tube to vaporize, restricting its use to suctions less than about 85 kPa.

The cup conductance and the gauge sensitivity determine the response time of the tensiometer. Ceramic tensiometer tips/cups for field use generally have a conductance in the order of 3×10^{-5} cm^2/s, which, in a vacuum gauge system, would have a response time of five seconds. A response time of 1 minute is adequate for most field applications.

The tensiometer is sensitive to very small changes in soil water, making it useful for irrigation scheduling. Irrigation scheduling is a process to determine when to irrigate and how much water to apply. Applying water in an untimely manner with too little or too much water can result in yield reductions. Overirrigation wastes water, costs dollars to pump, and may leach nutrients beyond the root zone. Tensiometers are particularly accurate at low tensions, which is the wettest part of the soil water range. Tensiometers are popular with growers of high-value crops, such as vegetables and fruits, on sandy soils.

A depth label is usually placed on the vacuum gauge or on the side of the tube to indicate the depth at which the ceramic tip will be set when installed, which is important for identification purposes. The soil suction reading on the vacuum gauge dial is an indication of soil water availability for plant use and does not need to be calibrated for salinity or temperature. The readings, however, have different meanings in terms of their use for irrigation scheduling depending on soil type. Table 1 and Fig. 2 provide some interpretation of tensiometer readings in relation to soil texture.

TENSIOMETERS WITH ELECTRONIC READER

The vacuum gauge tensiometer is not suitable for automated continuous data collection or for controlling irrigation system operation by switching it on or off. However, some models have electronic technology added to the tensiometer that allow them to be remotely read or automatically switch irrigation on or off.

The switching tensiometer models employ the standard vacuum gauge, to which a circular magnet is affixed to the hub of the indicator needle. This magnet is rotated by the action of the indicator needle as it moves according to drying/wetting cycle of the soil. The switch is affixed to the outside of the gauge face and can be adjusted for various readings where the switch will turn "on" The switch contains a reed switch that is held open by a small biasing magnet installed inside the switch. This magnet is of lesser power than the one mounted on the hub of the gauge indicator needle. As the indicator needle of

---

years ago. It was brought to the United States by Ashcroft Gauge Company 100+ years ago and still is the standard gauge used practically worldwide.

Table 1. Interpretations of Tensiometer Readings

Reading (Centibars)	Status	Explanation/Action
0	Saturated	Soil is saturated regardless of soil type. If readings persist, there is possible danger of waterlogged soils, a high water table, poor drainage and soil aeration, or the continuity of the water column in the tube may have broken.
5–10	Surplus water	Indicate a surplus of water for plant growth. Drainage continues, and persistent readings indicate poor drainage.
10–20	Field Capacity	Field capacity for all types of soils. Additional water will drain as deep percolation carrying nutrients without opportunity for plant use. Sandy soils, however, have very little storage capacity, and suction values increase rapidly as plants remove water past 15 to 20 centibars. For sensitive crops, like potato, rapid irrigation may be required before damaging stress can develop.
20–40	Irrigation range	Available water and aeration good for plant growth in fine- and medium-textured soils. Irrigation is not required for these soils at this range. Coarse-textured soils may require irrigation in the 20 to 30 centibar range and the finer sandy soils at 30 to 40 centibar ranges.
40–60		Usual range for starting irrigation. At 40 to 50 centibar, irrigation may need to be started for loamy soils. On clay soils (silty clay loams, silty clays, etc.), irrigation usually starts from 50 to 60 centibars. Heavy clay soils still have some available water. Irrigation, however, ensures maintaining readily available soil water. The stage of growth and type of crop will influence the decision.
70	Dry	Stress range. However, crop is not necessarily damaged. Some soil water is available in clay soils, but may be low for maximum production.
80		Top range of tensiometer accuracy; higher readings are possible, but tension within the water column inside tensiometer will break between 80 to 85 centibar, which has relationship to elevation of the area compared with mean sea level. At higher elevation, the continuity of water column inside the tube may break at a lower reading because of change in vapor pressure according to change in atmospheric pressure.

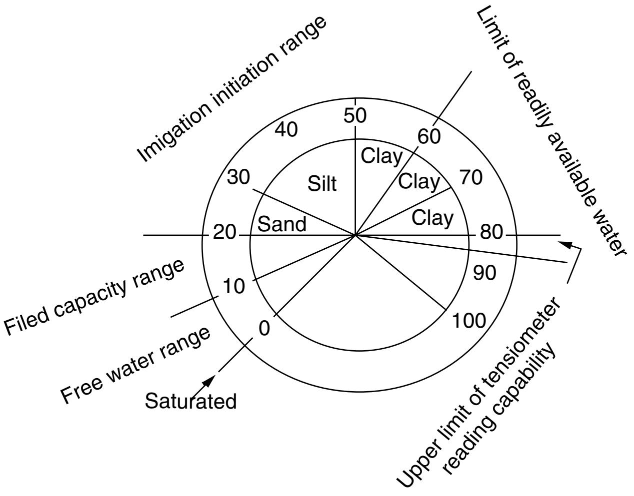

Figure 2. Schematic of the tensiometer vacuum gauge dial to show interpretation of readings in relation to soil type.

the tensiometer rises with the drying of soil to the point where the switch has been set, the magnet on the gauge overrides the small biasing magnet inside the switch, which causes the reed switch to close. The switch is wired into a low-voltage current loop from the control device to the solenoid valve. The switch closure completes the circuit and allows the solenoid valve to open and irrigation to start. When the soil rewets, the indicator needle drops down and the switch opens interrupting current loop from the control device that causes the solenoid valve to close, terminating irrigation.

For automated data collection, the standard vacuum gauge is replaced by a pressure transducer, where the partial vacuum reading is converted as current flow signal. The partial vacuum, caused by water moving out of tensiometer, causes a deflection in the face of the transducer, which excites the electronic elements of the device to generate the signal in a certain milliampere range equivalent to centibar or kPa linear scale. These changing signals coming from the transducer may be recorded by data logger or computer. The reading is continuous. The data acquisition computer may be programmed to start an irrigation system at a certain value. The automated versions of tensiometers enable remote operation of the irrigation system.

WORKING PRINCIPLE OF A TENSIOMETER

Soil water exists primarily as thin films around and between soil particles and is bound to soil particles by strong molecular forces. As the soil dries, the water films become thinner and become more tightly bound to soil matrix. This increase in tension within the films now in contact with the tensiometer causes water to be drawn from the ceramic tip. The withdrawal of water from the ceramic tip creates a partial vacuum within the tensiometer. Water continues to be drawn until the vacuum created inside the tensiometer equals the tension of the water films outside. At this point, equilibrium is reached and water ceases to flow. The vacuum gauge reading indicates the amount of suction or tension.

As water is added to the soil from rainfall or irrigation, the soil suction is reduced. The higher vacuum in the tensiometer causes soil water to be drawn into the tensiometer, and the vacuum will be reduced until a balance in tension is reached. The tensiometer continuously responds and maintains a balance with the soil water suction or tension, and the vacuum gauge indicates the amount of tension, hence the name tensiometer.

PREPARING THE TENSIOMETER FOR FIELD USE

Preparing a tensiometer for use starts with filling the reservoir and body tube with distilled water, taking care that the ceramic tip is wetted from one direction to avoid air entrapment in the finer pores of the ceramic tip. Allow the tensiometer to stand upright, and soon the tip will wet-up and free water will appear like sweat on the outer surface of the ceramic tip. Refill as necessary to dispel all the air from the tensiometer. A hand-operated vacuum pump may be used to help remove air from the tensiometer. Operating the pump when the tensiometer is filled with water and the tip is submerged in water helps remove gases from the pores of the ceramic tip and solution. After each pumping, refill the tensiometer completely with water and repeat until no more bubbles are observed. The tensiometer is then sealed by screwing the reservoir cap down securely.

After letting it stand in water in a bucket overnight, the tensiometer may be set outside upright in the air. The air will start drying the tip. The gauge should read 70 when the tip is air-dried. Repeat the wetting and drying cycle with the tensiometers that do not respond correctly the first time.

Tensiometers require checking prior to installation. As with any measurement device, proper care and maintenance are required. If the tensiometer was used previously, begin by washing and rinsing it inside and out. Residues on the porous ceramic tip that were not removed by washing may be removed by sanding lightly. Distilled water treated with three to five drops of chlorine bleach per gallon of water can be used to inhibit algae growth. Manufacturers also provide solutions for water treatment, which may be used according to direction. Distilled water available in a grocery store has been found to be adequate. If an excessive amount of air bubbles are noticed, then boiling may be helpful, but the remaining water must be stored in an airtight container.

If tensiometers are not to be installed immediately, cover the tips with a plastic bag to prevent evaporation or let them stand in a bucket of water until installed.

TENSIOMETER INSTALLATION

Site Selection

The number of tensiometer installation sites required will depend on the crops grown and field conditions. Fewer stations of tensiometers are needed when a single crop is grown in large blocks of uniform soil. If the soils are variable, different crops are to be grown and more stations are necessary. Stations need to be selected to represent an area, and care should be taken not to cause excessive compaction or destruction of plants during the installation process.

Except for very shallow-rooted crops, tensiometers are normally installed in groups of two or more to characterize the soil in the top half to three-quarters of the root zone. If the potential root zone is less than 12 inches, a single tensiometer may be installed in the center of the zone at 6 inches deep. With deeper-rooted crops, one tensiometer should be placed at the upper one-quarter of the rooting depth and another at the lower quarter point or three-quarters of the depth (see Fig. 3).

In deeply rooted crops, or situations where a distinct break in soil textures exists, three or more tensiometers may be needed. An example might be 18 inches of sand overlying a silty soil on which corn is grown and 3 to 4 feet of the root zone is to be managed. One tensiometer might be placed at 6 inches, a second at 18 inches, and a third at 2 to 3 feet. The differences in tension readings would make it possible to better assess the soil water conditions. Almost 70% of crop water is supplied by the top half and only 10% from lower one-fourth of the rooting depth. Irrigators, therefore, often manage only the top half or three-fourths of the root zone. Tensiometers should be long enough to reach the desired depth, and the diaphragm of the vacuum gauge must not touch the ground. The tensiometer should never be set in a hole or a depression.

Tensiometers may be installed using a soil auger or a probe. Manufacturers also provide simple coring tubes. Placement should be in a crop row to avoid traffic. Where furrow irrigation is used, the tensiometers may be angled slightly to place the tip under the furrow. The electronic tensiometers may require a cover to safeguard the electric connections from sprinkler or rainwater. If a valve cover box is used, the tensiometer tips need to be slanted out to be in the crop area. The hole should be small enough to create resistance to insertion of the tensiometer and the bottom shaped to form a close soil contact at the tip, which may be accomplished by returning a little loosened

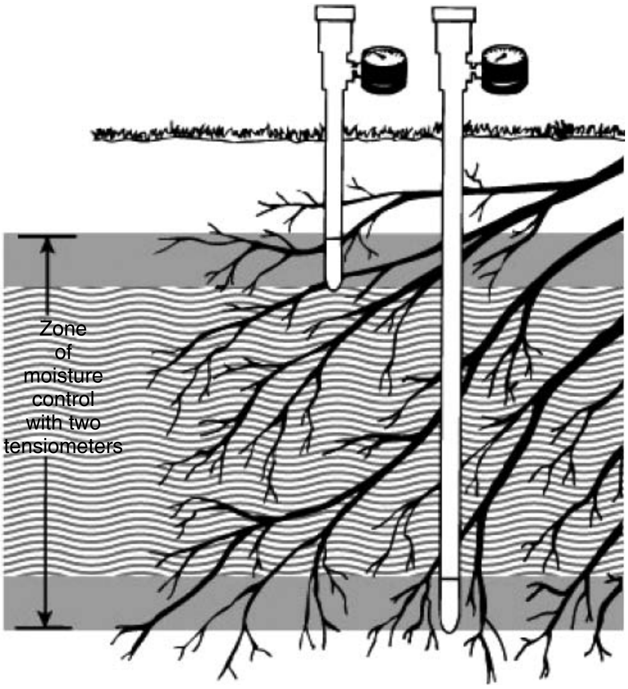

Figure 3. Sketch showing root zone moisture evaluation and control by two tensiometers.

portion of the soil from the depth of placement back into the hole and adding a little water. When the tensiometer is pushed for placement, the soft soil will move around the tip to conform to the rounded shape of the tip and make a good contact. A good contact is absolutely essential for the tensiometer to perform.

Tensiometers must be handled with care, because the tips may break if handled roughly. The depth label on the tensiometer will identify the root zone being monitored. Tensiometer locations need to be marked both in the row and at the edge of the field. A wooden stake painted bright or a metal rod with a colored flag attached are good markers. Locating tensiometers in tall crops can be a problem. A written log of the station locations is also recommended.

Servicing

Tensiometers are weatherproof, except for freezing, and generally require very little service. When first installed, there may be tiny air bubbles clinging to the sides of the body tube. However, after one cycle of soil water use, which creates a high vacuum, the bubbles will rise to the top and can be eliminated by refilling. The amount of bubbles will depend on the gas originally present in the vacuum gauge and the amount dissolved in the water. Servicing is best done soon after irrigation. Tensions are low, and air that may have been drawn into the cup at high tensions can be eliminated by refilling. Tensiometers return to equilibrium rapidly at low tensions. They also respond quickly to a very minute withdrawal of water from the system. If much air is drawn into a tensiometer at low tensions, the porous cup may be defective, and the tensiometer may need to be replaced. Some air entry is unavoidable. When using a number of tensiometers, watch for tensiometers that accumulate abnormal amounts of air. The colored fluid concentrate, supplied by the manufacturer for control of algae, helps to spot air bubbles in the tensiometer more easily.

Tensiometers should be removed from the field before freezing. The water can freeze and break the ceramic tip or the body, or damage the vacuum gauge. Tensiometers need to be emptied before long-term storage, which prevents salt deposition in the porous material with evaporation or rusting of the gauge.

Troubleshooting

A tensiometer that is out of water or leaking will remain at zero on the gauge, or the reading will fluctuate in the low suction range. Two or more successive zero readings may be a sign of a malfunction and should be investigated. If the gauge remains at zero, refill with water and use a hand pump, if available, to remove air. The tensiometer may have been empty because of dry soil. If the tip was dry, fine air bubbles will rise rapidly for several minutes and then cease. If larger bubbles rise and continue, a leak is indicated and the source should be determined. If the bubbles rise from the bottom, remove the tensiometer and replace the tip. If the bubbles enter from the side, the body tube may be cracked and should be fixed. If bubbles rise from the gauge, the leak may be in the gauge or the threaded connection. A leaky gauge needs to be replaced, but a threaded connection can be resealed. If no large bubbles rise and yet the reading remains at zero then the reservoir cap may be cracked or the seal may be defective. Inspect for 'O' ring if there needs to be one for a proper seal. In most cases, the trouble is easily corrected.

A damaged vacuum gauge may stick in one position or may not respond smoothly with changes in soil water. Check a suspect gauge against one known to be in good working order or replace the suspect gauge with a new one. Readings that are higher than expected, especially after irrigation, generally do not indicate a tensiometer failure. The irrigation water may not have penetrated to the depth of the tensiometer tip.

General Guidelines

- Place two or more tensiometers of different lengths near one another (a station), usually in the crop row. Two stations may be enough in a field with uniform soil and slope. One station should be near the start of an irrigation set and the other at the end of that irrigation set. The location of stations will depend on the type of irrigation system. With furrow irrigation, this location may be at the upper and lower quarter points of the first set in the field. For a center pivot, it may be in the outer and middle spans. If desired, one set may be established at the starting point of the pivot and another set at the endpoint.
- Tensiometer stations should be located in representative areas of the field. Do not position tensiometers in low spots or on knobs, and place them where the plant population is representative of the field.
- Tensiometer installation depth is determined by the active root zone of the crop. For example, for a corn crop on a deep soil, three tensiometers, installed at depths of 12, 24, and 36 inches, are recommended at each station.
- One should wait 24 hours after installing the tensiometer to obtain reliable readings. If the soil was dry at installation, irrigation or rainfall may be needed before obtaining satisfactory readings.
- Tensiometers should be installed as early as possible during the growing season and left in the field for the duration of the growing season. The roots of the crop must grow around the porous tip for reliable readings. Moving of the tensiometer during the growing season is not recommended.

LIMITATIONS OF TENSIOMETER

The major criticism of the tensiometer is that it functions reliably only in the wet range of soil water at readings of about 80 centibar or less. At higher readings, the porous tip may leak air and the gases will be drawn out of the water, and at low pressure, the water will vaporize causing discontinuation of the tension column or vacuum. The gauge reading will fall to zero, which is not as serious as it may seem, because most of the available water in coarse-textured soils, and about 50 percent or more in fine-textured soils, have already been used up at this range.

Another criticism is the price, which ranges from $35 to $60. When used in large quantities, the cost may seem prohibitive. Irrigation scheduling, however, has been shown to easily pay off through increased yields or reduced pumping cost. Proper handling may extend the useful life, and the cost may be spread over many seasons, making it cost effective.

Finally, the ceramic tip may gradually fill with precipitates because of soil water movement through the pores, which slows water transfer through the tip and increases the time required for the tensiometer to respond to a change in soil-water conditions. Some slowing does no harm, but if the response time becomes too slow, a new tip should be installed. The response time may be improved by rubbing the exterior of the tip with fine sand paper or soaking the tip in a mild acid solution. The amount of plugging depends on the soil-water chemistry and the manner of use.

Where tensiometers can be left in the ground, the tip porosity remains satisfactory for several years in most soils. However, each time the tensiometer is removed from the soil, tip life is reduced, which is particularly true if the soil is calcareous or saline. In extreme cases, where the tensiometer is installed and removed several times per season, the tip may need to be replaced after one year of use. To minimize this damage, a tip that is removed from the soil should be protected from drying until the tensiometer has been emptied, cleaned, and dried.

TILE DRAINAGE

Tara Reed
University of Wisconsin-Green Bay
Green Bay, Wisconsin

Tile (or relief) drains are porous pipes buried approximately 1 m beneath the surface of the soil. These pipes drain excess water from the land into nearby water bodies, most typically cisterns, open ditches, or streams. The most common use of tile drainage is to improve agricultural production by draining excess moisture from the plant root zone and allowing cultivation normally saturated soils. Subsurface tile drainage also lowers the water table and reduces soil salinity. Many of the world's productive agricultural areas, including 25% of the cropland in the United States, require artificial drainage (1). The area drained by subsurface drainage systems in the United States has increased since 1900 from less than 5 to almost 30 million hectares (2).

TILE DRAINS AND SURFACE WATER QUALITY

The impact of tile drainage on receiving water bodies varies (3). Site-specific variables such as soil characteristics, topography, climate, and precipitation patterns as well as the configuration of the drainage network and other land management practices determine the proportion of precipitation resulting in surface runoff and subsurface drainage and the overall impact on receiving water bodies. Because rainfall can penetrate further into the soil, surface water drainage is reduced in tile-drained fields compared to saturated soils (4). Depending on local conditions, subsurface drainage can reduce surface runoff (4,5).

Increased infiltration prolongs the time that water is in contact with the soil, thereby often reducing the transport of pollutants that are likely to adsorb to soil particles and increasing leaching of pollutants that are easily released (6). The difference between surface and subsurface water drainage may be minimized when there are direct surface water inlets into the subsurface drainage network or during particularly heavy rains, although Schwab et al. (5) found that good subsurface drainage decreased storm peak discharge by as much as 30%. On the other hand, management practices such as riparian buffer strips that are designed to reduce surface runoff usually do not target subsurface drainage. In these cases, tile drains may result in increased pollutant loading during storms (7).

Pesticides and Herbicides

Agricultural pesticides and herbicides enter surface waters by both surface runoff and subsurface drainage. Whether subsurface drainage decreases the amount of pollutant entering the waters depends on the hydrologic characteristics of the site and the particular chemical used. The two pollutant characteristics important to this discussion are persistence (the time it takes to degrade to a nonpollutant form) and soil sorption (the degree to which the pollutant adsorbs to soil particles). Persistent pesticide/herbicides are most likely to drain (either by surface runoff or subsurface drainage) into nearby receiving waters. Subsurface drainage is most likely to reduce the transport of pesticides and herbicides with high soil sorption qualities. Soil characteristics can influence sorption, which is greatest in strongly organic soils and those that have high clay content.

Pesticides are most commonly detected in subsurface flow during the first heavy rain following application. Other peak periods include soil thawing and tillage. However, the most important factors determining pesticide transport are precipitation volume and timing and soil type. Pesticide loss through tile drainage is usually substantially lower than loss through surface runoff. For an excellent review of tile drainage and pesticide transport, see Kladivko et al. (6).

Nutrient Pollution

Nitrogen. Much of the fertilizer applied to a field can be lost through tile drains. This is particularly true for nitrogen, which is more mobile than phosphorus in soil. Different species of nitrogen react differently. Tile drainage usually decreases losses of organic nitrogen and increases loss of nitrates and ammonium (3).

Depending on the form of fertilizer applied, in wet years, up to 70% of the nitrate applied as fertilizer can flow off fields through drainage systems (8). Manure application leads to significantly less nitrate loss to tile drains than commercial fertilizer application (9). Drain placement may also be important for reducing nitrogen losses through drain tiles. For example, shallow drains reduce nitrate leaching (10), and more frequent drain tile spacing reduces ammonium losses (11). Placement of tile drains in areas of preferential flow for manure or fertilizer applications can elevate effluent nutrient concentrations (12).

Phosphorus. Even though phosphorus (P) adsorbs strongly to soil particles, artificial drainage is an important contributor to nonpoint-source phosphorus pollution (13). Burke (14) reported phosphorus losses through drainage of approximately 33% of the amount applied in fertilizer. Other estimates range from 2% (15) to 17% of the applied phosphorus lost in drain tile effluent (16).

The greatest P loss from agricultural systems occurs during intensive rains (17). At these times, both surface flow and tile drain effluent P concentrations are highest (18). Tile drain discharge may be nonlinearly related to rainfall intensity. For example, in a small drainage system in Sweden, Ulen and Persson (19) found that half the yearly P export through drain tiles occurred during an average of 140 hours of episodic rain.

The ability of soils to retain P depends on soil P concentrations. As soils become P saturated, leaching becomes more prevalent. The point at which this happens varies in different soils ranging from 10–119 mg Olsen P kg^{-1} soil (20). The mass of phosphorus lost from agricultural systems via tile drains is small compared to nitrogen (21), but the impact on phosphorus-limited aquatic systems may be substantial.

TILE DRAINAGE AND MANAGEMENT

Agriculture in many areas would be difficult, if not impossible, without subsurface drainage. Research is ongoing into strategies for managing pollutants in drain tile effluent. Examples include controlled drainage where subsurface drainage systems are fitted with barriers to restrict drainage (3), subirrigation systems that both control drainage and provide increased irrigation to crops during nonsaturated periods (22), constructed treatment wetlands (23), and land retirement (24). Finding solutions that are both environmentally and agriculturally acceptable is critical to the survival of the world's waterways.

BIBLIOGRAPHY

1. Pavelis, G.A. (1987). Farm Drainage in the United States. In: *History, Status, and Prospects*. G.A. Pavelis (Ed.). USDA-ERS, Misc. Publ. 1455, pp. 110–136.
2. Shirmohammadi, A., Wenberg, R.D., Ritter, W.F., and Wright, F.S. (1993). Management of irrigation and drainage systems. In: *Integrated Perspectives*. R.G. Allen, (Ed.). ASCE, New York, pp. 531–41.
3. Skaggs, R.W., Breve, M.A., and Gilliam, J.W. (1994). Hydrologic and water quality impacts of agricultural drainage. *Crit. Rev. Environ. Sci. Technol.* **24**: 1–32.
4. Bengtson, R.L., Carter, C.E., McDaniel, V., and Halverson, B.E. (1984). Corn silage response to subsurface drainage in alluvial soil. *Trans. ASAE*. **27**: 1391.
5. Schwab, G.O., Fausey, N.R., Desmond, E.D., and Holman, J.R. (1985). Tile and surface drainage of clay soils. IV. Hydrologic performance with field crops (1973–80); V. Corn, Oats and Soybean Yields (1973–80); VI Water quality; Cross Moling over Tile Drains. Wooster, OH: OARDC, Res. Bull. 1166.
6. Kladivko, E.J., Brown, L.C., and Baker, J.L. (2001). Pesticide transport to subsurface tile drains in humid regions of North America. *Crit. Rev. Environ. Sci. Technol.* **31**: 1–62.
7. Reed-Andersen, T. (1999). *Flows of phosphorus on an agricultural landscape: implications for eutrophication and restoration of Lake Mendota*. PhD thesis. University of Wisconsin - Madison, Madison, WI, pp. 113–138.
8. Randall, G.W. and Iragavarapu, T.K. (1995). Impact of long-term tillage systems for continuous corn on nitrate leaching to tile drainage. *J. Environ. Qual.* **24**: 360–366.
9. Mitchell, J.K., McIsaac, G.F., Walker, S.E., and Hirschi, M.C. (2000). Nitrate in river and subsurface drainage flows from an east central Illinois watershed. *Trans. ASAE*. **43**: 337–342.
10. Gordon, R., Madani, T.A., Caldwell, K., Boyd, N., Astatkie, T., and Jamieson, R. (2000). Subsurface nitrate-N leaching loss as affected by drainage size and depth in a shallow slowly-permeable soil. *Can. Water Resourc. J.* **25**: 331–342.
11. Singh, M.A., Bhattacharya, K., Nair, T.V., and Singh, A.K. (2001). Ammonium losses through subsurface drainage effluent from rice fields of coastal saline sodic clay soils. *Water, Air, Soil Pollut.* **127**: 1–14.
12. Schellinger, G.R. and Clausen, J.C. (1992). Vegetative filter treatment of dairy barnyard runoff in cold regions. *J. Environ. Qual.* **21**: 40–45.
13. Sims, J.T., Simard, R.R., and Joern, B.C. (1998). Phosphorus loss in agricultural drainage: Historical perspective and current research. *J. Environ. Qual.* **27**: 227–299.
14. Burke, W. (1975). Fertilizer and other chemical losses in drainage water from Blanket Bog. *Irish Journal of Agricultural Research* **14**: 163–78.
15. Calvert, D.V. (1975). Nitrate, phosphate, and potassium movement into drainage lines under three soil management systems. *J. Environ. Qual.* **4**: 183–86.
16. Johnson, W.R., Ittihadieh, F., Daum, R.M., and Pil, A.F. (1965). Nitrogen and phosphorus in tile drainage effluent. *Soil Science Society Proceedings*, pp. 287–289.
17. Bengtson, R.L., Carter, C.E., Southwick, L.M., Willis, G.H., Fouss, J.L., and Rogers, J.S. (1993). Management of irrigation and drainage systems. In: *Integrated Perspectives*. R.G. Allen (Ed.). ASCE, New York, pp. 523–30
18. Phillips, P.A., Culley, J.L.B., Hore, F.R., and Patni, N.K. (1982). Dissolved inorganic nitrogen and phosphate concentrations in discharge from two agricultural catchments in Eastern Ontario. *Agric. Water Manag.* **5**: 29–40.
19. Ulen, B. and Persson, K. (1999). Field-scale phosphorus losses from a drained clay soil in Sweden. *Hydrological Processes* **13**: 2801–12.
20. Hesketh, N. and Brookes, P.C. (2000). Development of an indicator for risk of phosphorus leaching. *J. Environ. Qual.* **29**: 105–110.

21. Kladivko, E.J., Van Scoyoc, G.E., Monke, E.J., Oates, K.M., and Pask, W. (1991). Pesticide and nutrient movement into subsurface tile drains on silt loam soil in Indiana. *J. Environ. Qual.* **20**: 264–270.
22. L.A. Zucker and L.C. Brown (Eds.). (1998). Agricultural Drainage: Water Quality Impacts and Subsurface Drainage Studies in the Midwest. *Ohio State University Extension, Bulletin 871.*
23. Kovacic, D.A., David, M.B., Gentry, L.E., Starks, K.M., and Cooke, R.A. (2000). Effectiveness of constructed wetlands in reducing nitrogen and phosphorus export from agricultural tile drainage. *J. Environ. Qual.* **29**: 1262–1274.
24. Purkey, D.R. and Wallender, W.W. (2001). Drainage reduction under land retirement over shallow water table. *J. Irrig. Drain. Eng.* **127**: 1–7.

TILE DRAINAGE: IMPACTS, PLANT GROWTH, AND WATER TABLE LEVELS

Chandra Madramootoo
Macdonald Campus of McGill University
Ste-Anne de Bellevue, Quebec, Canada

Tile drainage is a method of artificial drainage, which lowers the water table and reduces water logging on productive agricultural lands. Improved drainage is an agricultural, rather than a land conversion, practice. The objective is to increase production efficiency, crop yields, and profitability on poorly drained agricultural lands. In the wetter, humid regions, precipitation will cause the water table to rise at, or near, the soil surface. If high amounts of precipitation occur during the growing season, the water table could rise within the crop root zone. The same situation occurs in the drier, arid, and semi-arid regions where irrigation is extensively practiced. Continual irrigation and canal seepage over the years lead to an accretion of the water table, and this could also be injurious to plant growth.

Every soil has some degree of natural drainage. However, if the rates of precipitation or irrigation exceed the natural drainage rate of the soil, the high water table conditions will impair crop growth, because of a lack of oxygen in the root zone. Artificial, or tile drainage, is therefore required to remove excess soil water and to lower the water table, to maintain an optimum soil–air environment within the crop root zone. Most of the world's highly productive agricultural soils require either surface drainage or tile drainage, or a combination of both, to maintain stable crop yields. It is estimated that there are about 200 million hectares of cropland worldwide with improved drainage. The benefits of drainage are given in Fig. 1.

- Ability to better manage and control the water table for enhanced crop productivity
- Reduced flooding
- Better soil water conditions for seedbed preparation and planting
- Earlier and more uniform seed germination
- Improved field machine trafficability
- More favorable air and salt balances in the crop root zone
- Higher crop yields and reduced yield variability
- Ability to plant higher value crops
- Higher farm incomes
- Increased land values
- Less soil erosion
- Better uptake of nutrients
- Improved soil structure
- Enhanced soil microbiological activity
- More timely field machine operations
- Ability to effectively lengthen the growing season, and make better use of lower heat units in northern regions
- Potential to improve water quality

Figure 1. Benefits of tile drainage.

Tile drainage (also known as subsurface pipe drainage) is the installation of horizontal, corrugated, perforated plastic pipes below the soil surface to enable water table drawdown. Parallel lateral pipes are normally of 75 to 100 mm in diameter and are installed on grades varying from 0.1% to 3%. The downstream ends of the laterals are connected to a buried collector pipe. The diameter of the collector increases with the area drained. The spacing between the lateral pipes depends on the soil type and the amount of water to be removed. The lateral pipes can be installed in a parallel layout, herringbone layout, or at random. Depending on topography, land formation, and proximity of a water receiving body, the collector may outlet by gravity to a watercourse, or into a sump. In the case of the latter, the discharge is then pumped to a lake or stream. Before the advent of plastic pipe in the late 1970s, clay tile pipes were used. However, plastic pipes are lighter and easier and faster to install. The drainage coefficient (or volume of water to be removed by the drainage system) in humid regions is about 10–15 mm/day.

One adaptation of tile drainage is the use of surface inlets to remove water ponded in surface depressions, or as conservation structures to reduce surface runoff and erosion on sloping lands. The drainage coefficient for these systems ranges from 15 to 20 mm/day.

In the irrigated regions, tile drainage not only controls the rise of the water table, but also reclaims salinized soils. The salinity of irrigated soils could increase with time, and this also reduces crop yields. Water in excess of the plant evapotranspiration requirement is applied during irrigation. This additional quantity of water is known as the leaching fraction. Salt is then leached from the root zone and removed from the field via the tile drains. Deeper drain installation ensures that salts do not rise to the root zone because of capillary action, which also prevents waterlogging of the root zone. The amount of surplus irrigation water to be removed by the tile drainage system is less than in humid regions. Consequently, lateral pipe drainage spacings are wider than in the humid regions. The drainage coefficient in irrigated regions varies from 2 to 5 mm/day.

Horizontal pipe drainage systems are usually installed with either a trenchless plow or a chain or wheel type trencher. The trenchless plow is widely used in Europe and North America and results in faster installation rates. The depth and grade of the pipe installation equipment is achieved with a horizontal laser grade control system. In cases in which installations are conducted in fine sand and silty soils, the pipes are wrapped with a synthetic geotextile, to minimize blockage of the drainage system by the soil particles.

Another form of subsurface drainage is known as vertical drainage or tubewell drainage. It primarily controls waterlogging and salinity in the irrigated regions. The primary purpose of tubewells is to abstract groundwater for irrigation, and consequently, the water table is lowered, and salinization caused by capillarity is minimized. This situation is ideal where the groundwater is not very brackish, or saline, and therefore suitable for irrigation. In areas where the groundwater is saline, then the pumped irrigation water must first be mixed with fresher, or sweeter, water in irrigation canals.

When subsurface pipe drainage reclaims salinized and waterlogged lands, it is viewed as an environmentally beneficial practice, because degraded lands that were once intensively irrigated and cropped are returned to their full productive potential.

Drainage installations result in both hydrologic and land use changes. On poorly drained lands, surplus water is either ponded on fields or moves as surface runoff to nearby open water bodies. On tile drained lands, water percolates below the root zone, which reduces surface ponding and runoff. Infiltrated water is stored in the soil profile and will eventually percolate to the deeper groundwater table.

Some of the most significant technological developments in tile drainage occurred in the period 1965–1975. Beginning in the 1980s, environmental concerns about drainage water quality emerged. High concentrations of nitrates were being measured at the outlet of tile systems. Drainage practitioners responded to these concerns with new technologies such as watertable management, controlled drainage, and subirrigation. Tile drainage is now regarded as a multifunctional water management practice, with systems being designed to not only serve crop needs, but also minimize environmental impacts.

Controlled drainage and subirrigation make use of the existing tile drainage system to raise the water table during certain periods of the year. By either restricting the outflow of water from the tile outlet or by pumping water back into the drainage system, particularly during the growing season, the crop can obtain water from the water table to meet its evapotranspiration requirement. This process helps to reduce the effects of drought stress during the growing season, thus increasing crop yields. An additional important benefit is that the amount of nitrates leaving the tile outlet is reduced, because of a process known as denitrification.

The benefits of tile drainage go beyond the control of excess soil water and salts in the crop root zone. Environmental and socioeconomic benefits are associated with vector control and public health. For example, drainage of stagnant water eliminates malaria, foot-rot in large animals, yellow fever, the liver fluke snail, and other waterborne diseases. The overall benefit is a healthier human population. Drainage also reduces or eliminates mildew infections and various root rots of plants.

MEASURING AND MODELING TREE AND STAND LEVEL TRANSPIRATION

J.M. Vose
G.J. Harvey
K.J. Elliott
B.D. Clinton

(from *Phytoremediation: Transformation and Control of Contaminants*, Wiley 2003)

SUMMARY OF PRACTICAL IMPLICATIONS

Transpiration is a key process in the application of phytoremediation to soil or groundwater pollutants. To be successful, vegetation must transpire enough water from the soil or groundwater to control or take up the contaminant. Transpiration is driven by a combination of abiotic (climate, soil water availability, and groundwater depth) and biotic (leaf area, stomatal functions, root amount and distribution, and hydraulic characteristics) that need to be evaluated when considering appropriate site and species combinations. The protocols are not trivial, but transpiration can be measured at a variety of scales using techniques such as direct measurements of sap flow on individual trees, eddy flux gradient analyses, or gauged watersheds. Alternatively, models can be used to estimate transpiration, but these usually require on-site calibration or parameterization to produce accurate predictions. Case study analyses across a range of site conditions and species indicate a maximum transpiration capacity of approximately 7.5×10^6 liters of water per hectare per year (8×10^5 gallons of water per acre per year), with a range of 1.5×10^6 to 7.5×10^6 liters per hectare per year (1.6×10^5 to 8×10^5 gallons per acre per year). Variation among sites is related to species, tree size, and stocking (i.e., vegetation density) differences. Application of a physiologically based and site-specific parameterized model suggests reasonable agreement between measured and predicted transpiration estimates for the Air Force Plant 4 site in central Texas.

IMPORTANCE OF ACCURATE MEASUREMENTS OF TRANSPIRATION

Transpiration—the amount of water used by a tree or stand of trees—is one of the key processes in the application of phytoremediation of soil water or groundwater pollutants. To be successful, native or planted vegetation must transpire enough water from the soil or groundwater layer containing the pollutant to control the transport or decrease the mass of contaminant. Hence, quantifying current and future transpiration and determining the principal location of water uptake by native and planted vegetation on the site must be the evaluation criteria for applying phytoremediation. Quantifying transpiration requires a thorough and accurate assessment of water use patterns such as, transpiration rates, depth of soil water uptake, interactions with climate, and soil water availability. Measuring current transpiration or predicting future transpiration is not trivial. Because transpiration is an integrated response of the atmosphere–plant–soil continuum, measurements and predictions of transpiration capacity must account for (1) variation in climatic driving variables (i.e., solar radiation, water vapor saturation deficit, precipitation, wind speed, and temperature), (2) structural and physiological (leaf stomatal function) characteristics of the vegetation (leaf surface area, and root area and extent), and (3) soil water dynamics (water-holding capacity, and permeability).

Evapotranspiration and transpiration are often used interchangeably, but these processes are different. Evapotranspiration includes the amount of water transpired by the vegetation, and losses due to evaporation of intercepted precipitation and soil surface evaporation. In forests, interception evaporation is a function of rainfall intensity and leaf and branch surface area, ranging from about 10 to 50 percent (1–3). In closed canopied forests, soil evaporation is a minor component of the overall water budget (3), but may become increasingly important in open stands. From a phytoremediation perspective, transpiration is the key factor to consider because interception evaporation does not involve soil water or groundwater.

The process of transpiration involves water movement through the soil, roots, stems, and leaves into the atmosphere in response to water potential gradients—always moving in the direction of smaller potential or negative gradients. Water potential is near zero when water is freely available and decreases to negative values when water becomes more limiting. The movement of water from the leaf interior to the atmosphere occurs through small openings in the leaf called stomata, which open and close in response to external (e.g., climatic factors) and internal (e.g., water potentials of leaves) driving variables. Species vary considerably in stomatal responses to these driving variables and provide opportunities for selecting species to optimize transpiration in different climatic environments.

Five methods are used to quantify transpiration: (1) precipitation minus runoff on gaged watersheds, (2) energy balance (e.g., Penman-Monteith equation), (3) eddy covariance, (4) hydrologic models, and (5) direct sap-flow measurements. The first three methods are integrated estimates for the entire vegetation–soil complex and provide estimates of evapotranspiration not transpiration. Hence, those methods do not directly partition water losses based on transpiration *versus* evaporation and provide no information on the source of water (i.e., shallow *versus* deep soil layers) for transpiration. Estimating transpiration with methods 1, 2, and 3 requires an independent analysis of the contribution of interception and soil surface evaporation. Hydrologic models vary considerably in complexity, ranging from very simple models [e.g., Thornthwaite (4) indices of potential evapotranspiration] to detailed physiologically based models that link vegetation, soils, and the atmosphere (3). In contrast, sap-flow measurements provide a direct measure of transpiration (after correcting for time lags) under field conditions at the individual tree level (5–7). However, modeling or other scaling approaches are required to extrapolate tree-level measurements to the stand.

In summary, there are numerous approaches to quantifying transpiration in native or plantation-derived vegetative ecosystems. However, these methods vary considerably in accuracy, in data and measurement requirements, and in the capability to predict future transpiration rates as stands develop. In this chapter, we review approaches to quantifying forest transpiration from the leaf level to the stand and discuss the pros and cons of different approaches. We then provide applications of a subset of these approaches from phytoremediation case studies in Texas, Colorado, and Florida.

OVERVIEW OF CONTROLS ON TRANSPIRATION

Transpiration rates vary considerably among species and geographic regions (Fig. 1). Which factors contribute to this variation? At large scales (i.e., regions), climate is an overriding control. The strong relationship between evapotranspiration and precipitation (Fig. 1) suggests that transpiration is principally limited by soil water supply. However, other climatic factors such as temperature, atmospheric vapor pressure deficit, and solar radiation also play important roles and interact with soil water availability and physiological status of the plants (Fig. 2). For example, one of the key effects of temperature is through the influence on the length of growing season, in which longer periods with temperatures above freezing promote longer leaf area duration and hence, surface area available for transpiration. Frozen or cold soils also restrict transpiration (9) by limiting the permeability of cell membranes (10,11). Solar radiation provides the

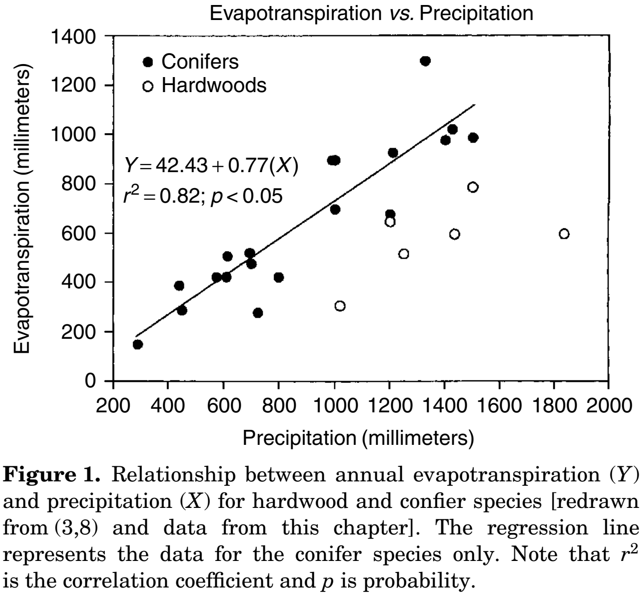

Figure 1. Relationship between annual evapotranspiration (Y) and precipitation (X) for hardwood and confier species [redrawn from (3,8) and data from this chapter]. The regression line represents the data for the conifer species only. Note that r^2 is the correlation coefficient and p is probability.

energy for transpiration and regulates stomatal opening. As a result, a strong relationship generally occurs between solar radiation and transpiration, estimated as sap flow in Fig. 2. Atmospheric vapor pressure deficit provides the gradient to which leaf-water vapor responds through the leaf stomata (Fig. 2), and wind speed has a direct influence on the leaf boundary layer (12). Optimal climatic conditions for transpiration include high soil water availability, high solar radiation, high vapor pressure deficits, warm temperatures for extended periods, and high wind speed. In most cases, these conditions do not occur simultaneously because increased soil water availability is usually a result of high rainfall that decreases solar radiation (due to increased cloud cover) and vapor pressure deficit (due to higher humidity). Species that have the ability to utilize deeper sources of soil or groundwater [i.e., phreatophytic vegetation such as poplar (*Populus* spp.) and willow (*Salix* spp.)] are an especially attractive option in hot, dry, and windy environments in the southwestern U.S., because transpired water can be derived from groundwater (13,14). Several studies have evaluated the influence of phreatophytes on surface and groundwater [e.g., (15–17)] from the perspective of negative impacts on streamflow and groundwater recharge. From a phytoremediation standpoint however, the high water consumption of phreatophytes has a positive effect to decrease aquifer recharge and influence the movement of contaminated shallow groundwater.

The structure, morphology, and physiological characteristics of the vegetation are also important regulators of transpiration. For example, at equal precipitation inputs, there are large differences in transpiration between conifer and hardwood species (Fig. 1), with hardwoods generally lower than conifers. Causes for these coarse scale differences are generally well known. The single greatest controlling factor is the quantity of leaves, expressed as leaf area index (in square meters per square meter). Site water availability and leaf area are related in that, sites with the

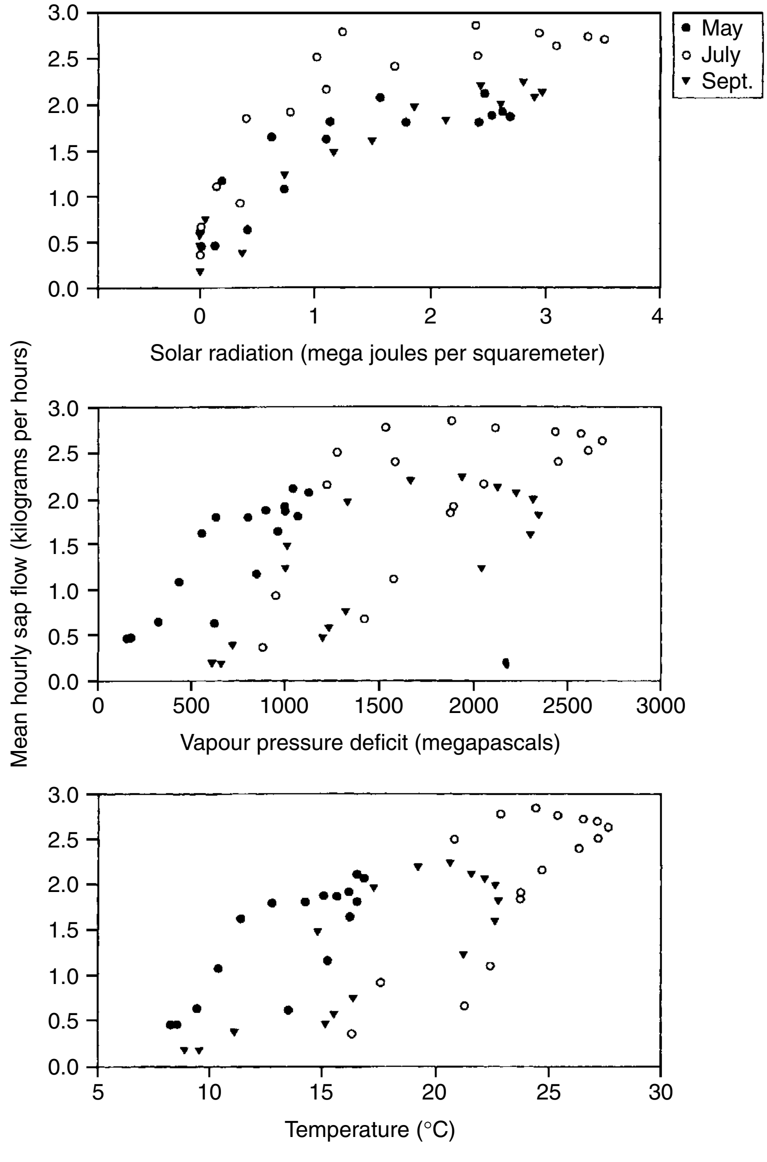

Figure 2. Mean hourly sap flows *versus* climatic driving variables for three seasonal measurement (May, July, and September) periods in central Colorado.

greatest water availability typically have the highest leaf area index (18,19), although nutrient availability (20,21) and temperature also play a role (22,23) in determining the maximum leaf area index. Watershed studies have documented strong relationships between leaf area and streamflow, with streamflow increasing exponentially as leaf area decreases (24,25). Because precipitation minus streamflow is an estimate of evapotranspiration at watershed scales, the implication is a direct control of stand-level transpiration by leaf area. Other structural and physiological factors regulating transpiration include the amount and permeability of sapwood and stomatal characteristics such as conductance and responsiveness to climatic variation and overall plant water status (i.e., water potential of leaves). Differences among species in leaf area, the rate of attainment of maximum leaf area, and physiological characteristics regulating the rate of water movement through the plant (sapwood amount and permeability, and stomatal conductance) provide opportunities for manipulating vegetation composition and structure to optimize transpiration. Optimal structural and physiological conditions for high transpiration amounts include rapid development of high leaf area, high stomatal conductance and sapwood permeability, and physiological characteristics that facilitate rapid responses to climatic conditions promoting transpiration.

Because transpiration is a function of root uptake from the soil and groundwater, soil characteristics are an important factor determining transpiration. Root growth and volume of soil occupied by roots are also important because water movement is slow when soils are drier than field capacity. Several factors determine soil water availability. First, the amount of precipitation entering the soil is a function of infiltration rate. Soils with low infiltration rates due to factors such as compaction or fine texture will have lower soil water availability because some precipitation may move across the soil surface in overland flow. Once in the soil, soil water availability is a function of water holding capacity and unsaturated hydraulic conductivity, both of which are determined by soil texture. Texture impacts water availability in different ways. Heavy clay soils (e.g., pore size less than 0.2 micrometer) have limited soil water availability because of very low rates of movement in the soil (i.e., conductivity) due to the fine pore space. In contrast, coarse textured sandy soils (e.g., pore size greater than 50 micrometers) have low water availability because of rapid drainage. Rooting volume and the presence or absence of restrictive layers are also important soil factors determining transpiration. For example, compacted soils provide a physical barrier to root growth, limiting root extension (26).

QUANTIFYING TRANSPIRATION

Leaf Level

Because water exits the plant primarily through leaf stomata (a small amount of cuticular transpiration may also occur in stems of some species), leaf–water relations are a key factor determining whole-plant transpiration (27). The concentration gradient of water vapor between the interior of the leaf and the atmosphere at the leaf boundary layer defines the maximum transpiration rate. Vapor exchange is also determined by the opening size of the stomata. When stomata are wide open, transpiration occurs at about 20 percent to 40 percent of the rate of evaporation of open water (28), whereas closed stomata limit transpiration to less than 1 percent of open water. Stomatal opening is controlled by guard cell turgor, which responds to light, temperature, vapor pressure, and water potential of the leaves. The rate of movement of water through the stomata is the stomatal conductance. The rate of stomatal response to climatic conditions varies by species, but generally reflects responses to current conditions, whereas stomatal responses to water potential in leaves may reflect previous climatic and environmental conditions.

Because of the tight linkage between transpiration, leaf stomatal conductance (hereafter referred to as leaf conductance), and water potential of leaves, knowledge of all three parameters is useful for evaluating transpiration capacity. For example, species that exhibit high leaf-level transpiration and conductance, and maintain high water potential in the leaves have the capacity to transpire large quantities of water. Similarly, the relationship between water potential in the leaves and conductance is often threshold dependent; i.e., species that maintain high leaf conductance at low water potential have the capacity to transpire more water under dry conditions (29). Because of the importance of factors such as leaf area index and distribution, sapwood amount and permeability, and the difficulty in extrapolating spatially and temporally from the leaf to stand level, there may be no direct correspondence between leaf-level transpiration and overall stand transpiration. Typically, leaf conductance and tree and stand-level transpiration are most highly related in young stands with simple canopy architecture (7), such as closely spaced, even-aged monocultures. However, as stands develop, the linkage between leaf conductance and tree or stand-level sap flow declines due to shifts in the importance of stomatal *versus* boundary layer conductance to total vapor phase conductance (30,31). Hence, leaf-level measurements should only be used as an indicator of transpiration capacity.

Tree Level

Transpiration at the whole-tree level represents the integrated movement of water vapor from all the leaves in the crown of the tree. As mentioned in the previous section, spatial and temporal variation severely limits extrapolation of individual leaf measurements to the tree, so more direct measurements at the tree level are required. Two approaches have typically been used. In a few instances, entire trees have been enclosed in a cuvette and the flux of water vapor calculated based on the rate of increase in humidity within the enclosure. This approach is severely limited by methodological constraints such as the size of trees, heat buildup within the cuvette, and alterations in the boundary layer and vapor pressure gradients.

Sap-flow rate and volume have also been used as an estimate of transpiration (32). Because of lags between water movement in the stem and leaf-level transpiration, sap flow is not a direct measure of transpiration, but can be corrected after accounting for lags (33,34). Typically, a 1- to 2-hour lag correction is applied to real time sap-flow data to account for this temporal difference (7,34).

Two sap-flow techniques have been utilized; heat balance and heat pulse. For the heat-balance approach, collars consisting of a heating element and thermocouples above and below the heating element are placed around the stem and the entire stem section is heated. Sap flow is calculated using the heat-balance principle based on the difference in temperature between thermocouples above and below the heated stem section, after subtracting for heat loss due to conduction by stemwood (35). An advantage of this approach is that it integrates sap flow along the entire stem and does not require an independent estimate of sapwood area. For larger trees, paired probes are inserted vertically into the sapwood (36). The upper probe is heated and both contain thermocouples. The probes measure heat dissipation, which increases with sap flow and the resultant cooling of the heat source, as the apparent thermal conductance of sapwood increases with sap velocity. To convert sap velocity to sap flow rate, the cross-sectional area of sapwood must also be determined. Typically, trees are cored and sapwood to heartwood ratios quantified. Because sap-flow probes measure sap flow velocity at only one location, multiple probes are required to adjust for the variation in sapwood thickness and permeability in the stem section. Despite this, unaccounted for variation in horizontal and vertical variation in sapwood thickness and permeability introduces some error into sap-flow estimates obtained with probes. The magnitude of error can be determined experimentally and corrected for in small trees by comparing sap flow with actual transpiration using procedures such as weighing lysimeters. In large trees, corrections are much more difficult and hence, predictions have more uncertainty. In contrast to the heat-balance method, the heat pulse method estimates sap flow based on the time lag between pulses of heat and the distance between the sensors (37).

Stand Level

While it is informative to understand transpiration at the leaf and tree level to help evaluate species and environments suitable for phytoremediation, stand-level transpiration ultimately determines how much soil water and groundwater are removed. However, unlike leaf and tree measurements, no methods directly measure stand transpiration. Instead, three indirect measurement approaches have been utilized. These approaches involve gaged watersheds, extrapolation of individual tree measurements, and eddy flux estimates. Gaged watersheds require a combination of well-defined watershed boundaries, tight bedrock, and well-constructed weirs or gages to provide accurate transpiration estimates. If these criteria are met, then evapotranspiration (ET) is estimated by the equation

$$ET = P - RO \pm \text{soil water storage} \qquad (1)$$

where P = precipitation and RO = runoff, determined from weirs or gages. Because P is a component of the equation, the accuracy of precipitation measurements will also influence evapotranspiration estimates. Changes in soil water storage are usually assumed to be negligible at annual time steps, although this is clearly not the case over shorter intervals. Hence, using this approach at time steps less than a year requires determining changes in soil water storage. Because evapotranspiration is estimated, interception evaporation must be determined and subtracted to estimate transpiration.

Extrapolating individual tree measurements to the stand can be done in a number of ways. For example, instruments that measure sap flow can be installed on trees representing the averaged sized tree and mean sap flow multiplied times the number of trees in the stand (i.e., a “mean-tree” approach). Considerable uncertainty in stand-level estimates can accompany this approach where sites are variable. Alternatively, relationships between tree diameter, sapwood area, or basal area and sap flow at the individual tree level can be applied to all trees. In both approaches, repeated sampling is required to account for seasonal variability.

The eddy flux method uses water vapor gradients at fixed intervals above and below the canopy to calculate evapotranspiration. The technique is based on the assumption that water vapor flux is proportional to the vertical gradient of water vapor between two measurement points (averaged over several minutes). Typically, measurements are conducted from towers extending through the canopy. To be useful for estimating transpiration of a particular stand, the stand must be large enough to encompass most of the footprint measured by the sensors. In many phytoremediation applications conceived as of 2003, the stands are too small for an eddy flux approach to be appropriate.

Modeling

The use of modeling provides a potentially powerful tool for predicting current transpiration of native or planted vegetation and for projecting future transpiration capacity as a function of stand development. At the coarsest level of forecasting, gross measures of plant water demand and use can be derived from empirical estimates of potential evapotranspiration (4,38). These approaches usually consider climate and soils to some extent, but do not consider vegetation effects such as leaf area index, rooting depth, or leaf-level physiological characteristics. Hence, empirical approaches are useful for gross estimates of transpiration, but have limited utility for evaluating actual effects on the groundwater. At the other extreme, detailed physiological models that link the soil–plant–atmosphere continuum provide much more accurate estimates of transpiration. Depending upon the structure, models may also provide estimates of specific uptake locations within the soil profile (3,39,40). Using detailed physiologically based models results in significantly greater data requirements. The most accurate application of these models requires site-specific estimates of soils, climate, and physiological characteristics of the major species on the site. However, large-scale application of detailed models with generalized parameters may provide estimates sufficiently accurate to be used in evaluating phytoremediation applications.

MEASURING AND MODELING TRANSPIRATION: CASE STUDY APPLICATIONS

Study Site Descriptions

Sap flow was measured at sites in Texas, Florida, and Colorado as components of larger studies evaluating the efficacy of using phytoremediation technology to clean up shallow groundwater contaminants. The north–central Texas study site was located, about 15 kilometers west of Fort Worth. The climate of this area is characterized as subhumid, with mild winters and hot, humid summers. The average annual precipitation is 80 centimeters per year with most rainfall occurring between May and October. Average annual temperature is 18.6 °C. Study plots were located on the U.S. Naval Air Station, which adjoins U.S. Air Force Plant 4. A plume containing trichloroethylene was detected in the terrace alluvial aquifer in 1985. To demonstrate phytoremediation potential, eastern cottonwood (*Populus deltoides* Marsh.) trees were planted in two plantations over the TCE plume. One plantation was planted with vegetative cuttings (whips) and the other with 1-year-old nursery grown seedlings. Each plantation was approximately 80 by 20 meters and located perpendicular to groundwater flow in the alluvial aquifer. Sap-flow measurements were conducted using the heat-balance method (collars) in the first and second year after plantation establishment.

The eastern Florida site was located in the city of Orlando. The climate of the area is humid, with mild winters and hot, humid summers. The average annual temperature is 22.6 °C and the average annual rainfall is 123 centimeters. Native vegetation of interest was located on the U.S. Naval Training Center. Trichloroethylene and tetrachloroethylene, which originated from a dry-cleaning facility that is no longer in operation, contaminate shallow groundwater. The plume extends under a 2-hectare forest and seepage wetland before reaching Lake Druid that borders the forest. A dense and diverse mix of overstory and understory species occur in the forest (density of 107 trees per hectare), with red bay [*Persea borbonia* (L.) Spreng.], camphor [*Cinnamomum camphora* (L.) Nees & Eberm.], slash pine and longleaf pine (*Pinus* spp.), sweet bay (*Magnolia virginiana* L.), and live oak and laurel oak (*Quercus* spp.) most abundant in the overstory. The most abundant understory species are skunk vine (*Paederia foetida* L.), saw palmetto [*Serenoa repens* (Bartr.) Small], cinnamon fern (*Osmunda cinnamomea* L.), and Christmas fern [*Polystichum acrostichoides* (Michx.) Schott.].

The central Colorado site is located approximately 20 kilometers southwest of Denver. The climate of the area is dry, with warm summers and cold winters. Annual precipitation averages approximately 44 centimeters, with 30 percent of this amount received in April and May. The average annual temperature is 12 °C. Study plots were located on the U.S. Air Force Plant PJKS. Trichloroethylene and dichloroethylene from a variety of sources contaminate the site. Measurements were conducted in two existing stands of natural vegetation: cottonwood–willow (*Populus* spp.–*Salix* spp.) and Gambel oak (*Quercus gambelii* Nutt.) The cottonwood–willow (*Populus* spp.–*Salix* spp.) stand is restricted to riparian areas (approximately 1 percent of the total land area of the site), while the Gambel oak (*Quercus gambelii*) stand is on more midslope locations (approximately 30 percent of the total land area of the site).

Methods

The sampling approach and methods varied among the three studies based on study objectives, species composition, and tree sizes. For the Texas study, sap flow from saplings in the plantation was estimated using sap-flow gauges (Dynamax Inc., Houston, TX) on 14 to 16 trees (divided equally among whips and 1-year-old trees) in May, June, July, August, and October over a 2-year period. During each measurement period, sap-flow measurements were taken every minute for 2 to 3 consecutive days. Data presented in this chapter represent averages of both plantations. In addition, sap flow was measured on nine larger native trees growing near the plantations using thermal dissipation probes (Dynamax, Inc., Houston, TX). Species sampled were: eastern cottonwood (*Populus deltoides* Marsh.), American elm (*Ulmus americana* L.), black willow (*Salix nigra* Marsh.), sugarberry [or large hackberry, (*Celtis laevigata* Willd.)], Eastern red cedar (*Juniperus virginiana* L.), and mesquite (*Prosopis pubescens* Benth.). At the end of sampling, increment cores were taken from the nine large trees for determining sapwood area.

For the Orlando study, sap flow was estimated using thermal dissipation probes installed on nine trees representative of major canopy species. Species sampled were: slash pine (*Pinus elliottii* Engelm.), longleaf pine (*Pinus palustris* Mill.), live oak (*Quercus virginiana* Mill.), laurel oak (*Quercus hemisphaerica* Bartram ex. Willd.), sweet bay (*Magnolia virginiana* L.), and camphor [*Cinnamomum camphora* (L.) Nees & Eberm.]. Two probe sets were installed into the sapwood on the north and south sides of sample trees, and sampling was conducted in November, March, and July for 2 to 3 consecutive days over a 1-year period. At the end of sampling, increment cores were taken and sapwood area determined.

For the Colorado study, sap flow was estimated using thermal dissipation probes on eight trees representing three species: eastern cottonwood (*Populus deltoides* Marsh.), narrow-leaf cottonwood (*Populus angustifolia* James.), and Gambel oak (*Quercus gambelii* Nutt.). Two probe sets were installed into the sapwood on the north and south side of sample trees and sampling was conducted in May, July, and September over a 1-year-period. At the end of sampling, increment cores were collected and sapwood area determined.

For all three studies, data were summarized to provide average hourly sap flow rates (kilograms per hour) or daily totals (kilograms per day). In addition, climate was measured at all three studies with climate stations located on-site. Measurements included: hourly rainfall (centimeters), wind speed (meters per second), solar radiation (Watts per square meter), temperature (°C), and relative humidity (percentage). Relative humidity and air temperature were used to calculate vapor pressure deficit (megapascals).

For the Texas plantation site, we parameterized and applied a mechanistic model of sap flow (PROSPER) and compared the results to sap flow measurements. Evapotranspiration at the Texas site was simulated because data were available to parameterize the model (7). The PROSPER model has been described in detail elsewhere (39,41), so only a general description is provided here. The PROSPER code is a phenomenological, one-dimensional model that links the atmosphere, vegetation, and soils. Plant and soil characteristics are combined into a single evapotranspiration surface that is characterized by a resistance to water vapor loss. This resistance is analogous to the relationship between stomatal resistance and water potential of the leaves and is a function of the water potential of the evapotranspiration surface. Evapotranspiration is predicted by a combined energy balance–aerodynamic method [Penman-Monteith equation modified as described in (42)] that is a function of the surface resistance to vapor loss described previously. The PROSPER model uses electrical network equations (41) to balance water allocation among vegetation and soil horizons. The flow of water within and between soil and plant is a function of soil hydraulic conductivity, soil water potential, root characteristics in each soil layer, and surface water potential. The PROSPER model predicts evapotranspiration, transpiration, and soil water distribution between soil layers daily, but monthly data are most accurate. The PROSPER model requires the following climatic data: solar radiation, precipitation, wind speed, air temperature, and vapor pressure. Initial model parameters include surface resistance to vapor loss, leaf area index, root distribution and surface area, soil moisture release, and several other parameters listed in Goldstein et al. (41).

Transpiration Estimates

Maximum transpiration rates for the study sites indicate large variation in transpiration potentials among sites (Table 1). On a per tree basis, rates ranged from 8 to 120 kilograms per tree per day. Much of this variation was related to differences in tree size that reflects differences in leaf area and sapwood area. For example, when pooling the data across the sites, a significant proportion of the variation in transpiration rates among and within sites can be explained by tree diameter (Fig. 3). Larger trees typically have greater sapwood volume resulting in more water transporting vessels (angiosperms) and tracheids (conifers) for sap-flow movement in the stems. Because leaf area is also related to sapwood area, larger trees will typically have greater leaf area index; and hence, greater surface area for transpiration.

When sap-flow rates are adjusted based on sapwood area (i.e., kilograms per day per square meter of sapwood), the variation in transpiration reflects species related differences in physiology (leaf, stem, and root), leaf area to sapwood area ratios, and site-dependent factors such as soil water availability and climate driving variables. Because species composition varies among sites and physiological and physical factors influence transpiration simultaneously, these studies

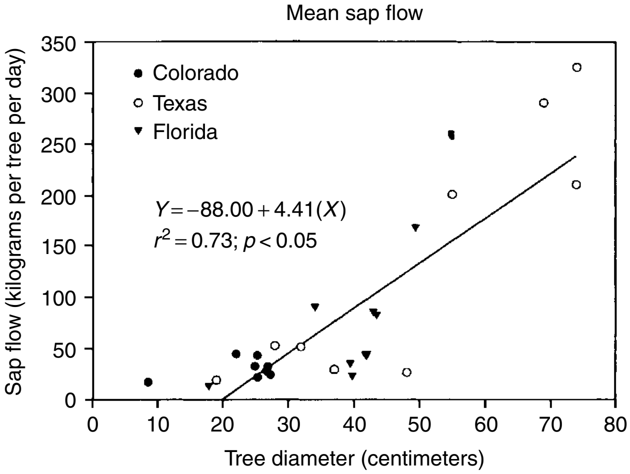

Figure 3. Total daily sap flows (Y) during peak transpiration periods (midsummer) *versus* tree diameter (X) across a range of species and site. Note that r^2 is the correlation coefficient and p is probability.

Table 1. Midsummer Peak Sap-Flow Rates Averaged Across Species and Measurement Days

		Sap-Flow Rates			
Site	Sapwood Area (Square Centimeters Per Tree)	Kilograms Per Day Per Square Meter of Sapwood	Kilograms Per Tree Per Day	Liters Per Hectare Per Year	Gallons Per Acre Per Year
Texas					
Plantation	30	2600	8	3 620 000	387 200
Native trees	820	1463	120	7 551 000	807 600
Colorado	234	1043	24	1 510 000	161 500
Florida	710	1535	109	6 859 000	733 600

cannot separate physiological and climatological effects; to do so requires an evaluation of transpiration rates of the same species and genotype in differing climatic and soil water availability conditions. For example, species sampled at the three sites represent a mixture of conifers, and ring porous and diffuse porous hardwood species, resulting in large differences in sapwood permeability and specific conductivity among sites and among species within sites (Fig. 4). In general, sap-flow velocity is lower in conifers and diffuse-porous species because sap flow moves through a number of annual rings, whereas water moves through only one or two annual rings in ring-porous species (11,43). Despite the limitations of the current approach, some notable patterns emerge when evaluating transpiration after adjusting for differences in sapwood area. For example, the cottonwood (*Populus deltoides*) plantation in Texas had the highest transpiration rate per unit of sapwood area, followed by the Florida stand, large trees in Texas, and the Colorado stand (Table 1). The high transpiration rate for cottonwood (*Populus deltoides*) in the plantation is a function of species characteristics that promote high transpiration, high leaf area per unit sapwood in the developing canopy, and access to shallow groundwater. In contrast, transpiration rates per unit of sapwood area were lowest in Colorado, even though the site contained cottonwood (*Populus deltoides*) and several of the measured trees occurred in the riparian zone. The combination of species composition and climate characteristics were not as conducive to high sap-flow rates per unit sapwood area relative to the other sites.

Using sap-flow techniques to predict actual stand transpiration requires frequent sampling to account for seasonal variation. Ideally, sap flow should be measured continuously for the entire growing season on a large number of trees. Because this approach is often impractical, an alternative is to measure sap flow at shorter frequencies and calculate bounds or maximum values as a tool to evaluate phytoremediation potentials. Because the sampling frequency varied among the case studies described here, we focused only on measurements during the highest transpiration period (midsummer). These estimates can be used as a “best-case scenario” approach—that is, if these rates occurred on the site, would transpiration be sufficient to control the plume? To estimate maximum potential transpiration at the stand level (i.e., kilograms per hectare or gallons per acre), we extrapolated the tree transpiration data (kilogram per tree per day) assuming a 180 days transpiration period and a stem density of 350 stems per hectare, except for the plantation where actual tree density was used (Table 1). We emphasize that these data provide estimates of maximum transpiration capacity under *in situ* climate conditions because the peak sap-flow rates were used in the extrapolation and previous studies have shown considerable seasonal variation in sap flow (7). The 350 stems per hectare is representative of a fully stocked stand under most forest conditions and is consistent with full canopy closure and maximum leaf area index.

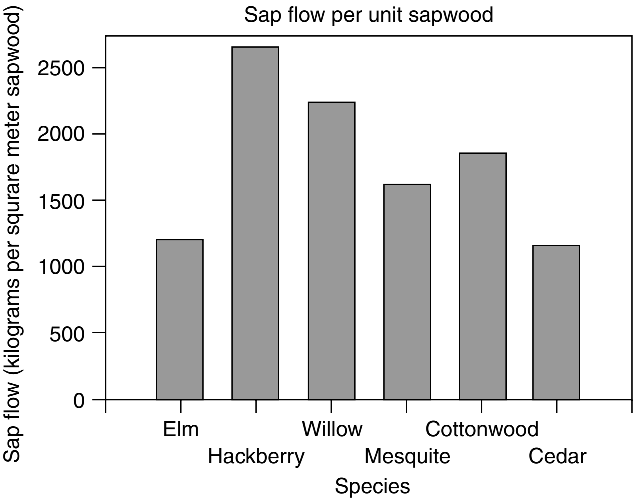

Figure 4. Mean growing season sap-flow rate per unit sapwood for six species in north–central Texas. Black willow (*Salix nigra* March.), eastern cottonwood (*Populus deltoides* Marsh.), and eastern red cedar (*Juniperus virginiana*, L.) are diffuse porous species, while American elm (*Ulmus americana* L.), hackberry (*Celtis laevigata* Willd.), and mesquite (*Prosopis pubescens* Benth.) are ring porous.

When comparing results from the sites with mature trees, the variation in maximum transpiration capacity is considerable. The Texas site has a maximum transpiration capacity of approximately 7.5×10^6 liters of water per hectare per year (8×10^5 gallons of water per acre per year) if the site was fully stocked with the sampled species. By contrast, the Colorado site has a maximum transpiration capacity of approximately 1.5×10^6 liters per hectare per year (1.6×10^5 gallons per acre per year). The plantation site in Texas currently has a maximum transpiration capacity of approximately 3.7×10^6 liters per hectare per year (4×10^5 gallons per acre per year). However, we anticipate that transpiration will equal or exceed the estimate from mature trees on the site (*versus* 7.5×10^6 liters per hectare per year or 8.0×10^5 gallons per acre per year) once the canopy develops and achieves the maximum leaf area.

Comparison of Measured versus Modeled Transpiration

A critical need for phytoremediation is the development and application of a tool to provide species and site-based estimates of transpiration. While a powerful tool for measuring transpiration from vegetation already on-site or quantifying transpiration of planted vegetation, sap-flow measurements at every phytoremediation site may not be practical. One potential tool for application across sites is the development or application of models. In most cases, however, models need to be calibrated or parameterized for specific site and species conditions. To evaluate the use of such a tool, we parameterized PROSPER for the Texas plantation using intensive site (soils, climate, and root distribution) and leaf-level measurements (stomatal characteristics and leaf area index) [see (7)]. We compared monthly transpiration estimates obtained with PROSPER to transpiration estimated from sap-flow measurements over a 2-year period (Fig. 5). Comparisons indicated generally good agreement between predicted and measured values, except during the late summer that was coincident with some of driest and hottest periods (August and September 1998) in our study. During this period, PROSPER predicted a considerable decline in transpiration, while measured values showed an increase. We attribute this discrepancy to an inability of PROSPER to adequately simulate root uptake from shallow groundwater during drought conditions, since the original formulation of PROSPER was designed to only simulate surface and soil water dynamics (39,41). The results of this comparison are consistent with other studies that have shown that PROSPER provides reasonable estimates of either evapotranspiration or transpiration (3,40). However, refinements in the subsurface water and groundwater hydrology and subsequent availability to tree roots might improve the predictive capability and usefulness as a phytoremediation evaluation tool.

CONCLUSIONS AND RECOMMENDATIONS

The importance of transpiration to the success of phytoremediation applications suggests that accurate estimates of current and potential transpiration should be a high priority when considering this approach for site management. Both measuring and modeling transpiration are important. Assessments require a substantial sampling commitment for direct measurements or parameterizing physiologically based models. This requires detailed knowledge of local site conditions and physiological parameters for the major species. For screening assessments, we recommend that published estimates be used to set the bounds for maximum transpiration capacity based on general climate and vegetation characteristics of the location. If these general transpiration rates are great enough to influence groundwater hydrology, then evaluations of current, enhanced (e.g., manipulating the structure and species composition of current vegetation), or new vegetation transpiration capacity should proceed. Technology and models exist to provide reasonable estimates and predictions of transpiration. However, the accuracy of the estimates depends on the investment in accounting for the spatial and temporal variation or in providing site and species-specific estimates for physiologically based transpiration models.

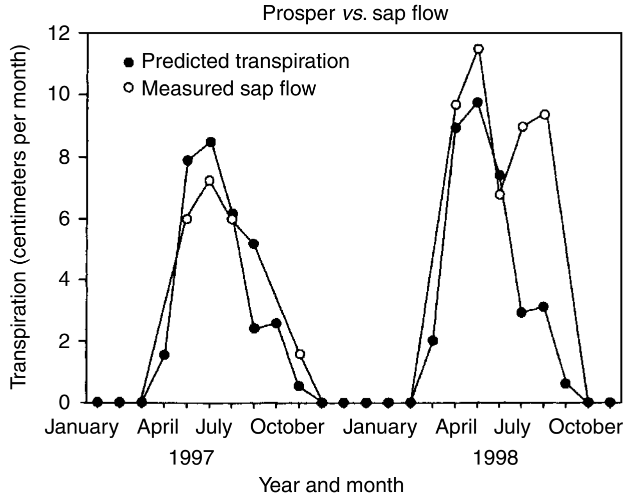

Figure 5. Comparison of measured sap flow and predicted transpiration from PROSPER for cottonwood (*Populus deltoides* Marsh.) plantations in north–central Texas.

Acknowledgments

This work was supported in part by the Department of Defense Environmental Security Technology Certification Program Project E95BRAB004, Gregory J. Harvey, Program Manager and in part by the Department of Defense, U.S. Navy (SOUTH-NAVFACENGCOM), Charleston, South Carolina. We thank Chris Sobek for her assistance in collecting sap-flow data, summarizing data, and generating graphs. Several other members of the Coweeta staff assisted with field sampling at various times, including: Bob McCollum, Nathan Stover, Patsy Clinton, Laura Phillips, and Josh Ward.

BIBLIOGRAPHY

1. Helvey, J.D. (1971). A summary of rainfall interception by certain conifers of North America. In: *Biological Effects of the Hydrologic Cycle*. E.J. Monke (Ed.). Proc. Third Int. Seminar for Hydrology Professors, Dept. of Agricultural Engineering. Purdue University, West Lafayette, Indiana, pp. 103–113.
2. Meyers, B.J. and Talsma, T. (1992). Site water balance and tree water status in irrigated and fertilized stands of *Pinus radiata*. *Forest Ecol. Manag.* **50**: 17–42.
3. Vose, J.M. and Swank, W.T. (1992). Water balances. In: *Atmospheric Deposition and Forest Nutrient Cycling: A Synthesis of the Integrated Forest Study*. D.W. Johnson and S.E. Lindberg (Eds.). *Ecological Studies* 91. Springer-Verlag, New York, pp. 27–49.
4. Thornthwaite, C.W. (1948). An approach towards a rationale classification of climate. *Geogr. Rev.* **38**: 55–94.
5. Hinckley, T.M., Brooks, J.R., Cermak, J., Ceulemans, R., Kucera, J., Meinzer, F.C., and Roberts, D.A. (1994). Water flux in a hybrid poplar stand. *Tree Physiol.* **14**: 1005–1018.
6. Martin, T.A., Brown, K.J., Cermak, J., Ceulemans, R., Kucera, J., Meinzer, F.C., Rombold, J.S., Sprugel, D.G., and Hinckley, T.M. (1997). Crown conductance and tree and stand transpiration in second-growth *Abies amabilis* forest. *Can. J. Forest Res.* **27**: 797–808.
7. Vose, J.M., Swank, W.T., Harvey, G.J., Clinton, B.D., and Sobek, C. (2000). Leaf water relations and sapflow in eastern cottonwood (*Populus deltoides* Bartr.) trees planted for phytoremediation of a groundwater pollutant. *Int. J. Phytorem.* **2**(1): 53–73.
8. Knight, D.H., Vose, J.M., Baldwin, V.C., Ewel, K.C., and Grodzinska, K. (1994). Contrasting patterns in pine forest ecosystems. *Ecol. Bull. (Copenhagen)* **43**: 9–19.
9. Fahey, T.J. (1979). The effect of night frost on the transpiration of *Pinus contorta* sp. *latifolia*. *Oecologia* **14**: 483–490.
10. Kaufmann, M.R. (1977). Soil temperature and the drying effects on water relations of *Pinus radiata*. *Can. J. Bot.* **55**: 2413–2418.
11. Kozlowski, T.T., Kramer, P.J., and Pallardy, S.G. (1991). *The Physiological Ecology of Woody Plants*. Academic Press, New York, p. 657.
12. Gates, D.M. (1980). *Biophysical Ecology*. Springer-Verlag, New York, p. 611.
13. Dawson, T.E. and Ehleringer, J.R. (1991). Streamside trees do not use stream water. *Nature* **350**(6316): 335–337.

14. Busch, D.E., Ingraham, N.L., and Smith, S.D. (1992). Water uptake in woody riparian phreatophytes of the southwestern United States: a stable isotope study. *Ecol. Appl.* **2**(4): 450–459.
15. Robinson, T.W. (1970). *Evapotranspiration by woody phreatophytes in the Humbolt River Valley near Winnemucca*, Nevada. U.S. Geological Survey Professional Paper 491–D, U.S. Government Printing Office, Washington, DC.
16. Van Hylckama, T.E.A. (1980). *Weather and evapotranspiration studies in a saltcedar thicket in Arizona*. U.S. Geological Survey Professional Paper 491–494, U.S. Government Printing Office, Washington, DC.
17. Allen, S.J., Hall, R.L., and Rossier, P.T. (1999) Transpiration by two poplar varieties grown as coppice for biomass production. *Tree Physiol.* **19**: 493–501.
18. Gholz, H.L. (1982). Environmental limits on aboveground net primary production, leaf area and biomass in vegetation zones of the Pacific Northwest. *Ecology* **53**: 469–481.
19. Long, J.N. and Smith, F.W. (1990). Determinants of stemwood production in *Pinus contorta* var. *latifolia* forests: the influence of site quality and stand structure. *J. Appl. Ecol.* **27**: 847–856.
20. Vose, J.M. and Allen, H.L. (1988). Leaf area, stemwood growth, and nutrition relationships in loblolly pine. *Forest Sci.* **34**: 546–563.
21. Colbert, S.R., Jokela, E.J., and Neary, D.G. (1990). Effects of annual fertilization and sustained weed control on dry matter partitioning, leaf area, and growth efficiency of juvenile loblolly and slash pine. *Forest Sci.* **36**: 995–1014.
22. Gholz, H.L. (1986). Canopy development and dynamics in relation to primary production. In: *Crown and Canopy Structure in Relation to Primary Production*. T. Fujimori and D. Whitehead (Eds.). Forest and Forest Products Research Institute, Ibaraki, Japan. pp. 224–242.
23. Cropper, W.P., Jr. and Gholz, H.L. (1994). Evaluating potential response mechanisms of a forest stand to fertilization and night temperature: a case study using *Pinus elliottii*. *Ecol. Bull.* (Copenhagen). **43**: 165–160.
24. Douglass, J.E. and Swank, W.T. (1975). *Effects of management practices on water quality and quantity: Coweeta Hydrologic Laboratory, North Carolina*. USDA Forest Service Gen. Tech. Rep. SE-13, Asheville, North Carolina.
25. Swank, W.T., Vose, J.M., and Elliott, K.J. (2000). Long-term hydrologic and water quality responses following commercial clearcutting of mixed hardwoods on a southern Appalachian catchment. *Forest Ecol. Manag.* **5267**: 1–16.
26. Heilman, P. (1981). Root penetration of Douglas-fir seedlings into compacted soil. *Forest Sci.* **27**: 660–666.
27. Schulze, E.-D. (1991). Water and nutrient interactions with plant water status. In: *Responses of Plants to Multiple Stresses*. H.A. Mooney, W.E. Winner, and E.J. Pell (Eds.). Academic Press, New York, pp. 89–101.
28. Waring, R.H. and Schlesinger, W.H. (1985). *Forest Ecosystems: Concepts and Management*. Academic Press, New York, p. 340.
29. Zhang, H., Simmonds, L.P., Morison, J.I.L., and Payne, D. (1997). Estimation of transpiration by single trees: comparison of sap flow measurements with a combination equation. *Agric. For. Meteorol.* **87**: 155–169.
30. Heilman, P.E., Hinckley, T.M., Roberts, D.A., and Ceulemans, R. (1996). Production physiology. In: *Biology of Populus and Its Implications for Management and Conservation*. R.F. Stettler, H.D. Bradshaw, Jr., P.E. Hielman, and T.M. Hinckley (Eds.). NRC Research Press, Ottawa, Canada, pp. 459–489.
31. Martin, T.A., Hinckley, T.M., Meinzer, F.C., and Sprugel, D.G. (1999). Boundary layer conductance, leaf temperature and transpiration in *Abies amabilis* branches. *Tree Physiol.* **19**: 435–443.
32. Steinberg, S.L., van Bavel, C.H.M., and McFarlan, M.J. (1989). A gauge to measure mass of flow of sap in stems and trunks of woody plants. *J. Am. Soc. Hort. Sci.* **114**: 466–472.
33. Schulze, E.-D., Cermak, J., Matyessek, R., Penka, M., Zimmerman, R., Vasicek, F., Gries, W., and Kucera, J. (1985). Canopy transpiration and water fluxes in the xylem of the trunk of *Larix* and *Picea*—a comparison of xylem flow, prorometer, and cuvette measurements. *Oecologia* **66**: 475–483.
34. Phillips, N., Nagchaudhuri, A., Oren, R., and Katul, G. (1997). Time constant for water transport in loblolly pine trees estimated from time series of evaporative demand and stem sapflow. *Trees* **11**: 412–419.
35. Baker, J.M. and van Bavel, C.H.M. (1987). Measurement of mass flow of water in stems of herbaceous plants. *Plant Cell Environ.* **10**: 777–782.
36. Granier, A. (1987). Evaluation of transpiration in a Douglas fir stand by means of sap flow measurements. *Tree Physiol.* **3**: 309–320.
37. Swanson, R.H. (1962). *An instrument for detecting sap movement in woody plants*. U.S. Department of Agriculture Forest Service Rocky Mountain Forest and Range Experiment. Station Paper. 68.
38. Mongaran, C. (1973). Economic feasibility of irrigating southern pines. *Water Resources Res.* **9**: 1485–1495.
39. Huff, D.W. and Swank, W.T. (1985). Modeling changes in forest evapotranspiration. In: *Hydrological Forecasting*. M.G. Anderson and T.P. Burt (Eds.). John Wiley & Sons, New York, pp. 121–151.
40. Vose, J.M. and Swank, W.T. (1994). Effects of long-term drought on the hydrology and growth of a white pine plantation in the southern Appalachians. *Forest Ecol. Manag.* **64**: 25–39.
41. Goldstein, R.A., Mankin, J.B., and Luxmoore, R.J. (1974). *Documentation of PROSPER: a model of atmosphere-plant-soil water flow*. EDFB/IBP-73/9, U.S. Department of Energy Oak Ridge National Laboratory, Oak Ridge, Tennessee.
42. Swift, L.W., Jr., Swank, W.T., Mankin, J.B. Luxmoore, R.J. and Golstein, R.A. (1975). Simulation of evapotranspiration and drainage from mature and clear-cut forest and young pine plantations. *Water Resources Res.* **11**(5): 667–673.
43. Kramer, P.J. and Kozlowski, T.T. (1979). *Physiology of Woody Plants*. Academic Press, New York, p. 811.

WATER LOGGING: TOPOGRAPHIC AND AGRICULTURAL IMPACTS

CHANDRA MADRAMOOTOO
Macdonald Campus of McGill University
Ste-Anne de Bellevue, Quebec, Canada

Water logging occurs on poorly drained soils. Under periods of heavy precipitation or snowmelt, water percolates very slowly, and soils become wet in short periods of time. Some soils have inherently low ability to infiltrate and hydraulic conductivity, especially heavier clay soils. Soil compaction because of heavy farm equipment further reduces bulk density, which in turn reduces the ability to infiltrate soil. The result is that the water table accretes and may eventually rise to the land surface. High-intensity rainfalls also lead to ponding of water on the soil surface. These conditions of surface ponding, a saturated soil profile, and gradual rise of the water table to the soil surface all lead to water logging. One very specific hydrogeologic condition that contributes to water logging is groundwater artesian pressure.

Land topographic features have an influence on water logging. On very flat lands, with low hydraulic gradients to rivers and watercourses, surface runoff is not easily drained. Water is also easily trapped in surface depressions on undulating lands. Low-lying coastal and riverine areas are very susceptible to flooding in the wet rainy periods. Towns, villages, and agricultural lands suffer from backwater effects. Infrastructural works sometimes exacerbates water logging. For example, during heavy rainfalls, water becomes trapped behind road and railway embankments. The situation could be extremely damaging under monsoon conditions in Asia, where widespread flooding occurs with loss of life and property.

Water logging also occurs in the arid and semiarid regions on poorly drained soils, because of excess irrigation water applications and canal seepage. It is estimated that about 25–30 million hectares of irrigated land suffer from water logging.

Globally, some areas of permanently wet or water logged regions exist, which are known ecologically as wetlands, and they serve very important environmental and hydrologic functions. Wetlands provide habitats for waterfowl, flora, fish, and animals. They also replenish groundwater, provide flood control, and filter pollutants. Organic soil deposits, or peat bogs, are also characterized by naturally high water tables.

It is estimated that over 200 million hectares of agricultural land worldwide suffer from poor drainage. Crops need an optimum soil-air environment in the root zone in which to thrive. Under high water table conditions or poor drainage, crop growth is retarded because of a lack of oxygen in the root zone. Root and shoot growth are affected by high water table conditions, as are other plant physiological conditions. For example, stem elongation is slowed, senescence and abscission of older leaves occur, rapid stomatal closure exists, and leaves begin to yellow. Crop failure is inevitable under prolonged flooding. Anaerobic conditions lead to reduced oxygen, and an accumulation of carbon dioxide, ethylene, other toxic substances, and byproducts in the soil profile, all of which are harmful to plant growth.

Some plants have the ability to withstand high water tables better than others. For example, rice is grown under flooded conditions. Some berry crops, e.g., highbush blueberries and cranberries grow under saturated soil conditions. Native vegetation and trees can be found in swamps, marshlands, and peat bogs. Crop tolerance to high water tables not only depends on crop type, but also on stage of crop growth. For example, most crops are likely to be killed by flooding during germination and early vegetative growth. As plants mature, they are slightly more tolerant to high water tables. However, vegetable crops are very sensitive to high water tables during their entire growth period. Deeper rooted crops, such as grasses, sugarcane, corn, shrubs, and trees, are somewhat hardier.

Apart from rice and some berry crops, it is impossible to undertake large-scale commercial plant and animal production without removal of excess soil water. Drainage is therefore essential for lowering the water table on poorly drained soils. The most common form of drainage is surface drainage. Open ditches are dug in fields to evacuate excess water. Lands are also shaped either by forming bedded fields or by a ridge and furrow system. In this way, crops can be grown on the raised beds or ridges. Other forms of land shaping include land leveling and smoothing, or grading the fields with a slope toward open drains. Surface depressions in fields can also be filled in by land leveling, and this will reduce ponded water on the land surface. Tile drainage, or subsurface pipe drainage, is another form of drainage for improving the productivity of poorly drained agricultural lands.

Tile drainage or tubewells are also used to lower the water table in irrigated lands. If seepage from irrigation canals is determined to be a major contributor to the water logging of adjacent lands, the canals can be lined, or buried pipe interceptor drains are installed adjacent to the canal. Relief wells are normally installed in lands affected by artesian uplift.

In the lowing coastal and riverine areas, flood control structures are built, which include levees, dykes, walls, and berms. Sluice gates or pumping stations are constructed to evacuate water from behind these structures to either the sea or rivers.

WEED CONTROL STRATEGIES

J.R. Newman
IACR-Centre for Aquatic Plant
Management
Reading Berkshire, United
Kingdom

The control of nuisance vegetation in watercourses is often necessary to ensure adequate flood defense for surrounding land, to provide facilities for recreational sporting activity, to provide better angling environments, to aid navigation, on public health grounds, for industrial uses, and to maintain adequate water supply and quality. In recent years, weed control for conservation has also become necessary to protect native habitats from invasion by alien, introduced, or nonnative species.

Aquatic weeds are different from terrestrial weeds in that they are usually a normal constituent of the aquatic ecosystem. Aquatic plants become weeds when they grow to excess and become a nuisance by blocking channels, reducing flow, or preventing other normal uses of the waterbody. Increased urbanization has resulted in an increase in the number of nonnative species now considered aquatic weeds, due primarily to escape of plants from domestic situations. The increased recreational pressure on enclosed inland waterways has also resulted in greater demands for aquatic plant management.

During the last 50 years, there has been a dramatic change in the type of aquatic weed problem. Emergent weeds, such as reeds and rushes, and floating leaved weeds, such as waterlilies, were the major problem, but eutrophication since the Second World War, as a result of intensification of agriculture, increased urbanization, and consequent sewage treatment requirements, has resulted in a shift to problems caused by filamentous and blue-green (cyanobacterial) algal blooms.

TYPES OF AQUATIC WEEDS

Although there are a large number of individual aquatic weed species, for the purposes of control, they can be divided into four major groups.

Emergent Plants

Emergent plants are rooted in the sediment at the margins of watercourses. They have stems and leaves that protrude above the water's surface. Growth is often limited by increasing water depth, making this type of weed common along the edges of water. Some common characteristics of these plants are long narrow leaves, and heights between 1 and 3 m. Species include *Phragmites communis* (common reed), *Typha latifolia* (bulrush), *Schoenoplectus lacustris* (common club rush), *Glyceria maxima* (reed sweet grass) and *Sparganium erectum* (common bur reed).

Smaller emergent and bankside plants, except grasses, tend to be broad-leaved weeds. These include *Rorippa nasturtium-aquaticum* (watercress) *Apium nodiflorum* (fool's watercress), *Berula erecta* (water parsnip), and others.

Other plants in this group are predominantly submerged but may form different types of leaves when they become emergent due to lowered water levels or grow above the water surface. These include some *Ranunculus* spp. (water crowfoots) and *Hippuris vulgaris* (mare's tail).

Plants that grow on banks at or above the water line should be included in this group. Common weeds in this class include *Lythrum salicaria* (purple loosestrife), *Epilobium hirsutum* (greater willow herb) and *Phalaris arundinacea* (reed canary grass).

A number of riparian weeds associated with watercourses can be included in this group. The dispersal of these species is furthered by flowing water. The alien invasive species *Fallopia japonica* (Japanese knotweed), *Heracleum mantegazzianum* (giant hogweed) and *Impatiens glandulifera* (Himalayan balsam) are included here.

Floating Plants

Free-Floating Plants. This group includes many of the worst types of aquatic weeds, such as *Azolla* spp. (water fern), *Eichhornia crassipes* (water hyacinth), *Lemna* spp. (duckweeds), *Pistia stratiotes* (water lettuce) and *Salvinia molesta* (giant water fern), all of which have been found in Britain. They are characterized primarily by very rapid vegetative reproduction and spread.

Floating-Leaved Plants. This type of plant is characterized by rooting in the sediment and forming floating leaves. They may be rooted at the margins and form floating mats over the surface of the water, such as *Hydrocotyle ranunculoides* (floating pennywort). This species has leaves held above the water surface. Plants rooted in deep water that have floating leaves held flat on the surface include *Potamogeton natans* (broad-leaved pondweed), *Nuphar lutea* (yellow water lily) and *Nymphaea alba* (white water lily).

Other plants in this group are predominantly submerged, but may form different types of leaves when they reach the water surface, such as the rosette type of leaf produced by *Callitriche* spp. (water starworts).

Submerged Plants

This group includes all submerged plants except algae. It can be divided into two sub-groups, those rooted in the sediment, such as *Elodea* spp. (waterweeds), *Myriophyllum spicatum* (spiked water milfoil), and *Potamogeton pectinatus* (fennel pondweed), or those free-floating below the water, such as *Ceratophyllum demersum* (rigid hornwort) and *Lemna trisulca* (ivy-leaved duckweed).

Algae

Algae can be divided into two types for the purpose of control, macrophytic types, including filamentous algae and charophytes, and unicellular types, including pea-soup algae and cyanobacteria or blue-green algae. Nuisance algae tend to grow best in nutrient-rich still or slow flowing water.

THE BIOLOGY AND ECOLOGY OF AQUATIC WEEDS

Seasonal Growth and Dispersal

Aquatic plants are usually a problem only during the summer months. Various forms of dieback or washout during autumn and winter ensure that very little aquatic vegetation is left to overwinter. However, nuisance aquatic weeds use several strategies to survive winter. These include seed production by annual riparian weeds and as an insurance policy by perennial submerged aquatic macrophytes. Others produce turions (specialized leaf buds) tolerant of anoxic conditions, for example, *P. pectinatus*, fennel pondweed. The majority of emergent macrophytes have dormant roots and rhizomes. Algae use spore production and *Azolla* and others can overwinter as intact plants, for example, *Callitriche* spp. and *Crassula helmsii*.

Aquatic weed growth usually starts when mean diel water temperature rises above 6°C; this is normally a week to a few weeks behind terrestrial plant growth. However, growth is rapid once started and stem elongation in *Ranunculus* spp. can produce 6 m of growth before flowering. Shortening day length usually initiates flowering so the majority of aquatic plants flower in or after the last week in June in the temperate Northern Hemisphere.

Stem fragmentation and vegetative reproduction is the most common form of dispersal within systems, and seed production is not considered important for maintaining weed populations. However, establishment of new weed populations is most likely to derive from establishment of seeds. Birds, animals, machinery, boats, and man can also spread viable fragments of weeds to new sites.

Characteristics of Nuisance Plants

Growth Rate. Species that have extremely fast growth rates tend to become greater problems than other species. For example, duckweed and *Azolla* can double their biomass in less than 4 days in good conditions.

Morphology. Species that produce dense masses of vegetation impede the flow of water. Others have rigid stems that may impede boat traffic, interfere with access to the water, and trap detritus, forming temporary dams that may cause flooding.

Dispersal Strategies. Plants that produce new plants by fragmentation are also very effective aquatic weeds. Often, mechanical control is responsible for increases in the distribution of these species in watercourses.

Toxicity and Taint. Some plants are toxic to animals, and others taint potable waters. Blue-green algae produce toxins that can kill fish, and other animals and also impart an unpleasant taste to drinking water.

CONTROL METHODS

Factors Affecting Choice

A number of factors influence the choice of weed management technique, including the cost of control. These factors are related to the uses of the water and the specific landscape situation of the waterbody. Often, a compromise between weed control efficacy and other considerations must be made.

Type of Control. In some cases, it is not desirable to attempt complete eradication of the weed population, and localized control is preferable. Selective control can be either by type of species or by, for example, clearing swims for anglers. Total control may be required where flood defense is an issue or where nonindigenous species are present. Temporary alleviation of a weed problem is sometimes possible by frequent mechanical control during the season.

Use of the Water. Certain weed control practices may not be compatible with the use of the water. The most obvious of these is when water is abstracted for drinking water supplies downstream of an intended herbicide application site. The limit of $0.1\,\mu g\,L^{-1}$ for individual pesticides set by the EU Drinking Water Directive imposes strict requirements on the distance between the site of application and the abstraction point. Dilution and flow rate affect this distance.

Human and Livestock Safety. This includes direct toxicity of herbicides to operators and water users, increases in turbidity due to stirring up of sediments, and unpleasant odors from rotting vegetation or herbicide residues in water. The safety of farm and other animals can be compromised by the direct toxicity of the herbicide, or, as is more likely, by increasing the palatability of treated poisonous plants.

Irrigation. The hazards to adjacent crops should be considered through irrigation with water containing pesticides or spray drift at the time of application and the spread of weeds such as *P. australis* into cropland when dredged spoil is deposited onto land adjacent to watercourses.

Industry. Weeds blocking intake sluices or the presence of pesticides in water may affect industrial processes requiring water.

Environmental Impacts. In general, the more effective a weed clearance operation, the greater the environmental impact. Deoxygenation of water, which affects fish, can occur when large volumes of cut or herbicide-treated weed are allowed to rot *in situ*. The number of invertebrates removed by mechanical control is estimated at approximately 1 million per m^3 of cut weed, and the number of vertebrates, mainly small fish, is estimated at about 40 per m^3 of weed removed. Care should be taken to avoid removing marginal vegetation during bird nesting times. Other indirect effects include destabilization of banks by removal of plant cover or by deposition of thick layers of cut vegetation on sloping banks.

MANAGEMENT OPTIONS

Mechanical Control

The following points should be considered when using mechanical weed control. Rafts of cut weed can drift downstream and block sluices, pumps, weirs, and other water control structures. Deoxygenation can occur if weed is left to rot in water, posing a risk to fish and invertebrate life. Rapid regrowth of cut weeds often necessitates a late season cut. In practice, the cost of mechanical control limits the frequency of the operation, thus a compromise between weed control and cost is usually necessary.

Cutting. Aquatic weeds can be cut manually or by using machines; the choice of technique depends on the scale of the problem to be tackled. Cutting by hand and raking cut material onto banks is not used widely now, although the majority of small drainage ditches used to be managed in this way. Now a wide range of mechanical cutting boats and weed bucket attachments for bank-based machines is available, but the principles of cutting and removing the cut weed remain unchanged. In general, the deeper the cut, the longer the weed control will last. Timing of the cut is critical. Good control can be achieved by regular cutting throughout the growing season, but this may not be possible for operational reasons. Cutting submerged weeds before the end of June will require a further cut toward the end of the season, and regrowth during summer tends to be faster than initial growth rates. The timing of cutting emergent and bankside weeds is less critical, but care should be taken to avoid bird nesting seasons.

Weed boats are used to cut submerged weeds in relatively deep and wide watercourses where large areas of weed require cutting. They usually have a front mounted reciprocating cutting bar which can cut to a depth of about 1.5 m. The work rate is estimated at approximately 1 km per day on a 10-m wide river. Cut weed must be removed from the water, either by a separate boat or by the action of a front rake on the cutting boat after cutting.

Bank-mounted equipment includes draglines, flail mowers, and dredging equipment. The work rate for cutting buckets and weed rakes is, on average, about 500 m day^{-1} on wide rivers.

Harvesting. The use of aquatic weed harvesters combines the cutting and collecting operation, therefore reducing time and increasing efficiency. They can and/or collect all types of weed, including filamentous algae and duckweeds and can deposit them at a collection point on the bank. The work rates for this type of equipment vary considerably with the density of weed and the distance traveled to offload the collected weed.

Uses for Cut Weed. There are a limited number of uses for cut weed; these include composting, reed for thatching, mulches, soil conditioners, and cattle forage. The material has low value, and processing should be done locally to minimize costs.

Chemical Control

Herbicides can be used to control a wide range of aquatic plants. They may be applied to emergent or surface-floating weeds by a foliar spray in much the same way as recommended for land plants. Submerged weeds and algae are treated by adding the herbicide to the water to build up an effective concentration in the water. This method will also control some floating and emergent species. Chemicals currently used to control aquatic plants are listed in Table 1.

Consideration should be given to the effects of herbicide treatment on the function of the waterbody. This not only involves the direct toxicity of the chemical but also indirect effects caused by decomposition of dead weeds. Retaining a proportion of the weed growth is recommended to maintain a habitat for invertebrates, fish, and other wildlife and to stabilize the ecosystem. Vegetation left *in situ* will take up nutrients and help to prevent the growth of algae. To minimize the effects of decomposing weed, herbicide labels have recommendations on treatment intervals that should be observed. It is normal to treat only 25% of an infested waterbody at one time, separated by the application interval. Consideration should be given to the irrigation interval when using herbicides in water. This is the time between application and the time when water is safe to use for irrigating of crops or watering livestock.

Application Methods

Application to Foliage. Recommendations for the correct dose are made in the form of weight of product or active ingredient per hectare. Herbicides are usually applied by knapsack or boat-mounted sprayer equipment. Using high-volume and low-pressure equipment to deliver a coarse spray with a minimum of small droplets is advisable to reduce the risk of spray drift. Localized or selective control can be achieved by the appropriate choice of herbicides and by spraying only those plants required. Water velocity and quality do not affect the treatment, but to avoid a buildup of chemical in the water, it is normal to apply herbicides in an upstream direction.

Application to Water. Recommendations usually refer to the theoretical concentration of active ingredient that would be achieved when the chemical has been evenly distributed throughout the water body but before any adsorption or degradation has occurred. This is usually expressed as parts per million (ppm or mg l^{-}1 or g m^{-3}). Some formulations may be applied on a rate per surface area basis because they sink onto the weeds or mud.

Table 1. Herbicides Suitable for Controlling Major Weed Groups

	Asulam	Copper	2,4-D amine	Dichlobenil	Diquat	Diquat Alginate	Endothall	Fluridone	Glyphosate	Maleic Hydrazide	Terbutryn
Trees and shrubs on banks									●		
Bracken and docks	●				○				●		
Broad-leaved weeds			●	○	○			○	●		
Grasses				○	○			●	●	●	
Reeds and sedges					○		●	●	●		
Floating-leaved plants				●	○	●	●	○	●		○
Free-floating plants			○		●			○	●		●
Submerged plants			○	●	●	●	●	●			●
Submerged plants (flowing water)						●	○				
Algae		●			●	●		○			●

● suitable for control; ○ short-term control or some species within groups not susceptible

The methods of application depend on the formulation. Herbicides should be spread as evenly as possible over the surface of the water. Granular formulations can be spread by hand using a suitable container, by a mechanical spreader, or by using an air-assisted blower. Liquid formulations should be diluted and applied by subsurface injection. This is usually achieved by trailing nozzles below the water as close to the top of the weed bed as possible, without disturbing the bottom sediment. Viscous gel formulations should be applied with specialized equipment.

Timing. Emergent and floating leaved plants should be treated from midsummer, when the leaves have formed fully. Annual weeds should be treated before flowers have set to prevent production of seeds.

Treatment of submerged weeds and algae is normally recommended in spring and early summer when weeds are actively growing. Treatment of floating-leaved plants is usually more successful before the floating leaves have formed. Later in the season, when large biomasses of weed have developed and particularly when the water is warm, there is a severe risk of deoxygenation of the water, after control, caused by decomposition of the dead weed. This effect is particularly acute when using terbutryn, as the mode of action blocks photosynthesis very quickly, but respiration continues.

An indirect hazard to fish may arise through possible effects on fauna that provide food for fish. These may be due to direct toxicity to invertebrates or more likely to loss of habitat after weed control operations. However, loss of habitat can lead to increased predation by fish. These effects are temporary and should be set against the disadvantages of dominant weed species reducing habitat structure for all aquatic fauna.

Biological Control

The decline in public acceptability of herbicide use in water, coupled with an increase in the costs of mechanical control has led to the development of several biological control agents for aquatic weeds.

Livestock. Horses, cows, goats, and sheep can be used to graze marginal vegetation. Care should be taken to avoid excessive poaching of the banks (although some poached banks are a valuable habitat), and fencing is required to keep animals from straying.

Chinese Grass Carp and Other Fish. Herbivorous fish are most common in Asia and South America. One species, the Chinese grass carp, *Ctenopharyngodon idella*, is available for weed control in enclosed situations. Triploid grass carp are used in the United State to ensure that the fish cannot breed. Licenses are usually required to introduce of these fish. Stocking densities vary depending on the type and severity of the weed problem, but usually a density of between 75 and 150 $\mathrm{kg\,ha^{-1}}$ will achieve adequate weed control. Bottom feeding fish stir up sediment and create turbid conditions in which submerged macrophytes cannot grow. These fish include primarily carp and bream.

Waterfowl. Waterfowl graze on some species of floating and submerged weed. When present in large numbers (e.g., swans on *Ranunculus*), they can cause significant losses of aquatic vegetation, which may not be desirable.

Pathogens and Insects. Classical biological control is not widely practiced in Europe, and not at all on aquatic and riparian weeds. Biological control agents survive only on their host species, meaning they cannot spread to ornamentals or related crop species. Most native plant species have a number of native host-specific pathogens or insect predators. There are effective biological control agents for many aquatic weed species (Table 2) established in other parts of the world, but none is used in Europe.

Biological control has been successful in about 45% of the attempts made on aquatic weeds. It is essentially a management tool designed to maintain the level of a pest

Table 2. Established Biological Control Agents for Aquatic Weeds

Weed Species	Biological Control Agent	Type	Region
Alternanthera philoxeroides	*Agasicles hygrophila*	Flea beetle	America
Azolla filiculoides	*Stenopelmus rufinasus*	Weevil	S. Africa
Eichhornia crassipes	*Neochetina eichhorniae*	Weevil	America Africa, India
	Neochetina bruchi	Weevil	America, Africa, India
	Cercospora piapori	Fungus	Africa
	Altenaria eichhorniae	Fungus	Africa
Hydrilla verticillata	*Hydrilla pakistanae*	Insect	America
	Hydrelli sarahae	Insect	
	Hydrellia pakistanae	Insect	
	Hydrellia balciunasi	Insect	
	Mycoleptodiscus terrestris	Fungus	
	Plectosporium tabacinum	Fungus	
	Cricotopus lebetis	Midge	
Myriophyllum aquaticum	*Lysathia* sp	Beetle	S. Africa
Myriophyllum spicatum	*Bagous subvittatus*	Insect	N. America
	Bagous myriophylli	Insect	
	Phytobius	Insect	
Salvinia molesta	*Cyrtobagous salviniae*	Weevil	Africa
	Neohydronomus affinis	Weevil	

at a tolerable level in the community without necessarily eradicating the weed species. It is a useful tool to employ in combination with other techniques such as mechanical or chemical control.

Environmental Control

Environmental control is the technique of changing the environment, either temporarily or permanently, to reduce the suitability of the habitat for the target weed. This technique tends to shift the weed species to other plants, which may or may not be desirable depending on the ultimate use of the waterbody in question.

Flow. Aquatic plants are divided into riverine or lacustrine species. Although not mutually exclusive, plants that tend to favor slow flowing environments do not grow in faster flowing water and vice versa. Alterations to flow can have marked effects on the composition of the aquatic plant community.

Water Chemistry. Reduction of the quantity of major plant nutrients entering water tends to reduce the biomass development of weed species but not necessarily the species composition of the weed community. Diversion or further treatment of sewage effluent to remove excess P has been effective in stabilizing river communities by removing the dominance of filamentous algae. Nutrient release from sediments is a major source of supply for lakes and other slow flowing watercourses, such as canals and drainage ditches. Even if all nutrients could be prevented from entering such systems, there would be sufficient nutrient supply in the sediment to provide for continued aquatic plant growth for more than 100 years in most cases.

Shade. Shading submerged aquatic macrophytes can be an effective method of control, especially in sensitive areas where chemical control is inappropriate. Shade can be produced by planting trees or shrubs on the banks of watercourses. However, this often interferes with access to the water and may not be appropriate in all situations. Opaque floating material can be used in still waters to achieve increased shading, although it must remain in place for at least 4 months to have a lasting effect. The use of dyes that absorb light at photosynthetic wavelengths is also widely used in amenity lakes and ponds. Increased turbidity from bottom feeding fish also excludes light, as does the action of powered boats, and these techniques are most effective in sluggish water where the sediment can remain suspended for long periods.

Burning. Emergent weeds, particularly the stiffer stemmed reeds, do not always collapse in the autumn and can form large masses of dead, standing material. In dry ditches and on banks, this material can sometimes be burned. This can be a useful way of reducing the bulk of plant material which might otherwise collapse during the winter and block channels. Burning can also be used to destroy cut material after drying. This is good way of disposing of poisonous plants, some of which remain toxic after death but become palatable to livestock.

Alterations in Water Level. Exposure of part, or all, of the bed of a water body by lowering the water level has been used successfully to manage aquatic vegetation. Control is achieved either by dehydration of the vegetation or by exposure to low temperatures. The process is sometimes termed drawdown or dewatering. It can also alter the character of the sediment which may reduce weed growth. However, in deeper water bodies, drawdown can allow weeds to establish in depths below their normal limit. Once established, they can grow up toward the surface, as the level is again raised, to remain in the higher light intensity. In these situations, drawdown can spread a weed problem into areas which would normally remain weed-free.

Some indication of the relative effectiveness of treatments can be see in Table 3. This is not a definitive guide for every situation, and expert advice should be sought before applying any technique to a particular weed problem to optimize efficacy and minimize costs.

The benefits assigned to each technique are for general guidance only, and the results obtained may be modified by specific site conditions.

Other Methods of Control

Many other techniques for controlling aquatic vegetation have been tried, which do not fit closely into any of the previous categories. These include the following: (1) the use of lasers to control emergent or floating weeds and ultrasound vibrations that can disrupt cells of submerged plants; (2) floating oil films that cause floating weeds to sink, and (3) increasing wave action that submerges floating weeds and increases turbidity thus suppressing submerged vegetation. There are also several biological control agents, including crayfish, other invertebrates, and wildfowl, that have been effective against some weeds. Most of these have been tested in tropical or subtropical conditions and, at present, the grass carp is the most effective biological control agent in more temperate regions. Some of the more effective alternative treatments are outlined below.

Magnetic Treatment of Flowing Water. The use of magnetic water treatment devices, it has been shown, affects the growth of algae. This is thought to be due to interference with the uptake and storage of calcium in algal cells. This treatment is most effective in closed recirculating systems.

Ultrasound. Ultrasonic devices are used widely in Europe to disrupt cells mainly of algae and biofilm bacteria in water storage reservoirs where herbicides cannot be used.

Miscellaneous Treatments. There are a number of alternative treatments for water that affect the nutrient concentration and, therefore availability, to aquatic plants. Most of these involve precipitation of phosphate in an insoluble form.

Table 3. The Relative Efficacy of Management Options

	Weed Type					
	Emergent		Floating-leaved		Submerged	
Management Option	Narrow-leaved	Broad-leaved	Rooted	Free-floating	Rooted	Algae
Mechanical						
Hand-cutting	○	○	○	**X**	○	**X**
Flail mower	○	○	**X**	**X**	**X**	**X**
Weed bucket	○	○	○	○	○	○
Weed boat	○	○	○	○	○	**X**
Harvester	○	○	○	○	○	○
Dredger	√	√	√	**X**	√	○
Chemical						
Foliar application	√	√	√	○	**X**	**X**
Applied to water	**X**	**X**	√	√	√	√
Biological						
Livestock	○	○	**X**	○	○	**X**
Waterfowl	**X**	**X**	**X**	√	○	○
Fish	**X**	**X**	√	√	√	√
Insects and pathogens	○	○	○	√	√	○
Biomanipulation	**X**	**X**	**X**	**X**	**X**	√
Environmental						
Shade	○	○	○	○	○	○
Nutrient manipulation	**X**	**X**	**X**	○	○	○

√ Effective control lasting at least one season
○ Moderate benefit only, or control lasting less than one season
X No useful effect

SCREEN FILTERS FOR MICROIRRIGATION

Dorota Z. Haman
University of Florida
Gainesville, Florida

Water quality is a major concern in the management of microirrigation systems (drip, microsprinkler, microspray). Emitters plugged by physical, chemical, or biological contaminants may create significant problems in everyday maintenance. Effective and reliable filtration is required for successful operation of microirrigation systems. There are different types of filters available for the removal of physical contaminants from irrigation water. Selection of a filter depends on the type and amount of contaminants in the irrigation water, the type of emitters, the size of the irrigation system, and the desired management practices. Screen filters should be a primary choice when water is pumped from a well where the only filtration requirement is to remove mineral particulate matter (Fig. 1). They are sufficient in the absence of organic contamination and in addition they are usually inexpensive and easy to maintain. To remove both particulate matter and organic growths, media filters (also called sand filters) may also be used. They are usually recommended when large amounts of algae or other organic contaminants are present.

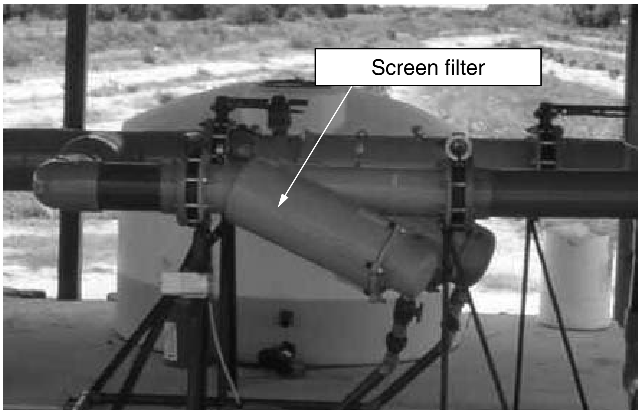

Figure 1. Screen filter installed in the main line of irrigation system.

Screen filters are simple and economical filtration devices of various shapes and sizes that are frequently used for the removal of physical contaminants. They can be made of metal, plastic, or synthetic cloth enclosed in a special housing. Screen filters are recommended for the removal of very fine sand or larger-sized inorganic debris from irrigation water. It is normally not effective to use screen filters for the removal of heavy loads of algae or other organic material typical in surface water supply, since screens clog rapidly, requiring too frequent cleaning to be practical.

When surface water is used for microirrigation, screens are often used as secondary filters, after organic matter has been removed with media (sand) filters. In this capacity they prevent washed out media from entering the irrigation system. When well water is used for irrigation, a screen filter may be used as a primary filter, or it may be secondary to a vortex sand separator, depending on the mineral particle load in the water.

Filtering screens are classified according to the number of openings per inch, with a standard wire size given for each screen size (Table 1). The minimum size of particle retained by a screen filter with a certain mesh can be determined from Table 1. A 200-mesh screen is usually recommended for drip type microirrigation where 150 mesh is often sufficient for microsprinklers and microsprays. Water filtered with a 200-mesh screen will contain only smaller particles of very fine sand, silt, and clay. These particles should freely move through the emitter openings unless flocculation, aggregation, and bridging are taking place due to unusual water chemistry. If clogging problems persist with 200-mesh filtration, chemical treatment of water may be necessary.

Generally, it is recommended to remove particles down to a size ten times smaller than the emitter's passageway so that grouping and bridging of particles will not cause clogging. The maximum particle size that can safely enter the irrigation system with a given emitting device should be provided by the manufacturer. Organic particles with a density approaching the density of water tend to group and bridge more easily. Particles heavier than water, typically mineral particles, may settle and collect in the low flow zones of the irrigation system.

The size of the filter is specified by its effective area, which is the area of the openings in the screen. It is specified in relation to cross-sectional area of the main pipe. A desirable ratio is 2 or more (area of openings much larger than the cross-sectional area of the pipe). The mesh size of the filter (opening size) will depend on the smallest particle size to be removed from the irrigation water.

Table 1. Representative Screen Mesh Numbers and the Corresponding Standard Opening Size Equivalents

Screen Filters		Removed Particles	
Screen Size, mesh	Opening Size, μm	Soil Classification	Particle Size, μm
4	4760	Very coarse sand	1000–2000
10	2000		
20	711	Coarse and medium sand	250–1000
40	420		
80	180	Fine sand	100–250
100	152		
120	125		
150	105		
180	89	Very fine sand	50–100
200	74		
270	53		
325	44	Silt	2–50
—[a]	—[a]	Clay	2 and less

[a] Screen filters are not normally used to remove this size particles.

When estimating the appropriate size of filter for a specific application, one should consider the quality of water needed, volume of water required to be passed through the filter between consecutive cleanings, filtration area of the filter screen, and allowable pressure drop through the filter. A screen filter can handle a large range of discharges. However, discharges that are large in relation to the filtering surfaces will result in greater pressure losses, shorter life of the filter, and the requirement for frequent cleaning.

If the irrigation water contains large amounts of sand or other inorganic contaminants, it is advisable to use vortex sand separators before the screen filters, which use centrifugal force to separate heavier particles away from the fine mesh cartridge. Heavy particles collect at the bottom of the filter, where they can be periodically flushed away.

Simple gravity screen filters are often used as primary filtration where an elevated water source from a canal or reservoir is available. Since the pressure loss through a gravity screen filter is very small, these filters are often used in systems where such losses must be minimized. They operate on the principle of gravity force, not water pressure. Because only atmospheric pressure is used, soft organic material is not forced through the screen. Gravity screens are easy to install and maintain. Usually, these filters are constructed as two chambers separated by a fine mesh screen, which can be cleaned manually or automatically. Most of them incorporate self-cleaning water jets.

If water is pumped from a surface source (lake, pond, stream, river, irrigation canal, or pit), aquatic plants, fish, and animals should be kept away from the intake pipe since the efficiency of pumping may be reduced by debris in the water entering the suction pipe. This can be accomplished with an intake screen. The design of the screen must allow for the separation of the debris from the water without pulling the debris against the screen. A horizontal screen placed below the water surface or an intake basket-type filter are often used for this purpose. Some of these screens are self-cleaning and require very little maintenance.

Small amounts of larger particles, such as lime rock flakes from the aquifer, can be removed using a Y-type strainer with replaceable screens and with clean-out faucets or valves. They can be flushed while in use if the required high flow for flushing is provided. Blowdown filters operate in the same way as other types of pressure screen filters. They are designed for easy cleaning, which is done by opening a valve that diverts the water flow through the screen, releasing trapped particles.

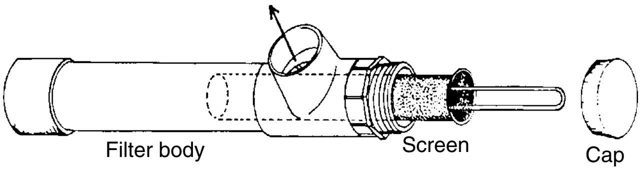

Figure 2. Schematics of a typical screen filter.

Occasionally, blowdown filters require manual cleaning to eliminate particles that have become lodged in the screen.

Screen filters can be cleaned manually by removing the screen and washing it with clean water (Fig. 2). However, dismantling is not very practical for use with waters containing high levels of contaminants, since it must be done very often. Blowdown screen filters greatly simplify the cleaning process. They are cleaned with the water diverted through the screen by opening a valve. This operation does not require disassembling. The screen does not have to be removed for routine cleaning of the filter, but it must be removed for occasional cleaning of the particles that become lodged in the screen.

The most convenient method for cleaning screen filters is automatic screen flushing triggered by critical pressure differential across the filter, which increases with buildup of debris on the screen. Automatic cleaning can also be set on a time schedule. Filtered water is always recommended for washing and backflushing. Additional safety devices such as secondary screen filters at the entrance to each irrigation system subunit are highly recommended.

XERISCAPE

Douglas F. Welsh
Texas A&M University
Collage Station, Texas

Much attention and controversy have surrounded the *xeriscape* concept of *landscaping* since its inception in 1981. The proper definition of xeriscape is "quality landscaping that conserves water and protects the environment." Above all things, it must be a quality design that balances the lawn area, shrub and flower plantings, and the hardscape (i.e., decks, patios and sidewalks). Landscapes composed of rocks or plastic flowers alone are not xeriscapes. Xeriscapes are in tune with the environment; therefore, xeriscape applies to the desert in southwestern United States as well as the semitropical southeast.

Xeriscape has seven basic principles:

- planning and design
- soil improvement
- appropriate plant selection
- practical turf areas
- efficient irrigation
- mulching
- appropriate maintenance

USE OF PRACTICAL TURF AREAS IN A XERISCAPE DESIGN

Of the seven principles, none has received more attention than practical turf areas. This principle, which concerns turfgrass in the landscape, has been shrouded in misinformation that has been touted as fact by "experts" in xeriscape, water supply, and turf culture.

The original turf-related principle established by the Denver originators of xeriscape was "limited turf use." For Denver and much of the arid West, the seemingly logical approach to reducing landscape water consumption was simply to reduce the use of turf. However, as the Xeriscape concept has matured and spread, the principle of limited turf use was increasingly scrutinized by horticulturists and turf experts. Today's xeriscape movement incorporates a more holistic approach to reducing turf irrigation, fully recognizing that the type of plant materials or irrigation in the landscape has as much of an effect on water consumption as the human factor and good landscape water management (Fig. 1).

THE NEED TO CHANGE ATTITUDES AND HABITS

Throughout the xeriscape movement, the evident truth is that plants do not waste water; people do. Another fact is that irrigation systems do not waste or save water; people do. The mission of xeriscape is clear: Change the attitudes and irrigation habits of professional and amateur landscape managers. Proper water management provides the greatest opportunity for *water conservation* in the landscape.

Xeriscape focuses on the use of turfgrass in the landscape because of the tremendous potential for irrigation water abuses in the name of maintaining green turfgrass. Within the traditional landscape, turfgrass has received the major share of total landscape irrigation because grass often makes up a large percentage of the total landscape. Through the principles of xeriscape, turf irrigation can be reduced and the many benefits of turfgrass can still be derived.

BENEFITS OF TURFGRASS IN THE LANDSCAPE

Turfgrass is an integral component of most landscapes. It is certainly the best recreational surface for children and athletes. Furthermore, it has a tremendous mitigating effect on the environment, reducing heat loads, noise and water and air pollution. A turfgrass lawn is second only to a virgin forest in the ability to harvest water and recharge groundwater resources. As a design component, turfgrass provides the landscape with unity and simplicity while inviting participation in it.

However, the fact remains that turfgrass is the highest user of irrigation water in the traditional landscape. This

Figure 1. The proper definition of xeriscape is "quality landscaping that conserves water and protects the environment." Above all things, it must be a quality design that balances the lawn area, shrub and flower plantings and the hardscape (i.e., decks, patios and sidewalks).

is significantly different from saying that turfgrass is the highest water-using plant in the landscape—which is not the case. The discrepancy between these two statements yields the most common misconception and misrepresentation in xeriscape, and it is therefore the basis of controversy and unproductive efforts. To resolve this controversy, some scientific and practical fundamentals of turfgrass are explained using actual xeriscape principles.

XERISCAPE PRINCIPLES FOR REDUCING TURFGRASS IRRIGATION

Specifically, xeriscape principles promote the following strategies to reduce turfgrass irrigation:

- Prepare soils for turf areas as carefully as any other planting area to use all the moisture available, promoting the plant's vigor and water-use efficiency.
- Place turf species in landscape zones based on water requirements.
- Select adapted turf species and varieties that have lower water demands.
- Irrigate turf in areas that provide function (i.e., recreational, aesthetic, foot traffic, dust and noise abatement, glare reduction, temperature mitigation).
- Use nonirrigated turf areas where appropriate.
- Irrigate turf based on true water needs.
- Decrease fertilization rates and properly schedule fertilization.

FINE-TUNING TURFGRASS XERISCAPE PRINCIPLES

In traditional landscape design, turfgrass makes up the major portion of landscapes. The tremendous square footage of turfgrass in a landscape accounts for the reason that turfgrass irrigation, as a percentage of total landscape irrigation, is so high. The "practical turf areas" guideline promotes the use of turf only in those areas of the landscape that provide function. In residential landscapes, a turf area is usually a necessity for recreation and entertainment. But turf should not be irrigated on narrow strips of land or other areas that are difficult to water (Fig. 2).

Good landscape water management begins with planning and design. By designing the landscape as zones based on plant water needs, turf can be appropriately placed for function, benefit and water efficiency. Zoning the landscape and irrigation system allows watering turfgrass on a more frequent schedule than shrubs. For established trees and shrubs, the irrigation strategy should use deep soil moisture and depend on natural rainfall to replenish soil moisture. When sufficient rainfall does not occur, supplemental irrigation of trees and shrubs may be required.

Another way to incorporate turf into the landscape and conserve water is simply not to irrigate. Many turfgrass species are drought-tolerant and can survive extreme drought conditions. The grass may turn brown for a while, but rainfall will green it up again. This approach may be unacceptable for many residential and commercial landscapes, but in the case of parklands, industrial sites and rights-of-way, brown turf may be acceptable.

Figure 2. In residential landscapes, turf areas are important for recreation and entertainment.

SELECTING THE PROPER TURFGRASS SPECIES AND VARIETIES

Wherever the landscape, selection of turfgrass species and varieties is of utmost importance. Extensive research has shown that there are significant differences in water requirements among turf species and even among varieties within species. The capacity of different turf species to avoid and resist drought also varies significantly. To help reduce landscape water requirements, xeriscape recommends selecting turfgrass varieties (and other landscape plants) that are both adapted to the area and have the lowest practical water requirements.

Landscape managers should be keenly aware of drought-stress indicators shown by turfgrass and other plants in the landscape, including a range of color changes, leaf curl, and wilting, and they should strive

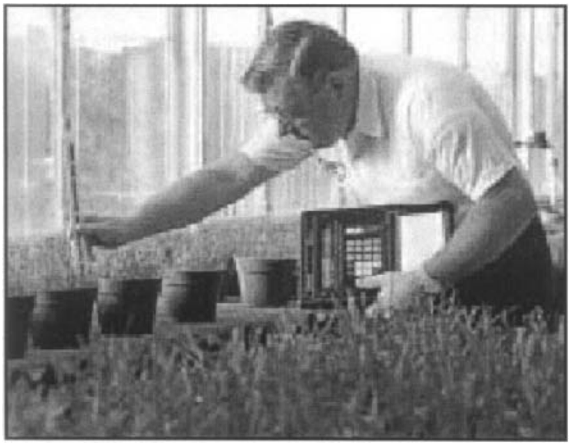

Figure 3. Wherever the landscape, selection of turfgrass species and varieties is of utmost importance. Extensive research has shown that there are significant differences in water requirements among turf species and even among varieties within species.

Figure 4. Attitudes and habits about turf are changing. For much of the arid West, the seemingly logical approach to reducing landscape water consumption was simply to reduce the use of turf. However, as the xeriscape concept has matured and spread, the principle of limited turf use was increasingly scrutinized by horticulturists and turf experts. Today's xeriscape movement incorporates a more holistic approach to reducing turf irrigation, fully recognizing that the type of plant materials or irrigation in the landscape has as much of an effect on water consumption as the human factor and good landscape water management.

to meet the water needs of each group of plants. By irrigating only when the plants require water versus by the calendar, the manager can dramatically reduce landscape water use.

The water requirements of turfgrasses can be minimized through specific horticultural practices. Decreasing fertilizer application rates and timely applications of slow-release fertilizers tend to reduce flushes of growth that can increase water requirements (Fig. 3).

THE XERISCAPE CHALLENGE

Xeriscape is a challenge and an opportunity for the "green" (landscape, turf and nursery) and "blue" (water utilities and agencies) industries. Through xeriscape, these two industries have been brought together to focus on landscape water use. Although this marriage has not always been easy, the best minds are prevailing in efforts to perfect and implement the xeriscape concept.

By embracing the xeriscape concept, including the principle of practical turf areas, the green and blue industries can continue to be recognized as good stewards of the environment (Fig. 4).

MEDIA FILTERS FOR MICROIRRIGATION

Dorota Z. Haman
University of Florida
Gainesville, Florida

To avoid plugging of microirrigation emitters, it is generally recommended to remove particles down to a size ten times smaller than the emitter's passageway so that grouping and bridging of particles will not cause clogging. The manufacturer should provide the maximum particle size that can safely enter the irrigation system with a given emitting device. Organic particles with a density approaching the density of water tend to group and bridge more easily.

Media (also called "sand") filters are well suited for removal of either organic or inorganic particles. Due to their three-dimensional nature, media filters have the ability to entrap large amounts of contaminants. They do not seal off as easily and therefore will not clog as often as filters that trap particles on their surface, such as screen filters.

Media filters used in microirrigation systems operate under pressure. They consist of fine gravel and "sand" (that actually is crushed granite or silica) of selected sizes placed in pressurized tanks (Fig. 1). The main body of the tank contains sand, which is the active filtering ingredient. The sand is placed on top of a thin layer of gravel that separates it from an outlet screen. Sharp-edged sand or crushed rocks are recommended for the filtering media since the sharp edges catch soft algal tissue. Crushed granite or silica graded into specific sizes for a particular system are commonly used porous media. It is not recommended to use natural sand as a filtering media since the particles are usually rounded and smooth. The size of media particles is very important. Too coarse particles result in poor filtration and possible clogging of the emitting devices, while too fine particles trigger unnecessary frequent backwashing of the filter.

Two factors describe the media used in the filter: uniformity coefficient and mean effective size. Uniformity coefficient reflects the range of sand sizes within the grade. It is desirable to keep sand particles as uniform in size as possible. The uniform size of filtering media assures better control of the filtration. Grading in size from fine to coarse causes premature clogging of the filter, since the nonuniform pores of the media also retain particles small enough to be tolerated by the emitters. The uniformity coefficient is represented by the ratio of the size of the screen opening that would pass 60% of the filter sand to the screen opening that will pass 10% of the same sand. For irrigation purposes a uniformity coefficient of 1.5 is considered adequate.

The mean effective sand size is the size of the screen opening that will pass 10% of the sand sample. It is a measure of the minimum sand size in the grade and therefore an indicator of the particle size that will be removed by the media. Some examples of sand media are shown in Table 1. It can also be seen that the quality

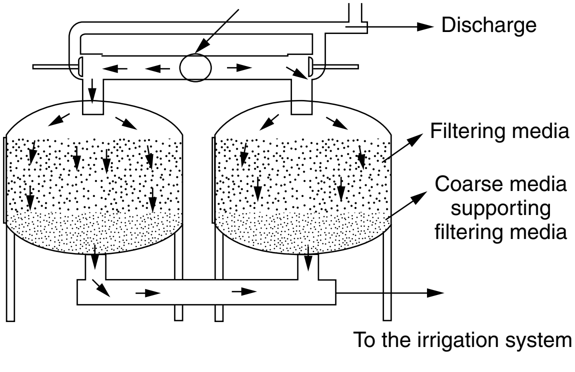

Figure 1. Schematic of media filters in filtration mode.

Table 1. Relationship Between Designated Media Numbers, Filtration Quality, and Equivalent Screen Sizes

		Removed Particles		Media Filters		
Corresponding Screen Size, mesh	Opening Size, μm	Soil Classification	Particle Size, μm	Media Number	Material	Filtration Quality, mesh
4	4760	Very coarse sand	1000–2000	8	granite	100–140
10	2000					
20	711	Coarse and medium sand	250–1000			
40	420					
80	180	Fine sand	100–250			
100	152					
120	125			11	Granite	140–200
150	105			16	Silica	140–200
180	89	Very fine sand	50–100	20	Silica	200–230
200	74					
270	53			30	Silica	230–400
325	44	Silt	2–50			
—[a]	—[a]	Clay	2 and less			

[a] Screens are not normally used to remove this size particles.

Figure 2. Media filters installed in a field.

of filtration increases with a smaller effective size of filtering media.

The design flow capacity of a media filter is expressed in liters per second of flow per square meter of surface area of the media normal to the direction of flow. The filter should be sized so it can handle the poorest water quality at a given site and provide the required flow for the functioning of the irrigation system (Fig. 2). The quality of irrigation water before filtration may vary with the seasons and weather conditions. This is especially true with regard to surface water but can also apply to well water in some cases.

The effectiveness of a filter is a measure of its ability to remove particles of a certain size. The effectiveness of filtration increases with a decreasing size of filtering media grain size (larger designation number). However, smaller media size will require more frequent cleaning.

Filtration effectiveness is also inversely proportional to the flow rate through the filter. The higher the flow rate, the lower the effectiveness of the media filters. Therefore, filter manufacturer's specifications and water quality samples should be used to select a filter for a specific application.

Media filters are cleaned by backwashing (Fig. 3). This operation consists of reversing the direction of water flow in the tank. Clean water is usually supplied from the second tank. The upward flow fluidizes the media and flushes collected contaminants. The backflush water is discharged and does not enter the irrigation system. The backwash flow must be carefully adjusted in order to provide sufficient cleaning without accidental removal of the media. If flows larger than those recommended are used to backflush the media, it is very likely that the media will be flushed from the filter by the backflow water. Only recommended flow rates for a given media should be used.

Most media filters are backflushed at prescheduled time intervals or by using automatic devices based on the pressure loss across the filter. For systems in which low quality water requires frequent backflushing, automatic cleaning is necessary to avoid problems.

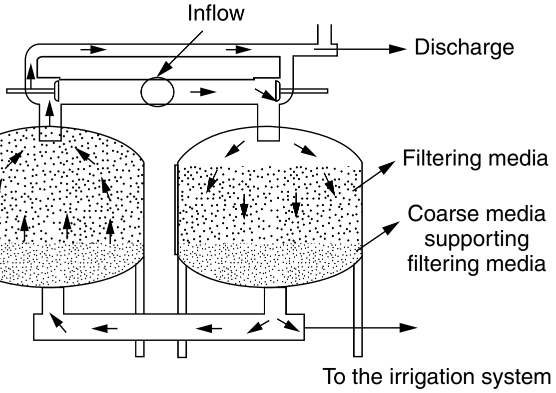

Figure 3. Schematic of the media filters in backflush mode.

As the filter is backflushed every time the pressure differential exceeds a predetermined value, large pressure drops in the irrigation system are avoided, maintaining the system's uniformity and efficiency. The pressure differential triggering backflushing depends on the pressure required for the proper functioning of the irrigation system. The irrigation system pump should be able to supply enough pressure to compensate for the pressure drop through the filter just before the filter is washed. Pump capacity must also be adequate to supply enough pressure and flow rate to flush the filter. Also, it is advisable in a large irrigation system to use a number of smaller tanks instead of a few large ones because of the higher backwash flow that large tanks require for good cleaning. Automatic backflush systems will eliminate sudden changes in water quality that can create problems if a filter is washed only at regular intervals. This is especially important in systems using surface water supplies with changing contamination levels.

AIR–SEA INTERACTION

S.A. HSU
Louisiana State University
Baton Rouge, Louisiana

INTRODUCTION

Air–sea interaction is, according to Geer (1), the interchange of energy (e.g., heat and kinetic energy) and mass (e.g., moisture and particles) that takes place across the active surface interface between the top layer of the ocean and the layer of air in contact with it and vice versa. The fluxes of momentum, heat, moisture, gas, and particulate matter at the air–water interface play important roles, for example, in environmental hydraulics and water–environment–health interactions, during low wind speeds before the onset of wave breaking, the exchange of air bubbles is limited. If this situation persists for a long time, algal blooms may develop, ultimately affecting water quality. On the other hand, during typhoon/hurricane conditions, the storm surge may affect the sewerage outflow at a greater depth than normal because of shoaling. In the area on the right-hand side of the storm track (in the Northern Hemisphere), runoff may also be blocked due to the surge, thus increasing the flood potential and saltwater intrusion.

Air–sea interaction encompasses vast scales in both spatial and temporal viewpoints, so only a few basic and applied topics are summarized here, including the parameterization of stability length, determination of friction velocity, wind–wave interaction, and the estimation of shoaling depth during storms. For more detailed laws and mechanisms in air–sea interaction, see Donelan (2) and recently Csanady (3); for air–sea exchange of gases and particles, see Liss (4) and most recently Donelan et al. (5); for the role of air–sea exchange in geochemical cycling, see Buat-Menard (6) and recently Liss (7); for larger scale air–sea interaction by La Niña and its impacts, see Glantz (8); for more physics, chemistry, and dynamics related to air–sea exchange, see Geernaert (9); and for wind–wave interaction, see Janssen (10).

PARAMETERIZATION OF THE STABILITY LENGTH

In the atmospheric boundary layer, the buoyancy length scale, L, also known as the Obukhov (or Monin–Obukhov) length, is a fundamental parameter that characterizes the "stability" of the surface layer (11). L describes the relative importance between the buoyancy effect (or thermal turbulence) and wind shear (or mechanical turbulence). According to Hsu and Blanchard (12), L can be parameterized as follows.

For unstable conditions (i.e., when $T_{\text{sea}} > T_{\text{air}}$),

$$\frac{z}{L} = -\frac{1000(T_{\text{sea}} - T_{\text{air}})(1 + \dfrac{0.07}{B})}{(T_{\text{air}} + 273.2)U_z^2} \tag{1}$$

For stable conditions ($T_{\text{air}} > T_{\text{sea}}$),

$$\frac{z}{L} = \frac{620(T_{\text{air}} - T_{\text{sea}})}{(T_{\text{air}} + 273.2)U_z^2} \tag{2}$$

According to Hsu (13,14),

$$B = 0.146(T_{\text{sea}} - T_{\text{air}})^{0.49} \tag{3}$$

where z is the height normally set to 10 m; T_{air} and T_{sea} stand for the air and sea temperatures, respectively; U_z is the wind speed at height z; and B is the Bowen ratio.

For operational and engineering applications, $z/L \leq -0.4$ is unstable, $|z/L| < 0.4$ is neutral, and $z/L \geq 0.4$ is stable.

PARAMETERIZATION OF THE ROUGHNESS LENGTH

The roughness parameter Z_0 can be computed based on the formula provided in Taylor and Yelland (15) that

$$\frac{Z_0}{H_s} = 1200\left(\frac{H_s}{L_p}\right)^{4.5} \tag{4}$$

and, for deep water waves,

$$L_p = \frac{gT_p^2}{2\pi} \tag{5}$$

where g is gravitational acceleration, H_s and L_p are the significant wave height and peak wavelength for the combined sea and swell spectrum, and T_p is its corresponding wave period. Note that H_s is defined as the average of the highest one-third of all wave heights during the 20-minute sampling period.

ESTIMATION OF THE FRICTION VELOCITY

The friction velocity (u_*) can be obtained as follows:

$$u_* = U_{10}C_d^{1/2} \tag{6}$$

where U_{10} is the wind speed at 10 m and C_d is the drag coefficient.

According to Amorocho and DeVries (16), the WAMDI Group (17), and Hsu (18), one may classify the air–sea interaction into three broad categories based on wave breaking conditions:

In light winds, $U_{10} < 7.5$ m/s, prior to the onset of wave breakers,

$$C_d = \left(\frac{u_*}{U_{10}}\right)^2 = 1.2875 * 10^{-3} \tag{7}$$

Both thermal and mechanical turbulence are important.

In moderate winds, $7.5 \leq U_{10} \leq 20$ m/s, the range after the onset but before the saturation of wave breakers,

$$C_d = (0.8 + 0.065U_{10}) * 10^{-3} \tag{8}$$

Mechanical turbulence is more important than thermal effects.

In strong winds, $U_{10} > 20$ m/s, after the saturation of wave breakers,

$$C_\mathrm{d} = 2.5 * 10^{-3} \tag{9}$$

Mechanical turbulence dominates.

ESTIMATING LATENT HEAT FLUX (OR EVAPORATION)

Using the parameter of the Bowen ratio supplied by Hsu (13,14), as shown in Eq. 3, the latent heat flux (H_latent) can be estimated as

$$H_\mathrm{latent}(\mathrm{W\,m^{-2}}) = \frac{1}{B}H_\mathrm{sensible} = \frac{1}{B}\rho_\mathrm{a}C_\mathrm{p}C_\mathrm{T}(T_\mathrm{sea} - T_\mathrm{air})U_{10} \tag{10}$$

where H_sensible is the sensible heat flux, ρ_a (=1.2 $\mathrm{kg\,m^{-3}}$) is the air density, C_p (=1004 $\mathrm{J\,kg^{-1}\,K^{-1}}$) is the specific heat at constant pressure for dry air, C_T ($=1.1 * 10^{-3}$) is the transfer coefficient for heat, ($T_\mathrm{sea} - T_\mathrm{air}$) is in K, and U_{10} in $\mathrm{m\,s^{-1}}$.

A latent heat flux of 1 $\mathrm{W\,m^{-2}}$ is equivalent to an evaporation rate of $3.56 * 10^{-3}$ $\mathrm{cm\,day^{-1}}$, so Eq. 10 can be used to estimate the evaporation rate.

ESTIMATING MAXIMUM SUSTAINED WIND SPEED DURING A HURRICANE

Under hurricane/typhoon conditions, intense air–sea interaction occurs. Beach erosion, engineering structures, storm surge, and sewerage outflow can all be affected, so this topic should deserve more attention than the deepwater environment. The very first subject related to a tropical cyclone is to estimate its maximum sustained wind speed at the standard height of 10 m (i.e., U_{10}). This is accomplished as follows.

From the cyclostrophic equation (i.e., centrifugal force = pressure gradient) (11),

$$\frac{U_\mathrm{a}^2}{\gamma} = \frac{1}{\rho_\mathrm{a}}\frac{\partial P}{\partial \gamma} = \frac{1}{\rho_\mathrm{a}}\frac{\Delta P}{\Delta \gamma} = \frac{1}{\rho_\mathrm{a}}\frac{P_\mathrm{n} - P_0}{\gamma - 0} \tag{11}$$

where U_a is the maximum sustained wind speed above the surface boundary layer, γ is the radius of the hurricane, $\partial P/\partial \gamma$ is the radial pressure gradient, P_n is the pressure outside the hurricane effect (1013 mb, the mean sea level pressure), and P_0 is the hurricane's minimum central pressure. Because $\rho_\mathrm{a} = 1.2\ \mathrm{kg\,m^{-3}}$, $\Delta P = (1013 - P_0)$ mb, and 1 mb = 100 $\mathrm{N\,m^{-2}}$ = 100 $\mathrm{kg\,m^{-1}\,s^{-2}}$, Eq. 11 becomes

$$U_\mathrm{a} = \left[\frac{100\ \mathrm{kg\,m^{-1}\,s^{-2}}}{1.2\ \mathrm{kg\,m^{-3}}}\right]^{1/2}\sqrt{\Delta P} = 9\sqrt{\Delta P} \tag{12}$$

According to Powell (19), $U_{10} = 0.7U_\mathrm{a}$; therefore

$$U_{10} = 6.3\sqrt{\Delta P} = 6.3(1013 - P_0)^{1/2} \tag{13}$$

where U_{10} is in $\mathrm{m\,s^{-1}}$ and ΔP in mb.

Equation 13 has been verified by Hsu (18). In 1985, during Hurricane Kate over the Gulf of Mexico, the U.S. National Data Buoy Center (NDBC) buoy #42003, located on the right-hand-side of the storm track near the radius of maximum wind, recorded a minimum sea-level pressure (P_0) of 957.1 mb. Therefore, $\Delta P = (1013 - 957.1) = 55.9$ mb. Substituting this value in Eq. 13, $U_{10} = 47.1\ \mathrm{m\,s^{-1}}$ which is in excellent agreement with the measured $U_{10} = 47.3\ \mathrm{m\,s^{-1}}$.

Another verification is provided in Fig. 1. According to Anthes (20, p. 22 and Fig. 2.8),

$$U_{10\gamma} = U_{10\,\mathrm{max}}\left(\frac{R}{\gamma}\right)^{0.5} \tag{14}$$

where $U_{10\gamma}$ is the sustained wind speed at a distance 10 m away from the storm center and $U_{10\,\mathrm{max}}$ is the maximum sustained wind at 10 m at the radius of maximum wind, R.

According to Hsu et al. (21), for operational applications,

$$\frac{R}{\gamma} = \ln\left(\frac{1013 - P_0}{P_\gamma - P_0}\right) \tag{15}$$

where P_γ is the pressure at a point located at a distance from the storm center and ln is the natural logarithm.

Substituting Eq. 15 in Eq. 14,

$$U_{10\gamma} = U_{10\,\mathrm{max}}\left[\ln\left(\frac{1013 - P_0}{P_\gamma - P_0}\right)\right]^{0.5} \tag{16}$$

During Hurricane Lili in 2002, the NDBC had two buoys, #42001 located near R, and #42003 located due east along 26 °N, approximately 280 km from 42001. The wind speed measurement at both buoys was 10 m. From the NDBC website (www.ndbc.noaa.gov/), at 20Z 2 October 2002 at #42001, $P_0 = 956.1$ mb. Substituting this P_0 in Equation 13, $U_{10\,\mathrm{max}} = 47.5\ \mathrm{m\,s^{-1}}$, in excellent agreement with the measured value of 47.2 $\mathrm{m\,s^{-1}}$ (=106 mph). Therefore, Eq. 13 is further verified. At the same time, $P_\gamma = 1011.1$ mb was measured at #42003. Substituting this P_γ in Eq. 16, we obtain $U_{10\gamma} = 8.8\ \mathrm{m\,s^{-1}}$. The measured $U_{10\gamma}$ at 42003 was 9.2 $\mathrm{m\,s^{-1}}$. The difference is only about 4%, so we conclude that Eqs. 13 and 16 can be used for nowcasting using the pressure measurements at P_γ and P_0 which are normally available via the official "Advisory" during a hurricane.

ESTIMATING MAXIMUM SIGNIFICANT WAVE HEIGHT DURING A HURRICANE

According to the USACE (22, p. 3–85, Eq. 3–64),

$$T_\mathrm{p} = 12.1\sqrt{\frac{H_\mathrm{s}}{g}} \tag{17}$$

$$\therefore\ \frac{H_\mathrm{s}}{gT_\mathrm{p}^2} = 0.0068 \tag{18}$$

More verification of Eq. 18 is provided in Hsu (23).

According to Hsu et al. (21), based on the evaluation of nine fetch-limited wind–wave interaction formulas, that provided by Donelan et al. (24) ranked best as follows:

$$\frac{gH_\mathrm{s}}{U_{10}^2} = 0.00958\left(\frac{gT_\mathrm{p}}{U_{10}}\right)^{1.65} \tag{19}$$

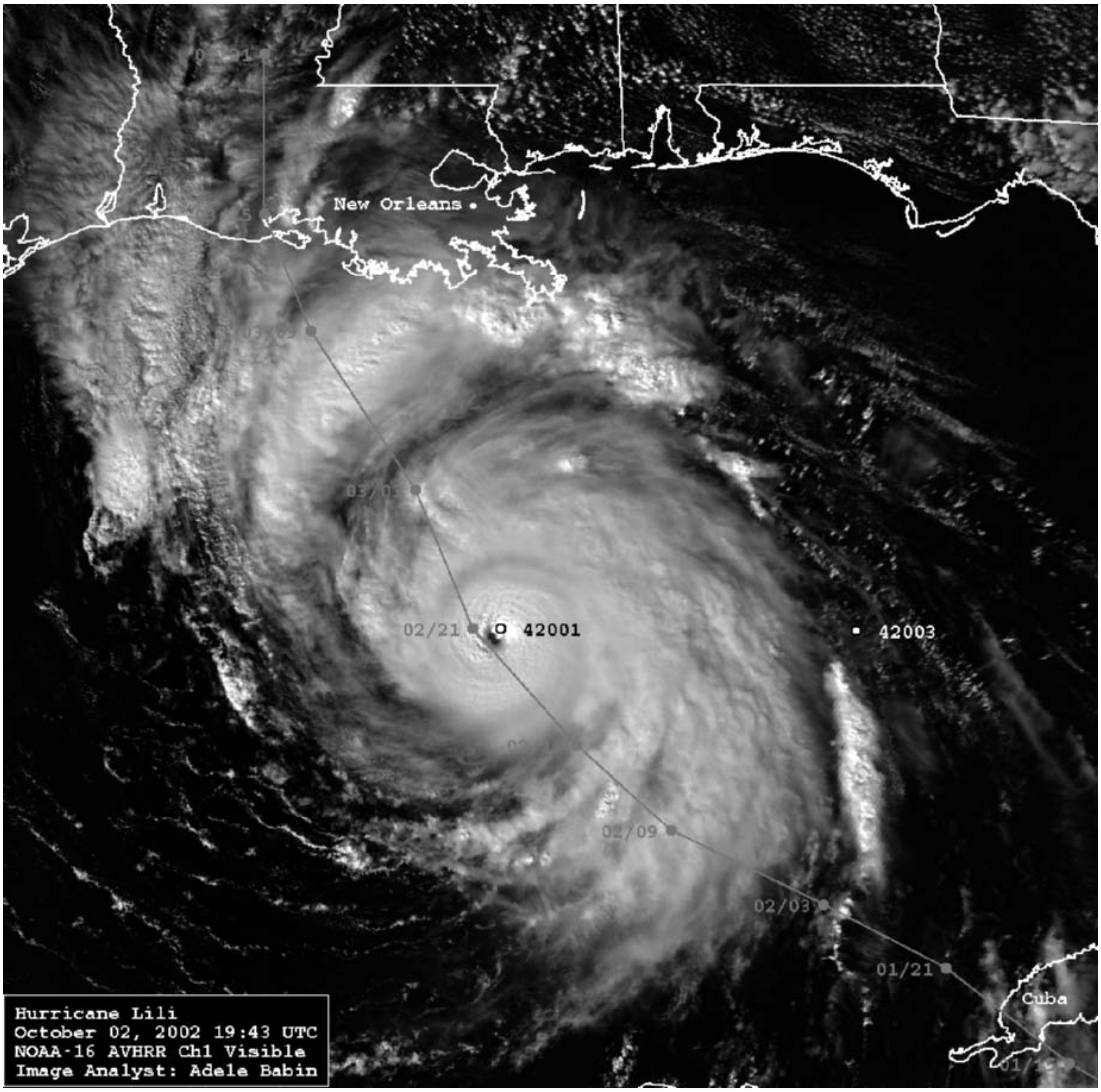

Figure 1. Satellite data (visible channel from NOAA-16) received and processed at the Earth Scan Lab, Louisiana State University, during Hurricane Lili (2002) in the Gulf of Mexico. The solid line represents the storm track. Data from NDBC buoys 42001 and 42003 are employed in this study. Note that the anemometers for both bouys were located at the standard 10 m height.

From Eqs. 18 and 19,

$$H_s = 0.00492\ U_{10}^2 \tag{20}$$

Substituting Eq. 13 in Eq. 20,

$$H_{s\ \max} = 0.20\ \Delta P \tag{21}$$

where $H_{s\ \max}$ is in meters and ΔP is in mb. Equation 21 is verified in Fig. 1. Buoy 42001, located near the radius of maximum wind, measured $P_0 = 956.1$ mb at 20Z 2 October 2002, so that $\Delta P = (1013 - 956.1) = 56.9$ mb. Substituting this value in Eq. 21, $H_{s\ \max} = 11.38$ m, which is in excellent agreement with that of 11.22 m measured at 21Z 2 October 2002 (within 1 hour after the measured minimal P_0).

ESTIMATING STORM SURGE AND SHOALING DEPTH

To estimate a typhoon/hurricane—generated storm surge (ΔS), and shoaling depth (D_{shoaling}), the following formulas are useful operationally, provided that the storm's minimum (or central) pressure near the sea surface (P_0) is known.

According to Hsu (23), for the storm surge in deep water before shoaling (i.e., when the waves feel the sea floor),

$$S_I = 0.070(1010 - P_0) \tag{22}$$

where S_I is the initial peak storm surge before shoaling. For the peak surge at the coast,

$$S_P = 0.070\,(1010 - P_0) * F_S * F_M \tag{23}$$

where F_S is a shoaling factor dependent on shelf topography and width and F_M is a correction factor for storm motion. Both F_S and F_M for certain areas are included in Hsu (23). A verification of Eq. 23 during Hurricane Georges in 1998 is also available in Hsu (23).

The shoaling depth is computed as follows:

From Taylor and Yelland (15), $D_{\text{shoaling}} = 0.2\ L_p$, and from Eq. 5,

$$D_{\text{shoaling}} = 0.2\,\frac{gT_p^2}{2\pi} = \frac{0.2\,H_s}{2\pi\left(\dfrac{H_s}{gT_p^2}\right)}$$

where $H_s/gT_p{}^2$ is called wave steepness, a useful parameter in coastal engineering.

From Hsu et al. (21) and under hurricane conditions from Eq. 18, $H_s/gT_p{}^2 = 0.0068$. Thus, from Eq. 21,

$$D_{\text{shoaling}} = 4.7\,H_s = 4.7 * 0.2(1013 - P_0)$$

$$\therefore \text{shoaling depth (meters)} \approx (1013 - P_0) \qquad (24)$$

CONCLUDING REMARKS

Although all formulas presented in this article are based on the open literature, they may need some verification before being applied to site-specific conditions. For example, Eq. 22 for the storm surge is for an open coast before shoaling. It needs to be adjusted for flooding at the coast due to different storm speeds and local bathymetry, as needed in Eq. 23.

BIBLIOGRAPHY

1. Geer, I.W. (Ed.). (1996). *Glossary of Weather and Climate With Related Oceanic and Hydrologic Terms*. American Meteorological Society, Boston, MA.
2. Donelan, M.A. (1990). Air-sea interaction. In: *The Sea*. Wiley, Vol. 9. Hoboken, NJ.
3. Csanady, G.T. (2001). *Air-Sea Interaction: Laws and Mechanisms*. Cambridge University Press, Cambridge, UK.
4. Liss, P.S. (1983). *Air-Sea Exchange of Gases and Particles*. Kluwer Academic, Springer, New York.
5. Donelan, M.A., Drennan, W.M., Saltman, E.S., and Wanninkhof, R. (Eds.). (2002). *Gas Transfer at Water Surfaces*. *Geophysical Monograph* #127, American Geophysical Society, Washington, DC.
6. Buat-Menard, P. (1986). *Role of Air-Sea Exchange in Geochemical Cycling*. Kluwer Academic, Springer, New York.
7. Liss, P.S. (1997). *Sea Surface and Global Change*. Cambridge University Press, Cambridge, UK.
8. Glantz, M.H. (2002). *La Niña and Its Impacts: Facts and Speculation*. Brookings Inst. Press, Washington, DC.
9. Geernaert, G.L. (1999). *Air-Sea Exchange: Physics, Chemistry & Dynamics*. Kluwer Academic, Springer, New York.
10. Janssen, P. (2004). *The Interaction of Ocean Waves and Wind*. Cambridge University Press, Cambridge, UK.
11. Hsu, S.A. (1988). *Coastal Meteorology*. Academic Press, San Diego, CA.
12. Hsu, S.A. and Blanchard, B.W. (2004). On the estimation of overwater buoyancy length from routine measurements. *Environ. Fluid Mech.*, in press.
13. Hsu, S.A. (1998). A relationship between the Bowen ratio and sea-air temperature difference under unstable conditions at sea. *J. Phys. Oceanogr.* **28**: 2222–2226.
14. Hsu, S.A. (1999). On the estimation of overwater Bowen ratio from sea-air temperature difference. *J. Phys. Oceanogr.* **29**: 1372–1373.
15. Taylor, P.K. and Yelland, M.J. (2001). The dependence of sea surface roughness on the height and steepness of the waves. *J. Phys. Oceanogr.* **31**: 572–590.
16. Amorocho, J. and DeVries, J.J. (1980). A new evaluation of the wind stress coefficient over water surfaces. *J. Geophys. Res.* **85**(C1): 433–442.
17. The WAMDI Group. (1988). The WAM model—a third generation ocean wave prediction model. *J. Phys. Oceanogr.* **18**: 1775–1810.
18. Hsu, S.A. (2003). Estimating overwater friction velocity and exponent of the power-law wind profile from gust factor during storms. *J. Waterways Port Coastal Ocean Eng.* **129**(4): 174–177.
19. Powell, M.D. (1982). The transition of the Hurricane Frederic boundary-layer wind field from the open Gulf of Mexico to landfall. *Mon. Weather Rev.* **110**: 1912–1932.
20. Anthes, R.A. (1982). *Tropical Cyclones, Their Evolution, Structure, and Effects*. *Meteorological Monographs Number* 41, American Meteorological Society, Boston, MA.
21. Hsu, S.A., Martin, M.F., Jr., and Blanchard, B.W. (2000). An evaluation of the USACE's deepwater wave prediction techniques under hurricane conditions during Georges in 1998. *J. Coastal Res.* **16**(3): 823–829.
22. U.S. Army Corps of Engineers. (1984). *Shore Protection Manual*. Vicksburg, MS.
23. Hsu, S.A. (2004). A wind-wave interaction explanation for Jelesnianski's open-ocean storm surge estimation using Hurricane Georges (1998) measurements. *Natl. Weather Dig.*, in press.
24. Donelan, M.A., Hamilton, J., and Hui, W.H. (1985). Directional spectra of wind-generated waves. *Philoso. Transa. Roy. Soc. London, Ser. A* **315**: 509–562.

NOAA'S ATLANTIC OCEANOGRAPHIC AND METEOROLOGICAL LABORATORY

National Oceanographic and
Atmospheric Administration
(NOAA)

June 18, 1999—The Atlantic Oceanographic and Meteorological Laboratory (AOML) in Miami, Florida, is one of 12 environmental research laboratories that work on environmental issues for NOAA's Office of Oceanic and Atmospheric Research (OAR). OAR research advances NOAA's ability to predict weather, helps monitor and provides understanding of climate and global change, as well as improve coastal ocean health.

AOML's mission is to conduct a basic and applied research program in oceanography, tropical meteorology, atmospheric and oceanic chemistry, and acoustics. The programs seek to understand the physical characteristics and processes of the ocean and the atmosphere, both individually and as a coupled system.

The principal focus of these investigations is to provide knowledge that may ultimately lead to improved prediction and forecasting of severe storms, better use and management of marine resources, better understanding of the factors affecting both climate and environmental quality, and improved ocean and weather services for the nation.

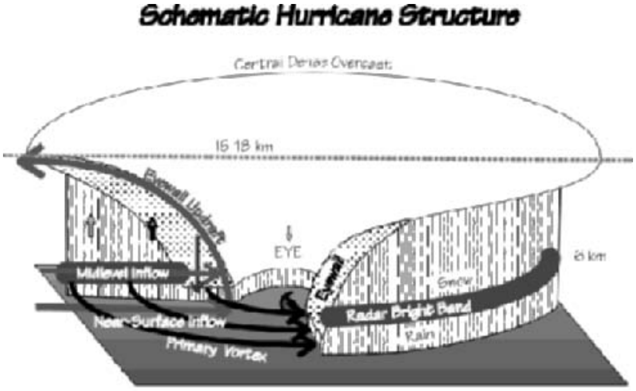

The inner core of a mature hurricane, showing the clear eye surrounded by the wall cloud, which is in turn surrounded by the overhanging Central Dense Overcast (CDO). The primary circulation is the air orbiting horizontally around an axis of rotation inside the eye. Superimposed upon the primary circulation is a secondary, in-up-out, circulation. The updrafts slope outward so that rain and snow or graupel fall from them.

Originally under the jurisdiction of the Environmental Science Services Administration (ESSA), the forerunner of NOAA, AOML was founded in Miami, Florida, in 1967. Several months after NOAA was established in 1970, groundbreaking began on a new 12-acre federally funded research facility on Virginia Key. AOML dedicated its new location on Feb. 9, 1973. It celebrated its 25th anniversary in 1998.

AOML has four main research divisions: Hurricane Research, Ocean Acoustics, Ocean Chemistry, and Physical Oceanography.

To learn more about AOML visit: http://www.aoml.noaa.gov/

HURRICANE RESEARCH DIVISION

An AOML/OAD Oceanographer prepares a 150 kHZ upward looking Acoustic Doppler Current Profiler (ADCP) for deployment 3 miles east of the Port of Miami. The 4 red transducers are covered with a mixture of cayenne pepper and silicone grease to prevent biofouling. The ADCP is used to obtain water column current profiles and provides such information as current speed and direction, information used in this project to determine optimal conditions for disposal of dredge materials.

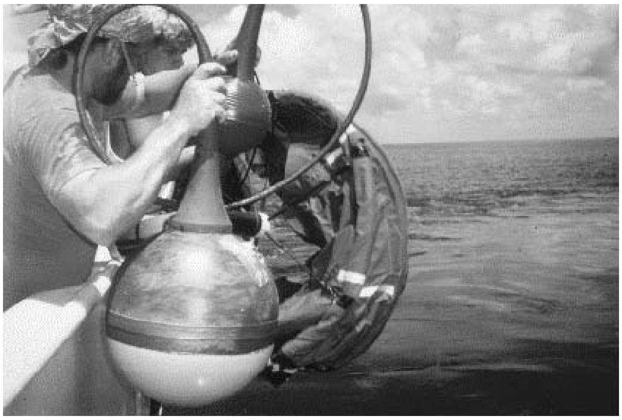

This image depicts the manual deployment of a global surface Lagrangian drifter (buoy). These devices measure sea surface temperature, surface currents, bathymetric pressure and wind. There are currently 756 of these instruments transmitting data back to AOML/PhOD from the oceanic regions all over the globe. The data re used for a variety of projects and programs.

The Hurricane Research Division (HRD) is NOAA's primary component for research on hurricanes. Its highest priority is improving the understanding and prediction of hurricane motion and intensity change. A key aspect of this work is the annual hurricane field program, supported by the NOAA Aircraft Operation's Center research/reconnaissance aircraft. Research teams analyze data from field programs, develop numerical hurricane models, conduct theoretical studies of hurricanes, prepare storm surge atlases, and study the tropical climate.

HRD works with the National Hurricane Center/Tropical Prediction Center in all phases of its research, the National Meteorological Center and the Geophysical Fluid Dynamics Laboratory—another of OAR's research labs—in research related to numerical modeling of hurricanes, and the National Severe Storms Laboratory—yet another OAR lab—in the study of landfalling hurricanes, as well as other NOAA groups, federal agencies, and universities in a variety of basic and applied research.

OCEAN ACOUSTICS DIVISION

The Ocean Acoustics Division (OAD) gathers, analyzes and reports coastal ocean data on human-related discharges and their potential environmental impacts. Additionally, OAD has an ongoing research program on the use of acoustics to measure coastal and deep ocean rainfall, an important element in calculating the global energy balance for climate monitoring and prediction. The Division works in cooperation with other federal, state, and local authorities to maximize research knowledge for use in economically and environmentally important projects in the coastal ocean.

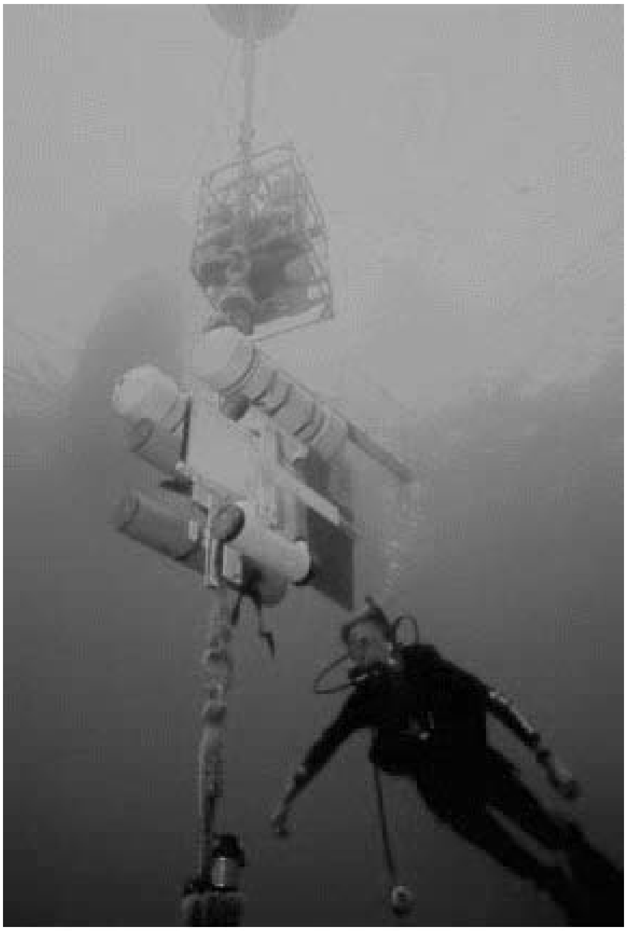

AOML/OCD and RSMAS scientists install multi-sensor optical/acoustic moorings off the Florida Keys in support of an ERL funded Pilot Study investigating the biology, physics and chemistry of Gulf Stream topographic and frontal eddies. The moorings provide time series data within an Ocean Surface Current Radar (OSCR) field.

OCEAN CHEMISTRY DIVISION

With a diverse scientific staff of marine, atmospheric, and geological chemists, as well as chemical, biological, and geological oceanographers, the Ocean Chemistry Division (OCD) is able to use multidisciplinary approaches to solve scientific research questions. The Division's work includes projects that are important in assessing the current and future effects of human activities on our coastal to deep ocean and atmospheric environments.

PHYSICAL OCEANOGRAPHY DIVISION

The Physical Oceanography Division (PhOD) provides and interprets oceanographic data and conducts research relevant to decadal climate change and coastal ecosystems. This research includes the dynamics of the ocean, its interaction with the atmosphere, and its role in climate and climate change. Data is collected from scientific expeditions, both in the deep ocean and in coastal regions. Satellite data is processed and incorporated into the analyses. PhOD manages the Global Ocean Observing (GOOS) Center, which manages the global collection, processing, and distribution of drifting buoy data and the information collected from ocean temperature profilers. This information is crucial to understanding and predicting shifts in weather patterns and the relationship of the ocean and the atmosphere as a coupled system.

LABORATORY EXPERIMENTS ON BIVALVE EXCRETION RATES OF NUTRIENTS

Paolo Magni
IMC—International Marine Centre
Torregrande-Oristano, Italy

Shigeru Montani
Hokkaido University
Hakodate, Japan

BACKGROUND

Benthic nutrient regeneration may be referred to as a new availability to the water column of significant amounts of nitrogen, phosphorus, and other nutrients, as a consequence of the metabolism of organic matter by the benthos (1). The processes of benthic nutrient regeneration in coastal marine systems are strongly influenced by the presence of abundant macrofauna (2–5). Correct evaluation of the biogenic flux of nutrients due to the excretory activity of infaunal species is therefore an important background of information to investigate the cycling of biophilic elements (nitrogen, phosphorus, and silicon).

In field studies, major drawbacks include the difficulty to discern between nutrient upward flux due to animal excretion and a number of local effects, such as microbial mineralization (4,6–8) and uptake (9–14), animal bioturbation and irrigation currents (15–19), tidal currents and wind-generated waves (20–22). Laboratory experiments on the animal excretion rates of nutrients under more controlled conditions represent a useful tool for quantifying the actual biogenic contribution by macrofauna to the total upward flux of nutrients from sediments. Nevertheless, these experiments have often restricted their investigations to ammonium (23–28) or, in a few cases, to ammonium and phosphate (29).

CASE STUDY

This study was conducted within a multidisciplinary project on the cycling of nutrients and organic matter in a tidal estuary in the Seto Inland Sea (30–37). Laboratory experiments were carried out on the excretion rates of ammonium, phosphate, and silicate by different size classes of the bivalves, *Ruditapes philippinarum* and *Musculista senhousia*. These species were selected as they were dominant on a sandflat of the estuary under investigation. An extrapolation of these results to a field community is presented in Temporal Scaling of Benthic Nutrient Regeneration in Bivalve-Dominated Tidal Flat. Both studies will be the basis of a third companion paper on the relationship between the temporal scaling of bivalve nutrient excretion and the seasonal change of nutrient concentrations in the porewater (Seasonal Coupling Between Intertidal Macrofauna and Sediment Column Porewater Nutrient Concentrations).

In these experiments, 2.5 L aquaria with and without (control) animals were employed and run on two different occasions for 24 hours. Each experiment consisted of a 10 h day (light) and a 10 h night (dark) treatment in which the changes in nutrient concentrations were measured every 2 h. Between the two (light and dark) treatments, a low tide lasting 2 h (like that approximately on the flat where animals were collected) was created, during which the experimental animals were not removed from sediments to keep the experiment continuous. The experimental setup and procedure are detailed in our associated paper where the bivalve excretion rates of ammonium and phosphate have been presented and discussed (35). We will extend this study to silicate, a nutrient species whose regeneration through animal excretory activity has been less investigated, either *in situ* (38,39) or in the laboratory (5,40). Table 1 includes some characteristics of the experimental animals, as well as the field-relevant (33) amount of algal food (*Thalassiosira* sp.) offered in four spikes during each experiment (35). In all treatments, there was a marked increase in all three nutrient concentrations, in the control (no animals), the increase was much more limited (i.e., silicate) or not observed (i.e., ammonium) (Fig. 1).

Based on the differences in nutrient concentrations between treatments and controls, relevant linear regression lines of five to six measurements were used to calculate the nutrient excretion rates for each size class of *R. philippinarum* and *M. senhousia* (Table 2). This approach may be a more reliable way to quantify the daily bivalve excretion, whereas previous similar experiments have been based on shorter incubations and/or the sole difference between initial and final values of nutrient concentrations (4,7,22,28,29). The data sets were subjected to ANOVA in a two-factor randomized complete block design, using the day/night variable as factor A, the time (hour) variable as factor B, and the size classes of each bivalve species as replicates (35). As found for ammonium in *R. philippinarum*, but not in *M. senhousia*, silicate excretion was significantly higher (57%, $p < 0.001$, $n = 36$) during the day than during the night, suggesting a possible effect of light on the excretory activity of this bivalve species. A comparison of nutrient excretion rates (μmol g^{-1} DW h^{-1}) of bivalve species obtained through *in situ* or laboratory experiments is given in Table 3.

According to the excretion rates of silicate found in our laboratory experiments, this study points to the importance of the excretory activity of these bivalve species to the biogenic regeneration of silicate. This aspect

Table 1. Animals Employed in the Laboratory Experiments and Experimental Conditions. Ind.: Number of Individuals[a,b]

Species		Size, mm	Ind., n	TOT, mg	DW, mg	Temp, °C	Chl*a*, (μg L^{-}1)	Expt, date
				Ruditapes philippinarum				
Size class I	AVG	9.4	12	197	9.9	19.6	26.3	May 1996
	SD	±1.4	–	±79	±4.0	±1.5	±8.7	–
Size class II	AVG	15.5	15	830	37.0	19.6	24.6	May 1996
	SD	±1.0	–	±174	±9.2	±1.5	±8.8	–
Size class III	AVG	18.9	9	1520	63.6	21.6	38.9	Sep 1996
	SD	±0.8	–	±149	±10.4	±0.3	±12.8	–
				Musculista senhousia				
Size class I	AVG	16.7	14	431	27.6	19.6	24.5	May 1996
	SD	±1.3	–	±117	±7.7	±1.5	±9.1	–
Size class II	AVG	23.5	8	1264	52.4	21.6	47.4	Sep 1996
	SD	±1.7	–	±172	±8.0	±0.3	±22.8	–

[a] Reproduced from Reference 35.

[b] TOT: mean (live) weight for each size class of the experimental animals; DW: mean dry soft-body weight for each size class of the experimental animals; Temp: experimental temperature; Chl a (Chlorophyll a) is the mean (AVG) ± standard deviation (SD) of four spikes of cultures of *Thalassiosira* sp. (Chl $a = 0.01 \times$ *Thalassiosira*sp. cell $+ 3.6$, $r^2 = 0.908$; $p < 0.001$, $n = 40$) for each day/night treatment

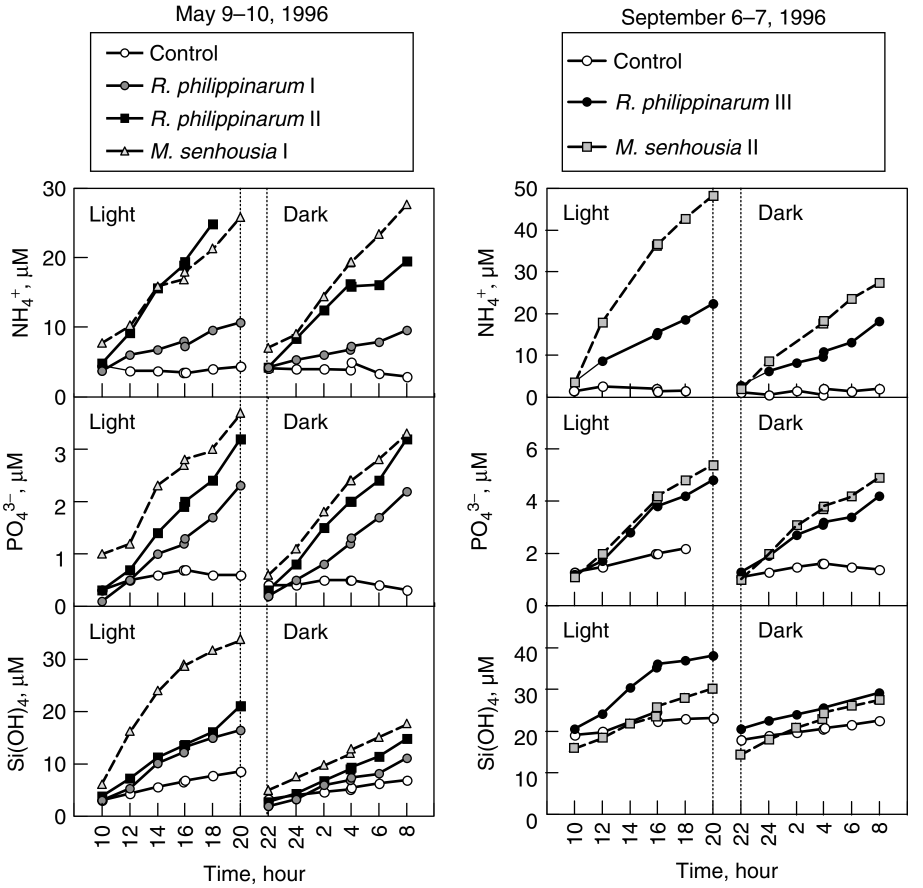

Figure 1. Laboratory experiments: changes of nutrient concentrations [ammonium, NH_4^+–N); phosphate, PO_4^{3-}–P; and silicate $Si(OH)_4$–Si] during the day (light) and night (dark) treatments. Vertical dashed lines: left line (time 20:00) indicates the end of the day (light) treatment; right line (time 22:00) indicates the start of the night (dark) treatment; between lines: low tide between treatments (based on Ref. 35).

Table 2. Nutrient Excretion Rate for Each Size class of *Ruditapes philippinarum* and *Musculista senhousia* During day/night Treatments (Experimental Temperature as in Table 1)[a]

	Nutrient Excretion Rate, μmol g^{-1}DW h^{-1}					
	NH_4^+		PO_4^{3-}		$Si(OH)_4$	
Species	Light	Dark	Light	Dark	Light	Dark
Ruditapes philippinarum						
Size class I	10.6	7.9	3.4	3.9	15.8	10.9
Size class II	9.6	5.8	1.0	1.1	4.1	3.3
Size class III	5.0	3.8	0.9	0.7	4.0	1.1
Light/dark mean	8.4[b]	5.8	1.8	1.9	8.0[b]	5.1
Total mean	7.1		1.9		6.6	
Musculista senhousia						
Size class I	9.3	11.4	1.2	1.5	14.5	4.8
Size class II	16.9	9.7	1.6	1.3	4.2	5.5
Light/dark mean	13.1	10.6	1.4	1.4	9.4	5.2
Total mean	11.8		1.4		7.3	

[a] Based on Reference 35.

[b] Mean day (light) excretion significantly higher (ANOVA $p < 0.001$) than night (dark) excretion (based on Reference 35).

has been more controversial than the contribution of bivalves to the regeneration of ammonium and phosphate. The extent of silicate excretion also varies considerably, depending on the bivalve species investigated and the environmental/experimental conditions employed.

In *in situ* experiments, Prins & Small (4) found no significant excretion of silicate by *Mytilus edulis* beds on an intertidal zone of the Westerschelde (The Netherlands). The occurrence of silicate fluxes was attributed to the possible increased rate of dissolution of silicate at higher temperature. Asmus et al. (38) in the eastern Wadden Sea (Germany) and Dame et al. (39) in the Easterschelde and the western Wadden Sea (The Netherlands) found high fluxes of silicate from mussel beds of *M. edulis*. Although, in both studies, the actual excretion rate of silicate by *M. edulis* was not estimated per biomass unit (e.g., μmol g^{-1} DW h^{-1}), Dame et al. (39) suggested that silicate release from mussel beds results from phytoplankton cells breaking down as they are metabolized by the mussels. Dame et al. (39) argued that the longer turnover time for silicate, compared to phosphate and ammonium, implies a lesser role for the mussel beds in recycling this nutrient species in the two estuaries under investigation.

In contrast, Asmus et al. (38) found rapid silicate release in the Sylt-flume study and suggested that the mussels are an accelerator in recycling biogenic silica. Similarly, a study on the nutrient excretion of *R. philippinarum* in core incubation experiments found that silicate regeneration was, on average, 9.2 times faster in the site farmed with clams (5). In mesocosm experiments using large tanks, Doering et al. (40) also found that the level of flux was elevated in the presence of another clam, *Mercenaria mercenaria*, by 86% and 57% for silicate and ammonium, respectively. Our results indicate

Table 3. Comparison of Nutrient Excretion Rates (μmol g^{-1} DW hour^{-1}) for Different Species of Mussels (m), Clams (c) and Oysters (o)[a]

Species and Study Area	Method[b]	NH_4^+	PO_4^{3-}	$Si(OH)_4$	T, °C	Reference
Mytilus edulis (m)						
Narragansett Bay, USA	Lab	3.1	nd[c]	nd[c]	15	41
Linher River, U.K.	Lab	4.9–34.6	nd	nd	11–21	23
Sound, DK	*In situ*	0.14–3.1	0.10–0.53	nd	0.7–18	42
Western Scheldt, NL	*In situ*	1.1	nd	nd	12	43
Musculista senhousia (m)						
Seto Inland Sea, JPN	Lab	9.3–16.9	1.2–1.6	–	18–22	35
Seto Inland Sea, JPN	Lab	9.3–16.9	1.2–1.6	4.2–14.5	18–22	This study
Modiolus demissus (m)						
Narragansett Bay, USA	Lab	3.58 ± 1.73	nd	nd	21	41
Great Sippewissett, USA	Lab	2.5	nd	nd	annual	2
Donax serra (m)						
Maitland River, S. Africa	*In situ*	0.35–8.1	nd	nd		44
Sundays River, S. Africa	Lab/*In situ*	2.2	nd	nd	15–25	26
Donax sordidus (m)						
Sundays River, S. Africa	Lab/*In situ*	2.9	nd	nd	15–25	26
Aspatharia wahlbergi (m)						
Lake Kariba, Zimbabwe	Lab	6.1	0.48	nd	25.2	29
Corbicula africana (c)						
Lake Kariba, Zimbabwe	Lab	12.9	nd	nd	25.2	29
Corbicula japonica (c)						
Lake Shinji, JPN	Lab	14.3	nd	nd	27	28
Mercenaria mercenaria (c)						
Delaware Bay, USA	Lab	0.9–1.5	nd	nd	20	45
Macoma balthica (c)						
Wadden Sea, DK	Lab	0.1[d]	nd	nd	13–15	16
Ruditapes philippinarum (c)						
Virgin Islands, USA	Lab	1.9–4.9	nd	nd	20.1	24
Moss Landing, USA	Lab	1–2.3	nd	nd	27.1	46
Moss Landing, USA	Lab	0.6–0.9	nd	nd	12, 15, 18	47
Marennes-Oléron, F	Lab	0.5–13	nd	nd	5–25	25
Hatchery, Ireland	Lab	0.16–1	nd	nd	18.8	27
Seto Inland Sea, JPN	Lab	3.8–10.6	0.7–3.9	—	18–22	35
Seto Inland Sea, JPN	Lab	3.8–10.6	0.7–3.9	1.1–15.8	18–22	This study
Crassostrea virginica (o)						
Delaware Bay, USA	Lab	0.5–0.9	nd	nd	20.1	24
Crassostrea gigas (o)						
North Brittany, F	*In situ*	0.28–6.6	nd	nd	27.1	48

[a] Based on Reference 35.

[b] Lab: laboratory experiments.

[c] nd: not determined.

[d] Excretion rate calculated as a wet soft-body weight.

well-balanced stoichiometric ratios among the nutrient species excreted by the bivalves (35), which can also be related to the use of the diatom *Thalassiosira* sp. as a food (35). This was aimed to approximate the actual field situation on the tidal flat, where abundant microalgal biomass, including resuspended benthic diatoms, is available to filter-feeders such as in *R. philippinarum* and *M. senhousia* (31,33). The high excretion rates of silicate of these two dominant bivalve species found in our laboratory experiments, together with those of ammonium and phosphate, suggest a major contribution of bivalve nutrient excretion to the upward flux of nutrients from sediments in our relevant study area. This will be the subject of the subsequent paper where we will apply these rates to the actual bivalve standing stock found in the field (Temporal Scaling of Benthic Nutrient Regeneration in Bivalve-Dominated Tidal Flat).

BIBLIOGRAPHY

1. Nixon, S.W., Oviatt, C.A., and Hale, S.S. (1976). *Nitrogen Regeneration and the Metabolism of Coastal Marine Bottom Communities. The Role of Terrestrial and Aquatic Organisms in Decomposition Process*. J. Anderson and A. MacFayden (Eds.). Blackwell, Oxford, pp. 269–283.
2. Jordan, T.E. and Valiela, I. (1982). A nitrogen budget of the ribbed mussel, *Geukensia demissa* and its significance in nitrogen flow in a New England salt marsh. *Limnol. Oceanogr.* **27**: 75–90.
3. Murphy, R.C., and Kremer, J.N. (1985). Bivalve contribution to benthic metabolism in a California lagoon. *Estuaries* **8**: 330–341.
4. Prins, T.C. and Smaal, A.C. (1994). The role of the blue mussel *Mytilus edulis* in the cycling of nutrients in Oosterschelde estuary (the Netherlands). *Hydrobiol.* **282/283**: 413–429.

5. Bartoli, M. et al. (2001). Impact of *Tapes philippinarum* on nutrient dynamics and benthic respiration in the Sacca di Goro. *Hydrobiol.* **455**: 203–212.
6. Falcao, M. and Vale, C. (1990). Study of the Rio Formosa ecosystem: benthic nutrient remineralization and tidal variability of nutrients in the water. *Hydrobiol.* **207**: 137–146.
7. Gardner, W.S., Briones, E.E., Kaegi, E.C., and Rowe, G.T. (1993). Ammonium excretion by benthic invertebrates and sediment–water nitrogen flux in the Gulf of Mexico near the Mississipi river outflow. *Estuaries* **16**: 799–808.
8. Smaal, A.C. and Zurburg, W. (1997). The uptake and release of suspended and dissolved material by oysters and mussels in Marennes-Oléron Bay. *Aquat. Living Resour.* **10**: 23–30.
9. Rizzo, W.M. (1990). Nutrient exchanges between the water column and a subtidal benthic microalgal community. *Estuaries* **13**: 219–226.
10. Sundbäck, K., Enoksson, V., Granéli, W., and Pettersson, K. (1991). Influence of sublittoral microphytobenthos on the oxygen and nutrient flux between sediment and water: a laboratory continuous-flow study. *Mar. Ecol. Prog. Ser.* **74**: 263–279.
11. van Duyl, F.C., van Raaphorst, W., and Kop, A.J. (1993). Benthic bacterial production and nutrient sediment-water exchange in sandy North Sea sediments. *Mar. Ecol. Prog. Ser.* **100**: 85–95.
12. Risgaard-Petersen, N., Rysgaard, S., Nielsen, L.P., and Revsbech, N.P. (1994). Diurnal variation of denitrification and nitrification in sediments colonized by benthic microphytes. *Limnol. Oceanogr.* **39**: 573–579.
13. Feuillet-Girard, M., Gouleau, D., Blanchard, G., and Joassard, L. (1997). Nutrient fluxes on an intertidal mudflat in Marennes-Oléron Bay, and influence of the emersion period. *Aquat. Living Resour.* **10**: 49–58.
14. Trimmer, M., Nedwell, D.B., Sivyer, D.B., and Malcolm, S.J. (1998). Nitrogen fluxes through the lower estuary of the river Great Ouse, England: the role of the bottom sediments. *Mar. Ecol. Prog. Ser.* **163**: 109–124.
15. Callender, E. and Hammond, D.E. (1982). Nutrient exchange across the sediment–water interface in the Potomac river estuary. *Estuar. Coast. Shelf Sci.* **15**: 395–413.
16. Henriksen, K., Rasmussen, M.B., and Jensen, A. (1983). Effect of bioturbation on microbial nitrogen transformations in the sediment and fluxes of ammonium and nitrate to the overlaying water. *Environ. Biogeochem.* **35**: 193–205.
17. Yamada, H. and Kayama, M. (1987). Liberation of nitrogenous compounds from bottom sediments and effect of bioturbation by small bivalve, *Theora lata* (Hinds). *Estuar. Coast. Shelf Sci.* **24**: 539–555.
18. Aller, R.C. (1988). Benthic fauna and biogeochemical processes: the role of borrow structures. In: *Nitrogen Cycling in Coastal Marine Environments*. T.H. Blackburn and J. Sørensen (Eds.). Wiley, pp. 301–338.
19. Luther, I.II. et al. (1998). Simultaneous measurements of O_2, Mn, Fe, I^-, and S(-II) in marine pore waters with a solid-state voltammetric microelectrode. *Limnol. Oceanogr.* **43**: 325–333.
20. Nakamura, Y. (1994). Effect of flow velocity on phosphate release from sediment. *Water Sci. Technol.* **30**: 263–272.
21. Miller-Way, T. and Twilley, T.T. (1996). Theory and operation of continuous flow systems for the study of benthic–pelagic coupling. *Mar. Ecol. Prog. Ser.* **140**: 257–269.
22. Asmus, R.M. et al. (1998). The role of water movement and spatial scaling for measurements of dissolved inorganic nitrogen fluxes in intertidal sediments. *Estuar. Coast. Shelf Sci.* **46**: 221–232.
23. Bayne, B.L. and Scullard, C. (1977). Rates of nitrogen excretion by species of *Mytilus* (Bivalvia: Mollusca). *J. Mar. Biol. Ass. U.K.* **57**: 355–369.
24. Langton, R.W., Haines, K.C., and Lyon, R.E. (1977). Ammonia-nitrogen production by the bivalve mollusc *Tapes japonica* and its recovery by the red seaweed *Hypnea musciformis* in a tropical mariculture system. *Helgol. Wiss Meeresunters* **30**: 217–229.
25. Goulletquer, P. et al. (1989). Ecophysiologie et bilan énergétique de la palourde japonaise d' élevage *Ruditapes philippinarum*. *J. Exp. Mar. Biol. Ecol.* **132**: 85–108.
26. Cockcroft, A.C. (1990). Nitrogen excretion by the surf zone bivalves *Donax serra* and *D. sordidus*. *Mar. Ecol. Prog. Ser.* **60**: 57–65.
27. Xie, Q. and Burnell, G.M. (1995). The effect of activity on the physiological rates of two clam species, *Tapes philippinarum* (Adams & Reeve) and *Tapes decussatus* (Linnaeus). *Biol. Environ. Proc. Royal Irish Acad.* **95B**: 217–223.
28. Nakamura, M., Yamamuro, M., Ishikawa, M., and Nishimura, H. (1988). Role of the bivalve *Corbicula japonica* in the nitrogen cycle in a mesohaline lagoon. *Mar. Biol.* **99**: 369–374.
29. Kiibus, M., and Kautsky, N. (1996). Respiration, nutrient excretion and filtratin rate of tropical freshwater mussels and their contribution to production and energy flow in Lake Kariba, Zimbabwe. *Hydrobiol.* **331**: 25–32.
30. Magni, P. (2000). Distribution of organic matter in a tidal estuary of the Seto Inland Sea, Japan, and its relationship with the macrozoobenthic communities. In: *Ad hoc Benthic Indicator Group—Results of Initial Planning Meeting*. IOC Technical Series No. 57, UNESCO, pp. 20–25.
31. Magni, P. and Montani, S. (1997). Development of benthic microalgal assemblages on an intertidal flat in the Seto Inland Sea, Japan: effects of environmental variability. *La mer* **35**: 137–148.
32. Magni, P. and Montani, S. (1998). Responses of intertidal and subtidal communities of the macrobenthos to organic load and oxygen depletion in the Seto Inland, Japan. *J. Rech. Océanogr.* **23**: 47–56.
33. Magni, P. and Montani, S. (2000). Physical and chemical variability in the lower intertidal zone of an estuary in the Seto Inland Sea, Japan: seasonal patterns of dissolved and particulate compounds. *Hydrobiol.* **432**: 9–23.
34. Magni, P., Abe, N., and Montani, S. (2000). Quantification of microphytobenthos biomass in intertidal sediments: layer-dependent variation of chlorophyll *a* content determined by spectrophotometric and HPLC methods. *La mer* **38**: 57–63.
35. Magni, P., Montani, S., Takada, C., and Tsutsumi, H. (2000). Temporal scaling and relevance of bivalve nutrient excretion on a tidal flat of the Seto Inland Sea, Japan. *Mar. Ecol. Progr. Ser.* **198**: 139–155.
36. Magni, P., Montani, S., and Tada, K. (2002). Semidiurnal dynamics of salinity, nutrients and suspended particulate matter in an estuary in the Seto Inland Sea, Japan, during a spring tide cycle. *J. Oceanogr.* **58**: 389–402.
37. Montani, S. et al. (1998). The effect of a tidal cycle on the dynamics of nutrients in a tidal estuary in the Seto Inland Sea, Japan. *J. Oceanogr.* **54**: 65–76.
38. Asmus, H., Asmus, R.M., and Reise, K. (1990). Exchange processes in an intertidal mussel bed: A Sylt-flume study in the Wadden Sea. *Berichte Biolog. Anstalt Helgol.* **6**: 78.
39. Dame, R.F. et al. (1991). The influence of mussel beds on nutrients in the western Wadden Sea and Eastern Scheldt estuaries. *Estuaries* **14**: 130–138.

40. Doering, P.H., Kelly, J.R., Oviatt, C.A., and Sowers, T. (1987). Effect of the hard clam *Mercenaria mercenaria* on benthic fluxes of inorganic nutrients and gases. *Mar. Biol.* **94**: 377–383.
41. Nixon, S.W., Oviatt, C.A., Garber, J., and Lee, V. (1976). Diel metabolism and nutrient dynamics in a salt marsh embayment. *Ecology* **57**: 740–750.
42. Schlüter, L. and Josefsen, S.B. (1994). Annual variation in condition, respiration and remineralisation of *Mytilus edulis* L. in the Sound, Denmark. *Helgol. Meeres.* **48**: 419–430.
43. Smaal, A.C., Vonck, A.P.M.A., and Bakker, M. (1997). Seasonal variation in physiological energetics of *Mytilus edulis* and *Cerastoderma edule* of different size classes. *J. Mar. Biol. Ass. U.K.* **77**: 817–838.
44. Prosh, R.M. and McLachlan, A. (1984). The regeneration of surf-zone nutrients by the sand mussel, *Donax Serra* Röding. *J. Exp. Mar. Biol. Ecol.* **80**: 221–233.
45. Snra, R.F. and Baggaley, A. (1976). Rate of excretion of ammonia by the hard clam *Mercenaria mercenaria* and the American oyster *Crassostrea virginica*. *Mar. Biol.* **36**: 251–258.
46. Mann, R. and Glomb, S.J. (1978). The effect of temperature on growth and ammonia excretion of the Manila clam *Tapes japonica*. *Estuar. Coast. Shelf Sci.* **6**: 335–339.
47. Mann, R. (1979). The effect of temperature on growth, physiology, and gametogenesis in the Manila clam *Tapes philippinarum* (Adams & Reeve, 1850). *J. Exp. Mar. Biol. Ecol.* **38**: 121–133.
48. Boucher, G. and Boucher-Rodoni, R. (1988). In situ measurements of respiratory metabolism and nitrogen fluxes at the interface of oyster beds. *Mar. Ecol. Prog. Ser.* **44**: 229–238.

TEMPORAL SCALING OF BENTHIC NUTRIENT REGENERATION IN BIVALVE-DOMINATED TIDAL FLAT

Paolo Magni
IMC—International Marine Centre
Torregrande-Oristano, Italy

Shigeru Montani
Hokkaido University
Hakodate, Japan

Beside light and temperature (1–5), nutrients such as ammonium ($NH_4{}^+$-N), phosphate ($PO_4{}^{3-}$-P), and silicate [$Si(OH)_4$-Si] are a key factor in controlling the growth, abundance, and structure of primary producers in the ocean (6,7). Hence, it is important to investigate the availability, sources, and distribution of these nutrient species, as well as their spatial and temporal scaling. Biological processes strongly influence nutrient regeneration in different marine systems. In the open ocean, an important portion of reduced N-forms (e.g., $NH_4{}^+$-N), for instance, is made available *in situ* from waste products of plankton metabolism (8,9) and supports the so-called "regenerated" primary production (10). In coastal marine ecosystems, benthic nutrient regeneration is a major driving force in cycling biophilic elements (e.g., N, P, and Si) (11–13) and abundant macrofauna, for example, dense assemblages of bivalves, it has been shown, play a major role in these processes (14–16).

BACKGROUND

The contribution of benthic macrofauna to the total upward flux of nutrients has been investigated (mostly for ammonium and, to a lesser extent, for phosphate) in many coastal and estuarine areas using several approaches. They include laboratory and mesocosm experiments (17–19). *In situ* benthic chambers and sediment core incubations (20–22), and open flow/tunnel systems (16,23–26). Measurements of macrofauna-influenced nutrient flux, however, often have temporal limitations as they are based on one or relatively few sampling occasions, and thus seasonal patterns are in most cases not known. In this article, we evaluate the magnitude and temporal scaling of biogenic flux of nutrients from intertidal sediments densely populated by bivalves, based on extrapolating nutrient excretion rates of dominant bivalves to a field community. In particular, we show that the seasonal pattern of nutrient fluxes can be strongly influenced by the animal standing stock and its temporal distribution. This is beside the effect and importance of variation in excretion rate due to animal physiological factors, such as seasonal cycles of gametogenesis, storage and use of body reserves, and water temperature (27,28). The nutrient species considered in this study include ammonium, phosphate, and silicate, for which we quantified in associated laboratory experiments Laboratory Experiments On Bivalve Excretion Rates of Nutrients the excretion rates of different size classes of two dominant bivalve species. The relevance of macrofaunal excretion in regenerating the inorganic forms of three major bioelements such as N, P, and Si is discussed.

MACROFAUNAL COMMUNITIES

We present here the macrofaunal composition and distribution at an individual station (Stn B5) of a transect line selected in a sandflat of the Seto Inland Sea of Japan (29). At this station, the porewater nutrient concentrations (ammonium, phosphate, and silicate) in the uppermost 10 cm of sediments were also investigated in parallel from January 1995 to April 1996. They will be the subject of a subsequent associated paper focusing on the relationship between the seasonal pattern of bivalve nutrient excretion, described here, and the seasonal variation of porewater nutrient concentrations Seasonal Coupling Between Intertidal Macrofauna and Sediment Column Porewater Nutrient Concentrations. The total density and biomass of macrofauna varied from 7,400 (July 1995) to 22,050 ind.m^{-2} (October 1995), and from 70.9 (July 1995) to 244 g DW m^{-2} (September 7, 1995), respectively (Fig. 1). The bivalves *Ruditapes philippinarum* and *Musculista senhousia* and the polychaetes *Ceratonereis erithraeensis* and *Cirriformia tentaculata* were dominant; they accounted for 60.5% and 94.7% of the total density and biomass, respectively. Remarkably, *R. philippinarum* and *M. senhousia* alone accounted for up to 83.3 ± 6.7% of the total biomass when this exceeded

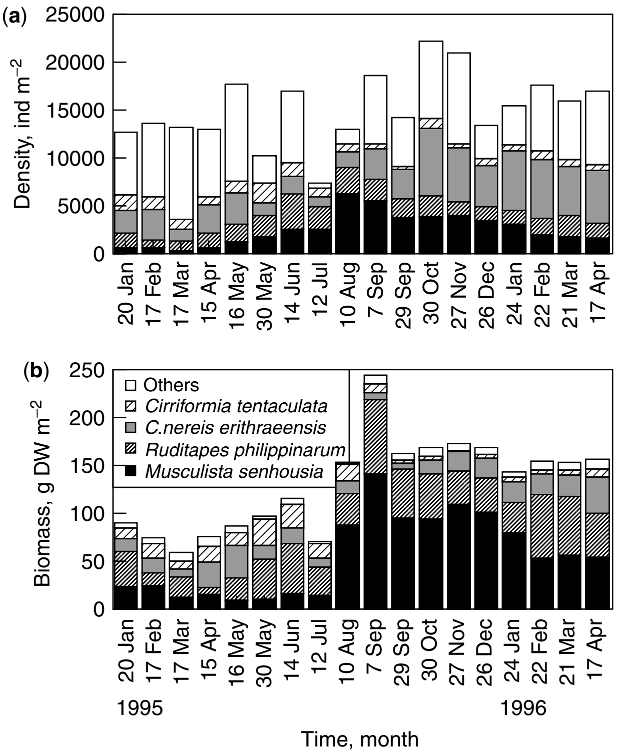

Figure 1. Seasonal variation of density (**a**) and biomass (**b**) of dominant macrozoobenthic species at Stn B5 (29). Note that in May 1995 and September 1995, sampling was carried out fortnightly.

120 g DW m^{-2}, from August 1995 till the end of the investigations. The high values of total macrofaunal and bivalve biomass may be regarded as a typical feature of many estuarine and intertidal areas, which are amongst the most productive systems in the ocean (30). In addition, biomass was markedly lower during the first half of the year (January 1995 to July 1995) than between late summer and winter; yet values progressively increased from early spring (March 1995) to early summer (June 1995). These marked temporal changes of macrofaunal communities reflect the high variability of these ecosystems. To estimate the magnitude and temporal scaling of biogenic nutrient excretion, we used an indirect approach. The mean excretion rates of ammonium, phosphate, and silicate for the two dominant bivalves *R. philippinarum* and *M. senhousia*, which were obtained in laboratory experiments Laboratory Experiments On Bivalve Excretion Rates of Nutrients, were applied to the relevant monthly biomass values found in the field. The ammonium and phosphate excretion rates of each bivalve species and their scaling to a field community have been extensively reported in Magni et al. (29). In this article, we applied these excretion rates to the bivalve biomass found at Stn B5 and extended this scaling to silicate, whose size-class dependent excretion rates are presented in Laboratory Experiments On Bivalve Excretion Rates of Nutrients. For silicate, we adopted the same temperature-dependent excretion rate factors as those used for ammonium and phosphate (Table 1).

BIVALVE NUTRIENT EXCRETION

The highest excretion rates of nutrients were estimated in September 7, 1995, up to a total of 50.2, 7.5, and 34.1 mmol m^{-2} d^{-1} for ammonium, phosphate, and silicate, respectively (Fig. 2). This corresponded to the period of highest biomass of both *R. philippinarum* and *M. senhousia*, which also accounted for the highest bivalve percentage (91.8%) of the total macrofaunal biomass. The lowest excretion rates occurred in April 1995 for *R. philippinarum* (lowest biomass on the same occasion), in March 1995 for *M. senhousia* (lowest biomass on May 16, 1995), and in February 1995 as the sum of the two bivalve species excretion rates. These latter rates were 4.1, 0.64, and 2.9 mmol m^{-2} d^{-1} for ammonium, phosphate, and silicate, respectively. The upward flux rates of nutrients obtained through this extrapolation of laboratory experiments on bivalve nutrient excretion to a field community are comparable to the highest biogenic releases reported for dense assemblages of bivalves such as oyster reefs (34) and mussel beds (16,35). This study also points to the importance of bivalve excretion to the biogenic regeneration of silicate, as previously suggested by field measurements indicating evidence of increased levels of silicate flux in the

Table 1. Adopted Temperature-Dependent Excretion Rates of Ammonium (NH_4^+-N), Phosphate (PO_4^{3-}-P), and Silicate [$Si(OH)_4$-Si] for *Ruditapes philippinarum* and *Musculista senhousia*

Period		Temperature, °C			Excretion Rate, μmol g^{-1}DW h^{-1}							
Month		Station			*R. philippinarum*				*M. senhousia*			
		B5-B1	H1	Y3	f^a	NH_4^+	PO_4^{3-}	$Si(OH)_4$	f^a	NH_4^+	PO_4^{3-}	$Si(OH)_4$
Dec, Jan,	AVG	4.7	5.9	10.4	0.6	4.3	1.1	4.0	0.5	5.9	0.7	3.7
Feb	SD	±1.8	±1.8	±2.1	–	–	–	–	–	–	–	–
Nov, Mar	AVG	11.7	11.9	15.2	0.9	6.4	1.7	5.9	0.7	8.3	1.0	5.1
–	SD	±0.9	±0.2	±3.0	–	–	–	–	–	–	–	–
Apr, May,	AVG	19.6	19.9	21.7	1	7.1	1.9	6.6	1	11.8	1.4	7.3
Jun, Oct	SD	±2.8	±4.0	±3.6	–	–	–	–	–	–	–	–
Jul, Aug,	AVG	27.7	28.0	26.9	0.9	6.4	1.7	5.9	1.2	14.2	1.7	8.8
Sep	SD	±2.3	±3.1	±1.1	–	–	–	–	–	–	–	–

[a] A factor (*f*) of 1 is used for the mean of the excretion rates obtained in laboratory experiments (see Laboratory Experiments On Bivalve Excretion Rates of Nutrients and based on Ref. 29).

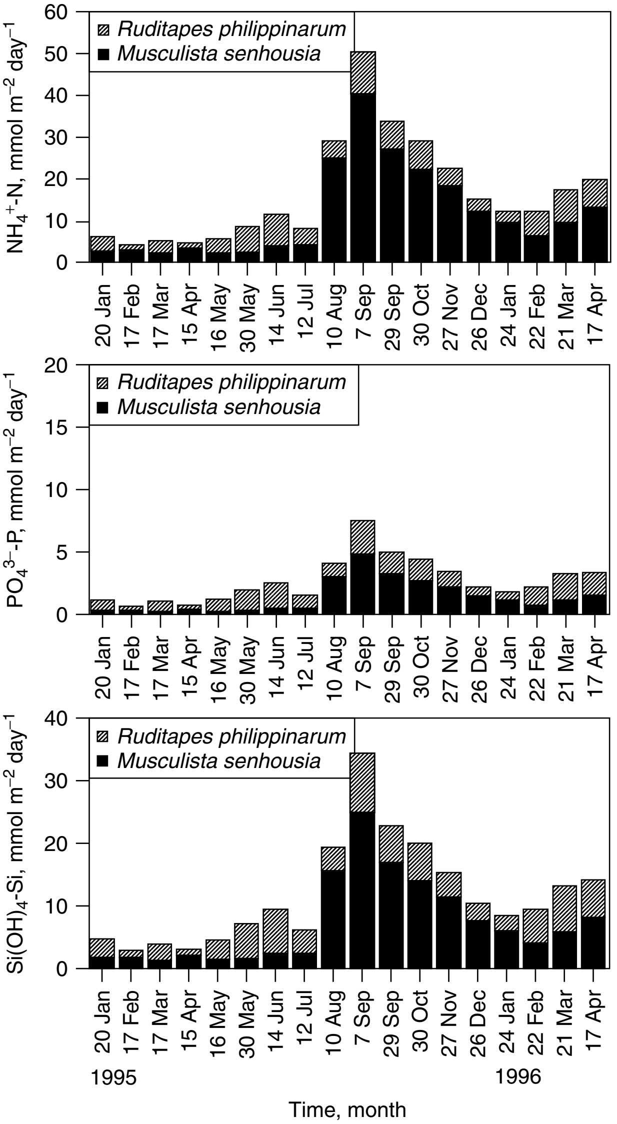

Figure 2. Magnitude and temporal scaling of ammonium (NH_4^+-N), phosphate (PO_4^{3-}-P), and silicate [$Si(OH)_4$-Si] excretion by *Ruditapes philippinarum* and *Musculista senhousia* in a field community.

presence of bivalves (17,24,36). The temporal scaling of bivalve nutrient excretion showed a marked seasonal pattern of large variations of nutrient flux, up to ca. 10-fold (*R. philppinarum*) and 20-fold (*M. senhousia*) between March–April 1995 and September 7, 1995, and a progressive decrease from late summer through winter. This approach may involve some limitations, such as the effect of differences between the bivalve performance in controlled laboratory experiments and that in the field and a relative approximation in adopting different excretion rates at temperatures other than those actually employed in the laboratory experiments (29). However, it indicates the strong influence of animal distribution on the magnitude and temporal scaling of biogenic nutrient regeneration due to bivalve excretion.

The great potential of this biogenic source of nutrients in cycling biophilic elements can also be highlighted by a comparison with the extent of benthic nutrient regeneration through diffusive flux. In particular, nutrient flux measured from nutrient concentrations in the porewater in adjacent intertidal and coastal areas was more than one order of magnitude lower; it varied from 0.2 to 1.5 mmol NH_4^+-N m^{-2} d^{-1} and from 0.01 to 0.05 mmol PO_4^{3-}-P m^{-2} d^{-1}. It can be inferred that a marked increase in biogenic nutrient regeneration is importantly controlled by the animal biomass increase (36) and has a major impact, acting as a positive feedback, on primary producers (41). These results indicate that abundant macrofauna and its excretory products play a primary role in benthic nutrient regeneration, are well balanced in their stoichiometric ratios, and thus act as a major factor to support primary production within the intertidal zone.

BIBLIOGRAPHY

1. Admiraal, W. and Peletier, H. (1980). Influence of seasonal variations of temperature and light on the growth rate of culture and natural population of intertidal diatoms. *Mar. Ecol. Progr. Ser.* **2**: 35–43.
2. Harrison, W.G. and Platt, T. (1980). Variations in assimilation number of coastal marine phytoplankton: effects of environmental co-variates. *J. Plank. Res.* **2**: 249–260.
3. Verhagen, J.H.G. and Nienhuis, P.H. (1983). A simulation model of production, seasonal changes in biomass and distribution of eelgrass (*Zoostera marina*) in Lake Grevelingen. *Mar. Ecol. Progr. Ser.* **10**: 187–195.
4. Giesen, W.B.J.T., van Katwijk, M.M., and den Hartog, C. (1990). Eelgrass condition and turbidity in the Dutch Wadden Sea. *Aquat. Bot.* **37**: 71–85.
5. Agustí, S. et al. (1994). Light harvesting among photosynthetic organisms. *Funct. Ecol.* **8**: 273–279.
6. Justic, D., Rabalais, N.N., and Turner, T. (1995). Stoichiometric nutrient balance and origin of coastal eutrophication. *Mar. Poll. Bull.* **30**: 41–46.
7. Valiela, I. et al. (1997). Macroalgal blooms in shallow estuaries: controls and ecophysiological and ecosystem consequences. *Limnol. Oceanogr.* **5**: 1105–1118.
8. Park, Y.C. and Carpenter, E.J. (1987). Ammonium regeneration and biomass of macrozooplankton and Ctenophores in Great South Bay, New York. *Estuaries* **4**: 316–320.
9. Wen, Y.H. and Peters, R.H. (1994). Empirical models of phosphorus and nitrogen excretion rates by zooplankton. *Limnol. Oceanogr.* **39**: 1669–1679.
10. Harrison, W.G. et al. (1992). Nitrogen dynamics at the VERTEX time-series site. *Deep-Sea Res.* **39**: 1535–1552.
11. Rowe, G.T., Clifford, C.H., Smith, K.L., and Hamilton, P.L. (1975). Benthic nutrient regeneration and its coupling to primary productivity in coastal waters. *Nature* **225**: 215–217.
12. Nixon, S.W. (1981). Remineralization and nutrient cycling in coastal marine ecosystems. In: *Nutrient Enrichment in Estuaries*. B. Nelson and L.E. Cronbin (Eds.). Humana Press, Clifton, NJ, pp. 111–138.
13. Klump, J.V. and Martens, C.S. (1983). Benthic nitrogen regeneration. In: *Nitrogen in the Marine Environment*. E.J. Carpenter and D.G. Capone (Eds.). Academic Press, New York, pp. 411–457.
14. Jordan, T.E. and Valiela, I. (1982). A nitrogen budget of the ribbed mussel, *Geukensia demissa*, and its significance in

nitrogen flow in a New England salt marsh. *Limnol. Oceanogr.* **27**: 75–90.

15. Murphy, R.C. and Kremer, J.N. (1985). Bivalve contribution to benthic metabolism in a California lagoon. *Estuaries* **8**: 330–341.
16. Prins, T.C. and Smaal, A.C. (1994). The role of the blue mussel *Mytilus edulis* in the cycling of nutrients in Oosterschelde estuary (the Netherlands). *Hydrobiol.* **282/283**: 413–429.
17. Doering, P.H., Kelly, J.R., Oviatt, C.A., and Sowers, T. (1987). Effect of the hard clam *Mercenaria mercenaria* on benthic fluxes of inorganic nutrients and gases. *Mar. Biol.* **94**: 377–383.
18. Nakamura, M., Yamamuro, M., Ishikawa, M., and Nishimura, H. (1988). Role of the bivalve *Corbicula japonica* in the nitrogen cycle in a mesohaline lagoon. *Mar. Biol.* **99**: 369–374.
19. Kiibus, M. and Kautsky, N. (1996). Respiration, nutrient excretion and filtration rate of tropical freshwater mussels and their contribution to production and energy flow in Lake Kariba, Zimbabwe. *Hydrobiol.* **331**: 25–32.
20. Yamada, H. and Kayama, M. (1987). Liberation of nitrogenous compounds from bottom sediments and effect of bioturbation by small bivalve, *Theora lata* (Hinds). *Estuar. Coast. Shelf Sci.* **24**: 539–555.
21. Gómez-Parra, A. and Forja, J.M. (1993). Benthic fluxes in Cadiz Bay (SW Spain). *Hydrobiol*. **252**: 23–34.
22. Yamamuro, M. and Koike, I. (1993). Nitrogen metabolism of the filter-feeding bivalve *Corbicula japonica* and its significance in primary production of a brackish lake in Japan. *Limnol. Oceanogr.* **35**: 997–1007.
23. Dame, R.F., Zingmark, R.G., and Haskin, E. (1984). Oyster reefs as processors of estuarine materials. *J. Exp. Mar. Biol. Ecol.* **83**: 239–247.
24. Asmus, H., Asmus, R.M., and Reise, K. (1990). Exchange processes in an intertidal mussel bed: a Sylt-flume study in the Wadden Sea. *Berichte Biolog. Anstalt Helgol.* **6**: 78.
25. Prins, T.C. and Smaal, A.C. (1990). Benthic-pelagic coupling: the release of inorganic nutrients by an intertidal bed of *Mytilus edulis*. In: *Trophic relationships in the marine environment. Proc. 24th Europ. Mar. Biol. Symp.*, Aberdeen Univ. Press, pp. 89–103.
26. Asmus, R.M., Asmus, H., Wille, A., Zubillaga, G.F., and Reise, K. (1994). Complementary oxygen and nutrient fluxes in seagrass beds and mussels banks? In: *Changes in Fluxes in Estuaries: Implications from Science to Management*. K.R. Dyer and R.J. Orth (Eds.). Olsen & Olsen, Fredensborg, pp. 227–238.
27. Bayne, B.L. and Scullard, C. (1977). Rates of nitrogen excretion by species of *Mytilus* (Bivalvia: Mollusca). *J. Mar. Biol. Ass. U.K.* **57**: 355–369.
28. Schlüter, L. and Josefsen, S.B. (1994). Annual variation in condition, respiration and remineralisation of *Mytilus edulis* L. in the Sound, Denmark. *Helgol. Meeres.* **48**: 419–430.
29. Magni, P., Montani, S., Takada, C., and Tsutsumi, H. (2000). Temporal scaling and relevance of bivalve nutrient excretion on a tidal flat of the Seto Inland Sea, Japan. *Mar. Ecol. Progr. Ser.* **198**: 139–155.
30. Heip, C.H.R. et al. (1995). Production and consumption of biological particles in temperate tidal estuaries. *Oceanogr. Mar. Biol. Ann. Rev.* **33**: 1–149.
31. Magni, P. and Montani, S. (1997). Development of benthic microalgal assemblages on a tidal flat in the Seto Inland Sea, Japan: effects of environmental variability. *La mer* **35**: 109–120.
32. Magni, P. and Montani, S. (1998). Responses of intertidal and subtidal communities of the macrobenthos to organic load and oxygen depletion in the Seto Inland Sea, Japan. *J. Rech. Océanogr.* **23**: 47–56.
33. Magni, P. and Montani, S. (2000). Physical and chemical variability in the lower intertidal zone of an estuary in the Seto Inland Sea, Japan: seasonal patterns of dissolved and particulate compounds. *Hydrobiol.* **432**: 9–23.
34. Dame, R.F., Wolaver, T.G., and Libes, S.M. (1985). The summer uptake and release of nitrogen by an intertidal oyster reef. *Neth. J. Sea Res.* **19**: 265–268.
35. Dame, R.F. et al. (1991). The influence of mussel beds on nutrients in the western Wadden Sea and Eastern Scheldt estuaries. *Estuaries* **14**: 130–138.
36. Bartoli, M. et al. (2001). Impact of *Tapes philippinarum* on nutrient dynamics and benthic respiration in the Sacca di Goro. *Hydrobiol.* **455**: 203–212.
37. Kuwae, T., Hosokawa, Y., and Eguchi, N. (1998). Dissolved inorganic nitrogen cycling in Banzu intertidal sand-flat, Japan. *Mangroves Salt Marshes* **2**: 167–175.
38. Matsukawa, Y., Sato, Y., and Sasaki, K. (1987). Benthic flux of nutrient salts on an intertidal flat. *Nippon Suisan Gakkaishi* **53**: 985–989.
39. Takayanagi, K. and Yamada, H. (1999). Effects of benthic flux on short term variations of nutrients in Aburatsubo Bay. *J. Oceanogr.* **55**: 463–469.
40. Yamamoto, T., Matsuda, O., Hashimoto, T., Imose, H., and Kitamura, T. (1998). Estimation of benthic fluxes of dissolved inorganic nitrogen and phosphorus from sediments of the Seto Inland Sea. *Umi to Kenkyu (Oceanogr. Soc. Japan)* **7**: 151–158 (in Japanese).
41. Peterson, B.J. and Heck, K.L. Jr. (2001). Positive interaction between suspension-feeding bivalves and seagrass—a facultative mutualism. *Mar. Ecol. Progr. Ser.* **213**: 143–155.

BREAKWATERS

Stefano Pagliara
Pietro Chiavaccini
Università di Pisa
Pisa, Italy

Breakwaters are coastal structures used to protect harbor and shore areas by dissipating and reflecting wave energy. They are built to

- — reduce wave disturbance in coastal and harbor areas and preserve related activities;
- — protect ships and boats from wave forces;
- — when located near shore, in the same direction as the coastline, they can stabilize the coastline, modifying cross-shore and long-shore sediment transport.

The choice of the type of structure depends on the availability of materials used, the characteristics of the incident wave, the bottom morphology, the geotechnical parameters of the soil, and the necessity of obtaining a flexible or rigid structure. Breakwaters can be classified as rubble-mound structures, vertical breakwaters, and floating breakwaters.

RUBBLE-MOUND STRUCTURES

The typical cross section of a rubble-mound breakwater is sketched in Fig. 1. It consists of different layers of stones. The center core is made up of quarry run. The external layer (armor) consists of large armor units, that can be either rock or specially designed concrete units (cubes, tetrapods, dolos). The breakwater crest is generally 1–2 m over the still water level (SWL). The crest width should be large enough to allow transport and installation of material during construction and when a repair is made (1).

Because of the relative dimensions of the units of the armor and those of the core, in some cases, it is necessary to build the breakwater as a filter of three or four layers (underlayers), so that the finer material of the core cannot be removed by the waves through the voids of the armor layer. To prevent removal of finer material, the filter must satisfy the following relations:

$$D_{15}(\textit{upper layer}) < 4 \div 5D_{85}(\textit{lower layer}) \tag{1}$$

$$D_{15}(\textit{upper layer}) < 20 \div 25D_{15}(\textit{lower layer}) \tag{2}$$

where D_{15} = nominal size that is exceeded by the 85% of the sample
D_{85} = nominal size that is exceeded by the 15% of the sample

A toe filter is necessary if the breakwater is built on erodible material. The toe filter prevents breaking waves from removing material from the base of the structure. If the breakwater is located in shallow water, the filter toe is exposed to extreme wave action. To avoid, or just limit, wave overtopping, it is possible to use a concrete structure (crown wall) located over the crest (Fig. 1).

When the breakwater is small and not high, it is possible to avoid using a center core. This kind of structure, of single sized stones, is called a "reef breakwater," and it is normally used for small submerged breakwaters (2).

Stability

The rubble-mound breakwater causes the dissipation of wave energy by generating eddies due to the breaking. The voids and the roughness of the structural material, as well as the permeability of the structure, are very important in the dissipation process. The wave energy entering through the structure creates shear stresses that can move the masses causing loss of stability.

Armor Stability. The stability formulas are based on experiments carried out on hydraulic models. One of the most used stability formulas is Hudson's (3) determined for a nonovertopping structure:

$$\frac{H}{\Delta D_{n50}^2} = (K \cot\alpha)^{1/3} \tag{3}$$

where H = characteristic height of the wave (H_s to $H_{1/10}$);
D_{n50} = equivalent cube length of median rock;
α = slope angle;
$\Delta = (\rho_s/\rho_w - 1)$ where ρ_s and ρ_w are rock density and water density, respectively;
K = stability coefficient (Tables 1 and 2).

The damage D represents the measure of the modification of the structure's profile under wave action. The damage can be defined by counting the number of rocks moved or by measuring the variation of the armor layer

Table 1. Values of the Stability Coefficient K for $H = H_s$[a]

		Damage D 0–5%	
Stone Shape	Placement	Breaking Waves	Nonbreaking Waves
Smooth, rounded	Random	2.1	2.4
Rough, angular	Random	3.5	4.0
Rough, angular	Special	4.8	5.5

[a] Slope $1.5 \le \cot\alpha \le 3.0$.

Table 2. Values of the Stability Coefficient K for $H = H_{1/10}$

		Damage D 0–5%			
		Breaking Waves		Non Breaking Waves	
Stone shape	Placement	Trunk	Head	Trunk	Head
Smooth, rounded	Random	1.2	1.1	2.4	1.9
Rough, angular	Random	2.0	1.3^a–1.9^b	4.0	2.3^a–3.2^b
Rough, angular	Special	5.8	5.3	7.0	6.4
Tetrapods	Random	7.0	3.5^a–5^b	8.0	4.0^a–6.0^b

[a] $\cot\alpha = 3$.
[b] $\cot\alpha = 1.5$.

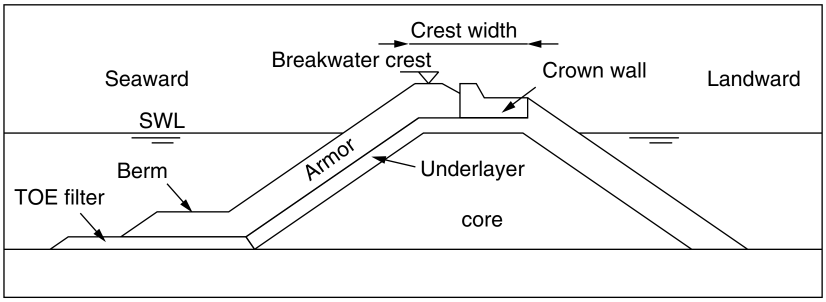

Figure 1. Typical section of a rubble-mound breakwater.

area (eroded area A of the damaged section). For this second case, Broderick (4) introduced a parameter (relative eroded area) defined as

$$S = \frac{A}{D_{n50}^2} \tag{4}$$

where D_{n50} is the nominal diameter, corresponding to 50% of the weight of the sample. The damage can be considered the number of masses of dimension equal to D_{n50} eroded in a strip of section of the same length. Zero damage means that there is nominally no removal of the armor units from the breakwater face.

The K value of Hudson's formula is different for the trunk and the head of the structure. The stones will be less stable on the head than on the trunk. In this case, K must be decreased by about 20%.

Van der Meer (5) derives expressions that include some additional parameters for an incident wave:

$$\frac{H_s}{\Delta D_{n50}} = 6.2 S^{0.2} P^{0.18} N_z^{-0.1} \xi_m^{-0.5} \quad \text{plunging waves } \xi_m < \xi_{cm} \tag{5}$$

$$\frac{H_s}{\Delta D_{n50}} = 1.0 S^{0.2} P^{-0.13} N_z^{-0.1} (\cot\alpha)^{0.5} \xi_m^{P} \quad \text{surging waves } \xi_m > \xi_{cm} \tag{6}$$

$$\xi_m = s_m^{-0.5} \tan\alpha$$

$$\xi_{mc} = [6.2 P^{0.31} (\tan\alpha)^{0.5}]^{1/(P+0.5)}$$

where S = relative eroded area (normally equal to 2);
P = notional permeability;
N_z = number of waves;
s_m = wave steepness $s_m = H_s/L_{\text{om}}$;
L_{om} = deepwater wavelength corresponding to the mean period.

For a homogeneous structure (no core, no filter, and stones of the same size), $P = 0.6$; a rock armor layer with a permeable core gives $P = 0.5$; an armor layer with filter on a permeable core gives $P = 0.4$. For a breakwater with an impermeable core, $P = 0.1$.

For overtopped and low crested structures, Van der Meer suggests multiplying D_{n50} by a reduction factor f_i defined by

$$f_i = \left(1.25 - 4.8 \frac{R_c}{H_s} \sqrt{\frac{s_{\text{op}}}{2\pi}}\right) \tag{7}$$

R_c is the freeboard, and $s_{\text{op}} = H_s/L_{\text{op}}$ (L_{op} is the deep water wavelength referred to the peak period). Expression (7) can be used in the range $0 < \left(\frac{R_c}{H_s}\sqrt{\frac{s_{\text{op}}}{2\pi}}\right) < 0.052$.

For submerged breakwaters, the following expression can be used (6):

$$\frac{h_c}{h} = (2.1 + 0.1S)\exp(0.14 N_s^*) \tag{8}$$

where h = water depth;
h'_c = height of the structure from the base;
S = relative eroded area;
$N_s^* = \frac{H_s}{\Delta D_{n50}} s_p^{-1/3}$ = spectral stability number.

Run Up and Overtopping

Run up is a phenomenon in which the incident crest wave runs up along a sloping structure to a level higher then the original wave crest. Together with overtopping, it plays, a very important role in the design of a rubble-mound structure because it depends on the characteristics of the structure (slope roughness, berm length, permeability).

Run up is expressed by $R_{u,x\%}$ that represents the level reached by the wave exceeded in x% of the cases by the incident wave. The run up level is referred to the SWL. For rubble-mound structures, Van der Meer's (6) formula is used:

$$\frac{R_{u,x\%}}{H_s} = a\xi_m, \qquad \xi_m < 1.5 \tag{9}$$

$$\frac{R_{u,x\%}}{H_s} = b\xi_m^c, \qquad \xi_m > 1.5 \tag{10}$$

These formulas are valid for breakwaters that have an impermeable or almost impermeable core ($P < 0.1$). If the breakwaters have a permeable core ($0.1 < P < 0.4$) Equations 9 and 10 become

$$\frac{R_{u,x\%}}{H_s} = d \tag{11}$$

It is usual in breakwater design to consider that $x\% = 2$; this means that $R_{u,2\%}$ is the run up exceeded by 2% of the waves. In this case, the values of the parameters of Equations (9,10), and 11 are $a = 0.96$, $b = 1.17$, $c = 0.46$, and $d = 1.97$.

ξ_m is the breaker parameter for deep water, corresponding to the mean period ($\xi_m = s_m^{-0.5}\tan\alpha$ where the symbols are explained in Equations 5 and 6).

In a low crest elevation, overtopping is allowed. Overtopping is the quantity of water passing over the crest of a structure per unit time, and it has the same dimensions of a discharge $Q(m^3/s)$, often expressed for unit length $q[m^3/(sm)]$.

A knowledge of overtopping is important in defining the necessary protection of the splash area and in assessing the risk to people or installations behind the breakwater. The amount of overtopping varies considerably from wave to wave; the overtopping discharge changes in time and space, and the greatest quantity is due to a small number of the incident waves. Wave overtopping for an impermeable rock armored slope structure with a crown wall can be expressed by the equation of Bradbury and Allsop (7), using the parameters of Aminti and Franco (8):

$$\frac{q}{gH_sT_{\text{om}}} = a\left[\left(\frac{R_c}{H_s}\right)^2 \sqrt{\frac{s_{\text{om}}}{2\pi}}\right]^{-b} \tag{12}$$

where S_{om} = deepwater wave steepness, based on mean period;
H_s = significant wave height;
R_c = crest freeboard relative to SWL in m;
T_{om} = deepwater wave mean period;
a, b = parameters as specified in Table 3.

G is the width (seaward) of the armor crest till the crown wall, and α is the slope of the armor layer.

Table 3. Coefficients for Equation 15 from Experimental Results

Armor Units	Cotα	G/H_s	a	b
Rock	2	1.10	17×10^{-8}	2.41
	2	1.85	19×10^{-8}	2.30
	2	2.60	2.3×10^{-8}	2.68
	1.33	1.10	5.0×10^{-8}	3.10
	1.33	1.85	6.8×10^{-8}	2.65
	1.33	2.60	3.1×10^{-8}	2.69
Tetrapods	2	1.10	1.9×10^{-8}	3.08
	2	1.65	1.3×10^{-8}	3.80
	2	2.6	1.1×10^{-8}	2.86
	1.33	1.10	5.6×10^{-8}	2.81
	1.33	1.85	1.7×10^{-8}	3.02
	1.33	2.60	0.92×10^{-8}	2.98

For rock armored permeable slopes, that have a theoretical permeability $P = 0.4$ and a berm in front of a crown wall, we can use the Pedersen and Burcharth formula (9):

$$\frac{qT_{\text{om}}}{L_{\text{om}}^2} = 3.2 \times 10^{-5} \left(\frac{H_s}{R_c}\right)^3 \frac{H_s^2}{A_c B \cot\alpha} \tag{13}$$

where A_c = level of the berm from SWL;
B = the width of the berm;
R_c = the level of the crest of the crown wall from SWL;
α = slope of armor.

Wave Reflection

Each coastal structure causes a wave reflection. Reflection plays a very important role because of the interaction between reflected and incident waves that can create a very confused sea, increasing the wave steepness. It is a problem especially at the entrance of an harbor because the steepness makes ship and boat maneuver very difficult. Besides, strong reflection increases the erosive force in front of the structure.

Rubble-mound breakwaters, which are permeable, rough, and sloping structures and structure of limited crest level, absorb a significant portion of the wave energy. For these structures, the reflection coefficient is small. For nonovertopped structures, that have a theoretical permeability P, we can use the following equation (10):

$$K_r = 0.071 P^{-0.082} (\cot\alpha)^{-0.62} S_{\text{op}}^{-0.46} \tag{14}$$

where K_r is the ratio of the reflected wave height and the incident wave height.

Wave Transmission

When energy passes over and through a breakwater, there is a wave transmission. The wave action in the landward side of the structure is smaller than in the seaward side. A wave is transmitted when a considerable amount of water overtops the structure and when the breakwater is very permeable and the wave period is relatively long. We define the coefficient of transmission as the ratio of transmitted to incident characteristic wave height or the ratio of the square of transmitted mean energy to incident mean wave energy:

$$K_t = \frac{H_{\text{st}}}{H_s} = \left(\frac{E_{\text{st}}}{E_s}\right)^{0.5} \tag{15}$$

For rock armored, low crested, submerged and reef breakwaters, we can use the Van der Meer and d'Angremond formula (11):

$$K_t = \left(0.031 \frac{H_s}{D_{n50}} - 0.24\right) \frac{R_c}{D_{n50}} + b \tag{16}$$

where $b = -5.42 s_{\text{op}} + 0.0323 \dfrac{H_s}{D_{n50}} - 0.0017 \left(\dfrac{B}{D_{n50}}\right)^{1.84} + 0.51$ for a conventional structure,
$b = -2.6 s_{\text{op}} - 0.05 \dfrac{H_s}{D_{n50}} + 0.85$ for a reef type structure,
H_s = significant wave height;
s_{om} = deepwater wave steepness, based on peak period;
R_c = crest freeboard relative to SWL, negative for submerged breakwaters;
B = width of crest;
D_{n50} = median of nominal diameter of rock for design conditions.

For conventional structures, K_t has a maximum of 0.75 and a minimum of 0.075, and for reef type structures, K_t varies between 0.15 and 0.6.

This formula can be used in the following range:

$$1 < \frac{H_s}{D_{n50}} < 6, \quad 0.01 < s_{\text{op}} < 0.05, \quad -2 < \frac{R_c}{D_{n50}} < 6$$

VERTICAL BREAKWATERS

Vertical breakwaters are rigid and not as flexible as rock armored ones. It is necessary to know very well the wave climate (direction, intensity, frequency, and duration of waves) to obtain a good design. A typical cross section of a vertical breakwater is sketched in Fig. 2a.

Forces on Vertical Structures

Nonbreaking Wave. A nonbreaking wave is perfectly reflected (coefficient of reflection = 1.0) and can be considered static because the period is longer than the natural period of oscillation of the structure. This stationary and static wave is called "clapotis." Assuming perfect reflection, the seaward wave has a height double the incident wave.

If the vertical breakwater is overtopped, a reduction of wave pressure occurs because it is truncated at the crest level. According to Sainflou (12), the pressures are (Fig. 2a)

$$p_1 = p_2 + \rho_w g h_s \left(\frac{H + \delta_0}{h_s + H + \delta_0}\right) \tag{17}$$

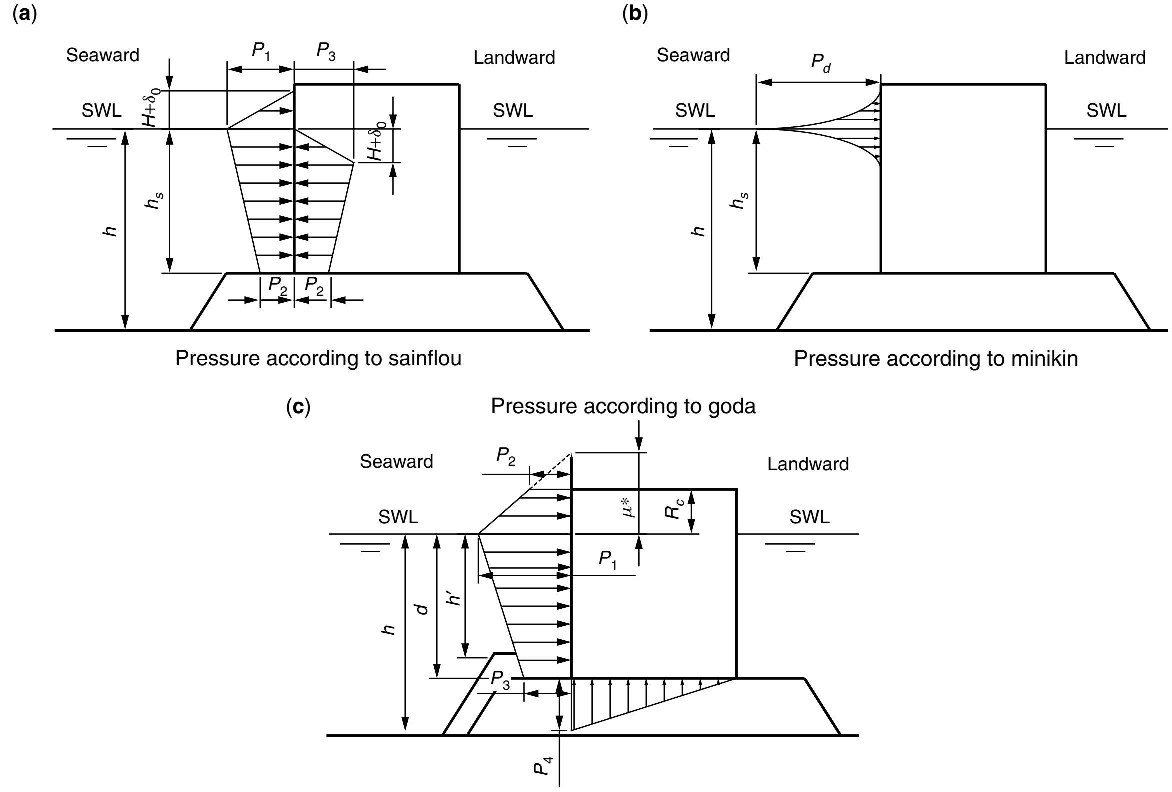

Pressure according to sainflou

Pressure according to minikin

Pressure according to goda

Figure 2. Pressure on vertical breakwaters.

$$p_2 = \frac{\rho_w g H}{\cosh(2\pi h_s/L_0)} \tag{18}$$

$$p_3 = \rho_w g(H - \delta_0) \tag{19}$$

$$\delta_0 = \frac{\pi H^2}{L_0}\coth(2\pi h_s/L_0) \tag{20}$$

where H = wave height;
p_1 = pressure at the still water level (SWL), corresponding to wave crest;
p_2 = pressure at the base of a vertical wall;
p_3 = pressure at the still water level, corresponding to a wave trough;
δ_0 = vertical shift in the wave crest and trough at the wall;
ρ_w = water density;
h_s = water depth at the foot of the structure;
L_0 = wavelength in deep water.

Breaking Waves. The worst condition for a vertical breakwater occurs when the water depth in front of the structure causes the wave to break. In this case, the dynamic component of the pressure must be considered, especially if the breaking is sudden, as happens for plunging waves. The forces generated by breaking are very high and a very short duration. According to Minikin (13), the dynamic pressure (Fig. 2b) is

$$p_d = 100\frac{h_t}{h_s L_0}(h_s + h_t)\rho_w g H \tag{21}$$

where H = wave height;
p_d = dynamic pressure at the still water level, corresponding to the wave crest;
ρ_w = water density;
h_s = water depth at the foot of the structure;
h_t = height of the caisson between the SWL and the berm;
L_0 = wavelength in deep water.

Another expression that can be used to determine the total pressure for breaking waves is (14)

$$p_1 = 0.5(1 + \cos\beta)(\alpha_1 + \alpha_2\cos^2\beta)\gamma H \tag{22}$$

$$p_2 = p_1\frac{\eta^* - h_c}{\eta^*} \tag{23}$$

$$p_3 = \alpha_3 p_1 \tag{24}$$

$$p_4 = 0.5(1 + \cos\beta)\alpha_1\alpha_3\gamma H \tag{25}$$

$$\eta^* = 0.75(1 + \cos\beta)H \tag{26}$$

$$\alpha_1 = 0.6 + 0.5\left(\frac{2kh}{\sinh 2kh}\right)^2 \tag{27}$$

$$\alpha_2 = \min\left\{\frac{h_b - d}{3h_b}\left(\frac{H}{d}\right)^2, \frac{2d}{H}\right\} \tag{28}$$

$$\alpha_3 = 1 - \frac{h}{h}\left(1 - \frac{1}{\cosh kh}\right) \tag{29}$$

$$H = \min(1.8H_s, H_f) \tag{30}$$

$$H_f = 0.18\frac{gT^2}{2\pi}\left(1-\exp\left\{\frac{3\pi^2 h_b}{gT^2}[1+15(\tan\vartheta)]^{4/3}\right\}\right) \quad (31)$$

where h_b = water depth at a distance of $5H_s$ from the vertical breakwater (seaward);
β = angle of incidence of the wave;
ϑ = bottom angle within the horizontal.

The other symbols are explained in Fig. 2c.

Ice Force. Ice exerts a force on both the landward and seaward sides. Ice force acts at the water level, and it can be as high as the crushing strength of the ice (1.5 Mpa) multiplied by the ice thickness. Even if ice is present in both sides of the structure, the worst case (ice acting only on one side) represents the design condition. Ice and waves do not occur simultaneously.

Wave Transmission

Wave transmission for a vertical breakwater is mainly the result of overtopping. Goda (15) has investigated the wave transmission of a vertical wall breakwater placed over a rock berm. He related $K_t = H_t/H_i$ to the relative freeboard $R_c/H(-1 < R_c/H < 1)$ and found this equation:

$$K_t = 0.2\left(\frac{h_v}{h_s}\right)^2 - 0.4\left(\frac{h_v}{h_s}\right) + 0.58 - 0.32\left(\frac{R_c}{H}\right) \quad (32)$$

where h_s = water depth at the foot of the structure;
h_v = height of the wall from the berm;
R_c = freeboard crest;
H = design wave height.

Wave Overtopping

The overtopping of a vertical breakwater can be determined by using Franco's formula (16):

$$\frac{q}{\sqrt{gH_s^3}} = a\exp\left(-b\frac{Rc}{\gamma_g\gamma_{b\sigma}H_s}\right) \quad (33)$$

where q = unit discharge (m^3/s per m);
a, b = coefficient depending on structure shape and on the water behavior at the seaward face; for a vertical plain wall, $a = 0.082$, $b = 3$;
$\gamma_{b\sigma}$ = depends on the direction of wave attack. For long crested waves,

$\gamma_{b\sigma} = \cos(\beta)$ for $0 \leq \beta \leq 37^\circ$,
$\gamma_{b\sigma} = 0.79$ for $37^\circ \leq \beta$.

For short crested waves,

$\gamma_{b\sigma} = 0.83$ for $0 \leq \beta \leq 20^\circ$
$\gamma_{b\sigma} = 0.83\cos(20-\beta)$ for $20^\circ \leq \beta$.

γ_g = depends on the type of the wall: for a plain wall $\gamma_g = 1$; for a plain wall with round hole perforations, $\gamma_g = 0.792$; for a plain wall with rectangular hole perforations, $\gamma_g = 0.722$; for a plain wall with rectangular hole perforations and open deck, $\gamma_g = 0.577$.

Concrete Caissons

The main component of a vertical breakwater is a cellular concrete caisson, a hollow box that is floated and *in situ*is filled with granular material to be sunk and placed on a berm prepared with sand or rock to form a stable structure (Fig. 3b).

Sometimes, the berm is a partial rubble-mound structure forming a composite breakwater. In some locations, the seaward face of the caisson is protected by sloping armor to avoid the breaking of incident waves in the front of the structure. The types of failure of a vertical caisson are shown in Fig. 3c. Normally, the loss of stability occurs from

— sliding of the caisson over the base;
— overturning of the caisson around its landward corner;
— loss of geotechnical stability of the base.

The forces acting are total wave pressure seaward (different for breaking or nonbreaking waves), hydrostatic force landward due to the still water depth, the buoyancy and the uplift force under the caisson, and the weight of the caisson. In some locations, ice force must be considered.

Sliding is verified if the total normal force multiplied by the coefficient of friction between the caisson and the berm is greater then 1.3 times the horizontal sliding force. Overturning is verified if the total stabilizing moment is greater then 1.5 times the overturning total moment. The point of overturning is the landward corner of the caisson. The last verification is to see if the resistance of the base (its critical value) is greater than the pressure due to the resultant of the acting forces. Stress is transmitted from the bottom of the caisson to the soil at an angle of 45°; this is ensured if the berm is large enough with respect to the caisson.

Sheet Piling Breakwater

Sheet piling structures are a particular type of vertical breakwater that consist of a series of piles engaged to form a continuous structure. The material most used is steel, but it possible to use reinforced concrete and timber. The choice depends on economic considerations and the availability of material. A vertical breakwater can be made of a single line of sheet piling, a double wall of sheet piles, or a cellular unit that has its the cavity filled with sand or rock. The head of a sheet piling is generally of rock, concrete slab, or paving. If steel piling is used, protection against corrosion should be provided.

The forces acting on sheet piling are the same as on caissons. Generally the worst conditions happen when the trough of the waves passes. In this case, the saturated backfill soil could cause the wall to fail seaward. The penetration of the sheet pile in to the soil has to develop the necessary resistance to cantilever action. The double wall ensures more resistance; the walls are as far apart as the depth of the water, and can be connected by cross-walls to form boxes. These boxes can be filled with sand or rocks to increase stability. A cellular sheet piling is necessary in deep water to stiffen the structure. Timber sheet pilings

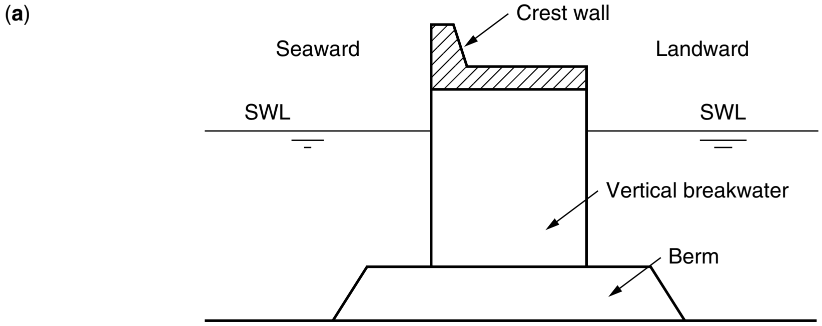

Typical vertical breakwater

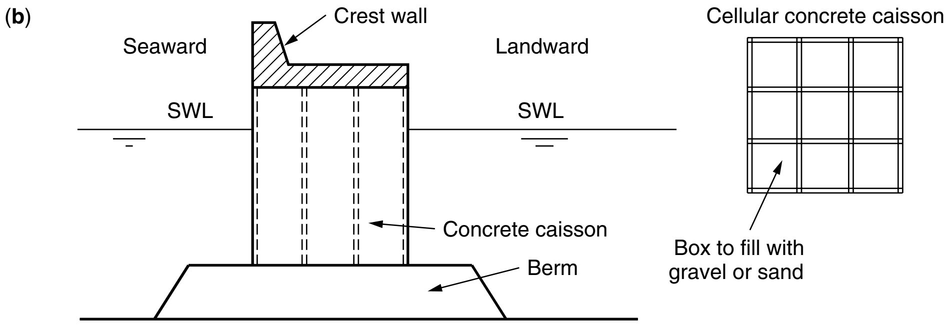

Concrete caisson vertical breakwater

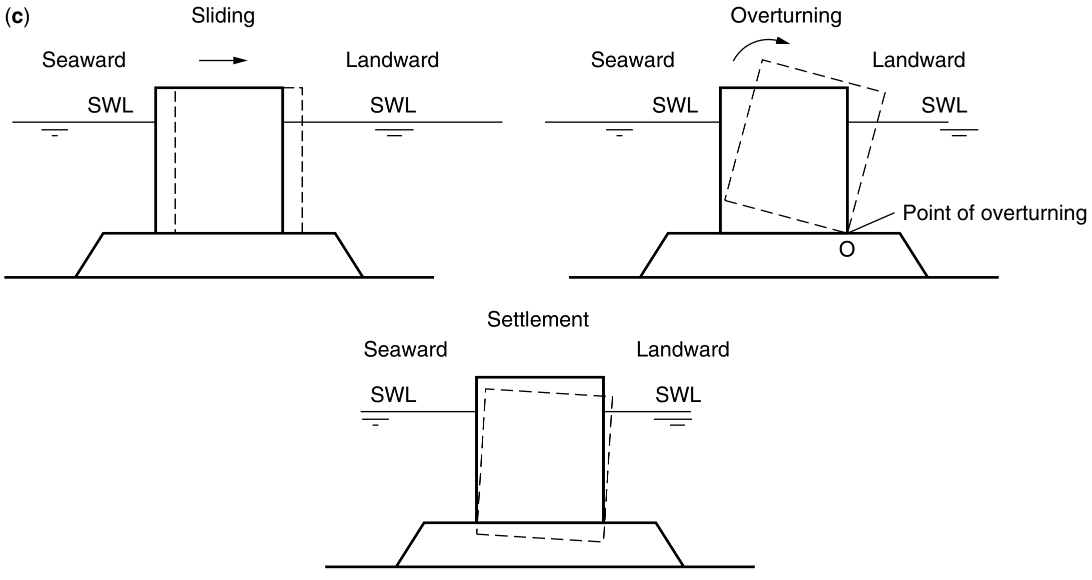

Types of failure

Figure 3. Vertical breakwater sections and types of failure.

are used in shallow water locations where wave conditions are moderate.

FLOATING BREAKWATERS

Floating breakwaters consist of composite structures that float on water (17). They can be used where waves are moderate and have a very short period. These structures are low in cost, insensitive to water depth and are easily portable, even if they provide less protection than fixed breakwaters. Generally, they consist of a box, pontoons, tethered floats, and floating docks affixed to piles. The floating breakwater reduces wave energy by reflection, dissipation interference, and conversion of the energy into a mono-oscillatory motion. The dimensions of floating breakwaters (the penetration depth, in particular) depend on the amount of energy that is necessary to dissipate and the mooring force on the anchor of the structure. The more the structure extends to the bottom the greater the reflection and the dissipation, but the mooring force also increases.

BIBLIOGRAPHY

1. CERC U.S. Corps of Eng. (1984). *Shore Protection Manual*. Vicksburg, MS.
2. U.S. Corps of Eng. (1992). EM110-2-2904 *Coastal Groins and Nearshore Breakwaters*. Washington, DC.
3. Hudson, R.Y. (1959). Laboratory investigation of rubble mound breakwaters. *J. of the Waterways and Harbors Division* **85**: 93–121.
4. Broderick, L.L. (1983). Riprap stability: A progress report. *Proceedings of costal Structures ASCE* **83**: 320–330.
5. Van der Meer, J.W. (1987). *Stability of Breakwaters Armor Layer, Co. Eng.* **11**.
6. Van der Meer, J.W. (1993). *Conceptual design of rubble mound breakwaters*, Rep. 483. Delft University, the Netherlands.
7. Bradbury, A.P. and Allsop, N.W. (1988). Hydraulic effects of breakwater crown ways. In: *Proceedings of the Breakwater '88 Conference, Institution of Civil Eng.* Thomas Telford Publishing, London, UK, pp. 385–396.
8. Aminti, P. and Franco, L. (1988). Wave overtopping on rubble mound breakwaters. *Proceedings of the 21st International Costal Engineering Conference ASCE* **1**: 770–781.
9. Pedersen, J. and Burchart, H.F. (1992). wave forces on crown walls. In: *Proceedings of the 23rd International Coastal Eng. Conference ASCE*. **2**: 1489–1502.
10. Postma, G.M. (1989). *Wave Reflection from Rock Scopes Under Random Wave Attack*. M.S Thesis Delft University of Technology, the Netherlands.
11. D'Angremond, K. and Van Roode, F.C. (2001). *Breakwaters and Closure Dams*. Delft University Press, the Netherlands, pp. 139–149.
12. Sainflou, M. (1928). *Treatise on Vertial Breakwaters*. Annal Des Ponts Et Chaussee, Paris, France.
13. Minikin, R.R. (1950). *Winds, Waves and Maritime Structures*. Griffin, London, pp. 38–39.
14. Goda, Y. (2000). *Random Seas and Design of Maritime Structure*. World Scientific, Singapore, pp. 134–139.
15. Goda, Y. (1969). *Reanalysis of Laboratory Data on Wave Transmission Over Breakwaters*. Report of the port and harbor research institute Vol. 8, no.3.
16. Franco, C. and Franco, L. (1999). Overtopping formulas for caisson breakwaters with nonbreaking 3D waves. *J. of Waterway, Port, Coastal and Ocean Eng.* **125**: 98–108.
17. U.S. Corps of Eng. (1996). EM110-2-2904 *Design of Breakwater and Jetties*. Washington, DC.

THE OCEAN IN CLIMATE

EDWARD S. SARACHIK

(from *The Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements*, Wiley 2003)

INTRODUCTION

Earth's present climate is intrinsically affected by the ocean—the climate without the ocean would be different in many essential ways: Without the evaporation of water from the sea surface, the hydrological cycle would be different; without ocean heat transport and uptake, the temperature distribution of the globe would be different; and without the biota in the ocean, the total amount of carbon in the atmosphere would be many time its current value. Yet, while we may appreciate the role of the ocean in climate, the difficulty and expense of making measurements below the ocean's surface has rendered the vast volume of the ocean a sort of *mare incognita*. Why is the ocean so important in Earth's climate, and which of its properties are of special significance for climate? How have we learned about the ocean and its role in climate, and what more do we need to know?

PROPERTIES OF THE OCEAN AND PROCESSES IN THE OCEAN

The ocean covers 70% of Earth's surface to an average depth of about 4 km. The mass of the ocean is 30 times and its heat capacity 120 times that of the atmosphere, and the ocean contains 80 times the carbon dioxide stored in the atmosphere.

The density of the ocean is controlled both by its temperature and by its salt content. Ocean density increases with salinity and decreases with temperature. Unlike fresh water, which has a maximum density at 4 °C (so that colder water and ice float on 4 °C water and the temperature at the bottom of a nonfrozen lake is 4 °C), water saltier than 26 parts per thousand of water is continuously denser as the temperature is lowered, and the temperature at the bottom of the ocean is closer to 1 °C.

Since heat always diffuses from warm to cool temperatures, why does not the temperature of the deep ocean eventually adjust and become the same as its surface temperature? Cold water constantly sinks at high latitudes (both in the Northern and Southern Hemispheres) and fills the deeper parts of the oceans with cold water so that the water at depth is always cold, even when the surface temperature is very warm. This circulation is called the thermohaline circulation.

About 7% of the ocean surface is covered by sea ice. Growth of sea ice radically changes the nature of the ocean surface: Sea ice reflects solar radiation, thereby preventing it from being absorbed by the surface and blocks the transfer of heat and moisture from the surface of the ocean to the atmosphere.

The average salinity in the global oceans is 34.7 parts per thousand salt to water by weight. As the total amount of salt in the ocean is constant, changes in salinity only occur because of additions and subtractions of fresh water. Salinity decreases as rain falls on the ocean or river water enters the ocean, and it increases as water evaporates from the surface of the ocean. As sea ice grows, it rejects salt into the ocean thereby increasing its salinity. Similarly, as ice melts, it dilutes the surrounding ocean and lowers its salinity. A specific parcel of water can either increase or decrease its salinity by mixing with parcels with different salinities.

HOW THE OCEAN INTERACTS WITH THE ATMOSPHERE TO AFFECT THE CLIMATE

The ocean interacts with the atmosphere at (or very near) the sea surface where the two media meet. Visible light can penetrate into the ocean several tens of meters, but heat, moisture, and momentum, carbon dioxide, and other gases exchange directly at the surface. Sea ice forms at the surface and helps to determine the local exchanges. The basic problem of the ocean in climate is to explain these interchanges and to determine those characteristics of the ocean that affect these exchanges.

The ocean may be considered to interact with the atmosphere in two distinct ways: passively and actively. It interacts passively when the ocean affects the atmosphere but does not change the essential manner in which the atmosphere is operating. An example of a passive interaction is the oceanic response to adding CO_2 to the atmosphere where the ocean simply delays the greenhouse warming of the atmosphere as heat and CO_2 enters the ocean.

Active interaction with the atmosphere produces results that would not otherwise be there—an example is where the warming of the atmosphere reduces the thermohaline circulation and produces a climate reaction that could not have been obtained without the essential interaction of the ocean. In particular, since the northern branch, say, of the thermohaline circulation brings cold water from high latitudes toward the equator, and since the water must be replaced by warm water that is cooled as it moves northward, the net effect of the thermohaline circulation is to transport heat northward and thereby warm the higher latitudes. As the atmosphere warms, the water becomes less dense both by the effect of temperature and by increased rainfall, a necessary concomitant of global warming. Since the atmosphere sees a reduced north–south temperature gradient at the sea surface, it reacts fundamentally differently than if the thermohaline circulation were at full strength.

Our present models of greenhouse warming have the ocean acting in both the active and passive modes—active when warming leads to a slowed thermohaline circulation and passive when heat and CO_2 simply enter the ocean surface and is therefore lost to the atmosphere. Another example of active interaction is El Niño, a phenomenon that would not exist were it not for the active interaction of the atmosphere and the ocean (see the chapter by Trenberth). The ocean also has been inferred (by examining the composition of ancient ice stored in the Greenland and Antarctic ice sheets) to have, and probably actively take part in causing, climatic variability on time scales anywhere from decades to a few thousand years, a type of variability not seen on Earth since it emerged from the last glacial maximum some 18,000 years ago.

MEASURING THE OCEAN

The ocean is remarkably poorly measured. While the global atmosphere is constantly probed and analyzed for the purposes of weather prediction, until recently no such imperative existed for the ocean. Our ability to measure the ocean is severely limited basically by the inability of radiation to penetrate very far into the ocean—this requires direct *in situ* measurements of the interior of the ocean. The world's oceanographic research fleet is small and incapable of monitoring the breadth and depth of the world's ocean, although valuable research measurements are constantly being taken at selected places. As a result of ocean observations, we know the basic pathways of water in the ocean, we have a good idea of the transports by the ocean, we have some idea of the basic mechanisms for much of the major ocean currents, and we have a good idea of how the exchanges at the surface are accomplished. Yet we cannot measure vertical velocity (it is far too small), and so we remain completely ignorant of the processes by which the interior of the ocean affects the surface. Similarly, we are ignorant of the basic processes of mixing and friction in the ocean, both basic to being able to model the ocean for climate purposes.

Major oceanographic programs have been conducted in the last decade (the World Ocean Circulation Experiment, WOCE) and (the Tropical Ocean–Global Atmosphere, TOGA), and while they have taught us much about the ocean circulation and El Niño, respectively, the basic lesson is that, unless we can make continuous long-term measurements beneath the surface of the ocean, it will forever remain unknown territory.

MODELING THE OCEAN

Because the ocean is poorly measured, and because much of what we need to know about the past and predict about the future cannot be directly known, only inferred, models have played a particularly important role in the development of oceanography and, in particular, the role of the ocean in climate.

The basic tool of climate studies is the coupled model, where the various components of the climate system—the atmosphere, ocean, cryosphere, biosphere, and chemosphere—are simultaneously and consistently coupled. The ocean component of such a coupled model exchanges its heat, fresh water, and momentum with the atmosphere at the sea surface. The test of the successful coupling of the atmosphere and ocean is the correct simulation of the time-varying sea surface temperature and surface winds, both of which are relatively easy to measure: Directly by ship or mooring, remotely by satellite, or by a combination of the two.

The development of off-line ocean-only models requires the heat, momentum, and freshwater forcing from the atmosphere to be known. Since precipitation and evaporation, in particular, are so poorly measured over the ocean, it is a continual struggle to know whether errors in the ocean model are due to errors in the model itself or errors in the forcing of the ocean by the atmosphere.

Ocean models themselves are relatively simple in concept: The known equations of water and salt are discretized and time stepped into the future. The discretization process requires a trade-off between fine resolution for accuracy and the need to simulate over long periods of time, which, because of limited computer resources, requires coarser resolution. While the equation of state of seawater relating density to salt, temperature, and pressure cannot be written down simply, it has, over the course of time, become known to high accuracy.

What makes ocean modeling difficult is the specification of those mixing processes that unavoidably cannot be resolved by whatever resolution is chosen. We are beginning to understand that enhanced small-scale mixing occurs near bottom topography and near boundaries: Purposeful release experiments, where a dye is released and then followed in time to see how the dye cloud evolves, has revealed this to us. Larger scale mixing, where parcels are interchanged because of the large-scale circulation (but still unresolved by the ocean models) itself is more problematic, but recent advances in parameterizing these unresolved mixing effects have shown promise.

THE FUTURE OF THE OCEAN IN CLIMATE

It is clear that the ocean is a crucial component of the climate system. Since so much of what is not known about the past and future of the climate system depends on active interactions with the ocean, it is clear that we have to learn more about its essential processes. How to go about learning about the ocean is the difficult question.

Direct measurements are very expensive, and satellites, while giving a global look at the entire ocean, see only its surface. Designs are currently underway for a Global Ocean Observing System (GOOS), but the cost of implementing such a system in toto is prohibitive, even if shared among the wealthier countries of the world.

It is likely that a combination of studies, perhaps conducted for entirely different purposes, will advance the field most rapidly. In particular, the advent of the El Niño–Southern Oscillation (ENSO) prediction, which requires subsurface ocean data as initial conditions, has made almost permanent the Tropical Atmosphere-Ocean (TAO) array in the tropical Pacific, giving an unprecedented and continuous view of a significant part of the tropical ocean. We may extend the reasoning to say that, where predictability is indicated and shows societal or economic value, the measurement systems to produce the initial data will almost certainly be implemented. The promise of predicting climate from seasons to a few years will expand the ocean-observing system considerably. Additional expansions will come from resource monitoring, pollution monitoring, and various types of monitoring for national security purposes. While monitoring for security has traditionally meant the data is classified, once taken, data can eventually reach the research arena—the vast amount of Soviet and U.S. data that was declassified after the end of the cold war has shown this.

Observations can be also combined with models to give "value-added" observations. Data at individual points in the ocean exist without reference to neighboring points unless models are used to dynamically interpolate the data to neighboring points using the equation of motion of a fluid. This so-called four-dimensional data assimilation is in the process of development and shows promise as a powerful way of optimally using the ocean data that can be taken.

Models can also be compared with other models. While this might seem sterile, fine-resolution models can be used to develop parameterizations of large-scale mixing for use in coarse-resolution ocean models that can be run the long times needed to participate in coupled model simulations of climate. Advances in computer power will ultimately allow successive refinements in resolution so that finer scale resolution models can be run directly.

We close by reemphasizing the crucial role that the ocean plays in climate and climate variability and the necessity to know more about the ocean for all aspects of the climate problem.

COASTAL WATERS

Susan-Marie Stedman
Silver Spring, Maryland

Coastal waters are usually defined as those waters in a zone whose landward boundary is the limit of tidal influence and whose seaward boundary is the edge of the continental shelf (or a depth of 200 m if there is no continental shelf). Salinity, temperature, turbidity, and other characteristics of coastal waters vary considerably in this transitional and very dynamic zone.

The effects of waves, wind, and other processes create two distinct layers in coastal waters: an upper surface layer known as the "mixed layer," which varies in thickness from tens to hundreds of meters thick, and a lower "deep layer." The mixed layer is warmer and less saline than the deep layer. The deep layer generally contains more nutrients due to the decomposition of organic matter that has fallen to the ocean bottom. The transition between the two layers, called the pycnocline, is usually very abrupt, occurring over a space of a meter or less (1).

The majority of coastal waters occur in three types of environments: the relatively low energy environments of estuaries and bays, the high-energy environments of waters along rocky shores and beaches, and the deeper waters of the continental shelf. All coastal waters are subject to the effects of tides, waves, and currents, to varying degrees. Coastal waters are also often affected by activities in coastal terrestrial areas, especially in parts of the world where urban and suburban development in coastal areas is intense.

PHYSICAL PROCESSES IN COASTAL WATERS

Waves

Waves in coastal waters are produced by the wind and modified by their interaction with the coastline itself. There are two general types of waves: chop, which are the short steep waves seen on windy days; and swell, which are the long gentle waves that are present even on calm days. Chop is produced by local wind conditions, whereas swell is produced by storms very distant from the coastline. Waves are often described by their period (the time between two successive crests passing the same point), length (the distance from one crest to another), height (the distance from the trough to the crest of the wave), and their speed, which is determined using the equation

$$C(\text{speed}) = \Lambda(\text{length})/T(\text{period})$$

As waves move from the deep ocean to shallow coastal waters their speed is affected by the depth of the water as defined by the equation $c = \sqrt{gH}$, where g is the acceleration due to gravity and H is the water depth. As waves approach the shore, the speed of the wave crests becomes greater than the speed of the wave troughs, creating the characteristic "breakers" that crash onto the shore.

If waves approach a shoreline at an angle they will be refracted: that is, they will bend so that they become more parallel to the shore. Wave refraction affects where wave energy is concentrated; for example, a protrusion of the shoreline will cause waves to bend toward it, concentrating wave energy in that area. Areas of wave convergence are usually subject to intense erosion, whereas areas of wave divergence tend to be depositional; thus the overall effect of waves is to straighten the coastline.

Tides

Tides are created by complex interactions between astronomical forces and the earth's geomorphology. One can begin to understand tides by imagining that the earth is covered with a layer of water with uniform depth. Gravitational and centrifugal forces associated with the moon and, to a lesser extent, the sun produce a bulge in the water at points both closest to the moon and on the opposite side of the earth. As the earth completes one rotation, a hypothetical point on the earth's surface will pass through a bulge twice, generating two high tides (and two low tides), which is known as a semidiurnal tide. Because the moon is rotating around the earth, each high tide theoretically occurs a bit later (approximately 30 min) than 12 h after the previous one. The sun's effect on tides produces the difference in tidal range seen during the lunar cycle. When the earth, moon, and sun are aligned, the tidal bulge is largest and the coast sees the highest high tides and the lowest low tides (both are known as "spring" tides). When the sun and moon are aligned at right angles with respect to each other and the earth, the tidal bulge is decreased, resulting in reduced tides (know as "neap" tides).

The actual dynamics of tides is much more complex than this simple theory alone can explain. Tides are affected not only by astronomical factors but also by the sizes, shapes, and depths of water bodies, by winds, by the earth's rotation, and by changes in barometric pressure. A wide range of tidal cycles and tidal ranges occurs around the world. Although most parts of the world have semidiurnal tides, some have only one high tide and one low tide per day (diurnal tides), and some experience a phenomenon known as "double low water," where two extreme low tides occur between high tides occurring 24 h apart (2). Tidal ranges vary not only with the lunar cycle but geographically, from a few centimeters in the Gulf of Mexico to the famous 12-m tides in the Bay of Fundy. However, at any given location along the coast, the tidal cycle is fairly constant, allowing prediction of tidal periodicity and height based on data from past years. As physical oceanographers become more adept at modeling tidal dynamics, predictions also can be made for areas without such records.

Currents

Both waves and tides, as well as the wind, generate currents in the ocean that circulate water between numerous and sometimes disparate oceanic and coastal environments. Where these currents are close to shore, they strongly influence not only the coastal waters but the shape of the coastline itself. Of most importance, however, is the way currents transport nutrients, sediments, and marine life along the coasts. Currents generally fall into three categories: tidal currents, drift currents, and a phenomenon known as upwelling.

Tidal Currents. Among the strongest forces affecting coastal waters are tides. Tidal currents are strongest where a constriction (such as a narrow tidal inlet) or other topographic feature (a valley in the continental shelf) concentrates the tidal flow. Tidal current velocities range between zero at slack tide (exact high or low tide) to midtide maximums that are generally in the range of 1–5 km/h. Under the Golden Gate Bridge in San Francisco, tidal currents sometimes reach 9 km/h, and at Seymour Narrows, British Columbia, they can reach 24 km/h (3). Tidal currents have long been considered a source of renewable energy and are being used to generate power in Canada, Russia, and France. These projects involve damming a portion of an estuary and using the force of the tide leaving the estuary to turn turbines. More recently, proposals have been made to place undersea "windmills" in areas of strong tidal currents to generate power without incurring the environmental cost of damming an estuary.

Drift Currents. Drift currents are caused by winds moving the surface layer of water. The strength of a drift current depends on wind speed, duration (how long the wind has been blowing), and fetch (how wide the body of water is). The velocity of drift currents rarely exceeds 1–2% of wind speed (1). Drift currents tend to move in the direction the wind is blowing but are also affected by the rotation of the earth, which creates a phenomenon known as the Coriolis effect. The Coriolis effect is best known for describing the direction in which a whirlpool or hurricane will rotate (clockwise in the northern hemisphere, counterclockwise in the southern hemisphere), but it also affects currents the same way. Thus, drift currents in the northern hemisphere will tend to curve to the right (and in the southern hemisphere to the left), giving rise to what is known as "Ekman drift." Finally, when a drift current encounters a coastline, it will tend to parallel the coastline in what is known as "littoral drift."

Upwelling. Upwelling refers to the vertical movement of water that occurs when the surface layer is pushed offshore

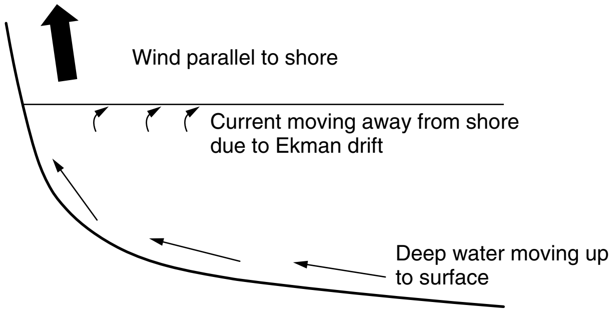

Figure 1. Upwelling along a coast.

by wind and Ekman drift, causing deeper waters to rise up and take its place (Fig. 1). While upwelling can be localized and transitory, prevailing winds create major upwelling areas along the coasts of Oregon in the United States, Peru in South America, northwest Africa, southwest Africa, and the western Indian Ocean. Where winds are steady, the upwelling is likewise fairly constant, for example, the Peruvian upwelling system; but where winds fluctuate (northwest Africa) or are only seasonal (the monsoon winds of the eastern Indian Ocean), upwelling is also periodic or seasonal. Upwelling introduces deep waters rich in nutrients into the coastal ecosystems, creating areas of high productivity and high concentration of marine life (4).

CHEMICAL CHARACTERISTICS OF COASTAL WATERS

Although fresh water is an important component of estuaries, ocean water is the primary component of most coastal waters. The primary dissolved constituents of ocean water (in greatest to lowest concentration) are calcium, magnesium, sodium, potassium, bicarbonate, sulfide, chloride, and bromide. These molecules combine in various ways to create salts, of which sodium chloride (table salt) is the most common. Salt content, or salinity, is measured in parts per thousand (ppt). The average salinity of ocean water is 35 ppt (5), but in coastal waters salinity can be reduced by the addition of fresh water from rivers, overland runoff, and/or groundwater discharge.

Ocean water also contains nutrients in the form of nitrogen and phosphate salts. The concentration of these nutrients is very low in the surface layer of the open ocean, only 0.001–0.08 mg/L N and 0.002–0.003 mg/L P, but can be 10–100 times higher in the deep layer. Oxygen, a very important constituent for biological activity, is dissolved in ocean water at concentrations from about 8 cm^3/L (for cold waters near the poles) to 4.5 cm^3/L (for warm waters at the equator). Other dissolved gases in ocean water include nitrogen, carbon dioxide, and hydrogen sulfide. The pH of ocean water varies between 7.9 and 8.3 (1).

With 44% of the world's population living within 150 km of the coast (6), human activities have a significant influence on the chemistry of coastal waters. Garbage, sewage, effluent from manufacturing plants, runoff from industrial and agricultural areas, as well as many other sources contribute nutrients, sediment, metals, pesticides, hydrocarbons, and other pollutants to coastal waters. Clumps of floating garbage, dead fish, and oiled beaches are some of the most visible effects of humans on coastal waters, but more recent concern has centered on "dead zones," areas where excess nutrients have led to overproduction of plant material, which in turn leads to oxygen depletion in a process known as "eutrophication" (7). Nutrient enrichment of coastal waters is a global problem that continues to grow as more countries increase their use of fossil fuels and inorganic fertilizers, the two principal sources of anthropogenic phosphorus and nitrogen (8).

TYPES OF COASTAL WATERS

Estuaries and Bays

An estuary is "a semi-enclosed coastal body of water which has a free connection with the open sea and within which sea water is measurably diluted with fresh water derived from land drainage" (9). Although the "land drainage" is not specified in Pritchard's classic definition, it is generally assumed that the freshwater source is a river, as is mentioned in many other definitions of an estuary (10). The tidal portion of streams that drain into estuaries is also considered to be part of the estuary. Bays are also semienclosed bodies of water, normally open to the ocean, but they differ from estuaries in not being associated with a river. Estuaries are much more common along the world's coasts than bays, although the term "bay" is often used in the name of an estuary, for example, Chesapeake Bay. Both estuaries and bays provide protected aquatic environments for marine life, but estuaries are more diverse and often more productive due to the range of habitats created by the salinity gradient and input of nutrients from streams. Dyer (11) described four physiographic types of estuaries: coastal plain estuaries (such as the drowned river valleys along the southeast coast of the United States), fjords (such as the submerged glacial valleys of Canada and New Zealand), lagoons (estuaries created by spits or bars), and tectonic estuaries (such as San Francisco Bay).

Rivers bring fresh water into an estuary, and tides bring saline water into the estuary. Where mixing is minimal, a layered water body is created with the less dense fresh water on top and the more dense saline water on the bottom (Fig. 2).

Turbulence in the water causes mixing of the two layers, which increases the thickness of the fresher upper layer as it moves from the landward part of the estuary toward

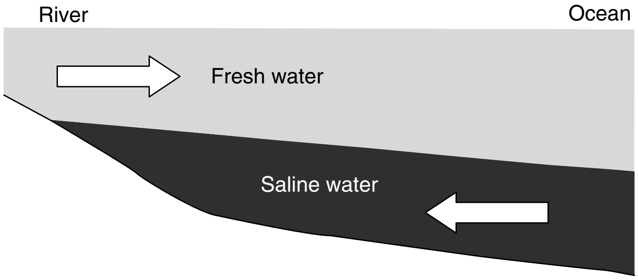

Figure 2. Layers of fresh water and saline water in a stratified estuary.

the ocean (4). The degree of mixing between the layers depends on the amount of freshwater input, the strength of tidal flow, and the effects of wind. Estuaries are often described as being one of three types: salt-wedge, partially mixed, or vertically homogeneous. When an estuary has a large freshwater input relative to tidal flow, the saline water tends to remain at the bottom with minimal mixing and the tip of this "salt wedge" changes location with the seasonal variation of freshwater flow. In partially mixed estuaries, the effects of tides and/or wind create a transition zone between the saline and fresh water layers. The mixing of the layers results in resuspension of sediments and nutrients brought into the estuary by the river and tide. In a vertically homogeneous estuary, there are no longer two water layers, but rather a single mixed body of water with a salinity gradient from low at the river mouth to high at the estuary's outlet to the ocean. In fjords, a sill at the mouth of the estuary traps saline water behind it, and mixing of fresh and saline water only occurs at the surface in water lying above the sill. The deeper waters behind the sill are replaced only when storm events disrupt this balance.

River discharge and tidal exchange both bring sediment and nutrients into an estuary. Estuaries tend to act as sediment traps due to two changes that occur when river discharge meets seawater. First, when clay particles in fresh water encounter saline water, the particles tend to flocculate (creating what is called a "turbidity maximum") and ultimately sink to the bottom of the estuary. Second, as rivers discharge fresh water at the surface of the estuary and tides bring saline water into the bottom of the estuary, current velocities decrease, causing sand particles to settle out. Complex hydrodynamics determine how much sediment stays in the estuary and how much is transported through the estuary to the ocean. Ultimately, sediments in estuaries undergo many cycles of transport by tides, currents, and waves before they find a permanent resting place (12).

Nutrients can be brought into estuaries in particular form, as dissolved constituents, or even via the atmosphere. Phosphorus, one important nutrient, has natural sources in soils and rocks. Weathering creates small particles with phosphorus bound to them that are carried into estuaries. Nitrogen, another important nutrient, is also found naturally in soils, but it is much more soluble than phosphorus and so tends to enter an estuary predominantly in dissolved form. Natural sources of nitrogen and phosphorus are generally small when compared to human-related sources such as agriculture (inorganic fertilizers), wastewater, and fossil fuel. Atmospheric deposition of nitrogen derived from fossil fuel combustion or animal waste is thought to be an important source of excess nutrients to estuaries in the United States (8).

Rocky Shores and Beaches

Rocky shores and beaches are high energy environments. Waves and currents keep the water almost constantly in motion and well aerated. The shoreface is a transition zone where waves become influenced by the bottom and change shape, increase in height, and realign to become more parallel to the shoreline. In the surf zone, currents are driven by the interaction of waves with the shoreface and with other waves. In some places the water draining off the shoreface creates very strong seaward-moving currents known as "rip currents." Offshore from the surf zone, currents are dominated by wind and tide energy but still tend to parallel the coastline.

Plants and animals must have special adaptations to thrive in these environments. Kelp and seaweed have strong attachment systems to keep them in place, and the dominant invertebrates have protective shells. Furthermore, the vigorous movement of water along rocky shores and beaches is essential to the survival of these organisms because they feed by filtering the water or snatching debris as it floats by. Sandy areas are the most dynamic because not only is the water constantly moving, but the substrate itself tends to move as well. Although many beaches seem devoid of plant or animal life, that appearance occurs only because most beach inhabitants are buried in the sand.

Shelf Waters

Seaward of estuaries and the coastline are the waters overlying the continental shelf. Tides continue to influence these waters, creating turbulence that results in a mixed water layer that can extend for tens of kilometers outward from the coast (2). Beyond this mixed layer, coastal waters become stratified into the surface and deep layers described earlier. Areas of high productivity tend to occur at the transition between mixed and stratified waters.

Coastal shelf waters are an important component of ocean circulation. Prevailing winds and the earth's rotation create a system of major currents that circle the ocean basins and in some cases move water from one basin to another. The strongest currents occur on the western sides of ocean basins, where they carry water toward the poles, and weaker currents occur on the eastern sides of ocean basins, where they carry water toward the equator. These major currents, which transport volumes ranging from 10 to 100 m^3/s (2), serve many important biological functions related to migration, dispersal of juvenile organisms, and transportation of nutrients. One of the best-known oceanic currents is the Gulf Stream, which occurs off the east coast of North America. Coastal waters also contain coastal boundary currents that trend parallel to the coastline, their overall direction determined by the shape of the coastline and the way winds tend to reorient perpendicular to the coast as they approach the coastline. For example, landward of the Gulf Stream in the western north Atlantic Ocean, a coastal boundary current flows south—opposite to the direction of Gulf Stream flow. Coastal boundary currents transport lesser volumes of water than oceanic currents—about 10 m^3/s on the average (2).

IMPORTANCE OF COASTAL WATERS

Coastal waters are extremely important for fisheries. With the exception of a few deep-water migratory species such as tuna, most other commercial and recreational marine fish (and shellfish) feed, reproduce, and are harvested in coastal waters. In the United States, coastal waters support a commercial fishing industry that in 2002 landed 4.3 million metric tons of fish and shellfish valued at $3.1 billion. The commercial fishing industry as a whole contributed $28.4 billion to the U.S. gross national product (GNP) in 2002 (13). Recreational fishing also produces billions of dollars of economic benefits per year (14). Coastal waters are also a very important habitat for marine mammals such as seals, manatees, and sea lions, some of which are listed as threatened or endangered under the U.S. Endangered Species Act.

Coastal waters are also very important to interstate and international commerce. Transport by water is generally the most economical and efficient means to move goods. U.S. ports and waterways transport more than 2 billion tons of domestic and foreign goods each year. The marine transportation industry contributes $742 billion to the U.S. GNP and provides approximately 13 million jobs (15).

Given the importance of coastal waters to fisheries, marine transportation, and numerous other uses that are significant from either an economic or ecological perspective, conflicts over coastal waters are common. Laws and regulations governing environmental protection and management of natural resources in the United States and other countries often require a cost–benefit analysis of proposed activities in coastal waters, or a determination that a proposal is in the best interests of the public. However, uncertainties about quantifying ecological values and predicting the response of coastal systems to change make it difficult to conduct these kinds of evaluations. The future health of coastal waters depends on a better understanding of their processes and their response to human activities.

BIBLIOGRAPHY

1. Barnabé, G. and Barnabé-Quet, R. (2000). *Ecology and Management of Coastal Waters: The Aquatic Environment*. Springer-Verlag, New York.
2. Beer, T. (1997). *Environmental Oceanography*. 2nd Edn. CRC Press, Boca Raton, FL.
3. National Imagery and Mapping Agency (NIMA). (1995). Tides and tidal currents. In: *The American Practical Navigator, Publ. No. 9*. NIMA, Bethesda, MD, pp. 143–164.
4. Mann, K.H. (2000). *Ecology of Coastal Waters with Implications for Management*. 2nd Edn. Blackwell Science, Malden, MA.
5. Dittmar, W. (1884). Report on the researches into the composition of ocean water collected by *H.M.S. Challenger* during the years 1873–1876. In: *Report on the Scientific results of the Voyage of the H.M.S. Challenger: Physics and Chemistry*. J. Murray (Ed.). Vol. 1. HMSO, London.
6. Cohen, J.E. et al. (1998). Estimates of coastal populations. *Science* **278**(5341): 1211–1212.
7. Nixon, S.W. (1995). Coastal marine eutrophication: a definition, social causes, and future concerns. *Ophelia* **41**: 199–219.
8. National Research Council (NRC). (2000). *Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution*. National Academy Press, Washington, DC.
9. Pritchard, D.W. (1967). What is an estuary: physical viewpoint. In: *Estuaries*. G.H. Lauf (Ed.). *AAAS* Publ. No. 83, Washington, DC, pp. 3–5.
10. Perillo, G.M.E. (1995). Geomorphology and sedimentology of estuaries: definitions and geomorphologic classifications of estuaries. *Dev. Sedimentol.* **53**.
11. Dyer, K. (1973). *Estuaries: A Physical Introduction*. Wiley-Interscience, Hoboken, NJ.
12. Nichols, M.M. and Biggs, R.B. (1985). Estuaries. In: *Coastal Sedimentary Environments*. 2nd Edn. R.A. Davis (Ed.). Springer Verlag, New York, pp. 77–186.
13. U.S. Department of Commerce (DOC). (2003). *Fisheries of the United States, 2002*. DOC, Silver Spring, MD.
14. American Sportfishing Association (ASA). (2002). *Sportfishing in America: Values of Our Traditional Pastime*. Alexandria, VA.
15. U.S. Department of Transportation (DOT). (1999). *An Assessment of the U.S. Marine Transportation System—A Report to Congress*. DOT, Washington, DC.

MARINE COLLOIDS

Peter H. Santschi
Texas A&M University
Galveston, Texas

Aquatic colloids, as nanoparticles and macromolecules in the 1 nm to 1 μm size range (1), differ fundamentally from solutions. In a solution, all the relevant interactions are comparable to the thermal energy, k*T*. Thus, even if water 'prefers' water, the tendency toward disorder—entropy—is dominant. In contrast, the colloidal state is unstable because of large interfacial energies and because particle–particle interactions are stronger than k*T*. In aquatic systems, physical techniques such as diffusion, permeation, light scattering, and viscosity can be used to characterize colloidal systems, chemical techniques can help to unravel their compositional complexities, and biological techniques can give insights into degradability, bioavailability, and stability. New insights into colloids in marine systems require application of state-of-the-art techniques and also paradigm shifts.

Most of the organic matter in the ocean is in the dissolved state, and 20–30% of the marine DOC is colloidal and subject to rapid turnover (2). Marine colloidal matter (COM) contains biologically resistant heteropolycondensations, such as humic-type molecules (geopolymers) that act as soaps, and products of carbon fixation and consumption, such as exopolymeric substances (EPS) that self-assemble (3,4) and thus act as glues (5,6, and references therein). Both geopolymers and biopolymers also contain strong metal-binding ligand groups, rendering them as metal-sequestering agents. Therefore, colloids in the ocean can act as vectors of metal removal and of solubilization, depending on the biomolecule, as well as modifiers of metal bioavailability. In addition to questions of metal bioavailability in the delicately balanced marine ecosystem, EPS in the marine colloidal pool are also important in their response to global change, as they provide some of the sinking material, that is, one of the conveyor belts of global carbon cycling. Last but not least, man-made nanoparticles, such as buckeyballs, carbon fibers, and nanotubes, which are potentially toxic to man, will make their way into the environment, and ultimately into the ocean, where they have unknown behavior and fate.

Colloids in seawater can be small and globular or long and fibrillar, as shown in Figure 1. Aquagenic fibrillar gel-forming microbial exudates are classified as “rigid”, other colloids, such as terrestrially derived (pedogenic) nanometer-sized humic acids, are more globular and are called “flexible” polymers (8–11). Both types of colloids are stabilized by metal ions such as Ca^{2+}, which holds individual humic acid molecules together as

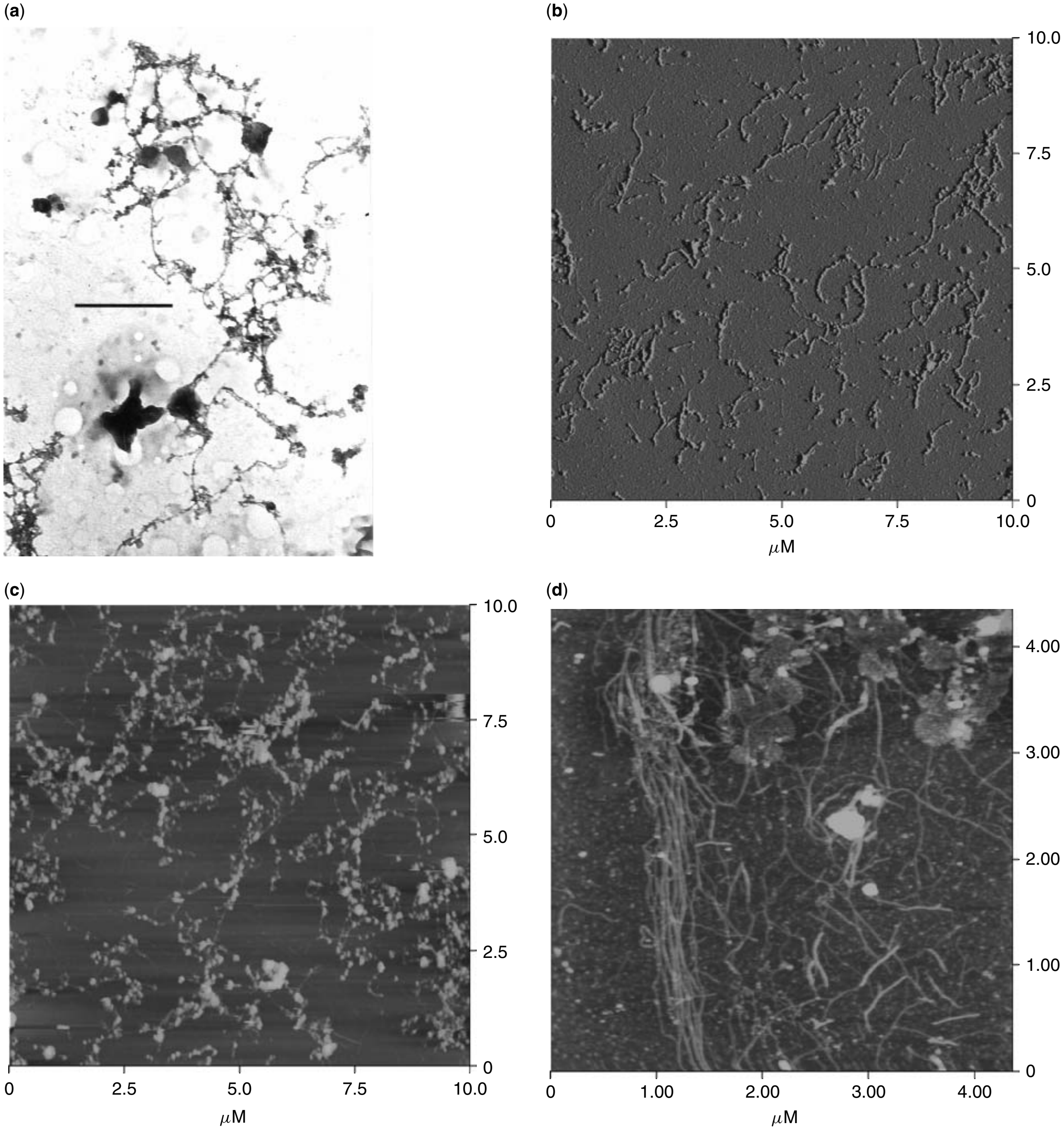

Figure 1. Fibrillar network of colloidal particles and fibrils (7) arranged in a pearl-on-necklace fashion. (**a**). Transmission electron microscopy (TEM) of stained fibrillar network, embedded in a hydrophilic resin, which is part of colloidal organic matter, COM (ultrafilter size range of 2–400 nm), sampled from Middle Atlantic Bight surface waters (courtesy of K. Wilkinson); (**b**). Atomic force microscopy (AFM) of sample from Middle Atlantic Bight surface water COM; Fibrils were also identified by AFM, albeit at lower abundance, in nepheloid layers at 2500 m but not in midwaters; (**c**). AFM of COM from Galveston Bay waters, TX. (**d**). AFM image of COM from surface waters of the Gulf of Mexico, which shows the potential for gel formation of fibrillar macromolecules at the edge of the mica support plate, where the gel-like aggregates likely formed as a consequence of drying.

aggregates (12) and gives rigidity to alpha-helical regions within exopolymeric fibrils (13). These exopolymeric fibrils are up to several 1000 nm in length, but only 1–3 nm wide, and can exist abundantly attached to cell surfaces or as free colloids (7,14–16). The degree of rigidity of fibrillar biopolymers can be studied using atomic force microscopy (AFM) or SANS (small angle neutron scattering). These fibrils can also be characterized by persistence length (defines a kind of stiff-chain-segment length for a linear polymer chain), end-to-end distance, and contour length (11). These measures allow predicting their physicochemical behavior in solution.

Marine colloids are commonly sampled for characterization using cross-flow ultrafiltration (17,18, and references therein) or field-flow fractionation techniques (19). Recent applications of ultrafiltration methods to the sampling of aquatic colloids have advanced our understanding of the abundance, distribution, physical conformation, chemical composition, and environmental behavior of colloids in aquatic systems, including freshwater, seawater, and groundwater. Applications of ultrafiltration to marine systems date back to the 1970s and 1980s, when mostly stirred cell ultrafiltration units with membrane disks were used. Applications of large volume cross-flow ultrafiltration systems to sampling marine colloids started in the early 1990s, and they are now commonly used in marine science (20,21).

In estuarine waters, many metals such as Cu, Ag, Fe, Hg, and Zn, have been found largely in colloidal form (5,18,22,23, and references therein). In marine environments, studies indicated that high molecular weight (HMW) COM has relatively high biological and chemical reactivity and short residence times in the water column (2,24). Therefore, colloids are also important intermediaries in organic carbon cycling in aquatic environments. Recent advances in sampling techniques now enable the isolation of large quantities (e.g., 100s of mg) of environmental colloids from large volumes of natural waters for elemental, isotopic, and molecular characterization (2,7,20,25–27), although there is still a debate about optimal sampling procedures (21,28, and references therein).

Due to their amphiphilic (or amphipathic) nature, aquatic colloids are not only charged, thus providing potentially strong chelating binding sites for trace metals and radionuclides, but also contain hydrophobic moieties for binding more hydrophobic organic trace compounds (5,29,30, and references therein). There are many potential metal binding sites in aquatic biomacromolecules, but most of them are occupied by major seawater cations, and much fewer with the much rarer trace metal ions. Binding sites include carboxyl, hydroxyl, phosphate, sulfate, amino, and sulfhydryl groups; A (or hard) metals prefer O over N and S binding sites, and B (or soft) metals prefer S (or N) over O binding sites (31). One would also expect to find that metal enrichments follow the Irving–Williams series, which has been documented for marine colloids by Guo et al. (32), and for bacterial EPS by Rudd et al. (33). Many of these sites are not exceptionally strong when acting alone, but the particular steric environment of the macromolecule can provide strongly chelating binding sites for trace metals when acting in concert (34,35). In addition, in the "egg-box" chelating environment (13,36), where Ca^{2+} stabilizes the alpha helix of acid polysaccharides (APS) such as alginic acid, Ca^{2+} can be replaced by other cations with similar radii.

EPS are usually rich in APS, and because much more is known about biopolymers such as APS that are studied in food science or biomedical science, where they are used as emulsifiers, stabilizers, coagulants (or anticoagulants) through their amphiphilic properties (37), the concepts and approaches in these scientific disciplines are very useful here. APS include alginates, carrageenans, xanthans, and hyaluronic acids, which contain building blocks that are also found in EPS and transparent exopolymeric particles (TEP) (14,38,39). One may therefore apply knowledge from these scientific disciplines to the biomolecular complexants in the marine colloidal pool, which are not as well studied. The physicochemical properties of EPS molecules are important because they can provide "hideouts" for micronutrients such as trace metals. EPS molecules contain APS and also proteins or lipids, which makes these molecules amphiphilic (1). The amphipilicity is caused by the generally hydrophilic properties of their acid polysaccharidic regions and the more hydrophobic properties of their protein or lipid-rich regions (37). Many of these EPS biopolymers also form gels (3,6) and thus, have properties of hydrocolloids.

Colloidal ligands that have surface-active properties and are found in the filter-passing fraction can be the cause of particle concentration effects on particle–water partition coefficients (K_d, the concentration ratio in particles and in solution) and kinetic constants (k_i) of trace metals or radionuclides, which have been documented in the literature for many years. Colloidal ligands can often be removed by coagulation (40,41) and flocculation (5). Such particle-concentration effects disappear, however, when one corrects for the presence of colloidal ligands (29). Values of the partition coefficients of the A-metal ^{234}Th(IV) to colloids (K_c) are generally similar to K_d values for particles (42), but this is not necessarily the case for B-metals such as Cu (43,44) where generally K_c is larger than K_d.

Aggregation rates of colloidally complexed metals show two distinctive reaction rates; a larger rate constant is consistent with Brownian "colloidal pumping" that occurs at a uniform rate and is faster at higher particle concentrations and a smaller rate constant that is different for different metals (22,45).

^{234}Th, which is generated *in situ* from the radioactive decay of ^{238}U and has a half-life of 24 days, can be used to derive residence times of Th(IV)-binding colloids. Average colloidal residence times (with respect to coagulation) range from fractions of days in estuaries to a few weeks in the surface ocean (2,20). Similar results are obtained in controlled laboratory coagulation experiments using radioactive metals bound to estuarine COM (23,45). These laboratory results also confirmed particle concentration effects on removal rate constants, previously demonstrated for Th(IV) (34,40,41).

Residence times that reflect colloidal organic carbon ages are longer, as much of the carbon is recycled. For example, Santschi et al. (2,7) and Guo et al. (42,46) obtained mostly "modern" (i.e., younger than 1950) apparent radiocarbon ages for high molecular weight colloids $\geq$10 kDa, but mostly older ages, for example, 400 to 4500 years, for low molecular weight COM $\geq$1 kDa in surface waters containing mostly biopolymers (2). In contrast, if the COM $\geq$1 kDa is enriched to 100% polysaccharides through repeated alcohol precipitation, modern radiocarbon ages are obtained for this material (7). In contrast to colloids from the euphotic zone, colloids from benthic nepheloid layers of estuaries and continental shelf or slope areas of the Gulf of Mexico and the Middle Atlantic Bight, which contain mostly geopolymers generated by hydrodynamic and physicochemical fractionation processes near the sediment–water interface during sediment resuspension (46,47), this age distribution is reversed. For example, apparent radiocarbon ages are up to 10 ky for $\geq$10 kDa colloids, and younger for $\geq$1 kDa colloids, indicating fundamental differences between biopolymers and geopolymers.

It thus appears that whenever colloids are generated by microorganisms in surface and bottom waters, their importance in trace metal binding and removal might be disproportionately higher than would be expected from the observed colloidal fractions of trace metals.

TEP-like biopolymers control the scavenging of metals and radionuclides and the coagulation/flocculation of particles in marine systems (5, and references therein) and also the early development of biofilms (15), which is largely due to their surface-active nature. These biopolymers can also initiate or modify precipitation of MnO_2 and FeOOH (48,49), SiO_2 (50), and $CaCO_3$ (51). Moreover, APS rich polymers in the extracellular milieu form flocs (52,53) and also bind extracellular enzymes in their active forms. This enzymatic activity allows trace metal binding (24,34) and modifies the solubility of associated molecules (14).

What do we know about the physicochemical parameters of colloidal organic matter in the ocean, for example, proton reactive sites, their pK_a values, electrophoretic mobilities, and relative hydrophobicities? Santschi et al. (2) reported a total concentration of 1.4 meq/g of proton reactive sites in marine colloids from the Gulf of Mexico, which is intermediate between that of humic and fulvic acids (10–16 meq/g) and polysaccharides (0.3–0.8 meq/g) (1). These proton-reactive sites are often assumed to have pK_a values of 2, 4, 6, 8, and 10 (54). Actual values of $pK_{a1} \leq 3$ have been determined for aquatic colloids (7,34,54–56), which could suggest sulfate or polyphosphate as possible ligands. Values for pK_{a2} are more variable and increase from about 4 to 9 across the salinity gradient in the Firth of Clyde estuary (55,56), suggesting that the nature of colloids changes dramatically across an estuarine salinity gradient. Interestingly, log K values, as well as fractions of metal-specific ligands (C_L) in the colloidal fraction (colloidal C_L/total C_L) of selected metals in coastal waters were significantly related to the pK_{a2} of the colloids; Cu showed a positive relationship for both log K and % colloidal C_L with pK_{a2}, whereas for Pb, a negative relationship was found (55). This is possibly due to the different affinities of these two metals to the two predominant forms of freshwater colloids (19), that is, the smaller organic carbon rich (enriched in Cu) and larger Fe-rich (enriched in Pb) colloids, both of which are also found in coastal waters (23).

In agreement with the pKa values, measured electrophoretic mobilities of aquatic colloids are negative. Electrophoretic mobilities, u_E, determined by dynamic light scattering in a ZetaSizer instrument at pH 8, were measured at about -2.5×10^{-8} m^2 V^{-1} s^{-1}, ranging from about $(-2$ to $-3.4) \times 10^{-8}$ m^2 V^{-1} s^{-1} (55–57). Their values were weakly related to the colloidal fraction of metal-specific ligands (colloidal C_L/total C_L); Cu and Pb again related in opposite ways.

In accordance with the relative high concentration of moderately strong acidic sites, for example, pK_{a1} values of $\sim$3 (55,56) or below for marine colloids (9), isoelectric points of about 2 to 4 for marine colloids have been determined by Quigley et al. (34) and Alvarado-Quiroz (35), suggesting that the surface charge is dominated by relatively strong acidic groups. Conversely, positively charged ligand groups such as amino groups are considerably rarer.

Marine colloids are negatively charged, suggesting hydrocolloid-like behavior, but that does not preclude coexistence with more hydrophobic sites, as has been said before. EPS and APS have been described as amphiphilic (1,37,58). Furthermore, it is likely that more hydrophobic ligands would be more particle-reactive and thus, have a shorter residence time in water. However, this has not yet been demonstrated in the field. Laboratory studies of bacterial attachment, however, show a clear relationship between attachment probability and relative hydrophobicity of bacterial surfaces (58,59), which are composed of EPS.

There are many things we do not know yet about marine colloids. For example, more than 50% of the organic matter in this pool remains uncharacterized at the molecular level (26,27,60), our knowledge about their tendency toward gel formation is still in its infancy (6), and we still do not have a quantitative assessment of the relative role of biodegradation (61) versus self-assembly (3) or aggregation contributing to the observed size distribution of marine colloids and particles (2,6,24). Even the molecular weight distribution of DOC is not as certain as one would conclude from the application of just one single method. For example, results from the application of state-of-the-art electrospray mass spectrometry techniques (62,63) indicate average molecular weights of DOC, normalized to charge, of less than 1 kDa, possibly due to fragmentation of carboxylic acid-rich compounds during electrospray ionization. On the other hand, much higher average molecular weights of 10^5 to 10^6 Da can be obtained from images of freshwater (9) or marine colloids (7) when using transmission electron microscopy (TEM) of stained specimens consisting of abundant fibrillar macromolecules 100s to 1000s of nm long, when embedded in hydrophilic resins, or AFM techniques of the same specimens mounted by adsorption or centrifugation onto mica surfaces that are flat on the atomic scale (Fig. 1). Thus, scientists interested in the properties of marine colloidal matter still have a long way to go until the role of this abundant material in seawater and self-assembles on the nanoscale (3) become better understood.

Acknowledgments

This review was supported, in parts, also by grants to the author from the U.S. National Science Foundation (NIRT-0210865 and OCE-0351559) and the Texas Institute of Oceanography.

BIBLIOGRAPHY

1. Buffle, J. (1990). *Complexation Reactions in Aquatic Systems: An Analytical Approach*. Ellis Horwood, New York.
2. Santschi, P.H. et al. (1995). Isotopic evidence for contemporary origin of high molecular weight organic matter in oceanic environments. *Geochim. Cosmochim. Acta* **59**: 625–631.
3. Chin, W.C., Orellana, M.V., and Verdugo, P. (1998). Spontaneous assembly of marine dissolved organic matter into polymer gels. *Nature* **391**: 568–572.
4. Kerner, M. et al. (2003). Self-organization of dissolved organic matter to micelle-like microparticles in river water. *Nature* **422**: 150–154.
5. Santschi, P.H., Burd, A.B., Gaillard, J-F., and Lazarides, A.A. (2004). Transport of materials and chemicals by nano-scale colloids and micro to macro-scale flocs in marine, freshwater and engineered systems. In: *Flocculation in Natural and Engineered Environmental Systems*. I.G. Droppo, G.G. Leppard, S.N. Liss, and T.G. Milligan (Eds.). CRC Press, London, UK in press.
6. Verdugo, P. et al. (2004). The oceanic gel phase: A bridge in the DOM-POM continuum. *Mar. Chem.*, in press.
7. Santschi, P.H. et al. (1998). Fibrillar polysaccharides in marine macromolecular organic matter as imaged by atomic force microscopy and transmission electron microscopy. *Limnol. Oceanogr.* **43**: 896–908.
8. Buffle et al. (1998).
9. Wilkinson, K.J., Balnois, E., Leppard, G.G., and Buffle, J. (1999). Characteristic features of the major components of freshwater colloidal organic matter revealed by transmission electron and atomic force microscopy. Colloids *Surf. A: Physicochemical Eng. Aspects* **155**: 287–310.
10. Balnois, E., Wilkinson, K.J., Lead, J.R., and Buffle, J. (1999). Atomic force microscopy of humic substances: Effects of pH and ionic strength. *Environ. Sci. Technol.* **33**: 3911–3917.
11. Balnois, E. et al. (2000). Conformations of succinoglycan as observed by atomic force microscopy. *Macromolecules* **33**: 7440–7447.
12. Simpson et al. (2000).
13. Moe, S., Draget, K., SkjÅk-Bræk, G., and Smidrød, O. (1995). Alginates. In: *Food Polysaccharides and their Application*. A. Stephen (Ed.). Marcel Dekker, New York, pp. 245–286.
14. Leppard, G. (1995). The characterization of algal and microbial mucilages and their aggregates in aquatic ecosystems. *Sci. Total Environ.* **165**: 103–131.
15. Leppard, G. (1997). Colloidal organic fibrils of acid polysaccharides in surface water: Electron-optical characteristics, activities and chemical estimates of abundance. *Colloids Surf. A: Physicochemical Eng. Aspects* **120**: 1–15.
16. Leppard, G.G. et al. (1997). A classification scheme for marine organic colloids in the Adriatic Sea: Colloid speciation by transmission electron microscopy. *Can. J. Fish. Aquat. Sci.* **54**: 2334–2349.
17. Guo, L., Wen, L., Tang, D., and Santschi, P.H. (2000). Re-examination of cross flow ultrafiltration for sampling aquatic colloids: evidence from molecular probes. *Mar. Chem.*
18. Guo et al. (2004).
19. Lyven, B. et al. (2003). Competition between iron- and carbon-based colloidal carriers for trace metals in a freshwater assessed using flow field- flow fractionation coupled to ICPMS, *Geochim. Cosmochim. Acta* **67**(20): 3791–3802.
20. Guo, L. and Santschi, P.H. (1997). Composition and cycling of colloids in marine environments. *Rev. Geophys.* **35**: 17–40.
21. Wells, M.L. (2002). Marine colloids and trace metals. In: *Biogeochemistry of Marine Dissolved Organic Matter*. D.A. Hansell and C.A. Carlson (Eds.). Academic Press, New York, pp. 367–404.
22. Wen, L-S., Santschi, P.H., Gill, G., and Paternostro, C. (1997). Colloidal and particulate silver in river and estuarine waters of Texas. *Environ. Sci. Technol.* **31**: 723–731.
23. Wen, L.S., Santschi, P.H., Paternostro, C., and Gill, G. (1999). Estuarine trace metal distributions in Galveston Bay I: Importance of colloidal forms in the speciation of the dissolved phase. *Mar. Chem.*, **63** (3–4): 185–212.
24. Santschi, P.H. et al. (2003). Control of acid polysaccharide production, and ^{234}Th and POC export fluxes by marine organisms. *Geophysical Res. Lett.* **30**(2), Art. No. 1044, dio 10.1029/2002GL016046.
25. Guo, L., Tanaka, N., Schell, D.M., and Santschi, P.H. (2003). Nitrogen and carbon isotopic composition of high-molecular-weight dissolved organic matter in marine environments. *Mar. Ecol. Prog. Ser.* **252**: 51–60.
26. Benner, R. (2002). Chemical composition and reactivity. In: *Biogeochemistry of Marine Dissolved Organic Matter*. D.A. Hansell and C.A. Carlson (Eds.). Academic Press, New York, pp. 59–90.
27. Aluwihare, L., Repeta, D., and Chen, R. (2002). Chemical composition and cycling of dissolved organic matter in the Mid-Atlantic Bight. *Deep Sea Res.* **II**(49): 4421–4437.
28. Guo, L., Santschi, P.H., and Warnken, K.W. (2000). Trace metal composition of colloidal organic matter in marine environments. *Mar. Chem.* **70**(4): 257–275.
29. Santschi, P.H., Lenhart, J.J., and Honeyman, B.D. (1997). Heterogeneous process affecting trace contaminant distribution in estuaries: the role of natural organic matter. *Mar. Chem.* **58**: 99–125.
30. Santschi, P.H., Guo, L., and Bianchi, T.S. (1999). Dissolved organic matter in estuaries of the Gulf of Mexico. In: *Biogeochemistry of Gulf of Mexico Estuaries*. T.S. Bianchi (Ed.). John Wiley & Sons, New York, pp. 269–299.
31. Stumm, W. and Morgan, J. (1996). *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters*. 3rd Edn. John Wiley & Sons, New York.
32. Guo et al. (2002).
33. Rudd, T., Sterritt, R., and Lester, J. (1984). Formation and conditional stability constants of complexes formed between heavy metals and bacterial extracellular polymers. *Water Res.* **18**: 379–384.
34. Quigley, M.S. et al. (2002). Importance of polysaccharides for ^{234}Th complexation to marine organic matter. *Limnol. Oceanogr.* **47**: 367–377.
35. Alvarado-Quiroz, N. (2004). Ph.D. Dissertation, Dept. of Oceanography, Texas A&M University, College Station, TX.
36. Kohn, R. (1972). Ion binding on polyuronides—alginate and pectin. *Pure Appl. Chem.* **42**: 371–397.
37. Dickinson, E. (2003). Hydrocolloids at interfaces and the influence on the properties of dispersed systems. *Food Hydrocolloids* **17**: 25–39.
38. Decho, A.W. (1990). Microbial exopolymer secretions in the ocean environments their role(s) in food webs and marine processes. *Oceanogr. Mar. Biol. Annu. Rev.* **28**: 73–153.

39. Passow, U. (2002). Transparent exopolymer particles (TEP) in aquatic environments. *Prog. Oceanogr.* **55**: 287–333.
40. Honeyman, B.D. and Santschi, P.H. (1989). Brownian-pumping model for oceanic trace metal scavenging: Evidence from Th isotopes. *J. Mar. Res.* **47**: 951–992.
41. Honeyman, B. and Santschi, P. (1991). Coupling of trace metal adsorption and particle aggregation: Kinetic and equilibrium studies using ^{59}Fe-labeled hematite. *Environ. Sci. Technol.* **25**: 1739–1747.
42. Guo, L., Santschi, P.H., and Baskaran, M. (1997). Interaction of thorium isotopes with colloidal organic matter in oceanic environments. *Colloids Surf. A* **120**: 255–272.
43. Muller, F. (1996). Interactions of copper, lead and cadmium with the dissolved, colloidal and particulate components of estuarine and coastal waters. *Mar. Chem.* **52**: 245–268.
44. Wells et al. (1998).
45. Stordal, M.C., Santschi, P.H., and Gill, G.A. (1996). Colloidal pumping: Evidence for the coagulation process using natural colloids tagged with ^{203}Hg. *Environ. Sci. Technol.* **30**: 3335.
46. Guo, L. et al. (1996). Cycling of high molecular weight dissolved organic matter in the Middle Atlantic Bight as revealed by carbon isotopic (^{13}C and ^{14}C) signatures. *Limnol. Oceanogr.* **41**: 1242–1252.
47. Guo and Santschi. (2000).
48. Cowen, J. and Bruland, K. (1985). Metal deposits associated with bacteria—implications for Fe and Mn marine biogeochemistry. *Deep Sea Res.* **32**: 253–272.
49. Taillefert et al. (2000).
50. Kinrade, S., Gillson, A., and Knight, C. (2002). Silicon-29 NMR evidence of a transient hexavalent silicon complex in the diatom *Navicula pelliculosa*. *J. Chem. Soc., Dalton Trans.* **3**: 307–309.
51. Leveille, R., Fyfe, W., and Longstaffe, F. (2000). Geomicrobiology of carbonate-silicate microbialites from Hawaiian basaltic sea caves. *Chem. Geol.* **169**: 339–355.
52. Alldredge, A.L., Passow, U., and Logan, B.E. (1993). The abundance and significance of a class of large, transparent organic particles in the ocean. *Deep-Sea Res.* **I**(40): 1131–1140.
53. Mopper et al. (1995).
54. Murphy, R.J., Lenhart, J.J., and Honeyman, B.D. (1999). The sorption of thorium (IV) and uranium (VI) to hematite in the presence of natural organic matter. *Colloids Surf. A: Physicochemical Eng. Aspects* **157**: 47–62.
55. Muller, F.L.L. (1998). Colloid/solution partitioning of metal-selective organic ligands, and its relevance to Cu, Pb and Cd cycling in the Firth of Clyde. *Estu. Coast. Shelfish Sci.* **46**: 419–437.
56. Muller, F. (1999). Evaluation of the effects of natural dissolved and colloidal organic ligands on the electrochemical lability of Cu, Pb, and Cd in the Arran Deep, Scotland. *Mar. Chem.* **67**: 43–60.
57. Lead, J., Hamilton-Taylor, J., Davison, W., and Harper, M. (1999). Trace metal sorption by natural particles and coarse colloids. *Geochim. Cosmochim. Acta* **63**: 1661–1670.
58. Stenstörm, T.A. (1989). Bacterial hydrophobicity, an overall parameter for the measurement of adhesion potential to soil particles. *Appl. Environ. Microbiol.* **55**: 142.
59. Ahimou, F. et al. (2001). Influence of electrical properties on the evaluation of the surface hydrophobicity of Bacillus subtilis. *J. Microbiol. Methods* **45**: 119–126.
60. Repeta et al. (2002).
61. Amon, R.M.W. and Benner, R. (1994). Rapid cycling of high molecular weight dissolved organic matter in the ocean. *Nature* **369**: 549–552.
62. Leenheer, J.A. et al. (2001). Molecular resolution and fragmentation of fulvic acid by electrospray ionization/multistage tandem mass spectrometry. *Anal. Chem.* **73**: 1461–1471.
63. Kujawinski, B. et al. (2002). The application of electrospray ionization mass spectrometry (ESI MS) to the structural characterization of natural organic matter. *Org. Geochem.* **33**: 171–180.

READING LIST

Gustafsson, O. and Gschwend, P. (1997). Aquatic colloids: concepts, definitions, and current challenges. *Limnol. Oceanogr.* **42**: 519–528.

DEEP WATER CORALS

NOAA Coral Reef Information
System (CORIS)

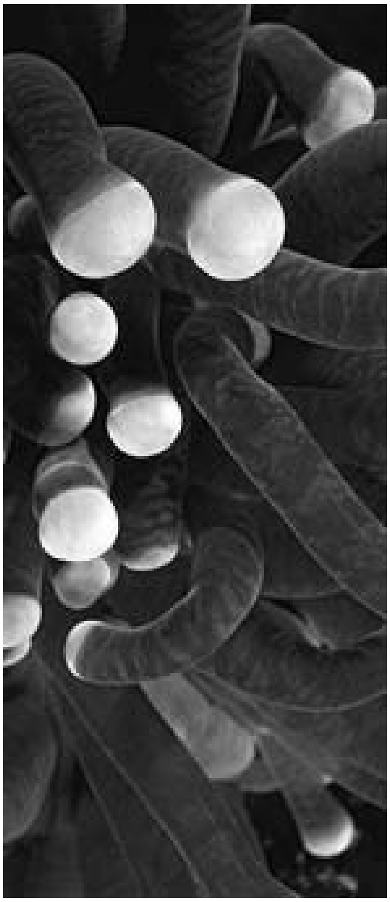

Many people are familiar with the coral reefs that thrive in shallow, well lighted, clear tropical waters where

---

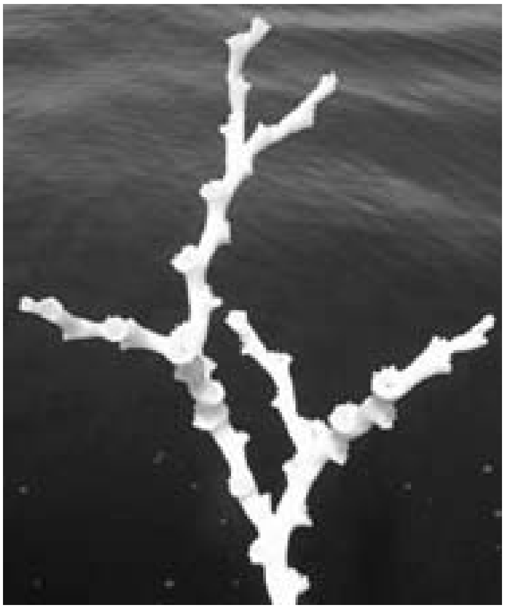

This healthy branch of *Lophelia* coral was sampled from deep ocean reefs off the coast of South Carolina. Unlike tropical species of coral, *Lophelia* possesses no symbiotic algae (zooxanthellae).

myriad colorful hard and soft corals provide habitat "infrastructure" for numerous invertebrates and fishes. The corals provide protection and cover, sources of nutrition, and sites for reproduction. Corals, however, also grow in the deep, cold sea. Although the existence of some of these deep-sea coral thickets has been known for several centuries, initially from pieces of broken corals brought up with fishing gear, scientists know little about their distribution, biology, behavior, and function as essential habitats for fishes and invertebrates.

Some deep-water corals (also called cold-water corals) do not form reefs exactly like those in tropical waters. Often, they form colonial aggregations called patches, mounds, banks, bioherms, massifs, thickets or groves. These aggregations are often still referred to as "reefs." Deep-water corals also provide crucial habitat and reproductive grounds for commercially important fisheries including sea bass, snapper, porgy, rock shrimp and calico shrimp, thus drawing the commercial fishing industry to these fragile areas.

Human activities constitute the most serious threat to these fragile corals. Destructive bottom fishing, as well as oil and gas exploration and exploitation have the potential to destroy large areas of coral habitat in a relatively short time. These activities create coral rubble, which is not a suitable habitat for fishes and invertebrates. In recent years, scientists have begun to study deep-water corals more closely, and some countries with deep-water corals in their territorial waters have begun to implement fishing restrictions in sensitive coral areas.

WHAT ARE DEEP-WATER CORALS?

Far fewer species of corals exist in cold, deep water than in the tropics. Three main groups of corals make up deep-water coral communities: hard (stony) corals of the Order Scleractinia, which form hard, ahermatypic reefs; black and horny corals of the Order Antipatharia; and soft corals of the order Alcyonacea, which includes the gorgonians (sea fans) (1). Deep-water corals are similar in some ways to the more familiar corals of shallow, tropical seas. Like their tropical equivalents, the hard corals develop sizeable reef structures that host rich and varied invertebrate and fish fauna. However, unlike their tropical cousins, which are typically found in waters above 70 m depth and at temperatures between 23° and 29 °C, deep-water corals live at depths just beneath the surface to the abyss (2000 m), where water temperatures may be as cold as 4 °C and utter darkness prevails.

At these depths, corals lack zooxanthellae. These symbiotic algae provide food for many shallow-water corals through photosynthesis. They also assist in the formation of the calcareous skeleton, and give most tropical corals their coloration. By contrast, the polyps of deep-water corals appear to be suspension feeders. They capture and consume organic detritus and plankton that are transported by strong, deep-sea currents. These corals are commonly found along bathymetric highs such as seamounts, ridges, pinnacles and mounds (2).

Deep-water corals range in size from small solitary colonies to huge, massive reef structures, which appear as oases of teeming life surrounded by more barren bathymetry. The gorgonians (sea fans) also range from small individuals to those with tree-like dimensions. The gorgonian, *Paragorgia arborea,* may grow in excess of three meters in length (3). Growth rates of branching deep-water coral species, such as *Lophelia* and *Oculina,* range

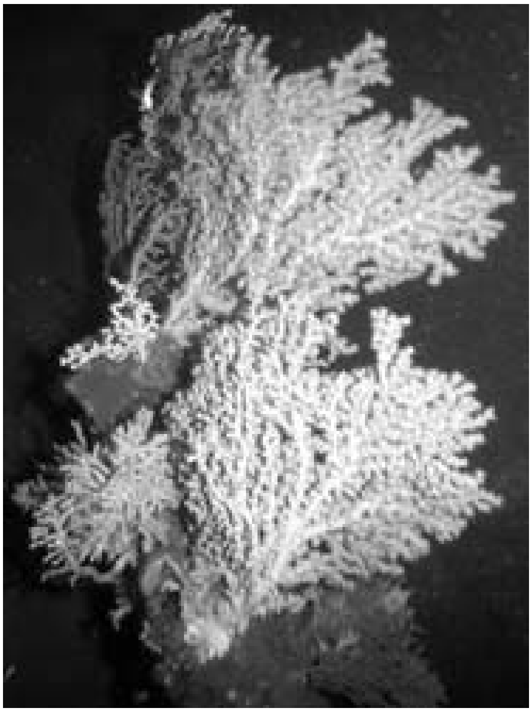

This thicket of *Paragorgia* corals was viewed by the deep-sea submersible *Alvin* at 1,043 m depth. *Photo*: Barbara Hecker.

from ~1.0–2.5 cm/yr, whereas branching shallow-water corals, such as *Acropora,* may exceed 10–20 cm/yr. Using coral age-dating methods, scientists have estimated that some living deep-water corals date back at least 10,000 years (4). However, little is known of their basic biology, including how they feed or their methods and timing of reproduction.

GENERAL DISTRIBUTION OF COLD-WATER CORALS

Deep-water corals are found globally, from coastal Antarctica to the Arctic Circle. In northern Atlantic waters, the principal coral species that contribute to reef formation are *Lophelia pertusa, Oculina varicosa, Madrepora oculata, Desmophyllum cristagalli, Enallopsammia rostrata, Solenosmilia variabilis*, and *Goniocorella dumosa*. Four of those genera (*Lophelia, Desmophyllum, Solenosmilia*, and *Goniocorella*) constitute the majority of known deep-water coral banks at depths of 400 to 700 m (5).

Madrepora oculata occurs as deep as 2,020 m and is one of a dozen species that occur globally and in all oceans, including the Subantarctic (6). Colonies of *Enallopsammia* contribute to the framework of deep-water coral banks found at depths of 600 to 800 m in the Straits of Florida (5).

TWO IMPORTANT DEEP-WATER CORALS

Two of the more significant deep-sea coral species are *Lophelia pertusa* and *Oculina varicosa*. These species form extensive deep-water communities that attract commercially important species of fishes, making them susceptible to destructive bottom trawling practices (7). Increased sedimentation places additional stress on corals.

Oil and gas exploration structures and activities, particularly in the North Sea and adjacent areas, also damage *Lophelia* communities. Subsequent oil and gas production activities may also introduce noxious substances into these areas.

LOPHELIA PERTUSA DISTRIBUTION

Lophelia pertusa is the most common aggregate-forming deep-water coral. Typically, it is found at depths between 200 and 1,000 m in the northeast Atlantic, the Mediterranean Sea, along the mid-Atlantic Ridge, the West African and Brazilian coasts, and along the eastern shores of North America (e.g., Nova Scotia, Blake Plateau, Florida Straits, Gulf of Mexico) as well as in parts of the Indian and Pacific Oceans. Like tropical coral reefs, *Lophelia* communities support diverse marine life, such as sponges, polychaete worms, mollusks, crustaceans, brittle stars, starfish, sea urchins, bryozoans, sea spiders, fishes, and many other vertebrate and invertebrate species.

Lophelia has been found most frequently on the northern European continental shelves between 200 and 1000 m, where temperatures range from 4° to 12 °C, but it has also been found at depths greater than 2,000 m. Once a colonial patch is established, it can spread over a broad area by growing on dead and broken pieces of coral (coral rubble). *Lophelia* has a linear extension of the polyps of about 10 mm per year. The reef structure has been estimated to grow about 1 mm per year (8). Scientists have also found *Lophelia* colonies on oil installations in the North Sea (9).

Lophelia pertusa can occur in a variety of structures and forms. DNA-based sequencing tests conducted at the University of Southampton Oceanography Centre, UK, have indicated that different morphological varieties of *Lophelia* all belong to the same species (10).

NOTED LOPHELIA AREAS

The world's largest known deep-water *Lophelia* coral complex is the Røst Reef. It lies in depths between 300 and 400 m west of Røst Island in the Lofoten archipelago, Norway. Discovered during a routine survey in May 2002, the reef is still largely intact. It covers an area approximately 40 km long by 3 km wide.

Relatively close by is the Sula Reef, located on the Sula Ridge, west of Trondheim on the mid-Norwegian Shelf,

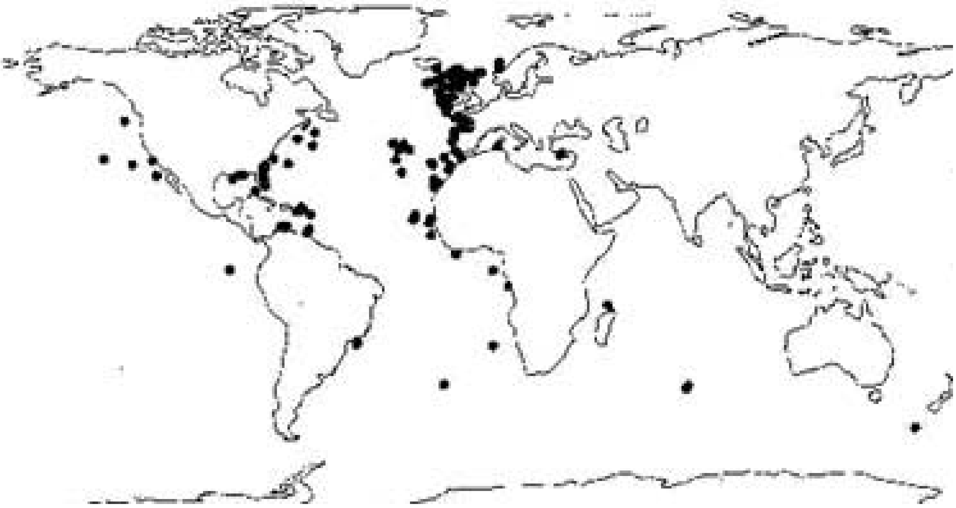

Global distribution of *Lophelia pertusa*. Image: Southampton Oceanography Centre, UK.

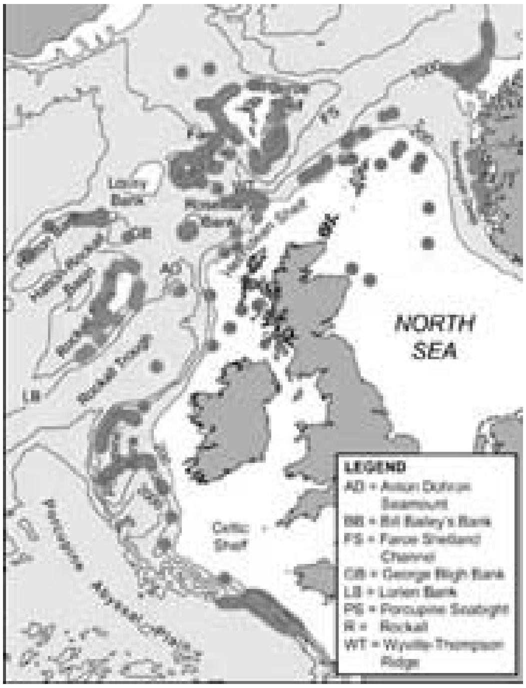

Reefs of *Lophelia pertusa* have been recorded on raised offshore seabeds in the north Atlantic off Britain and west Ireland. Image: Dr. Brian Bett, Southampton Oceanography Centre, UK.

at 200 to 300 m depth. It is estimated to be 13 km long, 700 m wide, and up to 35 m high (11), an area one-tenth the size of the 100 km^2 Røst Reef.

Discovered and mapped in 2002, Norway's 1,000-year-old and 2-km-long Tisler Reef lies in the Skagerrak—the submarine border between Norway and Sweden at a depth of 74 to 155 m. The Tisler Reef contains the world's only known yellow *Lophelia pertusa* corals. At present, Norway is the only European nation to enact laws protecting its *Lophelia* reefs from trawling, pollution, and oil and gas exploration. However, the European Commission has introduced an interim ban on trawling in the Darwin Mounds area, west of Scotland, in August 2003. A permanent ban on trawling is expected to follow.

Elsewhere in the northeastern Atlantic, *Lophelia* is found around the Faroe Islands, an island group between the Norwegian Sea and the Northeast Atlantic Ocean. At depths from 200 to 500 m, *Lophelia* is chiefly on the Rockall Bank and on the shelf break north and west of Scotland (12).

One of the most researched deep-water coral areas in the United Kingdom are the Darwin Mounds. The mounds were discovered in 1998 by the Atlantic Frontier Environmental Network (AFEN) while conducting large-scale regional surveys of the sea floor north of Scotland. They discovered two areas of hundreds of sand and cold-water coral mounds at depths of about 1,000 m, in the northeast corner of the Rockall Trough, approximately 185 km northwest of the northwest tip of Scotland. Named after the research vessel *Charles Darwin,* the Darwin Mounds have been extensively mapped using low-frequency side-scan sonar. They cover an area of approximately 100 km^2 and consist of two main fields—the Darwin Mounds East, with about 75 mounds, and the Darwin Mounds West, with about 150 mounds. Other mounds are scattered in adjacent areas. Each mound is about 100 m in diameter and 5 m high.

The tops of the mounds are covered with *Lophelia* corals and coral rubble, which attract other marine life. Side-scan sonar images show that the mounds appear to be sand volcanoes, each with a unique feature—a "tail." The tails are up to several hundred m long, all oriented downstream (13). The tails and the mounds are uniquely characterized by large congregations of deep-sea organisms called xenophyophores (*Syringammina fragilissima*), which are giant unicellular organisms that can grow up to 25 cm in diameter (14). Scientists are uncertain why these interesting organisms congregate in these areas. In addition, the *Lophelia* of the Darwin Mounds are growing on sand rather than hard substrate, an unusual condition unique to this area. Usually, coral larvae almost always settle and grow on hard substrates, such as dead coral skeletons or rock.

Lophelia corals exist in Irish waters as well (15). The Porcupine Seabight, the southern end of the Rockall Bank, and the shelf to the northwest of Donegal all exhibit large, mound-like *Lophelia* structures. One of them, the Theresa Mound, is particularly noted for its *Lophelia pertusa* and *Madrepora oculata* colonies. *Lophelia* reefs are also found along the U.S. East Coast at depths of 500 to 850 m along the base of the Florida-Hatteras slope. South of Cape Lookout, NC, rising from the flat sea bed of the Blake Plateau, is a band of ridges capped with thickets of *Lophelia*. These are the northernmost East Coast *Lophelia pertusa* growths. The coral mounds and ridges here rise as much as 150 m from the plateau plain. These *Lophelia* communities lie in unprotected areas of potential oil and gas exploration and cable-laying operations, rendering them vulnerable to future threats (16).

Finally, *Lophelia* is known to exist around the Bay of Biscay, the Canary Islands, Portugal, Madeira, the Azores, and the western basin of the Mediterranean Sea (13).

OCULINA VARICOSA DISTRIBUTION

Oculina varicosa is a branching ivory coral that forms giant but slow-growing, bushy thickets on pinnacles up to 30 m in height. The Oculina Banks, so named because they consist mostly of *Oculina varicosa*, exist in 50 to 100 m of water along the continental shelf edge about 26 to 50 km off of Florida's central east coast.

Discovered in 1975 by scientists from the Harbor Branch Oceanographic Institution conducting surveys of the continental shelf, *Oculina* thickets grow on a series of pinnacles and ridges extending from Fort Pierce to Daytona, Florida (7,17–19).

Like the *Lophelia* thickets, the Oculina Banks host a wide array of macroinvertebrates and fishes. They also are significant spawning grounds for commercially important species of food fishes including gag, scamp, red grouper, speckled hind, black sea bass, red porgy, rock shrimp, and the calico scallop (20).

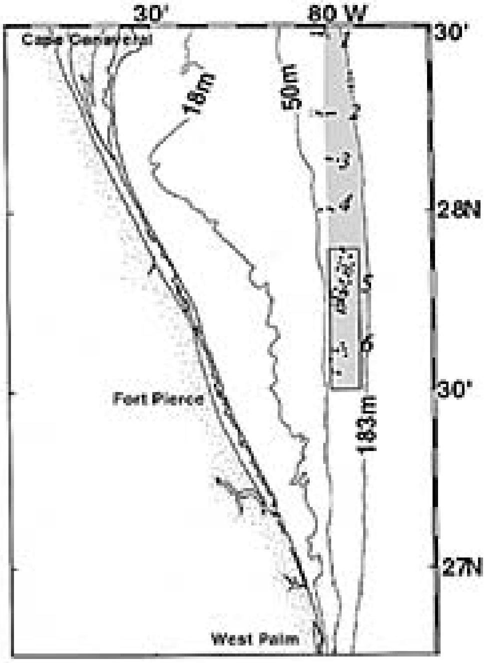

Chart of the Deep-water Oculina Coral Banks Marine Protected Area (MPA). The shaded area is the entire Oculina Habitat Area of Particular Concern (HAPC); the Experimental Oculina Research Reserve (EORR) is the smaller inset box. Recent dive sites from 2001 to 2003 include: 1) Cape Canaveral, 2) Cocoa Beach, 3) Eau Gallie, 4) Malabar, 5) Sebastian, and 6) Chapmans Lumps/Jeff's Reef. (Courtesy of: Dr. John K. Reed, Harbor Branch Oceanographic Institution).

THREATS TO LOPHELIA AND OCULINA CORALS

Both *Lophelia* and *Oculina* corals face uncertain futures. Until recently, *Lophelia* habitats remained undisturbed by human activity. However, as traditional fish stocks are depleted from northern European waters, bottom trawling has moved into deeper waters, where the gear has affected the coral beds. The towed nets break up the reef structure, kill the coral polyps and expose the reef to sediment by altering the hydrodynamic and sedimentary processes around the reef (21).

The fishes and invertebrates that depend on the coral structure lose their habitat and move out of the area. Damage may range from a decrease in the size of the coral habitat with a corresponding decrease in the abundance and biodiversity of the associated invertebrate and fish species, to the complete destruction of the coral habitat. The trawls also may resuspend sediments that, in turn, may smother corals growing downstream of the current. In addition, oil and gas exploration and extraction operations have begun to move into these deep-water areas, further threatening the resident corals.

Scientists estimate that within the Norwegian Exclusive Economic Zone, 30 to 50 percent of the total coral area of the Norwegian shelf has been damaged or destroyed by trawling (8). Scientists from the International Council for the Exploration of the Sea, the main provider of scientific advice to the European Commission on fisheries and environmental issues in the northeast Atlantic, have recommended that to protect the remaining deep-water coral groves, all of Europe's deep corals must be accurately mapped and then closed to fishing trawlers (22).

In 1999, the first complete mapping of the Sula Reef was carried out by the Norwegian Hydrographic Society, which used the latest available multibeam echosounder equipment to record both depth and backscatter data. The mapping was the product of joint cooperation between the Geological Survey of Norway and the Institute of Marine Research. That same year, the Norwegian Ministry of Fisheries issued regulations for the protection of the *Lophelia* reefs. An area of 1,000 km^2 at Sula, including the large reef, is now closed to bottom trawling. In 2000, an additional area, about 600 km^2 was closed. The Røst Reef, an area of about 300 km^2, was closed to bottom trawling in 2002.

Florida's Oculina Banks, once teeming with commercially important fish, now appear to be severely depleted of fish stocks (23). Much of the *Oculina* coral has been reduced to rubble, probably the result of a combination of destructive bottom trawling and natural causes like bioerosion and episodic die-offs. Regardless of the cause, the Oculina Banks now support a drastically reduced fishery because most of the significant spawning grounds have been destroyed (7,19).

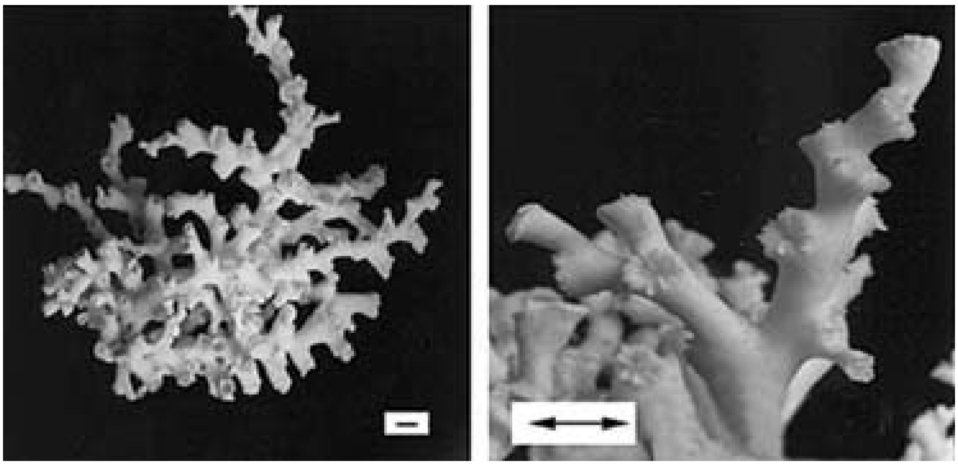

Two images of the deep sea coral *Lophelia pertusa*. The image on the left shows the complex growth structure of a small colony. On the right is a closeup of individual polyps. *Photo*: Ref. 7.

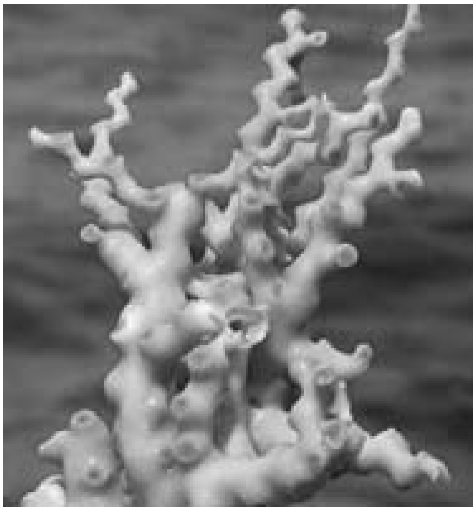

This beautiful species of *Lophelia* coral was collected off the coast of South Carolina.

Efforts to protect the Oculina Banks began in 1980, when scientists from Harbor Branch Oceanographic Institution initiated a call to implement protective measures for the area. Since then, progressively stricter legal protections have been implemented to facilitate a recovery. In 1984, a 315 km^2 portion of deep-water Oculina reef system was designated as a Habitat Area of Particular Concern by the South Atlantic Fishery Management Council (7,19), a designation that categorized it as an area of special ecological significance worthy of stricter regulatory and enforcement activity.

In 1994, after the area showed no significant recovery, part of the Oculina Banks was completely closed to bottom fishing in an area called the Experimental *Oculina* Research Reserve. In 1996, the South Atlantic Fisheries Management Council implemented additional protections in the reserve by prohibiting fishing vessels from dropping anchors, grapples, or attached chains there. In 1998, the council also designated the reserve as an Essential Fish Habitat. In 2000, the deep-water Oculina Marine Protected Area was extended to 1029 km^2 (7,19,24).

The Oculina Banks remain a hot spot for research and efforts to rehabilitate the coral (23). Scientists recently deployed concrete "reef balls" in the area in an attempt to attract fish and provide substrate for coral attachment, settlement and growth. They are cautiously optimistic about their initial restorative efforts in the reserve (23).

BIBLIOGRAPHY

1. Williams, G.C. 2001. Octocoral Research Center Web Site. Available on-line at http://www.calacademy.org/research/izg/orc_home.html
2. Southampton Oceanographic Centre. Coral ecosystems in deep water. Available on-line at http://www.soc.soton.ac.uk/GDD/DEEPSEAS/coralecosystems.html
3. Watling, L. 2001. Deep sea coral. Available on the NOAA Ocean Explorer Web site at http://oceanexplorer.noaa.gov/explorations/deepeast01/background/corals/corals.html.
4. Mayer, T. 2001. 2000 Years Under the Sea. Available on-line at http://exn.ca/Stories/2001/10/16/52.asp
5. Cairns, S. and G. Stanley. 1982. *Ahermatypic coral banks: Living and fossil counterparts*. Proceedings of the Fourth International Coral Reef Symposium, Manila (1981), 1: 611–618.
6. Cairns, S.D. 1982. *Antarctic and Subantarctic Scleractinia*. *Antarctic Research Series* 34: 74 pp.
7. Reed, J.K. 2002. *Comparison of deep-water coral reefs and lithoherms off southeastern U.S.A.* Hydrobiologia 471: 43–55.
8. Fossa, J.H., P.B. Mortensen, and D.M. Furevic. 2002. The deep-water coral *Lophelia pertusa* in Norwegian waters: distribution and fishery impacts. Hydrobiologia 417: 1–12. Available on-line at: http://www.imt.no/dokumenter/fossa.pdf
9. Bell, N. and J. Smith. 1999. Coral growing on North Sea oil rigs. *Nature* **402**: 601.
10. Rogers, A., M. Le Goff, and B. Stockley. Ecology of deep-water stoney corals (the Darwin Mounds). Available on-line at http://www.soton.ac.uk/~merg/corals.htm
11. Bellona Foundation. 2001. Coral reefs in Norwegian Waters. Available on-line at http://www.bellona.no/en/b3/biodiversity/coral_reefs/12784.html.
12. Tyler-Walters, H. 2003. *Lophelia* reefs. Marine Life Information Network: Biology and Sensitivity Key Information Sub-programme. Plymouth: Marine Biological Association of the United Kingdom. Available on-line at http://www.marlin.ac.uk/Bio_pages/Bio_BasicInfo_COR.Lop.htm

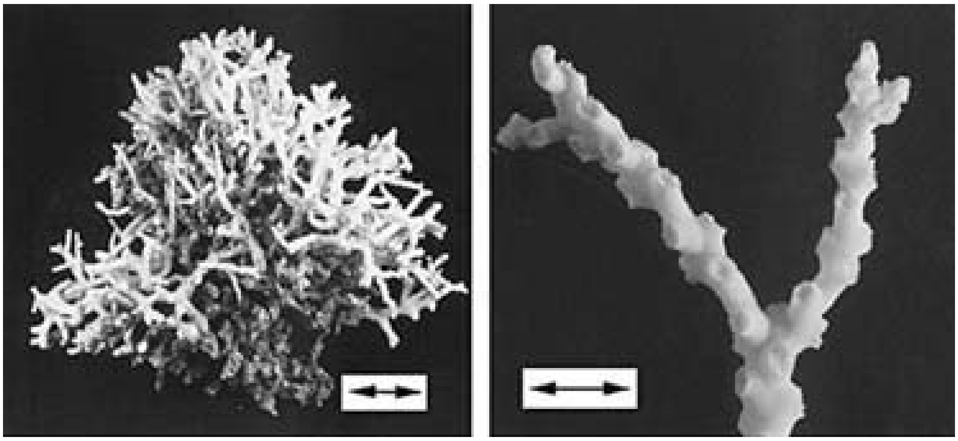

Two images of the deep-sea copral *Oculina varicosa*. The image on the left shows the complex growth structure of a small colony. On the right is a closeup of an individual branch. *Photo*: Ref. 7.

13. International Council for Exploration of the Sea (ICES). 2001a. Marine World, article on deep-sea coral. Available on-line at http://www.ices.dk/marineworld/deepseacoral.asp
14. International Council for Exploration of the Sea (ICES). 2002. *Advisory Committee on Ecosystems*. Report of the study group on mapping the occurrence of cold-water corals: 1–17.
15. Rogers, A.D. 1999. The biology of *Lophelia pertusa* (Linnaeus 1758) and other deep-water reef-forming corals and impacts from human activities. *International Review of Hydrobiology* **84**: 315–406.
16. Sulak, K. and S. Ross. 2001. A profile of the *Lophelia* reefs. Available on the NOAA Ocean Explorer Web site at http://oceanexplorer.noaa.gov/explorations/islands01/background/islands/sup10_lophelia.html.
17. Avent, R.M., M.E. King, and R.M. Gore. 1977. Topographic and faunal studies of shelf-edge prominences off the central eastern Florida coast. *Int. Revue ges.Hydrobiol.* **62**: 185–208.
18. Reed, J.K. 1981. *In Situ growth rates of the scleractinian coral Oculina varicosa occurring with zooxanthellae on 6-m reefs and without on 80-m banks*. pp. 201–206. in: W.J. Richards (ed.), Proceedings of Marine Recreational Fisheries Symposium.
19. Reed, J.K. 2002. Deep-water Oculina reefs of Florida: biology, impacts and management. *Hydrobiologia* **471**: 43–55.
20. Koenig, C., C.F. Coleman, and S. Brooke. Coral restoration in the Experimental *Oculina* Research Reserve of the South Atlantic. Available on-line at http://www.bio.fsu.edu/ifre/ifre_research_oculina.html
21. Hall-Spencer, J. et al. 2002. *Trawling damage to Northeast Atlantic ancient coral reefs*. Proc. R. Soc. London B.
22. International Council for Exploration of the Sea (ICES). 2001b. Close Europe's cold-water coral reefs to fishing. Available on-line at http://www.ices.dk/aboutus/pressrelease/coral.asp
23. Marine Protected Areas of the United States (MPA). 2002. Experimental *Oculina* Research Reserve. Available on-line at http://mpa.gov/mpadescriptive/cs_obrr.html
24. National Oceanic and Atmospheric Administration. 2000. *Final Rule, Amendment 4 to the Fishery Management Plan for Coral, Coral Reefs, and Live / Hard Bottom Habitats of the South Atlantic Regions (Coral FMP)*. Federal Register, Vol. 65, No. 115, June 14, 2000.

READING LIST

Atlantic Coral Ecosystem Study (ACES), 2000. Consortium of European scientists completed an environmental assessment of the status of Europe's deep-water corals, available on-line at: http://www.cool-corals.de

Hoskin, C.M., J.K. Reed, and D.H. Mook. 1987. *Sediments from a living shelf-edge reef and adjacent area off central eastern Florida*. Pp. 42–77, In: F, JMR, Maurrasse (ed.), Symposium on south Florida geology. Miami Geological Society. Memoirs 3.

International Council for Exploration of the Sea (ICES). 2001c: *Extract of the report of the Advisory Committee on Ecosystems 2002 on identification of areas where cold-water corals may be affected by fishing*.

National Oceanic and Atmospheric Administration. 1982. *Fishery Management Plan for Coral and Coral Reefs of the Gulf of Mexico and South Atlantic*. Gulf of Mexico and South Atlantic Fishery Management Councils.

Reed, J.K. and C.M. Hoskin. 1984. Studies of geological and biological processes at the shelf-edge with use of submersibles. In: *Undersea research and technology—Scientific application and future needs—abstracts*, *Symposia Series for Undersea Research*, Vol. 2, NOAA.

Reed, J.K., R.H. Gore, L.E. Scotto and K.A. Wilson. 1982. Community composition, structure, aerial and trophic relationships of decapods associated with shallow and deep-water *Oculina varicosa* coral reefs. *Bull. Mar. Sci.* **32**: 761–786.

MARINE DEBRIS ABATEMENT

JAVIER VELEZ-AROCHO
U.S. Environmental Protection Agency
Washington, District of Columbia

TRASH IN OUR OCEANS—YOU CAN BE PART OF THE SOLUTION

Marine debris, often called litter, has become a problem along shorelines, coastal waters, estuaries, and oceans throughout the world. It is defined as any man-made, solid material that enters our waterways directly (e.g., by dumping) or indirectly (e.g., washed out to sea via rivers, streams, storm drains, etc.). Objects ranging from detergent bottles, hazardous medical wastes, and discarded fishing line all qualify as marine debris. In addition to being unsightly, it poses a serious threat to everything with which it comes into contact. Marine debris can be life-threatening to marine organisms and humans and can wreak havoc on coastal communities and the fishing industry.

Sources of Marine Debris

There are two different sources from which debris pollutes our oceans. The first is from the land and includes users of

---

the beach, storm water-runoff, landfills, solid waste, rivers, and streams, floating structures, ill maintained garbage bins and dumps and litterbugs. Marine debris also comes from combined sewer overflows, and storm drains. Typical debris from these sources includes medical waste, street litter and sewage. Land-based sources cause 80% of the marine debris found on our beaches and waters.

The second source of marine debris is from ocean sources, and this type of debris includes galley waste and other trash from ships, recreational boaters and fishermen and offshore oil and gas exploration and production facilities.

Adding to this problem is the population influx along our nation's shores. More people means more paved area and wastes generated in coastal areas. These factors; combined with the growing demand for manufactured and packaged goods, have led to an increase in non-biodegradable solid wastes in our waterways.

The Beaches Environmental Assessment and Coastal Health Act (BEACH) of 2000

The BEACH Act was enacted on October 10, 2000, and it is designed to reduce the risk of disease to users of the Nation's coastal recreation waters. The act authorizes the EPA to award program development and implementation grants to eligible states, territories, tribes, and local governments to support microbiological testing and monitoring of coastal recreational waters, including the Great Lakes, that are adjacent to beaches or similar points of access used by the public. BEACH Act grants provide support for developing and implementing programs to notify the public of the potential for exposure to disease-causing microorganisms in coastal recreation waters. The act also authorizes EPA to provide technical assistance to States and local governments for the assessment and monitoring of floatable materials. In partially fulfilling that obligation, EPA has compiled the most current information to date on assessing and monitoring floatable materials in the document Assessing and Monitoring Floatable Debris.

The International Coastal Cleanup

The Ocean Conservancy, formerly known as the Center for Marine Conservation, established and maintains the annual International Coastal Cleanup (ICC) with support from EPA and other stakeholders. The first cleanup was in 1986 in Texas, and the campaign currently involves all of the states and territories of the United States and more than 100 countries around the world. The ICC is the largest volunteer environmental data-gathering effort and associated cleanup of coastal and underwater areas in the world. It takes place every year on the third Saturday in September. In 2001, over 140,000 people across the U.S. participated in the ICCC. They removed about 3.6 million pounds of debris from more than 7,700 miles of coasts, shorelines, and underwater sites.

National Marine Debris Monitoring Program

EPA along with other federal agencies helped to design the National Marine Debris Monitoring Program (NMDMP), and EPA is supporting The Ocean Conservancy's implementation of the study. NMDMP is designed to gather scientifically valid marine debris data following a rigorous statistical protocol. The NMDMP is designed to identify trends in the amounts of marine debris affecting the U.S. coastline and to determine the main sources of the debris. This scientific study is conducted every 28 days by teams of volunteers at randomly selected study sites along the U.S. coastline. The NMDMP requires, at a maximum, that 180 monitoring sites located along the coast of contiguous U.S. States and Alaska, Hawaii, Puerto Rico, and the U.S. Virgin Islands be fully operational. The program began in 1996 with the establishment of 40 monitoring sites ranging from the Texas/Mexico border to Port Everglades, Florida and included Puerto Rico and the U.S. Virgin Islands. To date 163 study sites have been designated and 128 sites are collecting data. The program will run for a 5-year period once all of the study sites have been established. For more information please visit: http://www.oceanconservancy.org/dynamic/learn/issues/debris/debris.htm.

Other Floatable Debris-Related Legislation

In response to a growing concern over floatable debris, governments have taken actions nationally, as well as internationally, to reduce discharges at their source. For example, intentional at-sea dumping of garbage generated on land became subject to international control in 1972 through the Convention on the Prevention of Marine Pollution by Dumping of Wastes and Other Matter,

commonly called the London Convention. Similarly, at-sea disposal of garbage generated during the routine operation of ships (e.g., garbage not deliberately carried to sea for the purpose of disposal) was addressed through a 1978 Protocol to the 1973 International Convention for the Prevention of Pollution by Ships (commonly called the MARPOL Convention).

In the Shore Protection Act of 1989 provides for controls on operations relating to the vessel transport of certain solid wastes (municipal or commercial waste) so that those wastes are not deposited in coastal waters. This act has provisions that reduce wastes being deposited in coastal waters during the transfer of waste from the waste source to the transport vessel to the waste receiving station.

As amended by the Water Quality Act of 1987, the Clean Water Act requires EPA to establish regulations that treat storm water and combined sewer overflows (CSOs) as point source discharges that must be regulated under the National Pollutant Discharge Elimination System (NPDES). The permits will prohibit non-storm water discharges into storm sewers while leading to improved control techniques and best management practices.

Debris Hazardous for Marine Animals

Each year millions of seabirds, sea turtles, fish, and marine mammals become entangled in marine debris or ingest plastics which they have mistaken for food. As many as 30,000 northern fur seals per year get caught in abandoned fishing nets and either drown or suffocate. Whales mistake plastic bags for squid, and birds may mistake plastic pellets for fish eggs. At other times, animals accidentally eat the plastic while feeding on natural food. According to recent estimates, 267 marine species have been reported entangled in or having ingested marine debris (Marine Mammal Commission, 1998). The plastic constricts the animals' movements, or kills the marine animals through starvation, exhaustion, or infection from deep wounds caused by tightening material. The animals may starve to death, because the plastic clogs their intestines preventing them from obtaining vital nutrients. Toxic substances present in plastics can cause death or reproductive failure in the fish, shellfish, and wildlife that use the habitat.

Polluted Waters Make Trouble for Humans, Too

Humans can also be directly affected by marine debris. Swimmers and divers can become entangled in abandoned netting and fishing lines like marine organisms. Beach users can be injured by stepping on broken glass, cans, needles or other litter. Appearance of debris, such as plastic, can also result in economic consequences. Floating debris, either as a floating slick or as dispersed items, is visually unappealing and can result in lost tourism revenues. New Jersey now spends $1,500,000 annually to clean up its beaches, and $40,000 to remove debris from the New York/New Jersey Harbor (YOTO Discussion Papers, 1998).

Fishing Industries and Communities Hit Hard by Marine Debris

Marine debris also acts as a navigational hazard to fishing and recreational boats by entangling propellers and clogging cooling water intake valves. Repairing boats damaged by marine debris are both time consuming and expensive. Fixing a small dent in a large, slow-moving vessel can take up to 2 days, costing the shipping company $30,000–40,000 per day in lost carrying fees, as well as up to $100,000 for the repair itself (Peter Wallace, personal contact). According to Japanese estimates, the Japanese fishing industry spent $4.1 billion (U.S.) on boat repairs in 1992 (YOTO Discussion Papers, 1998). Lost lobster traps cost New England fishing communities $250 million in

1978. These traps continue to catch lobsters and other marine organisms that are never harvested and sold; the communities' economies are therefore adversely affected.

Beach Raking

Mechanical beach raking, which is accomplished with a tractor and is used to remove debris from the shoreline, can help to remove floatable material from beaches and marine shorelines. However, it can also be harmful to aquatic vegetation, nesting birds, sea turtles, and other types of aquatic life. A study in Maine compared a raked beach and an adjacent natural beach to determine the effects of beach raking on vegetation. Beach raking not only prevents the natural re-vegetation process, but it reduces the integrity of the sand root mat just below the surface that is important in slowing beach erosion. Other problems include disturbance of vegetation if raking is conducted too, close to a dune. By removing seaweed, beach erosion can also be caused. Sand compaction is reduced when seaweed is removed, resulting in suspension of the sand in the water during high tides and contributing to loss of sand and erosion of the beach. Beach cleaning machines are harmful to nest birds and can destroy potential nesting sites, crush plover nests and chicks, and remove the plovers' natural wrack-line feeding habitat. To reduce the effects on nesting birds, beach raking should not be done during the nesting season.

Unseen Consequences

Once debris reaches coastal and ocean bottom, especially in areas with little current, it may continue to cause environmental problems. When plastic film and other debris settle on the bottom, it can suffocate immobile plants and animals, producing areas essentially devoid of life. In areas with some currents, such as coral reefs, debris can wrap around living coral, smothering the animals and breaking up their coraline structures.

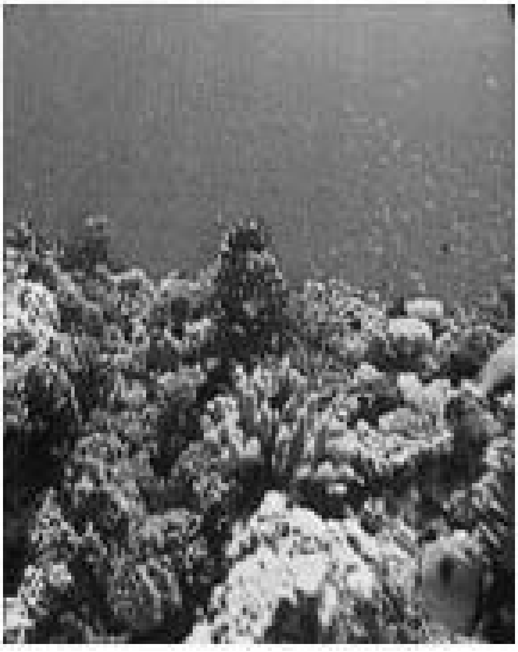

Trash Indicator of Bigger Problems

The typical floatable debris from Combined Sewer Overflows includes street litter, sewage (e.g., condoms, tampons, applicators), and medical items (e.g., syringes), resin pellets, and other material that might have washed into the storm drains or from land runoff. These materials or objects can make it unsafe to walk on the beaches, and pathogens or algae's blooms can make it unsafe to swim. Pollutants, such as toxic substances, can make it unsafe to eat the fish caught from the waters. Swimming in or ingesting waters which are contaminated with pathogens can result in human health problems such as, sore throat, gastroenteritis, meningitis or even encephalitis. Pathogens can also contaminate shellfish beds.

What Needs to Be Done

Reducing marine debris means reducing the amount of waste generated on land and disposing of it properly. Recycling can significantly reduce the amounts of litter reaching marine and coastal waters. Volunteer coastal cleanups and public education efforts can also help reduce the amount of debris reaching our waterways.

What You Can Do

- Dispose of trash properly. This will help reduce the amount of trash that is washed into our waterways from storm drains.
- Reduce, reuse and recycle. Visit the EPA's solid waste web site for more information.
- Cut the rings of plastic six-pack holders. This lowers the risk of entanglement to marine animals if the holders do make it out to sea.
- Participate in local beach, river or stream clean ups.
- Practice good housekeeping.
- Look for alternative materials or avoid excessive packaging when deciding on purchases.
- Educate others about marine debris.
- Get involved in your local area.

What the EPA and Other Partners Are Doing

- Providing resources to educate the public on the impact of marine debris and guidance for how to eliminate it.
- Providing resources and technical support for the development of studies to address the impact of marine debris.
- Providing guidance for the implementation of regulations addressing marine debris.
- Organizing and sponsoring cleanup efforts, like the U.S. and International Coastal Cleanups.

Whether directly or indirectly, many of our land-based activities ultimately impact and adversely affect the oceans. At the heart of the problem lie the great quantities of trash generated in our daily lives. Reusing and recycling efforts have helped, but a great deal more can be done. You can be part of the solution.

LARVAE AND SMALL SPECIES OF POLYCHAETES IN MARINE TOXICOLOGICAL TESTING

Jian-Wen Qiu
Ming Hung Wong
Hong Kong Baptist University
Kowloon Tong, Kowloon, Hong Kong

The ultimate goal of ecotoxicological testing is to provide data to predict potential impacts of pollutants on the ecosystem. To fulfill this goal, the tests should meet several criteria: the species should be ecologically relevant and sensitive to pollutants; the species should be readily available through field collection or laboratory culture; and the tests should yield reasonably rapid results at low cost. Although rarely consumed as seafood, polychaetes are a diverse group in marine ecosystems, typically consisting 35% to 75% of the total number of species in soft bottom coastal areas (1). They serve as prey for demersal fish and are important in the transfer of organic material and nutrients between the sediment and the overlying water. The use of polychaetes as bioassay organisms began in the 1960s, stimulated by the awareness that many pollutants were threatening the ecosystems and the surge in studies of polychaete life history. Reish and Gerlinger (2) is the most updated comprehensive review on the use of polychaetes in bioassays. Among the 8000 described species worldwide, only 49 polychaete species have been used in bioassays, of which all inhabit shallow waters. Most tests have been conducted with small species (*Neanthes arenaceodentata, Capitella capitata, Dinophilus gyrociliatus, Ophryotrocha labronica*, and *O. diadema*), which can be obtained from laboratory culture. Large polychaetes, i.e., *Nereis virens* and *Neanthes diversicolor*, have also been used. These species are usually obtained from field populations and have been used mostly in pollutant accumulation studies.

Polychaetes have been employed in acute, chronic, and life-cycle tests to determine the toxicity of metals, contaminated sediment, sediment pore water, municipal effluents, petroleum hydrocarbons, pesticides, and other synthetic organic chemicals. Although several species, such as *Capitella capitata, Dinophilus gyrociliatus, Ophryotrocha labronica*, and *O. diadema*, have been used in all test types, many other species have been used only in acute or chronic tests. The acute tests can be conducted with larvae, juveniles, and adults. During the test, the worms are usually not fed. When juveniles and adults are used, the acute test is usually run for 96 h. The test can be static, static renewal, or flow-through. In larval tests, however, the tests are usually static, with shorter exposure periods (24 or 48 h). Survivorship is usually the endpoint of acute tests. Chronic tests are usually conducted using juveniles or adults. The test protocol is the same as in acute test except the exposure duration is 10 to 28 days, during which the polychaetes are fed. The bioassay usually adopts a static renewal design. In addition to survival, sublethal effects such as cessation of feeding, defecation, or reduced growth are the endpoints. The life-cycle tests have been employed infrequently in the past. The test protocol is similar to that of the chronic test. In addition to survival and growth effects, reproductive parameters, such as eggs or juveniles produced, are the endpoints. Only small species have been used in life-cycle tests because their life cycle can be completed in less than one month. In recent years, many published toxicity studies have used small polychaete species in chronic or life-cycle tests. Examples are Bridges and Farrar (3), Anderson et al. (4), Linke-Gamenick et al. (5), McPherson et al. (6), and Nipper and Carr (7). There has also been new development in the use of polychaete larvae (8–10).

Below are several general conclusions regarding the use of polychaetes in bioassays. These conclusions are supported by examples.

1. *Polychaete Larvae are Very Sensitive to Pollutants*. In *Capitella capitata,* bifurcated abnormal larvae occurred at 10 ppb copper (11). In *Arenicola cristata*, abnormal larvae occurred at 0.75 ppb TBTO (12). In *Galeolaria caespitosa*, the 48-h EC50 for larval development ranged from 20 to 28 ppb copper (9).
2. *Polychaete Larvae are More Sensitive than the Adults to Pollutants*. In *Capitella capitata,* the 96-h LC50 for trochophores and adults was 0.22 and 7.5 ppm cadmium, respectively (13). In *Hydroides elegans*, the EC50 for trochophores was 29 ppb copper, whereas mortality of the adults occurred only when copper concentrations reached 500 ppb (10).
3. *Test Sensitivity Varies with Polychaete Species and Toxicant*. Reish and LeMay (14) compared the toxicity of seven metals with five polychaete species. They found that, for the same metal, different polychaete species gave very different results. For cadmium, the 96-h LC50 ranged from 2.6 ppm in *Pectinaria californiensis* to 14.1 ppm in *Neanthes arenaceodentata*. For the same species, sensitivity varies with toxicant. *Ophryotrocha labronica* was the most sensitive to copper, but it was the second most tolerant to chromium, and the most tolerant to mercury.
4. *The Chronic Test may not be More Sensitive than the Acute Test*. Reish and Gerlinger (2) noted that very limited data on the comparison of acute and chronic tests existed. After exposure of *Neanthes arenaceodentata* to 5.1 ppm cadmium for 20 days, no mortality existed, but the growth rate was significantly lower than that of the control. In another study with the same polychaete species, however, the 28-d LC50 values for mercury and copper were not lower than the 4-day values (15).
5. *Life-Cycle Tests may not be more Sensitive than Acute Tests*. At 100 ppb pentachlorophenol, *Dinophilus gyrociliatus* adults survived well, but the total number of eggs plus juveniles was significantly reduced (16). Similarly, at 50 ppb mercury, *Ophryotrocha diadema* survived well, but the number of offspring was reduced; in contrast, the 96-h LC50 for mercury was very close to the concentration that resulted in declined reproductive output (17).

BIBLIOGRAPHY

1. Qian, P.Y., Qiu, J.W., Kennish, R., and Reid, C.A. (2003). Recolonization of benthic infauna subsequent to capping of contaminated dredged material in East Sha Chau, Hong Kong. *Estuarine Coastal Shelf Sci.* **56**: 819–831.
2. Reish, D.J. and Gerlinger, T.V. (1997). A review on the toxicological studies with polychaetous annelids. *Bull Mar Sci.* **60**: 584–607.
3. Bridges, T.S. and Farrar, J.D. (1997). The influence of worm age, duration of exposure and endpoint selection on bioassay sensitivity for *Neanthes arenaceodentata* (Annelida: Polychaeta). *Environ Toxicol Chem.* **16**: 1650–1658.
4. Anderson, B.S. et al. (1998). Comparison of marine sediment toxicity test protocols for the amphipod *Rhepoxynius abronius* and the polychaete worm *Nereis (Neanthes) arenaceodentata*. *Environ Toxicol Chem.* **17**: 859–866.
5. Linke-Gamenick, I., Forbes, V.E., and Sibly, R.M. (1999). Density-dependent effects of a toxicant on life-history traits and population dynamics of a capitellid polychaete. *Mar Ecol Prog Ser.* **184**: 139–148.
6. McPherson, C.A. et al. (2002). Toxicity of 1,4-dichlorobenzene in sediments to juvenile polychaete worms. *Mar Pollut Bull.* **44**: 1405–1414.
7. Nipper, M. and Carr, R.S. (2003). Recent advances in the use of meifaunal polychaetes for ecotoxicological assessments. *Hydrologia.* **496**: 347–353.
8. Palau-Casellas, A. and Hutchinson, T.H. (1998). Acute toxicity of chlorinated organic chemicals to the embryos and larvae of the marine worm *Platynereis dumerilii* (Polychaeta: Nereidae). *Environ Toxicol Water Qual.* **13**: 149–155.
9. Ross, K.E. and Bidwell, J.R. (2001). A 48-h larval development toxicity test using the marine polychaete *Galeolaria caespitosa* Lamarck (Fam. Serpulidae). *Arch Environ Contam Toxicol.* **40**: 489–496.
10. Xie, Z.C., Wong, N.C., Qian, P.Y., and Qiu, J.W. (2004). Responses of polychaete *Hydroides elegans* life stages to copper stress. *Mar Ecol Prog Ser.* (in press).
11. Reish, D.J., Plitz, F.M., and Martin, J.M. (1974). Induction of abnormal polychaete worms. *Mar Poll Bull.* **51**: 125–126.
12. Walsh, G.E., Louie, M.K., McLaughlin, L.L., and Lores, E.M. (1986). Lugworm *Arenicola cristata* larvae in toxicity tests: survival and development when exposed to organotins. *Environ Toxicol Chem.* **5**: 749–754.
13. Reish, D.J., Martin, J.M., Plitz, F.M., and Word, J.Q. (1976). The effect of heavy metals on laboratory populations of two polychaetes with comparisons to the water quality conditions and standards in Southern California marine waters. *Water Res.* **10**: 299–302.
14. Reish, D.J. and LeMay, J.A. (1991). Toxicity and bioconcentration of metals and organic compounds by polychaetes. *Ophelia Suppl.* **5**: 653–660.
15. Reish, D.J. (1980). *The Effect of Different Pollutants on Ecologically Important Polychaete Worms*. U.S. Environmental Protection Agency. Environmental Research Laboratory, Narragansett, RI, EPA 600/3-80-053, p. 138.
16. Carr, R.S., Curran, M.D., and Mazurkiewicz, M. (1986). Evaluation of the archiannelid *Dinophilus gyrociliatus* for use in short-term life-cycle tests. *Environ Toxicol Chem.* **5**: 703–712.
17. Reish, D.J. and Carr, R.S. (1978). The effect of heavy metals on survival, reproduction, development and life cycles of two species of polychaetous annelids. *Mar Pollut Bull.* **9**: 24–27.

EL NIÑO: THE INTERANNUAL PREDICTION PROBLEM

Geophysical Fluid Dynamics
Laboratory—NOAA

Once every few years, the normally cool waters of the eastern tropical Pacific become unusually warm in a phenomenon known as El Niño. During El Niño years, disruptive weather patterns often occur over wide regions of the globe, including North America. For example, flooding rains can strike California, producing coastal erosion, mudslides, and crop damage, while droughts may occur in Australia and other regions. Along the South American coast, the local fishing industry is severely disrupted by the unusually warm ocean waters.

GFDL and Princeton University maintain active collaborative efforts aimed at understanding and predicting El Niño. George Philander of Princeton University is a pioneer in the field of El Niño simulation and theory. His work with GFDL's Ron Pacanowski established the viability of simulating El Niño using coupled ocean-atmosphere models and set the stage for attempts to use such models for El Niño prediction. At GFDL, Kikuro Miyakoda and Tony Rosati have led efforts to build a system to predict both El Niño and its impact on large-scale weather

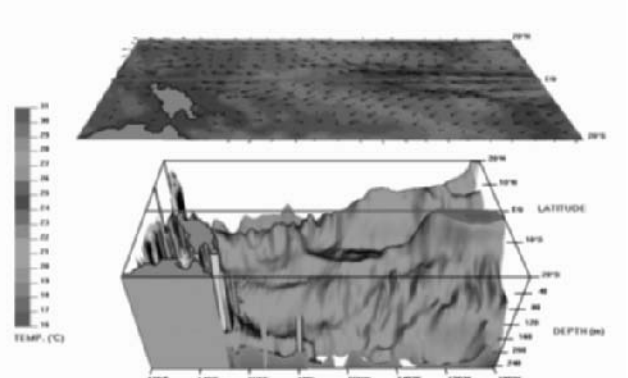

Modeled ocean temperature and surface current distribution over the tropical Pacific Ocean region obtained from a GFDL coupled ocean-atmosphere model used to predict El Niño/Southern Oscillation (ENSO). Lower (blue surface): three-dimensional depiction of the surface on which the ocean temperature is 20 °C. Undulations of this surface can be used to monitor ocean heat content changes associated with ENSO. During the cold phase of ENSO, this surface is deep in the warmer western tropical Pacific but rises toward the sea surface in the cooler eastern equatorial Pacific, as in the example shown. Upper: corresponding distribution of sea surface temperature and surface ocean currents. Gray regions depict land areas of New Guinea and northern Australia.

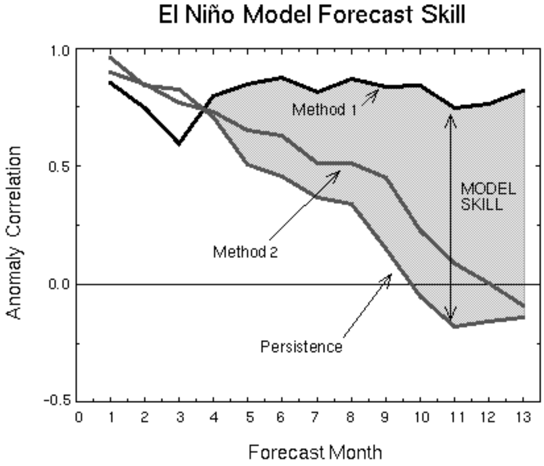

Model skill in the prediction of eastern tropical Pacific sea surface temperature anomalies for different forecast lead times. Higher values indicate greater skill. Based on an ensemble of 14 retrospective forecasts for 1982–1988 using a GFDL coupled ocean-atmosphere model. Forecast Method 1 uses both surface and sub-surface ocean data for initialization, whereas Forecast Method 2 uses only surface data. The results show the crucial importance of subsurface data for the success of the forecasts with this prediction system. [*Source*: Anthony Rosati, et al., Monthly Weather Review, submitted.]

patterns. Miyakoda has long been a leader in efforts to extend the limits of useful weather forecasts. Many of his research group's earlier modeling innovations had been incorporated into operational weather forecasting models at the National Meteorological Center, leading to significant improvements in the 3–5 day forecasts now available to the public.

Using coupled ocean-atmosphere models developed at GFDL and other research centers, scientists now have begun to make physically based predictions of El Niño conditions, with lead times of a year or more. These models have the potential to predict El Niño's effects on weather patterns over North America and other regions far removed from the tropical Pacific. Farmers, energy planners, and water resource managers are examples of those who would benefit from improved extended-range weather and short-term climate predictions.

GFDL'S MODULAR OCEAN MODEL

Detailed predictions of El Niño are made possible by the "Modular Ocean Model," a product of extensive research at GFDL. Through the efforts and foresight of Mike Cox, Keith Dixon, Ron Pacanowski, and Tony Rosati at GFDL, this ocean model is now a shared resource, used by hundreds of researchers worldwide. The model has become the mainstay of the oceanographic modeling community, not just in new operational El Niño predictions at the National Centers for Environmental Prediction, but for global climate change research and other applications as well.

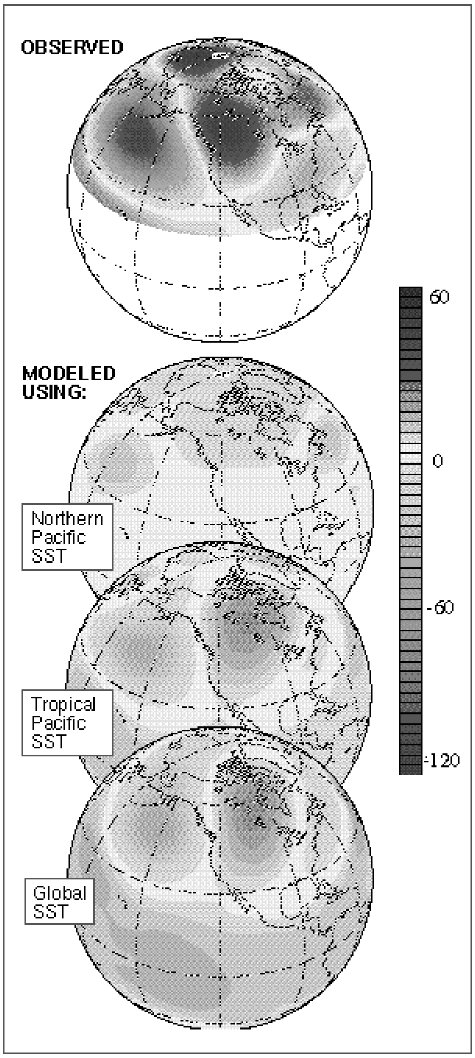

Observed and modeled atmospheric circulation anomalies over the Pacific Ocean and North America during six El Niño winters. The model simulations illustrate how long-range (1 year) predictions of tropical Pacific sea surface temperature may ultimately lead to improved long-range predictions of seasonal weather anomalies, even in regions remote from the tropical Pacific such as the continental United States. Shown are composite 500 mb geopotential height anomalies in meters. The three bottom diagrams are from atmospheric model experiments in which observed sea surface temperatures have been specified for different domains. [*Source*: Ngar-Cheung Lau and Mary Jo Nath, Journal of Climate, August 1994.]

RENEWABLE ENERGIES FROM THE OCEAN

Purnima Jalihal
National Institute of Ocean
Technology
Chennai, India

INTRODUCTION

Today renewable energies are the focus of attention because of the realization that fossil fuels are on the verge of extinction. Renewables like solar and wind energy have been studied thoroughly, and a lot of standard energy conversion devices have also been designed, tested, and proven. The ocean has a large potential in terms of the energy availability. However, harnessing the ocean energy forms is still in the infancy stage.

The oceans have several untapped features like varying temperature, waves, and tides. The mechanical motions and the thermal variation caused by solar radiation cause the oceans to have entrapped in them tremendous amounts of energy, which can be tapped. The energy tapped from these phenomena occurring in the oceans is renewable and therefore needs exploration. Although several attempts have been made to harness these energies and convert them into electricity, none of the methods are yet commercially viable. However, further studies in these areas and performance evaluation of pilot plants can lead to improvements in the economic and technical prospects of such plants. The three energy sources in the oceans are waves, tides, and thermal variation along the depth. This chapter gives a brief description of each type of energy and gives the Indian case studies for wave energy and ocean thermal energy conversion.

WAVE ENERGY

Waves result when wind blows over the ocean surface and energy gets transferred from the wind to the sea. The amount of energy transferred depends on the wind speed, the distance along which it affects the water, and the amount of time during which the wind blows. When a wave breaks at seashore, one can see the tremendous amount of energy released.

Wave energy is an irregular and oscillating low-frequency energy source that can be converted into usable electric energy. It varies from location to location and from season to season.

Harnessing wave energy involves complicated devices in which the mechanical motion of the pressure waves has to be converted to electrical energy. These devices can be generally categorized as follows:

1. Floating or pitching devices
2. Oscillating water columns (OWC)
3. Surge or focusing devices

The first type consists of mechanical linkages between floating or fixed objects. The oscillating water column type of devices transfer the energy from the waves to pneumatic power by alternate compression and decompression of air in a duct or chamber. This pneumatic power is converted to mechanical power and thereby to electric power through a generator. The surge devices channel large waves into a small area, which forces the height of waves to increase, and when this water passes through hydroelectric turbines on its way back to sea level, electricity is generated.

Wave Energy Plants Around the World

In Japan, research on wave energy began during the late 1970s. Several tests were conducted on OWCs, and the biggest focus has been on what is known as the "Mighty Whale" developed by Japan Marine Science and Technology (JAMSTEC). This OWC is probably the world's largest. However, Japan has been active in the area of the floating OWC and a concept called the Backward Bent Ducted Buoy has been developed. This is a floating conventional OWC but with the opening away from the wave direction. Norway has had research being carried out in wave energy for the past 25 years mainly by the Norwegean University of Science and Technology. They have been working on the Tapered Channel and OWC schemes. Portugal also is an active player in wave energy research. The Archimedes wave swing was invented here, and the concept uses the periodic wave pressure to cause the upper part of a buoy to move while the lower part stays in position. The United Kingdom has conducted research on the largest number of varied type of wave energy devices. The Shoreline OWC on Islay, the LIMPET OWC with Wells turbine, and the large OSPREY OWC are all the U.K.'s contributions to wave energy research. Some large prototypes like OSPREY failed before installation was completed. In Sweden, work has been carried out on a floating buoy with a submerged tube underneath. The tube is open to the sea at both ends, and there are some components for the power take off mechanism in the buoy. However, until now, no single plant has been completely commercialized and made repeatable. One must note that the Indian wave energy plant has been functional for the past 10 years, but it generates very low average powers. Details of this plant follow in a later section. Before we move on to the case study, a few details on the nature of waves will be discussed briefly.

Wave Theory

The average power P (W/m) in a regular sine wave per meter wave front of waves with height H and period T can be expressed as

$$P = \rho g^2 H^2 T / 8\pi$$

where ρ is the density of water and g is the acceleration caused by gravity. However, in practice, waves are far from ideal. In nature, waves are irregular and can be described by statistical models. If the wave conditions are measured, over 20 minutes, for example, the mean wave height H_m and the significant wave height H_s can be calculated. The significant wave height is defined as the average of the highest 33% of the waves. Under such circumstances, the wave power can be stated as:

$$P = 0.55 H^2{}_s T_z \text{ kW/m length of wave crest}$$

where T_z is the zero crossing period (1). It is very clear from the above equations that the energy in waves is dependent directly on the wave height and period, which makes it a variable and intermittent source of energy.

CASE STUDY

The Indian Wave Energy program

Indian wave energy research began in 1983 under the sponsorship of the Government's Department of

Figure 1. Wave energy plant.

Ocean Development. After about 6 years of research in the laboratory, the wave energy group and the Ocean Engineering Center, Indian Institute of Technology, Madras, coordinated a project, which involved the fabrication, and construction of the OWC-based wave energy system, which is shown in Fig. 1.

The structure consists of a heavy concrete caisson. The site selected was 45 m in front of the Vizhinjam Fisheries Harbor breakwater, off the Trivandrum Coast at the southern tip of the Indian Peninsula. The water depth at this location is 10 m, typical for fisheries breakwaters, and the average annual wave power potential is of the order of 15 kW/m. The already existing breakwater gives the approach needed for this structure. Figure 2 shows the sectional view of the caisson.

Stages of Research

Several stages of research were carried out on the power module by IIT, Chennai, and subsequently by the National Institute of Ocean Technology, India.

Stage I (1991–1994). The Constant chord Wells turbine coupled to a 150-kW slip ring-type induction generator was used. The major defect found in the system was the stalling behavior of the turbine at higher flow rate. Hence, efforts were made to carry out further studies on turbine and to reduce the no-load losses in the system. The studies resulted in the development of the stage II turbine.

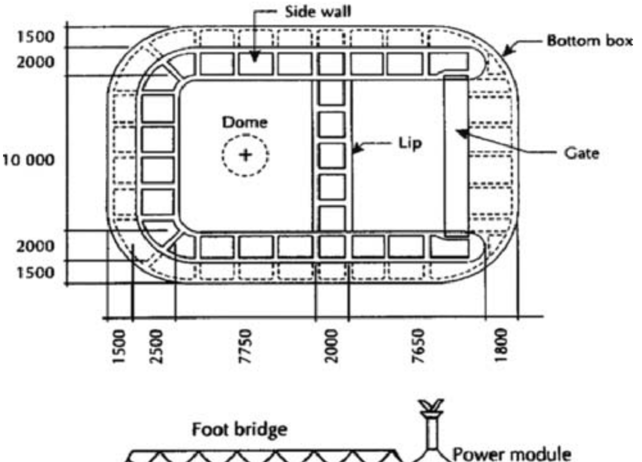

Figure 2. Cross-sectional view and elevation of wave energy device.

Stage II (1994–1997). A varying chord Wells turbine that has characteristics of self-pitch control, which can reduce the stalling effect, was installed in 1995. The data studies made on the system revealed that the efficiency of the turbine has been improved because of the reduced stalling and the selection of proper generator rating.

Stage III (1997–2000). An Impulse turbine with self-governing guide vanes coupled to a 55-kW generator was installed, and the studies revealed that the system was operable for all periods of the year. The problem with the impulse turbine is that the flipping guide vanes deteriorate the efficiency with time because of the corrosion of bearings in the marine environment, and a change in the geometry of vanes can result because of many cycles of operation. Hence, it was decided to change to fixed guide vanes.

Stage IV (2000–Present). A special variable-speed permanent magnet brushless alternator with constant power output is now used along with the impulse turbine with fixed guide vanes. The system can deliver power to charge a battery bank of 300 Ah capacity between 400 rpm and 1000 rpm. The objective of using a battery bank is to deliver a constant power requirement of 7 hp to a desalination system already installed at the site. These studies helped to understand the wave-to-wire conversion process and the pumping of power to the grid. The impulse turbine proved to give the highest efficiency, and this finding has led other wave energy researchers around the world to also consider using impulse turbine in place of wells turbine for wave energy conversion. It was decided to use the average power being generated by the plant to run a reverse osmosis-based desalination plant.

Wave-Powered Desalination System

The wave-powered desalination system consists of the impulse turbine, a special variable-speed alternator, and a battery backup. As the waves are random in nature, the alternator is a device that gives a constant DC voltage with varying input. When the wave heights are low, the turbine speed is low and may not be in the range within which the alternator generates voltage. In this case, the battery discharges and powers up the load. When more power is generated than is required by the load, the batteries are charged. This system was put in place after rigorous computer simulations and laboratory testing. The system is now operational and is now supplying water to the local village community. The system is self-sustaining in that energy from the sea generates freshwater out of seawater.

TIDAL ENERGY

Tides are caused by the gravitational attraction of the moon and the sun acting on the rotation of Earth. The relative motions of these bodies cause the surface of the sea to be raised and lowered periodically, according to several interacting cycles.

The tide moves a huge amount of water twice each day, and if we could harness it, large amounts of energy could be generated. However, converting it into useful electrical power is not easy. Tidal power works in a manner similar to a hydroelectric plant except that the dam is much larger. A huge dam or barrage is built across a river or a bay with a large difference between its low and high tides. When the tide goes in and out, the water flows through tunnels in the dam. The ebb and flow of the tides can turn a turbine. A major drawback of tidal power stations is that they can only generate when the tide is flowing in or out, i.e., only for 10 hours each day. However, some advantages are that there are no wastes like greenhouse gases, and production of electricity is reliable because tides are totally predictable.

Tidal energy potential has been investigated by several countries, notably, France, where a 240-MW demonstration plant was built on the Rance estuary during the 1960s and has now completed 30 years of successful operation. The Russians have built a small 400-kW device near Murmansk, which was later followed by a 17.4-MW experimental device, built by the Canadians at Annapolis on a small inlet off the Bay of Fundy. Small plants have also been installed in China. None of these countries have developed these plants any further. In the United Kingdom, a series of industrial consortia in collaboration with the Government have investigated the prospects for tidal energy on the Severn, Mersey and several smaller estuaries (2).

Every type of renewable energy project has its drawbacks, and tidal power plants have their problems. Obviously these stations would not be effective as the sole supplier of required energy (3). Also, some environmental issues develop with tidal energy. One issue is that the flooding of surrounding land can bring about much destruction. It could mean the relocation of settlement and agriculture plots. It can also change the composition of the habitat and threaten local inhabitants. Vegetation and marine life may not to adapt to the new water levels and may cease to grow or live where they previously existed. The main advantage of tidal energy is that it is nonpolluting. Hence, when designing tidal energy plants, these factors can be built into the design to make it economical.

India is also trying to venture into tidal energy in the Gulf of Khambhat region of Gujarat on the western side of India. Once the Gulf is closed, water levels within the reservoir can be controlled while the tidal fluctuation outside the reservoir continues, and hence, they can be harnessed for the generation of tidal energy.

OCEAN THERMAL ENERGY CONVERSION

The method of harnessing the energy by using the temperature difference between the surface of the ocean and at depths of around 1000 m is called ocean thermal energy conversion (OTEC) and is again a nonpolluting renewable energy source. It is capital intensive, but the unit cost comes down drastically as the rating becomes higher.

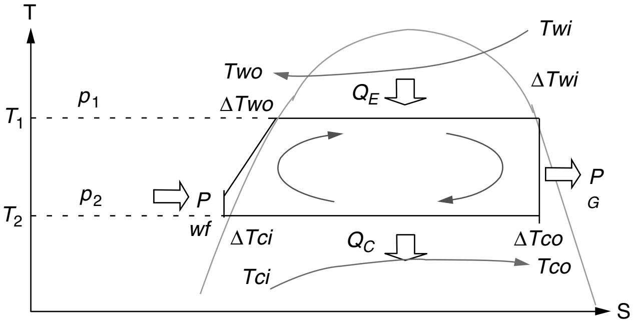

Figure 3. T-S Diagram for the closed Rankine cycle.

OTEC—Principle and Systems

OTEC uses the temperature-difference existing between warm surface seawater of around 27–29 °C and the cold deep-sea water of around 5–7 °C, which is available at a depth of 800 to 1000 m (4). Based on the power cycle, the OTEC system can be classified in to two open cycle and closed cycle systems. In the open cycle system, the warm surface water is flash-evaporated in a chamber maintained under high vacuum and the generated vapor drives a low-pressure turbine connected with the generator. The exhaust steam is condensed with cold seawater taken from the ocean depth. For an open cycle system, a condenser can be a means for freshwater production. The closed cycle system working on Rankine cycle uses a low boiling point liquid like R 134a or ammonia as the working fluid. The T-S diagram for the closed Rankine cycle is shown in Fig. 3.

The fluid is evaporated in a heat exchanger with the warm seawater from the ocean surface. After the vapor is expanded and drives the turbine, it is condensed by cold seawater from the depth of the ocean. This condensate is pumped back to the evaporator and recycled with the working fluid pump. A simplified flow diagram of the OTEC power cycle along with the Rankine cycle is shown below in Fig. 4.

Design Requirements of OTEC Plants

Choice of Working Fluid. The selection of equipment, sizing of piping, working conditions, and so on depend on the working fluid, which should have low-temperature boiling and condensation, high heat transfer coefficient, be environment friendly, safe, nontoxic, and so on. Some working fluids are ammonia, fluro carbon, tetra fluro ethane (HFL-1349), and propane. If safety is the most important issue, HFL-134a is the choice; however, ammonia is less expensive, easily available, and has superior thermodynamic properties. To handle ammonia, certain safety guidelines need to be followed carefully and strictly.

Choice of Material for Heat Exchangers. For an OTEC plant, the heat exchanger material should have corrosion resistance to seawater and compatibility with working

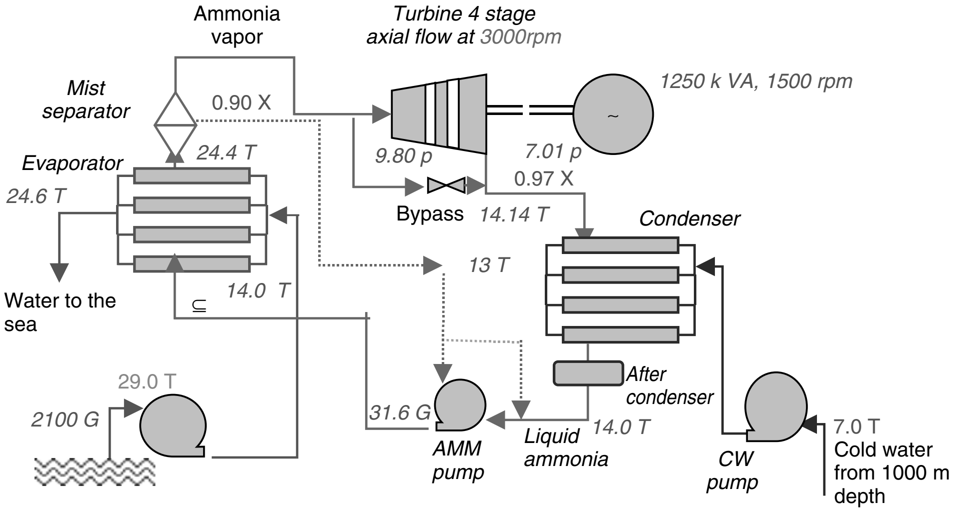

Figure 4. Energy balance diagram for 1-MW OTEC plant.

fluid. Materials that can be considered are aluminum, nickel, stainless steel, and titanium. Combinations like titanium with ammonia and aluminum with propane are workable.

Equipment Selection

Heat Exchangers. Heat exchangers are critical components, and the shell-tube, plate-fin, plate type of exchangers need to be compared. Considering volume required, fouling rates, heat transfer area, pressure drop, maximum working pressure, and other criteria, plate heat exchangers may have an edge over other types.

Turbine. The turbine generator system is another critical component of an OTEC plant, and the net power production depends on the turbine efficiency. Axial flow turbines give better adiabatic efficiency compared with radial flow turbines.

Sea Water Pumps. Seawater pumps for the warm and cold seawater consume about 30–40% of the gross power produced by the plant. The choice of material and characteristics to suit large discharge and low head should be considered for the selection.

Cold Water Pipe Material. The cold water pipe is usually around 1 km long and is subjected to varying hydrostatic pressures, temperatures, and forces caused by current drag and waves. In addition to the conventional steel pipes, other materials that can be considered are concrete, high-density polyethylene (HDPE), glass reinforced plastics (GRP), and so on. HDPE has some additional advantages in that it is buoyant and can float in water, which helps during towing and installation.

Other Criteria to be Considered

Station Keeping Systems. For floating OTEC plants, either the platform should be self-propelled or moored offshore. Design of the mooring depends on various environmental factors and the life expectancy of the plant.

Biofouling Control. In warm seawater, any solid surface begins to acquire a film in the form of a slime layer in a few hours. This process is called fouling, and it affects the heat exchanger performance when it grows within. It also increases drag on the pipes and vessels if not cleared for extended periods of time.

Some means of regular cleaning and addition of antifoulants like chlorine are measures to be taken to reduce the effects.

Power Transmission. For offshore plants, the power generated needs to be transmitted to shore and this requires underwater cables, which should have minimum transmission losses, resistance to water pressure, and such other characteristics. These requirements and more need to be catered to for designing an OTEC plant.

The most complex part of setting up a floating OTEC plant offshore is the deployment of the long cold water pipe and its connection to the barge. A HDPE pipe is buoyant, and it floats easily when towed on the surface. However, the lowering and the upending of the pipe requires several craft and tugs with cranes and other handling equipment. The mooring system should be designed well for operating as well as for storm conditions.

Case Study: The Indian 1-MW OTEC Plant

The National Institute of Ocean Technology is currently in the process of installing a floating barge-mounted OTEC plant off the Southern coast of India of gross rating 1 MW. A 1-km-long pipe made of high-density polyethylene will draw the cold water. The heat exchangers are of plate type and made of titanium. The turbine is a 1-MW ammonia turbine designed and fabricated for the first time. The working fluid is ammonia, and the Rankine cycle is being used. The power plant is housed on a non-self-propelled barge, which needs to be moored in 1000-m water depth. The configuration of the pipe/mooring system is shown in Fig. 5.

The motions of the barge under the environmental forces caused by wind, wave, and current affect the pipe and mooring behavior. Several analyses have been carried

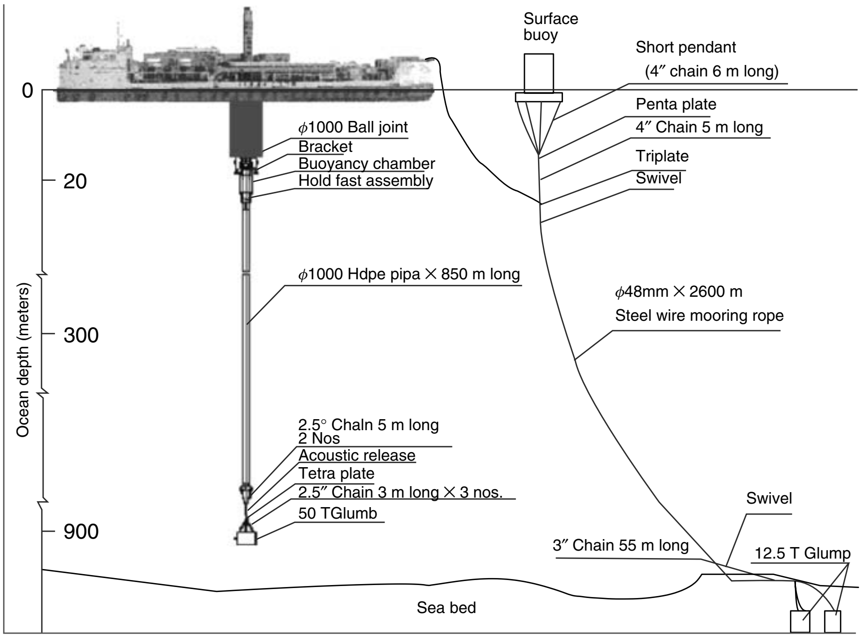

Figure 5. Pipe mooring configuration for 1-MW OTEC plant.

out to study the behavior of the system and arrive at the final design. The barge was specially designed for the OTEC plant, and some special features have been introduced in it, such as:

- A large 14-m-high retractable sump supporting the cold water pumps to meet the pump net positive suction head (NPSH) requirements.
- A quick-release mechanism to disconnect the mooring from the barge.
- A quick-release mechanism to disconnect the cold water pipe.

When commissioned, this will be the first megawatt-rating floating OTEC plant in the world.

CONCLUSIONS

Harnessing ocean energies in the various forms described in this article is challenging and needs several engineering and technological innovations. Most methods are still under study. However, with a focus toward overcoming the engineering complexities, a day will dawn when these renewable sources of energy will be harnessed and will replace the fossil fuels. The oceans are vast storehouses of energy, and man's endeavors and ingenuity will capture these ocean energy forms in such a way as to be environment friendly and convert energy-starved areas in the world to areas self-sufficient in energy.

Acknowledgment
The author expresses her gratitude to Ms. Sharmila N and Ms. Sri Lakshmi I of the National Institute of Ocean Technology, for helping with information gathering and word processing of this work.

BIBLIOGRAPHY

1. Shaw, R. (1982). *Wave Energy a Design Challenge*. Ellis Horwood Publishers, New York.
2. Available: http://europa.eu.int/comm/energy_transport/atlas/htmlu/tidalintro.html.
3. Available: http://ririessortment.com/tidal energy-rdrx.htm.
4. Avery, W.H. (1994). *Renewable Energy from the Ocean*. Oxford University Press, Oxford, UK.

ESTUARIAN WATERS

Priyanka K. Dissanayake
Moratuwa, Sri Lanka

IMPORTANCE OF ESTUARIES

Estuaries are in many ways the most complicated and variable of aquatic ecosystems. An estuary is a place where salt water from an ocean mixes with freshwater from the land and creates a unique and special environment in which marine species live, feed, and reproduce. Estuaries and intertidal regions are naturally exposed to stressful conditions.

An estuary and its immediate catchments form a complex system of ecological, physical, chemical, and social processes that interact in a highly involved and, at times, unexpected fashion. Estuaries are the end point for the detrimental effects of many upstream catchment activities. Materials carried from the land by rivers vary in quality and quantity, sometimes with strong seasonal patterns of high biological significance. Tidal oscillations cause vigorous reversals of flow. Inherent hydrographic patterns can lead to accumulation of materials and to upstream transport from the point of addition. The estuaries as recipients of wastes from rivers entering them and cities and industries along their shores are obviously more immediately susceptible to pollution damage than any other part of the marine system. Increased loads of nutrients and sediments from activities in the surrounding catchments are the two key threats to marine and estuarine water quality.

Estuaries are often used as disposal sites for unwanted biological or industrial products, usually referred to as waste. If such waste enters the estuarine ecosystem and is potentially harmful to life or has potentially detrimental effects on its health it is usually called a pollutant.

The usual way of controlling negative impacts of pollutant dispersal in estuaries is by monitoring its concentration in the water and making sure that the concentration does not exceed a certain level above which the pollutant is considered harmful. This problem consists of two parts, the *near field* and the *far field*. The near field problem considers technical details of pollutant outfall design and ways to inject the pollutant into the estuary from a point source so that its concentration is reduced as rapidly as possible. This is an engineering task, and therefore, the near field problem is not considered in oceanography. The far field problem studies the distribution of the pollutant through the entire estuary, starting from the disposal site not as a point source but as an extended source of uniform concentration across the estuary depth and width. The far field problem thus assumes that the engineering design of the pollutant outfall achieves reasonably uniform distribution over a sizeable section of the estuary and uses this situation as the starting point for the analysis.

PHYSICOCHEMICAL PROCESSES IN ESTUARIES

If the concentration of a pollutant where it is introduced into the estuary is known, its distribution through the estuary can be predicted from theoretical considerations. To understand how this is done, it is useful to begin with a look at the distribution of salt and freshwater in the estuary. Let us consider salt as a “pollutant” and follow its path through the estuary. Its source is at the estuary mouth, where its concentration is given by the oceanic salinity S_0. The salt enters the estuary with the net water movement of the lower layer. As it moves up the estuary, its concentration decreases through turbulent diffusion; in other words, the salt “pollutant” is diluted with freshwater that comes down from the river. We can normalize the salt “pollutant” distribution along the estuary by making it independent of its original “concentration” S_0 and define the salt–water fraction s as $s = S/S_0$, where S is the local depth averaged salinity at any arbitrary location in the estuary. This normalized salt concentration or saltwater fraction decreases from 1 at the mouth to 0 at the inner end of the estuary.

Similarly, we can consider the freshwater introduced by the river as a “pollutant” and follow its path through the estuary. Its source is at the inner end of the estuary, where its normalized concentration or freshwater fraction f is equal to 1. The freshwater then moves down through the estuary with the net water movement in the upper layer and reaches a concentration of zero at the estuary mouth. The estuary contains only saltwater and freshwater, so the saltwater fraction and the freshwater fraction always add up to 1 locally, which means that $f + s = 1$.

Not all pollutants are introduced into the estuary through the river or from the ocean. There are three classes of marine pollutants. Conservative pollutants are substances that are inert in the marine environment; their concentrations change only as a result of turbulent diffusion. (Salt and freshwater can be considered such substances.) Nonconservative pollutants undergo natural decay; their concentration depends on turbulent diffusion and also on the time elapsed since their introduction into the environment. The concentration of coupled nonconservative pollutants depends on turbulent diffusion and natural decay, both of which act to decrease their concentration, and on the availability of other substances in the environment, which allow the concentration to increase over time.

Conservative Pollutants

A conservative pollutant released at a point between the inner end and the mouth of the estuary spreads by turbulent diffusion in both directions, downstream and upstream. Its diffusive behavior is no different from the diffusive behavior of salt or freshwater. In the upstream direction, it diffuses in the same way as the salt water diffuses upstream; in the downstream direction, it follows the freshwater diffusion. If its concentration at the outlet is c_{out}, its concentration C in the estuary upstream from the release point is thus proportional to the salt fraction; downstream from the release point, it is proportional to the freshwater fraction:

$$\text{Upstream:}\quad C = c_{\text{out}}(S/S_{\text{out}})$$

$$\text{Downstream:}\quad C = c_{\text{out}}(f/f_{\text{out}})$$

Here, S_{out} is the vertically averaged salinity at the outlet location, and f_{out} is the freshwater fraction at the outlet. This shows that it is possible to predict the distribution of a pollutant in the entire estuary if the salinity distribution is known. The equation assumes steady-state conditions, continuous release of the pollutant at the constant concentration c_{out}.

Nonconservative Pollutants

The concentration of nonconservative pollutants decreases even in the absence of diffusion, through either biochemical or geochemical reaction. An example is the concentration of coliform bacteria released from a sewage outlet. The relationship between the concentration C of such a pollutant, the salinity S, and the fresh water fraction f is not as straightforward as for conservative pollutants, but concentration levels are always lower than those derived from the conservative case.

The concentration can be estimated by subdividing the estuary into compartments. If the compartments are chosen so that the ratio r of the freshwater volume V_{f} in the compartment and the volume of freshwater R introduced by the river into the estuary in one tidal cycle is constant ($r = V_f/R = \text{constant}$), it can be shown that the concentration of a nonconservative pollutant can be approximated by an equation for the concentration in each compartment:

$$\text{Upstream:}\quad C_p = C_{out}\frac{S}{S_{out}}\left[\frac{r}{1-(1-r)e^{-kT}}\right]^{n+1-p}$$

$$\text{Downstream:}\quad C_p = C_{out}\frac{f}{f_{out}}\left[\frac{r}{1-(1-r)e^{-kT}}\right]^{p+1-n}$$

In this equation, compartments are numbered from the inner end of the estuary, n is the compartment containing the outfall, and p is the compartment where the concentration is evaluated. T is the tidal period, and k is the decay constant for the pollutant. The larger k, the faster the decrease of the concentration over time. $k = 0$ represents the case where there is no independent decrease without diffusion, or the case of a conservative pollutant. Figure 1 compares this behavior with that of a conservative pollutant. Note that for $k = 0$, the equation for the concentration reduces to

$$\text{Upstream:}\quad C = c_{uot}(S/S_{out})$$

$$\text{Downstream:}\quad C = c_{out}(f/f_{out})$$

which is the same equation that was derived for the conservative pollutant.

Coupled Nonconservative Pollutants

Biochemical or geochemical processes reduce the concentration of a nonconservative pollutant, and they often also give rise to an increase in the concentration of another substance, which in turn may produce a third substance as a result of its own decay, and so on. Whether this chain of events constitutes pollution or not depends on the degree of harmfulness of each substance. Ammonium nitrogen, for example, is used as a fertilizer on land. Under natural conditions, it converts to nitrate (Fig. 2), which is also a nutrient.

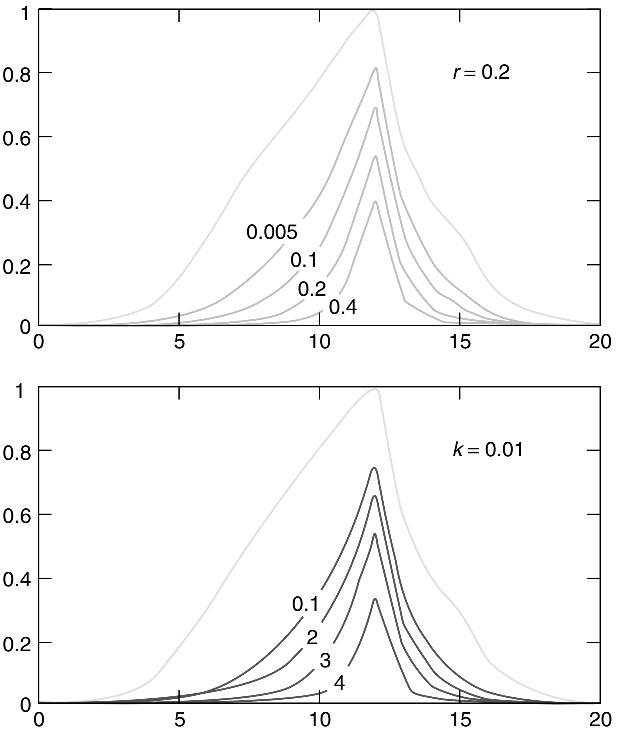

Figure 1. Behavior of conservative and nonconservative pollutants.

The introduction of nutrients into the marine environment does not automatically constitute pollution. But when nitrate and other nutrients are again mineralized in seawater, this process requires oxygen, which is not in unlimited supply in the ocean. Too much nutrient can lead to such a reduction of oxygen levels that the lack of oxygen can become a threat to marine life.

The indicators of nutrients and sediments are influenced by human activities from both diffuse inputs (urbanization and agriculture) and point discharges to the environment. The recommended indicators need to be biologically meaningful, yet readily collected and interpreted in relation to ecological conditions.

ROLE OF NUTRIENTS IN ESTUARIES

Nutrient Status

Six indicators of nutrients are recommended for assessing the status of marine and estuarine waters: dissolved inorganic nitrogen (DIN), total nitrogen (TN), dissolved inorganic phosphorous (DIP), total phosphorous (TP), chlorophyll a (Chl a), and dissolved oxygen (DO). Individual indicators have their own limitations, but when all six are used together, they provide a strong indication of nutrient status.

1. *DIN.* DIN is the combined concentration of nitrate, nitrite, and ammonia. DIN is the bioavailable fraction of the nitrogen in the water column.
2. *TN.* DIN may rapidly disappear as algae assimilate it, and so TN provides a better indication of the nitrogen cycling through the algal community and how much is dissolved in the water. TN is the sum of the concentrations of bioavailable (dissolved) and particulate forms of inorganic and organic nitrogen.

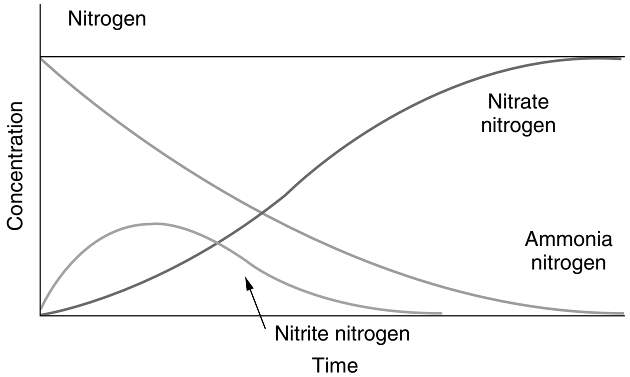

Figure 2. Behavior of a coupled nonconservative pollutant.

3. *DIP*. DIP is typically measured as the concentration of reactive orthophosphates, which is the form of phosphorous most readily bioavailable to aquatic plants.
4. *TP*. Like DIN, algae may rapidly assimilate DIP and so TP provides a better indication of the phosphorous cycling through the algal community and how much is dissolved in the water. TP is the sum of the concentrations of dissolved and particulate forms of inorganic and organic phosphorous.
5. *Chl a*. Chl a is a crude measure of phytoplankton biomass. Apart from light, phytoplankton are primarily limited by nutrient availability; Chl a is often used as an integrated, surrogate measure of nutrient status.
6. *DO*. DO, measured preferably as percent saturation (%), is the net balance of the amount of oxygen produced by photosynthesis, relative to that consumed by respiration and other forms of biological and chemical oxygen demand. Low DO (hypoxia) can arise following blooms of phytoplankton stimulated by nutrient enrichment. As the large mass of plant tissue decays, DO is lowered in the water column, and this can result in fish kills and nutrient inputs from the sediments.

Nutrients That Are Limiting Factors for Estuarine Phytoplankton

A limiting nutrient is a nutrient or trace element that is essential for phytoplankton to grow but is available in quantities smaller than required by the plants or algae to increase in abundance. Therefore, if more of a limiting nutrient is added to an aquatic ecosystem, larger algal populations develop until nutrient limitation or another environmental factor (e.g., light or water temperature) curtails production, although at a higher threshold than previously.

It is often said that nitrogen is the limiting nutrient in marine and coastal waters; however, the general assumption is often incorrect (1). Phosphorous, silica, and iron can also limit production in marine and coastal waters, and different trophic groups within the same ecosystem can be limited by different elements and nutrients (1).

Redfield Ratios. The recommended ratios of nutrients in estuaries are given by Redfield ratios, which are C:N:P = 106:16:1 for algae and Si:N:P = 20:106:16:1 for diatoms (1). Recently, Harris has suggested that a C:N:P ratio of 166:20:1 is probably more representative of the average phytoplankton composition (1). However, a review of the literature by Hecky and Kilham showed that the optimum N:P ratios for 14 freshwater and marine phytoplankton, for which data were available, ranged from 7–87 (1).

If the concentration of Si or N relative to P falls below these ratios (N:P<16:1, Si:P<20:1, these may be limiting nutrients (carbon is never limiting because it is available as a gas (CO_2); carbon in organic matter comes from dissolved CO_2 in water. The nitrogen levels control the rate of primary production. If the system is provided with high levels of nitrogen, algal blooms occur. Systems may be phosphorous limited, however, or become so when nitrogen concentrations are high and N:P>20:1(1). In such cases, excess phosphorous triggers eutrophic conditions. For optimum alagal growth, the N:P ratio should be in the range of 16–20:1(1). Increased reliance on the N:P ratio as a management tool is widely accepted by marine scientists and engineers to control nuisance algal blooms.

Another limitation on using N:P ratios is the considerable uncertainty over what measure of nitrogen and phosphorous should be used in calculating the N:P ratio. The ratio of TN:TP generally overestimates the algal available nitrogen to algal available phosphorous ratio aN:aP because organic phosphorous is more readily available than organic nitrogen and DIN:DRP underestimates aN:aP because DRP includes some organic phosphorous compounds in addition to orthophosphates (1). Therefore, the ratio of DIN:DIP gives a better estimate of the N:P ratio (1).

The recommended levels of total phosphorous in estuaries and coastal ecosystems to avoid algal blooms is 0.01–0.10 mg/*l* and 0.1–1 mg/l nitrogen (a 10:1 ratio of N:P) (1). It is recommended that total phosphorous levels in waterbodies should range within −0.005–0.15 mg/L to prevent eutrophication in estuarine waters; guidelines for protecting aquatic ecosystems recommend that phosphate values should not exceed 0.015 mg/L in estuaries (1).

ANZECC guidelines recommend the following concentration ranges to protect aquatic ecosystems: total N range from 0.01–0.10 mg/L for estuaries and 0.01–0.06 mg/L for coastal waters. The nitrate value should not exceed 0.1 mg/L for estuaries (1).

Regeneration of Nutrients (cycling)

Nitrogen Cycle in An Estuary (Fig. 3)

Phosphorous Cycle in An Estuary

$$\text{Organic P} \rightarrow \text{Inorganic P} \rightarrow \text{Organic P}$$

BUDGETARY MODELS FOR ESTUARIES

In science, models are tools that help us conceptualize, integrate, and generalize knowledge. “Budget models” are simple mass balance calculations of specific variables (such as water, salt, nitrogen, phosphorous, etc.) within defined geographic areas and defined periods.

The fluxes of materials to and from a system may be budgeted by many different procedures, but there are inherent similarities among these procedures. A budget describes the rate of material delivery to the system (‘inputs”), the rate of material removal from the system (“outputs”), and the rate of change of material mass within the system (“storage”). Some materials may undergo internal transformations. Such changes are sometimes referred to as “internal sources and sinks.” Figure 4 depicts the budget for any material.

Without further interpretation, budgets do not provide information on the processes, which account for the summed sources minus sinks (2). Some inferences about the actual processes can be drawn by comparisons among

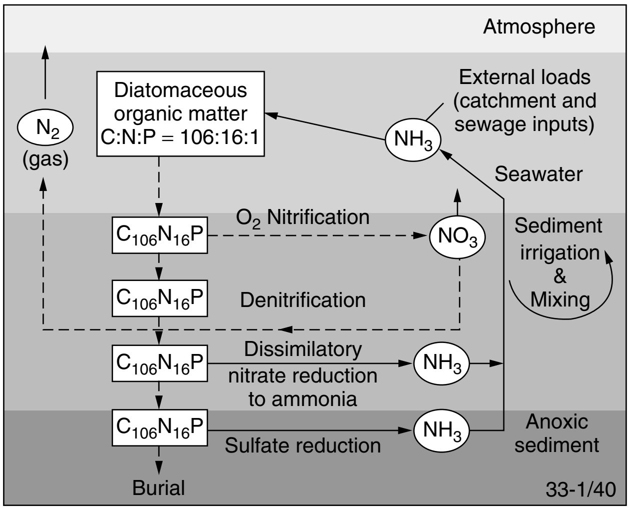

Figure 3. Nitrogen cycle in an estuary.

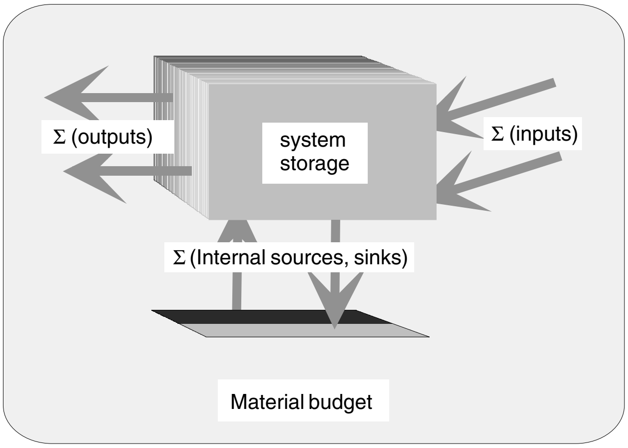

Figure 4. Budget for any material.

budgets for different materials. This approach is referred to as "stoichiometric analysis" of the fluxes.

It is also useful to describe such a budget in terms of this simple equation:

$$\frac{dM}{dt} = \sum \text{inputs} - \sum \text{outputs} + \sum (\text{sources} - \sin \text{ks})$$

where dM/dt represents the change in mass of any particular material in the system with respect to time (2).

In broad summary, water and salt budgets are used to estimate water exchange in coastal systems. Nutrient budgets (as a minimum, dissolved inorganic phosphorus and dissolved inorganic nitrogen) are also developed, and departure of the nutrient budgets from conservative behavior is a measure of net system biogeochemical fluxes.

Stoichiometric Analysis of Fluxes

In chemistry, "stoichiometry" is the study of the combination of elements in chemical reactions. Therefore, we use the term "stoichiometric analysis" to describe this approach to budgetary analysis. It should be kept in mind that the discussion to follow is based on the ΔY's for C, N, and P, as derived from nutrient budgets. The following is generally computed using the ΔY's.

- Net ecosystem metabolism, that is, the difference between primary production and respiration
- Net nitrogen fixation minus denitrification
- Net sulfate reduction
- Net $CaCO_3$ precipitation
- Net CO_2 gas flux across the air−sea interface.

NONCONSERVATIVE DISSOLVED INORGANIC P AND N FLUXES IN NEGOMBO ESTUARY, SRI LANKA, USING THE STANDARD LAND–OCEAN INTERACTION IN THE COASTAL ZONE (LOICZ) METHODOLOGY (3)

The Negombo estuary (7°4′–7° 12′ N: 79°47′–79°51′ E) is situated on the west coast of Sri Lanka. The estuary receives water from the Attanagalu Oya (ja-Ela and Dandugam Oya) drainage basin and is a dominant morphological feature of the watershed. The brackish water area is 3239 ha and is considered the estuarine part of the contiguous wetland system of the Muthurajawela Marsh Negombo Estuary. The main freshwater source, Attanagalu Oya, empties at Ja-Ela and Dandugam Oya at the southern tip of the estuary. In addition, the Hamilton canal is the connecting watercourse of the Kelani Estuary and the Negombo Estuary running parallel to the west coast from the north to the south along the Muthurajawela Marsh.

LOICZ, a one-box, one-layer model was used for this study, and the data available are applied on a seasonal basis (dry and wet). The data used are mentioned below and are the averages for dry and wet seasons:

1. river discharge
2. precipitation
3. evaporation
4. average salinity outside the system, within the system, and in the river discharge
5. dissolved inorganic phosphate (DIP) outside the system, within the system, and in the river discharge
6. dissolved inorganic nitrogen (DIN) outside the system, within the system, and in the river discharge

The nonconservative DIN and DIP fluxes are given in (Table 1).

$$[p - r] = -\Delta \text{DIP} \times \left(\frac{C}{P}\right)_{\text{part}}$$

$$\left[\text{nfix} - \text{denit}\right] = \Delta N_{\text{obs}} - \Delta N_{\text{exp}}$$

$$\left(\text{nfix} - \text{denit}\right) = \Delta N - \Delta P \times (N:P)_{\text{part}}$$

The system is an internal source for DIP and an internal sink for DIN in both the wet and dry seasons. In the dry season, the N:P ratio is less than the Redfield ratio of 16:1; therefore, the estuary is N limiting. In the wet season, the N:P ratio is greater than the Redfield ratio

Table 1. Nonconservative Dissolved Inorganic P and N Fluxes in Negombo Estuary (Sri Lanka) Considering a one–box, one Layer System[a]

Process	Dry Season Rate, (mmolm^{-2}d^{-1})	Wet season Rate, (mmolm^{-2}d^{-1})
Area	32.39 km^2	32.39 km^2
DDIP	+0.052	+0.008
DDIN	−0.104	−0.837
DIN:DIP	13	18
Stoichiometric Analysis		
(nfix-denit)phytoplankton[b]	−0.936	−0.965
(*p*-*r*) phytoplankton[c]	−5.512	−0.848
(nfix-denit)sea grass[b]	−1.664	−1.077
(*p*-*r*)sea grass[c]	−28.6	−4.4
nfix-denit)mangrove[b]	−0.676	−0.925
(*p*-*r*)mangrove[c]	−67.6	−10.4

[a] Reference (3).

[b] $(N{:}P)_{part}$ assumed to be 16 (plankton), 11(mangrove), and 30 (sea grass).

[c] $(C{:}P)_{part}$assumed to be 106(plankton), 1300 (mangrove), and 550 (sea grass).

of 16:1 but less than 20:1; therefore, the estuary has optimum nutrient concentrations, which support optimum phytoplankton growth.

The DDIP>0 in both seasons. This system with DDIP>0 is producing DIC via net respiration $(p - r)$ and is heterotrophic. The negative signs of $(p - r)$ calculated for phytoplankton, sea grass, and mangrove support this assumption if phytoplankton, sea grass, and mangrove dominate the input of organic matter from outside the system respectively. The (nfix−denit) is negative for phytoplankton as well as for sea grass and mangroves, indicating that denitrification is the main process taking place. The denitrificate rates are in the range of −0.676 to −1.664 for the dry season and −0.925 to −1.077 in the wet season.

Nitrogen is typically the nutrient that controls primary production in marine coastal ecosystems and, as such, is commonly implicated in the eutrophication of coastal waters. Denitrification is probably the most important nitrogen cycling pathway because it is one of the few natural processes that can counteract eutrophication. Up to 60 to 80% of the external nitrogen load delivered to coastal ecosystems may be lost to coupled sediment nitrification−denitrification. Despite its importance, sediment denitrification has rarely been measured in Sri Lankan coastal ecosystems.

Worldwide denitrification studies in tropical and subtropical estuaries where N_2 fluxes have been directly measured are also rare; most work has been carried out in temperate West European and North American systems. There are distinct biogeochemical differences between tropical/subtropical and temperate estuaries. Differences in tropical/subtropical estuaries most likely to influence denitrification rates include

generally low water column nitrogen concentrations,
different temperature and light regimes,
differences in benthic infauna,
shallower, so benthic processes that interact with denitrification such as benthic productivity are more important, and
episodically driven resulting in differences in the quality, timing, and delivery rates of carbon inputs to the sediments.

WATER QUALITY GOALS

In developing an effective strategy for mitigating the effects of nutrient overenrichment, one must understand the physical and ecological relationships that determine the extent and causes of nutrient overenrichment, along with societal objectives and behavioral responses. Societal objectives determine goals that a management strategy strives to achieve and is the bench mark against which it is evaluated.

The appropriate set of policies for any given estuary depends on the nutrient sources for that estuary. For example, if the main nutrient source is agriculture, a set of policies designed to promote the adoption of best management practices is required. These can be implemented at the local, regional, or national level. Alternatively, if atmospheric deposition is the main source of nutrients, policies that reduce atmospheric emissions of nitrogen are needed. The source of atmospheric nitrogen is often outside the local jurisdiction governing the estuary, so policies to combat this nutrient source must be implemented at the regional or national level.

Establishing Criteria and Standards

A water quality criterion is that concentration, quality, or intensive measure that, if achieved or maintained, allows or makes possible a specific water use. Water quality criteria are often the starting point in deriving standards, but criteria do not have a direct regulatory impact because they relate to the effect of pollution rather than its causes. A water quality standard is the translation of a water quality criterion into a legally enforceable ambient concentration, mass discharge, or effluent limitation expressed as a definite rule, measure, or limit for a particular water quality parameter. A standard may or may not be based on a criterion.

Traditionally, water quality standards have been absolute numbers as a concentration or discharge of a toxic substance that may not be exceeded or an oxygen concentration that must be maintained. However, water quality in a given waterbody can fluctuate as a result of random factors, such as weather and uncertainties in hydrologic processes. Eutrophication impacts include increased primary productivity; increased phytoplankton biomass; reduction in water clarity; increased incidents of low oxygen; changes in trophic structure, trophic interactions, and trophodynamics of phytoplankton, zooplankton, and benthic communities; damage to coral reefs; fish kills; reduced fisheries production; and decreased biotic diversity.

In general, the public must accept the need for improved receiving water quality and prevention measures. If the public does not perceive a problem, it is unlikely that elected officials will pursue this issue or that agency staff will have the resources or authority to implement solutions.

When setting up water quality standards, many other factors must also be considered (4)

- Are damages subject to threshold effects or are they continuous?
- During which season are impacts the greatest?
- Are sudden discharges controlled?
- What are the flushing and mixing conditions in the estuary?
- What timescales are involved?

BIBLIOGRAPHY

1. ANZECC/ARMCANZ. (October 2000). *Australian and New Zealand Guidelines for Fresh and Marine Water Quality*. Available: http://www.ea.gov.au/water/quality/nwqms/#quality.
2. Gorden, D.C., Jr., Bondrean, P.R., Mann, K.H., Ong, J.E., Silnert, W.L., Smith, S.V., Wattayakern, G., Wulff, F., and Yangi, T. (1996). *LOICZ Biogeochemical Modelling Guidelines*. LOICZ/R & S/95-5, vit p. 96 LOICZ, Texel, the Netherlands.
3. Dissanayake, U.A.P.K. (2003). *Strategies for the Improved Management of the Coastal Zone of Sri Lanka*. M.Sc. Thesis Environmental Management, University of Moratuwa, Sri Lanka.
4. Committee on the Causes and Management of Eutrophication. (2000). Ocean Studies Board, Water Science and Technology Board, National Research Council. *Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution*. National Academy Press, Washington, DC.

NOS/NMFS COOPERATIVE RESEARCH ON COASTAL FISHERIES AND HABITATS AT THE BEAUFORT LABORATORY

Jon Hare
NOAA Center for Coastal
Fisheries and Habitat Research

EARLY LIFE HISTORY OF FISHES

Laboratory researchers have studied fish early life history since the early 1900's. Work is conducted on egg, larval and juvenile stages of species ranging from anadromous shad to highly migratory tuna and billfishes. Topics of study include taxonomy, distribution, growth, predation, physical transport, behavior and stock identification.

Research on early life history stages at the Laboratory is supported by a fish rearing facility. A variety of species have been reared ranging from summer flounder, to black sea bass, to spotted sea trout. Fish from this facility are used in ecological, physiological, and behavioral studies.

FISHERIES OCEANOGRAPHY

Laboratory staff have been active in fisheries oceanography since the late-1970's. The goal of this research is to better understand the linkages between biological and physical oceanographic processes and fish population distribution, abundance and dynamics.

Recently, the Laboratory has been involved in the South Atlantic Bight Recruitment Experiment (SABRE), the purpose of which is to understand the effect of environmental variation on young menhaden survival, and ultimately, recruitment to the fishery. Employees have conducted egg and larval investigations on the continental shelf, have studied ingressing larvae at inlets, and have examined juveniles in estuarine nursery habitats. Coupling field observations, behavioral models and physical models has provided insights into the role that physical transport plays in controlling larval supply to inlets.

Current research in fisheries oceanography ranges from the effect of Langmuir circulation on plankton distributions, to the exchange of larval fishes across the Cape Hatteras zoogeographic boundary, to the role of the Charleston Bump as essential spawning and nursery habitat. The scales studied range from meters to hundreds of kilometers and from hours to decades, thereby allowing Laboratory researchers to link processes-oriented work with studies examining annual and decadal variation in fish populations.

HARMFUL ALGAL BLOOMS

Laboratory employees have been studying the "red tide" dinoflagellate, *Gymnodinium breve*, since a bloom in coastal North Carolina waters in 1987 caused both fish kills and shellfish closures. The source of *G. breve* cells is the shelf off the west coast of Florida, where blooms have occurred in 21 of the last 22 years. Local eddy circulation in the northeastern Gulf of Mexico and in the vicinity of the Dry Tortugas affects the retention and coastal distribution of blooms on the western Florida shelf as well as influences the entrainment of cells into the Loop Current-Florida Current-Gulf Stream systems. All blooms off the Florida Keys and in the South Atlantic Bight can be traced to the inoculation of shelf water with Gulf Stream water, as well as the proper physical conditions that support blooms. Monitoring and research of *G. breve* is continuing, as is work with other species that potentially cause harmful algal blooms.

The Laboratory's remote sensing capabilities contribute to the study of harmful algal blooms. The Laboratory serves as the Southeast Node of NOAA's CoastWatch program and researchers at the Laboratory are involved in the development of regional ocean color algorithms for the SeaWIFS satellite.

FLORIDA BAY-EVERGLADES NATIONAL PARK

Laboratory scientists initiated fisheries and habitat research in Florida Bay and adjacent waters in 1983. Past research has examined a variety of topics including larval fish distributions, ecology of juvenile spotted seatrout and gray snapper, evaluation of habitat use by juvenile fishes and invertebrates, nutrient limitation of seagrass growth and responses of both plant and animal communities to the die-off of seagrasses first identified in 1987.

Building on this past work, investigators at the Laboratory are conducting laboratory and field research as part of the Florida Bay Study. The overall goal of this work is to evaluate the relation between environmental and habitat change and the recruitment, growth and survivorship for fishery organisms in Florida Bay. Current research includes: examination of the influence of environment and habitat on larval and juvenile fish growth and survival; development of an individual-based energetic model for larval spotted seatrout; evaluation of the effect of habitat on the distribution of fishes and invertebrates and determination of the origin of elevated mercury levels in upper trophic levels.

SEAGRASS RESTORATION

Research on various aspects of seagrass ecology at the Laboratory has been continuous for 20 years. Areas of investigation have included development and dissemination of planting techniques monitoring protocols, and success criteria, as well as studies to determine the light requirements of seagrasses, functional equivalency of restored beds as compared to natural, undisturbed systems, and landscape scale studies regarding the temporal dynamics of seagrass bed pattern and distribution. Emphasis has been placed on research information transfer to managers, active participation in research projects, and litigation. The research approach has been to sustain a broad-based program covering a variety of ecological processes which allows the scientists to quickly adapt and respond to changing management concerns and issues.

MANAGEMENT AND RESEARCH USES OF GIS

Laboratory employees use Geographic Information Systems for various research and management activities. Menhaden researchers are examining the spatial and environmental aspects of menhaden catch and marine mammal researchers are studying the spatial interaction between fisheries and marine mammals.

As a more in depth example, researchers at the Laboratory are using GIS to map and to understand the distribution of aquatic seagrasses throughout the southeastern United States. Aquatic seagrass beds are important habitat to many fish and shellfish species. Linkage with the NOAA Coastal Assessment Framework associates the aquatic beds with coastal drainage basins. Within North Carolina, the vast majority of seagrass beds are inside the barrier islands of the Pamlico-Albemarle Estuarine Complex. Distribution of seagrass beds is related to annual extremes of salinity but other, more-local factors are also involved. The restricted abundance of seagrass beds along mainland shores reflects salinity stress, water quality, and water use problems from coastal development. These large-scale studies of seagrass are being linked with the smaller scale seagrass dynamics studies conducted by Laboratory employees.

MENHADEN RESEARCH

Scientists at the Beaufort Laboratory have been studying Atlantic menhaden since the 1950's and Gulf menhaden since the 1960's. In 1996, catch of both species totaled 1.8 million pounds and was valued at $94.2 million. Research has included studies of life history, movements, juvenile abundance, and patterns in the fishery. Current efforts include examination of the environmental factors that contribute to recruitment variability and continuation of the monitoring of fishery patterns.

Beginning in the late 1970's, menhaden boasts were asked to complete logbooks of daily purse-seine activities. For each purse-seine set, time, location, catch, and weather conditions are enumerated. Approximately 6,800 forms describing 23,000 purse seine sets are collected annually. Laboratory personnel also collect data on landings, fishing effort, and size and age of catch. These data are used for stock assessments, are helpful to industry personnel and are used by researchers to address various scientific questions. For example, these data are being used to examine the effects of hypoxia on Gulf menhaden distribution. Preliminary analyses indicate that catches during the summer of 1995 were low in hypoxic areas near shore. These data on adult fish distribution allow examination of the effects of hypoxia on higher tropic levels.

REEF FISH RESEARCH

Laboratory employees have studied reef fish off the southeastern United States since the late 1960's and have addressed both applied fishery issues and basic reef fish ecology. The Laboratory's Southeast Region Headboat Survey collects fisheries and biological data to support management activities. Approximately 165 vessels are monitored throughout the southeast, and in 1996, biological samples were collected from 37, 435 fishes of 124 species. These data are also used to examine patterns in the fishery and to study the structure and distribution of reef fish communities.

A variety of studies of reef fish ecology have been conducted by Laboratory employees using SCUBA. A temporal comparison off North Carolina suggests an increase in reef fish diversity and abundance possibly owing to climate change. The impact of removing piscivorous predators from reef areas is also being studied. Employees are continuing to examine changes in fish assemblages on managed and unmanaged reefs at Key Largo and participate in annual dives to assess reef fish populations at the Grays Reef National Marine Sanctuary off Sapelo Island, Georgia. Future research includes working with fisheries and GIS data to map essential reef fish habitat, and studying recruitment of larval fishes to reef habitats.

STOCK ASSESSMENT

Stock assessments are conducted by Laboratory employees for a variety of species in support of fisheries management. Stock assessments involve estimating population size-at-age, data which are collected by laboratory employees. From estimates of population size at age, age-specific fishing mortality for different levels of natural mortality can be calculated. Information on reproduction, such as age-specific maturity and fecundity, are used to calculate spawning potential ratios and to provide an indication of the level of fishing pressure experienced by the stock. The results of these assessments are provided to regional fishery management councils and commissions and include recommendations on how to either build or sustain healthy fish stocks along the Atlantic and Gulf coasts.

Laboratory employees conduct annual stock assessments for Atlantic menhaden and five year assessments for Gulf menhaden. Stock assessments are also completed for various reef fishes; recent assessments include red porgy, black sea bass, red snapper, vermilion snapper and scamp. Staff also conduct or provide advice on stock assessments for other fishery species in the southeast including blue crab, red drum, weakfish and wreckfish.

LINKING STOCK ASSESSMENTS WITH OCEANOGRAPHIC AND ECOLOGICAL PROCESSES

The ability to conduct stock assessment work and process-oriented work has allowed Laboratory researchers to further understand the processes that influence recruitment and stock abundance. For example, Gulf menhaden recruitment has been examined relative to Mississippi River discharge, a relation suggested by work done at the Laboratory during the GOMEX project. As another example, Atlantic menhaden recruitment has been related to the ingress of menhaden larvae into Beaufort Inlet. This time series of wintertime larval ingress at Beaufort Inlet has been used by participants in the SABRE program and data collection will continue.

MARINE MAMMALS

The Laboratory's research on marine mammals focuses on their interactions with fisheries, as well as life history and stock identification. Bottlenose dolphins are the primary species studied. Activities include recovery of stranded animals, live captures for tagging and data collection, photo-identification, and analysis of life-history samples. Results contribute to basic biological knowledge and to effective management and conservation.

As an example of some of the research conducted, dolphin 717 was freeze-branded during a live capture in 1995, and has since been regularly observed in the Beaufort area. Laboratory scientists know that she was a 3.5 year old female at the time of capture. Monitoring continues to determine when she will have her first calf as part of life history and recovery studies.

SEA TURTLES

Laboratory scientists continue an active program of sea turtle research. Surveys conducted since 1988 have underscored the importance of North Carolina's inshore waters, particularly the Pamlico-Albemarle Estuarine Complex, to juvenile loggerhead (*Caretta caretta*), green (*Chelonia mydas*), and Kemp's ridley (*Lepidochelys kempii*) sea turtles. Sea turtles are present in the Complex from April-December and during their emigration in early winter, are vulnerable to capture in pound nets set behind barrier islands. North Carolina pound nets are a passive gear that allow turtles to feed and to surface, to breathe.

Parts of the Pamlico-Albemarle Estuarine Complex were established as a pilot index area for sea turtles in 1995, using catch rates in pound nets as an index of abundance. The feasibility and methodology was established and the fishery was sampled again during fall 1996 and 1997. The goal of this project is to establish an index-abundance-area to monitor, assess, and predict the status of and impacts to sea turtles and their ecosystems. The cooperation of pound net fishermen in the area provides data not only on abundance, but also on movement and demography of these populations. In addition, blood samples are taken for analyses of health status, sex, and genetics.

PHRAGMITES RESEARCH

Researchers at the Laboratory are studying the use of *Phragmites* marsh by fish and shellfish species. Throughout the eastern United States many *Spartina alterniflora* salt marsh systems are being altered through *Phragmites australis* invasion. While information is lacking on the use of *P. australis* marsh by fish and shellfish species, increase in spatial distribution of *Phragmites* has concerned resource managers, who view such a shift as a degradation in habitat quality. Research conducted by Laboratory staff shows no significant differences in fish abundance or biomass between tidal *P. australis* and *S. alterniflora* marsh. Further, species diversity was higher in *P. australis* marsh. Stable isotope data from marsh fauna indicate that *P. australis* is an integral part of the estuarine food web.

MARINE CHEMISTRY

The Laboratory has a long history and continues research in marine chemistry. Although not highlighted here, the work bridges the gap between chemistry and habitat and fisheries issues. Studies include metals in the environment and their transfer through the food chain, the relation between iron and primary production, and otolith microchemistry as an indicator of past environmental exposure.

DISTRIBUTION AND DYNAMICS OF GAS HYDRATES IN THE MARINE ENVIRONMENT

Anne M. Trehu
Marta E. Torres
Oregon State University
Corvallis, Oregon

Gas hydrate is an ice-like substance that forms at low temperature and high pressure when adequate amounts

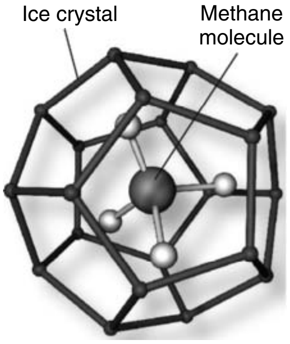

Figure 1. Schematic diagram of a gas hydrate molecule.

of water and low molecular weight gases, such as carbon dioxide, methane, and other hydrocarbons, are present (Fig. 1). In nature, gas hydrates containing primarily methane are widespread in the Arctic and in the shallow sediments on continental margins. Under certain geologic conditions, enough methane hydrate may be present for these deposits to represent a fossil fuel resource for the future. Concentrated deposits of gas hydrate near the seafloor form the substratum for unique assemblages of microbes that are only now being discovered; these microbes constitute the base of a foodweb for clams and other organisms, similar to those found at hydrothermal vent sites. Destabilization of gas hydrate in marine sediment may occur when artificial structures that heat up the surrounding sediments are installed; this may, in turn, result in catastrophic structural collapse. Because methane is a powerful greenhouse gas, sudden destabilization of gas hydrate on a large scale can have an impact on global climate. Indeed, past episodes of climate change linked to gas hydrate destabilization have been documented in the geologic record. For all of these reasons, researchers in many countries are trying to better understand the distribution and dynamics of gas hydrates in the marine environment.

NECESSARY CONDITIONS FOR FORMATION OF GAS HYDRATE IN NATURE

The temperature and pressure conditions required for gas hydrate to be chemically stable are found at the seafloor everywhere the water depth is greater than 300–500 m. (The minimum depth for gas hydrate stability decreases as the water temperature decreases.) As temperature increases beneath the seafloor at a rate that is controlled by the regional geothermal gradient, the effect of increasing temperature will eventually overtake the effect of increasing pressure, and gas hydrate will no longer be stable (Fig. 2). This balance between increasing pressure and temperature results in a gas hydrate stability zone beneath the seafloor that ranges in thickness from a few centimeters to several hundred meters. Based on these criteria, gas hydrate has the potential of forming almost everywhere beneath the continental slope and ocean basins.

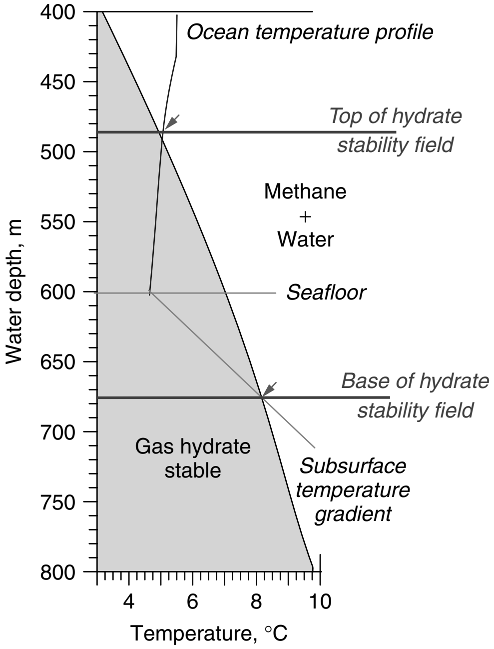

Figure 2. The temperature and pressure conditions under which gas hydrate is stable are constrained thermodynamically as shown by the stability boundary (shown here for pure methane hydrate in seawater). Hydrate is stable at low temperature and high pressure (which increases as water depth increases). This is shown by the light gray field. In this example, which depicts a site located at 600 m depth, the water temperature profile and the subseafloor temperature gradient are typical for the continental margin offshore Oregon. The top of the gas hydrate stability zone here is located within the water column at 485 m. Gas hydrates are stable in the sediments from the seafloor to a depth of 665 m, where the subsurface temperature becomes too high for methane hydrate to be stable.

Another critical condition that must be met for gas hydrate to form is that the concentration of gas present must exceed that required to stabilize the gas hydrate structure. Not enough methane is present in the sediments throughout most of the ocean basins. Gas hydrate is restricted to regions where organic material, which decomposes to form methane, is unusually abundant or where fluid flow transports gas into the gas hydrate stability zone from greater depth. These conditions are met on many continental margins (Fig. 3).

Other factors affecting gas hydrate stability are the salinity of the pore waters and the composition of the hydrate-forming gases that are present. Moreover, temperature and pressure conditions can change locally as a result of focused fluid flow, tectonic activity, or the introduction of artificial structures. It can change globally as a result of climatic perturbations. The distribution and abundance of gas hydrate in the marine environment are therefore constantly changing. New discoveries are being made every year, increasing the number of verified gas hydrate sites.

ORIGIN OF THE GAS IN GAS HYDRATE

The gas in gas hydrate has two possible origins. Methane can be generated in the upper few hundred meters by microbes (biogenic gas), or it can migrate into the

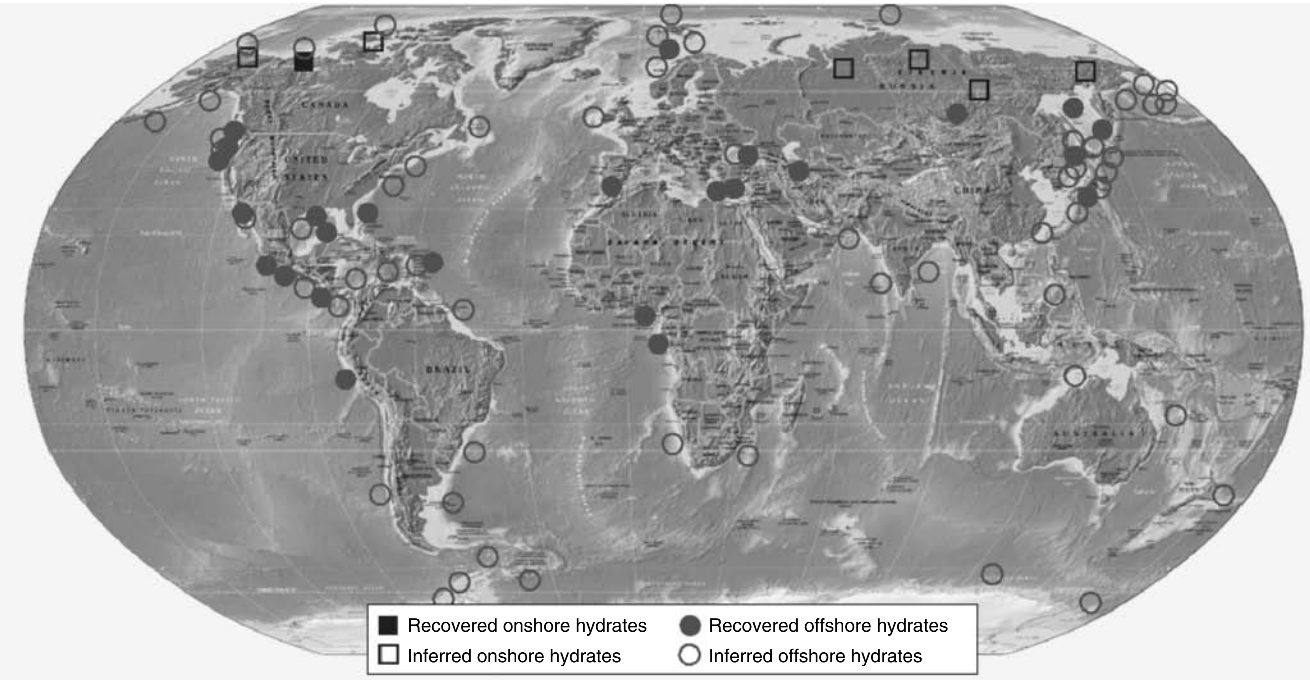

Figure 3. Global map showing locations where gas hydrate has been identified either through sample recovery or remote sensing. Figure courtesy of Alexei Milkov (BP America, Exploration and Production Technology Corp).

gas hydrate stability zone at greater depth. If the gas comes from subseafloor depths greater than about 2 km, it probably contains higher order hydrocarbon gases such as ethane and propane and has a distinctive carbon isotopic signature (thermogenic gas). Both sources of gas contribute to gas hydrate formation in marine sediments. Thermogenic gas is important in focused gas hydrate deposits, whereas biogenic gas dominates in the gas hydrates that are broadly distributed throughout large regions.

Figure 4. Burning the methane that is released as a gas hydrate sample decomposes. Hands belong to Robert Collier (College of Oceanic and Atmospheric Science, Oregon State University). Picture taken by one of the authors (M. Torres).

REMOTE SENSING FOR SUBMARINE GAS HYDRATES

Because methane hydrate is not stable at atmospheric pressure unless the temperature is below $-60\,^{\circ}$C, gas hydrate samples are rare. Most of the gas hydrate in a sediment sample is probably lost because the gas hydrate decomposes as the sample is recovered. Only large chunks of gas hydrate survive the trip from the seafloor to the earth's surface. Even these large chunks decompose rapidly after recovery, and the methane that is released by this decomposition can be burned, leading to the dramatic phenomenon of "burning ice" (Fig. 4). Special tools are required to sample and preserve natural gas hydrates for detailed studies.

Several geochemical and geophysical techniques are available for indirectly identifying the presence of gas hydrate *in situ*. One of the most widespread techniques is the use of seismic reflection data to detect the subseafloor boundary between sediments containing gas hydrate and sediments that contain free gas (Fig. 5). These data are analogous to an echo sounding of the earth beneath the seafloor. A sharp contrast in the elastic properties of sediments across this boundary results in a distinctive seismic reflection known as the bottom-simulating reflection (BSR). This reflection can be mapped over large areas of the ocean. Similarly, gas hydrates near the seafloor appear as strong reflections at the seafloor, although many other geologic features also produce seafloor "bright spots." Geophysical surveys are therefore a primary tool for determining whether gas hydrate is present in a "frontier" area. In many places, the presence of gas hydrate has been inferred from observations of a BSR in data that were originally

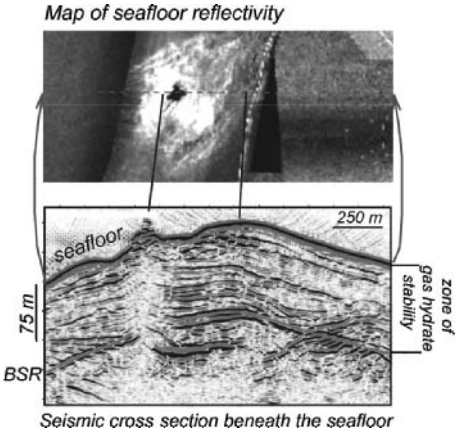

Figure 5. Data showing how the seafloor and subseafloor reflect high frequency sound waves. Such remote sensing strategies are used to survey the seafloor for the presence of gas hydrate. Bright (white) spots in the upper image indicate the presence at the seafloor of gas hydrate and associated carbonate. The presence of the reflection labeled BSR in the subsurface (lower image) indicates that gas hydrate is widespread deep in the sediments, even though massive gas hydrate at the seafloor is restricted to the summit of this structure. The seismic data can also be used to image the subsurface plumbing system that led to development of the hydrate deposit. Seafloor reflectivity data were acquired in 1999 by Chris Goldfinger (College of Oceanic and Atmospheric Science, Oregon State University). Seismic data were acquired by one of the authors (A. Trehu) in collaboration with Nathan Bangs (Institute for Geophysics, University of Texas).

acquired to address other problems in marine geology. These reflective patterns, however, are of limited use for determining how much gas hydrate is present and how it is distributed within the sediments. Researchers are developing new geophysical techniques to better quantify the distribution of gas hydrate from seismic and other remote-sensing techniques.

OBSERVATIONS OF GAS HYDRATE MOUNDS ON THE SEAFLOOR

Mounds of gas hydrates have been directly observed at the seafloor during submersible dives and by using deep-towed cameras and remotely operated vehicles (ROVs). Although many of these gas hydrate occurrences were originally identified on seafloor imaging data, some have been found by serendipity. For example, massive gas hydrate mounds offshore Vancouver Island, Canada, were first discovered when fishermen dragging for ground fish accidentally dredged some up in their nets (Fig. 6). Because gas hydrate is buoyant, their net rose rapidly to the sea surface when the gas hydrate broke free from the seafloor. Someone on board recognized that the bubbling, decomposing “ice” in the net was gas hydrate. Fortunately, no one was smoking!

Although these seafloor gas hydrate deposits may constitute only a small percentage of the total amount of gas hydrate present in marine sediments, they represent the most accessible and best studied deposits and are usually accompanied by complex and as yet poorly understood faunal assemblages. In all cases that have been studied in detail, these deposits contain a mixture of biogenic and thermogenic gas and result from geological structures that focus gas to the seafloor from a deep-seated source. They represent gas hydrate ‘sweet spots” that punctuate widespread diffuse deposits inferred from seismic data.

Vigorous expulsion of gas bubbles into the ocean is often observed above seafloor hydrate mounds, even when the mound lies within the thermodynamic gas hydrate stability field. The gas discharge, which is often episodic, provides evidence for the dynamic processes that lead to the formation of massive hydrate near the seafloor. It is likely that these deposits are ephemeral and that they are subject to destruction when fluid flow patterns change or when the seafloor is perturbed. It is possible that

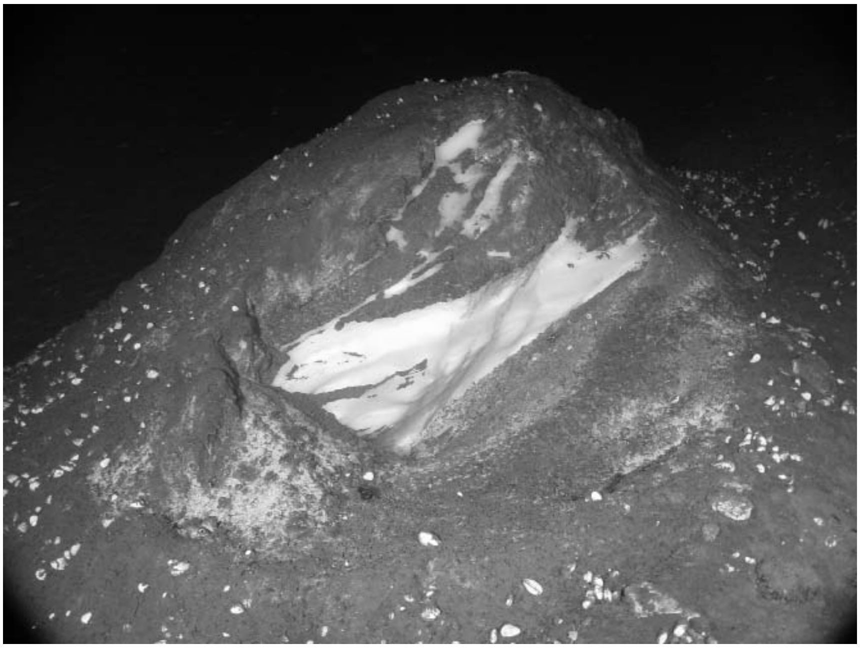

Figure 6. A thinly sedimented mound of gas hydrate exposed on the seafloor. White spots around the hydrate mound are clams, which are often associated with seafloor methane seeps. Photo was taken in August 2002 approximately 50 km offshore Vancouver Island during a survey by ROPOS (Remotely Operated Platform for Ocean Science), a Canadian research vehicle. Courtesy of Ross Chapman (University of Victoria).

large blocks of gas hydrate can spontaneously break away and rise to the sea surface, transporting methane from the seafloor to the atmosphere. Long-term observations of these deposits are needed to evaluate whether they represent a useable resource and whether they are an important factor in transferring methane from the ocean to the atmosphere.

SAMPLING GAS HYDRATES THROUGH DRILLING

More detailed information on the fine-scale distribution of gas hydrate in marine sediments can be obtained by drilling into the subsurface. Logging techniques permit measurement of geophysical and geochemical parameters throughout the entire borehole; coring allows for recovery of subsurface samples. Logging and coring to study gas hydrates is very expensive and has only been undertaken in a handful of places worldwide (Fig. 3). The Ocean Drilling Program (ODP), an international endeavor that enables the research community to sample up to a few kilometers beneath the Earth's surface to address problems of environmental change, earth history, and mantle dynamics, has dedicated several expeditions to study gas hydrates offshore Chile (Leg 141), the eastern United States (Leg 164), Japan (Leg 196), and the Pacific Northwest of the United States (Legs 146 and 204). The Japanese government has recently supported several expeditions to drill gas hydrates in the Nankai Trough offshore Japan. New programs are being planned to sample submarine gas hydrates in the Gulf of Mexico and elsewhere.

These drilling projects provide critical information needed to "ground truth" estimates of gas hydrate abundance and distribution inferred from remote sensing techniques. One conclusion from recent experiments is that the total amount of gas hydrate present in the globe, while still huge, may be less than previously thought. More importantly, these projects have revealed the "patchiness" of gas hydrate distribution beneath the seafloor and provide information on factors, such as sediment grain size, that may be controlling the heterogeneity in their distribution. They also provide unique information of the chemistry of the gases that form the gas hydrate and on the origin of these gases. This information is essential for understanding how gas hydrate deposits develop.

QUANTIFYING THE AMOUNT OF GAS HYDRATE IN BOREHOLES

To estimate the amount of gas hydrate *in situ*, several different types of measurements must be used because different techniques have different sensitivity and spatial resolution. The only way to directly measure the amount of gas in a core is to maintain the *in situ* pressure as it is recovered. This is technically difficult and can only be done for a limited number of samples in each drill hole. Another common technique is to measure the salinity or chloride concentration in the pore waters. Because salts are excluded when gas hydrate forms, the water that is released when it decomposes is fresh. Estimating the amount of hydrate present in the sediments using this technique requires knowledge of the salinity of the *in situ*pore water. This technique is limited by the need to extract pore water from the core, which is time consuming and disrupts the sample for other studies. Yet another technique is to measure the temperature of cores using a scanning infrared camera. Because gas hydrate decomposition is endothermic (takes up heat), locations in the core where gas hydrate is decomposing are up to several degrees colder than the rest of the core. This technique can efficiently and noninvasively sample the entire core but must be calibrated through comparison with other methods. Because core recovery is often incomplete, geophysical logging methods (which measure electrical resistivity, elastic wave velocity, density, and other physical parameters) are needed to sample the entire borehole. Models for interpreting these geophysical parameters obtained through logging must also be calibrated.

CONCLUSION

Gas hydrate is widespread in shallow sediments on the margins of the continents and represents an important, but often overlooked, component of the global carbon cycle. The distribution of gas hydrate is very heterogeneous. While gas hydrate probably represents only ~1–2% of the total volume of continental margin sediments that lie within the gas hydrate stability zone at the present time, concentrated deposits containing 10–30% gas hydrate have been found. These generally result from geologic conditions that lead to focusing of gas from a large volume in the subsurface into the gas hydrate stability zone.

The amount and distribution of gas hydrate in the seafloor, however, are probably not static. Instead, they change with time in response to changes in temperature, pressure, and fluid flow, which are, in turn, driven by tectonics (e.g., earthquakes and underwater mountain building) and global climate change. Rapid release of methane from gas hydrate in response to changes in ocean temperature and depth may affect global climate.

OCEANOGRAPHIC ENVIRONMENT OF GLACIER BAY

Alaska Biological Science
Center—U.S. Geological Survey

Oceanography describes one of the most fundamental physical aspects of a marine ecosystem. Glacier Bay exhibits a highly complex oceanographic regime within a small area. An understanding of many of the resource and research issues in Glacier Bay will not be possible without an understanding of the underlying oceanographic

---

processes causing the large spatial and annual variation found within the Bay.

The Bay is a recently (300 years ago) deglaciated fjord located within Glacier Bay National Park in Southeast Alaska. Glacier Bay is a fjord estuarine system that has multiple sills. These sills are often associated with contractions and are backed by very deep basins with tidewater glaciers and many streams. Glacier Bay experiences a large amount of runoff, high sedimentation, and large tidal variations. Melting occurs year-round, which is thought to fuel the estuarine circulation even through the winter. This runoff, and the presence of the tidewater glaciers makes the bay extremely cold. There are many small- and large-scale mixing and upwelling zones at sills, glacial faces, and streams. The complex topography and strong currents lead to highly variable salinity, temperature, sediment, productivity, light penetration, and current patterns within a small area. This complexity defies simple characterization or modeling based on other areas in Southeast Alaska. While several oceanographic studies have been conducted in Glacier Bay, these studies are contradictory and were of short duration and limited coverage, missing much of the spatial, seasonal and annual variation. In addition, some assumptions based on past studies have been contradicted by recent results. The constantly changing nature of the Bay may contribute to contradictions among past studies and between recent and historical results.

The primary data used in this study (see the Fjord Oceanographic Monitoring Handbook) were oceanographic surveys consisting of samples of 24 central-channel stations from the mouth of the bay to the heads of both the East and West Arm taken between 1992 and 2000. A oceanographic instrument (CTD) capable of recording depth, temperature, salinity, light penetration, amount of sediment, and amount of phytoplankton was used to obtain measurements at one-meter intervals throughout the water column to a depth of 300 m at each station. Surveys were conducted up to seven times a year in such a manner as to encompass the primary annual variation in oceanographic patterns.

Results from the current work indicate several shifts in the dominant paradigm of oceanographic understanding for this area. Primarily, deep-water renewal, and with it increased nutrient availability, is not limited to the winter months but can and probably does occur regularly in the spring/summer/fall at least during spring-tide series. We believe that Glacier Bay is not a traditional silled fjord estuary nor a plain fjord estuary like most estuaries in Southeast Alaska, but a combination of a stratified deep basin estuary and a tidally mixed estuary. This results in a complete blockage of estuarine circulation at the mouth of the bay; instead, mixing occurs by turbulent diffusion rather than buoyancy-driven entrainment as seen in a traditional estuary. In addition, where this turbulent water meets the deep stratified basin, a tidally-mixed front is created. This front is ideal for encouraging high primary productivity by phytoplankton. Mixing phenomena have now been observed to be much more extensive in Glacier Bay than previously thought. This, as with the increased deep-water renewal, almost certainly results in a more nutrient enhanced system. The Bay is now known to exhibit phenomenally high primary productivity, which is sustained throughout the summer in sharp contrast to other interior waters of Southeast Alaska and adjoining shelf and oceanic areas. The limits on this productivity appear to be caused not by grazing by zooplankton or nutrient limitation, but by light penetration. Light penetration in Glacier Bay is significantly reduced by both sedimentation phenomena related to the young nature of the terrestrial environment and to the high levels of the primary productivity itself. In addition to these major changes in understanding, we have identified several new phenomena that will need further study to understand their nature and significance.

As a result of the oceanographic monitoring program and the synthesis effort funded by the NPS there is now an extensive body of oceanographic and weather data integrated into the Glacier Bay information management system and available on CD-ROM. The Oceanographic Analyst Extension, a Geographic Information System (GIS) tool, has been created to allow viewing, analysis, and manipulation of these complex data in 3 and 4 dimensions.

NOS SANCTUARIES PROTECT NATION'S MARITIME HISTORY

NOAA National Ocean Service

Marine archaeology allows us to unravel the mysteries and riddles of shipwrecks and clarify our history. Each fragile wreck is a time capsule of our past. The remains of sunken vessels and their contents provide a distinctive window into the history of the ships and the lives of its passengers and crew. Using state-of-the-art technology, such as sonar and remotely operated vehicles, marine archaeologists are now able to conduct advanced explorations and create accurate depictions of wreck sites.

The National Ocean Service's National Marine Sanctuary Program plays a pivotal role in the exploration and preservation of the nation's maritime history. By protecting oceanic treasures within a sanctuary, marine archaeologists are able to uncover and preserve invaluable information and artifacts. Shipwrecks discovered in the sanctuaries, such as the *USS Monitor, Portland*, and those in Thunder Bay, are essential to understanding the importance of maritime history.

USS MONITOR

Undiscovered for more than 100 years, the *USS Monitor* was found in 1973 off the coast of North Carolina by scientists from Duke University using sidescan sonar technology. In 1862, this famous Civil War ship's battle with the Confederate ironclad *Virginia* revolutionized the

---

The crew on the deck of the *USS Monitor*. Members of the crew were lost when the *Monitor* sank in 1862. Human remains recently were found when the newly recovered turret was excavated last summer. (Monitor Collection, NOAA)

war at sea, and marked the end of an era of wooden-hulled sailing warships. The *Monitor* was the first of a class of low-freeboard, turreted war ships developed during the Civil War. Its revolving turret revolutionized naval warfare, and other features of the *Monitor* foreshadowed the future of naval technology.

The *Monitor* is more than an era frozen in time. Its dedication in 1975 as the first national marine sanctuary opened the door to advances in maritime archaeology and historic preservation. Since its discovery, numerous research expeditions have been conducted to study and recover the wreck. One of the most recent was in the *summer of 2002*. With the help of a claw-like structure known as the spider, the turret of the *Monitor* was recovered. Human remains also were found during the excavation of the turret in 2002. More than 100 artifacts also have been retrieved from the wreck, including the ship's anchor, condiment bottles, dinnerware, and a portion of leather bookbinding. These artifacts are now part of The Monitor Collection, which is housed in The Mariners' Museum Research Library. More information about the *USS Monitor* is available *from the Monitor National Marine Sanctuary*.

PORTLAND AT STELLWAGEN BANK

The coastal steam ship *Portland*, which rests on the sea floor of the *Stellwagen Bank National Marine Sanctuary*, sank Nov. 27, 1898, during the infamous "Portland Gale of 1898." All 192 passengers and crew were lost. With the help of state-of-the-art technology, the sanctuary obtained three new side-scan sonar images of the wreck. The images clearly show the side-by-side smoke stacks and the diamond-shaped metal walking beam that provided power to the side paddle wheels. The *Portland* has yet to be fully explored, but another mission is scheduled for September 2003. During this missions, scientists will use state-of-the-art technology including side-scan sonar and remote operating vehicles for exploration. View the *side-scan sonar images of the Portland* that were collected last year.

The *Monitor's* turret broke the surface during the summer 2002 expedition—the first time it had done so in 140 years

Stellwagen Bank, first discovered by Captain Henry Stellwagen in 1854, was especially important to navigators when they sailed at night or in poor weather. Mariners could gauge the distance to the sea floor and, when they saw the bank, would know that they were entering the Massachusetts Bay. The wealth of sea life present in the area also allowed for centuries of prosperous fishing fleets and other economic gains.

Stellwagen Bank was designated on Nov. 4, 1992, as the nation's 11th national marine sanctuary. Stretching 842 square miles, more than 1 million people each year visit the sanctuary. Many of them are intent on experiencing an encounter with the well-known Humpback whale.

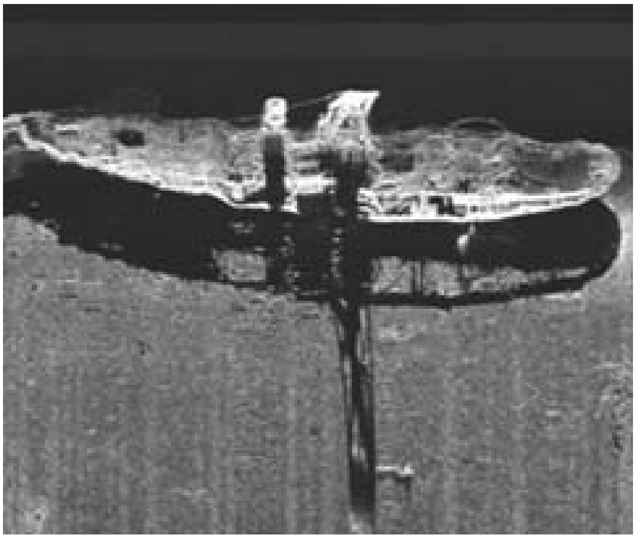

A sonar image of the *Portland*, taken in 2002

On June 30, 1906, the *New Orleans* was carrying coal when she collided with the *William R. Linn*. Al of the crew boarded the *Linn* safely before the pilothouse of he *New Orleans* blew off and the vessel sank. This sonar image of the wreck was taken in summer 2001

THUNDER BAY

The *Thunder Bay National Marine Sanctuary and Underwater Preserve*, dedicated on Oct. 7, 2000, is the first freshwater and Great Lakes sanctuary. More than 100 19th and 20th century shipwrecks rest in Thunder Bay and the surrounding waters of Lake Huron. For this reason, it has earned the nickname "Shipwreck Alley." Located off the coast of Alpena, Mich., Thunder Bay's collection of shipwrecks represents the diversity of vessels that navigated the Great Lakes, from wooden schooners to steel-hulled steamers.

These preserved vessels still have stories to tell of Great Lakes maritime history and commerce. Some of the wrecks have been broken apart and washed away, and others have recently been found and are waiting to be explored. Many wrecks are yet to be discovered but likely will be found as explorations continue. For more information, read more about *surveying Thunder Bay's Shipwreck Alley* in June 2001.

QUANTIFICATION OF ANOXIA AND HYPOXIA IN WATER BODIES

Gertrud K. Nürnberg
Freshwater Research
Baysville, Ontario, Canada

Hypoxic conditions (dissolved oxygen, DO, concentration below saturation) are widely spread in freshwater and saline environments. Especially in recent years, anthropogenic impacts have led to severe increases in estuarine and coastal anoxia [e.g., Gulf of Mexico (1) and European coast-lines (2)]. Diaz (3) describes 44 marine areas of moderate to severe hypoxia worldwide. Hypoxia and anoxia (lack of oxygen or zero concentration of DO) may even be more established in the freshwater environment, where it can occur naturally in lakes (4) or not (5) and where recent increases in organic and nutrient loading and changes in water flow have increased oxygen depletion in lakes (6), reservoirs (7), and large rivers (8,9).

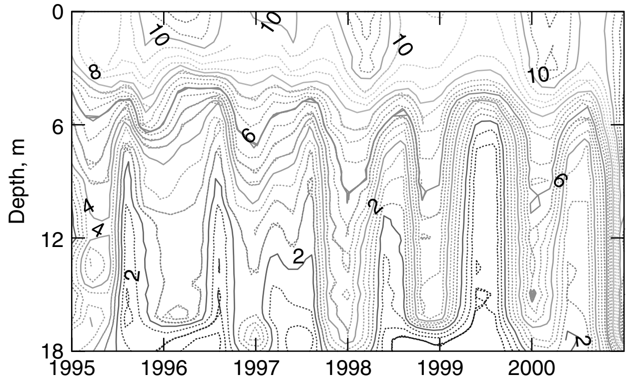

Figure 1. Dissolved oxygen isopleths for small urban Lake Wilcox, Ontario. The beginning of each year is indicated.

OCCURRENCE OF DO DEPLETION

An example isopleth (Fig. 1) for several years of DO data for an urban lake demonstrates DO variability with respect to space and time. The lake experiences severe anoxia in the hypolimnion because the bottom water is stagnant during the summer and fall. Entrainment of hypolimnetic anoxic water may lead to hypoxic conditions in the epilimnion in late summer. Such variation can be quantified by computing anoxic and hypoxic factors.

DEFINITION

The *anoxic* and *hypoxic factor* [AF (4), HF (10)] are calculating methods that quantify the extent and duration of anoxia and hypoxia based on oxygen profiles and morphometric (hypsographic) information. They are based on a series of measured oxygen profiles and morphometric data and can be computed for any water body (lake, reservoir, river, marine area). To render this index comparable across waters like other areal measures, for example, areal nutrient loads and fish yields, AF is corrected for (lake) surface area by simple division. Expressed this way, AF or HF is a ratio that represents the number of days in a year or season that a sediment area equal to the (lake) surface area is anoxic or hypoxic. Hence, its units are d/yr and d/season, summer or winter.

The exact steps in computing the specific factors are described in Fig. 2. The difference between the HF and the AF lies in the different threshold DO concentration used. The AF is based on the oxycline depth or its approximation; the HF is based on a certain threshold that indicates the hypoxic conditions of choice for that particular project, for example, 6.5 mg/L DO. To consider the mixed layer anoxia or hypoxia separately, epilimnetic factors (AF_{epi}, HF_{epi}) can be computed, as defined in Fig. 2.

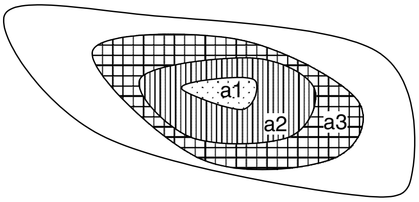

Figure 2. Computation and definition of anoxic and hypoxic factors. The sketch indicates different lake areas that become anoxic sequentially with time, where a_1 becomes anoxic first, then a_2, etc.

Determination of Anoxic or Hypoxic Factor

1. For the anoxic factor, the oxycline is determined at 1 or 2 mg/L dissolved oxygen (DO) concentration from (DO) profiles. The choice of the threshold values depends on the method of DO measurement. For example, 2 mg/L measured by a DO probe may actually coincide with anoxic conditions at the sediment surfaces located at that depth.
   For the hypoxic factor, the depth of the DO concentration threshold in question, for example, 6.5 mg/L, is determined.
2. The periods (total number n), for which the DO levels according to the chosen threshold are at approximately the same depth, are established.
3. The duration of each period of constant DO levels (t_i, in days) is multiplied by the corresponding area (a_i) and divided by the total surface area (A_{oi}) for the period (i). Especially in reservoirs where volumes and therefore areas change, is it important to use the surface area specific for the period. In water bodies without large volume changes, the average surface area A_o can be used instead.
4. These n terms, numbers of periods at different oxyclines, are then added up over the season or year, according to

$$\frac{t_1 \cdot a_1}{A_{o1}} + \frac{t_2 \cdot a_2}{A_{o2}} + \frac{t_3 \cdot a_3}{A_{o3}} + \cdots + \frac{t_n \cdot a_n}{A_{on}} = \sum_{i=1}^{n} \frac{t_i \cdot a_i}{A_{oi}}$$

Therefore AF and HF are defined as follows, where a lower DO threshold is used for AF (1 or 2 mg/L) than for HF (e.g., 6.5 mg/L):

$$\text{AF or HF} = \sum_{i=1}^{n} \frac{t_i \cdot a_i}{A_{oi}}$$

End of Summer Stratification. In stratified lakes, much of the oxygen depletion occurs in the fall when DO profiles may not always be available. The duration of the last period of anoxia or hypoxia can be estimated from the average summer hypolimnetic temperature according to a model that predicts the fall turnover date (or end of summer stratification) from the following equation, based on 92 worldwide lakes, $R^2 = 0.47, p < .0001$ (11):

$$\text{Log(fall turnover date)} = 2.62 - 0.116 \log(T) + 0.042 \log(z) - 0.002 \text{ latitude}$$

where the fall turnover date is the Julian day of the year; T is the average July, August temperature at circa 1 m above the bottom at the deepest location of the lake (C°); and z is mean depth, lake volume/lake surface area (m).

Seasonal Factors. The AF and HF can be determined for the stratification periods separately as winter (e.g., AF_{win} in d/winter) or summer factor (e.g., AF_{sum} in d/summer). When there is no winter anoxia, AF_{win} is zero, and AF_{sum} is equal to AF in d/yr.

Epilimnetic Factors. In stratified water bodies, it can be interesting to separate hypolimnetic from epilimnetic hypoxia, so that guidelines can be established specifically for the mixed zone. Therefore, epilimnetic factors (AF_{epi}, HF_{epi}) can be defined. These factors are determined by subtracting, for each period i, terms corresponding to anoxia or hypoxia in the seasonal stratified area below the thermocline ($a_{\text{thermo } i}$) from those for the whole water column before summation, according to

$$\text{AF}_{\text{epi}} \text{ or } \text{HF}_{\text{epi}} = \sum_{i=1}^{n} \left(\frac{t_i \cdot a_i}{A_{\text{o } i}} - \frac{t_i \cdot a_{\text{thermo } i}}{A_{\text{o } i}} \right)$$

Application

Internal Load Calculation. The anoxic factor was originally developed to determine the internal phosphorus load in lakes as the product of anoxic areal release rates (experimentally determined or predicted from sediment phosphorus concentration) and AF according to Eq. 1 (12).

$$\text{Internal load} = \text{Release rate} \times \text{AF}_{\text{sum}} \qquad (1)$$

where (summer) internal load is in mg TP per lake surface area (m^2) per summer, and release rate represents a summer average in mg TP released per anoxic sediment area per day of the anoxic summer period. If there is phosphorus release under ice in the winter, a release rate representative for the winter and AF_{win} are substituted in Eq. 1 to estimate the winter internal load.

Lake to Lake Variation. Anoxia can vary extensively between lakes. Even in geographically near areas like the Muskoka—Haliburton region on the Precambrian shield in southern Central Ontario, the AF ranged from 0 to 70 d/yr in relatively pristine lakes that were often small, but deep, whereas at least 50% had an AF > 10 d/yr. Lakes in more fertile regions, such as in the St. Lawrence–Great Lakes region, had consistently high AFs between 40 and 70 d/yr. Such differences in lakes and their AF indicate the need for different management strategies (5).

Quantification of anoxia by the AF has been proven useful in explaining lake to lake variation in anoxia. In particular, dependency of anoxia on nutrient concentration and morphometry was found and multiple regressions of the AF on average lake TP and TN concentrations and a morphometric ratio were highly significant in North American lakes (4) (Table 1, Eqs. A to C). These relationships can be used to model (or predict) the AF

Table 1. Relationships and Models Involving the Anoxic Factor (AF) that Demonstrate Between Lake Variation of Anoxia in North American Stratified Lakes[a]

Equation	Predicted	Regression Equation[b]	N	R^2	Source
A	AF_{sum}	$-35.4\ (5.1) + 44.2\ (4.3) \log(TP_{annual}) + 0.950\ (0.187)\ z/A^{0.5}$	73	0.65	(4)
B	AF_{sum}	$-36.2\ (5.2) + 50.1\ (4.4) \log(TP_{summer}) + 0.762\ (0.196)\ z/A^{0.5}$	70	0.67	(13)
C	AF_{sum}	$-173\ (25) + 72.5\ (9.7) \log(TN_{summer}) + 0.789\ (0.219)\ z/A_o^{0.5}$	61	0.54	unpubl. data
D	AF_{sum}	$-39.9\ (9.7) + 27.0\ (4.0) \log(\text{TP-Load})$	17	0.76	(4)
E	Fish species number	$0.97\ (2.42) - 1.53\ (0.49) \log(AF_{sum}+1) + 5.38\ (1.02) \log(A_o)$	52	0.51	(14)
F	Fish species number	$4.92\ (1.38) - 6.12\ (1.44) \log(AF_{win}+1)^b + 0.56\ (0.11)\ z$	32	0.71	(14)
G	Winterkill AF_{win}	$-1 + 10^{(0.091z+0.804)}$	–	–	Based on Eq. F

[a] AF_{sum}, summer AF (d/summer); AF_{win}, winter AF (d/winter); Winterkill-AF_{win}, value of AF_{win} above which a fish winterkill is likely for a given z; z, mean depth (m); A_o, lake surface area (km^2); TP_{annual}, annual water-column average TP (μg/L); TP_{summer}, epilimnetic summer TP (μg/L); fish species number, also called fish species richness is based on Central Ontario lakes.

[b] All regressions are highly significant at $p < .001$, except as noted by[b], $p < .01$. Standard errors of the regression coefficients are given in parenthesis.

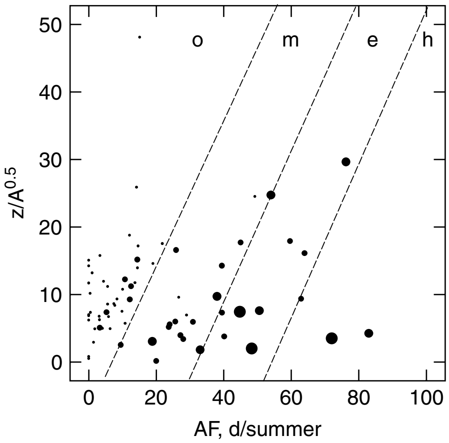

Figure 3. Chart to determine the lake trophic state from the anoxic factor (AF) and the morphometric variable $z/A_o^{0.5}$. Observed values from 70 North American lakes are presented. The symbol size indicates the four trophic states, oligotrophy, mesotrophy, eutrophy, and hypereutrophy (from small to large), and lines indicate the trophic state boundaries based on TP_{summer} in Equation B, Table 1).

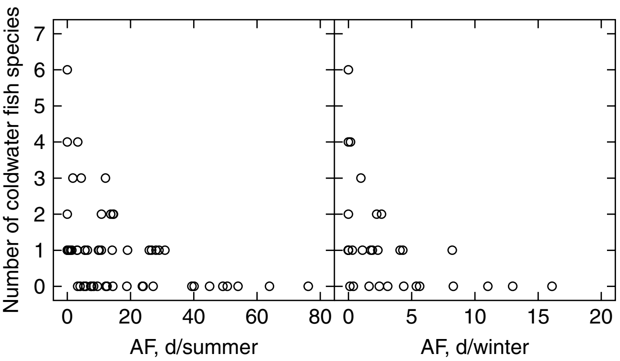

Figure 4. Number of South Central Ontario coldwater fish species in relation to AF_{sum} (left panel) and AF_{win} (right panel).

in lakes where DO data are not available or difficult to obtain, as in polymictic lakes. The AF describes the sediment surface that is overlain by anoxic water in stratified lakes, so it can be hypothesized that its predicted value resembles anoxic sediment surfaces in polymictic lakes. Such a modeled variable is useful in determining internal load according to Eq. 1 for shallow lakes, where release rates may be available but not the extent of anoxic sediment areas.

Trophic State Classification. The concept of the anoxic factor was used to establish limits for trophic state classification with respect to anoxia (13). When classified with respect to the trophic state, below 20 d/yr, the AF indicates oligotrophic conditions, 20 to 40 d/yr are usually found in mesotrophic lakes, 40–60 d/yr represents eutrophic conditions, and above 60 d/yr is typical of hypereutrophic conditions (13). A more exact measure of the trophic state would include lake morphometry, measured as $z/A^{0.5}$, because the deeper the lake compared to its area (the larger $z/A^{0.5}$), the larger the AF. Based on the relationship of the AF with the summer average TP and the morphometric ratio of Eq. B in Table 1, a chart was drawn to assist in evaluating a specific lake's trophic state (Fig. 3).

Fish species richness correlated with the AF in Central Ontario lakes after variability due to lake area was taken into account, and winter kill could be predicted by using a seasonal AF that quantifies anoxia under ice, and mean depth (14) (Table 1). In particular, cold water species including *Salmonidaes*, *Coregonidae*, and *Gadidae* were sensitive to summer and winter anoxia, and they occurred only when AF_{sum} was below 32 d/summer or AF_{win} was below 4.4 d/winter (Fig. 4).

Spatial Within-Lake Variability. In large and morphometrically distinct water bodies, the factors can be estimated separately for different sections. For example, the AF was quite different between the two basins in small (77 ha) Cinder Lake, Ontario, at 12 and 40 d/summer (4). Similarly pronounced were differences for the 100-km long, narrow (maximum 1 km wide) reservoir, Brownlee Reservoir, on the Snake River in Idaho and Oregon (15). The total reservoir had an AF of 68 d/yr on average between 1962 and 2000, ranging from 53 to 90 d/yr. But the AF of the 48-km long, deep lacustrine section was higher, 88 d/yr on average, ranging from 72 to 114 d/yr, because this section is stratified all summer long. The AF for the 52-km long, shallow riverine section was only 10 d/yr in 1999 and 14 d/yr in 2000 because it is often mixed and reaerated from atmospheric oxygen. Epilimnetic factors were smaller than factors for the whole water column, as expected. Severe anoxia in this reservoir is apparent because the long-term average lacustrine AF_{epi} was 10 d/yr, indicating that overall, an area in the lacustrine epilimnion equal to the lacustrine surface area is overlain by water below 2 mg/L DO for 10 days per summer. The average HF_{epi} of 58 d/yr indicates that the Oregon state criterion of 6.5 mg/L DO for the water column was exceeded in the lacustrine epilimnion for an equivalent of circa 2 months, on average.

Annual Variability. Year-to-year variability can be quite large and ranged from 72 to 114 d/yr AF and 102 to 215 d/yr HF in lacustrine Brownlee Reservoir between 1962 and 2000, from 45 to 68 d/yr AF in a small urban lake, Lake Wilcox, between 1987 and 2000, and from 0 to 10, 9 to 30, and 10 to 50 d/yr AF in three small glacial lakes on the Precambrian shield [Nürnberg, unpublished data and (4,5,10)].

Although the average AF can be predicted from TP and morphometry according to the equations in Table 1, TP concentration cannot always explain the variability in the AF between years, as the variation between annual average TP is often small. Instead, physical characteristics, especially hydrology in flow-managed reservoirs have been found to control the residual variance of DO depletion between years.

For example, several annual and seasonal hydrologic variables were significantly ($p < .01$) correlated with AF, HF, AF_{epi}, and HF_{epi} in the Snake River Brownlee Reservoir. In particular, spring (April–May), summer inflow (July–Sept) (Fig. 5), and late fall (Oct–Dec) flushing rate or its inverse, the residence time, were strongly correlated with the factors, so that they were smallest during periods of high inflow or flushing (10). Such a good inverse relationship of the fall outflow with anoxia and hypoxia can be expected because severe DO depletion happens mostly in the summer and fall; high flow rates out of the dam at that time would destabilize stratification in the lacustrine section, leading to decreased hypoxia. On the other hand, a strong inverse correlation for spring inflow with hypoxia can be explained by a delay in the onset of stratification.

Climatic Effects. The factors quantify annual trends, so climatic effects on DO depletion can be investigated. For example, the Pacific Northwest Index (PNI) was significantly correlated with HF and HF_{epi} [Fig. 6 (10)]. The PNI is a normalized index based on three terrestrial climate variables for the western North American continent: the air temperature at San Juan Islands, the total precipitation at Cedar Lake in the Cascade Mountains, and the snowpack depth on Mount Rainier (16).

Establishment of Criteria and Guidelines. The factors are particularly useful in setting DO criteria and guidelines because the quantification of hypoxia is flexibly based on individual threshold values and can be adapted to various DO levels. Typically, a threshold DO concentration is set by governmental agencies under which a water body is considered “impaired” and a plan to attain the guidelines, including the process of determining a TMDL (total maximum total load) (17), is initiated. These thresholds or acceptable levels vary for different water bodies, agencies, and intended usage but can all be accommodated by various definitions of the hypoxic factor.

Additional Applications. Furthermore, AF helped to compare methods to measure denitrification rates in lakes and determine the most appropriate one (18) and reveal the dependency of hypolimnetic anoxia in tropical African reservoirs on stage height and volume (19). The anoxic

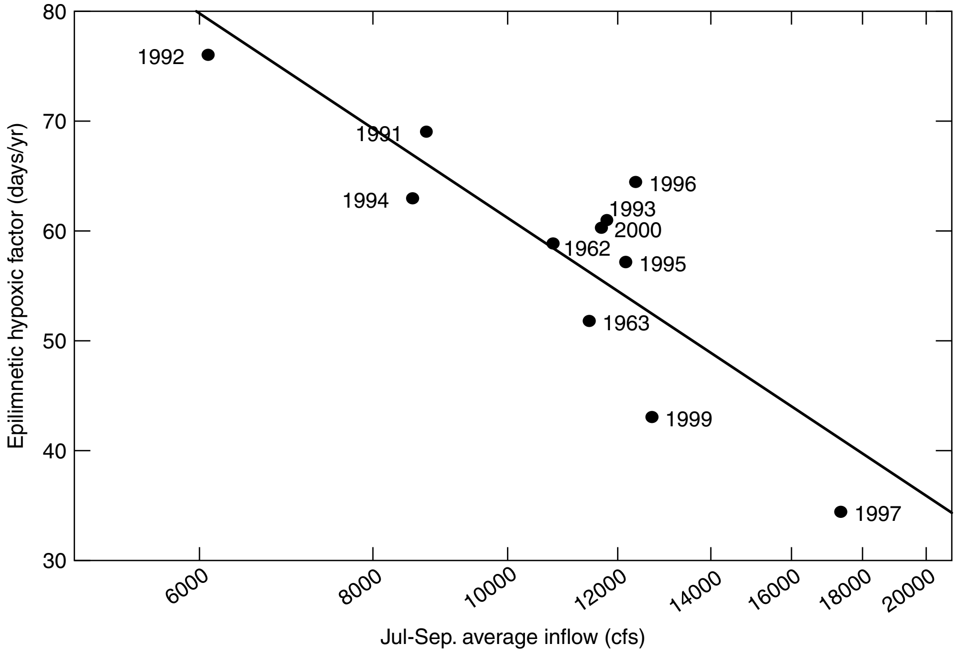

Figure 5. Lacustrine HF_{epi} of the Snake River Brownlee Reservoir, versus average summer inflow (July–Sept, computed from daily average flow values in cubic feet per second) between 1962 and 2000. The regression line is shown for $n = 11$, $p < .001, R^2 = 0.72$.

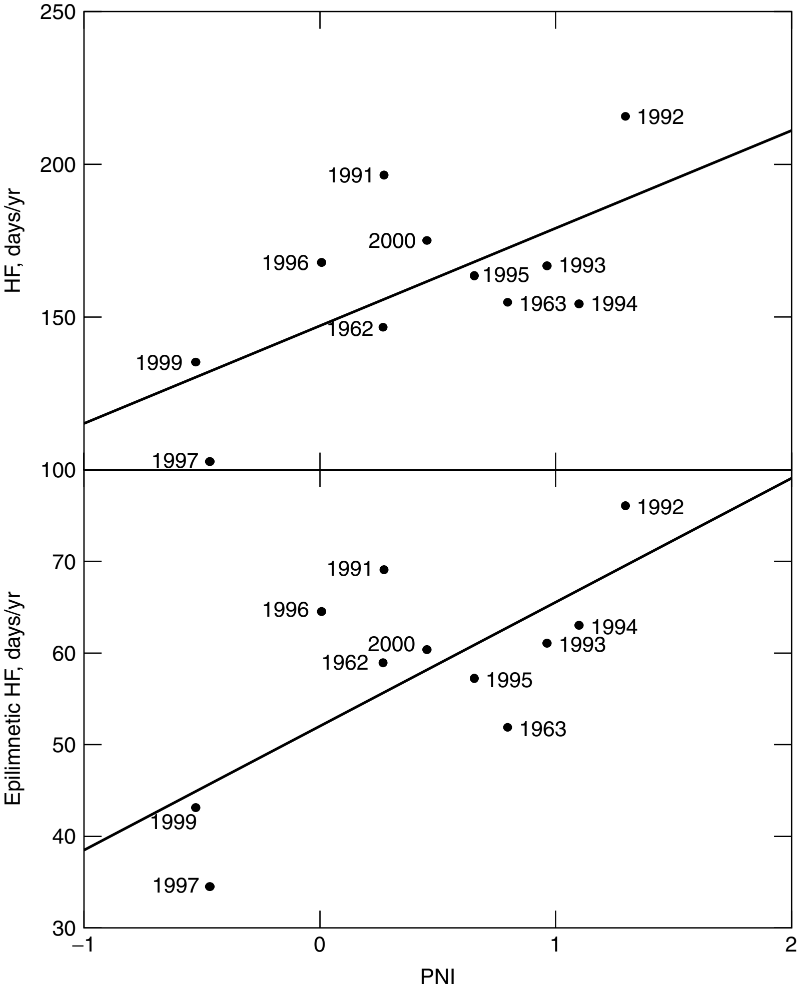

Figure 6. Lacustrine HF (top) and HF_{epi} (bottom) of Brownlee Reservoir versus the Pacific Northwest Climate Index (PNI). Regression lines are shown for $n = 11$, HF: $p < .05$, $R^2 = 0.41$; HF_{epi}: $p < .02, R^2 = 0.49$.

factor and its relationship with TP and morphometry were combined with information on fossil midge (*Chironomidae*) assemblages in paleolimnological studies to establish training sets and hindcast hypolimnetic anoxia for the past (20–22).

CONCLUSIONS

In general it can be concluded that the quantification of anoxia and hypoxia leads to increased knowledge of factors that control oxygen depletion in any freshwater and should be useful in estuarine and marine systems as well.

BIBLIOGRAPHY

1. Rabalais, N.N., Turner, R.E., and Wiseman, W.J. (2001). Hypoxia in the Gulf of Mexico. *J. Environ. Qual.* **30**: 320–329.
2. Johannessen, T. and Dahl, E. (1996). Declines in oxygen concentrations along the Norwegian Skagerrak Coast, 1927–1993: A signal of ecosystem changes due to eutrophication? *Limnol. Oceanogr.* **41**: 766–778.
3. Diaz, R.J. (2001). Overview of hypoxia around the world. *J. Environ. Qual.* **30**: 275–281.
4. Nürnberg, G.K. (1995). Quantifying anoxia in lakes. *Limnol. Oceanogr.* **40**: 1100–1111.
5. Nürnberg, G.K. (1997). Coping with water quality problems due to hypolimnetic anoxia in Central Ontario Lakes. *Water Quality Research Journal of Canada* **32**: 391–405.
6. Schertzer, W.M. and Sawchuk, A.M. (1990). Thermal structure of the Lower Great Lakes in a warm year: implications for the occurrence of hypolimnion anoxia. *Trans. Am. Fish. Soc.* **119**: 195–209.
7. Gordon, J.A. (1993). Dissolved oxygen in streams and reservoirs. *Water Environ. Res.* **65**: 571–573.
8. Quinn, J.M. and McFarlane, P.N. (1989). Epilithon and dissolved oxygen depletion in the Manawuta River, New Zealand: Simple models and management implications. *Water Res.* **23**: 825–832.
9. Chambers, P.A., Brown, S., Culp, H.M., and Lowell, R.B. (2000). Dissolved oxygen decline in ice-covered rivers of northern Alberta and its effects on aquatic biota. *J. Aquat. Ecosys. Stress Recov.* **8**: 27–38.
10. Nürnberg, G.K. (2002). Quantification of oxygen depletion in lakes and reservoirs with the hypoxic factor. *Lake and Reservoir Manage* **18**: 000–000.
11. Nürnberg, G.K. (1988). A simple model for predicting the date of fall turnover in thermally stratified lakes. *Limnol. Oceanogr.* **33**: 1190–1195.
12. Nürnberg, G.K. (1987). A comparison of internal phosphorus loads in lakes with anoxic hypolimnia: laboratory incubations

versus hypolimnetic phosphorus accumulation. *Limnol. Oceanogr.* **32**: 1160–1164.

13. Nürnberg, G.K. (1996). Trophic state of clear and colored, soft- and hardwater lakes with special consideration of nutrients, anoxia, phytoplankton and fish. *Lake and Reservoir Management* **12**: 432–447.
14. Nürnberg, G.K. (1995). Anoxic factor, a quantitative measure of anoxia and fish species richness in Central Ontario lakes. *Trans. Am. Fish. Soc.* **124**: 677–686.
15. Freshwater Research, Brown and Caldwell. (2001). *Assessment of Brownlee Reservoir Water Quality*. 2001, Technical Report for the City of Boise, ID.
16. Petersen, J.H. and Kitchell, J.F. (2001). Climate regimes and water temperature changes in the Columbia River: bioenergetic implications for predators of juvenile salmon. *Can. J. Fish. Aquat. Sci.* **58**: 1831–1841.
17. EPA. (1999). *Protocol for Developing Nutrient TMDLs*. U.S. EPA, Washington, DC.
18. Molot, L.A. and Dillon, P.J. (1993). Nitrogen mass balances and denitrification rates in Central Ontario lakes. *Biogeochem.* **20**: 195–212.
19. Townsend, S.A. (1999). The seasonal pattern of dissolved oxygen, and hypolimnetic deoxygenation, in two tropical Australian reservoirs. *Lake & Reservoirs: Research and Management* **4**: 41–53.
20. Quinlain, R., Smol, J.P., and Hall, R.I. (1998). Quantitative inferences of past hypolimnetic anoxia in south-central Ontario lakes using fossil midges (Diptera: Chironomidae). *Can. J. Fish. Aquat. Sci.* **55**: 587–596.
21. Little, J.L. and Smol, J.P. (2001). A chironomid-based model for inferring late-summer hypolimnetic oxygen in southeastern Ontario lakes. *J. Paleolimnol.* **26**: 259–270.
22. Clerk, S., Hall, R., Quinlain, R., and Smol, J.P. (2000). Quantitative inferences of past hypolimnetic anoxia and nutrient levels from a Canadian Precambrian Shield lake. *J. Paleolimnol.* **23**: 319–336.

FLOATING ICE

Aldo Conti
Frascati (RM)
Frascati, Italy

Floating ice, or sea ice, is very important because it regulates the surface exchange of heat, moisture, and salt between the atmosphere and the oceans. It is a high latitude phenomenon and can be found only in the Arctic ocean and in the oceans around Antarctica. Although the local amount of sea ice changes with the season, at any given time sea ice covers, globally, an area larger than the North American continent.

In the Arctic ocean, sea ice covers an average 14 to 16 million square kilometers in late winter and 7 to 9 million square kilometers at the end of summer. Here, the ice cover circulation is characterized by two large-scale motions. A strong transpolar drift current transports sea ice from the Bering Strait over the North Pole down to the east coast of Greenland. Then, much of the Arctic winter is characterized by a persistent high-pressure center over the western Arctic region, which stimulates a clockwise rotation of surface winds and waters, which accumulate a large amount of drifting sea ice that floats around the Beaufort Gyre.

In Antarctica, sea ice covers from 17 to 20 million square kilometers in late winter, but then the seasonal decrease is much larger, with only about 3 to 4 million square kilometers left at the end of summer. Here, sea ice is continually stirred by winds and ocean currents. Ice moves at an average speed of 10 to 20 kilometers per day, and it is divergent northward with divergences of up to 10% per day.

Floating ice can be divided into two categories, namely ice shelf and proper sea ice, which have a completely different origin. Ice shelves are floating masses of ice that originate from the land and are normally attached to the glacier that produced them, which is, in fact, terrestrial ice that has recently seeped onto the surface of the ocean. Icebergs are fragments of these ice shelves. The largest is the Ross Ice Shelf, in Antarctica, that covers an area of about 540,000 square kilometers, about the size of Texas. It contains a third of all the floating ice in the Antarctic oceans and is 90 meters thick at its hinge point and 200 meters at the front.

Sea ice is formed from ocean water that freezes. As the oceans are salty, freezing occurs at −1.8 °C. Sea ice itself is largely fresh, because the ocean salt, by a process called brine rejection, is expelled from the forming and consolidating ice. The resulting highly saline (and hence dense) water has an important influence on the ocean, overturning circulation. Fresh sea ice is formed by the cooling of the ocean as heat is lost into the atmosphere. The uppermost layer of the ocean is supercooled to slightly below the freezing point, at which time tiny ice platelets, known as frazil ice, form. Water movements push the frazil crystals against each other, causing the fragile arms of the crystals to break. Soon, a mixture of damaged crystals and crystal remnants occurs that results in an oily-looking ice soup on the surface of water, called grease ice. The name comes from its dark, oily appearance. In time, grease ice become frozen into a thin, transparent layer of ice. Still dark, like the grease ice, the first transparent stage is called dark nilas. It then grows thicker, thick enough to reflect light, and goes from dark to light. This solid white layer is referred to as light nilas. Water starts freezing at the bottom of the ice, increasing its thickness. When nilas becomes 30 cm thick, it is called first-year ice. In a single winter, first-year ice can grow up to 2.5 meters in thickness.

Sea ice influences the climate system in a number of ways, by altering the heat and chemical exchanges between the atmosphere and oceans. It has a high albedo; that is, sea ice is a strong reflector of solar radiation, and therefore energy. The growth and decay of sea ice are positive feedback processes: freezing and melting processes are self-promoting. When ice melts, for example, more solar energy is absorbed into the ocean surface waters, thereby promoting further melting. Alternately, more energy is reflected as ice growth occurs, restricting the solar radiation that warms the surface of the ocean. Less sea ice during one season may allow surface waters to store enough energy to inhibit sea ice formation in the following season. Sea ice also acts as an insulator, inhibiting exchanges of mass, heat, and momentum across the air-sea interface. The perpetual drift of sea ice extends these influences to regional and global scales. It influences the chemistry and circulation of the world's oceans by redistributing salt and freshwater and stimulating dense water formation, which facilitates overturn of the global ocean.

The seasonal sea ice cycle affects both human activities and biological habitats. For example, shipping vessels have to work quickly during periods of low ice concentration. Sea ice is the natural habitat of many mammals, like polar bears and seals, that hunt, feed, and breed on the ice. Scientists are worried that global warming might reduce ice formation, affecting the very limited food chain these animals depend on.

TECHNOLOGY DEVELOPMENT: HARDWARE DEVELOPMENT—MARINE INSTRUMENTATION LABORATORY (MIL)

Great Lakes Environmental
Research Laboratory (NOAA)

MIL's goal is to provide high quality instrumentation packages for oceanographic, biological, chemical, and water resources data collection. The highly skilled engineers and technicians of the Marine Instrumentation Laboratory (MIL) design electronics, develop schematics, manufacture circuit boards and develop computer software to support major systems design. MIL supports and maintains existing GLERL instrumentation, provides the hardware and software necessary for data acquisition systems, and provide various facility and ancillary support services.

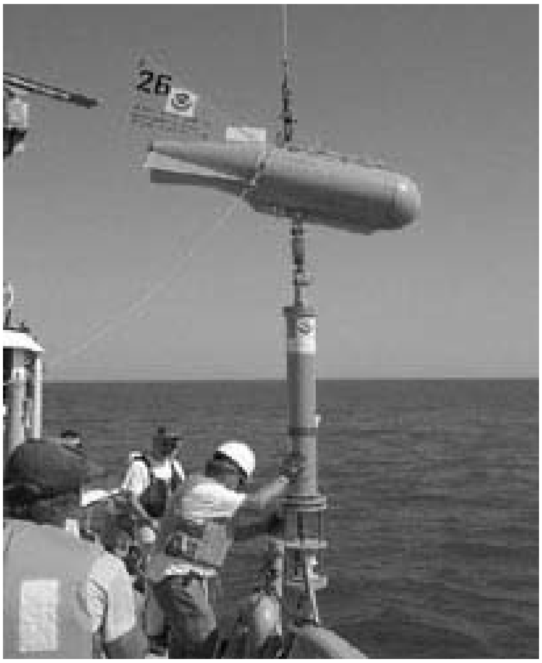

OCEANOGRAPHY

The purpose of this equipment is the measurement of water currents, temperature, water quality parameters, wave conditions and the sampling of water and sediments for subsequent analysis.

Focus on Current Measurements

MIL maintains a strong commitment to maintaining the current state of the art in this area. We currently maintain an inventory consisting of 8 RD ADCP's and several Sontek point current meters. We are currently also investigating the 3DACM current meter produced by Falmouth Scientific. MIL has also developed in house two series of satellite reporting Lagrangian drifters using Loran and GPS position sensing systems. Additionally MIL maintains a large active inventory of Vector Averaging Current Meters (VACM), Marsh-McBirney EM current meters and Neal Brown smart acoustic current meters (SACM). These aging systems are in the sunset phase of their utilization.

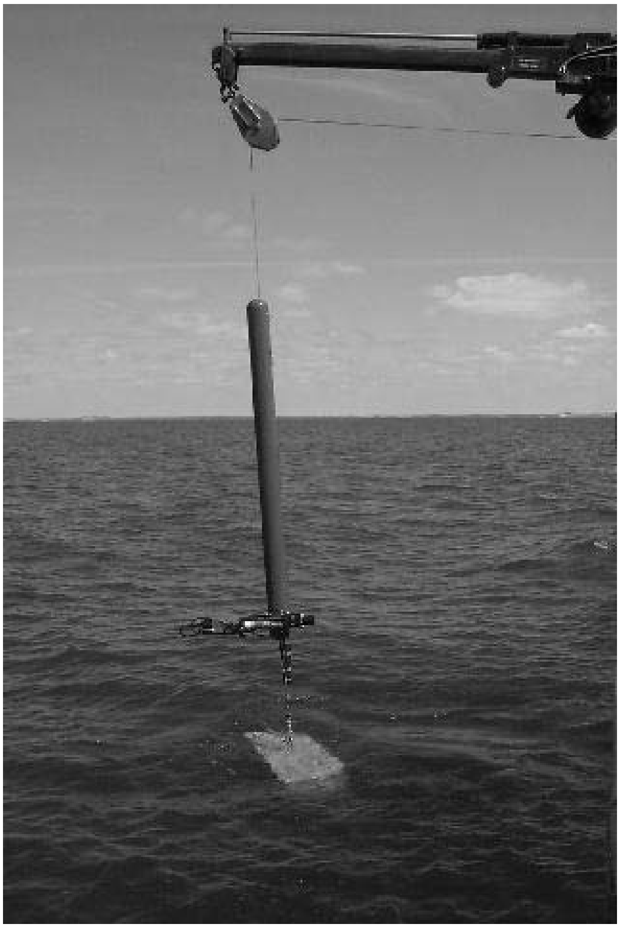

This article is a US Government work and, as such, is in the public domain in the United States of America.

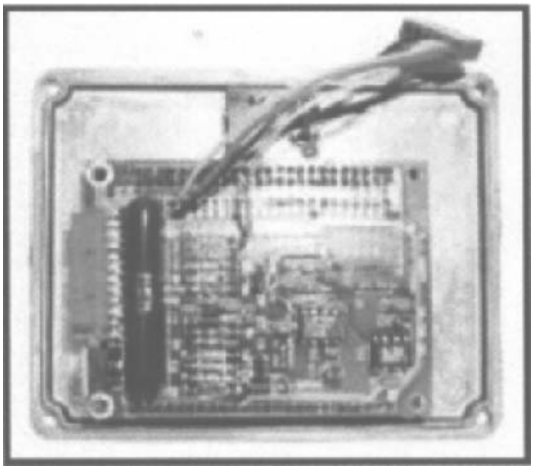

Autonomous CTD Profiler

The ACP was programmed to mechanically drive a CTD package up and down a cable through 40 meters of the water column at the rate of once per hour, taking measurements at sub-meter intervals. The ACP uses an electric motor to drive the CTD package up and down a wire rope cable. Rubber end stops attached to the cable determine the limits of travel and can be adjusted to meet the specific requirements of the experiment. The operation of the ACP is controlled by an embedded microprocessor and data is stored onto a compact flash disk. The flexibility of the programmable data logger allows specifying the interval of the cycles as well as the sample rate. The unit can sample and record data from the sensors in intervals as small as 10 cm.

Microcontrollers and Dataloggers

Recent advances in electronics have produced very small and economical computer driven controllers. These powerful tools have made it possible for instrument developers to design powerful measurement platforms with a minimum investment of time and capital. The controller recently developed for the MIL programmable sequential sediment sampler is shown.

Lagrangian Drifters

An important area of work for MIL has been the design and development of drifting current measurement systems. An example of an MIL designed drifter which used the GPS system to report locations to the ARGOS series satellites is shown. This unit is controlled with a Tattletale 2B datalogger. Currently we have 11 of these units in inventory. These units were produced significantly in advance of the availability of similar commercial units.

Sequential Sediment Samplers

MIL has designed in house a programmable 24 bottle suspended sediment sampler. This instrument is currently in it's third version. Early examples utilized an entirely in house designed controller. The current version uses a Tattletale TFX-11 datalogger. Currently there 22 units in active use with 14 committed to the EEGLE project in

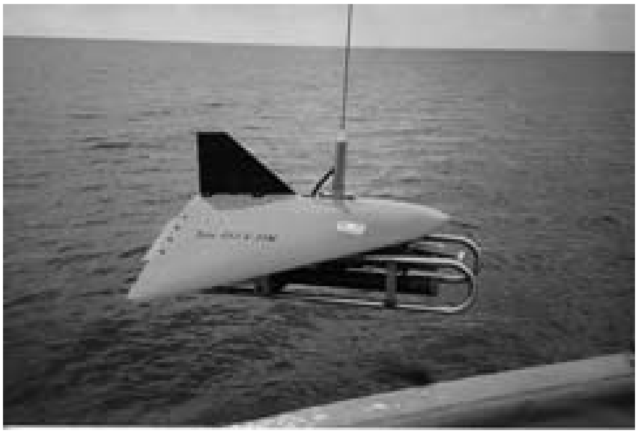

Web Cams:	Met Stations:
Muskegon, MI Alpena, MI Chicago, IL	Alpena, MI Muskegon, MI Saugatuck, MI Michigan City, IN Chicago, IL Kenosha, WI Milwaukee, WI

FY99. Several of these instruments have been distributed both nationally and internationally.

Plankton Survey System (PSS)

The Plankton Survey System (PSS) was assembled by GLERL's Marine Instrumentation Lab (MIL) in 1997 to measure distribution and abundance of zooplankton, tiny waterborne crustaceans that serve as a food base for various species of larval fish. It measures turbidity, chlorophyll a, photosynthetically active radiation (PAR), conductivity, temperature, and zooplankton distribution. The PSS and its onboard instruments has been successfully used on an ongoing basis to survey physical, chemical and biological conditions before, during and after Lake Michigan sediment resuspension events as part of the 5-year Episodic Events Great Lakes Experiment (EEGLE) program.

METEOROLOGY

WebCams and Met Stations

GLERL's Marine Instrumentation Laboratory has deployed and is maintaining a real-time network of shore-based meteorological instrument packages in Wisconsin, Michigan, Illinois, and Indiana. Stations in Illinois and Michigan also include web cameras. All meteorological observation stations measure and record wind speed, wind gust, wind direction, and air temperature. In addition, instruments in Chicago measure water levels, and instruments at Muskegon measure dew point, relative humidity, atmospheric pressure, and light level. This data as well as camera images are updated at 5-30 minute intervals and displayed on GLERL's website.

Lake Michigan Wireless Environmental Observatory

MIL is developing the first environmental observatory using wireless internet technology on the Great Lakes at GLERL's Lake Michigan Field Station in Muskegon, MI. When fully developed, the observatory will provide chemical, physical and biological data to support long-term research on the Great Lakes. These data will be available to the scientific and educational community via

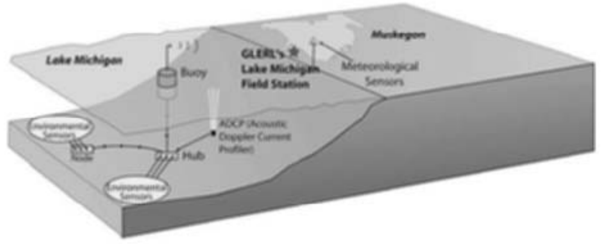

the internet through real-time observations. The goal of this project is to put in place the infrastructure needed to simplify sensor deployment and data acquisition to allow information access by scientific researchers, educators and the public. This is an important contribution to GLERL's leadership in supporting and promoting observation system development among Great Lakes universities and non-governmental organizations.

The environmental observatory consists of an offshore buoy connected to a hub that receives data from various environmental sensors such as an acoustic doppler current profiler. The data are then sent through a wireless link to an onshore receiver connected to the internet.

SEASONAL COUPLING BETWEEN INTERTIDAL MACROFAUNA AND SEDIMENT COLUMN POREWATER NUTRIENT CONCENTRATIONS

Paolo Magni
IMC-International Marine
Centre, Torregrande-Oristano,
Italy

Shigeru Montani
Hokkaido University
Hakodate, Japan

BACKGROUND

It is well established that dense bivalve assemblages play a major role in the processes of benthic nutrient regeneration in coastal marine ecosystems (1–7). In recent years, there is increasing evidence of the close relationship between nutrient concentrations in sediments and macrofauna-influenced upward flux of nutrients that enhances benthic primary productivity. This coupling has been shown, to varying degree, for a variety of benthic communities dominated by both suspension and deposit feeders. The former, including semi-infaunal mussel species, such as *Musculista senhousia* and *Modiolus americanus* (8,9), filter particulate matter from the water column; the latter, such as holothurians (10), consume high amounts of sediment bacteria, microalgae, and dead organic matter (11). Due to the complexity and multiplicity of mechanisms that govern nutrient cycling in coastal ecosystems, these studies highlight the need for integrated investigations on the variability of porewater nutrient concentrations and the composition, biology, and behavior of *in situ* macrofaunal species (12). There is still a noticeable lack of knowledge about the relationship between macrofauna and the temporal distribution and spatial variability of nutrients in sediments. Some works have conducted a seasonal study of porewater chemistry focusing on nitrogen organic compounds, such as dissolved free amino acids (13), and ammonium profiles and production (13–15). These studies have indicated that temperature has a pronounced effect on the porewater ammonium production rate and seasonal variation. In addition to the effect of environmental parameters that exhibit seasonal patterns and mineralization processes in sediments by bacteria, macrofauna also contribute to the total benthic metabolism by feeding, assimilation, and respiration. Recently, we have shown that the excretory activity of macrofauna strongly influences the magnitude and seasonal variability of the biogenic upward flux of nutrients (16; Temporal Scaling of Benthic Nutrient Regeneration in Bivalve-Dominated Tidal Flat). However, field studies on the relationship between porewater nutrient concentrations and macrofaunal communities come often from isolated surveys (17) or from transplantation/manipulation experiments (8,9). In particular, evidence of a coupling between the seasonal variability of nutrient concentrations in porewater and macrofauna-influenced upward flux of nutrients is still lacking.

SEASONALITY OF POREWATER NUTRIENT CONCENTRATIONS

We present here relevant results obtained from a multidisciplinary study carried out on a sandflat in the Seto Inland Sea (18–20). Extended information on the environmental characteristics and variability of this flat are reported in our associated papers (21–23). Emerged sediment temperature ranges widely through the year, from 2.8 °C (December) to 30.0 °C (July) in 1995 (Fig. 1).

At an individual sampling station, we investigated the macrofaunal communities (Temporal Scaling of Benthic Nutrient Regeneration in Bivalve-Dominated Tidal Flat) in parallel with the distribution of ammonium, phosphate, and silicate in porewater, monthly from January 1995 to April 1996. All these nutrient species showed marked spatial and temporal patterns and had a 10 to >30-fold variability between layers and months (Fig. 2). Ammonium varied from 75.1 (0–0.5 cm, January 1995) to

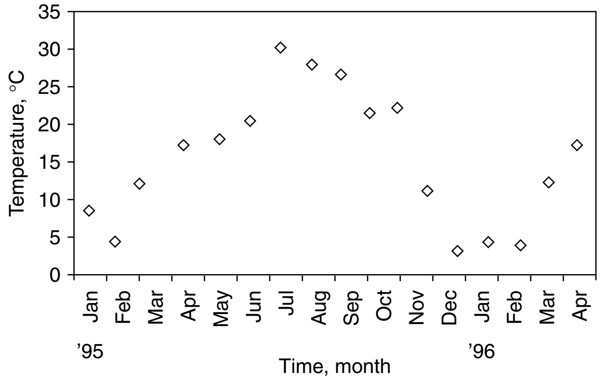

Figure 1. Seasonal variation of sediment temperature.

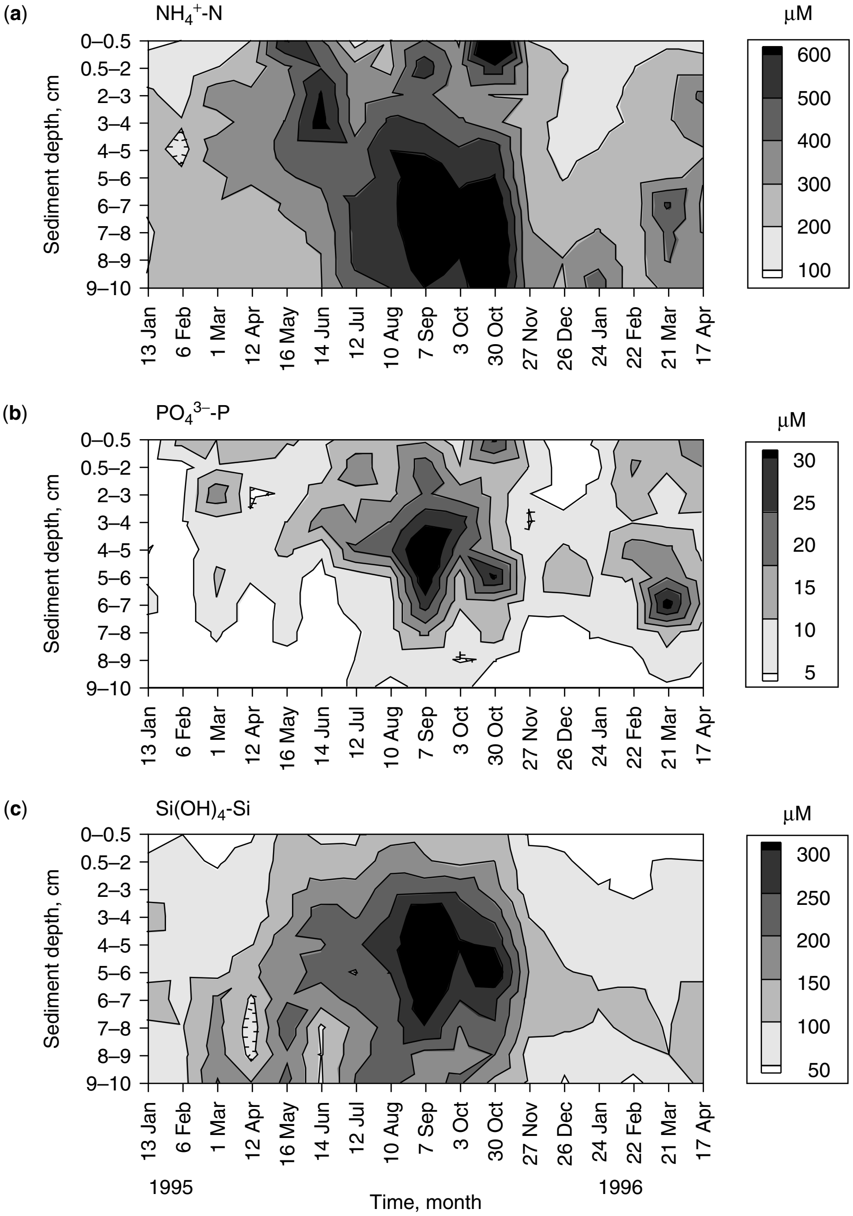

Figure 2. Spatial and seasonal variation of porewater ammonium ($NH_4{}^+$-N) (**a**), phosphate ($PO_4{}^{3-}$-P) (**b**), and silicate [$Si(OH)_4$-Si] (**c**) concentrations.

908 µM (0–0.5 cm, 30 October 1995), phosphate from 0.9 (0–0.5 cm, January 1995) to 36.9 µM (4–5 cm, September 1995), and silicate from 17.1 (0–0.5 cm, February 1996) to 379 µM (5–6 cm, 30 September 1995).

The spatial and temporal distribution of ammonium, phosphate, and silicate concentrations were consistent with each other. They were lowest in winter, progressively increased through spring and summer in the uppermost layers, and were highest between September and October 1995; a major increase occurred in intermediate layers (i.e., between 3 and 8 cm). Subsequently, minor but noticeable peaks of ammonium and phosphate concentrations were also found in March 1996, up to 518 and 32.7 µM at 6–7 cm, respectively.

For each sampling occasion and nutrient species, we summed the concentrations measured in each layer of the sediment column to be representative of an all-layer monthly pool expressed on a square meter basis. This was obtained by calculating the sediment porosity of each layer from the water content (weight loss on drying at 105°C for 20 h), assuming the bulk density of sediment particles as 2.5 g cm^{-3}. Each volume of porewater was subsequently multiplied by the relevant nutrient concentrations, which were finally expressed as areal depth-integrated values (mmol m^{-2}). Ammonium, phosphate, and silicate concentrations showed a strong correlation with each other (Fig. 3). These results suggest that similar and/or coincident processes may govern the spatial and seasonal variability of major inorganic forms of N, P, and Si in sediment porewater.

We were thus interested in assessing the existence of common environmental factors (i.e., temperature) and/or biological (i.e., macrofauna-influenced) processes that influence the variability of ammonium, phosphate, and silicate in porewater. Ammonium and silicate were highly correlated with temperature (Fig. 4). This could

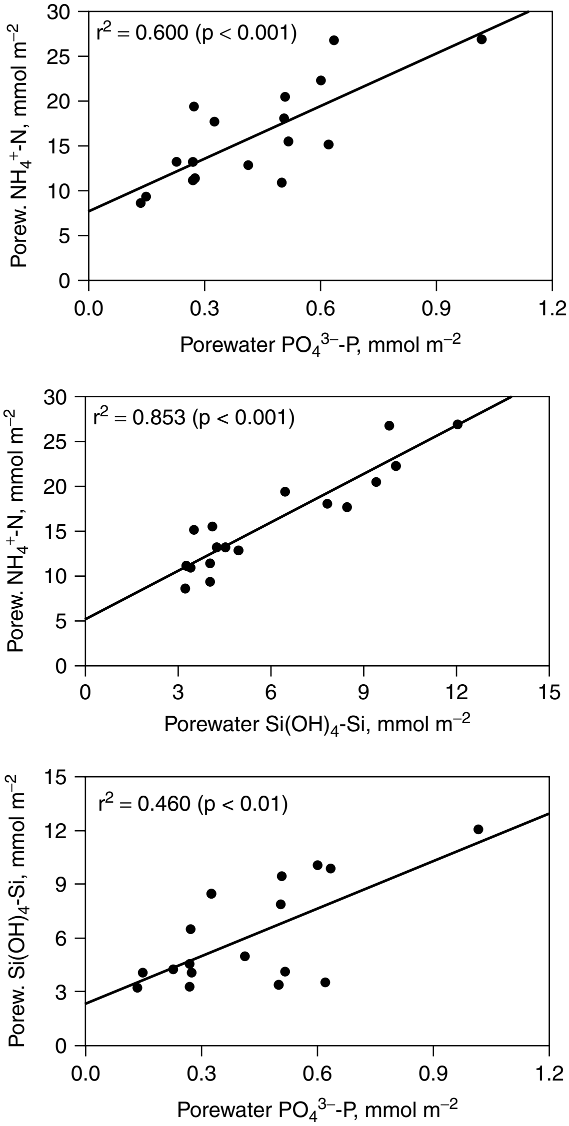

Figure 3. Relationship between porewater ammonium (NH_4^+-N), phosphate (PO_4^{3-}-P), and silicate [$Si(OH)_4$-Si] concentrations in the uppermost 10 cm of sediments.

be consistent with previous studies that have focused on the distribution of ammonium and found that it was strongly dependent on seasonal variations in temperature (13,24,25), whereas little combined information is available on the distribution of ammonium, phosphate, and silicate in the sediment column and relevant influencing factors. In our study, for instance, the correlation between temperature and phosphate was significant, yet rather weak. We then tested the hypothesis that the seasonal variation of all three nutrient species could be related to the activity of *in situ* benthic macrofauna.

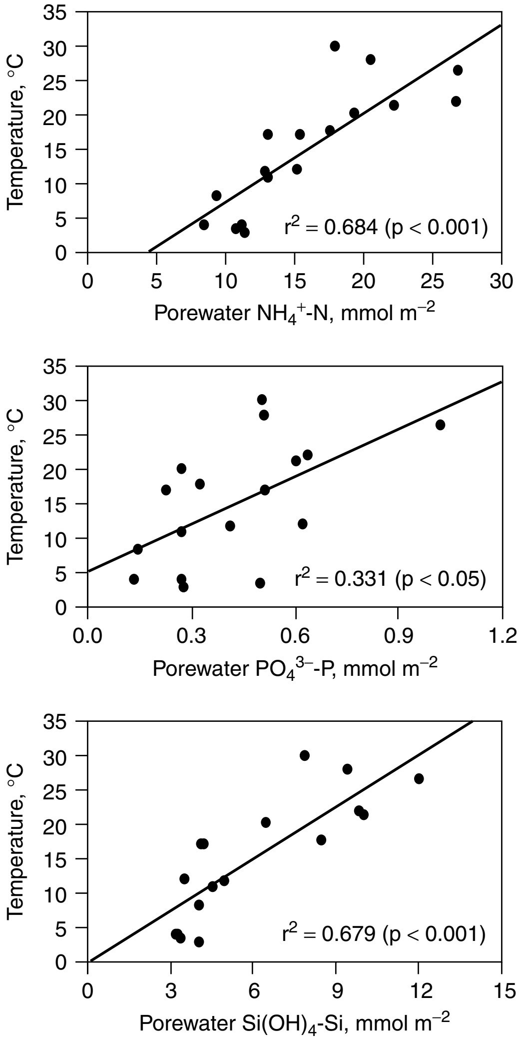

Figure 4. Relationship between porewater ammonium (NH_4^+-N), phosphate (PO_4^{3-}-P), and silicate [$Si(OH)_4$-Si] concentrations in the uppermost 10 cm sediments and temperature.

COUPLING WITH THE EXCRETORY ACTIVITY OF DOMINANT BIVALVES

We based our considerations on previous physiological measurements (Laboratory Experiments On Bivalve Excretion Rates of Nutrients) and scaling up (Temporal Scaling of Benthic Nutrient Regeneration in Bivalve-Dominated Tidal Flat) of nutrient excretion rates by two bivalve species dominant on this flat, *Ruditapes philippinarum* and *Musculista senhousia*. The plots of bivalve excretion rates of ammonium, phosphate, and silicate versus their relevant pool in the porewater showed a highly significant positive correlation in all cases (Fig. 5). These results suggest the importance of the physiological activity of the benthos on the seasonal variability of porewater chemistry. We conclude that the seasonal patterns of nutrient concentrations in the porewater are strongly coupled with the extent of biogenic regeneration of nutrients due to bivalve excretory activity. This study thus provides evidence of the influence of biological processes on the seasonal patterns of porewater nutrient distribution, suggesting a major role of macrofauna not only at the sediment–water interface, but also in the year-round processes that occur within sediments.

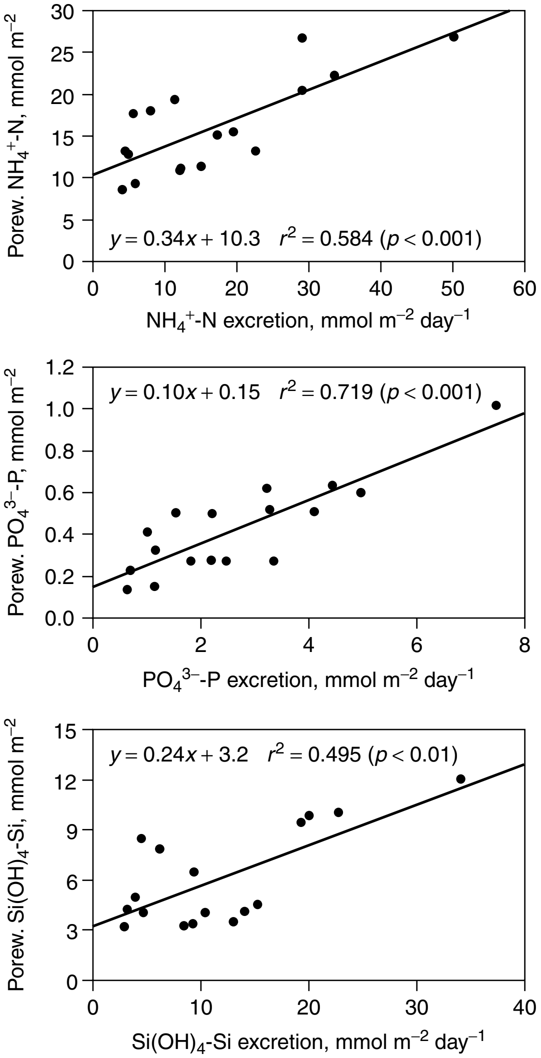

Figure 5. Relationship between porewater ammonium ($NH_4{}^+$-N), phosphate ($PO_4{}^{3-}$-P), and silicate [$Si(OH)_4$-Si] concentrations in the uppermost 10 cm of sediments and bivalve-influenced upward flux of those nutrients.

BIBLIOGRAPHY

1. Jordan, T.E. and Valiela, I. (1982). A nitrogen budget of the ribbed mussel, *Geukensia demissa*, and its significance in nitrogen flow in a New England salt marsh. *Limnol. Oceanogr.* **27**: 75–90.
2. Dame, R.F., Zingmark, R.G., and Haskin, E. (1984). Oyster reefs as processors of estuarine materials. *J. Exp. Mar. Biol. Ecol.* **83**: 239–247.
3. Murphy, R.C. and Kremer, J.N. (1985). Bivalve contribution to benthic metabolism in a California lagoon. *Estuaries* **8**: 330–341.
4. Boucher, G. and Boucher-Rodoni, R. (1988). In situ measurements of respiratory metabolism and nitrogen fluxes at the interface of oyster beds. *Mar. Ecol. Prog. Ser.* **44**: 229–238.
5. Nakamura, M., Yamamuro, M., Ishikawa, M., and Nishimura, H. (1988). Role of the bivalve *Corbicula japonica* in the nitrogen cycle in a mesohaline lagoon. *Mar. Biol.* **99**: 369–374.
6. Prins, T.C. and Smaal, A.C. (1990). Benthic-pelagic coupling: the release of inorganic nutrients by an intertidal bed of *Mytilus edulis*. In: *Trophic Relationships in the Marine Environment*. Proc. 24th Europ. Mar. Biol. Symp. Aberdeen, Univ. Press, pp. 89–103.
7. Yamamuro, M. and Koike, I. (1993). Nitrogen metabolism of the filter-feeding bivalve *Corbicula japonica* and its significance in primary production of a brackish lake in Japan. *Limnol. Oceanogr.* **35**: 997–1007.
8. Reusch, T.B.H. and Williams, S.L. (1998). Variable responses of native eelgrass *Zostera marina* to a non-indigenous bivalve *Musculista senhousia*. *Oecologia* **113**: 428–441.
9. Peterson, B.J. and Heck, K.L., Jr. (2001). Positive interaction between suspension-feeding bivalves and seagrass—a facultative mutualism. *Mar. Ecol. Progr. Ser.* **213**: 143–155.
10. Uthike, S. (2001). Interaction between sediment-feeders and microalgae on coral reefs: grazing losses versus production enhancement. *Mar. Ecol. Progr. Ser.* **210**: 125–138.
11. Uthike, S. (1999). Sediment bioturbation and impact of feeding activity of *Holothuria (Halodeima) atra* and *Stichopus chloronotus*, two sediment feeding holothurians, at Lizard Island, Great Barrier Reef. *Bull. Mar. Sci.* **64**: 129–141.
12. Christensen, B., Vedel, A., and Kristensen, E. (2000). Carbon and nitrogen fluxes in sediments inhabited by suspension-feeding (*Nereis diversicolor*) and non-suspension-feeding (*N. virens*) polychaetes. *Mar. Ecol. Prog. Ser.* **192**: 203–217.
13. Landén, A. and Hall, P.O.J. (1998). Seasonal variation of dissolved and adsorbed amino acids and ammonium in a near-shore marine sediment. *Mar. Ecol. Prog. Ser.* **170**: 67–84.
14. Blackburn, T.H. (1980). Seasonal variation in the rate of organic-N mineralization in oxic marine sediments. In: *Biogechimie de la matière organique à l'interface eau-sediment marine*. Édition du CNRS, Paris, pp. 173–183.
15. Laima, M.J.C. (1992). Extraction and seasonal variation of $NH_4{}^+$ pools in different types of coastal marine sediments. *Mar. Ecol. Prog. Ser.* **82**: 75–84.
16. Magni, P., Montani, S., Takada, C., and Tsutsumi, H. (2000). Temporal scaling and relevance of bivalve nutrient excretion on a tidal flat of the Seto Inland Sea, Japan. *Mar. Ecol. Progr. Ser.* **198**: 139–155.
17. Lomstein, B.A., Blackburn, T.H., and Henriksen, K. (1989). Aspects of nitrogen and carbon cycling in the northern Bering Shelf sediment: I. The significance of urea turnover in the mineralization of ammonium ion. *Mar. Ecol. Prog. Ser.* **57**: 237–248.
18. Magni, P. (1998). A multidisciplinary study on the dynamics of biophilic elements (C, N, P, Si) in a tidal estuary of the Seto Inland Sea, Japan: physico-chemical variability and macrozoobenthic communities. Ph.D. Thesis, The United Graduate School of Ehime University, Japan, p. 258.
19. Montani, S. et al. (1998). The effect of a tidal cycle on the dynamics of nutrients in a tidal estuary in the Seto Inland Sea, Japan. *J. Oceanogr.* **54**: 65–76.
20. Magni, P., Montani, S., and Tada, K. (2002). Semidiurnal dynamics of salinity, nutrients and suspended particulate matter in an estuary in the Seto Inland Sea, Japan, during a spring tide cycle. *J. Oceanogr.* **58**: 389–402.
21. Magni, P. and Montani, S. (1997). Development of benthic microalgal assemblages on an intertidal flat in the Seto Inland Sea, Japan: effects of environmental variability. *La mer.* **35**: 137–148.
22. Magni, P. and Montani, S. (1998). Responses of intertidal and subtidal communities of the macrobenthos to organic

load and oxygen depletion in the Seto Inland Sea, Japan. *J. Rech. Océanogr.* **23**: 47–56.

23. Magni, P. and Montani, S. (2000). Physical and chemical variability in the lower intertidal zone of an estuary in the Seto Inland Sea, Japan: seasonal patterns of dissolved and particulate compounds. *Hydrobiol.* **432**: 9–23.
24. Blackburn, T.H. and Henriksen, K. (1983). Nitrogen cycling in different types of sediments from Danish waters. *Limnol. Oceanogr.* **28**: 477–493.
25. Klump, J.V. and Martens, C.S. (1989). Seasonality of nutrient regeneration in an organic-rich coastal sediment: kinetic modeling of changing pore-water nutrient and sulfate distribution. *Limnol. Oceanogr.* **34**: 559–577.

MAPPING THE SEA FLOOR OF THE HISTORIC AREA REMEDIATION SITE (HARS) OFFSHORE OF NEW YORK CITY

Bradford Butman
U.S. Geological Survey

Repeated surveys using a multibeam mapping system document changes in the topography and distribution of sediments on the sea floor caused by placement of dredged material, remedial capping, and natural processes.

INTRODUCTION

The area offshore of New York City has been used for the disposal of dredged material for over a century. The area has also been used for the disposal of other materials such as acid waste, industrial waste, municipal sewage sludge, cellar dirt, and wood. Between 1976 and 1995, the New York Bight Dredged Material Disposal Site, also known as the Mud Dump Site (MDS), received on average about 6 million cubic yards of dredged material annually. In September 1997 the MDS was closed as a disposal site, and it and the surrounding area were designated as the Historic Area Remediation Site (HARS) (Figs. 1 and 2). The sea floor of the HARS, approximately 9 square nautical miles in area, currently is being remediated by placing a minimum 1-m-thick cap of clean dredged material on top of the surficial sediments that are contaminated from previous disposal of dredged and other materials. The U.S. Geological Survey (USGS) is working cooperatively with the U.S. Army Corps of Engineers (USACE) to map the sea floor geology of the HARS and changes in the characteristics of the surficial sediments over time.

HIGH-RESOLUTION SURVEYS OF THE SEA FLOOR OF THE HARS

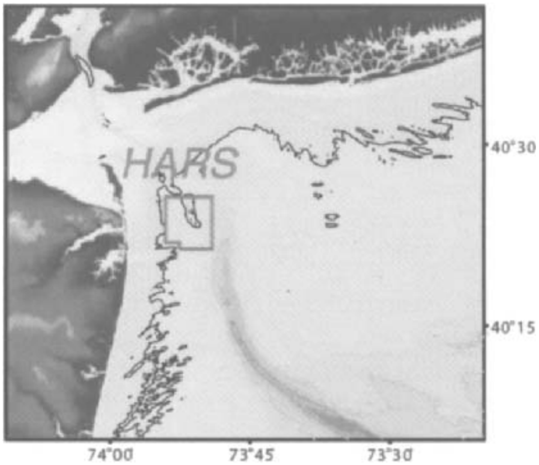

Figure 1. Map showing the area offshore of New York and New Jersey that has been used for the disposal of dredged materials and other wastes since the late 1800's. The Historic Area Remediation Site (HARS) is outlined in red.

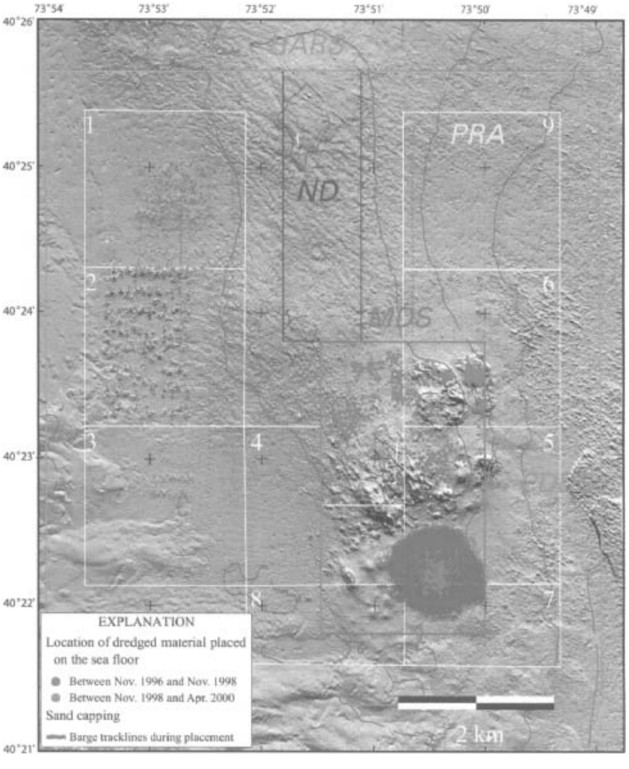

Figure 2. Shaded relief image of the Historic Area Remediation Site (HARS) in April 2000 showing the Primary Remediation Area (PRA, divided into nine cells), the no discharge zone (ND), the former Mud Dump Site (MDS), and the discontinued Cellar Dirt Disposal Site (CDDS). Companion images are shown in Figures 4 and 5. See text for a description of this image and major features. Bathymetric contour interval is 5 m (red lines).

Surveys of the HARS were conducted in November 1996 (prior to the closing of the Mud Dump Site), November 1998 (during early remediation of the HARS), and April 2000 (during continued remediation of the HARS) using a Simrad EM1000 multibeam mapping system (Fig. 3). Survey lines were run approximately 100 m apart to provide full coverage of the sea floor. The EM1000 measured the depth of water (to an accuracy of

---

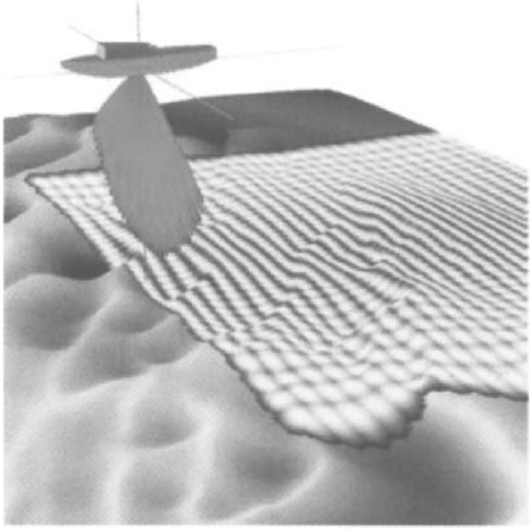

Figure 3. High-resolution multibeam mapping systems use sound from an array of transducers to measure water depth and sediment characteristics of the sea floor. The horizontal resolution of the maps is a few meters, providing an image of the sea floor topography and sediment properties somewhat comparable to an aerial photograph.

about 30 cm) as well as the intensity of sound reflected from the sea floor, which is referred to as backscatter intensity. High backscatter intensity generally indicates the presence of rocks and coarse-grained sediments, while low backscatter intensity indicates the presence of finer grained sediments. Direct observations using bottom photographs, video, and grab samples are needed to verify interpretations of the sea floor geology based on backscatter intensity.

IMAGES OF THE HARS SEA FLOOR

In this fact sheet, the topography and backscatter intensity data measured by the multibeam mapping system are presented in three types of images. Each of these images highlights different features and characteristics of the sea floor.

(1) A *shaded relief image* (Fig. 2) visually shows small topographic features (with relief of a few meters) that could not be effectively shown by contours alone at this scale. The image was created by vertically exaggerating the topography four times and then artificially illuminating the relief by a light source positioned 45° above the horizon from the north. In this image, topographic features are enhanced by strong illumination on the north-facing slopes and by shadows cast on the south-facing slopes.

(2) A *shaded relief image, colored by backscatter intensity*, combines the high-resolution view of topography with a measure of sediment characteristics (Figs. 4, 5B, and 6). In these images, the backscatter intensity is represented by a suite of eight colors ranging from blue, which represents low intensity (fine-grained sediments), to red, which represents high intensity (rock outcrops and coarse-grained sediments). These data are draped over the shaded relief image. The resultant image displays light and dark intensities within each color band that result from a feature's position with respect to the light source. For example, north-facing slopes, receiving strong illumination, show as a light intensity within a color band, whereas south-facing slopes, being in shadow, show as a dark intensity within a color band.

(3) A *shaded relief image, colored by bathymetry*, combines the high-resolution view of topography with color to show water depth (Fig. 5A).

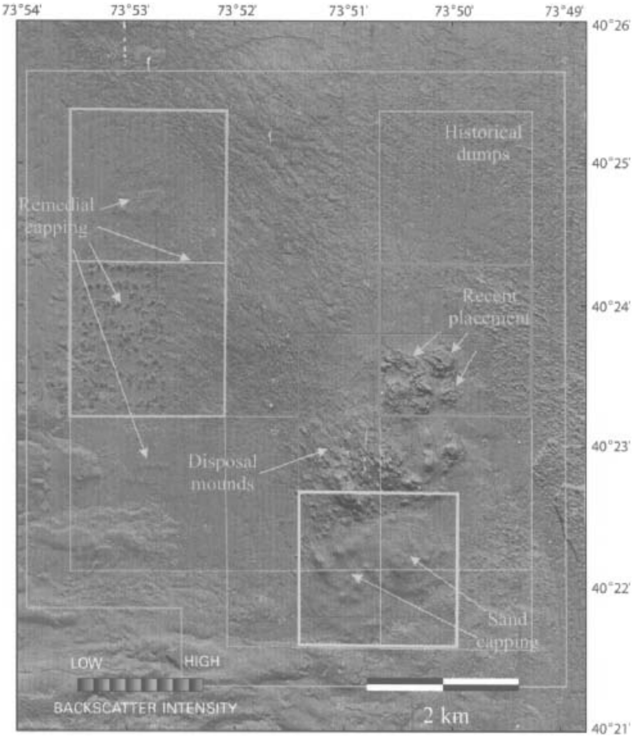

Figure 4. Pseudo-colored backscatter intensity and shaded relief map of the entire HARS in April 2000. The faint north-trending stripes run parallel to the survey tracklines and are artifacts of data collection and environmental conditions. The pink, green, and yellow boxes outline areas shown in Figure 6 to illustrate changes in backscatter intensity between 1996, 1998, and 2000. See text for a description of this image and major features. Bathymetric contour interval is 5 m (red lines).

THE SEA FLOOR OF THE HARS

Within the HARS, one of the most striking aspects of the sea floor is the variability in backscatter intensity and bottom morphology over distances of a few kilometers or less, caused by both natural and anthropogenic processes. This fact sheet presents companion images showing the sea floor of the HARS as mapped in April 2000 in plan view (Figs. 2, 4, and 6) and in perspective view (Fig. 5). Images of selected areas in 1996, 1998, and 2000 illustrate changes over time (Fig. 6). Major features of the sea floor of the HARS shown in these images include two

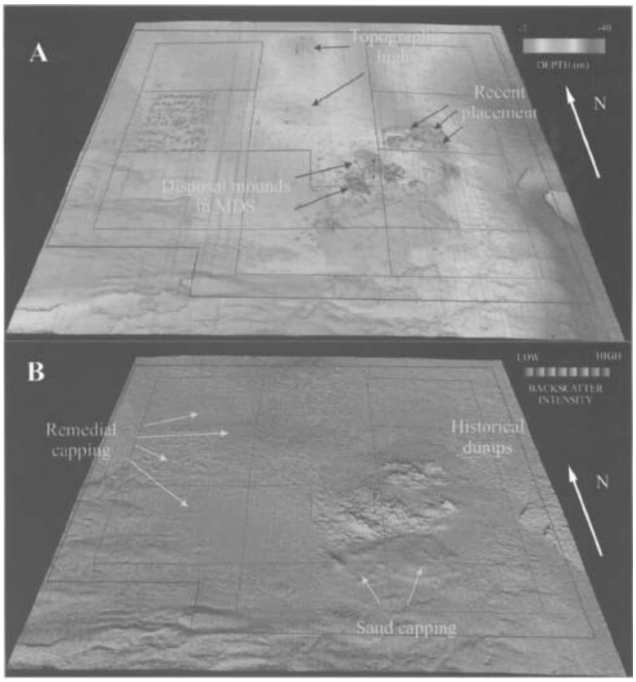

Figure 5. Perspective view of the Historic Area Remediation Site, looking from south to north, based on the multibeam survey carried out in April 2000. A, Shaded relief map with color-coded bathymetry. B, Backscatter intensity draped over shaded relief (see text for a description of the color scheme). The north-trending stripes, running parallel to the survey tracklines, are artifacts of data collection and environmental conditions. The topography, surface features, and the surficial sediments of the HARS have been heavily influenced by the disposal of dredged and other material in this region over the last century, and by recent remedial capping. See text for a description of these images and major features.

relatively smooth topographic highs composed of material dumped in the late 1800's and early 1900's ("Topographic highs" in Fig. 5A); mounds of material in the Mud Dump Site ("Disposal mounds in MDS" in Fig. 5A); two circular features where contaminated sediments were placed and then capped with sand, one in the late 1980's, and the other in 1997 ("Sand capping" in Figs. 4 and 5B and "Previous capping" and "New sand capping" in Fig. 6); material deposited between the November 1996 and November 1998 survey ("Recent placement" in Figs. 4 and 5A); many features about 50 m in size interpreted to be individual dumps of material ("Historical dumps" in Figs. 4 and 5B); and material placed as part of remediation activities ("Remedial capping" in Figs. 4, 5B, and 6).

CHANGES IN SURFICIAL PROPERTIES BETWEEN 1996, 1998, AND 2000

Comparison of the topography and backscatter intensity from the three multibeam surveys show how the area changed as a result of dredged material placed before the Mud Dump Site was closed and ongoing remediation of the HARS (see Fig. 2 for locations of placed material from USACE records).

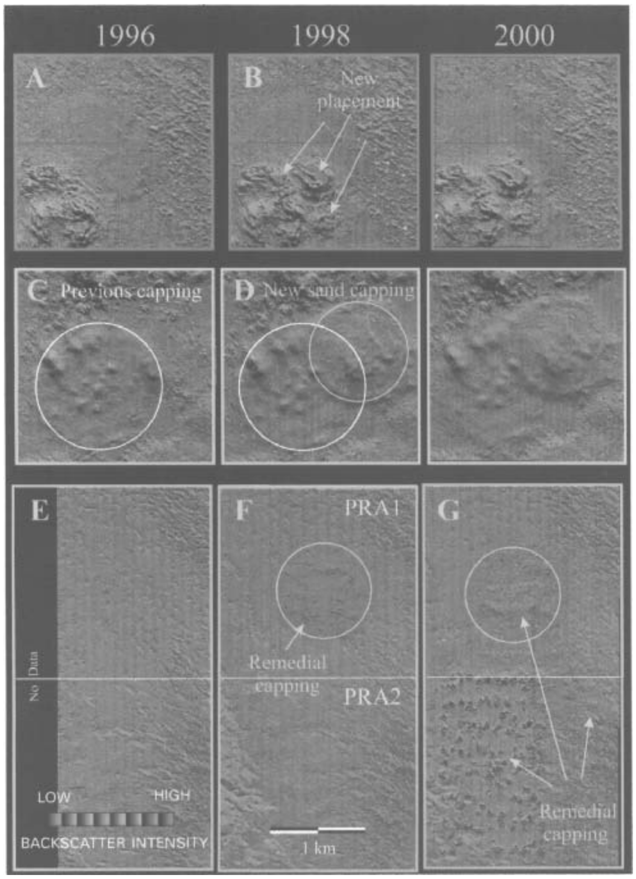

Figure 6. Pseudo-colored backscatter intensity and shaded relief map of parts of the HARS in 1996, 1998, and 2000 (see Fig. 4 for location). These images illustrate changes in the sea floor topography and backscatter intensity that occurred between 1996 and 1998 and between 1998 and 2000 caused by placement of dredged material and by remedial capping. See text for a description of these images and major features.

Between 1996 and 1998, changes include (1) mounds of medium backscatter intensity dredged material in the northeastern corner of the MDS, some as high as 6 m, placed between November 1996 and September 1997 (compare panels A and B, Fig. 6); (2) a circular area of low-backscatter intensity material about 1 km in diameter and 2 m thick in the southern part of the MDS associated with sand capping (compare panels C and D, Fig. 6); and (3) a circular area of low backscatter intensity material in PRA1 associated with remedial capping (compare panels E and F, Fig. 6).

Between 1998 and 2000, changes include (compare panels F and G, Fig. 6) (1) increased backscatter intensity in PRA1 due to additional placement of material and consolidation, de-watering, and possible winnowing of the previous cover; (2) a series of crater-like features in the western part of PRA2, 30 to 70 m long and on the order of 20 m wide with elevated rims and central depressions, that were apparently formed as remedial material impacted the soft sediments on the sea floor; and (3) an area of reduced backscatter intensity in the northeastern corner of PRA2 caused by the placement of remedial material.

Resolution limits of the multibeam system, and the amount of material placed over a relatively large area, preclude using the repeated topographic surveys for determining the amount of material placed on the sea floor. However these multibeam data clearly show the overall regional geology and, through comparison of topography and backscatter intensity, document the location of placed material and changes in sediment properties over time.

NOAA AND UNIVERSITY SCIENTISTS STUDY METHYL BROMIDE CYCLING IN THE NORTH PACIFIC

Shari Yvon-Lewis
Atlantic Oceanographic and Meteorological Laboratory

Kelly Goodwin
Sara Cotton
University of Miami

James Butler
Climate Monitoring and Diagnostics Laboratory

Daniel King
University of Colorado

Eric Saltzman
Ryszard Tokarczyk
University of Miami

Patricia Matrai
Brian Yocis
Eileen Loiseau
Bigelow Laboratory for Ocean Sciences

Georgina Sturrock
Commonwealth Scientific and Industrial Research Organization—Australia

As part of a study supported by both NASA and NOAA, scientists from two NOAA laboratories, three universities and CSIRO participated in a research cruise aboard the R/V Ronald H. Brown. The cruise departed Kwajalein, Republic of the Marshall Islands on 14 September 1999 and arrived in Seattle, Washington on 23 October 1999 with stops in Honolulu, Hawaii, Dutch Harbor, Alaska, and Kodiak, Alaska. The objective of this research effort was to obtain reliable measurements of the uptake and emission of methyl bromide and other climatically important halocarbons in tropical to temperate regions of the North Pacific Ocean.

Atmospheric methyl bromide (CH_3Br), which is of both natural and anthropogenic origin, has been identified as a Class I ozone-depleting substance in the amended and adjusted Montreal Protocol on Substances that Deplete Stratospheric Ozone. The role of the ocean in regulating the atmospheric burden of this gas is still somewhat uncertain. Methyl bromide is both produced and destroyed in the ocean through chemical and biological processes. The organisms or reactions that produce CH_3Br at rates sufficient to explain its observed concentrations are not known. Degradation has been shown to occur at rates that are faster than can be explained by known chemical degradation reactions, and evidence suggests that this additional degradation is bacterial consumption of CH_3Br. While recent measurements have shown that, on the whole, the ocean is a net sink for CH_3Br, measurement coverage to date has been limited and sporadic, which restricts our ability to map the spatial and temporal variations that are necessary for understanding how the system will respond to perturbations (e.g., Global Warming).

The measurements made during this cruise are designed to help improve our understanding of the role that the oceans play in the cycling of CH_3Br. The program involved instrumentation from two NOAA

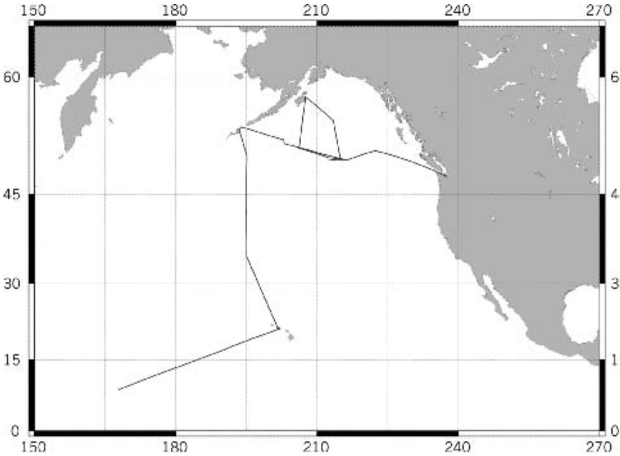

Figure 1. Cruise Track.

Figure 2. Scientists collecting water samples for production and degradation incubations.

---

laboratories and two universities. Measurements were made of the concentrations of CH_3Br and a suite of natural and anthropogenic halocarbons in the air and surface water, degradation rates of CH_3Br in the surface water, production rates of CH_3Br and other natural halocarbons in the surface water, and depth profiles of CH_3Br and other halocarbons. The combined results from these measurements will be used to constrain the budget of CH_3Br in these waters at this time of year. The relative importance of the biological and chemical processes will be examined for tropical and high latitudes. Attempts will also be made to extract relationships between the rates and concentrations measured and satellite measurements in order to develop proxies that can provide global coverage on shorter time scales. At this time, there is insufficient data to examine seasonal and long-term trends in net flux, production, or degradation. Until satellite measurable proxies can be found, additional research cruises are needed to reduce the uncertainty in the global net flux estimate and to map the spatial and temporal variations in the net fluxes, production rates, and degradation rates of CH_3Br and other climatically important halocarbons.

TIDALLY MEDIATED CHANGES IN NUTRIENT CONCENTRATIONS

PAOLO MAGNI
IMC—International Marine Centre
Torregrande-Oristano, Italy

SHIGERU MONTANI
Hokkaido University
Hakodate, Japan

Freshwater runoff during ebb flow and salt water intrusion during the flood may have a major effect on short-term changes in nutrient (ammonium, nitrate+nitrite, phosphate, and silicate) concentrations along an estuary. Time series hourly measurements conducted in a mixed-semidiurnal type estuary (i.e., characterized by two major lower and higher tidal levels) show that these changes are a strong function of both tidal state (e.g., low vs. high tide) and amplitude (e.g., neap vs. spring tide). In particular, the changes in nutrient concentrations are higher during ebb than during flood tide and largest between the lower low tide and the higher high tide of a spring tide. Finally, the importance of investigating simultaneously different stations along the estuarine spine is highlighted, in addition to studying the nutrient distribution based on selected salinity intervals which may reflect only the conditions at a particular tidal state.

BACKGROUND

An important aspect of the high variability of tidal estuaries is related to the effect of the tidal cycle on the physical and chemical characteristics of the water. In particular, on a timescale of hours, freshwater runoff during ebb flow and salt water intrusion during the flood may determine strong changes in salinity and dissolved and particulate compounds. Several studies in riverine and estuarine waters have investigated the distribution of nutrient (e.g., ammonium, nitrate, phosphate, and silicate) concentrations, based on the salinity gradient in the system at a particular tidal state (1–3). Accordingly, plots of nutrients versus salinity are often used to assess the source of different nutrient species, whether from inland, outside the estuary, or within it (3–7).

An evaluation of the distributional pattern of nutrients along an estuary has important ecological implications in relation to the cycling of biophilic elements, such as N (nitrogen), P (phosphorous), and Si (silicon). It is known that numerous processes influence the behavior of nutrients, whether they show conservative mixing or reflect removal or addition along an estuary. Ammonium consumption and ammonium oxidation, for instance, are predominant in the water column, whereas denitrification in sediments is responsible for nitrate removal from the water column (8–10). By contrast, bioturbation and excretion by abundant benthic animals may greatly contribute to the upward flux of regenerated nutrients, such as ammonium and phosphate, which in turn enhance primary production (11,12). Accordingly, it has been shown that regeneration processes within an estuary are consistent with often encountered nonconservative mixing of ammonium (4,8,13,14). This corresponds to the tendency of ammonium concentration to be high at midsalinity ranges, resulting in a poor correlation with salinity. In contrast, nitrate tends to show conservative behavior, as evidence of its riverine origin (2,14,15), although addition (15) or removal (7) is also found. Moreover, it must be considered that in some cases, nitrate versus salinity plots may fail to unravel active nitrate turnover, leading to an approximate balance of sources and sinks (16). As for silicate, a general pattern indicates that estuarine mixing of this nutrient species tends to be conservative (2,7,13). Yet, either silicate removal (7,17) or addition (4) occurs in relation to the development of an algal population in rivers or to a closer interaction with estuarine sediments, respectively, and varies with season (14). A major upward flux of silicate from sediments might also be related to the biological activity and excretory processes of abundant macrofaunal assemblages (18,19).

In addition to these general considerations, the distribution and cycling of nutrients depend strongly on the specific characteristics of each estuary, including water residence time and water depth, nutrient levels, and the extent of salt-water intrusion. Uncles and Stephens (3) showed that saline intrusion was a strong function of the tidal state and a weaker function of freshwater inflow. Accordingly, Balls (2) indicated that conservative mixing of phosphate, is a function of estuarine flushing time, as related to particle–water interaction and chemical speciation (20). In particular, phosphate removal at low salinities may be due to adsorption to iron and aluminium colloidal oxyhydroxides that aggregate and undergo sedimentation (7).

laboratories and two universities. Measurements were made of the concentrations of CH_3Br and a suite of natural and anthropogenic halocarbons in the air and surface water, degradation rates of CH_3Br in the surface water, production rates of CH_3Br and other natural halocarbons in the surface water, and depth profiles of CH_3Br and other halocarbons. The combined results from these measurements will be used to constrain the budget of CH_3Br in these waters at this time of year. The relative importance of the biological and chemical processes will be examined for tropical and high latitudes. Attempts will also be made to extract relationships between the rates and concentrations measured and satellite measurements in order to develop proxies that can provide global coverage on shorter time scales. At this time, there is insufficient data to examine seasonal and long-term trends in net flux, production, or degradation. Until satellite measurable proxies can be found, additional research cruises are needed to reduce the uncertainty in the global net flux estimate and to map the spatial and temporal variations in the net fluxes, production rates, and degradation rates of CH_3Br and other climatically important halocarbons.

TIDALLY MEDIATED CHANGES IN NUTRIENT CONCENTRATIONS

Paolo Magni
IMC—International Marine Centre
Torregrande-Oristano, Italy

Shigeru Montani
Hokkaido University
Hakodate, Japan

Freshwater runoff during ebb flow and salt water intrusion during the flood may have a major effect on short-term changes in nutrient (ammonium, nitrate+nitrite, phosphate, and silicate) concentrations along an estuary. Time series hourly measurements conducted in a mixed-semidiurnal type estuary (i.e., characterized by two major lower and higher tidal levels) show that these changes are a strong function of both tidal state (e.g., low vs. high tide) and amplitude (e.g., neap vs. spring tide). In particular, the changes in nutrient concentrations are higher during ebb than during flood tide and largest between the lower low tide and the higher high tide of a spring tide. Finally, the importance of investigating simultaneously different stations along the estuarine spine is highlighted, in addition to studying the nutrient distribution based on selected salinity intervals which may reflect only the conditions at a particular tidal state.

BACKGROUND

An important aspect of the high variability of tidal estuaries is related to the effect of the tidal cycle on the physical and chemical characteristics of the water. In particular, on a timescale of hours, freshwater runoff during ebb flow and salt water intrusion during the flood may determine strong changes in salinity and dissolved and particulate compounds. Several studies in riverine and estuarine waters have investigated the distribution of nutrient (e.g., ammonium, nitrate, phosphate, and silicate) concentrations, based on the salinity gradient in the system at a particular tidal state (1–3). Accordingly, plots of nutrients versus salinity are often used to assess the source of different nutrient species, whether from inland, outside the estuary, or within it (3–7).

An evaluation of the distributional pattern of nutrients along an estuary has important ecological implications in relation to the cycling of biophilic elements, such as N (nitrogen), P (phosphorous), and Si (silicon). It is known that numerous processes influence the behavior of nutrients, whether they show conservative mixing or reflect removal or addition along an estuary. Ammonium consumption and ammonium oxidation, for instance, are predominant in the water column, whereas denitrification in sediments is responsible for nitrate removal from the water column (8–10). By contrast, bioturbation and excretion by abundant benthic animals may greatly contribute to the upward flux of regenerated nutrients, such as ammonium and phosphate, which in turn enhance primary production (11,12). Accordingly, it has been shown that regeneration processes within an estuary are consistent with often encountered nonconservative mixing of ammonium (4,8,13,14). This corresponds to the tendency of ammonium concentration to be high at midsalinity ranges, resulting in a poor correlation with salinity. In contrast, nitrate tends to show conservative behavior, as evidence of its riverine origin (2,14,15), although addition (15) or removal (7) is also found. Moreover, it must be considered that in some cases, nitrate versus salinity plots may fail to unravel active nitrate turnover, leading to an approximate balance of sources and sinks (16). As for silicate, a general pattern indicates that estuarine mixing of this nutrient species tends to be conservative (2,7,13). Yet, either silicate removal (7,17) or addition (4) occurs in relation to the development of an algal population in rivers or to a closer interaction with estuarine sediments, respectively, and varies with season (14). A major upward flux of silicate from sediments might also be related to the biological activity and excretory processes of abundant macrofaunal assemblages (18,19).

In addition to these general considerations, the distribution and cycling of nutrients depend strongly on the specific characteristics of each estuary, including water residence time and water depth, nutrient levels, and the extent of salt-water intrusion. Uncles and Stephens (3) showed that saline intrusion was a strong function of the tidal state and a weaker function of freshwater inflow. Accordingly, Balls (2) indicated that conservative mixing of phosphate, is a function of estuarine flushing time, as related to particle–water interaction and chemical speciation (20). In particular, phosphate removal at low salinities may be due to adsorption to iron and aluminium colloidal oxyhydroxides that aggregate and undergo sedimentation (7).

Therefore, it is important that investigations of the distribution of nutrients as a function of salinity are conducted on proper spatial and temporal scales that take into account the extent and variability of salt-water intrusion. Time series surveys represent a valuable approach to quantifying such variability in estuarine waters (21,22). In particular, short-term monitoring surveys during a complete tidal cycle and simultaneous information at different stations can be a powerful tool for evaluating the extent of salt-water intrusion along the estuary and its direct impact on the water chemistry of an estuary. This approach may overcome the limitations of evaluating the distribution and behavior of nutrients based on a particular tidal state at arbitrarily selected salinity intervals. Major drawbacks when working across long distances to cover different salinity ranges include work time to travel between samplings, tidal-water displacements, variations in tidal velocities, and a possible wide range of river discharge. Simultaneous observations at different stations are expected to track changes in nutrient concentrations that occur along the estuarine spine in relation to the extent of salt-water intrusion over time.

Case study

This article reviews the results of 24-hour surveys in a tidal estuary of the Seto Inland Sea, Japan, during a spring tide of May 1995 (13,23). It is aimed to give an example of the effect of freshwater runoff and salt-water intrusion on the spatial and temporal variability of nutrient concentrations during a tidal cycle. The effect of a tide is also evaluated in relation to the *mixed-semidiurnal* behavior (i.e., with pronounced differences between two successive low and high tides) of the estuary. The fieldwork was conducted along a transect line of approximately 1.4 km linking the river to the rear to the subtidal zone. Multiprobe casts were used for hydrologic measurements. Nutrient concentrations were determined every hour in surface waters, simultaneously at a riverine, an intertidal, and a subtidal station, and every two hours at additional depths through the water column at the subtidal station. Details of the sampling scheme and analysis are given in Montani et al. (13) and Magni et al. (23).

At the beginning of the survey, there was a marked salinity gradient along the estuary; surface water salinity was 2.0, 18.5, and 31.3 psu at the riverine, intertidal, and subtidal stations, respectively (Fig. 1). As the lower low tide approached, salinity remained low at the riverine station (Fig. 1a), sharply decreased also at the intertidal station (Fig. 1b) and, subsequently, at the subtidal station (Fig. 1c). By contrast, soon after the lower low tide, salt-water intrusion rapidly caused an abrupt increase in salinity at both the intertidal and the riverine stations, up to >30 psu (Figs. 1a,b). During the flood, there was a homogeneous distribution of high-salinity water along the transect line. At the subtidal station, a major change in salinity as a function of the tidal cycle was also apparent, but restricted to the surface layer, whereas salinity remained constantly >31 psu below the surface (Figs. 1c,d).

The nutrient concentrations were also markedly affected by the tidal cycle. At the beginning of the survey, silicate and nitrate+nitrite concentrations were markedly higher at the riverine station than at the intertidal and subtidal stations, whereas the ammonium concentration was relatively higher at the intertidal station (Figs. 1a, b, c). Approaching the lower low tide, the nutrient concentrations in surface water increased rapidly, especially at the intertidal and subtidal stations. Differently at the riverine station, the ammonium concentrations remained low, suggesting no significant import of this nutrient species through freshwater inflow (Fig. 1a). At the subtidal station, the nutrient concentrations also showed a relatively consistent increase below the surface, yet progressively less noticeable with depth (Fig. 1d).

By contrast, during the flood, as high-salinity water flushed backward into the estuary, the nutrient concentrations dropped to the lowest values at all stations and depths (Figs. 1a, b, c, d); a 7.5-fold and 8.8-fold decrease of silicate and nitrate+nitrite concentrations, respectively, occurred at the riverine station. During the second part of the survey, after the higher high tide, both salinity and nutrient changes were less marked.

Figure 2 summarizes the relationships between salinity and nutrient concentrations. Salinity versus nutrient plots demonstrate that the distributional pattern of nutrients largely varied with station, depth, and the different nutrient species. In particular at the riverine station, silicate and nitrate+nitrite were negatively correlated with salinity; r^2 explained a large portion of total variance (i.e., $r^2 = 0.879$ and 0.796, respectively). By contrast, at this station, ammonium showed a positive correlation with salinity, whereas phosphate did not significantly correlate with salinity. Differently at the intertidal and subtidal stations, the concentrations of all nutrient species in surface waters were negatively and significantly correlated with salinity; levels of confidence varied from $p < .05$ (ammonium) to $p < .001$ (phosphate and silicate).

The variability of both salinity and nutrient concentrations was lowest at the subtidal station below the surface. This test data set (Fig. 2d) indicated that all nutrient species were correlated positively with salinity at a high level of significance ($p < .001$). Relevant plots also highlighted that, within such a restricted variability of salinity, silicate and nitrate+nitrite concentrations comprised narrower values than those of phosphate and ammonium (Fig. 2d). Accordingly, the model equation for the former two nutrient species explained a higher portion of the total variance ($r^2 = 0.423$, and $r^2 = 0.457$, respectively) than that explained by the latter ($r^2 = 0.221$, and $r^2 = 0.245$, respectively).

These results showed that the riverine input was a major source of silicate and, partially, nitrate+nitrite and phosphate. It was also apparent that the increase in nutrient concentrations at the intertidal station and subsequently at the subtidal station, was largest during the first part of the survey (Figs. 1b, c). Companion papers demonstrated that the intertidal zone also plays a major role in nutrient cycling, as a major site of nutrient regeneration within the estuary (12,14,24).

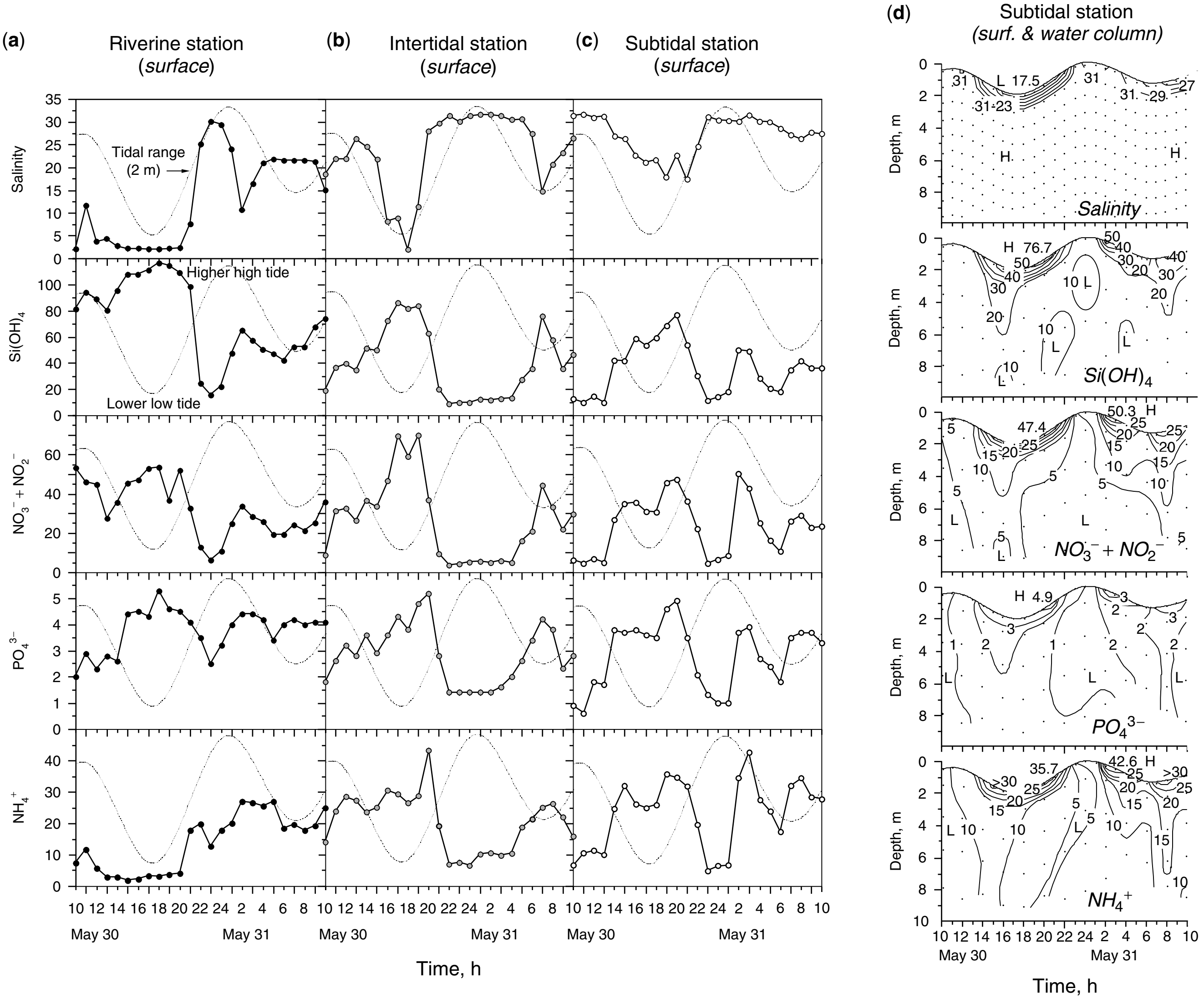

Figure 1. Time series of salinity (psu, practical salinity unit) and nutrient concentration (μM) in surface water at a riverine, intertidal, and subtidal station (Fig. 1a, b, c) and through the water column at the subtidal station (Fig. 1d) during a tidal cycle of a spring tide in a *mixed-semidiurnal* type estuary (Seto Inland Sea, Japan). Data sources: Fig. 1a, b, c from Montani et al. (13) (redrawn from Figs. 4 & 5); Fig. 1d: from Magni et al. (23) (adapted from Figs. 2 & 3).

CONCLUDING REMARKS

This article showed that the effect of salt-water intrusion on the dynamics of nutrients varies strongly both spatially (on relatively short distances) and temporally (on an hour timescale), and that this is much dependent on the tidal state. In particular, it was shown that the effect of tidal amplitude is important in determining the extent of the variations in nutrient concentrations, which were stronger between the lower low tide and the higher high tide. It also indicated that nutrient concentrations were higher during the ebb than during the flood and highest at the surface layer, as strongly correlated inversely with salinity. Finally, this work highlighted the importance of considering simultaneous investigations at different stations along the estuarine spine during a tidal cycle, especially on short distances, besides studying the nutrient dynamics based on selected salinity intervals that may reflect only the situation at a particular tidal state.

BIBLIOGRAPHY

1. Hernández-Ayón, J.M., Galindo-Bect, M.S., Flores-Baez, B.P., and Alvarez-Borrego, S. (1993). Nutrient concentrations are high in the turbid waters of the Colorado River delta. *Estuar. Coast. Shelf Sci.* **37**: 593–602.
2. Balls, P.W. (1994). Nutrient inputs to estuaries from nine Scottish east coast rivers; influence of estuarine processes on inputs to the North Sea. *Estuar. Coast. Shelf Sci.* **39**: 329–352.
3. Uncles, R.J. and Stephens, J.A. (1996). Salt intrusion in the Tweed Estuary. *Estuar. Coast. Shelf Sci.* **43**: 271–293.
4. Balls, P.W. (1992). Nutrient behaviour in two contrasting Scottish estuaries, the Forth and the Tay. *Oceanol. Acta* **15**: 261–277.

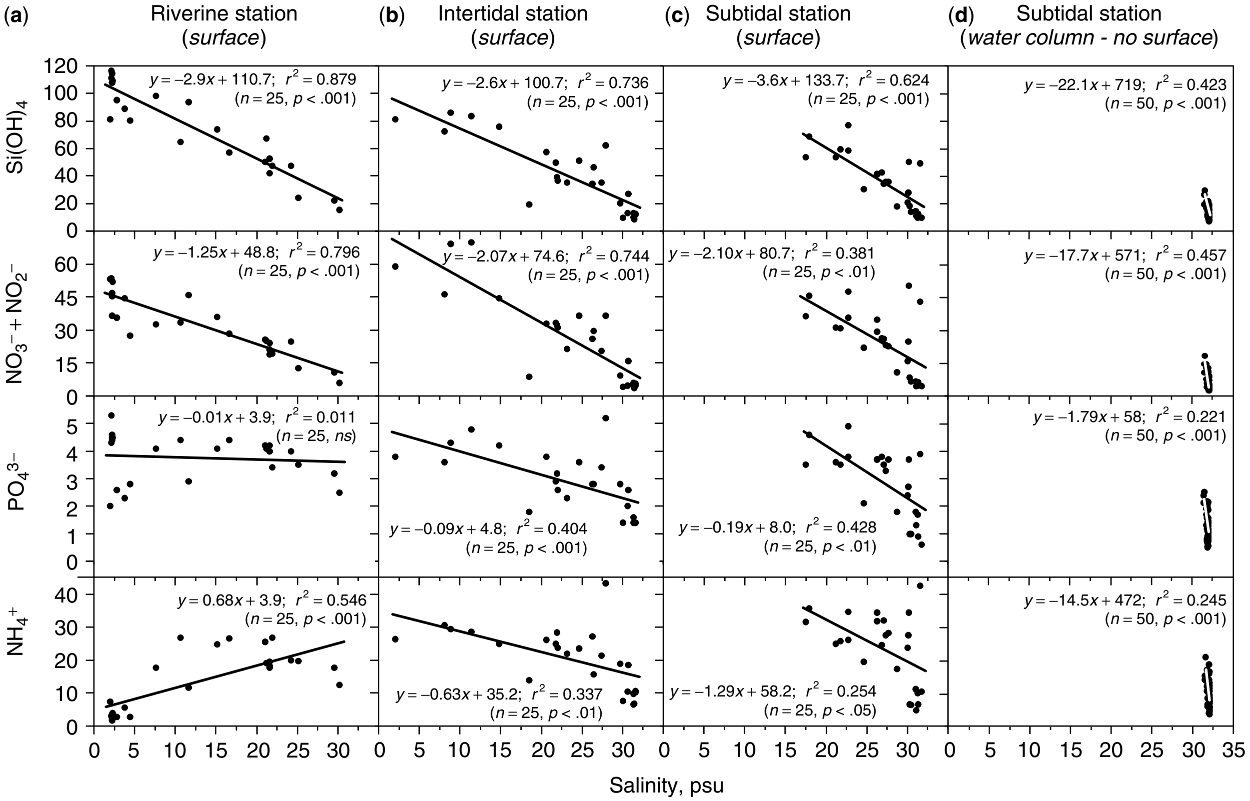

Figure 2. Plots of salinity (psu) versus nutrient concentration (μM) originated from the time series in Fig. 1. Data sources: Fig. 1a, b, c from Montani et al. (13) (redrawn from Figs. 8 to 11); Fig. 1d: after Magni et al. (23).

5. Clark, J.F., Simpson, H.J., Bopp, R.F., and Deck, B. (1992). Geochemistry and loading history of phosphate and silicate in the Hudson estuary. *Estuar. Coast. Shelf Sci.* **34**: 213–233.
6. Page, H.M., Petty, R.L., and Meade, D.E. (1995). Influence of watershed runoff on nutrient dynamics in a southern California salt marsh. *Estuar. Coast. Shelf Sci.* **41**: 163–180.
7. Eyre, B. and Twigg, C. (1997). Nutrient behaviour during post-flood recovery of the Richmond river estuary northern NSW, Australia. *Estuar. Coast. Shelf Sci.* **44**: 311–326.
8. Soetart, K. and Herman, P.M.J. (1995). Nitrogen dynamics in the westerschelde (the Netherlands) using a box model with fixed dispersion coefficients. *Hydrobiologia* **311**: 215–224.
9. Nixon, S.W. et al. (1996). The fate of nitrogen and phosphorus at the land–sea margin of the North Atlantic Ocean. *Biogeochem.* **35**: 141–180.
10. Nedwell, D.B., Jickells, T.D., Trimmer, M., and Sanders, R. (1999). Nutrients in estuaries. *Adv. Ecol. Res.* **29**: 43–92.
11. Herman, P.M.J., Middelburg, J.J., van de Koppel, J., and Heip, C.H.R. (1999). Ecology of estuarine macrobenthos. *Adv. Ecol. Res.* **29**: 195–240.
12. Magni, P., Montani, S., Takada, C., and Tsutsumi, H. (2000). Temporal scaling and relevance of bivalve nutrient excretion on a tidal flat of the Seto Inland Sea, Japan. *Mar. Ecol. Prog. Ser.* **198**: 139–155.
13. Montani, S. et al. (1998). The effect of a tidal cycle on the dynamics of nutrients in a tidal estuary in the Seto Inland Sea, Japan. *J. Oceanogr.* **54**: 65–76.
14. Magni, P. and Montani, S. (2000). Water chemistry variability in the lower intertidal zone of an estuary in the Seto Inland Sea, Japan: Seasonal patterns of nutrients and particulate compounds. *Hydrobiologia* **432**: 9–23.
15. Middelburg, J.J. and Nieuwenhuize, J. (2000). Uptake of dissolved inorganic nitrogen in turbid, tidal estuaries. *Mar. Ecol. Prog. Ser.* **192**: 79–88.
16. Middelburg, J.J. and Nieuwenhuize, J. (2001). Nitrogen isotope tracing of dissolved inorganic nitrogen behaviour in tidal estuaries. *Estuar. Coast. Shelf Sci.* **53**: 385–391.
17. Liss, P.S. and Pointon, M.J. (1973). Removal of dissolved boron and silicon during estuarine mixing of sea and river waters. *Geochim. Cosmochim. Acta* **37**: 1493–1498.
18. Bartoli, M. et al. (2001). Impact of *Tapes philippinarum* on nutrient dynamics and benthic respiration in the Sacca di Goro. *Hydrobiologia* **455**: 203–212.
19. Magni, P. and Montani, S. (2004). Magnitude and temporal scaling of biogenic nutrient regeneration in a bivalve-dominated tidal flat. In: *Encyclopedia of Water*. John Wiley & Sons, New York.
20. Froelich, P.N. (1988). Kinetic control of dissolved phosphate in natural rivers and estuaries: A primer on the phosphate buffer mechanism. *Limnol. Oceanogr.* **33**: 649–668.
21. Yin, K., Harrison, P.J., Pond, S., and Beamish, R.J. (1995). Entrainment of nitrate in the Fraser river estuary and its biological implications. II. Effects of spring *vs.* neap tide and river discharge. *Estuar. Coast. Shelf Sci.* **40**: 529–544.
22. Yin, K., Harrison, P.J., Pond, S., and Beamish, R.J. (1995). Entrainment of nitrate in the Fraser River estuary and its biological implications. III. Effects of winds. *Estuar. Coast. Shelf Sci.* **40**: 545–558.

23. Magni, P., Montani, S., and Tada, K. (2002). Semidiurnal dynamics of salinity, nutrients and suspended particulate matter in an estuary in the Seto Inland Sea, Japan, during a spring tide cycle. *J. Oceanogr.* **58**: 389–402.
24. Montani, S., Magni, P., and Abe, N. (2003). Seasonal and interannual patterns of intertidal microphytobenthos in combination with laboratory and areal production estimates. *Mar. Ecol. Prog. Ser.* **249**: 79–91.

THE ROLE OF OCEANS IN THE GLOBAL CYCLES OF CLIMATICALLY-ACTIVE TRACE-GASES

Robert C. Upstill-Goddard
University of Newcastle upon Tyne
Newcastle upon Tyne, United Kingdom

The oceans are important in the global cycles of a range of trace gases that influence atmospheric chemistry and climate. For some of these, the oceans are a net source to the atmosphere, whereas for others, they are a net sink. Major gases of interest are summarized in Table 1, along with their net flux directions and their principal roles in the troposphere and stratosphere. Carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O) are major greenhouse gases: CO_2 currently accounts for more than one-half of enhanced global warming, whereas CH_4 and N_2O, respectively, account for about 15% and 6% (1). Volatile sulphurs are also implicated in climate forcing and they play important roles in atmospheric chemistry. As a result of length restrictions, our discussion focuses only on these gases.

GLOBAL PARTITIONING OF ANTHROPOGENIC CO_2

Several mechanisms contribute to natural CO_2 cycling between the atmosphere and the Earth's surface. The largest natural exchanges occur through respiration and photosynthesis on land and in the oceans, and by solubility-driven uptake in the oceans, and the net result of these exchanges is a natural carbon cycle in overall balance. By comparison with these natural fluxes, the flux of anthropogenic CO_2, which is derived primarily from the burning of fossil fuels, cement production, deforestation, and other land-use changes, is rather small, $\sim 7 \pm 1$ Gt C per year (1) (1 Gt = 1 gigatonne = 10^{15} g), which is nevertheless large enough to significantly disturb the natural CO_2 cycle.

As a result of its unreactive nature, the residence time of CO_2 in the troposphere is of the order of 50–200 years, hence anthropogenic CO_2 tends to accumulate there. In fact, tropospheric CO_2 has risen from about 280 ppmv (parts per million by volume) preindustrially, as determined from ice core studies (2), to approaching 380 ppmv by mid-2004 (http://www.cmdl.noaa.gov/). However, this corresponds to somewhat less than one-half of the known anthropogenic release (1). During the 1980s (the latest decade for which estimates of all carbon sources and sinks are available), the mean rate of tropospheric CO_2 growth was only about 3.3 ± 0.1 Gt C per year, the remainder, about 3.7 ± 1.0 Gt C per year, having been absorbed by "sinks" located in the oceans and within the terrestrial biosphere (1). The fraction of anthropogenic CO_2 absorbed by these sinks is, however, not constant. Large fluctuations are evident in the continuous tropospheric records that date back to the late 1950s (1), and these fluctuations are believed to relate directly to short-term variations in global climate. For example, increased atmospheric growth rates correlate with El Niño climate warming events, whereas cooling periods, such as that which followed the eruption of Mount Pinatubo in the early 1990s, seem to be associated with reduced atmospheric growth. These variations are thought principally to reflect changes in the balance of terrestrial primary production (photosynthesis) versus respiration, decomposition, and the combustion of organic material (3).

High background variability has precluded directly measuring the relative magnitudes of the oceanic and terrestrial CO_2 sinks, hence they have hitherto been estimated using models, often with conflicting results (4). However, recent techniques based on measuring carbon and oxygen isotopes in air (5) now enable the partitioning of anthropogenic CO_2 between these reservoirs to be determined with greater certainty, and it is now generally agreed that the ocean and land sinks are of about the same magnitude (although the uncertainties are large), i.e., about 1.9 ± 0.6 Gt C per year (6).

Table 1. Some Atmospherically Active Trace Gases with Important Marine Sources or Sinks

Gas		Net Flux	Effect in the Atmosphere	Troposphere	Stratosphere
Carbon Dioxide	(CO_2)	Into Ocean	Infra-red activity	✓	
Methane	(CH_4)	Out of Ocean	Infra-red activity Atmospheric redox Ozone regulation	✓	
Nitrous Oxide	(N_2O)	Out of Ocean	Infra-red activity Ozone regulation	✓	
Carbon Monoxide	(CO)	Out of Ocean	Atmospheric redox	✓	
Dimethyl sulphide	(DMS)	Out of Ocean	Acidity, Cloud formation	✓	
Carbonyl Sulphide	(COS)	Out of Ocean	Sulphate aerosol (cooling)	✓	✓
Carbon Disulphide	(CS_2)	Out of Ocean	Source of COS	✓	✓
Organohalogens (Natural)		Out of Ocean	Atmospheric redox	✓	✓
Nonmethane hydrocarbons		Out of Ocean	Atmospheric redox		

MECHANISMS OF OCEAN CO_2 UPTAKE

Compared with many other atmospheric trace gases, CO_2 is rather soluble in seawater, where it occurs as three principal dissolved species that together comprise seawater dissolved inorganic carbon (DIC), which are bicarbonate (HCO_3^-, 91% of DIC), carbonate (CO_3^{2-}, 8% of DIC), and dissolved CO_2 (1% of DIC).

The capacity of the oceans to absorb atmospheric CO_2 is ultimately buffered by the DIC system. CO_2 uptake results in an increase both in the partial pressure of CO_2 (pCO_2) and in the concentration of HCO_3^-, which is produced through the reaction of CO_2 with CO_3^{2-}. With further CO_2 uptake, its conversion to HCO_3^- becomes limited by the decreasing CO_3^{2-} availability, the result being that more CO_2 remains in solution, further decreasing the capacity for CO_2 uptake at increasing tropospheric levels. As an illustration, for atmospheric CO_2 to rise by about 100 ppmv relative to today, the increase in seawater DIC would only be about 60% of that which has accompanied the approximately 100 ppmv rise in tropospheric CO_2 since the industrial revolution (1,6).

Studies based on the downward penetration of chemical tracers, such as ^{14}C, coupled with simple box models of ocean mixing (7), or more sophisticated General Circulation Models (8), show that tropospheric CO_2 equilibrates with the surface ocean mixed layer rather rapidly, on the order of a few years. In contrast, CO_2 penetration into the deep ocean interior is constrained by relatively slow vertical transport; the whole ocean equilibrates with the atmosphere on a timescale of more than 1000 years. As a result of this slow rate of downward mixing, with few exceptions, anthropogenic CO_2 has yet to penetrate below about 1000 m depth. Transport of CO_2 into the ocean interior occurs via the thermohaline circulation, in which cool surface waters in the high latitude North Atlantic (Greenland Sea) and Southern Oceans sequester CO_2 from the troposphere and sink because of their higher density. This downward transport is balanced in regions where deep water with a high capacity for CO_2 uptake wells up to the surface, such as occurs in parts of the tropics and the Southern Ocean. The sinking waters eventually spread laterally toward the subtropics, giving rise to a relatively uniform distribution of anthropogenic CO_2. Eventually, on the 1000-plus year timescale of deep ocean mixing, the deep CO_2-rich waters formed in this way will again well up to the surface, the principal site for this being the Equatorial Pacific. As this water warms during upwelling, its pCO2 will increase, resulting in CO_2 loss to the atmosphere by out gassing.

One other potential removal process for tropospheric CO_2 is through reaction with $CaCO_3$ contained in deep sea sediments; however, their response time to changes in tropospheric CO_2 is several thousand years (9).

According to some coupled ocean-atmosphere models, one possible consequence of global warming is an increase in the intensity of vertical stratification in the oceans (1), which would reduce the rate of surface to deep water mixing and, consequently, the uptake rate of tropospheric CO_2.

In addition to chemically and physically driven uptake, tropospheric CO_2 is also processed through the so-called "biological pump," in which organic matter produced via photosynthesis and cycled through the upper ocean food-web is ultimately transported downward via sinking organic particles or through vertical biomass migrations. This sinking flux of organic carbon is remineralized or respired back to inorganic carbon at depth with an accompanying release of dissolved inorganic nutrients (principally nitrate and phosphate). Model simulations suggest that without the biological pump, tropospheric CO_2 could be about 150 ppmv higher than at present (1). However, the likely response of the biological pump to increasing tropospheric CO_2 is uncertain. Most current evidence tends to discount increased productivity on the grounds of limitation by the supply of nutrients, which are seasonally depleted in most surface waters. However, extensive regions exist of the subarctic Pacific, equatorial Pacific, and Southern Ocean with abundant nitrate and phosphate throughout the year but with very low phytoplankton productivity, the so-called high-nutrient-low-chlorophyll (HNLC) regions. Their low productivity likely reflects a deficiency in the supply of a minor nutrient such as iron (10). Although predicting climate-induced changes in the supply rate of iron to the oceans is far from straightforward, these regions could conceivably play a significant role in the future ocean uptake of anthropogenic CO_2. It is also conceivable that global warming-induced changes in stratification described earlier could also modify the ocean's biological carbon cycle; however, the consequences of this are difficult to predict given the intrinsic complexity of the system (11).

THE GLOBAL ROLE OF CH_4 AND N_2O

As well as being important greenhouse gases, CH_4 and N_2O play important roles in atmospheric chemistry; N_2O is involved in the stratospheric cycling of NO_x (reactive nitrogen oxides) and ozone (12), and CH_4 takes part in reactions that govern levels of tropospheric ozone and hydroxyl radical (· OH) and stratospheric H_2O (13). Like CO_2, N_2O and CH_4 are currently increasing in the troposphere, both by about 0.3% per year (1). However, the role of the oceans in their global budgets differs from that for anthropogenic CO_2 in two fundamental ways. First, the oceans are a net source of tropospheric N_2O and CH_4, and second, for both, the marine source is one of several global sources whose relative magnitudes remain rather uncertain.

GLOBAL SOURCE UNCERTAINTIES

The marine system is one of two major N_2O sources, the other being terrestrial soils. Combustion, biomass burning, and fertilizers make additional minor contributions. Although the uncertainties are large, the oceans are thought to contribute around 20% of the natural global source (1); however, the contribution is larger with anthropogenic sources included (see below). For CH_4, many more global sources exist, in descending order of magnitude, these are: agriculture, wetlands, fossil fuels, biomass burning, termites, oceans, CH_4 hydrates, and landfills. However, the range of uncertainties is no better than it is for N_2O, ranging from ±100% (e.g., wetlands), to in excess of ±2000% (e.g., CH_4 hydrates) (1). According to these data, marine waters contribute only about 3% of the total CH_4 source to the troposphere, but this may be an underestimate (see below).

SOURCES OF SEAWATER N_2O

In seawater, N_2O develops as a byproduct during microbial nitrification (the conversion of dissolved ammonium, NH_4^+, to nitrate, NO_3^-) and as a reactive intermediate during microbial denitrification (the reduction of NO_3^- to gaseous nitrogen). Nitrification principally occurs in oxygenated waters, although it is inhibited by light, whereas denitrification is restricted to anoxic sediments and O_2-deficient waters. For these reasons, N_2O production is insignificant in most open ocean surface waters. In extremely O_2-deficient waters, N_2O can be consumed during denitrification because of its use as an electron acceptor by denitrifying bacteria (14). Coupling of the two processes, in which the NO_3^- developing from nitrification is consumed by denitrification, occurs both in marine sediments (15) and around the fringes of O_2-depleted waters in the open ocean (16).

The net rate of N_2O production by nitrification and denitrification is influenced by several factors in addition to dissolved O_2 availability, including the supply rates of NO_3^- and NH_4^+, the composition of the microbial ecosystem, and in sediments, physicochemical aspects such as porosity and grain size. Consequently, both processes show pronounced seasonal variability. These aspects, coupled with a nonuniform distribution of N_2O source regions, makes the total marine source of N_2O difficult to quantify.

Current data indicate around two-thirds of the marine source of tropospheric N_2O to derive from the open ocean, with the remainder coming from coastal waters. However, these estimates contain uncertainties resulting from incomplete spatial and seasonal sampling and difficulties related to estimating sea-to-air fluxes.

Open ocean emissions are approximately equally distributed between the northern and southern hemispheres. Most of the N_2O is located below the surface-mixed layer and highest concentrations occur at around 500–1000 m depth, where a high O_2 demand results from the bacterial decomposition of sinking organic particles. Exchange of this water with the atmosphere is usually slow, except during wintertime when the surface-mixed layer deepens because of cooling and wind-driven mixing, entraining waters from below. Regions experiencing strong seasonal upwelling are especially strong sources of tropospheric N_2O, which include the Tropical North Pacific, the Arabian Sea/northwestern Indian Ocean, the equatorial upwelling, and along the coasts of northwest Africa and western central and South America. In these areas, the upwellings bring N_2O-rich waters to the surface along with a plentiful supply of nutrients that fuels high primary productivity. The resultant large downward flux of organic particles gives rise to strongly O_2-deficient waters that replenish the deep water inventory of N_2O through nitrification-denitrification coupling (17). Some areas of the subtropical gyres and the North Atlantic seem to be weak sinks for tropospheric N_2O in winter and weak sources in summer.

The major coastal N_2O source regions are estuaries, open coastal shelf waters being generally at or close to equilibrium with tropospheric N_2O levels. Both water column nitrification associated with high suspended particle populations (18) and sediment denitrification (19) have been identified in estuaries.

Whereas oceanic N_2O emissions are considered to be wholly natural, significant N_2O in coastal waters appears to derive indirectly from anthropogenic activity. In particular, the use of fertilizers is reflected in enhanced transport of nitrogen to coastal waters and a consequent increase in N_2O production and tropospheric flux. As much as 90% of current estuarine N_2O emissions and 25% of continental shelf emissions may be anthropogenic and consistent with the geographic distribution of fertilizer use, human population, and atmospheric nitrogen deposition, more than 80% of these anthropogenic sources are located in the Northern Hemisphere mid-latitudes between 20°N and 66°N (20).

SOURCES OF SEAWATER CH_4

Like N_2O, seawater CH_4 has a microbial source (microbial methanogenesis). Methanogenesis is inhibited by dissolved O_2 and therefore usually occurs in anoxic sediments or in waters that are strongly O_2-depleted. Even so, CH_4 concentrations in the oxygenated surface ocean are on average about 30% above the tropospheric equilibrium value, most likely reflecting methanogenesis by O_2-tolerant methanogens inside bacterially maintained "anoxic microniches" in the guts of zooplankton and/or in particles (21). Consequently, the open ocean represents a small CH_4 source. In addition, regions of much higher CH_4 concentration occur in upwelling areas, associated with enhanced primary productivity as for N_2O (16,22). Based on the available open ocean data, the marine contribution to tropospheric CH_4 is about 10 Tg CH_4 per year, which is equivalent to about 3% of the total global source (1).

Coastal waters have much higher CH_4 concentrations than the open ocean but have, until recently, been excluded from global CH_4 source estimates because of a lack of data. For estuaries, values of 100 to 200 times the background equilibrium value are common. Such CH_4 levels reflect direct inputs from rivers, coastal seawater, underlying sediments, and *in situ* production/consumption from water column methanogenesis and microbial CH_4 oxidation. Recent work suggests that correctly accounting for these regions could increase the estimated marine CH_4 source by around 50% (23).

A potentially even larger CH_4 source may be geologically sourced CH_4 from natural marine seeps, which are most common on shallow continental shelves (24). Seeps are episodic in nature, and the CH_4 fluxes developing from them are predominantly by bubbles, which complicates making accurate measurements. As a result of CH_4 losses because of bubble dissolution and subsequent CH_4 oxidation in the water column, shallow water seeps are much more effective than deep water seeps at contributing CH_4 to the troposphere. Revised source estimates that take account of seep occurrences suggest that the total marine source of tropospheric CH_4 could be as much as 40 Tg yr^{-1} (25). If so, the oceans could be a much more important source of tropospheric CH_4 than previously thought.

VOLATILE SULPHUR COMPOUNDS

The marine system is an important source of sulphur globally; in particular, it is the principal source of biogenic atmospheric sulphate aerosol, which plays important roles in atmospheric chemistry and climate.

The predominant volatile sulphur in surface seawater is dimethyl sulphide (DMS) [$(CH_3)_2S$], which is a byproduct of algal metabolism and accounts for about 95% of marine sulphur emissions and 20% of total global sulphur emissions (1). DMS is rapidly oxidized by free radicals in the lower troposphere. Sulphur dioxide (SO_2) is the major reaction product and is subsequently transformed to sulphate aerosol through gas-to-particle reactions. This process is hypothesized to impact directly the radiative forcing of global climate, primarily through changes to cloud albedo (26). Carbonyl sulphide (COS) has also been implicated in global climate forcing. Compared with other volatile sulphurs, COS has a long tropospheric residence time of around 2–6 years. Consequently, it is transported into the stratosphere, where its photo-oxidation is believed to be an important source of sulphate aerosol, which is thought to impact Earth's radiation balance (27) and stratospheric ozone levels (28).

DIMETHYL SULPHIDE

A global database of more than 15,000 measurements of surface seawater (29) revealed distinct annual DMS cycles in the open ocean at mid to high latitudes. In the northern hemisphere, open ocean DMS increases during spring–summer, whereas in the southern hemisphere, concentrations peak six months later. These patterns relate to the timing of phytoplankton blooms and seasonal changes in mixed layer depth. In contrast, tropical regions show weak seasonality; DMS is elevated in the upwelling regions off western Africa and South America, but these concentrations are lower than those at high latitudes during summer. On coastal shelves, DMS is spatially and temporally variable, broadly correlating with seasonal primary productivity and the presence of algal blooms associated with upwelling at water mass boundaries (hydrographic fronts) (29). Similar concentrations are found in many estuaries, notable exceptions being those with high concentrations of suspended particles. Seasonal patterns are, however, rather different, with maximum concentrations occurring during late winter/early spring. The available data indicate that coastal regions may be larger emitters of DMS per unit area than the global mean, however, because of its much larger surface area, the open ocean is the most important source of tropospheric DMS (29).

CARBONYL SULPHIDE

Direct marine emissions are thought to account for about 20% of the global COS source (30), although the tropospheric oxidation of marine-derived DMS and carbon disulphide (CS_2) may increase this to as much as 55% (31). COS in seawater primarily develops from the photodecomposition of humic-like colored dissolved organic matter (CDOM) (32), although a small nonphotochemical source has also been inferred. The distribution of sea surface COS, therefore, corresponds closely to the concentration and reactivity of CDOM, which primarily derives from terrestrial sources. Consequently, COS concentrations are highest in estuaries, which are about an order of magnitude higher than those of adjacent coastal waters (32).

COS undergoes hydrolysis removal in seawater at a similar rate to that for its photo-production, hence COS shows a pronounced diel cycle in surface seawater with concentrations peaking in the early afternoon and declining to a minimum just before sunrise (33). Strong seasonal variation also exists. In the mid to high latitude open ocean the balance of photo-chemical production versus hydrolysis removal can lead to these regions becoming a seasonal sink for tropospheric COS (34). Net ocean uptake of COS has also been found in the subtropical ocean gyres (35). Taking account of these findings, the marine COS source is most likely dominated by the contribution from coastal and shelf areas.

BIBLIOGRAPHY

1. Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K., and Johnson, C.A. (Eds.). (2001). *Climate Change 2001: The Scientific Basis*. Cambridge University Press.
2. Barnola, J.M. et al. (1995). CO_2 evolution during the last millennium as recorded by Antarctic and Greenland ice. *Tellus* **47B**: 264–272.
3. Le Quéré, C., Orr, J., Monfray, P., and Aumont, O. (2000). Interannual variability in the oceanic sink of CO_2 from 1979 through 1997. *Glob Biogeochem Cyc.* **14**: 1247–1265.
4. Siengenthaler, U. and Sarmiento, J.L. (1993). Atmospheric carbon dioxide and the ocean. *Nature* **365**: 119–125.
5. Keeling, R.F. and Shertz, S.R. (1992). Seasonal and interannual variations in atmospheric oxygen and implications for the global carbon cycle. *Nature* **358**: 723.
6. Sarmiento, J.L. and Gruber, N. (2002). Sinks for anthropogenic carbon. *Physics Today* **55**: 30–36.
7. Broecker, W.S., Takahashi, T., Simpson, H.J., and Peng, T-H. (1979). Fate of fossil fuel carbon dioxide and the global carbon budget. *Science* **206**: 409–418.
8. Cox, P.M., Betts, R.A., Jones, C.D., Spall, S.A., and Totterdell, I.J. (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* **408**: 184.
9. Archer, D.E., Kheshgi, H., and Maier-Reimer, E. (1997). Multiple timescales for neutralization of fossil fuel CO_2. *Geophys Res. Lett.* **24**: 405–408.
10. Martin, J.H. (1992). Iron as a limiting factor. In: *Primary Productivity and Biogeochemical Cycles in the Sea*. P.G. Falkowski and A. Woodhead (Eds.). Plenum Press, New York, pp. 123–137.

11. Sarmiento, J.L., Hughes, T.M.C., Stouffer, R.J., and Manabe, S. (1998). Simulated response of the ocean carbon cycle to anthropogenic climate warming. *Nature* **393**: 245–249.
12. Nevison, C. and Holland, E. (1997). A re-examination of the role of anthropogenically fixed nitrogen on atmospheric N_2O and the stratospheric O_3 layer. *J. Geophys Res.* **10**: 15809–15820.
13. Crutzen, P.J. (1990). Methane's sinks and sources. *Nature* **350**: 390–381.
14. Cohen, Y. and Gordon, L.I. (1979). Nitrous oxide in the oxygen minimum of the of the eastern tropical north Pacific: Evidence for its consumption during denitrification and possible mechanisms for its production. *Deep Sea Res.* **25**: 509–524.
15. Vanraaphorst, W. et al. (1992). Nitrogen cycling in 2 types of sediments of the southern North-Sea (frisian front, broad fourteens)—field data and mesocosm results. *Neth. J. Sea Res.* **28**: 293–316.
16. Naqvi, S.W.A. et al. (1998). Budgetary and biogeochemical implications of N_2O isotope signatures in the Arabian Sea. *Nature* **394**: 462–464.
17. Suntharalingam, P. and Sarmiento, J.L. (2000). Factors governing the oceanic nitrous oxide distribution: Simulations with an ocean general circulation model. *Glob. Biogeochem. Cyc.* **14**: 429–454.
18. Barnes, J. and Owens, N.J.P. (1998). Denitrification and nitrous oxide concentrations in the Humber estuary, UK, and adjacent coastal zones. *Mar. Poll. Bul.* **37**: 247–260.
19. Dong, L.F., Nedwell, D.B., Underwood, G.J.C., Thornton, D.C.O., and Rusmana, I. (2002). Nitrous oxide formation in the Colne estuary, England: The central role of nitrite. *Appl. Environ. Microbiol.* **68**: 1240–1249.
20. Seitzinger, S.P. and Kroeze, C. (2000). Global distribution of N_2O emissions from aquatic systems: natural emissions and anthropogenic effects. *Chenmosphere* **2**: 267–279.
21. Lamontagne, R.A., Swinnerton, J.W., Linnenbom, V.J., and Smith, W.D. (1973). Methane concentrations in various marine environments. *J. Geophys. Res.* **78**: 5317–5324.
22. Upstill-Goddard, R.C., Barnes, J., and Owens, N.J.P. (1994). Nitrous oxide and methane during the 1994 SW monsoon in the Arabian Sea/noerthwestern Indian Ocean. *J. Geophys. Res.* **104**: 30,067–30,084.
23. Upstill-Goddard, R.C., Barnes, J., Frost, T., Punshon, S., and Owens, N.J.P. (2000). Methane in the southern North Sea: Low-salinity inputs, estuarine removal, and atmospheric flux. *Glob. Biogeochem. Cyc.* **14**: 1205–1217.
24. Judd, A.G. (2003). The global importance and context of methane escape from the seabed. *Geo-Marine Lett.* **23**: 147–154.
25. Kvenvolden, K.A., Lorenson, T.D., and Reeburgh, W.S. (2001). Attention turns to naturally occurring methane seepage. *EOS* **82**: 457.
26. Charlson, R.J., Lovelock, J.E., Andreae, M.O., and Warren, S.G. (1987). Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature* **326**: 655–661.
27. Turco, R.P., Whitten, R.C., Toon, O.B., Pollack, J.B., and Hamill, P. (1980). OCS, stratospheric aerosols and climate. *Nature* **283**: 283–286.
28. Chin, M. and Davies, D.D. (1995). A reanalysis of carbonyl sulphide as asource of stratospheric background sulphur aerosol. *J. Geophys. Res.* **100**: 8993–9005.
29. Kettle, A.J. et al. (1999). A global database of sea surface dimethylsulfide (DMS) measurements and a procedure to predict sea surface DMS as a function of latitude, longitude, and month. *Glob. Biogeochem. Cyc.* **13**: 399–444.
30. Watts, S.F. (2000). The mass budgets of carbonyl sulfide, dimethyl sulfide, carbon disulfide and hydrogen sulfide. *Atmos. Environ.* **34**: 761–779.
31. Kettle, A.J. et al. (2001). Assessing the flux of different volatile sulfur gases from the ocean to the atmosphere. *J. Geophys. Res.* **106**: 12193–12209.
32. Uher, G. and Andreae, M.O. (1997). Photochemical production of carbonyl sulfide in North Sea water: A process study. *Limnol. Oceanogr.* **42**: 432–442.
33. Ulshöfer, V.S., Flöck, O.R., Uher, G., and Andreae, M.O. (1996). Photochemical production and air-sea exchange of carbonyl sulfide in the eastern Mediterranean Sea. *Mar. Chem.* **53**: 25–39.
34. Ulshofer, V.S., Uher, G., and Andreae, M.O. (1995). Evidence for a winter sink of atmospheric carbonyl sulfide in the northeast Atlantic Ocean. *Geophys. Res. Lett.* **22**: 2601–2604.
35. Weiss, P.S., Johnson, J.E., Gammon, R.H., and Bates, T.S. (1995). A reevaluation of the open ocean source of carbonyl sulfide to the atmosphere. *J. Geophys. Res.* **100**: 23083–23092.

PACIFIC MARINE ENVIRONMENTAL LABORATORY—30 YEARS OF OBSERVING THE OCEAN

NOAA—Pacific Marine
Environmental Laboratory

Sept. 29, 2003—Although the NOAA Pacific Marine Environmental Laboratory in Seattle, Wash., celebrates its 30th anniversary this year, its staff has spent 43 years at sea. The figure of 15,654 days at sea was one of the many facts presented during the lab's anniversary celebration in August. That, along with 1,290 published journal articles and 352,000,000 hits on the PMEL Web page indicate that there's a lot going on out on Sand Point.

For two-thirds of its life, the lab has been under the direction of Eddie Bernard. An oceanographer by training, Bernard became director in 1983, a decade after the former Pacific Oceanographic Laboratory became PMEL. "We have dedicated people at PMEL who devote a lot of energy and creativity to the work we do," he said.

NATIONAL TSUNAMI MITIGATION PROGRAM

Some of that creativity and energy became evident when in 1994 the U.S. Senate asked NOAA to come up with a plan to reduce the risk of tsunamis to coastal residents. What resulted was the National Tsunami Hazard Mitigation Program, chaired by Bernard and composed of representatives from federal, state and local agencies from West Coast states, Alaska and Hawaii, working to save lives and property. "The National Tsunami Mitigation Program initiated by PMEL is a unique and effective partnership," said Rich Eisner of the California Governor's

---

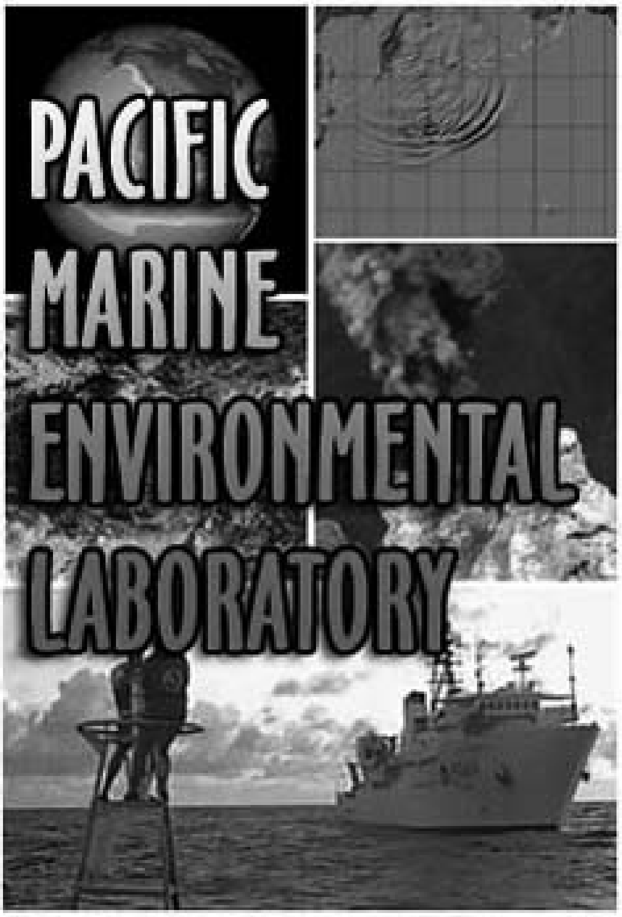

Office of Emergency Services, one of the program members. "The integration of science and mitigation policy, the warning centers, and local emergency management, and the application of new technologies fostered by PMEL have been successful beyond expectations." Among the technological innovations is the system of buoys along the West Coast that serve as warning devices, or, as Bernard calls them, "tsunameters." The tsunami program also includes a public education component that teaches coastal residents what to do in case of a possible tsunami. Signs now indicate evacuation routes and some coastal communities have been designated "TsunamiReady" for their efforts to educate and protect their residents.

UNDERWATER VOLCANOES AND VENTS

As one of NOAA's "wet" labs, PMEL focuses on a variety of ocean issues. When underwater volcanoes or vents, were first discovered in the Galapagos Islands 26 years ago, PMEL was among the first to start investigating these unusual underwater communities, where unique marine life thrive on the chemical soup spewed from the sea floor. "We may be taking drugs in the future made of enzymes that are more compatible with our bodies than synthetic compounds, which may have side effects," Bernard said.

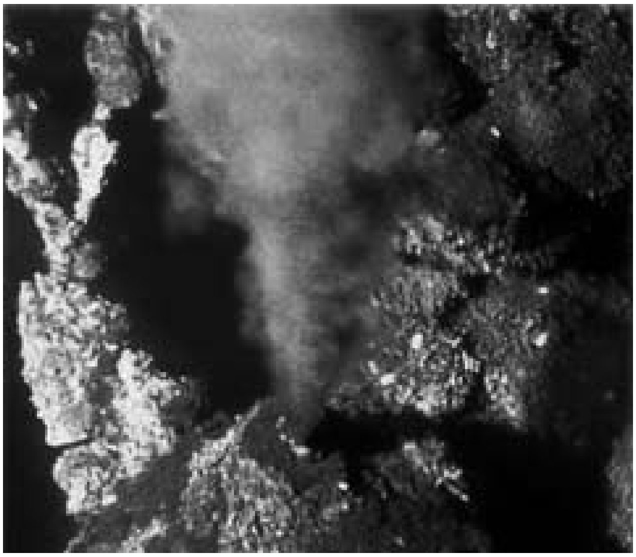

"What's spewing from the ocean floor could someday give us resistance to some new strains of infection."

FISHERIES OCEANOGRAPHY COORDINATED INVESTIGATIONS PROGRAM

PMEL began as a "small research laboratory with emphasis on water quality and environmental impact issues" in the waters off the West Coast extending to the equatorial Pacific Ocean. It now has an international reputation in many areas, especially its ability to collect ocean data and to work collaboratively in projects that cover many disciplines. One example is the Fisheries Oceanography Coordinated Investigations program that assists in forecasting fish stocks to help ensure a reliable supply and lower costs to consumers. "In 1985, Eddie Bernard took a big risk," said Doug DeMaster of the NOAA Marine Fisheries Service. That risk was offering to establish with the Alaska Fisheries Science Center and his counterpart, William Aron, a cutting-edge, applied

science program across NOAA line offices. Eighteen years after its inception, NOAA's FOCI has published more than 450 scientific articles and was awarded the Department of Commerce Bronze Medal in 2002 for "scientific achievements that have advanced fisheries oceanography and marine ecology and have contributed to building sustainable fisheries in the North Pacific." DeMaster noted, "Today, if you attend a FOCI meeting, you cannot tell which scientists are from NOAA Research and which are from NOAA Fisheries. In 1985, it took vision and courage to blur the lines between line offices. Today, it seems only natural."

PACIFIC TROPICAL ATMOSPHERE OCEAN ARRAY

Understanding the natural systems is a key element of the lab. "The ocean is dynamic, it moves all of the time," Bernard said. "We're now in the third generation of observing systems. In the equatorial Pacific, we have the world's longest continuous time series of open ocean data—25 years." The equatorial Pacific also proved to be the place to be if humans wanted early warning of El Niño events. El Niño is a disruption of the ocean-atmosphere system in the tropical Pacific having important consequences for weather around the globe. After the 1982–83 El Niño, considered the most intense in the 20th century, the challenge was given to develop some sort of early warning device so people could prepare for the devastating and beneficial aspects of an El Niño and its counterpart, La Niña. Once again, in 1994, PMEL harnessed the creativity and talent of dedicated scientists and came up with the Pacific Tropical Atmosphere Ocean (TAO) array, the world's largest ocean observing system. Bobbing in the Pacific are 70 buoys measuring and relaying surface wind, sea surface temperature, upper ocean temperatures and currents, air temperature and relative humidity in real-time via satellite. "We knew we were onto something when we linked data from the TAO buoys to the Internet and attracted millions of hits from all over the world," Bernard said. "All it takes is time, money and commitment."

UNDERWATER ACOUSTICAL MONITORING

Always eager to hear what the Earth has to say, PMEL scientists also listen to the planet via underwater acoustical monitoring. Using a variety of methods, including underwater hydrophones, PMEL listens for seismic activity, marine mammals and ship traffic. The systems also have picked up some so-far unidentified sounds. "People tend to think the ocean is quiet beneath the surface," said Christopher Fox, director of the ocean acoustics project. "But it's pretty noisy down there." Some things are easy to identify, Fox said, such as whales and ship traffic. But visitors to the ocean acoustics Web site can listen to such unidentified sounds that the lab has dubbed "Train," "Upsweep," "Whistle" and "Bloop."

After 30 years, PMEL knows that the Earth still holds countless tantalizing secrets. And PMEL scientists and staff are eager to unlock those secrets. "As the planet aspirates, it provides new opportunities," Bernard

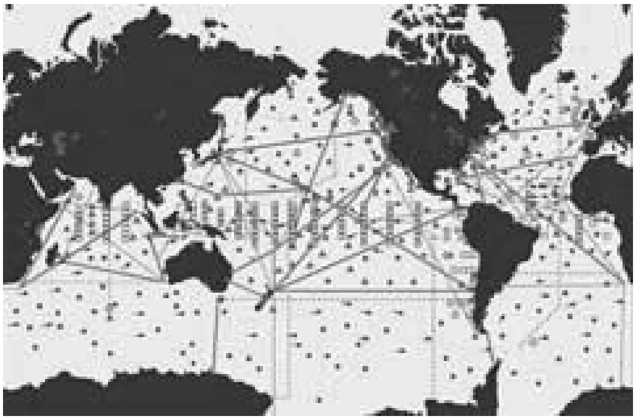

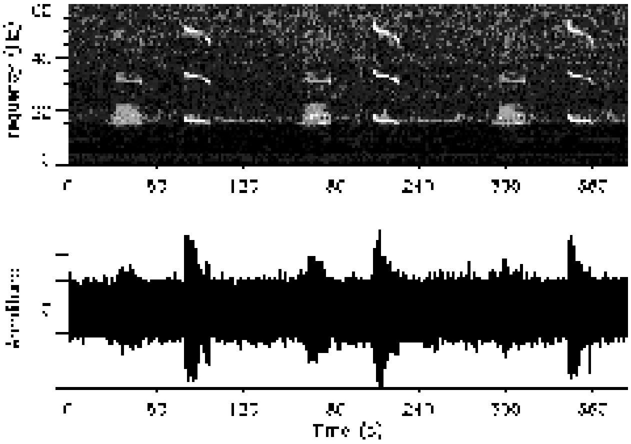

said. “It’s an ongoing science experiment with enormous challenges and rewards.”

RELEVANT WEB SITES

NOAA Pacific Marine Environmental Laboratory

Dr. Eddie Bernard, Director, Pacific Marine Environmental Laboratory

NOAA National Tsunami Hazard Mitigation Program

NOAA Vents Program

Fisheries Oceanography Coordinated Investigations program

NOAA Marine Fisheries Service

NOAA Alaska Fisheries Science Center

NOAA El Niño Theme Page

NOAA La Niña Theme Page

NOAA Pacific Tropical Atmosphere Ocean (TAO) Project

NOAA Vents Program: Underwater Acoustical Monitoring

NOAA’s ALASKA FISHERIES SCIENCE CENTER

Ocean Explorer: Sounds in the Sea

MEDIA CONTACT

Jana Goldman, NOAA Research, (301) 713–2483

SEAWATER TEMPERATURE ESTIMATES IN PALEOCEANOGRAPHY

Giuseppe Cortese
Alfred Wegener Institute for
Polar and Marine Research
(AWI)
Bremerhaven, Germany

BACKGROUND

Climatic issues, such as global warming, greenhouse gases release into the atmosphere, and the role mankind plays in affecting the climate of our planet have recently gained international interest.

As a result, political panels and decision-makers are starting to look with growing interest at the results of committees established with the purpose of analyzing past, present, and future climate change.

One of these panels [Intergovernmental Panel on Climate Change (IPCC)], established by the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP), regularly publishes reports (available also online at http://www.grida.no/climate/ipcc_tar/) intended for scientists, politicians, and anybody interested in and/or concerned by climate change.

The most powerful tools currently available to the scientific community for the analysis and prediction of climate are computer models, which simulate the functioning of Earth as a complex system. These models differ based on which components of the climate system (atmosphere, ocean, biosphere, cryosphere, etc.) they include and account for, their spatial scale and degree of detail, and on the different sets of physicochemical equations they use to describe and parameterize phenomena taking place at a subgrid spatial scale.

Regardless of setup and characteristics, though, models should be benchmarked against an observational base (e.g., to test how well the model reproduces a known distribution of a climatic variable). Models are moreover usually initialized by a set of observations (to let the incorporated physics modify the starting distribution).

Some of these datasets are readily available in the form of measurements performed over historic timescales: One of the best known examples is the atmospheric CO_2 concentration curve measured at Mauna Loa, Hawaii, since the late 1950s (1).

Other important climatic variables cannot be observed or measured directly, and indirect methods, so-called “proxies,” have been developed to obtain information on them. Some of the methods that estimate seawater temperature on geological timescales are described in the following.

INTRODUCTION

The heat balance of Earth is strongly affected, to a first approximation, by the amount of solar radiation reaching the top of the atmosphere. The motion of currents in the atmosphere and in the ocean redistributes this heat between low and high latitudes, as the low latitudes receive a higher amount of heat compared with the higher latitudes. The ocean can be regarded as the “thermostat” of such a heat machine, because water has a much higher heat capacity compared with air, and because the turnover time for the ocean (thousands of years) is several orders of magnitude higher than for the atmosphere (days). One of the main features of the large-scale circulation in the ocean is the formation of deep waters at high latitudes, in the Weddell and Ross Seas (around Antarctica), and in the Labrador and Greenland/Iceland Seas (in the northern North Atlantic), because of the cooling of enormous amounts of warm surface waters advected to these locations by ocean currents (e.g., the Gulf Stream in the North Atlantic). When these waters move to higher latitudes, they cool, become denser, and sink to the ocean depths, where they start a reversed journey, from high to low latitudes. During this cooling, they also release latent heat and moisture, which help to mitigate the climate of the coastal areas they fringe, including the whole of western Europe.

Indications exist, from geological records and model simulations (2), that during glacial times, this oceanic overturning did not operate with the same intensity as today, and/or that the places where deep waters formed were displaced compared with today. As seawater temperature, together with salinity, determines the density of seawater (and therefore it strongly influences its circulation and sinking characteristics), a variety of methods have been developed to determine this important climatic variable in the distant, geological past.

METHODS

Sediment cores collected from the bottom of the World Ocean represent an ideal archive to trace the history of the ocean. Some of these cores can be accurately dated, and often they contain the fossilized remains of organisms that lived in past oceans, and reacted to changes in their characteristics.

The most useful of these organisms, because of their high fossilization potential, species diversity, and sensitivity to a range of climatic variables, are micro- and nannofossils. As their name implies, these organisms range from nanometer to millimeter size and leave fossilizable body parts, with the most commonly used groups belonging to algae (diatoms, nannoplankton) and protozoans (radiolarians, foraminifera). Most species in these groups live close to the sea surface, and each of them is adapted to a particular set of sea surface water variables, including temperature, salinity, and macro- and micro-nutrient content. Several techniques have been developed to exploit this close link between species occurrences/abundances and past environmental variables. Our first global view of surface water temperatures in the ocean during the Last Glacial Maximum (ca. 18,000 years Before Present) came from one such technique, namely, Q-Mode Principal Component Analysis applied to foraminifer and radiolarian abundance data in sediment cores (3).

Semiquantitative Floral/Faunal Estimates

Some of the first attempts to obtain seawater temperature information from microfossil assemblages tried to derive this value directly from a formula. The latter was based on the presence/absence, in the fossil samples, of species having a well-known distribution in the modern ocean, and being representative of different climatic zones (e.g., equatorial, subtropical, temperate, etc.). This approach, initially developed for diatoms (4), has also been applied to radiolarians (5). Information on species diversities of modern planktonic foraminifera in the Indian Ocean have also been used, in the late 1970s, to estimate ocean paleotemperatures (6).

Transfer Functions

With the expansion of the knowledge about the biogeography of most plankton groups (and the ecological and environmental significance of many species), and the development of computers (allowing the implementation of more sophisticated algorithms and techniques), transfer functions made their breakthrough in paleoceanography. They were first described and applied to planktonic foraminifera by Imbrie and Kipp (7), who illustrated their utilization in paleoclimatology. The species assemblage present in a collection of modern sediment samples (calibration dataset), containing up to several hundreds of samples, is "simplified" into a limited number of faunal factors/assemblages. A multivariate regression is then used to calibrate these simplified assemblages to the desired environmental variable (e.g., surface seawater temperature) measured, in the water column, at the same locations where the surface sediment samples were recovered. The value of the environmental variable at a certain time in the past is estimated by calculating "pseudo-factors" for each of the past samples (i.e., the value the calibrated assemblages have in these samples) and replacing them in the calibration equation. One main assumption of this method is that the considered environmental variable (in our case, surface temperature) plays a major role in influencing the biogeographic distribution of the considered microfossil species, and that it is therefore responsible for most of the variance observed in the calibration dataset. Transfer functions have been successfully applied to past ocean surface temperature reconstructions by using planktonic foraminifera (8–14), radiolarians (15–22), and diatoms (23,24).

Geochemical Methods

Several methods have been recently developed to estimate past ocean temperatures by measuring chemical elements or organic compounds, which are either incorporated in the organism shell or produced during its life cycle. The general approach is to develop a calibration equation by growing the desired species in culture under a variety of environmental conditions, and measuring how its geochemical composition changes. Additionally, a calibration equation can also be obtained by analyzing extensive collections of recent sediments, covering areas where the environmental variables display a wide range. The resulting equation is then applied to fossil samples to derive the past value of the desired variable.

Mg/Ca

The relative amount of magnesium compared with calcium a certain planktonic foraminifera species incorporates in its test depends on temperature, and the Mg/Ca ratio can therefore also be used as a paleo-thermometer. The calibration equation is derived from laboratory culture studies, based on different species, including *Globigerinoides sacculifer* (25), *Globigerina bulloides*, and *Orbulina universa* (26). Revised calibrations and applications to geological records are being continuously developed (27–31). Another paleo-thermometer derived from elemental ratios in foraminifera tests is the Sr/Ca ratio (32).

Alkenones

Several species of nannoplankton, living in the shallower layer of the ocean (less than ca. 50 m depth), produce organic compounds named long-chain (C_{37}-C_{39}) unsaturated ketones, also known as alkenones. The two species being responsible for most of the alkenone production are *Emiliania huxleyi* and *Gephyrocapsa oceanica*. Although it is still not clear what the function of such compounds is, they have become a very useful tool for the estimation of past seawater temperature. It has been demonstrated (33,34) that the three different varieties of the C_{37} unsaturated compounds ($C_{37:2}$, $C_{37:3}$, $C_{37:4}$), characterized by 2, 3, or 4 double bonds, display abundance variations related to seawater temperature. The alkenone unsaturation ratio (35), $U^{k'}{}_{37} = [C_{37:2}]/([C_{37:2}] + [C_{37:3}])$, was calibrated to temperature (36) by growing *E. huxleyi* in culture under a range of temperatures. Since then, alkenones have been widely applied in paleoclimatology (37–45).

Other Methods

Other methods commonly used to estimate past ocean temperatures are as follows the modern analogue technique and its variations (46–51); artificial neural networks (52,53); and the δ^{18}O isotopic composition of planktonic foraminifera (54–58). The study of the isotopic signal stored in corals, although covering shorter time intervals than marine sediment cores, provides excellent temporal resolution archives (down to a few weeks), which allow to analyze seasonal climate variability (59,60).

BIBLIOGRAPHY

1. Keeling, C.D., Whorf, T.P., Wahlen, M., and van der Plicht, J. (1995). Interannual extremes in the rate of rise of atmospheric carbon dioxide since 1980. *Nature* **375**: 666–670.
2. Ganopolski, A. and Rahmstorf, S. (2001). Rapid changes of glacial climate simulated in a coupled climate model. *Nature* **409**: 153–158.
3. CLIMAP. (1976). The surface of the ice-age Earth. *Science* **191**: 1131–1137.
4. Kanaya, T. and Koizumi, I. (1966). Interpretation of diatom thanatocoenoses from the North Pacific applied to a study of core V20-130 (studies of a deep-sea core V20-120, part IV). *Sci. Rep. Tohoku Univ. Ser. 2.* **37**: 89–130.
5. Nigrini, C. (1970). Radiolarian assemblages in the North Pacific and their application to a study of Quaternary sediments in core V20-130. *Memoir of the Geological Society of America* **126**: 139–183.
6. Williams, D.F. and Johnson, W. (1975). Diversity of recent planktonic foraminifera in the southern Indian Ocean and late Pleistocene paleotemperatures. *Quaternary Res.* **5**: 237–250.
7. Imbrie, J. and Kipp, N.G. (1971). A new micropaleontological method for quantitative paleoclimatology: Application to a late Pleistocene Caribbean core. In: *Late Cenozoic Glacial Ages*. K. Turekian (Ed.). Yale University Press, New Haven, CT, pp. 71–181.
8. CLIMAP. (1981). *Seasonal Reconstructions of the Earth's Surface at the Last Glacial Maximum*. Geological Society of America, Map and Chart Series MC-36.
9. CLIMAP. (1984). The Last Interglacial ocean. *Quaternary Research* **21**: 123–224.
10. Mix, A.C. and Morey, A.E. (1996). Climate feedback and pleistocene variations in the atlantic south equatorial current. In: *The South Atlantic: Present and Past Circulation*. G. Wefer, W.H. Berger, G. Siedler, and D.J. Webb (Eds.). Springer-Verlag, Berlin, pp. 503–525.
11. Ortiz, J.D. and Mix, A.C. (1997). Comparison of Imbrie-Kipp transfer function and modern analog temperature estimates using sediment trap and core top foraminiferal faunas. *Paleoceanography* **12**(2): 175–190.
12. Mix, A.C., Morey, A.E., Pisias, N.G., and Hostetler, S.W. (1999). Foraminiferal faunal estimates of paleotemperature: Circumventing the no-analog problem yields cool ice age tropics. *Paleoceanography* **14**: 350–359.
13. Feldberg, M.J. and Mix, A.C. (2002). Sea-surface temperature estimates in the Southeast Pacific based on planktonic foraminiferal species; modern calibration and Last Glacial Maximum. *Marine Micropaleontol.* **44**: 1–29.
14. Niebler, H.S. et al. (2003). Sea-surface temperatures in the Equatorial and South Atlantic Ocean during the Last Glacial Maximum (23-19 ka). *Paleoceanography* **18**: 1069.
15. Moore, T.C., Jr. (1973). Late Pleistocene-Holocene oceanographic changes in the northeastern Pacific. *Quaternary Res.* **3**(1): 99–109.
16. Morley, J.J. (1979). A transfer function for estimating paleoceanographic conditions, based on deep-sea surface distribution of radiolarian assemblages in the South Atlantic. *Quaternary Res.* **12**: 381–395.
17. Pisias, N.G., Roelofs, A., and Weber, M. (1997). Radiolarian-based transfer functions for estimating mean surface ocean temperatures and seasonal range. *Paleoceanography* **12**(3): 365–379.
18. Bjørklund, K.R., Cortese, G., Swanberg, N., and Schrader, H.J. (1998). Radiolarian faunal provinces in surface sediments of the Greenland, Iceland and Norwegian (GIN) Seas. *Marine Micropaleontology* **35**: 105–140.
19. Abelmann, A., Brathauer, U., Gersonde, R., Sieger, R., and Zielinski, U. (1999). Radiolarian-based transfer function for the estimation of sea surface temperatures in the Southern Ocean (Atlantic sector). *Paleoceanography* **14**(3): 410–421.
20. Brathauer, U. and Abelmann, A. (1999). Late quaternary variations in sea surface temperatures and their relationship to orbital forcing recorded in the Southern Ocean (Atlantic sector). *Paleoceanography* **14**(2): 135–148.
21. Cortese, G. and Abelmann, A. (2002). Radiolarian-based paleotemperatures during the last 160 kyrs at ODP Site 1089 (Southern Ocean, Atlantic Sector). *Palaeogeography, Palaeoclimatology, Palaeoecology* **182**(3–4): 259–286.
22. Dolven, J.K., Cortese, G., and Bjørklund, K.R. (2002). A high-resolution radiolarian-derived paleotemperature record for the Late Pleistocene-Holocene in the Norwegian Sea. *Paleoceanography* **17**(4): 1072.
23. Koc-Karpuz, N. and Schrader, H.J. (1990). Surface sediment diatom distribution and Holocene paleotemperature variations in the Greenland, Iceland and Norwegian Sea. *Paleoceanography* **5**(4): 557–580.
24. Zielinski, U., Gersonde, R., Sieger, R., and Fütterer, D. (1998). Quaternary surface water temperature estimations: Calibration of a diatom transfer function for the Southern Ocean. *Paleoceanography* **13**: 365–383.
25. Nürnberg, D., Bijma, J., and Hembelen, C. (1996). Assessing the reliability of magnesium in foraminiferal calcite as a proxy for water mass temperatures. *Geochim. Cosmochim. Acta* **60**: 803–814.
26. Lea, D.W., Mashiotta, T.A., and Spero, H.J. (1999). Controls on magnesium and strontium uptake in planktonic foraminifera determined by live culturing. *Geochim. Cosmochim. Acta* **63**: 2369–2379.
27. Hastings, D.W., Russell, A.D., and Emerson, S.R. (1998). Foraminiferal magnesium in *G. sacculifer* as a paleotemperature proxy in the equatorial Atlantic and Caribbean surface oceans. *Paleoceanography* **13**(2): 161–169.
28. Mashiotta, T.A., Lea, D.W., and Spero, H.J. (1999). Glacial–interglacial changes in Subantarctic sea surface temperature and d18O-water using foraminiferal Mg. *Earth Planetary Sci. Lett.* **170**: 417–432.
29. Nürnberg, D., Müller, A., and Schneider, R.R. (2000). Paleo-sea surface temperature calculations in the equatorial east Atlantic from Mg/Ca ratios in planktic foraminifera: A comparison to sea surface temperature estimates from $U^{k'}{}_{37}$, oxygen isotopes, and foraminiferal transfer function. *Paleoceanography* **15**(1): 124–134.
30. Dekens, P.S., Lea, D.W., Pak, D.K., and Spero, H.J. (2002). Core top calibration of Mg/Ca in tropical foraminifera: Refining paleotemperature estimation. *Geochem. Geophys. Geosyst.* **3**(4): 10.1029/2001GC000200.
31. Lea, D.W., Pak, D.K., Peterson, L.C., and Hughen, K.A. (2003). Synchroneity of tropical and high-latitude Atlantic temperatures over the last glacial termination. *Science* **301**: 1361–1364.

32. Cronblad, H.G. and Malmgren, B.A. (1981). Climatically controlled variation of Sr and Mg in Quaternary planktonic foraminifera. *Nature* **291**: 61–64.
33. Brassell, S.C., Eglinton, G., Marlowe, I.T., Pflaumann, U., and Sarnthein, M. (1986). Molecular stratigraphy: A new tool for climatic assessment. *Nature* **320**: 129–133.
34. Prahl, F.G., Collier, R.B., Dymond, J., Lyle, M., and Sparrow, M.A. (1993). A biomarker perspective on prymnesiophyte productivity in the northeast Pacific Ocean. *Deep-Sea Res. I* **40**: 2061–2076.
35. Prahl, F.G., Mühlhausen, L.A., and Zahnle, D.L. (1988). Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions. *Geochim. Cosmochim. Acta* **52**: 2303–2310.
36. Prahl, F.G. and Wakeham, S.G. (1987). Calibration of unsaturation patterns in long-chain ketone compositions for palaeotemperature assessment. *Nature* **330**: 367–369.
37. Lyle, M.W., Prahl, F.G., and Sparrow, M.A. (1992). Upwelling and productivity changes inferred from a temperature record in the central equatorial Pacific. *Nature* **355**: 812–815.
38. Rostek, F.G. et al. (1993). Reconstructing sea surface temperature and salinity using δ^{18}O and alkenone records. *Nature* **364**: 319–321.
39. Zhao, M., Rosell, A., and Eglinton, G. (1993). Comparison of two $U^{k}{}_{37}$-sea surface temperature records for the last climatic cycle at ODP Site 658 from the sub-tropical Northeast Atlantic. *Palaeogeography, Palaeoclimatology, Palaeoecology* **103**: 57–65.
40. Ohkouchi, N., Kawamura, K., Nakamura, T., and Taira, A. (1994). Small changes in the sea surface temperature during the last 20,000 years: molecular evidence from the western tropical Pacific. *Geophys. Res. Lett.* **21**: 2207–2210.
41. Sikes, E.L. and Keigwin, L.D. (1994). Equatorial Atlantic sea surface temperature for the last 30 kyr: A comparison of $U^{k'}{}_{37}$, δ^{18}O, and foraminiferal assemblage temperature estimates. *Paleoceanography* **9**(1): 31–45.
42. Schneider, R.R., Müller, P.J., and Ruhland, G. (1995). Late quaternary surface circulation in the east equatorial South Atlantic: evidence from alkenone sea surface temperatures. *Paleoceanography* **10**(2): 197–219.
43. Schneider, R.R., Müller, P.J., Ruhland, G., Meinecke, G., Schmidt, H., and Wefer, G. (1996). Late quaternary surface temperatures and productivity in the East-Equatorial South Atlantic: Response to changes in Trade/Monsoon wind forcing and surface water advection. In: *The South Atlantic: Present and Past Circulation*. G. Wefer, W.H. Berger, G. Siedler, and D.J. Webb (Eds.). Springer-Verlag, Berlin, pp. 527–551.
44. Müller, P.J., Cepek, K., Ruhland, G., and Schneider, R.R. (1997). Alkenone and coccolithophorid species changes in Late Quaternary sediments from Walvis Ridge: Implications for the alkenone paleotemperature method. *Palaeogeography, Palaeoclimatology, Palaeoecology* **135**: 71–96.
45. Rühlemann, C., Mulitza, S., Müller, P.J., Wefer, G., and Zahn, R. (1999). Warming of the tropical Atlantic Ocean and slowdown of thermohaline circulation during the last deglaciation. *Nature* **402**: 511–514.
46. Hutson, W.H. (1980). The Agulhas Current during the late Pleistocene: analysis of modern faunal analogs. *Science* **207**: 64–66.
47. Prell, W.L. (1985). *The Stability of Low Latitude Sea Surface Temperatures: An Evaluation of the CLIMAP Reconstruction with Emphasis on the Positive SST Anomalies*. Rep. TR 025 U.S. Dept. of Energy, Washington, DC, pp. 1–60.
48. Pflaumann, U., Duprat, J., Pujol, C., and Labeyrie, L. (1996). SIMMAX: A modern analog technique to deduce Atlantic sea surface temperatures from planktonic foraminifera in deep-sea sediments. *Paleoceanography* **11**: 15–35.
49. Waelbroeck, C. et al. (1998). Improving past sea surface temperature estimates based on planktonic fossil faunas. *Paleoceanography* **13**(3): 272–283.
50. Hale, W. and Pflaumann, U. (1999). Sea-surface temperature estimations using a Modern Analog Technique with foraminiferal assemblages from Western Atlantic Quaternary sediments. In: *Use of Proxies in Paleoceanography—Examples from the South Atlantic*. G. Fischer and G. Wefer (Eds.). Springer, Berlin, pp. 69–90.
51. Hendy, I.L. and Kennett, J.P. (2000). Dansgaard-Oeschger cycles and the California Current System: Planktonic foraminiferal response to rapid climate change in Santa Barbara Basin, Ocean Drilling Program Hole 893A. *Paleoceanography* **15**: 30–42.
52. Malmgren, B.A. and Nordlund, U. (1997). Application of artificial neural networks to paleoceanographic data. *Palaeogeography, Palaeoclimatology, Palaeoecology* **136**: 359–373.
53. Malmgren, B.A., Kucera, M., Waelbroeck, C., and Nyberg, J. (2001). Comparison of statistical and artificial neural network techniques for estimating past sea-surface temperatures from planktonic foraminifer census data. *Paleoceanography* **16**(5): 520–530.
54. Emiliani, C. (1955). Pleistocene temperatures. *J. Geol.* **63**: 538–578.
55. Shackleton, N.J. and Opdyke, N.D. (1973). Oxygen isotope and palaeomagnetic stratigraphy of equatorial Pacific Core V28-238: Oxygen isotope temperatures and ice volumes on a 10^5 year and 10^6 year scale. *Quaternary Res.* **3**: 39–55.
56. Broecker, W.S. (1986). Oxygen isotope constraints on surface ocean temperatures. *Quaternary Research* **26**: 121–134.
57. Mulitza, S., Dürkoop, A., Hale, W., Wefer, G., and Niebler, H.S. (1997). Planktonic foraminifera as recorders of past surface-water stratification. *Geology* **25**: 335–338.
58. Bemis, B.E., Spero, H.J., Bijma, J., and Lea, D.W. (1998). Reevaluation of the oxygen isotopic composition of planktonic foraminifera: Experimental results and revised paleotemperature equations. *Paleoceanography* **13**(2): 150–160.
59. Felis, T. et al. (2000). A coral oxygen isotope record from the northern Red Sea documenting NAO, ENSO, and North Pacific teleconnections on Middle East climate variability since the year 1750. *Paleoceanography* **15**: 679–694.
60. Nyberg, J., Winter, A., Malmgren, B., and Christy, J. (2001). Surface temperatures in the eastern Caribbean during the 7th century AD average up to 4 °C cooler than present. *Eos Trans. AGU* **82**(47).

PHYSICAL OCEANOGRAPHY

Pacific Fisheries Environmental
Laboratory, NOAA

A great deal of the patterns and fluctuations observed in our living marine resources are attributable to the impact of physical processes in the environment on marine

---

ecosystems and their components. For this reason, PFEL places a strong emphasis on research that examines the role of the physical environmental variability on marine ecosystems in general and commercially important fish stocks specifically. The objectives of the physical oceanography task are:

- perform research on the temporal and spatial scales of environmental variability in eastern boundary current systems in relation to other marine ecosystems
- provide environmental input to SWFSC research programs, particularly the coastal groundfish program
- provide high quality marine information to the research community.

Research is performed at PFEL which integrates environmental and biological data sets, investigating the linkages between environmental variability and fluctuations in the abundance and distribution of marine populations on a continuum of scales (from global, basin-wide spatial scales to the scale of local upwelling centersand on time scales from decades down to days). Physical oceanography research is directed to:

- large-scale climatic variability
- environment/recruitment relationships in eastern boundary current ecosystems
- mesoscale (smaller scale) processes affecting coastal circulation and fisheries recruitment

Examples of research on the large-scale variability include studies of recurring temperature changes off the west coast of the U. S. and their effects on groundfish recruitment, and the investigation of environmental changes in the California Current region associated with recent ENSO events. Much of the mesoscale research focuses on relating environmental variability on day-to-year and 5–100 nm scales to patterns and events in the life history of groundfish (e.g., recruitment success). The physical oceanography research program is linked closely to those of the other tasks at PFEG and to research programs at the other SWFSC labs. PFEG scientists also are involved in numerous cooperative studies with oceanographers and fisheries scientists at many federal and state government, academic and privately supported research institutions.

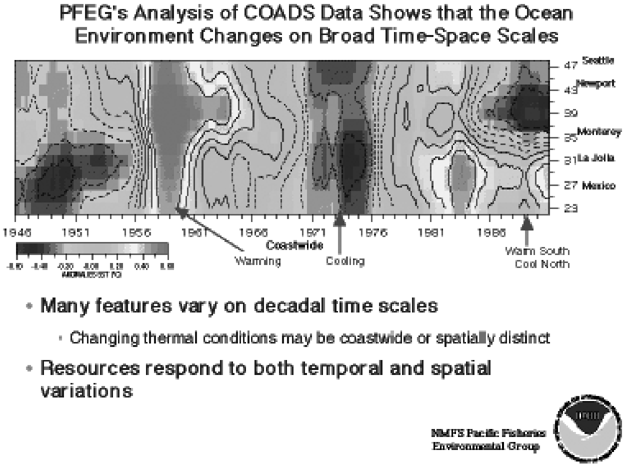

PFEG's Analysis of COADS Data Shows that the Ocean Environment Changes on Broad Time-Space Scales

- Many features vary on decadal time scales
  - Changing thermal conditions may be coastwide or spatially distinct
- Resources respond to both temporal and spatial variations

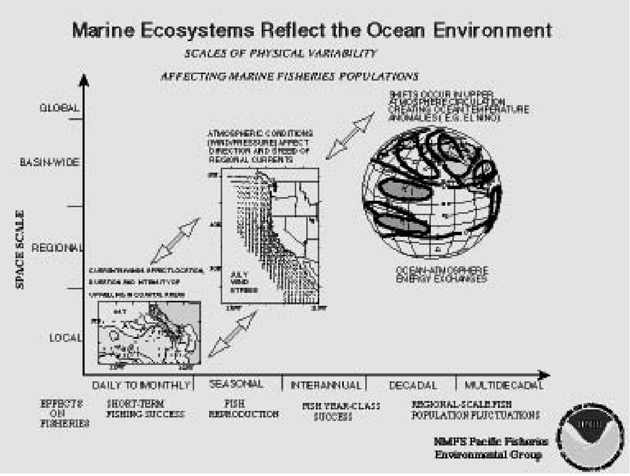

Marine Ecosystems Reflect the Ocean Environment

SCALES OF PHYSICAL VARIABILITY

AFFECTING MARINE FISHERIES POPULATIONS

Expertise in physical oceanography at PFEG and the linkages to the Navy's Fleet Numerical Meteorology and Oceanography Center (FNMOC) and Naval Postgraduate School, as well as numerous other government, academic, and private research facilities, has historically meant that this task serves regionally, nationally and internationally as a resource to other ocean scientists. Within the SWFSC, many cooperative research programs have been developed and planned. As an example, the task provides physical oceanographic expertise to the Tiburon Laboratory Rockfish Recruitment surveys each spring, to relate ocean variability off central California to rockfish recruitment. PFEG physical oceanographers are asked frequently to attend workshop and present seminars as experts on environmental-fishery linkages, and represent SWFSC, NMFS and NOAA on numerous committees and working groups.

COASTAL WATER POLLUTANTS

Upadhyayula V. K. Kumar
Choa Chu kang Ave-4
Singapore

POLLUTANT INPUT INTO THE MARINE ENVIRONMENT

Among all the diversity of human activities and sources of pollution, we can distinguish three main ways that pollutants enter the marine environment:

- direct discharge of effluents and solid wastes into the seas and oceans (industrial discharge, municipal waste discharge, coastal sewage, and others);
- land runoff into the coastal zone, mainly from rivers;
- atmospheric fallout of pollutants transferred by the air mass onto the sea surface.

Certainly, the relative contribution of each of these channels to the combined pollution input into the sea will be different for different substances and in different situations. At present, the signs and consequences of human activity can be found everywhere on Earth. One of the typical features of marine pollution is global spreading of a number of contaminants. Numerous data undoubtedly indicate the existence of a large-scale (global) field of background contamination of the hydrosphere.

Another important feature of marine pollution is the existence of increased pollution levels in the enclosed seas and coastal waters, compared with the open ocean.

Contamination levels also increase during the transition from the southern parts of all oceans to the north, where the main industrial centers and main pollution sources are concentrated. Besides the general distribution pattern of pollution sources, there are two other factors explaining the relative stability of global pollutant distribution in the world ocean: the relative confinement of large-scale water circulation within the limits of each hemisphere and the predominance of the zonal transport of the traces in the atmosphere.

Another distinctive and repeatedly registered feature of the general picture of contaminant distribution in the marine environment is their localization at the water–atmosphere and water–bottom sediment boundaries. Practically everywhere and for all trace components (primarily for oil hydrocarbons), their concentrations are considerably (usually hundreds and thousands of times) higher in the surface microlayer of water and in the upper layer of bottom sediments. These boundaries provide the biotopes for the communities of hyponeuston and benthos, respectively.

The existence of elevated levels of contaminants in zones of high bioproductivity is extremely alarming ecologically. These zones include the water layer up to 100 m from the water surface (photic layer) and boundaries of natural environments (water–atmosphere and water–bottom sediment, as previously mentioned), as well as enclosed seas, estuaries, and coastal and shelf waters. In particular, in shelf and coastal zones, which occupy 10% of the world ocean surface and less than 3% of its volume, the most intense processes of bioproduction, including the self-reproduction of the main living resources of the sea, take place. The main press of anthropogenic impact is also concentrated here.

The number and diversity of pollution components is growing as well. Contaminants that are globally distributed are combined here with hundreds and thousands of ingredients of local and regional distribution. Most of these substances are wastes and discharges from different local industries and activities.

Based on the extreme diversity of marine pollution components, the variety of their sources, the scales of distribution, and the degree of hazards, these pollutants can be classified in different ways, depending on their composition, toxicity, persistence, sources, volumes, and so on.

TYPES AND FORMS OF WATER POLLUTION

Water pollution is attributed to various sources and can be broadly divided into three categories: domestic, industrial, and radioactive wastes, which can be categorized in the following forms: (1) thermal pollution; (2) the addition of pathogenic organisms, leading to public health hazard; (3) oil pollution; (4) the addition of inert, insoluble mineral material; (5) the addition of biodegradable organic material that will result in the depletion or complete removal of dissolved oxygen; (6) toxicity due to the presence of synthetic organic compounds and salts of heavy metals; (7) enhanced eutrophication; (8) acid deposition or discharges; and (9) radioactivity.

Domestic Sewage

The objectionable features of domestic sewage are its population of pathogenic bacteria; its high content of organic matter, which gives it a high oxygen demand; its nutrient content, which gives it the potential of supporting large populations of algae and other plants, which in turn may be as objectionable as the sewage itself; and the obvious aesthetics. The first and last of these can be overcome by proper treatment in biological sewage treatment plants, but the effluent from these plants is generally rich in nutrients. So, the effluent discharged into nearby inshore waters may result in large objectionable weed crops and other plants. An additional difficulty of disposing of domestic sewage and even the effluents from sewage treatment plants in the sea is that the density of the sewage is invariably less than that of seawater. Thus, the sewage tends to float on the surface unless introduced in a region of strong currents, will mix with seawater, and is diluted only slowly.

Industrial Wastes

These can be highly varied in composition and present a variety of special problems. The wastes may be toxic to plants or animals. They may be highly acidic or basic. If they contain large quantities of organic matter, their BOD may be objectionably high. Surface-active ingredients such as detergents may cause objectionable foaming or disrupt normal bacterial populations. The settable chemicals tend to settle to the seabed, react with mineral content there, and change the entire infrastructure of the seabed itself, which is causing accumulation of heavy metals in the sediments of several seas.

Pathogens

A variety of pathogenic organisms, including viruses, bacteria, protozoa, and parasitic worms, exist in seawater and can cause diseases in plants, animals, and people. Impacts include human illness, seafood contamination, and recreational beach closures. Pathogens are discharged to coastal waters through both point and nonpoint sources, especially from insufficiently treated sewage that is released from septic systems on land and on ships and from agriculture and storm water runoff. Higher concentrations tend to occur after storms and related overflow of sewer systems, making it difficult to interpret trends and temporal fluctuations.

Nutrients

Important parameters for monitoring nutrient pollution in coastal waters include the following: nitrogen and phosphorus concentrations, maximum bottom dissolved oxygen levels, extent and duration of anoxic and hypoxic conditions, extent of SAV, chlorophyll-*a* concentrations, turbidity, and duration and extent of algal blooms (by type). Some parameters are important in assessing the vulnerability of an area to pollutants, such as nitrogen and phosphorus, or in determining baseline conditions of the area.

Oil

Petroleum residues can contaminate marine and coastal waters through various routes: accidental oil spills from tankers, pipelines, and exploration sites; regular shipping and exploration operations, such as exchange of ballast water; runoff from land; and municipal and industrial wastes. Although the main global impact is due to tar balls that interfere with recreational activities at beaches, the impact of petroleum hydrocarbon concentrations on marine organisms in the neuston zone in the ocean—particularly fish eggs and larvae—requires more attention. Large-scale oil spills from tankers often make the headlines; yet, nonpoint sources, such as regular maritime transportation operations and runoff from land, are actually considered the main contributors to the total oil discharge into the ocean, although conclusive statistics are lacking. Runoff and routine maintenance of the oil infrastructure, it is estimated, account for more than 70% of the total annual oil discharge into the ocean. Both the number and amount of *accidental* oil spills have been monitored and seem to have declined in the past decade. A single catastrophic event can, however, influence the statistics significantly *and* have a localized, yet tremendous impact on the ecosystem.

Tables 1 and 2 show the variety of oil pollution sources and give expert estimates of the scales of distribution and impact of each of these sources on the marine environment. Even though these estimates can vary up to one to two orders of magnitude (especially for natural oil sources, atmospheric input, and river runoff), the main anthropogenic flows of oil pollution into the marine environment come from land-based sources (refineries, municipal wastes, and river runoff) and transportation activity (tanker oil transportation and shipping). Polycyclic aromatic hydrocarbons (PAHs),

Table 1. Sources and Scale of Oil Pollution Input Into the Marine Environment

	Environment		Scale of Distribution and Impact		
Types and Source of Input	Hydrosphere	Atmosphere	Local	Regional	Global
Natural:					
Natural seeps and erosion of bottom sediments	+	−	+	?	−
Biosynthesis by marine organisms	+	−	+	+	+
Anthropogenic:					
Marine oil transportation (accidents, operational discharges from tankers, etc.)	+	−	+	+	?
Marine nontanker shipping (operational, accidental, and illegal discharges)	+	−	+	?	−
Offshore oil production (drilling discharges, accidents, etc.)	+	+	+	?	−
Onland sources: sewage waters	+	−	+	+	?
Onland sources: oil terminals	+	−	+	−	−
Onland sources: rivers, land runoff	+	−	+	+	?
Incomplete fuel combustion	−	+	+	+	?

Note: +, −, and ? mean, respectively, presence, absence, and uncertainty of corresponding parameters.

Table 2. Estimates of Global Inputs of Oil Pollution Into the Marine Environment (Thousands Tons/Year of Oil Hydrocarbon)

Source	1973*	1979**	1981*	1985***	1990***
Land-based sources:					
Urban runoff and discharges	2,500	2,100	1,080 (500–1,250)	34%	1,175 (50%)
Coastal refineries	200	60	100 (60–600)	–	–
Other coastal effluents	–	150	50 (50–200)	–	–
Oil transportation and shipping:					
Operational discharges from tankers	1,080	600	700 (400-1,500)	45%	564 (24%)
Tanker accidents	300	300	400 (300–400)	–	–
Losses from nontanker shipping	750	200	320 (200–600)	–	–
Offshore production discharges	80	60	50 (40–60)	2%	47 (2%)
Atmospheric fallout	600	600	300 (50–500)	10%	306 (13%)
Natural seeps	600	600	200 (20–2,000)	8%	259 (11%)
Total discharges	**6,110**	**4,670**	**3,200**	**100%**	**2,351**

especially benzo(a)pyrene, enter the marine environment mostly from atmospheric deposition.

Table 2 illustrates the general trend of declining total input of oil pollution into the world ocean over the years. The global situation reflected in this table certainly may differ at the regional level, which depends on natural conditions, the degree of coastal urbanization, the population density, industrial development, navigation, oil and gas production, and other activities. For example, in the North Sea, offshore production input reached 28% of the total input of oil pollution in last decade, instead of the "modest" 2% on the world scale shown in Table 2, which equaled the annual input of more than 23,000 tons of oil products at the background of their general changeable flow of 120,000–200,000 tons a year in the North Sea. One can expect similar situations in other regions of intensive offshore oil and gas development, for example, in the Gulf of Mexico, Red Sea, Persian Gulf, and Caspian Sea. The persistent pollution in oil production areas in the Caspian Sea or the amounts of annual discharges (about 40 million tons of produced waters polluted by oil products) during offshore drilling in the Gulf of Mexico. At the same time, no reliable balance estimates exist for these regions.

The continental shelf of the Gulf of Mexico is also distinctive for intense seepage of natural liquid and gaseous hydrocarbons, which can lead to formation of oil slicks and tar balls on the sea surface, which makes assessing and identifying anthropogenic oil pollution more difficult. In any case, the input of oil hydrocarbons from natural sources into the Gulf of Mexico is larger than in many other areas.

In the Baltic Sea, the Sea of Azov, and the Black Sea, the leading role in oil input most likely belongs to land-based sources, which are dominated by river inflow. The Danube River alone annually brings about 50,000 tons of oil to the Black Sea, half of the total oil input of about 100,000 tons.

Observations in the Caribbean basin, where annually up to 1 million tons of oil enter the marine environment, showed that about 50% of this amount came from tankers and other ships. In the Bay of Bengal and the Arabian Sea, oil pollution inputs from tanker and other ship discharges equal, respectively, 400,000 tons and 5 million tons of oil a year. The most intense tanker traffic exists in the Atlantic Ocean and its seas, which accounts for 38% of international maritime oil transportation. In the Indian and Pacific Oceans, this portion is, respectively, 34% and 28%.

Enforcing stricter requirements on activities accompanied by oil discharges led to a global decline of oil pollution inputs in the marine environment mentioned before. A number of dramatic events show the vulnerability of making an optimistic prognosis about decreasing oil pollution at the regional and global levels. For instance, catastrophic large-scale events took place in the Persian Gulf during and after the 1991 Gulf War. Between 0.5 and 1 million tons of oil were released into coastal waters. Besides, products of combustion of more than 70 million tons of oil and oil products were emitted into the atmosphere. Another large-scale accident occurred in Russia in September–November 1994. About 100,000 tons of oil were spilled on the territory of the Komi Republic, which threatened to cause severe oil pollution for the Pechora River basin and, possibly, Pechora Bay.

It must be remembered that catastrophes, in spite of the obvious consequences and all the attention they attract, are inferior to other sources of oil pollution in their scales and degree of environmental hazard. Land-based, oil-containing discharges and atmospheric deposition of products of incomplete combustion can accordingly give 50% and 13% of the total volume of oil pollution input into the world ocean (see Table 2). These diffuse sources continuously create relatively low but persistent chronic contamination across huge areas. Many aspects of the chemical composition and biological impacts of these contaminants remain unknown.

Persistent Organic Pollutants

Persistent organic pollutants (POPs) consist of a number of synthetic compounds, including industrial polychlorinated biphenyl's (PCBs); polychlorinated dioxins and furans; and pesticides, such as DDT, chlordane, and heptachlor, that do not exist naturally in the environment. A number of POPs often persist in the environment and accumulate through the food chain or in the sediment to a toxic level that is directly harmful to aquatic organisms and humans.

The marine environment collects contaminants from the air, but also from ocean currents, rivers discharging into the ocean, and sea ice that transports POP-laden particles.

Figure 1 presents some examples of contaminant levels in seawater. Hexachlorocyclohexane dominates the picture, except for Russian waters where PCB levels are high, up to 15 nanograms per liter in the Kara Sea. These high levels seem to mirror the high input of PCBs from Russian rivers.

Levels in seawater can also be used to shed light on the mechanisms that transport contaminants to the Arctic. Detailed measurements in the Bering and Chukchi Seas show that hexachlorocyclohexane levels in the water increase along a south–north gradient, which has been suggested as evidence for a cold-condensation theory; those semivolatile contaminants condense out of the atmosphere as temperatures drop. Less volatile contaminants, such as PCBs, DDT, and chlordane, were present at lower levels in the Bering and Chukchi Seas than in more temperate latitudes.

Concentrations of organic contaminants in Arctic marine sediments are, in general, extremely low compared with freshwater sediments, and are ten to a hundred times lower than in the Baltic Sea. The most apparent geographic trends are that concentrations of PCBs, hexachlorocyclohexane, and hexachlorobenzene are higher closer to the shore along the Norwegian coast than in the open sea. They are also higher in gulfs and river mouths along the Russian coast and around Svalbard.

Heavy Metals

Heavy metals exist naturally in the environment, and it is sometimes difficult to distinguish variations developing from anthropogenic inputs and those from the

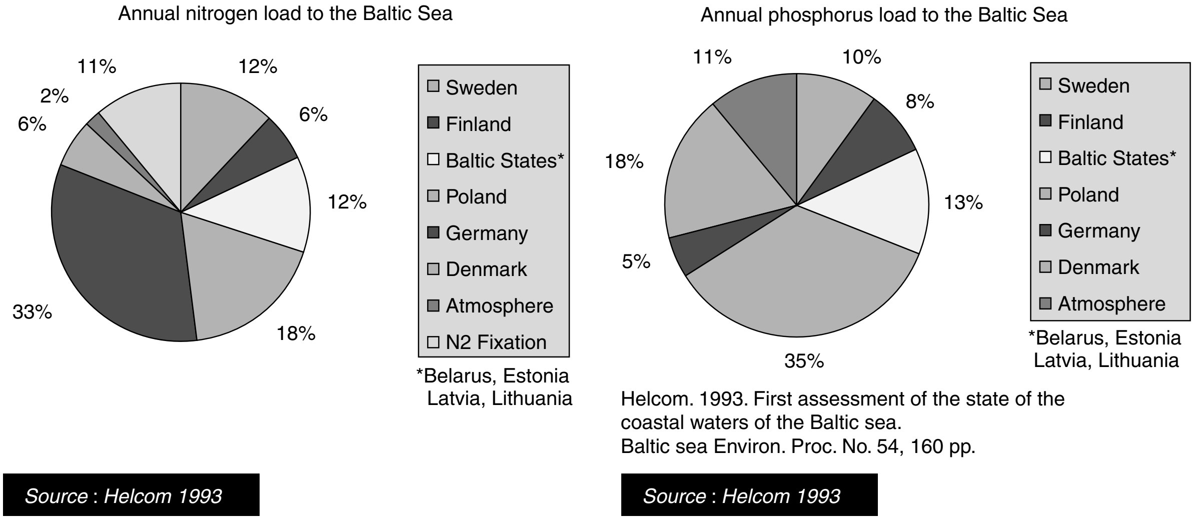

Helcom. 1993. First assessment of the state of the coastal waters of the Baltic sea. Baltic sea Environ. Proc. No. 54, 160 pp.

Source : *Helcom 1993*

Source : *Helcom 1993*

Figure 1. Nutrient release to coastal waters taking the Baltic Sea as an example.

natural hydrologic cycle and the atmosphere. Among the trace metals commonly monitored are cadmium, copper, mercury, lead, nickel, and zinc. When they accumulate through the food chain at moderate to high concentrations, some of these metals can affect the human nervous system.

The marine environment receives heavy metals from atmospheric deposition, river runoff, and local pollution. The relative importance of these sources will differ between regions. For example, rivers carrying metal-laden sediments deposit almost their entire load in the shelf seas, and only a minor portion reaches the deep ocean. Natural sources of metals are important and, in many cases, it is found that they are the main source to the marine environment.

Mining has contaminated ocean waters with several heavy metals. One documented example is in the fjord outside the Black Angel zinc mine in Greenland, where the levels of lead in the bottom water are up to 200 micrograms of lead per kilogram of water. These high lead levels are also reflected in seaweed, blue mussels, prawns, and in some fish; in capelin, lead levels are up to 5 micrograms per gram in the bone. However, no one has been able to document any biological effects in the fish. Cadmium levels in the water are also high, up to 2.5 micrograms per kilogram of water; but in contrast to lead, the animals in the fjord have cadmium levels close to background.

The cryolite mine in Ivittuut in southern Greenland has also contaminated the nearby water. Lead levels of 18 micrograms per kilogram of water have been measured. At Strathcona Sound in northern Baffin Bay, a lead-zinc mine has released lead, making concentrations in the fjord water one to two orders of magnitude higher than background concentrations in the open ocean. Some of the lead has also been taken up by seaweed and crustaceans. Outside a lead-zinc mine in east Greenland, shorthorn sculpins also have elevated levels of lead, whereas the fish outside the cryolite mine on southern Greenland have not been affected.

The mines at Ivittuut and Strathcona Sound have also contaminated their respective fjords with cadmium, but the levels are much lower than those outside the Black Angel mine. At these sites, the cadmium is not affecting the local sediment, nor are elevated levels found in nearby plants and animals.

Metal levels in Arctic Ocean water away from local sources are generally similar to global background levels. Today's global lead concentrations in oceans are generally more than ten times higher than those in prehistoric times. The levels are consistently higher in surface waters than in deeper layers. One might expect the lead levels in the upper Arctic sediments to mirror this increased long-distance transport, but this does not seem to be the case.

Recent seawater analyses from Pechora Bay and Kara and Laptev Seas show very high lead levels, ranging from 0.16 to 0.5 micrograms per kilogram of water. However, these data require confirmation before any conclusions are drawn.

Filter-feeders such as mussels take up lead from sediment particles. The concentration increases slightly with increasing shell length, indicating a moderate accumulation as the mussel ages. However, lead levels are low in crustaceans as well as in fish. The highest levels, 0.05 micrograms per gram of liver, have been recorded in Orkdalsfjorden in Norway.

Lead does not seem to accumulate in fish-eating birds or in marine mammals. In general, levels in marine mammals are low.

Cadmium Levels are High in Marine Biota. Cadmium levels in seawater fall within what could be considered natural background levels. Moreover, there is no indication from sediments that the levels have increased from preindustrial times, nor have temporal trends been detected.

An interesting phenomenon relating to cadmium is that its concentration increases farther away from the coast,

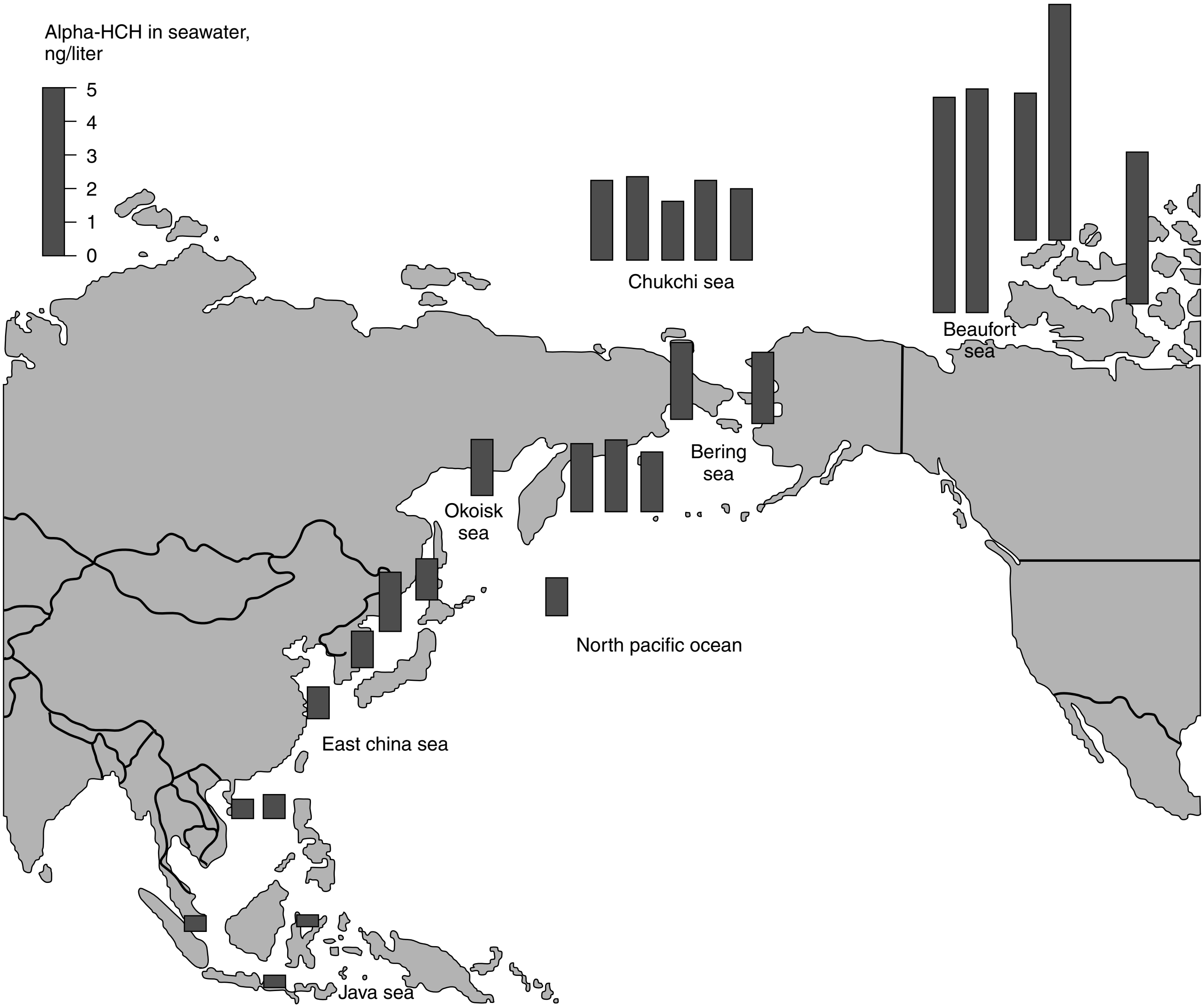

Figure 2. Alpha-hexachlorocyclohexane concentration in seawater increases from south to north, illustrating the cold-condensation effect (1).

which is probably connected with the change in salinity of the water. The result is that cadmium levels in both plants and animals are higher in the open ocean than in the inner region of large fjords, even when there are local sources contaminating the water. Cadmium accumulates with age in mussels and crustaceans. In general, the levels in crustaceans are higher than global background levels but show large variations.

Mercury levels are high and may be increasing. Several sets of data indicate that mercury levels are higher in the upper layers of Arctic marine sediments than in the layers representing preindustrial inputs see the upper right diagram of Fig. 3. Mercury is enriched even in the marine sediments taken at the North Pole. Natural processes may have caused these profiles, but they could indicate that human activities have increased the environmental mobility of existing stores of mercury.

Radioactive Wastes

These can somewhat be divided into high and low level wastes, depending on their activity. Radioactive wastes are characterized by losing their radioactivity with time. Some nuclides lose it quickly; others very slowly. A second consideration is that radioactive elements will enter the biological cycle and therefore the food web. High level radioactive wastes pose a complex problem in their disposal; the low level has been and is being disposed of directly in the sea. Reprocessing plants have added radionuclides to the sea.

Spent nuclear fuel is often processed to recover plutonium. Water used in reprocessing contains a mix of different radionuclides, and some of this contaminated water has been released routinely into the sea. In Europe, three reprocessing plants are relevant to the Arctic because of transport of radionuclides by ocean

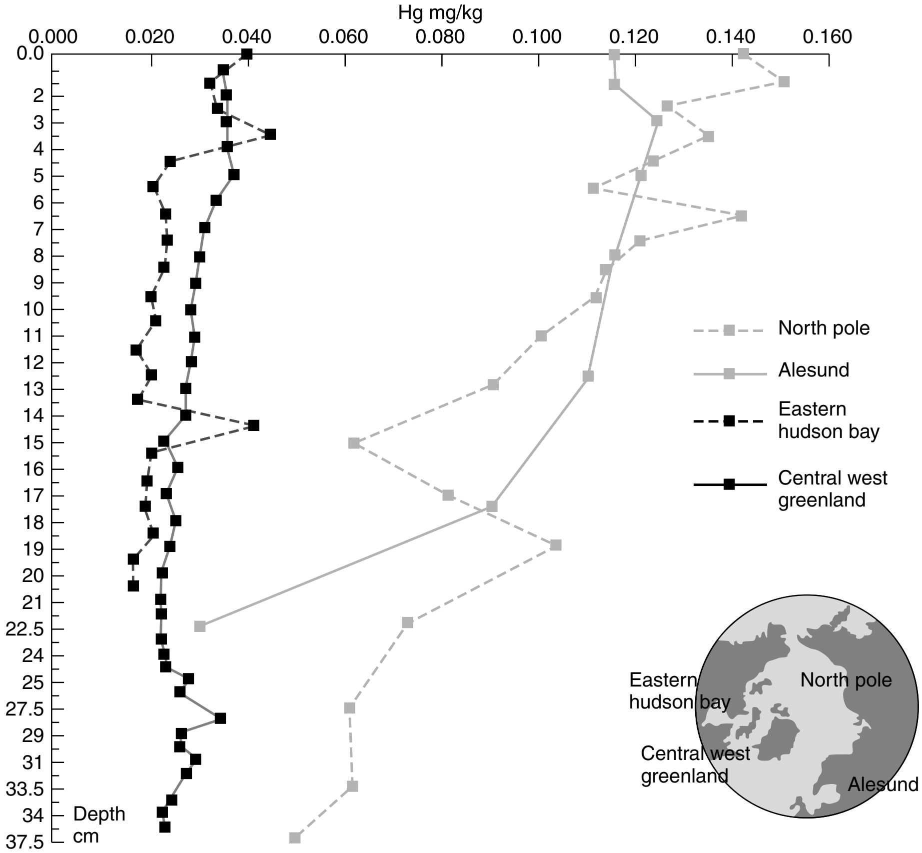

Figure 3. Mercury concentration at different depths in marine sediment cores (1).

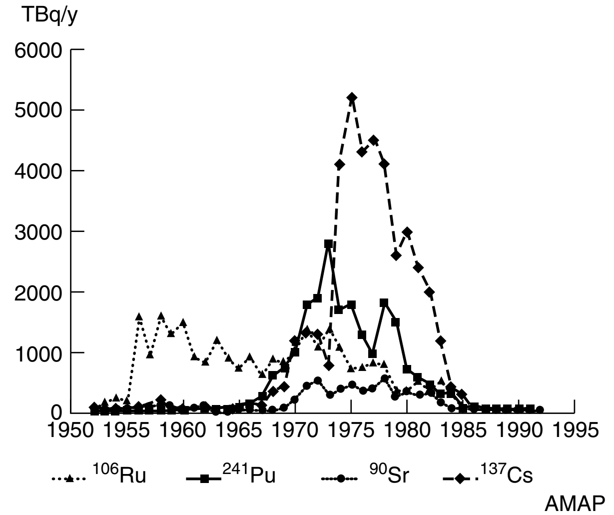

Figure 4. Discharges of beta-emitters from the Sellafield nuclear reprocessing plant, 10^{12} bequerels per year.

currents: Sellafield (formerly Windscale) in Cumbria on the northwest coast of England; La Hague near Cherbourg, France; and Dounreay in northeast Scotland. Sellafield has been the most important source of radionuclides to the Arctic marine environment because of the scale of its discharge. The effluent has been released into the sea and carried north by ocean currents. The releases, which started in 1952, are well documented. Figure 4 indicates the rates of liquid discharges from 1952 to 1992. Cesium-137 dominates. The peak of the release for most radionuclides was in the mid-to-late 1970s.

Underwater Weapon Tests Have Contaminated Chernaya Bay. Chernaya Bay is a fjord inlet, connected to the Barents Sea, on the southwestern coast of Novaya Zemlya. The former Soviet Union used the bay to conduct underwater tests of nuclear bombs in 1955 and 1957 and in the vicinity of the bay in 1961. As a result of these detonations, the bottom sediments of the bay are contaminated by elevated levels of radioactive plutonium and cesium, as well as other radioactive isotopes. However, the mobility of radionuclides in sediment is low and may at present cause only insignificant exposure of people. Exposure of biota is unknown.

Today, the inventory of plutonium in Chernaya Bay is similar to other sites of major plutonium contamination, such as the most contaminated areas of Bylot Sound (where a B-52 bomber crashed) and the Irish Sea in the vicinity of the Sellafield reprocessing plant.

Three underground nuclear detonations were carried out by the United States on Amchitka Island in the Aleutian Islands in 1965, 1969, and 1971. These detonations caused radioactive contamination of deep groundwater and rock around the shot cavities. Long-term monitoring activity is planned for this site to 2025. In 1996, aboveground radioactive contamination was detected at the site.

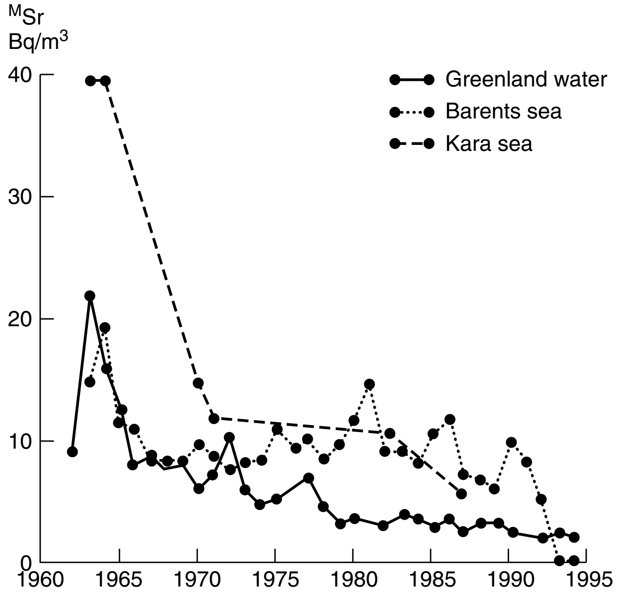

Figure 5. Time trends of activity concentrations of strontium-90 in seawater.

A major direct input of radionuclides into the marine environment has been from European nuclear reprocessing plants, particularly Sellafield on the shore of the Irish Sea. Currents transport the material along the Norwegian coast and into the Arctic Ocean. After 6 to 8 years, some of the contamination leaves the Arctic by way of the East Greenland Current, but much of it stays in the Arctic Basin much longer.

Environmental radiocesium levels have been measured since the early 1970s. As can be seen in Fig. 6, the releases of cesium-137 from Sellafield are virtually mirrored in the levels found in the Barents Sea after a transport time of 4 to 5 years. The peak in concentration in the early 1980s is probably the highest level that has ever occurred in that area of the ocean. The Chernobyl accident in 1986 added cesium to the Arctic Ocean and continues to do so via outflow from the Baltic Sea. Figure 6 shows the recent levels of cesium-137 in seawater around the Arctic. Strontium-90 has been measured in surface seawater collected around Greenland and the Barents Sea; see Fig. 5. During the past 35 years, levels in the waters around Greenland have decreased, approximately half was removed or decayed every 13.5 years. This value is probably representative of the Arctic Ocean as a whole. The highest levels of cesium-137 in people were recorded in the mid-1960s; see the Fig. 6. For the following 20 years, the human body burden decreased by a factor of 3 to 7. However, in 1986, the Chernobyl fallout changed the trend in areas directly affected by the accident.

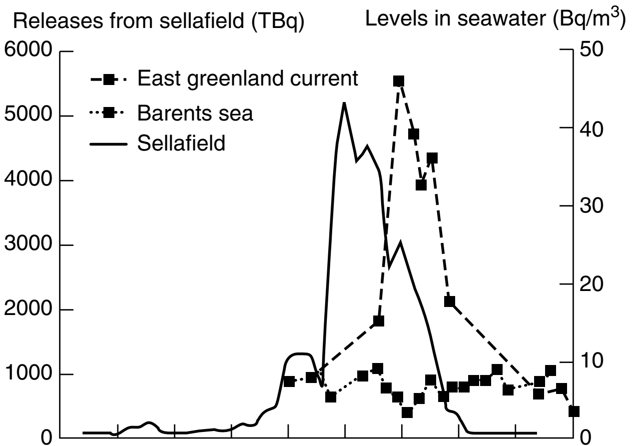

Figure 6. Releases of cesium-137 from Sellafield nuclear reprocessing plant (10^{12} bequerels per cubic meter).

Assessment of Receiving Waters

The effects of the ocean waste disposal are the result of a complex relation between two variables, concentration and time. The effect of the oceanic environment on the effluent is of critical significance. After discharge to the ocean, the effluent experiences changes in its physical and microbiological properties, which vary, each point as a function of time. An accurate prediction of pollution conditions or environmental impact depends on knowledge of the oceanographic conditions in the area. These conditions vary in time and space. A significant and representative quantitative judgement requires refined statistical analysis. A statistical description of receiving coastal waters should be based on adequate observation of an entire area for a sufficiently long period of time (at least 1 year). The factors operating should be recorded simultaneously to provide a comprehensive picture of the physical and microbiological properties of the area. The data collected should enable probability distributions of the variables to be derived to select a coherent and suitable set of design parameters. Continuous and periodic records should be taken to cover all typical oceanographic conditions at stations strategically located at different sites and should record the following:

(1) Currents (direction and speed) distributed in time/space to permit a comprehensive study of coastal water circulation patterns.
(2) A continuous record of tides and winds in the area.
(3) The density field obtained from time and space distribution of salinity and temperature enabling the stratification conditions of the coastal waters to be determined.
(4) The waves recorded periodically, covering the most critical conditions and enabling a probability distribution of the wave characteristics (height, period, and direction) to be derived for structural design.
(5) Periodic dispersion and diffusion experiments using dye tracers.
(6) Simultaneously with the oceanographic survey a program on bacterial concentration decrease should be carried out, covering typical oceanographic conditions. The experiments should be made *in situ*,

preferably in the existing continuous sewage field. Simultaneously sampling for bacteria and dye tracer direction will enable bacterial die-off rates and dilution to be investigated and the derivation of bacterial concentration decrease from both these effects. During concentration decrease experiments, the water temperature, solar radiation, and other climatic data should be collected.

Mainly tides and winds, to a lesser controlled extent by the density field, govern coastal currents near the shore. A comprehensive study of coastal circulation requires a continuous record of winds and tides and simultaneous measurements of vertical profiles of temperature, salinity, and currents. A detailed vector statistical analysis of the current field is an important requirement. The seasonal or climatological associations between oceanographic and meteorological data and *in situ* bacterial assays are of paramount importance; their interactions determine major design parameters. Statistical analysis of the sewage field is required to predict the seasonal variations in initial dilution and final height of the rise. This analysis should be based on simulations of alternative diffuser discharges at various density stratification conditions.

Oceanic Processes

To assess the oceanic equivalent dilution factors, the following formulations may be used.

Initial Dilution. The initial dilution, RI, and terminal rise height, Y_m, were estimated, assuming that the diffuser is a finite line source of buoyancy flux only, by the equation

$$\mathrm{RI} = \mathrm{FI} \cdot Y_m$$

where $\mathrm{FI} = f(F, \theta)U/q$; in which $f(f, \theta)$ is an experimental factor, which depends on θ and a type of Froude number $F = u3/B$. The current and density stratification are the oceanographic variables that directly interface in the initial dilution. We may consider a critical stratification condition in which this interface extends to a constant depth and the dilution layer increases linearly with the bottom slope according to a relation of the type $Y_m = a + b(x - x_0)$, where x is the distance from a horizontal plume perpendicular to the shore. The subsequent dilutions RM, represent a minor part of the overall receiving water reductions. Thus the total physical dilution can be evaluated as $\mathrm{RF} = \mathrm{RI} \cdot \mathrm{RM}$,

$$\mathrm{RF} = \mathrm{FI}(1 + \mathrm{FM}x)[a + b(x - x_0)]$$

All soluble pollutants experience this composite dilution factor. Small particulate matter will also disperse proportionately. Floatable matter may disperse to a lesser extent, remaining visually detectable and liable to be carried to shore by currents and surface winds. The removal of this material is required, and an initial dilution of 100 is required to provide sufficient emulsification for these materials.

Dilution Equivalent to Bacterial Concentration Decreases. Most experiments on bacterial concentration decrease have been shown to fit Chick's law very satisfactorily. $C(t) = C_0 \cdot 10 - t/T90$, where C_0 = coliform concentration at the origin, and T_{90} = time required for 90% reduction. Thus the dilution equivalent to the concentration decrease may be computed directly from

$$\mathrm{RB} = C_0/C(x) = 10'x/X_{90}$$

where $X_{90} = u \cdot T_{90}$ = distance required for the coliform concentration reduction of 90%.

Dilution Equivalent to Sedimentation. The disappearance rate of coliforms due to sedimentation of solids depends on the degree of removal in the treatment plant. Data collected by experts revealed the fact that T_{90} values decrease as the effluent TSS increases. Therefore, sedimentation effects are already incorporated in the field experimental results for the concentration decrease.

Dilution Equivalent to Treatment. The relation between removal rates and corresponding factor RT is $\mathrm{RT} = 100/[100 - T\ (\%)]$.

Dilution Equivalent to Disinfection. The dilution equivalent to disinfection is evaluated by the equation $\mathrm{RD} = 1/\mathrm{rd}$, where rd is the bacteria reduction factor. Chlorination of less treated sewage produces organic chlorine compounds, which are toxic and deleterious to the environment. Therefore, chlorination may compromise the already recognized ecologically beneficial effects of enrichment of coastal waters by the supply of nutrients and organic matter from those effluents.

Overall Equivalent Dilution. Assuming that $\mathrm{RS} = 1$, the overall equivalent dilution can be evaluated by the equation

$$\mathrm{R}_{\mathrm{total}} = \mathrm{RT} \cdot \mathrm{RD} \cdot \mathrm{RI} \cdot \mathrm{RM}(x) \cdot \mathrm{RB(x)}$$

$$\mathrm{R}_{\mathrm{total}} = \mathrm{RR} \cdot f(F, \theta)(bu/q)[a + b(x - x_0)] \cdot [1 + (KL/b_o)x] \cdot 10'x/u \cdot T_{90}$$

For a given set of parameters, $R_{\mathrm{total}} = \mathrm{CE/CP}$, Q and the oceanic parameters are $u(u, \theta), T_{90}$ and KL are the total equivalent approximation, a linear function of the diffuser length and an exponential quadratic function of the outfall length.

The Effect of Current. Separating the current factors including the current speed u from the above equation, a dilution function may be defined as $Fu = u \cdot 10'x/uT_{90}$, which represents the effect of current on overall oceanic dilution. The figure given shows curves of Fu as a function of u for the various values of the parameter $u_{90} = x/T_{90}$, that is, the velocity for a 90% concentration decrease. It may be seen that the curves show "inflections" that have been connected by the line. The inflection increases as u_{90} decreases or as T_{90} increases. This inflection divides the graph into two domains:

The initial dilution domain (linear)

The concentration decrease dilution factor dilution factor domain (exponential).

For strong currents, the initial dilution is the dominant factor, and for a given T_{90}, the only way to increase Fu is by initial dilution mixing depth, which can be attained by increasing the outfall length. For light currents, the dilution factor equivalent to die-off dominates.

Thus, the current data is collected, as it is useful in the design in

1. analysis of coastal circulation hydrodynamics;
2. prediction of initial dilution;
3. prediction of far field dilution and transport;
4. Prediction of waste field impaction probability.

Thus, a comprehensive oceanographic study is necessary to describe the characteristics of the receiving waters. Therefore, an adequate program of oceanographic investigation *in situ*, including bacterial concentration decrease, is obligatory.

The existence of elevated levels of contaminants in zones of high bioproductivity is extremely alarming ecologically. These zones include the water layer up to 100 m from the water surface (photic layer) and boundaries of natural environments (water–atmosphere and water–bottom sediment, as previously mentioned) as well as enclosed seas, estuaries, and coastal and shelf waters. In particular, in shelf and coastal zones, which occurs in 10% of the world ocean surface and less than 3% of its volume, the most intense processes of bioproduction, including self-reproduction of the main living resources of the sea, take place. The main press of anthropogenic impact is also concentrated here.

EFFECTS OF POLLUTANTS ON MARINE HABITAT

To estimate the hazard of different pollutants, we should take into account not only their hazardous properties, but other factors, too. These include the volumes of their input into the environment, the ways and scale of their distribution, the patterns of their behavior in water ecosystems, their ability to accumulate in living organisms, the stability of their composition, and other properties, such as the extreme diversity of marine pollution components, the variety of their sources, the scales of distribution, and the degree of hazards. Pollutants can be classified in different ways, depending on their composition, toxicity, persistence, sources, volumes, and so on.

To analyze large-scale pollution and its global effects, it is common to distinguish groups of the most widespread pollutants, which include chlorinated hydrocarbons, heavy metals, nutrients, oil hydrocarbons, surface-active substances, and artificial radionuclides. These substances form the so-called background contamination that exists now any place in the hydrosphere.

Depending on the type of impact on water organisms, communities, and ecosystems, the pollutants can be grouped in the following order of increasing hazard:

- substances that cause mechanical impacts (suspensions, films, solid wastes) that damage the respiratory organs, digestive system, and receptive ability;
- substances that provoke eutrophic effects (e.g., mineral compounds of nitrogen and phosphorus, and organic substances) that cause mass rapid growth of phytoplankton and disturb the balance, structure, and functions of water ecosystems;
- substances that have saprogenic properties (sewage with a high content of easily decomposing organic matter) that cause oxygen deficiency followed by mass mortality of water organisms and appearance of specific microflora;
- substances causing toxic effects (e.g., heavy metals, chlorinated hydrocarbons, dioxins, and furans) that damage the physiological processes and functions of reproduction, feeding, and respiration;
- substances with mutagenic properties (e.g., benzo(a)-pyrene and other polycyclic aromatic compounds, biphenyls, radionuclides) that cause carcinogenic, mutagenic, and teratogenic effects.

Some of these pollutants (especially chlorinated hydrocarbons) cause toxic and mutagenic effects. Others (decomposing organic substances) lead to eutrophic and saprogenic effects. Oil and oil products are a group of pollutants that have complex and diverse composition and various impacts on living organisms—from physical and physicochemical damage to carcinogenic effects.

Discharge of heated waters can change the structure and function of coastal marine communities. Impacts of fly ash from coal-fired power plants, hot salty water, and residual chlorine are also important. Dumping of fly ash in coastal waters and into the atmosphere has caused severe impacts on spinner dolphins and mangroves in an area of the south coast of India, and has reportedly changed the number of species of plankton.

Effects of Marine Oil Spills

Oil spills can have a serious economic impact on coastal activities and on those who exploit the resources of the sea. In most cases, such damage is temporary and is caused primarily by the physical properties of oil creating nuisance and hazardous conditions. The impact on marine life is compounded by toxicity and tainting effects resulting from the chemical composition of oil, as well as by the diversity and variability of biological systems and their sensitivity to oil pollution.

Biological Effects of Oil. Simply, the effects of oil on marine life are caused by either the physical nature of the oil (physical contamination and smothering) or by its chemical components (toxic effects and accumulation leading to tainting). Marine life may also be affected by cleanup operations or indirectly through physical damage to the habitats in which plants and animals live.

The main threat posed to living resources by the persistent residues of spilled oils and water-in-oil emulsions ("mousse") is physical smothering. The animals and plants most at risk are those that could come into contact with a contaminated sea surface: marine mammals and reptiles; birds that feed by diving or form flocks on the sea; marine life on shorelines; and animals and plants in mariculture facilities.

The most toxic components in oil tend to be those lost rapidly through evaporation when oil is spilled. Because of this, lethal concentrations of toxic components leading to large-scale mortality of marine life are relatively rare, localized, and short-lived. Sublethal effects that impair the ability of individual marine organisms to reproduce, grow, feed, or perform other functions can be caused by prolonged exposure to a concentration of oil or oil components far lower than will cause death. Sedentary animals in shallow waters such as oysters, mussels, and clams that routinely filter large volumes of seawater to extract food are especially likely to accumulate oil components. Although these components may not cause any immediate harm, their presence may render such animals unpalatable if they are consumed by humans, due to the presence of an oily taste or smell, which is a temporary problem as the components that cause the taint are lost when normal conditions are restored.

The ability of plants and animals to survive contamination by oil varies. The effects of an oil spill on a population or habitat must be viewed in relation to the stresses caused by other pollutants or by any exploitation of the resource. In view of the natural variability of animal and plant populations, it is usually extremely difficult to assess the effects of an oil spill and to determine when a habitat has recovered to its prespill state. In recognition of this problem, detailed prespill studies are sometimes undertaken to define the physical, chemical, and biological characteristics of a habitat and the pattern of natural variability. A more fruitful approach is to identify which specific resources of value might be affected by an oil spill and to restrict the study to meeting defined and realistic aims related to such resources.

Impact of Oil on Specific Marine Habitats. Within each habitat, a wide range of environmental conditions prevails, and often there is no clear division between one habitat and another.

Plankton is a term applied to floating plants and animals carried passively by water currents in the upper layers of the sea. Their sensitivity to oil pollution has been demonstrated experimentally. In the open sea, the rapid dilution of naturally dispersed oil and its soluble components, as well as the high natural mortality and patchy, irregular distribution of plankton, make significant effects unlikely.

In coastal areas, some marine mammals and reptiles, such as turtles, may be particularly vulnerable to adverse effects from oil contamination because of their need to surface to breathe and to leave the water to breed. Adult fish that live in nearshore waters and juveniles in shallow water nursery grounds may be at greater risk to exposure from dispersed or dissolved oil.

The risk of surface oil slicks affecting the sea bed in offshore waters is minimal. However, restrictions on the use of dispersants may be necessary near spawning grounds or in some sheltered, nearshore waters where the dilution capacity is poor.

The impact of oil on shorelines may be particularly great where large areas of rocks, sand, and mud are uncovered at low tide. The amenity value of beaches and rocky shores may require the use of rapid and effective cleanup techniques, which may not be compatible with the survival of plants and animals.

Marsh vegetation shows greater sensitivity to fresh light crude or light refined products, although weathered oils cause relatively little damage. Oiling of the lower portion of plants and their root systems can be lethal, whereas even a severe coating on leaves may be of little consequence especially if it occurs outside the growing season. In tropical regions, mangrove forests are widely distributed and replace salt marshes on sheltered coasts and in estuaries. Mangrove trees have complex breathing roots above the surface of the organically rich and oxygen-depleted muds in which they live. Oil may block the openings of the air breathing roots of mangroves or interfere with the trees' salt balance, causing leaves to drop and the trees to die. The root systems can be damaged by fresh oil that enters nearby animal burrows; the effect may persist for some time and inhibit recolonization by mangrove seedlings. Protection of wetlands, by responding to an oil spill at sea, should be a high priority because physical removal of oil from a marsh or from within a mangrove forest is extremely difficult.

Living coral grows on the calcified remains of dead coral colonies, which form overhangs, crevices, and other irregularities, inhabited by a rich variety of fish and other animals. If the living coral is destroyed, the reef itself may be subject to wave erosion. The proportion of toxic components, the duration of oil exposure, as well as the degree of other stresses, largely determine the effects of oil on corals and their associated fauna. The waters over most reefs are shallow and turbulent, and few cleanup techniques can be recommended.

Birds that congregate in large numbers on the sea or shorelines to breed, feed, or moult are particularly vulnerable to oil pollution. Although oil ingested by birds during preening may be lethal, the most common cause of death is from drowning, starvation, and loss of body heat following damage to plumage by oil.

Impact of Oil on Fisheries and Mariculture. An oil spill can directly damage the boats and gear used for catching or cultivating marine species. Floating equipment and fixed traps extending above the sea surface are more likely to become contaminated by floating oil, whereas submerged nets, pots, lines, and bottom trawls are usually well protected, provided they are not lifted through an oily sea surface. Experience from major spills has shown that the possibility of long-term effects on wild fish stocks is remote because the normal overproduction of eggs provides a reservoir to compensate for any localized losses.

Cultivated stocks are more at risk from an oil spill: Natural avoidance mechanisms may be prevented in the case of captive species, and oiling of cultivation equipment may provide a source for prolonged input of oil components and contamination of the organisms. The use of dispersants very close to mariculture facilities is ill advised because tainting by the chemical or by the dispersed oil droplets may result.

An oil spill can cause loss of market confidence because the public may be unwilling to purchase marine products from the region, irrespective of whether the seafood is actually tainted. Bans on fishing and harvesting marine products may be imposed following a spill to maintain market confidence and to protect fishing gear and catches from contamination.

Mercury levels in marine animals, including bivalves and crustaceans, are generally low, whereas mercury seems to accumulate in fish. The highest values in fish are from northern Canada.

For seals and whales, concentrations often exceed 0.5 micrograms per gram of muscle, especially in older individuals. Livers from ringed seals in the western Canadian Arctic have very high levels of mercury; up to 205 micrograms per gram of liver have been measured. Levels in livers of bearded seals from the Amundsen Gulf are higher than those of both global background and other Arctic areas, as are mercury levels in toothed whales and polar bears. Some of the highest levels, 280 microgram per gram liver (wet weight), have been recorded in pilot whales from the Faroe Islands.

The effects of these mercury levels on the animals are difficult to assess, because some of the mercury may be inactivated by high selenium levels. Moreover, the scientific focus so far has been on tissues relevant for human consumption, and very little information is available on the target organs for mercury, such as the brain. There are no effect studies from the Arctic. However, even for the most exposed animal populations in the western Canadian Arctic and in Greenland, selenium should be abundant enough to protect against mercury poisoning.

Mercury is a major concern because the levels in some animals high in the food chain indicate that the environmental load may have increased in recent years. For example, mercury levels in ringed seals from western Canada show that they accumulated mercury about three times faster during the late 1980s and early 1990s than in the early 1970s. Similar increases have been seen in ringed seals from northwest Greenland taken in 1984 and 1994 and in beluga livers from the western Canadian Arctic. Interpreting these findings is difficult because natural variations that may affect the trends are unknown. Moreover, other data, such as those from Atlantic walrus and ringed seal from central-east Greenland, have not indicated any temporal trends. Very little information is available on temporal trends in Arctic marine fish, but measurements from the Baltic Sea from 1980 to 1993 seem to confirm observations that mercury levels are increasing.

The seas are being polluted by organic and inorganic wastes from sewage, from agricultural and industrial wastes, and from runoff containing oil, hydrocarbons, and heavy metals. All of these contribute to sediment runoff and increased turbidity. Siltation of coral regions is also caused by excessive deforestation and land clearing for commercial crops. Construction and land reclamation has caused changes in water circulation and has increased sedimentation. On the coral reefs, there has been extensive overexploitation of resources by heavy fishing pressure, including very destructive methods such as blasting, coral mining, and cyanide poisoning for live fish collection.

Industrial Pollution

Industrial pollutants that affect coral reefs include nutrients from sewage and organic matter, fertilizer runoff, detergents containing phosphorus, and thermal discharge—the heated water from the cooling systems of power plants and other industries. These all cause nutrient overload, the growth of aquatic plant life, and depletion of dissolved oxygen, or eutrophication, which retards coral growth by decreasing light penetration and changing the dynamics of fish assemblages. Other industrial pollutants include heavy metals and other toxic substances.

The coral reefs bordering major cities throughout Southeast Asia have been largely destroyed. Pollution from oil refineries and drilling platforms, it has been shown, kill reef fish and have negative effects on growth rates, recruitment, and feeding of corals. Thermal pollution from hot water discharge from industrial areas is an additional threat to reef species, many of them cannot withstand sudden and drastic increases in temperature.

Sedimentation. Land-based human activities often cause sedimentation, a major source of reef degradation. As more people move to coastal cities on the South China Sea, there has been a big increase in construction and land reclamation. Land reclamation and sedimentation have been particularly intensive in Singapore. Land was reclaimed by dumping sand and dirt directly onto coral reef flats and shallow water.

These add to the erosion of beaches and sediment runoff that smothers corals and leads to the degradation of a reef. Increased sedimentation also leads to a change in the composition of marine fauna, favoring more resilient species. Sedimentation also comes from soil erosion from unsound agricultural practices, mismanagement of watersheds, exploitation of mangroves, land reclamation and construction, oil drilling, and dumping of terrestrial and marine mine tailings.

Overfishing. Overfishing is a force extremely destructive to corals in the South China Sea. Densities of fish are greatly decreased by overfishing. Coral is damaged by destructive fishing techniques and by removal for trade. It is estimated that 10–15% of the total fish yield in the Philippines comes from coral reef fisheries.

Fishing degrades a reef in several ways. Destructive and illegal fishing methods are common, especially in the Philippines, Indonesia, and Malaysia. These methods include dynamite blasting and cyanide fishing. Overfishing not only depletes fish stocks of target species but also changes the dynamics of the entire reef. Decreases in herbivores can lead to algal blooms that overtop coral growth and can cause mass mortality.

Blooms of noxious algae have increased in the past 20 years worldwide and are being blamed on inputs of excess nutrients due to human activities. Some of these noxious algae produce powerful nerve toxins that can cause massive fish kills or even kill a person who unsuspectingly eats shellfish that was harvested from waters tainted with toxic algae. The case of the "Cell from Hell" now blooming in East Coast waters (North Carolina, Virginia, Maryland) is especially noteworthy. Until recently, *Pfiesteria* was only a curiosity of academic specialists. In the past few years, this organism has been blamed for fish kills unprecedented in their size and has been linked to neurological damage in people who worked or swam in these waters (memory loss, learning difficulties, and decreases in white blood cell content upward of 20% have been recorded in people who were exposed to *Pfiesteria*). Blooms of *Pfiesteria* have been linked to nutrient enrichment of coastal waters due to nonpoint pollution from agriculture. Nutrients in waters allow huge population increases of toxic organisms in water that were unknown or rare. The U.S. EPA has pledged to adopt new standards for nutrient inputs to waters.

It is hard to imagine an organism more bizarre than *Pfiesteria*. When no fish prey are present, it goes into a cyst form and settles to the bottom, lying dormant in the sediments. It can also emerge to form an amoeba that feeds on algae in the water column, and even can become a photosynthetic plankton-like organism, except that it "steals" the chloroplasts of algae from its algal prey and uses photosynthesis only to supplement its nutrient supply in the water column. In the presence of certain species of fish, however, it becomes a "monster" predator capable of mass fish kills. As a "predatory" dinoflagellate, it produces different types of toxins that do an incredible array of damage to fish. Some toxins attack internal organs. Another works on the fish immune system. And one toxin actually strips the skin off of the fish. Those who have witnessed the power of *Pfiesteria* report thousands of fish flopping and thrashing on the water surface, and fish actually beaching themselves, fleeing the water as if on fire.

SUMMARY

The introduction by man, directly or indirectly, of substances or energy into the marine environment (including estuaries) results in such deleterious effects as harm to living resources, hazards to human health, and hindrance to marine activities, including fishing, impairment of quality for use of seawater, and reduction of amenities. As the uses of coastal waters and the ocean have increased, pollution of the ocean waters has increased in turn. River pollution has also had an impact on the ocean as the rivers transport material to the ocean and, as a result, make it the ultimate sink for the world's waste.

The following chart summarizes the sources of wastes and their effects.

Sources and Effects of Marine Pollution.

Type	Primary Source/Cause	Effect
Nutrients	Runoff approximately 50% sewage, 50% from forestry, farming, and other land use. Also airborne nitrogen oxides from power plants, cars, etc.	Feed algal blooms in coastal waters. Decomposing algae depletes water of oxygen, killing other marine life. Can spur algal blooms (red tides), releasing toxins that can kill fish and poison people.
Sediments	Erosion from mining, forestry, farming, and other land-use; coastal dredging and mining.	Cloud water; impede photosynthesis below surface waters. Clog gills of fish. Smother and bury coastal ecosystems. Carry toxins and excess nutrients.
Pathogens	Sewage, livestock.	Contaminate coastal swimming areas and seafood, spreading cholera, typhoid, and other diseases.
Alien species	Several thousand per day transported in ballast water; also spread through canals linking bodies of water and fishery enhancement projects.	Outcompete native species and reduce biological diversity. Introduce new marine diseases. Associated with increased incidence of red tides and other algal blooms. Problem in major ports.
Persistent toxins (PCBs, heavy metals, DDT, etc.)	Industrial discharge; wastewater discharge from cities; pesticides from farms, forests, home use etc.; seepage from landfills.	Poison or cause disease in coastal marine life, especially near major cities or industry. Contaminate seafood. Fat-soluble toxins that bioaccumulate in predators can cause disease and reproductive failure.

Oil	46% from cars, heavy machinery, industry, other land-based sources; 32% from oil tanker operations and other shipping; 13% from accidents at sea; also offshore oil drilling and natural seepage.	Low level contamination can kill larvae and cause disease in marine life. Oil slicks kill marine life, especially in coastal habitats. Tar balls from coagulated oil litter beaches and coastal habitat. Oil pollution is down 60% from 1981.
Plastics	Fishing nets; cargo and cruise ships; beach litter; wastes from plastics industry and landfills.	Discarded fishing gear continues to catch fish. Other plastic debris entangles marine life or is mistaken for food. Plastics litter beaches and coasts and may persist for 200 to 400 years.
Radioactive substances	Discarded nuclear submarine and military waste; atmospheric fallout; also industrial wastes.	Hot spots of radioactivity. Can enter food chain and cause disease in marine life. Concentrate in top predators and shellfish, which are eaten by people.
Thermal	Cooling water from power plants and industrial sites.	Kills off corals and other temperature-sensitive sedentary species. Displaces other marine life.

READING LIST

Brown, L. (1997). Can we raise grain yields fast enough? *Worldwatch* **10**(4): 9–17.

Engleman and LeRoy. (1996). *Sustaining Water: Population and the Future of Renewable Water Supplies*. Population Action International.

EPA. (1994). *National Water Quality Inventory, 1992: Report to Congress*. Office of Wetlands, Oceans and Watersheds, US EPA, Washington, DC.

Kane, H. (1997). Eco-farming in Fiji. *Worldwatch* **10**(4): 29–34.

Nixon, S. et al. The fate of nitrogen and phosphorus at the land-sea margin of the North Atlantic Ocean. *Biogeochemistry* **35**: 141–180.

NRC (National Research Council). (1993). *Managing Wastewater in Coastal Urban Areas*. NRC, Washington, DC.

NRC (National Research Council). (1994). *Priorities for Coastal Science*. NRC, Washington, DC.

Hotta, K. and Dutton, I.M. (Eds.). (1995). *Coastal Management in the Asia-Pacific Region: Issues and Approaches*. Japan International Marine Science and Technology Federation, Tokyo.

Web Sites

http://www.nos.noaa.gov/icri/state.html
http://www.cutter.com/osir/
http://www.igc.apc.org/wri/indictrs/reefrisk.htm
http://www.enn.com
http://www.env.gov.sg
Journals and Newspapers
Environet
The Straits Times
Science and Technology-The Hindu

TRACE ELEMENT POLLUTION

Bobby J. Presley
Texas A&M University
College Station, Texas

INTRODUCTION

All three words in the term “trace element pollution” need to be defined. Defining an element, that is a chemical element, is relatively straightforward. It is a substance consisting entirely of atoms having the same number of protons. All such substances are listed by name and symbol on the periodic charts of the elements found in many science textbooks. Definitions for “trace” and “pollution” are not so straightforward.

It is sometimes convenient to classify the chemical elements making up a complex substance or matrix into major, minor, and trace, based on their relative amounts (concentrations). Most authors designate those elements present in a matrix at one part per million or less by weight as trace elements. Although trace elements are in low concentration in the environment, they can be either essential or harmful to organisms, depending on the element and the circumstances.

A definition of pollution can be inferred from the widely accepted definition of marine pollution given by the United Nations Group of Experts on the Scientific Aspects of Marine Pollution (GESAMP). They say pollution is “the introduction by man, directly or indirectly, of substances or energy to the marine environment resulting in such deleterious effects as harm to living resources; hazards to human health; hindrance of marine activities including fishing; impairment of the quality for use of seawater; and reduction of amenities.” This definition can be applied to all environments, not just to the marine environment.

By the GESAMP definition, pollution must be harmful and must be caused by human activity. In some cases, it is easy to show both a human cause and harm to the environment. In other cases, one or both parts of the definition can be hard to prove. For example, it might be possible to show that the concentration of a trace element, such as mercury, is elevated above normal values in, for example, fish. It might be harder to show harm to the fish or to the consumer of the fish and it might be hard to show that the enrichment is because of human activity. In such a case, it might be better to say contamination has been documented, but not pollution.

ABUNDANCE, DISTRIBUTION, AND BEHAVIOR OF TRACE ELEMENTS

Trace elements occur naturally in the environment. Unlike pesticides, plastics, organic solvents, and other manufactured products, the mere presence of trace elements does not imply pollution, contamination, or even human activity. Sufficiently sensitive analytical techniques can detect some amount of all elements in almost any substance. To determine the amount accurately and precisely is, however, a challenge, and to decide what portion of the element is natural and what portion is because of human activity is an even bigger challenge.

In order to evaluate the significance of trace element occurrences in the environment, information is needed on:

1. Amounts (concentrations) in various compartments (air, soil, water, organisms, etc.).
2. Sources to the environment (both natural and human).
3. Transport mechanism and pathways between compartments (continents to oceans, water to organisms, etc.).
4. Transfer mechanism within compartments (shallow to deep water, fish gill surface to other organs, etc.).
5. Ultimate fate of the element (burial in sediments, mixing throughout the ocean, etc.).
6. Effect of the element on organisms (both short and long term).

As a first step in determining the significance of a given amount of a trace element in a particular environment, it is essential to know the ranges in concentration expected to occur naturally in various media (soil, water, organisms, etc.). Most trace elements were tied up in igneous rocks on the primitive earth. Table 1 gives average concentrations of selected elements in the uppermost part of the earth's crust along with data for organisms, soil, sediment, and water. These values can be considered to be "background" concentrations or at least concentrations not greatly influenced by human activity.

As igneous rocks "weather" to give soil and sediment some fraction of each element in the rock, it becomes dissolved in rivers, lakes, and the ocean. For most trace elements, however, a large fraction is retained in the solids during weathering. For this reason, the natural (background) trace element concentration in soils and sediments varies, depending on the igneous rocks they came from.

Many trace elements behave similarly during weathering, soil formation, erosion, and deposition. For example, fine-grained clay minerals become enriched in most trace elements and quartz sand and carbonate minerals become depleted. It is, therefore, important to have information on the grain size and mineralogy of soils and sediments when evaluating their trace element content. Thus, a given concentration, say 20 ppm Cu, might be background for a clay sediment but be contamination in a sand.

Reliable data on trace elements dissolved in rivers, lakes, and seawater is more difficult to produce than is data on rocks, soil, and sediment. Concentrations are generally a thousand or more times lower (Table 1), and this causes sensitivity and matrix interference problems for even the newest analytical instruments. It is also difficult to avoid contaminating samples with trace elements during

Table 1. Average Concentration of Selected Elements in Various Materials. Values for Sediment, Soil, Crust, Oysters and Fish are μg/g Dry Weight (ppm) Except Ca and Fe in Percent Dry Weight. Values for Water are μg/l (ppb)

Material	Ref.	As	Ba	Ca	Cr	Cu	Fe	Mn	Ni	Pb	V	Hg	Zn
Continental Crust	1	1.5	550	3.00	35	25	3.5	600	20.	20.	60	–	71
Gulf Coast Seds	2	6	–	–	44	10	1.8	330	16	15	–	0.050	60
Gulf Slope Seds	3	8.6	660	11.3	54	27	2.76	300	38	17	100	0.028	81
San Joaquin Soil	4	17.7		1.9	130	34.6		538	88	18.9	112	1.4	106
Sediment Criteria	5	55	–	–	145	390	–	–	50	110	–	1.3	270
Average Seawater	6	1.7			0.2	0.25	0.05	0.2	0.5	0.002	1.5	0.001	0.4
Miss. River Water	7		65			1.5	2.2	1.4	1.5	0.008	1.0		0.27
Seawater Criteria	8	69			1100	4.8			74	210		1.8	90
Fresh Water Criteria	9	340			16	13			470	65		1.4	120
Gulf Oysters	10	10.3			0.57	146	294	14.4	1.77	0.64		0.13	2150
Tuna Fish	11	3.8		100	0.75	2.8	72	0.6	0.5	0.5		4.1	17.4

1. Average metal levels in upper continental crust (95% igneous rock). Taylor and McClennon (1).
2. Median estuarine (inshore) surface sediment metal levels from the U.S. Environmental Protection Agency's northern Gulf of Mexico (GOM) Environmental Mapping and Assessment Program 1991–1993.
3. Average surface sediment metal levels observed among 43 stations ion the Gulf of Mexico Slope. BJ Presley, unpublished
4. Agricultural soil from the San Joaquin Valley, CA. US NIST Standard Reference Material #2709.
5. ER-M values from Long and Morgan (2), indicating sediment metal levels at which biological effects are often seen.
6. Average seawater values from Bruland, (3).
7. Average Mississippi River dissolved values from Shiller, (4).
8. Maximum Contaminant Concentration for seawater, US EPA, (5).
9. Maximum Contaminant Concentration for fresh water, US EPA, (5).
10. Average concentration in 485 pooled samples of 20 oysters each from the Gulf of Mexico. Presley (6).
11. Fillets from 16 Large Mediterranean Sea Tuna. International Atomic Energy Agency reference material # 350.

collection, storage, and analysis. For these reasons, nearly all of the dissolved trace element data published before 1970, and much of the recent data, is unreliable. Much of the published data for both fresh and seawater is too high by factors of up to 100. Only data that has been produced by a lab using a well-documented quality assurance program should be accepted.

Trace element concentrations in organisms are generally intermediate between those in sediments and those in water. However, concentrations vary widely with the specie of organism and with the specific organ within organisms. Livers, for example are enriched in some elements and kidneys in others and some species of organisms are highly enriched in one or another trace element. Trace element concentrations in organisms can also change with season, life stage, health, food supply, etc. Thus, identification of abnormal trace element concentrations is difficult unless reliable data is available for the same organism over a wide area and/or over some time period.

Good general compilations of reliable trace element data for organisms are not as available as are those for soil and sediment. General guidance can be acquired from Furness and Rainbow (7) or similar publications. Many recent journal articles give trace element data for specific organisms from specific locations, but it is not always clear which are background levels.

In addition to natural sources, many different human activities can add trace elements to the environment. Mining and metal processing are classic sources of contamination, but other manufacturing, transportation, and waste disposal practices can also be important. In the United States, the EPA's "Superfund" program has spent many millions of dollars to clean up trace element contamination at dozens of sites around the country. The kinds of practices that led to this gross contamination are very rare today. Environmental regulations and public pressure have caused industry to greatly reduce trace metal releases to the environment. However, as the world population grows, the Earth's surface is increasingly disturbed by agriculture, petroleum production, forestry, urban development, civil conflicts, and war, all of which make trace elements more available for uptake by humans and other organisms.

As discussed above, numerous possible sources exist of trace elements to the environment. It is almost always difficult to determine which possible source is most important for any given element at any given location. Although it is important to identify the source of trace elements, their environmental impact depends not on source but on concentration and behavior. Behavior, including mobility, transport, transfer, and biological uptake, depends strongly on the chemical and physical form of the trace element. In this respect, the size of the trace element specie or the particle with which it is associated is critical, as this will control its transport and settling behavior in air and water. A given trace element will behave differently physically, chemically, and biologically in each of its different forms, and it will partition itself among the various possible forms in response to environmental conditions. It is important to note that many trace elements are "particle-reactive" and will quickly associate with particles if added to the environment in a dissolved form. Trace element concentrations are almost always much higher in particles than in dissolved forms.

Although the behavior of a trace element, including its biological behavior, depends on its form, there is some form of most trace elements that will affect the health of organisms, including humans. At least 20 trace elements have long been known to be essential to health [e.g., (8)]. Diseases because of trace element deficiencies are well known among both humans and other organisms. A number of trace elements such as Cu, Ni, Zn, and Se are essential to life at very low concentrations but toxic at slightly higher concentrations. Good data on trace element concentrations in the environment are needed in order to know whether too little or too much of a giver element is present.

The toxic effects of both essential and nonessential trace elements are well known, in the case of As and Pb, human toxicity has been known for more than 2000 years. For other trace elements, toxic effects are less well recognized. In general, however, for all trace elements, an optimal concentration in the environment and in the organism gives optimal function (growth, reproduction, etc.) and higher or lower concentrations result in less than optimal function and possibly death.

In order for a trace element in the environment to have an effect on an organism, the element must, of course, be taken up by the organism. For plankton and other aquatic plants, this uptake is directly from solution, but for animals, some, or most, of the uptake might be from food or from ingestion of nonfood particles. In any case, at some point the trace element must be in a soluble form and be transferred across cell membranes and possibly transferred to some vital organ within the organism. The form of the trace element is very important in controlling these transfers and the resulting effects, but both environmental conditions (pH, temperature, etc.) and the type of organism and its condition (age, health, etc.) also play a role. Factors that influence the toxicity of trace elements have been discussed by Bryan (9), Luoma (10), and many others, and the large differences in sensitivity to trace elements exhibited by different organisms are well known [e.g., (11)].

ASSESSING BIOLOGICAL IMPACTS

The effect of a trace element on organisms depends on the abundance, distribution, and behavior of the trace element. As discussed above, these are difficult to determine and are subject to complex and incompletely understood processes. The environmental impact of waste disposal or other human activity is, therefore, often controversial. Environmental groups and industry often engage in public fingerpointing and lawsuits over specific activities. Often, more money is spent on lawyers than on attempts to scientifically document impacts. One reason for this is the difficulty in clearly documenting harmful effects in the field, especially at the population or ecosystem level. Laboratory toxicity testing is not easy, but it can usually show dose-response relationships that allow establishment of trace element concentrations above which harmful effects to a given organism are likely to result. Such results are, however, usually difficult to apply to the complex conditions in the field, as is discussed below.

The simplest laboratory toxicity tests are those that use death of the organism as the only indicator of effect. This crude measurement has been much criticized, but it does establish the rough relative toxicity for various trace elements to various organisms. This test will show, for example, that Cu is much more toxic to most plankton than is As. More subtle effects can also be sought in laboratory cultures of various organisms, for example, changes in metabolism, ability to reproduce, find food, grow, etc. A vast amount of literature exists on methods for detecting sublethal effects of toxins on organisms [e.g., (2,12)]. Different sublethal tests often give different results in rating the relative toxicities of different trace elements, but they have the potential for indicating possible long-term effects on organisms that might not show up in short-term acute tests.

Most laboratory toxicity tests use water as free from trace elements, complexing ligands, organic matter, etc. as is possible, so the response of the test organisms can be more clearly related to the trace element added in the test procedure. Consequently, a trace element is almost always less toxic in the environment than it is in the laboratory, because of complexing, adsorption, and other interactions in the environment.

Laboratory toxicity tests, even when they try to imitate the environment by using ambient water, multiple trace elements, varying salinities, temperatures, different life stages of organisms, etc., can never truly duplicate natural conditions. It is useful, then, to look for effects of pollutants in the environment, especially at the population level. This is, however, a difficult task, because of the natural temporal and spatial variability in abundance and health of organisms.

As a result of the relative expense and time involved in toxicity tests, and their sometimes ambiguous results, many environmental assessment programs seek only to determine concentrations of trace elements in the environment and to look for enrichments caused by human activity. If a trace element enrichment is detected, its significance can then be resolved by toxicity testing or detailed ecological field analysis. In any case, trace element enrichments could be sought in air, water, sediment, or organisms.

Water analysis might seem a logical way to detect trace element enrichments in the environment. Furthermore, the significance of trace element concentrations in water can be judged because the US EPA has published values for each element above which harm to organisms is likely (Table 1). However, ambient water, be it river water, groundwater, rainwater, or seawater, is notoriously hard to collect, store, and analyze for trace element content, as was discussed above. Water concentrations can also change over short time scales in some circumstances, which further complicates their use.

Soil or sediment can usually be more easily analyzed accurately for trace element content than can water. Soil and sediment also integrate trace element input over some time scale, so they don't need to be sampled as often as does water. Another advantage of sediment analysis is that it gives a historical view of pollutant input at sites where sediment is laid down layer by layer, year after year. Dates can be assigned to the different layers by use of radio-isotopes, pollen identification, or other means. Furthermore, sediment layers from prehistoric times give a background value for each trace element that can be compared with values in near-surface layers in order to quantify human-induced enrichments.

Recognizing gross sediment contamination is easy. Any sediment that is several-fold higher in a given trace element than the average crustal abundance of that element is contaminated unless some unusual mineralogy exists. However, it is harder to recognize subtle contamination because of difficulties in establishing an exact background concentration for a given location. Values from prehistoric depths in the sediment column are a possible background, as noted above. Another background is sediment well away from any point source of pollutant input. In using either of these methods, care should be taken to compare similar sediment types or to compare element to element concentration ratios rather than absolute concentrations (13).

Another problem with using sediment data is that only some unknown fraction of the trace element in the sediment is likely to be available to organisms, which has been much discussed in the literature [e.g., (14)], especially in conjunction with disposal of dredge spoil [e.g., (15)]. Many authors have suggested leaching sediments with dilute acids or other solutions [e.g., (14)] in order to remove only the trace element that could potentially be removed by an organism. Another suggestion that has been much debated is the ability of sulfide in the sediment to limit availability of trace elements to organisms [e.g., (16)]. Thus, although it generally agreed that only a fraction of the total trace element in sediments is available to organisms, no consensus exists on how to measure that fraction.

Long and Morgan [(2) and elsewhere] suggest another way to identify sediment that is potentially harmful to organisms because of chemical contamination. They compiled published matching biological health and chemical data from numerous field, laboratory, and modeling studies. The data was then ranked from the lowest to the highest contaminant concentration where any adverse biological effect was reported. From the ranking they derived two guideline concentrations for each contaminant. These two values separate the data into values that (1) rarely, (2) occasionally, or (3) frequently cause adverse biological effects. These derived values have been widely used in monitoring programs. See Table 1 for some of the actual values.

If both water and sediment offer analytical and data interpretation challenges, would it not be better to analyze organisms in order to assess trace element contamination? Certainly, advantages to this approach exist. For one thing, there is no question as to whether the element is available to organisms. For another, concentrations are often high enough to make analyses relatively easy, at least for common trace elements such as Cu and Zn. There are, however, problems, for example, deciding what organisms to analyze. It is not practical to analyze every organism at a given location, or even to analyze a representative specie from each major taxonomic group. What, then, should be analyzed?

Farrington (17) summarized the rationale for using common mussels (*Mytilus* sp.), various oyster species (*Crassostrea* and *Ostrea*), and other bivalves as "sentinel" organisms in monitoring studies in the marine environment. This approach has resulted in a very large worldwide data set for trace elements in bivalves. In the United States, the National Oceanic and Atmospheric Administration's "National Status and Trends Program" (NS&T) has been analyzing bivalves from the entire U.S. coastline since 1986 and has produced an especially useful and high-quality data set. As a result of the NS&T and similar data, bivalves should be the first choice for organisms to analyze in marine environmental monitoring programs. Many different kinds of plants and animals have been used in nonmarine environmental monitoring studies, everything from plankton and moss to polar bears. It all depends on what is available and the purpose of the monitoring. In general, organisms that have low natural variability in trace element concentration and are geographically widespread and easy to collect should be selected.

Data from good, long-term environmental monitoring programs can help answer the question "are things getting better or worse." Since strict environmental laws took effect in the United States in the 1970s, billions of dollars have been spent on pollution-control devices and cleanup of polluted sites. Have the efforts worked? (O'Connor (18,19) looked for temporal trends in the NS&T data discussed earlier. He found that for 2744 combinations of 14 chemicals and 196 collection sites over a 10-year time period, only 88 increases and 348 decreases in concentration are significant at the 95% confidence level. Chance alone predicts 69 increases and 69 decreases, so by this analysis, it is quite possible that no real increases occurred over that 10-year period and environmental quality along the U.S. coastline may have improved. This finding is consistent with observations of other environmental scientists. At least for chemical contaminants, environmental laws have worked and the U.S. environment is cleaner now than it was in the 1970s.

BIBLIOGRAPHY

1. Taylor, S.R. and McClennon, S.M. (1985). *The Continental Crust: Its Composition and Evolution*. Blackwell Scientific Publications, Oxford, p. 312.
2. Long, E.R. and Morgan, L.G. (1990). *The Potential for Biological Effects of Sediment-sorbed Contaminants Tested in the National Status and Trends Program*. NOAA Tech. Memo. NOS OMA 52, Rockville, MD, p. 167.
3. Bruland, K.W. (1983). Trace elements in sea-water. In: *Chemical Oceanography*. Vol. 8, J.P. Riley and R. Chester (Eds.). Academic Press, London, pp. 158–220.
4. Shiller, A.M. (1997). Dissolved trace elements in the Mississippi River: Seasonal Interannual and decadal variability. *Geochemica et Cosmochemica Acta* **61**: 4321–4330.
5. U.S. Environmental Protection Agency. (2002). *National Recommended Water Quality Criteria: 2002*. EPA-822-R-02-047, US EPA Office of Water, Washington, DC.
6. Presley, B.J. (1990). Trace metals in Gulf of Mexico oysters. *The Science of the total Environment* **97/98**: 551–593.
7. Furness, R.W. and Rainbow, P.S. (1990). *Heavy Metals in the Marine Environment*. Lewis Publishers, p. 264.
8. Mertz, W. (1981). The essential trace elements. *Science, V.* **213**: 1332–1338.
9. Bryan, G.W. (1976). Factors influenceing toxicity of heavy metals. In: *Marine Pollution*. R. Johnson (Ed.). Academic Press, London, pp. 185–302.
10. Luoma, S.N. (1983). Bioavailability of trace metals to aquatic organisms-A review. *Science of the Total Environment* **28**: 1–29
11. U.S. Environmental Protection Agency. (1987). *Biomonitoring to Achieve Control of Toxic Effluents*. EPA/625/8-87/013, US EPA Office of Water, Washington, DC.
12. U.S. Environmental Protection Agency. (1988). *Short-Term Methods For Estimating Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms*. Environmental Protection Agency, EPA 60014-87/028 Cincinnati, OH.
13. Trefry, J.H. and Presley, B.J. (1976). Heavy Metals in Sediments from San Antonio Bay and the Northwest Gulf of Mexico. *Environmental Geology* **1**: 283–294.
14. Campbell, P.G.C. (1988). *Biologically Available Metals in Sediments*. National Research Council of Canada, ISSN 0316-0114, Ottawa, Canada, p. 298.
15. Lake, J., Hoffman, G.L., and Schimmel, S.C. (1985). *Bioaccumulation of Contaminants from Black Rock Harbor Dredged Material by Mussels and Polychaetes*. U.S. Army, Corp of Engineers, Tech. Report D-85-2, Vicksburg, MS.
16. Allen, H.E., Fu, G., and Deng, B. (1993). Analysis of acid-volatile sulfide (AVS) and simultaneously extracted metals (SEM) for estimation of potential toxicity in aquatic sediments. Environ. *Toxicol. and Chem.* **12**: 1441–1453.
17. Farrington, J.W. (1983). Bivalves as sentinels of coastal chemical pollution; The mussel (oyster) watch. *Oceanus* **26**: 18–29.
18. O'Connor, T.P. (1998). Mussel Watch results from 1986 to 1996. *Mar. Poll. Bull.* **37**: 14–19
19. O'Connor, T.P. (1990). *Recent Trends in Environmental Quality: Results of the First Five Years of the NOAA Mussel Watch Project*. NOAA/ORCA, Silver Spring, MD.

CORAL REEFS AND YOUR COASTAL WATERSHED

U.S. Environmental Protection
Agency—Oceans and Coastal
Protection Division

Coral reefs are among the world's richest ecosystems, second only to tropical rain forests in plant and animal diversity. However, they are extremely sensitive environments that have special temperature, salinity, light, oxygen, and nutrient requirements. If environmental conditions fall outside the acceptable range of these requirements, the health and dynamics of a coral reef community can be severely disrupted. That's why coral communities are sensitive indicators of water quality

---

and the ecological health of the coastal watershed. They respond to alterations within the entire coastal watershed, such as changes in freshwater flows and nutrient inputs. Consequently, pollution from the destruction and alteration of surrounding coastal watersheds can directly affect the health and productivity of a coral reef.

WHAT ARE CORAL REEFS AND WHAT DO THEY DO?

Coral reef ecosystems are unique, biologically diverse systems recognized as valuable economic and environmental resources. Many people think coral reefs are made of plants or rocks, but they are actually made of animals! A coral polyp is a delicate, limestone-secreting animal. The limestone serves as a skeleton that either is embedded within the living tissue of the coral or encloses the animal. A coral reef is made up of colonies of these coral polyps. There are several benefits of coral reefs.

- Coral reefs are an important recreational and aesthetic resource for people visiting or living in coastal areas. People use coral reefs for fishing, underwater photography, scuba diving, and snorkeling.
- Coral reefs provide protection for harbors and beaches, which are often found behind reefs because the reefs provide natural protection from heavy wave action caused by coastal storms.
- Coral reefs are home to a number of species of fish and other marine species, including many that we rely on for food and economic purposes.
- Coral reefs also serve as a laboratory for students and scientists to study and learn about complex ecological and biological processes. In addition, the reefs yield many biological treasures that are increasingly being recognized as natural sources of biomedical chemicals.

SOME IMPACTS ON CORAL REEFS

Coral reef habitats are extremely sensitive to disturbances, such as various forms of pollution and physical contact. Pollution of coastal watersheds poses a threat to the existence of coral reefs. Impacts can result from activities occurring near the reef itself or from areas within the coastal watershed that drain to the reef. Disturbances and pollution can lead to diseases in coral such as bleaching (when the algae that give corals their color die). Natural occurrences, such as hurricanes, can adversely impact coral reefs through high-energy storm surges and the resulting resuspension of sediment. However, reefs are usually able to recover from natural disturbances.

People using the reef can have an adverse impact on reef resources. Portions of a coral reef can be broken by the impact of boat anchors and boat groundings. Divers and snorkelers can harm the reef by simply touching it or by removing the corals. Suntan oil from swimmers and snorkelers can harm or even kill sensitive corals. Dragging hooks, fishing line, and nets across the coral reef, as well as placing and recovering lobster traps on reefs, can be damaging. Overfishing also harms coral reefs by removing important species that eat the algae growing on corals. When these fish species are removed, the algae overgrow the corals, smothering them.

Marine debris, trash floating on the ocean or resting on the ocean floor, comes from many sources, including boaters, divers, improper disposal of trash on land, storm water runoff to rivers and streams, ships and other vessels, and offshore oil platforms. Marine debris can harm fish species and other aquatic organisms that use the reef. Trash that lands on the reef can kill corals by continually rubbing against it or smothering it.

An excessive amount of *nutrients* from improperly treated sewage, atmospheric deposition, agricultural and urban runoff, and cleaning products high in phosphates can harm coral reef habitats. In excess levels, nutrients overstimulate the growth of aquatic plants and algae. When nutrient levels increase, the delicate balance that exists between corals and algae is destroyed and the algae can overgrow the corals. When this situation is prolonged, the corals are smothered and die beneath the algal carpet. This, in turn, affects the fish and other aquatic organisms using the area, leading to a decrease in animal and plant diversity and affecting use of the water for fishing and swimming.

Some of the leading causes of nearshore coral decline can be related to *land development* and nearshore

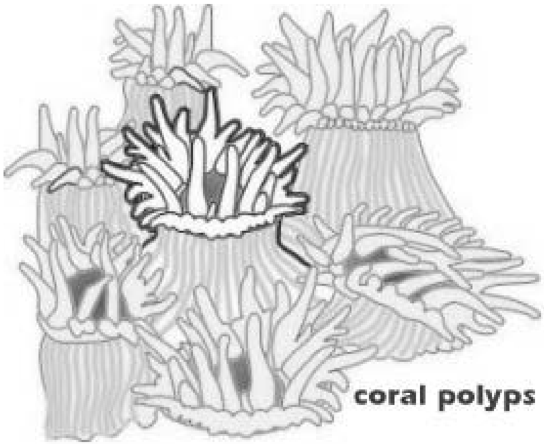

construction that are not environmentally sensitive. Sediment, silt, and other suspended solids wash off plowed fields, construction and logging sites, urban areas, strip-mined land, and eroded stream banks when it rains. Increases in coastal sediment are also caused by construction of seawalls, docks, and marinas; land-clearing; boats running through shallow waters, disturbing and suspending silts with their propellers; and snorkelers and divers kicking up sediment. Sediment can block sunlight that is essential for the survival of some corals, which live in a very close relationship with microscopic plants (algae) that require sunlight to survive. In addition, heavy sedimentation can bury corals, inhibiting their growth or killing them.

Pathogens are disease-causing microorganisms such as viruses, bacteria, and parasites. Pathogens are harmful to corals, causing disease and scarring in many species. These microorganisms enter water bodies from sources such as: inadequately treated sewage, storm water drains, septic systems, runoff from livestock pens, and boats that discharge sewage.

Coral reefs are vulnerable to the introduction of a wide variety of *toxic substances*, including metals (such as mercury and lead), toxic organic chemicals (such as PCBs and dioxin), pesticides, and herbicides found in sources like industrial discharges, urban and agricultural runoff, mining activities, and runoff from landfills. Some toxic substances bind to sediment and are transported to coastal waters through sedimentation. These toxic substances can cause scarring, death, or reproductive failure in fish, shellfish, and other marine organisms. In addition, they can accumulate in fish tissue, leading to fish consumption advisories. The sensitivity of corals makes them especially vulnerable to the introduction of toxic substances.

WHAT IS EPA DOING TO PROTECT CORAL REEFS?

International Coral Reef Initiative

In 1994, EPA, along with the State Department, the National Oceanic and Atmospheric Administration, and the Department of the Interior, formed an international coalition to coordinate information and bring higher visibility to the need for coral reef ecosystem preservation. The coalition became the *International Coral Reef Initiative* (ICRI), which now includes a membership of more than 90 countries.

EPA's Watershed Approach. EPA has joined with others to promote the Watershed Approach nationally as a means to further restore and maintain the physical, chemical, and biological quality of our nation's waters, including coral reefs. By addressing issues on a watershed scale, those areas that pose the greatest risk to human and ecological health can be targeted, several pollutants can be addressed at one time, the public can be involved in cleaning up the environment and protecting coral habitats, and integrated solutions for environmental protection can be considered. This is particularly important given the contribution of activities and sources of pollution within the larger watershed to the decline of coral reefs. Through the Watershed Approach, integrated coastal zone management tools and watershed concepts can be applied in the development of comprehensive management and conservation plans. The Watershed Approach aims

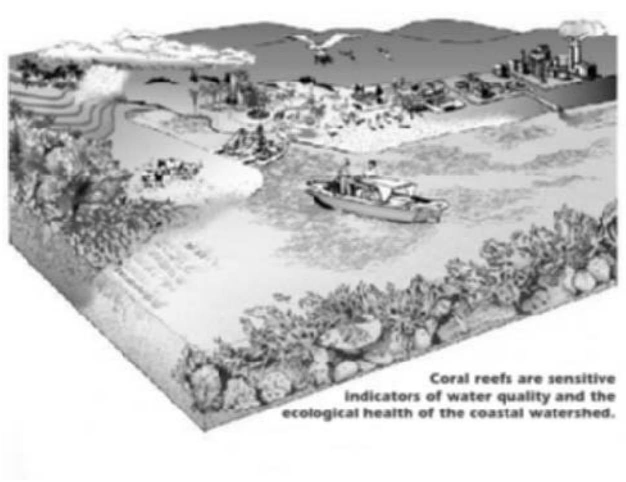

Coral reefs are sensitive indicators of water quality and the ecological health of the coastal watershed.

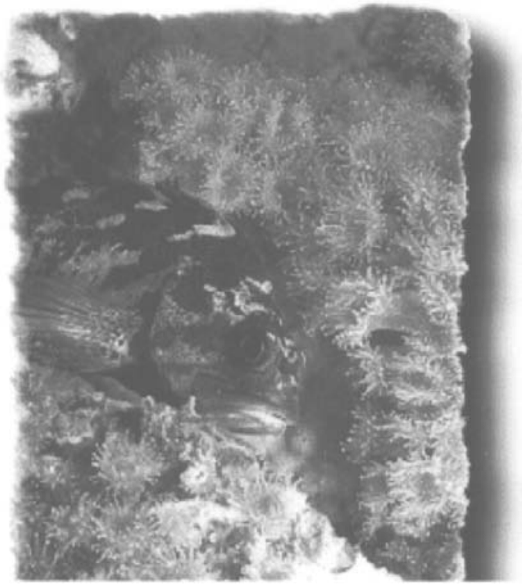

to determine protective approaches for controlling identified stressors to coral reef ecosystems. EPA's *Coastal Watershed Protection Strategy* specifically provides technical assistance and support to priority coastal watersheds, such as *National Estuary Programs* (NEPs) and other coastal waters identified by states.

Other EPA Programs. In assisting coastal states with the development of their *Coastal Nonpoint Pollution Control Programs*, EPA and other federal agencies developed guidance specifying management measures for sources of nonpoint pollution (diffuse runoff of pollutants) in coastal waters. In its program, a state or territory describes how it will implement nonpoint source pollution controls. EPA also works with other federal agencies to protect human health and coral reefs by reducing marine debris. The efforts include the establishment of the *National Marine Debris Monitoring Program*, which looks at the origins and amounts of marine debris deposited along U.S. coasts. EPA and the Coast Guard work together to regulate the transportation of municipal and commercial waste on vessels and to issue regulations for the manufacture, maintenance, and efficiency of marine sanitation devices (boat toilets), as well as the establishment of "no discharge zones" for vessel sewage. EPA also regulates the discharge of pollutants from facilities into sensitive marine waters. EPA assists states in the development of *water quality standards* designed to protect human health and aquatic life. This assistance includes the development of criteria for water quality that accurately reflects the most up-to-date scientific knowledge about the effects of pollutants on aquatic life, such as corals, and human health.

What Can You Do to Help Protect Coral Reefs?

You can do several things to help protect coral reefs and your coastal watershed:

- *Be Informed and Involved*. Learn about coral reefs and their importance to your coastal watershed. Participate in training or educational programs that focus on reef ecology. Be an informed consumer; ask the store owner or manager from what country the coral was taken and whether that country has a management plan to ensure that the harvest was legal and sustainable over time. Support the creation and maintenance of marine parks and reserves. Become a citizen volunteer. As a volunteer you might be involved in taking water quality measurements, tracking the progress of protection and restoration projects, or reporting special events like fish kills and storm damage. Volunteer for a reef cleanup or a beach cleanup. If you don't live near a coast, get involved in your local watershed program. Report dumping or other illegal activities.
- *Take Responsibility for Your Own Backyard*. Determine whether additional nutrients or pesticides are needed before you apply them, and look for alternatives to fertilizers and pesticides where the chance of runoff into surface waters might occur. Even if you live far from a coral reef ecosystem, these products might ultimately affect the waters that support coral. Consider selecting plants and grasses with low maintenance requirements. Water your lawn conservatively; the less water you use, the less runoff will eventually find its way into the oceans.
- *Practice Good Housekeeping*. Learn about procedures for disposing of harmful household wastes so they do not end up in sewage treatment plants that can't treat them or in landfills not designed to receive hazardous materials. Around the house, keep litter, pet waste, leaves, and grass clippings out of street gutters and storm drains to prevent their entrance into streams that might flow to reefs. Use the minimum amount of water needed when you wash your car to prevent waste and runoff. Never dump any household, automotive, or gardening wastes into a storm drain. They might end up on the reef. Take used motor oil, paints, and other hazardous household

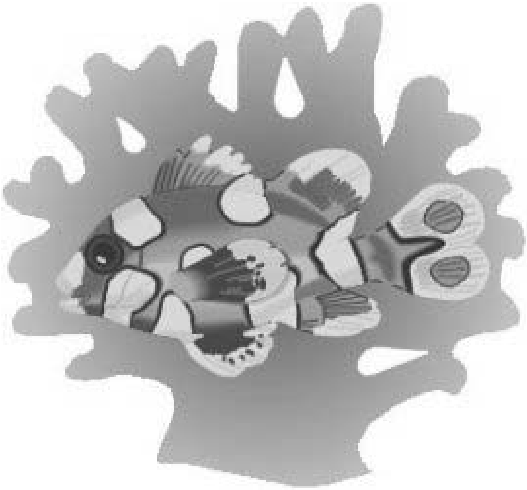

materials to proper collection sites such as approved service stations or designated landfills. Always follow label directions for the use and disposal of household chemicals. Keep your septic tank in good working order. The improper disposal of wastes and hazardous materials can lead to water quality problems and harm to the sensitive coral reef habitats.

- *Respect the Reef*. Help keep the reef healthy by following local guidelines, recommendations, regulations, and customs. If you dive, don't touch the coral. Keep your fins, gear, and hands away from the coral since this contact can hurt you and will damage the delicate coral animals. Stay off the bottom because stirred-up sediment can settle on corals and smother them. Avoid entering sensitive habitat areas with your boat or other motorized watercraft. Maintain your boat engine to prevent oil and gas leaks. Keep all waste produced during your excursions in a safe place to be disposed of properly when you're back on land. If you go boating near a coral reef, don't anchor your boat on the reef. Use mooring buoy systems if they are available. Maintain and use your marine sanitation devices properly. Conserve energy and keep your auto in good running condition. By conserving energy, harmful air emissions leading to air deposition are minimized.

SEA LEVEL AND CLIMATE

Richard Z. Poore
Christopher Tracey
U.S. Geological Survey
Reston, Virginia

Richard S. Williams, Jr.
U.S. Geological Survey
Woods Hole, Massachusetts

INTRODUCTION

Global sea level and the Earth's climate are closely linked. The Earth's climate has warmed about 1 °C (1.8 °F) during the last 100 years. As the climate has warmed following the end of a recent cold period known as the "Little Ice Age" in the 19th century, sea level has been rising about 1 to 2 millimeters per year due to the reduction in volume of ice caps, ice fields, and mountain glaciers in addition to the thermal expansion of ocean water. If present trends continue, including an increase in global temperatures caused by increased greenhouse-gas emissions, many of the world's mountain glaciers will disappear. For example, at the current rate of melting, all glaciers will be gone from Glacier National Park, Montana, by the middle of the next century (Fig. 1). In Iceland, about 11 percent of the island is covered by glaciers (mostly ice caps). If warming continues, Iceland's glaciers will decrease by 40 percent by 2100 and virtually disappear by 2200.

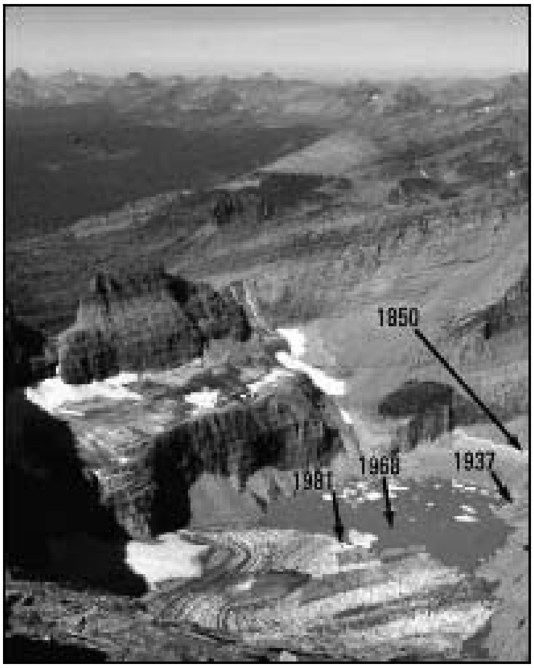

Figure 1. Grinnell Glacier in Glacier National Park, Montana; photograph by Carl H. Key, USGS, in 1981. The glacier has been retreating rapidly since the early 1900's. The arrows point to the former extent of the glacier in 1850, 1937, and 1968. Mountain glaciers are excellent monitors of climate change; the worldwide shrinkage of mountain glaciers is thought to be caused by a combination of a temperature increase from the Little Ice Age, which ended in the latter half of the 19th century, and increased greenhouse-gas emissions.

Most of the current global land ice mass is located in the Antarctic and Greenland ice sheets (Table 1). Complete melting of these ice sheets could lead to a sea-level rise of about 80 meters, whereas melting of all other glaciers could lead to a sea-level rise of only one-half meter.

GLACIAL-INTERGLACIAL CYCLES

Climate-related sea-level changes of the last century are very minor compared with the large changes in sea level that occur as climate oscillates between the cold and warm intervals that are part of the Earth's natural cycle of long-term climate change.

Table 1. Estimated Potential Maximum Sea-Level Rise from the Total Melting of Present-Day Glaciers. [Modified from Williams and Hall (1993). See also http://pubs.usgs.gov/factsheet/fs50-98/]

Location	Volume (km^3)	Potential Sea-Level Rise (m)
East Antarctic ice sheet	26,039,200	64.80
West Antarctic ice sheet	3,262,000	8.06
Antarctic Peninsula	227,100	.46
Greeland	2,620,000	6.55
All other ice caps, ice fields, and valley glaciers	180,000	.45
Total	32,328,300	80.32

During cold-climate intervals, known as glacial epochs or ice ages, sea level falls because of a shift in the global hydrologic cycle: water is evaporated from the oceans and stored on the continents as large ice sheets and expanded ice caps, ice fields, and mountain glaciers. Global sea level was about 125 meters below today's sea level at the last glacial maximum about 20,000 years ago (Fairbanks, 1989). As the climate warmed, sea level rose because the melting North American, Eurasian, South American, Greenland, and Antarctic ice sheets returned their stored water to the world's oceans. During the warmest intervals, called interglacial epochs, sea level is at its highest. Today we are living in the most recent interglacial, an interval that started about 10,000 years ago and is called the Holocene Epoch by geologists.

Sea levels during several previous interglacials were about 3 to as much as 20 meters higher than current sea level. The evidence comes from two different but complementary types of studies. One line of evidence is provided by old shoreline features (Fig. 2). Wave-cut terraces and beach deposits from regions as separate as the Caribbean and the North Slope of Alaska suggest higher sea levels during past interglacial times. A second line of evidence comes from sediments cored from below the existing Greenland and West Antarctic ice sheets. The fossils and chemical signals in the sediment cores indicate that both major ice sheets were greatly reduced from their current size or even completely melted one or more times in the recent geologic past. The precise timing and details of past sea-level history are still being debated, but there is clear evidence for past sea levels significantly higher than current sea level.

POTENTIAL SEA-LEVEL CHANGES

If Earth's climate continues to warm, then the volume of present-day ice sheets will decrease. Melting of the current Greenland ice sheet would result in a sea-level rise of about 6.5 meters; melting of the West Antarctic ice sheet would result in a sea-level rise of about 8 meters (Table 1). The West Antarctic ice sheet is especially vulnerable, because much of it is grounded below sea level. Small changes in global sea level or a rise in ocean temperatures could cause a breakup of the two buttressing ice shelves (Ronne/Filchner and Ross). The resulting surge of the West Antarctic ice sheet would lead to a rapid rise in global sea level.

Figure 2. Wave-cut terraces on San Clemente Island, California. Nearly horizontal surfaces, separated by step-like cliffs, were created during former intervals of high sea level; the highest terrace represents the oldest sea-level high stand. Because San Clemente Island is slowly rising, terraces cut during an interglacial continue to rise with the island during the following glacial interval. When sea level rises during the next interglacial, a new wave-cut terrace is eroded below the previous interglacial terrace. Geologists can calculate the height of the former high sea levels by knowing the tectonic uplift rate of the island. Photograph by Dan Muhs, USGS.

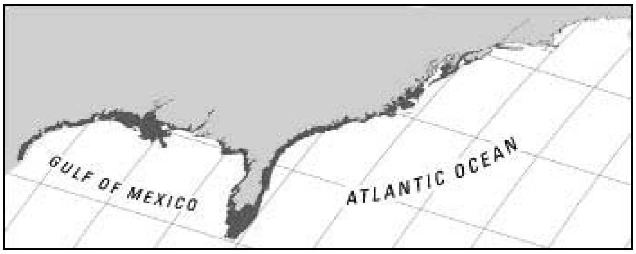

Figure 3. Red shows areas along the Gulf Coast and East Coast of the United States that would be flooded by a 10-meter rise in sea level. Population figures for 1996 (U.S. Bureau of the Census, unpublished data, 1998) indicate that a 10-meter rise in sea level would flood approximately 25 percent of the Nation's population.

Reduction of the West Antarctic and Greenland ice sheets similar to past reductions would cause sea level to rise 10 or more meters. A sea-level rise of 10 meters would flood about 25 percent of the U.S. population, with the major impact being mostly on the people and infrastructures in the Gulf and East Coast States (Fig. 3).

Researchers at the U.S. Geological Survey and elsewhere are investigating the magnitude and timing of sea-level changes during previous interglacial intervals. Better documentation and understanding of these past changes will improve our ability to estimate the potential for future large-scale changes in sea level.

READING LIST

Fairbanks, R.G. (1989). A 17,000-year glacio-eustatic sea level record; influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature* **342**(6250): 637–642.

Williams, R.S., and Hall, D.K. (1993). Glaciers, in Chapter on the cryo-sphere, in Gurney, R.J., Foster, J.L., and Parkinson, C.L., eds., *Atlas of Earth Observations Related to Global Change*. Cambridge University Press, Cambridge, UK, pp. 401–422.

THE PERMANENT SERVICE FOR MEAN SEA LEVEL

S. Jevrejeva
S. Holgate
P.L. Woodworth
Proudman Oceanographic Laboratory
Birkenhead, United Kingdom

Mean sea level (MSL) is the average level of the sea, relative to the level of the land on which the measurements are being made, recorded over an extended period such as a month, year, or the lunar nodal period of 18.6 years by an instrument called a tide gauge (or coastal sea level recorder). MSL data are used in a wide range of scientific applications including studies into climate change, ocean circulation variability and geology, as well as in practical applications such as surveying and the establishment of national leveling datums.

The Permanent Service for Mean Sea Level (PSMSL) is the global data bank for such MSL information, and it has since 1933 been responsible for the collection, publication, analysis, and interpretation of sea level data from the global network of tide gauges. It is based at the Proudman Oceanographic Laboratory (POL) in Liverpool, U.K. and is a member of the Federation of Astronomical and Geophysical Data Analysis Services (FAGS) established by the International Council for Science (ICSU). The PSMSL is supported by FAGS, the Intergovernmental Oceanographic Commission of the United Nations Educational, Scientific and Cultural Organisation (IOC/UNESCO), and NERC.

The database of the PSMSL contains almost 53,000 station-years of monthly and annual mean values of sea level from nearly 2000 tide gauge stations around the world received from almost 200 national authorities (see Fig. 1). On average, approximately 2000 station-years of data are entered into the database each year. All data are readily available from the PSMSL website: www.pol.ac.uk/psmsl.

Data for all stations are included in the PSMSL METRIC (or total) dataset. The METRIC monthly and annual means for any one station-year are necessarily required to be measured to a common datum, although, at this stage, datum continuity between years is not essential. The year-to-year datum checks become essential, however, if the data are subsequently to be included in the PSMSL 'Revised Local Reference (RLR)' component of the dataset. The RLR dataset contains records for which time series analysis of sea level changes can be performed. Long records from this dataset have been the basis of all analyses of secular changes in global sea level during the last century. The geographical distribution of longer RLR records contains significant geographical bias toward the Northern Hemisphere, a situation that is being rectified by means of international collaboration.

Aside from its central role of operation of the global sea level data bank, the PSMSL has a responsibility as a member of FAGS to provide the sea level community with as full a Service as possible with regard to the acquisition, analysis, and interpretation of sea level data. Consequently, the PSMSL provides a range of advice to tide gauge operators and data analysts. It has occupied a central planning and management role in the development of the Global Sea Level Observing System (GLOSS) of the IOC. Through GLOSS and via other routes, the PSMSL provides advice and training to national sea level authorities and individual sea level scientists and technologists. In addition to the provision of training materials (e.g., tide gauge operation manuals), the PSMSL supplies software packages suitable for tidal data analysis and quality control purposes.

In addition to the training courses associated with GLOSS, the PSMSL has every few years hosted important study groups and international conferences on sea level science. The study groups have concerned themselves with topics such as the use of global positioning system (GPS) receivers at tide gauge sites to determine the local rates of vertical land movement and have been held under the auspices of the International Association for the Physical Sciences of the Ocean (IAPSO) Commission on Mean Sea Level and Tides (CMSLT), the scientific body to which

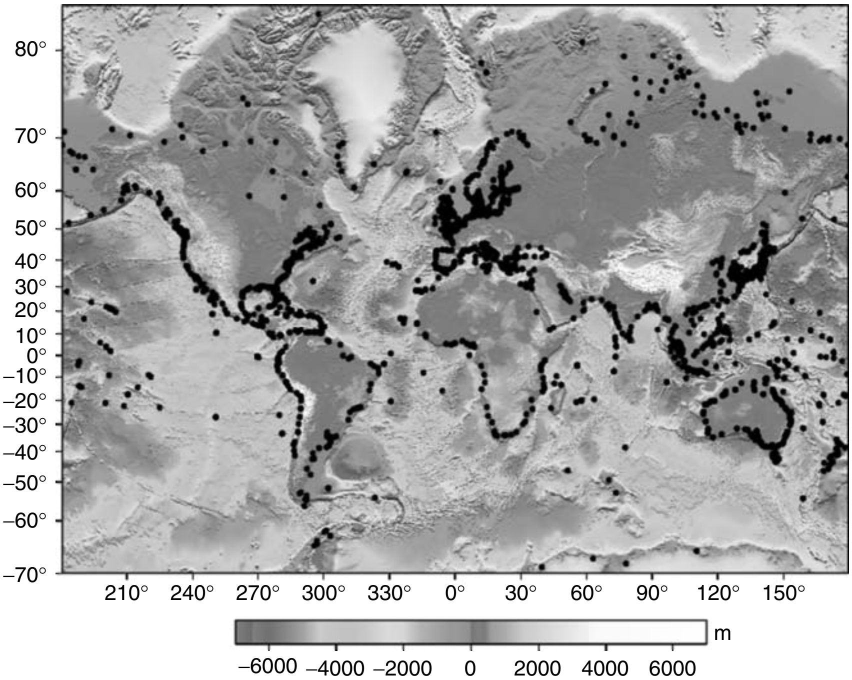

Figure 1. Distribution of PSMSL stations.

the PSMSL reports formally. The PSMSL hosted a major Symposium in Vienna in 1991 as part of the International Union of Geodesy and Geophysics (IUGG) Congress, an international conference at the Linnean Society in London in 1993 as part of its 60th anniversary celebrations, co-organized "Tidal Science 96" at the Royal Society in London in 1996, and took a major part in "A Celebration of UK Sea Level Science" at the Royal Society in 2004. The proceedings of each of these conferences have since been published. A further major conference is planned in 2008 for the PSMSL 75th anniversary.

Probably the most important recent scientific publications with which the PSMSL is associated are those of the First (1990), Second (1995), and Third (2000) Scientific Assessments of Intergovernmental Panel on Climate Change (IPCC). The PSMSL Director has been a lead author for the sea level chapters in each of the IPCC studies. Major conclusions have been that global sea level has indeed risen by approximately 10–20 cm during the past century and may rise by amounts several times larger during the next 100 years.

The PSMSL is conscious that developments in technology have expanded the field of sea and land level studies. During the 1990s, satellite radar altimetry and GPS recording become established techniques, whereas space gravity offers the potential for being an effective source of sea-level-related information in the future. Therefore, the PSMSL maintains full participation with altimeter and space gravity working groups in view of the importance of those techniques to sea level research. PSMSL personnel have Principal Investigator status for the TOPEX/Poseidon, Jason, ERS, and Envisat altimeter missions, in addition to the GRACE and GOCE space gravity projects. The major challenge for the future, to which the PSMSL is committed, is to see the established tide gauge and new space-based techniques closely linked within one coherent global sea level monitoring system.

READING LIST

Woodworth, P.L. and Player, R. (2003). The Permanent Service for Mean Sea Level: an update to the 21st century. *J. Coastal Res.* **19**: 287–295.

Woodworth, P.L., Aarup, T., Merrifield, M., Mitchum, G.T., and Le Provost, C. (2003). Measuring progress of the Global Sea Level Observing System. *EOS, Trans. Am. Geophys. Union*. **84**(50): 565.

MARINE AND ESTUARINE MICROALGAL SEDIMENT TOXICITY TESTS

Ignatio Moreno-Garrido
Luís M. Lubián
Julián Blasco
Institute of Marine Sciences of Analucia
Cadiz, Spain

The term microphytobentos refers to the microscopic algae that live on the submerged (temporary or not) floor in fresh water, estuarine, or marine environments. Microphytobenthos are mainly composed of mobile, pennate diatoms and cyanophytes. In marine and estuarine environments, those organisms can be found in habitats such as salt marshes, submerged vegetation beds, intertidal (sand or mud) flats, or subtidal sediments where light permits microalgal growth. As the presence of these photosynthetic organisms is not always evident, MacIntyre et al. (1) borrowed the term "secret garden" from the homonymous book published in 1888 and written by Frances Hodgson Burnett (1849–1874), in order to make a literary allusion to microphytobenthos. In fact, habitats where microphytobentos are the only primary producers are recognized as "unvegetated" areas, but the concentration of chlorophyll a in the upper 0.5 cm of the sediments where those organisms live generally exceed the depth-integrated chlorophyll in the entire overlying water column (2). In some cases, chlorophyll from microphytobenthos can be up to six orders of magnitude higher than that for the overlying water (3). Nevertheless, some authors estimate that primary production in the bottom would be lower than in the free plankton in spite of that exceptional data (4). But in some habitats, biomass from benthic microalgae can match or even exceed biomass of bacteria present in the same space (5). Thus, microphytobenthos necessarily play a key role in the benthic trophic webs (6–8). In certain biocenosis, the organisms of the microphytobenthos are the main—sometimes the only—source of carbon for grazers or bacteria (7,9,10). The microphytobenthos are also very important in relation to the stability of coastal and estuarine sediments. Although some cyanobacteria can show hydrophobicity as a mechanism to attach sediment particles or other cells (11), the main strategy of microphytobenthos to keep attached to the sediment is the production of agglutinant molecules [carbohydrates (CH) or exopolysaccharides (EPS)] (12–15). The presence of these molecules has important trophic implications, but their role in the maintenance of the structure of the upper part of the sediment is also important (16–19), because the film of adherent substances produced by microphytobenthos increases the sediment stability. Toxicants such as herbicides can alter microphytobenthic growth and the production of CH and EPS can be diminished. A loss of sediment stability can induce an increase in the turbidity of the water and a higher rate of deposition of fine particles on submerged higher plants or macroalgae, thus reducing even more the primary production of the whole and adjacent systems (20).

The importance of sediments in accumulation of xenobiotics in coastal and estuarine environments has been pointed out. Most chemical contaminants entering marine or estuarine environments eventually accumulate in sediments due to different reasons, including higher salinity values that diminish solubility of such substances in water. Sediment can act as a sink for these substances but also as a subsequent source for the same (25). On the other hand, and except in cases of extreme contamination, chemical data by themselves do not predict hazard (21,22). Thus, bioassays are needed to assess the potential toxicity of sediments.

In spite of the importance of microphytobenthos, few efforts have been made to develop standard toxicity test on these organisms. Guidelines from the SETAC (23) and recent revisions (24–27) about toxicity testing on benthic organisms offer good information on macro and meiofauna but completely ignore microphytobenthos. The reason for this "exclusion" cannot be found in the lack of importance of benthic microalgae, because of all the reasons expressed above. Probably the most important reasons for the scarcity of work on microphytobenthos ecotoxicology are the difficulties that this biological material cause.

First, it is not easy to efficiently remove microphytobenthos from sediments. Size and weight of microalgae on the sediment match part of the sediment particles mixed with them. There are descriptions of techniques that use the migration capacity of the microphytobenthic organisms in order to remove them from sediment. It is well known that microphytobenthos vertically migrate through the sediment as a function of the light and tide conditions. During low tide and light conditions, cells unbury and remain at the surface of the sediment, but during high tide or night conditions, microphytobenthic cells bury themselves again, in order to avoid being removed by the current or waves or grazed during nonphotosynthetic conditions (1,28,29). This vertical movement can be exploited to make cells migrate through a plankton net separated from the sediment by one or more lens tissue papers (8). Other authors improved on this method by covering the plankton net with a few millimeters of silica powder, where living mobile cells accumulate (30). Less effective seems to be the technique that uses the adherent capacity of microphytobenthos for removing cells by the disposition of a cover glass on the sediment (19). It is supposed that cells will attach to the cover glass, but it is not easy to ensure the efficiency and repeatability of the technique. Another approach is to take cores (made of Plexiglas, PVC, or other materials) (31) from the upper sediment and resuspend subsamples of the previously sliced sediment in order to directly count (and taxonomically identify) cells by light microscopy (4). This latter method ensures the integration of all species (motile or not). The use of fluorescence microscopy can help in distinguishing photosynthetic cells from debris: with a blue filter and a barrier filter of 530 nm, chlorophyll emits a bright red fluorescence that clearly reveals cells and facilitates their localization and count, something that is difficult if this technique is not used (32,33) (Fig. 1). Other techniques use density differences to separate diatoms from debris by centrifugation in a Percoll gradient (34). Although this technique seems to be good for isolating cells, the percentage of cell recovery is low (near 5% of total population in natural locations).

Due to the difficulties of handling cells among the sediment, several works were limited to analyzing photosynthetic pigments or photosynthesis in the upper sediments (35–37) as biomarkers for the microphytobenthic biomass. The analysis of degraded pigment molecules in sediments can be useful as a biomarker for grazing pressure (38).

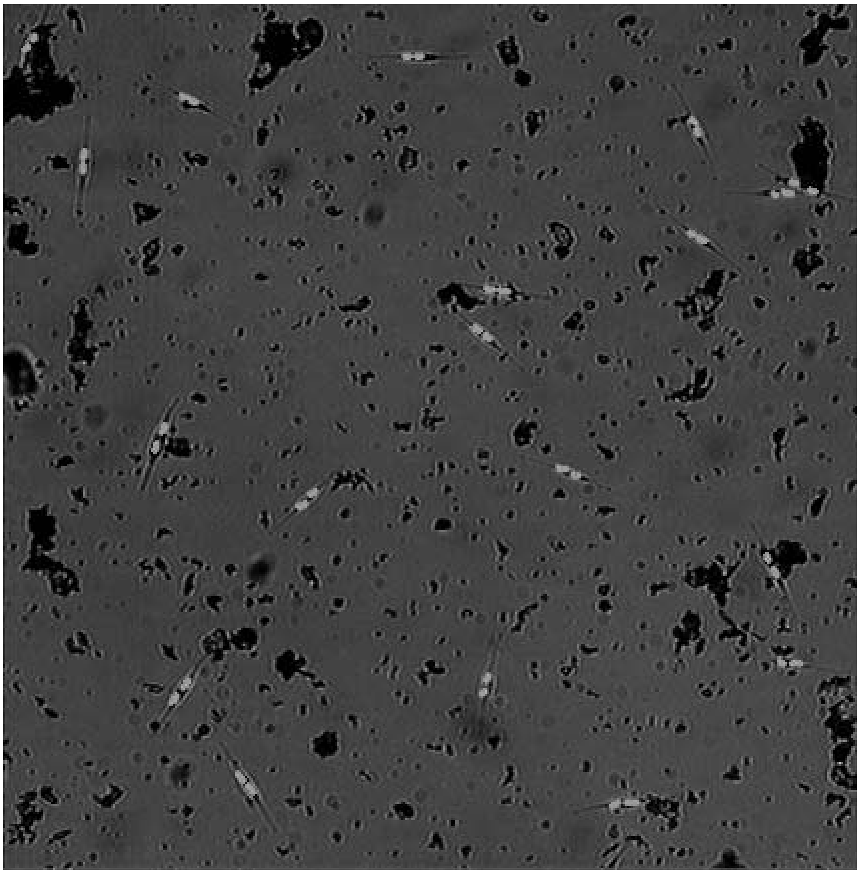

Figure 1. Cells of the diatom *Cylindrotheca closterium* among sediment particles, observed under fluorescence microscopy. Using a barrier filter of 530 nm, the two chloroplasts of each cell are bright red, facilitating the counting and location of the algal cells.

But difficulties in handling microphytobenthos do not end there. The disposition in "patches" of microphytobenthos in field locations is evident, providing a spatial heterogeneous distribution of the cells (39,40). This spatial distribution is sometimes conditioned by the presence of ridges and runnels (41) in the sediment surface, provoked by the natural dynamics of the ecosystem. This must be taken into account when estimations of microphytobenthic organism density are intended to be developed in actual locations (15). Delgado (4) described spatial heterogeneity in the delta of the Ebro River as being insignificant over distances between 10 cm and 10 m. Another consideration is that in emerged (low tide) conditions, there is a process of gradual compaction of sediment due to dewatering, which implies a higher density but a lower content of total pigments, exopolysaccharides, or individuals (42).

In spite of all this, *in situ* or *in vitro* bioassays involving microphytobenthos should be considered as powerful tools to determine potential toxicity of sediments (24). Some efforts have been made in this direction. Wong et al. (43) described a microalgal toxicity test on sediments from the coast of Hong Kong, but they used a planktonic (and not benthic) species (the chlorophyte *Dunaliella salina*) that, additionally, belongs to a genus that demonstrates strong resistance to toxicants. Abalde et al. (44) did not find growth inhibition of populations of *D. tertiolecta* at levels of 8 $\text{mg} \cdot \text{L}^{-1}$ Cu, and Moreno-Garrido et al. (45) found that *D. salina* was the most resistant species to Cu and Cd among the assayed microalgal strains. Each microalgal species shows different sensitivity for toxicants, but there are references showing similar (on the same order of magnitude) sensitivities for very different taxons (46). In this respect, *Isochrysis galbana*, *Cylindrotheca (Nitzschia) closterium*, and *Nannochloropsis species* showed similar responses to water soluble fraction of petroleum. Phytoplankton (free swimming or floating microalgae), periphyton (microalgae growing on solid substrate), and epipsammon (microalgae growing more or less attached to sand) showed comparable toxicity sensitivity responses to paraquat and simazine (47). Other works describe sediment toxicity tests on elutriates or extracts (48). Tolun et al. (49) described experiments where natural sediment toxicity tests on *Phaeodactylum tricornutum*, based on elutriates and bulk exposure, are compared: in the case of elutriates, the authors found different degrees of growth inhibition, although for direct exposure to sediments all organisms died, supporting the idea that direct exposure to sediment will give more realistic (or more sensitive) responses because part of the toxicants could not be extracted in elutriates or extracts (50). A very interesting approach to sediment toxicity tests on benthic algae was that made by Dahl and Blank (31), in which epipsammic communities were transported to a laboratory where they were kept and used in subsequent measurements of metabolic activities and short-term toxicity tests. Cairns et al. (51) defended the use of microorganisms in toxicity tests, because they can show very high sensitivities to toxicants and thus should be included in regulatory-proposed guidelines. But incorporation in those guidelines of multispecies tests that could be more "environmentally realistic" are very slow, fundamentally because of methodological questions (regarding replication and reproducibility). On the other hand, predictions based on multispecies tests are no more accurate than those based on monospecific bioassays, which are cheaper and more reproducible than a multispecies bioassay (51).

As far as we know, the first attempts to develop a standardized, repeatable protocol for sediment toxicity testing involving a microphytobenthic strain and direct exposure of microalgal cells to sediment have only recently taken place (32,33). In those works, populations of the benthic diatom *Cylindrotheca closterium* were exposed to sediment spiked with heavy metals or tensides. The test also considers the effect of particle size distribution on growth of the tested microalgal strain, which could mask actual responses of algae to present toxicants in other experiments such as those described by Tolun et al. (49). The test is simple, repeatable, and cheap, and it does not require special facilities other than those found in any laboratory. It is based on the 72-h algal growth inhibition test from OECD (52), adapted to sediments and marine or estuarine habitats. *Cylindrotheca closterium*, formerly known as *Nitzschia closterium*, demonstrated to be a good subject in other toxicity bioassays and there is a good pool of previous information about this species (53–58). This species is cosmopolitan for temperate coastal waters, ubiquitous, easy to handle in the laboratory, fast growing, sensitive to toxicants, and presents very low nutrient requirements. Nevertheless, other species have been assayed and compared with *C. closterium* in order to detect toxicity in natural sediments. EC_{50} values were calculated for three benthic diatoms exposed to sediment obtained from six different locations from Aveiro Lagoon (Portugal). Those values are shown in Table 1. Locations I, II, and VI were less toxic to microalgae than locations III, IV, and V, since *C. closterium* is slightly more sensitive than the other two species assayed for the majority of the samples. When a chemical analysis of the samples was performed and a similarity analysis carried out crossing toxicity values with possible substances involved (heavy metals, C, N, PCBs), it was found that some heavy metals (Sn, Zn, Hg, Cu, and Cr) had a great effect (Sn the greatest), as shown by the more than 50% of similarity between samples that showed significant growth inhibition for the benthic diatoms assayed (unpublished data).

Other recent work from Adams and Stauber (59) also describes a whole-sediment toxicity test on a benthic microalgal diatom (*Entomoneis* cf. *punctulata*). In this case, a flow cytometer is used to detect viability of living algae by the use of a fluorochrome (fluorescein diacetate,

Table 1. EC_{50} Values for Sediments from Six Locations at the Aveiro Lagoon (Portugal) and Three Benthic Diatoms

	Sample Sites					
Algal Species	I	II	III	IV	V	VI
P. tricornutum	N.I.[a]	N.I.	27	51	62	N.I
C. closterium	92	90	62	31	61	N.I.
Navicula sp.	N.I.	N.I.	64	71	65	N.I.

[a] If 100% of the sediment does not exhibit a growth inhibition value of 50%, EC_{50} is denoted as N.I. (no inhibition).

FDA). In this technique, used previously (60) to determine the toxic response in planktonic microalgal cells, FDA absorbed by living or dead cells is only hydrolyzed by nonspecific esterases inside the living cells. Hydrolyzed FDA is converted into fluorescein and can be detected by flow cytometry techniques. Although the FDA technique is less sensitive than growth inhibition measurements (60), flow cytometry opens a wide field of possibilities for microalgal toxicity testing on sediments, because several different fluorochromes can be used to measure quite different cellular parameters used as biomarkers.

BIBLIOGRAPHY

1. MacIntyre, H.L., Geider, R.J., and Miller, D.C. (1996). Microphytobenthos: the ecological role of the “secret garden” of unvegetated, shallow-water marine habitats. I. Distribution, abundance and primary production. *Estuaries* **19**(2A): 186–201.
2. Nelson, J.R. et al. (1999). Benthic microalgal biomass and irradiance at the sea floor on the continental shelf of the South Atlantic Bight: spatial and temporal variability and storm effects. *Continental Shelf Res.* **19**: 477–505.
3. Conde, D. et al. (1999). Comparison of the areal amount of chlorophyll *a* of planktonic and attached microalgae in shallow coastal lagoon. *Hydrobiologia* **408/409**: 285–291.
4. Delgado, M. (1989). Abundance and distribution of microphytobentos in the bays of Ebro delta (Spain). *Estuarine Coastal Shelf Sci.* **29**: 183–194.
5. La Rosa, T., Mirto, S., Mazzola, A., and Danovaro, R. (2001). Differential responses of benhtic microbes and meiofauna to fish-farm disturbance in coastal sediments. *Environ. Pollut.* **112**: 427–434.
6. Miller, D.C., Geider, R.J., and MacIntyre, H.L. (1996). Microphytobenthos: the ecological role of the “secret garden” of unvegetated, shallow-water marine habitats. II. Role in sediment stability and shallow-water food webs. *Estuaries* **19**(2A): 202–212.
7. Middelburg, J.J. et al. (2000). The fate of intertidal microphytobenthos carbon: an *in situ* ^{13}C labelling study. *Limnol. Oceanogr.* **45**(6): 1224–1234.
8. Absil, M.C.P. and Van Scheppingen, Y. (1996). Concentration of selected heavy metals in benthic diatoms and sediment in the Westerchelde Estuary. *Bull. Environ. Contam. Toxicol.* **56**: 1008–1015.
9. Puddu, A. et al. (2003). Bacterial uptake of DOM released from P-limited phytoplankton. *FEMS Microbiol. Ecol.* **46**(3): 257–268.
10. Dalsgaard, T. (2003). Benthic primary production and nutrient cycling in sediments with benthic microalgae and transient accumulation of macroalgae. *Limnol. Oceanogr.* **48**(6): 2138–2150.
11. Fattom, A. and Shilo, M. (1984). Hydrophobicity as an adhesion mechanism of benthic cyanobacteria. *Appl. Environ. Microbiol.* 135–143.
12. de Winder, B., Staats, N., Stal, L.J., and Paterson, D.M. (1999). Carbohydrate secretion by phototrophic communities in tidal sediments. *J. Sea Res.* **42**: 131–146.
13. Goto, N., Mitamura, O., and Terai, H. (2001). Biodegradation of photosynthetically produced extracellular organic carbon from intertidal benthic algae. *J. Exp. Mar. Biol. Ecol.* **257**: 73–86.
14. Staats, N., Stal, L.J., and Mur, L.R. (2000). Exopolysaccharide production by the epipelic diatom *Cylindrotheca closterium*: effects of nutrient conditions. *J. Exp. Mar. Biol. Ecol.* **249**: 13–27.
15. Taylor, I.S. and Paterson, D.M. (1998). Microspatial variation in carbohydrate concentrations with depth in the upper millimetres of intertidal cohesive sediments. *Estuarine Coastal Shelf Sci.* **46**: 359–370.
16. Austen, I., Andersen, T.J., and Edelvang, K. (1999). The influence of benthic diatoms and invertebrates on the erodability of an intertidal mudflat, the Danish Wadden Sea. *Estuarine Coastal Shelf Sci.* **49**: 99–111.
17. Blanchard, G.F. et al. (2000). The effect of geomorphological structures on potential biostabilisation by microphytobenthos on intertidal mudflats. *Continental Shelf Res.* **20**: 1243–1256.
18. de Brower, J.F.C., Bjelic, S., Deckere, E.M.G.T., and Stal, L.J. (2000). Interplay between biology and sedimentology in a mudflat (Biezelingse Ham, Westerchelde, The Netherlands). *Continental Shelf Res.* **20**: 1159–1177.
19. Madsen, K., Nilsson, P., and Sundbäck, K. (1993). The influence of benthic microalgae on the stability of a subtidal sediment. *J. Exp. Mar. Biol. Ecol.* **170**: 159–177.
20. Mason, C.F. et al. (2003). The role of herbicides in the erosion of salt marshes in eastern England. *Environ. Pollut.* **122**: 41–49.
21. Munawar, M. and Munawar, I.F. (1987). Phytoplankton bioassays for evaluating toxicity of *in situ* sediment contaminants. *Hydrobiologia* **149**: 87–105.
22. O’Connor, T.P. and Paul, J.F. (2000). Misfit between sediment toxicity and chemistry. *Mar. Pollut. Bull.* **40**(1): 59–64.
23. SETAC (Society of Environmental Toxicology and Chemistry—Europe). (1993). Guidance document on sediment toxicity tests and bioassays for freshwater and marine environments. In: *Workshop on Sediment Toxicity Assessment*. I.R. Hill, P. Matthiessen, and F. Heimbach (Eds.). November 1993, Renesse, The Netherlands.
24. Chapman, P.M. and Wang, F. (2001). Assessing contamination in estuaries. *Environ. Toxicol. Chem.* **20**(1): 3–22.
25. Lamberson, J.O., DeWitt, T.H., and Swartz, R.C. (1992). Assessment of sediment toxicity to marine benthos. In: *Sediment Toxicity Assessment*. G.A. Burton (Ed.). Lewis Publishers, Boca Raton, FL.
26. Traunspurger, W. and Drews, C. (1996). Toxicity analysis of freshwater and marine sediments with meio- and macrobenthic organisms: a review. *Hydrobiologia* **328**: 215–261.
27. Shin, P.K.S. and Lam, W.K.C. (2001). Development of a marine sediment pollution index. *Environ. Pollut.* **113**: 281–291.
28. Guarini, J.M., Blanchard, G.F., and Gros, P. (2000). Quantification of the microphytobenthic primary production in European intertidal mudflats—a modelling approach. *Continental Shelf Res.* **20**: 1771–1788.
29. Brunnet, C. and Lizon, F. (2003). Tidal and diel periodicities of size-fractionated phytoplankton pigment signatures at an offshore station in the southeastern English Channel. *Estuarine Coastal Shelf Sci.* **56**: 833–843.
30. Blanchard, G.F. et al. (2000). Measurement of ingestion rate of *Hydrobia ulvae* (Pennant) on intertidal epipelic microalgae: the effect of mud snail density. *J. Exp. Mar. Biol. Ecol.* **255**: 247–260.
31. Dahl, B. and Blank, H. (1996). Use of sand-living microalgal communities (epipsammon) in ecotoxicological testing. *Mar. Ecol. Prog. Ser.* **144**: 163–173.

32. Moreno-Garrido, I., Hampel, M., Lubián, L.M., and Blasco, J. (2003). Sediment toxicity tests using benthic marine microalgae *Cylindrotheca closterium* (Ehremberg) Lewin & Reimann (Bacillariophyceae). *Ecotoxicol. Environ. Safety* **54**: 290–295.
33. Moreno-Garrido, I., Hampel, M., Lubián, L.M., and Blasco, J. (2003). Marine benthic microalgae *Cylindrotheca closterium* (Ehremberg) Lewin & Reimann (Bacillariophyceae) as a tool for measuring toxicity of linear alkylbenzene sulfonate (LAS) in sediments. *Bull. Environ. Contam. Tox.* **70**(2): 242–247.
34. Sevecke, S., Morawietz, T., Krumbein, W.E., and Rhiel, E. (1997). Use of percoll gradient centrifugation for the isolation of diatoms from Wadden Sea sediments; diatom yields, species recoveries and population diversity. *Microbiologia* **13**(2): 173–184.
35. Bennet, A., Bianchi, T.S., Means, J.C., and Carman, K.R. (1999). The effects of polycyclic aromatic hydrocarbon and grazing on the abundance and composition of microphytobentos in salt marsh sediments (Pass Fourchon, LA). I. A microcosm experiment. *J. Exp. Mar. Biol. Ecol.* **242**: 1–20.
36. Bennet, A., Bianchi, T.S., and Means, J.C. (2000). The effects of PAH contamination and grazing on the abundance and composition of microphytobenthos in salt marsh sediments (Pass Fourchon, LA, USA): II: The use of plant pigments as biomarkers. *Estuarine Coastal Shelf Sci.* **50**: 425–439.
37. Light, B.R. and Beardall, J. (2001). Photosynthetic characteristics of sub-tidal benthic microalgal populations from a temperate, shallow water marine ecosystem. *Aquat. Bot.* **70**: 9–27.
38. Cartaxana, P., Jesus, B., and Brotas, V. (2003). Pheophorbide and pheophytin a-like pigments as useful markers for intertidal microphytobentos grazing by *Hydrobia ulvae*. *Estuarine Coastal Shelf Sci.* **58**: 293–297.
39. Seuront, L. and Spilmont, N. (2002). Self-organized criticality in intertidal microphytobentos path patterns. *Physica A* **313**: 513–539.
40. Rainey, M.P. et al. (2003). Mapping intertidal estuarine sediment grain size distributions through airborne remote sensing. *Remote Sensing Environ.* **86**: 480–490.
41. Taylor, I.S. and Paterson, D.M. (1998). Microspatial variation in carbohydrate concentrations with depth in the upper millimetres of intertidal cohesive sediments. *Estuarine Coastal Shelf Sci.* **46**: 359–370.
42. Perkins, R.G. et al. (2003). Changes in microphytobenthic chlorophyll *a* and EPS resulting from sediment compaction due to de-watering: opposing patterns in concentration and content. *Continental Shelf Res.* **23**: 575–586.
43. Wong, C.K.C., Cheung, R.Y.H., and Wong, M.H. (1999). Toxicological assessment of coastal sediment in Hong Kong using a flagellate, *Dunaliella tertiolecta*. *Environ. Pollut.* **105**: 175–183.
44. Abalde, J. et al. (1995). Response of the marine microalga *Dunaliella tertiolecta* (Chlorophyceae) to copper toxicity in short time experiments. *Bull. Environ. Contam. Toxicol.* **54**: 317–324.
45. Moreno-Garrido, I., Lubián, L.M., and Soares, A.M.V.M. (1999). Growth differences in cultured populations of *Brachionus plicatilis* Müller caused by heavy metal stress as function of microalgal diet. *Bull. Environ. Contamin. Tox.* **63**: 392–398.
46. Tsvetnenko, Y. and Evans, L. (2002). Improved approaches to ecotoxicity testing of petroleum products. *Mar. Pollut. Bull.* **45**: 148–156.
47. Bonilla, S., Conde, D., and Blank, H. (1998). The photosynthetic responses of marine phytoplankton, periphyton and epipsammon to the herbicides paraquat and simazine. *Ecotoxicology* **7**: 99–105.
48. Baun, A., Justensen, K.B., and Nyholm, N. (2002). Algal test with soil suspensions and elutriates: a comparative evaluation for PAH-contaminated soils. *Chemosphere* **46**: 251–258.
49. Tolun, L.G. et al. (2001). The pollution status and the toxicity of the surface sediments in Izmit Bay (Marmara Sea), Turkey. *Environ. Int.* 163–168.
50. Liß, W. and Ahlf, W. (1997). Evidence of whole-sediment, porewater and elutriate testing in toxicity assessment of contaminated sediments. *Ecotoxicol. Environ. Safety* **36**: 140–147.
51. Cairns, J., Jr., McCormick, P.V., and Belanger, S.E. (1993). Prospects for the continued development of environmentally-realistic toxicity tests using microorganisms. *J. Environ. Sci.* **5**(3): 253–268.
52. OECD (Organization for the Economic Cooperation and Development). (1998). *OECD Guideline for Testing of Chemicals: Alga, Growth Inhibition Test*.
53. Florence, T.M., Lumsden, B.G., and Fardy, J.J. (1983). Evaluation of some physicochemical techniques for the determination of the fraction of dissolved copper toxic to the marine diatom *Nitzschia closterium*. *Anal. Chim. Acta* **151**: 281–295.
54. Morris, E.P. and Kromkamp, J.C. (2003). Influence of temperature on the relationship between oxygen- and fluorescence-based esimates of photosynthetic parameters in a marine benthic diatom (*Cylindrotheca closterium*). *Eur. J. Phycol.* **38**: 133–142.
55. Rachlin, J.W., Warkentine, B., and Jensen, T.E. (1983). The growth response of the diatom *Chlorella saccarophyla, Navicula incerta* and *Nitzschia closterium* to selected concentrations of cadmium. *Bull. Torrey Bot. Club* **110**(2): 217–223.
56. Pablo, F., Stauber, J.L., and Buckney, R.T. (1997). Toxicity of cyanide and cyanide complexes to the marine diatom *Nitzschia closterium*. *Water Res.* **31**(10): 2435–2442.
57. Stauber, J.L. and Florence, T.M. (1985). The influence of iron on copper toxicity to the marine diatom, *Nitzschia closterium* (Ehrenberg) W. Smith. *Aquat. Toxicol.* **6**: 297–305.
58. Stauber, J.L. and Florence, T.M. (1990). Mechanism of toxicity of zinc to the marine diatom *Nitzschia closterium*. *Mar. Biol.* **105**: 519–524.
59. Adams, M.S. and Stauber, J.L. (2004). Development of a whole-sediment toxicity test using a benthic marine microalga. *Environ. Toxicol. Chem.* **23**(8): 1957–1968.
60. Hampel, M. et al. (2001). Acute toxicity of LAS homologues in marine microalgae. Esterase activity and inhibition growth as endpoints of toxicity. *Ecotoxicol. Environ. Safety* **48**: 287–292.

MARINE STOCK ENHANCEMENT TECHNIQUES

JANA DAVIS
Smithsonian Environmental
Research Center
Edgewater, Maryland

The United Nations estimates that 28% of all ocean fishery stocks are overexploited or severely depleted, and another 47% is fully exploited (1). Included in this statistic are species that are long-lived and short-lived, open-ocean and coastal, migratory and sedentary. About 75% of fishery stocks are unable to withstand further exploitation, so only 25% of stocks remain to satisfy the increasing human demand, as human population grows, for fishery products.

As a result of the pressure on fishery populations and the degradation of many of these resources, a suite of fishery management techniques have been developed. One such technique, stock enhancement, is broadly defined as the (human) activity by which the population of a species is increased. In contrast, more traditional management techniques generally focus on limiting human activities that reduce the population of a species.

Traditional measures may focus on reducing fishing effort, for example, by limiting the number of days fishermen can fish, restricting the amount of gear a fishermen can use, regulating the type of gear a fisherman can deploy, or setting restrictions on the types of individuals caught by size or gender. Not only does this latter method serve to reduce catch, it strives to shift the pressure away from more reproductively important individuals toward those with less likely contributions to the next generation over the course of their lifetimes.

GOALS OF STOCK ENHANCEMENT: A NONTRADITIONAL MANAGEMENT TECHNIQUE

The technique of stock enhancement, while often used in concert with more traditional catch-reducing techniques, does not describe limitations to human activities but rather the active improvement of a fishery stock. Enhancement efforts can encompass many approaches. The population can be bolstered both by increasing carrying capacity (the number of individuals that can be supported by their habitat) or by increasing the number of individuals themselves (2,3). The former method is effective only when the factor limiting population is related to habitat, food, or some other resource. The latter method, a special case of aquaculture, is effective only when the habitat can hold more individuals than it does at present.

The key to success in stock enhancement is to identify the factor(s) that limit the population in question and then increase the level of that factor. Incorrect identification of this factor could lead to wasted time, money, and energy. Consider the hypothetical case of a coral reef fish. If the fish population is not habitat-limited, but is recruitment-limited, adding more coral reef habitat would likely serve only to spread out the population spatially. No additional recruits would be available to settle on the new habitat substrate. On the other hand, if the fish is habitat-limited, adding additional individuals will not increase the ultimate population, as there would be no space for them to occupy.

Enhancement directed toward habitat-limited populations has included efforts to enhance stocks of both fish, especially salmonids, and invertebrates (4–7). Often these cases are referred to as “habitat restoration,” stressing the action rather than the consequence. Most often when people refer to “stock enhancement,” they are referring to hatchery-raised individuals, usually juveniles, that are released into the wild to bolster recruitment-limited populations directly. These added juveniles are raised outside of the system, for example, in, an aquaculture facility or a field pen. The juveniles may be offspring of wild parents collected from the field (8) or offspring from generations of parents held in captivity (9).

The goal of stock enhancement is to raise the individuals beyond the initial phase of high early life-history mortality and then release them into the wild. The stage at which they are released is determined through economic optimization, survivorship maximization, or a combination of both. Maintaining individuals, especially heterotrophs, in a hatchery or a pen is expensive. The expense tends to increase as organisms grow. One might calculate the optimum release size or age as a trade-off between the cost of maintaining an individual in the hatchery and the survivorship advantage that the hatchery offers. Often, as individuals grow, the survivorship advantage of the hatchery environment over the natural environment begins to reverse. Many organisms are cannibalistic when held in extremely high densities (10), a trend that increases as organisms grow. Disease transmission also becomes a problem at high densities. Lowering densities requires an expensive solution of creating more tank or pen space. Therefore, at a certain point, mortality rates are actually lower if the organisms are released into the wild than if held in the hatchery.

The optimum release point is different for each species. For example, Kemp’s ridley sea turtles are released at the age of 1 year (11). Blue crabs are released at the age of 2 to 3 months (12). Individuals that are never released but harvested in the hatchery are part of a complete aquaculture program such as, for example, farm-raised salmon, catfish, or shrimp (13,14).

In populations truly at risk, the goal of stock enhancement is not simply to provide more fish to catch. Ideally, after release, these hatchery-reared individuals contribute to the spawning stock and have a potential exponential impact on overall population generation after generation, depending on the degree of recruitment limitation. In these cases, a purely economic model to predict optimum size at release is not appropriate. Stock enhancement therefore has a potential role beyond fishery application and into the realm of threatened or endangered species protection.

STOCK ENHANCEMENT EXAMPLES

Stock enhancement techniques have been applied to many types of organisms worldwide. Finfish populations have been the most common recipients of hatchery-raised juveniles, including many salmon from both the Atlantic and Pacific (15,16), Japanese flounder (17), Hawaiian mullet (18), Nassau grouper (19), and Chesapeake Bay striped bass (20), just to name a few. Holothurians have been hatchery-raised for stock enhancement (21). Bivalves, such as quahogs (22) and soft-shell clams (23), and gastropods, such as abalones (24) and queen conch (25), have been the subject of stock enhancement efforts. Finally, crustaceans are beginning to receive more attention as possible stock enhancement targets; programs have been initiated for decades with American and

European lobsters (26,27), more recent programs for prawn (28) and the Japanese swimming crab (3), and a new exploratory program established for the Chesapeake Bay blue crab (8).

Stock enhancement efforts extend worldwide. In Japan, stock enhancement programs are in place for at least 34 finfish and 12 crustaceans (3). Programs relying on various specific techniques exist in South America, North America, Africa, Europe, Asia, Australia/New Zealand, and in the developed world as well as the developing world (29–31).

Specific techniques to optimize stock enhancement depend on the life-history traits of the species. Sessile organisms may be enhanced by seeding areas with young individuals (32). Depending on how closely the success of the enhancement effort is to be monitored, more mobile species may be released into isolated areas where they can be followed, similar to the idea of stocking freshwater lakes (8,28). For extremely mobile open-ocean species, juveniles are released into the open ocean (3), and even if tagged, large numbers cannot be followed over time.

PROBLEMS OF STOCK ENHANCEMENT

Enhancement has been a controversial management method for several reasons (15,29,33). First, hatchery-raised animals may not survive well in the wild. The hatchery environment can offer conditions very different from the natural environment: food is likely to be different, the method of foraging is different, and therefore the hatchery-raised organisms may have mortality-threatening inexperience with natural prey after their introduction into the wild. They may similarly be inexperienced in avoiding predators. Holding tanks may restrict movement, may have unnatural substrates, may have unnatural light regimes, or may have different flow regimes. As a result of all of these differences, stock enhancement efforts, often funded by the taxpayer, may not be successful (34). Many studies have noted differences in hatchery-raised and wild individuals in factors such as behavior, morphology, growth rates, and therefore survivorship (17,35,36).

A second concern is that hatchery-raised individuals may be too successful. Hatchery animals may compete with and displace wild animals (2,37). If at any point the population becomes habitat-limited, rather than recruitment-limited (or if the stock enhancement effort pushes the population over the carrying capacity), the survivorship of wild individuals may actually decrease. The ultimate result would then be a decline in the population of wild individuals, even if the overall population has increased.

A third concern is that hatchery-raised organisms may carry diseases into the wild, affecting the survivorship of wild individuals. Fourth, increases in stock size due to hatchery successes may provoke a rise in fishing effort and therefore greater pressure on the remaining wild individuals (34,38). Stock enhancement efforts are generally performed in concert with traditional management techniques, such as lowering catch and/or restricting catch to certain segments of the population. However, if the human population increases, pressure may be put on managers to reduce fishing regulations. Again, the ultimate result would be a decline in the survivorship rates of wild individuals, even if overall population has increased.

Finally, and perhaps the reason for the controversy generating the most attention in recent years is genetics, or "gene dilution." Even if efforts in the hatchery are expended to obtain brood stock from many different parents, genetic variability in the brood stock is unlikely to approach that of the wild parental stock. The offspring produced will be more similar to each other than the offspring of wild stock. Opponents of stock enhancement fear that releasing genetically relatively homogenous hatchery individuals will reduce genetic variability in wild populations as hatchery and wild individuals interbreed. The ultimate result after several to many generations might be a decline in population because lack of genetic diversity can limit a population's response to environmental change.

The controversy over stock enhancement has been fueled by the fact that most enhancement efforts have not been studied quantitatively (37,41,42). Quantitative, hypothesis-driven study did not begin until the late 1980s (43). However, now that the necessity to understand better the outcomes of enhancement programs has been recognized, in many cases the enhancement process has been refined. Methods to select better candidate species have been developed (44). Advances in tagging techniques have allowed better assessment and comparison of hatchery and wild animal survivorship (44). Most importantly, calls have been made for quantitative study of small-scale enhancement efforts before investment in large-scale programs begins (42,43).

OPTIMIZING STOCK ENHANCEMENT SUCCESS

Stock enhancement is becoming a more often commonly used method for addressing declines in fished species. Many of the problems identified above can be addressed to a comfortable degree before the program is initiated. For example, several programs are considering conditioning individuals to limit the differences between hatchery-raised and wild organisms before release into the wild (12,45,46). In this way, survivorship rates of hatchery organisms are increased, along with the output per enhancement program dollar spent.

In addition, carrying capacities of targeted release areas can be determined before programs are initiated to (1) determine whether the case is appropriate for stock enhancement and (2) determine optimum release densities of hatchery-raised organisms. Certain microhabitats within a targeted release region may have higher carrying capacity and greater food or refuge resources than others, and therefore distribution of hatchery-raised organisms can be optimized.

Methods of determining whether stock enhancement is successful are more difficult. Determining success requires knowledge about how well hatchery-raised organisms survive, how well wild organisms survive to allow comparison, and the contribution of hatchery organisms to the total population. As programs are developed, a wide range of survivorship and contribution values have been reported. For example, survivorship to fishery size was 3–4% for stocked panaeid shrimp (28), 21% for red drum in Texas (47), and up to 30% for Japanese flounder (3). Even among programs for the same species, values range widely. For example, in some European lobster programs, no hatchery-raised lobsters were recaptured in the fishery, and in others, 10–35% of landed lobsters were of hatchery origin (37).

Often the steps to quantify the success of a stock enhancement program take years of scientific study and require laborious study efforts, such as tagging and resampling individuals over time. Such efforts have been deemed mandatory by critics of stock enhancement programs before public monies are used to support these efforts to bolster fishery stocks, one of the world's most important natural resources.

BIBLIOGRAPHY

1. Food and Agriculture Organization of the United Nations. (2002). *The State of World Fisheries and Aquaculture*. United Nations, Rome, Italy.
2. Castro, K.M., Cobb, J.S., Wahle, R.A., and Catena, J. (2001). Habitat addition and stock enhancement for American lobsters, *Homarus americanus*. *Mar. Freshwater Res.* **52**: 1253–1261.
3. Masuda, R. and Tsukamoto, K. (1998). Stock enhancement in Japan: Review and perspective. *Bull. Mar. Sci.* **62**: 337–358.
4. Beamesderfer, R.C.P. and Farr, R.A. (1997). Alternatives for the protection and restoration of sturgeons and their habitat. *Environ. Biol. Fish.* **48**: 407–417.
5. Hendry, K., Cragge-Hine, D., O'Grady, M., Sambrook, H., and Stephen, A. (2003). Management of habitat for rehabilitation and enhancement of salmonid stocks. *Fish. Res.* **62**: 171–192.
6. Peterson, C.H., Grabowsi, J.H., and Powers, S.P. (2003). Estimated enhancement of fish production resulting from restoring oyster reef habitat: Quantitative valuation. *Mar. Ecol. Prog. Ser.* **264**: 149–264.
7. Rodwell, L.D., Barbier, E.B., Roberts, C.M., and McClanahan, T.R. (2003). The importance of habitat quality for marine reserve—fishery linkages. *Can. J. Fish. Aquat. Sci.* **60**: 171–181.
8. Davis, J.L.D., Young-Williams, A.C., Hines, A.H., and Zohar, Y. (in press). Assessing the feasibility of stock enhancement in the case of the Chesapeake Bay blue crab. *Can. J. Fish. Aquat. Sci.*
9. Iguchi, K., Watanabe, K., and Nishida, M. (1999). Reduced mitochondrial DNA variation in hatchery populations of ayu (*Plecoglossus altivelis*) cultured for multiple generations. *Aquaculture* **178**: 235–243.
10. Kestemont, P. et al. (2003). Size heterogeneity, cannibalism and competition in cultured predatory fish larvae: Biotic and abiotic influences. *Aquaculture* **227**: 333–356.
11. Caillouet, C.W., Fontaine, C.T., Manzella-Tirpak, S.A., and Shaver, D.J. (1995). Survival of head-started Kemp's ridley sea turtles (Lepidochelys kempii) released into the Gulf of Mexico or adjacent bays. *Chelate Conserv. Biol.* **1**: 285–292.
12. Davis, J.L.D., Eckert-Mills, M.G., Young-Williams, A.C., Hines, A.H., and Zohar, Y. (in press). Morphological conditioning of a hatchery-raised invertebrate, *Callinectes sapidus*, to improve field survivorship after release. *Aquaculture* **133**: 1–14.
13. Fleming, I.A. et al. (2000). Lifetime success and interactions of farm salmon invading a native population. *Proc. R. Soc. London Ser. B* **267**: 1517–1523.
14. Wirth, F.F. and Davis, K.J. (2003). Seafood dealers' shrimp-purchasing behavior and preferences with implications for United States shrimp farmers. *J. Shellfish Res.* **22**: 581–588.
15. Washington, P.M. and Koziol, A.M. (1993). Overview of the interactions and environmental impacts of hatchery practices on natural and artificial stocks of salmonids. *Fish. Res.* **18**: 105–122.
16. Thorpe, J.E. (1998). Salmonid life-history evolution as a constraint on marine stock enhancement. *Bull. Mar. Sci.* **62**: 465–475.
17. Kellison, G.T., Eggleston, D.B., and Burke, J.S. (2000). Comparative behaviour and survival of hatchery-reared versus wild summer flounder (*Paralichthys dentatus*). *Can. J. Fish. Aquat. Sci.* **57**: 1870–1877.
18. Leber, K.M., Blankenship, H.L., Arce, S.M., and Brennan, N.P. (1996). Influence of release season on size-dependent survival of cultured striped mullet, *Mugil cephalus*, in a Hawaiian estuary. *Fish. Bull.* **95**: 267–279.
19. Roberts, C.M., Quinn, N., Tucker, J.W., and Woodward, P.N. (1995). Introduction of hatchery-reared Nassau grouper to a coral reef environment. *N. Am. J. Fish. Manage.* **15**: 159–164.
20. Secor, D.H. and Houde, E.D. (1998). Use of larval stocking in restoration of Chesapeake Bay striped bass. *ICES J. Mar. Sci.* **55**: 228–239.
21. Battaglene, S.C., Seymour, J.E., and Ramofafia, C. (1999). Survival and growth of cultured juvenile sea cucumbers, *Holothuria scabra*. *Aquaculture* **178**: 293–322.
22. Walton, W.C. and Walton, W.C. (2001). Problems, predators, and perception: Management of quahog (hardclam), *Mercenaria mercenaria*, stock enhancement programs in southern New England. *J. Shellfish Res.* **20**: 127–134.
23. Beal, B.F., Kraus, F., and Gayle, M. (2002). Interactive effects of initial size, stocking density, and type of predator deterrent netting on survival and growth of cultured juveniles of the soft-shell clam, *Mya arenaria* L., in eastern Maine. *Aquaculture* **208**: 81–111.
24. Tegner, M.J. and Butler, R.A. (1985). The survival and mortality of seeded and native red abalones, *Haliotis rufescens*, on the Palos Verdes Penninsula. *Calif. Fish Game* **71**: 150–163.
25. Stoner, A.W. and Glazer, R.A. (1998). Variation in natural mortality: Implications for queen conch stock enhancement. *Bull. Mar. Sci.* **62**: 427–442.
26. Bannister, R.C.A. and Addison, J.T. (1998). Enhancing lobster stocks: A review of recent European methods, results, and future prospects. *Bull. Mar. Sci.* **62**: 369–387.
27. Beal, B.F., Chapman, S.R., Irvine, C., and Bayer, R.C. (1998). Lobster (Homarus americanus) culture in Maine: A community-based, fishermen-sponsored, public stock enhancement program. *Can. Ind. Rep. Fish. Aquat. Sci.* **244**: 47–54.
28. Davenport, J., Ekaratne, S.U.K., Walgama, S.A., Lee, D., and Hills, J.M. (1999). Successful stock enhancement of a lagoon prawn fishery at Rekawa, Sri Lanka, using cultured post-larvae of penaeid shrimp. *Aquaculture* **180**: 65–78.
29. Bell, J.D. and Gervis, M. (1999). New species for coastal aquaculture in the tropical Pacific: Constraints, prospects and considerations. *Aqua Int.* **7**: 207–223.
30. Cowx, I.G. (1999). An appraisal of stocking strategies in the light of developing country constraints. *Fish. Manage. Ecol.* **6**: 21–34.

31. Phillips, B.F. and Liddy, G.C. (2003). Recent developments in spiny lobster aquaculture. In: Phillips, B., Megrey, B.A., and Zhou, Y. (Eds). *Proc. 3rd World Fish. Congr.*: Feeding the world with fish in the next millenium—the balance between production and environment. American Fisheries Society Symposium Vol. 38, pp. 43–57.
32. Barbeau, M.A. et al. (1996). Dynamics of juvenile sea scallop (*Placopecten magellameus*) and their predators in bottom seeding trials in Lanenburg Bay, Nova Scotia. *Can J. Fish. Aquat. Sci.* **53**: 2494–2512.
33. Li, J. (1999). An appraisal of factors constraining the success of fish stock enhancement programmes. *Fish. Manage. Ecol.* **6**: 161–169.
34. Hilborn, R. (1998). The economic performance of marine stock enhancement projects. *Bull. Mar. Sci.* **62**: 661–674.
35. Mana, R.R. and Kawamura, G. (2002). A comparative study on morphological differences in the olfactory system of red sea bream (*Pagrus major*) and black sea bream (*Acanthopagrus schlegeli*) from wild and cultured stocks. *Aquaculture* **209**: 285–306.
36. Davis, J.L.D. et al. (2004). Differences between hatchery-raised and wild blue crabs (*Callinectes sapidus*): Implications for stock enhancement potential. *Trans. Am. Fish. Soc.*
37. Bannister, R.C.A., Addison, J.T., and Lovewell, S.R.J. (1994). Growth, movement, recapture rate and survival of hatchery-reared lobsters (*Homarus gammarus* (Linnaeus, 1758)) released into the wild on the English East Coast. *Crustaceana* **67**: 156–172.
38. Bannister, R.C.A. (2000). Lobster *Homarus gammarus* stock enhancement in the U.K.: Hatchery-reared juveniles do survive in the wild, but can they contribute significantly to ranching, enhancement, and management of lobster stocks? *Can. Spec. Publ. Fish. Aquat. Sci.* **130**: 23–32.
39. Tringali, M.D. and Bert, T.M. (1998). Risk to genetic effective population size should be an important consideration in fish stock-enhancement programs. *Bull. Mar. Sci.* **62**: 641–659.
40. Utter, F. (1998). Genetic problems of hatchery-reared progeny released into the wild, and how to deal with them. *Bull. Mar. Sci.* **62**: 623–640.
41. Heppell, S.S. and Crowder, L.B. (1998). Prognostic evaluation of enhancement programs using population models and life history analysis. *Bull. Mar. Sci.* **62**: 495–507.
42. Crowe, T.P. et al. (2002). Experimental evaluation of the use of hatchery-reared juveniles to enhance stocks of the topshell *Trochus niloticus* in Australia, Indonesia and Vanuatu. *Aquaculture* **206**: 175–197.
43. Leber, K.M. (1999). Rational for an experimental approach to stock enhancement. In: *Stock Enhancement and Sea Ranching*. B.R. Howell, E. Moksness, and T. Sväsand (Eds.). Blackwell Science, Oxford, pp. 63–75.
44. Blackenship, H.L. and Leber, K.M. (1995). A responsible approach to marine stock enhancement. *Am. Fish. Soc. Symp.* **15**: 167–175.
45. Olla, B.L. and Davis, M.W. (1989). The role of learning and stress in predator avoidance of hatchery-reared coho salmon (*Oncorhynchus kisutch*) juveniles. *Aquaculture* **76**: 209–214.
46. Berejikian, B.A. (1995). The effects of hatchery and wild ancestry and experience on the relative ability of steelhead trout fry (*Oncorhynchus mykiss*) to avoid a benthic predator. *Can. J. Fish. Aquat. Sci.* **52**: 2476–2482.
47. McEachron, L.W., Colura, R.L., Bumgaurdner, B.W., and Ward, R. (1998). Survival of stocked red drum in Texas. *Bull. Mar. Sci.* **62**: 359–368.

PHYSICAL AND CHEMICAL VARIABILITY OF TIDAL STREAMS

Paolo Magni
IMC—International Marine Centre
Torregrande-Oristano, Italy

Shigeru Montani
Hokkaido University
Hakodate, Japan

This article focuses on the variability of major hydrologic parameters of tidal streams that occur on different temporal and spatial scales within an individual estuary. Rapid changes in water temperature and salinity take place on a timescale of <1 hour, as a strong function of the tidal state (e.g., low vs. high tide) and amplitude (e.g., spring vs. neap tide). Large day–night fluctuations of dissolved oxygen (DO) concentration are also demonstrated and discussed in relation to the extent of *in situ* biological processes. In addition, the variability of these parameters is evaluated seasonally at controlled low tide. Finally, the year-round distribution of nutrient concentrations (i.e., ammonium, phosphate, silicate, and nitrate+nitrite) in low tide ebbing water is presented and compared with the variability in nutrient concentrations in the sediment porewater of an adjacent exposed sandflat.

SHORT-TERM (HOUR) VARIABILITY

A tidal stream crossing the interidal zone of an estuary in the Seto Inland Sea, Japan, was selected to describe daily fluctuations in water temperature, salinity, and DO concentration (1). Continuous measurements lasting 24 to 28 hours were made during two different tidal cycles, a spring tide of May 1995 and a neap tide of May 1996, using a CTD recorder (Alec AST-1000) placed about 10 cm from the bottom sediment (2,3).

Rapid and drastic changes in the selected parameters occurred during both surveys in 1995 (Fig. 1a) and 1996 (Fig. 1b). Major differences between the two surveys were also found, consistent with spring–neap tidal contrasts (4). Fluctuations in water depth were stronger in the 1995 survey than in the 1996 survey due to a larger amplitude in the tidal cycle during a spring tide (Fig. 1a). In particular, two major lower and higher tidal levels, typical of *mixed-semidiurnal* estuaries, occurred on May 30–31, 1995, as predicted by the local Maritime Safety Agency (2); the maximum tidal range was ca. 2 m (Fig. 1a).

Accordingly, the temporal changes in salinity were largest during the 1995 survey; salinity dropped from >31 to 2.0 psu, coincident with the lower low tide, and rapidly increased up to >31 psu again before the following higher high tide (Fig. 1a). At the subsequent higher low tide, the decrease in salinity was less marked, and minimum values were about 25 psu. An overall more restricted fluctuation of salinity occurred in the 1996

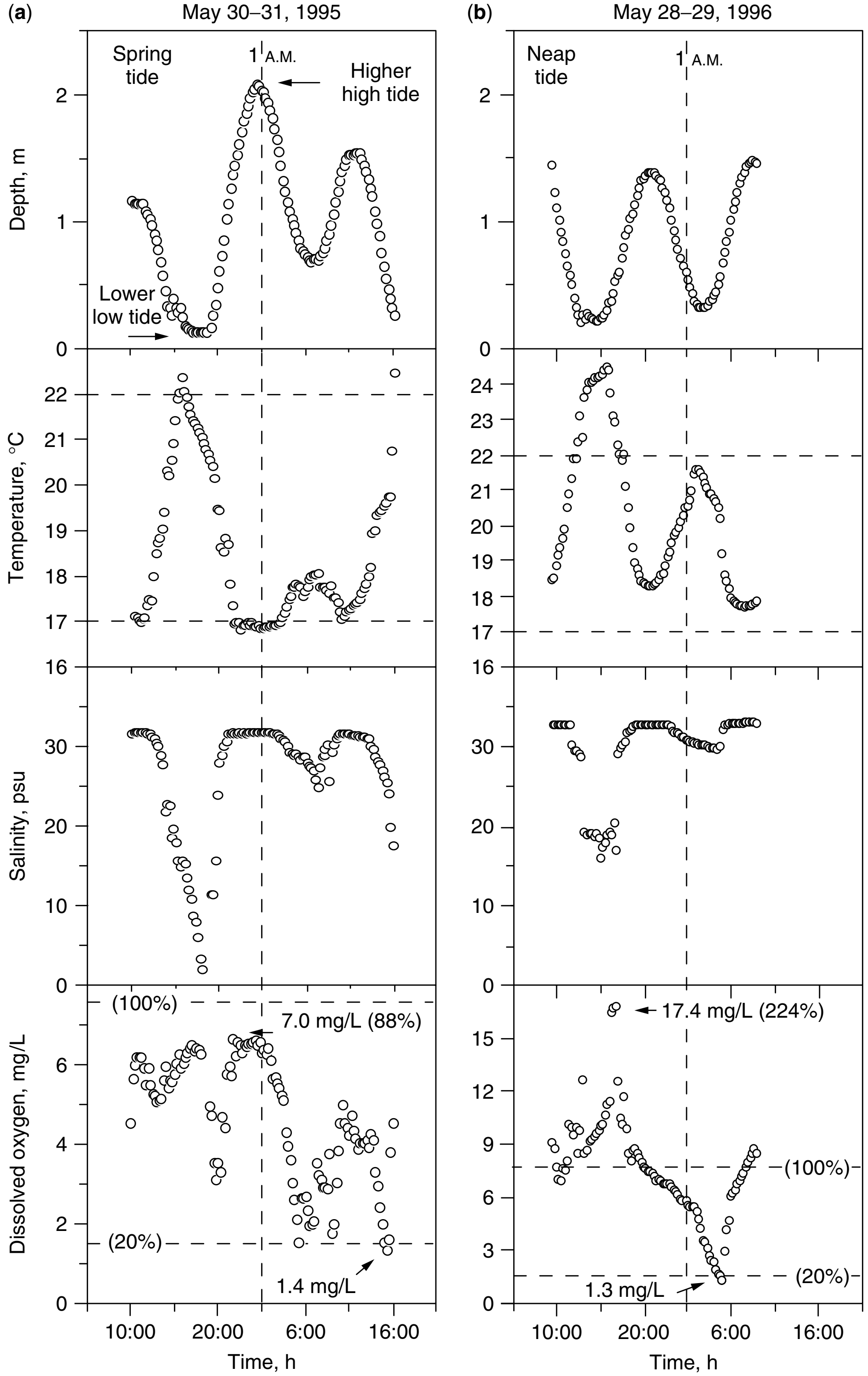

Figure 1. Daily fluctuations of water depth, temperature, salinity and dissolved oxygen (DO) concentration of a tidal stream in an estuarine intertidal zone (Seto Inland Sea, Japan). Measurements were made using a CTD cast placed ca. 10 cm from the bottom sediment at the end of May of 2 consecutive years (a: 1995; b: 1996). Relevant sensors were logged every 15 min. Notes: horizontal dashed lines in temperature and DO boxes individuate major differences within the same range of values (17–22 °C and 20–100% of air saturation, respectively) recorded in the 2 years. In DO boxes, absolute minimum and maximum values are also indicated. The vertical dashed line is arbitrarily depicted at 1 A.M. to highlight the temporal differences in tidal state and amplitude between the two surveys.

survey, consistent with the reduced amplitude of the tide (Fig. 1b) and a rainfall regime more limited in spring 1996 than in spring 1995 (1). These results show that saline intrusion is a strong function of tidal state (e.g., low vs. high tide) and amplitude (e.g., spring vs. neap tide). They also demonstrate that tidal streams may experience, within a few hours, strong changes in salinity that cause very different habitat conditions at the same location in an estuary. These may vary from oligohaline (0.51–5 psu) to euhaline (30.1–40 psu) conditions, in agreement with salinity ranges given by the Venice Conference (5).

The daily fluctuation in water temperature varied according to the tidal cycle and showed a temporal pattern opposite to that of salinity during both surveys (Fig. 1a, b). The range of temperature was relatively larger in May 1996 (17.7–24.5 °C) than in 1995 (16.9 °C–22.3 °C). In both cases, it was apparent that warmer waters were brought into the estuary by the freshwater runoff during ebb flow. These results are consistent with the period and site of measurements; physical processes of heat transfer in spring–summer (and water cooling in winter) are most effective in the upper and shallower riverine zone of the estuary (1).

Dissolved oxygen (DO) concentration was subjected to strong daily fluctuations, partly as a function of the tidal state (and water depth). The two surveys also showed major differences (Figs. 1a and b). In the 1995 survey, DO concentration was mostly within normoxic values, ranging from 1.4 mg/L to 7.0 mg/L. Differently in the 1996 survey, DO concentration showed a larger variation, in much higher values up to 17.4 mg/L.

Elevated DO concentrations are related to ecosystem processes of primary production (6). It is known that shallow lagoons and coastal areas dominated by seagrass or macroalgae are subjected to oversaturation of DO (i.e., >100% of air saturation), especially during the warm period (7,8). Viaroli et al. (7) reported DO oversaturation up to 150% in the near-bottom water of a coastal lagoon dominated by the macroalga *Ulva*. This was followed by the outbreak of a dystrophic crisis, complete anoxia through the water column at some stations of the lagoon. The development of large amounts of macroalgae (*Ulva* sp.) also tends to occur in the estuarine sand flat of this study coincident with increasing temperature and solar radiation during the spring. Extended beds of macroalgal biomass were present during the field measurements of the 1996 survey (personal observations). Accordingly, during the daytime measurements of May 28, DO concentration rapidly rose to oversaturation; a major increase of >200% of air saturation occurred between 16:00 and 16:30 (Fig. 1b), indicating a period of major oxygen production by macroalgae. Similar to the extremely high DO values found in this study, Piriou & Ménesguen (9) reported that the *in vitro* growth of *Ulva* under light saturation and nutrient enrichment raised the DO concentration to 22 mg/L, 4 hours after the experiment started.

By contrast, during nighttime, a progressive decrease in DO concentration occurred, down to hypoxic values of <1.4 mg/L or <14% of air saturation (10). Subsequently, at the next daylight of May 29, the DO concentration was back to normoxic values, reaching oversaturation (Fig. 1b). Differently during the 1995 survey, a much reduced macroalgal biomass was found that tended to be rapidly flushed away by heavy rainfall in spring 1995 (1). This was associated with a more a restricted range of DO concentration that never reached oversaturation during the daytime; whereas, similar to the 1996 survey, hypoxic values occurred during nighttime (Fig. 1a, b). Thus, the different extent of macroalgal coverage indicated how a large DO fluctuation might strongly depend on the extent of biological processes, such as photosynthesis and primary production in the presence of light and respiration at night.

SEASONAL VARIABILITY

This section presents the results of monthly measurements of water temperature, salinity, and DO concentration at controlled low tide on the same tidal stream where the short-term measurements of May 1995 and May 1996 were conducted (see above section and Fig. 1). Between January 1995 and April 1996, the water temperature varied from 3.6 °C ± 0.3 (December) to 29.0 °C ± 1.8 (July) (Fig. 2). Statistically significant differences were demonstrated for the specific geographical area (latitude 34°21′ N, longitude 43°21′E) between a warm period (May–October, water temperature 23.5 ± 4.2 °C) and a cold period (November–April, water temperature 10.0 ± 4.4 °C) (1).

The seasonal variability in low-tide salinity (Fig. 2) was consistent with the results obtained from the short-term measurements of May 1995 and May 1996 during a complete tidal cycle (Fig. 1). In particular, the salinity recorded during the seasonal survey was lowest on May 16, 1995 at 4.6 ± 1.5 psu (Fig. 2). Accordingly, the short-term measurements of May 1995 showed a salinity decrease during low tide to <5 psu (Fig. 1a). By contrast, reduced rainfall during the second half of 1995 and the first half of 1996 (1) corresponded to a progressive increase in salinity in the low-tide stream water during that period. In particular, salinity was >25 psu between the end of September 1995 and April 1996 and reached the highest values in November 1995 at 31.4 ± 1.1 psu (Fig. 2). Consistently, also during the short-term measurements of May 1996, the decrease in salinity during low tide was more limited than that found in May 1995 (Fig. 1a, b).

The DO concentration also varied greatly from 5.4 ± 1.7 mg/L (September 1995) to 15.1 ± 1.1 mg/L (April 1996)

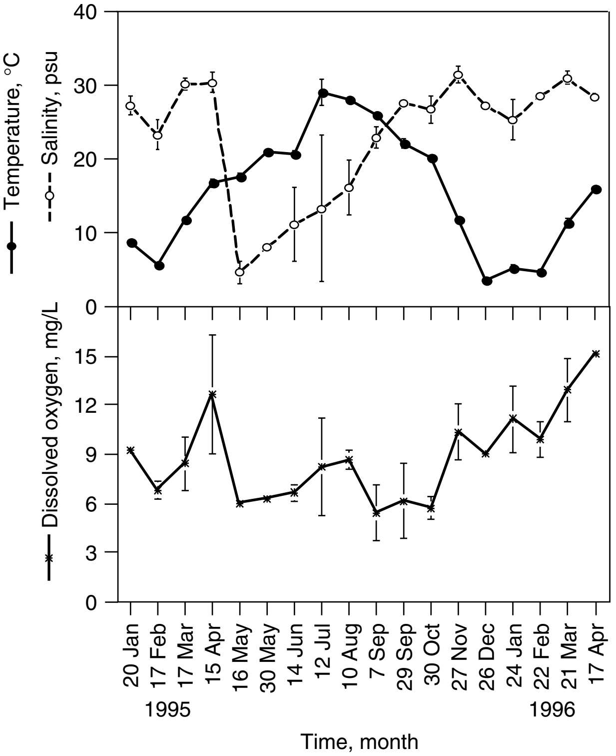

Figure 2. Seasonal variability in temperature, salinity, and dissolved oxygen (DO) concentrations in estuarine ebbing waters during a daytime low tide (Seto Inland Sea, Japan). Data source: Reference 1 (partially redrawn from Fig. 2, as mean values ± standard deviation of Stns. H1 and H2) (with kind permission of Kluwer Academic Publishers).

(Fig. 2). The temporal distribution of the DO concentration highlighted that, after a progressive increase from February 1995 (6.8 ± 0.5 mg/L) to April 1995 (12.6 ± 3.6 mg/L), a noticeable drop in DO concentration to 6–7 mg/L occurred in late spring, May 1995. This tended to persist throughout the following months (Fig. 2). In the second part of the seasonal survey, the DO concentration progressively increased again from relatively low values of 5.7 ± 0.7 mg/L^{-1} in October 1995 to 15.1 mg/L in April 1996 (Fig. 2). The temporal variability of the DO concentration was also consistent with the differences found during the short-term measurements. In particular, the DO level hardly reached oversaturation during the short-term measurements of spring 1995 (Fig. 1a), but it was up to >200% of air saturation in spring 1996 (Fig. 1b).

Nutrients

Nutrients, such as ammonium (NH_4^+-N), nitrate+nitrite ($NO_3^- + NO_2^-$-N), phosphate (PO_4^{3-}-P) and silicate [$Si(OH)_4$-Si], play a major role in controlling the growth, abundance, and structure of primary producers in the ocean (11,12). A qualitative and quantitative evaluation of their distribution represents therefore a fundamental step in assessing the dynamics of biophylic elements (e.g., carbon, nitrogen, phosphorus, and silicon) and the system processes of production and respiration (6). The variability in nutrient concentrations in transitional aquatic systems, such as estuaries and tidal streams, is related to the close interaction between water chemistry, tidal hydrology, and sedimentary processes (1,13,14).

Several approaches have been undertaken on different temporal and spatial scales to evaluate such variability. Time-series surveys of nutrient distribution have been carried out in relation to the ebb–flood tidal state (15–18) and the spring–neap tidal amplitude (4,19,20) or by controlled sampling regimes, such as around high tide (16,21). Other studies in coastal marine systems have used either intact cores (22,23) or *in situ* chambers (24) to quantify the nutrient flux at the sediment–water interface. However, a satisfactory evaluation of the actual physicochemical and biological processes from field enclosed experiments may be controversial because of the nonnatural state of the system, which is shielded from large water movements (25,26). This is particularly true in intertidal environments that are subjected to strong hydrodynamic tidal energy and are more rapidly and directly influenced by “external” variables (e.g., rainfall) and *in situ* benthic processes (e.g., nitrogen uptake by primary producers and release by macrozoobenthos) (27–33).

This work provides an example of the variability of nutrient concentrations in tidal stream water, consistent with large seasonal changes in temperature, salinity, and DO concentration (Fig. 2). A comparison with the nutrient concentrations in the sediment porewater from an adjacent sand flat exposed at low tide is given; it shows how major differences may occur between two closely interacting compartments of an estuary.

The seasonal distribution of nutrient concentrations in ebbing water and in sediment porewater is depicted in Fig. 3, and their annual range of variability is given in Table 1. Ammonium and phosphate concentrations

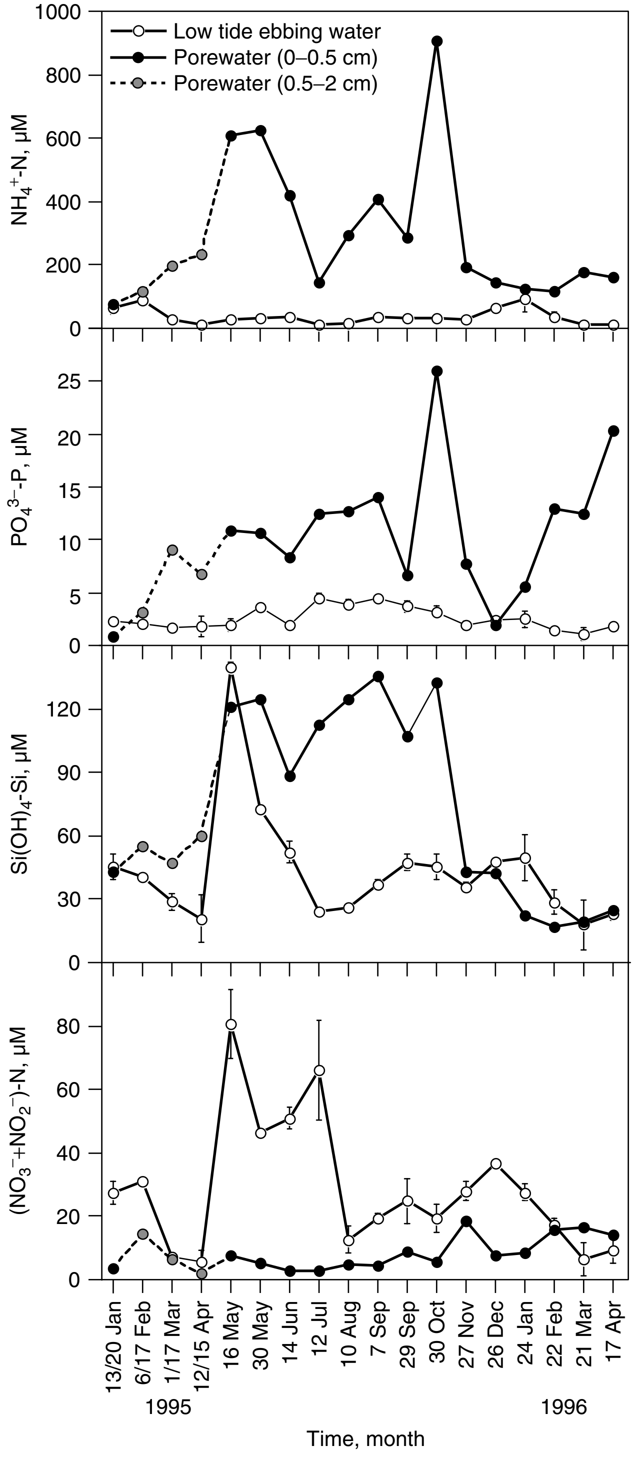

Figure 3. Seasonal variability in nutrient concentrations in estuarine ebbing waters and surface (0–0.5 cm) sediment porewater on an adjacent sand flat measured during a daytime low tide (Seto Inland Sea, Japan). Grey circles refer to porewater nutrient concentrations in subsurface (0.5–2 cm) sediments. On the horizontal axis, two different days from January 1995 to April 1995 refer to the dates of porewater (first day) and ebbing water (second day, see also Fig. 2) sampling. Data sources: ebbing water, data from Reference 1 (partially redrawn from Fig. 3, as mean values ± standard deviation of Stns. H1 and H2) (with kind permission of Kluwer academic Publishers).

Table 1. Annual Minimum and Maximum Values (± Standard Deviation and Period of The Year) of Nutrient Concentrations (μM) in Low Tide Ebbing Water and in the Porewater of Surface (0–0.5 cm) Sediments from an Adjacent Emerged Sand Flat (Seasonal Variation in Fig. 3)

	Ebbing Water		Sediment Porewater	
	Minimum	Maximum	Minimum	Maximum
NH_4^+	10.2 ± 6.8 (Apr '95)	90.8 ± 39.2 (Jan '96)	75.1 (Jan '95)	908 (Oct 30, '95)
PO_4^{3-}	1.1 ± 0.6 (Mar '96)	4.5 ± 0.4 (Jul '95)	0.9 (Jan '95)	26.0 (Oct 30, '95)
$Si(OH)_4$	17.9 ± 11.7 (Mar '96)	139.6 ± 2.5 (May 16, '95)	17.1 (Feb '96)	135.7 (Sept 7, '95)
$NO_3^- + NO_2^-$	5.7 ± 3.5 (Apr '95)	80.7 ± 11.0 (May 16, '95)	2.8 (Jul '95)	18.6 (Nov '95)

in ebbing water are much lower than those found in porewater in all seasons and on most occasions. However, a noticeable peak of ammonium concentration in ebbing water was observed during the colder months (December–February), coincident with a general decrease in ammonium concentration in the porewater. To some extent, this was also true for phosphate, which showed relatively lower values in porewater than in ebbing water in January 1995 and December 1995 (Fig. 3). Silicate concentration tended to follow a pattern similar to that of ammonium and phosphate, especially between spring and late summer. However, the extent of the differences in silicate concentration between porewater and ebbing water was more limited in autumn and spring. For the study area examined, it is suggested that a diffusive upward flux of nutrients from sediments, if solely based on the water–sediment concentration gradients (29,34) would be relatively limited, yet highly variable with season.

Different from the spatial distribution of ammonium, phosphate, and silicate, nitrate+nitrite concentration was higher in ebbing water than in porewater; major peaks occurred during periods of enhanced rainfall, as indicated by the sharp decrease in salinity on May 16 (Fig. 2). This is consistent with previous companion studies that demonstrated that freshwater input is a major contributor of nitrate+nitrite to the estuary of this work (1,2).

Finally, it can be highlighted that low nitrate concentration in porewater, further decreasing through the sediment column, commonly occurs due to numerous processes within sediment, such as nitrate ammonification and/or denitrification processes (35,36).

BIBLIOGRAPHY

1. Magni, P. and Montani, S. (2000). Water chemistry variability in the lower intertidal zone of an estuary in the Seto Inland Sea, Japan: Seasonal patterns of nutrients and particulate compounds. *Hydrobiologia* **432**: 9–23.
2. Montani, S. et al. (1998). The effect of a tidal cycle on the dynamics of nutrients in a tidal estuary in the Seto Inland Sea, Japan. *J. Oceanogr.* **54**: 65–76.
3. Magni, P., Montani, S., and Tada, K. (2002). Semidiurnal dynamics of salinity, nutrients and suspended particulate matter in an estuary in the Seto Inland Sea, Japan, during a spring tide cycle. *J. Oceanogr.* **58**: 389–402.
4. Vörösmarty, C.J. and Loder, T.C. III. (1994). Spring-neap tidal contrasts and nutrient dynamics in a marsh-dominated estuary. *Estuaries* **17**: 537–551.
5. Anonymous. (1959). Symposium on the classification of brackish waters, Venice. *Arch. Oceanogr. Limnol.* **2** (suppl.): 1.
6. D'Avanzo, C., Kremer, J.N., and Wainright, S.C. (1996). Ecosystem production and respiration in response to eutrophication in shallow temperate estuaries. *Mar. Ecol. Prog. Ser.* **141**: 263–274.
7. Viaroli, P., Pugnetti, A., and Ferrari, I. (1992). *Ulva rigida* and decomposition processes and related effects on nitrogen and phosphorus cycles in a coastal lagoon (Sacca di Goro, Po River Delta). In: Marine eutrophication and population dynamics. *Proc. 25th Eur. Mar. Biol. Symp.* G. Colombo et al. (Eds.). Ferrara, Italy, pp. 77–84.
8. Ziegler, S. and Benner, R. (1998). Ecosystem metabolism in a subtropical, seagrass-dominated lagoon. *Mar. Ecol. Prog. Ser.* **173**: 1–12.
9. Piriou, J-Y. and Ménesguen, A. (1992). Environmental factors controlling the *Ulva* sp. bloom in Brittany (France). In: Marine eutrophication and population dynamics. *Proc. 25th Eur. Mar. Biol. Symp.* G. Colombo et al. (Eds.). Ferrara, Italy, pp. 111–115.
10. Diaz, R.J. and Rosenberg, R. (1995). Marine benthic hypoxia: A review of its ecological effects and the behavioural responses of benthic macrofauna. *Oceanogr. Mar. Biol. Annu. Rev.* **33**: 245–303.
11. Justic, D., Rabalais, N.N., and Turner, T. (1995). Stoichiometric nutrient balance and origin of coastal eutrophication. *Mar. Poll. Bull.* **30**: 41–46.
12. Valiela, I., McClelland, J., Hauxwell, J., Beher, P.J., Hersh, D., and Foreman, K. (1997). Macroalgal blooms in shallow estuaries: Controls and ecophysiological and ecosystem consequences. *Limnol. Oceanogr.* **5**: 1105–1118.
13. Schubel, R.J. and Pritchard, D.W. (1986). Responses of upper Chesapeake Bay to variations in discharge of the Susquehanna River. *Estuaries* **9**: 236–249.
14. Page, H.M., Petty, R.L., and Meade, D.E. (1995). Influence of watershed runoff on nutrient dynamics in a southern California salt marsh. *Estuar. Coast. Shelf Sci.* **41**: 163–180.
15. Caffrey, J.N. and Day, J.W., Jr. (1986). Control of the variability of nutrients and suspended sediments in a Gulf

Coast estuary by climatic forcing and spring discharge of the Atchafalaya River. *Estuaries* **9**: 295–300.

16. Balls, P.W. (1994). Nutrient inputs to estuaries from nine Scottish east coast rivers; influence of estuarine processes on inputs to the North Sea. *Estuar. Coast. Shelf Sci.* **39**: 329–352.
17. Uncles, R.J. and Stephens, J.A. (1996). Salt intrusion in the Tweed Estuary. *Estuar. Coast. Shelf Sci.* **43**: 271–293.
18. Magni, P. and Montani, S. (2004). Tidally-mediated changes of nutrient concentrations. *Encyclopedia of Water (OC-136)*. John Wiley & Sons, New York.
19. Yin, K., Harrison, P.J., Pond, S., and Beamish, R.J. (1995). Entrainment of nitrate in the Fraser River estuary and its biological implications. II. Effects of spring *vs.* neap tide and river discharge. *Estuar. Coast. Shelf Sci.* **40**: 529–544.
20. Yin, K., Harrison, P.J., Pond, S., and Beamish, R.J. (1995). Entrainment of nitrate in the Fraser River estuary and its biological implications. III. Effects of winds. *Estuar. Coast. Shelf Sci.* **40**: 545–558.
21. Balls, P.W. (1992). Nutrient behaviour in two contrasting Scottish estuaries, the Forth and the Tay. *Oceanol. Acta* **15**: 261–277.
22. Cowan, J.L.W., Pennock, J.R., and Boyton, W.R. (1996). Seasonal and interannual patterns of sediment–water nutrient fluxes in Mobile Bay, Alabama (USA): Regulating factors and ecological significance. *Mar. Ecol. Progr. Ser.* **141**: 229–245.
23. MacIntyre, H.L. and Cullen, J.J. (1996). Primary production by suspended and benthic microalgae in a turbid estuary: Time-scale of variability in San Antonio Bay, Texas. *Mar. Ecol. Progr. Ser.* **145**: 245–268.
24. Yamada, H. and Kayama, M. (1987). Liberation of nitrogenous compounds from bottom sediments and effect of bioturbation by small bivalve, *Theora lata* (Hinds). *Estuar. Coast. Shelf Sci.* **24**: 539–555.
25. Nakamura, Y. (1994). Effect of flow velocity on phosphate release from sediment. *Water Sci. Technol.* **30**: 263–272.
26. Miller-Way, T. and Twilley, R.R. (1996). Theory and operation of continuous flow systems for the study of benthic–pelagic coupling. *Mar. Ecol. Progr. Ser.* **140**: 257–269.
27. Matsukawa, Y., Sato, Y., and Sasaki, K. (1987). Benthic flux of nutrient salts on an intertidal flat. *Nippon Suisan Gakkaishi* **53**: 985–989.
28. Sundbäck, K., Enoksson, V., Granéli, W., and Pettersson, K. (1991). Influence of sublittoral microphytobenthos on the oxygen and nutrient flux between sediment and water: A laboratory continuous-flow study. *Mar. Ecol. Prog. Ser.* **74**: 263–279.
29. Feuillet-Girard, M., Gouleau, D., Blanchard, G., and Joassard, L. (1997). Nutrient fluxes on an intertidal mudflat in Marennes-Oléron Bay, and influence of the emersion period. *Aquat. Living Resour.* **10**: 49–58.
30. Trimmer, M., Nedwell, D.B., Sivyer, D.B., and Malcolm, S.J. (1998). Nitrogen fluxes through the lower estuary of the river Great Ouse, England: The role of the bottom sediments. *Mar. Ecol. Prog. Ser.* **163**: 109–124.
31. Usui, T., Koike, I., and Ogura, N. (1998). Tidal effect on dynamics of pore water nitrate in intertidal sediment of a eutrophic estuary. *J. Oceanogr.* **54**: 205–216.
32. Oldham, C.E. and Lavery, P.S. (1999). Porewater nutrient fluxes in a shallow fetch-limited estuary. *Mar. Ecol. Prog. Ser.* **183**: 39–47.
33. Takayanagi, K. and Yamada, H. (1999). Effects of benthic flux on short term variations of nutrients in Aburatsubo Bay. *J. Oceanogr.* **55**: 463–469.
34. Barbanti, A., Ceccherelli, V.U., Frascari, F., Reggiani, G., and Rosso, G. (1992). Nutrient regeneration processes in bottom sediments in a Po delta lagoon (Italy) and the role of bioturbation in determining the fluxes at the sediment–water interface. *Hydrobiologia* **228**: 1–21.
35. Rysgaard, S., Christensen, P.B., and Nielsen, L.P. (1995). Seasonal variation in nitrification and denitrification in estuarine sediment colonized by benthic microalgae and bioturbating infauna. *Mar. Ecol. Prog. Ser.* **126**: 111–121.
36. Sundbäck, K., Miles, A., and Göransson, E. (2000). Nitrogen fluxes, denitrification and the role of microphytobenthos in microtidal shallow-water sediments: An annual study. *Mar. Ecol. Prog. Ser.* **200**: 59–76.

BLACK WATER TURNS THE TIDE ON FLORIDA CORAL

NASA—Goddard Space Flight Center

In early 2002, a patch of "black water" spanning over 60 miles in diameter formed off southwestern Florida and contributed to severe coral reef stress and death in the Florida Keys, according to results published from research funded by NASA, the U.S. Environmental Protection Agency and the National Oceanic and Atmospheric Administration (NOAA). The "black water" contained a high abundance of toxic and non-toxic microscopic plants.

Chuanmin Hu and other colleagues at the Institute for Marine Remote Sensing of the University of South Florida (USF), St. Petersburg, Fla., and colleagues from the Florida Fish and Wildlife Conservation Commission (FFWCC) and the University of Georgia, co-authored an article on this phenomenon that appeared as the cover

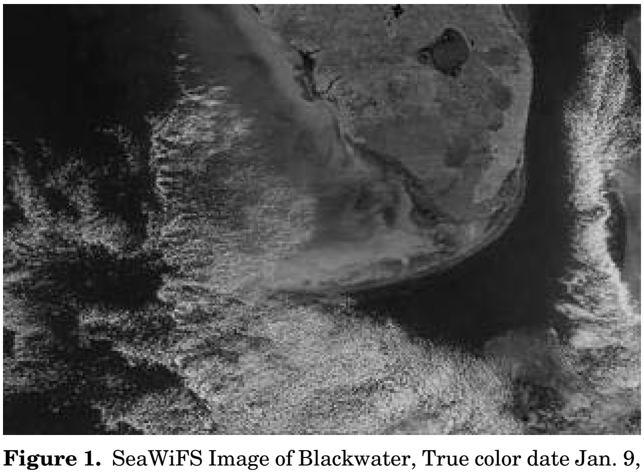

Figure 1. SeaWiFS Image of Blackwater, True color date Jan. 9, 2002. Courtesy The SeaWiFs Project and ORBIMAGE, Scientific Visualization Studio.

---

story of a recent issue of the American Geophysical Union's Geophysical Research Letters.

"The water appeared black in satellite imagery because the concentration of the microscopic plants and other dissolved matters were high," Hu said. Because plants and dissolved matter absorb sunlight, they reduce the amount of light normally reflected from the ocean. When a red-tide bloom occurs the water takes on various hues of red or brown. While not all microscopic plants contribute to red tides, the darker hue created by both the plankton and the harmful algal blooms made the water appear black when seen from the satellite.

When Hu and his colleagues examined the data collected by divers from the dark water area in the Florida Keys, they discovered a 70 percent decrease in stony coral cover, a 40 percent reduction of coral species, and a near-elimination of sponge colonies at two reef sites after the dark water passed. By examining satellite images and field survey data, the authors concluded that the coral reef ecosystem was stressed by microscopic organisms and toxins contained in the dark water.

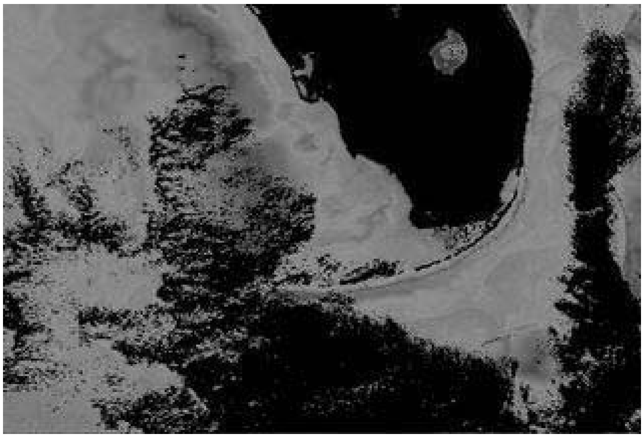

Figure 2. SeaWiFS Image of Blackwater, date Jan. 9, 2002 False color visualization showing chlorophyll. Reds and oranges represent high concentrations of chlorophyll. Courtesy The SeaWiFs Project and ORBIMAGE, Scientific Visualization Studio.

Figure 3. Microscopic image of a dinoflagellate, called Rhizosoleniacea, a cause for Red Tide. Rhizosoleniaceae blooms occur seasonally off Florida's Gulf coast due to freshwater outflow from the Everglades. Courtesy Florida Marine Research Institute.

The "black water" event caused alarm among local fishermen, divers, and the public, as the color of the water was unusual and fish seemed to avoid this large area of dark water. Satellite instruments such as the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) aboard Orbimage's SeaStar satellite and the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard NASA's Terra and Aqua satellites provide information on ocean color that allows scientists to monitor the health of the water and the shallow benthic (ocean bottom) environment. The SeaWiFS and MODIS measurements of the dark water led to a number of investigations to help clarify the issues and to provide answers to the public's concerns.

During January 2002, SeaWiFS detected the dark-colored water in the Florida Bight, just southwest of the Everglades. In fall 2001, the SeaWiFS images showed an extensive red tide off Florida's central west coast, near Charlotte Harbor.

Red tides occur every year off Florida and are known to cause fish kills, coral stress and mortality, and skin and respiratory problems in humans. They are caused by high concentration of microscopic plants called dinoflagellates. Other microorganisms called cyanobacteria can also cause harmful algal blooms. The waters containing this red tide migrated to the south along the coast.

Winter storms caused large amounts of fresh water to drain from the Everglades into Florida Bight (the curve in

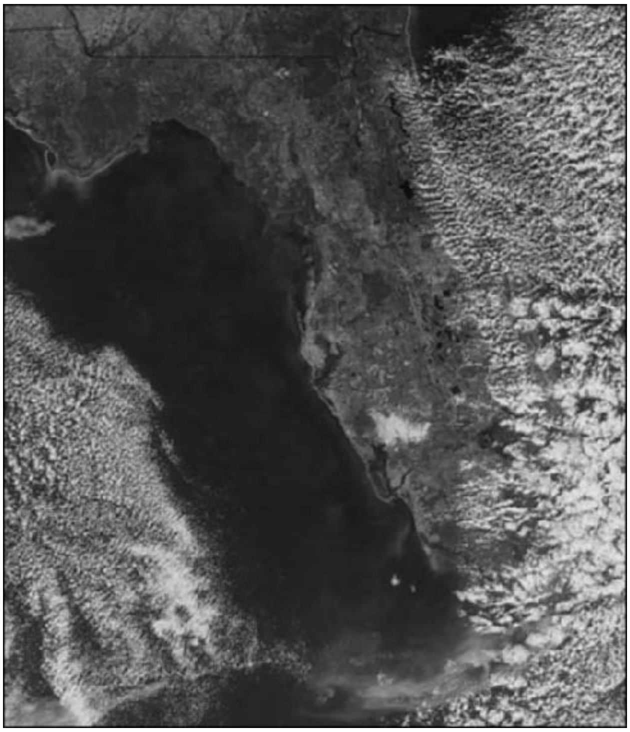

Figure 4. NASA's Terra satellite acquired this image of a red tide bloom along Florida's west coast December 22, 2001. Notice the dark reddish color of the ocean. Image courtesy Jacques Descloitres, MODIS Land Rapid Response Team at NASA GSFC.

the shoreline from the Keys north to Everglades National Park on the mainland), carrying high levels of nutrients such as silicate, phosphorus, and nitrogen to the sea. These caused a bloom of the microscopic marine plants known as diatoms in the same patch. The bloom turned the water dark and the "black water" patch re-circulated for several months in a slow clockwise motion off southwest Florida in the Florida Bight. Slowly, the dark water drifted farther south and toward the Florida Keys. By May 2002, the "black water" had moved through passages in the Florida Keys, dispersing into the Atlantic and the Gulf Stream.

Co-authors on this research article included Serge Andrefouet and Frank E. Muller-Karger of USF; Keith E. Hackett, Michael K. Callahan, and Jennifer L. Wheaton of FFWCC, St. Petersburg, Fla.; and James W. Porter of the University of Georgia, Athens, Ga.

NASA funded part of this research as part of its Earth Science mission to understand and protect our home planet. NASA's Earth Science Enterprise is dedicated to understanding the Earth as an integrated system and applying Earth System Science to improve prediction of climate, weather, and natural hazards using the unique vantage point of space.

SHALLOW WATER WAVES

Stefano Pagliara
Pietro Chiavaccini
Universitá di Pisa
Pisa, Italy

INTRODUCTION

Shallow water conditions happens when the ratio of water depth d to the wave length L is less than 1/25, i.e., $d/L < 0.04$. In this condition, there is a combination of high wave steepness, and ratios of wave height to water depth are close to unity. Shallow water waves approaching the coast are subjected to refraction, shoaling, and breaking.

THE SHALLOW WATER THEORIES

The Linear Theory

The linear wave theory (1), in which the nonlinear terms of the equation of motion are omitted, describes water motion by two nondimensional parameters, the relative depth d/L and the wave steepness H/L, which represent respectively the influence of the depth on wave characteristics (celerity, length, and height) and the measure of the wave dimension in respect of height. For large values of relative depth, the linear theory cannot be applied.

Wave profile in linear theory is defined by

$$\eta = a\cos(kx - \omega t) = a\cos\left(\frac{2\pi x}{L} - \frac{2\pi x}{T}\right) \tag{1}$$

The celerity is

$$C = \frac{L}{T} = \left(\frac{gT}{2\pi}\right)\tan h\left(\frac{2\pi d}{L}\right) = \sqrt{\left(\frac{gL}{2\pi}\right)\tan h\left(\frac{2\pi d}{L}\right)} \tag{2}$$

In shallow water, as $\frac{d}{L} \to 0 \tan h\left(\frac{2\pi d}{L}\right) \to \frac{2\pi d}{L}$ and celerity becomes dependent only on depth

$$C = \sqrt{gd} \tag{3}$$

Wave train propagates with a speed different from the celerity, called group velocity, defined by

$$C_G = nC = \frac{1}{2}\left[1 + \frac{2\pi d}{L}\frac{1}{\operatorname{senh}\left(\frac{2\pi d}{L}\right)}\right]C \tag{4}$$

In shallow water, the group velocity becomes $C_G = C$ because $n \to 1$, so wave train propagates with the same speed of the single waves of the train.

In reality, a wave approaching the coast is modified by the bottom: its height grows and loses characteristics of linearity, which implies a more detailed formulation.

The nonlinear theories, used more to describe shallow water waves, are the cnoidal wave theory and the solitary wave theory.

The Cnoidal Wave Theory

The theory of cnoidal waves was first developed by Korteweg and de Vries (2) and can be used to described unidirectional waves in shallow water, where $d/L < 1/8$ and $Ur > 20$. If the water height is small, the cnoidal wave profile becomes the sinusoidal one.

The cnoidal wave is periodic with the profile given by

$$\eta(x,t) = hCn^2\left[2K(\kappa)\left(\frac{x}{L} - \frac{t}{T}\right), \kappa\right] \tag{5}$$

where $K(k)$ is the complete elliptic integral of the first kind of modulus k, $Cn(r)$ is the Jacobian elliptic function, and h is the vertical coordinate of the water surface about the trough level at the horizontal coordinate. To simplify the use of *cnoidal* theory, Wiegel (3) developed graphs and table of easy appliance.

The Solitary Wave Theory

The solitary wave is a progressive wave consisting of a single crest and is not oscillatory as the other examined. The wave form is (4)

$$\eta(x,t) = H\sec h^2\left(\sqrt{\frac{3H}{4d^3}}(x - Ct)\right) \tag{6}$$

where C (wave celerity) is defined by

$$C = \sqrt{gd}\left(1 + \frac{H}{2d}\right) \tag{7}$$

According to the dispersion relation of the linear wave theory, the celerity of water waves in shallow water (Eq. 3) is smaller than solitary wave phase velocity because of the inclusion of terms that depend on H/d.

The solitary wave describes enough well waves approaching shallow water, even if wave period or wave length are not associated with the theory. When the wave pass from deep to shallow water their crests peak up and are separated by flat troughs appearing like a series of solitary waves. Solitary theory appears reasonable even if the periodicity is neglected, because in shallow water the period is not particularly significant, but rather the water depth becomes important.

As the solitary wave advances into shallow water, its height increases, the crests becomes greater and sharper, the trough becomes longer and flatter: in this condition, a wave is well represented by solitary waves. The solitary wave is not an oscillatory wave, as those obtained with other theories, but a translation one. Water particles, as a wave passes, are subjected to a translation in wave direction, whereas in the oscillatory waves it moves forward and backward, remaining after a period in its original position.

Wave profiles, according to the theories illustrated above, are presented in Fig. 1.

WAVE REFRACTION AND SHOALING

Waves passing from deep to shallow water are subjected to refraction in which the direction of their travel changes in such a way that approaching the coast the crests tends to become more parallel to the depth contours (Fig. 2). To determine the variation of the wave direction, Snell's law can be applied

$$\frac{\sin\theta_1}{C_1} = \frac{\sin\theta_2}{C_2} = \text{cons} \qquad (8)$$

where q_1 and q_2 are the angles between adjacent wave crests and the respective bottom contours, whereas C_1 and C_2 are the wave celerity at the two depths.

With a regular bottom (straight and parallel contours), the relation can be applied directly between the angle at any depth and the deep water angle approach.

With an irregular bottom, wave refraction may cause a spreading or a convergence of the wave energy. This effect can be easily illustrated taking in account the wave rays (Fig. 2), defined as the lines drawn normal to the wave crests and directed in the wave advance. If the wave rays spread, the wave crests become longer and the energy flux, assumed constant between two rays, must be extended over a greater length. The opposite (energy concentration) happens if the waves rays converge.

Actually, the calculations of the rays are made by software based on models such as a mild slope, parabolic, or Boussinesq wave model.

Another effect of the change in the wave length in shallow water is that the wave height increases. This effect is the consequence of the energy conservation in

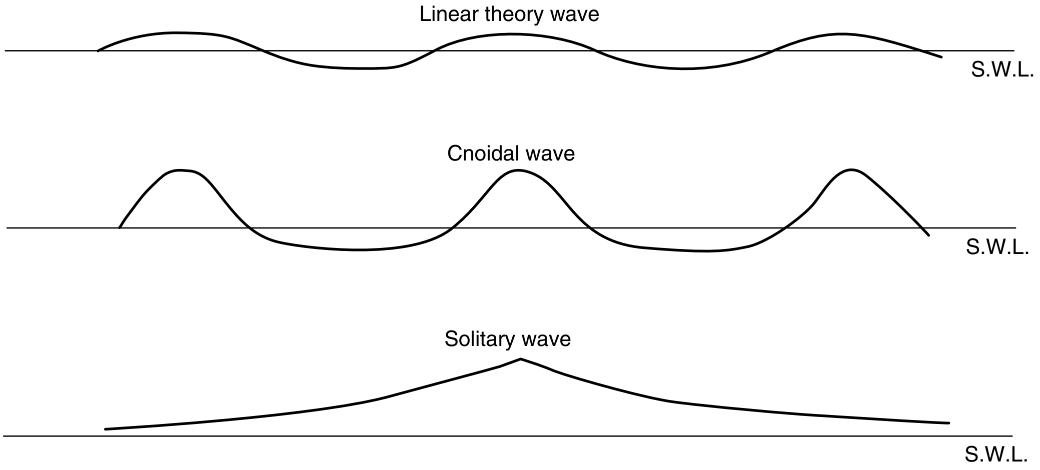

Figure 1. Wave profiles.

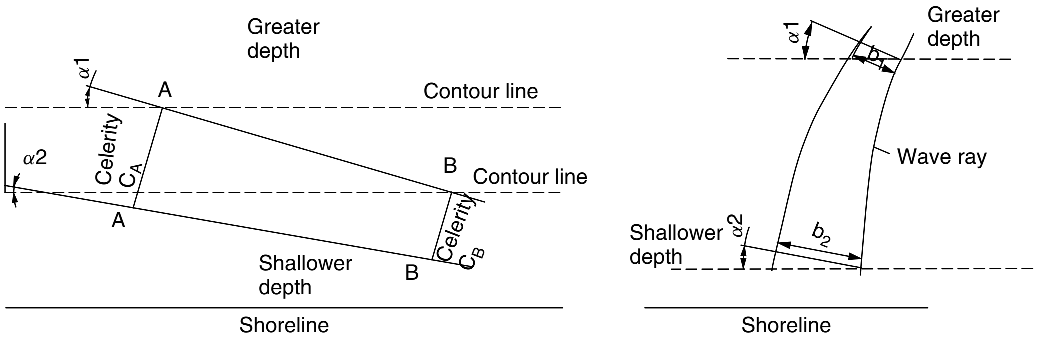

Figure 2. Wave refraction and wave rays.

concert with the decrease of the celerity approaching the shallow water. This phenomenon is referred to as shoaling.

The effects of shoaling and refraction in water can be expressed by the following formula:

$$H = H_0 K_s K_r \tag{9}$$

where H_0 is wave height in deep water, K_s is the shoaling coefficient

$$K_s = \sqrt{\frac{C_{G0}}{C_G}} = \sqrt{\frac{1}{2n \tan h \dfrac{2\pi d}{L}}} \tag{10}$$

with $n = \frac{1}{2}\left(1 + \frac{2\pi d}{L}\frac{1}{\operatorname{senh}\left(\frac{2\pi d}{L}\right)}\right)$ and K_r is the refraction coefficient

$$K_r = \sqrt{\frac{b_0}{b}} \tag{11}$$

where b_0 and b are the distance between two adjacent rays, respectively, in deep water and at a generic depth. For straight and parallel contours lines, the refraction coefficient becomes

$$K_r = \sqrt{\frac{\cos\theta_0}{\cos\theta}} \tag{12}$$

WAVE BREAKING

Waves shoaling causes the increasing of wave height until its physical limit because of steepness of wave H/L. When this limit is reached, the wave breaks and dissipates energy. Battjes (5) has shown that the breaking wave characteristics can be correlated to a parameter, called surf similarity x, which is defined as

$$\xi = \frac{\tan\beta}{\sqrt{\dfrac{H_0}{L_0}}} \tag{13}$$

where $\tan b$ is the beach slope and H_0 and L_0 are the wave height and length in deep water, respectively.

Three common type of breakers are recognized (Fig. 3):

- spilling ($x < 0.5$), in which each wave gradually peaks until the wave becomes unstable and cascades down as white water, bubbles, and foam;
- plunging ($0.5 < x < 3.3$), in which the shoreward face of the wave becomes vertical, curls over, and plunges forward and downward as an intact mass of water;
- surging ($3.3 < x < 5$), in which the base of the wave, while it is peaking up, finds the shore, and then the crest collapses and disappears.

When the surf similarity x is >5, reflection happens and no breaking occurs.

According to Galvin (6), a fourth breaker, called collapsing, intermediate between plunging and surging types, exists.

It is difficult to identify which type of breaker can verify, because it depends on the individual heights and interactions of the waves. Plunging and surging breakers can be seen on a beach during the same storm. However, spilling breakers are typical of very low sloping beaches with waves of high steepness values; plunging waves occur in steeper beaches and waves of intermediate steepness; surging, instead, is associated with high gradient beaches with waves of low steepness.

The breaking happens when the water particle velocity is greater than wave celerity. According to solitary wave theory, this condition is described by (7)

$$\left(\frac{H_b}{d_b}\right)_{\max} = 0.78 \tag{14}$$

where the subscript b denotes the breaking.

Laboratory tests pointed out that Equation (14) is verified more for oscillatory waves than solitary wave. However, it is considered fundamental to express the relation between the relative depth and the breaking condition.

Other parameters playing a role in wave breaking exist, such as beach slope and bottom roughness. An empirical relation considers the beach slope m (8):

$$\left(\frac{H_b}{d_b}\right)_{\max} = 0.75 + 25m - 112m^2 + 3870m^3 \tag{15}$$

in which as the slope increases, the breaking happens more and more nearshore.

Another expression was developed by Goda (9)

$$\left(\frac{H_b}{d_b}\right)_{\max} = 0.17\frac{L_0}{d_b}\left[1 - \exp\left(-\frac{15\pi d_b}{L_0}(1 + 15m^{4/3})\right)\right] \tag{16}$$

where L_0 is the deep water wave lenght.

For irregular waves (waves with different height and period), Kamphuis (10) proposed two criteria based on

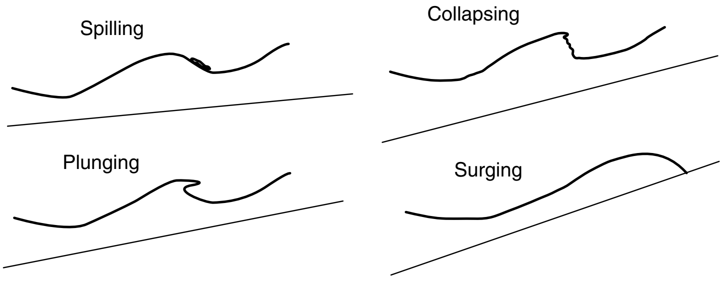

Figure 3. Breaking types.

extensive model tests

$$H_{sb} = 0.095e^{4m}L_{bp}\tan h\left(\frac{2\pi d_b}{L_{bp}}\right) \tag{17}$$

$$\frac{H_{sb}}{d_b} = 0.56e^{3.5m} \tag{18}$$

where H_{sb} is the significant wave breaking, L_{bp} the breaking wave length, and d_{b} is the breaking depth.

BIBLIOGRAPHY

1. Airy, G.B. (1845). Tides and waves. *Encyclopedia Metropolitana*. **192**: 241–396.
2. Korteweg, D.J. and de Vries, G. (1895). On the change of form of long waves advancing in a rectangular canal and on a type of long stationary wave. *Phil. Mag.* **5**: 442–443.
3. Wiegel. (1960).
4. Russell, J.S. (1884). *Report on waves. 14th Meeting of the British Association for the Advancement of Science*. pp. 311–390.
5. Battjes, J.A. (1974). Surf similarity. *Proceedings of the 14th Coastal Engineering Conference, ASCE*. pp. 466–497.
6. Galvin, C.J. (1968). Breaker type classification on three laboratory beaches. *J. Geophys. Res.* **73**: 3651–3659.
7. Munk, W.H. (1949). The solitary wave theory and its application to surfs problems. *Ann. New York Acad. Sci.* **51**: 376–424.
8. SPM. (1984).
9. Goda, Y. (1970). A synthesis of breaker indices. *Trans. Japan Soc. Civil Engrs.* **2**: 227–230.
10. Kamphuis, J.W. (1991). Wave transformation. *Co. Eng.* **15**: 173–184.

FURTHER READING

CERC U.S. Corps of Eng. (1984). *Shore Protection Manual*. Vicksburg, VA.

WATER WAVES

NAI KUANG LIANG
National Taiwan University
Taipei, Taiwan

INTRODUCTION

As one stands at the coast, an endlessly moving succession of irregular humps and hollows can be seen reaching to the shore. This is the water wave generated by wind. There are different kinds of water waves, which are driven by different forces. Beside wind waves, the tide and the tsunami are other well-known water waves. Different Water waves are distinguished by their wavelength, which is defined as the length between successive humps. The wavelength of a wind wave is shorter and is easily recognized.

A water wave is an important physical phenomenon in an ocean, sea, or lake. It influences beach morphology, maritime structures, and human activities very much. Because the wind varies tremendously, a typhoon or hurricane may generate huge waves. This storm usually causes disasters. Tidal and wind waves contain a lot of energy. Electric power can be generated by them.

BASIC CHARACTERISTICS OF A WATER WAVE

The water wave was studied mathematically in the nineteenth century as a form of oscillatory wave. In 1847, Stokes (1) published his famous paper entitled "On the theory of oscillatory waves." Stokes' wave theory has been widely used till now. The elements of an oscillatory wave are wave height, wave period, and wavelength. The wave height H is the distance between the wave crest and trough. The wavelength L is the distance between successive crests, and the wave period T the time difference between successive crests. In deep water, the water depth H is larger than the half of the wavelength L. After Stokes, L is a function of the wave period T in the following equation:

$$L = \frac{gT^2}{2\pi} \tag{1}$$

The wave celerity C is then

$$C = \frac{L}{T} = \frac{gT}{2\pi} \tag{2}$$

in which g is the earth's gravitational acceleration. Obviously, the celerity C is proportional to the wave period T. The water depth is less than $L/2$, so then the wave particle movement touches the bottom. Tides and tsunamis usually are shallow water waves or long waves. Equations 1 and 2 become the following:

$$L = \frac{gT^2}{2\pi}\tanh\left(\frac{2\pi h}{L}\right) \tag{3}$$

$$C = \frac{gT}{2\pi}\tanh\left(\frac{2\pi h}{L}\right) \tag{4}$$

A wave contains kinetic and potential energy. The average energy E per unit sea surface area is the following:

$$E = \frac{1}{8}\rho g H^2 \tag{5}$$

The wave energy propagates in a group velocity, Cg, which is given by the following equation (2):

$$Cg = \frac{C}{2}\left(1 + \frac{4\pi h/L}{\sinh 4\pi h/L}\right) \tag{6}$$

The wave energy flux P is as follows:

$$P = E \cdot Cg \tag{7}$$

A real ocean wave is not oscillatory, as Stokes' theory described, but irregular in a stochastic process. Pierson (3,4) introduced the technique of communication engineering to the ocean wave and proposed the random wave theory. The ocean wave would be a superposition of sinusoidal wave components of different directions, amplitudes, frequencies, and angular phases, in which the phase is a random variable of equal probability density between—π and π. Then the ocean wave can be represented by a power spectrum. However, for convenience, engineers usually use the significant wave to represent the ocean wind wave. For a group of N wave heights measured at a point, waves are ordered from the largest to the smallest and assigned a number from 1 to N to them. The significant wave height $H_{1/3}$ is defined as the average of the first (highest) $N/3$ wave heights (5). The order of the wave period is accompanied by its wave height as a pair. The significant wave period $T_{1/3}$ is defined as the average of the first $N/3$ wave periods.

WIND WAVE FORECASTING

During the Second World War, two famous American oceanographers, Dr. H.U. Sverdrup and Dr. W.H. Munk, were assigned by the U.S. Navy to develop a wind wave forecasting scheme for the Normandy landing operations. The work was originally completed in 1943 and classified and published in 1947(6). Later, the scheme was extensively patched and amended by Bretschneider(7). Therefore, the scheme has been named the SMB method. In the SMB method, the wind speed, wind duration, and fetch are the main parameters, where fetch is defined as the wind blowing distance in the water area.

Modern ocean wave modeling was initiated in 1956 and extensively developed and revised (8). So far it is still developing. The basic concept is the evolution of energy spectrum F governed by the energy balance equation:

$$\frac{df}{dt} = S_{\mathrm{in}} + S_{\mathrm{nl}} + S_{\mathrm{ds}} \tag{8}$$

in which S_{in} is the energy input flux from wind to wave spectrum components, S_{nl} the energy flux exchange due to nonlinear wave–wave interaction, and S_{ds} the energy flux output due to dissipation. The present operating ocean models are WOM, NWW3, etc.

TYPHOON WAVE

The typhoon or hurricane is an atmospheric eddy that originates in tropical or subtropical ocean regions. The typhoon wind speed is usually very high so that the typhoon wave plays a significant role in the design of coastal and offshore structures. Because the wind velocity changes rapidly, the generation process within typhoons is complicated. Parametric typhoon or hurricane wave prediction models were developed (9,10), which pointed out that the maximum wave exists at the right side of the typhoon center as one faces the forward direction.

The radius of a typhoon is about 200–400 km. When a typhoon is still far away from a location, the swell may arrive because the swell energy propagates usually faster than the typhoon moves. The typhoon can be regarded as a point source of wave generation. If Person A, who does not move, throws balls to Person B at a fixed time interval and the speed of the ball relative to the ground is constant, then Person B receives balls at the same time interval. If Person A moves toward Person B, Person B receives balls at a shorter time interval. If the ball is like energy, Person A throws more energy flux to Person B, as Person A moves toward Person B. This is the same as the well-known Doppler effect. As a whistling train moves toward an observer, the sound heard by the observer is higher in frequency than the actual whistle at the source. The observer receives more sound energy flux. As shown in Equation 7, wave energy flux P is wave energy $E \times$ group velocity Cg, which is proportional to the wave period. The wave energy is proportional to H^2 (Equation 5). Because the typhoon wave moving speed has little effect on the swell period, then the swell height is enhanced, as the typhoon approaches a station (11). There were cases where swell heights were larger than those inside the typhoon because the moving velocity was close to the group velocity of the swell.

BIBLIOGRAPHY

1. Stokes, G.G. (1847). On the theory of oscillatory waves. *Trans. Cambridge Philos. Soc.* **8**: 441.
2. Kinsman, B. (1965). *Wind Waves: The Generation and Propagation on the Ocean Surface*. Prentice-Hall, Englewood Cliffs, NJ.
3. Pierson, W.J., Jr. (1952). *A Unified Mathematical Theory for the Analysis, Propagation and Refraction of Storm Generated Ocean Surface Waves, Parts I and II*. New York University, College of Engineering, Research Division, Department of Meteorology and Oceanography prepared for the Beach Erosion Board, Department of the Army, and Office of Naval Research, Department of the Navy.
4. Pierson, W.J., Jr. (1955). Wind-generated gravity waves. In: *Advances in Geophysics*. Vol. 2. Academic Press, New York, pp. 93–178.
5. Dean, R.G. and Dalrymple, R.A. (1984). *Water Wave Mechanics for Engineers and Scientists*. Prentice-Hall, Englewood Cliffs, NJ.
6. Sverdrup, H.U. and Munk, W.H. (1947). *Wind, Sea and Swell: Theory of Relations for Forecasting*. U. S. Navy Hydrographic Office Pub. No. 601.
7. Bretschneider, C.L. (1952). The generation and decay of wind waves in deep water. *Trans. Am. Geophys. University* **33**(3): 381–389.
8. Komen, G.J., Cavaleri, L., Donelan, M., Hasselmann, K., Hasselmann, S., and Janssen, P.A.E.M. (1994). *Dynamics and Modeling of Ocean Waves*. Cambridge University Press.
9. Bretschneider, C.L. and Tamaye, E.E. (1976). Hurricane wind and wave forecasting techniques. *Proc. 15th Int. Coastal Eng. Conf.*, Honolulu, HI, pp. 202–237.
10. Young, I.R. (1988). Parametric hurricane wave prediction model. *J. Waterway Port Coastal Ocean Eng.* **114**(5): 637–652.
11. Liang, N.K. (2003). The typhoon swell Doppler effect. *Ocean Eng.* **30**: 1107–1115.

WOODS HOLE: THE EARLY YEARS

Northeast Fisheries Science
Center—NOAA

The beginning of Woods Hole dates back to the early 17th century. Five years before the settlement of Jamestown,

---

Virginia, and 18 years before the Pilgrims landed at Provincetown and Plymouth, Bartholomew Gosnold coasted along Cape Cod and Marthas Vineyard, and about May 31, 1602, he is believed to have landed at what is now known as Woods Hole. The Town of Falmouth, of which Woods Hole is presently a part, was first settled in 1659–61 when several persons were granted permission to purchase land. The date of the settlement of Woods Hole took place 17 years later. The town (Falmouth) was incorporated on June 4, 1686, and called Succonessett, the name which later, probably in 1694, was changed to Falmouth. On July 23, 1677, the land around Little Harbor of Woods Hole was divided among the 13 settlers in "lots of 60 acres upland to a share" and an "Indian deed" confirming the land title was signed by Job Notantico on July 15, 1679. Fishing, hunting, and sheep breeding were the principal occupations of the early settlers and their descendents. Later on a grist mill was built and salt was made by solar evaporation of sea water in pans built along the banks of Little Harbor.

These quiet, rural conditions, devoid of adventure, persisted until about 1815, when Woods Hole became an important whaling station from which ships operated on the high seas. The whaling industry in the United States became a very profitable business, and Woods Hole was a part of it. In 1854, the total receipts for the American whaling fleet amounted to $10.8 million, the largest part of this amount resulted from whaling carried out by Massachusetts captains. Woods Hole participated in these activities and prospered. It is known that between 1815 and 1860, not less than nine whaling ships were making port at the Bar Neck wharf, which was located where the U.S. Navy building of the Woods Hole Oceanographic Institution now stands. The place was busy processing oil and whalebone and outfitting ships. A bake house for making sea biscuits for long voyages stood next to the present "Old Stone Building" built in 1829 as a candle factory. This conspicuous old landmark on Water Street of Woods Hole, identified by an appropriate bronze plaque, now serves as a warehouse for the Marine Biological Laboratory for storing preserved zoological specimens. About 1860, whaling became less profitable and Woods Hole entered into the second phase of its economic life which was dominated by the establishment and operation of a new commercial venture known as the Pacific Guano Works.

During the years from 1863 to 1889, when the *Pacific Guano Works* was in operation the life of Woods Hole centered around the plant which was built at Long Neck near the entrance to what is known now as Penzance Point. Many large sailing vessels carrying sulphur from Italy, nitrate of soda from Chile, potash from Germany, and many schooners under the American flag loaded with guano and phosphorus from the Pacific Coast of South America were anchored in Great Harbor waiting for their turn to unload their cargoes. The number of laborers regularly employed by the Guano Company varied from 150 to 200 men, mostly Irishmen brought in under contract. Several local fishermen found additional employment as pilots for guano ships. The company maintained a store where various goods such as leather, lead pipe, tin, coal, wood, and other items were bought and sold. The store acted also as a labor housing agency. Through efforts of the business manager of the Guano Company, the Old Colony Railroad was persuaded to extend its branch from Monument Beach to Woods Hole. The establishment of well-organized and reliable transportation to Boston was an important factor in the future life of the community.

The Pacific Guano Works was established by the shipping merchants of Boston who were seeking cargo for the return voyage of their ships. The guano deposits of one of the Pacific islands seemed to furnish this opportunity. As soon as the joint stock company was organized in 1859 with the capital of $1 million, arrangements were made almost immediately by which the newly formed concern came into possession and control of Howland Island. This island is located in the middle of the Pacific Ocean at longitude 177 deg. W., a short distance north of the Equator, about 1, 500 miles true south from Midway Island of the Hawaiian archipelago. At the same time appropriate plant and docking facilities were built at Woods Hole and 33 large sailing ships became available for hauling guano. Unlike the well-known guano islands off the coast of Peru, Howland Island is located in the zone of abundant rainfall. Consequently, the guano deposits of the island were leached of organic components and consisted of highly concentrated phosphate of lime.

Fertilizer produced by the company was made by restoring the lost organic matter of the phosphate rock

1887 Map of Woods Holl (Fisheries is bottom left) by adding the right proportion of organic constituents which were obtained from menhaden, pogy, and other industrial fish which abound in Cape Cod waters. The rock was pulverized and purified by washing; fish brought in by local fishermen were first pressed to extract oil, and the residue digested with sulphuric acid, washed, and dried. Acid was produced locally from sulphur imported from Sicily, and the digestion of fish flesh was carried out in large lead-lined vats. The plant was well equipped with machinery needed for the process and even had a chemical laboratory where chemists made the necessary analyses. Various sheds for storage and drying, barracks for laborers, and a business office completed the facilities.

When the deposits of phosphate rock on Howland Island were exhausted, the company acquired title to the Greater and Lesser Swan Islands from the U.S. Government. These islands are located in the Caribbean Sea at latitude 17 deg. N. and longitude 83 deg. W. off the coast of Honduras. The islands are only 400 miles from Key West, Florida, and 500 miles from New Orleans. They contained good-quality phosphate rock and being much closer to Woods Hole greatly reduced the voyage time and cost of delivery. Further expansion of the company consisted in the acquisition of Chisolm's Island near the coast of South Carolina, construction of a plant for cracking and washing phosphate rock on the Ball River side of the island, and establishment of a processing plant in Charleston, S.C. From the initial production (in 1865) of 7, 540 sacks of fertilizer weighing 200 pounds each, the output reached 11, 420 tons in 1871 and continued to grow until the combined annual production in 1879 of the works at Woods Hole and Charleston reached from 40, 000 to 45, 000 tons of guano fertilizer.

Spencer Baird, Secretary of the Smithsonian Instution and first commissioner of the U.S. Commission of Fish and Fisheries arrived in Woods Hole in 1871. Baird was greatly impressed by the idea of utilizing menhaden and other fishes for the production of guano fertilizer and considered it a worthwhile project. In a letter dated October 18, 1875, to John M. Glidden, treasurer of the Pacific Guano Works Company, Baird urged him "to make a display of your wares at the centennial (in Philadelphia), as this is one of the most important interests in the United States." He writes further that "there is no species (of fish) worked up elsewhere comparable to the movement with the menhaden, or pogy, as to numbers and the percentage of oil. The combination, too, of the pogy scrap with the South Carolina phosphates and the guanos of the West Indies and of the PacificA are also quite novel, and as being especially an American industry, are eminently worthy of full appreciation."

While the scientists, agriculturalists, and stockholders of the company thought very highly of the guano works, the existence of a malodorous plant was not appreciated by the residents of Woods Hole who suffered from a strongly offensive odor whenever the wind was from the west. Woods Hole might have continued to grow as one of the factory towns of Massachusetts but, fortunately for the progress of science and good fortune of its residents (except those who invested their savings in the shares of Pacific Guano Works), the company began to decline and became bankrupt in 1889.

Cessation of business and heavy monetary losses brought financial disaster to many residents of Woods Hole. The gloom prevailing in the village after the closing of the guano works began to dissipate, however, with the development of Woods Hole as a place of scientific research and with the increasing tourist trade. The factory buildings were torn down, the chimney which dominated the Woods Hole landscape was dynamited, and over 100,000 pounds of lead lining the acid chambers were salvaged. Large cement vats and the remnants of the old wharf remained; in the following years the latter became a favored place for summer biologists to collect interesting marine animals and plants.

The years from 1871 to the death of Baird in 1887 were the formative period of the new era of *Woods Hole as a scientific center*. In historical documents and in old books the present name Woods Hole is spelled in a different way. The old name "Woods Holl" is considered by some historians of Cape Cod to be a relic of the times prior to the 17th century when the Norsemen visited the coast. The "Holl", supposed to be the Norse word for "hill", is found in the old records. The early settlers gave the name "Hole" to inlets or to passages between the islands, such as "Robinson's Hole" between Naushon and Pasque Islands, or "Quick's Hole" between Pasque and Nashawena Islands, and Woods' Hole between the mainland and Nonamesset Island. In 1877 the Postmaster General ordered the restoration of the original spelling "Wood's Holl", which remained in force until 1896 when the United States Post Office changed it back to Woods

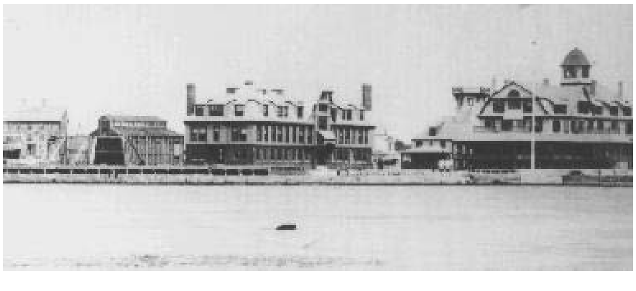

Hole and eliminated the apostrophe in Wood's. The change was regretted by the old timers and by C. O. Whitman who had given the specific name "hollensis" to some local animals he described.

At the time of his arrival at Woods Hole in 1871, Baird was well known to the scientific circles of this country and abroad as a naturalist, student of classification and distribution of mammals and birds, and as a tireless collector of zoological specimens. He maintained voluminous correspondence with the scientists in the United States and Europe, and was Permanent Secretary of the recently organized American Association for the Advancement of Science. To the general public he was known as a contributor to a science column in the New York Herald and author of many popular magazine articles. His newly acquired responsibilities as Commissioner of Fisheries greatly added to his primary duties as Assistant Secretary of the Smithsonian Institution which was primarily responsible for the establishment of the National Museum in Washington. As a scientist, Baird belonged to the time of Louis Agassiz, Th. H. Huxley, and Charles Darwin. Like Agassiz he attended medical college but never completed his studies, although the degree of M. D. honoris causa was later conferred upon him by the Philadelphia Medical College.

In the words of Charles F. Holder, "he was a typical American of the heroic type. A man of many parts, virtues, and intellectual graces, and of all the zoologists science has given the world.... he was most prolific in works of practical value to man and humanity."

Commissioner Baird attended many Congressional hearings and conferences with state officials and fishermen at which the probable causes of the decline of fisheries were discussed and various corrective measures suggested. From the lengthy and frequently heated discussions and evidence presented by the fishermen and other persons familiar with the fisheries problems, he became convinced that an alarmingly rapid decrease in the catches of fish had continued for the last 15 or 20 years. Such a decline was particularly noticeable in the case of scup, tautog, and sea bass in the waters of Vineyard Sound. It was logical, therefore, that the new Commissioner of Fisheries would select for his initial activities the New England coastal area where the fishing industry was of greatest importance as a politico-economical factor. Woods Hole, however, was not a significant fishing center.

In the "Fisheries and Fishlng Industry of the United States" prepared and edited by Goode (1884–87) for the 1880 Census, the fishing activity at Woods Hole is described in the following words: "Of the male inhabitants only seven are regularly engaged in fishing, the remainder being employed in the guano factory, in farming and other minor pursuits.... There is one ship carpenter in Wood's Holl, but he finds employment in his legitimate business only at long intervals. Of sailmakers, riggers, caulkers, and other artisans there are none. Four men are employed by Mr. Spindel, during the height of the fishing season, in icing and boxing fish. The boat fishery is carried on by seven men from April until September, inclusive. Only three species of fish are usually taken, namely, scup, tautog, and sea bass. The total catch of each fisherman is about 15 barrels, or about 2400 pounds. In addition about 6,720 lobsters are annually taken."

Before selecting a location for permanent headquarters for the work on fishery management and conservation, Baird undertook extensive explorations of the fishing grounds off the entire New England Coast. Section 2 of the Joint Resolution Number 8 of Congress gave the Commissioner full authority to carry out the necessary research. In part it reads as follows "and further resolved, That it shall be the duty of the said Commissioner to prosecute investigations and inquiries on the subject, with the view of ascertaining whether any and what diminution in the number of the food-fishes of the coast and the lakes of the United States has taken place; and, if so, to what causes the same is due; and also, whether any and what protective, prohibitory, or precautionary measures should be adopted in the premises; and to report upon the same

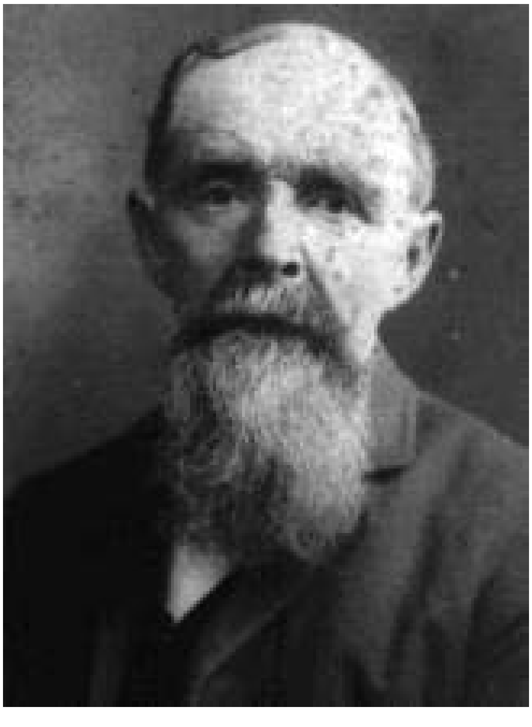

to Congress." Section 4 of the same Resolution contains an important clause which authorizes the Commissioner of Fisheries "to take or cause to be taken, at all times, in the waters of the seacoast of the United States, where the tide ebbs and flows, and also in the waters of the lakes, such fish or specimens thereof as many in his judgement, from time to time, be needful or proper for the conduct of his duties as aforesaid, any law, custom, or useage of any State to the contrary notwithstanding."

The significant words "where the tide ebbs and flows" were interpreted by Baird in a very broad scientific sense which extended the authority for his investigations to the offshore areas of the open ocean.

Pounds and weirs were most frequently accused by the public as destructive methods of fishing responsible for the decline in the abundance of food fishes along the coast. Although Baird gave very serious consideration to the possible destructiveness of fixed nets, traps, pounds, pots, fish weirs, and other stationary apparatus, he was fully aware of the complexity of the factors which may cause the decline in fish populations. He discusses this difficult problem in a paper entitled "Report on the condition of the sea fisheries of the south coast of New England" and published as the first section of the voluminous First Report of the Commissioner of Fish and Fisheries for 1871. Of the causes which may have contributed to the decrease of summer shore fisheries of the south side of Massachusetts and Rhode Island, a fact which he considered as well established by the testimonies of competent persons, he lists the following: (1) decrease or disappearance of the food of commercial fishes; (2) migration of fishes to other localities; (3) epidemic diseases and "peculiar atmospheric agencies, such as heat, cold, etc."; (4) destruction by other fishes; (5) man's activities resulting in the pollution of water, in overfishing, and the use of improper apparatus.

The biologist of today will recognize in this statement Baird's broad philosophical approach to the major problem of fishery biology. The outlined program combined oceanographical and meteorological investigations with the studies of biology, ecology, parasitology, and population dynamics of various fish species. Baird's program of research is as comprehensive and valid today as it was 90 years ago.

No time was lost in initiating this program. Woods Hole was selected as the base of the sea coast operations during the first summer and Vinal N. Edwards became the first permanent federal employee of the fisheries service. In spite of the insignificance of local fisheries, this locality offered a number of advantages which were recognized by Baird. Communication with Boston, New York, and Washington was good and promised to be better with the expected opening of the railroad branch in 1872. Being centrally located in relation to principal fishing grounds of New England and having good dock facilities and water of sufficient depth for sea going vessels, Woods Hole was a suitable base for visiting the offshore grounds. Furthermore, it was believed that the alleged decrease in food fishes was most clearly manifested in the region around Vineyard Sound. The small yacht Mazeppa of the New Bedford Custom House and the revenue-cutter Moccasin attached to the custom-house at Newport, R.I., were placed at the disposal of Baird; and the Light-House Board granted permission to occupy some vacant buildings and the wharf at the buoy-station on the west bank of Little Harbor. The Secretary of the Navy came to Baird's assistance by placing at his command a small steam launch which belonged to the Boston Navy Yard and by giving many condemned powder tanks which could be used for the preservation of specimens. Nets, dredges, tanks, and other gear were provided by the Smithsonian Institution. Cooperation of the various governmental agencies was authorized by Congress which in Section 3 of the Resolution specified that "the heads of the Executive Departments be, and they are hereby directed to cause to be rendered all necessary and practicable aid to the said Commissioner in the prosecution of the investigations and inquiries aforesaid."

This provision of the law was of great value. It is apparent, however, that the success in obtaining cooperation authorized by law depended a great deal on the personal characteristics of Baird, his great ability of getting along with people, and his remarkable power of persuasion, These qualifications played the major role in his success in organizing the Commission's work and also in obtaining the cooperation of scientists as well as that of fishermen and businessmen.

The investigation during the first summer consisted primarily in collecting large numbers of fishes and studying their spawning, rate of growth, distribution, and food. In the course of this work nearly all the fish pounds and traps, some 30 in number, in the vicinity of Woods Hole, were visited and their location recorded. There was no difficulty in obtaining the owners' permission to examine these installations and to collect the needed specimens. Altogether 106 species of fish were secured, photographed, and preserved for the National Museum. Of this number 20 or more species had not previously been known from Massachusetts waters. Information gained in this manner was supplemented by the testimonies of various fishermen who presented their ideas either for or against the use of traps and pounds. Among them was Isaiah Spindel, who at the request of Baird, prepared a description of a pound net used at Woods Hole and explained its operation. In the following years Spindel became an influential member of the group of local citizens who supported Baird's plan of establishing a permanent marine station at Woods Hole.

The ship Moccasin under the command of J. G. Baker was engaged in taking samples of plankton animals, in determining the extent of beds of mussels, starfish, and other bottom invertebrates, and in making temperature observations.

One of the principal collaborators in the studies conducted at Woods Hole in 1871 was A. E. Verrill of Yale University, a professor whom Baird appointed as his assistant and placed in charge of the investigations of marine invertebrates. Dredging for bottom animals during the first summer was carried out on a relatively small scale from a chartered sailing yacht Mollie and a smaller vessel used in the immediate vicinity of Woods Hole. Extensive collections were made by wading on tidal flats exposed at low water.

Zoological work attracted considerable interest among the biologists of this country. Many of them stopped at Woods Hole for greater or lesser periods and were encouraged by Baird to use the facilities of the Fish Commission. The group included such well known men as *L. Agassiz*, A. Hyatt, W. G. Farlow, Theodore Gill, Gruyure Jeffries of England, and many others. The first year's work extended until the early part of October. Before returning to Washington, Baird commissioned Vinal N. Edwards of Woods Hole to continue the investigation as far as possible. By the end of the first year a general plan of study of the natural histories of the fishes and the effect of fishing on fish populations was prepared with the assistance of the well-known ichthyologist, Theodore N. Gill. His old "Catalogue of the fishes of the Eastern Coast of North America from Greenland to Georgia", was revised and the next text including the recently collected data concerning the Massachusetts fishes, appeared in the First Report of the U.S. Commissioner of Fish and Fisheries. The plan of investigation suggested by Gill was adopted by Baird as a guide for the work of his associates for the purpose of "securing greater precision in the inquiries." The plan is composed of 15 sections, such as Geographical distribution, Abundiance, Reproduction, etc., with detailed subdivisions under each one. A questionnaire containing 88 different items was included in order to facilitate the inquiries conducted among the fishermen. The scope of the highly comprehensive program is complete enough to be useful today; marine biologists of today would probably only rephrase it, using modern terminology. During the first year of operations conducted at Woods Hole, Baird and his associates laid down the foundation of the new branch of science which we now call fishery biology or fishery science.

AN ANALYSIS OF THE INTERNATIONAL MARITIME ORGANIZATION–LONDON CONVENTION ANNUAL OCEAN DUMPING REPORTS

Christine Dickenson
Iver W. Duedall
Florida Institute of Technology
Melbourne, Florida

Trends are analyzed for types and quantities of permitted wastes, primarily dredged material, sewage sludge, and industrial waste, to be dumped at sea by member countries to the London Convention (LC) from 1992 and 1995 through 1998. In 1972, the Inter-Governmental Conference on the Dumping of Wastes at Sea led to the creation of the London Dumping Convention (now called the London Convention) to help regulate the dumping of wastes at sea. The act of dumping, as defined by the LC, is international disposal at sea of any material and in any form, from vessels, aircraft, platforms, or other artificial structures. The first consultative meeting of the LC contracting parties was held in 1976 by the Inter-Governmental Maritime Consultative Organization [now called the International Maritime Organization (IMO)]. During this meeting, the procedure for the reporting of permits issued, on an annual basis, for dumping at sea by contracting parties, was determined. The IMO-LC annual reports on permitted wastes list the number of permits issued by member countries, the types and quantities of wastes permitted for dumping at sea, and the location and designation of dump sites. Now with nearly 25 years of dumping records available, we are able to see trends in permitted dumping activity. In 1976, the first year of permitted dumping records, the combined amount of permitted wastes was nearly 150 million tons. In the last four years (1995 through 1998), where data are available, the total amount of wastes permitted to be dumped by the LC contracting parties as between 300 and 350 million tons. Currently, a majority of these wastes are being disposed of in the East Asian Seas and the North Sea and the largest quantity of waste being dumped is dredged material. Although the disposal of wastes at sea is considered to be a major issue, it is only responsible for about 10% of the total anthropogenic contaminants entering the ocean. Unfortunately, the long-term impacts of this dumping in the ocean are still largely unknown.

INTRODUCTION

"Historically, most coastal countries used the sea for waste disposal. It was generally the most economic way to manage waste, since land usually had, and still has, a high price tag while the sea has no private owner in the normal sense. In addition, dilution processes served the illusion that dumping at sea does not cause any permanent damage. So why risk contaminating land or drinking water with wastes if the sea was close by?" (1). "Accurate worldwide records on the amounts of wastes disposed at sea prior to 1976 are virtually impossible to obtain" (1). However, as a result of the international activities leading to the creation of the London Dumping Convention (LC) in 1972, information is now available on the number of permits issued by many countries for disposal at sea, their dumpsite locations, and the kinds and quantities of wastes that have been dumped (1).

According to the LC, the most common form of ocean dumping today is disposal from ships or barges (1,2). Wastes are loaded on these vessels and then taken to the dumpsites. Dumpsites are chosen based on the kind of waste and the ocean's properties at each site (2). Liquid wastes are generally disposed of in more dispersive environments, where mixing will rapidly dilute the dumped material. Solid wastes, on the other hand, are usually disposed of in less dispersive near-shore sites to keep the solids confined.

Here we specifically analyze trends in types and quantities of permitted wastes to be dumped at sea by member countries to the LC from 1992 and 1995 through 1998. The types of wastes that have been dumped in the ocean include industrial waste, sewage sludge, dredged material, incineration at sea, and radioactive wastes. Industrial waste may include both liquid and solid wastes and it may contain such items as acid-iron waste, fish-processing liquids, metal refinery wastes, and gas pipeline flushing wastes (2). An overall reduction in the dumping of these types of wastes has been achieved over the years by switching to alternative disposal methods, reusing wastes, and creating cleaner production technologies. The dumping of industrial wastes at sea has been prohibited since 1996 (3). "Sewage sludge is an anaerobic waste product from treatment of municipal wastewater. The sludge is in aqueous form containing about 3% suspended particles by weight" (2). Alternatives for the disposal of sewage sludge at sea include incineration, deposit on land, and agriculture use (3). "Dredged materials range from clean sands to heavily contaminated fine grained materials" (2). "Physical properties of dredged materials, including grain size, bulk density, water content, and geotechnical characteristics, are especially variable due to the kind or type of sediment being dredged, which is itself dependent on geological and watershed characteristics, as well as to the operational procedures used in dredging and disposal" (2). Incineration at sea is defined as the burning of liquid chlorinated hydrocarbons as well as other halogenated compounds in which all ash is dumped into the sea (3). This type of dumping was phased out early in 1991. The dumping of radioactive wastes, however, might be the most harmful practice of all. Although the dumping of high level radioactive wastes has never been allowed under the London Convention, it has still occurred in some cases, and even though a moratorium was placed on the dumping of low level radioactive wastes in 1983, this type of dumping still occurred (3,4). It was not until 1994 that this act became legally binding. Finally, the other waste category includes such wastes as inert geological materials, decommissioned vessels, scrap metals, and fish wastes (3).

METHODS

Annual reports on permitted dumping at sea by contracting parties were obtained from the IMO-LC (5). There is approximately a three-year delay before the most recent year's records are put into report format and published. Older records that only existed in hard copy form had to be entered into Excel spreadsheets to create a new database. Early on, it was decided that all wastes would be reported in tons to make comparisons from year to year easier. Many of the older reports had wastes recorded in cubic meters. These records were converted into tons by multiplying the cubic meters by the density of the waste (1.3 g/cm^3 for dredged material and other varying factors as indicated by each permit). Newer records, however, were received in Excel format and, for the most part, they were already converted into tons, making them very easy to add to the database.

The primary categories of wastes of importance to this study include dredged material, sewage sludge, industrial waste, and other matter. From the 1992 and 1995 through 1998 reports (5), it was desired to find out which waste was being dumped in the largest quantity, as well as which countries were dumping the most wastes and which water bodies were receiving the most wastes. The more recent data is compared to earlier reports from 1976 to 1985. Comparisons of the different sources of pollutants in the oceans and comparisons between land dumping and sea dumping are also made. Then the past, as well as the future, of ocean dumping can be assessed.

RESULTS

Tables in the IMO-LC annual reports (5) on permitted wastes (IMO, 1992, 1995–1998) list the number of permits issued to member countries, the types and quantities of wastes permitted for disposal at sea, and the location and designation of dumping sites. It is important to know that these reports are reflective of what has been permitted to be dumped and not of what has actually been dumped. The accuracy of these records, therefore, is somewhat questionable (1). It is also important to remember that not all of the contracting parties report their activities every year and some reports from contracting parties may be incomplete (2). For example, in 1995, thirty-eight of the seventy-five contracting parties registered did not report their activities. There is also no way of knowing how much noncontracting parties are dumping (2).

In 1992, 473 million tons were reported as permitted for disposal at sea. In the last four years (1995 through 1998) where data are available, the total amount of wastes permitted to be dumped by the LC contracting parties was 351 million tons in 1995, 312 million tons in 1996, 309 million tons in 1997, and 348 million tons in 1998. Figure 1 illustrates these data.

As seen in Fig. 2, a majority of these wastes are being disposed of in the East Asian Seas (Fig. 3) and the North Sea (Fig. 4). The "other areas" category is represented mostly by the United States and its disposal in the Gulf of Mexico (Fig. 5).

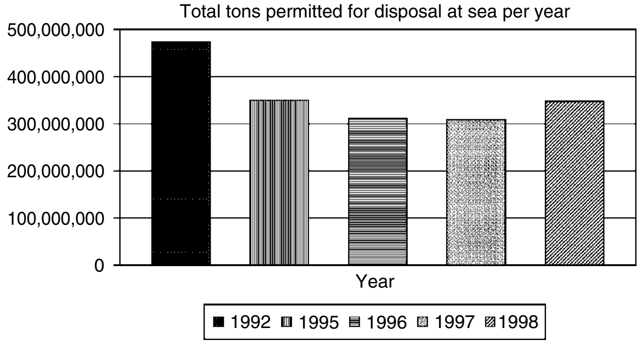

Figure 1. London Convention permitted dumping from 1992 and 1995 through 1998.

The countries responsible for dumping the most wastes, in descending order, are China, the United States, Hong Kong, the United Kingdom, Germany, and Belgium (Table 1).

As shown in Fig. 6, the largest quantity of waste being dumped is dredged material. At 327 million tons in 1998, it is the largest quantity by far, with sewage sludge coming in as the next highest quantity with only 15 million tons.

DISCUSSION

In 1976, the first year of permitted dumping records, the combined amount of permitted wastes was nearly 150 million tons (1). By 1985, this amount had grown to approximately 300 million tons (1). This apparent doubling, however, partly corresponds to an increase in the number of contracting parties to the LC. It is also important to remember that not all of the contracting parties report their activities every year and some reports from contracting parties may be incomplete. Therefore, it is not possible to provide a highly accurate interpretation of these data and the reader should be cautioned that any analysis must be considered approximate.

Figure 3. China and Hong Kong dumping in the East Asian Seas. (From Ref. 6).

Looking back to the data in Fig. 1, waste disposal in the ocean seems to continue to increase into the early 1990s with 473 million tons in 1992. However, by 1995 through 1998 this amount had leveled off to between 300 and 350 million tons. This decrease could be representative of the many changes in policy occurring within the LC in the 1990s to intentionally decrease the amount of waste disposed of in the ocean (3). It is also known that a major dredging project took place in Hong Kong in the early 1990s in order to expand the airport there (3). This may explain why the amount of waste dumped in 1992 was so high. The East Asian Seas section of Fig. 2 supports this theory, with data from 1992 being much higher than in the following years. Yearly fluctuations like this occur due to the variation in maintenance dredging and new works associated with shipping activities (3).

According to the IMO-LC web page, "in early 1991, incineration at sea operations came to a halt, ahead of

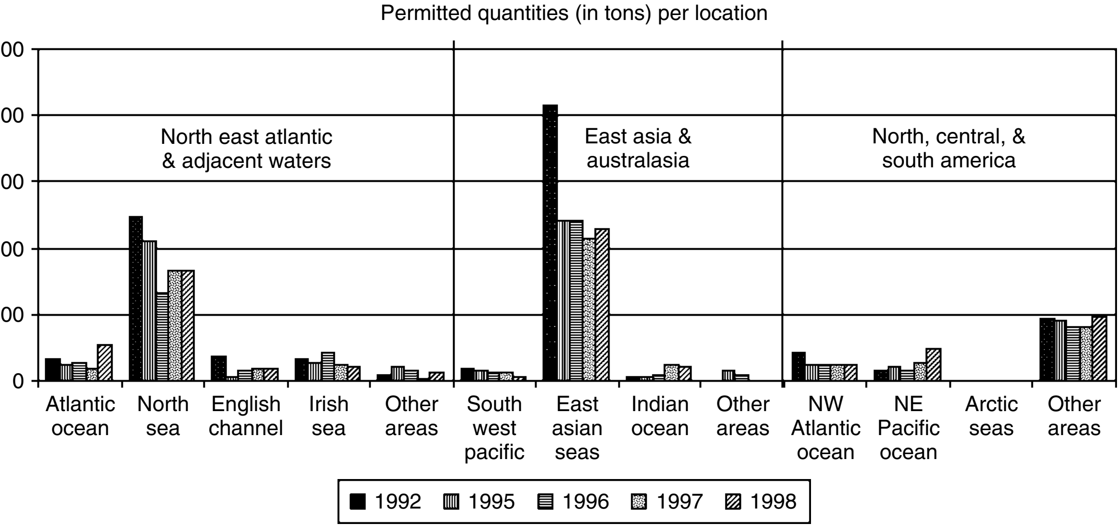

Figure 2. Locations of dumping sites with their respective amounts of wastes dumped.

Figure 4. The United Kingdom, Germany, and Belgium dumping in the North Sea. (From Ref. 6).

the agreed global deadline of December 31, 1992.... In 1990, contracting parties to the LC agreed to phase-out sea disposal of industrial waste effective by January 1, 1996.... In 1993, contracting parties started a detailed review of the LC, leading to the adoption of a few crucial amendments.... These amendments consolidated in a legally binding manner the prohibition to dump all radioactive wastes or other radioactive matter and of industrial wastes, the latter as per January 1, 1996, as well as the prohibition of incineration at sea of industrial wastes and of sewage sludge" (3).

Table 1. Highest Combined Tons for 1992 and 1995–1998

Rank	Country	Tons
1	China	295,482,060
2	United States	293,272,190
3	Hong Kong	285,437,056
4	United Kingdom	246,759,318
5	Germany	132,149,000
6	Belgium	127,507,436

According to reports by contracting parties, no permits for the dumping of industrial waste have been issued since 1996. Before this phase-out, Japan and the Republic of Korea were responsible for most of the industrial waste being dumped (3). The amount of sewage sludge being dumped at sea decreased in the early 1990s, reflecting the phase-out of this practice by several countries, Ireland and the United Kingdom being the most recent. Currently, only Japan, the Philippines, and the Republic of Korea dispose of sewage sludge at sea (3).

Figure 5. The United States dumping in the Gulf of Mexico and the northwest Atlantic. (From Ref. 6).

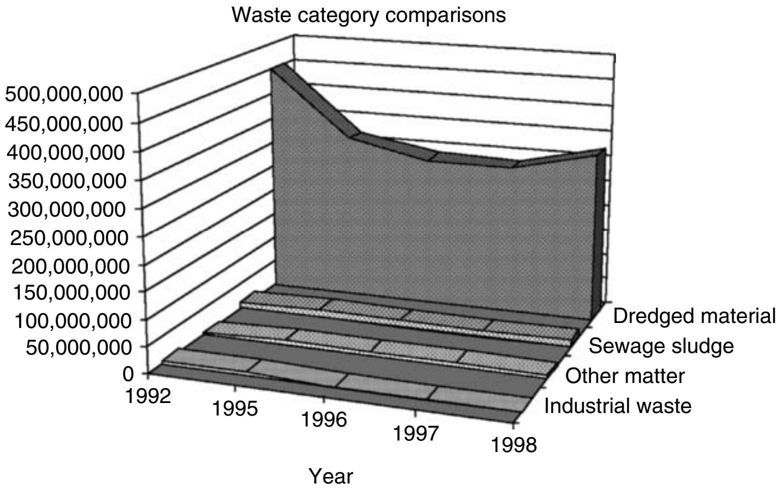

Figure 6. Dredged material, by far, makes up the largest quantity of waste being dumped at sea.

Dredged material makes up the majority of what is being dumped at sea (see Fig. 6). Unlike other wastes, the amount of dredged material dumped each year tends to fluctuate greatly due to variable maintenance dredging and new works associated with shipping activities (3). For example, there were huge dredging efforts going on in Hong Kong with the extension of the airport there in 1992. One hundred forty-four million tons of dredged material were permitted to be disposed of at sea that year by Hong Kong. Arguably, this could be responsible for the 1994 anomaly of the combined total of 473 million tons, meaning that the average amount of wastes dumped annually actually leveled out to between 300 and 350 million tons much earlier than the later 1990s.

The interactions of the wastes dumped at sea with seawater and the toxicity of these wastes to organisms, including humans, are of great importance to the scientific community. Although the long-term effects of disposal at sea are unknown, many studies have looked at some of the short-term effects, particularly in relation to radioactive wastes that have been dumped. Studies conducted in the Arctic by the IAEA and the U.S. Congress Office of Technology Assessment found that (1) releases from dumped objects were small and confined to the immediate vicinity of the dumped objects, (2) projected future doses to the general public were small, (3) doses to marine organisms were insignificant in the context of populations, and (4) remediation on radiological grounds was not warranted (7). A study by Hill et al. (8) assesses the possible effects of 40 years of dredging, and how it might have affected benthos in the Irish Sea. Zooplankton and micronekton communities, off eastern Tasmania, as studied by Bradford et al. (9), were found to be affected by the presence of a warm-core eddy instead of jarosite wastes dumped within the vicinity. As de La Fayette (7) said, it is difficult to draw conclusions from data that has been collected because more research, monitoring, and prevention projects are still needed for us to understand all of the factors that are involved in knowing how, when, and where ocean stored wastes might affect us and the environment.

Although such concern is felt about disposal at sea, it is not the largest contributor of pollutants to the oceans. As seen in Fig. 7, the IMO-LC puts it in fourth place, making up only 10% of the total pollutants in the oceans. Runoff and land-based discharges are the largest source, making up 44%. Even maritime transportation, at an estimated 12%, pollutes the oceans more than dumping does.

In an attempt to compare ocean dumping to disposal in landfills, we look at the United States. According to Zero Waste America (10), in 2001, 409 million tons of municipal waste were generated in the United States. Of that, 278 million tons were disposed of in landfills. Table 2 illustrates dumping on land and in the ocean by the United States during 1992 and 1995 through 1998.

Those five years of ocean dumping, totaling 293 million tons, is only 15 million tons more than what was disposed of in landfills in 2001 alone. In conclusion, we note that much more waste is being disposed of in landfills than in the ocean. Therefore, shouldn't we be more concerned about the land?

CONCLUSION

Although the environmental impacts of dumping at sea are still largely unknown, it is comforting that the LC is moving in a positive direction toward more regulations for the better protection of our oceans. With industrial

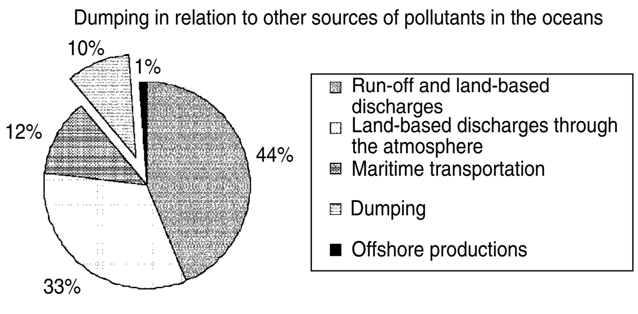

Figure 7. At 10%, dumping ranks fourth compared to the other sources of pollutants in the ocean.

Table 2. Ocean Dumping versus Land Dumping in the United States

Year	Tons in Millions	
	Land	Ocean
1992	241	67
1995	249	58
1996	238	46
1997	236	53
1998	238	69

Data taken from Reference 10.

waste already phased out and sewage sludge on its way to becoming completely phased out, that leaves dredged material and "other matter" as the future of ocean dumping. Other matter mostly makes up inert geological material from mining and excavations; bulky wastes such as steel equipment, scrap metal, and concrete; fish wastes; obsolete ammunition and explosives; discontinued oil platforms; spoiled food; and other random wastes. This category may also eventually be phased out, but dredged material will most likely continue to be dumped since this form of waste, in the normal sense, either came "from the ocean floor" or somewhere close to it being that two-thirds of dredged material is connected with maintenance operations to keep harbors, rivers, and other waterways from being blocked up. Unfortunately, according to the IMO-LC, about 10% of dredged material is moderately to heavily contaminated from a variety of sources including shipping, industrial and municipal discharges, and land runoff. Whether we find somewhere to dispose of dredged material on land, or we continue to dispose of it in the sea, it will continue to affect our environment.

Acknowledgments
We are grateful to Rene Coenen, Office of the London Convention, International Maritime Organization, for kindly providing the LC dumping records. We are also thankful to Ruth Caulk and Melissa Sheffer for assisting in the preparation of the ocean dumping spreadsheets.

BIBLIOGRAPHY

1. Duedall, I.W. (1990). A brief history of ocean disposal. *Oceanus* **33**(2): 29.
2. Connell, D.W. and Hawker, D.W. (1992). *Pollution in Tropical Aquatic Systems*. CRC Press, Boca Raton, FL.
3. London Convention. (2003). Available: http://www.london-convention.org.
4. Ahnert, A. and Borowski, C. (2000). Environmental risk assessment of anthropogenic activity in the deep sea. *J. Aquat. Ecosys. Stress Recov.* **7**: 299–315.
5. IMO. (1972). *Convention on the Prevention of Marine Pollution by Dumping of Wastes and Other Matter, London Convention 1972*. Final Reports on Permits Issued in 1992, 1995, 1996, 1997, and 1998.
6. Karls, F. (1999). *World History: The Human Experience—The Modern Era*. Glencoe/McGraw-Hill, New York.
7. de La Fayette, L. (1998). The London Convention 1972: preparing for the future. *Int. J. Mar. Coastal Law* **13**(4): 515–536.
8. Hill, A.S. et al. (1999). Changes in Irish Sea benthos: possible effects of 40 years of dredging. *Estuarine Coastal Shelf Sci.* **48**: 739–750.
9. Bradford, X. et al. (1999).
10. Zero Waste America. (2001). Available: http://www.zerowasteamerica.org/Statistics.htm.

MARINE SOURCES OF HALOCARBONS

Robert M. Moore
Dalhousie University
Halifax, Nova Scotia, Canada

Halogenated gases are important in the atmosphere by virtue of their ability to carry chlorine, bromine, and iodine to high altitude, where they can act as efficient catalysts of ozone destruction. The effect on ozone of chlorine release in the stratosphere by certain manufactured chlorofluorocarbons is well known, but naturally occurring halogen carriers account for a part of the background cycle of ozone breakdown. The oceans constitute an enormous reservoir of dissolved halogens in the form of halide ions, but the transport of this material to the upper atmosphere is inefficient because sea salt particles, introduced to the atmosphere as sea spray from breaking waves and bubble-bursting, have short lifetimes, being readily washed out by rain. For this reason the relatively minute concentrations of dissolved halogenated gases, such as methyl chloride occurring in surface seawater at concentrations around 10^{-10} M, have the potential to drive significant transport of chlorine into the upper atmosphere. Since the role of anthropogenic chlorine and bromine gases in stratospheric ozone depletion was recognized, there has been renewed interest in the ocean as a source of halogenated trace gases. The emission of iodine compounds from the ocean has more recently been shown to be potentially important in aerosol production as well as in affecting ozone concentrations over the ocean (1).

The fluxes of trace gases such as those between the ocean and atmosphere is most commonly calculated from concentration measurements in the surface ocean and the overlying atmosphere together with empirical relationships between gas exchange coefficients and wind velocity (2). Supersaturation of a gas at the sea surface with respect to the atmosphere will support an outward flux, the magnitude of which is strongly dependent on wind speed, and to a lesser extent on temperature. Gas fluxes can be highly variable spatially and seasonally, and, in general, estimated global fluxes will have a substantial degree of uncertainty due to the sparseness of the concentration measurements combined with a large uncertainty in the exchange velocity. Ideally, gas fluxes would be calculable based on a firm understanding of the processes governing both the production and loss of the gas in the upper ocean and the processes controlling gas exchange. This remains a distant goal, and current studies are directed at developing an understanding of these processes.

Some of the first measurements made of halogenated methanes in seawater suggested that the ocean could be the predominant source of methyl chloride accounting for most of the 4 million tonnes that must be added annually to the atmosphere to support the observed concentrations (3). More recently, it has been demonstrated that, in the case of methyl chloride, the ocean is just one of many sources (including biomass burning, woodrot fungi, and vegetation) and can probably account for only about 15% of the total flux to the atmosphere (4). In contrast, the ocean is still recognized to be the most important source of atmospheric methyl iodide.

Impetus for the study of sources of methyl bromide came with the discovery of the efficiency with which bromine could catalyze stratospheric ozone loss, working either directly or in concert with chlorine. This led to concern about the possible role of methyl bromide, manufactured for use as a fumigant, in observed ozone loss. The existence of natural sources and sinks for methyl bromide made it more difficult to quantify the deleterious effect of the anthropogenic component. Extensive studies of the distribution of methyl bromide in the oceans and of its flux across the air–sea interface in different regions demonstrated the existence of both production and loss processes within the ocean (5). The balance between these determines whether a particular region will emit CH_3Br to the atmosphere or take it up. A major oceanic sink results from the reaction between CH_3Br and chloride ion, the rate being strongly dependent on temperature. Other losses include chemical hydrolysis and also biological uptake. Bromine is carried to the atmosphere by many other gases produced in the ocean, notably bromoform and dibromomethane (6).

Early studies pointed to an association between certain halogenated compounds in seawater, such as methyl iodide, and coastal beds of macrophytes. The more general phenomenon of organohalogen production by seaweeds has long been known, and a wide range of halogenated organic compounds of varying molecular weight and complexity have been identified (7). The focus here is on halogenated compounds that play a role in atmospheric chemistry, so these are typically volatile compounds. Two such compounds that have been shown to have major sources in macrophytes are di- and tribromomethane (CH_2Br_2 and $CHBr_3$). The latter has been found to occur within certain seaweeds at concentrations up to the percent level. It is apparent that haloperoxidase enzymes are involved in the biosynthesis of organohalogens by macrophytes.

In spite of the variety and abundance of halogenated methanes in seaweeds and the flux of halogen to the atmosphere that can thus be supported in the coastal zone, the overall importance of these algae on an oceanwide scale is small on account of their limited distribution (8), although, in the case of bromoform and dibromomethane, Carpenter and Liss, (9) estimated that macrophytes are a major source. For this reason there has been interest in the possible production of the same compounds by marine microalgae, which are ubiquitous in the sunlit waters of the world's oceans. An approach that has been adopted with some success to study this question has been the use of laboratory cultures of specific algae. This has demonstrated the capacity of a range of phytoplankton of different algal classes to produce compounds such as CH_3I, CH_2Br_2, $CHBr_3$, and CH_2I_2 (10,11), although species vary in the spectrum of compounds that they produce as well as in the production rates. There has been some progress in elucidating the production mechanism of these compounds with the identification of haloperoxidase enzymes in a few phytoplankton species (11). However, much uncertainty remains regarding the oceanic source of methyl chloride, bromide, and iodide. For although there are published studies that point to the capacity of some phytoplankton to produce methyl halides (12), the measured rates when normalized to biomass concentrations have typically been found to be quite inadequate to account for the observed fluxes from the ocean into the atmosphere (10). This may be due in part to the limitations of laboratory studies of phytoplankton cultures: relatively few species can be grown successfully in the laboratory and the conditions are very dissimilar to those in the ocean itself. Alternative possible explanations include the involvement of microbial processes (13), zooplankton, and photochemistry, or a combination of sources. One laboratory study has provided some evidence for the production of methyl iodide in seawater irradiated with simulated sunlight (14), but it has yet to be demonstrated that this process is significant in the ocean.

Open ocean studies of halocarbon distributions show that there is no simple correlation with phytoplankton biomass, quantified as chlorophyll *a* concentration. This means that the production of a particular compound cannot be attributed uniformly to all species of phytoplankton. More success has been obtained in studies that measured a series of photosynthetic pigments and evidence is forthcoming from one such study that CH_3Br has a source in Prymnesiophytes (15).

Measurements of dissolved methyl halides in the ocean typically show relatively high concentrations in the surface mixed layer, frequently with a maximum directly beneath, declining to levels at or near detection limits in deep waters (16). Such profiles are broadly consistent with a source at or near the surface and net consumption at greater depths. The maximum may be explained as occurring in a zone of production beneath the surface mixed layer that is poorly ventilated, so with reduced emission to the atmosphere. However, certain halocarbons have quite different distributions with depth. The chlorofluorocarbons such as CFC11 and CFC12 may have higher concentrations below the surface and at intermediate depths, with levels diminishing to low values in some of the deepest waters of the ocean. These distributions are well understood to be the result of transport of these anthropogenic gases from the atmosphere either by direct exchange at the surface or via downward mixing and advection of cold, dense, ventilated waters to intermediate and abyssal depths. There is now evidence that some halocarbons that have relatively short atmospheric lifetimes (e.g., dichloromethane and tri- and tetrachloroethylene) are also introduced to the ocean by the same processes, and that these gases have much longer lifetimes in the ocean than in the atmosphere. This apparently explains the presence of dichloromethane in the near bottom waters of the Labrador Sea (17).

Chloroform is likely to undergo the same processes but may have an additional *in situ* source.

Loss processes occurring in the ocean can be inferred for most organohalogens by their lower concentrations at depth than at the surface. They have been studied most thoroughly for methyl bromide and, for this compound, are known to include bacterial removal and chemical loss through hydrolysis and reaction with chloride ion (18). Other loss processes affecting a range of halogenated compounds apparently exist in waters that are depleted in dissolved oxygen (19). It should be noted a loss process for one halocarbon might represent a source for another: thus, reaction of both CH_3Br and CH_3I with chloride ion yields CH_3Cl. The photocatalyzed loss of CH_2I_2 is a source of CH_2ICl.

Continued measurement of halocarbons both in the ocean and atmosphere will improve our knowledge of the magnitude, geographic distribution, and seasonality of their fluxes, but a greater challenge is likely to be identifying more accurately their sources. That knowledge is essential if we hope to be able to predict how the fluxes of these atmospherically reactive gases may change in the future.

BIBLIOGRAPHY

1. O'Dowd, C.D. et al. (2002). Marine aerosol formation from biogenic iodine emissions. *Nature* **417**: 632–636.
2. Wanninkhof, R. (1992). Relationship between wind speed and gas exchange over the ocean. *J. Geophys. Res.* **97**: 7373–7382.
3. Harper, D.B. (2000). The global chloromethane cycle: biosynthesis, biodegradation and metabolic role. *Nat. Prod. Rep.* **17**: 337–348.
4. Moore, R.M. (2000). The solubility of a suite of low molecular weight organochlorine compounds in seawater and implications for estimating the marine source of methyl chloride to the atmosphere. *Chemosphere: Global Change Sci.* **2**: 95–99.
5. Lobert, J.M. et al. (1995). *Science* **267**: 1002–1005.
6. Quack, B. and Wallace, D.W. (2003). Air–sea flux of bromoform: controls, rates, and implications. *Global Biogeochem. Cycles* **17**(1): 1023–1029.
7. Gribble, G.W. (2003). The diversity of naturally produced organohalogens. In: *The Handbook of Environmental Chemistry*. Vol. 3, Part P, G.W. Gribble (Ed.). Springer, New York, pp. 1–15.
8. Baker, J.M. et al. (2001). Emissions of CH_3Br, organochlorines and organoiodines from temperate macroalgae. *Chemosphere Global Change Sci.* **3**: 93–106.
9. Carpenter, L.J. and Liss, P.S. (2000). On temperate sources of bromoform and other reactive organic bromine gases. *J. Geophys. Res.* **105**: 20,539–20,547.
10. Manley, S.L. and de la Cuesta, J. (1997). Methyl iodide production from marine phytoplankton cultures. *Limnol. Oceanogr.* **42**: 142–147.
11. Moore, R.M., Webb, M., Tokarczyk, R., and Wever, R. (1996). Bromoperoxidase and iodoperoxidase enzymes and production of halogenated methanes in marine diatom cultures. *J. Geophys. Res.* **101**: 20,899–20,908.
12. Scarratt, M.G. and Moore, R.M. (1996). Production of methyl chloride and methyl bromide in laboratory cultures of marine phytoplankton. *Mar. Chem.* **54**: 263–272.
13. Amachi, S., Kamagata, Y., Kanagawa, T., and Muramatsu, Y. (2001). Bacteria mediate methylation of iodine in marine and terrestrial environments. *Appl. Environ. Microbiol.* **67**: 2718–1722.
14. Moore, R.M. and Zafiriou, O.C. (1994). Photochemical production of methyl iodide in seawater. *J. Geophys. Res.* **99**: 16,415–16,420.
15. Baker, J.M. et al. (1999). Biological production of methyl bromide in the coastal waters of the North Sea and open ocean of the northeast Atlantic. *Mar. Chem.* **64**: 267–285.
16. Moore, R.M. and Groszko, W. (1999). Methyl iodide distribution in the ocean and fluxes to the atmosphere. *J. Geophys. Res.* **104**: 11,163–11,171.
17. Moore, R.M. (2004). Dichloromethane in N. Atlantic Ocean waters. *J. Geophys. Res.* **109**: 9004.
18. King, D.B. and Saltzman, E.S. (1997). Removal of methyl bromide in coastal seawater: chemical and biological rates. *J. Geophys. Res.* **102**: 18,715–18,721.
19. Tanhua, T., Fogelqvist, E., and Bastürk, Ö. (1996). Reduction of volatile halocarbons in anoxic seawater, results from a study in the Black Sea. *Mar. Chem.* **54**: 159–170.

FOOD CHAIN/FOODWEB/FOOD CYCLE

ANDREW JUHL
Lamont–Doherty Earth
Observatory of Columbia
University
Palisades, New York

A representation of the feeding relationships of the organisms, or groups of organisms, within an ecological community will be shown here. By showing which organisms feed on which other organisms, the pathways of energy flow through the ecosystem can be followed. "Food chain" is an older term and is currently less used than "foodweb" or "food cycle." The newer terms were coined to acknowledge the complexity of feeding relationships within most ecosystems. Examples of marine planktonic and intertidal foodwebs will be used below to illustrate specific points, but the concepts are applicable to any ecosystem.

The simplest food chain is organized in a strict hierarchy with primary producers (organisms that generate organic matter by fixing inorganic carbon, usually through photosynthesis) eaten by herbivores (primary consumers), the herbivores eaten by carnivores (secondary consumers), on to tertiary consumers, quaternary consumers, and soon. Each level within the food chain is termed a trophic level. Figure 1 shows an example of a very simple food chain for a planktonic ecosystem where phytoplankton are the primary producers, eaten by zooplankton herbivores. Zooplankton, in turn, are eaten by fish. Each group of organisms is represented by a shape, and the feeding connection and direction of energy flow are shown with an arrow. Simple food chains have heuristic value and are the basis of many quantitative ecosystem models. However, they can be criticized for oversimplification and for missing major groups of organisms.

One step toward a more realistic picture of an ecosystem is to disaggregate the trophic levels. Ideally, each level

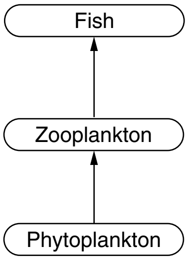

Figure 1. A simple food chain for a planktonic ecosystem.

could be broken down into its constituent species. One problem to consider is whether rare species should be depicted. A food chain diagram could include any feeding relationship that occurs within a system, whether common or unusual. This could lead to so many boxes and connecting arrows that any heuristic value would be jeopardized. In addition, the required information to fully disaggregate species groups may not be available. For example, microorganisms can be difficult to distinguish to the species level. Practical considerations often lead researchers to lump microorganisms into size classes and functional groups (e.g., primary producers >2 μm, 2–20 μm, <20 μm). In practice, the degree of aggregation of species is dictated by the objectives that motivate construction of the food chain diagram.

It is also more realistic to deemphasize the hierarchy implied by trophic levels. Deviations from strict trophic levels are common in aquatic ecosystems. For example, omnivorous species can prey on different trophic levels. A planktonic copepod may sometimes feed on small zooplankton, acting as a secondary consumer. The same copepod may also feed on phytoplankton, thereby acting as a primary consumer (1). Mixotrophy is another complication, common in planktonic ecosystems. Mixotrophic organisms can operate as primary producers, yet they may also consume other primary producers, or even consumers (2). Some aquatic organisms can occupy different trophic levels during different developmental stages. For example, eggs and early larval stages of the copepod *Acartia tonsa* are fed upon by the heterotrophic dinoflagellate *Protoperidinium* cf. *divergens* (3). However, adult *A. tonsa* feed voraciously on the dinoflagellate (4). This is an example of an "ontogenetic niche," or a niche that changes as the organism develops (5). Accommodating such complexities is one reason why the term "foodweb" is generally preferred over "food chain" (see Fig. 2). The many arrows necessary to accurately depict feeding relationships within most ecosystems resemble a web more than a linear chain.

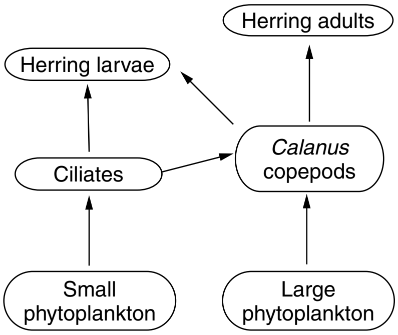

Figure 2. A more complex foodweb for the planktonic ecosystem in Fig. 1. Some groups are taxonomically identified, others are separated by size. The copepods are omnivorous, feeding on both phytoplankton and ciliates. Different feeding roles are shown for adult and larval herring.

Decomposers are frequently neglected when foodweb diagrams are constructed (6,7), and yet organic production from all trophic levels eventually cycles through decomposers. As the importance of heterotrophic bacteria and microzooplankton to planktonic biogeochemical cycles began to be realized during the 1970s and 1980s, depictions of plankton foodwebs began to include bacterial decomposers and their consumers in what has become known as the "microbial loop" (8,9) (see Fig. 3). Including decomposers in foodwebs of other ecosystems would certainly be more realistic, although it may not be necessary for every research goal.

QUANTITATIVE FOODWEBS

Even a complicated foodweb diagram is a qualitative description of an ecosystem and may only be representative of a single point in time. One approach to making the foodweb quantitative is to include an estimate of the energy flux along each pathway (10). Energy produced at the foodweb base will be dissipated at successive trophic levels. Knowing or assuming the energy dissipation at each step allows an estimate of the maximum production possible at higher trophic levels. For example, maximum potential fisheries production of the world's oceans has been derived from global estimates of planktonic primary production (11,12). In these analyses, different areas of the world's oceans were assumed to have simple food chains of varying length, resulting in variable efficiency of energy transfer until a trophic level with harvestable biomass was reached. Ocean regions with short food chains had greater fishery production per unit primary production than areas with long food chains.

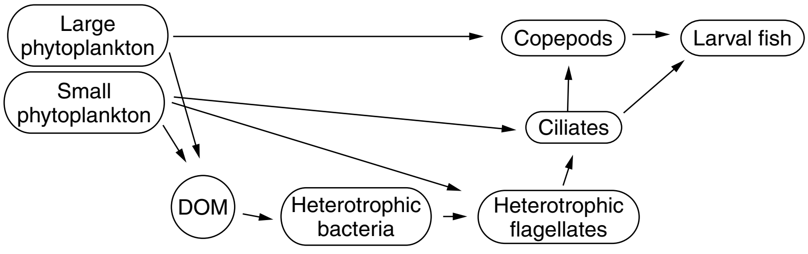

Figure 3. A planktonic foodweb that includes a representation of the "microbial loop" as an alternate pathway from phytoplankton to larger organisms. The loop, shown with thicker arrows, begins with phytoplankton producing dissolved organic matter (DOM). DOM feeds heterotrophic bacteria that are preyed upon by small, unicellular predators (heterotrophic flagellates and ciliates). These protists are food for copepods and larval fishes.

The biological components of many ecosystem models can also be viewed as attempts to make quantitative foodwebs. Models of plankton ecosystems (13,14) are usually not based on energy transfer. Instead, fluxes to and from organisms, and the standing biomass of the organisms, are expressed in units of an important element, often nitrogen, phosphorus, or carbon. Nevertheless, the flux pathways between organisms describe a food chain or foodweb.

Assigning interaction strengths to particular pathways is another approach to making foodwebs less qualitative. One can define strong or weak interactions within a foodweb that are not necessarily related to the flux of energy or material mediated by those interactions. A strong interaction is indicated if removal of one component species leads to pronounced change in the overall community (15). Alternately, removal of a species may cause little change, indicating a weak interaction. The definition of pronounced change is subjective; however, local extinctions and other large shifts in species diversity of a community would likely be acceptable demonstration that a pronounced change had occurred. For example, Paine (16) described two foodwebs for Pacific rocky intertidal communities, one for a coastal region of Washington State, the other for an Alaskan bay. The same 44 species or groups were present in both areas but the interaction strengths were quite different. In Washington, experimental removal of the starfish *Pisaster ochraceus* led to competitive elimination of multiple species of sessile invertebrates and benthic algae. The excluded species were replaced by the blue mussel, *Mytilus californicus*. When the starfish was present, it prevented the mussel from excluding the other species. Thus, the strong interaction between *Pisaster* and *Mytilus* was critical to maintaining a diverse intertidal community, although *Pisaster* preyed on many species and the bulk of its energy came through paths without *Mytilus*. Additional strong interactions within the foodweb were found, but removal of some species had little or no effect on the remaining community. By contrast, in Alaska, *Pisaster* only weakly interacted with other species.

A predator like *Pisaster* (in Washington), whose presence maintains community diversity by preventing competitive exclusion, is known as a “keystone” predator (17). Many other aquatic predators have since been described as keystone predators, including sea otters (18), mosquito larvae (19), and largemouth bass (20).

GENERAL PROPERTIES OF FOODWEBS

In the search for general rules of ecology, it is not surprising that compilations of documented foodwebs have been studied to determine whether foodwebs of different ecosystems share common properties. Certainly, the basic trophic structure with primary producers, one or more levels of consumers, and decomposers is common to vastly different ecosystems. Examples of other properties that might follow general rules include the ratio of predatory to prey species, the number of links per species, the fraction of total species in each trophic level, and the fraction of omnivorous species (21–23). Empirical foodwebs have also been investigated with respect to more abstract metrics, such as connectance—the ratio of actual links within a foodweb to the number of topologically possible links (21,24,25). However, the search for general properties within compilations of foodwebs has been criticized because, as mentioned above, the individual foodwebs used as raw data in such studies are frequently incomplete and overaggregated (6,7,26). The problems associated with documented foodwebs cannot be denied; however, it has been argued that some rules are sufficiently robust to overcome the biases (27).

The search for general rules relating to foodwebs is also approached using theoretical foodweb models. A theoretical foodweb uses a system of equations to represent the interactions of different organisms within an idealized ecosystem. Properties such as the number of links per species and connectance can be modified in different versions of the model. The stability of each resultant system of equations can be examined as various properties of the foodweb are modified (24,27,28). Theoretical foodwebs rapidly provide results for experiments that are difficult or impossible to conduct with real organisms. However, if the system of equations used does not adequately represent a natural condition, extrapolating results to the real world may be problematic.

Whether qualitative or quantitative, foodwebs are models of how the organisms within a community interact. They focus on feeding relationships, neglecting other modes of interaction, such as competition. Like any model, they are open to a variety of criticisms. It is important to recognize that it is not desirable for a model to contain all the complexities of nature; there would otherwise be no advantage to creating the model. Rather, a model should capture those aspects of nature that are necessary for understanding the question at hand. So long as that context is maintained, foodwebs will continue to be one of the most useful tools for teaching, studying, and understanding ecological communities.

BIBLIOGRAPHY

1. Landry, M.R. (1981). Switching between herbivory and carnivory by the planktonic marine copepod *Calanus pacificus*. *Mar. Biol.* **65**: 77–82.
2. Porter, X. (1988). Phagotrophic phytoflagellates in microbial food webs. *Hydrobiologia* **159**: 89–97.
3. Jeong, H.J. (1994). Predation by the heterotrophic dinoflagellate *Protoperidinium cf. divergens* on copepod eggs and early naupliar stages. *Mar. Ecol. Prog. Ser.* **114**: 203–208.
4. Jeong, H.J. (1994). Predation effects of the calanoid copepod *Acartia tonsa* on a population of the heterotrophic dinoflagellate *Protoperidinium cf. divergens* in the presence of co-occurring red-tide dinoflagellate prey. *Mar. Ecol. Prog. Ser.* **111**: 87–97.
5. Werner, E.E. (1984). The ontogenetic niche and species interactions in size-structured populations. *Annu. Rev. Ecol. System.* 393–425.
6. Tobin, J.E. (1991). Scientific correspondence. *Nature* **352**: 482.
7. Wardle, D.A. (1991). Scientific correspondence. *Nature* **352**: 482.
8. Pomeroy, L.R. (1974). The ocean’s food web, a changing paradigm. *Bioscience* **24**: 499–503.

9. Azam, F., Fenchel, T., Meyer-Reil, L., and Thingstad, F. (1983). The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* **10**: 257–263.
10. Paine, R.T. (1966). Food web complexity and species diversity. *Am. Nat.* **100**: 65–75.
11. Ryther, J.H. (1969). Photosynthesis and fish production in the sea. *Science* **166**: 72–76.
12. Pauly, D. and Christensen, V. (1995). Primary production required to sustain global fisheries. *Nature* **374**: 255–257.
13. Biktashev, V.N., Brindley, J., and Horwood, J.W. (2003). Phytoplankton blooms and fish recruitment rate. *J. Plankton Res.* **25**: 21–33.
14. Hakanson, L., Boulion, V.V., and Ostapenia, A.P. (2003). The influence of biomanipulations (fish removal) on the structure of lake food webs, case studies using the LakeWeb-model. *Aquat. Ecol.* **37**: 87–99.
15. MacArthur, R.H. (1972). Strong, or weak, interactions? *Trans. Conn. Acad. Arts Sci.* **44**: 177–188.
16. Paine, R.T. (1980). Food webs: linkage, interaction strength and community infrastructure. *J. Anim. Ecol.* **49**: 667–685.
17. Paine, R.T. (1969). A note on trophic complexity and community stability. *Am. Nat.* **103**: 91–93.
18. Estes, J.A. and Palmisano, J.F. (1975). Sea otters: their role in structuring nearshore communities. *Science* **185**: 1058–1060.
19. Cochran-Stafira, D.L. and von Ende, C.N. (1998). Integrating bacteria into food webs: studies with *Sarracenia purpurea inquilines*. *Ecology* **79**: 880–898.
20. Schindler, D.E., Hodgson, J.R., and Kitchell, J.F. (1997). Density-dependent changes in individual foraging specialization of largemouth bass. *Oecologia* **110**: 592–600.
21. Pimm, S.L., Lawton, J.H., and Cohen, J.E. (1991). Food web patterns and their consequences. *Nature* **350**: 669–674.
22. Havens, K. (1992). Scale and structure in natural food webs. *Science* **257**: 1107–1109.
23. Link, J. (2002). Does food web theory work for marine ecosystems? *Mar. Ecol. Prog. Ser.* **230**: 1–9.
24. May, R.M. (1972). Will a large complex system be stable? *Nature* **238**: 413–414.
25. Rejmanek, M. and Stary, P. (1979). Connectance in real biotic communities and critical values for stability of model ecosystems. *Nature* **280**: 311–313.
26. Paine, R.T. (1988). Food webs: road maps of interactions or grist for theoretical development? *Ecology* **69**: 1648–1654.
27. Pimm, S.L. and Lawton, J.H. (1977). Number of trophic levels in ecological communities. *Nature* **268**: 329–331.
28. Sugihara, G., Schoenly, K., and Trombla, A. (1989). Scale invariance in food web properties. *Science* **245**: 48–52.
29. Pimm, S.L. and Lawton, J.H. (1978). On feeding on more than one trophic level. *Nature* **275**: 542–544.

PLANKTON

Andrew Juhl
Lamont–Doherty Earth
Observatory of Columbia
University
Palisades, New York

INTRODUCTION

Plankton are aquatic organisms living in the water column that drift, float, or swim only weakly. The term derives from the Greek word *planktos*, meaning "to drift or "to wander." It is a descriptive classification, implying a certain type of lifestyle, rather than characteristics associated with taxonomy, size, or ecological role. The following information focuses on marine plankton, but much of the information is directly applicable to freshwater plankton.

Plankton can be contrasted with nekton, aquatic organisms capable of vigorous swimming (such as large fishes, squids, and marine mammals), and with benthos, those aquatic organisms that live on or at the bottom. It is generally thought that organisms classified as plankton do not have much ability to control their horizontal position relative to large-scale currents. However, there are no hard rules defining how strongly an organism may swim and still be considered planktonic. The boundaries between plankton, nekton, and benthos are nebulous. In fact, some organisms can be placed in each of the groups during different developmental stages. For example, the blue crab, *Callinectes sapidus*, has a planktonic larval stage (1). After 30+ days as plankton, these larvae settle to the bottom, becoming benthic juveniles (2). Although they spend most of their adult life on the bottom, blue crab adults swim surprisingly well (*Callinectes* translates to "beautiful swimmer") and females will traverse long distances to spawn (3). Despite a loose definition, plankton is a meaningful term that describes one of the most common aquatic lifestyles.

CATEGORIES OF PLANKTON

The organisms that comprise plankton are exceedingly diverse. Many Archaea and Eubacteria (the two major taxonomic groups of bacteria), most algal divisions, and most animal phyla have some representatives that are planktonic. Plankton also range in size from less than 1 μm (bacterioplankton) to several meters across (some jellyfish and colonies of planktonic tunicates). The range of taxa and sizes results in very different ecological roles for different plankton. Several categorization schemes are commonly used to organize the diversity of plankton into groups that share particular properties.

Phytoplankton, Zooplankton, and Bacterioplankton

Phytoplankton are photosynthetic plankton, mostly single-celled, microscopic algae that use light energy to fix carbon dioxide dissolved in water. This process makes solar energy available to them for their own biochemical needs. When they are eaten by other organisms, and when they release dissolved organic matter into the water, the energy becomes available to nonphotosynthetic organisms. Phytoplankton are therefore considered the base of most aquatic foodwebs (primary producers). Phytoplankton commonly include photosynthetic prokaryotes (cyanobacteria) and photosynthetic eukaryotes, such as diatoms, dinoflagellates, cryptophytes, and chlorophytes.

Zooplankton are eukaryotic, planktonic heterotrophs. They are usually phagotrophic, meaning that they ingest particles. However, there are examples of osmotrophic zooplankton (feeding on dissolved organic matter) (4).

Also, many eggs and larvae that are temporarily planktonic do not feed, but only use stored energy reserves. Both examples would still be considered zooplankton. Common zooplankton include single-celled, heterotrophic protists, such as ciliates and choanoflagellates, and multicellular animals, such as copepods and euphausiids.

Gelatinous zooplankton are a subcategory of zooplankton. Although taxonomically diverse, they share a body composition that is jellylike, mostly lacking hard, structural components. Their bodies have very high water content and are often transparent. Cnidarian medusae known as “jellyfish” are the most commonly known examples. Additional examples of gelatinous zooplankton are found in the ctenophores, mollusks (heteropods), and urochordates (planktonic tunicates such as salps, doliolids, and larvaceans). Although there are small gelatinous forms, the largest zooplankton are all gelatinous. Because of their fragile body plan, gelatinous zooplankton are difficult to collect quantitatively. Their importance to marine ecosystems may therefore be underestimated. Gelatinous zooplankton are rare in fresh water.

Bacterioplankton, or planktonic bacteria, includes representatives of both Eubacteria and Archaea. Photosynthetic bacteria are sometimes included within this category, creating an overlap with phytoplankton. Generally, however, the category implies heterotrophic bacteria. Heterotrophic bacterioplankton are osmotrophs. They either take up dissolved organic matter freely available in the water, or they use external enzymes to solubilize organic matter from particles, making it available for transport into their cells.

Holoplankton Versus Meroplankton

Organisms that are planktonic their entire lives are holoplankton. The term derives from the Greek word *holos*, meaning “entirely.” Meroplankton are organisms that are planktonic during only a portion of their lives. The Greek word *meros* means “mixed.” Most meroplankton are larval forms that are benthos or nekton as adults. For example, many fish larvae are planktonic. As they grow and their swimming ability improves, they eventually become nekton. Similarly, many mollusks have planktonic larvae that eventually settle to the bottom and become benthos. Because ecological conditions for the plankton are often very different from those facing the adults, meroplanktonic larvae often do not resemble the adults. In numerous cases, meroplanktonic larval stages were first scientifically described as distinct species; only later was the connection to the adult stage realized. Organisms that are entirely planktonic except for a benthic resting stage, such as some cladocerans or dinoflagellates, are still generally considered holoplanktonic.

Plankton are also commonly separated based on size. Viruses are abundant in aquatic environments and may be the smallest plankton, for those who consider viruses to be living. Bacterioplankton are the smallest cellular plankton and may be less than 1 μm in length. The largest plankton are cniderian medusae (jellyfish) that may be more than a meter across the bell, with tentacles that extend over many meters. Planktonic tunicate colonies (especially Pyrosomes) can also be many meters across, although the individuals within the colony are only millimeters to centimeters in length.

Size categories are arbitrary, but Sieburth (5) described the following categories currently used by many researchers: picoplankton (organisms 0.2–2 μm in length), nanoplankton (2–20 μm), microplankton (20–200 μm), mesoplankton (0.2–20 mm), macroplankton (2–20 cm), and megaplankton (20–200 cm). Viruses would be considered femtoplankton. These size categories are often combined with other categories. For example, picophytoplankton and mesozooplankton are common terms in oceanographic literature.

COLLECTING AND STUDYING PLANKTON

In general, the number of individuals per unit volume of water increases as planktonic organisms decrease in size. Nevertheless, natural waters are typically a very dilute suspension of even the smallest organisms and concentrating the organisms is often necessary for analyses. High numbers of small organisms (pico-, nano-, microplankton) can be collected simply by taking a water sample with a bucket or a hydrographic bottle. Concentrating such small organisms by filtration is often done after water collection. The filter pore size used determines the smallest organisms collected on the filter surface. Use of different filter pore sizes allows separation of the plankton sample into several size classes. Larger organisms (microplankton and up) may be too rare for quantitative collection in a whole-water sample. They are often collected by dragging plankton nets through the water. Plankton nets typically have a defined mesh size so that only organisms larger than the mesh will be collected. Plankton nets therefore size fractionate and concentrate before collection. A larger mesh net will filter a greater volume of water before it becomes clogged and ineffective. Larger mesh nets can also be dragged through the water more rapidly. It is therefore important to match the net mesh with the size, concentration, and swimming ability of the organisms one desires to collect. Because of their small size, taxonomic identification of most plankton requires microscopic examination after collection.

Buckets, bottles, nets, and microscopes are still used to study plankton, but these traditional approaches are now complemented with a myriad of advanced techniques. For example, on the smallest scales, methods of molecular biology can be used to study gene expression in individual cells of picoplankton (6). On the largest scales, the global distribution of phytoplankton can be studied in satellite images (7).

SPATIAL PATTERNS OF PLANKTON

The dominant large-scale horizontal pattern in plankton abundance is a decline as one goes away from the coast. One may have to travel tens to hundreds of miles offshore for this pattern to become clear. Nevertheless, the central areas of the oceans that are furthest from land generally have the lowest abundance of plankton, while coastal areas have the highest. This is true for all types of plankton (8–10). However, it should be kept in mind that many species of plankton are adapted to live in open-ocean environments. Therefore, certain planktonic species may actually be more abundant in the open ocean than nearer the coast. Exceptions to the general offshore pattern in plankton abundance can be caused by mesoscale eddies (11,12), oceanographic fronts (13,14), upwelling zones (15), and other phenomena (16). Although areas far from shore tend to have low plankton abundance, the diversity of species can be very high by comparison to coastal locations (17).

By definition, plankton have a limited ability to control their horizontal movement. Horizontal patterns of plankton abundance are therefore regulated by environmental factors external to the organism: currents, light, temperature, nutrient availability, and so on. These factors also contribute to vertical patterns in plankton abundance. However, individual behavior is also important in forming vertical patterns because many planktonic organisms do have the ability to control their vertical position within the water column.

Because light from the sun is rapidly absorbed by water, and because phytoplankton need light to photosynthesize, actively growing phytoplankton are restricted to the upper portion of the water column. In coastal waters, sufficient light for photosynthesis may only penetrate the top few meters (or less). In these conditions, vertical phytoplankton abundance often directly correlates with light; the highest phytoplankton abundance is at the surface, with a sharp decline as depth increases (9). This pattern is mostly caused by higher phytoplankton growth where more light is available. In addition, some phytoplankton can swim enough to change their vertical position in the water column (by as much as 10 m/d) (18). Similarly, other phytoplankton species are positively buoyant due to gas or lipid inclusions. Therefore, high phytoplankton abundance near the surface can sometimes be due to upward movement of phytoplankton (19,20). By contrast, in the open ocean, while sufficient light for photosynthesis is generally found within the upper 100–200 m, phytoplankton abundance is usually low near the surface. Highest phytoplankton abundance in open-ocean waters often occurs in a subsurface layer (9,21). Although light levels are always highest near the surface, phytoplankton also require inorganic nutrients to grow, just like other plants. In offshore regions, low phytoplankton abundance near the surface is generally attributable to low nutrient availability. Greater phytoplankton stocks can be supported at deeper depths where both nutrients and light are available.

Different phytoplankton species are often found at different depths (22,23) and may be specialized to live under the conditions associated with certain depth strata (24,25). However, vertical mixing from tides, winds, surface cooling, internal waves, and other physical processes can homogenize the vertical patterns of phytoplankton abundance and species composition.

Because zooplankton do not require light, they can be found at all ocean depths. However, greater zooplankton abundance is generally found in the surface waters where phytoplankton occur. This should not imply that zooplankton and phytoplankton abundances are necessarily directly related. While many zooplankton feed on phytoplankton, many feed on bacteria or other zooplankton. Correlations between zooplankton and phytoplankton abundance patterns can be direct and causal for some groups of zooplankton, that is, for zooplankton that feed on phytoplankton. However, other zooplankton are only indirectly, or coincidentally, associated with phytoplankton.

Many common mesozooplankton, especially crustacean mesozooplankton such as copepods and euphausiids, vertically migrate between the deeper and upper water layers on a daily basis. Such diel vertical migrations of zooplankton may cover several hundred meters each way (26). The most compelling explanation for this behavior is that the zooplankton migrate down during the day to avoid predation, then rise to the upper layers at night to feed (27,28). At night, migrating zooplankton are less likely to be detected by predators in the surface layers that depend on sight (such as fishes) to catch prey. Darkness, whether due to depth or time of day, provides one of the few opportunities to hide in the relatively featureless, planktonic habitat. Although daily vertical migrations are common in both fresh and marine waters throughout the world, it should be pointed out that most zooplankton species probably do not vertically migrate. Many species live out their lives associated with a particular depth strata (29). For example, Marlow and Miller (30) found that only around 10% of zooplankton species in the oceanic Gulf of Alaska vertically migrated. However, those vertically migrating species made up a large fraction of the total zooplankton biomass. Even species that do vertically migrate can change their migration pattern in response to local conditions (31) or over the course of development (32).

TEMPORAL PATTERNS OF PLANKTON

The abundance of plankton varies seasonally in many places, especially in higher latitude waters where seasonality is most pronounced. A pattern observed in many temperate lakes, estuaries, and the open North Atlantic Ocean is the spring phytoplankton bloom (33–35). The cycle begins with low phytoplankton abundance in winter. In spring or early summer, phytoplankton abundance increases dramatically (a bloom). The zooplankton biomass often also increases, although there may be a time lag between the increase in phytoplankton and the later increase in zooplankton. Eventually, the stocks of both phyto- and zooplankton decline to moderate levels that are sustained through the summer. A second bloom can occur in autumn, before plankton levels decline again for the winter.

Much research effort has focused on describing and understanding seasonal patterns of plankton exemplified by the spring-bloom cycle. It provides the best known temporal pattern in plankton stocks. However, in many lower latitude waters and the open North Pacific Ocean, seasonal cycles are either absent (36), subtle (37,38), or complicated by the impact of episodic events such as upwelling, mesoscale eddies, or storms (11,15,39).

Temporal patterns of plankton in polar seas are often dictated by the timing of sea ice melt (40,41).

GLOBAL IMPORTANCE OF PLANKTON

Although most planktonic organisms are individually small and seemingly ineffectual, plankton are important in aggregate because of their ubiquitous occurrence on this mostly water-covered planet. Outstanding examples of the global importance of plankton include their roles in the production of aquatic resources that humans value, in the Earth's climate, in oxygen production, and in sediment formation. Aquatic ecology and global biogeochemical processes cannot be understood without considering the role of plankton.

As the base of the aquatic foodweb, plankton are a critical food resource for aquatic organisms that are valuable to humans. The connection may be direct, as when valuable species such as oysters and anchovies feed on phytoplankton. The connection may also be indirect, for instance, if a commercially valuable predatory fish, such as a tuna, eats the anchovies. Either way, the occurrence of valued species is often coincident with the distribution of plankton. Large, motile species such as large fishes, seabirds, and whales concentrate their activities in locations where plankton are most abundant (42–44). This dependence of larger organisms on plankton means that commercial fisheries yield can be extrapolated from production of the plankton (45,46). Areas with higher plankton production have much greater fisheries yield.

Plankton influence the Earth's climate by removing carbon dioxide from the atmosphere. Carbon dioxide is the main atmospheric gas responsible for the "greenhouse" effect that regulates the Earth's temperature. Photosynthesis by phytoplankton fixes carbon dioxide from the atmosphere, transforming the carbon into organic compounds. Most of the carbon fixed by phytoplankton during photosynthesis is grazed by zooplankton and respired, releasing carbon dioxide back to the atmosphere (47,48). Nevertheless, some of the fixed carbon sinks to the deep ocean as the cells and bodies of plankton, zooplankton fecal matter, and other organic aggregates. Although the standing biomass of phytoplankton is less than 1% of the biomass of terrestrial plants (49), the amount of carbon fixed by phytoplankton and land plants is roughly equal (49,50). Because carbon fixation is coupled to the release of oxygen, phytoplankton are also responsible for approximately half of the Earth's oxygen production. It is estimated that one-third of the carbon fixed by phytoplankton eventually sinks to the deep ocean, effectively removing the carbon from the atmosphere for thousands of years (51). Determining the amount of carbon removed by plankton from the atmosphere and transported to the deep ocean, and predicting how that process may change in the future, are major questions in climate research.

Phytoplankton also influence climate by releasing sulfur compounds to the atmosphere. They are thought to be the main natural source of volatile sulfur compounds to the troposphere (52), mostly in the form of dimethylsulfide. The pathways leading to dimethylsulfide release from the oceans are incompletely understood, but phytoplankton primarily release the precursors to dimethylsulfide when their cells are ruptured during autolysis, viral attack, or grazing by zooplankton. Enzymes of phyto- and bacterioplankton then transform the released precursors to dimethylsulfide, which is released to the atmosphere. The sulfur aerosols stemming from dimethylsulfide of phytoplankton scatter incoming solar radiation. These aerosols also act as cloud condensation nuclei, further increasing atmospheric backscatter of solar radiation. Therefore, phytoplankton (and their interactions with other planktonic organisms) alter cloud formation and the amount of solar radiation reaching the Earth's surface (53).

A fraction of the carbon fixed by phytoplankton becomes a component of deep-sea sediments, removing it from the atmosphere for millions of years (54). Other inorganic structural components of plankton, such as the silicate tests of diatoms and radiolarians, and the calcareous shells of coccolithophores, pteropods, and foraminifera, also contribute to ocean bottom sediments. Over 60% of the ocean bottom is covered with sediments that are primarily planktonic in origin (55). Over geological time, such sea-bottom sediments, formed in previous ages, have been uplifted. Thus, even the peaks of mountains can be composed of material that originated from plankton.

BIBLIOGRAPHY

1. Epifanio, C.E. (1995). Transport of blue crab (*Callinectes sapidus*) larvae in waters off mid-Atlantic states. *Bull. Mar. Sci.* **57**: 713–725.
2. von Montfrans, J. et al. (1995). Settlement of blue crab postlarvae in Western North Atlantic estuaries. *Bull. Mar. Sci.* **57**: 834–854.
3. Tankersley, R.A., Wieber, M.G., Sigala, M.A., and Kachurak, K.A. (1998). Migratory behavior of ovigerous blue crabs *Callinectes sapidus*: evidence for selective tidal-stream transport. *Biol. Bull.* **195**: 168–173.
4. Manahan, D.T. (1990). Adaptations by invertebrate larvae for nutrient acquisition from seawater. *Am. Zool.* **30**: 147–160.
5. Sieburth, J.M. (1979). *Sea Microbes*. Oxford University Press, New York.
6. Bird, C. and Wyman, M. (2003). Nitrate/nitrite assimilation system of the marine picoplanktonic cyanobacterium *Synechococcus* sp. strain WH 8103: effect of nitrogen source and availability on gene expression. *Appl. Environ. Microbiol.* **69**: 7009–7018.
7. Yoder, J.A. and Kenelly, M.A. (2003). Seasonal and ENSO variability in global ocean phytoplankton chlorophyll derived from 4 years of SeaWiFS measurements. *Global Biogeochem. Cycles* **17**: 1112.
8. Reid, J.L., Brinton, E., Fleminger, A., Venrick, E.L., and McGowan, J.A. (1978). Ocean circulation and marine life. In: *Advances in Oceanography*. H. Charnock and G. Deacon (Eds.). Plenum Publishing, New York, pp. 65–130.
9. Mullin, M.M. (1986). Spatial and temporal scales and patterns. In: *Plankton Dynamics of the Southern California Bight*. R.W. Eppley (Ed.). Springer, New York, pp. 216–273.
10. Cavender-Bares, K.K., Karl, D.M., and Chisholm, S.W. (2001). Nutrient gradients in the western North Atlantic Ocean: relationship to microbial community structure and

comparison to patterns in the Pacific Ocean. *Deep-Sea Res. I* **48**: 2373–2395.

11. The Ring Group. (1981). Gulf Stream cold-core rings: their physics, chemistry, and biology. *Science* **212**: 1091–1100.
12. Falkowski, P.G., Ziemann, D., Kolber, Z., and Bienfang, P.K. (1991). Role of eddy pumping in enhancing primary production in the ocean. *Nature* **352**: 55–58.
13. Marra, J., Houghton, R.W., and Garside, C. (1990). Phytoplankton growth at the shelf-break front in the Middle-Atlantic Bight. *J. Mar. Res.* **48**: 851–868.
14. Franks, P.J.S. (1992). Sink or swim: accumulation of biomass at fronts. *Mar. Ecol. Prog. Ser.* **82**: 1–12.
15. Small, L.F. and Menzies, D.W. (1981). Patterns of primary productivity and biomass in a coastal upwelling region. *Deep-Sea Res.* **28A**: 123–149.
16. Ryan, J.P., Polito, P.S., Strutton, P.G., and Chavez, F.P. (2002). Unusual large-scale phytoplankton blooms in the equatorial Pacific. *Prog. Oceanogr.* **55**: 263–285.
17. McGowan, J.A. and Walker, P.W. (1993). Pelagic diversity patterns. In: *Species Diversity Patterns in Ecological Communities*. R.E. Ricklefs and D. Schluter (Eds.). University of Chicago Press, Chicago, IL, pp. 203–214.
18. Kamykowski, D. (1995). Trajectories of autotrophic marine dinoflagellates. *J. Phycol.* **31**: 200–208.
19. Eppley, R.W., Holm-Hansen, O., and Strickland, J.D.H. (1968). Some observations on the vertical migration of dinoflagellates. *J. Phycol.* **4**: 333–340.
20. Ault, T.R. (2000). Vertical migration by the marine dinoflagellate *Prorocentrum triestinium* maximises photosynthetic yield. *Oecologia* **125**: 466–475.
21. Cullen, J.J. and Eppley, R.W. (1981). Chlorophyll maximum layers of the Southern California Bight and possible mechanisms of their formation and maintenance. *Oceanol. Acta* **4**: 23–32.
22. Venrick, E.L. (1982). Phytoplankton in an oligotrophic ocean: observations and questions. *Ecol. Monogr.* **52**: 129–154.
23. Chisholm, S.W., Olsen, R.J., Zettler, E.R., Goericke, R., Waterbury, J.B., and Welschmeyer, N.A. (1988). A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334**: 340–343.
24. Takahashi, M., Ichimura, S., Kishino, M., and Okami, N. (1989). Shade and chromatic adaptation of phytoplankton photosynthesis in a thermally stratified sea. *Mar. Biol.* **100**: 401–409.
25. Ferris, M.J. and Palenik, B. (1998). Niche adaptation in ocean cyanobacteria. *Nature* **396**: 226–228.
26. Enright, J.T. (1977). Copepods in a hurry: sustained high-speed upward migration. *Limnol. Oceanogr.* **22**: 118–125.
27. Bollens, S. and Frost, B.W. (1989). Predator-induced diel vertical migration in a planktonic copepod. *J. Plankton Res.* **11**: 1047–1065.
28. DeRobertis, A., Jaffe, J.S., and Ohman, M.D. (2000). Size-dependent visual predation risk and the timing of vertical migration in zooplankton. *Limnol. Oceanogr.* **45**: 1838–1844.
29. McGowan, J.A. and Walker, P.W. (1979). Structure in the copepod community of the North Pacific Central gyre. *Ecol. Monogr.* **49**: 195–226.
30. Marlow, C.J. and Miller, C.B. (1975). Patterns of vertical distribution and migration of zooplankton at Ocean Station 'P'. *Limnol. Oceanogr.* **20**: 824–844.
31. Ohman, M.D. (1990). The demographic benefits of diel vertical migration by zooplankton. *Ecol. Monogr.* **60**: 257–281.
32. Osgood, K.E. and Frost, B.W. (1994). Ontogenetic diel vertical migration behaviors of the marine planktonic copepods *Calanus pacificus* and *Metridia lucens*. *Mar. Ecol. Prog. Ser.* **104**: 13–25.
33. Sverdrup, H.U. (1953). On the conditions for the vernal blooming of phytoplankton. *J. Cons. Int. Explor. Mer.* **18**: 287–295.
34. Heinrich, A.K. (1962). The life histories of plankton animals and seasonal cycles of plankton communities in the oceans. *J. Cons. Int. Explor. Mer.* **27**: 15–24.
35. Lund, J.W.G. (1965). The ecology of the freshwater phytoplankton. *Biol. Rev.* **40**: 231–293.
36. Hayward, T.L., Venrick, E.L., and McGowan, J.A. (1983). Environmental heterogeneity and plankton community structure in the central North Pacific. *J. Mar. Res.* **41**: 711–729.
37. Frost, B.W. (1993). A modelling study of processes regulating plankton standing stock and production in the open subarctic Pacific Ocean. *Prog. Oceanogr.* **32**: 17–56.
38. Venrick, E.L. (1993). Phytoplankton seasonality in the central North Pacific: the endless summer reconsidered. *Limnol. Oceanogr.* **38**: 1135–1149.
39. Chang, J., Chung, C., and Gong, G. (1996). Influences of cyclones on chlorophyll *a* concentration and *Synechococcus* abundance in a subtropical western Pacific coastal ecosystem. *Mar. Ecol. Prog. Ser.* **140**: 199–205.
40. Park, G.M., Yang, S.R., Kang, S-H., Chung, K.H., and Shim, J.H. (1999). Phytoplankton biomass and primary production in the marginal ice zone of the northwestern Weddell Sea during austral summer. *Polar Biol.* **21**: 251–261.
41. Rat'kova, T.N. and Wassman, P. (2002). Seasonal variation and spatial distributions of phyto- and protozooplankton in the central Barents Sea. *J. Mar. Sys.* **38**: 47–75.
42. Hyrenbach, K.D., Fernandez, P., and Anderson, D.J. (2002). Oceanographic habitats of two sympatric North Pacific albatrosses during the breeding season. *Mar. Ecol. Prog. Ser.* **233**: 283–301.
43. Baumgartner, M.F. and Mate, B.R. (2003). Summertime foraging ecology of North Atlantic right whales. *Mar. Ecol. Prog. Ser.* **264**: 123–135.
44. Beaugrand, G. and Reid, P.C. (2003). Long-term changes in phytoplankton, zooplankton and salmon related to climate. *Global Change Biol.* **9**: 801–817.
45. Ryther, J.H. (1969). Photosynthesis and fish production in the sea. *Science* **166**: 72–76.
46. Pauly, D. and Christensen, V. (1995). Primary production required to sustain global fisheries. *Nature* **374**: 255–257.
47. Calbet, A. (2001). Mesozooplankton grazing effect on primary production: a global comparative analysis in marine ecosystems. *Limnol. Oceanogr.* **46**: 1824–1830.
48. Calbet, A. and Landry, M.R. (2004). Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnol. Oceanogr.* **49**: 51–57.
49. Field, C.B., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P.G. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**: 237–240.
50. Longhurst, A., Sathyendranath, S., Platt, T., and Caverhill, C. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.* **17**: 1245–1271.
51. Falkowski, P.G., Barber, R.T., and Smetacek, V. (1998). Biogeochemical controls and feedbacks on ocean primary production. *Science* **281**: 200–206.

52. Simo, R. (2001). Production of atmospheric sulfur by oceanic plankton: biogeochemical, ecological and evolutionary links. *Trends Ecol. Evol.* **16**: 287–294.

53. Steinke, M., Malin, G., and Liss, P.S. (2002). Trophic interactions in the sea: an ecological role for climate relevant volatiles? *J. Phycol.* **38**: 630–638.

54. Raven, J.A. and Falkowski, P.G. (1999). Oceanic sinks for atmospheric CO_2. *Plant Cell Environ.* **22**: 741–755.

55. Kennett, J.P. (1982). *Marine Geology*. Prentice-Hall, Englewood Cliffs, NJ.

MAJOR IONS IN SEAWATER

Marie de Angelis
Humboldt State University
Arcata, California

Seawater contains all of the natural elements. However, its composition is dominated by relatively few ionic species. These species, which comprise >99.8% of the mass of the total dissolved salts in the ocean, are known as the major ions and are defined as dissolved species having a concentration in seawater of >1 mg kg^{-1} (1 ppm) (Table 1). Na^+ and Cl^- alone account for 86% of all dissolved chemical species in seawater.

The source of Cl^-, and most anions in seawater, is *outgassing* of the Earth's interior, or *volcanic emissions*. The majority of the major cations are derived from the effects of continental weathering and are delivered to the ocean via rivers. The gross composition of seawater is the result of the partitioning of elements between continental rock and seawater over geologic time. The major ions represent the more soluble elements, which have preferentially partitioned into seawater.

The major ions are found to occur in nearly constant ratios to each other throughout most of the world's oceans. That is, although the salinity or the total amount of salt dissolved in seawater varies from location to location in the ocean, the ratios of the major ions to each other remains nearly constant, which is true from ocean to ocean as well as from surface to deep waters (Table 2).

Table 1. Major Ions in Seawater

Major Ion	g kg^{-1} Seawater at Salinity of 35.000 ppt	% by wt	g ion/g Cl^-
Chloride, Cl^-	19.353	55.30	1
Sodium, Na^+	10.781	30.77	0.5561
Sulfate, $SO_4{}^{2-}$	2.712	7.75	0.1400
Magnesium, Mg^{2+}	1.284	3.69	0.0668
Calcium, Ca^{2+}	0.4119	1.18	0.0213
Potassium, K^+	0.399	1.14	0.0206
Bicarbonate, $HCO_3{}^{2-}$	0.126	0.41	0.0075
Bromide, Br^-	0.0673	0.19	0.0035
Boron, B (as $B(OH)_3$	0.0257	0.013	0.00024
Strontium, Sr^{2+}	0.00794	0.023	0.00041
Fluoride, F^-	0.00130	0.0037	0.00006

Concentrations are presented in terms of g of ion per kg of seawater, % by weight in seawater, and the weight ratio g ion per g Cl^-.

Table 2. Ion to Chloride Ratios in Various Oceans and Selected Seas for Na^+ and K^+

Ocean or Depth Interval	Na^+/Cl^-	K^+/Cl^-
Atlantic	0.5552	0.0206
Pacific	0.5555	0.0206
Indian	0.5554	0.0207
Red Sea	0.5563	0.0206
Mediterranean	0.5557	0.0206
0–100 m	0.5554	0.0206
700–1500 m	0.5557	0.0206
>1500 m	0.5555	0.0206
MEAN	0.5555$^+$/−0.0007	0.0206$^+$/−0.0002

Also shown are ion to chloride ratios for selected depth intervals in the world ocean. Mean values are for the world ocean over all depth intervals.

This observation is known as "The Rule of *Constant Proportions*" or conservative behavior. As Cl^- is the single most abundant ionic species in seawater, the Rule of Constant Proportion is usually expressed as the ratio of major ions to chloride ion (e.g., Na^+/Cl^-). Exceptions to this rule exist for calcium (Ca^{2+}), strontium (Sr^{2+}), and bicarbonate ($HCO_3{}^{2-}$) as a small fraction of the total concentration of these species participate in biological reactions resulting in slight variations of ion/Cl^- ratios between surface and deep water.

The *residence time* (τ) of a chemical species in the ocean can be defined as the average time an individual atom for a given element remains in seawater before being permanently removed. Assuming that the major source of dissolved salts to the ocean is from riverine input, residence time of a given ion can be calculated as the total mass (in g) of ion in the ocean (M_{ocean}) divided by the mass of ion delivered by rivers annually ($Flux_{in}$) (in g yr^{-1}).

$$\tau\,(\text{years}) = \frac{M_{ocean}}{Flux_{in}}$$

Flux $_{in}$ can be estimated from the average concentration of the ion in the world's rivers (in g L^{-1}) multiplied by the total volume of river water entering the ocean annually (in L yr^{-1}). M_{ocean} can be estimated from mean seawater concentrations (in g L^{-1}) multiplied by the volume of the ocean (in L). The residence times for the major ions calculated by this method are extremely large, on the order of millions to hundreds of millions of years (Table 3).

The long residence times of major ions reflects the relatively low reactivity of these chemical species. Except

Table 3. Mean Seawater and River Water Concentrations and Residence Times for Several Major Ions

Species	Mean SW (mg L^{-1})	Mean RW (mg L^{-1})	$\frac{SW}{RW}$	τ (10^6 years)
Cl^-	19, 350	5.75	3, 365	123
Na^+	10, 760	5.25	2, 090	75
Mg^{2+}	1, 294	3.35	386	14
$SO_4{}^{2-}$	2, 712	8.25	338	12
K^+	399	1.3	307	11
Ca^{2+}	412	13.4	31	1.1
Br^-	67	0.02	3, 350	123

for Ca^{2+}, Sr^{2+}, and $HCO_3{}^{2-}$, no biological processes exist that accelerate the removal of these species to the oceanic sediment. The main removal mechanisms for major ions are kinetically slow reactions such as precipitation or incorporation into clay minerals in the sediment or oceanic crust. Thus, major ions are conservative or remain in constant proportion to each other because the major ions have very long residence times relative to either the residence time of water in the ocean (40,000 years) or the oceanic circulation time (1000 years). Even if the salt content (salinity) of seawater changes from location to location in the ocean, the ratios of the major ions do not change because they react on much longer time scales than the removal or addition of water by precipitation or evaporation. As the residence time of major ions greatly exceeds the average oceanic circulation time of approximately 1000 years, no change is seen in major ion concentrations from ocean to ocean or between surface water and deep water

The major ion composition of river water and seawater are different (Table 3). The composition of present-day river water reflects the weathering of continental rock, whose composition reflects the long-term partitioning of chemical species between the continents and seawater. The composition of seawater with respect to its major ion composition is not believed to have changed over the past several million years despite the constant input of river water of different composition. The oceans are believed to be at steady-state equilibrium with respect to the major ions. That is, the composition of seawater remains constant because the flux of major ions into the ocean equals the flux of major ions out of the ocean.

Although the major source of major ions is river input, hydrothermal circulation of seawater through oceanic basalt, particularly at relatively low temperatures, may be a significant source of some major cations (Na^+, K^+). To achieve steady-state equilibrium, the sum of the sinks or fluxes out of major ions must balance the input from rivers. Several processes, most occurring at extremely slow rates over large areas of the oceanic seafloor, have been identified as sinks for major ions. These processes include: (a) Cation exchange in which clay minerals within the oceanic sediment as well as clays delivered to the oceans by rivers exchange cations in seawater to form new clay minerals relatively enriched in Na, K, and Mg at the expense of Ca. (b) Trapping and eventual burial of seawater within interstitial water of marine sediments. This process removes the more concentrated ions in seawater (Na, Cl). (c) Evaporite formation during some periods of Earth's history, in which large deposits of minerals derived from seawater have formed when seawater was trapped and evaporated from shallow, closed basins. Although the areal extent of such basins is limited at the present time, formation of evaporites is an important removal mechanism for some major ions, including Na, Cl, and $SO_4{}^{2-}$, over geologic time. (d) *Sea spray* transported to land can result in net removal of Na and Cl. (e) *Reverse weathering* involving reactions between ions in seawater and cation-poor aluminosilicates derived from continental weathering. This process results in the formation of new cation-rich clay minerals and CO_2, resulting in the net removal of Na, K, and Mg from seawater. (f) *Hydrothermal circulation* involving the reaction of major ions between seawater and oceanic basalt is a net sink for Mg^{2+} and $SO_4{}^{2-}$.

READING LIST

Mackenzie, F.T. and Garrels, R.M. (1966). Chemical balance between rivers and oceans. *Amer. J. Sci.* **264**: 507–525.

Holland, H.D. (1984). *The Chemical Evolution of the Atmosphere and Oceans*. Princeton University Press, Princeton, NJ, p. 584.

Pilson, M.E.Q. (1998). *An Introduction to the Chemistry of the Sea*. Prentice-Hall, Englewood Cliffs, NJ, p. 431.

TSUNAMI

Tsunami (also called Seismic Sea Wave, and popularly, Tidal Wave), an ocean wave produced by a submarine earthquake, landslide, or volcanic eruption. These waves may reach enormous dimensions and have sufficient energy to travel across entire oceans. From the area of the disturbance, the waves will travel outward in all directions, much like the ripples caused by throwing a rock into a pond. The time between wave crests may be from 5 to 90 minutes, and the wave speed in the open ocean will average 450 miles per hour.

Tsunamis reaching heights of more than 100 feet have been recorded. As the waves approach the shallow coastal waters, they appear normal and the speed decreases. Then, as the tsunami nears the coastline, it may grow to great height and smash into the shore, causing much destruction.

1. Tsunamis are caused by an underwater disturbance—usually an undersea earthquake. Landslides, volcanic eruptions, and even meteorites can also generate a tsunami.
2. Tsunamis can originate hundreds or even thousands of miles away from coastal areas. Local geography may intensify the effect of a tsunami. Areas at greatest risk are less than 50 feet above sea level and within one mile of the shoreline.
3. People who are near the seashore during a strong earthquake should listen to a radio for a tsunami warning and be ready to evacuate at once to higher ground.
4. Rapid changes in the water level are an indication of an approaching tsunami.
5. Tsunamis arrive as a series of successive "crests" (high water levels) and "troughs" (low water levels). These successive crests and troughs can occur anywhere from 5 to 90 minutes apart. They usually occur 10 to 45 minutes apart.

The Tsunami Warning System, a cooperative international organization and operated by the United States Weather Service, has been in operation since the 1940s. The headquarters of the center is located in Hawaii. An associated Alaska Regional Tsunami Warning System is located in Alaska. Tsunami prediction essentially

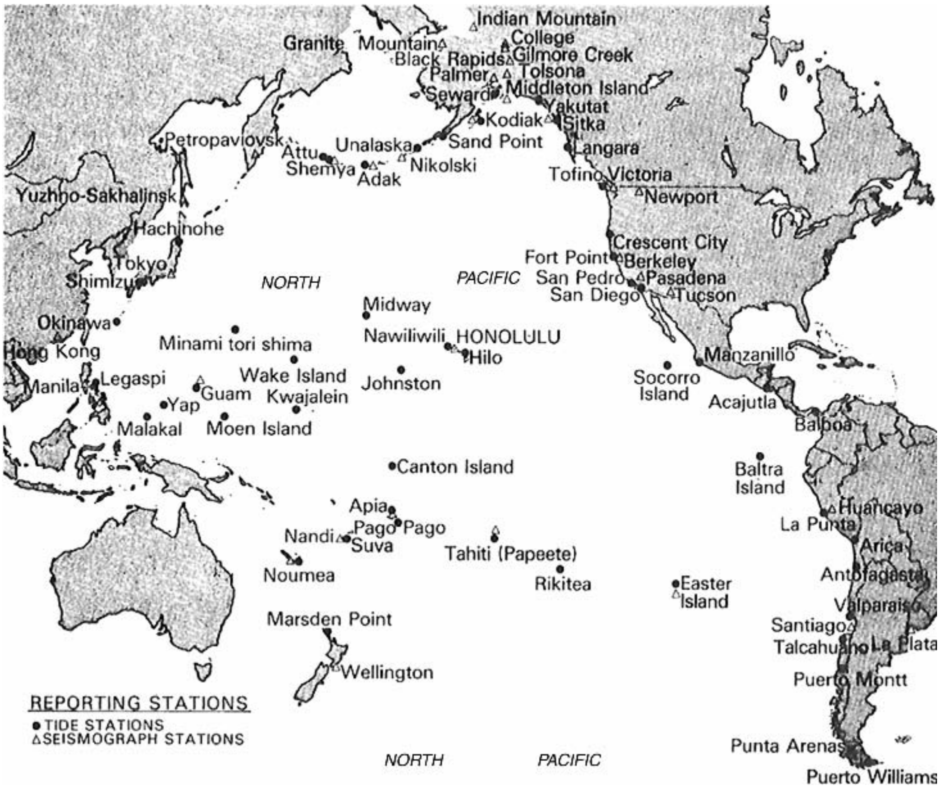

Figure 1. Network of tide and seismograph stations that are part of the Tsunami Warning System, headquartered in Hawaii. (*National Oceanic and Atmospheric Administration.*).

commences with earthquake monitoring and prediction information. Inputs from these systems are linked with information from a series of tide monitoring installations. Locations of tide stations and of seismograph stations in the Tsunami Warning System are shown in Fig. 1.

When inputs indicate conditions are favorable for a tsunami, a watch is issued for the probable affected area. Warnings are issued when readings from various tidal stations appear to match the seismographic information. Because of the complexity of the factors involved and a large degree of uncertainty nearly always present, there is a tendency to issue watches and warnings as a safety precaution even though a tidal wave of significance may not develop. Unfortunately, after awhile, persons in likely areas to be affected grow callous to watches and tend to ignore them. A tsunami that hit Hilo, Hawaii in 1960 killed 60 residents even though they had been warned of the coming event an ample 6 hours in advance of the strike.

This, however, was not the reason for the most devastating Tsunami in history which occurred on December 26, 2004 in the Indian Ocean as a result of a 9.0 earthquake off the shore of Banda Aceh, Indonesia. The resulting Tsunami quickly hit the Indonesian coastline, but hours later struck Sri Lanka, India, The Maldives, and later yet, Keyna and Somalia on the East coast of Africa. More than 150,000 lives were lost because no network existed to communicate the likely result of the 9.0 quake in the Indian Ocean. No such event had occurred in this area of the world in over 120 years since Kakatoa erupted in 1883.

Twenty three earthquake monitoring stations picked up the seismic shocks in Indonesia itself, and the U.S. Geological Survey's worldwide monitoring system, with 120 stations, pinpointed the quake immediately. However, on that fateful Sunday, the few warning telephone calls that were made went unanswered. Thus, lack of communication, more than lack of technology, caused the extreme loss of life.

The material destruction, however, touching so many countries, was a result of the movement of the Earth's crust beneath the Indian Ocean extending along a fracture believed to have been 600 miles in length, ultimately creating a tidal wave nearly 1000 miles long.

Such horror can never again touch the Indian Ocean nations as they continue to build a Tsunami Watch System containing all the elements of the early warning capability that has existed in the Pacific Ocean since the middle of the 20th century.

In the late 1970s, scientists suggested an improved method for making tsunami predictions. For a number of years, specialists have suggested that better analysis and interpretation of seismic waves produced by earthquakes may improve the prediction of tsunamis. Seismic waves range from very short-period waves that result from the sharp snap of rocks under high stress to very long-period waves, due mainly to the slower movements of large sections of the ocean floor. Many researchers believe that tsunamis result mainly from the vertical movement of these large blocks, leading to a tentative conclusion that the strength of seismic waves of very long period may be the best criterion for an earthquake's ability to generate a tsunami.

Part of the problem is that most seismographs installed in the system are not very sensitive to very long-period waves, and thus a given earthquake cannot be analyzed effectively in terms of its potential for producing a tsunami. Equipment has been refined so that, today, shorter period waves are used to locate an earthquake and 20-second waves are used to calculate the magnitude. However, some scientists feel that the true magnitude of some earthquakes can only be determined by measuring the characteristics of longer waves, such as 100-second waves. Brune and Kanamori (University of California at San Diego) have observed that the Chilean earthquake of 1960 had a magnitude of 8.3 when calculated on the basis of 20-second waves, but its magnitude was 9.5, or more than 10 times larger in wave amplitude and more than 60 times larger in energy released, when calculated by Kanamori's method, which attempts to include the energy release represented by very long-period seismic waves. Other scientists are coming to the viewpoint that many warnings could be omitted if predictions were based on longer waves. Two of the first long-period seismographs incorporated in the Tsunami Warning System were installed in Hawaii and on the Russian island of Yuzhno-Sakhalinsk, which is northeast of Vladivostok. Later, an installation was made at the Alaska Warning Center in Palmer. Whether the latest reasoning proves successful must await a number of years of experience with the earthquakes of the future and the resulting tsunamis.

CONTINUING TSUNAMI RESEARCH

During the 2000s, research pertaining to the fundamentals of tsunamis and the development of mathematical models of the phenomenon continues. Considerable attention is being directed to specific regions, including the west coasts of Mexico and Chile, the southwestern shelf of Kamchatka (Russia) and, in the United States, the generation of tsunamis in the Alaskan bight and in the Cascadia Subduction Zone off the west coasts of Washington (Puget Sound) and Oregon.

Research also is being directed toward the development of simple and more economic warning systems, particularly in the interest of the developing countries, the coasts of which border on the Pacific Ocean.

The Tsunami Warning System previously described requires millions of dollars for equipment, maintenance, and operation, well beyond the means of some countries. Also, some scientists believe that more localized equipment installations could possibly serve local shore communities better while costing less. These observations, however, do not challenge the need and validity for the larger tsunami network. The National Oceanic and Atmospheric Administration (NOAA) has developed a system costing in the range of $20,000 that can be installed and operated by non experts. The system has been undergoing trials at Valparaiso, Chile, a city that has been struck by nearly 20 tsunamis within the past two centuries. A sensor (accelerometer) is installed in bedrock under the city and can measure tectonic activity in excess of 7.0 on the Richter scale. These measurements are interlocked with level sensors.

Researchers B.F. Atwater and A.L. Moore (University of Washington), in their attempts to model earthquake and tsunami activity in the area over the last thousand years, have reported what they believe to have been a large earthquake on the Seattle fault some time between 1000 and 1100 years ago. The researchers report, "Water surged from Puget Sound, overrunning tidal marshes and mantling them with centimeters of sand. One overrun site is 10 km northwest of downtown Seattle; another is on Whidbey Island, some 30 km farther north. Neither site has been widely mantled with sand at any other time in the past 2000 years. Deposition of the sand coincided—to the year or less—with abrupt, probably tectonic subsidence at the Seattle site and with landsliding into nearby Lake Washington. These findings show that a tsunami was generated in Puget Sound, and they tend to confirm that a large shallow earthquake occurred in the Seattle area about 1000 years ago."

Simulations of tsunamis from great earthquakes on the Cascadia subduction zone have been carried out by M. Ng, P.H. Leblond, and T.S. Murty (University of British Columbia). A numerical model has been used to simulate and assess the hazards of a tsunami generated by a hypothetical earthquake of magnitude 8.5 associated with rupture of the northern sections of the subduction zone. The model indicates that wave amplitudes on the outer coast are closely related to the magnitude of sea-bottom displacement (5 meters). The researchers observe, "Some amplification, up to a factor of 3, may occur in some coastal embayments. Wave amplitudes in the protected waters of Puget Sound and the Straits of Georgia are predicted to be only about one-fifth of those estimated on the outer coast."

READING LIST

Atwater, B.F. and Moore, A.L (December 4, 1991). A tsunami about 1000 years ago in Puget Sound, Washington. *Science*: 1614.

Bernard, E.N. (Ed.). (July 1989). *Tsunami Hazard: A Practical Guide for Tsunami Hazard Reduction*, (Papers from Symposium at Novosibirsk, Russia). International Union of Geodesy and Geophysics, New York.

Bryant, E. (2001). *Tsunami: The Underrated Hazard*. Cambridge University Press, New York.

Collins, E. (February 1988). Wave Watch. *Sci. Amer:* 28.

Kubota, I. (Spring 1987). Japan's Weather Service and the Sea. *Oceanus*: 71.

Lander, J.F., Lockridge, P.A., and Kozuch, M.J. (1997). *Tsunamis Affecting the West Coast of the United States, 1806–1992*. DIANE Publishing Company, Collingdale, PA.

Lander, J.F. (1997). *Tsunamis Affecting Alaska, 1737–1996*. DIANE Publishing Company, Collingdale, PA.

Ng, M.K-F., Leblond, P.H., and Murty, T.S. (November 30, 1990). Simulation of Tsunamis from Great Earthquakes on the Cascade Subduction Zone. *Science*: 1248.

Prager, E.J. (1999). *Furious Earth: The Science and Nature of Earthquakes, Volcanoes, and Tsunamis*. McGraw-Hill Professional Book Group, New York.

Soloviev, S.L., Kim, K.S., Solovieva, O.N. et al. (2000). *Tsunamis in the Mediterranean Sea, 2000 B.C.–2000 A.D.* Kluwer Academic Publishers, Norwell, MA.

Staff. (1997). *Long-Wave Runup Models*. World Scientific Publishing Company, Inc., River Edge, NJ.

Tsuchiya, Y. and Shuto, N. (1995). *Tsunami: Progress in Prediction, Disaster Prevention and Warning*. Kluwer Academic Publishers, Norwell, MA.

WEB REFERENCES

International Tsunami Information Center (ITIC): http://www.shoa.cl/oceano/itic/frontpage.html.

National Tsunami Hazard Mitigation Program: http://www.pmel.noaa.gov/tsunami/.

Tsunami Links: http://www.pmel.noaa.gov/tsunami-hazard/links.html.

BALLOONING AND METEOROLOGY IN THE TWENTIETH CENTURY

Linda Voss
U.S. Centennial of Flight
Commission

Balloons are ideal for gathering meteorological information and have been used for that purpose throughout their history. *Meteorological* measurements of wind and air pressure have gone hand in hand with the earliest balloon launches and continue today. Balloons can climb through the denser air close to the Earth to the thinner air in the upper atmosphere and collect data about wind, the different layers of the atmosphere, and weather conditions as they travel.

The first meteorological balloon sondes, or "registering balloons," were flown in France in 1892. These balloons were relatively large, several thousand cubic feet, and carried instruments to record barometric pressure (barometers), temperature (thermometers), and humidity (hygrometers) data from the upper atmosphere. They were open at the base of the balloon and were inflated with a lifting gas, which could be hydrogen, helium, ammonia, or methane. The lifting gas in the balloon exited through the opening as the balloon expanded during its ascent and the air became thinner and the pressure dropped. At the end of the day, as the lifting gas cooled and took up less space, the balloon descended very slowly. The meteorologists had to wait until the balloon descended all the way to Earth to

Two men performing balloon tests for the U.S. Weather Bureau.

A balloon equipped for meteorological observations. A German balloon ascent in the late 1800s. 17 Balloon Equipped for Meteorological Observations.

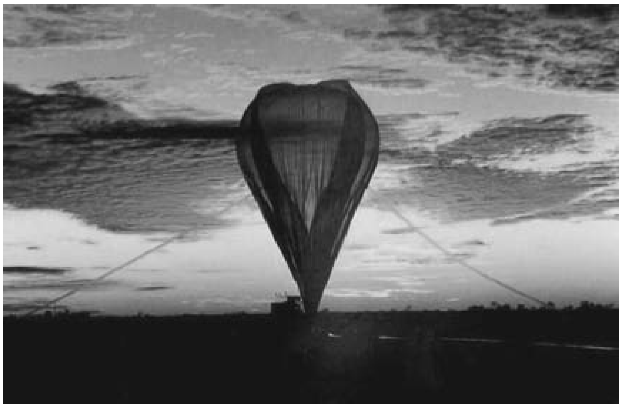

A zero-pressure balloon being inflated at Alice Springs, Australia.

Weather balloons are used daily to carry meteorological instruments to 20 miles (30 kilometers) and above into the atmosphere to measure temperature, pressure, humidity, and winds. The balloons are made of rubber and weigh up to 2.2 pounds (one kilogram). More than 200,000 weather soundings are made with such balloons worldwide each year.

FIG. 2. FIRST ASCENSION IN AMERICA, ST. LOUIS, SEPTEMBER 15, 1904

Preparing to launch America"s first "ballon-sonde." Since this first launch on September 15, 1904, in St. Louis, Missouri, literally millions of weather balloons have been launched by the National Weather Service and its predecessor organization. From: The Principles of Aerography, by Alexander McAdle, 1917.

retrieve their instruments, which often had drifted up to 700 miles (1,126 kilometers) from their launch point.

The German meteorologist Assmann solved the problem of drifting balloons and retrieval of instruments in 1892 by introducing closed rubber balloons that burst when they reached a high altitude, dropping the instruments to Earth by parachute much closer to the launch site. These balloons also had fairly constant rates of ascent and descent for more accurate temperature readings. Assmann also invented a psychrometer, a type of hygrometer used to measure humidity in the air that laboratories generally use.

In the 1930s, meteorologists were able to get continuous atmospheric data from balloons when the radiosonde was developed. A radiosonde is a small, radio transmitter that broadcasts or radios measurements from a group of instruments. Balloons, usually unmanned, carry the transmitter and instruments into the upper atmosphere. The radiosonde transmits data to Earth while measuring humidity, temperature, and pressure conditions.

Today, three types of balloons are commonly used for meteorologic research.

Assmann's rubber, or neoprene, balloon is used for measuring vertical columns in the atmosphere, called vertical soundings. The balloon, inflated with a gas that causes the balloon to rise, stretches as it climbs into thin air, usually to around 90,000 feet (27,400 meters). Data is taken as the balloon rises. When the balloon has expanded from three to six times its original length (its volume will have increased 30 to 200 times its original amount), it bursts. The instruments float to Earth under a small parachute. The neoprene balloon can either carry radiosondes that transmit meteorological information or be tracked as a pilot balloon, a small balloon sent aloft to show wind speed and direction. Around the world, balloons equipped with radiosondes make thousands of soundings of the winds, temperature, pressure, and humidity in the upper atmosphere each day. But these balloons are launched and tracked from land, which limits what the radiosondes can measure to less than one-third of the Earth's surface.

Zero-pressure plastic (usually polyethylene) balloons were first launched in 1958. They carry scientific instruments to a predetermined atmospheric density level. Zero-pressure balloons are the best for extremely high altitudes because the balloons can be lighter and stress on them can be distributed over the surface of the balloon.

About the same time, the Air Force Cambridge Research Laboratories (AFCRL) started working on super-pressure balloons, which were made from mylar. The development of mylar plastic films and advances in electronic miniaturization made constant-altitude balloons possible.

Mylar is a plastic that can withstand great internal pressure. The mylar super-pressure balloon does not expand as it rises, and it is sealed to prevent the release of gas as the balloon rises. By the time the balloon reaches the altitude where its density equals that of the atmosphere, the gas has become pressurized because the heat of the sun increases the internal gas pressure. However, because mylar can withstand great internal pressure, the volume of the balloon remains the same. By carefully calculating the weight of the balloon and whatever it is carrying, the altitude at which the balloon will achieve equilibrium and float can be calculated. As long as the pressure inside the balloon remains the same, it will remain at that altitude.

These balloons could be launched to remain aloft at specified altitudes for weeks or months at a time.

Moreover, satellites could be used to track and request data from many balloons in the atmosphere to obtain a simultaneous picture of atmospheric conditions all over the globe. Another advantage of super-pressure balloons is that, since they transmit their data to satellites, they can gather data from over oceans as well as land, which is a limitation of balloons equipped with radiosondes.

The AFCRL program resulted in the Global Horizontal Sounding Technique (GHOST) balloon system. With GHOST, meteorologists at last achieved their goal of semipermanent platforms floating high in the atmosphere.

Eighty-eight GHOST balloons were launched starting in March of 1966. The GHOST balloons and their French counterpart, EOLE, (the name Clement Ader used for one of his aircraft—named after the Greek god of the wind) used strong, plastic super-pressure balloons to trace air circulation patterns by drifting with the wind at constant density altitudes. Many super-pressure balloons were aloft at a time, grouped at constant density levels. Each balloon had a sensing device and transmitting system for gathering information on its position and weather data and transmitted atmospheric and weather data to weather satellites. They first transmitted their data to the NASA *Nimbus-4* meteorological satellite in 1970.

In 1966, a GHOST balloon circled the Earth in 10 days at 42,000 feet (12,801 meters). By 1973, NASA had orbited scientific instrument packages aboard sealed balloons at altitudes up to 78,000 feet (23,774 meters). Other GHOST balloons remained aloft for up to a year. The program lasted for 10 years.

The ultimate of the super-pressure balloons was the balloon satellite *Echo I.* Launched into space in 1960, the balloon inflated to a sealed volume by residual air, benzoic acid, and a chemical called anthraquinone.

Constant-altitude, super-pressure balloons continue to fly over the oceans and land surfaces. These balloons have been relied on for decades to provide extensive knowledge of global meteorology and improve worldwide weather forecasting.

READING LIST

Crouch, Tom D. *The Eagle Aloft: Two Centuries of the Balloon in America*. Washington, DC.: Smithsonian Institution Press, 1983.

Tannenbaum, Beulah and Harold E. *Making and Using Your Own Weather Station*. New York. Venture Books, 1989.

Vaeth, J. Gordon. "When the Race for Space Began." *U.S. Naval Institute Proceedings*. August 1961.

ONLINE REFERENCES

Cowens, John. "Building a Psychrometer," Greenwood Elementary, La Grande, OR 97850. September 28, 1993. http://explorer.scrtec.org/explorer/explorer-db/html/783750680-447DED81.html

"Hygrometer," Microsoft® Encarta® Online Encyclopedia 2000. 1997–2000 Microsoft Corporation. http://encarta.msn.com.

Lally, Vincent E. "Balloon: Modern Scientific Ballooning," Microsoft® Encarta® Online Encyclopedia 2000. 1997–2000 Microsoft Corporation. http://www.encarta.msn.com//find/concise.asp?ti=761576250&sid=3#s3. http://www.encarta.msn.com.

"Measuring the Weather," USAToday.com. http://www.usatoday.com/weather/wmeasur0.htm.

"Meteorology," Microsoft® Encarta® Online Encyclopedia 2000. 1997–2000 Microsoft Corporation. http://encarta.msn.com.

Warner, Lucy, ed. "Forecasts: Observing and Modeling the Global Atmosphere," *UCAR at 25*. University Center for Atmospheric Research. Boulder, Colorado. Oct. 17, 2000. http://www.ucar.edu/communications/ucar25/forecasts.html.

"Weather Instruments to Make." http://asd-www.larc.nasa.gov/SCOOL/psychrometer.html

To find out how you can make your own psychrometer, link to The CERES S'COOL Project at http://asd-www.larc.nasa.gov/SCOOL. The link has lots of information about making weather observations. Making a Psychrometer is in the Table of Contents or go to http://explorer.scrtec.org/explorer/explorer-db/html/783750680-447DED81.html.

To learn more about weather instruments and even set up your own weather station to report to the U.S. National Weather Service, go to http://www.usatoday.com/weather/wmeasur0.htm.

Educational Organization	Standard Designation (where applicable)	Content of Standard
National Science Education Standards	Content Standard D	Students should develop an understanding of energy in the earth system.
International Technology Education Association	Standard 9	Students will develop an understanding of engineering design.
International Technology Education Association	Standard 10	Students will develop an understanding of the role of trouble shooting, research and development, innovation, and experimentation in problem solving.
International Technology Education Association	Standard 7	Students will develop an understanding of the influence of technology on history.

BAROMETRIC EFFICIENCY

Todd Rasmussen
The University of Georgia
Athens, Georgia

Water levels in wells are often observed to fluctuate as the air pressure changes. Blaise Pascal described this effect in 1663 (1) and was the first to attribute the water level changes in wells to changes in atmospheric pressure. He noted that water levels declined as the barometric pressure increased, and vice versa.

The barometric efficiency, BE, is used to relate changes in water levels, ΔW, to changes in barometric (air) pressure, ΔB:

$$BE = -\frac{\Delta W}{\Delta B} \tag{1}$$

where consistent units (e.g., feet of water, mm of Hg, hPa) are used for both water levels and barometric pressure (2,3).

To understand why this relationship exists, we can conceive of an aquifer that is entirely isolated from the atmosphere. Such an aquifer maintains a constant total head, $H = H_o$, within the aquifer and is entirely unaffected by atmospheric pressure changes. Water level elevations, W, in the aquifer are measured in an open borehole.

The total head within the well is the sum of the water level elevation plus the barometric pressure exerted on the water surface, $H = W + B$. If the total head within the well is equal to the total head within the aquifer, $H_o = W + B$, and the total head within the aquifer is constant, then the water level varies inversely with barometric pressure, $W = H_o - B$.

The change in water level is just the negative of the change in barometric pressure, $\Delta W = \Delta(H_o - B) = -\Delta B$, so that the barometric efficiency is 100 percent, $BE = -\Delta W/\Delta B = -(-\Delta B)/\Delta B = 1$.

A second example assumes that the total head in the hypothetical aquifer increases with barometric pressure, $H = H_o + B$, so that the head in the aquifer goes up and down over time. In this case, the water levels in the well remain unchanged, and the barometric efficiency of the aquifer is zero.

These two extremes are shown in Fig. 1. Note that the barometric pressure is identical in both cases. In the left figure, the barometric efficiency is 100%, so that the water level varies inversely with barometric pressure. The total head is the sum of the barometric pressure and water levels, so it remains unchanged. In the right figure, the barometric efficiency is zero, so that the water level is unaffected by barometric pressure changes and the total head varies directly with barometric pressure.

Most aquifers lie between these two extremes, however, and the actual response depends, in large part, on whether the aquifer is confined or unconfined. Both of these cases are described below.

CONFINED AQUIFERS

For confined aquifers, Jacob (4) showed that increasing the load on the ground surface increases the load on the aquifer. This additional weight is either carried by the mineral grains or by the water within the aquifer pores. If the entire weight is borne by the mineral grains, and these grains do not deform with the increase in load, then the total head within the aquifer remains unchanged, and the barometric efficiency is 100%.

The barometric efficiency is smaller if some of the weight is carried by the fluid. A very low barometric efficiency in confined aquifers occurs when the fluid within the aquifer bears most of the weight. Examples of such aquifers include poorly consolidated sedimentary aquifers or horizontal fractures that extend great distances.

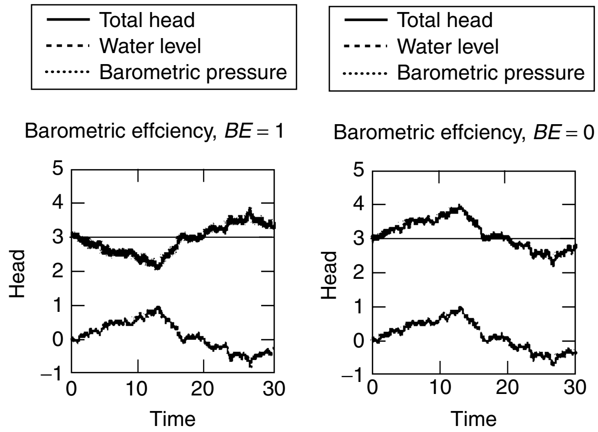

Figure 1. Effect of barometric pressure on water levels in wells. Both figures show barometric pressure (lower, dashed line), total head (solid line), and water levels (dotted line). Left figure shows effects when the barometric efficiency is 100%, $BE = 1$, and right figure shows effects when the barometric efficiency is zero, $BE = 0$.

Figure 2 shows the extreme conditions. In the left figure, the entire increase in load caused by an increase in barometric pressure is carried by the mineral grains. In this case, the pore fluids are not affected by the increase in load, the total head remains unchanged, and water levels drop in an amount equal to the increase in barometric pressure. In the right figure, the mineral grains do not carry the load and the fluid carries the increased load, causing an increase in total head.

In confined aquifers, the barometric response is virtually instantaneous. A change in barometric pressure should cause an immediate change in water levels in wells. Water levels in large-diameter wells may not respond immediately, however, because of the time required for water levels to adjust to the new level. Instead of a rapid

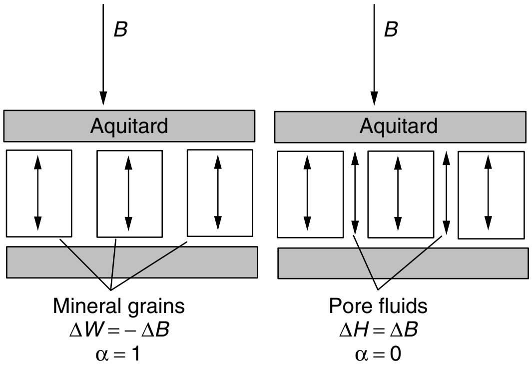

Figure 2. Effect of barometric pressure on total head in confined aquifers. The mineral grains carry the load in the left figure, whereas the pore fluids carry the load in the right figure. Most confined aquifers, however, are intermediate between these two extremes.

response to pressure changes, a slow response may be found in these wells.

This phenomena is called *borehole storage* and is a function of the diameter of the borehole, the length of the screened interval relative to the aquifer thickness, the aquifer hydraulic properties (e.g., transmissivity and storativity), and the rate at which the barometric pressure is changing.

The effects of borehole storage can be eliminated by placing a packer in the well below the water surface, but above the screened interval, which eliminates the need for the water level to change in response to changes in barometric pressure. Some change in volume may result from the slight compressibility of water, but this effect is very small. A gauge-type pressure sensor (i.e., internally vented to the atmosphere) is then placed below the packer to measure the water pressure.

Several techniques are available for estimating the confined aquifer barometric efficiency, including linear regression and Clark's Method (5). Clark's method is an unbiased technique for estimating the barometric efficiency that performs well when a water level trend is present in the data (6). Regression provides a better estimate, but only when the trend function can be accurately specified (7).

UNCONFINED AQUIFERS

Barometric pressure changes commonly do not affect water levels in wells located in shallow, unconfined aquifers, which is because the air pressure moves rapidly through the unsaturated zone and causes an immediate increase in total head within the aquifer.

In deeper unconfined aquifers, however, the typical response to barometric pressure changes is to see an immediate inverse response (i.e., an increase in barometric pressure causes an equivalent and opposite water level response), followed by a slow decay back to the original water level (5,8,9).

To understand this complex response, we focus on how total head responds to barometric pressure changes. Initially, no increase in total head within the unconfined aquifer occurs because the water held within pores is not confined by an overlying confining unit. The total head within the aquifer can rise over time, however, as air diffuses downward through the unsaturated zone.

The total head in the unconfined aquifer increases once the air pressure change reaches the water table. The time required for the total head to respond to barometric pressure changes is a function of the depth of the water table and the air diffusivity within the unsaturated zone. The air diffusivity is higher in coarse, dry soils and is lower in wet soils or in soils with a low air permeability.

As the water level response in observation wells is the total head minus the barometric pressure, the fact that the total head does not change initially means that a rapid water-level response occurs to barometric pressure changes. In fact, the barometric efficiency is one in wells with a deep unsaturated zone—or where the air permeability of the unsaturated zone is low—followed by a gradual decrease as air diffuses downward to the water table.

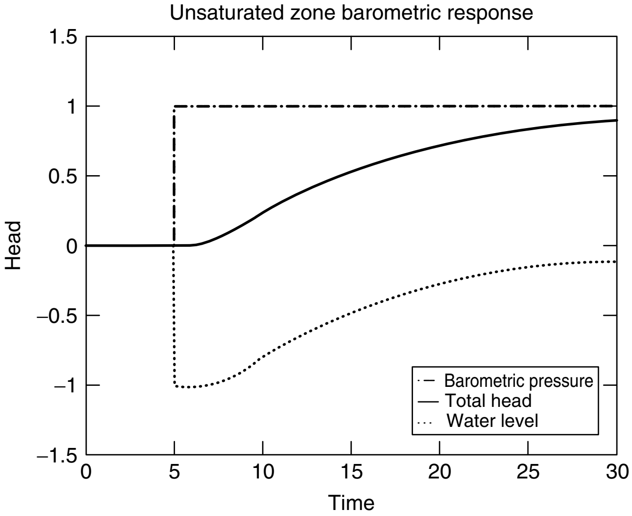

Figure 3. Effect of barometric pressure on total head and water levels in confined aquifers. The barometric pressure in this example takes a step from zero to one at $t = 10$. The total head slowly rises as the barometric pressure makes its way through the unsaturated zone. The water levels initially fall to maintain the constant total head, but then return to normal once the barometric pressure signal reaches the water table.

This concept is illustrated graphically in Fig. 3. Note that the barometric pressure is simplified to a simple step change from zero to one. The total head does not initially respond, but slowly rises over time as the barometric pressure moves through the unsaturated zone and eventually reaches the water table. The observed water level in a well placed in this aquifer responds immediately (as long as borehole storage can be neglected), but eventually decays back to its original level

The barometric efficiency of unconfined aquifers is not single valued like the confined aquifer barometric efficiency. Instead, the response includes a delay that can be estimated using regression deconvolution (5,9). The resulting response can be used to estimate the unsaturated air diffusivity of the unsaturated zone.

BIBLIOGRAPHY

1. Pascal, B. (1973). *The Physical Treatises of Pascal.* Octagon Books, New York.
2. Bear, J. (1972). *Dynamics of Fluids in Porous Media.* Dover Publications, New York.
3. Walton, W.C. (1970). *Groundwater Resource Evaluation.* McGraw-Hill, New York.
4. Jacob, C.E. (1940). On the flow of water in an elastic artesian aquifer. *Amer. Geophys. Union Trans.* **21**: 574–586.
5. Rasmussen, T.C. and Crawford, L.A. (1997). Identifying and removing barometric pressure effects in confined and unconfined aquifers. *Ground Water.* **35**(3): 502–511.
6. Clark, W.E. (1967). Computing the barometric efficiency of a well. *J. Hydraul. Divi. Amer. Soc. Civil Engin.* **93**(HY4): 93–98.
7. Davis, D.R. and Rasmussen, T.C. (1993). A comparison of linear regression with Clark's method for estimating barometric

efficiency of confined aquifers. *Water Resources Res.* **29**(6): 1849–1854.

8. Weeks, E.P. (1979). Barometric fluctuations in wells tapping deep unconfined aquifers. *Water Resources Res.* **15**(5): 1167–1176.
9. Spane, F.A., Jr. (2002). Considering barometric pressure in groundwater flow investigations. *Water Resources Res.* **38**(6): 1078.

READING LIST

Spane, F.A., Jr. and Mercer, R.B. (1985). *HEADCO: A Program for Converting Observed Water Levels and Pressure Measurements to Formation Pressure and Standard Hydraulic Head*. Pacific Northwest National Laboratory, Rockwell Hanford Operations, Richland, WA. Report PNL-10835.

CERES: UNDERSTANDING THE EARTH'S CLOUDS AND CLIMATE

NASA—Langley
Research Center

The Clouds and the Earth's Radiant Energy System (CERES) instrument is one of several launched aboard the Earth Observing System's (EOS) Aqua spacecraft in 2002. Scientists use observations from the CERES instrument to study the energy exchanged between the Sun; the Earth's atmosphere, surface and clouds; and outer space.

The CERES Aqua instruments are the fourth and fifth CERES instruments in orbit. NASA launched the first CERES instrument aboard the Tropical Rainfall Measuring Mission satellite or TRMM in November 1997. Results of the TRMM mission show that the first CERES provided better measurement capabilities than any previous satellite instrument of its kind. Two other CERES instruments are currently orbiting the Earth on the EOS Terra spacecraft, launched in late 1999. Early CERES Terra results give new insights into the effects of clouds on climate and how the climate system changes from decade to decade. Two CERES instruments on each of the Terra and Aqua spacecraft will provide global coverage of energy radiated and reflected from the Earth. Scientists use measurements from both satellites' orbits to improve observations of the daily cycle of radiated energy.

NASA Langley Research Center manages the CERES mission. Langley's highly successful Earth Radiation Budget Experiment (ERBE) provided the foundation for the design of the CERES instrument. ERBE used three satellites to provide global energy measurements from 1984 through the 1990s. The TRW Space & Electronics Group in Redondo Beach, Calif., built all six CERES instruments.

WHAT CERES WILL MEASURE

CERES measures the energy at the top of the atmosphere, as well as estimate energy levels in the atmosphere and at the Earth's surface. Using information from very high resolution cloud imaging instruments on the same spacecraft, CERES will determines cloud properties, including altitude, thickness, and the size of the cloud particles. All of these measurements are critical for advancing the understanding of the Earth's total climate system and the accuracy of climate prediction models.

BALANCING THE EARTH'S ENERGY BUDGET

The balance between Earth's incoming and outgoing energy controls daily weather and climate (long-term weather patterns). Sunlight or solar energy is the planet's only incoming energy source. Heat emitted from the sunlight reflected by the Earth's surface, atmosphere and clouds make up the planet's outgoing energy. Scientists have been working for decades to understand this critical energy balance, called the Earth's "energy budget."

The energy received from the Sun is at short wavelengths, while the energy emitted by the surface of the Earth, the atmosphere and clouds is at long wavelengths. Greenhouse gases in the atmosphere absorb the long wavelength energy or heat emitted by the Earth. Increases in the amounts of greenhouse gases produced by both natural processes or human activities can lead to a warming of the Earth's surface. Such changes may, in turn, alter the planet's daily weather and climate.

Clouds and small particles in the atmosphere called aerosols also reflect some sunlight back into space. Major sources of aerosols include windblown dust, emissions from the burning of fossil fuels, such as gasoline, and the burning of forests and agricultural fields.

CLOUD EFFECTS

One of the most intriguing questions facing climate modelers today is how clouds affect the Earth's climate

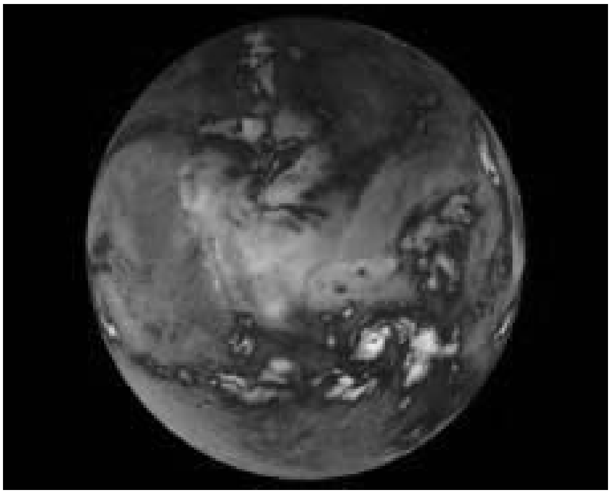

CERES detects low (blue and white) to high (yellow) amounts of emitted heat

---

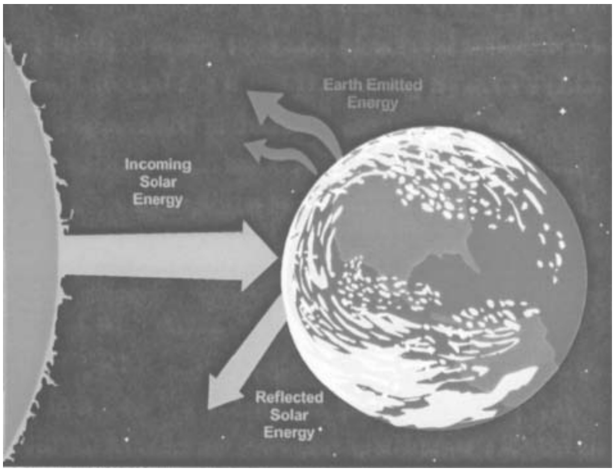

Earth's radiation budget is the balance between incoming and outgoing energy

and vice versa. The U.S. Global Change Research Program classifies understanding the role of clouds and the Earth's energy budget as one of its highest scientific priorities. Understanding cloud effects requires a detailed knowledge of how clouds absorb and reflect sunlight, as well as how they absorb and re-emit outgoing heat emitted by the planet. For example, low, thick clouds primarily reflect incoming solar energy back to space causing cooling. Thin, high clouds, however, primarily trap outgoing heat and produce warming. To date, satellite studies have found that clouds have an overall cooling effect on the Earth.

Analyses of satellite data also indicate that clouds which form over water are very different from clouds which form over land. These differences affect the way clouds reflect sunlight back into space and how much heat emitted from the Earth the clouds absorb and re-emit. For example, over the equator in the eastern Pacific Ocean during El Niño events, there is a significant decrease in the amount of energy emitted by the Earth due to increased cloudiness. El Niño events occur when portions of the eastern Pacific Ocean become considerably warmer than normal, causing an increase in cloudiness over the region. These changes can affect weather patterns around the world.

WATER VAPOR EFFECTS

Water vapor in the atmosphere also impacts our daily weather and climate, though scientists are only beginning to understand how this complex mechanism works. Water vapor acts like a greenhouse gas and absorbs outgoing heat to warm the Earth. Because water vapor also condenses to make clouds, additional water vapor in the atmosphere also may increase the amount of clouds.

FUTURE MISSIONS

One additional CERES instrument is available to fill the gap between Aqua and the next generation of highly accurate Earth radiation budget measurements. These

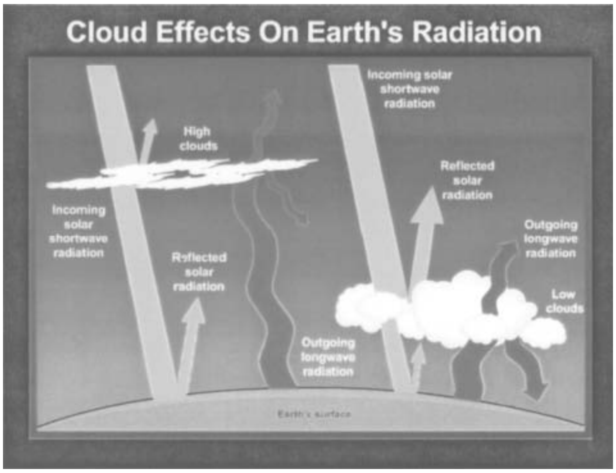

observations are expected to be made on the National Polar-orbiting Operational Environmental Satellite System (NPOESS) starting around 2010. To continue the 22-year record of global energy measurements, the next CERES mission should launch in 2007.

EDUCATIONAL OUTREACH

As a CERES instrument passes overhead, students worldwide are observing clouds and then sending their observations to NASA Langley's Atmospheric Sciences Data Center (ASDC). At the ASDC, scientists store data for further analysis by the CERES science team. The student observations are part of a global educational outreach program called the Students' Cloud Observations On-Line (S'COOL) project. Since the project began five years ago, S'COOL has reached over 1,000 schools in all 50 states and 57 other countries on five continents.

COMMERCIAL APPLICATIONS

CERES supports commercial applications by providing data about weather and sunlight at the Earth's surface for the renewable energy industry via an innovative Web site (http://eosweb.larc.nasa.gov/sse/). The Surface Meteorology and Solar Energy Project maintains the site. In the first three years of operation, the number of registered users of the Web site, including major energy companies, financial institutions and federal agencies, has grown to over 2,000 from nearly 100 countries. With 35,000 hits per month since January 2001, SSE is the most accessed Web site at the ASDC.

CHINOOK

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

Wind is defined as the movement of air. Although we commonly define wind as a gentle breeze or a harsh gust of cold air, there is a wind phenomenon that can increase the temperature instead of lowering it. It is called a Chinook wind. Chinooks are most commonly associated with the Rocky Mountain range in North American but can also be found in the Swiss Alps and the Andes. They can increase temperatures high enough to melt the snow in their path as they travel down the mountainside. The Chinook wind falls under the classification of katabatic wind, which means that it moves downhill.

The name “Chinook” was taken from the Chinook Indians that lived along the Rocky Mountains until the early 1800s when the tribe became extinct due to disease. Although the tribe died off, their legends and tales live on. One such legend is that of the Chinook wind, which in its literal translation means “snow-eater.”

It is possible for a Chinook to take place anytime during the year, but its effects are much more dramatic during the winter months. Chinook winds cause dramatic increases in the temperature on the eastern side of the Rockies and can send temperatures into the fifties and sixties. A temperature change of this magnitude can take anywhere from an hour to a day. The heat produced is a reaction formed from the change of gas to liquid in the atmosphere. These warm gusts of air then cause the evaporation of any snow on the ground, hence the name snow-eater. There is no definite length of time that a Chinook will last. On average, it can last hours or days.

Chinooks are the end result of the warm moist air of the Pacific Ocean moving up and down the Rocky Mountains. A westerly wind collects the warm moist air in the Pacific Ocean and carries it to the coast where it meets the western side of the Rockies. As the air makes its way up the mountain, the air becomes cooler, and the precipitation is squeezed out of it. Through this process warm, dry air is produced as it makes its way down the eastern side of the mountain. This process is repeated over and over again on each mountain that is in the way of the wind and each time produces warmer, dryer air. The air that makes its way down the last mountain is extremely dry and warm. This wind is called the Chinook wind. These warm gusts of air can reach speeds up to 100 mph.

The results from Chinook winds are both positive and negative. Chinook winds cause evaporation of the snow covering the ground, so the length of the grazing season is extended. A longer grazing season decreases the need to stock up on feed for animals. A negative result is the decreased amount of precipitation due to the quick evaporation of snow. Less precipitation causes hardships in planting.

READING LIST

Alberta Agriculture, Food, and Rural Development. (2002). *Weather in Alberta.* Available: http://www.agric.gov.ab.ca/agdex/000/7100001b.html. (March 11).

The Alberta Traveller. (2002). *Chinook Winds.* Available: http://www.traveller.babelfish.com/weather_guides/Chinook.html. (March 11).

Black Hills Weather. (2002). *The Snow Eating Wind.* Available: http://www.blackhillsweather.com/chinook.html. (March 11).

Ho, I. (2002). *Snow-Eater. The Catalyst.* Available: http://www.brown.edu/Students/Catalyst/archive/spring99/ch.html. (March 11).

Under the Chinook Arch. http://www.nucleus.com/~cowboy/Misc/Chinooks.html. (11 March 2002).

GLOBAL CLIMATE CHANGE

Geophysical Fluid Dynamics
Laboratory—NOAA

In 1967, two GFDL scientists, Syukuro Manabe and Richard Wetherald, published what is now regarded as the first credible calculation of the effect of increased carbon dioxide on the climate. They calculated that a doubling of atmospheric carbon dioxide would warm the earth’s surface by about 2 °C. This result laid the foundation for what has become an international, multi-disciplinary research effort on global warming.

Manabe, in collaboration with oceanographer Kirk Bryan and other scientists at GFDL, has continued to lead the international effort to develop the coupled ocean-atmosphere climate models that are crucial to understanding and predicting the impact of greenhouse gases.

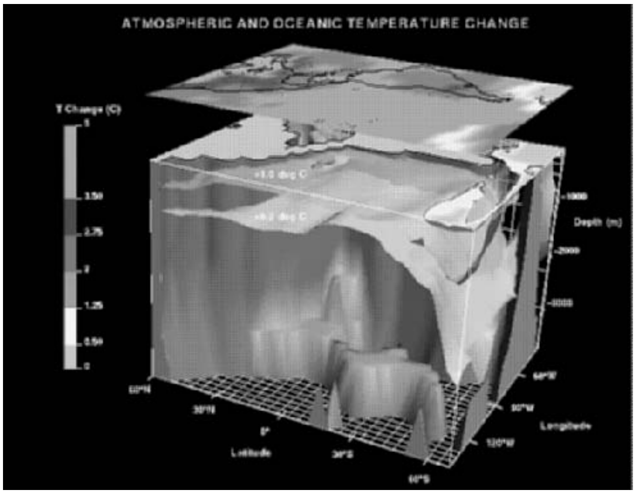

Three-dimensional view of projected surface air temperature and ocean warming (°C) due to greenhouse gases as calculated by a low-resolution GFDL coupled ocean-atmosphere climate model. The top panel shows the surface air temperature change over North and South America and surrounding regions. The three-dimensional box illustrates the depth to which a 1 °C and 0.2 °C warming has penetrated in the model’s Pacific Ocean. The gray surface depicts the model’s ocean floor. Note the deep mixing of the ocean warming signal in the southern hemisphere ocean near Antarctica. The temperature changes are projections of the warming due to greenhouse gases by the latter half of the twenty-first century in the absence of further increases in sulfate aerosol forcing. Results shown are based on years 61–80 of a transient CO_2 increase experiment (+1% per year compounded). [*Source*: adapted from Syukuro Manabe and Ronald Stouffer, Nature, 15 July 1993.]

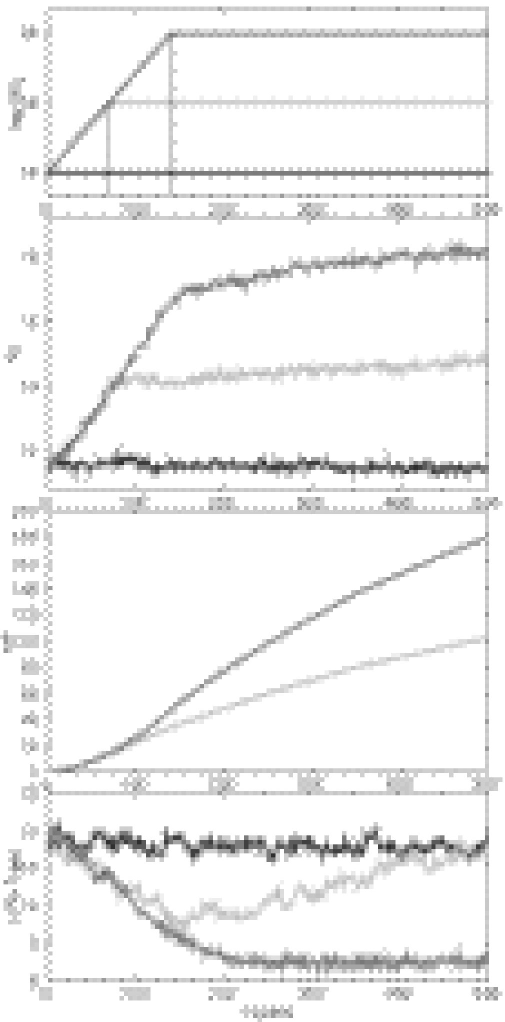

Impact of increasing CO_2 on the earth's climate as simulated in a GFDL coupled ocean-atmosphere climate model. Shown are timeseries of: a) prescribed CO_2 concentration on a logarithmic scale in comparison to present levels; b) global mean surface air temperature (°C); c) global mean increase of sea level (cm) due to thermal expansion; and d) intensity of the North Atlantic Ocean's meridional overturning circulation (10**6 m**3/sec). The labels "Control", "$2xCO_2$", and "$4xCO_2$" refer to separate experiments where CO_2 either remains constant (Control), or increases at 1% per year (compounded) to double ($2xCO_2$) or quadruple ($4xCO_2$) the current concentration. Note that the sea level rise estimates do not include the effect of melted continental ice sheets. With this effect included, the total rise could be larger by a substantial factor. [*Source*: Syukuro Manabe and Ronald Stouffer, Nature, 15 July 1993.]

A PROBLEM FOR CENTURIES TO COME?

In a recent paper, published 26 years after Manabe's pioneering one-dimensional CO_2 sensitivity study, he and Ron Stouffer of GFDL used a three-dimensional coupled ocean-atmosphere model to examine possible CO_2-induced climate changes over several centuries. Earlier studies had focused on shorter time horizons.

In their scenario, CO_2 quadruples over a period of 140 years, then no longer increases. This perturbation is enough to cause the ocean's global thermohaline circulation to almost disappear in the model. The global thermohaline circulation is important because it is responsible for a large portion of the heat transport from the tropics to higher latitudes in the present climate. In addition, sea level continues rising steadily for centuries after the CO_2 increase is halted. From this perspective, global warming can no longer be viewed as just as a problem of our own lifetimes, but as a legacy—with uncertain consequences—now being passed forward to many future generations.

EVALUATING CLIMATE MODELS

GFDL scientists, including Tony Broccoli and Tom Delworth, are searching for innovative ways to evaluate climate models and to distinguish between human-induced climate change and natural climate variability. Measurements of the current climate, historical observations, and glimpses of earth's climate during the ice ages and other past climates all provide opportunities to test climate models. Through research on climate models and observations, scientists at GFDL will continue to evaluate and refine the climate models that will be needed to help answer critical policy-relevant questions about global warming and its consequences.

OBSERVATIONS OF CLIMATE AND GLOBAL CHANGE FROM REAL-TIME MEASUREMENTS

David R. Easterling
Thomas R. Karl
(from *Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements*, Wiley 2003)

INTRODUCTION

Is the planet getting warmer?
Is the hydrologic cycle changing?
Is the atmospheric/oceanic circulation changing?
Is the weather and climate becoming more extreme or variable?
Is the radiative forcing of the climate changing?

These are the fundamental questions that must be answered to determine if climate change is occurring. However, providing answers is difficult due to an inadequate or nonexistent worldwide climate observing system. Each of these apparently simple questions are quite complex because of the multivariate aspects of each question and because the spatial and temporally sampling required to address adequately each question must be considered on a global scale. A brief review of our ability to answer these questions reveals many successes, but points to some glaring inadequacies that must be addressed in any attempt to understand, predict, or assess issues related to climate and global change.

IS THE PLANET GETTING WARMER?

There is no doubt that measurements show that near-surface air temperatures are increasing. Best estimates suggest that the warming is around 0.6 °C (+−0.2 °C) since the late nineteenth century (1). Furthermore, it appears that the decade of the 1990s was the warmest decade since the 1860s, and possibly for the last 1000 years. Although there remain questions regarding the adequacy of this estimate, confidence in the robustness of this warming trend is increasing (1). Some of the problems that must be accounted for include changes in the method of measuring land and marine surface air temperatures from ships, buoys, land surface stations as well as changes in instrumentation, instrument exposures and sampling times, and urbanization effects. However, recent work evaluating the effectiveness of corrections of sea surface temperatures for time-dependent biases, and further evaluation of urban warming effects on the global temperature record have increased confidence in these results. Furthermore, by consideration of other temperature-sensitive variables, e.g., snow cover, glaciers, sea level and even some proxy non-real-time measurements such as ground temperatures from boreholes, increases our confidence in the conclusion that the planet has indeed warmed. However, one problem that must be addressed is that the measurements we rely upon to calculate global changes of temperature have never been collected for that purpose, but rather to aid in navigation, agriculture, commerce, and in recent decades for weather forecasting. For this reason there remain uncertainties about important details of the past temperature increase and our capabilities for future monitoring of the climate. The IPCC (1) has summarized latest known changes in the temperature record, which are summarized in Fig. 1.

Global-scale measurements of layer averaged atmospheric temperatures and sea surface temperatures from instruments aboard satellites have greatly aided our ability to monitor global temperature change (2–4), but the situation is far from satisfactory (Hurrell and Trenberth, 1996). Changes in satellite temporal sampling (e.g., orbital drift), changes in atmospheric composition (e.g., volcanic emissions), and technical difficulties related to overcoming surface emissivity variability are some of the problems that must be accounted for, and reduce the confidence that can be placed on these measurements (5). Nonetheless, the space-based measurements have shown, with high confidence, that stratospheric temperatures have decreased over the past two decades. Although perhaps not as much as suggested by the measurements from weather balloons, since it is now known that the data from these balloons high in the atmosphere have an inadvertent temporal bias due to improvements in shielding from direct and reflected solar radiation (6).

IS THE HYDROLOGIC CYCLE CHANGING?

The source term for the hydrologic water balance, precipitation, has been measured for over two centuries in some locations, but even today it is acknowledged that in many parts of the world we still cannot reliably measure true precipitation (7). For example, annual biases of more than 50% due to rain gauge undercatch are not uncommon in cold climates (8), and, even for more moderate climates, precipitation is believed to be underestimated by 10 to 15% (9). Progressive improvements in instrumentation, such as the introduction of wind shields on rain gauges, have also introduced time-varying biases (8). Satellite-derived measurements of precipitation have provided the only large-scale ocean coverage of precipitation. Although they are comprehensive estimates of large-scale spatial precipitation variability over the oceans where few measurements exist, problems inherent in developing precipitation estimates hinder our ability to have much confidence in global-scale decadal changes. For example, even the landmark work of Spencer (10) in estimating worldwide ocean precipitation using the microwave sounding unit aboard the National Oceanic and Atmospheric Administration (NOAA) polar orbiting satellites has several limitations. The observations are limited to ocean coverage and hindered by the requirement of an unfrozen ocean. They do not adequately measure solid precipitation, have low spatial resolution, and are affected by the diurnal sampling inadequacies associated with polar orbiters, e.g., limited overflight capability. Blended satellite/*in situ* estimates also show promise (11); however, there are still limitations, including a lack of long-term measurements necessary for climate change studies.

Information about past changes in land surface precipitation, similar to temperature, has been compared with other hydrologic data, such as changes in stream flow, to ascertain the robustness of the documented changes of precipitation. Figure 1 summarizes some of the more important changes of precipitation, such as the increase in the mid to high latitude precipitation and the decrease in subtropical precipitation. Evidence also suggests that much of the increase of precipitation in mid to high latitudes arises from increased autumn and early winter precipitation in much of North America and Europe. Figure 2 depicts the spatial aspects of this change, reflecting rather large-scale coherent patterns of change during the twentieth century.

Other changes related to the hydrologic cycle are summarized in Fig. 1. The confidence is low for many of the changes, and it is particularly disconcerting relative to the role of clouds and water vapor in climate feedback effects.* Observations of cloud amount long have been made by surface-based human observations and more recently by satellite. In the United States, however, human observers have been replaced by automated measurements, and neither surface-based or spaced-based data sets have proven to be entirely satisfactory for detecting changes in clouds. Polar orbiting satellites have an enormous difficulty to overcome related to sampling aliasing and satellite drift (12). For human observers changes in observer schedules, observing biases, and incomplete sampling have created major problems in data

---

*An enhancement or diminution of global temperature increases or decreases due to other causes.

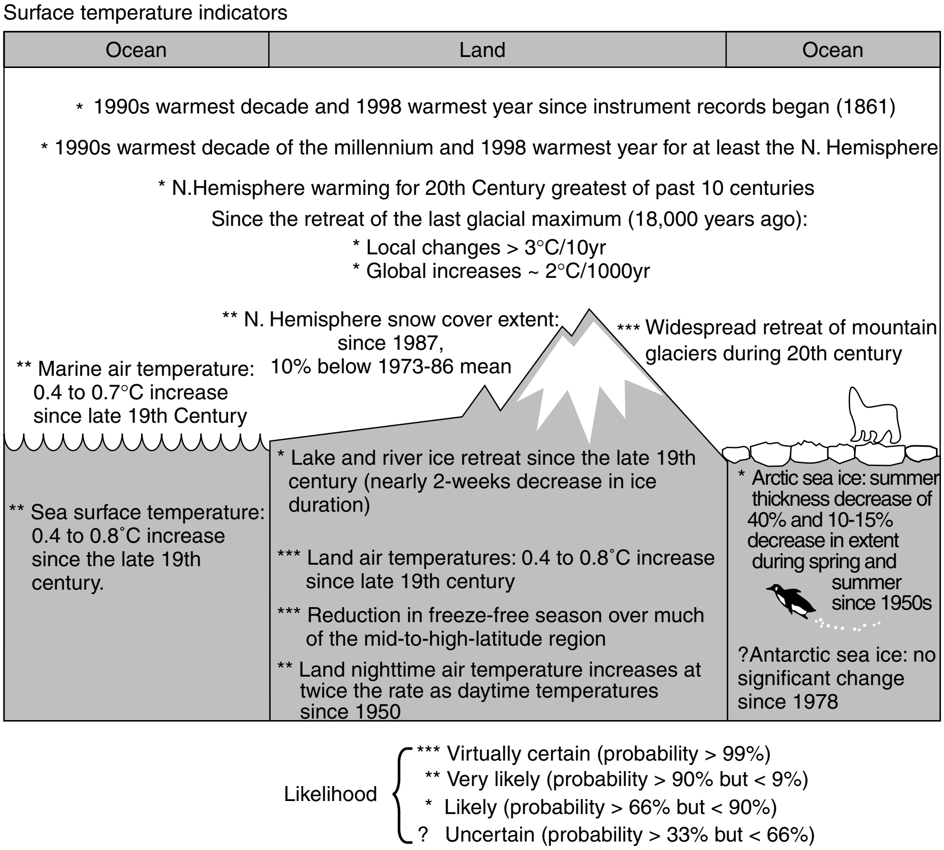

Figure 1. Schematic of observed variations of selected indictors regarding (**a**) temperature and (**b**) the hydrologic cycle (based on Ref. 1).

interpretations, now compounded by a change to new automated measurements at many stations. Nonetheless, there is still some confidence (but low) that global cloud amounts have tended to increase. On a regional basis this is supported by reductions in evaporation as measured by pan evaporimeters over the past several decades in Russia and the United States, and a worldwide reduction in the land surface diurnal temperature range. Moreover, an increase in water vapor has been documented over much of North America and in the tropics (1).

Changes in water vapor are very important for understanding climate change since water vapor is the most important greenhouse gas in the atmosphere. The measurement of changes in atmospheric water vapor is hampered by both data processing and instrumental difficulties for both weather balloon and satellite retrievals. The latter also suffers from discontinuities among successive satellites and errors introduced by changes in orbits and calibrations. Upper tropospheric water vapor is particularly important for climate feedbacks, but, as yet, little can be said about how it has varied over the course of the past few decades.

IS THE ATMOSPHERIC/OCEANIC CIRCULATION CHANGING?

Surprisingly, there is a considerable lack of reliable information about changes in atmospheric circulation, even though it is of daily concern to much of the world since it relates to day-to-day weather changes. Analyses of circulation are performed every day on a routine basis, but the analysis schemes have changed over time, making them of limited use for monitoring climate change. Moreover, even the recent reanalysis efforts by the world's major numerical weather prediction centers, whereby the analysis scheme is fixed over the historical record, contains time-varying biases because of the introduction of data with time-dependent biases and a changing mix of data (e.g., introducing satellite data) over the course of the reanalysis (13). Even less information is available on measured changes and variations in ocean circulation.

A few major atmospheric circulation features have been reasonably well measured because they can be represented by rather simple indices. This includes the El Niño–Southern Oscillation (ENSO) index, the North Atlantic Oscillation (NAO) index, and the Pacific–North American (PNA) circulation pattern index. There are interesting decadal and multidecadal variation, but it is too early to detect any long-term trends. Evidence exists that ENSO has varied in period, recurrence interval, and strength of impact. A rather abrupt change in ENSO and other aspects of atmospheric circulation seems to have occurred around 1976–1977. More frequent ENSOs with rare excursions into its other extreme (La Niña) became much more prevalent. Anomalous circulation regimes associated with ENSO and large-amplitude PNA patterns persisted in the North Pacific

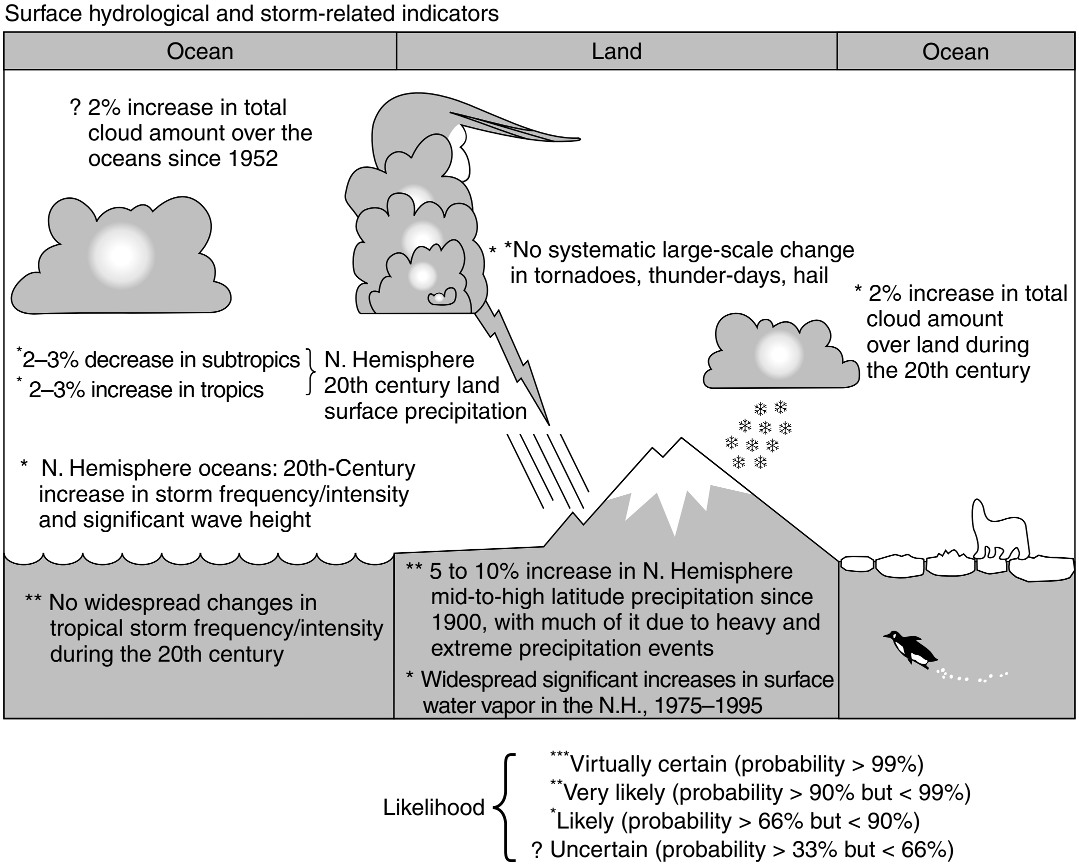

Figure 2. Precipitation trends over land 1900–1999. Trend is expressed in percent per century (relative to the mean precipitation from 1961–1990) and magnitude of trend is represented by area of circle with green reflecting increases and brown decreases of precipitation.

from the late 1970s into the late 1980s, affecting temperature anomalies. Moreover, the NAO has been persistent in its association with strong westerlies into the European continent from the late 1980s until very recently when it abruptly shifted. As a result, temperature anomalies and storminess in Europe have abruptly changed over the past 2 years compared to the past 7 or 8 years.

Increases in the strength of the Southern Hemisphere circumpolar vortex during the 1980s have been documented (14,15) using station sea level pressure data. This increase was associated with a delayed breakdown in the stratospheric polar vortex and ozone deficit in the Antarctic spring. A near-global sea level pressure data set has been used to identify changes in circulation patterns in the Indian Ocean. Allan et al. (16) and Salinger et al. (17) find that circulation patterns in the periods 1870–1900 and 1950–1990 were more meridional than those in the 1900–1950 period, indicating intensified circulation around anticyclones. These changes may be related to changes in the amplitude of longwave troughs to the south and west of Australia and the Tasman Sea/New Zealand area and a subsequent decrease in precipitation in Southwest Australia (18,19).

IS THE WEATHER AND CLIMATE BECOMING MORE EXTREME OR VARIABLE?

Climate and weather extremes are of great interest. Due to inadequate monitoring as well as prohibitively expensive access to weather and climate data held by the world's national weather and environmental agencies, only limited reliable information is available about large-scale changes in extreme weather or climate variability. The time-dependent biases that affect climate means are even more difficult to effectively eliminate from the extremes of the distributions of various weather and climate elements. There are a few areas, however, where regional and global changes in weather and climate extremes have been reasonably well documented (20).

Interannual temperature variability has not changed significantly over the past century. On shorter time scales and higher frequencies, e.g., days to a week, temperature variability may have decreased across much of the Northern Hemisphere (8). Related to the decrease in high-frequency temperature variability there has been a tendency for fewer low-temperature extremes, but widespread changes in extreme high temperatures have not been noted.

Trends in intense rainfall have been examined for a variety of countries. Some evidence suggests an increase in intense rainfalls (United States, tropical Australia, Japan, and Mexico), but analyses are far from complete and subject to many discontinuities in the record. The strongest increases in extreme precipitation are documented in the United States and Australia (21).

Intense tropical cyclone activity may have decreased in the North Atlantic, the one basin with reasonably consistent tropical cyclone data over the twentieth century, but even here data prior to World War II is difficult to assess regarding tropical cyclone strength. Elsewhere, tropical cyclone data do not reveal any long-term trends, or if they do they are most likely a result of inconsistent analyses. Changes in meteorological assimilation schemes have complicated the interpretations of changes in extratropical cyclone frequency. In some regions, such as the North Atlantic, a clear trend in activity has been noted, as also in significant wave heights in the northern half of the North Atlantic. In contrast, decreases in storm frequency and wave heights have been noted in the south half of the North Atlantic over the past few decades. These changes are also reflected in the prolonged positive excursions of the NAO since the 1970s.

IS THE RADIATIVE FORCING OF THE PLANET CHANGING?

Understanding requires a time history of forcing global change. The atmospheric concentration of CO_2, an important greenhouse gas because of its long atmospheric residence time and relatively high atmospheric concentration, has increased substantively over the past few decades. This is quite certain as revealed by precise measurements made at the South Pole and at Mauna Loa Observatory since the late 1950s, and from a number of stations distributed globally that began operating in subsequent decades. Since atmospheric carbon dioxide is a long-lived atmospheric constituent and it is well mixed in the atmosphere, a moderate number of well-placed stations operating for the primary purpose of monitoring seasonal to decadal changes provides a very robust estimate of global changes in carbon dioxide.

To understand the causes of the increase of atmospheric carbon dioxide, the carbon cycle and the anthropogenic carbon budget must be balanced. Balancing the carbon budget requires estimates of the sources of carbon from anthropogenic emissions from fossil fuel and cement production, as well as the net emission from changes in land use (e.g., deforestation). These estimates are derived from a combination of modeling, sample measurements, and high-resolution satellite imagery. It also requires measurements for the storage in the atmosphere, the ocean uptake, and uptake by forest regrowth, the CO_2 and nitrogen fertilization effect on vegetation, as well as any operating climate feedback effects (e.g., the increase in vegetation due to increased temperatures). Many of these factors are still uncertain because of a paucity of ecosystem measurements over a sustained period of time. Anthropogenic emissions from the burning of fossil fuel and cement production are the primary cause of the atmospheric increase.

Several other radiatively important anthropogenic atmospheric trace constituents have been measured for the past few decades. These measurements have confirmed significant increases in atmospheric concentrations of CH_4, N_2O, and the halocarbons including the stratospheric ozone destructive agent of the chloroflourocarbons and the bromocarbons. Because of their long lifetimes, a few well-placed high-quality *in situ* stations have provided a good estimate of global change. Stratospheric ozone depletion has been monitored both by satellite and ozonesondes. Both observing systems have been crucial in ascertaining changes of stratospheric ozone that was originally of interest, not because of its role as a radiative forcing agent, but its ability to absorb ultraviolet (UV) radiation prior to reaching Earth's surface. The combination of the surface- and space-based observing systems has enabled much more precise measurements than either system could provide by itself. Over the past few years much of the ozonesonde data and satellite data has been improved using information about past calibration methods, in part because of differences in trends between the two observing systems.

Figure 3 depicts the IPCC (9) best estimate of the radiative forcing associated with various atmospheric constituents. Unfortunately, measurements of most of the forcings other than those already discussed have low or very low confidence, not only because of our uncertainty about their role in the physical climate system, but because we have not adequately monitored their change. For example, estimates of changes in sulfate aerosol concentrations are derived from model estimates of source emissions, not actual atmospheric concentrations. The problem is complicated because of the spatially varying concentrations of sulfate due to its short atmospheric lifetime. Another example is measurements of solar irradiance, which have been taken by balloons and rockets for several decades, but continuous measurements of top-of-the-atmosphere solar irradiance did not begin until the late 1970s with the *Nimbus*

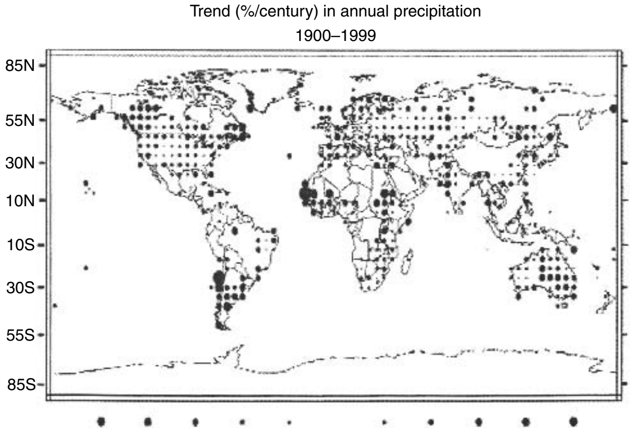

Figure 3. Estimates of globally and annually averaged radiative forcing (in W/m_{-2}) for a number of agents due to changes in concentrations of greenhouse gases and aerosols and natural changes in solar output from 1750 to the present day. Error bars are depicted for all forcings (from Ref. 1).

7 and the Solar Maximum Mission satellites. There are significant absolute differences in total irradiance between satellites, emphasizing the critical need for overlaps between satellites and absolute calibration of the irradiance measurements to determine decadal changes. Spectrally resolved measurements will be a key element in our ability to model the effects of solar variability, but at the present time no long-term commitment has been made to take these measurements. Another important forcing that is estimated through measured, modeled, and estimated changes in optical depth relates to the aerosols injected high into the atmosphere by major volcanic eruptions. The aerosols from these volcanoes are sporadic and usually persist in the atmosphere for at most a few years. Improved measurements of aerosol size distribution and composition will help better understand this agent of climate change.

WHAT CAN WE DO TO IMPROVE OUR ABILITY TO DETECT CLIMATE AND GLOBAL CHANGE?

Even after extensive reworking of past data, in many instances we are incapable of resolving important aspects concerning climate and global change. Virtually every monitoring system and data set requires better data quality, continuity, and fewer time-varying biases if we expect to conclusively answer questions about how the planet has changed, because of the need to rely on observations that were never intended to be used to monitor the physical characteristics of the planet of the course of decades. Long-term monitoring, capable of resolving decade-to-century-scale changes, requires different strategies of operation.

In situ measurements are currently in a state of decay, decline, or rapid poorly documented change due to the introduction of automated measurements without adequate precaution to understand the difference between old and new observing systems. Satellite-based systems alone will not and cannot provide all the necessary measurements. Much wiser implementation and monitoring practices must be adopted for both space-based and surface-based observing systems in order to adequately understand global change. The establishment of the Global Climate Observing System (GCOS) is a high priority (22), and continued encouragement by the World Meteorological Organization (WMO) of a full implementation of this system in all countries is critical. Furthermore, in the context of the GCOS, a number of steps can be taken to improve our ability to monitor climate and global change.

These include:

1. Prior to implementing changes to existing environmental monitoring systems or introducing new observing systems, standard practice should include an assessment of the impact of these changes on our ability to monitor environmental variations and changes.
2. Overlapping measurements in time and space of both the old and new observing systems should be standard practice for critical environmental variables.
3. Calibration, validation, and knowledge of instrument, station, and/or platform history are essential for data interpretation and use. Changes in instrument sampling time, local environmental conditions, and any other factors pertinent to the interpretation of the observations and measurements should be recorded as a mandatory part of the observing routine and be archived with the original data. The algorithms used to process observations must be well documented and available to the scientific community. Documentation of changes and improvements in the algorithms should be carried along with the data throughout the data archiving process.
4. The capability must be established to routinely assess the quality and homogeneity of the historical

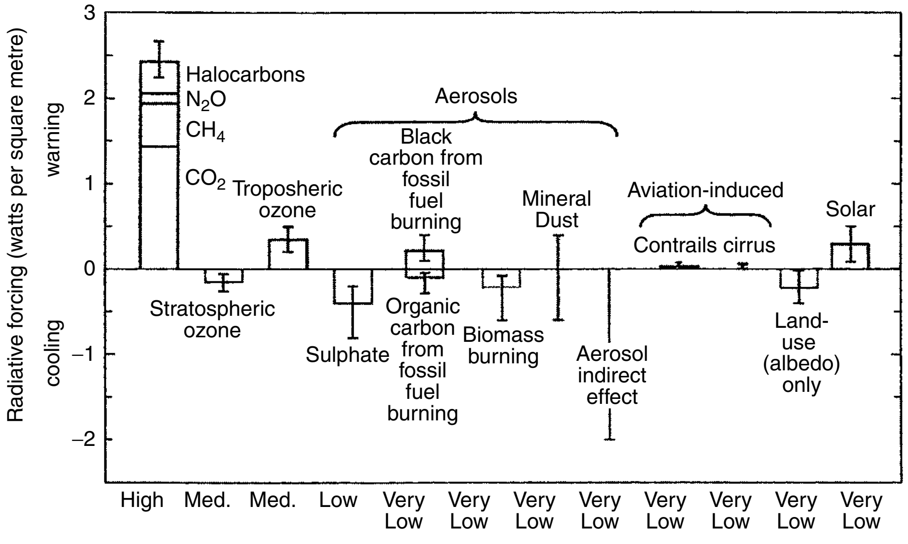

Figure 4. Global, annual-mean radiative forcings (Wm^{-2}) due to a number of agents for the period from pre-industrial (1750) to the present. The height of the vertical bar denotes the central or "best" estimate, no bar indicates that it is not possible to provide a "best" estimate. The vertical line indicates an estimate of the uncertainty range and the level of scientific understanding is a subjective judgement about the reliability of the forcing estimate based on such factors as assumptions, degree of knowledge of the physical/chemical mechanisms, etc. (From Ref. 1).

database for monitoring environmental variations and change, including long-term high-resolution data capable of resolving important extreme environmental events.

5. Environmental assessments that require knowledge of environmental variations and change should be well integrated into a global observing system strategy.
6. Observations with a long uninterrupted record should be maintained. Every effort should be made to protect the data sets that document long-term homogeneous observations. Long term may be a century or more. A list of prioritized sites or observations based on their contribution to long-term environmental monitoring should be developed for each element.
7. Data-poor regions, variables, regions sensitive to change, and key measurements with inadequate temporal resolution should be given the highest priority in the design and implementation of new environmental observing systems.
8. Network designers, operators, and instrument engineers must be provided environmental monitoring requirements at the outset of network design. This is particularly important because most observing systems have been designed for purposes other than long-term monitoring. Instruments must have adequate accuracy with biases small enough to resolve environmental variations and changes of primary interest.
9. Much of the development of new observation capabilities and much of the evidence supporting the value of these observations stem from research-oriented needs or programs. Stable, long-term commitments to these observations, and a clear transition plan from research to operations, are two requirements in the development of adequate environmental monitoring capabilities.
10. Data management systems that facilitate access, use, and interpretation are essential. Freedom of access, low cost, mechanisms that facilitate use (directories, catalogs, browse capabilities, availability of metadata on station histories, algorithm accessibility and documentation, on-line accessibility to data, etc.), and quality control should guide data management. International cooperation is critical for successful management of data used to monitor long-term environmental change and variability.

BIBLIOGRAPHY

1. IPCC (2001). *Climate Change, 2001: The Scientific Basis. Contribution of Working Group 1 to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. J.T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (Eds.). Cambridge University Press, New York.
2. Spencer, R.W. and Christy, J.R. (1992). Precision and radiosonde validation of satellite gridpoint temperature anomalies, Part I: MSU channel 2. *J. Climate* **5**: 847–857.
3. Spencer, R.W. and Christy, J.R. (1992). Precision and radiosonde validation of satellite gridpoint temperature anomalies, Part II: A tropospheric retrieval and trends during 1979–90. *J. Climate* **5**: 858–866.
4. Reynolds, R.W. (1988). A real-time global sea surface temperature analysis. *J. Climate* **1**: 75–86.
5. National Research Council (NRC) (2000). *Reconciling Observations of Global Temperature Change, Report of the Panel on Reconciling Temperature Observations*. National Academy Press, Washington, DC.
6. Luers, J.K. and Eskridge, R.E. (1995). Temperature corrections for the VIZ and Vaisala radiosondes. *Appl. Meteor.* **34**: 1241–1253.
7. Sevruk, B. (1982). Methods of correcting for systematic error in point precipitation measurements for operational use. *Hydrology Rep.* **21**. World Meteorological Organization, Geneva, 589.
8. Karl, T.R., Knight, R.W., and Plummer, N. (1995). Trends in high-frequency climate variability in the twentieth century. *Nature* **377**: 217–220.
9. IPCC (1992). *Climate Change, 1992, Supplementary Report*, WMO/UNEP. J.T. Houghton, B.A. Callander, and S.K. Varney (Eds.). Cambridge University Press, New York, pp. 62–64.
10. Spencer, R.W. (1993). Global oceanic precipitation from the MSU during 1979–1992 and comparison to other climatologies. *J. Climate* **6**: 1301–1326.
11. Huffman, G.J., Adler, R.F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., and Schneider, U. (1997). The global precipitation climatology project (GPCP) combined precipitation dataset. *Bull. Am. Meteor. Soc.* **78**: 5–20.
12. Rossow, W.B. and Cairns, B. (1995). Monitoring changes in clouds. *Climatic Change* **31**: 175–217.
13. Trenberth, K.E. and Guillemot, C.J. (1997). Evaluation of the atmospheric moisture and hydrologic cycle in the NCEP reanalysis. *Clim. Dyn.* **14**: 213–231.
14. van Loon, H., Kidson, J.W., and Mullan, A.B. (1993). Decadal variation of the annual cycle in the Australian dataset. *J. Climate* **6**: 1227–1231.
15. Hurrell, J.W. and van Loon, H. (1994). A modulation of the atmospheric annual cycle in the Southern Hemisphere. *Tellus* **46A**: 325–338.
16. Allan, R.J., Lindesay, J.A., and Reason, C.J.C. (1995). Multidecadal variability in the climate system over the Indian Ocean region during the austral summer. *J. Climate* **8**: 1853–1873.
17. Salinger, M.J., Allan, R., Bindoff, N., Hannah, J., Lavery, B., Leleu, L., Lin, Z., Lindesay, J., MacVeigh, J.P., Nicholls, N., Plummer, N., and Torok, S. (1995). Observed variability and change in climate and sea level in Oceania. In: *Greenhouse: Coping with Climate Change*. W.J. Bouma, G.I. Pearman, and M.R. Manning (Eds.). CSIRO, Melbourne, Australia, 100–126.
18. Nicholls, N. and Lavery, B. (1992). Australian rainfall trends during the twentieth century. *Int. J. Climatology* **12**: 153–163.
19. Allan, R.J. and Haylock, M.R. (1993). Circulation features associated with the winter rainfall decrease in southwestern Australia. *J. Climate* **6**: 1356–1367.
20. Easterling, D.R., Meehl, G., Parmesan, C., Changnon, S., Karl, T., and Mearns, L. (2000). Climate extremes: observations, modeling, and impacts. *Science* **289**: 2068–2074.

21. Easterling, D.R., Evans, J.L., Groisman, P.Ya., Karl, T.R., Kunkel, K.E., and Ambenje, P. (2000). Observed variability and trends in extreme climate events: A brief review. *Bull. Am. Meteor. Soc.* **81**: 417–426.
22. Spence, T. and Townsend, J. (1995). The Global Climate Observing System (GCOS). *Climatic Change* **31**: 131–134.

READING LIST

Elliott, W.P. (1995). On detecting long-term changes in atmospheric moisture. *Climatic Change* **31**: 219–237.

Groisman, P.Y., and Legates, D.R. (1995). Documenting and detecting long-term precipitation trends: Where we are and what should be done. *Climatic Change* **31**: 471–492.

Hurrell, J.W. and Trenberth, K.E. (1996). Satellite versus surface estimates of air temperature since 1979. *J. Climate* **9**: 2222–2232.

OVERVIEW: THE CLIMATE SYSTEM

Robert E. Dickinson

(from *The Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements*, Wiley 2003)

The climate system consists of the atmosphere, cryosphere, oceans, and land interacting through physical, chemical, and biological processes. Key ingredients are the hydrological and energy exchanges between subsystems through radiative, convective, and fluid dynamical mechanisms. Climate involves changes on seasonal, year-to-year, and decadal or longer periods in contrast to day-to-day weather changes. However, extreme events and other statistical measures are as, or more, important than simple averages. Climate is seen to impact human activities most directly through the occurrence of extremes. The frequency of particular threshold extremes, as, for example, the number of days with maximum temperatures above 100 °F, can change substantially with shifts in climate averages.

THE ATMOSPHERE

The atmosphere is described by winds, pressures, temperatures, and the distribution of various substances in gaseous, liquid, and solid forms. Water is the most important of these substances. Also important are the various other radiatively active (“greenhouse”) gases, including carbon dioxide and liquid or solid aerosol particulates. Most of the mass of the atmosphere is in the troposphere, which is comprised of the layers from the surface to about 12 km (8 km in high latitudes to 16 km at the equator) where the temperature decreases with altitude. The top of the troposphere is called the tropopause. Overlying this is the stratosphere, where temperatures increase with altitude to about 50 km or so (Fig. 1). The tropospheric temperature decreases with altitude are maintained by vertical mixing driven by moist and dry convection. The temperature increases with altitude in the stratosphere in response to increasing heating per the unit mass by ozone absorption of ultraviolet radiation. The variation of temperature structure with latitude is indicated in Fig. 2. The troposphere is deepest in the tropics because most thunderstorms occur there. Because of this depth and stirring by thunderstorms, the coldest part of the atmosphere is the tropical tropopause. In the lower troposphere temperatures generally decrease from the equator to pole, but warmest temperatures shift toward the summer hemisphere, especially in July. Longitudinally averaged winds are shown in Fig. 3. Because of the geostrophic balance between wind and pressures, winds increase with altitude where temperature decreases with latitude. Conversely, above about 8 km, where temperatures decrease toward the tropical tropopause, the zonal winds decrease with altitude. The core of maximum winds is referred to as the jet stream. The jet stream undergoes large wavelike oscillations in longitude and so is usually stronger at a given latitude than in its longitudinal average. These waves are especially noticeable in the winter hemisphere as illustrated in Fig. 4.

GLOBAL AVERAGE ENERGY BALANCE

Solar radiation of about 342 W/m^{-2} entering Earth’s atmosphere is absorbed and scattered by molecules. The major gaseous absorbers of solar radiation are water vapor in the troposphere and ozone in the stratosphere. Clouds and aerosols likewise scatter and absorb. Clouds are the dominant scatterer and so substantially enhance the overall planetary reflected radiation, whose ratio to incident solar radiation, about 0.31, is referred to as *albedo*. Thermal infrared radiation, referred to as *longwave*, is controlled by clouds, water vapor, and other greenhouse gases. Figure 5 (4) illustrates a recent estimate of the various terms contributing to the global

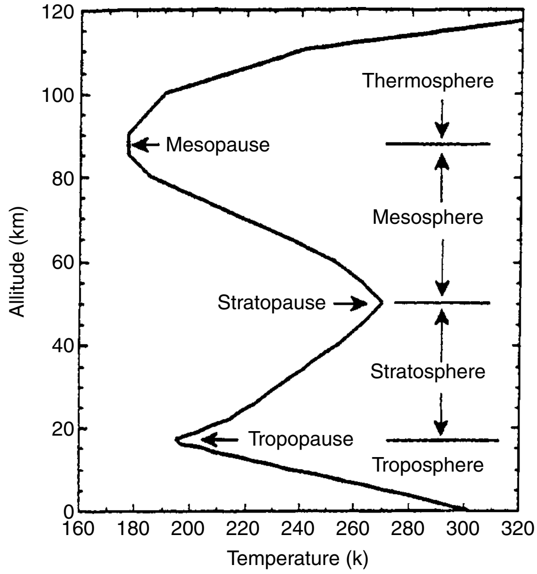

Figure 1. Main zones of the atmosphere defined according to the temperature profile of the standard atmosphere profile at 15°N for annual-mean conditions (1).

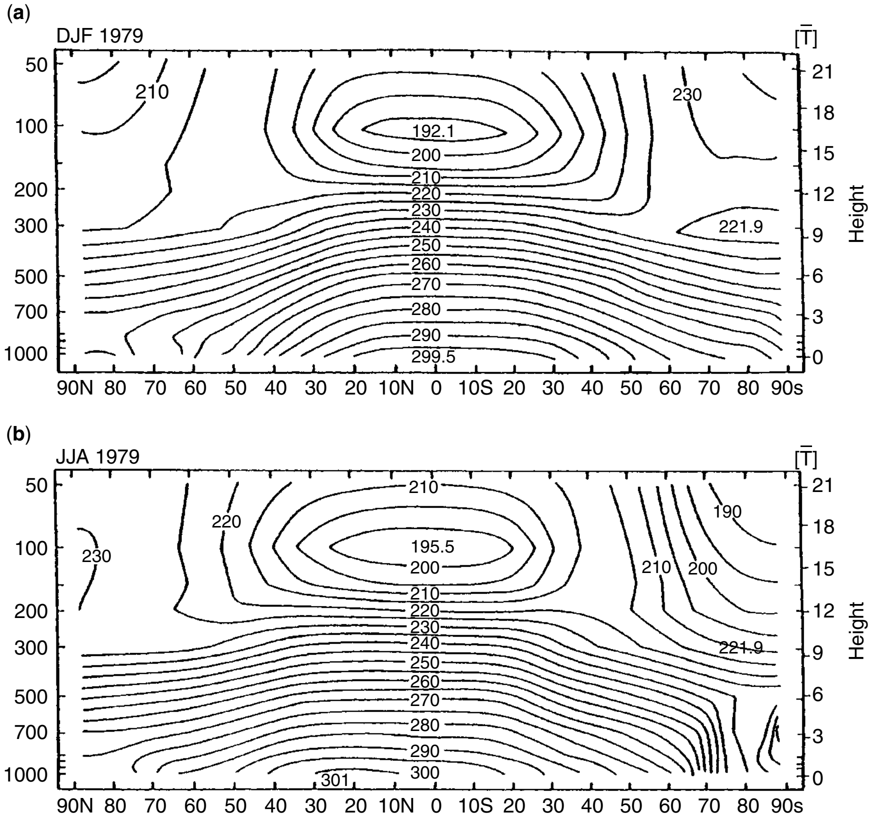

Figure 2. Zonal mean meridional cross sections of temperature during two seasons. (*a*) December 1978–February 1979 and (*b*) June–August 1979. Height scale is approximate and the contour interval is 5 K (2).

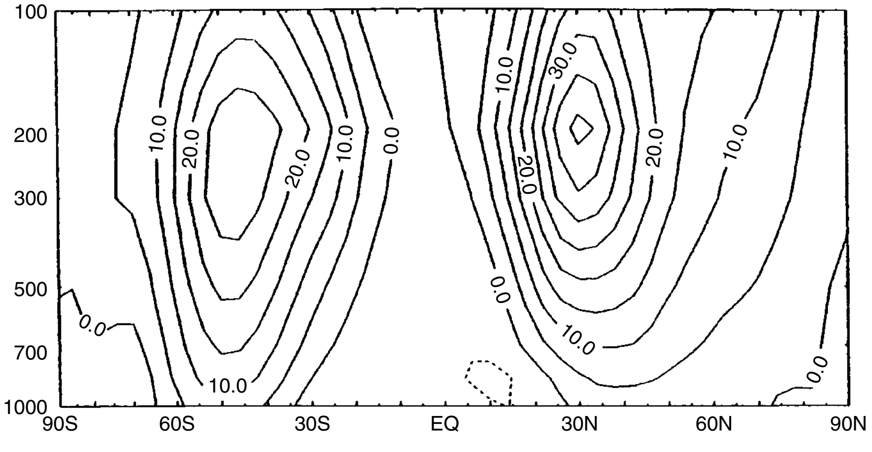

Figure 3. Meridional cross sections of longitudinally averaged zonal wind (top panels, m/s) for DJF (3).

energy balance. The latent heat from global average precipitation of about 1.0 m per year is the dominant nonradiative heating term in the atmosphere.

Because of the seasonally varying geometry of Earth relative to the sun, and the differences in cloudiness and surface albedos, there are substantial variations in the distribution of absorbed solar radiation at the surface and in the atmosphere, as likewise in the transfer of latent heat from the surface to the atmosphere. This heterogeneous distribution of atmospheric heating drives atmospheric wind systems, either directly or through the creation of available potential energy, which is utilized to maintain random occurrences of various kinds of instabilities, such as thunder-storms and wintertime cyclonic storm systems. These dynamical systems hence act to redistribute energy within the atmosphere and so determine the distributions of temperature and water vapor. Likewise, the balances at the surface between fluxes of radiative, latent, and thermal energies determine surface temperatures and soil moistures. The properties of the near-surface air we live in are determined by a combination of surface and atmospheric properties, according to processes of the atmospheric boundary layer. Thus climatic anomalies in surface air may occur either because of some shift in atmospheric circulation patterns or through some modification of surface properties such as those accompanying deforestation or the development of an urban area.

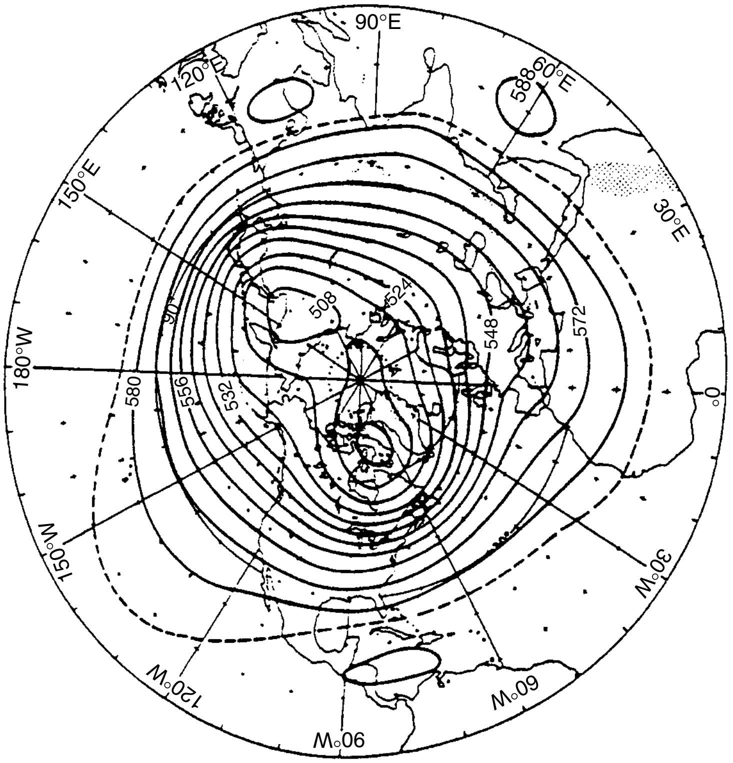

Figure 4. Mean 500-mb contours in January, Northern Hemisphere. Heights shown in tens of meters (3).

THE ATMOSPHERIC BOUNDARY LAYER

The term *boundary layer* is applied in fluid dynamics to layers of fluid or gas, usually relatively thin, determining the transition between some boundary and the rest of the fluid. The atmospheric boundary layer is the extent of atmosphere that is mixed by convective and mechanical stirring originating at Earth's surface. Such stirring is commonly experienced by airplane travelers as the bumps that occur during takeoff or landing, especially in the afternoon, or as waves at higher levels in flying over mountainous regions. The daytime continental boundary layer, extending up to several kilometers in height, is most developed and vigorously mixed, being the extent to which the daytime heating of the surface drives convective overturning of the atmosphere. The land cools radiatively at night, strongly stabilizing the atmosphere against convection, but a residual boundary layer extends up to about 100 m stirred by the flow of air over the underlying rough surface. This diurnal variation of fluxes over the ocean is much weaker and the boundary layer is of intermediate height. The temperature of the atmosphere, when stirred by dry mixing, decreases at a rate of 9.8 K/km. Above the boundary layer, temperatures decrease less rapidly with height, so that the atmosphere is stable to dry convection. A layer of clouds commonly forms at the top of the daytime and oceanic boundary layers and contributes to the convection creating the boundary layer through its radiative cooling (convection results from either heating at the bottom of a fluid or cooling at its top). Also, at times the clouds forming near the top of the boundary layer can be unstable to moist convection, and so convect upward through a deep column such as in a thunderstorm.

ATMOSPHERIC HYDROLOGICAL CYCLE

The storage, transport, and phase changes of water at the surface and in the atmosphere are referred to as the hydrological cycle. As already alluded to, the hydrological cycle is closely linked to and driven by various energy exchange processes at the surface and in the atmosphere. On the scale of continents, water is moved from oceans to land by atmospheric winds, to be carried back to the oceans by streams and rivers as elements of the land hydrological cycle. Most of the water in the atmosphere is in its vapor phase, but water that is near saturation vapor pressure (relative humidity of 100%) converts to droplets or ice crystals depending on temperature and details of cloud physics. These droplets and crystals fall out of the atmosphere as precipitation. The water lost is replenished by evaporation of water at the surface and by vertical and horizontal transport within the atmosphere. Consequently, much of the troposphere has humidities not much below saturation. Saturation vapor pressure increases rapidly with temperature (about 10% per kelvin of change). Hence, as illustrated in Fig. 6, the climatological concentrations of water vapor vary from several percent or more when going from near-surface air to a few parts per million near the tropical tropopause. Water vapor concentrations in the stratosphere are close to that of the tropical tropopause, probably because much of the air in the lower stratosphere has been pumped through the tropical tropopause by moist convection.

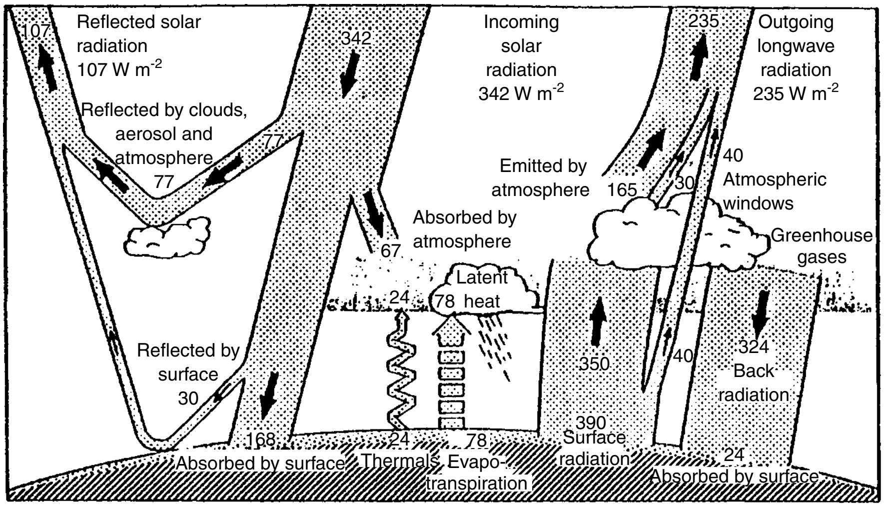

Figure 5. Earth's annual global mean energy budget based on the present study. Units are W/m^2 (4).

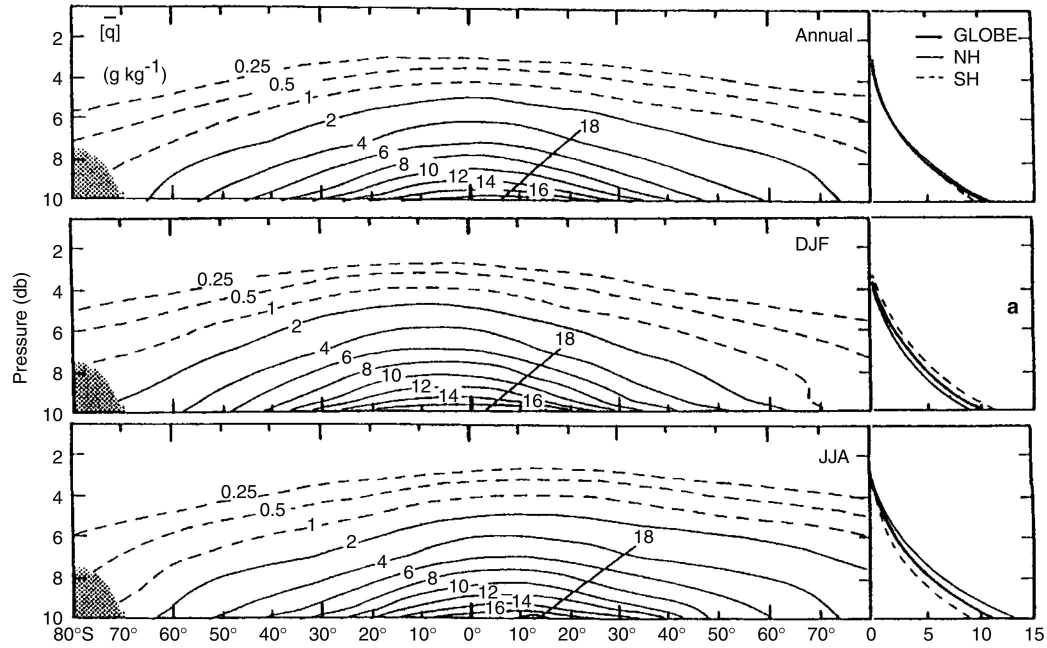

Figure 6. Zonal mean cross sections of the specific humidity in g/kg for annual, DJF, and JJA mean conditions. Vertical profiles of hemispheric and global mean values are shown on the right (5).

CLIMATE OF THE STRATOSPHERE

The dominant radiative processes in the stratosphere are the heating by absorption of solar ultra violet (UV) radiation and cooling by thermal infrared emission from carbon dioxide and, to a lesser extent, ozone molecules. The stratospheric absorption of UV largely determines how much harmful UV reaches the surface. Ozone in the upper troposphere and lower stratosphere additionally adds heat by absorption of thermal emission from the warmer surface and lower layers. The stratosphere, furthermore, enhances the greenhouse warming of CO_2 in the troposphere through substantial downward thermal emissions to the troposphere.

How changes of ozone change stratospheric and tropospheric radiative heating depends on the amounts of overlying ozone and, for thermal effect, on pressure and radiative upwelling depending on underlying temperatures.

Besides radiative processes, stratospheric climate is characterized by its temperature and wind patterns and by the chemical composition of its trace gases. At mid-stratosphere, temperature increases from winter pole to summer pole with an accompanying eastward jet stream in the winter hemisphere extending upward from the tropospheric jet steam. This wind configuration allows planetary wave disturbances to propagate into the stratosphere, contributing significant temporal and longitudinal variabilities. Conversely, the westward jet, found in the summer stratosphere attenuates wave disturbances from below, and so is largely zonally symmetric, changing only with the seasonal heating patterns.

THE CRYOSPHERE

The term *cryosphere* refers to the components of the climate system dominated by water in its frozen phase, that is, in high latitudes and extratropical winter conditions. Elements include snow, its distribution and depths, sea ice, its distribution and properties, high-latitude ice caps, and temperate glaciers. The largest volume of frozen water is stored in ice caps, and glaciers. This storage acts to remove water from the oceans. How it changes with climate change is, hence, of interest for determining changing sea levels.

Ice is highly reflective of sunlight, especially in crystal form. The loss of solar heating because of this high albedo acts to substantially reduce high-latitude temperatures especially in spring and early summer where near-maximum solar radiation sees white snow-covered surfaces. This high albedo can be substantially masked by cloud cover and, over land, tall vegetation such as conifer forests.

THE OCEAN

Oceans are a major factor in determining surface temperatures and fluxes of water into the atmosphere. They store, release, and transport thermal energy, in particular, warming the atmosphere under wintertime and high-latitude conditions, and cooling it under summer and tropical conditions.

How the oceans carry out these services depends on processes coupling them to the atmosphere. Atmospheric winds push the oceans into wind-driven circulation systems. Net surface heating or cooling, evaporation, and precipitation determine oceanic densities through controlling temperature and salinity, hence oceanic buoyancy. This net distribution of buoyancy forcing drives “thermohaline” overturning of the ocean, which acts to transport heat. Climate of the surface layers of the ocean includes the depth to which waters are stirred by waves and net heating or cooling. Heating acts to generate shallow warm stable layers, while cooling deepens the surface mixed layers. Under some conditions, convective overturning of cold and/or high-salinity water can penetrate to near the ocean bottom.

BIBLIOGRAPHY

1. Hartmann, D.L. (1994). Atmospheric temperature, Chapter 1.2. In: *Global Physical Climatology*. Academic Press, San Diego.
2. Grotjahn, R. (1993). Zonal average observations, Chapter 3. In: *Global Atmospheric Circulations: Observations and Theories*. Oxford University Press, New York.
3. Holton, J.R. (1992). The observed structure of extratropical circulations, Chapter 6.1. In: *An Introduction to Dynamic Meteorology*. Academic, San Diego.
4. Kiehl, J.T. and Trenberth, K.E. (1997). Earth's annual global mean energy budget. *J. Clim.* **78**: 197–208.
5. Peixoto, J.P. and Oort, A.H. (1992). Observed atmospheric branch of the hydrological cycle, Chapter 12.3. *Physics of Climate*. American Institute of Physics, New York.

CLIMATE AND SOCIETY

Michael H. Glantz

(from *The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts*, Wiley 2003)

At the turn of the twentieth century, scholars who wrote about the interplay between climate and society did so based on their perceptions of climate as a boundary constraint for the development prospects of a society. Perceptions of climate were used as an excuse to dominate societies in Africa, Asia, and Latin America. As a result, climate–society studies soon became viewed as a colonial ploy to control populations in developing areas in the tropics. Perhaps the most cited book in this regard was written by Ellsworth Huntington, *Climate and Civilization*, published in 1915 (1). In his view, inhabitants of the tropics were destined to accept lower levels of economic and social development because their climate setting was not conducive to lively (i.e., productive) human activity or an aggressive work ethic. According to Huntington, tropical climate was the main culprit causing people in the tropics to be less productive than people in temperate regions. Huntington argued that the temperate climate has an energizing effect on humans. With the growing belief that such an argument was racist in intent, Huntington's work was challenged, and discussion of the various ways in which climate might influence human behavior was stifled for decades, notwithstanding a few notable exceptions. One such exception is entitled *Climate and the Energy of Nations* (2) in which Markham referred to the “air-conditioning revolution,” a revolution based on the development and spread of a new technology into the tropical areas. Markham asserted that technology brings islands of temperate-zone climate into the tropics, thereby generating a more aggressive work ethic.

Following the end of World War II and the onset of the Cold War between Soviet-style communism and Western capitalism and democracy, attention of governments turned to Cold War conflicts, avoidance of nuclear war, searches for allies, and decolonization. The major Cold War nations were in a competition to show that *their* approach to economic development was the only way for the newly independent countries to follow. A main stated objective was their intent to assist these countries to become food secure based on the nation's resources. Consideration of climate was making its way back into the discussions of economic development in developing countries. Once again interest was raised with regard to climatic constraints to economic development in tropical countries.

In the 1950s and 1960s, attention focused on decolonization and political development of the newly independent states (e.g., 3). In the mid-1970s, a World Bank report about the economic prospects for developing countries—*The Tropics and Economic Development: A Provocative Inquiry into the Poverty of Nations*—hinted at the economic, social, and political problems caused by climate variability from one year to the next. Its author (4) noted that recurrent droughts in northeast Brazil are a chronic constraint on the region's economic development prospects. His reference to interannual climate variability was brief and unelaborated. However, climate as a boundary constraint was starting to give way to climate as something that societies might be able to forecast and cope with, at least in its extremes.

In the 1970s, attention focused on how the vagaries of weather exposed hundreds of millions of people to hunger and, depending on the socioeconomic situation in a particular country, to famine as well (e.g., 5,6). Thus, there was a growing number of examples of the notion that climate was not really a boundary constraint to the level of development that a people or culture could expect to attain. This notion began to give way to the belief that variability in climate, from one year to the next or one decade to the next, could be coped with so as to soften the impacts of climate variability and weather extremes on agriculture and livestock and, more generally, on the productivity of the land's surface (e.g., 7,8).

Recall that the 1970s was a disruptive decade with respect to climate: 5 years of drought in the West African Sahel (5); failure of the Soviet harvest and subsequent large-scale, low-cost grain purchases by the Soviet Union in the early 1970s (9); the global food crisis (10); talk of a possible return to an ice age (e.g., 11,12); the Ethiopian famine (13); drought-related coups in sub-Saharan Africa; drought in the wheat-producing Canadian prairie provinces (7); the first drop in global fish catches since the end of World War II (10), and so forth.

A devastating 5-year drought from 1968 to 1973 in the West African Sahel and its associated death and environmental destruction in the region drew attention to the impacts on household and village responses to prolonged, multiyear droughts. Widespread droughts around the globe in 1972–1973, famines in West Africa and Ethiopia, blamed for the most part on an El Niño event, along with the drop in fish landings, prompted the U.N. Secretary General to convene a series of UN-sponsored world conferences on food (1974), population (1974), human settlements (1976), water (1977), desertification (1977), climate (1979) and technology (1979).

Thus, toward the middle of the 1970s, at least five new major climate-related scientific issues emerged: the effect of chlorofluorocarbons (CFCs) on the ozone layer in the stratosphere, talk of an impending Ice Age suddenly shifted to talk of a human-induced global warming, acid rain, desertification, and El Niño. Each of these issues raised interest in climate–society interactions to higher levels among researchers in different disciplines, government agencies, economic sectors, the media, and the public. Societies around the globe responded (and continue to respond) in different ways to each of these climate-related issues. For example, desertification is an environmental issue that is of great concern to African countries.

North Americans, however, refused to accept the view that desertification could occur in the U.S. West as a result of mismanagement of the land's surface, while noting that *desertification* was the plight of poor developing countries in Africa. The term desertification first appeared in a report on the destruction of dry forests in central Africa by a French forester (14). Since then, the concept of desertification has been expanded to include such land degradation processes as soil erosion, wind deflation, soil salinization, water logging, livestock overgrazing, and soil trampling. While many of these processes were exposed during the prolonged drought in the West African Sahel and then labeled as desertification, it is not difficult to show that similar processes also take place in the U.S. West.

The acid rain issue was addressed in the United States with the implementation by the U.S. Congress of a decade-long national assessment called NAPAP (National Acid Precipitation Assessment Program). Stratospheric ozone depletion was addressed globally in the 1980s with the development of international legal instruments culminating in the Montreal Protocol of 1987 and, later, amendments to it (15).

It was in the early 1970s, 1972–1973 to be exact, that an El Niño event (defined briefly as an invasion of warm water from the Western Pacific into the central and eastern equatorial Pacific Ocean) attracted global attention. An event in 1982–1983, the biggest in a century until that time, captured the full attention of the scientific community and various governments as a natural phenomenon that spawned hazards around the globe. Such hazards included, but were not limited to, droughts, floods, frosts, fires and food shortages, famine, and disease. Ever since the mid-1970s, research funding of El Niño–related research has been growing along with international interest in the phenomenon and its societal and environmental impacts. The extraordinary El Niño event of 1997–1998 helped to make El Niño and its cold counterpart, La Niña, household words throughout much of the world. Only at the end of the twentieth century did La Niña events become of serious interest to the El Niño research and forecasting communities (16). This belated interest is even more surprising given the scientific observation that tropical storms and hurricanes in the Atlantic Basin and in the Gulf of Mexico tend to increase in number during La Niña events and drop in number during El Niño events.

Global warming is an environmental issue that arose out of discussions and governmental and scientific concerns about the possibility of a global cooling. It was first suggested in 1896 by Swedish chemist Arrhenius (17,18) that the burning of coal by human societies would add enough extra carbon dioxide into the air to eventually heat up Earth's atmosphere by a few degrees Celsius. This issue was revisited in the 1930s by Callendar (19), who thought that a human-induced global warming of the atmosphere could stave off the imminent recurrence of an ice age. The issue was again revisited in the 1950s when global warming was looked at in neutral terms, as an experiment that societies were performing on the chemistry of the atmosphere, for which the outcome is unknown (20).

It was not until the mid-1970s that human-induced global warming began to be viewed as an adverse event for future generations of human societies and ecosystems that might not be able to adapt to the rate of warming expected to occur. The cause of the warming was attributed to the increasing amounts of greenhouse gases (CO_2, CFCs, CH_4, NO_x, collectively referred to as GHGs) being emitted into the atmosphere as a result of human activities. Carbon dioxide is a product of the burning of fossil fuels, and its amount in the atmosphere has been rising since the onset of the Industrial Revolution in the late 1700s. Tropical deforestation also contributes carbon dioxide to the atmosphere. Tropical forests have served as sinks for carbon dioxide, pulling it out of the air and storing it. When trees are felled, decompose or burned, the stored carbon is emitted into the air.

Chlorofluorocarbons (CFCs), a greenhouse gas as well as a stratospheric "ozone eater," are man-made chemicals first discovered in the 1920s for use as a refrigerant. Methane resulting from livestock rearing (e.g., cattle, pigs) and from the increasing number of landfills is another greenhouse gas. Nitrous oxides are used by farmers in fertilizers and have been widely applied to agricultural lands around the globe in increasing amounts since the end of World War II. Of these major greenhouse gases, carbon dioxide is seen at the main culprit in the global warming debate.

Current scientific research suggests that the level of climate change that might be expected (at current rates of greenhouse gas emissions) is on the order of 1.5 to 4.5 °C by the end of the twenty-first century (21–23). Concerned with the prospects of a changing global climate, many nations have come together to call for a technical assessment of the state of the science through the Intergovernmental Panel on Climate Change (IPCC).

The degree of warming, however, is dependent on numerous factors: the rate at which GHGs continue to be emitted into the atmosphere, the shift by societies to alternative energy sources, the rate of tropical deforestation, the residence time of GHGs in the atmosphere (several of these gases will remain in the atmosphere for decades to centuries), the development of methods to sequester carbon (i.e., taking it from the atmosphere and binding it in some way in Earth's land surface, vegetation, or oceans), and so forth. Some degree of global warming is inevitable, given the residence time of the GHGs already emitted into the atmosphere. This means that societies around the globe, from local to national, must attempt to ascertain how a warmer global climate regime might affect regional and local climates. Will there be more extreme climate events (such as droughts, floods, frosts, fires) or fewer? These societies must also seriously consider nationally, as well as collectively in cooperation with other countries, the most effective way(s) to cope with the potentially adverse impacts of some degree of human-induced global warming.

Coping mechanisms for climate change likely to occur decades in the future can be divided into three categories: preventive, mitigative and adaptive measures. *Preventive* measures are designed to prevent the increased buildup of GHGs in human-induced global warming, acid rain, desertification, and El Niño. Each of these issues raised interest in climate–society interactions to higher levels among researchers in different disciplines, government agencies, economic sectors, the media, and the public. Societies around the globe responded (and continue to respond) in different ways to each of these climate-related issues. For example, desertification is an environmental issue that is of great concern to African countries.

Climate-related surprise is not a black-and-white condition. People are hardly ever either totally surprised or never surprised. There are shades of surprise with regard to human responses to the same climate-related event. They can be hardly surprised, mildly surprised, somewhat surprised, very surprised, extremely surprised or totally surprised (NB: each of these examples was taken from the scientific literature). Myers (24, p. 358) introduced the interesting notion of "semisurprised." Thus, surprise may best be described in "fuzzy" terms with the degree of surprise dependent on several intervening variables such as personal experience, core beliefs, expectations, or knowledge about a phenomenon or about a geographic location.

One could argue that there are knowable as well as unknowable surprises (25). Knowable refers to the fact that some climate surprises can be anticipated (24). For example, certain parts of the globe are drought prone. It is known that drought will likely recur. What is not known is exactly when it will take place, how long it will last, how intense it will be, or where its most devastating impacts are likely to occur. El Niño is in this category. While we have now come to expect these events to recur, we do not know when that will happen or what it will be like. The uncertainty then cascades down the "impacts chain," and as we speculate about likely impacts of an El Niño, the degree of uncertainty will increase.

Take, for example, the 1997–1998 El Niño. Even with the best monitoring and observing system in the world focused on minute changes in various aspects of the tropical Pacific Ocean, forecasters and modelers were unable to predict the onset of one of the biggest El Niño events in the past 100 years. Nor were they able to predict the course of development of that event. They were better than in earlier times, however, at predicting some of its impacts on societies in certain parts of the globe, especially those where the influences of changes in the sea surface temperatures in the tropical Pacific are known to be strong.

Societies (and their scientists) are on a learning curve with regard to the various ways that climate variability and climate change might affect climate-related human activities. They must avoid becoming complacent as a result of a belief that they fully understand atmospheric processes or their impacts. They must accept that there will be climate surprises in the future, even if the global climate does not change. They must learn from past experiences on how best to cope with the vagaries of climate (26).

Many countries now realize that climate-related problems do not stop at international boundaries. There are many transboundary issues that demand regional (if not international) cooperation, given that countries share river basins, inland seas, airsheds, the global atmosphere as well as the onslaught and impacts of extreme meteorological events such as droughts, floods, and tropical storms.

While climate-related anomalies cannot be prevented, societal preparation for, and response to, their adverse impacts can be improved through better knowledge of the direct and indirect ways in which atmospheric processes interact with human activities and ecological processes. The enhancement of such knowledge will lead to better forecasts as well as better computer modeling of the interactions among land, sea, and air. A society forewarned of climate-related hazards is forearmed to cope with those hazards more effectively.

BIBLIOGRAPHY

1. Huntington, E. *Civilization and Climate*. Yale University Press, New Haven, Connecticut, 1915, rev, ed. 1924, reprinted by University Press of the Pacific, 2001.
2. Markham, S.F. (1944). *Climate and the Energy of Nations*. Oxford University Press, London.
3. Pye, L. (1966). *Aspects of Political Development*. Little, Brown and Co., Boston, Massachusetts.
4. Kamarck, A.M. (1976). *The Tropics and Economic Development: A Provocative Inquiry into the Poverty of Nations*. The Johns Hopkins University Press, Baltimore, Maryland.
5. Glantz, M.H. (Ed.). (1976). *The Politics of Natural Disaster: The Case of the Sahel Drought*. Praeger Press, New York.
6. Sen, A. (1981). *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford University Press, Oxford, UK.
7. Glantz, M.H. (Ed.). (1977). *Desertification: Environmental Degradation in and around Arid Lands*. Westview Press, Boulder, Colorado.
8. Hare, K. (1977). Connections between climate and desertification. *Environ. Conserv.* **4**(2): 82.
9. Trager. (1975).
10. Brown, L., and Eckholm, E.P. (1974). *By Bread Alone*. Praeger Press, New York.
11. Ponte, L. (1976). *The Cooling*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
12. *Weather Conspiracy: The Coming of the New Ice Age*, Ballentine Books, New York, compiled by The Impact Team, 1977.
13. Wolde Mariam, M. (1984). *Rural Vulnerability to Famine in Ethiopia, 195777*, Vikas Publishing House, New Delhi.
14. Aubreville, A. (1949). *Climate, Forests and Desertification in Tropical Africa*, Sociéte d'Editions Géographiques, Maritimes et Coloniales.
15. Benedick, R.E. (1998). *Ozone Diplomacy: New Directions in Safeguarding the Planet*. Harvard University Press, Cambridge, Massachusetts, Enlarged Edition.
16. Glantz, M.H. (Ed.). (2002). *La Niña and Its Impacts: Facts and Speculation*. United Nations University Press, Tokyo, Japan.
17. Arrhenius, S. (1896). On the influence of carbonic acid in the air upon the temperature of the ground. *Philos. Mag.* **41**: 237–276.
18. Arrhenius, S. (1908). *Worlds in the Making*. Harper & Brothers, New York.
19. Callendar, G.S. (1938). The artificial production of carbon dioxide and its influence on temperatures. *Q. J. Roy. Meteor. Soc.* **64**: 223–237.
20. Revelle, R.R., and Suess, H.E. (1957). Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric CO_2 during the past decades. *Tellus* **9**: 18–27.
21. IPCC (1990). *Climate Change: The IPCC Scientific Assessment*. Cambridge University Press, Cambridge, UK.
22. IPCC (1996). *Climate Change 1995: The Science of Climate Change*, Contribution of Working Group I to Second Assessment Report, Cambridge University Press, Cambridge, UK.
23. IPCC (Intergovernmental Panel on Climate Change). (2001). *Climate Change 2001: Impacts, Adaptation, and Vulnerability*, Contribution of Working Group II to Third Assessment Report, Cambridge University Press, Cambridge, UK.
24. Myers, N. (1995). Environmental unknowns. *Science* **269**: 358–360.
25. Streets, D.G., and Glantz, M.H. (2000). Exploring the concept of climate surprise. *Global Environ. Chang.* **10**: 97–107.
26. Glantz, M.H. (Ed.). (1988). *Societal Responses to Regional Climatic Change: Forecasting by Analogy*. Westview Special Study, Boulder, Colorado.

READING LIST

Smith, J.B., Bhatti, N., Menzhulin, G., Benioff, R., Bodyko, M.I., Campos, M., Jallow, B., and Rijsberman, F. (Eds.). (1996). *Adapting to Climate Change: An International Perspective*. Springer-Verlag, New York.

WHAT IS CLIMATOLOGY?

National Drought Mitigation
Center

"CLIMATE IS WHAT YOU EXPECT. WEATHER IS WHAT YOU GET."

Weather is the condition of the atmosphere over a brief period of time. For example, we speak of today's weather or the weather this week. Climate represents the composite of day-to-day weather over a longer period of time.

People in Minneapolis–St. Paul expect a white Christmas, and people in New Orleans expect very warm, humid summers. And a traveler going to Orlando, Florida, in March will not pack the same kind of clothing as a traveler going to Vail, Colorado, in March. These examples show how climate influences our daily lives. Additionally:

- Our houses are designed based on the climate where we live.
- Farmers make plans based on the length of the growing season from the last killing freeze in the spring to the first freeze in the fall.
- Utility companies base power supplies on what they expect to be the maximum need for heating in the winter and the maximum need for cooling in the summer.

A climatologist attempts to discover and explain the impacts of climate so that society can plan its activities, design its buildings and infrastructure, and anticipate the effects of adverse conditions. Although climate is not weather, it is defined by the same terms, such as temperature, precipitation, wind, and solar radiation.

Climate is usually defined by what is expected or "normal", which climatologists traditionally interpret as the 30-year mean. By itself, "normal" can be misleading unless we also understand the concept of variability. For example, many people consider sunny, idyllic days normal in southern California. History and climatology tell us that this is not the full story. Although sunny weather is frequently associated with southern California, severe floods have had a significant impact there, including major floods in 1862 and 1868, shortly after California became a state. When you also factor in severe droughts, most recently those of 1987–94, a more correct statement would be that precipitation in southern California is highly variable, and that rain is most likely between October and April.

The misconception that weather is usually normal becomes a serious problem when you consider that weather, in one form or another, is the source of water for irrigation, drinking, power supply, industry, wildlife habitat, and other uses. To ensure that our water supply, livelihoods, and lives are secure, it is essential that planners anticipate variation in weather, and that they recognize that drought and flood are both inevitable parts of the normal range of weather.

HOW DOES CLIMATOLOGY HELP US PREPARE FOR DROUGHT?

Climatology provides benchmarks, such as the drought of record. The drought of record is the drought remembered as having the greatest impact on a region. Most of us are not consciously aware of how much the climate fluctuates from one decade or century to the next. One way for reservoir managers, municipal water suppliers, and other planners to check reality is to compare expectations of water supply against a region's drought of record. But caution is necessary here: although the weather conditions could recur, the impacts would likely be very different. For most of the country, the drought of record was 30 to 60 years ago, and population concentrations and water use patterns have shifted substantially since then. Planners need to consider and watch out for a variety of problems and misconceptions.

Specifically, climatology answers crucial questions such as:

- How often does drought occur in this region?
- How severe have the droughts been?
- How widespread have the droughts been?
- How long have the droughts lasted?

Examining water supplies and understanding the impact of past droughts help planners anticipate the effects of drought:

- What would happen if the drought of record occurred here now?
- Who are the major water users in the community, state, or region?
- Where does our water supply come from and how would the supplies be affected by a drought of record?
- What hydrological, agricultural, and socioeconomic impacts have been associated with the various droughts?
- How can we prepare for the next drought of record?

CLOUD SEEDING

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

Cloud seeding is the deliberate treatment of certain clouds or cloud systems to affect the precipitation process within those clouds. The treating of clouds can alter weather conditions, thus cloud seeding is also called weather modification. The technology that led to cloud seeding has been developed only during the past 60 years. Yet as this technology becomes more efficient, increased worldwide application of cloud seeding for practical use is likely to continue into the future.

THE CLOUD SEEDING PROCESS

To understand how cloud seeding works, one must understand some facts about the weather. All air in the atmosphere contains moisture. Warm air rises from the earth's surface and begins to cool. As the air cools, the moisture condenses into tiny droplets that make up clouds. A cloud is almost 100% air. The tiny droplets composing clouds are not heavy enough to fall to the ground until they merge with millions of other droplets at temperatures below 32 °F and interact with dust, salt, or sand particles.

One type of cloud seeding, known as cold cloud seeding, introduces silver iodide and other agents to enhance ice crystallization in clouds colder than 32 °F. This is often called as the static seeding effect. Once the droplets within the clouds freeze, the resulting ice crystals grow at the expense of the water droplets surrounding them, a process called sublimation. Other crystals grow through contact with neighboring droplets; this is known as riming. Through these two processes, the tiny crystals form snowflakes. If these snowflakes are heavy enough, they fall from the clouds and, depending on the temperatures below, will come to the earth as snow, raindrops, or a mixture.

The other type of cloud seeding, warm cloud seeding, produces rainfall from clouds that are above 32 °F. This process involves introducing additional condensation nuclei (salt particles), which cause additional water droplets to condense within the clouds. If the collision of these particles makes the droplets heavy enough, precipitation can fall from the clouds. In each case, it will usually take the clouds 20 to 30 minutes to produce rain, making it crucial to monitor cloud movement.

USES OF WEATHER MODIFICATION

The primary goal of most weather modification projects is to increase levels of precipitation or to suppress fog or hail. Water agencies, local municipalities, farmers, ranchers, hydroelectric power facility operators, ski resort owners, and others sponsor cloud seeding activities. To date, cloud seeding has successfully stimulated precipitation in more than 50 countries. More than half of the United States now has some type of regulation concerning cloud seeding programs.

More recently, cloud seeding has been used to suppress certain damaging effects of weather. Airports have employed programs to disperse fog levels and increase flight visibility. And in areas damaged by hail, programs have been undertaken to decrease hailstorms.

As technology continues to enhance the weather modification process, it can be assumed that more entities will sponsor cloud seeding programs. Cloud seeding technology is highly portable and very flexible to changing weather conditions. For those who use it, cost/benefit ratios are typically very favorable.

EFFECTIVENESS OF CLOUD SEEDING

From the earliest experiment, which produced just a few droplets of rain, cloud seeding has progressed to make some significant and valuable impacts on weather. Various studies have documented the effects of different programs. For the most part, these programs are overwhelmingly successful. To augment precipitation, well-designed and well-conducted projects yield an average winter precipitation of 5 to 20% more in continental regions and 5 to 30% more in coastal areas; they have yielded as much as 100% more in warm weather. Increases depend on a variety of factors, including spatial coverage of suitable cloud systems and the frequency of different systems.

ENVIRONMENTAL HEALTH CONCERNS OF WEATHER MODIFICATION

To date, no significant environmental problems have been attributed to cloud seeding programs, though government agencies, private firms, and research institutions have conducted many studies. Researchers believe that any negative effects are minimal because relatively little seeding material is used compared to the large surface area that is targeted. For example, the most common seeding material, silver iodide, will yield a concentration of less than 0.1 microgram per liter in rainwater or snow. The U.S. Public Health Service states that the acceptable amount of silver iodide in water is 50 micrograms per liter.

One other common question about cloud seeding is whether the stimulation of rainfall in one area results in decreased rainfall in other areas. It does not. Clouds are inefficient in the way they gather and distribute moisture. They never gather or release all the moisture that is available; rather, clouds gather only about 1% of the moisture in the atmosphere. Even if a seeding program tripled the efficiency of cloud formation, the cloud would still contain only about 3% of the moisture, leaving 97% available. Furthermore, some analyses of precipitation data downwind from seeding projects have indicated small increases in precipitation. To date, no scientific studies have shown that some areas receive precipitation at the expense of their neighbors.

FUTURE OF CLOUD SEEDING

Despite all of the advances made during the past 60 years, weather modification continues to receive relatively little support. Many states and the federal government acknowledge and regulate cloud seeding, but few provide significant funding for weather modification studies or projects. Most scientists insist that more research must be done and more data must be gathered before weather-changing programs garner public trust on a larger scale. However, most researchers are optimistic that new studies and technological advances will continue to advance the science of weather modification.

READING LIST

American Society of Civil Engineers. (1995). *Guidelines for Cloud Seeding to Augment Precipitation.* ASCE Manuals and Reports on Engineering Practice No. 81.

Stauffer, N.E. and Williams, K. (2000). *Utah Cloud Seeding Program, Increased Runoff/Cost Analysis.* Technical Report, Utah Department of Natural Resources, Division of Water Resources.

Weather Modification Advisory Board. (1978). *The Management of Weather Resources. Report to the Secretary of Commerce*, 2 volumes.

Klien, D.A. (1978). *Environmental Impacts of Artificial Ice Nucleating Agents*. Dowden, Hutchinson & Ross, Inc., Stroudsburg, PA.

Brown, K.J., Elliot, R.D., and Edelstein, M.W. (1978). *Transactions of Total-Area Effects of Weather Modification.* Report to the National Science Foundation on a workshop held August 8–12, 1977, Fort Collins, CO.

Cloud Seeding (Weather Modification) Frequently Asked Questions—http://www.xmission.com/~nawc/wmfaq.html.

The Science of Cloud Seeding—http://twri.tamu.edu/twripubs/WtrResrc/v20n2/text-1.htm.

Cloud Seeding Circus.com—www.cloudseedingcircus.com.

Al Weather Modification Page—www.atmos-inc.com/weamod.html.

CONDENSATION

Aldo Conti
Frascati (RM), Italy

Condensation is the physical process by which a vapor changes to its liquid state. Condensation happens when the temperature of the vapor decreases below the so-called dew point. In physics, this sort of process is called a phase change because the matter involved changes its state (i.e., from gaseous to liquid, in this case). Water condensation is evident in our atmosphere; it produces fog, mist, dew, clouds, and rain, depending on the conditions. It can be seen on a cold sheet of glass, where condensed water forms a maze of droplets.

Condensation depends mainly on temperature, and it is a process that happens at the molecular level. When the temperature is high, the molecules in a vapor have plenty of energy and collide at high speed, which means that the molecules bounce immediately and do not stay together long enough to establish a bond. In this situation, the vapor remains. But when the temperature, and therefore the speed, decreases, then the molecules can stick to each other. The result is a droplet of liquid. Condensation is ruled not only by temperature, but by pressure as well, which is the reason why one speaks about dew point and not dew temperature. As a general rule, the dew point temperature increases with pressure.

Condensation, with evaporation, is very important in the water cycle of the earth. By condensation, water falls back as rain, hail, or snow and becomes available again for human use.

COSMIC WATER

D.L. Marrin
Hanalei, Hawaii

Once believed by science to be the substance that distinguished the earth from the rest of the universe, it is now understood that water is ubiquitous in the cosmos—not only as ice and vapor, but also as liquid. Sophisticated scientific instruments can detect cosmic water on the basis of the light, or other electromagnetic, waves that it absorbs and emits (1). Unlike the legendary "waters of chaos" that gave rise to the material world, water's component hydrogen and oxygen atoms owe their existence to the Big Bang and to the stars, respectively.

HYDROGEN AND OXYGEN

Hydrogen is both the simplest and the most abundant atom in the universe. It represents about 75% of the atomic mass in the cosmos. The word *hydrogen* literally means "water-forming" according to the Greek language from which it is derived. Hydrogen atoms are generally traced back to the so-called Big Bang, when a tremendous amount of energy was released and subsequently expanded into what we call our universe. As the newborn universe began to cool, subatomic and atomic particles (e.g., quarks, protons, electrons) were initially created and later drawn together by a number of fundamental forces to form atoms. Possessing one proton and one electron, hydrogen is believed to have been the first atom created. As more hydrogen atoms were created in the early universe, they coalesced into dense gas clouds that contained much of the conventional matter.

Oxygen is the third most abundant atom in the cosmos—behind hydrogen and helium. Because helium is a very inert (nonreactive) atom, water is sometimes described as an interaction between the two most abundant "reactive" atoms in the cosmos. Unlike hydrogen, the origins of the oxygen atom are rooted in dying stars rather than in the Big Bang. As stars near the end of their stellar life, they begin to cool and to switch from hydrogen to helium as a source for nuclear fusion. The cooling stars then enter a phase where they become increasingly dense as intense gravitational energy compresses them into an extremely unstable state that may explode during the final stages of compression (2). This explosion or *supernova* releases the outer layer of the star, which contains many common heavier atoms (e.g., oxygen, carbon), into space as interstellar clouds. Dust grains comprising these clouds are composed of both silicate (oxygen-rich) and carbonaceous (carbon-rich) minerals that are available to react with hydrogen (1).

INTERSTELLAR SPACE

In the interstellar realms of galaxies, water exists predominantly as ice—adhering to the tiny particles that comprise the ubiquitous dust and gas clouds. Water is the primary molecular ice attached to these particles, although methane, carbon monoxide, and water–ammonia mixtures may also be present depending on physical conditions in the gas clouds (3). Water ice in the 10 K temperatures and vacuum conditions of interstellar space is what physical chemists refer to as *amorphous ice* or *glassy water*; it is relatively unstructured compared to the highly crystalline ices that are formed at higher temperatures (e.g., those characteristic of the earth's surface and atmosphere). Amorphous ice is so unstructured that it can flow, not unlike a viscous version of liquid water (4). Some astrophysicists posit that the simple organic molecules responsible for biological life may have been created in this strange ice (4). As interstellar temperatures rise above 150 K (as often occurs near stars), amorphous ice irreversibly transits to more familiar crystalline ice. The various phases and molecular structures of water as a function of temperature are shown in Table 1.

Although water's component atoms are plentiful in interstellar dust and gas clouds, creation of molecular water requires either converting O and OH species directly to water ice on the surfaces of dust grains or producing water vapor via heat energy—usually in the form of stellar radiation (5). The latter process requires that water vapor adhere to dust grains, where the newly formed water molecule is protected from the same ionizing radiation that created it. Scientists currently believe that stars facilitate the creation of water vapor and also that water vapor assists in the birthing of stars. Stars are being born and are dying on an ongoing basis, such that star birthing regions (e.g., the Orion Cloud Complex of the Milky Way Galaxy) generate up to 20% of a galaxy's luminosity as gas and dust clouds are gravitationally compressed into newborn stars. Recent data indicate that these cloud complexes contain an extremely high concentration of water vapor, which has been estimated of the order of 1 part in 2000 (6). The superabundance of water in stellar nurseries (about 20 times greater than that in similar interstellar clouds)

Table 1. Representative Temperature Ranges and Cosmic Locations for the Three Phases of Water

Location	Temperature, K[a]	Phase	Comments
Stellar surfaces	4000 to 50,000	None	Water molecules do not exist. Only hydrogen and oxygen atoms or plasma.
Stellar/planetoid surfaces	<3500	Vapor	Water exists at the surface of cool stars and in cooler regions of some hot stars.
Some planetoids	273 to 373 (at 1 atmosphere)	Liquid	Few places in the cosmos possess the requisite temperatures and pressures.
Interplanetary/interstellar space	>100	Solid (crystalline)	Crystalline ices possess both a cubic and a more common hexagonal structure.
Interstellar space	10 to 150	Solid (amorphous)	The depths of interstellar space are cold enough to produce glassy water.

[a] These temperatures, which are measured on the Kelvin scale, are only approximate and vary depending on environmental conditions (e.g., pressure, rate of cooling or heating).

may permit the gas and dust to cool sufficiently so that condensation can proceed and stars are eventually formed.

As hot winds (in the form of shock waves) are sent out during stellar birthing, the cloud must be cooled—initially by molecular hydrogen and subsequently by water and other simple gases (7). Water vapor is created during the interstellar cloud shocks as oxygen reacts explosively with hydrogen, causing the water vapor concentration to increase substantially during star birthing. Scientists have theorized one of two eventual fates for water created in star birthing. One is that the intense heat of the fledgling star rapidly dissociates water into its component atoms. The other is that the water is deposited on dust grains that later form the star's planetoids. The origin of earthly water is usually attributed to either this second process or the impact of large comets, which are believed to have been more prevalent during the earth's early history.

STARS

Two of the brightest supergiants in our galaxy, Betelgeuse and Antares, have water in their *photospheres*, which constitutes the visible portion of a star (8). A star's photosphere is where its gases transit from opaque to transparent, permitting us to see the stars that are located closest to the earth. This stellar water is actually present within the star itself, not simply as a component of the surrounding dust and gas cloud from which the star was birthed. Aging supergiant stars release massive amounts of water as they die; however, the exact source and role of this water are not yet known.

In addition to cool stars, water has been discovered in the photosphere of at least one hot or *main sequence* star—the Sun. Water cannot exist on the surface of the Sun, where temperatures of 6000 K dissociate the water molecule into its component hydrogen and oxygen atoms. Water can exist in the dark centers of sunspots, where temperatures are less than 3500 K (9). Sunspots are relatively calm solar regions where strong magnetic fields filter out the energy emanating from the intense interior, rendering them both the coolest and darkest regions of the Sun. Water is a major player in determining a star's *radiative opacity*, which describes the extent to which light escapes from stars into interstellar space (10). In this role, water impedes the outward flow of radiation from stars by absorbing energy within certain wavelengths and, thus, renders the star more opaque than it would otherwise appear.

COMETS AND METEORS

Comets are one of the few interstellar objects that are commonly associated with water, predominantly as ice. Comets are composed primarily of water ice that incorporates many of the other simple molecules in interstellar dust and gas clouds (e.g., carbon monoxide, methane, ammonia). Comets are most easily recognized by their unmistakable tails, which can extend millions to hundreds of millions of kilometers behind the icy body of the comet. The tail consists of dust and ionized particles (mostly water ice) that are always transported away from the Sun by the *solar wind*, which is an ionized stream of particles consisting predominantly of protons and electrons. The ionization of water ices is the primary mechanism influencing the properties of a comet's tail, including the steam jets that release tons of water vapor per second from the comet. It is now believed that these steam jets result from solar-induced changes in ice's molecular geometry (e.g., a transition from the amorphous to crystalline structure). Large comets are generally accepted as a source of planetary water, but controversy surrounds the hypothesis that many small comets hitting the planet's upper atmosphere also contribute significantly to the volume of water on the earth.

The first liquid water in the solar system, it was projected, made its appearance on meteors just 20 million years after our Sun and its debris emerged from the interstellar dust and gas cloud (11). Although liquid water is rarely present on the surface of meteors today, the chemical interaction of water with primitive rocks produced carbonate minerals, suggesting that the chemical processes of water evaporation and condensation were among the earliest in the solar system. Recently, a small meteorite found in southwestern North America contained actual liquid water within its salt crystals, which were believed to have been created from the original interstellar cloud that gave rise to the solar system.

PLANETOIDS

Most planet-sized bodies in our solar system (and probably in others) are now known or suspected to contain water in some form. A number of recent missions have revealed a Martian landscape that almost certainly indicates the large-scale flow of liquid water. The surface features of Mars (e.g., flood plains, river beds, mud deposits) suggest the recent presence of liquid water and also the mineralogy of Martian rocks could have resulted only from aqueous processes. Moreover, it has been hypothesized that Mars may have also once possessed surface oceans. The Jovian moon, Europa, is another of the solar system planetoids that probably contains liquid water located tens of kilometers beneath its icy surface. The liquid water underlying Europa's surface ice is believed to be an ocean containing saltwater that may be similar in composition to the seawater of the earth's oceans. Unlike the earth, the heat required to maintain water in a liquid phase on Europa is believed to originate from an internal source such as volcanic activity rather than from the heat of the Sun.

Data collected from the Infrared Space Observatory indicate the presence of water in the upper atmospheres of our solar system's gas giant planets and on one of Saturn's moons (1). The source of water in these planet's atmospheres is attributed to comets or to water-containing interplanetary dust. Based on recently developed techniques for measuring a suite of stellar characteristics (e.g., orbital velocity, position, brightness), the search for planets has been extended beyond our solar system to other star systems in the galaxy (12). Various planets have been identified orbiting stars in the constellations of Leo, Pegasus, Virgo, and Ursa Major that probably possess surface temperatures ranging from slightly less than 100 °C down to −100 °C. Planets or moons characterized by this temperature range could possess water in solid, gaseous, and liquid phases.

BIBLIOGRAPHY

1. Salama, A. and Kessler, M. (2000). *ISO and Cosmic Water*. European Space Agency Bulletin 104, pp. 31–38.
2. Lewis, J. (1995). *Physics and Chemistry of the Solar System*. Academic, San Diego, pp. 30–31.
3. Tielens, A.G.G.M., Hagen, W., and Greenberg, J.M. (1983). Interstellar ice. *J. Phys. Chem.* **87**: 4220–4229.
4. Blake, D. and Jenniskens, P. (2001). The ice of life. *Sci. Am.* **August**: 44–51.
5. O'Neill, P.T. and Williams, D.A. (1999). Interstellar water and interstellar ice. *Astrophys. Space Sci.* **266**: 539–548.
6. Cowen, R. (1998). Unveiling the hidden universe. *Sci. News* **153**: 328–330.
7. Harwit, M., Neufield, D.A., Melnick, G.J., and Kaufman, M.J. (1998). Thermal water vapor emission from shocked regions in Orion. *Astrophys. J. Lett.* **497**: 105–108.
8. Jennings, D.E. and Sada, P.V. (1998). Water in Betelgeuse and Antares. *Science* **279**: 844–847.
9. Oka, T. (1997). Water on the sun: Molecules everywhere. *Science* **277**: 328–329.
10. Tsuji, T. (1986). Molecules in stars. *Annu. Rev. Astron. Astrophys.* **24**: 89–125.
11. Endress, M., Zinner, E., and Bischoff, A. (1996). Early aqueous activity on primitive meteorite parent bodies. *Nature* **379**: 701–703.
12. Boss, A. (1996). Extrasolar planets. *Phys. Today* **September**: 32–38.

THE WATER CYCLE

NASA—Goddard Space Flight
Center

INTRODUCTION

As seen from space, one of the most unique features of our home planet is the water, in both liquid and frozen forms, that covers approximately 75% of the Earth's surface. Believed to have initially arrived on the surface through the emissions of ancient volcanoes, geologic evidence suggests that large amounts of water have likely flowed on Earth for the past 3.8 billion years, most of its existence. As a vital substance that sets the Earth apart from the rest of the planets in our solar system, water is a necessary ingredient for the development and nourishment of life.

HYDROLOGIC HISTORY

The notion that water is continually circulating from the ocean to the atmosphere to the land and back again to the ocean has interested scholars through most of recorded history. In Book 21 of the lliad, Homer (ca. 810 B.C.) wrote of "the deep-flowing Oceanus, from which flow all rivers and every sea and all springs and deep wells." Thales (ca. 640 B.C.–ca. 546 B.C.) and Plato (ca. 427 B.C.–347 B.C.) also alluded to the water cycle when they wrote that all waters returned by various routes to the sea. But it wasn't until many centuries later that scientific measurements confirmed the existence of a water (or hydrologic) cycle. Seventeenth century French physicists Pierre Perrault (1608–1680) and Edmond Mariotte (1620–1684) separately made crude precipitation measurements in the Seine River basin in France and correlated these measurements with the discharge of the river to demonstrate that quantities of rainfall and snow were adequate to support the river's flow.

WATER, WATER, EVERYWHERE

Water is everywhere on Earth and is the only known substance that can naturally exist as a gas, liquid, and solid within the relatively small range of air temperatures and pressures found at the Earth's surface. In all, the Earth's water content is about 1.39 billion cubic kilometers (331 million cubic miles) and the vast bulk of it, about 96.5%, is in the global oceans. Approximately 1.7% is

---

Table 1. One Estimate of Global Water Distribution. Estimates of Groundwater are Particularly Difficult and Vary Widely Amongst Sources, with the Value in this Table Being Near the High End of the Range. Using the Values in This Table, Groundwater Constitutes Approximately 30% of Fresh Water, Whereas Ice (Including Ice Caps, Glaciers, Permanent Snow, Ground Ice, and Permafrost) Constitute Approximately 70% of Fresh Water. With Other Estimates, Groundwater is Sometimes Listed as 22% and Ice as 78% of Fresh Water

	Volume (1000 km^3)	Percent of Total Water	Percent of Fresh Water
Oceans, Seas, & Bays	1,338,000	96.5	—
Ice caps, Glaciers, & Permanent Snow	24,064	1.74	68.7
Groundwater	23,400	1.7	—
Fresh	(10,530)	(0.76)	30.1
Saline	(12,870)	(0.94)	—
Soil Moisture	16.5	0.001	0.05
Ground Ice & Permafrost	300	0.022	0.86
Lakes	176.4	0.013	—
Fresh	(91.0)	(0.007)	0.26
Saline	(85.4)	(0.006)	—
Atmosphere	12.9	0.001	0.04
Swamp Water	11.47	0.0008	0.03
Rivers	2.12	0.0002	0.006
Biological Water	1.12	0.0001	0.003
Total	**1,385,984**	**100.0**	**100.0**

Source: Gleick, P.H., 1996: Water resources. In *Encyclopedia of Climate and Weather*, ed. by S. H. Schneider, Oxford University Press, New York, vol. 2, pp. 817–823.

stored in the polar icecaps, glaciers, and permanent snow, and another 1.7% is stored in groundwater, lakes, rivers, streams, and soil. Finally, a thousandth of 1% exists as water vapor in the Earth's atmosphere (Table 1).

A MULTI-PHASED JOURNEY

The hydrologic cycle describes the pilgrimage of water as water molecules make their way from the Earth's surface to the atmosphere, and back again. This gigantic system, powered by energy from the sun, is a continuous exchange of moisture between the oceans, the atmosphere, and the land (Fig. 1).

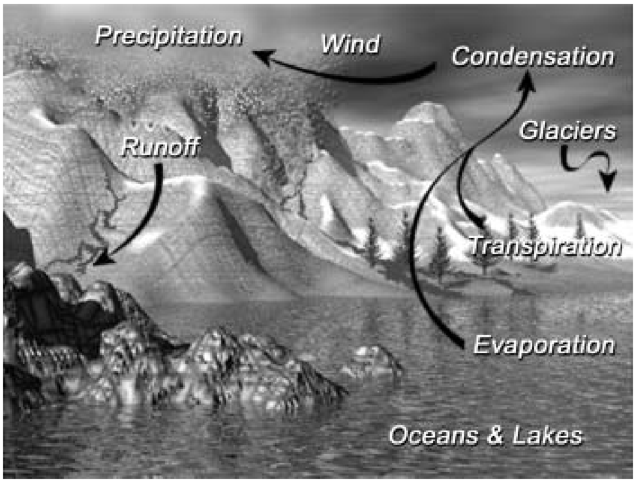

Figure 1. In the hydrologic cycle, individual water molecules travel between the oceans, water vapor in the atmosphere, water and ice on the land, and underground water.

Studies have revealed that the oceans, seas, and other bodies of water (lakes, rivers, streams) provide nearly 90% of the moisture in our atmosphere. Liquid water leaves these sources as a result of evaporation, the process by which water changes from a liquid to a gas. In addition, a very small portion of water vapor enters the atmosphere through sublimation, the process by which water changes from a solid (ice or snow) to a gas. (The gradual shrinking of snow banks, even though the temperature remains below the freezing point, results from sublimation.) The remaining 10% of the moisture found in the atmosphere is released by plants through transpiration (Fig. 2). Plants take in water through their root systems to deliver nutrients to their leaves, then release it through small pores, called stomates, found on the undersides of their leaves. Together, evaporation, sublimation, and transpiration, plus volcanic emissions, account for all the water vapor in the atmosphere. While evaporation from the oceans is the primary vehicle for driving the surface-to-atmosphere portion of the hydrologic cycle, transpiration is also significant. For example, a cornfield 1 acre in size can transpire as much as 4000 gallons of water every day.

After the water enters the lower atmosphere, rising air currents carry it upward, often high into the atmosphere,

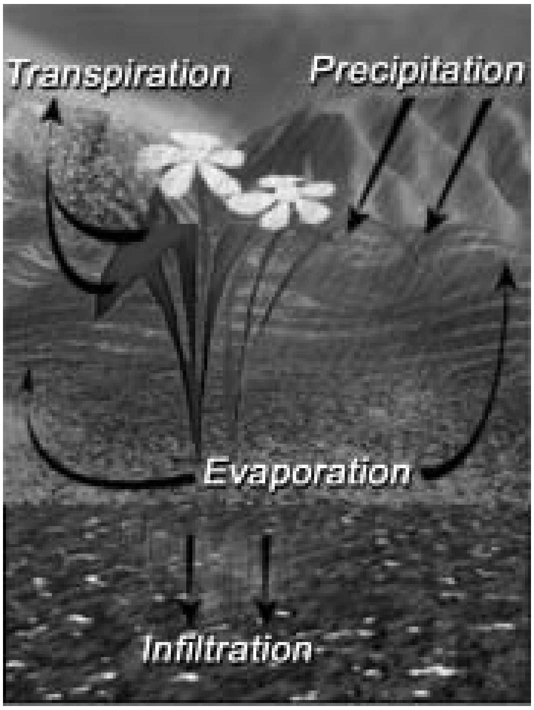

Figure 2. Plants return water to the atmosphere through transpiration. In this process, water evaporates from pores in the plant's leaves, after being drawn, along with nutrients, from the root system through the plant.

where the air cools and loses its capacity to support water vapor. As a result, the excess water vapor condenses (i.e., changes from a gas to a liquid) to form cloud droplets, which can eventually grow and produce precipitation (including rain, snow, sleet, freezing rain, and hail), the primary mechanism for transporting water from the atmosphere back to the Earth's surface.

When precipitation falls over the land surface, it follows various routes. Some of it evaporates, returning to the atmosphere, and some seeps into the ground (as soil moisture or groundwater). Groundwater is found in two layers of the soil, the "zone of aeration," where gaps in the soil are filled with both air and water, and, further down, the "zone of saturation," where the gaps are completely filled with water. The boundary between the two zones is known as the water table, which rises or falls as the amount of groundwater increases or decreases (Fig. 3). The rest of the water runs off into rivers and streams, and almost all of this water eventually flows into the oceans or other bodies of water, where the cycle begins anew (or, more accurately, continues). At different stages of the cycle, some of the water is intercepted by humans or other life forms.

Even though the amount of water in the atmosphere is only 12,900 cubic kilometers (a minute fraction of Earth's total water supply that, if completely rained out, would cover the Earth's surface to a depth of only 2.5 centimeters), some 495,000 cubic kilometers of water are cycled through the atmosphere every year, enough to uniformly cover the Earth's surface to a depth of 97 centimeters. Because water continually evaporates, condenses, and precipitates, with evaporation on a global basis approximately equaling global precipitation, the total amount of water vapor in the atmosphere remains approximately the same over time. However, over the continents, precipitation routinely exceeds evaporation, and conversely, over the oceans, evaporation exceeds precipitation. In the case of the oceans, the routine excess of evaporation over precipitation would eventually leave the oceans empty if they were not being replenished by additional means. Not only are they being replenished, largely through runoff from the land areas, but, over the past 100 years, they have been over-replenished, with sea level around the globe rising by a small amount. Sea level rises both because of warming of the oceans, causing water expansion and thereby a volume increase, and because of a greater mass of water entering the ocean than the amount leaving it through evaporation or other means. A primary cause for increased mass of water entering the ocean is the calving or melting of land ice (ice sheets and glaciers).

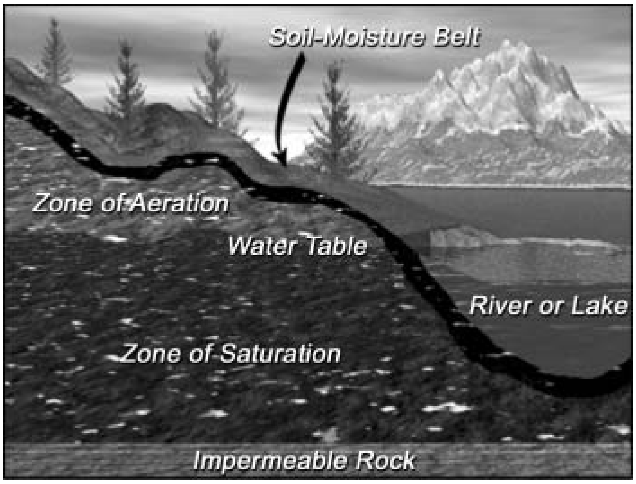

Figure 3. The water table is the top of the zone of saturation and intersects the land surface at lakes and streams. Above the water table lies the zone of aeration and soil moisture belt, which supplies much of the water needed by plants.

Throughout the hydrologic cycle, there are an endless number of paths that a water molecule might follow. Water at the bottom of Lake Superior may eventually fall as rain in Massachusetts. Runoff from the Massachusetts rain may drain into the Atlantic Ocean and circulate northeastward toward Iceland, destined to become part of a floe of sea ice, or, after evaporation to the atmosphere and precipitation as snow, part of a glacier. Water molecules can take an immense variety of routes and branching trails that lead them again and again through the three phases of ice, liquid water, and water vapor. For instance, the water molecules that once fell 100 years ago as rain on your great grandparents' farmhouse in Iowa might now be falling as snow on your driveway in California.

THE WATER CYCLE AND CLIMATE CHANGE

Amongst the highest priorities in Earth science and environmental policy issues confronting society are the potential changes in the Earth's water cycle due to climate change. The science community now generally agrees that the Earth's climate will undergo changes in response to natural variability, including solar variability, and to increasing concentrations of greenhouse gases and aerosols. Furthermore, agreement is widespread that these changes may profoundly affect atmospheric water vapor concentrations, clouds, and precipitation patterns. For example, a warmer climate, directly leading to increased evaporation, may well accelerate the hydrologic cycle, resulting in an increase in the amount of moisture circulating through the atmosphere. Many uncertainties remain, however, as illustrated by the inconsistent results given by current climate models regarding the future distribution of precipitation.

THE AQUA MISSION AND THE WATER CYCLE

As mentioned earlier, the hydrologic cycle involves evaporation, transpiration, condensation, precipitation, and runoff. NASA's Aqua satellite monitors many aspects of the role of water in the Earth's systems, and will do so at spatial and temporal scales appropriate to foster a more detailed understanding of each of the processes that contribute to the hydrologic cycle. These data and the analyses of them nurture the development and refinement of hydrologic process models and a corresponding improvement in regional and global climate models, with a direct anticipated benefit of more-accurate weather and climate forecasts.

Aqua's contributions to monitoring water in the Earth's environment involves all six of Aqua's instruments: the Atmospheric Infrared Sounder (AIRS), the Advanced Microwave Sounding Unit (AMSU), the Humidity Sounder for Brazil (HSB), the Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E), the Moderate Resolution Imaging Spectroradiometer (MODIS), and Clouds and the Earth's Radiant Energy System (CERES). The AIRS/AMSU/HSB combination provides more-accurate space-based measurements of atmospheric temperature and water vapor than have ever been obtained before, with the highest vertical resolution to date as well. Since water vapor is the Earth's primary greenhouse gas and contributes significantly to uncertainties in projections of future global warming, it is critical to understand how it varies in the Earth system.

The water in clouds is examined with MODIS, CERES, and AIRS data; and global precipitation is monitored with AMSR-E. The cloud data includes the height and areal coverages of clouds, the liquid water content, and the sizes of cloud droplets and ice particles, the latter sizes being important to the understanding of the optical properties of clouds and their contribution to the Earth's albedo (reflectivity). HSB and AMSR-E, both making measurements at microwave wavelengths, have the ability to see through clouds and detect the rainfall under them, furthering the understanding of how water is cycled through the atmosphere.

Frozen water in the oceans, in the form of sea ice, is examined with both AMSR-E and MODIS data, the former allowing routine monitoring of sea ice at a coarse resolution and the latter providing greater spatial resolution but only under cloud-free conditions. Sea ice can insulate the underlying liquid water against heat loss to the often frigid overlying polar atmosphere and also reflects sunlight that would otherwise be available to warm the ocean. AMSR-E measurements allow the routine derivation of sea ice concentrations in both polar regions, through taking advantage of the marked contrast in microwave emissions of sea ice and liquid water. This continues with improved resolution and accuracy, a 22-year satellite record of changes in the extent of polar ice. MODIS, with its finer resolution, permits the identification of individual ice floes, when unobscured by clouds.

AMSR-E and MODIS also provide monitoring of snow coverage over land, another key indicator of climate change. Here too, the AMSR-E allows routine monitoring of the snow, irrespective of cloud cover, but at a coarse spatial resolution, while MODIS obtains data with much greater spatial detail under cloud-free conditions.

As for liquid water on land, AMSR-E provides an indication of soil moisture, which is crucial for the maintenance of land vegetation, including agricultural crops. AMSR-E's monitoring of soil moisture globally should permit, for example, the early identification of signs of drought episodes.

THE AQUA SPACECRAFT

Aqua is a major mission of the Earth Observing System (EOS), an international program centered in NASA's Earth Science Enterprise to study the Earth in detail from the unique vantage point of space. Focused on key measurements identified by a consensus of U.S. and international scientists, EOS is further enabling studies of the complex interactions amongst the Earth's land, ocean, air, ice and biological systems.

The Aqua spacecraft circles the Earth in an orbit that ascends across the equator each day at 1:30 p.m. local time and passes very close to the poles, complementing the 10:30 a.m. measurements being made by Terra, the first of the EOS spacecraft, launched in December 1999. The instrument complement on Aqua is designed to provide information on a great many processes and components of the Earth system, including cloud formation, precipitation, water vapor, air temperature, cloud radiative properties, sea surface temperature, surface wind speeds, sea ice concentration and temperature, snow cover, soil moisture, and land and ocean vegetation. The individual swaths of measurements will be compiled into global images, with global coverage of many variables being obtained as frequently as every two days or, with the help of numerical models, combined every 6 or 12 hours into comprehensive representations of the Earth's atmospheric circulation and surface properties. In combination with measurements from other polar orbiting satellites, Aqua measurements also provide accurate monthly-mean climate assessments that can be compared with and assimilated into computer model simulations of the Earth's climate.

The Earth Observing System has three major components: the EOS spacecraft, an advanced ground-based computer network for processing, storing, and distributing the collected data (the EOS Data and Information System); and teams of scientists and applications specialists who study the data and help users in universities, industry, and the public apply it to issues ranging from weather forecasting and climate prediction to agriculture and urban planning.

CYCLONES

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

Cyclones are hazardous weather conditions characterized by extreme gusts of wind moving in a circular pattern, low pressure, and intense rain. They form over tropical or subtropical waters; however, some can reach land, where they have in the past inflicted great amounts of damage on buildings and communities.

Based on wind strength, cyclones are given various names. Tropical depressions have maximum surface winds of less than 39 mph. Those cyclones whose maximum winds are between 39 and 74 mph are known as tropical storms. Upon attaining 74-mph surface winds, a cyclone is known as one of a number of names that are all regional equivalents of the same type of storm.

The names of cyclones are typhoon (in the NW Pacific Ocean), hurricane (in the North Atlantic, NE Pacific, and South Pacific Ocean), tropical cyclone (in the SW Indian

Ocean), severe cyclonic storm (in the North Indian Ocean), and severe tropical cyclone (in the SW Pacific and SE Indian Oceans). Cyclones always build over tropical seas. There are seven basins in which the majority of tropical cyclones form: the Atlantic Basin, the North Indian Basin, the Southwest Indian Basin, the Southeast Indian-Australian Basin, the Australian-Southwest Pacific Basin, the Northwest Pacific Basin, and the Northeast Pacific Basin. Heat gives cyclones their energy, and thus the water over which a cyclone forms must be at least as warm as 80°F (i.e., tropical). Other conditions necessary for a cyclone include a fast cooling atmosphere, moisture in the middle troposphere area, a distance of more than 300 miles from the equator to be influenced by the Coriolis force, and minimal vertical wind shear. This vertical wind shear is the result of differences between winds in the lower and upper portions of the atmosphere. However, the main contributor to the formation of a cyclone is a disturbance in the form of a thunderstorm or group of showers. When all of these factors come together, conditions are right for a tropical cyclone, although the presence of each of these factors does not guarantee the formation of a cyclone. Cyclones are spontaneous; a minute variation in one variable can be the difference between a hurricane and a thunderstorm.

The circular area in the center of a cyclone, known as the "eye," has conditions quite different from those in the region surrounding it. Calmness and a light breeze characterize the eye. Temperatures are generally higher than those in the surrounding area, and the sky is usually very clear.

Since the 1970s, cyclones have been regularly named by various people and agencies, depending on where the cyclones originated. They are named to ensure ease of description among newscasters, weatherpeople, and the general public. Names are often taken from rotating lists, one name for each letter of the alphabet. When an exceptional cyclone occurs, its name is taken out of use ("retired") to avoid confusion.

Strong cyclones often cause extensive damage to whatever they encounter. Damage can range from crop destruction to total devastation of building structures, depending on the severity of the cyclone. They become most dangerous as they hit land and spawn tornadoes. These are formed when tropical cyclones begin to lose their power. The diameter of a tropical cyclone is measured in kilometers; the diameter of a tornado is measured in meters. In 1980, one of the most destructive cyclone-spawned tornadoes in the United States caused nearly $100,000,000 worth of damage to the Austin, Texas, area. Along with destruction, cyclones (and the subsequent tornadoes) cause death. Large objects lifted from the path of the extreme wind are tossed about like lethal weapons. In 1964, twenty-two people were killed by a tornado that hit the Los Angeles area in the United States. The most damaging cyclone in history was Hurricane Andrew, which caused over $26 billion in damage to the Southeastern United States. The most deadly cyclone ever may have been the Bangladesh Cyclone, which killed at least 300,000 people in 1970. One of the ways in which a hurricane will cause damage is through the storm surge. This is the name given to the phenomenon whereby sea level in the area of a cyclone rises due to cyclonic winds.

Once a cyclone reaches 74 mph maximum winds and is considered a hurricane, it's intensity is rated by the Saffir–Simpson Scale, used since the 1970s by the National Oceanic and Atmospheric Administration (NOAA). Hurricanes are referred to as category one, category two, category three, category four, or category five. Category one hurricanes have winds between 74 and 95 mph. They cause negligible damage to buildings and other structures, although they can damage mobile homes and road signs. They carry a storm surge of about 5 feet. Category two hurricanes have maximum winds of 96–110 mph. They can cause minor damage to buildings and sizable damage to mobile homes and trees. Their storm surge is usually around 7 feet. Category three hurricanes have winds from 111–130 mph. They cause noticeable damage to buildings and trees; buildings near the shoreline are often destroyed by flooding. Evacuation of certain low-lying areas can be necessary for category three hurricanes, and their storm surge ranges from 9–12 feet. Category four hurricanes have winds between 131 and 155 mph. These hurricanes can cause massive damage to smaller structures. They destroy trees and mobile homes utterly. Evacuation of large areas can be necessary, as the storm surge can reach up to 18 feet. Category five hurricanes are the most powerful; they have maximum sustained winds of more than 155 mph. Residential as well as industrial structures are often destroyed. Damage is catastrophic, and major evacuations usually take place. The storm surge exceeds 18 feet above normal.

Cyclones can range from relatively simple tropical storms to devastating hurricanes whose winds swirl at furious speeds. Their effects can be disastrous and long-lasting.

READING LIST

Bureau of Meteorology. (2002). *Surviving Cyclones.* Available: http://www.bom.gov.au/info/cyclone/. (March 13).

Department of Public Safety. (2002). "*General Hurricane Information.*" Available: http://www.escambia-emergency.com/geninfo.asp. (March 13).

Landsea, Christopher W. *FAQ: Hurricanes, Typhoons, and Tropical Cyclones.* Available: http://www.aoml.noaa.gov/hrd/tcfaq/tcfaqA.html. (March 13).

National Weather Service Houston/Galveston. (2002). *Tropical Cyclone FAQ.* Available: http://www.srh.noaa.gov/hgx/index3/tropical3/geninfo3.htm. (March 13).

Hurricane Research Division: FAQ Available: http://www.aoml.noaa.gov/hrd/tcfaq/F1.html.

Hurricane Research Division: FAQ Available: http://www.aoml.noaa.gov/hrd/tcfaq/E7.html.

Hurricane Research Division: FAQ Available: http://www.aoml.noaa.gov/hrd/tcfaq/B1.html.

Hurricane Research Division: FAQ Available: http://www.aoml.noaa.gov/hrd/tcfaq/A10.html.

Hurricane Research Division: FAQ Available: http://www.aoml.noaa.gov/hrd/tcfaq/A1.html.

Wonders of Weather. Available: http://school.discovery.com/lessonplans/programs/wondersofweather/vocab.html.

The Saffir–Simpson Hurricane Scale. Available: http://www.nhc.noaa.gov/aboutsshs.shtml.

WATER CYCLE

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

The water cycle—a series of steps through which virtually all water on earth is constantly cycling. The names of these steps vary, but they generally follow this order: storage, evaporation, transpiration, condensation, precipitation, infiltration, and runoff.

Due to the nature of the water cycle, there is really no "beginning" or "end", but for simplicity, the cycle can be said to begin at the storage stage. Storage is the storing of water on, in, and above the earth. There are about 340 million cubic miles of water on the earth. This water is stored in many places: glaciers, seas, rivers, lakes, polar ice caps, and all living things on the planet. The largest percentage of water, approximately 97.25%, exists as salt water in the ocean. Water is also stored in the earth's atmosphere and as groundwater. Groundwater in the top few miles of the earth's crust is easily obtainable, whereas groundwater further down is chemically attached to rocks and usually cannot be accessed.

Evaporation is an essential way in which water is transferred from the earth to the atmosphere. During evaporation, water molecules change from a liquid state directly into a gaseous state, known as water vapor. Changes in temperature and air pressure on the earth cause millions of gallons of water to evaporate into the atmosphere each day. The majority of the water that is transferred into the atmosphere is primarily in a liquid state before the change; however, a small percentage of new water vapor comes from ice. The process of water changing state from a solid directly to a gas is called sublimation. Sublimation occurs on glaciers and polar ice sheets. When the sun strikes these massive ice structures, the temperature and pressure allow no room for a liquid, only a solid and a gas.

Transpiration is another way that water is transferred into the atmosphere. Although similar in process to evaporation, transpiration differs because the medium from which the water comes is not earth, it is plants. Transpiration is the process through which water absorbed from the ground by plant roots evaporates into the atmosphere. After the plants absorb the water from the soil, the water moves up through the veins of the plant, eventually reaching the leaves (or the plant's equivalent). These leaves have pores, or stomata, which allow water to evaporate. Large forests can release massive amounts of water into the atmosphere through transpiration.

Once the water has been transferred into the atmosphere by either evaporation or transpiration, the molecules will eventually condense. This process is called condensation. Condensation is the reciprocal process of evaporation. During condensation, water changes state from a gas to a liquid. Once again, a change in either temperature or pressure initiates the change. Condensation is most visible in the formation of dew, fog, and clouds. Dew usually appears in the morning and is water that has condensed on any solid object during the rapid temperature change that occurs from day to night. If the water vapor is dense enough around objects when the temperature change happens, dew is formed. Fog is really a low level cloud formation. When the conditions are right, fog is formed on or near the ground. The only difference between fog and cloud formations is their locations with reference to the earth. Clouds form when water vapor in the atmosphere encounters a critical temperature or air pressure, causing the water vapor to change state into a liquid. Water molecules then "stick together" causing physical cloud structures to emerge. When the molecules in the cloud that are stuck together reach a certain critical mass, they release the moisture causing precipitation to begin.

Precipitation is the process by which water is transferred from the medium of the atmosphere to the medium of the earth. Precipitation literally comes in all shapes and sizes, including rain, snow, and ice. The differentiating factor that determines the form water takes when precipitating is air temperature. The lower the temperature, the more likely it is that the precipitation will be either snow or ice (the closer the temperature is to the freezing point of water, 32 °F or 0 °C the more likely). Droplets of rain form around small particles of dust or dirt to form a cohesive raindrop as a vehicle for the trip to the earth.

Regardless of the form it takes on its descent, water will take one of two paths once it arrives on the earth's crust. It will either seep into the ground or flow into larger bodies of water such as streams, seas, oceans, and rivers. The former process is infiltration, and the latter is runoff. Infiltration restores groundwater to the water table that has been lost from wells. During the infiltration process, the water is purified by cascading amongst rocks and minerals that draw impurities out of the water.

Water that does not infiltrate the earth's crust runs off into streams, rivers, seas, and oceans. Runoff replenishes bodies of water that have lost water due to evaporation. Runoff can occur either above the land or below it. Large amounts of water can run across the land immediately after precipitation. Smaller amounts of water run through the ground and along the water table until they reach a larger body of water.

So thusly does the water complete its cycle throughout the different media of the earth. This process is essential for the survival of life on the earth because it allows water to permeate every accessible region of the crust and atmosphere, thereby allowing life to flourish wherever water exists.

READING LIST

Hydrologic Cycle. (2002). Available: http://www.und.nodak.edu/instruct/eng/fkarner/pages/cycle.htm. (April 1).

The Water Cycle: Scientific Concepts. (1995–1998). Available: http://mbgnet.mobot.org/fresh/cycle/concepts.htm. (March 28, 2002).

Gedzelman, S.D. (2001). The Water Cycle. In: *Microsoft Encarta Encyclopedia Deluxe 2001*.

DEGREE DAY METHOD

Carlos D. Messina
University of Florida
Gainesville, Florida

An adequate description and prediction of plant development is critical for understanding plant growth and its responses to the environment. Only then we can assess the impact of a changing environment on plant productivity and survival, understand plant adaptation, design adequate agricultural production systems, and optimize natural resources management. Of particular importance for water resource management are the effects of plant ontogeny on plant water use and productivity throughout the regulation of leaf area. This determines the patterns of plant water demand and the partitioning between transpiration and evaporation components. Therefore, plant development dictates the selection of cultivars that best fit the water availability patterns under dryland conditions, and the amount, frequency and timing for irrigation otherwise.

Plant development is responsive to environmental cues such as photoperiod, water availability, and temperature. Figure 1a shows an example of the close association between ambient temperature and plant development for boll growth duration in cotton. Rameur first suggested in 1735 that the duration of any developmental stage was related to temperature and that this duration can be characterized by a thermometric constant, which can be predicted by the sum of daily air temperatures (2). Since then, several methods based on the concept of normalizing time by temperature to predict developmental rates were developed (2,3) and applied in life sciences. Although we can formulate some hypothesis to explain the relationship between plant development and temperature based on the effects of temperature on circadian oscillations (4), lipid composition and membranes fluidity, and cell division (5), the scientific basis of these empirical methods remains elusive. It was shown, however, that degree day based methods significantly improved the description and prediction of phenological events relative to other approaches based on the time of the year or the number of days (3).

Because of the close coupling between the timescale of plants and temperature, Ritchie and NeSmith (6) proposed that the most appropriate term to describe plant development would be "thermal time" or degree days (DD), whose units are °C-day. The degree day method is based on the effects of temperature on developmental rates, rather than on the duration of a phase. Figure 1b shows the linear relationship between temperature and rate of development, the reciprocal of the duration of the phase in days. The base temperature (T_b), temperature at which the rate of development is zero, is obtained by extrapolating this linear relationship to the intersection with the x-axis:

$$T_\mathrm{b} = -\frac{y_0}{\alpha}$$

where y_0 is the intercept and α is the slope of the linear regression equation (Fig. 1b). Then, the rate of development (RD) can be calculated as

$$\mathrm{RD}(t) = \alpha\ (T_\mathrm{a} - T_\mathrm{b})$$

where T_a is the mean temperature. Integrating the rate of development in time,

$$\int \mathrm{RD}\, dt = \int \alpha\ (T_\mathrm{a} - T_\mathrm{b})\, dt$$

and considering a daily time step ($\Delta t = 1$) for integration, we can estimate the cumulative development (CD) at time n (t_n), provided that $t_0 = 0$, as

$$\mathrm{CD} = \alpha \sum_{n} (T_\mathrm{a} - T_\mathrm{b})$$

For constant temperatures, given that development is complete when CD equals one, the degree days above a certain base temperature for a given phase are

$$\mathrm{DD} = (T_\mathrm{a} - T_\mathrm{b})\, n$$

or

$$\mathrm{DD} = \frac{1}{\alpha}$$

For the example of cotton shown in Fig. 1, boll growth requires 1000 °C-days above a base temperature

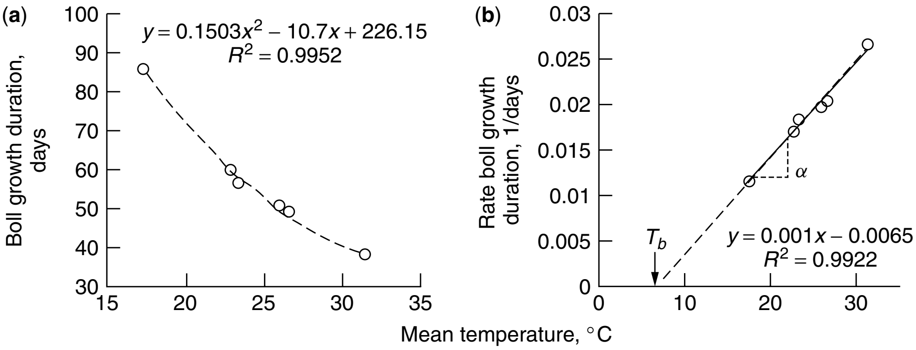

Figure 1. Responses of boll growth duration in cotton (*Gossypium hirsutum*, L.) development to air temperature. (**a**) Boll growth duration, (**b**) rate of boll growth duration (1/d). T_b denotes base temperature (see below), and α is the slope of the regression line. Data is from (1).

of 6.5 °C to reach maturity. Although this formulation for calculating DD is useful for estimating the requirements of the plant to complete a given developmental phase, in natural environments, temperature fluctuates and makes this procedure inadequate. The following canonical form to calculate DD is frequently used:

$$\mathrm{DD} = \sum_{n} \left[\left(\frac{T_{\max} + T_{\min}}{2} \right) - T_{\mathrm{b}} \right]$$

where $T_{\max}$ and $T_{\min}$ are maximum and minimum daily temperatures, n is the number of days of temperature observations, and $[(T_{\max} + T_{\min})/2] = T_{\mathrm{b}}$ when $[(T_{\max} + T_{\min})/2] < T_{\mathrm{b}}$. An alternative method, frequently used in calculating DD for corn, compares $T_{\max}$ and $T_{\min}$ with T_{b} individually before calculating their average. Significant discrepancies in calculated DD between these procedures may arise under some circumstances (3).

These procedures for calculating *DD* are adequate, provided that (1) the daily temperature does not fall below T_{b} and does not exceed an upper threshold for a significant part of the day, (2) the temperature of the growing plant tissue is the same as the air temperature, and (3) the response of the rate of development to temperature is linear over the range of temperature that the crop experiences (5). When these assumptions are not satisfied, alternative approaches are required. For example, McMaster et al. (5) and Vinocur and Ritichie (7) used soil or apex temperatures to calculate DD because the growing points were close to the soil rather than at the height where the temperature is normally measured. In simulation models such as CERES (8), hourly temperatures are approximated from $T_{\max}$ and $T_{\min}$, and if the temperature is below T_{b}, DD is set to zero for that part of the day. For a thorough discussion of the limitations of the degree day method, see Ritchie and NeSmith (5).

The degree day method can be found in the literature in a different mathematical form under the concept of physiological day. This model is generally used to simulate development in legumes and is particularly adequate for incorporating the effects of photoperiod on development (9). The rate of development is calculated as the product of two functions:

$$\mathrm{RD}(t) = f(P) \times f(T)$$

One function accounts for the effects of photoperiod (P) and the other for the effects of temperature (T) on development. The model predicts relative development; the maximum rate of development is standardized to 1.0. At optimum temperature (T_{opt}) and photoperiod ($P < P_{\min}$ in short day species and $P > P_{\max}$ in long day species), the rate of progress in calendar days equals the rate of progress in physiological days. When conditions deviate from the optimum, the rate of progress per day decreases and becomes a fraction of a physiological day. Then, cumulative development is measured in photothermal time units. The effects of the photoperiod on the rate of development are nonlinear:

$$\begin{aligned}
f(P) &= 0, && \text{if } P > P_{\max} \\
f(P) &= 1 - \frac{P - P_{\min}}{P_{\max} - P_{\min}}, && \text{if } P_{\min} < P \le P_{\max} \\
f(P) &= 1, && \text{if } P \le P_{\min}
\end{aligned}$$

In short day species, the relative rate of development is maximum below a threshold $P_{\min}$, above which the development decreases to zero at $P_{\max}$. Both $P_{\min}$ and $P_{\max}$ are characteristic of the species and cultivar (Table 1). A similar function is used to describe the effects of temperature:

$$\begin{aligned}
f(T) &= 0, && \text{if } T \le T_{\mathrm{b}} \\
f(T) &= \frac{T - T_{\mathrm{b}}}{T_{\mathrm{opt1}} - T_{\mathrm{b}}}, && \text{if } T_{\mathrm{b}} < T < T_{\mathrm{opt1}} \\
f(T) &= 1, && \text{if } T_{\mathrm{opt2}} \ge T \ge T_{\mathrm{opt1}} \\
f(T) &= \mathbf{1} - \frac{T - T_{\mathrm{opt2}}}{T_{\mathrm{upper}} - T_{\mathrm{opt2}}}, && \text{if } T_{\mathrm{upper}} > T > T_{\mathrm{opt2}} \\
f(T) &= 0, && \text{if } T \ge T_{\mathrm{upper}}
\end{aligned}$$

where T_{opt1} and T_{opt2} define a plateau whose rate of development is maximum. For temperatures above T_{opt2} and below T_{opt1}, the rate of development decreases linearly to zero at T_{upper} and T_{b}, respectively. Table 1 shows characteristic values for selected cereals and legumes.

In this chapter, a derivation of the degree day method is shown along with a mathematical variant that incorporates the effects of the photoperiod. The literature about models that uses degree days is vast, it is not

Table 1. Cardinal Temperatures and Photoperiods for Selected Models of Cereals and Legumes[a]

	Temperature, °C				Photoperiod, h	
	T_{b}	T_{opt1}	T_{opt2}	T_{upper}	$P_{\min}$	$P_{\max}$
Legumes						
Soybean						
Vegetative	7	28	35	45		
Early reproductive	6	26	30	45	11.7–14.6	15.5–21.0
Late reproductive	−48	26	34	45		
Bean						
Vegetative	4	27	35	45		—
Early reproductive	5	22	35	45	12.2–13.2	—
Late reproductive	0	18	35	45		
Peanut[b]						
Vegetative	11	28	30	55		
Early reproductive	11	28	28	55	—	—
Late reproductive	5	26	26	55		
Cereals						
Wheat & Barley	0	15	—	—	20	—
Millet	10	36	—	—	12	125–150[c]
Rice	9	—	—	—	11.7–12.8	35–189[c]
Maize	8	34	—	—	12.5	0.3–0.8[d]
Sorghum	8	34	—	—	12.5–13.6	30–90[c]

[a] Adapted from Reference 10.

[b] insensitive to photoperiod.

[c] DD (°C-day) per hour increase in photoperiod.

[d] Units are in days per hour increase in photoperiod (11).

the intent to review it here. However, the rationale and concept presented here covers most of the models currently used. Summerfield et al. (12) discuss additive photothermal models and Jones et al. (13) provide further details on multiplicative ones. Ritchie and NeSmith (6) and McMaster and Wilhelm (3) discuss some limitations of the degree day method. As emphasized in these previous papers, to obtain reliable and accurate predictions or descriptions of plant development, the parameters and the method used for their estimation must be consistent.

BIBLIOGRAPHY

1. Reddy, K.R., Davidonis, G.H., Johnson, A.S., and Vinyard, B.T. (1999). Temperature regime and carbon dioxide enrichment alter cotton roll development and fiber properties. *Agron. J.* **91**: 851–858.
2. Wang, J.Y. (1960). A critique of the heat unit approach to plant-response studies. *Ecology* **41**: 785–790.
3. McMaster, G.S. and Wilhelm, W.W. (1997). Growing degree-days: one equation, two interpretations. *Agric. Forest Meteorol.* **87**: 291–300.
4. Michael, T.P., Salome', P.A., and McClung, C. (2003). Two *arabidopsis* circadian oscillators can be distinguished by differential temperature sensitivity. *Proc. Nat. Acad. Sci. USA* **100**: 6878–6883.
5. McMaster, G.S. et al. (2003). Spring wheat leaf appearance and temperature: extending the paradigm? *Ann. Bot.* **91**: 697–705.
6. Ritchie, J.T. and NeSmith, D.S. (1991). Temperature and crop development. In: *Modeling Plant and Soil Systems*. R.J. Hanks and J.T. Ritchie (Eds.). *Agronomy* 31. American Society of Agronomy, Madison, WI, pp. 5–29.
7. Vinocur, M.G. and Ritchie, J.T. (2001). Maize leaf development biases caused by air-apex temperature differences. *Agron. J.* **93**: 767–772.
8. Ritchie, J.T., Singh, U., Godwin, D.C., and Bowen, W.T. (1998). Cereal growth, development and yield. In: *Understanding Options for Agricultural Production*. G.Y. Tsuji, G. Hoogenboom, and P.K. Thornton (Eds.). Kluwer Academic, Dordrecht, the Netherlands, pp. 79–98.
9. Boote, K.J., Jones, J.W., and Hoogenboom, G. (1998). Simulation of crop growth: CROPGRO model. In: *Agricultural Systems Modeling and Simulation*. R.M. Peart and R.B. Curry (Eds.). Marcel Dekker, New York, pp. 651–693.
10. Jones, J.W., Tsuji, G.Y., Hoogenboom, G., Hunt, L.A., Thornton, P.K., Wilkens, P.W., Imamura, D.T., Bowen, W.T., and Singh, U. (1998). Decision support system for agrotechnology transfer. In: *Understanding Options for Agricultural Production*. G.Y. Tsuji, G. Hoogenboom, and P.K. Thornton (Eds.). Kluwer Academic, Dordrecht, the Netherlands, pp. 157–177.
11. Kiniry, J.R. (1991). Maize phasic development. In: *Modeling Plant and Soil Systems*. R.J. Hanks and J.T. Ritchie (Eds.). *Agronomy* 31. American Society of Agronomy, Madison, WI, pp. 55–69.
12. Summerfield, R.J. et al. (1993). Towards the reliable prediction of time to flowering in 6 annual crops .2. *soybean* (*Glycine max*). *Exp. Agric.* **29**: 253–289.
13. Jones, J.W., Boote, K.J., Jagtap, S.S., and Mishoe, J.W. (1991). Soybean development. In: *Modeling Plant and Soil Systems*. R.J. Hanks and J.T. Ritchie (Eds.). *Agronomy* 31. American Society of Agronomy, Madison, WI, pp. 71–90.

DESERTIFICATION

Aldo Conti
Frascati (RM), Italy

Desertification is a degradation of the top layer of the soil that reduces its ability to support plant life and to produce food. As a result, the soil becomes dusty and dry, and it is easily carried away by erosion, which affects wild and domestic animals, wild plants and crops and, finally, humans. Desertification can be the result of human activities or of natural climate changes. In the latter case, the process is very slow and can take several thousand years to produce its effects. A common misconception about desertification is that it spreads from a desert core. The truth is that land degradation can occur where land abuse has become excessive. If it is not stopped in time, desertification spreads from that spot. Eventually, many of these spots merge and form a large homogeneous area. Another misconception is that desertification is the result of droughts. In fact, well-managed land can recover from even a long period of drought with very little adverse effect as soon as rains return. But it is true that droughts can increase the pace of desertification already taking place.

Desertification is a term that has been in use since at least 1949, when Aubreville, a perceptive and well-informed botanist and ecologist, published a book on *Climate, Forets, et Desertification de l'Afrique Tropicale*. Aubreville observed desertification in tropical Africa and understood immediately that the culprit was not the Sahara desert gaining land. He noticed instead that the main reasons behind desertification were tree cutting, indiscriminate use of fire to clear the land, and cultivation.

Many processes can lead to desertification. Logging, for instance, makes the soil unstable on mountain slopes. Eventually, all the soil runs down as a dangerous landslide or mud river and leaves behind exposed rocks, unable to support any life. What is left behind are barren mountains, particularly evident in China. In Europe, one of the leading causes of desertification is overgrazing. In the past, wild animals used to move following the rainfall, always grazing the richest areas. In modern times, the use of fences prevented these movements, and the result was heavy overgrazing that left the soil exposed to erosion. In many areas, desertification is the result of agriculture. A typical example of this is salinization of soil that happens normally when the soil is overirrigated. The water that is not used by plants evaporates, leaving behind salts that concentrate in the soil. Eventually, the concentration of salts in the soil becomes so high that plants cannot survive, again exposing the land to erosion. In same cases, desertification is the more direct result of urbanization, mining, and recreational activities. In any case, the adverse effects are still the same.

Nowadays, desertification is a serious problem that affects, according to some estimates, up to 30% of dry lands. Worldwide, desertification is making approximately 12 million hectares useless for cultivation every year. But land degradation and desertification are by no means new problems, despite the attention focused on them in recent years. During the first conquest of Africa, it was normal to clear patches of land with fire and then use them to grow crops. After three or four seasons, the land was depleted of nutrients and unable to support any plant life. Moreover, there is some historical evidence that serious and extensive land deterioration was already occurring several centuries before. It started in arid regions, and it had three epicenters: the Mediterranean Sea and other places where destructive changes in soil and plant cover had occurred, but were small in extent or not well known.

Luckily, desertification in many areas has been stopped, but very little effort has been made to restore the land to its original productivity. Today, desertification can be defeated using techniques already known, if financial resources are available and the political will to act is present. For instance, only in the last few decades, satellite images have allowed a better understanding and monitoring of the problem on a large scale.

There are several possible remedies available, even at the local level. The first is to avoid cutting trees or, at least, to replace them with new ones. Plants, a major soil stabilizer, can alone stop erosion. Moreover, the use of available local water and ways to control the salinity of the soil can be very effective. On this topic, genetic engineering is trying to help. Scientists are working on the development of crops that can survive higher salinity, both as a way to use the land and produce food and as a way to save lost soils. Curiously, one of the remedies until now used against desertification is to pollute the soil. In Iran, oil is sprayed over semiarid land with crops. The oil covers the seedlings, retains the moisture, and prevents them from being blown away by the wind until they grow large enough.

As stated at the beginning, desertification is sometimes a slow natural process. As an example, it should be enough to say that 20,000 years ago, the Sahara Desert was a lush forest. This is proved by the fossils of the animals that used to live in this forest. Moreover, pictures taken using radar from the Space Shuttle allow us to identify numerous dry riverbeds under a few meters of sand. This is already happening again due to the global warming on the earth. Many deserts are expanding, even though it is not possible to find a human reason for that. Desertification seems to threaten, in particular, all countries in the Mediterranean region. The coasts of North Africa are already disappearing into the sand, but the same might start happening in Europe. The heat wave of 2003 caused lots of problems. In Italy, many crops were heavily affected, and in the whole Mediterranean region, but in particular in Spain and Portugal, large areas of forest were destroyed by fires.

DEW

Aldo Conti
Frascati (RM), Italy

Dew forms when water condenses on objects on the surface near the ground and forms a thin layer or many droplets. Dew forms normally during the night when the air temperature decreases and approaches the dew point. Objects on the surface cool down, too, by radiative cooling, facilitating the condensation of small droplets of water. Dew formation is helped by the high humidity of the bottom layer of air, close to the surface. This layer can supply the needed water and prevent the evaporation of the dew already deposited. Strong winds can inhibit the formation of dew. Turbulence mixes a larger layer of air and creates a more homogeneous distribution of humidity and heat, thus preventing the formation of the right conditions near the ground. Dew forms more easily on surfaces that cool efficiently, like metals, which is the reason that cars are often covered with dew in the morning. But dew is often seen on grass or plants because their transpiration creates a thin layer of very humid air.

Dew is very important in the ecology of many deserts, especially those along the western coasts of Africa and South America. In some of them, water from condensation of dew caused by night cooling often exceeds that of rainfall. But dew can be an interesting source of water for human consumption, too. For instance, data recorded in the Negev desert of Israel, a country where water is really scarce, have shown that dew falls for 200 days each year. Moreover, dew is plentiful on many little islands that, surrounded by water, have high humidity but where no water dwells. Some islands of the Mediterranean Sea suffer from a chronic water shortage, which is the reason that one of the first dew-collecting plants has been built in Ajaccio (Corsica Island in France). It is also interesting to describe an old project studied in India and unfortunately never actually realized. The idea was to pump water at 4.5 °C from the sea at a depth of 500 meters. The pumping scheme called for the use of four 1.2-meter pipes and wind-powered pumps. A heat exchanger of 130,000 square meters could then condense, every day, more than 600 cubic meters of dew.

But many animals and plants have already learned how to survive on dew. A place particularly interesting for this kind of adaptation is the Namib desert in Namibia (Africa). Here, many insects in the early morning sit on top of sand dunes trying to catch some dew (or fog) on their wings and legs. In particular, the beetle *Stenocara* has even modified the surface of its hard shell surface to trap mist more efficiently. This beetle shell is covered with bumps whose peaks are smooth like glass and attract water. The troughs around the bumps are covered with a wax that repels water. The water is therefore collected by the peaks, and when a droplet is big enough to touch the water repelling valley it rolls down to the animal's mouth. Some researchers think that the easy trick of the animal can be used in dew collecting devices. A similar surface can be prepared in many ways, even with an ink-jet printer that sprays hydrophilic ink onto an acetate sheet. Dew-collecting devices have the problem of making the water run off the collecting surface.

Another Namibian creature well adapted to survive on dew is the plant *Welwitschia mirabilis*. Despite its look, it is a close relative of pine trees. It has a short trunk split in two and from each side grows a single leaf that can be several meters long. Welwitschias are among the oldest plants on the earth, and some of them are 1500 years old. The trick that allows this plant to survive in one of the driest deserts of the planet is a two-part root system. Welwitschias have a long root that goes deep in the soil trying to reach the water table and extends for several meters away from the plant. These roots collect all the dew that forms on the top layer of soil. Moreover, all the water that condenses on the leaves forms big drops that eventually run down, so that the plant can water its own roots.

DEW DESERTS

Giora J. Kidron
The Hebrew University of
Jerusalem
Jerusalem, Israel

The term dew deserts denotes deserts that receive dew precipitation throughout the year. Dew is defined as condensation of atmospheric vapor and does not include vapor condensation that stems from the wet ground, i.e., distillation. We term dew desert as a desert that receives at least 10 mm of dew in >0.1 mm of daily precipitation. As a result of dew precipitation, primary production in dew deserts is much higher, affecting the entire food chain.

Like the Atacama Desert in South America, the Namib Desert in South Africa, and the Sonora Desert in Baja California (North America), dew only affects those portions of the desert close to a sea or ocean, i.e., to a large body of water that can serve as an adequate vapor source. As most deserts occupy mainly the interior of the continents, the term applies only to a small part of near sea deserts.

At present, information regarding the extent of dew precipitation within deserts is scarce. Most of the available information pertains to the heart of the Negev Desert where annual precipitation is 90–100 mm (1). There, a total of 33 mm was annually measured (2). Nevertheless, based on rain isohyets and the proximity to a sea or ocean (3), one may hypothesize that parts of the Sahara and the Arabian Peninsula are in fact dew deserts. Certain areas within the fog deserts may also be dew deserts (4–6).

Although daily dew amounts may usually range between 0.1 and 0.3 mm and the total annual sum may be much smaller than that of rain precipitation, the occurrence of a steady and constant source may be of great importance in arid and semiarid zones (7–9). As the threshold for organism activity was found to be 0.03 mm of dew (10), amounts of 0.1–0.3 mm are still sufficient to allow growth activity.

Although small, the dew amounts received may allow the establishment and distribution of many species, otherwise not capable of inhabiting the site. Consequently, higher biomass and higher species diversity may characterize the dew deserts. Biomass increase per millimeter of dew may be much higher than that calculated for rain (7) and that extrapolated, by adding the addition of water from dew to that of rain. The dew may thus be an important source of moisture for the primary food chain in arid and semiarid zones (11).

Many groups of micro-organisms were hypothesized to use dew. Among them, cyanobacteria, hypolithic algae and cyanobacteria that occupy the underside of stones (12,13), and lichens (14,15). Among the lichens are the endolithic lichens, embedded within calcite or dolomite crystals at the upper 1–5 mm of the rock and stone surface (16); epilithic and epedaphic lichens that dwell on the rock and soil surface, respectively (17); foliose (18); and fruticose lichens (18,19). In addition, bugs, ants, and beetles were reported to use dew directly by drinking it (20), whereas isopods and beetles were reported to use it indirectly by consuming wet food (20). When wet, snails in the Negev Desert were reported to successfully graze on endolithic lichens. By doing so, the snails disintegrate the limestone, a weathering phenomenon that can take place only upon rock moistening (21).

Although cyanobacteria, hypolithic algae, and cyanobacteria, epedaphic, endolithic, epilithic, foliose, and fruticose lichens were all hypothesized to use dew, positive evidence for dew use under field conditions was obtained only from endolithic, epilithic, foliose, and fruticose lichens. These lichens were shown to photosynthesize for 2–4 h in the morning following dew in the Negev Desert (Fig. 1) (10,18,22). One should note that whereas lithobiontic lichens (inhabiting rocks and stones) were shown to use dew (22), dew use by micro-organisms inhabiting the soil is still controversial. Whereas Lange et al. (22) and Veste et al. (23) showed the use of dew by epedaphic lichens inhabiting soil and sand, respectively, and indirect evidence, expressed by the development of sexual and vegetative reproduction organs, were also monitored in mosses in the Negev Desert (24), the use of dew by endedaphic cyanobacteria inhabiting sand is not certain. According to Jacobs et al. (25), dew may moisten the surface. However, other reports indicate that dew moistening of the soil is rare (24,26). According to Bunnenberg and Kühn (26), an amount of 0.13 mm of dew at 9 cm above ground amounted to 0.03 mm only at the surface. Similarly, Kidron et al. (24) showed that average dew precipitation of 0.1 mm measured on glass plates at 0.7 cm above ground amounted to only 0.034 mm at the surface. As cyanobacteria need liquid water for growth (27), and the necessary threshold for net photosynthesis was 0.1 mm (28), only rarely sufficient dew will moisten the surface to reach >0.1 mm, and thus the contribution of dew to the growth and development of endedaphic cyanobacteria may be negligible.

The uncertainty regarding the use of dew by certain micro-organisms and at certain habitats is mainly linked to two factors: (a) Does dew condense at all habitats? (b) Is the amount supplied sufficient for use and for net carbon gain? As for the second question, rapid dew evaporation during the morning may result in a net loss of carbon, as was shown for the lichen *Ramalina maciformis* following mornings of low dew precipitation (10).

Whereas the efficiency to which different micro-organisms may use dew may be species-dependent, no controversy exists as to the perquisite conditions necessary, i.e., the capability of the dew to condense at their habitat. This ability depends, of course, on the existence of sufficient moisture in the air. It also depends

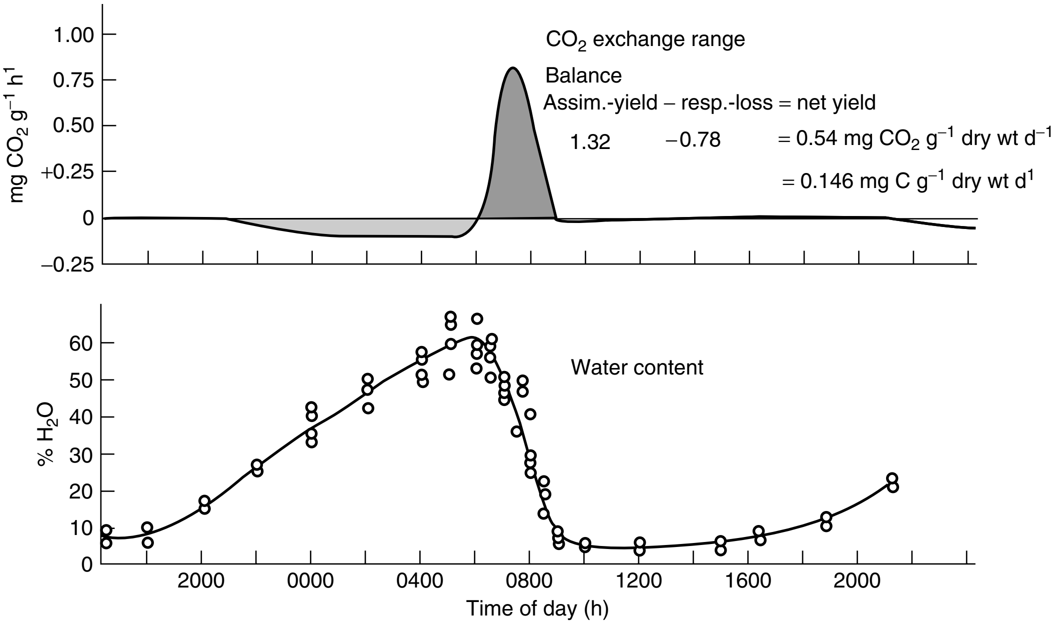

Figure 1. Apparent CO_2 exchange rate (above) and the thalli water content (below) of *Ramalina maciformis* during 24 h in September 1967. Net carbon yield resulted in 0.146 mg C per gram of dry weight a day (18).

on substrate temperatures that will dictate whether the dew point temperature, necessary for vapor condensation, will be reached.

Once reached, dew condensation will take place. As dew was found to condense at a more or less constant rate (29), the time during which the dew point is reached may dictate, to a large extent, the overall amount of dew that will be condensed. Furthermore, as phototrophic microorganisms necessitate light hours for photosynthesis, the length of time during which dew is available may dictate, to a large extent, the overall net gain of organic carbon owing to dew use. Thus, variables that may affect dew condensation and dew duration should be considered once dew availability in different habitats is examined. Factors such as topographical elevation, height above ground, aspect, location along the slope, and angle may all affect dew condensation.

Recent dew measurements carried in the Negev Desert aimed to examine the role of the above-mentioned factors. Thus, in order to obtain continuous dew amounts and duration, a simple and inexpensive method was adapted. The Cloth-Plate Method (CPM) consists of 10 × 10 × 0.2 glass plates glued at their bottom to 10 × 10 × 0.5 cm plywood plates, thus creating an identical substratum (30). Velvet-like cloth (6 × 6 × 0.15 cm) are attached each afternoon to the center of the glass plate and collected throughout the following morning into glass jars that are immediately sealed and then weighed in a nearby lab, oven dried (in 70 °C) and then weighed again, and their moisture content is calculated. By placing plates next to each other within a certain habitat, the CPM facilitates inexpensive large-scale continuous measurements (Fig. 2). The readings are also not affected by wind, as might be the case with some other devices (31).

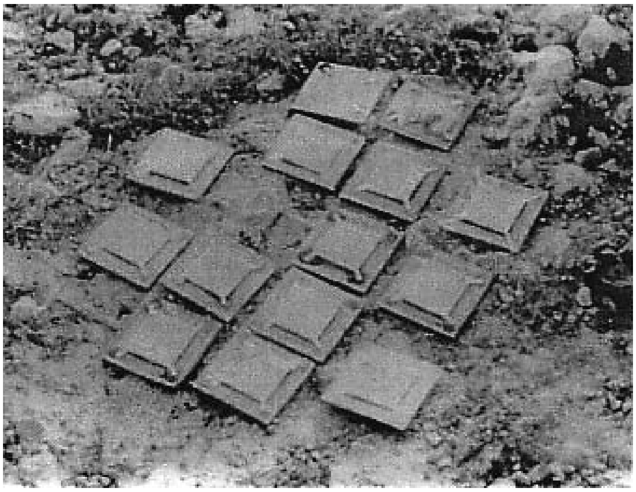

Figure 2. Dew measurements by the CPM in the Negev Highlands.

The CPM was used to assess the possible role of altitude and distance from the sea. When simultaneous measurements at three locations at 250, 550, and 1000 m above sea level being 37, 55, and 98 km from the Mediterranean Sea were carried out in the Negev Desert, positive correlations between dew and fog amounts and altitude were found (Fig. 3). Dew precipitation, as well as fog precipitation, increased with altitude, with the most elevated location receiving more than twice the amount obtained at the topographically lowest location, which was the case, although the most elevated location was also the farthest away from the Mediterranean Sea, i.e., from a vapor source. The data thus indicate that, within 100 km from the sea, topographical height may compensate for the increase in distance from the vapor source. In this regard, one should note that elevation plays a major role in controlling dew and fog precipitation also in the Atacama Desert (4).

The CPM facilitates careful analysis of dew precipitation. When continuous dew measurements were carried out, a typical condensation and evaporation pattern was obtained during the morning hours (Fig. 4). Dew condensation was found to continue also after dawn and sunrise, explained by radiation-induced air turbulence (30). As for the relationship between dew amounts and height above ground, dew values showed high variability (33,34). When dew was measured on a hilltop in the Negev Highlands at 0.7, 10, 20, 30, 40, and 50 cm above ground, maximal values were recorded at 10 cm above ground (Fig. 5). The findings reflected two effects: the warming effect of the soil as a result of a nocturnal heat flux that rises from the deeper horizons of the soil, responsible for the decrease in dew values near the ground (31,34,35), and the air turbulence at height, resulting from higher wind velocities, responsible for lower dew values farther away from the ground (35,36).

High variability in dew amounts and duration was also found in the Negev Desert when dew was measured at 0.7 and 40 cm above ground along limestone slopes (approximately 50 m long) of four aspects (north, south, east, and west) within a second-order drainage basin (Fig. 6). Whereas at 0.7 cm above ground, the hilltops and the bottom parts of the northern and western aspects received the highest amounts, the wadi beds received approximately half these values (Fig. 6a). The lowest values were received at the south-facing midslope (located at the lee side of the prevailing north north-west winds), being approximately a quarter of the maximal values recorded. At 40 cm above ground, a reverse trend was obtained with the south-facing midslope and the wadi beds receiving the highest values (Fig. 7). And thus, whereas the 0.7 cm above-ground results were not in accordance with the classical model that predicts high dew values at the wadi beds (because of nocturnal katabatic wind) and higher amounts at the lee side of the wind, i.e., at the south-facing slope (because of undisturbed inversion), the 40 cm above ground measurements corresponded to the classical model. The discrepancy was explained by the overwhelming impact of the rock surface temperatures on the dew values at 0.7 cm above ground, with south-facing rock surfaces being 3–8 degrees higher than the north-facing rock surfaces throughout the night. As for the higher amounts of dew received at 0.7 cm above ground at the hilltops in comparison with the wadi

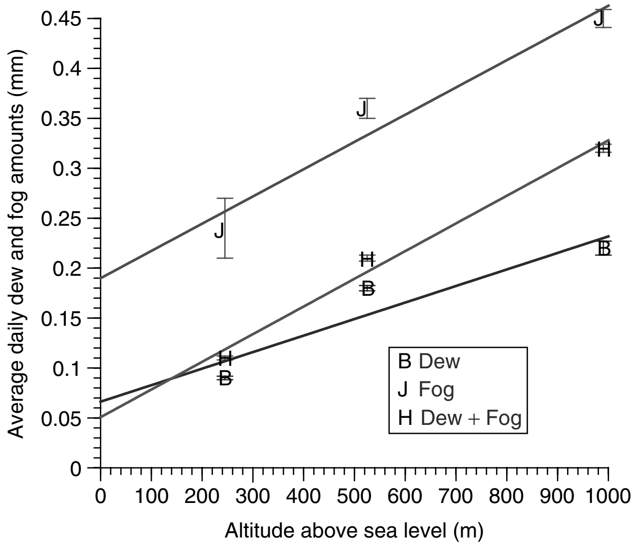

Figure 3. The relationships between dew and fog precipitation with altitude in the Negev Desert (32).

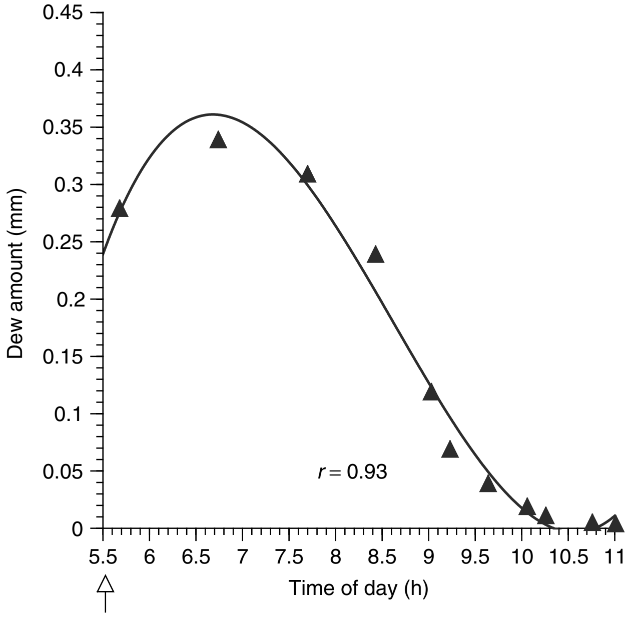

Figure 4. A typical condensation and evaporation pattern during the morning hours (30).

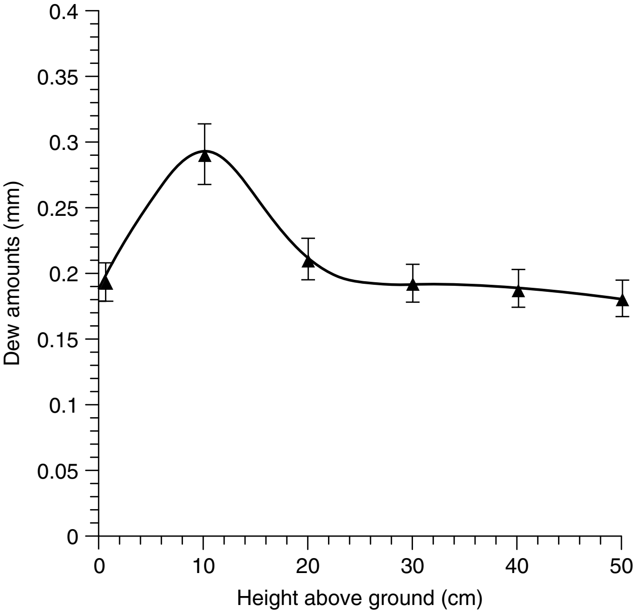

Figure 5. Dew amounts with height above ground (30).

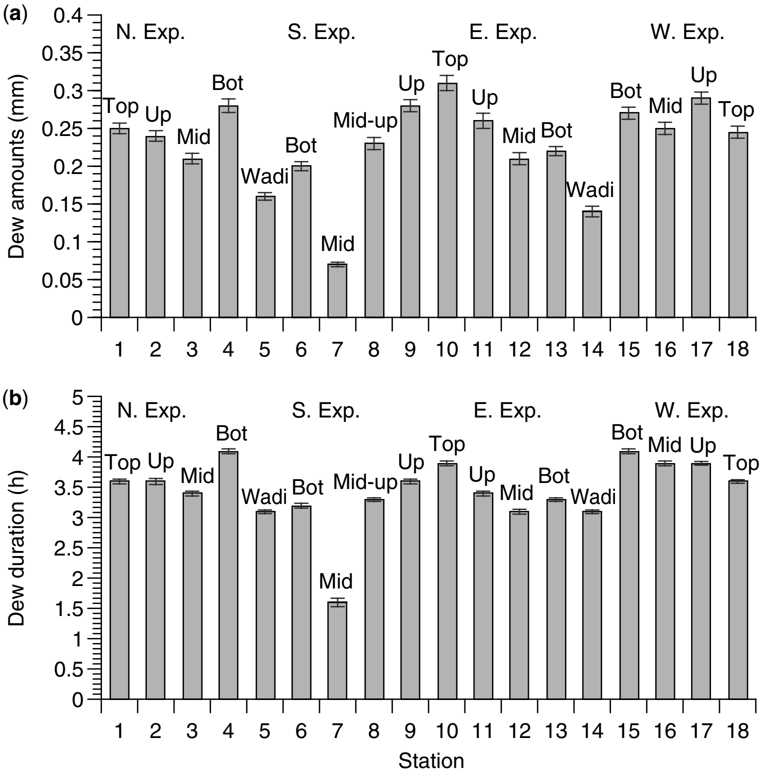

Figure 6. Dew amounts (**a**) and duration (**b**) within 18 stations located at 4 aspects of a second-order drainage basin in the Negev Desert. Top = hilltop, Up = Upper slope, Mid = Midslope, Bot = Bottom slope (37).

beds, it was hypothesized that the afternoon winds act as a cooling agent, facilitating a much earlier drop in surface temperatures (and thus reaching much faster the necessary dew point temperature) at the wind-exposed hilltops (37).

Aspect and slope location also dictate dew duration, seen as being of even greater ecological importance to organisms than the maximal amount (38–40). Generally, a positive correlation between dew amounts and duration was found (34,37), and thus, by affecting dew amounts, aspect and slope position also affect dew duration (Fig. 6b). Nevertheless, dew duration was also affected by slope location and aspect beyond these relations. Although condensation was found to take place in all habitats also after sunrise (for usually 0.5–1.0 h after sunrise), condensation at the sun-sheltered habitats of the bottom north- and west-facing slopes continued for up to 1.5 h following sunrise (41). As a result of the higher maximal values and the limited desiccation effect of the sun in these habitats, these habitats were characterized by longer dew duration.

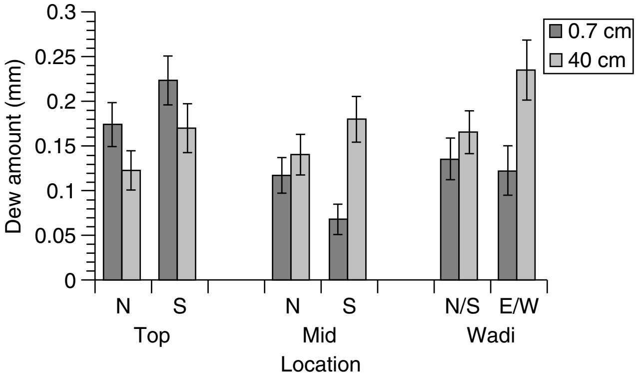

Figure 7. Dew measurements at 0.7 and 40 cm above ground at three habitats within a second-order drainage basin in the Negev Desert. N = northern exposure, S = southern exposure, Wadi N/S = wadi between the northern and southern exposure, Wadi E/W = wadi between the eastern and western exposures (37).

The substrate angle is another factor found to affect dew amounts and duration. When cloths were attached to 50 × 50 × 10 cm wooden boxes, having sides of different angles (30, 45, 60, 75, and 90°) placed at different aspects (facing north, south, west, and east) on a hilltop, a decrease in dew amounts with an increase in angle, from 30° to 90°, was found and dew amounts were positively correlated with $\cos(\theta)$ (Fig. 8). Thus, dew amounts obtained at an angle of 90° was approximately a quarter of the values obtained at a horizontal surface and at an angle of 30°, both of which received similar values. This difference was explained by the slower rate of nocturnal cooling that takes place with an increase in angle in accordance with the lower proportion of sky seen by the substrate (43). No preferential condensation in accordance with aspect was found. Nevertheless, dew duration was aspect-dependent with a decrease in duration following the order west>north>south>east (Fig. 9). Thus, daylight dew duration was approximately double at the west-facing sun-sheltered angle than at the sun-exposed east-facing aspect.

The variability in dew amounts and duration may have important consequences for micro-organisms. For

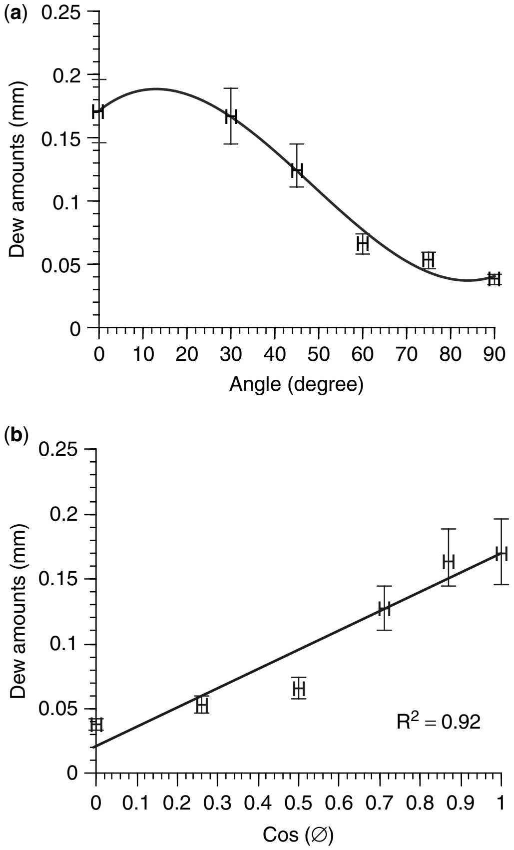

Figure 8. The relationships between dew amounts with (**a**) angle (in degrees) and (**b**) cos(θ) as measured in the Negev Desert (42).

most micro-organisms, the ability to use dew is primarily a function of the amount condensed and the duration during which dew is available, both of which are habitat-dependent. Thus, when cloths were attached to different substrates, high variability in dew amounts was received on loose cobbles, partially embedded cobbles, and rock surfaces (Fig. 10). Loose cobbles that rapidly cool at night were found to receive approximately twice the amounts of embedded cobbles and more than 4 times the amounts condensed on a nearby rock (44). Whereas the cobbles were primarily inhabited by endolithic lichens, the rock surfaces were mainly inhabited by epilithic lichens. Similarly, lichen growth on loose cobbles was controlled by angle-induced dew. Thus, the top of loose cobbles, inhabited by lichens, received approximately twice the amount condensed on the uninhabited side of the cobble (45). Consequently, dew contribution to the ecosystem biomass may be highly important. For instance, in research conducted by Kappen et al. (22,46) in the Negev Desert, the biomass of lichens (with *Ramalina maciformis* predominating) in the northern aspect was over 200 g m^{-2}, three to five times as much than at the other aspects. Although some of the differences can certainly be attributed also to use of rain-induced moisture, the fact that the main water source for *Ramalina maciformis* and other endolithic and epilithic lichens in the Negev Desert was dew pointed to the importance of dew on lichen biomass. The difference in lichen cover of the sun-sheltered and the sun-exposed slopes points also to the role of dew in lichen distribution (47). As lichens may serve as food for snails, the whole food chain may be affected by dew precipitation.

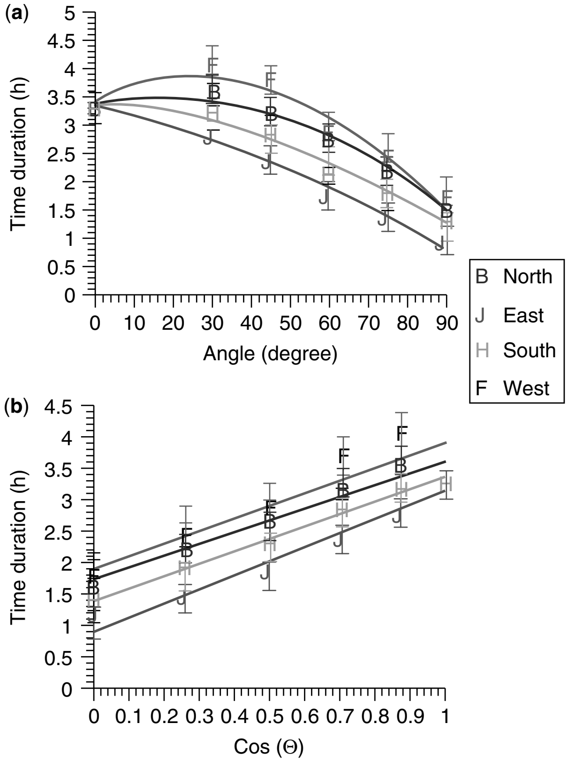

Figure 9. The relationships between dew duration with (**a**) angle (in degrees) and (**b**) cos(θ) in the Negev Desert (42).

Dew precipitation may also affect vascular plants. The effect may not always be positive, as dew may hasten plant fungal infection (7,48). Dew may preferentially accumulate on leaves because of their lower temperature (0.5–2 °C lower in comparison to the ambient air) (35,36) and thus, wet the leaves for several hours during the morning (34).

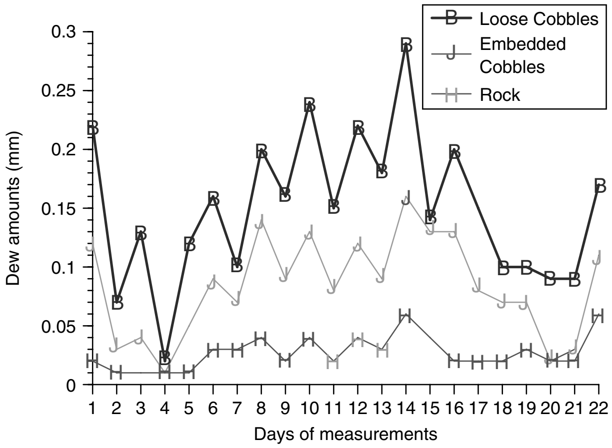

Figure 10. Simultaneous dew amounts as obtained on loose and embedded cobbles inhabited by endolithic lichens, and rock surface inhabited by epilithic lichens (44).

However, no conclusive evidence of direct dew use by vascular plants is yet available, and the research conducted on this topic is controversial (7,8,31,49,50). Dew, however, may have an indirect effect, facilitating stomata opening (51), longer hours of photosynthesis (39), reducing transpiration (7), and affecting flowering (7). Dew may facilitate seedling survival (52), recovery from wilting (34), plant growth (34), and yield (34,53). By affecting micro-organisms and plant growth, and by facilitating insect survival (20), dew also indirectly affects the nutrient cycling (54) and soil-forming processes (55), which in turn has an important impact on the ecosystem and may thus call for a specific awareness of the dew deserts.

BIBLIOGRAPHY

1. Rosenan, N. and Gilad, M. (1985). Meteorological data. In: *Atlas of Israel*. Carta, Jerusalem.
2. Evenari, M. (1981). Ecology of the Negev Desert, a critical review of our knowledge. In: *Developments in Arid Zone Ecology and Environmental Quality*. H. Shuval (Ed.). Balaban ISS, Philadelphia, PA, pp. 1–33.
3. Reitan, C.H. and Green, C.R. (1968). Appraisal of research on weather and climate of desert environments. In: *Deserts of the World, An Appraisal of Research Into Their Physical and Biological Environments*. W.G. McGinnies, B.J. Goldman, and P. Paylore (Eds.). The University of Arizona Press, Tucson, AZ, pp. 19–92.
4. Redon, J. and Lange, O.L. (1983). Epiphytic lichens in the region of a Chilean "fog oasis" (Fray Jorge). 1. Distributional patterns and habitat conditions. *Flora.* **174**: 213–243.
5. Lange, O.L., Meyer, A., Ullmann, I., and Zellner, H. (1991). Microclimate conditions, water content and photosynthesis of lichens in the coastal fog zone of the Namib Desert: measurements in the fall. *Flora.* **185**: 233–266.
6. Lange, O.L., Meyer, A., Zellner, H., and Heber, U. (1994). Photosynthesis and water relations of lichen soil crusts: field measurements in the coastal fog zone of the Namib Desert. *Functional Ecol.* **8**: 253–264.
7. Wallin, G.R. (1967). Agrometeorological aspects of dew. *Agric. Meteorol.* **4**: 85–102.
8. Evenari, M., Shanan, L., and Tadmor, N. (1971). *The Negev, the Challenge of a Desert*. Harvard Univ. Press, Cambridge, MA, p. 345.
9. Noy-Meir, I. (1973). Desert ecosystems: environment and producers. *Ann. Rev. Ecol. Syst.* **4**: 25–51.
10. Kappen, L. etal. (1979). Ecophysiological investigations on lichens of the Negev Desert, IV: Annual course of the photosynthetic production of *Ramalina maciformis* (Del.) Bory. *Flora.* **168**: 85–108.
11. Shachak, M. and Steinberger, Y. (1980). An algae-desert snail food chain: energy flow and soil turnover. *Oecologia.* **46**: 402–411.
12. Berner, T. (1974). *The Ecophysiology of the Hypolithic Algae in the Negev Highlands of Israel*. Ph.D. Thesis, The Hebrew University of Jerusalem.
13. Berner, T. and Evenari, E. (1978). The influence of temperature and light penetration on the abundance of the hypolithic algae in the Negev Desert of Israel. *Oecologia.* **33**: 255–260.
14. Friedmann, E.I., Lipkin, Y., and Ocampo-Paus, R. (1967). Desert algae of the Negev. *Phycologia.* **6**: 185–195.
15. Friedmann, E.I. and Galun, M. (1974). Desert algae, lichens and fungi. In: *Desert Biology II*. G.W. Brown (Ed.). Academic Press, New York, pp. 165–212.
16. Fry, E.I. (1922). Some types of endolithic lichens. *Ann. Bot.* **35**: 541–562.
17. Golubic, S., Freiedmann, I., and Schneider, J. (1981). The lithobiontic ecological niche, with special reference to microorganisms. *J. Sedimentary Petrol.* **51**: 475–478.
18. Lange, O.L., Schulze, E-D., and Koch, W. (1970). Ecophysiological investigations on lichens of the Negev Desert, II: CO_2 gas exchange and water conservation of *Ramalina maciformis* (Del.) Bory in its natural habitat during the summer dry period (Technical translation 1655. National Research Council of Canada). *Flora.* **159**: 38–62.
19. Lange, O.L. (1969). Ecophysiological investigations on lichens of the Negev Desert. I. CO_2 gas exchange of *Ramalina maciformis* (Del.) Bory under controlled conditions in the laboratory (Technical translation 1654. National Research Council of Canada). *Flora.* **158**: 324–359.
20. Broza, M. (1979). Dew, fog and hygroscopic food as a source of water for desert arthropods. *J. Arid. Environ.* **2**: 43–49.
21. Shachak, M., Jones, C.G., and Granot, Y. (1987). Herbivory in rocks and the weathering of a desert. *Science* **236**: 1098–1099.
22. Lange, O.L., Schulze, E-D., and Koch, W. (1970). Ecophysiological investigations on lichens of the Negev Desert, III: CO_2 gas exchange and water metabolism of crustose and foliose lichens in their natural habitat during the summer dry period. *Flora.* **159**: 525–538.
23. Veste, M., Littmann, T., and Friedrich, H. (2001). Microclimate boundary conditions for activity of soil lichen crusts in sand dunes of the north-western Negev Desert, Israel. *Flora.* **196**: 465–474.
24. Kidron, G.J., Hernstadt, I., and Barzilay, E. (2002). The role of dew as a moisture source for sand microbiotic crusts in the Negev Desert, Israel. *J. Arid. Environ.* **52**: 517–533.
25. Jacobs, A.F.G., Heusinkveld, B.G., and Berkowicz, S.M. (2000). Dew measurements along a longitudinal sand dune transect, Negev Desert, Israel. *Int. J. Biometeorol.* **43**: 184–190.
26. Bunnenberg, C. and Kühn, W. (1977). Application of the β-absorption method to measure dew on soil and plant surfaces. *Int. J. Appl. Radiation Isotopes* **28**: 751–754.
27. Lange, O.L., Kilian, E., and Ziegler, H. (1986). Water vapor uptake and photosynthesis of lichens: performance differences in species with green and blue-green algae as phycobionts. *Oecologia* **71**: 104–110.
28. Lange, O.L. et al. (1992). Taxonomic composition and photosynthetic characteristics of the "biological soil crusts" covering sand dunes in the Western Negev Desert. *Func. Ecol.* **6**: 519–527.
29. Zangvil, A. (1996). Six years of dew observation in the Negev Desert, Israel. *J. Arid. Environ.* **32**: 361–372.
30. Kidron, G.J. (1998). A simple weighing method for dew and fog measurements. *Weather* **53**: 428–433.
31. Noffsinger, T.L. (Ed.). (1965). Survey of techniques for measuring dew. In: *Humidity and Moisture, Measurement and Control in Science and Industry*. A. Wexler (Ed.). Reinhold Publishing Corporation, New York, pp. 523–531.
32. Kidron, G.J. (1999). Altitude dependent dew and fog in the Negev desert, Israel. *Agric. Forest Meteorol.* **96**: 1–8.
33. Lloyd, M.G. (1961). The contribution of dew to the summer water budget of Northern Idaho. *Bull. Am. Meteorol. Soc.* **42**: 572–580.
34. Duvdevani, S. (1964). Dew in Israel and its effect on plants. *Soil Sci.* **98**: 14–21.
35. Long, I.F. (1958). Some observations on dew. *Meteorol. Mag.* **87**: 161–168.

36. Monteith, J.L. (1957). Dew. *Quart. J. Royal Meteorol. Soc.* **83**: 322–341.
37. Kidron, G.J., Yair, A., and Danin, A. (2000). Dew variability within a small arid drainage basin in the Negev highlands, Israel. *Quart. J. Royal Meteorol. Soc.* **126**: 63–80.
38. Lange, O.L., Geiger, I.L., and Schulze, E-D. (1977). Ecophysiological investigations on lichens in the Negev Desert. V. A model to simulate net photosynthesis and respiration of *Ramalina macifomis*. *Oecologia.* **28**: 247–259.
39. Kappen, L. et al. (1980). Ecophysiological investigations on lichens of the Negev Desert, VII: The influence of the habitat exposure on dew imbibition and photosynthetic productivity. *Flora.* **169**: 216–229.
40. Lange, O.L. and Tenhunen, J.D. (1982). Water relations and photosynthesis of desert lichens. *J. Hattori. Bot. Lab.* **53**: 309–313.
41. Kidron, G.J. (2000). Analysis of dew precipitation in three habitats within a small arid drainage basin, Negev Highlands, Israel. *Atmos. Res.* **55**: 257–270.
42. Kidron, G.J. (in press). Angle and aspect dependent dew precipitation in the Negev Desert, Israel. *J. Hydrol.*
43. Oke, T.R. (1978). *Boundary Layer Climates*. John Wiley & Sons, New York, p. 372.
44. Kidron, G.J. (2000). Dew moisture regime of endolithic and epilithic lichens inhabiting calcareous cobbles and rock outcrops, Negev Desert, Israel. *Flora.* **195**: 145–153.
45. Kidron, G.J. (2002). Causes of two patterns of lichen colonization on cobbles in the Negev Desert, Israel. *Lichenologist.* **34**: 71–80.
46. Kappen, L., Lange, O.L., Schulze, E-D., Evenari, M., and Buschbom, U. (1975). Primary production of lower plants (lichens) in the desert and its physiological basis. In: *Photosynthesis and Productivity in Different Environments*. J.P. Cooper (Ed.). Cambridge University Press, Cambridge, UK, pp. 133–143.
47. Danin, A. and Garty, J. (1983). Distribution of cyanobacteria and lichens on hillsides of the Negev Highlands and their impact on biogenic weathering. *Z. Geomorph.* **27**: 423–444.
48. Duvdevani, S., Reichert, I., and Palti, J. (1946). The development of downy and powdery mildew of Cucumbers as related to dew and other environmental factors. *Palestine J. Bot.* (Rehovot Series). **2**: 127–151.
49. Stone, E.C. (1957). Dew as an ecological factor. I. A review of the literature. *Ecology* **38**: 407–413.
50. Waisel, Y. (1958). Dew absorption by plants of arid zones. *Bull. Res. Counc. Israel* **6D**: 180–186.
51. Schulze, E-D. et al. (1972). Stomatal responses to changes in humidity in plants growing in the desert. *Planta* **108**: 259–270.
52. Stone, E.C. (1957). Dew as an ecological factor. II. The effect of artificial dew on the survival of *Pinus ponderosa* and associated species. *Ecology* **38**: 414–422.
53. Rotem, J. and Reichert, J. (1964). Dew - a principal moisture factor enabling early blight epidemics in a semi-arid region of Israel. *Plant Disease Reptr.* **48**: 211–215.
54. Jones, C.G. and Shachak, M. (1990). Fertilization of the desert soil by rock-eating snails. *Nature* **345**: 839–841.
55. Shachak, M., Jones, C.G., and Brand, S. (1995). The role of animals in an arid ecosystem: Snails and isopods as controllers of soil formation, erosion and desalinization. *Adv. GeoEcol.* **28**: 37–50.

DEW POINT

Aldo Conti
Frascati (RM), Italy

The dew point is the temperature to which air must be cooled at constant pressure until it reaches saturation and the water, or any other vapor, begins condensing. The result is the formation of fog. When the dew point falls below freezing, it is called the frost point. In this case, water forms ice crystals directly. The dew point is also a good indicator of the amount of water contained in the air. The more humid the air, the higher the dew temperature. The dew point is often reached in the evening, when the air cools down, which is why fog is more common in the evening, or why one normally finds that dew or frost are common in the morning but disappear as soon as the temperature increases.

Dew point and relative humidity can be measured with an instrument called a wet- and dry-bulb psychrometer. The dry bulb is a normal thermometer that measures the actual air temperature. The wet bulb is another thermometer, but has a bulb wrapped in a piece of cloth dampened by a string that dips into a bottle of distilled water. Unless the air is so humid that it is very close to saturation, the wet bulb thermometer measures a lower temperature, because it is cooled by the evaporation of water. The difference between the dry-bulb and wet-bulb temperatures is called the wet-bulb depression. It is then possible to calculate the dew point as a function of the wet-bulb depression and of the dry-bulb temperature.

To calculate the dew point starting from the psychrometer readings, the first step is to calculate the saturation vapor pressure in millibars, corresponding to the dry- and wet-bulb temperature:

$$E_{\mathrm{s}} = 6.11 \times 10.0[7.5T/(237.7 + T)]$$

$$E_{\mathrm{swb}} = 6.11 \times 10.0[7.5T_{\mathrm{wb}}/(237.7 + T_{\mathrm{wb}})]$$

where T and T_{wb} are the readings of the dry- and wet-bulb temperatures.

Now we are ready to calculate the actual mixing ratio of the air:

$$W = [(T - T_{\mathrm{wb}})(C_{\mathrm{p}}) - L_{\mathrm{v}}(E_{\mathrm{swb}}/P)]/[-(T - T_{\mathrm{wb}})(C_{\mathrm{pv}}) - L_{\mathrm{v}}]$$

where C_{p} = specific heat of dry air at constant pressure (J/g)~1.005 J/g
C_{pv} = specific heat of water vapor at constant pressure (J/g)~4.186 J/g
L_{v} = Latent heat of vaporization (J/g)~2500 J/g
T = air temperature in °C
T_{wb} = wet bulb temperature in °C
E_{swb} = saturation vapor pressure at the wet-bulb temperature (mb)
P = atmospheric pressure at the surface ~1013 mb at sea level

We can now use the following formula to obtain the saturation mixing ratio for the air:

$$W_{\mathrm{s}} = E_{\mathrm{s}}/P$$

$$\mathrm{RH} = W/W_{\mathrm{s}}$$

Now we can use the relative humidity to calculate the actual vapor pressure (E) of the air as follows:

$$E = \mathrm{RH} \times E_{\mathrm{s}}$$

The dew point temperature is then

$$T_{\mathrm{d}} = [-430.22 + 237.7 \times \ln(E)]/[-\ln(E) + 19.08]$$

Let us now assume a psychrometer that gives the following readings at a pressure of 1013 mbar:

$$T = 30\,^{\circ}\mathrm{C}$$

$$T_{\mathrm{wb}} = 20\,^{\circ}\mathrm{C}$$

Then,

$$E_{\mathrm{swb}} = 23.34 \text{ mbar}$$

$$E_{\mathrm{s}} = 42.31 \text{ mbar}$$

The mixing ratio becomes

$$W = 0.019$$

$$W_{\mathrm{s}} = 0.042$$

The relative humidity is then,

$$\mathrm{RH} = 0.45 \text{ (or 45\%)}$$

The actual vapor pressure is then,

$$\mathrm{E} = 18.95 \text{ mbar}$$

and the dew point is

$$T_{\mathrm{d}} = 16.7\,^{\circ}\mathrm{C}$$

Instead of going through all these formulas, it is possible to use tables that give the dew point temperature as a function of the readings for different pressures.

DROUGHTS

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

A drought is a period of time during which weather is in one way or another excessively dry. A drought's severity is determined by the degree of moisture deficiency in the affected area, as measured by various methods. Droughts are a standard, recurring, and normal feature of long-term climate. There are four generally accepted categories of drought: meteorological, socioeconomic, hydrologic, and agricultural.

Meteorological droughts involve a lack of precipitation during a given time period. Their specific definition depends mostly on the region in question; a lack of precipitation for some areas would be excessive precipitation in others. Some regions receive rain year-round, and some receive virtually none. Likewise, some regions receive precipitation consistently throughout the year, and others have seasonal precipitation patterns. A departure from any of these climatological characteristics as abnormally low precipitation can cause a meteorological drought. Meteorological droughts are usually measured in terms of their duration and in contrast to established precipitation averages for specified time periods for a specified area. This type of drought generally precedes the other three types.

Agricultural droughts need not involve a severe lack of precipitation. They are marked by damage to crops and plants due to insufficient water supplies. Because water from rain and snow is more likely to run into creeks and rivers instead of filtering into the soil due to human intervention, agricultural droughts can sometimes occur in spite of noticeable rainfall. Depending on a plant's species and stage of growth, it needs varying amounts of water from the soil; when the majority of crops cannot obtain the amount of water they need, an area is said to be undergoing an agricultural drought. This type of drought is most damaging to farmers and those in poor countries, and it generally comes after the onset of a meteorological drought.

A hydrologic drought involves water shortages in reservoirs, lakes, and streams due to lack of precipitation. These droughts are measured by stream flow and water levels in local reservoirs. Local climate is not the only factor that causes a hydrologic drought; altered land usage, dams, and degraded soil quality can all have effects as well. Reduced precipitation in one area can cause a hydrologic drought in another because various areas are hydrologically connected by their rivers, lakes, and other bodies of water. The primary cause of this type of drought is generally a meteorological drought; hydrologic droughts usually come some time after agricultural droughts.

The final category of drought is socioeconomic. This type of drought, like the other three, involves a lack of precipitation. It occurs when reduced precipitation causes a shortage of water as related to the demands of the local populace. The socioeconomic drought is the only type of drought that has a major effect on the general population, and it is generally the last of the four types to occur.

Aside from lack of rain, the three most important variables to look at during a drought are air temperature, humidity, and circulation patterns in the atmosphere. Atmospheric circulation patterns that produce little or no precipitation are often associated with droughts. Climate abnormalities are another component of a drought. Precipitation levels that would be considered a drought in one area are normal in another that has a different climate. To understand the cause of a drought, scientists study the circulation patterns of the atmosphere across global distances, using computer-generated atmospheric simulator models.

Droughts also have direct effects on communities that experience this phenomenon. Both water quality and cost are negatively affected by droughts. Changes in the taste and odor of drinking water can also occur. Those who draw their water from wells may have to drill more deeply to reach the lowered level of the water table.

During a drought, water conservation is important. Restrictions on watering gardens, washing cars, and using sprinkler systems are applied when a drought emergency or warning is in effect. There are generally three stages of drought preparedness: drought watches, drought warnings, and drought emergencies. During a drought watch, no restrictions are enforced, but citizens are encouraged to curb water usage because the possibility of drought has been established for an area. During a drought warning, restrictions are still not enforced, but citizens are instructed to reduce water usage further in response to impending drought conditions. The final stage is a drought emergency. This is the most serious stage, often involving mandatory water usage guidelines to ensure that sufficient water is available for critical needs. At this stage, officials try to avoid local shortages by evenly distributing the water available.

Droughts can have a great impact in many ways. There are three main types of impact they can have: economic, social, and environmental. A lack of water in the soil for crops can cause farmers great financial difficulty. The resulting decrease in crop production has a ripple effect on the economy, causing short-term and long-term problems. One of the biggest short-term problems is unemployment. Certain industries can also be affected by long-term problems. For instance, loss of tax revenue can hinder tourism. Brush and trees can become very dry, and an outbreak of destructive fires can result. Drought can harm the logging industry, fisheries, and hydroelectric power generation, all economic effects. It can also cause insect problems in agriculture, as well as erosion and disease. Environmental effects include reduced biodiversity, degraded air quality, and perhaps most importantly, devastating forest fires. Social impacts include induced emigration, rampant famine, and, indirectly, greater poverty. These are merely a few of the multitude of impacts, direct and indirect, that drought can have.

READING LIST

All About Droughts. (2002). Available: http://205.156.54.206/om/drought.htm. (March 13).

Hanson, Ronald L. Evaporation and Droughts, *U.S. Geological Survey*. (2002). Available: http://geochange.er.usgs.gov/sw/changes/natural/er/. (March 13).

Impacts of Drought. (2002). Available: http://enso.unl.edu/ndmc/enigma/impacts.htm. (March 13).

McNab, A. and Karl, T. Climate and Droughts, *National Oceanic and Atmospheric Administration*. (2002). Available: http://geochange.er.usgs.gov/sw/changes/natural/drought/. (March 13).

Stormfax Guide to Droughts, *Stormfax*. Available: http://www.stormfax.com/drought.htm.

Understanding Your Risk: Impacts of Drought. Available: http://www.drought.unl.edu/risk/impacts.htm.

All About Droughts. Available: http://www.nws.noaa.gov/om/drought.htm.

Understanding and Defining Drought. Available: http://www.drought.unl.edu/whatis/concept.htm.

Drought Information Center. Available: http://www.dep.state.pa.us/dep/subject/hotopics/drought/drtterm.htm.

DROUGHT INDICES

Michael J. Hayes
Climate Impacts
Specialist—National Drought
Mitigation Center

INTRODUCTION

Drought indices assimilate thousands of bits of data on rainfall, snowpack, streamflow, and other water supply indicators into a comprehensible big picture. A drought index value is typically a single number, far more useful than raw data for decision making.

There are several indices that measure how much precipitation for a given period of time has deviated from historically established norms. Although none of the major indices is inherently superior to the rest in all circumstances, some indices are better suited than others for certain uses. For example, the Palmer Drought Severity Index has been widely used by the U.S. Department of Agriculture to determine when to grant emergency drought assistance, but the Palmer is better when working with large areas of uniform topography. Western states, with mountainous terrain and the resulting complex regional microclimates, find it useful to supplement Palmer values with other indices such as the Surface Water Supply Index, which takes snowpack and other unique conditions into account.

The National Drought Mitigation Center is using a newer index, the Standardized Precipitation Index, to monitor moisture supply conditions. Distinguishing traits of this index are that it identifies emerging droughts months sooner than the Palmer Index and that it is computed on various time scales.

Most water supply planners find it useful to consult one or more indices before making a decision. What follows is an introduction to each of the major drought indices in use in the United States and in Australia.

PERCENT OF NORMAL

Overview: The percent of normal is a simple calculation well suited to the needs of TV weathercasters and general audiences.

Pros: Quite effective for comparing a single region or season.

Cons: Easily misunderstood, as normal is a mathematical construct that does not necessarily correspond with what we expect the weather to be.

The percent of normal precipitation is one of the simplest measurements of rainfall for a location. Analyses using

---

the percent of normal are very effective when used for a single region or a single season. Percent of normal is also easily misunderstood and gives different indications of conditions, depending on the location and season. It is calculated by dividing actual precipitation by normal precipitation—typically considered to be a 30-year mean—and multiplying by 100%. This can be calculated for a variety of time scales. Usually these time scales range from a single month to a group of months representing a particular season, to an annual or water year. Normal precipitation for a specific location is considered to be 100%.

One of the disadvantages of using the percent of normal precipitation is that the mean, or average, precipitation is often not the same as the median precipitation, which is the value exceeded by 50% of the precipitation occurrences in a long-term climate record. The reason for this is that precipitation on monthly or seasonal scales does not have a normal distribution. Use of the percent of normal comparison implies a normal distribution where the mean and median are considered to be the same. An example of the confusion this could create can be illustrated by the long-term precipitation record in Melbourne, Australia, for the month of January. The median January precipitation is 36.0 mm (1.4 in.), meaning that in half the years less than 36.0 mm is recorded, and in half the years more than 36.0 mm is recorded. However, a monthly January total of 36.0 mm would be only 75% of normal when compared to the mean, which is often considered to be quite dry. Because of the variety in the precipitation records over time and location, there is no way to determine the frequency of the departures from normal or compare different locations. This makes it difficult to link a value of a departure with a specific impact occurring as a result of the departure, inhibiting attempts to mitigate the risks of drought based on the departures from normal and form a plan of response (1).

STANDARDIZED PRECIPITATION INDEX (SPI)

Overview: The SPI is an index based on the probability of precipitation for any time scale.
Who uses it: Many drought planners appreciate the SPI's versatility.
Pros: The SPI can be computed for different time scales, can provide early warning of drought and help assess drought severity, and is less complex than the Palmer.
Cons: Values based on preliminary data may change.
Developed by: T.B. McKee, N.J. Doesken, and J. Kleist, Colorado State University, 1993.
Monthly maps: http://www.drought.unl.edu/monitor/spi.htm; http://www.wrcc.dri.edu/spi/spi.html.

SPI Values	
2.0+	extremely wet
1.5 to 1.99	very wet
1.0 to 1.49	moderately wet
−.99 to .99	near normal
−1.0 to −1.49	moderately dry
−1.5 to −1.99	severely dry
−2 and less	extremely dry

The understanding that a deficit of precipitation has different impacts on groundwater, reservoir storage, soil moisture, snowpack, and streamflow led McKee, Doesken, and Kleist to develop the Standardized Precipitation Index (SPI) in 1993. The SPI was designed to quantify the precipitation deficit for multiple time scales. These time scales reflect the impact of drought on the availability of the different water resources. Soil moisture conditions respond to precipitation anomalies on a relatively short scale. Groundwater, streamflow, and reservoir storage reflect the longer-term precipitation anomalies. For these reasons, McKee et al. (2) originally calculated the SPI for 3-, 6-, 12-, 24-, and 48-month time scales.

The SPI calculation for any location is based on the long-term precipitation record for a desired period. This long-term record is fitted to a probability distribution, which is then transformed into a normal distribution so that the mean SPI for the location and desired period is zero (3). Positive SPI values indicate greater than median precipitation, and negative values indicate less than median precipitation. Because the SPI is normalized, wetter and drier climates can be represented in the same way, and wet periods can also be monitored using the SPI.

McKee et al. (2) used the classification system shown in the SPI values table to define drought intensities resulting from the SPI. McKee et al. (2) also defined the criteria for a drought event for any of the time scales. A drought event occurs any time the SPI is continuously negative and reaches an intensity of −1.0 or less. The event ends when the SPI becomes positive. Each drought event, therefore, has a duration defined by its beginning and end, and an intensity for each month that the event continues. The positive sum of the SPI for all the months within a drought event can be termed the drought's "magnitude".

Based on an analysis of stations across Colorado, McKee determined that the SPI is in mild drought 24% of the time, in moderate drought 9.2% of the time, in severe drought 4.4% of the time, and in extreme drought 2.3% of the time (2). Because the SPI is standardized, these percentages are expected from a normal distribution of the SPI. The 2.3% of SPI values within the "Extreme Drought" category is a percentage that is typically expected for an "extreme" event (Wilhite 1995). In contrast, the Palmer Index reaches its "extreme" category more than 10% of the time across portions of the central Great Plains. This standardization allows the SPI to determine the rarity of a current drought, as well as the probability of the precipitation necessary to end the current drought (2).

The SPI has been used operationally to monitor conditions across Colorado since 1994 (4). Monthly maps of the SPI for Colorado can be found on the Colorado State University website (http://ulysses.atmos.colostate.edu/SPI.html). It is also being monitored at the climate division level for the contiguous United States by the National Drought

Mitigation Center and the Western Regional Climate Center (WRCC). You can download the SPI program and sample files here.

PALMER DROUGHT SEVERITY INDEX (THE PALMER; PDSI)

Overview: The Palmer is a soil moisture algorithm calibrated for relatively homogeneous regions.

Who uses it: Many U.S. government agencies and states rely on the Palmer to trigger drought relief programs.

Pros: The first comprehensive drought index developed in the United States.

Cons: Palmer values may lag emerging droughts by several months; less well suited for mountainous land or areas of frequent climatic extremes; complex—has an unspecified, built-in time scale that can be misleading.

Developed by: W.C. Palmer, 1965.

Weekly maps: http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/regional_monitoring/palmer.gif.

Palmer Classifications	
4.0 or more	extremely wet
3.0 to 3.99	very wet
2.0 to 2.99	moderately wet
1.0 to 1.99	slightly wet
0.5 to 0.99	incipient wet spell
0.49 to −0.49	near normal
−0.5 to −0.99	incipient dry spell
−1.0 to −1.99	mild drought
−2.0 to −2.99	moderate drought
−3.0 to −3.99	severe drought
−4.0 or less	extreme drought

In 1965, W.C. Palmer developed an index to measure the departure of the moisture supply (5). Palmer based his index on the supply-and-demand concept of the water balance equation, taking into account more than just the precipitation deficit at specific locations. The objective of the Palmer Drought Severity Index (PDSI), as this index is now called, was to provide measurements of moisture conditions that were standardized so that comparisons using the index could be made between locations and between months (5).

The PDSI is a meteorological drought index, and it responds to weather conditions that have been abnormally dry or abnormally wet. When conditions change from dry to normal or wet, for example, the drought measured by the PDSI ends without taking into account streamflow, lake and reservoir levels, and other longer-term hydrologic impacts (6). The PDSI is calculated based on precipitation and temperature data, as well as the local Available Water Content (AWC) of the soil. From the inputs, all the basic terms of the water balance equation can be determined, including evapotranspiration, soil recharge, runoff, and moisture loss from the surface layer. Human impacts on the water balance, such as irrigation, are not considered. Complete descriptions of the equations can be found in the original study by Palmer (5) and in the more recent analysis by Alley (7).

Palmer developed the PDSI to include the duration of a drought (or wet spell). His motivation was as follows: an abnormally wet month in the middle of a long-term drought should not have a major impact on the index, or a series of months with near-normal precipitation following a serious drought does not mean that the drought is over. Therefore, Palmer developed criteria for determining when a drought or a wet spell begins and ends, which adjust the PDSI accordingly. Palmer (5) described this effort and gave examples, and it is also described in detail by Alley (7). In near-real time, Palmer's index is no longer a meteorological index but becomes a hydrological index referred to as the Palmer Hydrological Drought Index (PHDI) because it is based on moisture inflow (precipitation), outflow, and storage, and does not take into account the long-term trend (6).

In 1989, a modified method to compute the PDSI was begun operationally (8). This modified PDSI differs from the PDSI during transition periods between dry and wet spells. Because of the similarities between these Palmer indices, the terms *Palmer Index* and *Palmer Drought Index* have been used to describe general characteristics of the indices.

The Palmer Index varies roughly between −6.0 and +6.0. Palmer arbitrarily selected the classification scale of moisture conditions based on his original study areas in central Iowa and western Kansas (5). Ideally, the Palmer Index is designed so that a −4.0 in South Carolina has the same meaning in terms of the moisture departure from a climatological normal as a −4.0 in Idaho (7). The Palmer Index has typically been calculated on a monthly basis, and a long-term archive of the monthly PDSI values for every climate division in the United States exists with the National Climatic Data Center from 1895 through the present. In addition, weekly Palmer Index values (actually modified PDSI values) are calculated for the climate divisions during every growing season and are available in the *Weekly Weather and Crop Bulletin*. These weekly Palmer Index maps are also available on the World Wide Web from the Climate Prediction Center at http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/regional_monitoring/palmer.gif.

The Palmer Index is popular and has been widely used for a variety of applications across the United States. It is most effective measuring impacts sensitive to soil moisture conditions, such as agriculture (1). It has also been useful as a drought monitoring tool and has been used to trigger actions associated with drought contingency plans (1). Alley (7) identified three positive characteristics of the Palmer Index that contribute to its popularity: (1) it provides decision makers with a measurement of the abnormality of recent weather for a region; (2) it provides an opportunity to place current conditions in historical perspective; and (3) it provides spatial and temporal representations of historical droughts. Several states, including New York, Colorado, Idaho, and Utah, use the Palmer Index as one part of their drought monitoring systems.

There are considerable limitations when using the Palmer Index, and these are described in detail by Alley (7) and Karl and Knight (6). Drawbacks of the Palmer Index include:

- The values quantifying the intensity of drought and signaling the beginning and end of a drought or wet spell were arbitrarily selected based on Palmer's study of central Iowa and western Kansas and have little scientific meaning.
- The Palmer Index is sensitive to the AWC of a soil type. Thus, applying the index for a climate division may be too general.
- The two soil layers within the water balance computations are simplified and may not be accurately representative of a location.
- Snowfall, snow cover, and frozen ground are not included in the index. All precipitation is treated as rain, so that the timing of PDSI or PHDI values may be inaccurate in the winter and spring months in regions where snow occurs.
- The natural lag between when precipitation falls and the resulting runoff is not considered. In addition, no runoff is allowed to take place in the model until the water capacity of the surface and subsurface soil layers is full, leading to an underestimation of runoff.
- Potential evapotranspiration is estimated using the Thornthwaite method. This technique has wide acceptance, but it is still only an approximation.

Several other researchers have presented additional limitations of the Palmer Index. McKee et al. (4) suggested that the PDSI is designed for agriculture but does not accurately represent the hydrological impacts resulting from longer droughts. Also, the Palmer Index is applied within the United States but has little acceptance elsewhere (9). One explanation for this is provided by Smith et al. (10), who suggested that it does not do well in regions where there are extremes in the variability of rainfall or runoff. Examples in Australia and South Africa were given. Another weakness in the Palmer Index is that the "extreme" and "severe" classifications of drought occur with a greater frequency in some parts of the country than in others (1). "Extreme" droughts in the Great Plains occur with a frequency greater than 10%. This limits the accuracy of comparing the intensity of droughts between two regions and makes planning response actions based on a certain intensity more difficult.

CROP MOISTURE INDEX (CMI)

Description: A Palmer derivative, the CMI reflects moisture supply in the short term across major crop-producing regions and is not intended to assess long-term droughts.
Pros: Identifies potential agricultural droughts.
Developed by: W.C. Palmer, 1968.
Weekly maps: http://www.cpc.ncep.noaa.gov/products/_analysis_monitoring/regional_monitoring/cmi.gif.

The Crop Moisture Index (CMI) uses a meteorological approach to monitor week-to-week crop conditions. It was developed by Palmer (11) from procedures within the calculation of the PDSI. Whereas the PDSI monitors long-term meteorological wet and dry spells, the CMI was designed to evaluate short-term moisture conditions across major crop-producing regions. It is based on the mean temperature and total precipitation for each week within a climate division, as well as the CMI value from the previous week. The CMI responds rapidly to changing conditions, and it is weighted by location and time so that maps, which commonly display the weekly CMI across the United States, can be used to compare moisture conditions at different locations. Weekly maps of the CMI are available as part of the USDA/JAWF Weekly Weather and Crop Bulletin (http://www.usda.gov/oce/waob/jawf/wwcb.html).

Because it is designed to monitor short-term moisture conditions affecting a developing crop, the CMI is not a good long-term drought monitoring tool. The CMI's rapid response to changing short-term conditions may provide misleading information about long-term conditions. For example, a beneficial rainfall during a drought may allow the CMI value to indicate adequate moisture conditions, while the long-term drought at that location persists. Another characteristic of the CMI that limits its use as a long-term drought monitoring tool is that the CMI typically begins and ends each growing season near zero. This limitation prevents the CMI from being used to monitor moisture conditions outside the general growing season, especially in droughts that extend over several years. The CMI also may not be applicable during seed germination at the beginning of a specific crop's growing season.

SURFACE WATER SUPPLY INDEX (SWSI; PRONOUNCED "SWAZEE")

Description: The SWSI is designed to complement the Palmer in the state of Colorado, where mountain snowpack is a key element of water supply; calculated by river basin, based on snowpack, streamflow, precipitation, and reservoir storage.
Pros: Represents water supply conditions unique to each basin.
Cons: Changing a data collection station or water management requires that new algorithms be calculated, and the index is unique to each basin, which limits interbasin comparisons.
Developed by: Shafer and Dezman, 1982.

The Surface Water Supply Index (SWSI) was developed by Shafer and Dezman (12) to complement the Palmer Index for moisture conditions across the state of Colorado. The Palmer Index is basically a soil moisture algorithm calibrated for relatively homogeneous regions, but it is not designed for large topographic variations across a region and it does not account for snow accumulation and subsequent runoff. Shafer and Dezman designed the SWSI to be an indicator of surface water conditions and described the index as "mountain water dependent", in which mountain snowpack is a major component.

The objective of the SWSI was to incorporate both hydrological and climatological features into a single index value resembling the Palmer Index for each major river basin in the state of Colorado (Shafer and Dezman 1982). These values would be standardized to allow comparisons between basins. Four inputs are required within the SWSI: snowpack, streamflow, precipitation, and reservoir storage. Because it is dependent on the season, the SWSI is computed with only snowpack, precipitation, and reservoir storage in the winter. During the summer months, streamflow replaces snowpack as a component within the SWSI equation.

The procedure to determine the SWSI for a particular basin is as follows: monthly data are collected and summed for all the precipitation stations, reservoirs, and snowpack/streamflow measuring stations over the basin. Each summed component is normalized using a frequency analysis gathered from a long-term data set. The probability of non-exceedence—the probability that subsequent sums of that component will not be greater than the current sum—is determined for each component based on the frequency analysis. This allows comparisons of the probabilities to be made between the components. Each component has a weight assigned to it depending on its typical contribution to the surface water within that basin, and these weighted components are summed to determine a SWSI value representing the entire basin. Like the Palmer Index, the SWSI is centered on zero and has a range between −4.2 and +4.2.

The SWSI has been used, along with the Palmer Index, to trigger the activation and deactivation of the Colorado Drought Plan. One of its advantages is that it is simple to calculate and gives a representative measurement of surface water supplies across the state. It has been modified and applied in other western states as well. These states include Oregon, Montana, Idaho, and Utah. Monthly SWSI maps for Montana are available from the Montana Natural Resource Information System (http://nris.state.mt.us/wis/SWSInteractive/).

Several characteristics of the SWSI limit its application. Because the SWSI calculation is unique to each basin or region, it is difficult to compare SWSI values between basins or regions (13). Within a particular basin or region, discontinuing any station means that new stations need to be added to the system and new frequency distributions need to be determined for that component. Additional changes in the water management within a basin, such as flow diversions or new reservoirs, mean that the entire SWSI algorithm for that basin needs to be redeveloped to account for changes in the weight of each component. Thus, it is difficult to maintain a homogeneous time series of the index (8). Extreme events also cause a problem if the events are beyond the historical time series, and the index will need to be reevaluated to include these events within the frequency distribution of a basin component.

RECLAMATION DROUGHT INDEX

Description: Like the SWSI, the RDI is calculated at the river basin level, incorporating temperature as well as precipitation, snowpack, streamflow, and reservoir levels as input.

Who uses it: The Bureau of Reclamation; the State of Oklahoma as part of their drought plan.

Pros: By including a temperature component, it also accounts for evaporation.

Cons: Because the index is unique to each river basin, interbasin comparisons are limited.

Developed by: The Bureau of Reclamation, as a trigger to release drought emergency relief funds.

RDI Classifications	
4.0 or more	extremely wet
1.5 to 4.0	moderately wet
1 to 1.5	normal to mild wetness
0 to −1.5	normal to mild drought
−1.5 to −4.0	moderate drought
−4.0 or less	extreme drought

The Reclamation Drought Index (RDI) was recently developed as a tool for defining drought severity and duration, and for predicting the onset and end of periods of drought. The impetus to devise the RDI came from the Reclamation States Drought Assistance Act of 1988, which allows states to seek assistance from the Bureau of Reclamation to mitigate the effects of drought.

Like the SWSI, the RDI is calculated at a river basin level, and it incorporates the supply components of precipitation, snowpack, streamflow, and reservoir levels. The RDI differs from the SWSI in that it builds a temperature-based demand component and a duration into the index. The RDI is adaptable to each particular region and its main strength is its ability to account for both climate and water supply factors.

Oklahoma has developed its own version of the RDI and plans to use the index as one tool within the monitoring system designated in the state's drought plan. The RDI values and severity designations are similar to the SPI, PDSI, and SWSI.

DECILES

Description: Groups monthly precipitation occurrences into deciles so that, by definition, "much lower than normal" weather cannot occur more often than 20% of the time.

Who Uses It: Australia.

Decile Classifications	
Deciles 1–2: lowest 20%	much below normal
Deciles 3–4: next lowest 20%	below normal
deciles 5–6: middle 20%	near normal
deciles 7–8: next highest 20%	above normal
deciles 9–10: highest 20%	much above normal

Pros: Provides an accurate statistical measurement of precipitation.
Cons: Accurate calculations require a long climatic data record.
Developed by: Gibbs and Maher, 1967.

Arranging monthly precipitation data into deciles is another drought-monitoring technique. It was developed by Gibbs and Maher (14) to avoid some of the weaknesses within the "percent of normal" approach. The technique they developed divided the distribution of occurrences over a long-term precipitation record into tenths of the distribution. They called each of these categories a *decile*. The first decile is the rainfall amount not exceeded by the lowest 10% of the precipitation occurrences. The second decile is the precipitation amount not exceeded by the lowest 20% of occurrences. These deciles continue until the rainfall amount identified by the tenth decile is the largest precipitation amount within the long-term record. By definition, the fifth decile is the median, and it is the precipitation amount not exceeded by 50% of the occurrences over the period of record. The deciles are grouped into five classifications.

The decile method was selected as the meteorological measurement of drought within the Australian Drought Watch System because it is relatively simple to calculate and requires less data and fewer assumptions than the Palmer Drought Severity Index (10). In this system, farmers and ranchers can only request government assistance if the drought is shown to be an event that occurs only once in 20–25 years (deciles 1 and 2 over a 100-year record) and has lasted longer than 12 months (15). This uniformity in drought classifications, unlike a system based on the percent of normal precipitation, has assisted Australian authorities in determining appropriate drought responses. One disadvantage of the decile system is that a long climatological record is needed to calculate the deciles accurately.

BIBLIOGRAPHY

1. Willeke, G., Hosking, J.R.M., Wallis, J.R., and Guttman, N.B. (1994). *The National Drought Atlas*. Institute for Water Resources Report 94–NDS–4, U.S. Army Corps of Engineers.
2. McKee, T.B., Doesken, N.J., and Kleist, J. (1993). The relationship of drought frequency and duration to time scales. Preprints, 8th Conference on Applied Climatology, pp. 179–184. January 17–22, Anaheim, CA.
3. Edwards, D.C. and McKee, T.B. (1997). Characteristics of 20th century drought in the United States at multiple time scales. Climatology Report Number 97–2, Colorado State University, Fort Collins, CO.
4. McKee, T.B., Doesken, N.J., and Kleist, J. (1995). Drought monitoring with multiple time scales. Preprints, 9th Conference on Applied Climatology, pp. 233–236. January 15–20, Dallas, TX.
5. Palmer, W.C. (1965). Meteorological drought. Research Paper No. 45, U.S. Department of Commerce Weather Bureau, Washington, DC.
6. Karl, T.R. and Knight, R.W. (1985). Atlas of Monthly Palmer Hydrological Drought Indices (1931–1983) for the Contiguous United States. Historical Climatology Series 3–7, National Climatic Data Center, Asheville, North Carolina.
7. Alley, W.M. (1984). The palmer drought severity index: limitations and assumptions. *Journal of Climate and Applied Meteorology* **23**: 1100–1109.
8. Heddinghaus, T.R. and Sabol, P. (1991). A review of the palmer drought severity index and where do we go from here? In: *Proc. 7th Conf. on Applied Climatology*, pp. 242–246. American Meteorological Society, Boston.
9. Kogan, F.N. (1995). Droughts of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data. *Bulletin of the American Meteorological Society* **76**(5): 655–668.
10. Smith, D.I., Hutchinson, M.F., and McArthur, R.J. (1993). Australian climatic and agricultural drought: Payments and policy. *Drought Network News* **5**(3): 11–12.
11. Palmer, W.C. (1968). Keeping track of crop moisture conditions, nationwide: The new Crop Moisture Index. *Weatherwise* **21**: 156–161.
12. Shafer, B.A. and Dezman, L.E. (1982). Development of a Surface Water Supply Index (SWSI) to assess the severity of drought conditions in snowpack runoff areas. In: *Proceedings of the Western Snow Conference*. Colorado State University, Fort Collins, CO, pp. 164–175.
13. Doesken, N.J., McKee, T.B., and Kleist, J. (1991). Development of a surface water supply index for the western United States. Climatology Report Number 91–3, Colorado State University, Fort Collins, CO.
14. Gibbs, W.J. and Maher, J.V. (1967). Rainfall deciles as drought indicators. *Bureau of Meteorology Bulletin* No. 48, Commonwealth of Australia, Melbourne.
15. White, D.H. and O'Meagher, B. (1995). Coping with exceptional droughts in Australia. *Drought Network News* **7**(2): 13–17.

READING LIST

Gommes, R. and Petrassi, F. (1994). Rainfall variability and drought in Sub-Saharan Africa since 1960. Agrometeorology Series Working Paper No. 9, Food and Agriculture Organization, Rome, Italy.

Le Houérou, H.N., Popov, G.F., and See, L. (1993). Agrobioclimatic classification of Africa. Agrometeorology Series Working Paper No. 6, Food and Agriculture Organization, Rome, Italy.

Wilhite, D.A. (1995). Developing a precipitation-based index to assess climatic conditions across Nebraska. Final report submitted to the Natural Resources Commission, Lincoln, NE.

Wilhite, D.A. and Glantz, M.H. (1985). Understanding the drought phenomenon: The role of definitions. *Water International* **10**(3): 111–120.

THE EARTH OBSERVING SYSTEM: AQUA

NASA—Goddard Space
Flight Center

EARTH SYSTEM SCIENCE

Beginning in the 1960s, NASA pioneered the study of the atmosphere from the unique perspective of space

---

with the launch of its Television Infrared Observation Satellite (TIROS-1). Thanks to new satellite and computer technologies, it is now possible to study the Earth as a global system. Through their research, scientists are better understanding and improving their forecasting of short-term weather phenomena.

Long-term weather and climate prediction is a greater challenge that requires the collection of better data over longer periods. Since climate changes occur over vast ranges of space and time, their causes and effects are often difficult to measure and understand. Scientists must obtain long-term data if they are to reach a full understanding of the interactions among the Earth's physical and biological systems. NASA's Earth Observing System (EOS) will help us to understand the complex links among air, land, water and life within the Earth system.

WHAT IS AQUA?

NASA's commitment to studying the Earth as a global system continues with the Aqua spacecraft (originally called EOS PM-1), representing a key contribution by NASA to the U.S. Global Change Research Program. Aqua carries six state-of-the-art instruments to observe the Earth's oceans, atmosphere, land, ice and snow covers, and vegetation, providing high measurement accuracy, spatial detail, and temporal frequency. This comprehensive approach to data collection enables scientists to study the interactions among the four spheres of the Earth system—the oceans, land, atmosphere, and biosphere.

Aqua, Latin for "water," is named for the large amount of information that the Aqua spacecraft will collect about the Earth's water cycle. In particular, the Aqua data will include information on water vapor and clouds in the atmosphere, precipitation from the atmosphere, soil wetness on the land, glacial ice on the land, sea ice in the oceans, snow cover on both land and sea ice, and surface waters throughout the world's oceans, bays, and lakes. Such information will help scientists improve the quantification of the global water cycle and examine such issues as whether or not the cycling of water might be accelerating.

In addition to information about the water cycle, Aqua also provides information on many additional elements of the Earth system. For instance, Aqua enables studies of the fluxes of radiation from the Sun and from the Earth that combine to constitute the Earth's radiation balance. It also enables studies of small particles in the atmosphere termed "aerosols" and such trace gases in the atmosphere as ozone, carbon monoxide, and methane. The trace gases each have a potential contribution to global warming, whereas the aerosols are more likely to have a cooling effect. Aqua also provides observations on vegetation cover on the land, phytoplankton and dissolved organic matter in the oceans, and the temperatures of the air, land, and water. All of these measurements have the potential to contribute to improved understanding of the changes occurring in the global climate and the role of the interactions among the various elements of the climate system.

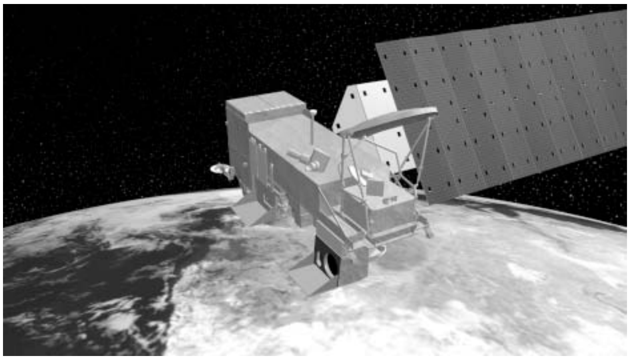

One of the most exciting of the potential practical benefits likely to derive from the Aqua data is improved weather forecasting. Aqua carries a sophisticated sounding system that allows determination of atmospheric temperatures around the world to an accuracy of 1° Celsius in 1-km-thick layers throughout the troposphere, the lowest portion of the atmosphere. The troposphere extends to an altitude of about 10–15 km, depending on location, and contains most of the global cloud cover. The anticipated 1° Celsius accuracy far exceeds current accuracies from satellite observations and, in conjunction with the moisture profiles also obtainable from the Aqua sounding system, offers the potential of improved weather fore-casting. NASA is working with the U.S. National Oceanic and Atmospheric Administration and the European Centre for Medium-Range Weather Forecasts to facilitate the incorporation of the Aqua data in their weather forecasting efforts.

INTERNATIONAL COLLABORATION

Aqua is a joint project of the United States, Japan, and Brazil.

THE SPACECRAFT

The spacecraft was designed and built by TRW in Redondo Beach, California. Aqua is based on TRW's modular, standardized AB1200 common spacecraft bus. This design features common subsystems scalable to the mission-specific needs of Aqua as well as future missions. Instrument payloads can be attached on a "mix and match" basis without changing the overall design or subsystem support requirements.

THE INSTRUMENTS

The Atmospheric Infrared Sounder (AIRS), built by BAE Systems, was provided by NASA's Jet Propulsion Laboratory in Pasadena, California. AIRS is the highlighted instrument in the AIRS/AMSU-A/HSB triplet centered on measuring humidity, temperature, cloud properties, and the amounts of greenhouse gases throughout the atmosphere. AIRS/AMSU-A/HSB will improve weather forecasting, establish the connection between severe weather and climate change, examine whether the global water cycle is accelerating, and detect the effects of greenhouse gases.

The Advanced Microwave Scanning Radiometer for EOS (AMSR-E), built by Mitsubishi Electronics Corporation, was provided by Japan's National Space Development Agency. AMSR-E measures precipitation rate, cloud water, water vapor, sea surface winds, sea surface temperature, ice, snow, and soil moisture.

The Advanced Microwave Sounding Unit (AMSU-A), built by Aerojet and provided by NASA's Goddard Space Flight Center (GSFC) in Greenbelt, Maryland, obtains temperature profiles in the upper atmosphere (especially the stratosphere) and will provide a cloud-filtering capability for tropospheric temperature observations. The EOS AMSU-A is part of the closely coupled AIRS/AMSU-A/HSB triplet.

The Clouds and the Earth's Radiant Energy System (CERES), built by TRW, was provided by NASA's Langley Research Center in Hampton, Virginia. This instrument measures the Earth's total thermal radiation budget, and, in combination with Moderate Resolution Imaging Spectro-radiometer (MODIS) data, provides detailed information about clouds. The first CERES instrument was launched on the Tropical Rainfall Measuring Mission (TRMM) satellite in November 1997; the second and third CERES instruments were launched on the Terra satellite in December 1999; and the fourth and fifth CERES instruments was on board the Aqua satellite launched in May 2002. The pairs of CERES on both Terra and Aqua satellites allow coincident measurements by one CERES scanning in lines perpendicular to the path of the satellite and by the other CERES scanning in lines at various angles with respect to the satellite's path.

The Humidity Sounder for Brazil (HSB), built by Matra-Marconi, was provided by Brazil's Instituto Nacional de Pesquisas Espaciais, the Brazilian Institute for Space Research. The HSB obtains humidity profiles throughout the atmosphere. The HSB is the instrument in the AIRS/AMSU-A/HSB suite that allows humidity measurements even under conditions of heavy cloudiness and haze.

MODIS, built by Raytheon Santa Barbara Remote Sensing, was provided by GSFC. MODIS is a 36-band spectroradiometer measuring visible and infrared radiation and obtaining data that is used to derive products ranging from vegetation, land surface cover, and ocean chlorophyll fluorescence, to cloud and aerosol properties, fire occurrence, snow cover on the land, and sea ice cover on the oceans. The first MODIS instrument was launched on board the Terra satellite.

Aqua was launched in May 2002 aboard a Delta 7920-10L launch vehicle from Vandenberg Air Force Base, California. The stowed spacecraft is 8.8 ft (2.68 m) × 8.2 ft (2.49 m) × 21.3 ft (6.49 m). Deployed, Aqua is 15.8 ft (4.81 m) × 54.8 ft (16.70 m) × 26.4 ft (8.04 m). The spacecraft, at launch, weighed 6784 lbs with a full propellant load of 508 lbs and is powered by 4.6 kilowatts of electric power from its solar array.

Aqua was launched into a circular 680-km orbit. Over a period of days after separation from the launch vehicle, it was commanded by the ground to raise its orbit to the prescribed 705-km (438-mile) orbit. This was necessary in order to allow for proper phasing of Aqua with other spacecraft in orbit and the polar ground stations used for communications. The spacecraft was ultimately be positioned in a near-polar (98°) orbit around the Earth in synchronization with the Sun, with its path over the ground ascending across the equator at the same local time every day, approximately 1:30 p.m. The early afternoon observation time contrasts with the 10:30–10:45 a.m. equatorial crossing time (descending in this case) of the Terra satellite. The two daytime crossing times account for why the Terra and Aqua satellites were originally named "EOS AM" and "EOS PM," respectively. The combination of morning and afternoon observations allows studies concerning the diurnal variability of many of the parameters discussed above.

MANAGEMENT

Overall management of the Aqua mission is located at GSFC, which is managing the integration and testing of the spacecraft. The Aqua data is processed, archived, and distributed using distributed components of the Earth Observing System Data and Information System (EOSDIS). EOSDIS also provides the mission operations systems that perform the functions of command and control of the spacecraft and the instruments. NASA's Kennedy Space Center is responsible for the launch operations, including Boeing's Delta launch vehicle and the prelaunch integrated processing facility. The U.S. Air Force is responsible for all range-related matters. GSFC manages EOS for NASA's Earth Science Enterprise (ESE), headquartered in Washington, DC.

DATA PROCESSING AND DISTRIBUTION

Aqua provides a major part of a 15-year environmental dataset focusing on global change. The Aqua instruments produce more than 750 gigabytes of data per day, which is equivalent to 75 personal computer hard disks at 10 gigabytes each per day. This massive amount of information is handled using EOSDIS, in addition to its present accumulation of nearly 3000 gigabytes per day.

EOSDIS provides the high-performance computing resources needed to process, store, and rapidly transmit terabytes (thousands of gigabytes) of the incoming data every day. EOSDIS has several distributed sites that perform these functions: Distributed Active Archive Centers (DAACs) that process, store and distribute the data, and Science Investigator-led Processing Systems that process the data and send them to the DAACs for storage and distribution. EOSDIS uses an "open" architecture to allow insertion of new technology while enabling the system to support the changing mission and science needs throughout the EOS Program.

GOALS AND OBJECTIVES

NASA's ESE identified several high-priority measurements that EOS should make to facilitate a better understanding of the components of the Earth system—the atmosphere, the land, the oceans, the polar ice caps, and the global energy budget. The specific objectives of Aqua include:

- producing high-spectral resolution obtaining 1 K/1 km global root-mean-square temperature profile accuracy in the troposphere by 1 year after launch;
- extending the improved TRMM rainfall characterization to the extra tropics;
- producing global sea surface temperature daily maps under nearly all sky conditions for a minimum of 1 year;
- producing large-scale global soil moisture distribution for regions with low vegetation;
- producing calibrated global observations of the Earth's continents and ocean surfaces 150 days after the mission is declared operational;
- capturing and documenting three seasonal cycles of terrestrial and marine ecosystems and atmospheric and cloud properties;
- producing three sets of seasonal/annual Earth radiation budget records;
- producing improved measurements of the diurnal cycle of radiation by combining Aqua measurements with Terra measurements for months of overlap;
- producing combined cloud property and radiation balance data to allow improved studies of the role of clouds in the climate system; and,
- capturing, processing, archiving, and distributing Aqua data products, by 150 days after the mission is declared operational.

A NEW PERSPECTIVE

Complemented by Terra, aircraft and ground-based measurements, Aqua data enable scientists to distinguish between natural and human-induced changes. The EOS series of spacecraft are the cornerstone of NASA's ESE, a long-term research effort to study the Earth as a global environment.

More information on EOS and the science related to it can be found at the EOS Project Science Office website at http://eospso.gsfc.nasa.gov and at the Earth Observatory website at http://earthobservatory.nasa.gov. Further information on Aqua can be found at http://aqua.nasa.gov

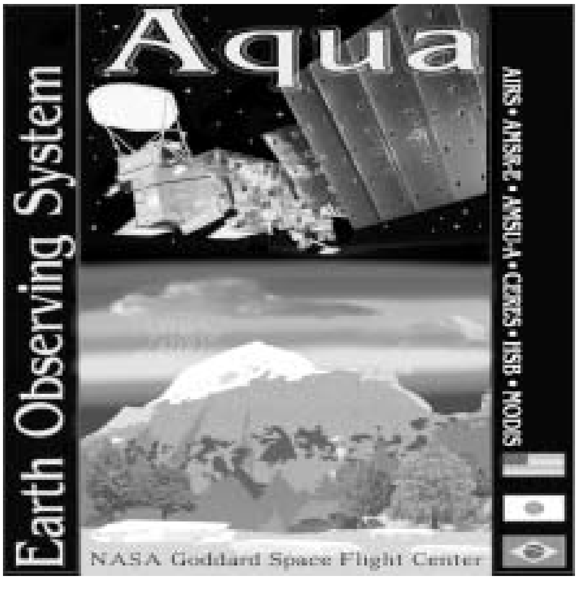

ENTROPY THEORY FOR HYDROLOGIC MODELING

Vijay P. Singh
Louisiana State University
Baton Rouge, Louisiana

INTRODUCTION

Entropy theory has recently been employed in a broad range of applications in hydrology, and new applications continue to unfold. This paper revisits entropy theory and its application to hydrologic modeling.

Hydrologic systems are inherently spatial and complex, and our understanding of these systems is less than complete. Many of the systems are either fully stochastic or part stochastic and part deterministic. Their stochastic nature can be attributed to randomness in one or more of the following components that constitute them: (1) system structure (geometry), (2) system dynamics, (3) forcing functions (sources and sinks), and (4) initial and boundary conditions. As a result, a stochastic description of these systems is needed, and entropy theory enables the development of such a description.

Engineering decisions concerning hydrologic systems are frequently made with less than adequate information. Such decisions may often be based on experience, professional judgment, rules of thumb, crude analyses, safety factors, or probabilistic methods. Usually, decision-making under uncertainty tends to be relatively conservative. Quite often, sufficient data are not available to describe the random behavior of such systems. Although probabilistic methods allow for a more explicit and quantitative accounting of uncertainty, their major difficulty occurs due to the lack of sufficient or complete data. Small sample sizes and limited information render estimation of probability distributions of system variables by conventional methods difficult. This problem can be alleviated by using entropy theory that enables determining the least biased probability distributions based on limited knowledge and data. Where the shortage of data is widely rampant, as is normally the case in many countries, entropy theory is particularly appealing.

Since the development of entropy theory by Shannon in the late 1940s and of the principle of maximum entropy (POME) by Jaynes in the late 1950s, there has been a proliferation in applications of entropy. The real impetus to entropy-based modeling in hydrology was, however, provided in the early 1970s, a great variety of entropy-based applications have since been reported, and new applications continue to unfold. This article aims to revisit entropy theory and to underscore its usefulness for both modeling and decision-making in hydrology.

ENTROPY THEORY

Entropy theory is comprised of three main parts: (1) Shannon entropy, (2) the principle of maximum entropy, and (3) the principle of minimum cross entropy. Before discussing these parts, it will be instructive to discuss briefly the meaning of entropy.

Meaning of Entropy

Entropy originated in physics. It is an extensive property like mass, energy, volume, momentum, charge, or number of atoms of chemical species, but unlike these quantities, it does not obey a conservation law. The entropy of a system is an extensive property, so the total entropy of the system equals the sum of the entropies of individual parts. The most probable distribution of energy in a system is the one that corresponds to the maximum entropy of the system. This occurs under the condition of dynamic equilibrium. During evolution toward a stationary state, the rate of entropy production per unit mass should be minimum, compatible with external constraints. This is the Prigogin principle. In thermodynamics, entropy is decomposed into two parts: (1) entropy exchanged between the system and its surroundings and (2) entropy produced in the system. According to the second law of thermodynamics, the entropy of a closed and isolated system always tends to increase. In hydraulics, entropy is a measure of the amount of irrecoverable flow energy that the hydraulic system expands to overcome friction. The system converts a portion of its mechanical energy to heat energy which then is dissipated to the external environment. Thus, the process equation in hydraulics expressing the energy (or head) loss, it can be argued, originates in the entropy concept.

Entropy has been employed in thermodynamics as a measure of the degree of ignorance about the true state of a system. If there were no energy loss in a hydraulic system, the system would be orderly and organized. The energy loss and its causes make the system disorderly and chaotic. Thus, entropy can be interpreted as a measure of the amount of chaos within a system. Algebraically, it is proportional to the logarithm of the probability of the state of the system. The constant of proportionality is the Boltzmann constant, and this defines Boltzmann entropy.

Shannon Entropy

Shannon (1) developed the entropy theory for expressing information or uncertainty. To understand the informational aspect of entropy, we perform an experiment on a random variable X. There may be n possible outcomes $x_1, x_2, \ldots, x_n$, whose probabilities are $p_1, p_2, \ldots, p_n$; $P(X = x_1) = p_1, P(X = x_2) = p_2, \ldots, P(X = x_n) = p_n$. These outcomes can be described by

$$P(X) = (p_1, p_2, \ldots, p_n);\ \sum_i^n p_i = 1; p_i \geq 0, i = 1, 2, \ldots, n \tag{1}$$

If this experiment is repeated, the same outcome is not likely, implying that there is uncertainty as to the outcome of the experiment. Based on one's knowledge about the outcomes, the uncertainty can be more or less. For example, the total number of outcomes is a piece of information, and the number of those outcomes of nonzero probability is another piece of information. The probability distribution of the outcomes, if known, provides a certain amount of information. Shannon (1) defined a quantitative measure of uncertainty associated with a probability distribution or the information content of the distribution in terms of entropy, $H(P)$ or $H(X)$, called Shannon entropy or informational entropy as

$$H(X) = H(P) = -\sum_{i=1}^{n} p_i \ln p_i = E[-\ln p] \tag{2}$$

If the random variable X is continuous, then Shannon entropy is expressed as

$$\begin{aligned} H(X) &= -\int_0^\infty f(x) \ln[f(x)]\, dx \\ &= -\int \ln[f(x)]\, dF(x) = E[-\ln f(x)] \end{aligned} \tag{3}$$

where $f(x)$ is the probability density function (PDF) of X, $F(x)$ is the cumulative probability distribution function of X, and E [.] is the expectation of [.].

Thus, entropy is a measure of the amount of uncertainty represented by the probability distribution and is a measure of the amount of chaos or of the lack of information about a system. If complete information is available, entropy $= 0$; otherwise, it is greater than zero. The uncertainty can be quantified using entropy by taking into account all different kinds of available information. Shannon entropy is the weighted Boltzmann entropy.

Principle of Maximum Entropy

In search of an appropriate probability distribution for a given random variable, entropy should be maximized. In practice, however, it is common that some information about the random variable is available. The chosen probability distribution should then be consistent with the given information. There can be more than one distribution consistent with the given information. From all such distributions, we should choose the distribution that has the highest entropy. To that end, Jaynes (2) formulated the principle of maximum entropy (POME), a full account of it is presented in a treatise by Levine and Tribus (3). According to POME, the minimally prejudiced assignment of probabilities is that which maximizes entropy subject to the given information, that is, POME takes into account all of the given information and at the same time avoids considering of any information that is not given.

If no information about the random variable is available, then all outcomes are equally likely, that is, $p_i = 1/n$, $i = 1, 2, 3, \ldots, n$. It can be shown that Shannon entropy is maximum in this case and may serve as an upper bound of entropy for all cases involving some information. In a more general case, let the information available about P or X be

$$p_i \geq 0, \sum_{i=1}^{n} p_i = 1 \tag{4}$$

and

$$\sum_{i=1}^{n} g_r(x_i) p_i = a_r \quad r = 1, 2, \ldots, m \tag{5}$$

where m is the number of constraints, $m + 1 \leq n$ and g_r is the rth constraint. Equations 4 and 5 are not sufficient to determine P uniquely. Therefore, there can be many distributions that satisfy Eqs. 4 and 5. According to POME, there will be only one distribution that will correspond to the maximum value of entropy, and this distribution can be determined using the method of Lagrange multipliers which will have the following form:

$$p_i = \exp[-\lambda_0 - \lambda_1 g_1(x_i) - \lambda_2 g_2(x_i) \ldots - \lambda_m g_m(x_i)] \quad i = 1, 2, \ldots, n \tag{6}$$

where λ_i, $i = 0, 1, 2, \ldots, m$, are Lagrange multipliers that are determined by using the information specified by Eqs. 4 and 5.

Because the POME-based distribution is favored over those with less entropy among those that satisfy the given constraints, according to Shannon entropy as an information measure, entropy defines a kind of measure on the space of probability distributions. Intuitively, distributions of higher entropy represent more disorder, are smoother, are more probable, are less predictable, or assume less. The POME-based distribution is maximally noncommittal with regard to missing information and does not require invoking ergodic hypotheses.

Principle of Minimum Cross Entropy

According to Laplace's principle of insufficient reason, all outcomes of an experiment should be considered equally likely, unless there is information to the contrary. On the basis of intuition, experience, or theory, a random variable may have an a priori probability distribution. Then, Shannon entropy is maximum when the probability distribution of the random variable is one which is as close to the a priori distribution as possible. This is called the principle of minimum cross entropy (POMCE) which minimizes Bayesian entropy (4). This is equivalent to maximizing Shannon entropy.

To explain POMCE, let us suppose that we guess a probability distribution for a random variable x as $Q = \{q_1, q_2, \ldots, q_n\}$ based on intuition, experience, or theory. This constitutes the prior information in the form of a prior distribution. To verify our guess, we take some observations $X = (x_1, x_2, \ldots, x_n)$ and compute some moments of the distribution. To derive the distribution $P = \{p_1, p_2, \ldots, p_n\}$ of X, we take all the given information and make the distribution as near to our intuition and experience as possible. Thus, POMCE is expressed as

$$D(P, Q) = \sum_{i=1}^{n} p_i \ln \frac{p_i}{q_i} \tag{7}$$

where the cross entropy D is minimized. If no a priori distribution is available and if, according to Laplace's principle of insufficient reason, Q is chosen as a uniform distribution U, then Eq. 7 takes the form

$$D(P, U) = \sum_{i=1}^{n} p_i \ln \left[\frac{p_i}{1/n} \right] = \ln n \left(\sum_{i=1}^{n} p_i \ln p_i \right) \tag{8}$$

Hence, minimizing $D(P, U)$ is equivalent to maximizing Shannon entropy. Because D is a convex function, its local minimum is its global minimum. Thus, a posterior distribution P is obtained by combining a prior Q with the specified constraints. The distribution P minimizes the cross (or relative) entropy with respect to Q, defined by Eq. 7, where the entropy of Q is defined as in Eq. 2. Cross-entropy minimization results asymptotically from Bayes' theorem.

JOINT ENTROPY, CONDITIONAL ENTROPY, AND TRANSINFORMATION

If there are two random variables X and Y whose probability distributions are $P(x) = \{p_1, p_2, \ldots, p_n\}$ and $Q(y) = \{q_1, q_2, \ldots, q_n\}$, which are independent, then Shannon entropy of the joint distribution of X and Y is the sum of the entropies of the marginal distributions expressed as

$$H(P, Q) = H(X, Y) = H(P) + H(Q) = H(X) + H(Y) \tag{9}$$

If the two random variables are dependent, then Shannon entropy of the joint distribution is the sum of the marginal entropy of one variable and the conditional entropy of the other variable conditioned on the realization of the first. Expressed algebraically,

$$H(X, Y) = H(X) + H(Y \mid X) \tag{10}$$

where $H(Y|X)$ is the conditional entropy of Y conditioned on X. The conditional entropy can be defined as

$$H(X|Y) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) \ln(p(x_i \mid y_j) \tag{11}$$

It is seen that if X and Y are independent, then Eq. 10 reduces to Eq. 9. Furthermore, the joint entropy of dependent X and Y will be less than or equal to the joint entropy of independent X and Y, that is, $H(X, Y) \leq H(X) + H(Y)$. The difference between these two entropies defines transinformation $T(X, Y)$ or $T(P, Q)$ expressed as

$$T(X, Y) = H(X) + H(Y) - H(X, Y) \tag{12}$$

Transinformation represents the amount of information common to both X and Y. If X and Y are independent, $T(X, Y) = 0$. Substitution of Eq. 10 in Eq. 12 yields

$$T(X, Y) = H(Y) - H(Y|X) \tag{13}$$

Equation 13 states that stochastic dependence reduces the entropy of Y.

ENTROPY AS A MODELING TOOL

Although entropy theory has been applied in recent years to a variety of problems in hydrology, its potential as a decision-making tool has not been fully exploited. A brief discussion follows highlighting this potential. Fundamental to the concepts presented below is the need for probability distributions that can be derived by using entropy theory.

Information Content of Data

One frequently encounters a situation in which one exercises freedom of choice, evaluates uncertainty, or measures information gain or loss. The freedom of choice, uncertainty, disorder, information content, or information gain or loss has been variously measured by relative entropy, redundancy, and conditional and joint entropies employing conditional and joint probabilities. As an example, in the analysis of empirical data, the variance has often been interpreted as a measure of uncertainty and as revealing gain or loss in information. However, entropy is another measure of dispersion—an alternative to variance. This suggests that it is possible to determine the variance whenever it is possible to determine entropy measures, but the reverse is not necessarily true. However, variance is not the appropriate measure if the sample size is small.

To measure correlation or dependence between any two variables, an informational coefficient of correlation r_0 is defined as a function of transinformation, T_0, as

$$r_0 = [1 - \exp(-2T_0)]^{0.5} \tag{14}$$

The transinformation, given by Eq. 14, expresses the upper limit of common information between two variables and represents the level of dependence (or association) between the variables. It represents the upper limit of transferable information between the variables, and its measure is given by r_0. The ordinary correlation coefficient r measures the amount of information transferred between variables under specified assumptions, such as linearity and normality. An inference similar to that of the ordinary correlation coefficient, r, can be drawn by defining the amount (in percent) of transferred information by the ratio T/T_0, where T can be computed in terms of ordinary r.

Criteria for Model Selection

Usually, there are more models than one needs, and so a model has to be chosen. Akaike (5) formulated a criterion, called the Akaike information criterion (AIC), for selecting the best model from amongst several models as

$$\text{AIC} = 2\log(\text{maximized likelihood}) + 2k \tag{15}$$

AIC provides a method of model identification and can be expressed as minus twice the logarithm of the maximum likelihood plus twice the number of parameters used to find the best model. The maximum likelihood and entropy are uniquely related. When there are several models, the model that gives the minimum value of AIC should be selected. When the maximum likelihood is identical for two models, the model that has the smaller number of parameters should be selected, for that will lead to a smaller AIC and comply with the principle of parsimony.

Hypothesis Testing

Another important application of entropy theory is testing of hypotheses (6). By using Bayes' theorem in logarithmic form, an evidence function is defined for comparing two hypotheses. The evidence in favor of a hypothesis over its competitor is the difference between the respective entropies of the competition and the hypothesis under test. Defining surprisal as the negative of the logarithm of the probability, the mean surprisal for a set of observations is expressed. Therefore, the evidence function for two hypotheses is obtained as the difference between the two values of the mean surprisal multiplied by the number of observations.

Risk Assessment

In common language, risk is the possibility of loss or injury and the degree of probability of such loss. Rational decision-making requires a clear and quantitative way of expressing risk. In general, risk cannot be avoided, and a choice has to be made between risks. There are different types of risk, such as business risk, social risk, economic risk, safety risk, investment risk, and occupational risk. To put risk in proper perspective, it is useful to clarify the distinction between risk, uncertainty, and hazard.

The notion of risk involves both uncertainty and some kind of loss or damage. Uncertainty reflects the variability of our state of knowledge or state of confidence in a prior evaluation. Thus, risk is the sum of uncertainty plus damage. Hazard is commonly defined as a source of danger that involves a scenario identification (e.g., failure of a dam) and a measure of the consequence of that scenario or a measure of the ensuing damage. Risk encompasses the likelihood of converting that source into the actual delivery of loss, injury, or some form of damage. Thus, risk is the ratio of hazard to safeguards. By increasing safeguards, risk can be reduced, but it is never zero. Awareness of risk reduces risk, so awareness is part of safeguards. Qualitatively, risk is subjective and is relative to the observer. Risk involves the probability of a scenario and its consequence resulting from the occurrence of the scenario. Thus, one can say that risk is probability and consequence. Kaplan and Garrick (7) analyzed risk using entropy.

HYDROLOGIC MODELING USING ENTROPY THEORY

A historical perspective on entropy applications in environmental and water resources is given in Singh and Fiorentino (8) and Singh (9). Harmancioglu and Singh (10) discussed the use of entropy in water resources. A brief synopsis of entropy-based applications follows.

Derivation of Probability Distributions

Frequency distributions that satisfy the given information are often needed. Entropy theory is ideally suited to that end. POME has been employed to derive a variety of distributions; some have found wide applications in environmental and water resources. Many of these distributions have been summarized in Singh and Fiorentino (8) and by Singh (9). Let $p(x)$ be the probability distribution of X that is to be determined. The information on X is available in terms of constraints given by Eq. 2. Then, the entropy-based distribution is given by Eq. 6. Substitution of Eq. 5 in Eq. 2 yields

$$\exp(\lambda_0) = Z = \sum_{i=1}^{n} \exp\left[-\sum_{j=1}^{m} \lambda_j g_j(x_i) \right] \qquad (16)$$

where Z is called the partition function and λ_0 is the zeroth Lagrange multiplier. The Lagrange parameters are obtained by differentiating Eq. 16 with respect to the Lagrange multipliers:

$$\begin{aligned} \frac{\partial \lambda_0}{\partial \lambda_j} &= -a_j = E[g_j], \quad j = 1, 2, 3, \ldots, m \\ \frac{\partial^2 \lambda_0}{\partial \lambda_j^2} &= \mathrm{Var}[g_j] \\ \frac{\partial^2 \lambda_0}{\partial \lambda_j \partial \lambda_k} &= \mathrm{Cov}[g_j, g_k] \\ \frac{\partial^3 \lambda_0}{\partial \lambda_j^3} &= -\mu_3[g_j] \end{aligned} \qquad (17)$$

where E[.] is the expectation, Var [.] is the variance, Cov [.] is the covariance, and is the third moment about the centroid, all for g_j.

When there are no constraints, then POME yields a uniform distribution. As more constraints are introduced, the distribution becomes more peaked and possibly skewed. In this way, entropy reduces from a maximum for the uniform distribution to zero when the system is fully deterministic.

Parameter Estimation

It is desirable to estimate parameters of a distribution in terms of the given constraints. Entropy theory accomplishes precisely that. Singh (11) described POME-based estimation for a number of probability distributions used in hydrology and environmental and water resources. He also discussed a comparison of the POME-based method with the methods of moments, maximum likelihood estimation, and some others. The comparison shows that the POME-based method is comparable to some methods and is better than others.

Entropy-Spectral Analysis for Flow Forecasting

Maximum entropy spectral analysis (MESA) was introduced by Burg (12). It has several advantages over conventional spectral analysis methods. It has short and smooth spectra that have high-degree resolutions (13). The statistical characteristics that are used in stochastic model identification can also be estimated using MESA, thus permitting integration of spectral analysis and computations related to stochastic model development. Ulrych and Clayton (14) reviewed the principles of MESA and the closely related problem of autoregressive time series modeling. Shore (15) presented a comprehensive discussion of minimum cross-entropy spectral analysis.

The relationship between spectrum $W(f)$ with frequency f of a stationary process $x(t)$ and entropy $H(f)$ can be expressed as

$$H(f) = \frac{1}{2}\ln(2\,w) + \frac{1}{4\,w}\int_{w}^{+w} \ln[W(f)]\,df \qquad (18)$$

where w is the frequency band. Equation 18 is maximized subject to the constraint equations given as autocorrelations until log m:

$$\rho(n) = \int_{w}^{+w} W(f)\exp(i_2\pi f n \Delta t)\,df, \qquad m \le n \le +m \quad (19)$$

where Δt is the sampling time interval and $i = (-1)^{1/2}$. Maximization of Eq. 19 is equivalent to maximizing

$$H(f) = \int_{w}^{+w} \ln[W(f)]\,df \qquad (20)$$

which is known as Burg entropy. The spectrum $W(f)$ can be expressed in terms of the Fourier series as

$$W(f) = \frac{1}{2\,w}\sum_{n=\infty}^{\infty} \rho(n)\exp(i2\pi n f \Delta t) \qquad (21)$$

Substituting Eq. 21 in Eq. 20 and maximizing lead to MESA.

Jaynes (16) has shown that MESA and other methods of spectral analysis, such as Schuster, Blackman–Tukey, maximum likelihood, Bayesian, and autoregressive (AR, ARMA, or ARIMA) models are not in conflict and that AR models are a special case of MESA. Krstanovic and Singh (17,18) employed MESA for long-term stream flow forecasting. Krstanovic and Singh (19,20) extended the MESA method to develop a real-time flood forecasting model. Padmanabhan and Rao (21,22) applied MESA to analyze rainfall and river flow time series. Rao et al. (23) compared a number of spectral analysis methods with MESA and found that MESA is superior. Eilbert and Christensen (24) analyzed annual hydrologic forecasts for central California and found that dry years might be more predictable than wet years. Dalezios and Tyraskis (25) employed MESA to analyze multiple precipitation time series.

Regional Precipitation Analysis and Forecasting

The Burg algorithm or MESA can be applied to identify and interpret multistation precipitation data sets and to explore spectral features that lead to a better understanding of rainfall structure in space and time (25).

Then, multistation rainfall time series can be extrapolated to develop regional forecasting capabilities.

Grouping of River Flow Regimes

An objective grouping of flow regimes into regime types can be employed as a diagnostic tool for interpreting the results of climate models and flow sensitivity analyses. By minimizing an entropy-based objective function (such as minimum cross entropy), a hierarchical aggregation of monthly flow series into flow regime types can, therefore, be effectively performed, which will satisfy chosen discriminating criteria. Such an approach was developed by Krasovskaia (26) who applied it to a regional river flow sample for Scandinavia for two different formulations of discriminating criteria.

Basin Geomorphology

Entropy plays a fundamental role in characterizing landscape. Using entropy theory for the morphological analysis of river basin networks, Fiorentino et al. (27) found that the connection between entropy and the mean basin elevation is linearly related to basin entropy. Similarly, the relation between the fall in elevation from the source to the outlet of the main channel and the entropy of its drainage basin, it was found, is linear and so also was the case between the elevation of a node and the logarithm of its distance from the source. When a basin was ordered following the Horton–Strahler ordering scheme, a linear relation was found between drainage basin entropy and basin order. This relation can be characterized as a measure of basin network complexity. Basin entropy, it was also found, is linearly related to the logarithm of the magnitude of the basin network. This relation led to a nonlinear relation between the network diameter and magnitude where the exponent, it was found, is related to the fractal dimension of the drainage network.

Design of Hydrologic Networks

The purpose of measurement networks is to gather information in terms of data. Fundamental to evaluating these networks is the ability to determine if the networks are gathering the needed information optimally. Entropy theory is a natural tool for that determination. Krstanovic and Singh (28,29) employed the theory for space and time evaluation of rainfall networks in Louisiana. The decision whether to keep or to eliminate a rain-gauge was based entirely on the reduction or gain of information at that gauge. Yang and Burn (30) employed a measure of information flow, called directional information transfer index (DIT), between gauging stations in the network. The value of DIT varies from zero, where no information is transmitted and the stations are independent, to one where no information is lost and the stations are fully dependent. Between two stations of one pair, the station that has the higher DIT value should be retained because of its greater capability of inferring information at the other side.

Rating Curve

Moramarco and Singh (31) employed entropy theory to develop a method for reconstructing the discharge hydrograph at a river section where only water level is monitored and discharge is recorded at another upstream section. The method, which is based on the assumption that lateral inflows are negligible, has two parameters linked to remotely observed discharge and permits, without using a flood routing procedure and without the need of a rating curve at a local site, relating the local river stage to the hydraulic condition at a remote upstream section.

CONCLUDING REMARKS

Entropy theory permits determining of the least biased probability distribution of a random variable, subject to the available information. It suggests whether or not the available information is adequate and, if not, then additional information should be sought. In this way, it brings the model, the modeler, and the decision-maker closer. As an objective measure of information or uncertainty, entropy theory allows communicating with nature, as illustrated by its application to the design of data acquisition systems, the design of environmental and hydrologic networks, and the assessment of the reliability of these systems or networks. In a similar vein, it helps better understand the physics or science of natural systems, such as landscape evolution, geomorphology, and hydrodynamics. A wide variety of seemingly disparate or dissimilar problems can be meaningfully solved by using entropy.

BIBLIOGRAPHY

1. Shannon, C.E. (1948). A mathematical theory of communications, I and II. *Bell System Technical Journal* **27**: 379–443.
2. Jaynes, E.T. (1957). Information theory and statistical mechanics, I. *Physical Review* **106**: 620–630.
3. Levine, R.D. and Tribus, M. (Eds.). (1978). *The Maximum Entropy Formalism*. The MIT Press, Cambridge, MA, p. 498.
4. Kullback, S. and Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics* **22**: 79–86.
5. Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Proceedings, 2nd International Symposium on Information Theory*. B.N. Petrov and F. Csaki. (Eds.). Publishing House of the Hungarian Academy of Sciences, Budapest, Hungary.
6. Tribus, M. (1969). *Rational Description: Decision and Designs*. Pergamon Press, New York.
7. Kaplan, S. and Garrick, B.J. (1981). On the quantitative definition of risk. *Risk Analysis* **1**(1): 11–27.
8. Singh, V.P. and Fiorentino, M. (Eds.). (1992). *Entropy and Energy Dissipation in Water Resources*. Kluwer Academic Publishers, Dordrecht, the Netherlands.
9. Singh, V.P. (1998). *Entropy-Based Parameter Estimation in Hydrology*. Kluwer Academic Publishers, Boston, MA.
10. Harmancioglu, N.B. and Singh, V.P. (1998). Entropy in environmental and water resources. In: *Encyclopedia of Hydrology and Water Resources*. R.W. Hershey and R.W. Fairbridge (Eds.). Kluwer Academic Publishers, Boston, MA, pp. 225–241.
11. Singh, V.P. (1998). The use of entropy in hydrology and water resources. *Hydrological Processes* **11**: 587–626.

12. Burg, J.P. (1975). *Maximum entropy spectral analysis*. Unpublished Ph.D. thesis, Stanford University, Palo Alto, CA, p. 123.
13. Fougere, P.F., Zawalick, E.J., and Radoski, H.R. (1976). Spontaneous life splitting in maximum entropy power spectrum analysis. *Physics of the Earth and Planetary Interiors* **12**: 201–207.
14. Ulrych, T. and Clayton, R.W. (1976). Time series modeling and maximum entropy. *Physics of the Earth and Planetary Interiors* **12**: 188–199.
15. Shore, J.E. (1979). *Minimum cross-entropy spectral analysis*. NRL Memorandum Report 3921, Naval Research Laboratory, Washington, DC.
16. Jaynes, E.T. (1982). On the rationale of maximum entropy methods. *Proceedings of the IEEE* **70**: 939–952.
17. Krstanovic, P.F. and Singh, V.P. (1991). A univariate model for long-term streamflow forecasting: 1. Development. *Stochastic Hydrology and Hydraulics* **5**: 173–188.
18. Krstanovic, P.F. and Singh, V.P. (1991). A univariate model for long-term streamflow forecasting: 2. Application. *Stochastic Hydrology and Hydraulics* **5**: 189–205.
19. Krstanovic, P.F. and Singh, V.P. (1993). A real-time flood forecasting model based on maximum entropy spectral analysis: 1. Development. *Water Resources Management* **7**: 109–129.
20. Krstanovic, P.F. and Singh, V.P. (1993). A real-time flood forecasting model based on maximum entropy spectral analysis: 2. Application. *Water Resources Management* **7**: 131–151.
21. Padmanabhan, G. and Rao, A.R. (1986). Maximum entropy spectra of some rainfall and river flow time series from southern and central India. *Theoretical and Applied Climatology* **37**: 63–73.
22. Padmanabhan, G. and Rao, A.R. (1988). Maximum entropy spectral analysis of hydrologic data. *Water Resources Research* **24**(9): 1591–1533.
23. Rao, A.R., Padmanabhan, G., and Kashyap, R.L. (1980). Comparison of recently developed methods of spectral analysis. *Proceedings, Third International Symposium on Stochastic Hydraulics*. Tokyo, Japan, pp. 165–175.
24. Eilbert, R.F. and Christensen, R.O. (1983). Performance of the entropy minimax hydrological forecasts for California water years 1948–1977. *Journal of Climate & Applied Meteorology* **22**: 1654–1657.
25. Dalezios, N.R. and Tyraskis, P.A. (1989). Maximum entropy spectra for regional precipitation analysis and forecasting. *Journal of Hydrology* **109**: 25–42.
26. Krasovskaia, R. (1997). Entropy-based grouping of river flow regimes. *Journal of Hydrology* **202**: 173–1191.
27. Fiorentino, M., Claps, P., and Singh, V.P. (1993). An entropy-based morphological analysis of river basin networks. *Water Resources Research* **29**(4): 1215–1224.
28. Krstanovic, P.F. and Singh, V.P. (1992). Evaluation of rainfall networks using entropy: 1. Theoretical development. *Water Resources Management* **6**: 279–293.
29. Krstanovic, P.F. and Singh, V.P. (1992). Evaluation of rainfall networks using entropy: II. Application. *Water Resources Management* **6**: 295–314.
30. Yang, Y. and Burn, D.H. (1984). An entropy approach to data collection network design. *Journal of Hydrology* **157**: 307–324.
31. Moramarco, T. and Singh, V.P. (2001). Simple method for relating local stage and remote discharge. *Journal of Hydrologic Engineering, ASCE* **6**(1): 78–81.

EVAPORATION

THEODORE A. ENDRENY
SUNY-ESF
Syracuse, New York

Evaporation of water is a solar energy driven phase change from liquid to vapor that maintains the hydrologic cycle by transferring liquid water at the earth's surface to water vapor in the atmosphere, where it may lift, condense, and precipitate to earth as liquid water. In this discussion, evaporation is defined to include the closely associated transpiration process as a subcategory. The coupled processes are often called as evapotranspiration, or ET, where transpiration focuses on the transport of liquid water through the plant roots, stem, and leaf prior to evaporation through the leave's stomata. Further, when the discussion does not make a clear distinction between potential and actual evaporation, potential should be assumed. Potential evaporation refers to the amount of water that available energy and diffusion processes can transfer into atmospheric vapor, which is typically greater then the amount actually transferred due to limits on soil water volumes and resistances in the path.

PHYSICAL CONTROLS ON EVAPORATION

Evaporation is effectively a two-step process that first, requires that water changes phase to a vapor state and second, requires that the vapor is transported by advection and/or diffusion into unsaturated air. The phase change alone does not completely satisfy the requirement for evaporation when dynamic equilibrium exists at the boundary between liquid and vapor and condensation of the vapor saturated air returns liquid water to the surface to maintain no net water loss. Transport, therefore, ensures removal of the water vapor and a net loss of heat and mass from the liquid surface. In the vapor state, evaporated water is invisible to the human eye, which detects wavelengths between 0.4 and 0.7 μm, but is detectable in other areas of the electromagnetic spectrum. As such, though clouds are derived from and contain evaporated water, they are not vapor and instead reveal condensed water droplets that geometrically scatter light.

The first step in changing the phase of water from liquid to vapor requires an input of solar energy, which is stored at the surface and supports nighttime evaporation. Nearly 52% of solar energy absorbed at the earth's surface is used for vaporization. This energy is called latent heat of vaporization, or λ, and is a function of water temperature. When the phase changes directly from frozen water to vapor, known as sublimation, a greater amount of energy is required. Phase changes require energy to separate the hydrogen bond based, attractive intermolecular forces holding water molecules in an organized pattern and close proximity. Latent heat is a name that suggests dormant, or invisible, and is used to indicate that, unlike the measurable effect of sensible heat on air temperature, the solar energy used in evaporation remains 'hidden' from thermometer measurement.

As an example, at the standard pressure of 1013.3 milibar (mb), heating liquid water from 0 to 100 °C requires 4186.8 joules (J) of heat per kilogram of water per degree C temperature change and is detectable with a thermometer. In contrast, the latent heat of vaporization converting 100 °C water into 100 °C vapor does not have a measurable impact on water temperature. The latent heat, in megajoules per kilogram ($\mathrm{MJ\,kg^{-1}}$), required per °C of water temperature, T, is given as

$$\lambda = 2.501 - 0.002361 \cdot T$$

It is apparent that the majority of heat input to evaporate water, whether at 0 °C or 99 °C, is that for the phase change. A relatively constant 1350 $\mathrm{Js^{-1}\,m^{-2}}$ stream of solar energy entering the earth's upper atmosphere provides energy to evaporate water just like a stove can boil and evaporate water. When 50% of this solar constant strikes the earth's surface after atmospheric attenuation (e.g., scattering and reflection), it requires approximately 1 hour and 10 minutes to generate the latent heat needed to evaporate 1 kg of water at 20 °C.

Latent heat, though not detectable as affecting air temperature, is stored with the vapor in greater vibrational, rotational, and translational movement of vapor molecules. The removal of this vapor from the liquid, therefore, causes a measurable loss of energy and temperature from the remaining liquid, producing a cooling effect. A familiar example of this process is that wet skin, from sweat or a shower, cools faster in moving rather than still air because the wind speeds evaporation that takes heat from the body. Knowing that heat is used for evaporation, it is now clear that a covered pot will boil faster than a counterpart uncovered pot do to a lower loss of heat to the net evaporation of water. Heat stored in vapor is later released back into the environment when the vapor vibrational speeds slow, and it condenses into water, called latent heat of condensation. When vapor passes directly to solid frozen water, called deposition, a greater amount of heat stored in rotational, vibrational, and translational molecular movement is released into the environment.

The second physical step in evaporation is transport by advection and/or diffusion, which provides the net movement of water molecules from the liquid water surface of soil, plants, or lakes to atmospheric vapor. Vapor and wind gradients exert the principal controls on removal of water vapor beyond the saturated layer of air that maintains condensation–evaporation equilibrium dynamics. Fundamental barriers to transport beyond the layer of dynamic equilibrium include stagnant air and saturated air above the evaporating surface, conditions readily created within and at the surface of soil and plant systems. Net evaporation therefore increases with steeper wind and saturation gradients, which are defined as the change in wind or saturation with distance above the evaporating surface. Work on fluid velocity and turbulence gradients in the mid-1900s by Prandtl and von Karman has been used to estimate momentum, sensible heat, and vapor transport from wind speed measurements.

Estimates of atmospheric wind and vapor conditions above the evaporating surface provide important data for estimating wind and saturation gradients and predicting barriers to vapor transport and net evaporation. Meteorological stations are frequently equipped with anemometers and thermometers at 2 meters (m) above the ground surface to help establish the wind and vapor gradients controlling evaporation. Wind profiles, it is assumed, begin with stagnant air at the no-slip boundary, or zero-plane displacement height, and increase logarithmically. In a landscape broken by tree canopies, a 2 m wind measurement may be inadequate to represent observations. Research has shown that dynamic turbulence and eddies created by such forested heterogeneity result in increased wind and evaporation rates that exceed the estimated atmospheric potential.

Based on Dalton's work on individual pressures of multiple atmospheric gases summing to the observed atmospheric pressure, vapor is often reported as a partial pressure and can be derived from measurements of temperature. In the following equation, vapor pressure is reported in kilopascals (kPa) and temperature in °C. Initially dry warm air can absorb more water than initially dry cold air before reaching saturation.

$$e = 0.6108 \exp\left(\frac{17.27\,T}{237.3 + T}\right)$$

The dry-bulb temperature is used to estimate the total amount of vapor the air could absorb prior to saturation; the dew point temperature represents the temperature to which the air must cool for total saturation. When the dry-bulb and dew point temperatures are equal, the air is fully saturated. The ratio of actual to saturated vapor pressures is the relative humidity. The dew point temperature is estimated by using a psychrometer that measures the difference between dry-bulb and wet-bulb thermometers, called the wet-bulb depression, together with lookup tables relating wet-bulb depression to dew point temperature. The dry-bulb is a normal thermometer measuring air temperature, but a moist piece of cloth typically covers the wet-bulb, and evaporation of the water from the cloth causes the temperature to drop. Chilled mirrors hygrometers, hair hygrometers, and vapor pressure sensors are also used to detect the vapor content in the air.

EVAPORATION MEASUREMENT AND ESTIMATION

Evaporation is fundamental to both energy and water balances, yet despite the importance of evaporation to hydrologic assessment of the paths, quantities, and quality of water in the lithosphere, biosphere, and atmosphere, the complexity of the process has prevented easy or exact techniques for measuring and estimating it. The relative accuracy of yearly river basin evaporation estimates is high, as the estimate time frame and spatial area become smaller, but the simple application of energy and water balance models to solve for evaporation becomes less tractable. A variety of measurement and estimation techniques have been developed for these smaller scales, such as hourly, daily, and monthly evaporation from reservoirs, farm fields, and single plants.

In general, evaporative fluxes from the land surface are more difficult to measure than from open water, given that an immeasurable number of irregular, tiny and unique soil and leaf surfaces are involved in this phase change and that suction gradients draw water to this evaporating interface from unobserved reserves of unknown volume. Fluxes from open water, though relatively homogeneous, still provide challenges when subsurface inflows and outflows are poorly understood and significant and wind and water advected energy influencing evaporation is heterogeneously distributed. Hence, the numerous methods developed for estimating evaporation are categorized based on the type of surface, availability of water, and the importance of stored energy, water-advected energy, and air-advected energy.

Actual evaporation can be measured by using a water balance approach, a turbulent-transfer approach, a potential evaporation approach, or a water quality approach. The water balance approach can function with measurements of a mass balance being kept for a water pan, such as a the Class-A Pan of the National Weather Service, a soil and plant system, such as in a weighing lysimeter, or of a small, enclosed atmosphere. Turbulent-transfer methods, which derive evaporation from estimates of momentum or heat flux, can provide estimates for larger heterogeneous areas, but assume that the air sampled by the field instrumentation used for the Bowen ratio or eddy-correlation method is well mixed to represent the upwind land area by a given fetch. Water quality methods include techniques that track concentrations of dissolved solids, which enrich when evaporation removes the water solvent, and isotope tracer studies that show heavier isotopes are enriched by preferred evaporation of lighter isotopes. Sap flow monitoring in trees provides another technique to measure the flux of water from the ground to the atmosphere.

Mathematical estimates of evaporation rates have been approached by using an equally wide variety of techniques and include temperature-based, aerodynamic-based, radiation-based, and combination-based methods. Temperature-based methods, such as the monthly time-step Thornthwaite equation, use air temperature and length of day, as well as an assumed humidity, to compute the potential evaporation, and they have been adapted to suit several different climates and regions. Aerodynamic methods assume that solar radiation is not limiting and consider only wind speed and turbulence as controls on the transport of water vapor away from the surface. Radiation-based methods likewise assume that wind turbulence and eddies are not limiting and use measured incoming radiation and the latent heat of vaporization to compute evaporation flux. A popular form of this equation is the Priestly–Taylor, which increased it by a factor of 30% to account for added aerodynamic transfer. The combination method, known most extensively for the Penman–Monteith equation, uses air temperature, net radiation, wind speed, and relative humidity vapor gradients to derive minute by minute and daily evaporation rates.

HYDROLOGIC IMPACTS OF EVAPORATION

Observations of terrestrial river discharge reveal that more water precipitates on land than evaporates from land and that more water evaporates from oceans than precipitates on oceans. Precipitation totals may vary from year to year, but evaporative demands are rather steady, which creates a greater relative fluctuation in river discharge than in precipitation. This is illustrated by considering a 20% decrease in annual precipitation from 100 to 80 cm, where 50 cm went to evaporation in both years, and discharge dropped by 40% from 50 to 30 cm. Evaporation and its impact on liquid and vapor water volumes and the partitioning of solar energy into latent and sensible heat create and maintain a range of climatic conditions, from microclimates on the scale of a tree canopy to macroclimates that describe the global distribution of plants.

The volume of water evaporated from the ocean and land surface is greatest at the meteorological equator, or intertropical convergence zone (ITCZ), and smallest at the poles, which is the result of a similar longitudinal distribution of insolation intensity. Sinking Hadley cell air at the 30° latitude belts, which warms to absorb greater amounts of water vapor, is the cause of a belt of deserts in this region. The distribution of incoming solar radiation, which is greatest at the equator and smallest at the poles, is the driving force explaining the meridional (across lines of latitude) distribution of evaporation. Wind transport of this evaporated vapor from the equatorial region to the midlatitudes and poles, where latent energy is released to the atmosphere as sensible heat during condensation, is one of only a few processes that help to maintain the earth's energy balance.

Global water balance numbers reveal that a relatively small volume of evaporated vapor resides in the atmosphere. The earth's atmosphere has a volume of 12,900 km^3, and contains just 0.001% of all global water. As evaporated vapor, it receives 71,000 $\text{km}^3\,\text{yr}^{-1}$ from land, 1000 $\text{km}^3\,\text{yr}^{-1}$ from lakes, and 505,000 $\text{km}^3\,\text{yr}^{-1}$ from oceans, and this flux rate into its total volume equals a residence time of 8.2 days, or just over a week before evaporated water precipitates. The atmospheric vapor precipitates at a volumetric rate of 577 $\text{km}^3\,\text{yr}^{-1}$, of which 119,000 $\text{km}^3\,\text{yr}^{-1}$ falls on land. Observation and estimation of river discharge at 47,000 $\text{km}^3\,\text{yr}^{-1}$ was used to deduce the amount evaporated from land, which is 61%

Table 1. Continental Average Estimated Evaporation

Continent	Area, km^2	Evaporation, $\text{mm}\,\text{yr}^{-1}$	Evaporation, %
Antarctica	14,100,000	28	17
Europe	10,000,000	375	57
Asia	44,100,000	420	60
South America	17,900,000	946	60
North America	24,100,000	403	62
Africa	29,800,000	582	84
Australia	7,600,000	420	94
Total Land	148,900,000	480	64

Table 2. Average Estimated Evaporation Based on Studies of Precipitation and Runoff

Watershed by River Name	Continent [Nation(s)]	Basin Size, km^2	Evaporation^a, %
Brahmaputra	Asia (Tibet/Bangladesh)	589,000	35
Irrawaddy	Asia (Burma)	431,000	40
Yangtzekiang	Asia (China)	1,970,000	50
Amazon	South America (six nations)	7,180,000	53
Orinoco	South America (Venezuela)	1,086,000	54
Lena	Asia (Russia)	2,430,000	54
Mekong	Asia (China)	795,000	57
Yenesei	Asia (Russia)	2,599,000	58
Ganges	Asia (China)	1,073,000	58
Saint Lawrence	North America (Canada, U.S.)	1,030,000	67
Amur	Asia (Russia)	1,843,000	68
Congo	Africa (7 nations)	3,822,000	75
Ob	Asia (Russia)	2,950,000	76
Mississippi	North America (U.S.)	3,224,000	79
La Plata	South America (five nations)	2,650,000	80
Average	–	2,224,800	60

[a] Evaporation may include basin transfers because it was not directly measured but derived from precipitation and runoff measurements.

evaporating, and this differs from the continental average of 64% reported in Table 1.

For the continental United States, which contains deserts in Arizona and rain forests in Washington, the average annual percentage of precipitation converted to evaporation is approximately 62%. This value is similar to global patterns but varies considerably from that measured on other continents (see Table 1) and larger watersheds (see Table 2).

The agricultural impact of evaporation is both the cause of nutrient uptake and growth in plants, as well as the loss of soil water and plant stress. Maintenance of optimal water levels in the soil, called field capacity, when gravitational water has drained, is the goal of many irrigation projects. If irrigation causes evaporation to exceed local precipitation, then salts will be drawn to the soil surface, which often creates osmotic gradients at the root interface that kill the agricultural crop. Agricultural irrigation to satisfy the high evaporation demand of sunny agricultural land, such as California's Central Valley, has become a direct competitor for use as a public water supply.

READING LIST

Shuttleworth, W.J. (1992). Evaporation. In: *Handbook of Hydrology*. D.R. Maidment (Ed.). McGraw Hill, New York.

Dingman, S.L. (1994). *Physical Hydrology*. Prentice-Hall, Englewood Cliffs, NJ.

EVAPOTRANSPIRATION

JOSE O. PAYERO
University of Nebraska-Lincoln
North Platte, Nebraska

Liquid water from a surface can be transformed into water vapor by either evaporation or by transpiration. *Evaporation* is the process for converting liquid water to water vapor and removing it from the evaporating surface. *Transpiration* is the vaporization of water contained in plant tissues and the removal of vapor to the atmosphere through leaf stomata (1,2). Evaporation and transpiration occur simultaneously in cropped surfaces, and there is no easy way of quantifying the magnitude of each component. For practical applications, such as irrigation scheduling and irrigation system design, it has been more useful to consider both processes combined. The combination of these two processes is called *evapotranspiration*. The proportion of each component in a cropped surface is affected by factors such as vegetative cover, available water in the soil, and surface wetness. For an annual crop like corn, evaporation is the dominant component of evapotranspiration at the beginning of the season when the soil surface is exposed to solar radiation. As the crop grows and the canopy covers the surface, evaporation is minimized, and transpiration becomes the dominant component.

IMPORTANCE OF EVAPOTRANSPIRATION

Evapotranspiration is an important process in agriculture and other natural sciences, such as hydrology, because it is an important component of the hydrologic cycle. It represents the water that is effectively lost from the earth's surface, and can no longer be controlled by humans. This type of water loss is often called consumptive use. Other types of processes that usually cause water losses from a given area on the earth's surface, such as runoff and deep percolation, do not involve a change in the physical state of water and therefore, water can still be controlled to some degree by humans.

Plants use water as a solvent and transport mechanism for nutrients and other chemicals, as a reagent for the chemical reactions involved in their physiological processes (such as photosynthesis), and as a component of cell cytoplasm, which allows plant tissues to stay turgid. Most of the water consumed by plants, however, is used in evapotranspiration. Evapotranspiration has the important function of regulating the temperature of plants, keeping them cool within a temperature range that favors growth. When the water supply in the soil is limited, for instance, plants respond by closing their stomata. This restricts the rate of evapotranspiration, and the temperature of the canopy tends to increase (3). This increase in canopy temperature has been used to estimate the rate of evapotranspiration of crops and as a way to detect crop water stress for irrigation scheduling (4). Because most of the water consumed by plants is lost in evapotranspiration, it takes a considerable amount of water for a crop to produce one unit weight of dry matter, as shown for different crops in (Fig. 1).

Researchers have shown that crop yield is often linearly related to crop evapotranspiration, up to the point where yield is limited by factors other than water (6). Therefore, if evapotranspiration is limited, yield is usually reduced. For this reason, in regions where rainfall is not sufficient to provide enough water for crops to keep evapotranspirating at a nonlimited rate, irrigation is required to obtain adequate crop yields. The nonlimited rate of evapotranspiration, however, can also be maintained by applying excess water. Application of excess water, however, has been linked to undesired side effects such as drainage problems, salinization of soils, soil erosion, and pollution of surface and groundwaters. In places where irrigation water needs to be pumped, pumping excess water also represents higher production cost. Therefore, it is considered ideal to schedule irrigation according to crop water needs. This requires, among other things, good knowledge of crop evapotranspiration rates.

MEASURING EVAPOTRANSPIRATION

Measuring the rate of evapotranspiration of crops and other surfaces is complex and is a subject that has attracted considerable research. Many methods have been devised to measure and estimate evapotranspiration. *Methods for measuring* evapotranspiration include the use of lysimeters, scintillometers, micrometeorological techniques such as the Bowen ratio and eddy covariance methods, and measurement of sap flow. Methods to measure evapotranspiration usually require expensive instrumentation and specialized training, and therefore have been confined mostly to research applications. Because of this, considerable effort has been devoted to develop methods for estimating evapotranspiration.

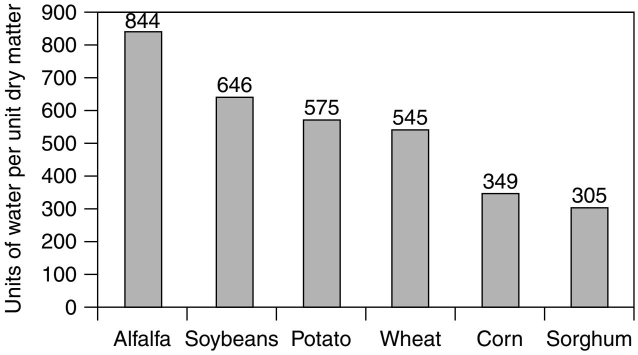

Figure 1. Unit weight of evapotranspirated water needed per unit weight of dry matter produced by different crops (adapted from Ref. 5).

ESTIMATING EVAPOTRANSPIRATION

Methods for estimating evapotranspiration include the use of different kinds of devices and a variety of equations. Devices such as evaporation pans, atmometers, and soil moisture monitoring devices have long been used to estimate evapotranspiration. The use of equations, however, is by far the most common method to estimate evapotranspiration. A great variety of equations have been proposed though the years, ranging from very simple to very complex models (1,7).

Simple models usually try to estimate evapotranspiration based on empirical relationships involving one or several meteorological variables. Complex models consider the physics governing the evapotranspiration process and try to include all factors that significantly contribute to the process (1,8). Complex models can be further divided into single-layer models and multiple-layer models. Single-layer models, such as the Penman–Monteith model, consider the crop canopy as a “big leaf,” taking into account only processes that occur between the top of the canopy and the atmosphere. Multiple-layer models, on the other hand, also take into account those processes that take place below the crop canopy. Multiple-layer models are more theoretically sound than single-layer models, but their complexity makes them impractical for widespread use. Because a single-layer model is sufficiently accurate for most practical applications and relatively simple to apply, it has been proposed as the recommended method for estimating evapotranspiration (9).

Considerable effort has been made to estimate evapotranspiration using inputs obtained from remote sensing platforms, such as satellites or airplanes (1,10,11), and others have even tried to estimate evapotranspiration by measuring the flux of stable isotopes (12). Evapotranspiration, however, is more often estimated from equations that use meteorological data as input, as well as inputs that describe the characteristics of the evaporating surface. This is a convenient method because meteorological data are readily available in most places. Most nations and states support a network of meteorological stations and offer this information to the public in various ways. Meteorological data commonly used to estimate evapotranspiration include solar radiation, air temperature, relative humidity, and wind speed.

A detailed procedure for calculating evapotranspiration has been described by Allen et al. (1). The method involves a two-step process. One step consists of calculating the evapotranspiration rate of a reference crop, either clipped grass or alfalfa, which is usually known as reference evapotranspiration. In older literature, this was also called potential evapotranspiration. It represents a measure of the evaporating demand of the atmosphere for a short, cropped surface that effectively covers the ground, is growing healthily, and is not short of water, that is, a condition in which transpiration is not limited by stress, and the evapotranspiration demand of the atmosphere is met. The second step involves adjusting the reference evapotranspiration to match the conditions of the specific surface or crop being considered. This is done by calculating an adjustment factor, usually known as the crop coefficient. Depending on the accuracy required, calculating crop coefficients can also be a simple or complex process (7,13). Multiplying the reference evapotranspiration by the crop coefficient then results in the evapotranspiration rate for the crop or surface in question.

REQUIREMENTS FOR EVAPOTRANSPIRATION

Procedures used to estimate evapotranspiration try to simplify the complexities of the physical and physiological processes that affect evapotranspiration to a manageable number of quantifiable variables. They try to recognize that for the evapotranspiration process to take place, it is necessary to have

1. energy
2. water
3. space in the atmosphere to hold the water vapor
4. a transport mechanism for the water vapor to move from the surface to the atmosphere.

Evapotranspiration is an energy-driven process. It takes approximately 2.45 megajoules of energy to evaporate 1 kilogram of water at 20 °C. The sun supplies the energy needed for evapotranspiration from the earth's surface. Part of the solar energy that reaches the evaporating surface, however, is reflected back to the atmosphere and cannot be used for evapotranspiration. Of the energy that stays on the evaporating surface, known as net radiation, not all is used in evapotranspiration. The energy balance of a surface also includes energy that is absorbed or released by the soil (soil heat flux), by the air (latent heat flux), and that used in evapotranspiration (latent heat flux). All of these types of heat fluxes take place simultaneously from a given surface, and their proportions depend on the characteristics of the surface and weather conditions.

The amount of energy used for evapotranspiration includes short-wave radiation that comes directly from the sun, long-wave radiation or heat that comes from the surface, and advective heat that is transported horizontally by hot wind to the evaporating surface. The amount of energy available for evapotranspiration varies with latitude, day of the year, time of day, atmospheric conditions, and the characteristics of the surface itself. Figure 2 shows how the theoretical clear-sky solar radiation for different latitudes varies throughout the year. Figure 3 shows how measured solar radiation varies during the day and the effect of cloudiness in reducing the amount of solar energy that reaches the surface. As a general rule, the higher the amount of energy available at the evaporating surface, the higher the evapotranspiration rate, assuming that all other conditions required for evapotranspiration are met.

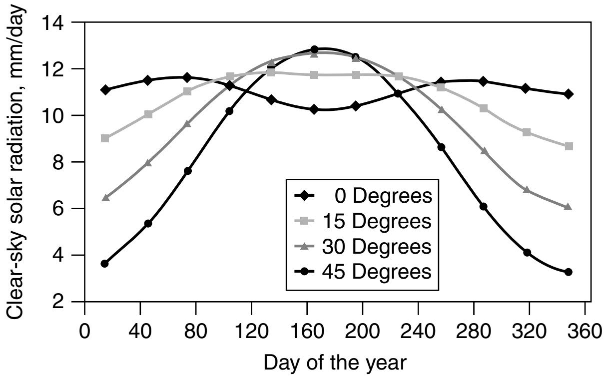

Figure 2. Theoretical clear-sky solar radiation values for different northern latitudes and day of the year at sea level.

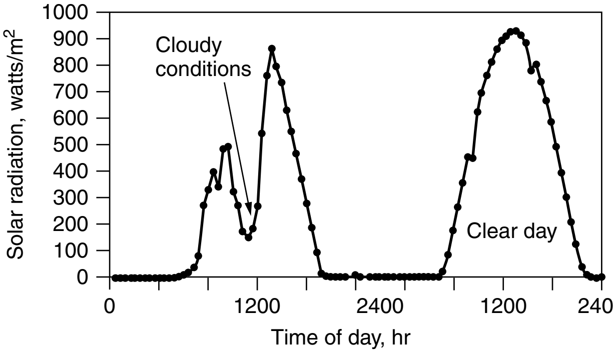

Figure 3. Measured solar radiation in North Platte, NE, on July 15 and 16, 2001.

Water is the next essential requirement for evapotranspiration. If water is not available, the energy available to evaporate water is then used to heat the air and the soil. Under desert conditions, for instance, where energy is usually plentiful, but water is normally very limited, the evapotranspiration rate can be very small. Most of the energy is converted to sensible heat flux, soil heat flux, and a very small portion or none of the energy is converted to latent heat flux. Similarly, when crops are under water stress, the evapotranspiration rate is reduced.

Space in the atmosphere to hold the water vapor is also needed for evapotranspiration to proceed. When water is evaporated from a surface, the water vapor will travel to the atmosphere where it will be stored, provided that the atmosphere is not saturated. If the atmosphere is already saturated, it will not be able to store the additional water vapor, and evapotranspiration will be restricted. The drier the air in contact with the evaporating surface, the more evapotranspiration is enhanced. For this reason, the humidity of the air, usually expressed as relative humidity, is an important factor to consider when estimating evapotranspiration.

The last requirement for evapotranspiration is a transport mechanism for moving water vapor from the surface to the atmosphere. There are two basic transport mechanisms for water vapor. The first is turbulence.

Turbulence is created when air moves horizontally over a rough surface. The friction created by the contact of the air with the rough elements of the surface creates a vertical component of the wind speed, which creates eddies of different sizes. The size of these eddies depends on the roughness of the surface and on the magnitude of the wind speed. These eddies carry the water vapor to the atmosphere. The other transport mechanism is buoyancy. Hot air is less dense that cold air; as the air close to the surface becomes hotter than the air above, it tends to ascend, carrying with it the water vapor. Turbulence is the most important transport mechanism in most instances, so wind speed is a very important factor in determining evapotranspiration. The higher the wind speed and the rougher the surface, the more turbulence, and the higher the evapotranspiration rate will be.

MAGNITUDE OF EVAPOTRANSPIRATION

Evapotranspiration rates are most commonly expressed in units of water depth per unit time, such as millimeters per day (mm day^{-1}) or inches per month (in months^{-1}). Since it takes energy to evaporate water, water depths can also be expressed in terms of energy received per unit area. Therefore, evaportranspiration is often expressed in units of energy per unit area per unit time, such as watts per squared meter (w m^{-2}), or megajoules per squared meter per day (MJ m^{-2} day^{-1}). It can, however, also be expressed in units of energy per unit area, such as watts m^{-2}, or MJ m^{-2} day^{-1}. Many of the factors affecting evapotranspiration are so dynamic that the magnitude of the evapotranspiration rates for a given surface will vary from day to day, from place to place, and throughout the day. Figure 4 shows the calculated daily evapotranspiration rate for corn in North Platte, Nebraska, during the 2000 growing season. It shows the typical large variations in evapotranspiration rate that can be expected from day to day, as a result of normal daily changes in weather conditions. It also shows a seasonal pattern, as a response to the seasonal changes in available energy and to the changing water demand of the crop during its growing cycle.

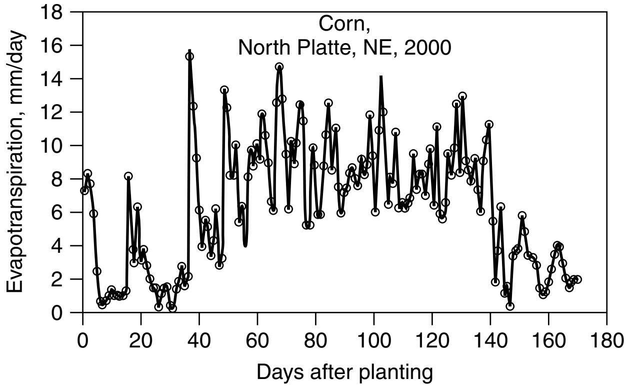

Figure 4. Calculated daily evapotranspiration for corn in North Platte, NE, during the year 2000.

BIBLIOGRAPHY

1. Allen, R.G., Pereira, L.S., Raes, D., and Smith, M. (1998). *Crop evapotranspiration—guidelines for computing crop water requirements*. Irrigation and Drainage Paper No. 56, FAO, Rome, Italy.
2. Allen, R.G. et al. (2002). Evapotranspiration from a satellite-based surface energy balance for the Snake Plain Aquifer in Idaho. *Proceedings of the 2002 USCID/EWRI Conference*. July 9–12, San Luis Obispo, CA, pp. 167–178.
3. Gates, D.M. (1964). Leaf temperature and transpiration. *Agronomy Journal* **56**(3): 273–278.
4. Yazar, A., Howell, T.A., Dusek, D.A., and Copeland, K.S. (1999). Evaluation of Crop Water Stress Index for LEPA irrigated corn. *Irrigation Science* **18**: 171–180.
5. Hoeft, R.G., Nafziger, E.D., Johnson, R.R., and Aldrich, S.R. (2000). *Modern Corn and Soybean Production*, 1st Edn. MCSP Publications, Champaign, IL.
6. Schneekloth, J.P. et al. (1991) Crop rotations with full and limited irrigation and dryland management. *Transactions of the ASAE* **34**(6): 2372–2380.
7. Doorenbos, J. and Pruitt, W.O. (Eds.). (1977). *Crop Water Requirements*. Irrigation and Drainage Paper No. 24, FAO, Rome, Italy.
8. Suttleworth, W.J. and Wallace, J.S. (1985). Evaporation from sparse crops—an energy combination theory. *Quarterly Journal of the Royal Meteorological Society* **111**: 839–853.
9. EWRI. (2001). *The ASCE Standardized Reference Evapotranspiration Equation*. Environmental and Water Resource Institute of the American Society of Civil Engineers. Standardization of Reference Evapotranspiration Task Committee Report. Available: http://www.kimberly.uidaho.edu/water/asceewri/.
10. Moran, M.S. et al. (1989). Mapping surface energy balance components by combining Landsat Thematic Mapper and ground-based meteorological data. *Remote Sensing of Environment* **30**: 77–87.
11. Payero, J.O. (1997). *Estimating Evapotranspiration of Reference Crops Using the Remote Sensing Approach*. Ph.D. Dissertation, Utah State University, Logan, UT.
12. Yakir, D. and Sternberg, L.daS.L. (2000). The use of stable isotopes to study ecosystem gas exchange. *Oecologia* **123**: 297–311.
13. Wright, J.L. (1982). New evapotranspiration crop coefficients. *Proceedings of the ASCE* **108**: 57–75.

FOG

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

Fog is a cloud that materializes near the ground or over water. It is a reaction that occurs when the temperature near the ground cools to the temperature required to produce dew and causes the water vapor in the air to become visible in the form of a cloud of precipitation. There are different types of fog, which occur when different variables are involved. The two conditions necessary for fog formation are mild or no winds and air temperatures that are equivalent to the temperature at the dew point. Mild or no winds are important because obstructions in the way of a strong wind can cause it to blow in an up and down motion, which brings warmer air down and pushes the colder air up. Cool air is a requirement for fog formation so wind prevents it from happening.

Fog, like most weather hazards, can cause serious problems for society. The only time this condition is not applicable is when fog forms over water due to the absence of obstructions. In this case, the reverse reaction occurs, and the fog worsens as the wind grows faster. Fall nights are often said to contain the perfect conditions for producing fog. At night, the ground cools which also cools the air directly above it. This reaction causes droplets of water to become suspended in the air resulting in the creation of fog. However, these perfect conditions don't last forever. By morning, the heat from the sun begins to warm the ground, and the fog evaporates.

There are many types of fog, but the four most common are evaporation fog, upslope fog, precipitation fog, and radiation fog which is commonly called ground fog. Evaporation fog is found mostly out at sea, which is why it is also known as "sea fog." It is the reaction of moist air moving over colder water. The fog conditions worsen as the wind speed increases. Upslope fog is produced by moist air that is sent by a strong wind up into mountainous regions. Precipitation fog is created when precipitation in the form of rain or snow hits drier air and causes water vapor to materialize instantly. Radiation fog is seen at night when the ground is cool. The air that comes in contact with the ground also becomes cool and creates water vapor which soaks the ground and causes the formation of dew.

Like most weather hazards, fog can have severe impacts on society. Fog can cause health problems in polluted areas because the water vapor produced can become acidic. Driving through a thick cloud severely reduces a person's visibility and is a major contributor to car accidents. Even worse than the occasional car accident due to fog are the aircraft and boating accidents that can occur. Throughout history, fog has been known as the silent murderer that has taken hundreds of lives at sea in catastrophic boating accidents.

READING LIST

Joseph, P. *Weather Wisdom.* (2002). Available: http://www.touch-tmj4.com/4weather/wxwisdom/fog/fogtypes.asp. (March 19).

Malan, J. *Weather Wisdom.* (2002). Available: http://www.touch-tmj4.com/4weather/wxwisdom/fog/fog.asp. (March 19).

Malan, J. *Weather Wisdom.* (2002). Available: http://www.touch-tmj4.com/4weather/wxwisdom/fog/fogdisasters.asp. (March 19).

Parsons, L. *Fog.* (2002). Available: http://www.activeangler.com/articles/safety/articles/lee_parsons/fog.asp. (March 19).

Understanding Clouds and Fog. USA Today. (2002). Available: http://www.usatoday.com/weather/wcloud0.htm. (March 19).

Williams, J. *Fall nights often perfect for forming fog. USA Today*. (2002). Available: http://www.usatoday.com/weather/wrfog.htm. (March 19).

COASTAL FOG ALONG THE NORTHERN GULF OF MEXICO

Timothy Erickson
National Weather Service New Orleans/Baton Rouge Forecast Office
Slidell, Louisiana

INTRODUCTION

Coastal fog is a major problem for all traffic along the United States' coastlines. Lives and large monetary losses have occurred because of coastal fog. These losses have been realized over land, in the air, and on the water. Huge strides have been taken to understand and combat the coastal fog forecasting problem by the National Weather Service. Studies and research continue to improve these fog forecasting techniques. Tens of thousands of people use the Gulf of Mexico as their home or as part of their occupation in many different ways, and billions of dollars in products and property are carried and moved through the channels and river systems to and from the Gulf of Mexico each year. All are affected by coastal fog many times throughout the fall, winter, and spring. This research project took place along the northern Gulf of Mexico from the upper Texas coast to the Mississippi coastline.

METHODOLOGY

The synoptic and mesoscale patterns used in this research were from the fall, winter, and spring of 1998, 1999, 2000, and 2001. Fog was defined as water droplets suspended in the air reducing visibility to one half-mile or less. Shower and thunderstorm activity reducing visibility to these levels was not used. No records exist for fog contributions to the hydrologic cycle. Water contributions caused by fog are dismissed as false tips in rain gauges. The reference to "boundary layer" in this article will be the layer of atmosphere from the surface to the base of the lowest inversion.

Variables used in this research were surface pressure, rainfall, moisture advection, wind direction, wind speed, water temperatures, ambient temperature, and ambient dew point temperature from land observations and production platforms in the northern Gulf of Mexico. Parameters used from previous studies by Johnson and Graschel (1) were air temperature, dew point temperature, wind direction, wind speed, ceiling heights, and visibilities from oil and gas production platforms at an average altitude of 35 ms over the northern Gulf of Mexico. Gulf of Mexico sea surface temperatures were provided by the Tropical Prediction Center oceanographer in Miami, Florida.

SCALES OCCURRING WITH COASTAL FOG

Mesoscale—Widespread

Horizontal range would normally be from 50 to several hundred miles.

Microscale—Locally

Horizontal range could be at a point close to 50 miles.

CONDITIONS THAT HELP PRODUCE COASTAL FOG

Radiative

Heat is transmitted via long wave radiation (infrared) away from Earth. This radiation causes nocturnal cooling of the ground and subsequent atmospheric boundary layer. Radiational cooling is responsible for developing a stable layer near the surface by cooling the layer of air in contact with the ground while leaving a relatively warmer layer immediately above this cooler layer. This stability by warmer air above cooler air with increasing altitude is called an inversion. This feature is important in forecasting the depth of coastal fog. As the radiative process continues, the layer will cool further until saturation is achieved and fog can be produced, given the pressure does not change much during the process.

Coastal Interface. Temperature and moisture gradients are strong where the ocean meets land, especially at night during the fall, winter, and spring. During prime radiational cooling processes, this interface can become moist and stable horizontally from land to water when there is no forcing. These conditions were found when a thin ribbon of fog developed at this interface of land and water. The fog stretched linearly for more than 100 miles and vertically for hundreds of feet. Studies on how this type of fog develops are ongoing but horizontal skew-t sounding-like profiles may be a tremendous help in what is occurring at this land–water junction when this fog forms. No horizontal sounding-like profiles exist, and this is the first mention of such a profile. But if they were available, they could give some valuable insight not only into the coastal "ribbon" fog formation but also where thunderstorms are more likely to develop along with many other variables.

Frontal

This is another inversion producing process. Coastal fog can be developed with the assistance of cold, warm, or stationary fronts.

Cold Frontal. An inversion develops by cooler air displacing warmer air at the surface. The interface where these air masses meet in the vertical is known as a frontal inversion.

Coastal fog formation under a frontal inversion is nonexistent with fast-moving cold fronts, because of stronger winds that cause a deeply mixed layer, along with dry, cold air advection behind the front.

When warm air is lifted by a cold front, it causes displacement and mid-altitude air ahead of the front becomes subsident. This process causes compressional warming, which forms another inversion well ahead of the cold front. As the surface cools at night, the boundary layer becomes stable well ahead of the front. If there has been rain ahead of the cold front with a prefrontal trough, there will be plenty of moisture and a strong inversion that may develop some dense coastal fog. Even without a prefrontal trough or rainfall, the subsidence inversion and gulf moisture ahead of a slow-moving cold front will sometimes be enough to provide conditions promoting dense coastal fog.

Warm Frontal. Relatively cool saturated or nearly saturated air ahead of a warm front is gently displaced by warmer air. The cool air is normally not very deep (100–300 ms) and is topped by warmer air at relatively low altitudes, which causes an inversion to form where the two air masses meet in the vertical. Cooler air cannot hold as much moisture as warm air. As the air ahead of the warm front is already moisture laden, cooler, and stable, water vapor condenses to small particles, which causes fog to form. This result happens frequently along the coast during the cooler months of the year and sometimes far inland. A similar scenario along the coast was also described by Hsu (2). The dynamics of warmer, higher dew point air flowing over cooler waters was described by Kotsch (3) and Mullan (4).

Coastal fog can occur behind warm fronts but is not as common. The air is, by definition of a warm front, warmer than the air ahead of it. The warmer air expands and therefore can hold more water vapor, which will normally dissipate any fog behind a warm front if no overwhelming positive moisture flux is occurring in the warm sector. When this flux occurs, it can be far too great for even the warm air to hold and the water vapor will condense, which causes fog to form. During research, this result occurred only when an inversion was still present behind a warm front.

Stationary and Slow Moving Fronts. These types of fronts are aggressive at developing coastal fog, which can occur in two ways ahead of a cold front that becomes stationary or slow moving. One is when moisture is not displaced and wind speeds are very light behind the trough preceding the cold front. The second is when the wind fetch is well over the marine environment and it brings warm moist air back over cooler waters or land. Warm frontal fog conditions were explained above.

Two unusual fog days occurred as post-cold frontal events. A cold front passed through the southern Louisiana coastal region, and dew points and dry bulb temperatures cooled. Coastal fog then began to develop over many sites from east central Texas to south central Louisiana. The air mass change was more negative for the dry bulb temperature than for the dew point temperature, and the cold front slowed from about 10 knots to less than 5 knots as it reached the coastline. This decrease caused the air behind the front to become saturated. Frontal forcing weakened, and consequently northerly winds weakened and created perfect conditions for coastal fog development. Post-cold frontal fog is rare, but when it does occur, visibilities can plummet to less than one quarter of a mile quickly.

Hsu's (2) description on frontal fog production is similar to this research. Johnson and Gracshel (1) called frontal fog "mixing fog."

Maritime (Sea Fog)

When fog develops via any process over any large marine area, it is called marine or sea fog. Marine fog may be stationary or be advected elsewhere. It is during the advection process that coastal fog occurs.

Dormant. As the cooler boundary layer air over land "drains" offshore, it causes stability to decrease, which makes mixing possible. The warm gulf gives off water vapor to the cooler air, which causes the vapor to condense, making water droplets. This process quickly deteriorates when only little "drainage" is occurring. Latent heat released because of condensation warms the marine boundary layer too fast, which makes the stability in this layer neutral. Fog will form in the near shore waters if this latent heating can be exhausted through the top of the maritime boundary layer. Fog development caused by this process does not occur often. When it does, the fog stays near the coast, rarely moving anywhere because of the very slow, almost laminar flow of air moving off land. When this type of fog affects the shoreline, it is also called coastal fog. The other processes mentioned above may produce nonadvective marine fog that stays in the near shore waters.

Advancing. This process in which coastal fog is achieved is the most common. The air in the boundary layer is well mixed but does not entrain air from above the inversion. Fog that forms either onshore or offshore by any process and moves elsewhere is an advective fog. When fog advects from one offshore location to another or to the coast, it is said to be advancing. The most common type of advancing fog is found with post-strong cold front events. Winds from these fronts move the top layer of warm shelf water away from the coast. This water is replaced by cooler upwelled water. A strong temperature gradient exists where the cooler upwelled water meets the warm water offshore. When the next cold front produces return flow from the warmer water offshore, moisture-laden air saturates while moving over the cooler water. Extensive areas of fog or very low cloud ceilings form before moving inland, which causes extremely low visibilities stretching for miles along the coast. When the moisture came in the form of low clouds during research, the cloud deck would often descend to the surface producing very dense coastal fog as well. As warm air overlays cooler water, it produces a stable boundary layer. No vertical mixing occurs by air moving up or down. Mixing was produced solely by wind advancing the fog. The wind did not reach through the inversion layer, and therefore, no entrainment occurred.

A similar scenario was recorded by Binhau (5) and Hsu (2). Their contributions showed this environment to be stable, with surface winds from the southeast through southwest. This scenario holds true for all locations along the northern Gulf of Mexico.

Advective

Anytime fog forms elsewhere and is forced to another location, it is a type of advection fog. A well-mixed boundary layer also exists with advective fog, but again, air from above the inversion is not mixed downward. In some cases, the atmosphere will be capable of supporting fog but it cannot produce it. When this happens, fog may be brought into the area by wind or by the slow movement of an entire layer of air. It is important that this air mass does not mix with air from above the inversion. If this mixture occurs, dry, relatively warmer air will be mixed into the boundary layer, causing the fog to erode.

Conglomerate of Two or More Types from Above

This fog producer is the most common. Two or more of the above processes usually develop fog that forms almost anywhere on Earth. One major variable may occur at the time of fog formation, but most of the time fog is supported by another equal or weaker variable, for example, frontal-induced marine fog. That is, fog develops over the marine environment by a cold front, which brings warm, moist air back over cooler waters. The fog develops mainly because of the marine environment, where it derives its moisture, but it could only do so as a result of forcing by the cold front.

VARIABLES PRESENT DURING COASTAL FOG PRODUCTION

Variables Always Present During Coastal Fog Formation

Negative or Neutral Omega Within or Just Above the Boundary Layer. When lift occurred during the research, fog would dissipate or simply not form. The lift causes mixing through the inversion, which brings dry, relatively warm air into the boundary layer.

Weak or No Positive Vorticity in the Boundary Layer. Coastal fog developed under boundary layer-negative vorticity regimes, but it would not develop under moderate-to-strong positive vorticity in the boundary layer. When weak vorticity occurred within or at the top of the boundary layer, fog would lift and become a low-level cloud ceiling. As the vorticity center moved past, the ceiling would once again descend to the surface, which caused coastal fog to form.

Inversion. An inversion is always present during any and all fog development and duration.

Variables Present During Research for Each Condition

Radiation Fog Variables

Winds 0 to 3 Knots. Radiation fog events are created without mixing through the inversion. Wind greater than 3 knots was found to create too much mixing and dissipated coastal fog during radiative events.

Moisture Advection or Rainfall Within 36 Hours. Moisture is needed for any fog to develop. With no moisture advection, it was found that moisture input from rainfall would be sufficient inside 36 hours. This field was dependant on several variables. These variables included amount of rainfall, ground moisture, insolation, and amount of rainfall coupled with timing. The best results were found with light rain episodes during the early morning with strong insolation during the day.

Neutral or Negative Omega. Neutral or negative lift from some height above down to the top of the boundary layer was always found when coastal fog formed during this research. Positive lift caused mixing through the inversion, which dissipated the fog.

Weak-to-No Boundary Layer Positive Vorticity. Coastal fog formed under negative and neutral vorticity regimes. Coastal fog was either displaced or was not present when moderate-to-strong positive vorticity was found within the boundary layer. Coastal fog was also present with weak positive vorticity, but a few interesting findings occurred. As a weak positive vorticity center moved through the boundary layer, coastal fog would lift, which created a low-level ceiling from 100 to 400 feet. When the vorticity maxima passed, the low-level ceiling would descend to the ground, which caused coastal fog to return. This phenomenon was known for creating “bouncy” fog conditions where visibilities would swing wildly from as little as 0 to as much as 4 miles.

Clear Skies. Clear skies were the overwhelming majority of sky conditions experienced during radiational coastal fog events. The minority sky condition consisted of very high thin cirrus clouds. No radiational coastal fog events occurred during any other cloud conditions.

Outside Downtown Areas. Radiational coastal fog events during this study were found outside the downtown areas of cities. The heat island effect was enough to dissipate any fog trying to form inside these areas.

Moisture Advection or Rain Within 36 Hours. The highest frequency of coastal fog during frontal regimes was found when a prefrontal trough passed. Moisture was input by both the front causing return flow from the marine environment and the prefrontal trough causing rainfall.

Clear Skies or Very High Cloud Ceiling. High cloud ceilings were noted several times throughout the research when frontal-induced coastal fog developed. Winds of 10 to 20 knots were noted above a shallow inversion, which allowed heat from the boundary layer to escape and be carried away.

Surface Winds of 0 to 12 Knots. During radiative conditions, coastal fog only formed when wind speeds were 0 to 3 knots. During advective or advancing conditions, coastal fog was carried to or along the coast when wind speeds were 4 to 12 knots.

No coastal fog formed during this research when wind speeds were greater than 12 knots.

Onshore Winds of At Least 4 Knots and Not More Than 12 Knots. During marine-induced coastal fog, winds were necessary to advance the fog from the marine environment to the coast. But wind speeds greater than 12 knots caused the fog to dissipate by mixing with air above the inversion. When these winds did not reach the top of the boundary layer, no mixing occurred through the inversion and the fog would simply lift, developing a low-level cloud ceiling.

No Moisture Advection. Surprisingly, moisture preconditioning was not necessary when marine fog caused coastal fog. The marine fog was the moisture advection.

Clear or Very High Ceilings. Several cases showed coastal fog production via a low-level cloud deck. When heat was capable of radiating through the inversion at the top of the low-level clouds, the cloud deck lowered or deepened toward the surface through the night, finally reaching the ground, which caused coastal fog to form. It was also noted that the lower the cloud deck, the shorter the time frame coastal fog would develop. A very general rule of thumb was realized when the boundary layer was capable of supporting fog. On average, it took about an hour for the cloud ceiling to descend 100 feet. Cloud bases higher than 1000 feet never reached the surface during marine-induced coastal fog events.

Temperature. Johnson and Graschel (1) found temperature differences of several variables to be important when maritime fog developed. As indicated in their article, “Sea Fog and Stratus: A Major Aviation and Marine Hazard In The Northern Gulf Of Mexico,” the differences between water and air temperatures, as well as between water and dew point temperatures, were the most important variables producing marine fog when the right atmospheric conditions were in place. The graphs below show these parameters versus relative humidity (RH) values.

Johnson and Graschel’s study (1) as well as this project found RH values of 98% or greater always present with coastal fog.

Figure 1 shows the continental shelf region along the northern Gulf of Mexico, which is shown at the 200-m depth contour. It is also the region where a cool water temperature of 20 °C (68 °F) or less was found to be critical for marine fog development in the northern gulf during the right atmospheric conditions. These findings may not be the same at other locations around the globe because fog development depends on temperature gradients, over water and/or land, which are relative. Water temperature findings close to these were also accomplished by Binhau (5).

Figure 2 shows RH versus (Ta - Tw), where RH is relative humidity and (Ta - Tw) is the difference between the ambient air temperature and the water temperature.

In Fig. 3, in regard to the positive (+) numbers, when the water temperature is cooler than the air temperature, the air must be moving between 4 knots to as high as 12 knots for coastal fog to form. During this process, the moisture-laden warmer air loses its heat to the cooler water. As the air cools to its dew point, condensation takes place. Latent heat is released during the condensation process. The air can lose this added heat to the cooler water below and through the top of the boundary layer

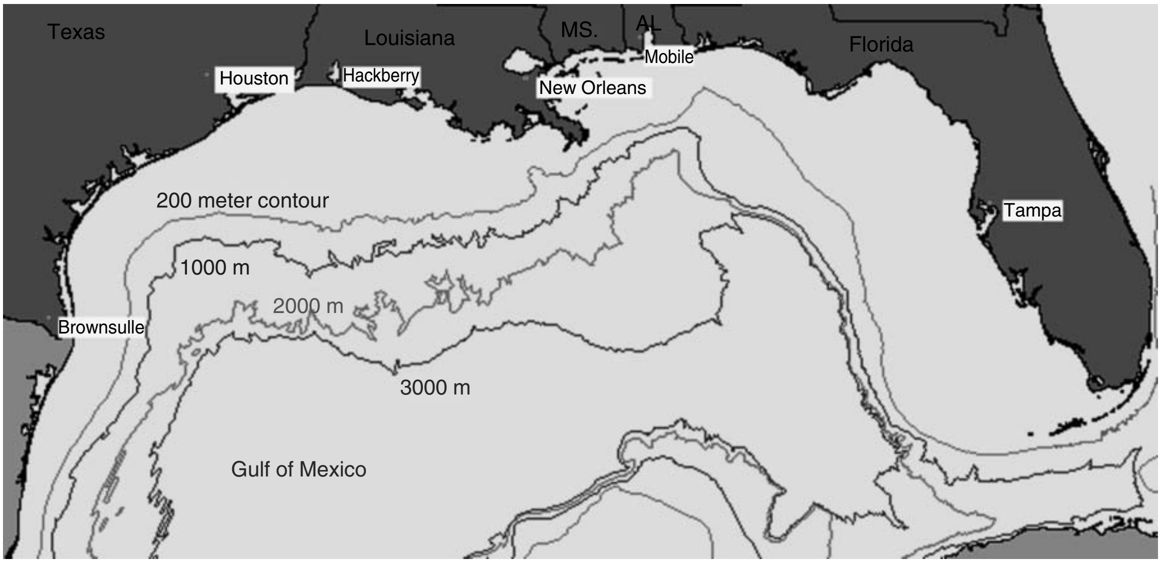

Figure 1. Continental Shelf-Northern Gulf at Mexico.

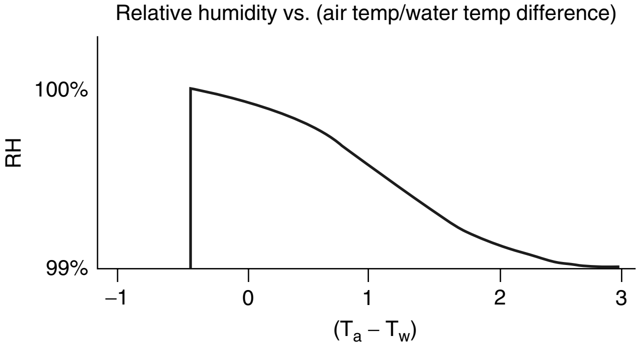

Figure 2. This graph shows the relationship between air temperature and water temperature when coastal fog developed with respect to the lowest relative humidity value found during research.

via long wave radiation, so fog can begin to form. This process is efficient and is the main reason for coastal fog developing from advancing marine fog.

The positive numbers in Fig. 3 show a stable boundary layer also described by Binhau (5) and Hsu (2).

In Fig. 4, in regard to the negative (−) numbers, when the water temperature is warmer than the air temperature, the air must be calm or moving no faster than 3 knots for coastal fog to form. During this process, when coastal tidal marshes exist, a column of cooler dry air will have a net gain of water vapor, which causes the dew point temperature to rise. When the air and dew point temperature are very close, condensation will begin. As a result of condensation, latent heat is released. The column will need to lose this heat through the top of the boundary layer before fog will begin to form. Even though marine fog occurs by this process, it is not an efficient coastal fog producer unless fog initially forms along the coast. Most of the time, fog will develop over the marine environment and be stationary or drift farther offshore.

Negative numbers in Fig. 4 show an unstable boundary layer as described by Binhau (5). Binhau showed that this type of fog may be produced over the marine environment even with northerly winds of 30 knots. But with winds of this magnitude, coastal fog will never develop along the northern gulf coast because the winds would force it out to sea.

Figure 5 shows RH versus (Td − Tw), where RH is relative humidity and (Td − Tw) is the difference between the ambient dew point temperature and the water temperature.

In Fig. 6, in regard to the positive (+) numbers, it is important to remember that the atmosphere as well as the ocean is always trying to reach a state of equilibrium.

When the ambient dew point temperature is higher than the water temperature, water molecules can easily move to and become a part of the ocean surface. Water molecules find it hard to break away from the waters' surface during these conditions, and therefore, a net moisture flux from air to water occurs. The air temperature is always equal to or greater than the dew point. Hence, there will be a transfer of heat from the air to the water as well. This process cools the air temperature, but saturation is difficult to achieve because there is a net loss of water vapor to the water surface. The air continues to cool and dry until temperatures of the water, the air, and the dew point equal or become very close. This process eventually causes saturation and can, but rarely does, cause coastal fog to form when atmospheric conditions are right and (Td—Tw) is zero or very close. Normal occurrences of fog on the positive side of Fig. 6 are when fog develops elsewhere and advances into the area.

In Fig. 7, in regard to the negative(−) numbers, when the water temperature is warmer than the dew point, water molecules can easily break away from the water surface to the air. Regardless of the air temperature,

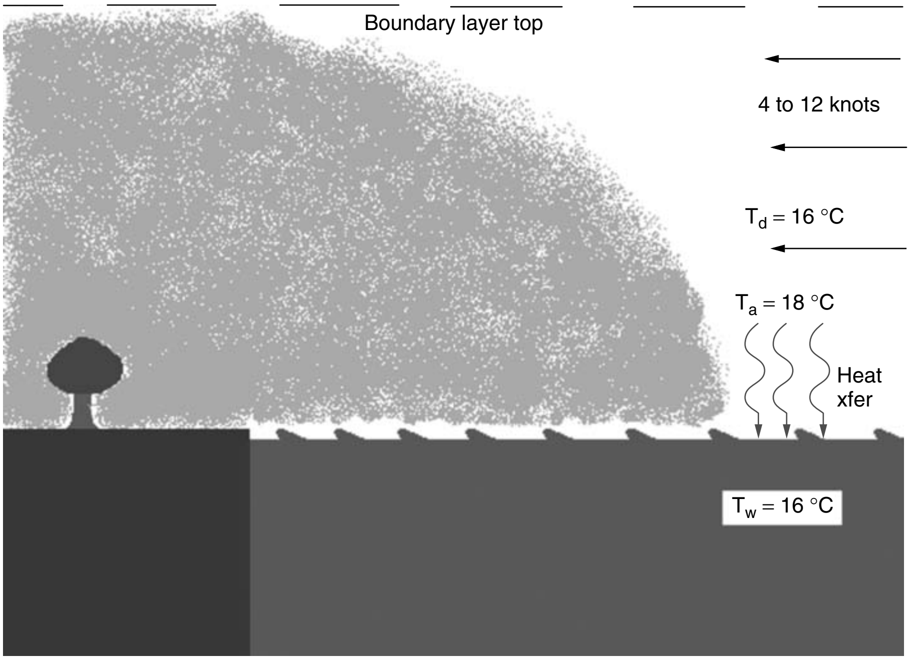

Figure 3. Stable Boundary Layer.

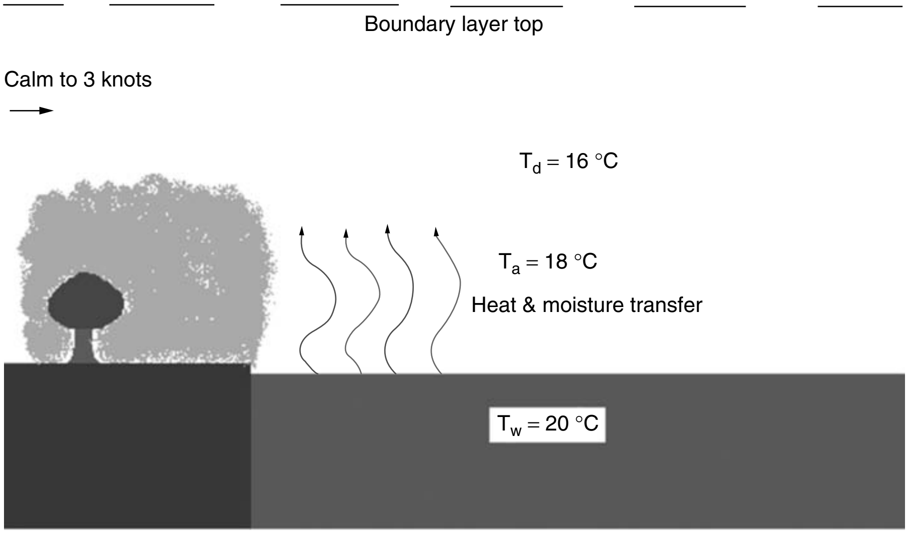

Figure 4. This picture shows a light offshore wind, dew point temperature (Td), air temperature (Ta), and water temperature (Tw).

the water will try to modify it until the water and air temperatures are equal. Even if saturation is achieved, moisture will continue to be added to the air until the dew point is equal to the water temperature. As moisture is added to the air at saturation, condensation begins and latent heat is released to the air until the temperatures of the water, dew point, and air are equal. This saturation can easily develop fog near or offshore. Normally, marine fog produced by this process occurs with little airflow (0–3 knots). After marine fog forms, it can be easily forced to the coast as winds increase ahead of the next cold front. This coastal fog producer is aggressive.

As Fig. 7 depicts, this process occurs with a minimal separation of air, dew point, and water temperatures. When separations are too large, there may be too much

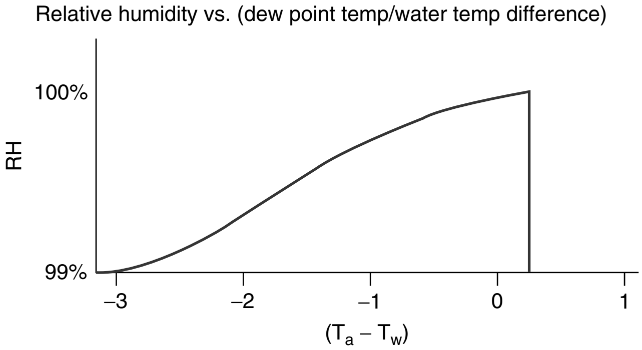

Figure 5. This graph shows the relationship between dew point temperature and water temperature when coastal fog developed with respect to the lowest RH value found during research.

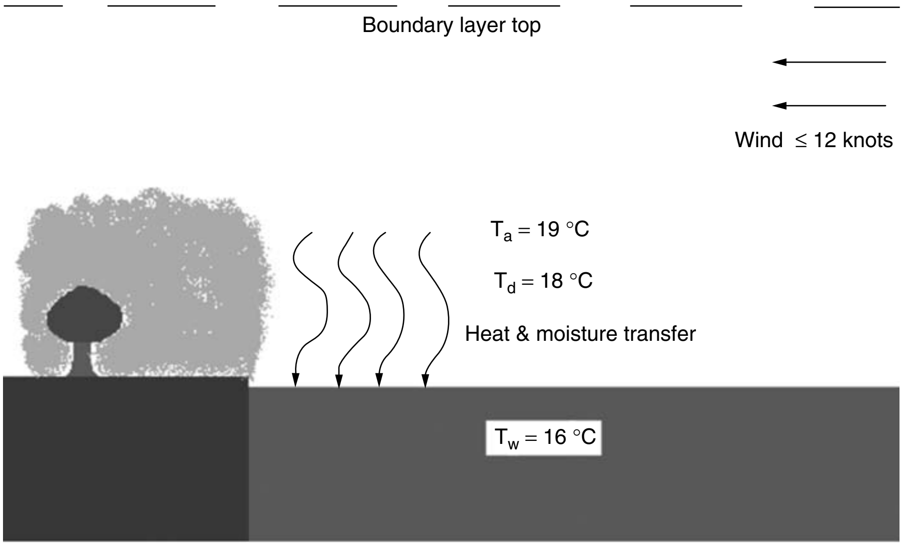

Figure 6. This picture shows winds of 12 knots or less, air temperature (Ta), dew point temperature (Td), and water temperature (Tw).

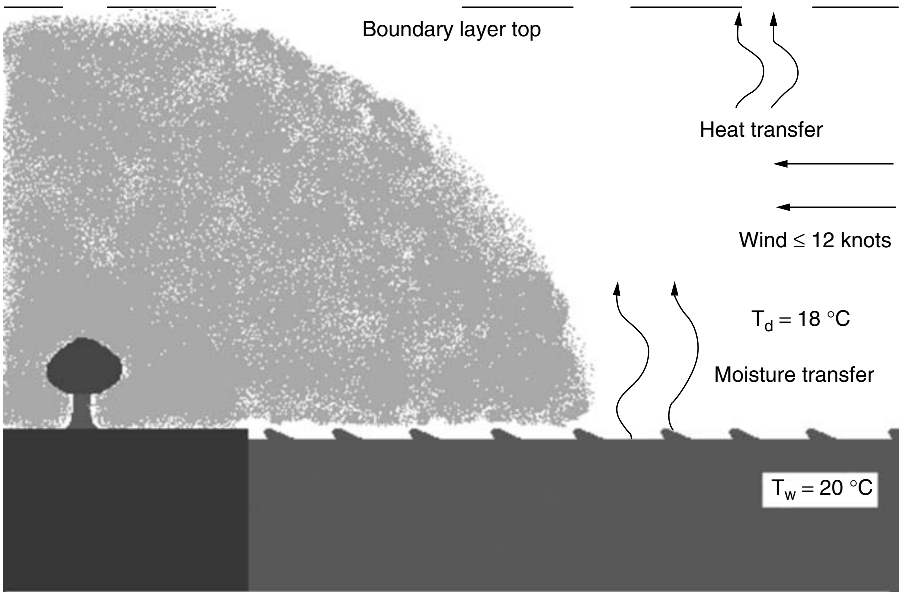

Figure 7. This picture shows wind speeds less than or equal to 12 knots, dew point temperature (Td), and water temperature (Tw).

dry air to modify before the next system brings this air mass back to shore. If the air is moving more than 3 knots, there will be a continual replacement of dry air and/or air from above the inversion may be mixed into the boundary layer and saturation will not be achieved.

Advective Fog Variables

Winds of At Least 4 Knots and Not More Than 12 Knots. Advective fog, which causes coastal fog, is different from advancing marine fog in a couple of ways, i.e., where the fog developed and the wind fetch. Fog that develops onshore and moves to the coast is only known as advective fog. Whenever fog moves to the coast, the wind direction will always be in a direction from the fog to the coastal location. All other variables that are needed for marine fog to advance to the coast are also needed for advective fog.

No Moisture Advection. Even though moisture may be in place, advective fog does not need to have a premoistened atmosphere for coastal fog to be produced. Moisture advection can be induced by the fog moving to the coast.

Clear or Very High Cloud Ceilings. Cloud conditions were found to be the same as for marine fog.

Conglomerate Fog Variables. The variables for each condition associated with coastal fog development have to be present when coastal fog forms under more than one condition.

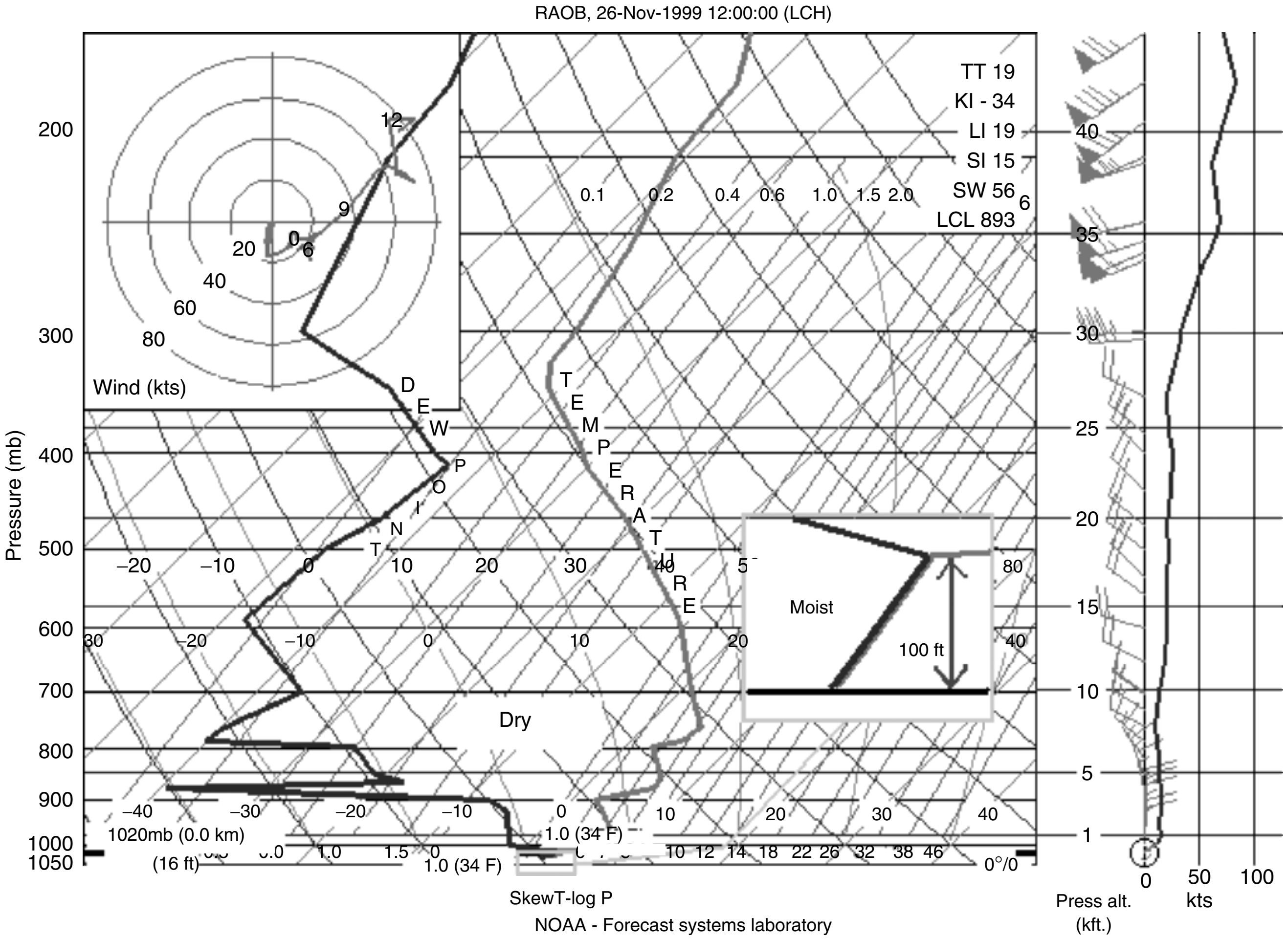

Figure 8. Sounding Profile-Radiative.

SOUNDING PROFILES PRESENT DURING COASTAL FOG

A sounding profile similar to Fig. 8 was always present shortly before and during radiative coastal fog events. This actual sounding shows a very shallow (~100 feet) moist layer capped by three strong inversions.

A sounding profile similar to Fig. 9 was always present shortly before and during advective and maritime coastal fog events. This actual sounding shows a very deep (~1000 feet) moist layer capped by a moderate inversion.

Fog rarely develops beneath the type of inversion in Fig. 10. This actual sounding shows a deep moist layer capped by a frontal inversion; coastal fog occurred before the passage of this cold front, but it quickly dissipated by the time this sounding was taken.

Frontal-induced and conglomerate fog was found with all three types of sounding profiles. Frontal-induced coastal fog occurred more often under the first and second profile types.

Sounding profiles are important for forecasting the depth of coastal fog or any type of fog. Maritime, advective, and frontal-induced coastal fog can only be as deep as the height from the surface to the base of the inversion. These types of fog produced the deepest fog along the coast as well as inland, and consequently they took longer to dissipate.

During research, most radiation-induced coastal fog formed only under a low-level inversion. The inversion could be as high as 100 feet or as low as a few feet from the ground. The inversion was frequently strong. Air parcels were not able to penetrate the inversion, but it would not stop radiative heat transfer from the surface via long-wave radiation. Radiative fog did not always follow the depth rule. It could be from the surface to the base of the inversion deep or as shallow as a few inches, even when the inversion was much higher.

An interesting find during the project was horizontal stability and moisture profiles during radiative fog formation conditions along the coast. The horizontal profile shown below from the coast along with the radiative vertical profile from inland locations were both present when a thin “ribbon” of fog formed along the coast (see Fig. 11). This fog was found to run linearly along the coast and stretched upward for hundreds of feet. The inland vertical sounding did not support such a high

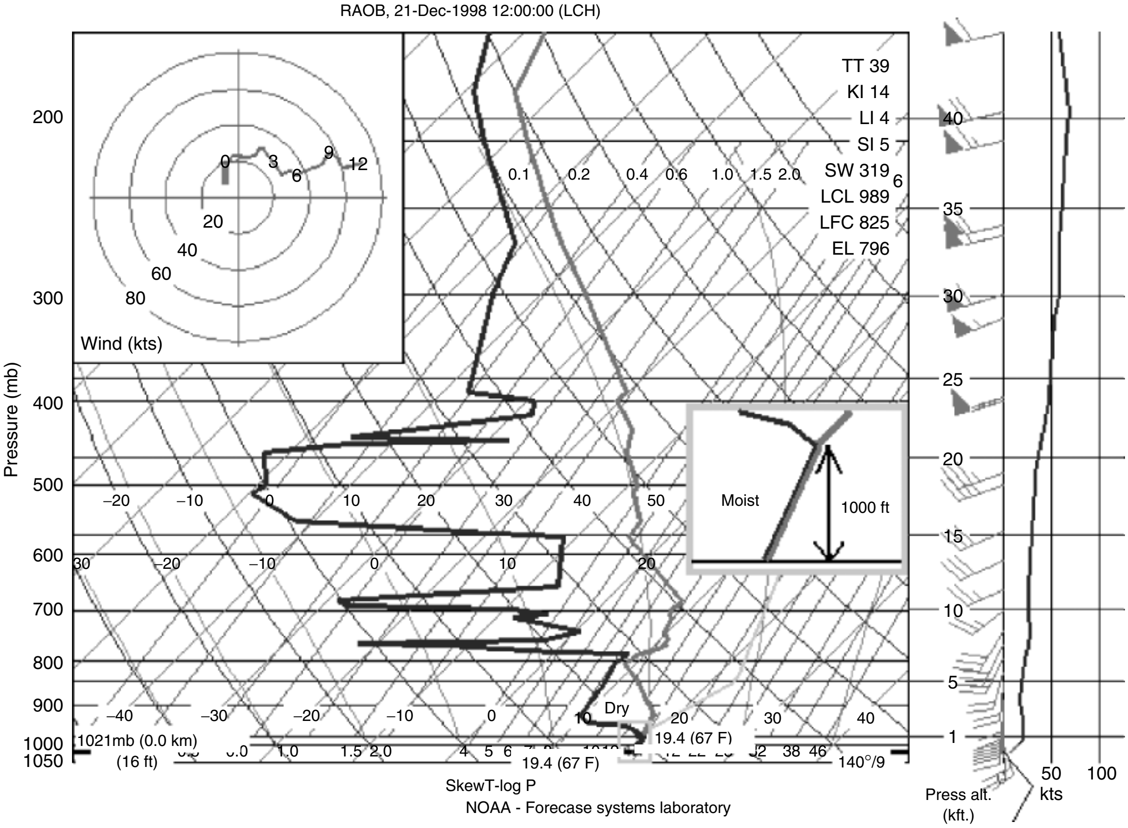

Figure 9. Sounding Profile-Advective.

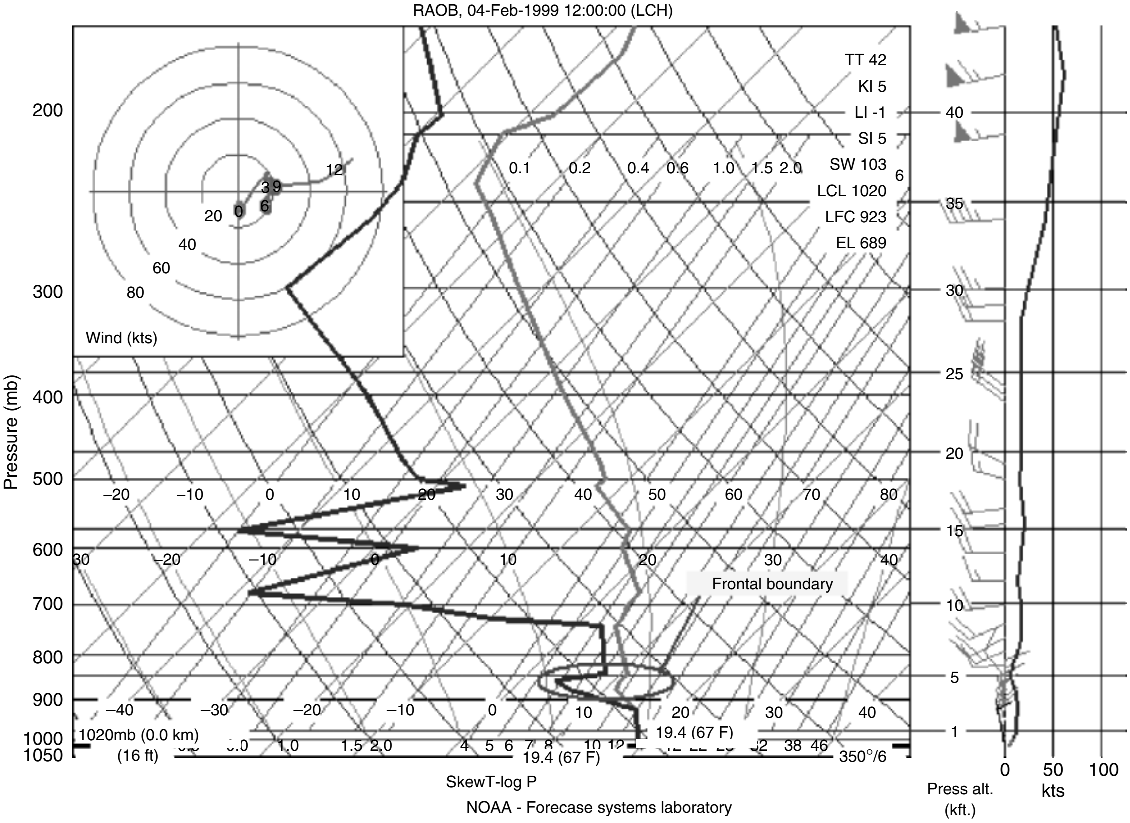

Figure 10. Sounding Profile-Inversion.

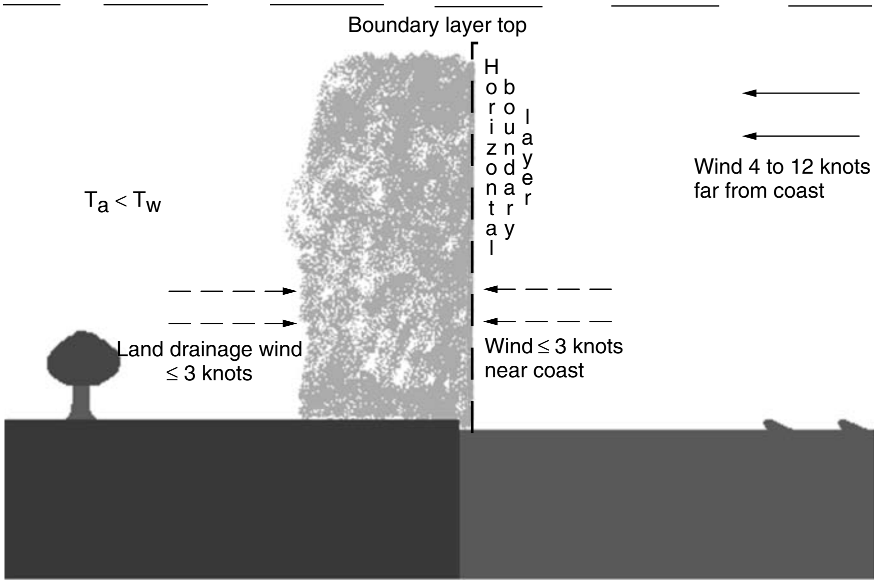

Figure 11. This illustration shows a hypothetical representation on how a very thin ribbon of fog develops at the coast.

depth of coastal fog, but this vertical profile may have been different along the coast because strong moisture and temperature gradients existed. Research into this phenomenon is ongoing.

CONCLUSION

Fog has never received the attention it deserves because it does not make for an explosive story like hurricanes and tornadoes. But fog is blamed for large monetary losses as well as property losses. Each year, fog is blamed for indirectly taking more lives than hurricanes in the United States. Fog also donates a small but significant amount of water to the hydrologic cycle. In some places, such as the western high coast region of South America, fog is the only way insects and grasses receive water.

When fog develops, there is always an inversion in place, which means the boundary layer is disconnected from the remaining atmosphere above with respect to mixing. This process is called decoupling. When the inversion erodes and mixing resumes through this layer, it is said to be coupled. As a result, pollutants released to the environment will remain in the boundary layer during decoupled conditions and will mix out during coupled conditions. Petrochemical plants and other facilities-producing pollutants that are dispersed to the environment can use fog as an indicator for when not to release waste products.

Results during research show there are numerous variables and observations from the microscale environment to consider when forecasting coastal fog conditions. Current National Weather Service numerical models do not solve for microscale conditions, and therefore, forecasters must rely on pattern recognition to resolve these issues.

BIBLIOGRAPHY

1. Johnson, G.A. and Graschel, J. (1990). *Sea Fog and Stratus: A Major Aviation and Marine Hazard in the Northern Gulf of Mexico.*
2. Hsu, S.A. (1988). *Coastal Meteorology*. Academic Press, New York, p. 52.
3. Kotsch, W.J. (1983). *Weather for the Mariners*. Naval Institute Press, Annapolis, MD.
4. Mullan, D.S. (1984). Low stratus and sea fog in the North Irish Sea. *Weather (GB)* **39**: 235–239.
5. Binhau, W. (1985). *Sea Fog*. Springer-Verlag, New York.

RAIN FORESTS

Aldo Conti
Frascati (RM), Italy

Rain forests are, by definition, those forests that receive more than 2500 mm of rain each year. Rain forests are characterized by very dense vegetation dominated by tall trees and huge biodiversity. Rain forests exist in many parts of the planet, but most of them are along the equator, where the weather is stable throughout the year and there is never a dry season. Rain forests do not have seasons at all. The amount of rain is almost constant during the year, and the temperature seldom dips below 16 °C. Rain forests cover 7% of the earth's land surface and 2% of its total surface, but are home to more than half of all animal and plant species. Despite the fact that rain forests cover less than 10% of the earth, they support a third of its plant matter. The largest tropical rain forest in the world is the Amazon Rain Forest, which lies in the countries of Brazil, Bolivia, Peru, Ecuador, and Colombia. There are rain forests in Africa, mainly in the Congo, and in Oceania.

The large amount of rain also creates some of the biggest rivers and flood plains of the world.

Unfortunately, rain forests are in danger. They lie mainly in poor countries, where the economic situation forces people to use all resources, which is why most rain forests have shrunk dramatically in size over the last few decades. What is in danger is their huge biodiversity. Some researchers estimate that the Amazon Forest alone might host up to 10 million species of animals, mainly insects. At the rate the forest is disappearing, most of the estimates say that we are losing something like 100 of these species every day, before even being described by science.

Most of the animals in rain forests have adapted to live in the upper level of the canopy, where food is really plentiful. The constant weather means that there are flowers and fruits at any time of the year. The largest group of animals is that of insects that can easily climb trees and often have a highly symbiotic relationship with the plant life of the forest. Ants and termites are the most abundant animal of every rain forest. Animals are very important for the survival of the forest. Underneath the canopy, the wind is not strong enough to disperse efficiently the seeds produced by plants. Plants rely therefore on insects, that very often pick up the seeds and drop them some distance away. But the most important animals for seed dispersal are birds. Birds in rain forests eat mainly fruits, and the seeds can pass through their digestive system unscathed. By the time a bird excretes its load of seeds, it has normally flown a long distance from the plant, which ensures a high level of genetic mixing, healthy and helpful for plants. There are even some seeds that do not germinate unless they have gone through the digestive system of birds. More than one-fourth of all bird species in the world today live in tropical rain forests.

To have an idea of the importance of rain forests, some figures might be interesting. A single square meter of rain forest supports between 45 and 80 kg of biomass, far more than any other biome. This biomass gains more importance considering that most of it is made of carbon removed from our atmosphere. One hectare of rain forest can contain 200 species of trees and more than 40,000 species of insects. In Panama, scientists discovered fully 80% of the world's currently known beetle species on only 19 trees. Once, researchers discovered over 600 new species of beetle by studying a single species of tree.

Although not as rich in species as their Asian or Amazonian counterparts, African rain forests contain more than half of that continent's animal and plant species, even though they cover less than 7% of its total land area.

Because of the huge biodiversity, most species have evolved to occupy very specialized niches of the environment, which means that many species depend on each other and cannot survive without each other. Deforestation and many other human activities disrupt these complicated relationships.

Rain forests could play a crucial role in feeding the whole world's population. Many vegetables and fruits such as bananas and peppers that we consume come from rain forests. Peanuts come from rain forests, as do many drinks (coffee, tea, cola), oils (palm, coconut), flavorings (cocoa, vanilla), and other foods (beans, grains, fish). And many more vegetables are still there ready to be discovered. Moreover, researchers have identified over 200 plants that produce potential cancer fighting substances. And this considering that only 1% of plants have been intensively screened for such properties. Tropical rain forests do not offer only goods. They are a vital part of the hydrogeologic cycle of the planet and act as a global air purifier, absorbing huge amounts of carbon dioxide and releasing oxygen.

Despite their importance, rain forests everywhere are exposed to huge threats. Often, forests are cleared with fire to make room for cultivation. One plot is used for a few years until the soil is exhausted, and then farmers move on to clear another patch, putting the lushest forests in danger of desertification. But other industrial interests, including timber and mining, are taking advantage of rain forests. Part of the danger comes also from animal and plant species introduced from other environments. All these activities result, every year, in a rather large loss of rain forest. It is difficult to estimate the extent of the damage, as data are not plentiful or reliable. Nevertheless, it is true that in some countries, like Madagascar, the whole forest has almost disappeared in a few decades. As a result, human activities might be inducing the most important mass extinction since the fall of dinosaurs 65 million years ago. According to some research, up to 10% of the world's species might disappear in the next 25 years. But the truth is that over 50% of the earth's plants and animal are in danger. Nearly 20% of known endangered vertebrates are threatened by introduced species. Cultures are going extinct, too. Since the turn of the century, 90 tribes of indigenous peoples have been wiped out in Brazil alone. The pace of annihilation is increasing; 26 of those tribes were killed or scattered in the past decade. Everything should be done to halt this loss that many scientists think might affect the earth's climate, too, on a global scale.

FROST

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

Essentially, a type of dew, frost is ice formed by the condensation of atmospheric water vapor on a surface. It generally occurs at night, and when frost forms, it can sometimes be seen in patterns of ice crystals. Frost can be extremely damaging to outdoor crops and plants. Researchers and weather services throughout the world monitor the effects of damaging frosts. Frost forms on any surface, including cars, grass, and buildings. Frost forms through one of two processes: the formation of dew that subsequently freezes, or deposition, which is the process wherein a gas changes to a solid. In frost formation, this gas is water vapor, and the solid state is ice crystals. These two processes will occur when the air is saturated. Saturation occurs when the air holds as much water vapor as is possible at its temperature and pressure. The temperature at which frost is deposited is known as the frost point temperature.

Condensing water vapor must have something upon which to condense. If the temperature of the ground is below 0 °C, then the deposit will initially be as dew. Over time, this dew will eventually freeze, forming frost. However, if the dew point of the air is below 0 °C, the deposit will be hoarfrost, which is ice that forms directly through deposition, without initially forming as dew. Hoarfrost is also known as black frost, because unlike regular frost, it is not visible as white crystals (normal frost is called white frost).

It is possible for frost to form even if the air temperature is above freezing. Frost formation depends solely on the air's dew point. However, the temperature of a surface affects whether or not dew, frost, or hoarfrost form because colder objects radiate less heat into the air surrounding them, keeping that air's dew point down. The formation of frost is also governed by a process known as radiational cooling. Frost formation requires a surface temperature below 0 °C, so cold surfaces are necessary. At night, certain surfaces will cool much faster than the surrounding air and other surfaces because all objects radiate heat at all times. During the day, objects generally recoup any lost energy from energy received from the Sun. However, at night, objects no longer receive heat radiated by the Sun, and so less energy is generally received, resulting in a rapidly decreasing nighttime temperature. Frost is more likely to form on clear nights because surfaces cool faster when no clouds radiate heat. On cloudy nights, the clouds may radiate enough heat to surfaces to prevent frost from forming.

Frost typically forms under conditions of light or no wind and sufficiently cold temperature. Winds cause air turbulence, and this turbulence mixes the air, which inhibits frost formation. Typically, frost will form more easily overnight because temperatures tend to be lower and the air moves more slowly than it does during the day. Due to radiational cooling, frost forms less often in areas where many buildings, trees, and other objects are; it also forms less often near bodies of water.

Multiple factors are used by scientists and meteorologists to determine whether or not frost will occur on a given night. One is the general weather of that day/night. The situation most favorable to frost formation is a cloudy day followed by a clear night because clouds prevent the Sun from adequately heating the soil. Humidity is also used. If the dew point is over 5.5 °C at night, frosts are unlikely. If it is below 2.2 °C, a frost is highly probable.

In areas where frost forms, local weather services will designate the type of frost that might be deposited in the region. These designations are light, heavy, and killing. A light frost will have no destructive effects on vegetation. A heavy frost is a significant deposit but is not likely to affect the staple vegetation of a region. A killing frost is severely destructive to vegetation and can decimate an entire crop. In the United States, frost warnings will be issued only until October 15th west of a line from Frederick, Maryland, to Charlottesville, Virginia. East of this line, warnings are not issued past November 1st. In the spring, frost warnings are issued only if there is a possibility that crops and other plants could be damaged.

Killing frosts are monitored by weather services, as regions attempt to predict possible arrival dates, so that crop producers can better prepare for their arrival. Various methods can be employed to diffuse the effects of a harmful frost, such as placing small heating systems throughout a crop area, continuously sprinkling water on crops throughout the night, or, on a smaller scale, placing simple screening cloths around a home garden. Farmers and scientists are continuing to experiment with new methods to avoid the harmful effects of black frost on vegetation. Frost harms plants by forming ice in and around cells. The water around the cells is purer than that inside the cell, so it will usually freeze first. Those plants that have greater quantities of solutes within their cells are more resistant to frost and can withstand cold temperatures more easily than those plants with little solute in their water. Plant damage from frost is determined by the type of plant, the stage of growth of the plant, and the length of time the temperature is low enough for frost formation.

READING LIST

The Encyclopedia of the Atmospheric Environment on-line. Available: http://www.doc.mmu.ac.uk/aric/eae/Weather/Older/Frost.html.

The Weather Doctor. Available: http://www.islandnet.com/∼see/weather/whys/frost.htm.

About Frost by Steve Horstmeyer. Available: http://www.shorstmeyer.com/wxfaqs/frost/frost.html.

UCLA Department of Atmospheric Science on-line. Available: http://www.atmos.ucla.edu/web/.

Iowa State University—On-line Case Studies. Available: http://www.iitap.iastate.edu/jhodson/idot/frost/jan1/frost1.html.

Dew and Frost. Available: http://www.nssl.noaa.gov/∼cortinas/1014/112_2.html.

ASK—A—SCIENTIST Archive, Argonne National Laboratory. Available: http://newton.dep.anl.gov/askasci/wea00.htm.

When Does The Last Frost Usually Occur? Available: http://www.crh.noaa.gov/gjt/frost.htm.

How Does Frost Injury Occur? Available: http://www.gov.on.ca/OMAFRA/english/crops/field/news/croppest/cp0798_w.htm.

Sterling Watch/Warning Definitions. Available: http://www.erh.noaa.gov/er/lwx/Defined/.

Cooling Air. Available: http://www.doc.mmu.ac.uk/aric/eae/Weather/Older/Cooling_Air.html.

Frosty Cars. Available: http://www.weatherwise.org/qr/qry.frostycar.html.

Frost Damage, Control, and Prevention—Fruit and Vines. Available: http://www.pir.sa.gov.au/pages/agriculture/horticulture/frostdamagefs.pdf.

Frost Risk Assessment and Damage. Available: http://www.pir.sa.gov.au/pages/agriculture/horticulture/frost_risk.htm:sectID=445&tempID=11.

FROST DAMAGE

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

Frost is ice formed by the condensation of atmospheric water vapor on a surface. When low temperatures are present in a certain region, the potential for frost damage to plant life exists in that region. Frost damage can injure plants permanently or slow plant development. Several factors dictate the extent and severity of frost damage to plants in any given region.

Frost damage can have profound effects on agriculture and ecosystems. Frost is most likely to form on a cold, clear night that has been preceded by a cold and cloudy day. Lack of wind is vital to a developing frost. The little heat that was stored in the ground during the day will dissipate more quickly overnight. If the dew point has dropped below freezing, the result, is a heavy frost. As conditions change from day to night, an air temperature drop of 1 °C per hour is a signal that a frost is likely to develop.

Even if the conditions of freezing temperatures and a calm night are present, some areas will be more susceptible to frost damage than others. Fields that have lighter soils, which dry out faster, fail to insulate the soil below. This will prevent the natural warming by soil radiation of the air directly above the soil surface and increase the likelihood that a heavy frost will form. Recently cultivated fields will suffer the same fate. Low-lying valleys, where cold dense air cannot be affected by winds, can be heavily damaged. Areas of a field that are close to the edge of a crop formation are also susceptible. In these areas, grass near a crop formation acts as a blanket or insulator, preventing warm soil from heating cold air directly above it. Finally, areas that have recently been treated with herbicides are more susceptible to frost damage. Herbicide stress on a plant can be compounded by cold stress from the weather, increasing the possibility of frost damage.

Frost damage is evidenced by a variety of symptoms on plants. Plant leaves are the best indicators of frost damage. Frost damage on the youngest leaves of a plant (the top leaves) is often seen as a burn on the tips of the leaves. More severe damage is a darkening of the entire leaf. In the most extreme form, the entire plant takes on a black appearance. This darkening is evidence that the frost penetrated and destroyed cell membranes of the plant. Severe darkening of this kind is sometimes referred to as a "killing frost." Whether a killing frost has set in on plant life may not be noticeable until a day or two after the frost. If the plant has turned almost completely brown, chances of recovery are not good. However, a closer look may reveal that the lower part of the plant, or the pseudostem, is still green, a good sign that some recovery from the damage may be possible. The three most important factors in determining the ability of plant life to withstand damaging frost are the plant's maturity, its health prior to a frost, and the weather immediately following the frost. Susceptibility to frost damage increases as plant development increases. As the growing season progresses, the chances of weather conditions conducive to a damaging frost decrease. Young plants are less susceptible to frost damage that will lead to a plant's death because their growing points are still below ground, insulated from freezing temperatures. However, should frost injury occur at this young stage, it could severely delay growth as the season progresses and affect the overall harvest. More mature plants present more opportunities for frost damage. Mature plants have more exposed leaves and may have growing points above the earth's surface. Damage to the outside of a mature plant can constrict future growth. Plant health prior to a frost also determines the ability of a plant to recover. If plants have been continually exposed to cold stress, herbicides, excessive moisture, or disease, even the most minimal frost can be debilitating. Finally, the weather following frost damage plays an important role in the plant's recovery. If warm temperatures follow frost damage, a plant's ability to recover increases.

Worldwide, 5–15% of all agricultural production is lost to frost damage each year. Frost can also cause a loss of food supplies for an animal species by killing leaves, seeds, and fruits.

READING LIST

www.epa.gov—*United States Environmental Protection Agency*—Office of Pesticide Programs.

www.sciencedaily.com/releases/2000/10/001017073120.htm—*Science Daily Magazine*—Climate Change Shifts Frost Seasons & Plant Growth.

www.agric.gov.ab.ca/crops—*Alberta Agriculture, Food, And Rural Development*—Frost Damage to Cereals.

www.pioneer.com/usa/crop—*Pioneer Hi-Bred International, Inc.*—Microclimatic Effects on Frost Damage, Early Season Frost Damage to Corn.

www.pir.sa.gov.au/pages/agriculture—*PIRSA Agriculture*—Frost Risk Assessment and Damage.

Carter, P.R. (1995). Late spring frost and post-frost clipping effect on corn growth and yield. *J. Prod. Agric.*.

Kunkel, K.E. and Hollinger, S.E. (1995). Late spring freezes in the central USA: Climatological aspects. *J. Prod. Agric.*

Nielsen, R.L. (1999). Assessing frost damage to young corn. Purdue pest management and crop production newsletter. *Purdue Univ.*, 27 May 1999.

THE GLOBAL WATER CYCLE

U.S. Global Change Research
Program

USGCRP-supported research on the global water cycle focuses on: (1) the effects of large-scale changes in land use and climate on the capacity of societies to provide adequate supplies of clean water; and (2) how natural processes and human activities influence the distribution and quality of water within the Earth system and to what extent the resultant changes are predictable. Specific areas include:

---

identifying trends in the intensity of the water cycle and determining the causes of these changes (including feedback effects of clouds on the global water and energy budgets as well as the global climate system); predicting precipitation and evaporation on timescales of months to years and longer; and modeling physical/biological processes and human use of water, to facilitate efficient water resources management.

The USGCRP budget includes $311 million in FY 2003 for research and observations related primarily to the Global Water Cycle. The Global Water Cycle program studies the movements and transformations of water, energy, and water-borne materials through the Earth system and their interactions with ecosystems. The movements and transformations of water are important because they appear to control the variability of the Earth's climate and they provide an essential resource for the development of civilization and the Earth's environment. Figure 1 schematically illustrates the movements and transformations. This cycling involves water in all three of its phases—solid, liquid, and gaseous—and exchanges large amounts of energy as water moves and undergoes phase changes. Therefore, the water cycle operates necessarily on a broad continuum of time and spatial scales.

Water vapor is a greenhouse gas that maintains temperatures in a range required by life on Earth. Many of the uncertainties in the current projections of the effects of the atmospheric buildup of carbon dioxide are related to the feedbacks between the climate and the water cycle. While warmer temperatures enable the atmosphere to hold more water leading to further warming, the complex interactions among changing cloudiness, precipitation patterns, land cover, and decreasing snow and ice cover have limited the quantitative understanding of the links between water and climate warming.

Water is not evenly distributed over the globe, nor is it always accessible for human use. Society is becoming more vulnerable to variations in the water cycle as a result of expanding populations and increasing water use. The increasing demands for water accompanied by the growing economic losses from droughts and floods place pressure on the science community to develop the knowledge and tools needed to manage our limited water resources more effectively. There are large potential paybacks from increased investments in scientific research to improve the monitoring and prediction of the global water cycle variations and in water management applications.

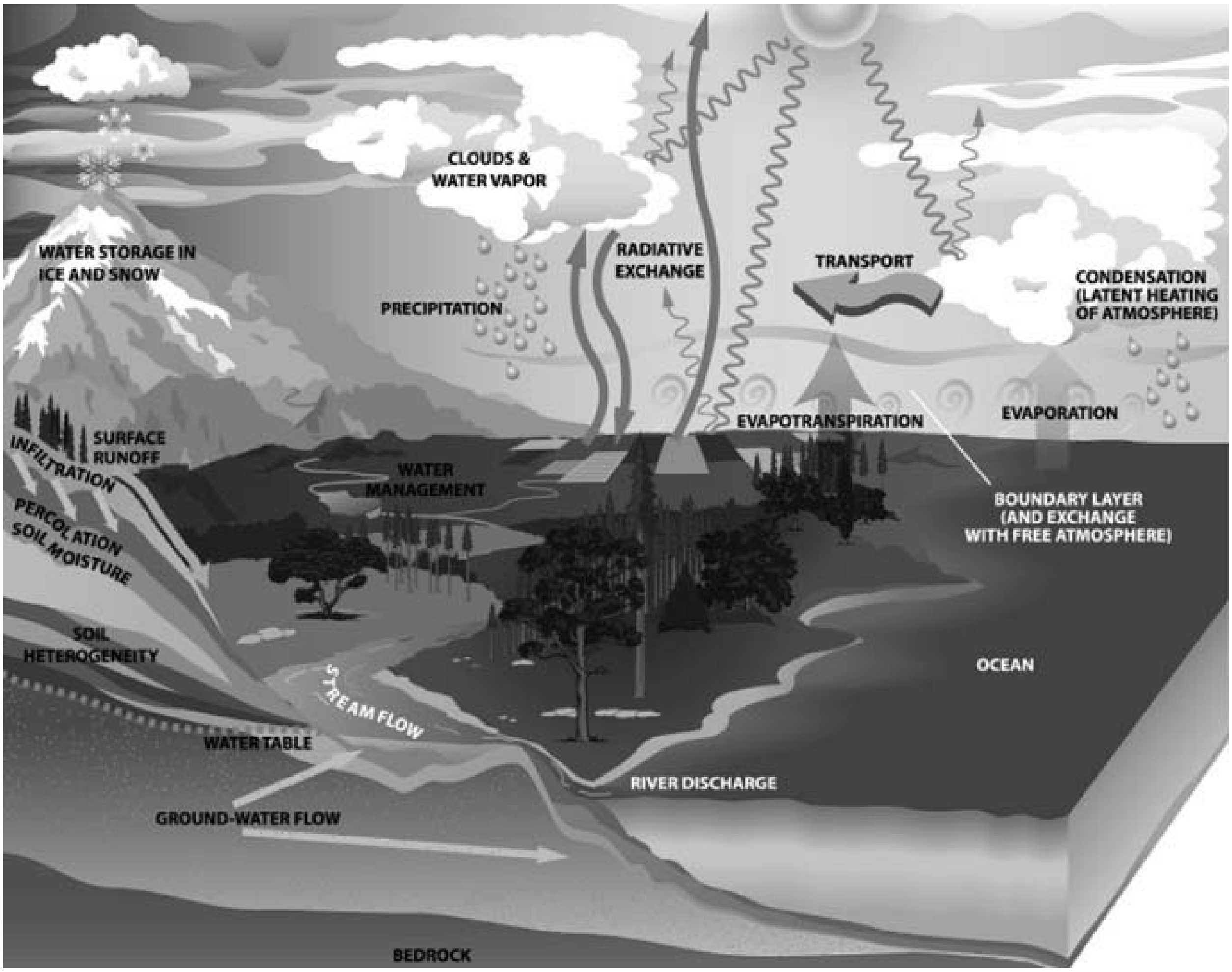

Figure 1. Conceptualization of the water cycle.

On a national basis, near-crisis situations have occurred in several dry southwestern river basins, including the Colorado and Rio Grande, where over-allocation has taken place. Recent drought conditions and rapid development in these basins have exposed the intensity of competition that exists over the available water resources. The development of a capability to predict where water management crises will emerge due to a drought or extended flood conditions is a priority for the Global Water Cycle program. The ability to provide probabilistic forecasts of rainfall and snowfall at various time and space scales is at the center of all potential applications of climate change science and climate information systems. The program has research activities directed at developing experimental predictions that will ultimately benefit society through better protection of human health and assets, and more efficient water system management and infrastructure planning.

Human activity is an integral part of the water cycle. A recent USGCRP-commissioned report, *A Plan for a New Science Initiative on the Global Water Cycle*, issued in 2001, concluded that, among other priorities, there is a pressing need to determine the causes of water cycle variations on both global and regional scales, and to what extent these variations are induced by human activities. In view of this emerging link between water science and water resource issues, the USGCRP global water cycle strategic plan addresses two major questions: (1) What are the effects of large-scale changes in land use and climate on the capacity of societies to provide adequate supplies of clean water, and (2) how do natural processes and human activities influence the distribution and quality of water within the Earth system and to what extent are resultant changes predictable?

Stakeholders are helping to define the Global Water Cycle program at the catchment and larger river basins scales. Users are interested in better forecasts of precipitation, runoff, and soil moisture. Reservoir management decisions require forecast lead times of up to seasons and, in some cases, years. For planning reservoirs, dam recommissioning, and water control infrastructure, and developing new proposals for water law, projections of water variability are required on the decadal to century timescales.

The USGCRP Global Water Cycle program focuses on characterizing, explaining, and predicting variability and long-term changes in the global water cycle and their impacts. To address the issues arising from the intimate role of the water cycle in controlling climate variability on seasonal to multidecadal timescales, the program investigates the pathways of water movement between the biosphere and surface hydrologic systems, the atmosphere, and the oceans, as well as feedback processes between climate, weather, and biogeochemical cycles. Because the biosphere is a substantial regulator of the Earth's carbon cycle, the global water cycle maintains a considerable influence upon the global pathways of carbon. Globally, the cycling of water and its associated energy and nutrient exchanges among the atmosphere, ocean, and land determine the Earth's climate and cause much of climate's natural variability.

A critical contribution of the USGCRP to Federal water activities lies in the benefits that come from drawing together the wide range of programs and expertise from different agencies with the capabilities of the academic community to address these complex issues. The elements of the management structure that the USGCRP has put in place during the past year include: (1) Interagency Global Water Cycle working group, (2) Global Water Cycle scientific steering group, and (3) Global Water Cycle program office. The linkages between the global water cycle, the global carbon cycle, and climate will be explored in the coming year through this strengthened program management structure.

SEE ALSO:

Water Cycle [also available: PDF Version]. Chapter 5 from the S*trategic Plan for the Climate Change Science Program* (*July 2003)*. See also the draft white paper, *The Global Water Cycle and Its Role in Climate and Global Change* [PDF] (posted 27 Nov 2002).

Water Cycle. Presentation from Breakout Session 8 of the *US Climate Change Science Program: Planning Workshop for Scientists and Stakeholders*, 3–5 December 2002, Washington, DC.

Climate Variability—Atmospheric Composition—Water Cycle. Presentation from Breakout Session 19 of the *US Climate Change Science Program: Planning Workshop for Scientists and Stakeholders*, 3–5 December 2002, Washington, DC.

GROUND-BASED GPS METEOROLOGY AT FSL

Seth I. Gutman
Kirk L. Holub
NOAA Forecast Systems
Laboratory

INTRODUCTION

Water vapor is one of the most important components of the Earth's atmosphere. It is the source of precipitation, and its latent heat is a critical ingredient in the dynamics of most major weather events. As a greenhouse gas, water vapor also plays a critical role in the global climate system: it absorbs and radiates energy from the sun and affects the formation of clouds and aerosols and the chemistry of the lower atmosphere. Despite its importance in climate and weather prediction, water vapor has been one of the most poorly measured and least understood components of the Earth's atmosphere. Researchers at FSL and elsewhere are utilizing recent technology to reverse this situation.

The ability to use the Global Positioning System (GPS) to make accurate refractivity measurements under all weather conditions has led to the development of a promising new meteorological observing system for NOAA. The

---

first and most mature application of ground-based GPS meteorology involves the measurement of integrated (total column) precipitable water vapor (IPW) in the atmosphere. The GPS-IPW technique is more advantageous than conventional water vapor observing systems because of its low-cost, high-measurement accuracy, all weather operability, and long-term measurement stability. Further, GPS-IPW requires no external calibration, operates unattended for long periods with high reliability, and is easily maintained. Since GPS-IPW measurements are compatible with satellite data retrievals, they provide an independent method for calibrating and validating global satellite observations.

These positive attributes, however, are accompanied with one major disadvantage: GPS-IPW provides no direct information about the vertical distribution of water vapor in the atmosphere. In an attempt to mitigate this deficiency, researchers at government laboratories and universities around the world are investigating the best ways to use GPS-IPW as a "proxy quantity" for moisture profiles in weather forecasting.

In this article we discuss how IPW is now calculated from GPS signal delays and the potential use of slant-path measurements in numerical weather prediction models. Preliminary results of the effect of GPS-IPW on numerical weather prediction, the demonstration network, data and product availability, and plans for the operational network are also described.

CALCULATING IPW FROM GPS SIGNAL DELAYS

GPS signals are delayed as they pass through the Earth's atmosphere (Fig. 1). The signal delay caused by the presence of free electrons in the ionosphere makes the largest contribution to the total atmospheric delay. Because the ionosphere is a dispersive medium, the velocity of the GPS signals is frequency dependent and its impact can be effectively eliminated by using dual frequency receivers.

Below the ionosphere, in the electrically neutral portion of the atmosphere, refraction (that is, slowing and bending) of the GPS signal is caused by changes in temperature, pressure, and water vapor. Most of this delay occurs in the troposphere, which extends from about 9 km at the poles to about 13 km at the equator. The primarily tropospheric delay consists of a hydrostatic (or dry) component caused by the mass of the atmosphere and a wet component (the wet delay) caused by the dipole moment of the water vapor molecule. The contributions of the wet and dry components of the tropospheric signal delay are in the same proportion as the wet and dry components of the atmosphere.

FSL currently collects GPS observations from a demonstration network of 55 sites (Fig. 2) and processes them to produce IPW measurements every 30 minutes using the scheme shown in Fig. 3.

The first step in obtaining IPW from GPS observations is to determine the zenith-scaled delay caused by the neutral atmosphere. This delay is commonly referred to as the zenith tropospheric delay (ZTD), and is calculated from carrier phase and range observations made by networks of GPS receivers. The calculation is made using GPS analysis software such as GAMIT (GPS At MIT), which in addition to the GPS observations, requires improved satellite orbits and parameters describing the orientation of the Earth in space and time. Next, the ZTD is separated into its wet and dry components using additional observations made by collocated surface meteorological sensors. The zenith- scaled hydrostatic delay (ZHD) is caused by the mass of the atmosphere directly above the site and can be estimated with great accuracy from a surface pressure measurement. The wet signal delay (ZWP) is caused by water vapor along the paths of the radio signals to all satellites in view, about 6 to 8 with the current GPS satellite constellation.

ZWP is calculated simply by subtracting the hydrostatic delay from the tropospheric delay. The resulting wet delay can be mapped into IPW with an error of about 5 degrees using a quantity that is proportional to the mean vapor pressure-weighted temperature of the atmosphere (Tm). Tm may be estimated from a climate model, the surface temperature derived from a numerical weather prediction model, or measured directly using remote sensing techniques. FSL is planning to utilize model-derived Tm estimates operationally.

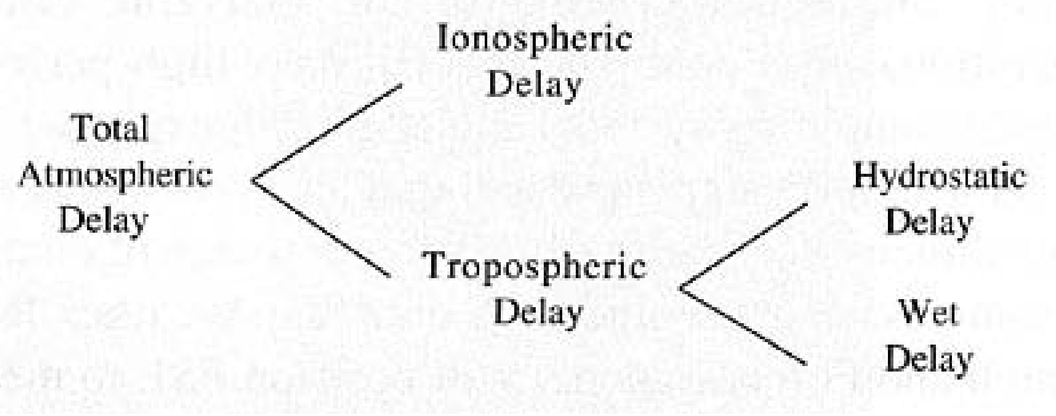

Figure 1. Signal delays caused by the atmosphere.

Figure 2. A map of the NOAA-FSL Global Positioning System Integrated Precipitable Water (GPS-IPW) Demonstration Network (55 sites) as of October 1999.

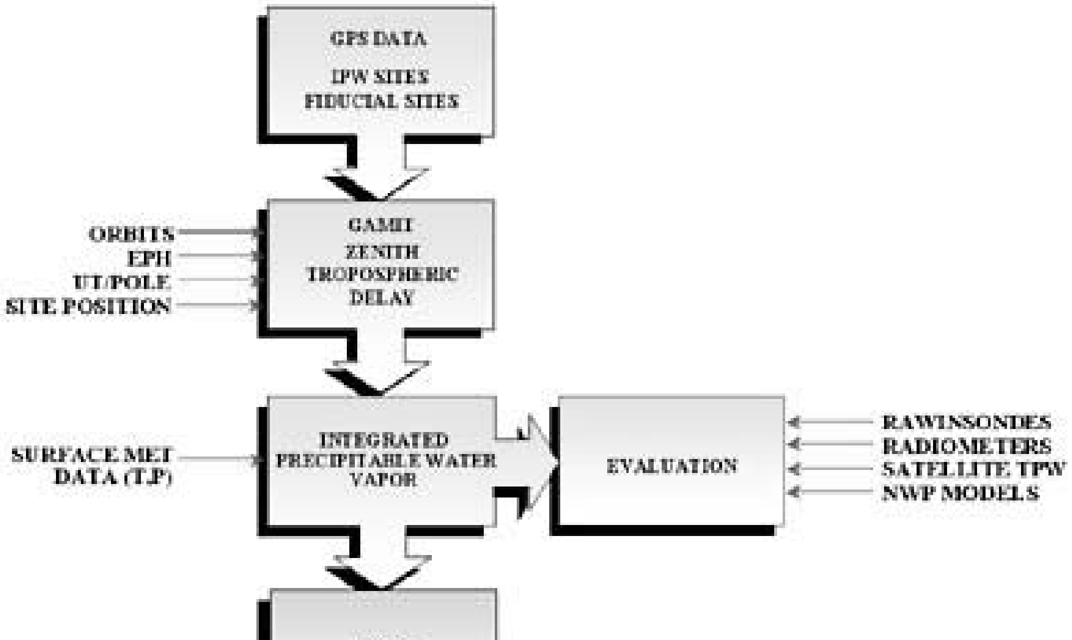

Figure 3. The FSL-developed data processing scheme used to produce IPW measurements.

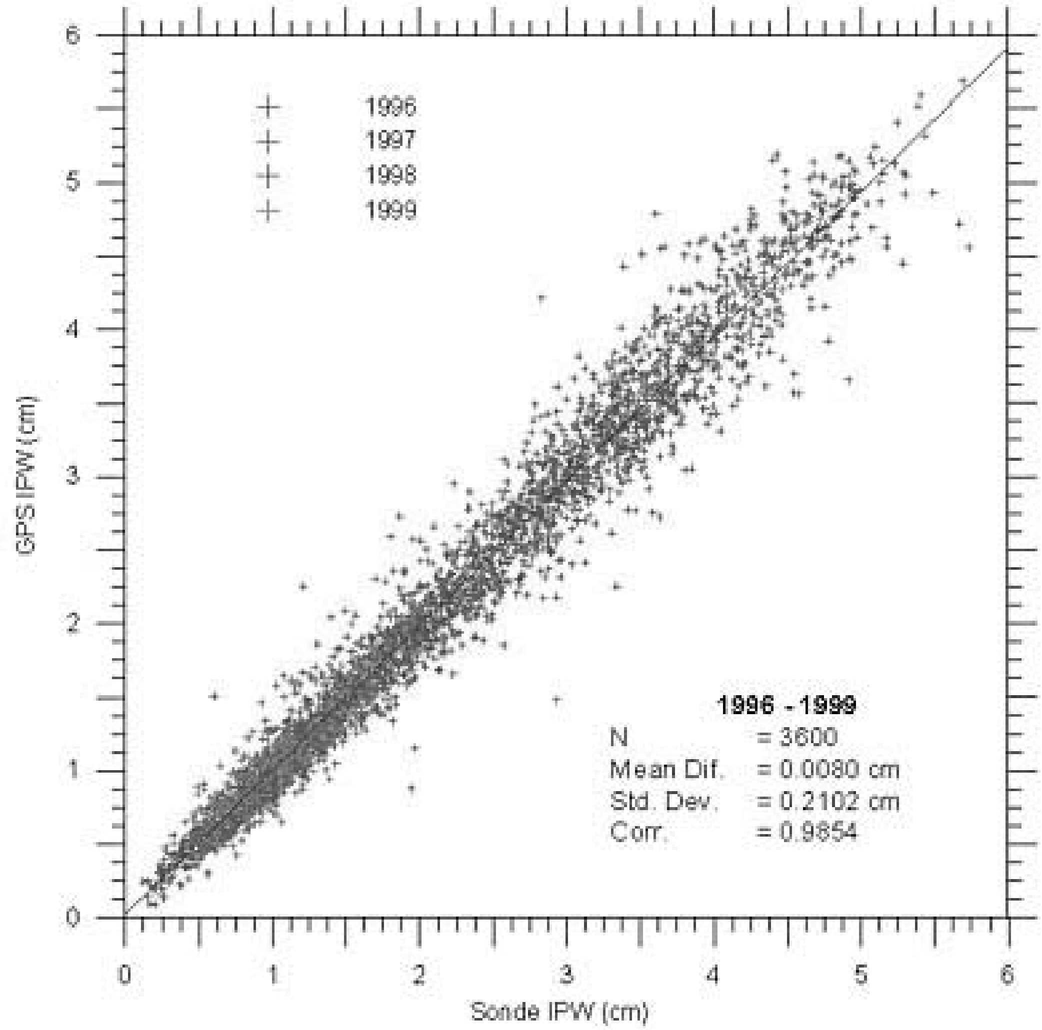

Figure 4. Scatterplot of GPS and rawinsonde observations of integrated precipitable water vapor at the ARM CART site near Lamont, Oklahoma, between January 1996 and September 1999.

Integrated precipitable water calculated from GPS signal delays is physically identical to integrated measurements or retrievals made by other upper-air observing systems including rawinsondes, ground-based microwave water vapor radiometers, or satellite microwave and infrared instruments including sounders and interferometers. Comparisons of GPS and radiosonde-derived total column water vapor have been carried out continuously since 1996 under all weather conditions at the DOE Atmospheric Radiation Measurement site near Lamont, Oklahoma. Results from 3600 comparisons (to September 1999) indicate a mean difference of 0.08 mm and a standard deviation of 2.1 mm (Fig. 4).

SLANT-PATH SIGNAL DELAY MEASUREMENTS

Recent investigations by FSL director A.E. MacDonald and Yuanfu Xie (of the Forecast Research Division) of the potential use of line-of sight estimates of path-integrated water vapor (derived from slant-path GPS signal delay measurements) to retrieve the 3-D moisture field have been very interesting and potentially significant. The experiments involve assimilating simulated slant-path moisture measurements from a wide area network of closely spaced stations into the Quasi-Nonhydrostatic (QNH) model using variational techniques. In recent research, their simulations indicate that it may be possible to recover the three- dimensional structure of the moisture field from a densely spaced network of ground-based GPS receivers making a single line-of- sight, or slant path, measurement of the signal delay to all satellites in view. The configuration of the GPS satellite constellation as seen from Boulder, Colorado, between 1200 and 1300 UTC on 28 September 1999 is shown in Fig. 5. A GPS satellite moves across the sky at the rate of about 30 degrees per hour. Although 10 satellites are visible above the horizon in this example, six to eight would be more typical at any one time.

Making a slant-path signal delay measurement with the same accuracy as a zenith-scaled measurement is

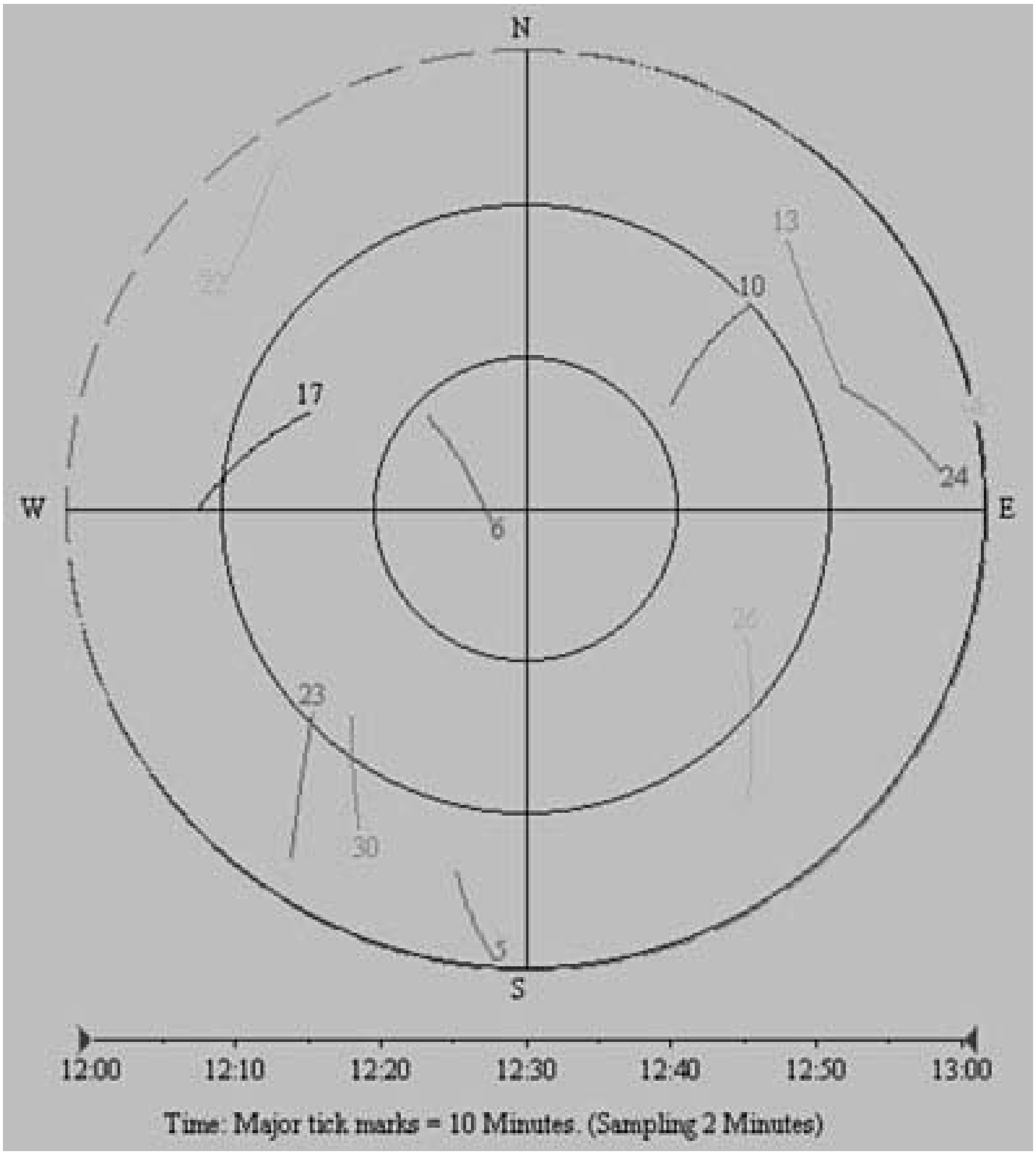

Figure 5. Configuration of the GPS satellite constellation as seen from Boulder between 1200 and 1300 UTC 28 September 1999.

not trivial. The sources of measurement error that are successfully managed through geodetic modeling of the zenith tropospheric signal delay will have to be dealt with in other ways. Although some of the most important information about the structure of the atmosphere can be obtained from low-angle observations, measurement errors increase significantly along with the negative impact of multipath reflections from nearby obstacles as satellites approach the horizon. One way to mitigate these problems is to utilize advanced GPS receivers and antennas that maximize the ability to track satellites under all conditions and reject multipath reflections. Unfortunately, not all problems can be eliminated through the selection of hardware, and advanced data processing techniques will be needed as well.

Research at Scripps Institution and the University of Hawaii into ways to monitor the accuracy of GPS orbit predictions suggests that these techniques can also be used to reduce systematic errors in slant-path signal delay or refractivity measurements to individual satellites.

EFFECT OF GPS-IPW DATA ON THE ACCURACY OF NUMERICAL WEATHER PREDICTION

Since 1997, parallel runs with and without GPS have been carried out using the research version of the Rapid Update Cycle (RUC-2) model to assess how GPS-IPW data affect the accuracy of numerical weather prediction. Results from the first two years using optimal interpolation techniques have been encouraging despite the fact that the observations came from only a limited number of widely spaced sites. Model runs using data acquired from more sites over a larger area through September 1999 confirm improvements

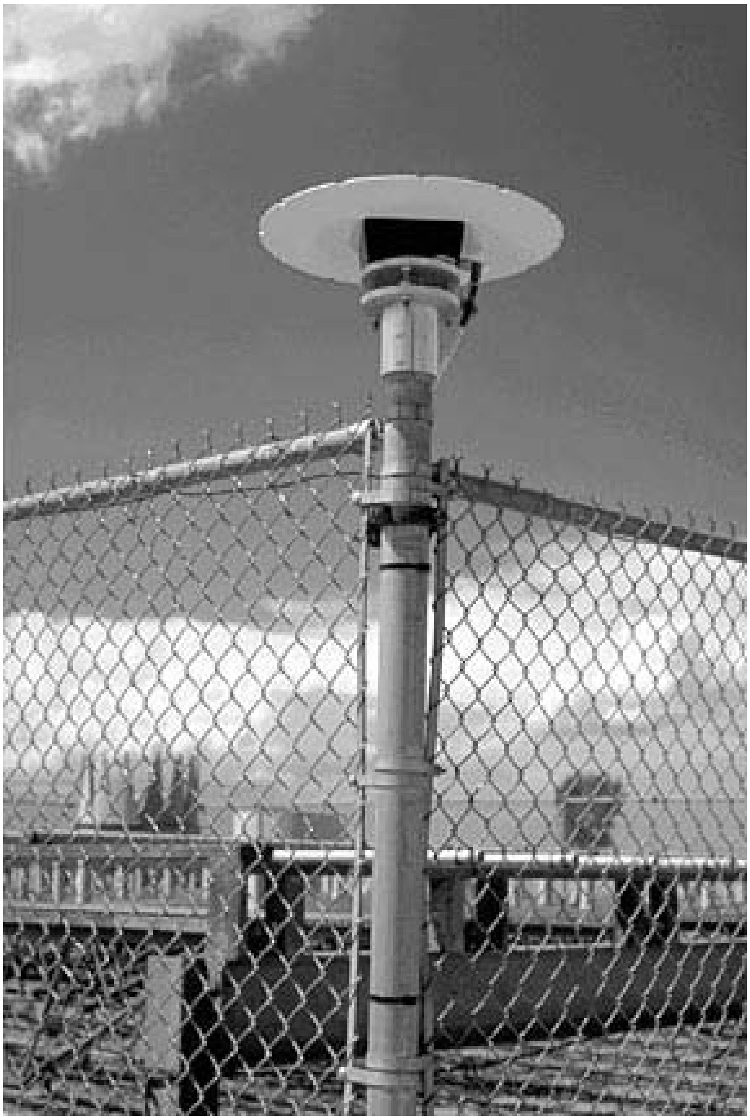

Figure 6. GPW-IPW installation at the NOAA Profiler Network site at Platteville, Colorado.

in forecast accuracy, especially under conditions of active weather. Therefore, NOAA meteorologists expect significant improvements in short-term forecasts of clouds, precipitation, and severe weather when high-resolution numerical weather prediction models routinely use data from a nationwide network of GPS-IPW systems in conjunction with data from other observing systems and advanced data assimilation techniques.

The decision to implement ground-based GPS Meteorology (GPS-Met) as a next-generation upper-air observing system will be supported in part by promising assessments such as this one. In anticipation of a favorable decision, network design and implementation options for a national network of ground-based GPS- IPW systems are being evaluated at FSL.

GPS-IPW DEMONSTRATION NETWORK

The rapid development of the GPS-IPW Demonstration Network for meteorological remote sensing has been made possible by a fortuitous synergy with the positioning and navigational applications of GPS by the U.S. Coast Guard and U.S. Department of Transportation. As of October 1999, the data acquisition component of the demonstration network consisted of 55 GPS-IPW systems operating in the continental United States and Alaska.

Thirty-four systems are currently installed at NOAA Profiler Network (NPN) sites, seven at sites belonging to other NOAA organizations or institutions affiliated with NOAA, 11 belong to the U.S. Coast Guard Maritime Differential GPS (DGPS) system, and three are at the Department of Transportation Nationwide Differential GPS facilities. Typical sites from each organization are illustrated in Figs. 6–9.

In addition to supporting the assessment of GPS as a possible next-generation upper-air observing system, the GPS-IPW Demonstration Network is designed to help NOAA accomplish the following tasks:

- Evaluate the engineering and scientific bases of ground- based GPS-Met, including advanced data acquisition and processing techniques.

Figure 7. GPS-IPW installation at the Scripps Institution of Oceanography at La Jolla, California.

- Develop and test strategies to build, monitor, operate and maintain large networks of GPS reference stations for meteorological remote sensing.
- Develop techniques to acquire, process, and quality control GPS observations and data products.
- Provide observations and derived meteorological products to users (such as forecasters, modelers, researchers) and data archives.
- Transfer ground-based GPS-Met technologies to operational use in weather forecasting and climate monitoring.
- Other possible applications under investigation include calibration and validation of environmental satellite data and improved positioning and navigation services.

All ground-based observing systems in the GPS-IPW Demonstration Network consist of dual-frequency GPS receivers and antennas, and collocated surface meteorological sensors.

These systems are located at sites where shelter, power, and communications are available to operate and collect data from the instruments, and transmit these data in real or near real time to one of two locations. The generalized flow of data and products from the network is illustrated in Fig. 10.

DATA AND PRODUCT AVAILABILITY

GPS and surface meteorological observations from the GPS-IPW Demonstration Network sites are available to the general public in near real time through the NOAA National Geodetic Survey. Information and raw data may be acquired via the Web.

Processed data, including GPS signal delays and integrated precipitable water vapor, are available shortly after improved NAVSTAR GPS satellite orbits and Earth Orientation Parameters are available from one of the International GPS Service (IGS) tracking stations. This usually occurs within 24 hours of the close of the day, but efforts to accelerate the process and make improved orbits available within 1–3 hours are well underway. IPW and other products may be acquired from the FSL Demonstration Division, *GPS-Met Observing Systems Branch*.

Figure 8. GPS-IPW installation at the U.S. Coast Guard Differential GPS site at Cape Canaveral Air Force Station.

PLANS FOR NETWORK EXPANSION AND IMPROVEMENT

Our primary goals in 2000 are to continue to expand the demonstration network, demonstrate distributed data processing using low-cost PCs instead of high-end workstations, and implement real-time data processing.

Expansion of the Demonstration Network. Now that all NOAA Profiler Network sites have been equipped with GPS-IPW systems, expansion of the network will mostly proceed by installing GPS Surface Observing System (GSOS) packages at U.S. Coast Guard Maritime DGPS and Department of Transportation NDGPS sites. Depending on the availability of funds and the status of interagency agreements under review, during the next year we hope to install 21 new systems at DGPS sites (mostly in the Mississippi Valley and Great Lakes regions), 11 NDGPS sites, and one at the Department of Energy ARM facility at Point Barrow, Alaska (Fig. 11).

Implementation of Real-time Data Processing. We define real-time data processing as acquiring and processing GPS and ancillary observations to yield signal delay or IPW calculations within a single numerical weather prediction assimilation cycle. In the case of the Rapid Update Cycle, running operationally at the National Centers for Environmental Prediction, this is approximately 75 minutes.

Real-time data processing techniques are being tested and evaluated in a collaborative effort involving FSL, the Scripps Permanent Orbit Array Center, and the University of Hawaii at Manoa. Techniques involve acquiring data from a subset of the IGS global tracking network and using these observations to produce an improved retrospective orbit with only about 2-hour latency. An orbit prediction that covers the data gap is also made, and it is this short-term prediction that is used to calculate IPW. In theory, the error in a prediction that spans only 2 or 3 hours will be proportionally less than an error made over an interval of 36–48 hours.

Real-time quality control techniques are also under evaluation. The most promising involve continuous monitoring of the relative positions of a number of sites, and using these data to infer changes in orbit

Figure 9. GPS-IPW installation at the DOT National Differential GPS site at Whitney, Nebraska.

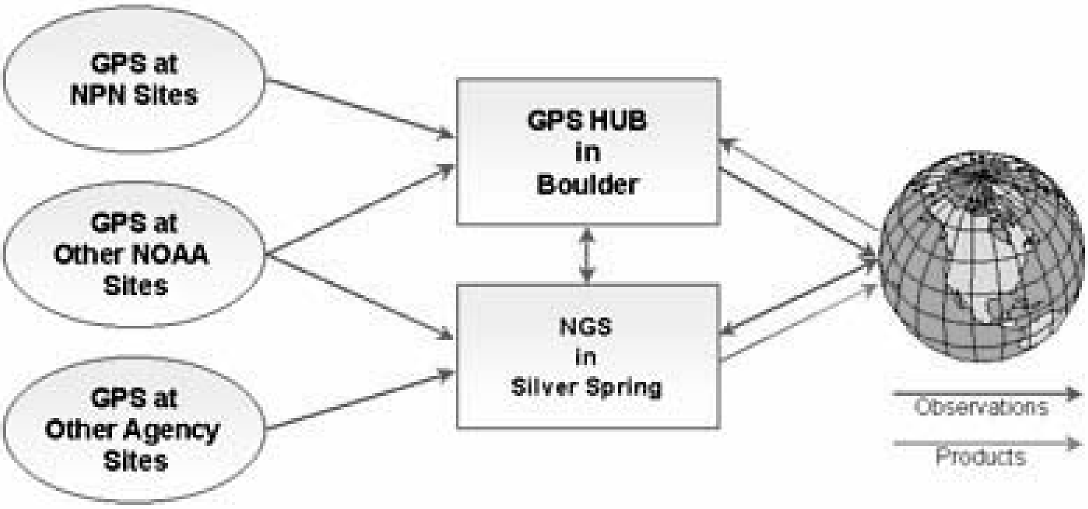

Figure 10. Flow of data and products from the GPS-IPW Demonstration Network.

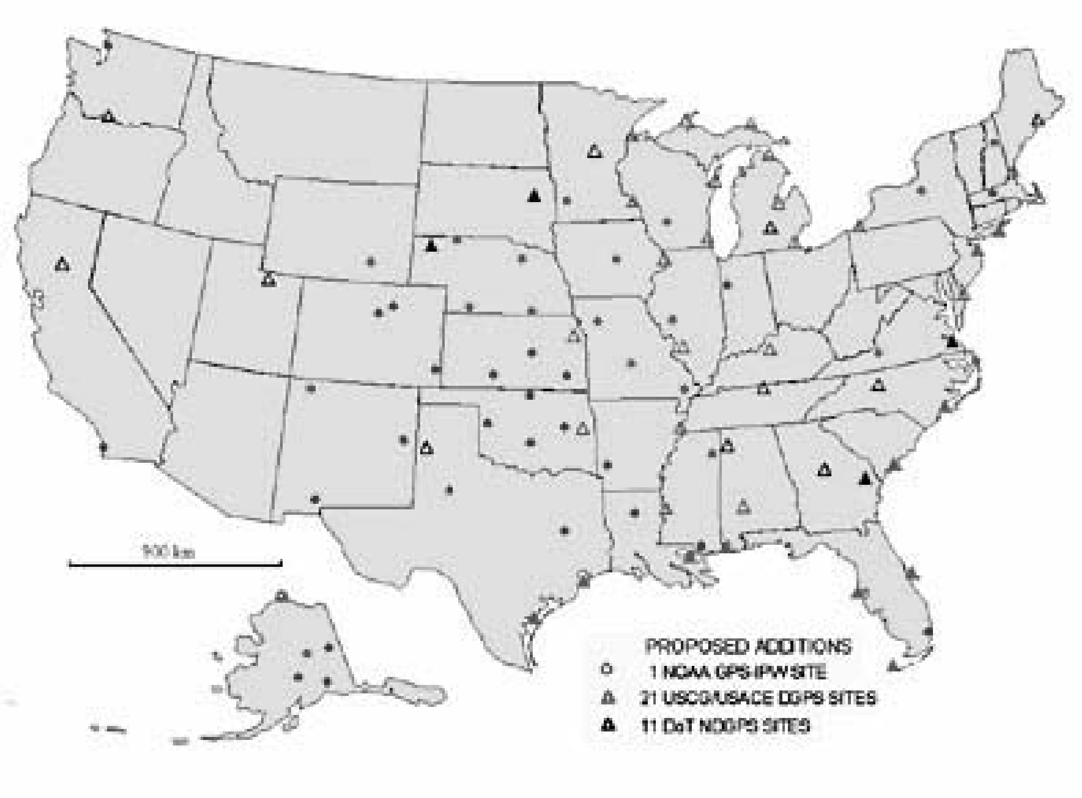

Figure 11. Planned expansion of the Demonstration Network to about 92 sites during 2000.

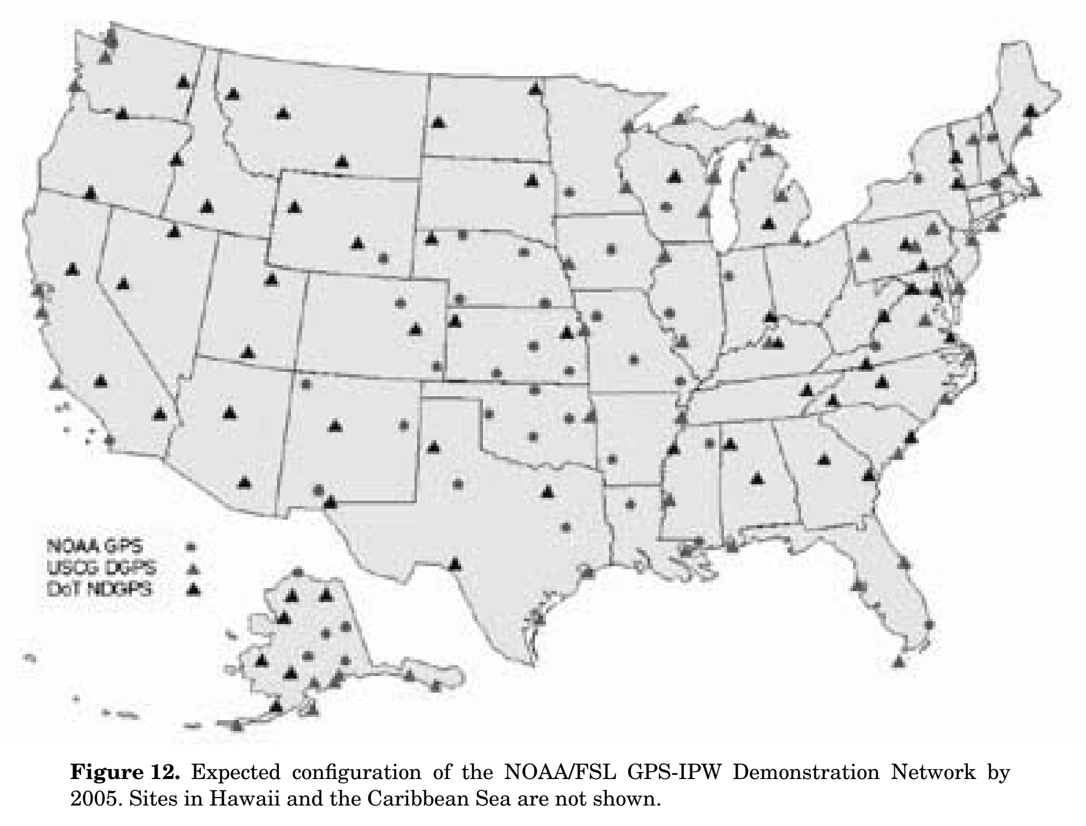

Figure 12. Expected configuration of the NOAA/FSL GPS-IPW Demonstration Network by 2005. Sites in Hawaii and the Caribbean Sea are not shown.

Figure 13. An Automated Surface Observing System (ASOS) installation at Cape Hatteras, North Carolina.

Figure 14. GPS receiver placement on top of FSL's new office building, the David Skaggs Research Center.

Next Page

Figure 15. A map of the AWIPS offices.

accuracy for specific satellites in the constellation. When a problem is encountered, the satellites are removed temporarily from the ephemerides until an updated orbit can be produced.

Distributed Data Processing. Recent advances in low-cost PC processor speed and memory will be utilized to perform data processing in a fully distributed environment. During this year we have demonstrated the ability to partition a large network into smaller subnetworks, process each independently in substantially less time with no significant loss of accuracy and precision.

Operational GPS-IPW Network Implementation Strategy. The expansion of the GPS-IPW Demonstration Network to about 200 sites, and the transition from retrospective to real-time data processing will enable us to assess the impact of these data on weather forecast accuracy. Based on the results of these studies, a decision to implement ground-based GPS-IPW as a next-generation upper-air observing system for NOAA is expected. The following plan has been developed to expand the demonstration network to an operational network of about 1000 sites with an average station spacing of somewhat less than 100 km (Fig. 12).

- Once transition of the GPS-IPW Demonstration Network to operational status has become a reality, receivers and antennas will be upgraded. Communications will transfer from FTS-2000 to the AWIPS [Advanced Weather Interactive Processing System] communications systems via the Internet.
- GPS receivers will be added to about 800 Automated Surface Observing (ASOS) sites (sample site shown in Fig. 13). The reason for collocating GPS at ASOS sites is to take advantage of the surface meteorological data and site infrastructure, including shelter, power, data communications, field maintenance, and logistics support. This will minimize implementation time and life-cycle cost. The GPS antenna installation at a typical ASOS site will resemble the one at FSL's new location, the David Skaggs Research Center (Fig. 14).
- Data processing hardware, software, and training will be provided to all AWIPS offices (Fig. 15).

[Editor's Note: More information on the topics covered here is available by contacting Seth Gutman, who can provide copies of published articles which include a list of references.]

CLIMATE AND WATER BALANCE ON THE ISLAND OF HAWAII

James O. Juvik
D.C. Singleton
G.G. Clarke
University of Hawaii
Hilo, Hawaii

INTRODUCTION

The island of Hawaii, with a surface area of only 10,455 km, exhibits a spectacular range of climatic

---

Previous Page

Figure 15. A map of the AWIPS offices.

accuracy for specific satellites in the constellation. When a problem is encountered, the satellites are removed temporarily from the ephemerides until an updated orbit can be produced.

Distributed Data Processing. Recent advances in low-cost PC processor speed and memory will be utilized to perform data processing in a fully distributed environment. During this year we have demonstrated the ability to partition a large network into smaller subnetworks, process each independently in substantially less time with no significant loss of accuracy and precision.

Operational GPS-IPW Network Implementation Strategy. The expansion of the GPS-IPW Demonstration Network to about 200 sites, and the transition from retrospective to real-time data processing will enable us to assess the impact of these data on weather forecast accuracy. Based on the results of these studies, a decision to implement ground-based GPS-IPW as a next-generation upper-air observing system for NOAA is expected. The following plan has been developed to expand the demonstration network to an operational network of about 1000 sites with an average station spacing of somewhat less than 100 km (Fig. 12).

- Once transition of the GPS-IPW Demonstration Network to operational status has become a reality, receivers and antennas will be upgraded. Communications will transfer from FTS-2000 to the AWIPS [Advanced Weather Interactive Processing System] communications systems via the Internet.
- GPS receivers will be added to about 800 Automated Surface Observing (ASOS) sites (sample site shown in Fig. 13). The reason for collocating GPS at ASOS sites is to take advantage of the surface meteorological data and site infrastructure, including shelter, power, data communications, field maintenance, and logistics support. This will minimize implementation time and life-cycle cost. The GPS antenna installation at a typical ASOS site will resemble the one at FSL's new location, the David Skaggs Research Center (Fig. 14).
- Data processing hardware, software, and training will be provided to all AWIPS offices (Fig. 15).

[Editor's Note: More information on the topics covered here is available by contacting Seth Gutman, who can provide copies of published articles which include a list of references.]

CLIMATE AND WATER BALANCE ON THE ISLAND OF HAWAII

James O. Juvik
D.C. Singleton
G.G. Clarke
University of Hawaii
Hilo, Hawaii

INTRODUCTION

The island of Hawaii, with a surface area of only 10,455 km, exhibits a spectacular range of climatic

---

diversity comparable with that found on large continents. Three major factors contribute to this climatic diversity:

1. Topographic relief. The volcanic mountains of Mauna Kea and Mauna Loa reach summit elevations of 4,205 m and 4,168 m, respectively. The attitudinal range provides for a diversity of temperatures, and the mountains themselves are barriers that induce orographic precipitation.
2. Large-scale synoptic wind field. The strong and persistent northeast trade winds interact with the island topography to produce distinctive windward and leeward climates. The associated upper-level trade wind inversion exerts a particularly strong control on mountain precipitation gradients.
3. Local circulation. Differential heating and cooling of the land, water, mountain, and lowland areas on Hawaii give rise to localized wind regimes which add to the island's climatic diversity.

KOPPEN CLIMATIC ZONES

Integrating the attitudinal temperature gradients with the annual, seasonal, and spatially variable rainfall regimes results in a diverse combination of climatic environments. The Koppen climate classification uses monthly temperature and precipitation characteristics in a descriptive system that distinguishes broad regional and global climatic zones. The system has been often criticized for its empirical approach and lack of emphasis on "dynamic processes" (e.g., 1); however, as a "first approximation" the Koppen classification offers useful insights into regional climatic patterns.

Four broad Koppen climatic zones are distinguished on the island of Hawaii. They are organized primarily as concentric attitudinal bands on the mountain slopes. Figure 2 illustrates the spatial distribution of these Koppen climatic types on the island. The map was constructed on the basis of temperature (absolute or extrapolated) and precipitation data from 55 island stations. Discussions of the zones follow.

Humid Tropical Zone (A climates)

Characterized by warm temperatures throughout the year and relatively high annual rainfall, humid tropical climates occupy the lower slopes of the island from sea level to about 450 m (slightly higher in warmer areas of leeward

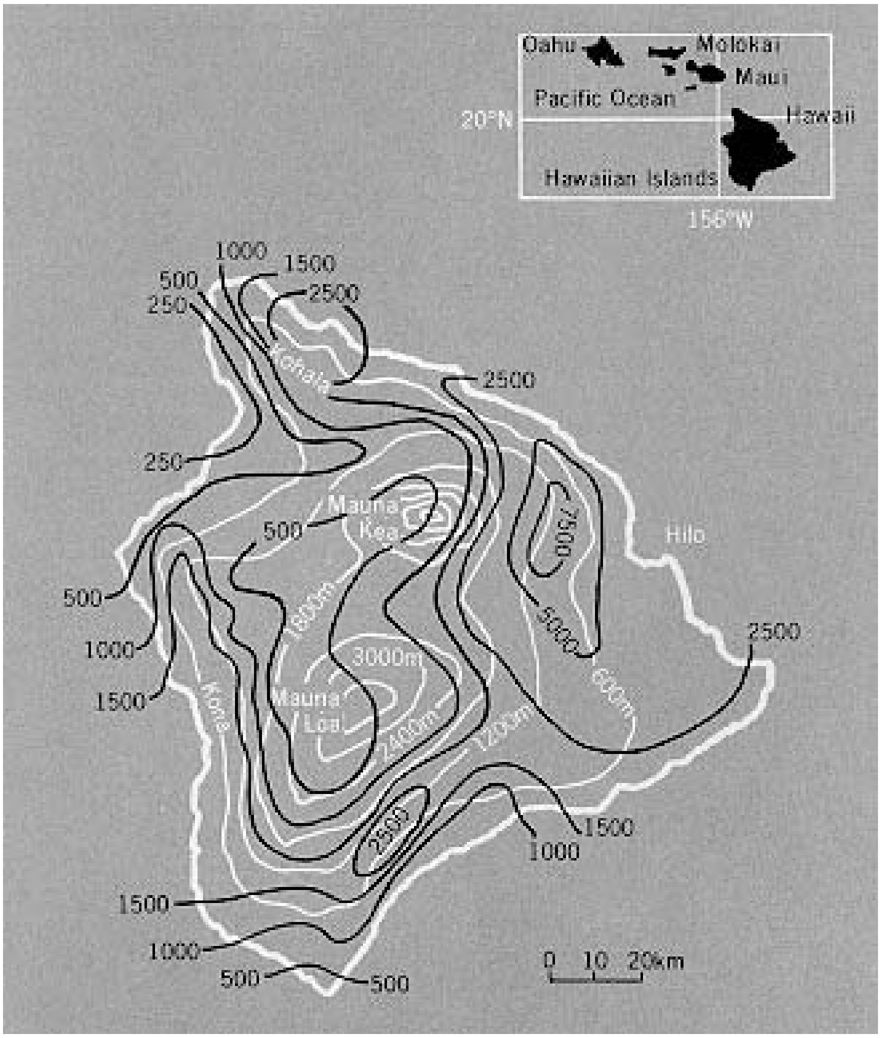

Figure 1. Distribution contours of mean annual rainfall (mm), superimposed on topographic map of the island of Hawaii. (Redrawn from 2,3).

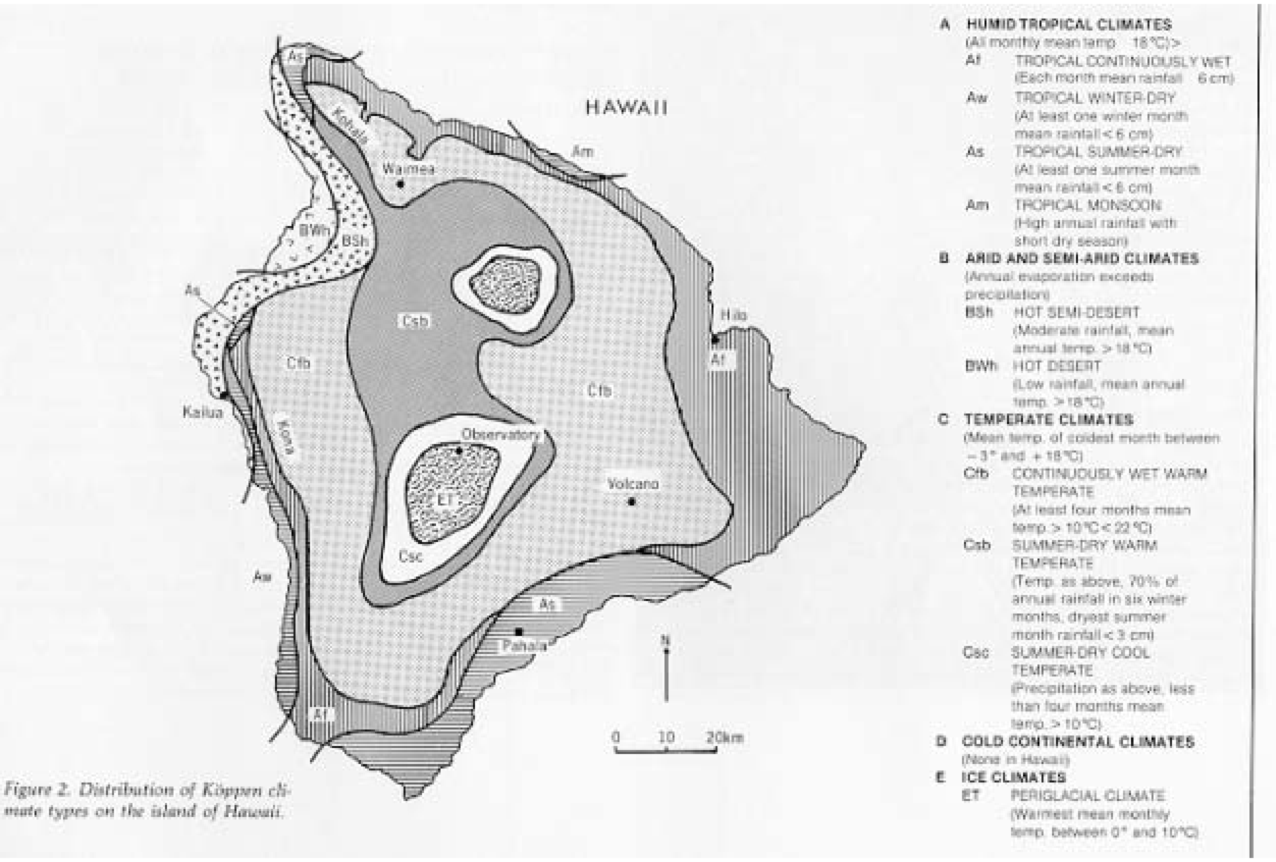

Figure 2. Distribution of Köppen climate types on the island of Hawaii.

Kona). This tropical zone may be further differentiated on the basis of rainfall seasonality. Large-scale synoptic disturbances in winter (mid-latitude cyclonic storms) produce substantial rainfall that is to some extent independent of slope aspect or elevation, and as a result most locations on the island exhibit an absolute winter maximum in rainfall. However, windward areas of Hawaii also receive substantial orographic rainfall throughout the year, with the result that there is no distinct dry season (Af climate). Lowland areas on the island that are transitional in location between windward and leeward receive less orographic rainfall (since they are not oriented normal to trade wind flow) and exhibit a distinctive summer dry season (As climate). Humid summer-dry climates are not common anywhere in the world, since for most tropical locations rainfall is at a maximum in the summer, the result of increased convective instability in the high-sun period. Outside of Hawaii the As climate type occurs only in southern Madras (India) and adjacent northern Sri Lanka.

The leeward or Kona coast of Hawaii contains the only extensive area of summer maximum rainfall in the Hawaiian archipelago (Aw, winter-dry climate). Isolated from the prevailing trade wind flow by intervening high mountains, the Kona coast's dominant circulation pattern is formed by a localized land-sea breeze regime. Increased land surface temperatures in summer strengthen the daily sea breeze regime and increase convective instability, leading to a high frequency of afternoon thundershowers. The vertical structure necessary for thundershower development is further assured by the high mountains, which exclude the trade wind aloft and limit the potential for strong vertical wind shear. The presence of a strong shearing force would otherwise tend to destabilize these leeward convectional cells. Although there is generally a summer rainfall maximum throughout Kona, the Aw climate gives way to Af at elevations above 400 m, where, by virtue of general orographic position, there is adequate precipitation in all months.

Arid and Semi Arid Zones (B climates)

A classic rain shadow desert exists on the leeward side of the Kohala mountains. Smaller and lower (maximum elevation 1,670 m) than Mauna Kea and Mauna Loa, the Kohala mountains are incapable of blocking out trade wind flow to leeward. Having become depleted of moisture during windward ascent, the trades warm adiabatically to leeward, promoting a hot, arid zone. With only 190 mm of annual rainfall, Kawaihae on the leeward Kohala coast is the driest location in the Hawaiian archipelago. The Koppen system distinguishes two climatic subtypes, the true desert (BWh climate) and the semidesert (BSh climate) on the basis of relative aridity. In leeward Kohala the true desert gives way to semi-desert at higher elevations.

---

[1] The Temperate (C) and Polar (E) climates as originally proposed by Koppen were not applied in high-altitude tropical environments, which, because of their orographic complexity,

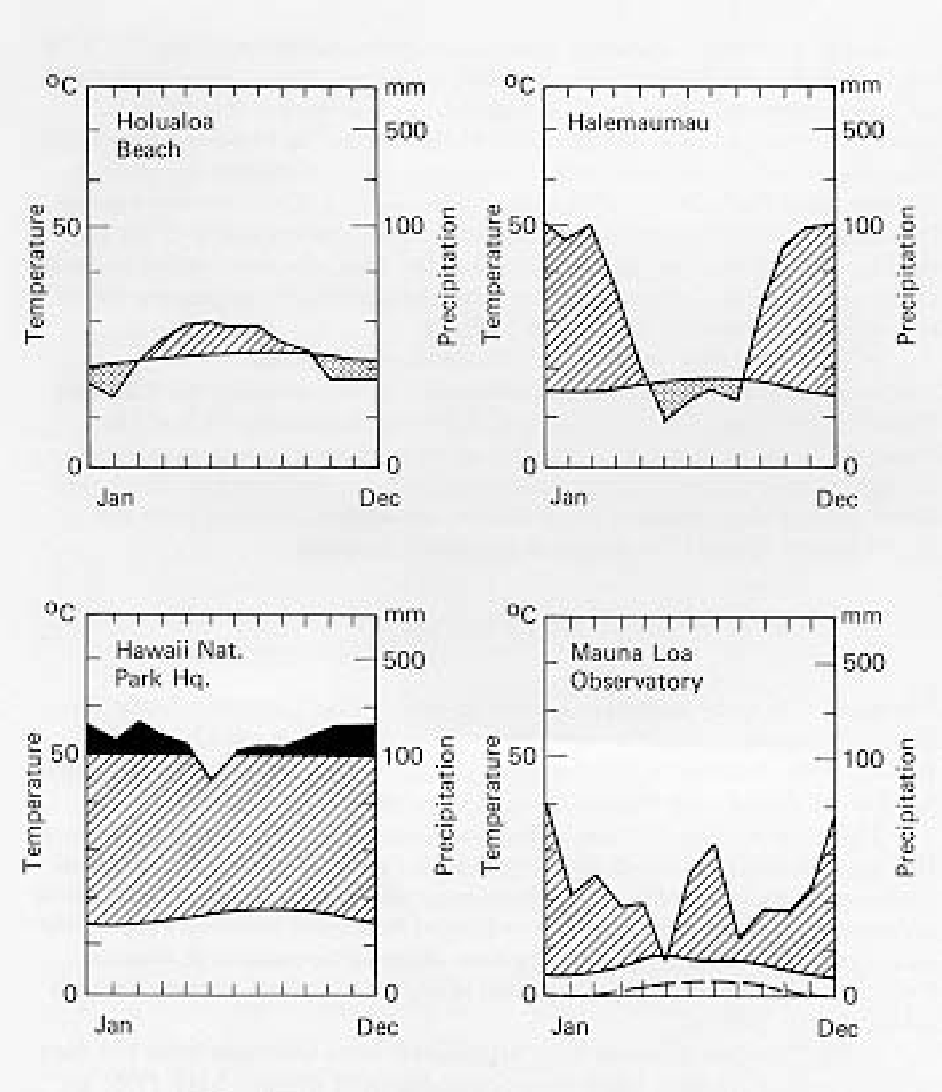

Figure 3. "Walter" climate diagrams for four Hawaii island stations (4).

Temperate Zone (C climates)[1]

Average air temperature in Hawaii decreases with altitude at the rate of about 0.55øC/100 m (Price, 1973). When the criteria of the Koppen classification are used, at elevations above 400–500 m on the mountain slopes, tropical climates grade to temperate as a result of decreasing average temperatures. As a result of the moderating influence of altitude, almost two-thirds of the "tropical" island of Hawaii possesses a temperate upland climate. The majority of this zone is characterized by warm summers and adequate precipitation in all months (Cfb climate). Except for the absence of a stronger season variability the upland Hawaii climates are analogous to those of similar Koppen designation in Pacific coastal areas of North America. Ascending orographic clouds compressed between the rising mountain slope and an upper-air temperature inversion produce frequent ground level mountain fog, an important moisture source for upland vegetation (Juvik and Perreira, 1974). At still higher elevations on Mauna Kea and Mauna Loa (above 2,000 m) there is a tendency toward summer drought. The increased strength and frequency of the trade wind inversion in summer (modal elevation 1,800 m) inhibits the vertical penetration of orographic clouds and precipitation to the higher slopes. This summer-dry zone (Gsb climate) also occurs at a lower elevation in leeward Kohala, Mauna Kea, and Mauna Loa, where summer orographic precipitation is largely absent. Above 2,500 m on both Mauna Loa and Mauna Kea the summer-dry regime changes from warm to cool (Gc climate).

Alpine (periglacial) Zone (E climates)

Above 3,200-m level on Mauna Kea and Mauna Loa all months have a mean temperature below 10 °C, and the climates are classified as periglacial (ET). Nighttime freezing is common throughout the year. Although it

---

lack of meteorological data, and absence of strong seasonality, were simply designated as highland climates (H). In more recent global and regional climatic maps of the tropics, highland C and E climatic areas are frequently portrayed in order to show approximate attitudinal analogs for these broad latitudinal climatic zones.

exhibits a winter maximum, annual rainfall is very low (200–400 mm) and variable. Above the 3,500-m level, winter snowfall accounts for a substantial portion of the seasonal precipitation. Koppen used the 10 °C (warmest month) boundary to separate the C and E climates on the basis that trees will not normally grow where mean temperatures fall below this level. Hence E climates characterized the treeless arctic tundra. The upper tree line of Mauna Kea (3,000 m) corresponds fairly closely to the C/E boundary mapped in Fig. 2. On Mauna Loa the tree line is much lower for edaphic reasons (recent lava).

WATER BALANCE

The preceding discussion of Koppen climatic zones on Hawaii provides a general overview of the dramatic range in regional climatic diversity found on the island. However, this descriptive approach says little about the direct linkage of climate to physical and biological processes at the earth/atmospheric interface.

An integration of seasonal moisture supply (precipitation) with the evaporation and transpiration demands of the environment (determined primarily by solar energy inputs) provides an index of moisture surplus or deficit. Such indices can illuminate direct process/response relationships between climate and the terrestrial ecosystem.

In an initial survey of water balance climatology on the island of Hawaii, Mueller-Dombois (4) constructed a series of climate diagrams 21 stations (see Fig. 3). This type of diagram, popularized by Walter and Lieth (5), portrays seasonal curves of mean monthly temperature and precipitation. According to Muller-Dombois (6) an index of precipitation efficiency is built into the diagrams by making one degree of temperature (Celsius) equal to two millimeters of precipitation in the scaling of the two ordinates. This is based on the assumption that monthly potential evapotranspiration (in millimeters) is roughly equal to twice the mean monthly temperature (7). Wherever the precipitation curve drops below the temperature curve, a drought season is indicated. Thus the graph is transformed into a water balance diagram with the temperature curve interpreted as an index of potential evapotranspiration.

A serious problem inherent in this graphing technique is the tendency to approximate evapotranspiration with a simple linear function of air temperature (i.e., the 2:1 ratio). Chang (8,9) has reviewed the problems of temperature-based estimations of potential evapotranspiration and points out that solar radiation rather than temperature is the primary forcing function in the evaporative process. Temperature-based estimates of evaporation implicitly assume a strong correlation between temperature and solar radiation. In Hawaii, as a result of advection and the buffering effect of the surrounding marine environment, there is generally poor correlation between solar radiation and temperature. In Hilo, for example, the range in mean monthly air temperature is only 1.4 °C between June (24.2 °C) and December (22.8 °C). By contrast, the receipt of solar radiation in June (563 Ly; see solar radiation data for 1965 from Ref. 10) is more than twice that in December (263 Ly).

It is obvious from the above comparison that temperature-based estimates of evapotranspiration cannot be expected to portray realistically the seasonal fluctuations implied in the radiation data. However, in the absence of a dense network of solar radiation monitoring stations on the island, upon which more sophisticated spatial modeling of evapotranspiration might be based, it is necessary to revert to some form of temperature-derived estimation in a “first approximation” of water balance regimes.

Thornthwaite (11) has developed perhaps the most widely adopted method of estimating potential evapotranspiration. His empirical formula is based essentially on air temperature:

$$E = 1.6\ (10T/I)^a. \tag{1}$$

Potential evapotranspiration E is computed from mean monthly temperature T and an empiric “heat index” I, which itself is an exponential function of temperature; a is a constant.

To obtain mean monthly evapotranspiration, the values derived from eq. (1) are corrected for mean daylength and number of days in the month. The Thornthwaite equation, although subject to the general limitations of all temperature-based methods, might be expected to give better results than the Walter method in Hawaii,

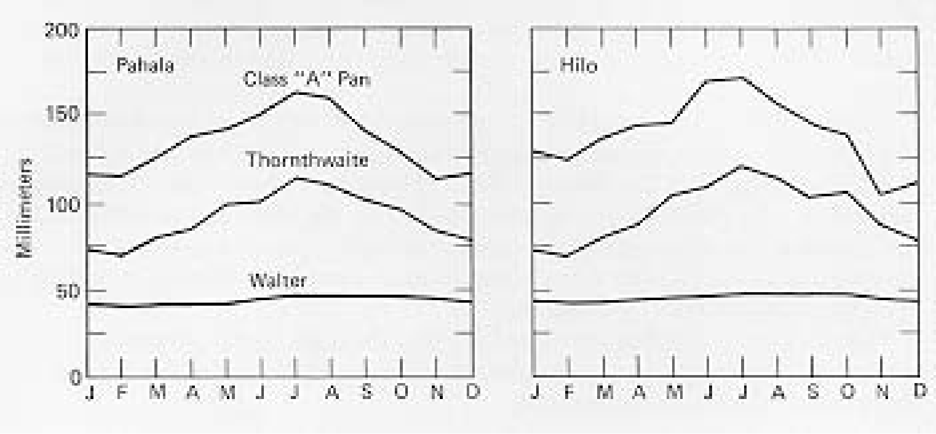

Figure 4. Monthly estimated potential evapotranspiration and measured pan evaporation for Hilo and Pahala data.

since potential evapotranspiration is expressed as an exponential rather than linear function of temperature.

In Fig. 4, monthly values of potential evapotranspiration derived by both the Walter and the Thornthwaite methods have been plotted along with class "A" pan-evaporation data for Hilo and Pahala. It is evident that Walter grossly underestimates pan evaporation (here assumed to be equal to potential evapotranspiration) and also fails to detect the seasonal rhythm apparent in the pan data. Thornthwaite also underestimates pan evaporation but does so in a fairly consistent manner and achieves a strong covariation with the pan data in seasonal rhythm. This suggests that the Thornthwaite method might be useful in Hawaii if a correction factor could be derived to compensate for the consistent underestimation exhibited in Fig. 4.

In Fig. 5, monthly values of the Thornthwaite potential evapotranspiration estimate are plotted against pan evaporation for Hilo and Pahala. With a regression coefficient of 0.844, approximately 71% of the observed variation in pan evaporation can be explained by variation in the Thornthwaite estimate. (The regression coefficient is significant at 0.01 level.) Potential evapotranspiration Y can thus be reasonably predicted from the Thornthwaite values X by the linear regression equation

$$Y = 42.8 + 1.016\,(X) \tag{2}$$

Before eq. (2) can be applied as a general (island-wide) correction factor for the Thornthwaite potential evapotranspiration estimate, it must be verified that the relationship established in Fig. 5 (for two lowland locations) is equally valid for mid- and high-altitude areas of the island.

There are no class "A" pan evaporation data for inland mountain areas of Hawaii with which the lowland-derived correction factor might be compared. However, Juvik and Clarke (12) have accumulated limited experimental data on mountain evaporation gradients in Hawaii Volcanoes National Park on the east flank of Mauna Loa.

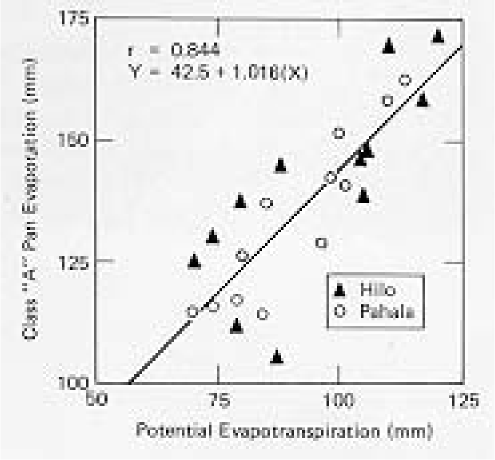

Figure 5. Relationship between monthly class "A" pan evaporation and estimated monthly potential evapotranspiration (Thornthwaite) for Hilo and Pahala data.

These data were obtained by using four constant-level pan evaporimeters (Fig. 6) situated along an attitudinal transect between sea level and 2,000 m.

In Fig. 7, measured mean daily evaporative rates (averages from 133 days of simultaneous readings taken from September 1974 through May 1975) are plotted against elevation. There is a clear linear decrease in evaporation over the attitudinal range surveyed (approximately 0.72 mm/day/1,000 m). Figure 7 also shows the corrected (eq. 2) Thornthwaite potential evapotranspiration values (mean of 9 months, September to May) derived from temperature-recording stations that occur near the evaporimeter transect. There is good agreement between the Thornthwaite and the evaporimeter values (differences range from 1% to 12%), largely because air temperature also decreases linearly with elevation. On the basis of the close agreement in Fig. 7, the corrected Thornthwaite estimate was considered acceptable to use for all areas of the island

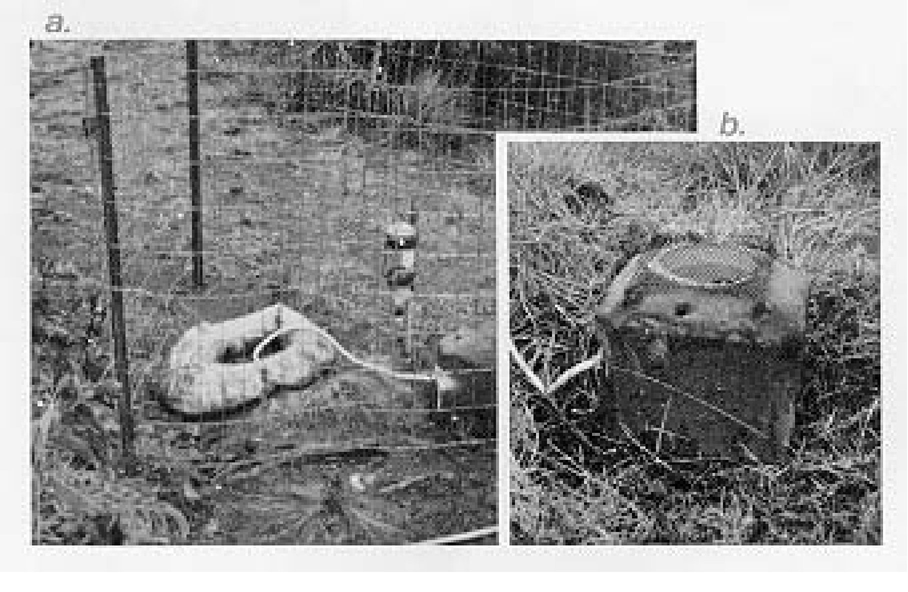

Figure 6. Constant-level pan evaporimeter. a) field installation with inner tube reservoir; b) evaporimeter detail. Note foam insulation around evaporation pan.

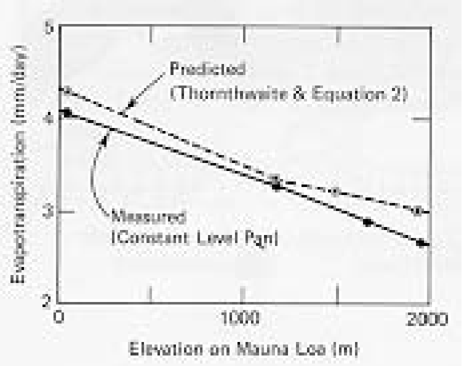

Figure 7. Relationship between measured and predicted evapotranspiration along an attitudinal transect in Hawaii Volcanoes National Park on the east flank of Mauna Loa (pan data from 12).

in the derivation of monthly and annual potential evapotranspiration from temperature data.

Corrected Thornthwaite estimates of monthly and annual potential evapotranspiration were computed from standard tables (13) and eq. (2), for 30 stations on the island of Hawaii. The evapotranspiration data were then integrated with monthly precipitation values to produce seasonal water balance diagrams (Fig. 8). Because some of the water surplus received in the wet season is stored as soil moisture for utilization during dry periods, the computation of seasonal water balance must incorporate a parameter describing the moisture storage capacity of the soil.

On the geologically youthful island of Hawaii, soils are not generally well developed except for limited areas where ash deposits occur in high-rainfall zones. Recent lava flows exhibiting little or no soil development cover substantial portions of the island. The average depth to bedrock for 75 different Hawaii island soil types and subtypes has been calculated at 0.89 m with only moderate variation (14), lending quantitative credence to this stated geological youthfulness of the island. Soil moisture storage capacity has not been well studied for most Hawaiian soil types. A value of 125 mm/m is the average moisture capacity for ten different soil types for which data are available. If this is assumed to be representative, then the average soil moisture storage capacity for all Hawaii island locations would be 111.2 mm (i.e., 0.89 × 125). In the construction of water balance diagrams for all island stations this value was rounded off to 100 mm so that standard moisture depletion tables could be employed in water balance calculations (13).

The 30 water balance diagrams constructed for the island depict both steep gradients and pronounced regional differences in seasonal moisture surplus and deficit. In Fig. 8 the difference between annual precipitation and potential evapotranspiration has been mapped in four zones:

1. Annual surplus exceeding 1,000 mm. This zone comprises 20% (2,100 km^2) of the island area and is restricted to the high-rainfall regions of windward Mauna Kea, Mauna Loa, and the summit area of Kohala. The annual moisture surplus in this zone

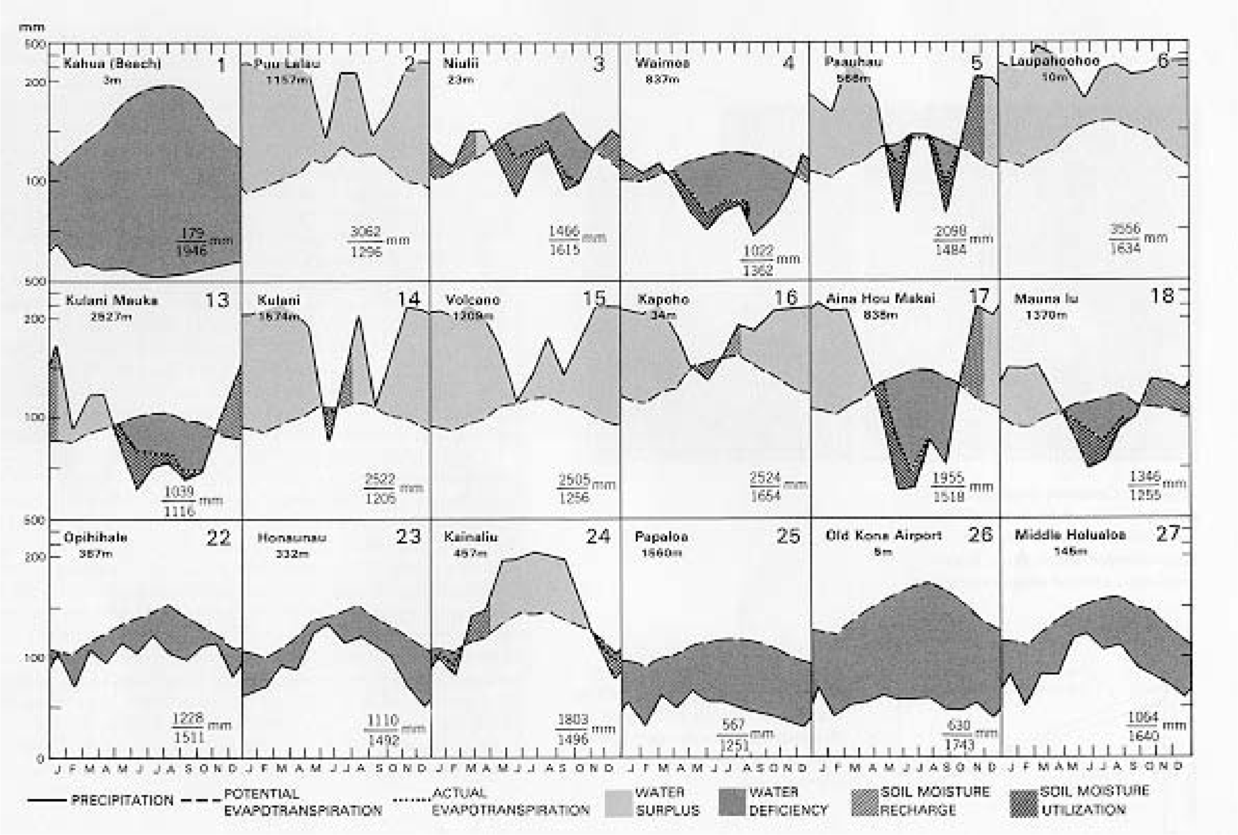

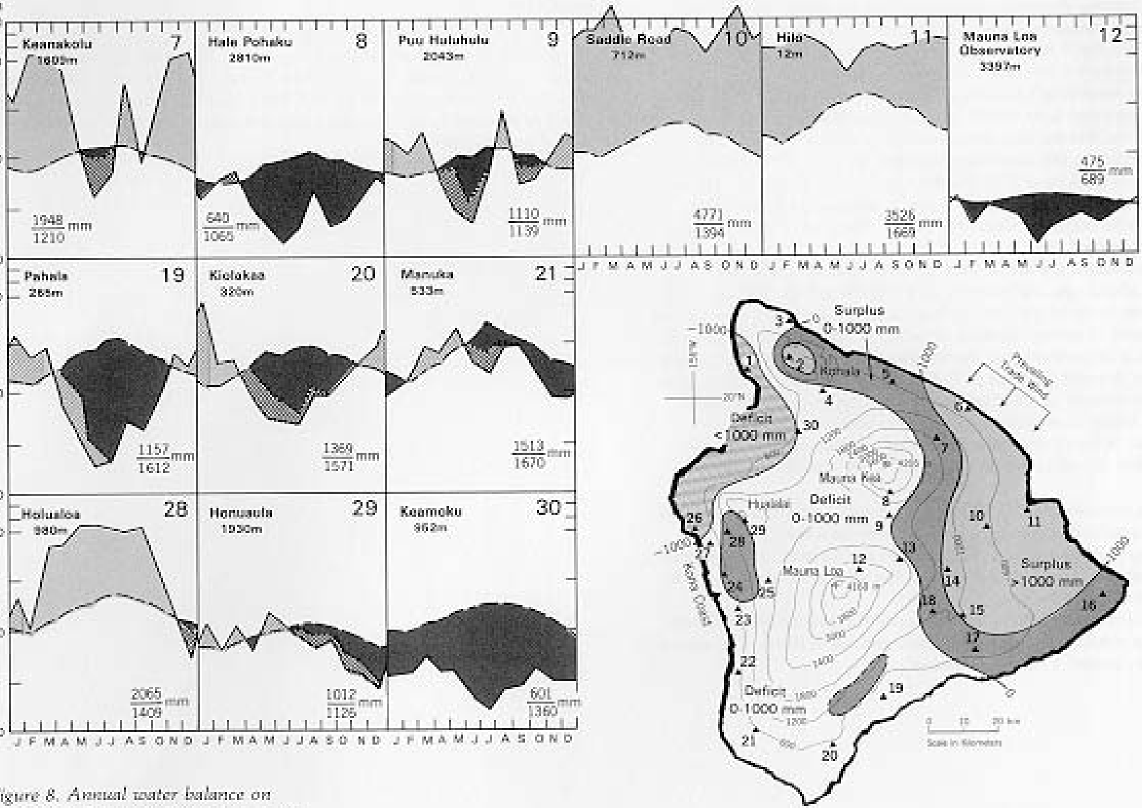

Figure 8. Annual water balance on the island of Hawaii. Ratios are mean annual precipitation over mean annual potential evapotranspiration in millimeters.

ranges as high as 3,377 mm (station 10) at middle elevations. All stations within this zone (stations 2, 6, 10, 11, 14, and 15) exhibit an absolute winter maximum in precipitation, and a secondary summer maximum also occurs at elevated stations where summer orographic precipitation is exaggerated (e.g., stations 2 and 14).

2. Annual surplus between 0 and 1,000 mm. This zone comprises 21% (2,200 km^2) of the island area and extends from middle to high elevations on the windward slopes down to sea level in those areas where slope aspect is not oriented perpendicular to prevailing trade wind flow, and thus the orographic rainfall component is diminished. For the windward stations there is typically a moderate summer drought (stations 5, 7, 9, 13, and 17) from 2 to 5 months long. The increased strength of the trade wind inversion in particular limits summer rainfall at higher elevations. The localized core area of high convectional rainfall in Kona also falls within this moisture zone. However, here the deficit period occurs in winter (stations 24, 28, and 29) and is not severe.
3. Annual deficit between 0 and 1,000 mm. This zone comprises 54% (5,600 km^2) of the island area and occupies a predominantly leeward location on Kohala, Mauna Kea and Mauna Loa. The drought period may be concentrated in either the summer (on the windward side) or the winter (on the Kona side) and is typically 6 to 12 months long.
4. Annual deficit exceeding 1,000 mm. This zone comprising 5% (550 km^2) of the island area is restricted to leeward Kohala and north Kona. The annual moisture deficit may exceed 1,900 mm (station 1).

SUMMARY

The Thornthwaite water balance diagrams and map demonstrate graphically the tremendous climatic diversity on the island of Hawaii. Although the Koppen map (Fig. 2) shows only a relatively small portion of the island to be arid or semi-arid (BWh and BSh), from the water balance analysis it is evident that nearly 60% (zones 3 and 4 above) of the island experiences an annual moisture deficit.

Acknowledgments

This research was supported in part by grants from the Hawaii Natural History Association and the U.S. Department of Interior, Office of Water Resources Research.

BIBLIOGRAPHY

1. Carter, D.B. and Mather, J.R. (1966). Climatic classification and environmental biology. *Publ. Climatol.* **19**(4): 305–390.
2. Taliaferro, W.J. (1959). *Rainfall of the Hawaiian Islands*. Hawaii Water Authority, p. 394.

3. State of Hawaii (1970). *An inventory of basic water resources dataÄisland of Hawaii*. Rep. R34, Dept. of Land and Natural Resources, Honolulu, Hawaii, p. 188.
4. Mueller-Dombois, D. (1966). Climate. Chap. IV. In: *Atlas for Bioecology Studies in Hawaii Volcanoes National Park*. Maxwell S. Doty and D. Mueller-Dombois. (Eds.). U.S. National Park Service, p. 507. (Republished as Hawaii Agr. Exp. Sta. Bull. 89, 1970.)
5. Walter, H. and Lieth, H. (1960). *Klimadiagram-Weltatlas*, Jena.
6. Mueller-Dombois, D. (1976). *The Major Vegetation Types and Ecological Zones in Hawaii Volcanoes National Park and Their Application in Park Management and Research*. In Proceedings of the First Conference in Natural Sciences, Hawaii Volcanoes National Park, edited by S. W. Smith, Department of Botany, University of Hawaii, Honolulu, 149–161.
7. Gaussen. (1954).
8. Chang, J.-H. (1959). An evaluation of the 1948 Thornthwaite classification. *Ann. Assoc. Amer. Geogr.* **49**(1): 24–30.
9. Chang, J.-H. (1968). *Climate and Agriculture*. Aldine Publishing Co., Chicago, p. 304.
10. Lof, G.O.C., Duffie, J.A., and Smith, C.O. (1966). *World Distribution of Solar Radiation*. Rep. 21, Solar Energy Laboratory, Univ. of Wis., p. 59. (plus maps).
11. Thornthwaite, C.W. (1948). An approach toward a rational classification of climate. *Geogr. Rev.* **38**: 55–94.
12. Juvik, J.O. and Clarke, G.G. (1976). *Topoclimatic Gradients in Hawaii Volcanoes National Park (abstract)*. In Proceedings of the First Conference in Natural Sciences, Hawaii Volcanoes National Park, edited by S. W. Smith, Department of Botany, University of Hawaii, Honolulu, p. 113.
13. Thornthwaite, C.W., and J.R. Mather (1957). Instructions and tables for computing potential evapotranspiration and the water balance. *Publ. Climatol.* **10**(3): 1–311.
14. Sato, H.H. et al. (1973). *Soil Survey of Island of Hawaii, State of Hawaii, U.S.D.A.* Soil Conservation Service, p. 115. (plus maps).

READING LIST

Juvik, J.O. and Perreira, D.J. (1974). Fog interception on Mauna Loa, Hawaii. *Proc. Assoc. Amer. Geogr.* **6**: 22–25.

Price, S. (1973). Climate. In: *Atlas of Hawaii*. R.W. Armstrong. (Ed.). University of Hawaii Press, Honolulu, pp. 53–60.

HEAT OF VAPORIZATION

Naraine Persaud
Virginia Polytechnic Institute
and State University
Blacksburg, Virginia

HEAT OF VAPORIZATION

The heat of vaporization (here denoted as L_v) is defined as the heat added (or given off) when unit mass undergoes isobaric phase transformation in any closed two-phase, one-component liquid/vapor system. In engineering and meteorology, L_v is used in a restricted sense to mean the heat of vaporization of the two-phase liquid water/water vapor system. Although much of the subsequent discussion focuses on L_v, for this specific system as an example, the concepts covered are universally applicable to all fluid/vapor systems.

As illustrated below (Fig. 1), for the liquid water/water vapor system, L_v represents the heat gained when unit mass of water in the system evaporates in the isobaric phase transformation H_2O (liquid) → H_2O (vapor). For the reverse phase change H_2O (vapor) → H_2O (liquid), i.e., condensation, L_v is lost from the system. This seemingly simple phase transition H_2O (liquid) ↔ H_2O (vapor) is the fundamental driving process of the earth's hydrological cycle, the working principle of the steam engine that ushered humanity into the industrial revolution along with its (often negative) social and environmental pollution consequences, and the physical mechanism that maintains the body temperature of plants and warm-blooded animals.

In general, the state of any closed two-phase, one-component system is defined by the state variables temperature (T in °K), saturation vapor pressure (P in Pascal), and volume (V in m^3). The behavior of any such system (generally termed as PVT systems) is usually represented as a family of experimental constant temperature curves (isotherms) on a P-V coordinate plane called an Amagat–Andrews diagram. The general shape of these experimental isotherms is illustrated below (Fig. 2).

For the liquid water/water vapor system, the liquid and vapor phases coexist in equilibrium only at P-V coordinates between 2 and 3 along an isotherm, provided

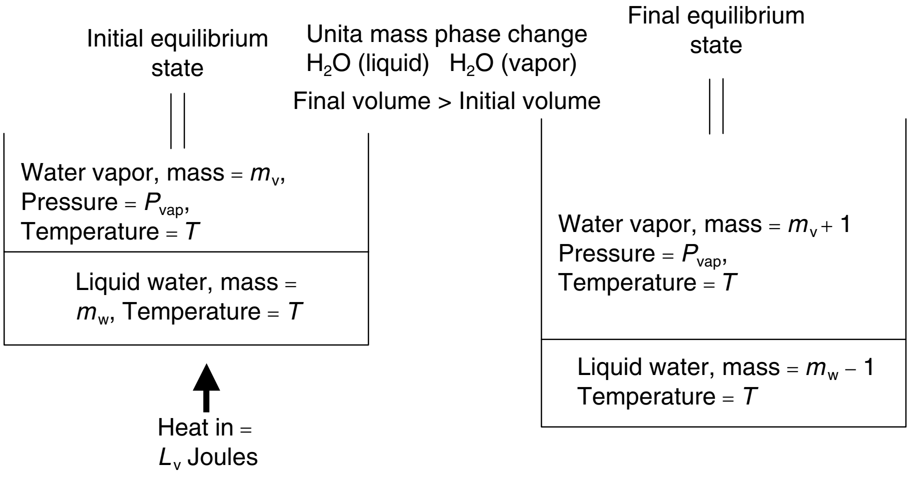

Figure 1. Liquid Water/Water Vapor System.

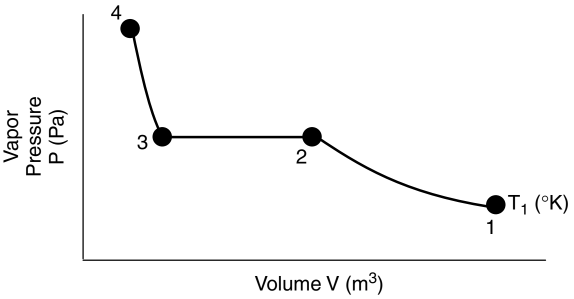

Figure 2. Amagat–Andrews Diagram.

that T_1 is above the triple point temperature of water (0.01 °C, i.e., the temperature at which ice, liquid water, and water vapor can coexist in equilibrium), and below the critical temperature (374 °C, i.e., the temperature above which it is impossible to produce condensation by increasing the pressure). Between 2 and 1, the system can exist as vapor only, and as liquid between 3 and 4. Thus the liquid-vapor phase transition at a given temperature can only take place at constant pressure or vice-versa. Consequently, as shown in Fig. 1, isobaric liquid vaporization and condensation is necessarily an isothermal process, implying that the triple-point saturation vapor pressure is fixed (it is 611 Pa), and so is the saturation vapor pressure at the critical temperature (it is 2.21×10^7 Pa $= 218.2$ atm). Similar isotherms and parameters exist for all liquid/vapor systems.

This observed behavior of closed two-phase, one-component systems is of course predicted by Gibb's Phase Rule (1), namely $F + N = C + 2$, where $F =$ degrees of freedom, i.e., the smallest number of intensive variables (such as pressure, temperature, concentration of components in each phase) that must be specified to completely describe the state of the system; $N =$ number of phases, i.e., distinct subsystems of uniform chemical composition and physical properties; and $C =$ the number of components, i.e., the number of independent chemical constituents meaning those constituents whose concentration can be varied independently in the different phases. In a liquid/vapor system $P = 2$ and $C = 1$, and therefore $F = 1$, implying only one intensive variable is needed to specify the state of the system. Therefore, temperature and pressure cannot be fixed independently. For the liquid water/water vapor system, this means physically that at a given temperature between the triple point and critical temperature, water vapor will evaporate or condense to achieve the equilibrium saturation vapor pressure as would be evidenced in a complete Amagat–Andrews diagram for water (2,3).

The earth's atmosphere and oceans can be considered as a vast closed two-phase liquid water/moist air system. Consequently, for most practical engineering and meteorological applications, one is interested in the heat of vaporization of the liquid water/moist air system rather than a pure liquid water/water vapor system. Fortunately, the presence of the other gases (collectively called dry air) in the liquid water/moist air system has negligible effect on the saturation vapor pressure. The reason is that the dry air component in the liquid water/moist air system remains unchanged and is always in the gaseous state during phase transition at temperatures and pressures of practical interest. Therefore, it can be considered as a closed subsystem as opposed to the open liquid water and water vapor subsystems. Consequently, results obtained from an analysis of the thermodynamics of the pure system are applicable to the natural liquid water/moist air system.

Energy conservation required under the first law of thermodynamics implies that heat (Q) exchanged reversibly with the surroundings between equilibrium states of any closed two-phase PVT system is consumed by any internal energy change (ΔU) of the liquid and vapor phases associated with the mass change from one phase to the other, and any mechanical work ($\pm P\Delta V$) realized as the volume of the system increases (positive work) or decreases (negative work). Stated mathematically, $Q = \Delta U + P\Delta V$, or in a differential form, $\delta Q = dU + pdV$. Here, P is the saturation vapor pressure (P_{vap} in Fig. 1). As entropy (S) is defined as Q/T, then $\delta Q = TdS$. The first law can therefore be restated in terms of exact differentials as $TdS = dU + PdV$. Dividing by dV at constant T and rearranging gives $dU/dV = T(dS/dV) - P$. Using the Maxwell relation $(\partial S/\partial V)_{\mathrm{T}} = (\partial P/\partial T)_{\mathrm{V}}$ (1), dS/dV can be replaced (for fixed T) by dP/dT. The equation becomes $dU/dV = T(dP/dT) - P$.

At a given pressure and temperature, the internal energy of the system (U in Joules) can be partitioned as $m_{\mathrm{w}}u_{\mathrm{w}} + m_{\mathrm{vap}}u_{\mathrm{vap}}$, where m_{w}, m_{vap} and u_{w}, u_{vap} represent the masses and the specific internal energies (internal energy per unit mass in J kg^{-1}) of the water and water vapor in the system. Similarly, the volume (V in m^3) of the system can be partitioned as $m_{\mathrm{w}}v_{\mathrm{w}} + m_{\mathrm{vap}}v_{\mathrm{vap}}$, where v_{w}, v_{vap} represent the specific volume (volume per unit mass in J kg^{-1}) of the water and water vapor in the system. If, as illustrated in Fig. 1, the system internal energy changes by ΔU from U to $U + \Delta U$ as a result of L_{v} Joules of heat absorption to convert unit mass of water to water vapor, then $U + \Delta U = (m_{\mathrm{w}} - 1)u_{\mathrm{w}} + (m_{\mathrm{w}} + 1)u_{\mathrm{vap}}$, and therefore $\Delta U = (u_{\mathrm{w}} - u_{\mathrm{vap}})$. Similar reasoning shows that, if the volume changes from V to $V + \Delta V$ in the process, then $\Delta V = (v_{\mathrm{vap}} - v_{\mathrm{w}})$. The mechanical work because of volume change is $P\Delta V$, where P (the saturation vapor pressure) is a constant at a fixed temperature. Therefore, the heat absorbed (or released) by the system for isobaric phase transition of unit mass in the liquid water/water vapor system ($L_{\mathrm{v}}$by definition) $= \Delta U + P\Delta V$. Dividing by ΔV gives $\Delta U/\Delta V = (L_{\mathrm{v}}/\Delta V) - P$. Substituting $\Delta V = (v_{\mathrm{vap}} - v_{\mathrm{w}})$ gives $\Delta U/\Delta V = [L_{\mathrm{v}}/(v_{\mathrm{vap}} - v_{\mathrm{w}})] - P$, or in a differential form, $dU/dV = [L_{\mathrm{v}}/(v_{\mathrm{vap}} - v_{\mathrm{w}})] - P$. [It should be noted that because enthalpy (H) is defined as $H = U + PV$, then $\Delta U + P\Delta V = \Delta H$, and therefore L_{v} is the same as the specific enthalpy change ($\Delta h = \Delta H$ per unit mass) for phase transition of unit mass in the liquid water/water vapor system.]

Combining the results for dU/dV from the two previous paragraphs gives $T(dP/dT) - P = [L_{\mathrm{v}}/(v_{\mathrm{vap}} - v_{\mathrm{w}})] - P$, and therefore $T(dP/dT) = L_{\mathrm{v}}/(v_{\mathrm{vap}} - v_{\mathrm{w}})$, which can be rearranged to obtain the general forms of the Clapeyron (Emile Clapeyron, 1799–1864) equation $dP/dT = L_{\mathrm{v}}/[T(v_{\mathrm{vap}} - v_{\mathrm{w}})] = \Delta h/[T(v_{\mathrm{vap}} - v_{\mathrm{w}})]$ or $L_{\mathrm{v}} = [T(v_{\mathrm{vap}} - v_{\mathrm{w}})]dP/dT$.

The Clapeyron equation can be used to obtain L_v at a given temperature T for any liquid provided one can obtain values of $(v_{\text{vap}} - v_w)$ and an accurate representation of dP/dT. Values of $(v_{\text{vap}} - v_w)$ can be obtained from tabulated measurements (3). Alternatively, because $v_{\text{vap}} \gg v_w$ at the low pressures, $(v_{\text{vap}} - v_w)$ can be taken as equal to v_{vap}. Assuming further that at low pressures water vapor behavior closely approximates that of an ideal gas, then $v_{\text{vap}} = RT/P$, where R is the specific gas constant for water $= 8.314/0.018 = 461.9$ J kg^{-1} °K^{-1}. The Clapeyron equation becomes $L_v = (RT^2/P)dP/dT$, and this form is referred to as the Clausius–Clapeyron equation (Rudolph Clausius, 1822–1888).

Unfortunately, no single function has been shown to adequately represent vapor pressure data for various liquid/vapor systems over wide ranges of T. For the liquid water/water vapor system, values of dP/dT for a given value of T can be obtained by finite differencing of tabulated values. As an example, consider using the above equation to estimate L_v for water at human body temperature of 37 °C. Tabulated values are $v_w = 1.007$ cm$^3 g^{-1}$ = 0.01007 m^3 kg^{-1}, $v_{\text{vap}} = 22{,}760$ L kg^{-1}, = 22.760 m^3 kg^{-1}, $P = 5.940$ kPa at 36 °C and 6.624 kPa at 38 °C, which gives $(v_{\text{vap}} - v_w) = 22.760 - 0.01007 = 22.75$ m^3 kg^{-1}. By finite differencing, $dP/dT \approx \Delta P/\Delta T = (6.624 - 5.940)/(38 - 36) = 0.342$ kPa °C^{-1} = 0.342 kPa °K^{-1} = 342 Pa °K^{-1} (a temperature difference of 1 °C is the same as a difference of 1°K). At $T = 37 + 273.16 = 310.16$°K, $L_v = 310.16$°K (22.75 m^3 kg^{-1}) ×342 Pa °K^{-1} = 2399 kJ kg^{-1}. The actual value (tabulated as Δh) is 2414 kJ kg^{-1}, an error of <1%.

Sometimes it is more efficient to have semianalytical expressions for $P(T)$ and hence dP/dT. Examples of such expressions for $P(T)$ are the Antoine equation, $\ln P = A + [B/(T + C)]$, and the Rankine equation, $\ln P = A + B/T + C \ln T$, where T is (usually) in °K, and A, B, C are empirical parameters obtained by fitting these equations to experimental data on phase equilibrium for a specific liquid/vapor system for various temperature ranges. Antoine's equation coefficients for the vapor pressure of quite a number of chemical compounds are available online in the NIST Chemistry Webbook (3). This reference lists Antoine coefficients for liquid water based on the vapor pressure data of Bridgeman and Aldrich (4) over several temperature ranges as:

Antoine Equation Coefficient	Temperature Range (°K)			
	273 to 303	304 to 333	334 to 363	344 to 373
A	5.40221	5.20389	5.0768	5.08354
B	1838.675	1733.926	1659.793	1663.125
C	−31.737	−39.485	−45.854	−45.622

These coefficients apply to Antoine's equation in the form $\log_{10} P = A - [B/(T + C)]$ with P in bar and T in °K. At 37 °C, this equation gives $P = 6.280$ kPa compared with the measured tabulated value of 6.274 kPa. The derivative of Antoine equation in this case is $dP/dT = P[2.3026B/(T + C)^2]$, which gives dP/dT at 37 °C $= 6.280 \times 0.0545 = 0.342$ kPa °K^{-1}, which is the same as obtained by finite differencing of tabulated values.

For the liquid water/water vapor system, the equation for $P(T)$ developed by Richards (5) covers a wide temperature range and gives estimates for dP/dT close to those obtained by finite differencing of tabulated data. This equation is:

$$P(T) = ae^{f(T_r)}$$

where P is in millibar (mb), e is the base of natural logarithms, $T_r = 1 - \dfrac{373.15}{T}$, T is the temperature in °K, and $a = 1013.25$. The function $f(T_r) = 13.3185T_r - 1.9760\ (T_r)^2 - 0.6445\ (T_r)^3 - 0.1299\ (T_r)^4$. The derivative (with respect to T) is:

$$\frac{dP}{dT} = P(\frac{373.15}{T^2})(\frac{df(T_r)}{dT_r})$$

At 37 °C, $T_r = -0.2031$, the calculated value of $P = 62.78$, $mb = 6.278$ kPa compared with the tabulated measured value of 6.274 kPa. Using $P = 62.78$, mb gives $dP/dT = 62.78 \times 0.00388 \times \{13.3185 - (2 \times 1.9760 \times -0.2031) - [3 \times 0.6445 \times (-0.2031)^2] - [4 \times 0.1299 \times (-0.2031)^3]\} = 0.2434 \times \{13.3185 + 0.8026 - 0.0797 + 0.0044\} = 0.2434 \times 14.046 = 3.419$ mb °K^{-1} = 341.9 Pa °K^{-1}, practically the same as the value of 342 Pa °K^{-1} obtained above from finite differencing of tabulated data.

Applying the Clausius–Clapeyron equation with $P = 6.280$ kPa at 37 °C by the Antoine equation, gives $L_v = 461.9 \times [(310.16)^2/6.280] \times 0.342 = 2{,}420$ kJ kg^{-1} compared with the actual value of 2414 kJ kg^{-1}.

BIBLIOGRAPHY

1. Cengel, Y.A. and Boles, M.A. (2002). *Thermodynamics: An Engineering Approach*, 4th Edn. McGraw-Hill Series in Mechanical Engineering, Boston, MA.
2. Wagner, W. and Pruss, A. (2002). The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use. *J. Phys. Chem. Ref. Data* **31**: 387–535.
3. Lemmon, E.W., McLinden, M.O., and Friend, D.G. (2003). Thermophysical properties of fluid systems. In: *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*. P.J. Linstrom and W.G. Mallard (Eds.). National Institute of Standards and Technology, Gaithersburg, MD. Available: http://webbook.nist.gov.
4. Bridgeman, O.C. and Aldrich, E.W. (1964). Vapor pressure tables for water. *J. Heat Transfer* **86**: 279–286.
5. Richards, J.M. (1971). Simple expression for the saturation vapor pressure of water in the range −50 to 140 °C. *Brit. J. Appl. Phys.* **4**: L15–L18.

HYDROLOGIC HISTORY, PROBLEMS, AND PERSPECTIVES

Jason A. Hubbart
University of Idaho
Moscow, Idaho

HYDROLOGIC HISTORY

Since humans first settled along the banks of lakes and rivers, there has been a keen interest in the proper management of fresh water resources both as a necessity for life as well as to avoid potential health hazards. Along the Indus in Pakistan, the Tigris and Euphrates in Mesopotamia, the Hwang Ho in China, and the Nile in Egypt, the first "hydraulic engineers" created canals, levees, dams, subsurface water conduits, and wells as early as 5000–6000 years ago. Hydrologic information became vital to these early civilizations. The flow rates and yields of rivers were monitored by the Egyptians as early as 3800 years ago, and rainfall measuring instruments were first used approximately 2400 years ago by Kautilya of India (1).

The idea that a global hydrologic cycle actually exists dates back at least 3000 years when early Greek philosophers such as Thales, Anaxagoras, Herodotus, Hippocrates, Plato, and Aristotle conceptualized the basic ideas governing the process (2). Many initial ideas established by the Greeks about the hydrologic cycle were quite reasonable. However, many of the initial mechanisms concerning the routes by which water returned from the sea and entered rivers were devoid of as much logic. Despite apparent gaps in hydrologic mechanisms, the Romans developed aqueduct systems reflecting an extensive practical understanding of hydrology and hydraulics and did so using the basic hydrologic ideas established and passed along by the Greeks (3). During the Renaissance, Leonardo da Vinci (1500) in France proclaimed, on the basis of field observations, that the waters in rivers come from precipitation. During that time, any unrealistic mechanisms proclaimed by the Greek philosophers concerning the hydrologic cycle were either refuted or modified (4,5). In the seventeenth century, the modern scientific approach to studying the hydrologic cycle was assumed by the Frenchmen Pierre Perault and Edme Marriotte. By the 1670s and 1680s, they had published data and calculations to support the contention that precipitation was the precursor to streamflow. By 1700, Edmun Halley, an English scientist, added to the work of Perault and Marriotte by estimating the quantity of water involved in the hydrologic cycle of the Mediterranean Sea and surrounding lands.

Substantial progress was made during the eighteenth century in applications of mathematics, fluid mechanics, and hydraulics by scientists like Pitot, Bernoulli, Euler, Chezy, and other European professionals. The term "hydrology" arrived at its current meaning around 1750, and by 1800, the work of English physicist and chemist John Dalton solidified the current understanding of the global hydrologic cycle (2).

Until the 1800s, the physical processes governing groundwater flow had confounded scientists and created barriers to understanding the hydrologic cycle. These barriers were eliminated in 1856, when the French engineer, Henry Darcy, introduced his law describing flow through porous media. Other advances in the hydrologic sciences were made throughout the 1800s. Poiseuille, DuPuit, DuBoys, Stokes, Manning, Reynolds, and others contributed substantially to fluid mechanics, hydraulics, and sediment transport during this period of time (3). Also during the 1800s, literary publications began to surface, increasingly frequently in the last half of the century. Many works examined relationships between precipitation and stream flow out of necessity for engineers designing bridges and other structures. During this time, the close association between hydrology and civil engineering was established. Daniel Mead published the first English-language text in hydrology in 1904, and Adolf Meyer followed with his text in 1919. Both texts were written for civil engineers (1). The association of hydrology and civil engineering established during this time, it has been argued, both enhanced as well as possibly inhibited the development of hydrology as a science (6).

The first half of the twentieth century saw great advancements in the hydrologic sciences starting with the addition of the Section of Scientific Hydrology in the International Union of Geodesy and Geophysics in 1922. This was followed by the addition of the Hydrology Section of the American Geophysical Union in 1930. These were the first formal recognitions of the scientific status of hydrology. Many individuals contributed substantially in their areas of hydrologic expertise during the early and middle decades of the twentieth century, including A. Hazen, E. J. Gumbel, H. E. Hurst, and W. B. Langbein in statistical applications of hydrologic data; O. E. Meinzer, C. V. Theis, C. S. Slichter, and M. K. Hubbert who contributed to the development of theoretical and practical aspects of groundwater hydraulics; L. Prandtl, T. Von Karman, H. Rouse, V. T. Chow, G. K. Gilbert, and H. A. Einstein in sediment transport and stream hydraulics; R. E. Horton and L. B. Leopold who contributed greatly to runoff processes and quantitative geomorphology; W. Thornthwaite and H. E. Penman in furthering the understanding of hydroclimatologic processes and modeling evapotranspiration; and A. Wolman and R. S. Garrels who contributed greatly to understanding and modeling water quality. It was not until the 1960s that detailed field studies attempting to understand the physical processes by which water enters streams began to emerge (3).

At the emergence of the twenty-first century, many new breakthroughs in the hydrologic sciences are imminent. In the forthcoming years, breakthroughs will describe the relationships between hydrologic regimes to current and future climate change and the effects of hydrologic processes on landform development. New findings will also include modeling regional evapotranspiration rates and geomorphologic water transport (1).

Population growth and land use changes have resulted in sizable impacts on surface water and groundwater resources. Increasing demands for fresh water has led to water stress in many parts of the world. Increasing quantities of water are used for agriculture, and increasing water quantities withdrawn from subsurface water bodies and large rivers have had significant influences on the water cycle. Population growth and land use changes have led to polluting of surface waters and groundwater resources as well as soil deterioration through erosion, salinization, and of course pollution. The ability to meet challenges of increased demand and to propose mitigating strategies will require intensive studies looking into how population growth and associated land use change and water demands influence the local, regional, and global water cycles.

HYDROLOGIC PROBLEMS AND PERSPECTIVES

Many specific scientific problems or challenges related to hydrologic study exist. These include assessing the impacts that growing populations have on water demand, land use change on groundwater reserves, and river discharges. Once successfully concluded, these assessments must be used as tools for predicting the extent of expected water stress in the coming decades due to changes in the supply and demand of water resources. Scientists must strive to understand the extent to which increased evaporation and transpiration through increased withdrawal of groundwater and surface water will lead to changes in local and global water cycles. There is a growing need for improved understanding of pollutant related flows in groundwaters and the self-cleaning capacities of aquifers and rivers related to different pollutants. Thus, further work is needed to describe the impacts of population growth and land use change on the quality of groundwater reserves, soils, and rivers and also what the effects may be on the natural ecosystems that depend on them.

Studies will be needed to understand landscape processes better, including hill slope forms, channel networks, and the processes responsible for transporting water, sediments, and pollutants. Hydrologic research in past decades has focused attention mainly on modeling the dynamic response of channel networks. Progress in understanding catchment scale flow will require that relevant hill slope flow also be carefully studied. This will likely require collaborative research approaches among hydrologists, geomorphologists, geologists, pedologists, and others to identify scaling relationships in flow domain structures.

There remains a need for improved understanding of the interactions between climate change and the hydrologic cycle. Further understanding and documentation is required to predict apparent changes that have led to an intensification of the hydrologic cycle due to climate change. Critical questions need to be asked concerning current changes observed in water availability and whether critical regions can be identified where the interactions between water availability and human demand become problematic. Conducting this type of research necessitates better understanding of past climates and paleohydrologic behavior. At present, the ability to predict changes in the fluxes of the water cycle under various scenarios of climate change is poor. Therefore, it is desirable to create improved methods of portraying the interactions between water systems and natural and human induced environments.

Due to the inherently complex nature of hydrologic systems, the characterization and subsequent modeling of these systems is a continual challenge to scientists. There are a number of unresolved issues that continue to impede the ability of scientists to analyze and predict behavioral changes of hydrologic systems. One of the most predominant features of hydrologic systems is the spatial heterogeneities and temporal variabilities that occur perpetually and persistently on multiple scales. Modeling physical processes encumbered by such diversity is a constant challenge. For these reasons, theoretical upscaling and downscaling attempts to develop quantitative links among process descriptions at various hydrologic scales has been so arduous. New or modified algorithms need to be devised, which will govern observed variabilities and define constitutive relations for each system. In the past, many empirical observations made on simple systems have been often erroneously applied to more complicated systems. Hydrologic systems are quite complex, and equations valid at one scale are not necessarily valid on higher scales, so one may question the applicability of theories that are originally developed for a simple system and vice versa.

Moreover, hydrologic processes are nonlinear and there tend to be strong couplings among them. As a result, the effect of one process is influenced by the occurrence of other processes, and a negligible effect of one process can lead to significant impacts on another. Clearly, there is a need to develop rigorous physical theories to describe the synergistic relationships of coupled nonlinear hydrologic systems on various scales. Intense study on a small scale is not leading to reliable working policy for addressing watershed processes critical to broad scale water management. As such, work needs to be undertaken designed specifically to improve our ability to model and predict water flows and movement. Finally, these models must also be appropriate and applicable to modeling and accurately predicting of flows and outputs in ungauged basins.

Hydroelectric dams, important for generating electricity, have transformed large fluvial systems into chained lakes. These large waterworks have generated a number of substantial environmental disruptions, including the erection of what are often unsurpassable obstacles for fish, explosive growth of floating aquatic plants, and eutrophication of dam reservoirs. The large development projects that connect watersheds and improve continental navigational networks also affect wetlands. Timber harvesting can likewise have negative effects on the production and regulation of water flows because increased soil erosion rates may increase the amount of suspended sediments, which can affect the quality of the water resources and the functioning of dams and reservoirs. On this basis, studies are needed to clarify the most appropriate management practices for harvesting timber.

Water resource management faces many other monumental problems. Academia needs to understand the current crisis completely to become innovative in producing new technologies and in advocating arrangements to satisfy demand. The sciences must be able to integrate and upscale. The field of ecohydrology (ecology and hydrology) is a testimony to the potential success of integrating sciences and should serve as a model for future interdisciplinary cooperatives. However, from a historical context, academia is easily drawn apart to the point that each discipline often fails to identify with statements of findings produced by other disciplines. The public is poorly served by divided or polarized science. Hydrologists are not unique integrators amongst scientists, but interdisciplinary research is required in hydrology to solve the current water-related problems. The goal of strengthening the identity of hydrology as a discipline is commensurate with hydrology's desire to serve the needs of society better. For example, the general public is much more aware of water problems (too much, too little, too dirty) than of the discipline of hydrology and how it relates to water problems. However, people agree that objectives intended to strengthen the hydrologic sciences will result in hydrology that is better equipped to serve the needs of society. These are therefore compatible ideals, well suited to serve both coalitions.

Hydrologic resources management has evolved through incremental refinements to its current point by using information on precipitation, streamflows, water demands, and other processes based to a high degree on a measurement technology that is often archaic in design. Now, scientific advances offer technologies that can be used to add much more information and provide much greater detail. The academic community is often the source of new science and of new technology designed to organize and disseminate information better. Academia is also often the source of courses and literature intended to inform people and make them comfortable with information resources. The world has a fixed quantity of water to use to satisfy a widening diversity and deepening intensity of human and environmental needs. Therefore, academic must make a concerted effort to inform the general public of the limitations of hydrologic resources.

A system for water resources management must also be sustainable to succeed. It is a major challenge to preserve options in a dynamic world where changes are caused by anthropogenic activities, environmental evolution, and geologic processes. These are the challenges that await educators, scientists, and water resource managers in the future.

Ultimately, a management system that addresses periodic severe crises may be less sustainable than that designed to mitigate long-term contingencies. The greatest difficulty lies in protecting future options as social preferences change. Regardless of origin, scientists, educators, politicians, and lawmakers must all cooperate to meet these challenges.

Hydrology and water use are obviously considered from a variety of perspectives, including social, economic, legal, scientific, and environmental, to determine how differing viewpoints affect the quality and quantity of water supplies. Researchers study pollution carried by water, which includes oceans, rivers, streams, rain, snow, and ice, and devise methods to clean and control it. Some study weather related problems such as flood forecasting, drought management, acid rain, and global warming. Others manage water resources so that the goals of all water users are achieved efficiently while protecting the environment. To extenuate and manage water quality issues adequately, they must be dealt with from all perspectives. Some water quality problems that can be common to entire regions include toxic contamination from industry, waste disposal, and eutrophication from human sewage. Bacterial pollution of water supplies is a perpetual dilemma that has adverse effects on human health. The major concerns centered on high bacterial and organic loads range from poor quality drinking water, eutrophication, and disappearance of aquatic life to food contamination and the prevalence of waterborne diseases.

Rivers can be polluted from a number of different sources. The largest contributors include industrial and urban wastewaters from large cities, wastewater from mining industries, and agricultural runoff. As a result, many river water resources today are chemically and biologically contaminated. Large quantities of agricultural contaminants are likewise disposed of in streams flowing to the sea, where there has been found clear evidence of elevated levels of phosphorous, nitrates, potassium, pesticides such as dichlorodiphenyltrichloroethane (DDT), and highly organic effluents. These elevated levels of pollutant concentrations caused by the indiscriminate discharge of highly pollutant-loaded effluents into waterbodies are also seen in deep underground aquifers, which are often contaminated by the same negligence.

Experiences during the past century have contained many useful lessons that are now being used to shape water management practices to address effectively the threats of water shortages and pollution that are growing worldwide. A half-century ago, water resources planning focused on building facilities to capture and deliver water for beneficial uses and to contain floods to reduce damage. Planners were certain that full resource development would foster economic growth and serve broad social needs. Governments built projects that proved financially costly, environmentally destructive, and politically divisive. Environmental harm increased as projects drained rivers and wetlands and converted forests, meadows, and deserts into fields and cities. Political differences deepened as jurisdictions sought to develop their own water resources with little regard for the needs of neighbors. Disputes mounted among towns, states, and nations. As water resources were more fully developed, projects became larger and more costly, and their adverse environmental impacts became more severe. Fiscal and environmental concerns caused the political process to halt construction as society looked for alternative approaches to basic human needs. Now, water withdrawals are approaching the upper limit to what nature can supply, and freshwater resources are being reduced by pollution and threatened by climate change.

Throughout history, humans have been concerned with the proper management of hydrologic resources. This task has been less than facile simply because water availability and population induced threats vary greatly among locations and over time. There is a need to think more creatively to gain deeper understanding of the inherent heterogeneities and diversities within hydrologic systems. Proper management of water resources will need to address the problems and challenges of growing regional and global populations with complex societies but will also need to sustain the needs of the earth's natural ecosystems.

BIBLIOGRAPHY

1. Eagleson, P.E. et al. (1991). *Opportunities in the Hydrologic Sciences*. National Academy Press, Washington, DC, pp. 39–45.
2. Nace, R.L. (1974). *General Evolution of the Concept of the Hydrological Cycle. Three Centuries of Scientific Hydrology*. UNESCO-World Meteorological Organization-International Association of Hydrological Sciences, Paris, pp. 40–48.
3. Dingman, S.L. (1994). *Physical Hydrology*, 1st Edn. Prentice-Hall, Upper Saddle River, NJ, pp. 4–5.
4. Adams, F.D. (1938). *The Birth and Development of the Geological Sciences*. Williams and Wilkins, Baltimore, MD, p. 8.
5. Biswas, A.K. (1970). *History of Hydrology*. North-Holland, Amsterdam, pp. 83–84.
6. Klemes, V. (1986). Dilettantism in hydrology: Transition of destiny. *Water Resour. Res.* **22**: 177–188.

HUMIDITY—ABSOLUTE

Naraine Persaud
Virginia Polytechnic Institute
and State University
Blacksburg, Virginia

HUMIDITY—ABSOLUTE

Absolute humidity is one of many measures to quantify the water vapor content of the atmosphere. It is defined as the mass (m_v) of water vapor per unit volume of moist air. It can be denoted as ρ_{v} with SI units of g (vapor) m^{-3} moist air or in corresponding engineering units as lb ft^{-3} or (more commonly) as grain ft^{-3}. Conversion factors are $1\ \mathrm{lb} = 453.6\ \mathrm{g} = 7000$ grain and $1\ \mathrm{m}^3 = 35.31\ \mathrm{ft}^3$. Other terms in use for this quantity are vapor concentration and density of water vapor in a moist air mixture. Absolute humidity is quite often incorrectly used to mean specific humidity, which is the mass of water vapor per unit mass of moist air, and sometimes to mean the mixing ratio, which is the mass of water vapor per unit mass of dry air. Absolute humidity, specific humidity, and mixing ratio are important in the study of evaporation from water bodies, crop transpiration, soil-water balances, and in ventilation and air conditioning engineering applications. These interrelated measures are rarely measured directly.

Absolute humidity can be estimated indirectly in terms of psychrometric measures of atmospheric water vapor content and known physical parameters. The key assumption in making such estimates is that the mixture of gases in moist atmospheric air behaves ideally. With this assumption, a volume (Vm^3) of moist air at a given pressure (P in Pa; usually at $P = 1.013 \times 10^5 Pa = 1$ atm) and temperature T°K obeys the relationship $\mathrm{PV} = \mathrm{nRT}$, where n is the number of moles of mixture and R is the ideal gas constant $= 8.314\ \mathrm{J\ mol^{-1}\ {}^\circ K^{-1}}$. The number of moles of mixture can be taken as the mass of mixture/mean molecular weight (M_{wm}) of the mixture, which gives $\mathrm{PV} = \mathrm{mRT}/M_{wm} = mR * T$, where $R* = R/M_{wm}$ and is called the specific ideal gas constant. Dividing both sides by V gives $P = mR * T/V = \rho_m\ R * T$, where $\rho_m = m/V$ is the density of the moist air.

The partial pressure P_i of the ith constituent in a volume V of an ideal gas mixture at temperature T can be defined as the pressure it would exert if it alone occupied the volume V, implying that $P_iV = n_iRT$. As for the entire mixture, the total pressure $PV = nRT$, which implies (by summation over i) that $P = \sum P_i$ and (by division) that $P_i/P = n_i/n$. $P = \sum P_i$ is known as Dalton's Law of Partial Pressures. Similarly, the partial volume V_i of the ith constituent T can be defined as the volume it would occupy if exerted and if it alone existed at the total pressure of the mixture. This definition implies $PV_i = n_iRT$, and therefore $V = \sum V_i$ (know as Amagat's Law) and $V_i/V = n_i/n$. Combinations of these results show that $P_i/P = V_i/V = n_i/n$. As the water vapor content by volume in moist air rarely exceeds 0.04, this result implies that its partial pressure rarely exceeds 0.04 atm.

Application of the foregoing ideal relations to estimate absolute humidity from psychrometric meterological observation would be straightforward if the composition of the atmospheric air were constant. However, this is not so, because the water vapor content by volume (i.e., volume water vapor per unit volume of atmospheric air) is the most variable constituent of the atmosphere. It can vary from as little as $2 \times 10^{-6}\%$ (at high altitudes, in extremely cold conditions such as Antartica, and in deserts) to between 4 and 5% (in warm equatorial regions close to large waterbodies). On the other hand, the other constituents of atmospheric air are relatively constant. Therefore, a quantity measured on a volume V of moist air can be regarded as a combination of the quantity measured on the water vapor (V_v) and dry air (V_d) volume components. Thus, the mass m_m of moist air = mass water vapor m_v + mass of dry air m_d. Similarly, the density of moist air $\rho_m = m_m/V = m_v/V + m_d/V = \rho_v + \rho_d$ where ρ_v (the partial density of the vapor component) is by definition the absolute humidity, and ρ_d is the partial density of dry air in the vapor/dry air mixture. The quantity ρ_v/ρ_m represents the mass of water vapor per unit mass of moist air, i.e., the specific humidity. The total moles $n = n_v + n_d$, and therefore, because mass = moles times molecular weight, $\rho_m = [n_vM_{wv}/V] + [n_dM_{wd}/V]$, where M_{wv} and M_{wd} represent molecular weight of water (18.016 g mol^{-1}) and dry air (28.97 g mol^{-1}), respectively. As the molecular weight of water is less than that of dry air, the foregoing indicates that the density of a given volume V of moist air would be less than the density of the same volume of dry air ($<1.29\ \mathrm{kg\ m^{-3}}$). Indeed, if this were not so, instead of our poetry-inspiring cloudy skies, most of the earth would be blanketed by dense surface fogs most of the time.

Pursuing this line of reasoning, the total pressure P of a volume V of moist air at temperature T would be equal to $P_v + P_d$, where P_v and P_d are the partial pressures of the water vapor and dry air. P_v is usually denoted as e, implying $P_d = P - e$. Applying the ideal gas relations above for a volume V of moist air at atmospheric pressure P and temperature T, $e = \rho_vR_vT$, and therefore ρ_v(the absolute humidity) $= e/R_vT = eM_{wv}/RT$, because $\mathrm{R}_v = R/M_{wv}$ is the specific gas constant for water vapor. Dividing numerator and denominator by M_{wd} (molecular weight of dry air) gives:

$$\rho_v = \frac{e\left(\dfrac{M_{wv}}{M_{wd}}\right)}{R_v T} = \frac{0.622e}{R_d T}$$

where R_d is the specific gas constant for dry air (=287.04 J kg^{-1} °K^{-1}).

Thus, at a given temperature T (in °K), the absolute humidity can be obtained if the vapor pressure e were known.

Observations of humidity are widely reported as relative humidity (RH), which is strictly defined as the quotient of the mixing ratio (w) of actual (vapor unsaturated) air at a temperature T and the mixing ratio ($w*$) in vapor-saturated air at the same temperature. The mixing ratio w is defined as the mass of water vapor per unit mass of dry air, and therefore $w = \rho_v/\rho_d$. Therefore, $RH = \mathrm{w}/w*$. As at a given T, $\rho_d = P_d/(R_d T)$ and $P_d = P - e$, therefore from the above equation for ρ_v, $w = 0.622e/(P - e)$. At saturation, the mixing ratio $w* = 0.622e*/(P - e*)$, where $e*$ is the saturation vapor pressure at temperature T. As e and $e*$ are, in general, small compared with P, $w \approx 0.622e/P$, and $w* \approx 0.622e*/P$, implying that $w/w*$ is very nearly equal to $e/e*$, i.e., the ratio of the vapor pressure in actual (vapor unsaturated) air at a temperature T to the vapor pressure in vapor-saturated air at the same temperature. Therefore, e can be obtained from reported RH values if $e*$ were known.

Many equations exist for estimating $e*$ at a given temperature T°K. Based on thermodynamic theory, $e*$ is a function of T and can be obtained by integrating the Clausius–Clapeyron equation $de*/dT = (0.622\,Le*)/(R_d T^2)$, where L is the latent heat of evaporation (J kg^{-1}) at temperature T, R is the specific gas constant for dry air = 287.04 J kg^{-1} °K^{-1}, and T is the absolute temperature (°K). As L is a function of T, the Clausius–Clapeyron equation can be integrated for a given function $L(T)$ to obtain $e*(T)$. One such integral is that of Goff and Gratch (1) and Goff (2) that has become the basis for the published Smithsonian and World Meterorological Organization (WMO) standard tables of $e*(T)$. Gibbins (3) reviewed this and many other equations developed for estimating $e*(T)$. One developed by Richards (4) gives values that differ by no more than 0.01% of the tabulated Goff–Gratch standard. This equation is:

$$e*(T) = ae^{f(T_r)}$$

where $e*$ is in millibar (mb), e is the base of natural logarithms, $T_r = 1 - \dfrac{373.15}{T + 273.15}$, T is the temperature in °C, and $a = 1013.25$. The function $f(T_r) = 13.3185T_r - 1.9760\ (T_r)^2 - 0.6445\ (T_r)^3 - 0.1299\ (T_r)^4$. For example, at 20°C, $T_r = -0.2729$, $f(T_r) = -3.7694$, and $e*(20°\mathrm{C}) = 23.372$ mb (=2337.2 Pa) compared with the tabulated standard value of 23.373 mb.

Another computationally less tedious equation for $e*(T)$ commonly used in ventilation and air conditioning engineering calculations is the Magnus equation (also referred to as the Magnus–Teten formula) valid for air temperatures between 0–50°C:

$$e*(T°\mathrm{C}) = 6.112\ \mathrm{e}^{\frac{17.62T}{243.12+T}}$$

where T is the temperature in °C and $e*$ is in mb. This formula gives $e*(20°\mathrm{C}) = 23.325$ mb.

BIBLIOGRAPHY

1. Goff, R.M. and Gratch, S. (1946). Low pressure properties of water from −160 to 212°F. *Trans. Amer. Soc. Heating Ventilation Engineers* **52**: 95–121.
2. Goff, J.A. (1957). Saturation pressure of water on the new Kelvin temperature scale. *Trans. Amer. Soc. Heating Ventilation Engineers* **63**: 347–354.
3. Gibbins, C.J. (1990). A survey and comparison of relationships for the determination of the saturation vapour pressure over plane surfaces of pure water and of pure ice. *Ann. Geophysica* **8**: 859–886.
4. Richards, J.M. (1971). Simple expression for the saturation vapor pressure of water in the range −50 to 140°C. *Brit. J. Appl. Phys.* **4**: L15–L18.

RELATIVE HUMIDITY

CLIMENT RAMIS
ROMUALDO ROMERO
SERGIO ALONSO
University of the Balearic
Palma de Mallorca, Spain

INTRODUCTION

If samples of air collected at different times of the year and over different points of the earth between the surface and 80 km altitude are analyzed, it is observed that the proportions of gases that form air are, except for a few components, nearly constant in all samples.

That fixed composition is nitrogen (N_2), the most abundant gas with 78.08% in volume concentration, oxygen (O_2) with 20.95%, and argon (Ar) with 0.93%. Other less abundant gases, known as trace gases, are helium (He), neon (Ne), hydrogen (H_2), carbon dioxide (CO_2), ozone (O_3), and water vapor (H_2O), accounting for all together less than 0.04% in volume concentration. The atmospheric layer that has this composition is called the homosphere.

Among the trace gases, ozone shows a particular vertical distribution. The greater concentrations are found at altitudes between 20 and 35 km. This layer is known as the ozonosphere. The concentrations near the earth's surface and over 40 km are very low. On the other hand, the ozone concentration is at a maximum in spring and at a minimum in autumn in the Northern Hemisphere.

Water vapor distribution exhibits great spatial and temporal variability. At any given point, the quantity of water vapor varies continuously, and it is always different from one place to another. An image from the Meteosat satellite in the water vapor (WV) channel permits visualizing the spatial variability (Fig. 1). This spatial variability is particularly high in the vertical, such that about 75% of the water vapor in the atmosphere is in the lower 3000 m, and practically all of this trace gas is trapped in the lowest atmospheric layer or troposphere (10–12 km deep). The high spatial and temporal variability of water vapor requires the definition of variables that, having a simple formulation, permit us to quantify its abundance.

Water itself is a very important component because it can exist in three states (solid, liquid, and gas) within the atmospheric temperature range. Many processes in the atmosphere can cause water to change state. Particularly relevant is when water vapor condenses into liquid water because this process is accompanied by a release of heat that can be absorbed by the air. Similarly, when liquid water freezes or water vapor sublimes, latent heat is released. On the contrary, ice melting, ice sublimation, and liquid evaporation imply heat absorption. The quantification of water vapor content is therefore some measurement of the latent heat available to the air, which, if released, can strongly affect the atmospheric stability to vertical motions and dynamic processes such as the development of pressure systems.

In the remainder of this article, some indexes that are used to quantify water vapor content in the atmosphere are described.

VAPOR PRESSURE AND SATURATION VAPOR PRESSURE

The air is a mixture of gases. The pressure exerted by air is the sum of the partial pressures exerted by each of its components (Dalton's law). The partial pressure exerted by water vapor, denoted as vapor pressure (e), is then a measure of the water vapor content in the atmosphere. A state can be represented by a point in a diagram whose coordinates are (temperature, vapor pressure) = (T, e).

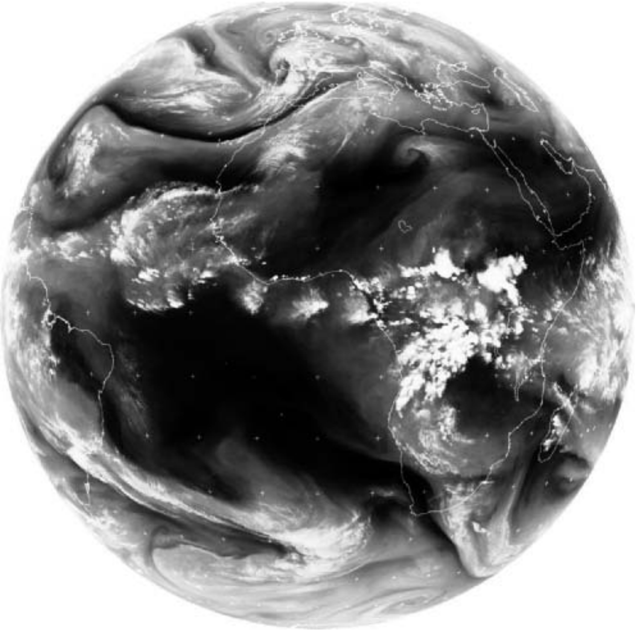

Figure 1. WV image from Meteosat satellite on November 13, 2002 at 1500 UTC (from Dundee University, UK).

For each temperature, a maximum vapor pressure exists that is known as the saturation vapor pressure (e_s). That saturation value cannot be surpassed, and if water vapor in excess occurs, condensation or sublimation results. Figure 2 shows the state curve (T, e_s), with respect to liquid water for temperatures above 0 °C and with respect to supercooled water (e_s) and ice (e_i) below 0 °C. The values of saturation vapor pressure are determined experimentally. A list of values as a function of temperature can be found in Smithsonian Meteorological Table (2). Table 1 includes some e_s values at several temperatures.

Efforts have been carried out to adjust the curve (T, e_s) with respect to liquid water (above and below 0 °C) by an analytic formula. Following is the expression of Tetens (3),

$$e_s(t) = 6.11 * 10\left(\frac{7.5t}{t + 237.3}\right) \tag{1}$$

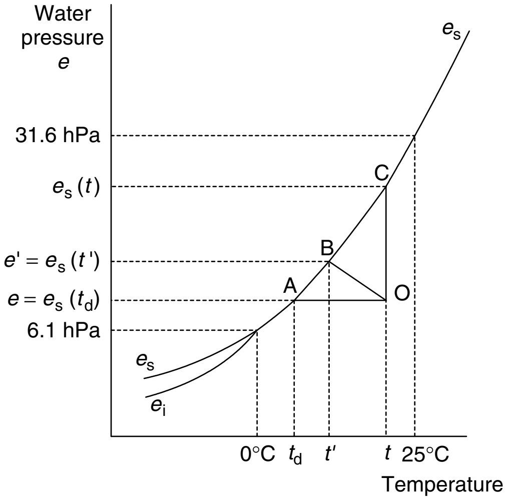

Figure 2. State curve of saturation vapor pressure as a function of temperature. e_s = pressure over a plane surface of liquid water; e_i = pressure over a plane surface of ice. See text for additional information contained in the figure (Diagram is based on Fig. 2 of Reference 1).

Table 1. Values of Saturation Vapor Pressure Over a Plane Surface of Liquid Water at Different Temperatures

t, °C	e_s, hPa
−20	1.2
−10	2.9
0	6.11
10	12.28
20	23.38
30	42.43
40	73.76
50	123.30

Figure 3. Psychrometer. Note the wet thermometer with its bulb covered by wet muslin (photograph courtesy of INM, Spain).

that of Bolton (4),

$$e_s(t) = 6.112 * \exp\left(\frac{17.67t}{t + 243.5}\right) \tag{2}$$

and the modified Clausius–Clapeyron formula,

$$\ln e_s = 53.67957 - \frac{6473.70}{t + 273.16} - 4.8451 \ln(t + 273.16) \tag{3}$$

where t is the temperature in °C and e_s, is the saturation vapor pressure in hPa. The Boltons expression is precise to 0.3% in the range −35 °C to 35 °C and the modified Clausius–Clapeyron method to 0.01% between −40 °C and 40 °C. Both intervals cover a wide spectrum of temperatures in which condensation phenomena are important. Other more complex formulas for saturation vapor pressure on liquid water and on ice can be obtained by polynomial adjustments (5).

Given an air parcel of state (T, e) represented by point O in Fig. 2, two very simple processes permit the air to reach saturation. Suppose that at constant temperature (t), water vapor is added to the parcel so that the vapor pressure is increased. This process eventually leads to saturation, that is, the vapor pressure would reach $e_s(t)$ (point C in the diagram). Analogously, the temperature could be diminished while maintaining the water pressure constant and a lower temperature (t_d) would be reached in which the water pressure would equal the saturation value, that is, $e_s(t_d) = e$ (point A). The temperature at which this happens is known as the “dew point.” If the temperature decreased further, the vapor pressure would equally decrease following curve e_s and water vapor would condense.

Another saturation mechanism would consist of adding liquid water to the air parcel. That liquid water would evaporate, getting the necessary heat from the same air. The temperature would diminish at the same time that the vapor pressure would increase. Thus, with enough liquid water available, saturation would eventually be reached at some temperature t' (point B in Fig. 2). That temperature, known as the “wet bulb temperature,” is the smallest possible temperature an air mass can acquire by evaporation of water in its body.

DETERMINATION OF VAPOR PRESSURE

The vapor pressure at the surface in the real atmosphere is determined indirectly by the psychrometer. That instrument is composed of two identical thermometers placed in a stationary screen box; one has its bulb dampened in muslin (Fig. 3), called the wet thermometer. The temperature difference between the dry and the wet thermometers depends on the water vapor content of the atmosphere; the smaller the vapor content, the greater the temperature difference. Assuming that the air circulating near the wet thermometer effectively reaches the saturation vapor pressure according to the process explained above, its temperature is simply t′ (Fig. 2), and we refer to the dry bulb temperature as t. The following relation is satisfied:

$$e = e_s(t') - \frac{8c_p}{5Lv}p(t - t') \tag{4}$$

where c_p ($c_p = 1005$ Jkg^{-1}K^{-1}) is the specific heat of dry air at constant pressure, L_v ($L_v = 2501 \times 10^6$ J kg^{-1}) the latent heat of condensation, and p is the pressure, which permits determination of the vapor pressure e from the dry and wet bulb thermometer readings. Commonly, tables with double entries of t' and $t - t'$ are constructed which permit determination of e (2).

RELATIVE HUMIDITY

Once a vapor pressure value is determined, the whole picture is not complete unless that value is compared to the saturation vapor pressure at the temperature of the air.

Relative humidity (h) is thus defined as the ratio (expressed in a percentage) of the vapor pressure e and the saturation vapour pressure $e_s(t)$ at the corresponding temperature (1,6):

$$h = 100\frac{e}{e_s(t)} \tag{5}$$

Keeping in mind the dew point definition introduced earlier,

$$h = 100\frac{e_s(t_d)}{e_s(t)} \tag{6}$$

and $e_s(t_d)$ can never be greater than $e_s(t)$. For an interpretation, see Fig. 2. When $h = 100\%$, saturation has been reached and condensation is produced.

Some surfaces exposed to the environment often acquire a temperature that is different from that of the air. For those surfaces, an equivalent relative humidity can be defined as

$$h = 100\frac{e}{e_s(t_s)} = 100\frac{e_s(t_d)}{e_s(t_s)} \tag{7}$$

where t_s is now the temperature of the surface. A surface cooler than the air ($t_s < t$), for example, would then promote saturation more easily than the surrounding air. This explains why dew sometimes forms on plant leaves although 100% relative humidity has not been reached in the atmosphere. Analogous expressions could be defined considering the saturation pressure with respect to ice.

Hygrometers are instruments that measure relative humidity directly. Digital hygrometers are generally based on some change in the electrical properties of some material as a function of the relative humidity of the air, whereas analogic hygrometers use the variation of length that the human hair (especially that of a blond woman!) experiences with a change in relative humidity.

If the relative humidity is known, the dew point can be determined by inverting some of the formulas for calculating the saturation vapor pressure (Eqs. 1–3). Conversion tables are also available for that purpose (2).

OTHER HUMIDITY INDEXES

There are several other humidity indexes apart from relative humidity. They are often more useful than the latter for adequate formulation of the moist thermodynamic processes that occur in the atmosphere. The following are emphasized:

Mixing Ratio (*r*)

It is defined as the quantity of water vapor per unit mass of dry air. Its units are therefore g g^{-1} but generally it is expressed in g kg^{-1}:

$$r = \frac{m_v}{m_d} \tag{8}$$

Both the water vapor and dry air portions of the sample, which have partial pressures e and $p - e$, respectively, comply with the equation for the ideal gases; then,

$$e = \frac{m_v}{V}R_vT \tag{9}$$

$$p - e = \frac{m_d}{V}R_dT \tag{10}$$

where R_v and R_d are the specific gas constants of water vapor and dry air, respectively, V is the volume, and T the temperature. Combining Eqs. 9 and 10,

$$r = \frac{m_v}{m_d} = \frac{R_d}{R_v}\frac{e}{p - e} = \varepsilon\frac{e}{p - e} \tag{11}$$

and using the constants $R_d = 287.04$ J kg^{-1} K^{-1} and $R_v = 461.50$ J kg^{-1} K^{-1}, Eq. 11 is

$$r = 0.622\frac{e}{p - e} \tag{12}$$

Within the interval of temperatures and atmospheric pressures, $e << p$, and therefore

$$r \cong 0.622\frac{e}{p} \tag{13}$$

If the vapor pressure corresponds to its saturation value, then a saturation mixing ratio is defined as

$$r_s \cong 0.622\frac{e_s}{p} \tag{14}$$

that depends only on pressure and temperature.

Using Eqs. 13 and 14, the relative humidity is

$$h = 100\frac{e}{e_s(t)} \cong 100\frac{r}{r_s(t)} \tag{15}$$

Therefore, the numerical values of h deduced from vapor pressure or from the mixing ratio are practically the same.

The exact relation between the relative humidity and the mixing ratio is

$$h = 100\frac{r}{r_s(t)}\frac{1+\dfrac{r_s(t)}{\varepsilon}}{1+\dfrac{r}{\varepsilon}} \tag{16}$$

Specific Humidity (*q*)

Specific humidity is defined as the relation between the vapor mass and the mass of moist air that contains this vapor, that is, the quantity of water vapor per unit mass of moist air:

$$q = \frac{m_v}{m_r + m_d} \tag{17}$$

The units are g g^{-1} although it is generally expressed in g kg^{-1}.

Using the same formulation introduced previously for the mixing ratio,

$$q = \varepsilon\frac{e}{p - e(1-\varepsilon)} \tag{18}$$

And according to the same previous arguments,

$$q \cong \varepsilon\frac{e}{p} \tag{19}$$

and therefore

$$h \cong 100\frac{q}{q_s(t)} \tag{20}$$

The same comments noted earlier for the mixing ratio (r) are valid for the specific humidity in relation to relative humidity.

Knowledge of the specific humidity permits us to calculate the specific parameters for humid air, for example, the specific heat

$$c_p' = (1-q)c_p + qc_v \tag{22}$$

where c_p and c_v are the specific heats of dry air and water vapor at constant pressure, respectively ($c_p = 1005$ J $kg^{-1}K^{-1}$, $c_v = 1.85\ 10^3$ J $kg^{-1}K^{-1}$); and the specific gas constant

$$R' = (1-q)R_d + qR_v \tag{23}$$

BIBLIOGRAPHY

1. McIntosh, D.H. and Thom, A.S. (1978). *Essentials of Meteorology*. Wykeham.
2. (1980). *Smithsonian Meteorological Tables*, 5th Edn.
3. Murray, F.W. (1967). On the computation of saturation vapour pressure. *J. Appl. Meteorol.* **6**: 203–204.
4. Bolton, D. (1980). The computation of equivalent potential temperature. *Mon. Weather Rev.* **108**: 1046–1053.
5. Flatan, P.J., Walks, R.L., and Cotton, W.R. (1992). Polynomial fits to saturation vapour pressure. *J. Appl. Meteorol.* **31**: 1507–1513.
6. Wallace, J.M. and Hobbs, P.V. (1977). *Atmospheric Science. An Introductory Survey*. Academic Press, San Diego, CA.

HURRICANES: MODELING NATURE'S FURY

Geophysical Fluid Dynamics
Laboratory—NOAA

Few natural phenomena evoke a greater sense of awe and mystery than hurricanes, the powerful and destructive tropical storms that often menace coastal regions. Scientists at GFDL, led by Yoshio Kurihara, have worked steadily for two decades to develop physically based models that can be used to understand the basic mechanisms producing hurricanes and to predict the movement and intensity of individual storms.

A PREDICTION BREAKTHROUGH

Improvements in hurricane prediction have been made possible by improved observations, models, and methods for incorporating observations into the forecast models. Kurihara and collaborators Bob Tuleya and Morris Bender realized that current weather analyses were inherently inadequate as "initial conditions" for such models. They developed a method to circumvent much

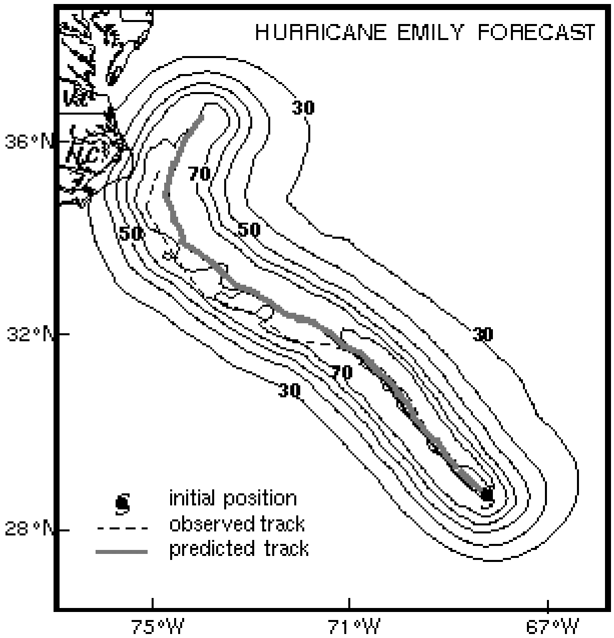

Three-day forecast of Hurricane Emily's track (red line) and intensity produced by the GFDL hurricane model. Thin contours show the maximum predicted wind speed (in knots) during the forecast period. The forecast was available at 8 a.m., Sunday, Aug. 29, 1993. The hurricane caused extensive damage along North Carolina's Outer Banks beginning late Tuesday afternoon (Aug. 31) before recurving to the northeast and out to sea, as predicted by the GFDL model

---

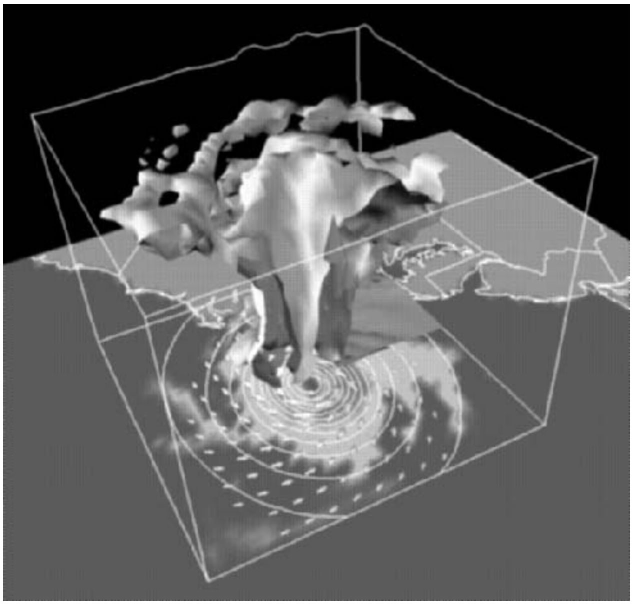

Three-dimensional view of Hurricane Emily near the North Carolina coast as modeled by the dynamical hurricane prediction system developed at GFDL. The view is a 48-hour forecast (valid 8 p.m., August 31, 1993) obtained from the GFDL model. This forecast system provides improved hurricane track predictions roughly two days in advance of hurricane conditions and has the potential to provide useful forecasts of storm intensity and precipitation as well. In the case of Hurricane Emily, the model successfully predicted the occurrence of hurricane-force winds in the Cape Hatteras, N.C. vicinity. In the figure, the three-dimensional "cloud-like" volume (which is sliced vertically through the storm center) depicts the 91% relative humidity surface; the colored shading at the earth's surface shows precipitation (higher values are red); red and white arrows show vertical motions and surface winds, respectively; white contours depict sea-level pressure; the dark blue surface and solid green background represent the Atlantic Ocean and the U.S. East Coast, respectively. Note the "eye-wall" in the relative humidity surface at the storm center and the heavy precipitation north of the storm

of the initial data problem and achieved a breakthrough in hurricane prediction.

The new forecast system achieved a series of remarkably successful forecasts, compared to other available methods for Hurricanes Andrew, Iniki, and Emily. For example, while conventional models predicted that Hurricane Emily (August 1993) would make landfall in the vicinity of Georgia, the GFDL forecast system correctly predicted that the hurricane would deliver a glancing blow to North Carolina's Outer Banks before veering back out to sea. Based on its impressive performance during the 1993 and 1994 hurricane seasons, the model has been adopted as the operational hurricane prediction model at the National Meteorological Center.

HURRICANE EMILY FORECAST

Given the expense and inconvenience of hurricane evacuations and emergency preparations, the public benefits greatly if unnecessary hurricane warnings can be avoided. Furthermore, residents are likely to take hurricane warnings more seriously if the forecasts are shown to be more reliable. With potential damage from a single storm exceeding $25 billion (e.g., Hurricane Andrew in 1992), reliable short-term preparations, while expensive to undertake, can result in large net savings by the public.

A NEW COASTAL FORECAST SYSTEM

In a complementary effort, GFDL is cooperating with scientists at Princeton University and NOAA's National Meteorological Center (NMC) and National Ocean Service (NOS) to develop a coastal forecast system for the East Coast. The system is designed to provide nowcasts and forecasts of currents, storm surges, and other information to coastal residents, the fishing and marine transport industries, and government users. The forecast system uses a high-resolution ocean model–developed by Princeton's George Mellor in collaboration with GFDL–and an atmospheric model from NMC.

HYDROLOGIC CYCLE

Sharad K. Jain
National Institute of Hydrology
Roorkee, India

Vijay P. Singh
Louisiana State University
Baton Rouge, , Louisiana

The hydrologic cycle, the endless circulation of water between the earth, the oceans, and the atmosphere, is fundamental to the study of the waters of the earth. The cycle entails a huge turnover of water driven by the energy of the Sun and gravity. It has a profound influence on the landscape and the earth's climate. The major components of the hydrologic cycle are precipitation, evapotranspiration, interception, infiltration, overland and channel flow, and groundwater flow. Quantitatively, the hydrologic cycle is represented by a mass balance or continuity equation. The relative significance of the terms of this equation depends on the space and timescale. Human activities and the hydrologic cycle are interactive and influence the earth's climate.

INTRODUCTION

The hydrologic cycle is a fundamental concept in hydrology and is amongst a number of cycles known to be operating in nature, such as the carbon cycle, the nitrogen cycle, and other biogeochemical cycles. The National Research Council (1) defines the hydrologic cycle as "the pathway of water as it moves in its various phases to the atmosphere, to the earth, over and through the land, to the ocean and back to the atmosphere." This cycle has no beginning or end, and water is present in all three states, solid, liquid,

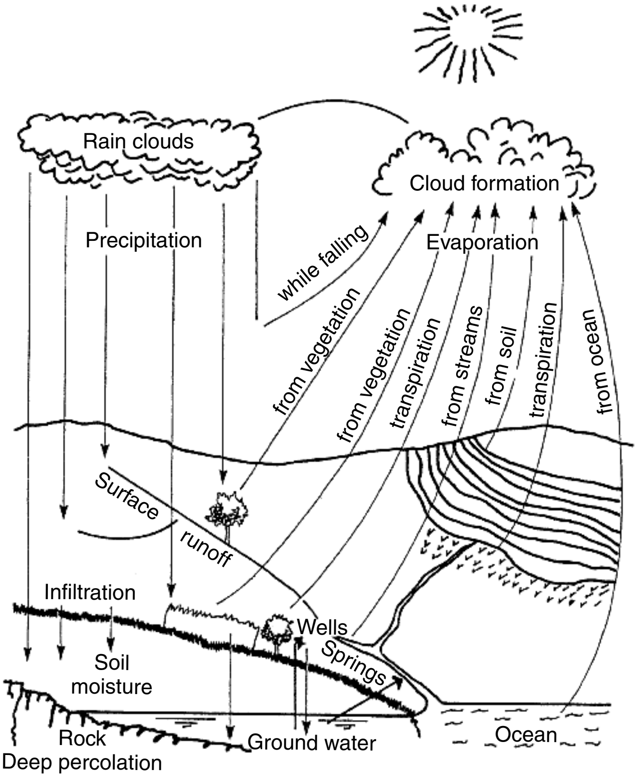

Figure 1. Pictorial representation of the hydrologic cycle (2).

and gas, in the cycle. A pictorial representation of the hydrologic cycle is given in Fig. 1.

The hydrologic cycle, also known as the water cycle, is a concept that considers the processes of motion, loss, and recharge of the earth's waters. It connects the atmosphere and two storages of the earth system: the oceans and the landsphere (lithosphere and pedosphere). The water that is evaporated from the earth and the oceans enters the atmosphere. Water leaves the atmosphere through precipitation. The oceans receive water from the atmosphere by precipitation and from the landsphere through stream flow and groundwater flow. The only way that water leaves oceans is by evaporation. The landsphere receives water by precipitation. The water leaves this sphere by evaporation, transpiration, evapotranspiration, stream flow, and groundwater flow. Evaporation and precipitation are the processes that take place in the vertical plane, whereas stream flow and groundwater flow occur mostly in the horizontal plane.

The exchange of water among the oceans, land, and the atmosphere has been termed 'the turnover' by Shiklomanov (3). This turnover affects the global patterns of the movement of ocean waters and gases in the atmosphere, thereby greatly influencing climate. Besides, water is a very good solvent, and hence, geochemistry is an integral part of the hydrologic cycle. The hydrologic cycle is, thus, the integrating process for the fluxes of water, energy, and the chemical elements (4). Usually, rain and snow are considered the purest form of water, although they may also be mixed with pollutants in the atmosphere. During the journey on the earth, many chemical compounds are mixed with water, and consequently water quality undergoes a change.

The hydrologic cycle can also be visualized as a perpetual distillation and pumping system. In this endless circulation of water, the glaciers and snow packs are replenished, the quantity of river water is replenished, and its quality is restored. From the view point of water use, the land phase of the hydrologic cycle is the most important. The necessity and utility of studying the hydrologic cycle arises because water is essential for the survival of life and is an important input in many economic activities. But water of the desired quality may not be available when needed.

In view of the complexities and extensive coverage, the study of the complete hydrologic cycle is truly interdisciplinary. For instance, the atmospheric part is studied by meteorologists, the pedospheric part by soil scientists, the lithosphere by geologists, and the part pertaining to oceans falls in the domain of oceanographers. The domain of a hydrologist is confined mainly to the land phase, and this in itself is a very large domain. A host of other professionals may be interested in studying the hydrologic cycle, including energy utility managers, chemists, agricultural engineers, public health officers, and inland navigation managers. It is hard to draw lines demarcating various domains in such an interdisciplinary subject, and the boundaries are often blurred. Teams of people from various specialties often work together on a problem dealing with the components of the hydrologic cycle.

COMPONENTS OF THE HYDROLOGIC CYCLE

The hydrologic cycle can be subdivided into three major systems: The oceans are the major reservoir and source of water, the atmosphere functions as the carrier and deliverer of water, and the land is the user of water. The amount of water available at a particular place changes with time because of changes in the supply and delivery. Globally, water movement is a closed system, but locally, it is an open system. A systems representation of the hydrologic cycle is shown in Fig. 2.

The major components of the hydrologic cycle are precipitation (rainfall, snowfall, hale, sleet, fog, dew, and drizzle), interception, depression storage, evaporation, transpiration, infiltration, percolation, moisture storage in the unsaturated zone, and runoff (surface runoff, interflow, and base flow).

Evaporation of water takes place from the oceans and the land surface mainly due to solar energy. The moisture moves in the atmosphere as water vapor which precipitates on land surface or oceans as rain, snow, hail, or sleet. Part of this precipitation is intercepted by vegetation or buildings. Of the amount reaching the land surface, part infiltrates into the soil, and the remaining water runs off the land surface to join streams. These streams finally discharge into the ocean. Some of the infiltrated water percolates deep to join groundwater, and some comes back to the streams or appears on the surface as springs.

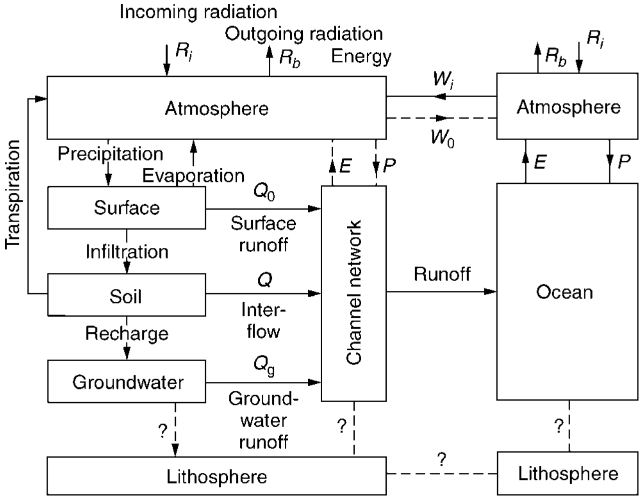

Figure 2. Systems representation of the hydrologic cycle (5).

This immense movement of water is driven mainly by solar energy: the excess of incoming radiation over outgoing radiation. Therefore, the Sun is the prime mover of the hydrologic cycle. The energy for evaporating water from streams, lakes, ponds, oceans, and other open water bodies comes from the Sun. A substantial quantity of moisture is added to the atmosphere by transpiration of water from vegetation. Living beings also supply water vapor to the atmosphere through perspiration. Gravity has an important role in the movement of water on the earth's surface, and anthropogenic activities also have an increasingly important influence on water movement.

An interesting feature of the hydrologic cycle is that at some point in each phase, there usually occur (1) transportation of water, (2) temporary storage, and (3) a change of state. For example, in the atmospheric phase, there is vapor flow, vapor storage in the atmosphere, and condensation or formation of precipitation created by a change from vapor to either the liquid or solid state. Moreover, in the atmosphere, water is in the vapor form, whereas it is mostly (saline) liquid in the oceans.

Scales for Studying the Hydrologic Cycle

Depending on the purpose of study, the hydrologic cycle is studied on many different spatial scales. Horton (6) clearly recognized the diversity of scales when he stated *Any natural, exposed surface may be considered as a unit area on which the hydrologic cycle operates. This includes, for example, an isolated tree, even a single leaf or twig of a growing plant, the roof of a building, the drainage basin of a river system or any of its tributaries, an undrained glacial depression, a swamp, a glacier, a polar ice-cap, a group of sand dunes, a desert playa, a lake, an ocean, or the Earth as a whole* (Quoted in Ref. 4). A qualitative representation of the hydrologic cycle is given in Fig. 3.

From the view point of hydrologic studies, two scales are readily distinct, the global scale and the catchment scale.

Global Scale. From a global perspective, the hydrologic cycle is comprised of three major systems: the oceans, the atmosphere, and the landsphere. Precipitation, runoff, and evaporation are the principal processes that transmit water from one system to the other, as illustrated in Fig. 4. This illustration depicts a global geophysical view of the hydrologic cycle and shows the interactions among the earth (lithosphere), the oceans (hydrosphere), and the atmosphere. Study on the global scale is necessary to understand the global fluxes and global circulation patterns. The results of these studies form important inputs to water resource planning for a national or regional water resource assessment, weather forecasting, and study of climate changes. These results may also form the boundary conditions of small-scale models/applications.

Catchment Scale. While studying the hydrologic cycle on a catchment scale, the spatial coverage can range from a few square km to thousands of square km. The timescale could be a storm lasting for a few hours to a study spanning many years. When the water movement of the earth system is considered, three systems can be recognized: the land (surface) system, the subsurface system, and the aquifer (or geologic) system. The dominant processes of the hydrologic cycle of the land system, are precipitation, evapotranspiration, infiltration, and surface runoff. The land system itself comprises three subsystems: the vegetation subsystem, the structural subsystem, and the soil subsystem. These subsystems subtract water from precipitation through interception, depression, and detention storage. This water is either lost to the atmospheric system or enters the subsurface system. The exchange of water among these subsystems takes place through infiltration, exfiltration, percolation, and capillary rise. Figure 5 shows the schematic of the hydrologic cycle on a global scale in the earth system and a microscale view of the cycle in the land system.

Timescales in the Hydrologic Cycle. The time required for the movement of water through various components of the hydrologic cycle varies considerably. The velocity of stream flow is much higher than the velocity of groundwater. The time-step size for an analysis depends on the purpose of study, the availability of data, and how detailed the study is. The estimated periods of renewal of water resources in water bodies on the earth are given in Table 1. The time step should be sufficiently small so that the variations in the processes can be captured in sufficient detail, but at the same time, it should not put undue burden on data collection and computational efforts.

The range of the spatial and temporal dimensions of many processes related to the hydrologic cycle is shown in Fig. 6.

MATHEMATICAL REPRESENTATION OF THE HYDROLOGIC CYCLE

The quantities of water going through the various components of the hydrologic cycle can be evaluated by the so-called hydrologic equation, which is a simple spatially lumped continuity or water budget equation:

$$I - Q = \Delta S \tag{1}$$

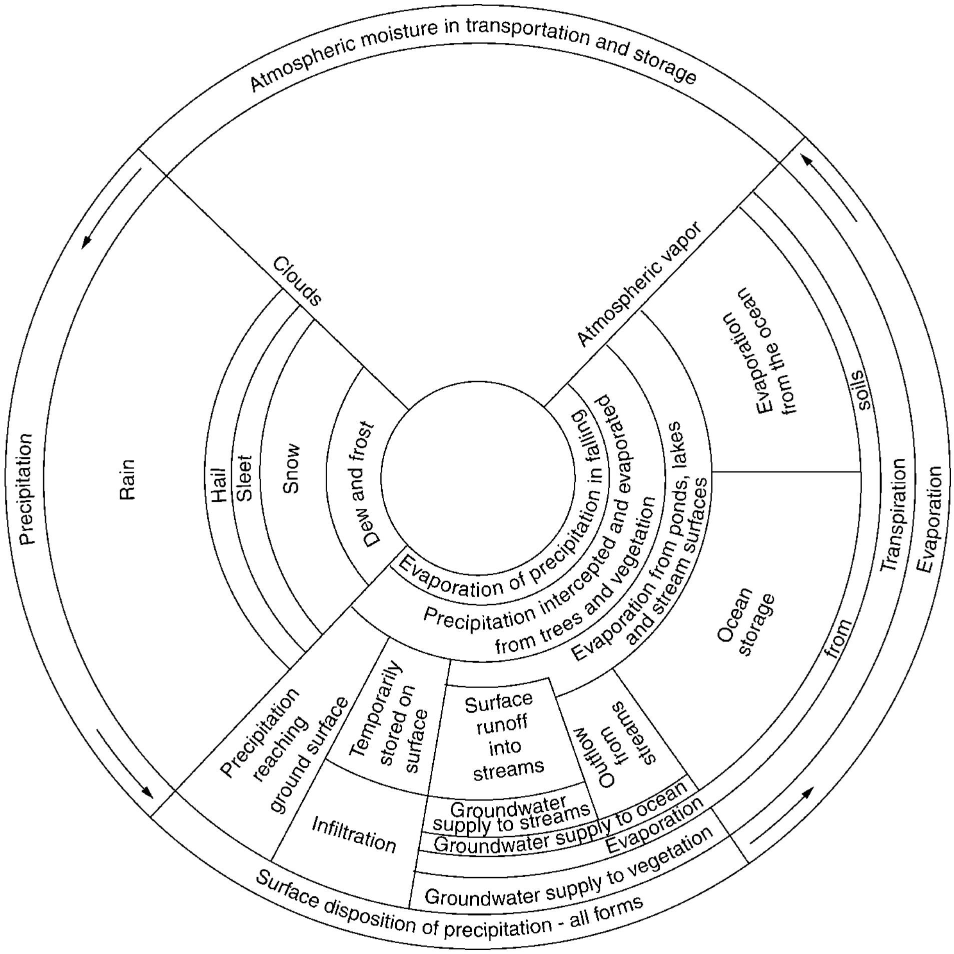

Figure 3. Qualitative representation of the hydrologic cycle (6).

Table 1. Periods of Water Resource Renewal on the Earth[a]

Water of Hydrosphere	Period of Renewal
World ocean	2500 years
Ground water	1400 years
Polar ice	9700 years
Mountain glaciers	1600 years
Ground ice of the permafrost zone	10000 years
Lakes	17 years
Bogs	5 years
Soil moisture	1 year
Channel network	16 days
Atmospheric moisture	8 days
Biological water	Several hours

[a] Reference 3.

where I = inflow of water to a given area during any given time period, Q = outflow of water from the area during the selected time period, and ΔS = change in storage of water in the given area during the time period. If I and Q vary continuously with time, then Eq. 1 can be written as

$$dS(t)/dt = I(t) - Q(t) \tag{2}$$

By integrating, this equation can also be written as

$$\int dS(t) = \int [I(t) - Q(t)]\,dt$$

or

$$S(t) - S(0) = \int_0^t I(t)dt - \int_0^t Q(t)dt$$
$$= V_I(t) - V_0(t) \tag{3}$$

where $S(0)$ is the initial storage at time $t = 0$, $S(t)$ is the storage at time t, and $V_0(t)$ and $V_I(t)$ are volumes of outflow and inflow at time t. Each of the terms of this lumped equation is the result of a number of other terms. They can be subdivided and even eliminated from the equation, depending on the temporal and spatial scale of the study. The continuity equation is one of the governing equations of almost all hydrologic problems. For a watershed, Eq. 1 may be written as

$$P + Q_{\mathrm{SI}} + Q_{\mathrm{GI}} - E - Q_{\mathrm{SO}} - Q_{\mathrm{GO}} - \Delta S - \varepsilon = 0 \tag{4}$$

where P is precipitation, Q_{SI} is surface inflow, Q_{GI} is groundwater inflow, E is evaporation from the watershed, Q_{SO} is surface water outflow, Q_{GO} is groundwater outflow, ΔS is the change in the storage of water in the watershed, and ε is a discrepancy term. For large watersheds, Q_{GI}

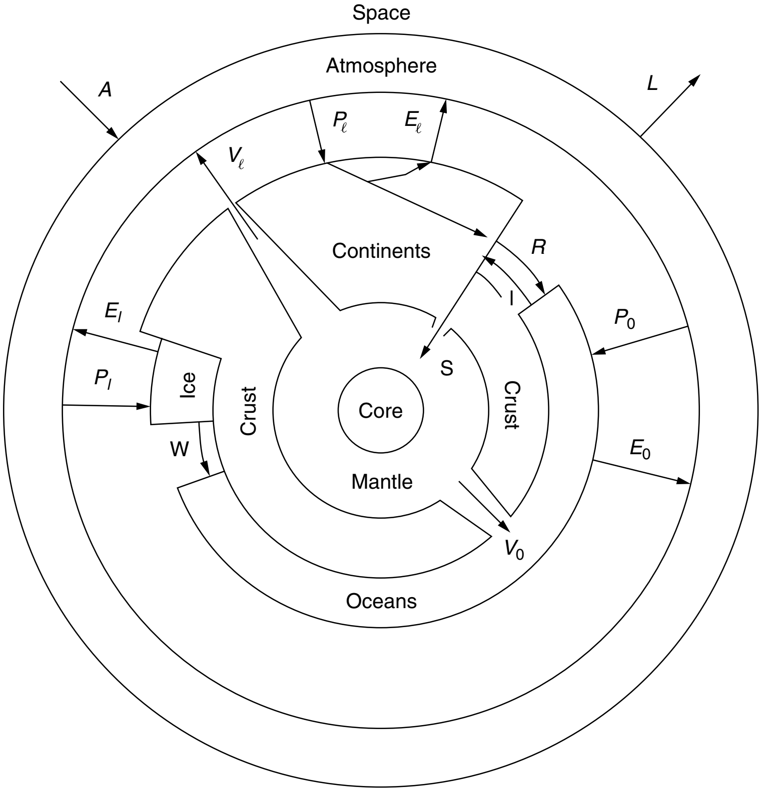

A = Additions of water from space
E_0 = Evaporation from oceans
E_I = Evaporation (i.e., sublimation) from ice
E_ℓ = Evapotranspiration from land
I = Intrusion of seawater into continental aquifers
L = Loss of water to space
P_0 = Precipitation on oceans
P_I = Precipitation on ice
P_ℓ = Precipitation on land
R = Runoff from continents
S = Subduction of water - containing crust
V_0 = Volcanic venting to oceans
V_ℓ = Volcanic venting to atmosphere
W = Wastage of ice sheets to ocean

Figure 4. The hydrologic cycle as a global geophysical process (4).

and Q_{GO} are usually negligible. The discrepancy term is included in Eq. 4 because the sum of all other terms may not be zero due to measurement errors and/or simplifying assumptions. However, a small value of the discrepancy term does not necessarily mean that all other terms have been correctly measured/estimated.

Depending on the specific problem, the terms of Eq. 1 may be further subdivided. For example, when applying the hydrologic equation for short time intervals, the change in the total water storage (ΔS) may be subdivided into several parts: changes of moisture storage in the soil (ΔM), in aquifers (ΔG), in lakes and reservoirs (ΔL), in river channels (ΔS_C), in glaciers (ΔS_G), and in snow cover (ΔS_S). Thus, ΔS can be expressed as

$$\Delta S = \Delta M + \Delta G + \Delta L + \Delta S_C + \Delta S_G + \Delta S_S \qquad (5)$$

The hydrologic equation may be applied to any time interval; computation of the mean annual water balance for a basin is the simplest because it is possible to disregard changes in water storage in the basin (ΔS), which are difficult to measure and compute. In general, the shorter the time interval, the more stringent are the requirements for measuring or computing the components, and the more subdivided are the terms of Eq. 1. This results in a complex equation which is difficult to close with acceptable errors.

The hydrologic equation may be applied to areas of any size, but the complexity of computation depends greatly on the extent of the area under study. The smaller the area, the more complicated is its water balance because it is difficult to estimate components of the equation. Finally, the components of the hydrologic equation may be expressed in terms of the mean depth of water (mm), as a volume of water (m^3), or as flow rates (m^3/s or mm/s).

Global Water Balance

According to estimates (7), the annual average depth of precipitation on the land surface is about 108×10^3 km^3. Out of this, about 61×10^3 km^3 is returned to the atmosphere as evapotranspiration, and the runoff from land to oceans is 47×10^3 km^3. As far as the water balance of oceans is concerned, the depth of precipitation over them is about 410×10^3 km^3, 47×10^3 km^3 of water is received as runoff from the land, and 457×10^3 km^3 is lost as evaporation. If we consider the water balance of the atmosphere, 457×10^3 km^3 water is received as evaporation from oceans and 61×10^3 km^3 from land. The precipitation over oceans is

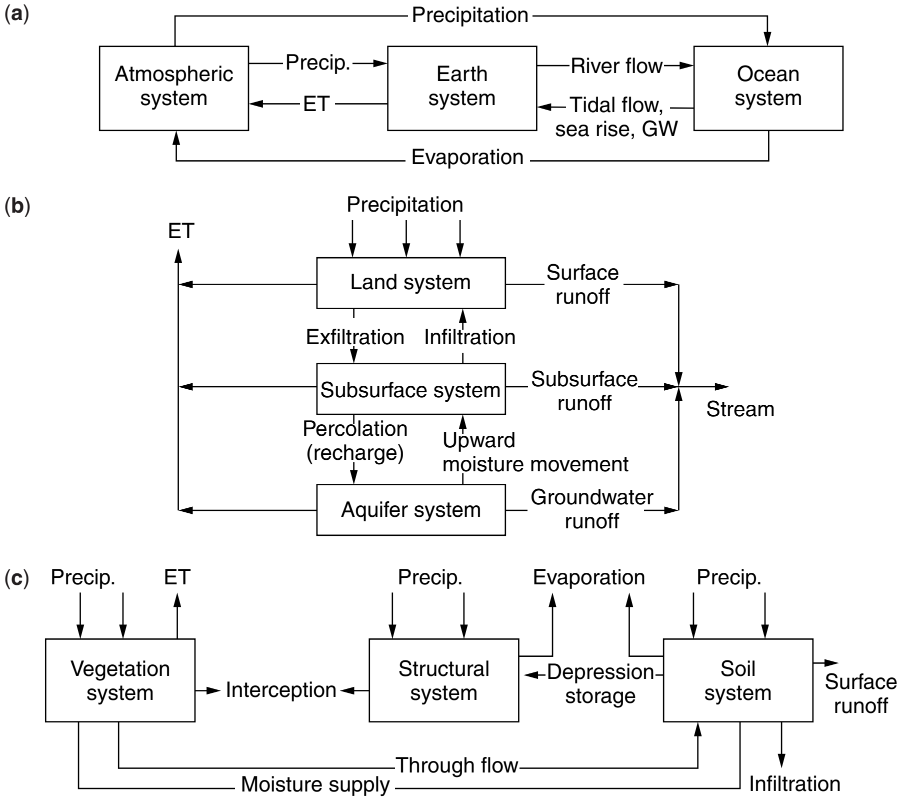

Figure 5. (**a**) A global schematic of the hydrologic cycle (2). (**b**) A schematic of the hydrologic cycle of the earth system (2). (**c**) A detailed schematic of the hydrologic cycle in the land system (2).

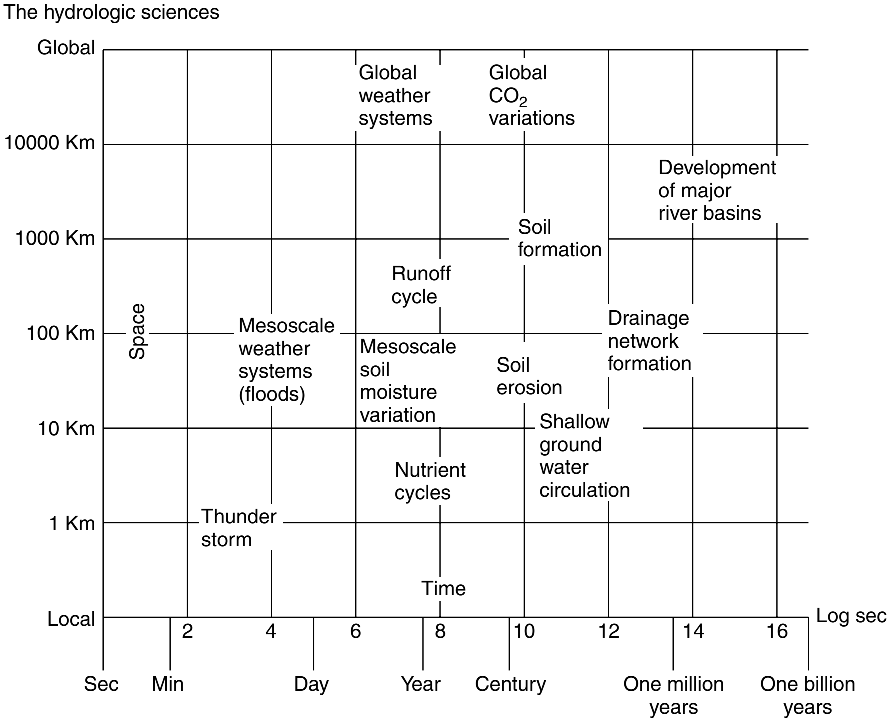

Figure 6. Illustrative range of process scales (4).

410×10^3 km^3, and it is 108×10^3 km^3 over land. Table 2 gives quantitative values of water in various forms on the earth.

The storages and fluxes of the hydrologic cycle are shown in Fig. 7.

INFLUENCE OF HUMAN ACTIVITIES AND LAND-USE CHANGES ON THE HYDROLOGIC CYCLE

Watersheds are subjected to many types of changes, major or minor, for various reasons. Some of them are natural

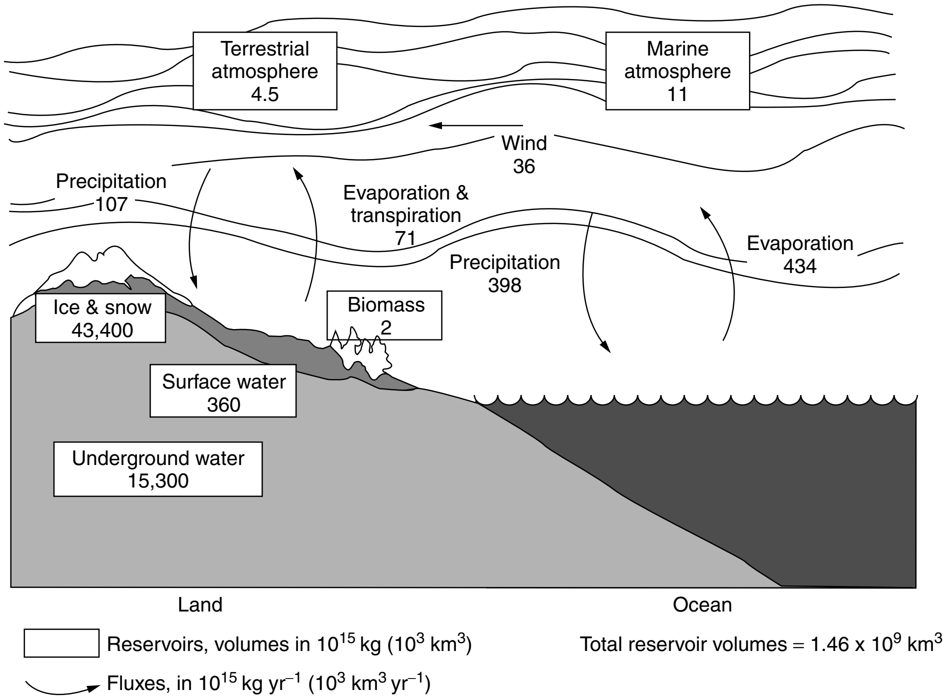

Figure 7. Volumes and fluxes of the global hydrologic cycle (9).

changes, and some are due to human activities. Watershed changes affect virtually all elements of the hydrologic cycle. The quality of water is significantly deteriorating at many places due to industrial and agricultural activities. There has been a growing need to quantify the impact of major human-induced changes on the hydrologic cycle to anticipate and minimize the potential environmental detriment and to satisfy the water resource requirements of society. Even if the water of adequate quantity were present at a place, its use may be limited because of poor quality. The classical and the modern viewpoints of the human role in the hydrologic cycle are shown in Fig. 8.

Most watershed changes can be distinguished as point changes and nonpoint changes. Structural changes, such as dam construction, channel improvement, and detention storage are examples of point changes and affect watershed response in terms of evaporation, seepage, and residence time. Forestry, agriculture, mining, and urbanization are nonpoint land-use changes that affect catchment response. A qualitative discussion of the hydrologic consequences of watershed changes is in order.

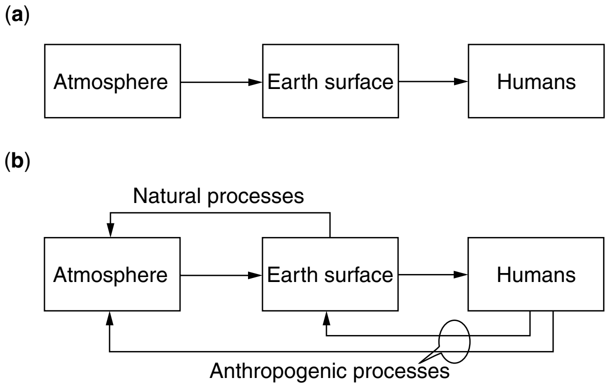

Figure 8. (**a**) Classical viewpoint of the human role in the hydrologic cycle. (**b**) Modern viewpoint of the human role in the hydrologic cycle (1).

Effects of Agricultural Changes

These changes imply that a land area that was forested or barren earlier, is now being cultivated. As a result, the vegetal cover changes; the land slope may be altered a little bit, soil crusting and infiltration characteristics change, and artificial bounds may be placed. The effect of these changes on the hydrologic regime is pronounced and may be multiplicative. Water may be withdrawn from the groundwater zone, or canal irrigation may be introduced leading to noticeable changes in water table behavior. The impacts are also noticed in evapotranspiration, overland flow, channel flow, and infiltration. Fertilizers, pesticides, and insecticides that are applied to crops affect the quality of runoff from agricultural areas.

Effects of Urbanization

A land area that was being used for forestry or agriculture might be transformed into an urban area where houses, roads, parks, parking lots, and sewers are constructed. A large increase in the paved (impervious) surface takes place which considerably reduces infiltration, and the removal of storm water is accelerated. Urban development usually increases the volume and peak of direct runoff for a given rainfall. The time of travel of water is reduced, resulting in lower lag time and a lower time of concentration. However, it is possible to control these effects by providing detention storage or changing the landscape and sizing the storm drains. In brief, the hydrologic effects of urbanization are

Table 2. Water Reserves in Various Phases of the Hydrologic Cycle[a]

	Distribution Area, 10^3 km^2	Volume, 10^3 km^3	Layer, m	Percentage of Global Reserves	
				Of Total Water	Of Freshwater
World ocean	361,300	1,338,000	3,700	96.5	–
Groundwater	134,800	23,400	174	1.7	–
Freshwater	–	10,530	78	0.76	30.1
Soil moisture	–	16.5	0.2	0.001	0.05
Glaciers and permanent snow cover	16,227	24,064	1,463	1.74	68.7
Antarctic	13,980	21,600	1,546	1.56	61.7
Greenland	1,802	2,340	1,298	0.17	6.68
Arctic islands	226	83.5	369	0.006	0.24
Mountainous regions	224	40.6	181	0.003	0.12
Ground ice/permafrost	21,000	300	14	0.022	0.86
Water reserves in lakes	2,058.7	176.4	85.7	0.013	–
Fresh	1,236.4	91	73.6	0.007	0.26
Saline	822.3	85.4	103.8	0.006	–
Swamp water	2,682.6	1.47	4.28	0.0008	0.03
River flows	148,800	2.12	0.014	0.0002	0.006
Biological water	510,000	1.12	0.002	0.0001	0.003
Atmospheric water	510,000	12.9	0.025	0.001	0.04
Total water reserves	510,000	1,385,984	2.718	100	–
Total freshwater reserves	148,800	35,029	235	2.53	100

[a] Reference 8.

1. increased water withdrawals from surface and subsurface water bodies; sometimes demands exceed the available natural resources;
2. increased peak flow and diminishing baseflow of streams;
3. reduced infiltration;
4. increased pollution of rivers and aquifers that endanger the ecology;
5. increased withdrawals from ground water; and
6. changes in local microclimate.

Effects of Forest Activities

These activities may be directed toward planting trees as well as cutting them. The immediate effect of forest activities is change in vegetal cover. When a forested area is deforested and forest litter removed, the interception of precipitation is virtually eliminated. Litter removal changes the infiltration capacity of soil and has a pronounced effect on raindrop impact and resulting soil erosion. When vegetation is lost, evapotranspiration is generally decreased. These changes amount to increased production of direct runoff, reduced surface roughness, and decreased recharge to groundwater. The hydrograph of direct runoff rises more quickly because of the reduced time to peak. However, when additional trees are planted in an area, the effect is reversed, though the impact takes place gradually as the trees grow.

Effects of Structural Changes

Typical structural changes include a dam, a weir, and channel improvement. A dam–reservoir is constructed for many purposes. Regardless of its intended function, it does affect the hydrology of the stream on which it is built. In general, the peak of the outflow from a reservoir is less and the flow may be more even than in the preproject condition. The volume of flow downstream may be considerably less in the afterproject scenario if the reservoir water is diverted elsewhere.

In contrast with reservoirs, channel improvements have an opposite effect. For example, decreasing channel roughness increases flow velocity and hence peak discharge for the same channel size. The removal of vegetation, lining of the channel, and proper maintenance can greatly reduce roughness. Other alterations, such as straightening the channel, maintenance of bands, or increasing slope, significantly affect travel time and flow velocity. Depending on the bed material, infiltration through the bed and banks also modifies flow characteristics.

Impact of Climate Changes on the Hydrologic Cycle

An analysis of measured data series as well as paleoclimatological records suggests that the climate of the earth may be undergoing a significant change. The climatic change may have important impacts on the hydrologic cycle (the converse is also true). The change may be taking place due to natural and human causes. The major changes that could affect climate are changes in vegetation (changes in evapotranspiration, soil moisture, albedo, and radiation balance), increased use of water for day-to-day needs, burning of fossil fuels, and industrial activities. Large-scale water transfers from one basin to another may also cause climate change in the long run. The change manifests itself in changed patterns of spatial and temporal variability in the components of the hydrologic cycle.

The increased emission of greenhouse gases is believed to be the cause of a gradual increase in the earth's temperature. An increase in the temperature of the atmosphere would lead to higher evapotranspiration; changes in precipitation patterns, timing, and distribution; melting of polar ice caps; and recession of glaciers. Higher melting of polar ice and glaciers and thermal expansion of sea water will cause a rise in the seawater level and inundation of islands of low elevations as well as cities adjacent to seas. Another possibility is that an increase in temperature may mean more precipitation, some of which will be snow at the poles, leading to an additional accumulation of ice. The coupled atmosphere–ocean general circulation models are widely used to study the response of climate to various changes. Notwithstanding a large number of studies, what is going to happen and where? is not known with a sufficient degree of certainty and accuracy.

SUMMARY AND CONCLUSIONS

Water is central to the environment and is considered the driver of nature. All studies of water are concerned with some part of the hydrologic cycle. The occurrence and availability of water are closely linked to human development. As a result of many human activities, the features/characteristics of many components of the hydrologic cycle are undergoing changes that are likely to have long-term effects on climate. It is necessary to predict the magnitude and extent of these changes so that the required ameliorative measures can be initiated in time.

BIBLIOGRAPHY

1. NRC. (1982). *Scientific Basis of Water Resource Management*. National Research Council, National Academy Press, Washington, DC.
2. Singh, V.P. (1992). *Elementary Hydrology*. Prentice-Hall, Englewood Cliffs, NJ.
3. Shiklomanov, I.A. (1999). *World Water Resources: Modern Assessment and Outlook for the 21st Century*. (Prepared in the framework of IHP, UNESCO), State Hydrology Institute, St. Petersburg.
4. NRC. (1991). Opportunities in the Hydrologic Sciences. *Committee on 'Opportunities in the Hydrologic Sciences' of National Research Council*. National Academy Press, Washington, DC.
5. Singh, V.P. (1988). *Hydrologic Systems*. Vol. I. Prentice-Hall, Englewood Cliffs, NJ.
6. Horton, R.E. (1931). *The Field, Scope, and Status of the Science of Hydrology*. pp. 189–202. in Trans. AGU, Reports and Papers, Hydrology. National Research Council, Washington, DC.
7. Seckler et al. (1998). *World Water Demand and Supply, 1990 to 2025: Scenarios and Issues*. Research Report 19. International Water Management Institute. Colombo, Sri Lanka.
8. UNESCO. (1978). *World Water Balance and Water Resources of the Earth*. United Nations Educational, Scientific, and Cultural Organization, Paris.
9. NRC. (1986). *Global Change in Geosphere-Biosphere*. National Research Council, National Academy Press, Washington, DC.

HYDROSPHERE

Yoseph Negusse Araya
Open University
Milton Keynes, United Kingdom

INTRODUCTION

The hydrosphere [Greek *hydor* water and *sphera* sphere] refers to the water on or surrounding the surface of the globe, as distinguished from those of the lithosphere (the solid upper crust of the earth) and the atmosphere (the air surrounding the earth). More specifically, the hydrosphere includes the region that includes all the earth's liquid water, frozen and floating ice, water in the upper layer of soil, and the small amounts of water vapor in the earth's atmosphere. The hydrosphere is the major setting for the earth's hydrologic cycle [see Hydrologic Cycle].

ORIGIN OF WATER ON EARTH

The most significant feature of the earth, in contrast to our neighboring planets, is the liquid water that covers more than two-thirds of the planet's surface. This water came about during the early days of the formation of the earth, when the earth's surface cooled down and the oxygen and hydroxides contained in the accreted material, diffused toward the surface. These gases then cooled and condensed to form the earth's oceans. It is believed that since then, there has been little loss or gain in the overall quantity of the hydrosphere, despite minor fluctuations such as gain from continued degassing and infalling comets and loss in the upper layers of the atmosphere caused by ultraviolet light breaking up water molecules.

DISTRIBUTION AND QUANTITY OF WATER ACROSS THE GLOBE

The earth's water has six major reservoirs in which water resides. These include the oceans, the atmosphere (split into two reservoirs, one over the land and one over the oceans), surface water (including water in lakes, streams, and the water held in the soil), groundwater (water held in the pore spaces of rocks below the surface), and snow and ice. The locations of some major reservoirs on earth are shown in Fig. 1.

The approximate contribution of the different components of the reservoirs to the hydrosphere, the annual recycled volumes, and the average replacement periods are shown in Table 1.

Table 1 highlights the enormous disparity between the huge volume of salt water and the tiny fraction of freshwater and, in addition, the long residence time of polar ice and groundwater, as opposed to the brief period for which water remains in the atmosphere. Some 96.5% of the total volume of the world's water, it is estimated, exists in the oceans and only 2.5% as freshwater. Of this freshwater, nearly 70% occurs in the ice sheets and

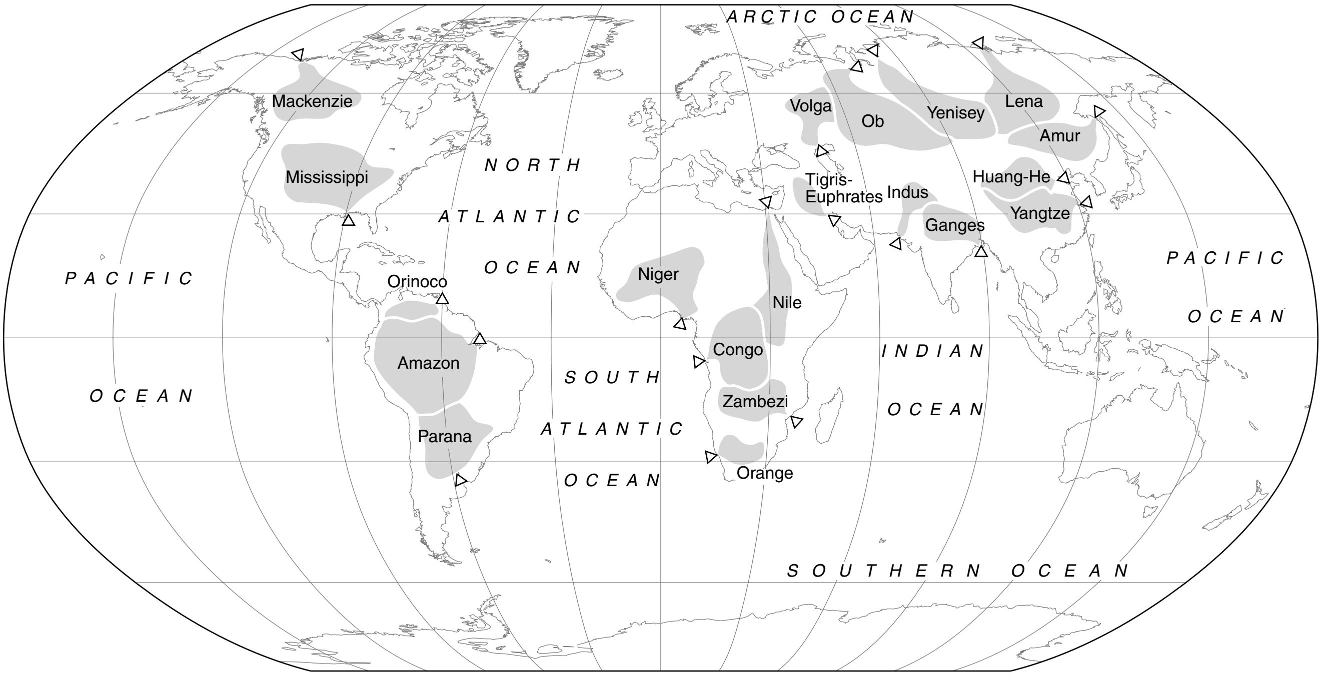

Figure 1. The locations of some major global water reservoirs: oceans and surface water drainage basins (1).

Table 1. The Distribution of Water Across the Globe[a]

Location	Volume 10^3 km^3	% of Total Volume in Hydrosphere	% of Freshwater	Volume Recycled Annually km^3	Renewal Period, Years
Ocean	1,338,000	96.5	—	505,000	2,500
Groundwater (gravity and capillary)	23,400[b]	1.7		16,700	1,400
Predominantly fresh groundwater	10,530	0.76	30.1		
Soil moisture	16.5	0.001	0.05	16,500	1
Glaciers and permanent snow cover	24,064	1.74	68.7		
Antarctica	21,600	1.56	61.7		
Greenland	2,340	0.17	6.68	2,477	9,700
Arctic Islands	83.5	0.006	0.24		
Mountainous regions	40.6	0.003	0.12	25	1,600
Ground ice (permafrost)	300	0.022	0.86	30	10,000
Water in lakes	176.4	0.013	—	10,376	17
Fresh	91.0	0.007	0.26		
Salt	85.4	0.006	—		
Marshes and swamps	11.5	0.0008	0.03	2294	5
River water	2.12	0.0002	0.006	43,000	16 days
Biological water	1.12	0.0001	0.003		—
Water in the atmosphere	12.9	0.001	0.04	600,000	8 days
Total volume in the hydrosphere	1,386,000	100	—		
Total freshwater	35,029.2	2.53	100		

[a] Reference 2.

[b] Excluding groundwater in the Antarctic estimated at 2 million km^3, including predominantly freshwater of about 1 million km^3.

glaciers in the Antarctic, Greenland and in mountainous areas, whereas a little less than 30%, it is calculated, is stored as groundwater in the world's aquifers.

Water moves through the reservoirs by a variety of processes, at different rates, and for unique residence times within any reservoir. This flow of water constitutes the Earth's hydrologic cycle [see Hydrologic Cycle]. A brief summary of the major processes involved in this movement, along with the amount of water transferred per unit time are shown in Fig. 2.

THE BIOGEOCHEMISTRY OF THE HYDROSPHERE

The quality of natural water in the various reservoirs of the hydrosphere depends on a number of interrelated factors. These factors include geology, climate, topography, biological processes, land use, and the time for which the water has been in residence. Table 2 gives a comparison of major elements in selected reservoirs.

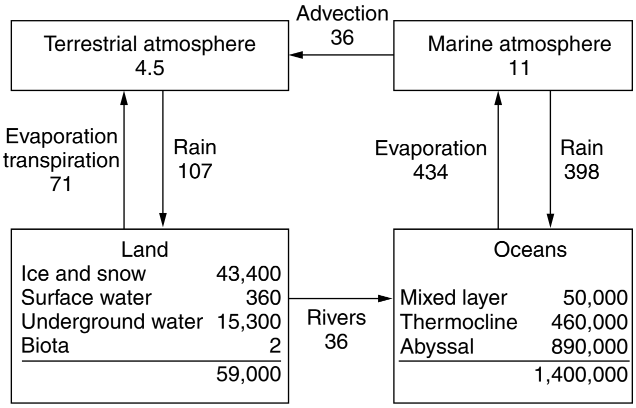

Figure 2. Estimates of global water reservoirs (in 10^{15} kg and 10^{15} kg/yr) global water cycle fluxes (3).

Table 2. Chemistry of Some Hydrospheric Components (in Parts per Million—ppm)[a]

Major Element	Average Seawater	Average Natural River Water	Average Rainwater
Chloride (Cl^-)	19,000	5.75	3.79
Sodium (Na^+)	10,500	5.15	1.98
Sulphate ($SO_4{}^{2-}$)	2,700	8.25	0.58
Magnesium (Mg^{2+})	1,350	3.35	0.27
Calcium (Ca^{2+})	410	13.4	0.09
Potassium (K^+)	390	1.3	0.3
Bicarbonate ($HCO_3{}^-$)	142	52	0.12
Bromide (Br^-)	67	0.02	
Strontium (Sr^{2+})	8	0.03	
Silica (SiO_2)	6.4	10.4	—
Boron (B)	4.5	0.01	
Fluoride (F^-)	1.3	0.1	

[a] References 1 and 4.

Rainwater has a low concentration of nutrients compared to the other reservoirs because it originates as evaporated water vapor and also has a relatively short residence time in the atmosphere. Even so, it is never pure. The major constituents originate from dissolution of aerosol particles, which are formed from natural processes, such as evaporation of sea spray or human activities, such as burning of fossil fuels. Naturally rain water has a slightly acid pH (about 5.5). This results from the formation of mild carbonic acid, when rainwater reacts with atmospheric carbon dioxide:

$$CO_2 + H_2O \longrightarrow H_2CO_3$$

In areas of high emission of sulfur dioxide or nitrogen oxide gases from industrial activities or fossil fuel burning, hydrolysis of rainwater may result in more acidic rain and a pH as low as 4 [see Acid Rain].

River waters have an intermediate concentration of ions compared to that of rainwater and oceans. The main factor controlling the composition or river water is the weathering reaction between rainfall and rocks through which this water passes. An example is that of calcite in limestone, which reacts with carbonic acid of rainfall, as

$$CaCO_3 + H_2CO_3 \longrightarrow Ca^{2+} + 2HCO_3{}^-$$

Lakes also have an intermediate concentration of ions compared to those of river and seawater. Lake waters constitute a reservoir of freshwater, and their composition depends on four factors: the hydrology (e.g., the relative importance of groundwater or surface water inputs, evaporation), the surrounding geology (e.g., carbonate rocks or granite), temperature-driven circulation patterns, and anthropogenic factors (e.g., acid rain, agricultural fertilizers). In some instances, evaporation of water from lakes formed in closed basins may result in a high concentration of salts, as opposed to areas of high rainfall.

Sea and ocean waters are dominated by sodium and chloride ions, followed by sulfate and magnesium. Surface sea water is alkaline at an average pH of about 8. Seawater tends to have a more or less uniform composition in the major elements. But concentrations of minor constituents, including trace and heavy metals and nutrients, vary with depth and location, resulting in marked differences in biological productivity. Organisms living on the surface of the seawater are also involved in changes in its composition via removal of nutrients and breakdown of organic matter at different depths.

Ground water composition is the result of the rock type in which it is confined (e.g., limy is to calcium as argillaceous is to silica); the chemical processes of dissolution, hydrolysis, oxidation-reduction; and biological processes. Moreover, anthropogenic contaminants such as excess fertilizers and heavy metals may also affect the composition of groundwater.

Ice is a pure solid and has thus only few impurities in its structure, but particulate matter and gases may be trapped within it. Analysis of successively trapped gases or other anthropogenic substances such as carbon dioxide in polar ice caps, has been used to study consecutive changes in the atmospheric composition of past times.

EFFECT OF HUMAN BEINGS ON THE HYDROSPHERE

During the last 200 or so years, the sharp rise in population, urbanization, industrial development, and intensification of agricultural practices have combined to affect most natural waterbodies of the earth. This is due to the transport of waste products from those activities by surface water, groundwater, and the atmosphere. The scale and intensity of this pollution vary considerably; there are global problems such as the presence of heavy metals, regional problems such as acid rain, and much more localized ones such as groundwater contamination.

Overall, globally, organic material from domestic sewage, municipal waste, and agroindustrial effluent is the most widespread pollutant. The sewage contains

Table 3. The World's Major Water Quality Issues[a]

Issue Scale	Waterbodies Polluted[b,c]	Sector Affected	Time Lag Between Cause and Effect	Effects Extent
Organic pollution	Rivers++ Lakes + Groundwater +	Aquatic environment	< 1 year	Local to district
Pathogens	Rivers ++ Lakes + Groundwater +	Health ++	< 1 year	Local
Salinization	Groundwater ++ Rivers +	Most uses Aquatic environment Health	1–10 years	District to region
Nitrate	Rivers + Lakes + Groundwater ++	Health	> 10 years	District to region
Heavy metals	All bodies	Health Aquatic environment Ocean fluxes	< 1 to > 10 years	Local to global
Organics	All bodies	Health Aquatic environment Ocean fluxes	1–10 years	Local to global
Acidification	Rivers ++ Lakes ++ Groundwater +	Health Aquatic environment	> 10 years	District to region
Eutrophication	Lakes ++ Rivers +	Aquatic environment Most uses Ocean fluxes	> 10 years	Local
Sediment load (increase and decrease)	Rivers + Lakes	Aquatic environment Most uses Ocean fluxes	1–10 years	Regional
Diversion, dams	Rivers + Lakes + Groundwater ++	Aquatic environment Most uses	1–10 years	District to region

[a] References 5 and 6.

[b] +: serious on global scale.

[c] ++: very serious issue on global scale.

pathogens that lead to disease and mortality among the populations using this water.

Moreover, this organic material has also high concentrations of nutrients, particularly nitrogen and phosphorus, which cause eutrophication (i.e., nutrient enrichment) of lakes and reservoirs. This eutrophication promotes abnormal plant growth and oxygen depletion, which destroys aquatic ecosystems. Excess fertilizers from agricultural production areas also have similar consequences.

Acidification of surface waters as a result of acid rain has adverse effects on aquatic life and also human health. Salinization, the high concentration of salts in the soils of irrigated areas, as a result of poor drainage and high evaporative loss also causes water pollution. Sediments in the form of suspended load may affect physical structures, for example, silting up of dams and damage to aquatic life. Table 3 is a summary of these problems by the type of waterbodies polluted and the extent and reach of the effects.

In general, it may be noted that the effect of pollution can reach far beyond the vicinities of its origin. Moreover, the effects may not be noticed until substantial time has elapsed and it is too late. This calls for constant monitoring and control strategies to save the earth's most precious resource, water.

BIBLIOGRAPHY

1. Ernst, W.G. (Ed.). (2000). *Earth Systems, Processes and Issues*. Cambridge University Press, Cambridge, UK.
2. Shiklomanov, I.A. (in press). *World Water Resources at the Beginning of the 21st Century*. Cambridge University Press, Cambridge, UK.
3. Chahine, M.T. (1992). The hydrological cycle and its influence on climate. *Nature* **359**: 373–380.
4. Meybeck, M. (1979). Concentration des Eaux Fluviales en Eléments Majeurs et Apports en Solution aux Océans. *Revue de Géologie Dynamique et de Géographie Physique* **21**: 215–246
5. United Nations Environment Programme. (1991). *Freshwater Pollution*. Nairobi, United Nations Environment Programme/Global Environment Monitoring System Environment.
6. World Health Organization. (1991). *GEMS / Water 1990–2000. The Challenge Ahead*. UNEP/WHO/UNESCO/WMO Programme on Global Water Quality Monitoring and Assessment, Geneva.

READING LIST

World Water Assessment Programme. (2003). *Water for People, Water for Life: The United Nations World Water Development Report*. UNESCO, Berghahn Books, Paris.

HYDROLOGIC CYCLE, WATER RESOURCES, AND SOCIETY

JASON A. HUBBART
University of Idaho
Moscow, Idaho

THE HYDROLOGIC CYCLE

Water is more or less constantly moving and changing from one state to another (solid, liquid, or vapor) while interacting with the physical processes in the atmosphere, lithosphere, and biosphere. These changes and movements of water are linked together into a grand scheme called the hydrologic cycle. Components of the hydrologic cycle include the water vapor and clouds in the atmosphere, and also liquid surface waters (oceans, lakes, and streams) on continents as well as groundwater. Other important components of the hydrologic cycle include glacial ice held on continents and water in biomass. Plants and animals are about 70% water by volume. Water evaporates in enormous quantities from the oceans and then falls as precipitation either on land or ocean. That portion which falls on land evaporates, evapotranspires, runs off, or infiltrates by some measure. Eventually, all water ends up back in the ocean. Between the various stages of the hydrologic cycle, water moves between temporary storage areas often called reservoirs. These movements are controlled by climatic conditions, which include rain, snow, wind, and other meteorologic processes.

Various pathways exist in the hydrologic cycle (Fig. 1). In nonvegetated systems, interception processes usually do not occur; however, within vegetated systems, precipitation can be intercepted and lost to evaporation. Some precipitation may reach the ground by stemflow but most often reaches the soil by a process called throughfall. Precipitation that reaches the soil surface can then infiltrate and recharge soil waters. From there, it can percolate through soil to deeper soil layers or groundwater supplies. The rate of evaporation from soils or surface waters (including oceans) can be influenced by several factors, including temperature, relative humidity, and wind.

Warmer air can hold more water vapor, thereby generating increased evaporation rates. Drier air can accept much more water vapor than air that is nearly saturated, and windy conditions generally accelerate evaporation. Transpiration from plants takes place via water loss from the stomata, the gas exchange organs on the leaves of plants. The stomata also act as an important cooling mechanism for plants by means of latent heat absorbed and transferred from the plants. This process also supplies the tensional forces necessary to help draw water into roots from the surrounding soil.

Transpiration is often lumped with evaporation and is then called evapotranspiration. Besides evapotranspiration, water can be converted to water vapor in the atmosphere by sublimation. When water is heated enough, it becomes water vapor, which results in evaporation. Sublimation is the result of ice going directly from a solid to water vapor without passing through a liquid state.

Air can hold only a certain amount of water vapor. The amount depends largely upon temperature; thus, as mentioned, hotter air can hold more water vapor than colder air. When air is saturated with water vapor, the water vapor condenses into droplets of water, forming clouds. When the droplets gain enough mass, they fall to the ground as rain or snow. Water that falls as snow or rain evaporates, sublimates, runs off the ground surface (runoff), or soaks into the ground (infiltration).

Whether the water runs off or infiltrates is controlled by a number of mechanisms including precipitation rate, soil water content, slope, and vegetation.

As rates of precipitation increase, runoff rates tend to increase. In extreme cases, rapid cloudbursts of precipitation can result in flash floods and potential land slides. Prior to rainfall, the relative amount of soil water content can dictate the amount of water that soil can hold before saturation. Generally, more water will infiltrate into dry soil than into wet soil, and more water will run off steeper slopes than off moderate slopes. Vegetation can counteract this process by holding more water than hill slopes by themselves. Hillsides, which have undergone timber harvesting or have been burned, often suffer severely from mudslides and flooding. Runoff processes within the hydrologic cycle include stream flow and groundwater flow, each of which eventually results in flow to ocean reservoirs. Lakes and streams may occur in low points of the landscape due to streamflow, or where the groundwater table emerges above the surface. Other contributors of the runoff process include melting of glacial ice and calving of icebergs into the sea.

Globally, precipitation exceeds evapotranspiration over continents, evaporation exceeds precipitation over ocean reservoirs, and excess water moves from continents to oceans as runoff. The ratio of precipitation rate to evaporation rate can vary dramatically between regions. For example, in tropical rainforests, precipitation may greatly exceed evapotranspiration (1).

Often, rates of hydrologic activity are measured in terms of “mean residence time,” or the average amount of time that water remains in its various states. For example, the mean residence time of a water molecule in the atmosphere is very short, usually from days to a week or two. Water tends to move quickly through plant and animal tissues but can be stored in cells for a much longer period of time. It can take weeks to months for water to move through surface drainage networks, depending on the complexity of the geomorphologic network. Water may be stored for months to years in soil water, and individual water molecules may remain in deep groundwater, glaciers, and ocean basins for decades to 10,000 or more years.

Traditionally, quantities of water transferred within the global hydrologic cycle are described in terms of cubic kilometers (km^3) by virtue of the immense quantities held in respective reservoirs. Each km^3 of water contains 10^{12} liters and weighs 10^{15} grams (1). Often, changes in the water cycle are discussed in terms of depth, for example, centimeters per year (cm/yr) or millimeters per hour (mm/hr). Annually, approximately 100 cm of water are evaporated from the surface area of oceans.

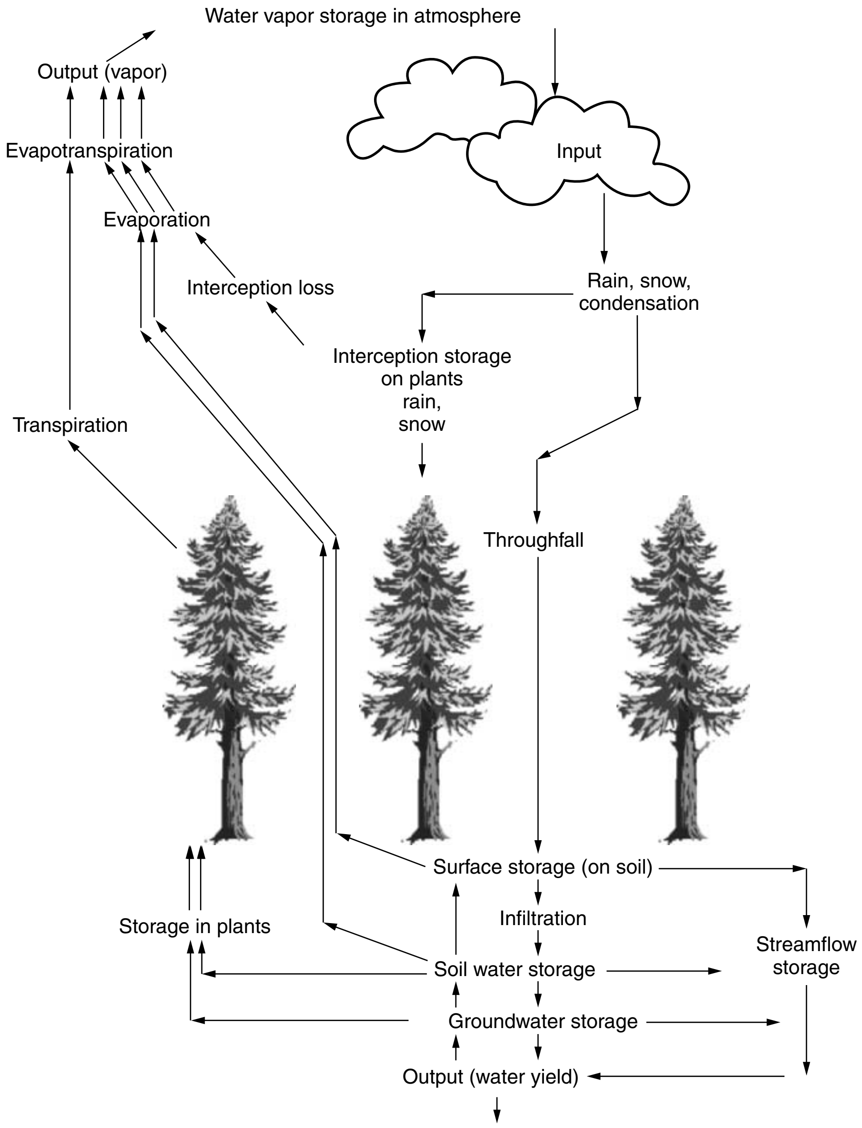

Figure 1. The hydrologic cycle, consisting of water storage compartments and the solid, liquid, and gaseas phases associated within and between each compartment.

Oceans, which also happen to be the dominant pool in the global water cycle, contain approximately 1,350,000 km^3 of water. Of this, approximately 425,000 km^3 are evaporated to the atmosphere each year, 385,000 km^3 are returned directly to the oceans via precipitation, and 40,000 km^3 are delivered to land by rainfall (net transport to land). At any given time, the atmosphere contains about 13,000 km^3 water, or approximately 0.3 cm of rainfall. The large quantities of annual water movement through the atmosphere result in 111,000 km^3 being precipitated upon land and vegetation and 71,000 km^3, either evaporated or transpired (evapotranspiration). Polar and glacial ice reservoirs account for approximately 33,000,000 km^3 water, water held in soils is equal to approximately 122,000 km^3, and groundwaters hold approximately 15,300,000 km^3 (1–3).

WATER RESOURCES

Water resources refer to the supply of groundwater and surface water in a given area. Water resources may also reference the current or potential value of the resource to the community and the environment. The maximum rate at which water is potentially available for human use and management is often considered the best measure of the total water resources of a given region. Approximately 30% of the world's fresh water is in liquid form and therefore potentially accessible for human use and management at any given time (4). The rest is either locked up in polar or glacial ice or water vapor. Of the 30%, almost all is held in groundwaters. Historically, attempts to develop global assessments of available water resources have resulted in limited applicability. The usefulness of resulting aggregated quantities, based on streamflow and population calculations, which lead to measurements in relative abundance and shortages of water regionally, have often been unreliable. The extreme difficulty in preparing a global assessment stems from the general lack of sufficient and reliable information on water availability, quality, and water use in many areas of the world.

Efforts to balance supply and demand and plans for a sustainable future are severely hampered by this lack of reliable information. Studies of water resources leading to meaningful assessments have been found realistic only if conducted regionally or locally. Only then has proper accounting of seasonal and interannual variability of stream flow as well as interactions between groundwater and surface water been appropriately accounted for. Likewise, only then have the potentials for reusing the water as it proceeds downstream and the balance of in-stream and withdrawal uses been appropriately managed. Reduced scale also allows for the importance of water quality in determining suitability for use and, most importantly, the realistic evaluation of social, economic, and political factors that help determine per capita water use (5–8).

Despite these constraints, some attempts to describe global water use have led to reasonable conclusions. For example, estimates indicate that since 1900, global water withdrawal has increased about ninefold and per capita withdrawal has quadrupled. Globally, the largest use of water is for irrigation (70%); industry uses 20%, and 10% is for direct human consumption (9). As a result, humans now withdraw about 35% of the world's reliable runoff. At least another 20% of this runoff is left in streams to transport goods by boats, dilute pollution, and sustain fisheries and wildlife (10). Obviously, these percentages may vary greatly from one region to another, depending on natural precipitation and the degree of development and human population in the region (9). There may be further variance because the distribution of water resources over the landmass of the earth is uneven and unrelated to population demographics or economic development. Because of these complexities and constraints, humans have attempted to increase available water resources by increasing precipitation in various anthropogenic manners, including cloud seeding. Humans have also attempted to decrease evapotranspiration by modifying vegetation, sometimes genetically. The effects of these manipulations are usually minimal and most often temporary but unfortunately often have serious environmental, social, economic, and legal ramifications (4).

SOCIETY

Water is central to survival. Without it, plant and animal life would be impossible. Water is a central component of the earth's ecosystems and provides important controls on weather and climate. Water is likewise central to economic well-being via delivery of direct precipitation and agricultural irrigation, forestry, navigation, waste processing, and hydroelectricity. There are more than 250 international river basins, and in each basin, water moving across diplomatic boundaries can create sensitive and potentially dangerous political challenges. Our present understanding of climate and the consequences of future climate change (whether natural or anthropogenic), together with population growth and economic development, necessitate that water resources are of increasing interest and importance for the foreseeable future.

Reflecting the impacts of population growth alone, water use is increasing everywhere. If per capita consumption of water resources continues to rise at its current rate, humans could be using over 90% of all available freshwater by the year 2025, leaving just 10% for all other living organisms. This forecast does nothing to reflect the high likelihood of climate change. Many studies point to serious consequences for water resources due to climate change, especially in arid regions. For instance, a 1 °C to 2 °C increase in the mean annual air temperature and a 10% decrease in precipitation can create a 40 to 70% reduction in annual runoff in regions that already experience insufficient moisture. The same is true for water resources of large economic regions. American scientists have calculated that a 2 °C climate warming and a 10% reduction in precipitation, can decrease water resources by 1.5 to 2.0-fold in those regions of the country located in arid climate zones alone.

It is well known that the severity of water management problems is determined by the ratio of available water resources to demand. Based on predicted climate changes, plans for further development in irrigated agriculture, reservoir construction, and most water-consuming types of industry will have to be revised. All such revisions will need to result in the greatest positive impacts on hot arid regions. These regions already have difficulties with water supplies and undergo persistent conflicts between different water users and consumers. Regardless, as water use responds to changing demand with some degree of inertia, regional water availability will be determined primarily by the changing water regime within a given territory. The water management systems of river basins will be complex and will have to be extremely flexible and capable of efficient control over water resources under different climate scenarios. In this respect, the regions with large capabilities in river runoff regulation, as is the case in many parts of the United States of America, Canada, and Europe, will have considerable advantages relative to solving water supply and flood regulation, as opposed to those regions with more natural river systems such as regions located in southern and southeastern Asia and South America.

Relatively little account has been taken of the qualitative or quantitative depletion of water resources due to increasing pollution of natural freshwater sources. On this basis alone, all water resource estimates could be construed as optimistic. The major sources of intensive pollution of waterways and waterbodies are contaminated industrial and municipal wastewater as well as water runoff originating from irrigated areas. This problem can be no more acute than in the industrially developed and densely populated regions where relatively little wastewater purification takes place. It is estimated that in 1995, the volume of wastewater was 326 km^3/year in Europe, 431 km^3/year in North America, 590 km^3/year in Asia, and 55 km^3/year in Africa. Many of these countries discharge large portions of their wastewater containing harmful substances into the hydrologic system with no preliminary purification. Prime water resources are thus polluted and their subsequent use becomes unsuitable, especially as potable supplies. Every cubic meter of contaminated wastewater discharged into water bodies and watercourses can potentially spoil between eight and ten cubic meters of pure water. This means that most parts of the world are already facing the threat of catastrophic depletion of water resources due to pollution alone.

Forecasts of future population, industry, and power generation growth are used to predict water withdrawal and water availability. The values obtained for 2025 are ±10–15% for those regions of predominantly developed countries and ±20–25% for regions with predominantly developing countries. To be able to plan for and properly manage these potential future demands for fresh water, it will be vital to achieve close cooperation between scientists and water resource managers. Scientists from different countries and international organizations will have to collaborate and deal with many problems related to hydrology, climatology, and other related disciplines in an cooperative effort to understand better the complex use and protection of water resources.

In human history, water has been central to society and remains so today. Over time, we have grown more aware of the accelerated growth of industry and population correlated to increased uses of water and how that use relates to the earth's fixed supply. In recent times, it has become clearer that human prosperity and prospects for survival vary with the amount and distribution of fresh, unpolluted water. Each year there are millions more humans but no more water than before (11). The same amount of water exists now as when the earth was formed. Ultimately, although finite in quantity, water is a renewable resource. There can be little argument that exponential population growth coupled with an increasing demand for freshwater is resulting in increasing pressures on the resource. The continued degradation, abuse, and pollution of the environment and its resources may cause irremediable damage to our own health and that of the planet on which all life depends. Despite strong feedback mechanisms between hydrologic and meteorologic processes, which have historically led to poor representations of hydrologic processes, increased understanding will lead to improved management strategies for the future and the various processes of the hydrologic cycle.

BIBLIOGRAPHY

1. Schlesinger, W.H. (1997). *Biogeochemistry, An Analysis of Global Change*. Academic Press, San Diego, CA, pp. 346–349.
2. L'vovitch, M.I. (1973). The global water balance. *earth observing soc. (EOS), Trans. Am. Geophys. Union* **54**: 28–42.
3. Chahine, M.T. (1992). The hydrologic cycle and its influence on climate. *Nature* **359**: 373–380.
4. Dingman, S.L. (1994). *Physical Hydrology*, 1st Edn. Prentice-Hall, Upper Saddle River, NJ, pp. 67–68.
5. L'vovitch, M.I. (1974). *World Water Resources and Their Future*. Translated by R.L. Nace. American Geophysical Union, Washington, DC, p. 56.
6. Barney, G.O. (Ed.). (1980). *The Global 2000 Report to the President of the United States*. Pergamon Press, New York, pp. 328–331.
7. Laurenson, E.M. (1987). Assessment of water resources—group report. In: *Resources and World Development*. D.J. McLaren and B.J. Skinner (Eds.). Wiley-Interscience, New York, p. 16.
8. Shiklomanov, I.A. (1989). Climate and water resources. *Hydrol. Sci. J.* **34**: 495–529.
9. Wright, R.T. and Nebel, B.J. (2002). *Environmental Science Toward a Sustainable Future*, 8th Edn. Prentice-Hall, Upper Saddle River, NJ, pp. 214–217.
10. Miller, G.T. (2004). *Living in the Environment. Principles, Connections, and Solutions*, 13th Edn. Brooks/Cole, Pacific Grove, CA, p. 314.
11. Ward, A.D. and Elliot, W.J. (Eds.). (1995). *Environmental Hydrology*. Lewis, CRC Press, Boca Raton, FL, p. 9.

ISOHYETAL METHOD

Sharad K. Jain
National Institute of Hydrology
Roorkee, India

Vijay P. Singh
Louisiana State University
Baton Rouge, Louisiana

The isohyetal method is used to estimate the mean precipitation across an area by drawing lines of equal precipitation. The method uses topographic and other data to yield reliable estimates.

INTRODUCTION

An accurate assessment of the mean areal precipitation is needed in many hydrologic analyses. Precipitation observations from gages are point measurements, and precipitation exhibits appreciable spatial variation over relatively short distances. Numerous methods of computing areal precipitation from point measurements have been proposed, such as the arithmetic average method, the Thiessen polygon method, the Isohyetal method, and kriging. The choice of any method depends on the quality and nature of data, precision required (1), the availability of time and resources, and the preference of the analyst.

Let the precipitation data be available at n stations, spread over an area and P_i be the observed depth of precipitation at the ith station. Using a linear interpolation technique, an estimate of the precipitation across the area can be expressed by

$$P* = \sum_{i=1}^{n} P_i W_i \tag{1}$$

where W_i is the weight at the ith station. Spatial averaging techniques differ in the method of evaluating these weights. The weights of an optimal interpolation technique are chosen so that the variance of error in estimation is a minimum.

The simplest technique for computing the average precipitation depth across a catchment is to take an arithmetic average of the observed depths at gauges within the area for the time period of concern. If the gauges are relatively uniformly distributed across the catchment and the precipitation depths do not vary greatly, this technique yields good results. However, gauges are not always uniformly distributed, there may be large variations in precipitation depths, and topography and orography influence storm movement and precipitation. For these reasons, the arithmetic average is not a preferred method; a technique that can take into account the spatial variability of precipitation is needed.

ISOHYETAL METHOD

Isohyets are contours of equal precipitation analogous to contour lines on a topographic map. In the isohyetal method, precipitation values are plotted at their respective stations on a suitable base map, and isohyets are drawn to create an isohyetal map. Isohyetal lines are based on interpolation between rain gauge stations. While constructing isohyets, it is assumed that rainfall between two stations varies linearly, unless abrupt changes in topography indicate otherwise.

Obviously, this method works best when there are a number of rain gauges. In regions of little or no physiographic influence, isohyetal contours may be drawn taking into account the spacing of stations, the quality, and the variability of the data. In regions of pronounced orography, where precipitation is influenced by topography, analysts take into consideration the orographic effects and storm orientation to interpolate between station values, thus making full use of their knowledge of orographic effects and storm morphology. The areas enclosed by two successive isohyets are multiplied by the average of the isohyets to obtain the volume of precipitation needed to compute the average rainfall depth across the basin. However, a more accurate value of mean areal rainfall is obtained if the lengths of the isohyets are measured as well as the area between the isohyets. If b is the length of the lower value isohyet B, a is the length of the higher value isohyet A, and Δi is the isohyet interval (A–B), then it can be shown that mean rainfall r for the segment is

$$r = B + \frac{\Delta i(2a + b)}{3(a + b)} \tag{2}$$

An isohyetal map resembles a topographical map, except that it contains contours of equal precipitation depths. The topographic parameters that are used to draw isohyets are the elevation of the station, the proximity to mountains and their slope, and the proximity to valleys. If linear interpolation between stations is used, the results will be essentially the same as those obtained by methods, such as the Thiessen polygon method. Clearly, the accuracy of the method depends high on the skill of the analyst; an improper analysis may lead to serious errors. This technique of computing the average precipitation is laborious, if done manually, because the precipitation values need to be interpolated, isohyetal lines are to be drawn, and areas are to be measured for each storm.

Steps of the Isohyetal Method

1. The location of each station is plotted on a map of the catchment.
2. The observed depth of precipitation at each station is marked on the map.
3. The increment of isohyets is chosen, depending on the variation of precipitation values and the scale of the map.
4. Isohyets at selected increments are plotted. In flat areas, linear interpolation between observed precipitation depths is performed. If the area has mountains, valleys, etc., isohyets are drawn keeping in view the influence of these features on precipitation.
5. Starting either from the outermost or the innermost isohyet, the average precipitation between two consecutive isohyets is obtained as the arithmetic average of the value of two isohyets. For example, between the 10 and 20 cm isohyets, the average precipitation would be 15 cm.
6. The area between the two isohyets is measured by using a planimeter or by counting the number of boxes (if the map is drawn on graph paper) and using the scale of the map.
7. Multiplying the area between isohyets by the average precipitation gives the volume of precipitation for that area. The total catchment precipitation is the sum of all such values.
8. The average depth of precipitation is obtained by dividing the sum by the area of the catchment:

$$P = \frac{\sum_{i=1}^{n} P_i A_i}{\sum_{i=1}^{n} A_i} \tag{3}$$

where A_i is the area between successive isohyets and P_i is the average precipitation across this area.

Figure 1 shows an isohyetal map for a catchment that was drawn using point rainfall data. The area enclosed by each isohyet was calculated as given in Table 1. To compute the average catchment rainfall, the average isohyetal value is worked out for each isohyet (the maximum observed rainfall was 108 cm and the minimum 38 cm). The average value multiplied by the area enclosed by the corresponding isohyets gives the volume of precipitation for that area. The volumes for different areas are summed and divided by the area of the catchment to get the average catchment rainfall. The computations are shown in Table 1.

Features of the Isohyetal Method

Important features of the isohyetal method are

- This method is widely regarded as the most accurate method for determining the average precipitation across an area.

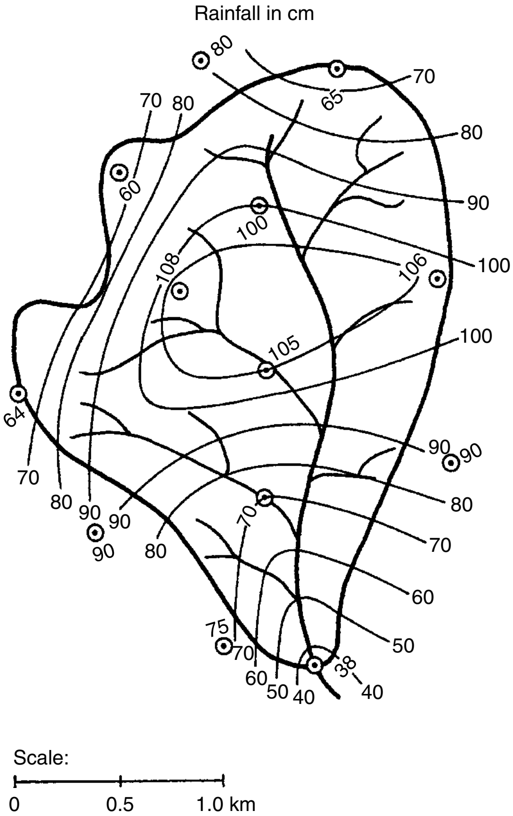

Figure 1. The isohyetal method for computing mean areal rainfall.

Table 1. Estimate of Mean Areal Rainfall by the Isohyetal Method

Isohyet Value, cm	Average Value, cm	Area Enclosed, km^2	Net Area, km^2	Rainfall Volume, km^2-cm
105	106.5	0.79	0.79	84.14
100	102.5	1.52	0.73	74.83
90	95	2.57	1.05	99.75
80	85	3.47	0.90	76.50
70	75	4.50	1.03	77.25
60	65	5.18	0.68	44.20
50	55	5.39	0.21	2.20
<40	39	5.41	0.02	0.78
	Total		5.41	459.65

Average catchment rainfall = 459.65/5.41 = 84.96 cm

- The isohyetal method permits using available pertinent data, such as orographic features and storm morphology.
- This method is subjective because drawing isohyets depends on the individual analyst. Consequently, there is a potential for mistakes.
- The method is time-consuming.
- The isohyetal method is usually applied in the past tense. It is not generally applied in the present to provide inputs for modeling or forecasting.
- This method can be used for spatial averaging of other variables, such as evapotranspiration and snow depth.

There are many computer softwares (2,3) that can be used to draw an isohyetal map from point data. Most of these use a statistical technique known as kriging. The use of software relieves the hydrologist of the tedious job of drawing a map, interpolating precipitation from point data, and drawing isohyets. But most software follow a mathematical algorithm and do not use topographic or storm features. While manually drawing isohyets, most people use judgment rather than actual calculations in interpolation. An alternative is using a computer to draw the isohyetal map which can be refined by the hydrologist later using additional information, experience, and judgment. This results in a considerable saving of time and provides better results.

BIBLIOGRAPHY

1. Singh, V.P. (1989). *Hydrologic Systems*. Vol. II, *Watershed Modeling*. Prentice-Hall, New Jersey.
2. Diskin, M.H. and Davis, P.R. (1970). *A Computer Program for Interpolation and Plotting of Isohyetal Points*. Agricultural Research Service, ARS 41–177, United States Department of Agriculture, Washington, DC.
3. Shearman, R.J. and Salter, P.M. (1975). An objective rainfall interpolation and mapping technique. *Hydrological Sciences Bulletin* **20**(3): 353–363.

WHAT ABOUT METEOROLOGY?

Pacific Northwest National
Laboratory—Shrub-Steppe
Ecology Series

BACKGROUND INFORMATION

Meteorology is the science of the atmosphere. The word meteorology comes from the Greek meteors, which

---

means lofty or elevated and logos, meaning discourse or description. The field of meteorology includes both weather and climate and is concerned with the physical, dynamical (a force producing change or motion), and chemical state of the Earth's atmosphere and interactions between the atmosphere and the Earth's surface. Humans' earliest observations of climate and weather date to hunters, gatherers, and farmers, who studied signs of changing weather in the sky. Today, scientists who study weather are called meteorologists.

Weather and climate are closely related, but big differences exist between them. Weather refers to the condition of the atmosphere at a given time and location. It describes daily local changes in temperature, humidity, wind, clouds, and precipitation (rain, snow). Climate, in contrast, describes the average of these conditions over a long period of time–month, year, or more (annual precipitation, monthly mean, high, and low temperatures) and variability.

WHAT ABOUT THE COLUMBIA BASIN'S CLIMATE?

As you know, the climate of the Columbia Basin tends to be warm and dry. The Cascade Mountains to the west of the Basin greatly influence our climate by means of their "rain shadow" effect, which causes moisture (rain or snow) to be released from the clouds before they reach the area. This limits the amount of precipitation we receive each year. For example, the average rainfall in the lower Columbia Basin is 6 to 7 inches per year, which includes about 13 inches of snow (13 inches of snow equals about 1.5 inches of rain). Because the mountains significantly reduce the amount of moisture in the air, fewer clouds exist in the Columbia Basin, which means more sunshine and very warm temperatures from late spring through early autumn. The average annual temperature recorded at the Hanford Meteorology Station–located on the Hanford Site near Richland–is 53.3 degrees Fahrenheit. The warmest month is July, which averages 76.4 degrees Fahrenheit, while the coolest month is January, which averages 30.5 degrees Fahrenheit. The hottest temperature ever recorded at Hanford was 113 degrees Fahrenheit on August 4, 1961; the coldest was −23 degrees Fahrenheit on February 1 and 3, 1950.

HOW IS WIND CREATED?

In the Columbia Basin, we are all familiar with wind, the kind that can whip up to 50 miles per hour, blowing dust from nearby fields, tumbling trees, and tearing roofs off buildings. But how is wind created? Scientists say it has to do with air pressure. If you have seen what happens to a tire if it's punctured or a balloon when you blow it up and then hold the neck and let the air come back out, you get an idea of this pressure. Air inside a tire or balloon is at high pressure, and air outside of the tire or balloon is at lower pressure. Air comes out of the high-pressure tire or balloon into the lower pressure air of its surroundings. A similar occurrence happens in the Earth's atmosphere. Air moves from regions of high pressure to regions of low pressure (though not in a straight line). This movement of air is what we know as wind. The greater the difference between high pressure and low pressure, the stronger the wind.

WHAT ARE CLOUDS, FOG, RAIN, AND SNOW?

Clouds are formed as air rises into the atmosphere. They are created by very tiny water droplets (it takes a million cloud droplets to make one raindrop). Scientists classify clouds by their shape and by their height (low, medium, and high). Stratus and cumulus are two kinds of low-level clouds. Altostratus and altocumulus are medium-level clouds, and cirrus and cirrostratus are high-level clouds.

Fog is a stratus-type cloud that reaches the ground. As cloud droplets grow, they become heavy and eventually may become so heavy they fall to the ground as rain. In parts of the world where air temperatures are cold enough, water droplets in the clouds freeze into ice crystals and join together into snowflakes. When the snowflakes become big and heavy enough they fall from the clouds–and if the air near the ground remains cold enough (usually near or below freezing) they remain as snowflakes. If the air above and near the ground is too warm, they melt and fall as rain.

WHAT DO METEOROLOGISTS DO?

To become a meteorologist you must go to college and take courses in mathematics, physics, chemistry, and many different courses in meteorology or atmospheric sciences. Meteorologists may have a variety of different jobs. Some meteorologists that many people call "weathermen" may study day-to-day weather events and provide forecasts for the public. Other meteorologists may fly in research aircraft to study clouds, rain, or hurricanes. Meteorologists conducting tornado research may use cars and trucks to try to "chase" these violent storms so they can understand them and provide better forecasts to save lives. Some meteorologists also work mainly in offices with computers. They develop computer models to try to better understand the Earth's atmosphere. Others, primarily interested in research, may work in colleges and universities or in government or private laboratories. Those involved in weather forecasting may work for the government (such as the National Weather Service) or for private weather companies or television/radio stations.

> "Science is constructed of facts as a house is of stones. But a collection of facts is no more a science than a heap of stones is a house."-Henri Poincare
> Keep in mind this fact sheet is intended to be used only as background information to support your effort to encourage inquiry-based science, which parallels the way scientists uncover knowledge and solve problems.

SUGGESTED ACTIVITIES

1. You can have fun determining the monthly average(s) for your community.
   Were you surprised to read that the average temperature at the Hanford Meteorology Station

in July was 76.4 degrees Fahrenheit and 30 degrees Fahrenheit in January? Did you think the average temperature would be higher for our area? Lower?

Here's How: You will need a calendar, pencil, calculator, and access to the local newspaper. Each day check the paper for high and low temperature from the previous day.

1. Record both the high and low temperatures on your calendar.
2. At the end of the month, add all the high temperatures for your community and divide by the number of days in the month. Do the same for the low temperatures.
3. What is the average high? Low?

If you track the weather for a year then you can compute a 1-year average (which describes the climate of your community). What was the maximum temperature for the month? For the year? The minimum temperature for the month? For the year? How much hotter than the average was the hottest day? The coolest night? Ask a friend or relative in another city to keep same data on the temperatures in their community. How do the temperatures they record differ from yours? What are some of factors that might influence the temperature?

2. Do you know about weather in other areas of the world?

You will need a local newspaper, a world map, pencil, and paper that has a grid on it (see example below). Each member of your class or family should select a city from the map to track (make sure the city is one whose temperature and other aspects of weather are reported in your paper. As a group, decide how often you will check the temperature, precipitation, humidity, cloudiness, and windiness in the cities you have selected. Remember to record your data on the data sheet. If you are not familiar with the symbols used for the weather in the newspaper, check the key or talk to a teacher or your parent(s). Watch for news on weather phenomena in your city. Share what you have learned about your city at a weekly class or family meteorological meeting. How could this information help you plan a trip? Decide where to live? (Adapted from Nick Walker, The "Weather Dude" Seattle, Washington)

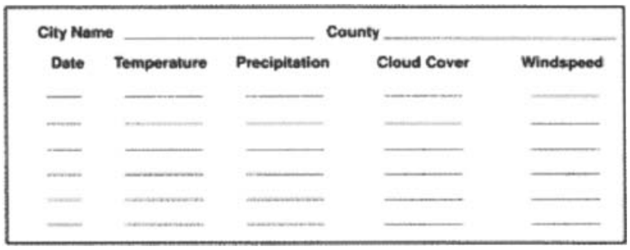

City Name ____________ County ____________

Date	Temperature	Precipitation	Cloud Cover	Windspeed

3. Did you ever wonder how the weather and climate impact plants and animals?

Pick a site in your backyard, or in a park that you can visit on a regular basis. Take along a notebook and sit quietly in the spot you picked. Record in your journal the date, time of day, and the weather (What is the temperature? Is it cloudy, windy, humid?). What animals do you see (insects, birds, fish, mammals)? What are they doing? Where do they go in winter? Can they be found in other habitats? If you can identify the birds, note their names in your notebook. If you don't know the types of birds you see quickly sketch the birds and note any special markings. Later look them up in a bird guide and add the names to your notebook. What plants do you see? Are shoots just starting to emerge from the ground? Are the leaves frostbitten? Visit your favorite spot throughout the year, in different kinds of weather and record your observations in your notebook. Weather and climate influence the life cycle of plants and animals but what other factors influence the plants and animals you observed? How would those plants and animals be affected if the climate changed?

Acknowledgments

Development and printing was funded by the Partnership for Arid Lands Stewardship (http://www.pnl.gov/pals). Written by: Dana Hoitink. Series Editor: Georganne O'Connor; Design: WinSome Design; Printing: Eagle Printing and Graphic Design. First Printing: April 1999.

BIBLIOGRAPHY

1. ***Cloudy with a Chance of Meatballs.*** 1982. Judith Barrett, Judi Barrett. Simon & Schuster. New York.
2. ***Full Option Science System (FOSS),*** Encyclopedia Britannica Educational Corp. Chicago, ***Air and Weather Module,*** Grades K-2.
3. ***How the Weather Works.*** Michael Allaby. 1995. Reader's Digest Association Incorporated. Pleasantville, New York.
4. ***National Audubon Society Field Guide to North American Weather.*** 1991. David M. Ludlum, Alfred A. Knopf. New York.
5. ***Nature Journaling.*** Clare Walker Leslie and Charles Roth. 1998. Storey Books. Pownal, Vermont.
6. ***Puddle Jumpers: Fun Weather Projects for Kids.*** Jennifer Storey Gillis. 1996. Storey Books. Pownal, Vermont.
7. ***Ranger Rick's NatureScope,*** Wild About Weather. National Wildlife Federation. Washington, DC.
8. ***Science and Technology for Children (STC).*** Carolina Biological Supply Company. Burlington, North Carolina, Weather Unit - Grades K-2.
9. ***Weather: Nature Company Discoveries Library.*** Time-Life Books. 1996. San Francisco, California.
10. ***Weather,*** Brian Cosgrove. 1991. Eyewitness Books, Knopf. New York.

WEB SITES

1. **The Hanford Meteorological Station:** http://etd.pnl.gov:2080/HMS/
2. **The "Weather Dude":** http://www.wxdude.com/
3. **The NOAA Weather Page:** http://www.esdim.noaa.gov/weather_page.html

4. **EarthWatch—Weather on Demand:** http://www.earthwatch.com/
5. **How the Weather Works**—http://www.weatherworks.com/
6. **NOAA Weather Education**—http://www.nws.noaa.gov/om/edures.htm
7. **The Weather Network**—http://www.theweathernetwork.com/

BASIC RESEARCH FOR MILITARY APPLICATIONS

W.D. Bach, Jr.

(from *The Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements*, Wiley 2003)

> If you know the enemy and know yourself, your victory will never be endangered; if you know Heaven and know Earth, you may make your victory complete
>
> —Sun Tzu, The Art of War X, 31 circa 500 BC

The ancient Chinese general succinctly summarizes necessary ingredients for success in war. Understanding that "Heaven" represents the atmospheric environment, the need to know the weather is deeply rooted in military preparation and tactics. History is replete with examples of commanders using weather as an ally or suffering its bad effects. Even with today's modern technology, the dream of an all-weather military has not become a reality. Thus the military continues to seek ways to use the atmosphere as a "combat multiplier" in the order of battle.

MILITARY PERSPECTIVE

The military must understand the atmosphere in which it operates. That understanding will ultimately depend on scientific understanding of the atmospheric processes, an ability to use all of the information available, and an ability to synthesize and display that information in a quickly understandable fashion.

The military's needs for understanding the atmosphere's behavior has changed as a result of geopolitics and advancing technology. On the battlefield, future enemies are more apt to resort to chemical or biological attacks as a preferred method of mass destruction. The atmospheric boundary layer is the pathway of the attack. Increased reliance on "smart" weapons means that turbulent and turbid atmospheric effects will influence electromagnetic and acoustic signals propagating through the boundary layer. Intelligence preparation of the battlespace, a key to successful strategies, depends on full knowledge of the enemy, weather, and terrain. It requires an ability to estimate atmospheric details at specific locations and future times to maximize strategic advantages that weather presents a commander. It also avoids hazards and strategic disadvantages.

A convergent theme of basic research in the atmosphere is the interdependence of research progress in acoustic and electromagnetic propagation with the measurement and modeling of the dynamical boundary layer. The scattering of propagating energy occurs because of fine structure changes in the index of refraction in the turbulent atmosphere. The connection between the two—the refractive index structure function parameter, C_n^2—is proportional to the dissipation rate of the turbulent kinetic energy.

Military operations are often in time (seconds to hours) and space (millimeters to tens of kilometers) scales that are not addressed by conventional weather forecasting techniques. Atmospheric information is required for meteorological data-denied areas. Furthermore the real conditions are inhomogeneous at various scales. Significant adverse effects may have short lifetimes at high resolution. Such breadths of scales make accurate quantification of important boundary layer processes very difficult. Carefully planned and executed experiments in the uncontrolled laboratory of the atmosphere are needed for almost every phase of the research. The combined range of atmospheric conditions and potential propagation types and frequencies that are of interest to the military is too broad to condense into a few nomograms. The propagation studies require intensive and appropriate meteorological data to understand the atmospheric effects. Understanding of heterogeneous atmospheric fields arising from inhomogeneous conditions requires measurements achievable only by remote sensors.

Chemical and biological agents constitute major threats to military field units as well as to civil populations. Electromagnetic propagation and scattering are the principal means of remote or in-place detection and identification of agents. Current models of atmospheric transport and diffusion of the agents over stable, neutral, and convective conditions are marginally useful and based on 40-year-old relationships. Significantly improved models of the transport and diffusion are needed to estimate concentrations and fluctuations at various time and space scales for simulated training, actual combat, and for environmental air quality.

RESEARCH ISSUES

Basic research for the military in the atmospheric sciences comes under two broad, interdependent scientific efforts:

Atmospheric wave propagation, which is concerned about the effects of the atmosphere, as a propagating medium, on the transmission of electromagnetic (EM) or acoustic energy from a source to a receptor, and atmospheric dynamics, which is concerned with quantifying the present and future state of the atmosphere.

Atmospheric Wave Propagation

Electromagnetic and acoustic propagation and scattering is a large subject that addresses problems in both characterizing the battlespace and in detecting targets for engagement and situational awareness. Although there has been much previous research in the general area of wave propagation and scattering, there remains a pressing need for research that specifically address issues related to propagation and scattering in realistic atmospheric environments.

Electro-optical and infrared propagation in the atmosphere is limited by turbulence and by molecular and aerosol extinction by both absorption and scattering. The relative importance of turbulence and extinction is dependent on the wavelength of the radiation and the atmospheric conditions. Modern, high-resolution imaging systems are often turbulence limited, rather than diffraction limited, impairing long-range, high-resolution target acquisition, recognition, and identification. Laser-based systems are limited in their ability to retain spatiotemporal coherence and effectively focus at tactical ranges. Like acoustic signals, electromagnetic signals are primarily useful for target detection or ranging and for remote sensing of the atmosphere itself. Although much research has been done in the area of electromagnetic extinction, and current models are generally good, applications of these models for remote-sensing purposes is an area of active research. A large body of research exists in the area of atmospheric turbulence effects for electromagnetic propagation, but with less conclusive results. Hence understanding of atmospheric turbulence for optical and infrared propagation is still an active area of research.

Understanding of the interaction between propagation and turbulence is rather straightforward in the theoretical sense. Practically, the time scales of the EM propagation interactions with the air are much shorter, $O(10^{-9}\text{ s})$ along a path, than the capability to locally sample the atmospheric turbulence or density fluctuations, $O(10^{-1}\text{ s})$ and to characterize the turbulence spectra, $O(10^{+3}\text{ s})$. Acoustic interactions also occur on short time scales. Second, the turbulent structure of the atmosphere between the source and detector is largely unknown, so the scale of the disturbance is undetermined. Furthermore, with models, grid volume samples are realized at $O(10^{0}\text{ s})$ in large eddy simulation models and at $O(10^{2}\text{ s})$ in fine mesoscale models. To make the connection between propagation effects, turbulence characteristics, and atmospheric models, new measurement techniques are needed. In careful field campaigns to relate propagation to atmospheric conditions, the biggest problem is independent ground truth.

Atmospheric Dynamics

The military emphasis on global mobility requires weather forecasting support at all scales from global to engagement scales. In recent years, the explosion in computer capacity enabled highly complex numerical weather prediction models on most of these scales. Model accuracy has improved even as complexities are added. Model results have become more accepted as a representation of the atmospheric environment. As computation power increases, finer and finer scales of motion are represented in smaller grid volumes. Lately, large eddy simulations represent dynamics at grid sizes of a few meters in domains of 5 km^2 by 1 km deep.

As the models go to finer scales, the variability imposed by large-scale synoptic and mesoscale influences the Atmospheric Boundary Layer (ABL) is modified by both the cyclical nature of solar radiation and metamorphosis due to the stochastic behavior of clouds and other natural processes as well as anthropogenic causes. The resulting (stable and unstable) boundary layers are sufficiently different that current models of one state do not adequately capture the essential physics of the other or the transition from one state to the other. Accurate predictions become more difficult, principally because the atmosphere is poorly represented. Data are lacking at the scales of the model resolutions. Parameterizations required to close the set of equations are inadequate. Observations are lacking to describe the four-dimensional fields of the forecast variables at the model resolution. The models are unlikely to improve until better theories of small-scale behavior are implemented and observations capable of testing the theories are available.

For the military, techniques to represent the inhomogeneous boundary layer in all environments—urban, forest, mountains, marine, desert—is absolutely essential for mission performance. In most cases, this must be done with limited meteorological data. Furthermore, the military is more frequently interested in the effects—visibility, trafficability (following rain), ceiling—than in the "weather" itself. Nevertheless, without high-quality, dependable models that represent the real atmosphere in time and space, the effects will not be representative.

Basic research attempts to improve modeling capability by increasing the knowledge base of the processes of the atmosphere. Until new measurement capabilities are developed and tested, our ability to characterize the turbulent environment—affecting propagation and dispersion of materials—will be severely limited.

PROTOTYPE MEASUREMENT SYSTEMS

Military basic research has participated in several new techniques to measure winds and turbulence effects in the atmospheric boundary layer. These techniques concentrate on sampling a volume of the atmosphere on the time scales of (at least) the significant energy containing eddies of the atmosphere.

The Turbulent Eddy Profiler, developed at the University of Massachusetts, Amherst, is a 90-element phased-array receiving antenna operating with a 915-MHz 25-kW transmitter. It is designed to measure clear air echoes from refractive index fluctuations. The received signal at each element is saved for postprocessing. These signals are combined to form 40 individual beams simultaneously pointing in different directions every 2 s. In each 30-m range gate of each beam, the intensity of the return, the radial wind speed is computed from the Doppler shift, and the spread of the Doppler spectrum is calculated. These data are displayed to show a four-dimensional evolution of refractive index structures in the boundary layer. Further processing gives estimates of the three components of the wind velocity. Combining the wind vectors with refractive index structures has shown significant horizontal convergence occurring with strong refractive index structures. At 500 m above the ground the volume represented is approximately a 30-m cube.

The Volume Imaging Lidar at the University of Wisconsin measures backscatter from atmospheric aerosols in three dimensions at 7.5-m range resolutions. The evolution of boundary layer structures can be seen in a variety of experiments. Capabilities to estimate horizontal winds at selected altitudes have been developed and demonstrated.

An eye-safe, scanning Doppler lidar operating at 2 μm has been jointly developed with NOAA/ERL/ETL to measure radial wind speeds at 30-m resolution. The error of the measurement is < 0.2 m/s. Various scanning approaches have shown several different evolutions of the morning transition and breakdown of a low-level jet.

Another lidar system has been developed at the University of Iowa to measure the horizontal wind in the boundary layer at about 5-m increments every minute. Teamed with existing FMCW radars, having comparable vertical resolution, should add to meteorological understanding of layers of high refractive index.

CURRENT MEASUREMENT CAPABILITIES

Reliable measures of atmospheric temperature and moisture at high resolution in space and time are still lacking. Some progress has been made but does not yet achieve the resolutions of the wind measurements.

Prototype and existing wind/wind field data do not yet measure the turbulence. Although the Doppler spectrum width is an indicator of the eddy dissipation rate, few researchers or equipment developers report the variable or its spatial variability. To date, none have done so operationally.

CONCLUSIONS

Some progress is being made by the military to develop necessary instrumentation to measured atmospheric fields at high time and spatial resolution. This is motivated by the need to improve the theory and subsequently the models of atmospheric motions at small scales to provide the armed services with reliable models on which to gain superiority over the adversary and successfully complete the mission.

UNCERTAINTIES IN RAINFALL–RUNOFF MODELING

Patrick Willems
Hydraulics Laboratory
Leuven, Belgium

WHY DEAL WITH RUNOFF PREDICTION ERRORS OR UNCERTAINTIES?

Errors and uncertainties in the prediction of rainfall–runoff are often substantial. Consideration of these errors is therefore crucially important. When water decision-making is based on the rainfall–runoff prediction, for instance, by means of rainfall–runoff, uncertainties affect model outcomes and, consequently, the decisions. Hence, good modeling practice provides not only the results of model predictions but also the accuracy of these results. It supplies decision makers with important additional information on the uncertainty in the data and information that they use as a basis for their decisions. After quantification of the runoff prediction errors, water policies can be set up for which the efficiency can be guaranteed up to specified acceptable risk levels. Combined with the uncertainty analysis, a sensitivity analysis provides the modeler and decision maker with information about the importance of various types of model limitations and sources of uncertainty.

RAINFALL INPUT ERRORS AND UNCERTAINTIES

Rainfall is the driving force of the rainfall–runoff process. Therefore, both for watersheds and for urban drainage catchments, the rainfall input errors are one of the most important sources of uncertainty in the runoff prediction process. These errors have therefore been studied extensively in the literature. In general, they can be classified as rainfall measurement errors and errors by the simplification in the real spatial variability of the rainfall.

Rainfall Measurement Errors

Both rain gauge and radar data are subject to errors in the registration of the rainfall intensity (1–6). For rain gauges, this might be caused by:

- *Rain Losses by Wind Effects, Position, and Shelter*. Rain gauges installed above the ground level or in the lee of trees or buildings typically record less rainfall because of stronger wind effects. In comparison with rain gauges installed at ground level in a pit, gauges at higher elevations can have wind-induced losses up to 5–15% (7–9). The presence of a wind shield on the gauge can reduce this influence (10–13).
- *Rain Losses by Wetting and Evaporation*. All gauges using collecting funnels have losses of rain water remaining on the funnel surface without passing into the collecting chamber, which subsequently evaporates.
- *For Tipping Bucket Rain Gauges: Loss of Water During the Tipping Movements of the Buckets*. This bias can be corrected by calibration of the rain gauge (14–18).
- *For Tipping Bucket Rain Gauges: Resolution Error*. This is the error because of resolution of the recorded rainfall intensity, where the resolution corresponds to the volume of one tipping-bucket. This type of error is only important for small rainfall volumes.

Simplification in the Real Temporal and Spatial Variability of Rainfall

Rainfall–runoff models require simplification in the real spatial distribution of the rainfall. Lumped conceptual models require a larger averaging of the rainfall over the catchment in comparison with semidistributed and fully distributed models. The accuracy of areal rainfall estimates strongly depends on the simplification of the rainfall variability in both space and time. For the spatial dimension, the degree of simplification depends on the rain gauge network density more than on the interpolation method. The effect of the rain gauge network density was first studied by McGuinness (19) and afterwards by Bras and Rodriguez-Iturbe (20); Dawdy and Bergman (21); Wilson et al. (22); Troutman (23); Bastin et al. (24); Schilling (25); Lebel et al. (26); Fontaine (27); Peters-Lidard and Wood (28); Roux et al. (29); Yoo and Ha (30); and Lau and Sharpe (31). Areal reduction factors (or more generally, areal correction factors) have been developed for bias correction when areal rainfall estimates are based on rain gauge data (32,33). Uncertainty estimates for such areal correction factors can be found in Willems (34). They are based on simulations with a spatial rainfall generator. One of the problems with the accuracy of the areal correction factors is locating the desired point within the given area. If the point is away from the center of the area, the correction factors may be vastly different in comparison with correction factors based on historical storms located centrally in the catchment. Correction factors based on individual storm events also may not necessarily look like the climatological rainfall patterns for the whole catchment region. The climatological patterns indeed have to be considered with a reasonable level of confidence (35).

When radar data is available, this data can be used in addition to the rain gauge data. Radar data mainly provides detailed information on the spatial coverage and movement of the rain storms. The rainfall intensity registrations are less accurate, mainly because of the ambiguity of Z–R relationships for conversion of radar reflectivity to rainfall rate. Radar data, thus, can be used in a complementary way to the rain gauge data to increase the accuracy of the spatial interpolation (36,37). The accuracy and benefits of radar data depend, however, largely on the type of the radar and on the data processing (38–41).

PROPAGATION OF INPUT ERRORS THROUGH THE MODEL

To calculate the propagation of the rainfall input errors and related uncertainties to the uncertainty in the runoff prediction, different techniques can be used (the same techniques are also used for propagation of errors or uncertainties on other input variables and parameters). The most popular technique is the Monte Carlo method [either using Simple Random Sampling, Importance Sampling (e.g., 42), or Latin Hypercube methods]. These techniques also allow performing sensitivity analysis, on which the relative contribution of uncertain inputs to the output uncertainty is obtained. Alternatives to the Monte Carlo technique are based on analytical variance calculation. Only the first two moments of the error distributions are considered ("first order second method" reliability analysis methods; 43–47). These methods usually involve some degree of simplification of the model response behavior.

The significance of rainfall input on runoff hydrograph properties, such as runoff volumes, time to peak runoff, and peak runoff rates, has been studied extensively in the literature (21,22,48–54).

In a limited number of studies, the importance of the rainfall input errors in comparison with other uncertainty sources is investigated. For urban drainage stormwater modeling, Schilling and Fuchs (55), Lei and Schilling (56), Willems and Berlamont (57), and Van Mameren and Clemens (58) have made such assessment, and concluded that the input of rainfall and the paved catchment area (which is not so accurately known in urban drainage modeling) are by far the most important uncertainty sources. For watersheds, Andréassian et al. (59) have shown that rainfall–runoff models are more sensitive to the spatial rainfall distribution than to the spatial distribution of watershed (land-surface) parameters.

UNCERTAINTIES IN INPUT DATA AND MODEL PARAMETERS

Other inputs in rainfall–runoff models consist of evapotranspiration estimates, but also in geometric catchment characteristics (at least input of the catchment area is needed), and for distributed rainfall–runoff models, geological, soil type, and land-use maps.

When runoff predictions are based on modeling, the uncertainty on the model results does not only originate from model input uncertainties, but also from uncertainties on the model parameters and the model structure.

Model parameters are most often derived from calibration on the basis of simultaneous observations for the model input and model output variables. Parameter uncertainties will then develop when erroneous and/or limited time series of measurements are used for calibration or when the calibration procedure is not optimal. When inputs are biased (systematic deviation from their real values), the calibrated model parameters might also be biased. This problem has been shown, for instance, by Troutman (23,60,61) and Chaubey et al. (62), investigating the effect of rainfall input errors on runoff volumes. It is obvious that parameter uncertainties decrease when a longer time series of observations is used for calibration.

Parameter uncertainties strongly depend on the parameter calibration process. Optimization algorithms are most often used for parameter calibration (63–69). These algorithms also provide information on the confidence limits of the calibrated model parameter values, as they are based on statistical considerations (see the linear regression theory in any basic statistical handbook as the most simple example). The shuffled complex evolution algorithm (SCE-UA) is one of the most advanced algorithms (70,71). Other techniques are based on maximum likelihood estimation [in combination with Kalman filter in the recursive parameter calibration techniques of Jakeman et al. (72) and Young (73)]. Parameter uncertainty estimates were calculated and discussed for hydrological models by Johnston and Pilgrim (74), Kuczera (75), and Kuczera and Mroczkowski (76).

Goodness-of-fit statistics used in the optimization are the mean squared error, the coefficient of efficiency (77), Young's Information Criterion (73), etc.; for an overview, see Legates and McCabe (78).

Parameter optimization may have difficulties because of the different sources of model errors. For complex models, it might be true that different parameter sets give equally acceptable goodness-of-fits. This problem together with related issues are discussed by several authors, and are denoted by the terms "non-uniqueness," "ambiguity," "equifinality," "nonidentifiability," "overparameterization," etc. (79–81). It may, in addition to other parameter optimization problems (such as the presence of many local optima in the objective function), explain underestimations in the parameter uncertainty estimates.

One method for inferring parameter uncertainties (and also for total model uncertainty), at the same time addressing the above-mentioned problem, is the generalized likelihood uncertainty estimator (GLUE) proposed by Beven and Binley (82). After it was used in a number of studies presented in the hydrological literature (e.g., 83–85), it became more and more popular. GLUE rejects the existence of an optimum model and parameter set and assumed that multiple likely models and parameter sets exist [the "equifinality" concept of Beven (81,86,87)]. This approach is opposed to the more classic approach of identifying the true or optimal model in face of uncertainty. Goodness-of-fit statistics in the GLUE approach are used as likelihood measures. A number of parameter sets is generated via sampling, and the performance of each trial assessed through a likelihood measure. When the likelihood measure reaches a lower limit, then the tried parameter set is called nonbehavioral and is rejected. To the behavioral sets, likelihood-based weights can be applied and parameter and model output uncertainties assessed (88). The problem of the GLUE approach is, however, that is does not account for input uncertainties (89), apart from other statistical problems that make the GLUE inappropriate for uncertainty estimation [e.g., subjective likelihood functions are used and no formal representation is made of the model error (66,90)]. It can, at most, be used for identifying the most sensitive parameters of the model.

For all cases where models are calibrated, the uncertainty in individual parameter values strongly depends on the model sensitivity to each specific parameter. Sensitivity analysis was therefore conducted in many studies.

Input and parameters strongly depend on the scale (temporal and spatial). Often, because of averaging effects, the uncertainty is likely to be much smaller for a larger scale average than for a smaller scale average (see also 91).

Model structure uncertainties, as the last type of model uncertainties considered, originate from the modeler's limited understanding of the physical reality. These can be assessed by comparison of different model structures for the hydrological system under study, or based on the analysis of the model residuals (considering model structure uncertainties as rest term after subtraction of the input and parameter uncertainties; see next sections).

MODEL RESIDUALS

The total uncertainty in the model results (the model output) can be quantified by calculation of the standard deviation or variance (squared standard deviation) of the model residuals (differences between modeled and observed runoff values). This variance can be further split up in the mean error (the bias or systematic deviation in the model output) and the random error (representing the uncertainty after bias correction). An illustration is given in the scatterplot of Fig. 1, where residuals are shown for all hydrographs in a runoff series. The mean deviation and the random error standard deviation are represented by the solid and dotted lines. Both may depend on the runoff value; typically, higher errors are found for higher runoff values. Transformation of both axes in the scatterplot can be done to normalize the residuals—to derive a normal distribution for the residuals (Gaussian residuals) independently of the runoff value (homoscedastic residuals). The Box–Cox (BC) transformation (92) is most flexible; as—depending on its parameter value—it can cover the whole range of weak and strong transformations. The strongest transformation in this range is the logarithmic transformation. The logarithmic transformation is useful in case the standard deviation of the residuals is shown proportional to the runoff value. In this case, the model has a constant relative error. In most practical cases, neither the absolute error nor the relative error is constant, and a BC transformation is needed (see applications by 93). Another transformation applied to reach Gaussian residuals is the normal quantile transform (94). In the case of heteroscedastic residuals, dependence between the residuals and the runoff value has to be considered (95–97), which can be done in the easiest way by the use of a meta-Gaussian model where both the residuals and the runoff discharges are transformed into a Gaussian distribution, and the correlation between both is considered (89,94).

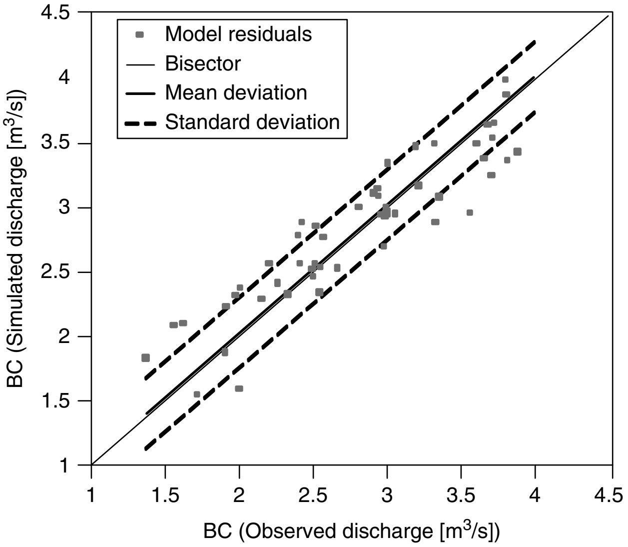

Figure 1. Scatterplot showing model residuals.

SERIAL DEPENDENCE

When runoff series with small timesteps are considered (e.g., hourly or smaller than the recession constants of one of the subflows in the catchment), residuals calculated from runoff values at different timesteps will have a serial dependence. Often, this serial dependence is higher for dry periods in comparison with wet periods, which is caused by the longer recession constant of the baseflow component in comparison with the quick runoff components. To derive independent residuals, a prior selection can be made of nearly independent values from the flow series. These can be peak values during subsequent quick-flow hydrographs, cumulative volumes during these hydrographs, low-flow values during subsequent baseflow recession periods, etc.

Other approaches use the full time series of residuals and describe both the amplitude and the serial correlation by means of stochastic models (98–101). These stochastic models are needed, for instance, on the basis of the recursive parameter estimation techniques.

ERRORS ON RUNOFF OBSERVATIONS

The total residual variance is larger than the total variance of the errors in the model output. Errors on the observed runoff values indeed also take part in the residual variance. These runoff discharges are most often not directly measured, but calculated from river discharge observations. River discharges are most often closely related to the average rainfall–runoff discharge upstream of the river flow monitoring station. Effects of routing along the river and flooding (over bank storage), may explain differences between both. As river flow monitoring stations are only available at a limited number of locations, they are lacking in providing the modeler with a detailed spatial coverage of runoff data. Errors in runoff observations can thus be classified in the following categories:

- errors in the flow data (by uncertainties in the rating curve for limnigraphic stations or in the accuracy of the flow monitoring device);
- differences between the flow at the monitoring station and the upstream runoff because of routing and flooding influences;
- differences between the runoff at a single point along the catchment or river and the average runoff upstream of the flow monitoring station (only relevant for distributed hydrological models).

MODEL ERROR

After compensation for the runoff observation errors, the total variance of the errors in the modeled runoff can be derived from the total residual variance. After calculation of the contribution of the input and parameter uncertainties (by propagation through the model), the remaining uncertainty can then be considered as the model structure uncertainty. When observations are available for some of the submodel output variables, the model structure uncertainties can be decomposed for different submodels.

Such a separate description of the different types of model uncertainties has the advantage that the contributions of the different sources of uncertainty to the total uncertainty in the model output can be quantified and compared. The comparison of the uncertainties resulting from the data, on the one hand, and the uncertainties resulting from the model structure, on the other hand, is especially valuable. Ragas et al. (102) called these two types of uncertainties respectively "operational" and "fundamental" uncertainties. As a general rule, a balance between the two is recommended. Whenever the operational uncertainties dominate, much more attention should be given to data collection than to model structure improvement in an attempt to improve the model results. Whenever the fundamental uncertainties dominate, the reverse is recommended. By comparing the contributions of the different uncertainty sources, efficient measures can thus be determined to reduce the total uncertainty in the model results.

IGNORANCE

Previous paragraphs only consider "quantifiable uncertainty." The real uncertainty may, however, be larger than this quantifiable uncertainty. It indeed might occur that the system structure is depending on specific influences, which were not observed in the past, but which may occur in an unpredictable manner in the future, which is a specific and special type of "lack of knowledge" that we cannot include in our uncertainty predictions because it is not quantifiable. Such additional nonquantifiable uncertainty is usually referred to as "ignorance." Ignorance occurs when we are missing relevant knowledge. Two types of ignorance have to be considered (see also 103):

- *Recognized Ignorance or Accepted Ignorance:* we realize and accept that we are ignorant, and communicate about this;
- *Total Ignorance:* we do not realize this lack of knowledge and are in complete ignorance (or "ignorance about the fact that we are ignorant").

Our confidence in a model may range from being certain (called "determinism" by Harremoës, see Fig. 2) to accepting that we are ignorant (zero confidence,

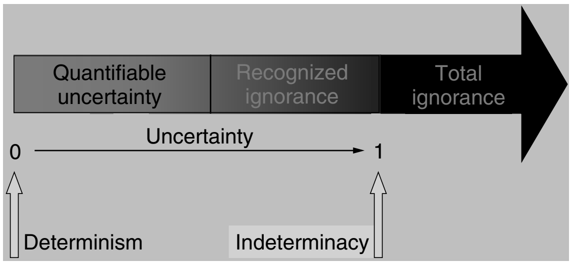

Figure 2. Illustration of the level of uncertainty from the ideal of what we know to total ignorance (103).

“indeterminacy” in Fig. 2). Regardless of our state of confidence, we may be correct or incorrect as a result of ignorance. Nonquantifiable uncertainties, thus, also may contribute to the runoff prediction uncertainty, and may, in some cases, be more serious than the quantifiable uncertainties. Therefore, a need to communicate these uncertainties exists in the decision support (104). When essential processes are not incorporated in the structure of the hydrological model because of insufficient knowledge about the processes, then it would be wrong to forget about this recognizable ignorance and not to communicate this potential source of uncertainty. The communication still can be supported by quantitative information, by running some scenarios on the unknown processes, and to report on the potential results (e.g., epistemic uncertainty, comparing a set of competing models based on different assumptions for the unknown processes).

BIBLIOGRAPHY

1. Larson, L.W. and Peck, E.L. (1974). Accuracy of precipitation measurements for hydrological modelling. *Water Resources Res.* **10**(4): 857–863.
2. Sevruk, B. (1989). Reliability of precipitation measurements. In: *Proceedings of the International Workshop on Precipitation Measurement*. B. Sevruk (Ed.). World Meteorological Organization, Geneva, pp. 13–19.
3. Goodison, B.E., Sevruk, B., and Klemm, S. (1989). WMO solid precipitation measurement intercomparison: Objectives, methodology, analysis, Atmospheric Deposition. *IAHS Publ.* **179**: 57–64.
4. Groisman, P.Y., Koknaeva, V.V., Belokrylova, T.A., and Karl, T.R. (1991). Overcoming biases of precipitation measurement: A history of the USSR experience. *Bull. Am. Meteorol. Soc.* **71**(11): 1725–1732.
5. Groisman, P.Y. and Legates, D.R. (1994). The accuracy of United States precipitation data. *Bull. Am. Meteorol. Soc.* **75**(2): 215–227.
6. Peck, E.L. (1997). Quality of hydrometerological data in cold regimes. *J. Am. Water. Resour. Assoc.* **32**(1): 125–134.
7. Braak, C. (1945). *Wind Influence on rain gauge measurements (in Dutch)*. Royal Dutch Meteorological Institute of the Netherlands, Report 102 ’48. Mededeelingen en Verhandelingen’, Rijksuitgeverij, ’s-Gravenhage, the Netherlands.
8. Neff, E.L. (1977). How much rain does a rain gage gage?. *J. Hydrol.* **35**: 213–220.
9. Sevruk, B. (1996). Adjustment of tipping-bucket precipitation gauge measurements. *Atmospheric Res.* **42**: 237–246.
10. Alter, C.J. (1937). Shielded storage precipitation gauges. *Mon. Weather Rev.* **65**: 262–265.
11. Larkin, H.H. (1947). A comparison of the Alter and Nipher shields for precipitation gages. *Bull. Am. Meteorol. Soc.* **28**(4): 200–201.
12. Yang, D. et al. (1999). Quantification of precipitation measurement discontinuity induced by wind shields on national gauges. *Water Resources Res.* **35**(2): 491–508.
13. Duchon, C.E. and Essenberg, G.R. (2001). Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields. *Water Resources Res.* **37**(12): 3253–3263.
14. Marsalek, J. (1981). Calibration of the tipping bucket raingauge. *J. Hydrol.* **53**: 343–354.
15. Adami, A. and Da Deppo, L. (1985). On the systematic errors of tipping bucket recording raingauges. *Proceedings of the International Workshop on the Correction of Precipitation Measurements*. World Meteorological Organization, Instruments and Observing Methods Report, no. 24, Geneva, pp. 27–30.
16. Niemczynowicz, J. (1986). The dynamic calibration of tipping-bucket raingauges. *Nordic. Hydrol.* **17**: 203–214.
17. Luyckx, G. and Berlamont, J. (2001). Simplified method to correct rainfall measurements from tipping bucket rain gauges. *Proceedings of the World Water and Environmental Resources (EWRI) Congress—Urban Drainage Modeling (UDM) Symposium*. Orlando, May 20–24.
18. La Barbera, P., Lanza, L.G., and Stagi, L. (2002). Tipping bucket mechanical errors and their influence on rainfall statistics and extremes. *Water Sci. Technol.* **45**(2): 1–9.
19. McGuinness, J.L. (1963). Accuracy of estimating watershed mean rainfall. *J. Geophys. Res.* **68**(16): 4763–4767.
20. Bras, R.L. and Rodriguez-Iturbe, I. (1976). Rainfall network design for runoff prediction. *Water Resources Res.* **12**(6): 1197–1208.
21. Dawdy, D.R. and Bergmann, J.M. (1969). Effect of rainfall variability on streamflow simulation. *Water Resources Res.* **5**(5): 958–966.
22. Wilson, C.B., Valdes, J.B., and Rodriguez-Iturbe, I. (1979). On the influence of the spatial distribution of rainfall on storm runoff. *Water Resources Res.* **15**(2): 321–328.
23. Troutman, B.M. (1983). Runoff prediction errors and bias in parameter estimation induced by spatial variability of precipitation. *Water Resources Res.* **19**(3): 791–810.
24. Bastin, G., Lorent, B., Duque, C., and Gevers, M. (1984). Optimal estimation of the average areal rainfall and optimal selection of raingauge locations. *Water Resources Res.* **20**(4): 463–470.
25. Schilling, W. (1984). Effect of spatial rainfall distribution on sewer flows. *Water Sci. Technol.* **16**(8/9): 177–188.
26. Lebel, T., Bastin, G., Obled, C., and Creutin, J.D. (1987). On the accuracy of areal rainfall estimation: a case study. *Water Resources Res.* **23**(11): 2123–2134.
27. Fontaine, T.A. (1991). Predicting measurement error of areal mean precipitation during extreme events. *Water Resour. Bull.* **27**(3): 509–520.
28. Peters-Lidard, C.D. and Wood, E.F. (1994). Estimating storm areal average rainfall intensity in field experiments. *Water Resources Res.* **30**(7): 2119–2131.
29. Roux, C., Guillon, A., and Comblez, A. (1995). Space-time heterogeneities of rainfalls on runoff over urban catchments. *Water Sci. Technol.* **32**(1): 209–215.
30. Yoo, C. and Ha, E. (2002). Basin average rainfall and its sampling error. *Water Resources Res.* **38**(11): 1259, doi:0.1029/2001WR001081.
31. Lau, J. and Sharpe, W. (2003). Comparison between distributed point rainfall measurements and spatially correct rainfall data in urban drainage runoff and quantity predictions. *Proceedings of the 6th International Workshop on Precipitation in Urban Areas*. Pontresina, December 4–7.
32. Rodriquez-Iturbe, I. and Mejia, J.M. (1974). On the transformation of point rainfall to areal rainfall. *Water Resources Res.* **10**(4): 729–735.
33. Lebel, T. and Laborde, J.P. (1988). A geostatistical approach for areal rainfall statistics assessment. *Stochast. Hydrol. Hydraul.* **2**: 245–261.

34. Willems, P. (2001). Stochastic description of the rainfall input errors in lumped hydrological models. *Stochastic Environ. Res. Risk Assess.* **15**: 132–152.
35. Silverman, Rogers, and Dahl. (1981). On the sampling variance of raingauge networks. *J. Appl. Meteorol.* **20**: 1468–1478.
36. Grimes, D.I.F., Pardo-Igúzquiza, E., and Bonifacio, R. (1999). Optimal areal rainfall estimation using raingauges and satellite data. *J. Hydrol.* **222**: 93–108.
37. Sokol, Z. (2003). Utilization of regression models for rainfall estimates using radar-derived rainfall data and rain gauge data. *J. Hydrol.* **278**: 144–152.
38. Joss, J. and Waldvogel, A. (1990). Precipitation measurements and hydrology, In: *Radar in Meteorology*. D. Atlas (Ed.). Amer. Meteor. Society, pp. 577–606.
39. O'Connel, P.E. and Todini, E. (1996). Modelling of rainfall, flow and mass transport in hydrological systems: an overview. *J. Hydrol.* **175**: 3–16.
40. Collier, Ch.C. (1996). *Applications of Weather Radar Systems*. Praxis Publishing Ltd., p. 390.
41. Harrison, D.L., Driscoll, S.J., and Kitchen, M. (2000). Improving precipitation estimates from weather radar using quality control and correction methods. *Meteorol. Appl.* **6**: 135–144.
42. Kuczera, G. and Parent, E. (1998). Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *J. Hydrol.* **211**: 69–85.
43. Wood, E.F. (1976). An analysis of the effects of parameter uncertainty in deterministic hydrologic models. *Water Resources Res.* **12**(5): 925–932.
44. Garen, D.C. and Burges, S.J. (1981). Approximate error bounds for simulated hydrographs. *J. Hydraul. Divis ASCE.* **107**(HY11): 1519–1534.
45. Yen, B.C., Chang, S.T., and Melching, C.S. (1986). First-order reliability analysis. In: *Stochastic and Risk Analysis in Hydraulic Engineering*. B.C. Yen (Ed.). Water Resources Publications, Littleton, CO, pp. 1–34.
46. Kuczera, G. (1988). On the validity of first-order prediction limits for conceptual hydrologic models. *J. Hydrol.* **103**: 229–247.
47. Melching, C.S., Yen, B.C., and Wenzel, H.G. (1990). A reliability estimation in modeling watershed runoff with uncertainties. *Water Resources Res.* **26**(10): 2275–2286.
48. Singh, V.P. (1977). Sensitivity of some runoff models to errors in rainfall excess. *J. Hydrol.* **33**: 301–318.
49. Beven, K.J. and Hornberger, G.M. (1982). Assessing the effect of rainfall pattern of precipitation on modelling stream flow hydrographs. *Water Resources Bull.* **18**(5): 823–829.
50. Hamlin, M.J. (1983). The significance of rainfall in the study of hydrological processes at basin scale. *J. Hydrol.* **65**: 73–94.
51. Corradini, C. and Singh, V.P. (1985). Effect of spatial variability of effective rainfall on direct runoff by geomorphologic approach. *J. Hydrol.* **81**: 27–43.
52. Seliga, T.A., Aron, G., Aydin, K., and White, E. (1992). Storm Runoff Simulation Using Radar Rainfall Rates and A Unit Hydrograph Model (SYN-HYD) Applied to GREVE Watershed. In: *American Meteorological Society, 25th International Conference on Radar Hydrology*, pp. 587–590.
53. Faures, J., Goodrich, D.C., Woolhiser, D.A., and Sorooshian, S. (1995). Impact of small-scale spatial variability on runoff modeling. *J. Hydrol.* **173**: 309–326.
54. Shah, S.M.S., O'Connell, P.E., and Hosking, J.R.M. (1996). Modelling the effects of spatial variability in rainfall on catchment response. 1. Formulation and calibration of a stochastic rainfall field model. *J. Hydrol.* **175**: 67–88.
55. Schilling, W. and Fuchs, L. (1986). Errors in stormwater modelling—a quantitative assessment. *J. Hydraul. Engineer.* **112**(2): 111–123.
56. Lei, J.H. and Schilling, W. (1994). Parameter uncertainty propagation analysis for urban rainfall runoff modelling. *Water Sci. Technol.* **29**(1-2): 145–154.
57. Willems, P. and Berlamont, J. (1999). Probabilistic modelling of sewer system overflow emissions. *Water Sci. Technol.* **39**(9): 47–54.
58. Van Mameren, H. and Clemens, F. (1997). Guidelines for hydrodynamic calculations on urban drainage in the Netherlands: overview and principles. *Water Sci. Technol.* **36**(8-9): 247–252.
59. Andréassian, V. et al. (2004). Impact of spatial aggregation of inputs and parameters on the efficiency of rainfall-runoff models: A theoretical study using chimera watersheds. *Water Resources Res.* **40**: W05209, doi:10.1029/2003WR002854.
60. Troutman, B.M. (1982). An analysis of input errors in precipitation-runoff models using regression with errors in the independent variables. *Water Resources Res.* **18**(4): 947–964.
61. Troutman, B.M. (1985). Errors and parameter estimation in precipitation-runoff modelling, 1. Theory. *Water Resources Res.* **21**(8): 1195–1213.
62. Chaubey, I., Haan, C.T., Grunwald, S., and Salisbury, J.M. (1999). Uncertainty in the model parameters due to spatial variability of rainfall. *J. Hydrol.* **220**: 48–61.
63. Duan, Q., Gupta, V.K., and Sorooshian, S. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Res.* **28**: 1015–1031.
64. Gan, T.Y. and Biftu, G.F. (1996). Automatic calibration of conceptual rainfall-runoff models: optimisation algorithms, catchment conditions, and model structure. *Water Resources Res.* **32**(12): 3513–3524.
65. Gupta, H.V., Sorooshian, S., and Yapo, P.O. (1998). Towards improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Water Resources Res.* **34**: 751–763.
66. Thiemann, M., Torsset, M., Gupta, H., and Sorooshian, S. (2001). Bayesian recursive parameter estimation for hydrologic models. *Water Resources Res.* **37**: 2521–2535.
67. Beven, K. and Young, P. (2003). Comment on "Bayesian recursive parameter estimation for hydrologic models" by M. Thiemann, M. Torsset, H. Gupta, and S. Sorooshian. *Water Resources Res.* **39**(5): COM 1-1–COM 1-4.
68. Gupta, H.M., Thiemann, M., Torsset, M., and Sorooshian, S. (2003). Reply to comment by K. Beven and P. Young on "Bayesian recursive parameter estimation for hydrologic models," *Water Resources Res.* **39**(5): COM 2-1–COM 2-5–.
69. Yapo, P.O., Gupta, H.V., and Sorooshian, S. (1996). Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. *J. Hydrol.* **181**: 23–48.
70. Duan, Q. (1994).
71. Vrugt, J.A., Gupta, H.V., Bouten, W., and Sorooshian, S. (2003). A shuffled complex evolution Metropolis algorithm for optimization and uncertainty assessment of hydrological model parameters. *Water Resources Res.* **39**(6): 1201, doi:10.1029/2002WR001642.
72. Jakeman, A.J., Littlewood, I.G., and Hornberger, P.G. (1990). Computation of the instantaneous unit hydrograph

and identifiable component of flow with application to two small upland catchments. *J. Hydrol.* **117**: 275–300.

73. Young, P. (1986). Time series methods and recursive estimation in hydrological systems analysis, In: *River Flow Modelling and Forecasting*. Craijenhoff, D.A. and Moll, J.R. (Eds.). Reidel, Dordrecht, the Netherlands, pp. 129–180.
74. Johnston, P.R. and Pilgrim, D. (1976). Parameter optimisation for watershed models. *Water Resources Res.* **12**(3): 477–486.
75. Kuczera, G. (1983). Improved parameter inference in catchment models: 1. evaluating parameter uncertainty. *Water Resources Res.* **19**: 1151–1162.
76. Kuczera, G. and Mroczkowski, M. (1998). Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Water Resources Res.* **34**: 1481–1489.
77. Nash, J.E. and Sutcliffe, J.V. (1970). River flow forecasting through conceptual models, 1. A discussion of principles. *J. Hydrol.* **10**: 282–290.
78. Legates, D.R. and McCabe, G.J. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimate model validation. *Water Resources Res.* **35**(1): 233–241.
79. Gupta, V.K. and Sorooshian, S. (1983). Uniqueness and observability of conceptual rainfall-runoff model parameters: the percolation process examined. *Water Resources Res.* **19**(1): 269–276.
80. Beck, M.B. and Halfon, E. (1991). Uncertainty, identifiability and the propagation of prediction errors: a case study of Lake Ontario. *J. Forecasting*. **10**: 135–162.
81. Beven, K.J. (1993). Prophesy, reality and uncertainty in distributed hydrological modelling. *Adv. Water Resour.* **16**: 41–51.
82. Beven, K.J. and Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*. **6**: 279–298.
83. Cameron, D.S., Beven, K.J., Tawn, J., Blazkova, S., and Naden, P. (1999). Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *J. Hydrol.* **219**: 169–187.
84. Cameron, D.S., Beven, K.J., and Naden, P. (2000). Flood frequency estimation by continuous simulation under climate change (with uncertainty). *Hydrol. Earth Syst. Sci.* **4**: 393–405.
85. Blazkova, S. and Beven, K.J. (2002). Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty). *Water Resources Res.* **38**: 1–14.
86. Beven, K.J. (2001). How far can we go in distributed hydrological modelling. *Hydrol. Earth Systems Sci.* **5**: 1–12.
87. Beven, K.J. (2002). Towards an alternative blueprint for a physically based digitally simulated hydrologic response modeling system. *Hydrol. Processes* **16**: 189–206.
88. Freer et al. (1996).
89. Montanari, A. and Brath, A. (2004). A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. *Water Resources Res.* **40**: W01106, doi:10.1029/2003WR002540.
90. Clarke. (1994).
91. Heuvelink, G.B.M. (1998). *Error Propagation in Environmental Modelling with GIS*. Taylor & Francis, London, p. 127.
92. Box, G.E.P. and Cox, D.R. (1964). An analysis of transformations. *J. Royal Statistical Soc.* 211–243, discussion 244–252.
93. Thyer, M., Kuczera, G. and Wang, Q.J. (2002). Quantifying parameter uncertainty in stochastic models using the Box-Cox transformation. *J. Hydrol.* **265**(1-4): 246–257.
94. Kelly, K.S. and Krzysztofowicz, R. (1997). A bivariate meta-Gaussian density for use in hydrology. *Stochast Hydrol. Hydraul.* **11**: 17–31.
95. Sorooshian, S. (1980). Parameter estimation of rainfall-runoff models with heteroscedastic streamflow errors: the noninformative case. *J. Hydrol.* **52**: 127–138.
96. Sorooshian, S. and Dracup, J.A. (1980). Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: correlated and heteroscedastic error cases. *Water Resources Res.* **16**: 430–442.
97. Sorooshian, S. and Gupta, V.K. (1983). Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models: influence of calibration data variability and length on model credibility. *Water Resources Res.* **19**(1): 251–259.
98. Whitley, R.J. and Hromadka, T.V. II (1991). Evaluating uncertainty in design storm runoff. *Water Resources Res.* **27**(10): 2779–2784.
99. Jakeman, A.J. and Hornberger, G.M. (1993). How much complexity is warranted in a rainfall-runoff model. *Water Resources Res.* **29**(8): 2637–2649.
100. Young, P. and Beven, K.J. (1994). Data-based mechanistic modelling and the rainfall-flow nonlinearity. *Environmetrics* **5**: 335–363.
101. Lees, M.J., Price, N., Wheater, H.S., and Peach, D. (1998). A rainfall-runoff simulation model for South Galway, Ireland, In: *Hydrology in a Changing Environment*. Vol. 3. H. Wheater and C. Kirby (Eds.). John Wiley & Sons, New York, pp. 93–104.
102. Ragas, A.M.J., Haans, J.L.M., and Leuven, R.S.W.E. (1997). Selecting water quality models for discharge permitting. *Eur. Water Pollut. Control* **7**(5): 59–67.
103. Harremoës, P. (2003). The role of uncertainty in application of integrated urban water modeling. *Proceedings of the IMUG Conference*, Tilburg, April 23–25.
104. van der Sluijs, J.P. et al. (2003). *RIVM/MNP Guidance for Uncertainty Assessment and Communication: Detailed Guidance (RIVM/MNP Guidance for Uncertainty Assessment and Communication Series*. Vol. 3. Utrecht University, Utrecht.

READING LIST

Fankhauser, R. (1998). Influence of systematic errors from tipping bucket raingauges on recorded rainfall data. *Water Sci. Technol.* **37**(11): 121–129.

Melching, C.S., Yen, B.C., and Wenzel, H.G. (1990). A reliability estimation in modeling watershed runoff with uncertainties. *Water Resources Res.* **26**(10): 2275–2286.

Luyckx, G., Willems, P., and Berlamont, J. (1998). Influence of the spatial variability of rainfall on sewer system design, In: *Hydrology in a Changing Environment*. Vol. 3. H. Wheater and C. Kirby (Eds.). John Wiley & Sons, New York, pp. 339–349.

Radwan, M., Willems, P., and Berlamont, J. (2004). Sensitivity and uncertainty analysis for river water quality modelling. *J. Hydroinformat.* **06.2**: 83–99.

Willems, P. (2000). *Probabilistic Immission Modelling of Receiving Surface Waters*. Ph.D thesis, Faculty of Engineering, K. U. Leuven, Belgium.

MONSOON

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

Cyclic air masses accompanied by either much or little rain, carried by winds that reverse direction semiannually, are called monsoons. The word monsoon comes from the Arabic word “mausim,” which is translated as “season.” Monsoons occur over almost every continent, but they come in varying degrees of strength, as judged by rainfall production and wind speed. The most powerful and well-known monsoons occur in India and Southeast Asia.

Monsoons are caused by a difference in temperature, and thus air pressure, between landmasses and oceans. They are especially strong in Southeast Asia because the area is geographically conducive to monsoon weather. There is much more land in the Northern Hemisphere, namely, the Asian continent, and there is much more water in the Southern Hemisphere, namely the Indian Ocean. Due to the earth’s tilt as it revolves around the Sun, the Northern Hemisphere receives more sunlight during the summer, and the Southern Hemisphere receives more sunlight during the winter. This means that there is always a marked difference between the air temperature over land and the air temperature over the sea in Southern Asia and India. The two variations that result are called wet and dry monsoons. Locally, they are called wet season (summer) and dry season (winter). Wet monsoons occur when the land in the Northern Hemisphere becomes heated by extensive sunlight. Due to convection, the warm air over land rises into the atmosphere, and cooler ocean air slides in to take its place. This cool air carries vast amounts of moisture, which is dropped on the landmass, in this case India and SE Asia.

Dry monsoons occur in the winter when the reverse occurs, and the oceans are heated by sunlight. The air over them rises, and cool air from the land leaves to take its place. This cold air takes moisture with it; thus, there is little rainfall over land.

The precipitation released during the wet season can fall quickly and unpredictably, sometimes resulting in floods, but usually resulting in enormous quantities of much needed rainfall for the farmers and residents of Southeast Asia. Due to this fact, monsoons are integral to the lives and well-being of virtually all of the people who live in India and Southeast Asia, as they have been for millennia. However, despite their importance to the area, they can cause great damage at times, as well. It is not uncommon for abnormal monsoon conditions to result in massive flooding or extended droughts. In fact, a 1987 Indian drought due to variations in the local monsoons was one of the worst of the twentieth century. And parts of India can receive 40 feet of rain or more in a span of less than 4 months.

The monsoon seasons are the distinctive feature of the Southeast Asian climate, but they do exist to lesser degrees elsewhere in the world. Africa, Australia, and the Southwestern United States all undergo minor monsoon seasons during the year. Generally, their monsoon weather is less pronounced than that of Southeast Asia because they do not have its unique geographical properties.

READING LIST

“The Asian Monsoons”. (April 16, 2002). Available: http://www.uea.ac.uk/~f821864/monsoon.html.

“The Indian Monsoon And Its Impact On Agriculture”. (April 16, 2002). Available: http://www.usda.gov/oce/waob/jawf/profiles/specials/monsoon/monsoon.htm.

“Monsoon”. (April 17, 2002). Available: http://www.encyclopedia.com/html/m/monsoon.asp.

“Monsoon”. (April 17, 2002). Available: http://www.pbs.org/wnet/nature/monsoon/html/intro.html.

“Monsoons”. (April 14, 2002). Available: http://library.thinkquest.org/C003603/english/monsoons/glossary.shtml?tqskip1 = 1& tqtime = 0412.

Rosenberg, M. “Monsoons”. (April 14, 2002). Available: http://geography.about.com/library/weekly/aa102599.htm.

AFRICAN MONSOONS

Mike Bettwy
Goddard Space Flight Center,
NASA

In the world of weather, we’ve all heard of the “butterfly effect,” that is, if a butterfly flaps its wings in China, it will affect the weather in New York a few weeks later. While this may be an exaggeration, a new study offers more evidence of how seasonal weather patterns can impact conditions thousands of miles away.

One cyclic weather pattern is the monsoon—a seasonal reversal in wind direction—caused by the fact that land and ocean heat up at different rates. Land quickly responds to heating, while oceans, with a much larger heat capacity,

This photograph of heavy rainfall in Ghana was taken August 19, 2001 during an ongoing NASA/USAID project. The West African monsoon season typically lasts from May-October, with the largest totals observed during June-September. Click on image to enlarge. Credit: Distinguished Professor, Dr. James W. Jones, University of Florida

---

take longer to react. The difference in temperature leads to a difference in pressure, ultimately creating a wind shift which changes rainfall patterns. The monsoon system in West Africa is especially important, as it brings vital rainfall to support life and maintain local economies.

Now, new research by NASA and University of Maryland scientists suggests that these naturally occurring annual swings in temperature and pressure create two distinct monsoons in Africa, one in late spring and another in mid-summer. Using data from satellites, including NASA's Tropical Rainfall Measuring Mission (TRMM), researchers were able to detect a distinct northward "jump" in the monsoon from late spring to mid-summer.

In the first episode, large differences in sea surface temperatures between coastal waters and water further out to sea enhances rainfall on the West African Coast. By July, a second monsoon appears. This downpour is related to African Easterly Waves, low-pressure systems that form on the African continent during the summer months. On average, about 60 waves are generated over North Africa each year, and previous research suggests they may kick off small circulations that develop into hurricanes over the North Atlantic Ocean.

Between May and October, about two Easterly Waves per week travel from the African continent onto the cool Eastern Atlantic. These waves generally fall apart, but their remnants do survive and head towards the Western Atlantic and Caribbean, where they regenerate. Only about one in ten Easterly Waves survive to develop into gale-force tropical storms or full-fledged hurricanes. But, almost 85 percent of major hurricanes (wind speeds 111 mph or greater) began as Easterly Waves. Some studies have also suggested that many of the tropical cyclones in the Eastern Pacific Ocean can also be traced back to Africa.

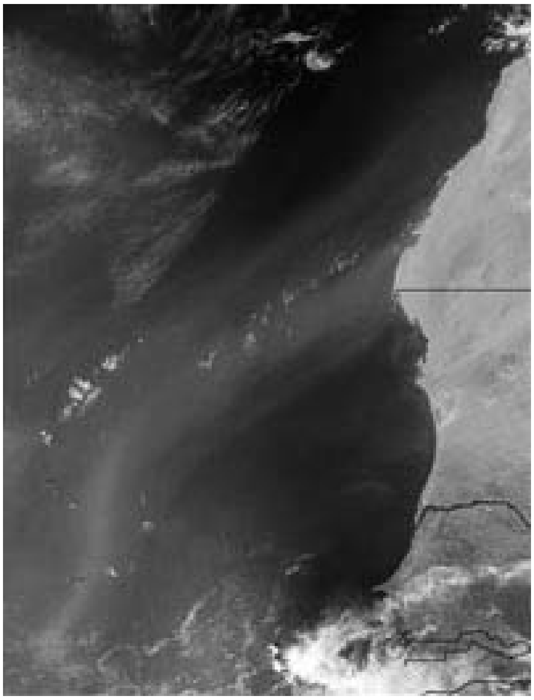

The Moderate Resolution Imaging Spectroradiometer (MODIS), aboard NASA's Aqua satellite, captured this true-color image of Saharan Desert dust blowing southwestward off the coasts of Morocco and Mauritania in West Africa. The light brown plume can be seen wending its way over the Cape Verde Islands in the Eastern Atlantic Ocean. Click on image to enlarge. Credit: NASA/GSFC

While not all Easterly Waves will form into hurricanes, Florida, like Africa, receives a large chunk of its annual rainfall from these storm systems. But other waves that originate to the north of Ghana are dry and pick up debris and dust from Africa's Sahara Desert. A strong high-pressure system over the Atlantic Ocean, known as the Bermuda High, helps steer these waves toward the Florida coast. Easterly Waves tend to be drier in El Niño years, which means they're less likely to produce hurricanes, but more likely to bring pollutants and dust toward the United States.

PERMANENT FROST

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

INTRODUCTION

The last geographic remnant of the last great Ice Age is permafrost, or permanently frozen ground. Permafrost exists today in many places on the earth, including the northernmost parts of America, Eurasia, and the Antarctic. In the coldest regions, permafrost has developed and existed for millions of years. Today, it covers about 20% of the earth's surface.

Frozen soil or rock that remains at a temperature of 0 °C (32 °F) or below for 2 or more years is permafrost. Permafrost is defined solely by temperature and is not determined by the amount of snow cover, soil moisture content, or its location. Permafrost may contain more than 30% ice or no ice at all, and may be topped by meters of snow cover or no snow at all. The defining characteristic of permafrost is its temperature.

REMAINS OF THE GREAT ICE AGES

What remains from the last Ice Age is nearly one-fifth of the Earth's surface still permanently frozen. During the past 2 million years, the great ice sheets that covered much of the earth's Northern Hemisphere prevented the surface below them from thawing. This freeze seeped deeper and deeper into the earth, forming thick layers of permanently frozen ground, permafrost. As time passed, much of this permafrost thawed due to the end of the Ice Age. As the earth's surface began to warm again, permafrost levels declined. However, across the Northern Hemisphere and in Antarctica, much of the ground remains permanently frozen, even during the summer months. Other permafrost forms when the ground cools enough during the winter to prevent it from thawing in the following summer.

Though the determining characteristic of permafrost is ground temperature, the determining cause is not air temperature but atmospheric climate. Temperature instead affects the distribution and thickness of permafrost.

GEOGRAPHY OF PERMAFROST

Permafrost is located throughout the Northern and Southern Hemispheres. The largest continuous mass of permafrost in the Southern Hemisphere is beneath the thick ice of Antarctica. This colossal mass alone comprises about 9% of the earth's surface. Permafrost is more widespread in the Northern Hemisphere. It is found as far north as Greenland and as far south as the heights of the Himalayas. Most of the Northern Hemisphere's permafrost, which makes up 23% of the hemisphere, occurs in the Eastern Hemisphere in parts of Siberia, eastern Russia, northern parts of Mongolia and China, and in the Tibetan Plateau. Other significant permafrost comprises most of Alaska (85% of the state) and parts of northwestern Canada. More than 50% of the surface of Canada is permafrost. Farther north, permafrost levels decline rapidly as land gives way to the Arctic Ocean; yet, a shelf of permafrost has been discovered even under the sea.

Permafrost is separated into two categories: continuous and discontinuous. Continuous implies permafrost everywhere except underneath bodies of water. Discontinuous permafrost is broken into two subdivisions: widespread and sporadic. In the widespread permafrost zone, 50–90% of the land is permafrost. In the sporadic zone, only 10–50% of the ground is permafrost.

LIFE IN A PERMAFROST REGION

A widely held misconception is that a permafrost zone is barren and inhospitable to any form of plant or animal life. On the contrary, the Arctic permafrost zone is home to a variety of mammals, birds, insects, and evergreen forest plant life that covers much of the Northern Hemisphere's permafrost zone.

In Arctic summers, the terrain is alive with animal life. Insects and other invertebrates inhabit the shallow wetlands. Birds feed from these insects and the plant life; they are provided ideal breeding and nesting grounds during the summer months. Snow geese and tundra swans are among the more common visitors of more than 100 bird species, which also include some hawks, owls, and cranes. Polar bears and caribou are also common in the Arctic's permafrost region. Musk oxen, other species of bears, several rodent species, foxes, squirrels, and hares often accompany these Arctic inhabitants. However, when the brief summer comes to a close and the frigid cold reasserts itself upon the region, most animal life leaves in what is essentially a mass exodus.

The trees and shrubs that grow in permafrost forests are much smaller than those in the south and are far more resistant to freezing conditions. Forests tops are evidenced by species of spruce, birch, and aspen trees, and are carpeted below by shrubs, heath, and moss.

PERMAFROST FEARS

Scientists fear that changes to the permafrost regions pose a serious threat to humans and animals. The Earth's natural warming, or human-influenced global warming, has begun to melt the thick permafrost. Also, removal of vegetation and organic cover, as well as forest fires, can also have temperature-raising effects on permafrost regions, causing further thawing of the Permafrost. If this problem continues, the permafrost could turn into a slush mixture. This would cause permanent damage to buildings, roads, pipelines, and other infrastructure in the Arctic region. In particular, oil pipelines, originally set in solid permafrost, would be disrupted. The results could be disastrous, releasing oil into an already fragile ecosystem. Scientists believe this chain of events has begun, but admit that their investigations still remain at a very early stage. However, in Alaska and other Arctic regions, builders are already taking more precautions before building structures on permafrost.

READING LIST

Detecting Permafrost ‘til the Cows Come Home. Alaska Science Forum. (1989). Available: http://www.gi.alaska.edu.

Scientists Fear Cycle of Global Warming. ABC News. (2001). Available: http://www.abcnews.com.

Permafrost at the Geological Survey of Canada. The Geological Survey of Canada. (2001). Available: http://sts.gsc.nrcan.gc.ca/permafrost/general.htm.

The Frozen Underground. The Weather Notebook. (2002). Available: http://www.weathernotebook.org.

New permafrost data may affect climate models study. University of Delaware Update. (2002). Available: http://www.udel.edu.

Heidorn, K.C. (2002). The Weather Doctor Almanac. Ph.D. Available: http://www.islandnet.com.

United States Environmental Protection Agency online database. (2002). Available: http://www.epa.gov.

Available: http://www.geography.uc.edu.

Seifert, R. (1994). *Permafrost: A Building Problem in Alaska*. The University of Alaska Fairbanks, Cooperative Extension Service.

Brown, R.J.E. (1967). *Permafrost Investigation in British Columbia and Yukon Territory*. National Research Council of Canada, Division of Building Research. Technical Paper No. 253, Ottawa, Canada.

RADAR USE IN RAINFALL MEASUREMENTS

Theodore A. Endreny
SUNY-ESF
Syracuse, New York

Meteorologists, hydrologists, and other environmental scientists use microwave radar (radio detection and ranging) techniques to observe remotely the location, type, and intensity of atmospheric precipitation. Doppler radar units provide an additional measurement of storm movement and wind rotational strength. Hence, microwave radar units are used for remotely sensing precipitation and winds.

Radar operates by emitting electromagnetic radiation into the air at various scan heights and then detecting the part of that radiation that is reflected back by hydrometeors (e.g., backscattered) to the radar device and considered reflectivity. Hydrometeors refer to liquid and frozen precipitation, but backscattering can also occur from large flocks of birds, insects, and other large airborne particles. Based on the density, size, and shape of the hydrometeors (e.g., spherical liquid vs. angular frozen crystals) and the polarization of the radar beam, the algorithms processing the radar signal can determine precipitation type and intensity.

RAINFALL AS ENERGY SPECTRUM

The wavelengths of radiation that are most commonly used for measuring precipitation are 3 cm (X-band), 5 cm (C-band), and 10 cm (S-band). National Weather Service (NWS) Weather Surveillance Radar 1988 Doppler (WSR-88D) can detect the radiation at various distances. Beam focusing algorithms are used to ensure that the beam penetrates a limited volume and distance and its signal-to-noise ratio is maximized. When radar beams are backscattered from intense precipitation, detection rises up to a 258-km radius from the radar unit, but reflectivity in weak to moderate precipitation usually limits detection to a 147-km radius.

WSR-88D Doppler radar stations were deployed in the 1990s to replace a network of non-Doppler radar units that were first deployed in the 1950s, after rapid advances in radar during WWII. The WSR-88D (also known as NEXRAD, for next generation weather radar) stations consist of a radar data acquisition unit (RDA) and data post-processing units. The RDA contains the antenna, pedestal, radome (a white spherical radar antenna cover), transmitter, receiver, and signal processor to transmit and then receive the microwave signal. The process of transmitting a signal, receiving and listening for a return signal, and then transmitting the next signal, takes place nearly 1300 times each second. For the 5-minute period to rotate 360° and survey precipitation, the WSR-88D transmits for only 0.6 seconds but receives for 4 minutes and 59.4 seconds to conduct surveillance of beam reflectivity.

Circular scans of 360° about the horizon by the WSR-88D referred to as volume coverage patterns (VCPs), are used to monitor the surrounding volume of the atmosphere. VCPs vary with weather forecasting needs, and there are currently two clear-air mode and two precipitation-mode VCPs. The differences between VCPs affect both the height and length of atmosphere sampled, which are inversely related such that scans that sample at greater heights will be limited to shorter ranges from the radar unit.

In both clear-air mode VCPs, the radar antenna transmits at five tilt angles, which are 0.5°, 1.5°, 2.5°, 3.5° and 4.5° (known as the 1st, 2nd, 3rd, 4th, and 5th tilt) above the horizon, completing a full VCP scan in 10 minutes (see Fig. 1). In both clear-air mode VCPs, two sweeps are taken at the 1st and 2nd tilt, the first sweep to monitor radar reflectivity and the second to monitor Doppler velocity, and a single sweep is taken for the 3rd to 5th tilt when reflectivity and velocity monitoring are combined. Distinguishing the two clear-air mode VCPs is the 3rd tilt, where one mode conducts two sweeps and the other mode one sweep, to monitor reflectivity and velocity. After the radar performs a full 360° revolution at each tilt angle, the radar has completed a full volume scan, and the radar processing algorithms can compute the precipitation rates. Each tilt has a limited range of detection from the radar, which is greatest at the 1st tilt and decreases with higher tilts. Precipitation VCPs begin tilts at 0.5° and end tilts at 19.5° to sample a greater volume of atmosphere. In one precipitation-mode VCP, a total of 14 tilts is sampled in 5 minutes (see Fig. 2); in the other, slower VCP, 9 tilts are scanned in 6 minutes (see Fig. 3).

Ground-based Doppler radar precipitation estimates are often overestimates by a factor of 2 due to their inability to scan the lower volume without ground interference. They often extend a prediction of upper level precipitation readings and do not account for evaporation in the lower 1 km of the atmosphere. Gauge measurements

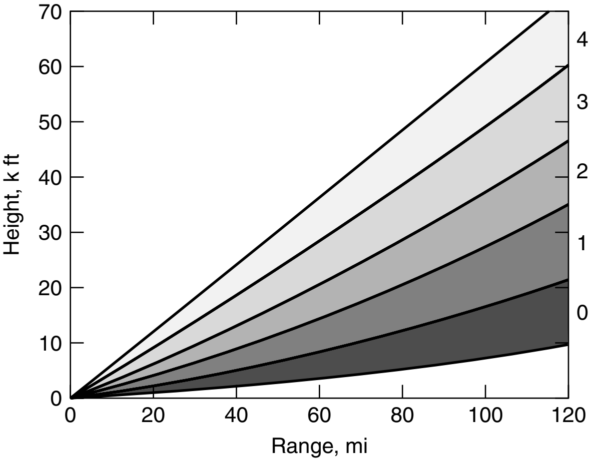

Figure 1. Clear-air mode VCP scan tilt and heights (thousand feet) for 0 to 120 mile ranges.

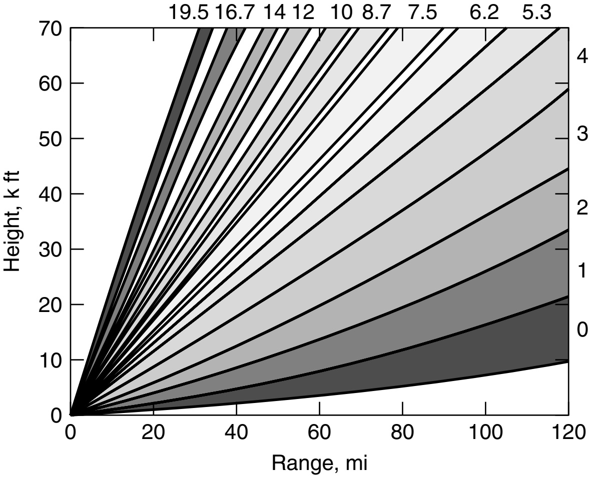

Figure 2. Precipitation-mode VCP scan tilt and heights (thousand feet) for 0 to 120 mile ranges in fast mode.

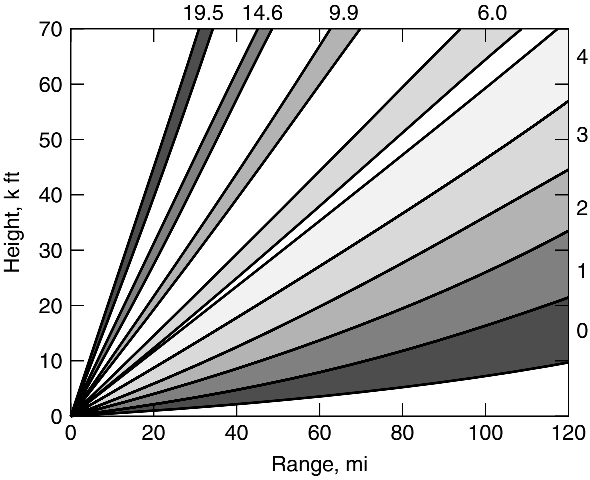

Figure 3. Precipitation-mode VCP scan tilt and heights (thousand feet) for 0 to 120 mile ranges in slow mode.

recorded at the ground may be used to correct some of this bias, and the adjusted radar estimates can provide better spatial estimates than traditional Thiessen polygon and other methods of interpolating point measurements. Extensive research on removing radar bias is ongoing.

Separating transmitting, attenuating, and receiving aspects of the radar is most helpful in understanding Doppler radar. The pulsed-Doppler radar is transmitted by coordinating oscillation and amplification. After the stabilized local oscillator (STALO) generates a nearly perfect sinusoidal continuous wave (cw) signal, which is modulated on and off, it is then amplified to create the powerful microwave signal. These pulses are phase coherent such that their interaction with a moving hydrometeor could displace the phase and allow for Doppler interpretation of hydrometeor movement path and velocity. The microwave beam is manipulated to have a fixed diameter, and an antenna is used to focus the radiation into a narrow angular region, or beam width, containing the peak radiation intensity.

Radar attenuation describes the power loss in the microwave beam due to absorption and scatter during its path through the atmosphere. The total power extracted from the wave is proportional to the combined absorbance and total-scatter cross section encountered and is determined by the highly sensitive radar receiving antennas.

The radar receiver detects the radar reflectivity factor, or Z, which is given units of $\text{mm}^6\ \text{m}^{-3}$, and is proportional to the sample volume of the radar, the number of hydrometeors within that volume, and the 6th power of the diameter of the hydrometeors. Determining the diameter of the hydrometeors is one of the most sensitive steps in determining rainfall rates, for there is often more than one diameter in the volume scanned. Much of the work in this area stems from observations made by Marshall and Palmer in 1948, determining the raindrop diameter for different raindrop size distributions. The Z factor is most commonly expressed as decibels Z, or dBZ, which is 10 times the $\log_{10}$ of Z.

Rainfall rates, R, are computed from the receiver measured Z factor using an empirical power law expressed as $R = \alpha Z^{\beta}$, where the α and β constants are determined from collecting actual rainfall data. For different climates, different values of α and β have been determined; values for α range from 0.013 to 0.028, and those for β range from 0.62 to 0.76. Currently, the NWS uses four different combinations of α and β to estimate rainfall rates for different weather systems. Methods for determining snowfall rates, which are expressed as S, are not as stable as for rainfall rates. Rainfall rates for WSR-88D radars associated with detected dBZ values are provided in Table 1. Note that most bird, insect, and dust detections are usually less than 30 dBZ.

Rainfall products are initially stored in binary HRAP units, which are part of a spherical coordinate system that records intensities. When HRAP units are transformed to a Cartesian system, typically presented at spatial scales of 1 to 4 km bins of average rainfall rates, for time periods per volumetric scan (about 5 minutely), 15 minutely, 1 hourly, 3 hourly, 6 hourly, and storm total precipitation. These rainfall rate products can be viewed on-line in spherical coordinates. They also can be downloaded and reconfigured or purchased from various vendors in the Cartesian coordinate system.

Products available from the WSR-88D radar units linked to the Internet provide “real-time” reflectivity images. These products are categorized as a base reflectivity image from the 0.5° tilt and a composite reflectivity image showing the greatest values sampled from all tilts. The base reflectivity maximum short-range distance provides a 227-km sample radius, and its maximum long-range distance is a 456-km sample radius; both have a spatial resolution at 1 km × 1 km grids. The composite reflectivity maximum short-range distance provides a 227-km sample radius at 2 km × 2 km grids, and the long-range distance provides a 419-km sample radius, but spatial resolution is at a 4 km × 4 km grid.

Algorithms process the data from the integrated volume scan and search for anomalous propagation (AP), such as backscatter from a nonmoving target such as the ground. One method of detecting AP is to look for an abrupt decrease in backscatter between the 1st and 2nd scan, which would indicate that the 1st scan captured surface terrain. Other methods may compare signal strength between different polarization angles or concurrently analyze strength with knowledge of the zero-degree isotherm, or radiosonde determined elevation above which precipitation should be frozen. Another algorithm

Table 1. Rainfall Rates

Reflectivity dBZ	Rain Rate in h^{-1}	Rain Rate cm h^{-1}
65	16.0	40.6
60	8.0	20.3
55	4.0	10.2
52	2.5	6.4
47	1.3	3.2
41	0.5	1.3
36	0.3	0.6
30	0.1	0.3
20	Trace	Trace

technique is to discriminate hail from rainfall or snow by capping the maximum dBZ at 50 or 53.

Radar measured rainfall is also observed from space by the Precipitation Radar instrument, owned by the Japanese, which rides aboard the Tropical Rainfall Measuring Mission (TRMM), a joint U.S. and Japanese satellite. This satellite, which has a low earth orbit of 350 km, was launched in November 1997 to measure precipitation in tropical areas (approximately 30°N and S of the equator) at a resolution of 4.3 km horizontal at 250 m vertical increments and a swath width of 220 km. It is sensitive to a rain rate of 0.5 mm h^{-1}.

Numerous private companies offer rainfall radar, but WSR-88D units are nationally deployed by the NWS, under the federal Department of Commerce and National Atmospheric and Oceanic Administration (NOAA), as well as the Department of Defense and the Federal Aviation Administration. The NWS-88D radar units are deployed in every state, except Connecticut, New Hampshire, and Rhode Island, and in the territories of Guam and Puerto Rico. Most states contain more than one unit; the NWS placed them so that their observed volumes would overlap and create complete coverage for the United States. Daily archives of NEXRAD data files can be downloaded on the Internet at: http://www.nws.noaa.gov/radfiles.html. Current NEXRAD observations are viewable at: http://weather.noaa.gov/radar/mosaic/DS.p19r0/ar.us.conus.shtml.

READING LIST

Doviak, R.J. and Zrnic, D.S. (1993). *Doppler Radar and Weather Observations*. Academic Press, San Diego, CA.

NOAA National Weather Service. Radar Information. Available: http://weather.noaa.gov/radar/radinfo/radinfo.html.

Smith, J.A. (1996). *Radar Hydro Meteorology (CIV 591)*. Princeton University, Princeton, NJ.

RAINFALL

JAMES A. SMITH

(from *The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts*, Wiley 2003)

The capability to measure rainfall advanced dramatically in the last quarter of the twentieth century. The advances have been paced by remote-sensing technologies including both ground-based weather radar and satellite-borne instruments. The most dramatic developments have centered around the capability to monitor precipitation globally from satellite sensors. This measurement capability provides a variety of avenues for hydroclimatological analysis and forecasting. Advances in ground-based radar technologies and deployment of dense networks of rain gages has enhanced the ability to measure rainfall at short time scales (less than 1 h) and small spatial scales (less than 1 km). These time and space scales are often most relevant for water management applications. A brief summary of rainfall measurement and analysis capabilities is presented in the following three sections and organized by the three principal measurement technologies: rain gage, radar, and satellite.

RAIN GAGES

Networks of rain gages play a key role in hydrologic applications ranging from flood forecasting to design of high-hazard structures and water supply management. A wide variety of recording and nonrecording rain gages are used for hydrologic applications. Review and discussion of rain gage technologies are presented in the work by Sumner (1).

There exist several inherent sources of error that affect all types of rain gages. All rain gages suffer from errors due to modification of the wind field by the gage [see Robinson and Rodda (2) for detailed discussions]. The magnitude of errors depends on wind speed, siting characteristics, and type of precipitation (3–9). Rain gage measurement of rainfall is especially difficult in a variety of settings, including mountain ridges, forests, and water bodies. Measurement errors for snow are typically much larger than for rain and are generally in the form of catch deficiencies (3).

Rain gage networks serve as the basis for climatological assessments of precipitation that are used for a wide range of applications (see, e.g., 10). Three of the principal types of climatological analyses that are used for water management applications are illustrated in Figs. 1 to 3. Assessments of average rainfall conditions, in a variety of forms, are central to activities involving industrial, municipal, and agricultural water use. Mean annual precipitation is shown in Fig. 1 [see also Groisman and Legates (3) for a discussion of biases in rain gage analyses of mean precipitation]. Global assessments of continental precipitation have been developed from rain gage observations by Legates (11) [see also Legates and Wilmott (12)]. Precipitation frequency analysis plays a central role in engineering design problems, especially in urban areas (13). The 15-min, 100-year rainfall magnitude for the United States (14) is illustrated in Fig. 2. The network of gages that have the temporal resolution to provide short-term precipitation frequency analyses, such as those in Fig. 2, is far less dense than the rain gage network used to produce mean annual precipitation maps. Consequently, it is difficult to assess the true geographic variability of extreme rainfall rates. It is likely that geographic features, such as mountains and land–water boundaries, exert a pronounced influence on the frequency of extreme rainfall rates. The density of the network, however, is not adequate to resolve these geographic variations. Design of high-hazard structures, such as spillways on major dams, is determined through probable maximum precipitation (PMP) analyses (15,16). Rain gage data sets, in the form of storm catalogs, play a central role in PMP analyses. Storm catalogs for PMP analyses consist of gage observations from specific events. Consequently,

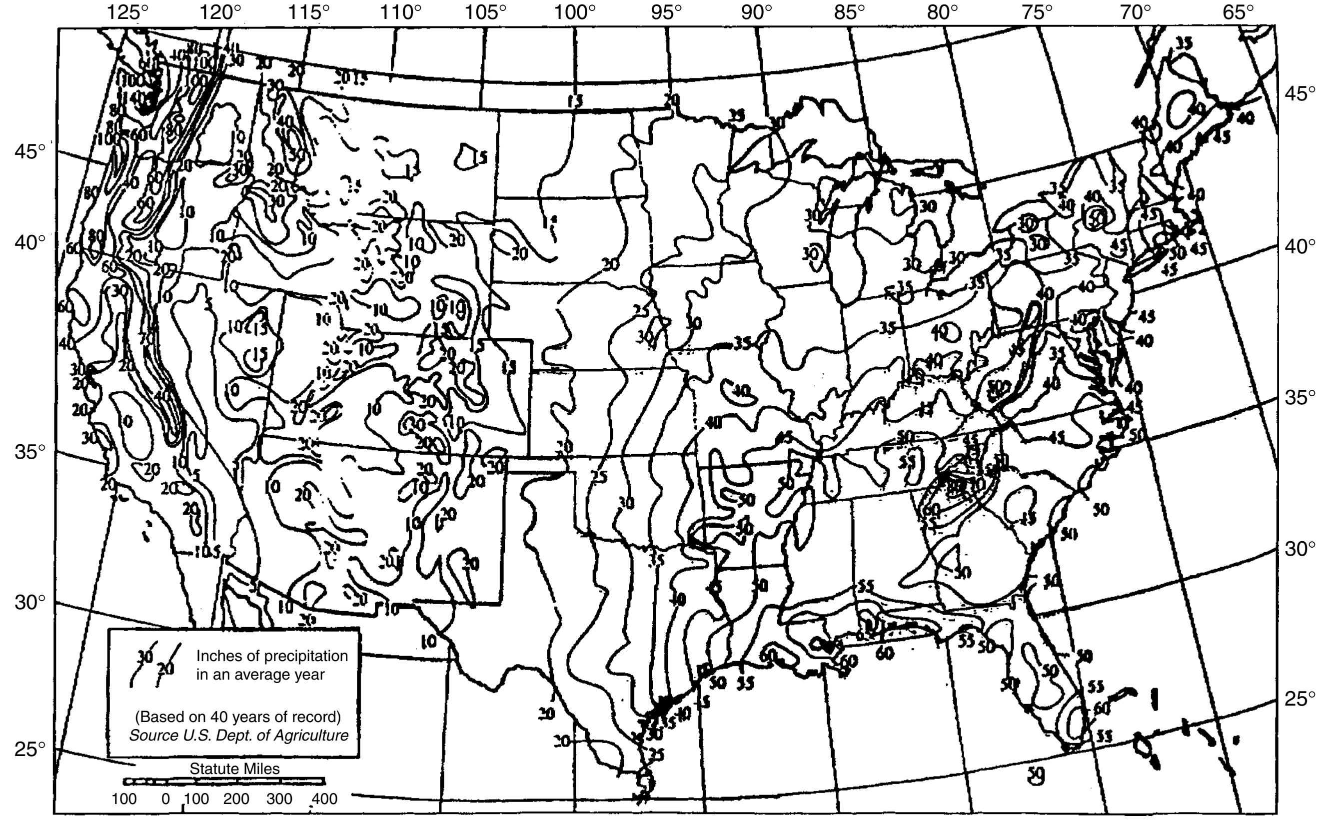

Figure 1. Mean annual precipitation (inches) for the United States from rain gage observations.

the density of gage observations in regions experiencing catastrophic rainfall is critical for PMP analyses. The 6-h, 200 mi^2 PMP for the eastern United States is shown in Fig. 3. The greatest uncertainties in PMP analyses are for small areas (less than 200 mi^2), short time periods (6 h and less), and for regions of complex terrain (17).

RADAR

Implementation of the NEXRAD (next-generation weather radar) system of WSR-88D (weather surveillance radar—1988 Doppler) radars has resulted in dramatic advances in rainfall measurement capabilities for the United States (18). Operational National Weather Service (NWS) rainfall products derived from WSR-88D observations provide rainfall analyses for the United States at 1-h time resolution and spatial resolution of approximately 4 km (19).

The hourly digital product (HDP) rainfall estimates are created by the WSR-88D radar product generator on a 131×131, 4-km grid centered at each radar site. The range over which rainfall products are constructed for each site is approximately 230 km. The algorithm used to construct this product (20) consists of the following steps: (1) quality control, including identification and elimination of anomalous propagation returns, (2) conversion of radar reflectivity factor to rainfall rate through a Z-R relationship, (3) correction for range effects, (4) aggregation of rainfall estimates to hourly, 4-km grid scale, and (5) bias correction using rain gage observations. The HDP product is the base rainfall product from the NEXRAD system. Detailed assessments of HDP algorithm performance are presented in Smith et al. (21) and Baeck and Smith (22) [see also Joss and Waldvogel (23), Wilson and Brandes (24), and Anagnostou and Krajewski (25)].

In a second stage of WSR-88D rainfall processing, multisensor precipitation analyses employ rain gage observations and the 4-km HDP rainfall fields in an optimal estimation framework developed by Krajewski (26) and Seo (27,28). These rainfall fields are subsequently composited into a regional mosaic. The regions that comprise the individual mosaics correspond to the watershed boundaries that delimit the NWS River Forecast Center areas of coverage. Algorithms used for mosaicking of multiple, overlapping radar coverages are described in Seo et al. (29). A national, hourly precipitation analysis is produced at the National Centers for Environmental Prediction (NCEP).

The national 4-km, hourly rainfall mosaic produced by NWS from rain gage and WSR-88D rainfall products will provide an important source of rainfall information for climatological analyses, especially as the observing period

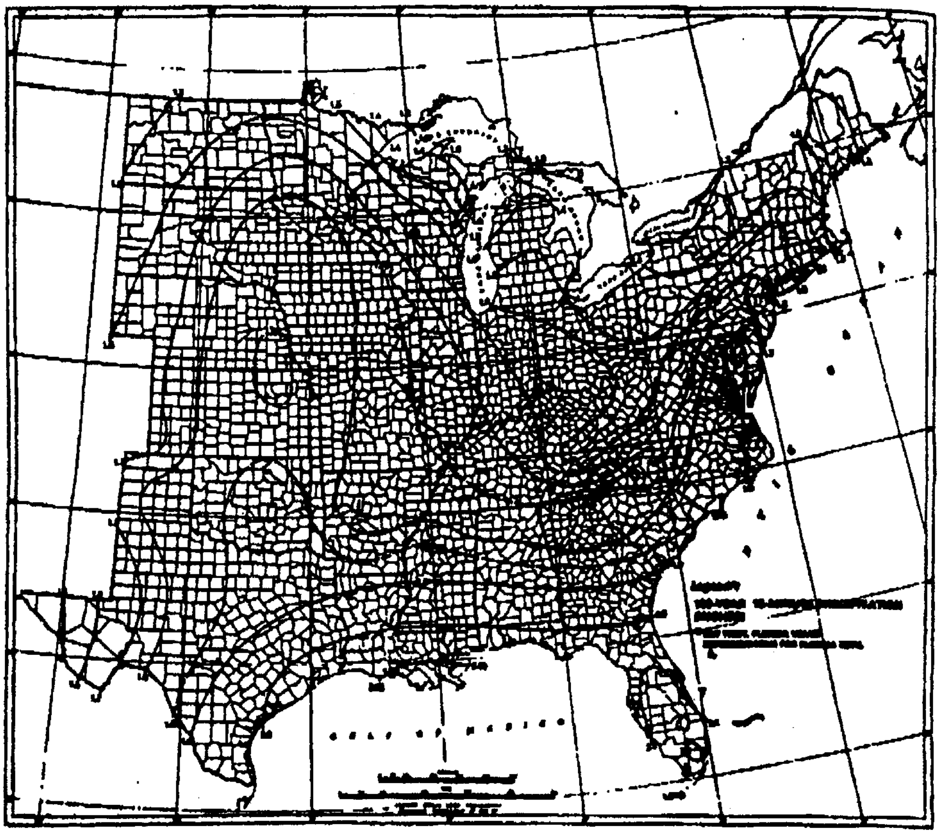

Figure 2. The 100-year, 15-min rainfall magnitudes (inches) for the United States east of the Rocky Mountains.

increases. Radar observations have not generally served as the basis for climatological analyses of rainfall [see, however, Baeck and Smith (30) for an exception]. Issues of bias in radar rainfall estimation must be addressed for radar-based rainfall databases to be most useful for climatological studies (21).

Radar polarimetric measurements (31–34), which utilize the capability of radar to transmit and receive electromagnetic radiation at alternating polarization, hold promise for providing significant improvements in rainfall estimates. Polarization measurements have been shown to be quite useful for quality control algorithms, including detection of bright band, hail, and AP [anomalous propagation (of radar waves, due to sharp gradients of water and air density)], as well as for algorithms for estimating rainfall rate (31,35). The NEXRAD network was designed for eventual implementation of polarization measurements by the WSR-88D.

Radar has provided a significant component of the observational basis for studying storms that produce extreme rainfall. Chappell (36) and Doswell et al. (37) summarize key elements of heavy rainfall producing storms with particular emphasis on storms that produce large point rainfall accumulations through small net storm motion [see also Maddox et al. (38)]. These storms have been termed quasi-stationary convective systems (36). Houze et al. (39) provide a detailed summary of radar-derived storm structure for severe thunderstorms in the central United States [see also Perica and Foufoula-Georgiou (40) and Steiner et al. (41)].

WSR-88D observations, and the rainfall products derived from these observations, have provided a new playing field for hydrologic application and science. Many hydrologic problems that were previously not possible to address due to an absence of information concerning rainfall, have been attacked from an observational perspective. Numerous examples can be drawn from flood hydrology. Figure. 5 illustrates a storm total rainfall analysis constructed for the rapidan storm of June 27, 1995 (42). More than 600 mm of rain fell on the east slope of the Virginia Blue Ridge during a 12-h period resulting in record unit discharge for the United States east of the Mississippi River and catastrophic landslides and debris flows. Fluvial and geomorphic impacts of the rapidan storm rival those described in the classic study by Hack and Goodlett (43) for the June 1949 storm in the Shenandoah Mountains. The chief difference between studies of the 1949 and 1995 storms is rainfall measurement at the 1-km horizontal scale and 6-min time scale for the 1995 storm that allows direct assessment of hydrologic processes.

SATELLITE

Satellite-borne instruments have proven useful for monitoring precipitating cloud system since the 1960s. Steady

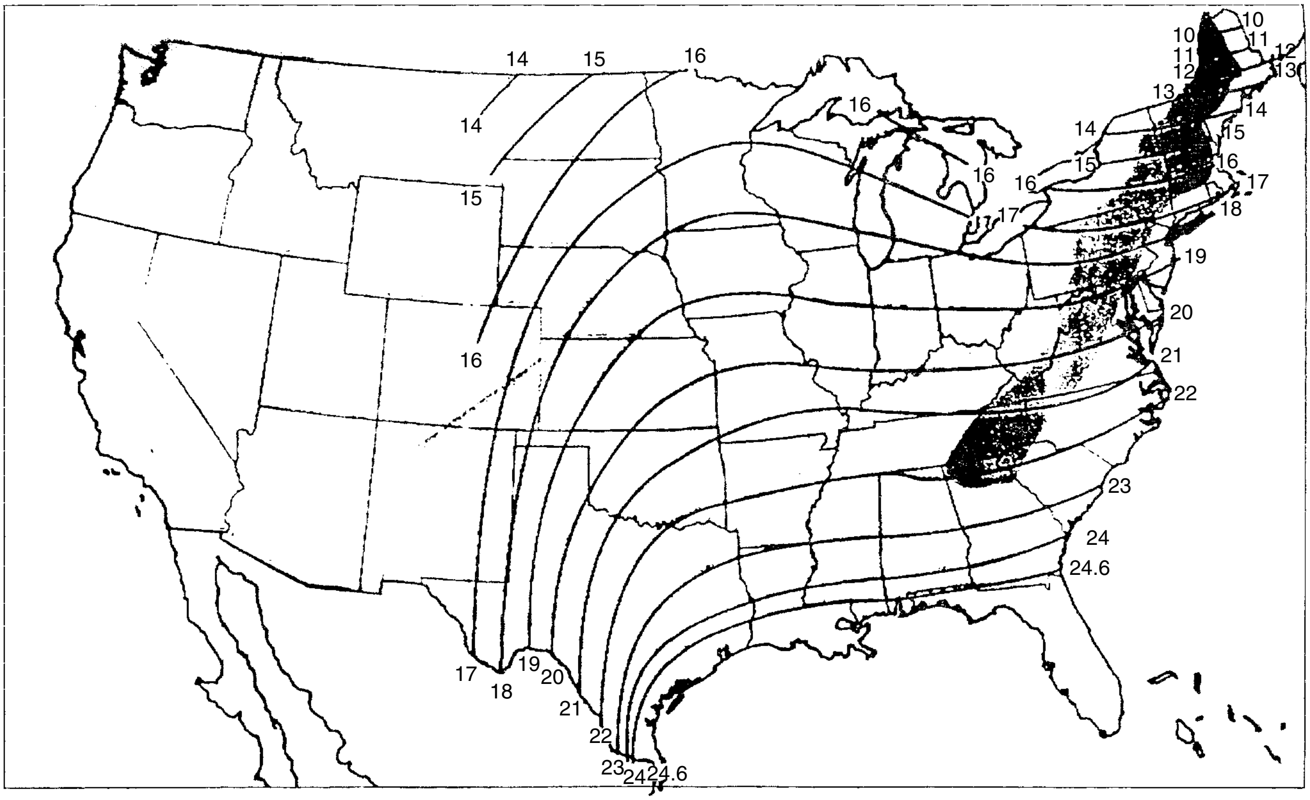

Figure 3. The 6-h, 200 mi^2 PMP magnitudes (inches) for the eastern United States.

progress has been made in developing algorithms for removing rainfall accumulations from passive satellite observations in the microwave (44,45) and infrared (46–48) portions of the electromagnetic spectrum. This progress is reflected in rapidly advancing capabilities for hydroclimatological analysis (49–53). Techniques for quantitative precipitation estimation from satellite sensors are reviewed below.

A geostationary, infrared-based satellite algorithm (46) has been developed and implemented for heavy rainfall measurement. This algorithm is specifically designed for deep, moist convective systems. Estimated precipitation rates are based on the cloud top temperature obtained from the 10.7-μm infrared channel. The empirical equations used to relate cloud top temperature and rainfall rate were calibrated from radar data sets consisting of observations from thunderstorm systems. A moisture correction factor obtained from the precipitable water and mean relative humidity is used to adjust the estimates for different moist environments. The technique of relating rain rate and cloud top temperature tends to overestimate the rain area in some cases and rain rate in others. The technique is also subject to underestimation of rain rates in warm cloud top environments and overestimation of cold top storms in strong wind shear environments.

The *Tropical Rainfall Measuring Mission* (*TRMM*) satellite (54) is designed to measure tropical precipitation and its variation. With the inclusion of a precipitation radar, TRMM provides the first opportunity to estimate the vertical profile of the latent heat that is released through condensation. The TRMM rainfall data will be particularly important for studies of the global hydrological cycle and for testing the ability of climate models to simulate climate accurately on the seasonal time scale.

The TRMM instruments for rainfall observation consist of a precipitation radar, a multifrequency microwave radiometer, and a visible and infrared (VIS/IR) radiometer. The precipitation radar provides measurements of the three-dimensional rainfall distribution over both land and ocean. The precipitation radar will permit the measurement of rain over land where passive microwave channels have difficulty. The horizontal resolution is approximately 4 km, the range resolution is 250 m, and the scanning swath width is 220 km. The multichannel microwave radiometer provides information on vertically integrated precipitation, its areal distribution, and its intensity. Rainfall analyses using the microwave radiometer are best suited for open ocean conditions. The visible infrared (IR) scanner provides high-resolution information on cloud coverage, type, and cloud top temperatures and serves as the link between these data and the long and virtually continuous coverage by the geosynchronous meteorological satellites. The instrument, with a swath width of 720 km, will provide cloud distributions by type and height and rain estimates from brightness temperatures at a horizontal resolution of approximately 2 km.

Satellite IR observations from geostationary satellites have been used extensively for assessing the climatology of extreme rainfall producing storms. An extensive

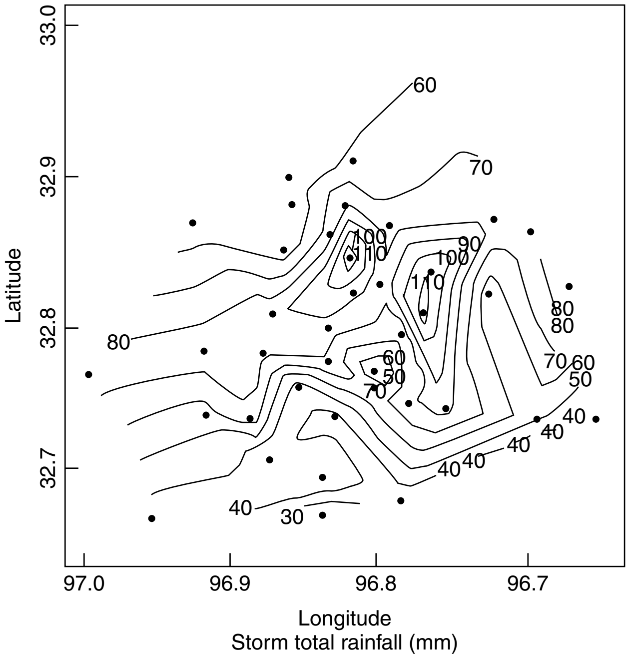

Figure 4. Storm total rainfall (mm) from the Dallas Metropolitan Area mesonet for the Dallas hailstorm of May 5, 1995. The dimensions of the surrounding box are approximately 30 × 30 km.

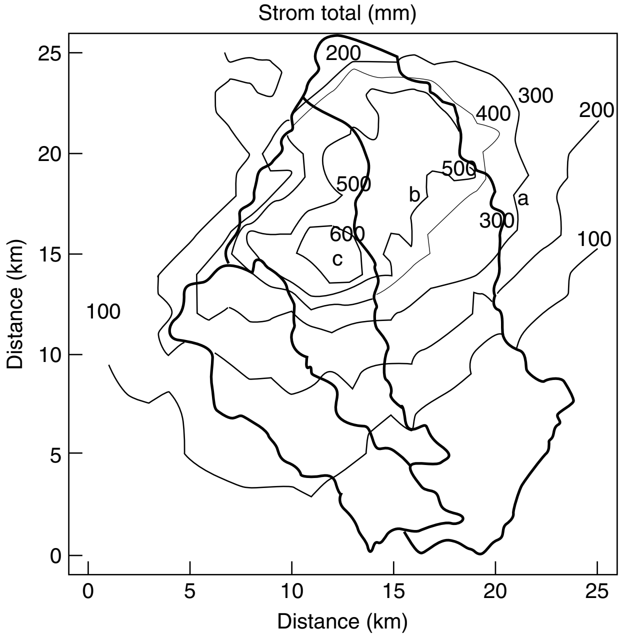

Figure 5. Storm total rainfall (mm) from the Sterling, Virginia, WSR-88D for the rapidan storm of June 27, 1995. The basin boundary for the 295-km^2 watershed and boundaries for 3 subwatersheds are shown in solid lines.

climatology has been developed for mesoscale convective complexes (55) based on IR-based assessments of cloud properties. Numerous studies have examined the links between mesoscale convective complexes (MCCs), and the more general category of mesoscale convective systems, and heavy rainfall [see Houze (56)].

BIBLIOGRAPHY

1. Sumner, G. (1988). Precipitation measurement and observation, in *Precipitation: Process and Analysis*. Wiley, Chichester Chapter 7.
2. Robinson, A.C., and Rodda, J.C. (1969). Rain, wind and the aerodynamic characteristics of raingauges. *Met. Mat.* **98**: 113–120.
3. Groisman, P.Y., and Legates, D.R. (1994). The accuracy of United States precipitation data. *Bull. Am. Meteor. Soc.* **75**(2): 215–227.
4. Sevruk, B. *Methods of correction for systematic error in point precipitation measurement for operational use*, WMO Publication No. 589, OHR No. 21, 1982.
5. Sevruk, B. *Wind-induced measurement error for high intensity rains*, Proceedings, WMO/IAHS International Workshop on Precipitation Measurements, St. Moritz, Switzerland, 1989, pp. 199–204.
6. Nystuen, J.A., Proni, J.R., Black, P.G., and Wilkerson, J.C. (1996). A comparison of automatic rain gauges. *J. Atmos. Oceanic Technol.* **13**: 62–73.
7. Steiner, M., Smith, J.A., Burges, S.J., Alonso, C.V., and R.W. Darden (1999). Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Water Resour. Res.* **35**(8): 2487–2503.
8. McCollum, J.R., and Krajewski, W.F. (1998). Uncertainty of monthly rainfall estimates from rain gauges in the Global Precipitation Climatology Project. *Water Resour. Res.* **34**(10): 2647–2654.
9. Larson, L. and Peck, E. (1974). Accuracy of precipitation measurements for hydrologic modeling. *Water Resour. Res.* **10**(4): 857–863.
10. Frei, C. and Schaer, C. (1998). A precipitation climatology of the Alps from high-resolution raingauge observations. *Int. J. Climatol.* **18**(8): 873–900.
11. Legates, D.R. (1987). A climatology of global precipitation. *Climatology* **40**(1), .
12. Legates, D.R. and Willmott, C.J. (1990). Mean Seasonal and Spatial Variability in Gauge-Corrected, Global Precipitation. *Int. J. Climatol.* **10**(2): 111–127.
13. Urbonas, B.R., and Roesner, L.A. (1992). Hydrologic design for urban drainage and flood control. In: *Handbook of Hydrology*. D. Maidment (Ed.). McGraw-Hill, New York.
14. Frederick, R.H., Myers, V.A., and Auciello, E.P. *Five-to 60-minute Precipitation Frequency for the Eastern and Central United States*. NOAA Technical Memo, NWS Hydro-35, June 1977.
15. Hansen, M. (1987). Probable maximum precipitation for design floods in the United States. *J. Hydrol.* **96**: 267–278.
16. World Meteorological Organization (WMO) (1986). *Manual for estimation of Probable Maximum Precipitation*, 2nd Edn. WMO-No. 332, WMO, Geneva, Switzerland.
17. National Research Council. (1994). *Estimating Bounds on Extreme Precipitation Events*. National Academy Press, Washington, DC.
18. Klazura, G.E., and Imy, D.A. (1993). A description of the initial set of analysis products available from the NEXRAD WSR-88D System. *Bull. Am. Meteor. Soc.* **74**: 1293–1311.
19. Hudlow, M.D., Smith, J.A., Walton, M.L., and Shedd, R.C. (1991). NEXRAD—New era in hydrometeorology. In: *Hydrological Applications of Weather Radar*. I. Cluckie and C. Collier (Eds.). Ellis-Norwood, New York, pp. 602–612.

20. Fulton, R.A., Breidenbach, J.P., Seo, D.-J., and Miller, D.A. (1998). The WSR-88D rainfall algorithm. *Weather Forecast.* **13**(2): 377–395.
21. Smith, J.A., Seo, D.-J., Baeck, M.L., and Hudlow, M.D. (1996). An intercomparison study of NEXRAD precipitation estimates. *Water Resour. Res.* **32**(7): 2035–2045.
22. Baeck, M.L., and Smith, J.A. (1998). Rainfall estimation by the WSR-88D for heavy rainfall events. *Weather Forecast.* **13**: 413–436.
23. Joss, J. and Waldvogel, A. (1989). Precipitation measurement and hydrology: A review. In: *Battan Memorial and Radar Conference*. David Atlas (Ed.). American Meteorological Society, Boston.
24. Wilson, J.W., and Brandes, E.A. (1979). Radar measurement of rainfall—a summary. *Bull. Am. Meteor. Soc.* **60**: 1048–1058.
25. Anagnostou, E., and Krajewski, W.F. (1998). Calibration of the WSR-88D precipitation processing subsystem. *Weather Forecast.* **13**(2): 396–406.
26. Krajewski, W.F. (1987). Co-kriging of radar and rain gage data. *J. Geophys. Res.* **92**(D8): 9571–9580.
27. Seo, D.-J. (1998). Optimal estimation of rainfall fields using radar and rain gage data. *J. Hydrol.* **208**(1,2), pp. 25–36.
28. Seo, D.-J. (1998). Real-time estimation of rainfall fields using rain gage data under fractional coverage conditions. *J. Hydrol.* **208**(1,2), pp. 37–52.
29. Seo, D.J., Fulton, R., and Breidenbach, J. (1998). *Rainfall estimation in the WSR-88D era for operational hydrologic forecasting in the National Weather Service, Symposium on Hydrology, Phoenix*. American Meteorological Society, pp. J60–J62.
30. Baeck, M.L., and Smith, J.A. (1995). Climatological analysis of manually digitized radar data for the United States. *Water Resour. Res.* **31**(12): 3033–3049.
31. Zrnic, D.S. (1996). Weather radar polarimetry—trends toward operational applications. *Bull. Am. Meteorol. Soc.* **77**(7): 1529–1534.
32. Zrnic, D.S., and Ryzhkov, A.V. (1996). Advantages of rain measurements using specific differential phase. *J. Atmos. Ocean. Tech.* **13**(2): 454–464.
33. Ryzhkov, A.V., and Zrnic, D.S. (1996). Assessment of rainfall measurement that uses specific differential phase. *J. Appl. Meteorol.* **35**(11): 2080–2090.
34. Aydin, K., Bringi, V.N., and Liu, L. (1995). Rain rate estimation in the presence of hail using S-band specific differential phase and other radar parameters. *J. Appl. Meteor.* **34**: 404–410.
35. Petersen, W.A., Carey, L.D., Rutledge, S.A., Knievel, J.C., N.J. Doesken, Johnson, R.H., McKee, T.B., Vonder Haar, T., and Weaver, J.F. (1999). Mesoscale and radar observations of the Fort Collins flash flood of 28 July 1997. *Bull. Am. Meteor. Soc.* **80**(2): 191–216.
36. Chappell, C. (1989). Quasistationary convective events. In: *Mesoscale Meteorology*. P. Ray (Ed.). American Meteorological Society, Boston.
37. Doswell III, C.A., Brooks, H.E., and Maddox, R.A. (1996). Flash flood forecasting: An ingredients-based methodology. *Weather Forecast.* **11**(4): 560–581.
38. Maddox, R., Chappell, C., and Hoxit, L. (1979). Synoptic and meso-α scale aspects of flash flood events. *Bull. Am. Meteor. Soc.* **60**(2): 115–123.
39. Houze, R.A., Smull, B.F., and Dodge, P. (1990). Mesoscale organization of springtime rainstorms in Oklahoma. *Monthly Weather Rev.* **118**: 613–654.
40. Perica, S., and Foufoula-Georgiou, E. (1996). Linkage of scaling and thermodynamic parameters of rainfall: Results from midlatitude mesoscale convective systems. *J. Geophys. Res.* **101**(D3): 7431–7448.
41. Steiner, M., Houze, R.A., Jr., and Yuter, S.E. (1995). Climatological characterization of three-dimensional storm structure from operational radar and raingage data. *J. Appl. Meteor.* **34**: 1978–2007.
42. Smith, J.A., Baeck, M.L., Steiner, M., and Miller, A.J. (1996). Catastrophic rainfall from an upslope thunderstorm in the Central Appalachians: the Rapidan Storm of June 27, 1995. *Water Resour. Res.* **32**(10): 3099–3113.
43. Hack, J. and Goodlett, J. *Geomorphology and Forest Ecology of a Mountain Region in the Central Appalachians*, *U.S. Geological Survey Professional Paper 347*, 1960.
44. Negri, A.J., Adler, R.F., Nelkin, E.J., and Huffman, G.J. (1994). Regional rainfall climatologies derived from special sensor microwave imager (SSM/I) data. *Bull. Am. Meteor. Soc.* **75**: 1165–1182.
45. Adler, R.F., Huffman, G.J., and Keehn, P.R. (1994). Global tropical rain estimates from microwave-adjusted geosynchronous IR data. *Remote Sensing Rev.* **11**: 125–152.
46. Vicente, G.A., and Scofield, R.A. (1997). *Real time rainfall rate estimates derived from the GOES-8/9 satellites for flash flood forecasting, numerical modeling and operational hydrology*, Preprints, 13th Conference on Hydrology, Long Beach, CA, American Meteorological Society, pp. J115–J118.
47. Huffman, G.J., Adler, R.F., Rudolf, B., Schneider, U., and P.R. Keehn (1995). A technique for combining satellite-based estimates raingauge analysis, and NWP model precipitation information into global precipitation estimates. *J. Climate* **8**(5): 1284–1295.
48. Adler, R.F., and Negri, A.J. (1988). A satellite infrared technique to estimate tropical convective and stratiform rainfall. *J. Appl. Meteor.* **27**: 30–51.
49. Krajewski, W.F., Ciach, G.J., and McCollum, J.R. (2000). Initial validation of the global precipitation climatology project monthly rainfall over the United States. *J. Appl. Meteorol.* **39**(7): 1071–1086.
50. Adler, R.F., Huffman, G.J., Bolvin, D.T., Curtis, S., and Nelkin, E.J. (2000). Tropical rainfall distributions determined using TRMM combined with other satellite and rain gauge information. *J. Appl. Meteor.* **39**(12): 2007–2023.
51. Huffman, G.J., Adler, R.F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., and Schneider, U. (1997). The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Am. Meteor. Soc.* **78**(1): 5–20.
52. Huffman, G.J., Adler, R.F., Morrissey, M.M., Bolvin, D.T., Curtis, S., Joyce, R., McGavock, B., and Susskind, J. (2001). Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeor.* **2**(1): 36–50.
53. Kummerow, C. (2000). The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. Appl. Meteor.* **39**(12): 1965–1982.
54. Simpson, J., Adler, R.F., and North, G. (1988). A proposed tropical rainfall measuring mission (TRMM). *Bull. Am. Meteor. Soc.* **69**: 278–295.
55. Maddox, R.A. (1980). Mesoscale convective complexes. *Bull. Am. Meteor. Soc.* **61**(11): 1374–1387.
56. Houze, R.A.(1993). *Cloud Dynamics*. Academic, New York.

READING LIST

Barros, A.P. and Lettenmaier, D.P. (1994). Dynamic modeling of orographically induced precipitation. *Rev. Geophys.* **32**(3): 265–284.

Seo, D.J., Breidenbach, J., and Johnson, E. (1999). Real-time estimation of mean field bias in radar rainfall data. *J. Hydrol.* **223**(3,4): 131–147.

RAINFALL AND RUNOFF

Theodore A. Endreny
SUNY-ESF
Syracuse, New York

Rainfall is generated by multiple weather patterns, summarized into frontal, orographic, and convective, that affect the duration and intensity of storm accumulation. Climate data analyses reveal regional and local trends in the seasonal and annual rainfall patterns of duration and intensity and have been reported in map and tabular form for durations of 5 minutes to 24 hours and a variety of return periods in Technical Papers 40 and 49 by the U.S. Weather Bureau and in HYDRO-35 by the National Weather Service. A large variation in rainfall spatial and temporal duration and intensity has been observed for large and small area, long and short time events, and a combined system of drop-size spectrometers, recording rain gauges, and microwave radar have helped to capture this detail nationally at the resolution of 5 minutely volume scans for 1 km^2 bins, aggregated to rain-gauge-corrected hourly totals at 4 km^2 bins. Internal warm and cold cloud precipitation dynamics, low-level evaporation of falling precipitation, and distinction between rain, sleet, freezing rain, and snow still provide technological challenges. Time series plots of rainfall are referred to as hyetographs, maps of equal rainfall amounts are known as isopleths, and National Weather Service NEXRAD Doppler radar gauge-corrected hourly 4 km^2 grid based rainfall maps are known as Stage III products. The spatial aggregation of rainfall point estimates from rain gauges is performed using surface fitting, weighting, and interpolation schemes, including the method of Thiessen polygons.

RAINFALL OBSERVATION AND ESTIMATION

Estimates of expected precipitation intensities, as a function of duration and return period (frequency), are referred to as intensity-depth-frequency (IDF) curves (see Fig. 1), and are computed using climatic precipitation data or regional coefficients, as found in such methods as HYDRO-35. Engineering analyses of the failure of various structures, such as drainage culverts and retention basins, are examined using the IDF curve for storms of regulated frequency and duration. The time of concentration, or the time for water to travel from its hydraulically most distant watershed point to the outlet, is used to design peak runoff concentration. Because the rainfall intensity diminishes with duration, the watershed will generate the greatest instantaneous volume of runoff for the time of concentration. Probable maximum precipitation (PMP) rates represent known and estimated regional maximum storm volumes for a given duration and are required inputs in the human safety analysis of critical infrastructure, such as nuclear power plants and large reservoirs and dams where significant loss of life is possible in cases of failure.

The temporal persistence or absence of sequential rainfall affects the earth's surface antecedent moisture and response to the current rainfall, as do surface slopes, ground cover, and roughness. Surface geologic, geographic, and biological features and controls on rainfall partition percentages and timing are relatively steady or predictable across time; surface antecedent moisture is ephemeral and has greater variability in partitioning between surface storage, groundwater recharge, and surface and subsurface runoff to rivers, lakes, and oceans. Antecedent moisture is categorized for the growing and dormant seasons, and in a Natural Resources Conservation Service (NRCS) method, antecedent moisture is a function of the previous 5-day precipitation totals. At the annual time step, however, soil moisture is relatively constant, absent periods of drought. As such, the fact that perennial, intermittent, and ephemeral streams drain most coastal areas indicates that rainfall exceeds evaporation and that runoff is a unique land feature. The atmosphere precipitates nearly 116,000 km^3 yr^{-1} over land surfaces, but only 44,700 km^3 yr^{-1} drains to the oceans in rivers, and only 2,200 km^3 yr^{-1} drains in groundwater. The difference is evaporated mostly from the land surface.

RAINFALL PARTITIONING AT THE LAND SURFACE

Landscape rather than plot analysis is the preferred method for predicting the partitioning of rainfall into runoff, which can travel as surface, unsaturated subsurface (through flow and interflow), and saturated subsurface flow toward the receiving water. Several theories on the partitioning and pathways of rainfall runoff exist, some of which are illustrated in Fig. 2, which isolate the interaction of rainfall and soil saturation conditions. In general, Hortonian overland flow expects infiltration to be exceeded uniformly in space and time and predicts an entire hillslope of overland runoff. This theory does not explain observed conditions in most areas, but it works reasonably well in many dry western U.S. states. Betson flow expects infiltration to have some spatial variation where it is exceeded by rainfall, and was derived from mathematical analysis and not located in the hillslope. In separate field studies by Hewlett and Dunne, this phenomenon of a dynamic spatial and temporal zone of runoff was observed and called the concept of variable source area (VSA). Whether this runoff originates as the result of rainfall rates exceeding infiltration, known as infiltration excess overland flow, or as the result of saturation from below, preventing infiltration, known as saturation excess overland flow, remains difficult to distinguish. In contrast, assuming that the system is a gaining stream where the terrestrial water table is above the streambed, then base flow is relatively constant for a storm runoff.

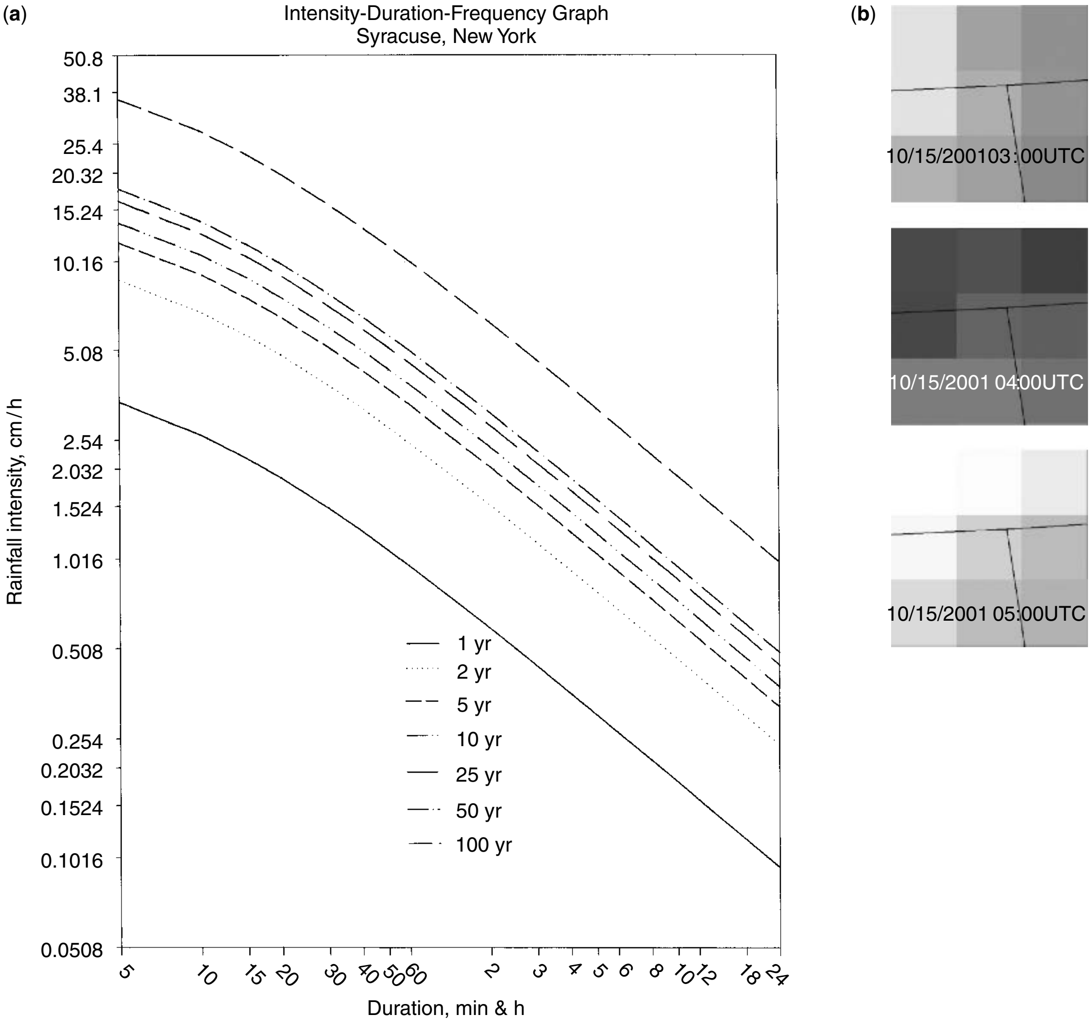

Figure 1. (**a**) The intensity–duration–frequency (IDF) curve for Syracuse, NY, where intensity decreases with duration, but increases with return period, or frequency; (**b**) Illustrations of space and time varying rainfall intensity, represented by 4km x 4km NEXRAD Stage III precipitation grids, where white indicates 0 mm/hour and dark gray indicates 14.5 mm/hour. The total area of 3 x 3 grids is 144 km^2, located at 45 degrees Latitude North by 74 Degrees Longitude West. The dark lines connect at the point where Franklin and Clinton County, New York border Ontario, Canada.

The distinction between infiltration excess and saturation excess has been examined theoretically, comparing observed precipitation rates, infiltration capacity rates, and hydraulic conductivity values, longer period (hourly) averaged precipitation rates seldom exceed infiltration capacity. Shorter period rates (5 minutely) of intense convective storms often do exceed soil infiltration capacity. In arid climates and agricultural areas where soils crust, infiltration is significantly retarded, and precipitation excess often generates runoff. In most other pervious areas, long-duration surface runoff is typically generated by saturation from below, where the local or perched water table has reached the surface. Observations that distinguish between these two events have been advanced using increased spatial and temporal remote sensing monitoring technologies of soil moisture and rainfall dynamics. As watershed rainfall runoff studies increasingly consider urban and suburban dynamics, a greater percentage of the watershed is paved, making the distinction of a runoff-generating mechanism a moot point.

RUNOFF OBSERVATION AND ESTIMATION

Runoff generated from rainfall can be expressed as a volume, a volume per time, as a depth per watershed area, or as a depth per watershed area per time. The ratio of the total runoff and total observed rainfall per time period gives the runoff ratio, which may be presented for climatological or storm periods and has an average U.S. value of 38%. Table 1 provides runoff ratio values for several global regions, and Table 2 provides values for several river basins, which range from 65 to 20%. It is apparent that not all rainfall is converted into storm runoff, and this was first quantified in the 1670s on the upper Seine River in France, but that significant percentages go toward recharging soils and groundwater supplies. The amount of rainfall that is converted into runoff is labeled the effective precipitation, which is an important distinction from observed (e.g., rain gauge measured) precipitation. The depth difference between observed and effective rainfall is explained by the filling by rainfall of a series of storages, including canopy interception, depression storage, detention storage, and soil moisture storage. Areas where observed precipitation is equal to effective, or in stream observed runoff, include rain on water and rain on impervious surfaces (e.g., pavement, rooftops, frozen soil) that are directly connected to the point of measured discharge.

Runoff in channels, or discharge, is measurable using a variety of techniques, including monitoring dilution of injected tracers; monitoring stage height per time for discharge calibrated control structures; monitoring total volumes of discharge per time; and monitoring depth integrated velocities with physical floats, impellers, and Doppler signals together with cross-sectional areas. A runoff time series follows a traditional trend of rising and falling discharge per time, known as a hydrograph. The hydrograph is often separated to distinguish different flow contributions, first delineating a steady volume of base flow or groundwater inflow mostly unaffected by storm dynamics and second, delineating the overlying pulse of event flow. The event flow is potentially separated into separate types of response by using tracers and mixing models, such as slow and fast from the through flow, interflow, pressure wave release, macropore or pipe flow, surface flow, and in-stream rainfall.

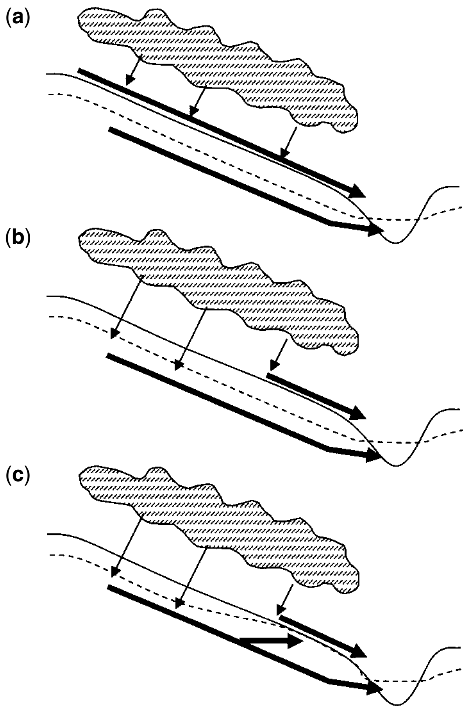

Figure 2. In all figures, the thin solid line is the hillslope surface, the dashed line is the watertable and stream surface, the hashed cloud is rain, the thin arrows are rainfall with or without infiltration, and the thick arrows are surface and subsurface runoff. (**a**) Hortonian concept for overland runoff; the entire hillslope contributes, and base flow is below the water table; (**b**) Betson concept of a with partial hillslope area contributions, but not derived from observations of variable areas; and (**c**) Dunne and Hewlett concept of a variable source area where the water table rises to the ground surface and generates overland runoff.

ENGINEERING VERSUS SCIENTIFIC METHODS

Discharge as total storm runoff, peak storm runoff, or storm time series runoff is estimated by using an equally wide array of methods, and each uses a specific method that first distinguishes the effective precipitation. The most common engineering approach to predicting storm total runoff is the curve number, or SCS, method that partitions observed rainfall into effective rainfall for 6 to 24 hour storms based on antecedent moisture, land cover type, management condition, and hydraulic soil type, which ranges from high infiltration A soils to poor infiltration D soils. The most common engineering approaches to predicting storm peak runoff are either by the curve number based TR-55 method or the runoff coefficient based rational method, which uses a land cover type runoff coefficient to convert observed rainfall intensity into an effective intensity. The most common engineering approach to predicting time series runoff is the unit hydrograph, which requires a user input of the effective rainfall across time, provides a transfer function known as a convolution integral technique to output volume per time discharge, and represents the observed rising and falling limbs of incrementally changing watershed contribution. Unit hydrographs are created for unit depth rainfall (e.g., 1 cm) of distinct duration and based on the principle that runoff response is linear and additive, are weighted by greater or lesser effective precipitation amounts. Empirical, synthetic, and geomorphic techniques exist to create unit hydrographs.

Alternative analytic approaches exist to partition observed rainfall into estimated runoff and, through validation studies on several watersheds, are increasingly demonstrated as accurate. These approaches are based on solving a series of equations for estimating interception volumes, filling detention, depression, and retention storages, partitioning soil infiltration volumes, and possibly accounting for evaporation rates, where initial estimates of storages are either input or acquired from model spin ups where evaporation is a controlling process. As the watershed size increases, effective values for observed precipitation and parameters for soil and plant processes present greater model input challenges. Distinctions between these mechanistic models include their representation of the watershed in space, as either distributed or lumped, and their simulation through time, as either storm or continuous. New modeling techniques are experimenting with parameter optimization through genetic algorithms that mutate and select the most robust parameter strings, and neural networks that seek optimal, but nonphysical, solutions that have best met the objective functions, which typically minimize differences between observed and predicted runoff. Popular rainfall runoff models include those used for operational forecasting

Table 1. Continental Average Estimated Runoff

Continent	Area km^2	Runoff mm yr^{-1}	Runoff Ratio %
Antarctica	14,100,000	28	83
Europe	10,000,000	375	43
Asia	44,100,000	420	40
South America	17,900,000	946	40
North America	24,100,000	403	38
Africa	29,800,000	582	16
Australia	7,600,000	420	6
Total Land	148,900,000	480	36

Table 2. Average Estimated Runoff Ratio for Major River Basins

Watershed by River Name	Continent [Nation(s)]	Basin Size km^2	Runoff Ratio %
Brahmaputra	Asia (Tibet/Bangladesh)	589,000	65
Irrawaddy	Asia (Burma)	431,000	60
Yangtzekiang	Asia (China)	1,970,000	50
Amazon	South America (six nations)	7,180,000	47
Orinoco	South America (Venezuela)	1,086,000	46
Lena	Asia (Russia)	2,430,000	46
Mekong	Asia (China)	795,000	43
Yenesei	Asia (Russia)	2,599,000	42
Ganges	Asia (China)	1,073,000	42
Saint Lawrence	North America (Canada, U.S.)	1,030,000	33
Amur	Asia (Russia)	1,843,000	32
Congo	Africa (seven nations)	3,822,000	25
Ob	Asia (Russia)	2,950,000	24
Mississippi	North America (U.S.)	3,224,000	21
La Plata	South America (five nations)	2,650,000	20
Average		2,224,800	40

by the National Weather Service, such as the SAC-SMR (Sacramento Soil Moisture Accounting) or on those capturing governing topographic controls on generating runoff, such as TOPMODEL. A partial list of agency and research based surface water models is provided by the USGS Surface Water Modeling Clearinghouse website.

RUNOFF ROUTING

Explicit routing of runoff across the landscape, through a retention basin, and along a section of channel upstream of the gauging station may provide additional insighting time-based runoff estimates. In each case, routing has been examined using fully dynamic equations of energy and continuity, written in their complete form (e.g., St. Venant) or neglecting acceleration and pressure gradients (e.g., kinematic), solved using numerical methods of finite element and finite difference. Simpler solutions that still consider the effects of friction, such as the Darcy–Weisbach and Manning equations have been used to route flow across land, as have stand-alone continuity equations that simply push water from cell to cell at a given slope-based velocity. The algorithms used to partition flow in grid-based networks of the watershed terrain map (e.g., digital elevation map files with draped land cover) are varied, ranging from a single path to flow into a maximum of all eight neighboring lower cells. Routing across triangular irregular network grids and polygons is also used. Flow through a reservoir, known as level-pool routing, and along a channel (e.g., the Muskingum method), attenuates the inflow hydrograph based on increasing storage surface area and stage-discharge relationships and wave celerity and propagation rates, respectively. Both methods are presented in most hydraulic engineering handbooks.

SPECIAL TOPICS: WATER YEAR, RUNOFF QUANTITY AND RUNOFF QUALITY

Annual time series analysis of precipitation hyetographs and runoff hydrographs reveals trends that give rise to regional characterization of U.S. hydrologic zones. The hyetograph analysis considers rain and liquid equivalent snowfall. West of the 100th meridian, rainfall is typically less than 20 inches annually. California has increased precipitation during the winter when a Pacific blocking high moves northward; southwestern states have summer monsoon rains; central, midwestern, and southeastern states have increased precipitation during summer convective showers, and eastern states have rather steady annual precipitation. Runoff often follows separate trends, given the land surface seasonal function in partitioning rainfall. Snow accumulation in the winter often keeps winter runoff low; snowmelt in the spring and higher altitudes in the summer release a pulse of the largest stream flow volumes. Spring and summer vegetation growth creates soil moisture deficits that allow for larger soil storage of rainfall and reduced stream flow. The fall dieback of vegetation reduces evaporation rates, and the soil storage fills and remains filled, which increases stream flow. The annual rainfall–runoff cycle that leads to a renewal of soil moisture reserves and increased runoff has led to the creation of the water year, beginning on October 1. Anomalies to this pattern exist, and events such as warm rain on snow, which melts the pack in midwinter but does not infiltrate into the frozen soils, can create huge winter floods. Likewise, the global ocean and atmospheric patterns set in place by systems such as the El Niño southern oscillation can cause southeastern spring and summer droughts when there are usually heavy rains.

Special topics in runoff hydrology include flood and drought analysis as well as water quality, which are useful in human safety, infrastructure design, and environmental management. The Water Resources Council 1972 method continues to provide guidance on fitting annual maximum discharge observations onto a Type III Pearson distribution to forecast return interval floods that exceed observed record lengths. Although 10, 50, and 100 year events have significant discharge rates and therefore erosive power, channel patterns are shaped by flows of the more frequent 1.5 return intervals. Drought analysis is critical to sizing drinking water reservoirs and wastewater treatment discharge dilution plans and are typically based on 7-day, 10-year low flows, or for ungauged rivers, on regional relationships derived from similar streams that were gauged. Water quality has become more directly tied to rainfall runoff analysis after the successful control of many municipal and industrial point source pollution dischargers, leaving the nonpoint source (NPS) pollution from agricultural fields, commercial parking lots, urban streets, and residential lawns. NPS runoff from these sources, it has been observed, explains the large fraction of soil, nutrient, chemical, metal, organic, and bacterial loads entering the 40% of national waterways that are still noncompliant with Clean Water Act requirements for fishing and swimming access.

READING LIST

Dingman, S.L. (1994). *Physical Hydrology*. Prentice-Hall, Englewood Cliffs, NJ.

Mays, L.W. (2001). *Water Resources Engineering*. John Wiley & Sons, New York.

REMOTE SENSING OF APPLICATIONS IN HYDROLOGY

Jessica Black
NOAA/CREST CCNY
New York, New York

Theodore A. Endreny
SUNY-ESF
Syracuse, New York

REMOTE SENSING DEFINED

Remote sensing is the science of observing and analyzing an area, object, or phenomenon based on data collected by sensors physically separated from the target. The sensors are mounted on platforms such as ground-based, airplane, or satellite platforms. Physical radiation transfer is the foundation for remote sensing. Typically, the observation is of radiation reflected or emitted from the object of interest. There are two types of remote sensing: active and passive. In active remote sensing, the object is illuminated by radiation of a known wavelength (or frequency); the backscatter from the object is recorded by the sensor. In passive remote sensing, the sensor records radiation generated from natural sources, such as the sun, or radiation emitted from the earth or atmosphere. Radiation that the sensor receives has a spectral response that reveals properties of the atmosphere and object of interest. By extending remote sensing applications to improve the spatial and temporal resolution of certain key water resource variables, scientists can better understand and quantify water resource processes.

METHODOLOGIES

Qualitative and quantitative approaches for analyzing remotely sensed data provide discrete and continuous information about an area. The qualitative approach separates observed spectral data into different classes and generates map-based images showing the spatial distribution of the classes. For example, if trying to map vegetation in an area that has certain known vegetation classes, the qualitative approach produces a map where every pixel is assigned a vegetation class or a ‘no class.’

The quantitative approach provides numerical information about the area or phenomenon of interest. For example, by computing the normalized difference vegetation index (NDVI) using visible and near-infrared data and generating a map where a value is attached to every pixel, a map user has an idea how dense the vegetation is and where the denser areas lie.

For either of these approaches, several different classification methods can be used. All classification methods fall under two general categories: supervised classification and unsupervised classification. Supervised classification requires the scientist to collect data from the area of interest and assign these data to pixels within the image. Examples of supervised classification methods are the parallelepiped method and the minimum distance method. Examples of unsupervised classification are the chain method and the ISODATA method.

RESOLUTION AND SPECTRAL REGIONS

To determine the type of data required for a remote sensing application, scientists need to identify and quantify the key data characteristics of the phenomenon of interest. These characteristics may include the temporal resolution, spatial resolution, radiometric resolution, and spectral resolution. All four characteristics determine the platform, sensor, and capital necessary for the project.

The spatial resolution determines the minimal size of objects that can be seen from the sensor; the smaller the area covered, the higher the resolution. The temporal resolution is the time between data acquisitions across the

Table 1. Resolution Types and Potential Applications

Resolution Type	Resolution	Applications	Sensor Examples
Spatial	10–100 m	Local hydrological processes	ASAR; MSS
Spatial	1–5 km	Local meteorologic processes	AVHRR/2; MODIS; SSM/I; ERBE
Spatial	>25 km	Basin-wide studies; global studies	
Temporal	Every 30 min	Meteorologic studies	VHRR; VISSR
Temporal	2/day	Meteorologic studies	AVHRR
Temporal	Every 2 weeks	Land use	ETM+
Temporal	Once	Maps	

same area. The radiometric resolution is the sensitivity of a sensor to differences in signal strength as it records the incoming radiant flux. The radiometric resolution is often determined by the number of bits (6 bits, 8 bits, etc.), available for storing the data. The spectral resolution is the range of wavelengths detected by the selected sensor (1). Different wavelengths or frequencies can be used for different applications. For examples of different uses for different resolutions and spectral regions, see Tables 1 and 2.

Almost all spectral regions have been used for passive remote sensing, ranging from the visible (VIS), infrared (IR), and microwave (MW) regions (Table 3); however, active remote sensing has been limited to microwave regions. Each of these regions breaks down to smaller regions; some of these regions are more commonly used than others. For example, remote sensors will capture red, blue, and green within the visible region more often than purple/violet or yellow because of the way certain physical environments reflect these particular wavelengths. Within the IR region, near-infrared (NIR) and thermal-infrared (TIR) radiation react to or are emitted from certain components of the physical environment more than other wavelengths in that region. The bands of microwave radiation wavelengths and frequencies are denoted by the letters K-, W-, X-, C-, S-, and L-.

Table 2. Spectral Bands, Bandwidths, and Their Potential Applications[a]

Spectral Band	Bandwidth	Applications	Sensor Examples
Blue	0.45–0.50 μm	Water penetration, land use, vegetation characteristics, sediment	ETM+; ASTER; MISR; TM
Green	0.50–0.60 μm	Green reflectance of healthy vegetation	AATSR; ASTER; MISR; TM
Red	0.60–0.70 μm	Vegetation discrimination because of red chlorophyll absorption	AATSR; ASTER; AVHRR/2; MISR
Panchromatic	0.50–0.75 μm	Mapping, land use, stereo pairs	ASTER;
Near-infrared	0.75–0.90 μm	Biomass, crop identification, soil crop, land–water boundaries	AATSR; ASTER; AVHRR/2
Midinfrared	1.5–1.75 μm	Plant turgidity, droughts, clouds, snow-ice discrimination	AATSR
Thermal infrared	8–12.5 μm	Relative temperature, thermal discharges, vegetation classification, moisture studies, thermal inertia	AATSR; AVHRR/2
Microwave: Shortwave	0.1–5 cm	Snow cover, depth, vegetation water content	ALT Duel-frequency Radar Altimeter; MWR
Microwave: Long wave	5–24 cm	Melting snow, soil moisture, water–land boundaries, penetrate vegetation	SAR; AMI-SAR

[a] Compiled based on Reference 2, p. 11, Table 1.1.

Table 3. Wavelengths and Frequencies for VIS, IR, and MW

Region	Wavelength (λ)	Frequency (v)
Visible: R,B,G	0.4–0.7 (μm)	750,000–428,571 GHz
Infrared: NIR, TIR	0.7–100 (μm)	428,571–3000 GHz
Microwave shortwave: K-, W-band	0.1–5 (cm)	300–6 GHz
Microwave long wave: X-, C-, S-, L-band	5–24 (cm)	6–1.25 GHz

Basic Terms and Equations

Basic terms and equations fundamental to the theory of remote sensing applications in water resources are given below. Wavelength, λ (m), is the length of the cycle of vibration, given as

$$\lambda = \frac{c}{v} \tag{1}$$

where c is the speed of light, approximated as 3×10^8 m/s, and v is the frequency of vibration (1/s). Note that the speed of light is adjusted for density differences between space and the earth's atmosphere. Calculation of brightness temperature (3), T_B, is based on the measured excitance, or vibration, from a narrow spectral band, as a function of emissivity, ε, and temperature, T:

$$T_{\mathrm{B}}(\lambda) = \varepsilon(\lambda)T \tag{2}$$

The terms backscatter (3), emission (3) and reflectance (3) refer to different types of energy received by a spaceborne sensor from the targeted area. Backscatter is the reflectance of microwave radiation from the target, defined by

$$\frac{P_{\mathrm{r}}}{P_{\mathrm{r}}} = \frac{\lambda^2 G_{\mathrm{t}} G_{\mathrm{r}}}{(4\pi)^3 R_{\mathrm{t}}^2 R_{\mathrm{r}}^2}\sigma \tag{3}$$

where P_{r} is the received power, P_{t} is the transmitted power, λ is the wavelength of the radar beam, G is the antenna gain, R is the distance (range) between target and antenna, and σ is the radar scattering cross section. Reflectance occurs when radiation bounces off the target and is redirected. Specular reflectance occurs when ~100% is redirected in the same direction, and diffuse reflectance occurs when 0% is redirected in the same direction. Reflectance is defined by

$$r_{\mathrm{s},\lambda} = \frac{\pi L_{\mathrm{s},\lambda}}{(E_{\mathrm{dir},\lambda} + E_{\mathrm{dif},\lambda})} \tag{4}$$

where $r_{\mathrm{s},\lambda}$ is the spectral surface reflectance, $L_{\mathrm{s},\lambda}$ is the reflected radiance at the surface, $E_{\mathrm{dir},\lambda}$ is the direct irradiance, and $E_{\mathrm{dif},\lambda}$ is the diffuse irradiance. Excitance (emission) is the radiation that comes from the earth's surface, such as heat, and is defined by

$$L_{\infty,\lambda} = L_{0,\lambda} t_\lambda(0,\infty) + \int_{z=0}^{z=\infty} L_{\mathrm{B},\lambda}[T(x)] \frac{\partial t_\lambda(z,\infty)}{\partial z} dz \tag{5}$$

where t_λ is the spectral transmissivity, $L_{\mathrm{B},\lambda}$ is the spectral blackbody radiance, $L_{0,\lambda}$ is the surface emitted radiance, and $L_{\infty,\lambda}$ is the upwelling radiance.

Water resource remote sensing applications covered in this article include evapotranspiration, flood mapping, precipitation, soil moisture, surface snow, waterbodies, and wetland detection. Each topic is discussed in detail, and basic algorithms are provided where possible. These algorithms, adapted from different situations, may have limitations, when applied to new situations. These limitations, noted and explained in the article, should be considered when applying the equations to different environments.

EVAPOTRANSPIRATION

Evaporation is the physical change of liquid water to a vapor. Transpiration is the evaporation of water from plant tissues into the atmosphere. Evapotranspiration (ET) is the combination of these two processes. Remote sensing of evapotranspiration remains an imperfect science in part because scientists are still largely unable to measure evapotranspiration directly. Scientists must rely on creating evapotranspiration algorithms and defining the variables within the algorithm using remote sensing. Even with these algorithms, scientists still require ground data for variables not determined through remote sensing methods (4).

The two spectral regions utilized for ET are VIS and IR (4). Scientists can analyze two types of ET, actual ET and potential ET, using different methods. Actual ET, the ET that occurs in any real situation, becomes more difficult to quantify in remote sensing, compared to potential ET(PET), the ET that occurs in when the soil/vegetated area has 100% soil moisture.

Currently, many algorithms exist to calculate ET. For this article, two equations were selected that calculate some idealized evaporation rates (potential and reference crop), and then different methods are explained for calculating actual ET. The first equation, developed by Priestley and Taylor (5), calculates the reference crop evaporation assuming that approximately 30% of the evaporation is driven by turbulent transfer. This equation depends upon remotely sensed thermal properties.

$$E_{\mathrm{rc}} = \alpha \frac{\Delta}{\Delta + \gamma}(R_{\mathrm{n}} - G) \tag{6}$$

where E is the evaporation rate for the reference crop; R_{n} is the net radiation; G is the ground heat flux; Δ is the slope of the vapor pressure versus temperature curve; and α is the empirical evaporation constant, determined as 1.26 or dependent on soil moisture (6,7).

The second equation, derived by Choudhury (8), calculates global potential evaporation. The data for this equation need to be collected from a variety of satellites that measure visible and thermal bands (4):

$$E_0 = \frac{\Delta R_{\mathrm{ni}} + \rho c_{\mathrm{p}} \dfrac{D}{c r_{\mathrm{e}}}}{\Delta + \gamma \dfrac{(r_{\mathrm{s}} + r_{\mathrm{h}})}{r_{\mathrm{e}}}} (\mathrm{mmd}^{-1}) \tag{7}$$

where Δ (hPa K^{-1}) is the differential of saturated vapor pressure with respect to temperature evaluated at air temperature T_{a} (K), γ (hPa K^{-1}) is the thermodynamic psychrometric constant, R_{ni} (mm d^{-1}) is the thermal net radiation, ρ is the density of air, c_{p} is the specific heat of air, D (hPa) is the vapor pressure deficit, r_{s}(s m^{-1}) is the surface resistance, r_{H} (s m^{-1}) is the aerodynamic resistance for heat transfer from a virtual source height where the canopy is at a reference level in the atmosphere, r_{e}(s m^{-1}) is the effective resistance to heat transfer obtained by putting the aerodynamic resistance r_{H} and the resistance to long-wave radiative transfer r_{R} in parallel, and c (mm d^{-1}) is the constant equal to the product of latent heat of evaporation and density of water.

The Choudhury (8) equation, based only on PET from grassy areas, needs to be adjusted before application to other types of land cover (9,10):

$$K_{\mathrm{c}} = \frac{E_{\mathrm{m}}}{E_0} \tag{8}$$

where K_{c} is the crop coefficient; E_{m} is the basic evaporation equation result with changes in albedo, aerodynamic effective resistance, and the minimum surface resistance relevant to the crop/landcover to be studied; and E_0 is the basic evaporation equation result.

To calculate actual evapotranspiration, scientists employ numerical models because of the complexity of the data requirements. One method, the integration of soil-vegetation-atmosphere-transfer (SVAT) models with planetary boundary layer (PBL) models, have the advantage of calculating ET from a higher altitude, thereby reducing surface spatial heterogeneity (4). Some problems to consider when tackling the calculation of actual evapotranspiration are the difficulty of validation and the inability to derive air humidity from remotely sensed data (4).

FLOOD MAPPING

Flood mapping is the process of delineating areas that may be flooded or have been flooded. Remote sensing flood mapping deals mostly with areas that have been flooded. The spectral regions used for this type of mapping are NIR (only for cloud-free conditions) and MW. NIR produces maps where the waterbodies are dark in contrast to surrounding soil and vegetation. Barber et al. (11) determined that C-HH data (active microwave data) also produced flood delineation maps collecting specular reflection, as long as the incident angles were greater than 45°. Active microwave sensors have an advantage over NIR sensors for flood mapping because they can penetrate clouds (12).

PRECIPITATION: RAINFALL AND SNOWFALL

The detection and estimation of precipitation can be either ground-based or space-based. All three spectral regions are used, VIS, IR, and MW, and each spectral region requires a different approach to calculating precipitation. Common instruments for ground-based rainfall and snowfall estimation are weather radars. These instruments take vertical measurements of snow/rainstorms and receive the backscattering from the individual drops or flakes. Spaceborne instruments include infrared and visible sensors (13).

For ground-based weather radars, the theoretical basis for rainfall estimation consists of the following equations (13,14):

$$Z = \int_0^{\infty} D^6 N_{\mathrm{V}}(D) dD \tag{9}$$

$$R = 6\pi \times 10^{-4} \int_0^{\infty} D^3 v(D) N_{\mathrm{V}}(D) dD \tag{10}$$

$$v(D) = cD^{\gamma} \tag{11}$$

where Z is the radar reflectivity factor, $N_{\mathrm{V}}(D)\ dD$ is the drop size distribution per volume of air, D is the spherical raindrop diameter, R is the precipitation rate, $v(D)$ represents the functional relationship between the raindrop terminal fall speed in still air and the equivalent spherical raindrop diameter, $c = 3.778$ (if v is expressed in m s^{-1} and D in mm), and γ is 0.67. A limitation of Equation (9) is that raindrops are not round, they are oblate. To take this limitation into account, an adjustment is made in the original Z reflectivity equation where dual polarization is used. A limitation of this method is that single point measurements for the differential reflectivity may not properly represent rainfall at the land surface (13).

The relationship between the radar reflectivity factor, the radar properties, and backscatter is defined by the following equation (15):

$$\overline{P_{\mathrm{r}}} = C\frac{|K|^2}{r^2}Z \tag{12}$$

where $\overline{P_{\mathrm{r}}}$ is the mean power (W) received from raindrops at range r (km), C is the radar constant and $|K|^2$ is a coefficient related to the dielectric constant of water (~ 0.93). The limitations of this equation are that it holds for Rayleigh scattering theory, where targets are small compared to radar wavelength, perfect radar calibration is necessary, and attenuation is absent (14). The relationship between Z and R (sometimes denoted as S for snowfall), the precipitation rate, is the following:

$$Z = AR^B \tag{13}$$

where A and B are coefficients that vary with climatic and location characteristics but are independent of R (Table 4). This equation is theoretically based upon Eq. 14 and empirically based on numerical evidence (15). There are two methods for calculating A and B: regression analysis and parameterization using the raindrop size distribution (Eqs. 9–11) (14).

For snowfall, Z_{e} is considered instead of Z because of the physical differences between snowflakes and water drops (16):

$$Z_{\mathrm{e}} = \frac{|K|_{\mathrm{i}}{}^2}{|K|_{\mathrm{W}}{}^2} \cdot Z \tag{14}$$

where Z_{e} is the radar reflectivity of snow/hail, $|K|_{\mathrm{i}}{}^2$ is the dielectric constant for ice and $|K|_{\mathrm{W}}{}^2$ is the dielectric constant for water. The 'correct' values for $|K|_{\mathrm{i}}{}^2$ depend on the particle size selected (or determined). Some examples of selected particle sizes and their respective dielectrics can be found in Table 5. A typical value for the dielectric constant for water ($|K|_{\mathrm{W}}{}^2$) is 0.93.

Two major physical differences exist between snow and rain that make Eq. 13 less accurate for snowfall. Upwelling of snow due to the snowflake density/size

Table 4. Empirical Constant Values and Respective Precipitation[a]

A	B	Precipitation Type	Reference
140	1.5	Drizzle	(19)
250	1.5	Widespread rain	(19)
200	1.6	Stratiform rain	(20)
31	1.71	Orographic rain	(21)
500	1.5	Thunderstorm rain	(19)
486	1.37	Thunderstorm rain	(22)
2000	2	Aggregate snowflakes	(23)
1780	2.21	Snowflakes	(24)
1050	2	Snowfall (X-Band)	(25)
229	1.09	Snowfall (X-Band)	(26)
427	1.09	Snowfall (X-Band)	(27)

[a] Adapted from Reference 13 pg 113, Table 6.2; $Z(\mathrm{mm}^6\ \mathrm{m}^{-3})$, $R(\mathrm{mm\ h}^{-1})$

Table 5. Snowfall Diameter Types and Respective Dielectric Constants $|K|_i^2$

Diameter Size/Type	Dielectric Constant Value
Melted deep diameters	0.208
Ice sphere diameters	0.176

and the inhomogeneity of snowflake type and density do not follow the basic assumptions required for the Z–R relationship and decrease its overall accuracy (17). However, Collier (13) states that with careful quality control and adjustment of radar measurements, snowfall estimations can be as accurate as rainfall up to 50 km from the radar site.

For infrared satellite measurements, two algorithms commonly used are the cloud indexing algorithm and the life history algorithm. There are also several modeling methods that use artificial neural networking (ANN) in combination with remote sensing data. Cloud indexing is based on the idea of a classification type where the brightness temperature of every pixel is calculated and compared to a threshold brightness temperature. The threshold brightness temperature represents the temperature below which a cloud pixel precipitates. The algorithm is adjusted using empirically derived constants. One of the major limitations of this algorithm is that it is relevant only to cumuliform clouds and precipitating clouds. It is not relevant to cold nonprecipitating clouds like cirrus. However, this adjustment can be overcome by using wavelengths of 10.8 and 12 μm to differentiate between the two types of clouds (18). Another major limitation is calibration and validation due to the lack of spatial and temporal correlation between rain gauges and the algorithm. Rain gauges are point measurements whereas the algorithm measures over space. An alternative to rain gauges for calibration/validation is ground-based radar, but it too has inaccuracies (13):

$$\mathrm{PI} = A_0 + \sum_i A_i\, S_i(\mathrm{TBB}_i) \qquad (15)$$

where $\mathrm{TBB_i}$ is the Pixel brightness temperature, T_0 is the brightness temperature threshold that is determined empirically, A_0 is a constant that is determined empirically, A_i is a constant that is determined empirically, S is the surface area, which is a function of brightness temperature, and PI is the precipitation index.

The life history method takes into account the growth and decay of a particular cloud or storm, and rain rates are adjusted to the cloud's stage of growth. The variables and limitations of this equation are similar to those of the cloud indexing equation (13).

$$\mathrm{PI} = A_0 + A \cdot S(\mathrm{TBB}) + A' \frac{d}{dt} S(\mathrm{TBB}) \qquad (16)$$

Modeling programs that use the IR spectral region, such as precipitation estimation from remotely sensed information using artificial neural networks (PERSIANN), are alternative methods of calculating rainfall and snowfall more precisely than the simple algorithms above.

The use of the passive MW spectral region in precipitation detection and estimation is a hot topic in research now and is producing very favorable results. Microwave estimation of precipitation is favorable because it measures precipitation more directly than IR, where the radiation is absorbed or scattered by the precipitation. The distinction between the two methods of estimation, absorption/emission and scattering, is important because both have limitations. Absorption/emission of microwave radiation by liquid water is relatively high and increases as frequency increases. This increase in emission drowns out signals of light precipitation. Scattering occurs when ice particles are in the lower parts of the cloud before the particles become liquid. However, the existence of ice particles must then be connected to the physical precipitation in some manner (13).

SOIL MOISTURE

Soil moisture is defined as the water retained in the soil after infiltration. The spectral regions used for remote sensing are MW and TIR. The spatial resolution for this work is very important, and is one of the reasons that the microwave region is preferentially used by hydrologists. The use of thermal infrared and passive microwave requires more algorithmic work compared to active microwave. Generally, soil moisture is measured by developing an algorithm that contains variables quantifiable through remote sensing analysis (28). For TIR, the following equations can be used:

$$\mathrm{DT_s} = T_\mathrm{s}(\mathrm{PM}) - T_\mathrm{s}(\mathrm{AM}) = f\left(\frac{1}{D}\right) \qquad (17)$$

$$D = w \cdot Q_\mathrm{c} \cdot k \qquad (18)$$

where $\mathrm{DT_s}$ is the diurnal temperature, the difference between the afternoon surface temperature $T_\mathrm{s}(\mathrm{PM})$ and the early morning surface temperature $T_\mathrm{s}(\mathrm{AM})$; D is the diurnal inertia; w is the day length; Q_c is the volumetric heat capacity, and k is the thermal capacity.

Based on this equation, a scientist would then develop a relationship between the temperature change and the soil moisture. The limitations of this equation are that it can be used only for bare soil. Clouds, surface topography, and local meteorologic conditions can decrease accuracy due to their absorption of thermal radiation. The penetration depth that can be considered is small because the thermal radiation is not penetrating but emanating (28).

Microwave soil moisture equations are based on dielectric constants or the emanation of microwaves from the soil surface. Passive microwave soil moisture calculations require the brightness temperature, ground emissivity, and the dielectric constant. Scientists will then use the final temperature brightness outcome to develop an empirical relationship to known soil moisture values (28,29):

$$T_{\mathrm{B}} = t(H) \cdot \lfloor rT_{\mathrm{sky}} + (1-r)T_{\mathrm{soil}} \rfloor + T_{\mathrm{atm}} \tag{19}$$

where T_{B} is the brightness temperature, $t(H)$ is the atmospheric transmissivity for a radiometer at height H above the soil, r is the smooth surface reflectivity, T_{soil} is the thermometric temperature of the T_{soil}, T_{atm} is the average thermometric temperature of the atmosphere, and T_{sky} is the contribution from the reflected sky brightness.

Active microwave soil moisture measurement has problems similar to passive, requiring an empirical development of a relationship between backscatter and soil moisture. The spatial resolution of active microwave sensors has an advantage over passive and thermal infrared sensors. Both types of microwave remote sensing can also be used for change detection (28). Some limitations to consider for microwave sensing are surface roughness and measurement depth. Surface roughness is spatially and temporally variable, which affects measurements on the watershed scale (12). Measurement depth is important for calculating the correct dielectric constant (28).

SURFACE SNOW

Surface snow is important as a source of freshwater for many areas. Scientists are interested in many characteristics of surface snow that will help them to estimate quantity and time to runoff. The most important characteristics are identification of extent, snow-covered area (SCA), depth, density, and snow water equivalent (SWE). Each of these characteristics requires different spectral data, and each can use different techniques during analysis.

Surface snow extent can be measured using VIS and MW from satellite platforms. The method employed for snow extent is detection. Each spectral region has advantages and limitations. The visible region usually has good resolution (depending upon the type of satellite) but cannot provide data at night and is not good for distinguishing between clouds and snow.

Passive microwaves (SSM/I) are limited to very large areas due to poor resolution (~25 km) and detect only dry snow. Passive microwave data are also hard to validate due to large pixel size. Active microwave data, such as SAR, can discriminate between wet and dry snow with a high resolution but have trouble with forested areas (30).

Surface snow water equivalent can be calculated using microwave and gamma radiation. The gamma radiation is detected by using an airplane platform, and microwaves are detected by using an airplane or satellite platform. The following equation was developed for gamma radiation calculation of SWE. When using gamma radiation to calculate SWE, soil moisture can attenuate the radiation, and this affects the accuracy (20,31).

$$\mathrm{SWE} = \frac{1}{A}\left(\ln\frac{C_0}{C} - \ln\frac{100 + 1.11M}{100 + 1.11M_0}\right) \tag{20}$$

where SWE (g/cm^2) is the snow water equivalent, C and C_0 are uncollided terrestrial gamma counts over snow and bare ground, M and M_0 are the percent soil moisture in snow covered and bare soil areas, and A (cm^2/g) is the radiation alternation coefficient in water.

The SWE based on passive microwave data is a function of brightness temperature and empirical constants. These constants allow the "moving" algorithm from location to location. When calculating SWE, the effect of vegetation on backscattering and emissions must be considered. Forested areas tend to decrease the accuracy of results. A major limitation of passive microwave snow detection and SWE estimation is that passive microwaves sensors cannot detect wet snow. Wet snow increases brightness temperatures much more than dry snow, thereby decreasing the accuracy of the original equation. Even slightly wet snow is hard to detect with passive microwaves.

Active microwaves, such as SAR, can detect wet snow but cannot distinguish dry snow from no-snow areas. This suggests that multiband methods have a future in SWE determination, but the sensor technology may not be available (20,32).

$$\mathrm{SWE} = A + B\left(\frac{T_{\mathrm{B}}(f_1) - T_{\mathrm{B}}(f_2)}{(f_2 - f_1)}\right) \tag{21}$$

where A is the offset of the regression of the brightness temperature difference, B is the slope of the regression of the brightness temperature difference, f_2 is the high scattering channel, commonly 37 GHz, f_1 is the low scattering channel, commonly 18 or 19 GHz, and SWE (mm) is the snow water equivalent.

WATERBODIES: DETECTION, COLOR, CLARITY, AND QUALITY

Remote sensing can be used to observe several different characteristics of waterbodies such as waterbody detection; waterbody perimeter calculation; and waterbody clarity, color, and quality. Waterbody detection is the process of determining the existence of water bodies within an area, whether they are ephemeral or permanent. The regions used for this type of detection are VIS, IR, and MW. IR is absorbed by water, so waterbodies show up as black areas on false colored IR images, and when using

VIS, radiation is reflected (12). Passive microwave emissions can be used to calculate the brightness temperature, and this allows distinguishing between water and floating ice. Open water has a lower brightness temperature (12).

Examples of satellite sensors and platforms for waterbody detection are SSM/I, Landsat TM, and NOAA AVHRR (LIC). A problem to consider with the spectral regions of VIS and IR is that VIS and IR are constrained by clouds. A general limitation to consider is that the area to be analyzed should exceed the spatial resolution by a factor of 10 (33).

To determine water quality, clarity, or color requires using the same spectral regions, VIS and IR, but with a different method of analysis. Using linear and nonlinear equations based on the radiance, reflectance, or energy of the radiation, scientists relate these to the parameter of interest. There are three approaches for creating this relationship: empirical, semiempirical, and analytical. The empirical methods use measured spectral properties and measured water quality parameters. The semiempirical methods use spectral/optical characteristics to approximate the appropriate wavelength for the water quality parameter. The analytical method uses optical properties of the water quality parameters to develop a model for the spectral quality of the water, and this model is introduced into an equation that relates the remote sensing data, the water quality parameters, and the optical parameters. Some examples of parameters that can be measured are suspended sediments, chlorophylls (algae), DOM (humus), oil, and temperature (Table 6) (34). An example of an empirical/semiempirical equation is

$$Y = A + BX \tag{22}$$

or

$$Y = AB^X \tag{23}$$

where Y is the measured radiance, reflectance, or energy; X is the water equality parameter of interest; A is the empirically derived constant; and B is an empirically derived constant (35). An example of an analytical equation is

$$R = r_\mathrm{i}\frac{b}{a+b} = r_i\omega_b \tag{24}$$

where R is the reflectance, r_i is the radiance to reflectance conversion, a is the absorption, b is the backscattering, and ω_b is the backscattering albedo (34,36).

WETLAND DETECTION

The detection of wetlands has become extremely important for environmental agencies as more and more shoreline areas become developed. The main difficulty in identifying wetlands is the submerged aquatic vegetation. This vegetation can make water detection difficult with VIS and near- or midinfrared. However, thermal infrared allows comparison between aquatic vegetation and terrestrial vegetation because the water/vegetation combination has higher thermal inertia compared to that of land (12). The major limitation of this spectral region is that when the surrounding area is supersaturated with water, the thermal infrared comparison is difficult (12). Another method is to use active microwave radar data, such as SAR, and a qualitative approach to distinguish plants from open water and land.

ADDITIONAL TOPICS

Topics not covered in this article, but remain important in the area of water resources and hydrology, are land cover mapping, area delineation, slope delineation, flood forecasting, bathometry, ice and glaciers, and permafrost. Land cover mapping was discussed briefly

Table 6. Water Quality Parameters and Spectral Bands

Water Quality Parameter	Spectral Region	Note
Suspended sediments	700–800 μm(38)	Other scientists have found that optimum wavelength depends on concentration(37).
Chlorophyll (algae)	0.69 μm (red edge)	Distinction between chlorophyll and suspended sediment is difficult given the spectral data available; however, one should concentrate on the relationship between chlorophyll-a and the red edge of the spectrum(34)
Temperature	10.4–12.5 μm	
Oils	8–14 μm	Most commonly used sensors are those that are placed on airborne platforms, e.g., planes, because immediate and constant coverage is required for an oil spill; however, Landsat TM can also be used for monitoring. Spatial resolution is very important for patches and windrows. Infrared thermal remote sensing is most important for cleanups; ultraviolet is important as well because thin layers of oil emit high ultraviolet radiation patterns; passive microwave is good because oil emits more microwaves than water, but spatial resolution is poor (34).

Satellites useful for Water Resources[a]

Variable	Satellite	Central Wavelength or Frequency Required for Observation	Resolution	Coverage
Snow-covered area (SCA)	NOAA	0.62, 10.80 μm	1 km	2 per day
	SPOT	0.59, 0.69, 0.89 μm	10–25 m	16 day[b]
	GOES	0.64 μm (visible)	2 km	2 per hour
Snow depth	Nimbus 7	37 GHz	30 km	2 per hour
	DMSP			
Snow water equivalent (SWE)	SSM/I			
	MOS-1, MSR	19.3 GHz, 37.0 GHz	25 km	2 per day
		23.0 GHz, 31 GHz	23–32 km	2 per day
	ERS-1,2	C-band(5.3 GHz)V V	30 km	35 days
Wet snow	RADARSAT	SAR C-BAND(5.3 GHz) H H SAR	8–25 m	16 days[c]
Surface temperature	NOAA	12.0 μm (band 4)	1 km	2 per day
	Landsat TM	12.5 μm (band 6)	80 m	16 days
Evapotranspiration	NOAA	0.62, 0.91, 10.80, 12.0 μm	1 km	2 per day
	GOES	0.64,11.5 μm	2–8 km	2 per hour
	Meteosat			
	GOES	0.65 μm (visible)		
Precipitation	TRMM	Infrared	3 km	2 per hour
	Landsat TM	0.52, 0.60, 0.96, 0.90, 1.75, 2.35, 12.5 μm	30 m	16 days
Land cover/land use and vegetation	Landsat	0.55, 0.65, 0.75, 0.9 μm	80 m	16 days
	MSS NOAA	0.62, 0.91 μm (visible)	1 km	2 per day
	SPOT	0.59, 0.69, 0.89 μm	10–25 m	16 day[b]
	ERS-1,2	C-band (5.3 GHz) V V	30 m	35 days
Soil Moisture	RADARSAT	C-band (5.3 GHz) H H	8–25 m	16 days[c]
Groundwater	Landsat	0.95 μm (band 7, near IR)	80 m	16 days
	SPOT			
	ERS-1,2	0.59, 0.69, 0.89, μm	10–25 m	16 days[b]
	Radarsat	C-band (5.3 GHz) V V	30 m	35 days
	Landsat TM	C-band (5.3 GHz) H H	8–25 m	16 days[c]
	Landsat	0.52, 0.60, 0.69, 0.90, 1.75, 2.35, 12.5 μm	30 m	16 days
Surface Water	MSS	0.55, 0.65, 0.75, 0.9 μm	80 m	16 days

[a] Compiled from Reference (12).

[b] Coverage frequency can be up to 2–3 days in northern latitudes using the off-nadir pointing capability of the satellite.

[c] Coverage frequency can be increased depending on incidence angle.

in the opening of this article. Terrain elevation analysis from imagery remotely obtained by stereopairs and interferometry provides the main method for area and slope delineation. Flood forecasting may occur by real-time monitoring of water surface elevations with light detection and ranging (LIDAR) or by simulation of soil moisture dynamics, land surface runoff, and river routing with remotely sensed inputs of vegetation, terrain, and precipitation.

CONCLUSION

There are several limitations of remote sensing of water resources to keep in mind while delineating a water resource application. The first limitation, that parameters are never measured directly, cannot be solved easily but rather calls for an increase in confidence in calibration and validation methods for the equation and the ability to select equations/algorithms or methods proper for the particular study (2). A lack of suitable algorithms for available data can also hinder studies. To a certain extent, remote sensing is new for water resources, so data required for specific algorithms may not be accessible. Calculating parameters such as evaporation may require expensive mapping programs that cannot be afforded by certain research institutes and universities. Finally, much of VIS and IR remote sensing depends on the weather, such as the presence of clouds, and other uncontrollable factors.

APPENDIX

Common Acronyms

AVHRR	Advanced very high resolution radiometer
DEM	Digital elevation model
EOS	Earth Observing System
ERS-1,2	European Resources Satellite 1,2
GIS	Geographic Information Systems
IR	Infrared radiation
NOAA	National Oceanic and Atmosphere Administration
NDVI	Normalized difference vegetation index
PBL	Planetary boundary layer model
RADARSAT	Radar satellite
SAR	Synthetic aperture radar
SPOT	Système Probatoire pour l'Observation de la Terre
SSM/I	Special sensor microwave imager

SWE	Snow water equivalent
TIR	Thermal infrared radiation
VIS	Visible radiation

BIBLIOGRAPHY

1. Jensen, J.R. (1996). *Introductory Digital Image Processing: A Remote Sensing Perspective*, 2nd Edn. Prentice-Hall, Upper Saddle River, NJ.
2. Schultz, G.A. and Engman, E.T. (2000). *Remote Sensing in Hydrology and Water Management*. Springer, New York.
3. Rott, H. (2000). Physical principles and technical aspects of remote sensing. In: *Remote Sensing in Hydrology and Water Management*. G.A. Shultz and E.T. Engman (Eds.). Springer, New York.
4. Meneti, M. (2000). Evaporation. In: *Remote Sensing in Hydrology and Water Management*. G.A. Shultz and E.T. Engman (Eds.). Springer, New York.
5. Priestly, C.H.B. and Taylor, R.J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Weather Rev.* **100**: 81–92.
6. Barton, I.J. (1978). A case study comparison of microwave radiometer measurements over bare and vegetated surfaces. *J. Geophys. Res.* **83**: 3513–3517.
7. Davies, J.A. and Allen, C.D. (1973). Equilibrium, potential and actual evaporation from cropped surfaces in southern Ontario. *J. Appl. Meteorol.* **12**: 649–657.
8. Choudhury, B.J. (1997). Global pattern of potential evaporation calculated from the Penman–Monteith equation using satellite and assimilated data. *Remote Sensing Environ.* **61**: 64–81.
9. D'Urso, G. and Meneti, M. (1996). Performance indicators for the statistical evaluation of digital image classifications. *ISPRS J. Photogr. Remote Sensing* **51**(2): 78–90.
10. D'Urso, G. and Meneti, M. (1995). Mapping crop coefficients in irrigated areas from Landsat TM images. *SPIE Int. Soc. Opt. Eng.* **2585**: 41–47.
11. Barber, D.G., Hocheim, K.P., Dixon, R., Mosscrop, D.R., and McMullan, M.J. (1996). The role of Earth observation technologies in flood mapping: A Manitoba case study. *Can. J. Remote Sensing* **22**(1): 137–143.
12. Pietroniro, A. and Leconte, R. (2000). A review of Canadian remote sensing applications in hydrology, 1995–1999. *Hydrol. Processes* **12**: 1641–1666.
13. Collier, C.G. (2000). Precipitation. In: *Remote Sensing in Hydrology and Water Management*. G.A. Shultz and E.T. Engman (Eds.). Springer, New York.
14. Uijlenhoet, R. (2001). Raindrop size distributions and radar reflectivity-rain rate relationships for radar hydrology. *Hydrol. Earth Syst. Sci.* **5**(4): 615–627.
15. Battan, L.J. (1973). *Radar Observation of the Atmosphere*. The University of Chicago Press, Chicago, IL.
16. Smith, P.L. (1984). Equivalent radar reflectivity factors for snow and ice particles. *J. Climate Appl. Meteorol.* **23**(8): 1258–1260.
17. Matrosov, S.Y. (1998). A dual-wavelength radar method to measure snowfall rate. *J. Appl. Meteorol.* **37**(11): 1510–1521.
18. Kidder, S. and Vonder Harr, T. (1995). *Satellite Meterology*. Academic Press, New York, pp. 307–343.
19. Joss, J., Schram, K., Thams, J.C., and Waldvogel, A. (1970). *On the Quantitative Determination of Precipitation by Radar*. Wissenschaftliche Mitteilung Nr 63, Zurich, Eidegenossische Kommision zum Studium der Hagelbildung und der Hagelabwehr.
20. Marshall, J.S. and (Mc)Palmer, W. (1948). The distribution of raindrops with size. *J. Meteorol.* **5**: 165–166.
21. Blanchard, D.C. (1953). Raindrop size distribution in Hawaiian rains. *J. Meteorol.* **10**: 457–473.
22. Jones, D.M.A. (1956) Raindrop size distribution in Hawaiian rains. *J. Meteorol.* **10**: 457–473.
23. Gunn, K. and Marshall, J. (1958). The distribution with size of aggregate snowflakes. *J. Meteorol.* **15**: 452–561.
24. Sekhon, R.S. and Srivastava, R.C. (1970). Snow size spectra and radar reflectivity. *J. Atm. Sci.* **27**: 229–267.
25. Puhakka, T. (1975). *On the Dependence of Z-R Relation on the Temperature in Snowfall. Preprints*. 16th *Radar Meterol. Conf.*, Houston, TX, American Meteorology Society, pp. 504–507.
26. Boucher, R.J. and Wieler, J.G. (1985). Radar determination of snowfall rate and accumulation. *J. Climate Appl. Meteorol.* **15**: 452–461.
27. Fujiyoshi, Y., Endoh, T., Yamada, T., Tsuboki, K., Tachbana, Y., and Wakaham, G. (1990). Determination of a Z–R relationship for snowfall using a radar and high sensitivity snow gauges. *J. Appl. Meteorol.* **29**: 147–152.
28. Engman, E.T. (2000). Soil Moisture. In: *Remote Sensing in Hydrology and Water Management*. G.A. Shultz and E.T. Engman (Eds.). Springer, New York.
29. Schmugge, T. (1990). Measurements of surface soil moisture and temperature. In: *Remote Sensing of Biosphere Functioning*. R.J. Hobbs and H.A. Mooney (Eds.). Springer-Verlag, New York.
30. Rango, A., Walker, A.E., and Goodison, B.E. (2000). Snow and ice. In: *Remote Sensing in Hydrology and Water Management*. G.A. Shultz and E.T. Engman (Eds.). Springer, New York.
31. Carroll, T.R. (1990). Airborne and Satellite Data Used to Map Snow Cover Operationally in the U.S. and Canada. *Proc. Int. Symp. Remote Sensing Water Resour.*, Enschede, The Netherlands, **147**: 155.
32. Rott, H. (1993). Capabilities of Microwave Sensors for Monitoring Areal Extent and Physical Properties of the Snowpack. *Proc. NATO Adv. Res. Workshop Global Environ. Change Land Surf. Processes Hydrol.*, Tuscon, AZ.
33. USEPA. (1999). *Remotely Sensing Water Resources: an EPA/NASA Workshop*. Office of Water (4503F), United States Environmental Protection Agency, Washington, DC.
34. Ritchie, J.C. and Schiebe, F.R. (2000). Water quality. In: *Remote Sensing in Hydrology and Water Management*. G.A. Shultz and E.T. Engman (Eds.). Springer, New York.
35. Morel, A. and Gordon, H.R. (1980). Report of the working group on water color. *Boundary Layer Meteorol.* **18**: 343–355.
36. Dekker, A.G., Malthus, T.J., and Hoogenboom, H.J. (1995). The remote sensing of inland water quality. In: *Advances in Remote Sensing*. F.M. Danson and S.E. Plummer (Eds.). John Wiley, Chichester, pp. 123–142.
37. Curran, P.J. and Novo, E.M.M. (1988). The relationship between suspended sediment concentration and remotely sensed spectral radiance: A review. *J. Coastal Res.* **4**: 351–368.
38. Ritchie, J.C., Schiebe, F.R., McHenry, J.R., Wilson, R.B., and May, J. (1976). Sun angle, reflected solar radiation and suspended sediments in north Missippi reservoirs. In: *Remote Sensing of Earth Resources Vol. IV*. F. Shahrokhi (Ed.). University of Tennessee Space Institute, Tullahoma, TN.

ATMOSPHERIC SCIENTISTS

U.S. Department of Labor,
Bureau of Labor Statistics

SIGNIFICANT POINTS

- Almost 4 in 10 atmospheric scientists work for the Federal Government, which is the largest employer of such workers.
- A bachelor's degree in meteorology, or in a closely related field with courses in meteorology, is the minimum educational requirement; a master's degree is necessary for some positions, and a Ph.D. is required for most research positions.

NATURE OF THE WORK

Atmospheric science is the study of the atmosphere—the blanket of air covering the Earth. *Atmospheric scientists*, commonly called *meteorologists*, study the atmosphere's physical characteristics, motions, and processes, and the way in which it affects the rest of our environment. The best known application of this knowledge is in forecasting the weather. Aside from predicting the weather, scientists also attempt to identify and interpret climate trends, understand past weather, and analyze today's weather. However, weather information and meteorological research also are applied in air-pollution control, agriculture, forestry, air and sea transportation, defense, and the study of possible trends in the Earth's climate, such as global warming, droughts, or ozone depletion.

Atmospheric scientists who forecast the weather, known professionally as *operational meteorologists*, are the largest group of specialists. They study information on air pressure, temperature, humidity, and wind velocity; and apply physical and mathematical relationships to make short-range and long-range weather forecasts. Their data come from weather satellites, weather radars, sensors, and weather stations in many parts of the world. Meteorologists use sophisticated computer models of the world's atmosphere to make long-term, short-term, and local-area forecasts. More accurate instruments for measuring and observing weather conditions, as well as high-speed computers to process and analyze weather data, have revolutionized weather forecasting. Using satellite data, climate theory, and sophisticated computer models of the world's atmosphere, meteorologists can more effectively interpret the results of these models to make local-area weather predictions. These forecasts inform not only the general public, but also those who need accurate weather information for both economic and safety reasons, such as the shipping, air transportation, agriculture, fishing, forestry, and utilities industries.

The use of weather balloons, launched a few times a day to measure wind, temperature, and humidity in the upper atmosphere, is currently supplemented by sophisticated atmospheric monitoring equipment that transmits data as frequently as every few minutes. Doppler radar, for example, can detect airflow patterns in violent storm systems—allowing forecasters to better predict tornadoes and other hazardous winds, and to monitor the storms' direction and intensity. Combined radar and satellite observations allow meteorologists to predict flash floods.

Some atmospheric scientists work in research. *Physical meteorologists*, for example, study the atmosphere's chemical and physical properties; the transmission of light, sound, and radio waves; and the transfer of energy in the atmosphere. They also study factors affecting the formation of clouds, rain, and snow; the dispersal of air pollutants over urban areas; and other weather phenomena, such as the mechanics of severe storms. *Synoptic meteorologists* develop new tools for weather forecasting using computers and sophisticated mathematical models of atmospheric activity. *Climatologists* study climactic variations spanning hundreds or even millions of years. They also may collect, analyze, and interpret past records of wind, rainfall, sunshine, and temperature in specific areas or regions. Their studies are used to design buildings, plan heating and cooling systems, and aid in effective land use and agricultural production. Environmental problems, such as pollution and shortages of fresh water, have widened the scope of the meteorological profession. *Environmental meteorologists* study these problems and may evaluate and report on air quality for environmental impact statements. Other research meteorologists examine the most effective ways to control or diminish air pollution.

WORKING CONDITIONS

Most weather stations operate around the clock, 7 days a week. Jobs in such facilities usually involve night, weekend, and holiday work, often with rotating shifts. During weather emergencies, such as hurricanes, operational meteorologists may work overtime. Operational meteorologists also are often under pressure to meet forecast deadlines. Weather stations are found everywhere—at airports, in or near cities, and in isolated and remote areas. Some atmospheric scientists also spend time observing weather conditions and collecting data from aircraft. Weather forecasters who work for radio or television stations broadcast their reports from station studios, and may work evenings and weekends. Meteorologists in smaller weather offices often work alone; in larger ones, they work as part of a team. Meteorologists not involved in forecasting tasks work regular hours, usually in offices. Those who work for private consulting firms or for companies analyzing and monitoring emissions to improve air quality usually work with other scientists or engineers; fieldwork and travel may be common for these workers.

EMPLOYMENT

Atmospheric scientists held about 7,700 jobs in 2002. The Federal Government was the largest single employer

---

of civilian meteorologists, accounting for about 2,900. The National Oceanic and Atmospheric Administration (NOAA) employed most Federal meteorologists in National Weather Service stations throughout the Nation; the remainder of NOAA's meteorologists worked mainly in research and development or management. The U.S. Department of Defense employed several hundred civilian meteorologists. Others worked for professional, scientific, and technical services firms, including private weather consulting services; radio and television broadcasting; air carriers; and State government.

Although several hundred people teach atmospheric science and related courses in college and university departments of meteorology or atmospheric science, physics, earth science, or geophysics, these individuals are classified as college or university faculty, rather than atmospheric scientists. (See the statement on **postsecondary teachers** elsewhere in the *Handbook*.)

In addition to civilian meteorologists, hundreds of Armed Forces members are involved in forecasting and other meteorological work. (See the statement on **job opportunities in the Armed Forces** elsewhere in the *Handbook*.)

TRAINING, OTHER QUALIFICATIONS, AND ADVANCEMENT

A bachelor's degree in meteorology or atmospheric science, or in a closely related field with courses in meteorology, usually is the minimum educational requirement for an entry-level position as an atmospheric scientist.

The preferred educational requirement for entry-level meteorologists in the Federal Government is a bachelor's degree—not necessarily in meteorology—with at least 24 semester hours of meteorology courses, including 6 hours in the analysis and prediction of weather systems, 6 hours of atmospheric dynamics and thermodynamics, 3 hours of physical meteorology, and 2 hours of remote sensing of the atmosphere or instrumentation. Other required courses include 3 semester hours of ordinary differential equations, 6 hours of college physics, and at least 9 hours of courses appropriate for a physical science major—such as statistics, chemistry, physical oceanography, physical climatology, physical hydrology, radiative transfer, aeronomy, advanced thermodynamics, advanced electricity and magnetism, light and optics, and computer science. Sometimes, a combination of education and appropriate experience may be substituted for a degree.

Although positions in operational meteorology are available for those with only a bachelor's degree, obtaining a second bachelor's degree or a master's degree enhances employment opportunities and advancement potential. A master's degree usually is necessary for conducting applied research and development, and a Ph.D. is required for most basic research positions. Students planning on a career in research and development need not necessarily major in atmospheric science or meteorology as an undergraduate. In fact, a bachelor's degree in mathematics, physics, or engineering provides excellent preparation for graduate study in atmospheric science.

Because atmospheric science is a small field, relatively few colleges and universities offer degrees in meteorology or atmospheric science, although many departments of physics, earth science, geography, and geophysics offer atmospheric science and related courses. Prospective students should make certain that courses required by the National Weather Service and other employers are offered at the college they are considering. Computer science courses, additional meteorology courses, a strong background in mathematics and physics, and good communication skills are important to prospective employers. Many programs combine the study of meteorology with another field, such as agriculture, oceanography, engineering, or physics. For example, hydrometeorology is the blending of hydrology (the science of Earth's water) and meteorology, and is the field concerned with the effect of precipitation on the hydrologic cycle and the environment. Students who wish to become broadcast meteorologists for radio or television stations should develop excellent communication skills through courses in speech, journalism, and related fields. Those interested in air quality work should take courses in chemistry and supplement their technical training with coursework in policy or government affairs. Prospective meteorologists seeking opportunities at weather consulting firms should possess knowledge of business, statistics, and economics, as an increasing emphasis is being placed on long-range seasonal forecasting to assist businesses.

Beginning atmospheric scientists often do routine data collection, computation, or analysis, and some basic forecasting. Entry-level operational meteorologists in the Federal Government usually are placed in intern positions for training and experience. During this period, they learn about the Weather Service's forecasting equipment and procedures, and rotate to different offices to learn about various weather systems. After completing the training period, they are assigned a permanent duty station. Experienced meteorologists may advance to supervisory or administrative jobs, or may handle more complex forecasting jobs. After several years of experience, some meteorologists establish their own weather consulting services.

The American Meteorological Society offers professional certification of consulting meteorologists, administered by a Board of Certified Consulting Meteorologists. Applicants must meet formal education requirements (though not necessarily have a college degree), pass an examination to demonstrate thorough meteorological knowledge, have a minimum of 5 years of experience or a combination of experience plus an advanced degree, and provide character references from fellow professionals.

JOB OUTLOOK

Employment of atmospheric scientists is projected to increase about as fast as the average for all occupations through 2012. The National Weather Service has completed an extensive modernization of its weather forecasting equipment and finished all hiring of meteorologists needed to staff the upgraded stations. The Service has no plans to increase the number of weather stations or the number of meteorologists in existing stations. Employment of meteorologists in other Federal agencies is expected to remain stable.

On the other hand, job opportunities for atmospheric scientists in private industry are expected to be better than for those in the Federal Government over the 2002-12 period. As research leads to continuing improvements in weather forecasting, demand should grow for private weather consulting firms to provide more detailed information than has formerly been available, especially to weather-sensitive industries. Farmers, commodity investors, radio and television stations, and utilities, transportation, and construction firms can greatly benefit from additional weather information more closely targeted to their needs than the general information provided by the National Weather Service. Additionally, research on seasonal and other long-range forecasting is yielding positive results, which should spur demand for more atmospheric scientists to interpret these forecasts and advise weather-sensitive industries. However, because many customers for private weather services are in industries sensitive to fluctuations in the economy, the sales and growth of private weather services depend on the health of the economy.

There will continue to be demand for atmospheric scientists to analyze and monitor the dispersion of pollutants into the air to ensure compliance with Federal environmental regulations outlined in the Clean Air Act of 1990, but related employment increases are expected to be small. Opportunities in broadcasting are rare and highly competitive, making for very few job openings.

EARNINGS

Median annual earnings of atmospheric scientists in 2002 were $60,200. The middle 50 percent earned between $39,970 and $76,880. The lowest 10 percent earned less than $30,220, and the highest 10 percent earned more than $92,430.

The average salary for meteorologists in nonsupervisory, supervisory, and managerial positions employed by the Federal Government was about $74,528 in 2003. Meteorologists in the Federal Government with a bachelor's degree and no experience received a starting salary of $23,442 or $29,037, depending on their college grades. Those with a master's degree could start at $35,519 or $42,976; those with the Ph.D., at $51,508. Beginning salaries for all degree levels are slightly higher in areas of the country where the prevailing local pay level is higher.

RELATED OCCUPATIONS

Workers in other occupations concerned with the physical environment include **environmental scientists and geoscientists, physicists and astronomers, mathematicians**, and **civil, chemical**, and **environmental engineers**.

SOURCES OF ADDITIONAL INFORMATION

Information about careers in meteorology is available on the Internet from:

- American Meteorological Society. Internet: http://www.ametsoc.org/AMS. Phone: (617)-227-2425.

Information on obtaining a meteorologist position with the Federal Government is available from the U.S. Office of Personnel Management through a telephone-based system. Consult your telephone directory under U.S. Government for a local number or call (703) 724−1850; Federal Relay Service: (800) 877−8339. The first number is not toll-free, and charges may result. Information also is available from the Internet site: http://www.usajobs.opm.gov/.

OOH ONET CODES

19−2021.00

Suggested citation: Bureau of Labor Statistics, U.S. Department of Labor, *Occupational Outlook Handbook, 2004-05 Edition*, Atmospheric Scientists, on the Internet at http://www.bls.gov/oco/ocos051.htm (visited August 14, 2004).

RAIN SIMULATORS

Joseph Holden
University of Leeds
Leeds, United Kingdom

Rain simulators have been extensively used in hydrological, pedological, and geomorphological problems (1−13). Simulators allow much greater control of the rainfall variable for experimental purposes and have been used in both field and laboratory applications (13). The amount, intensity, and duration of rain can be controlled, along with other parameters such as drop-size distribution and water chemistry, to varying degrees, depending on the system of application. Unlike a ring infiltrometer, the surface does not pond immediately but will only do so at some later stage, if the input is great enough. The time to ponding will depend on the application rate as well as on the hydraulic properties of the soil. Surface runoff can be collected to determine infiltration rates (by subtracting surface runoff rates from application rates) and erosion, either as a whole sample for smaller plots or in a subsampling strategy for larger plots.

TYPES OF RAINFALL SIMULATORS AND CHOICE OF APPROPRIATE EQUIPMENT

There is a range of simulator designs, and a database catalogue has been developed incorporating design information by the British Geomorphological Research Group (www.geog.le.ac.uk/bgrg). Some other designs are published in Costin and Gilmour (14); Romkens et al. (15); Bowyer-Bower and Burt (16); Esteves et al. (17); and Waddington and Devito (18). There are two rainfall simulator systems: spray nozzles (including rotating sprays) and drip-screens. Spray systems generally supply rainfall in pulses to the ground (14), and rotating sprays deliver rainfall across a large surface area although intensity usually decreases with distance from the rotating nozzle. The spray-type systems often provide rainfall at terminal velocities which approach that of natural rainfall. Drip systems use hypodermic syringes (15) or other drop formers across a fixed grid to produce rainfall across relatively small surface areas. The drop formers are often not raised high enough to allow representative terminal velocities, but they generally allow a constant rainfall rate whose drop sizes are more easily controlled than in the spray systems (16).

Various factors, including money, time available, the purpose of the experiment, and the local climatic conditions whose simulation is desired, influence the choice of simulator. Table 1 lists some of the factors of importance in choosing a type of rainfall simulator and the most suitable type of simulator for each factor. The amount of water used to provide rainfall at any given intensity is generally much less for a drip-screen design than for spray-type simulators (16). This may be crucially important in remote arid or semiarid areas. The use of a drip-type simulator generally means, however, that smaller plot areas will be used for analysis. Using appropriate pressure regulators, rain simulators in the field or laboratory can be attached to a mains water supply. It usual to control the amount and intensity of rain during a simulation by using a pressure regulator, but often a simple tap connected to a calibrated manometer board will suffice (16).

Table 1. Desirable Characteristics of a Rainfall Simulator

Desired Attribute	Best Suited Simulator
Easily transportable	Drip-type
Efficient use of water	Drip-type
Protection from the wind to increase accuracy of rainfall application volumes into selected plot	Drip-type
Reproduction of high intensity rainfall to reflect expected natural conditions	Spray-type
Reproduction of low intensity rainfall to reflect expected natural conditions	Drip-type
Long duration rainfall	Drip-type
Achievement of 'natural' terminal velocity	Spray-type
Control over drop-size distribution	Drip-type (but not very flexible)
Attainment of desired drop size relative to rainfall intensity (i.e., flexibility)	Spray-type
Accurate replication of rainfall parameters	Drip-type
Uniformity of rainfall across the whole plot	Drip-type
Large plot area	Spray-type
Adaptability to difficult terrain and vegetation	Spray-type

Drop-Size Distribution

Drops of a uniform size formed by a rain simulator do not represent those in natural rainfall. If the experiment requires simulating a natural drop-size distribution (19–21), then it is often necessary to have a wire mesh that scatters and breaks up drops into different sizes. For drip-type simulators, Bowyer-Bower and Burt (16) found that suspending a wire mesh below the drop former of 3 mm × 3 mm spacing and 1 mm diameter wire was suitable. Salles et al. (22) demonstrated that it is necessary to sample at least 10,000 drops to estimate the drop-size distribution from a rain simulator at an accuracy of 3% or less. Several methods are available for examining drop-size distribution, including light diffraction and filter paper collection. An economic and simple approach is to allow drops to fall into finely sieved flour and then to bake the pellets and sieve them through meshes of different diameters. Larger intensity rainfall is associated with larger raindrop size (21), but it is often difficult to replicate changing drop sizes as intensity is altered using rain simulators, although spray-type simulators are more flexible.

Achievement of Terminal Velocity and Energy Characteristics of the Rain

Where erosion or soil crusting processes are the focus of study, it is essential to know the kinetic energy of the raindrops produced by a rain simulator (23). Kinetic energy is crucial for soil detachment and transport. Natural rain normally falls from great heights by which time the droplets have achieved terminal velocity (maximum velocity in air—they cannot fall any faster because air resistance prevents any further acceleration due to gravity). Droplets that fall at terminal velocity have more kinetic energy than those falling from small heights. Larger drops require a greater height to reach their terminal velocity. Droplets of 1 mm in diameter require around 2.5 m, and droplets of 2 mm in diameter require around 6 m to reach terminal velocity. Thus, if a rainfall simulator is not positioned high enough, then the kinetic energy of the drops will be significantly less than that of natural rainfall. Knowledge of the drop size distribution is therefore important for determining the rain splash energy provided.

Spatial and Temporal Coverage of a Plot

For experiments that are concerned only with some aggregate or mean effect of simulated rain, then variations in rainfall delivery across a plot may be unimportant. Where rainfall simulation is being used to study small-scale processes that are themselves spatially variable (such as rill initiation), then knowledge of the simulator's inherent variability is vital (24). Rainfall delivery may be measured quite simply by placing a series of collectors across a trial plot and recording rainfall delivery to the ground across a range of rainfall intensities and durations. Many spray systems have greater intensity rains closer to the spray nozzle, and intensity declines as the distance from the nozzle increases.

Typical Measurements

Depending on the nature of the experiment, it is usually necessary to create a bounded plot when simulating rain. This is particularly important if a parameter of interest is infiltration. It is usual to bound the plot on three sides and then to collect the surface runoff from the unbounded fourth (downslope) side of the plot. The bounding prevents leakage out of the plot, so that more accurate estimates of infiltration rates can be made. Measurement of the surface runoff can indicate the infiltration rate within the plot, as long as the application rate is known: At steady state, Infiltration = Rainfall intensity—Surface Runoff. Units of length are used for the three variables in the equation, and so one must remember to divide the surface runoff volumes by the plot area. There will be a time lag in response so that in the early stages of the experiment when the soil surface is wetting and water is stored in detention pools, then the previous equation will not hold. The equation also ignores evaporation and storage on vegetation surfaces. If the interest is in infiltration rates, it is often necessary to run the experiment long enough to allow runoff rates to stabilize. Early in the experiment, runoff rates are likely to increase over time, but in theory should eventually stabilize if the rainfall intensity remains constant. It should then be possible to apply the theoretical infiltration curve of Philip (25) which is an exponential curve fitted through the first and last data point of the order of $I = A + Bt^{-0.5}$, where I = infiltration rate at time t and A and B are constants. This equation may be solved using simultaneous equations (i.e., equations at two different times) to allow constants A and B to be evaluated.

In practice, however, surface runoff from the plot on most slopes of interest will rarely stabilize completely because of surface ponding and episodic cut and fill of microtopographical features on the surface and the burst out of water followed by a period of pool refill and microchannel change (12). It may therefore be better to use a mean value across a sample of data points to estimate infiltration. If the rainfall intensity is lower than the infiltration rate into the surface of investigation, then it may be that no surface runoff is observed. If surface runoff is collected, it is also possible to collect sediment, solutes, and microorganisms moved by the runoff.

Typically, results are scaled up to the hillslope and sometimes the catchment scale. However, this should be done only with great caution. First, most hillslopes have a high natural heterogeneity such that plots may not be representative. Second, coupling of water, sediment, nutrients, or microorganisms within a plot may not be the same as on a hillslope. Sediment moved from the top to bottom of a plot, for example, may not move from the top to bottom of a hillslope during a real storm because there may be natural zones of deposition and erosion and larger hillslopes may not provide the good coupling seen in a plot-scale experiment.

Almost all field rainfall simulator experiments have collected runoff solely from the surface layer and have disregarded any lateral through-flow in deeper soil layers. However, using an appropriate apparatus, it is possible to investigate soil water flow mechanisms using a rain simulator by digging a pit at the lower end of the plot and collecting runoff from different layers of the subsurface (12). However, pit disturbance may change the local hydrologic gradients resulting in unwanted changes of the natural processes, particularly, if the layers of investigation are deep below the surface.

IMPORTANCE OF RAINFALL INTENSITY USED

Some rainfall simulators can reproduce natural rainfall typical of tropical storms across 50-m plots (e.g., Reference 17 reproduced 70 mm h^{-1} storms), whereas others can reproduce gentle rains typical of temperate parts of the world whose intensities are as low as 3 mm h^{-1} (12). The choice of rainfall intensity used in a simulation may be very important, and many hydrologic studies have ignored this issue. Most studies have simulated only one rainfall intensity. Some studies, however, found that as rainfall intensity is increased on a simulation plot, the steady infiltration rate also increases (12,26). This is very important and has major implications for conclusions drawn from experiments. There are likely to be physical reasons for this phenomenon, including increased ponding development on the soil surface as rain intensity is increased that leads to a greater hydraulic head that forces more water into the soil (27–29); increased ponding that results in dispersing and suspending soil crusts and surface particles, allowing pores to unblock; and most soil plots are nonuniform; some parts allow much more infiltration than others (30). Therefore, if rainfall intensity is increased, then proportionally more water may infiltrate into particular parts of a plot. This will result in an increased plot-averaged value of the infiltration rate. Thus, those experiments that have simulated unrealistic rainfall intensities (e.g., 150 mm h^{-1} in the British Isles where rainfall very rarely exceeds 12 mm h^{-1}) may have unrealistic results. The results from an intense rainfall simulation may assume that the infiltration capacity of the soil is 80 mm h^{-1}, for example, but if a more realistic intensity had been used, then the result may be an infiltration capacity of 6 mm h^{-1}. Thus, conclusions drawn about runoff production in soil water processes, irrigation, and erosion may be misleading, and care should be taken in choosing an appropriately representative range of rainfall intensities.

BIBLIOGRAPHY

1. De Ploey, J., Savat, J., and Moeyersons, J. (1976). The differential impact of some soil factors on flow, runoff creep and rainwash. *Earth Surface Processes* **1**: 151–161.
2. Lusby, G.C. (1977). Determination of runoff and sediment yield by rainfall simulation. In: *Research Techniques, Erodibility and Sediment Delivery*. T.J. Toy (Ed.). Geobooks, Norwich, pp. 19–30.
3. Bork, H.R. and Rohdenburg, H. (1981). Rainfall simulation in southeast Spain: analysis of overland flow and infiltration. In: *Soil Conservation, Problems and Prospects*. R.P.C. Morgan (Ed.). John Wiley, Chichester, pp. 293–302.
4. Imeson, A.C. (1983). Studies of erosion threshold in semi-arid lands: field measurements of soil loss and infiltration in northern Morocco. In: *Rainfall Simulation, Runoff and Soil Erosion*. J. De Ploey (Ed.). Catena Supplement 4: pp. 79–89.

5. Pilgrim, D.H. and Huff, D.D. (1983). Suspended sediment in rapid subsurface stormflow on a large field plot. *Earth Surface Processes* **8**: 451–453.
6. Imeson, A.C. and Kwaad, F.J.P.M. (1990). The response of tilled soils to wetting by rainfall and the dynamic character of soil erodibility. In: *Soil Erosion on Agricultural Land*. J. Boardman, I.D.L. Foster, and J.A. Dearing (Eds.). John Wiley, Chichester, pp. 3–14.
7. Morgan, R.P.C. (1995). *Soil Erosion and Conservation*. Longman, London.
8. Bergkamp, G. (1998). A hierarchical view of the interactions of runoff and infiltration with vegetation and microtopography in semiarid shrublands. *Catena* **33**: 201–220.
9. Cerda, A. (1998). Relationships between climate and soil hydrological and erosional characteristics along climatic gradients in Mediterranean limestone areas. *Geomorphology* **25**: 123–134.
10. Foster, I.D.L., Fullen, M.A., Brandsma, R.T., and Chapman, A.S. (2000). Drip-screen rainfall simulators for hydro- and pedo-geomorphological research: The Coventry experience. *Earth Surface Processes and Landforms* **25**: 691–708.
11. Duiker, S.W., Flanagan, D.C., and Lal, R. (2001). Erodibility and infiltration characteristics of five major soils of southwest Spain. *Catena* **45**: 103–121.
12. Holden, J. and Burt, T.P. (2002). Infiltration, runoff and sediment production in blanket peat catchments: implications of field rainfall simulation experiments. *Hydrological Processes*. **16**: 2537–2557.
13. Holden, J. and Burt, T.P. (2002). Laboratory experiments on drought and runoff in blanket peat. *European Journal of Soil Science* **53**: 675–689.
14. Costin, A.B. and Gilmour, D.A. (1970). Portable rainfall simulator and plot unit for use in field studies of infiltration, runoff and erosion. *Journal of Applied Ecology* **7**: 193–200.
15. Romkens, M.J.M., Glenn, L.F., Nelson, D.W., and Roth, C.B. (1975). A laboratory rainfall simulator for infiltration and soil detachment studies. *Soil Science Society of America Proceedings* **39**: 158–160.
16. Bowyer-Bower, T.A.S. and Burt, T.P. (1989). Rainfall simulators for investigating soil response to rainfall. *Soil Technology* **2**: 1–16.
17. Esteves, M. et al. (2000). The ‘EMIRE’ large rainfall simulator: design and field testing. *Earth Surface Processes and Landforms* **25**: 681–690.
18. Waddington, J.M. and Devito, K.J. (2001). Portable irrigation system for studying hillslope and wetland runoff generation processes. *Hydrological Processes* **15**: 281–287.
19. Laws, J.O. and Parsons, D.A. (1943). The relation of raindrop size to intensity. *Transactions of the American Geophysical Union* **24**: 452–60.
20. Best, A.C. (1950). The size distribution of raindrops. *Quarterly Journal of the Royal Meteorological Society* **76**: 16–36.
21. Hudson, N.W. (1971). *Soil Conservation*. Batsford, London.
22. Salles, C., Poesen, J., and Borselli, L. (1999). Measurement of simulated drop size distribution with an optical spectro pluvimeter: sample size considerations. *Earth Surface Processes and Landforms* **24**: 545–556.
23. Epema, G.F. and Riezebos, H.Th. (1983). Fall velocity of waterdrops at different heights as a factor influencing erosivity of simulated rain. In: *Rainfall Simulation, Runoff and Soil Erosion*. J. DePloey (Ed.). Catena Supplement 4: pp. 2–17.
24. Lascelles, B., Favis-Mortlock, D.T., Parsons, A.J., and Guerra, A.J.T. (2000). Spatial and temporal variation in two rainfall simulators: implications for spatially explicit rainfall simulation experiments. *Earth Surface Processes and Landforms* **25**: 709–721.
25. Philip, J.R. (1957). The theory of infiltration 1. The infiltration equation and its solution. *Soil Science* **83**: 345–357.
26. Bowyer-Bower, T.A.S. (1993). Effects of rainfall intensity and antecedent moisture on the steady-state infiltration rate in a semi-arid region. *Soil Use and Management* **9**: 69–76.
27. Schiff, L. (1953). The effect of surface head on infiltration rates based on the performance of ring infiltrometers and ponds. *Transactions of the American Geophysical Union* **34**: 257–266.
28. Philip, J.R. (1958). The theory of infiltration 6. Effect of water depth over soil. *Soil Science* **85**: 278–286.
29. Schmid, B.H. (1989). On overland flow modelling: can rainfall excess be treated as independent of flow depth? *Journal of Hydrology* **107**: 1–8.
30. Hawkins, R.H. (1982). Interpretations of source area variability in rainfall-runoff relations. In: *Rainfall-runoff relationship. Proceedings of the International Symposium on rainfall runoff modelling held May 18–21, 1981, Mississippi State University*. Water Resources Publications, Littleton, CO, 303–324.

SNOW DENSITY

Pratap Singh
National Institute of Hydrology
Roorkee, , India

Density is one of the fundamental and important characteristics of snow. It is defined as the mass of snow per unit volume. To determine the density of snow, a core sample of snow is extracted from the accumulated snow on the ground using a snow sampler of known volume. The extracted snow sample is weighed and its mass is determined. The density of the sampled snow is determined by dividing the mass by the volume. The most common unit used for representing the density is g/cc or kg/m^3.

Freshly fallen snow is very light and has a very low density. The density of fresh snow varies from about 0.01 to 0.20 g/cc, but a rule of thumb often used for the average density of fresh snowfall is 0.1 g/cc when actual measurements are not made. Once snow accumulates on the ground and stays over time, the density of the snow increases. Metamorphic changes in snow, which involve readjustment of the shape and size of crystals, contribute to changes in the density of snow. Metamorphism also results from the compaction of snow caused by the pressure of overlying layers of snow and can also be accelerated by strong winds, warm temperatures, and intermittent melting of the snow cover. The density of snow changes significantly during the first 2–3 months when transformation of fresh snow into fine granular snow continues and recrystallization processes dominate. During a period of time, snow first transforms into firn, defined as snow that has survived for at least one melt season. When climatic conditions allow snow to remain

Table 1. Typical Densities of Snow in Different Forms[a]

Snow Type	Density g/cc
New snow at low temperature in calm air	0.01–0.03
New snow immediately after falling in calm air	0.05–0.07
Damp new snow	0.10–0.20
Settled snow	0.20–0.30
Depth hoar	0.20–0.30
Windpacked snow	0.35–0.40
Firn	0.40–0.65
Very wet snow and firn	0.70–0.80
Glacier ice	0.85–0.95

[a] Ref. 1.

for several years, it gradually transforms into glacier ice. Glacier ice is impermeable to air and water; some air is present only as bubbles. Glacier ice is the most compacted form of snow. It attains the maximum density at this stage, which is very close to the density of water, 1.0 g/cc. The density of snow at different stages is shown in Table 1.

To estimate the amount of water stored as snow in a basin, information on the density and depth of the snowpack is needed. Usually, such information is collected through snow surveys, which are organized to obtain the density and depth of snow at representative points in the basin. Snow density is determined for the total depth of snowpack because it varies with depth. The average density of a snowpack is determined by taking the average of the snow density obtained at different depth levels of the snowpack. The snow water equivalent at any point can be obtained by multiplying the snow density by the depth. In addition to the texture and wetness of snow, information on the density of snow near the surface is also required to ascertain the suitability of snow for skiing, which is one of the most important recreational activities in Europe and North America.

BIBLIOGRAPHY

1. Singh, P. and Singh, V.P. (2001). *Snow and Glacier Hydrology*. Kluwer Academic Publishers, Dordrecht, The Netherlands, p. 221.

FLATTOP MOUNTAIN SNOTEL SNOWPACK: WATER YEAR 2004

Global Change Research
Program—U.S. Geological
Survey

These materials prepared by the Global Change Research Program—U.S. Geological Survey, Northern Rocky Mountain, Glacier Field Station. Dan Fagre, Program Coordinator, email: dan_fagre@usgs.gov. For more information, visit the Glacier Field Station—Global Change Research Program web site.

The Flattop Mountain SNOTEL (SNOw TELemetry) station is one of nearly 600 similar stations operated throughout the western United States by the Natural Resources Conservations Service (NRCS, U.S. Department of Agriculture). These stations measure and record Snow Water Equivalent (SWE), which is the weight of snow water equivalent to inches of water. Most SNOTEL sites also measure temperature and precipitation; many now measure snow depth as well. Various agencies and organizations use the data from the stations to forecast water availability, floods, and avalanche hazard.

The Flattop Mountain SNOTEL station has operated since October 1st, 1969. To help calculate water storage and availability during agricultural growing seasons, the data are compiled by water year; a water year runs from 1 October - 30 September. SWE records from the station extend for 33 water years (1970–2002). Precipitation records from the station start water year 1979, with temperature records starting water year 1983.

The Flattop Mountain SNOTEL station is located at approximately 6300 feet in elevation on Flattop Mountain, a high plateau between the Lewis and Livingston Ranges in Glacier National Park, Montana. The site is three air miles south of the Continental Divide. During the winter, complex combinations of weather and terrain determine snowfall at the site. Westerly weather systems predominate, bringing moisture from the Gulf of Alaska or Pacific Ocean. Less frequent northerly systems spill drier Arctic air through passes on the Continental Divide as they slide south along the Rocky Mountain Front. Flattop Mountain's position between the Livingston and Lewis Ranges amplifies the effects of these large-scale weather systems; the two ranges rise 2–4000 feet above the site and orographically wring moisture from both westerly and upslope easterly storms. Flattop Mountain is a useful indicator of snowfall throughout Glacier National Park because it is subject to the factors that influence conditions elsewhere in the park.

The graph below shows the current water year SWE at Flattop Mountain along with snow depth and average daily temp:

Since the last update of this webpage (May 10th), the snowpack at Flattop Mtn. SNOTEL station continued its annual melt. Conditions in May and early June have generally been wet, with average temperatures. The station recorded. 1" or more of precipitation on 27 of 44 days, for a total of 8.3" of precipitation, 5.2" of that

---

during May. Nearly all of that fell as rain; the station has measured just 1” of SWE gain in May and .5” in June. The highest one-day rainfall was .9” on May 28th, while the longest stretch of days with precipitation was May 25th–May 30th. It was followed by the longest period without precipitation, May 31st–June 4th. The precipitation of the past four weeks is common for this time of year; the station typically shows a secondary peak in precipitation known locally as the “June Monsoon.”

The station recorded the warmest temperatures early in the month. After a mid-month winter storm, temperatures remained cool. Maximum and minimum temperatures for May were 16.7 C (60 F) on May 1st and -8 C (17.6 F) on May 12th. The highest mean daily temperature for the month was 8.2 C (47.8 F) on May 7th. The mean monthly temperature was 3.7 C (38.6 F), which is identical to the mean monthly temperature for May for the 21 years (1982–2003) that the station has recorded temperatures.

The average temperatures for the month ended a trend of above-average temperatures that dominated the winter. November has so far been the only month of the 2004 Water Year with mean temperatures below average. From December through April, mean monthly temperatures have remained above average, with December and January warmer than the mean (2.6 C/4.7 F and .5 C/1 F respectively), and February and March substantially warmer than average (3.2 C/5.8 F for February; 5 C/9 F for March). The difference between 2004 and the 21-year average narrowed in April and closed in May. The Mean Monthly Temperature Comparison Graph below shows this pattern. While the graph shows a relatively normal curve to the temperatures, the curve appears shifted a month forward. That is, early winter temperatures are typical mid-winter temperature, and late winter/early spring temperatures are typical of late spring. A. pdf version of this graph is available at Mean Monthly Temperature Graph.

The warm temperatures of the winter had a significant effect on the snowpack at the station. Despite a dry February, SWE gain for February-April was 15.9”, which is average. The warm temperatures, however, have resulted in 7” of SWE loss in the same period, leaving the snowpack considerably below average when it peaked on April 20th, a week earlier than the 34-year average of April 27th. The SWE total was 37.4”, or 81% of the 34-year average of 46.74”. In addition to the 7” of SWE loss in March and April, the station lost 1.2” of SWE in October, for a total loss of 8.2”—nearly enough to make up the 9.3 inch deficit on April 20th. It’s thus reasonable to conclude that the below average snowpack at Flattop Mtn. SNOTEL is primarily the result of warm temperatures rather than lack of precipitation. A graph showing the monthly SWE gain at Flattop Mountain SNOTEL is below; a. pdf version of the graph can be found at Monthly SWE 2004.

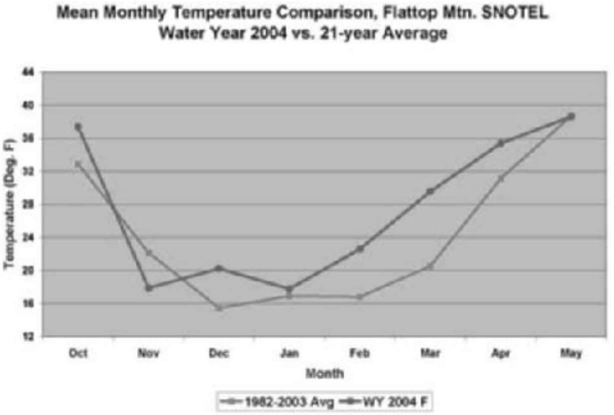

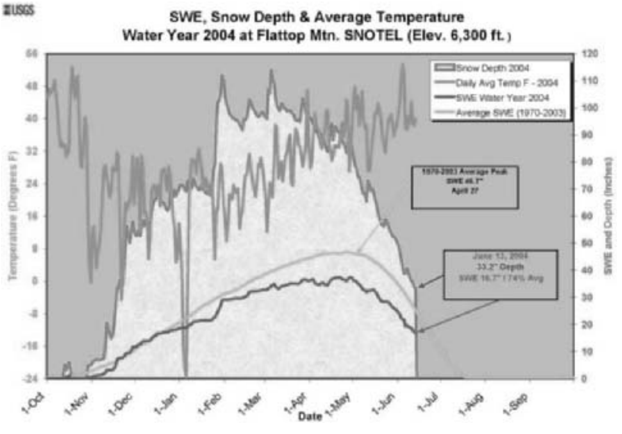

Since April 30th, the snow pillow at Flattop has recorded 22” of SWE loss and 1.5” of SWE gain. The station measured 13” of SWE loss and 1” of SWE gain in May. The greatest SWE loss occurred May 18th (1.1”) and June 3rd (1.4”). As of midnight, June 14th, the SWE total at Flattop was 16.7”, which represents 74% of the 34-year average of 25.3”. Snow depth was 33”. Despite the below-average SWE total, conditions at the station are well within the range of the past 34 water years. The maximum SWE recorded on this day (Julian Day 165) was 48.6” in 1974; the minimum was 0” in 1987, a 48.6” range. All told, there have been 12 years in the 34-year history of the station in which SWE total on this date was lower. The SWE total on this date last year was 22.6”, with a snow depth of 44.5”. A graph comparing SWE for significant water years is below.

A more detailed. PDF version of this graph is available for printing or viewing at SWE Comparison Graph.

In its June10th El Nino/ENSO Discussion, the National Weather Service Climate Prediction Center (CPC) reported that oceanic and atmospheric conditions In May throughout the Pacific Basin were typical of neutral phase of the El Nino Oscillation. The CPC reported that Sea Surface Temperature (SST) anomalies have been warmer than average in the western and central equatorial Pacific and cooler than average in the eastern equatorial Pacific. These anomalies have led to precipitation and wind speed gradients across the equatorial Pacific, with precipitation and wind speeds generally higher in the western portion of the equatorial Pacific. Forecasts of conditions for the next 3 months are split, with some models predicting neutral conditions and others El Nino conditions. The CPC concluded that ENSO-neutral conditions are more likely through August 2004, while noting that trends after that are uncertain.

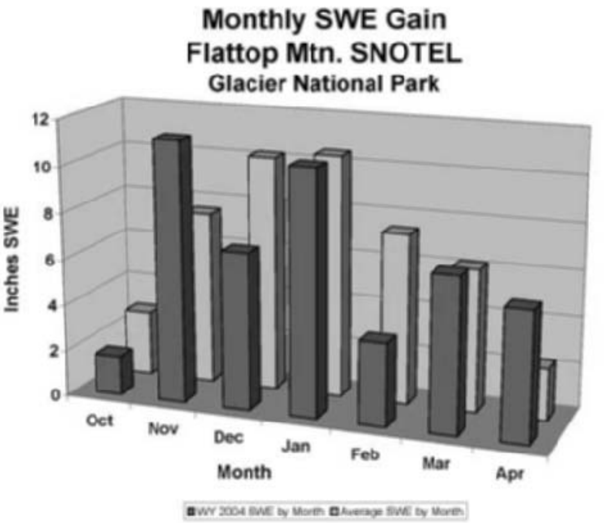

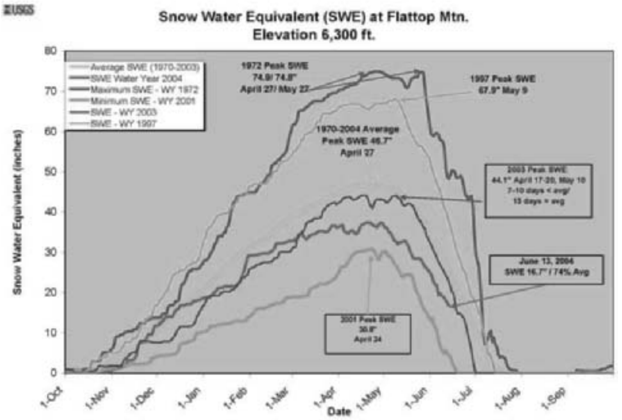

For more information, see the climate prediction center's El Nino/ENSO Discussion.

Also available for viewing are graphs for Water Year 2003, Water Year 2002, Water Year 2001, Water Year 2000, and Water Year 1999.

For details on how snowpack affects the spring opening of Logan Pass, visit our **Spring Opening of the Going-to-the-Sun Road** web page (Logan Pass is located at 6646 feet elevation along the upper edge of the McDonald drainage, approximately 10 miles southeast of the Flattop Mountain SNOTEL station).

- For the most recent SNOTEL data, graphs, and maps of snow distribution visit the Montana NRCS - Snow, Water, and Climate Services Snow Products web page—they provide *snow depth!* data for the Flattop Mountain SNOTEL station under the Daily Products-Reports heading. Also, their detailed map of Montana SNOTEL sites is a great resource.
- The NRCS, National Water & Climate Center provides access to all SNOTEL data as well as in-depth publications describing the SNOTEL program.
- Avalanche forecasts for Flathead National Forest, Kootenai National Forest, & Glacier National Park can be found at the Glacier Country Avalanche Center, which also links to avalanche advisories for areas throughout the western US & Canada.
- For current & historic river level data, visit the USGS - Montana Current Streamflow Conditions web page.

SNOW AND SNOWMELT

Pratap Singh
National Institute of Hydrology
Roorkee, India

Snow is a solid form of precipitation. The formation of snow crystals inside a cloud depends on the presence of ice nuclei, water vapor, motion of the cloud, and the temperature of the cloud. Snow crystals are formed when the cloud temperature is below freezing. These crystals experience diffusional, accretional, and aggregational growth inside the cloud. At temperatures greater than about −5 °C, crystals usually stick together, and aggregation of ice crystals forms snowflakes. Snowflakes obtain their maximum size when the temperature of a cloud is near 0 °C. Aggregation of ice crystals ceases at temperatures below −20 °C. Due to changes in the ambient condition of the cloud, the size and form of the crystals change when they descend through a cloud. In a typical cloud, a snowflake of 1 mm diameter can grow to 10 mm in about 20 minutes. Because of the irregular shapes of snowflakes, measurement of their linear dimensions is difficult. The maximum diameter of snowflakes may range from 0.1 mm to several centimeters.

Snow falls from the clouds mainly in the form of branched hexagonal crystals or stars. Other observed shapes include hexagonal rods, needles, and plates. Snowflakes have a large surface area and fall slowly compared with raindrops. Therefore, snowflakes scavenge more atmospheric aerosols than rain. Precipitation as snow occurs when the atmospheric temperature is near 0 °C or less. Compared to rainfall, snowfall occurs more uniformly, but its distribution on the ground is highly influenced by wind. To obtain the water equivalent of snow, the actual depth of snow unaffected by wind and its density are needed.

Accumulation of snow on the ground leads to the development of a snowpack, which stores water in the form of snow. Depending upon climatic conditions, snow on the ground can stay for a period as long as several months. The global distribution of snow shows that a major portion of the Northern Hemisphere is covered by snow during winter. Snow has great scientific interest because of its practical importance and utility in many fields, especially, in water resources, climate, and winter sports. For many countries, such as the United States, Canada, India, China, Pakistan, Afghanistan, Russia, Nepal, and European countries, snowmelt runoff is a vital source of water for drinking, irrigation, and hydropower generation. Several important rivers, such as the Columbia River in the United States and Canada, the Rhine River in Europe, and the Indus River in Asia, get substantial contributions from snowmelt runoff. Melting snow provides more than 70% of the water supply for the western United States (1).

In general, snow starts melting in the spring when the temperature begins rising. The depth and extent of the snowpack is reduced due to snow melting, and water is released from the snowpack. Snowmelt runoff estimates are needed for forecasting seasonal water yields, river regulation, reservoir operation, determination of design floods, design of hydrologic and hydraulic structures, and planning flood control programs. The snowmelt is estimated either using the energy balance approach or the temperature index method. The energy balance method requires information on radiation energy, sensible and latent heat, energy transferred through rainfall onto the snow, and heat conduction from the ground to the snowpack. When all the components for an energy balance computation are known, the melt rate can be expressed as (2)

$$M = \frac{1000\,Q_m}{\rho_w L \beta} \tag{1}$$

where M is the depth of melt water (mm d^{-1}), Q_m is the net energy flux available for melting (kJm^{-2} d^{-1}), L is the latent heat of fusion (333.5 kJ kg^{-1}), ρ_w is the density of water (1000 kg m^{-3}), and β is the thermal quality of the snow. The thermal quality of snow depends on the amount of free water content (generally 3–5%) and the temperature of the snowpack. Assuming $\beta = 0.97$ for a thermally ripened snowpack, Eq. 1 reduces to

$$M = 0.0031\,Q_m \tag{2}$$

Snowmelt follows an almost diurnal pattern of energy availability, and thus, it is possible to determine the diurnal variation in snowmelt by applying the diurnal distribution of energy received on the snow surface. The albedo, the most important parameter controlling the absorption of solar radiation, is very high for snow. For fresh snow, the albedo is about 0.80–0.90, suggesting that most of the short-wave radiation is reflected back to the atmosphere from the snow surface.

In the absence of detailed energy data for computing snowmelt, the temperature index method is considered the best substitute for the energy balance and is widely used. The air temperature, expressed as degree-days, is used for snowmelt computation. Temperatures are the most readily available data; therefore, the temperature index method is extensively used. The most common expression for estimating snowmelt using temperature is given as

$$M = D(T_i - T_b) \tag{3}$$

where M is the depth of melt water (mm d^{-1}) produced in unit time, D is the degree-day factor (mm °C^{-1} d^{-1}), T_i is the index air temperature (°C), and T_b is the base temperature (usually, 0 °C). Daily mean temperature is most commonly used as an index of temperature for snowmelt. The degree-day factor is used to convert the degree-days to the depth of snowmelt. A wide range of degree-day factors (0.7–9.2 mm °C^{-1} d^{-1}) has been reported in the literature. However, the majority of the reported values of the degree-day factor range from 3–5 mm °C^{-1} d^{-1} (2). The degree-day factor is influenced by the physical properties of the snow and, therefore, changes with time. Usually, temperature data used for melt estimation are available at a few locations in a basin. These available temperatures can be interpolated/extrapolated to different altitudes using the temperature lapse rate (usually, 0.65 °C/100 m).

Snow melts on the upper surface of the snowpack exposed to the atmosphere. In some cases, a little melting is also possible from the lower surface of the snowpack due to ground heat transfer. Snow is a porous medium, and melt water generated on the surface percolates through the snowpack before appearing as runoff. In the beginning of the melt season, the upper part of the snowpack may be warm (0 °C), but the lower part may be cold (subfreezing). Under such conditions, melt water produced on the surface may freeze in the snowpack to raise the temperature to 0°C. Unless the whole snowpack is isothermal at 0°C, no melt runoff is produced.

BIBLIOGRAPHY

1. Chang, A.T.C., Foster, J.L., Gloersen, P., Campbell, W.J., Josberger, E.G., Rango A., and Danes, Z.F. (1987). Estimating snowpack parameters in the Colorado River basin. In: *Proc. Large Scale Effects of Seasonal Snow Cover*. IAHS Publ. No. 166, pp. 343–353.
2. Singh, P. and Singh, V.P. (2001). *Snow and Glacier Hydrology*. Kluwer Academic Publishers, Dordrecht, The Netherlands, p. 221.

SNOW SURVEYS

Pratap Singh
National Institute of Hydrology
Roorkee, India

The basic objective of conducting snow surveys is assessing the spatial distribution of the snow water equivalent accumulated in the basin at a specific time. The distribution of snow is affected by several factors, including topography, orientation, vegetation, and wind speed. During strong winds, the drifting and blowing of snow create a highly uneven spatial distribution of snow depth. Therefore, to account for uneven distribution of snow in a basin, snow surveys are needed. Snow surveys are primarily a field activity that include determining the depth and vertically integrated density. An estimate of the snow water equivalent at a site is computed by multiplying the snow depth by the density of the accumulated snow. Snow surveys are made at regular intervals during both accumulation and ablation periods. Snow surveys conducted just after winter are considered especially important because they provide information about the total snow water equivalent accumulated during the winter. Such information helps in assessing the water resource potential of the deposited snow and is used for forecasting the volume of water expected from snowmelt.

The distribution of the snow water equivalent is needed as input to snowmelt models used for predicting runoff. In addition to hydrologic applications, agricultural, transportation, and recreational activities also require such data (1).

To carry out snow surveys, first, a snow course is established in the basin. The snow course represents a preselected line of marked sampling points in the basin. It is selected by considering representative sites to provide a reliable index of the water stored as snow across the entire basin. Based on the topographical features of the basin, the length of the snow course and a suitable number of observations are chosen for snow surveying. Once the snow course is established in the basin, measurements are repeated every year on the same snow course. Each sampling point of a snow course is located by measuring its distance from a reference point marked on a map. In mountainous regions, snow is distributed much more unevenly and therefore, the length of a snow course is shorter, say, only a few hundred meters, compared to flat regions. Moreover, sampling points are closely located in mountainous areas. The average of the snow water equivalent at all sampling points on the snow course gives information on the availability of water in the basin in the form of snow. Depending on the area covered by surveying, surveys can be made by surveyors using snowshoes, skis, or snow vehicles.

The most common instruments used for snow surveying include a graduated tube that has a snow cutter fixed at its lower end and a balance for determining the weight of the snow sample. The extracted core is retained in the sampler, and the sampler plus core is weighed. The weight of the empty sampler is predetermined. Therefore, the difference in weight with and without the core gives the value of the snow water equivalent of the core. Some instruments, such as the spring balance, give a direct value of the snow water equivalent of a sample when it is weighed. Proper documentation of records is made after sampling the site. In addition to the depth, density, and snow water equivalent for each sampling site, the snow course number and details of site (latitude, longitude, and elevation), weather conditions, vegetation, etc., are also recorded. Such data help in interpreting data collected. The date and timing of the beginning and end of the snow surveys and name of the snow surveyor are also recorded. A complete set of information should be prepared in the desired format just after finishing snow surveys.

An alternative method used for surveying the snow is based on the detection of natural gamma radiation emitted from the ground surface. The gamma radiation flux near the ground originates primarily from natural radioisotopes in the soil. Potassium, uranium, and thallium in the soil are the main sources of radiation. Typically, more than 95% of the gamma radiation is emitted from the top 20 cm of soil. Snow on the ground attenuates the emission of natural gamma radiation from the earth's surface, and the magnitude of attenuation provides an estimate of the snow water equivalent. A greater water equivalent of the snow on the ground will provide a higher attenuation of radiation. Therefore, presnow or no-snow measurements of gamma radiation level are required along the traverse to be followed for surveying. This approach can be used for a ground survey as well as for an aerial survey by aircraft. In a ground survey, a hand carried detector provides a means for measuring the averaged water equivalent along the length of the snow course. The equipment includes a portable gamma-ray spectrometer that uses a small scintillation crystal to measure the rays across a wide spectrum. Recently, radars are also being used for estimating snow water equivalent in basins. Such a radar is installed on a small carrier attached to a snow scooter, that is driven over the snow following the snow course.

In aerial natural gamma radiation surveys, a gamma-ray sensor and a recording device are installed in a low-flying aircraft. Flying of aircraft along the predetermined line is very essential in this type of survey because a relatively small deviation from the flight line may introduce errors due to radiation from other types of rocks. Thus, the flight line must be selected very carefully, so that the aircraft can follow it from year to year. Such aerial surveys may be made at altitudes up to 300 m but are most accurate at the lowest altitude possible for the terrain because of significant atmospheric attenuation of the radiation. This approach is considered the most reliable for level terrain but can also be used for hilly areas where elevation differences up to 400 m can be surveyed. There are some advantages and disadvantages of this approach. The disadvantage is that the aerial survey can be used only to measure the water equivalent of 10 to 100 mm of show (2). The advantage of natural gamma radiation method is that manual observations of the depth, density, or snow water equivalent are not required and a larger area can be surveyed, in a short time.

BIBLIOGRAPHY

1. Singh, P. and Singh, V.P. (2001). *Snow and Glacier Hydrology*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
2. World Meteorological Organisation. (1981). *Guide to Hydrological Practices. Vol. I, Data Acquisition and Processing.* WMO No. 168.

A STATISTICAL APPROACH TO CRITICAL STORM PERIOD ANALYSIS

F. Ruiz
M. Abad
A. Rodríguez-Ramírez
L.M. Cáceres
J. Rodríguez Vidal
Universidad de Huelva. Avda.
de las Fuerzas Armadas
Huelva

R. Pino
J.M. Muñoz
Universidad de Sevilla. Avda.
Reina Mercedes
Sevilla

INTRODUCTION

Periodicity analysis is a natural and integral part of meteorological studies on different timescales. Numerous studies have focused on the possible cyclicity of several meteorologic data and/or phenomena, such as wind velocities, atmospheric pressures, sunspots, hurricanes, and currents (1–4).

In numerous coastal areas, winter storms are common meteorologic phenomena characterized by wind measuring 10 or higher on the Beaufort scale (89 km per hour, 55 miles per hour). Storm periods are a significant problem for coastal managers; they cause considerable property damage to beach–nearshore profiles with tourist-economic uses, fisheries, harbors, private vacation homes, and public earthworks (5). Their effects may vary over successive sectors of the same coast with different hydrodynamic features (6,7). These events increase coastal erosion and move sands rapidly offshore, whereas lower energetic conditions may cause gradual beach accretion (8). Consequently, managers need tools (i.e., event periodicity, sediment transport, modeling approach) to evaluate the possible consequences (9–11).

In this article, we analyze some mathematical procedures used in time series analysis. These robust methods are applied to calculate an approach to critical storm periodicity on the southwestern Spanish coast.

THE STATISTICAL BASIS FOR CRITICAL STORM PERIOD ANALYSIS

Introduction: The Predicting Nature

Statistical techniques permit us to analyze the structure and behavior of natural phenomena, their relationships, and temporal evolution. In this last case, high variability levels appear in the time series analysis, and consequently some statistical approaches may be useful to study their periodicity. The following step is prediction, but sometimes it is not possible to reach this objective because of great variability in the variables that describe the nature of a reliable prediction or insufficient levels of correlation between both explanatory variables and objective variables.

Even though adequate levels of prediction cannot be reached, statistical methodology allows us to obtain useful conclusions on the natural phenomena under study. The analysis of a “storm period” is a particular case of this type of study. Prediction is difficult, but it is possible to reach conclusions about past behavior. To this end, time series and spectral analysis are especially useful.

Spectral Analysis

Spectral analysis is a set of statistical and mathematical methods for identifying hidden periodicities in a time series. Moreover, spectral analysis is concerned with estimating the spectrum over the whole range of frequencies (or their inverses, the time periods). These techniques are widely used in electrical engineering, physics, computer science, geology, and environmental sciences (12–15).

From a mathematical viewpoint, spectral analysis is a modification of Fourier analysis to suit it for stochastic rather than deterministic functions of time. Fourier analysis is oriented to approximate a function by a sum of sine and cosine terms that is called the Fourier series representation. We distinguish two classes of spectral techniques, univariate and bivariate.

Univariate Spectral Analysis. The Wiener–Khintchine theorem states that the variance of any stationary stochastic process can be expressed as the integral in the interval (0, p) of a function called the spectral density function (or spectrum); so a spike of this function identifies a frequency with an important contribution to the variance, and a periodicity is associated with that frequency. Basically, univariate spectral analysis computes an estimate of the spectrum from the Fourier series representation of the available time series, and a graphical representation, called a periodogram (Fig. 1), is used to identify possible spikes in the underlying spectral density function (16).

This method has been used to calculate the periodicity of the heart rate and locomotor activity of several organisms (17), orbital cyclicities in geologic sequences (18), and seismology (19). In meteorology, it may be applied to the relationships among wind and waves (20) and analysis of the periodicity and persistence of daily rainfalls (21).

Bivariate Spectral Analysis. When we try to identify periodic movements common to a pair of time series, we need bivariate spectral analysis techniques. The definition of a cross-spectrum lets us generalize the univariate spectrum of this situation, although its interpretation is not so clear. Several statistical measures can be computed from this cross-spectrum, but some controversy is associated with the selection of the one most appropriate. We can remark on one of these measures for its easy interpretation: coherence. It measures the square of the linear correlation between two time series at a given frequency. It takes values in [0,1], and the closer it is to one, the more closely related are the two process at that frequency, so identifying a common periodic term.

In meteorology, these techniques have been used in investigating tropospheric ozone formation and decomposition processes (22), correlation of sea levels and atmospheric variables (23), analysis of the interactions between biosphere and atmosphere (24), and monitoring present-day climatic conditions (25).

A CASE STUDY: THE RECENT STORM RECORD ON THE SOUTHWESTERN SPANISH COAST

Study Area

The Huelva coast (SW Spain) is composed of large sandy beaches (145 km long), interrupted only by estuarine mouths (Fig. 2: Guadiana, Piedras, Tinto-Odiel, and Guadalquivir). The littoral morphology of this area is linked mainly to a mesotidal regime (2 m mean range), medium wave energy, and the fluvial sedimentary inputs of the Guadiana (144 m^3/s) and Guadalquivir (185 m^3/s) Rivers, a littoral net sediment flow oriented toward the east (26–29).

This coast is heavily “defended” by modern protective groins, causing either total (Vila Real do Santo Antonio,

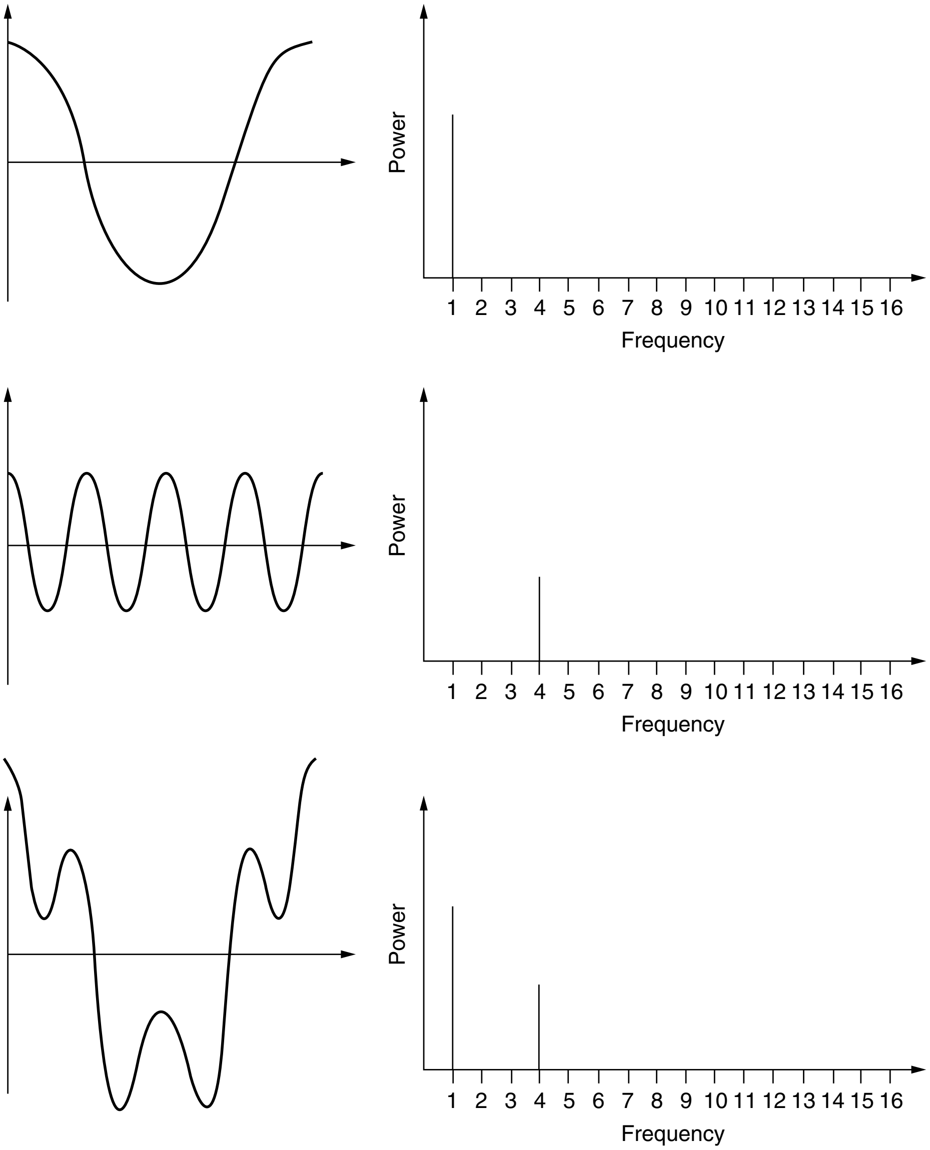

Figure 1. Fourier power spectra (right) resulting from various idealized data sets (left). Modified from Reference 15.

Huelva) or partial (Isla Cristina, Punta Umbria) obstacles to sedimentary transition. Consequently, two sedimentary units (Ayamonte-Punta Umbría and Mazagón-Doñana) may be delimited in this human-altered area, directly related to the effects of the main groins (30).

Each unit includes two different zones: (1) beaches affected by erosive processes (Ayamonte, Isla Cristina, La Antilla, Punta Umbría, Mazagón, Matalascañas) and (2) prograding spits (El Rompido, Doñana). These beaches constitute an important tourist area and are visited by more than 1 million people between May and September. This "sun and beach" tourism needs continuous management during the winter months due to periodic storm damage. The causes of this periodicity have not been studied but may be initially related to climate variables, such as the North Atlantic Oscillation Index (NAO; 31). The NAO is an index that represents the differences in atmospheric pressure at sea level between the Azores and Iceland (32). Positive values are associated with low cyclone activity in southern Europe and conversely (33).

The Winter Storm Record

The recent record (1956–1996) of critical storm periods was obtained from two daily journals (Odiel and ABC), wave heights were recovered from the Huelva harbour archives, and the wind speeds were supplied by the Spanish National Institute of Meteorology at the Huelva station (period 1961–1996). Data were normalized to standard scales (Fig. 3).

Between 1956 and 1996, eight winter storm periods on the southwestern Spanish coast were identified (Fig. 3, C): 1962–1963, 1969–1970, 1972–1973, 1978–1979, 1981–1982, 1987–1988, 1989–1990, and 1995–1996. In most cases, these high-energy periods are concentrated in December and January. An additional periodicity may be inferred in the fair weather conditions, the intervals separating two consecutive storm periods (5). After the

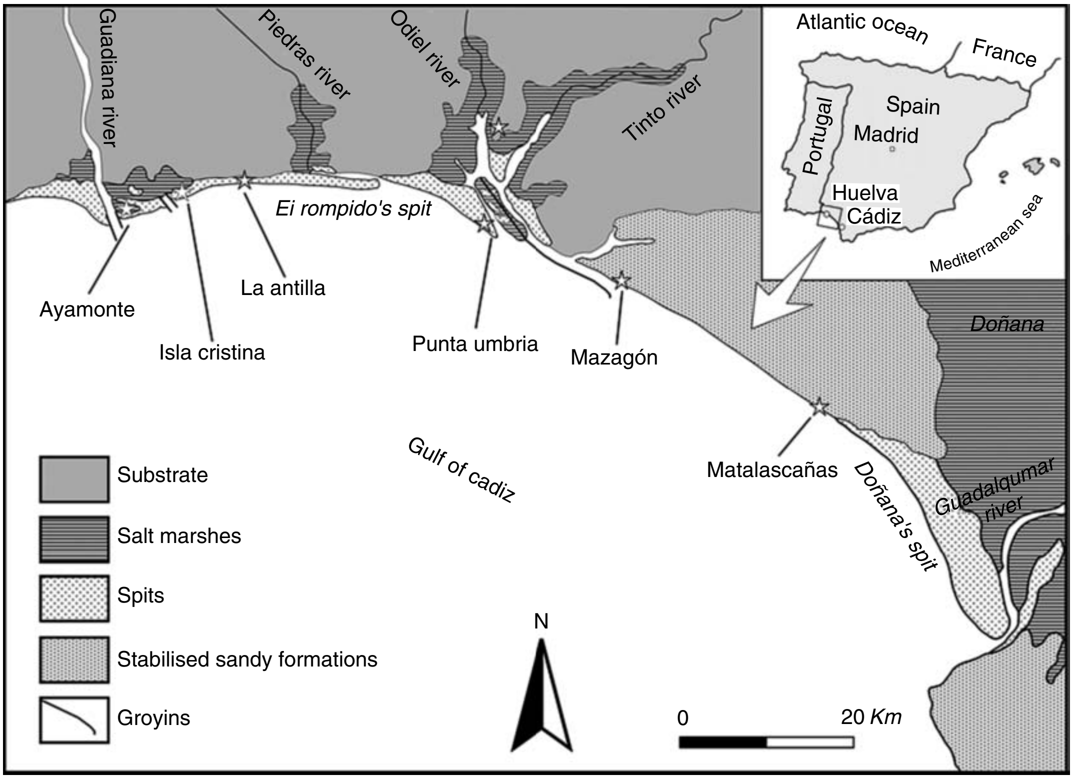

Figure 2. Location map, including the main tourist beaches, spits and groins. 1: Plio-Pleistocene substrate; 2: salt marshes; 3: spits; 4: stabilized sandy formations. Main groins. a: Vila Real do Santo Antonio; b: Huelva.

first storm period (1962–1963), a long interval (6–7 years) was found until the next period, followed by a shorter cycle (2–3 years).

An analysis of the storm features indicates a direct relation between the direction and speed of the prevailing winds and the causes of these events. In all cases, these winds come from the third quadrant at high to very high speed (Fig. 3: 8–12) on the Beaufort scale. Consequently, a more detailed study of the third quadrant velocities is needed to establish the periodicity of the storm wind.

Meteorologic Variables and Storm Periodicity

The application of the Fourier transformation to the time series of autumn–winter wind speeds permits us to distinguish two more probabilistic frequencies (Fig. 4a: 6–7 years and 9–10 years), which are also found in the storm series. In most cases, a 9-year poststorm period contains two storms (period 2 and period 3), situated at 6–7 and 9–10 years after period 1, respectively. This initial coincidence would indicate a possible statistical interrelationship between both variables, but the small record of recent storm periods permits only a first approximation.

This hypothesis was tested in the period 1996–2000 by analyzing the new record of storm periods. A high-energy period (1998–1999) happened in agreement with the hypothetical storm sequence, which predicts new storms in 1995–1996 (period 2) and 1997–1998 (period 3) after 1989 (period 1). In this hypothetical 9-year sequence, the next storm period will be between 2004 and 2005, although a new positive date does not totally confirm the initial hypothesis. We will need 20–30 storm periods (a century, at least) for an adequate, statistical test of this correlation.

This possible periodicity may be related to cyclical changes of the cyclone regime, which are expressed in NAO values (Fig. 4a). In the analysis of the wind–NAO spectral series, the 3-year frequency shows the higher probabilities (Fig. 4b), indicating some degree of common behavior between these two variables.

This frequency agrees with the two main frequencies (6 and 9–10 years) obtained for the wind series. Moreover, an additional coincidence was found by contrasting the evolution of the NAO values and the storm record. Except for the 1989–1990 winter, the remaining storm periods were characterized by negative values of the NAO index, indicating a possible relation between storms and high cyclone phases. In the future, the NAO values collected during storm periods may even be positive due to the increasing tendency of this variable in the last decades.

ADDITIONAL MANAGEMENT IMPLICATIONS

A continuous, yearly study in 20–30 years will be interesting to contrast the first approach now presented to estimate the periodicity of critical storm periods

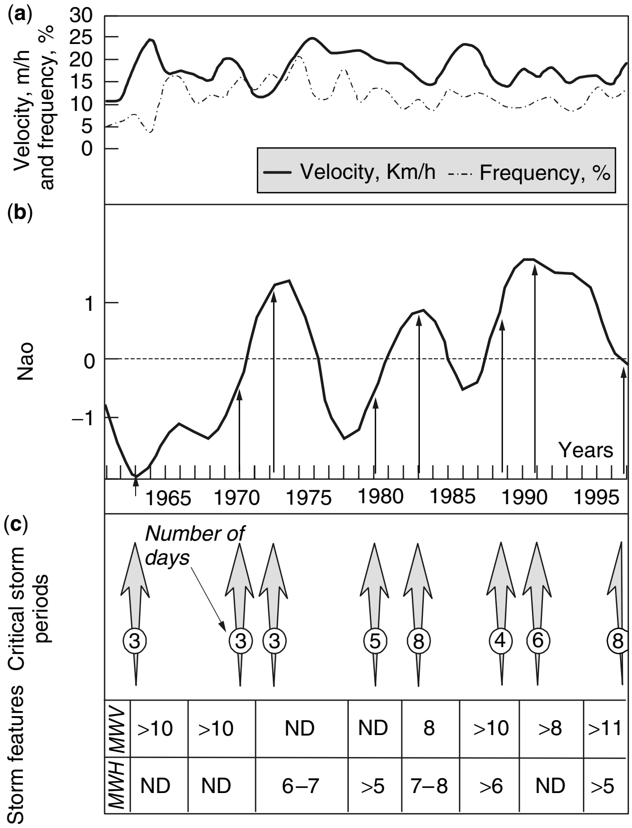

Figure 3. (**a**) Velocity and frequency of the third quadrant winds (period 1961–1996); (**b**) NAO record; (**c**) Critical storm periods between 1960 and 1996, with the maximum wind velocities (MWV; Beaufort scale) and maximum wave heights (MWH, in meters). ND: No data.

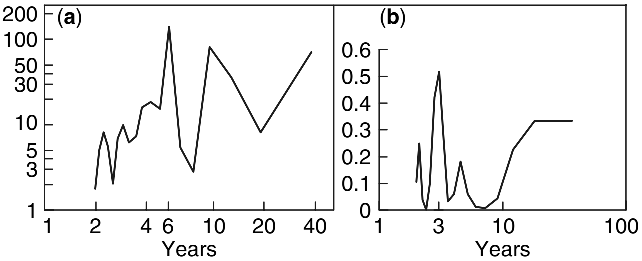

Figure 4. Periodicity diagram of the third quadrant wind speeds; Coherence values between NAO and wind speeds (X axis on log scale).

and its consequence more rigorously. Some additional investigations are necessary:

1. A short-term, periodic analysis of beach–nearshore profiles at the main beaches, with special attention to significant changes during these storms.
2. A first approach to long-term coastal development in examining the roles of waves, tides, storms, and sediments in coastal dynamics, as pointed out by Carter (34).
3. An evaluation of the sedimentary effects produced by the two main groins of the sector (Vila Real do Santo Antonio and Huelva), with indications of possible new emplacements and/or changes in their characteristics (mainly length and alignment).
4. Coastal strategies, including future redistribution of the littoral space (protection zones, demolition of shorefront residences, and relocation of both public infrastructures and private homes).

Taking into account the total damage occasioned by storms in the 1989–1999 period (probably up to $ 20 million), a moderate inversion ($2 million/year) may be very profitable in the next decade.

CONCLUSIONS

In the last 40 years (1956–1996), the analysis of storm periods suffered by the Huelva littoral (SW Spain) indicates a concentration in the winter months (December–January) and statistical behavior similar to third quadrant wind speeds. After the first storm period registered (1962–1963), two new high-energy periods were found in the following 10-year intervals: period 2, at 6 years and period 3, at 9–10 years. The first test of this hypothesis was positive in the 1997–2000 period.

This first approach holds out the prospect of a possible statistical prediction of critical storm periodicity in the next decades, in connection with new investigations into meteorologic changes, sedimentological studies, and coastal management strategies in this area.

Acknowledgments
The financial support of the DGYCIT (projects AMB99-0226-C03-03 and BTE2000-1153) and the Almonte Corporation are gratefully acknowledged. This article is a contribution to IGCP 396 and 437.

BIBLIOGRAPHY

1. Gazis, P.R. (1996). Long-term enhancements in solar wind speed. *J. Geophys. Res.* **101**: 415.
2. Elsner, J.B., Kara, A.B., and Owens, M.A. (1999). Fluctuations in North Atlantic hurricane frequency. *J. Climate* **12**: 427–437.
3. Rodbell, D.T. et al. (1999). An approximately 15,000 year record of El Niño-driven alluviation in southwestern Ecuador. *Science* **283**: 516–520.
4. Artigas et al. (2004).
5. Lee, G., Nicholls, R.J., and Birkemeier, W. (1998). Storm-driven variability of the beach-nearshore profile at Duck, North Carolina, USA, 1981–1991. *Mar. Geol.* **148**: 163–177.
6. Ballesta, M., Morales, J.A., and Acosta, E.A. (1998). Efecto erosivo de los temporales del invierno 1995–1996 sobre la playa de Mazagón (Huelva, SO de España). *Rev. Soc. Geol. España* **11**: 285–296.
7. Rodríguez-Ramírez, A., Cáceres, L.M., Rodríguez Vidal, J., and Cantano, M. (2000). Relación entre clima y génesis de crestas/surcos de playa en los últimos cuarenta años (Huelva, Golfo de Cádiz). *Rev. Cuat. Geomorf.* **14**: 109–113.
8. Komar, P.D. (1976). *Beach Processes and Sedimentation*. Prentice-Hall, Upper Saddle River, NJ.

9. Ahyerre, M., Chebbo, G., Tassin, B., and Gaume, E. (1998). Storm water quality modelling, an ambitious objective?. *Water Sci. Technol.* **37**: 205–213.
10. Gönnert, G. (1999). The analysis of storm surge climate change along the German coast during the 20th century. *Quat. Int.* **56**: 115–121.
11. Williams, J.J. and Rose, C.P. (2001). Measured and predicted rates of sediment transport in storm conditions. *Mar. Geol.* **179**: 121–133.
12. Jenkins, G.M. and Watts, D.G. (1968). *Spectral Analysis and Its Applications*. Holden-Day, London.
13. Cohn, B.P. and Robinson, J.E. (1976). A forecast model for Great Lakes water levels. *J. Geol.* **84**: 455–465.
14. Chatfield, C. (1989). *The Analysis of Time Series: An Introduction*. Chapman and Hall, London.
15. Swan, A.R.H. and Sandilands, M. (1995). *Introduction to Geological Data Analysis*. Blackwell Science, Malden, MA.
16. Kokoszka, P. and Mikosch, T. (2000). The periodogram at Fourier frequencies. *Stochastic Processes Their Appl.* **86**: 49–79.
17. Bojsen, B.H., Witthøfft, H., Styrishave, B., and Andersen, O. (1998). *In situ* studies on heart rate and locomotor activity in the freshwater crayfish, *Astacus astacus* (L.) in relation to natural fluctuations in temperature and light intensity. *Freshwater Biol.* **39**: 455–465.
18. Scarparo Cunha, A.A. and Koutsoukos, E.A.M. (2001). Orbital cyclicity in a Turonian sequence of the Cotinguiba Formation, Sergipe Basin, NE Brazil. *Cretaceous Res.* **22**: 529–548.
19. Baisch, S. and Bokelmann, G.H.R. (1999). Spectral analysis with incomplete time series: An example from seismology. *Comput. Geosciences* **25**: 739–750.
20. Beard, A.G., Williams, P.J.S., Mitchell, N.J., and Muller, H.G. (2001). A spectral climatology of planetary waves and tidal variability. *J. Atmos. Solar-Terrestrial Phys.* **63**: 801–811.
21. Kottegoda, N.T., Natale, L., and Raiteri, E. (2004). Some considerations of periodicity and persistence in daily rainfalls. *J. Hydrol.* **296**: 23–37.
22. Sebald, L., Treffeisen, R., Reimer, E., and Hies, T. (2000). Spectral analysis of air pollutants. Part 2: Ozone time series. *Atmos. Environ.* **34**: 3503–3509.
23. Fenoglio-Marc, L. (2001). Analysis and representation of regional sea-level variability from altimetry and atmospheric-oceanic data. *Geophys. J. Int.* **145**: 1–18.
24. Baldocchi, D., Falge, E., and Wilson, K. (2001). A spectral analysis of biosphere-atmosphere trace gas flux densities and meteorological variables across hour to multi-year time scales. *Agric. Forest Meteorol.* **107**: 1–27.
25. Sondag, F. et al. (2003). Monitoring present day climatic conditions in tropical caves using an Environmental Data Adquisition System (EDAS). *J. Hydrol.* **273**: 103–118.
26. Vanney, J.R. (1970). *L'hydrologie du Bas Guadalquivir. Serv. Publ. Dpto. Geograf. Aplic.*, CSIC, Madrid.
27. CEEPYC. (1979). *Centro de Estudios y Experimentación de Puertos y Costas "Ramón Iribarren"*. Plan de estudio de la dinámica litoral de la Provincia de Huelva. Informe Dirección General de Puertos y Costas, Madrid.
28. Borrego, J. and Pendón, J.G. (1989). Influencia de la actividad humana sobre la evolución sedimentaria de un sector de la ría de Huelva (España). *Geolis* **3**: 125–131.
29. Cuena, G.J. (1991). *Proyecto de regeneración de las playas de Isla Cristina*. Informe del Servicio de Costas, Ministerio de Obras Públicas y Turismo.
30. Ojeda, J. (1989). La dinámica litoral reciente de la Costa Occidental de Andalucía. In: F. Díaz del Olmo and J. Rodríguez Vidal (Eds.). *El Cuaternario en Andalucía Occidental.*; *AEQUA Monografías* **1**: 123–132.
31. Hurrell, J.W. (1996). Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature. *Geophys. Res. Lett.* **23**: 665–668.
32. Kutzbach, J.E. (1970). Large-scale features of monthly mean Northern Hemisphere anomaly maps of sea-level pressure. *Mon. Weather Rev.* **98**: 708–716.
33. Rodwell, M.J., Rowell, D.P., and Folland, C.K. (1999). Ocean forcing of the wintertime North Atlantic Oscillation and European climate. *Nature* **398**: 320–323.
34. Carter, R.W.G. (1988). *Coastal Environments*. Academic Press, London.

SUBLIMATION

Edinara Adelaide Boss
Artur-Nogueira-SP, Brazil

CONCEPT

Sublimation occurs when the matter passes directly from the solid phase to the vapor phase, as is shown in Fig. 1. Examples include dry ice, how a frost-free refrigerator works, and naphthalene.

Sublimation is an endothermic process:

$$\Delta H_{\text{sublimation}} > 0$$

The heat required to change the phase of a pure substance is given by

$$Q = m \cdot L$$

where m is the mass of the pure substance and L is the latent heat.

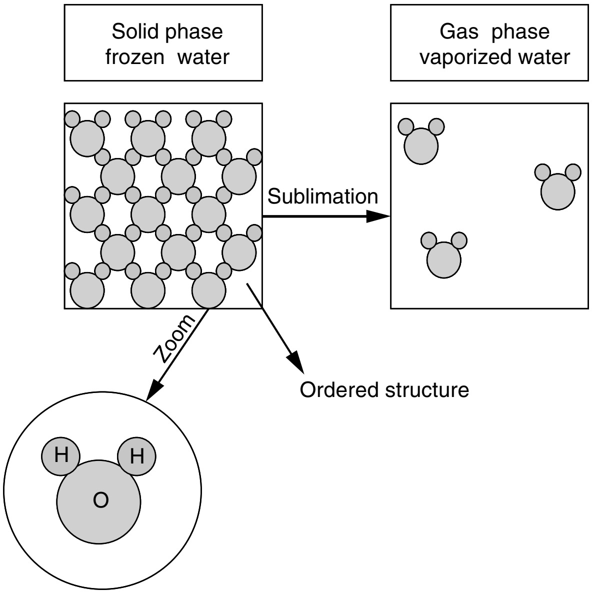

Figure 1. Water sublimation.

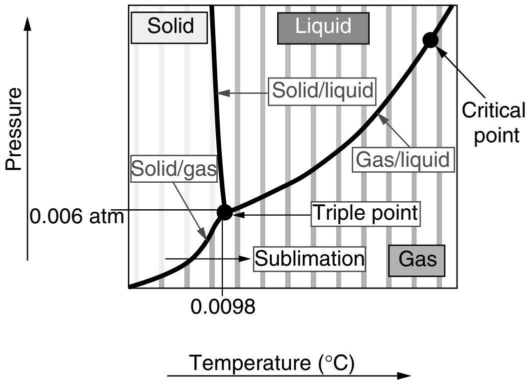

Figure 2. Phase diagram of pure water.

The water latent heat for sublimation is 2.83×10^6 J/Kg.

Figure 2 presents the water phase diagram, which usually shows the pressure versus the temperature.

The diagram is divided into three areas, which represent the solid, liquid, and gaseous phases.

Triple point is a particular condition of temperature and pressure, which the three phases for a substance are in equilibrium.

Above the critical point, the liquid and gas phases are indistinguishable. The critical temperature is the minimum temperature for liquefaction of a gas using pressure. The critical pressure is the pressure required for liquefaction. The critical temperature of water is 374 °C, and the water critical pressure is 218 atm.

Sublimation phenomenon happens below the triple point in the water phase diagram. The sublimation curve is the line between the solid and gas phases.

At the triple point of the water phase diagram, the temperature is 0.0098 °C and the pressure is 0.006 atm.

APPLICATION

Freeze-drying is a process by which a water or aqueous substance is removed from a frozen material or frozen solution by sublimation. Figure 3 shows an example of a frozen solution in a freeze-drying tray.

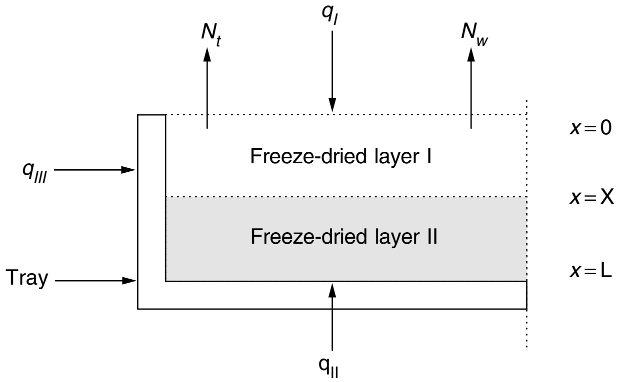

Figure 3. Diagram of a frozen material on a tray during freeze-drying. X is the position of the sublimation interface between the freeze-dried (layer I) and the frozen material (layer II).

Here q_{I}, q_{II}, and q_{III} are the heat, which can be supplied to the surface by conduction, convection, or radiation. N_{w} and N_{t} represent the mass flow of the water vapor and the total mass flow, respectively, in the dried layer.

This process involves the following three periods: freezing, primary drying, and secondary drying periods (1) as shown in Fig. 4.

A batch freeze-dryer is shown in Fig. 5. This equipment is composed of a stainless steel cylinder chamber with eight connectors for connection of the bottles. Above the cylinder chamber, a cover is built by transparent material. At the cylinder, there is a system with three trays. The vacuum pump localizes outside the freeze-dryer.

The freeze-drying process has a wide range of applications, including pharmaceutical products, foodstuffs (the organoleptic properties are important and have to be maintained), and other industrial byproducts (conservation of living microorganism, dehydration, or concentration of heat labile).

Another example of freeze-drying is the soluble coffee process, which produces a porous product by sublimation. This porous material is eager for water.

The performance of the overall freeze-drying process depends significantly on the freezing stage. The material to be processed is cooled to a temperature below the solidification. The shape of the pores, the pore size distribution, and the pore connectivity of the porous network of the dried layer formed by the sublimation of the frozen water during the primary drying stage depend on the ice crystals formed during the freezing stage.

The solvent is removed by sublimation under a vacuum pressure, and heat addition in the primary drying stage is carried out. A significant amount of the latent heat of sublimation is also kept when the water molecules sublime and enter the vapor phase. Because of this result, the temperature of the frozen product is reduced. Then it is necessary to supply heat to the product. The heat could be provided by conduction, convection, and/or radiation.

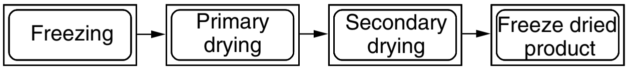

Figure 4. Freeze-drying periods.

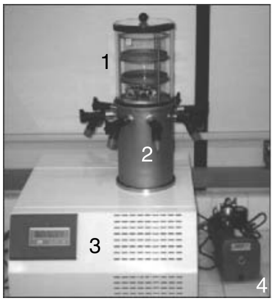

1. Glass camera with trays of inoxidable steel.
2. Manifold-exits of silicon and serpentine interns.
3. System of control of refrigeration and pressure.
4. Bomb of vacuous.

Figure 5. Freeze-drying equipment.

During the secondary drying stage the solvent is removed from the chamber and a small amount of sorbed water can be removed by desorption.

A dynamic mathematical model for the freeze-drying stages was proposed by Boss et al. (2). This model represents the process well.

The low temperature of the freeze-drying process and the low water activity resulting from this process minimize some reactions of degradation such as Maillard's reaction, protein denaturation, and enzymatic reaction (3). Color and flavor are important quality attributes in food, and the freeze-drying process preserves more of the pigments and reduces more of the transport rates than does the conventional drying processes. The hot air drying causes irreversible damages to the texture and incomplete rehydration. The freeze-drying process proposes good preservation of the original product characteristics and an increase in the stability of the product during the storage.

BIBLIOGRAPHY

1. Boss, E.A., Maciel Filho, R., and Toledo, E.C.V. (2004). Real time optimization for freeze-drying process. In: *14th European Symposium on Computer Aided Process Engineering—Escape 14*, Portugal, May.
2. Boss, E.A., Costa, N.A., Maciel Filho, R., and Toledo, E.C.V. (2004). Freeze drying process: real time model and optimization. *Chemical Engineering and Processing*. Available: http://www.sciencedirect.com/science/journal/0255270/.
3. Liapis, A.I., Millman, M.J., and Marchello, J.M. (1985). An analysis of the lyophilization process using a sorption-sublimation model and various operational policies. *AIChE J.* **31**(10): 1594–1604.

TRANSPIRATION

Koichi Miyashita
Okayama University
Kurashiki, Japan

Transpiration is the evaporation of water from the plant to the atmosphere. The moved water speed in a plant by the transpiration depends on a vapor pressure deficit that is the difference of a water vapor pressure between the atmosphere and the leaf surface. Two types of the transpiration exist. One is the stomatal transpiration through stomata, and the other is the cuticular transpiration through a cuticular. When the stomata open, the cuticular transpiration is much less than the stomatal transpiration. However, the cuticular transpiration may become large when the stomata are almost closed or when the leaf is cracked and damaged. Figure 1 shows that the water vapor in a stomatal aperture diffuses to the atmosphere through the stomata and the boundary layer of a leaf, in the case of the open stomata.

The transpiration rate from a leaf (Tr) can be generally expressed as

$$Tr = \frac{c_p \rho}{\gamma L} \frac{e_s - e_a}{r_s + r_a}$$

where C_p is the specific heat capacity, ρ is the density of air, γ is the psychrometer constant, L is the latent heat, e_s is the water vapor pressure of the stomatal aperture, e_a is the water vapor pressure of the atmosphere, r_s is the stomatal resistance, and r_a is the leaf boundary layer resistance that is effected by a leaf size and a shape.

The transpiration rate (Tr) is proportional to the difference ($e_s - e_a$) of a water vapor pressure between the stomata and the atmosphere. Usually, in the case of a plant with the water fully supplied, the water vapor pressure of the stomatal aperture is the saturated water vapor pressure at a leaf temperature. The more atmosphere dries, the more ($e_s - e_a$) become large. Then, the transpiration rate increases.

In general, the plant transpiration has the diurnal pattern. It is great in the daytime. On the other hand, the transpiration is small at the nighttime. In the daytime, the stomatal resistance (r_s) becomes small because of the opening of stomata, depending on photosynthesis activities by the solar radiation. Moreover, the exchange of a water vapor becomes active because of the increase of the wind speed on a leaf surface. Then, the leaf boundary layer resistance (r_a) decreases. Therefore, the transpiration rate (Tr) becomes large for the reason that the transpiration is inversely proportional to the sum total ($r_s + r_a$) of a stomatal resistance (r_s) and a leaf boundary layer resistance (r_a).

Meteorological factors significantly affect the transpiration. Several experimental results have been published about the relationships between the meteorological factors and the transpiration. Transpiration completely depends on the physiological activities as long as stomatal aperture does not change. However, the plant may regulate stomatal opening and transpiration. A temporary reduction of stomatal aperture is often caused by a decrease of light intensity, dry air, water deficit, high temperature, and toxic gases. Haseba (1) showed by a simulation analysis between meteorological factors and physiological activity that the transpiration rate increases with increasing a wind speed under a high air temperature, a low relative humidity, a low solar radiation, and a large stomatal conductance and that decreases in an adverse condition. He also showed that the transpiration may not change between two extreme conditions even if a wind speed increases.

The increase in a solar radiation and an air temperature makes the transpiration increase because the saturated water vapor pressure in the stomatal pore (e_s) increases by rising leaf temperature. On the other hand, the transpiration rate is inhibited as the relative humidity of the ambient air is great and the water vapor deficit decreases.

For example, transpiration is also influenced by deficit of the soil moisture. Transpiration does not change if the soil moisture slightly decreases. However, if the soil water content greatly decreases, to keep the water content plant regulate stomatal aperture decreases transpiration.

Transpiration is necessary for the growth of plants. The flow of water through the plant by transpiration increases the absorption of minerals from the soil. Changes in the water balance and the amount of water available in the soil

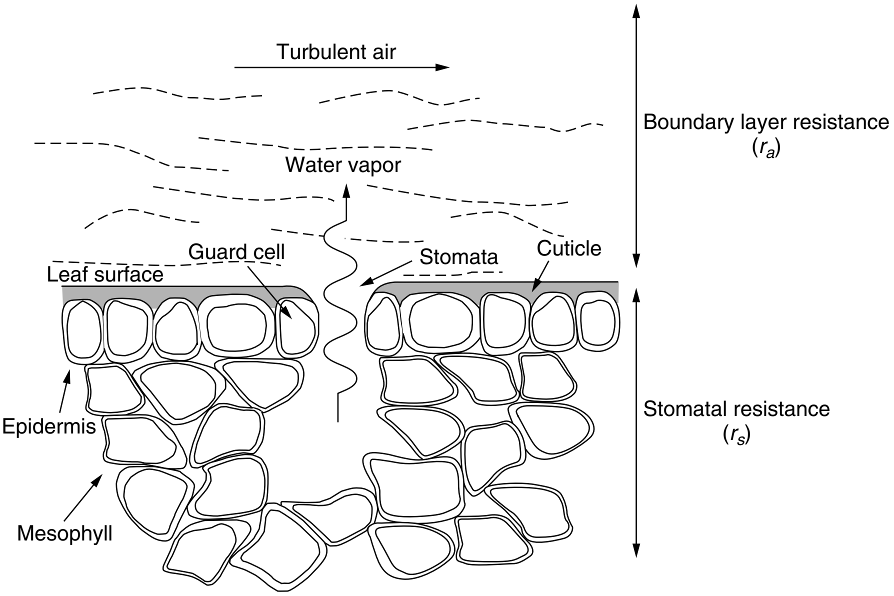

Figure 1. Conceptual diagram of the transpiration from a leaf in case of the open stomata. The curvilinear arrow is the flow of a water vapor from the stomatal aperture to the atmosphere.

can be crucial for plants. There is a rising concern that the worldwide water resources may be exhausted by global warming. Under the drought stresses, a plant does not only decrease the transpiration along with the decrease of stomatal conductance, but also it often may cause damage. Miyashita et al. (2) showed that the recovery level of transpiration in the kidney bean after drought stress gradually decreased as the periods of water stress.

The other effect of transpiration is to decrease the leaf temperature. In particular, the adjustability of a leaf temperature prevents too high of a temperature of a plant from the strong solar radiation. It is well known that the leaf temperatures are different among the plant species even if meteorological conditions are the same, which is mainly the difference of transpiration.

The differences (Ts-Ta) between leaf temperature (Ts) of worldwide barley cultivars (north America cvs, north African cvs, Chinese cvs, Ethiopian cvs, southwest Asiatic cvs, Japanese cvs, Korean cvs, Nepalese cvs, Turkish cvs, and European cvs) and ambient air temperatures (Ta) were measured under the same field (3). The temperature differences (Ts-Ta) of barley cultivars in the habitats with the cool meteorological phenomena such as Europe were large under high-temperature conditions. These show the differences of capability in the transpiration among the geographical varieties. Not only the plant species, but also the relationships between transpiration and the meteorological conditions at the places of origin are recognized.

For the plant canopy, such as a forest or an agricultural land, the transport of water vapor from the canopy to the atmosphere includes both the transpiration from plants and the evaporation from the wetted surfaces of plants and soils. Total of transpiration and evaporation is known and described as evapotranspiration. Transpiration of plant canopy is approximately proportional to the green mass, although the transpiration rate of individual leaves decreases with increasing leaf densities because of shielding of radiation and wind. Therefore, transpiration from the canopy is affected by development of the foliar mass. Canopy transpiration has the seasonal and diurnal patterns. Total annual and daily transpiration of stands of plant in various climatic regions, which are obtained from data of numerous authors, are summarized (4). Community transpiration of tropical rainforest is $1500 \sim 2000$ mm/year, that of grain fields $400 \sim 500$ mm/year and that of open vegetation in arid desert $0.01 \sim 0.4$ mm/day. Under similar climatic conditions, forests with greater mass transpire more than grass lands, which in turn transpire more than heath. In addition to accurate observational data, the detailed simulation models, which include all basic hydrological parameters, are now available in calculation of stand transpiration. These quantitative data about transpiration of plant vegetations are important for forestry and landscape management, irrigation management, and prediction of climatic change.

As high CO_2 in the future decreases stomatal aperture (or stomatal conductance), transpiration may decrease if CO_2 increases and leaf temperature is constant. However, high leaf temperature at higher ambient temperature may partly counteract the expected suppression of transpiration. Despite higher leaf temperature, water-use efficiency (the ratio of photosynthesis to transpiration) increases at high CO_2 mainly because of increased photosynthesis (5).

MEASUREMENT OF TRANSPIRATION

Transpiration in the single leaf, the individual plant, and the canopy level measured by many methods is as follows: (1) The weighing method. The total weight of a plant with a pot is precisely measured at a constant time. The amount of a drop of a weight is comparable with the transpiration. As similar methods, there are the leaf cutting method for a single leaf and the lysimeter method for many plants. (2) The chamber method. The plant is covered by a transparent box with an inlet and an outlet port for the air flow. The humidity of an outlet port increases by the transpiration if an air flows. Transpiration rate can be calculated from the difference of an absolute humidity and an air flow rate. This method can be widely used for a single leaf, an individual plant, and a canopy. (3) The sapflow method. There is a transpiration stream within a plant vessel from a root to a leaf. It is in proportion to the increasing amount of a stem temperature when a stem is warmed by a constant quantity of heat. This method is used for plants (tomato, soybean, sugarcane, sweet potato, and etc) or trees. (4) Micrometeorological method. Evapotranspiration from a canopy is measured by a micrometeorological observation. Also, the Bowen ratio energy balance method, the aerodynamic method, the turbulent transport method, and so on are available.

BIBLIOGRAPHY

1. Haseba, T. (1973). Agro-meteorological studies of transpiration in plant leaf in relation to the environment. *Mem. Coll. Agr., Ehime Univ.* **18**: 1–141.
2. Miyashita, K., Tanakamaru, S., Maitani, T., and Kimura, K. (In press). Recovery responses of photosynthesis, transpiration, and stomatal conductance in kidney bean following drought stress. *Environ. Exp. Bot.*
3. Miyashita, K., Maitani, T., and Takeda, K. (1999). Comparison of surface temperature of barley varieties under high temperature conditions. *J. Agric. Meteorol.* **55**(4): 323–328.
4. Larcher, W. (1995). *Physiological Plant Ecology*, 3rd Edn. Springer, Berlin, pp. 264–275.
5. Imai, K. and Okamoto-Sato, M. (1991). Effects of temperature on Co2 dependence of gas exchanges in C3 and C4 crop plants. *Jpn. J. Crop Sci.* **60**: 139–145.

WATERSPOUT

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

A funnel-shaped weather disturbance which exists over a body of water, a waterspout is often referred to as a tornado on water. This is accurate for one type of waterspout, the aptly named tornadic waterspout. It is somewhat inaccurate with regard to the much more common type of waterspout, the fair weather waterspout.

Tornadic waterspouts are tornadoes over water. These are stronger, larger, and faster than fair weather waterspouts. However, they do tend to form over land as true tornadoes before moving over water. Fair weather waterspouts form differently from tornadoes. Rather than developing downward, they begin at the surface, building up toward the sky.

For waterspouts to form, a series of things must occur. First, a small portion of air over the surface of water must become heated. As this warm air begins to rise, the cool air around it starts to form a small vortex of rotating air. This vortex of swirling air is carried upward by rising water vapor; waterspouts usually form over warm water because there is more water vapor present. Upon rising to an overhead cloud mass, the vortex forms a funnel cloud. At this point, a fair weather waterspout begins its five-stage life cycle.

The first stage of fair weather waterspout development involves the formation of a light spot on the surface of the water, which is surrounded by a darker area. The second stage is the formation of a spiral pattern on the surface around the dark spot. This spiral is colored lightly in some places and darkly in others. The third stage is when the funnel winds begin to pick up speed. The waterspout becomes more visible as sea spray is picked up from below.

This water swirls around the dark spot, known as a cascade. Like a hurricane, it appears to center around an eye. The fourth stage is the stage at which the waterspout is most powerful. Winds can be 60 mph or more. The funnel cloud can be hundreds or even thousands of feet high. Almost all the destruction caused by a waterspout occurs during this stage. The fifth and final stage is the decay of the waterspout. Often, rain begins to fall from the top of the waterspout. As these vortices require warm air to sustain themselves, the resulting downdraft of cool air speeds the process of dissipation of the waterspout. The entire life span of a fair weather waterspout is not often longer than 15 minutes.

Fair weather waterspouts can cause some damage, but it is always over water; they are rarely a threat upon reaching land. Tornadic waterspouts, however, are much more destructive. Their winds can exceed 150 miles per hour, they can last a half-hour or longer, and they are much more likely to appear during severe weather disturbances, such as tropical storms, tropical depressions, and hurricanes. These waterspouts can sink smaller ships and can cause extensive damage to sails and masts on larger ships. However, most waterspouts merely result in large amounts of rain.

READING LIST

The Five Stage Life Cycle of A Waterspout. (8 April 2002). Available: http://www.ziplink.net/users/jfrantis/cycle_ws.htm.

Smith, Bruce B. Waterspouts. (8 April 2002). Available: http://www.crh.noaa.gov/apx/science/spouts/waterspouts.htm.

Tornadoes: Waterspouts. (8 April 2002). Available: http://library.thinkquest.org/C003603/english/tornadoes/waterspouts.shtml.

Waterspouts. (8 April 2002). Available: http://www.srh.noaa.gov/eyw/HTML/spoutweb.htm.

Williams, J. *Waterspouts: Tornadoes Over Water*. USA Today. (8 April 2002). Available: http://www.usatoday.com/weather/wspouts.htm.

UNITED STATES WEATHER BUREAU

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

From 1890–1970, the United States Weather Bureau provided various weather-related services to the nation. Weather analysis had been considered for hundreds of years, but real progress was not made until the nineteenth century, when techniques and technology began to make accurate and useful weather forecasting possible, the main purpose of the United States Weather Bureau, its predecessors, and its successor, the National Weather Service. The first national American weather service was initiated on February 9, 1870, when President Ulysses S. Grant signed into existence what was then known as The Division of Telegrams and Reports for the Benefit of Commerce. This service began in November 1870. It lasted until 1890, when it was renamed the United States Weather Bureau, and Congress reassigned it from the Department of War to the Department of Agriculture at the behest of President Benjamin Harrison. Then, in 1940, President Franklin Delano Roosevelt reassigned the Weather Bureau again, this time to its current home, the Department of Commerce. The United States Weather Bureau existed until 1970, when its name was again changed, this time by President Richard Nixon, to the National Weather Service, the name it is known by today.

During the 1800s and early 1900s, servicemen for the Weather Bureau often communicated by telegraph, sharing readings of atmospheric pressure, temperature, and the like. They also communicated with seafaring ships in this manner, receiving the first weather report from a ship in 1905. From 1898 to 1933, the Weather Bureau also used kites regularly to ascertain atmospheric conditions. As technology progressed, so did the tools used by the Weather Bureau. Starting in 1909, the Bureau began using weather balloons, a more versatile tool, for weather measurements. And in 1925, it began to use airplanes in Washington, DC., as a more precise tool. The Air Commerce Act of 1926 make the Bureau responsible for providing weather-related services to civil aviators.

Besides forecasting daily weather, the Bureau also made progress in forecasting, tracking, and predicting more specific types of inclement weather. In 1943, an airplane was flown into the eye of a hurricane for the first time, a technique that now provides for hurricane readings that could not otherwise be obtained. In 1948, members of the Air Weather Service issued the first official tornado warning. And in 1952, the Bureau launched a program to forecast and warn of severe locally thunderstorms.

Throughout the Bureau's 70 years of existence, the United States government also tried to incorporate it into an international network of weather forecasting. In 1951, the World Meteorological Organization was formed, and at its head was U.S. Weather Bureau President Francis W. Reichelderfer. Ten years later, President John F. Kennedy called for a worldwide weather forecasting system in his State of the Union address. This all came together in 1963, when the World Weather Watch was established. This program still ensures today the free exchange of data, analysis, and equipment among WMO members (including the NWS), allowing for better weather forecasting and more standardized methods and measurements.

READING LIST

<http://www.randomhouse.com/features/isaacsstorm/science/history.html>

<http://www.nws.mbay.net/history.html>

<http://205.156.54.206/ pa/special/history/125thsk.htm>

<http://www.gfz-potsdam.de/ewc98/abstract/mildner.html>

<http://www.crh.noaa.gov/lmk/history1.htm>

WEATHER FORECASTING THROUGH THE AGES

NASA—Goddard Space Flight
Center

INTRODUCTION

Imagine a rotating sphere that is 12,800 kilometers (8000 miles) in diameter, has a bumpy surface, is surrounded by a 40-kilometer-deep mixture of different gases whose concentrations vary both spatially and over time, and is heated, along with its surrounding gases, by a nuclear reactor 150 million kilometers (93 million miles) away. Imagine also that this sphere is revolving around the nuclear reactor and that some locations are heated more during one part of the revolution and others during another part of the revolution. And imagine that this mixture of gases continually receives inputs from the surface below, generally calmly but sometimes through violent and highly localized injections. Then, imagine that after watching the gaseous mixture, you are expected to predict its state at one location on the sphere one, two, or more days into the future. This is essentially the task encountered day by day by a weather forecaster (Ryan, Bulletin of the American Meteorological Society, 1982).

EARLY HISTORY

The art of weather forecasting began with early civilizations using reoccurring astronomical and meteorological events to help them monitor seasonal changes in the weather. Around 650 B.C., the Babylonians tried to predict short-term weather changes based on the appearance of clouds and optical phenomena such as haloes. By 300 B.C., Chinese astronomers had developed a calendar that divided the year into 24 festivals, each festival associated with a different type of weather.

Around 340 B.C., the Greek philosopher Aristotle wrote *Meteorologica*, a philosophical treatise that included theories about the formation of rain, clouds, hail, wind, thunder, lightning, and hurricanes. In addition, topics such as astronomy, geography and chemistry were also addressed. Aristotle made some remarkably acute observations concerning the weather, along with some significant errors, and his four-volume text was considered by many to be the authority on weather theory for almost 2000 years. Although many of Aristotle's claims were

---

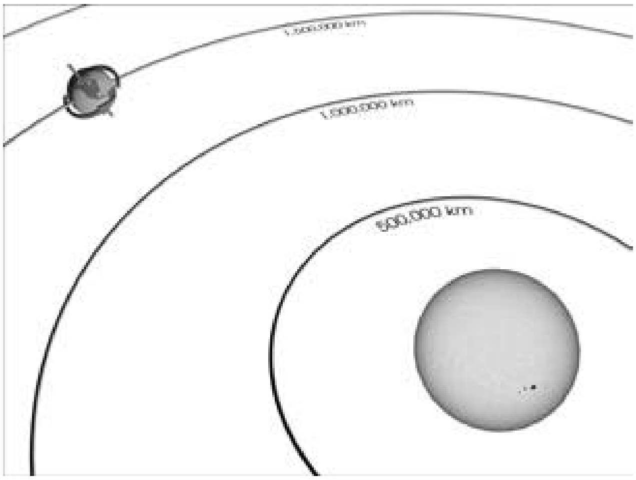

Earth rotates on its axis once every 23 hours, 56 minutes, and completes one revolution around the sun every 365.25 days

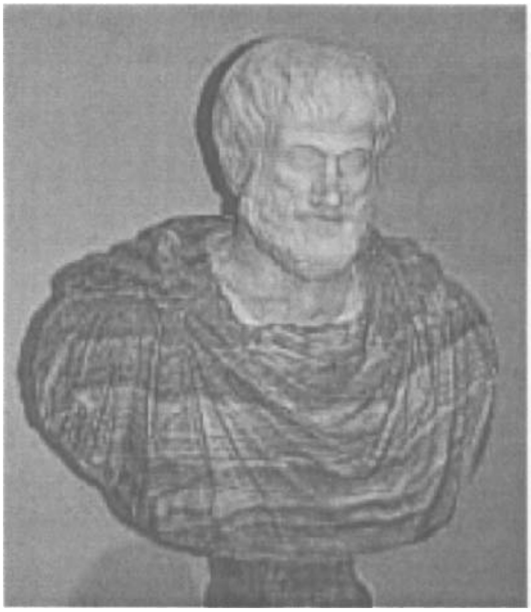

Aristotle, as sculpted by the Greek sculptor Lysippos

erroneous, it was not until about the 17th century that many of his ideas were overthrown.

Throughout the centuries, attempts have been made to produce forecasts based on weather lore and personal observations. However, by the end of the Renaissance, it had become increasingly evident that the speculations of the natural philosophers were inadequate and that greater knowledge was necessary to further our understanding of the atmosphere. In order to do this, instruments were needed to measure the properties of the atmosphere, such as moisture, temperature, and pressure. The first known design in western civilization for a hygrometer, an instrument to measure the humidity of air, was described by Nicholas Cusa (c. 1401–1464, German) in the mid-fifteenth century. Galileo Galilei (1564–1642, Italian) invented an early thermometer in 1592 or shortly thereafter; and Evangelista Torricelli (1608–1647, Italian) invented the barometer for measuring atmospheric pressure in 1643.

While these meteorological instruments were being refined during the seventeenth through nineteenth centuries, other related observational, theoretical, and technological developments also contributed to our knowledge of the atmosphere; and individuals at scattered locations began to make and record atmospheric measurements. The invention of the telegraph and the emergence of telegraph networks in the mid-nineteenth century allowed the routine transmission of weather observations to and from observers and compilers. Using these data, crude weather maps were drawn and surface wind patterns and storm systems could be identified and studied. Weather-observing stations began appearing all across the globe, eventually spawning the birth of synoptic weather forecasting, based on the compilation and analysis of many observations taken simultaneously over a wide area, in the 1860s.

With the formation of regional and global meteorological observation networks in the nineteenth and twentieth centuries, more data were becoming available for observation-based weather forecasting. A great stride in monitoring weather at high altitudes was made in the 1920s with the invention of the radiosonde. Small lightweight boxes equipped with weather instruments and a radio transmitter, radiosondes are carried high into the atmosphere by a hydrogen or helium-filled balloon that ascends to an altitude of about 30 kilometers before bursting. During the ascent, these instruments transmit temperature, moisture, and pressure data (called soundings) back to a ground station. There, the data are processed and made available for constructing weather maps or insertion into computer models for weather prediction. Today, radiosondes are launched every 12 hours from hundreds of ground stations all over the world.

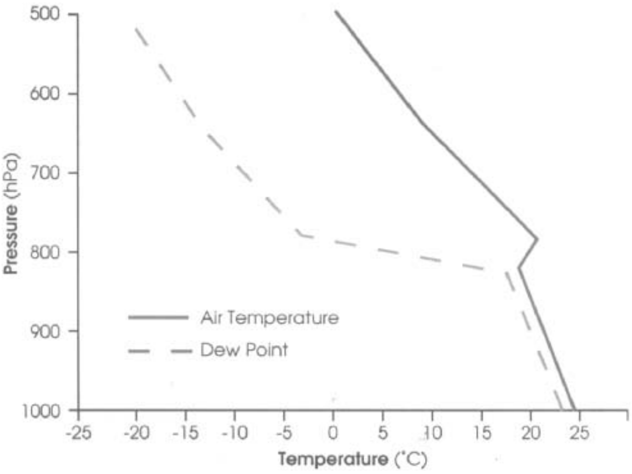

A schematic sounding of air temperature and dewpoint derived from radiosonde data. This sample schematic sounding includes a temperature "inversion" (temperatures increasing with height) at about 800 hPa and reflects atmospheric conditions that frequently precede the development of severe thunderstorms and possibly tornadoes. [1 hectoPascal (hPa) = 1 millibar (mb).]

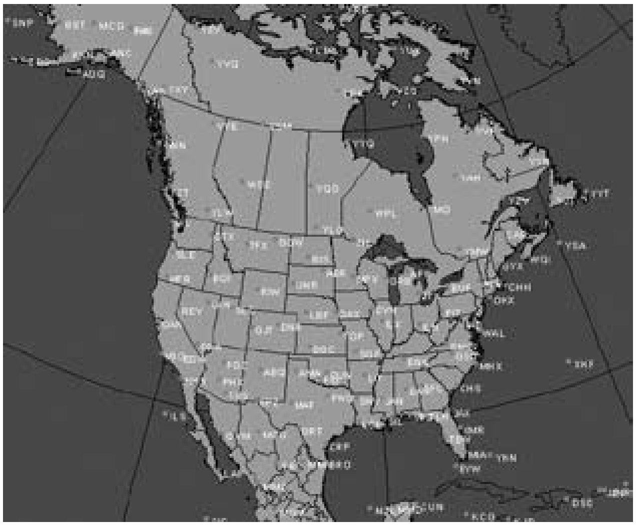

The North American network of upper air ground stations, each indicated by a three-letter station identifier. Radiosondes are launched and tracked from each location every twelve hours

TOWARDS NUMERICAL PREDICTION

Over the past few centuries, physical laws governing aspects of the atmosphere have been expressed and refined through mathematical equations. The idea of numerical weather forecasting—predicting the weather by solving mathematical equations—was formulated in 1904 by Vilhelm Bjerknes (1862–1951, Norwegian) and developed by British mathematician Lewis Fry Richardson (1881–1953, British). Despite the advances made by Richardson, it took him, working alone, several months to produce a wildly inaccurate six-hour forecast for an area near Munich, Germany. In fact, some of the changes predicted in Richardson's forecast could never occur under any known terrestrial conditions. Adding to the failure of this effort, a six-hour forecast is not particularly useful if it takes weeks to produce.

Courageously, Richardson reported his results in his book *Weather Prediction by Numerical Process*, published in 1922. In one of the chapters of this work, Richardson describes a scheme for predicting the weather before it actually happens, a scheme involving a roomful of people, each computing separate sections of the equations, and a system for transmitting the results as needed from one part of the room to another. Unfortunately, Richardson's estimated number of human calculators needed to keep pace with weather developments was 64,000, all located in one very large room.

Richardson's work highlighted the obvious fact that a large number of calculations had to be made very rapidly in order to produce a timely forecast. In the late 1940s, using one of the earliest modern computers, significant progress toward more practical numerical weather forecasts was made by a team of meteorologists and mathematicians at the Institute for Advanced Study (IAS) in Princeton, New Jersey. Mathematician John von Neumann (1903–1957, Hungarian-American) directed the construction of the computer and put together a team of scientists led by Jule Charney (1917–1981, American) to apply the computer to weather forecasting. Charney determined that the impracticality of Richardson's methods could be overcome by using the new computers and a revised set of equations, filtering out sound and gravity waves in order to simplify the calculations and focus on the phenomena of most importance to predicting the evolution of continent-scale weather systems. In April 1950, Charney's group made a series of successful 24-hour forecasts over North America, and by the mid-1950s, numerical forecasts were being made on a regular basis.

MODERN TOOLS OF THE TRADE

Modern technology, particularly computers and weather satellites, and the availability of data provided by coordinated meteorological observing networks, has resulted in enormous improvements in the accuracy of weather forecasting. Satellites, in particular, have given forecasters routine access to observations and data from remote areas of the globe. On April 1, 1960, the polar-orbiting satellite TIROS 1 (the first in the series of Television and Infrared Observation Satellites) was launched. Although the spacecraft operated for only 78 days, meteorologists worldwide were ecstatic over the pictures of the Earth and its cloud cover that TIROS relayed back to the ground.

Over the past 40 years, satellite sensor technology has advanced enormously. In addition to providing visual images, satellites can also provide data that allow calculation of atmospheric temperature and moisture profiles and other environmental variables. This is done using a variety of instruments, among them atmospheric

British mathematician Lewis Fry Richardson (1881–1953). (Photo credit: Bassano and Vandyk Studios.)

The first picture of Earth from a weather satellite, taken by the TIROS-1 satellite on April 1, 1960. Although primitive in comparison with the images we now receive from satellites, this first picture was a major advance

sounders, which measure quantities at various levels in atmospheric columns. The data retrieved from sounder measurements taken from a satellite can be made similar to radiosonde observations, with the major advantage that the satellite data are more complete spatially, filling in gaps between weather ground stations, which often are hundreds or even thousands of kilometers apart.

THE AQUA SPACECRAFT

In 2002, the Atmospheric Infrared Sounder (AIRS), the Advanced Microwave Sounding Unit (AMSU), and the Humidity Sounder for Brazil (HSB) were launched together on NASA's Earth Observing System (EOS) Aqua satellite, a satellite that also carries a Moderate-Resolution Imaging Spectroradiometer (MODIS), two Clouds and the Earth's Radiant Energy System (CERES) sensors, and an Advanced Microwave Scanning Radiometer for EOS (AMSR-E). The AIRS/AMSU/HSB sounding system provides the capability for determining atmospheric temperature and moisture more accurately than ever before from space-based measurements. These measurements provide to the National Oceanic and Atmospheric Administration (NOAA), the European Centre for Medium-Range Weather Forecasts (ECMWF), and the weather community at large for assimilation into operational numerical weather prediction models. It is expected that assimilation of global AIRS/AMSU/HSB data, complementing other operational observations, should lead to a substantial improvement in the accuracy of mid- and long-range weather forecasts.

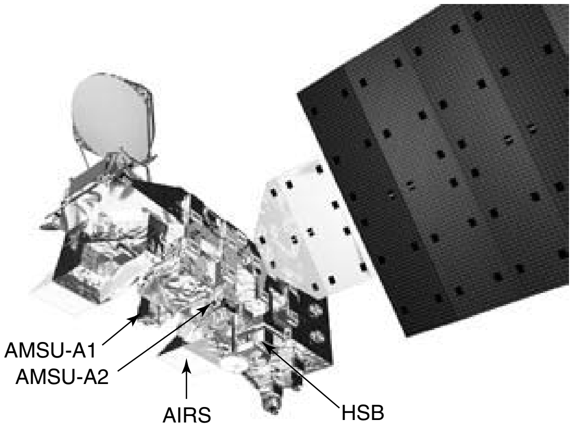

The Aqua spacecraft, with pointers to the AMSU-A1, AMSU-A2, AIRS, and HSB instruments

The AIRS/AMSU/HSB instrument suite builds on the technical heritage of NOAA's High Resolution Infrared Sounder (HIRS) and Microwave Sounding Unit (MSU). The HIRS/MSU combination was the National Weather Service's (NWS's) operational weather sounding system for nearly twenty years, flying on numerous NOAA polar orbiting satellites. This system was enhanced in the late 1990s by the replacement of the four-channel MSU by a 20-channel AMSU, which includes Aqua's AMSU and HSB channels. Looking toward further improvements in weather forecasts, the NWS has set measurement requirements for temperature at an accuracy of 1 °C in layers 1 km thick and humidity at an accuracy of 20% in layers 2 km thick in the troposphere (the lower part of the atmosphere, where weather systems are of most relevance to human life and property). AIRS/AMSU/HSB will meet these requirements, allowing meteorologists to improve and extend their predictions and reduce the number of significant prediction mistakes, like failing to predict a major storm prior to a few hours before its arrival.

BENEFITS TO SOCIETY

The economic benefits of more accurate weather forecasts are immense. For example, a more accurate 24-hour forecast of heavy rain and thunderstorms along a cold front could allow airline dispatchers enough time to reroute their airplanes appropriately and thereby help alleviate costly delays. Being able to pinpoint a wintertime low temperature in Florida could be the deciding factor in whether orange grove farmers make the correct decision regarding deployment or non-deployment of freeze prevention methods to save their crops. Better information on wind patterns could aid the National Hurricane Center in producing a more accurate forecast of a hurricane's track and might enable a reasoned decision to be made regarding evacuating thousands of families out of Miami or Jacksonville. In military operations, there is a considerable historical record of instances when weather conditions have altered the course of battles, including examples when accurate forecasting has been a deciding factor in one side's victory.

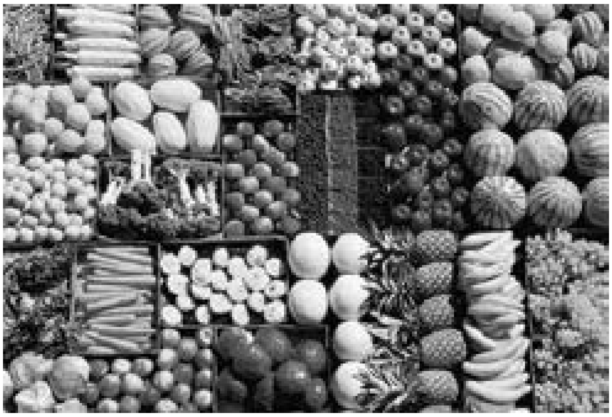

Sample crops that could benefit from more accurate weather forecasts aiding farmers in selecting harvesting times and in protecting their crops from freezing temperatures

CLOSING

Only fifty years ago, weather forecasting was an art, derived from the inspired interpretation of data from a loose array of land-based observing stations, balloons, and aircraft. Since then it has evolved substantially, based on an array of satellite and other observations and sophisticated computer models simulating the atmosphere and sometimes additional elements of the Earth's climate system. All this has been made possible by advances in satellite technology, a sweeping acceleration in worldwide communications, and overwhelming increases in computing power. Aqua's AIRS/AMSU/HSB combination should further these advances, enabling more accurate predictions over longer periods.

OVERVIEW OF WEATHER SYSTEMS

John W. Nielsen-Gammon
(from *Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements*, Wiley 2003)

INTRODUCTION

Meteorology is the only science in which it is common for a practitioner to appear every day on the evening news. Weather forecasting has immediate interest to people because it affects their day-to-day lives. The impact is great because weather itself changes from day to day. If weather were all good, or all bad, perhaps no one would notice—when's the last time you thought about the quality of the ground under your feet?

FOUNDATIONS OF WEATHER FORECASTING

The attempt to understand the weather is driven largely by the desire to be able to forecast the weather. However, the atmosphere is a complicated dynamical system, and an understanding of even the basic physical principles came only recently. Before that, scientists either attempted to understand the physical causes of weather events or recorded "weather signs," rules for weather forecasting. Because the physical understanding was incorrect, and most weather signs had no physical basis, neither approach was very successful. True progress in the study of the atmosphere that would lead to success in weather forecasting required a complete understanding of physical principles and a comprehensive set of observations of the atmosphere.

Aristotle's *Meteorologica*, written around 340 BC, was the leading book on meteorological principles in the Western world for two millenia. With his theories on the fundamental physical nature of the universe, Aristotle sought to explain a wide range of meteorological phenomena. For example, he argued that wind is caused by the hot, dry exhalation of the earth when struck by sunlight. A leading book on weather signs, *De Signis Tempestatum*, written by Aristotle's pupil Theophrastus, focused on such maxims as "A dog rolling on the ground is a sign of violent storm" and "Reddish sky at sunrise foretells rain." Since some of these maxims are still in use today, it appears that Theophrastus's work has outlived that of Aristotle.

Astronomers during the Middle Ages often argued that weather could be predicted by careful examination of the sky, clouds, and stars. This method of weather forecasting is sometimes successful, but it more closely resembled astrology than modern-day meteorology. Astrological forecasting schemes grew more advanced and intricate with time, while the scientific study of weather systems made little or no progress.

Scientific progress required basic observational data for describing weather systems. Only with good observations could theories be tested and the correct nature of the atmosphere be determined. The primary tools for observation of weather conditions are the thermometer (for temperature), the barometer (for air pressure), and the hygrometer (for humidity). Prior to and during the Middle Ages, these basic weather instruments did not exist. These instruments were invented in the sixteenth and seventeenth centuries and refined throughout the eighteenth century. The resulting observations of the atmosphere eventually established the basis for meteorological breakthroughs of the nineteenth and twentieth centuries.

Meanwhile, Isaac Newton had laid the foundation for the development of physical laws governing the motion of objects. The great mathematician Leonhard Euler of Germany, in 1755, rewrote those laws in a form that applied to a continuous fluid such as air or water. But Euler's laws were incomplete. They described how air can be neither created nor destroyed and how air accelerates (and wind blows) in response to forces acting on it. Missing from Euler's equations were the crucial relationship between temperature and pressure and the consequences of evaporation and condensation.

With the new instruments, the study of air became an experimental science. Through numerous laboratory experiments, scientists gradually discovered the fundamental physical laws governing the behavior of gases. In particular, the missing key relationship, the first law of thermodynamics, became known from experiments during the first half of the nineteenth century.

Now that the basic physical laws were known, a comprehensive set of observations of the atmosphere was necessary for further progress. Weather observations from a single careful observer were insufficient, but dozens of observations taken simultaneously across Europe or the United States gave a much clearer picture of the distribution of weather elements within a winter storm. Such "synoptic" observations were also used directly in weather forecasting. As rapid long-distance communication became possible through the telegraph, individual weather systems could be tracked and therefore forecasted. The coordinated observations that have taken place since the middle of the nineteenth century form the foundation of modern weather forecasting and our understanding of weather systems of all scales.

DEVELOPING AN UNDERSTANDING OF WEATHER SYSTEMS

Progress in understanding meteorology in the nineteenth century was largely based on thermodynamics. James Pollard Espy of the United States conducted a variety of laboratory experiments involving the condensation of water vapor in air. In 1841, based on his experiments and those of others, Espy correctly described the basic principle of thunderstorm formation: Condensation of water vapor within ascending air causes the air to become warmer than its surroundings and thus to continue rising. Espy went on to assert that the resulting lower pressure near the ground accounts for the low pressure observed in large-scale storm systems.

Espy's theory lay dormant until later in the nineteenth century, when the complete first law of thermodynamics was established. With this development, scientists were able to put Espy's theory of convection on a solid mathematical footing. Karl Theodor Reye of Germany, for example, determined the specific conditions under which such ascending air would be unstable. Their ideas form our understanding of the fundamental nature of individual thunderstorms to this day.

Espy also suggested that upward motion and the release of latent heat by evaporation might be the driving force for large weather systems. The widespread development of synoptic observations finally established that surface winds spiral inward toward a large-scale low-pressure system. The strong temperature contrasts within low-pressure systems suggested that the ascent would be driven, at least initially, by a current of warm air impinging on a current of cold air and rising. The rising air, and the resulting latent heating from condensation, would cause surface pressures to fall. This theory for the cause of low-pressure systems gained rapid acceptance for both theoretical and observational reasons in the 1870s.

A leading proponent of this thermal theory of cyclones was the American William Ferrel. Earlier, Ferrel had completed the application of Euler's equations to the atmosphere by formally including the effect of Earth's rotation as the Coriolis force. Ferrel also saw a clear analog to the role of equatorial convection in the general circulation, which drives ascent in the tropics and descent in the subtropics.

Unfortunately for its proponents, the thermal theory was wrong. Coordinated mountaintop observations in Europe, analyzed by Julius von Hann, showed that above the ground the centers of low-pressure systems were actually cooler than the centers of high-pressure systems. This observation illustrates the self-limiting nature of moist convection: The downward motion outside the convective cell causes warming of that surrounding air, so that eventually the instability is eliminated.

WEATHER AND ENERGY

The above discussion leads to a key question for the understanding of weather systems: How do weather systems grow and maintain themselves? As noted above, the basic dynamics of ordinary moist convection were well established, both conceptually and mathematically, by the end of the nineteenth century. Convection was also understood to be an important driver of the general circulation, or at least the tropical and subtropical circulation known as the Hadley cells.

By the turn of the century, the work of Hann and others showed that convection was not the energy source for extratropical cyclones, the midlatitude migratory low-pressure systems. Attention shifted from the role of latent heat release to the role of the large horizontal temperature gradients that were systematically observed within midlatitude low-pressure systems. In 1903, Max Margules, born in the Ukraine and working in Austria, calculated the amount of kinetic energy that could be obtained from the rising of hot air and the sinking of cold air in a low-pressure system. He found that the amount of energy that could be converted into kinetic energy was comparable to the actual amount of kinetic energy in a mature storm system. Margules had identified the correct energy source for extratropical cyclones. The process by which cyclones form and move was described by researchers working under Vilhelm Bjerknes in Bergen, Norway, shortly after World War I. Finally, a comprehensive theory by Jule Charney in 1947 explained that the structure and intensity of low-pressure systems was a consequence of the growth of unstable eddies on a large-scale horizontal temperature gradient.

Hurricanes did not have the prominent horizontal temperature variations found in midlatitude weather systems; indeed, many hurricanes seemed almost perfectly symmetrical. The prominence of convection throughout the hurricane, particularly within the eyewall, suggested that convection was fundamentally important in the development and maintenance of hurricanes. But how did the hurricane organize itself or maintain itself against large-scale subsidence within its environment? The currently accepted theory, published by Kerry Emanuel and Richard Rotunno in 1986 and 1987, relies on radiative cooling of the subsiding air at large distances from the hurricane. Emanuel also noted that as air spirals inward toward the eye it would become more unstable because it would be gaining heat and moisture from the sea surface at a progressively lower pressure.

Theories satisfactorily describing organized convection such as mesoscale convective systems and supercells also have been slow to develop. One reason for this is that no observing systems could accurately describe the structure of organized convection until the development of radar. Indeed, the term “mesoscale” was coined specifically to refer to in-between sizes (10 to 500 km) that were too large to be adequately observed at single locations and too small to be resolved by existing observing networks. The widespread use of weather radar in the 1950s helped fill the observation gap. The development of Doppler radar (for measuring winds within precipitating systems) for research purposes in the 1970s and as part of a national network in the 1990s helped even more. With comprehensive radar observations, it became clear that the long lifetime of organized convection was due to a storm keeping its updraft close to the leading edge of its cold, low-level outflow. The storm could then take advantage of the ascent caused by the cold air undercutting warm air without having its supply of warm air cut off completely. But even radar was not sufficient. The development of the first dynamical descriptions of supercells and squall lines in the 1980s, by Joseph Klemp, Richard Rotunno, Robert Wilhelmson, and Morris Weisman, relied upon numerical computer-generated simulations of the phenomena to provide an artificial data set unavailable from observations.

While the above discussion has focused on self-contained weather systems, certain “weather producers” may be thought of as byproducts of these weather systems. Many forms of severe weather, such as hail, lightning, and tornadoes, are essentially side effects of convection (organized or otherwise) but have little or no influence on the convection itself. Other mesoscale phenomena, such as sea breezes, upslope precipitation, and downslope windstorms, do not represent instabilities at all. Instead they are called “forced” weather phenomena because they are driven by such external features as solar heating gradients and topographic obstacles to the large-scale flow.

FORECASTING

The massive strides in weather forecasting during the past 50 years are due in large part to our growing understanding of the nature and dynamics of important weather systems. Forecasting of many phenomena has evolved from an exercise in extrapolation to specific predictions of the evolution of weather systems to computer simulations that accurately forecast the evolution of weather systems. Currently, the most skillful weather forecasts involve the prediction of large-scale extratropical weather systems.

One very important advance in our ability to forecast large-scale weather systems was the development of the omega equation, first presented in simplified form by Richard Sutcliffe in 1947. The omega equation is named after the Greek symbol that represents the change in pressure of an air parcel. In its original form, the omega equation is based upon a simplification of the equations of motion that assumes that all motions are large-scale and evolve slowly. When this approximation is made, it is possible to diagnose vertical motion entirely from large-scale pressure and temperature variations. Vertical motion is important for weather prediction because it is fundamentally related to clouds and precipitation, but large-scale vertical motion also can be used to infer the evolution of low-level and upper-level wind patterns. Vertical motion became the key to understanding the evolution of weather. In forecast offices, maps were designed and widely distributed that made it easy to look at the large-scale fields and diagnose vertical motion.

A second important advance was the development of numerical weather prediction, or NWP. NWP, discussed more extensively in the chapter by Kalnay in the dynamics part of this *Handbook*, has completely transformed the forecasting of large-scale weather systems. In the past, forecasters relied on their diagnosis of the current weather patterns to make forecasts. Nowadays, computers can make much more accurate forecasts than humans alone, and the best possible forecast is obtained by humans working with computer forecast output. The forecaster’s task has become one of identifying likely errors in the model forecast, based on the forecaster’s knowledge of systematic errors within the model, errors in the initial analysis, and computer forecast scenarios that run counter to experience.

In contrast to large-scale weather systems, individual convective storms are not directly simulated, nor accurately forecasted, by the numerical weather prediction models currently in use. Forecasting convective storms still relies heavily on extrapolation and a sound knowledge of what sorts of storms are likely to be produced by certain larger-scale weather conditions. The modern era of forecasting severe weather began in 1948, when two Air Force weather forecasters (Maj. Ernest Fawbush and Capt. Robert Miller) had responsibility for forecasting at Tinker Air Force Base near Oklahoma City. After a damaging tornado struck the base in 1948, the two meteorologists were given the task of identifying the days when tornadoes were likely to strike the base. We now know that forecasting the specific path of a tornado is essentially impossible, but when a similar low-pressure system evolved a few days later, the forecasters were compelled by the base commander to use a newly minted severe weather warning system to issue a forecast of a possible tornado at the base. Amazingly, the forecast came true, and preventive measures taken before the second tornado struck the base prevented considerable casualties to aircraft and personnel. This weather forecast, possibly the most serendipitous in the history of humankind, led directly to modern tools for diagnosing the likelihood of severe weather from large-scale conditions and laid the foundation for modern tornado forecasting.

CONCLUSION

This historical review has emphasized that progress in our understanding of the weather has required four ingredients: the need to understand the science of the atmosphere before one can hope to make accurate forecasts; a complete knowledge of the basic physical underpinnings of atmospheric behavior; observations adequate to test the theories and point the way toward new ones; and numerical models to provide simulations of the atmosphere more complete than any set of observations. The following chapters of this part of the *Handbook* discuss the current understanding of the weather phenomena discussed above, improvements in which have lead to vastly improved weather forecasts.

READING LIST

Frisinger, H.H. (1983). *The History of Meteorology: To 1800*. American Meteorological Society, Boston, MA.

Kutzbach, G. (1979). *The Thermal Theory of Cyclones: A History of Meteorological Thought in the Nineteenth Century*. American Meteorological Society, Boston, MA.

Miller, R.C., and Crisp, C.A. (1999). The first operational Tornado Forecast—twenty million to one. *Weather and Forecasting* **14**(4): 479–483.

UNIT HYDROGRAPH THEORY

Vijay P. Singh
Louisiana State University
Baton Rouge, Louisiana

INTRODUCTION

In 1932, Sherman (1) developed the concept of the unit hydrograph (UH) for determining the surface or direct runoff hydrograph (DRH) from the effective rainfall hyetograph (ERH). The unit hydrograph, or simply unit graph, of a watershed is defined as the DRH resulting from one unit (1 inch or 1 centimeter) of effective rainfall (ER) that occurs uniformly across the watershed at a uniform rate during a unit period of time. This unit period is not necessarily equal to unity; it can be any finite period up to the time of concentration. This unit period of ER is the period for which the UH is determined for a watershed. As soon as this period changes, so does the UH for the watershed. This means that there can be as many UHs for a watershed as periods of ER. Often used are a 1-hour UH, 6-hour UH, 12-hour UH, or a 1-day UH. Here the duration of 1 hour, 6 hours, 12 hours, or 1 day is not the duration for which the UH occurs, but it is the duration of the ER for which the UH is derived. The ER is assumed constant during its duration D for which the UH is derived, so the effective rainfall intensity (ERI) is $1/D$.

Since its inception, the unit hydrograph concept has been widely applied in surface-water hydrology. The principle underlying the UH has considerable appeal and has, therefore, been applied to a wide range of problems in environmental and water resources. For example, Rendon-Herrero (2) employed the UH concept to derive unit sediment graphs. Williams (3) extended the same concept to derive instantaneous unit sediment graphs. Thomann (4) used it to derive water quality graphs, and Dooge (5) employed it for groundwater recharge. It has also been used by Snyder and Asmussen (6) for subsurface flow, by Venetis (7) for groundwater flow, by Singh et al. (8) for irrigation flow, and by Westestrom and Singh (9) for snowmelt runoff.

The objective of this chapter is to describe the unit hydrograph theory and illustrate its application to build hydrologic models in four select areas.

UNIT HYDROGRAPH THEORY

The fundamental premise of unit hydrograph theory is that the hydrologic system is linear and time invariant. Although the assumptions implied or made in the theory are, strictly speaking, seldom true, the theory, nevertheless, does yield reasonable results in practice. The theory is applicable to a wide range of problems in hydrology and water resources.

The unit hydrograph is based on these postulates: (1) For a given drainage basin, the duration of surface runoff is constant for all uniform-intensity storms of the same duration, regardless of differences in the total volume of surface runoff. (2) For a given drainage basin, if two uniform-intensity storms of the same length produce different total volumes of surface runoff, then the rates of surface runoff at corresponding times, after the beginning of the two storms, are in the same proportion to each other as their total volumes of surface runoff. (3) The time distribution of surface runoff from a given storm period is independent of concurrent runoff from antecedent storm periods.

Postulates 1 and 2 together make up the principle of proportionality. In other words, if the duration of the ER is fixed but its volume changes, then the duration of the ERH does not change, but its ordinates do in proportion to the volume of ER. The third postulate is the principle of superposition, which allows the decomposition of a complex ERH into rectangular blocks or pulses and then superimposing, on one another, hydrographs of these rectangular pulses, each of steady intensity, to obtain the total DRH.

As stated in the definition of the unit hydrograph, it is assumed that the ER is uniformly distributed within its duration. This means that the rainfall intensity is uniform throughout the watershed while rain falls. The other assumption is that it is uniformly distributed throughout the watershed. This means that uniform runoff occurs from all parts of the watershed. The ER is defined as that portion of rainfall that produces surface runoff, so all abstractions must be subtracted from rainfall, and for the ER to be uniform, the abstractions must be uniform throughout the watershed.

These postulates, together with the definition and the assumptions, constitute what is now referred to as the unit hydrograph theory. The theory fundamentally assumes that the watershed is linear and time invariant, that is, the DRH is derived from the ERH by a linear operation. Once the assumptions of linearity and time invariance are recognized, the theory can be applied to any linear, time-invariant system (10,11) and is not restricted only to computing the surface runoff hydrograph. For this reason, it has been employed to derive the unit sediment hydrograph, the unit pollutant graph, the unit groundwater hydrograph, and the unit interflow hydrograph, among others. Thus, in a generic sense, the behavior of a linear, time-invariant system that receives input and produces output can be described by the theory. By conservation of mass, the volume of input must be equal to the volume of output.

The dimensions of the UH can be expressed in two ways. First, if the UH is derived from the DRH due to an ER of specified duration whose depth is not unity, then the dimensions of UH ordinates will be equal to the dimensions of runoff divided by the dimension of the depth of ER. Runoff can, however, be expressed either in terms of discharge in volumetric rate units or specific discharge in volumetric rate per unit area or intensity units. Therefore, the dimensions of UH ordinates are either the dimensions of discharge per unit depth or the dimensions of discharge per unit area per unit depth. Let L denote the dimension of length and t the dimension of time. Then, the two forms of the dimensions of UH ordinates are (1) L^2/t and (2) $1/t$.

LIMITATIONS OF THE UNIT HYDROGRAPH THEORY

The unit hydrograph theory is only an approximation, and its postulates are not strictly valid. Nevertheless, it is a useful practical tool and is amenable to linear mathematics whose methods are simple and easily understood.

An ER of a specified duration seldom occurs uniformly across a watershed of reasonable size. Spatial variation generally becomes more pronounced as the watershed size increases. Rainstorms that produce intense rainfall usually do not extend across large areas. The nonuniform distribution of rainfall can cause variation in the hydrograph shape. However, errors in the UH due to nonuniformity in areal distribution can be minimized by decomposing the watershed into subwatersheds, each small enough to ensure approximately uniform areal distribution, by deriving the UH for each subwatershed, and then by appropriately combining these UHs to develop the UH for the entire watershed.

An ER seldom occurs uniformly, even for as short a duration as 5 minutes. However, the effect of temporal variation of rainfall intensity on the UH shape depends principally on the watershed size. For example, rainfall bursts lasting only a few minutes may produce well-defined peaks in the runoff hydrograph from small watersheds but may cause little change in the runoff hydrograph shape from large watersheds. The effect of temporal variation can be minimized by employing the instantaneous unit hydrograph (IUH).

All watersheds in nature are nonlinear; some are more linear, and some are less. They are linear only by assumption. If the hydrographs from storms of the same duration are compared, their ordinates are not in proportion to their corresponding runoff amounts. The peaks of the UHs for small rainfalls are usually lower than those for larger ones.

MATHEMATICAL REPRESENTATION

To avoid ambiguity in interpreting the UH, it must be emphasized that it corresponds to a particular duration of an ER and has a volume of one unit (i.e., 1 inch or 1 cm). Because the ER, it is assumed, occurs uniformly, its duration is sufficient to define its uniform intensity. This duration is referred to as the unit storm duration and may change from one application to another. Thus, for a given watershed, there can be a large number of UHs; each corresponds to a specific duration of ER.

Let an ERH of D-hour duration be denoted as

$$I(t) = I, 0 \le t \prec D; \quad I(t) = 0, \quad t \ge D \tag{1}$$

where I is a constant ER intensity (L/t), the D-hour UH is $h(D, t)$ and the resulting DRH is $Q(t)$. Then, applying the principle of proportionality, one obtains

$$Q(t) = h(D, t)I(t)D = h(D, t)ID \tag{2}$$

The quantity ID denotes the volume of ER. This equation states that the DRH is proportional to the ER volume. When, $ID = 1$, the DRH and UH are numerically the same.

Let an ERH be a complex hyetograph composed of an n number of pulses, each of duration D hours. Let the intensity of pulse j be $I_j, j = 1, 2, 3, \ldots, n$, and the DRH of pulse j be $Q_j(t)$. If the D-hour UH is $h(D, t)$, then the DRH due to the complex hyetograph can be expressed as

$$Q(t) = \sum_{j=1}^{n} Q_j[t - (j-1)D] \tag{3}$$

Now, note that Q_1 starts at $t = 0$, Q_2 starts at $t - D$, and Q_j starts at $t - (j-1)D$. The individual DRHs can be expressed as

$$\begin{aligned}
&Q_1(t) = h(D, t)I_1 D \\
&Q_2(t - D) = h(D, t - D)I_2 D, \qquad h(D, t - D) = 0, t \le D \\
&Q_3(t - 2D) = h(D, t - 2D)I_2 D, \quad h(D, t - 2D) = 0, t \le 2D \\
&\vdots \\
&Q_n[t - (n-1)D] = h[D, t - (n-1)D]I_n D \\
&\qquad\qquad\qquad h[D, t - (n-1)D] = 0, t \le (n-1)D
\end{aligned} \tag{4}$$

Thus, the DRH can be expressed as

$$Q(t) = \sum_{j=1}^{n} h[D, t - (j-1)D]I_j D \tag{5}$$

The DRH at any time t is the result of the ERH pulses up to that time but not subsequent to it, so Equation 5 must be modified as

$$Q(t) = \sum_{j=1}^{t} h[D, t - (j-1)D]DI_j D \tag{6}$$

For simplicity, time t can be taken as a discrete variable that takes on values at D time intervals.

The difficulty that develops from the dependence of the UH on the duration of D of the ER is circumvented by letting D tend to be instantaneous. The UH so obtained is called the instantaneous unit hydrograph (IUH). Thus, the IUH, $h(0, t) = h(t)$ is a hypothetical UH due to the ER whose duration tends to zero as a limit, but whose volume remains unity (say, 1 cm). Clearly, the IUH is independent of the duration of ER. Mathematically,

$$h(t) = h(0, t) = \lim_{D\to 0} h(D, t) \tag{7}$$

and

$$\delta(t) = \lim_{D\to 0} I(t, D)D \tag{8}$$

where $\delta(t)$ is the Dirac delta function at $t = 0$. Physically, this function can be thought of as a spike of infinitesimally small thickness and infinitely large height such that the area under the spike is 1. Therefore, Eq. 6 can be expressed as

$$Q(t) = \int_0^t h(t-\tau)I(\tau)d\tau, \quad h(\tau) = 0, \quad \tau \leq 0 \tag{9}$$

Because mathematical operations are linear, Eq. 9 can also be written as

$$Q(t) = \int_0^t h(\tau)I(t-\tau)d\tau \tag{10}$$

where

$$h(t) \geq 0 \quad \text{for any } t \geq 0$$

$$\lim_{t\to\infty} h(t) = 0$$

and

$$\int_0^\infty h(s)ds = 1$$

Experimental evidence shows that

$$\int (t - \bar{t})h(t)dt \succ 0 \tag{11}$$

where $\bar{t}$ is the centroid of the IUH located at the t axis:

$$\bar{t} = \int_0^\infty th(t)dt \tag{12}$$

Equation 12 states that the IUH must be skewed to the right, that is, its peak must be to the left of the centroid, and a long tail extends to the right.

If the duration of the ER is indefinitely long and its intensity is one unit per unit of time (say, 1 cm/h), then the hydrograph so obtained is termed the summation hydrograph or S-hydrograph. This hydrograph assumes a deformed S-shape, and its ordinates ultimately approach the rate of ER in the limit or at the time of equilibrium. If the ERH is broken down into pulses each of duration D hours, then the SH will be obtained by superimposing the DRHs due to individual pulses. For this ERH, the SH, $U(t)$, is obtained from Eq. 6 as

$$U(t) = \sum_{j=1}^{t} h[D, t-(j-1)D]D = \sum_{j=1}^{t} Q[D, t-(j-1)D] \tag{13}$$

Clearly, if D tends to 0, $h[D, t-(j-1)D]$ tends to $h(t)$, the lower limit for summation will approach the origin, and

$$U(t) = \int_0^t h(s)ds \tag{14}$$

Now, the relations between IUH, UH, and SH can be derived. From the definitions of the UH and IUH, one can write

$$h(D, t) = \frac{1}{D}\int_{t-D}^{t} h(s)ds \tag{15}$$

Similarly, the relation between UH and SH can be expressed as

$$h(D, t) = \frac{1}{D}[U(t) - U(t-D)] \tag{16}$$

As D tends to 0,

$$h(t) = \frac{dU(t)}{dt} \tag{17}$$

It may be instructive to summarize that the three hydrographs—IUH, UH, and SH—are ascribed to three distinct characteristics of the effective rainfall—volume, duration, and intensity—as shown in Table 1:

Also note that the time to the peak of $h(D, t)$ can be determined by differentiating Eq. 16 with respect to t and equating the derivative to zero:

$$h(t) - h(t-D) = 0 \tag{18}$$

Equation 18 states that the peak of $h(D, t)$ occurs when the ordinate of the IUH is equal to the ordinate at D time units earlier. The SH attains its maximum value at a time equal to D hours less than the time base of the initial $h(D, t)$.

UNIT HYDROGRAPH MODELS

The unit hydrograph theory has been extensively applied to develop a range of hydrologic models in surface water hydrology, groundwater hydrology, and water quality hydrology. Four illustrative examples are presented.

Rainfall–Runoff Models

The unit hydrograph theory was developed for rainfall–runoff modeling and has been most extensively applied therein. There is a large range of UH-based models, including empirical, conceptual, and physically based (10). However, for economy of space, only a simple conceptual model will be presented to illustrate the power of the UH theory. To develop a conceptual model,

Table 1. Effective Rainfall Characteristics

Hydrograph	Volume	Duration	Intensity
IUH	Unity	Zero	Indefinite
UH	Unity	Finite	1/duration
SH	Indefinite	Indefinite	Unity

a watershed is represented by a network of linear elements, such as reservoirs and channels. This network can be defined using the laws of geomorphology or arbitrarily to mimic best the flow pattern in a watershed. Rainfall is reduced to effective rainfall by subtracting infiltration. Then, an instantaneous ER is routed through the network of reservoirs and/or channels to obtain the IUH of the watershed. By convoluting this IUH with the ERH, the DRH is obtained.

A linear reservoir can be defined by a linear relation between storage S and surface runoff Q as

$$S = KQ \tag{19}$$

where K is the residence time. The water balance of the reservoir can be expressed as

$$\frac{dS}{dt} = I(t) - Q(t) \tag{20}$$

where I is ER and t is time. Substituting Eq. 20 in Eq. 19 yields

$$\frac{dQ}{dt} + \frac{1}{K}Q = \frac{I}{K} \tag{21}$$

For an instantaneous $I(t)$, represented by $\delta(t)$, $Q(t)$ will reduce to $h(t)$. Therefore, Eq. 21 can be expressed as

$$\frac{dh}{dt} + \frac{1}{K}h = \frac{\delta(t)}{K} \tag{22}$$

For $h(0) = 0$, the solution of Eq. 22 is

$$h(t) = \frac{1}{K}\exp(-t/K) \tag{23}$$

If n linear reservoirs are arranged in series, then the IUH of this cascade is

$$h(t) = \frac{1}{K\Gamma(n)}\left(\frac{t}{k}\right)\exp\left(-\frac{t}{K}\right) \tag{24}$$

Equation 24 is a gamma function and is known as the Nash model (12).

A linear channel is defined by a linear rating curve or a linear relation between flow cross section $A(t)$ and discharge Q:

$$A = C(x)Q \tag{25}$$

where $C(x)$ is a constant at a channel section, varies with distance, and has the connotation of the inverse of the velocity; and x is the distance along the channel. The continuity equation for a channel in one-dimensional form can be written as

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = 0 \tag{26}$$

Substituting Eq. 25 in Eq. 26 produces

$$\frac{\partial Q}{\partial x} + C(x)\frac{\partial Q}{\partial t} = 0 \tag{27}$$

The inflow to the channel is upstream and can be denoted as $Q(x_0, t)$. A linear channel simply translates the flow by the time equal to the travel time, T. Thus, the solution of Eq. 27 is

$$Q(x_1, t) = Q(x_0, t - T), \quad T = C(x_1 - x_0) \tag{28}$$

where x_1 denotes the downstream end of the channel reach. If the inflow is given by the delta function $\delta(t)$ at the upstream end of the channel, the IUH of the channel is

$$h(t) = \delta(t - T) \tag{29}$$

The linear reservoirs and linear channels can be arranged in a way that best represents the watershed. Dooge (13) employed a network of linear channels and reservoirs to represent the watershed. All operations are linear, so the IUH of the network can be derived with little difficulty. For any $I(t)$, the DRH can be determined by convolving the IUH with $I(t)$ as

$$Q(t) = \int_0^t h(t - \tau)I(\tau)d\tau \tag{30}$$

For practical purposes, Eq. 30 can also be expressed in discrete form.

Channel Flow Routing

Channel flow is routed by using either a systems-based approach or a hydraulic approach. The unit hydrograph theory is applicable to both approaches if they are linear. The systems-based approach based on the Muskingum method is the most popular. This method is comprised of Eq. 20 and a linear storage–discharge relation expressed as

$$S = k[wI + (1 - w)Q] \tag{31}$$

where S is the storage of water within the channel reach, I is the inflow rate, Q is the outflow, and w and k are routing parameters. For a delta function inflow, $\delta(t)$, the IUH of the Muskingum method is obtained by substituting Eq. 31 in Eq. 20 and solving the combined equation:

$$h(t) = -\frac{w}{1 - w}\delta(t) + \frac{1}{k(1 - w)^2}\exp\left[-\frac{t}{k(1 - w)}\right] \tag{32}$$

where $I(0)$ is $I(t)$ at $t = 0$. By convoluting Eq. 32 with $I(t)$, the outflow discharge $Q(t)$ can be obtained.

On the other hand, in a hydraulic approach, St. Venant equations are employed. The unit hydrograph theory has been applied to several linearized forms of these equations. One simple approximation is the kinematic diffusion approximation (14,15) expressed as

$$\left(gh_0 - \frac{u_0^2}{4}\right)\frac{\partial^2 Q}{\partial x^2} = 3gS_0\frac{\partial Q}{\partial x} + \frac{2gS_0}{u_0}\frac{\partial Q}{\partial t} \tag{33}$$

where h_0 is the flow depth at the reference discharge, u_0 is the velocity at the reference discharge, Q is the perturbation discharge around the reference discharge, g is the acceleration due to gravity, S_0 is the bed slope, x is the distance along the channel, and t is time. The reference discharge may correspond to steady uniform flow. For instantaneous flow at the upstream end of the channel reach, $\delta(t)$, the IUH of the channel reach described by Eq. 33 is

$$h(t) = \frac{x}{\sqrt{4\pi Dt^2}} \exp\left[-\frac{(x-at)^2}{4Dt}\right] \tag{34}$$

where $a = 1.5\,u_0$, and

$$D = \frac{Q_0}{2S_0}\left(1 - \frac{F^2}{4}\right), \quad F = \frac{u_0}{\sqrt{gh_0}} \tag{35}$$

where F is the Froude number corresponding to the reference discharge and D is hydraulic diffusivity. By convoluting Eq. 34 with the upstream discharge, the outflow (or routed) discharge can be obtained.

Groundwater Flow

The unit hydrograph theory has been applied to model groundwater recharge, stream–aquifer interaction, aquifer response due to pumping, and a whole host of groundwater problems. Only the problem of stream–aquifer interaction will be discussed to illustrate an application of the unit hydrograph theory in groundwater hydrology. To that end, suppose that a seminifinite aquifer interacts with a stream. The aquifer is confined, homogeneous, and isotropic. The governing equation for one-dimensional flow can be written as

$$\frac{\partial^2 H(x,t)}{\partial x^2} = \frac{1}{D}\frac{\partial H(x,t)}{\partial t} \tag{36}$$

where H is the hydraulic head and D is the aquifer diffusivity equal to the ratio of transmissivity (T) to storativity S_c. The initial and boundary conditions can be written as

$$h(x,0) = 0 \tag{37a}$$

$$h(0,t) = I(0,t) \tag{37b}$$

$$h(0,t) \to 0, \text{ as } x \to \infty \tag{37c}$$

The IUH of Eq. 36 can be obtained using the Laplace transform (16) as

$$h(x,t) = \frac{x}{2t}\left[\frac{S_c}{\pi t T}\right]^5 \exp\left[-\frac{S_c}{T}\frac{x^2}{4t}\right] \tag{38}$$

Equation 38, when convoluted with $I(0,t)$, will yield the aquifer response.

SEDIMENT TRANSPORT

Analogous to the IUH, in the instantaneous unit sediment graph (IUSG), h_S is defined as the distribution of sediment from an instantaneous burst of rainfall that produces one unit of runoff and is considered the product of the IUH and the sediment concentration distribution $C(t)$:

$$h_S(t) = h(t)C(t) \tag{39}$$

In the IUSG, $C(t)$, it is assumed, varies linearly with V_Q, the amount of effective rainfall or direct runoff. Analogous to the ERH, Chen and Kuo (17) defined the effective sediment erosion intensity per unit depth (e.g., foot) (ESEI), S_e. Then, similar to the surface runoff hydrograph given by Equation 6, the sediment discharge, Q_S, can be expressed as

$$Q_S(t) = \sum_{j=1}^{n} h_S[D, t-(j-1)D]D)S_e(t)D \tag{40}$$

where D is the time interval of each pulse. The dimensions of Q_S and h_S are FT^{-1} and T^{-1}, respectively. A value of 1 hour or less may be appropriate for D. Singh and Chen (18) established that the sediment yield, Y, and the direct runoff volume are related as

$$Y = aV_Q^b \tag{41}$$

where a and b are parameters. A similar relation exists between $Q(t)$ and $Q_S(t)$. As $Q(t)$ varies with ERI, so does $Q_S(t)$ with ESEI. In other words, the effective rainfall intensity or depth produces direct runoff; in turn, the direct runoff produces the direct sediment yield. Conceptually, therefore the ERI generates its counterpart ESEI which when combined with IUSG produces a sediment graph. The relation between $S_e(t)$ and the effective rainfall is of the same form as that between $Q_S(t)$ and $Q(t)$. The sum of S_e ordinates must equal Y as that of ER ordinates must equal V_Q.

CONCLUSIONS

The unit hydrograph theory is a powerful tool that is applicable to a wide range of problems in hydrology and environmental engineering. In practice, this theory is still routinely applied. Although most hydrologic and environmental systems are nonlinear, their linear part dominates in most cases. That is why this theory works so well.

BIBLIOGRAPHY

1. Sherman, L.K. (1932). Stream flow from rainfall by the unit graph method. *Engineering News Record* **108**: 501–505.
2. Rendon-Herrero, O. (1978). Unit sediment graph. *Water Resources Research* **14**(5): 889–901.
3. Williams, J.R. (1978). A sediment graph model based on an instantaneous unit sediment graph. *Water Resources Research* **14**(4): 659–664.
4. Thomann, R.V. (1972). *Systems Analysis and Water Quality Management*. McGraw-Hill Book Company, New York.
5. Dooge, J.C.I. (1960). The routing of groundwater recharge through typical elements of linear storage. *IASH Bulletin* **52**: 286–300.
6. Snyder, W.M. and Asmussen, L.E. (1973). Subsurface hydrograph analysis by convolution. *Journal of Irrigation and Drainage Division* ASCE **98**(IR3): 405–418.
7. Venetis, C. (1968). On the impulse response of an aquifer. *IAHS Bulletin* **13**(3): 136–139.

8. Singh, V.P., Scarlatos, P.D., and Prasad, S.N. (1990). An improved Lewis-Milne equation for the advance phase of border irrigation. *Irrigation Science* **11**: 1–6.
9. Westerstrom, G. and Singh, V.P. (2000). An investigation of snowmelt runoff on experimental plots in Lulea, Sweden. *Hydrological Processes* **14**: 1869–1885.
10. Singh, V.P. (1988). *Hydrologic Systems: Vol. 1. Rainfall-Runoff Modeling*. Prentice Hall, Englewood Cliffs, NJ.
11. Singh, V.P. (1989). *Hydrologic Systems: Vol. 2. Watershed Modeling*. Prentice Hall, Englewood Cliffs, NJ.
12. Nash, J.E. (1957). *The Form of the Instantaneous Unit Hydrograph* IASH Publication, **45**(3): 114–121.
13. Dooge, J.C.I. (1959). A general theory of the unit hydrograph. *Journal of Geophysical Research* **64**(2): 241–256.
14. Dooge, J.C.I. (1973). *Linear Theory of Hydrologic Systems*. Technical Bulletin No. 1468, p. 327, U.S. Department of Agriculture, Agricultural Service, Beltsville, MD.
15. Singh, V.P. (1996). *Kinematic Wave Modeling in Water Resources: Surface Water Hydrology*. John Wiley & Sons, New York.
16. Hall, F.R. and Moench, A.F. (1972). Application of the convolution equation to stream-aquifer relationships. *Water Resources Research* **8**(2): 487–493.
17. Chen, V.J. and Kuo, C.Y. (1986). A study of synthetic sediment graphs for ungaged watershed. *Journal of Hydrology* **84**: 35–54.
18. Singh, V.P. and Chen, V.J. (1982). On the relation between sediment yield and runoff volume. In: *Modeling Components of Hydrologic Cycle*. V.P. Singh (Ed.). Water Resources Publications, Littleton, CO, pp. 555–570.

WEATHER AND THE ATMOSPHERE

David P. Stern
Goddard Space Flight Center,
NASA

An optional extension of section (S-1) "Sunlight and the Earth." A more detailed (but qualitative) discussion of heat flow in the atmosphere and related weather processes, including the roles of buoyancy, convection and humidity.

> Around and around goes the wind, and from its circuits returns the wind.
>
> *Ecclesiastes, ch. 1, v.6*

PRESSURE AND CONVECTION

Let us start with the **flow of air.** Suppose a "parcel of air" is heated near the ground (by **conduction** of heat, the flow of heat due to direct contact). Heat makes it expand, it becomes less dense than the surrounding air and **buoyant,** and it rises like a hot-air balloon (or like a drop of oil in a bottle of water).

At the higher levels of the atmosphere, this warm bubble again gives up its heat (to other flows or perhaps to cold space), cools down, and other bubbles coming from below push it to the side, where it descends again. (diagram on the board). Such a circulating flow is called **convection.**

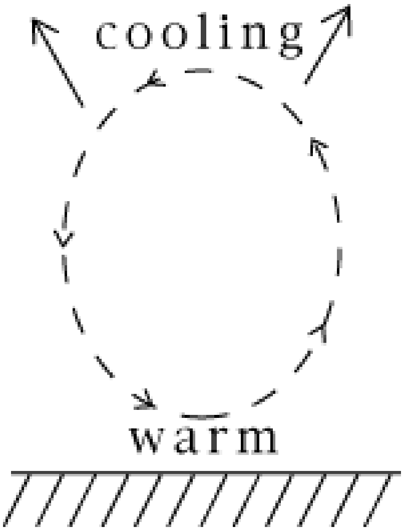

More generally, convection is any flow which

1. picks up heat at one place,
2. drops it at another, and
3. is driven by this transport of heat.

The important thing to remember when dealing with convective flows, is that the **higher** one is in the atmosphere, the lower are the **pressure and density** of the air. What compresses it is the **weight of air above it** which it must support. On top of Mt. Everest, less air is piled up on top and therefore the pressure is lower.

At ground level, the compressing weight of the atmosphere amounts to about 1 kilogram on each square centimeter. That pressure does not bother our bodies, because the air inside us is at the same pressure, and the fluids of the body (like blood) do not compress easily. For the same reason, fish have no problems with depth—even at a depth of 100 meters, with a pressure 11 times larger (10 kilogram water above each square centimeter, plus the weight of the atmosphere) they feel no discomfort.

(Divers too can stand such pressure, provided the air they breathe is similarly compressed—except that the mixture must be changed, otherwise they may take in too much oxygen, and too much nitrogen is dissolved in their bloodstream.)

At an altitude of about 5 kilometers, only **half** the atmosphere is above us, the other half is below, so only the weight of half the atmosphere must be supported, and the pressure is reduced to one half.

By "Boyle's law" (named for Robert Boyle, 1627–91), the **density** is also reduced to one half (ignoring any variation in temperature). Rising an **additional** 5 kilometers, the pressure again falls by half, to 1/4 of what it was on the

---

ground, and at 15 kilometers, it is halved again to about 1/8. All this is approximate and depends on temperature, but the trend should be clear.

The cabin of a jetliner flying at 10 km must be sealed and pressurized, because passengers breathing air at 1/4 the sea-level density would be starved for oxygen and might lose consciousness. On the very rare occasions when a jetliner loses its pressure, masks connected to oxygen canisters drop down automatically, allowing the passengers to breathe normally while the pilot quickly descends to a lower altitude.

WEATHER

First, **a word of caution:** what follows below is **a very crude and simplified** picture of a much more complicated process. While the circulation in the drawing below resembles the one proposed by Hadley in 1735, the actual situation is much more involved.

(a) Local Weather

When the atmosphere is stable, the higher we go, the **cooler** the air is.

Air is **warmest** near the ground, which absorbs receives heat from sunlight. It is **coldest** above the level where jetliners fly, at 10–15 kilometers, the region from where it **radiates** most of its heat into space. That is why mountaintops are cold and the highest mountains have snow on their tops.

(Still higher layers may get quite hot again, by absorbing UV and "extreme UV," but they have little effect on what goes on below them).

How Exactly Does This Happen?

Suppose some "parcel of air" (dry air, for now—humidity is an additional factor, considered later) is heated by the ground and rises. Higher up the pressure is lower, so the air expands: but expansion **cools** it.

Similarly, if for some reason the parcel is blown down, is compressed again and **heated** by the compression. Such up-and-down motions happen all the time, and the net result is that when conditions are stable, the temperature drops at a **steady rate** as we go higher.

The motion of the rising parcel of air depends on its surroundings. It always cools by expansion—but is it still warmer than the still air around it? **If it is, it continues to rise; if not, it stops.** As will be seen, this is where the **humidity** of air has an important effect.

[On an ordinary day, direct heating by the ground only carries the air a few hundred meters, perhaps a kilometer, creating above the ground a "boundary layer" with many small convective flows. Large scale motions like thunderstorms usually occur higher up (see below).]

(b) Global Weather

Convection also operates on a global scale. The greatest heating occurs near the equator, and air heated there **rises and flows poleward,** to cooler regions of the Earth.

The rotation of the Earth greatly modifies this flow, by the *Coriolis effect*, as explained below.

- At the equator the east-west motion of the air **matches** that of the equator.
- **Away** from the equator, however, the Earth's surface is **closer** to the rotation axis, the distance it covers in 24 hours is **smaller,** and so its west-to-east speed is **slower**. If the air moving **away** from the equator (right side of the drawing above) **persists** at its original west-to-east speed, it will **overtake** the local surface, making the predominant winds blow from **west to east** ("westerlies").
- The cooler air returns equatorward (middle part of the drawing above) at **lower** altitudes, completing the loop. If it **still** kept its original west-to-east speed, it would **again** match the local rotation of the equator.
- Actually, the air loses velocity to friction with the ground. Therefore, the time it returns to the equator, it **lags behind** the rotating ground, and the average wind is **easterly**.

In the age of sailing ships, sea-captains took advantage of this system. Sailing from Spain to America, they would

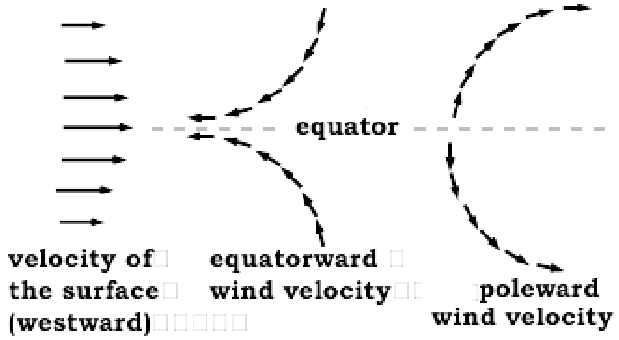

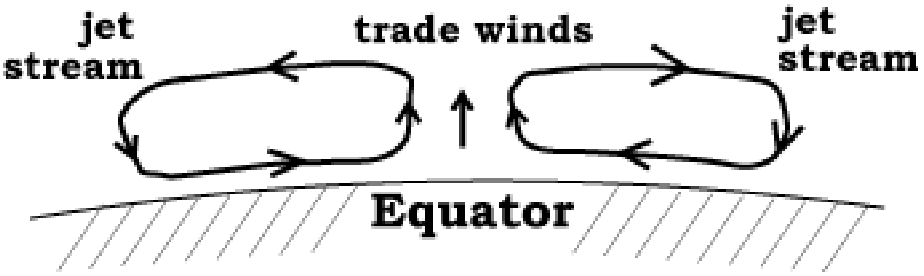

go closer to the equator, a more southern route that took advantage of the easterly "trade winds." Sailing back home they would go further north and use the westerlies. The many Spanish wrecks off Florida, some containing quite rich cargo, were lost on this home voyage back to Spain, loaded with gold and silver from Mexico and South America.

Jetliners flying across the US cannot swing quite as far. However, when flying west to east, their pilots often exploit the fast core of the westerlies, known as the **jet stream** and flowing at high altitudes. Flying westward they try to avoid the jet stream.

WATER VAPOR

Instead of heating the Earth, sunlight can **evaporate water** from it—especially from the oceans, which cover most of the Earth's surface. Humid air may be viewed as containing **additional energy**, invested by the Sun when its heat evaporated the water. While heat drives convection, humidity may **amplify** it.

Hot humid air is what drives thunderstorms, and a warm ocean surface is also the traditional birthplace of violent tropical storms, known as hurricanes in America and typhoons in Asia.

We look at two examples of humidity in action.

1. In a **thunderstorm**, hot humid air rises, as in ordinary convection. As it rises to regions of lower pressure, it cools by expanding. However, cold air cannot hold as much humidity as warm air, and the extra water is therefore squeezed out. In moderate convection it forms **clouds** (as in the second example below), but in a vigorous thunderstorm, there is too much of it and it turns into **rain**.

Giving up water **heats** the air, or rather, slows down its cooling, because the heat invested by the Sun in evaporating the water is now passed back to the air. As a result, the rising air is still warmer than the air layers around it, and it **continues to rise vigorously**. It squeezes out still more rain and forms the tall thunderstorm clouds, which pilots know well to avoid.

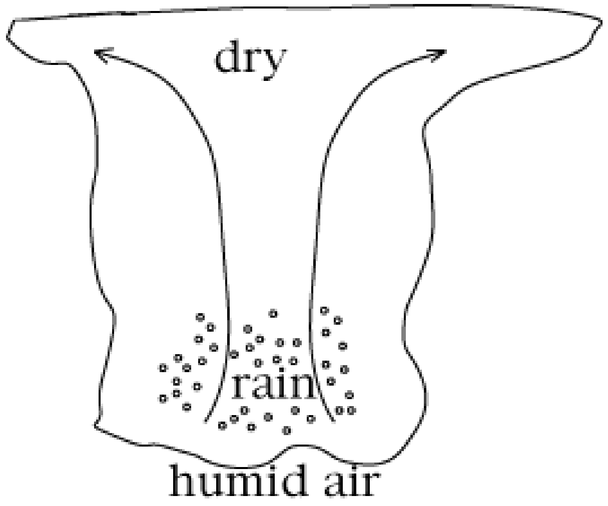

Thunderstorm convection

(In very vigorous thunderstorms, the "updrafts" of rising air may rise so quickly that they blow raindrops into the higher and colder sections of the cloud, where they **freeze,** producing **hail.** Some hailstones are picked up again and again, adding layers of ice on each upwards journey. That is how unusually big hailstones can form.)

2. On a hot clear day, many fluffy small **clouds** may form. A light plane flies across the land, and every time it comes under a cloud, the pilot feels that it is **bodily lifted.** What is happening?

—Here is the reason. The heating of the ground by the Sun has created many small convection currents rising upwards. Their air contains humidity, not enough for a serious thunderstorm, but enough to produce small clouds when droplets of water condense as the rising air cools.

The small clouds mark the top of the "boundary layer" near the ground, with many small circulating flows. Each cloud sits on top of a rising convection current, which lifts the airplane as it flies across it. Since "what goes up must come down" the pilot can also expect downdrafts between the clouds, where the air goes down again, as part of the circulating convection. Such up-and-down motions may make passengers on low-flying airplanes quite airsick!

Further Exploring. Strictly for those interested in history: the original 1896 article in which Svante Arrhenius proposed the greenhouse effect. Note that carbon dioxide was then called "carbonic acid."

Tidbit. Who first wrote "**Everybody talks about the weather, but nobody does anything about it**"? Most would claim it was Mark Twain (just type the first 5 words into a search engine and see!), but it ain't necessarily so. It looks like Twain's style, but actually the words first appeared in an editorial in the Hartford (Conn.) Courant on 24 August 1897, written by **Charles Dudley Warner.**

Warner was a good friend of Twain, who himself had lived in Hartford for many years (he left before 1897). He was a newspaperman in Hartford and the two had collaborated on an 1873 book "The Gilded Age." Warner is also remembered for other quotes, e.g., "Politics make strange bedfellows."

Concerning the quote, see on the web http://www.m-w.com/wftw/99aug/082499.htm. and more about Charles Dudley Warner (and Twain) at http://courant.ctnow.com/projects/twain/warner.htm.

VAPOR PRESSURE

Naraine Persaud
Virginia Polytechnic Institute
and State University
Blacksburg, Virginia

Vapor pressure is the partial pressure exerted by the vapor of a substance in the gaseous phase of a closed heterogeneous system at equilibrium with condensed phases (i.e., solid or liquid) of the substance present in the system. It is the macroscopic measure of the tendency of the molecules of a given species in the condensed phases of the system to escape into the gaseous phase. A heterogeneous system means a mixture of two or more distinct, physically or chemically homogeneous phases. Mixture implies the phases are mechanically separable; distinct implies presence of unequivocally defined physical boundaries (interfaces) between phases; and homogeneous implies each phase is uniform throughout chemically and physically. The simplest possible conceptualization for a heterogeneous system is any closed two-phase system consisting of the solid or liquid of a pure substance in equilibrium with its vapor at a given temperature. In more complex heterogeneous systems, the condensed phases may be completely miscible or immiscible liquids, aqueous solutions, discontinuous liquid aerosols, or solid/liquid emulsions, and the vapors may be mixed with other vapors and gases present in the gaseous phase. Although the entire system is closed (permitting only energy transfer) with respect to its surroundings, the subsystems (phases) are open, permitting both energy and mass transfer between phases.

The equilibrium vapor pressure is of interest and importance in many applications encountered in water/atmospheric science and engineering such as quantifying water vapor pressure in the air above ice or liquid water, above an aqueous solution, and in the presence of curved air–water interfaces (menisci, bubbles, and droplets). Covering all these facets of vapor pressure would necessitate a discussion that far exceeds the limits of an encyclopedia article on vapor pressure. Our subsequent discussion will therefore focus on the fundamental relationships that can be universally applied to quantitatively describe and predict phase equilibrium in any closed heterogeneous system—in particular, a discussion of the Gibbs chemical potential and phase rule. We will then demonstrate applications of these principles to the simplest (and perhaps the most common) case of the vapor pressure in a two-phase system of a pure substance with specific reference to the solid/liquid/vapor system for pure water.

In a quintessential contribution to classical thermodynamics (entitled "On the equilibrium of heterogeneous substances") published between 1875 and 1876 in the *Transactions* of *the Connecticut Academy of Arts and Sciences*, Josiah Willard Gibbs (1839–1903) formulated seminal concepts governing equilibrium in *any* closed heterogeneous system (1,2). Specifically, Gibbs addressed reversible changes in the equilibrium state of closed heterogeneous systems at uniform temperature and pressure due to changes in the chemical composition or physical state of any of the phases through chemical reactions or mass transfer between phases.

From the first law of thermodynamics, for a reversible change in equilibrium state of a closed system, the differential change (gain) in the internal energy $dU = \delta q + \delta w$, where δq and δw represent the small amounts of the (positive) total heat or total work transferred into the system. Heat or work transferred out of the system is negative, implying a loss in U. The total work can be divided into any PV work (δw_{PV}) done on the system ($\delta w_{PV} = -P\,dV$ since dV is negative), and any (reversible) non-PV work done on the system ($\delta w_{\mathrm{n}PV}$). Since dS (entropy change) is defined as $\delta q/T$, then $\delta q = T\,dS$ and $dU = T\,dS - P\,dV + \delta w_{\mathrm{n}PV}$. In the absence of any external force fields and surface effects, Gibbs identified the non-PV work with changes in composition of each phase of the system. Gibbs therefore extended the internal energy function U for pure substances to include composition variables as $U(S, V, n_1, n_2, \ldots, n_c)$, where $n_1, n_2, \ldots, n_C$ represent the minimum number (C) of variables (termed as components) required to characterize the composition of a given phase j (of J phases), specified either as number of moles (n_i), mole fractions ($x_i = n_i/n_{\mathrm{total}}$), or mass fractions. Mathematically, the C variables represent the basis vectors of the compositional space of a phase, and $(n_1, n_2, \ldots, n_C)$ can be written as a compositional vector $\mathbf{n}_i$. Since dU is an exact differential, for the entire multiphase system, $dU = (\partial U/\partial S)_{V,ni}dS + (\partial U/\partial V)_{S,ni}dV + \sum_j \sum_i (\partial U/\partial n_i)_{S,V,\mathbf{n}_j}dn_i$, where the summation is over $j = 1, 2, \ldots, J$ phases and $i = 1, 2, \ldots, C$ components. Comparison to $dU = T\,dS - P\,dV + \delta w_{\mathrm{n}PV}$ implies $(\partial U/\partial S)_{V,ni} = T$, $(\partial U/\partial V)_{S,ni} = -P$, and $\sum_j \sum_i (\partial U/\partial n_i)_{S,V,\mathbf{n}_j}dn_i = \delta w_{\mathrm{n}PV}$.

The intensive partial molar internal energies $\partial U/\partial n_i$ are a measure of how much the internal energy of the system changes if one changed the ith species in any phase by a small amount (dn_i mole) while keeping the amounts of the other species (denoted as vector $\mathbf{n}_j$) together with S, V in the phase constant. For a given phase, Gibbs termed the intensive partial molar quantity ($\partial U/\partial n_i$) the chemical potential (function) of the ith component in the phase and denoted it as μ_i. For a given phase $\mu_i = f(S, V, \mathbf{n}_j)$ represents the slope of the $U(S, V, \mathbf{n}_j)$ space for that phase. Physically, the chemical potential represents the generalized force for defining energy (work) due to changes (displacements) in composition ($\mathbf{n}_j$) of any of the phases due to reaction or mass transfer. Since they represent partial molar quantities, μ_i values depend on the composition of the phase. Chemical potential gradients driving such non-PV work transfer between phases or between system and surroundings are analogous to temperature and pressure gradients that drive energy transfer as heat and PV work.

In order to transform the system coordinates from (S, V) to the more convenient and practical (P, T), one can apply a Legendre transformation on the basic $U(S, V)$ state function in order to eliminate S and V in favor of P and T. Gibbs applied the Legendre transformation to develop the state potential function $G = U + PV - TS$ defined on the five primitive thermodynamic state functions (internal energy U, pressure P, volume V, temperature T, and entropy S). The total differential of this function is $dG = d(U + PV - TS) = dU + P\,dV + V\,dP - T\,dS - S\,dT$. Substituting $dU = T\,dS - P\,dV + \delta w_{\mathrm{n}PV}$ gives $dG = -S\,dT + V\,dP + \delta w_{\mathrm{n}PV}$. Therefore, if some (reversible) process changes the system equilibrium state at constant temperature and pressure, $dG = \delta w_{\mathrm{n}PV}$ and $\int dG = \Delta G = w_{\mathrm{n}PV}$ represents the work exclusive of PV work that must be done on the system (or by the system if ΔG is negative) to achieve the change in equilibrium state. Gibbs termed the state function G as the free energy of the system. Since the state function H (enthalpy) was defined as $H = U + PV$, the free energy function is formally defined as $G = H - TS$, and is therefore alternatively called the free enthalpy. Since for any phase $dG \quad (P, T, \mathbf{n}_i) = (\partial G/\partial P)_{T,\mathbf{n}_i} dP + (\partial G/\partial T)_{P,\mathbf{n}_i} dT + \sum_i (\partial G/\partial n_i)_{P,T\mathbf{n}_j} dn_i$, then by comparison with the above total differentials dG and dU, $(\partial G/\partial P)_{T,\mathbf{n}_i} = V$, $(\partial G/\partial T)_{P,\mathbf{n}_i} = -S$, and $\mu_i = (\partial U/\partial n_i)_{S,V,\mathbf{n}_j} = (\partial G/\partial n_i)_{P,T,\mathbf{n}_j}$ for a given phase.

As the four extensive thermodynamic energy functions (i.e., U, G, H, and A) present the same physical information in different coordinates, Legendre transformations made on $U(S,V,\mathbf{n}_i)$ can be used to develop the other two energy functions, enthalpy $H(S, P, \mathbf{n}_i)$ and Helmholtz free energy $A(T, V, \mathbf{n}_i)$. The partial molar values of these functions give equivalent definitions for $\mu_i = (\partial H/\partial n_i)_{S,P,\mathbf{n}_j} = (\partial A/\partial n_i)_{T,V,\mathbf{n}_j}$. In the case of a single pure substance $\mathbf{n}_j$ has only one element, and the value of μ for a given phase is simply the value of the state energy function divided by the moles of the pure substance comprising the phase.

Extensive thermodynamic equilibrium state functions (the four energy functions included) of the form $f(x_1, x_2, \ldots)$ are first-order homogeneous functions of degree one, meaning that $f(\lambda x_1, \lambda x_2, \ldots) = \lambda f(x_1, x_2, \ldots)$, where λ is an arbitrary constant scaling all *extensive* variables in the basis vector set $\mathbf{x}_i$. Applying Euler's theorem for first-order homogeneous functions in this case would give $f(x_1, x_2, \ldots) = \sum_i (\partial f/\partial x_i)\ x_i$. As an example, applying these concepts for the Gibbs energy function $G(T, P, \mathbf{n}_i)$ gives (since T and P are intensive variables), $G(T, P, \lambda\mathbf{n}_i) = \lambda G(T, P, \mathbf{n}_i)$ and by Euler's theorem $G(T, P, \mathbf{n}_i) = \sum_i n_i (\partial G/\partial n_i)_{P,T,\mathbf{n}_j} = \sum_i n_i \mu_i$. Therefore, $dG = \sum_i n_i d\mu_i + \sum_i \mu_i dn_i = V\,dP - S\,dT + \sum_i \mu_i dn_i$. Subtraction of these two expressions for dG gives $V\,dP - S\,dT + \sum_i n_i d\mu_i = 0$. This result is called the Gibbs–Duhem equation. Similar "Gibbs–Duhem" equations can be derived for any extensive state function of the form $f(x_1, x_2, \mathbf{n}_i)$ using Euler's theorem (although strictly this name refers to the Gibbs function). Such results are essential starting points in the analysis of solution/vapor systems.

For internal equilibrium (meaning nothing changes macroscopically) in any heterogeneous system, it is intuitive (albeit there are formal proofs) that the intensive state functions P, T, and μ_i (the chemical potential of each component) must be uniform throughout the system. Achieving uniformity of μ_i implies that mass transfers between phases occur such that the phase with the smaller chemical potential will grow at the expense of the phase with the larger chemical potential until equilibrium is achieved. At equilibrium since T, P, μ_i are fixed, the system cannot deliver non-PV work and therefore $dG = 0 = \sum_i \sum_j \mu_{ij} dn_{ij}$, where the summation is over $j = 1, 2, \ldots, J$ phases and $i = 1, 2, \ldots, C$ components.

Among thermodynamic intensive state functions, the chemical potential function $\mu(P, T, \mathbf{n}_i)$ serves as a powerful tool in analyzing processes (i.e., changes in equilibrium state) in heterogeneous systems of variable composition, especially those involving phase changes, chemical reactions, or electrochemical and surface phenomena. At equilibrium, $\mu(P, T, \mathbf{n}_i)$ defines a surface (sometimes called the Gibbs surface) representing all possible equilibrium states of the system. Each point on the Gibbs surface represents the minimum value of $G(P, T, \mathbf{n}_i)$ at a given T, P. From the relationships for $G(T, P, \mathbf{n}_i)$ in the preceding paragraphs, and Euler's condition for exactness, $(\partial\mu/\partial T)_{P,\mathbf{n}_i} = \partial/\partial n_i (\partial G/\partial T)_{P,\mathbf{n}_j} = -(\partial S/\partial n_i)_{P,\mathbf{n}_j} = -s$ (partial molar entropy) and similarly $(\partial\mu/\partial P)_{T,\mathbf{n}_i} = (\partial^2 G/\partial n_i\ \partial P)_{T,\mathbf{n}_j} = v$ (partial molar volume) for a given phase.

Gibbs also quantified the minimum number (D) of unconstrained intensive variables (such as T, P, μ_i) required to unambiguously define the equilibrium state of a closed heterogeneous system. Each such variable constituted one of D "internal degrees of freedom" of the system. As discussed above, components (C) are distinct chemical species that can be independently varied and whose total quantities are conserved; phases (J) are the coexisting mechanically separable, distinct, homogeneous subsystems.

Gibbs considered a system containing C components present in J phases designated $1, 2, 3, \ldots, J$. In each phase, in addition to P and T, there are $C - 1$ possible independent composition variables since the sum of the moles in that phase is fixed. The total *possible* number of degrees of freedom describing the system is therefore $2J + J(C - 1) = JC + J$. If the system is in a state of internal equilibrium, P, T, and μ_i are equal in all phases. These constitute restraints that reduce the total possible degrees of freedom. Equality of P and T in each phase provides $2(J - 1)$ restraints. Equality of μ_i in all phases imposes $J - 1$ restraints for each component (meaning that for the ith component one can write $J - 1$ equalities for that μ_i), and therefore $C(J - 1)$ restraints for C components. The difference between the potential degrees of freedom and the restraints gives the actual degrees of freedom of the system. This gives $D = JC + J - (2J - 2 + JC - C) = JC + J - 2J + 2 - JC + C = C - J + 2$. This result is the well-known Gibbs phase equilibrium rule and is universally applicable to all heterogeneous systems at equilibrium. D (sometimes called the variance) is the number of unconstrained variables, that is, intensive quantities that can be independently varied without changing the number of phases. It may be interesting to note that it is no coincidence that this result resembles Descartes's topological formula ($F = 2 + E - V$) relating the number of faces (F), edges (E), and vertices (V) of simple polyhedra. In this context, the phase rule characterizes the topology of the Gibbs surface for the phases at equilibrium.

The foregoing concepts and relationships are *sine qua non* to the understanding and quantifying of all heterogeneous phase equilibrium systems. As an example, let us consider the experimental phase equilibrium for a pure ice/liquid water/water vapor system. It can be represented as the locus of points in a vapor pressure (P in pascal) versus temperature (T in kelvin) coordinate plane representing states for which these phases are in equilibrium. Such a plot, termed a phase diagram, for water is shown in Fig. 1. There are three regions in the phase diagram corresponding to the three phases of water. The phase diagram shows three possible phase transitions liquid ↔ vapor (vaporization), solid (ice) ↔ liquid (fusion or melting), and solid (ice) ↔vapor (sublimation). Applying the phase rule to Fig. 1, all three phases can coexist at only one combination of P and T known as the triple point. At the triple point the system has no degree of freedom, implying that the triple point is unique for all pure substances. The two lines separating the vapor from the liquid and solid phases specify the pressure at which the vapor of the substance is in equilibrium with its liquid or solid phases at a given temperature. For example, along the sublimation line, ice and vapor can coexist in equilibrium. Along the phase equilibrium lines the system has one degree of freedom, and hence the vapor pressure is fixed if T is specified and vice versa. Since the equilibrium vapor pressure for a given system may be established at any temperature, and because vapor formation is endothermic, an increase in temperature will tend to shift the equilibrium state toward higher vapor pressure and vice versa. Therefore, vapor pressure is necessarily an increasing function of the temperature of the system. The temperature at which the vapor pressure exactly equals the atmospheric pressure ($1\ \text{atm} = 1.103 \times 10^5$ Pa) is called the normal boiling point of the substance (100 °C for water). Above the critical temperature (374 °C), it is not possible to produce condensation by increasing the pressure, and the system has two degrees of freedom (i.e., P and T). Similar phase diagrams exist for all pure substances except that the fusion line usually tilts to the right of the triple point. Water is one of few substances (another is bismuth) whose melting temperature is lowered as P increases so that its fusion line tilts to the left of the triple point.

For a single-component phase diagram (as in Fig. 1), application of the foregoing general equilibrium relationships are greatly simplified. A question of practical importance is: What is the relation $P(T)$ between vapor pressure (P) and T along the equilibrium sublimation and vaporization curves? Using subscript α to denote any of the condensed phases (solid, liquid) and the vapor phase as β, then at any point along the phase transition curves $\mu_\alpha = \mu_\beta$. Applying Taylor's theorem to expand $\mu_\alpha\ (T, P)$ and $\mu_\beta\ (T, P)$ in the neighborhood of the same arbitrary point P_0, T_0 [i.e., about $\mu_\alpha\ (T_0, P_0) = \mu_\beta\ (T_0, P_0)$] on the phase transition curve by one gets $\mu_\alpha\ (T_0 + dT, P_0 + dP) = \mu_\alpha\ (T_0, P_0) + (\partial\mu_\alpha/\partial T)_P dT + (\partial\mu_\alpha/\partial P)_T dP \ldots$ equal to the similar expansion for $\mu_\beta\ (T, P)$ about $\mu_\beta\ (T_0, P_0)$. Equating the terms involving dP and dT in these expansions gives $dP/dT = -(\partial\mu_\alpha/\partial T - \partial\mu_\beta/\partial T)/(\partial\mu_\beta/\partial P - \partial\mu_\alpha/\partial P)$. As shown in a previous paragraph $(\partial\mu_\alpha/\partial T)_{P,n\beta} = -s_\alpha$ and $(\partial\mu_\alpha/\partial P)_{T,n\beta} = v_\alpha$, where s and v represent partial molar entropy and volume.

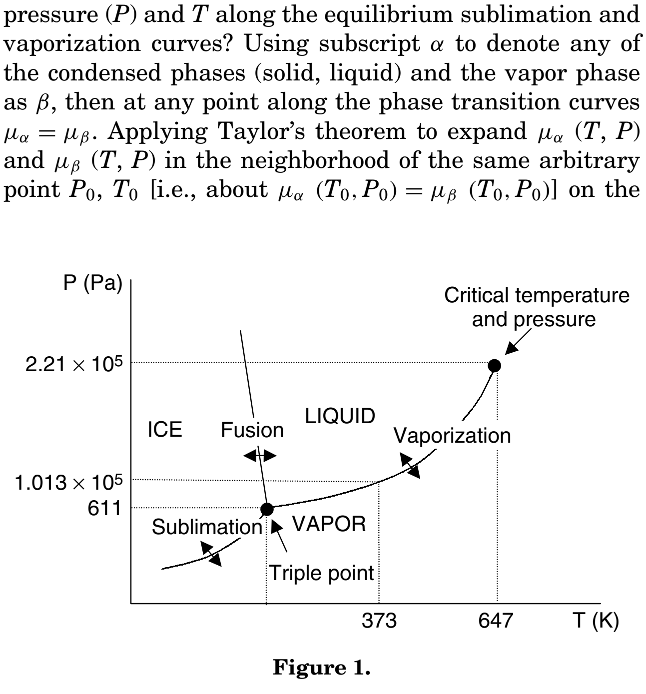

Figure 1.

However, in the special case of only one component, partial molar values of extensive state functions are the same as their molar values, and therefore $-s_\alpha = -S_\alpha$ and $v_\alpha = V_\alpha$, where S_α and V_α denote molar entropy and volume. $-S_\beta$ and V_β are similarly defined, and therefore $dP/dT = (S_\beta - S_\alpha)/(V_\beta - V_\alpha)$. This important general equation for the slope of the phase equilibrium line in phase diagrams for single-component systems was first proposed by French physicist and engineer Emile Clapeyron in 1834. As, by definition, $(S_\beta - S_\alpha) = (H_\beta - H_\alpha)/T$, then $dP/dT = (H_\beta - H_\alpha)/[T(V_\beta - V_\alpha)]$. By convention, extensive property changes (such as ΔS, ΔH, ΔG) for phase transition are taken such that their values are positive. Along the vaporization curve in Fig. 1, $\Delta H_{\text{vap}} = H_{\text{vap}} - H_{\text{liquid}}$ and $\Delta V_{\text{vap}} = V_{\text{vap}} - V_{\text{liquid}}$. Similarly, along the sublimation curve, $\Delta H_{\text{sub}} = H_{\text{vap}} - H_{\text{solid}}$ and $\Delta V_{\text{sub}} = V_{\text{vap}} - V_{\text{solid}}$.

The foregoing reasoning assumes that the gaseous phase consists only of the vapor. If, however, other gases are present (e.g., water vapor in air), this would increase the external pressure on the condensed phase and tend to increase mass transfer to the vapor phase. For this system, $\mu_\alpha\ (P)|_T = \mu_\beta\ (P)|_T$ at equilibrium (for fixed T). Taylor's series expansion, described above, reduces to $dP_\alpha\ (\partial\mu_\alpha/\partial P)_T = dP_\beta\ (\partial\mu_\beta/\partial P)_T$ and therefore $dP_\alpha/dP_\beta = V_\beta/V_\alpha$, where V is the molar volume, P_β is the vapor pressure, and P_α is the total pressure exerted on the condensed phase. This assumes that the only effect of the other gases is increasing the total pressure on the condensed phase, meaning they behave as a closed subsystem with respect to the vapor phase. V_β (molar volume of the vapor phase) is usually $>> V_\alpha$ (molar volume of the condensed phase) and therefore V_α can be assumed constant. Assuming ideal behavior of the vapor phase, $V_\beta = RT/P_\beta$. Substitution and rearranging gives $dP_\beta/P_\beta = (V_\alpha dP_\alpha)/(RT)$. Integration between $(P_\beta)_1$ at $(P_\alpha)_1$ and $(P_\beta)_2$ at $(P_\alpha)_2$ gives $\ln[(P_\beta)_2/(P_\beta)_1] = [V_\alpha/RT][(P_\alpha)_2 - (P_\alpha)_1]$. For water, using $V_\alpha = 18\ \text{cm}^{-3} \cdot \text{mol}^{-1} = 1.8 \times 10^{-7}\ \text{m}^3 \cdot \text{mol}^{-1}$ at 277.16 K (4 °C), $R = 8.314\ \text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$, and 1 atm ($1.013 \times 10^5$ Pa) change in external pressure, we find the right-hand side of the solution $= 0.000791$. This implies the ratio of the vapor pressures $= e^{0.000791} = 1.00079$ and therefore an increase in vapor pressure of 0.079%. This shows that the effect of the presence of other "inert" gases that increase the external pressure on the condensed phase is generally small and can be ignored for most cases of practical interest.

The Clapeyron equation can be integrated exactly if functions $(\Delta H)_{\text{tr}}(T, P)$ and $(\Delta V)_{\text{tr}}(T, P)$ can be defined. Here, the subscript tr denotes either solid/vapor or liquid/vapor transition. Such functions are usually semianalytical at best. They are often applicable over short ranges of T and P and therefore values for pure substances are usually presented as tables (3). These tabulations usually list H, S, and V for the pure phases as values per unit mass obtained by dividing the molar values by the molecular weight in appropriate mass units.

Reasonable approximations can be made if one assumes (1) $V_{\text{vap}} >> V_{\text{liquid}}$ or V_{solid}, (2) ΔH_{tr} is constant over the range of temperature of interest, and (3) the vapor behaves as an ideal gas. In this case, $v_{\text{vap}} = RT/P$, where v_{vap} is the specific volume of the vapor ($\text{m}^3{\cdot}\text{kg}^{-1}$) and R ($\text{J}{\cdot}\text{kg}^{-1}{\cdot}\text{K}^{-1}$) is the specific ideal gas constant ($R = 8.314\ \text{J}{\cdot}\text{mol}^{-1}{\cdot}\text{K}^{-1}$/molecular weight of the vapor in $\text{kg}{\cdot}\text{mol}^{-1}$). After some rearrangement, the Clapeyron equation becomes $dP/dT = d(\ln P)/d(1/T) = -(\Delta h)_{\text{tr}}/R$, where $(\Delta h)_{\text{tr}}$ is the specific enthalpy change ($\text{J}{\cdot}\text{kg}^{-1}$), that is, the latent heat for the phase transition. Integration of this approximation of the Clapeyron equation (called the Clausius–Clapeyron equation) between $P = P_1$ at $T = T_1$ and $P = P_2$ at $T = T_2$ gives $\ln(P_2/P_1) = (\Delta H_{\text{tr}}/R)\,(1/T_1 - 1/T_2)$. Indefinite integration gives $\ln P = [-(\Delta h)_{\text{tr}}/(RT)] + C$ (an integration constant). This implies that if observations of $\ln P$ are plotted versus $1/T$, the (negative) slope of the line multiplied by $(-R)$ gives an estimate of $(\Delta h)_{\text{tr}}$ over a short temperature range of interest. Discontinuities exist as the upper limit of the temperature range approaches the triple point or the critical temperature.

In practical applications over wider ranges of temperature, empirical equations are preferred (4). For water, the Goff–Gratch (5–7) and Magnus–Teten (8) equations for $P(T)$ are commonly used in the atmospheric sciences. The Goff–Gratch function for water is regarded as the standard (9) and is given by $P(T) = ke^{f(\theta)}$, where $\theta = 373.16/T$ and

$$f(\theta) = a\theta + b\ln(\theta) + ce^{d/\theta} + fe^{g\theta}$$

where P is in mb, T is in K, $k = 7.95357242 \times 10^{10}$, $a = -18.1973$, $b = 5.0281$, $c = -70242.1852$, $d = -26.1205$, $f = 58.0692$, and $g = -8.0395$.

The corresponding function for ice is $P(T) = ke^{f(\theta)}$, where $\theta = 273.16/T$ and

$$f(\theta) = a\theta + b\ln(\theta) + c/\theta$$

where P is in mb, T is in K, $k = 5.75185606 \times 10^{10}$, $a = -20.9470$, $b = -3.5665$, and $c = -2.0189$.

The Magnus–Teten function is $P(T) = ae^{f(T)}$ with $a = 6.1078$ (P at triple point) and $f(T) = c_1\ (T - 273.16)/(T - c_2)$, where P is in mb and T is in K. For water $c_1 = 17.2694$ and $c_2 = 35.86$. For ice $c_1 = 21.8746$ and $c_2 = 7.66$

Values of P at 10 °C = 283.16 K for water are 12.272 mb by the Goff–Gratch equation and 12.279 mb by the Magnus–Teten equation. Values of P at −10 °C = 263.16 K for ice are 2.597 by the Goff–Gratch equation. The corresponding value by the Magnus–Teten formula is 2.595 mb.

BIBLIOGRAPHY

1. Gibbs, J.W. (1928). *The Collected Works of J. Willard Gibbs*. Longmans, Green and Co., New York.
2. Wheeler, L.P. (1951). *Josiah Willard Gibbs, the History of a Great Mind*. Yale University Press, New Haven, CT.
3. Lemmon, E.W., McLinden, M.O., and Friend, D.G. (2003). Thermophysical properties of fluid systems. In: *NIST Chemistry WebBook*. P.J. Linstrom and W.G. Mallard (Eds.). *NIST Standard Reference Database*, Number 69, National Institute of Standards and Technology, Gaithersburg, MD. Available: http://webbook.nist.gov.
4. Gibbins, C.J. (1990). A survey and comparison of relationships for the determination of the saturation vapour pressure over plane surfaces of pure water and of pure ice. *Ann. Geophys.* **8**: 859–886.
5. Goff, R.M. and Gratch, S. (1946). Low pressure properties of water from −160 to 212 °F. *Trans. Am. Soc. Heating Ventilation Eng.* **52**: 95–121
6. Goff, J.A. (1957). Saturation pressure of water on the new Kelvin temperature scale. *Trans. Am. Soc. Heating Ventilation Eng.* **63**: 347–354.
7. Fukuta, N. and Gramada, C.M. (2003). Vapor pressure measurement of supercooled water. *J. Atmos. Sci.* **60**: 1871–1875.
8. Murray, F.W. (1967). On the computation of saturation vapor pressure. *J. Appl. Meteorol.* **6**: 203–204.
9. World Meteorological Organization. (1988). *WMO Technical Regulations: General Meteorological Standards and Recommended Practices*, *WMO* No. 49. WMO, Geneva.

ADIABATIC COOLING

Naraine Persaud
Virginia Polytechnic Institute
and State University
Blacksburg, Virginia

When any closed gaseous system expands against an external pressure with no compensatory input of heat from its surroundings to offset the pressure–volume work done by the system, the expansion results in adiabatic cooling of the gas. The vice versa process results in adiabatic heating. The term adiabatic (from the Greek word meaning "not passing through") implies an idealized process and a perfectly insulated system. Real, imperfectly insulated systems would permit some compensatory heat flux. However, if the effect of such fluxes on the system temperature is small relative to the cooling effect, the expansion can be considered as adiabatic for practical purposes.

Adiabatic cooling and heating is the principal mechanism underlying the continual turbulent vertical mixing (autoconvection) in the earth's troposphere, that is, the lowest 9 km (polar regions) to 18 km (tropics) thick atmospheric layer comprising about 80% of the atmospheric mass where all weather processes occur. It is especially important in the planetary boundary layer (i.e., the lowest 1–2 km thick layer) where surface effects on vertical motions are more manifest. Cloud formation and precipitation, observed temperature decrease with height, and energy and mass transfers in the troposphere all depend on adiabatic cooling. This thermophysical process is therefore sine qua non to the continuity of the earth's hydrological cycle. The focus here is on the physics of adiabatic cooling in the troposphere and the resulting temperature decreases with elevation.

A global average of close to 30% of incoming solar short-wave radiation is absorbed at the earth's surface layers. The resulting increase in the temperature of these layers and heat transfer to the air is the most important driver for turbulent vertical convection and associated adiabatic cooling in the troposphere. The earth's atmosphere is a quasiuniform thermophysical system, meaning that pressure gradients rather than the pressure of the atmosphere as a whole can be defined. To facilitate the analysis of tropospheric adiabatic cooling, the concept of an air parcel is introduced. A parcel (or less commonly an eddy) is a discrete closed volume of tropospheric air with uniform properties that is small enough so that the pressure at any point is the same. This implies that, although each parcel has its own individual structure and evolution in time, the observed state and dynamics of tropospheric air is the net result of the motion of these parcels. Because of the many forces involved in such motions, applying Newton's second law to describe the hydrodynamics of these parcels in the earth's atmosphere would result in a complex set of three-dimensional vector differential equations in spherical coordinates. However, only the relatively simpler vertical component of the vector equations is required for analysis of adiabatic cooling. This component is dominated by the vertical pressure and gravitational forces (1).

In order to understand the interplay of these forces, consider a prizm-shaped, initially motionless parcel in a *dry* atmosphere air above the earth's surface as shown in Fig. 1.

It is assumed that tropospheric air is continuously stratified meaning pressure (P) and density (ρ_d) are continuous and decreasing functions of vertical height (z) above the surface. The subscript d refers to dry air. The volume of the prizm (V) is ($\Delta x\ \Delta y\ \Delta z$). The net force on the prizm is $-V\rho_d g$ (downward weight) + $(\Delta x\ \Delta y)(dP/dz)\ \Delta x$ (upward net pressure buoyancy force). Since the prizm is assumed static, the net force per unit mass (acceleration) = 0 and therefore $(1/\rho_d)\ dP/dz + g = 0$. This equation (termed the hydrostatic or aerostatic equation) states the necessary condition for vertical hydrostatic equilibrium in the troposphere, namely, that gravity must be balanced by the vertical component of the atmospheric pressure gradient force.

Any disturbance of this hydrostatic equilibrium would result in vertical motion of the parcel. In particular, if the density of the parcel (ρ_p) is different from that of the surrounding air (ρ_0), the parcel would accelerate. The buoyancy acceleration is $g[(\rho_0/\rho_p) - 1]$ and is positive (upward) if $\rho < \rho_0$ and vice versa (assuming the vertical pressure gradient force is fixed). Substituting $g = 9.81\ \text{m}\cdot\text{s}^{-2}$, a 10% decrease in the density of the parcel would result in an acceleration of close to $1\ \text{m}\cdot\text{s}^{-2}$. Such buoyancy accelerations are the direct result of temperature changes caused by heating (from surface or otherwise) of air parcels. Assuming dry air behaves as an ideal gas $\rho_d = P/R_dT$ [where R_d is the specific gas constant for dry air = $8.314\ \text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}/0.289$ kg (mol dry air)$^{-1}$ and T is the absolute temperature (K)], the buoyancy acceleration can be written as $g[(T/T_0) - 1]$, where T and T_0 are the temperatures of the parcel and surrounding air at a given elevation.

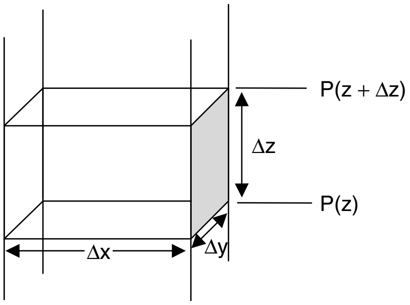

Figure 1. Schematic representation of a retarding basin.

For purposes of analysis of adiabatic cooling, an air parcel is considered as a perfectly insulated bubble of air submersed in the atmosphere, although in reality it could exchange heat with its surroundings. The thermal conductivity of dry air (k_{da} in $\text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$ or $\text{cal}\cdot\text{s}^{-1}\cdot\text{cm}^{-1}\cdot\text{K}^{-1}$) at 1-atm pressure is linearly related to temperature over the range 263–473 K. It increases from $0.023823\ \text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$ ($0.0000568\ \text{cal}\cdot\text{s}^{-1}\cdot\text{cm}^{-1}\cdot\text{K}^{-1}$) at 273.16 K by $0.0000712\ \text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$ ($0.0019\ \text{cal}\cdot\text{s}^{-1}\cdot\text{cm}^{-1}\cdot\text{K}^{-1}$) for each kelvin unit (or °C) change in temperature. The conductivity of moist air is always less than or equal to k_{da}, the difference widening in nonlinear manner with increasing temperature and increasing humidity. Upward acceleration and accompanying expansion of the parcel occur relatively quickly, and because of the low thermal conductivity values of moist air, any heat transfer to the parcel is negligible. This assumption makes possible analysis of the thermal physics of an ascending air parcel as an adiabatic cooling process.

From the integral form of the first law of thermodynamics, a change in internal energy (ΔU) of a parcel equals absorption (gain) or release (loss) of heat (q) minus mechanical work done *by* the parcel (loss) or work done *on* the parcel (gain). Assuming expansion/compression takes place reversibly (i.e., ideally balanced quasistatic equilibrium is always maintained during the process), the mechanical pressure–volume (PV) work done by the parcel against the external pressure (P) is given as $P(\Delta V)$, where ΔV is the volume change (positive for expansion). Thus, in general, $q = \Delta U + P(\Delta V)$. The ideal gas equation of state for unit mass of dry air is $PV = R_dT$. Assuming reversible adiabatic expansion (i.e., process is isentropic), $q = 0$ and $\Delta U = -P(\Delta V) = -R(\Delta T)$ for reversible adiabatic cooling of an ideal gas.

For dry air (a homogeneous gaseous mixture) U in general is a function of absolute temperature (T) and P. However, for an ideal gas, U is a function of T only as demonstrated by James Prescott Joule (1818–1889). Also, U is a state function implying that ΔU is the same for a given temperature change (ΔT) of the parcel, regardless of how ΔT was achieved. Two possibilities are discussed below, namely, changing T at constant volume or at constant pressure.

The specific heat of any substance is the heat input required to cause unit temperature change in unit mass of the substance. In general, specific heat is a function of temperature. The specific heat for gases at constant volume and constant pressure are defined as $C_v(T) = (dq/dT)|_v$ and $C_p(T) = dq/dT|_p$. From the first law, for dry air $C_v < C_p$ since "extra" heat is required at constant pressure for expansion work done against the surrounding pressure. Over tropospheric temperature ranges of practical importance, the specific heats for dry air (denoted as C_{vd} and C_{pd}) can be considered as constants.

Several important results follow from applying the first law to the analysis of adiabatic reversible cooling of a dry air parcel in a quasiuniform atmosphere assuming ideal behavior and constant C_{vd} and C_{pd}. Applying $\Delta U = q - P(\Delta V)$ to the heating of unit mass dry air through ΔT, at constant P, implies that $C_{pd}(\Delta T) - \Delta U = P(\Delta V) = R_d(\Delta T)$. But ideal behavior implies that ΔU can be taken as $C_{vd}(\Delta T)$, and therefore $C_{pd}(\Delta T) - C_{vd}(\Delta T) = R(\Delta T)$, which simplifies to $C_{pd} - C_{vd} = R_d$.

Taking the total differential of the equation of state for unit mass of dry air assuming ideal behavior (i.e., $PV = R_dT$) gives $V\,dP + P\,dV = R_d dT$ or $P\,dV = R_d dT - V\,dP$. Substituting this, together with $R_d = C_{pd} - C_{vd}$, and $dU = C_{vd}dT$ into the differential form of the first law (i.e., $\delta q = dU - P\,dV$) gives $\delta q = C_{vd}dT - V\,dP + (C_{pd} - C_{vd})dT = C_{pd}dT - V\,dP$. The result $\delta q = C_{pd}dT - V\,dP$ is in effect a restatement of the differential form of the first law for dry air.

For an adiabatic ideal gas process $\delta q = 0$ and $C_{pd}dT = V\,dP = [(R_dT)/P]dP$. Separating variables gives $d(\ln T) = (R_d/C_{pd})d(\ln P)$. Integrating between $P = P_0$ at $T = T_0$ and $P = P$ at $T = T$, and rearranging gives $T/T_0 = (P/P_0)^{R_dC_{pd}}$. This result is one of the three adiabatic ideal gas process P–V–T relationships known as Poisson equations (Siméon Denis Poisson, 1781–1840). It describes the quasistatic reversible adiabatic expansion for an ideal gas with constant specific heats from an initial temperature T_0 and pressure P_0 to a new temperature T and pressure P. Usually the ratio $\gamma_d = C_{pd}/C_{vd}$ is defined. Dividing $C_{pd} - C_{vd} = R_d$ by C_{pd} and rearranging gives $C_{pd} = (R_d\gamma_d)/(\gamma_d - 1) = R_d/\kappa_d$ with $\kappa_d = (\gamma_d - 1)/\gamma_d$. For dry air $R_d = 8.314\ \text{J·mol}^{-1}\text{·K}^{-1}/0.029\ \text{kg (mol dry air)}^{-1} = 286.7\ \text{J·kg}^{-1}\text{·K}^{-1}$, $C_{pd} = 1004\ \text{J·kg}^{-1}\text{·K}^{-1}$, and $\kappa_d = 0.286$.

The result $T/T_0 = (P/P_0)^{\kappa_d}$ can be solved for $T_0 = T(P/P_0)^{\kappa_d} = \theta_d$. This definition of θ_d with $P_0 = 1$ bar (1000 mb) is termed the potential temperature. It is the temperature that the parcel initially at T, P would assume when changed adiabatically to 1000 mb. From its definition, $\ln\theta = \ln T + \kappa_d\ [\ln P_0 - \ln P]$ and $d(\ln\theta) = d(\ln T) - \kappa_d\ \text{d}(\ln P)$. Dividing $\delta q = C_{pd}dT - [(RT)/P]dP$ by $C_{pd}T$ gives $\delta q/C_{pd}T = d(\ln T) - \kappa_d d(\ln P)$ and therefore $d(\ln\theta_d) = \delta q/C_{pd}T$. As $\delta q = 0$ for an adiabatic process, $d(\ln\theta_d) = 0$, implying the potential temperature is a parameter that remains invariant during the adiabatic ascent of a parcel. As $T = (\theta_d/P_0)P^{0.286}$, plots of T versus $P^{0.286}$ for different θ_d result in a family of straight lines.

Using $\delta q = C_{pd}dT - V\,dP$ with $\delta q = 0$ for adiabatic cooling, gives $dT/dP = V/C_{pd} = 1/(\rho_dC_{pd})$ because V per unit mass (specific volume) is the reciprocal of the density (ρ_d). As shown above, for hydrostatic equilibrium $dP = -\rho_dg\,dz$. Substituting this result for dP in the previous equation gives $-dT/(\rho_dg\,dz) = 1/(\rho_dC_{pd})$ or $dT/dz = -g/C_{pd} = -(g\kappa_d)/R_d$. Writing this result as $C_{pd}dT = -g\,dz$ shows that adiabatic cooling can be interpreted physically as the dry air parcel trading its internal energy ($C_{pd}dT$) to bootstrap its ascent in the troposphere to acquire gravitational potential energy ($-g\,dz$). This interpretation is appropriate since as already discussed the atmosphere is a quasiuniform thermophysical system.

The temperature of an adiabatically cooled parcel of dry (ideally behaved) air ascending in a tropospheric air layer that is in vertical hydrostatic equilibrium therefore decreases linearly with altitude with slope $= -(g\kappa_d)/R_d$. This important theoretical result is known as the dry adiabatic lapse rate (DALR). For $\gamma_d = 1.4$, $\kappa_d = 0.286$, and together with $R_d = 286.7\ \text{J·kg}^{-1}\text{·K}^{-1}$, $g = 9.81\ \text{m·s}^{-2}$, gives $dT/dz = 0.0098\ \text{K·m}^{-1}$. Assuming a layer of dry air in the troposphere is rapidly and perfectly mixed by vertical convection of a large number of air parcels rising and falling adiabatically, the temperature would be expected to decrease linearly by approximately 1 K (same as 1 °C) per 100 m change in elevation. The layer is isentropic (constant entropy with elevation) and therefore the potential temperature (θ_d) is constant.

The DALR is a theoretical value achievable only in a turbulent well-mixed layer of dry tropospheric air. However, real tropospheric air is always moist because a varying amount of water vapor is always present. To facilitate including this factor in analysis of the behavior of rising moist parcels, the water vapor and dry air are treated as distinct components of a (one-phase) ideal gas mixture. In this case, three independent variables are needed to define an equation of state for moist air, namely, two (usually P,T) of the P–V–T variables together with a (humidity) variable that measures the mass ratio of water vapor in the air. Among several possible measures of this ratio the specific humidity (r_q) is the one traditionally used. It is defined as the mass of water vapor per unit mass of moist air.

Consider a volume (V) of moist air of mass (m_m) = mass water vapor (m_v) + mass of dry air (m_d). The subscripts m, v, d denote properties associated with the moist mixture, vapor, and dry air, respectively. The density $\rho_m = m_m/V = m_v/V + m_d/V = \rho_v + \rho_d$ (i.e., the sum of the partial densities), and therefore specific humidity $(r_q) = \rho_v/\rho_m$. The total pressure P of the volume V of moist air at temperature T (K) would be equal to $P_v + P_d$ (i.e., the sum of the partial pressures). P_v is usually denoted as e, implying $P_d = P - e$. Using ideal gas relations, $e = \rho_vR_vT$ and $P - e = \rho_dR_dT$, where $R_v = R/M_v$, $R_d = R/M_d$ denote the specific gas constants for water vapor and dry air, $R = 8.314\ \text{J·mol}^{-1}\text{·K}^{-1}$ is the universal gas constant, and M is molecular weight ($M_v = 0.018\ \text{kg·mol}^{-1}$ and $M_d = 0.029\ \text{kg·mol}^{-1}$). This gives $\rho_v = e/(R_vT) = (eM_{wv})/(RT)$ and $\rho_d = (P - e)/(R_dT)$ Dividing numerator and denominator of the right-hand side for ρ_v by M_d gives $\rho_v = (0.622\ e)/(R_dT)$, where $0.622 = M_v/M_d$ (usually denoted as ε), and R_d is the specific gas constant for dry air ($= 286.7\ \text{J·kg}^{-1}\text{·K}^{-1}$). Summing ρ_v and ρ_d, rearranging algebraically, and simplifying gives $\rho_m = [P/(R_dT)]\{1 - [e\ (1 - \varepsilon)/P]\}$. This shows that density of moist air is slightly less than the density of dry air $[= P/(R_dT)]$ at a given pressure and temperature, a fact first noted by Sir Isaac Newton in 1717 in his book entitled *Optiks*. The value of the expression in curly brackets increases with decreasing P for a given e. Taking a maximum possible value of $e = 4500$ Pa at mean sea level with $P = 1$ atm $= 1.013 \times 10^5$ Pa, and $\varepsilon = 0.622$ gives $1 - [e\ (1 - \varepsilon)/P] = 0.96$ and implies $\rho_m \approx \rho_d$. Using these results gives, after substituting and simplifying, $r_q = \rho_v/\rho_m = (\varepsilon e)/\{P - [(1 - \varepsilon)e]\} \approx \varepsilon e/P$. The mixing ratio r_w defined as mass vapor per unit mass dry air in the mixture $= \rho_v/\rho_d = (\varepsilon e)/(P - e) \approx r_q$.

Eliminating e between ρ_v and ρ_d, rearranging algebraically, and simplifying gives the equation of state for moist air as $P(T, r_q) = \rho_m R_d T(1 + 0.61\ r_q)$. This is equivalent to the ideal gas equation for dry air with $T(r_q) = T(1 + 0.61\ r_q)$ or with $R_m = R_d(1 + 0.61\ r_q)$. In the former case, $T(r_q)$ is termed the virtual temperature (denoted as T_v) and the moist air equation of state is written as $P = \rho_m R_d T_v$. T_v is the temperature dry air would assume in order to have the same density as moist air with specified values of P, T, and r_q.

The specific heats C_{pm}, C_{vm} for moist air would also be functions of r_q. Partitioning the heat for a temperature change (ΔT) in mass (m_m) of moist air at a given r_q expanding at constant pressure gives $m_m C_{pm}(\Delta T) = m_v C_{pv}(\Delta T) + m_d C_{pd}(\Delta T)$. Dividing by $m_m(\Delta T)$ gives $C_{pm} = r_q C_{pv} + (1 - r_q)C_{pd}$ since r_q is by definition m_v/m_m and $m_d = m_m - m_v$. Similar reasoning gives $C_{vm} = r_q C_{vv} + (1 - r_q)C_{vd}$. Values given in the Smithsonian Meteorological Tables (2) are (in $\text{J·kg}^{-1}\text{·K}^{-1}$): $C_{pv} = 1846$, $C_{pd} = 1005$, $C_{vv} = 1386$, and $C_{vd} = 716$. Substituting these values and simplifying gives $C_{pm} = C_{pd}(1 + 0.84\ r_q)$ and $C_{vm} = C_{vd}(1 + 0.94\ r_q)$. From these $\gamma_m = C_{pm}/C_{vm} = \gamma\,[(1 + 0.84\,r_q)/(1 + 0.94\,r_q)]$, where $\gamma = C_{pd}/C_{vd} = 1.4$ is the ratio of the specific heats for dry air as described previously. Multiplying numerator and denominator of the term in square brackets by $(1 - 0.94\ r_q)$ and neglecting r_q raised to powers >1 gives $\gamma_m \approx \gamma(1 - 0.1\ r_q)$. Similarly, $\kappa_m = R_m/C_{pm} = [R_d(1 + 0.61 r_q)]/[C_{pd}(1 + 0.84\,r_q)]$. Multiplying numerator and denominator of the term in square brackets by $(1 - 0.84\ r_q)$ and neglecting r_q raised to power >1 gives $\kappa_m \approx \kappa_d(1 - 0.23\ r_q)$, where $\kappa_d = R_d/C_{pd}$.

γ_m is lower for moist air than dry air but does not affect the DALR appreciably as long as the air does not become saturated. For example, using the foregoing results gives the unsaturated moist adiabatic lapse rate (MALR) as $dT/dz = -(g\kappa_m)/R_m$, where $\kappa_m = (\gamma_m - 1)/\gamma_m$ and $R_m = R_d(1 + 0.61\,r_q)$. At 20 °C and 50% saturation the vapor pressure ≈ 1170 Pa. Using the equations for ρ_v and ρ_m with $R_d = 286.7\ \text{J·kg}^{-1}\text{·K}^{-1}$ and $P = 1.013 \times 10^5$, Pa gives $r_q = \rho_v/\rho_m = 0.0072$ kg vapor (kg moist air)$^{-1}$, $\gamma_m = 1.4 \times 0.999 = 1.399$, $\kappa_m = 0.285$, $R_m = 288.0\ \text{J·kg}^{-1}\text{·K}^{-1}$, and MALR $= -0.0097\ \text{K·m}^{-1}$ compared with $-0.0098\ \text{K·m}^{-1}$ for the DALR. As was done for dry air, the potential temperature invariant for moist unsaturated air (θ_m) can be defined as $\theta_m = T[P/P_0]^{\kappa_m}$, where as discussed above, $\kappa_m \approx \kappa_d(1 - 0.23\ r_q)$ with $\kappa_d = R_d/C_{pd}$.

If a rising moist air parcel becomes cooled to the dew point, condensation occurs (if condensation nuclei are present). The elevation at which this occurs is called the lifting condensation level (LCL). The dew point temperature at the LCL (T_{dL}) is different than at the surface because, although $r_q(\approx \varepsilon e/P)$ remains the same, the vapor pressure decreases as the ascending parcel expands. The dew point (T_{d0}) at the surface depends on the observed ambient temperature (T) and relative humidity $= (e/e_s)|_T$ of the air. The relationship can be obtained using the Clausius–Clapeyron equation $de_s/e_s = (L\,dT)/(R_v T^2)$ for the slope of the equilibrium saturation vapor pressure (e_s) versus temperature curve, where L is the latent heat of vaporization and $R_v (= 461.8\ \text{J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1})$ is the specific gas constant for water vapor (3). L (J·kg^{-1}) is usually a weak function of temperature approximated as $L = 3.071 \times 10^6 - 2134\ T$ K. L is usually taken as a constant value $= 2.5 \times 10^6\ \text{J·kg}^{-1}$ corresponding to $T = 273$ K. Integrating the Clausius–Clapeyron equation between $e_s = e$ at T to $e_s = e_s$ at T_d gives (after substituting for L and R_v), $T - T_{d0} = -(1.845 \times 10^{-4} T)T_{d0} \ln[(e/e_s)_T]$. When required, reasonable estimates of $e_s(T)$ can be obtained by integrating the Clausius–Clapeyron equation on the half-open interval using the triple point of water ($e_s = 611$ Pa at $T = 273.13$ K) as the lower bound. After substituting numerical values for L and R_v, rearranging and simplifying the result is $\ln e_s$(in Pa) $= \ln 611 + [19.83 - (5417/T)]$.

The foregoing implies T_{dL} depends on the surface T_{d0} and on the dew point lapse rate (DPLR). Logarithmic differentiation of $r_{qs} = \varepsilon e_s/P$ gives $dr_{qs}/r_{qs} = de_s/e_s - dP/P = 0$ and therefore $de_s/e_s = dP/P$. Substituting the Clausius–Clapeyron equation for de_s/e_s and the aerostatic equation dP/P gives $L/[R_v(T_d)^2]\,dT_d = (-g/R_d T)dz$, where T_d is the dew point of the ascending parcel at elevation z and T is the corresponding air temperature, and $R_d \approx R_m$. Substituting numerical values $L = 2.5 \times 10^6\ \text{K·kg}^{-1}$, $R_v = 461.8\ \text{J·kg}^{-1}\text{·K}^{-1}$, $R_d = 286.7\ \text{J·kg}^{-1}\text{·K}^{-1}$, and $g = 9.81\ \text{m·s}^{-1}$ gives dT_d/dz (in $\text{K} \cdot \text{m}^{-1}$) $= -(T_d)^2/(0.158 \times 10^6)T$. The LCL occurs when T by cooling at the DALR $= T_d$ by cooling at the DPLR. Therefore, the LCL is obtained by solving $T_0 - 0.0098\,\text{LCL} = T_{d0} - [(T_d)^2/(0.158 \times 10^6 T)]$ LCL, where T_0 is the surface air temperature, which gives LCL $= (T_0 - T_{d0})/\{0.0098 - [(T_d)^2/(0.158 \times 10^6 T)]\}$. T_d and T are functions of elevation but in practice are considered equal to T_{d0} and T_0, respectively, in estimating the LCL. For example, given $T_0 = 20\,^\circ\text{C}$ (293 K) and $T_{d0} = 1\,^\circ\text{C}$ (274 K), the LCL $= 19/\{0.0098 - [(274)^2/(0.158 \times 10^6 \times 293)]\} = 2.323$ km. Alternatively, it has been shown that $[(T_d)^2/(0.158 \times 10^6 T)]$ ranges from about 1.7 to 1.9 K·km^{-1} (4). Using an average value of 1.8 K·km^{-1}, the LCL can be estimated as LCL (km) $= (T_0 - T_{d0})/8$. Using this alternate formula in the above example gives LCL $= 19/8 = 2.38$ km.

For ascent above the LCL, latent heat of condensation is released into the parcel, but the parcel remains saturated at the given temperature. A change $d(r_{qs})$ in r_{qs} (where the subscript s denotes saturation) would result in latent heat released into the parcel $= -L\,d(r_{qs})$, where L denotes the latent heat of vaporization. As was done previously, substitution of the hydrostatic equation $dP = -\rho_s g\,dz$ into the first law as $\delta q = C_{ps}dT - V\,dP$ gives (for unit mass) $\delta q = C_{ps}dT + g\,dz$. Since no water is lost from the parcel $\rho_s = \rho_m = P/R_m T$. Substituting $-Ld(r_{qs})$ for δq gives

$C_{ps}dT + g\,dz = -L\,d(r_q)$. Assuming almost all of the heat released by condensation goes to heating the dry air component in the parcel would imply $C_{ps} \approx C_{pd}$. Making the substitution, dividing throughout by $C_{pd}dz$, and rearranging gives $dT/dz = -(L/C_{pd})(dr_{qs}/dz) - g/C_{pd}$. This is the saturated adiabatic lapse rate (SALR). Since r_{qs} is a function of T and P, dividing its total differential (dr_{qs}) by dz gives $dr_{qs}/dz = [(\partial r_{qs}/\partial P)]_T(dP/dz) + [(\partial r_{qs}/\partial T)]_P(dT/dz)$. Substituting this result with $dP/dz = -\rho_m g$ into the SALR expression, rearranging, and simplifying gives

$$\text{SALR} = \frac{-g}{C_{pd}}\left(\frac{[1 - \rho_m L(\partial r_{qs}/\partial P)_T]}{[1 + (L/C_{pd})(\partial r_{qs}/\partial T)_P]}\right)$$

Differentiating $r_{qs} \approx \varepsilon e_s/P$, where $e_s(T)$ is the saturation vapor pressure gives $[(\partial r_{qs}/\partial P)]_T = -(\varepsilon e_s)/P^2$ and $[(\partial r_{qs}/\partial T)]_P = (\varepsilon/P)(de_s/dT)$. Assuming constant L and ideal gas behavior, the Clausius–Clapeyron equation (3) for the slope of the saturation vapor pressure versus temperature curve is $de_s/dT = Le_s/(R_v T^2)$, where R_v is the specific gas constant for water vapor. Substitution of these results together with $\rho_m \approx \rho_d = P/R_d T$ into the equation for the SALR and simplifying gives a more practical form as (4):

$$\text{SALR} = \frac{-g}{C_{pd}}\left(\frac{\left(1 + \dfrac{Lr_{qs}}{R_d T}\right)}{\left(1 + \dfrac{L^2 r_{qs}}{C_{pd} R_v T^2}\right)}\right)$$

The value of $(Lr_{qs}/R_d T)$ is usually small. In many applications this term is neglected, resulting in the simpler expression usually encountered in atmospheric science literature. The SALR represents the saturated lapse rate above the LCL. It is assumed that the system remains closed, that is, r_q remains the same (all the water and heat from condensation remains within the ascending parcel), and the ascent continues as a reversible saturated-adiabatic process. However, the system would no longer remain closed if (as may be expected) some or all of the water and heat from condensation at a given elevation is immediately released from the parcel. This would imply an irreversible and nonadiabatic process. However, the heat lost is small relative to the heat content of the parcel and the ascent in this case is termed as pseudoadiabatic (meaning "as if it were adiabatic"). The SALR given above is considered equal to the pseudoadiabatic rate for all practical purposes.

However, the SALR is considerably lower than the DALR $\approx$ MALR and is a function of temperature. The SALR is about 4 $\text{K}\cdot\text{km}^{-1}$ for humid air masses near the surface. It increases to about 9 $\text{K}\cdot\text{km}^{-1}$ as the humidity decreases with increasing elevation. An average value is about 5.4 $\text{K}\cdot\text{km}^{-1}$. Like its counterparts θ_d and θ_m for dry and unsaturated air, the potential temperature invariant for saturated adiabatic ascent (θ_e termed the equivalent potential temperature) is defined as $d(\ln\theta_e) = \delta q/(C_{pd}T) = (-L\,dr_{qs})/(C_{pd}T) = -d[(Lr_{qs})/(C_{pd}T)]$. Integrating (on the half-open interval with lower bound $\theta_e = \theta_d$ at $r_{qs} = 0$) and exponentiating gives $\theta_e = \theta_d e^{[Lr_{qs}/(C_{pd}T)]}$. Physically, θ_e is the temperature attained if moist air at temperature T is expanded pseudoadiabatically, until all the water has condensed out and is removed from the parcel, and then is compressed adiabatically to the standard pressure of 1000 mb.

The theoretical lapse rates developed in the foregoing discussion serve as reference values for inferring tropospheric conditions from the observed environmental lapse rate (ELR) at a given location. Departures of the ELR from these reference lapse rates provide a measure of the stability (i.e., the autoconvective propensity or the tendency for vertical motions induced by density stratification) in a given layer of tropospheric air.

Below the LCL, when the ELR = DALR $\approx$ MALR (lapse rate is adiabatic) a rising parcel will always be at the same temperature as its surroundings. Therefore, it would have the same density and would neither tend to return nor continue its displacement. The stability of a tropospheric layer with these conditions prevailing is termed neutral. If the ELR $<$ DALR (lapse rate is subadiabatic) an ascending parcel will quickly become colder (more dense) than the surrounding air and will return to its original position. The layer is termed absolutely stable. Stable air resists vertical movement. If the reverse is true and the ELR $>$ DALR (lapse rate is superadiabatic) an ascending parcel will become warmer (less dense) than the surrounding air and will continue to rise. Except in the heated layer next to the surface, such instabilities do not persist since they are rapidly counteracted by vigorous autoconvection and vertical mixing as fast as they develop. In terms of θ_d, these stability conditions correspond to stable ($\theta_d < 0$), neutral ($\theta_d = 0$), and unstable ($\theta_d > 0$). In a similar manner, saturated air parcels will be stable, neutral, or unstable with respect to vertical displacements depending on whether the ELR is $<$, $=$, or $>$ the SALR or $\theta_e <, =, \text{or} > 0$.

The tropospheric lapse rate for the standard atmosphere (5) is $-6.5\,^\circ\text{C}\cdot\text{km}^{-1}$. Parameters for the standard atmosphere have been specified by the International Civil Aeronautical Organization (ICAO) and are needed for standardization in aviation, aeronautics, and meterology. Using subscript sa to denote the standard atmosphere, solving $-0.0065\ \text{K}\cdot\text{m}^{-1} = -[g(\gamma_{sa} - 1)/(R_d\gamma_{sa})]$ for γ_{sa}, with $g = 9.81\ \text{m}\cdot\text{s}^{-2}$ and $R_d = 286.7\ \text{J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$, gives $\gamma_{sa} = 1.23$ compared to the value of $\gamma_d = 1.4$ for dry air. Other parameters for the standard atmosphere are surface temperature = 288.15 K, pressure = 101, 325 Pa, density = $1.225\ \text{kg}\cdot\text{m}^{-3}$, and mean molecular weight = $28.964\ \text{g}\cdot\text{mol}^{-1}$.

The standard lapse rate implies $T(z) = T_0 - \beta z$, where $T(z)$ is the temperature (K) at elevation z (m), T_0 is the surface temperature, and $\beta = 0.0065\ \text{K}\cdot\text{m}^{-1}$. Assuming ideal gas behavior, substituting $T = T_0 - \beta z$ into the aerostatic equation $dP/dz = (-Pg/RT)dz$, separating variables, and integrating on a half-open interval with lower bound $P = P_0$ at $T = T_0$ gives, after some rearranging, $\beta z/T_0 = 1 - (P/P_0)^{R\beta/g}$. These relationships can be used to calculate P, T of tropospheric air for the standard atmosphere at any given elevation. As an example, for $z = 10$ km and $T_0 = 288.15$ K, $T = 288.15\ \text{K} - (0.0065\ \text{K}\cdot\text{m}^{-1} \times 10^4\text{m}) = 223.2$ K. Substituting $T_0 = 288.15$ K, $P_0 = 1.01325 \times 10^5$ Pa, $R = 286.7\ \text{J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$, and $\beta = 0.0065\ \text{K}\cdot\text{m}^{-1}$ gives P at 10 km and $P = 0.26$ atm with $T = 223$ K.

BIBLIOGRAPHY

1. Brown, R.A. (1990). *Fluid Mechanics of the Atmosphere*. Academic Press, San Diego, CA.
2. List, R.J. (1971). *Smithsonian Meterological Tables*, 6th Edn. Smithsonian Institution Press, Washington, DC.
3. Salby, M.L. (1996). *Fundamentals of Atmospheric Physics*. Academic Press, San Diego, CA.
4. Tsonis, A.A. (2002). *An Introduction to Atmospheric Thermodynamics*. Cambridge University Press, Cambridge, UK.
5. NOAA. (1976). *The U.S. Standard Atmosphere*. U.S. Government Printing Office, Washington, DC (available from National Technical Information Office, Springfield, VA, product number ADA-035-6000).

RAIN AND ROCKS: THE RECIPE FOR RIVER WATER CHEMISTRY

Mohammed Riajul Islam
University of Idaho
Moscow, Idaho

INTRODUCTION

The dissolved and suspended loads of most of the world's major rivers are dominated by weathering processes. Particularly, chemical weathering of rocks and soils is one of the essential processes in the geochemical cycling of elements in rivers (1,2) as well as in the continent–river–ocean system (3). The climatic factors such as moisture and temperature mainly affect the weathering processes (2,4). The most soluble ions that will end up in the ocean through rivers are released because of interactions between water and rocks at the earth's surface (5–7). Weathering's importance has led to a huge number of detailed geochemical studies on the world's major river systems, including the Ganges-Brahmaputra (3,8), other major Himalayan rivers (9,10), the Amazon (11), the Mackenzie (12), rivers of the Canadian Shield (5,12), the Congo (13–15), some large Chinese rivers (16,17), the Mississippi, Indus, and some other U.S. rivers (18,19), the Loire River in France (20), and the upper Niger river (21).

Aerosols are the primary source of chemical elements transported by rainwater (22), which are removed from the atmosphere by rainout or by washout. Atmospheric aerosols (sea salt, crustal dust, and biogenic aerosols) have been found as the major sources of dissolved species in river water (14,23,24). Dissolved components from rainwater come from those originating from sea salt aerosols; from those originating from terrestrial aerosols (soil dust, biological emissions); and from those originating from anthropogenic sources (industry, agriculture, burning of fossil fuels, etc.) (25).

Along with the geological characterization of the watershed, meteorological parameters are critically important in evaluating rainfall effects, as demonstrated by Grosbois et al. (20). These investigators have demonstrated, in accordance with Negrel and Roy (25), that the Loire River basin was characterized by its diversity of air-mass trajectories divided into four major origins: (1) westerly, which originates from the Atlantic (26); (2) northeasterly to northwesterly, which originates from the North Atlantic and North Sea and passes over Great Britain and the industrialized countries from France to Eastern Europe; (3) easterly, with a continental origin and a geochemical signature influenced by polluted countries and a large forested area; and (4) southeasterly to southwesterly, which originates in the Mediterranean Sea and is influenced by natural Saharan aerosols and pollution from Spain.

The radiogenic isotope ^{87}Sr is produced because of the radioactive decay of ^{87}Rb isotope. The existence of ^{87}Sr—which is often expressed as the ^{87}Sr/^{86}Sr ratio in a given geological material—is a function of its age and Rb/Sr ratio. As a result, Sr released by aerosols having different origins results in different ^{87}Sr/^{86}Sr ratios in rainwater. The Sr variations are caused mainly because of the mixture of different sources with different Sr isotope compositions (25). Sr isotope ratios in precipitation have been reported by a wide number of investigators from various parts of the world (27–30).

METHODOLOGY OF STUDYING ROCKS–RECHARGE INTERACTIONS AND RAINWATER CHEMISTRY

Geology of the Study Areas

Consideration of geology is crucial in assessing the link between river water chemistry and local bedrock geochemistry, as demonstrated by English et al. (31) in their Seti river studies in the Himalayan area, which receives high rainfall (32). Over 70% of the TDS in the Seti River is derived from carbonate weathering in the Tethyan sedimentary series/Greater Himalayan series and upper Nawakhot Group of the Lower Himalayan as bedrock and sediment control riverine TDS and Sr. The investigators also found certain rock types in the river bedload that have been derived from related bedrock because of weathering. Furthermore, tributary waters display unique characteristics directly related to the local geology (i.e., bedrocks, climate, hydrology, etc.). The river watersheds, tributaries, local rainfall, humidity, and temperature data must be considered for the study of rocks–recharge interactions as well as river water chemistry. Table 1 shows the effects of various local bedrocks on river water chemistry in different regions.

Sampling

Systematic sampling of precipitation (25), river and tributary waters, bedrock, and river bed sediments would identify geochemical controls on stream chemistry (3,9,10,20,31).

The dissolved elements in stream waters provide information on the source of the atmospheric contribution (36). The marine contribution decreases with increasing distance from the ocean, whereas terrestrial contributions become dominant. However, it is very difficult to identify different kinds of sources (e.g., terrestrial aerosols, anthropogenic aerosols), unless isotopes (e.g., ^{87}Sr/^{86}Sr)

Table 1. Chemical Compositions of Surface Water in Various Rivers, μmol/L

Rivers (Ref.)	Location	pH	TDS, mg/L	Ca^{2+}	Mg^{2+}	Na^{+}	K^{+}	Cl^{-}	SO_4^{2-}	NO_3^{-}
Yaou catchment	Precipitation—dry season	6.1	3.8	5	4.2	33	5.1	23.1	6.4	6.4
(French Guiana) (33)	Precipitation—wet season	6.1	1.7	3	2.9	9.1	3.1	8.7	3.7	0.6
Ganges, Bangladesh (34)	Rajshahi—August	7.6	125	466	201	175	69	94	76	0.3
	Rajshahi—March	8.5	374	1418	612	619	110	234	64	2.7
Brahmaputra, Bangladesh (34)	Chilmari—August	7.6	105	393	168	104	52	25	55	bdl[a]
	Chilmari—March	8.5	145	590	214	196	75	53	149	1.6
Trisuli, India (34)	Betrawati—April	—	70	324	49	67	39	19	46	bdl
	Betrawati—August	—	107	418	136	200	43	49	133	1.9
Nyong basin river, Cameroon (35)	Ayos	5.5	—	1314	975	758	1809	829	966	bdl
	Mefou	6.5	—	3856	1097	3677	4833	3295	2994	1195
Seti River, Nepal (31)	Seti above Liyangwan	—	87	531	222	82	44	40	127	—
	Seti below Taru Gad	—	91	501	300	80	38	29	122	—

[a] bdl, below detection limit.

are obtained from atmospheric contributions. The major and trace element data of sediments does provide clues to source area (37).

For rainwater collection, an automatic precipitation sampler is used (25). The basic requirements for the collector are automatic detection and collection of rainfall, dry fallout elimination, collection of frozen precipitation, and avoidance of sample contamination, particularly for trace element analysis.

Chemical Analysis: Rainwater

Rainwater is analyzed for conductivity, temperature, and pH—all determined on site. Electrical conductivity and water temperature are determined with a microprocessor conductivity meter. pH is measured with any standard pH meter. For major and trace elemental analyses, AAS, ICP-MS, and IC (ion chromatography) are used. Bicarbonate is analyzed by HCl titration and Gran's method (25).

Chemical separation and mass spectrometry procedures for Sr followed the standard method used at BRGM. Strontium is separated by using a cation exchange column (i.e., DOWEX AG50X8) with 2 *N* HCl as eluant. After chemical separation, one-fifth of each sample is analyzed using a Finnigan MAT 262 multiple collector mass spectrometer. The $^{87}Sr/^{86}Sr$ ratios are normalized to $^{86}Sr/^{88}Sr = 0.1194$.

Bedrock and river sediments can be analyzed for major and trace elements using XRF, ICP-MS, and AAS.

RESULTS AND INTERPRETATIONS

Meteorological data are taken generally from local meteorological offices. Average air mass trajectories and weather patterns enable one to identify air mass origins. For example, Negrel and Roy (25) demonstrated the limits of the four air-mass trajectories and four directions of the air-mass movements in the Massif Central area in France. Rainwater chemistry and Sr isotope values are determined for the evaluation of rainwater. The mean pH value of rainwater is 5.22 ± 0.65 (ranges 4.29–6.2) as determined in Massif Central, France (25), whereas the pH of pristine rainwater is moderately acidic with a theoretical value of 5.7 (4). Any higher values reported in France could be the result of dissolution of windblown dust with high $CaCO_3$ content. Rainwater often has a pH value below 5.7 because of the presence of H_2SO_4, weak organic acids, anthropogenic emission of H_2SO_4, and/or HNO_3. The concentrations of elements in rainwater vary greatly with the variation of sampling period (25). In France, the ion imbalance between cations and anions in rainwater is well studied (26,38,39). Significant temporal variation in $^{87}Sr/^{86}Sr$ ratios is also present (709198 to 0.713141) and is particularly associated with changes in seasons. Table 2 indicates the influence of meteorological and anthropogenic effects on rainwater chemistry.

Quantification of Marine and Terrestrial Contributions Using Concentration Data

A sodium (Na) reference (hereafter referred to as Y_{ref}) is used to determine sea salt (ss) and terrestrial contributions in rainwaters (rw) (46–48). To calculate the contribution of the sea salt component (ss) with seawater (sw) characteristics, the following equation is used:

$$X_{\mathrm{ss}} = Y_{\mathrm{ref}} x (X/Y)_{\mathrm{sw}} \quad (1)$$

The contribution of the terrestrial component (t) is the difference between the total composition of rainwater (rw) and the sea salt (ss) contribution:

$$X_{\mathrm{t}} = X_{\mathrm{rw}} - X_{\mathrm{ss}} \quad (2)$$

In calculating the contribution for a given element *X*, the main obstacle to resolve is the selection of the reference species (marine or terrestrial). An element to be used as a marine or terrestrial reference must be free from fractionation during aerosol formation, transportation, or deposition by rain impaction. In addition, the element must be absolutely of marine or terrestrial origin.

Table 2. Chemical Compositions of Representative Rainwater in Various Regions, μequiv/L

Regions (Ref.)	pH	Cl^-	SO_4^{2-}	NO_3^-	NH_4^+	Ca^{2+}	Mg^{2+}	Na^+	K^+
Buyukcekmece, Istanbul (mean) (40)	5.36	215.11	231.89	78.41	16.41	473.6	259.84	113.13	73.6
Patras, Greece (41)	5.16	114.3	46.1	19.4	16.3	98.5	30.4	90.2	6.6
Montseny, Spain (42)	6.4	28.4	46.1	20.1	22.9	57.5	4	9.8	22.3
Sardinia, Italy (43)	5.18	322	90	29	25	70	77	252	17
Silent Valley, India (mean) (44)	—	43	20	21	3	43	14	46	4
Morvan, France (45)	5.25	14.7	29.4	18.8	57	12.4	6.4	13.3	—

Generally, Na is used as a reference element for sea salt and Mg for coastal or marine rains (4). However, both Na and Mg contents in rainwater can be influenced by nonmarine inputs (48–50). Chlorine could be used as a reference if the Cl/Na ratios of rainwater do not differ from those of seawater (1.16 mol/mol) (4).

Calcium comes mainly from a natural terrestrial source, and a marine source may not be significant for K and Sr. Nitrate (NO_3) and SO_4 come mainly as a result of human activities (e.g., fuel combustion) (4). However, major element chemistry cannot characterize perfectly the various terrestrial end members. For this reason, Sr isotope compositions are used to get additional information regarding the different possible terrestrial sources and their respective proportions.

To determine the possible terrestrial inputs, Losno et al. (38) used the Al content in rainwater to determine Na from terrestrial input (t), according to the following equation:

$$\mathrm{Na_t} = \mathrm{Al_{rw}}(\mathrm{Na/Al})_\mathrm{t} \tag{3}$$

Strontium Isotope Systematics

From the mixture of sea salt and terrestrial dust, the terrestrial isotope ratio $(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})_\mathrm{t}$ in each rainwater sample can be calculated using Sr isotope systematics (30). The binary mixing equation for these two end members is demonstrated by

$$(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})_\mathrm{rw} = \alpha(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})_\mathrm{sw} + (1-\alpha)(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})_\mathrm{t} \tag{4}$$

where $(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})_\mathrm{rw}$ is the measured isotope ratio in rainwater, $(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})_\mathrm{sw}$ is the isotope ratio of seawater (0.70917 ± 0.00001), $(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})_\mathrm{t}$ is the crustal isotope ratio, and α is $\mathrm{Sr_{marine}}/\mathrm{Sr_{terrestrial}} + \mathrm{Sr_{marine}}$ (determined in Eqs. 1–3).

Hence, we obtain

$$(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})_\mathrm{t} = \{(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})_\mathrm{rw} - \alpha(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})_\mathrm{sw}\}/(1-\alpha) \tag{5}$$

The terrestrial ratios in the rain samples significantly vary, which is indicative of terrestrial aerosol heterogeneities because isotopic variations in the crustal source demonstrate the different terrestrial components with different isotope ratios in the mixture. The isotopic composition of Sr in dusts can be calculated by studying $^{87}\mathrm{Sr}/^{86}\mathrm{Sr}$ ratios of river water. For example, on a carbonate watershed, the river water's $^{87}\mathrm{Sr}/^{86}\mathrm{Sr}$ ratio is the same as that of the rocks (25).

Identification and Quantification of Natural and Anthropogenic Inputs in River Water

Attention must be given to the different contributions of the different end-member inputs to the total dissolved load, as any calculation of the export rate of chemical species (natural or anthropogenic origin) requires the quantification of each end member. This problem can be solved using a mass balance calculation.

In general, a mass balance equation for the budget can be expressed as follows (51,52):

$$\begin{aligned} &\text{rock weathering} + \text{atmospheric input} \\ &\quad + \text{human activities input} \\ &\quad\quad = \text{dissolved load} + \text{solid load} \end{aligned}$$

Determination of an output–input budget (53,54) and determination of element origins are required for the mass balance approach. The mass balance approach may evaluate the contribution of atmospheric inputs, anthropogenic impacts, and rock weathering processes.

Atmospheric Inputs Correction

Dissolved species in surface water may contain a significant amount of solutes from rainfall (51). The target of the atmospheric input correction is to quantify and subtract the portion of the elements carried by rainwater in the chemical composition of the river water. Knowledge of the chemical composition of the rainfall on the total watershed is required for the quantification of atmospheric input because of rainfall (24,51).

Classically (51), for any element Z in the river $[(\mathrm{Z})_\mathrm{r}]$, correction of the atmospheric contribution to a river (r) is estimated by reference to the Cl concentration, called $(\mathrm{Cl})_\mathrm{ref}$, multiplied by the Z/Cl ratios of rainwater (rw):

$$(\mathrm{Z})_\mathrm{rw} = (\mathrm{Cl})_\mathrm{ref} \cdot (\mathrm{Z/Cl})_\mathrm{rw} \tag{6}$$

$$(\mathrm{Z})_\mathrm{r,rw\ corrected} = (\mathrm{Z})_\mathrm{initial\ r} - (\mathrm{Z})_\mathrm{rw} \tag{7}$$

Atmospheric chloride ions originate from sea salt and human activities. Except for salt rock, Cl release because of rock weathering has not been demonstrated clearly (55). Moreover, through the hydrological cycle, Cl behaves conservatively (55), so it is used as a reference of atmospheric inputs in many unpolluted hydrosystems (3,56). However, chloride ions can also have an anthropogenic origin (57). For the mass balance equations, the number of Cl ions issued from rainwater, known as $(\mathrm{Cl})_\mathrm{ref}$, has to be determined. The $(\mathrm{Cl})_\mathrm{ref}$ is calculated with each mean weighted chloride content for the sampling spots multiplied by the concentration factor F of each region. The concentration effect of evapotranspiration is represented by the factor F and is related to total amount of rainwater P (in mm) and the evapotranspiration process E (in mm) by the equation

$$F = P/(P - E) \tag{8}$$

The general equation to calculate $(\mathrm{Cl})_{\mathrm{ref}}$ is

$$(\mathrm{Cl})_{\mathrm{ref}} = (\mathrm{Cl})_{\mathrm{SA}} \cdot 0.25 \cdot \mathrm{F}_{\mathrm{SA}} + (\mathrm{Cl})_{\mathrm{MC}} \cdot 0.75 \cdot \mathrm{F}_{\mathrm{MC}} \tag{9}$$

where $(\mathrm{Cl})_{\mathrm{SA}}$ and $(\mathrm{Cl})_{\mathrm{MC}}$ correspond to the weighted concentration of chloride in study areas.

The residual chloride $(\mathrm{Cl})_{\mathrm{res}}$ in the river is attributed only to human activities:

$$(\mathrm{Cl})_{\mathrm{res}} = (\mathrm{Cl})_{\mathrm{measured}} - (\mathrm{Cl})_{\mathrm{ref}} \tag{10}$$

Budgets of Human Activities

When the influence of rainwater is subtracted from the river water composition, the general mass balance equation indicates that the dissolved load carried out by the river results from rock weathering and human activities. It is needed to quantify the influence of anthropogenic inputs in order to evaluate the chemical weathering on the river watershed (20).

It is widely known that phosphorus and nitrogen forms and potassium and chloride ions are mainly derived from agricultural fertilizers, animal waste, and municipal and industrial sewage (58–60). These species are indicative of both population and agricultural impacts, whereas sulfates are bacterial oxidation products of some industrial products (61,62).

The contributions of the different end members to the dissolved load can be calculated using the estimation of chemical signature of the pure end members. For each chemical species, the contributions for the silicate (SIL), carbonate (CAR), and anthropogenic (ANT1, ANT2) end members are calculated for each sample of river water using the following equation, as demonstrated by Grosbois et al. (20):

$$[\mathrm{X}]_{\mathrm{riv}} = a[\mathrm{X}]_{\mathrm{SIL}} + b[\mathrm{X}]_{\mathrm{CAR}} + c[\mathrm{X}]_{\mathrm{ANT1}} + d[\mathrm{X}]_{\mathrm{ANT2}} \tag{11}$$

Conditionally, $a + b + c + d = 1$, where X represents the amounts (in μmol/L) of X in the SIL, CAR, ANT1, and ANT2 end members, and a, b, c, and d represent the proportions (%) of each end-member contribution.

Natural and Anthropogenic Export in the River Watershed

Using the mixing equation (Eq. 11), it is possible to determine natural and anthropogenic exports from the river watershed. The export rates ω (in t/d) because of rock weathering or anthropogenic additions can be calculated for each end member (SIL, CAR, ANT1, and ANT2) and for each hydrological period as follows:

$$\omega = \Sigma(\alpha[\mathrm{Ca}^{2+}] + \alpha[\mathrm{Mg}^{2+}] + \alpha[\mathrm{K}^{+}] + \alpha[\mathrm{Na}^{+}] + \alpha[\mathrm{HCO_3}^{-}] + \alpha[\mathrm{Si}]) \cdot W \tag{12}$$

where [X] is the concentration (g/L), W represents the discharge (m^3/s) measured at the sampling points during the considered hydrological period, and α represents the proportion of an end member for the sample (20). The rock weathering and anthropogenic rates are calculated and ω results behave differently in accordance with the hydrological period.

Contribution of the Lithology

Comparing $^{87}\mathrm{Sr}/^{86}\mathrm{Sr}$ and elemental concentration ratios (e.g., Ca/Na, Mg/Na, and $\mathrm{HCO_3}$/Na) for the rivers draining different monolithological watersheds (i.e., silicate, carbonate, etc.), the respective contribution of each lithology can be deciphered (14,63). The following methodology can be used to correct all ratios for atmospheric input:

$$\mathrm{X}_{\mathrm{atm}} = \mathrm{Cl}_{\mathrm{r}}(\mathrm{X}/\mathrm{Cl})_{\mathrm{rw}}$$

where $\mathrm{X}_{\mathrm{atm}}$ is the percentage of element coming from the atmosphere, Cl_{r} is the chloride concentration in the river, and $(\mathrm{X}/\mathrm{Cl})_{\mathrm{rw}}$ is the ratio measured in the local rainwater (64).

CONCLUSIONS

Chemical weathering of rocks and soils is one of the essential processes in the geochemical cycling of elements in rivers. Although climatic factors such as moisture and temperature play a major role in releasing those elements, atmospheric aerosols have been identified as a major source of dissolved species in river water. Rainwater is significantly impacted by sea salt aerosols, terrestrial aerosols, and anthropogenic sources. For the study of rocks–recharge interactions as well as river water chemistry, the river watersheds, tributaries, local rainfall, humidity, and temperature must be considered. Furthermore, attention must be given to the different contributions of the different end-member inputs, either natural or anthropogenic, to the total dissolved load, as any calculation of the export rate of chemical species requires the quantification of each end member.

BIBLIOGRAPHY

1. Garrels, R.M. and McEnzie, F.T. (1971). *Evolution of Sedimentary Rocks*. W.W. Norton, New York.
2. Drever, J.I. (1988). *The Geochemistry of Natural Waters*, 2nd Edn. Prentice-Hall, Englewood Cliffs, NJ.
3. Sarin, M.M. et al. (1989). Major ion chemistry of the Ganga-Brahmaputra River system, weathering processes and fluxes to the Bay of Bengal. *Geochim. Cosmochim. Acta* **53**: 997–1009.
4. Berner-Kay, E. and Berner, R.A. (1987). *The Global Water Cycle. Geochemistry and Environment*. Prentice-Hall, Englewood Cliffs, NJ.
5. Millot, R., Gaillardet, J., Dupre, B., and Allegre, C.J. (2002). The global control of silicate weathering rates and the coupling with physical erosion: new insights from rivers of the Canadian Shield. *Earth Planet. Sci. Lett.* **196**: 83–98.
6. Krishnaswami, S. et al. (1992). Strontium isotopes and rubidium in the Ganga-Brahmaputra River system: weathering in

the Himalaya, fluxes to the Bay of Bengal and contributions to the evolution of oceanic ^{87}Sr/^{86}Sr. *Earth Planet. Sci. Lett.* **109**: 243–253.

7. Derry, L.A. and France-Lanord, C. (1996). Neogene Himalayan weathering history and river ^{87}Sr/^{86}Sr: impact on the marine Sr record. *Earth Planet. Sci. Lett.* **142**: 59–72.
8. Galy, A., France-Lanord, C., and Derry, L.A. (1999). Th strontium isotopic budget of Himalayan rivers in Nepal and Bangladesh. *Geochim. Cosmochim. Acta* **63**: 1905–1925.
9. Dalai, T.K., Singh, S.K., Trivedi, J.R., and Krishnaswami, S. (2002). Dissolved rhenium in the Yamuna River system and the Ganga in the Himalaya: role of black shale weathering on the budgets of Re, Os, and U in rivers and CO_2 in the atmosphere. *Geochim. Cosmochim. Acta* **66**: 29–43.
10. Quade, J., English, N., and DeCelles, P.G. (2003). Silicate versus carbonate weathering in the Himalaya: a comparison of the Arun and Seti River watersheds. *Chem. Geol.* **202**: 275–296.
11. Gibbs, R.J. (1967). The geochemistry of the Amazon River system: Part I. The factors that control the salinity and the composition and concentration of the suspended solids. *Geol. Soc. Am. Bull.* **78**: 1203–1232.
12. Millot, R., Gaillardet, J., Dupre, B., and Allegre, C.J. (2003). Northern latitude chemical weathering rates: clues from the Mackenzie River basin, Canadian. *Geochim. Cosmochim. Acta* **67**: 1305–1329.
13. Probst, J-L. et al. (1992). Dissolved major elements exported by the Congo and the Ubangi Rivers during the period 1987–1989. *J. Hydrol.* **135**: 237–257.
14. Negrel, P., Allegre, C.J., Dupre, B., and Lewin, E. (1993). Erosion sources determined by inversion of major and trace element ratios and Sr isotopic ratios in river water. The Congo basin case. *Earth Planet. Sci. Lett.* **120**: 59–76.
15. Gaillardet, J., Dupre, B., and Allegre, C.J. (1995). A global geochemical mass budget applied to the Congo Basin rivers: erosion rate and continental crust composition. *Geochim. Cosmochim. Acta* **59**: 3469–3485.
16. Ming-hui, H., Stallard, R.F., and Edmond, J.M. (1982). Major ion chemistry of some large Chinese rivers. *Nature* **298**: 550–553.
17. Han, G. and Liu, C-Q. (2004). Water geochemistry controlled by carbonate dissolution: a study of the river waters draining karst-dominated terrain, Guizhou Province, China. *Chem. Geol.* **204**: 1–21.
18. Keasler, K.M. and Loveland, W.D. (1982). Rare earth elemental concentrations in some Pacific Northwest rivers. *Earth Planet. Sci. Res.* **61**: 68–72.
19. Goldstein, S.J. and Jacobsen, S.B. (1988). Rare earth elements in river waters. *Earth Planet. Sci. Res.* **89**: 35–47.
20. Grosbois, C., Negrel, Ph., Fouillac, C., and Grimaud, D. (2000). Dissolved load of the Loire River: chemical and isotopic characterization. *Chem. Geol.* **170**: 179–201.
21. Picouet, C., Dupre, B., Orange, D., and Valladon, M. (2002). Major and trace element geochemistry in the upper Niger River (Mali): physical and chemical weathering rates and CO_2 consumption. *Chem. Geol.* **185**: 93–124.
22. Junge, C.E. (1963). *Air Chemistry and Radioactivity*. Academic Press, New York.
23. Gibbs, R.J. (1970). Mechanisms controlling world water chemistry. *Science* **170**: 1088–1090.
24. Likens, E.G., Boremann, F.H., Pierce, R.S., Eaton, J.S., and Johnson, N.M. (1977). *Biogeochemistry of Forested Ecosystem*. Springer-Verlag, Berlin.
25. Negrel, Ph. and Roy, S. (1998). Rain chemistry in the Massif Central (France): a strontium isotopic and major elements study. *Appl. Geochem.* **13**: 941–952.
26. Jaffrezo, J.L. (1987). Étude du lessivage des aerosols atmospheriques per les precipitations en milieu urbain. Ph.D. thesis. Université de Paris VII.
27. Graustein, W.C. and Armstrong, R.L. (1983). The use of ^{87}Sr/^{86}Sr ratios to measure atmospheric transport into forested watershed. *Science* **219**: 298–392.
28. Gosz, J.R. and Moore, D.I. (1989). Strontium isotope studies of atmospheric inputs to forested watersheds in New Mexico. *Biogeochemistry* **8**: 115–134.
29. Dupre, B., Negrel, Ph., Seimbille, F., and Allegre, C.J. (1994). ^{87}Sr/^{86}Sr ratios variation during a rain event. *Atmos. Environ.* **28**: 617–620.
30. Herut, B., Starinsky, A., and Kartz, A. (1993). Strontium in rainwater from Israel, sources, isotopes and chemistry. *Earth Planet. Sci. Lett.* **120**: 77–84.
31. English, N.B., Quade, J., DeCelles, P.G., and Garzione, C.N. (2000). Geologic control of Sr and major element chemistry in Himalayan rivers, Nepal. *Geochim. Cosmochim. Acta* **64**: 2549–2566.
32. Central Bureau of Statistics. (1998). *Statistical Pocket Book, Nepal*. His Majesty's Government National Planning Commission Secretariat, Kathmandu, Nepal.
33. Freyssinet, P. and Farah, A.S. (2000). Geochemical mass balance and weathering rates of ultramafic schists in Amazonia. *Chem. Geol.* **170**: 133–151.
34. Galy, A. and France-Lanord, C. (1999). Weathering processes in the Ganges-Brahmaputra basin and the riverine alkalinity budget. *Chem. Geol.* **159**: 31–60.
35. Viers, J. et al. (2000). Major and trace element abundances, and strontium isotopes in the Nyong basin rivers (Cameroon): constraints on chemical weathering processes and elements transport mechanisms in humid tropical environments. *Chem. Geol.* **169**: 211–241.
36. Elbaz-Poulichet, F. et al. (1999). Trace element geochemistry in the upper Amazon drainage basin (Bolivia). *Chem. Geol.* **157**: 319–334.
37. Vital, H. and Stattegger, K. (2000). Major and trace elements of stream sediments from the lowermost Amazon River. *Chem. Geol.* **168**: 151–168.
38. Losno, R., Bergametti, G., Carlier, P., and Mouvier, G. (1991). Major ions in marine rain water with attention to sources of alkaline and acidic species. *Atmos. Environ.* **25A**: 763–770.
39. Galloway, J.N., Likens, G.E., Keene, W.C., and Miller, J.N. (1982). The composition of precipitation in remote areas of the world. *J. Geophys. Res.* **87**: 8771–8786.
40. Basak, B. and Alagha, O. (2004). The chemical composition of rain water over Buyukcekmece Lake, Istanbul. *Atmos. Res.* **71**: 275–288.
41. Glavas, S. and Moschonas, N. (2002). Origin of observed acidic–alkaline rains in a wet-only precipitation study in a Mediterranean coastal site, Patras, Greece. *Atmos. Environ.* **36**: 3089–3099.
42. Aliva, A. and Alarcon, M. (1999). Relationship between precipitation chemistry and meteorological situations at a rural site in NE Spain. *Atmos. Environ.* **33**: 1663–1667.
43. Le Bolloch, O. and Guerzoni, S. (1995). Acid and alkaline deposition in precipitation of the western coast of Sardinia, Central Mediterranean (40 °N, 8 °E). *Water Air Soil Pollut.* **85**: 2155–2160.

44. Rao, P.S.P. et al. (1995). Rain water and throughfall chemistry in the Silent Valley forest in south India. *Atmos. Environ.* **29**: 2025–2029.
45. Plaisance, H., Coddeville, P., Roussel, I., and Guillermo, R. (1996). A qualitative determination of the source locations of precipitation constituents in Morvan, France. *Environ. Technol.* **17**: 977–986.
46. Hofmann, E.J., Hofmann, G.L., and Duce, R.A. (1974). Chemical fractionation of alkaline earth metals in atmospheric particulate matter over the North Atlantic. *J. Geophys. Res.* **8**: 675–688.
47. Bonsang, B., Nguyen, B.C., Gaudry, A., and Lambert, G. (1980). Sulfate enrichment in marine aerosols owing to biogenic gaseous sulfur compounds. *J. Geophys. Res.* **85**: 7410–7416.
48. Suzuki, T. and Tsunogai, S. (1988). Origin of calcium in aerosols over the western North Pacific. *J. Atmos. Chem.* **6**: 363–374.
49. Ichikuni, M. (1978). Calcite as a source of access Ca in rain water. *J. Geophys. Res.* **83**: 67249–67688.
50. Keene, W.C.A.A., Pszenny, P., Galloway, J., and Hawley, M.E. (1986). Sea-salt corrections and interpretations of constituent ratios in marine precipitation. *J. Geophys. Res.* **91**: 6647–6658.
51. Meybeck, M. (1983). Atmospheric inputs and river transport of dissolved substances. *IAHS Publ.* **141**: 173–192.
52. Drever, J.I. and Zobrist, J. (1992). Chemical weathering of silicate rocks as a function of elevation in the Southern Swiss Alps. *Geochim. Cosmochim. Acta* **56**: 3209–3216.
53. Velbel, M.A. (1985). Geochemical mass balances and weathering rates in forested watersheds of the Sothern Blue Ridge. *J. Sci.* **285**: 901–930.
54. Mast, M.A. and Drever, J.I. (1990). Chemical weathering in the Loch Vale watershed, Rocky Mountain National Park, Colorado. *Water Resour. Res.* **26**: 2971–2978.
55. Meybeck, M. (1979). Concentrations des eaux fluviales en elements majeurs et apports en solution aux oceans. *Rev. Geol. Dyn. Geogr. Phys.* **21**: 215–246.
56. Gaillardet, J., Dupre, B., Allegre, C.J., and Negrel, P. (1997). Chemical and physical denudation in the Amazon River basin. *Chem. Geol.* **142**: 141–173.
57. Sherwood, W.C. (1989). Chloride loading in the South Fork of the Shenandoah River, Virginia, U.S.A. *Environ. Geol. Water Sci.* **14**(2): 99–106.
58. Meybeck, M. (1986). Composition chimique des ruisseaux non pollues en France. *Sci. Geol. Bull.* **39**(1): 3–77.
59. Dojlido, J.R. and Best, G.A. (1996). Chemistry of water and pollution. *J. Hydrol.* **180**: 1–4.
60. Ramos, C. (1996). Effect of agricultural practices on the nitrogen losses to the environment. *Fert. Environ.* 355–361.
61. Mackenzie, R.T. and Garrels, R.M. (1966). Chemical mass balance between rivers and the oceans. *Am. J. Sci.* **264**: 507–525.
62. Probst, J-L. (1986). Dissolved and suspended matter transported by the Girou River (France): mechanical and chemical erosion rates in calcareous molasses basin. *Hydrol. Sci. J.* **31**: 1–3.
63. Gaillardet, J., Dupre, B., Louvat, P., and Allegre, C.J. (1999). Global silicate weathering and CO_2 consumption rates deduced from the chemistry of large rivers. *Chem. Geol.* **159**: 3–30.
64. Freydier, R. (1997). Chimie des precipitations en Afrique intertropicale: cations et anions majeurs, elements traces et acides organiques. These de doctorat. Université Paul Sabatier, Talouse, France.

READING LIST

Berner-Kay, E. and Berner, R.A. (1996). *Global Environment, Water, Air and Geochemical Cycles*. Prentice-Hall, Englewood Cliffs, NJ.

PHYSICS AND CHEMISTRY OF WATER

ACID RAIN AND SOCIETY

Paulette Middleton

(from *The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts*, Wiley 2003)

INTRODUCTION

Acid rain is one of many manifestations of how actions of society can have adverse effects on human health and welfare. Now more than ever before, the breadth of socioeconomic as well as environmental impacts associated with air pollutants, connections among pollutant contributions to these many impacts, and the implications of these connections are being recognized for policy making and development of management strategies.

It can no longer be argued that it is very costly to mediate acid rain and related air quality concerns. Assessments are beginning to suggest that multiple benefits associated with addressing acid rain in combination with other issues outweigh the costs of control of key responsible pollutants. In addition, when innovative strategies, which include market trading and incentives for conservation and use of clean fuels, are initiated, the costs of pollutant management become even lower. As factors that are not easily quantified monetarily are considered more directly in assessments, the benefits become even greater.

Recent analyses show that the implementation of the acid-rain-related part of the 1990 Clean Air Act Amendments has resulted in reductions in acidity in the north-eastern United States. Improvements in acid-related impacts have also been suggested. However, projections of future conditions over the next 20 to 50 years suggest that, unless more dramatic steps are taken, the overall burden of harmful pollutants could continue to rise in general in different parts of the country. Dramatic reductions in sulfur oxides, the current main contributors to acidity, alone are probably not enough. Planned reductions in nitrogen oxides may or may not be adequate. Similarly, continued close monitoring, if not increased management, of volatile organic compounds and fine particulates, and reassessment of their importance to acid rain and related concerns will be important over the next few years.

Of course, all of these issues are part of the bigger international picture of energy, environment, and economy. While the United States may be more aggressively and wisely addressing acid rain and related issues here at home, the projections for future fossil fuel use worldwide must be considered for the sake of regional air quality as well as the global climate condition. Capital investment in cleaner technologies such as renewable energy and the promotion of conservation strategies worldwide could bring long-term environmental and economic benefits that far surpass the initial costs. The alternative, continued growth in fossil fuel usage in developing countries, could exacerbate air quality problems that already exist in many of these areas and, in the long term, could cause the same acid rain damages experienced in many parts of North America and western Europe. In addition, it would contribute to adverse long-term carbon-dioxide-induced climate change.

As a contribution to our understanding of the atmospheric pollution problem and our role in the solution, this chapter summarizes:

- Acid rain and its relationship with other major issues
- U.S. response to the acid rain issue
- Current assessment of progress on reducing the effects of acid rain
- Projections and speculation on the future of acid rain

ACID RAIN: THE PHENOMENON

The relationships between chemical emissions into the atmosphere and the effects of the chemicals on various ecosystems, human health, and materials are highly complex. Many harmful chemicals (i.e., air pollutants) chemically interact to form other pollutants, which are perhaps even more harmful than the originally emitted chemicals. The most prominent examples of these dangerous chemical products are acid rain, ozone, and aerosols. In addition, many air pollutants are thought to interact in a synergistic fashion to cause even more harm as a group rather than individually. An example of the possible synergism is the hypothesized impact of acid rain and ozone on forest ecosystems. Understanding the causes and effects of acid rain and related air quality issues has become an important mission of atmospheric scientists around the globe.

DEFINITION OF ACID RAIN

Acid rain is the general term used to describe the removal, by rainfall, of acidic pollutants from the atmosphere. Acids also can be removed by other forms of precipitation, such as snow or fog. Acid pollutants may also fall as dry particles or gases that form acids when later combined with moisture. The term *acid deposition* is used to include all the possible forms of acid pollutant removal from the atmosphere, but *acid rain* remains the popular term.

The majority of the deposited acids are nitric acid and sulfuric acid. In some of the more rural regions of the world, organic acids also are important. In very remote areas where the level of acid is low, the “natural background” acid is carbonic acid, which is associated with carbon dioxide in the air. The overall acidity of precipitation also depends on the basic (or alkaline) constituents of the precipitation. Major bases include ammonia and geologic materials, such as dust and fly ash.

Acidity is measured in terms of a pH scale, which is a measure of the log of the hydrogen ion concentration in the precipitation. The scale runs from 0 to 14 with 0 being very acidic and 14 being very alkaline. A midscale value of 7 is considered neutral. A change in 1 pH unit indicates 10-fold increase or decrease in acidity. Unpolluted rain water is considered to have a pH of about 5.6. This acidity is assumed to contain only carbonic acid. In the highly polluted eastern states of the United States., the average acidity of water has a pH between 4 and 5. Even in some remote areas of the world, pH values of 5.2 have been found. These acidity levels suggest that there is a long-range transport of nitrogen and sulfur chemicals.

SOURCES OF ACIDITY

Atmospheric acids mainly are produced in the air as a result of complex chemical reactions of the acid precursor gases. Direct emissions of acids such as sulfuric acid, hydrogen chloride, and hydrogen fluoride have been estimated but are not thought to play a significant role in the acidic deposition processes (1). Sources of the harmful chemical precursors affecting the acidity of deposited materials can be natural or human caused. The three pollutants of most concern in the acidic deposition process are sulfur dioxide (SO_2), nitrogen oxides (NO_x) and volatile organic compounds (VOCs). Fossil-fuel-based power plants and motor vehicles are the major sources of all of these acid precursor pollutants in industrialized areas of the world (2,3).

Sulfur gases are primarily emitted from point sources, involving the combustion of coal in particular. Natural sources of sulfur gases include sea spray, volcanoes, and biologic activity. These sources, however, are at least a factor of 10 *less* than the human-caused emission for major industrial areas such as the United States. Nitrogen gases also result primarily from human activities involving fossil-fuel-derived energy use related to transportation and utilities. Major natural sources include soils and lightning and are thought to make up a significant portion of the overall emission totals, more than has been estimated for the sulfur gases. Estimates of these levels, however, are uncertain (1). The VOCs that produce the organic acids and influence the chemistry producing sulfuric and nitric acids also come mainly from automobile use. However, for the VOCs, natural production from vegetation can be quite significant. In highly vegetated, low industrialized regions natural sources become the dominant producers of VOCs.

Estimates of alkaline particulate and ammonia emissions indicate that there is a high potential for acid neutralization in some parts of the United States. The estimates, however, are subject to a high degree of uncertainty (1). On a global scale, emissions of these important acid neutralizers are among the least well-known chemical emissions (2).

EFFECTS OF ACID RAIN

The effects of acid deposition are the subject of continuing controversy. The northeastern United States has experienced the worst reported impacts in the United States (1). Severe damages attributed to acid rain also have been documented for parts of western Europe (4).

The most sensitive systems to acid deposition are poorly buffered lakes and streams. Buffering capacity refers to the availability of alkaline minerals from soil or rocks to neutralize the acids. When minerals are dissolved in a lake, buffering is able to diminish acid effects. However, this buffering ability or alkalinity can be used up with additions of acidic pollutants. Low alkalinity lakes have the greatest potential for damage, since their neutralizing minerals can be quickly depleted.

Vegetation is exposed to wet acidic deposition through rain, snow, and by direct contact with low, acid-laden clouds. There is currently no widespread forest or crop damage in the United States related to these possible pathways. However, cloud acidity, together with a complex combination of other factors (e.g., ozone, soil acidification, climate) contribute to reduced cold tolerance in high-elevation spruce in the eastern United States and in Europe. This can contribute to damage to trees above cloud level during winters with particularly low temperatures.

Adverse effects on forests in other regions of the world are associated with ozone, as is the case with high-elevation pines in California, or they are closely related to localized soil nutrient deficiencies, as is the case with sugar maples in eastern Canada. Acidic deposition may increase leaching rates of important base cations, principally magnesium and calcium, in forest solids and may be a contributing factor in sugar maple decline in some areas.

Generally, controlled experiments on trees and crops have indicated that ozone, at concentrations near ambient levels, adversely affects forests and crops primarily by growth reduction. Other controlled experiments have demonstrated that normal levels of atmospheric sulfur and nitrogen deposition cause no negative direct effects. Some areas actually may benefit through nutrient enrichment by nitrogen and sulfur deposition.

Computer models project that continued acidic deposition could result in long-term deficiencies of nutrients in some soils. However, currently, there is no evidence to indicate that forest health in general is currently affected by nutrient deficiency or will be affected in the next half century.

Air pollution and acidic deposition contribute to the corrosion of metals and deterioration of stone in buildings, statues, and other cultural resources. Although air pollution is an important concern for cultural objects, the magnitude of its effect on construction materials has been difficult to assess. Many construction materials have protective coatings such as paints; therefore, maintenance and service of protective coatings have an important role in determining the ultimate impact of air pollutants. Paints may also be affected by ambient levels of air pollution.

Another related side effect of acid rain is visibility degradation. Fine particles in the atmosphere containing sulfate, nitrate, and other chemical constituents, which when deposited are associated with acid deposition, cause visibility degradation while in the air. These fine particles have been the major factor in the reduction of visibility in rural and urban areas in the eastern United States since the beginning of the century. In the U.S. West, visibility degradation is being reported in major urban areas and in national parks and wilderness areas.

Direct adverse effects of these pollutants on humans occur largely through the respiratory system. Sensitive populations with existing respiratory or cardiovascular problems, such as those with asthma, are especially susceptible. These effects have been mainly associated with the acid precursor gases and ozone. Studies of the effects of acidic aerosols, composed primarily of nitric acid, ammonium bisulfate, and sulfuric acid, are still relatively new. It has been found that on rare occasions acid levels approach 10 times the long-term mean levels for sulfuric acid. Although substantial uncertainty exists, the body of data raises concern that acidic aerosols alone, or in concert with other pollutants, may be contributing to health effects in exposed populations at current concentration levels.

Human health also can be affected indirectly by pollutants related to acid deposition. People who eat large amounts of fish from acidic lakes or streams may experience exposure to methylmercury in some regions of the country. Drinking water from acidic sources may contain significantly elevated levels of lead. It is unlikely, however, that exposure to humans by this pathway is important, except in isolated cases.

SOCIAL RESPONSE TO ACID RAIN

Historically, toxic effects have been observed in populations acutely exposed to high concentrations of air pollutants. As early as the Middle Ages in London prohibitions on coal burning were instituted in response to perceived health effects of mixtures of dust, soot, and fog. The industrial revolution brought the air pollution issue to the United States, where air quality management continued to be considered local in nature into the twentieth century.

The severity of air pollution impacts became very obvious during the London "killer fog" of 1952, when a mixture of particulates, sulfur dioxide, and acidic fog was associated with severe respiratory effects and approximately 4000 deaths. Emergence of air pollution as a public health issue in the 1950s, as a result of this and other deadly episodes, led to the development of federally funded research programs, culminating in the Clean Air Act and in the establishment of the Environmental Protection Agency (EPA) in 1970. These were major stimuli for the establishment of the U.S. National Ambient Air Quality Standards (NAAQS) that today restrict the atmospheric concentration of pollutants such as sulfur dioxide, nitrogen oxides, and ozone.

Other countries around the world have been developing institutional responses to the threat to human health of air pollution. Air pollution effects on the environment, however, had been slower to be recognized as a serious issue. Acid rain and its ecological effects were first documented in England at the end of the nineteenth century and became regional issues in northwestern Europe and the northeastern United States and eastern Canada only recently—in the late 1960s. During this period and into the 1970s, the mounting anecdotal evidence of the effects of acid rain on aquatic and terrestrial ecosystems launched acid rain as perhaps the first pollution threat to the environment to receive international attention.

The origins of the pressures to regulate acid rain in the United States were primarily twofold. First, Canada protested, lobbied, and publicized its contention that major environmental damage was occurring in its eastern provinces because of acid deposition, and that the major sources of acid precursors were in the United States Second, elected officials and citizens in the eastern and New England states echoed the same concerns, elevating the acid rain controversy to the level of a growing interregional conflict between receptor states and polluting states (5).

The U.S. responses to these concerns took the form of federal research and eventually control programs. The first step was the Acid Precipitation Act of 1980, which created the National Acid Precipitation Assessment Program (NAPAP). During its first 10 years, the research and periodic assessments conducted by NAPAP improved the understanding of the scientific processes and effects of acid deposition. The monitoring and research conducted in the 1980s and the subsequent integrated assessment completed in 1990 provided the scientific knowledge base for Title IV, the Acid Deposition Control Program, of the 1990 Clean Air Act Amendments.

Title IV is designed to reduce the adverse effects of acid deposition through the reductions in annual emissions of sulfur dioxide (SO_2) and nitrogen oxides (NO_x), the precursors to acid rain. Recognizing that the principal sources of acid rain precursors in the atmosphere are emissions from the combustion of fossil fuels, control measures were initiated to reduce emissions from electric utilities. However, rather than the traditional command-and-control approach to regulation, alternative methods of compliance were allowed. These methods included technological adaptation (e.g., scrubbers, higher-efficiency boilers), fuel-switching, and an innovative SO_2 emissions allowance trading program. This represented the first national effort to use market-based incentives to achieve environmental goals.

Due to the innovative nature of using market-based incentives for environmental regulation, Congress set up a mechanism for checking how well trading was working. As part of this activity, Congress asked NAPAP to assess the costs and economic impacts of the acid deposition control program as well as the effectiveness and benefits associated with the various human health and welfare effects. The effects included visibility, materials, and cultural resources damages and ecosystem effects. NAPAP was also asked to consider the deposition levels needed to protect sensitive ecosystems. The results of the assessment of Title IV are to be reported to Congress quadrennially, beginning with the 1996 Report to Congress (6).

CURRENT CONDITIONS

As of the completion of the first report to Congress, several observations have been made regarding the success of Title IV. It appears that the market-based approach has lowered compliance costs. Costs are lower than expected, probably due to a number of factors such as railroad deregulation, technological innovation, and lower operating costs for scrubbers. In addition, all affected utilities have fulfilled the compliance requirements of Title IV. In the first annual reconciliation of allowances and emissions, SO_2 allowances matched or exceeded SO_2 emissions. NO_x reductions have not been as dramatic. This is expected since mandates on NO_x reductions are not in place yet. However, NO_x emissions from all sources in 1995 were 1.5 million tons below 1980 levels. Utilities were responsible for 53% of that reduction.

Statistically significant reductions in acidity and sulfate in precipitation were reported at monitoring sites in the Midwest, mid-Atlantic, and northeastern United States. There is no real evidence of statistically significant decreases in nitrate concentration. Changes in aquatic ecosystems have not yet been detected. However, over the last 15 years, lakes and streams throughout many areas of the United States have experienced decreases in sulfate concentration in response to decreased emissions. While there is some evidence of recovery from acidification in New England, Adirondack lakes continue to acidify, suggesting that additional reductions may be needed in these areas (6).

Sulfur and nitrogen deposition has caused adverse impacts on certain sensitive forest ecosystems in the United States, with high-elevation spruce in the eastern United States being most sensitive. Other sensitive forests are apparently not experiencing the same effects in mortality and growth, at least for now, but some of the same processes appear to be slowly occurring.

The leaching of soil nutrients by continued acidic deposition is a gradual process that will eventually impact forest nutrition and growth in many areas. The recent reductions in sulfur should result in some small immediate improvements in sensitive forests, but large improvements will be slow to occur.

Reduced emissions of sulfur oxides are expected to reduce sulfate concentration and its contribution to haze. It is difficult to assess the extent to which recent reductions have contributed to changes in visibility over the past few years since meteorological and other factors determine the overall changes in visibility. Information is needed over the long term.

The recent reductions in SO_x and NO_x emissions are expected to reduce fine particulates and, as a result, lead to improved human health. It is suggested that reduced emissions will lead to a reduction in premature mortality from cardiovascular and respiratory causes and to a dramatic reduction in the number of asthma symptom days.

One difficulty in determining effects at this time is that many impacts have a response times that are longer than the few years since the passage of Title IV. Visibility and acute health effects can be detected on the order of hours to days. Episodic aquatic effects and soil and plant processes in the forest ecosystem respond on the order of days and weeks to months. Chronic human health, chronic aquatic effects, and forest health, on the other hand, indicate response times on the order of years to decades. Effects on forest solid nutrient reserves and effects on materials begin to show up on the order of decades to centuries. These latter effects are more on the order of climate change impacts response times.

The difference in response times, of course, makes an evaluation of actions taken in the early 1990s difficult to quantify. Improvements in health and visibility can serve as indicators of positive change. However, as already noted, even changes in visibility cannot be directly attributed to sulfate reductions alone, since other factors such as meteorological variability play a role in determining visibility changes especially in the humid part of the eastern United States.

KEEPING A BROAD BASIS OF ASSESSMENT AND ACTION

To review, acid deposition is an end product of a complex series of interactions among atmospheric chemical species emitted by both natural and human sources. For policy assessment purposes, the most important groups of chemical species are compounds containing sulfur and nitrogen compounds that are emitted from factories, power plants, and automobiles based on fossil-fuel combustion. In addition, volatile organic carbon compounds and fine particles play a role in modulating chemical processes and acidity. Some key compounds remain unchanged in the atmosphere and some are neutralized, but others are oxidized into more acidic forms through a complicated series of chemical, meteorological, physical, and biological interactions.

Decisions about the control of acid deposition must deal with the environmental impacts of estimated future emission levels as well as present levels. Projections depend on many complex and interacting socioeconomic factors. The predictability of how rapidly and to what extent fossil fuels will be replaced by clearer and safer fuels (society may change transportation and other energy use habits) and the inter-relationships among countries of the world becoming driving influences for these changes is highly uncertain. Given the demonstrated value of examining multiple causes and effects together, it will continue to be important to keep the base of assessment broad in spite of the uncertainties. The elements of such assessments are illustrated in Fig. 1.

On issues related to acid rain, other policy discussions going on throughout North America also illustrate the growing awareness of the interconnections, as well as the need to capitalize on the relationship in developing strategies for the future. For example, EPA, through the Federal Advisory Committee Act (FACA), is leading the development of combined ozone, particulate matter (aerosols), and regional haze implementation program rules and guidance. Other activities at the regional level, such as the Western Governors' Association (WGA) Air Quality Initiative, the Ozone Transport and Analysis Group (OTAG), the Southern Appalachian Mountain

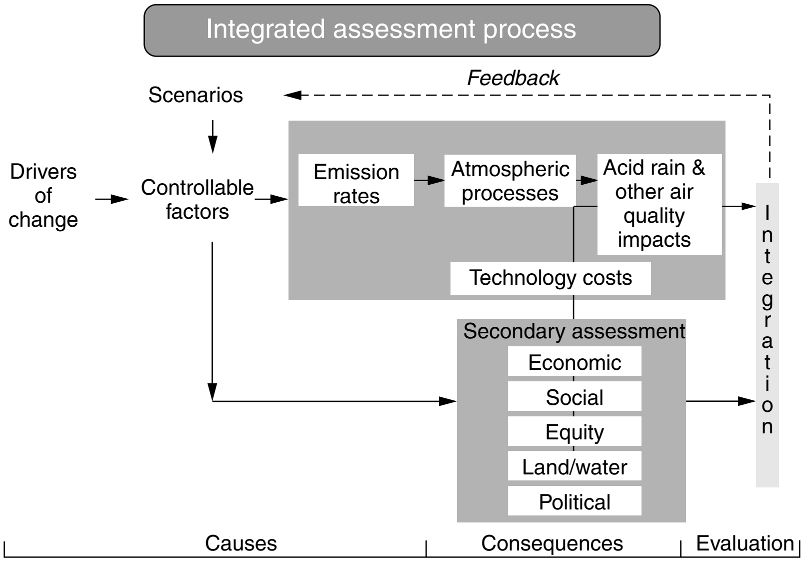

Figure 1. Important elements to consider in an assessment of acid rain in the context of other important environmental, energy, and economic concerns.

Initiative (SAMI), the Southern Oxidant Study (SOS), and the North American Research Strategy for Tropospheric Ozone (NARSTO), are addressing various science and policy issues associated with ozone, particulate matter, and regional haze.

On a continental scale, the Commission on Environmental Cooperation (CEC) is working on North American strategies for addressing transboundary concerns, which include ozone, particulate matter, regional haze, and acid rain along with other hazardous pollutants. Finally, on a much broader scale, the U.S. global climate change program is addressing regional climate assessment for areas throughout the United States. The regional air quality concerns addressed by FACA and the climate concerns to be considered in these discussions are closely linked through the development of aerosols that can influence climate on regional scales as well as produce other problems. On a larger policy implementation level, the two are linked through the development of energy strategies aimed at reducing greenhouse gas emissions and the emission of other more traditionally harmful air pollutants.

All of these programs and approaches share the same fundamental concerns. The role of natural or background processes and the role of chemical interactions in determining the levels of impacts in different regions continue to be fundamental overarching scientific questions. The implementation issues of emissions trading versus pollution prevention versus technological controls are also part of each aspect of the various debates. Assessments of trade-offs for any decisions made with respect to any of these issues must consider the less quantifiable, and sometimes more uncertain, impacts associated with heath, social impacts and values, equity, and related environmental concerns about water and soil quality as well as air. Given these strong interconnections, it is important to make the best use of research and policy-making resources across organizations addressing acid rain and related issues, where energy and the environment are key factors.

BIBLIOGRAPHY

1. U.S. National Acid Precipitation Assessment Program (NAPAP). (1991). *Acidic Deposition: State of Science and Technology, Summary Report of the U.S. National Acid Precipitation Assessment Program*. Washington, DC.
2. Graedel, T.E. et al. (1993). A compilation of inventories of emissions to the atmosphere. *Global Biogeochem. Cycles* **7**: 1–26.
3. Middleton, P. (1995). Sources of air pollutants. In: *Composition, Chemistry, and Climate of the Atmosphere*. Van Nostrand Reinhold, New York.
4. Graedel, T.E. and Crutzen, P.J. (1989). The changing atmosphere. *Sci. Am.* **261**(Sept.): 58.
5. Rhodes, S.L., and Middleton, P. (1983). The complex challenge of controlling acid rain. *Environment* **25**: 6–9, 31–38.
6. U.S. National Acid Precipitation Assessment Program (NAPAP). (1998). *NAPAP 1996 Report to Congress*, http://www.nnic.noaa.gov/CENR/NAPAP/NAPAP.96.htm.

ADSORPTION CAPACITY OF ACTIVATED CARBON FOR WATER PURIFICATION

Carl L. Yaws
Prasad K. Narasimhan
Lamar University
Beaumont, Texas

Adsorption on activated carbon is an effective method for water purification. In this article, results for the adsorption capacity of activated carbon for removing organic chemicals from water are presented for a wide variety of substances—alcohols, ketones, aldehydes, acetates, benzenes, chloroalkanes, chloro-olefins, and bromoalkanes. The results use the adsorption capacity as a function of the adsorption potential of the chemical in the aqueous liquid. The correlation constants are displayed in an easy to use tabular format that is especially applicable to rapid engineering usage.

Correlation results and experimental data are in favorable agreement.

INTRODUCTION

Physical and thermodynamic data for organic chemicals are especially helpful to engineers and scientists in industry. In particular, capacity data for adsorption of organic chemicals on activated carbon is becoming increasingly important in engineering and environmental studies because of more and more stringent governmental regulations involving substances in water.

In this article, the objective is to provide results for the adsorption capacity of activated carbon for removing organic chemicals from water. Such results are usable in engineering and environmental studies. As an example of such usage, capacity data from the correlation are useful in the engineering design of carbon adsorption systems for water purification.

CORRELATION FOR ADSORPTION CAPACITY

Adsorption capacity was correlated as a function of adsorption potential (Polyani theory) of a chemical in the aqueous phase using the following equation:

$$\log Q = A + B\ T/V \log\ (x_s/x) \qquad (1)$$

Q = adsorption capacity at equilibrium, g compound/100 g carbon
A and B = regression coefficients
T = temperature, K
V = molar volume of liquid, cm^3/mol
x_s = solubility in water, parts per million by weight, ppm(wt)
x = concentration in water, parts per million by weight, ppm(wt)

The coefficients (A and B) are given in Table 1 for a wide variety of organic chemicals in water—alcohols, ketones, aldehydes, acetates, benzenes, chloroalkanes, chloro-olefins, and bromoalkanes. The table also provides the representative formula and an example compound for the chemical families. The correlation should not be used for other chemical families.

The coefficients were determined from regression of available data. In preparing the correlation, a literature search was conducted to identify data source publications (1–10). The excellent compilations by Arbuckle (1); Sonthier, Crittenden, and Summer (2); and Speth (3,4) were used for adsorption capacity. The compilations by Howard and Meylan (5); Mackay, Shiu, and Ma (6); Verschueren (7); Yalkowsky (8); and Yaws (9) were used for water solubility. Liquid molar volume was ascertained from data compilations by Yaws (9,10). The publications were screened and copies of appropriate data were made. These data were then keyed into the computer to provide a database of adsorption capacity at different adsorption potentials in the aqueous phase. The database also served as a basis to check the accuracy of the correlation.

Graphs of adsorption capacity as a function of the adsorption potential are shown in Figs. 1–5, for the substances. The graphs discloses favorable agreement of correlation results and experimental data.

In general, experimental data available in the literature are very limited for the adsorption capacity of activated carbon for organic compounds in aqueous liquids at low concentrations (parts per million range). In view of the very limited experimental data, estimates in the tabulation should be considered rough approximations. Adsorption capacity is also complicated by differences in carbon, lots of carbon, source of water, and concentration ranges (4). If initial feasibility studies using the correlation are favorable, follow-up experimental determination of equilibrium adsorption capacity is recommended.

Examples

The correlations for water solubility are useful in engineering applications involving process wastewater. Examples are shown below.

Table 1. Adsorption Capacity of Activated Carbon—Aqueous System[a]

$\log_{10} Q = A + B\ T/V \log_{10}(x_s/x)$ ($Q - g$ of Compound/100 g of carbon, x – ppm(wt))

Chemical Family	A	B	Formula	Example Compound
Alcohols	1.7250	−0.1210	$C_nH_{2n+2}O$	1-Butanol
Ketones	1.7250	−0.1210	$C_nH_{2n}O$	Diethyl ketone
Aldehydes	1.7250	−0.1210	$C_nH_{2n}O$	Butyraldehyde
Acetates	1.7250	−0.1210	$C_nH_{2n}O_2$	Butyl acetate
Benzenes	1.8893	−0.1199	$C_{n+6}H_{2n+6}$	Ethyl benzene
Chloroalkanes	2.6021	−0.1931	$C_nH_{2n+2-x}Cl_x$	Chloroform
Chloro-olefins	2.2175	−0.1599	$C_nH_{2n-x}Cl_x$	Trichloroethylene
Bromoalkanes	3.0607	−0.1991	$C_nH_{2n+2-x}Br_x$	Bromoform

[a] F 400 Activated carbon, Calgon.
Q—adsorption capacity at equilibrium, g compound/100 g carbon.
x_s—solubility in water at 25 °C, parts per million by weight, ppm(wt).
x—concentration in water at 25 °C, parts per million by weight, ppm(wt).
T—temperature, K.
V—molar volume of liquid, cm^3/mole.
A and B—regression coefficients.

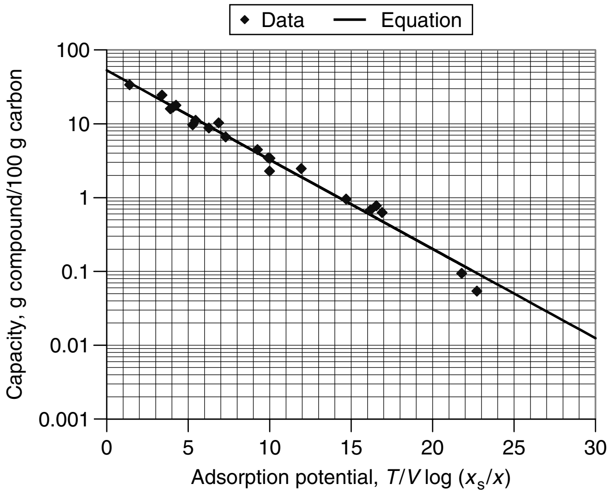

Figure 1. Adsorption capacity of activated carbon for organic oxygen chemicals (alcohols, ketones, aldehydes, and acetates) in water.

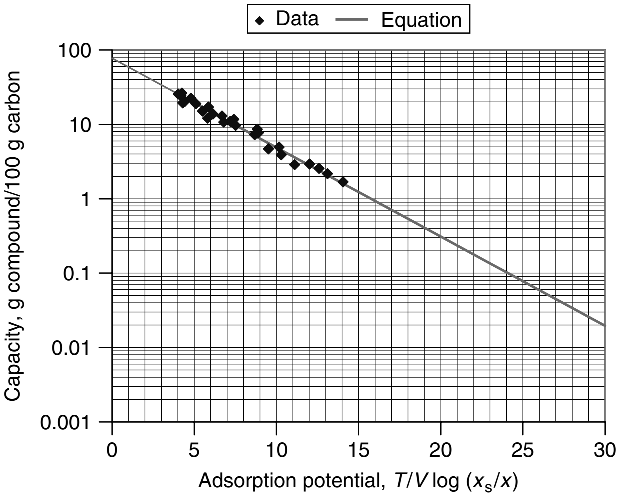

Figure 2. Adsorption capacity of activated carbon for benzenes in water.

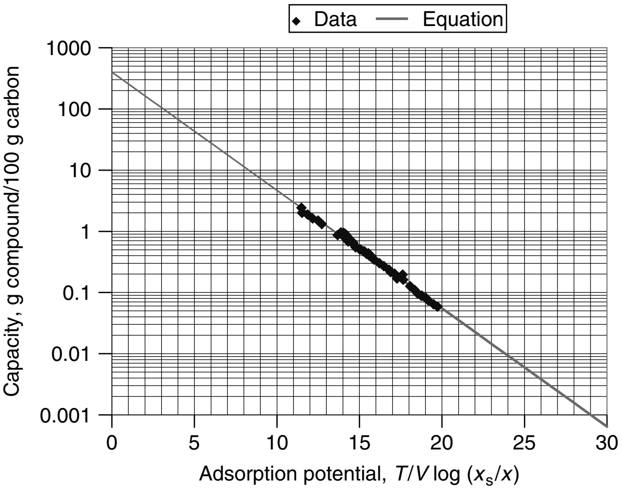

Figure 3. Adsorption capacity of activated carbon for chloroalkanes in water.

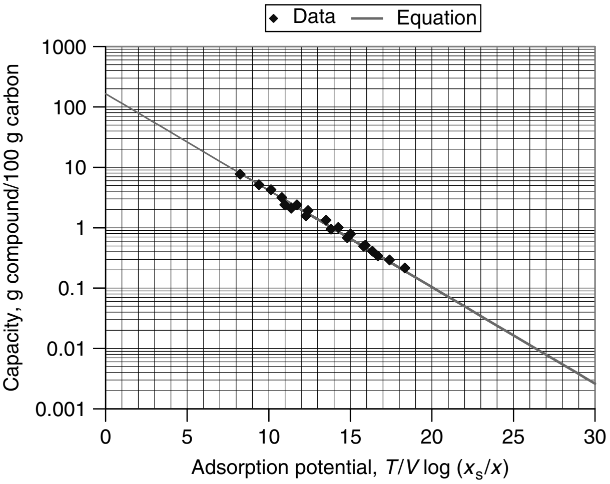

Figure 4. Adsorption capacity of activated carbon for chloro Olefins in water.

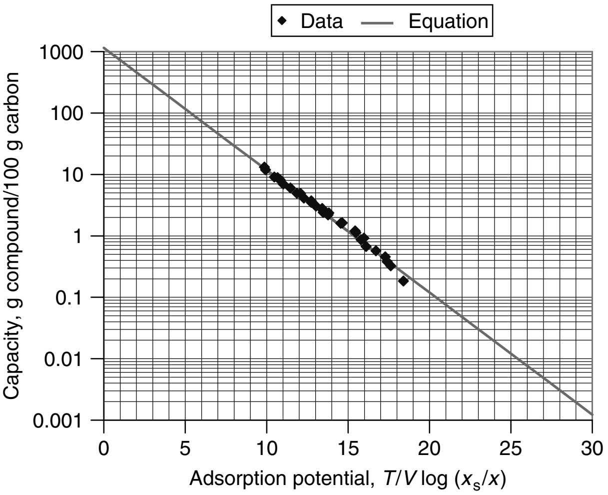

Figure 5. Adsorption capacity of activated carbon for bromoalkanes in water.

Example 1. A chemical spill of *n*-hexanol occurs in a body of water at ambient conditions. Additional data are

Concentration in water, ppm(wt)	96.31
Temperature, K	298.15
Molecular weight, g/mol	102.177
Liquid density, g/cm^3	0.816
Liquid molar volume, cm^3/mol	125.216
Solubility in water, ppm(wt)	5875

Estimate the adsorption capacity of activated carbon to purify the water.

The correlation for adsorption capacity may be used. Substitution of the coefficients for alcohols and above data in the correlation equation yields

$$\begin{aligned}\log_{10} Q &= 1.7250 - 0.1210 \\ &\quad \times [298.15/125.216 \log_{10}(5{,}875/96.31)] \\ &= 1.21063 \\ Q &= 16.24 \text{ g } n\text{-hexanol}/100 \text{ g carbon.}\end{aligned}$$

The calculated value and data compare favorably (16.24 vs. 17.79; deviation = 1.55/17.79 = 8.7%).

Example 2. A chemical spill of toluene occurs in a body of water at ambient conditions. Additional data are

Concentration in water, ppm(wt)	2.211
Temperature, K	298.15
Molecular weight, g/mol	92.141
Liquid density, g/cm^3	0.865
Liquid molar volume, cm^3/mol	106.521
Solubility in water, ppm(wt)	542.4

Estimate the adsorption capacity of activated carbon to purify the water.

The correlation for adsorption capacity may be used. Substitution of the coefficients for benzenes and the above data in the correlation equation yields

$$\begin{aligned}\log_{10} Q &= 1.8893 - 0.1199 \\ &\quad \times [298.15/106.521 \log_{10}(542.4/2.211)] \\ &= 1.0874 \\ Q &= 12.23 \text{ g toluene/100 g carbon}\end{aligned}$$

The calculated value and data compare favorably (16.23 vs. 12.96; deviation = 0.73/12.96 = 5.6%).

Acknowledgments
The Texas Hazardous Waste Research Center (Lamar University, Beaumont, Texas) provided partial support for this work.

BIBLIOGRAPHY

1. Arbuckle, W.B. (1981). *Environ. Sci. Technol.* **15**(7): 812.
2. Sontheimer, H., Crittenden, J.C., and Summers, R.S. (1988). *Activated Carbon for Water Treatment*, 2nd Edn. DVGW—Forschungsstelle, FRG.
3. Speth, T.F. (1986). Masters Thesis, Michigan Technological University, Houghton, MI.
4. Speth, T.F. and Miller, J.R. (1990). *J. AWWA* **72**(February).
5. Howard, P.H. and Meylan, W.M. (Eds.). (1997). *Handbook of Physical Properties of Organic Chemicals*. CRC Press, Boca Raton, FL.
6. Mackay, D., Shiu, W.Y., and Ma, K.C. (1992, 1992, 1993, 1995, 1997). *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*. Vols. 1–56. Lewis, New York.
7. Verschueren, K. (1996, 2001). *Handbook of Environmental Data on Organic Chemicals*, 3rd and 4th Edn. Van Nostrand Reinhold, New York.
8. Yalkowsky, S.H. (1990–2004). *Aquasol Database*. University of Arizona, Tucson, AZ.
9. Yaws, C.L. (1999). *Chemical Properties Handbook*. McGraw-Hill, New York.
10. Yaws, C.L. (2003). Yaws Handbook of Thermodynamic and Physical Properties of Chemical Compounds. electronic ed. www.knovel.com, Knovel Corporation, Norwich, NY.

ADSORPTION OF ORGANIC COMPOUNDS

Suparna Mukherji
IIT Bombay
Powai, Mumbai, India

INTRODUCTION

Organic compounds in the aqueous phase have a strong tendency to interact with surfaces, which results in their immobilization on the surfaces and consequent removal from the aqueous phase. This tendency is particularly high for nonpolar hydrophobic organic compounds (HOCs) characterized by low aqueous solubility. In water treatment applications, low levels of HOC contamination of both natural and anthropogenic origin are effectively removed by passage through a bed packed with activated carbon. In the subsurface environment, organic compounds in groundwater have a strong tendency to associate with soil. Thus, in contaminated zones soils and sediments are heavily loaded with organic contaminants, and the loading on these natural geosorbents per unit mass is eventually significantly higher than the aqueous phase concentration. The desorption of sorbed contaminants from soil acts as a long-term source, contaminating groundwater over decades. The association of HOCs with soil is typically termed sorption, which is inclusive of both absorption (partitioning) and adsorption (interface accumulation) (1).

Although sorption is primarily considered to be a reversible process, it significantly affects the fate and transport of organic compounds (2). It affects transformation processes such as photolysis and biodegradation. Thus, sorbed compounds are not readily available to microorganisms even though they may be readily biodegradable from aqueous solution. Aside from the fraction of the organics that are irreversibly bound to the surfaces, through reactions such as polymerization reactions, the reversibly bound fraction can desorb back into aqueous solution. However, the rate of desorption varies with the degree of penetration into the surface sites on the solids. A part of the reversibly bound fraction desorbs rapidly, while the remaining desorbs at a very slow rate over an extended period of time. This is responsible for the hysteresis and aging phenomena often reported in the literature. Hysteresis implies that isotherms relating the loading of a compound on a sorbent to aqueous phase concentration at equilibrium obtained in adsorption studies are not valid for desorption. Aging is the phenomenon of slow desorption observed for sorbents that have been exposed to pollutants over a long period of time. Thus, desorption experiments in laboratory contaminated soil may not be representative of field contaminated soil.

EQUILIBRIUM CONSIDERATIONS IN SORPTION OF HYDROPHOBIC ORGANIC COMPOUNDS ON GEOSORBENTS

Adsorption Isotherms

The maximum loading of an organic compound on a sorbent (q_e, mg/g) is a function of the solution phase concentration (C_e, mg/L) at equilibrium. This relationship is defined for a constant temperature and is known as an isotherm. Since the characteristics of the geosorbents are highly variable, each geosorbent yields a unique isotherm that needs to be obtained experimentally. Isotherms are obtained by setting up multiple batch reactors (sealed bottles) in which a fixed volume of the aqueous phase (V, L) containing varying concentrations of the organic compound (C_o, mg/L) is exposed to a fixed mass of sorbent (M, g). Parallel control reactors without sorbent are also included to determine the system losses. The reactors are tumbled end over end until equilibrium is achieved, and the loading on the sorbent (q_e) is determined by liquid phase mass balance as shown in Eq. 1 (1):

$$q_e = \frac{(C_o - C_e)}{(M/V)} \tag{1}$$

Equation 2 indicates how q_e can be determined after loss correction; it assumes system losses are linearly related to C_e and yields a constant slope S, which is estimated using the controls (3):

$$q_e = \frac{[C_o - C_e(1 + S/V)]}{(M/V)} \tag{2}$$

The experimentally obtained isotherms are fitted to models, such that the model and its parameters indicate the nature of the relationship between q_e and C_e. Three models most commonly used for fitting sorption data are the linear model, the Langmuir model, and the Freundlich model. The model expression, parameters, and the linear transformation typically used for parameter estimation are provided in Table 1 (1).

The constant K_D in the linear model can be determined by one-parameter linear regression. The constants Q_a^o and b in the Langmuir model and the constants K_F and n in the Freundlich model can be determined by two-parameter linear regression by using appropriate data transformations. However, a direct estimation of model parameters by nonlinear regression, which can readily be performed using statistical packages, is more desirable since the transformations performed distort the error structure, thus violating a basic assumption inherent in least squares linear regression (4). The Langmuir model assumes monolayer sorption, reversible sorption, and equal energy sorption sites. It is linear at low concentrations and levels off at a high concentration and shows the characteristic shape depicted in Fig. 1. The Freundlich model is empirically derived and typically provides the best fit to data since it allows for multilayer sorption and heterogeneous sites with variable energies (1). The characteristic shape of this isotherm for various n values is demonstrated in Fig. 1. Low n values, significantly lower than unity, are indicative of greater heterogeneity in site energy distribution and more favorable sorption for low C_e values. The K_F value reflects the sorption capacity.

Sorption in Various Geosorbent Domains

Sorption of nonpolar organic compounds on geosorbents can occur in various domains that are distributed within

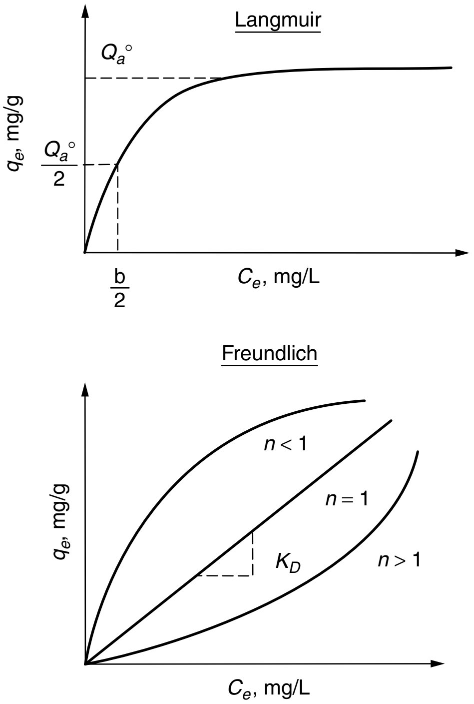

Figure 1. Characteristic shape of the Langmuir and Freundlich isotherm.

Table 1. Isotherm Models and Linear Transformations for Parameter Estimation

Model	Equation	Parameters	Linearization[a] ($y = a + mx$) and Parameter Estimation
Linear	$q_e = K_D C_e$ $K_D = f_{OC} K_{OC}$	K_D, distribution coefficient; f_{OC}, fraction organic carbon in sorbent; K_{OC}, organic carbon normalized coefficient	$a = 0, K_D = m$
Langmuir	$q_e = \dfrac{Q_a^o b C_e}{1 + bC_e}$	Q_a^o, monolayer sorption capacity; b, relates to sorption energetics	Reciprocal transform: $\dfrac{1}{q_e}$ versus $\dfrac{1}{C_e}$ $Q_a^o = 1/a;\ b = a/m$
Freundlich	$q_e = K_F C_e^n$	K_F, relates to sorption capacity; n, relates to sorption energetics	Log transform: $\log q_e$ versus $\log C_e$ $K_F = 10^a;\ n = m$

[a] Here a is the intercept and m is slope.

the geosorbent. Thus, the overall q_e is the sum total of q_e in the various domains. The domains have been conceptualized as the amorphous organic matter domain, the condensed organic matter domain, and the mineral domain. While organic matter is typically considered to be of natural origin, organic matter of anthropogenic origin can also exert a significant influence when present (5). Such organic matter includes nonaqueous phase liquids and combustion residues such as soot.

Sorption in Amorphous and Condensed Organic Matter Domains. It is well established that soil organic matter has a strong influence on the sorption of organic compounds and a linear correlation exists over a wide range of organic carbon content from about 0.1% to 20%. The distribution coefficient K_D for sorption of an organic compound on soil (assuming linear isotherm) can thus be normalized using the fraction organic carbon on a soil to yield the organic carbon normalized coefficient (K_{OC}). The K_{OC} value is strongly dependent on the organic compound of interest and correlates strongly with the octanol–water partition coefficient, K_{OW}, and the solubility S of the organic compounds. Linear correlations between log K_{OC} and log K_{OW} have been developed for specific groups of compounds as summarized in various reviews and handbooks (6,7). However, these values should be used as preliminary estimates.

It is well established that the linear model is often inadequate and the Freundlich model provides better fits to data, particularly when data is obtained over a wide C_e range. For phenanthrene sorption on various sorbents, Weber et al. (8) have demonstrated how the individually measured (single point) K_D and corresponding K_{OC} values vary as a function of C_e, when sorption is well characterized by the Freundlich isotherm. This has been attributed to the heterogeneity of sorption sites. The organic matter in the sorbent includes proteins, lignin, cellulose, and macromolecules that are altered over geologic time scales. This organic matter originates from organisms, such as microorganisms (algae and bacteria), zooplanktons and phytoplanktons, and vascular plants, depending on the depositional environment of the sorbents.

The linear model was originally proposed for sorption of organic compounds on various soils. Soil organic matter is proposed to be amorphous in nature and is often termed "soft" carbon, which is easily oxidized (e.g., humic and fulvic acids). Sorption of organics on soft carbon is like a partitioning phenomenon; that is, organic contaminants are distributed over the entire volume of soil organic matter. This is essentially a physisorption process characterized by low sorption energies. The K_{OC} values that are determined from correlations based on K_{OW} and aqueous solubility widely reported in the literature are only representative of sorption in the soft organic matter fraction.

As the organic matter gets aged by weathering processes over geologic time scales, indicated by a lowering in the O/C and H/C ratios, it gets more condensed, the degree of cross-linking increases, its aromaticity increases, and it becomes "hard" and insoluble (difficult to oxidize), for example, kerogen present in shales. Sorption on condensed organic matter is a true adsorption phenomenon, characterized by accumulation of organic compounds at the interface. The condensed organic matter has a significantly higher sorption capacity for organic compounds compared to the amorphous organic matter, and the sorption nonlinearity increases (characterized by decrease in Freundlich exponent n) with increase in the degree of condensation as diagenesis progresses. The sorption energies are also higher as in chemisorption processes. The distributed reactivity model proposed by Weber et al. (9) indicates that sorption in geosorbents is the sum total of sorption in the individual components; however, shale particles in a subsurface sand, rich in condensed organic carbon, were found to dominate the sorptive interactions. The failure of subsurface remediation technologies designed using linear isotherms have been attributed to increased nonlinearity.

Physisorption involving weak intermolecular forces is a reversible process. Thus, sorption on soft amorphous carbon typically present in soil is reversible. The adsorption isotherm and desorption isotherm are coincident and both are linear, as shown in Fig. 2. For sorption on hard condensed carbon such as kerogen, the adsorption and desorption isotherms are not coincident and significant hysteresis exists as demonstrated in Fig. 2. Moreover, both sorption and desorption are characterized by significant nonlinearity.

Sorption in Mineral Domain. The mineral surfaces exert greater influence on the sorption of polar organic compounds and the sorption is strongly dependent on the solution pH, pK_a values of the polar organic compound, and pH$_{\mathrm{PZC}}$ (point of zero charge of the sorbent). For nonpolar organic compounds, although the influence of organic matter content and nature dominates the sorptive interaction, mineral surfaces can also play a role, particularly when the organic matter content is less than

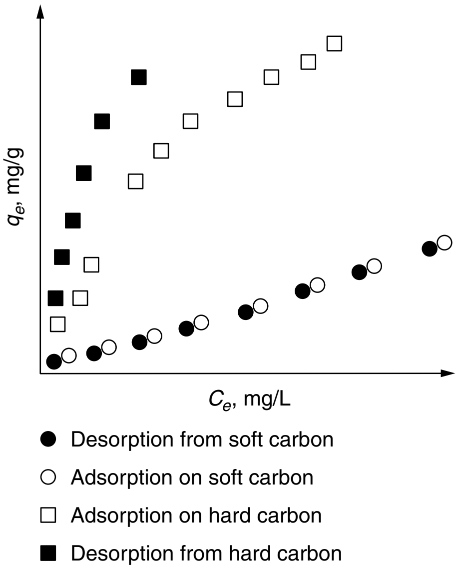

Figure 2. Adsorption and desorption isotherms on hard and soft carbon.

0.1–0.2%. Aquatic systems and subsurface environments are often characterized by low organic matter content. Moreover, the clay liners used in landfills can play a role in the retardation of organic compounds by virtue of sorptive interactions.

Sorption on minerals is dependent on the area of the mineral surface per unit mass of the sorbent. Area of the exposed mineral surface is higher when the organic carbon content is lower, since the natural organic matter competes with the organic compound for attaching on to the mineral surface. Thus, porous silica and the finer clay fraction exert a greater influence. Overall sorption also depends on the type of mineral surfaces that are present—clays (montmorillonite, kaolinite), silica, alumina, and iron oxide. Competition between sorption on natural organic matter and on mineral surfaces may depend on the proportion of swelling clays, the determining parameter being the ratio of clay surface area to the fraction of organic carbon. The interactions on mineral surfaces are typically due to van der Waals forces, that is, dipole–dipole, dipole–induced dipole, and hydrogen bonding interactions. Sometimes the interactions are stronger; chemisorption and clay colloid induced polymerization dominate sorption. Such specific binding of organics on the mineral surfaces may cause sorption to appear irreversible (7).

Factors Affecting Sorption Equilibrium

Factors other than geosorbent characteristics that affect sorption have been discussed elaborately by Schwarzenbach et al. (7). These include the following:

- *Temperature*: Adsorption is typically an exothermic process and K_D decreases with increasing temperature. Heat of adsorption associated with physisorption is typically a few hundred calories per degree per mole.
- *Nature of the Compound*: Within a group of organic compounds, K_D and K_{OC} increase as the aqueous solubility decreases and as the octanol–water partition coefficient K_{OW} increases. These linear free energy relationships have been widely utilized for obtaining preliminary estimates of the linear sorption coefficient.
- *Solution pH*: The effect of pH on the sorption of nonpolar organic compounds is not very significant. For ionizable organics the effect is significant.
- *Salinity*: As salinity increases, the aqueous solubility of organic compounds decreases; thus, K_D increases. However, typical levels of dissolved salts present in aquatic environments do not cause much impact.
- *Organic Colloids*: Colloids are macromolecules that remain in aqueous solution. Thus, in aqueous solution the organic compound is present as truly dissolved species and also as species associated with colloidal macromolecules. This causes an apparent lowering in the K_D values.

RATE CONSIDERATIONS IN SORPTION OF HYDROPHOBIC ORGANIC COMPOUNDS

The instantaneous local equilibrium assumption with respect to sorption processes is often made for modeling the transport of organic compounds through subsurface systems. These transport models assume that the sorption kinetics is so rapid that rate considerations can be ignored and that the sorption isotherm is linear (constant K_D). With these assumptions the net effect of sorption is manifested by a reduction in the advective velocity and dispersion coefficient by a constant factor, referred as the retardation factor (2). The retardation factor (R) can be estimated if K_D and porous media characteristics, such as the porosity (ε) and solid density (ρ_s) or bulk density (ρ_B), are known, as shown in Eq. 3:

$$R = 1 + \frac{\rho_B K_D}{\varepsilon} = 1 + \left(\frac{1-\varepsilon}{\varepsilon}\right)\rho_s K_D \tag{3}$$

However, the assumption of instantaneous equilibrium is typically a poor assumption and this often leads to failure of the models in predicting the experimentally observed fate and transport phenomenon. Nonequilibrium sorption can affect the transport of contaminants by causing effects such as early breakthrough, decrease in peak breakthrough concentration, and tailing/asymmetry in the concentration profile.

Sorption rates are typically slow at the intraparticle level. The various mechanisms implicated in slow sorption include (1,10) low film transfer coefficient (k_f) for mass transfer across the boundary layer separating the sorbent from the aqueous phase; slow diffusion through natural organic matter matrices along the sorbent surfaces due to low surface diffusion coefficient (D_S); and slow diffusion through tortuous nanopores due to low pore diffusion coefficient (D_p). However, the relative contribution of the various mechanisms is difficult to predict. The adsorption–desorption bond energies may also play a role in determining the process kinetics. High energy adsorption sites within the organic matter matrix and in the nanopores may be responsible for slow sorption kinetics. A conceptualization of factors affecting adsorption rates is illustrated in Fig. 3. Film transfer and intraparticle diffusion constitute resistances in series, and since intraparticle diffusion is typically slower, it has a dominant impact on the slow kinetics. The two intraparticle diffusion mechanisms—surface and pore diffusion—represent parallel mechanisms, and the faster of the two is thought to control overall mass transfer.

Typically, the sorption and desorption of organics on geosorbents is characterized by a rapid phase, followed by a slow phase—that is, it is bimodal. The magnitude of the slow fraction is quite significant. Pignatello and Xing (10) have effectively illustrated how slow sorption can affect the experimentally determined K_D values. With increasing contact time in equilibrium studies from short (1–3 days) to long (several months), the K_D was demonstrated to increase from 30% to tenfold for a range of organic compounds. The time to reach true equilibrium in batch systems may vary from 1 wk to several months.

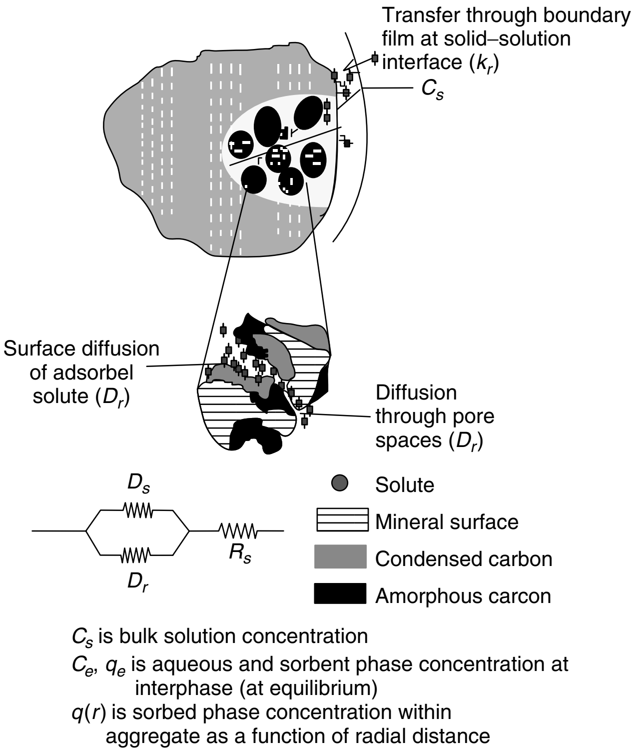

C_s is bulk solution concentration
C_e, q_e is aqueous and sorbent phase concentration at interphase (at equilibrium)
$q(r)$ is sorbed phase concentration within aggregate as a function of radial distance

Figure 3. Rate considerations in sorption processes on a geosorbent.

Slow sorption leads to a chemical fraction that resists desorption. If sorption is allowed to occur over a long time period (aging), the contribution of slow sorption and a desorption resistant fraction increases; that is, the slow fraction following a rapid release may vary from 10% to 96% depending on aging time and the type of geosorbent. The desorption resistant fraction is persistent in natural environments and is also difficult to extract. Thus, for historically contaminated field samples, it is difficult to ensure complete recovery by the extraction techniques employed. Hatzinger and Alexander (11) effectively demonstrated how the biodegradation of phenanthrene and 4-nitrophenol was significantly reduced in aged soils.

Sorption–desorption kinetics is significantly affected by diffusion processes; however, Fickian diffusion alone cannot explain experimental and field-scale observations. It has been reported that the slow fraction is inversely dependent on the initial concentration. This is partly explained by isotherm nonlinearity (Freundlich exponent significantly lower than unity); however, it also appears that intraparticle retardation increases as the concentration within a particle declines. Moreover, sorption is reported to be kinetically hysteretic—the rate of filling the slow sites appears to be faster than the rate at which they are emptied, thus indicating an apparent irreversibility. Several researchers have used stochastic modeling for describing the slow sorption phenomenon on geosorbents; that is, the rate constants and diffusion coefficients are assumed to be distributed according to a statistical density function (e.g., two-parameter gamma distribution).

BIBLIOGRAPHY

1. Weber, W.J., Jr. and DiGiano, F.A. (1996). *Process Dynamics in Environmental Systems*. John Wiley & Sons, Hoboken, NJ.
2. Weber, J.J., Jr., McGinley, P.M., and Katz, L.E. (1991). Sorption phenomena in subsurface systems: concepts, models and effects on contaminant fate and transport. *Water Res.* **25**: 499–528.
3. Lion, W.L., Sauffer, T.B., and MacIntyre, W.G. (1990). Sorption of hydrophobic compounds on aquifer materials: analysis methods and the effect of organic carbon. *J. Contam. Hydrol.* **5**: 215–234.
4. Berthouex, P.M. and Brown, L.C. (1994). *Statistics for Environmental Engineers*. Lewis Publishers CRC Press, Boca Raton, FL, pp. 273–278.
5. Luthy, R.G. et al. (1997). Sequestration of hydrophobic organic contaminants by geosorbents. *Environ. Sci. Technol.* **31**: 3341–3347.
6. Lyman, W.J. (1990). Adsorption coefficient for soils and sediments. In: *Handbook of Chemical Property Estimation Methods*. W.J. Lyman, W.F. Reehl, and D.H. Rosenblatt (Eds.). ACS, Washington, DC, Chap. 4.
7. Schwarzenbach, R.P., Gschwend, P.M., and Imboden, D.M. (1993). *Environmental Organic Chemistry*. John Wiley & Sons, Hoboken, NJ, Chap. 11.
8. Weber, W.J., Jr., Huang, W., and LeBoeuf, E.J. (1999). Geosorbent organic matter and its relationship to the binding and sequestration of organic contaminants. *Colloids Surf. A* **151**: 167–179.
9. Weber, W.J., Jr., McGinley, P.M., and Katz, L.E. (1992). A distributed reactivity model for sorption by soils and sediments. 1. Conceptual basis and equilibrium assessments. *Environ. Sci. Technol.* **26**: 1955–1962.
10. Pignatello, J.J. and Xing, B. (1996). Mechanisms of slow sorption of organic chemicals to natural particles. *Environ. Sci. Technol.* **30**: 1–11.
11. Hatzinger, P.B. and Alexander, M. (1995). Effect of aging of chemicals in soil on their biodegradability and extractability. *Environ. Sci. Technol.* **29**: 537–547.

AGE DATING OLD GROUNDWATER

Gholam A. Kazemi
Shahrood University of Technology
Shahrood, Iran

INTRODUCTION

First, two closely interrelated terms, groundwater age and groundwater residence time, need to be explained and differentiated. Groundwater age is the amount of time that has lapsed since the water molecule, whose age is of concern, was recharged into the system till this molecule is sampled for age-dating. Groundwater residence time is the time it takes for a single molecule of water to travel from the recharge area to the discharge area of the aquifer, the time interval between infiltrating into, and exfiltrating out of, the subsurface media. Groundwater age and residence time are comparable to the age and lifetime of a person, respectively.

Groundwater age and residence time are illustrated in Figs. 1a,1b. Note that some researchers mix up these two

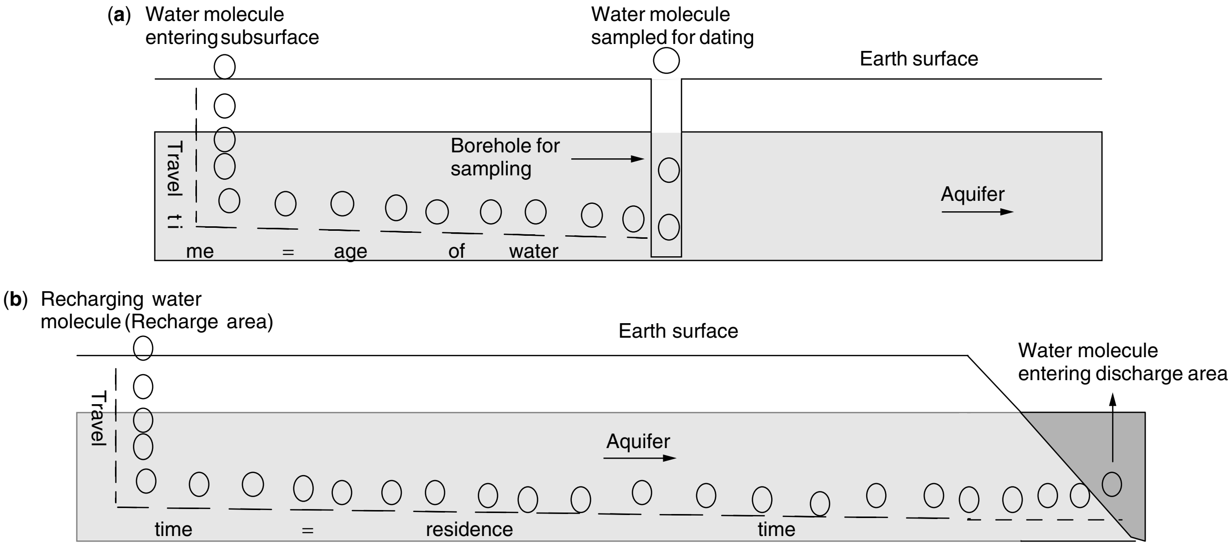

Figure 1. (**a**) Illustration of groundwater age (**b**) Illustration of groundwater residence time.

concepts and do not clearly differentiate between them. Mazor and Nativ (1, p. 219), for instance, refer to the age of groundwater as hydraulic age: “Hydraulic ages relate to travel time, the time needed for a volume of groundwater to move across a flow path Δl.” Similarly, their definition of residence time “or isotopic age” is no different from the definition of “groundwater age” given by many other researchers. A number of workers use slightly different terms for groundwater age such as mean age, transit time, turnover time (2) or mean travel time (3). A recent useful discussion on groundwater age is that by Bethke and Johnson (4).

DISCUSSION

There is neither clear definition nor many citations of the term “old groundwaters” in the literature (in contrast to “young groundwaters”). Examples of limited available references include Coplen (5), Zhao et al. (6), Le Gal La Salle et al. (7), and Park et al. (8).

Old groundwaters may be defined as those that are too old to be dated using techniques like ^{3}H, CFCs, ^{85}Kr, ^{36}Cl/Cl and ^{3}H–^{3}He. These techniques are specifically applied to age-date young groundwater. Four methods have been in use for dating old groundwaters. In the order of the age range they can measure, these include (1) ^{39}Ar, (2) ^{14}C, (3) ^{36}Cl, and (4) ^{4}He. A brief description of each dating technique and example(s) from each method follow. Note that other less known/used methods such as ^{2}H, ^{18}O, ^{32}Si, and ^{81}Kr for dating old groundwater have received very little attention (9).

^{39}Ar

Argon-39 whose half-life is 239 years can be used to date groundwater ages that range from 200 to 1000 years (10). This method has not been used widely.

^{14}C

The ^{14}C method is the most extensively used method for dating old groundwater. ^{14}C, also known as radiocarbon, has a half-life of 5730 years; it can enter groundwater systems by various geochemical and biological processes. The concentration of ^{14}C in soil CO_2 is 100 pmc (percent modern carbon) in prenuclear bomb times (pre-1952) and up to 200 pmc (about 130 pmc, on average) in postbomb years (11). Once water reaches the saturated zone of an aquifer, it is isolated from the atmosphere, and its ^{14}C decays. ^{14}C measurement is useful for dating groundwater ranging in age from 200 to 50,000 years. Values of 90 pmc or more, indicate post-1952 water, and values less than 50 pmc indicates ages up to 50,000 years.

Examples

1. The age of palaeochannel groundwaters of Western Australia is estimated at 20,000 years (12).
2. The ages of groundwater in the upper part of the Mojave River Basin, California, USA, and in the regional aquifer close to the Mojave River were estimated at less than 2400 and 20,000 years, respectively (13).
3. Le Gal La Salle et al. (7) computed that the age of groundwater in the Iullemeden Basin (Niger) is older than 25,000 years.
4. The age of groundwater brines in the Dead Sea area, determined by Yechieli et al., was 9000 years (14).
5. Kazemi (15) calculated that the age of groundwater in various parts of the Buckinbah Creek Catchment, Eastern Australia, is between 970 and 13,000 years.

^{36}Cl

^{36}Cl, whose half-life is 301,000 years, is useful for dating very old groundwater. It is produced naturally in the deep subsurface through thermal neutron capture on the stable nuclide chlorine-35; at and near the land surface by spallation and neutron capture related to secondary cosmic particle bombardment; by cosmic particle interaction with atmospheric gases, particularly argon; and artificially through the activation of ^{35}Cl in sea water by the high flux of neutrons from testing of fusion devices in the western Pacific starting in late 1952 and extending through mid-1958 (16).

Rainfall dissolves atmospheric ^{36}Cl, and carries it to the subsurface where decays, and its initial value is reduced according to the following equation:

$$C = C_0 e^{\lambda t}$$

where C is the concentration of ^{36}Cl at any time t, C_0 is the initial concentration of ^{36}Cl, and λ is the decay constant of $2.25 \times 10^{-6} \alpha^{-1}$. The longer ^{36}Cl resides in the aquifer, the less is its concentration, and hence the older the ambient groundwater. The subsurface production of ^{36}Cl is not large enough to affect the age calculations (8). The age range of the ^{36}Cl method is from 50,000–1,200,000 years.

Example. The age of soil water in Southern Nevada, USA, was estimated at up to 120,000 years (17).

Limitation. If chlorinity of the water sample exceeds ~75–150 mg/kg, the ^{36}Cl method cannot be used for groundwater dating (8). Note that the ratio of ^{36}Cl/Cl is used for dating young groundwater (18).

^{4}He

Helium accumulates in basin groundwater from radioactive decay of the uranium and thorium in the basin strata, as well as from any flux into the basin from the underlying crust and mantle (6). If the rate at which the *in situ* production and basal flux have supplied helium to flowing groundwater is known, then it is possible to calculate the length of time that the groundwater of known helium content has resided in the subsurface. The age range of this method is 10^4 to 10^8 years and due to semiquantitative nature of the method, it is referred to by some researchers as an age indicator rather than a dating method (11).

Example. Using this method, the age of the Baden region of Switzerland was estimated at several hundred thousand years (10).

BIBLIOGRAPHY

1. Mazor, E. and Nativ, R. (1992). Hydraulic calculation of groundwater flow velocity and age: Examination of the basic premises. *J. Hydrol.* **138**: 211–222.
2. Maloszewski, P. and Zuber, A. (1982). Determining the turnover time of groundwater systems with the aid of environmental tracers. 1. Models and their applicability. *J. Hydrol.* **57**: 207–231.
3. Portniaguine, O. and Solomon, D.K. (1998). Parameter estimation using groundwater age and head data, Cape Cod, Massachusetts. *Water Resour. Res.* **34**(4): 637–645.
4. Bethke, C.M. and Johnson, T.M. (2002). Paradox of groundwater age: Correction. *Geology* **30**(4): 385–388.
5. Coplen, T.B. (1993). Uses of environmental isotopes. In: *Regional Ground-water Quality*. W.M. Alley (Ed.). ITP, New York, pp. 255–293.
6. Zhao, X. et al. (1998). Controls on the distribution and isotopic composition of helium in deep ground-water flows. *Geology* **26**(4): 291–294.
7. Le Gal La Salle, C., Fontes, J.C., Andrews, J.N., Schroeter, P., Karbo, A., and Fifield, K.L. (1995). Old groundwater circulation in the Iullemeden basin (Niger): Preliminary results of an isotopic study. In: *Application of Tracers in Arid-Zone Hydrology*. E.M. Adar and C. Leibundgut (Eds.). *IAHS publications* no. 232, IAEA, Vienna, pp. 129–139.
8. Park, J., Bethke, C.M., Torgerson, T., and Johnson, T.M. (2002). Transport modeling applied to the interpretation of groundwater ^{36}Cl age. *Water Resour. Res.* **38**(5): 1–15.
9. Plummer, L.N., Michel, R.L., Thurman, E.M., and Glynn, P.D. (1993). Environmental tracers for age dating young groundwater. In: *Regional Ground-Water Quality*. W.M. Alley (Ed.). ITP, New York, pp. 227–254.
10. Mazor, E., Jaffe, F.C., Fluck, J., and Dubois, J.D. (1986). Tritium corrected ^{14}C and atmospheric noble gas corrected ^{4}He applied to deduce ages of mixed groundwaters: Examples from the Baden region, Switzerland. *Geochim. Cosmochim. Acta* **50**: 1611–1618.
11. Mazor, E. (1991). *Applied Chemical and Isotopic Groundwater Hydrology*. Open University Press, Milton Keynes.
12. Turner, J.V., Rosen, M.R., Fifield, L.K., and Allan, G.L. (1995). Chloride-36 in hypersaline palaeochannel groundwaters of Western Australia. In: *Application of Tracers in Arid-Zone Hydrology*. E.M. Adar and C. Leibundgut (Eds.). *IAHS publications* no. 232, IAEA, Vienna, pp. 15–33.
13. Izbicki, J.A., Martin, P., and Michel, R.L. (1995). Source, movement and age of groundwater in the upper part of the Mojave River Basin, California, USA. In: *Application of Tracers in Arid-Zone Hydrology*. E.M. Adar and C. Leibundgut (Eds.). *IAHS publications* no. 232, IAEA, Vienna, pp. 43–56.
14. Yechieli, Y., Ronen, D., and Kaufman, A. (1996). The source and age of groundwater brines in the Dead Sea area, as deduced from ^{36}Cl and ^{14}C. *Geochimica Cosmochim. Acta* **60**(11): 1909–1916.
15. Kazemi, G.A. (1999). *Groundwater Factors in the Management of Dryland and Stream Salinity in the Upper Macquarie Valley, New South Wales, Australia*. Unpublished Ph.D. Thesis, University of Technology, Sydney, Australia.
16. Davis, S.N., Moysey, S., Cecil, L.D., and Zreda, M. (2003). Chlorine-36 in groundwater of the United States: Empirical data. *Hydrogeol. J.* **11**: 217–227.
17. Tyler, S.W., Chapman, J.B., Conrad, S.H., and Hammermeister, D. (1995). Palaeoclimate response of a deep vadose zone in southern Nevada, USA, as inferred from soil water tracers. In: *Application of Tracers in Arid-Zone Hydrology*. E.M. Adar and C. Leibundgut (Eds.). *IAHS publications* no. 232, IAEA, Vienna, pp. 351–361.
18. Cook, P.G. and Solomon, D.K. (1997). Recent advances in dating young groundwater: Chlorofluorocarbons, ^{3}H/^{3}He and ^{85}Kr. *J. Hydrol.* **191**: 245–265.

AMMONIA

Robert M. Burgess
U.S. Environmental Protection
Agency
Narragansett, Rhode Island

Ammonia is a primary breakdown product of nitrogen-containing organic matter in aquatic environments. Because a great deal of organic matter and waste is introduced into the hydrologic cycle by natural and anthropogenic processes, ammonia can be found in aquatic environments, especially those exposed to large amounts of human or domestic animal wastes. Ammonia is also of great importance in understanding the global cycling of several key elements, including nitrogen and carbon. Ammonia is one of the compound uniquely critical to life which when not present in adequate amounts in aquatic environments can be associated with negative biological effects and when in excess can result in adverse ecological and toxic effects. When insufficient ammonia is available to plants, which assimilate ammonia to grow, they may experience metabolic problems due to nutrient limitations. Conversely, excess ammonia is commonly associated with eutrophic conditions in water bodies and toxicity to some organisms. Further, nitrous oxide (N_2O) a molecule involved in the environmental cycling of ammonia is considered a potential 'greenhouse gas' which can also destroy stratospheric ozone. Because of these adverse effects, ammonia is frequently considered a pollutant in aquatic systems.

BACKGROUND

The word 'ammonia' first came into chemical-related use in approximately 1799 and was derived from the Egyptian deity, Ammon. Ammonia referred to the gas emitted by camel excreta in the vicinity of temples honoring Ammon. The form of ammonia known as ammonium ($NH_4{}^+$) was named by the eminent Swedish chemist Berzelius in 1808.

Today, it is known that ammonia consists of one nitrogen atom and three hydrogen atoms (NH_3) and has a molecular weight of 17. Ammonia is formed from the bacterial decomposition of organic matter containing nitrogen. For example, the degradation of the simple organic molecule urea [$(NH_2)_2CO$] in water:

$$(NH_2)_2CO(l) + H_2O(l) \longrightarrow 2\,NH_3(g) + CO_2(g) \qquad (1)$$

results in the formation of the gases, un-ionized ammonia (NH_3) and carbon dioxide (CO_2). In the same way, complex organic molecules in industrial and municipal wastes and plant and animal matter also form NH_3 when degraded. In aqueous solution, ammonia exists in two forms as a function of pH, ionic strength, and temperature:

$$NH_4{}^+ + OH^- \longleftrightarrow NH_3 + H_2O \qquad (2)$$

Ionic strength and temperature influence ammonia speciation, but pH is the controlling variable. Consequently, with a base dissociation constant (pK_b) of 9.5, at a pH less than 9.5, most of the ammonia in solution is ionic ammonium ($NH_4{}^+$); at a pH greater than 9.5, the majority is NH_3. Consequently, in seawater and freshwater, the majority of ammonia is in the $NH_4{}^+$ form.

CYCLING

As noted above, ammonia is a key component of the greater global nitrogen cycle. As shown in Fig. 1, the fixation of nitrogen gas (N_2) forms NH_3, and the ammonification of organic nitrogen also results in forming ammonia. Ammonia is assimilated by plants and used to form organic nitrogen molecules, including amino acids, and then proteins. Conversely, ammonia can be formed by the ammonification of amino acids and proteins. Through nitrification, ammonia can also be oxidized to form nitrite ($NO_2{}^-$) which is further oxidized to form nitrate ($NO_3{}^-$). Nitrate is another nitrogen molecule that is assimilated by plants. Anaerobic conditions retard nitrification and cause an elevation of NH_3 concentrations. Denitrification of nitrite and nitrate results in forming N_2 and several nitrogen oxides, including N_2O, which completes the nitrogen cycle. Nitrification, denitrification, and ammonification are all reactions that occur abiotically but can also be moderated by bacteria. In nature, fixation requires microbial mediation. Synthetic fixation can be performed using the process developed by Haber and Bosch in 1910–1914 to produce fertilizer. Energetically, the path leading to the formation of organic nitrogen compounds is most favored compared to the reverse path resulting in nitrogen gas (Fig.1).

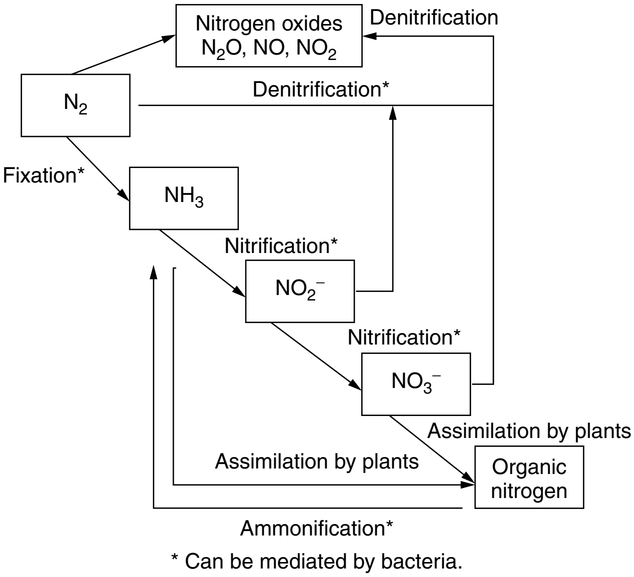

Figure 1. Nitrogen cycle.

Equation 1 gives an example of NH_3 formed by ammonification. The formation of NH_3 by fixation of N_2, as shown in Equation 3, is more complicated and involves linking the nitrogen cycle to the carbon cycle. In this reaction, a simple carbon molecule (CH_2O) is oxidized to form CO_2; the inclusion of N_2 and an exchange of hydrogen ions (H^+) and electrons (e^-) results in ammonium formation:

$$CH_2O(s) + H_2O(l) + N_2(g) + 8H^+ + 6e^- \longrightarrow CO_2(g) + 4H^+ + 4e^- + 2\,NH_4{}^+(l) \qquad (3)$$

Without the involvement of the carbon cycle, this reaction would not be thermodynamically favored, and N_2 fixation would not take place. Because Equation 3 is thermodynamically favored, microbes mediate the reaction and acquire energy.

SOURCES AND CONCENTRATIONS

Several studies have attempted to quantify ammonia sources and concentrations around the world. In these studies, the atmosphere is considered the ultimate sink for global ammonia, although, it should be noted, other studies speculate that the oceans are the ultimate sink. Primary sources of ammonia to the global environment from four different studies are listed in Table 1. Despite the different information used to assemble the table, the sources show remarkable similarity. The range of total ammonia released to the environment is 45 to 75 TgN yr^{-1}. Clearly, domesticated animals are the largest sources of ammonia globally, followed by soil, the oceans, and fertilizers. Lesser amounts of ammonia originate from biomass burning and human waste, and even lesser from fuel combustion and automobiles. The primary sources of ammonia are all systems containing bacteria that significantly enhance the nitrogen fixation and ammonification reactions discussed before.

As may be expected, concentrations of ammonia in aquatic systems vary based on sources and environmental conditions (Table 2).

Overlying fresh and marine waters that are well mixed, aerobic, and lack a direct source of ammonia maintain relatively low concentrations in the high parts per billion range ($\mu g\ L^{-1}$ or ppb); sediment interstitial waters, especially those that are anoxic, support elevated levels in the lower parts per million range ($mg\ L^{-1}$ or ppm). Sewage effluents and sludges can have extremely high concentrations of ammonia from human and animal waste. As a concentrated form of nitrogen-containing organic matter, sludge is potentially very elevated in concentrations of ammonia. Because of the elevated concentrations of ammonia in sewage effluents, waste treatment facilities frequently aerate the liquid to enhance nitrification and counter anaerobic conditions. Finally, rainwater can have elevated ammonia concentrations as a result of the diffusion of NH_3 into precipitation from the atmosphere.

Table 1. Sources of Ammonia in the Environment

	Amount, TgN yr^{-1}[a]			
Source	Study 1	Study 2	Study 3	Study 4
Domestic animals	32	26	24.5	22
Ocean surface	13	—	7	8
Undisturbed soils	10	15	5.1	6
Fertilizers	9	3	6.4	9
Biomass burning	5	5	2	6
Human waste	4	3	—	3
Coal combustion	2	< 2	—	0.1
Automobiles	0.2	0.2	—	0.2
Total	75	54	45	54

[a] Teragrams nitrogen per year (tera = 10^{12}).

Table 2. Ammonia Concentrations in Aquatic Samples

Aquatic sample	Total Ammonia Concentration ($mg\ L^{-1}$)
Coastal marine waters	0.24–0.58
Freshwaters	<1.0
Marine sediment interstitial waters	1–11
Freshwater sediment interstitial waters	1–53
Sewage effluent	16–33
Sewage sludge	1000–2030[a] ($mg\ kg^{-1}$)
Rainwater	0.05–2.60[a]

[a] Expressed as NH_4^+.

EUTROPHICATION

Eutrophication is the process of enriching a water body with nutrients (see *Eutrophication*). In marine systems, nitrogen-containing compounds (e.g., NO_3^-, NH_3) most often cause eutrophic conditions, in freshwater systems, phosphorus-containing compounds are currently considered most important. In most instances of eutrophication, anthropogenic sources of nutrients, including fertilizers and domestic animals, are suspected as the causes. Consequently, eutrophication is considered a major environmental problem in areas where human populations are present. Eutrophication causes several adverse environmental effects. The first effect is the occurrence of large algal blooms. These blooms can reduce light penetration into the water throughout the system which may adversely affect other vegetative life. Next, the accumulation of decaying organic matter enhances the growth of bacteria which causes a decline in water column dissolved oxygen concentrations. The resulting anaerobic conditions can kill aquatic animals and foster an increase in ammonia concentrations in the water and adjacent sediments. Because anaerobic environments support the persistence of ammonia, it is frequently associated with eutrophic conditions. Thus, the presence of ammonia perpetuates the potential for eutrophic conditions to arise. Another adverse effect of eutrophication and elevated concentrations of ammonia is volatilization of NH_3 to the atmosphere followed by transport and deposition in other regions. The occurrence of eutrophication has many implications, including severe disruptions of ecological processes and potential ammonia-related toxicity to aquatic organisms.

TOXICITY

Ammonia toxicity to aquatic organisms is extremely well documented. Table 3 has a small collection of selected ammonia toxicity values. The toxicity of ammonia is dominated by the NH_3 form; NH_4^+ demonstrates little contribution to overall toxicity. Because the concentration of NH_3 is a direct function of solution pH, very frequently, the toxicity of total ammonia to organisms is observed to increase as pH increases. This trend is explained by the commensurate increase in NH_3 concentration as pH increases. Interestingly, the toxicity of NH_3 actually decreases as pH increases, indicating that NH_3 is less toxic at higher pHs.

Ammonia toxicity, as measured by effects on survival, tends to occur in the low parts per million concentration range, although this varies by species. Sublethal effects to

Table 3. Ammonia toxicity to aquatic organisms

Species	Medium	End Point	Toxicity Value[a] (mg/L)
Mysid shrimp (*Americamysis bahia*)	Marine	Mortality	2.82
Mysid shrimp (*Americamysis bahia*)	Marine	Sublethal	0.23
American lobster (*Homarus americanus*)	Marine	Mortality	2.21
Inland silverside (*Menidia beryllina*)	Marine	Survival	0.88
Inland silverside (*Menidia beryllina*)	Marine	Sublethal	0.06
Striped bass (*Morone saxatilus*)	Marine	Mortality	0.91
Winter flounder (*Pseudpleuronectes americanus*)	Marine	Mortality	0.53
Cladoceran (*Daphnia magna*)	Freshwater	Mortality	2.69
Cladoceran (*Daphnia magna*)	Freshwater	Sublethal	0.63
Rainbow trout (*Oncorhynchus mykiss*)	Freshwater	Mortality	0.39
Rainbow trout (*Oncorhynchus mykiss*)	Freshwater	Sublethal	0.02
Smallmouth bass (*Micropterus dolomieui*)	Freshwater	Mortality	1.01
Smallmouth bass (*Micropterus dolomieui*)	Freshwater	Sublethal	0.60

[a] Mortality values are LC_{50}s; sublethal values are EC_{50}s.

aquatic organisms, including effects on growth and reproduction, occur in the high parts per billion concentration range but also vary by species. In general, fish species in both fresh and marine waters are the most sensitive to ammonia, as compared to invertebrate species. For example, Table 3 shows that in marine waters, the inland silverside (a minnow) is approximately three times more sensitive to ammonia than the mysid shrimp, based on a survival end point. Similarly, in freshwater, the rainbow trout is about seven times more sensitive than a cladoceran (Table 3). These differences in sensitivity are also apparent based on sublethal effects; for example, sublethal effects in trout occur at concentrations 30 times lower than sublethal effects in a cladoceran (Table 3)! The extreme sensitivity of fish species to ammonia has significant implications for the aquaculture and fisheries industries with regard to ensuring that their wastes do not adversely affect their products.

Because nitrogen-containing organic matter accumulates in sedimentary aquatic environments, elevated concentrations of ammonia are often found in sediments. This is especially true in eutrophic settings enriched with anthropogenically derived nutrients. Consequently, concerns about ammonia toxicity are frequently focused on sediment-dwelling organisms and not the water column species in aquatic systems. Of course, if there is evidence that ammonia from point (e.g., sewage treatment plant) or nonpoint sources is being released into aquatic systems, ammonia toxicity may be expected.

ANALYSIS

The two most common ways to measure ammonia in aqueous solution are by spectrometry and ion-selective electrodes. Both methods function well in freshwater and seawater, including clean and polluted samples. The spectroscopic method involves a reaction between ammonia and phenol and sodium hypochlorite with sodium nitroprusside to form indophenol blue which is measured by spectrometry. A variation on this method substitutes sodium salicylate for the phenol, eliminating the production of *o*-chlorophenol, a toxic substance, and resulting in comparable sensitivity. In the ion-selective method, changes in the measured electrode potential are compared to a reference potential. These potentials are proportional to the amount of NH_3 in the sample solution allowing for a determination of ammonia concentration. Both methods have advantages and disadvantages. The spectrometric method is the more sensitive technique, detection limits are in the mid- to high parts per billion range, but it suffers from being a relatively slow method which uses somewhat hazardous reagents. The ion-selective method is relatively fast allowing analyses to be completed in minutes but is relatively insensitive; detection limits are in the upper parts per billion to lower parts per million range. Both methods also differ in terms of the form of ammonia measured. Spectrometric methods measure NH_4^+ whereas the ion-selective method measures NH_3. Given that the two forms of ammonia behave differently in the environment, it is important to understand which is being measured.

READING LIST

Historical

Lewis, R.J. (1993). *Hawley's Condensed Chemical Dictionary*. Van Nostrand Reinhold Company, New York, p. 1275.

Partington, J.R. (1989). *A Short History of Chemistry*. Dover, New York, p. 415.

General Chemistry

Hampson, B.L. (1977). Relationship between total ammonia and free ammonia in terrestrial and ocean waters. *Journal of Cons. in Exploration of the Mer.* **37**: 117–122.

Khoo, K.H., Culberson, C.H., and Bates, R.G. (1977). Thermodynamics of the dissociation of ammonium ion in seawater from 5 to 40 °C. *Journal of Solution Chemistry* **6**: 281–290.

O'Neill, P. (1985). *Environmental Chemistry*. George Allen and Unwin, Boston, MA, p. 232.

Sources and Concentrations

Bini, C. and Bresolin, F. (1998). Soil acidification by acid rain in forest ecosystems: a case study in northern Italy. *The Science of the Total Environment* **222**: 1–15.

Friedl, G., Dinkel, C., and Wehrli, B. (1998). Benthic fluxes of nutrients in the northwestern Black Sea. *Marine Chemistry* **62**: 77–88.

Park, G.S. and Park, S.Y. (2000). Long-term trends and temporal heterogeneity of water quality in tidally mixed estuarine waters. *Marine Pollution Bulletin* **40**: 1201–1209.

Phillips, B.M., Anderson, B.S., and Hunt, J.W. (1997). Measurement and distribution of interstitial and overlying water ammonia and hydrogen sulfide in sediment toxicity tests. *Marine Environmental Research* **44**: 117–126.

Rosenfeld, J.K. (1979). Ammonium adsorption in nearshore anoxic sediments. *Limnology and Oceanography* **24**: 356–364.

Sarda, N. and Burton, G.A. (1995). Ammonia variation in sediments: spatial, temporal and method-related effects. *Environmental Toxicology and Chemistry* **14**: 1499–1506.

Schlesinger, W.H. and Hartley, A.E. (1992). A global budget for atmospheric NH_3. *Biogeochemistry* **15**: 191–211.

Valiela, I. (1984). *Marine Ecological Processes*. Springer-Verlag, New York, p. 546.

Wetzel, R.G. (1983). *Limnology*. W.B. CBS College Publishing, Saunders Company, New York, p. 767.

Wong, J.W.C., Li, K., Fang, M., and Su, D.C. (2001). Toxicity evaluation of sewage sludges in Hong Kong. *Environment International* **27**: 373–380.

Eutrophication and Toxicity

Ankley, G.T., Katko, A., and Arthur, J.W. (1990). Identification of ammonia as an important sediment-associated toxicant in the Lower Fox River and Green Bay, Wisconsin. *Environmental Toxicology and Chemistry* **9**: 313–322.

Burgess, R.M. et al. (2003). Removal of ammonia toxicity in marine sediment TIEs: a comparison of ulva lactuca, zeolite and aeration methods. *Marine Pollution Bulletin* **46**: 607–618.

Erickson, R.J. (1985). An evaluation of mathematical models for the effects of pH and temperature on ammonia toxicity to aquatic organisms. *Water Research* **19**: 1047–1058.

Kohn, N.P. et al. (1994). Acute toxicity of ammonia to four species of marine amphipod. *Marine Environmental Research* **38**: 1–15.

Morell, J.M. et al. (2001). Nitrous oxide fluxes in Caribbean and tropical Atlantic waters: evidence for near surface production. *Marine Chemistry* **74**: 131–143.

National Research Council. (2000). *Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution*. National Academy of Sciences, National Academy Press, Washington, DC, p. 377.

Russo, R.C. (1985). Ammonia, Nitrite, and Nitrate. In: *Fundamentals of Aquatic Toxicology*. G.M. Rand and S.R. Petrocelli (Eds.). Hemisphere Publishing Corporation, New York, p. 666.

U.S. Environmental Protection Agency. (1989). *Ambient Water Quality Criteria for Ammonia (Saltwater)—1989*. EPA 440-5-88-004. Office of Regulation and Standards, Washington, DC.

U.S. Environmental Protection Agency. (1998). *1998 Update of Ambient Water Quality Criteria for Ammonia*. EPA 822-R-98-008. Office of Water, Washington, DC.

Viessman, W. and Hammer, M.J. (1985). *Water Supply and Pollution Control*. Harper & Row Publishers, New York, p. 797.

Analytical Chemistry

Bower, C.E. and Holm-Hansen, T. (1980). A salicylate-hypochlorite method for determining ammonia in seawater. *Canadian Journal of Fisheries and Aquatic Science* **37**: 794–798.

Gilbert, T.R. and Clay, A.M. (1973). Determination of ammonia in aquaria and in sea water using the ammonia electrode. *Analytical Chemistry* **45**: 1757–1759.

BERYLLIUM IN WATER

JAY CLAUSEN
AMEC Earth and
Environmental, Inc.
Westford, Massachusetts

INTRODUCTION

Beryllium was discovered by Vauquelin in 1798 as an oxide in beryl and emeralds. The metal was isolated in 1828 by Bussy and Wohler but its use was limited until it was discovered that a mixture with copper produced an alloy six times stronger than steel (1). Beryllium is a Group 2 (IIA) element found naturally as a rare alkaline earth metal in the earth's crust and is the lightest metal. As a pure metal, beryllium has a steel gray color and is present in a variety of materials, such as coal, oil, rock, soil, and volcanic dust. It can be found concentrated in igneous intrusive rocks such as granites, specifically beryllium-rich pegmatite. Beryllium ions are small enough to substitute for silicon in igneous rocks. Beryllium is found in some 50 mineral species; the most important are beryl [$Be_3Al_2(SiO_3)_6$], bertrandite [$Be_4Si_2O_7(OH)_2$], chrysoberyl [$BeAl_2O_4$], and phenacite [Be_2SiO_4]. These minerals can be present at elevated concentrations in water due to the weathering of the host rock. Beryllium, however, does not dissolve readily under natural conditions. For this reason, precious forms of beryl, better known as either aquamarine (blue or blue-green) or emerald (green), can be found in nature.

In addition to its presence in its natural state, beryllium can also be redistributed into the environment through anthropogenic activities such as burning coal (2–4) and oil and disposal of metal manufacturing waste (5). Prior to the 1940s, there were few applications for beryllium. Specialized metal components requiring lightweight and strength have resulted in its increased usage during the last 60 years. Most of the beryllium ore that is mined is converted into alloys (mixtures of metals). Beryllium can be found in computer, aircraft, spacecraft, and automotive parts; gyroscopes; sports equipment; nuclear reactor reflectors; X-ray machines; mirrors; and other specialized engineering components.

Beryllium metal and its salts are extremely toxic. Inhalation of fine beryllium dust can result in the lung disease, berylliosis, also known as chronic beryllium disease, and possibly cancer (6–9). The U.S. Department of Health and Human Services lists beryllium as a Class A carcinogen (2). As a consequence, beryllium is typically included as one of several metals analyzed as part of environmental investigations.

Numerous analytical methods are available for analyzing beryllium: inductively coupled plasma-mass spectrometry (ICP-MS), optical emission spectrometry (ICP-OES), atomic absorption spectrometry (AAS), colorimetric, and fluorimetric methods (10,11). A detailed listing of the colorimetric methods can be found in Taylor et al. (12). The preferred method for soil and water analysis in the environmental business is ICP-MS following United

States Environmental Protection Agency (USEPA) methods 200.7 and 6010. The ICP-MS instrumentation has greater sensitivity and reduced interference from other metals, compared to other analytical methods. A typical reporting limit of beryllium in environmental investigations using the ICP-MS method is 5 ppb; typical method of detection limits for soil and water are 0.1 μg/kg and 0.7 μg/L, respectively. AAS is also an USEPA approved process following method 7090 or 7091. Due to beryllium's low concentration and subsequent interference with other metals in environmental samples, a number of separation and preconcentration methods have been developed using chelating agents, ion-exchange resins, sorbants, and surfactants.

CHEMISTRY

Most beryllium oxides and hydroxides have very low solubilities, although the chloride and sulfate oxides are very soluble. However, the chloride and sulfate oxides hydrolyze to relatively insoluble hydroxide species. Baes and Mesmer (13) found that beryllium hydroxide solubilities range from 0.9 to 100 μg/L in the pH range of 6 to 8.5. Iron, manganese, and zinc beryllium silicates are believed more soluble because they contain divalent sulfur, which can be easily oxidized (14).

The dominant dissolved species found under environmental conditions are Be^{2+}, $BeOH^+$, $Be(OH)_2$, and $Be(OH)_3^-$. Thermodynamic data for the dominant beryllium species are presented in Table 1.

Figure 1, a phase diagram for beryllium under normal temperatures and atmospheric conditions at a concentration of 90 μg/L, suggests the solid species of BeO dissolves to form Be^{2+} under acidic conditions. Increasing the concentration of beryllium in solution results in a shift of the BeO to the Be^{2+} transition zone at higher pH. However, this is contrary to known environmental conditions, whereby beryllium can form insoluble mineral species such as beryl, emerald, and bertrandite. Brookins (18) suggested that the discrepancy between thermodynamic calculations and reality may be attributable to the presence of metastable species, which affect the solubility of the various beryllium species.

Table 1. Thermodynamic Data for Beryllium

Species[a,b]	$\Delta G^0{}_f$, kcal/gfw[c]	Reference
Be (metal)	0	15
Be^{2+} (aq)	− 90.76	16
BeO (s)	−138.70	16
$BeO_2{}^{2-}$ (aq)	−152.99	16
$Be_2O_3{}^{2-}$ (aq)	−298	17
Be_2O^{2+} (aq)	−218	17
$Be(OH)_2$ (s)	−194.79	16
$BeO \cdot Be(OH)_{2(s)}$	−338	17
BeS (s)	−55.9	17
$BeCl_2$ (s)	−111.8	17
$BeSO_4$ (s)	−260.2	17

[a] Aq: aqueous.

[b] s: solid.

[c] Kcal/gfw–kilo calories per gram formula weight.

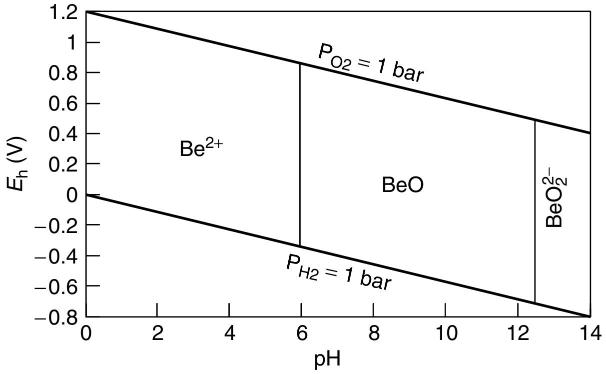

Figure 1. E_h–pH diagram for the Be–O–H system at a concentration of 90 μg/L under standard temperature and pressure.

SOURCE MATERIAL

Beryllium is a widely distributed metal that has a crustal average of 3 ppm, and its average concentration is 5 ppm in granitic rock, according to Krauskopf (19). The weathering product of rock is typically a kaolinite or hydrolysate, which accumulates in the B-soil horizon. Hydrobiotite, a feldspar mineral, weathers to montmorillonite, which sorbs beryllium (20). Beryllium concentrations in soil often correlate with aluminum and magnesium levels because both metals can be involved in cation exchange reactions with montmorillonite minerals. Berg et al. (21) found a correlation between aluminum and beryllium in a study near an aluminum smelter. Similarly, a study by Asami and Kubota (22) found a weak correlation, 0.505 at a level of significance of 0.1% for n=113 samples, between beryllium and aluminum for soils collected from several different countries. An earlier study by Asami and Fukazawa (23) found a similar relationship for soils collected in Japan near a beryllium manufacturing facility.

A number of studies have been conducted to determine the range of beryllium concentrations in surficial soils (0.3 to 6 ppm) throughout the world (22,24–34). A study by Anderson et al. (35) of four eastern U.S. Piedmont soils suggested higher beryllium concentration ranges for surface soil of 6.3 to 13.3 and 7 to 30.5 mg/kg for deeper soils. However, the number of samples collected was very small and insufficient to develop statistically meaningful values. In a study in Germany, Grote et al. (25) found beryllium levels range from 0.05 to 1 mg/kg for rural soils and 1 to 2.5 mg/kg for urban soils. This study suggests that elevated levels of beryllium can be found as a result of anthropogenic activities.

Fishbein (36) reports beryllium can be found in coal at concentrations ranging from 0.1 to 1000 ppm and in oil at a concentration of 0.8 ppb. Data of beryllium concentration in different types of U.S. coal by weight are presented in USEPA (37) and summarized in Taylor et al. (12). Weathering of soil and rock and combustion of oil and coal can account for the presence of beryllium in water.

Surface Water

A number of studies have been conducted on surface waters such as lakes, rivers, and seawater (Table 2).

Table 2. Concentration of Beryllium in Various Waters

Source of Water	Location[a]	n	Mean, μg/L	Range, μg/L[b]	Reference
Seawater	NA	NA	0.0006	NA	38
Seawater	NA	NA	NA	0.00036–0.009	39
River water	NA	NA	NA	ND–0.01	40
River water	US/Canada	59	NA	ND–0.22	41
River water	US	NA	<0.05	ND–0.28	42
River water	US, downgradient of manufacturing plants	NA	NA	18–21	36
River water	Czechoslovakia—Be rich bedrock, acid rain	39	11.97	2–24	43
Surface water	US	1577	0.19	0.01–1.22	44
Surface water	US	NA	NA	ND–1.2	42
Surface water	Eastern US/Siberia	NA	NA	0.1–0.9	45
Surface water	USSR	NA	NA	ND–0.1	20
Surface water	Czechoslovakia	3028	0.0054	0.0022–65	46
Finished water	US	380	0.1	0.02–0.17	44
Finished water	US	NA	NA	0.01–0.7	47
Finished water	US	NA	NA	ND–0.17	42
Finished water	US	NA	0.2	0.01–1.2	36
Bottle water	Europe	NA	NA	0–2.5	11
Groundwater	USSR, granitic area	NA	NA	0.01–30	20
Groundwater	India, downgradient of Be manufacturing plant	NA	NA	0.001–0.002	24

[a] NA: not available.
[b] ND: nondetectable.

In some cases, it appears that "finished water," treated drinking water, was analyzed. However, the studies in many cases do not indicate whether the original source of water was a river, lake, or groundwater. A number of studies also state that surface water was sampled but do not elaborate if this consisted of river or lake water.

A study of 59 surface water bodies and 15 rivers in the United States and Canada found a maximum beryllium concentration of 0.22 μg/L (41). However, the study did not specify which waterbodies were studied and whether raw or finished water was analyzed. Similarly, Kopp (44) analyzed 1577 surface water samples from the United States and found a concentration range of 0.01 to 1.22 μg/L and a mean of 0.19 μg/L. Griffitts et al. (45) reported a range of beryllium of 0.1 to 0.9 μg/L in eastern U.S. and Siberian surface waters. Soviet surface water rarely exceeded 0.1 μg/L, according to Grigor'yev (20).

In Hem's (48) review of several studies, it was concluded "that river waters normally can be expected to carry no more than a few tenths of a microgram per liter of beryllium in solution." However, elevated beryllium concentrations were found in streams where the beryllium-rich source rock was being attacked by acid rain (46). Skirivan et al. (43) suggest that the concentration of fluoride ions and pH significantly control beryllium concentrations in surface waters.

Studies by Skirivan et al. (46) and Burba et al. (49) suggest beryllium can be sorbed onto suspended organic matter. The work of Grigor'yev (20) found between 50 and 73% of the total beryllium in river water samples is associated with suspended solids, which supports the Burba et al. (49) findings. However, some streams at low pH contain dissolved beryllium. Upon discharge of low pH stream water into a higher pH pond, the mobile beryllium was scavenged from the water by solid species containing hydrated metal oxides (aluminum, iron, and manganese) residing on organic matter (43,46,50,51). At higher pH, the beryllium complexes of $BeOH^+$ and $Be(OH)_2$ form, which then can be adsorbed onto colloidal particles (52). Skirivan et al. (43) and Veseley et al. (46) found the highest beryllium concentrations in surface water were associated with low pH waters. At a pH less than 4, beryllium is principally found as the Be^{2+} cation (43,52).

Fishbein (36) found beryllium levels of 18 to 21 μg/L from effluent samples collected downgradient of power plants and metal manufacturing facilities, although the locations were not specified. This study suggests that elevated beryllium levels in surface water are a possibility near metal manufacturing facilities.

Groundwater

Very limited information is available on the concentration of beryllium in groundwater in the general literature. Data is likely to be available from environmental studies of potential contaminated sites; however, this information is not readily accessible.

No general regional baseline studies of beryllium in groundwater have been published in the peer-reviewed literature. However, Grigor'yev (20) studied groundwater from a granitic intrusive area and found levels range from 0.01 to 30 μg/L. Additionally, a study by Korkish et al. (11) was conducted with European bottled water, which likely originated as ground-fed spring water. The Korkish et al. (11) study found beryllium levels range from nondetectable to 2.25 μg/L. Burba et al. (49) present data on two mineral water samples collected in Germany, which

Table 3. Beryllium Groundwater Concentrations for Paducah, KY[a]

Formation	Depth	n	Mean, μg/L	Range, μg/L
Upper continental recharge system	shallow	83	0.92	0.2–3
Regional gravel aquifer	mid	237	1.0	0.04–4.4
McNairy	deep	14	0.76	0.04–2.2

[a] Reference 53.

again presumably have a groundwater origin. A sample from Essen had a beryllium concentration of 9 μg/L, and a sample from Bad Pyrmont had a level of 55.8 μg/L.

To place the previous discussion in context, the US EPA maximum contaminant level (MCL) for beryllium is 4 μg/L. The MCL is the highest level of a contaminant that is allowed in drinking water and is an enforceable standard. The drinking water equivalent level (DWEL) for beryllium is 70 μg/L. A DWEL is a lifetime exposure concentration protective of adverse, noncancer health effects, which assumes that all exposure to a contaminant is from drinking water.

Several studies of potential contamination sites have been conducted where beryllium was one of several metals analyzed. One specific study focused on a beryllium manufacturing plant in India (34). This study found groundwater concentrations ranged from 0.001 to 0.002 μg/L, which suggests the beryllium levels represent background conditions.

Another detailed groundwater study (Table 3) was conducted at an uranium gaseous diffusion facility and included beryllium data for groundwater in a shallow unconfined system, a mid-depth confined aquifer, and a deep confined aquifer (53). The pH of this low ionic strength groundwater ranged from 5.3 to 7.9 and was oxidizing. No known source of beryllium existed at this facility, so the beryllium concentrations were likely representative of ambient background conditions.

CONCLUSION

The available information on the presence of beryllium in water samples is spotty, many of the studies lack specific information on type of water sample (lake, river, or groundwater), where, when, and why samples were collected. However, taken as a whole, the studies indicate that the presence of beryllium in water samples should be low at concentrations in the tenths to single ppb range. Beryllium water concentrations near landfills, metal manufacturing sites, coal regions, or locations of beryllium-enriched pegmatitic rock can have elevated levels, typically at the tens of ppb level. Higher beryllium water concentrations are limited by the low solubility of most mineral species.

BIBLIOGRAPHY

1. Walsh, K.A. (1973). Beryllium. In: *The Encyclopedia of Chemistry*, 3rd Edn. Van Nostrand Reinhold, New York.
2. USDHHS. (2002). 10th Report on Carcinogens. U.S. Department of Health and Human Services (USDHHS), Public Health Service, National Toxicology Program. Available: http://ehis.niehs.nih.gov (accessed July 2003).
3. Nriagu, J.O. (1990). Global metal pollution. *Environment* **32**: 7–33.
4. Nriagu, J.O. and Pacyna, J.M. (1988). Quantitative assessment of worldwide contamination of air, water, and soils by trace metals. *Nature* **333**: 135–139.
5. Bhat, P.N. and Pillai, K.C. (1997). Leachability and immobilization of beryllium from solid waste (red-mud) generated in processing beryl. *Water Air Soil Pollut.* **94**: 297–306.
6. Eisenbud, M. (1998). The standard for control of chronic beryllium disease. *Appl. Occup. Environ. Hyg.* **12**: 25–31.
7. Saltini, C. et al. (1998). Immunogenetic basis of environmental lung disease: Lessons from the berylliosis model. *Eur. Respir. J.* **12**: 1463–1475.
8. Reeves, A.L. (1995). Toxicity of detrimental metal ions, Beryllium. In: *Handbook of Metal–Ligand Interactions, Biological Fluids: Bioorganic Medicine*. Vol. II. Dekker, New York, pp. 772–779.
9. Hofele, J., Sietz, M., Grote, M., and Janssen, E. (1994). Untersuchungen zu berylliumgehalten von kartoffeln und boden. *Agribiological Res.* **47**: 3–4.
10. Sato, A. (1981). *Practice On Furnace Atomic Absorption Spectrometry*. Kodansha, Tokyo, Japan.
11. Korkisch, J., Sorio, A., and Steffan, I. (1976). Atomic-absorption determination of beryllium in liquid environmental samples. *Talanta* **28**: 289–294.
12. Taylor, T.P. et al. (2003). Beryllium in the environment: A review. *J. Environ. Sci. Health* **A38**: 439–469.
13. Baes, C.F., Jr. and Mesmer, R.E. (1976). *The Hydrolysis of Cations*. John Wiley & Sons, New York.
14. Hormann, P.K. (1969). Beryllium. In: *Handbook of Geochemistry*. K.H. Edepohl (Ed.). Springer-Verlag, New York.
15. Rossini, F.D., Wagman, D.D., Evans, W.H., Levine, S., and Jaffe, I. (1952). *Selected Values of Chemical Thermodynamic Properties*. National Bureau Standards Circular 500. Washington, DC.
16. Wagman, D.D. et al. (1982). The NBS tables of chemical thermodynamic properties, selected value for inorganic C1 and C2 organic substances in SI units. *J. Phys. Chem. Reference Data.* **2**: 392.
17. Latimer, W.M. (1952). *Oxidation Potentials*, 2nd Edn. Prentice-Hall, New York.
18. Brookins, D.G. (1988). *E_h-pH Diagrams for Geochemistry*. Spinger-Verlag, New York.
19. Krauskopf, K.B. (1979). *Introduction to Geochemistry*, 2nd Edn. McGraw-Hill, New York.
20. Grigor'yev, N.A. (1986). Distribution of beryllium at the surface of the earth. *Int. Geol. Rev.* **28**: 127–179 and **28**: 326–371.
21. Berg, T., Royset, O., Steines, E., and Vadset, M. (1995). Atmospheric trace element deposition: principal component analysis of ICP-MS data from moss samples. *Environ. Pollut.* **88**: 67–77.
22. Asami, T. and Kubota, M. (1995). Background levels of beryllium in several countries. *Environ. Geochem. Health* **17**: 32–38.
23. Asami, T. and Fukazawa, F. (1981). Beryllium pollution of soils surrounding beryllium factories. *J. Sci. Soil Manure* **52**: 1–7.

24. Bhat, P.N. and Pillai, K.C. (1996). Beryllium in environmental air, water, and soil. *Water Air Soil Pollut.* **95**: 133–146.
25. Grote, M., Sietz, M., Hofele, J., and Kowalsik, J. (1997). Investigations on the phytoavailability of beryllium and vanadium. *Fresenius Environ. Bull.* **6**: 037–042.
26. Kabata-Pendia, A. and Pendias, H. (1992). *Trace Elements in Soils and Plants.* 2nd Edn. CRC Press, Boca Raton, FL.
27. Schacklette, H.T. and Boerngen, J.G. (1984). *Element Concentrations in Soils and Other Surficial Materials of the Conterminous United States*. USGS Professional Paper 1270. Washington, DC.
28. Bowen, H.J.M. (1979). *Environmental Chemistry of the Elements*. Academic Press, New York, p. 60.
29. Ure, A.M. and Bacon, J.R. (1978). Comprehensive analysis of soils and rocks by spark-source mass spectrometry. *Analyst* **103**: 807.
30. Conner, J.J. and Shacklette, H.T. (1975). *Background Geochemistry of Some Rocks, Soils, Plants, and Vegetables in the Conterminous United States*. USGS Professional Paper 574-F. Washington, DC, pp. 33–35.
31. Meehan, W.R. and Smythe, L.E. (1967). Occurrence of beryllium as a trace element in environmental materials. *Environ. Sci. Technol.* **1**: 839–844.
32. Gorham, F. and Swaine, D.J. (1965). The influence of oxidizing and reducing conditions upon the distribution of some elements in lake sediments. *Limnol. Oceanogr.* **10**: 268–279.
33. Vinogradov, A.P. (1959). *The Geochemistry of Rare and Dispersed Chemical Elements in Soils.* 2nd Edn. Consultants Bureau, New York.
34. Swaine, D.J. (1955). *The Trace Element Content of Soils*. Commonwealth Agricultural Bureau, Farnham, UK.
35. Anderson, M.A., Bertsch, P.M., and Miller, W.P. (1990). Beryllium in selected southeastern soils. *J. Environ. Qual.* **19**: 347–348.
36. Fishbein, L. (1984). Overview of analysis of carcinogenic and/or mutagenic metals in biological and environmental samples, I. Arsenic, beryllium, cadmium, chromium, and selenium. *Int. J. Environ. Anal. Chem.* **50**: 874–880.
37. USEPA. (1998). *Study of Hazardous Air Pollutant Emissions from Electric Utility Steam Generating Units—Final Report to Congress*. U.S. Environmental Protection Agency, Research Triangle Park, NC.
38. Goldberg, E.D., Broecker, W.S., Gross, M.G., and Turekian, K.K. (1971). *Radioactivity in the Marine Environment*. National Academy of Science, Washington, DC, pp. 137–146.
39. Measures, C.I. and Edmond, J.M. (1986). Determination of beryllium in natural waters in real time using electron capture detection gas chromatography. *Anal. Chem.* **58**: 2065–2069.
40. Merrill, J.R., Lyden, E.F.X., Honda, M., and Arnold, J.R. (1960). The sedimentary geochemistry of beryllium isotopes. *Geochim. Cosmochim Acta.* **18**: 108–129.
41. Duram, W.H. and Haffty, J. (1961). *Occurrence of Minor Elements in Water*. USGS Circular 445. Washington, DC.
42. SDWC. (1977). *Drinking Water and Health*. Safe Drinking Water Committee of the National Academy of Sciences, Washington, DC.
43. Skirivan, P., Stastny, M., Kotkova, P., and Burain, M. (1996). Partition of beryllium and several other trace elements in surface waters. *Scientia Agriculturae Bohemica* **27**: 131–145.
44. Kopp, J.F. (1969). *The occurrence of trace elements in water. Proc. 3rd Annu. Conf. Trace Substances Environ. Health.* University of Missouri, Columbia, MD, pp. 59–73.
45. Griffits, W.R., Allaway, W.H., and Groh, D.H. (1977). The relation of other substances and trace elements to health and disease. In: *Beryllium In Geochemistry and the Environment*. Vol. II. National Academy of Science, Washington, DC.
46. Veseley, J., Benes, P., and Sevcik, K. (1989). Occurrence and speciation of beryllium in acidified freshwaters. *Water Res.* **23**: 711–717.
47. APHA. (1976). *Standard Methods for the Examination of Water and Waste Water*, 13th Edn. American Public Health Association, Washington, DC.
48. Hem, J.D. (1989). *Study and Interpretation of the Chemical Characteristics of Natural Water*, 3rd Edn. USGS Water-Supply Paper 2254. Washington, DC.
49. Burba, F., Willmer, P.G., Betz, M., and Fuchs, F. (1983). Atomic absorption determination of beryllium traces in natural waters. *Int. J. Environ. Anal. Chem.* **13**: 177–191.
50. Borovec, Z., Tolar, V., and Mraz, L. (1994). The state of accumulation and source relations of Co, Ni, Cr, Be, and Zn in silt-clay sediments of the Labe river. *Geology Carp. Ser. Clays.* **45**: 3–9.
51. Borovec, Z. (1993). Partitioning of silver, beryllium, and molybdenum among chemical fractions in the sediments from the lake (Elbe) river in central Bohemia, Czech Republic. *GeoJournal* **29**: 359–364.
52. Benes, P. and Jiranek, V. (1974). Adsorption and colloidal behavior of carrier-free ^{7}Be in aqueous solutions. *Radiochimica Acta* **21**: 49–53.
53. Moore, G.K. (1995). *Inorganic Soil and Groundwater Chemistry Near Paducah Gaseous Diffusion Plant, Paducah, Kentucky*. Oak Ridge National Laboratory, Oak Ridge, TN.

READING LIST

Bashir, W. and Paull, B. (2001). Sensitive and selective ion chromatographic method for the determination of trace beryllium in water samples. *J. Chromatogr. Anal.* **910**: 301–309.

Burguera, J.L. et al. (2000). Determination of beryllium in natural and waste waters using on-line flow-injection preconcentration by precipitation/dissolution for electrothermal atomic absorption spectrometry. *Talanta* **52**: 27–37.

Capitan, F. et al. (1989). Determination of beryllium in water by ion-exchange spectrofluorimetry. *Analyst* **114**: 969–973.

Hoshino, H., Monura, T., Nakano, K., and Yotsuyanagi, T. (1996). Complexation of Be(II) ion with 1-(2,4-dihydroxy-1-phenylazo)-8-hydroxy-3,6-napthalenedisulfonate as the chemical basis of the selective detection of ultratrace Be by reversed-phase high-performance liquid chromatography. *Anal. Chem.* **68**: 1960–1965.

Kubova, J., Nevoral, V., and Sresko, V. (1994). Determination of beryllium trace contents in mineral waters after preconcentration on a chelating ion-exchanger. *Fresenius J. Anal. Chem.* **348**: 287–290.

Majee, S. and Das, J. (1998). Use of chelating ion exchanger containing acetoacetanilide group selective separation of beryllium. *Indian J. Chem.* **27A**: 983–986.

Narayana, R. and Khopkar, S.M. (1985). Separation of beryllium(II) from other metallic ions with aliquot 336S by extraction chromatography of its malonato complex. *Analyst* **110**: 1295–1298.

Rose, A.W., Hawkes, H.E., and Webb, J.S. (1979). *Geochemistry in Mineral Exploration*. Academic Press, London.

Singh, H.B., Agnihotri, N.K., and Singh, V.K. (1998). Determination of trace amounts of beryllium using derivative spectrophotometry in non-ionic micellar medium. *Talanta* **47**: 1287–1296.

Voloschick, I.N., Litvina, M.L., and Rudenko, B.A. (1995). Ion chromatographic determination of beryllium in rock and waste waters with a chelating sorbent and conductimetric detection. *J. Chromatogr. Anal.* **706**: 315–319.

DISSOLVED ORGANIC CARBON

Jamie R. Lead
University of Birmingham
Birmingham, United Kingdom

DISSOLVED ORGANIC CARBON

Of the three words in the title of this entry, little explanation is needed for the last two. The first, however, needs some clarification. The term has traditionally been used to denote material that has been filtered through a certain type of standard membrane (0.45 μm nominal pore size). Material passing through the filter (filtrate) was termed the dissolved phase, and material retained by the filter (retentate) was termed the solid phase. This filtration step was performed for a number of reasons such as reduction in the complexity of the sample and partial sterilization of the filtrate due to the removal of microorganisms. The procedure is still widely used in the scientific and regulatory community. However, it has long been recognized as an inadequate definition because of the presence of colloidal carbon (1–1000 nm in size) and the artifacts inherent in filtration (1,2). The use of a single filtration step thus leads to results that are neither thermodynamically accurate nor biogeochemically relevant. There have been a number of theoretical and practical attempts to delimit the colloidal and dissolved phases. Now, the main practical delimiter is based on ultrafiltration using a nominal pore size of 1,000–10,000 daltons (ca 1 nm) (5). This is a more useful definition for understanding biogeochemical processes but remains entirely operational. In this article, the term 'dissolved' will be used in the wider sense to incorporate a fraction of the colloidal material.

The importance of dissolved organic carbon (DOC) lies in its role in the biogeochemical carbon cycle, as a necessary basis for aquatic microbial processes and in aquatic chemistry (6). Although the global burden of carbon is dominated by sediments (7), dissolved organic carbon is of obvious importance as carbon is the basis of life, an energy and nutrient source for microorganisms. DOC is also of significance in trace pollutant binding (and consequently in pollutant transport and bioavailability), in the charge and aggregation behavior of colloids, in weathering, in pH buffering, and in acid–base and redox chemistry of natural waters. It thus has a large and essential role in determining the chemical composition of natural waters. Although microbiological and chemical aspects are frequently taught and researched in isolation, it may be impossible to understand the composition and regulation of natural waters without simultaneously considering the microbiological and chemical processes because they are so closely interlinked (8). Any distinction between the two processes is therefore likely to be arbitrary and misleading.

The chemical structure of much of dissolved organic carbon is poorly understood because of its complexity, its variability over both space and time, and the ease with which its structure is changed by sampling and analytical procedures (1,9). As DOC is physicochemically diverse, a number of partially successful subclassifications have been used based on size, chemistry, and source. However, for most DOC, it is impossible to define either an exclusive structural or functional definition or classification system, as has been done for biochemical polymers. Here, we will discuss two types of DOC, humic substances and biologically produced chemicals. In addition, anthropogenic pollutants may also be important, for instance, in human and ecological health, although not in terms of total environmental loads of DOC.

Humic substances (HS) are often the dominant type of DOC, particularly in 'organic-rich' waters such as marshes and in dystrophic waters where microbial activity is limited. HS are a general class of complex, biogenic, macromolecular species. They are heterogeneous and polydisperse and do not have one single structure but exhibit a range of characteristic properties such as molar mass and charging behavior. It has been suggested that heterogeneity is their essential, defining characteristic (10). They are the products of biological decomposition. However, they differ from the biologically produced chemicals of the next section because they have undergone a substantial degree of chemical and biological degradation (humification), resulting in a kinetically stable product resistant to further degradation. They are defined operationally by the extraction system used to separate them from other components such as clays, oxides, and nonhumic organic material. It is important to realize that they are not defined either by structure or function because their structure is too complex and they have no clearly defined 'function' in the biochemical sense. HS are usually divided into three main groups based on their relative solubilities. It is hoped that these reflect to some extent the solubility of HS in the environment and properties such as solid:solution partitioning and transport. In any case, the division results in a simplification of the HS which has analytical advantages. The three fractions are

fulvic acid—soluble at all pHs.
humic acid—soluble at high pH, precipitated in acid (pH 2).
humin—insoluble in acid and base.

The extraction and fractionation scheme have been standardized by the International Humic Substances Society (IHSS). In general terms and considering their heterogeneity, fulvic acids and aquatic HS are more soluble, have a greater density of functional groups, carry a greater charge, and have a lower C:O ratio than humic acid or terrestrial HS, respectively. HS are important in a number of processes and have been much studied. Information has been collected on properties such as their acid–base, metal binding and redox chemistry, photochemistry, structure, functional group content, molar mass, and role as a nutrient. However, a fundamental, unanswered question relates to the extraction and fractionation scheme: To what extent are their properties significantly modified by the extraction system and therefore how are extracted HS different from 'natural' HS?

Biologically produced chemicals consist of a number of well-defined chemicals produced directly from microbial activity. These can further divided into (1) simple monomers and (2) more complex polymers. The monomers include amino acids and monosaccharides that are relatively simple compounds whose structure can be ascertained precisely. In the main, they are produced and consumed rapidly and thus have a short residence time (a few days), low concentration, and are of relatively minor importance. Their concentrations depend on the relative rate of supply and the rate of demand, both controlled primarily by microbial activity. The compounds may be 'free' or bound to humic substances. As for HS, their concentration and fraction of total DOC are variable, depending on the water type and season. More complex, biologically produced polymers, sometimes denoted as biopolymers, are less well studied than HS. However, they can have concentrations similar to HS and can make up 30–40% of total DOC, particularly in eutrophic waters. They consist primarily of proteins and polysaccharides from sources such as plant and microbial cell walls, exudates secreted by microbes, and internal cellular proteins and polysaccharides from microbes after cell lysis. It is possible that HS and biopolymers have differing roles in aquatic systems. For instance, it is well known that HS tends to stabilize colloids and particles and reduce aggregation. However, it is possible that the biopolymers tend to increase aggregation by bridging mechanisms (11) that have important consequences for the chemical and biological composition of natural waters.

In summary, DOC, including some colloidal OC, is clearly extremely important in biological and chemical activity and interactions in the environment. However, despite the progress made and the vast amount of research performed, many fundamental questions await answers.

BIBLIOGRAPHY

1. Lead, J.R., Davison, W., Hamilton-Taylor, J., and Buffle, J. (1997). **3**: 219–232.
2. Horowitz, A.J. et al. (1996). *Environ. Sci. Technol.* **30**: 954–963.
3. Gustafsson, O. and Gschwend, P.M. (1997). *Limnology and Oceanography* **42**: 519–528.
4. Stumm, W. (1993). *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **73**: 1–19.
5. Martin, J-M., Dai, M-H., and Cauwet, G. (1995). *Limnology and Oceanography* **40**: 119–131.
6. Thurman, E.M. (1985). *Organic Geochemistry in Natural Waters*. Martinus Nijhoff/Dr W. Junk, The Netherlands.
7. Butcher, S.S., Charlson, R.J., Orians, G.H., and Wolfe, G.V. (Eds.). (1992). *Global Biogeochemical Cycles*. Academic Press, London.
8. Buffle, J. and de Vitre, R.R. (1994). *Chemical and Biological Regulation in Aquatic Systems*. Lewis Publishers, New York.
9. Wilkinson, K.J., Balnois, E., Leppard, G.G., and Buffle, J. (1999). *Colloids and Surfaces A: Physicochemical and Engineering Aspects*. Vol. 155, pp. 287–310.
10. MacCarthy and Suffet. (1989). *Chemical Modeling in Aqueous Species*. American Chemical Society Symposium.
11. Buffle, J. et al. (1998). *Environmental Science and Technology* **32**: 2887–2899.

MECHANISMS OF WATER ADSORPTION ON CARBONS

Piotr A. Gauden
Artur P. Terzyk
Gerhard Rychlicki
Sylwester Furmaniak
Radosław P. Wesołowski
Marek Wiśniewski
N. Copernicus University
Toruń, Poland

Piotr Kowalczyk
Faculty of Science, Chiba University
Chiba, Japan

The properties of water are of importance in many scientific disciplines such as chemistry, biology, geology, nanotechnology, and materials technology. Moreover, the adsorption of water on activated carbon (AC) is an important topic in many different areas of science and technology because water is the most common solvent in nature. The adsorption and phase behavior of polar fluids in carbon pores has been studied extensively, but our understanding of the adsorption of water on carbonaceous materials is still incomplete (1–4).

In recent years, a number of experimental and simulation studies of adsorption of water in pores have appeared in the literature. Some studies have assumed that the adsorption behavior of water in graphite pores is hydrophobic. Although principally hydrophobic adsorbents may contain significant numbers of adsorption centers that can interact with water, it is generally believed that the combination of weak carbon–water dispersive attractions and strong water–water associative interactions are responsible for the complex behavior of water confined in carbonaceous pores.

On the other hand, surface heterogeneity can be introduced by incorporating some polar groups on the surface and those active sites lead to enhanced water adsorption. The loss of water–water H-bonds in graphitic pores is compensated for by the ability of water molecules to form H-bonds with the surface sites in activated carbon pores; the result is a change in the molecular level adsorption phenomenon from capillary condensation in graphite pores to continuous pore filling in activated carbon pores. This leads to an important practical conclusion that water adsorption measurements can be used to determine the number of surface-active centers.

The next important problem is the structure of water adsorbed in pores. Some experimental X-ray diffraction (XRD) (5), differential scanning calorimetry (DSC) (6), dielectric relaxation spectroscopy (DRS) (6), and theoretical studies (6–8) showed the presence of a well-ordered structure of water adsorbed in the graphite micropores at ambient temperature. The adsorbed water has a more ordered structure than bulk liquid, but less than ice, and the structure of water molecules depends on the width of pores (5,8).

One of the most spectacular methods is calorimetry (2,9–22). Careful measurements of the energetic effects accompanying water adsorption on carbon blacks and carbons were studied by scientists from the so-called "Russian school of adsorption" in the 1950s (15,16) and the 1980s (17,18) and by others (9–11,13,14,19–22). Heats of adsorption of water indicate that adsorption is a strong function of surface chemistry. The primary high-energy sites have a significant influence on adsorption and the enthalpy of adsorption at low relative pressures. Three components can be delineated according to measured differential heats of adsorption: (1) chemical adsorption, (2) physical adsorption, and (3) condensation and/or micropore filling. From the point of view of the thermodynamic verification of theoretical models, the first is the most interesting and important. Chemisorption of water on carbons has been previously reported (2,9–11,17,20–22). Summing up, analysis of the experimental data points out that significant discrepancies exist between the role of specific surface groups (e.g., various types, location, density) and their energy of interaction with water molecules and a quantitative description of the isotherm over the entire relative pressure range that is universal for all carbons with low, medium, and high densities of surface groups.

Recently, knowledge of the water–carbon intermolecular interaction and adsorption of water in the microporous region of ACs has been greatly increased by quantum methods (QM) and molecular simulations. A natural starting point for computer simulations is the development of a model potential for the interaction between water and benzene (23,24). Although model potentials already exist for the water/benzene and water/graphite systems (3,23,25), the reliability of these models has not been widely tested. The chemistry and reaction mechanisms of adsorption, as well as of intercalation, are less known, and ab initio QM methods provide a way to examine these phenomena on the atomic level.

The existing water/graphite model potentials give binding energies for a single water monomer interacting with the surface ranging from −6.91 to −18.00 kJ/mol (25). There also exists an experimental value of −15.07 kJ/mol for the binding energy (26). However, these values seem somewhat low in the light of the most recent high-level ab initio results for the water–benzene system, for which the electronic binding energy has been estimated at −19.32 kJ/mol (27). A large basis set at MP2 level (second-order Mřller–Plesset perturbation theory) calculations on a series of fragment models up to $C_{96}H_{24}$ in size was used by Feller and Jordan (28) to estimate the electronic binding energy of an individual water molecule to single-layer graphite (−24.28 ± 1.67 kJ/mol). The interaction between water molecules and graphite structures was clearly stronger with the $C_{96}H_{24}$ model than for the smaller ones. The opposite result was observed for a smaller basis and the Hartree–Fock method (28). These calculations highlight the difficulty of approaching the complete basis set limit in extended systems, such as graphite, using traditional atom-centered basis sets. Feller and Jordan shown (29) that the estimated water–graphite binding energy is slightly larger than that for the interaction between two water molecules (−20.92 ± 0.42 kcal/mol) (30). At the same time, this interaction is significantly smaller than the average electronic binding energy of water monomers in bulk water, where the hydrogen bonding network involves each water molecule in (up to) four hydrogen bonds (30). Nonetheless, the estimated water–graphite interaction is sufficiently large to lead to the conclusion that in carbon, the interaction of water with the graphite walls plays a significant role in determining the structure of the confined water. On the other hand, in the water–benzene complex, the important attractive interactions are the dipole-quadrupole, dipole-induced dipole (induction), and dispersion contributions (23,24,27).

Ruuska and Pakkanen (28) investigated interactions between water molecules and one- and two-layer graphite using the ab initio Hartree–Fock (HF) method. Three one-layer models were used as a reference to investigate the influence of the adsorption site (on top, bridge, and hollow), the orientation of the water molecule, and the size of the model. Only minor differences were found in the three adsorption sites. The behavior of water molecules between two graphite planes was examined using the two-layer model. The coronene planes were moved from a distance of 9 to 3 Å, and the water molecule was placed in the middle of the layers. It is clear from results published by Ruuska and Pakkanen (28) that a water molecule requires relatively much space because it doubles the interlayer distance. Additionally, the features of the two-layer graphite with a water molecule between are similar to those of graphite intercalates.

A more complex problem considered by QM is the interaction of water with a graphite-like surface containing oxygenated hydrophilic groups. It is well known that adsorption is strongly dependent on the presence and arrangement of hydrophilic groups. Tarasevich and Aksenenko (31) used the semiempirical PM3 method to study the interaction of very few water molecules with a partially oxidized graphite surface. The main conclusion of this work is that the water molecules are adsorbed at the hydrophilic center (i.e., the carboxyl) with the formation of microclusters consisting (at initial coverages) of two to three molecules. In the light of these results, Tarasevich and Aksenenko (31) suggested that water adsorption measurements cannot be used to determine the number of active centers even at hydrophobic, and, far less, at hydrophilic surfaces, and methanol was recommended for this purpose.

Hamad and co-workers (32) reported the quantum chemical characterization of a hydrophilic site modeled by a carboxyl group attached to one of the carbon atoms in the graphite layer. They supply values of point charges at the

–COOH sites that have been used to derive a pair potential for the water-active site interaction. Picaud et al. (33) presented molecular dynamics simulations based on a classical potential issued from QM calculations (30) to provide a view of the dynamics and the structure of a water layer on a model surface that consists of a planar graphitic layer plus different numbers of active carboxylic sites. The results of these simulations show that the –COOH groups act as strong trapping sites for a few water molecules that then become nucleation centers for other water molecules that form larger aggregates tied to the –COOH.

The strong effect of PACs on water adsorption was indicated on the basis of the computer simulations (3,8,33–45). The adsorption of water molecules on a hydrophobic surface has been investigated (8,34,35,38–41,44). On the other hand, several types of polar oxygen-containing sites [e.g., carboxyl (42) carbonyl (34,42,43), and hydroxyl (37,42) groups or H-sites (40)] were placed on the surface of carbon with different densities and local distributions to determine the influence of each of these factors on the adsorption of water. Computer simulation models are still being developed and more and more complicated systems have been analyzed. To provide a realistic pore representation and a description of water adsorption, recently, Thompson and Gubbins (3,43,45) suggested the application of a model based on reverse Monte Carlo. In this model, surface sites have been added at random points on the edges of graphene microcrystals possessing random sizes and structures. The major drawback to advanced numerical calculations is the computing power and time needed to simulate isotherms. Therefore, uncomplicated analytic approaches are often favored to simplify consideration and cut down significantly the computation time, for instance, so-called local isotherms obtained from computer simulation are described by a theoretical model [for example, proposed by Talu and Meunier (46)] (37).

Muller and Gubbins (36) showed, that for nonactivated carbon slit pores almost no adsorption occurs until a sharp vertical rise in the adsorption curve is seen and the pore fills suddenly (i.e., capillary condensation occurs). If the surface is doped by active sites, the adsorption isotherm changes drastically, strong bonds are formed between these sites and water molecules, and these adsorbed molecules become nucleation sites for other water molecules to adhere. Thus, these results confirm the role of primary surface sites in the enhancement of water adsorption. Furthermore, Jorge et al. (42) showed that the local distribution of primary sites (carboxyl, hydroxyl, and carbonyl) has a strong effect on the low-pressure part of the adsorption isotherm, whereas the overall site density affects mainly the vapor–liquid phase transition. The type of oxygen-containing group was not of critical importance because more complex groups can effectively be represented by simpler sites. On the other hand, McCallum et al. (37) proposed two alternative mechanisms for water adsorption onto activated carbon walls due to surface sites regularly arranged or a random array (the same site density in both cases was assumed). Five stages are observed from the analysis of both types of simulated adsorption isotherms. Moreover, their behavior is similar to experimental ones.

The theory describing the mechanism of water adsorption on carbonaceous materials proposed by Dubinin and Serpinsky (DS) (1,3,16,47) is probably the most famous. Although the fundamentals of this model were developed 50 years ago, it is still widely used mainly due to its simplicity. Dubinin et al. (16) introduced a phenomenological model of adsorption of water molecules on energy privileged sites which act as primary adsorption centers for water on a predominantly hydrophobic carbon surface. On the other hand, water molecules adsorbed on these sites produce new centers (called elsewhere secondary ones) for further adsorption and cluster formation. Note that the hyperbolic behavior of this equation leads to an infinite adsorbed water concentration at higher pressures. It is well known that the DS model describes well only the data measured on nonporous surfaces and on strongly hydrophobic carbons. The next main limitation is the assumption of unlimited adsorption space. Therefore, the original DS equation has been improved by Dubinin et al. (47), Barton et al. (48–50), and others (10,11,51–55) by taking into consideration new functions describing a decrease in the number of secondary adsorption sites.

On the other hand, it is well known that the simplest method leading to an increase in the affinity of a carbon surface for water vapor is, for example, oxidation of this surface. In such a case, a more hydrophilic surface chemical composition leads to a rise in adsorption, especially in the initial range of the water isotherm (an increase in the enthalpy of adsorption and enthalpy of immersion in water are observed simultaneously (2,9–11,15,17,19–21,56)). For hydrophilic surfaces, DS-like equations (1,3,10,11,51–55) are insufficient and, for example, the D'Arcy and Watt (DW) adsorption isotherm can be applied (57) (however, some attempts to improve the DS model to correct the description of this type of experimental data were recently proposed by Gauden and Terzyk (52,53)). In the original DW model, it is assumed that sorption of water takes place independently on strongly (high-energy) binding primary sites (which are hydrophilic groups) and to a less extent on weakly (low-energy) binding primary sites (located elsewhere on the carbon surface). Moreover, the term describing multilayer adsorption on the so-called secondary centers (which need not be completely occupied before monolayer formation) is assumed in the D'Arcy and Watt isotherm. As shown by Evans et al. (58,59) and others (1,10,11,60,61), the DW equation satisfactorily describes water adsorption data on different adsorbents (not only on carbons). Note that in the opinion of Barton and co-workers (59), as well as in the opinion of the authors of the current study (54,55), it appears that D'Arcy and Watt's isotherm contains some inconsistencies. A serious flaw in this simple treatment is the assumption that Langmuir and the original DS isotherms describe the adsorption of water molecules on the adsorbent independently (i.e., they are connected with various centers). Moreover, the further disadvantage of this model is an unsatisfactory description of experimental adsorption data at higher pressures and enthalpy of adsorption data (10,11,51,54,55) due to the assumption of one kind of surface center on the carbon surface. Therefore, we recently generalized the original D'Arcy and Watt equation taking into account the influence of the concentration of the surface primary sites on the amount of the secondary ones (54,55). Next, we assumed that the adsorbent has regions of heterogeneity on the molecular level (i.e., the various functional groups and/or defects on the surface of the adsorbent). Moreover, water molecules can adsorb on all types of high-energy primary sites obeying the Langmuir isotherm. Summing up, the multisite generalized D'Arcy and Watt (MSGDW) model was proposed (54,55). It can be noticed that finite adsorption space is taken into account by the term $(1 - ka^n)$. From the analysis of the isotherm generated as well as the enthalpy plots, it is seen that these curves are similar to those measured using adsorption calorimetry (10,11,18,20,54,55,63,65), especially at low coverages.

Note that the theories proposed by Dubinin, Serpinsky and D'Arcy and Watt do not give information on the dimensions of adsorbed water clusters (however, it is obvious that their sizes are confined to the available adsorption space). On the other hand, it is well known that water molecules in the bulk state form dimers (64); however, the number of dimerized molecules is not significant at room temperature and standard pressure. In a hydrophobic micropore, water molecules are expected to be more dimerized due to the compressed condition there. In narrow pores, the dimers grow to a greater cluster, which has been used to model the elemental structure of water (1,8,44,65). On the other hand, the pore width is limited, and so the size of the clusters is restricted by the finite pore dimension. Iiyama et al. (66) stated that water forms clusters containing at least five water molecules per cluster.

Different authors, for example, Do and Do (67) and Neitsch et al. (68), applied the above mentioned fundamental information about the size of water clusters to develop new water adsorption models. As a result, the model proposed by Do and Do (67) is based on the growth of a water cluster around surface functional groups, the creation of a pentamer, and penetration of these clusters into the micropore (they remain in the fine pores if they have sufficiently high dispersion energy). Note that the Do and Do model can be considered a so called "hybrid" (similarly to the DW and MSGDW equations) in which the first term describes adsorption on primary sites (it is a BET-type) and the second term refers to desorption from the secondary sites and the adsorption of the pentamer into the micropores (so the equation describes a micropore filling mechanism). Neitsch et al. (68) modified this model. They assumed that the size of the water clusters (pentameter in the original Do and Do model) is the best fitting parameter.

Another well-known model describing water adsorption was proposed by Talu and Meunier (46). Their theoretical isotherm equation assumed the following mechanism of adsorption of water molecules: at low pressures, the behavior of the system is controlled by vertical interactions of a molecule to the surface and the molecules are adsorbed on active sites. At intermediate coverages where the inflection point occurs, the adsorbed molecules form clusters via hydrogen bonding. At high surface coverages where the plateau is observed, the adsorption is limited by micropore filling.

Malakhov and Volkov (69) proposed a cooperative multimolecular adsorption theory describing adsorption of alcohols on glassy polymers. This model was successfully adopted by Rutherford to water adsorption on activated carbons (70). In this theory, the movement of adsorbate molecules to a single adsorption site is taken into account. The cooperative influence is induced by the adsorbed molecule (primary stage) which promotes the entry of other molecules to adjacent centers. In the secondary stage, the model considers a unit triad where a central adsorption site promotes adsorption on either side of the central site. Additionally, secondary centers are defined as side associates forming from the primary surface sites and hence allowing the formation of dimers, trimers, and more complicated clusters. Summing up, the cooperative multimolecular adsorption theory is based on the assuming the existence of one kind of primary and three kinds of secondary centers. However, this theory is not valid to describe adsorption at high relative pressures (70).

Several predictive models have been proposed (in our opinion the most important are mentioned above). Unfortunately, none of them satisfactorily describe the isotherms and the corresponding differential heat of adsorption over the whole pressure range. Moreover, some fundamental aspects are not taken into account in the theoretical considerations: the role of specific surface groups (e.g., various type, location, density), their interaction with water molecules, and the structure of water confined in pores.

BIBLIOGRAPHY

1. Mowla, D., Do, D.D., and Kaneko, K. (2003). *Chem. Phys. Carbon* **28**: 229.
2. Vartapetyan, R.Sh. and Voloshchuk, A.M. (1995). *Usp. Khim.* **64**: 1055. (in Russian).
3. Brennan, J.K., Bandosz, T.J., Thomson, K.T., and Gubbins, K.E. (2001). *Colloids Surf. A* **187–188**: 539.
4. Terzyk, A.P. (2003). *J. Colloid Interf. Sci.* **268**: 301.
5. Iiyama, T., Nishikawa, K., Suzuki, T., and Kaneko, K. (1997). *Chem. Phys. Lett.* **274**: 152.
6. Sliwinska-Bartkowiak, M. et al. (2001). *Phys. Chem. Chem. Phys.* **3**: 1179.
7. Sanfelix, P.C., Holloway, S., Kolasinki, K.W., and Darling, G.R. (2003). *Surf. Sci.* **532**: 166.
8. Ohba, T., Kanoh, H., and Kaneko, K. (2004). *J. Am. Chem. Soc.* **126**: 1560.
9. Tsutsumi, K. and Mizoe, K. (1989). *Colloids Surf. A* **37**: 29.
10. Terzyk, A.P., Gauden, P.A., and Rychlicki, G. (1999). *Colloid Surf. A* **148**: 271.
11. Terzyk, A.P., Rychlicki, G., and Gauden, P.A. (2001). *Colloid Surf. A* **179**: 39.
12. Vartapetyan, R.Sh., Voloshchuk, A.M., Isirikyan, A.A., Olyakov, N.S., and Tarasevich, Yu.I. (1995). *Colloids Surf. A* **101**: 227.
13. Stoeckli, F., Jakubov, T., and Lavanchy, A. (1994). *J. Chem. Soc., Faraday Trans.* **90**: 783.

14. Carrasco-Marin, F., Meuden, A., Ceneteno, T.A., Stoeckli, F., and Moreno-Castilla, C. (1997). *J. Chem. Soc. Faraday Trans.* **93**: 2211.
15. Avgul, N.N., Dzigit, O.M., Kiselev, A.V., and Sherbakov, K.D. (1953). *Izv. AN SSSR* **2**: 105. (in Russian).
16. Dubinin, M.M., Zaverina, E.D., and Serpinsky, V.V. (1955). *Trans. Faraday Soc.* 1760. (in Russian).
17. Dubinin, M.M. and Isrikikyan, A.A. (1989). *Izv. AN SSSR* **10**: 2183. (in Russian).
18. Vartapetyan, R.Sh., Voloshchuk, A.M., Dubinin, M.M., and Yakubov, T.S. (1987). *Izv. AN SSSR* **9**: 1934. (in Russian).
19. Miura, K. and Morimoto, T. (1988). *Langmuir* **4**: 1283.
20. Rychlicki, G. (1989). *Polish J. Chem.* **63**: 225.
21. Philips, J., Kelly, D., Radovic, L., and Xie, F. (2000). *J. Phys. Chem. B* **104**: 8170.
22. Groszek, A.J. (2001). *Carbon* **39**: 1857.
23. Dang, L.X. and Feller, D. (2000). *J. Phys. Chem. B* **104**: 4403.
24. Tarakeshwar, P. et al. (1999). *J. Chem. Phys.* **111**: 5838.
25. Vernov, A. and Steele, W.A. (1992). *Langmuir* **8**: 155.
26. Avgul, N.N. and Kiselev, A.V. (1970). *Chem. Phys. Carbon* **6**: 1.
27. Feller, D. (1999). *J. Phys. Chem.* **103**: 7558.
28. Ruuska, H. and Pakkanen, T.A. (2003). *Carbon* **41**: 699.
29. Feller, D. and Jordan, K.D. (2000). *J. Phys. Chem. A* **104**: 9971.
30. Feyereisen, M.W., Feller, D., and Dixon, D.A. (1996). *J. Phys. Chem.* **100**: 2993.
31. Tarasevich, Y.I. and Aksenenko, E.V. (2003). *Colloids Surf. A* **215**: 285.
32. Hamad, S., Mejias, J.A., Lago, S., Picaud, S., and Hoang, P.N.M. (2004). *J. Phys. Chem. B* **108**: 5405.
33. Picaud, S., Hoang, P.N.M., Hamad, S., Mejias, J.A., and Lago, S. (2004). *J. Phys. Chem. B* **108**: 5410.
34. Malenkov, G.G. and Dubinin, M.M. (1984). *Izv. AN SSSR* **6**: 1217. (in Russian).
35. Slasli, A.M., Jorge, M., Stoeckli, F., and Seaton, N.A. (2003). *Carbon* **41**: 479.
36. Muller, E.A. and Gubbins, K.E. (1998). *Carbon* **36**: 1433.
37. McCallum, C.L., Bandosz, T.J., McGrother, S.C., Muller, E.A., and Gubbins, K.E. (1999). *Langmuir* **15**: 533.
38. Striolo, A., Chialvo, A.A., Cummings, P.T., and Gubbins, K.E. (2003). *Langmuir* **19**: 8583.
39. Brodskaya, E.N., Zakharov, V.V., and Laaksonen, A. (2001). *Langmuir* **17**: 4443.
40. Galle, J. and Vortler, H.L. (2001). *Surface Sci.* **481**: 39–53.
41. Spohr, E., Trokhymchuk, A., and Henderson, D. (1998). *J. Electroanal. Chem.* **450**: 281.
42. Jorge, M., Schumacher, C., and Seaton, N.A. (2002). *Langmuir* **18**: 9296.
43. Brennan, J.K., Thomson, K.T., and Gubbins, K.E. (2002). *Langmuir* **18**: 5438.
44. Kimura, T. et al. (2004). *J. Phys. Chem. B* **108**: 14043.
45. Thompson, K.T. and Gubbins, K.E. (2000). *Langmuir* **16**: 5761.
46. Talu, O. and Meunier, F. (1996). *AIChE J.* **42**: 809.
47. Dubinin, M.M. and Serpinsky, V.V. (1981). *Carbon* **19**: 402.
48. Barton, S.S., Evans, M.J.B., and MacDonald, J.A.F. (1991). *Carbon* **29**: 1099.
49. Barton, S.S., Evans, M.J.B., and MacDonald, J.A.F. (1992). *Carbon* **30**: 123.
50. Barton, S.S., Evans, M.J.B., Liang, S., and MacDonald, J.A.F. (1996). *Carbon* **34**: 975.
51. Wesołowski, R.P., Gauden, P.A., Terzyk, A.P., and Furmaniak, S. (2004). *Carbon* **53**: 42.
52. Terzyk, A.P. and Gauden, P.A. (2002). *4th Int. Symp. Coal Struc.—Struc. Reactivity Carbonaceous Mater.*, Gliwice, Poland.
53. Gauden, P.A. *J. Colloid Interface Sci.*, in press.
54. Gauden, P.A., Furmaniak, S., Terzyk, A.P., Rychlicki, G., and Wesołowski, R. (2004). *ISSHAC—5 Proc.* Gdańsk, p. 138.
55. Furmaniak, S., Gauden, P.A., Terzyk, A.P., Wesołowski, R.P., and Rychlicki, G. *Carbon* submitted.
56. Barton, S.S. and Koresh, J.E. (1983). *J. Chem. Soc., Faraday Trans.* **79**: 1147.
57. D'Arcy, R.L. and Watt, I.C. (1970). *Trans. Faraday Soc.* **66**: 1236.
58. Evans, M.J.B. (1987). *Carbon* **25**: 81.
59. Barton, S.S., Evans, M.J.B., and MacDonald, J.A.F. (1994). *Langmuir* **10**: 4250.
60. Venkataramani, B. (1993). *Langmuir* **9**: 3026.
61. Sychev, M. et al. (2001). *Microporous Mesoporous Mater.* **47**: 311.
62. Vartapetyan, R.Sh., Voloshchuk, A.M., Dubinin, M.M., Chekulina, L.A., and Chepurnoi, S.G. (1986). *Izv. AN SSSR* **11**: 2407. (in Russian).
63. Dubinin, M.M. (1981). *Izv. AN SSSR* **1**: 9. (in Russian).
64. Curtis, L.A., Frurip, D.J., and Blender, M. (1979). *J. Chem. Phys.* **71**: 2703.
65. Kaneko, K., Hanzawa, Y., Iiyama, T., Kanda, T., and Suzuki, T. (1999). *Adsorption* **5**: 7.
66. Iiyama, T., Ruike, M., and Kaneko, K. (2000). *Chem. Phys. Lett.* **331**: 359.
67. Do, D.D. and Do, H.D. (2000). *Carbon* **38**: 767.
68. Neitsch, M., Heschel, W., and Suckow, M. (2001). *Carbon* **39**: 1421.
69. Malakhov, A.O. and Volkov, V.V. (2000). *Polymer Sci. Ser. A* **42**: 1120.
70. Rutherford, S.W. (2003). *Carbon* **41**: 622.

THE EFFECT OF CARBON SURFACE CHEMICAL COMPOSITION ON THE MECHANISM OF PHENOL ADSORPTION FROM AQUEOUS SOLUTIONS

Artur P. Terzyk
Gerhard Rychlicki
Piotr A. Gauden
N. Copernicus University
Toruń, Poland

Piotr Kowalczyk
Faculty of Science,
Chiba University
Chiba, Japan

Experimental studies of solute adsorption from dilute solutions, especially from aqueous dilute solutions, are carried out in many laboratories, yet the theoretical foundations of this phenomenon are still insufficiently solved (1). Adsorption of phenols on carbons is an important subject especially due to ecological aspects. Therefore, an enormous number of papers has been published in this field. The most important reviews dealing with this problem were recently published by Radovic and co-workers (2) and by Moreno - Castilla (3). The authors consider the changes in the surface chemistry of carbons with changes in pH, and discuss different mechanisms of adsorption proposed in the literature.

The main regularities associated with adsorption of phenol on carbons are as follows (2):

- (usually observed) a decrease in adsorption with a rise in surface acidity (as measured by Boehm titration) and with surface oxidation, problems in determining real equilibrium;
- chemisorption is inhibited by the presence of oxygen surface complexes; therefore, the sites responsible for phenol chemisorption are carbon sites of the active surface area, that is, oxygen free sites located at the edges of graphene layers whereas physisorption occurs all over the surface;
- irreversible adsorption is caused by oxidative coupling in the presence of molecular oxygen.

The authors mentioned pointed out the importance of the papers by Coughlin and Ezra (4) who found that adsorption of phenol decreases after oxidation of carbon, and they suggested that surface oxygen complexes increase the affinity of water to carbon and, therefore, solvent molecules can effectively block some micropores. Magne and Walker (5) suggested that phenol adsorbs by physisorption and chemisorption, but they did not define the mechanism of adsorption. Mattson and co-workers (6), Kipling (7), and Puri (8) emphasized the important role of surface carbonyls (the interaction of the π electrons of the benzene ring with the partial positive charge of the carbonyl carbon atom) in the adsorption of phenol and benzene. Coughlin and Ezra (4) suggested that oxygen chemisorbed on carbon appears to influence phenol adsorption only at low concentrations, when the molecules are thought to be adsorbed parallel to the surface due to nonpolar forces operating over the entire phenol molecules. At higher concentrations, phenol molecules are probably adsorbed in the vertical or "end-on" position due to intramolecular interaction with the hydroxyl groups directed away from the carbon surface. Under these conditions, surface oxygen has little influence on phenol adsorption. They also suggested that the decrease in phenol adsorption after oxidation can be caused by the removal of electrons from the π electron system of the basal planes as a result of oxygen chemisorption. They pointed out that adsorption of phenol takes place between the phenol aromatic ring and the π electron system of carbon basal planes.

In their paper, Radovic et al. (2) pointed out the two possible factors influencing the mechanisms of phenol adsorption on carbons. Citing the paper of Coughlin and Ezra (4) and Mattson and co-workers (9), they concluded that the mechanism of phenol adsorption is determined by so-called "$\pi-\pi$ interactions" and "donor–acceptor complex" formation. The first mechanism assumes that oxygen atoms bound to the carbon reduce the π electron density and weaken the dispersion forces between phenol π electrons ring and the π electrons of carbons. Second, one postulates that the adsorption mechanism is based on the formation of donor–acceptor complexes between the surface carbonyl groups (electron donors) and the aromatic rings of phenol acting as the acceptor and the decrease in phenol adsorption after carbon oxidation can be due to the oxidation of surface carbonyls (9).

The energetics of phenol adsorption was studied by Kiselev and co-workers (10) who considered the effect of the substitution of the benzene ring on the enthalpy of phenol adsorption on graphitized carbon black at "zero" coverage. They showed that the substitution of hydrogen for an OH group in the benzene ring increases the potential energy of interaction of the phenol molecule with the surface of graphitized carbon by about 15 kJ/mole and the substitution for an NH_2 group by about 14 kJ/mole.

The study by Jankowska and co-workers (11) paid attention to the role of basic groups in phenol adsorption, showing that they play a dominant role in this process. They also pointed out that the rise in the ratio of the immersion heat of carbon in benzene to the immersion heat in water leads to a rise in phenol adsorption. For that reason, their study is of extreme importance from the point of view of the adsorption mechanism of phenol on carbons. The authors concluded that phenol adsorption is limited by the micropore volume of the carbons, and the actual position beneath this limit is determined by their acid–base characteristics. The study by these authors performed with Koganovski (12) also pointed out that in the range of low equilibrium concentrations, the pore volume of carbon and also the ratio of acidic/basic group concentration determine phenol adsorption. The larger this ratio, the lower the adsorption. The results of Koganovski and co-workers were also confirmed in the paper by Tarasevich and Rak (13) showing that the modification of coal leading to the introduction of basic surface sites favors the adsorption of weak acid in micropores. There was also Speculation about the role of surface functionalities in phenol adsorption on carbons at infinite dilution. Biniak and co-workers (14) studied the isotherms of phenol adsorption on different carbons, and the main attention was paid to the correlation between the value of adsorbability and acidic properties of carbon surface; however, a linear relationship between these two values was possible only if it was assumed that the acidic groups diminish phenol adsorbability with different contributions. They also pointed out that moderately and slightly acidic groups are the most important ones influencing the mechanism of phenol adsorption on carbons.

Zawadzki (15), applying FTIR spectroscopy, showed that oxidation of the carbon surface does not always lead to a decrease in phenol adsorption. Contrary to others, he showed that phenol uptake increases after oxidation. He explained his results by showing that surface oxidation increases the binding energy between phenol and the carbon surface. The interaction with the surface is partially a specific one and occurs with the participation of π electrons of the adsorbed molecules. The main role is played by hydrogen bonds between phenol and surface oxides. Therefore, he concluded that the effect of the competition between water and phenol in the adsorption on surface-active sites is negligible and cannot result in a decrease in phenol adsorption after oxidation by nitric acid.

Calleja and co-workers (16) compared the different models of diffusion showing that the pore diffusion and linear models describe the kinetic adsorption curves of phenol on carbons satisfactorily. The study by Halhouli and co-workers (17) is one rarely met in the literature, where the effect of temperature on phenol adsorption was studied. They showed that the decrease in pH has no effect on the adsorption capacity of charcoal (unspecified by them) and the rise in the temperature increases the adsorption of phenol.

Singh and co-workers (18) pointed out the role of hydrogen bonding between phenolic hydrogen and chemisorbed oxygen present as oxygen surface groups. They also calculated the diameter of the phenol molecule as 0.62 nm. It is interesting that the computer simulation studies of liquid phenol (19) led to the diameter of 0.6 nm.

Loughlin and co-workers reported differences in the kinetics of phenol adsorption on carbons in the presence or absence of oxygen and proposed a model (20). They found that phenol diffusion is pore and surface diffusion and the reaction between phenol and oxygen is observed to be first order. They explained the results obtained as follows: both the molecules of oxygen and phenol diffuse to the surface of carbon and are adsorbed. Reaction occurs on the carbon surface and is catalyzed by it. If the order is assumed to be first, the dissociation of the oxygen molecules occurs before their reaction with phenol. On the contrary, if the order is second, the adsorbed oxygen molecules react directly with phenol.

Abuzaid and Nakhla (21) continued the study on this subject showing that temperature increases the rates of adsorption as well as phenol polymerization. They found values of diffusion energy around 15 kJ/mole, of the same magnitude as determined by others (22,23). The detailed study of phenol adsorption in oxic and anoxic conditions together with the results published by other authors led Terzyk and co-workers to the free-radical mechanism of phenol polymerization on carbons (22–24). The authors mentioned also concluded that the incorporation of surface oxygen groups, called “lactones” in Boehm’s method, leads to a decrease in the amount of chemically bound phenol.

Barton and co-workers (25) studied the effect of chemical heterogeneity on phenol adsorption. They showed that the fit of adsorption data described by the Dubinin–Radushkevich (DR) equation is good, and the characteristic adsorption energy values varied from 19–22 kJ/mole. Stoeckli et al. (26,27) showed that the Dubinin–Astakhov (DA) adsorption isotherm equation is applicable for describing phenol adsorption data on a series of microporous carbons and that the parameter n of this equation can be successfully approximated as 4. Following this idea, Terzyk developed a relationship between the characteristic energy of phenol adsorption ($n = 4$) and the average micropore diameter of carbon slit-like micropores (22,23).

The results of Barton et al. (25) supported the previous study by Singh (28) showing that phenol and aniline are more strongly adsorbed (from water) by basic than by acidic carbon. Their explanation is as follows: the process of phenol adsorption via hydrogen bonding with acidic surface sites should be accompanied by an energetically unfavored process of water–phenol bond breaking. Therefore, adsorption takes place mainly by the formation of electron donor–acceptor complexes between basic sites (principally π electron-rich regions on the basal planes of the carbon and the aromatic ring of the adsorbate). It appears, however, that further adsorption takes place by pore filling. The presence of water adsorbed by oxygen groups decreases the adsorption of phenol. The rise in phenol adsorption after oxidation as well as alkylation of coal was found by Jaoui et al. (29).

An interesting study of phenol and benzene adsorption on zeolites was recently done by Klemm and co-workers (30). They showed that the most important interactions in phenol adsorption are coulombic with Na^+ cations. An interesting study of the adsorption of phenols, stressing the role of solute–solvent interactions, was done by Dargaville and co-workers (31). Laszlo and co-workers (32) pointed out that phenol molecules and also water are present on the surface; therefore, the competition between them should be taken into account in calculating the cross-sectional area of adsorbed phenol; moreover, some pores can be inaccessible to adsorbed phenol–solvent aggregates.

Interesting results have been recently published by Nevskaia and co-workers (33). They showed that for nonoxidized carbon, the mechanism of adsorption is based mainly on the interaction between π electrons in phenol and π electrons of carbon; after oxidation, donor–acceptor complexes arise between basic surface oxygen groups and phenol aromatic rings. The study of the kinetics of phenol adsorption performed by the same group (34) showed that for nonmodified carbon, diffusion within the porous structure controls the adsorption rate; for oxidized carbon, the controlling step is the adsorbate–surface interaction, and a shorter time is required to reach the adsorption equilibrium in comparison with the original carbon.

Mahajan et al. (35) pointed out that both water and phenol can form hydrogen bonds with surface groups. Salvador and Merchan (36), studying the thermal desorption of phenolic compounds adsorbed on carbons, concluded that the values of adsorption energies are in the range of the energy of hydrogen bonds. Moreover, they concluded that this energy does not increase with ring substitution but rather with an increase in the electrophilic nature of the substituents.

A very important paper confirming this point of view was also published recently by Franz and co-workers (37). The authors emphasized the importance of the hydrogen bonding created by the molecules of solute and surface groups and suggested a mixed mechanism between the π–π dispersion interactions and bonding of water molecules to surface groups (mainly “carboxylic”). They also postulated that this effect is different for different solute molecules and it is absent in phenol adsorption from water solutions (i.e., water adsorption on active surface groups predominates). The final conclusion was that Mattson's mechanism is not a driving force for adsorption of aromatics on carbons.

Terzyk (38), studying the adsorption of a series of organic compounds on carbons, concluded that the effect of carbon surface chemistry on the adsorption of organics decreases with a rise in temperature. Thus it seems obvious that, differences in the mechanisms of adsorption (more accurately the domination of one of them), observed by different authors are caused by the temperature dependence of the energy of intermolecular interactions. The dispersion forces are temperature independent, whereas the electrostatic interactions between dipoles of adsorbed molecules (and solvent) and surface polar groups generally decrease with a rise in temperature. This implies that the measurements of the temperature dependence of adsorption can be a very useful tool for elaborating the mechanisms of this process. The study published recently by Garcia - Araya and co-workers (39) pointed out the importance of papers showing the temperature dependence of adsorption from solutions on carbons. Costa et al. (40) attributed the rise in phenol adsorption with temperature to an activated adsorption effect, Chiou and Manes (41) concluded that this results from packing effects, and Mattson and co-workers (6) that it is chemisorption. In the opinion of the authors of this study, chemisorption is usually small, so the temperature dependence of adsorption is mainly due to the decrease in the amount of electrostatically interacting molecules with surface groups located at the entrances to small micropores where the so-called "primary micropore filling" process occurs.

A more general mechanism of adsorption was proposed by Radovic and co-workers (2). They pointed out the importance of the pH of the solution and the surface chemistry of carbon, as well as the relative importance of electrostatic and dispersive adsorbate/adsorbent interactions. They also concluded that even though different surface functional groups are responsible for the amphoteric character of carbon, their specific identification is less important than the determination of the isoelectric point and the point of zero charge. The authors concluded that adsorption of aromatic anions occurs primarily on the basal plane of carbon and the electron-withdrawing effects of surface groups suppress the interaction of basal planes with the adsorbate's aromatic rings. Adsorption of aliphatic anions also occurs predominantly on the basal plane, and the same electron-withdrawing effects enhance the interaction with carboxyl anions. Although electrostatic interactions are important, the $\pi-\pi$ dispersion interactions appear to be dominant in the adsorption of aromatic solutes. They also pointed out that if dispersion forces dominate the functionalization (of the carbon adsorbent or the aromatic adsorbate), which decreases the π electron density, this leads to suppressed or weaker adsorption.

In his review, Moreno-Castilla (3) also pointed out the importance of the irreversible adsorption of phenols on carbons. This process was studied extensively in the pioneering paper of Grant and King (42) and by Tamon et al. (43,44) and others (22,24). The occurrence of this effect complicates the observation of regularities during the adsorption process.

It can be seen that despite an enormous number of papers, there are still unsolved problems:

The role of porosity; some authors implied that the adsorption of phenol is mainly controlled by the porous structure, and others thought that it is not. There is no correlation between phenol uptake and total surface area of carbons, and the details of the role of surface functional groups in phenol adsorption are still unknown; although different general mechanisms of phenol adsorption have been proposed, we still lack the general adsorption mechanism taking into account the "micropore filling" process and also answering the question, exactly what types of surface functional groups influence the value of phenol adsorption beneath the previously mentioned limit of micropore filling. The irreversibility of phenol adsorption on carbons is fairly well established, however, the detailed mechanism of this process as well as the well documented answer to the question, what types of surface functional groups (and why) are responsible for this process are still unknown. The influence of surface functional groups on the kinetics of phenol adsorption has not yet been determined, that is it is not known what types of surface functionalities determine the kinetics of adsorption of that compound.

Recently, Terzyk (22) presented the results of phenol adsorption–desorption (and kinetic) measurements at three temperatures and at two pH levels on a series of microporous activated carbons. Comparing the results of phenol adsorption with those determined for paracetamol, acetanilide, and aniline, it was shown that at a low equilibrium concentration, phenol adsorption decreases with the basic and carboxylic group content of the carbon surface. A linear correlation was observed between the average adsorbability of phenol and the total concentration of surface bases and carboxyls [the same author showed that the results for two Polish adsorbents follow the same line (23)]. The plot of the results tabulated in the paper by Biniak and co-workers (14) shows a similar tendency Therefore, carboxyls, that is, the most acidic and some of the basic groups, play the most important role in phenol adsorption at low concentrations. For adsorption at the acidic pH level, the value of adsorbability was correlated with the enthalpy of carbon immersion in water (therefore, with the total content of surface acidic groups and the pH of the zero point charge). Consequently, at an acidic pH, where all the carbons studied are below the pH of the point of zero charge, the more polar the surface of carbon, the lower the adsorbability of phenol. Therefore, it can be presumed that at infinite dilution and neutral pH, the positive surface charge resulting from adsorption of protons by the strongest surface bases leads to an increase in water adsorption and blocks some more active sites. The same effect is caused by undissociated surface acidic groups, and both of them decrease phenol adsorbability, compared to that observed for nonmodified carbon. The same author showed that in the range of micropore filling the pore blocking effect by phenol molecules dominates, and this effect weakens with a rise in temperature. Thus, the mechanism of phenol adsorption process is mixed, and

$\pi - \pi$ interactions between the phenol aromatic ring and pore walls are "screened" by the blocking of pore entrances by water and phenol molecules (interactions *via* hydrogen bonds) (38).

BIBLIOGRAPHY

1. Jaroniec, M. and Madey, R. (1988). *Physical Adsorption on Heterogeneous Solids*. Elsevier, Amsterdam, p. 251.
2. Radovic, L.R., Moreno-Castilla, C., and Rivera-Utrilla, J. (2001). *Chem. Phys. Carbon* **27**: 227.
3. Moreno-Castilla, C. (2004). *Carbon* **42**: 83.
4. Coughlin, R.W. and Ezra, F.S. (1968). *Environ. Sci. Technol.* **2**: 291.
5. Magne, P. and Walker, P.L. (1986). *Carbon* **24**: 101.
6. Mattson, J.S., Mark, H.B., Jr., Malbin, M.D., Weber, W.J., Jr., and Crittenden, J.C. (1969). *J. Colloid Interface Sci.* **31**: 116.
7. Kipling, J.J. (1967). *Porous Carbon Solids*. R.L. Bond (Ed.). Academic Press, London, p. 65.
8. Puri, B.R. (1970). *Chem. Phys. Carbon* **6**: 191.
9. Mattson, J.S. and Mark, H.B. (1971). *Activated Carbon. Surface Chemistry and Adsorption from Solution*. Marcel Dekker, New York, p. 209.
10. Beljakova, L.D., Kiselev, A.V., and Kovaleva, N.V. (1966). *Zhurn. Fiz. Khim.* **40**: 1494.
11. Jankowska, H., Świątkowski, A., Gelbin, D., Radeke, K.H., and Seidel, A. (1986). *Ext. Abstr. Int. Carbon Conf. Carbon 86*, Baden–Baden, Germany, p. 397.
12. Koganovski, A.M., Zaidel, A., and Radeke, K.H. (1987). *Khim. Technol. Vody* **6**: 500.
13. Tarasevich, Y.I. and Rak, V.S. (1988). *Kolloid. Zhurn* **60**: 84.
14. Biniak, S., Kaźmierczak, J., and Świątkowski, A. (1989). *Adsorption Sci. Technol.* **6**: 182.
15. Zawadzki, J. (1989). *Chem. Phys. Carbon* **21**: 147.
16. Calleja, G., Serna, J., and Rodriguez, J. (1993). *Carbon* **31**: 691.
17. Halhouli, K.A., Darwish, N.A., and Al-Dhoon, N.M. (1995). *Sep. Sci. Technol.* **30**: 3313.
18. Singh, B., Madhusudhanan, S., Dubey, V., Nath, R., and Rao, N.B.S.N. (1996). *Carbon* **34**: 327.
19. Mooney, D.A., Muller-Plathe, F., and Kremer, K. (1998). *Chem. Phys. Lett.* **294**: 135.
20. Loughlin, K.F., Hassan, M.M., Ekhator, E.O., and Nakhla, G.F. (1996). In *Fundamentals of Adsorption*. M.D. LeVan (Ed.). Kluwer, Boston, MA, p. 537.
21. Abuzaid, N.S. and Nakhla, G.F. (1997). *Sep. Sci. Technol.* **32**: 1255.
22. Terzyk, A.P. (2003). *J. Colloid Interface Sci.* **268**: 301.
23. Terzyk, A.P. (2003). *Adsorption Sci. Technol.* **21**: 539.
24. Terzyk, A.P., Rychlicki, G., and Gauden, P.A. (2003). *Carbon 03*. Extended Abstracts, Oviedo, Spain.
25. Barton, S.S., Evans, M.J.B., and MacDonald, J.A.F. (1997). *Polish J. Chem.* **71**: 651.
26. Stoeckli, F., Lopez–Ramon, M.V., and Moreno–Castilla, C. (2001). *Langmuir* **17**: 3301.
27. Stoeckli, F. and Hugi-Cleary, D. (2001). *Russ. Chem. Bull.* **50**: 2060.
28. Singh, D.H. (1971). *Indian J. Chem.* **9**: 1369.
29. Jaoui, M., Gruber, R., Cagniant, D., Nosyrev., I.E., Rogalski, M., and Achard, C. (1997). *Proc. ICCS'97*. P&W Druck and Verlag, Essen, p. 1831.
30. Klemm, E., Wang, J., and Emig, G. (1998). *Microporous Mesoporous Mater.* **26**: 11.
31. Dargaville, T.R., Guerzoni, F.N., Looney, M.G., and Solomon, D.H. (1996). *J. Colloid Interface Sci.* **182**: 17.
32. Laszlo, K., Botta, A., Nagy, L.G., and Cabasso, I. (1999). *Colloids Surf. A.* **151**: 311.
33. Nevskaia, D.M. and Guerro-Ruiz, A. (2001). *J. Colloid Interface Sci.* **234**: 316.
34. Nevskaia, D.M., Santianes, A., Munoz, V., and Guerro-Ruiz, A. (1999). *Carbon* **37**: 1065.
35. Mahajan, O.P., Moreno-Castilla, C., and Walker, P.L. (1980). *Sep. Sci. Technol.* **15**: 1733.
36. Salvador, F. and Merchan, M.D. (1996). *Carbon* **12**: 1543.
37. Franz, M., Arafat, H.A., and Pinto, N.G. (2000). *Carbon* **38**: 1807.
38. Terzyk, A.P. (2004). *J. Colloid Interface Sci.* **275**: 9.
39. Garcia-Araya, J.F., Beltran, F.J., Alvares, P., and Masa, F.J. (2003). *Adsorption* **9**: 107.
40. Costa, E., Calleja, G., and Marijuan, L. (1989). *Adsorption Sci. Technol.* **5**: 213.
41. Chiou, C.C.T. and Manes, M. (1974). *J. Phys. Chem.* **78**: 622.
42. Grant, T.M. and King, C.J. (1990). *Ind. Eng. Chem. Res.* **29**: 264.
43. Tamon, H. and Okazaki, M. (1996). *J. Colloid Interface Sci.* **179**: 181.
44. Tamon, H., Atsushi, M., and Okazaki, M. (1996). *J. Colloid Interface Sci.* **177**: 384.

CARBONATE GEOCHEMISTRY

Kenneth F. Steele
University of Arkansas
Fayetteville, Arkansas

Calcite ($CaCO_3$) is one of the most common minerals at or near the surface of the earth and thus is one of the main contributors to carbonate geochemistry. Less common carbonate minerals include dolomite [$MgCa\ (CO_3)_2$]; aragonite, a calcite polymorph (mineral of the same composition as calcite but having a different atomic structure); azurite and malachite (copper hydroxycarbonate minerals); siderite ($FeCO_3$); and rhodochrosite ($MnCO_3$) which can be important spatially and economically.

Although calcite does occur in igneous and metamorphic rocks, the mineral is present dominantly in the sedimentary environment which is the focus of this article. Calcite may be precipitated to form a relatively pure monomineralic rock (limestone), be present as a cement binding sediments into indurated rock, occurs simply as a trace mineral in rocks, and composes many shells and fossils. Weathering of calcite is relatively easy and depends largely on the amount of carbonic acid present. Other carbonate minerals, except dolomite, generally are dissolved and precipitated in a manner geochemically similar to calcite.

CALCITE SOLUBILITY PRODUCTS

The solubility of $CaCO_3$ depends on which polymorph is considered and the reference for the equilibrium constant.

For example, at 25 °C, the ion concentration products (K_{sp}) for calcite and aragonite (the most common polymorphs of $CaCO_3$) are $10^{-8.48}$ and $10^{-8.34}$, respectively (1). Krauskopf and Bird (2) provide calcite and aragonite K_{sp} values of $10^{-8.35}$ and $10^{-8.22}$, respectively. The K_{sp} values for calcite solubility from Krauskopf and Bird (2) are 35% greater than those from APHA (1). A similar relationship exists for aragonite; Krauskopf and Bird (2) present K_{sp} values for aragonite that are 31% higher. Variation in solubility experiments can yield significant differences in the K_{sp}.

Temperature

Counter to the typical increase of mineral solubility with increasing temperature observed for most minerals, many carbonate minerals are more soluble in cold water. For example, the K_{sp} for calcite at 0 °C and at 50 °C are $10^{-8.02}$ and $10^{-8.63}$, respectively (3). These values represent about a fourfold difference in calcite solubility caused by temperature alone. A thorough discussion of the thermodynamic reactions of carbonate minerals solubilities is beyond the scope of this article. A simplified explanation why carbonate minerals are more soluble in cold water is that the dissolution reaction for carbonate minerals is exothermic, which results in higher temperatures favoring the solid phase over dissolved ions. A more detailed discussion for this unusual behavior can be found in Langmuir (4).

In addition to an increase in solubility products with decreasing temperature, carbon dioxide is more soluble at lower temperatures, further favoring carbonate mineral dissolution in cooler environments. More carbonic acid is present in cold water at a given partial pressure of carbon dioxide (p_{CO_2}), and carbonic acid concentration is the controlling factor for the solubility of carbonate minerals under natural conditions. Carbon dioxide is obtained from the air or from decomposition of organic matter that releases carbon dioxide which reacts with water to form carbonic acid. These factors combine to form the calcite compensation depth (about 4500 m) below which calcite dissolves (2).

When geothermal waters reach the surface of the earth and precipitate tufa, a calcium carbonate rock, some reaction must overcome the temperature decrease, or these waters would dissolve calcite rather than deposit tufa as the water cools. The required reaction is the loss of CO_2 due to lower p_{CO_2} pressure at the earth's surface and subsequent decrease in the amount of carbonic acid, which controls the amount of dissolved calcite (see Calcite Solubility below).

Pressure

The increased partial pressure of CO_2 near the surface of the earth increases the amount of CO_2 that dissolves in water, therefore, increasing the solubility of carbonate minerals such as calcite. Pressure alone does not affect the solubility of calcite as much as the effect of temperature. Nonetheless, where pressure is large, for example, deep within the ocean, its effect alone can increase calcite solubility about twofold (2).

Ionic Strength

Other ions in solution shield Ca^{2+} and CO_3^{2-} ions from interacting and precipitating. Faure (5) describes this situation as "... the activity of the ions in electrolyte solutions is less than their concentration. It is plausible to expect that the interference by the other ions increases with their concentrations and charges." Activity can be defined as the concentration of an ion at zero ionic strength. The effect of the concentration and charge of ions is represented by the ionic strength (I) of the solution:

$$I = 0.5\ \Sigma\ m_i z_i^2 \tag{1}$$

where m_i is the molar concentration of an ion (i) and z is the charge of the ion.

Ionic strength can be use to calculate the activity coefficient γ which relates molar concentration to activity; γ is 1 or less.

$$m_i(\gamma) = a \tag{2}$$

where a is the activity of the ion.

γ can be calculated from the Debye–Hückel equation (Eq. 3) or other similar equations that use I and some constants that relate to the dielectric constant of water and temperature (A and B) and effective size (a) of ions. Values of the constants can be obtained from Faure (5).

$$-\log\gamma = \frac{Az^2 I^{1/2}}{I + aBI^{1/2}} \tag{3}$$

SATURATION INDEX

Waters tend to precipitate $CaCO_3$ when oversaturated with $CaCO_3$, and tend to dissolve $CaCO_3$ if undersaturated with $CaCO_3$. The Saturation Index (SI) is the most widely used method to determine the amount of calcite that will be precipitated or dissolved. Many water chemistry computer programs calculate RS (relative saturation) which is related to SI as

$$\mathrm{SI} = \log \mathrm{RS} \tag{4}$$

where RS = ratio of $CaCO_3$ activity product (K_a) to $CaCO_3$ solubility product (K_{sp}).

SI values of 1 indicate saturation, negative values indicate undersaturation, and positive values indicate oversaturation.

MEASURED ALKALINITY

Alkalinity is the measure of the acid-consuming ability (power) of a solution expressed as mg/L $CaCO_3$. In nature, the main ions that neutralize H^+ ions are HCO_3^- and CO_3^{2-} (6). Although OH^-, PO_4^{3-} and other negative ions can neutralize H^+ ions, generally the concentrations of these ions is so small in the natural environment that HCO_3^- and CO_3^{2-} are the only significant acid-neutralizing ions. In some situations, especially in anthropogenically influenced situations, OH^- can be important (e.g., the impact of cement on water chemistry); thus, OH^- is included in the following calculations. Alkalinity is usually determined by titrating a 50 or 100 mL water sample with 0.02 N H_2SO_4 to a pH of 8.3 for CO_3^{2-} and/or OH^-. This 8.3 pH end point is often referred to as the phenolphthalein (or *P*) end point because this indicator changes colors at pH 8.3. Titration continues past 8.3 to a pH of about 4.5 as any HCO_3^- (originally present or formed from partial neutralization of CO_3^{2-}) is neutralized. The pH of 4.5 can vary from 4.3–4.9 depending on the presence and amounts $CaCO_3$ and certain anions (1). This final end point is referred to as total alkalinity. These pH titration end points are based on inflection points on the titration curve of pH versus milliliters of acid added to the water sample titration curve for neutralization of the bases. The amount of acid necessary to reach the phenolphthalein end point is equivalent to the amount of base ($CaCO_3$) neutralized and is referred to as *P* alkalinity. The amount of $CaCO_3$ to neutralize "completely" or "totally" the H^+ ions added is referred to as *T* alkalinity. Thus, *T* must be larger than or equal to *P*. There are five alkalinity conditions possible: (1) only HCO_3^-, (2) only CO_3^{2-}, (3) only OH^-, (4) HCO_3^- plus CO_3^{2-}, and (5) CO_3^{2-} plus OH^-. Bicarbonate and OH^- ions cannot coexist under typical natural conditions because these ions react to form CO_3^{2-} ions and H_2O. These five situations are discussed briefly below.

Bicarbonate Alkalinity. If the initial pH of a sample is less than 8.3, there is only bicarbonate alkalinity and no *P* alkalinity. Or one can state that there is only *T* alkalinity (the amount of acid necessary to lower the pH below 4.5).

Carbonate Alkalinity. There are two abrupt changes in pH during alkalinity titration, one at 8.3 and the other at 4.5. If only CO_3^{2-} ions are present, the amount of acid necessary to lower the pH below 8.3 is one-half the total acid (i.e., *T*) necessary to lower the pH below 4.5. The first abrupt pH change at 8.3 represents the conversion of carbonate to HCO_3^- (i.e., half of the carbonate has been neutralized), and the second abrupt pH change at 4.5 represents neutralization of the HCO_3^- created from carbonate. If only CO_3^{2-} ions are present, $T = 2P$.

Hydroxide Alkalinity. If the initial pH is greater than 8.3 and the addition of acid rapidly lowers the pH below 4.5, there is only hydroxide alkalinity and thus $T = P$.

Bicarbonate Plus Carbonate Alkalinity. Because P represents the conversion of CO_3^{2-} to HCO_3^-, the same amount of acid (*P*) is required to neutralize the HCO_3^- ions formed from the CO_3^{2-} ions; 2*P* represents complete neutralization of CO_3^{2-} ions present. In addition, HCO_3^- originally present is included in *T*, which means that $T > 2P$. Because CO_3^{2-} ions are present, there will be two abrupt changes in pH (at 8.3 and 4.5).

Carbonate Plus Hydroxide Alkalinity. When both CO_3^{2-} and OH^- ions are present, the amount of acid required to lower the pH below 8.3 is that necessary to convert CO_3^{2-} to HCO_3^- plus that necessary to neutralize OH^- ions (i.e., the OH^- milliequivalent/L (meq/L) concentration is equal to T minus CO_3^{2-} meq/L). In addition, T-P is equal to 1/2 the CO_3^{2-} (the amount of CO_3^{2-} converted to HCO_3^{2-} concentration) or CO_3^{2-} meq/L equals 2(T-P). These two relationships can be combined to yield OH^- meq/L equals to 2P-T. These conditions constrain the *P* and *T* titration relationship to $T < 2P$.

Table 1. Alkalinity Determined by Titration

Result of Titration	Bicarbonate Alkalinity	Carbonate Alkalinity	Hydroxide Alkalinity
$P = 0$	T	0	0
$2P = T$	0	T	0
$P = T$	0	0	T
$P > 0.5T$ or $T > 2P$	$T - 2P$	$2P$	0
$P < 0.5T$ or $T < 2P$	0	$2(T - P)$	$2P - T$

Table 1 summarizes the equations necessary to determine the portion of alkalinity due to HCO_3^-, CO_3^{2-} ions, and OH^-. At pH 9, the OH^- concentration is only 1.0 mg/L as $CaCO_3$ and can generally be neglected in alkalinity determinations for water whose pH is less than 9.0.

CALCULATION OF BICARBONATE, CARBONATE, AND HYDROXIDE ALKALINITY

Bicarbonate, carbonate, and hydroxide alkalinity can be calculated from total alkalinity (TA) if the pH is known. These concentrations are often calculated, especially for carbonate and hydroxide ion concentrations that are generally low and difficult to measure. The ion concentrations must be given in molar (M) concentrations, which are indicated by brackets in this article. For most natural situations, Eq. 5, which takes into account charge differences, is adequate for determining the various sources of alkalinity:

$$2\mathrm{TA} = [\mathrm{HCO_3^-}] + 2[\mathrm{CO_3^{2-}}] + [\mathrm{OH^-}] \tag{5}$$

TA is commonly expressed as mg/L $CaCO_3$ and must be converted to moles/L (M) to make the following calculations. The conversion of mg/L to M $CaCO_3$ is accomplished by using Eq. 6

$$(\mathrm{TA\ moles/L\ CaCO_3}) = \frac{\mathrm{mg/L\ CaCO_3}}{100{,}000\ \mathrm{mg/mole}} \tag{6}$$

Rearranging the second disassociation constant for carbonic acid (Eq. 7) in terms of $[CO_3^{2-}]$ yields Eq. 8:

$$K_2 = \frac{[\mathrm{H^+}][\mathrm{CO_3^{2-}}]}{[\mathrm{HCO_3^-}]} = 10^{-10.3} \tag{7}$$

$$[\mathrm{CO_3^{2-}}] = \frac{K_2[\mathrm{HCO_3^-}]}{[\mathrm{H^+}]} \tag{8}$$

Substitution of Eq. 8 for $[CO_3^{2-}]$ in Eq. 5 and expressing $[OH^-]$ in terms of $\frac{K_w}{[H^+]}$, based on Eq. 9, yields Eq. 10.

$$[\mathrm{OH^-}] = \frac{K_w}{[\mathrm{H^+}]} = \frac{10^{-14}}{[\mathrm{H^+}]} \tag{9}$$

$$2\mathrm{TA} = [\mathrm{HCO_3^-}] + 2\left[\frac{K_2[\mathrm{HCO_3^-}]}{[\mathrm{H^+}]}\right] + \frac{K_w}{[\mathrm{H^+}]} \tag{10}$$

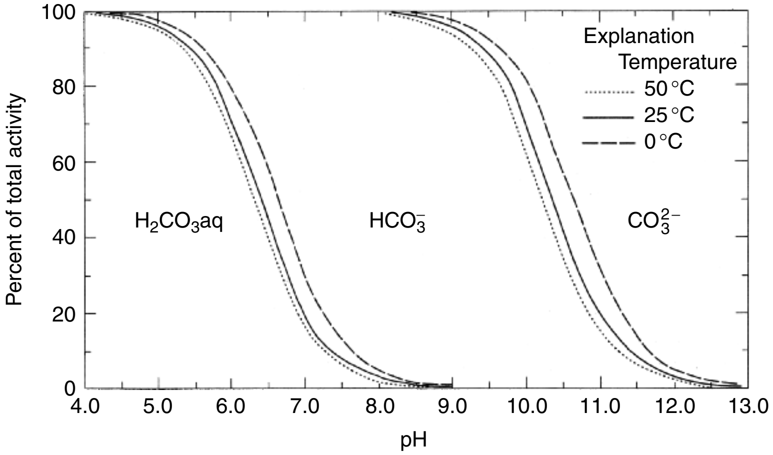

Figure 1. Plot of percent of total activity of carbonate species versus pH at 0, 25 and 50 °C. From Hem (7).

Rearranging Equation 10 to solve for HCO_3^- (Eq. 12) is carried out by the following step (Eq. 11) and expressing K_w as exponent of 10 and negative pH as exponent of 10.

$$2TA - \frac{K_w}{[H^+]} = [HCO_3^-]\left(1 + \frac{2K_2}{[H^+]}\right) \tag{11}$$

$$[HCO_3^-] = \frac{2TA - 10^{K_w + pH}}{1 + 2K_2 10^{pH}} \tag{12}$$

Once $[HCO_3^-]$ has been calculated (or determined by titration), this concentration can be substituted in Equation 8 and $[CO_3^{2-}]$ can be calculated. To determine the hydroxide alkalinity, only the pH must be known (Eq. 9).

To convert these molar alkalinity concentrations (Eqs. 8, 9, and 12) to mg/L, the molar concentrations must be multiplied by the molecular weight of $CaCO_3$ (100,000 mg/mole) and for $[HCO_3^-]$ and $[OH^-]$, a factor of 1/2 is required because 1 mole of $CaCO_3$ is equivalent to 2 moles of HCO_3^- and 2 moles of OH^-, whereas 1 mole of CO_3^{2-} is equivalent to 1 mole of $CaCO_3$.

These are the equations necessary for calculating bicarbonate, carbonate, and hydroxide alkalinities as $CaCO_3$ mg/L:

Bicarbonate alkalinity ($CaCO_3$ mg/L)

$$= 50{,}000\ \left(\frac{2TA - 10^{K_w + pH}}{1 + 2K_2 10^{pH}}\right) \tag{13}$$

Carbonate alkalinity ($CaCO_3$ mg/L)

$$= 100{,}000\ (K_2[HCO_3^-](10^{pH})) \tag{14}$$

Hydroxide alkalinity ($CaCO_3$ mg/L)

$$= 50{,}000\ (10^{K_w + pH}) \tag{15}$$

CARBONATE SPECIES

Concentrations of the carbonate species $[H_2CO_3]$, $[HCO_3^-]$ and $[CO_3^{2-}]$ can be calculated if the total dissolved inorganic carbonate (DIC) concentration and pH are known. [DIC] is often expressed as HCO_3^- because it is the dominant ion at typical pH values (Fig. 1). Figure 1 shows the concentrations of the carbonate species for an aqueous solution with respect to the pH and temperature of the solution. Note that activities (concentrations) are higher at lower temperatures. The pertinent equilibrium reactions are the two disassociation constants for H_2CO_3 (Eqs. 8 and 16).

$$K_1 = \frac{[H^+][HCO_3^-]}{[H_2CO_3]} = 10^{-6.4} \tag{16}$$

$$[DIC] = [H_2CO_3] + [HCO_3^-] + [CO_3^{2-}] \tag{17}$$

By rearranging K_1 (Eq. 16) to express $[H_2CO_3]$ in terms of $[HCO_3^-]$, by rearranging K_2 (Eq. 7) to express $[CO_3^{2-}]$ in terms of $[HCO_3^-]$, and then substituting these values in Equation 17, the following equation is obtained:

$$\begin{aligned}[DIC] = {} & [[H^+]\ [HCO_3^-](K_1^{-1})] \\ & + [HCO_3^-] + [[H^+]^{-1}\ [HCO_3^-](K_2)]\end{aligned} \tag{18}$$

Factoring out $[HCO_3^-]$, rearranging Eq. 16 to solve for $[HCO_3^-]$, and expressing $[H^+]$ as pH (negative exponent of 10) yields the following:

$$[HCO_3^-] = \left[\frac{[DIC]}{(10^{-pH - K_1}) + (10^{pH + K_2}) + 1}\right] \tag{19}$$

Because K_1, K_2, and K_w are temperature dependent, the concentrations of the carbonate species must be determined by using the appropriate values for a specific temperature.

CALCITE SOLUBILITY

If only calcite were present in an aqueous solution with no carbon dioxide, one could determine the solubility of calcite by obtaining the square root of the solubility product for calcite (Eq. 20). Reported K_{sp} values for calcite vary with the reference and range from $K_{sp} = 10^{-8.3}$ to $K_{sp}\ 10^{-8.5}$ at 25 °C (1,2). The value of $10^{-8.3}$ will be used throughout this article.

$$[CaCO_3] = [Ca^{2+}][CO_3^{2-}] = K_{sp} = 10^{-8.3} \tag{20}$$

Hydrolysis

Although one can calculate the solubility of Ca^{2+} as $10^{-4.15}$ M or 6.8×10^{-5} from Equation 20, the CO_3^{2-} ion will undergo hydrolysis (Eq. 21):

$$CaCO_3 + H_2O = Ca^{2+} + HCO_3^- + OH^- \qquad (21)$$

$$K = [Ca^{2+}][HCO_3^-][OH^-] = 10^{-12} \qquad (22)$$

Inspection of Eq. 21 indicates that all three product ions are equal. The cube root of 10^{-12}, yields a concentration of 10^{-4} M for all three ions; thus, the pH of the solution is 10. At this high pH, there is appreciable disassociation of HCO_3^- to CO_3^{2-} and H^+ (which is quickly neutralized). Then these CO_3^{2-} ions undergo hydrolysis producing equal amounts of HCO_3^- and OH^- (Eq. 23).

$$CO_3^{2-} + H_2O = HCO_3^- + OH^- \qquad (23)$$

Furthermore, water with a pH of 10 has a $[CO_3^{2-}]/[HCO_3^-]$ ratio of 1/2 (from Eq. 7). The total carbonate in this situation is the sum of HCO_3^- and CO_3^{2-} (H_2CO_3 is negligible at this high pH—see Fig. 1), and total carbonate is equal to Ca^{2+} (Eq. 24).

$$\Sigma[CO_3^{2-}] = [HCO_3^-] + [CO_3^{2-}] = [Ca^{2+}] \qquad (24)$$

Because $\dfrac{[CO_3^{2-}]}{[HCO_3^-]} = 0.5$, Eq. 7 can be modified to

$$[Ca^{2+}] = [HCO_3^-] + 0.5\,[HCO_3^-] = 1.5\,[HCO_3^-] \qquad (25)$$

Instead of Ca^{2+} having the same concentration as HCO_3^- and OH^-, Ca^{2+} will be equal to 1.5 $[HCO_3^-]$ and 1.5 $[OH^-]$ because $[HCO_3^-] = [OH^-]$ (Eq. 23).

Equation 22 is then modified to

$$K = [Ca^{2+}]\,1.5\,[HCO_3^-]\,1.5\,[OH^-] \qquad K = 2.25 \times 10^{-12} \qquad (26)$$

Because $[Ca^{2+}] = 1.5[HCO_3^-] = 1.5[OH^-]$, the Ca^{2+} concentration can be estimated by the cube root of 2.25×10^{-12} (Eq. 26) or 1.3×10^{-4} (Eq. 26). M (equivalent to 5.2 mg/L). This value is about twice that obtained in Equation 20 and represents the lowest equilibrium concentration of Ca^{2+} for pure water *not* in contact with CO_2. Using this method results in a pH change from 10.00 to 9.94. Recalculation using this method based on the new pH of 9.94 yields a 2% change in $[Ca^{2+}]$ and a new pH of 9.93. Another iteration of these calculations yields a $[Ca^{2+}]$ difference of 1% and the pH remains 9.93.

Carbonic Acid

The main factor controlling carbonate solubility in nature is the amount (partial pressure) of CO_2 that forms H_2CO_3 (Eq. 27). The partial pressure of CO_2 in the atmosphere is 0.0003 or 0.03% of the atmospheric pressure at the surface of the earth (i.e., 1 bar or atmosphere). Partial pressure is equal to moles per liter (5); thus, there is $10^{-5.0}$ M H_2CO_3 in water in contact with the atmosphere (Eqs. 26 and 27).

$$CO_2 + H_2O = H_2CO_3 \qquad (27)$$

$$K = \frac{[H_2CO_3]}{[CO_2)(H_2O]} = \frac{[H_2CO_3]}{(0.0003)(1)} = 10^{-1.5} \qquad (28)$$

$$[H_2CO_3] = 3 \times 10^{-4}(10^{-1.5}) = 10^{-5.0}\text{M} \qquad (29)$$

In most natural environments, the appropriate equation for calcite solubility is Equation 30 where CO_2 is present:

$$CaCO_3 + H_2CO_3 = Ca^{2+} + 2HCO_3^- \qquad (30)$$

Equation 29 (2) takes into account the presence of CO_2 at the earth's surface which forms carbonic acid (Eq. 27) that readily dissolves $CaCO_3$:

$$K = \frac{[Ca^{2+}][HCO_3^-]^2}{[H_2CO_3]} = 10^{-4.4} \qquad (31)$$

$[HCO_3^-] = 2[Ca^{2+}]$ from inspection of Eq. 30; therefore,

$$K = 10^{-4.4} = \frac{[Ca^{2+}][2Ca^{2+}]^2}{[H_2CO_3]} = \frac{4[Ca^{2+}]^3}{10^{-5.0}} \qquad (32)$$

The Ca^{2+} concentration calculated from Equation 32 is 4.7×10^{-4} M (equivalent to 18.8 mg/L). This is the solubility of calcite at surface conditions, 25 °C and 1 atmosphere pressure, which provides a "minimum" concentration of atmospheric carbonic acid. More CO_2 can be dissolved in cold water, and the CO_2 partial pressure in soil can be rather high, leading to increased calcite solubility. Decaying vegetation in soil can have a CO_2 partial pressure in excess of 0.1 atmospheres (2) which can significantly raise H_2CO_3 concentrations of infiltrating meteoric water. The solubility of calcite in water in contact with 0.1 atmosphere of CO_2 can be calculated as in Eqs. 28 and 32, producing a calcite solubility of 3.2×10^{-3} M (equivalent to 128 mg/L Ca^{2+}). This value is not an absolute maximum but does give an approximate upper limit for Ca^{2+} concentrations from limestone near the surface of the earth. Other acids, for example, sulfuric from the oxidation of sulfide minerals such as pyrite (FeS_2), can also have a significant effect on calcite solubility locally.

CAVE FORMATION

Carbonic acid-rich water forms limestone caves, which is the most common type of cave. When the water table is high, carbonic acid-rich water dissolves the limestone (calcite). Later when the water table drops, a "void" filled with air is formed and a large void is called a cave. Smaller amounts of water rich in Ca^{2+} and HCO_3^- may continue to flow through the void. When these waters enter the void, CO_2 partial pressure decreases and CO_2 is released. This degassing of CO_2 drives the reactions in Equations 27 and 30 to the left, leading to the precipitation of calcite and formation of stalagmites, stalactites, and other cave features. Although evaporation may play a small role in the deposition of calcite, the loss of CO_2 is the most important factor.

HARDNESS

Water hardness is a measure of how difficult it is for soap to lather; most of the divalent ions in water must be precipitated by soap before soap can produce suds. The most abundant divalent ions in nature are Ca^{2+} and Mg^{2+}, although Fe^{2+} and other divalent ions can be important in some waters. The major source of Ca^{2+} and Mg^{2+} is carbonate minerals; therefore, hardness is expressed as mg/L $CaCO_3$. The range in hardness values with descriptive terms is shown in Table 2.

In the past, hardness was determined by a titration analysis of divalent ions (7). Today, hardness is calculated by converting Ca^{2+} and Mg^{2+} molar concentrations to mg/L $CaCO_3$ by multiplying the mg/L concentrations of the two ions by the inverse of the equivalent weights of the ions and summing the two concentrations (Eq. 31):

$$\text{Hardness as mg/L CaCO}_3 = 2.497\ (\text{Ca}^{2+}\ \text{mg/L}) + 4.118\ (\text{Mg}^{2+}\ \text{mg/L}) \qquad (33)$$

Some scientists consider hardness more of a water quality (use) factor that relates to scaling in hot water heaters and in industrial setting, rather than an important geochemical factor. However, comparison of hardness and alkalinity (both expressed as mg/L $CaCO_3$) can yield helpful water chemistry information. If hardness exceeds alkalinity, the excess is termed “noncarbonate” hardness, meaning that there are noncarbonate mineral sources of Ca^{2+} and Mg^{2+} ions. The remainder of the hardness is considered “carbonate” hardness derived from carbonate minerals. In this discussion, it is assumed that there is “no” hydroxide alkalinity. If alkalinity exceeds hardness, there must be carbonate mineral sources that do not contain Ca^{2+} and/or Mg^{2+}, for example, $NaHCO_3$, or cation exchange has occurred with clay minerals (e.g., Ca^{2+} replacing Na^+).

Table 2. Hardness Ranges. Units are in mg/L of Calcium Carbonate[a]

0–60	Soft
51–120	Moderately hard
121–180	Hard
>180	Very hard

[a] Reference 7.

DOLOMITE

Dolomite has a very small K_{sp}; values range from 10^{-17} to 10^{-19} (2). The K_{sp} is too small to produce large thicknesses of dolomite and furthermore there is no true dolomite being precipitated today. Some “protodolomite” is precipitated in restricted environments, such as evaporite lagoons. Dolomite should not exist, and yet there are thick sequences of dolostone throughout the world and throughout the geologic record—hence, the “dolomite problem.”

Dolostone characteristically preserves fossils poorly, is coarse grained, and commonly has cavities/pore spaces, which all indicate replacement of limestone as the origin of the rock. In addition, it is known that saline water, high pH, a low Ca^{2+}/Mg^{2+} ratio, and elevated temperatures all favor dolostone formation. Based on these factors, various marine and deep water reflux models have been proposed, as well as mixing models for fresh water and sea water, but no “perfect” model has been developed for the formation of dolomite. The dolomite atomic structure is highly ordered which may affect the amount of time necessary to develop the structure. Dissolution of dolomite also is complicated because dolomite undergoes incongruent dissolution as Mg^{2+} dissolves before Ca^{2+} (2). At any rate, the formation and dissolution of dolomite is complex and not well understood in comparison with other carbonate minerals.

BIBLIOGRAPHY

1. Eaton, A.D., Clesceri, L.S., and Greenberg, A.E. (Eds.). (1995). *Standard Methods for the Examination of Water and Wastewater*. American Public Health Association, Washington DC, pp. 2-25–2-38 and 4-17–4-18.
2. Krauskopf, K.B. and Bird, D.K. (1995). *Introduction to Geochemistry*, 3rd Edn. McGraw-Hill, New York.
3. Garrels, R.M. and Christ, C.L. (1965). *Solutions, Minerals and Equilibria*. Harper & Row, New York.
4. Langmuir, D. (1997). *Aqueous Environmental Geochemistry*. Prentice-Hall, Upper Sanddle River, NJ.
5. Faure, G. (1998). *Principles and Applications of Geochemistry*, 2nd Edn. Prentice-Hall, Upper Saddle River, NJ.
6. Drever, J.I. (1997). *The Geochemistry of Natural Waters-Surface and Groundwater Environments*, 3rd Edn. Prentice-Hall, Upper Saddle River, NJ.
7. Hem, J.D. (1985). *Study and Interpretation of the Chemical Characteristics of Natural Water*. 3rd Edn. U.S. Geological Survey Water Supply Paper 2254, U.S. Geological Survey, Alexandria, VA, p. 263.

CARBONATE IN NATURAL WATERS

Robert Y. Ning
King Lee Technologies
San Diego, California

INTRODUCTION

Carbonate is the divalent anion resulting from complete neutralization of carbonic acid. In natural waters, it is in dynamic equilibrium with carbon dioxide gas, carbonic acid, bicarbonate anion, and solid forms of insoluble carbonate minerals—mainly calcium carbonate:

$$CO_2(g) + H_2O \Longleftrightarrow H_2CO_3 \Longleftrightarrow HCO_3{}^- \Longleftrightarrow CO_3{}^{2-} \Longleftrightarrow \text{Carbonate(s)}$$

Various aspects of the CO_2–carbonate systems in natural waters have been extensively reviewed (1–6). Practical interests lie in the areas of geochemistry, biochemistry, water treatment and usage, and environment.

GEOCHEMISTRY

Carbonate Dissolution

Typically, limestone [$CaCO_3$] and dolomite [$CaMg(CO_3)_2$] rocks are dissolved by infiltrating rainwater which has taken up biogenic CO_2 in the soil zone. As this solution percolates downward, continuously dissolving solid carbonates, the hydrogeochemical system will change from “open” with respect to a CO_2 gas phase in the unsaturated zone to “closed” with respect to that phase in the saturated zone. Thus, the solution passes through various stages of a system, which are intermediate between the limiting cases, in general. In modern times, the dissolution of carbonate rocks is enhanced by the action of acids of sulfur and nitrogen oxides from atmospheric pollution.

Carbonate Deposition

Carbonate rocks, mainly aragonite and calcite, constitute about 11% of the total sediments of the earth’s crust. Bicarbonate salts are relatively soluble in water, but calcium carbonate is not. Freshwater does not sustain extreme supersaturation of calcium carbonate. In most cases, the phase precipitated is calcite—the polymorph of calcium carbonate that has the lowest solubility product. Carbonate solubility in seawater is governed essentially by the same parameters, but kinetically the precipitation of calcium carbonate is complicated by high concentrations of Mg^{2+} and the actions of marine organisms.

pH Buffer

The pH of natural waters on land is controlled mainly by the bicarbonate–carbonate system. Most of the CO_2 absorbed into the ocean is in the form of HCO_3^- (93%) and CO_3^{2-} (6%). The pH of the upper regions of the sea is controlled by the carbonate as well as the borate systems. In the lower regions of the sea, pH appears to be controlled by the precipitation and solution of various minerals. The bicarbonate alkalinity of natural waters has the following typical ranges: rainwater 0 to 2.5 ppm (mg/L), river and groundwater 6 to 600 ppm, and seawater 140 to 159 ppm.

Geochemical Carbon Cycle

The atmospheric CO_2 concentration of about 0.04% on a timescale of million of years is controlled mainly by a combination of degassing from the earth’s interior balanced by the uptake of carbon dioxide via the weathering of calcium and magnesium silicate minerals (7,8). Carbon dioxide in water as carbonic acid is converted to dissolved calcium and magnesium bicarbonate on land during weathering. This carbon form is stored in lakes and aquifers or carried to the sea by rivers and precipitated there as calcium and magnesium carbonate minerals. The overall reactions shown here were first formulated by Urey (8) in 1952.

$$CO_2 + CaSiO_3 \Longleftrightarrow CaCO_3 + SiO_2$$
$$CO_2 + MgSiO_3 \Longleftrightarrow MgCO_3 + SiO_2$$

The reactions from left to right represent carbon dioxide absorption via weathering. The reactions from right to left represent degassing due to the thermal decarbonation of carbonates after burial at sufficient depth in the earth’s crust. Added to the Urey reactions should also be the organic carbon subcycle. It exerts an important additional control on the atmospheric carbon dioxide level. This involves the removal of carbon dioxide from the atmosphere through photosynthesis by organisms, the burial of organic matter in sediments, and the weathering of organic matter (or its thermal decomposition) which returns carbon dioxide. The overall reaction is represented as

$$6\ CO_2 + 6\ H_2O \Longleftrightarrow (CH_2O)_6 + 6\ O_2$$

Together, these reactions regulate carbon dioxide on a geologic timescale.

BIOCHEMICAL ASPECTS

Biological Pump

Life has developed on the earth in the water phase and in general requires buffering of the pH at around 7.0. Here carbonate–bicarbonate plays a major role. The process of photosynthetic fixation of CO_2 in the photic zone of the upper ocean in the presence of sunlight and nutrients is balanced by the export of organic carbon (soft tissues, secretions) and carbonate (bones, shells) out of the mixed layer into the deep ocean by gravity. Chemically, the “biological pump” is composed of the organic carbon pump and carbonate pump (2):

$$CO_2 + H_2O + \text{energy} \Longleftrightarrow (CH_2O)_X \ \text{(organic carbon)} + O_2$$
$$Ca^{2+} + 2\ HCO_3^- \Longleftrightarrow CaCO_3 \ \text{(bones, shells)} + H_2O + CO_2$$

The organic carbon pump sequesters CO_2 from the mixed surface layer and dumps it as particulate organic carbon into the deeper ocean, sometimes seen as white flakes of “marine snow.” The carbonate pump, on the other hand, acts in the direction opposite to the organic carbon pump by producing CO_2.

Oceanic Chalk

Much of the calcium carbonate is in the form of hard parts in calcareous organisms (e.g., foraminifera, cocolithophores). Other marine carbonates (aragonite, magnesian, calcite) are metastable phases whose origin is in the reluctance of the Mg^{2+} to be dehydrated and converted to dolomite, [$MgCa(CO_3)_2$] and magnesite,

[$MgCO_3$]. Carbonate sediments in sunlit waters of tropical seas originate as organically secreted solids. They occur either as skeletal debris, as organically secreted carbonate mud, or in various combinations of the two. More than 70% of the total $CaCO_3$ contained in large reef systems is in the form of fine, unconsolidated sand deposited in thick beds across large areas adjacent to a living reef frame (4). In addition to skeletal particles, nonskeletal carbonate particles are important, for example, ooids, pellets, pelloids, intraclasts, and grapestones that produce thick carbonate deposits in the rock record. Enormous petroleum reservoirs discovered in carbonate rocks (Devonian of Canada, Pennsylvanian and Permian strata of New Mexico and Texas, Mesozoic and Cenozoic deposits in Middle East) point to the intimate association of organic carbon with carbonate deposits.

Biological Carbonates

Calcite and aragonite, two polymorphs of $CaCO_3$ represent the bulk of the carbonates formed in biological systems. The solubility products of these two forms are 4.7×10^{-9} (calcite) and 6.9×10^{-9} (aragonite), which shows that aragonite is less stable than calcite. Surface seawater to a depth of 3,000 m is commonly supersaturated with respect to aragonite and calcite (6). This disequilibrium state is attributed to a series of inhibitors, most notably magnesium and dissolved organic matter, that interfere with the nucleation of a $CaCO_3$ seed crystal. By the same token, it appears to offer an optimal condition for the biocontrolled crystallization of $CaCO_3$ to sustain an abundance of marine life.

WATER TREATMENT AND UTILIZATION

Carbonate Scaling

Relevant to water treatment is the fact that natural waters contain high concentrations of soluble bicarbonate salts in dynamic equilibrium with carbonic acid (dissolved carbon dioxide) and insoluble carbonate salts. To estimate the amount of the three forms in water, it is important to know that carbonic acid (H_2CO_3) is a divalent acid whose $pK_1 = 6.35$ and $pK_2 = 10.33$. At pH = 6.35, the un-ionized acid and the bicarbonate ($HCO_3{}^-$) ion exist in equal concentrations. pH is a logarithmic scale, so at pH = 7.35 (one unit higher), the solution would contain 90% bicarbonate and 10% carbonic acid, and at pH = 8.35, the solution would contain 99% bicarbonate and 1% carbonic acid. The pH of 8.2–8.4 (phenolphthalein end point in alkalinity titrations) is also significant in that above this pH, carbonate ($CO_3{}^{2-}$) ions begin to appear. At pH = 9.3, about 10% of the bicarbonate ions would have converted to carbonate ion, and at pH = 10.33, carbonate and bicarbonate ion concentrations would be equal. Calcium carbonate is quite insoluble (14 mg/L water at 25°C), so natural waters are usually found in the pH range of 6–8; bicarbonate ions range between 6 to 600 mg/L, and calcium ranges from 2 to 200 mg/L. Seawater representing 98% of blended natural waters on the surface of the earth typically contains 410 mg/L of Ca^{2+} ions and 143 mg/L of $HCO_3{}^-$ ions. Magnesium carbonate is about 100 times more soluble than calcium carbonate, so the magnesium ion concentration in seawater is much higher (about 1310 mg/L). Carbonate scaling potentials in surface waters are focused mainly on calcium carbonate crystallization because it is the first carbonate scale to form under treatment or use (9,10).

Calcium Carbonate Scaling Indexes

The Langelier saturation index (LSI) and the form modified for high salinity water called the Stiff and Davis stability index (S&DSI) are widely used to predict calcium carbonate scaling potentials in cooling systems and boiler waters (9,10). These indexes are widely applied for predicting the same in reverse osmosis systems (11). The formulas for calculating the indexes (9–11) are given here:

$$\mathrm{LSI} = \mathrm{pH} - \mathrm{pH_s}$$

where $\mathrm{pH_s}$ is the pH at which the water is saturated in calcium carbonate and is calculated by the following:

$$\mathrm{pH_s} = (9.3 + A + B) - (C + D)$$

where

$A = (\mathrm{Log_{10}}[\mathrm{TDS}] - 1)/10$
$B = -13.12 \times \mathrm{Log_{10}}(^\circ\mathrm{C} + 273) + 34.55$
$C = \mathrm{Log_{10}}[\mathrm{Ca^{2+} as\ CaCO_3}] - 0.4$
$D = \mathrm{Log_{10}}[\mathrm{alkalinity\ as\ CaCO_3}]$

In this model, we see that calcium carbonate saturation is affected by salinity (total dissolved solids, TDS) and temperature along with calcium and bicarbonate ion concentrations. The Langelier index loses its accuracy in predicting the solubility of calcium carbonate in water when salt concentrations exceed 4,000 mg/L (11). Stiff and Davis modified Langelier's formula for highly saline waters by using empirically derived values. The Stiff and Davis stability index (S&DSI) is calculated as follows:

$$\mathrm{S\&DSI} = \mathrm{pH} - p_{\mathrm{Ca}} - p_{\mathrm{ALK}} - K$$

where

p_{Ca} is the negative log of the calcium concentration expressed as molarity;
p_{ALK} is the negative log of the total alkalinity concentration expressed as molarity;
K is a constant whose value depends on the water temperature and ionic strength.

At an index of less than 0, water will have a very limited scaling potential. At increasingly positive index values, the scaling potential increases. By projecting the pH, TDS, calcium, and alkalinity concentrations, calcium carbonate scaling indices can be calculated.

Carbonate Scale Control with Acid versus Antiscalant

Carbonate scale control in the purification or use of water and wastewater typically involve using either acid to reduce pH and carbonate concentration or antiscalant to inhibit the growth of calcium carbonate crystals, or a combination of the two. Sodium hexametaphosphate

(SHMP) is a generic antiscalant widely used due to its low cost. It acts, as do the newer organic antiscalants, by inhibiting the growth of calcium carbonate seed crystals (11,12). The limitations of SHMP are that it is stable only as an anhydrous solid. It is difficult to dissolve in the plant, and once dissolved in water, it slowly hydrolyzes to phosphate ions which precipitate calcium, magnesium, aluminum, iron, and other ions as insoluble phosphate salts; in addition, it is a stimulant for microbial growth. As an antiscalant for simultaneous control of other types of scales, it is much less effective than the newer types of organic antiscalants. Great advances have been made in the development and use of organic antiscalants. In a line of six distinctly differentiated organic antiscalants, all classified under ANSI-NSF Standard 60 for drinking water production by reverse osmosis (13), the maximum calcium carbonate scaling control possible ranges from LSI = 2.8 to 3.2 and S&DSI = 3.9 to 4.5. Under these high limits for calcium carbonate control, the differentiated antiscalants are selected more on the basis of simultaneous control of calcium, strontium, and barium sulfates, calcium fluoride, silica, iron, aluminum, and heavy metal scales. In each case when used, acidification of feedwater is not necessary.

ENVIRONMENTAL IMPACT

Concern for the environmental impact of the natural CO_2 balance in the biosphere is raised by the modern trend of deforestation, industrialization, burning of fossil fuels, and population explosion. Understanding the natural processes is the starting point for monitoring environmental changes and modulating human activities.

Natural Balance

The earth's atmosphere contains approximately $2,000 \times 10^9$ tons of CO_2, the oceans $130,000 \times 10^9$ tons, and the rocks of the earth together with fossil fuels approximately $40,000 \times 10^9$ tons. The concentration of CO_2 in the atmosphere is maintained at a relatively constant value of about 0.03% by weight by an annual exchange between the oceans and the air of approximately 100×10^9 tons of CO_2 gas. Fluctuations of CO_2 in the atmosphere occur due to photosynthesis in daylight hours, which fixes approximately 60×10^9 tons of CO_2 per year, and to biological respiration and decay, which release about the same amount.

Ocean Storage of CO_2

It is known that CO_2 will form solid hydrates with water at temperatures below 10°C and pressures higher than 4.4 MPa (450 m depth); at 0°C, the hydrate is stable at pressures higher than 1.2 MPa (125 m depth). The cold deep waters of the world's oceans are unsaturated with respect to CO_2 and have an enormous and under used capacity to dissolve CO_2. The high dissolved inorganic carbon level of the oceans (38,000 Gt C) would be little changed, even if they were to gain all the carbon in the known fossil fuel reserves (4,000 Gt C). The idea of injecting human-produced CO_2 directly into the deep ocean has attracted much interest (2).

BIBLIOGRAPHY

1. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry*, 3rd Edn. John Wiley & Sons, New York.
2. Wong, C.S. and Hirai, S. (1997). *Ocean Storage of Carbon Dioxide*. International Energy Assoc., Greenhouse Gas R&D Programme, Cheltenham, UK. Available: http://www.ieagreen.org.uk.
3. Matthess, G., Frimmel, F., Hirsch, P., Schulz, H.D., and Usdowski, H-E. (Eds.). (1992). *Progress in Hydrogeochemistry*. Springer-Verlag, New York.
4. Bhattacharyya, A. and Friedman, G.M. (Eds.). (1983). *Modern Carbonate Environments*. Hutchinson Ross Publishing Co., Stroudsburg, PA.
5. Loewenthal, R.E. and Marais, G.V.R. (1976). *Carbonate Chemistry of Aquatic Systems: Theory & Application*. Ann Arbor Science, Ann Arbor, MI.
6. Degens, E.T. (1976). *Inorganic Biochemistry*. *Topics in Current Chemistry*. Vol. 64, Springer-Verlag, New York.
7. Berner, R.A. (1995). A.G. Hogbom and the development of the concept of the geochemical carbon cycle. *American Journal of Science* **295**: 491–495.
8. Urey, H.C. (1952). *The Planets: Their Origin and Development*. Yale University Press, New Haven, CT, p. 242.
9. Kemmer, F.N. (Ed.). (1988). *The Nalco Handbook*. 2nd Edn. McGraw Hill, New York, pp. 4.10–4.18 and 6.4–6.6.
10. Betz Handbook of Industrial Water Conditioning. (1991), 9th Edn. Betz Laboratories, Trevose, PA, pp. 183, 184 and 223–225.
11. Byrne, W. (1995). *Reverse Osmosis: A Practical Guide for Industrial Users*. Tall Oaks Publishing Inc., Littleton, CO, pp. 111–117.
12. Ning, R.Y. (1999). Operational Characteristics of Reverse Osmosis Process Chemistry. *Ultrapure Water* **16**(4): 39–42.
13. *The six antiscalants are available from King Lee Technologies (San Diego, CA) under the Pretreat Plus™ name.* General information can be found on the web site: http://www.kingleetech.com.

CHLORINE-36 AND VERY OLD GROUNDWATERS

Gholam A. Kazemi
Shahrood University of Technology
Shahrood, Iran

INTRODUCTION

Chlorine has nine isotopes whose mass numbers range from 32 to 40, but only three of these occur naturally. The remaining six have half-lives shorter than 1 hour, (1), and this is why they do not occur in nature. The three abundant isotopes of chlorine include two stable isotopes, chlorine-35 and chlorine-37, with 75.53% and 24.47% abundances, respectively, (2) and one radioactive isotope, chlorine-36, with a half-life of 301,000 ± 2000 years (some references ± 4000 years). The natural mixture of these three isotopes makes up environmental chlorine.

PRODUCTION OF CHLORINE-36

Chlorine-36 is produced by the following processes:

1. cosmic ray splitting of ^{40}Ar in the atmosphere:

   $$^{40}\mathrm{Ar} + \mathrm{x} \longrightarrow {}^{36}\mathrm{Cl} + \mathrm{x}' + \alpha$$

   The amount of ^{36}Cl produced by this process depends mainly on the geographic latitude and is greater in middle latitudes. For example, the fallout rate of ^{36}Cl at Stripa in Sweden at 57 °C latitude is 14±3 atoms $\mathrm{m}^{-2}\ \mathrm{s}^{-1}$ (3), the mean atmospheric flux of ^{36}Cl over the United States is 19.6 ± 4.5 atoms $\mathrm{m}^{-2}\ \mathrm{s}^{-1}$, and the mean global ^{36}Cl flux is 30.5 ± 7 atoms $\mathrm{m}^{-2}\ \mathrm{s}^{-1}$ (4).
2. neutron activation of stable ^{35}Cl in the subsurface:

   $$^{35}\mathrm{Cl} + {}_0{}^1\mathrm{n} \rightarrow {}^{36}\mathrm{Cl}$$

   The rate of production of ^{36}Cl by this process varies depending on the type of rocks, minerals, and solutions, and the availability of neutron sources (neutrons are produced by bombardment of rocks with alpha particles as well as through spontaneous fission of ^{238}U). For instance, the ^{36}Cl production rate at the top of Murray Group Limestone was calculated by Davie et al. (1) (using the following equation) as 0.09 atoms/cm^3/year.

   $$\mathrm{P}_{36} = {}^{36}\mathrm{Cl}/{}^{40}\mathrm{Ca} \times \rho/\mathrm{MW} \times N_\mathrm{A} \times \log_\mathrm{e}{}^2/t_{1/2}$$

   where P_{36} is production rate of ^{36}Cl, ^{36}Cl/^{40}Ca is the equilibrium ratio of ^{36}Cl to ^{40}Ca, ρ is the density of the limestone, MW is the molecular weight of limestone, N_A is Avogadro's number, and $t_{1/2}$ is the half-life of ^{36}Cl.
3. Weapon testing of fusion devices starting in late 1952 and extending through mid-1958 has led to the production of up to 70,000 atoms $\mathrm{m}^{-2}\ \mathrm{s}^{-1}$ (5). Chlorine-36 produced by this phenomenon is used as a marker to age-date young groundwaters that are less than 50 years old (6,7).

It has been recently suggested that the operation of two nuclear reactors in the United States has led to the production of ^{36}Cl (8).

MEASUREMENT OF ^{36}CL CONTENT OF WATER SAMPLES AND REPORTING THE RESULTS

To analyze water samples for ^{36}Cl content, the following steps are undertaken (4,8): A specified volume of a water sample, depending on its chlorine concentration, is selected (the selected volume should contain a minimum of 2 mg chlorine). In some studies, sample volumes of up to 4 liters have been collected. The sample is then prepared for target loading. This means preparing about 20–350 mg pure AgCl and pressing it into copper sample holders. The samples are prepared for target loading by adding $\mathrm{AgNO_3}$ solution to the samples to precipitate AgCl. The precipitate (AgCl) is then purified by dissolving it with $\mathrm{NH_4OH}$. Next, $\mathrm{Ba(NO_3)_2}$ is added to precipitate unwanted sulfur as $\mathrm{BaSO_4}$. $\mathrm{HNO_3}$ is added next to neutralize the solution and in the last stage, the AgCl is reprecipitated with $\mathrm{AgNO_3}$ for target loading. If the $\mathrm{Cl^-}$ concentration of the sample is very low, a carrier must be added to it. The carrier may be AgBr of very low sulfur and chlorine content. The prepared water samples are then analyzed for ^{36}Cl content using an accelerator mass spectrometer (AMS). Usually, standard and blank samples are tested to identify the reproducibility and accuracy of the measurements. Davie et al. (1) reported that ^{36}Cl measurements by AMS have a long reproducibility of about 6% and a sensitivity for ^{36}Cl/Cl of a few parts in 10^{15}.

The measured chlorine-36 is reported in two ways:

1. as a ratio of $^{36}\mathrm{Cl/Cl} \times 10^{15}$ (the number of ^{36}Cl atoms to total number of chlorine atoms). The multiplication factor (10^{15}) is used because the actual concentration of ^{36}Cl is very small.
2. as atoms of chlorine-36 per liter of water using the following equation (8):

$$^{36}\mathrm{Cl\ (in\ atoms\ per\ liter)} = 1.699 \times 10^4 \times B \times R$$

where B is the concentration of chlorine in mg/l and R is the (^{36}Cl/total Cl) × 10^{15} ratio. Therefore, to determine the concentration of ^{36}Cl, the concentration of $\mathrm{Cl^-}$ in water samples must be measured by conventional methods such as titration or ion chromatography.

Example

The $^{36}\mathrm{Cl/Cl} \times 10^{15}$ of a water sample is 550 and its chlorine concentration is 200 mg/l. What is the ^{36}Cl concentration in this sample?

$$^{36}\mathrm{Cl\ (in\ atoms\ per\ liter)} = 1.699 \times 10^4 \times B \times R$$

^{36}Cl (in atoms per liter) = $1.699 \times 10^4 \times 200$ mg/l $\times 550 = 186.89 \times 10^7$ atoms/liter (see Table 1 for more examples).

APPLICATIONS OF CHLORINE-36 AND PRINCIPLES OF THE ^{36}CL DATING METHOD

Chlorine-36, in this article, is treated as a tool to age-date very old groundwaters. However, it has many other applications such as the geochronological study of young volcanic rocks (9), estimation of groundwater recharge (10,11), calculation of dispersivity (12) and a mixing study of two waterbodies (13).

The principle of this method is simple. It is based on the radioactive decay of ^{36}Cl in the subsurface groundwater system. Above the earth's surface, chlorine-36 produced (atmospheric ^{36}Cl) with an initial value, C_0, enters groundwater by rainwater infiltration. After time t, it decays and reaches a new concentration, C, according to the following equation:

$$C = C_0 e^{-\lambda t}$$

Table 1. ^{36}Cl and Cl^{-} Measurements of Murray Mallee Samples[a]

Well	^{36}Cl/Cl × 10^{-15b}	Cl^{-} mg/l	^{36}Cl^{-} Atoms/l × 10^{6b}
1	19 + 4	264	85 + 18
2	22 + 4	551	205 + 37
3	35 + 6	361	215 + 37
4	32 + 4	576	315 + 39
5	28 + 6	1067	510 + 110
6	3p17	314	195 + 37
7	31 + 7	475	250 + 56
8	31 + 7	468	245 + 55
9	43 + 7	804	585 + 95
10	22 + 8	747	280 + 100
11	32 + 15	752	410 + 190
12	42 + 7	2023	1450 + 240
13	21 + 6	3191	1140 + 330
14	39 + 7	2573	1710 + 310
15	33 + 6	1607	900 + 160
16	22 + 4	5327	1990 + 360
17	13 + 4	32,600	7200 + 2200
18	22 + 6	5566	2080 + 570

[a] From Reference 1, with permission from Geoscience Australia.
[b] The number after + is the error range.

From the initial concentration, C_0, and the present concentration, C, one can calculate the length of time that the ^{36}Cl has resided in the subsurface groundwater system. For example, if the concentration of ^{36}Cl in a groundwater sample is 2000 atoms/liter and its initial concentration (in the infiltrating rainwater) was 3000 atoms/liter, then the time "t" would be 176,136 years. This means that the water molecule, which contain the ^{36}Cl in question, entered the groundwater system 176,136 years ago.

However, many inaccuracies and problems surround C_0 and C values, the initial and the present concentration of ^{36}Cl, respectively. We do not accurately know the concentration of ^{36}Cl in the rainwater when it entered the subsurface system. Davis et al. (8) present six ways to find the C_0 value and argue that all contain weaknesses and inaccuracies. These approaches include:

1. calculation of theoretical cosmogenic production and fallout,
2. measuring ^{36}Cl in present-day atmospheric precipitation and using it as C_0,
3. assuming that shallow groundwater contains a record of C_0,
4. extracting ^{36}Cl from vertical depth profiles in desert soils,
5. recovering ^{36}Cl from cores of glacial age, and
6. calculating subsurface production of ^{36}Cl for water that has been isolated from the atmosphere for more than 1 million years.

The situation is no better with regard to C. The assumption in using the decay equation is that the ^{36}Cl atoms which enter subsurface groundwater system behave as an isolated packet when they migrate through the flow system (piston flow theory). As explained above, ^{36}Cl is produced in the atmosphere, and it is also produced in the subsurface. There are other subsurface sources of ^{36}Cl such as what is called "dead chlorine," chlorine present in fluid inclusions; ancient formation waters; saline water from compacting clays; and chlorine from salt beds and evaporites. In addition, subsurface mixing (mixing of low and high ^{36}Cl content waters), cross formational flow, diffusion between aquitard and aquifer, and dilution and evaporation processes all unite to complicate the job of finding which C should be used in the decay equation, the C that is measured in the laboratory or the C that is obtained when the contributions/effects of all the above factors have been eliminated.

Apart from ^{4}He which is considered an age indicator not a dating method (14), ^{36}Cl is the only method to age-date very old groundwaters quantitatively. However, in only a very few cases were groundwater ages satisfactorily determined. In many instances, the chlorine-36 dating study was not successful due to many factors such as the uncertainty in estimating C_0, the effect of dilution/evaporation processes on the value of C, and the much longer half life of ^{36}Cl compared with the age of the groundwater (1,15,16). Other counterproductive factors include relatively recent introduction of the method, less interest in very old groundwaters, a limited number of well-defined regional aquifers worldwide for which this method is appropriate, many analytical considerations and the unavailability of equipment in a large number of countries, difficulty in interpreting the results (initial value problem, local interferences, many sources of ^{36}Cl, etc,), and inherent inadequacies [for example, Park et al. (19) show that if the chlorinity of the water sample exceeds ~75–150 mg/kg, the $^{\mathrm{Cl}}36$ method cannot be used for groundwater dating]. These may be the reasons that an important text book such as that by Mazor (14) does not include a section on ^{36}Cl to describe it as an age-dating tool. A valuable recent paper by Park et al. (17) describes the principles and limitations of age-dating very old groundwater by ^{36}Cl.

CASE STUDIES

The work by Davis and Schaeffer (18) was the first attempt to apply ^{36}Cl isotopes to environmental problems. However, Davie et al. (1) argue that extensive application of ^{36}Cl in environmental studies started only after the introduction of the accelerator mass spectrometry (AMS) technique for detecting ^{36}Cl by Elmore and his colleagues in 1979 (19). Because earlier methods of analysis such as decay counting required 13 g of chlorine, a difficult task, and it had much less sensitivity than AMS.

Examples of aquifers subjected to chlorine-36 dating include the Great Artesian Basin (GAB) of Australia (20,21); the fractured metamorphic rock aquifer in the Clare Valley, South Australia (22); the Milk River aquifer of the western Canada sedimentary basin (23,24); the Aquia aquifer in the Maryland Coastal Plain (25); the palaeochannel groundwater of Kalgoorlie in Western Australia (15); and the groundwater brines of the Dead Sea area (16). Of these aquifers, GAB is perhaps the most studied as far as chlorine-36 is concerned. A list of directly related publications dealing with chlorine-36 in various aquifers around the globe is included at the end of this article for further consultation. It is interesting to note that all publications dealing with ^{36}Cl have been produced by groups of authors; there is no single author publication, which may be an indication of the complexity of the matter.

One of the early studies dealing with ^{36}Cl in groundwater is that by Davie et al. (1), in which the ^{36}Cl content of limestone groundwater in the Murray Basin of Australia was measured. The results of their analysis, as an example, are presented in Table 1. This publication was selected for detailed explanation because it has not been cited elsewhere and it is a valuable study. Table 1 shows that the ^{36}Cl/Cl ratios of the Murray Mallee samples are low, generally lower than those reported by Davis et al. (8) for groundwaters of the United States (It may be inappropriate to compare the two study areas because of the large difference in their sizes). This study showed that the ratio of ^{36}Cl/Cl in the aquifer increases along the flow path most probably because of significant localized recharge.

The age of the deep Jurassic confined aquifer in the GAB was estimated by Bentley et al. (20) at $(1200 \pm 500) \times 10^3$ years (much older than the hydraulic age). Notle et al. (23) measured the ^{36}Cl concentration of 12 samples from the Milk River aquifer, Alberta, Canada, and found that the ages of the groundwater ranged from 0.6–2 Ma at a distance of 80 km from the recharge area.

By analyzing 17 samples from water wells and seeps, Turner et al. (15) showed that the ^{36}Cl/Cl ratios and ^{36}Cl content of palaeochannel groundwater in Kalgoorlie, Western Australia, ranged from $32–129 \times 10^{-15}$ and 1460 to 79,000 atoms/l, respectively. Similarly, the ^{36}Cl/Cl ratios of 17 samples from the Continental Terminal aquifer in Niger calculated by Le Gal La Salle et al. (26) ranged from $3–27 \times 10^{-15}$.

Davis et al. (8) measured the concentration of ^{36}Cl in the groundwater of the United States to determine what they call the "preanthropogenic levels" (natural = pristine = background level) of ^{36}Cl in groundwater. This study which was subsequently completed by Moysey et al. (4) is the most comprehensive study worldwide of ^{36}Cl in groundwater so far; a total of 183 water samples from unpolluted wells and springs were analyzed. Davis et al. (8) showed that natural ratios of ^{36}Cl/Cl are lowest near the coast and increase to a maximum in the central Rocky Mountains of the United States (less than 50×10^{-15} in Florida to more than 1200×10^{-15} in the Rocky Mountains). They attribute this pattern to an inland decrease in the stable chlorine content of atmospheric precipitation; the increase in the ratio of ^{36}Cl/Cl inland is due to an inland decrease in stable chlorine, not an increase in ^{36}Cl.

BIBLIOGRAPHY

1. Davie, R.F. et al. (1989). Chlorine-36 measurements in the Murray Basin- preliminary results from the Victorian and South Australian Mallee region. *BMR J. Aust. Geol. Geophys.* **11**(2/3): 261–272.
2. Marques, J.M., Andrade, M., Aires-Barros, L., Graca, R.C., Eggenkamp, H.G.M., and Antunes da Silva, M. (2001). 87Sr/86Sr and 37Cl/35Cl signatures of CO2-rich mineral waters (N-Portugal): Preliminary results. In: *New Approaches Characterizing Groundwater Flow*. K.P. Seiler and S. Wohnlich (Eds.). Swets and Zeitlinger Lisse, Lisse, Vol. 2, pp. 1025–1029.
3. Bowen, R. (1988). *Isotopes in the Earth Sciences*. Elsevier, London.
4. Moysey, S., Davis, S.N., and Cecil, L.D. (2003). The distribution of meteoric ^{36}Cl/Cl in the United States: A comparison of models. *Hydrogeol. J.* Online Publication, 10.1007/s10040-003-0287-z.
5. Bentley, H.W., Phillips, F.M., and Davis, S.N. (1986). Chlorine-36 in the terrestrial environment. In: *Handbook of Environmental Isotope Geochemistry*. P. Fritz and J.C. Fontes (Eds.). Elsevier, New York.
6. Coplen, T.B. (1993). Uses of environmental isotopes. In: *Regional Ground-Water Quality*. W.M. Alley (Ed.). ITP, New York, pp. 255–293.
7. Cook, P.G. and Solomon, D.K. (1997). Recent advances in dating young groundwater: chlorofluorocarbons, ^{3}H/^{3}He and ^{85}Kr. *J. Hydrol.* **191**: 245–265.
8. Davis, S.N., Moysey, S., Cecil, L.D., and Zreda, M. (2003). Chlorine-36 in groundwater of the United States: Empirical data. *Hydrogeol. J.* **11**: 217–227.
9. Leavy, B.D., Philips, F.M., Elmore, D., Kubic, P.W., and Gladney, E. (1987). Measurement of ^{36}Cl/Cl in young volcanic rocks: An application of accelerator mass spectrometry to geochronology. *Nucl. Instrum. Methods Phys. Res.* **B29**: 246–250.
10. Cook, P.G. and Robinson, N.I. (2002). Estimating groundwater recharge in fractured rocks from environmental ^{3}H and ^{36}Cl, Clare Valley, South Australia. *Water Resour. Res.* **38**(8): 10.1029/2001WR000772.
11. Cecil, L.D., Beasley, T.M., Pittman, J.R., Michel, R.L., Kubik, P.W., Sharma, P., Fehn, U., and Gove, H.E. (1992). Water infiltration rate in the unsaturated zone at the Idaho National Engineering Laboratory estimated from chlorine-36 and tritium profiles and neutron logging. In: *Water Rock Interaction*. Vol. 1, Y.F. Kharaka and A.S. Maest (Eds.). Balkema, Rotterdam, pp. 709–714.
12. Cecil, L.D., Welhan, J.A., Green, J.R., Frape, S.K., and Sudicky, E.R. (2000). Use of chlorine-36 to determine regional scale aquifer dispersivity, eastern Snake River Plain aquifer, Idaho/USA. *Nucl. Instrum. Methods Phys. Res.* **B172**: 679–687.
13. Calf, G.E., Bird, J.R., Kellet, J.R., and Evans, W.R. (1998). *Origins of Chlorine Variation in the Murray Basin- Using Environmental Chlorine-36*. Australian Bureau of Mineral Resources, Geology and Geophysics Record 1988/43, Groundwater series No. 16, pp. 28–30.
14. Mazor, E. (1991). *Applied Chemical and Isotopic Groundwater Hydrology*. Open University Press, Melton Keynes.
15. Turner, J.V., Rosen, M.R., Fifield, L.K., and Allan, G.L. (1995). Chlorine-36 in Hypersaline Palaeochannel Groundwaters of Western Australia. In: *Application of Tracers in Arid-zone Hydrology*. E.M. Adar and C. Leibundgut (Eds.). *IAHS publication* no. 232, Vienna, pp. 15–33.
16. Yechieli, Y., Ronen, D., and Kaufman, A. (1996). The source and age of groundwater brines in the Dead Sea area, as deduced from ^{36}Cl and ^{14}C. *Geochim. Cosmochim. Acta* **60**(11): 1909–1916.
17. Park, J., Bethke, C.M., Torgerson, T., and Johnson, T.M. (2002). Transport modeling applied to the interpretation of groundwater ^{36}Cl age. *Water Resour. Res.* **38**(5): 10.1029/2001WR000399.

18. Davis, R.F. and Schaeffer, O.A. (1955). Chlorine-36 in nature. *Ann. NY Acad. Sci.* **62**: 105–122.
19. Elmore, D. et al. (1979). Analysis of ^{36}Cl in environmental water samples using an electrostatic accelerator. *Nature* **272**: 22–25; errata 246.
20. Bentley, H.W. et al. (1986). Chlorine-36 dating of very old groundwater 1. The Great Artesian Basin, Australia. *Water Resour. Res.* **22**: 1991–2001.
21. Love, A.J., Herczeg, A.L., Sampson, L., Cresswell, R.G., and Fifield, K.L. (2000). Sources of chlorine and implications for ^{36}Cl dating of old groundwater, southwestern Great Artesian Basin. *Water Resour. Res.* **36**: 1561–1574.
22. Cook, P.G. (2003). Groundwater ages in fractured rock aquifers. *Proc. Int. Conf. Groundwater Fractured Rocks*, Prague, IHP-VI, series on groundwater no. 7, pp. 139–140.
23. Notle, E. et al. (1991). Measurements and interpretations of ^{36}Cl in groundwater, Milk River aquifer, Alberta, Canada. *Appl. Geochem.* **6**: 435–445.
24. Phillips, F.M., Bentley, H.W., Davis, S.N., Elmore, D., and Swanick, G.B. (1986). Chlorine-36 dating of very old groundwater, 2, Milk River Aquifer, Alberta, Canada. *Water Resour. Res.* **22**: 2003–2016.
25. Purdy, C.B. et al. (1996). Aquia Aquifer dissolved Cl^{-} and ^{36}Cl/Cl: Implications for flow velocities. *Water Resour. Res.* **32**: 1163–1171.
26. Le Gal La Salle et al. (1995). Old groundwater circulation in the Iullemeden basin (Niger): Preliminary results of an isotopic study. USA. In: *Application of Tracers in Arid-zone Hydrology*. E.M. Adar and C. Leibundgut (Eds.). *IAHS publication* no. 232, IAEA, Vienna, pp. 129–139.

READING LIST

Andrews, J.N., Fontes, J.C., Michelot, J.L., and Elmore, D. (1986). Insitu neutron flux, ^{36}Cl production and groundwater evolution in crystalline rocks at Stripa, Sweden. *Earth Planet Sci. Lett.* **77**: 49–58.

Fritz, P. and Fontes, J.C. (Eds.). (1986). *Handbook of Environmental Isotope Geochemistry*. Elsevier, New York.

Davis, S.N. et al. (2000). Chlorine-36 in groundwater containing low chlorine concentrations. *Groundwater* **38**: 912–921.

Hendry, M.J., Wassennar, L.I., and Kotzer, T. (2000). Chlorine and chlorine isotopes (^{36}Cl and δ^{37}Cl) as tracers of solute migration in a thick, clay rich aquitard system. *Water Resour. Res.* **36**: 285–296.

Torgersen, T. and Phillips, F.M. (1993). Reply to "Comment on chlorine 36 dating of very old groundwater, 3. Further results on the Great Artesian Basin, Australia by T. Torgersen et al., " by J.N. Andrews and J.C. Fontes. *Water Resour. Res.* **29**: 1875–1877.

CHLOROFLUOROCARBONS (CFCs)

Gholam A. Kazemi
Shahrood University of
Technology
Shahrood, Iran

INTRODUCTION

The following paragraph, extracted mostly from a comprehensive paper by Badr et al. (4), is a brief yet valuable introduction to chlorofluorocarbons (CFCs). Thereafter, an overview of CFC applications in hydrologic studies, especially for age-dating young groundwater, is presented.

Before 1930, the refrigerants commonly used were ammonia (R-717, the most popular refrigerant in the industrial and heavy commercial sectors), carbon dioxide (R-747), ethyl chloride (R-160), isobutane (R-600a), methyl chloride (R-40), methylene chloride (R-30), and sulfur dioxide (R-764). Because of either, toxicity or fire hazard, the need for a safer refrigerant became clear. The new refrigerant had to be nonflammable, have a boiling point between −40 and 0 °C, be highly stable, and low in toxicity. In response to these requirements, the CFC family of refrigerants was developed. Dichloromonofluoromethane (R-21) was the first CFC developed that had attractive thermodynamic properties for use as a refrigerant, but due to its low toxicity, stability, and inertness, it had never been used commercially as a refrigerant. The first CFC suitable for commercial applications was dichlorodifluoromethane (R-12), whose discovery was announced in 1930 at the American Chemical Society meeting in Atlanta, Georgia. CFCs have been among the most useful chemical compounds ever developed. There is no known natural source of CFCs (5).

In addition to their uses as working fluids in refrigerators and air-conditioning systems, CFCs have also been employed, since 1950, as blowing agents for foams and plastics, aerosol propellants, solvents for cleaning precision and delicate electronic equipment, as well as for dry cleaning, in sterilizers for surgical instruments and catheters, and as working media in Rankine-cycle engines for waste-heat recovery systems. By 1974, approximately one million tonnes per year of CFCs were being produced worldwide. Within the European Countries (EC), the main uses for CFCs (according to 1988 consumption figures) are as aerosol propellants (38%), foam-blowing agents (35%), solvents (17%), and refrigerants in cooling systems (10%). In the USA, CFCs are used for the following purposes: foam-blowing (34%), refrigeration and air conditioning (35.5%), solvents (18%), sterilizers (5.4%), and other applications (7.0%). CFCs along with some other gases harm the earth's atmosphere by two phenomena: (1) depletion of the ozone layer and (2) the greenhouse effect (or global warming). A major cause of environmental concern for CFCs is their long atmospheric lifetime (55 years for CFC-11 and 140 years for CFC-12). Hydrochlorofluorocarbons (HCFCs) are now replacing CFCs and recent data suggests that the growth rate of CFCs concentration in the atmosphere has begun to decrease (6).

The concentration of CFCs in the atmosphere is measured regularly at a few stations worldwide. The atmospheric concentrations for CFC-11, CFC-12, and CFC-113, those species that are used to age-date young groundwaters (less than 50 years old), are illustrated in Fig. 1. Industrial use of CFC-12 started in the 1930s, CFC-11 in the 1950s, and CFC-113 in the 1970s (7). Atmospheric concentrations of these species show little spatial variation; only 10% variation is observed among average concentrations in Ireland, Oregon, Barbados,

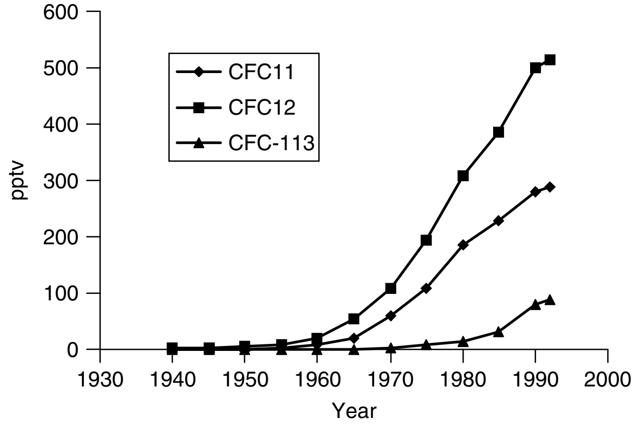

Figure 1. Atmospheric concentrations of CFC-11, CFC-12, and CFC-113 (pptv) from 1940 to 1992. (*Source*: Smethie et al., 1992, cited in Ref. 8. redrawn)

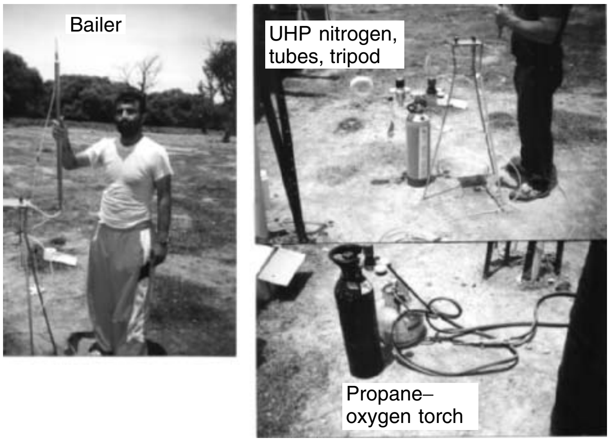

Figure 2. Apparatus for sampling groundwater for CFC analysis.

Samoa, and Tasmania (source: Cunnold et al., 1994, cited in Ref. 8).

CFCs in hydrological studies—Age dating young groundwaters

The use of CFCs in hydrology began when Thompson et al. (1) discussed the usefulness of a special type of fluorocarbon known as Freon 11, as a tracer in both groundwater and surface water hydrology. They found good agreement between CCl_3F data and hydrologic and tritium age data. Brown (2) also examined the use of fluorocarbons as groundwater tracers through soil column studies. There was no significant use of or development in CFC dating techniques during the 1980s, possibly due to the availability of other methods such as the tritium method and the low resolution of the available CFC analytical techniques. From 1992, the application of CFCs in groundwater studies, especially for age dating and determination of rainfall recharge became important (3). In groundwater textbooks, not many contain a section on the CFC age-dating technique. One text that deals briefly with the topic is by Domenico and Schwartz (9). A book chapter by Plummer et al. (10) is a valuable resource that covers most aspects of the CFC age-dating technique.

SAMPLING GROUNDWATER FOR CFCS

In sampling groundwater for CFC analysis, care must be exercised to prevent the sample from contacting air. The average CFC concentration in the air is about 300 times that in groundwater. Therefore, any small contamination from air will have a very large impact on the concentration of CFCs in the groundwater sample. A sophisticated sampling apparatus like that developed by CSIRO (Commonwealth Scientific and Industrial Research Organization), Australia, is necessary to meet this condition (Fig. 2). Prior to sampling, about three volumes of well water are purged. A 6-mm nylon tube and a stainless steel bailer are sent downhole to obtain the sample. Both nylon tube and bailer are initially flushed with ultrahigh purity (UHP) nitrogen to remove any air from the system. The sample containers are 62 cm^3 borosilicate glass ampules that are flushed, prior to filling, with UHP nitrogen for two minutes. After filling, the ampules are flame sealed using a propane–oxygen torch. The CFC content of the samples is measured using a purge-and-trap gas chromatograph with an electron capture detector (ECD). An illustrative and practical text that deals with sampling for CFCs is that by Kazemi (11).

In another method of sampling for CFCs, the ground water is pumped via a copper tube for a considerable length of time. While the water is flowing through the tube, a small length of the tube is cut and flame sealed quickly to avoid contamination from air. One other method of sampling involves pumping a large volume of groundwater into a big container and then sampling groundwater from the water inside the container.

HOW TO CALCULATE CFC AGE

Apparent CFC ages are obtained by converting measured CFC concentrations in groundwater to equivalent air concentrations using known solubility relationships (12) and the recharge temperature (13). Therefore, the first step is to measure the CFC concentration in groundwater (usually less than 400 pg/kg). The second step is to determine the air temperature at which the rainfall recharge occurs. This is usually average summer temperature (for regions of summer dominated rainfall), average winter temperature (for regions of winter dominated rainfall), and mean annual air temperature (for regions of all year round rainfall). The third step is to calculate the solubility of the CFCs at the temperature considered (summer, winter, or annual) by using the Warner and Weiss (12) relationship. The final step is to use all three above and the following formula to obtain the equivalent atmospheric concentration (EAC) of CFCs:

$$\mathrm{EAC} = \frac{\mathrm{CFCs_{gw}}}{S \times 121} \text{(adapted from Ref. 13)}$$

where EAC is the equivalent atmospheric concentration, $\mathrm{CFCs_{gw}}$ is the concentration of CFCs in the groundwater, S is solubility, and 121 is the molecular weight of CFCs in units of g/mol. The EAC is then compared to the graphs of CFC concentrations in the atmosphere (Fig.1), and the year when the recharge occurred is determined.

LIMITATIONS AND POSSIBLE SOURCES OF ERROR IN THE CFC DATING TECHNIQUE

Error in Estimating Recharge Temperature and Excess Air

Temperature plays a major role in the solubility of CFCs in rainwater. If the air temperature at the time of recharge is overestimated, then the solubility of CFCs that are calculated will be greater than real values. This leads to overestimation of groundwater ages. Note that the error in estimating the recharge temperature is minimal for waters recharged before 1985 (3).

Excess air is the air bubbles trapped during recharge. The CFC content of these air bubbles is slowly dissolved in the groundwater, leading to a rise in the CFC concentration in groundwater. This results in underestimating groundwater ages.

Thick Unsaturated Zone (More Than 5 Meters) and the Associated Error

It is assumed in the CFC technique that the concentration of CFCs immediately above the water table is the same as the CFC concentration in air. If there is a thick unsaturated zone and the soil water content is high, then CFCs may be dissolved preferentially into the soil water. Consequently, the CFC concentration above the water table differs from the CFC concentration in air. Furthermore, there will be a lag between the time that rainwater, containing CFCs, leaves the earth surface and the time that it enters the groundwater. This phenomena leads to overestimating the groundwater age.

Contamination During Sampling

There is always a possibility that groundwater could be contaminated during sampling. Apart from air contamination, CFCs may enter the sample through sampling equipment.

Microbial Degradation of CFCs in the Aquifer

Experiments by Oremland et al. (14) have shown that CFCs are degraded by microbial activity especially in the soil zone. However, the magnitude of degradation is low. If CFCs are decomposed to a large extent, then it is difficult to obtain the true age of the groundwater.

Sorption and Desorption of CFCs from the Aquifer Matrix, Diffusion, and dispersion

Russell and Thompson (15) showed that CFCs are adsorbed to and released from soil particles. This has two implications for the CFC dating technique:

1. Rainwater infiltrating through the soil zone adsorbs some CFCs from soil particles (in addition to the CFCs gained from the atmosphere), and hence it is enriched in CFCs.
2. Some of the CFC content of infiltrating rainwater may be adsorbed to soil particles, and hence it is depleted in CFCs.

In the saturated zone, CFCs are adsorbed to the aquifer matrix. Some time later, the adsorbed CFCs may be released to the groundwater system. Diffusion and dispersion of CFCs in groundwater leads to underestimation of CFC concentration in groundwater and hence overestimation of groundwater age.

Mixing of Water of Different Ages

The problem of mixing of groundwaters from different sources and different aquifers that surrounds most of the hydrological calculations affects the CFCs technique, too. Groundwater in fractured rock aquifers is the most problematic because of the large difference between the age of fracture water and the age of matrix water (15). This problem may present the biggest challenge to the CFC dating technique but could be alleviated, to some extent, by comparing CFC ages with hydraulic ages or ages obtained by other techniques.

Uncertainty in Estimating the Atmospheric Input Function

There is always some uncertainty in estimating the historical concentration of CFCs in the atmosphere at a particular site, especially if the site is far from CFC measurement stations. This could lead to underestimation or overestimation of groundwater ages. However, the nearly homogeneous concentration of CFCs in the atmosphere across long distances reduces the uncertainty.

EXAMPLES OF THE APPLICATION OF CFCS IN GROUNDWATER STUDIES

Reilly et al. used CFC and tritium dated ages as well as a numerical simulation technique to quantify the groundwater flow rate, recharge ages of shallow groundwater, and mixing properties of the groundwater system in Locust Grove, Maryland, U.S.A. They showed that the two methods together (the age-dating method and the simulation technique) enabled a coherent explanation of the flow paths and rates of movement; at the same time, they indicated weaknesses in understanding the system. They also showed that the numerical simulations in the absence of the environmental tracer information provided nonunique and therefore, uncertain estimates of the quantities of water and flow paths within the groundwater system.

Bohlke and Denver (18) used CFC age dating along with chemical and isotopic techniques to resolve the history and fate of nitrate contamination in two agricultural watersheds of the Atlantic Coastal Plain, Maryland, USA.

Cook and Solomon (19) showed that only the thick unsaturated zone would have a major influence on the result of CFC dating. Cook et al. (20) showed that the problem of excess air and degradation of CFCs are two factors that may negatively affect CFC dating results. They also concluded that

1. CFC-11 appears to be degraded both in the highly organic unsaturated zone and below the 3.5-m depth in the aquifer.
2. Degradation of CFC-11 resulted in apparent ages that greatly overestimate groundwater travel times at all depths.
3. There is no evidence for degradation of either CFC-12 or CFC-113, whereas CFC-113 shows evidence of retardation. It was, therefore, decided that the best age tracer that matches hydraulic ages is CFC-12.

Cook et al. (21) used CFCs age data in County Musgrave on Eyre Peninsula and the Clare Valley, both in South Australia, to determine the recharge rate to the shallow groundwater. The results showed that the recharge rate to the aquifer was about 20–50 mm per year.

Szabo et al. (22) dated the shallow groundwater in the southern New Jersey coastal plain using CFCs and $^3H/^3He$. They found an excellent correlation between ages obtained from the two methods and also with those obtained by a groundwater flow model.

Cook and Solomon (8) compared the results of CFC, $^3H/^3He$, and ^{85}Kr measurements to date young groundwater, to estimate the groundwater flow path, and to determine vertical and horizontal flow velocities. They noted that in most cases the results from the three methods are in good agreement. However, where there is a difference between the ages, the difference could be used to

1. obtain other information on flow processes;
2. determine the mixing of water of different ages; and
3. identify a fracture flow system because of the different rates at which these tracers diffuse between fractures and matrix.

The residence times of groundwater seepage to streams in the New Jersey coastal plain were determined by Modica et al. (23) using CFC data. They found a negative correlation between the age of groundwater and its nitrate concentration.

CFC groundwater ages proved to be a helpful tool for managing salinized catchments in Eastern Australia (24).

Having measured the CFC ages of groundwater from a number of aquifers in Eastern Australia, Kazemi (11) showed that the groundwater does not become older along the flow line. He also showed a negative correlation between the age of groundwater and the maximum fluctuation in the water table.

MacDonald et al. (25) used CFC concentration in the groundwater of the Permian Aquifer in Dumfries, Scotland, to identify trends in groundwater quality. They clearly demonstrated a positive correlation between the percentage of modern water and the concentration of nitrate in the aquifer; the higher the percentage of modern water in the aquifer, the higher its nitrate concentration.

WORLDWIDE CFC MEASUREMENT LABORATORIES AND COSTS

As of 1999, the following organizations were able to measure CFC concentration in groundwater (11):

1. The Center for Groundwater Studies (CGS), Adelaide, Australia
2. United States Geological Survey, Reston, Virginia
3. Institute fur Umveltphysik, University of Heidelberg, Germany
4. Geological Survey, Switzerland
5. Institute of Geological and Nuclear Studies, New Zealand
6. University of Waterloo, Canada.

The cost of analyzing groundwater for CFCs was about US $100 per sample as of 1999. The total cost of a single analysis is about double considering the expenses of (1) the labor, (2) the ultra high purity (UHP) C size nitrogen gas cylinder, (3) the oxygen and fuel (for sealing ampules), and (4) the shipment of samples to laboratories and hiring of sampling equipment (11).

BIBLIOGRAPHY

1. Thompson, G.M., Hayes, J.M., and Davis, S.N. (1974). Fluorocarbon tracers in hydrology. *Geophys. Res. Lett.* **1** (4): 177–180.
2. Brown, J.D. (1980). *Evaluation of Fluorocarbon Compounds as Groundwater Tracers: Soil Column Studies*. M.Sc Thesis, University of Arizona, Tucson, AZ.
3. Busenberg, E. and Plummer, L.N. (1992). Use of chlorofluorocarbons (CCl_3F and CCl_2F2 as hydrologic tracers and age-dating tools: The alluvium and terrace system of central Oklahoma. *Water Resour. Res.* **28**(9): 2257–2283.
4. Badr, O., Probert, D.S., and O'Callaghan, P.W. (1990). Chlorofluorocarbons and the environment: Scientific, economic, social, and political issues. *Appl. Energy* **37**: 247–327.
5. Lovelock, J.E. (1971). Atmospheric fluorine compounds as indicators of air movements. *Nature* **230**: 379.
6. Schoeberl, M.R. (1999). Ozone and stratospheric chemistry. In: *EOS Science Plan*. M.D. King (Ed.). NASA, Greenbelt, MD, pp. 309–337.
7. Plummer, L.N. and Busenberg, E. (1999). Chlorofluorocarbons. In: *Environmental Tracers in Subsurface Hydrology*. P.G. Cook and A.L. Herczeg (Eds.). Kluwer, Dordrecht, The Netherlands, pp. 441–478.
8. Cook, P.G. and Solomon, D.K. (1997). Recent advances in dating young groundwater: Chlorofluorocarbons, $^3H/^3He$ and ^{85}Kr. *J. Hydrol.* **191**: 245–265.
9. Domenico, P.A. and Schwartz, F.W. (1998). *Physical and Chemical Hydrogeology*, 2nd Edn. John Wiley & Sons, New York.
10. Plummer, L.N., Michel, R.L., Thurman, E.M., and Glynn, P.D. (1993). Environmental tracers for age dating young groundwater. In: *Regional Ground-Water Quality*. W.M. Alley (Ed.). ITP, New York, pp. 227–254.
11. Kazemi, G.A. (1999). *Groundwater Factors in the Management of Dryland and Stream Salinity in the Upper Macquarie Valley, New South Wales, Australia*. Unpublished Ph.D. Thesis, University of Technology, Sydney, Australia.
12. Warner, M.J. and Weiss, R.F. (1985). Solubilities of chlorofluorocarbons 11 and 12 in water and seawater. *Deep Sea Res.* **32**: 1485–1497.)
13. Cook, P.G. and Herczeg, A.L. (1998). Groundwater chemical methods for recharge studies. Part 2 of the *Basics of Recharge and Discharge*. L. Zhang (Ed.). CSIRO Publishing, Australia.

14. Oremland, R.S., Lonergan, D.J., Culbertson, C.W., and Lovely, D.R. (1996). Microbial degradation of hydrochlorofluorocarbons ($CHCl_2F$ and $CHCl_2CF_3$) in soils and sediments. *Appl. Environ. Microbiol.* 1818–1821.
15. Russell, A.D. and Thompson, G.M. (1983). Mechanisms leading to enrichment of the atmospheric fluorocarbons CCl_3F and CCl_2F_2 in groundwater. *Water Resour. Res.* **19**: 57–60.
16. Cook, P.G. (2003). Groundwater Ages in Fractured Rock Aquifers. In: *Proc. Int. Conf. Groundwater Fractured Rocks*. Prague, Krasny-Hrkal-Bruthans (Eds.). IHP-VI, series on groundwater no. 7, pp. 139–140.
17. Reilly, T.E., Plummer, L.N., Phillips, P.J., and Busenberg, E. (1994) The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer. *Water Resour. Res.* **30** (2): 421–433.
18. Bohlke, J.K. and Denvor, J.M. (1995). Combined use of groundwater dating, chemical and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, Atlantic Coastal Plain, Maryland. *Water Resour. Res.* **31** (9): 2319–2339.
19. Cook, P.G. and Solomon, D.K. (1995). Transport of atmospheric trace gases to the water table: Implications for groundwater dating with chlorofluorocarbons and krypton-85. *Water Resour. Res.* **31** (2): 263–270.
20. Cook, P.G. et al. (1995). Chlorofluorocarbons as tracers of groundwater transport processes in a shallow, silty sand aquifer. *Water Resour. Res.* **31** (3): 425–434.
21. Cook, P.G., Love, A., and Dowie, J. (1996). Recharge estimation of shallow groundwater using CFC age dating. *Miner. Energy South Aust. J.* **3**: 32–33.
22. Szabo, Z. et al. (1996). Age dating of shallow groundwater with chlorofluorocarbons, tritium, helium and flow path analysis, southern New Jersey coastal plain. *Water Resour. Res.* **32** (4): 1023–1038.
23. Modica, E., Buxton, H.T., and Plummer, L.N. (1998). Evaluating the source and residence times of groundwater seepage to streams, New Jersey coastal plain. *Water Resour. Res.* **34**(11): 2797–2810.
24. Kazemi, G.A., Milne-Home, W.A., and Keshwan, M. (1998). Groundwater age dating with CFC in the Macquarie River Basin, NSW, Australia; Implications for dryland salinity management. In: *Proc. IAH 28th Congr. 'Gambling Groundwater'*, Las Vegas, NV, Sep–Oct. 1998, pp. 549–554.
25. MacDonald, A.M., Darling, W.G., Ball, D.F., and Oster, H. (2003). Identifying trends in groundwater quality using residence time indicators: An example from the Permian Aquifer of Dumfries, Scotland. *Hydrogeol. J.* **11** (4): 504–517.

COAGULATION AND FLOCCULATION

KIRSTEN EXALL
National Water Research
Institute
Burlington, Ontario, Canada

INTRODUCTION

Removal of suspended or colloidal solids is an integral component of both drinking water and wastewater treatment. Natural waterbodies from which drinking water is obtained contain a variety of particulates, including clays, silt, organic detritus, and microorganisms. Municipal and industrial wastewaters contain high concentrations of solids, as well as nutrients, organics, and pathogens. Many contaminants of concern to human and environmental health are commonly associated with particles (e.g., certain heavy metals and polynuclear aromatic hydrocarbons, PAHs) or are of colloidal size themselves (e.g., pathogenic microorganisms such as *Cryptosporidia*) (2). Removal of suspended or colloidal particles is therefore necessary to reduce a variety of health hazards.

The addition of chemical coagulants to aid in removing solids and organic matter is a fundamental step in conventional drinking water treatment. In wastewater treatment, coagulants may be added to enhance primary settling (known as chemically enhanced primary treatment or advanced primary treatment) or to improve separation of sludge in the secondary clarifier. Additionally, flocculants may be applied during sludge thickening and dewatering; this application has recently been discussed by Bache (1). This article focuses on chemical coagulant addition and subsequent flocculation processes during drinking water and wastewater treatment; various factors that influence the effectiveness of these processes are discussed.

COAGULATION AND FLOCCULATION

Colloidal Stability

Colloidal particles, as broadly defined, exist in the size range of nanometers to micrometers. Such particles do not readily settle out of suspension in natural water or wastewater due to very low settling velocities and do not aggregate quickly due to electrostatic or steric stabilizing influences. Many colloidal materials in natural water and wastewater have negatively charged surfaces, which lead to the formation of an electrical double layer (the charged surface with adsorbed ions and a more diffuse, surrounding ion cloud). Repulsion between like-charged surfaces of neighboring colloids imparts electrostatic stability to the particles. The degree of stability depends on the magnitude of the surface charge, which can be indirectly measured by the zeta potential of the particle. Further discussion of the origins, impact, and modeling of the electrical double layer and electrostatic stabilization (as they relate to water treatment) can be found in Letterman et al. (2).

Steric stabilization may occur when polymer segments adsorb on a particle's surface, leaving loops and tails extending into solution. Interactions between adsorbed polymers on nearby colloids then lead to repulsion, as mixing or compression of the extended polymer segments increases local concentration of those segments (which may directly repel one another) and reduces the volume available to each segment (a reduction in entropy). Steric stability in natural waters can arise from the adsorption of materials such as humic substances on colloid surfaces (2).

MECHANISMS OF COAGULATION AND FLOCCULATION

Most colloids are not indefinitely stable and aggregate after some time. Coagulation in water treatment uses physicochemical processes to accelerate the kinetics of destabilization (coagulation) and aggregation (flocculation). Coagulation involves adding a chemical agent to destabilize colloids in the water and requires intensive mixing to disperse the coagulant rapidly. Flocculation, on the other hand, the formation of aggregates, or flocs, of the destabilized colloids, requires gentle mixing to allow interparticle collisions to form flocs without shearing the aggregates (2). Separation of the flocs and other particulate matter from the water may then occur by gravity settling (sedimentation), dissolved air flotation, or filtration (3).

Coagulation processes include double layer compression, charge neutralization, sweep coagulation, and complex formation; these processes are described below:

- *Double layer compression* occurs on addition of simple salts (e.g., Na^+) to a colloidal suspension; this mechanism of colloid destabilization is impractical for water and wastewater treatment because the salt concentrations required are high and particle aggregation is slow.
- *Charge neutralization* is a common mechanism of contaminant removal employed in water treatment, whereby cationic (positively charged) coagulant species adsorb to the negatively charged particle surface, neutralizing the surface charge and thereby reducing electrostatic stability. Addition of excess coagulant, or overdosing, can result in the reversal of surface charge and restabilization of the dispersion (2). In the case of polymeric coagulants, neutralization is described by the electrostatic charge patch model, where polymer species adsorb to portions of the colloid surface, resulting in the formation of charge-neutral or positively charged "patches," which cause electrostatic attraction and aggregation of colloidal species (4,5).
- In *sweep coagulation*, voluminous metal hydroxide flocs bring about removal of colloidal contaminants through enmeshment as they move through the suspension. The relationship between floc structure and operating problems during sweep coagulation has been reviewed by Bache et al. (6).
- Dissolved contaminants, such as organic species or phosphate, may also be removed by *formation of insoluble complexes* or by adsorption to flocs (2).

Flocs are aggregates of primary particles bound together by intermolecular forces. Flocculation, or floc growth, occurs through interparticle collisions. Both particle collision frequency and collision efficiency (sticking) influence the formation of flocs; floc strength and mixing intensity, or shear, influence the final floc size (7). Polymer flocculants (described below) can improve flocculation through adsorption and *interparticle bridging*. In this mechanism of flocculation, segments of a polyelectrolyte are bound to one colloid particle, but loops and dangling chains of the polymer extend into solution and can attach to nearby particles, leading to an increase in the size of the agglomerate (4,5). It may be difficult to distinguish between charge neutralization and bridging when using polymers for flocculation, because both processes can take place simultaneously (8).

COAGULANTS AND FLOCCULANTS

Coagulation has long been used as an aid to sedimentation in water and wastewater treatment. Common coagulants include both metal salts and charged polymers (also termed polyelectrolytes); flocculants are typically composed of high molecular mass polymers.

Inorganic Coagulants

When simple inorganic coagulants, such as aluminum sulfate (alum) or ferric chloride, are added to water, the salts dissociate, and the metal ions undergo hydrolysis to form a variety of hydroxoaluminum or hydroxoiron species. In these reactions, the acidity of the metal cation causes deprotonation of waters of hydration. The resultant species may be cationic, anionic, or neutral and can include an insoluble metal hydroxide; the distribution of species depends on the concentration of the metal ion, pH, and other materials in the water.

Successive reactions in a solution of an aluminum salt lead to the formation of the species Al^{3+}, $Al(OH)^{2+}$, $Al(OH)_2{}^+$, $Al(OH)_3$, and $Al(OH)_4{}^-$. Note that for the sake of simplicity, the waters of hydration (e.g., $Al(H_2O)_6{}^{3+}$) are conventionally not shown. The cationic species can destabilize colloids through surface charge neutralization. At higher concentrations or pH, both Al^{3+} and Fe^{3+} are converted to insoluble hydroxides, which form voluminous, amorphous flocs. These flocs are responsible for sweep coagulation and can enmesh colloidal material as they pass through the suspension. Depending on the conditions of hydrolysis, polynuclear aluminum species (those containing more than one aluminum ion) may also be formed. These include dimeric [e.g., $Al_2(OH)_2{}^{4+}$], trimeric [$Al_3(OH)_4{}^{5+}$], and tridecameric ($Al_{13}O_4(OH)_{24}{}^{7+}$) forms (9). Concerns exist over the potential health effects of aluminum in drinking water, so coagulation conditions should be chosen to minimize residual levels of the metal in the finished water (2). Similar reactions occur when using iron salts, and equivalent mononuclear and polynuclear iron species exist (2,9).

Prehydrolyzed Coagulants

Various prehydrolyzed coagulants have been prepared and investigated, including polyaluminum chloride (PACl) (e.g., 10,11), polyaluminum silicate sulfate (PASS) (12), polyaluminum sulfate (PAS, also known as polyaluminum hydroxysulfate, PAHS) (e.g., 13,14), and various iron-based and mixed Fe-Al coagulants (15,16). PACl is the most commonly used of these (9). Prehydrolyzed coagulants are formed by partial neutralization of a simple salt solution, such as aluminum chloride, by a base. The degree of prehydrolysis is expressed as a hydroxide-to-metal ion ratio and can be expressed as a percent basicity (10), where

$$\text{Basicity}\ (\%) = 100 * [OH^-]/(3 * [M_T]) \qquad (1)$$

Prehydrolysis favors the formation of soluble polynuclear species, such as the tridecamer (9,10). The rate of particle destabilization is faster with PACl than with alum, and the prehydrolyzed product, PACl, is less dependent on pH for efficient coagulation. The performance of prehydrolyzed aluminum coagulants is also less affected by low temperatures than that of alum [9, and references therein]. Though effective in organic matter removal, prehydrolyzed coagulants are not superior to alum in this regard (9) and may even display higher sensitivity to organic matter concentration (19). Other reported benefits of prehydrolyzed coagulants have included lower aluminum residuals, reduced sludge volume, less alkalinity consumption, and improved settling efficiency (9,18).

Organic Polymers

Synthetic polymers are commonly used as coagulants or flocculants in water and wastewater treatment (4,20,21). Typical polymers include polyacrylamide (PAM), polydiallyldimethyl ammonium chloride (PolyDADMAC), and epichlorohydrindimethylamine (Epi-DMA). Many polymers are derivatized (chemically altered) to contain charged functional groups, such as quaternary ammonium groups, to improve coagulating characteristics (4). The behavior of a polymer is governed by the nature of its charge, whether cationic (positively charged), anionic (negatively charged), zwitterionic (containing both positively and negatively charged groups), or nonionic. Although anionic and nonionic polymers are effective in removing colloidal contaminants under some conditions and may be used as flocculants in conjunction with inorganic coagulants, the polymers used as primary coagulants in water and wastewater treatment are most often cationic (2,4,20).

Synthetic polymers may be used as either coagulants or flocculants, depending on the molecular character and mode of action on suspended solids. Polymer coagulants typically have a low molecular mass and high charge density and commonly act on colloids to bring about coagulation through charge neutralization. In this case, removal is considered to proceed by the electrostatic charge patch model (discussed above). As with inorganic coagulants, charge reversal and restabilization is possible with polymer coagulants. Polymer flocculants, on the other hand, generally have very high molecular mass and low charge density and typically act through a bridging mechanism. Flocculants are also commonly known as coagulant aids because they are often used in conjunction with inorganic or polymeric coagulants (4,5).

The effectiveness of polymers (particularly the large molecular mass flocculants) can depend strongly on their conformation, which in turn depends on the preparation of polymer solution. Many flocculants are supplied in concentrated, highly entangled forms, often as dispersions or emulsions in mineral oils. The polymer must be dissolved in water and at least partially disentangled, or "activated," for optimal flocculation behavior. The intensity of mixing during dispersion can impact the dissolution process; a balance must be achieved between rapid dispersion of the polymer and rupture of polymer chains under high shear conditions (8). Other factors that may influence the effectiveness of a polymer solution include other ions in the water (8) and the age of a polymer solution (22).

Concerns over the use of synthetic polymers in water and wastewater treatment have been raised with regard to the toxicity of residual trace levels in the effluent and formation of toxic by-products on reaction of residues with disinfectants (e.g., 23,24). Polymers are expected to be largely partitioned into the solid phase, but the amount of residual polymer remaining in the liquid stream depends on the dosage. Careful dosage selection and control is therefore needed. Other concerns exist with respect to contaminant monomer concentrations (e.g., acrylamide in polyacrylamide formulations); many countries regulate the composition and allowed dosages of polymers used in drinking water treatment (2,25,26).

Natural Coagulants

Natural products, including chitosan (a derivative of chitin, found in the shells of crustaceans), starch derivatives, seed extracts of such plants as *Moringa oleifera*, tannin- and lignin-based materials, and alginate have also been studied for use in municipal water or wastewater treatment (4,27). These products have been examined primarily for use in developing countries, but often display many of the advantages of synthetic polymeric coagulants and flocculants, while providing the additional benefits of biodegradability and reduced toxicity (27).

Ballasted Flocculation

Under conditions that would lead to the formation of poorly settling flocs, such as low solids and high organics concentrations, ballasted flocculation may be used. In this technique, additional solids are added with (or shortly after) the coagulant to improve floc formation and settling. Polymers may also be dosed to improve adhesion of the forming flocs to the solids, and the solids may be recovered and recycled through the process (4,28). Various solids have been used as ballast, including bentonite, fly ash, powdered magnetite, and microsand (3,29).

ORGANIC MATTER REMOVAL

Organic species exist in many forms in natural waters and wastewaters and, even when not directly harmful to human health, can be carriers of contaminants such as pesticides and can interfere with treatment processes. Reaction of organic matter (OM) with disinfectants may also lead to the formation of harmful disinfection by-products, including trihalomethanes (2). In drinking water treatment, the term "enhanced coagulation" refers to the practice of optimizing coagulation conditions for removing natural organic matter (NOM), which is typically accomplished through increased coagulant dosage and adjustment of pH (30).

In general, the presence of organic matter increases the required dosage of a metal salt coagulant. Once the coagulant demand is met, turbidity and NOM removal occur simultaneously (31–33). A linear relationship often exists between OM concentration and the coagulant dose required for coagulation. The presence of such a "stoichiometric" relationship between the concentration of humic substances and alum dosage was first described by Hall and Packham (31) and has since been observed with various coagulants and organic species (e.g., 33,34).

The nature of the organic matter present affects its removal by coagulation. Many researchers (e.g., 32,33) have observed preferential removals of hydrophobic, dissolved organic carbon compounds over hydrophilic organic fractions and of higher molecular mass organic matter over lower molecular mass compounds.

Metal ion concentration and pH are important factors governing the mechanism of OM removal by coagulation. Two distinct regions of organic matter or color removal have been observed (32,34,35), one near pH 5.5 and one centered near pH 7. The two domains have been explained in terms of removal mechanisms. The lower pH and aluminum concentration region is thought to represent removal by charge neutralization/precipitation, where cationic aluminum species interact directly with OM to bring about removal (2,32). The higher pH/higher coagulant dosage domain is thought to indicate sweep coagulation, where organic species adsorb directly to the surface of formed $Al(OH)_3$ flocs (32,35). Removal of OM may not occur by a single, simple mechanism, however. It has been suggested that even the OM that is ultimately removed by adsorption to $Al(OH)_3$ surfaces may first be complexed by cationic coagulant species (17,33).

The presence of organic matter also increases the required dosage of polymer coagulants and flocculants, thereby greatly increasing the cost of treatment (4,36). As with metal salts, the more hydrophobic OM fractions are more easily removed by polymer, and increasing polymer charge density improves OM removal (37). Polymer molecular mass also influences organic matter removal, although the effect is inconsistent (e.g., 37,38).

FACTORS AFFECTING EFFICIENCY OF COAGULATION AND FLOCCULATION

Many variables may influence the progression of coagulation and flocculation processes, but the primary factors include pH, temperature, mixing conditions, and other ions.

pH Dependence

The speciation of aluminum and iron in aqueous solution are highly dependent on pH. Under acidic conditions, cationic coagulant species (e.g., $Al(OH)^{2+}$) are dominant; in highly basic solution, the anionic hydroxoaluminum or iron form ($Al(OH)_4{}^-$ or $Fe(OH)_4{}^-$, respectively) predominates. At intermediate pH, both the pH and metal ion concentration affect the speciation of aluminum or iron, in turn influencing the interactions between the coagulant species and colloidal solids or organic matter (2). The optimum pH for solids removal and organic matter removal may differ slightly and depends on the mechanism of coagulation (32,34,35).

To minimize residual aluminum or iron in treated water, it is preferable to conduct coagulation under conditions of minimum metal hydroxide solubility. Aluminum hydroxide exhibits a minimum in solubility at 25 °C at approximately pH 6.3; the minimum solubility of iron at the same temperature is at approximately pH 8 (2). In general, synthetic polymers are much less sensitive to pH than metal salt coagulants (4,20).

Temperature

It has long been known that cold temperatures increase the coagulant dose required and change the nature of the flocs produced during coagulation by simple salts (39). Many reasons for the temperature effects have been suggested, including changes in water viscosity (which could affect mixing efficiency), change in the rate and extent of hydrolysis, and changes in the size and strength of flocs formed at lower temperatures (40,41).

The effect of pH varies in cold waters; the optimum pH value for coagulation and flocculation changes with both coagulant concentration and temperature (39). Kang and Cleasby (40) noted that the value of pK_w is temperature-dependent, and adjusting the pH of the solution so that a constant pOH value was held in cold and warm waters was found partially effective in reducing the impact of low temperature on flocculation kinetics.

Prehydrolyzed coagulants are generally less susceptible to temperature changes than are simple aluminum or iron salts (11,18). Synthetic polymers are also not as sensitive to water temperature as metal salt coagulants, although storage of polymer solutions at elevated temperatures may increase the rate of degradation and lowers the effective polymer concentration (22).

Mixing Effects

Coagulation and flocculation processes require both a rapid mix phase to achieve complete and uniform distribution of the coagulant in the water to be treated and a slow mix phase, which promotes interparticle collisions so that flocs may form and grow. Both mixing phases influence the efficiency of contaminant removal.

Although the intensity of rapid mixing is not particularly important for sweep coagulation processes (where contaminant removal occurs by physical interaction between amorphous metal hydroxide flocs and colloids), charge neutralization with metal salt coagulants requires contact between the rapidly changing cationic aluminum or iron species and the colloids in the water before hydrolysis is complete and the metal hydroxide solid forms. For effective charge neutralization, the coagulant species must therefore be distributed throughout the water sample as quickly as possible. Optimization of the rapid mix parameters improves turbidity removal considerably for a given water sample and lowers the required coagulant dosage (42). To optimize the rapid mixing phase of treatment, the intensity of mixing (often expressed as the velocity gradient, G), duration of mixing, and container geometry should all be considered (44).

Increased agitation accelerates floc formation and aggregation through interparticle collisions; if mixing is too vigorous, however, turbulent shear forces cause the breakup of flocs. Interparticle collisions occur through transport by fluid shear, which is controlled by characteristics of the slow mixing phase of treatment, such as vessel geometry, flow patterns, and the turbulence and intensity of mixing (2). Floc breakup may occur through surface erosion of primary particles or fracture of the relatively fragile floc aggregate and is dependent on the intensity of the slow mix phase. The competing processes of aggregation and floc breakup result in a stable floc size for a given set of formation conditions (7).

As discussed above, mixing intensity is significant in the activation of polymer flocculants (8). It is also an important parameter in the application of polymer coagulants and flocculants. In general, higher intensity mixing (higher energy input) is required for the dispersion of polymers than for metal salts, but care must be taken not to shear the polymer (20,43).

Other ions

A number of aspects of water quality impact coagulation processes, including the alkalinity of the water. At very low alkalinity, small additions of an acidic metal coagulant (such as alum) can depress the pH of the water below that of efficient coagulation; at very high alkalinity, coagulant demand may be extremely high (32).

Some anions, such as fluoride, phosphate and certain organic acids, can strongly complex cationic metal species, inhibiting coagulation by reducing free cations in solution (2). The presence of divalent cations, such as Ca^{2+}, typically broadens the pH range for effective organic matter removal and lowers the coagulant dosage required. This is likely due to the complexation of such cations with anionic organic functional groups, lowering the stability of the organic species and thereby decreasing demand on the coagulant (32,45). Sulfate may enhance coagulation by suppressing charge reversal or by catalyzing the kinetics of precipitation through the formation of outer sphere complexes with aluminum ions (45). The presence of silicate or sulfate during prehydrolysis of aluminum salts may also influence aluminum speciation, thereby altering the coagulant composition and activity (12).

The selection and effectiveness of polymers may also be affected by the ionic strength of the system. The addition of simple salts (such as NaCl) causes a contraction in polymer conformation, reducing the effectiveness of polymer flocculants (5).

The conflicting effects of many variables make prediction of coagulant demand for a given water type difficult, so the required dosage of coagulant or flocculant is typically determined empirically by laboratory jar tests.

SUMMARY

Coagulants and flocculants are commonly used in the destabilization and aggregation of suspended solids during water and wastewater treatment. Both inorganic salts and organic polymers may be used as coagulants, and various synthetic and natural organic polymers are applied as flocculants. Organic matter removal is another common objective of coagulation during drinking water treatment, although it generally requires a higher coagulant dosage. Other factors that influence the effectiveness of coagulants and flocculants include pH, temperature, mixing conditions, and other ions in the water. To maximize the efficiency of contaminant removal while minimizing residuals, careful coagulant or flocculant dose selection and control are essential.

BIBLIOGRAPHY

1. Bache, D.H. (2002). Polymers in alum sludge dewatering: developments and issues. In: *Chemical Water and Wastewater Treatment VII*. H.H. Hahn, E. Hoffmann, and H. Odegaard (Eds.). IWA, Cornwall, pp. 287–296.
2. Letterman, R.D., Amirtharajah, A., and O'Melia, C.R. (1999). Coagulation and flocculation. In: *Water Quality and Treatment: A Handbook of Community Water Supplies*, 5th Edn. R.D. Letterman (Ed.). McGraw-Hill, Toronto, pp. 6.1–6.66.
3. Gregory, R., Zabel, T.F., and Edzwald, J.K. (1999). Sedimentation and flotation. In: *Water Quality and Treatment: A Handbook of Community Water Supplies*, 5th Edn. R.D. Letterman (Ed.). McGraw-Hill, Toronto, pp. 7.1–7.87.
4. Bolto, B.A. (1995). Soluble polymers in water purification. *Prog. Polym. Sci.* **20**: 987–1041.
5. Mortimer, D.A. (1990). Synthetic polyelectrolytes—a review. *Polym. Int.* **25**(1): 29–41.
6. Bache, D.H. et al. (1999). Sweep coagulation: structures, mechanisms and practice. *J. Water SRT: Aqua* **48**(5): 201–210.
7. Gregory, J. and Yukselen, M.A. (2002). Break-up and reformation of flocs formed by hydrolyzing coagulants and polymeric flocculants. In: *Chemical Water and Wastewater Treatment VII*. H.H. Hahn, E. Hoffmann, and H. Odegaard (Eds.). IWA, Cornwall, pp. 29–38.
8. Kim, Y.H. (1990). Importance of flocculant preparation for use in solid/liquid separation. In: *Solid/Liquid Separation: Waste Management and Productivity Enhancement*. H.S. Muralidhara (Ed.). Battelle Press, Columbus, OH, pp. 289–304.
9. Gregory, J. and Duan, J. (2001). Hydrolyzing metal salts as coagulants. *Pure Appl. Chem.* **73**(12): 2017–2026.
10. Pernitsky, D.J. and Edzwald, J.K. (2003). Solubility of polyaluminum coagulants. *J. Water SRT: Aqua* **52**(6): 395–406.
11. Van Benschoten, J.E. and Edzwald, J.K. (1990). Chemical aspects of coagulation using aluminum salts I. Hydrolytic reactions of alum and polyaluminum chloride. *Water Res.* **24**(12): 1519–1526.
12. Boisvert, C. and Jolicoeur, J.P. (1999). Influences of sulfate and/or silicate present in partially prehydrolyzed Al(III) flocculants on Al(III) speciation in diluted solutions. *Colloid Surf. A: Physicochem. Eng. Aspects* **155**(2–3): 161–170.
13. Shen, Y-H. and Dempsey, B.A. (1998). Fractions of aluminum coagulants as a function of pH. *Environ. Technol.* **19**(8): 845–850.
14. Koether, M.C., Deutschman, J.E., and vanLoon, G.W. (1997). Low-cost polymeric aluminum coagulant. *J. Environ. Eng.* **123**(9): 859–864.
15. Gray, K.A., Yao, C., and O'Melia, C.R. (1995). Inorganic metal polymers: Preparation and characterization. *J. AWWA* **87**(4): 136–146.
16. Jiang, J. and Graham, N.J.D. (1998). Preliminary evaluation of the performance of new pre-polymerized inorganic coagulants for lowland surface water treatment. *Water Sci. Technol.* **37**(2): 121–128.

17. Edzwald, J.K. (1993). Coagulation in drinking water treatment: Particles, organics and coagulants. *Water Sci. Technol.* **27**(11): 21–35.
18. Lind, C. (1994). Reduce residuals with PACl coagulants. *Public Works* **Dec. 1994**: 43–47.
19. Exall, K. and vanLoon, G.W. (2000). Using coagulants to remove organic matter. *J. AWWA* **92**(11): 93–102.
20. Fettig, J., Ratnaweera, H., and Odegaard, H. (1991). Synthetic organic polymers as primary coagulants in wastewater treatment. *Water Supply* **9**: 19–26.
21. Bolto, B.A. et al. (1996). The use of soluble organic polymers in waste treatment. *Water Sci. Technol.* **34**(9): 117–124.
22. Gehr, R. and Kalluri, R. (1983). The effects of short-term storage and elevated temperatures on the flocculation activity of aqueous polymer solutions. *Water Pollut. Res. J. Can.* **18**: 23–43.
23. Rowland, C.D., Burton, G.A., and Morrison, S.M. (2000). Implication of polymer toxicity in a municipal wastewater effluent. *Environ. Toxicol. Chem.* **19**(8): 2136–2139.
24. Lee, J-F., Liao, P-M., Tseng, D-H., and Wen, P-T. (1998). Behavior of organic polymers in drinking water purification. *Chemosphere* **37**(6): 1045–1061.
25. Dentel, S.K., Chang, L-L., Raudenbush, D.L., Junnier, R.W., and Abu-Orf, M.M. (2000). Analysis and fate of polymers in wastewater treatment. Water Environment Research Federation, WERF Report 94-REM-2.
26. Letterman, R.D. and Pero, R.W. (1990). Contaminants in polyelectrolytes used in water treatment. *J. AWWA* **82**(11): 87–97.
27. Kawamura, S. (1991). Effectiveness of natural polyelectrolytes in water treatment. *J. AWWA* **83**(10): 88–91.
28. Desjardins, C., Koudjonou, B., and Desjardins, R. (1999). Laboratory study of ballasted flocculation. *Water Res.* **36**: 744–754.
29. Ding, Y., Dresnack, R., and Chan, P.C. (2001). Mechanism and feasibility of colloidal particle removal by a microcarrier process. *Int. J. Environ. Stud.* **58**: 535–548.
30. Edzwald, J.K. and Tobiason, J.E. (1999). Enhanced coagulation: US requirements and a broader view. *Water Sci. Technol.* **40**(9): 63–70.
31. Hall, E.S. and Packham, R.F. (1965). Coagulation of organic color with hydrolyzing coagulants. *J. AWWA* **58**(9): 1149–1166.
32. Randtke, S.R. (1988). Organic contaminant removal by coagulation and related process combinations. *J. AWWA* **80**(5): 40–56.
33. Kuo, C-J. and Amy, G.L. (1988). Factors affecting coagulation with aluminum sulfate II. Dissolved organic matter removal. *Water Res.* **22**(7): 863–872.
34. Van Benschoten, J.E. and Edzwald, J.K. (1990). Chemical aspects of coagulation using aluminum salts II. Coagulation of fulvic acid using alum and polyaluminum chloride. *Water Res.* **24**(12): 1527–1535.
35. Dempsey, B.A. (1984). Removal of naturally occurring compounds by coagulation and sedimentation. *CRC Crit. Rev. Environ. Control* **14**(4): 311–331.
36. Narkis, N. and Rebhun, M. (1997). Flocculation in presence of organic macromolecules of natural water and secondary effluents. *Water Sci. Technol.* **36**(4): 85–91.
37. Bolto, B.A. et al. (1999). Experimental evaluation of cationic polyelectrolytes for removing natural organic matter from water. *Water Sci. Technol.* **40**(9): 71–79.
38. Kam, S-K. and Gregory, J. (2001). The interaction of humic substances with cationic polyelectrolytes. *Water Res.* **35**(15): 3557–3566.
39. Camp, T.R., Root, D.A., and Bhoota, B.V. (1940). Effects of temperature on rate of floc formation. *J. AWWA* **32**(11): 1913–1927.
40. Kang, L-S. and Cleasby, J.L. (1995). Temperature effects on flocculation kinetics using Fe(III) coagulant. *J. Environ. Eng.* **121**(12): 893–901.
41. Hanson, A.T. and Cleasby, J.L. (1990). The effects of temperature on turbulent flocculation: Fluid dynamics and chemistry. *J. AWWA* **82**(11): 56–73.
42. Rossini, M., Garrido, J.G., and Galluzzo, M. (1999). Optimization of the coagulation—flocculation treatment: Influence of rapid mix parameters. *Water Res.* **33**(8): 1817–1826.
43. Young, S., Stanley, S.J., and Smith, D.W. (2000). Effect of mixing on the kinetics of polymer-aided flocculation. *J. Water SRT: Aqua* **49**(1): 1–8.
44. Mhaisalkar, V.A., Paramasivam, R., and Bhole, A.G. (1991). Optimizing physical parameters of rapid mix design for coagulation—flocculation of turbid waters. *Water Res.* **25**(1): 43–52.
45. Hundt, T.R. and O'Melia, C.R. (1988). Aluminum—fulvic acid interactions: mechanisms and applications. *J. AWWA* **80**(4): 176–186.

CONDUCTIVITY-ELECTRIC

David M. Gray
Mettler-Toledo Thornton Inc.
Bedford, Massachusetts

Electrical conductivity is the ability of a material to carry electrical current. In water, it is generally used as a measure of mineral or other ionic concentration. Conductivity is a measure of the purity of water or the concentration of ionized chemicals in water. However, conductivity is only a quantitative measurement: it responds to all ionic content and cannot distinguish particular conductive materials in the presence of others. Only ionizable materials will contribute to conductivity; materials such as sugars or oils are not conductive.

In a metal conductor, electrical current is the flow of electrons and is called *electronic* conductance. In water, electrical current is carried by ions because electrons do not pass through water by themselves. This is *electrolytic* conductance.

When a voltage is applied between two inert electrodes immersed in a solution, any ions between them will be attracted by the electrode that has an opposite charge. Ions will move between electrodes and produce a current depending on the electrical resistance of the solution. This is the basis of conductivity measurement—an application of Ohm's law. For high purity waters, it is common to express conductivity as its reciprocal, resistivity.

To prevent altering a sample by major ionic movement and electrochemical reaction at the electrodes, alternating current is always used for measurement. Using ac, the polarity changes frequently enough that ions do not move or react significantly. Measuring systems must control the voltage, frequency, and current density to minimize errors due to electrode polarization and capacitance. Modern instrumentation may change one or more of these variables automatically, depending on the conductivity range being measured.

UNITS OF MEASURE

Conductivity is the conductance of a standardized volume of water of 1 cm^2 cross-sectional area and electrodes spaced 1 cm apart, as shown in Fig. 1. The same results will be obtained with various geometries as long as the ratio of length/area is equal to 1 cm^{-1}, the cell constant. Units of measure are related as listed in Table 1. Most conductivity values are quite small, so unit multiplier prefixes of milli (m, 10^{-3}) and micro (μ, 10^{-6}) are commonly used.

The standardized geometry of the cell constant ensures that the conductivity measurement is a property of the sample and not of the sensor. The cell constant is determined with precision by calibration in standard solutions of known conductivity. Accepted standard potassium chloride solutions are established in an ASTM standard method. Standard reference materials are also available from NIST.

Cell constants other than 1 cm^{-1} may be used as long as the measuring instrument readout is normalized accordingly. A lower cell constant sensor is needed to enable the measuring instrument to measure accurately in low conductivity (high resistivity) samples. Higher cell constants are needed to measure in high conductivity samples. The exact requirements depend on the measuring instrument.

Most practical cells do not use the parallel plate electrode arrangement of Fig. 1. They have greater durability and allow more convenient installation in other arrangements. For example, typical pure water sensors for on-line measurement use concentric electrodes that maintain the spacing and geometry for 0.01 to 0.1/cm constant. A variety of two-electrode process conductivity sensors is illustrated in Fig. 2.

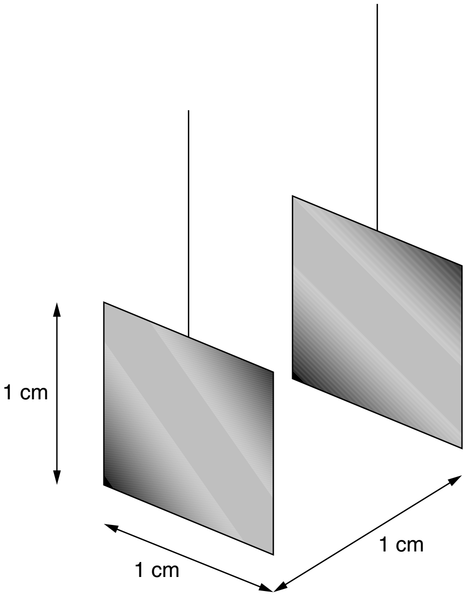

Figure 1. Cell constant = 1 cm electrode spacing divided by 1 cm^2 cross-sectional area of sample.

Table 1. Units Related to Conductivity Measurement

Measurement	Application	Units
Resistance	Electrical circuit	ohm (Ω)
Conductance	Electrical circuit	ohm^{-1} (Ω^{-1}) = siemens (S) = mho (now obsolete)
Resistivity	High purity water	ohm · cm (Ω· cm)
Conductivity	Most water samples	siemens/cm (S/cm) = mhos/cm (now obsolete), siemens/m (S/m)[a]

[a] Most users employ units of S/cm. However, SI conductivity units used in some parts of the world are S/m which can easily be confused. 1 S/cm = 100 S/m.

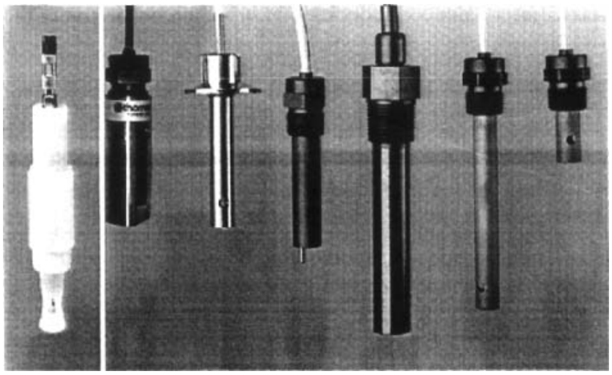

Figure 2. Conventional two-electrode conductivity sensors: 0.1 cm^{-1} in retractable housing, 0.1 cm^{-1} in flow chamber, sanitary 0.1 cm^{-1}, 10 cm^{-1} insertion, 50 cm^{-1} insertion, long 0.1 cm^{-1} insertion, short 0.1 cm^{-1} insertion.

The conductivity of very dilute solutions can be calculated from physical chemistry data based on Equation 1 which sums the conductivity contribution of all ions in the solution:

$$\Lambda = \rho \Sigma(\lambda_i c_i) \tag{1}$$

Λ = conductivity
ρ = density of water
λ_i = equivalent ionic conductance of ion 'i'
c_i = concentration of ion 'i'

TEMPERATURE EFFECTS

Conductivity is affected by temperature because water becomes less viscous and ions can move more easily at higher temperatures. Conventionally, conductivity measurements are referenced to 25 °C, though occasionally a 20 °C reference is used. The variation with temperature is apparent in Equation 1 because λ_i and, to a lesser degree, ρ are temperature dependent. The conductivity of most ions increases by about 2.2% of their value per °C which allows for simple temperature compensation. This is suitable for most midrange conductivity measurements. Very low and very high conductivity samples require special handling of temperature effects.

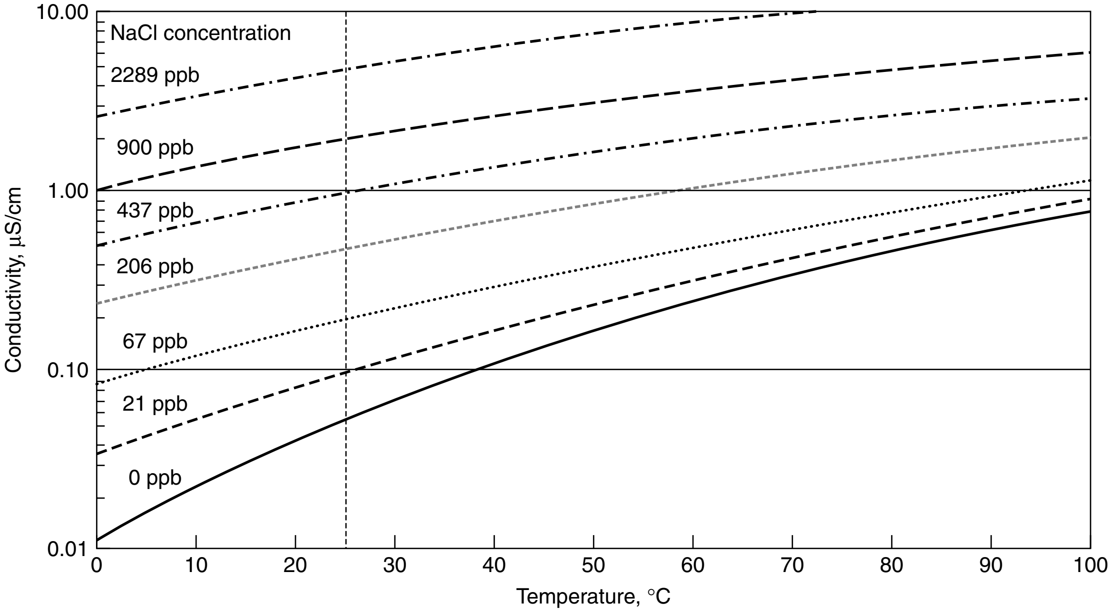

Figure 3. High purity water conductivity vs. temperature.

As water is purified below 5 µS/cm (>0.5 MΩ·cm), the small amount of hydrogen (H^+) and hydroxide (OH^-) ions from water itself become a significant part of total conductivity, that is, the concentrations, c_i, in Eq. 1, of H^+ and OH^- must be included. Water self-ionizes to a much greater degree at higher temperatures, producing a higher concentration of ions. Those ions are also more conductive. The temperature coefficient increases to 4–7%/°C, depending both on purity level and temperature; both effects are nonlinear. These effects are evident in Fig. 3. Specialized high purity temperature compensation must be employed for accurate measurements in this low conductivity range.

Highly concentrated solutions also exhibit deviant behavior. In modest concentration ranges, conductivity is roughly proportional to concentration. At higher concentrations, Eq. 1 does not apply, and ionic interference restricts the mobility of the ions. Conductivity levels off and in many cases decreases as concentration increases, as shown in Fig. 4. Measurement near the peaks of these curves is ambiguous and cannot be relied on for inference of concentration because there are two possible values.

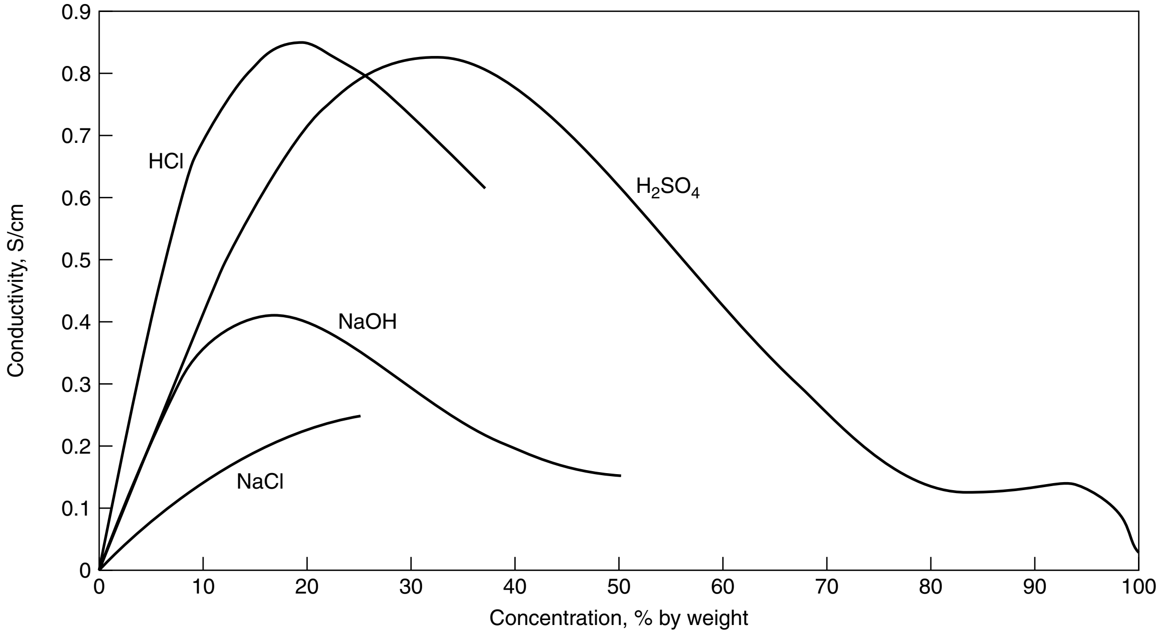

Figure 4. Conductivity vs. concentration at 25°C.

Highly conductive solutions also have a lower temperature influence, typically less than 2%/°C.

ALTERNATIVE MEASUREMENT TECHNOLOGIES

Figure 1 and its description refer to the conventional two-electrode sensor and measurement technique. Four-electrode conductivity measurement uses a sensor incorporating four electrodes. It is useful for highly conductive and/or dirty water samples which would foul the surfaces or plug the narrow passages of conventional high constant two-electrode sensors.

Four-electrode measurement applies ac through the sample via two outer drive electrodes, as shown in Fig. 5. These electrodes may become fouled, and the circuit will compensate to maintain the ac current level constant. Two inner measuring electrodes are used to sense the voltage drop through the portion of solution between them. The circuit makes a high impedance ac voltage measurement, drawing negligible current and making it much less affected by additional resistance due to fouling of the measuring electrode surfaces. Sensors for four-electrode conductivity measurement are shown in Fig. 6.

Inductive (also known as noncontact, electrodeless, or toroidal) conductivity measurement is made without any direct electrical contact with the sample. The sensor consists of two parallel coils sealed within a doughnut-shaped insulated probe, as shown in Figs. 7 and 8. The instrument energizes one coil with ac. A weak ac current is induced in the surrounding sample, depending on its conductivity. That current, in turn, induces a signal in the measured coil which provides the measurement signal. The sample acts like the core of a transformer. A temperature sensor is incorporated into the probe body to enable compensation in the instrument.

Because there are no electrodes in contact with the sample, extreme fouling conditions can be tolerated. Coatings can cause errors only if there is so much accumulation that they reduce the diameter of the hole.

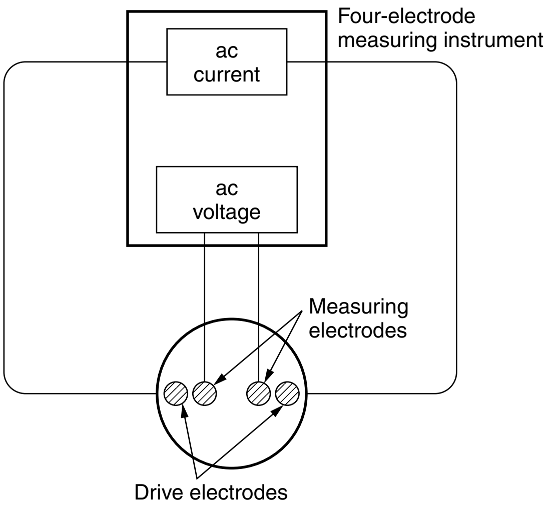

Figure 5. Four-electrode conductivity measurement.

Figure 6. Four-electrode conductivity sensors.

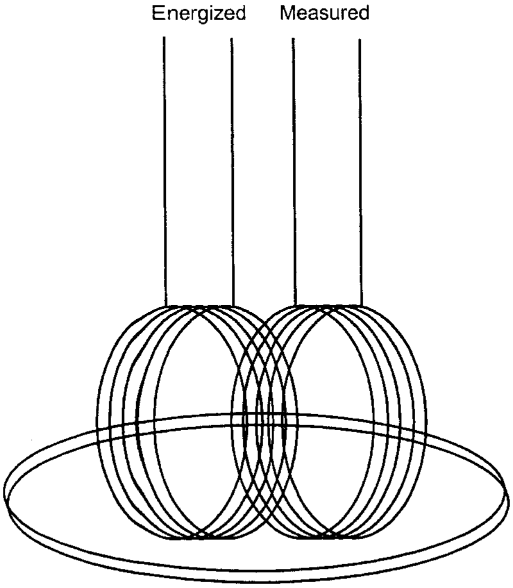

Figure 7. Inductive conductivity measurement.

Care is needed in installating an inductive sensor to allow for the specified spacing around it. Otherwise, the cross section of sample immediately around the probe (cell constant) will be affected, and recalibration may be required.

TOTAL DISSOLVED SOLIDS (TDS)

TDS is sometimes inferred from conductivity and is reported in units of parts per million. However, the relationship of conductivity and concentration is not standardized and to be meaningful, should be specified whenever TDS units are used. Typical conversions are based on sodium chloride (which may also be called salinity) at approximately 0.5 ppm TDS per μS/cm. Alternatively, a "natural water" mineral composition including bicarbonates would have a conversion of 0.6–0.7 ppm TDS per μS/cm. Conversions may also be slightly nonlinear with concentration.

Figure 8. Inductive conductivity measurement equipment.

READING LIST

Standard Test Methods for Electrical Conductivity and Resistivity of Water. D1125, American Society for Testing and Materials, W. Conshohocken, PA.

Certificate of Analysis, Aqueous Electrolytic Conductance Standard Reference Materials 3190-3193, 3198-3199. National Institute of Standards and Technology, U.S. Department of Commerce, Gaithersburg, MD.

Standard Test Method for Electrical Conductivity and Resistivity of a Flowing High Purity Water Sample. D5391, American Society for Testing and Materials, W. Conshohocken, PA.

Weast, R.C. (Ed.). (1978). *CRC Handbook of Chemistry and Physics*, 59th Edn. CRC Press, West Palm Beach, FL, pp. D265–D314.

CONSERVATION AND THE WATER CYCLE

Natural Resources
Conservation Service

Water is probably the natural resource we all know best. All of us have had firsthand experience with it in its many forms—rain, hail, snow, ice, steam, fog, dew.

Yet, in spite of our daily use of it, water is probably the natural resource we least understand. How does water get into the clouds, and what happens to it when it reaches the earth? Why is there sometimes too much and other times too little? And, most important, is there enough for all the plants, and all the animals, and all the people?

Water covers nearly three fourths of the earth; most is sea water. But sea water contains minerals and other substances, including those that make it salty, that are harmful to most land plants and animals. Still it is from the vast salty reservoirs, the seas and oceans, that most of our precipitation comes—no longer salty or mineral laden. Water moves from clouds to land and back to the ocean in a never ending cycle. This is the water cycle, or the hydrologic cycle.

Ocean water evaporates into the atmosphere, leaving impurities behind, and moves across the earth as water vapor. Water in lakes, ponds, rivers, and streams also evaporates and joins the moisture in the atmosphere. Soil, plants, people, and animals, and even factories, automobiles, tractors, and planes, contribute moisture. A small part of this moisture, or water vapor, is visible to us as fog, mist, or clouds. Water vapor condenses and falls to earth as rain, snow, sleet, or hail, depending on region, climate, season, and topography.

Every year about 80,000 cubic miles of water evaporates from oceans and about 15,000 cubic miles from land sources. Since the amounts of water evaporated and precipitated are almost the same, about 95,000 cubic miles of water are moving between earth and sky at all times.

Storms at sea return to the oceans much of the water evaporated from the oceans, so land areas get only about 24,000 cubic miles of water as precipitation. Precipitation on the land averages 26 inches a year, but it is not evenly distributed. Some places get less than 1 inch and others more than 400 inches.

The United States gets about 30 inches a year, or about 4300 billion gallons a day. Total streamflow from surface and underground sources is about 8.5 inches a year, or about 1200 billion gallons a day. This is the amount available for human use—homes, industry, irrigation, recreation.

The difference between precipitation and streamflow -21.5 inches a year, or 3100 billion gallons a day—is the amount returned to the atmosphere as vapor. It is roughly 70 percent of the total water supply. It includes the water used by plants.

People can exist on a gallon or so of water a day for drinking, cooking, and washing though we seldom do or have to. In medieval times people probably used no more than 3 to 5 gallons a day. In the 18th century, especially in Western nations, people were using about 95 gallons a day. At present in the United States, people use about 1500 gallons a day for their needs and comforts including recreation, cooling, food production, and industrial supply.

When water hits the ground some soaks into the soil, and the rest runs off over the surface. The water that soaks into the soil sustains plant and animal life in the soil. Some seeps to underground reservoirs. Almost all of this water eventually enters the cycle once more.

People can alter the water cycle but little, so their primary supply of water is firmly fixed. But we can manage and conserve water as it becomes available—when it falls on the land. If we fail to do so we lose the values that water has when used wisely.

Water management begins with soil management. Because our water supply comes to us as precipitation falling on the land, the fate of each drop of rain, each snowflake, each hailstone depends largely on where it falls—on the kind of soil and its cover.

A rainstorm or a heavy shower on bare soil loosens soil particles, and runoff—the water that does not soak into the soil—carries these particles away. This action, soil erosion by water, repeated many times ruins land for most uses. Erosion, furthermore, is the source of sediment that fills streams, pollutes water, kills aquatic life, and shortens the useful life of dams and reservoirs.

Falling rain erodes any raw earth surface. Bare, plowed farmland, cleared areas going into housing developments, and highway fills and banks are especially vulnerable.

In cities and suburbs, where much of the land is paved or covered—streets, buildings, shopping centers, airport runways—rainwater runs off as much as 10 times faster than on unpaved land. Since this water cannot soak into the soil, it flows rapidly down storm drains or through sewer systems, contributing to floods and often carrying debris and other pollutants to streams.

Grass, trees, bushes, shrubs, and even weeds help break the force of raindrops and hold the soil in place. Where cultivated crops are grown, plowing and planting on the contour, terraces, and grassed waterways to carry surplus water from the fields are some of the conservation measures that slow running water. Stubble mulching protects the soil when it has no growing cover. Small dams on upper tributaries in a watershed help control runoff and help solve problems of too much water one time and not enough another time.

Throughout the world the need for water continues to increase. Population growth brings demands for more water. Per capita use of water, especially in industrialized countries, is increasing rapidly.

It is our management of the precipitation available to us that determines whether or not we have both the quantity and the quality of water to meet our needs.

It is our obligation to return water to streams, lakes, and oceans as clean as possible and with the least waste.

THE WATER CYCLE is an endless process of water circulation going on throughout the world.

To trace the movement of water through the cycle, begin at the far right of the diagram. There the sun's energy is transferring water from the sea and earth to the atmosphere in the form of water vapor. The soil and inland water bodies through *evaporation* and plants through *transpiration* add large amounts of water vapor to the atmosphere, but most of it comes from the oceans. Man, animals, and machines add small amounts by means of *respiration* and *combustion*.

Air masses (top of diagram) carry the water vapor across the earth, and the water vapor condenses into *precipitation*.

At the left, precipitation falls as rain, snow, sleet, hail. Some evaporates while falling and returns to the atmosphere. A small amount is intercepted and held by plants or by buildings, automobiles, and other structures and machines until it evaporates back into the atmosphere.

Most of the precipitation soaks into the soil; the part that doesn't run to the sea by way of streams and rivers. Ground water gets there more slowly.

Misuse and poor management of the soil will decrease the amount of water that soaks into the soil and increase the amount that runs off over the surface. Runoff on bare land leads to erosion. Grass, trees, and other plants hold the soil in place and slow the runoff, allowing more water to soak into the soil.

Some of the water that soaks into the soil is used by plants. Part of it percolates beyond the reach of plant roots to the water table, to underground reservoirs, and to springs and artesian wells.

Runoff on its way to the sea can be intercepted and stored for industrial or household use, and it can be diverted for irrigation.

Little water has been added or lost through the ages. The water cycle prevails in all places and at all times with neither beginning nor end.

DEFLUROIDATION

Eli Dahi
Environmental Development
Corporation
Søborg, Denmark

DEFINITION

Defluoridation refers to methods of water treatment that reduce the concentration of fluoride in the water, normally in order to make it safe for human consumption. Some water treatments that have the capacity of reducing the fluoride concentration along with most other anions, or anions and cations, in the water are not considered defluoridation methods. Thus, general demineralizing methods like distillation, reverse osmosis, electrodialysis, and resin deanionization, which are able to remove fluoride fully or partly from the water, are not considered defluoridation methods. On the other hand, methods that only remove fluoride without any addition or reduction of other parameters are not yet discovered, which is why the expression "fluoride removal" lacks precision. Defluoridation is used to characterize methods that reduce the fluoride ion specifically, without other ***major*** changes to the quality of the treated water.

OBJECTIVES AND EFFECTS

The objective of defluoridation is to prevent or to mitigate ***endemic fluorosis***. The term endemic fluorosis covers a wide variety of clinical manifestations of accumulative fluoride poisoning, mainly because of long-term consumption of relatively high concentration fluoride water. In mild cases, fluorosis is observed as a more or less repulsive mottling of teeth, ***dental fluorosis*** (Fig. 1). More severe cases are known as ***skeletal fluorosis*** (Fig. 2), including deformities like knock knees (genu varum), bowleg (genu valgum), antero-posterior bowing of tibia, humpback (kyphosis), exostosis, and others.

Also, ***nonskeletal fluorosis*** includes swollen joints, muscular tenderness, rheumatism-like pain especially in the joints, systematic rigidity, stiffness, and impaired movement, all symptoms that give a feeling of rapid aging. Also, gastric problems, burning sensation during urination, mental retardation, and paraplegia are reported as clinical manifestation of fluorosis.

SAFE CONCENTRATIONS

According to WHO, the “optimum” and the maximum water concentrations are **0.8 and 1.5 mgF/L**, respectively. These guidelines are adopted on the background of studies carried out mainly in low fluoride areas and in the decades before the widespread fluoridation of toothpaste. Presently, in a fluorotic area, most professionals consider **0.5 and 1.0 mg/L** as the desired and maximum allowable concentrations, respectively. In a defluoridation process, the aim is to treat down to 0.5 mg/L, whereas 1 or 1.5 mg/L can be tolerated as the saturation level. Experience, as well as mass balances, show that once the drinking and cooking water is within these limits, the users do not need to worry about the daily intake of fluoride from other sources like food and normal use of toothpaste, even when residing in a fluorotic area.

Figure 1. Dental fluorosis.

Figure 2. Skeletal fluorosis. (**a**) Stage I: swollen joints, pain, and rigidity. (**b**) Stage II: deformed but still moving. (**c**) crippled, only moving short and with difficulties and pain.

FLUORIDE OCCURRENCE

Fluoride occurs in all natural waters. Ocean waters contain about 1 mg/L, whereas most fresh surface and subsurface waters contain less than 0.5 mg/L. Groundwater is, in general, subject to more variation at higher range. In fluorotic areas, groundwater and/or surface waters may contain up to 20–30 mg/L. As an extreme case, the concentration of 90 mg/L has been recorded in one Rift Valley Lake, where the water is soft but alkaline and the evaporation is high. Fluoride-rich waters that are subject to defluoridation for human consumption are in the range 1.5–20 mg/L.

APPROPRIATE TECHNOLOGY

Defluoridation differs from most other water treatments in four main aspects:

First, because fluoride does not deteriorate the piped water quality technically or oganoleptically, it is normally only required for the small part of the water that is used for drinking and cooking.

Second, defluoridation of large quantities of water, out of which only a small part is used for human consumption, is environmentally unsound, which is because of accumulation of correspondingly large quantities of toxic sludge that would create a new problem, more severe to deal with.

Third, most large cities and towns do not need defluoridation, because they are historically built in sites privileged to have good water resources.

Fourth, defluoridation is easily avoided in industrialized countries, as they have no difficulties in prospecting and transporting low fluoride water.

APPROPRIATE CRITERIA

As a result of the above-mentioned four factors, defluoridation is often required in small scale and in rural areas in developing countries. As such, many defluoridation approaches have been launched, proclaiming success without unbiased field proof and thorough optimization. Further confusion comes with the fact that the process or the technical setup that may work in one context of socio-economic and environmental conditions may fail in another. Local availability and acceptability of the required materials, fluoride contamination level, and water quality are major factors to be considered when selecting the process and the design that minimizes the capital and running costs.

THREE OPERATIONAL LEVELS

Defluoridation technology has to be simple, affordable, reliable, and operational at least at three different levels:

1. Domestic level, e.g., in the kitchen where piped water may not be available, see Fig. 3.
2. Institutional level, e.g., in schools or a working place where piped water may be available and the defluoridator is protected, see Fig. 4.
3. Community level, e.g., in a village or a market where the defluoridator is supplied with water often

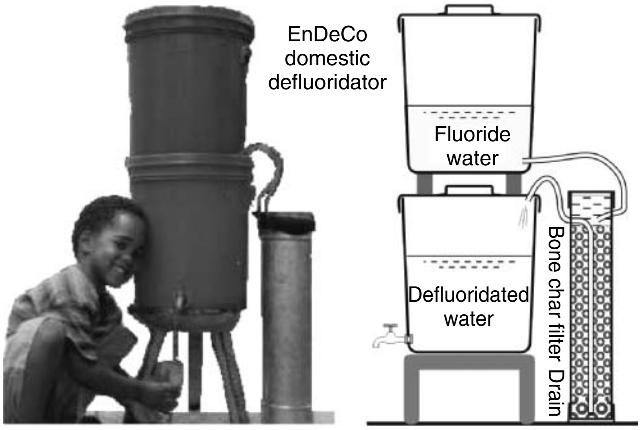

Figure 3. An example of field-tested and optimized domestic defluoridator. It can treat 3000 L of water containing 5 mgF/L, suitable for one family and one-half yearly recharge.

intermittently through a pipe or a hand or power pump, see Fig. 5.

TREATMENT PRINCIPLES

Defluoridation is carried out in one of three different processes:

1. A sorption process in a plug flow filter column with a medium that has a certain capacity of absorption, adsorption, or ion exchange of the fluoride. This process requires recharge or regeneration of the medium upon saturation. Numerous media are known to have defluoridation properties, see Table 1. However, because of capacity, availability, and subsequent water quality limitations, only bone char and activated alumina are worth mentioning. The plants shown in Figs. 1–3 are based on bone char sorption, which is the process of choice if acceptable

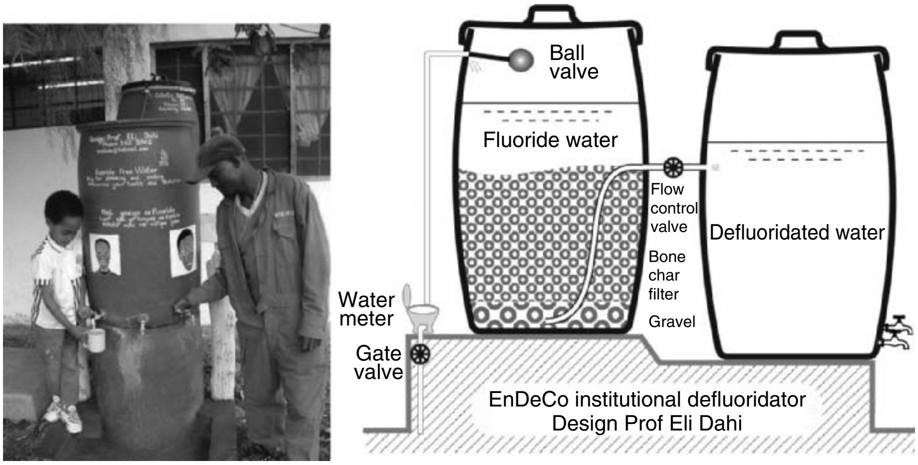

Figure 4. The EnDeCo Institutional Defluoridator. It can be made at any desired capacity, normally in the range 20–160 m^3 for a raw water fluoride of 5 mg/L.

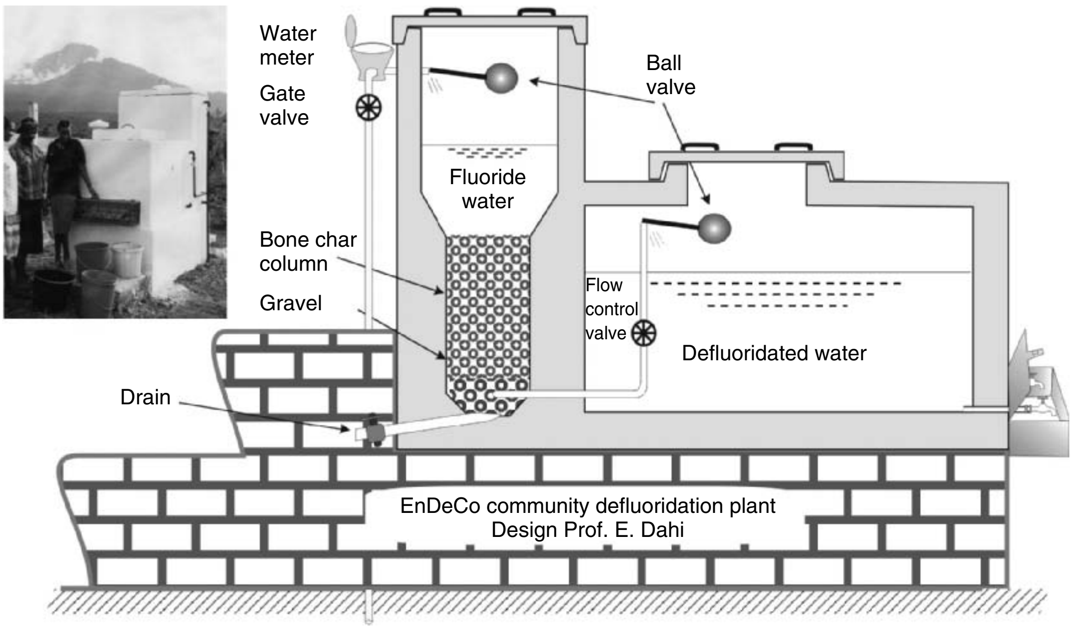

Figure 5. An example of a community defluoridator that has been operating continuously since 1994.

Table 1. Some Natural Media Capable to Sorb Fluoride

Magnesite	Apophyllite	Natrolite
Stilbite	Clinoptilolite	Gibbsite
Goethite	Kaolinite	Halloysite
Bentonite	Verimiculite	Zeolite(s)
Serpentine	Alkalkine soil	Aiken soil
Acidic clay	Kaolinitic clay	China clay
Fuller's earth	Diatomaceous earth	Ando soil

to the community. Alternatively, the same setups can be used in the activated alumina process, which is often preferred, e.g., in strictly vegetarian societies that consider the use of (cow) bone char as unethical.

2. A coprecipitation process in batch containers, where the precipitating chemicals are totally mixed with the raw water. This process requires filling of water and chemicals, mixing and settling, and subsequent withdrawal of the treated water and the produced sludge. In the so-called Nalgonda technique, a high dosage of alum and lime are used in a coagulation/sedimentation process, where the precipitation of aluminium hydroxide instantly binds a part of the water fluoride. A domestic approach for the Nalgonda process is given in Fig. 6.
3. Contact precipitation is a process where the water is mixed with the precipitating chemicals, calcium and phosphate ions, flowing into a catalytic filter column. So far, this process is only known to operate in a fluoride-saturated bone char, being the catalyst, dosed with calcium chloride and sodium dihydrogen phosphate. Contact precipitation is considered to be the process of choice in the future, when the local availability of bone/bone char reaches its limit.

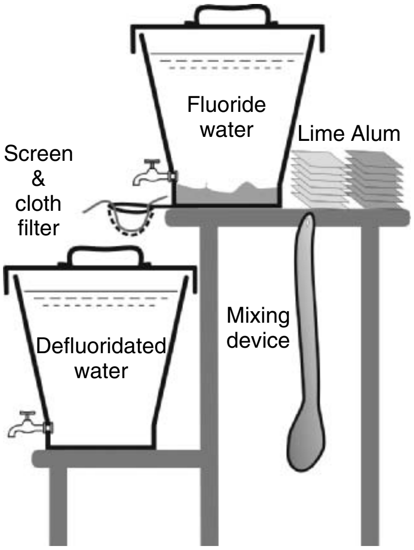

Figure 6. The Nalgonda process as optimized for use in the Rift Valley.

SORPTION DESIGN

The sorption process may be designed on the basis of daily water demand, capacity of the medium, the raw water contamination level, proposed length of operation period, and the bulk density of the medium. The required column treatment capacity and the amount of medium required in a filter column are calculated (Eqs. 1 and 2). From the bulk volume of the medium (Eq. 3), the dimensions of the column are derived (Eqs. 4 and 5):

$$\kappa = D \cdot P \tag{1}$$

$$M = \frac{D \cdot P \cdot (F_r - F_t)}{1000 \cdot C} \tag{2}$$

$$V_m = \frac{M}{\sigma} \tag{3}$$

$$d = \sqrt{\frac{4000 \cdot V_m}{\pi \cdot h}} \tag{4}$$

$$H = 1.2 \cdot h \tag{5}$$

where D is daily water demand, L/day.
C is capacity of the medium, g/kg.
F_r is raw water contamination level, mgF/L.
F_t is treated water concentration, (0.5) mgF/L.
P is proposed length of operation period, days
σ is bulk density of the medium, kg/L.
κ is required column capacity, L.
M is amount of medium required in a filter column, kg.
V_m is bulk volume of the medium, L.
d is inner diameter of the filter column, cm.
h is height of medium in the column, cm.
H is height of the column, cm.

Unfortunately, the capacity of the medium, and its purity, is subject to batch variation. Further, it depends on the fluoride concentration in the raw water. In practice, however, the capacities of 4, 1, and 0.03 g/kg are for good qualities of bone char, activated alumina, and activated clay, respectively.

NALGONDA PROCESS DESIGN

In the Nalgonda and similar processes, the container used has to have a volume well over the daily usage, which will allow for comfortable mixing in the container as well as include provision for the sludge water loss. A plastic bucket of 20 L is often sufficient for the daily use of a family. About 18 L can be treated at a time, out of which about 16 can be used.

The Nalgonda process design is complicated and must be confirmed empirically, which is mainly because of the coprecipitation that lacks stoichiometry, the great variation in media quality and solubility, and the efficiency being dependent on the raw water quality, in particular the fluoride concentration, pH, and the alkalinity. The best possible estimation is based on the Freundlich equation:

$$m = \frac{V_b \cdot (F_i - F_o)}{\alpha \cdot F_o^{1/\beta}} \tag{6}$$

where m is amount of alum required for a daily treatment of a batch, in g.
F_r is fluoride concentration in the raw water, in mg/L.
F_t is residual fluoride concentration in the treated water, in mg/L.
V_b is volume of water to be treated in batch, L.
α is sorption capacity constant, $L^{(1-1/\beta)}$ $mg^{2/\beta}g^{-1}$.
β is sorption intensity constant.

Any resulting pH between 6.2 and 7.6 is adequate. For pH $= 6.7$ and required residual fluoride between 1 and 1.5 mg/L, $\alpha = 6$ and $\beta = 1.33$. The amount of lime required may be 20–50% of the alum dosage.

Alum and lime are added simultaneously to the raw water bucket where it is dissolved/suspended by stirring. The operator should stir fast while counting to 60, i.e., about 1 minute, and then slowly while counting to 300, i.e., about 5 minutes. The mixture is left for settling for about one hour. The treated water is then tapped through the cloth into the treated water bucket from where it is stored for daily drinking and cooking. This separation of the water from the sludge is essential in order to avoid escape of the toxic aluminium to the treated water and in order to avoid the detachment of removed fluoride from the aluminium hydroxide flocs during storage.

TESTING OR METERING

The coprecipitation technique, although arbitrary in nature, is reliable in the sense that the same removal is obtained if the procedure followed exactly.

On the other hand, the process results in poor removal efficiency, and it demands burdensome daily preparation.

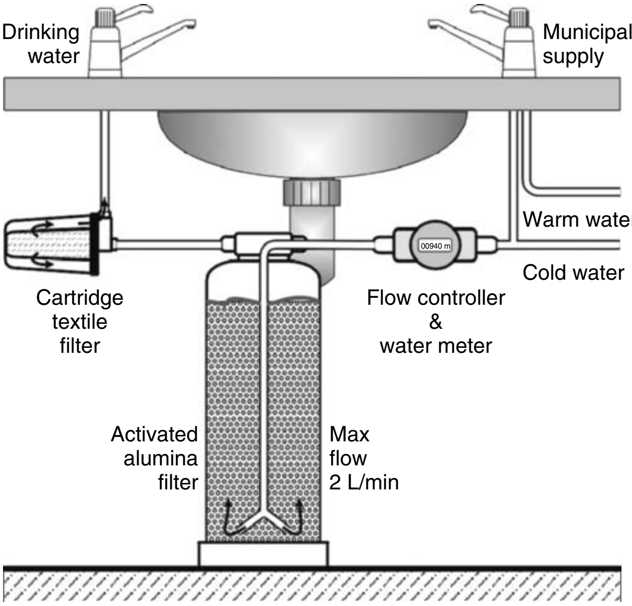

Figure 7. A safe and user-friendly setup of defluoridation based on sorption, in case of activated alumin. The system, containing 8 kg AA, is commercially available in Scandinavia and declared to be able to treat 2000 L of water containing 5 mg/L.

The sorption process is less burdensome. It is also able to provide high removal efficiencies, even at high fluoride levels in the raw water. However, the operator has to watch for the break through the column at saturation. As fluoride is oganoleptically neutral, this cannot be done without either regular testing of the treated water fluoride or checking the accumulated water flow.

Testing the fluoride concentration in the treated water was impossible without expensive laboratory equipment. Now, simple, reliable, and cheap test kits are available.

Figure 7 shows a user-friendly setup for household defluoridation. The secondary cartridge filter ensures that no toxic aluminium can escape with the treated water, and the water meter ensures that accumulated water flow can be compared with declared capacity.

READING LIST

E. Dahi et al. (Eds.). (1995, 1997, 2000, 2004). *Proceedings of 1st–4th International Workshops on Fluorosis Prevention and Defluoridation of Water*. International Society of Fluoride Research.

DEUTERIUM

Gholam A. Kazemi
Shahrood University of
Technology
Shahrood, Iran

INTRODUCTION

The nucleus, the chief ingredient of the atom, consists of two main types of particles: protons (positive charge) and neutrons (neutral). The number of neutrons is either equal or slightly greater than the number of protons. Some elements have only one type of atom, but most have several types. Isotopes of an element have the same number of protons in the atomic nucleus but a different number of neutrons; this results in the same atomic number but different atomic weight. The average atomic weights of the elements have been determined from two sets of data: the atomic weight of each isotope and the relative abundances of the isotopes (1). Isotopes with a higher number of neutrons are commonly referred to as heavy isotopes of that element.

HYDROGEN ISOTOPES

The hydrogen atom is believed to have been created during the “Big Bang” by proton–electron combinations (2). Together with an oxygen atom, it forms the water molecule and has three isotopes: ^1H (common hydrogen), ^2H (deuterium, D), and ^3H (tritium, T). Deuterium, discovered by Urey et al. (3) in 1932, is much heavier than ^1H. Common hydrogen has one proton, deuterium has one proton and one neutron, and tritium has one proton and two neutrons. Deuterium (also referred to as the heavy stable isotope of hydrogen) and common hydrogen are stable, but tritium is radioactive. In seawater, the relative abundances of hydrogen isotopes are H (99.984%), D (0.016%), and T (5×10^{-5}%) (1). In the hydrosphere, there is only one deuterium atom per 6700 H atoms (4). Note that the arena of this article is limited to deuterium as it is a part of the water molecule. Hence, deuterium that may exist in other materials such as methane will not be dealt with here.

ANALYZING WATER SAMPLES FOR DEUTERIUM AND REPORTING THE RESULTS

To analyze water samples for deuterium, first the sample is reduced to H_2. There are a variety of techniques to achieve this; most have been reviewed by Vaughn et al. (5). A common method is the use of zinc in quartz ampoules at 700 °C. The analysis is then carried out using a mass spectrometer such as a Finnigan MAT 250 triple collector mass spectrometer (6). The precision attainable in many laboratories in sample preparation and analysis of water samples for deuterium is 1‰ (7). The subsidized cost of analysis for deuterium as of 1999 is about A$35 per sample (approximately US$25) (8).

An automated system for high precision deuterium analysis on large numbers of water samples was introduced by Vaughn et al. (5) at the Institute of Arctic and Alpine Research at the University of Colorado. This system does not require sample preparation and is capable of extended unattended operation, an attractive option when analyzing thousands of samples during a weekend.

The results of deuterium analysis are expressed in delta (δ) units per mil (parts per thousand or ‰) differences relative to an arbitrary standard known as standard mean ocean water (SMOW):

$$\delta‰ = [(R - R_{\text{standard}})/R_{\text{standard}})] \times 1000$$

where R is the isotopic ratio ($^2H/^1H$) of the sample and R_{standard} is the isotopic ratio ($^2H/^1H$) of the standard (9). Negative δD values represent water that has less than SMOW δD (i.e., they are isotopically depleted), and positive values represent water that has more than SMOW δD (i.e., they are isotopically enriched). In this system, therefore, δD of SMOW is zero. However, a more recent unit to express these isotopes has been proposed (10). In this system, the hydrogen isotopic ratios are reported relative to VSMOW (Vienna standard mean ocean water) on a normalized scale, so that the (δ^2H) of SLAP (standard light Antarctic precipitation) relative to VSMOW is −428‰ (10). Therefore, the δ^2H of VSMOW is 0‰.

DEUTERIUM IN THE GLOBAL WATER CYCLE

At the first onset of the water cycle, the deuterium content of the ocean water is increased when water evaporates to form clouds (see Isotope Fractionation article for more information). Thus, water vapor in the atmosphere (in the form of clouds or in any other form) has less deuterium compared to ocean water. The evaporated water (clouds) is then condensed into precipitation to form rivers, surface waters, lakes, and groundwater, all of which is isotopically lighter than ocean water, unless subjected to considerable evaporation. The final destination of surface and ground water that has less deuterium than ocean water is the ocean that was originally enriched in deuterium due to evaporation. Therefore, the deuterium content of ocean water remains constant over time. [However, the early ocean was plausibly deuterium-depleted in comparison with the present-day ocean, see Lecuyer et al. (11)]. Therefore, it is evident that the deuterium content of water in any part of the hydrologic cycle is influenced by a variety of parameters; the most important is evaporation. The factors that influence the deuterium content of atmospheric precipitation are described here (7):

Altitude Effect. Generally the higher the altitude, the more depleted the precipitation (−0.15 to −0.50‰ per 100-m elevation);

Latitude Effect. The deuterium content decreases as latitude increases.

Continental Effect. The deuterium content decreases inland from the coast.

Seasonal Variation. The deuterium content of winter rainfall is less than that of summer rains.

Amount Effect. The greater the rainfall, the more depleted the deuterium.

Apparent Temperature Effect. This effect is similar to the effect of latitude because the higher the latitude, the lower the temperature. Therefore, it is not easy to find the cause of the deuterium difference, whether deuterium depletion is due to the decrease in temperature or due to the increase in latitude. Because of this complexity, Mazor (1) ignored the latitude effect and did not mention it as a factor that influences deuterium content.

Having considered the many factors above, it is now easy to realize that atmospheric precipitation in each part of the world is unique in its deuterium content. Harmon Craig (12) studied the deuterium and ^{18}O content of a large number of precipitation samples worldwide. He drew a line to express the relationship between these two isotopes in a global perspective which is now known as the Craig Line, the World Meteoric Water Line (WMWL), or sometimes, Global Meteoric Water Line (GMWL). The Craig line equation is $\delta D = 8\delta^{18}O + 10$. In this equation, 8 (coefficient) is the slope of the line, and 10 (the intercept) is called the deuterium excess ($d = \delta D - 8\delta^{18}O$). The lower the slope, the higher the effect of evaporation on the water samples studied. A typical evaporation line has a slope of five. The higher the deuterium excess, the higher the effect of evaporative enrichment on the water samples in question. Generally speaking, samples plotted to the right of the Craig line are isotopically enriched, and those plotted to the left are depleted. However, it should be stressed that each region has its own meteoric water line, a local meteoric water line (LMWL), that should be prepared by analyzing the deuterium and ^{18}O content of a number of local atmospheric precipitation samples. The equation of a number of LMWLs for various parts of the world is given by Mazor (1).

APPLICATIONS OF DEUTERIUM AS A TRACER—ISOTOPIC HYDROLOGY

In almost all water related studies, both deuterium and ^{18}O are measured simultaneously. The applications of stable isotopes of oxygen ($\delta^{18}O$) and hydrogen (^2H) in water studies, especially for analysis of flow path and fractionation processes dates back to the early 1950s (9). Undoubtedly, measurement of deuterium in water, whether it is river, lake, ground, rain, or porewater is the fundamental of an important field of knowledge "isotopic hydrology," a technique or science whose many applications have been summarized by Payne (13). The textbook by Mazor (1) and the recent works by Clarke and Fritz (14) and Cook and Herczeg (15) deal extensively with the topic.

To interpret the deuterium data, the most common procedure is to plot δD and $\delta^{18}O$ on a diagram on which the Craig line and, if possible, LMWL are also illustrated. It is then possible to interpret the data, for example drawing an evaporation line, calculating the slope of the line, and determining the deuterium excess. As an example, the results of an isotopic hydrology study by Kazemi (8) on the Buckinbah Creek Watershed in Eastern Australia are presented in Table 1 and Fig. 1. As can be seen from Fig. 1, only two samples show the effect of evaporative enrichment. From Fig. 1, it can also be inferred that the deep groundwater, compared to shallow groundwater, has been subjected generally to less enrichment by evaporation. Note that the approximate positions of summer and winter precipitation samples are also shown in Fig. 1.

The deuterium content of water samples is measured to (1) find the cause of groundwater salinization (16,17); (2) estimate groundwater inflow into wetlands (18,19); (3) separate the streamflow hydrograph into baseflow and an event water component (20,21); (4) calculate surface evaporation via a stable isotope evaporation model (22); (5) study the mechanism, timing, and area of recharge to groundwater (23–26); (6) study the pattern of groundwater circulation and identify the hydrogeologic characteristics of karst groundwater flow (27); (7) identify the origin of groundwater, whether groundwater was recharged in modern time or in the Holocene (28); (8) study the infiltration processes in clayey soils (29); (9) determine groundwater discharge from forests (30); and (10) be used as a tracer in groundwater artificial recharge schemes (31,32).

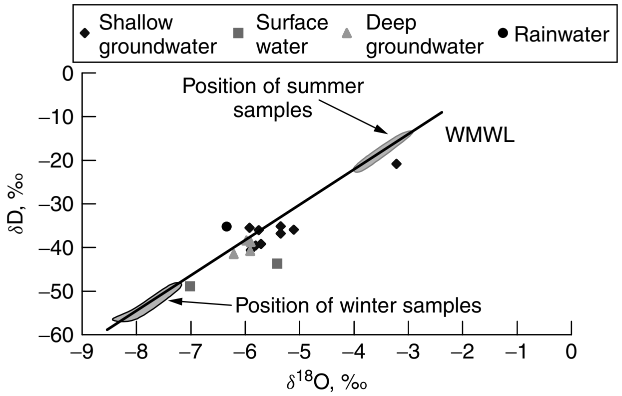

Figure 1. Deuterium versus oxygen-18 plot of water samples from the Buckinbah Creek Watershed, NSW, Australia (8). The Craig (World Meteoric Water Line) is also illustrated.

Case Studies

A large number of case studies are regularly published through the International Atomic Energy Agency (IAEA) and International Association of Hydrological Sciences publications, IAHS (33) and IAEA (34). Examples of the deuterium content of groundwater from the Amman–Zarkan Basin in Jordan are −32 to −26‰ and from the Aladag karstic aquifer in Turkey −100.1 to −58.4‰ (33).

BIBLIOGRAPHY

1. Mazor, E. (1991). *Applied Chemical and Isotopic Groundwater Hydrology*. Open University Press, Milton Keynes.
2. Ottonello, G. (1997). *Principles of Geochemistry*. Columbia University Press, New York.
3. Urey, H.C., Brickwedde, F.G., and Murphy, G.M. (1932). An isotope of hydrogen of mass 2 and its concentration (abstract). *Phys. Rev.* **39**: 864.
4. Drever, J.I. (1997). *The Geochemistry of Natural Waters: Surface and Groundwater Environments*. Prentice-Hall, Englewood Cliffs, NJ.
5. Vaughn, B.H. et al. (1998). An automated system for hydrogen isotope analysis of water. *Chem. Geol.* **152**: 309–319.
6. Jorgensen, N.O. and Heinemeier, J. (2003). Isotope Studies of Saline Groundwater in a Sandstone Aquifer on Bornholm, Denmark. *Proc. Int. Conf. Groundwater Fractured Rocks*, Prague, IHP-VI, series on groundwater no. 7, pp. 205–206.
7. Coplen, T.B. (1993). Uses of environmental isotopes. In: *Regional Ground-Water Quality*. W.M. Alley (Ed.). ITP, New York, pp. 255–293.
8. Kazemi, G.A. (1999). *Groundwater Factors in the Management of Dryland and Stream Salinity in the Upper Macquarie Valley, New South Wales, Australia*. Unpublished Ph.D. Thesis, University of Technology, Sydney, Australia.

Table 1. The Results of Stable Isotopes Analysis of Surface, Ground, and Rainwater Samples in the Buckinbah Creek Watershed, NSW, Australia[a]

Site[b]	Sgw	Sgw	Sgw	Sgw	Sgw	Sgw	Sgw	Sgw	Sgw	Sgw	Sw	Sw	Dgw	Dgw	Dgw	Dgw	Rw
$\delta^{18}O$‰	−5.8	−5.7	−5.8	−5.4	−5.9	−5.4	−5.9	−5.1	−5.9	−3.2	−7	−5.4	−5.9	−6	−5.9	−6.21	−6.34
δD ‰	−40	−39	−36	−37		−35	−36	−36	−41	−21	−49	−44	−41	−38	−38.7	−41.5	−35.2

[a] Reference 7.

[b] Sgw: Shallow groundwater; Dgw: Deep groundwater; Sw: Surface water; Rw: Rainwater.

9. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
10. Hornberger, G.M. (1995). New manuscript guidelines for the reporting of stable hydrogen, carbon, and oxygen isotopic ratio data. *Water Resour. Res.* **31**(12): 2895.
11. Lecuyer, C., Gillet, P., and Robert, F. (1998). The hydrogen isotope composition of seawater and the global water cycle. *Chem. Geol.* **145**: 249–261.
12. Craig, H. (1961). Isotopic variations in meteoric water. *Science* **133**: 1702–1703.
13. Payne, B.R. (1988). The status of isotope hydrology today. *J Hydrol.* **100**: 207–237
14. Clarke, I. and Fritz, P. (1997). *Environmental Isotopes in Hydrogeology*. Lewis Publishers, Boca Raton, FL.
15. Cook, P.G. and Herczeg, A.L. (Eds.). (1999). *Environmental Tracers in Subsurface Hydrology*. Kluwer, Dordrecht, The Netherlands.
16. Stimson, J. et al. (1993). Causes of groundwater salinisation in a low lying area of Cochabamba Valley, Bolivia. In: *Applications of Isotope Techniques in Studying Past and Current Environmental Changes in the Hydrosphere and Atmosphere*. IAEA publication, Vienna, pp. 185–199.
17. Kazemi, G.A. and Milne-Home, W.A. (1999). The application of strontium, oxygen-18 and deuterium isotopes in the estimation of the solute budget and hydrology of a small watershed in Eastern Australia. *Eos Trans. AGU* **80**(17): Spring Meet. Suppl. S139.
18. Hunt, R.J., Krabbenhoft, D.P., and Anderson, M.P. (1996). Groundwater inflow measurements in wetland systems. *Water Resour. Res.* **32**(3): 495–507.
19. Krabbenhoft, D.P., Bowser, C.J., Kendall, C., and Gat, J.R. (1994). Use of oxygen-18 and deuterium to assess the hydrology of groundwater-lake systems. In: *Environmental Chemistry of Lake and Reservoirs*. American Chemical Society, Washington, DC, pp. 66–90.
20. Hooper, R.P. and Shoemaker, C.A. (1986) A comparison of chemical and isotopic hydrograph separation. *Water Resour. Res.* **22**(10): 1444–1454.
21. Gunyakti, A., Altinbilek, H.D., and Qasem, H. (1995). Study of small basin hydrology in Anatolia using environmental isotope analysis. *Hydrolog. Sci. J.* **40**(1): 43–53.
22. Meyers, J.B., Swart, P.K., and Meyers, J.L. (1993). Geochemical evidence for groundwater behavior in an unconfined aquifer, South Florida. *J. Hydrol.* **148**: 249–272.
23. Sami, K. (1992). Recharge mechanisms and geochemical processes in a semi-arid sedimentary basin, Eastern Cape, South Africa. *J. Hydrol.* **139**: 27–48.
24. Leaney, F.W. and Herczeg, A.L. (1995). Regional recharge to a karst aquifer estimated from chemical and isotopic composition of diffuse and localized recharge, South Australia. *J. Hydrol.* **164**: 363–387.
25. Feurdean, L., Feurdean, V., Zamfirescu, F., and Sarbu, S. (2003). Evidence and constraints revealed by environmental stable isotopes in the chemoautotrophic system of the Movile Cave, Romania. *Proc. Int. Conf. Groundwater Fractured Rocks*, Prague, IHP-VI, series on groundwater no. 7, pp. 201–202.
26. Feurdean, L., Feurdean, V., and Oraseanu, I. (2001). The use of deuterium as natural tracer to point out groundwater flow from karst area. Case study. In: *New Approaches Characterizing Groundwater Flow*. K.P. Seiler and S. Wohnlich (Eds.). Swets and Zeitlinger, Lisse, pp. 59–63.
27. Bayari, C.S. and Gunay, G. (1995). Combined use of environmental isotopic and hydrochemical data in differentiation of groundwater flow patterns through the Aladag karstic aquifer, Turkey. In: *Application of Tracers in Arid-Zone Hydrology*. IAHS publication no. 232, pp. 99–117.
28. Guendouz, A. et al. (2003). Hydrogeochemical and isotopic evolution of water in the Complex Terminal aquifer in the Algerian Sahara. *Hydrogeol. J.* **11**: 483–495.
29. Mathieu, R. and Bariac, T. (1996). An isotopic study (^{2}H and ^{18}O) of water movements in clayey soils under a semiarid climate. *Water Resour. Res.* **32**(4): 779–789.
30. Thorburn, P.J., Hatton, T.J., and Walker, G.R. (1993). Combining measurements of transpiration and stable isotopes of water to determine groundwater discharge from forests. *J. Hydrol.* **150**: 563–581.
31. Jorgensen, N.O and Helleberg, B.B. (2002). Stable isotopes (^{2}H and ^{18}O) and chloride as environmental tracers in a study of artificial recharge in Denmark. In: *Management of Aquifer Recharge for Sustainability*. P. Dillon (Ed.). Balkema, Lisse, pp. 245–250.
32. Le Gal La Salle, C. et al. (2002). Isotope contribution to geochemical investigations for aquifer storage and recovery. In: *Management of Aquifer Recharge for Sustainability*. P. Dillon (Ed.). Balkema, Lisse, pp. 265–268.
33. IAHS. (1995). *Application of Tracers in Arid Zone Hydrology*. IAHS Pub No. 232, p. 542.
34. IAEA. (1979). *Isotopes in Lake Studies*. International Atomic Energy Agency, Vienna, Austria.

DISTILLED WATER

Aldo Conti
Frascati (RM), Italy

Water is never pure in nature. It normally contains some salts and minerals in solution and even bacteria and pollutants. Most of these substances, salts and minerals, in particular come from rocks and soils with which the water has been in contact. All these impurities make water useless for some purposes. It must be cleaned of pollutants and bacteria for human consumption, for instance. But for some other uses, even salts and minerals are dangerous. At home, in our vapor iron, water should be free of calcium carbonate that forms a layer on the heating resistor and eventually breaks it. The easiest method of purification is distillation, during which water is heated until it turns into vapor. The vapor is then collected on a cold surface, where it condenses. Because salts and other impurities do not vaporize, they normally remain in the boiling solution. Distilled water is never perfectly pure, because there are substances with nearly the same boiling point as that of water and some liquid droplets can be carried by the steam. Nevertheless, distilled water can be 99.9% pure and, if needed, can be distilled a second time. Distillation is very efficient in removing heavy metals, radionuclides, and solid particles, and it kills bacteria and viruses.

From a chemical point of view, even the purest distilled water is always a solution containing a very small amount of hydroxide (OH^-) and hydronium (H_3O^+) ions, which happens because water, like many other solvents, reacts with itself. At room temperature, one can expect one molecule in every ten million to dissociate and donate a hydrogen atom to another molecule.

Distilled water does not remain pure for long. If left exposed to air, it absorbs carbon dioxide very efficiently and reacts with it producing carbonic acid (H_2CO_3), which is why distilled water often becomes acidic. If one leaves distilled water exposed to air for 24 hours, a pH between 5.5 and 6.0 can be expected. To neutralize this acidity, it is sufficient to boil the water for 5 to 10 minutes, so that it releases the dissolved carbon dioxide, bringing the pH very close to 7.

The acidity induced by carbon dioxide is enough to make distilled water unhealthy for human consumption. Moreover, water contains electrolytes (sodium, potassium, chloride) and many other elements, such as magnesium, that are essential for our body. Most, if not all, of these substances can be found in food, too, but nevertheless, drinking large amounts of distilled water is not advisable.

ELECTRICITY AS A FLUID

David P. Stern
Mauricio Peredo
Goddard Space Flight Center,
NASA

The preceding sections did not follow the conventional order:

First **electric currents** were brought up, in connection with Oersted's work.

Then **electric charges,** in connection with electrons.

That may be OK, because most of us are **familiar** with currents and charges. **Currents** run electric lights, radios, TVs, clocks and appliances in the home, and **"static" electric charge** causes papers and clothes to cling together in dry weather. An electric current is essentially the continuous flow of electric charge. **This section** goes a little further, to the concepts of **voltage** and of the **electric field**.

OHM'S LAW

It is easier to understand electricity if we regard electric charge as a sort of **fluid,** like water, as scientists did for the first 200 years. Yes, it does consist of individual electrons, but those are so small that any large charge behaves like a continuous fluid. In the same way sand pours like a fluid, and water in a glass is usually regarded as a fluid, even though it consists of individual molecules.

Using a pump we can push water through a pipe around a closed circuit (top drawing). The rate at which it flows past any point in the pipe–measured in **gallons** (or liters) **per second**–depends on the **pressure** produced by the pump (measured in pounds per square inch, or kilograms per square centimeter). More accurately, it depends on the **pressure difference** between the entrance to the pipe (left) and the exit from the pipe (right).

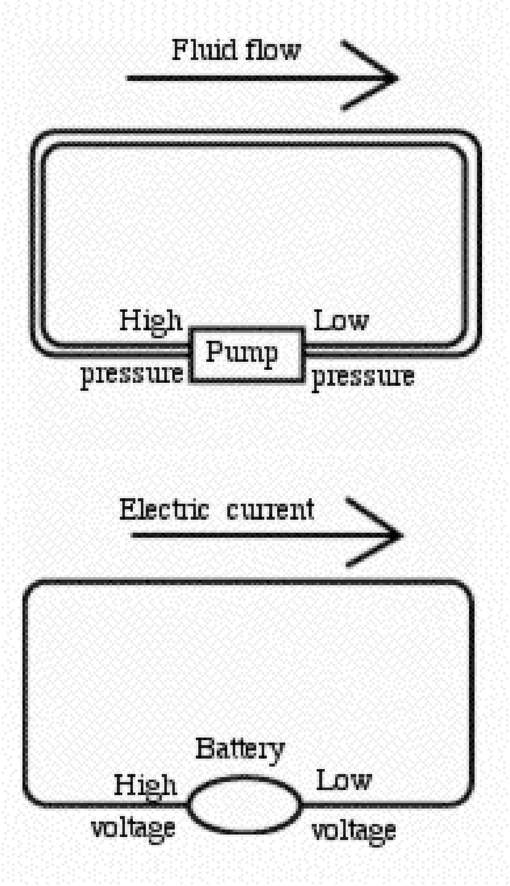

The greater the pressure difference, the greater the flow. In addition, given a certain pressure difference, a **fatter** pipe will carry more water, and a **longer** one will resist the flow and carry less.

ALL THIS MIRRORS EXACTLY THE BEHAVIOR OF THE ELECTRIC FLUID

No pipes here, though: electricity in our homes and appliances usually flows in metal wires, most often of copper. Electrons in a metal can jump from atom to atom, and that way carry negative charge around the circuit.

Like a fluid, they are driven by a kind of electric pressure, known as **voltage,** because it is measured in units known as **volts,** named after the Italian scientist Alessandro Volta. An electric battery produces (by a chemical process) a **voltage difference V** between its two ends, and therefore acts like a pump (bottom drawing).

The **electric current I** flows from high voltage to low voltage and is measured in units known as **Amperes,** named for André-Marie Ampere whom we met in section #2. And as with water, we expect that if we increase the driving voltage V, the driven current I will also increase. In fact, the two are pretty much proportional: double the

---

voltage, and you get double the current. That relation is known as **Ohm's Law,** after Georg Ohm who first formulated it.

Ohm's law (with some extra details concerning the length and thickness of the wire) is usually among the first things taught in electricity classes, and many students therefore view it as one of the fundamental laws of electricity. **It isn't.** It holds quite well for metal wires, but as will be shown in section #7a, it fails badly in fluorescent tubes–while in space currents exist which flow without any voltage driving them (section #10a).

ELECTRIC FIELDS

One small caution here. By long tradition, the direction of the electric current is defined as the direction in which **positive** charges move. We may blame Ben Franklin for deciding–by pure guess–what kind of electricity is called "positive" and which "negative." A century later it was found that most electric currents were carried by negative electrons, which move in the opposite direction.

One may therefore argue that the flow direction which should be assigned to electric fluid is really the **opposite** of what we say it is. But it's much too late to change the old convention.

In our homes, electric currents and effects of electricity are usually channeled along insulated wires. **In 3-dimensional space,** on the other hand, electric phenomena tend to spread out. If the way electric current flows in the home resembles water flow in pipes, then in space the flow is often like ocean currents or air motion in the atmosphere, spread out in 2 and even 3 dimensions.

In a wire, voltage depends only on one dimension–on the distance along the wire. **In space,** every point can have its own voltage. Currents that flow may depend on such voltages, but don't look to Ohm's law for guidance, because the flow of current is primarily dictated by magnetic fields, which makes a difference.

The **3-dimensional voltage distribution** is often called the **electric field.**

Compare:

In a **magnetic field** the **direction of the field** is that of the force. If **isolated N magnetic poles** existed, they too could do so, by moving in the direction of the field, while S poles would move in the opposite direction. The poles at the ends of a compass needle move this way, and thus line up the needle in the direction of the field.

Similarly

An **electric field** is a region where electric forces can be felt by charged objects, and the **direction of the field** is the one in which **positive** charges would move. A positive ion moves in the direction of the field, a negative electron moves in the opposite direction.

If **both** magnetic and electric fields are present, the motion of ions and electrons gets complicated. **That** however is left for a later section.

ANALYSIS OF AQUEOUS SOLUTIONS USING ELECTROSPRAY IONIZATION MASS SPECTROMETRY (ESI MS)

William Henderson
University of Waikato
Hamilton, New Zealand

INTRODUCTION

Electrospray ionization mass spectrometry was developed primarily by Fenn and co-workers in the late 1980s for the analysis of macromolecules such as polyethers and proteins (1). Since then, the technique has undergone an exponential growth in applications, in many areas of chemical and biochemical sciences. A considerable number of textbooks (2,3) and major reviews (4) describing the experimental basis of the technique (5–8) and applications have appeared (8–10) and ESI MS has been widely used for the analysis of aqueous solutions.

THE ESI TECHNIQUE

The key features of the technique are:

1. The ability to analyze materials directly from solution. Almost any solvent, including water, can be used, depending on the solubility of the analyte under study.
2. Ionization is very gentle; this is of great importance for thermally labile analytes such as proteins or weakly bound metal complexes.
3. The technique provides information on mixtures of compounds, which may be in equilibrium in solution, and is applicable to paramagnetic materials (such as many transition metal and lanthanide ions and complexes thereof).
4. The technique is well suited for the analysis of both positive- and negative-ion species.
5. As with all mass spectrometry techniques, miniscule amounts of sample are required, and detection limits are usually very low.

In electrospray ionization, a solution of the compound is nebulized, at atmospheric pressure, from a metal capillary held at a high potential. Nebulization is assisted by an inert gas, typically nitrogen. This produces a spray of charged liquid droplets that, through processes of droplet fission, solvent evaporation, and ion evaporation, produce gas phase ions. The ions are then detected using one of a number of mass analyzers (such as quadrupole or ion trap). Fragmentation of the initially formed ions can be achieved by the technique of collision-induced dissociation (CID), whereby ions undergo collisions with gas and/or solvent molecules. The application of a *cone voltage*—a potential difference that accelerates the ions—is one such method that can effect fragmentation of the original ions.

For analyte species that are already charged, the electrospray process simply transfers existing solution ions into the gas phase. For analytes that are uncharged, ionization must occur for the material to be detected by the mass spectrometer; a number of processes may operate, depending on the actual system, but can include protonation, deprotonation, metallation (e.g., attachment of adventitious alkali metal cations present in aqueous solutions stored in glass), and oxidation.

APPLICATIONS OF ESI MS TO THE ANALYSIS OF AQUEOUS SOLUTIONS

Aqueous solutions are highly amenable to analysis by ESI MS. In many cases, an organic cosolvent is added, to reduce the surface tension of the resulting solvent mixture, which aids in spray formation (7). The surface activity of an analyte ion can affect its ESI response; a high surface activity will result in a high response in the ESI mass spectrum. As an example, an equimolar solution of Cs^+ and $C_{10}H_{21}NMe_3{}^+$ in 1:1 methanol–water gives a substantially more intense signal for the high surface activity ammonium cation (11).

ESI MS can be used for both qualitative and quantitative analysis. As an example, the negative-ion electrospray mass spectrum of an aqueous solution of $NaBrO_3$ (containing 5% methanol) is shown in Fig. 1. Under gentle ionization conditions (cone voltage 30 V), the BrO_3^- ion is exclusively observed at *m/z* (mass-to-charge) 127, but at a high cone voltage (120 V), fragmentation occurs, giving BrO_2^- (*m/z* 111), BrO^- (*m/z* 95), and Br^- (*m/z* 79). The distinctive two-line pattern for each ion arises due to the isotopic signature of bromine (^{79}Br 100%, ^{81}Br 97.3% relative abundances) and *m/z* values are given for the ^{79}Br isotopomer. Electrospray mass spectrometry has been used to quantify levels of oxyhalide anions in aqueous solutions; for example, bromate (12) and other oxyhalide anions (13) can be measured at sub-ppb levels.

For a metal salt M^+ X^- dissolved in water or an organic solvent–water mixture, the ions that are typically observed over a wide mass range are ion clusters of the type $[(MX)_n + M]^+$ in positive-ion mode and $[(MX)_n + X]^-$ in negative-ion mode; salts containing doubly charged ions behave in a similar manner (14).

The tendency for ions to retain a hydration (or more generally a solvation) sphere in the gas phase can easily be investigated using ESI MS. As an example, ESI MS of the sodium salts of the doubly charged sulfur oxyanions SO_4^{2-} (sulfate), $S_2O_6^{2-}$ (dithionate), and $S_2O_8{}^{2-}$ (peroxydisulfate) reveals differences in the gas phase stabilities of these ions. The unsolvated $SO_4{}^{2-}$ ion cannot be detected, and instead a series of hydrated ions $[SO_4 + nH_2O]^{2-}$ ions are observed with $n = 5–16$. Ions with $n < 3$ are unstable and undergo charge separation to give $HSO_4{}^-$ and OH^-. However, for $S_2O_6{}^{2-}$ and $S_2O_8{}^{2-}$, the unsolvated ions can be observed in the gas phase, because of the greater charge delocalization, and hence lower charge density of these ions, which results in higher gas phase stabilities (15).

Metal cations also have a strong tendency to retain a hydration sphere, with divalent cations being more

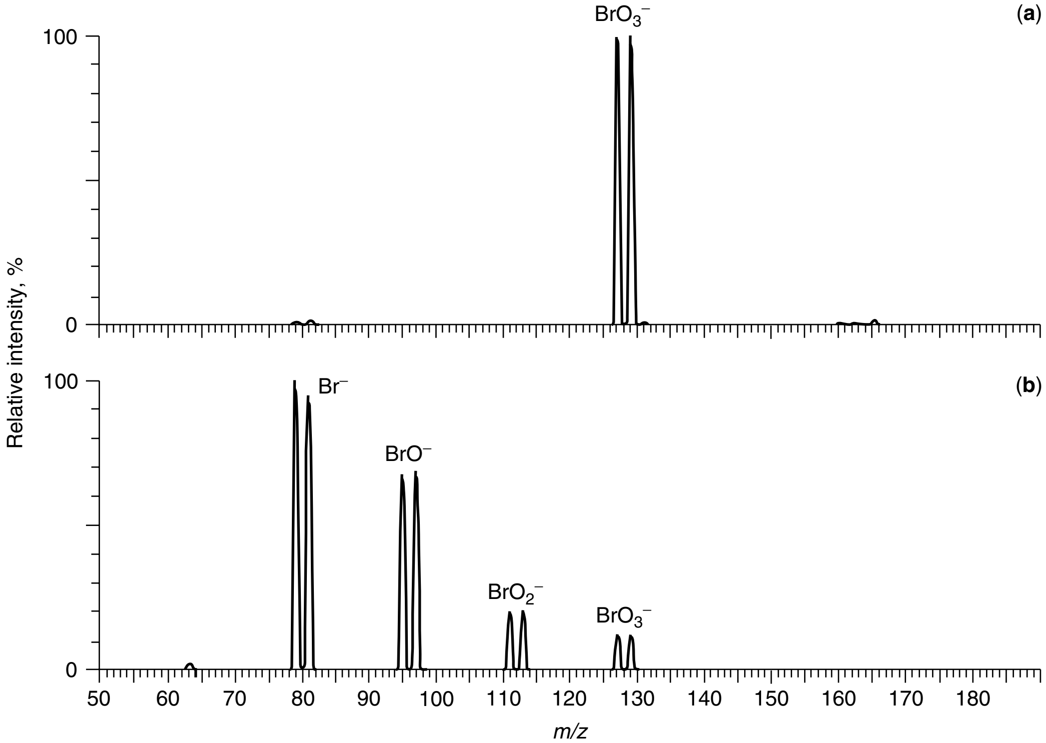

Figure 1. Negative-ion electrospray mass spectra of an aqueous solution of potassium bromate, $KBrO_3$, containing 5% of methanol to aid in spray stabilization. Spectrum (**a**) is recorded at a cone voltage of 30 V and shows exclusively the $BrO_3{}^-$ ion, while spectrum (**b**) is recorded at 120 V and additionally shows the series of fragment ions $BrO_2{}^-$, BrO^-, and Br^-.

strongly solvated than monocations, as expected. For example, $[M(H_2O)_n]^{2+}$ ions have been generated for a range of divalent metal ions, and ion–water binding energies were determined (16). For trivalent cations, and to a lesser extent for divalent cations, charge reduction will occur in the gas phase if the hydration sphere is unable to stabilize the charge on the metal cation. For example, for a solvated trivalent cation

$$[M(H_2O)_n]^{3+} \rightarrow [M(OH)(H_2O)_{n-2}]^{2+} + H_3O^+$$

In the case of copper(II) systems, reduction to copper(I) is often seen. For other metals, under forcing CID conditions, further charge reduction to low valent metal ion species can occur (17): for example, the lanthanide ions give $[Ln(OH)(H_2O)_n]^{2+}$, which give ions LnO^+, $LnOH^{2+}$, and ultimately Ln^+ (18). This experimental regime is often referred to as *bare metal* or *elemental* mode, and mass analysis of the resulting ions can be used to provide isotopic information on the elemental ion concerned.

In addition to the applications described above, the list below gives a selection of some applications of ESI MS in the analysis of aqueous (or mixed aqueous–organic) solutions:

Biochemical materials of many types (19), including proteins (20)
Polyoxometallate anions (21)
Aqueous polysilicate anions (22)
Polyhalide (23) and polyselenide (24) anions
Polyphosphates (25)
Sulfur oxyanions (15,26)
Chromium(III) and chromium(VI) species (27)
Oxyhalide anions $EO_n{}^-$ ($E = Cl$, Br, I; $n = 0–4$) (12,13,28)

BIBLIOGRAPHY

1. Whitehouse, C.M., Dreyer, R.N., Yamashita, M., and Fenn, J.B. (1985). *Anal. Chem.* **57**: 675–679.
2. Gross, J.H. (2004). *Mass Spectrometry—A Textbook*. Springer, New York.
3. Cole, R.B. (Ed.). (1997). *Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation and Applications*. Wiley-Interscience, Hoboken, NJ.
4. Gross, M.L. and Caprioli, R.M. (2003–2006). *The Encyclopedia of Mass Spectrometry*. Elsevier, New York.
5. Kebarle, P. (2000). *J. Mass Spectrom.* **35**: 804–817.
6. Kebarle, P. and Tang, L. (1993). *Anal. Chem.* **65**: 972A–986A.
7. Cech, N.B. and Enke, C.G. (2001). *Mass Spectrom. Rev.* **20**: 362–387.
8. Stewart, I.I. (1999). *Spectrochim. Acta* **54B**: 1649–1695.
9. Fenn, J.B. et al. (1990). *Mass Spectrom. Rev.* **9**: 37–70.
10. Colton, R., D'Agostino, A., and Traeger, J.C. (1995). *Mass Spectrom. Rev.* **14**: 79–106.
11. Zhou, S. and Cook, K.D. (2001). *J. Am. Soc. Mass Spectrom.* **12**: 206–214.
12. Charles, L., Pépin, D., and Casetta, B. (1996). *Anal. Chem.* **68**: 2554–2558.
13. Charles, L. and Pépin, D. (1998). *Anal. Chem.* **70**: 353–359; Barnett, D.A. and Horlick, G. (1997). *J. Analyt. Atom. Spectrom.* **12**: 497–501.
14. Hao, C. et al. (2001). *J. Mass Spectrom.* **36**: 79–96; Hao, C. and March, R.E. (2001). *J. Mass Spectrom.* **36**: 509–521.
15. Blades, A.T. and Kebarle, P. (1994). *J. Am. Chem. Soc.* **116**: 10761–10766.
16. Blades, A.T., Jayaweera, P., Ikonomou, M.G., and Kebarle, P. (1990). *J. Chem. Phys.* **92**: 5900–5906; Jayaweera, P., Blades, A.T., Ikonomou, M.G., and Kebarle, P. (1990). *J. Am. Chem. Soc.* **112**: 2452–2454; Peschke, M., Blades, A.T., and Kebarle, P. (1999). *Int. J. Mass Spectrom.* **185–187**: 685–699.
17. Cheng, Z.L., Siu, K.W.M., Guevremont, R., and Berman, S.S. (1992). *Org. Mass Spectrom.* **27**: 1370–1376.
18. Stewart, I.I. and Horlick, G. (1994). *Anal. Chem.* **66**: 3983–3993.
19. Gaskell, S.J. (1997). *J. Mass Spectrom.* **32**: 677–688.
20. Mann, M. (1990). *Org. Mass Spectrom.* **25**: 575–587.
21. Deery, M.J., Howarth, O.W., and Jennings, K.R. (1997). *J. Chem. Soc. Dalton Trans.*: 4783–4788; Lau, T-C., Wang, J., Guevremont, R., and Siu, K.W.M. (1995). *J. Chem. Soc. Chem. Commun*: 877–878.
22. Bussian, P. et al. (2000). *Angew. Chem. Int. Ed. Engl.* **39**: 3901–3905.
23. McIndoe, J.S. and Tuck, D.G. (2003). *Dalton Trans.*: 244–248.
24. Raymond, C.C., Dick, D.L., and Dorhout, P.K. (1997). *Inorg. Chem.* **36**: 2678–2681.
25. Choi, B.K., Hercules, D.M., and Houalla, M. (2000). *Anal. Chem.* **72**: 5087–5091.
26. Stewart, I.I., Barnett, D.A., and Horlick, G. (1996). *J. Analyt. Atom. Spectrom.* **11**: 877–886.
27. Gwizdala, A.B. III, Johnson, S.K., Mollah, S., and Houk, R.S. (1997). *J. Analyt. Atom. Spectrom.* **12**: 503–506.
28. Charles, L. and Pépin, D. (1998). *Anal. Chem.* **70**: 353–359.

FENTON'S REACTION AND GROUNDWATER REMEDIATION

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Hydrogen peroxide, an *in situ* chemical oxidizer in the form of Fenton's Reaction, can rapidly treat waste waters, sludges, and contaminated soils and groundwater. The effect of the Fenton's Reaction is organic contaminant destruction, toxicity reduction, and biodegradability improvement. In addition, removal of biological oxygen demand (BOD), chemical oxygen demand (COD), odor, and color occur. If Fenton's Reaction is allowed to be carried to completion, then all organic molecules will ultimately be broken into carbon dioxide and water. Fenton's Reaction is a nonselective oxidizer. Fenton's Reaction also called Fenton's Reagent or Fenton's Chemistry in the literature.

Soils contaminated with toxic and recalcitrant organic wastes such as petroleum hydrocarbons and chlorinated solvents can be effectively treated with Fenton's Reaction (1–3). *In situ* chemical oxidation (ISCO) uses contact chemistry of the oxidizing agent to react with petroleum hydrocarbons, volatile organic compounds, munitions, certain pesticides, and wood preservatives in the vadose zone or in the groundwater. The gasoline additive, methyl tertiary butyl ether (MtBE), has been shown to break down with Fenton's Reaction (4–6).

HISTORY

Hydrogen peroxide is one of the earliest chemical oxidants to be discovered and used in industry. It was discovered in the late 1700s and was commercialized in the early 1800s. Hydrogen peroxide works as a remedial treatment chemical in two ways: free radical production and direct chemical oxidation using hydrogen peroxide.

A British Professor H. J. H. Fenton (7,8) described the exothermic and somewhat violent reaction of hydrogen peroxide with iron salts (ferrous sulfate). A very thick and heavy brown foam forms early in the reaction with soil. Fenton's chemistry or Fenton's Reaction uses a transition metal catalyst or an acid to enhance the oxidation chemical reaction of hydrogen peroxide by producing the hydroxyl radical. The hydroxyl radical is a powerful oxidation agent, second behind fluorine.

During the 1930s, the reaction mechanisms were described, although even today, the details are not completely understood because of the large number of side reactions. Commercial reactors have been available for the water treatment industry over the past 20 years. Within the past decade, *in situ* groundwater treatment with Fenton's Reaction have been developed. Within the last several years, commercialization of these methods has become more commonplace.

REACTIONS

Fenton's Reaction produces the hydroxyl free radical, which cleaves and oxidizes organic compounds such as gasoline resulting in producing successively smaller organic compounds. The intermediate compounds of the Fenton's Reaction are generally mono-carboxylic and di-carboxylic acids. These products are nontoxic, naturally occurring, and can be oxidized further to carbon dioxide and water (9).

The breakdown products of the Fenton's Reaction as well as the hydrogen peroxide and iron do not produce additional organic compounds or inorganic solids such as permanganate and dichromate, because there are no carbon atoms in the treatment chemicals. Although many organic compounds can be oxidized with Fenton's Reaction, compounds resulting from the oxidation products of larger molecules, such as small chlorinated alkanes, n-paraffins, and short-chain carboxylic acids, appear to resist further destruction by the Fenton's Reaction. Laboratory bench tests are highly recommended to define feasibility and optimal chemistry on a site-by-site basis.

For *in situ* chemical oxidation, the metal catalyst is usually provided by iron oxides within the soil or fill material, or added separately as a solubilized iron salt, such as iron sulfate. In addition, pH adjustment using an acid such as sulfuric (H_2SO_4) is common because the chemical oxidation is more rapid and efficient under lower pH conditions (pH 2–4 is optimal). The supportive chemical processes that essentially result in the destruction of petroleum hydrocarbons, and other volatile organic compounds are well documented (1–3).

EQUATIONS

When chemical oxidant H_2O_2 is injected at concentrations of 10% to 35% into the subsurface, it decomposes readily into reactive hydroxyl radicals ($OH^\bullet$) and water. The hydroxyl radical ($OH^\bullet$) in the subsurface can rapidly mineralize hydrocarbon, solvent, and other contaminants to water and carbon dioxide. This reaction is enhanced in the presence of iron. Iron is naturally occurring in soil and groundwater or can be added during the injection process, if needed. The reaction is based on the principle of Fenton's Reaction where iron and hydrogen peroxide react to form hydroxyl radicals and other byproducts as shown in (1) (1–3).

$$
\begin{gathered}
Fe^{2+} + H_2O_2 \rightarrow OH^\bullet + OH^- + Fe^{3+} \\
Fe^{3+} + H_2O_2 \rightarrow Fe^{2+} + HO_2/O_2{}^- + H^+ \\
Fe^{3+} + HO_2/O_2 \rightarrow Fe^{2+} + O_2 + H^+ \\
\text{(iron catalyst gets regenerated)} \\
OH^\bullet + H_2O_2 \rightarrow HO_2/O_2{}^- + H_2O \\
OH^\bullet + Fe^{2+} \rightarrow Fe^{3+} + OH^- \\
HO_2/O_2{}^- + Fe^{2+} + H^+ \rightarrow Fe^{3+} + H_2O_2 \qquad (1)
\end{gathered}
$$

$OH^\bullet$ = hydroxyl radical

The process is self-replicating because the reaction of ferric iron with hydrogen peroxide to generate the perhydroxyl radical also occurs (10). Equation (2) is another way to view the reaction:

$$Fe^{3+} + H_2O_2 \rightarrow H^+ + \bullet OOH^- + Fe^{2+} \qquad (2)$$

The perhydroxyl radical is a less powerful oxidant than is the hydroxyl radical, but it does generate further ferrous ions that in turn stimulate further reaction with the hydrogen peroxide to produce more hydroxyl radicals (10).

Chlorinated solvents such as trichloroethylene (TCE) and tetrachloroethylene (PCE) have been the focus of *in situ* applications of Fenton's Reaction. The double bonds, C=C, that characterize chlorinated ethenes are more reactive than are the single C–C bonds of chlorinated ethanes. Therefore, PCE and TCE are more susceptible to chemical oxidation than is TCA. Although these chemicals are susceptible to chemical oxidation, relative resistance to oxidation goes from highest to lowest: PCE, TCE, vinyl chloride, phenanthrene, benzene, and hexane. The oxidation reaction for a common solvent, TCE, forms several unstable daughter products (epoxides), then breaks down to ketones and aldehydes, finally yielding carbon dioxide, water, and chloride ions (11). The oxidation of TCE is shown in (3):

$$4OH^\bullet + C_2HCl_3 \rightarrow 2CO_2{}^- + 3Cl + 5H^+ \qquad (3)$$

The hydroxyl radical that attacks the carbon–hydrogen bonds is capable of degrading many chlorinated solvents, chloroalkenes, esters, aromatics, pesticides, and other recalcitrant compounds such as MTBE, PCP, and PCB (2). The Fenton's chemistry reaction is highly complex. The iron cycles between the Fe(II) and Fe(III) oxidation states yield the hydroxyl radical and other byproducts (11).

Residual H_2O_2 not used in the oxidation process breaks down to water and oxygen in a matter of hours. In addition to the reaction described in (2), there are a large number of competing reactions including the free radical scavengers, most importantly, carbonate and bicarbonate alkalinity, that will greatly affect the overall reaction scheme. In addition, H_2O_2 can serve as an oxygen source for microbes in the subsurface to enhance biodegradation of contaminants.

IN SITU ISSUES

Fenton's Reaction has been effective in the relatively low concentrations in waste water treatment plants, but not in *in situ* applications. In this scenario, solid oxygen demand is limited in the treatment water and competing reactions are small by comparison with those in soils. Soil contains competing reactions that will react with the hydrogen peroxide, including organic carbon, unoxidized minerals, and biologically produced chemicals. Therefore Fenton's Reaction treatment plans in the field may take significantly more applications than the bench testing might indicate because of the competing reactions that occur in the soil and groundwater. More information about hydrogen peroxide decomposition in the presence of aquifer material is described in Miller and Valentine (12).

Although Fenton's chemistry has been documented for well over 100 years, it has been employee safety and handling issues that have kept large numbers of environmental contractors from using *in situ* chemical oxidation technology. Although this technology can be used safely, significant safety planning, worker training, personal protective equipment, on-site supervision, and monitoring must be an integral part of all oxidation projects. The main challenges for *in situ* applications are delivery of the oxidants, which are difficult to handle, and pH control of the subsurface environment where the reaction must occur. Since most soil has adequate concentrations of naturally occurring iron, addition of iron is not needed at many sites. The typical iron to hydrogen peroxide ratios is 1:5 to 10 wt/wt. Iron levels less than 25 to 50 mg/L can require excessive reaction times (10 to 24 hours) (13). Improper control of the Fenton's reaction will only generate oxygen and water, not the high intensity Fenton's oxidation chemistry.

DELIVERY OPTIONS

Delivery of the Fenton's Reaction chemistry requires close spacing of the injection ports/wells because of the rapid nature of the reactions. The liquids are generally pumped using pressures of 100 to 600 psi for probe rods or 1000 to 3000 psi using lances. Pouring hydrogen peroxide into wells or trenches may work for treating water in those areas, but because of the rapid reaction rates, pressure is needed to move the peroxide away from the injection location. The high pressure is also needed to ensure mixing in the subsurface and a reasonable radius of influence. Fenton's Reaction, although powerful, is not generally recommended or cost effective for low-concentration groundwater plumes where the concentrations of organics is less than 5 mg/L. Variations in the effective porosity, permeability, lithologic heterogeneities, BOD, COD, pH, and contaminant mass affect the outcome of *in situ* applications of Fenton's Reaction.

Hydrogen peroxide reacts in an optimal manner in lower pH settings, with lower alkalinity readings. In some cases, strong acids, most notably sulfuric acid, lower the pH. Trace chloride from chlorinated compounds will likely combine with sodium or calcium ions to form salts or with hydrogen to form weak acids. Careful evaluation of soil and water chemistry with a bench test with soil and water samples is recommended before the start of any injection process. Because of the rapid reaction time of the Fenton's Reaction, subsurface spacing of injection ports must be relatively close. The lower the effective porosity of the treatment zone, the closer the injection port spacing. Based on field experience, clays and silts, which are problematic to remediate *in situ*, typically require 0.6 to 1.0 m spacing, whereas injection ports for clean sand and gravels can be placed at 2.0 to 3.0 m spacing.

Chemical compatibility of the injection equipment components and safety procedures become critical with the injection of strong acids, bases, oxidants, and other chemicals. Although handling hydrogen peroxide and other oxidants requires significant safety training and planning, the oxidant is effective at remediation of a variety of organic contaminants and is relatively inexpensive. The reaction time for hydrogen peroxide in the subsurface is usually within seconds to minutes rising under ideal conditions to hours. A rise in temperature in the subsurface illustrates the exothermic nature of Fenton's Reaction process.

Based on field research, the optimum reaction temperature is relatively low, as measured on the reaction foam in the field using infrared thermal meters in 35 °C to 41 °C. With increases in temperature beyond about 57 °C, the peroxide becomes more volatile as it starts to produce a wispy white vapor. The gas is water vapor and carbon dioxide, the products of chemical oxidation. Subsurface reaction temperatures are measured with an infrared thermal meter. Reaction temperatures in the range of 82 °C to 93 °C are explosive and unsafe. Temperatures of subsurface chemical reactions can be monitored and lowered by adding water, lowering concentrations of the catalyst or oxidant, and reducing injection pressures. Effective porosity of the treatment zone may be reduced by the formation of metal oxide precipitates.

CASE STUDIES

Fenton's Reaction has been successful on a variety of sites throughout the United States since the early to mid-1990s. Reductions of organic contaminants such as DNAPLs in full-scale treatment projects are as high as 90% to 99%, depending on the contaminant, the subsurface variables, and the starting concentrations. Specific Fenton's Reaction case examples are included in a variety of publications, including U.S. EPA, (14), U.S. Navy (9), U.S. Navy (15), and Jacobs and Testa (16).

BIBLIOGRAPHY

1. Watts, R.J., Smith, B.R., and Miller, G.C. (1991). Treatment of octachlorodibenzo-p-dioxin (OCDD) in surface soils using catalyzed hydrogen peroxide. *Chemosphere* **23**: 949–956.
2. Watts, R.J. et al. (1992). Hydrogen peroxide for physicochemically degrading petroleum-contaminated soils. *Remediation* **2**: 413.
3. Watts, R.J. and Stanton, P.C. (1994) *Process Conditions for the Total Oxidation of Hydrocarbons in the Catalyzed Hydrogen Peroxide Treatment of Contaminated Soils*. WA-RD, Vol. 337.1, Washington State Department of Transportation, Olympia, WA.
4. Jacobs, J., Guertin, J., and Herron, C. (Eds.). (2001). *MTBE and its Effect on Soil and Groundwater Resources*. Lewis Publishers/CRC Press, Boca Raton, FL.
5. Leethem, J.T. (2002). In-situ chemical oxidation of MTBE: A case study of the successful remediation of a large gasoline release. *Contaminated Soil Sediment and Water* **July/August**: 70–75.
6. Kelly, K.L., Marley, M.C., and Sperry, K.L. (2002). In-situ chemical oxidation of MTBE. *Contaminated Soil Sediment and Water*. **July/August**: 36–40.
7. Fenton, H.J.H. (1893). *Proc. Chem. Soc.* **9**: 113.
8. Fenton, H.J.H. (1894). *J. Chem. Soc.* **65**: 899.
9. U.S. Navy. (2001). In Situ Chemical Oxidation Process using Fenton's Reagent, Navy Pollution Abatement Ashore Technology Demonstration/Validation Program. Available: http://p2ashore.nfsec.navy.mil/cgi-bin/project_descriptions/cfpr_results.cfm?PROJECT.
10. Vance, D.B. (2002). New Applications in Redox Reactions for In-Situ Groundwater Remediation. Available: http://2the4net/ThreeRedox.htm.
11. Suthersan, S.S. (2002). *Natural and Enhanced Remediation Systems*. CRC Press/Lewis Publishers, Boca Raton, FL.
12. Miller, C.M. and Valentine, R.L. (1995). Hydrogen peroxide decomposition and quinoline degradation in the presence of aquifer material. *Water Resources* **29**(10): 2353–2359.
13. U.S. Peroxide. (2002). Reference Library Peroxide Applications: Fenton's Reagent, Iron-Catalyzed Hydrogen Peroxide. Available: http://www.h2o2.com/applications/industrial-wastewater/fentonsreagent.html.
14. U.S. EPA. (2000). Abstracts of Remediation Case Studies, Vol. 4, EPA 542-R-00-006. Available: http://www.epa.gov/tio/download/
frtr/abstractsvol4.pdf.
15. U.S. Navy. (2003). Available: http://enviro.nfesc.navy.mil/erb/restoration/technologies/remed/phys_chem/phc-43.asp.
16. Jacobs, J. and Testa, S.M. (2003). Design considerations for in-situ chemical oxidation using high pressure jetting technology. *AEHS Soil, Sediment and Water* **March/April**: 51–60.

READING LIST

Watts, R.J., Udell, M.D., Rauch, P.A., and Leung, S.W. (1990). Treatment of Pentachlorophenol Contaminated Soils Using Fenton's Reaction. *Hazardous Waste Hazardous Materials* **7**: 335–345.

WHERE WATER FLOATS

NASA Marshall Space Flight Center

The **Fluid Physics Research Program** is implemented at the Lewis Research Center in Cleveland, Ohio and is managed by the Microgravity Research Program Office located at Marshall Space Flight Center in Huntsville, Alabama.

A fluid is any material that flows in response to an applied force, and assumes the shape of its container. Examples include liquids, gases, and granular systems such as soil and sand.

Fluid physics is vital to understanding, controlling, and improving all of our industrial and natural processes. The engines used to propel a car or an airplane, the shape of the wings of an airplane that allow it to fly, the operations of boilers that generate steam used to produce over 90% of the world's electric power, the understanding of how cholesterol is transported in our bloodstream and effects heart disease, and how pollutants are transported and dispersed in air and water are just a few examples of how fluid physics affects our everyday life and forms the very basis for an industrial society. Areas of NASA's research in fluid physics include the following:

Complex Fluids are fluids containing structures or aggregates on supermolecular scales; examples include

---

colloids, emulsions, liquid crystals, granular solids, and polymer solutions.

Multiphase Flow and Heat Transfer includes the study of processes such as boiling, condensation, flows and heat transfer in fluid-fluid and fluid-solid mixtures.

Interfacial Phenomena investigates the shape and dynamics of solid-liquid-gas interfaces such as during a coating process, the liquid-gas interface as in partially filled containers, and the effect of energy and mass transfer on these interfaces.

Dynamics and Instabilities deal with fluid flows and stability of thin and thick films, geological and biological flows, liquid bridges, drops and bubbles, and electrochemical and chemical transport. Pattern formation and turbulence are other interesting areas investigated under this topic.

FRESHWATER

Gholam A. Kazemi
Shahrood University of Technology
Shahrood, Iran

INTRODUCTION

Water is an important solvent and is sometimes called the universal solvent (1). It dissolves, to the extent possible, the material with which it comes in contact. Salinity is a measure of various constituents dissolved in water, including solids, liquids, and gases. Total dissolved solids (TDS), another form of expressing the amount of dissolved ingredients in water, is the sum of total solids that have been dissolved in water [note that TDS at $180\,^\circ\mathrm{C} = \text{Sum of ions} + \mathrm{SiO_2} - 0.5082 \times \mathrm{mg/l\,HCO_3}$; (2)]. However, many references do not differentiate between TDS and salinity because the absolute majority of the dissolved constituents of water are usually the solids. Both parameters have units of mg/l or ppm. These two units, ppm and mg/l, are almost the same except for high salinity waters for which $1\ \mathrm{mg/l} = 1\ \mathrm{ppm} \times$ density of water. The salinity of groundwater ranges from 25 mg/l in quartzite springs to more than 300,000 mg/l in brine (3), and the salinities of distilled water, River Murray in South Australia, seawater, and the Dead Sea are 0, 553, 35,000, and 385,000 mg/l, respectively (4).

Waters are classified into various types based on their salinity or TDS content. Two widely used classification systems are illustrated in Table 1. The Australian Water Resources Council (AWRC) (1976) (5) has devised a considerably different classification, shown in Table 2, which can be regarded as one of limited local use.

DEFINITION

Based on the two widely used classification systems above, freshwaters are those whose salinity or TDS is less than 1000 mg/l. Interestingly, freshwater has identical definitions in both systems; other water types, for example, brines, have slightly to considerably different meanings. Despite this equality, Manning (8) believes that criteria for terms such as fresh, when applied to water quality, are not universal. He writes (p. 221) “It depends on the individual tastes and on the purpose for the water. One person’s unacceptable saline water might be another’s only drinking water.”

In a generalized sense, freshwaters may be regarded as all waters in the world, excluding saline waters of ocean, seas, and saline lakes. Based on this, global freshwater resources are estimated at $36.02 \times 10^6\ \mathrm{km}^3$ accounting for about 2.6% of the total world water of $1.348 \times 10^9\ \mathrm{km}^3$ (9). If one adopts this generalization, then, rivers, freshwater lakes, most groundwaters, and soil moisture all are eligible to be categorized as freshwater resources.

The term “freshwater,” may be misleading. It may suggest that freshwaters are those that are suitable for drinking, but this is not the case. In other words, freshwater is an arbitrary chemical denomination and does not strictly imply suitability of this type of water for drinking. Various guidelines, such as those established by the United States, European Community, and World Health Organization, suggest a TDS of less than 500 mg/l to 1500 mg/l for drinking water, which does not completely conform with the TDS of freshwater. For instance, a particular type of freshwater may not be suitable for drinking because one or more of its constituents exceeds the desirable criteria. Water from an irrigation field,

Table 2. AWRC Classification of Water

Water Type	TDS
Fresh	<500 mg/l
Marginal	500–1,000
Brackish	1000–3000
Saline	>3000

Table 1. Water Types Based on TDS Value—(Different Classification Systems)

Water Type[a]	TDS mg/l[a]	Water Type[b]	TDS mg/l[b]
Fresh	0–1000	Fresh	0–1000
Brackish	1000–10,000	Slightly saline	1,000–3,000
Saline	10,000–100,000	Moderately saline	3,000–10,000
Brine	>100,000	Very saline	10,000–35,000
		Briny	>35,000

[a] *Source*: Reference 6, in Reference 3.
[b] *Source*: Reference 7—proposed by USGS.

for example, may be chemically classified as fresh, but due to its high nitrate, phosphate, or toxic components is more likely to be rejected as a source of drinking water. Similarly, water that flows through, or emerges from, geologic media containing sulfur deposits will not be suitable for drinking, though chemically fresh, because of its low pH value. Hundreds or thousands of similar cases can be described. However, in normal conditions and in the majority of cases, so-called freshwater meets drinking water standards and can be consumed.

SOME GENERAL POINTS ABOUT FRESHWATER

1. All types of wet atmospheric precipitations such as rain, snow, and hail turn into freshwater.
2. Surface waters are generally lower in TDS than groundwaters and, thus, are more likely to be classified as freshwater compared with groundwater. It is not incorrect to say that river waters are generally fresh.
3. Various groundwaters (excluding connate water, intruded sea water, magmatic and geothermal water, and water affected by salt leaching) generally have a TDS of less than 10,000 mg/l (10). Thus, they may or may not be fresh.
4. Generally the electrical conductivity (EC) of groundwater in hard rock regions is 120 μS/cm or less (11) and therefore, is fresh.
5. Water emanating from upland catchments and water of lowland rivers that have typical ECs less than 700 μS/cm(12) are fresh.
6. Igneous, volcanic, carbonate, and metamorphic rocks generally yield freshwater.
7. Freshwater may turn into other types of water (brackish, saline, etc) by various means, including human-induced pollution, sea/saltwater intrusions, overexploitation, and other natural causes.

BIBLIOGRAPHY

1. Kemmer, F.N. (1988). *The NALCO Water Handbook*, 2nd Edn. McGraw-Hill, New York.
2. Hounslow, A.W. (1995). *Water Quality Data: Analysis and Interpretation*. Lewis, New York.
3. Todd, D.K. (1980). *Groundwater Hydrology*, 2nd Edn. John Wiley & Sons, New York.
4. Taylor, S. (1993). *Dryland Salinity: Introductory Extension Notes*. Department of Conservation and Land Management of New South Wales, Australia.
5. Australian Water Resources Council. (1976). *Review of Australia's Water Resources 1975*. AGPS Canberra, for the Department of National Resources.
6. Carroll, D. (1962). *Rainwater as a Chemical Agent of Geologic Processes—A Review*. US Geological Survey Water-Supply Paper 1535.
7. Hem, J.D. (1985). *Study and Interpretation of the Chemical Characteristics of Natural Waters*, 3rd Edn. U.S. Geological Survey Water-Supply Paper 2254, Denver, CO.
8. Manning, J.C. (1997). *Applied Principles of Hydrology*. Prentice-Hall, Upper Saddle River, NJ.
9. McDonald, A. and Kay, D. (1988). *Water Resources: Issues and Strategies*. Longman Scientific and Technical, New York.
10. Rail, C.D. (2000). *Groundwater Contamination*. Vol. 1: *Contamination, Sources and Hydrology*. Technomic, Lancaster, PA.
11. Bender, S. (2003). *Low Mineralized Water: Challenge for Hydrochemical Investigations in Hard-rock Area. Proc. Int. Conf. Groundwater Fractured Rocks*, Prague, IHP-VI, series on groundwater no. 7, pp. 195–196.
12. Tebbutt, T.H.Y. (1995). *Principles of Water Quality Control*. Reprint of 4th Edn. Butterworth Heinemann, Oxford.

DISSOLVED GASES

M. Boström
Linköping University
Linköping, Sweden

B.W. Ninham
University of Regensburg
Regensburg, Germany
and
Australian National University
Canberra, Australia

Gases dissolved in water play an important role in solution chemistry and in biological processes in solutions. The story that we will present of dissolved gases in water is in some sense twofold. We begin with the more conventional description and a general discussion of gases dissolved in groundwater, in seawater, and in blood, and decompression sickness. We then give a short introduction to a new and rapidly developing scientific field within colloid and solution chemistry that has demonstrated a profound role for dissolved gas in, for example, colloidal interactions, hydrophobic interactions, conductance, surface tension, and bubble fusion. We argue strongly that different scientific disciplines that consider dissolved gases in very different ways can benefit tremendously from better knowledge of work in other disciplines.

DISSOLVED GASES IN GROUNDWATER

Dissolved gases and dissolved organic matter (hydrophobic acids, bases, and neutrals) are ubiquitous in water. A large number of dissolved organic materials are composed of humic acids that are heterogeneous organic macromolecules. These have an important role in binding of metals and pesticides. Humic substance concentrations range from 20 μg/L in groundwaters to 30 mg/L in surface waters. Groundwater also contains dissolved gases due to exposure to the atmosphere, gas production reactions with minerals, and biological activity. These gases are certainly at least as important as dissolved matter. The most common gases are those that constitute the atmosphere; O_2, N_2, and CO_2. Methane, H_2S, N_2O, and other gases are also present in groundwater due to biogeochemical processes. Dissolved gases can have a tremendous impact. The presence of dissolved H_2S in water can, for example, give rise to the smell of rotten eggs, and above an H_2S concentration of 1 mg/L, the water is no longer suitable for consumption. Dissolved oxygen, on the other hand, is vital for biology. The concentration of dissolved oxygen (at 20 °C) at saturation is 9.1 mg/L. This is not that far above the minimum concentration of dissolved oxygen that can maintain aquatic life (which is around 5 mg/L). Aerobic bacteria can break down oxygen and leave an oxygen concentration close to this lower limit (it may even become too low to sustain aquatic life) (1). Another major dissolved gas is carbon dioxide. The total concentration of carbon oxides ($[CO_2(aq)] + [H_2CO_3] + [HCO_3^-] + [CO_3^{2-}]$) is around 2.3 mM in seawater, 0.1–5 mM in river water, 0.5–8 mM in groundwater, and 10–50 μM in rainwater (2).

DISSOLVED GASES IN THE OCEANS

A large number of measurements of the fluxes of gases between the atmosphere and the ocean have been carried out to determine gases dissolved in seawater (3,4). The first determination of N_2 dissolved in Atlantic seawater was performed in 1938 (5). A large number of methods have been used to measure dissolved gases in water (e.g., Winkler method, gasometric methods, mass spectroscopic methods (6), and gas chromatographic methods (4,7)). The mechanisms of gas exchange in the ocean are not entirely understood. Studies of gas exchange between air and water are often based on the simple theoretical surface film model developed by Liss and Slater (3). However, it is a complicated system and more recent attempts have been made to include different complications. Gas exchange depends, for example, on the partial pressure of a gas in air, its solubility coefficient, concentration in bulk surface water, wind velocity (8), temperature, salinity, molecular diffusivity, bubbles, and bubble size (9). Another complicating factor for oxygen and carbon dioxide dissolved in the ocean is that biological processes affect them. Oxygen profiles as a function of depth under the ocean surface are very different compared to the profile of N_2 (4) because O_2 is produced in the surface euphotic zone by photosynthesis and consumed by decomposition of organic matter in the aphotic zone.

The composition of gases in seawater (or in any other aqueous solution) under solution equilibria differ from their atmospheric composition. Dissolved gases that have been studied include carbon dioxide (CO_2), carbon monoxide (CO), nitrous oxide (N_2O), nitrogen (N_2), oxygen (O_2), methane (CH_4), sulfur dioxide (SO_2), dimethyl sulphide (DMS), and argon (Ar). Many of these gases are related to global warming, and there has been some discussion of the ocean as a source of methane to the atmosphere (4,10). Some of these dissolved gases are also contained in air bubbles. For example, Nakayama et al. (4) report values around 0.6 mM/kg for N_2, 0.35 mM/kg for O_2, and 16.5 μM/kg for Ar in the surface (0–50 m depth) region of the north Pacific Ocean and a mean value of 18.4 μmol/kg of dissolved air bubbles in the upper 30 m of water (4).

DISSOLVED GASES IN BIOLOGICAL SYSTEMS

Dissolved gases are very important in biological systems, and all forms of life visible to the human eye require oxygen. Hemoglobin is an important transport protein embedded in the cell membrane of the red blood cell. Its main function is to collect oxygen in the lungs, where it is abundant, and deliver it to the tissues around the body. O_2 that is used by most organisms is converted to CO_2. The carbon dioxide that diffuses in the blood is mainly in the form of dissolved CO_2 because conversion to carbonic acid is a comparably slow process. Long before this conversion takes place, it has usually diffused into the red cell. Once inside, conversion to carbonic acid takes place well buffered by the hemoglobin in the cell wall. While O_2 is delivered to the tissues, CO_2 is absorbed and transported out of the system (11).

Large volumes of gas dissolve in the tissues of divers. Gases such as nitrogen and helium are not metabolized by our bodies and need to be eliminated during slow decompression. Gradual decompression for several days is needed following extended human dives to pressures above 20 atm to avoid excessive gas supersaturation and bubble formation that may lead to decompression sickness (12,13). An interesting biochemical method to tackle decompression sickness has been investigated in experiments on rats. Suspensions of the human colonic microbe *Methanobrevibacter smithii* were introduced into the rats. This microbe that converts 4 mol of H_2 and 1 mol carbon dioxide (which is a gas that is abundant in tissues) to 1 mol methane and water, significantly reduces the risk of decompression sickness in rats (12). Decompression sickness begins at a much lower gas pressure than traditionally anticipated (less than 2 atm). The most likely explanation for this has only recently emerged and is discussed in the following section.

THE ROLE OF DISSOLVED GASES IN COLLOID SCIENCE AND BIOCHEMISTRY

Literature has accumulated during the last decade in physical chemistry journals focusing on the effects of dissolved gases on, for example, bubble fusion (14), emulsion stability (15,16), colloid interactions (17), hydrophobic forces (18,19), and optical cavitation (20). Bubble fusion depends strongly on salt concentration and ionic species but obviously also on the amount of dissolved gas (14). A mechanism has been proposed for hydrophobic forces that are based on surface adsorption and nucleation of dissolved gases in water. Suggestions have been made that dissolved gas molecules can be involved in propagating perturbations induced by hydrophobic surfaces. This could possibly give rise to the extraordinarily long-range hydrophobic interaction observed. There is clear evidence that degassing a system can significantly reduce both the magnitude and range of the hydrophobic interaction. Bubble formation occurs much more readily near hydrophobic surfaces. We have plenty of hydrophobic surfaces in our bodies, which is highly relevant for the understanding of decompression sickness. It is a plausible explanation why decompression sickness begins at a much lower pressure than expected and why physical exertion before decompression can increase the severity. Exertion is likely to damage tissue mildly and increase the number of hydrophobic sites where such cavitation can occur (14). The interested reader is referred to the references suggested, but we would like to mention one very interesting proposal made by Karaman et al. (15). They argued, quite convincingly, that there might be an unanticipated role for dissolved gases in biochemistry. Chemical reactivity can depend on dissolved gas through enhanced radical production in microbubbles formed from gas aggregates. Experiments in molecular biology have revealed that negatively charged restriction enzymes lock onto negatively charged DNA and cut it at particular sequences. Enzymes and DNA have the same charge, so this cannot be explained in terms of electrostatics, and the source of the energy to drive the catalysis is not explained. If cavitation occurs on contact, free radical $OH^\bullet$ production follows, and the transfer of an electron to $Cl^\bullet$ does provide sufficient energy. Experiments reveal that above 0.2 M NaCl, where cavitation is inhibited, the enzyme loses its catalytic property entirely and reversibly. Below 0.2 M NaCl, the addition of vitamin C, a well-known free radical scavenger, totally inhibits the enzyme. Dissolved gas can play an important role in harnessing weak molecular forces cooperatively to influence vital biochemical reactions. Interestingly, the cutting efficiency of DNA by enzymes also depends on the choice of background salt following a so-called Hofmeister series (see article on the "Hofmeister Effect" in this volume), in particular on the supposedly irrelevant choice of negative co-ions (21). One can show that this experimental finding further strengthens the proposed mechanism for the activation energy.

THE ROLE OF DISSOLVED GASES IN SURFACE TENSION OF ELECTROLYTES

In the past, there has been a discussion of the ion specificity found for the surface tension of salt solution. In the air–water interfacial region (which is roughly the first 6 Å closest to the air–water interface), there is a gradual transition of the relative amount of water and dissolved gases. This means that an ion near the interface experiences a force that originates from the change in solvation energy when the ion moves into the surface region (with its increasing concentration of dissolved gas and decreasing concentration of water molecules). This is besides previously discussed image potentials and dispersion potentials (originating from the polarizability of ions). Inclusion of solvation energy effects acting on ions near the air–water interface gives exactly the right order of magnitude and the right ion specific sequence for the surface tension of different salt solutions (22). If one furthermore includes dispersion potentials acting on ions, one improves the result further. There are clearly other effects that influence the surface tension of salt solutions, but it is clear that solvation energy effects, related to the interfacial profiles of dissolved gases and water molecules, together with ionic dispersion potentials between ions and the interface, play vital parts. References and a discussion of ion specific effects can be found in the article "Hofmeister Effects" in this volume.

Some attempts have been made to model theoretically the influence of dissolved gases, but there is in most cases still a long way to go until it is firmly understood. But it is very clear from experiments that dissolved gases play a very important, sometimes vital, role in a large number of different chemical and biological systems in water.

BIBLIOGRAPHY

1. Lester, J.N. and Birkett, J.W. (1999). *Microbiology and Chemistry for Environmental Scientists and Engineers*. E & F.N. Spon, London.
2. Stumm, W. and Morgan, J.J. (1981). *Aquatic Chemistry: An Introduction Emphasizing Chemical Equilibria in Natural Waters*. John Wiley & Sons, New York.
3. Liss, P.S. and Slater, P.G. (1974). *Nature* **247**.
4. Nakayama, N., Watanabe, S., and Tsunogai, S.J. (2002). *Oceanography* **58**: 775.
5. Rakestraw, N.W. and Emmel, V.M.J. (1938). *Mar. Res.* **1**: 207.
6. Emerson, S., Quay, P.D., Stump, C., Wilber, D., and Knox, M. (1991). *Biogeochem. Cycles* **5**: 49.
7. Gamo, T. and Horibe, Y. (1980). *Bull. Chem. Soc. Jpn.* **53**: 2839.
8. Liss, P.S. and Merlivat, L. (1986). In: *The Role of Air-Sea Exchange in Geochemical Cycling*. P. Buat-Menard (Ed.). Kluwer, Dordrecht, The Netherlands.
9. Merlivat, L. and Memery, L. (1983). *J. Geophys. Res.* **88**: 707.
10. Bates, T.S., Kelly, K.C., Johnson, J.E., and Gammon, R.H. (1996). *J. Geophys. Res.* **101**: 6953.
11. Zubay, G.L. (1998). *Biochemistry*, 4th Edn. Wm. C. Brown, Dubuque, IA.
12. Weathersby, P.K., Homer, L.D., and Flynn, E.T. (1982). *J. Appl. Physiol.* **53**: 940.
13. Kayar, S., Miller, T., Wolin, M., Aukhert, E., Axley, M., and Kiesow, L. (1998). *Am. J. Physiol.* **275**: R677.
14. Craig, V.S.J., Ninham, B.W., and Pashley, R.M. (1993). *J. Phys. Chem.* **97**: 10192.
15. Karaman, M.E., Ninham, B.W., and Pashley, R.M. (1996). *J. Phys. Chem.* **100**: 15503.
16. Maeda, N., Rosenberg, K.J., Israelachvili, J.N., and Pashley, R.M. (2004). *Langmuir* **20**: 3129.
17. Alfridsson, M., Ninham, B.W., and Wall, S. (2000). *Langmuir* **16**: 10087.
18. Craig, V.S.J., Ninham, B.W., and Pashley, R.M. (1999). *Langmuir* **15**: 1562.
19. Wennerström, H. (2003). *J. Phys. Chem. B* **107**: 13772.
20. Bunkin, N.F. et al. (1997). *Langmuir* **13**: 3024.
21. Kim, H-K., Tuite, E., Norden, B., and Ninham, B.W. (2001). *Eur. Phys. J. E* **4**: 411.
22. Boström et al. (2004). Submitted.

HARD WATER

Gholam A. Kazemi
Shahrood University of
Technology
Shahrood, Iran

INTRODUCTION

In defining hard water, first it is necessary to describe hardness, a closely related chemical property of water.

Hardness is defined by some as a property of water that prevents lather formation with soap and produces scale in hot water systems (1); by others, in simpler terms, as the soap consuming capacity of water (2); and in broader terms, by Freeze and Cherry (3), as the "metallic ion content of water which reacts with sodium soaps to produce solid soaps or scummy residue and which reacts with negative ions, when the water is evaporated in boilers, to produce solid boiler scale."

The ions which result in, or produce, hardness include Mg^{+2}, Ca^{+2}, Sr^{+2}, Fe^{+2}, and Mn^{+2}, and to a lesser extent, Ba^{+2} and Zn^{+2} and other divalent ions. The first two are the most influential. There are different forms of hardness (4):

1. *Total hardness*: Ca and Mg expressed as $CaCO_3$
2. *Calcium hardness*: Ca expressed as $CaCO_3$
3. *Magnesium hardness*: Mg expressed as $CaCO_3$
4. *Temporary hardness*: Equals alkalinity, if alkalinity is less than total hardness
5. *Permanent hardness*: The amount of hardness greater than temporary hardness
6. *Noncarbonate hardness*: Permanent hardness

Of the six different forms of hardness, the most important and the one that is commonly known as hardness, is total hardness. It is calculated from the following simplified equation:

$$\text{Hardness} = 2.5\ \text{Ca} + 4.1\ \text{Mg} \quad (\text{units in mg/l})$$

Note that there are equations to calculate each of the six different forms of hardness.
Example: The Ca and Mg concentrations in a water sample are 20 and 15 mg/l, respectively. What is its hardness?

$$\text{Hardness} = 2.5 \times 20 + 4.1 \times 15 = 111.5\ \text{mg/l}$$

As stated, each of the six forms of hardness can be calculated by analyzing a water sample for the relevant ion(s), followed by using the relevant equation. However, there are techniques for measuring total hardness directly such as (a) a standard solution of soap, that was in use for many years (2); and (b) Schwarzenbach et al.'s method of colorimetric titration (5), which is sometime referred to as the EDTA method.

The author believes that earlier, the measurement of hardness was a common and an inseparable practice in water analysis. To agree with this statement, one needs to look at a relatively old handbook and read, "The common denominator of the majority of water problems is hardness" (Ref. 2, p. 4.1, first published in 1979). Somehow, in contrast, recent hydrochemical analysis of water samples does not usually include a measurement of hardness. For instance, there is not a single report of hardness in the hundreds of hydrochemistry papers presented at the International Association of Hydrogeolgists (IAH) conferences in Prague (2003) and in Munich (2001) and in the recent past (1997–2001) annual symposia of the Geological Society of Iran. Other aspects of water chemistry and water biology, such as nutrients, heavy metals, bacteriological species, chlorinated hydrocarbons and many more contaminants have become so widespread and are considered more important and more dangerous than hardness. Rejection of water for drinking or other uses, these days, is many times more likely to be due to its harmful DDT level than because of its high hardness value. One other possible reason is that hardness and hard water topics are only important in a narrow field of water, e.g., hot water boilers.

The second point is that industrialization, urbanization, agricultural practices, and most environmentally unfriendly activities are not likely to result in an increase in the hardness of water resources. These activities do not usually generate hardness producing ions such as Ca^{+2} or Mg^{+2}, though, they may produce other trivial (in terms of hardness) ions such as Mn^{+2} and Fe^{+2}.

DEFINITION AND SPECIFICATIONS OF HARD WATER

Todd (6, p. 282) reports that Hippocrates (460–354 B.C.) was the first to use the term hard and soft for water when he, in his treatises on public hygiene, wrote "consider the waters which the inhabitants use, whether they be marshy and soft, or hard and running from elevated and rocky situations, and then if saltish and unfit for cooking... for water contributes much to health." However, it is probably safe to say that hard and soft water in Hippocrates' writings are likely to mean different things than what we today call hard and soft water.

The definition of hard water varies slightly among researchers. The most frequently used classification system for hardness is that of Sawyer and McCarty (7), presented in Table 1. There is another slightly different classification system by Hem (8), which is referred to in some publications as the classification of the U.S. Geological Survey (9).

Based on Table 1, hard water is defined as water whose hardness is about 120 to 300 mg $CaCO_3$/l. However, some researchers argue that there is no exact definition for determining hard or soft water, because it depends, at least partially, on what one is accustomed to (9). What is hard water for one person, may be moderately hard for another and vice versa. The very hard water category in Table 1 can be grouped into the hard water category as it has the specifications of hard water. It is, however, unclear to which category one can assign moderately hard water, should hard water and soft water be the only categories available.

Table 1. Different Types of Water Based on Hardness

Water Type	Hardness, mg $CaCO_3$/l	
Soft	below 75[a]	Below 60[b]
Moderately hard	75–150[a]	60–120[b]
Hard	150–300[a]	120–180[b]
Very hard	Over 300[a]	Over 180[b]

[a] Reference 7.
[b] Reference 8.

John C. Manning (9, p. 228), a very well known authority in the field of hydrology for his Manning's roughness coefficient, succinctly yet nicely describes the characteristics of hard water; "Really hard waters, are unmistakable. If you have ever stayed in a motel supplied with hard water and tried to take a shower with a tiny bar of soap many motels supply their guests, you know how frustrating it is to see the soap disappear before you are probably lathered and bathed. On the other hand most people like water with a little hardness. It usually tastes better, and it washes the soap off when you bathe. Very soft water feels slick on your body, as though the soap were still there. Water with a little hardness, which is to say with calcium and magnesium ions, is favored for irrigation, too, as it tends to maintain higher soil permeability and allow infiltration to take place."

The explanations given above can be summarized as "those waters that require considerable amounts of soap to produce a foam or lather and that also produce scale in hot water pipes, heaters, boilers, and other high water temperature units, are called hard water." By the end of this article, the author confesses that the reader may become confused as what is hard water and what is hardness.

The economic disadvantages of hard water include increased soap consumption and higher fuel costs for boilers due to poor heat conduction resulting from scale formation. The term scale describes a continuous, adherent layer of foreign material formed on the waterside of a surface through which heat is exchanged (5). The thermal conductivity of $CaCO_3$ scale is as low as 0.0022 cal/s.cm. °C, 418 times lower than that of copper which is 0.92 cal/s.cm. °C. Hence, precipitation of $CaCO_3$ due to use of hard water significantly reduces the thermal conductivity of boilers. As a result, guidelines have been produced to set the limit for hardness of waters used in various industries; a few are presented as examples in Table 2. In a similar fashion, 80 mg/l of hardness is set as the limit for domestic water (10).

It is generally accepted that there is no health hazard associated with hard water (1). However, a recent study in Taiwan (11) shows that there is a significant negative relationship between drinking water hardness and colon cancer mortality, an important finding for the Taiwanese water industry and human health. This is in line with Tebbutt's (1) report that there is some statistical evidence to suggest that artificially softened waters may increase the incidence of some forms of heart disease. All these confirm Muss's (12) observation that in the United States, death rates from heart and circulatory diseases are lower in states where public water supplies are harder.

Table 2. Recommended Hardness Limit for Industrial Process Water[a]

Use	Hardness Limit, mg $CaCO_3$/l
Carbonated beverages	200–250
Confectionery	75
Dairy	180
Food canning and freezing	25–400 (depends on the type of food)
Food equipment washing	10
Food processing, general	10–250
Laundering	0–50
Paper and pulp, fine	100 (Ca hardness, 50)
Paper, groundwood	200
Paper, kraft, bleached	100
Paper, kraft, unbleached	200
Paper, soda, and sulfate pulps	100

[a] Reference 6.

WHAT TO DO WITH HARD WATERS

When needed, there are a few methods to change a hard water to a soft water to suit the purpose, an exercise called water softening (1):

1. *Lime softening*: This method involves adding lime and is suitable for calcium hardness of the carbonate form. In a typical experiment, total hardness and carbonate hardness of 150 and 125 mg/l are reduced, after softening, to 65 and 40 mg/l, respectively.
2. *Lime—soda softening*: This technique involves adding soda ash (Na_2CO_3), and is suited to all forms of calcium hardness. In a typical experiment, total hardness and carbonate hardness of 200 and 100 mg/l are reduced, after softening, to 40 and 40 mg/l, respectively.
3. *Excess—lime softening*: This is suitable for magnesium carbonate hardness. Chlorination may be unnecessary after such softening, but carbonation is necessary.
4. *Excess—lime soda softening*: This method involves adding lime and soda ash. It is a complicated process which is suited to all forms of magnesium hardness and produces a large volume of sludge.
5. *Ion exchange*: This method takes place through three approaches: (a) zeolite, (b) hydrogen cycle cation exchangers, and (c) strong ion exchanger. The Na^+ ions of zeolites replace Ca^{+2} and Mg^{+2} ions in the water, a reaction like ion exchange between Na_2clay and water. A natural sodium cycle zeolite has an exchange capacity of about 200 gram equivalents/m^3. Hydrogen cycle cation exchangers exchange all cations for hydrogen and so, produce zero-hardness water. A strong anion exchanger removes all ions.

HARDNESS OF WATER—WORLDWIDE CASES

Limestone and dolostone aquifers that contain gypsum and anhydrite in their matrix may yield hard water, whose hardness is from 200–300 mg/l or more (10). Groundwaters are generally higher in hardness than surface waters (4). The hardness of groundwater in Akwa Ibom State, Nigeria, ranges from 1.2–65 mg/l (13). In the Jornada del Muetro Basin of New Mexico, the groundwater hardness reaches 260 mg/l (14). The hardness of shallow groundwater in Balatina and Arpineni in the Republic of Moldovia reaches 55.5 mg-ecv/l (15), and in the Damman

Aquifer of Bahrain, reaches 3115 mg/l and an average of 1296 mg/l (16). The Mississippi River at Vicksburg has a hardness of 172 mg/l (2).

BIBLIOGRAPHY

1. Tebbutt, T.H.Y. (1995). *Principles of Water Quality Control*. Reprint of 4th Edn. Butterworth Heinemann, Oxford, UK.
2. Kemmer, F.N. (1988). *The NALCO Water Handbook*, 2nd Edn. McGraw-Hill, New York.
3. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
4. Hounslow, A.W. (1995). *Water Quality Data: Analysis and Interpretation*. Lewis, New York.
5. McCoy, J.W. (1981). *The Chemical Treatment of Boiler Water*. Chemical Publishing, New York.
6. Todd, D.K. (1980). *Groundwater Hydrology*, 2nd Edn. Jonn Wiley & Sons, New York.
7. Sawyer, C.N. and McCarthy, P.L. (1967). *Chemistry for Sanitary Engineers*, 2nd Edn. McGraw-Hill, New York.
8. Hem, J.D. (1985). *Study and Interpretation of the Chemical Characteristics of Natural Waters*, 3rd Edn. U.S. Geological Survey Water-Supply Paper 2254, Denver, CO.
9. Manning, J.C. (1997). *Applied Principles of Hydrology*. Prentice-Hall, Upper Saddle River, NJ.
10. Bell, F.G. (1998). *Environmental Geology, Principles and Practice*. Blackwell Science, Cambridge, UK.
11. Yang, C.Y. and Hung, C.F. (1998). Colon cancer mortality and total hardness levels in Taiwan's drinking water. *Arch. Environ. Contamination Toxicol.* **35**(1): 148–151.
12. Muss, D.L. (1962). Relationship between water quality and deaths from cardiovascular disease. *J. Am. Water Works Assoc.* **54**: 1371–1378.
13. Akujieze, C.N., Coker, S.J.L., and Oteze, G.E. (2003). Groundwater in Nigeria—a millennium experience—distribution, practice, problems and solutions. *Hydrogeol. J.* **11**(2): 259–274.
14. Schulze-Makuch, D., Goodell, P., Kretzschmar, T., and Kennedy, J.F. (2003). Microbial and chemical characterization of a groundwater flow system in an intermountain basin of southern New Mexico, USA. *Hydrogeol. J.* **11**(3): 401–412.
15. Melian, R., Myrlian, N., Gouriev, A., Moraru, C., and Radstake, F. (1999). Groundwater quality and rural drinking-water supplies in the Republic of Moldova. *Hydrogeol. J.* **7**(2): 188–196.
16. Zuber, W.K. (1999). The Damman aquifer in Bahrain - Hydrochemical characterization and alternatives for management of groundwater quality. *Hydrogeol. J.* **7**(2): 197–208.

AN ANALYSIS OF THE IMPACT OF WATER ON HEALTH AND AGING: IS ALL WATER THE SAME?

HOWARD REMINICK
Ohno Institute on Water and Health
Willoughby, Ohio

The human body was created to be in balance or return to balance. We call this homeostasis. Our body relies on its chemical processing plant to keep a metabolic balance. The body knows when it is being depleted of essential minerals and nutrients, which creates imbalance. When this happens, the body has to draw from its reserves. But this can only be temporary, until these reserves are replenished.

Our body knows when its energy level is low. If nothing is done, if we continue to ignore the body's call for help, we let our body begin on the path of its destruction. When the body is neglected, it becomes vulnerable to attack from invaders from the outside. It also is open to attack from the inside by free radicals. Bacteria, which are essential for our survival, suddenly turn on us and destroy us.

The body reacts to regain its balance by calling on the immune system. Sometimes it is too late if the body has been in a contaminated state too long, which can result from the poor quality of water we put in our body every day.

Water has the power to transfer information from cell to cell and bring back balance from imbalance. Water is vital to maintain life and organization of our cells. Organized cells are healthy, productive cells. To create harmony and balance, whether in our bodies or in the environment, water must also be balanced.

The body knows when it is drying out. The body can keep its ability to remain balanced only if it is given enough good water. Not just clean water, but water that can provide a good internal environment for the body to function and carry on its purpose to remain balanced or return to being balanced.

Water holds some of the most basic and, therefore, significant secrets of life. This is not a new idea. All ancient civilizations described this in their attempt to explain the structure of life. Even then as today, many believe that water within the cell has a "vital force." Cells are the fundamental units of life, so then they are responsible for all life processes and are the force that permeates all nature. If that force is available in any one substance, then it must be in water.

The earth is 70% water. Our body is 70% water. A cell is 70% water. DNA is 70% water. Is this just a coincidence? About 3 billion years ago, life was born in water. The first life form existed in pure, clean water. There is a direct connection between the quality, content and even the structure of the water in our body that can make the difference between sickness and health.

One of the most important books written, which explains the necessity of water for health is *Your Body's Many Cries For Water*, by F. Batmanghelidj, a medical doctor, who presents the most compete description, and further warns us that this relationship is not being addressed by general medical practice (1). He attributes the basis of all disease to dehydration and overly prescribed medications. He also warns that dehydration is a world epidemic, which can be controlled by attention to the body's hydration needs.

Understanding disease and aging begins with knowing what our body needs throughout its lifetime to maintain health and vitality. Premature aging can result from loss of water volume without adequate replacement or from contaminated water in our bodies. Because of environmental toxins and chemicals, food and water sources have been depleted of essential minerals. As a result, the body's chemistry becomes unbalanced, which creates a breakdown in the body's internal environment, making it susceptible to disease and rapid aging.

The body must receive at least 2.5 liters of water daily to replace loss. Although beverages such as coffee, tea, soda, and beer contain water, they do not count as hydration replacement. These beverages actually use water to be processed in our body, which increases dehydration. Only water can replace water in our body.

Symptoms of the body's deterioration begin to appear when the body loses 5% of its total water requirements. In a healthy, normal adult, this appears as fatigue and general discomfort. Although an adult can survive a 5% water loss, this can be fatal to an infant. In elderly persons, a 5% water loss may not be fatal, but the body chemistry becomes abnormal, especially the percent of electrolytes being overbalanced by sodium. Continuous water loss over time will speed up aging. Lost water must be replaced (Fig. 1).

Because water is an excellent solvent, vital nutrients are absorbed in the cells. The body must receive the best quality of water available to ensure that its internal environment is in the best operational state. This allows the body to be at its best to respond to distress and imbalance by making corrections, returning the body to homeostasis.

Every function such as digestion, absorption, elimination of waste products, and production of glucose to provide energy is maintained if the body has water that is not only pure, but bio-active. This bioactivity is created by magnetic influence throughout the body. Minerals absorbed in the body's water are carried throughout the body to become ionizing agents for all cells. The effect of biomagnetism on hemoglobin in the blood is critical. Hemoglobin consists of iron, which is the catalyst in its reaction to the body's biomagnetism in its role of stimulating nerve and muscle impulses.

It should now be evident that there are essential issues about the water in our body that must be considered if we are to remain as healthy as possible. First and foremost, all water is not the same. Although all water is H_2O, that is where it ends. And there is no way to fool the body. Water must be able to prevent toxins and chemical substances from accumulating and creating destructive influences in cells. Water must bring all minerals and nutrients required for metabolism into each cell and remove any substance that can damage the cell.

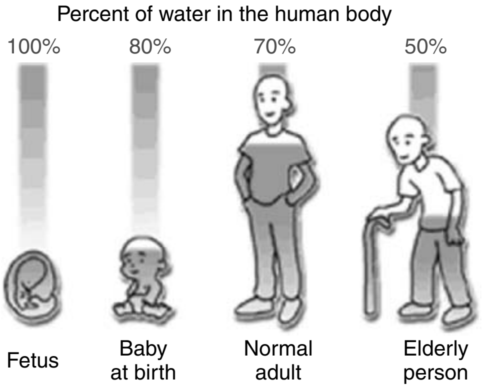

Figure 1. Percent of water in the human body.

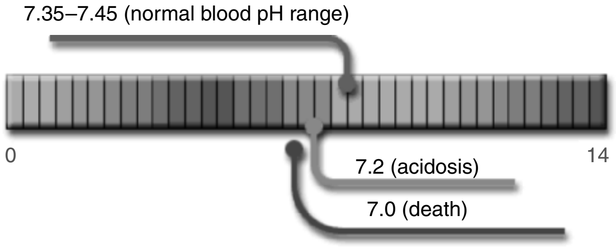

Figure 2. 7.35–7.45 (normal blood pH range).

Another critical need in the body is the pH balance of blood and body fluids. Normal blood pH has a very small window of normalcy (Fig. 2). It must range between 7.35 and 7.45. There must be an adequate amount of oxygen in the blood. Too little will diminish all organic function. Too much can cause oxidation, which is what happens when an apple is left exposed. Because of the ever-increasing levels of pollution and toxicity in our environment, which includes food and water supplies, our blood tends to be acidic.

THE INTERACTIVE EFFECT OF MAGNETIZATION AND WATER ON CELL BIOLOGY

Everything that happens in our bodies every day of our lives depends on the presence of magnetism. Magnetism in our body is created by the magnetic field, which dominates and controls the stability of our planet Earth. It also controls the stability of all life on Earth.

Does magnetization of the water do anything to change its basic physical properties? The answer is yes. The changes are very small but measurable. Magnetized water exhibits a change in the strength with which water molecules are able to bind with one another. Magnetized water can disrupt hydrogen bonding in water. Water, charged with magnetic energy, allows minerals and other organic substances to provide intermediate pathways of flow for magnetic energy to be directed to where it is needed in the body. The fact that magnetized water possesses altered binding forces between water molecules and various proteins, salts, and mineral substances appears to endow it with the ability to dissolve buildups of various types of salts and minerals.

There is a fascinating fact about the level of energy in moving water in our body and its effect on ionization. Because all living things are known to function largely by the influence of bio-magnetism, it seems reasonable to provide the body with water that is magnetically activated. Normally, water forms large molecular clusters, which slows movement and absorption into cells. This can be seen by microscopic analysis [most water molecules will aggregate around the rim of a slide (Figs. 3, 4, and 5). Naturally magnetized water will disperse throughout the slide, indicating free flow]. Because water molecules, which are not adequately magnetized, will lose order when they are not adequately magnetized, this can affect intracellular hydration.

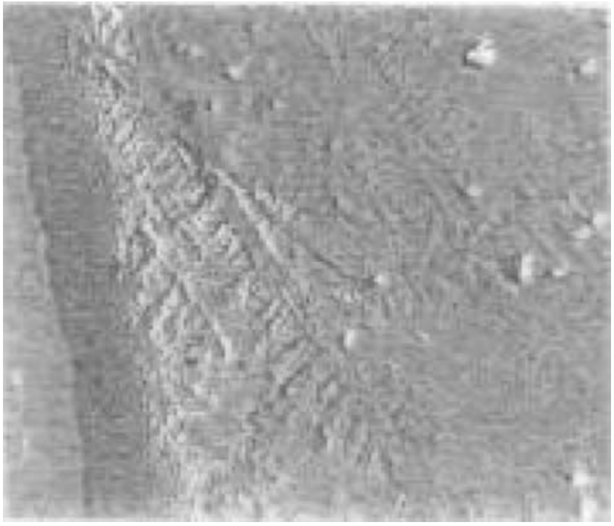

Figure 3. Microscopy photo of an unmagnetized water molecular structure. Note sparse formation of crystals and thick molecular clustering appearing as a mass and gathered at rim of slide.

Figure 4. Microscopy photo of naturally magnetized water. Note large number and density of crystals and individually fragmented molecules, remaining evenly dispersed on slide.

Water has long been known to be affected by magnetism. Magnetization can significantly reduce the surface tension of water and increase the pH, making it more alkaline. Water molecules are known to have a small interaction with magnetism, creating chains of organized molecules, which can influence the biochemistry of the water in the body. The intricate biological requirements needed to maintain electrochemical balance have been reported in books and journals about human physiology. Magnetically activated water may increase the percentage of intracellular hydration, which is a major factor in controlling the aging process and aging-related diseases.

RELATIONSHIP OF MAGNETISM TO CELL STABILITY

What is a living cell? It appears as one large molecule tied together by a matrix of water molecules. Like the atoms that compose them, they are tied together by ionic bonds, consisting of positive and negative charges on their outer surfaces. What creates these bonds?

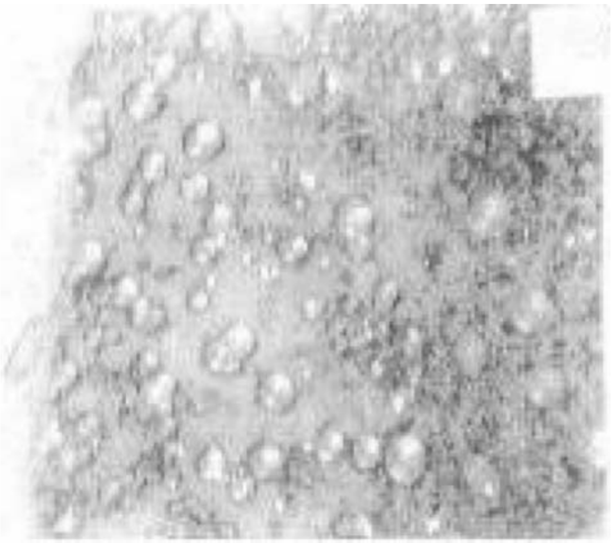

Figure 5. Microscopy photo showing detailed individual molecules of naturally magnetized water. They will not re-form in large clusters over time.

If salts like sodium chloride are added to water, the sodium and chloride ions (Na^+ and Cl^-) immediately become surrounded by water molecules to dissipate their charges. In doing so, adjacent water molecules take on the charges to form H_2ONa^+, which possesses the unique property of being able to transfer the small positively charged nucleus of one of its hydrogen atoms to a neighboring water molecule. Then the neighboring molecule becomes positively charged as H_3O^+, due to the extra hydrogen proton bonded to its oxygen atom. This charged molecule transfers one of its protons to another neighbor, cascading in a linear fashion to water molecules surrounding negatively charged ions. Thus, the stability of a particular molecule is the electromagnetic bond of energies between atoms.

Magnetism affects the electron stability of cells. Electrons of a cell are established in pairs. Think of an unstable element moving around the body, looking for an electron to make itself stable. It will steal an electron from wherever it can, at the expense of a stable cell, which needs paired electrons to remain stable. When a cell loses an electron from a pair, it will lose energy and die, sacrificing itself to the "thief," which is how bacteria and cancer strive. This is believed to be the activity creating free radicals, that is, strong, unstable magnetic activity as a result of a cell with electron imbalance aggressively seeking another electron, thereby creating imbalance in another cell.

One of the first things we learn about the complex molecules of living cells is the way in which their atoms are bonded together. Magnetism takes place when electrons orbit around their atoms. Electrons have a property called spin, which makes them act like miniature magnets with two opposite poles. If neighboring electrons spin so their poles are aligned in the same direction, they become magnetic. The electrons circling around atoms resist attempts to disturb their usual orbits, so they can create their own magnetic field.

The molecular parts which compose the cells of our bodies are much the same as those that composed the earliest forms of life on Earth. If we wish to comprehend the composition of life and its molecular roots and determine how these molecules move and flow in a spontaneous and miraculous fashion to sustain and perpetuate themselves, then, like nature, we must begin with water. Genetics defines which molecules are produced; water defines how they move and associate with each other.

Modern scientific technology has revealed the basis of molecular biology in the cell, its DNA, and how critical water is to its order. If DNA is dehydrated, like other bonding components of the cell, it loses its orderly, repeating structures and takes on more random shapes and forms. In other words, DNA loses its normal, functional properties.

It is now clear that all information needed to direct molecules into their functional shapes is contained in the molecules themselves in association with each other and the surrounding water medium. DNA, operating in harmony with water, determines shapes that are adopted as they are produced and move within the cell. It appears that DNA molecules themselves are the “code” and water the medium where the code is expressed.

As long as environment and temperature are maintained within the cell and there is adequate magnetization to keep cell biology functioning normally, then cellular reproduction, along with normal DNA coding, will continue. If, however, the content and effectiveness of the water inside and outside the cell contain destructive toxins, the degradation and deterioration of cells will increase quickly, resulting in rapid aging and disease.

THE EFFECT OF WATER ON CELLS

Fluid fluctuations affect blood volume and cellular function, which can be life threatening. When the amount of fluid within various cellular compartments is relatively constant, there is an exchange of solutes and water between compartments to maintain unique compositions. Individuals with more body fat have proportionately less total body water and are more susceptible to fluid imbalances that cause hydration.

The total body water percentage decreases with age, resulting in inadequate cellular hydration. Most critical is the decrease in intracellular fluid, which is due, in part, to increased fat and decreased muscle, and decreased ability to regulate the sodium and water balance. With age, kidney function becomes less efficient in producing urine, and responses for conserving sodium weaken.

There are two major issues about the water in our body that must be resolved if the body is to remain in a state of harmony and stability. First, the water must be able to prevent toxins and chemical substances from accumulating and creating destructive influences in the cells. Water must bring all minerals and nutrients required for metabolism into the cell and remove any substances that can damage the cell. Also, the cell wall must be protected from damage and invasion. Second, water is involved in every function of the body, so it must act as a conductor of electrochemical activity, such as neurotransmission, by moving water from one nerve cell to another smoothly and effectively.

When cells are studied under magnetic resonance imaging (MRI), water is observed to act very differently from when it is not exposed to a magnetic field. Molecules are re-formed in smaller clusters in a linear arrangement. The molecules are lined up and move in and out of the cells easily. In contrast, when water is photographed after removal from MRI exposure, molecules are randomly ordered and cluster with neighboring water molecules to form large, molecular clusters. (Interestingly, when cancer cells are studied under MRI influence, their water molecules are also organized and calm, in contrast to their aggressive, violent movement without MRI exposure.) However, naturally magnetized water remains organized in small clusters, whether exposed to MRI or not.

THE PASSAGE OF WATER IN AND OUT OF CELLS

Movement of water occurs because of hydrostatic pressure and osmosis in capillaries. As water flows from artery to vein, changes in pressure create movement across membranes. Any changes in pressure allow proteins, minerals, and other nutrients carried by the blood to escape into spaces between vessels and deprive the cells of vital needs to sustain life. In addition, when the water in the blood is acidic and contaminated with chemicals, it enters the cells and changes their structure, including changes in DNA, which is how disease begins and attacks structures in the body.

Osmosis is the diffusion of water across a membrane that permits the free passage of water but not solutes. Molecules of solutes (sugars, amino acids, and ions) dissolved in water are in constant motion, which creates diffusion, the random movement toward zones where the concentration of the molecules is lower. This tends to distribute molecules uniformly. Because of diffusion, both solvent (water) and solute molecules in a cell will move from areas where concentration is greater to an area where their concentration is less.

As water molecules continue to diffuse inward toward the area of lower concentration of unbound water molecules, pressure within the cell increases, which is called hydrostatic pressure. This is the force that must be applied to stop osmotic movement of water across a membrane. When two areas are separated by a membrane, what happens depends on whether the molecule can pass freely through the membrane; most solutes in cells cannot do so.

Solutes are water soluble and are kept within the cell. Water molecules, however, can diffuse across the membrane into the cell, thus diluting the high concentration of solutes within the cell so that it matches more closely the lower concentration in the solution outside the cell (extracellular matrix). Eventually, however, the cell will reach an equilibrium, and the osmotic force driving water inward is counterbalanced exactly by the hydrostatic pressure driving water out. As a result, the water and solute concentrations will be the same in both compartments.

The nature and proportion of solutes in the body's water determine its effectiveness in hydrating and detoxifying cells so that a homeostatic, internal environment can be maintained to ensure proper cell metabolism, which requires bioenergy. A key component in dispensing bioenergy is the interactive effect of water and biomagnetism. This interactive bioenergy is seen in the effectiveness of osmosis and the osmolarity of the water in the extracellular matrix, as well as in the cells. Through the influence of the body's bio-magnetism, greater penetration of ions through the cell membrane, transported by water, will increase the health and stability of cells.

Osmosis is created by a magnetic force in our body, which keeps the movement in balance. Osmotic activity occurs as a result of osmolarity, the total concentration of all solute particles in a solution. The extent to which water's concentration is decreased by solutes depends on the number, not the type, of solute particles in a solution, because one molecule or one ion of solute displaces one water molecule. Thus, a solution's osmolarity is a count of the number of solute particles (including ions) dissolved in one liter of water and is seen in the solution's ability to cause osmosis. The measurement used to describe the solute concentration of body fluids such as water is the milliosmol. The normal osmolality of body fluids should average around 270 milliosmoles.

Osmolarity controls the distribution and movement of water between body compartments. The osmolality of intracellular and extracellular fluid tends to equalize and, therefore, provides a measure of body fluid concentration and the body's hydration status. Osmolarity measures the effective change in pressure for water, assuming that all the solutes are impenetrable. When a solute is added to water, the volume is expanded and includes the original water plus the volume occupied by the solute particles. In measuring osmolarity, the volume of water is, therefore, reduced by an amount equal to the volume of added solute.

When equal volumes of solutions of different osmolarity are separated by a membrane that is permeable to all molecules in the system, net diffusion (tendency of the molecules or ions to scatter evenly) of both solute and water occurs, each moving down its own concentration gradient. A solution that changes the tone or shape of a cell by altering its internal water volume is described by its osmotic activity, or tonicity. Tonicity is the tendency of a solution to resist expansion of intracellular fluid volume.

Solutions with concentrations of nonpenetrating solutes equal to those found in cells are isotonic, or have the same tonicity. An isotonic solution has the same osmolality as the intracellular fluid (ICF) or the extracellular fluid (ECF), which is required to maintain normal cell metabolism. An example of an isotonic solution is 5% dextrose in water and a normal saline solution of 0.9% NaCl. This solution is isotonic because there is no net movement of the active particles in the solution into cells and the particles are not metabolized. Extracellular fluids and most intravenous solutions are isotonic.

Solutions that contain fewer nonpenetrating solutes are called hypotonic. Cells placed in hypotonic solutions plump up rapidly as water rushes into them. Distilled water represents the most extreme example of hypotonicity. Because it contains no solutes, distilled water pulls water into cells without proper resistance or osmotic pressure and causes them to swell or burst.

A hypertonic solution has a concentration greater than 285 to 294 milliosmoles. An example is a 3% saline solution. Water is pulled out of the cells so that the cells shrink. Tonicity is critical to cell metabolism when correcting water and solute imbalances by administering different types of replacement solutions, such as those used in IV therapy solutions.

Balanced osmolarity can be the difference between the life and death of an organism. An example is seawater, overbalanced with sodium ions, or distilled water, which can be deadly if infused because it has no osmolarity. If distilled water is present on both sides of a permeable membrane, no net osmosis occurs, even though water molecules continue to move in both directions through the membrane. If, however, the solute concentration on both sides of the membrane differs, water concentration differs as well, because as the solute concentration increases, water concentration decreases.

There is a consideration when determining osmotic activity. In living plant cells, as water diffuses into the cell, the point is reached where the hydrostatic pressure (the back pressure exerted by water against the membrane) within the cell is equal to its osmotic pressure (tendency to resist further net water entry). Generally, the higher the amount of nonpenetrable solutes within the cell, the higher the osmotic pressure and the greater the hydrostatic pressure that must be present to resist further net water entry. Osmotic pressure is also necessary to control water loss.

However, these pressure changes do not occur in living animal cells, which lack rigid walls and are surrounded only by flexible plasma membranes. Osmotic imbalances cause animal cells to swell or shrink, due to net water gain or loss, until the solute concentration on both sides of the plasma membrane is the same or the membrane is stretched to its limit. Many molecules, particularly intercellular proteins and selected ions, are prevented from diffusing through the plasma membrane. Any changes in their concentration produce changes in water concentration on the two sides of the membrane and result in net loss or gain of water by the cell.

THE EXTRACELL MATRIX AND WATER

Our cells are continuously bathed in an extracellular fluid called interstitial, found outside the cell. Because it contains such a large variety of ingredients, such as amino acids, sugars, fatty acids, vitamins, hormones, neurotransmitters, and salts, it is sometimes referred to as the "soup" that feeds the cells. It is also the dumping ground for waste products removed from the cells. To remain healthy, each cell must absorb substances from this "soup" in the exact amounts it needs at specific times and reject the rest.

The human body can maintain homeostasis only if the extracellular fluid (ECF) volume is closely regulated to maintain blood pressure and salt balance. Its osmolarity must be regulated to prevent swelling or shrinking of the cells. The maintenance of water balance is important in regulating ECF osmolarity. To maintain stability, the ECF must respond quickly to changes in volume and osmolarity (2).

In humans, kidneys can spare water by eliminating wastes in concentrated urine. This urinary concentration relies on water reabsorption driven by an osmotic gradient formed along the collecting ducts, consisting of NaCl and urea at similar osmolarities. Survival of renal cells is induced by NaCl, but not urea, because urea is a membrane-permeant solute. With excess saline in the tissues, there is a decrease in water excretion to bring the salt concentration back to normal, which allows the cell concentration to remain constant and prevent shrinking by fluid moving out of the cells to reach equilibrium within the cells.

The environment of a cell is directly affected by the environment outside the cell. Cells have a reciprocal relationship with their environment. The environment surrounding a cell consists of a structured, basic substance, called the extracellular matrix (ECM). This enormous space of connective tissues has a significant effect on determining genetic outcomes of the cell. Water is its primary regulation system. The ion composition of the elements in the extracellular space is activated in water.

Understanding how a cell maintains its integrity requires understanding the role of the extracellular matrix in the body in the health and longevity of its internal environment. Most body cells are in contact with a jelly-like substance composed of proteins and polysaccharides. These molecules are secreted by the cells into an organized mesh in the extracellular matrix, where they serve as a sort of glue that helps hold cells together. In reference to the cells, this can be viewed as the mortar between bricks of a wall. Depending on the time of day, this "mortar" can be gelatin-like or liquid because of the impact of the sympathetic and parasympathetic nervous systems, which perform different functions at different times of the day.

During the 12-hour "dumping cycle" (3:00 A.M. to 3:00 P.M.), controlled by the sympathetic nervous system, the "mortar" is gelatose to remove and retain wastes from cells and is acidic. During the 12-hour "feeding cycle" (3:00 P.M. to 3:00 A.M.), controlled by the parasympathetic nervous system, the "mortar" is liquid to transport nutrients to cells and is alkaline. It is the effect of magnetically activated water in the extracell matrix that assures an optimal delivery system and that these cycles function properly with pH balance (Fig. 6).

The ECM is the largest system in the body. It is virtually responsible for all vital functions that take place within the body. All cells depend on an intact ECM for their existence. It regulates the cell's environment and is part of every aspect of the immune system because the lymphatic system is connected with it. It takes care of cell nutrition through internal circulation of water and

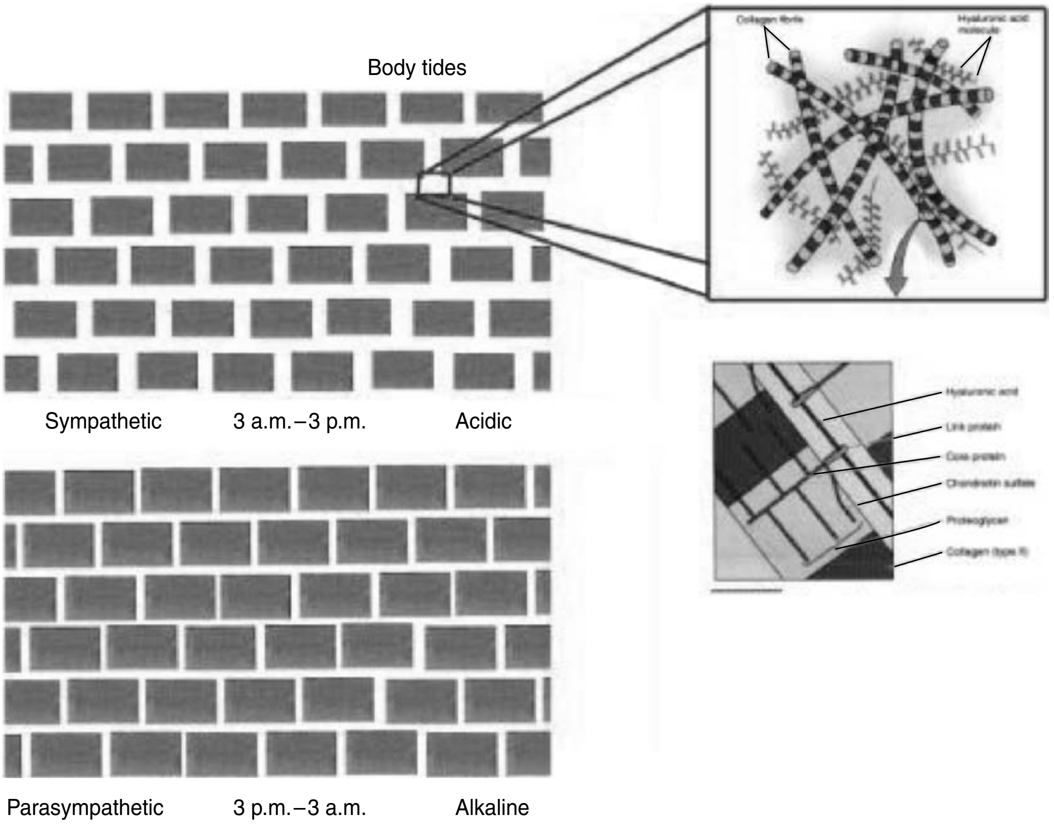

Figure 6. ECM Cycle.

removal of wastes. Disease originates in dysfunctions in this space, as it affects organs and other systems throughout the body.

More than having a supporting and space-filling function, the ECM provides regeneration for specific organ functions, as well as mediation of nerve and vascular functions. Everything that comes out of the blood takes a complicated route through the connective tissue to the cell and then into the lymphatic system. It is even more vital to the health of the human body because the nerves and blood vessels do not come into direct contact with the cells at any point; therefore, the ECM becomes the mediator for nervous and nutritional flow. As connective tissue is really a living part of the body, everything attributable to the life and health of the body, and all its basic functions, has to be attributed to the extracellular matrix.

The extracellular matrix provides the communication of connective tissue over the entire organism. The entire field of activity and information related to communication is the extracellular fluid. Connective tissue is found everywhere between the bloodstream and the principal structural parts. Everything that comes out of the blood takes a complicated route through the connective tissue to the primary, functioning cell and then into the lymphatic system, which is connected with it. It takes care of cell nutrition through internal circulation of water and the removal of wastes. It regulates the cell's environment and is part of every defense process.

The extracellular matrix is connected to the endocrine gland system via capillaries and the central nervous system via the peripheral nerve endings. Both systems are connected to each other in the brain stem, so superior regulatory centers can be influenced by the ECM. The advantage of the interconnection of these systems results in a significant increase in the performance capacity of the ECM, which can be attained only by highly functional, interconnected sets of systems, functioning as one.

When looking for the source of disease and aging, we must first look at the water in our bodies. I would like to repeat the statement that led me to search for the causes of aging diseases: If the body is composed of 70% water, then the condition of the water in our body can determine how well our body maintains its state of health. The structure of all living tissue is predominately water, a plain and simple substance. Knowing the importance of water in the body can help us learn how to keep the body from breaking down. Further, understanding that a living organism is composed mostly of water should help us better appreciate the need to supply our body with the best water possible.

The signature of life is found in the water we drink every day. A body can remain in a state of health only if it is given enough water, not just clean water, but water that has the structure, properties, and activity to provide a good internal environment for a body to function at its best. We become sick, not just from dehydration, but from the inadequacy of the water we drink. The question becomes, "is all water the same?" The answer is a definite no. Only water that is bioactive with magnetic energy can provide ideal hydration.

After years of investigation on the biological effects of magnetized water on human health, it became apparent that all magnetized water is not the same, nor does each behave the same in the body. Water that is magnetized with devices displays structural changes initially but will return to its original formation because every water source is unique; each has its own frequency, its own signal, due to its unique memory. Artificially altering water, therefore, will not create a permanent structural change. In all of our studies, only water that is naturally magnetized, which creates the water's memory, will perform as the optimal hydration source consistently and aid the body in maintaining a better internal environment, resulting in better health and more natural aging.

To learn more about the phenomenon of naturally magnetized water, go to http://www.ohno.org.

BIBLIOGRAPHY

1. Batmanghelidj, F. (1997). *Your Body's Many Cries for Water*. Global Health Solutions, Vienna, VA.
2. Sherwood, L. (1997). *Human Physiology: From Cells to Systems*, 3rd Edn. Wadsworth, Belmont, CA.

READING LIST

Abrams, A. (1913). *Human Energy*. Borderland Sciences, San Francisco, CA.

Ackerman, J.M. (2001). The Biophysics of the Vascular Autonomic Signal and Healing, *New Perspectives, Journal of the Center for Frontier Sciences of Temple University* **10**(2).

Collins, J.C. (2000). *Water: The Vital Force of Life*. Molecular Presentations, Kinderhook, NY.

Collins, J.C. (1991). *The Matrix of Life*. Molecular Presentations, East Greenbush, NY.

Davis, R.D. and Rawls, W.C., Jr. (1996). *Magnetism and Its Effects on the Living System*. Acres, Metairie, LA.

Dubrov, A.P. (2001). Russian Advances in the Study of the Effects of Environmental Factors and Electromagnetic Medicine. *New Perspectives, Journal of the Center for Frontier Sciences of Temple University* **10**(2).

Ganong, W.F. (1993). *Review of Medical Physiology*, 16th Edn. Appleton & Lange, Norwalk, CT.

Gerber, R. (1996). *Vibrational Medicine*. Bear, Santa Fe, NM.

Marieb, E.N. (1995). *Human Anatomy and Physiology*, 3rd Edn. Benjamin/Cummings, Redwood City, CA.

McCance, K.L. and Huether, S.E. (1994). *Pathophysiology: The Biologic Basis for Disease in Adults and Children*. Mosby, St. Louis, MD.

Oschman, J.L. (2001). Energy Medicine: The Scientific Basis. *New Perspectives, Journal of the Center for Frontier Sciences of Temple University* **10**(2).

Oschman, J.L. (2000). *Energy Medicine: The Scientific Basis*. Churchill Livingstone, Edinburgh, UK.

Osyka, F. and Edmonds. (1991). *Critical Necessity of Negative/Positive Ion Balance: The Ion Effect*. Alpine Industries, Greeneville, TN.

Philpott, W.H. and Kalita, D.K. (2000). *Magnet Therapy*. Alternative Medicine.com Books, Tiburon, CA.

Pischinger, A. (1975). *Matrix and Matrix Regulation*. Haug International, Brussels.

HEAVY WATER

Gholam A. Kazemi
Shahrood University of
Technology
Shahrood, Iran

> Heavy, heavy water
> Won't wash away the sins of the father
> Unholy, holy water
> Leading us like lambs to the (slaughter)
>
> James Young
> (URL Ref. http://www.styxnet.com/styxlyrics/hw2.htm)

INTRODUCTION

This article should be studied in conjunction with the closely related articles on Deuterium and Isotope Fractionation. In some ways, it might have been slightly more appropriate if the Deuterium and Heavy Water articles were combined into one. For example, the whole introduction and some parts of the text of both articles overlap, making it more or less difficult to draw a sharp boundary between the two.

Each of the atoms of hydrogen and oxygen, which form the water molecule, has three isotopes; ^{1}H, ^{2}H (D), ^{3}H (T) for hydrogen and ^{16}O, ^{17}O, and ^{18}O for oxygen. The average terrestrial abundances of these isotopes are given in Table 1 (1).

The abundances of these isotopes in the hydrosphere, biosphere, and lithosphere and in different stages of the water cycle vary slightly from these figures.

Considering the six isotopes of hydrogen and oxygen, 18 combinations of H–O–H or "18 varieties of water" are possible (2,3). These and their respective molecular weights are shown in Table 2.

Table 1. Average (%) Terrestrial Abundances of Hydrogen and Oxygen Isotopes[a]

^{1}H (Proteum - stable)	99.985	^{16}O (stable)	99.762
^{2}H (Deuterium - stable)	0.015	^{17}O (stable)	0.038
^{3}H (Tritium -radioactive[b])	$< 10^{-14c}$	^{18}O (stable)	0.200

[a] Reference 1.
[b] $t_{1/2} = 12.3$ years
[c] Some references give 10^{-14}–10^{-16}.

Of the above eighteen varieties of waters, $^{1}H_2{}^{16}O$ has a mass of 18 and is the most common; it is generally known as "water" or sometime called ordinary or light water. Deuterium oxide ($D_2{}^{16}O$) has a mass of 20 and is known as "heavy water." Some references use the term "semiheavy water" for $HD^{16}O$. Washburn and Urey (4) were the first to enrich deuterium oxide in water in 1932. Some of the 18 varieties of water are extremely rare. An interesting, though trivial, point is that the molecular weight of the most common form of water $^{1}H_2{}^{16}O$ (18), equals the number of different varieties of water, 18.

CHARACTERISTICS OF HEAVY WATER

The physicochemical properties of heavy water and deuterium are compared with those of water and hydrogen in Table 3.

The extent of dissociation of heavy water is about half that of ordinary water:

$$D_2O \rightarrow D^+ + OD^-$$

and

$$H_2O \rightarrow H^+ + OH^-$$

Table 2. Eighteen Varieties of Water and Their Respective Masses

$H_2{}^{16}O$	18	$D_2{}^{16}O$	20	$T_2{}^{16}O$	22	$HD^{16}O$	19	$HT^{16}O$	20	$DT^{16}O$	21
$H_2{}^{17}O$	19	$D_2{}^{17}O$	21	$T_2{}^{17}O$	23	$HD^{17}O$	20	$HT^{17}O$	21	$DT^{17}O$	22
$H_2{}^{18}O$	20	$D_2{}^{18}O$	22	$T_2{}^{18}O$	24	$HD^{18}O$	21	$HT^{18}O$	22	$DT^{18}O$	23

Table 3. Physicochemical Properties of Heavy Water, Water, Deuterium, and Hydrogen[a]

Characteristics	Heavy water	Water	Deuterium	Hydrogen
Density at 25 °C (g/ml)	1.1045	0.99707		0.0899
Boiling point, °C	101.42	100	−249.55	−252.77
Freezing point, °C	3.8	0	−254.5	−259.2
Molecular volume at 20 °C	18.092	18.016		
Viscosity, poise at 20 °C	12.6	10.09		
Solubility of NaCl, g/100 g	30.5	35.9		
Critical pressure, atm			16.432	12.807
Evaporational enthalpy, kJ/mol	46.5	40.7		
Dielectric constant	78.25[b]	78.54[b]	1.277[c]	1.228[c]
Ionization potential, eV	12.6	12.6	15.46	15.427
Bond length at 298 K, Å	104.45	0.958 angle	0.7416	0.7461
Bond strength kJ/mol	—	Diatomic	443.546	436.002

[a] References 5–7. Note that in many references, the boiling and freezing points of H_2 and D_2 are given in K.
[b] At 20°K.
[c] At 25 °C.

This means that a bond to deuterium breaks 18 times more slowly than the same bond to hydrogen (6). This is the reason that the vapor pressure of heavy water is lower than that of ordinary water.

The amazing behavior of heavy water is its inertness in normal metabolic processes. For example, rats will die of thirst if given heavy water exclusively, and plants will not grow in pure heavy water (8). Experiments with mice showed that heavy water, due to a slightly different reaction rate compared to ordinary water, inhibits mitosis and causes progressive damage of tissues that need rapid regeneration. Human drinking of up to 25% heavy water is not harmful, but the main problem is consequent lack of enough ordinary water in the body. After several days of ingesting only heavy water, the body fluids contain about 50% heavy water, and at this point symptoms begin due to the decrease in cell division rates of rapidly dividing tissues, such as hair roots and stomach linings. Drinking heavy water may lead to the remission of aggressive cancer cell, but the, effect predicted is not great enough to make this a useful therapy (URL Ref. http://en2.wikipedia.org/wiki/Heavy_water#Toxicity).

In slight contrast to these publications, Kushner et al. (9) report that, so long as it has not been used in nuclear reactors, D_2O itself poses little risk to the environment because:

1. D_2O itself is not toxic to animals, except under exceptional and very improbable circumstances;
2. D_2O is nonradioactive; and
3. The amount of the ^3H, found in heavy water is very low, <6 kBq/kg D_2O (<0.15 μCi/kg D_2O), barely above background levels.

More information on the biological aspects of heavy water may be found in Katz (10). Dahl (11) provides good coverage of many aspects of heavy water, including discovery, uses, production techniques, its relationship to developments in nuclear physics and chemistry, social history, and a chronology of events relating to heavy water.

APPLICATIONS OF HEAVY WATER

Heavy water is used (1) as a neutron moderator, (2) to detect neutrinos, and (3) in the biology/biotechnology field. Biotechnological applications of heavy water are in the early stages of research and have not yet received widespread attention.

Neutron Moderator

The poem from James Young, prefacing this article, is directed toward this application of heavy water, its use in nuclear reactors. In nuclear reactors, an important factor in producing a chain reaction in uranium is the energy distribution of neutrons, which peaks near 1 MeV. To reduce the neutrons energies from 1 MeV to thermal energies (so they can react with the uranium in the reactor) would require at least 2000 collisions (with something else). The most common materials used to slow down neutron energy are heavy water and graphite (12). A huge nuclear reactor that uses heavy water as a moderator is CANDU in Canada (Fig. 1). The advantage of heavy water is that it has a low probability of absorbing neutrons. Ordinary water also acts as a moderator but because light water absorbs neutrons, reactors using light water must use enriched uranium rather than natural uranium; otherwise criticality is impossible. A nation that has a sufficiently powerful heavy water reactor can use it to turn uranium into bomb-usable plutonium, entirely bypassing uranium enrichment and all of the related technological infrastructure. Heavy water reactors have been used for this purpose by India, Israel, Pakistan, and North Korea. Due to its usefulness in nuclear

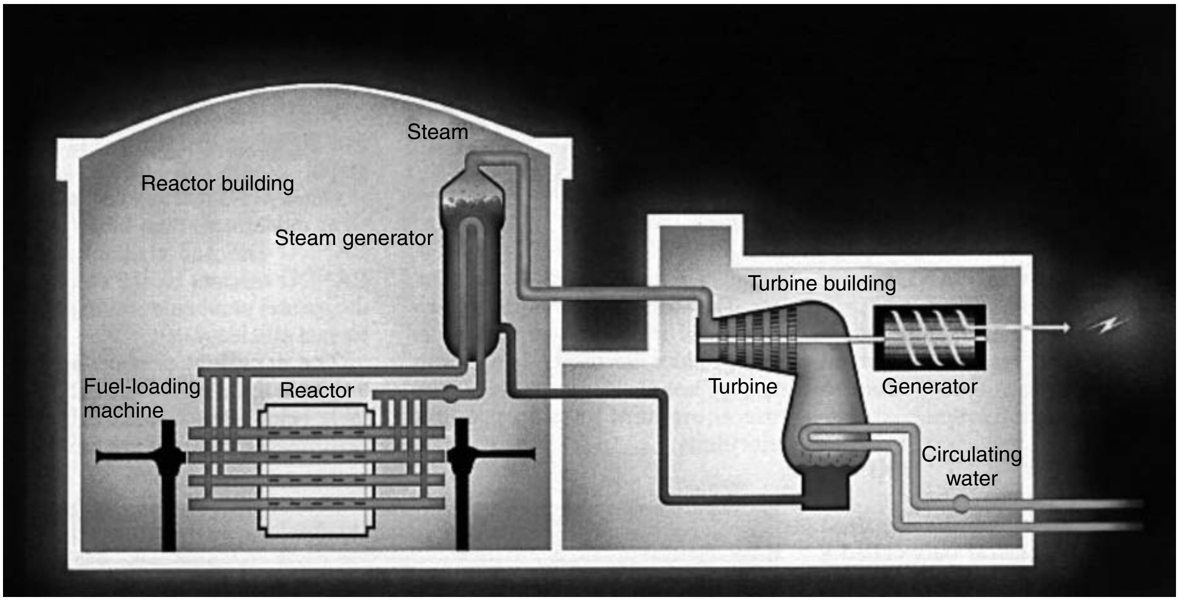

Figure 1. The layout of the CANDU heavy water reactor in Canada (URL Ref http://www.nucleartourist.com/type/candu.htm).

weapons programs, heavy water is subject to government control in several countries. In addition, heavy-water-moderated reactors can be used to make tritium (URL Ref. http://www.fas.org/nuke/intro/nuke/heavy.htm).

Neutrino Detector

The neutrino, discovered in 1956, is an elementary particle that has no charge, zero rest mass and travels at the speed of light (13). The detection of this particle is a difficult and complicated task. The Sudbury Neutrino Observatory (SNO), near Sudbury, Ontario, is a heavy-water Cherenkov detector that is designed to detect neutrinos produced by fusion reactions in the Sun. It is located 6800 feet underground in an old mine (to shield it from cosmic rays) and uses 1000 tonnes of heavy water on loan from Atomic Energy of Canada Limited (AECL), contained in a 12-meter diameter acrylic vessel. Neutrinos react with heavy water (D_2O) to produce flashes of light called Cherenkov radiation. This light is then detected by an array of 9600 photomultiplier tubes mounted on a snoball surrounding the heavy water vessel. The detector laboratory is extremely clean to reduce background signals from radioactive elements present in the mine dust which would otherwise hide the very weak signal from neutrinos (http://sps.la.asu.edu/~steve/paper/sno/node6.html).

Biological and Biotechnological Applications of Heavy Water

Kushner et al. (9) describe the potentially extensive use of heavy water and deuterated compounds in biotechnological applications. They report as follows:

1. Many microorganisms can grow in "pure" (99.6–99.8%) D_2O; usually, after a period of adaptation, they can produce a very large number of deuterated compounds that have many applications in optical technology, in deuterated drugs and fibers, etc.
2. Deuterated glycerol has interesting potential uses as a precursor of transparent plastics and other compounds and can be formed by *Dunaliella* spp, a salt-tolerant alga that grows in D_2O.
3. Heavy water has some direct therapeutic uses in both animals and humans. Nontoxic levels of heavy water, for instance, reduce induced hypertension and associated changes in rats.
4. Heavy water has been widely used in measuring body water space in children and in adults at various stages of nutrition and exercise.
5. Applications of deuterium in environmental studies are numerous. (These are dealt with in the Deuterium article.)

PRODUCTION OF HEAVY WATER

Heavy water occurs naturally in limited, yet precisely unknown, quantities. Production of heavy water includes separating these limited quantities (not actually making them) from other ordinary water molecules. In the hydrosphere, one in every 6700 molecules of water contains deuterium in its composition. If ^{16}O is the other atom of all of these "one in 6700" molecules, then the ratio of heavy water/water in the hydrosphere is 1/6700. This ratio, though most likely realistic, may be smaller because some of the oxygen atoms in water molecules that contain deuterium are ^{17}O and ^{18}O. In other words, not all deuterium isotopes join with ^{16}O to form heavy water; some are combined with ^{17}O and ^{18}O to form other varieties of water depicted in Table 2.

The discovery of deuterium in 1932 by Urey and his colleagues (14), may be considered the onset of the artificial production of heavy water. However, Gilreath (15, p. 61) in this regards, writes, "In 1933, G. N. Lewis and his collaborators at the University of California succeeded, by the continuous electrolysis of a very large amount of water obtained from an old commercial electrolytic cell, in preparing about 0.1 cc of water which contained practically only deuterium atoms of mass 2. The specific gravity of this water was about 10% greater than that of ordinary water which was in accordance with the molecular weight of D_2O." This paragraph may indicate that the *first drop* of heavy water was produced in 1933.

There are a few methods for artificially enriching the deuterium content of ordinary water and, thus, producing heavy water. These include distillation, electrolysis, and chemical exchange.

Distillation

The foundation of this method is based on the difference between the vapor pressures and boiling points of H and D. This was the principle used by Urey to discover deuterium. The most efficient process for enriching heavy water by this method uses a fractionating column in a continuous process, in which an upward-directed stream of vapor is kept in contact with a downward-directed stream of liquid; the two streams constantly exchange molecules. The molecules of H_2O have a relatively greater tendency to enter the vapor stream, and the molecules of D_2O enter the liquid (15). In 1943, the first plant based on distillation was built to produce heavy water for use as a moderator in nuclear reactors. However, the plant operated for less than a year only because (1) it was found that the chemical exchange method is more efficient and (2) it was learned that graphite is superior to D_2O as a moderator (15).

Electrolysis

The principle of this method, a common method for laboratory scale production of heavy water and described in many chemistry textbooks, is simple. Passage of an electric current through water leads to splitting water into hydrogen and oxygen. The deuterium atoms move more slowly because they are heavier and, so, it is possible to separate them. Finally, the resulting deuterium is reacted with oxygen to form heavy water. The electrolytic technique consists of up to seven stages, depending on the purity of heavy water required; the higher the purity of the required heavy water, the greater the number of stages of the process. This technique takes a vast amount of electrical energy. In a typical experiment, 2400 liters of ordinary water produce 83 ml of D_2O that is 99% pure (6).

Chemical Exchange

Minute differences between chemical properties and the reaction rates of hydrogen and deuterium form the foundation of this method. In this process, water flows from the top of a tower that contains a catalyst, downward to meet a mixture of steam and hydrogen moving upward. The catalyst promotes a rapid isotopic exchange reaction between the hydrogen and the water molecules in the steam; hence, there is a concentration of deuterium in the steam in the form of heavy water molecules. The flow of water from the top of tower causes condensation of the steam and, consequently, enrichment of deuterium in the outflow from the bottom of the tower. The outflow is then electrolyzed to increase its deuterium content and it is, then, repeatedly reprocessed in the system to achieve a high concentration of D_2O (15). The refined/advanced version of this method is now known as the Girdler Sulfide process (9—see the following section about the Bruce Heavy Water Plant for more information about this method).

WORLDWIDE PRODUCTION OF HEAVY WATER (THE FOLLOWING INFORMATION HAS BEEN GATHERED FROM WEB SITES)

As can be inferred from the previous explanations, a huge technical infrastructure is required to produce heavy water, and, thus, only a few countries such as Argentina, Canada, India, and Norway have the capacity to produce significant amounts.

The first commercial heavy water plant was the Norsk Hydro facility in Norway (built in 1934, capacity 12 metric tons per year); this is the plant which was attacked by the Allies to deny heavy water to Germany during World War II.

India is the world's second largest producer of heavy water through its Heavy Water Board of the Department of Energy (URL Ref: http://www.heavywaterboard.org), but accidents and shutdowns may have led to limitations on its production rate.

The Bruce Heavy Water Plant in Ontario, Canada (Fig. 2), built in 1979, has a capacity of 700 metric tons/year; it is the world's largest producer of D^2O (1979; 700 metric tons/year). It uses the Girdler Sulfide (GS) process which incorporates a double cascade in each step. In the upper ("cold," 30–40 °C) section, deuterium from hydrogen sulfide preferentially migrates into water. In the lower ("hot," 120–140 °C) section, deuterium preferentially migrates from water into hydrogen sulfide. An appropriate cascade arrangement actually accomplishes enrichment. In the first stage, the gas is enriched from 0.015% deuterium to 0.07%. The second column enriches this to 0.35%, and the third column achieves an enrichment between 10% and 30% deuterium. This product is sent to a distillation unit for finishing to 99.75% "reactor-grade" heavy water. Only about one-fifth of the deuterium in the plant feed water becomes heavy water product. The production of a single pound of heavy water requires 340,000 pounds of feed water. Heavy water produced by Bruce is used in CANDU reactors.

Figure 2. Bruce Heavy Water Plant in Ontario, Canada (Toronto Star, Online Edition, Dec. 2, 2003).

The Bruce Power Plant, which presently operates in the shutdown mode, was going to be decommissioned a few years ago, but its license was renewed in 2002 for 2 more years (Nuclear Energy Overview, 4 November 2002 cited in World Nuclear Association, WNA, News Briefing 02.45, 30 Oct to 5 Nov. 2002). The latest news headline about Bruce Power Plant, appeared in Toronto Star online edition, Dec. 2, 2003:

"Safety commission finds 'serious event' at Bruce

A defect in the backup shutdown system of a reactor at the Bruce B nuclear generating station went undetected for more than 3 weeks in December, according to a newly released report by the Canadian Nuclear Safety Commission."

There are a few website publications which claim that some countries, Iran, for example, are pursuing plans for heavy water reactors. One such report is that by J. Boureston and C. Mahaffey from Janes International Security which can be found at: http://www.janes.com/security/international_security/news/jir/jir031114_1_n.shtml. There are also locality maps of various heavy water reactors and production plants, photographs, statistics, and many other heavy-water-related publications available at various websites. A useful site is that of the Heavy Water Board of India at http://www.heavywaterboard.org.

BIBLIOGRAPHY

1. IUPAC. (1992). Isotope compositions of the elements-1989. *Pure Appl. Chem.* **63**: 991–1002.
2. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
3. Hem, J.D. (1985). *Study and Interpretation of the Chemical Characteristics of Natural Waters*, 3rd Edn. U.S. Geological Survey Water-Supply Paper 2254, Denver, CO.
4. Washburn, E.W. and Urey, H.C. (1932). *Proc. Natl. Acad. Sci.* **18**: 496.
5. Lagowski, J.J. (1973). *Modern Inorganic Chemistry*. Marcel Dekker, New York.

6. Sienko, M.J. and Plane, R.A. (1979). *Chemistry: Principles and Applications*. McGraw-Hill, New York.
7. Attendorn, H.G. and Bowen, R.N.C. (1997). *Radioactive and Stable Isotope Geology*. Chapman & Hall, London.
8. Bowen, R. (1988). *Isotopes in the Earth Sciences*. Elsevier, London.
9. Kushner, D.J., Baker, A., and Dunstall, T.G. (1996). Biotechnological Potential of Heavy Water and Deuterated Compounds. *Proc. Papers 1996 Risk Assessment Res. Symp.*, Chapter 2.
10. Katz, J.J. (1960). The biology of heavy water. *Sci. Am.* **July**: 106–115.
11. Dahl, P.F. (1999). *Heavy Water and the Wartime Race for Nuclear Energy*. Institute of Physics, London.
12. Yang, F. and Hamilton, G.H. (1996). *Modern Atomic and Nuclear Physics*, International Editions. McGraw-Hill, New York.
13. Goldstein, A.M. (1986). *A Dictionary of Physics*. CBS Publishers and distributors, Delhi, India.
14. Urey, H.C., Brickwedde, F.G., and Murphy, G.M. (1932). An isotope of hydrogen of mass 2 and its concentration (abstract). *Phys. Rev.* **39**: 864.
15. Gilreath, E.S. (1985). *Fundamental Concepts of Inorganic Chemistry*. 19th printing, International Student Edition. McGraw-Hill, Singapore.

HENRY'S LAW

Roger C. Viadero Jr.
West Virginia University
Morgantown, West Virginia

OVERVIEW AND THEORY

William Henry (1774–1836) established the fundamental relationship to define the partitioning of a soluble (or partially soluble) chemical species between gas and liquid phases, at equilibrium, in a dilute aqueous solution. For the transfer of a chemical species, C, from the gas phase to the aqueous phase, presented schematically in Fig. 1, the following chemical reaction and dimensionless equilibrium constant, k, can be written for a dilute solution:

$$C_g \xleftrightarrow{k} C_{aq} \tag{1}$$

$$k = \frac{[C_{aq}]}{[C_g]} \tag{2}$$

where $[C_g]$ = gas phase concentration (mol/L) of C and $[C_{aq}]$ = aqueous phase concentration (mol/L) of C.

Upon recognizing that the molar concentration of a gas is related to the partial pressure of that gas via the ideal gas law presented in Eq. 3, the expression for the dimensionless equilibrium constant can be rewritten as in Eq. 4:

$$C_g = \frac{n}{V} = \frac{p_g}{RT} \tag{3}$$

$$k = \frac{[C_{aq}]}{[C_g]} = \frac{[C_{aq}],}{\dfrac{p_g}{RT}} \tag{4}$$

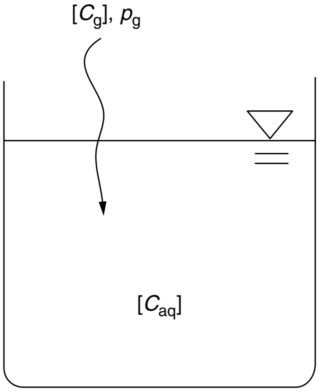

Figure 1. Transfer of chemical species C from the gas phase to the aqueous phase.

where n = number of moles of C, V = volume occupied by the gas (L), p_g = partial pressure of the gas (atm), R = ideal gas constant (0.082057 atm/L/K/mol) and T = temperature of the gas (K). By rearranging the relationship presented in Eq. 4, a familiar form of Henry's law is realized:

$$C_{aq} = k_H p_g \tag{5}$$

where k_H = Henry's law constant (mol/L/atm). The dimensionless equilibrium constants, k, and k_H, are related according to Eq. 6:

$$k_H = \frac{k}{RT} \tag{6}$$

Values of k_H for representative inorganic chemical species are presented in Table 1 (1,2). There are such a large number of volatile or semivolatile organic chemicals of environmental importance; readers are referred to the following representative supporting literature for detailed enumerations of Henry's law constants for organic species (3–10).

When using tabulated values of Henry's law constants or dimensionless equilibrium constants, care must be taken to examine the direction of transport being considered because the direction of reaction affects the units of k_H. For example, if the development of k_H presented above were reformulated for volatilization of C (from the aqueous to the gas phase), values of k_H would have units of atm L/mol and would thus be the inverse of the case developed above. The inconsistency encountered in the units of k_H is due partially to the

Table 1. k_H Values for Representative Inorganic Chemical Species[a]

Reaction	k_H (mol/L/atm)[b]
$CO(g) \leftrightarrow CO(aq)$	9.55×10^{-4}
$SO_2(g) \leftrightarrow SO_2\ (aq)$	1.23
$NH_3(g) \leftrightarrow NH_3\ (aq)$	57.5
$NO_2(g) \leftrightarrow NO_2\ (aq)$	1.00×10^{-2}
$O_3(g) \leftrightarrow O_3(aq)$	9.33×10^{-3}

[a] References 1 and 2.
[b] At 25 °C and 1 atm.

difference in conventions used by engineers and chemists and general lack of consensus with regard to a standard for the direction of transport (1,11).

LIMITATIONS OF HENRY's LAW

Three main limitations must be considered when applying Henry's law: (1) the state of equilibrium, (2) the solute concentration, and (3) the veracity of using standard data for calculations in contaminated water systems. The application of Henry's law is limited to cases where chemical equilibrium has been reached in a dilute solution (<1% solute concentration). However, the concentrations of most water constituents are normally less than 1%, so the application of Henry's law is typically valid. Tabulated values of k_H are generally based on experiments using pure water. As a result, the application of such data to contaminated water systems can yield unpredictable results (1). Under such circumstances, experiments may be needed to determine the appropriate value of k_H for the contaminated solution. Similar care should be taken when working with solvent systems other than water (e.g., when determining the solubility of a gas in an organic solvent).

APPLICATION OF HENRY'S LAW

By evaluating the relative magnitude of k_H for particular species, it is possible to learn about the solubility and general tendency of those species to partition between water and air. For instance, in the transport of a soluble chemical species from the gas phase to the aqueous phase, a compound that has a high k_H would be more soluble than one that has a lower k_H and would partition into the water phase where it may be transported downgradient or adsorb onto solid surfaces. In contrast, a species that has a low k_H would be less water soluble and tend to remain in the gas phase.

The surface of the earth is dominated by an air/water interface, so the interaction between gas and aqueous phase chemicals is extremely important in understanding the origins, fate, and transport of water and dissolved contaminants through the hydrogeochemical cycle.

Consider a sample of pure water which is completely isolated from the atmosphere. At standard conditions, the pH of the water would be neutral (pH = 7). However, when the pure water is allowed to reach equilibrium with an atmosphere containing $CO_2(g)$, the following series of reactions occurs:

$$CO_2(g) \xleftrightarrow{k_H} CO_2(aq) \tag{7a}$$

$$CO_2(g) + H_2O \xleftrightarrow{k_1'} \Leftrightarrow H_2CO_3{}^* \tag{7b}$$

$$H_2CO_3{}^* \xleftrightarrow{k_1} H^+ + HCO_3{}^- \tag{7c}$$

$$HCO_3{}^- \xleftrightarrow{k_2} H^+ + CO_3{}^{2-} \tag{7d}$$

At standard conditions, the partial pressure of $CO_2(g)$ is 3.162×10^{-4} atm (12). $CO_2(aq)$ and carbonic acid, H_2CO_3, act as weak acids and the equilibrium pH of the water is ~5.6. In this case, the atmosphere serves as a constant source of CO_2, and the total concentration of carbonate species increases with pH as presented in Fig. 2, where

$$C_{T_{CO_3}} = [H_2CO_3{}^*] + [HCO_3^-] + [CO_3^{2-}] \tag{8}$$

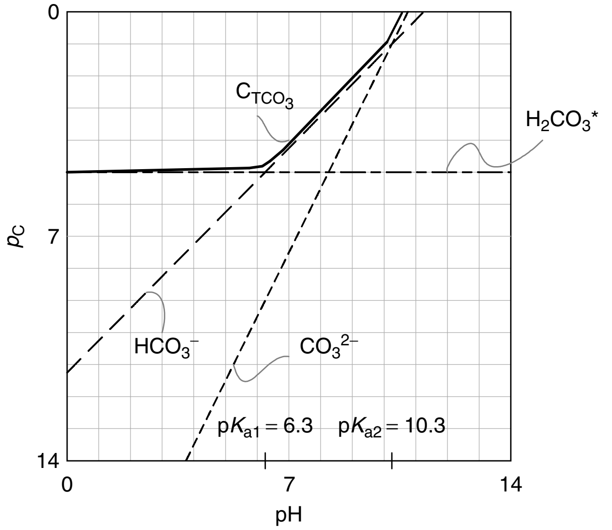

Figure 2. Plot of p_C versus pH for an aqueous system in equilibrium with atmospheric $CO_2(g)$ at standard conditions.

Further, any change in the partial pressure of CO_2 would result in a corresponding shift in the equilibrium pH. Failure to account for the presence and effects of $CO_2(g)$ on the pH would result in an entirely incorrect interpretation of acid–base chemistry which would then adversely impact conclusions drawn relative to environmentally important chemical reactions (e.g., precipitation/dissolution reactions, complexation reactions, and oxidation/reduction reactions).

BIBLIOGRAPHY

1. Stumm, W. and Morgan J. (1996). *Aquatic Chemistry*, 3rd Edn. John Wiley & Sons, New York.
2. Mackay, D., Shiu, W., and Ma, K. (1995). *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*, Vol. IV: *Oxygen, Nitrogen, and Sulfur Containing Compounds*. Lewis, Boca Raton, FL.
3. Burkhard, L., Armstrong, D., and Andren, A. (1985). Henry's law constants for the polychlorinated biphenyls. *Environ. Sci. Technol.* **19**: 590–596.
4. Mackay, D. and Shiu, W. (1981). A critical review of Henry's law constants for chemicals of environmental interest. *J. Phys. Chem. Ref. Data* **10**: 1175–1199.
5. Mackay, D., Shiu, W., and Ma, K. (1992). *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*, Vol. I: *Monoaromatic Hydrocarbons, Chlorobenzenes, and PCBs*. Lewis, Boca Raton, FL.
6. Mackay, D., Shiu, W., and Ma, K. (1992). *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*, Vol. II: *Polynuclear Aromatic Hydrocarbons, Polychlorinated Dioxins, and Dibenzofurans*. Lewis, Boca Raton, FL.
7. Mackay, D., Shiu, W., and Ma, K. (1993). *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate*

for Organic Chemicals, Vol. III: *Volatile Organic Chemicals*. Lewis, Boca Raton, FL.

8. Murphy, T., Mullin, M., and Meyer, J. (1987). Equilibration of polychlorinated biphenyls and toxaphene with air and water. *Environ. Sci. Technol.* **21**: 155–162.
9. Rice, C., Chernyak, S., and McConnell, L. (1997). Henry's law constants for pesticides measured as a function of temperature and salinity. *J. Agric. Food Chem.* **45**: 2291–2298.
10. U.S. EPA. (1998). *How to determine Henry's law constants, F_m Values, F_r Values, and F_e values for organic compounds, Appendix J to part 60*. U.S. Environmental Protection Agency, Washington, DC.
11. Benjamin, M. (2002). *Water Chemistry*. McGraw-Hill. New York.
12. Snoeyink, V. and Jenkins, D. (1980). *Water Chemistry*. John Wiley & Sons. New York.

HOFMEISTER EFFECTS

M. Boström
Linköping University
Linköping, Sweden

W. Kunz
University of Regensburg
Regensburg, Germany

B.W. Ninham
Australian National University
Canberra, Australia
and
University of Regensburg
Regensburg, Germany

During the 1880s and 1890s, Franz Hofmeister (Professor of Pharmacology at the Charles University in Prague) and some of his coworkers demonstrated that salts with common cation (positive ion) but differing in anion (negative ion) have different effectiveness in stabilizing protein suspensions (1,2). They can be arranged in a sequence that seemed later to be universal. Reference (2) includes an English translation of two of Hofmeister's papers. As can be seen in Fig. 1, depending on the choice of salt, different concentrations are required to precipitate a certain concentration of proteins from whole egg white (3). Hofmeister effects, or Hofmeister sequences, refer to the relative effectiveness of anions or cations on a wide range of phenomena. They have remained unexplained by theories of physical chemistry until recently. But, as we will discuss later, some encouraging steps have recently been taken toward resolving the problem. Understanding Hofmeister effects is essential in understanding how different salt ions in water influence solution chemistry and are critical to biological processes that are ubiquitous.

Hofmeister effects have been observed in an enormous number of systems: solubility of salts; solubility of oxygen; surface tension of salt solutions (4); pH measurements (5); buffers; ion binding to micelles, proteins (6), and membranes; conformational changes of rhodopsin (membrane bound protein acting as light receptor) (7); transport across human red blood cells (8); molecular forces (9,10); and colloid stability. Another example is the cutting efficiency of DNA in buffered salt solutions that not only follows a Hofmeister series: A change in supposedly irrelevant choice of buffer has the ability to reverse the Hofmeister series (11). A corresponding change in buffer has recently been demonstrated to reverse, in the same way, the Hofmeister series in pH measurements. Even bacterial cell growth follows the Hofmeister series (12).

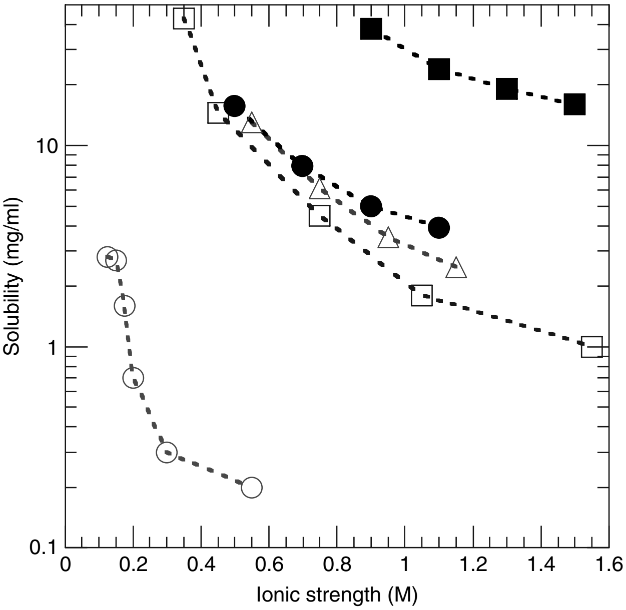

Figure 1. Variation, as a function of salt type, of hen-egg-white protein (lysozyme) solubility. Depending on the choice of salt, different concentrations are required to precipitate a certain concentration of proteins from whole egg white. Data taken from Ref. 3 are shown for NH_4OAC (filled squares), NH_4Cl (filled circles), KCl (open triangle), NaCl (open square), and KSCN (open circle).

Understanding Hofmeister effects is especially important for biological and medical applications. Although many biochemical functions require specific ions at certain concentrations, too high concentrations of, for example, NaCl can have deleterious effects on enzymatic activities (13,14). One important biological example is the human red blood cell that is sensitive to salt. During gas exchange, involved in buffering of blood and in O_2 and CO_2 transport in blood, the chloride ion moves between the plasma and the red blood cell intracellular fluid (15). It is well known experimentally that transport of Cl^-, Na^+, K^+, and water molecules through the cell membrane, and the corresponding cell volume changes, depend strongly on the choice of background anions (8). The rate constant for chloride exchange is, e.g., lowered to 6.5% of its original value when 120 mM of the chloride ions are replaced with thiocyanate ions. The interactions of ions with cell membranes and with membrane bound proteins (hemoglobin) determine key regulatory mechanisms. Taboy et al. (16), investigating hemoglobin, found a larger redox potential in the presence of NO_3^- than in a Cl^- salt solution (with a maximum redox potential around a salt concentration of 0.2 M). This finding is interesting because it demonstrates that different anions have a different affinity for hemoglobin and may provide useful insights as to why the affinity of chloride ions is larger for deoxyhemoglobin than for oxyhemoglobin (17). This finding has important biological consequences, but the reasons have not been clear. Even though it is a different story, it is worth noticing that the difference between two typical Hofmeister anions, thiocyanate and chloride, has also been exploited in detoxification of anticancer agents (18).

These systematic Hofmeister series demonstrate that common underlying interactions are involved, which is encouraging and strongly suggests that general and truly predictive theories may not be far away. Clearly experiments and theory that explain these effects are valuable and can provide insights into the underlying mechanisms that control vital biological systems. Having explained why it is important to understand Hofmeister effects, we now proceed to discuss some recent theoretical attempts to understand it.

The traditional hand waving "explanation" for the Hofmeister effects has been that the ions exert their effects indirectly by changing the water structure. People have used words such as: "hard and soft ions," "hydrophobic and hydrophilic ions," "lyotropic series," "chaotropic and cosmotropic ions," "salting-in," and "salting-out." When words like these and simple theories have proven insufficient to accommodate experimental results, a large number of additional effects and interactions have been invoked (accompanied by numerous adjustable fitting parameters). Although "understanding" of experiments has occasionally been claimed, there has been no predictability.

Until recently, the theories used were mainly based on electrostatic interactions between charged particles. These theories failed to explain why a change from one monovalent salt to another (i.e., salts with the same charges) can have huge effects. The main attempts to explain Hofmeister phenomena involved extension of theory to allow different "hydrated ion sizes," parameters that vary from one situation to another. However, such attempts do not work and Hofmeister series do not correlate directly with any additional physical or chemical parameter. The problem is that there must be a coupling, or cooperative effects, which are not so easy to characterize. An understanding has only recently emerged that the polarizability of ions, neglected in previous theories, plays an essential role. Polarizability is a quantity that measures the response of a particle to a perturbing electromagnetic field. Although two charged particles can interact via electrostatic forces, a quantum mechanical attraction also exists, the van der Waals force, between two charged neutral particles that is related to the polarizabilities of the two particles. This van der Waals force is also present for the interaction between charged particles but has usually been assumed to be negligibly small compared with the electrostatic force. Although electrostatic forces often dominate at low salt concentrations, van der Waal forces become important at high salt concentrations where the electrostatic forces are screened. Ninham et al. (19) demonstrated that van der Waals forces between ion and interface often dominate over electrostatic forces above and around the salt concentration we have in our blood (i.e., 0.15 molar salt), which is exactly the concentration regime where Hofmeister effects are observed. As can be seen in Fig. 2, ion-specific bubble fusion is, for instance, inhibited, following an abrupt transition with increased salt concentration, in sodium chloride solutions above the salt concentration we have in our blood (20). This is why ocean water is foamy, but fresh water is not. Today the salt concentration of the ocean is around 0.6 M, but the concentration in the Permian Sea was close to what we have in our blood, which means that even a small decrease in salt concentration would have had dramatic effects on bubble fusion. To have a sufficiently high salt concentration in the Permian Sea must have been a critical condition when early life emerged. A related topic is that the local salt concentration, e.g., in the Baltic Sea, may have dropped suddenly below the critical 0.15 M concentration after ice melted at the end of ice ages. A speculation on how this may have caused mass extinction of plankton can be found in the references (1).

At present, two parallel and complementary lines of work demonstrate that both surface effects and bulk effects are important for the understanding of Hofmeister effects. We will first briefly discuss surface related effects. The theoretical papers by Boström et al. have focused on surface phenomena, i.e., the van der Waals interaction between ions and interface. Boström et al. (21) demonstrated, e.g., that this is one important reason why the charge of lysozyme (hen-egg-white protein) at pH = 4.5 is higher in 0.1 M NaSCN than in 0.1 M NaCl. The negatively charged thiocyanate anions (SCN^-) are more attracted toward the protein, by larger van der Waals forces, than are the chloride ions. A higher local

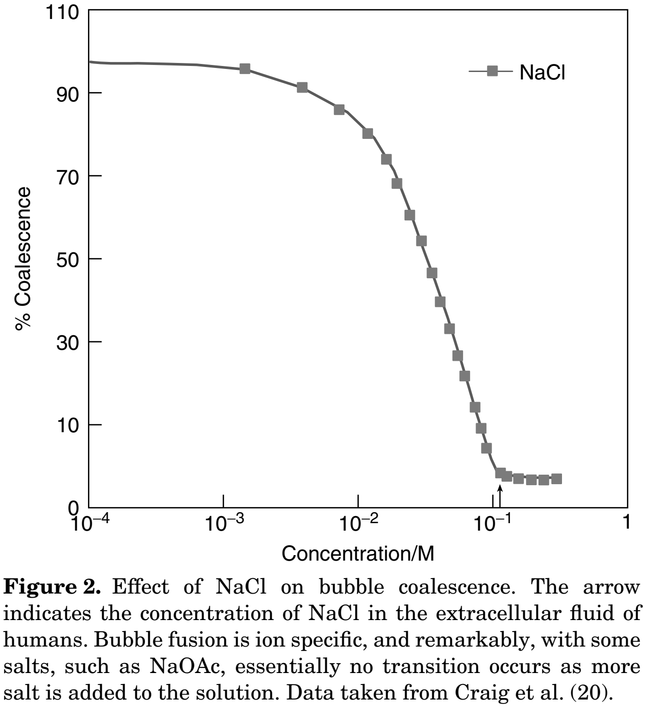

Figure 2. Effect of NaCl on bubble coalescence. The arrow indicates the concentration of NaCl in the extracellular fluid of humans. Bubble fusion is ion specific, and remarkably, with some salts, such as NaOAc, essentially no transition occurs as more salt is added to the solution. Data taken from Craig et al. (20).

anion concentration with SCN^- ions near the protein surface gives rise to a larger local concentration of hydronium ions (H_3O^+) at the protein surface, which in turn leads to more hydrogen ions being bound to the protein charge groups (i.e., a larger protein charge) in a thiocyanate (or iodide) salt compared with a chloride salt. Recent extension by Boström et al. of this theoretical work (investigating theoretically solutions with different buffers and proteins) has demonstrated that this also influences the ion-specific buffer capacity of a lysozyme solution at typical protein concentrations. Remarkably, in as far as both pH in solution and charge of hen-egg-white protein is concerned, one can often to a good approximation replace a certain concentration of NaCl with a smaller concentration of NaSCN and find the same final result. A similar result has been observed in experiments on enzymatic activity by P. Baudin et al. and in pH measurements in protein (cytochrome c) solutions by P. Baglioni et al. This kind of surface related effects has also offered a partial explanation (22) for the experimental observation by Bonneté et al. (6) that the interaction between two lysozyme proteins in salt solutions follows a Hofmeister series (the interaction is, e.g., more attractive in $NaNO_3$ solutions than in NaCl solutions).

The other line of work has focused on interactions in the bulk solution between ions and between ion and water molecules. It is important to stress that van der Waals forces are only one consequence of the ion polarizability. Other consequences are dipole-induced dipole and induced dipole-induced dipole forces between ions and between ions and water molecules. These forces have been in focus in molecular simulations. It has been repeatedly demonstrated that the static polarizability of ions must be included in computer molecular simulations to get reasonable results for the solvation energy of ions in water (23,24). This, however, is not the entire story. To get correct results for the activity coefficients (25) and solvation energies (26) of ions in water, strong evidence exists that one has to include the entire dynamic ionic polarizability. In several important contributions (27–29), attempts have been made to include forces that originate from the static polarizability correctly in computer simulations that include water molecules. These ion–ion and ion–water interactions are also important for the air–water interface. The air–water interface, and the surface tension of salt solutions, is more complicated than one might initially have guessed and not yet fully understood. Promising simulations (28) seem to give the right trend in Hofmeister sequence for the surface tension change with added salt. But these simulations give results that differ greatly from experimental surface tension changes with added salt. This is also true for pure water where simulations without fitting parameters give incorrect results for the surface tension and freezing temperature. One reason for these problems must be that simulations so far have not included many body dispersion forces properly. But simulations together with experiments are still useful tools for further understanding. As demonstrated by Jungwirth and Tobias (28), some highly polarizable anions, such as bromide, may, contrary to conventional wisdom (which is based on electrostatic theories), go to the air–water interface. It has also been demonstrated that one important effect is that the water molecules near an air–water interface are not isotropically ordered (28), which effectively gives an anisotropic film near the interface that strongly influences the interaction between ions and the air–water interface. These interactions (ion–ion, ion–water molecule, ion interface) influence each other in a self-consistent manner, which makes this a complicated but solvable problem. It is clearly important to take into account all of the complementary aspects discussed here for further progress. Our most recent results suggest a key role for interfacial dissolved gases, and the corresponding change in ion solvation energy, behind the observed surface tension changes (see article on DISSOLVED GASES). Theories, accompanied by refined simulations and experiments, are on the way that may substantially improve the understanding of the Hofmeister effect in the near future. It is also worth remarking that until now no theoretical work has included the effect of dissolved gasses properly. These gases are known to influence colloid interactions (30), bubble–bubble interaction (20), conductance, pH, and most likely the surface tension of salt solutions. It is not only CO_2 that is important but also other dissolved gasses. The interested reader can find more details about Hofmeister effects in different reviews (1,31–34).

BIBLIOGRAPHY

1. Kunz, W., Lo Nostro, P., and Ninham, B.W. (2004). *Curr. Op. Colloid Int. Sci.* **9**: 1.
2. Kunz, W., Henle, J., and Ninham, B.W. (2004). *Curr. Op. Colloid Int. Sci.* **9**: 19.
3. Ries-Kautt, M.M. and Ducruix, A.F. (1989). *J. Biol. Chem.* **264**: 745.
4. Boström, M., Williams, D., and Ninham, B.W. (2001). *Langmuir* **17**: 4475.
5. Boström, M. et al. (2003). *J. Phys. Chem. B* **207**: 2875.
6. Bonneté, F., Finet, S., and Tardieu, A. (1999). *J. Crystal Growth* **196**: 403.
7. Vogel, R. (2004). *Curr. Op. Colloid Int. Sci.* **9**: 133.
8. Payne, J.A., Lytle, C., and McManus, T. (1990). *J. Am. J. Physiol.* **259**: C819.
9. Pashley, R. et al. (1986). *J. Phys. Chem.* **90**: 1637.
10. Boström, M., Williams, D., and Ninham, B.W. (2001). *Phys. Rev. Lett.* **87**: 168103.
11. Kim, H-K., Tuite, E., Norden, B., and Ninham, B.W. (2001). *Eur. Phys. J. E* **4**: 411.
12. Lo Nostro, P. et al. (2004). *Curr. Op. Colloid Int. Sci.* **9**: 97.
13. Yancey, P. et al. (1982). *Science* **217**: 1214.
14. Baudin, P. et al. (2004). *Curr. Op. Colloid Int. Sci.* **9**: 43.
15. Prange, H. et al. (2001). *J. Appl. Pysiol.* **91**: 33.
16. Taboy, C. et al. (2000). *J. Biol. Chem.* **275**: 39048.
17. Rollema, H., de Bruin, S., Janssen, L., and van Os, G. (1975). *J. Biol. Chem.* **250**: 1333.
18. Hovinen, J., Petterson-Fernholm, T., Lahti, M., and Vilpo. (1998). *J. Chem. Res. Toxicol.* **11**: 1377.
19. Ninham, B.W. and Yaminsky, V. (1997). *Langmuir* **13**: 2097.
20. Craig, V., Ninham, B.W., and Pashley, R. (1993). *Nature* **364**: 317.

21. Boström, M., Williams, D., and Ninham, B.W. (2003). *Biophys. J.* **85**: 686.
22. Tavares, F.W., Bratko, D., Blanch, H., and Prausnitz, J.M. (2004). *J. Phys. Chem. B* **108**: 9228.
23. Stuart, S.J. and Berne, B.J. (1996). *J. Phys. Chem.* **100**: 11934.
24. Grossfield, A., Ren, P., and Ponder, J.W. (2003). *J. Am. Chem. Soc.* **125**: 15671.
25. Kunz, W., Belloni, L., Bernard, O., and Ninham, B.W. (2004). *J. Phys. Chem. B* **108**: 2398.
26. Boström, M. and Ninham, B.W. (2004). *J. Phys. Chem. B* **108**: 12593.
27. Ahn-Ercan, G., Krienke, H., and Kunz, W. (2004). *Curr. Op. Colloid Int. Sci.* **9**: 92.
28. Jungwirth, P. and Tobias, D. (2002). *J. Phys. Chem. B.* **106**: 6361.
29. Netz, R.R. (2004). *Curr. Op. Colloid Int. Sci.* **9**: 192.
30. Alfridsson, M., Ninham, B.W., and Wall, S. (2000). *Langmuir* **16**: 10087.
31. Cacace, M.G., Landau, E.M., and Ramsden, J.J. (1997). *Q. Rev. Biophys.* **30**: 241.
32. Collins, K.D. and Washabaugh, M.W. (1985). *Q. Rev. Biophys.* **18**: 323.
33. Leontidis, E. (2002). *Curr. Op. Colloid Int. Sci.* **7**: 81.
34. Abernethy, J.L. (1967). *Proc. California Association of Chem. Teachers* **44**: 177.

CLATHRATE HYDRATES

Carolyn Ann Koh
E. Dendy Sloan
Colorado School of Mines
Golden, Colorado

DEFINITION OF CLATHRATE HYDRATES

Clathrate hydrates are crystalline inclusion compounds consisting of a network of polyhedral water cages (host), which enclathrate small hydrophobic molecules (guests; typically <0.9 nm). Clathrate hydrates can be described as solid solutions (1), such that the empty hydrate lattice is unstable with respect to ice. When guest molecules occupy the water cages, the hydrate lattice is stabilized by van der Waals interaction forces. Typically, a water cage has a maximum occupancy of one guest molecule.

Simple clathrate hydrates contain only one type of guest molecule within the hydrate lattice, while *binary clathrate hydrates* contain two different types of guest molecules. *Mixed clathrate hydrates* refer to hydrates containing two or more different guest species. When the guest species are gas molecules, such as methane, ethane, or carbon dioxide, clathrate hydrates are also called *gas hydrates*. It is not necessary for all the cages in these clathrate hydrate structures to be occupied; for example, methane hydrate can be prepared with 90% of the cages occupied by methane (2). The cage filling depends on pressure, temperature, and the nature of the guest species. Hence, these structures are often referred to as "nonstoichiometric compounds."

COMMON CLATHRATE HYDRATE STRUCTURES

The three most common types of clathrate hydrate structures are sI, sII, and sH hydrate. sI hydrate predominates in natural environments, while sII hydrate is found mainly in artificial environments (3). sH hydrate may occur in either environment but is less common than either sI or sII hydrate. These three structures have distinct crystallographic properties (see Table 1 and References 2 and 4 for further details). Both sI and sII hydrates have cubic crystal structures (Fig. 1), while sH hydrate has a hexagonal crystal structure. sI and sII hydrates contain two different types of water cages (Fig. 2). The water cages are described by the general notation X^n, where X is the number of sides of a cage face and n is the number of cage faces having these X sides. For example, the pentagonal dodecahedral 5^{12} water cage is composed of 12 *five*-membered water rings. A unit cell of sI hydrate contains two small (5^{12}) + six large tetracaidecahedral ($5^{12}6^2$) cages. A unit cell of sII hydrate contains six small (5^{12}) + eight large hexacaidecahedral ($5^{12}6^4$) cages (sII). A unit cell of sH hydrate is comprised of three types of water cages: three small (5^{12}) cages, two irregular dodecahedral ($4^35^66^3$) cages, and a single large icosahedral ($5^{12}6^8$) cage (Fig. 2).

The type of clathrate hydrate structure that forms depends on the size of the guest molecule; for example, CH_4 and CO_2 both form sI hydrate; C_3H_8 and C_4H_8O (tetrahydrofuran, THF) form sII hydrate; while a larger guest molecule such as methylcyclohexane in the presence of methane forms sH hydrate (2,4). Table 2 illustrates the range of different guest molecules that form sI, sII,

Table 1. Crystallographic Properties of the Common Clathrate Hydrate Structures

Property	sI	sII	sH
Space group	*Pm3n*	*Fd3m*	*P6/mmm*
Unit cell parameters (nm)	$a = 1.20$	$a = 1.70$	$a = 1.21, c = 1.01$
Average cavity radius (nm) [number of cavities per unit cell (cavity type)]	$0.395[2(5^{12})]$(S) $0.433[6(5^{12}6^2)]$(L)	$0.391[16(5^{12})]$(S) $0.473[8(5^{12}6^4)]$(L)	$0.391[3(5^{12})]$(S) $0.406[2(4^35^66^3)]$(S) $0.571[1(5^{12}6^8)]$(L)
Symmetry of cages	M$\underline{3}$(S), $\underline{4}$2m(L)	3m (S), $\underline{4}$3m (L)	Mmm(S), $\underline{6}$m2(S), 6/mmm(L)
Ratio of numbers of small/large cavities	0.33	2	5
Number of water molecules per unit cell	46	136	34
General unit cell formula	$X \cdot 5.75H_2O$	$X \cdot 17H_2O$	$5(X,Y) \cdot 34H_2O$

L, large cage; S, small cage; X,Y, guest molecules.

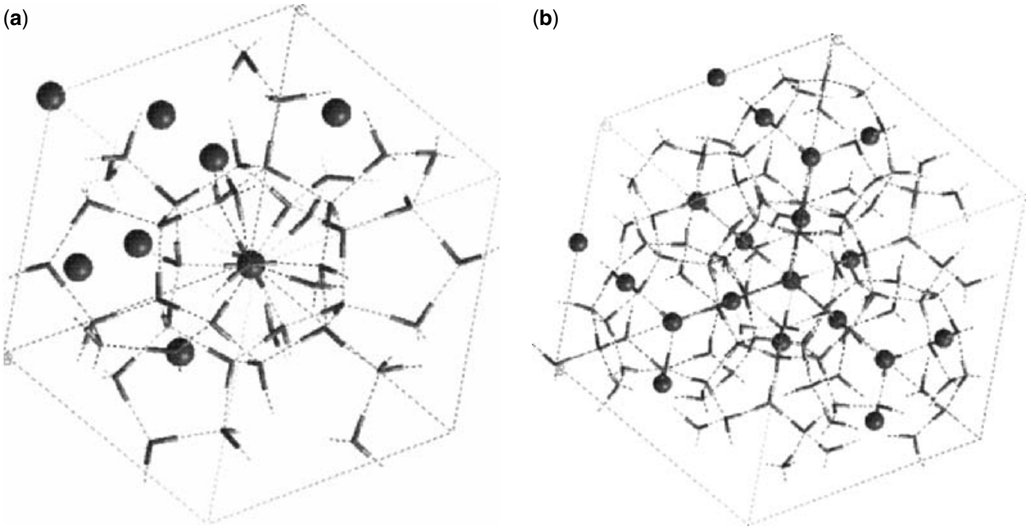

Figure 1. The partial unit cell structures of (**a**) sI hydrate and (**b**) sII hydrate. Dotted lines represent hydrogen bonds between water molecules. Spheres represent guest molecules.

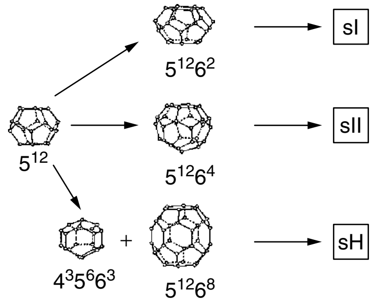

Figure 2. Water cages of sI, sII, and sH hydrates.

and sH hydrates. Typically, high pressure (3–11 MPa) and low temperature (275–285°K) conditions are required to stabilize gas hydrate systems. Table 2 gives the dissociation pressures of different clathrate hydrates and the occupancy of the corresponding hydrate guest molecules at around 277°K (the temperature at which most natural and industrial hydrates are found). Clathrate hydrates will be stable at 277°K and at a pressure above the dissociation pressure. Although CH_4 and C_2H_6 are both sI hydrate formers, a binary CH_4/C_2H_6 mixture can exhibit sI/sII transitions with varying pressure and/or composition. In contrast, a binary CH_4/CO_2 mixture (where both pure components are sI hydrate formers) forms only sI hydrate (Table 2).

NOVEL CLATHRATE HYDRATE STRUCTURES

Between the time that clathrate hydrates were first discovered in 1810 (6) and upto 1996, only three main structures were identified. Yet, within the last eight years (1997–2004), there have been a number of exciting reports of new clathrate hydrate structures. In addition to the discovery of multiple occupancies of the hydrate water cages, at least three new clathrate hydrate structures have been discovered. Some of these new structures are discussed next.

As mentioned earlier, typically only one guest molecule is accommodated within each clathrate hydrate cage. However, recent measurements and computer calculations have demonstrated that clathrate hydrate structures can consist of water cages containing more than one guest molecule. Specifically, neutron diffraction measurements show that double occupancy of the large cage of sII hydrate can occur for both nitrogen hydrate and oxygen hydrate at higher pressures (7). Similar conclusions have been derived from molecular dynamics computer calculations (8). Multiple cage occupancy has also been proposed in sI, sII, and sH argon hydrates from first principle and lattice dynamics calculations (9).

Multiple occupancy of the water cages in H_2 hydrate was discovered using neutron and X-ray diffraction and Raman spectroscopy by Mao et al. (10,11). H_2 molecules were reported to exhibit double and quadruple occupancy in the small and large cages of sII hydrate, respectively (10,11). However, recently Lokshin and co-workers in collaboration with the Mao group (10) reported that D_2 molecules only singly occupy the small cages of D_2 hydrate (12). This discovery that H_2 can form a clathrate hydrate is especially noteworthy since it was previously assumed that small molecules such as H_2 and He are too small to stabilize a clathrate hydrate structure. The pure H_2 hydrate structure is typically formed at very high pressures (near 1 GPa). However, the addition of a promoter molecule, for example, THF, to form a binary H_2/THF hydrate will stabilize H_2 within the sII framework at pressures two orders of magnitude lower than that in pure hydrogen hydrates (13).

A complex clathrate hydrate structure, 1.67 choline hydroxide-tetra-*n*-propylammonium fluoride$\cdot 30.33H_2O$, was identified in 1999 using single-crystal X-ray diffraction (14). A space group of *R*-3 was assigned to this structure. This new clathrate hydrate consists of alternating stacks of sH and sII hydrates. Unlike

Table 2. Different Hydrate Guest Molecules and Their Clathrate Hydrate Structures, Dissociation Pressures, and Water Cage Occupancy

Guest Molecule	Structure Formed	Occupancy of Water Cages[a]	P,kPa; T, K[b]
CH_4	sI	5^{12} (86%) and $5^{12}6^2$ (95%)	4170, 277.6
C_2H_6	sI	5^{12} (0.3%) and $5^{12}6^2$ (97%)	814, 277.6
CO_2	sI	5^{12} (70%) and $5^{12}6^2$ (98%)	3847, 277.6
C_3H_8	sII	$5^{12}6^4$ (∼ 100%)	455, 277.8
C_2H_4O (EO)[c]	sI	$5^{12}6^2$ (∼ 100%)	101.3, 284.2
C_4H_8O (THF)[c]	sII	$5^{12}6^4$ (∼ 100%)	101.3, 277.5
CH_4+ C_2H_6	Binary sI/sII	sII−C_1: 5^{12} (79%), $5^{12}6^2$ (18%)	90.4% C_1: 2096, 277.6
		sII−C_2: $5^{12}6^2$ (80%)	
		sI−C_1: 5^{12} (63%), $5^{12}6^2$ (3%)	56.4% C_1: 1564,[d] 277.6
		sI−C_2: 5^{12} (0.2%), $5^{12}6^2$ (97%)	
CH_4+ CO_2	Binary sI	CH_4: 5^{12} (66%), $5^{12}6^2$ (41%)	66% C1: 2840, 277.0
		CO_2: 5^{12} (14%), $5^{12}6^2$ (55%)	
Methylcyclohexane + CH_4	sH	5^{12} (CH_4), $4^35^66^3$ (CH_4), $5^{12}6^8$ (mch)	2137, 277.6
t-Butylmethyl ether + CH_4	sH	5^{12} (CH_4), $4^35^66^3$ (CH_4), $5^{12}6^8$ (*t*-btme)	2550, 277.5

[a] A Gibbs energy minimization program, CSMGem, has been used to estimate the occupancy of each water cage (5).

[b] Values are from experimental data extracted from Reference 4.

[c] Dissociation temperatures are given for these compounds since they are stable at atmospheric pressure.

[d] Dissociation pressure calculated from CSMGem (5).

the common (sI, sII, and sH) hydrates, which contain hydrophobic guest molecules, the choline hydroxide guest in this complex structure exhibits both hydrophobic and hydrophilic modes of hydration. Choline hydroxide occupies the large cages ($5^{12}6^4$ and $5^{12}6^8$), displacing one of the host water molecules of the $5^{12}6^8$ cages. The tetra-*n*-propylammonium guest was reported to occupy fused supercages ($4^35^{30}6^3$ and 5^{36}). It was suggested that the closed-cage version of this structure could occur in natural and industrial environments with natural gas molecules as the guest species.

Contrary to previous reports of bromine forming a number of different hydrate structures, single-crystal X-ray diffraction measurements by Udachin et al. (15) showed that bromine hydrate only forms a tetragonal structure of space group $P4_2/mmm$ (cf. Ref. 16). The composition of bromine hydrate was found to be $Br_2 \cdot nH_2O$, where $n = 8.62–10.68$. The tetragonal framework is comprised of small 5^{12} cages and large $5^{12}6^2$ and $5^{12}6^3$ cages. The large cages are occupied by varying amounts of bromine molecules. The small cages were either vacant or partially occupied by O_2 or N_2 molecules, which were incorporated during crystallization of the bromine/water solution in air.

sT clathrate hydrate, a trigonal structure of space group *P*321, was discovered in 2001 (17). The structure was determined using single-crystal X-ray diffraction and NMR spectroscopy. All three types of large cages ($5^{12}6^3$, $5^{12}6^2$, and $4^15^{10}6^3$) in the structure are occupied by dimethyl ether guest molecules. The ratio of small to large cages is 0.25, which is significantly smaller than other known structures. As a consequence, this structure is most efficient at minimizing void space when the small cages are empty and has a higher density ($1.074\ g{\cdot}cm^{-3}$) than sII or sI hydrate. The authors suggest that this higher density may result in sT being favored over sI and sII for intermediate sized guests at high pressure. However, they also acknowledge that other factors need to be accounted for, such as the degree of ring strain of the four-membered water rings.

Methane hydrate has been shown from neutron and X-ray diffraction to undergo structural transformations to unconventional structures at elevated pressures (2,18,19). sI hydrate transforms to MH-II, a hexagonal structure, at around 1 GPa, which further transforms at around 2 GPa to an orthorhombic ice-filled structure MH-III. The ice-filled structure of methane hydrate was found to be stable up to 42 GPa (19).

Similarly, THF hydrate has been reported to undergo a structural transition from sII hydrate to an unusual structure at high pressure (0.8 GPa) (20). From neutron diffraction, the new high pressure phase of THF hydrate was assigned to an orthorhombic structure of space group *Pnma*. The water framework is reported to be formed exclusively from one type of 14-hedron cage, $4^45^46^6$. The unit cell stoichiometry of this new structure is 4(THF-d_8)·$24D_2O$.

CLATHRATE HYDRATES IN NATURAL AND INDUSTRIAL ENVIRONMENTS

Clathrate hydrates, containing predominantly biogenic methane, occur naturally in ocean sediments and permafrost regions (2–4). The amount of carbon in these natural gas hydrate deposits is estimated to be twice that of combined fossil fuel reserves [1 m^3 of methane hydrate in which 90% of the cages are occupied contains the equivalent of 156 m^3 of methane under standard conditions (STP)]. Significant efforts have been made over recent years to evaluate the feasibility of recovering these natural gas hydrate deposits. Oceanic gas hydrate deposits are currently considered to be too disperse and expensive to warrant near-term exploration (3). Conversely, permafrost deposits have significantly higher hydrate concentrations than hydrates found in ocean sediments. As a consequence, pilot drilling and characterization are being performed in Alaska and Canada in the permafrost, and in Japan in the ocean (3). However, extensive research and development is still required before these natural gas hydrate deposits can be produced safely and economically.

Natural gas hydrates also occur in gas and oil subsea pipelines, when the fluid in the pipeline reaches the temperature and pressure conditions that are favorable for hydrate formation. This can result in blockage of the pipelines and, as a consequence, catastrophic economic loss and ecological and safety risks (2,4). The typical method used to prevent hydrate formation within subsea pipelines is to add a thermodynamic inhibitor, which shifts the hydrate formation conditions to lower temperatures and/or higher pressures (2,4).

However, in many deepwater production scenarios, thermodynamic inhibition can become uneconomical and even prohibitive due to the high concentrations required under these conditions. Therefore, new technologies are currently being developed to control hydrate formation within these deepwater pipelines. These new methods include:

1. The addition of low dosage hydrate inhibitors (LDHIs) that are effective at concentrations below about 1 vol%. There are two broad classes of LDHIs: kinetic hydrate inhibitors (KHIs) and antiagglomerant inhibitors (AAs). KHIs operate by delaying nucleation and/or crystal growth. AAs prevent hydrate crystals from agglomerating and forming a blockage by maintaining the hydrates in the form of a suspended slurry, which allows fluid flow to occur unimpeded.
2. “Cold flow,” whereby hydrates could be pumped through the pipeline without the need of inhibitors (21). Sintef-BP researchers have reported that the addition of water to a flow of dry hydrate results in the formation of further dry hydrate. It is suggested that this implies that if there is a cold flow of oil/gas containing dry hydrates, warm well streams containing water can then be added, with the water from the well being converted to dry hydrates.

Numerous technological applications have been proposed for clathrate hydrates. Highlights of some of these applications are given as follows. The discovery of the ability to store H_2 in a pure H_2 hydrate at high pressures by Mao et al. (10,11) and more recently within a binary H_2/THF hydrate at lower pressures (13) presents an innovative technology for hydrogen storage. This technology seems particularly feasible since the successful proof-of-concept that hydrogen can be stored at mild conditions (of T and P) within a binary clathrate hydrate lattice using promoter molecules (13). The advantage of using clathrate hydrate materials for hydrogen storage is that the hydrogen is stored in molecular form, without the need of a chemical reaction for its release (13). This increases the overall energy efficiency of the system. Furthermore, all stored H_2 should be completely released by decomposition of the hydrate at near-ambient temperature and pressure conditions. The storage of methane within a sI hydrate framework is also currently being evaluated as an alternative method to fuel ships or to transport stranded gas (3).

Methane hydrate has also been considered a potential source for climate change (3). The clathrate gun hypothesis (also known as the “late Quaternary climate change”) suggests that methane released from methane hydrate around 15,000 years ago caused immense global warming (22). However, this hypothesis remains controversial.

The use of gas hydrates for cool storage has been explored since the early 1980s (23). The R141b gas hydrate (sII hydrate) was selected as a cool-storage system, since the R141b refrigerant offers significant advantages over Freon hydrate formers, including a lower saturation pressure and lower cost. R141b hydrate can be formed under mild conditions of 0.1 MPa and around 278°K. A further potential application of clathrate hydrates is CO_2 sequestration. This would involve injecting liquid CO_2 (droplets) into the ocean at depths of greater than 500 m where a solid interface of CO_2 hydrate will form between CO_2 and water (24,25). Injecting CO_2 into deep marine sediments would not be expected to result in long-term effects to biological systems and could circumvent the problem of eventually releasing the disposed CO_2 into the atmosphere.

Other proposed applications of gas hydrates include a range of separation processes, such as the treatment of aqueous and gaseous pollutant streams [e.g., recovery of chlorinated hydrocarbons from contaminated groundwater (26) or carbon dioxide from flue gases (27), respectively], isotopic fractionation of oxygen and hydrogen (28), and marine desalination (29).

CONCLUSION

The increasing utilization of molecular tools, such as X-ray and neutron diffraction, Raman and NMR spectroscopy, and computer simulation calculations, is producing a wealth of new and interesting findings on the structural and dynamic properties of clathrate hydrates. Further understanding of the hydrate structural transitions occurring over a range of temperature, pressure, and composition and the kinetics of these transitions will be important in advancing the range of new technologies being developed based on clathrate hydrates. This new molecular-level insight of hydrate structure and dynamics could have implications in the broader fields of inclusion compounds and crystal growth processes.

Acknowledgment
Thanks to K. Hester for his help with the CSMGem calculations.

BIBLIOGRAPHY

1. van der Waals, J.H. and Platteeau, J.C. (1959). Clathrate solutions. *Adv. Chem. Phys.* **2**: 1.
2. Koh, C.A. (2002). Towards a fundamental understanding of natural gas hydrates. *Chem. Soc. Rev.* **31**: 157–167.
3. Sloan, E.D. (2003). Fundamental principles and applications of natural gas hydrates. *Nature* **426**: 353–359.
4. Sloan, E.D. (1998). *Clathrate Hydrates of Natural Gases*, 2nd Edn. Marcel Dekker, New York.
5. CSMGem Program, Colorado School of Mines.
6. Davy, H. (1811). *Philos. Trans. R. Soc. London* **101**: 1–35.

7. Chazallon, B. and Kuhs, W.F. (2002). *In situ* structural properties of N_2-, O_2-, and air-clathrates by neutron diffraction. *J. Chem. Phys.* **117**: 308–320.
8. van Klaveren, E.P. et al. (2001). Molecular dynamics simulation study of the properties of doubly occupied N_2 clathrate hydrates. *J. Chem. Phys.* **115**: 10500–10508.
9. Inerbaev, T.M. et al. (2004). Theoretical study of clathrate hydrates with multiple occupation. *J. Incl. Phen. Macrocyclic Chem.* **48**: 55–60.
10. Mao, W.L. et al. (2002). Hydrogen clusters in clathrate hydrates. *Science* **297**: 2247–2249.
11. Mao, W.L. and Mao, H.K. (2004). Hydrogen storage in molecular compounds. *Proc. Natl. Acad. Sci.* **101**: 708.
12. Lokshin, K.A. et al. (2004). Structure and dynamics of hydrogen molecules in the novel clathrate hydrate by high pressure neutron diffraction. *Phys. Rev. Lett.* **93**: 125503–125504.
13. Floruuse, L.J. et al. (2004). Stable low pressure hydrogen clusters stored in a binary clathrate hydrate. *Science*.
14. Udachin, K.U. and Ripmeester, J.A. (1999). A complex clathrate hydrate structure showing bimodal guest hydration. *Nature* **397**: 420–423.
15. Udachin, K.A., Enright, G.D., Ratcliffe, C.I., and Ripmeester, J.A. (1997). Structure, stoichiometry, and morphology of bromine hydrate. *J. Am. Chem. Soc.* **119**: 11481–11486.
16. Allen, K.W. and Jeffrey, G.A. (1963). *J. Chem. Phys.* **38**, 2304.
17. Udachin, K.A., Ratcliffe, C.I., and Ripmeester, J.A. (2001). A dense and efficient clathrate hydrate structure and unusual cages. *Angew. Chem. Int. Ed.* **40**: 1303–1305.
18. Loveday, J.S. et al. (2001). Transition from cage clathrate to filled ice: the structure of methane hydrate III. *Phys. Rev. Lett.* **87**: 215501–215504.
19. Hirai, H. et al. (2003). Retention of filled ice structure of methane hydrate up to 42 GPa. *Phys. Rev. B* **68**: 172102–172104.
20. Kurnosov, A.V. et al. (2004). New clathrate hydrate structure: high-pressure tetrahydrofuran hydrate with one type of cavity. *Angew. Chem. Int. Ed.* **43**: 2922–2924.
21. Evensen, K. (2004). Cold flow techniques set to thaw out hope for tie backs. *Upstream Technol.* **76**.
22. Kennett, J.P., Cannariato, G., Hendy, I.L., and Behl, R.J. (2003). *Methane Hydrates in Quaternary Climate Change: The Clathrate Gun Hypothesis*. American Geophysical Union, Washington, DC.
23. Bi, Y. et al. (2004). Influence of volumetric-flow rate in the crystallizer on the gas hydrate cool-storage process in a new gas hydrate cool-storage system. *Appl. Energy* **78**: 111–121.
24. Rehder, G. et al. (2004). Dissolution rates of pure CH_4 hydrate and CO_2 hydrate in undersaturated seawater at 1000-m depth. *Geochim. Cosmochim. Acta* **68**: 285.
25. Radhakrishnan R. and Trout, B.L. (2002). A new approach for studying nucleation phenomena using molecular simulations: application to CO_2 hydrate clathrates. *J. Chem. Phys.* **117**: 1786.
26. Seo, Y. and Lee, H. (2001). A new hydrate-based recovery process for removing chlorinated hydrocarbons from aqueous solutions. *Environ. Sci. Technol.* **35**: 3386–3390.
27. Kang, S-P. and Lee, H. (2000). *Environ. Sci. Technol.* **34**: 4397–4400.
28. Maekawa, T. (2004). Experimental study on isotopic fractionation in water during gas hydrate formation. *Geochem. J.* **38**: 129–138.
29. Max, M.D. (2002). Desalination using positively buoyant or negatively buoyant/assisted buoyancy hydrate. US Patent 6,497,794.

HYDRATION

Carolyn Ann Koh
Colorado School of Mines
Golden, Colorado

DEFINITION OF HYDRATION

Hydration can be defined as the process by which water is added to produce a *hydrate*. The process is reversible, that is, water can be reextracted from the hydrate. A second meaning of *hydration* involves a *hydration reaction*. This is where water is permanently and chemically combined with a reactant in a way that it can no longer be reextracted. The second meaning of hydration will not be considered here.

Hydration to form a hydrate can occur when water (*solvent*) interacts with *solute* molecules to form a solution, which involves *hydrophobic hydration*. Hydrophobic hydration is the main focus here.

HYDRATION OF HYDROPHOBIC MOLECULES

Hydration of nonpolar molecules is usually associated with hydrophobic effects in which water cavities or hydration shells are formed. This is accompanied by a loss in entropy and aggregation of the nonpolar (hydrophobic) molecules. Hence, ordering of the water hydration shell around hydrophobic molecules has been attributed to "clathrate" behavior in which the water hydration shell is dominated by pentagons compared to bulk liquid water (1). However, thermodynamic studies (2) suggest that although there are larger numbers of pentagons in the solvation shell compared to the bulk, the hydration shell also contains significant numbers of hexagons and larger polygons. It was suggested that the existence of larger polygons in the hydration shell indicates that the clathrate analogy is too simple to explain water organization along hydrophobic surfaces (2).

The water hydration shell around nonpolar molecules, such as methane, has been studied before, during and after gas clathrate hydrate formation using neutron diffraction with H/D isotopic substitution coupled with empirical potential structure refinement (EPSR) computer simulations (3–5). Gas clathrate hydrates are crystalline inclusion compounds that are formed when water and nonpolar molecules, such as methane, come into contact at high pressure and low temperature (6). Figure 1 illustrates the expansion of the water hydration shell and increased ordering of the second hydration shell around methane after clathrate hydrate formation.

Hydrophobic behavior has been observed for a concentrated solution of 0.86 mole fraction of tertiary butanol in water (7). This behavior is generally seen in more dilute solutions. A solution of tertiary butanol in water presents

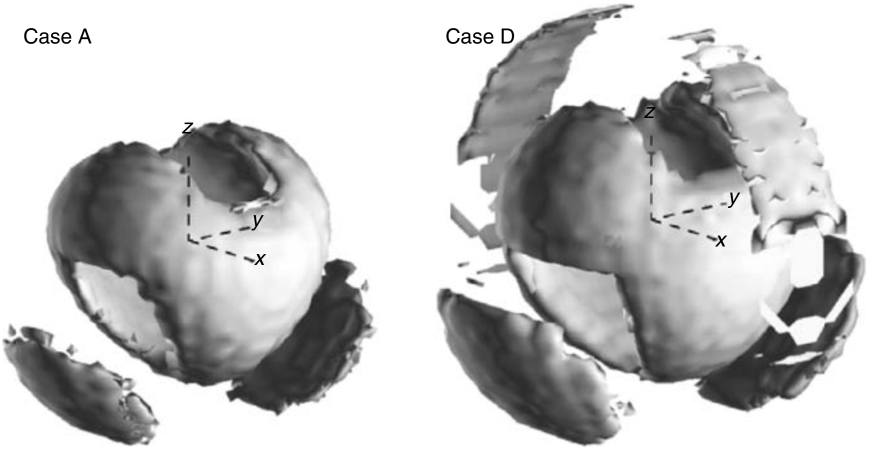

Figure 1. The hydration shell structure around methane: Case A, before hydrate formation (180 bar, 18 °C); case D, after hydrate formation (180 bar, 4 °C).

a classic example of an ampliphile–water solution. Neutron diffraction with H/D isotopic substitution coupled with EPSR computer simulations was used to examine the intermolecular correlations between alcohol and water molecules. The results revealed that "water pockets" are created within the concentrated alcohol–water solution, which reduce the degree of alcohol–alcohol direct hydrogen bonding. The pockets were found, on average, to contain between two and three water molecules. These water molecules exhibit a strong tendency to interact with the polar hydroxyl groups on the alcohol. As a result, the nonpolar groups became more tightly packed (i.e., exhibiting hydrophobic association) compared to that in the pure liquid alcohol. In the case of more dilute solutions of tertiary butanol in water (0.06 mole fraction of alcohol), the structural data do not support the conventional view that hydrophobic processes are dominated by perturbations of the first hydration shell. Conversely, significant changes are observed in the second hydration shell of water, with a "compression and tightening" of water molecules in the second shell (8).

A two-moment information theory was developed to describe the hydrophobic effect (9). This theory was shown to have a clear connection with the molecular principles of statistical thermodynamics. It was suggested that this theory could also be applied to describe hydrophobic effects on biopolymer structure in aqueous solution.

HYDRATION OF IONS

Concentrated solutions of sodium hydroxide in water have been studied using neutron diffraction with H/D isotopic substitution and EPSR computer simulations. The solute was found to affect the tetrahedral network of hydrogen-bonded water molecules (cf. when high pressure is applied to pure water). The results also indicated that there was a competition between the hydrogen-bonding interactions and Coulomb forces in determining the orientation of water molecules within the cation solvation shell (10).

The effect of adding sodium chloride as a salting-out agent to a dilute solution of tertiary butanol in water (0.02 mole fraction of alcohol) has also been studied using neutron diffraction (11). Contrary to previous understanding on salting-out (12), these measurements show that an anion bridge is formed between the polar ends of nearby alcohol molecules. Hence, there is significant enhancement of polar–polar interactions of the alcohol hydroxyl groups. As a result of the anion bridge formation, further hydrophobic interactions occur between the nonpolar/nonpolar groups.

CONCLUSION

Hydration of nonpolar molecules and salts and hydrophobic hydration have been studied in detail over the last couple of decades using computer simulation and theory. Recently, microscopic tools such as neutron diffraction with H/D isotopic substitution have been applied to study hydration of nonpolar molecules and salts. Neutron diffraction measurements coupled with molecular simulation provide a powerful tool to verify existing models of hydration of nonpolar molecules and salts, as well as combinations of these species.

Acknowledgment
The author acknowledges funding support from Chevron Texaco Energy Technology Company, EPSRC, the ISIS facility, and the CSM Hydrate Consortium. Thanks also to P. Buchanan, N. Aldiwan, and Y. Lui.

BIBLIOGRAPHY

1. Franks, F. (1973). *Water. A Comprehensive Treatise*. Vol. 2, Plenum, New York, pp. 1–48.
2. Head-Gordon, T. (1995). Is water structure around hydrophobic groups clathrate-like? *Proc. Natl. Acad. Sci.* **92**: 8308–8312.
3. Koh, C.A. et al. (2000). Water ordering around methane during hydrate formation. *J. Chem. Phys.* **113**: 6390–6397.
4. Buchanan, P. et al. (2003). *In situ* neutron diffraction studies of methane hydrate formation and decomposition. *J. Chem. Eng. Data* **48**: 778–782.
5. Koh, C.A. (2002). Towards a fundamental understanding of natural gas hydrates. *Chem. Soc. Rev.* **31**: 157–167.
6. Koh, C.A. and Sloan, E.D. (2005). Clathrate hydrates. In: *Encyclopedia of Water*. John Wiley& Sons, Hoboken, NJ.
7. Bowron, D.T. and Moreno, S.D. (2002). The structure of a concentrated aqueous solution of tertiary butanol: water

pockets and resulting perturbations. *J. Chem. Phys.* **117**: 3753–3761.

8. Bowron, D.T., Soper, A.K., and Finney, J.L. (2001). Temperature dependence of the structure of a 0.06 mole fraction tertiary butanol–water solution. *J. Chem. Phys.* **114**: 6203–6219.
9. Hummer, G. et al. (1996). An information theory model of hydrophobic interactions. *Proc. Natl. Acad. Sci.* **93**: 8951–8955.
10. Botti, A. et al. (2004). The microscopic structure of concentrated NaOH solutions. *J. Chem. Phys.* **120**: 10154–10162.
11. Bowron, D.T. and Finney, J.L. (2003). Structure of a salt–amphiphile–water solution and the mechanism of salting out. *J. Chem. Phys.* **118**: 8357–8372.
12. Sacco, A., De Cillis, F.M., and Holz, M.J. (1998). *Chem. Soc. Faraday Trans.* **94**: 2089.

THE MIRAGE OF THE H_2 ECONOMY

Reuel Shinnar
The City College of the CUNY
New York

In a recent paper (1), I analyzed the new "hydrogen economy" and explained why, for most uses, it makes no sense to convert fuels, electricity, or any other form of energy to hydrogen, only to use the hydrogen as a fuel for fuel cells. As H_2 is not a resource but an inefficient energy carrier, it would appear that the H_2 economy is just a mirage. If alternative energy sources are ever developed, an economy based on electricity would be cheaper (by a factor of three) and much easier to adopt than H_2. Oddly, I had covered the same topic twenty years ago in a paper that analyzed the hydrogen economy of the 1970s (2). Then, the investment in H_2 research was about 10 billion wasted dollars before it was phased out. In this column, I will try to explain why the scientific establishment engages in such obviously futile behavior and why this mirage attracts so many followers.

In the mid 1970s, I became aware of a strange phenomenon. Many leading scientific journals were suddenly giving enormous coverage to hydrogen, the fuel of the future; a new term, hydrogen economy, was born. International meetings attracted thousands of participants and the United States, Germany, and Japan funded the necessary development with vast resources. The central idea, put forward by the International Atomic Energy Research Center in Vienna, used a high-temperature nuclear reactor as the heat source. In this reactor, pressurized helium was used as the coolant. To generate electricity, the hot helium was first expanded through a turbine and then used for steam generation before being fed back to the reactor.

The "thermochemical cycle," the name applied to this new process, does not include a turbine, and the heat from cooling the hot helium is used to drive a set of closed-cycle chemical reactions; the net result is that water is split into hydrogen and oxygen. Following is a hypothetical example of such a reaction:

$$\mathrm{MeO + Heat \longrightarrow Me + 1/2O_2}$$
$$\mathrm{Me + H_2O \longrightarrow H_2 + MeO + Heat}$$

where Me is any metal. No such reaction was ever discovered, and all cycles proposed were quite complex. One does not need to be an experienced cost estimator to realize that such a cycle would be very expensive and have a low thermal efficiency.

The H_2 was supposed to be used to replace natural gas in existing pipelines. Several obvious questions were not asked:

1. A pipeline for methane cannot be used to transport H_2. To avoid leakage, one needs totally different valves and connectors. Furthermore, as the volume of H_2 is triple that of methane, one needs larger pipes and different compressors. Of course, the same factors would apply to the distribution of the gases to customers. None of the valves or controls would be tight enough, and all burners would have to be changed.
2. As nuclear reactors are available, why not use electricity instead of methane for most applications? Electricity is far more efficient than H_2 or methane (by a factor of two) for most uses. Let me illustrate. Today, in the United States and the world, a large amount of electricity is generated from natural gas. This electricity, on a BTU basis, costs by a factor of about 2.2 more than the natural gas feed. The fact that we do it obviously means that electricity is more valuable than natural gas. At the same price, it would have driven out natural gas. H_2 as a fuel has an even lower value than natural gas. If we now convert electricity to H_2, we again lose at least one-third of the energy. Thus, H_2 made from electricity has a value about four times lower than the electricity. This simple calculation illustrates the stupidity of the idea to convert nuclear or solar energy to hydrogen. At that time, no one was foolhardy enough to suggest using H_2 generated by a nuclear reactor as fuel for a fuel cell. Furthermore, an electricity distribution system is already in place and can grow with demand.
3. Why not make H_2 by electrolysis, a well-known process? Nobody asked why—in contradiction to our cumulative technical experience—a set of chemical reactions should generate a large increase in ΔT more cheaply or efficiently than electrolysis. The main item in the cost of electricity is the nuclear reactor, and if a thermochemical cycle is less efficient, it is inherently more expensive.
4. The cost of switching to hydrogen is astronomical. To produce the equivalent of one million BTU of hydrogen per day by electrolysis would require a 25-kilowatt nuclear reactor at an investment cost of \$150,000, which means that a standard nuclear reactor of one gigawatt costing \$6 billion will be able to supply 40,000 million BTU/day of H_2 the

pockets and resulting perturbations. *J. Chem. Phys.* **117**: 3753–3761.

8. Bowron, D.T., Soper, A.K., and Finney, J.L. (2001). Temperature dependence of the structure of a 0.06 mole fraction tertiary butanol–water solution. *J. Chem. Phys.* **114**: 6203–6219.
9. Hummer, G. et al. (1996). An information theory model of hydrophobic interactions. *Proc. Natl. Acad. Sci.* **93**: 8951–8955.
10. Botti, A. et al. (2004). The microscopic structure of concentrated NaOH solutions. *J. Chem. Phys.* **120**: 10154–10162.
11. Bowron, D.T. and Finney, J.L. (2003). Structure of a salt–amphiphile–water solution and the mechanism of salting out. *J. Chem. Phys.* **118**: 8357–8372.
12. Sacco, A., De Cillis, F.M., and Holz, M.J. (1998). *Chem. Soc. Faraday Trans.* **94**: 2089.

THE MIRAGE OF THE H_2 ECONOMY

Reuel Shinnar
The City College of the CUNY
New York

In a recent paper (1), I analyzed the new “hydrogen economy” and explained why, for most uses, it makes no sense to convert fuels, electricity, or any other form of energy to hydrogen, only to use the hydrogen as a fuel for fuel cells. As H_2 is not a resource but an inefficient energy carrier, it would appear that the H_2 economy is just a mirage. If alternative energy sources are ever developed, an economy based on electricity would be cheaper (by a factor of three) and much easier to adopt than H_2. Oddly, I had covered the same topic twenty years ago in a paper that analyzed the hydrogen economy of the 1970s (2). Then, the investment in H_2 research was about 10 billion wasted dollars before it was phased out. In this column, I will try to explain why the scientific establishment engages in such obviously futile behavior and why this mirage attracts so many followers.

In the mid 1970s, I became aware of a strange phenomenon. Many leading scientific journals were suddenly giving enormous coverage to hydrogen, the fuel of the future; a new term, hydrogen economy, was born. International meetings attracted thousands of participants and the United States, Germany, and Japan funded the necessary development with vast resources. The central idea, put forward by the International Atomic Energy Research Center in Vienna, used a high-temperature nuclear reactor as the heat source. In this reactor, pressurized helium was used as the coolant. To generate electricity, the hot helium was first expanded through a turbine and then used for steam generation before being fed back to the reactor.

The “thermochemical cycle,” the name applied to this new process, does not include a turbine, and the heat from cooling the hot helium is used to drive a set of closed-cycle chemical reactions; the net result is that water is split into hydrogen and oxygen.

Following is a hypothetical example of such a reaction:

$$MeO + Heat \longrightarrow Me + 1/2O_2$$
$$Me + H_2O \longrightarrow H_2 + MeO + Heat$$

where Me is any metal. No such reaction was ever discovered, and all cycles proposed were quite complex. One does not need to be an experienced cost estimator to realize that such a cycle would be very expensive and have a low thermal efficiency.

The H_2 was supposed to be used to replace natural gas in existing pipelines. Several obvious questions were not asked:

1. A pipeline for methane cannot be used to transport H_2. To avoid leakage, one needs totally different valves and connectors. Furthermore, as the volume of H_2 is triple that of methane, one needs larger pipes and different compressors. Of course, the same factors would apply to the distribution of the gases to customers. None of the valves or controls would be tight enough, and all burners would have to be changed.
2. As nuclear reactors are available, why not use electricity instead of methane for most applications? Electricity is far more efficient than H_2 or methane (by a factor of two) for most uses. Let me illustrate. Today, in the United States and the world, a large amount of electricity is generated from natural gas. This electricity, on a BTU basis, costs by a factor of about 2.2 more than the natural gas feed. The fact that we do it obviously means that electricity is more valuable than natural gas. At the same price, it would have driven out natural gas. H_2 as a fuel has an even lower value than natural gas. If we now convert electricity to H_2, we again lose at least one-third of the energy. Thus, H_2 made from electricity has a value about four times lower than the electricity. This simple calculation illustrates the stupidity of the idea to convert nuclear or solar energy to hydrogen. At that time, no one was foolhardy enough to suggest using H_2 generated by a nuclear reactor as fuel for a fuel cell. Furthermore, an electricity distribution system is already in place and can grow with demand.
3. Why not make H_2 by electrolysis, a well-known process? Nobody asked why—in contradiction to our cumulative technical experience—a set of chemical reactions should generate a large increase in ΔT more cheaply or efficiently than electrolysis. The main item in the cost of electricity is the nuclear reactor, and if a thermochemical cycle is less efficient, it is inherently more expensive.
4. The cost of switching to hydrogen is astronomical. To produce the equivalent of one million BTU of hydrogen per day by electrolysis would require a 25-kilowatt nuclear reactor at an investment cost of $150,000, which means that a standard nuclear reactor of one gigawatt costing $6 billion will be able to supply 40,000 million BTU/day of H_2 the

equivalent of 6700 barrels of oils. Building the capacity for supplying a new pipeline that could carry 12% of the current U.S. capacity of natural gas would have to be built from the start at a cost of close to $1 trillion. For thermochemical cycles, the cost would be significantly higher. All of this must be invested before a pipeline can be started. The same energy needs could be supplied by electrical energy at only one-quarter the cost because electrical energy has a higher efficiency for almost all purposes. The investment for electricity can be spread over many years because a grid is already in existence. Furthermore, just the energy from the nuclear reactions would cost $50 per million BTU H_2.

5. The most dangerous of all fuels, pressurized H_2 ignites with an invisible flame. Its explosive limits are very wide, and its minimum ignition energy is ten times smaller than that of any other gaseous fuel. This problem occurs not only when H_2 is being distributed or used, but even more so when it is generated in a thermochemical cycle. When H_2 is produced by electrolysis, one can locate the plant far away from the nuclear reactor generating the electricity, which, however, is not true for helium, which must be pumped in such large quantities that it cannot be sent over long distances at reasonable cost. Society has accepted the risks of explosion in conventional fuels but tries to limit exposure. Thus, even a small propane storage tank may not be transported through a tunnel, but what level of risk from a potential explosion at a nuclear power plant complex is acceptable? None of these critical questions were asked then—nor are they being asked today.

These questions have been on my mind for a long time. In the “age of technology,” how could the scientific community have been misled by such obvious nonsense? And why had so many engineering firms and large companies come up with totally unrealistic cost estimates, at least by a magnitude too low, and thermal efficiency estimates that are unrealistically high?

The intellectual climate of the time can be illustrated by a small story. *Science* carried a six-page article, featured on the cover, about a new cycle developed by General Electric that promised to produce H_2 at $2.50/million BTU H_2 (at that time it would have cost $7.00/million BTU to produce H_2 from oil). For fun, I flow-sheeted the process and published the mass and heat balance in a letter to *Science* (3) suggesting that if GE can reduce costs to $2.50, they should forget about hydrogen and change the economy. By comparison, us poor chemical engineers could not achieve such results at a cost of $250/million BTU H_2. Three months later, I was invited to GE's central laboratory where the Director of Chemical Research offered me a consulting job. The head of the lab was a chemist who had read about the need for a thermochemical cycle, and he actually developed the first one that worked. His engineers had given him this cost estimate. When my letter was published, he asked Corporate Engineering to make their own estimate and they confirmed my numbers. This story illustrates a key aspect of the hydrogen economy. Many excellent scientists lack the skills needed to deal with economics and costs. Later, I published another paper demonstrating, from elementary design principles, that any thermochemical process must, by necessity, cost much more than electrolysis (2).

It is true that many scientists could have erred because they were not knowledgeable about process economics. However, I also came to realize that other forces were driving the program. Although they are inherently safer, high-temperature nuclear reactors of the type required by thermochemical cycles were deemed less competitive and more costly than boiling water reactors, which created an opportunity for their developers to seek more government support. A large nuclear research community was hungry for grants and contracts.

Interestingly and unfortunately, the episode of the 1970s has been ignored or forgotten. However, before we place large national resources behind an effort to again deal with such major problems as renewable energy, global warming, and energy independence, we have to understand what went wrong the first time a so-called “hydrogen economy” was attempted.

The paper I just published is about the reincarnation of the hydrogen economy. This time, all the proposed processes are feasible, but the program itself seems to make even less sense. As before, the main emphasis is on the replacement of conventional fuels with a national distribution system of H_2 at a tremendous loss in efficiency in the production and transmission steps. As before, the same intractable safety and cost problems still persist.

Similar driving forces remain at work. Large research budgets are projected to find a use for a technology in which DOE has invested huge sums of money to little or no avail. This time, it is fuel cells powered by hydrogen, a technology that is good for space applications, where H_2 is used as a rocket fuel, and therefore available in small quantities, and other specialty applications, but useless for applications that couple H_2 generation with the generation of electricity. Such systems have half the efficiency of combined cycle turbines, cost over five times more (1,4), and create at least double the greenhouse emission, probably much more.

Today, there even is talk about switching cars from gasoline to H_2, which makes the transition problem even more difficult as the new supplying system cannot start locally, and this would also require the use of pressurized H_2 at 6000 psi, a pressure very seldom used in industry. Nobody mentions that in industry an H_2 tank with the capacity proposed for a family car requires storage in a special room with a blowout wall. Any technician approaching such a pressurized tank to check the valves swings a wooden broom to prevent being fried by an invisible flame which may come from a leaking valve, as above 1500 psi, an H_2 leak is self-igniting (1). For a gas station, a minimum-sized storage tank would have the explosive power of two thousand tons of TNT or two of the largest bombs used by the Air Force. A terrorist could just open a valve and detonate a small bomb after 20 seconds.

The most obvious fallacy of the new H_2 economy, however, is that it solves all of our problems (5,6). For those of us who accept the first and second laws of thermodynamics, it is obvious that the use of H_2 would increase the use of other energy resources. Arguments are made that, once we have a hydrogen distribution system, we could capture the CO_2 produced and sequester it. However, we have no idea how to sequester such vast amounts of CO_2 safely. A successful policy can only be formulated after all the elements are solved and a good deal is known about costs. At present, the United States releases 100 million tons of sequestered CO_2 back into the atmosphere each year.

In a system in which the basic science is known, costs are reduced by innovative ideas that develop during the implementation stage of a project rather than by more research. The cost of liquid natural gas plants and combined cycle power plants decreased in the last ten years by a factor of two even though the basic technology did not change because many were built to compete in the market. So, research, to be effective, needs to be coupled with a clear plan for implementation.

In order to promote the use of fuel cells or other new technologies, which are by one to two orders of magnitude more expensive than available technologies, the research community has invented the myth of "learning curves." In the last 40 years, computers became, through miniaturization, cheaper by several magnitudes, but this option does not apply to energy systems. In the last 20 years, the cost of solar cells was reduced by less than a factor of two despite a tremendous research budget. Better engineering and mass production can reduce the cost of most new technologies by a factor of two to three. However, this would require the building of large plants based on competitive bidding, not a new research breakthrough. Large research budgets over the past 40 years have not made the cost of fuel cells more competitive.

The alternative sources that really reduce CO_2 emissions include solar and nuclear energy and biomass. The use of H_2 just makes every one of these options several times more expensive. Solar and nuclear energy produce electricity, and, for most purposes, electricity is far more efficient to use than H_2. To provide electricity to the motor in a H_2 car requires three to four times the solar or nuclear capacity required for providing the electricity directly via a battery. Although we don't have good electric cars yet, Toyota has just come out with a hybrid car with a plug-in battery large enough for a 40-mile drive. Such a car could reduce total gasoline consumption by 80%, but electric cars would only reduce greenhouse emissions if the electricity were generated by solar energy. As for biomass, which is a valuable but limited future resource, conversion by fermentation to alcohol followed by conversion to gasoline or diesel is far cheaper and thermally more efficient than conversion to hydrogen.

In the early 1990s, an Israeli company, Luz (7), built solar electric power plants using a high-temperature heat transfer fluid, which included sufficient storage to supply Los Angeles with 300 megawatts of electricity on a reliable basis. The Luz plant had another major advantage; the design allowed for conventional fuel to be used as a backup because, even in the desert, 10% of the days may be rainy. To be profitable, all Luz needed was a 50% tax break on the investment. When Luz became large, the tax breaks were taken away. If Luz had been allowed to build a few gigawatts, the price would have come down and other companies would have come in; we might have had affordable solar energy today.

We have today the resources and the technology to do something useful about global warming if we really had the resolve. For example, for \$50 billion a year over twenty years, we could replace 40% of coal power plants with solar power plants using proven technology; we just have to engineer them better and provide them with sufficient storage capacity. We are the only western nation who has enough sun and enough area for solar power. Alternatively, we could eliminate about 80% of our gasoline used for cars. We could easily afford to do both, but we have to realize that no research can make solar, nuclear, or other alternative energy competitive with cheap oil or gas.

Our society has solved far more difficult large-scale technical problems successfully, e.g., space travel and the most advanced defense weapons in the world. Applying our capabilities to global warming or to resource depletion to preserve a livable world to our grandchildren would require not only conviction but a strong political will to do so.

To give a recent example, thirty years ago, we faced the problem of increasing pollution from our coal power plants. We had the technology (scrubbers) to reduce the pollution substantially. It would have cost about \$20–30 billion, but politics and the influence of senators from coal-producing states prevented the adoption of such a solution. Instead, we spent \$20 billion on "clean coal" research programs that failed to produce results that could help solve this problem. However, the "clean coal" effort did provide large research budgets along with the illusion that we were accomplishing something. Thirty years later, the same plants are producing more electricity and consuming more coal—still poisoning the fish we eat with mercury, causing smog that blocks us from seeing the sky, destroying our forests, and, most importantly, impairing the health of all of us more than any other form of pollution.

The private car is another example of the obstacles that a serious national policy faces. In the 1970s, in response to the Arab oil embargo, oil price increases, and an alleged likelihood that the oil resource base would become depleted within the next several decades, Presidents Nixon, Ford, and Carter proposed new energy policy initiatives that included the goal of improving auto fuel economy. These initiatives resulted in the Corporate Average Fuel Economy standards (CAFÉ), which required that the fleet fuel economy of new light-duty vehicles (5500 lbs or less) reach 27.5 mpg by a certain, date a 30% reduction in gasoline demand. This standard limited the ability of the companies to produce large cars and stationwagons. The same provision set lower standards in mileage and emissions for light-duty trucks weighing less than 5500 lbs.

Nine years ago, the car industry realized that one could produce large cars (SUVs) that fit all the definitions of a light truck, such that one should be able to fold all seats to the floor to provide a major space occupying the car and several others. It was a great success. As light trucks convertible to luxury passenger cars do not count as passenger cars for the CAFE laws and are not considered in the fleet average, this allowed car companies again to produce gas guzzlers. Eight years later, 30% of the passenger cars sold in the United States are SUVs, and all the efficiency gains of 30 years have been wiped out.

As we face the potential of global warming and resource depletion that might endanger the future of our grandchildren, we escape our responsibilities by pretending that all problems can be solved better and at lower cost through the mirage of more research. Thus, the H_2 mirage allows car companies to continue to produce gas-guzzling, highly polluting SUVs while touting the glorious future, an "H_2" car that will miraculously be energy efficient and pollution-free. And, as in the 70s, thousands in the scientific and political communities are promoting this mirage to the public.

Thus, mirages like the "hydrogen economy" play an important psychological function in our society, and many in the technical community play along. The role engineers and scientists should be playing today is to educate the public and the politicians about the real options. We should be reminded of an old cartoon in the Pogo series that bears the caption "We have met the enemy and he is us."

BIBLIOGRAPHY

1. Shinnar, R. (2003). The hydrogen economy, fuel cells, and electric cars. *Technology in Society* **25**(4): 453–576.
2. Shinnar, R., Shapira, D., and Zakai, S. (1981). A thermochemical and hybrid cycles for hydrogen production. A differential economic comparison with electrolytes. *IEC Process Design & Development* **20**: 581.
3. Shinnar, R. (1975). Thermochemical hydrogen generation heat requirements and costs. *Science* **188**: 1036.
4. Shinnar, R. (2003). Fuel-cell faux pas. *Chemical & Engineering News* **81**(14): 4.
5. Rifkin, J. (2003). The hydrogen economy. *Environmental Magazine*. January.
6. Bush, G.W. (2003). *State of the Union Address*. January 28.
7. Shinnar, R. (1993). The rise and fall of luz. *ChemTech* **23**: 50–53.

READING LIST

Energy Research of the Department of Energy. (2001). *Was It Worth It?* National Research Council.

Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards. (2002). National Research Council.

HYDROGEN ION

Gholam A. Kazemi
Shahrood University of Technology
Shahrood, Iran

INTRODUCTION

This article includes an introduction to the ion and its properties followed by an overview of the characteristics of the hydrogen atom. The latter has been extracted mostly from Sienko and Plane (1). This article should be studied in conjunction with the pH and Hydronium Ion articles.

An ion in general terms can be any grouping of one or more atoms that carries a charge (1). The characteristics of an ion include:

1. the sign (type) of charge: positive (cation) or negative (anion);
2. the number of charges or valence (z): monovalent divalent, etc;
3. the ionic radius (*r*): the distance from the center of the nucleus to the outer shell of the atomic particle;
4. the ionic potential: *z/r*, the ratio of the charge of the ion to the ionic radius;
5. the electronegativity: the power of the ion to attract electrons;
6. the molar volume: the volume of one mole (cm^3/mol); and
7. the ionic conductance: the ability to carry electrical current.

When an ion is surrounded by many other ions/molecules such as in a solution, its other properties such as ionic strength and ion activity coefficient come to light. Ottonello (2) is a major reference where many properties of various ions are tabulated.

Hydrogen, the lightest of all gases, is colorless, odorless, and tasteless. Together with helium, it occupies the first period of the periodic chart of elements and is the most abundant element in the solar system. After helium, it has the second lowest boiling temperature of all elements (2). Because of its low boiling point, liquid hydrogen is used as a cryogenic fluid (to produce low temperatures), and because of its high heat of combustion (120 kJ/g), it is a valuable rocket fuel. Molecular hydrogen, H_2, can combine directly with most elements. Mixtures of H_2 and O_2 can be explosive. Much of the industrial consumption of hydrogen is in hydrogenation reactions, the addition of H_2 to other molecules. Water and power are required to produce hydrogen. In its compounds, hydrogen is found in three oxidation states, +1, −1, and 0. Hydrogen is such an important element that it has its own association. The National Hydrogen Association (WWW. Hydrogenus.org), for example, was formed in 1989 (in the U.S.) to act as a focal point in hydrogen related matters. Anyone interested in water related matters should know the important characteristics of hydrogen because of its place in the water molecule.

DEFINITION OF A HYDROGEN ION

A hydrogen ion, H^+, is a monovalent cation that occurs when a hydrogen atom loses its only electron:

$$H \longrightarrow H^+ + e^-$$

Therefore, the hydrogen ion has only one proton and no electron. This is why "proton" is another name for hydrogen ion. Some references state that there is no such ion as a hydrogen ion and what occurs, is actually a hydronium ion, H_3O^+, rather than a hydrogen ion (3). The hydride ion, H^-, occurs when a hydride such as NaH dissociates:

$$\mathrm{NaH} \longrightarrow \mathrm{Na^+ + H^-}$$

In this reaction, the hydride ion radius is 1.46 Å. Molecular hydrogen is formed from the reaction of a hydride ion and a hydrogen ion (4):

$$\mathrm{H^+ + H^- \longrightarrow H_2}$$

$H_3{}^+$ is an ion that is artificially generated by the interaction of H^+ and H_2. For instance, generation of $H_3{}^+$ is the most critical specification of the ion source in a mass spectrometer (5).

CHARACTERISTICS OF THE HYDROGEN ION

Some major characteristics of the hydrogen ion are presented in Table 1.

In chemistry and geochemistry, it is assumed that the standard heat capacity, C_p^0; standard partial molal Gibbs free energy from the elements, G_f^0; standard partial molal enthalpy of formation from the elements, H_f^0; standard partial molal entropy, S_0^p; and standard partial molal volume, V_p^0, of the hydrogen ion equal zero at all temperatures and pressures (2,7).

MEASUREMENT OF HYDROGEN ION CONCENTRATION

More than 100 years ago, Arrhenius suggested that acid properties are attributable to the presence of an easily detachable hydrogen ion, H^+ (1). Now, everyone knows that the concentration of hydrogen ion in any solution is a measure of the solution's acidity or pH ($\mathrm{pH} = -\log \mathrm{H^+}$). Thus, the measurement of H^+ concentration is better described in the pH article of the encyclopedia and is only briefly dealt with here.

The H^+ ion is the easiest ion concentration to measure. This is done using a hydrogen electrode, which is formed by bubbling H_2 over a platinum electrode surface (7). A variety of these electrodes are commercially available; Metrohm pH meter Model 744 and Orion pH meter Model 420A are examples. Measurement of pH cannot be more accurate than 0.02 pH units because the standard solutions or buffers which have been calibrated by the U.S. Bureau of Standards (NBS) have an absolute accuracy in the range of ±0.01 to 0.02 pH units (7). In field conditions, where disturbing factors are ample, the accuracy is usually ±0.05 pH units. Colorimetry is one older method of measuring H^+ concentration.

Table 1. Notable Characteristics of the Hydrogen Ion, H^{+a}

Type	Cation
Valence	1
Radius[b]	—
Ionic potential (IP)[c]	—
Electronegativity	2.2
Molar volume	0 (This is based on convention)
Ionic conductance	349.8 $\mathrm{cm^2 \times (equiv\Omega)^{-1}}$
Charge to mass ratio (q/m)	9.6×10^7 C/kg

[a] References 4, 6, 7.

[b] It is difficult to define because the electron of the hydrogen atom has gone, and only the proton remains. So, there is no outer shell left to measure its distance to the proton.

[c] It cannot be calculated because there is no specified ionic radius.

Some Examples

Pure neutral water at 25 °C has a pH of 7, a H^+ concentration of 10^{-7} mg/l. A 1-molal solution of HCl will have an H^+ concentration of 1 and a 0.01-molal solution will have an H^+ concentration of 10^{-2} mg/l. The concentration of hydrogen ion in a 0.01-molal solution of a weak such as H_2CO_3 is only $10^{-4.2}$ mg/l, and it is $10^{-10.8}$ mg/l in an aqueous solution of pH 3.2.

Many authors prefer to use the term "hydrogen ion activity" rather than hydrogen ion concentration. The reason that the "pH" concept has been developed is the very low concentration of hydrogen ion and the difficulty in comprehending such low, sometime ambiguous, numbers. In solutions, the concentration of H^+ never reaches zero; there is no solution that is 100% basic or 100% acidic.

PRODUCTION OF HYDROGEN ION IN WATER

Hydrogen ion is produced in natural water from the three following main sources as well as from humic and fluvic acids, volcanic gases, acid rain, and short-chain organic acids present in some oil field brines (8).

1. Hydrolysis

$$\mathrm{H_2O \longrightarrow H^+ + OH^-}$$

As stated above, however, it might be more appropriate to write the previous equation in the form:

$$\mathrm{2H_2O \longrightarrow H_3O^+ + OH^-}$$

2. Dissociation

$$\mathrm{H_2CO_3 \longrightarrow H^+ + HCO_3{}^-}$$

3. Oxidation

$$\mathrm{2FeS_2 + 7.5O_2 + 7H_2O \longrightarrow 2Fe(OH)_3 + 8H^+ + 4SO_4{}^{2-}}$$

IMPORTANCE OF THE HYDROGEN ION

Some of the important functions and behaviors of the hydrogen ion are described as following. Considering the huge and universal importance of H^+, these should be regarded only as examples.

1. The radius of the hydrogen ion is very close to the size of the neutron. This similarity is the principle used to design and develop compensated neutron

logs (CNL) and neutron probes. The former is used as an important tool for calculating the porosity of petroleum reservoirs and aquifers (9,10), and the latter is of extensive use in agricultural practices for measuring soil moisture (11).

2. The concentration of hydrogen ion, pH, is a master variable of the groundwater system because the hydrogen ion participates in many important chemical reactions such as (12):

   a: mineral dissolution—precipitation

   $$CaCO_3 + H^+ = Ca^{2+} + HCO_3{}^-$$

   b: aqueous complexation

   $$Fe^{3+} + H_2O = Fe\,(OH)^{2+} + H^+$$

   c: adsorption −desorption

   $$\equiv FeOH + Cu^+ = \equiv FeOCu + H^+$$

3. The hydrogen ion plays a major role in denitrification through which various forms of nitrogen in the soil revert to N_2 (6):

   $$4NO_3{}^- + 4H^+ \longrightarrow 2N_2 + 5O_2 + 2H_2O$$

   It also plays a major role in redox reactions and acid−base reactions.

4. In sanitary engineering and water treatment, pH is an important factor in chemical coagulation, disinfection, water softening, corrosion control, and biological processes (13).
5. The most remarkable function of hydrogen ion is its control of the acidity or alkalinity of solutions.

BIBLIOGRAPHY

1. Sienko, M.J. and Plane, R.A. (1979). *Chemistry: Principles and Applications*. McGraw-Hill, New York.
2. Ottonello, G. (1997). *Principles of Geochemistry*. Columbia University Press, New York.
3. Fetter, C.W. (1988). *Applied Hydrogeology*. 2nd Edn. CBS Publishers and Distributors, Delhi.
4. Lagowski, J.J. (1973). *Modern Inorganic Chemistry*. Marcel Dekker, New York.
5. Attendorn, H.G. and Bowen, R.N.C. (1997). *Radioactive and Stable Isotope Geology*. Chapman & Hall, London.
6. Evangelou, V.P. (1998). *Environmental Soil and Water Chemistry: Principles and Applications*. John Wiley & Sons, New York.
7. Langmuir, D. (1997). *Aqueous Environmental Geochemistry*. Prentice-Hall, Upper Saddle River, NJ.
8. Hounslow, A.W. (1995). *Water Quality Data: Analysis and Interpretation*. Lewis, New York.
9. Hearst, J.R., Nelson, P.H., and Paillet, F.L. (Eds.). (2001). *Well Logging for Physical Properties: A Handbook for Geophysicists, Geologists, and Engineers*, 2nd Edn. John Wiley & Sons, New York.
10. Todd, D.K. (1980). *Groundwater Hydrology*, 2nd Edn. John Wiley & Sons, New York.
11. Wilson, E.M. (1984). *Engineering Hydrology*, 3rd Edn. Macmillan, London.
12. Deutsch, W.J. (1997). *Groundwater Geochemistry: Fundamentals and Applications to Contamination*. Lewis, Boca Raton, FL.
13. Tebbutt, T.H.Y. (1995). *Principles of Water Quality Control*. Reprint of 4th Edn. Butterworth Heinemann, Oxford.

THE HYDRONIUM ION

Carl W. David
University of Connecticut
Storrs, Connecticut

Naked protons cannot exist in aqueous solution. Even when one writes

$$HCl(g) \rightarrow H^+(aq) + Cl^-(aq) \tag{1}$$

one knows that the $H^+(aq)$ is a representation of a significantly more complicated moiety (or, to be more precise, set of moieties). The simplest representation of $H^+(aq)$ is the hydronium ion (sometimes also referred to as the oxonium ion or the hydroxonium ion),

$H_3O^+(aq)$, but this is just the first in a series the next few members of which are $H_5O_2{}^+$, $H_7O_3{}^+$, $H_9O_4{}^+$, and in summary $(H_2O)_nH_3O^+$, the hydrated hydronium ions (Fig. 1).

The hydronium ion is isoelectronic with ammonia (NH_3) and therefore, ought to be pyramidal with the oxygen at the apex of the pyramid and the three protons disposed below, with symmetry C_{3v}. Like ammonia, it should invert, with the O[or N] passing through the plane of the three protons from one side to the other. In the planar transition state, when all four nuclei are coplanar, the symmetry should be D_{3h}. Also, like ammonia, there should be a barrier to this umbrella inversion; appropriate tunneling has been observed spectroscopically (ν_2). It is thought that the O–H distance is about 0.96 Å, and the H–O–H angle about 115°. The barrier to inversion, a saddle point on the energy hypersurface, is thought to be between 1 and 5 kcal/mol (Fig. 2).

The simplest occurrence of the hydronium ion can be seen in the gas phase water dimer, $(H_2O)_2$. The proton transfer reaction inside the dimer would be

$$H_2O\text{–}H_2O(g) \rightarrow H_3O^+(g) : OH^-(g) \tag{2}$$

forming an ion pair (Fig. 3).

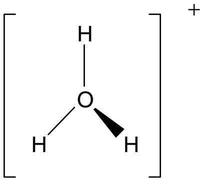

Figure 1. The hydronium cation.

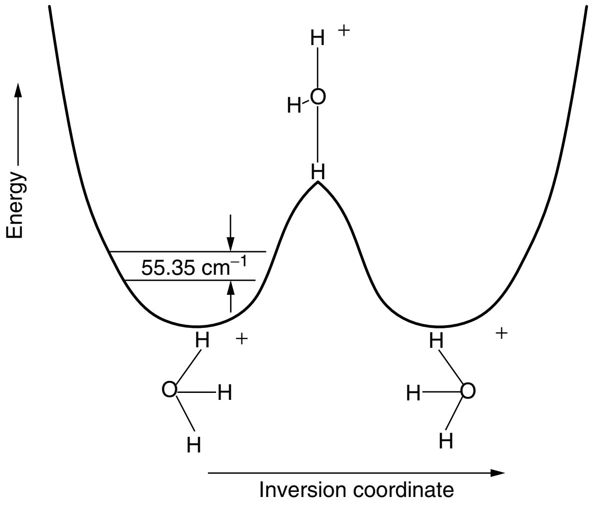

Figure 2. Potential energy surface for the hydronium ion's inversion coordinate.

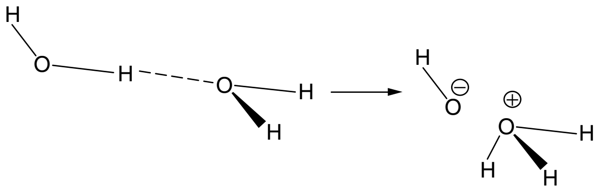

Figure 3. Dimer water forming an ion pair through proton transfer.

In the dimer, one molecule of water is a proton donor to the other (which is a proton acceptor). The former water is a Bronsted acid, the latter a Bronsted base. On the line of the hydrogen bond, $HO–H_{donor}::::OH_2$, the movement of the donor proton from one oxygen to another is equivalent to the proton transfer (above). The O–O distance in the dimer is about 3 Å, and the OH distances are about 1 Å, so the shift of 1 Å in a proton's position (between two oxygens) creates an $OH^-:::H_3O^+$ ion pair.

Hydrated versions of this ionization process illustrate that the actual process in liquid water is not cleanly definable. Figure 4 shows the results of three water molecules involved in ionization, whereas Fig. 5 shows the results with four water molecules.

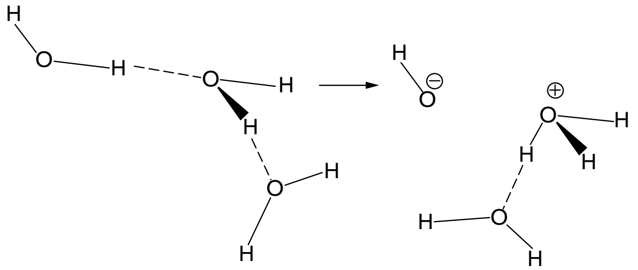

Figure 4. With three molecules of water, one can form a hydrated hydronium cation and a hydroxide anion.

Clearly, in liquid water, the hydroxide and hydronium ions are solvated but not statically enough to allow definition of a moiety.

In aqueous media, for the analog of the proton transfer reaction, one could write

$$2H_2O(l) \rightarrow H_3O^+(aq) + OH^-(aq) \tag{3}$$

the prototype of all acid–base chemistry in an aqueous medium.

In aqueous acidic media, the hydronium ion is rarely written. Instead, the older notation is employed. For instance, for acetic acid, CH_3COOH, one would write

$$CH_3COOH(aq) \Leftrightarrow H^+(aq) + CH_3COO^-(aq) \tag{4}$$

partial dissociation into a proton and a acetate anion, rather than the more accurate

$$H_2O(l) + CH_3COOH(aq) \Leftrightarrow H_3O^+(aq) + CH_3COO^-(aq) \tag{5}$$

The conductance properties of water are attributable in part to the hydronium cation and its peculiar properties. The Grotthuss mechanism for proton transfer in ice employs hydronium ions hydrogen bound to adjacent water molecules, in which, like the dimer, an excess proton migrates (or tunnels) from one oxygen to an adjacent one, moving only a tiny distance, whereupon the charge has shifted enormously (one O–O distance).

A similar mechanism for $OH^-(aq)$ charge migration most likely also obtains. Here a water donor adjacent to the hydroxide anion shifts its donating proton to the hydroxide, thereby itself becoming a hydroxide anion (Fig. 6).

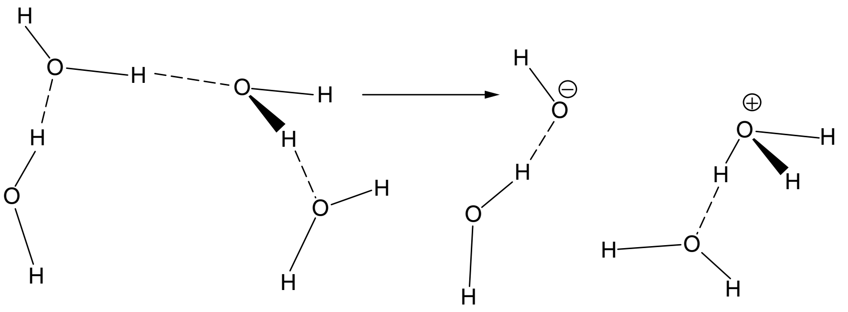

Figure 5. With four waters, hydrated hydroxide and hydrated hydronium ions are possible.

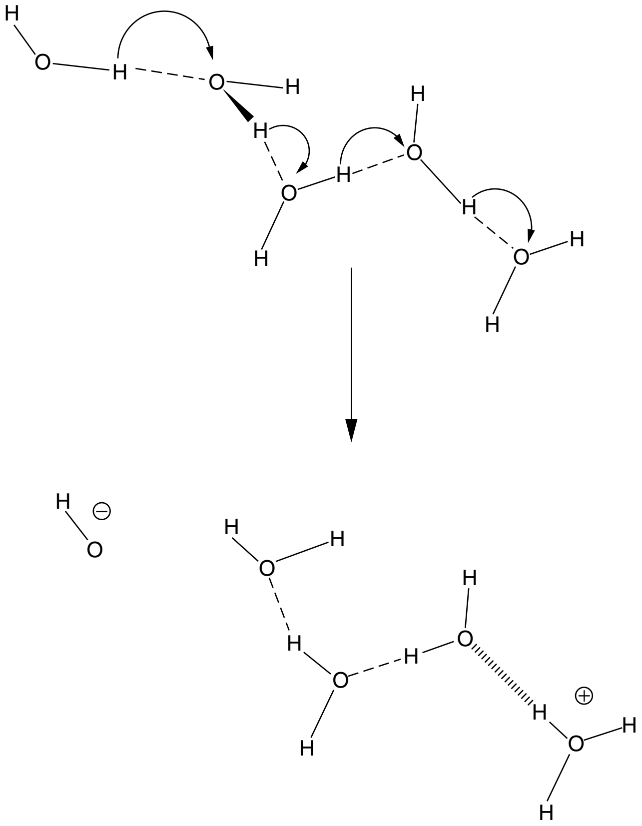

Figure 6. The Grotthuss mechanism for positive and negative charge creation and migration in aqueous solution.

INFILTRATION AND SOIL MOISTURE PROCESSES

PAUL R. HOUSER

(from *The Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts*, Wiley 2003)

Infiltration is the process of water entry from surface sources such as rainfall, snowmelt, or irrigation into the soil. The infiltration process is a component in the overall unsaturated *redistribution* process (Fig. 1) (1) that results in *soil moisture* availability for use by vegetation transpiration, exfiltration (or evaporation) processes, chemical transport, and groundwater recharge. Soil moisture, in turn, controls the partitioning of subsequent precipitation into infiltration and runoff, and the partitioning of available energy between sensible and latent heat flux.

Because of the importance of soil moisture on multiple processes, its definition can be elusive (2); however, it is most often described as moisture in the unsaturated surface layers (first 1 to 2 m) of soil that can interact with the atmosphere through evapotranspiration and precipitation (3).

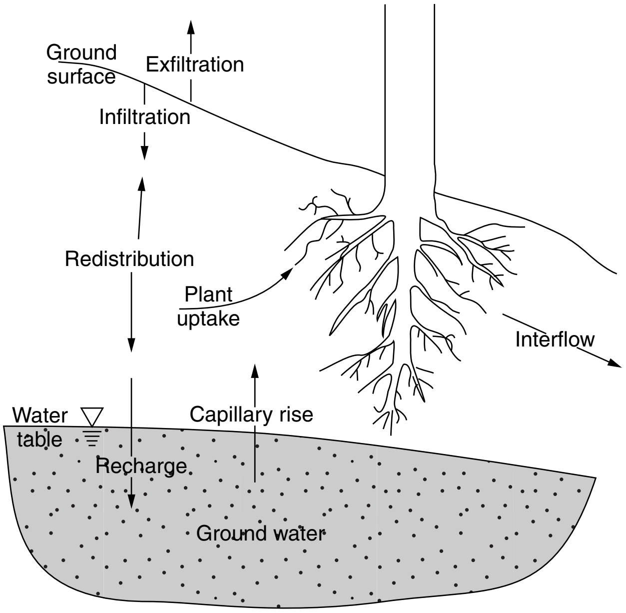

Figure 1. Unsaturated zone definition and active processes (1).

CONTROLS ON INFILTRATION AND SOIL MOISTURE

To characterize soil moisture and infiltration, the physical controls on these processes must be considered. The primary soil controls will be considered in this chapter; however, other factors such as soil chemistry, thickness, soil layering or horizons, and preferential flow paths, as well as vegetation cover, tillage, roughness, topography, temperature, and rainfall intensity also exert important controls (4).

A soil's particle size distribution has a large impact on its hydraulic properties. Soil particles less than 2 mm in diameter are divided into three texture groups (sand, silt, and clay) that help to classify broad soil types and soil water responses (Fig. 2) (5). The type of clay and the coarse material over 2 mm in diameter can also have a significant impact on soil water properties. An overview of methods for determining particle size properties is given by Gee and Bauder (6).

Bulk density, ρ_b (M/L^3) is the ratio of the weight of dry solids to the bulk volume of the soil, and *porosity*, ϕ (M^3/M^3), is the total volume occupied by pores per unit volume of soil:

$$\phi = \frac{V_a + V_w}{V_s} = 1 - \frac{\rho_b}{\rho_m} \tag{1}$$

where V_s (L^3) is the total volume of soil, V_a (L^3) is the volume of air, V_w (L^3) is the volume of water, and ρ_m (ML^{-3}) is the particle density (normally about 2.65 g/cm^3).

The volumetric water content, or soil moisture, θ (L^3 L^{-3}) is the ratio of water volume to soil volume:

$$\theta = \frac{V_w}{V_s} = \frac{W_w}{W_d}\frac{\rho_b}{\rho_w} \tag{2}$$

where W_w (M) is the weight of water, W_d (M) is the weight of dry soil, and ρ_w (M/L^3) is the density of water.

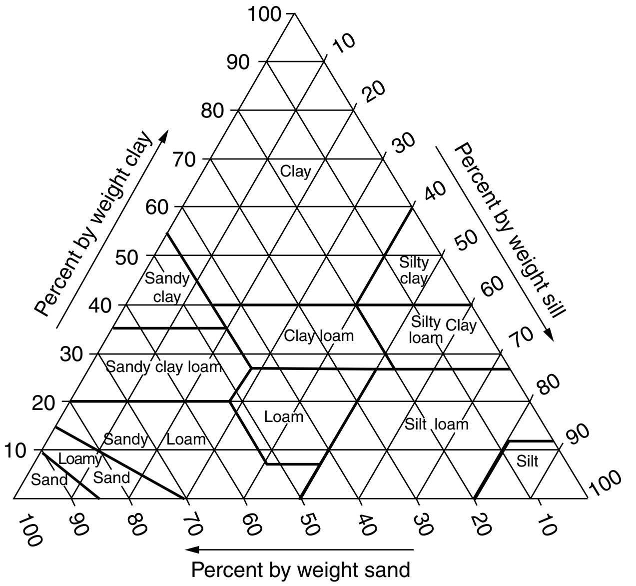

Figure 2. Soil textural triangle describing the relationship between texture and particle size distribution (5).

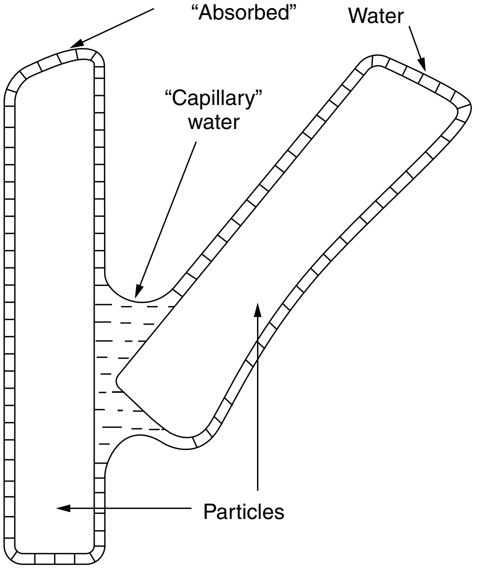

Figure 3. Capillarity and adsorption combine to produce suction (7).

Soil moisture can vary in both time and space, with a theoretical range from 0 to φ, but for natural soils the range is significantly reduced due to isolated pore space and tightly held or "adsorbed" water (Fig. 3) (7). If a soil is saturated, then allowed to drain until the remaining water held by surface tension is in equilibrium with gravitational forces, it is at *field capacity*, θ_f. Vegetation can remove water from the soil until the *permanent wilting point*, θ_w, is reached. Therefore, the *available water content* for plant use, $\theta_a = \theta_f - \theta_w$. Typical ranges of porosity, field capacity, and wilting point for different soils are given in Fig. 4 (8).

In unsaturated soils, water is held in the soil against gravity by surface tension (Fig. 3). This tension, suction, or *matric potential*, ψ (L), increases as the radii of curvature of the meniscus or water content decreases (Fig. 5) (9). Matric potential is expressed in reference to atmospheric pressure, so for saturated soil $\psi = 0$ and for unsaturated soil $\psi < 0$.

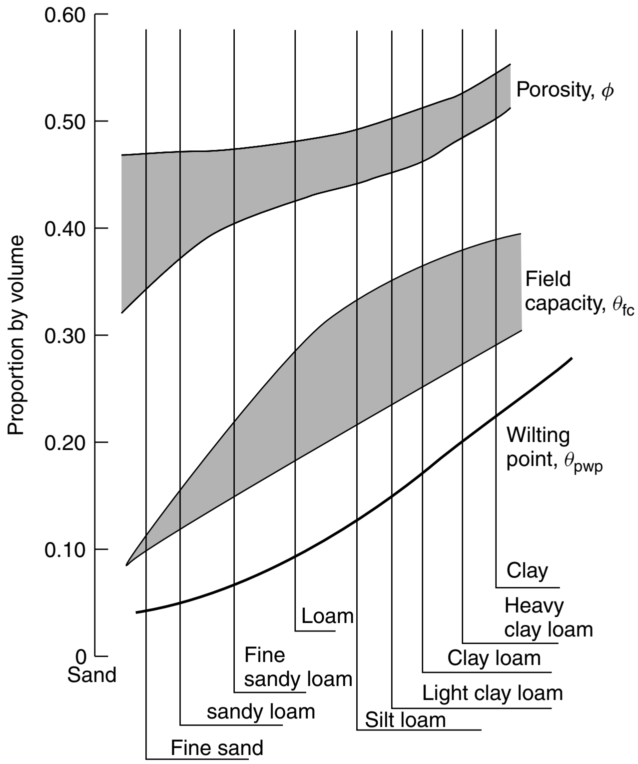

Figure 4. Water holding properties of various soils (8).

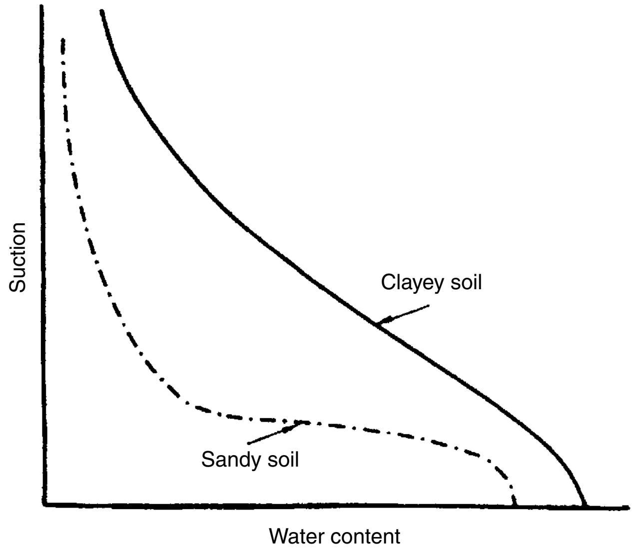

Figure 5. Effect of texture on water retention characteristics (9).

The *hydraulic conductivity*, *K* (L/T), is a measure of the ability of the soil to transmit water that varies nonlinearly over a large range depending on both soil properties and water content (Fig. 6) (10). Many laboratory and field hydraulic conductivity measurement methods exist for

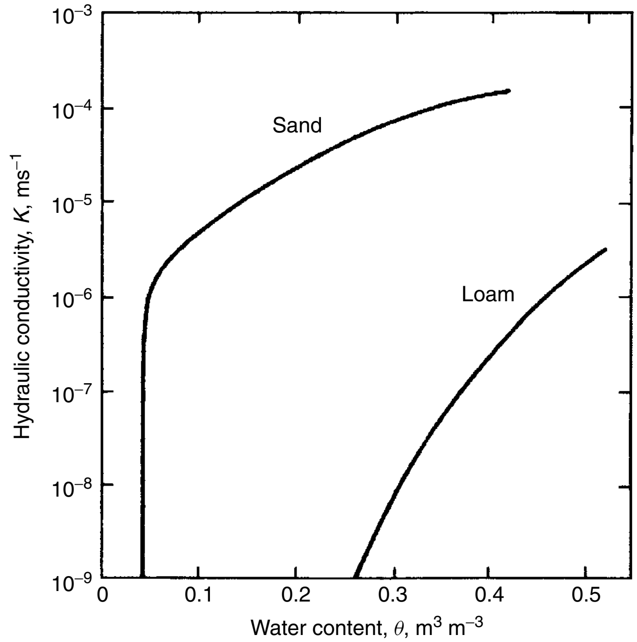

Figure 6. Effect of texture and soil moisture on hydraulic conductivity (10).

use with various soils; see Bouwer and Jackson (11) or Green et al. (12) for details.

Soil water content can significantly impact infiltration by (1) increasing the hydraulic conductivity, which increases infiltration, and (2) reducing the surface tension that draws moisture into the soil, which reduces infiltration. The net effect of these impacts depends on the water content itself, the water input rate, and duration and the distribution of hydraulic conductivity.

The *water retention characteristic* describes a soil's ability to store and release water and is defined by the relationship between soil moisture and the matric potential (Fig. 5). This is a power function relationship that has been described by Brooks and Corey (13) and Van Genuchten, (14) among others. The water tension characteristic is usually measured in air pressure chambers where the water content of a soil sample can be monitored over a wide pressure range (15).

The water retention relationship may actually change between drying and wetting due to the entrapment of air in soil pores (Fig. 7) (16). For practical applications, this effect, called *hysteresis*, is usually neglected (17).

PRINCIPLES OF SOIL WATER MOVEMENT

Through experiments on saturated water flow through sand beds, Darcy (18) found that the rate of flow, Q (L^3/T), through a cross-sectional area A (L^2), is directly proportional to head loss (e.g., water elevation difference), ΔH (L), and inversely to the flow path length, Δl (L):

$$Q = KA\frac{\Delta H}{\Delta l} \tag{3}$$

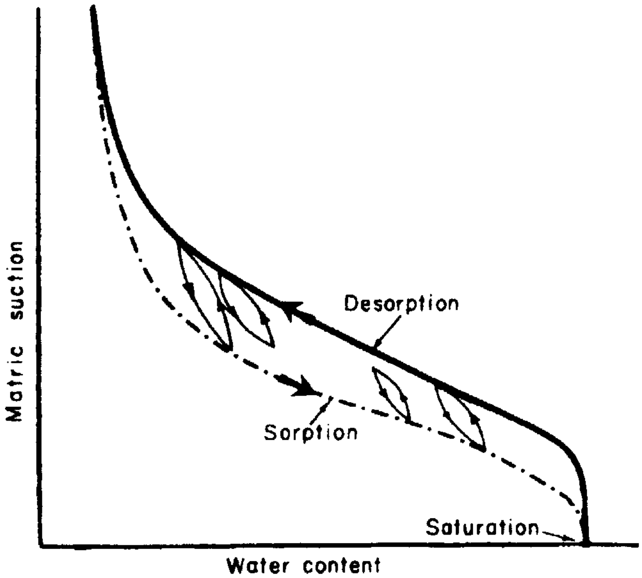

Figure 7. Changes in water retention characteristics between sorption and desorption (16).

Combining *Darcy's law* with the law of conservation of mass results in a description of unsaturated flow called *Richards equation* (19):

$$\frac{\partial\theta}{\partial t} = \frac{\partial}{\partial z}\left(\frac{K}{C}\frac{\partial\theta}{\partial z}\right) - \frac{\partial K}{\partial z} \tag{4}$$

where $C = -\partial\theta/\partial\psi$ is the water content change in a unit soil volume per unit matric potential, ψ change. The Richards equation is the basis for most simulations of infiltration and redistribution of water in unsaturated soil. Using some approximations, analytical solutions of the Richards equation are available (20,21) that show good agreement with observations (22). The Richards equation is based on saturated flow theory, and does not account for all of the processes active in natural systems, so it may not always perform well (23).

INFILTRATION ESTIMATION

Some basic principles that govern the movement of water into the soil can be used to predict infiltration. The *infiltration capacity*, f(L), is the maximum rate that a soil in a given condition can absorb water and generally decreases as soil moisture increases. If the *rainfall rate* is less than the infiltration capacity, then infiltration proceeds at the capacity rate. However, if the rainfall rate exceeds the infiltration capacity, then infiltration proceeds at the capacity rate, and the excess rainfall ponds on the surface or runs off. As the time from the onset of rainfall increases, infiltration rates decrease due to soil moisture increases, raindrop impact, and the clogging of soil pores, until a steady-state infiltration rate is reached (Fig. 8) (24). Existing infiltration models use empirical, approximate, or physical approaches to predict infiltration.(25)

Empirical. Empirical infiltration models generally utilize a mathematical function whose shape as a function of

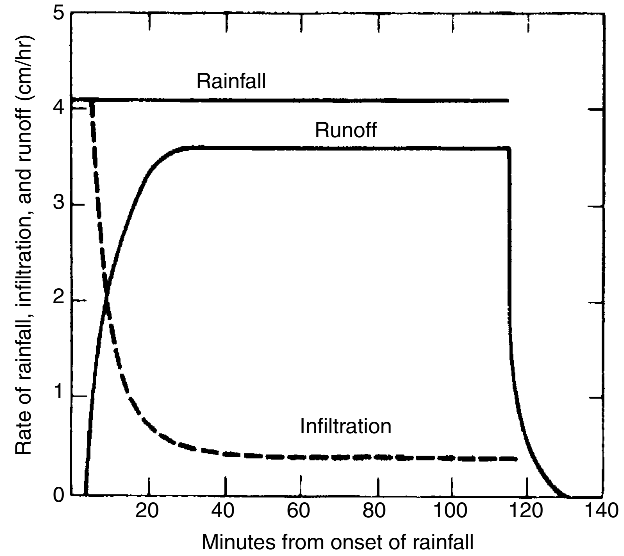

Figure 8. Idealized relationship between rainfall, infiltration, and runoff rates (24).

time, t, matches observations and then attempts a physical explanation of the process.

Kostiakov (26) proposed the simple infiltration rate, f(L/T) model:

$$f = \alpha\gamma t^{\alpha-1} \tag{5}$$

where α and γ are constants that have no particular meaning and must be evaluated by fitting the model to experimental data.

Horton's (27) infiltration model has been widely used in hydrologic simulation. It relates infiltration capacity to initial infiltration rate, and f_0, the constant infiltration rate at large times, f_c:

$$f = f_c + (f_0 - f_c)e^{-\beta t} \tag{6}$$

where β is a soil parameter describing the rate of decrease of infiltration.

Approximate. Analysis approximations to the Richards equation are possible if several simplifying assumptions are made. Most approximate infiltration models treat the soil as a semi-infinite medium, with the soil saturating above a wetting front.

Green and Ampt (28) assumed in a soil with constant hydraulic properties, the matric potential at the moving wetting front is constant, leading to a discontinuous change in soil moisture at the wetting front:

$$f = K\left[1 + \frac{(\varphi - \theta_i)S_f}{F}\right] \tag{7}$$

where S_f (L) is the effective suction at the wetting front, θ_i is the initial water content, and F (L) is the accumulated infiltration.

Phillip (29) proposed that the first two terms in a series of powers of $t^{1/2}$ could be used to approximate infiltration:

$$f = \tfrac{1}{2}St^{1/2} + A \tag{8}$$

where S is a parameter called sorptivity, t is time from ponding, and A is a constant that depends on soil properties. In this model, the infiltration rate approaches a constant equal to the hydraulic conductivity at the surface water content, and the wetting front advances without changing its shape and approaches a constant velocity.

Physical. Recent advances in numerical methods and computing has facilitated the practical application of the Richards equation to realistic flow problems. Such packages can simulate water infiltration and redistribution using the Richards equation and including precipitation, runoff, drainage, evaporation, and transpiration processes (30).

INFILTRATION MEASUREMENT

Infiltration rates can be measured at a point using a variety of methods described here, each appropriate for certain conditions. However, because of the large temporal and spatial variability of infiltration processes, catchment average infiltration rates may be desired, which can be obtained through the water balance analysis of rainfall–runoff observations (31).

Ring Infiltrometer. This simple method is most appropriate for flood irrigation or pond seepage infiltration. A cylindrical metal ring is sealed at the surface and flooded. Intake measurements are recorded until steady-state conditions are reached (32). If the effects of lateral flow are significant, then a double-ring infiltrometer can be used. Due to ponding conditions within the ring, observed infiltration rates are often higher than under natural conditions (33).

Sprinkler Infiltrometer. This method is appropriate for quantifying infiltration from rainfall. Artificial rainfall simulators are used to deliver a specified rainfall rate to a well-defined plot. Runoff from the plot is measured, allowing computation of the infiltration rate (34,35).

Tension Infiltrometer. The tension or disk infiltrometer employs a soil contact plate and a water column that is used to control the matric potential of the infiltrating water. By varying the tension, the effect of different size macropores can be determined (36,37).

Furrow Infiltrometer. This method is useful if information on infiltration of flowing water in irrigation furrows is desired. Either the water added to a small section of blocked off furrow to maintain a constant depth or the inflow—outflow of a furrow segment can be monitored to determine the infiltration characteristics of the system (38).

SOIL MOISTURE MEASUREMENT

Soil water content can be determined directly using gravimetric techniques or indirectly by inferring it from a property of the soil (39,40).

Gravimetric. The oven-drying soil moisture measurement technique is the standard for calibration of all other methods but is time consuming and destructive. The method involves obtaining a wet soil sample weight, drying the sample at 105 °C for 24 h, then obtaining the dry sample weight [see Eq. (2)].

Neutron Thermalization. High-energy neutrons are emitted by a radioactive source into the soil and are preferentially slowed by hydrogen atoms. The number of slow neutrons returning to the detector are a measure of soil moisture.

Gamma Attenuation. The attenuation in soil of gamma rays emitted from caesium-137 is directly related to soil density. If the soil's bulk density is assumed to be constant, then changes in attenuation reflect changes in soil moisture (41).

Time-Domain Reflectometry (TDR). TDR measures the soil's dielectric constant, which is directly related to soil moisture, by measuring the transmit time of a voltage pulse applied to a soil probe.

Tensiometric Techniques. This method measures the capillary or moisture potential through a liquid-filled porous cup connected to a vacuum gage. Conversion to soil moisture requires knowledge of the water retention characteristic.

Resistance. The electrical resistance or conductivity of a porous block (nylon, fiberglass, or gypsum) imbedded in the soil depends primarily on the water content of the block. However, because of salinity and temperature sensitivity, measurements of these sensors are of limited accuracy (42).

Heat Dissipation. Changes in the thermal conductivity of a porous block imbedded in the soil depend primarily on the water content of the block. The dissipation of a heat pulse applied to the block can be monitored using thermistors, then the soil water content can be determined from calibration information.

Remote Sensing. Soil moisture can be remotely sensed with just about any frequency where there is little atmospheric absorption (43). But, it is generally accepted that long wavelength, passive microwave sensors have the best chance of obtaining soil moisture measurements that contain little error introduced by vegetation and roughness and offer great potential to remotely sense soil moisture content with depth due to differential microwave absorption with varying dielectric constant (44).

SPATIAL AND TEMPORAL VARIABILITY

Natural soils exhibit considerable spatial heterogeneity in both the horizontal and vertical directions, and at all distance scales from the pore to the continent to a degree that it is difficult to capture this variability in routine measurements [45,46]. This large variations in soil properties, infiltration, and soil moisture over relatively small areas makes it difficult to transfer the understanding of processes developed at a point to catchment scales. Many hydrological models assume that a single spatially representative average soil property can be used to characterize catchment (or even larger) scale processes. It is clear from the nonlinear character of soil water processes that catchment average infiltration cannot be computed based on catchment average soil properties. It is also clear that the physical meaning of a soil property, say porosity, is relative to the volume over which it is averaged [47]. However, there is a need to understand and reduce this complexity for the purposes of prediction and management. Several approaches, including dividing the catchment into hydrologically similar subareas [48], various statistical approaches [49], and scaling and similarity theory [50,51] have made headway toward an understanding of infiltration and soil moisture spatial variability, but are not being widely used in practical applications.

One of the most important recent findings in this regard is the scale invariance of soil water movement(i.e., the Richards equation) remain uniform across spatial scales [52]. This new understanding of the underlying symmetry of the Richards equation may help to facilitate a workable scale invariant analytical soil water dynamical model.

Finally, there is acontinuing need for the observation of soil properties, soil moisture, and infiltration processes at multiple scales to facilitate understanding and prediction of these complex and socially significant processes. It is likely that remote sensing of soil moisture and other land surface factors will be instrumental in this respect.

BIBLIOGRAPHY

1. Dingman, S.L. (1994). *Physical Hydrology*. Prentice-Hall, Englewood Cliffs, NJ, p. 211.
2. Lawford, R.G. (1992). An Overview of soil moisture and its role in the climate system. In: F.J. Eley, R. Granger and L. Martin (Eds.). *Soil Moisture; Modeling and Monitoring for Regional Planning*, National Hydrology research Centre Symposium No.9 Poceedings, pp. 1–12.
3. Schmugge, T., Jackson, T.J., and McKim, H.L. (1979). Survey of in-situ and remote sensing methods for soil moisture determination. In: *Satellite Hydrology*. American Water Resources Association.
4. Rawls, W.J., Ahuja, L.R., Brakensiek, D.L., and Shirmohammadi, A. (1993). Infiltration and soil water movement. In: *Handbook of Hydrology*. D.R. Maidment (Ed.). McGraw-Hill, New York, pp. 5.1–5.51.
5. Hillel, D. (1982). *Introduction to Soil Physics*. Academic, New York, p. 29.
6. Gee, G.W., and Bauder, J.W. (1986). Particle size analysis. In: *Methods of Soil Analysis, Part I—Physical and Mineralogical Methods*, 2nd Edn. A. Klute (Ed.). *American Society of Agronomy Monograph* 9, pp. 383–411.
7. Hillel, D. (1980). *Fundamentals of Soil Physics*. Academic, New York, p. 143.
8. Dunne, T. and Leopold, L. (1978). *Water in Environmental Planning*. W.H. Freeman, New York, p. 175.
9. Hillel, D. (1980). *Fundamentals of Soil Physics*. Academic, New York, p. 150.
10. Marshall, T.J., and Holmes, J.W. (1988). *Soil Physics*, 2nd Edn. Cambridge University Press, p. 87.
11. Bouwer, H. and Jackson, R.D. (1974). Determining soil properties. In: *Drainage for Agriculture*. J. van Schilfgaard (Ed.). American Society of Agronomy, Madison, WI, pp. 611–672.

12. Green, R.E., Ahuja, L.R., and Chong, S.K. (1986). Hydraulic conductivity, diffusivity, and sorptivity of unsaturated soils—field methods. In: *Methods of Soil Analysis, Part I—Physical and Mineralogical Methods*, 2nd Edn. A. Klute (Ed.). American Society of Agronomy Monograph 9, pp. 771–798.
13. Brooks, R.H. and Corey, A.T. (1964). *Hydraulic Properties of Porous Media*, Hydrology Paper 3, Colorado State University, Fort Collins, CO.
14. Van Genuchten, M.Th. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* **44**: 892–898.
15. Cassel, D.K. and Klute, A. (1986). Water potential: Tensiometry. In: *Methods of Soil Analysis, Part I—Physical and Mineralogical Methods*, 2nd Edn. A. Klute (Ed.). American Society of Agronomy Monograph 9, pp. 563–596.
16. Hillel, D. (1980). *Fundamentals of Soil Physics*. Academic, New York, p. 153.
17. Rawls, W.J., Ahuja, L.R., Brakensiek, D.L., and Shirmohammadi, A. (1993). Infiltration and Soil Water Movement. In: *Handbook of Hydrology*. D.R. Maidment (Ed.). McGraw-Hill, New York, pp. 5.5–5.6.
18. Darcy, H. (1856). *Les fontaines publiques de la ville de Dijon*. Dalmont, Paris.
19. Richards, L.A. (1931). Capillary conduction of liquids in porous mediums. *Physics* **1**: 318–333.
20. Kuhnel, V., Dooge, V.C.I., Sander, G.C., and O'Kane, J.P.J. (1990). Duration of atmosphere-controlled and soil-controlled phases of infiltration for constant rainfall at a soil surface. *Ann. Geophys.* **8**: 11–20.
21. Sposito, G. *Recent advances associated with soil water in the unsaturated zone*. In: *Reviews of Geophysics*, Supplement, U.S.National Report to International Union of Geodesy and Geophysics 1991–1994, pp. 1059–1065.
22. Whisler, F.D. and Bouwer, H. (1970). A comparison of methods for calculating vertical drainage and infiltration in soils. *J. Hydrol.* **10**: 1–19.
23. Nielsen, D.R., Biggar, J.W., and Davidson, J.M. (1962). Experimental consideration of diffusion analysis in unsaturated flow problems. *Soil Soc. Am. Proc.* **26**: 107–111.
24. Dunne, T. and Leopold, L. (1978). *Water in Environmental Planning*, W.H. Freeman, New York, p. 169.
25. Rawls, W.J., Ahuja, L.R., Brakensiek, D.L., and Shirmohammadi, A. (1993). Infiltration and soil water movement. In: *Handbook of Hydrology*. D.R. Maidment (Ed.). McGraw-Hill, New York, pp. 5.21–5.23.
26. Kostiakov, A.N. (1932). On the dynamics of the coefficient of water-percolation in soils and on the necessity for studying it from a dynamic point of view for purposes of amelioration. *Trans. Sixth Comm. Intern. Soil. Sci. Soc. Russian, part A*, 17–21.
27. Horton, R.E. (1940). An approach toward a physical interpretation of infiltration-capacity. *Soil Sci. Soc. Am. J.* **5**: 399–417.
28. Green, W.H., and Ampt, G.A. (1911). Studies on soil physics: 1. Flow of air and water through soils. *J. Agric. Sci.* **4**: 1–24.
29. Phillip, J.R. (1957). The theory of infiltration: 1. The infiltration equation and its solution. *Soil Sci.* **83**: 345–357.
30. Ross, O.J. (1990). Efficient numerical methods for infiltration using Richards equation. *Water Resour. Res.* **26**: 279–290.
31. Soil Conservation Service, Hydrology. (1972). In: *SCS National Engineering Handbook*. U.S. Department of Agriculture, Washington, DC, Sec. 4.
32. Bouwer, H. (1986). Intake rate: Cylinder infiltrometer. In: *Methods of Soil Analysis, Part I—Physical and Mineralogical Methods*, 2nd Edn. A. Klute (Ed.). *American Society of Agronomy Monograph* 9, pp. 825–844.
33. Lukens, R.P. (Ed.). (1981). *Annual Book of ASTM Standards*, Part 19: *Soil and Rock, Building Stones*, pp. 509–514.
34. Agassi, M., Shainberg, I., and Morin, J. (1988). Effects on seal properties of changes on drop energy and water salinity during a continuous rainstorm. *Aust. J. Soil Res.* **26**: 1–10.
35. Peterson, A. and Bubenzer, G. (1986). Intake rate sprinkler infiltrometer. In: *Methods of Soil Analysis, Part I—Physical and Mineralogical Methods*, 2nd Edn. A. Klute (Ed.). *American Society of Agronomy Monograph* 9, pp. 845–870.
36. Ankeny, M.D., Kaspar, T.C., and Horton, R. (1988). Design for an automated tension infiltrometer. *Soil Sci. Soc. Am. J.* **52**: 893–896.
37. Perrouix, K.M. and White, I. (1988). Designs of disc permeameters. *Soil Sci. Soc. Am. J.* **52**: 1205–1215.
38. Kincaid, D.C. (1986). Intake rate: Border and furrow. In: *Methods of Soil Analysis, Part I—Physical and mineralogical Methods*, 2nd Edn. A. Klute (Ed.). *American Society of Agronomy Monograph* 9, pp. 871–887.
39. Gardner, W.H. (1986). Water content. In: *Methods of Soil Analysis, Part I—Physical and Mineralogical Methods*, 2nd Edn. A. Klute (Ed.). *American Society of Agronomy Monograph* 9, pp. 493–544.
40. Schmugge, T.J., Jackson, T.J., and McKim, H.L. (1980). Survey of methods for soil moisture determination. *Water Resour. Res.* **16**(6): 961–979.
41. Jury, W.A., Gardner, W.R., and Gardner, W.H. (1991). *Soil Physics*, 5th Edn. Wiley, pp. 45–47.
42. Hillel, D. (1982). *Introduction to Soil Physics*. Academic, New York, 1982, pp. 57–89.
43. Schmugge, T.J. (1985). Remote sensing of soil moisture. In: *Hydrological Forecasting*. John Wiley & Sons, New York.
44. Ulaby, F.T., Bradley, G.A., and Dobson, M.C. (1979). Potential application of satellite radar to monitor soil moisture. In: *Satellite Hydrology*, *American Water Resources Association*, pp. 363–370.
45. Jury, W.A., Gardner, W.R., and Gardner, W.H. (1991). *Soil Physics*, 5th Edn. John Wiley & Sons, New York, pp. 268–293.
46. Dingman, S.L. (1994). *Physical Hydrology*. Prentice-Hall, Englewood Cliffs, NJ, pp. 247–250.
47. Bear, J. (1972). *Dynamics of Fluids in Porous Media*. Elsevier, New York.
48. Springer, E.P. and Cundy, T.W. (1987). Field-scale evaluation of infiltration parameters from soil texture for hydrologic analysis. *Water Resour. Res.* **23**: 325–334.
49. Berndtsson, R., and Larson, M. (1987). Spatial variability of infiltration in a semi-arid environment. *J. Hydrol.* **90**: 117–133.
50. Sharma, M.L., Gander, G.A., and Hunt, G.C. (1980). Spatial variability of infiltration in a watershed. *J. Hydrol.* **45**: 101–122.
51. Wood, E.F., Sivaplan, M., and Beven, K. (1990). Similarity and scale in catchment storm response. *Rev. Geophys.* **28**: 1–18.
52. Sposito, G. (1995). *Recent advances associated with soil water in the unsaturated zone*. In: *Reviews of Geophysics*, Supplement, U.S. National Report to International Union of Geodesy and Geophysics 1991–1994, pp. 1059–1065.

ION EXCHANGE AND INORGANIC ADSORPTION

K. Kadirvelu
Jyotsna Goel
Centre for Fire, Explosives, and
Environmental Safety
Timarpur, India

The presence of inorganic ions in water is of major concern for water quality because they can interfere with legitimate use of water for various applications. Inorganic ions, particularly heavy metals, are seriously questionable because of their toxicity and threat to human life and the environment.

Ion exchange and adsorption share so many common features in regard to removal of inorganic ions from water that they can be grouped together as unified processes for inorganic ion treatment. These processes involve the increase in concentration of a particular component at the surface or interface between two phases.

The objective of this article is to enumerate application of adsorption and ion exchange processes for inorganic ion removal from water. This article presents a review of potential adsorbents and ion exchangers for inorganic ion removal. It introduces important topics regarding inorganic ion removal by ion exchange and adsorption processes. Factors effecting metal removal are also presented and discussed in sections of the article. Research results from batch studies are presented to illustrate basic concepts. Accordingly, emphasis is placed on providing a review of adsorbents and ion exchangers available for inorganic ion removal.

ADSORPTION AND ION EXCHANGE: IN GENERAL

Adsorption involves, in general, the accumulation (or depletion) of solute molecules at an interface (including gas-liquid interfaces, as in foam fractionation, and liquid-liquid interfaces, as in detergency). Adsorbent surfaces are often physically and/or chemically heterogeneous, and bonding energies may very widely from one site to another. In discussing the fundamentals of adsorption, it is useful to distinguish between physical adsorption, involving only relatively weak intermolecular forces, and chemisorption, which involves, essentially, the formation of a chemical bond between the sorbate molecule and the surface of the adsorbent. The general features, which distinguish physical adsorption and chemisorption, are presented in Table 1.

Table 1. Distinction Between Physical Adsorption and Chemisorption (1)

Physical Adsorption	Chemical Adsorption
Low heat of adsorption (<2 or 3 times latent heat of evaporation.)	High heat of adsorption (> or 3 times latent heat of evaporation.)
Nonspecific.	Highly specific.
Monolayer or multilayer.	Monolayer only.
No dissociation of adsorbed species. Only significant at relatively low temperatures.	May involve dissociation. Possible over a wide range of temperature.
Rapid, nonactivated, reversible.	Activated, may be slow and irreversible.
No electron transfers, although polarization of sorbate may occur.	Electron transfer leading to bond formation between sorbate and surface.

Ion exchange is an effective means of removing heavy metals from wastewaters. It is a reversible chemical reaction, where the removal of heavy metals is accomplished by the exchange of ions on the resin for those in wastewater. In ion exchange, ions of positive charges (cations) and negative charges (anions) from the fluid (from aqueous solution) replace dissimilar ions of the same charge in solids. The ion exchanger contains permanently found functional groups of opposite charge-type. Cation exchangers generally contain bound sulfonic acid groups; less commonly, groups are carboxylic, phosphonic, and phosphinic, and so some resins involve quaternary ammonium groups (strongly basic) and amino groups (weakly basic).

Most ion exchangers in large-scale use are based on synthetic resins—either preformed and then chemically reacted as styrene or formed from active monomers. Natural zeolites were the first ion exchangers, and natural and synthetic zeolites are in use today.

Ion exchange may be thought of as a reversible reaction and of chemically equivalent quantities. A common example of ion exchange is the familiar water-softening reaction

$$\mathrm{Ca}^{++} + 2\mathrm{NaR} \Leftrightarrow \mathrm{CaR}_2 + 2\mathrm{Na}^+ \qquad (1)$$

where R represents a stationary univalent anionic site in the electrolyte network of the exchanger phase.

FACTORS AFFECTING INORGANIC IONS REMOVAL BY ADSORPTION

Factors affecting inorganic ion removal by adsorbent include (a) nature of the adsorbent (b) solution pH, (c) presence of foreign ions, and (d) temperature.

Nature of the Adsorbent

Surface Area and Pore Structure. Surface area with active functional groups is one of the principle characteristics affecting the adsorption capacity of an adsorbent. The adsorption capacity of a solid adsorbent is generally proportional to the specific surface area of the mass adsorbent; that is, the adsorption of certain solutes increases with an increasing surface area. Table 2 shows the physical properties of a few adsorbents that have been already employed for inorganic removal.

However, the specific surface area alone is frequently inadequate to explain the adsorption capacity of porous solids, such as activated carbons, for different solutes. In view of this, pore size distribution in activated carbon indicates that microspores contribute a major portion of the specific surface area.

Table 2. Properties of Adsorbents Employed for Inorganic Ion Removal (2–6)

Property	Calgon CAL	Coir-Pith Carbon	AC from *Arudo Donax*	AC Cloth	Carbon Aerogel
Surface area (m^2/g)	1100–1300	575	1142	1689	700
Porosity (%)	65.0	93.11	76.0	66.5	36.2
Micropore Volume (g/ml)	0.12	0.0913	0.47	0.964	0.10
Bulk density (g/ml)	0.56	0.12	0.5	0.13	0.644
Moisture content (%)	2.0	5.88	9.3	2.0	18.0
Ash content (%)	7.0	7.87	21.56	13.0	3.5

Depending on their preparation, active carbons also contain larger pores, known as macro and mesopores in the classification proposed by Dubinin (7) and now adopted by the International Union of Pure and Applied Chemistry (IUPAC) (8). The definition of the different types of pores is based on their width w, which represents the distance between the walls of a slit-shaped pore or the radius of a cylindrical pore. One distinguishes micropores, for which w does not exceed 2 nm; macropores, for which w exceeds about 50 nm; and mesopores of intermediate size (2.0–50 nm).

A linear relationship has also been demonstrated between the specific surface area and sorbate molecular size (9).

Chemistry of Adsorbent Surface. The nature and the presence of specific functional groups on the surface of the adsorbent impart certain characteristics that affect the adsorption process. It is now known that surface functional groups are formed during the carbon (adsorbent) activation process.

Attempts have been made to identify and estimate the surface oxygen chemical structures (functional groups) using several physical and chemical techniques, which include neutralization of acid and base (known as Boehm's method); desorption of the oxide layer; potentiometer, thermometric, and radiometric titrations; direct analysis of oxide layers by specific chemical reaction; Infrared (IR) spectroscopy; and X-ray photoelectron spectroscopy (XPS). The surface morphology and metal ion distribution adsorbent samples can be visualized via scanning electron microscopy (SEM). Figure 1 shows SEM micrograph for carbon aerogel equilibrated with Pb (II). SEM of activated carbon surfaces before and after the adsorbent was equilibrated with the metal ion solution clearly shows the presence of Pb (II) ions.

As a result of these investigations, various researchers have postulated the existence of functional groups, such as carboxyic, phenolic, lactonic, aldehydes, ketones, quinones, hydroquinone, anhydrides, and ethereal structures. Also, analysis of physicochemical properties of adsorbent ions provides a deep insight into adsorption mechanism involved during the adsorption process.

Solution pH

Extensive research investigations have been reported for adsorption of inorganic ions at different solution pH values (9–11). The dependence of adsorbent's adsorption capacity for inorganic ion removal from aqueous solution pH has been attributed to changes in surface chemical characteristics of adsorbent and species distribution of the inorganic ions varying with solution pH. In general, the cationic species removal will increase with increasing solution pH (Fig. 2), whereas in the case of inorganic ions present as anionic species, removal generally increases with decreasing pH, which is mainly because of electrostatic repulsion, which can be understood by removal of Cr (VI) by various adsorbents at lower pH range.

Specifically, pH affects the status of the outer hydration sheaths of the metal ion, metal speciation of inorganic ion in aqueous medium, complexation and solubility, and the electrochemical behavior of the carbon surface.

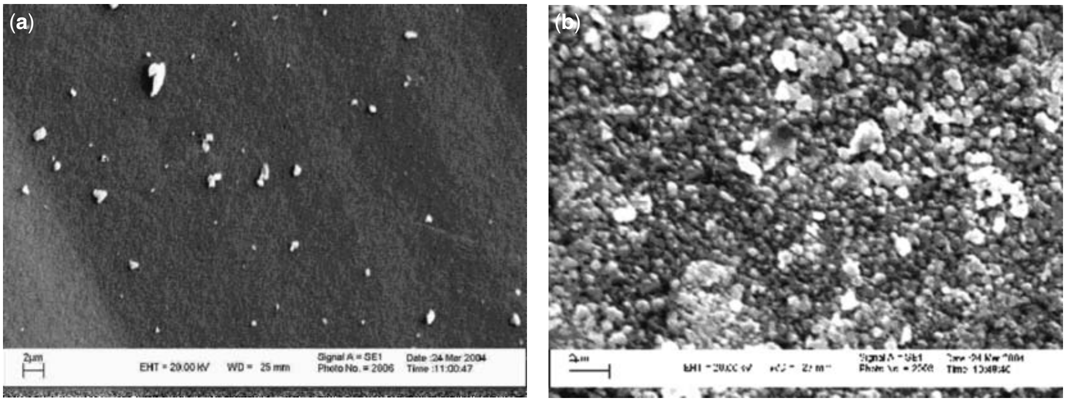

Figure 1. (**a**) SEM of carbon aerogel; (**b**) SEM of carbon aerogel loaded with lead.

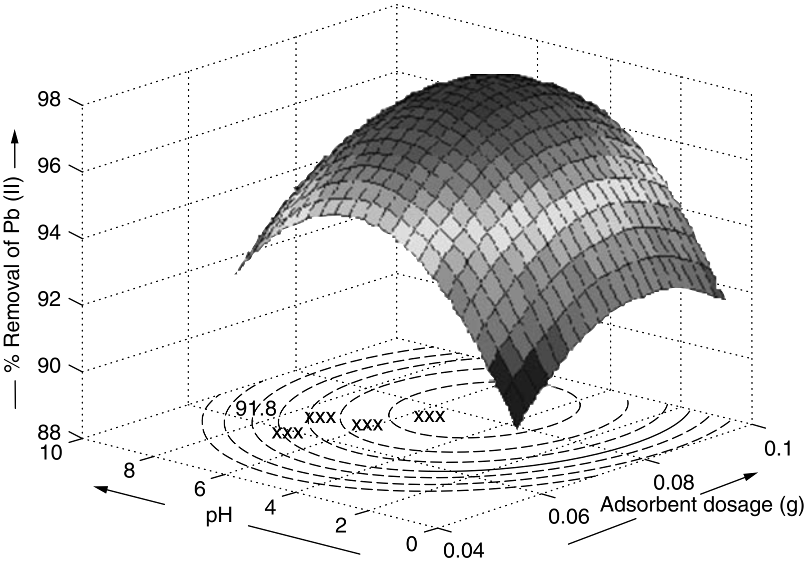

Figure 2. Plot for percent removal of Pb (II) versus pH and adsorbent dosage using carbon aerogel as adsorbent.

Figure 3 depicts possible inorganic ion removal mechanisms for carbon aerogel in relationship to pH, at which solution precipitation begins for Pb (II), Hg (II), and Ni (II). Figure 3 shows clearly that at any pH, the removal of metal ions is greater by adsorption in the presence of carbon aerogel than by precipitation in the absence of carbon. A rapid increase in percentage removal of metal ions is observed after pH 5.4 and 7.5 for Pb (II) and Ni (II), respectively, because of precipitation in the absence of carbon aerogel.

The adsorption of metal ions depends on both the nature of the adsorbent surface and the species distribution of the metal ions in the aqueous solution, which is, in turn, governed by solution pH. Removal noticed in a pH range of 3–5 for Pb (II) and Hg (II), because of their partial hydrolysis resulting in the formation of hydrolyzed metal ion species $[PbOH]^+$, $[HgOH]^+$ Pb $(OH)_2$, and Hg $(OH)_2$. In the case of Ni (II), the maximum adsorption range was shifted to 4–7, which may be because of formation of $[NiOH]^+$ and Ni $(OH)_2$ at these pH (12). Low solubilities of hydrolysis metal ion species may be another reason for maximum adsorption. At pH values below pH_{prec}, adsorption was hypothesized to be the primary removal mechanism. Also, it has been reported that metal removal by activated carbon was found to be inversely proportional to the solubility of the metal (13).

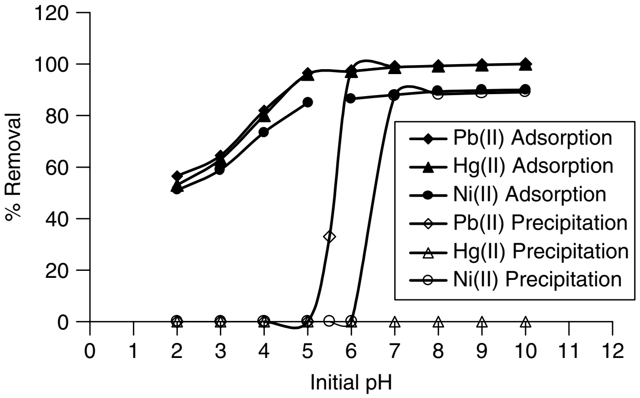

Figure 3. Effect of solution pH on the removal of lead, mercury, and nickel using carbon aerogel.

Namasivayam and Kadirvelu (14) reported an ion exchange phenomenon for Hg (II) removal by the formation of the following surface complexes when Hg (II) is present in the solution with the functional groups present on GAC (ACG-50).

$$2(C_XOH^+) + Hg^{2+} \longrightarrow (C_XO)_2Hg^{2+} + 2H^+ \qquad (2)$$

According to the surface complexation theory, Hg (II) removal increases as pH function increases can be explained on the basis of a decrease in adsorption competition between proton and metal ion species for the adsorbent surface sites by the decrease in positive surface charge, which results in a lesser columbic repulsion forward of the Hg (II) ion.

Presence of Foreign Ions

The presence of complexing ligands and competing adsorbates can alter metal removal from that observed in the mono-component system. The degree of competition is dependent on the type and concentration of competing ions, number of surface sites, and affinity of the surface for adsorbate and the ionic property of adsorbate. Presence of other cocations can influence the adsorption capacity of adsorbent with respect to one adsorbate. This phenomenon is illustrated in Fig. 4, where the adsorption equilibrium curve for Hg (II) removal by GAC is considerably effected by the presence of other metal ions like Cu (II) and Pb (II) (15,16). Similar results have been reported for the effect of EDTA on lead removal using GAC column studies (17).

Temperature Effect

As the process of adsorption is spontaneous, it is accompanied by a decrease in the system's free energy. A

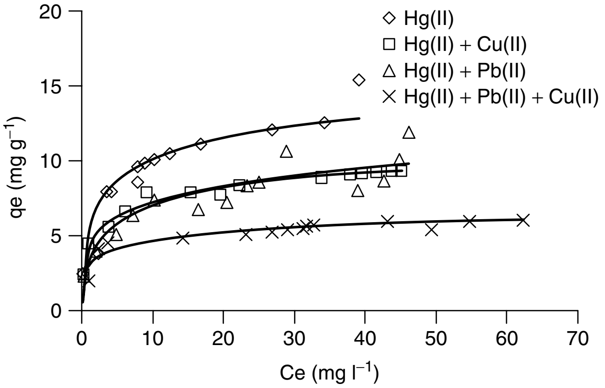

Figure 4. Equilibrium curve for Hg (II) ion in binary and tertiary system with Cu (II) and Pb (II) ions using GAC as adsorbent.

decrease in entropy always exists because of loss of degrees of freedom of the solute in passing from the dissolved state to the adsorbate state. It follows from the relationship

$$\Delta G = \Delta H - T\Delta S \tag{3}$$

that the adsorption process must always be exothermic; that is ΔH must have negative value regardless of the nature of interaction forces. An increase in temperature, therefore, will result in a reduction of the equilibrium adsorption capacity, whereas lower temperature will favor an increased capacity.

As reported by J. Goel et al. (6), according to the adsorption isotherm, the amount of Hg (II) by carbon aerogel increases with increasing temperature. Values of the thermodynamic parameters, such as ΔG^0, ΔH^0, and ΔS^0, describing mercury uptake by carbon aerogel were calculated using the following thermodynamic equations:

$$\Delta G^0 = -RT.\ln K \tag{4}$$

$$\Delta H^0 = [R \times T_1 \times T_2/(T_2 - T_1)] \times \text{Ln}\ (K_2/K_1) \tag{5}$$

$$\Delta S^0 = (\Delta H^0 - \Delta G^0)/T \tag{6}$$

where R is the gas constant and K, K_1, and K_2 are the equilibrium constant at temperature T, T_1, and T_2, respectively. The equilibrium constants was calculated from:

$$K = C_{\text{eq,s}}/C_{\text{eq,1}} \tag{7}$$

where $C_{\text{eq,s}}$ and $C_{\text{eq,1}}$ were the equilibrium concentration of Hg (II) in solution and on the adsorbent, respectively. Values found for ΔG^0, ΔH^0, and ΔS^0, as presented in Table 3, are indicative of the spontaneous nature of the uptake process.

The positive ΔH^0 value confirms the endothermic nature of the sorption process. As diffusion is an endothermic process, it would be expected that increased solution temperature would result in increased uptake of Hg (II) ions from aqueous solution.

Anoop Krishnan et al. (18) has reported that the endothermic nature of the adsorption process, where uptake of mercury on the bagasse piths activated carbon, is favored by increase in temperature. Panayotova (19) and Raji et al. (20) have also presented alike thermodynamics for the uptake of copper and mercury, respectively.

Table 3. Thermodynamic Parameters for the Mercury Adsorption on Carbon Aerogel

Temperature (K)	% Removal	ΔG^0 (J mol^{-1})	ΔH^0 (Jmol^{-1})	ΔS^0 (Jmol$^{-1}K^{-1}$)
293	60	−987.714	+25708.82	+91.11444
303	68	−1898.86	+18990.88	+68.94302
313	73	−2588.29	+32921.47	+113.4497
323	80	−3722.79	+46023.63	+154.0137
333	87	−5262.92	+28133.9	+100.2908
343	90	−6265.83	—	—

ADSORPTION MECHANISM—ION EXCHANGE PROCESS

With some of the adsorbents, like coir pith carbon, adsorption mechanisms have been explained based on the ion exchange model (12). A pure carbon surface is considered to be nonpolar, but in actual practice, some carbon–oxygen complexes C_x O, CO_x and C_xO_2 are usually present (21–23). The surface oxygen complexes hydrolyze water molecules as shown below:

$$C_xO + 2H_2O \longrightarrow C_xOH_2{}^{2+} + 2OH^- \tag{8}$$

$$CO_x + xH_2O \longrightarrow C(OH)^x + xOH^- \tag{9}$$

$$C_xO^2 + H_2O \longrightarrow C_xO^{2+} + 2OH^- \tag{10}$$

As the carbon is prepared with H_2SO_4, $(NH_4)_2S_2O_8$, and $NaHCO_3$, groups such as C_xONa^+, $C_xONa_2{}^{2+}$, C_xSO_3Na, and C_xONH_4 are also present. Na^+ in the above groups also exchange with H^+ in the medium as follows:

$$C_xONa^+ + H^+ \longrightarrow C_xOH^+ + Na^+ \tag{11}$$

$$C_xONa^+ + 2\ H^+ \longrightarrow C_xOH_2{}^{+2} + Na^+ \tag{12}$$

$$C_xONa_2{}^{2+} + H^+ \longrightarrow C_xOH_2{}^{+2} + 2\ Na^+ \tag{13}$$

$$CxSO_3Na + H^+ \longrightarrow C_xSO_3H + Na^+ \tag{14}$$

Excess of Na^+ was introduced into carbon where it was washed with $NaHCO_3$ to neutralize any free sulfuric acid (H_2SO_4) during the preparation of carbon. When metal ion is present in solution, its adsorption will free some H^+ and the pH increase will be lower than in the blank. At the same time, Na^+ will also be released according to reactions as given below:

$$2C_xOH^+ + M^{2+} \longrightarrow (C_xO)_2M^{2+} + 2H^+ \tag{15}$$

$$C_XOH_2{}^{2+} + M^{2+} \longrightarrow C_XOM^2 + 2H^+ \tag{16}$$

$$2C_xONa^+ + M^{2+} \longrightarrow (C_xO)_2M^{2+} + 2Na^+ \tag{17}$$

$$C_XONa_2{}^{2+} + M^{2+} \longrightarrow C_XOM^2 + 2Na^+ \tag{18}$$

$$2C_xSO_3H^+ + M^{2+} \longrightarrow (C_xSO_3)_2M^+ + 2H^+ \tag{19}$$

$$2C_xSO_3Na + M^{2+} \longrightarrow (C_xSO_3)M + 2Na^+ \tag{20}$$

which indicates that apart from exchange of H^+ ions on the adsorbent with M^{2+} ions, significant amount of Na^+ ions, which were present in the adsorbent, were also exchanged with M^{2+} ions.

ION EXCHANGERS FOR INORGANIC ION REMOVAL

By ion exchange, undesirable ions are replaced by others, which do not contribute to contamination of the environment. The method is technologically simple and enables efficient removal of even traces of impurities from solutions. Examples of selective removal of heavy metal ions by ion exchange are presented in Table 4. They include removal of Pb (II), Hg (II), Cd (II) Ni (II), V (IV, V), Cr (III, VI), Cu (II), and Zn (II) from water and industrial wastewaters by various means of modern type of ion exchangers.

By ion exchange, either all ions can be removed from a solution or substances are separated. Therefore, selective removal of ionic contamination and complete deionization can be distinguished. The choice between selective removal and complete deionization depends mainly on the composition of solution and on the extent of decontamination required.

Hg (II) ions sorption on the ion exchanger Imac TMR (Styrenedivinylbenzene Copolymer) can be described by means of the following reactions:

$$2\text{R–SH} + \text{Hg}^{2+} \longrightarrow \text{R–S–Hg–S–R} + 2\text{H}^+ \qquad (21)$$

$$\text{RSH} + \text{HgCl}^+ \longrightarrow \text{R–S–HgCl} + \text{H}^+ \qquad (22)$$

where R is the styrenedivinylbenzene copolymer. After Hg (II) ions sorption, the ion exchanger Imac TMR can be regenerated by means of the concentrated hydrochloric acids.

Metal ions and other cationic impurities are removed on the cation exchanger, whereas chromates are removed on the anion exchangers:

$$\text{nRSO}_3\text{H} + \text{M}^+ \longrightarrow \text{nRM}^{n+} + \text{nH}^+ \qquad (23)$$

$$2\text{ROM} + \text{CrO}_4{}^{2-} \longrightarrow \text{R}_2\text{CrO}_4 + 2\text{OH}^- \qquad (24)$$

$$\text{R}_2\text{CrO}_4 + \text{CrO}_4{}^{2-} + \text{H}^+ \longrightarrow \text{R}_2\text{Cr}_2\text{O}_7 + \text{OH}^- \qquad (25)$$

The NaOH solutions are applied for anion exchanger regeneration. Further developments of conventional ion exchangers, like polystyrenesulphonate cation exchangers and strongly basic anion exchangers, were aimed at large exchange capacity and better chemical, thermal, and mechanical stability. Kinetic capability of ion exchangers could be improved by means of macroporous ion exchangers developed in the 1970s. The range of application of the ion exchange method was remarkably extended, first by invention and application of new organic and inorganic ion exchangers.

ADSORBENTS FOR INORGANIC ION REMOVAL

Large numbers of adsorbents have been developed so far for the removal of inorganic ions, which primarily

Table 4. Selective Removal of Heavy Metal Ions by Ion Exchangers

Metal Ion	Ion Exchangers	Composition and Functional Groups
Pb (II)	Leawtit TP 207 (24)	Chelating ion exchanger with functional iminodiacetate group
	Ethanotricarboxylate, 1,1-ddicarboxylate-2-etanophosphonate (25)	Gel type Copolymer Of Vinyl benzene Chloride, Styrene and Divinyl Benzene
	Polyacrylamidoglycol Resin (26)	Amide, OH and carboxylic groups
	Clinoptilolite (27)	Aluminosilicate structure
Hg (II)	Imac TMR	Styrenedivinylbenzene Copolymer
	Varian BTAM (Nike), Varian BTKM, Varian BSM, Ionac SR-3 (28–31)	Containing nitrogen and sulphur functional group (22), Thiocarbamate group (23), Mercapathile group (24), Thiourea functional group (25)
	Dowex A-1 (32)	Iminodiacetate group
Cd (II)	Dowex 50 W-X4	Polystyrenesulphonic cation exchanger
	Diphonix (33)	Sulphonic groups and diphosphonic groups
	KB-2 and KB-2-12P (34)	Polyacrylate carboxylic cation exchangers
	Phillipsite and Chabazite (35)	Natural zeolite from volcanic rocks of Central Italy
Ni (II) and V (IV, V)	Wofatit CA-20 (36)	Macro porous carboxylic cation exchanger
	Chelex −100 (37)	Iminodiacaetae ion exchanger with Na/H/Ni in exchange system
	Sodium titnate, $Na_4Ti_9O_{20}$ (38)	Layered material with exchangeable Na ions in the titanium oxide layer.
Cr (III, VI)	Lewatit MP-500A (39)	Basic anion exchanger
	Amberlite IR-120 (40)	Polystyrenesulphone cation exchanger

Table 5. Summary Table of Maximum Reported Adsorption Capacities (mg/g)

Sorbent	Source	Cd (II)	Cr (III)	Cr (VI)	Hg (II)	Pb (II)
Bark	(41–44)	32	19.45		400	182
CPEI cotton	(45)				1000	
Chitin	Cr, Hg, Pb (46)	558	92		100	796
Chitosan	(47, 48)	16.50		57	1123	58
Clay	(49, 50)	28				116
Dead biomass	(51)			43.0		
Leaf mould	(46–52)				150	1587
Lignin	Cd, Cr, Hg and Pb (46)	87	17		632	1587
Modified wool	(53)	46.65				
Moss	(54, 55)	5.058	76	43.9	16.2	135
Peat	(54, 55)					
Sawdust	(56)			16.05		230
Seaweed	(57, 58)	215				344
Xanthate	(59, 60)	33.27	19.67		1.149	18
Coir pith carbon	(3, 61)	94.0			154.0	264.0
Fly ash	(62)	198			.073	444.0
Red mud	(63)	67.0			7.1	165.0
Peanut hull carbon	(64)	89.0			220.0	210.0
Furnace Sludge	(65)	7.4				68.0

include activated carbon prepared from a variety of materials of biological origin, industrial wastes, dead biomass, organic wastes, etc. Besides, several low-cost adsorbents such as chitin, chitosin, clay, fly ash, peat, moss, zeolites, xanthates, seaweed/algae/alginate, lignin, and bark/tannin-rich material have also been explored for heavy metal containing wastewaters (37–56). Table 5 presents some of the highest adsorption capacity reported for various low-cost adsorbents.

Although a significant number of low-cost adsorbent materials from sources, the commercial activated carbon (CAC) is still being used intensively to date. Based on its size and shape, activated carbon is classified as follows: Powder (PAC), granular (GAC), Fibre (ACF), and cloth (ACC). As a result of the different sources of raw materials, the extent of chemical activation, and the physicochemical characteristics, each type of activated carbon has its specific application with respect to physicochemical properties, as well as depending on the inherent advantage and disadvantages in wastewater treatment.

Acknowledgment
Authors are grateful to Sh. A. K. Kapoor, Director, Dr. Chitra Rajagopal, Scientist “F”, CFEES, DRDO, for granting permission to write this article and their constant encouragement and support.

BIBLIOGRAPHY

1. Ruthven, D.M. (1984). John Wiley & Sons, New York.
2. Netzer, A. and Hughes, D.E. (1984). *Water Res.* **18**: 927.
3. Kadirvelu, K. and Namasivayam, C. (2000). *Environ. Technol.* **21**: 1091–1097.
4. Basso, M. C., Cerella, E.G., and Cukierman, A.L. (2002). *Ind. Eng. Chem. Res.* **41**: 180–189.
5. Kadirvelu, K., Faur-Brasquet, C., and Le Cloirec, P. (2000). *Langmuir* **16**: 8404–8409.
6. Goel, J., Kadirvelu, K., Rajagopal, C., and Garg, V. K. (2004). *Carbon*. Elsevier Press, Amsterdam, In press.
7. Dubinin, M. M. (1983). *Carbon* **19**: 321.
8. IUPAC. (1972). Manual of symbols and terminology, Appendix 2, Pt.1, colloids and surface Chemistry. *Pure and Appl. Chem.* **3**: 578.
9. Calgon Corporation. *Basic Concept of Adsorption on Activated Carbon*. Calgon, Pittsburgh, PA.
10. Kadirvelu, K., Thamaraiselvi, K., and Namasivayam, C. (2001). *Sep. and Purifi. Technol.* **24**: 497–505.
11. Ricou, P., Lecuyer, I., and Le Cloirec, P. (1998). *Environ. Technol.* **19**: 1005–1062.
12. Kadirvelu, K. (1998). Characterization of coir pith carbon and its utilization in the treatment of metal bearing wastewater. Ph.D. Thesis, Bharthiar University, Coimbatore, India, pp. 59–60.
13. Sigworth, E.A. and Smith, S.B. (1972). *J. AWWA* **64**: 386–391.
14. Namasivayam, C. and Kadirvelu, K. (1999). *Carbon* **37**: 243–256.
15. Goel, J., Kadirvelu, K., and Rajagopal, C. (2004). *Ads. Sci. Tech.* **22**: 257–274.
16. Goel, J., Kadirvelu, K., and Rajagopal, C. (2004). *Environ. Technol.*
17. Reed, B.E. and Arunachalam, S. (1995). *ASCE J. Env. Eng.* **121**(9): 653–663.
18. Panayotova, M.I. (2001). *Waste Mgmt.* **21**: 671–676.
19. Anoop Krishnan, K. and Anirudhan, T.S. (2002). *J. Haz. Mater.* **B92**: 161–183.
20. Raji, C. and Anirudhan, T.S. (1996). Removal of Hg (II) from aqueous solution by sorption on polymerized sawdust. *Ind. J. Chem. Tech.* **3**: 49–54.
21. Huang, C.P. and Ostovic, F.B. (1978). *J. Env. Eng. Div. ASCI.* **104**: 863–878.
22. Sharma, D.C. and Forster, C.F. (1994). *Bioresource Technol.* **47**: 257–264.
23. Yenki, M. K.N. and Natarajan, G. S. (1991). *Separation Sci. Technol.* **26**: 661–674.

24. Bayer, A.G. (2000). Leverkusen.
25. Trochimczuk, W.A. (1999). *Europ. Polym. J.* **35**: 1457–1464.
26. Rivas, B.L., Maturana, H.A., Villegas, S., and Pereira, E. (1998). *Polym. Bull.* **40**: 721–728.
27. Petruzzelli, D., Pagano, M., and Tiravanti, G. (1995). *Ind. Eng. Chem. Res.* **34**: 2612–2617.
28. Bolto, B.A., Pawlowski, L.E., and Spoon, F.N. Ltd. (1987). London.
29. Warshawsky, A. and Erlich-Rogozinski, S. (1977). *Microm. J.* **22**: 362–368.
30. Dorfne, K. (1991). Walter de Gruyter, Berlin.
31. Inoue, K. et al. (1993). *Bunseki Kagaku.* **42**: 725–731.
32. Rengan, K. (1997). *J. Radioanal. Nucl. Chem.* **219**: 211–215.
33. Hubicki, Z., Jakowicz, A., and Lodyga, A. Application of the ions from waters and sewages. In: *Adsorption and Environmental Protection, Studies in Surface Science and Catalysis*. A. Dabrowski (Ed.). Vol. 120. Elsevier, New York.
34. Kauspedience, D., Snukiskis, J., and Gefeniene, A. (1998). *J. Radioanal. Nucl. Chem.* **229**: 129–132.
35. Colella, C., De Gennaro, M., Langella, A., and Pansini, M. (1993). Cadmium removal from wastewater using chabazide and phillipsite. In: *Natural Zeolites 93*. D.W. Ming and F.A. Mupton (Eds.). Brockport, NY, p. 377.
36. Halle, K., Fischwasser, K., and Fenz, B. (1982). *Technol. Umweltschutz.* **25**: 1201032.
37. Leinonen, H. and Lehto, J. (2000). *Rect. Funct. Polym.* **43**: 1–6.
38. Lehto, J. et al. (1996). *J. Radio Nucl. Chem.* **208**: 435–443.
39. Untea, I., Tudorache, E., and Florea, C. (2000). *Revista de Chimie.* **51**: 123–126.
40. Drela, I. (1996). *Rudy Met. Niezelaz.* **41**: 321–323.
41. Randell, J.M., Bermann, R.L., Garrett, V., and Waiss, A.C. (1974). *J. Foorest Prod. J.* **24**(9): 80–84.
42. Teles de Vasconcelos, L.A. and Gonzalez Beca, C.G. (1994). *Eur. Water Poll. Control* **4**(1): 41–51.
43. Masri, M.S. and Friedman, M. (1974). *J. Appl. Polymer Sci.* **18**: 2367–2377.
44. Randall, J.M., Hautala, E., and Waiss, A.C., Jr. (1974). Removal and recycling of heavy metals ions from mining and industrial waste water streams with agricultural byproducts. *Proceedings of the Fourth Mineral Waste Utilization Symposium*, Chicago, IL, pp. 329–334.
45. Roberts, E.J. and Rowland, S.P. (1973). *Env. Sc. Technol.* **7**(6): 552–555.
46. Masri, M.S., Reuter, F.W., and Friedman, M. (1974). *J. Appl. Poly. Sc.* **18**: 675–681.
47. Prades, E.G. et al. (1994). *J. Chem. Tech. Biotechnol.* **59**: 289–295.
48. Cadena, F., Rizvi, R., and Peters, R.W. (1990). Feasibility studies for the removal of heavy metals from solution using tailored bentonite. In hazardous and industrial wastes. *Proceedings of the 22nd Mid-Atlantic Industrial Waste Conference*. Drexel University, Philadelphia, PA, pp. 77–94.
49. Matis, K.A. and Zoublious, A.I. (1994). *Biotech. Bioeng.* **44**: 354–360.
50. Niu, H., Xu, X.S., and Wang, J.H. (1993). *Biotech. Bioeng.* **42**: 785–787.
51. Sharma, D.C. and Forster, C.F. (1994). *Biores. Technol.* **49**: 31–40.
52. Srivastava, S.K., Singh, A.K., and Sharma, A. (1994). *Environ. Technol.* **15**: 353–361.
53. Low, K.S., Lee, C.K., Cheong, F.Y., and Lim, G.S. (1993). *J. Environ. Sci. Health A* **28**(8): 1689–1704.
54. Tong, P., Baba, Y., Adachi, Y., and Kawaza, K. (1991). *Chem Lett.* **9**: 1529–1532.
55. Kertman, S.V., Kertman, G.M., and Chibrikova, Z.H. (1993). *J. Appl. Chem. USSR* **66**(2): 465–466.
56. Dikshit, V.P. (1989). *Nat. Acad. Sci. Lett.* **12**(12): 419–421.
57. Holan, Z.R., Volesky, B.N., and Prasetyo, I. (1993). *Biotech. Bioeng.* **41**(8): 819–825.
58. Leush, A., Holan, Z.R., and Volesky, B. (1995). *J. Chem. Tech. Biotechnol.* **62**: 279–288.
59. Bricka, R.M. and Hill, D.O. (1989). Metal immobilization by solidification of hydroxide and xanthate sludges. In: *Environmental Aspects of Stabilisation and Solidifications of Hazardous and Radioactive Wastes*. P.L. Cote and T.M. Gilliam (Eds.). ASTM, Philadelphia, PA, pp. 257–272.
60. Flynn, C.M., Jr., Carnahan, T.G., and Lindastrom, R.E. (1980). Adsorption of heavy metal ions by xanthated sawdust. Report of Investigation #8427. United States. Bureau of Mines.
61. Kadirvelu, K. and Namasivayam, C. (2003). *Advances in Environmental Research*.
62. Apak, R., Tutem, E., Hugul, M., and Hizal, J. (1998). *Water Res.* **32**: 430–440.
63. Suemitsu, R., Venishi, R., Akashi, I., and Nakana, M. (1986). *J. Appl. Poly. Sci.* **31**: 75–83.
64. Periasamy, K. (1993). Adsorptive removal of heavy metals from synthetic indisustrial wastewater using peanut hull carbon. Ph.D. Thesis, Bharthiar University, Coimbatore, India.
65. Lopez-Delgado, C.P. and Girgis, B.S. (1998). *Water Res.* **32**: 1169–1177.

IRON

William R. Roy
Illinois State Geological Survey
Champaign, Illinois

CHEMISTRY OF IRON

The average concentration of iron in the earth's crust is about 5%, and it occurs in a wide variety of oxide, carbonate, silicate, and sulfide compounds either as metallic ores or in rock-forming minerals. Iron also occurs naturally in water from the dissolution of iron-containing minerals. Iron in water can occur in a dissolved form, a colloidal form that is combined with organic matter, and as suspended, iron-containing particles. Dissolved iron occurs in water in two oxidation states: ferrous iron, Fe (II), and ferric iron, Fe (III). In solution, iron can occur as a cation (Fe^{2+}) and in combination with other ions to form complexes and neutral ion pairs. Iron does not generally occur as the trivalent cation (Fe^{3+}) in significant concentrations in water. Trivalent iron generally exists in combination with other constituents. Depending on the occurrence of anions, such as sulfate, carbonate, and phosphate, and the pH and reduction-oxidation potential (E_h), the major iron species found in natural waters include a sulfate complex ($FeSO_4{}^0$),

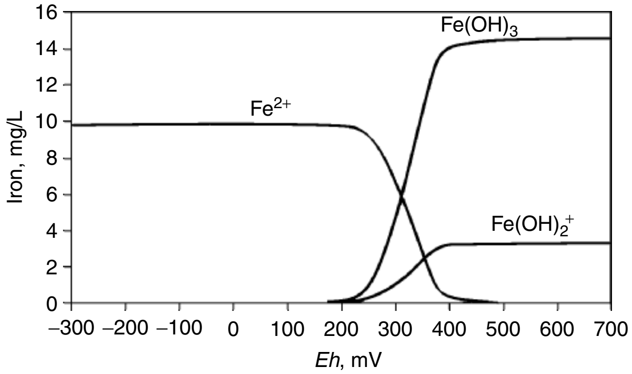

Figure 1. The distribution of 10 mg/L of iron in water at 15 °C and pH 7.0 as a function of oxidation-reduction potential (*Eh*).

a carbonate form [$Fe(HCO_3)_2$], phosphate complexes ($FeHPO_4{}^0, FeH_2PO_4{}^+$), and various hydroxide species such as $FeOH^+$, $Fe(OH)_2{}^+$, $Fe(OH)_3{}^0$, $Fe(OH)_4{}^-$, and $FeOH^{2+}$. The relative proportion of ferrous iron and ferric iron depends on the E_h of the water; ferrous iron is the dominant form in oxygen-poor water (reducing conditions). In oxygen-rich water (oxidizing conditions), ferric forms of iron dominate. For example, in a solution containing 10 mg/L of iron [Fig. 1], about 98% of the iron would occur as Fe^{2+} in a relatively reducing system. At an *Eh* greater (i.e., more oxidizing) than about 200 mV, ferrous iron is converted to the ferric forms $Fe(OH)_3{}^0$ and $Fe(OH)_2{}^+$. At an *Eh* of 700 mV, about 96% of the iron occurs in ferric forms. When a groundwater sample is collected, it may initially appear clear. However, if the sample is allowed to react with air and oxidize, the ferrous iron in the initially clear, reduced sample will be converted to ferric forms:

$$4Fe^{2+} + O_2 + 8OH^- + 2H_2O \rightleftharpoons 4Fe(OH)_3 \quad (1)$$

As a consequence, the sample may develop a reddish-brown color, and iron-containing particles may settle at the bottom of the container. The rate at which ferrous iron is oxidized depends on temperature, pH, and the amount of dissolved oxygen and iron.

Microorganisms can influence the oxidation state and distribution of iron in water. Iron can be used as an energy source via oxidation by microbial activity or reduced because iron can act as an electron acceptor. For example, *Crenothrix* and *Leptothrix* are diverse groups of iron-precipitating bacteria that are widely distributed in freshwater and soil. After ferrous iron is oxidized to the sparingly soluble ferric form, it may precipitate as ferric hydroxide. The bacterial cell sheaths can become coated with the precipitate, and the resulting gelatinous mass can clog water pipes and drains. There are bacteria and fungi that can reduce ferric iron. However, much less is known about iron-reducing microorganisms (1). The *Geobacter* are a group of bacteria that can reduce ferrous iron.

HEALTH AND REGULATORY ISSUES

Iron is not toxic in human consumption. Iron is an essential element for the formation of hemoglobin. The recommended dietary allowance for iron in healthy adults is 10 mg per day for men and 15 mg per day for premenopausal women (2). Dissolved iron in potable water is more of a nuisance than a potential heath issue. It can impart an unpleasant metallic taste to drinking water and alter the taste of food cooked in the water. The taste threshold for ferrous iron is about 0.1 mg/L. Concentrations exceeding 0.3 mg/L of iron can stain plumbing and clothing (3). The secondary maximum contaminant level for public water supplies is 0.3 mg/L, and 1.0 mg/L for freshwater aquatic life (4). However, bioassays by Roy et al. (5) suggested that concentrations less than 1.0 mg/L may be toxic to freshwater aquatic life.

LEVELS OF IRON IN WATER

The amount of dissolved iron in water depends on pH and *Eh*, but concentrations of less than 0.01 to 10 mg/L are typical for natural waters (6). Well-aerated surface water usually contains less than 0.05 mg iron/L. Iron concentrations in aquifers in Illinois have ranged from <0.05–21 mg/L (7). Iron in acid-mine drainage from coal mines has ranged from 10–5,625 mg/L (5,6).

SOLUBILITY OF IRON IN WATER

The amount of dissolved iron in water can be limited by the solubility of iron-containing solid phases. Schwab and Lindsay (8) concluded that when $pH + pE$ (pE is related to the redox potential *Eh* by $pE = (EhF)/(2.3RT)$ where $R =$ gas constant, $T =$ absolute temperature, and $F =$ Faraday constant) is less than 8.5, the concentration of ferrous iron in groundwater is limited by the solubility of siderite, $FeCO_3$:

$$Fe^{2+} + CO_3{}^{2-} \rightleftharpoons FeCO_3 \quad (2)$$

When $pE + pH$ is greater than 8.5, the solubility of ferrosic hydroxide, $Fe_3(OH)_8$, limits the amount of ferrous iron:

$$2Fe^{3+} + Fe^{2+} + 8H_2O \rightleftharpoons Fe_3(OH)_8 + 8H^+ \quad (3)$$

When the concentration of iron is greater than the solubility product of either siderite or ferrosic hydroxide, the iron precipitates, from solution. Ferrosic hydroxide is thermodynamically unstable, and other minerals may form, causing ferrosic hydroxide to dissolve. For example, amorphous ferric hydroxide, [$Fe(OH)_{3,a}$] and goethite (α-FeOOH) are less soluble than ferrosic hydroxide and given time, may control iron solubility. Moreover, amorphous ferric hydroxide can dehydrate (9) to form the more stable mineral hematite (Fe_2O_3):

$$2\,Fe(OH)_{3,a} \rightleftharpoons Fe_2O_3 + 3H_2O \quad (4)$$

The prevalent form of iron in surface water in the iron–carbon dioxide–water system can be illustrated by constructing a pH–p*E* stability diagram [Fig. 2]. This type of diagram illustrates which forms of iron are most thermodynamically stable at various values of pH

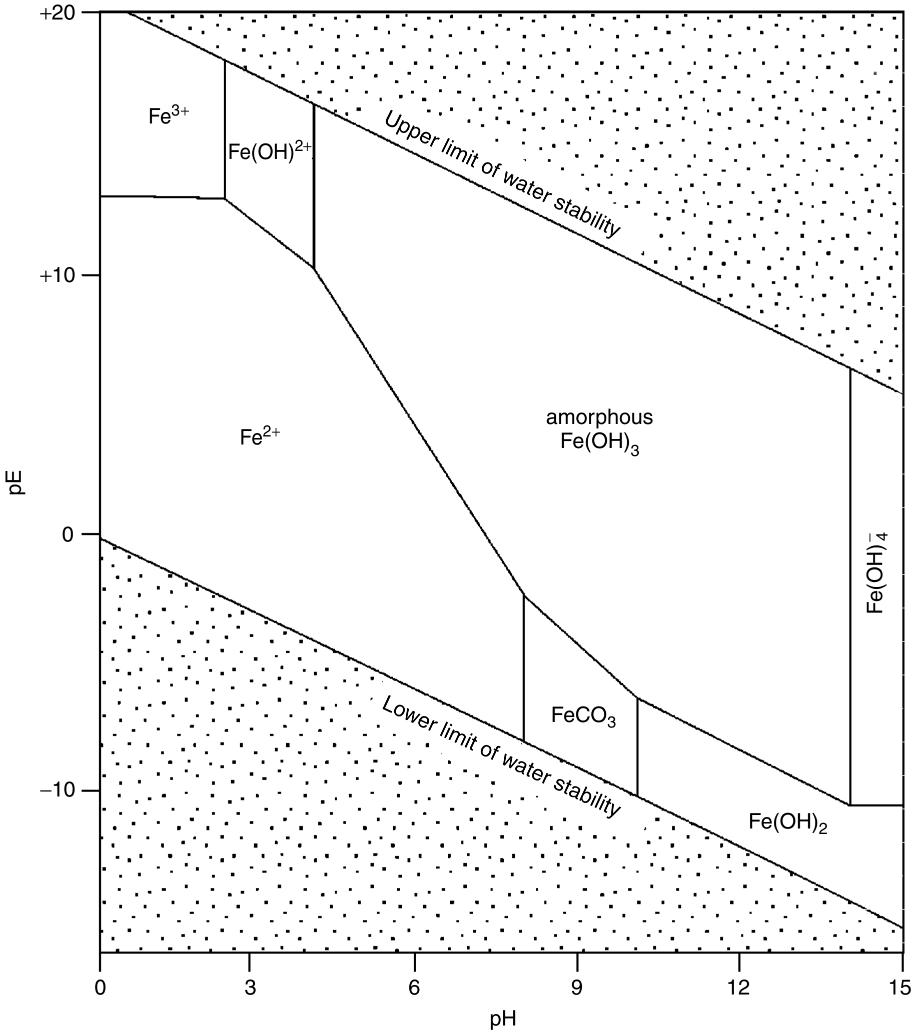

Figure 2. pE–pH diagram for the iron–carbon dioxide–water system at 25 °C and 1 atm (11).

and pE rather than at specific concentrations. Ferrous iron is the stable form in equilibrium with relatively acidic and reducing waters, whereas ferric iron, as both Fe^{3+} and $FeOH^{2+}$, is in equilibrium with acidic and oxidizing waters. In relatively less oxidizing and alkaline waters, the most thermodynamically stable forms may be amorphous iron hydroxide, siderite ($FeCO_3$), or solid ferrous hydroxide.

The concentration of iron in water may also be influenced by sorption–desorption interactions with mineral surfaces. Iron oxides and hydroxides have been studied as sorbents for other metals, but ionic iron is subject to the same electrostatic interactions that attract cations to clays and organic matter. Ferric iron can be sorbed by silica, clay minerals, and iron and manganese oxides (10). The sorption of ferrous iron has not been studied in detail.

METHODS FOR DETERMINING IRON

The amount of iron dissolved in water can be determined quantitatively by atomic absorption spectroscopy (AAS), atomic emission spectroscopy, and by colorimetric methods (12). In the AAS method, a water sample is aspirated into an air-acetylene flame and atomized. A monochromatic beam of light at a wavelength corresponding to iron is passed through the flame. The amount of light energy sorbed is proportional to the amount of iron present. When analyzed at a wavelength of 248.3 nm, the optimal concentration range for iron is about 0.3 to 10 mg/L.

Dissolved iron can be determined by argon-supported inductively coupled plasma (ICP) spectroscopy. In ICP, the liquid sample is aspirated into a high-temperature (6000–8000 K) flame, and atomic emission is measured to determine the concentration of the iron present. The concentration of iron is proportional to the intensity of the emissions. Iron has many such emission lines in its spectrum, but the wavelength of 259.94 nm is commonly used for concentrations ranging from about 0.007 to 100 mg/L.

Dissolved iron also can be determined by the phenanthroline method. In this method, iron in solution is reduced to the ferrous form, and then allowed to react with 1, 10-phenanthroline to form a reddish-orange product. The absorbance of the complex is measured in the range of 460–520 nm. Both AAS and ICP methods determine total iron, but the amount of ferric and ferrous iron can be determined separately by the phenanthroline method (13).

METHODS FOR REMOVING IRON FROM WATER

The presence of large concentrations of dissolved iron is a common water-quality problem, especially for groundwater (1). Methods for removing iron are summarized in Lehr et al. (3) and Vigneswaran and Visvanathan (2). The oxidation of soluble, ferrous iron into relatively insoluble ferric iron, followed by filtration of the suspended mass is common to many of the methods. Depending on its concentration in the untreated water, the oxidation of ferric iron can be promoted by aeration, catalytic oxidation with manganese oxides, and chlorination. Household water softeners are designed to remove calcium and magnesium from hard water using an ion exchange resin, but they can also remove dissolved iron. Lime softening of municipal water supplies can also promote the precipitation of ferric iron. The presence of iron-fixing bacteria may necessitate the use of relatively large amounts of bleach (“shot chlorination”) to reduce the amount of iron in the final product.

BIBLIOGRAPHY

1. Chapelle, F.H. (2001). *Ground-Water Microbiology and Geochemistry*. John Wiley & Sons, New York.
2. Vigneswaran, S. and Visvanathan, C. (1995). *Water Treatment Processes. Simple Options*. CRC Press, Boca Raton, FL.
3. Lehr, J.H., Gass, T.E., Pettyjohn, W.A., and DeMarre, J. (1980). *Domestic Water Treatment*. McGraw-Hill, New York, pp. 172–178.
4. U.S. EPA. (1999). *National Recommended Water Quality Criteria–Correction*. EPA 822-Z-99-001.
5. Roy, W.R., Skelly, T.M., Krapac, I.G., Griffin, R.A., Dickerson, D.R., Schuller, R.M., Suloway, J.J., and Shimp, N.F. (1984). *Geochemical Properties and Toxicological Effects on Aquatic Life of Coal Wastes*. Environmental Geology 107, Illinois State Geological Survey, Champaign, IL, p. 124.
6. Patterson, J.W. (1975). *Wastewater Treatment Technology*. Ann Arbor Science, Ann Arbor, MI.
7. Warner, K.L. (2000). *Analysis of Nutrients, Selected Inorganic Constituents, and Trace Elements in Water from Illinois Community-Supply Wells, 1984–91*. Water-Resources Investigations Report 99-4152, U.S. Geological Survey.
8. Schwab, A.P. and Lindsay, W.L. (1983). Effect of redox and pH on the solubility and availability of iron. *Soil Sci. Soc. Am. J.* **47**:201–205.
9. Garrels, R.M. and Christ, C.L. (1965). *Solutions, Minerals, and Equilibria*. Freeman Cooper, San Francisco, CA.
10. Anderson, M.A. and Rubin, A.J. (1981). *Adsorption of Inorganics at Solid–Liquid Interfaces*. Ann Arbor Science, Ann Arbor, MI.
11. Stumm, W. and Morgan, J.J. (1981). *Aquatic Chemistry*. John Wiley & Sons, New York.
12. American Public Health Association. (1992). *Standard Methods for the Examination of Water and Wastewater*, 18th Edn. Washington, DC, Section 3500. Iron, 3-65–3-68.
13. Stucki, J.W. (1981). The quantitative assay of minerals for Fe^{2+} and Fe^{3+} using 1, 10-phenanthroline: II. A photochemical method. *Soil Sci. Soc. Am. J.* **45**: 638–641.

ISOTOPES

Aldo Conti
Frascati (RM), Italy

Isotopes are atoms with the same atomic number, **Z**, but different atomic weight, **A**. Different isotopes, therefore, have the same chemical but different physical properties, such as melting and boiling point. The name isotope means “at the same place” and comes from the periodic table, where different isotopes of the same element are at the same place. The scientific way of labeling isotopes reflects this. They have the symbol of the corresponding element, but followed by a hyphen and the number of nucleons (neutrons and protons) of the isotope. So, C-12 is the normal carbon (six protons and six neutrons) and C-14 is a heavy unstable isotope (six protons and eight neutrons).

Isotopes of the same element share the same chemical properties because they depend only on the number of protons, Z. But different isotopes can have very different weights, which is particularly evident in hydrogen. Normal hydrogen has only one proton and no neutrons. Deuterium, or H-2, has one neutron and its weight is nearly twice that of hydrogen. Then, tritium, H-3, has two neutrons and is three times as heavy, which is the reason why, when deuterium combines with oxygen, the resulting molecule is called heavy water. Due to the extra weight of deuterium, heavy water boils at 101.42 °C (214.56 °F) and freezes at 3.81 °C (38.86 °F), at standard pressure. The difference in weight affects the speed of chemical reactions involving water. For this reason, heavy water is not really safe to drink. Experiments performed with mice have shown that large amounts of heavy water in the body can inhibit mitosis, which affects rapidly dividing tissues, such as the stomach lining.

Isotopes do not differ only in physical properties. The extra neutrons make some of the nuclei highly unstable, which is the reason that some isotopes are radioactive. A typical example is uranium. Its isotopes, such as U-235, decay radioactively following a long chain, which leads to lead. Some of this decay involves a neutron becoming a proton by emitting an electron. In this case, the nucleus becomes a new element, with the atomic number increased by one. The energy emitted during the decay can be harvested in nuclear plants or used destructively in atomic bombs.

The instability of their nuclei makes isotopes useful for many applications. A famous one is carbon dating. Carbon comes mainly in two isotopes, the normal C-12 and the heavy, unstable, C-14. When an organism grows, it uses them in the same proportion in which they are found in the environment. After the death of the organism, the ratio changes continuously, as the heavy C-14 decays. C-14 has a half life of 5730 years, which means that after this period half of it has gone. By this method, it is possible to date any biological remains or artifact up to about 50,000 years old. For older objects, the amount of C-14 becomes too small to be measured reliably. Isotopes are often used as labels. By replacing a stable atom in a molecule with an unstable radioactive isotope, it becomes possible to follow the chemical reactions the molecule undergoes. Our ability to detect radioactive isotopes makes them very useful in the fields of medicine and biology, where they are used as tracers.

ISOTOPE FRACTIONATION

Gholam A. Kazemi
Shahrood University of
Technology
Shahrood, Iran

INTRODUCTION

The following paragraph extracted from the major textbook by Bowen (1) describes fractionation as it is dealt with in general chemistry. The subsequent paragraph defines fractionation in general hydrochemistry. Then, isotope fractionation is discussed.

In general chemistry, fractionation is the separation of a mixture, usually of chemically related components, into fractions that possess different properties. A mixture of liquids having different boiling points, may be separated by collecting fractions that boil at different temperatures in a fractionating column. This is a long, vertical column containing plates or bubble caps attached to a still. As a consequence of internal reflux, a gradual separation occurs between high and low boiling point fractions of the liquid mixture.

In hydrochemistry, the fractionation factor is a ratio which shows the depletion or addition of ions in rainwater compared to seawater (2). In other words, the fractionation factor is the ratio of two ions in rainwater to the ratio of the same ions in seawater. For example, for Cl^- relative to Na^+, the fractionation factor F_{Na} is (2):

$$F_{\mathrm{Na}} = \frac{(\mathrm{Cl/Na})_{\mathrm{rainwater}}}{(\mathrm{Cl/Na})_{\mathrm{seawater}}}$$

In this equation, the units of concentration can be mg/l, mmol/l, meq/l, etc.

Example: What is the fractionation factor of Na for a rainwater sample whose chloride and sodium concentrations are 1 and 0.5 mg/l, respectively?

$$\mathrm{Cl/Na}_{\mathrm{rainwater\ sample}} = 1/0.5 = 2$$

$$\mathrm{Cl/Na}_{\mathrm{seawater}} = 1.81$$

$$F_{\mathrm{Na}} = \frac{(2)}{(1.81)} = 1.11$$

A fractionation factor greater than unity indicates that the rainwater sample is depleted (in comparison to seawater) in the ion in question. If it is less than unity, it shows the enrichment (in comparison to seawater) of the sample in the ion. Appelo and Postma (2) also propose the fractionation factor as a quantitative approach to find the activity of a single species that forms part of a given total concentration, for example, for $\mathrm{HCO_3}$, $m_{\mathrm{HCO_3}^-} = \alpha \times \Sigma\mathrm{CO_2}$, where α is the fractionation factor. Some researchers use “enrichment factor (Ef)” in an equation slightly different from the fractionation factor to deduce source/s of solutes in atmospheric precipitation (3,4). If the Ef is greater than unity, the ion in question (in the sample) is derived partly or wholly from a source other than oceanic origin. If it is less than unity, the ion could have been derived wholly from oceanic sources:

$$\mathrm{Ef} = \| \frac{(\mathrm{Ion/Cl})_{\mathrm{sample}}}{(\mathrm{Ion/Cl})_{\mathrm{seawater}}}$$

In this example, the enrichment factor equals 0.91.

$$\mathrm{Ef} = \frac{(0.5)}{(0.55)} = 0.91$$

DEFINITION

Isotope fractionation is defined by Domenico and Shwartz (5) as a change in an isotopic ratio (for example, $^{18}\mathrm{O}/^{16}\mathrm{O}$) because of a chemical reaction (for example, water changes to water vapor). It is also defined by Coplen (6) as the fractionation or partitioning of isotopes by physical or chemical processes which is proportional to the differences in their masses. Coplen also classifies isotopic fractionation processes into two categories, chemical and physical isotopic fractionation processes: physical processes such as ultrafiltration or gaseous diffusion of ions or molecules. Chemical processes are those that lead to redistribution of the isotopes of an element among phases or chemical species and are divided into two broad categories equilibrium isotopic reactions and kinetic isotopic reactions (6).

Gilreath (7) in a simple yet valuable text describes the fundamentals of isotope fractionation in detail and refers to it as “isotope separation.” Anyone intending to deal with isotope fractionation is advised to consult this text first. The point of primary importance is that fractionation of isotopes is important in the geologic environment *only* for elements of low atomic number. Bowen (1) describes three main processes which cause fractionation: (1) isotope exchange reactions; (2) unidirectional reactions; and (3) physical processes such as evaporation, condensation, crystallization, melting, adsorption, desorption, and diffusion. Isotopes may also be fractionated in the laboratory by a variety of means (URL Ref. http://www.c14dating.com/frac.html). One method that may result in considerable fractionation during isotopic analysis is the recently introduced technique of secondary-ion mass spectrometry (SIMS) (8).

The following are the factors that control the rate of isotope fractionation:

1. Fractionation is minimized when reactions are fast (9).
2. The higher the mass difference between various isotopes of the element, the more effective the fractionation process. Deuterium, for example, has the largest mass difference from common hydrogen compared to isotopes of other elements. Hence, it is subject to greater fractionation.

3. Isotope fractionation decreases as temperature increases—at finite temperatures, all isotopic species are well mixed (6).
4. Isotopic exchange between minerals and water in deep, basinal flow systems or in geothermal systems induced by high temperature helps the fractionation process to proceed (5).
5. In evaporation from a pond, fractionation is more efficient if the vapor produced is constantly removed, for example, by wind (10).

Fractionation Factor

The fractionation factor is the ratio of isotopes (for example, $^{18}O/^{16}O$) in a liquid phase to the same ratio in the vapor phase;

$$\alpha = \frac{(^{18}\mathrm{O}/^{16}\mathrm{O})_{\mathrm{liquid}}}{(^{18}\mathrm{O}/^{16}\mathrm{O})_{\mathrm{vapor}}} \tag{11}$$

or more clearly/generally defined by Haldorsen et al. (11) "as the increase or decrease of a given isotope in a substance that undergoes a phase change or simply in terms of delta notation:

$$\alpha = \frac{1000 + \delta^{18}\mathrm{O}_{\mathrm{liquid}}}{1000 + \delta^{18}\mathrm{O}_{\mathrm{vapor}}}$$

Example: In the condensation of water vapor to liquid, typical values of $\delta^{18}O$ for the liquid and vapor are, -5% and -14%, respectively. Therefore, the fractionation factor is,

$$\alpha = \frac{1000 + (-5)}{1000 + (-14)} = \frac{995}{986} = 1.0092$$

Each fractionating process results in a specific fractionation factor. For example, the oxygen fractionation factor in freezing differs from that from condensation. The former and the latter are 1.0111 and 1.003, respectively. This means that freezing is more effective than condensation for isotope fractionation; freezing leads to stronger fractionation than condensation.

Fractionation factors approach unity as temperature increases; when $\alpha = 1$, the isotope ratios become equal in both phases being considered, and hence no further isotopic fractionation can occur (1). Higher fractionation factors mean more effective fractionation processes.

Isotopic Fractionation Curve

A curve defines the equilibrium isotope fractionation between coexisting phases as a function of temperature (1). It usually takes the form of an exponentially decreasing curve; there is a negative correlation between isotope fractionation and temperature. The shape of this curve is different for different isotopes.

Fractionation of Various Environmental Isotopes

As can be inferred from the information provided so far, isotope fractionation is an extensive topic in chemistry/geochemistry and has been dealt with in detail in many references such as Bowen (1), Gat (12), Gilreath (7), and Ottonello (13); the last is a thorough mathematical treatment of the topic. Here, in the *Encyclopedia of Water*, the discussion of isotope fractionation is limited to the isotopes of hydrogen and oxygen, specifically, the atoms that form the water molecule, as well as other isotopes that are of frequent use in water related/hydrologic studies. The latter group includes isotopes of the elements C, N, Cl, S, Sr, and less used U, Th, and Si. These are, often, referred to as environmental isotopes (6). Some workers such as Haldorsen et al. (11) classify the environmental isotopes into two groups: hydrologic isotopes (O, D, T) and biochemical isotopes (C, N, S). Although this classification covers only a limited number of environmental isotopes, it is sound for fractionation. For the first group, the main fractionating processes are physical or chemical reactions, biological/bacteriological/microbial processes are dominant in the fractionation of the second group.

Before discussing all environmental isotopes one by one, it should be pointed out that in some cases, considerable isotope fractionation has led to the usefulness of those isotopes in hydrologic studies, for example, H and O isotopes (enrichment and depletion processes are the fundamentals for studying lake–groundwater interaction, recharge temperature, and many other applications). In contrast, in some other cases, lack of isotope fractionation has been helpful. Strontium isotopes, for example, are used to study the origin of solutes in the streams because they are not subjected to substantial fractionation.

H and O

Hydrogen and oxygen atoms have three isotopes each (see also Heavy Water and Deuterium articles): ^{1}H, ^{2}H, ^{3}H, and ^{16}O, ^{17}O, and ^{18}O. Water molecules containing the heavy isotopes are less likely to evaporate and more likely to condense than those containing lighter isotopes. Hence, when a body of water evaporates, the vapor phase is depleted and the residue, the liquid phase, is enriched in heavy isotopes, ^{2}H, ^{3}H, ^{17}O, and ^{18}O. During melting/freezing of water, heavy isotopes are slightly more inclined to freeze and less inclined to melt than lighter isotopes, making the melt water depleted in ^{18}O and deuterium compared to the snowpack. Ultrafiltration (passage of water through shale micropores) enriches the residue (not passed through water) in O and H isotopes (6), but transpiration is not a fractionating process. The principles of hydrogen isotope fractionation are the scientific foundation for producing heavy water, which is used in the nuclear industry.

The fractionation factor for deuterium is greater than that for oxygen-18 because of a larger difference between deuterium and hydrogen masses compared to the difference between ^{16}O and ^{18}O masses. The fractionation factors of these two elements at various temperatures are calculated via the following equations (1):
For oxygen,

$$1000 \ln\,(\alpha\ ^{18}\mathrm{O}) = 2.644 - 3.026 \times 10^{3}/T + 1.534 \times 10^{6}/T^{2}$$

For hydrogen,

$$\ln\,(\alpha\ D) = 0.0771 + 13.436 \times 10^{3}/T^{2}$$

(note that the temperature is in K).

Example:

The fractionation factor for ^{18}O at 293.15 K (20 °C) is

$$\ln(\alpha\ ^{18}\mathrm{O}) = (2.644 - 3.026 \times 10^3/293.15 + 1.534 \times 10^6/293.15^2)/1000 = 1.01000$$

The fractionation factor for ^{18}O during calcite precipitation from water is 1.0349, which means that calcite is 35% more enriched in ^{18}O relative to the water from which it precipitates. However, not all precipitated calcite is richer in ^{18}O compared to "parent water." For instance, if calcite precipitation is due to freezing, the ^{18}O content of calcite is lower (11).

C

Oxidation–reduction processes in which the heavier isotope is concentrated in the more oxidized forms are important in carbon fractionation. The most influential factor in carbon isotope fractionation is selective uptake of lighter CO_2 by plants. During photosynthesis, plants discriminate against the heavier isotopes of carbon; the isotope ^{13}C in organic matter (soil) is depleted by 1.8% compared to its natural ratios in the atmosphere. It is believed that all organisms discriminate against ^{14}C about twice as much as against ^{13}C and the ratio between the stable ^{12}C and ^{13}C atoms can be used to correct for the initial depletion of ^{14}C. The amount of isotope fractionation depends on the photosynthetic pathway used by the plant. Plants are divided into three categories in this regard (11). At high temperatures, equilibrium fractionation of isotopes may occur between CH_4 and CO_2, and this is sensitive to changes in pH, oxygen fugacity, and ionic strength as well as to changes in temperature (13). As an example, the fractionation factor for ^{13}C in $CaCO_3$ in equilibrium with CO_2 is 1.01074.

Cl

The author has not come across a remark in the literature about the occurrence of fractionation in chlorine isotopes. This may have happened because chloride is a highly mobile, soluble, and geochemically conservative (is not removed from solution by mineral interaction or secondary mineral formation) element and has a relatively high mass.

N

Fractionation of nitrogen isotopes occurs mostly through biological processes and some isotope exchange reactions. Kinetic isotope fractionation often dominates nitrogen reactions and because of this, during denitrification, δ^{15}N values of residual nitrate increase exponentially as concentrations decrease (6). Bowen (1) reports that conversion of ammonia to nitrate (nitrification) by nitrifying organisms produces isotope fractionation up to 20% and even greater values (up to 30%) are associated with denitrification.

S

Sulfur in near-surface environments frequently changes between the oxidized and the reduced forms, so it acquires a wide variety of ^{34}S/^{32}S ratios. Isotope fractionation occurs in many processes of the sulfur cycle. The principal reaction leading to the fractionation of sulfur isotopes in nature is the reduction of sulfate ion by anaerobic bacteria (13). Bacterial reactions are important because they are kinetic, fast, favor the lighter isotope, and result in depleting ^{34}S in natural biogenic H_2S and hence enrich the remaining sulfate (11). Brucher et al. (14) studied isotope fractionation during bacterial sulfate reduction and found that in addition to the variations in the rates of bacterial sulfate reduction, genetic and physiological differences between various genera of sulfate-reduction bacteria strongly influence isotope fractionation.

In the precipitation of sulfide minerals from solution, the fractionation and enrichment of ^{34}S depend on the relative bond strengths and the temperature of solution (13). Sulfide minerals in newly deposited sediment can be enriched in ^{32}S by about 50% compared with associated marine sulfate (1). The equilibrium isotope fractionation factor between SO_4^{2-} and H_2S is 1.074 and is 1.020 between water and dissolved sulfate at 88 °C (6). During dissolution of evaporites, the largest reservoir of sulfate, no sulfur fractionation takes place.

Sr

Because of its relatively high mass, no detectable isotope fractionation accompanies precipitation of Sr within minerals (6).

ARTIFICIAL ISOTOPE FRACTIONATION FOR "SEPARATION OF ISOTOPES"

The most interesting and applied part of the isotope fractionation topic is artificial separation ("fractionation") of isotopes, a subject that is scientifically simple, but, in engineering terms very complicated. Historically, the first successful attempt to separate isotopes was that by Aston (the inventor of the mass spectrograph and the Nobel prize winner in 1922) who separated neon into two fractions of different atomic weights by repeated diffusion. The topic gained importance with the discovery, in 1939, that ^{235}U is a fissionable material and in succeeding years that uranium-235 is the key material for producing nuclear weapons and power (7).

VARIOUS TECHNIQUES FOR SEPARATING ISOTOPES ARTIFICIALLY

As stated before, separating isotopes is an intensive complicated engineering task, and only a few countries have the capacity. Four principles that constitute the foundation of isotope separation should be considered when devising a separation process (7):

1. At a given temperature, gaseous molecules, regardless of mass, have the same average kinetic energy.

2. The gravitational or centrifugal forces acting on a particle of matter are directly proportional to its mass.
3. The electric or magnetic forces affecting ionized atoms or molecules follow Coulomb's laws for these forces.
4. Among the lighter elements, isotopic species may vary slightly in their chemical reactivities.

Considering these four principles, a variety of techniques have been developed for isotopic enrichment (7). Each of these techniques has advantages and disadvantages and is suitable for specific types of conditions and isotope(s).

1. The gaseous-diffusion method is used for separating and enriching uranium isotopes. The production plant which uses this technique is huge; the huge diffusion plant at Oak Ridge, USA, in operation from 1945, represents one of the greatest technical achievements in the history of humankind.
2. The electromagnetic method is suitable only where cheap electricity is available.
3. The thermal diffusion method is good for laboratory-scale production.
4. The centrifuge method was not developed into a large-scale production plant.
5. The electrolytic method is used only for separating hydrogen and deuterium.
6. The chemical exchange method is used mostly for separating C and N.
7. The distillation method is only for enriching deuterium.

A sideline and an interesting coincidence is that while this article was being written, the Board of Governors of International Atomic Energy Agency (IAEA) postponed its decision on Iran's (the home country of the author) uranium enrichment and nuclear program activities (CNN Online edition, Nov. 22, 2003). The uranium enrichment and nuclear activities program of Iran has been in the news headlines for quite a long time.

BIBLIOGRAPHY

1. Bowen, R. (1988). *Isotopes in the Earth Sciences*. Elsevier, London.
2. Appelo, C.A.J. and Postma, D. (1999). *Geochemistry, Groundwater and Pollution*. Balkema, Rotterdam.
3. Kazemi, G.A. and Mehdizadeh, H. (2001). Solute content and chemical composition of atmospheric precipitation of Shahrood Region, *Proc. of 5th Iranian Geol. Convention*, Tehran, September 2001, pp. 401–405 [in Persian].
4. Panettiere, P., Cortecci, G., Dinelli, E, Bencini, A., and Guidi, M. (2000). Chemistry and sulfur isotopic composition of precipitation at Bologna, Italy. *Appl. Geochem.* **15**: 1455–1467.
5. Domenico, A.P. and Shwartz, F.W. (1991). *Physical and Chemical Hydrogeology*. John Wiley & Sons, New York.
6. Coplen, T.B. (1993). Uses of environmental isotopes New York. In: *Regional Ground-Water Quality*. W.M. Alley (Ed.). ITP, pp. 255–293.
7. Gilreath, E.S. (1985). *Fundamental Concepts of Inorganic Chemistry*. 19th printing, International Student Edition. McGraw-Hill, Singapore.
8. MacRae, N.D. (1995). Secondary-ion mass spectrometry and geology. *Can. Mineralogist* **33**: 219–236.
9. Langmuir, D. (1997). *Aqueous Environmental Geochemistry*. Prentice-Hall, Upper Saddle River, NJ.
10. Mazor, E. (1991). *Applied Chemical and Isotopic Groundwater Hydrology*. Open University Press, Milton Keynes.
11. Haldorsen, S., Riise, G., Swanson, B., and Sletten, R.S. (1997) Environmental isotopes as tracers in catchments. In: *Geochemical Processes, Weathering and Groundwater Recharge in Catchments*. O.M. Saether and P.D. Caritat (Eds.). Balkema, Rotterdam, pp. 185–210.
12. Gat, J.R. (1981). Isotopic fractionation. In: *Stable Isotope Hydrology, Deuterium and Oxygen-18 in the Water Cycle*. J.R. Gat and R. Gonfiantini (Eds.). IAEA, technical report series, No. 210, pp. 21–34.
13. Krauskopf, K.B. (1979). *Introduction to Geochemistry*, 2nd Edn. McGraw-Hill, New York.
14. Bruchert, V., Knoblauch, C., and Jorgensen, B.B. (1999). Sulfur isotope fractionation during bacterial sulfate reduction. In: *Geochemistry of Earth's Surface*. H. Armannsson (Ed.). Balkema, Rotterdam, pp. 303–305.

MARIOTTE BOTTLE—USE IN HYDROLOGY

JOSEPH HOLDEN
University of Leeds
Leeds, United Kingdom

A Mariotte bottle is a device that allows constant flow of a fluid from a container, even when the fluid level in the container changes. It is named after Edme Mariotte (1620–1684), a French physicist, who, independently from Boyle, discovered the law that states that the volume of a gas varies inversely with its pressure. Mariotte bottles have a wide range of applications in engineering and science (such as fuel supply) and are used in a variety of hydrologic applications ranging from river discharge gauging to infiltration measurement.

Figure 1 shows the general design of a Mariotte bottle. It consists of a bottle that has an escape pipe or tap to allow fluid to exit and a vent tube to allow air to enter the bottle. The bottle is sealed so that the only location where air can enter is at that bottom of the vent tube, b. The fluid (such as water) exits through a tap (or any sort of escape tube) at c, at a rate that is controlled by the size of the escape tube and the head of water. The head of water, h will always be equal to the distance between the outlet tap and the bottom of the vent tube. Therefore, regardless of the water level inside the bottle, there will be a constant head of water.

If the vent tube is kept at a constant level, then as a result of the reduction of the fluid in the bottle, the air above it expands and its pressure drops. When the pressure of the air has dropped so that it plus the pressure of the water column from c to b (the vent tube opening) is less than the atmospheric pressure of the air outside the Mariotte bottle (and therefore also inside the vent tube), then air enters the Mariotte bottle at b and rises through the fluid. Therefore, the pressure at c will always be equal

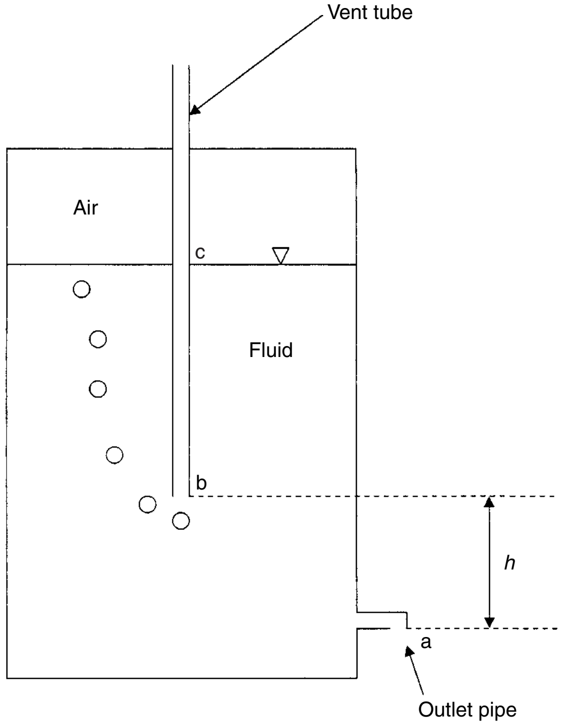

Figure 1. A Mariotte bottle.

to atmospheric which is also equal to the pressure at b. The excess pressure, under which the fluid is ejected, is the pressure of the fluid between the levels b and a. It remains constant until the fluid drops to the level of b, and the discharge from the escape pipe will remain constant until then. Thus, Mariotte bottles can be used wherever there is an application for a constant flow of water from a supply reservoir.

If the vent tube is raised to a greater height, the head will increase, and therefore, the rate of water flowing out of the bottle will increase. It is therefore possible to control the rate of water release from the bottle by controlling the head of water, and this is, in turn, controlled by the height of the base of the vent tube above the outlet at a. If the vent tube is moved down to the level of the escape opening, a, the flow stops because then the height of the column becomes zero. The discharge also ceases if the vent tube is sealed preventing air from entering the bottle.

There is a wide range of common hydrologic applications for Mariotte bottles. They are used frequently for adding a constant rate of solution (such as a dye or tracer) to watercourses. Most commonly this is for discharge gauging where, if a constant rate of solution (such as sodium chloride solution) is added to river, then the discharge of the river can be calculated based on the equation,

$$Q = q(C_i - C_d)/(C_d - C_b)$$

where Q is the river discharge, q is the rate of injection of tracer solution, C_i is the concentration of the tracer, C_b is the background concentration in the river, and C_d is the downstream sampled concentration of the tracer.

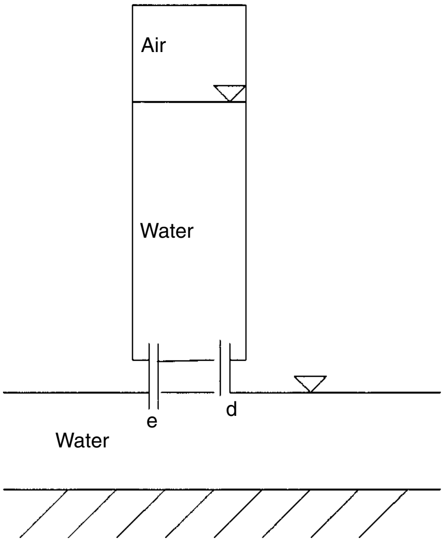

Figure 2. Mariotte bottle used to keep a constant water level above the ground surface.

It is often useful to use a Mariotte bottle to supply a constant rate of water and also, if the bottle is graduated, then it can be used to determine how much water has been supplied during any given period. Frequently, Mariotte bottles are used in instruments designed to measure soil infiltration rates such as rain simulators or other types of infiltrometers where a constant rate of water supply to the soil surface might be required.

In constant head infiltrometers (see entry for INFILTROMETERS), the Mariotte principle is used to ensure that there is a constant depth of ponded water above the soil surface. This is done by placing the vent tube for the Mariotte bottle at the same end of the bottle as the water exit pipe as in Fig. 2. Whenever the water level on the ground drops below the height of the base of the vent tube, d, air is allowed to enter the water supply bottle, and water can be released from e until the water level on the ground has risen back to the base of the vent tube. In this way, a constant height or head of water can be maintained on a soil surface. This principle can also be used for irrigation to maintain a constant surface water level or a constant groundwater level.

MARS EXPLORATION ROVER MISSION

GUY WEBSTER
Jet Propulsion Laboratory
Pasadena, California

DONALD SAVAGE
NASA Headquarters
Washington

OPPORTUNITY ROVER FINDS STRONG EVIDENCE MERIDIANI PLANUM WAS WET

Scientists have concluded the part of Mars that NASA's Opportunity rover is exploring was soaking wet in the past.

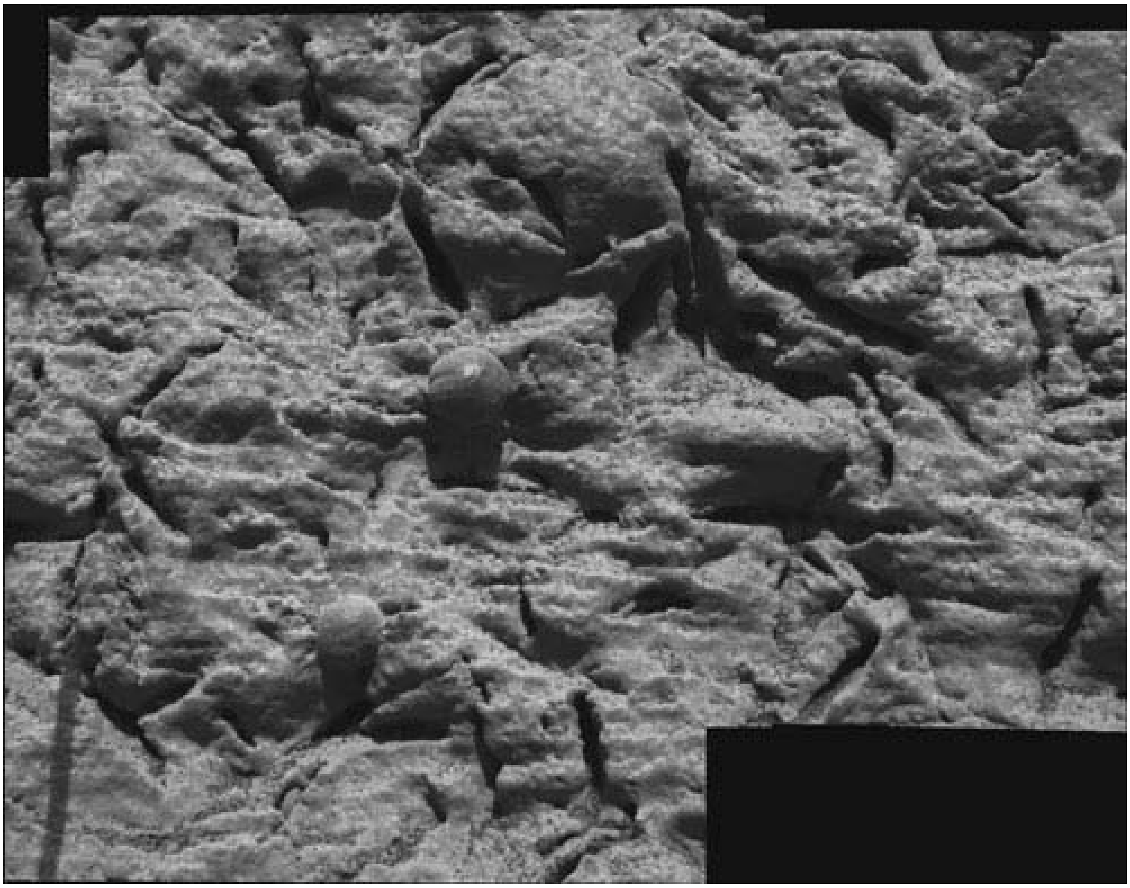

This image, taken by Opportunity's microscopic imager, shows a portion of the rock outcrop at Meridiani Planum, Mars, dubbed "Guadalupe." View Opportunity images from this press release.

Evidence the rover found in a rock outcrop led scientists to the conclusion. Clues from the rocks' composition, such as the presence of sulfates, and the rocks' physical appearance, such as niches where crystals grew, helped make the case for a watery history.

"Liquid water once flowed through these rocks. It changed their texture, and it changed their chemistry," said Dr. Steve Squyres of Cornell University, Ithaca, N.Y., principal investigator for the science instruments on Opportunity and its twin, Spirit. "We've been able to read the tell-tale clues the water left behind, giving us confidence in that conclusion."

Dr. James Garvin, lead scientist for Mars and lunar exploration at NASA Headquarters, Washington, said, "NASA launched the Mars Exploration Rover mission specifically to check whether at least one part of Mars ever had a persistently wet environment that could possibly have been hospitable to life. Today we have strong evidence for an exciting answer: Yes."

Opportunity has more work ahead. It will try to determine whether, besides being exposed to water after they formed, the rocks may have originally been laid down by minerals precipitating out of solution at the bottom of a salty lake or sea.

The first views Opportunity sent of its landing site in Mars' Meridiani Planum region five weeks ago delighted researchers at NASA's Jet Propulsion Laboratory, Pasadena, Calif., because of the good fortune to have the spacecraft arrive next to an exposed slice of bedrock on the inner slope of a small crater.

The robotic field geologist has spent most of the past three weeks surveying the whole outcrop, and then turning back for close-up inspection of selected portions. The rover found a very high concentration of sulfur in the outcrop with its alpha particle X-ray spectrometer, which identifies chemical elements in a sample.

"The chemical form of this sulfur appears to be in magnesium, iron or other sulfate salts," said Dr. Benton Clark of Lockheed Martin Space Systems, Denver. "Elements that can form chloride or even bromide salts have also been detected."

At the same location, the rover's Mössbauer spectrometer, which identifies iron-bearing minerals, detected a hydrated iron sulfate mineral called jarosite. Germany provided both the alpha particle X-ray spectrometer and the Mössbauer spectrometer. Opportunity's miniature thermal emission spectrometer has also provided evidence for sulfates.

On Earth, rocks with as much salt as this Mars rock either have formed in water or, after formation, have been highly altered by long exposures to water. Jarosite may point to the rock's wet history having been in an acidic lake or an acidic hot springs environment.

The water evidence from the rocks' physical appearance comes in at least three categories, said Dr. John Grotzinger, sedimentary geologist from the Massachusetts Institute of Technology, Cambridge: indentations called "vugs," spherules and crossbedding.

Pictures from the rover's panoramic camera and microscopic imager reveal the target rock, dubbed "El Capitan," is thoroughly pocked with indentations about a centimeter (0.4 inch) long and one-fourth or less that wide, with apparently random orientations. This distinctive texture is familiar to geologists as the sites where crystals of salt minerals form within rocks that sit in briny water. When the crystals later disappear, either by erosion or by dissolving in less-salty water, the voids left behind are called vugs, and in this case they conform to the geometry of possible former evaporite minerals.

Round particles the size of BBs are embedded in the outcrop. From shape alone, these spherules might be formed from volcanic eruptions, from lofting of molten droplets by a meteor impact, or from accumulation of minerals coming out of solution inside a porous, water-soaked rock. Opportunity's observations that the spherules are not concentrated at particular layers in the outcrop weigh against a volcanic or impact origin, but do not completely rule out those origins.

Layers in the rock that lie at an angle to the main layers, a pattern called crossbedding, can result from the action of wind or water. Preliminary views by Opportunity hint the crossbedding bears hallmarks of water action, such as the small scale of the crossbedding and possible concave patterns formed by sinuous crestlines of underwater ridges.

The images obtained to date are not adequate for a definitive answer. So scientists plan to maneuver Opportunity closer to the features for a better look. "We have tantalizing clues, and we're planning to evaluate this possibility in the near future," Grotzinger said.

JPL, a division of the California Institute of Technology in Pasadena, manages the Mars Exploration Rover project for NASA's Office of Space Science, Washington.

For information about NASA and the Mars mission on the Internet, visit http://www.nasa.gov/.

Images and additional information about the project are also available at http://marsrovers.jpl.nasa.gov/ and http://marsrovers.jpl.nasa.gov/relocate.html?relocate= http://athena.cornell.edu/.

http://marsrovers.jpl.nasa.gov and http://athena.cornell .edu.

REMOVAL OF ORGANIC MICROPOLLUTANTS AND METAL IONS FROM AQUEOUS SOLUTIONS BY ACTIVATED CARBONS

Basu Saha
Loughborough University
Loughborough, United Kingdom

INTRODUCTION

World population growth has triggered agricultural and industrial growth for the sustenance of the day-to-day demands of the population. However, the increase in agricultural and industrial growth has also led to a corresponding growth in pollution related to the two activities. The use of pesticides to boost food production has caused increased contamination of drinking water, thus putting pressure on governments to invest in the provision of clean water. The same can also be said about the increase in industrial activities, which has seen, among many of the pollutants, the increase of heavy metals in the water bodies. Since the installation of water pollution control schemes are not profit-generating activities, there is a lot of ongoing research into finding efficient but inexpensive ways of separating pollutants from drinking water.

The ban on the use of atrazine for nonagricultural purposes in the United Kingdom in 1993 has resulted in the introduction of alternative herbicides. Among them is benazolin, which can be a potential groundwater pollutant owing to its high water solubility when compared with that of atrazine. A report compiled by the Environment Agency (1) revealed that, for the first time in England and Wales, benazolin has become one of the 10 most frequently occurring herbicides in surface water. The potential harmful effects of benazolin as a herbicide necessitates that it must be removed from water bodies if present at concentrations higher than that required by environmental regulations. However, conventional water treatment processes, such as filtration, disinfection, coagulation, and sedimentation, have been reported to be more effective in removing less water-soluble and easily degradable pesticides (2). Ozone technology is a potentially attractive technique in wastewater treatment. However, in some cases, the ozone degradation products can be equally harmful. Lambert et al. (3) have investigated degradation efficiency of benazolin and other herbicides using ozone and a combination of ozone and hydrogen peroxide. They realized that for benazolin, doses of 1.0 mg O_3/l, 2.0 mg O_3/l, and 3.1 mg O_3/l, degradation efficiencies of 55%, 68%, and 88% could be achieved, respectively. The combination of ozone and hydrogen peroxide further increased the degradation efficiency, but they recommended complementary techniques for other herbicides. Adsorption on activated carbon is one of the well-established and effective techniques for the removal of herbicides from water. The adsorptive efficiency of activated carbon is because of its well-developed internal pore structure, surface area, and surface reactivity. Typically, activated carbon consists of interwoven micropores, mesopores, and macropores, and these are responsible for the high surface area of the carbons. The existence of surface functional groups in an activated carbon matrix means that they can be manipulated by thermal or chemical treatments to produce adsorbents that are tailored for particular functions.

As wastewater streams consist of a mixture of pollutants that can generally be classified as organic and inorganic pollutants, tailoring an adsorbent to remove both classes of pollutants would give a significant process advantage. Among the inorganic pollutants, heavy metals form a major part in the classification. Heavy metals constitute a class of environmental pollutants that are of particular concern in the treatment of industrial wastewater because of their potential danger to humans and the environment. Therefore, they must be treated if present at concentrations higher than those specified by the environmental regulations.

Activated carbons have been proven to be an effective adsorbent for the removal of organic compounds from water (4,5). Research has also shown that they have a potential to remove heavy metals, especially in its modified state. For example, a number of authors have oxidized activated carbons to enhance sorption of heavy metals (6–9). The oxidation processes have mainly been by use of nitric acid, ozone, and air. The three methods have always produced the same results, that of producing weakly acidic functionality on the surface of the carbon. Saha et al. (10) compared the metal sorption capacity of a conventional carbon and its oxidized form. The adsorption of herbicides and trace metal ions onto activated carbon is discussed in this chapter. In particular, we discuss the effect of modification of activated carbon on the uptake of copper and benazolin. The effect of post-treatment of the oxidized carbons on the uptake of the target pollutants is also highlighted. The adsorbents were characterized in a bid to understand their chemical and physical properties, which gives a better understanding of the nature of the association of the target pollutants and the adsorbents.

BACKGROUND

Herbicides have undoubtedly contributed to the growth of agricultural productivity. For example, it has been reported that a $4 billion U.S. investment in herbicide control saved $16 billion U.S. in crops every year (11). However, the benefit realized from the use of herbicides does not factor in the negative effects they have on human health.

As a result of their negative impact on human health and the environment, stringent legislation has been imposed on the purity of drinking water. The European Union (E.U.) agreed to have the drinking water Directive Admissible Concentration for single herbicide or pesticide at a maximum of 0.1 μg/L. In Great Britain, the National Rivers Authority (12) showed that eulan and permethrin (mainly in effluents from the textile industry) had exceeded their Environmental Quality Standard (EQS) limits. Also, diazinon, a sheep dip insecticide, had also exceeded its EQS limits. One major concern was that the most commonly used, herbicides atrazine, diuron, bentazone, isoproturon, and mecoprop exceeded the standard 0.1 μg/L in surface waters, which has been and always will be a major concern simply because most groundwater sources used for drinking water have no treatment facilities designed to remove herbicides. Also, the most common water treatment processes, such as sedimentation, disinfection, coagulation, and filtration, can effectively remove only selected herbicides, especially those that have low solubility in water. On the other hand, advanced water treatment processes (mainly using adsorption onto activated carbon) have proved to be the most efficient and reliable method for the removal of aqueous dissolved organic herbicides. Hence, a lot of studies have gone into trying to improve the removal efficiency of the existing adsorbents. In improving the adsorption capacity of adsorbents, attention has been turned on to the number of variables that affect the adsorption capacity of the activated carbons.

As pollution in wastewater is because of a cocktail of pollutants, the use of activated carbon could be a viable way of simultaneously adsorbing pollutants of a different nature, which has a further advantage of cutting down on the number of wastewater processing units. Hence, the objective of this research is to investigate the effect of simultaneous adsorption of a heavy metal (copper) and a selected herbicide (benazolin) on as-received and surface-modified activated carbons. The adsorbents chosen for this work are as-received F400 supplied by Chemviron (USA) and various surface-modified samples of F400 that were prepared in the author's laboratory. Benazolin is highly soluble in water (solubility 600 mg/l) when compared with atrazine (solubility 33 mg/l). Properties, structural formula, and three-dimensional representation of benazolin are given in Table 1 and Fig. 1, respectively. Copper has been chosen as a representative of heavy metals. In order to understand the sorption mechanism of the target pollutants, physical and chemical characterization were performed for all the adsorbents. The other objective of the study is to investigate the mechanism of adsorption of the chosen pollutants. If the mechanism is fully understood, it is possible to produce tailor-made engineered adsorbents for various purposes in water treatment.

EXPERIMENTAL

Adsorbent Materials

A coal-based commercial granular activated carbon, F400 was used for surface modification with nitric acid. The

Table 1. Properties and Structural Formula of Benazolin

Herbicide	Chemical Formula	Structural Formula	Solubility in Water (mg/L)	pKa
Benazolin	$C_9H_6ClNO_3S$	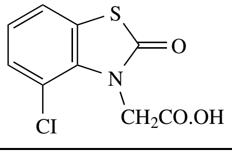	600	3.04

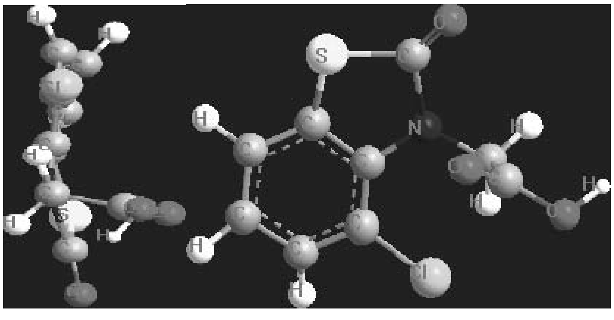

Figure 1. 3-D molecular representation of benazolin.

carbon sample was obtained from Chemviron (USA). Before the oxidation process, the carbon was washed in distilled water and dried at 383°K for 24 hours. The dried sample was then subjected to different modes of modification using part of a method described by Saha et al. (7) and Chingombe et al. (13).

Sample Preparation

A known amount of the washed and dried sample was introduced into the reaction vessel that contained nitric acid and distilled water at a ratio of 1:1 (v/v). The reaction mixture was heated to 363°K, and the reaction was allowed to continue for 6 hours with continued stirring. The reaction solution was removed and a fresh reaction mixture of nitric acid and water was introduced. The reaction was allowed to continue for an additional 3 hours under the same reaction conditions. The oxidized carbon was then washed with distilled water until no further change in pH could be detected.

The resultant water-washed samples (after oxidation) were divided into three portions. The first portion was denoted AC1, and this is a sample that was washed with water after the oxidation reaction. AC2 was a sample that was washed with 0.1 M sodium hydroxide to remove humic acids that were formed during the oxidation process. The washing was continued until no further coloration from the humic acids could be detected. Reconditioning of the sample to the hydrogen form was performed using 0.1 M hydrochloric acid, and further washing with distilled water was done until the pH of the supernatant solution stabilized to a pH of about 4.0. The conversion was deemed complete when the inlet and the outlet pH of the liquid flowing in the column were the same. The third sample, denoted AC3, was a result of heating the third portion of the oxidized sample at a temperature of 580°K under vacuum of 2 mmHg for 12 hours to remove humic acids.

ANALYSIS OF METAL ION AND HERBICIDE

Copper

The copper ion content was analyzed using a Varian SpectraAA-200 atomic absorption spectrophotometer (AAS). The AAS was set on the flame mode using air-acetylene mixture.

HPLC Analysis Method

As preconcentration of benazolin before analyzing on the HPLC was required, a trace analysis method had to be developed for the quantification of the herbicide. A Hewlette Packard 100 series HPLC consisting of a diode array detector, column thermostat, auto sampler, and a binary pump was used for analysis. Jones Chromatograph supplied a Genesis C18 silica column (4 μm, 150 × 3.0 mm) combined with a 1 cm Genesis guard column. The detector, flow rate, and injection volume were set at 222 nm, 0.88 ml/min, and 100 μl, respectively, and this gave the retention time of benazolin at 2.01 minutes. The column temperature was also set at 311°K. The eluent consisted of a solution of 40% acetonitrile and 60% buffer (10 mmol KH_2PO_4 at pH 3). The calibration curve for benazolin showed that correlations were linear over the range (0–8 ng/μl).

Solid phase extraction (SPE) technique was employed for the preconcentration of benazolin before analysis. The complete procedure for SPE is given in Table 2.

Table 2. Procedures for Solid Phase Extraction (SPE)

Conditioning	500 mg, s-triazine, 6 ml SPE columns were conditioned by passing 10 ml of HPLC-grade methanol followed by 10 ml of 2% MeOH in water under gravity.
Extraction	500 ml sample solution were passed through the column at a flow rate of 1 ml/min.
Elution	6 ml of acetonitrile was allowed to soak in the columns for 20 minutes before being passed through at approximately 0.5 ml/min and the eluent collected in 7 ml vials. The columns were then blown dry with nitrogen and the solvent was evaporated to dryness using a stream of nitrogen.
Reconstitution	500 μl of 40% acetonitrile, 60% 10 mmol KH_2PO_4 buffer at pH 3.0 was added to the vials to redissolve the analytes. The vials were shaken vigorously and their contents transferred to HPLC auto-sample vials.

CHARACTERIZATION OF ADSORBENTS

All samples were characterized to investigate the effect of the treatment and the chemical as well as the physical characteristics of these materials. The characterization included scanning electron micrographs (SEM), Brunauer-Emmett-Teller (BET) and Langmuir surface area measurements, pore size distribution, Fourier Transform Infra Red spectroscopy (FTIR) analysis, sodium capacity measurement, pH titration, zeta potential measurements, and X-ray photoelectron spectroscopy (XPS) analysis and elemental analysis. The detailed characterization procedure and the results have been reported by Chingombe et al. (13). The results indicated that a significant change occurred in the structure of the precursor by oxidation and subsequent treatments. Characterization of the adsorbents revealed that modification of the carbon by oxidation resulted in the introduction of weakly acidic functional groups, and the presence of such groups was confirmed by FTIR, pH titration, zeta potential measurements, and sodium capacity results. Sodium capacity was enhanced by a factor of up to 8 for the oxidized carbons compared with as-received samples.

Batch Adsorption Experiments

The sorption of copper on the F400 carbon range is shown in Fig. 2. The results clearly show that AC1 and AC2 have the same adsorptive capacity for copper when compared with AC3 and unoxidized F400. However, among the adsorbents employed, as-received F400 showed the least adsorptive capacity. The enhanced capacity observed on AC1, AC2, and AC3 could be as a result of the oxidation process, which generated carboxylic functional groups as evidenced in the FTIR spectra and the high sodium capacity of the above sorbents. These groups

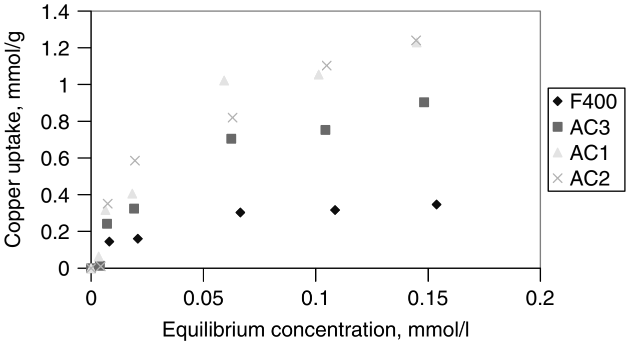

Figure 2. Equilibrium sorption results of copper on F400 and the oxidized carbons.

would be dissociated in aqueous solution, and they are responsible for copper sorption. The oxidized and heat-treated sample AC3 showed reduced capacity compared with AC1 and AC2, which could be because of the decomposition of the carboxylic acid groups during the heat treatment. Toledo-Bautista et al. (14) made similar observations when they adsorbed Cr(III) on a range of conventional and modified carbons.

The formation of metal surface complexes on the oxidized carbon involving cooperative action seems quite likely. Approximate calculation of oxygenated functional group density per unit area for F400(ox) yields a value of 0.02 functional groups per square angstrom. Given that the oxidized carbons evaluated in the present study possess a large proportion of pores in the region of 10–20 angstrom and the diameter of the hydrated metal ions is approximately 8 angstrom (15), it is reasonable to assume a cooperative binding mechanism (see Fig. 3).

Figure 4 shows the adsorption of benazolin on all the adsorbents under study. The results clearly show that benazolin adsorption on F400 is slightly higher than AC2 and reasonably higher than AC1 and AC3. However, AC1 and AC3 seem to have similar sorptive performance, whereas AC2 showed a higher performance than the other oxidized carbons, which can be expected because there was a significant loss in micropores for samples AC1 and AC3 as obtained by the pore size distribution analysis (13). It is generally accepted that sorption of an organic molecule takes place mainly in the micropore region of the adsorbents, because in this region, the contact points between the adsorbent and the molecules is higher than in mesopores and macropores. Also, it is reported that an overlap of potential forces exists in the micropores, which results in increased adsorption forces. Therefore, naturally, the capacity of AC1 and AC3 is expected to be low. Another factor in the difference between the adsorption capability of the adsorbents is the electrostatic interaction between the adsorbent and the adsorbate.

The zeta potential results show that the isoelectric point (IEP) of the oxidized carbons are significantly lower (1.5) than that of the conventional carbon (6.0) because of the introduction of weakly carboxylic acid groups with dissociation between pH 2 and 6. As the experiments are performed at pH 5, it would imply that the surface of the oxidized carbons is predominantly negative because of the dissociation of the carboxylic groups. Also, the pK_a value of benazolin (3.04) is lower than the pH of the solution experiment, and this condition would render mutual repulsion between the adsorbent and the adsorbate. Adsorption of benazolin is more favorable on F400 because of its high IEP. Streat and Horner (4), when comparing adsorption of some herbicides (including benazolin) on F400 and hypercross-linked polymers, realized that F400 outperformed the polymeric sorbents. They postulated

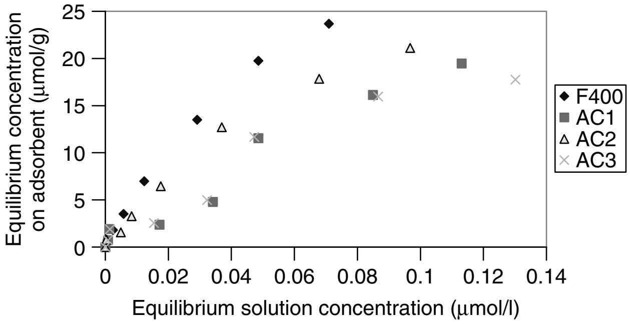

Figure 4. Adsorption of benazolin on F400 and the oxidized carbons.

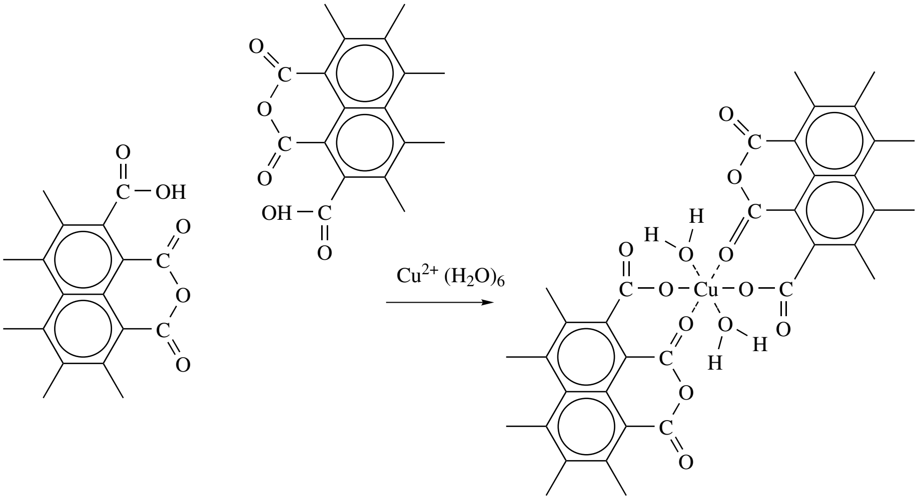

Figure 3. Postulated complexation reaction between copper (II) and oxidized carbon surface.

that adsorption of benazolin on F400 was through the dispersion interaction of the π electrons in the respective aromatic systems by a donor-acceptor mechanism. They further noted that the chloro (–Cl) group on benazolin is an electron-withdrawing group, therefore implying that a reduction of electron density occurs in the π ring system of the herbicide. Hence, benazolin would act as an electron acceptor in the donor-acceptor mechanism involving F400. This observation is consistent with investigations carried out by Radovic et al. (16) when they considered the adsorption of aromatics possessing electron-withdrawing and electron-donating functional groups on conventional and chemically modified activated carbons. Chemical modification involved samples that were subjected to nitration and others that were oxidized. Heat-treated carbons were also considered. They used nitrobenzene and aniline as the target adsorbates. Nitrobenzene contains nitro (–NO_2) group, which is known to be electron withdrawing, whereas aniline possesses an amino (–NH_2) group, which is electron donating. As nitration and oxidation reduces the π-electron density in the graphene layers of the carbon surface, which would in turn favor adsorption of aromatics with electron donor functional groups, it was realized that maximum adsorption uptake of aniline (with electron donor group) were attained on oxidized carbon surfaces. On the other hand, maximum adsorption uptake for nitrobenzene were found on heat-treated carbon surfaces (which possess higher π-electron density than the oxidized carbons). Hence, from the view point of Radovic et al. (16), it would be reasonable to postulate that benazolin adsorption is favored by adsorbents with high π-electron density. However, in the oxidized carbons employed in this work, such a mechanism would be highly unlikely because of the introduction of oxygen groups in the form of carboxylic acids. The carboxylic acid functional groups are electron withdrawing, and this condition reduces the π-electron density in the ring systems of the carbon. Therefore, the dispersion interaction between the herbicide and the oxidized carbon is reduced. Hence, unoxidized F400 would be expected to have a higher benazolin capacity than the oxidized carbons.

Figures 5 and 6 show the adsorption isotherms of benazolin in the presence of copper at 5 ppm and 10 ppm for unoxidized F400 and AC1 and AC3 samples, respectively. For both cases, it is apparent that benazolin uptake is distinctly affected in the presence of copper ions. Also, Fig. 7 shows the effect of the initial concentration of benazolin on the uptake of copper. The results clearly show that as the initial concentration of benazolin increases, a corresponding decrease occurs in the uptake of copper up to the point when the initial concentration of benazolin is at approximately 20 ppb, that is when the uptake of copper normalizes. The immediate conclusion that could be drawn from the two sets of results is that a competition exists for the same adsorption sites by copper and benazolin. As mentioned previously, the adsorption of benazolin is favored through electron donor-acceptor type of a mechanism, which involves π electron in the ring systems of the carbon. Biniak et al. (6) showed that copper ions had the capacity to interact with the π electron in the ring systems of the carbon. In their work, they employed XPS and FTIR to study the state of the copper adsorbed on different carbon surfaces. They investigated three types of modified activated carbons. One of the modified carbons was obtained by annealing a conventional carbon in a vacuum at 1000°K. This carbon had a low oxygen content of 2.72% compared with oxidized and ammonia-treated carbons, which had 10.10% and 6.20%, respectively. The annealed carbon was virtually devoid of acidic functional groups. XPS and FTIR results indicated that the dominant mechanisms of copper adsorption on heat-treated active carbons could be dipole-dipole (π-d) interactions between graphene layers and metal ionic species as well as spontaneous electrochemical reduction of copper ions. Hence, it is expected that benazolin and copper would compete for the same sites derived from the π electrons in the graphene layers of the modified adsorbents. However, conclusive evidence of

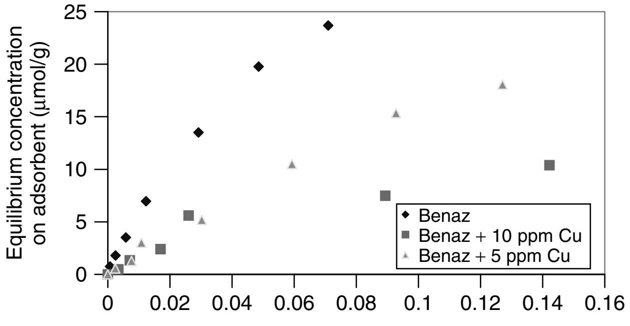

Figure 5. Simultaneous adsorption of copper and benazolin on F400.

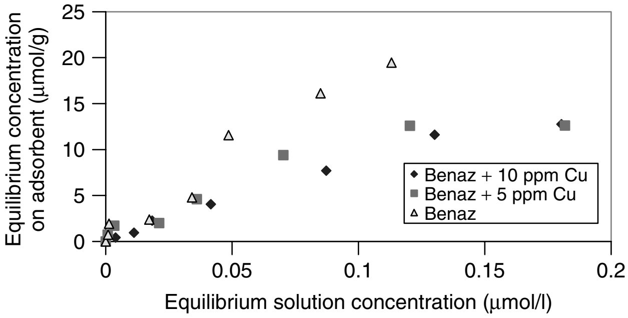

Figure 6. Simultaneous adsorption of copper and benazolin on AC1.

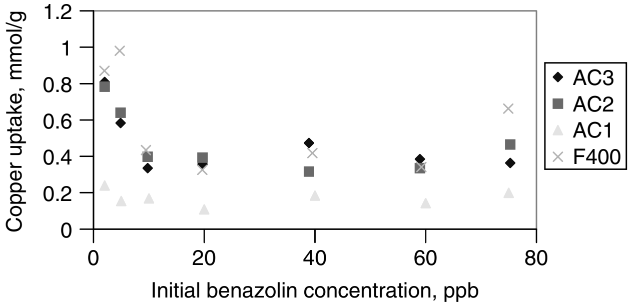

Figure 7. Effect of the initial concentration of benazolin on copper uptake.

such competition would only be proved by employing an annealed carbon adsorbent.

CONCLUSIONS

Oxidation of the conventional carbon and post-treatment of the resultant carbons had an influence on their uptake for copper and benazolin. Oxidation resulted in the introduction of weakly acidic carboxylic functionality, which markedly enhanced the uptake of copper ions. The isoelectric point (IEP) was also reduced as a result of oxidation. Oxidation of the carbon resulted in altered pore size distribution. Reduction in microporosity was observed, and it also explained the low uptake of benazolin by oxidized sample compared with as-received F400. However, post-treatment by washing with sodium hydroxide seemed to open up the pores for AC2 sample. Simultaneous adsorption of benazolin and copper showed that direct competition existed for the active sites, and possibly the two adsorbates targeted the π electron cloud on the graphene layers of the adsorbent. Copper uptake by the conventional and modified samples decreased in the presence of benazolin and vice versa. The decrease in metal uptake was dependent on the concentrations of metals in multicomponent aqueous solutions. The effect of copper on benazolin uptake was more pronounced than the effect of benazolin on copper uptake from water. Further work is necessary to fully understand the precise mechanisms of adsorption of complex aromatic molecules onto carbon, and this forms the basis of our continued work. Research is in progress in the author's laboratory to provide conclusive evidence on the interaction of these adsorbates with the modified carbon surface.

The environment is challenged by other micropollutants, namely polychlorinated hydrocarbons, organic oxygen compounds, surfactants, pesticides, aliphatic intermediates developing from the chemical industry, and, of course, endocrine disrupting compounds (EDCs). Further study of activated carbon continues to explore for potential solutions to these important problems.

Acknowledgments

I would like to express my sincere gratitude to Professor Michael Streat for his support, continuous encouragement, and advice for conducting the research in the area of adsorption and ion exchange and development of tailored sorbent materials for environmental remediation. I am indebted to Dr. Hadi Tai and Purazen Chingombe for their meticulous experimental work. Funding from European Community, EPSRC, Severn Trent Water Ltd., and Commonwealth Scholarship Commission is gratefully acknowledged.

BIBLIOGRAPHY

1. Environment Agency. (2002). *The Annual Report of the Environment Agency pesticide monitoring programme*. UK.
2. Kouras, A., Zouboulis, A., Samara, C., and Kouimtzis, T. (1998). *Environ. Pollut.* **103**: 193–202.
3. Lambert, S.D., Graham, N.J.D., and Croll, B.T. (1996). *Ozone Sci. Engineer.* **18**: 251–269.
4. Streat, M. and Horner, D.J. (2000). *Trans. IChemE.* **78**(B2): 363–382.
5. Streat, M., Malik, D.J., and Saha, B. (2004). In: *Ion Exchange and Solvent Extraction*. Vol. 16. A.K. SenGupta and Y. Marcus (Eds.). Marcel Dekker, New York, pp. 1–84.
6. Biniak, S., Pakula, M., Szymanski, G.S., and Swiatkowski, A. (1999). *Langmuir* **15**: 6117–6122.
7. Saha, B., Tai, M.H., and Streat, M. (2001). *Trans. IChemE.* **79**(Part B): 211–217.
8. Jia, Y.F., Xiao, B., and Thomas, K.M. (2002). *Langmuir* **18**: 470–478.
9. Saha, B., Tai, M.H., and Streat, M. (2003). *Chem. Eng. Res. Des.* **81**: 1343–1353.
10. Saha, B., Tai, M.H., and Streat, M. (2001). *Trans. IChemE.* **79**(Part B): 345–351.
11. Department of the Environment. (1996). *Pesticides in Water*. Department of the Environment Report, The Working Party on the Incidence of Pesticides in Water.
12. National Rivers Authority. (1998). Pesticides in the aquatic environment. *Water Quality Series* **26**: 12.
13. Chingombe, P., Saha, B., and Wakeman, R.J. (2004). *16th International Congress of Chemical and Process Engineering*, Prague, Czech Republic.
14. Toledo-Bautista, I., Utrilla-Rivera, J., Garcia-Ferro, M.A., and Castillo-Moreno, C. (1994). *Carbon* **32**(1): 93–100.
15. Nightingale, E.R.J. (1959). *Phys. Chem.* **63**: 1381–1386.
16. Radovic, L.R. et al. (1997). *Carbon*. **35**(9): 1339–1348.

MOLECULAR NETWORK DYNAMICS

D.L. Marrin
Hanalei, Hawaii

Perhaps the twentieth century's most startling discovery about liquid water is that it functions as a vast interconnected network, rather than as a random collection of independent molecules. Individual water molecules serve as network *elements*, which constitute the building blocks for the fundamental water network. Elements in the water network are connected to one another via magnetic-type linkages known as *hydrogen bonds*. In solid water (ice), each water molecule hydrogen bonds with all four of its nearest neighbors in forming a perfect *tetrahedron*, which is a three-sided pyramid comprising the most basic molecular geometry of water (see Fig. 1). As ice melts into liquid water, some of the hydrogen bonds between water molecules are broken (causing the tetrahedron to distort), whereas the remaining bonds transition to an ultra-dynamic state whereby they are switched as rapidly as a trillion times per second. The dynamics of this unique molecular network are believed to underlie many of water's anomalous properties.

NETWORK BEHAVIORS

This frantic switching of hydrogen bonds, which constitutes the underlying dynamism of water's molecular network, permits water to flow and to behave as a liquid even though it retains much of the molecular geometry of a solid. Physical chemists have long modeled liquid water as a space-filling network of individual water molecules, within

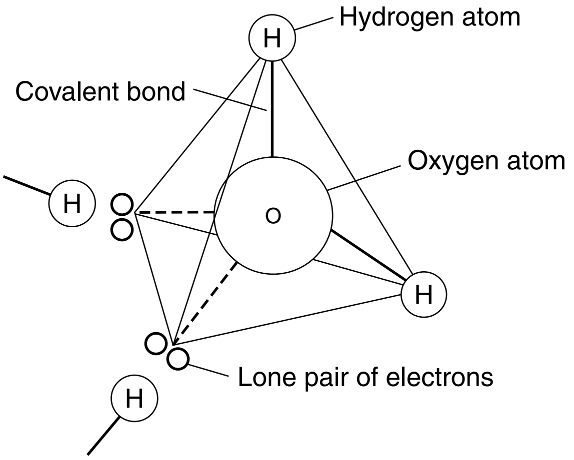

Figure 1. This diagram represents the simplest tetrahedral unit of water's H-bonded network (1). Two of the four hydrogen atoms comprising the tetrahedron are covalently bonded to the central oxygen atom (forming an individual H_2O molecule), whereas the remaining two hydrogen atoms (donated by each of two neighboring water molecules) are hydrogen bonded to the oxygen's lone pairs of electrons. Every water molecule in the network can act as both a hydrogen donor and a hydrogen acceptor in bonding with its nearest neighbors.

which all potential hydrogen bonds were described as either unbroken (ON) or broken (OFF) (2). More recently, scientists have characterized the shuffling (i.e., breaking and forming) of hydrogen bonds comprising liquid water's network as a nonlinear process and, consequently, have been unable to decipher the rules that govern the changing of molecular interconnections. This inability does not imply that the switching rules are haphazard—only that science's observational and mathematical techniques are currently overwhelmed by water's dynamism. Even scientists who study much simpler computer-generated systems have found it difficult to identify switching rules solely by analyzing network patterns and behaviors. Instead, researchers arbitrarily set the switching rules for these artificial systems and then observe corresponding network behaviors.

Whereas network hierarchies within liquid water have only recently been investigated, structural hierarchies are well described. Water exhibits structural self-organization on at least three distinct levels: molecules, clusters of molecules (e.g., hydration shells), and clusters of clusters (e.g., micelles). Self-organization implies that water molecules or molecular clusters arrange themselves into increasingly complex states. It is this self-organization that reportedly permits a network of simple interconnected elements, along with their switching rules, to display coherent behaviors (3). Although the primary water network is held together by short-range hydrogen bonds between neighboring molecules, it may either behave as an extensive system with global coherence or as individual smaller regions with local order (4). These two behaviors have been classically linked to network patterns displaying chaotic or disordered (i.e., bulk liquid water) and more ordered (i.e., water clusters) properties, respectively.

BIOLOGICAL STRUCTURE AND FUNCTION

Molecular biologists originally considered water to be merely an inert solvent within which biomolecules (e.g., proteins and nucleic acids) structured and activated themselves. More recently, water has been perceived as a key structural and functional component of many biomolecules, whose interaction with the surrounding environment is often mediated through *hydration shells*, which are immense water networks surrounding the biomolecules (5). Water releases its hydrogen bonds with biomolecular components so that protein or DNA assembly may begin. Conformational changes in the biomolecule then restructure both the internal (bound or integral) and external (hydrating) water that, in turn, facilitates more changes in the biomolecule. This iterative process appears to repeat until assembly and folding mechanisms yield a bioactive crystal.

In addition to the structural importance of water in biomolecules, the insertion and removal of water also has profound effects on their functioning. One of the most remarkable functions that proteins perform is that of a biochemical catalyst or *enzyme*, which permits chemical reactions to proceed at rates that are millions of times faster than would occur under ambient conditions. Water appears to play several roles in the functioning of enzymes. First, water molecules located at the interface of the protein-substrate complex often provide the free energy that actually binds the enzyme to its substrate. Second, water may be excluded from the active site until the moment of catalysis; at which time, it is allowed to enter the active site and reconfigure hydrogen bonds (permitting the reaction to proceed). Finally, the enzymes actually structure the water network or networks around them, thus enhancing the rate of catalysis (6).

In addition to its role as a structural component of biomolecule, water hydrates DNA as it would any other solute in an aqueous solution. It appears that water is most highly ordered adjacent to the DNA molecule and is less ordered with each successive hydration shell. This highly ordered water is sometimes referred to as *vicinal* or biological water (6). As was the case for proteins, each subunit of the DNA seems to have its own water network that interacts with other networks and, through successive hydration shells, with bulk water. If the geometry of the DNA molecule is contorted, the hydration shells and associated geometric clusters seem to be likewise affected. Similarly, the DNA molecule may contort in response to changes initiated by the water that hydrates it. The reasons for twisting and folding DNA molecules may include stuffing it into chromosomes or unraveling it during transcription and replication.

COMPLEX DYNAMICS

Scientists working with the pentagonal water cluster, which serves as a basic building block for many large hydration shells, observed that water's vibrational states are caused by a combination of bending and stretching motions of its bonds. Molecular vibrations are expressed in the units of cycles per second, or *hertz*. Although water's covalent bonds vibrate at about 10^{14} hertz (one-hundredth of a picosecond per cycle), its hydrogen bonds vibrate within the lower frequency range of 10^{12} to 10^{13} hertz (one-tenth to one picosecond per cycle). The time frames for one cycle of these bond vibrations are shorter than those measured for the exchange hydrogen bonds in water's bulk liquid network (more than one picosecond). One *picosecond* equals a time period of one-trillionth of a second (10^{-12} seconds). In addition to the dynamics associated with vibrating and exchanging bonds, water's molecular network performs several concurrent activities that have been clocked at both faster and slower rates.

As an example, *ultrafast vibrational energy* is transferred in water on timescales less than one-tenth of a picosecond (7). This ultrafast transfer could play a role in liquid water's transporting vibrational energy among certain solutes (e.g., proteins, DNA) and, perhaps, in transferring information within water's local network. Recent spectroscopy studies have suggested that liquid water's network may also function on timescales of less than a *femtosecond* (one-thousandth of a picosecond), which is at least 100 times faster than the ultrafast vibrational energy. Moreover, some of the studies indicate that most molecules in the liquid network form only two (rather than three or four) hydrogen bonds with their neighbors, perhaps creating long chains of molecules within a disordered cluster network (8). In contrast to these sub-picosecond dynamics, the reorientation of water molecules in a spatially restricted environment (e.g., the interface of a micelle) occurs on the much slower timescale of hundreds to thousands of picoseconds. Although further research is required to substantiate the recent observations, models describing both the structure and dynamics of water's molecular network are rapidly evolving.

SEAWATER

Thus far, the discussion of water's network has dealt with either pure water or dilute aqueous solutions in which solutes comprise only a small fraction (e.g., less than 0.1%) of the total solution. By contrast, the salts (ions) in seawater account for about 3.5% of the total solution. Unlike the structure of pure water's bulk phases and clusters, the molecular structure of seawater is poorly described as a result of difficulties associated with employing the investigative techniques that routinely probe the structure of pure water networks. Moreover, the high concentration of ions in seawater is hypothesized to significantly affect the water network. As the percentage of water molecules comprising hydration shells increases, the network structure of bulk water is compromised (see Fig. 2). A salting-out effect is observed at very high ion concentrations, when the interaction between water molecules and ions is sufficiently weakened to permit the formation of salt crystals.

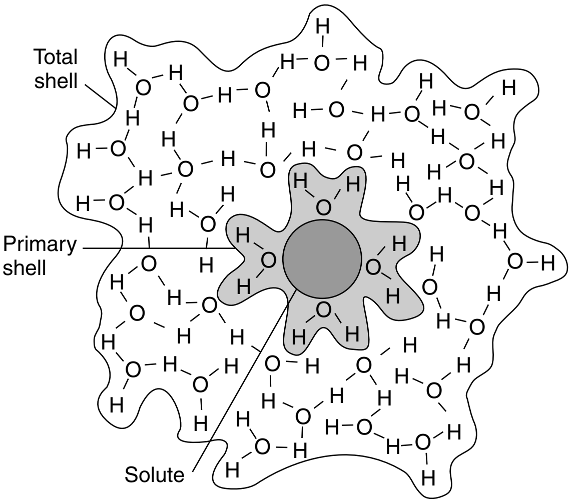

Figure 2. This diagram illustrates the primary and total hydration shells surrounding an ionic, or electrically charged, solute (1). In this case, a positively charged solute (e.g., a sodium ion in seawater) attracts the negatively charged oxygen atoms of surrounding water molecules. A solute's effect on the orientation of water molecules in the liquid network is greatest within the primary shell and diminishes as a function of distance from the hydrated ion. Theoretically, the total hydration shell includes all water molecules in the network that are affected by the solute.

The most pertinent changes in the molecular water network that have been observed or hypothesized in transitioning from dilute aqueous solutions to seawater include (1) the aforementioned ion–ion interactions, (2) overlapping hydration shells in which a single water molecule simultaneously hydrates more than one ion, and (3) varying orientations of the water molecules that comprise ion hydration shells (9). These factors are believed to profoundly affect the structure and dynamics of the water network comprising the world's oceans, which contain 97% of the planetary water. Perhaps scientists will better understand the oceans when the secrets of seawater's molecular network dynamics are more fully revealed.

BIBLIOGRAPHY

1. Marrin, D.L. (2002). West. In: *Universal Water*. Inner Ocean Publ., Maui, HI, pp. 100–114.
2. Rice, S. (1975). Conjectures on the structure of amorphous solid and liquid water. *Topics Curr. Chem.* **60**: 109–200.
3. Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind*. MIT Press, Cambridge, UK, pp. 87–90.
4. Blumberg, R.L., Stanley, H.E., Geiger, A., and Mausbach, P. (1984). Connectivity of hydrogen bonds in liquid water. *J. Chem. Phys.* **80**: 5230–524.
5. Westhof, E. (1993). Structural water bridges in nucleic acids. In: *Water and Biological Macromolecules*. E. Westhof (Ed.). CRC Press, Boca Raton, FL, Chapter 7.
6. Clegg, J. (1979). Metabolism and the intracellular environment: the vicinal-water network model. In: *Cell-Associated Water*. W. Drost-Hansen and J. Clegg (Eds.). Academic Press, New York, pp. 363–413.
7. Nitzan, A. (1999). Ultrafast relaxation in water. *Nature* **402**: 472–475.
8. Zubavicus, Y. and Grunze, M. (2004). New insights into the structure of water with ultrafast probes. *Science* **304**: 974–976.
9. Robinson, G., Wilse, S.B.Z., Singh, S., and Evans, M. (1996). *Water in Biology, Chemistry and Physics*. World Scientific, Singapore, pp. 245–259.

IN SITU CHEMICAL MONITORING

Melissa A. Singer Pressman
Joseph H. Aldstadt, III
Genetic Technologies, Inc.
Testing Institute
Waukesha, Wisconsin

INTRODUCTION

For routine monitoring of environmental systems, *in situ* measurements offer several distinct advantages over laboratory analysis (1). *In situ* analysis allows much higher rates of sample throughput. As a result of the rapid dynamics of most environmental systems, it is important to collect frequent samples to prevent undersampling of the phenomena under study, which is essential to allow for a more accurate characterization of the true dynamics of the system and, therefore, a better understanding of the biogeochemical processes involved (2). Additionally, *in situ* analysis eliminates sample contamination that inevitably develops during the transport and storage of samples to the laboratory. Furthermore, current research efforts in several laboratories have demonstrated that relatively inexpensive monitors can be fielded that can self-calibrate, can be remotely controlled, and can provide reliable measurements in realistic environments (3). In particularly, it is the inability to provide high data quality at trace levels in complex matrices that has prevented *in situ* measurements from gaining more favor.

In situ monitors have been classified broadly as either "chemical sensors" or "chemical analyzers." Chemical sensors depend on diffusion to bring the analyte to the transducer surface. This type of sensor is mechanically simple, sensitive, and practical from the standpoint of low power requirements and small footprint, but often suffers from poor selectivity and the inability to self-calibrate. More importantly, sensors are based on reversible chemical reactions, i.e., the interaction of the analyte with the transducer surface, which makes it difficult for these sensors to detect environmentally relevant levels; that is, to detect trace levels of a chemical species, the equilibrium binding (formation) constants must be very large. For the reaction of analyte A with ligand B to form a complex AB, the rate equations are (3):

$$aA + bB \rightarrow ABrate = -(d/dt)\ [A] = -(d/dt)\ [B]$$

$$= k_1[A]^{\mathrm{a}}[B]^{\mathrm{b}}$$

$$xAB \rightarrow A + Brate = -(d/dt)\ [AB] = k_{-1}[AB]^{\mathrm{x}}$$

The equilibrium constant for the dissociation of the complex is equal to the ratio of the rate constants for the dissociation of the analyte-ligand complex and the formation of the complex, as shown in the following equation (3):

$$K_D = k_{-1}/k_1$$

If the complex dissociates slowly, the time required to return the signal to a steady baseline ("clear down" time) is introduced. One way to avoid this problem is to continuously renew the reagent at the sensor surface. This strategy depends on molecular diffusion to transport the analyte to and from the sensor surface; this type of sensor is called a "renewable reagent sensor" (4–6).

Chemical analyzers, on the other hand, depend on *active* transport of the analyte into and away from the monitor. Although analyzers lack the mechanical simplicity of sensors, the addition of valves and pumps in the design can more importantly provide a means by which to accommodate self-calibration of the monitor. The addition of these mechanical capabilities also allows for multistep, complex chemistries to be used as a way to increase sensitivity and selectivity of the monitor.

Continuous flow analysis (CFA) is a more general term than so-called "chemical analyzers," referring to any measurement process in which the concentration of an analyte is determined based on a steady flow of sample through the instrument (7). Nonchromatographic CFA-based measurement systems are widely used for industrial process control and environmental monitoring. As the instrument is often placed in a hostile environment in these types of applications, the relative simplicity of the CFA approach is attractive (8). Initial attempts to use CFA-based instruments were direct adaptations of benchtop (i.e., laboratory) CFA systems that were ruggedized (e.g., water-proofed) and taken to the sampling site (9). This approach was burdensome, expensive, and far from robust.

An example of an early effort of this type is illustrated by the work of Elliott et al. in using a segmented continuous flow analyzer, often called an "autoanalyzer," for shipboard chemical analyses (10). These researchers monitored nitrate concentrations in the Everglades using a commercially available autoanalyzer that they modified for the nitrate derivatization chemistry. In their design, air bubbles were created in-line to segment the liquid sample stream as a means to prevent uncontrolled dispersion

More recently, *in situ* CFA monitors have been developed that are not merely ruggedized bench-top instruments. One of the more interesting instruments, and the model which inspired this research, was described in work published in 1994 by Jannasch et al. at the Moss Landing Marine Laboratory in Monterey, CA (11). This group of chemical oceanographers was interested in accurately measuring nitrate ion in seawater with high sampling frequency. To accomplish this goal, they designed a CFA monitor "from the ground up" and incorporated several novel features. In doing so, they created a new paradigm in the science of chemical measurements.

In the system developed by Jannasch et al., a well-characterized, irreversible colorimetric method was used to measure nitrate. The dissolved nitrate was determined by reduction to nitrite in a neutrally buffered (imidazole) solution by an in-line copperized cadmium surface. The nitrite then merged with an incoming reagent stream containing sulfanilamide (*p*-aminobenzenesulfonamide or

PABSA) in acid to undergo reaction to form a diazonium ion. This species was then reacted with *N*-(1-naphthyl)ethylenediamine dihydrochloride (NED) to form an azo dye of high molar absorptivity that was detected by molecular absorbance spectroscopy.

The novel *in situ* monitor that the Moss Landing group developed used nonmechanical osmotic pumps to both deliver reagents to the system and to propel the continuous flow sample stream through the microscale manifold (∼1 mm diameter conduits) (11). For the two reagent osmotic pumps (i.e., for PABSA and NED), the flow was driven by the osmotic pressure differential developed across the semipermeable membrane that separated a saturated brine solution from a solution with deionized water (12). The pressure created from the diffusion of deionized water across the semipermeable membranes resulted in compression of inner bags containing the reagents and their pulse-free delivery to the manifold (at ∼4 μL per hr). For the sample pump, conventional osmotic pumps were modified by removing the inner bags and mounting the rigid permeable membranes in an enclosed container. Four of these modified osmotic pumps were then coupled to draw sample through the manifold at 12 μL per hr.

Microscale flow manifolds can be fabricated by "cold-forging" PVC: A wire bent in the desired flow microconduit configuration was pressed onto the surface of a flat-milled sheet of PVC. Following removal of the wire, ∼0.5-mm diameter holes were drilled into the plate for connection of components. The optical system in the Moss Landing nitrate analyzer used a simple light-emitting diode (LED) source with photodiode detector, with data collected by a low-power programmable datalogger. The monitor exhibited a detection limit of 0.11 μM and a linear response to 20 μM nitrate. Measurements were taken every 2 min and the response time was ∼30 min. Quality control samples were measured automatically every 48 hours. Jannasch et al. demonstrated the *in situ* monitor's stability by operating it continuously for month-long periods in large tanks at the Monterey Bay Aquarium.

The Moss Landing group's research incorporated several novel features—mainly the microscale flow design using nonmechanical osmotic pumps for reagent delivery and sample intake—into an *in situ* monitor capable of frequent and reliable measurements in environmentally complex samples. In addition, through the periodic addition of nitrate standards, the monitor automatically self-calibrated by correcting for baseline drift. Another advantage to this system was its flexibility; using different irreversible chemistries, the *in situ* monitor could be readily adapted to measure other analytes for optical detection and quantitation. A drawback of the system was that it detected relatively high levels of analyte (i.e., micromolar levels), albeit levels that are relevant to studying this important nutrient. Concerns over particles in the water eventually clogging the microconduit channels are less serious, because small particles (approx. <50 μm) cannot be drawn into the manifold at the low flow rates used.

In 1999, Byrne and coworkers at the University of South Florida developed a monitor based on long pathlength absorbance spectroscopy (LPAS), which could be used for the determination of nitrate and several other biologically important chemicals in seawater at much lower levels. These workers used the same well-characterized chemistry as Jannasch et al., but they benefited from a 1-m liquid core waveguide as the absorbance flow cell. Another difference with the Byrne group's approach was in the use of mechanical (peristaltic) pumps instead of osmotic pumps—resulting in a five-fold increase in response time. This monitor was field tested recently in the Gulf of Mexico, and detection limits for nitrate were observed to be approximately 2 nM (13,14).

The next key development was reported recently by Worsfold and coworkers at the University of Plymouth. A portable *in situ* monitor for the determination of trace phosphate in surface waters was developed (15,16). This work demonstrated a monitor that was built with a forged plate similar in design to the work done by Jannasch et al. The phosphate monitor had a throughput of ∼225 determination per hr and a detection limit of 0.15 μM with 2.0% RSD ($n = 9$) reported. This monitor was evaluated by mapping the phosphate concentration in Port Phillip Bay in Australia for a brief (2 hr) period. The agreement between determinations obtained using the *in situ* monitor and data obtained using manual sampling aboard the ship followed by laboratory analysis was within ±10%.

INSTRUMENT DEVELOPMENT FOR *IN SITU* MONITORING OF CR(VI)

For the *in situ* measurement of relatively low levels of important biological nutrients such as nitrate or phosphate, the work of the Jannasch, Byrne, and Worsfold groups (and others) was pioneering. However, clearly the next challenge in developing *in situ* monitoring instruments was the ability to measure relevant levels of trace metal ions in surface waters by the *in situ* approach. Hexavalent Cr was chosen as the target analyte because it is regulated at nanomolar levels (low parts per million) in surface waters. In addition, the metropolitan-industrial areas of southern Lake Michigan, specifically the Milwaukee Harbor area, contain industrial sources that generate a concentration of total Cr in surface water that ranges from 0.1 to greater than 100 μg/L (17). These reasons, coupled with the well-characterized spectrometric chemistry (18–21), made Cr a logical choice as a representative analyte to study for the development of an *in situ* monitor that measures trace metals in the environment.

A benchtop method for determining hexavalent chromium in high particulate-containing surface waters by sequential injection (SI) was developed (22). For utility in particulate-rich waters, the relative performance of two membrane-based methods for sampling was compared. The first membrane approach was based on a commercial design known as the "supported capillary membrane sampler (SCMS)" that uses tubular membranes; the second approach was based on a conventional parallel-plate dialyzer (PPD) design that uses planar membranes. The membranes were evaluated using the colorimetric method for the determination of hexavalent chromium by complexation with 1,5-diphenylcarbazide (DPC). Thin-walled (~200 μm) microporous (pore size ~0.2 μm) polypropylene membranes were equilibrated with DPC during each sampling period. Formation of the DPC-Cr(VI) complex allowed for efficient membrane transport; without the membrane, Cr(VI) transport decreased by ~90%. Factors optimized included reagent concentration, sampling time, flow rate, and spectrophotometric conditions. Optimal conditions for the reagents were 2.00 mM DPC and 0.100 M nitric acid, and sampling times of 600 s and 900 s for the planar and tubular designs, respectively. The planar (PPD) design increased the sensitivity relative to the tubular (SCMS) design by ~225%. The PPD-SI method was applied to the determination of dissolved Cr(VI) in high particulate-containing surface water samples. Figures-of-merit included a detection limit of <20 μg/L, precision of 1.1% relative standard deviation at 100 μg/L ($n = 4$), and selectivity for dissolved Cr(VI) in several surface water samples with high levels of particulate matter and potential interferences.

The optimized method conditions from the SI studies was then applied to the design, fabrication, and testing of an *in situ* remote monitor for the measurement of sub-part per million levels in surface waters (23,24). The monitor was based on a continuous flow analysis (CFA) design in which a narrow conduit (1.0 mm) manifold was hot-forged between Plexiglas plates. A miniature low-power peristaltic pump was used to draw sample (1.5 mL/hr) into the manifold where chromium (VI) in the sample was derivatized to its diphenylcarbazide complex. The absorbance of the complex was then measured in a light-emitting diode/photodiode detection cell. Solenoid-actuated osmotic pumps were used to introduce quality control standards at user-determined periodicities to achieve acceptable data quality. Figures-of-merit included a detection limit of 0.25 mg/L (7 min cycle time), precision of 2.5% relative standard deviation at 1 mg/L ($n = 3$), and selectivity for dissolved Cr(VI) in several surface water samples with moderate levels of particulate matter. Continuous testing of the monitor over a three-week period using Milwaukee Harbor water fortified with Cr(VI) yielded results that were in close agreement (±5%) with an ICP-MS reference method.

BIBLIOGRAPHY

1. Hanrahan, G., Ussher, S., Gledhill, M., Achterberg, E.P., and Worsfold, P.J. (2002). High temporal and spatial resolution environmental monitoring using flow injection with spectroscopic detection. *TrAC, Trends Analytic Chem.* **21**: 233.
2. Rawls, R. (1998). Wanted: instruments to view ocean's dynamics. *Chem. Engineer. News*. **31**.
3. Johnson, K.S. and Jannasch, H.W. (1994). Analytical chemistry under the sea surface: monitoring ocean chemistry *in situ*. *Naval Research Rev.* **3**: 4.
4. DeGrandpre, M.D., Baehr, M.M., and Hammar, T.R. (1999). Calibration-free optical chemical sensors. *Analytical Chem.* **71**: 1152.
5. Inman, S.M., Stromvall, E.J., and Lieberman, S.H. (1989). Pressurized membrane indicator system for fluorogenic-based fiber-optic chemical sensors. *Analytica. Chimica. Acta.* **217**: 249.
6. Degrandpre, M.D. (1993). Measurement of seawater pCO2 using a renewable-reagent fiber optic sensor with colorimetric detection. *Analytical Chem.* **65**: 331.
7. Ruzicka, J. and Hansen, E.H. (1988). *Flow Injection Analysis*, 2nd Edn. John Wiley & Sons, New York.
8. Johnson, K.S., Coale, K.H., and Jannasch, H.W. (1992). Analytical chemistry in oceanography. *Anal. Chem.* **64**: 1065A.
9. Valcarcel, M., Cardenas, S., and Gallego, M. (2002). Continuous flow systems for rapid sample screening. *Trac-Trends Analytic Chem.* **21**: 251.
10. Elliott, C.L., Snyder, G.H., and Cisar, J.L. (1989). A modified Autoanalyzer-II method for the determination of NO_3^- in water using a hollow-Cd reduction coil. *Communications in Soil Science and Plant Analysis* **20**: 1873.
11. Jannasch, H.W., Johnson, K.S., and Sakamoto, C.M. (1994). Submersible, osmotically pumped analyzers for continuous determination of nitrate in-situ. *Analytic Chem.* **66**: 3352.
12. Pump, A.O. (2000). *Technical Note*. Alzet, Inc.
13. Kaltenbacher, E.A., Byrne, R.H., and Steimle, E.T. (2001). Design and applications of a chemical sensor compatible with autonomous ocean-sampling networks. *IEEE J. Oceanic Engineer.* **26**: 667.
14. Bryne, R.K. (1999). Autonomous *in situ* analysis of the upper ocean. *Sea Technol.* **71**.
15. Hanrahan, G., Gledhill, M., Fletcher, P.J., and Worsfold, P.J. (2001). High temporal resolution field monitoring of phosphate in the River Frome using flow injection with diode array detection. *Analytica. Chimica. Acta.* **440**: 55.
16. Lyddy-Meaney, A.J., Ellis, P.S., Worsfold, P.J., Butler, E.C.V., and McKelvie, I.D. (2002). A compact flow injection analysis system for surface mapping of phosphate in marine waters. *Talanta.* **58**: 1043.
17. Environmental Defense Fund. (2004). Available: http://www.scorecard.org/.
18. Miller-Ihli, N.J. (1992). Chromium. *Technol. Instrum. Anal. Chem.* **12**: 373.
19. Lynch, T.K. and Wilson, J. (1984). Speciation of metals in solution by flow injection analysis. Part 1: Sequential spectrophotometric and atomic absorption detectors. *Analyst.* **109**: 839.
20. Losi, M.E., Amrhein, C., and Frankenberger, W.T. (1994). *Rev. Environ. Contamin. Toxicol.* **136**: 91.
21. Kotas, J. and Stasicka, Z. (2000). Chromium occurrence in the environment and methods of its speciation. *Environ. Pollut.* **107**: 263.
22. Pressman, M.A. and Aldstadt, J.H. (2003). A comparison of diffusion samplers for the determination of hexavalent chromium by sequential injection spectrophotometry. *Microchem. J.* **74**: 47–57.
23. Pressman, M.A. (2003). *Development of Novel Analytical Techniques and Methods for the Determination of Trace Chromium in the Environment*, Ph.D. dissertation, University of Wisconsin-Milwaukee.
24. Pressman, M.A. and Aldstadt, J.H. (2005). *Development and Testing of an in situ Chemical Monitor for Quantitation of Hexavalent Chromium in Natural Waters*, in preparation.

NITROGEN

Kenneth F. Steele
Kristofor R. Brye
University of Arkansas
Fayetteville, Arkansas

INTRODUCTION

Nitrogen (N) is ubiquitous in the environment and in the correct form is an essential nutrient for plants and animals. About 26.5% of Earth's nitrogen is in rocks and essentially all of the remainder (73%) is in the atmosphere as an inert monotomic gas (N_2) that comprises 78% of the atmosphere (1). Only a small amount of nitrogen is present in the hydrosphere and biosphere, but nitrogen is a crucial nutrient there. For nitrogen to be used by organisms, it must be in a usable form; it must be combined or "fixed" with oxygen or hydrogen. The most common fixed form of nitrogen in ground water is the anionic oxide—nitrate (NO_3^-), and the second most common form is the cationic form–ammonium (NH_4^+). Approximately 85% of the nitrogen in stream water is in the form of organic N and most of the rest (NO_3^- and NH_4^+) is derived from the decomposition of organic matter (2). N_2 is converted to organic nitrogen predominantly by microorganisms and certain plants. About 96% of the nitrogen in soil is bound in organic matter, and nitrate is a product of organic matter degradation (3). Nitrogen occurs in five oxidation states, −3, 0, +1, +3 and +5, that are primarily the result of bacteria-induced oxidation. The nitrogen species are listed in Table 1.

NITROGEN CYCLE

Fixation by plants can occur; leguminous plants convert N_2 to organic nitrogen compounds, but most plants cannot use N_2 directly. Most plants require atmospheric nitrogen converted to ammonium or nitrate ions (4). Organic nitrogen is not directly available to plant roots and has to be mineralized in the soil by a chain of microbial processes (Fig. 1). The release of ammonium and ammonia (total ammonia) from organic matter is referred to as mineralization (3) or ammonification. Total ammonia is readily oxidized to nitrate. The conversion of nitrogen gas to fixed form and from fixed form to nitrogen gas includes four main conversion reactions that form the essential parts of the nitrogen cycle (Fig. 1):

1. fixation
2. ammonification
3. nitrification
4. denitrification

Table 1. Common Forms of Nitrogen in the Environment and Typical Concentrations of Nitrogen

Species	Oxidation State	Typical Source	Typical Concentration in Water as N
N_2	0	Atmosphere	14 mg/L
NO_3^-	+5	Oxidation of organic matter/ atmosphere	<0.6 mg/L background to 250 mg/L for contaminated sources
NO_2^-	+3	Intermediate form in nitrification, denitrification, and reduction of NO_3^-	<0.1 mg/L
NH_3	−3	Fixation of N_2	<0.01 mg/L
NH_4^+	−3	Ionization of NH_3	<0.01 mg/L
N_2O	+1	Reduction of NO_2^-	Negligible

Fixation

For N_2 gas to be used by plants, it must first be fixed with hydrogen to form ammonium ions. Fixation is a microbial process, and even the few types of plants that can use N_2 directly are ultimately dependent on microbes also. For example, legumes can use N_2 because of the symbiotic relationship between the plant and *Rhizobium* bacteria that live in nodules in the plant roots. Other means of fixation include conversion of atmospheric N_2 to nitrate by lightning and industrial processes, such as the Haber–Bosch process for producing ammonia (NH_3 gas) (4).

Ammonification

Ammonification (also referred to as mineralization), the process of converting organic nitrogen (e.g., urea) to ammonium (Eq. 1), generally occurs during decomposition of animal and plant matter (4). Dissolution of ammonium carbonate releases ammonium ions:

$$\underset{\text{Urea}}{H_2NCONH_2} + 2H_2O \longrightarrow \underset{\text{Ammonium carbonate}}{(NH_4)_2CO_3} \quad (1)$$

As indicated by Eqs. 2 and 3, the relationship between NH_3 and NH_4^+ is pH dependent. Under alkaline conditions (pH > 9), formation of NH_3 is favored due to neutralization of NH_4^+ by hydroxide (OH^-) (Eq. 2). In acid conditions (Eq. 3), ammonia gas is transformed into

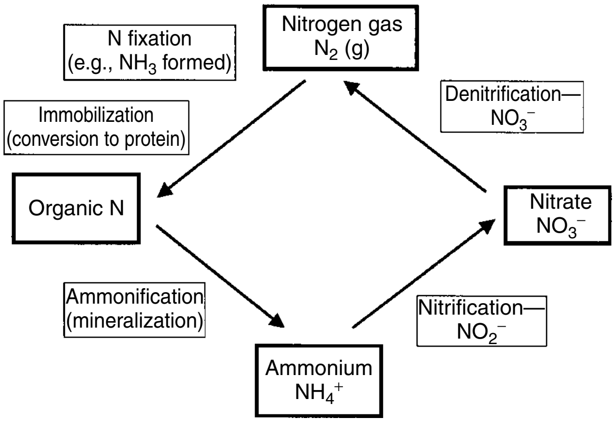

Figure 1. Simplified nitrogen cycle.

ammonium ions. For example, at 15 °C, ammonia gas makes up 0.027% of total ammonia ($NH_3 + NH_4{}^+$) at a pH of 6.0, but makes up 2.7% of total ammonia at a pH of 8.0. Although 2.7% is a small amount, it is a hundredfold increase of ammonia (5).

$$NH_3 + H_2O = NH_4{}^+ + OH^- \quad (2)$$

$$NH_3 + H^+ = NH_4{}^+ \quad (3)$$

Other factors that affect the concentration of NH_3 in water are temperature and ionic strength. The concentration of NH_3 increases as temperature increases and decreases as ionic strength increases. At 15 °C and pH 8.0, the percent NH_3 in solution is 2.7%, whereas, at 30 °C and pH 8.0, the concentration increases to 7.5% (5). The effect of ionic strength is minor for fresh to slightly saline waters.

Nitrification

Nitrification is the biological oxidation of ammonium ions to nitrite followed by conversion of nitrite to nitrate. Biological oxidation is carried out by *Nitrosomonas* and *Nitrobacter* bacteria (6,7) as shown in Eqs. 4 and 5. In addition to these two bacteria, (8) indicate that several different autotrophic organisms mediate nitrification:

$$2NH_4{}^+ + 3O_2 \xrightarrow[\text{bacteria}]{\textit{Nitrosomonas}} 2NO_2{}^-(\text{nitrite}) + 4H^+ + 2H_2O \quad (4)$$

$$2NO_2{}^- + O_2 \xrightarrow[\text{bacteria}]{\textit{Nitrobacter}} 2NO_3{}^-(\text{nitrate}) \quad (5)$$

Because oxidation of nitrite to nitrate is a rapid reaction and nitrite is metastable, nitrite is not abundant (Table 1) (9). Note that if both oxygen and nitrate are present in soil, bacteria preferentially use oxygen because it yields more energy (1). Therefore, nitrification typically occurs in oxic conditions, whereas denitrification occurs in anoxic conditions.

Denitrification

Denitrification is the biological reduction of nitrate to nitrogen gas by heterotrophic bacteria, such as *Pseudomonas denitrificans* and *Thiobacillus denitrificans* (8). These bacteria require organic carbon for energy (Equations 6 and 7).

$$NO_3{}^- + \text{organic carbon} \xrightarrow{\text{Heterotrophic bacteria}} NO_2{}^- \quad (6)$$

$$NO_2{}^- + \text{organic carbon} \xrightarrow{\text{Heterotrophic bacteria}} N_2 + CO_2 + H_2O \quad (7)$$

Because oxidation of organic carbon by oxygen yields more energy, generally the oxygen concentration must be very low (anoxic conditions) for nitrate to be used (4,6).

Analytical Terminology

Analytically, organic nitrogen (proteins, peptides, nucleic acid, urea, and other compounds) and ammonia are determined together and are referred to as "Kjeldjahl nitrogen," a term that refers to the analytical method used. By expressing all nitrogen species as N, the concentrations of the various species can be added and subtracted. Organic N (−3 charge) is determined by analyzing ammonia by other methods and subtracting this value from the Kjeldhal nitrogen concentration (10).

Inorganic N is the sum of nitrate, nitrite, and total ammonia (ammonia plus ammonium). Some analytical methods determine the sum of nitrate and nitrite, but the results are sometimes incorrectly reported as nitrate because nitrite concentrations are considered very low. Total N is the sum of organic N and inorganic N.

AQUATIC NITROGEN

Despite the large proportion of nitrogen in the atmosphere, its abundance in water is low-measured as a few mg/L N, mostly in the inert form of nitrogen. Typically, the most environmentally important form of nitrogen is nitrate; however, dissolved organic nitrogen makes up the majority of total dissolved nitrogen in watersheds in New England (11). Note that for primary producers ammonium may be more available to organisms than nitrate (12). The various common species of nitrogen in the environment and the estimated concentration of each are presented in Table 1. Mueller and others (13) have demonstrated that background concentrations of nitrate as N are less that 0.6 mg/L.

Nitrogen species are subject to oxidation/reduction (redox) and reactions with hydrogen ions (pH). In Fig. 2, the $NO_3{}^-/NO_2{}^-$ pe–pH boundary is dotted to indicate that $NO_2{}^-$ is metastable. Nitrate, at a pe above the $NO_3{}^-/NO_2{}^-$ boundary, is reduced at the $NO_3{}^-/N_2$ boundary to form nitrogen. Thus, there is little $NO_3{}^-$ available to form $NO_2{}^-$. Because nitrite is a metastable species, it occurs only in low concentrations in natural waters where denitrification is occurring (9). Nitrate is present in the N_2 field, but under natural conditions, it is present in concentrations less than those of N_2. Ammonium ions are the dominant nitrogen species at low pe and pH. Ammonia is the dominant species above pH 9 (Fig. 2) and thus is not present in significant concentrations in natural waters; most natural waters have a pH range of 5–9. Dissolved organic N can also be relatively mobile.

Mobility

Nitrate is very soluble in water and because of its negative charge, is not adsorbed by humus or clay. Nitrate not taken up by plants can be transported from the upper soil zone to lower soil zones and eventually to groundwater. Ammonium ions are not nearly as mobile as nitrate ions, in part, because their positive charge allows their transport to be retarded by adsorption on humus and clay. In some situations, ammonium ions can be transported to deep soil horizons (3).

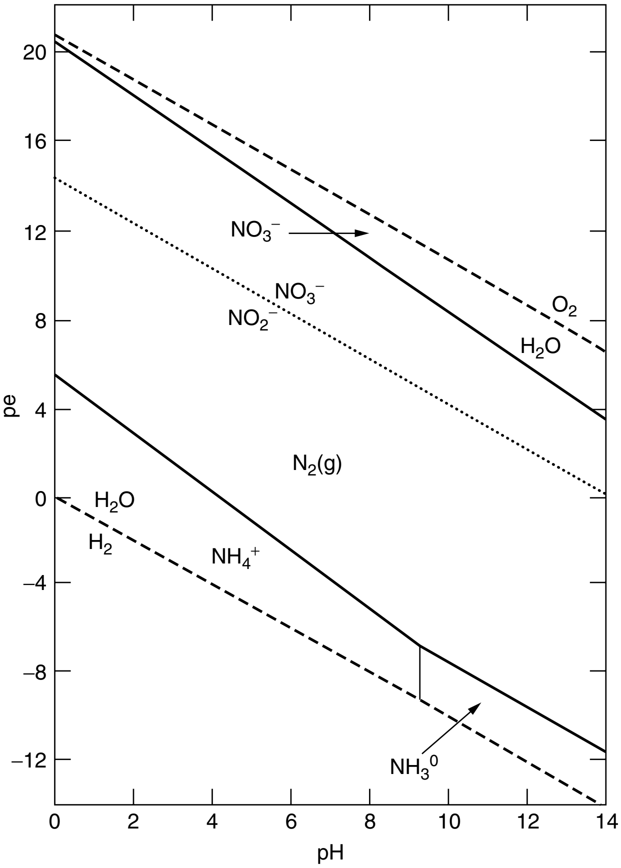

Figure 2. Stability fields for common nitrogen species in natural waters. The diagram is based on the following conditions: temperature of 25 °C, partial pressures of H_2 and O_2 are each 1 atmosphere, partial pressure of N_2 is 0.77 atmosphere, and total aqueous N concentration is 0.001 moles per liter.

Toxicity

Ammonia is a pungent, colorless gas that is very soluble in water and is toxic to fish at concentrations greater than 0.2 mg/L. Dissolved ammonia reacts with water to form ammonium ions, which are not nearly as toxic (1/50th) as the gas (3). The ammonia–ammonium equilibrium is controlled to a large extent by temperature and pH; only highly alkaline waters favor the formation of ammonia. The pH range of most natural waters is such that ammonium ions are dominant compared to ammonia gas (Fig. 2). The U.S. Environmental Protection Agency bases its ammonia criteria for water on a complex set of equations that include the effect of temperature and pH on ammonia concentration. The equations were developed to account for acute and chronic toxicity to particular types of fish. In addition, the criteria are based on specific periods of time and the frequency of exceeding these criteria (14).

Typically nitrate is toxic to warm-blooded animals only where large amounts of nitrite are formed by the reduction of nitrate. Nitrate can be reduced in the gastrointestinal tract to nitrite, especially in infants. Nitrite enters the bloodstream and reacts with hemoglobin to produce methemoglobin that reduces the number of hemoglobin cells available for oxygen exchange. These reactions can be life threatening for children less than 3 years old. Babies who consume water with nitrate concentrations as low as 1 mg/L as N risk methemoglobinia or "blue baby disease" because of the impairment of oxygen transport. There are also indications that nitrate (as related to nitroso compounds) may be associated with gastric cancer and perhaps liver and kidney cancer. The relative consumption of nitrate from water compared to consumption of dietary nitrate is a factor that concerns researchers attempting to determine the risk of cancer from water (15,16). The maximum concentration limit (MCL) of 10 mg/L N for nitrate in drinking water is based on minimizing methemoglobinia and is considered safe concerning cancer.

CONTAMINATION

The causes of pollution of water are divided into point sources and nonpoint sources. Nitrogen contamination of water is generally related to nitrate from sewage and septic systems (point sources) and from agricultural sources. The agricultural sources are commercial and animal manure fertilizers (nonpoint sources) (17,18) and from confined animal operations (point sources). Agriculture is generally considered the main source of nitrogen pollution of water (19). During rainstorms, streams often exhibit a dilution of nitrate concentration that indicates that most of the nitrate is transported to these streams by groundwater. Excessive nutrients in runoff can result in eutrophication of surface waters, for example, the Neuse River in North Carolina, U.S.A. (20). Best management practices have been developed to minimize the negative impacts of fertilizers (17).

ISOTOPES

There are two stable isotopes of nitrogen, ^{14}N and ^{15}N. Although the whole earth abundance of ^{14}N is 99.63‰ of total nitrogen, due to fractionation, the ^{15}N:^{14}N ratio is 0.003677 (21) in the atmosphere. Differences in the ratio of the two isotopes in the various nitrogen species are a result of the faster chemical reaction of the lighter isotope than the heavier isotope. Isotopic ratios of nitrogen are expressed as

$$\delta^{15}\mathrm{N} = [(R_{\mathrm{sample}}/R_{\mathrm{air}}) - 1] \times 1000$$

where R is the abundance ratio of ^{15}N to ^{14}N. The international standard for δ^{15}N is air, which by definition has a δ^{15}N of 0‰. Values greater than 0‰ indicate enrichment in ^{15}N, and values less than 0‰ indicate depletion of ^{15}N. Sources of nitrate have been determined using δ^{15}N. The premise is that agricultural activity and sewage (including septic tank effluent) may have distinct δ^{15}N values. A definitive determination of source is sometimes thwarted by the possibility of overlapping δ^{15}N values between possible sources and changes in δ^{15} N due to volatilization of NH_3 from the soil and denitrification of nitrate in water. The origin of the nitrate in water can be found by using δ^{18}O to recognize different sources of nitrate and the influence of denitrification on δ^{15}N (21,22). A number of studies have had success using N and often O isotopes to determine the source of nitrate contamination (23,24).

BIBLIOGRAPHY

1. Hitchon, B., Perkins, E.H., and Gunter, W.D. (1999). *Introduction to Ground Water Geochemistry*. Geoscience, Alberta, Canada, pp. 195–200.
2. Berner, E.K. and Berner, R.A. (1987). *The Global Water Cycle*. Prentice-Hall, Inc., Englewood Cliffs, NJ, p. 397.
3. Canter, L.W. (1996). *Nitrates in Groundwater*. CRC Lewis, Boca Raton, FL.
4. Jeurgens-Gschwind, S. (1989). Ground water nitrates in other developed countries (Europe)–relationships to land use patterns. In: *Nitrogen Management and Ground Water Protection*. R.F. Follett (Ed.). Elsevier, New York, pp. 75–138.
5. U.S. EPA. (1976). *Quality Criteria for Water*. U.S. Environmental Protection Agency, Washington, DC, pp. 10–13.
6. Manahan, S. (1999). *Environmental Chemistry*. CRC Press, Boca Raton, FL.
7. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry*, 3rd Edn. John Wiley & Sons, New York, pp. 927–934.
8. Kendall, C. and Aravena, R. (1999). Nitrate isotopes in groundwater systems. In: *Environmental Tracers in Subsurface Hydrology*. P.G. Cook and A.L. Herczeg (Eds.). Kluwer Academic, Boston, pp. 261–298.
9. Kehew, A.E. (2001). *Applied Chemical Hydrogeology*. Prentice-Hall, Upper Saddle River, NJ.
10. Eaton, A.D., Lenore, C.S., and Greenberg, A.E. (Eds.). (1995). *Standard Methods for the Examination of Water and Wastewater*. American Public Health Association, Washington, DC, pp. 4-75–4-96.
11. Campbell, J.L. et al. (2000). Dissolved organic nitrogen budgets for upland, forested ecosystems in New England. *Biogeochemistry* **49**: 123–142.
12. Peterson, B.J. et al. (2001). Control of nitrogen export from watersheds by headwater streams. *Science* **292**: 86–90.
13. Mueller, D.K., Hamilton, P.A., Helsel, D.R., Hitt, K.J., and Ruddy, B.C. (1995). Nutrients in ground water and surface water of the United States—an analysis of data through 1992. U.S. Geological Survey Water-Resources Investigations Report 95-4031, p. 74.
14. U.S. EPA. (1999). *1999 Update of Ambient Water Quality Criteria for Ammonia. U.S. Environmental Protection Agency*. National Service Center for Environmental Publications, Cincinnati, OH.
15. Bogardi, I. and Kuzelka, R.D. (Eds.). (1991). *Nitrate Contamination: Exposure, Consequence and Control*. NATO ASI Series. Springer-Verlag, New York.
16. Follett, R.F. and Walker, D.J. (1989). Ground water quality concerns about nitrogen. In: *Nitrogen Management and Ground Water Protection*. R.F. Follett (Ed.). Elsevier, New York, pp. 1–22.
17. Steele, K.F. (Ed.). (1995). *Animal Waste and the Land–Water Interface*. CRC Press, Lewis, Boca Raton, FL.
18. Steele, K.F. and McCalister, W.K. (1991). Potential nitrate pollution of ground water in limestone terrain by poultry litter, Ozark Region, USA. In: *Nitrate Contamination: Exposure, Consequence and Control*. I. Bogardi and R.D. Kuzelka (Eds.). NATO ASI Series, Springer-Verlag, New York, pp. 209–218.
19. Spalding, R. and Exner, M. (1993). Occurrence of nitrate in groundwater—A review. *J. Environ. Qual.* **22**: 392–402.
20. Paerl, H.W., Pinckney, J.L., Fear, J.M., and Peierls, B.L. (1998). Ecosystem responses to internal and watershed organic matter loading: Consequences for hypoxia in the eutrophying Neuse River estuary, North Carolina, USA. *Mar. Ecol. Prog. Ser.* **166**: 17–25.
21. Clark, I.D. and Fritz, P. (1997). *Environmental Isotopes in Hydrogeology*. Lewis, CRC Press, Boca Raton, FL, pp. 150–152.
22. Ostrom, N.E., Peggy, H., and Ostrom. (1999). Nitrogen isotopes. In: *Encyclopedia of Geochemistry*. C.P. Marshall and R.W. Fairbridge (Eds.). Kluwer Academic, Dordrecht, The Netherlands, pp. 431–433.
23. Karr, J.D., Showers, Hinson, W.J., and Thomas, H. (2002). Nitrate source identification using d ^{15}N in a ground water plume near an intensive swine operation. *Ground Water Monitoring and Remediation* 68–75.
24. Aravena, R., Evans, M.L., and Cherry, J.A. (1996). Stable isotopes of oxygen and nitrogen in source identification of nitrate from septic systems. *Ground Water* **31**: 180–186.

OSMOSIS-DIFFUSION OF SOLVENT OR CAUSED BY DIFFUSION OF SOLUTES?

Karl Erik Zachariassen
Norwegian University of
Science and Technology
Trondheim, Norway

Osmosis is the movement of water through a semipermeable membrane separating two solutions to the solution with the higher solute concentration. The membrane is semipermeable in the sense that it is permeable to water, but not to the solutes. The osmotic movement of water can be blocked by the application on the high concentration side of a hydrostatic pressure, which counteracts the so-called osmotic pressure that moves the water through the membrane.

The prevailing theory ascribes osmosis to a diffusion of water molecules through the membrane down a water concentration gradient across the membrane. The solutes are supposed to dilute the water, and the dilution will be strongest at the side with the highest solute concentration.

However, several investigators claim that this explanation fails to explain important characteristics of osmotic phenomena. Osmotic movement of water is considerably faster than should be expected from a diffusive movement of water. In addition, the water concentration of solutions of different solutes are not correlated with the osmotic activities of the solutions. Furthermore, the osmotic pressure is related to the total concentrations of dissolved particles, and van t'Hoff pointed out that the number of solute particles and the solution volume are related to the osmotic pressure in the same way as the number of molecules, pressure, and volume of a gas, i.e., osmotic phenomena, follow the universal gas laws. No theory has been proposed to explain the agreement between osmotic phenomena and gas laws on the basis of diffusion of water down a water concentration gradient.

In 1903, George Hulett (1) published a theory that explains the agreement between osmotic phenomena and gas laws and that is in agreement with the other observations on osmotic phenomena. Hulett emphasized that the solute particles diffuse inside solutions in the same way as gas molecules diffuse within the space in which the gas is distributed. In the same manner, as gas molecules create a pressure when they collide with the gas boundary, solute particles create a pressure when they collide with the boundary of the solution. Both for gases and solutions this pressure will tend to cause the volume to expand, which is easy for a gas if the boundaries can yield to the pressure and allow expansion of the volume. However, the water in which the solutes of a solution are distributed is essentially incompressible, and the expansive pressure created by the diffusing solutes cannot lead to an expansion unless more water can enter the solution and increase the solution volume. The semipermeable membrane of osmotic processes allows such an influx of water from a solution where a lower solute concentration creates a lower expanding pressure, whereas the solutes that create the pressure are prevented from passing. Water will subsequently move across the membrane until the pressures on both sides are equal.

According to the gas laws, an ideal 1 molar solution (1 mole of particles dissolved in a volume of 1 L), which is separated from pure water by a semipermeable, will create a pressure of 22.1 atm, which corresponds to a water column of impressive 224 m. It will have a melting point of −1.86 °C, and a solution with this osmotic concentration is said to be a 1 osmolar solution.

Proportionality exists among the osmolality, the osmotic pressure, and the melting point depression. This proportional relationship with osmotic pressure is also displayed by vapor pressure and boiling point elevation. Because of the exact proportionality that exists among these parameters, they are frequently referred to as the colligative parameters (from the Latin *colligare* = bind together).

A closer inspection of the relationship between mole-based concentration measures and the colligative properties reveals that it is the molal concentration that is most closely related to the colligative parameters. For this reason, some investigators prefer the term *osmolality*.

UNITS FOR OSMOTIC CONCENTRATIONS

Some confusion exists about the units for osmotic concentration. The officially approved unit for osmolality is Osmoles/kg water. Some investigators find that this unit is misleading, because it invites us to consider an Osmole as a defined amount of solute, which when diluted in 1 kg of water, gives an osmolality of 1 Osmole/kg water. Accordingly, when diluted in 2 kg of water, the same osmole of solutes should give an osmolality of 0.5 Osmole/kg water, but this is never the case. An Osmole is not a defined amount of solute, and the concept of osmolality has a meaning only for solutions. For this reason, some investigators use the unit Osm, where the letter *m* emphasizes the close relationship between osmotic and molal concentrations, which have *m* as the official unit.

To add to the confusion, the official unit for osmolality has led some authors to introduce Osmoles/L solution as a unit for osmolarity. However, the melting point is claimed to be −1.86 °C both for a 1 osmolal and for a 1 osmolar solution. Hence, the distinction between units for osmolality and osmolarity is merely semantic and of no practical consequence. Osmolality and osmolarity are different terms for the same thing, and furthermore, they are just quasi-measures for concentration. Solutions of different solutes with the same molar or molal concentrations all have different osmolalities, and for a given type of solute, no exact proportionality exists between molal concentrations on one side and colligative parameters on the other, which is at least in part because the amount of hydration water associated with each solute particle differs substantially from substance to substance, and for dissolved salts, the degree of dissociation varies from substance to substance and with their concentration.

Although osmolality is not a precise chemical concentration measure, it is a useful tool for comparing total solute concentrations of fluids in physiology, and it is a useful measure to control the chemical solute concentrations of body fluids and solutions.

BIBLIOGRAPHY

1. Hulett, G. (1903). Beziehung zwischen negativem Druck und osmootischen Druck. *Z. Phys. Chem.* **42**: 353–368.

Hammel, H.T. and Scholander, P.F. (1976). *Osmosis and Tensile Solvent*. Springer Verlag, Berlin.

PARTITIONING AND BIOAVAILABILITY

Jim Philp
Fiona Stainsby
Sandra Dunbar
Napier University
Edinburgh, Scotland, United Kingdom

INTRODUCTION

As soon as a chemical is spilled it becomes a pollutant, but the fate of that pollutant depends to a large degree on its chemistry, which determines how it will partition to various “compartments,” or phases, of the environment. This idea is enshrined in risk-based strategies for assessing contaminated land and water. In many cases a contaminant spill on land quickly becomes a water pollution issue, be it surface or groundwater. As many countries are heavily reliant on groundwater as a source of potable water, investigating the fate of pollutants is an essential prerequisite to any remedial action plan for contaminant cleanup.

Partitioning is governed by some fundamental properties of chemicals, but partitioning should be viewed in relation to the toxicity of the chemical in question and its degradability in the various environmental compartments. In this article, the important properties of chemicals that relate to partitioning are examined, and the theory is reinforced by example. The examples chosen are benzene, phenol, and some chlorinated phenols. All of these compounds are important industrial chemicals.

Bioavailability is a simple concept, but practically it is a complex issue that remains to be resolved. The bioavailability concept is also enshrined in the risk assessment approach to contaminated land and water. If a pollutant is present in soil or water in high concentration, but is not available to the biota, then surely it presents minimal risk. Regardless of the apparent simplicity of the concept, there is great urgency to find methods for quantifying bioavailability. Some approaches to solving this quandary are discussed.

PROPERTIES THAT DETERMINE PARTITIONING

The compartments of the environment relevant to partitioning are air, water, soil, and sediment and, through bioaccumulation/magnification through food chains, biota. Partitioning into the biota is beyond the scope of this article, and it will deal principally with partitioning to nonliving compartments. Figure 1 represents a common scenario in environmental remediation, contamination from old gas stations.

Vapor Pressure

This is the partial pressure of a chemical in a gas phase that is in equilibrium with the pure liquid or solid phase of the chemical. Vapor pressure is greatly influenced by temperature, and figures are usually quoted at a constant temperature (usually 20 °C or 25 °C). If the volume of the containing vessel is also known, the vapor pressure can be used to calculate the concentration of the chemical in the headspace of the vessel from the ideal gas law:

$$\frac{n}{V} = \frac{P}{RT} \tag{1}$$

where P = vapor pressure (atm)
V = volume
n = number of moles of the chemical
R = gas constant
T = absolute temperature

Another way to think of vapor pressure is that it is the pressure that a vapor exerts on the atmosphere: thus it will be clear that it is the driving force for volatilization. Vapor pressures of organic compounds are often compared to the vapor pressure of mercury (atmospheric pressure is 760 mmHg). Thus the vapor pressure of benzene is 87.2 mmHg at 25 °C. Volatilization potential can be misleading in the absence of information on the water solubility of the compound, however. A compound that has a low vapor pressure may still have a high tendency to "escape" if the water solubility is low. For example in Table 1, 2,4,6-trichlorophenol has a much lower vapor pressure than 3-chlorophenol, and yet its tendency to escape to air is higher. This can be explained when the water solubilities are compared: the much lower value for 2,4,6-trichlorophenol results in it having a Henry's constant about a magnitude order higher than 3-chlorophenol, even though its vapor pressure is much lower.

Henry's Law Constant

To study partition phenomena, then, Henry's law constant is more valuable than vapor pressure. Henry's law constant is the ratio of the equilibrium concentration of a compound in air and its equilibrium concentration in water. It can therefore be regarded as the partition coefficient between air and water, and is the best indicator of the tendency of a chemical to volatilize from water. Henry's law constant can be written as:

$$H = \frac{P}{S} \tag{2}$$

where H = Henry's law constant (atm-m^3/mol)
P = vapor pressure (atm)
S = water solubility (mol/m^3)

Water Solubility

This is of central importance to understanding partition phenomena in the environment. It is the maximum, or

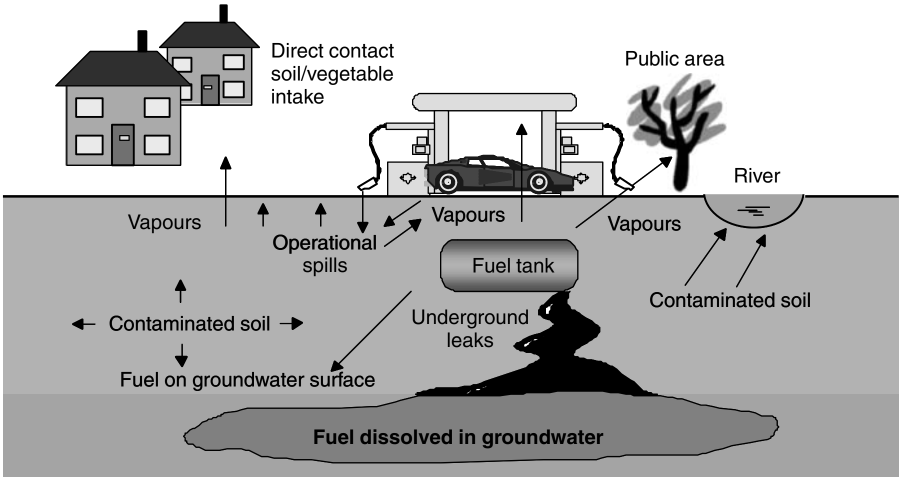

Figure 1. Partitioning of fuel as a contaminant from a gas station that has a leaking underground storage tank. Because fuels are multicomponent liquids, different components partition to different compartments of the environment.

Table 1. Important Partition Variables for a Range of Compounds. The Figures were Computed from U.S. EPA Software (EPI Suite v.3.11 and ECOSAR)

	Vapor Pressure, mmHg, 25 °C	Henry's Constant	Water Solubility, mg/L	Log K_{ow}	Log K_{oc}	Level III Fugacity Mass Amount, %	Toxicity Fish 96-h LC_{50}, mg/L	Biodeg. Probability, MITI Non-Linear
Benzene	87.2	5.55×10^{-3}	1790	2.13	2.22	Air: 37.6 Water: 48.4 Soil: 13.8 Sediment: 0.20	59.2	.73
Phenol	0.323	3.33×10^{-7}	77, 900	1.46	2.43	Air: 1.33 Water: 47.3 Soil: 51.3 Sediment: 0.098	27.7	.71
3-chloro-phenol	0.119	3.45×10^{-7}	26, 000	2.50	2.64	Air: 1.35 Water: 39.4 Soil: 59.1 Sediment: 0.21	15	0.39
2,4,6-trichloro-phenol	0.00536	2.6×10^{-6}	800	3.69	3.07	Air: 2.55 Water: 22.9 Soil: 72.4 Sediment: 2.17	3.7	0.043
Pentachloro-phenol	1.08×10^{-5}	2.45×10^{-8}	14	5.12	3.53	Air: 0.0665 Water: 5.59 Soil: 70.9 Sediment: 23.4	0.8	.0031

saturation, concentration of a compound that dissolves in water. It also is significantly affected by temperature. As the water solubility of compounds is defined by a host of factors, comparisons should always be made at a constant temperature, usually 25 °C. In natural water systems, the concentration of pollutant chemicals is usually rather low, making this especially important.

Polarity refers to the extent to which charge is unevenly distributed within a molecule. Water itself is highly polar. Therefore charged or highly polar organic and inorganic molecules, hydrophilic molecules, are readily soluble in water. Conversely, nonpolar, hydrophobic molecules find solution in water a less energetically favorable state, and they have low water solubility. In the examples in Table 1, benzene has moderate water solubility, but the addition of the polar hydroxyl group makes phenol much more water soluble. The solubility of any solute in any solvent is mediated by forces of attraction; the main ones are:

- van der Waals forces;
- hydrogen bonding;
- dipole–dipole interactions.

The size and shape of a molecule is also critical in determining its overall water solubility. Large molecules tend to have lower water solubility because they have a higher molar volume. The addition of halogens to an organic molecule might be expected to increase the water solubility of the new molecule as halogens are electronegative (electron-withdrawing) species and would thus tend to increase the polarity of the molecule. However, in the increasingly chlorinated series of phenols in Table 1, this is not the case: increasing chlorination very markedly decreases the solubility of chlorophenols because the halogens increase the molecular volume of organic compounds (1).

Other chemical functional groups affect water solubility in rather specific ways. The presence of polar functional groups tends to increase the water solubility of molecules compared to unsubstituted hydrocarbons: alcohols, amines, ethers, ketones, and organic acids are good examples. Even with these examples, though, a general rule of increasing molecular size and decreasing water solubility still exists. The short-chain fatty acid acetic acid is much more water soluble than a long-chain fatty acid. The additions to the molecule are not more polar functional groups but simply more nonpolar $C–H_2$ groups that make the molecule larger but less polar.

The situation is further complicated by more specific chemical reactions. The chlorophenols in Table 1 illustrate this point. Chlorophenols behave as weak acids when dissolved in water due to the loss of a proton from the hydroxyl group to leave a phenoxide ion. The more chlorines that are added to the ring, the more electrons are withdrawn from the electron-rich phenoxide ion. This lowers the phenoxide negative charge and therefore lowers its ability to hold a proton. Thus the more chlorines on the ring, the more easily a proton is released, and the greater the acidity. The ionized form is usually several orders of magnitude more water soluble than the neutral un-ionized form. The relative proportion of each species is highly dependent on pH.

Log K_{ow}, a Measure of Hydrophobicity

Polarity and water solubility do not give direct measures of the hydrophobicity of a compound. A hydrophobic compound might be expected to dissolve in a lipophilic phase, and this is the basis for the octanol–water partition coefficient, K_{ow}. This is the ratio of a chemical's concentration, at equilibrium, when dissolved in octanol and in water, when both phases are saturated with each other.

This test was devised by the pharmaceutical industry to study partitioning of drugs as it roughly equates to the partitioning between water and body fat. It has proven very useful in predicting how chemicals partition in the environment. Large molecules of low polarity have a high K_{ow} and are more likely to partition to solids in the environment. Small, polar molecules dissolve in water rather than octanol, have a low K_{ow}, are much less likely to partition to solids in the environment, and so are more likely to remain in an aqueous phase:

$$K_{ow} = \frac{\begin{array}{c}\text{concentration of pollutant in}\\ \text{water-saturated octanol (mg/L)}\end{array}}{\begin{array}{c}\text{concentration of pollutant}\\ \text{in octanol-saturated water (mg/L)}\end{array}} \qquad (3)$$

The range of values for K_{ow} is very wide, and tabulated values are more often expressed as log K_{ow}. In the examples in Table 1, it can be seen that, as the molecule becomes larger through increasing chlorination, the water solubility decreases and log K_{ow} increases.

Log K_{oc}, Partitioning to Soil

The most basic statement of the composition of soil is that it contains sand, clay, organic matter, water, and air. Although typically the organic matter content of soil is in the range of 0.1 to 5%, this is where organic pollutants sorb, not to the soil as a whole. The soil adsorption coefficient, K_{oc}, takes account of this. For a liquid organic pollutant,

$$K_{oc}\ (\text{mL/g}) = \frac{\begin{array}{c}\text{mass of pollutant sorbed}\\ \text{to soil organic matter (mg/g)}\end{array}}{\begin{array}{c}\text{mass of pollutant in}\\ \text{the aqueous phase (mg/mL)}\end{array}} \qquad (4)$$

The range of values for K_{oc} is very wide, and therefore the use of log K_{oc} is preferred. The soil partition coefficient is greater, the more hydrophobic the pollutant. However, it has been measured for far fewer chemicals than K_{ow}. It might be expected to be a more realistic measure than K_{ow}, but as data are relatively rare for K_{oc}, a series of correlation equations has been developed that relate K_{oc} to K_{ow}. The increase in log K_{ow} seen in Table 1 as chlorination increases is mirrored by increasing log K_{oc}.

Fugacity

The term fugacity literally means the "tendency to flee," and refers to a molecule's driving force to escape from the compartment in which it currently exists. With knowledge of all of a chemical's partition coefficients, it is possible to predict its relative concentration in air, water, and soil at equilibrium. The gradient of fugacity between two compartments (e.g., soil and soil porewater) determines the potential for the chemical to move from one compartment to the other. Fugacity therefore has units of pressure, but this can be related to the chemical concentration through modeling, and thus the distribution through various compartments predicted (Table 1).

THE RELATIONSHIP TO TOXICITY AND BIODEGRADATION

Hydrophobicity and Toxicity

Particularly K_{ow} and log K_{ow} have become increasingly used as predictors of the behavior of pollutants in the environment. It has been correlated with toxicity. The most common mode of action regarding the toxicity of industrial pollutants is probably narcosis. Nonpolar narcosis results from the perturbation of cellular membranes (2) due to the entry of hydrophobic (lipid-soluble) molecules into the phospholipid bilayer of the membrane. Narcosis has been shown to be a nonspecific physiological effect which is independent of chemical structure, and log K_{ow} is highly correlated with nonpolar narcosis toxicity (3).

Phenols have greater toxicity than predicted by nonpolar narcosis (4). They act by polar narcosis, which probably results from a strong hydrogen-bonding group that makes the compound more polar. Nevertheless, there is still a strong correlation between the toxicity of these compounds and log K_{ow}. Pentachlorophenol has greater toxicity as a result of a further mechanism: its acidity makes it a weak acid respiratory uncoupler (WARU). Terada (5) noted that WARUs possess an acid-dissociable moiety (e.g., a hydroxyl group), a bulky hydrophobic moiety, such as an aromatic ring, and multiple electronegative groups, such as halogens. During weak acid uncoupling, ATP synthesis is inhibited with no effect on the respiratory chain (6) because electron transport and phosphorylation reactions are uncoupled from each other.

Hydrophobicity and Biodegradability

It is believed that only molecules of hydrocarbons that are dissolved in the aqueous phase are available for intracellular metabolism (7). The rate at which a particular organic compound dissolves in water is critical to both the toxicity and biodegradability of the compound as this governs the rate of transfer to the organism (in biodegradability, microorganisms are important). The rate of transfer is determined by the equilibrium and actual concentration in the bulk phase and the aqueous phase. This is central to the concept of bioavailability.

From the compounds in Table 1, the increasing degree of chlorination decreases the rate of biodegradation. Increasing chlorination also leads to a marked drop in water solubility and a corresponding increase in hydrophobicity, which influence biodegradability. The toxicity of the compound also adversely affects biodegradation. The interaction between toxicity and biodegradation is not well understood, but they are clearly linked and competing processes: at low concentrations, a compound may be biodegradable, but this is arrested at higher concentrations by toxicity. If the concentration is very low for biodegradation, then this may not be enough to support microbial growth and reproduction, the concept of threshold elaborated in detail by Alexander (8).

BIOAVAILABILITY

Bioavailability is a concept recently borrowed from nutrition to explain the behavior of chemical pollutants in the environment (9). There are several definitions in circulation. In its simplest form (10), bioavailability is defined as "the amount of contaminant present that can be readily taken up by microorganisms and degraded." But this ignores bioavailability to higher organisms. Another definition that takes account of this and alludes to the risk assessment practice (11) is "a measure of the potential of a chemical for entry into ecological or human receptors."

The bioavailability quandary mentioned previously has been accurately articulated by the International Union of Pure and Applied Chemistry (IUPAC).

> "In the IUPAC Conference of Pesticide Chemistry in London in 1998, there was a controversial discussion on the topic whether agrochemicals in the environment are really hazardous. The conventional wisdom is that chemicals in soil are available to microorganisms, plant roots, and soil fauna like earthworms and animals via dermal exposure. Then bioaccumulation through the food web may induce indirect exposure to higher organisms. National governments are simply reducing the threshold levels of chemicals in the environments for larger safety margins in their guidelines. However, the question raised at the Congress was that chemical residues in the environment are not always bioavailable, so that *the actual exposure of biota to chemicals is rather different from the amount (concentration) present in the environment* (authors' italics). In addition, the persistence and efficacy of agrochemicals are affected by their bioavailability in soil. There is a lack of comprehensive understanding of the bioavailability of chemicals with different chemical characteristics, such as non-polar organochlorines, polar agrochemicals and etc. Thus, there is a strong need to clarify the scientific basis for bioavailability: definition, estimation methods and affecting factors."

Measuring Bioavailability

Measuring or even estimating bioavailability is problematic as it is affected by many individual and interacting conditions relating to soil and water chemistry, pollutant chemistry and partitioning, and biological transformation and concentration factors (12). Moreover, the "aging" process brings about a time-dependent decrease in bioavailability as compounds become sequestered in soil over time by ill-understood mechanisms (13). The techniques currently being researched can be categorized broadly as direct or indirect and biological or chemical (14), although, as pointed out by these authors, only organisms can determine whether a chemical is bioavailable, so direct chemical methods are not possible.

Direct Biological Methods. These are techniques that measure the actual amount of a chemical taken up by a target organism, and this may ultimately be the most accurate measure of bioavailability, although it is by no means routine. For example, the determination of chemical levels in earthworms is relatively expensive and time-consuming. The great strength of these direct techniques, however, is that they integrate all the biotic and abiotic modifying factors of chemical bioavailability (14).

Indirect Biological Methods. It is possible to observe a response to a chemical in an organism without actually measuring the concentration of the chemical. A range of responses is possible, such as lethality, enzyme induction or inhibition, and reproductive effects. These are regarded as indirect biological methods because the effect may be quantifiable in the organism, but the chemical concentration remains unknown.

A promising, novel approach is the use of genetically modified microorganisms that can detect and quantify specific pollutants. This type of biosensor is based on the highly specific genetic control mechanisms used by microorganisms to ensure that specific proteins are expressed only when needed; for example, to detoxify a particular toxic substance. This control is exerted by inducible promoters, consisting of a specific DNA sequence upstream of the genes to be controlled and a DNA-binding protein that either activates or prevents transcription in response to the presence or absence of the target compound. Biosensors of this type are easily generated by fusing such a controllable promoter to a reporter gene that generates a detectable signal when the promoter is activated (15). The most popular reporter is the bacterial *lux* system that produces light (Fig. 2), because light can be easily quantified and, due to its rarity in biology, there is no interference from the background biochemistry of the host organism.

Such biosensors can be both highly specific and responsive to very low concentrations, and the microbial cell membrane makes this approach truly a measure of bioavailability. There are limitations to developing such biosensors for detecting organic pollutants as often the

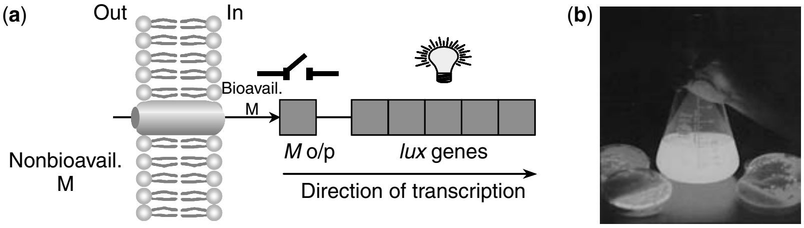

Figure 2. (**a**) Schematic of a genetically modified whole bacterial cell biosensor for detecting bioavailable metal. The metal may be expected to be transported into the bacterial cell via a specific transporter protein mediating uptake and efflux. If the metal is in a nonbioavailable form, for example, as a precipitate, it will not be transported. The interaction of the metal with the specific metal regulatory element initiates transcription resulting in synthesis of the luciferase enzyme and the production of light. (**b**) The naturally bioluminescent bacterium *Vibrio fischeri* producing light from colonies and liquid culture.

required specificity is lacking. For heavy metals, specificity may not be a problem. Nevertheless, careful selection of the host strain is required for a variety of reasons (16). Such biosensors offer the prospect of a much more routine measure of bioavailability.

Indirect Chemical Methods. Indirect chemical methods usually involve extracting a fraction of the chemical (metals or organics) from a soil; extractability is defined by the chemical itself, the nature of the extractant(s), and the experimental conditions applied. The origins of this approach are in sequential extraction procedures (17) that attempt to quantify the speciation of toxicants into those weakly bound and those strongly bound to the soil matrix. These tests often assume that the weakly bound toxicants are those that are more bioavailable, although evidence for this is incomplete. A recent advance in this technique is sequential accelerated solvent extraction (18) that reduces sample preparation time and maintains relatively constant extraction conditions. A technique under development is the use of selective supercritical fluid extraction (19).

A number of methods are based around the human gastrointestinal tract (the physiologically based extraction test, PBET; 20). To mimic human conditions, they incorporate gastric juices and enzymes, mix at 37 °C, and use soil residence times similar to those found in children after ingestion of food. Analysis of all solutions produced in the PBET extractions is completed by analytical chemistry techniques under matrix-matched conditions.

However sophisticated, any such technique is ultimately limited by the lack of an organism. There will be no universal chemical test for bioavailability as the phenomenon is organism- and even species-dependent (21).

BIBLIOGRAPHY

1. Yalkowsky, S.H. and Valvani, S.C. (1979). Solubilities and partitioning. 2. Relationships between aqueous solubilities, partition coefficients and molecular surface areas of rigid aromatic hydrocarbons. *J. Chem. Eng. Data* **24**: 127–129.
2. Cronin, M.T.D. and Schultz, T.W. (1998). Structure-toxicity relationships for three mechanisms of action of toxicity to *Vibrio fischeri*. *Ecotoxicol. Environ. Saf.* **39**: 65–69.
3. Cronin, M.T.D. and Dearden, J.C. (1995). Review: QSAR in toxicology. 1. Prediction of aquatic toxicity. *Quant. Struct. Act. Relationships* **14**: 1–7.
4. Cronin, M.T.D. and Schultz, T.W. (1997). Validation of *Vibrio fischeri* acute toxicity data: mechanism of action-based QSARs for non-polar narcotics and polar narcotic phenols. *Sci. Total Environ.* **204**: 75–88.
5. Terada, H. (1990). Uncouplers of oxidative phosphorylation. *Environ. Health Perspect.* **87**: 213–218.
6. Schultz, T.W. and Cronin, M.T.D. (1997). Quantitative structure-activity relationships for weak acid respiratory uncouplers to *Vibrio fischeri*. *Environ. Toxicol. Chem.* **16**: 357–360.
7. Sikkema, J., de Bont, J.A.M., and Poolman, B. (1995). Mechanisms of membrane toxicity of hydrocarbons. *Microbiolo. Rev.* **59**: 201–222.
8. Alexander, M. (1999). Threshold. In: *Biodegradation and Bioremediation*. Academic Press, San Diego, pp. 105–116.
9. Madsen, E.L. (2003). *Report on Bioavailability of Chemical Wastes with Respect to the Potential for Soil Bioremediation*. U.S. Environmental Protection Agency report *EPA/600/R-03/076*.
10. Maier, R. (2000). Bioavailability and its importance to bioremediation. In: *International Society for Environmental Biotechnology: Environmental Monitoring and Biodiagnostics*. J.J. Valdes (Ed.). Kluwer, Boston, pp. 59–78.
11. Loehr, R.C. and Webster, M.T. (1997). Effects of treatment on contaminant availability, mobility, and toxicity. In: *Environmentally Acceptable Endpoints in Soil*. D.G. Linz and D.V. Nakles (Eds.). American Academy of Environmental Engineers, Annapolis, MD, pp. 137–386.
12. Hund-Rinke, K. and Kördel, W. (2003). Underlying issues in bioaccessibility and bioavailability: experimental methods. *Ecotoxicol. Environ. Saf.* **56**: 52–62.
13. Chung, N. and Alexander, M. (2002). Effect of soil properties on bioavailability and extractability of phenanthrene and atrazine sequestered in soil. *Chemosphere* **48**: 109–115.
14. Lanno, R., Wells, J., Conder, J., Bradham, K., and Basta, N. (2004). The bioavailability of chemicals in soil for earthworms. *Ecotoxicol. Environ. Saf.* **57**: 39–47.
15. Philp, J.C., French, C., Wiles, S., Bell, J.M.L., Whiteley, A.S., and Bailey, M.J. (in Press). Wastewater toxicity assessment by whole cell biosensor. In: *Handbook of Environmental Chemistry*. D. Barceló (Ed.). Vol. 5. Water Pollution, Emerging Organic Pollutants in Wastewaters, Springer-Verlag, Berlin.
16. Rensing, C. and Maier, R.M. (2003). Issues underlying use of biosensors to measure metal bioavailability. *Ecotoxicol. Environ. Saf.* **56**: 140–147.
17. Tessier, A., Campbell, P.G.C., and Bisson, M. (1979). Sequential extraction procedure for the speciation of particulate trace metals. *Anal. Chem.* **51**: 844–851.
18. Tao, S., Guo, L.Q., Wang, X.J., Liu, W.X., Ju, T.Z., Dawson, R., Cao, J., Xu, F.L., and Li, B.G. (2004). Use of sequential ASE extraction to evaluate the bioavailability of DDT and its metabolites to wheat roots in soils with various organic carbon contents. *Sci. Total Environ.* **320**: 1–9.
19. Björklund, N.T., Bøward, S., Pilorz, K., Mathiasson, L., and Hawthorne, S.B. (2000). Introducing selective supercritical fluid extraction as a new tool for determining sorption/desorption behavior and bioavailability of persistent organic pollutants in sediment. *J. Biochem. Biophys. Methods* **43**: 295–311.
20. Ruby, M.V., Davis, A., Schoof, R., Eberle, S., and Sellstone, C.M. (1996). Estimation of lead and arsenic bioavailability using a physiologically based extraction test. *Environ. Sci. Technol.* **30**: 422–430.
21. Reid, B.J., Jones, K.C., and Semple, K.T. (2000). Bioavailability of persistent organic pollutants in soils and sediments—a perspective on mechanisms, consequences and assessment. *Environ. Pollut.* **108**: 103–112.

READING LIST

Hemond, H.F. and Fechner-Levy, E.J. (2000). *Chemical Fate and Transport in the Environment*. Academic Press, San Diego.

Landis, W.G. and Yu, M-H. (2003). *Introduction to Environmental Toxicology*, 3rd Edn. Lewis, Boca Raton, FL.

Nirmalakhandan, N.N. and Spreece, R.E. (1988). Prediction of aqueous solubility of organic compounds based on molecular structure. *Environ. Sci. Technol.* **22**: 328–338.

Watts, R.J. (1997). *Hazardous Wastes: Sources, Pathways, Receptors*. John Wiley & Sons, New York.

PHYSICAL PROPERTIES

KEVIN S. MCLEARY

STRUCTURE

Water is a molecule composed of two atoms of hydrogen bonded to an atom of oxygen. It has a volume of 2.97 × 10^{-11} μ^3 (29.7 Å^3) (1). As a result of the hybridization of the oxygen atom, the molecule is bent at an angle of 104.45°. Due to the nonlinear structure and to the difference in electronegativity between oxygen (3.44) and hydrogen (2.20), the molecule has a dipole moment of 1.85 Debye (2).

Water molecules are held together by intermolecular attractions. In water, the intermolecular attractions are hydrogen bonds. The negatively charged oxygen atom in one molecule of water is electrically attracted to the positively charged hydrogen atom of another molecule. Hydrogen bonding governs the structure of water up to the boiling point (2).

SPECIFIC WEIGHT (γ)

The specific weight of a substance is its weight per unit of volume. In U.S. Customary Units, it is often expressed as pounds-force per cubic foot (3). In SI units, it is often expressed in kN/m^3. Although a value of 62.4 pounds-force per cubic foot (9.81 kN/m^3) is commonly used for the majority of engineering purposes, the specific weight of water varies with temperature, from 62.42 pounds-force per cubic foot (9.805 kN/m^3) at 32 °F (0 °C) to 59.8 pounds-force per cubic foot (9.399 kN/m^3) at 212 °F (100 °C) (4). Table 1 shows the variability in the specific weight of water with temperature in U.S. Customary Units. Table 2 shows the variability in the specific weight of water with temperature in SI Units.

Specific weight is also a function of gravity, and it varies slightly with elevation. One cubic foot of water at 32 °F at sea level has a specific weight of 62.42 pounds-force per cubic foot (9.805 kN/m^3); the same volume of water at 10,000 feet above sea level has a specific weight of 62.36 pounds-force per cubic foot (9.796 kN/m^3). Because the earth is not a perfect sphere, but bulges at the equator, specific weight is also a weak function of latitude. However, these variations are small enough to warrant the use of a constant value of 62.4 pounds-force per cubic foot (9.81 kN/m^3) at all elevations and latitudes (4).

DENSITY (ρ)

The density of a fluid is defined either as its mass per unit of volume or as the ratio of its specific weight to the acceleration due to gravity. Density ranks among the most important hydraulic properties of water (1). It is a factor in the settling velocity of particles and in the filtration of water through a porous medium. A fluid's Reynolds number, the most important dimensionless number in hydraulics, is a direct function of density. The Reynolds number is used in calculating friction losses, pipe sizing, and pumping requirements

In U.S. Customary Units, density is expressed in slugs/ft^3, and in SI units, it is expressed in kg/m^3. Like specific weight, density varies with temperature. The density of pure water varies from 1.940 slugs/ft^3 (999.8 kg/m^3) at 32 °F (0 °C) to 1.860 slugs/ft^3 (958.4 kg/m^3) at 212 °F (100 °C). Tables 1 and 2 show the variation of density in U.S. Customary Units and SI units, respectively.

SPECIFIC GRAVITY

Specific gravity is defined as the ratio of the density of one fluid at a specified temperature to that of a reference fluid at a reference temperature. For liquids, the reference liquid is usually water, although varying reference temperatures have been used, for example, 39° and 68 °F (4° and 20 °C). Although specification of the reference temperature is required to ensure accuracy, it is often not given because the density of water varies by only 0.2% over the usual temperature range (5). The specific gravity of water is usually taken as 1.0.

ABSOLUTE (DYNAMIC) VISCOSITY (μ)

The viscosity of a fluid is expressed as a measure of the resistance to flow. Mathematically, it is defined as the ratio of the shear stress of a fluid to its velocity gradient or as the proportionality constant between the shear stress and the velocity gradient:

$$\frac{F}{A} = \mu \frac{dv}{dy} \tag{1}$$

Viscosity can be measured by taking two plates of area A, which are separated by a fluid whose thickness is y. The bottom plate is held stationary. The top plate is moved at a constant velocity v. The force required to maintain the constant velocity is measured and used to calculate the viscosity μ of the fluid.

The chief cause of viscosity in fluids is molecular cohesion. Molecular cohesion decreases as temperature is increased, therefore, viscosity also decreases as temperature increases. Changes in pressure do not affect viscosity (5). Impurities also exert little influence on viscosity (1). Tables 1 and 2 show the effect of temperature on viscosity.

KINEMATIC VISCOSITY (ν)

There are many times in fluid mechanics and hydraulics when the fluid's absolute viscosity is divided by its density. To ease calculations, this term has been defined as kinematic viscosity because it has kinematic units of length and time, ft^2/sec or m^2/s.

Both absolute viscosity and density are a function of temperature, so kinematic viscosity is also a function of temperature. Tables 1 and 2 show how kinematic viscosity varies with temperature.

Table 1. Physical Properties of Water in U.S. Customary Units[a]

Temp., °F	Specific Weight, γ, $\mathrm{lb_f/ft^3}$	Density, ρ, $\mathrm{slug/ft^3}$	Modulus of Elasticity[b] $E/10^3$, $\mathrm{lb_f/in.^2}$	Dynamic Viscosity, $\mu \times 10^5$, $\mathrm{lb_f \cdot sec/ft^2}$	Kinematic Viscosity, $\nu \times 10^5$, $\mathrm{ft^2/sec}$	Surface Tension,[c] σ, $\mathrm{lb_f/ft}$	Vapor Pressure, p_v, $\mathrm{lb_f/in.^2}$
32	62.42	1.940	287	3.746	1.931	0.00518	0.09
40	62.43	1.940	296	3.229	1.664	0.00614	0.12
50	62.41	1.940	305	2.735	1.410	0.00509	0.18
60	62.37	1.938	313	2.359	1.217	0.00504	0.26
70	62.30	1.936	319	2.050	1.059	0.00498	0.36
80	62.22	1.934	324	1.799	0.930	0.00492	0.51
90	62.11	1.931	328	1.595	0.826	0.00486	0.70
100	62.00	1.927	331	1.424	0.739	0.00480	0.95
110	61.86	1.923	332	1.284	0.667	0.00473	1.27
120	61.71	1.918	332	1.168	0.609	0.00467	1.69
130	61.55	1.913	331	1.069	0.558	0.00460	2.22
140	61.38	1.908	330	0.981	0.514	0.00454	2.89
150	61.20	1.902	328	0.905	0.476	0.00447	3.72
160	61.00	1.896	326	0.838	0.442	0.00441	4.74
170	60.80	1.890	322	0.780	0.413	0.00434	5.99
180	60.58	1.883	318	0.726	0.385	0.00427	7.51
190	60.36	1.876	313	0.678	0.362	0.00420	9.34
200	60.12	1.868	308	0.637	0.341	0.00413	11.52
212	59.83	1.860	300	0.593	0.319	0.00404	14.70

[a] References 3 and 7.
[b] The modulus of elasticity is given at atmospheric pressure.
[c] Surface tension is given for water in contact with air.

Table 2. Physical Properties of Water in SI Units[a]

Temp., °C	Specific Weight, γ, $\mathrm{kN/m^3}$	Density, ρ, $\mathrm{kg/m^3}$	Modulus of Elasticity,[b] $E/10^6$, $\mathrm{kN/m^2}$	Dynamic Viscosity, $\mu \times 10^3$, $\mathrm{N{\cdot}s/m^2}$	Kinematic Viscosity, $\nu \times 10^6$, $\mathrm{m^2/s}$	Surface Tension,[c] σ, N/m	Vapor Pressure, p_v, $\mathrm{kN/m^2}$
0	9.805	999.8	1.98	1.781	1.785	0.0765	0.61
5	9.807	1000.0	2.05	1.518	1.519	0.0749	0.87
10	9.804	999.7	2.10	1.307	1.306	0.0742	1.23
15	9.798	999.1	2.15	1.139	1.139	0.0735	1.70
20	9.789	998.2	2.17	1.002	1.003	0.0728	2.34
25	9.777	997.0	2.22	0.890	0.893	0.0720	3.17
30	9.764	995.7	2.25	0.798	0.800	0.0712	4.24
40	9.730	992.2	2.28	0.653	0.658	0.0696	7.38
50	9.689	988.0	2.29	0.547	0.553	0.0679	12.33
60	9.642	983.2	2.28	0.466	0.474	0.0662	19.92
70	9.589	977.8	2.25	0.404	0.413	0.0644	31.16
80	9.530	971.8	2.20	0.354	0.364	0.0626	47.34
90	9.466	965.3	2.14	0.315	0.326	0.0608	70.10
100	9.399	958.4	2.07	0.282	0.294	0.0589	101.33

[a] References 3 and 7.
[b] The modulus of elasticity is given at atmospheric pressure.
[c] Surface tension is given for water in contact with air.

SURFACE TENSION (σ)

Surface tension is caused by the attractive forces between molecules in a liquid. In a bulk liquid, each molecule is attracted to its neighboring molecules, and the total force on the molecule is zero. However, a molecule at the surface experiences an attraction to other molecules deeper in the liquid, but there are no molecules outside to balance these attractions. The molecule at the surface experiences a net attractive force toward the bulk liquid.

Surface tension is the amount of work necessary to increase the surface area of the liquid by one unit. It has units of force per unit area ($\mathrm{ft\text{-}lb_f/ft^2}$, or $\mathrm{lb_f/ft}$) and decreases as the temperature of the liquid increases. It is also a function of the gas contacting the liquid (5). Table 1 gives the variation of surface tension of water in contact with air as a function of temperature in U.S. Customary Units, and Table 2 shows the variation in SI units.

VAPOR PRESSURE (P_V)

Water is a collection of a large number of molecules in a liquid phase that move about at random. The molecules collide with each other and rebound in different directions. Each molecule moves with a certain velocity (energy). Molecules on the surface of a liquid moving away from the bulk liquid have no other molecule to collide with. If such a molecule has sufficient energy to escape the intermolecular attractions, it may escape the liquid into the vapor space above the liquid. It evaporates. Evaporated water molecules also move about at random. Some of the water vapor condenses into the liquid water phase.

If the water were placed in a closed container at a constant temperature and the air evacuated from the vapor space above the liquid, one could measure the pressure exerted by the water vapor. The pressure would increase until equilibrium is attained—when the rate of molecules evaporating and the rate of molecules condensing are equal. The pressure of the water vapor then is constant. This is the vapor pressure of the water for the given temperature. Table 1 gives the vapor pressure of water in psi. Table 2 gives the vapor pressure in kN/m^2.

MODULUS OF ELASTICITY (*E*)

Although water is usually considered incompressible, it is slightly compressible. The modulus of elasticity is defined as the increase in pressure that results in a specific percentage change in volume:

$$E = \frac{\Delta p}{\Delta V/V} \tag{2}$$

where Δp is the increase in pressure, ΔV is the resulting change in volume, and V is the original volume of water.

The modulus of elasticity has units of pressure and is a function of temperature. Tables 1 and 2 show the variation of the modulus of elasticity with temperature.

SPEED OF SOUND (*C*)

Sound is the propagation of a pressure wave through a material and is dependent on the material's modulus of elasticity and density. The speed of sound, c, in a liquid is given by

$$c = \sqrt{\frac{Eg_c}{\rho}} \tag{3}$$

The speed of sound in water is important in calculating the pressure increase due to water hammer. When the velocity in a pipe is decreased suddenly, the pipe experiences an increase in pressure. The increase in pressure, Δp, caused by a decrease in velocity of Δv, is

$$\Delta p = \frac{\rho c \Delta v}{g_c} \tag{4}$$

HEAT CAPACITY (C_P)

Heat capacity is defined as the amount of heat needed to raise the temperature of a given amount of substance by 1°. It is usually expressed in units of heat per unit of substance per degree. The term "heat capacity" is often used for this quantity of heat on a molar basis. If a mass basis is used, the term "specific heat" is frequently used.

The heat capacity of water is the basis of the unit of heat in both the English and SI systems. Therefore, the heat capacity (specific heat) of water is 1 Btu/pound-mass/°F in the English system and 1 cal/g/°C in the SI system.

Heat capacities of liquids are not functions of temperature, except at high reduced temperatures. (A reduced temperature is a thermodynamic quantity relating the temperature of a material to its critical temperature. The critical temperature for water is 647.1 K.) (6). At reduced temperatures exceeding 0.7–0.8, liquid heat capacities are large and are strong functions of temperature.

BIBLIOGRAPHY

1. Fair, G.M., Geyer, J.C., and Okun, D.A. *Water and Wastewater Engineering*. John Wiley & Sons, New York.
2. Atkins, P.W. *Physical Chemistry*, 2nd Edn. W.H. Freeman, San Francisco.
3. Metcalf & Eddy, Inc. *Wastewater Engineering: Treatment, Disposal, and Reuse*, 3rd Edn. McGraw-Hill, New York.
4. F.S. Merritt (Ed.). *Standard Handbook for Civil Engineers*, 3rd Edn. McGraw-Hill, New York.
5. Lindeburg. *Civil Engineering Reference Manual*.
6. Smith, J.M. and van Ness, H.C. *Introduction to Chemical Engineering Thermodynamics*, 3rd Edn. McGraw-Hill, New York.
7. Culp/Wesner/Culp. *Handbook of Public Water Systems*. Van Nostrand-Reinhold, New York.

ENVIRONMENTAL PHOTOCHEMISTRY IN SURFACE WATERS

Mark A. Young
University of Iowa
Iowa City, Iowa

INTRODUCTION

Surface waters make up a large fraction of the available terrestrial surface area, accounting for over 70% of the total area with most of that consisting of ocean waters. A correspondingly large fraction of the incoming solar flux will impinge on natural waters, penetrating to a depth determined by the scattering and absorption properties of the water system. While most of the short wavelength solar radiation is attenuated in the atmosphere, mainly by absorbing species such as O_2 and O_3, the surface flux of energetic photons in the near-UV and visible remains appreciable. While water itself is not photoactive, the irradiated volume of typical surface waters contains a plethora of potential chromophores. The chromophoric material consists of dissolved species and suspended particulate matter that can have strong electronic absorptions that significantly overlap with the available solar flux. Absorption of solar photons can initiate a rich variety of photochemical processes with sufficient kinetic rates to manifest a substantial impact on the chemistry taking place in the aqueous environment. Photochemically driven chemistry, in turn, can have enormous implications for global geochemical cycles, climate, and biogenic processes, making photochemistry a critical consideration in the study of environmental processes.

BACKGROUND

Natural surface waters can be either saltwater (marine, or ocean, waters) or freshwater (continental waters). Within each of these reservoirs, widely varying environments are found: remote and coastal areas in the oceans, and rivers, lakes, streams, bogs/marshes on the continents, as well as interfacial environments, such as estuaries. In addition to differences in salinity and pH, there will be variations in the types and concentrations of chromophoric species residing in each reservoir. Potential chromophores include dissolved organic material (DOM), inorganic NO_x^- ions, metal complexes, anthropogenic pollutants, and colloidal matter consisting of crustal minerals and particulate organic material (POM). Organic chromophores may also adsorb on, or complex with, the surface of mineral colloids and participate in charge transfer processes. Likewise, metals can become incorporated into organic material. The distinction between dissolved and colloidal material is not sharp and often depends on the analytical procedure used to assay concentrations. The relative importance of the various chromophores will depend on the details of the aqueous environment. Chromophores can absorb solar photons and undergo photochemical transformation while yielding a variety of primary and secondary reactive species, as depicted schematically in Fig. 1. The photogenerated intermediates can, in turn, react with other species present in the water at rates that are often diffusion controlled to form final products. In some instances, the chromophore may be regenerated, resulting in a photocatalytic cycle that can be highly efficient (Fig. 1).

Excitation of a chromophore can initiate a number of photophysical and photochemical processes that are outlined in the following equations:

$$S + h\nu \xrightarrow{J} S^* \quad \text{excitation;} \tag{1}$$

$$S^* \xrightarrow{k_L, k_I} S\ (+h\nu) \quad \text{luminescence, internal conversion, and intersystem crossing;} \tag{2}$$

$$S^* + Q \xrightarrow{k_Q} S + Q \quad \text{quenching;} \tag{3}$$

$$S^* \longrightarrow S^+ + e^- \quad \text{photoionization;} \tag{4}$$

$$S^* \xrightarrow[{[R]}]{k_R} P \quad \text{reaction;} \tag{5}$$

$$S^* + R \xrightarrow{k_{ET}} S + R^* \quad \text{energy transfer;} \tag{6}$$

$$S^* + R \xrightarrow{k_{CT}} S^{+/-} + R^{-/+} \quad \text{charge transfer.} \tag{7}$$

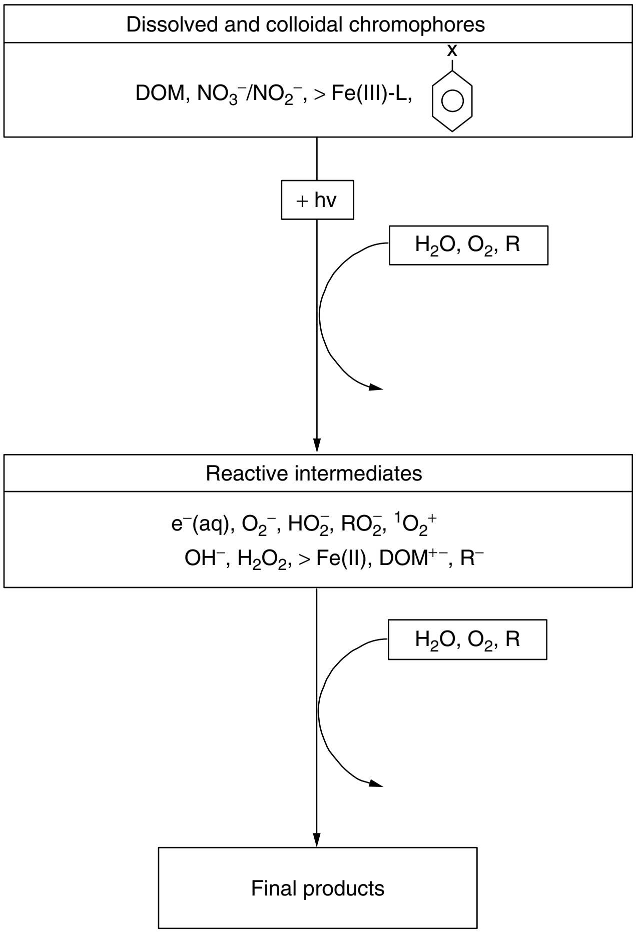

Figure 1. Schematic representation of photochemical processes in surface waters. Chromophores consist of DOM, nitrate/nitrite ions, dissolved and colloidal metal complexes [represented by Fe(III)], and anthropogenic pollutants (such as substituted aromatics). Reactive intermediates consist of solvated electrons, various radical species, both cationic and anionic, peroxides, and reduced metal centers.

Photophysical processes, luminescence (fluorescence and phosphorescence), internal conversion, intersystem crossing—Eq. 2—and collisional quenching—Eq. 3—will return the chromophore to the electronic ground state and convert the photon energy to heat. The excited state may also ionize, Eq. 4, or react, Eq. 5, converting the photon energy into chemical potential. The reaction could be unimolecular, such as dissociation, or bimolecular in nature, involving other components of the water system. Products might include the reactive intermediates in Fig. 1, which can then participate in secondary chemistry. Energy transfer, or sensitization, Eq. 6, could also induce a concurrent change in the electronic spin state, yielding a reactive triplet excited state.

The photochemical sequence is initiated by absorption of a photon, Eq. 1, with a photolytic rate constant, J. From a simple consideration of Beer's Law, J is seen to depend on the incident solar flux, I, and the (molar, decadic) extinction coefficient of the chromophore, ε. The relative efficiency, or yield, of a reaction can be specified by the quantum yield, ϕ, which is the number of product species formed divided by the number of photons absorbed. The absorption coefficient and quantum yield are generally wavelength dependent. The flux depends on wavelength as well as the solar zenith angle and other factors, such as latitude, longitude, and meteorological conditions, such as cloud cover. The rate constant for a photoinduced process can then be written as an integral over the wavelength region of interest:

$$J = 2.303 \int \varepsilon(\lambda)\phi(\lambda)I(\lambda)d\lambda = 2.303 \sum_i \overline{\varepsilon}(\lambda_i)\overline{\phi}(\lambda_i)\overline{I}(\lambda_i)\Delta\lambda_i \tag{8}$$

The flux in this case is termed the spectral actinic flux and is a spherically averaged irradiance that takes into account radiation incident on the chromophore over 4π steradians. For practical purposes, the integral is often evaluated as a discrete sum over small wavelength intervals, as expressed in the last term of Eq. 8, where $\overline{\varepsilon}$, $\overline{\phi}$, and $\overline{I}$ represent average quantities over the interval $\Delta\lambda_i$ centered at λ_i.

The solar actinic flux at noon (zenith angle = 0°) for a northern latitude (47.4°N) is shown in Fig. 2. The flux is expressed in terms of the photon flux per unit wavelength interval with units of einstein $\cdot$ m^{-2} $\cdot$ s^{-1} $\cdot$ nm^{-1}, where an einstein corresponds to 1 mole of photons (Avogadros's number, $N_A = 6.022 \times 10^{23}$ mol^{-1}). The UV-B radiation (280–320 nm) is strongly attenuated by the ozone layer, effectively limiting the useful actinic flux to about 300 nm. Shorter wavelength radiation does reach the surface but at 293 nm, the flux has decreased by more than two orders of magnitude relative to 300 nm. UV-A (320–400 nm) is not strongly absorbed by ozone and the flux increases up to the maximum in the visible at approximately 575 nm. A photon has a molar energy of

$$\overline{E} = N_A \frac{hc}{\lambda} \text{ J} \cdot \text{einstein}^{-1} \tag{9}$$

A 300-nm photon has an energy of 399 kJ $\cdot$ einstein^{-1}, which is on the order of typical bond enthalpies (e.g., C–C, 348 kJ; O–O, 146 kJ; C–OH, 322 kJ; benzylic C–H, 368 kJ; phenolic O–H, 362 kJ; C–Cl, 342 kJ). Over the wavelength range of photochemical interest, 300–500 nm, Fig. 2 indicates that the noon-time flux to surface waters is about 4 einstein $\cdot$ m^{-2} $\cdot$ h^{-1}. Typical penetration depths range from 1 to 20 m and so the molar rate of photon absorption is $0.2–4.2 \times 10^{-3}$ einstein$\cdot$L$^{-1}\cdot$h^{-1}. Dissolved organic carbon (DOC) is the major absorbing species in most natural waters (see Fig. 2) and has an approximate average concentration of 5 mg C$\cdot$L^{-1}. While highly diverse in nature, an average chromophore molecular weight of 120 g C$\cdot$mol^{-1} (10 carbon atoms) can be assumed and so the absorber concentration is approximately 0.04 mM. The estimated photon absorption rate for a well mixed volume would correspond to an average excitation rate of 5–100 h^{-1}, meaning that a near-surface chromophore is excited approximately once each minute (2,3). Furthermore, it is estimated that the accumulated annual radiation dose is about 1300 times larger (2). Given the time scales for transport and chemistry in surface waters, it is clear that photochemically driven reactions will be an important component of chemical cycles in the hydrosphere (Fig. 2).

REACTIVE INTERMEDIATES

A variety of reactive intermediates have been identified from the irradiation of natural waters at relevant solar wavelengths (see Fig. 1). Selected properties for some of the important species are listed in Table 1, along with estimated steady-state concentrations in typical surface waters (generally corresponding to a noon-time solar flux). Under conditions of constant solar illumination, pseudo-first-order kinetics result and the steady-state concentration for an intermediate, $[I]_{SS}$, can be used in conjunction with the second-order rate constant, k, for a particular reaction to calculate the reaction half-life as

$$\tau = \frac{\ln 2}{k[I]_{SS}} \tag{10}$$

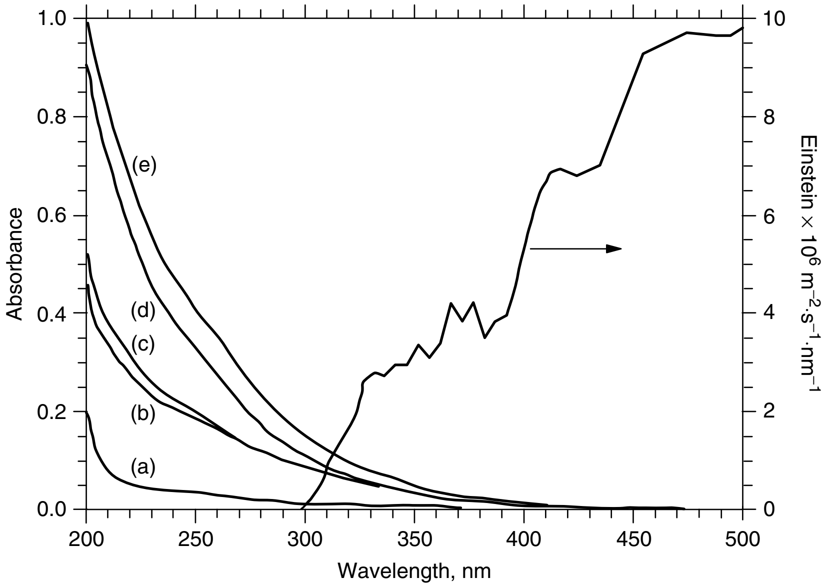

Figure 2. Left ordinate: Absorption spectra of surface water samples with varying concentrations and sources of DOC in mg C$\cdot$L^{-1}: (**a**) = 3.0, (**b**) = 7.8, (**c**) = 13.4, (**d**) = 13.4, (**e**) = 15.4. Path length is 1.0 cm. Data are adapted from Reference 7. Right ordinate: Solar actinic flux for noon-time sun (zenith angle = 0°) for a northern, midlatitude. Data is from Reference 1.

and, thus, estimate the importance of a given reaction pathway. For example, half-lives for reactions with phenol, a representative organic, aromatic compound, have been included in Table 1.

Primary transient species are produced rapidly, on the order of <1 ns. These transients can then react with other components in the aquatic environment to produce secondary products. Besides H_2O, one of the most important interactions involves dissolved oxygen, which, in well aerated surface waters, will be present in concentrations determined by Henry's Law,

$$[O_2] = K_H P_{O_2} \tag{11}$$

where $K_H = 1.3 \times 10^{-3}\ M^{-1} \cdot atm^{-1}$ is the Henry's Law constant (8). The mixing ratio of O_2 in the atmosphere is 21%, yielding a concentration of $[O_2] = 0.27$ mM at 298 K. Primary radicals will react with dissolved O_2 at diffusion controlled rates (diffusion controlled rate constant, $k_D \approx 7 \times 10^9\ M^{-1} \cdot s^{-1}$ at 298 K) corresponding to a time scale of $\approx 0.4\ \mu s$, which will limit the lifetime of many photogenerated transients (the "oxygen wall") (9).

e^-(aq)

Electrons generated from photoionization of DOM are rapidly solvated in aqueous environment to produce, e^-(aq). The solvated electron reacts with dissolved oxygen to form the superoxide,

$$e^-(aq) + O_2 \xrightarrow{k} O_2^- \tag{12}$$

where $k \approx 2 \times 10^{10}\ M^{-1} \cdot s^{-1}$ [which is larger than the diffusion limit due to a tunneling mechanism for e^-(aq) transport] (4,5). The time scale for Eq. 12 is on the order of 200 ns. The short lifetime coupled with measured typical production rates leads to a very small steady-state concentration of $[e^-(aq)]_{SS} \approx 10^{-17}$–$10^{-15}$ M (2,9). While reaction with e^-(aq) has been shown to degrade organic pollutants (2,10), it is unlikely to play an important role in photochemical processes in most surface waters.

O_2^-

Superoxide, O_2^-, can be formed from reaction with e^-(aq), as above, in charge transfer reactions with sufficiently reducing species, and via electron trapping at the surface of semiconductor colloids. Photolysis of DOM is a major source of superoxide but photolysis of trace metal complexes may also contribute to O_2^- formation (11). Superoxide is a mild reductant/oxidant (Table 1) and can participate in both one-electron oxidation and reduction reactions. The steady-state concentration is uncertain but from limited formation rate measurements, the concentration is $[O_2^-]_{SS} \approx 10^{-9}$–$10^{-7}$ M (9,10). The relatively high concentration suggests that superoxide may be important in photochemical degradation of various species (7). However, O_2^- is unstable in water with respect to disproportionation to form hydroperoxyl radical and hydrogen peroxide (3,12)

$$O_2^- + H^+ \overset{K}{\rightleftharpoons} HO_2 \quad \log K = -4.8 \tag{13}$$

$$HO_2/O_2^- + H^+ \longrightarrow H_2O_2 + O_2 \tag{14}$$

$RO_2^\bullet$

Peroxyl radicals can be formed from the photolysis of DOM and metal complexes. Formation rates for RO_2 species are thought to linearly scale with the DOM concentration in typical waters (2), but the reactivity of the radical is highly dependent on the type of humic substance involved, which is itself not well characterized. These uncertainties make it difficult to estimate a steady-state concentration for peroxyl radicals. HO_2 is a mild oxidant/reductant (Table 1) but DOM-derived peroxyl radicals exhibit significant

Table 1. Selected Properties of Some Important Reactive Intermediates in Aqueous Photochemistry[a]

	Radical				
Property	e^-(aq)	O_2^-	HO_2	OH	$^1O_2^*$
ΔG_f^0 (kJ·mol^{-1})	276[b]			13[b]	
ΔH_f^0 (kJ·mol^{-1})	277[b]	80[c]	138[c]	−7[b]	94.3[d]
E^0 (V)[e]	−2.9[f]	−0.33[g]	0.79[h]	1.9[i]	0.65[j]
$[I]_{SS}$ (M)	10^{-17}–10^{-15}	10^{-9}–10^{-7}		10^{-19}–10^{-17}	10^{-14}–10^{-13}
k (M^{-1}·s^{-1})[k]	3.0(7)	5.8(2)	8.5(3)	1.0(10)	0.2(7)
τ (days × 10^3)	0.3–30	0.0001–0.01		0.08–8	0.04–0.4

[a] The rate constant data is for reaction with C_6H_5OH (phenol) except for the HO_2 radical, which is for reaction with 1,4-$C_6H_4(OH)_2$.

[b] Data from Reference 4.

[c] Data from Reference 5.

[d] Gas phase value.

[e] Data from Reference 6.

[f] $H_2O + e^- \rightarrow e^-(aq)$

[g] $O_2 + e^- \rightarrow O_2^-$

[h] $HO_2 + e^- \rightarrow HO_2^-$

[i] $OH + e^- \rightarrow OH^-$

[j] $^1O_2^* + e^- \rightarrow O_2^-$

[k] Number in parentheses indicates the exponent value. Data from Reference 7.

reactivity with alkylphenols (13) and so may be an important reaction channel for some pollutants.

$^1O_2$*

Singlet oxygen, $^1O_2^*$, is an electronically excited state of oxygen with 94.3 kJ of excess energy. Triplet excited ^{3}DOM*, generated by intersystem crossing from the initially excited singlet state, can sensitize the formation of $^1O_2^*$ through energy transfer, Eq. 6. Most of the singlet oxygen will be rapidly quenched back to the triplet ground state by H_2O ($\tau \approx 4\ \mu s$) and only a fraction of the $^1O_2^*$ will be available for subsequent reaction (2). The production of $^1O_2^*$ generally scales with the amount of DOC and steady-state concentrations of $0.3–3.0 \times 10^{-14}\ M$ per mg·L^{-1} of DOC have been measured for typical surface waters (14). Singlet oxygen is a reactive electrophile, particularly toward cyclic 1,3 dienes, phenolate anions, and some polynuclear aromatics (2,14,15) and can play an important role in aqueous photochemistry.

OH

The hydroxyl radical is a powerful oxidant (Table 1) that exhibits high reactivity toward a range of organic species. It is mainly formed from photodissociation of NO_3^-/NO_2^- but also via photolysis of metal-hydroxides, decomposition of H_2O_2, and DOM photochemistry. The hydroxyl radical is rapidly scavenged by DOM and the steady-state concentration of OH in surface waters is generally small, $[OH]_{SS} \approx 10^{-19}–10^{-17}$ M (9,10). The OH radical readily reacts with many organic species, via abstraction and addition to double bonds, but the low steady-state concentration suggests that hydroxyl will not be the dominant photooxidant for most surface waters. For example, a "typical" OH rate constant (4) is $5 \times 10^9\ M^{-1} \cdot s^{-1}$, corresponding to a reaction half-life of $\tau \approx 160$ d at the higher OH concentrations. However, for environments with a larger concentrations of NO_x^- relative to DOC, hydroxyl radical reactions can be important pathways due to near diffusion-controlled OH rate constants (2).

H_2O_2

Hydrogen peroxide is a relatively stable, nonradical species produced from photochemical processes in natural surface waters. It is not a primary product but, rather, is generated in secondary reactions such as the disproportionation of superoxide, Eqs. 13 and 14. Hydrogen peroxide absorbs only weakly in the UV ($\varepsilon = 1\ M^{-1} \cdot cm^{-1}$ at 310 nm) (16), and decomposition to yield hydroxyl radicals may be dominated by biological mediation and redox cycles involving trace metals

$$H_2O_2 \longrightarrow 2OH \tag{15}$$

The slow decomposition of H_2O_2 leads to relatively high steady-state concentrations in surface waters, $[H_2O_2]_{SS} = 10^{-8}–10^{-7}$ M (2,9).

CHROMOPHORES

NO_3^-/NO_2^-

Inorganic nitrate and nitrite ions are perhaps the most well defined and easily understood photoactive component of natural waters. Nitrate concentration in surface waters is on the order of 1 mg·L^{-1} although it may be substantially higher in waters subject to an influx of agricultural drainage due to the heavy use of nitrogen fertilizers. The absorption band maxima for nitrate and nitrite are at 302 nm ($\varepsilon = 3.1\ M^{-1}{\cdot}cm^{-1}$) and 352 nm ($\varepsilon = 9.6\ M^{-1}{\cdot}cm^{-1}$), respectively (17,18). Ultraviolet excitation of nitrite leads to photodissociation (18):

$$NO_2^- + h\nu \longrightarrow NO + {}^\bullet O^- \tag{16}$$

The $^\bullet O^-$ species, the conjugate base of hydroxyl radical, is rapidly protonated at pH < 12 [pK_a($^\bullet$OH) = 11.9] (4):

$${}^\bullet O^- + H^+ \longrightarrow {}^\bullet OH \tag{17}$$

The quantum yield for $^\bullet$OH production from nitrite varies from 0.015 to 0.08 over the wavelength range 298–371 nm (18,19). Photolysis of the nitrate ion yields either nitrite ion or nitrogen dioxide, which can then reform nitrate and nitrite:

$$NO_3^- + h\nu \longrightarrow NO_2^- + O \tag{18}$$

$$NO_3^- + h\nu \longrightarrow {}^\bullet NO_2 + {}^\bullet O^- \tag{19}$$

$$2{}^\bullet NO_2 + H_2O \longrightarrow NO_2^- + NO_3^- + 2H^+ \tag{20}$$

For wavelengths ≥290 nm, the quantum yield for formation of $^\bullet$OH from NO_3^-, via Eq. 19, is in the range of $9.2–17 \times 10^{-3}$ (17,18). Hydroxyl is a highly reactive radical and a strong oxidizing agent (see above). Various laboratory studies have shown that trace organic molecules, such as biologically refractory pollutants, are degraded by irradiation in the presence of nitrate/nitrite (17–19).

DOM

Dissolved organic matter (DOM) is a major and ubiquitous component of natural surface waters, having numerous sources such as soils, sediments, and aqueous biological activity (20). DOM concentrations are often expressed in terms of the dissolved organic carbon (DOC) fraction, which comprises about 60% of the total DOM. Continental waters have typical levels of 1–10 mg C·L^{-1} DOC, although some waters, bogs, marshes, and shallow streams, for instance, can have concentrations exceeding 50 mg C·L^{-1}. Marine DOC is about 1 mg C·L^{-1} and represents an enormous carbon reservoir, approximately equal in size to the CO_2 reservoir in the atmosphere (2). DOM is nominally defined as the organic content of water that passes through a filter of some pore size, 0.1–0.5 μm, and so will have a colloidal component. The DOC in surface waters can be divided into two fractions; a lower molecular weight nonhumic fraction and higher molecular weight humic component. The nonhumic fraction is about 10–60% of the DOC and is made up of small compounds of mainly biogenic origin, such as peptides, carbohydrates, fats, and chlorophyll, which are subject to rapid biological degradation. In addition, anthropogenically derived species, such as biocides (often halogenated compounds) and polycyclic aromatic hydrocarbons (PAHs, from combustion processes) are present. The higher molecular weight fraction (>300 Da), comprising 40–90% of the DOC, is termed humic matter (HM). Humic matter is a complex material composed of condensed biopolymers and is much more biologically refractory. HM is mildly acidic, has a significant aromatic content, and is highly functionalized with carboxyl, phenol, hydroxyl, carbonyl, ether, and ester functional groups present. HM can be further divided based on solubility properties into humic acids (HAs), which precipitate under acidic conditions, and fulvic acids (FAs), which are soluble at all pH values. FAs are more highly substituted and have a higher oxygen content than HAs and also have a lower average molecular weight.

DOM is a highly colored material with an absorption spectrum that overlaps significantly with the solar actinic flux (see Fig. 2). The UV absorption of surface waters has been found to be highly correlated with the DOC content (10,21,22). Studies have shown that photoexcited DOM can yield many of the radical species outlined in Fig. 1 and Table 1 through primary and secondary processes (9). The radical production rate varies with the character of the DOM and the concentration, as well as the location and type of surface water environment. One study found that coastal marine waters yielded radical production rates of $0.18–1.8\ \mu\mathrm{M\cdot d^{-1}}$ in the summer due to irradiation of humic material in the surface layer (to a depth of 1.5 m) (23). The active spectral region was determined to be 300–385 nm, suggesting a quantum yield for radical production of $1–2 \times 10^{-3}$. Measurement of light absorption rates in highly colored, organic-rich waters that exhibit significant absorption at longer, visible wavelengths suggest the radical production rate will be orders of magnitude higher (9). These radicals dominate the photochemistry of most natural water samples, particularly in marine environments.

The DOM will be altered by photolysis, yielding a cationic radical product, for instance, but it can also be attacked by the photogenerated radicals. DOM decomposition can lead to photobleaching, allowing solar radiation to penetrate deeper into the water and impacting microbial photosynthesis. DOM is generally biologically refractory, resulting in long carbon reservoir residence times, estimated at 6000 years for deep ocean DOM (24). Photodecomposition of DOM may be an important removal pathway with significant implications for the global carbon cycle (9,15,24,25). The lower molecular weight degradation products are biologically labile and can provide a major fraction of the biological carbon uptake in certain surface waters (21,25).

Fe(III/II)

Iron is an abundant crustal metal and is one of the most important trace metals in natural water systems. The Fe-(hydrous)oxides (e.g., α-FeOOH, goethite; α-Fe_2O_3, hematite; γ-FeOOH, lepidocrocite; $Fe(OH)_3$, ferrihydrite) may be present as dissolved species but are predominantly in colloidal form. Iron has two common oxidation states, Fe(III) and Fe(II), and participates in redox reactions that are an important component of aqueous chemical cycles. Fe(III) can form strong complexes with naturally occurring organic ligands, such as polycarboxylate anions (e.g., oxalate, malonate, citrate), and humic and fulvic acids (11,26). Photoexcitation of a Fe(III) complex promotes a ligand to metal charge transfer process yielding a free, oxidized ligand and a reduced, Fe(II), center. Subsequent reactions generate reactive intermediates (11,12), as schematically depicted below for the oxalato ($C_2O_4{}^{2-}$) ligand:

$$>\mathrm{Fe(III)C_2O_4^-} + h\nu \longrightarrow\ >\mathrm{Fe(II)} + \mathrm{C_2O_4^-} \quad (21)$$

$$\mathrm{C_2O_4^-} \longrightarrow \mathrm{CO_2} + \mathrm{CO_2^-} \quad (22)$$

$$\mathrm{CO_2^-} + \mathrm{O_2} \longrightarrow \mathrm{CO_2} + \mathrm{O_2^-} \quad (23)$$

$$>\mathrm{Fe(II)} + \mathrm{O_2} \longrightarrow\ >\mathrm{Fe(III)} + \mathrm{O_2^-} \quad (24)$$

$$>\mathrm{Fe(II)} + \mathrm{H_2O_2} \longrightarrow\ >\mathrm{Fe(III)} + \mathrm{OH} \quad (25)$$

where ">Fe(III/II)" represents either an iron colloid surface site or a dissolved species. The superoxide will form hydrogen peroxide and hydroperoxyl radicals, as discussed above. In colloidal material, the surface Fe(II) is generally oxidized back to Fe(III) by the ligand or O_2. However, due to weakened binding with neighbor atoms, the reduced species may also be released to the solution (3,26). The result is photodissolution of the mineral colloid and an increase in soluble iron species, facilitating redox reactions and promoting microbial activity.

BIBLIOGRAPHY

1. Finlayson-Pitts, B.J. and Pitts, J.N., Jr. (1986). *Atmospheric Chemistry: Fundamentals and Experimental Techniques*. Wiley-Interscience, Hoboken, NJ.
2. Hoigne, J., Faust, B.C., Haag, W.R., Scully, J.F.E., and Zepp, R.G. (1989). Aquatic humic substances as sources and sinks of photochemically produced transient reactants. In: *Aquatic Humic Substances*. Vol. 219, I.H. Suffet and P. MacCarthy (Eds.). American Chemical Society, Washington, DC, pp. 363–381.
3. Stumm, W. and Morgan, J.J. (1996), 3rd Edn. John Wiley & Sons, New York.
4. Buxton, G.V., Greenstock, C.L., Helman, W.P., and Ross, A.B. (1988). Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms and hydroxyl radicals (OH/O^-) in aqueous solution. *J. Phys. Chem. Ref. Data* **17**: 513–886.
5. Spinks, J.W.T. and Woods, R.J. (1990). *An Introduction to Radiation Chemistry*, 3rd Edn. John Wiley & Sons, Hoboken, NJ.
6. Wardman, P. (1989). Reduction potentials of one-electron couples involving free radicals in aqeous solution. *J. Phys. Chem. Ref. Data* **18**: 1637–1755.
7. NDRL Radiation Chemistry Data Center. Available: http://www.rcdc.nd.edu/.
8. Seinfeld, J.H. and Pandis, S.N. (1998). *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons, Hoboken, NJ.

9. Mopper, K. and Kieber, D.J. (2000). Marine photochemistry and its impact on carbon cycling. In: *The Effects of UV Radiation in the Marine Environment*. Vol. 10. S.D. Mora, S. Demers, and M. Vernet (Eds.). Cambridge University Press, Cambridge, UK, p. 101–129.
10. Cooper, W.J., Zika, R.G., Petasne, R.G., and Fischer, A.M. (1989). Sunlight-induced photochemistry of humic substances in natural waters: major reactive species. In: *Aquatic Humic Substances*. Vol. 219. I.H. Suffet and P. MacCarthy (Eds.). American Chemical Society, Washington, DC, pp. 333–362.
11. Faust, B.C. (1994). A review of the photochemical redox reactions of iron(III) species in atmospheric, oceanic, and surface waters: influences on geochemical cycles and oxidant formation. In: *Aquatic and Surface Photochemistry*. G.R. Helz, R.G. Zepp, and D.G. Crosby (Eds.). Lewis Publishers, Boca Raton, FL, pp. 3–37.
12. Zuo, Y. and Hoigne, J. (1992). Formation of hydrogen peroxide and depletion of oxalic acid in atmospheric water by photolysis of iron(III)-oxalato compounds. *Environ. Sci. Technol.* **26**: 1014–1022.
13. Faust, B.C. and Hoigne, J. (1987). Sensitized photooxidation of phenols by fulvic acid and in natural waters. *Environ. Sci. Technol.* **21**: 957–964.
14. Haag, W.R. and Hoigne, J. (1986). Singlet oxygen in surface waters. 3. Photochemical formation and steady-state concentration in various types of waters. *Environ. Sci. Technol.* **20**: 341–348.
15. Aguer, J-P., Richard, C., and Andreux, F. (1997). Comparison of the photoinductive properties of commercial, synthetic and soil-extracted humic substances. *J. Photochem. Photobiol. A* **103**: 163–168.
16. Bolton, J.R. and Cater, S.R. (1994). Homogeneous photodegradation of pollutants in contaminated water: an introduction. In: *Aquatic and Surface Photochemistry*. G.R. Helz, R.G. Zepp, and D.G. Crosby (Eds.). Lewis Publishers, Boca Raton, FL, pp. 467–490.
17. Zepp, R.G., Hoigne, J., and Bader, H. (1987). Nitrate-induced photooxidation of trace organic chemicals in water. *Environ. Sci. Technol.* **21**: 443–450.
18. Alif, A. and Boule, P. (1991). Photochemistry and the environment part XIV. Phototransformation of nitrophenols induced by excitation of nitrite and nitrate ions. *J. Photochem. Photobiol. A* **59**: 357–367.
19. Marbury, S.A. and Crosby, D.G. (1994). The relationship of hydroxyl reactivity to pesticide persistence. In: *Aquatic and Surface Photochemistry*. G.R. Helz, R.G. Zepp, and D.G. Crosby (Eds.). Lewis Publishers, Boca Raton, FL, pp. 149–161.
20. Beckett, R. (1990). The surface chemistry of humic substances in aquatic systems. In: *Surface and Colloid Chemistry in Natural Waters and Water Treatment*. R. Beckett (Ed.). Plenum Press, New York, pp. 3–20.
21. Kieber, R.J., Zhou, X., and Mopper, K. (1990). Formation of carbonyl compounds from UV-induced photodegradation of humic substances in natural waters: fate of riverine carbon in the sea. *Limnol. Oceanogr.* **35**: 1503–1515.
22. Miller, W.L. (1994). Recent advances in the photochemistry of natural dissolved organic matter. In: *Aquatic and Surface Photochemistry*. G.R. Helz, R.G. Zepp, and D.G. Crosby (Eds.). Lewis Publishers, Boca Raton, FL, pp. 111–127.
23. Zafiriou, O.C. and Dister, B. (1991). Photochemical free radical production rates: Gulf of Maine and Woods Hole-Miami transect. *J. Geophys. Res.* **96**: 4939–4945.
24. Mopper, K. et al. (1991). Photochemical degradation of dissolved organic carbon and its impact on the oceanic carbon cycle. *Nature* **353**: 60–62.
25. Kieber, D.J., McDaniel, J., and Mopper, K. (1989). Photochemical source of biological substrates in sea water: implications for carbon cycling. *Nature* **341**: 637–639.
26. Sulzberger, B., Laubscher, H., and Karametaxas, G. (1994). Photoredox reactions at the surface of iron(III) (hydr)oxides. In: *Aquatic and Surface Photochemistry*. R. Helz, R.G. Zepp, and D.G. Crosby (Eds.). Lewis Publishers, Boca Raton, FL, pp. 53–73.

ISOTOPE EXCHANGE IN GAS-WATER REACTIONS

TOM WYDRZYNSKI
WARWICK HILLIER
The Australian National University
Canberra, Australia

INTRODUCTION

Molecular oxygen began to accumulate in Earth's atmosphere about 2.2–2.7 billion years ago, and it is widely accepted that it originated from the photosynthetic oxidation of water (1). The accumulation of O_2 in the atmosphere marked the most important event in the evolution of life on Earth as the atmosphere changed from a highly reducing to an oxidizing condition. All preexisting life forms had to either find an anaerobic ecological niche or adapt to the ever-increasing O_2 levels of the atmosphere.

Today on Earth, the photosynthetic O_2 generators are widely distributed on land and in the oceans in the form of plants, algae, and cyanobacteria. The current atmospheric pool of O_2 contains about 4×10^{19} mols and takes about 3.7 million years to turnover. The unique enzyme complex that produces the O_2 by photosynthetic organisms is called photosystem II (PSII), which acts as a water/plastoquinone oxido-reductase (2). PSII contains a special photochemical reaction center, which at excitation can generate strong oxidants that can be coupled to the splitting of two water molecules into O_2, four protons, and four electrons. The protein motif within PSII that catalyzes the water oxidation reaction contains a cluster of four Mn ions, a Ca ion, and a redox-active tyrosine and is often referred to as the O_2 evolving complex (OEC).

The oxidation of water by PSII is driven photochemically by the capture of light quanta (3,4). The chlorophyll binding proteins associated with peripheral regions of the PSII complex serve as light-harvesting antenna and direct spectrally diverse quanta to the special reaction center chlorophyll cluster called P680. At absorption of a light quantum, P680 generates an excited state with its pheophytin partner (P680-Pheo)*, which in turn is in equilibrium with the exciton in the antenna and the radical pair $P680^+$-$Pheo^-$. The time course for this reaction has been under some debate, but it is in the order of 3–20 ps. The primary charge separation is rapidly followed by a charge stabilization event, in which the electron is rapidly transferred to a tightly bound plastoquinone molecule, Q_A, in 300−500 ps, which results in the formation of $P680^+$-Pheo Q_A^-. The electron hole in $P680^+$ is subsequently filled with an electron derived from a redox-active tyrosine, Y_Z, located in the D1 protein subunit within some 40−380 ns, which results in the formation of Y_Z^{ox} P680-Pheo Q_A^-. The electron at Q_A^- is then transferred to a second bound plastoquinone molecule, Q_B, in 100−200 μs, whereas Y_Z^{ox} is reduced by the Mn_4Ca cluster with S-state-dependent kinetics (see below) in the range of 30−1300 μs.

One key finding for understanding the water oxidation reaction is the observation that for dark-adapted, PSII-containing samples illuminated by a series of brief, saturating light flashes, O_2 is released with a damped periodicity of four. The peak O_2 yields occur on the third, seventh, eleventh, and so on flashes and eventually damp out after the third or fourth cycle (5, see Figs. 2B and 2C). The phenomenological model to explain this observation was developed by Kok et al. (6,7) and entails a cyclic reaction sequence in which the OEC traverses through five intermediate states called the S states. Beginning in S_0 and progressing to S_4, each S state is advanced by a single quantum event at P680. When it reaches S_4, O_2 is released within 1−2 ms, S_0 is regenerated, and the cycle begins anew (see Fig. 1).

The damping in the O_2 flash yield oscillations (which is caused by the desynchronization of the forward S state transitions within a given population of O_2 evolving centers) is explained by a certain probability for a "miss" (α) to occur during a flash in which an electron hole fails to accumulate in the OEC and, depending on the flash duration, by a certain probability for a "double hit" (β) to occur in which a double excitation event takes place. The maximum O_2 yield on the third flash is explained by the condition that the cycle initially starts in the S_1 state after complete dark-adaptation. The higher S states (S_2 and S_3) exhibit lifetimes of tens of seconds to several minutes, depending on the temperature, and deactivate to S_1 if not advanced by another incoming quantum, whereas the S_0 state slowly relaxes in the dark to the S_1 state (8). The S_4 state being metastable reacts immediately to produce O_2 and the S_0 state.

DURING WHICH STEP IN THE S STATE CYCLE DOES THE SUBSTRATE WATER BIND?

In the original Kok et al. hypothesis, it was implied that the two substrate water molecules entered the reaction sequence during the last step, just before O_2 release. Although this notion has been considered throughout the literature, the most recent models for the OEC generally predict that the substrate water binds to the catalytic site at the beginning of the S state cycle and that a concerted four-electron event occurs during the S_3 to S_4 to S_0 transition to create the O–O bond (9,10). However, definitive experimental evidence to show at which step in the S state cycle the substrate water binds and the O–O bond forms has until recently remained elusive.

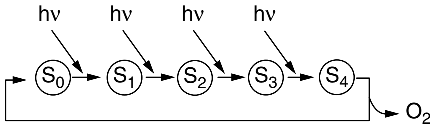

Figure 1. OEC Cycle.

The binding of the substrate water to the catalytic site can be probed by oxygen isotope exchange measurements between the solvent water and the photogenerated O_2 using mass spectrometry. Basically, the measurement involves the transfer of a photosynthetic sample into labeled water of known oxygen isotopic composition (i.e., ^{18}O-enriched water) and then determining the extent of isotopic incorporation into the photogenerated O_2.

In the first oxygen isotope exchange measurements (11−13), the fastest detectable rate of exchange was limited to ∼30 s, mainly because of the long stabilization time of the "open" chamber system that introduced isotopically labeled water to the sample. However, in the more recent measurements from our group, the stabilization time was greatly improved, by a factor of nearly 5000, through the introduction of a "closed" chamber system (14,15). The results from this experimental approach revealed for the first time that one substrate water molecule is bound to the catalytic site throughout the S-state cycle, whereas the second substrate water molecule is bound in at least the S_2 and S_3 states, and that the O–O bond indeed forms during the last step of the reaction sequence. This article documents the experimental approach and the arguments that come to these conclusions.

RAPID ^{18}O EXCHANGE MEASUREMENTS DURING PHOTOSYNTHETIC O_2 EVOLUTION

To probe the substrate water bound at the OEC, excess ^{18}O isotope (added as ^{18}O-labeled water) is introduced into the sample to set up a new isotopic distribution, because natural abundance ^{18}O in the solvent water is low at about 0.2%. The effect of the new isotopic distribution on the photogenerated O_2 can then be measured at $m/e = 32$, $m/e = 34$, and $m/e = 36$ with an in-line mass spectrometer (e.g., the VG Instruments IsoPrime stable isotope spectrometer Manchester UK) for the $^{16}O^{16}O$, $^{16}O^{18}O$, and $^{18}O^{18}O$ di-oxygen products, respectively. From these measurements, the extent of ^{18}O incorporation during the O_2 evolving process can be determined. Figure 2(a) shows the O_2 flash yields at these three masses for spinach thylakoid membrane samples after the injection of an aliquot of $H_2{}^{18}O$ (e.g., 98.5% enriched water from ISOTECH, Miamisberg, Ohio) that gives a final ^{18}O isotopic enrichment in the sample of $\varepsilon \sim 12\%$. In this case, after complete equilibration, the absolute yields (as determined by the extrapolated lines shown in the figure for the third flash) at the three masses follow closely the theoretical isotope equilibrium distribution: $32{:}34{:}36 = (1-\varepsilon)^2 : 2\varepsilon(1-\varepsilon) : \varepsilon^2 = 100\%$, which indicates that no major isotopic discrimination occurs during the O_2 evolving process, as shown earlier from steady-state measurements (16).

Figure 2(b) shows a plot of the normalized O_2 yields at $m/e = 32$ as a function of flash number. The oscillatory

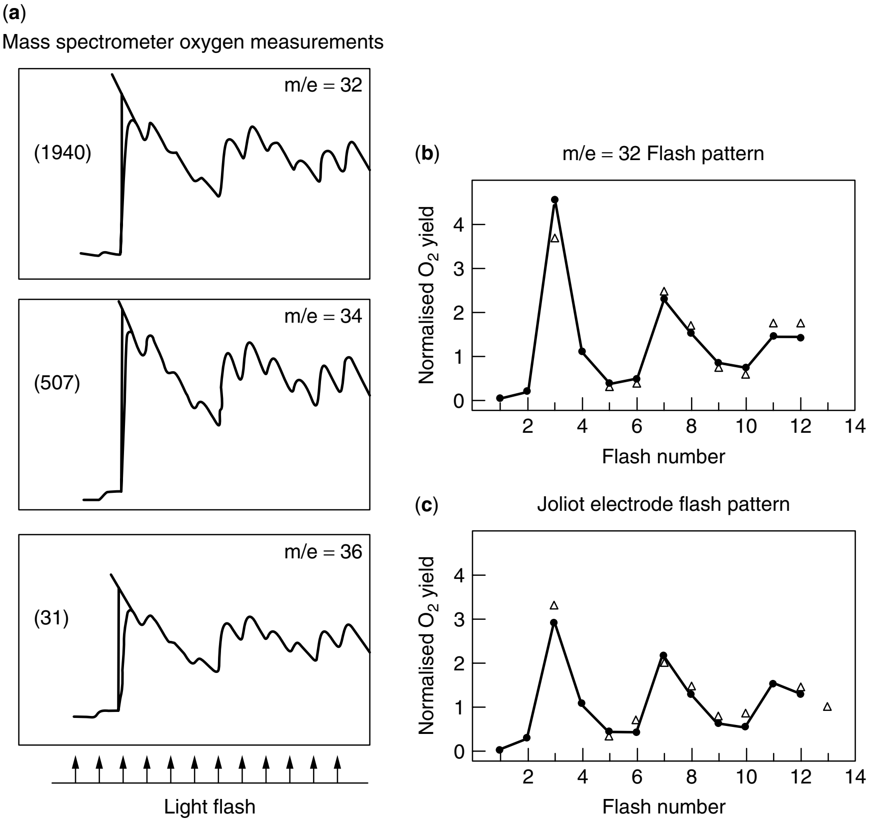

Figure 2.

pattern in the mass spectrometric measurements follows very closely the flash pattern in Fig. 2(c), which was obtained by the conventional amperometric method using a so-called Joliot-type oxygen electrode. Both types of measurements give rise to similar Kok et al. parameters, which indicates that the two techniques are measuring the same phenomenon.

To get information on the binding affinities of the substrate water in the OEC, the rate of isotope exchange between ^{18}O-labeled water and the photogenerated O_2 needs to be determined. To achieve this, our group designed a "closed" chamber system in which the sample volume is considerably reduced (160 μl) and a stirrer is incorporated to facilitate rapid mixing. By employing a solenoid-actuated Hamilton syringe, ^{18}O-labeled water can be rapidly injected and equilibrated with the sample within 6–8 ms. Thus, by measuring the extent of isotope incorporation as a function of the time after injection of the ^{18}O-labeled water, the rate of isotope exchange can be determined.

To evaluate the S-state dependence on the rate of isotope exchange, the "closed" chamber is optically coupled to a bank of three xenon flash lamps (e.g., FX-193 lamp, 4 μF @ 1 kV capacitor from EG & G, Salem, MA) via a 3-to-1 fiber optic to variously illuminate the sample with brief (FWHH ~8 μs), saturating light flashes. The samples are first given a series of activating flashes to preset the sample into a particular S state (see below). Then, after the rapid addition and equilibration of the ^{18}O-labeled water, a subsequent turnover flash sequence (spaced 5–10 ms apart, depending on S state, see below) is applied to photogenerate O_2. As the actual instrumental response time is relatively slow (being about 10 s because of the diffusion of the O_2 gas across a semipermeable membrane into the inlet line from the sample chamber), the flash spacings of a subsequent flash sequence that normalizes the oxygen signals are increased, typically to 20 s. As such, to retard the deactivation reactions of the higher S states during these flash spacings (see above), the temperature of the sample is reduced (usually to 10 °C), which is accomplished by enclosing the sample chamber in a cooling block. Details of the setup have been published earlier (63,64) and are given in Fig. 3.

S-STATE DEPENDENCE ON THE RATE OF ISOTOPE EXCHANGE BETWEEN ^{18}O-LABELED WATER AND THE PHOTOGENERATED O_2

To determine S-state dependence on the rate of isotope exchange, the mass measurements are made as a function of time between the addition and equilibration of ^{18}O-labeled water with the sample in a particular S-state and the first peak yield of the photogenerated O_2. Thus, for example, for the S_3 state, an S_1-enriched sample suspended in $H_2{}^{16}O$ is given two preflashes to advance the OEC into the S_3 state. $H_2{}^{18}O$ is then rapidly injected and followed by a variable delay time (Δt), which allows for the time-dependent isotopic exchange to occur between the solvent water and the substrate water. A single flash is then given, and the oxygen signals at m/e = 34 and m/e = 36 are measured. The oxygen signals are then collected on subsequent flashes (in which complete

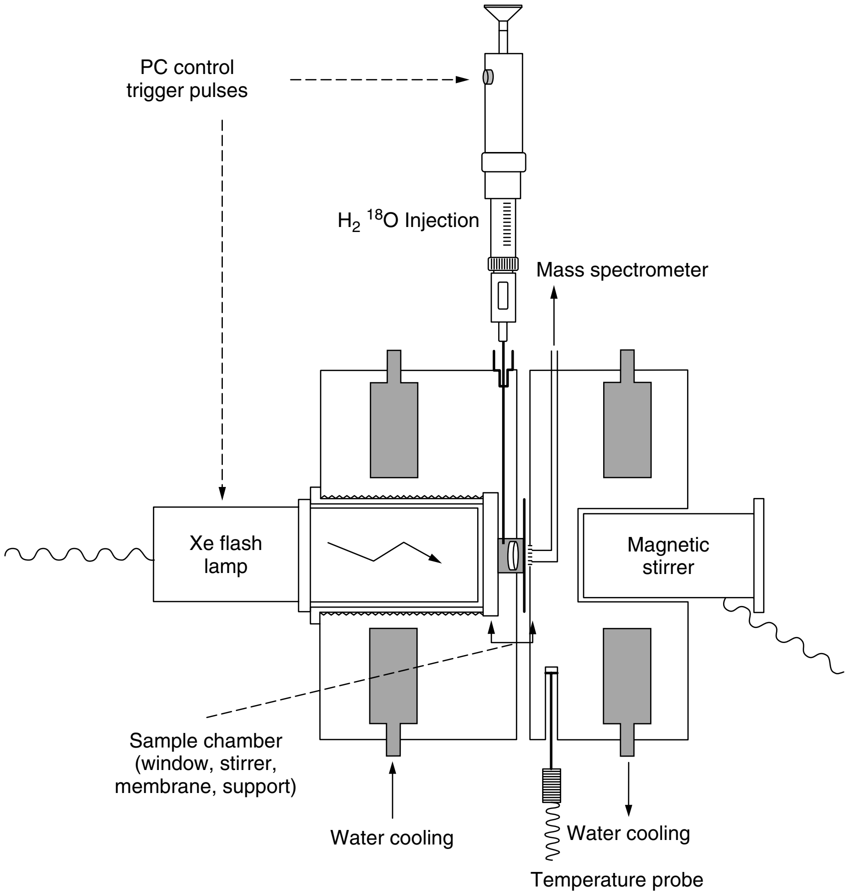

Figure 3.

exchange has occurred because of the 20 s flash spacings) to normalize the data among the different sample aliquots measured at different Δt values.

The complete ^{18}O exchange kinetics were determined for the different S states using the injection/flash protocols depicted in Fig. 4.

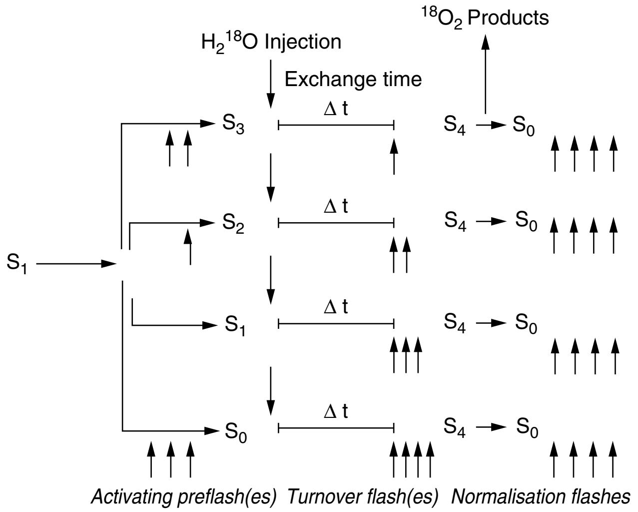

Figure 4.

The basic protocol involves the following procedure. Before loading, the samples are thawed in the dark at room temperature and given one saturating preflash followed by a 10-min dark period at room temperature to enrich the S_1 population. The S_1-enriched sample is then loaded into the sample chamber in the dark, equilibrated to 10 °C for 10–12 min, and then given several activating flashes (spaced 1 s apart). At a given S state, $H_2{}^{18}O$ is rapidly injected into the chamber ($t_{1/2\ (\mathrm{mixing})} \sim 4$ ms) and a variable delay time (4 ms < Δt <200 s) is provided to allow the water bound in the OEC to exchange with the ^{18}O-enriched solvent water. After the delay time, several rapid turnover flashes (spaced at 5, 10, and 10 ms apart for S_2, S_1, and S_0 states, respectively) are then applied to advance the sample through to the S_3 to S_4 to S_0 transition, where the mass ratio of the photogenerated O_2 (Y_M) is determined. After this process, several normalization flashes (spaced 20 s apart) are given and their separate O_2 yields ($Y_{n(1\rightarrow 4)}$) are measured. The Y_M is corrected for background O_2 contributions, double hits, and the injection response as described earlier (17,18) to give Y_C.

The resolvable ^{18}O exchange kinetics for spinach thylakoid samples as a function of the S-states are shown in Fig. 5. Here, the corrected, normalized yields of O_2 (Y_C) measured at $m/e = 34$ for the mixed labeled $^{16}O^{18}O$ product (left side of the figure) and at $m/e = 36$ for the double labeled $^{18}O^{18}O$ product (right side of the figure) are presented. Each point is a plot of Yc at a particular delay time (Δt) between the injection of $H_2{}^{18}O$ into the sample and the turnover flash sequences that photogenerated O_2. The m/e = 36 data exhibit only a single kinetic phase and are fit with a simple exponential function:

$$^{36}Y = (1 - \exp(-^{36}k\ t)) \tag{1}$$

In contrast, the m/e = 34 data on the left side in Fig. 5 reveals two distinct kinetic phases. The insets show expanded time ordinates for the fast phases. The two phases, however, are unequal in amplitude, with the fast phase constituting slightly more than half of the total signal. The basis for this difference in amplitude is well explained by the enrichment condition for two independent, exchanging sites. As the apparent kinetics of the two phases differ by at least a factor of 10, the fast phase of exchange is virtually complete before the slow phase begins. Thus, at short Δt only one substrate water molecule is exchanging at the catalytic site, which means at a typical ^{18}O enrichment of $\varepsilon = 12\%$, the mass distribution at 32:34:36 for the two oxygen isotopes will be 88:12:0. On the other hand, at longer Δt, when the second substrate water molecule is also exchanging, the mass distribution will be 77.44:21.12:1.44. Therefore, the relative contributions of the fast and slow phases will be unequal, with the fast phase representing ~57% (i.e., 12/21.12) of the total amplitude and the slow phase ~43%. The 0.57:0.43 distribution between the two phases is found consistently in the m/e = 34 data. Thus, the exchange kinetics in this case is fit exactly by the sum of two exponentials:

$$^{34}Y = 0.43(1 - \exp(-^{34}k_1\ t)) + 0.57(1 - \exp(-^{34}k_2\ t) \tag{2}$$

The solid lines in Fig. 5 represent the best fits to the data points based on Eqs 1 and 2 for the m/e = 36 and m/e = 34 data, respectively. The exchange rates from the corresponding kinetic fits at 10 °C are given in Table 1. The entire S-state results for the thylakoids have been presented earlier (17,18).

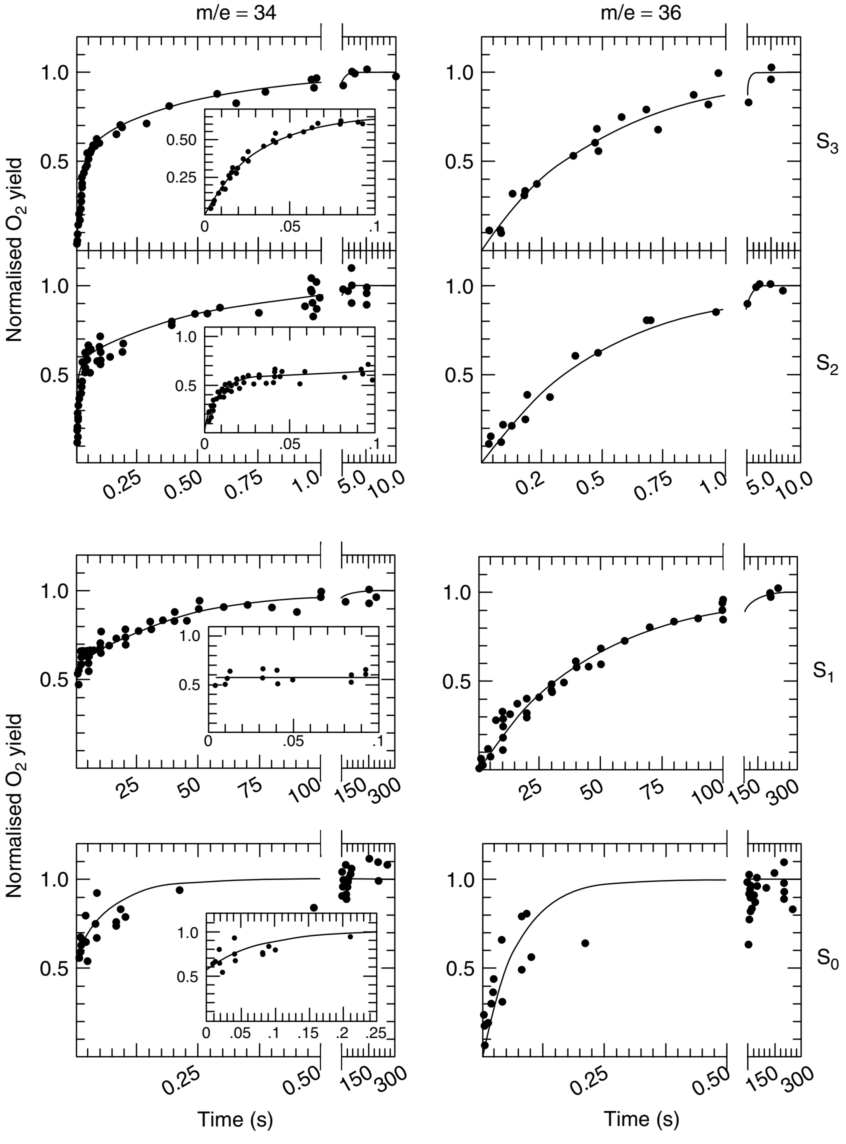

Figure 5.

Table 1. Rate Constants for the ^{18}O Exchange in Spinach Thylakoid Samples as a Function of the S States at 10 °C

S-State	^{36}k (s^{-1})	$^{34}k_1$ (s^{-1})	$^{34}k_2$ (s^{-1})
S_3	2.1 ± 0.2	1.9 ± 0.2	36.8 ± 1.9
S_2	2.2 ± 0.1	1.9 ± 0.3	120 ± 6
S_1	0.022 ± 0.002	0.021 ± 0.002	>175
S_0	18 ± 3	8 ± 2	>175

Adapted from Ref. 18.

IMPLICATIONS OF SUBSTRATE WATER BINDING FOR THE MECHANISM OF O_2 EVOLUTION

Several striking features are found in the rate constants listed in Table 1. In each S state, the slow phase of exchange was resolvable, whereas the fast phase of exchange could be resolved only in the S_2 and S_3 states. For the S_0 and S_1 states, the fast phase of exchange is equal to or faster than the response time of the currently available experimental setup ($\sim 100 - 175\ s^{-1}$). Most importantly, the rate constants for the slow phase kinetics in the m/e = 34 data are virtually identical to the rate constants for the monophasic kinetics in the m/e = 36 data. This last result clearly indicates that the extent of ^{18}O incorporation into the photogenerated O_2 is limited throughout the S-state cycle by the substrate water molecule undergoing the slow isotope exchange process and that the biphasic behavior in the m/e = 34 data does not originate from PSII heterogeneity. Rather the two kinetic components in the m/e = 34 data reflect the binding of the substrate water at two separate sites.

The most relevant question to ask is at which S state does the substrate water bind. Based on the kinetic resolution of our current mass spectrometric setup, a measurable ^{18}O exchange kinetic will conclusively show the existence of bound water in a particular S state, which is justified by the observations that the O_2 release on the S_3 to S_4 to S_0 transition takes place in ~2 ms, whereas the rest of the S-state advancement from S_0 to S_3 takes at most another 1 ms (8). The catalytic site thus can make one complete turnover in ~3 ms. Therefore, any ^{18}O exchange that is slower than 3 ms would be indicative of bound water. From the rate constants for the slow phase of ^{18}O exchange listed in Table 1 for the m/e = 34 data, it is clear that one substrate water molecule is bound to the catalytic site through the S-state cycle. Interestingly, the results for the fast phase of exchange show that the second substrate water molecule is bound in at least the S_2 and S_3 states, but they are inconclusive for the S_0 and S_1 states because the fast exchange kinetics in these S states are not resolvable by our current setup (i.e., faster than 6–8 ms).

Even though the measurable exchange rates listed in Table 1 are slower than the total turnover time of the catalytic site, the magnitudes of these exchange rates in general are all remarkably fast, where the slowest exchange is measured in the S_1 state with $t_{1/2}$ ~50 s. Exchange rates in this magnitude range indicate relatively labile binding sites and can be interpreted for two overall mechanisms: (1) exchange of water bound to a metal site or (2) diffusional or isotope equilibration processes, where the solvent water must be translocated to the binding site across a physical barrier.

Rates of whole water exchange at a metal site can span a tremendous range (as much as 18 orders of magnitude), and examples for various transition metals have been summarized earlier (19–22). The most significant factors governing water exchange at a metal site are the charge and ionic radius of the metal center and any electronic occupancy of d orbitals. As a general rule, water exchange rates will decrease (by several orders of magnitude) as the metal center (M^n) is oxidized and charge is increased; i.e., k_{ex} (M^n-OH_2) > k_{ex} (M^{n+1}-OH_2). Furthermore, metal-bound water can undergo deprotonation in response to an increase in charge (and acidity) at the metal center. Thus, it may also be expected for water exchange to decrease at deprotonation for a given oxidation state; i.e., k_{ex} (M^n-OH_2) > k_{ex} (M^n-OH). In the case of the deprotonation of a neighboring ligand (L), the opposite effect may occur because the effective charge of the metal center is reduced, As such the rate of water exchange will tend to increase; i.e., k_{ex} (HL-M^n-OH_2) < (L-M^n-OH_2). The protonation state will thus be defined by the pK of the aquo metal complex.

Any number of possible intermediates in the water oxidation reaction can exist, and it is important to appreciate that the ^{18}O exchange measurements could reflect not only whole water and hydroxyl exchange but also the exchange of μ-oxo (–O–), peroxo (–O–O–), and oxo (=O) intermediates. Indeed various models in the literature on O_2 evolution invoke one or more of these intermediates (9,10). Although a dearth still exists in studies of ^{18}O exchange in model compounds, it is expected that the isotope exchange in μ-oxo (–O–), perxo (–O–O–), and oxo (=O) intermediates is much slower (in the order of hours to days) than whole water exchange because of tighter bonding and lack of protonation.

The other type of mechanism that could limit the ^{18}O exchange could be diffusion or isotope equilibration across a physical barrier. The catalytic site for O_2 evolution in PSII is located within the protein domain away from the solvent interface. In this situation, the structural barriers provided by the protein matrix may be mediated through specific water chains or water channels (23), something akin to what has been proposed for the cytochrome c oxidase (24,25). Such a water channel may provide an important function in optimizing O_2 formation by controlling the accessibility of the water substrates and O–O bond formation at the catalytic site, thereby minimizing side reactions that lead to the formation of reactive oxygen intermediates (23). The inability to resolve the fast phase of exchange in the S_0 and S_1, states (Table 1) could indicate that the second substrate water molecule only enters the reaction sequence at the S_2 state. However, such a possibility can only be clarified by resolution of the fast phase kinetics or by the progressive improvement in the resolution of crystal structure of PSII (26–28).

The S-state dependence of the ^{18}O exchange kinetics for the substrate water are clearly complex, but these results place limits on the types of mechanisms that may be involved in the oxidation of water during photosynthesis. Although the S-state-dependent changes in the kinetics may in the first instance be interpreted for metal-centered oxidations and/or substrate or other ligand protonation/deprotonation events, second-order effects mediated by the protein environment, such as H-bonding networks and structural conformations, will also have to be considered. To sort out the fine details, additional measurements will have to be made under different biochemical conditions and in mutants of the PSII. This work is currently in progress.

Acknowledgments

The authors gratefully acknowledge the various contributions to this research by Johannes Messinger, Murray Badger, and Garth Hendry. This research was supported in part by a Human Frontier Science Program grant (RGP 29/2002).

BIBLIOGRAPHY

1. Blankenship, R.E. and Hartman, H. (1998). *T.I.B.S.* **23**: 94–97.
2. Wydrzynski, T. and Satoh, K. (Eds.). (In press). *Photosystem II: The Water/Plastoquinone Oxido-Reductase in Photosynthesis*. Springer-Kluwer Academic Publishers, Dordrecht, The Netherlands.
3. Diner, B.A. and Rappaport, F. (2002). *Annu. Rev. Plant. Biol.* **53**: 551–580.
4. Barter, L.M.C., Durrant, J.R., and Klug, D.R. (2003). *Proc. Natl. Acad. Sci. USA.* **100**: 946–951.
5. Joliot, P., Barbieri, G., and Chabaud, R. (1969). *Photochem. Photobiol.* **10**: 309–329.
6. Kok, B., Forbush, B., and McGloin, M. (1970). *Photochem. Photobiol.* **11**: 457–475.
7. Forbush, B., Kok, B., and McGloin, M. (1971). *Photochem. Photobiol.* **14**: 307–321.
8. Joliot, P. and Kok, B. (1978). *Bioenergtics of Photosynthesis*. Govindjee (Ed.). Academic Press, New York.
9. Nugent, J.H.A. (Ed.). (2001). Special Issue: Photosynthetic Water Oxidation. *Biochim. Biophys. Acta* **1503**.
10. Messinger, J. and Lubitz, W. (Eds.). Special Issue: Photosynthetic water oxidation. *Phys. Chem. Chem. Phys.* (In press).
11. Radmer, R. and Ollinger, O. (1980). *F.E.B.S. Lett.* **110**: 57–61.
12. Radmer, R. and Ollinger, O. (1985). *F.E.B.S. Lett.* **195**: 285–289.
13. Bader, K.P., Thibault, P., and Schmid, G.H. (1987). *Biochim. Biophys. Acta* **893**: 564–571.
14. Messinger, J., Badger, M., and Wydrzynski, T. (1995). *Proc. Natl. Acad. Sci. USA* **92**: 3209–3213.
15. Hillier, W., Messinger, J., and Wydrzynski, T. (1998). *Biochemistry* **37**: 16908–16914.
16. Guy, R.D., Fogel, M.L., and Berry, J.A. (1993). *Plant. Phys.* **101**: 37–47.
17. Hillier, W. and Wydrzynski, T. (2000). *Biochemistry* **39**: 4399–4405.
18. Hillier, W. and Wydrzynski, T. (In press). *Phys.Chem. Chem. Phys.*
19. Hillier, W. and Wydrzynski, T. (2001). *Biochim. Biophys. Acta* **1503**: 197–209.
20. Hendry, G. and Wydrzynski, T. (2003). *Biochemistry* **42**: 6209–6217.
21. Lincoln, S.F. and Merbach, A.E. (1995). *Adv. Inorg. Chem.* **42**: 1–88.
22. Richens, D.T. (1997). *The Chemistry of Aqua Ions*. John Wiley, West Sussex, UK.
23. Wydrzynski, T., Hillier, W., and Messinger, J. (1996). *Physiol. Plant.* **96**: 342–350.
24. Iwata, S., Ostermeier, C., Ludwig, B., and Michel, H. (1995). *Nature* **376**: 660–669.
25. Tsukihara, T. et al. (1996). *Science* **272**: 1136–1144.
26. Zouni, A. et al. (2001). *Nature* **409**: 739–743.
27. Kamiya, N. and Shen, J-R. (2003). *Proc. Natl. Acad. Sci. USA.* **100**: 98–103.
28. Ferreira, K.N. et al. (2004). *Science* **303**: 1831–1838.

RADON IN WATER

Babs A. Makinde-Odusola
Riverside, California

Some groundwater contains the naturally occurring radioactive gas, radon-222 (radon). Radon dissolved in groundwater is released into indoor air when plumbed in water is agitated indoors. Radon released from domestic water typically contributes less than 2% of total indoor radon. This article describes the health risks from radon in domestic water and the efforts within the United States to regulate radon in drinking water. The U.S. National Research Council (NRC) found that the primary health risks from radon in domestic water are lung cancer and stomach cancer, primarily to ever-smokers. Radon concentration in air could reach levels considered unsafe for underground miners at some water treatment facilities that aerate water as part of the treatment process. Inhalation of radon gas and its progeny may result in lung cancer (inhalation risk). An estimated 11% of the health risk is from direct ingestion of radon dissolved in drinking water (ingestion risk). The NRC found that radon is a greater source of cancer risk than any other regulated chemical in drinking water. Some ecological studies contradict using the linear nonthreshold (LNT) model to assess the health risks from exposure to radon at low levels typical of indoor air. The U.S. Environmental Protection Agency (EPA) relied on the LNT model to assess health risks and propose standards and goal for radon in drinking water. Some water utilities could opt for treatment to comply with the radon in drinking water standard. Some other water utilities might avoid treatment by adopting a multimedia approach to radon mitigation. The multimedia approach could be more cost-effective and achieve greater health risk reduction. Noneconomic considerations may influence that decision. Health risk reduction and cost analyses (HRRCA) suggest that most of the benefits would accrue to ever-smokers and smaller water systems could incur most of the costs.

INTRODUCTION

Radon, in this article, refers to the specific isotope, radon-222, the daughter product of radium-226 and a member of the uranium-238 decay series. The decay scheme for uranium-238 is illustrated in Drago (1) and

NRC (2). Uranium is found in trace amount in soils everywhere. Radon escapes as a gas from soils into groundwater, indoor air, and outdoor (ambient) air. The U.S. Geological Survey (3) reported that soil air within the United States contains between 200 and 2,000 pCi/L ($1\ \text{pCi/L} = 37\ \text{Bq}\,\text{m}^{-3}$) of radon. Radon is a chemically inert, colorless, and odorless gas whose half-life is about 3.82 days. Radon emits alpha (α) particles as it decays to a series of solid short-lived radionucleotides, such as polonium-218 and lead-214. Those physical properties are important in understanding the health risks from exposure to radon and the mitigation of such risks.

The United Nations Scientific Committee on the Effects of Atomic Radiation (4) estimated that about 50% of the average annual effective dose from natural sources of ionizing radiation worldwide is from exposure to the alpha-particles produced by radon and its progeny. Potential contributors of radon to indoor air include soil gas, natural gas, building materials, and to a lesser extent—domestic water. Radon emanating from soils under homes is the primary source of radon in indoor air, which can range from <1 to 3,000 pCi/L (3).

Radon in drinking water is typically of concern to sources that rely on groundwater, where there had not been sufficient opportunity for significant decay nor release into the atmosphere before reaching indoor plumbing. A portion of the radon in plumbed water used indoors is released into the air, especially during agitation (e.g., laundering, dishwashing, humidifying, toilet flushing, etc.), which increments the level of radon in indoor air. The NRC (2) estimated that the average transfer coefficient is about 0.01% (1 in 10,000) of radon in water; that is 10,000 pCi/L of radon in water would increase the level of radon in indoor air by 1 pCi/L. The U.S. EPA (5) estimated that less than 2% of radon in indoor air emanates from domestic water. This author prefers to emphasize “domestic water” because of the widespread use of bottled water for drinking and because the NRC (2) found that the major part of the health risk is from inhalation of radon and its progeny.

Some countries have policies to reduce public exposure to radon, especially in indoor environments. Cole (6) provides the historical development of radon risk policies in the United States. The U.S. Indoor Radon Abatement Act of 1988 established the goal of reducing indoor radon concentrations to natural ambient level. The NRC (2) estimated that the arithmetic national ambient average level of radon was 0.4 pCi/L. The U.S. EPA established a voluntary action level of 4 pCi/L for radon in indoor air. The indoor radon action level in other countries is as high as 27 pCi/L (7). Indoor radon guidelines are 22 pCi/L and 11 pCi/L in Canada and the European Union, respectively (7). The U.S. EPA (5) reported that more than 2500 lives were “saved” through the year 2000 since the inception of its voluntary indoor radon abatement program in 1986.

Many countries are developing policies to reduce exposure to radon in domestic water. The World Health Organization (WHO) is creating a database of national radon regulations [http://www.who.int/ionizing_radiation/env/radon/en/]. The rest of this article focuses on efforts within the United States. The NRC (2) found that radon is a greater source of cancer risk than any other regulated chemical in drinking water. Evaluating, and where necessary, mitigating the level of indoor radon is common during real estate transactions in the northeastern United States, even though compliance with the action level is voluntary. That practice could extend to testing for radon in privately owned wells after the EPA finalizes the maximum contaminant level (MCL) for radon in drinking water. Regulating radon in domestic water had been very controversial in the United States and dates back to 1986. In 1986, Congress directed the EPA to establish a mandatory standard for radon in drinking water as part of the amendments to the Safe Drinking Water Act (SDWA). The EPA (8) proposed the radon in drinking water rule using the risk management approach typically used for drinking water contaminants. Amendments to the SDWA in 1996 directed the EPA to withdraw the proposed rule of 1991 and repropose after an independent risk assessment by the NRC. The amendments include provisions for a multimedia approach to achieving an equal or greater level of health risk reduction. In 1999, the EPA (5) proposed 300 pCi/L and 4,000 pCi/L as the MCL and alternate MCL, respectively, for radon in drinking water. The EPA is expected to finalize the radon in drinking water rule after 2004. Key inputs and major milestones in developing the radon rule are illustrated in Fig. 1 of Article MW-66 (9).

OCCURRENCE OF RADON IN DOMESTIC WATER

Radon levels in water, like those in ambient air and indoor air, vary temporally and spatially. The EPA (5) analyzed data from the Safe Drinking Water Information System (SWDIS) to estimate the nationwide occurrence of radon in community water systems (CWS). The EPA (10) found geographic disparity in its occurrence. New England CWS had the highest arithmetic and geometric mean concentrations of 2933 pCi/L and 1214 pCi/L, respectively. Some privately owned wells, not monitored nor regulated by the EPA, might contain radon at much higher levels. The USGS (3) reported that the radon level in groundwater ranges from about 100 pCi/L to nearly 3,000,000 pCi/L. The EPA (5) estimated the population-weighted national average concentration of radon in CWS at 213 pCi/L, and a credible range of 213 to 240 pCi/L. In general, smaller CWS that rely on groundwater have the highest levels of radon. Radon levels exceeded 300 pCi/L at more than 51% and 16% of the CWS serving fewer than 100 people and 10,000 people, respectively (5). Radon levels exceeded 300 pCi/L and 4000 pCi/L in the domestic water of more than 16 million people and 77,000 people, respectively (5). The EPA (5) and NRC (2) have additional information on the occurrence and spatial distribution of radon in domestic water, indoor air, and ambient air. SWDIS would contain additional radon occurrence data when the radon rule is finalized and CWS begin compliance monitoring.

Radon levels at a particular well could fluctuate significantly. The EPA (5) proposed a monitoring regime that would require CWS using groundwater to monitor for radon initially at each entry point to the distribution system, quarterly for 1 year. Compliance would be based

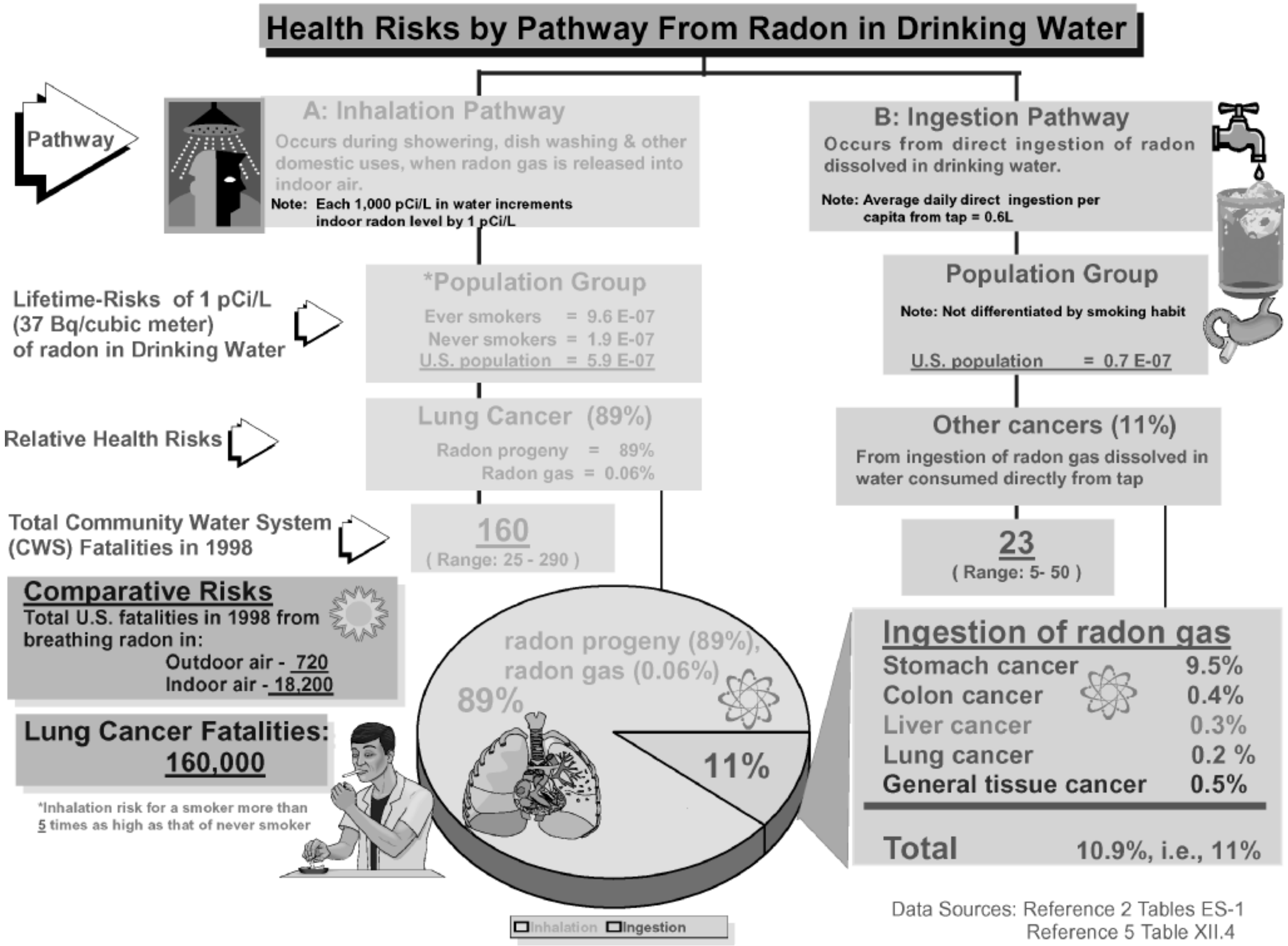

Figure 1. Estimated fatalities in the U.S. by pathway from radon in drinking water provided by CWS.

on the average of four consecutive quarterly samples. Zikovsky and Chah (11) found that concentrations of radon are lognormal in distribution and that the arithmetic mean of concentrations might create a false impression about exposure levels. The EPA (5) estimated that 26.8% of sources whose radon concentrations averaged between 200 and 300 pCi/L during the first four quarterly measurements would have long-term exceedances of 300 pCi/L. Field et al. (12) found higher levels of radon in domestic water at the point of use than prior to distribution because of increments from decay of the radium-226 deposits within the older portions of water distribution systems. The EPA (5) identified man-made sources of radon in groundwater to include releases from wastes from radium or thorium processing and from phosphate and uranium mining and milling operations.

RISK ASSESSMENT OF RADON IN DOMESTIC WATER

The health risks from radon in domestic water are primarily lung cancer and stomach cancer. Radon is a human carcinogen based on epidemiological studies of underground miners, cellular mutagenesis studies, and experimental research in animals (13). A linear nonthreshold (LNT) model is often used to assess the health risks of exposure to radon. The use of the LNT model to assess the health effects from exposure to radon at typical indoor concentrations is controversial. The National Council on Radiation Protection and Measurements (14) evaluated the evidence for and against the LNT dose–response hypothesis and found that the LNT model is more plausible. The NCRP (14) did not exclude other dose–response relationships because of growing evidence that adaptive responses, bystander effects, and other variables may modify the dose–response relationship. The NRC Committee on the Biological Effects of Ionizing Radiation (BEIR) VI did not exclude the possibility of a threshold relationship at the very low levels of radon exposure typically found in indoor air. The NRC (2) supported using the LNT relationship to derive quantitative estimates of cancer risks in the absence of credible evidence to the contrary. The NRC (2) recommended reconsideration of the LNT assumption when the “probability of an exposed cell becoming transformed” and the “latent period before malignant transformation is complete” are known. Lorimore et al. (15) observed the potential for interaction between irradiated and nonirradiated cells in the production of genetic damage and its significant implications for risk estimation. Brenner and Sachs (16) suggested that bystander effects from alpha-particle damage could cause downwardly curving dose–response relations and inverse dose-rate effects.

Parsons (17) presented a biological model of radiation hormesis that invalidated the LNT. Cohen (18) tested the LNT model by analyzing lung cancer mortality rates and average radon concentration in homes in 1601 U.S. counties while addressing problems typical of ecological studies. Cohen (18) found a "strong tendency for lung cancer rates to decrease with increasing radon exposure in sharp contrast to the increase" expected from the LNT model. The NRC (2) recommended investigating the significant negative trend of county lung cancer mortality with increasing home radon concentration established in the ecological study by Cohen (18). Puskin (19) found similar "strongly negative correlations for cancers strongly linked to cigarette smoking" but did not find "such correlations" for cancers not linked to smoking. Puskin (19) explained that the ecological correlation (18) "provides no substantial evidence for a protective effect of low level radon exposure." Van Pelt (2003) found that altitude (elevation) explains some of the inverse association between radon and lung cancer. Van Pelt (2003) postulated that the decrease in lung cancer is "caused by the carcinogenic effect of higher absolute oxygen concentration in the inspired air at lower."

The most direct way to assess the effects of radon in homes is through case-control studies (20). Krewski et al. (13) summarizes case-control studies of residential radon and lung cancer in North America. Field and Duport (21) presented points and counterpoints regarding lung cancer risks from exposure to residential radon. The NRC (2) reviewed available epidemiological studies of indoor radon and concluded "it is low and difficult, if not impossible to detect given the high background lung-cancer mortality in the populations studied."

Inhalation of radon gas, especially its progeny, causes lung cancer primarily among ever-smokers. BEIR VI (20) evaluated the risks of domestic radon to ever-smokers and never-smokers separately. CDC (22) defines ever-smokers as people who had smoked at least 100 cigarettes in their entire lifetimes. Alavanja (23) found many similarities between the pathologic lung-cancer mechanisms observed among those exposed to tobacco smoke and those exposed to radon gas.

NRC (2) conducted the most comprehensive review of the health risks of radon in domestic water. That review included epidemiological studies, BEIR studies, and earlier EPA (8,24) risk assessments. Both BEIR VI (20) and the NRC (2) reviewed the mechanistic information on alpha-particle-induced carcinogenesis. The NRC (2) postulated that the biological effects of exposure to radon at the low level of typical indoor air are initiated by the passage of single alpha-particles that have very high linear energy transfer (LET). Hei et al. (25) described the mutagenesis of a single alpha-particle in mammalian cells. Genes involved at low dose are different from those involved at higher doses and adaptive responses in cells exposed to low-level ionizing radiation allow the cells to better withstand a much higher dose later (priming effect) [http://www.llnl.gov/str/JulAug03/Wyrobek.html].

The NRC (2) developed a new model of ingested radon diffusion because its review of published ingestion risks found a wide range in the estimates. The uncertainties were mainly due to the assumed model of radon absorption in the gastrointestinal tract. The NRC (2) noted the lack of data on the extent to which radon diffuses through the stomach wall and recommended research to reduce the uncertainty regarding ingestion risks. Details of the mathematical models of radon diffusion in the stomach and the physiologically based pharmacokinetic (PBPK) model of radon behavior in the body can be found in the NRC (2). The NRC (2) estimated the ingestion risks using estimates of the amount of water ingested, the fraction of radon lost from water before ingestion, and calculations of the dose absorbed by tissues exposed. The NRC (2) estimated the confidence limits on the ingestion risk but noted, "the actual risk from ingested radon could be as low as zero depending on the validity of the LNT dose–response hypothesis." Yamaoka el al. (2004) found inhalation of radon gas from radon enriched water induced a small amount of active oxygen in the blood, decreased the levels of lipid peroxide and total cholesterol in the blood of osteoarthritis patients alleviating pain related to peroxidation reactions and immune depression.

Table 1 is a summary of the risks estimated by the NRC (2). Table 1 shows that males have higher lifetime risks than females and that the inhalation risk to ever-smokers is five times as high as those for never-smokers of either sex. Notice that the aggregate health risk to the general population is thrice as high as the risk for never-smokers due to the inclusion of ever-smokers. Figure 1 shows the relative distribution of cancer fatalities by exposure pathways. Notice in Fig.1 that 89% of the health risks are from inhalation of the radioactive by-products of radon released from water into indoor air. About 11% of the health risk is from direct ingestion of tap water (0.6 L/day directly from a tap). An earlier risk assessment by the EPA (5) estimated that the proportion of ingestion and inhalation risks were 53% and 47%, respectively.

Both the NRC (2) and the EPA (5) found that the hazards of exposure to radon during showering are likely to be small because the buildup of radon progeny is less than 4% of the maximum possible value during a typical shower that lasts 10 minutes. The NRC (2) found it highly unlikely that radon in drinking water has substantial teratologic or reproductive effects. The NRC (2) found that information was insufficient to determine radon risks to "sensitive" sub-populations, such as children. However, about 30% of fatal stomach cancers are attributed to exposure before age 10 (2). The NRC (2) suggested that there might be an opportunity interval for lung-cancer risk reduction from exposure during childhood. Bilban and Vaupotic (26) found higher chromosome aberrations in pupils at elementary schools that had high indoor radon levels. Fewer than 34% of lung cancer patients are diagnosed before age 65 (27).

Many researchers have estimated fatalities from radon in domestic water. Mills (28) estimated 60, 400, and 5000 annual fatalities from radon in surface water, public groundwater, and private wells sources, respectively. The NRC (2) used the risk estimates (Table 1) and information regarding the occurrence of radon in CWS, daily water

Table 1. Lifetime and Unit Risks Posed by Exposure to Radon in Drinking Water at 1 pCi/L (37 Bqm^{-3})[a]

Exposure Pathway	Lifetime Risk[b] Male	Female	U.S. Population	Unit Risk U.S. Population
Inhalation (ever-smokers)	1.15×10^{-6}	7.40×10^{-7}	9.62×10^{-7}	1.31×10^{-8}
Inhalation (never-smokers)	2.18×10^{-7}	1.48×10^{-7}	1.85×10^{-7}	2.43×10^{-9}
Inhalation (population)	7.77×10^{-7}	4.44×10^{-7}	5.92×10^{-7}	7.90×10^{-9}
Ingestion* (population)	5.55×10^{-8}	8.51×10^{-8}	7×10^{-8}	9×10^{-10}
Total risk (inhalation and ingestion)	8.14×10^{-7}	5.18×10^{-7}	6.66×10^{-7}	8.89×10^{-9}
*Breakdown of Ingestion Risks				
Stomach			5.92×10^{-8}	
Colon			2.18×10^{-9}	
Liver			2.15×10^{-9}	
Lung			1.26×10^{-9}	
General tissue			2.92×10^{-9}	
Total ingestion			7×10^{-8}	

[a] Data sources: Reference 2, pp. 16, 17; Reference 5, Table X11.3.

[b] Annual unit risk can be obtained by dividing lifetime unit risks by the expected lifetime for each population category. Expected lifetime in years: Ever-smokers = 73.7; never-smokers = 76.1; combined population = 74.9.

Table 2. Health Risks from Radon in Drinking Water

a. Estimated Annual U.S. Fatalities Due to Radon in *Drinking Water

Health Risk	Source of Estimate	Central Estimate	Range
Lung cancer	NRC (2)	160	25–290
Lung cancer	EPA (5)	148	70–263
Stomach cancer	NRC (2)	23	5–50
Stomach cancer	EPA (5)	18	3–38

b. Occurrence of Lung Cancer due to Radon by Exposure Pathway

Pathway	Central Estimate	Range
Outdoor air (2)	720	120–300
Indoor air (5)	18200	3,000–33,000
*Drinking Water (2)	160	25–290
*Drinking Water (5)	148	70–263

c. Comparative Health Risks of Radon in *Drinking Water

Health Risk	Annual Fatalities (Central Estimates) All Causes	*Drinking Water	% of All Causes Due to *Drinking Water
Lung cancer (2)	160,000	160	0.10%
Stomach cancer (2)	14,000	18	0.13%

Data sources: NRC(2)page 18; EPA(5).

Fatalities based on 1998 data.

*Potable water delivered by a community water system (CWS).

intake, and smoking patterns among several factors to derive estimates of cancer fatalities in the United States. Table 2 summarizes the estimated U.S. fatalities due to radon in domestic water relative to other sources. The NRC estimated that in 1998 radon from soil gas caused about 18,200 lung cancer fatalities in the United States. The NRC (2) also estimated that 160 lung cancer fatalities were due to radon in domestic water. Notice that an estimated 720 fatalities were from an unavoidable background level of radon in outdoor air. The NRC (2) estimated that 23 stomach cancer fatalities were from ingesting radon in drinking water. The EPA (5) estimated that radon in domestic water caused 148 lung cancer and 18 stomach cancer fatalities annually. The EPA (5) also estimated (screening level) that an additional 36 fatalities annually were from radon in water delivered outside residences, at schools, offices, etc.

RISK MANAGEMENT

The EPA (8) first proposed 300 pCi/L and 0 pCi/L as the maximum contaminant level (MCL) and MCL goal (MCLG), respectively, for radon in drinking water. Many criticized the proposed rule because it was not considered cost-effective. In addition, a multimedia approach could be more cost-effective and achieve greater health risk reduction. In 1996, as part of the amendments to the SDWA, Congress directed the EPA to withdraw the proposed rule and re-propose using an updated risk assessment provided by the NRC. Congress also directed the EPA to develop an alternate MCL (AMCL) and guidelines for multimedia mitigation (MMM) programs (MMMP). The NRC (2) defined the AMCL as the concentration of radon in water that would cause an increase of radon in indoor air that is no greater than the level of radon naturally present in outdoor air. The NRC (2) determined that the AMCL should be about 4000 pCi/L based on the national average outdoor air radon concentration of 0.4 pCi/L and the water to air transfer factor of 10,000. EPA is only to establish an AMCL if the MCL is more stringent than the water equivalent of the national average of radon in outdoor air.

The EPA (5) used NRC risk estimates (Table 1) to repropose the radon rule (MCL = 300 pCi/L; MCLG = 0 pCi/L, AMCL = 4000 pCi/L). CWS whose radon levels are within the range of the MCL and the AMCL have the option of complying with either the MCL or the less stringent AMCL. CWS that choose to comply with the AMCL must participate in an approved and verifiable MMMP that achieves at least an equivalent level of risk reduction as treatment to the MCL. CWS may opt to develop individual MMMP in accordance with EPA guidelines if the primacy state chooses not to develop an MMMP. The EPA (5) proposed four criteria required for approval of MMMP: involve the public in developing MMMP; include quantitative goals to measure the effectiveness of MMMP; include implementation plans, strategies and specific activities for achieving quantitative goals; and include plans for measuring and reporting results. Examples of MMMP include mitigation of radon levels in existing construction and use of radon-resistant techniques for new construction. CWS that choose the AMCL approach are not required to treat, but to participate in approved MMMP. The implemented MMMP might not reduce the ingestion risk, but reduce the inhalation risks, though not necessarily in the same community, nor in the same homes. Some CWS systems may opt to purchase water from other sources or consolidate with other CWS in lieu of treating or implementing the MMMP.

HEALTH RISK REDUCTION AND COST–BENEFIT ANALYSES (HRRCA)

Radon is the first drinking water contaminant for which the EPA (27) used the new cost–benefit analyses framework. The HRCCA (27) included detailed cost–benefit analyses by water system size for each considered MCL and MMMP adoption. In estimating benefits, the EPA (27) used the value of $5.8 million (1997 dollars) as the value of statistical life (VSL) for each avoided cancer fatality and willingness to pay (WTP) of $536,000 for nonfatal cancers. The EPA (5) estimated that the cost for MMMP is $0.7 million per avoided lung cancer (the estimated average costs of indoor radon screening for all homes and mitigation cost for each home at or above the 4-pCi/L indoor air action level). The EPA (5) estimated annualized national costs at $408 million and monetized annual benefits at $362 million if 100% of impacted CWS choose to comply with the proposed MCL and assuming costs of capital of 7%. The estimated national annualized costs and monetized benefits are $60 million and $301 million, respectively, under a scenario that assumes that 95% of states implement MMMP and 90% of the CWS in each state comply with the AMCL. The EPA (5) prefers that states adopt the AMCL approach and develop MMMP but the NRC (2) noted that noneconomic considerations, such as equity, could be deciding factors in whether to undertake MMMP. The EPA (5) estimated that annual MCL compliance costs per household are about $270 and $10 for CWS serving 25 to 100 people and more than 100,000 respectively. The HRRCA shows that small CWS would incur most of the treatment costs but that the majority benefits would go to larger systems under the MCL approach. Some have argued that the EPA underestimated the costs of compliance and overstated the benefits. Costs could be higher because of unique local conditions and environmental permitting. Hahn and Burnett (29) found that the proposed radon rule was unjustifiable based on cost–benefit analyses.

TREATMENT TECHNOLOGIES

CWS may choose to comply with either the MCL or, under certain circumstances, with the AMCL. Treatment of radon in domestic water will reduce the health risks from both direct ingestion and inhalation through a decrement. Deb (30) reported on the decrements in indoor radon level after water treatment in three New Hampshire communities. Deb (30) found that significant reduction in the indoor radon level, occurred only in homes served by the CWS where water (radon concentration >20,000 pCi/L in water) was the major source of radon in indoor air.

The EPA (5) identified aeration as the best available technology (BAT) for removing radon from domestic water in large CWS. Article DW-32 (31) contains additional information regarding radon removal from drinking water. Aeration treatment methods such as a packed tower, diffused bubble, spray, slat tray, and cascade aeration are very efficient in removing radon from water because of the volatility of radon gas. Liquid-phase granular activated carbon (GAC) removes radon by adsorption but requires longer empty bed contact time (EBCT) ranging from 20 to 130 minutes, and its performance depends on water quality (1). The EPA (5) identified point-of-entry (POE) GAC as one of the affordable small systems compliance technologies (SSCTs). Drago (1) reported removal efficiencies of 20% to greater than 99% at 34 water treatment facilities and found that construction costs for either BAT are site specific. Either treatment technology requires disinfection after treatment, which EPA believes, will reduce the microbial risk and provide the additional benefit of complying with the proposed groundwater rule (32). The EPA (5,27) found that high levels of dissolved iron and manganese co-occurring with radon in water may complicate and increase radon treatment costs.

The EPA (5) reviewed the permitting and other environmental issues of either technology. Beta-gamma radiation from radon progeny might pose concern to GAC operators (1). A major concern of aeration technologies is off-gas emissions. Fisher et al. (33) found that radon concentration in air could reach levels considered unsafe for underground miners at some water treatment facilities that aerate water as part of the treatment process. The NRC (2) and EPA (5) recommended monitoring for radon in air around facilities treating groundwater for any contaminant and mitigation, if necessary, to reduce exposure. The EPA (5) modeled risks from potential air emissions of radon off-gas from aeration facilities based on data from 20 CWS whose radon levels ranged from 1000 pCi/L to 16,000 pCi/L. The EPA (5) estimated a maximum lifetime risk of 2×10^{-5} and a crude national estimate of 0.09 cancer fatalities per year if all CWS comply with the MCL. An air dispersion modeling study for the East Valley Water District, San Bernardino, California, suggests that the PTA (Packed Tower Aeration) downwind risk estimates could exceed the local regulatory risk threshold (34). The CDM (34) model predicted that a maximum individual cancer risk at some locations near the PTA could exceed 10^{-5} and 0.2 as the mean cancer burden for the service area. Preliminary design parameters for PTA to mitigate risks from radon emissions were not feasible from design/construct perspectives, for example, stack height up to 200 feet (34).

CONCLUSIONS

The preferred risk management model for radon in domestic water differs from those of other regulated drinking water contaminants because

- Amendments to the SDWA in 1996 require the EPA Administrator to take into account the costs and benefits of control programs for radon from other sources in setting the MCL.
- The major source of radon in indoor air is not regulated. Domestic water, which contributes on average less than 2% of radon in indoor air, would be regulated. Areas that have the highest indoor radon level may not coincide with service areas of CWS that have the highest level of radon in domestic water.
- The majority of health risks is not from direct ingestion of water, but rather from radon released during domestic use. Drinking water standards are usually set based on risks primarily due to ingestion.
- The majority of health risks are due to synergy with voluntary lifestyle choice—smoking. An ongoing public health antismoking campaign may result in significant reductions in the population of ever-smokers, thus reducing overall population unit risks. Though the EPA did not specifically categorize ever-smokers as a “sensitive subpopulation,” their inclusion in risk analyses and benefit analyses significantly increased the benefits of adopting the rule.
- CWS may choose to comply with the AMCL rather than the MCL. In some scenarios, health benefits from risk reduction might accrue to communities outside the service area of the CWS. Health benefits might accrue to selected individuals rather than uniform risk reduction to all consumers.

Ayotte et al. (35) estimated that a 0.05% reduction in the prevalence of smoking would prevent as many deaths from lung cancer as mitigation of indoor radon. The NRC (2) noted “except in situations where concentrations of radon in water are very high, the reduction of radon in water will generally not make a substantial reduction in the total radon-related health risks to occupants of dwellings served by the water supply.” The NRC (2) expressed similar concerns about the lack of data regarding the “overall effects of radon-resistant construction methods on reducing concentrations of radon in indoor air radon concentrations.”

BIBLIOGRAPHY

1. Drago, J.A. (1998). *Critical Assessment of Radon Removal Systems for Drinking Water Supplies*. American Water Works Association Research Foundation, Denver, CO.
2. National Research Council. (1999). *Risk Assessment of Radon in Drinking Water*. National Academy Press, Washington, DC.
3. U.S. Geological Survey. (2000). *The Geology of Radon*. U.S. Geological Survey, Denver, CO.
4. United Nations Scientific Committee on the Effects of Atomic Radiation. (1993). *Report to the General Assembly*. United Nations, New York. Available: http://www.unscear.org/reports/2000_1.html
5. U.S. Environmental Protection Agency. (1999). *National Primary Drinking Water Regulations; Radon-222: Proposed Rule*. 40 CFR Parts 141 and 142. Federal Register (November 2, 1999) **64**(211): 59246–59378. [64 FR 59246]. U.S. Govt. Printing Office, Washington, DC.
6. Cole, L.A. (1993). *Element of Risk: The Politics of Radon*. AAAS Press, Washington, DC.
7. World Health Organization. (2002). *Radon and Health*. Final Information Sheet. October 2002. Available: http://www.who.int/ionizing_radiation/env/Radon_Info_sheet.pdf
8. U.S. Environmental Protection Agency. (1991). *National Primary Drinking Water Regulations; Radionuclides: Proposed Rule*. 40 CFR Parts 141 and 142. Federal Register (July 18, 1991) **56**(138): 33050–33127. [56 FR 33050]. Govt. Printing Office, Washington, DC.
9. Makinde-Odusola, B. (2004). MW-66 drinking water quality standards—United States. *Encyclopedia of Water*. John Wiley & Sons, New York.
10. U.S. Environmental Protection Agency. (1999). *Methods Occurrence, and Monitoring Document for Radon*, Draft. U.S. Environmental Protection Agency, Office of Ground Water and Drinking Water.
11. Zikovsky, L. and Chah, B. (1990). The lognormal distribution of radon concentration in ground water. *Ground Water* **28**(5): 673–676.
12. Field, R.W., Fisher, E.L., Valentine, R.L., and Kross, B.C. (1995). Radium-bearing pipe scale deposits: implications for national waterborne radon sampling methods. *Am. J. Public Health* **85**: 567–570.

13. Krewski, D. et al. (2002). A combined analysis of North-American case control studies of residential radon and lung cancer: An update. *Radiat. Res.* **158**(6): 785–790.
14. National Council on Radiation Protection and Measurements. (2001). *Evaluation of the Linear -Nonthreshold Dose-Response. Model for Ionizing Radiation NCRP Report No. 136*. NCRP, Bethesda, MD. Available: http://www.ncrp.com/rpt136.html.
15. Lorimore, S.A. et al. (1998). Chromosomal instability in the descendants of unirradiated surviving cells after a-particle irradiation. *Proc. Natl. Acad. Sci. USA* **95**: 5730–5733.
16. Brenner, D.J. and Sachs, R.K. (2003). Domestic radon risks may be dominated by bystander effects—but the risks are unlikely to be greater than we thought. *Health Phys.* **85**(1): 103.
17. Parsons, P.A. (2002). Radiation hormesis: Challenging LNT theory via ecological and evolutionary considerations. *Health Phys.* **82**(4): 513.
18. Cohen, B.L. (1995). Test of the linear-no threshold theory of radiation carcinogenesis for inhaled radon decay products. *Health Phys.* **68**(2): 157–174.
19. Puskin, J.S. (2003). Smoking as a confounder in ecologic correlations of cancer mortality rates with average county radon levels. *Health Phys.* **84**(4): 526–532.
20. National Research Council. (1999). *Health Effects of Exposure to Radon (BEIR VI)*. National Academy Press, Washington, DC.
21. Field, R.W. and Duport, P.J. (2003). Exposure to residential radon causes lung cancer. W.R. Hendee, Moderator, Point/Counterpoint. *Med. Phys.* **30**(4): 485–488.
22. Centers for Disease Control. (1995). *Morbidity Mortality Wkly. Rep.*, Cigarette smoking among adults—United States 1993.
23. Alavanja, M.C. (2002). Biologic damage resulting from exposure to tobacco smoke and from radon: Implication for preventive interventions. *Oncogene* **21**(48): 7365–7375.
24. U.S. Environmental Protection Agency. (1994). *Report to the United States Congress on Radon in Drinking Water, Multimedia Risk and Cost Assessment of Radon*. Office of Drinking Water, Washington, DC.
25. Hei, T.K. et al. (1997). Mutagenic effects of a single and an exact number of α particles in mammalian cells. *Proc. Natl. Acad. Sci. USA* **94**: 3765–3770.
26. Bilban, M. and Vaupotic, J. (2001). Chromosome aberrations study of pupils in high radon level elementary school. *Health Phys.* **80**(2): 157–163.
27. U.S. Environmental Protection Agency. (1999). *Health Risk Reduction and Cost Analysis for Radon in Drinking Water Regulations: Notice and Request for Comments and Announcement of Stakeholder Meeting*. 40 CFR Parts 141 and 142. Federal Register (February 26, 1999) **64**(38): 9559–9599. [64 FR 9559], U.S. Govt. Printing Office, Washington, DC.
28. Mills, W.A. (1990). Risk assessment and control management of radon in drinking water. In: *Radon, Radium and Uranium in Drinking Water*. C.R. Cothern and P.A. Rebers (Eds.). Lewis, Chelsea, MI, pp. 27–37.
29. Hahn, R.W. and Burnett, J.K. (2001). *The EPA's Radon Rule: A Case Study in How Not to Regulate Risks*. Regulatory Analysis 01-01. AEI-Brookings Joint Center for Regulatory Studies. Available: http://www.aei.brookings.org
30. Deb, A.K. (1992). *Contribution of Waterborne Radon to Home Air Quality*. American Water Works Association Research Foundation, Denver, CO.
31. DW-32. (2004). Methods of Reducing Radon in Drinking Water. *Encyclopedia of Water*. John Wiley & Sons, New York.
32. U.S. Environmental Protection Agency. (2000). *National Primary Drinking Water Regulations: Groundwater Rule; Proposed Rule 40 CFR Parts 141 and 142*. Federal Register (FR) **65**(91): 30194–30274). Available: http://www.access.gpo.gov/su_docs/aces/aces140.html (accessed Jan 2001).
33. Fisher, E.L., Fuortes, L.J., and Field, R.W. (1996). Occupational exposure of water-plant operators to high concentrations of radon-222 Gas. *J. Occup. Environ. Med.* **38**: 759–764.
34. CDM. (1999). *Radon Removal Preliminary Cost Analysis and Risk Assessment*. Prepared for the East Valley Water District, San Bernardino, CA.
35. Ayotte, P. et al. (1998). Indoor exposure to ^{222}Rn: A public health perspective. *Health Phys.* **75**(3): 297–302.
36. Yamaoka et al. (2004). Study on biologic effects of radon and thermal therapy on osteoarthritis. *Journal of Pain 2004* **5**: 20–25.
37. Van Pelt, W.R. (2003). Epidemiological associations among lung cancer, radon exposure and elevation above sea level–a reassessment of Cohen's county level radon study. *Health Physics* **85**(4): 397–403.

Note: All referenced URLs were valid as of July 2003.

SILICA IN NATURAL WATERS

Robert Y. Ning
King Lee Technologies
San Diego, California

INTRODUCTION

Silica (1–4) and silicates (5–7) are generic names given to the family of silicon dioxide compounds derived from the dehydration–polymerization of silicic acid [$Si(OH)_4$]. Silicic acid, variously called orthosilicic acid, monosilicic acid, or silicic acid monomer, is a weak tetravalent acid that has pK_a values of 9.9, 11.8, and 12. Although silicic acid is prevalent in natural waters and is readily available from acidification of commercial crystalline sodium silicate or metasilicate (8), it has a great propensity to polymerize by dehydration to form Si–O–Si anhydride bonds:

$$\begin{aligned} \mathrm{n\ Si(OH)_4} &\longrightarrow \mathrm{(OH)_3Si{-}O{-}Si(OH)_3\ dimer} \\ &\longrightarrow \text{Oligomers} \longrightarrow \text{Colloidal polymers} \\ &\longrightarrow \mathrm{(SiO_2)_n} \end{aligned}$$

Thus silicon dioxide (SiO_2, silica) is the final and most stable product in normal conditions of silico-oxygen acid polymerization. The molecule of silica represented by the formula SiO_2 is polymeric and is more accurately represented by the formula $(SiO_2)_n$, where n is a large number.

The rate of silicic acid polymerization is strongly pH-dependent and is influenced by silicic acid concentration. It is very fast in neutral and slightly alkaline solutions, and extremely slow at low pH values of 2 to 3. Dilute polysilicate solutions that are undersaturated with respect to amorphous silica depolymerize to monosilicic acid.

Silicates imply more complex forms of silica in which silicic acid co-polymerizes with hydroxides of other metal elements covering almost the entire periodic table. This gives rise to the immense variety of rocks found on earth, which are copolymers of silicic acid in ultimate dehydrated forms. In solution, silicic acid readily reacts with alcohols and other hydroxyl-group containing organic compounds to form anhydrides. This provides for a great diversity of products and interactions in the biosphere.

CLASSIFICATION

Silica in natural waters can generally be classified in three categories: reactive soluble, nonreactive soluble (colloidal), and nonreactive insoluble (particulate). Reactivity refers to staining by molybdate reagent in the colorimetric assay of dissolved silica. The rate of this staining reaction is quantitatively reproducible only with low molecular weight polysilicic acids and varies inversely with the degree of polymerization of silicic acid (9). Monosilicic acid reacts with molybdic acid under assay conditions within 2 minutes, and disilicic acid reacts completely in about 10 minutes. Higher polysilicic acids require much longer times for substantial but incomplete reactions. After removal of suspended insoluble silica by filtration using typically a 0.45 micron filter, total dissolved silica (reactive + colloidal) can be measured by atomic emission or absorption spectroscopy. The colloidal silica value is the difference between the total and reactive silica. Normal detection limits of reactive silica in 25 mL samples are 0.1 mg/L. A greater sensitivity of 0.001 mg/L is possible in 100 mL sample sizes (9).

Agglomerated silica and silicates can reach a size, in solution, where they are no longer ions, but are actually colloidal particles that can remain in suspension indefinitely and are very difficult to remove by filtration.

GEOCHEMISTRY

Silica (SiO_2) in nature occurs in the crystalline forms of quartz, cristobalite, and tridymite, and also in the form of amorphous silica and glass. In number of atoms as well as weight, silicon is exceeded only by oxygen in the earth's crust. More than 95% of the volume of the earth's crust is composed of quartz and a small number of other rock-forming silicates.

Second only to carbon, silicon forms the largest number of compounds with other elements. The large number of carbon compounds is due to the fact that the bond energies of the C–C, C–O, and C–H bonds are of about equal magnitude, and they will therefore be formed with about the same probability. In contrast, the bond energy of the Si–O bond is considerably higher than that of the Si–H bond and more than twice that of the Si–Si bond. As a consequence, instead of the common C–C–C chains of carbon chemistry, chains of the type Si–O–Si–O are the skeletons of silicon chemistry (5).

Typical concentrations of dissolved silica in natural waters are:

Lakes	1–3 ppm
Major rivers	3–15 ppm
Seawater	1–10 ppm
Wells	2–60 ppm
Wells in volcanic and oil fields	50–300 ppm

In river and seawater, silica is not found exclusively in the colloidal state, but also in molecular solution, because its saturation concentrations are not attained. Despite the propensity for monosilicic acid to polymerize, it has been shown that silica in seawater is nearly exclusively in the monosilicic acid form in concentrations in the range of <0.006 to 10.1 ppm (11, 12). The concentration is low at the surface and increases with depth. Samples of lake and river water contain only monosilicic acid; there was no evidence of the presence of dimeric silicic acid (13). Silica in ground waters also appears to be in low molecular weight forms rather than the colloidal form (14). The concentration of particulate silica in seawaters of several oceans has been determined as 0.0006 to 0.43 ppm SiO_2 (12). It is thought that it is mainly of biogenic origin. Colloidal silica is depolymerized by seawater, which is considered at least one order of magnitude undersaturated with respect to amorphous silica.

DISSOLUTION OF SILICATE ROCKS

Pure silica such as quartz is completely dehydrated silicic acid. Depolymerization by rehydration as represented below represents dissolution to form soluble silicic acid:

$$SiO_2(s) + 2H_2O \longrightarrow Si(OH)_4$$

The dissolution rate depends on temperature and pH, it is minimum at pH = 2 to 3 and rises rising about 100-fold at pH = 11. Heat-accelerated hydrolysis would be expected in volcanic regions.

Silicic acid is also generated from the weathering of silicate rocks. Silicates are basically dehydrated forms of metal ion salts of silicic acid, which can be decomposed by carbonic acid. The following two reactions represent important dissolution processes of silicates (1):

$$2NaAlSi_3O_8 + 2CO_2 + 6H_2O \longrightarrow Al_2Si_4O_{10}(OH)_2 + 2Na^+ + 2HCO_3^- + 2Si(OH)_4$$

$$KMgFe_2AlSi_3O_{10}(OH)_2 + 1/2O_2 + 3CO_2 + 11H_2O \longrightarrow Al(OH)_3 + 2Fe(OH)_3 + K^+ + Mg^{2+} + 3HCO_3^- + 3Si(OH)_4$$

Field measurements of the weathering rate on a per square meter catchment area basis from different geographic areas are remarkably similar: 10^{-2} to 10^{-1} moles $m^{-2}yr^{-1}$. It is estimated that 10^5 m^2 surface area of mineral grains active in weathering is available per m^2 of geographic area and that only a few monolayers of rock silicates are dissolved per year (1). The silica concentrations on porewater increase continuously from the earth's surface to a depth of 3 meters. The silica concentration decreases slightly with depth either by precipitation or—under the influence of groundwater—by dilution. The total annual chemical load delivered from the continents to the oceans is estimated at 39.3×10^{14} g of which SiO_2 amounts to 4.26×10^{14} g (15).

DEPOSITION OF SILICA AND SILICATES

Most natural waters are undersaturated in amorphous silica and silicates, but unusual conditions found in geothermal waters (16) and industrial concentration or uses of waters do lead to deposition of silica and silicates. Biochemical extraction and deposition of silica by living organisms (see below) constitute a major deposition pathway.

Up to a pH of approximately 9, the solubility of amorphous silica (SiO_2) is relatively constant. It has been reported by various investigators as 100 to 150 ppm at 25 °C; the dissolved species is $Si(OH)_4$. Amorphous silica is metastable relative to quartz. The rate of crystallization of quartz is so slow at ordinary temperatures that the solubility of amorphous silica represents the upper limit of dissolved aqueous silica.

The precipitation of amorphous silica and silicates apparently proceeds in a series of steps. Polymerization by dehydration accompanied by cross-linking reactions and aggregation by van der Waal forces leads to negatively charged silica colloids. Further aggregation eventually leads to soft then hard gels.

At suitably high pHs, if $Fe(OH)_3$, $Al(OH)_3$, $Ca(OH)_2$, and $Mg(OH)_2$ are present, these hydroxides readily participate in the copolymerization resulting in the incorporation of Fe, Al, Ca, and Mg into complex anhydrous silicate structures. The strong tendency of magnesium hydroxide to react with silica, even in the monomeric state, is shown by the fact that the addition of 300 ppm of $Mg(OH)_2$ to water will reduce the soluble silica content from 42 to 0.1 ppm. The order of precipitability starting with the most readily precipitated metal is copper, zinc, manganese, cadmium, lead, nickel, silver, magnesium, and calcium. Metals that react in acidic solutions with silicates are Fe^{3+}, Co^{2+}, Zn^{2+}, and Cu^{2+} (4). Aluminum seems to occupy a unique position in combining with silica. Aluminosilicate gels (zeolites) can be prepared that have ion-exchange properties. A gel composition of $5SiO_2{:}Al_2O_3{:}Na_2O$ has an exchange capacity of 6% of its weight of CaO, and almost all of the sodium ions are exchangeable. Colloidal silica and silicates are active as flocculants and precipitate with other natural colloidal matter in rivers, lakes and the sea.

GEOCHEMICAL SILICA CYCLE

Geologically, the silica cycle is intimately associated with the carbonate cycle (see Carbonate in Natural Waters). The atmospheric CO_2 concentration of about 0.04% on a timescale of millions of years is controlled mainly by a combination of degassing from the earth's interior balanced by the uptake of carbon dioxide via the weathering of calcium and magnesium silicate minerals (17,18). The carbon dioxide in water as carbonic acid is converted to dissolved calcium and magnesium bicarbonate on land during weathering. This carbon form is stored in lakes and aquifers or is carried to the sea by rivers and precipitated there as calcium and magnesium carbonate minerals. The overall reactions shown here were first formulated by Urey (18) in 1952.

$$CO_2 + CaSiO_3 \Longleftrightarrow CaCO_3 + SiO_2$$

$$CO_2 + MgSiO_3 \Longleftrightarrow MgCO_3 + SiO_2$$

The reactions from left to right represent carbon dioxide absorption via weathering. The reactions from right to left represent degassing due to thermal decarbonation of carbonates after burial at sufficient depths in the earth's crust. Added to the Urey reactions should also be the organic carbon subcycle. It exerts an important additional control on the atmospheric carbon dioxide level. This involves the removal of carbon dioxide from the atmosphere through photosynthesis by organisms, the burial of organic matter in sediments, and the weathering of organic matter (or its thermal decomposition) which returns the carbon dioxide. The overall reaction is represented as:

$$6CO_2 + 6H_2O \Longleftrightarrow (CH_2O)_6 + 6O_2$$

Together, these reactions regulate carbon dioxide and silica on a geologic timescale.

BIOCHEMICAL ASPECTS

The importance of silicon for life on the earth results primarily from the fact that the fertility of soil relies in large part on the ability of clay minerals to absorb and release water and several cations that are indispensable for plant nutrition. This process is fundamental to the life of plants and based on them, to the life of those animals which in turn feed upon plants. Data on the physiology of silicon is limited, except that in a number of species, its presence is needed for normal growth.

Dissolved silica compounds play a substantial role in the cells of living organisms. Large amounts of silica are found in horsetail, rice, feather grass, reed and bamboo, where it contributes to the strength of the leaves and stems, and in the skeletons of diatoms which consist of very pure SiO_2. Silicon compounds are vital components in the metabolism of many bacteria, especially those that live in hot springs. The bacterium *Proteus mirabilis* even substitutes silicon for phosphorous in phospholipids. Silicon is also present as a trace element in higher animals and in humans, where it constitutes about 10 mg per kg live weight. It is essential in the cells of connective tissue and is involved in the biosynthesis of collagen, the substance that forms hair and nails, and in the formation of bony tissue.

The enrichment of silicon in mineralized tissues of unicellular organisms, such as diatoms and radiolaria, and a few metaozoa, such as sponges and gastropods is spectacular (19). The mineral phase is always amorphous silica in various stages of hydration as $SiO_2 \cdot nH_2O$. Individual diatom populations require a certain level of monosilicic acid, and declining ambient silica concentrations in natural environments may influence the sequence of seasonal successions. The present day ocean contains a few ppm (parts per million) of dissolved silica; diatoms are often cited as the primary agent of silica removal from the world's ocean and keeping a low silicic acid level. Others favor inorganic extraction as the main regulating mechanism because common silicates take up dissolved silica from silica-rich seawater. The data suggest that the world ocean was low in silicon for most of its history.

BIBLIOGRAPHY

1. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry*. 3rd Edn. John Wiley & Sons, New York.
2. Iler, R.K. (1979). *The Chemistry of Silica*. John Wiley & Sons, New York.
3. Sosman, R.B. (1965). *The Phases of Silica*. University Microfilms International, Ann Arbor, MI (originally from Rutgers University Press, New Brunswick, NJ).
4. Iler, R.K. (1955). *The Colloidal Chemistry of Silica and Silicates*. Cornell University Press, Ithaca, NY.
5. Liebau, F. (1985). *Structural Chemistry of Silicates*. Springer-Verlag, New York.
6. Babushkin, V.I., Matveyev, G.M., and Mchedlov-Petrossyan, O.P. (1985). *Thermodynamics of Silicates*. Springer-Verlag, New York.
7. Eitel, W. (1954). *The Physical Chemistry of the Silicates*. University of Chicago Press, Chicago, IL.
8. Alexander, G.B. (1953). *J. Amer. Chem. Soc.* **75**: 2887.
9. Alexander, G.B. (1953). *J. Amer. Chem. Soc.* **75**: 5655.
10. Kemmer, F.N. (Ed.). (1988). *The Nalco Water Handbook*. 2nd Edn. McGraw-Hill, New York.
11. Isshiki, K., Sohrin, Y., and Nakayama, E. (1991). *Marine Chemistry* **32**: 1.
12. Kido, K. and Nishimura. (1975). *Deep-Sea Research* **22**: 323.
13. Okada, T. and Kuwamoto, T. (1985). *Anal. Chem.* **57**: 258.
14. George, S., Steinberg, S.M., and Hodge, V. (2000). *Chemosphere* **40**: 57.
15. Matthess, G., Frimmel, F., Hirsch, P., Schulz, H.D., and Usdowski, H-E. (Eds.). (1992). *Progress in Hydrogeochemistry*. Springer-Verlag, New York.
16. Henley, R.W. (1983). *Geothermics* **12**: 307.
17. Berner, R.A. (1995). *Am. J. Science* **295**: 491.
18. Urey, H.C. (1952). *The Planets: Their Origin and Development*. Yale Univ. Press, New Haven, CT, p. 242.
19. Degens, E.T. (1976). *Topics in Current Chemistry*. Vol. 64, *Inorganic Biochemistry*. Springer-Verlag, New York.

SODIUM IN NATURAL WATERS

NITISH PRIYADARSHI
Ranchi University
Ranchi, Jharkhand, India

INTRODUCTION

Sodium is a light alkali metal which floats on water in its pure state. However, it is very chemically active and is rarely found in its pure state. All natural waters contain sodium; nearly all sodium compounds readily dissolve in water, and it naturally leaches from rocks and soils.

The sodium ion, the sixth most abundant metallic ion in the earth's crust, is a natural constituent of both food and water from several sources. Most water supplies contain less than 20 mg of sodium per liter, but in some countries, levels can exceed 250 mg/liter. Saline intrusion, mineral deposits, seawater spray, sewage effluents, industrial brines, chemicals used in paper processing, soap, glass, and glass-wool industries, food processing, slaughtering and meat packing industries and salt used in road deicing all contribute significant quantities of sodium to water. In addition, water treatment chemicals, such as sodium fluoride, sodium bicarbonate, and sodium hypochlorite can together result in sodium levels as high as 30 mg/liter. Domestic water softeners can give levels of over 300 mg/liter. A much more important source of sodium in fluvial waters is the input from leaching of halite from sedimentary or evaporite deposits. More than 40% of the sodium in fluvial waters comes from this source. Sodium is also derived from the weathering of silicate minerals, principally sodic plagioclase feldspar. This source accounts for more than 20% of the sodium in river water. It is also an important element in mica, amphibole, and pyroxene.

A Na^+ restricted diet is recommended to patients suffering from hypertension or congenital heart disease; in such cases, the intake of high Na^+ through drinking water may prove critical (1). The current drinking water standard for sodium is 20 milligrams per liter.

SODIUM IN NATURAL WATER

Groundwaters

The Na content of groundwaters is a function of (1) weathering of Na plagioclase from bedrock, followed by (2) exchange of Ca^{2+} for Na^+ on the surfaces of newly formed clay minerals. Feth et al. (2) considered the weathering reactions which supplies Na^+ to groundwater as:

$$\begin{aligned}&3Na_{0.66}Ca_{0.33}Al_{1.33}Si_{2.66}O_8 + 14H_2O + 4CO_2\\&\quad\rightarrow 2H_4Al_2Si_2O_9 + 2Na^+ + Ca^{2+} + 4H_4SiO_4 + 4HCO_3^-\\&\quad\text{(kaolinite)}\end{aligned}$$

and

$$\begin{aligned}&6Na_{0.66}Ca_{0.33}Al_{1.33}Si_{2.66}O_8 + 6CO_2 + 18H_2O\\&\quad\rightarrow 3Na_{0.66}Al_{2.66}Si_{3.33}O_{10}(OH)_2 + 2Ca^{2+} + 2Na^+\\&\quad + 6H_4SiO_4 + 6HCO_3{}^-\text{(montmorillonite)}\end{aligned}$$

These theoretical calculations apply well to the actual minerals in the rocks and soils of the area studied (western U.S.)

Garrels (3) suggested that the Na^+ content of groundwater would rise due to weathering of plagioclase feldspars. If the system is closed, the waters eventually proceed through kaolinite stability. If calcite is precipitating, this montmorillonite will be sodic and hold the Na^+

content of the water relatively steady. If calcite is not precipitating, Ca^{2+} will be used by the montmorillonite, and the Na^+ content of the groundwater will continue to rise. Low concentrations of Na^+ ions are associated with the minerals, kaolinite and quartz (4).

Sathyanarayanan and Periakali (5), in their study of Na concentration in groundwater in Ultrabasic and Peninsular Gneissic Rocks of Tamil Nadu State of India, suggested that weathering of plagioclase feldspar and biotite in the Peninsular gneiss could be the source of Na in groundwater.

In short, the Na^+ content of groundwater is a function of weatherable minerals, pH, reaction time versus drainage time, the supply of Ca^{2+} to form Ca montmorillonite, sluggish drainage conditions, intensive irrigation practices, extensive use of chemical fertilizers, inadequate sanitary facilities, and precipitation dominance controlling water chemistry (6–9). In addition, groundwater can contain large amounts of Na^+ from the dissolution of evaporites, from saltwater intrusion, and smaller amounts from cyclic sea salts in original precipitation.

Rivers

The concentration of elements in river water varies with time at any single sampling site as a function of discharge, tributary supply, and groundwater discharge. In areas where rivers are draining evaporite deposits, the Na content can be very high. For example, streams in the Brazos River basin (Texas) have as much as 99,800 ppm Na and an average of 35,900 ppm (10).

The unweighted average Na content of 639 rivers and streams is 38.6 ± 5.9 ppm ($s/\sqrt{n}$). The range is 0.1 to 1.540 ppm, and the median is 6.1 ppm Na (11,12). The mean content may be used as a worldwide estimate, but the population is actually polymodal with a mean for humid and temperate weathering zones, a mean for arid climates, and an extremely high mean for rivers draining evaporite deposits (13). Comparison shows that river water and seawater are opposite in character for Na concentration. In seawater, Na>Mg>Ca; in average river water, Ca>Na>Mg (14). Livingstone's (11) weighted mean Na^+ content of world rivers was 6.3 ppm.

Lake Waters

The sodium content of lake water is a function of river and groundwater supply followed by increase due to evaporation where this process exceeds input. The evaporation may reach a stage of Na mineral precipitation, which occurs at a high level of Na^+.

Seawater

The sodium released into solution during weathering is transported by rivers to seas/oceans and represents a primary cyclic reservoir. Early knowledge of the average composition of seawater was provided by Dittmar in 1884; he carefully analyzed 77 water samples, representative of all oceans and taken both from the surface and from the depths. These samples had been collected on the voyage around the world of H.M.S. Challenger (1872–1876). He determined halides, sulfate, magnesium, calcium, and potassium. Sodium was calculated by difference by subtracting the sums of the chemical equivalents of the negative and positive ions. This procedure was followed because Dittmar was unable to achieve satisfactory direct determinations for sodium. He estimated 30.61% sodium in sea water.

In the valuable treatise, The Data of Geochemistry, (1924), F. W. Clarke calculated that the 300,000,000 cubic miles of oceanic waters owed their salinity to 4,800,000 cubic miles of dissolved salts which had been derived from the igneous rocks of the earth's crust. He confined his computations to the sodium (Na) percentages, reckoning an average of 1.14% in seawater and 2.90% of this element in igneous rocks. The average concentration of sodium in seawater is 1.1×10^7 g/L (14). Earlier, Goldberg (15) estimated 10,500 ppm Na^+ in seawater. Calculations of Garrels and Thompson (16) indicated that 99% of the Na in seawater exists as uncomplexed Na^+.

TREATMENT

Sodium may be removed from a water supply by ion exchange. The ion exchange resin in the water softener is enriched with calcium and removes the sodium from the water. Other demineralization techniques, such as reverse osmosis, distillation, electrodialysis and freeze–thaw processes, should be effective for removing sodium from water. If a large supply of water is required, it may be necessary, if feasible, to blend it with another water source low in sodium.

In addition to water treatment, another method of removing sodium from drinking water would require the construction or reconstruction of a water well. This frequently involves installing additional casing beyond the depth normally required by regulations.

BIBLIOGRAPHY

1. Holden, W.S. (1970). *Water Treatment and Examination*. J & Churchill, London.
2. Feth, J.H., Roberson, C.E., and Polzer, W.L. (1964). *Sources of Mineral Constituents in Water from Granite Rocks of the Sierra Nevada*. U.S. Geological Survey Water Supply Papers, 1535-I, pp. 1–70.
3. Garrels, R.M. (1967). Genesis of some groundwaters from igneous rocks. In: *Researches in Geochemistry*, 2nd Edn. P.H. Abelson (Ed.). John Wiley & Sons, New York, pp. 405–420.
4. Garrels, R.M. (1976). A survey of low temperature water-mineral relations. In: *Interpretation of Environmental Isotope and Hydrochemical Data in Groundwater Hydrology*. International Atomic Energy Agency, Vienna, Austria.
5. Sathyanarayanan, M. and Periakali, P. (2003). Geochemistry of groundwater in Ultrabasic and Peninsular gneissic rocks, Salem District, Tamil Nadu. *J. Geol. Soc. India* **62**(1): 63–73.
6. Gibbs, R.J. (1970). Mechanisms controlling world water chemistry. *Science*. **170**: 1088–1090.
7. Prakasa Rao, J. (1997). *Environmental Evaluation of Groundwater Quality of a Developing Urban Area of Andhra Pradesh, India*. Unpublished. Ph.D. Thesis, Andhra University, Visakhapatnam, p. 104.

8. Subba Rao, N. (1998). Groundwater quality in crystalline terrain of Guntur District, Andhra Pradesh. *Visakha Sci. J.* **2**: 51–54.
9. Ballukrya, P.N. and Ravi, R. (1999) Characterization of groundwater in the unconfined aquifers of Chennai city, India: Part I- Hydrogeochemistry. *J. Geol. Soc. India* **54**: 1–11.
10. Baker, R.C., Hughes, L.S., and Yeast, I.D. (1964). *Natural Sources of Salinity in the Brazos River, Texas*. U.S. Geological Survey Water Supply Papers. 1669-cc, pp. 1–81.
11. Livingstone, D.A. (1963) *Chemical Composition of Rivers and Lakes*. U.S. Geological Survey Professor. Papers 440-G, pp. 1–64.
12. Anderson, P.W. and George, J.R. (1966). *Water-Quality Characteristics of New Jersey Streams*. U.S. Geological Survey Water Supply Papers 1819-G, pp. 1–48.
13. Wedepohl, K.H. (1970). *Handbook of Geochemistry*. Vol. II/2, Springer-Verlag, Berlin.
14. Mason, B. and Moore, C.B. (1985). *Principles of Geochemistry*, 4th Edn. Wiley Eastern Limited, New Delhi, p. 242.
15. Goldberg, E.D. (1957). Biogeochemistry of trace metals. In: *Ecology*, vol.1. *Treatise on Marine Ecology and Paleocology*. J.W. Hedgpeth (Ed.). Geological Society America Member **67**, pp. 345–358.
16. Garrels, R.M. and Thompson, M.E. (1962). A chemical model for sea water at 25 °C and one atmosphere total pressure. *Am. J. Sci.* **260**: 57–66.

SOFT WATER

Gholam A. Kazemi
Shahrood University of
Technology
Shahrood, Iran

INTRODUCTION

At the outset, it should be pointed out that soft and hard waters are two sides of a coin. Hence, the respective articles, this one and the Hard Water article are alike in many aspects. In defining soft water, it is necessary first to describe hardness, a closely related chemical property of water. Hardness is defined by some as a property of water that prevents lather formation with soap and produces scale in hot water systems, (1); by others, as the soap consuming capacity of water (2); and in broader terms by Freeze and Cherry (3), as the "metallic ion content of water which reacts with sodium soaps to produce solid soaps or scummy residue and which reacts with negative ions, when the water is evaporated in boilers, to produce solid boiler scale."

The ions which result in, or produce, hardness include Mg^{+2}, Ca^{+2}, Sr^{+2}, Fe^{+2}, and Mn^{+2}, and to a lesser extent, Ba^{+2} and Zn^{+2} and other divalent ions. The first two of these, are the most influential. There are different forms of hardness (4):

1. Total hardness: Ca and Mg expressed as $CaCO_3$
2. Calcium hardness: Ca expressed as $CaCO_3$
3. Magnesium hardness: Mg expressed as $CaCO_3$
4. Temporary hardness: Equals alkalinity, if alkalinity is less than total hardness
5. Permanent hardness: The amount of hardness greater than temporary hardness
6. Noncarbonate hardness: Permanent hardness

Of the six different forms of hardness, the most important and the one that is commonly known as hardness, is total hardness. It is calculated by the following simplified equation:

$$\text{Hardness} = 2.5\ \text{Ca} + 4.1\ \text{Mg (units in mg/l)}$$

Note that there are equations to calculate each of the six different forms of hardness.
Example: The Ca and Mg concentrations in a water sample are 20 and 15 mg/l, respectively. What is its hardness?

$$\text{Hardness} = 2.5 \times 20 + 4.1 \times 15 = 111.5\ \text{mg/l}$$

As stated, each of the six forms of hardness can be calculated by analyzing a water sample for the relevant ion(s), followed by using the relevant equation. However, there are techniques to measure total hardness directly such as (1) a standard solution of soap, that was in use for many years (2); and (2) Schwarzenbach et al's. method of colorimetric titration (5), which is sometime referred to as the EDTA method.

The author believes that earlier, measuring hardness was a common and an inseparable practice in water analysis. To agree with this statement, one needs to look at a relatively old handbook and read, "*The common denominator of the majority of water problems is hardness*" (2, p. 4.1, first published in 1979). Somehow, in contrast, recent hydrochemical analysis of water samples does not usually include a measurement of hardness. For instance, there is not a single report of hardness in the hundreds of hydrochemistry papers presented at the International Association of Hydrogeolgists (IAH) conferences in Prague (2003) and in Munich (2001) and in the recent past (1997–2001) annual symposia of the Geological Society of Iran. This occurred, perhaps, because other aspects of water chemistry and water biology, such as nutrients, heavy metals, bacteriological species, chlorinated hydrocarbons, and many contaminants have become so widespread that they are considered more important and more hazardous than hardness. This is to say that rejecting water for drinking or other uses, these days, is many times more likely to be due to its harmful DDT level than because of its high hardness value. One other possible reason is that hardness and soft water topics are important only in some narrow fields of water, e.g., hot water boilers.

The second point is that industrialization, urbanization, agricultural practices, and most environmentally unfriendly activities are not likely to result in an increase in the hardness of water resources because these activities do not usually generate hardness-producing ions such as Ca^{2+} or Mg^{2+}, though, they may produce other trivial (in terms of hardness) ions such as Mn^{2+} and Fe^{2+}.

The author, here, wishes to draw the reader's attention to one other point and pose a question as well. He believes that "softness" would have had an identical application and a similar meaning to hardness if it would have been there. Therefore, it is interesting to speculate as to what was the rationale for use of the term "hardness", not softness. One reason might be that hard water is generally easier to notice because of its unwanted and disappointing characteristics and our predecessors looked for something to describe it. In other words, the term hardness was created to explain easily noticeable hard water, not to describe soft water. This is why there is no such practice as water hardening, changing soft water to hard water, in contrast to the well-established exercise of water softening which changes hard waters to soft waters.

DEFINITION AND SPECIFICATIONS OF SOFT WATER

Todd (6, p. 282) reports that Hippocrates (460–354 B.C.) was the first to use the terms hard and soft for water, when in his treatises on public hygiene, he wrote, "consider the waters which the inhabitants use, whether they be marshy and soft, or hard and running from elevated and rocky situations, and then if saltish and unfit for cooking... for water contributes much to health." However, it is probably safe to say that hard and soft water in Hippocrates' writings, are likely to mean different things from what we today call hard and soft water.

The most frequently used classification system for water based on hardness is that of Sawyer and McCarty (7), presented in Table 1. There is another slightly different classification system by Hem (8), which is referred to in some publications as the classification of the U.S. Geological Survey (9).

The definition of soft water varies slightly among researchers. Based on Table 1, soft water is defined as water whose hardness is less than 60 or 75 mg $CaCO_3$/l. This means that waters with less than 15 mg Ca^{2+}/l and 10 mg Mg^{2+}/l are soft; higher concentrations in waters belong to the other categories such as hard water or very hard water. However, some researchers argue that there is no exact definition for determining soft or hard water because it depends, at least partially, on what one is accustomed to (9). What is soft water for one person, may be moderately hard for another and vice versa. The very hard water category in Table 1 can be grouped into the hard water category as it has the specifications of hard water. It is, however, unclear to which category one can assign moderately hard water, if hard water and soft water are the only categories available.

Table 1. Different Types of Water Based on Hardness

Water Type	Hardness mg $CaCO_3$/l	
Soft	below 75[a]	Below 60[b]
Moderately hard	75–150[a]	60–120[b]
Hard	150–300[a]	120–180[b]
Very hard	Over 300[a]	Over 180[b]

[a] Reference 7.
[b] Reference 8.

Manning (9), discredits soft water and states that very soft water when used for bathing feels slick on the body, as though the soap is still there. He also explains that most people prefer slightly hard water because of its taste and its ability to wash the soap off the body. Soft waters are not very good for irrigation either because they may lead to a lower infiltration rate compared with slightly hard water (9). In contrast to hard water, soft water produces a foam or lather easily but does not produce scale in hot water pipes, heaters, boilers, and other high water temperature units. Therefore, using soft water is an economically worthwhile practice because it leads to a decrease in soap consumption and lowers fuel costs for boilers due to good heat conduction.

The recommended water for domestic use should have a hardness of 80 mg/l; i.e., it should be soft (10). This is why water softening is a common practice where the water supply has a hardness greater than 80–100 mg/l (3). Soft water is also required for some industrial processes such as food equipment washing, confectionery, food canning and freezing, food processing, and laundering (6).

SOFT WATER AND HEALTH

There is a general perception among the public that soft water is healthier to drink. Studies, though not satisfactorily conclusive, suggest the opposite, especially with regard to artificially softened water. For instance, Tebbutt (1995) (1) reports that there is some statistical evidence to suggest that artificially softened waters may increase the incidence of some forms of heart disease. Also, a recent study in Taiwan (11) shows that there is a significant negative relationship between drinking water hardness and colon cancer mortality, a finding considered important for the Taiwanese water industry and human health. All these confirm Muss's (12) observation that in the United States, death rates from heart and circulatory diseases are lower in states where public water supplies are harder. Furthermore, artificially softened water may taste salty because its sodium content increases in the softening process.

AVAILABILITY OF NATURALLY SOFT WATER

Soft water may exist naturally or may be produced artificially, as stated before. Atmospheric precipitations (snow and rain) are soft waters. Similar to these are waters that emanate from upland catchments with an average total hardness of 10 mg/l (1). In contrast, most groundwaters, excluding those that occur in fractured igneous rocks, are hard. All ground water and surface water that pass through, or are in contact with, carbonate and other Ca–Mg bearing rocks are hard. Lowland rivers with a hardness of 200 mg/l (1) are also hard. Therefore, only a limited percentage of freshwater resources of the world, groundwaters and river waters, is soft.

Examples of naturally occurring soft waters include the Vilnius well fields groundwater (intermorainic and alluvial deposit aquifer), Vilnius District, Lithuania, that have unbelievably low hardness values of 3.9–8.5 mg/l (13), groundwater in Akwa Ibom State of Nigeria (fluviovolcanic rocks) with a hardness of 1.2–65 mg/l (Akujieze, 2003) (14), and shallow basement groundwater of Lithembwe dambo, Malawi, with 1–14 mg/l of calcium and 1–7 mg/l of magnesium (15).

BIBLIOGRAPHY

1. Tebbutt, T.H.Y. (1995). *Principles of Water Quality Control*. Reprint of 4th Edn. Butterworth Heinemann, Oxford, UK.
2. Kemmer, F.N. (1988). *The NALCO Water Handbook*, 2nd Edn. McGraw-Hill, New York.
3. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
4. Hounslow, A.W. (1995). *Water Quality Data: Analysis and Interpretation*. Lewis, New York.
5. McCoy, J.W. (1981). *The Chemical Treatment of Boiler*. Water Chemical Publishing, New York.
6. Todd, D.K. (1980). *Groundwater Hydrology*, 2nd Edn. John Wiley & Sons, New York.
7. Sawyer, C.N. and McCarthy, P.L. (1967). *Chemistry for Sanitary Engineers*, 2nd Edn. McGraw-Hill, New York.
8. Hem, J.D. (1985). *Study and Interpretation of the Chemical Characteristics of Natural Waters*, 3rd Edn. U.S. Geological Survey Water Supply Paper 2254, Denver, CO.
9. Manning, J.C. (1997). *Applied Principles of Hydrology*. Prentice Hall, Upper Saddle River, NJ.
10. Bell, F.G. (1998). *Environmental Geology, Principles and Practice*. Blackwell Science, Cambridge, UK.
11. Yang, C.Y. and Hung, C.F. (1998). Colon cancer mortality and total hardness levels in Taiwan's drinking water. *Arch. Environ. Contamination Toxicol.* **35**(1): 148–151.
12. Muss, D.L. (1962). Relationship between water quality and deaths from cardiovascular disease. *J. Am. Water Works Assoc.* **54**: 1371–1378.
13. Klimas, A.A. (1995). Impacts of urbanization and protection of water resources in the Vilnius District, Lithuania. *Hydrogeol. J.* **3**(1): 24–35.
14. Akujieze, C.N., Coker, S.J.L., and Oteze, G.E. (2003). Groundwater in Nigeria—a millennium experience—distribution, practice, problems and solutions. *Hydrogeol. J.* **11**(2): 259–274.
15. Chilton, P.J. and Foster, S.D. (1995). Hydrogeological characterization and water-supply potential of basement aquifers in tropical Africa. *Hydrogeol. J.* **3**(1): 36–49.

SOLUBILITY OF CHEMICALS IN WATER

Carl L. Yaws
Prasad K. Narasimhan
Helen H. Lou
Lamar University
Beaumont, Texas

Ralph W. Pike
Louisiana State University
Baton Rouge, Louisiana

Correlations are presented for the solubility of chemicals in water. A wide variety of compounds are covered—alcohols, ketones, ethers, aldehydes, acids, acetates, amines, monochloroalkanes, and monobromoalkanes. The new correlation provides reliable solubility values down to very low concentrations (parts per million range). The correlation is based on the boiling point of the compound. The correlation and experimental data are in favorable agreement. The results are useful in process engineering for wastewater minimization.

IMPORTANCE OF WATER SOLUBILITY

The solubility of compounds in water is very important. This importance will increase in the future because of heath, safety, and environmental considerations. The following brief discussion illustrates the importance even at very low concentrations.

For health involving human exposure to substances in air, the threshold limit value (TLV) for *sec*-butanol in air is given as 100 ppm (parts per million, vol) by the Occupational Safety and Health Act (1). A concentration of only 0.0005 mol fraction of *sec*-butanol in water provides about 236 ppm of *sec*-butanol in air at the air–water interface. This concentration of 236 ppm in air exceeds the threshold limit value of 100 ppm.

For safety, the lower explosion limit (LEL) for *sec*-butanol in air is given as 1.7% by Yaws (2). A concentration of only 0.04 mol fraction of *sec*-butanol in water provides about 1.9% of *sec*-butanol in air at the air–water interface. This concentration of 1.9% in air exceeds the lower explosion limit of 1.7%.

For environmental, consider a spill of *sec*-butanol in contact with water. The water will become saturated with *sec*-butanol. At saturation, the solubility of *sec*-butanol in water is about 0.181 weight fraction or 0.0508 mol fraction (3). This concentration of only 0.0508 mol fraction at saturation will provide about 24,000 ppm or 2.5% of *sec*-butanol in air at the air–water interface. This concentration of 24,000 ppm or 2.4% greatly exceeds both the threshold limit value of 100 ppm and the lower explosion limit of 1.7%.

This brief discussion (using *sec*-butanol as an example) indicates that very low concentrations of a chemical in water can provide concentrations in air at the air–water interface that exceed the threshold limit value for human exposure and the lower explosion limit for flammability.

CORRELATION FOR WATER SOLUBILITY

In earlier work by Yaws and co-workers (2), the water solubility of hydrocarbons was correlated as a function of the boiling point of the compound. In this article, it was determined that the boiling point method is also applicable to the correlation of the water solubility of chemicals:

$$\log_{10}(S) = A + B\,T_B + C\,T_B^2 + D\,T_B^3 \qquad (1)$$

where S = solubility in water at 25 °C, ppm by weight,
T_B = boiling point of compound, K
A, B, C, and D = regression coefficients

Table 1. Solubility in Water[a]

	$\log_{10} S = A + B\,T_B + C\,T_B^2 + DT_B^3$				(S, ppm(wt), T_B, K)			
Chemical Family	A	B	C	D	T_{BMIN}	T_{BMAX}	Formula	Example Compound
Alcohols	45.6398	−2.3859E−01	4.8739E−04	−3.7160E−07	370	620	$C_nH_{2n+2}O$	*n*-Pentanol
Ketones	45.2000	−2.3859E−01	4.8739E−04	−3.7160E−07	350	590	$C_nH_{2n}O$	Diethyl ketone
Ethers	11.4000	−2.2000E−02	0.0000E+00	0.0000E+00	300	560	$C_nH_{2n+2}O$	Dipropyl ether
Aldehydes	12.4200	−2.2000E−02	0.0000E+00	0.0000E+00	320	560	$C_nH_{2n}O$	1-Hexanal
Acids	14.3000	−2.2000E−02	0.0000E+00	0.0000E+00	440	560	$C_nH_{2n}O_2$	*n*-Hexanoic acid
Acetates	13.0000	−2.3382E−02	0.0000E+00	0.0000E+00	330	600	$C_nH_{2n}O_2$	Ethyl acetate
Primary amines	11.4300	−1.8000E−02	0.0000E+00	0.0000E+00	400	560	$C_nH_{2n+3}N$	*n*-Hexyl amine
Secondary amines	11.3600	−1.8000E−02	0.0000E+00	0.0000E+00	360	560	$C_nH_{2n+3}N$	Dipropyl amine
Tertiary amines	11.8200	−2.0000E−02	0.0000E+00	0.0000E+00	360	560	$C_nH_{2n+3}N$	Triethyl amine
Monochloroalkanes	−7.4500	1.0050E−01	−2.7288E−04	1.9987E−07	300	590	$C_nH_{2n+1}Cl$	2-Chlorobutane
Monobromoalkanes	−7.1700	1.0050E−01	−2.7288E−04	1.9987E−07	300	590	$C_nH_{2n+1}Br$	1-Bromobutane

[a] S-solubility in water at 25 C, parts per million by weight, ppm(wt).
A, B, C, and D-regression coefficients.
T_B-boiling point temperature of compound, K.
T_{BMAX}-maximum temperature, K.
T_{BMIN}-minimum temperature, K.

The range of boiling point depends on the chemical family.

The coefficients (A, B, C, and D) are given in Table 1 for a wide variety of chemicals—alcohols, ketones, ethers, aldehydes, acids, acetates, amines, monochloroalkanes, and monobromoalkanes. The table also provides the boiling point range (T_{BMIN}, T_{BMAX}) for each chemical family. The correlation is applicable to compounds whose boiling points are in this range. The correlation should not be used for compounds whose boiling points are outside the boiling point range.

The coefficients were determined from regression of available data. In preparing the correlation, a literature search was conducted to identify data source publications (2–13). The excellent compilations by Horvath (5–6); Howard and Meylan (7); Mackay, Shiu, and Ma (9); Verschueren (11); Yalkowsky (3); and Yaws (2) were used extensively. The publications were screened and copies of appropriate data were made. These data were then keyed in to the computer to provide a database for which experimental data are available. The database also served as a basis to check the accuracy of the correlation.

Graphs of water solubility versus boiling point are shown in Figs. 1–11 for the chemical families. The graphs indicate favorable agreement of correlation values and experimental data.

For ketones, the graphs disclose that more data are available for alcohols. The graphs, also disclose that the data for ketones appear to be below and approximately parallel to the more extensive data for alcohols. This suggests that the solubility of ketones is lower and approximately parallel to the water solubility of alcohols.

For the other oxygenated compounds (ethers, aldehydes, acids, and acetates), the graphs disclose that the data for oxygenated compounds appear to be above and approximately parallel to the data for benzenes. This suggests that the solubility of oxygenated compounds is higher and approximately parallel to the water solubility of benzenes.

For the nitrogen compounds (primary amines, secondary amines, and tertiary amines), the graphs disclose

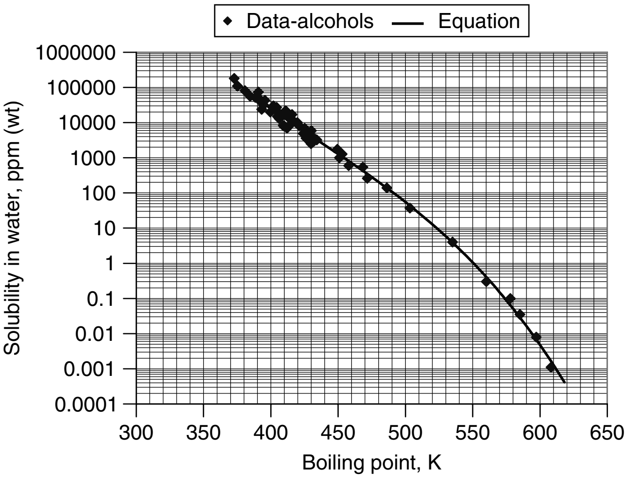

Figure 1. Solubility of alcohols in water.

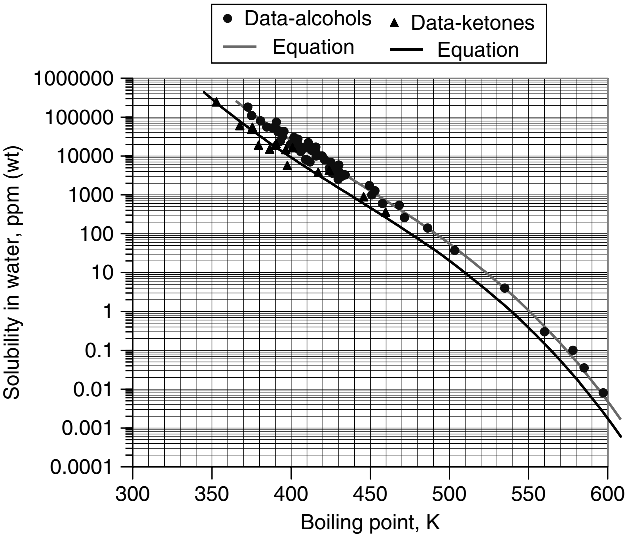

Figure 2. Solubility of ketones in water.

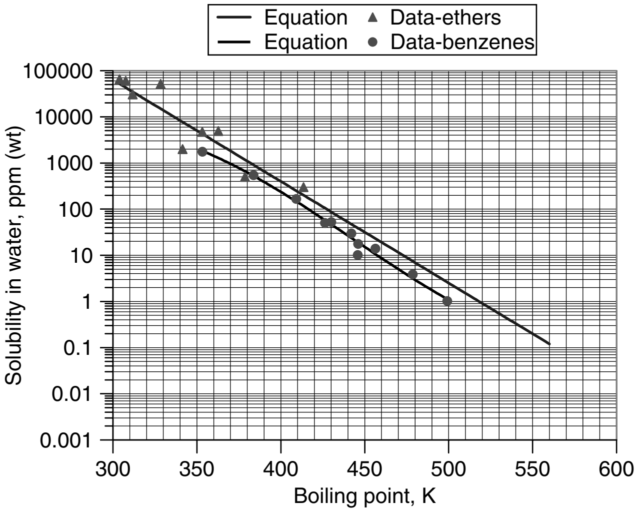

Figure 3. Solubility of ethers in water.

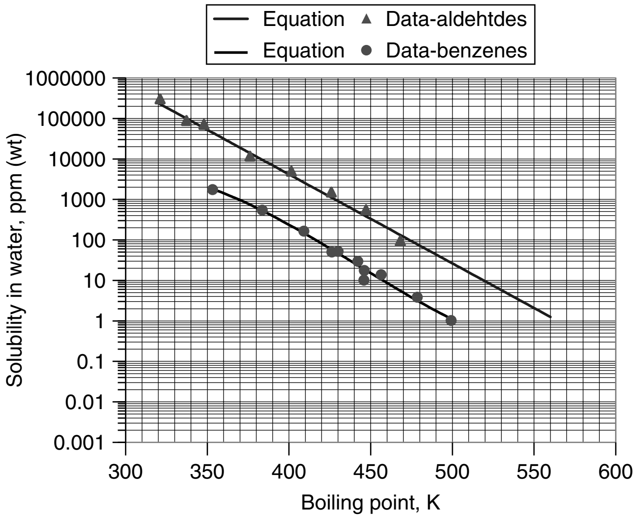

Figure 4. Solubility of aldehydes in water.

that the data for amines appear to be above and approximately parallel to the data for benzenes. This suggests that the solubility of amines is higher and approximately parallel to the water solubility of benzenes.

For monochloro and monobromoalkanes, the graphs disclose that the data appear to be lower and approximately parallel to the data for benzenes. This suggests that the solubility of monochloro- and monobromoalkanes is lower and approximately parallel to the water solubility of benzenes.

EXAMPLES

The correlations for water solubility are useful in engineering applications involving process wastewater. Examples are shown below.

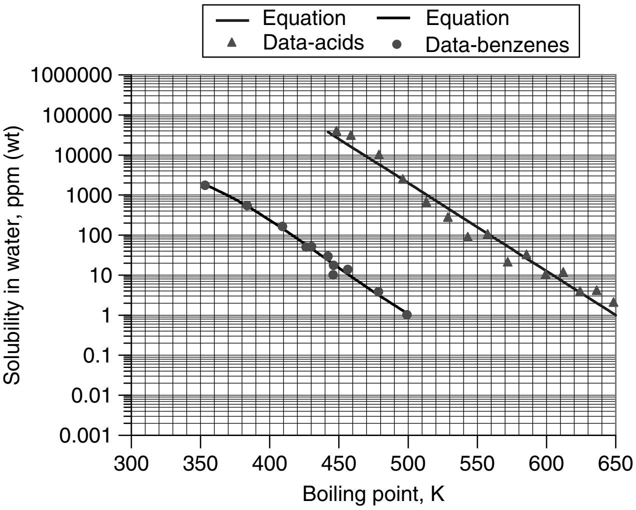

Figure 5. Solubility of acids in water.

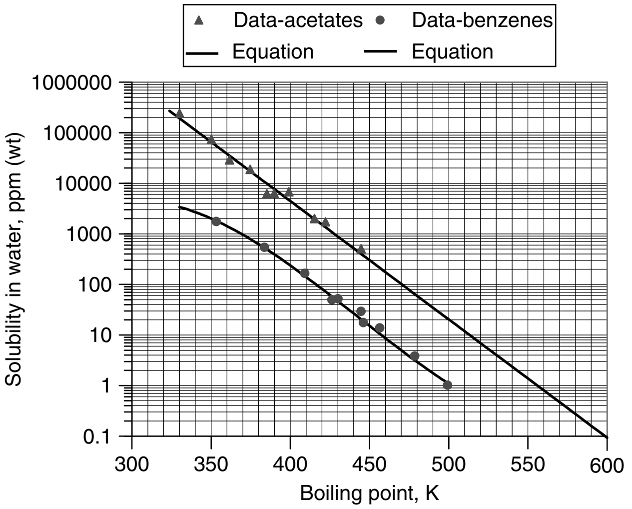

Figure 6. Solubility of acetates in water.

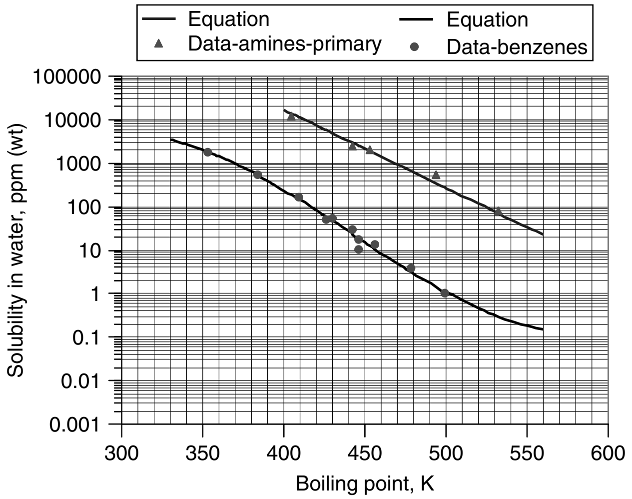

Figure 7. Solubility of primary amines in water.

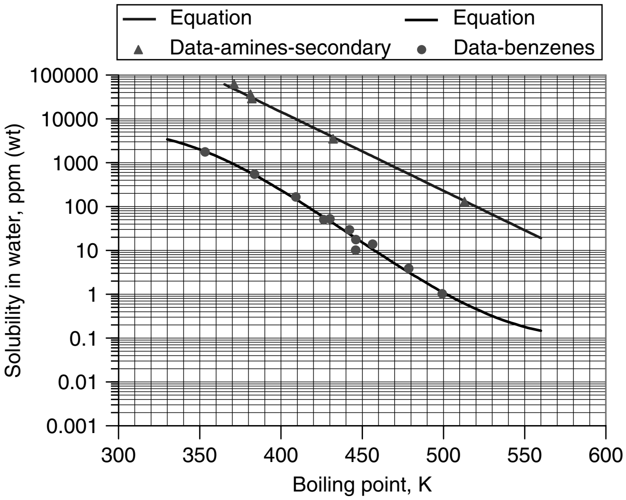

Figure 8. Solubility of secondary amines in water.

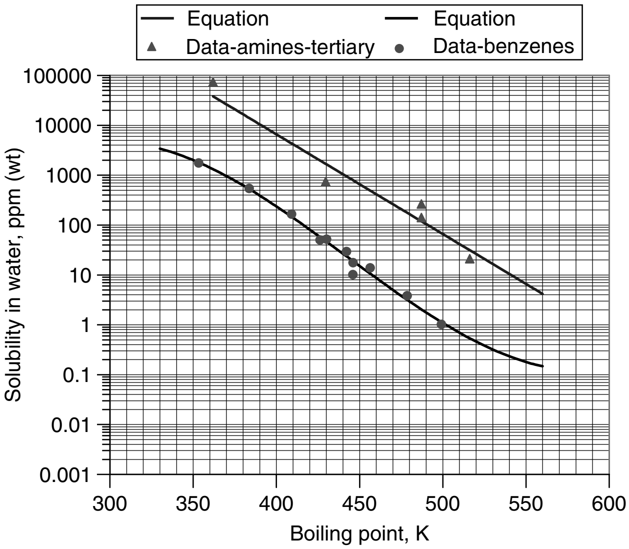

Figure 9. Solubility of tertiary amines in water.

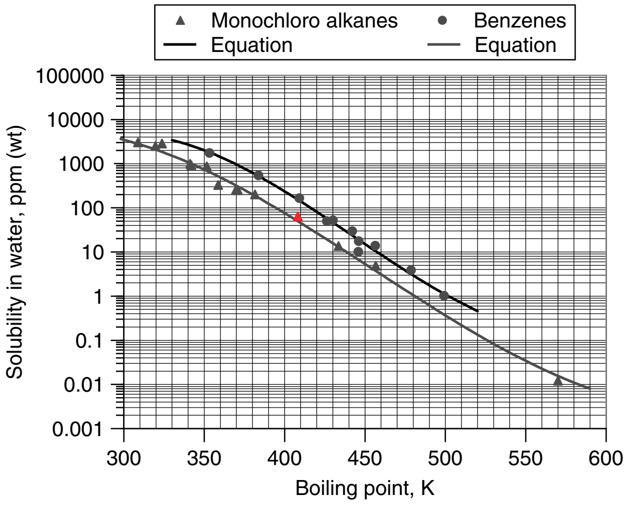

Figure 10. Solubility of monochloroalkanes in water.

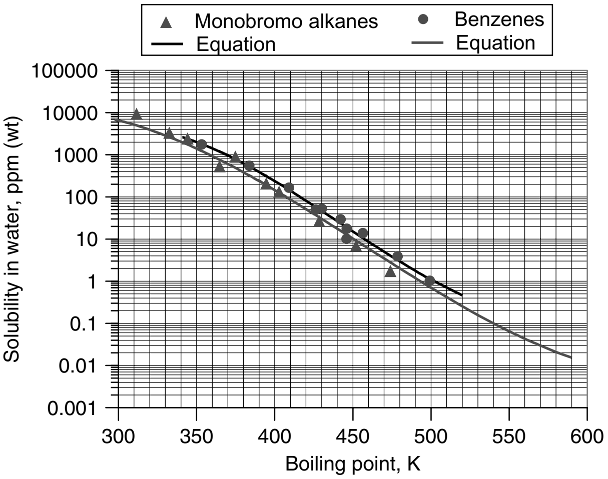

Figure 11. Solubility of monobromoalkanes in water.

Example. A chemical spill of dodecanol occurs into a body of water at ambient conditions. Estimate the concentration in the water at saturation.

The correlation for alcohols maybe used to determine the solubility in water. Substitution of the coefficients and boiling point of dodecanol in the correlation equation yields

$$\begin{aligned}\log_{10}(S) &= A + B\,T_B + C\,T_B^2 + D\,T_B^3\\ &= 45.6398 - 2.3859\text{E}{-01}(535.00)\\ &\quad + 4.8739\text{E}{-04}(535.00)^2 - 3.7160\text{E}{-07}(535.00)^3\\ \text{S} &= 3.93 \text{ ppm (wt)}.\end{aligned}$$

The calculated value and data compare favorably (3.93 vs. 4.00; deviation = 0.07/5 = 1.8%).

Example. A chemical spill of diethyl ketone occurs into a body of water at ambient conditions. Estimate the concentration in the water at saturation.

The correlation for ketones maybe used to determine the solubility in water. Substitution of the coefficients and boiling point of diethyl ketone in the correlation equation yields

$$\begin{aligned}\log_{10}(S) &= A + B\,T_B + C\,T_B^2 + D\,T_B^3\\ &= 45.2000 - 2.3859\text{E}{-01}\ (375.14)\\ &\quad + 4.8739\text{E}{-04}(375.14)^2 - 3.7160\text{E}{-07}\ (375.14)^3\\ \text{S} &= 45,257 \text{ ppm (wt)}.\end{aligned}$$

The calculated value and data compare favorably (45,257 vs. 48,235; deviation = 1,758/48,235 = 3.6%).

Acknowledgment

The Texas Hazardous Waste Research Center (Lamar University, Beaumont, Texas) provided partial support for this work.

BIBLIOGRAPHY

1. *Guide to Occupational Exposure Values*. (1997). American Conference of Governmental Industrial Hygienists, Cincinnati, OH.

2. Yaws, C.L. (1999). *Chemical Properties Handbook*. McGraw-Hill, New York.
3. Yalkowsky, S.H. (1990–2004). *Aquasol Database*. University of Arizona, Tucson, AZ.
4. Hine, J. and Mukherjee, P.K. (1975). *J. Org. Chem.* **40**: 292.
5. Horvath, A.L. and Getzen, F.W. (1999). *J. Phys. Chem. Ref. Data* **28**(3): 649–777.
6. Horvath, A.L. (1982). *Halogenated Hydrocarbons: Solubility-Miscibility with Water*. Marcel Dekker, New York.
7. Howard, P.H. and Meylan, W.M. (Eds.). (1997). *Handbook of Physical Properties of Organic Chemicals*. CRC Press, Boca Raton, FL.
8. *Liquid–Liquid Equilibrium Data Collection*. (1979). Dechema Chemistry Data Series, Frankfurt/Main, Germany, Vol. V, part 1.
9. Mackay, D., Shiu, W.Y., and Ma, K.C. (1992, 1992, 1993, 1995, 1997). *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*. Vols. 1–5, Lewis, New York.
10. Suzuki, T. (1991). *J. Computer-Aided Molecular Design* **5**: 149.
11. Verschueren, K. (1996, 2001). *Handbook of Environmental Data on Organic Chemicals*, 3rd and 4th Edn. Van Nostrand Reinhold, New York.
12. Donald, M., Shiu, W-Y., and Ma, K-C. (1995). *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*. Vol. 4, Lewis, New York.
13. (2001). *Handbook of Environmental Data on Organic Chemicals*, 4th Edn. John Wiley & Sons, New York.

SOLUBILITY OF HYDROCARBONS IN SALT WATER

Carl L. Yaws
Prasad K. Narasimhan
Lamar University
Beaumont, Texas

INTRODUCTION

Physical and thermodynamic property data are required to design and operate industrial processes. In particular, the water solubility of hydrocarbons in the sea and ocean is becoming increasingly important because of more and more stringent regulations regarding the environment.

In this article, results are presented for the solubility of hydrocarbons in salt water. The results are applicable to the complete range of salt concentrations, including water without salt to water saturated with salt. The results are intended for use in engineering applications. As an example of such usage, solubility values from the correlation are useful in determining the distribution of a hydrocarbon spill upon contact with seawater. Solubility values at other salt concentrations may also be ascertained.

CORRELATION OF SALT WATER SOLUBILITY

The correlation of the solubility of hydrocarbons in salt water containing salt is based on a series expansion in salt concentration:

$$\log_{10} S = A + B\,X + C\,X^2 \tag{1}$$

where S = solubility in salt water at 25 °C, parts per million by weight, ppm(wt)
X = concentration of salt (NaCl) in water, parts per million by weight, ppm(wt)
$A, B,$ and C = regression coefficients

The coefficients (A, B, and C) are given in Table 1. The values in the table were determined from regression of the data sources (1–10) for water solubility. The compilations by Price (7) and Sutton and Calder (8) were especially helpful. The values presented are applicable to a wide variety of hydrocarbons (alkanes, cyclopentanes, cyclohexanes, benzenes (no and single substitution), and benzenes (double and triple substitution). The range of application for the correlations is about C_5 to C_{16} compounds.

Graphs of water solubility are shown in Figs. 1–3 for representative hydrocarbons (pentane, methylcyclopentane, and benzene). In the figures, solubility values are plotted at salt concentrations ranging from water without salt to water saturated with salt. The graphs discloses favorable agreement of correlation values and experimental data.

Examples

The correlations for solubility are useful in engineering applications involving hydrocarbons in contact with water containing salt. Examples are shown below.

Example 1. A chemical spill of *n*-pentane occurs in a body of saltwater [salt = 10,000 ppm(wt)] at ambient conditions. Estimate the concentration in the saltwater.

The values in Table 1 are used for the coefficients. Introducing the boiling point of *n*-pentane to calculate A and using the tabulated values for B and C provides

$$\begin{aligned} A &= -17.7916 + 1.7865\text{E}{-01}\,T_B - 5.0255\text{E}{-04}\,T_B^{\,2} + 4.1218\text{E}{-07}\,T_B^{\,3} \\ &= -17.7916 + 1.7865\text{E}{-01}(309.22) - 5.0255\text{E}{-04}(309.22)^2 + 4.1218\text{E}{-07}(309.22)^3 \\ &= 1.5850 \end{aligned}$$

$$B = -4.5956\text{E}{-06}$$

$$C = 2.2978\text{E}{-12}$$

Substitution of these coefficients and the salt concentration in the correlation equation for solubility in saltwater yields

$$\begin{aligned} \log_{10} S &= A + B\,X + C\,X^2 \\ &= 1.5850 - 4.5956\text{E}{-06}(10{,}000) + 2.2978\text{E}{-12}(10{,}000)^2 \\ &= 1.5393 \\ S &= 34.62 \text{ ppm (wt)} \end{aligned}$$

Table 1. Solubility in Salt Water[a]

	$\log_{10} S = A + BX + CX^2$	S-solubility, ppm (wt), X(salt) – ppm(wt)	
Chemical Family	A, B, and C	Formula	Example Compound
Alkanes	$A = -17.7916 + 1.7865\text{E}{-01} T_B - 5.0255\text{E}{-04}\ T_B^2 + 4.1218\text{E}{-07}\ T_B^3$ $B = -4.5956\text{E}{-06}$ $C = 2.2978\text{E}{-12}$	C_nH_{2n+2}	n-pentane
Cyclopentanes	$A = 9.55 - 2.2987\text{E}{-02} T_B$ $B = -4.5987\text{E}{-06}$ $C = 2.2993\text{E}{-12}$	C_nH_{2n}	Methyl cyclopentane
Cyclohexanes	$A = 9.75 - 2.2987\text{E}{-02} T_B$ $B = -4.5987\text{E}{-06}$ $C = 2.2993\text{E}{-12}$	C_nH_{2n}	Cyclohexane
Benzenes (no and single substitution)	$A = 5.725 + 3.6778\text{E}{-03}\ T_B - 3.0346\text{E}{-05}\ T_B^2$ $B = -3.2714\text{E}{-06}$ $C = 1.6357\text{E}{-16}$	$C_{n+6}H_{2n+6}$	Benzene
Benzenes (double and triple substitution)	$A = 5.375 + 3.6778\text{E}{-03}\ T_B - 3.0346\text{E}{-05}\ T_B^2$ $B = -3.2714\text{E}{-06}$ $C = 1.6357\text{E}{-16}$	$C_{n+6}H_{2n+6}$	Benzene

[a] S-solubility in salt water at 25 C, parts per million by weight, ppm(wt).
X-salt (NaCl) concentration in water, ppm(wt).
Salt concentration range is $X = 0$ to 358,700 (saturated).
Salt in seawater is 34,472 ppm(wt).
A, B, and C—regression coefficients.
T_B—boiling point of compound, K.
T_{BMAX}—maximum temperature, $K = 560$.
T_{BMIN}—minimum temperature, $K = 300$.
Correlation should not be used for compounds having boiling points outside T_BMAX and T_BMIN.
Range of application for correlation is about C_5 to C_{16}.

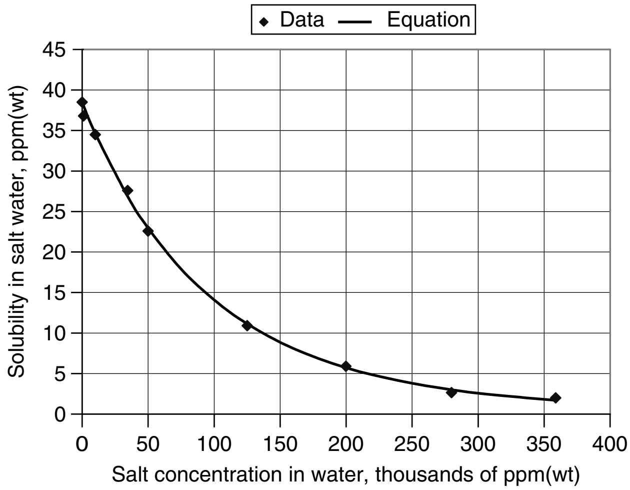

Figure 1. Solubility of n-pentane in salt water.

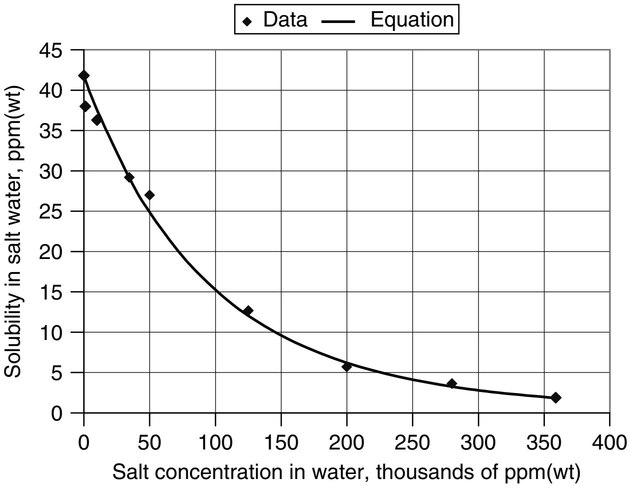

Figure 2. Solubility of methyl cyclopentane in salt water.

The calculated value and data compare favorably (34.62 vs. 34.50; deviation = 0.12/34.50 = 0.3%).

Example 2. A chemical spill of benzene occurs in seawater [salt = 34,472 ppm(wt)] at ambient conditions. Estimate the concentration in the seawater.

The values in Table 1 are used for the coefficients. Introducing the boiling point of benzene to calculate A and using the tabulated values for B and C provides

$$\begin{aligned} A &= 5.725 + 3.6778\text{E}{-03}\ T_B - 3.0346\text{E}{-05}\ T_B^2 \\ &= 5.725 + 3.6778\text{E}{-03}\ (353.31) \\ &\quad - 3.0346\text{E}{-05}\ (353.31)^2 \\ &= 3.2364 \end{aligned}$$

$$B = -3.2714\text{E}{-06}$$

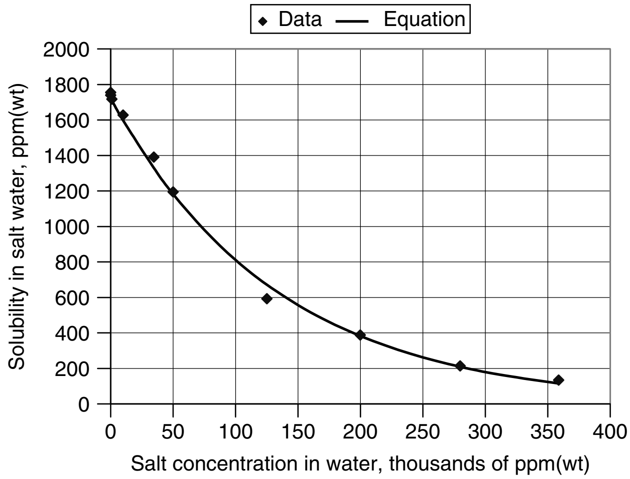

Figure 3. Solubility of benzene in salt water.

$C = 1.6357\text{E}{-16}$.

Substitution of these coefficients and the salt concentration in the correlation equation for solubility in saltwater yields

$$\begin{aligned}\log_{10} S &= A + B\,X + C\,X^2\\ &= 3.2364 - 3.2714\text{E}{-06}(34{,}472)\\ &\quad + 1.6357\text{E}{-16}(34{,}472)^2\\ &= 3.1236\\ S &= 1329.24 \text{ ppm (wt)}\end{aligned}$$

The calculated value and data compare favorably (1329.24 vs. 1391.00; deviation = 61.76/1391.00 = 4.4%).

Acknowledgment
The Texas Hazardous Waste Research Center (Lamar University, Beaumont, Texas) provided partial support for this work.

BIBLIOGRAPHY

1. Donald, M., Shiu, W-Y., and Ma, K-C. (1995). *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*. Vol. 4, Lewis, New York.
2. (2001). *Handbook of Environmental Data on Organic Chemicals*, 4th Edn. John Wiley & Sons, New York.
3. Hine, J. and Mukherjee, P.K. (1975). *J. Org. Chem.* **40**: 292.
4. Howard, P.H. and Meylan, W.M. (Eds.). (1997). *Handbook of Physical Properties of Organic Chemicals*. CRC Press, Boca Raton, FL.
5. Mackay, D., Shiu, W-Y., and Ma, K-C. (1992, 1992, 1993, 1995, 1997). *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*. Vols. 1–5, Lewis, New York.
6. Price, L.C. (1976). *Bull. Am. Assoc. Pet. Geologists* **60**(2): 213.
7. Sutton, C. and Calder, J.A. (1974). *Environ. Sci. Technol.* **8**(7): 654.
8. Suzuki, T. (1991). *J. Comput.-Aided Molecular Design* **5**: 149.
9. Verschueren, K. (1996, 2001). *Handbook of Environmental Data on Organic Chemicals*, 3rd and 4th Edn. Van Nostrand Reinhold, New York.
10. Yalkowsky, S.H. (1990–2004). *Aquasol Database*. University of Arizona, Tucson, AZ.
11. Yaws, C.L. (1999). *Chemical Properties Handbook*. McGraw-Hill, New York.

SOLUBILITY OF HYDROCARBONS AND SULFUR COMPOUNDS IN WATER

Carl L. Yaws
Prasad K. Narasimhan
Helen H. Lou
Lamar University
Beaumont, Texas

Ralph W. Pike
Louisiana State University
Baton Rouge, Louisiana

Correlations are presented for the solubility of hydrocarbons and organic sulfur compounds in water. A wide variety of compounds are covered—alkanes, olefins, diolefins, acetylenes, cyclopentanes, cyclohexanes, benzenes, mercaptans, thiophenes, and sulfides. The new correlation provides reliable solubility values down to very low concentrations (parts per million range). The correlation is based on the boiling point of the compound. Correlation and experimental data are in favorable agreement. The results are useful in process engineering for wastewater minimization.

IMPORTANCE OF WATER SOLUBILITY

The solubility of compounds in water is very important. This importance will increase in the future because of heath, safety, and environmental considerations. The following brief discussion illustrates the importance even at very low concentrations.

For health involving human exposure to substances in air, the threshold limit value (TLV) for *n*-pentane in air is given as 600 ppm (parts per million, vol) by the Occupational Safety and Health Act (1). A concentration of only 0.0000001 mol fraction of *n*-pentane in water provides about 7000 ppm of *n*-pentane in air at the air–water interface. This concentration of 7000 ppm in air exceeds the threshold limit value of 600 ppm.

For safety, the lower explosion limit (LEL) for *n*-pentane in air is given as 1.4% by Yaws (2). A concentration of only 0.000001 mol fraction of *n*-pentane in water provides about 7% of *n*-pentane in air at the air–water interface. This concentration of 7% in air exceeds the lower explosion limit of 1.7%.

For environmental, consider a spill of *n*-pentane in contact with water. The water will become saturated. At saturation, the solubility of *n*-pentane in water is about 0.0000385 weight fraction or 0.00000916 mol fraction, as given by Yalkowsky (3). This concentration of only

0.00000916 mol fraction at saturation will provide about 644,000 ppm or 64.4% of *n*-pentane in air at the air–water interface. This concentration of 644,000 ppm or 64.46% greatly exceeds both the threshold limit value of 600 ppm and the lower explosion limit of 1.4%.

This brief discussion (using *n*-pentane as an example) indicates that very low concentrations (ppm or less) of compounds in water can provide concentrations in air at the air–water interface that exceed the threshold limit value for human exposure and the lower explosion limit for flammability.

CORRELATION FOR WATER SOLUBILITY

In earlier work by Yaws and coworkers (2), the water solubility of various chemical types was correlated as a function of the boiling point of the compound. In this article it was determined that the boiling point method is also applicable to correlation of the water solubility of hydrocarbons and organic sulfur compounds:

$$\log_{10}(S) = A + BT_B + CT_B^2 + DT_B^3 \qquad (1)$$

where S = solubility in water at 25 °C, ppm by weight,

T_B = boiling point of compound, K

A, B, C, and D = regression coefficients

The range for boiling point is 298 K to about 560 K.

The coefficients (A, B, C, and D) are given in Table 1 for a wide variety of hydrocarbons and sulfur compounds—alkanes, olefins, diolefins, acetylenes, cyclopentanes, cyclohexanes, benzenes, mercaptans, thiophenes, and sulfides. The table also provides the boiling point range (T_{BMIN}, T_{BMAX}) for which the correlation is applicable. The correlation should not be used for compounds whose boiling points are outside the boiling point range.

The coefficients were determined from regression of available data. In preparing the correlation, a literature search was conducted to identify data source publications (2–10). The excellent compilations by Howard and Meylan (5); Mackay, Shiu, and Ma (6); Verschueren (8); Yalkowsky (3); and Yaws (2) were used extensively. The publications were screened and copies of appropriate data were made. These data were then keyed in to the computer to provide a database for which experimental data are available. The database also served as a basis to check the accuracy of the correlation.

Graphs of water solubility versus boiling point are shown in Figs. 1–8 for the compound families. The graphs indicate favorable agreement of correlation values and experimental data.

For olefins and diolefins, the graphs disclose that more data are available for alkanes. The graphs also disclose that the data for olefins and diolefins appear to be above and approximately parallel to the more extensive data for alkanes. This suggests that the solubility of olefins and diolefins is higher and approximately parallel to the water solubility of alkanes.

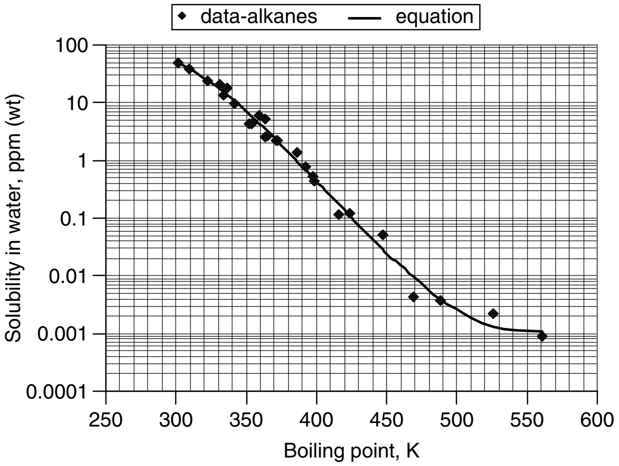

Figure 1. Solubility of alkanes in water.

Table 1. Solubility in Water[a]

Chemical family	A	B	C	D	T_{BMIN}	T_{BMAX}	Formula	Example Compound
$\log_{10} S = A + BT_B + CT_B^2 + DT_B^3$				(S - ppm(wt), $T_B - K$)				
Alkanes	−17.6520	1.77811E−01	−5.00907E−04	4.11124E−07	298	560	C_nH_{2n+2}	*n*-Hexane
Olefins	−17.0300	1.77811E−01	−5.00907E−04	4.11124E−07	298	560	C_nH_{2n}	1-Hexene
Diolefins	−16.5610	1.77811E−01	−5.00907E−04	4.11124E−07	298	560	C_nH_{2n-2}	1,5-Hexadiene
Acetylenes	−15.8350	1.77811E−01	−5.00907E−04	4.11124E−07	298	560	$C_{n+2}H_{2n+2}$	1-Hexyne
Cyclohexanes	−16.7000	1.77811E−01	−5.00907E−04	4.11124E−07	298	560	C_nH_{2n}	Methyl cyclohexane
Cyclopentanes	−16.9000	1.77811E−01	−5.00907E−04	4.11124E−07	298	560	C_nH_{2n}	Methyl cyclopentane
Benzenes (single substitution)	−24.0080	2.21196E−01	−5.55632E−04	4.18830E−07	298	560	$C_{n+6}H_{2n+6}$	Toluene
Benzenes (multiple substitution)	−23.6500	2.21196E−01	−5.55632E−04	4.18830E−07	298	560	$C_{n+6}H_{2n+6}$	*p*-Xylene
Mercaptans	−6.9000	1.00500E−01	−2.72880E−04	1.99870E−07	298	560	$C_nH_{2n+2}S$	Propyl mercaptan
Thiophenes	−6.8500	1.00500E−01	−2.72880E−04	1.99870E−07	298	560	$C_nH_{2n-4}S$	2-Methyl thiophene
Sulfides	−6.5390	1.00500E−01	−2.72880E−04	1.99870E−07	298	560	$C_nH_{2n+2}S$	Ethyl methyl sulfide

[a] S = solubility in water at 25 C, parts per million by weight, ppm(wt).

A, B, C, and D = regression coefficients.

T_B = boiling point of compound, K.

T_{BMAX}-maximum temperature, K.

T_{BMIN}-minimum temperature, K.

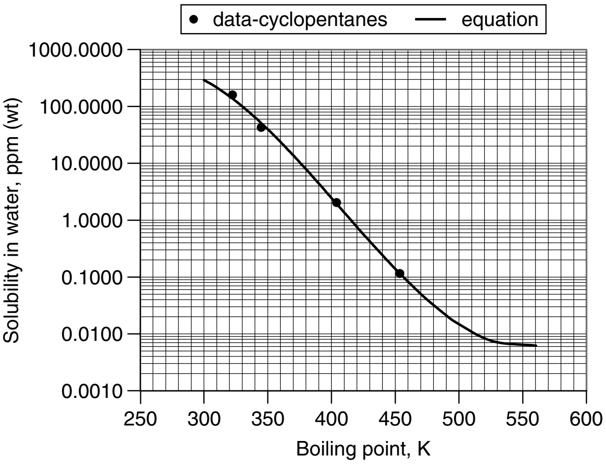

Figure 2. Solubility of cyclopentanes in water.

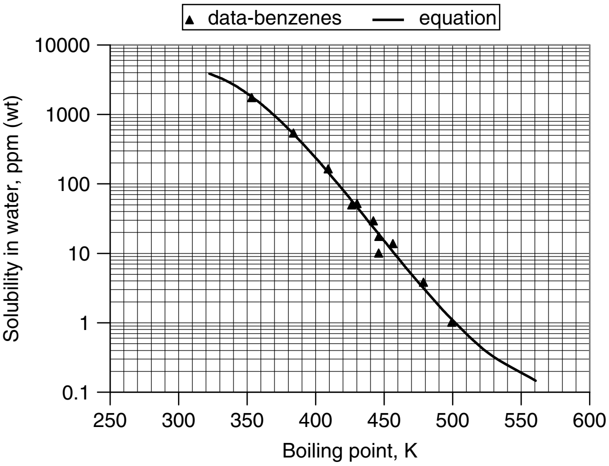

Figure 3. Solubility of benzenes in water.

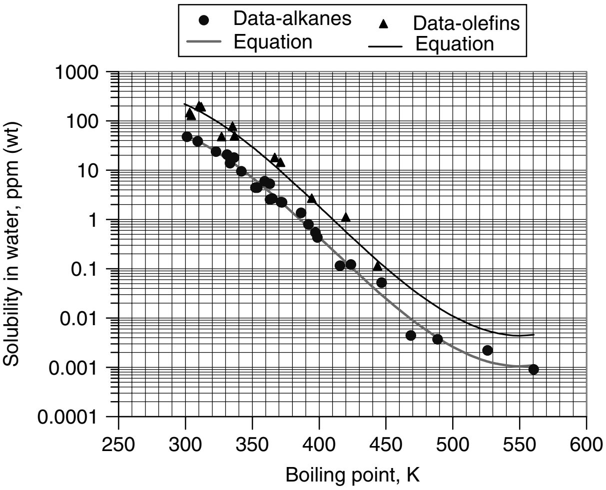

Figure 4. Solubility of olefins in water.

For sulfur compounds (mercaptans, thiopheolefins, and diolefins), the graphs disclose that more data are available for benzenes. The graphs also, disclose that the data for sulfur compounds appear to be slightly above and

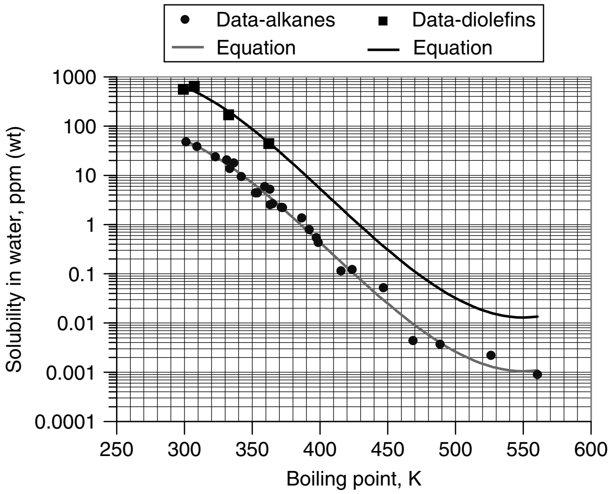

Figure 5. Solubility of diolefins in water.

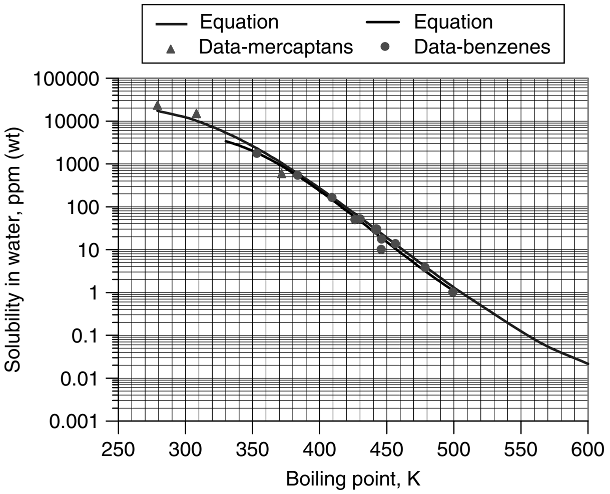

Figure 6. Solubility of mercaptans in water.

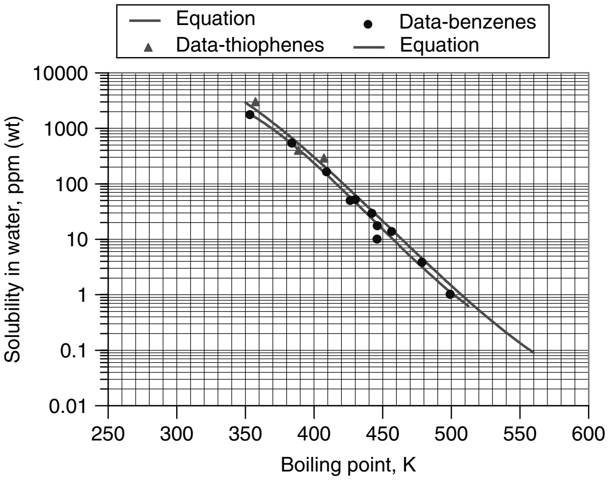

Figure 7. Solubility of thiophenes in water.

approximately parallel to the more extensive data for benzenes. This suggests that the solubility of sulfur

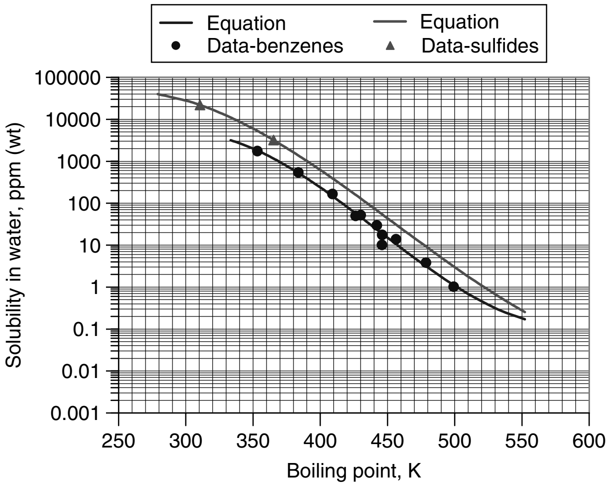

Figure 8. Solubility of sulfides in water.

compounds is higher and approximately parallel to the water solubility of benzenes.

Examples

The correlations for water solubility are useful in engineering applications involving process wastewater. Examples are shown below.

Example 1. A chemical spill of *n*-hexane occurs into a body of water at ambient conditions. Estimate the concentration in the water at saturation.

The correlation for alkanes maybe used to determine the solubility in water. Substitution of the coefficients and boiling point of *n*-hexane in the correlation equation yields:

$$\begin{aligned}\log_{10}(S) &= A + BT_{\mathrm{B}} + CT_{\mathrm{B}}^2 + DT_{\mathrm{B}}^3 \\ &= -17.6520 + 1.77811\mathrm{E}{-01}(341.88) \\ &\quad - 5.00907\mathrm{E}{-04}(341.88)^2 \\ &\quad + 4.11124\mathrm{E}{-07}(341.88)^3 \\ S &= 10.46 \text{ ppm (wt)}\end{aligned}$$

The calculated value and data compare favorably (10.46 vs 9.47; deviation = 0.99/9.47 = 10.4%).

Example 2. A chemical spill of toluene occurs into a body of water at ambient conditions. Estimate the concentration in the water at saturation.

The correlation for benzenes (single substitution) maybe used to determine the solubility in water. Substitution of the coefficients and boiling point of toluene in the correlation equation yields

$$\begin{aligned}\log_{10}(S) &= A + BT_{\mathrm{B}} + CT_{\mathrm{B}}^2 + DT_{\mathrm{B}}^3 \\ &= -24.0080 + 2.21196\mathrm{E}{-01} * (383.78) \\ &\quad - 5.55632\mathrm{E}{-04} * (383.78)^2 \\ &\quad + 4.18830\mathrm{E}{-07} * (383.78)^3 \\ S &= 524.68 \text{ ppm (wt)}\end{aligned}$$

The calculated value and data compare favorably (524.68 vs 542.4; deviation = 17.72/542.4 = 3.3%).

Acknowledgments

The Texas Hazardous Waste Research Center (Lamar University, Beaumont, Texas) provided partial support for this work. Portions of this material appeared in *Oil & Gas Journal* **89**; 79 (April 8, 1991), *Chem. Eng.*, **100**; 122 (Oct., 1993), *Oil & Gas Journal* **89**: 86 (Sept. 16, 1991), *Oil & Gas Journal* **95** (35): 80 (Aug. 29, 1994), and *Chem. Eng.* **110**: 60 (Aug., 2003). These portions are reprinted by special permission.

BIBLIOGRAPHY

1. *Guide to Occupational Exposure Values*. (1997). American Conference of Governmental Industrial Hygienists, Cincinnati, OH.
2. Yaws, C.L. (1999). *Chemical Properties Handbook*. McGraw-Hill, New York.
3. Yalkowsky, S.H. (1990–2004). *Aquasol Database*. University of Arizona, Tucson, AZ.
4. Hine, J. and Mukherjee, P.K. (1975). *J. Org. Chem.* **40**: 292.
5. Howard, P.H. and Meylan, W.M. (Eds.). (1997). *Handbook of Physical Properties of Organic Chemicals*. CRC Press, Boca Raton, FL.
6. Mackay, D., Shiu, W-Y., and Ma, K-C. (1992, 1992, 1993, 1995, 1997). *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*. Vols. 1–5, Lewis, New York.
7. Suzuki, T. (1991). *J. Comput.-Aided Molecular Design* **5**: 149.
8. Verschueren, K. (1996, 2001). *Handbook of Environmental Data on Organic Chemicals*, 3rd and 4th Edn. Van Nostrand Reinhold, New York.
9. Donald, M., Shiu, W-Y., and Ma, K-C. (1995). *Illustrated Handbook of Physical–Chemical Properties and Environmental Fate for Organic Chemicals*. Vol. 4, Lewis Publishers, New York.
10. (2001). *Handbook of Environmental Data on Organic Chemicals*, 4th Edn. John Wiley & Sons, New York.

SORPTION KINETICS

Th.D. Karapantsios
M.X. Loukidou
K.A. Matis
Aristotle University
Thessaloniki, Greece

INTRODUCTION

Sorption is, by definition, a general term describing the attachment of charged species (such as toxic metal ions) from a solution to a coexisting solid surface. Sorption kinetics may be controlled by several independent processes that can act in series or in parallel (see Fig. 1). These processes fall in one of the following general categories:

1. bulk diffusion,
2. external mass transfer (film diffusion),

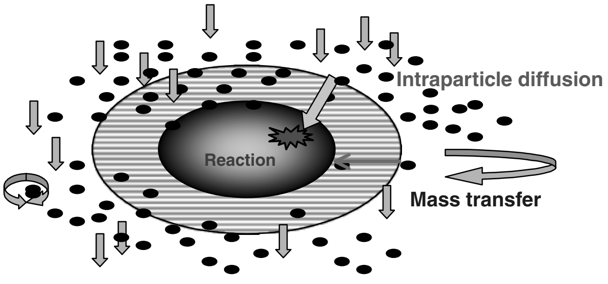

Figure 1. Illustration of the mechanism during sorption to describe suitably and model the respective process kinetics.

3. chemical reaction (chemisorption), and
4. intraparticle diffusion.

Kinetic analyses allow estimating sorption rates and also lead to suitable rate expressions characteristic of possible reaction mechanisms (1–3).

At sufficiently high agitation speed in a reaction vessel, the bulk diffusion step can be safely ignored because then sorption onto sorbent particles is decoupled from mass transfer in the bulk mixture. Apart from that, it is quite common that more than one process can contribute to system performance at the same time. In this case, the extensive interrelationships between the various equations make the overall kinetic model exceedingly complicated to evaluate. A rather simplifying approach to circumvent this problem is to assume that each one of the co-current processes dominate over the others (i.e., the rate-controlling step) at specific time regimes of the process and then study them independently (4); this approach was also adopted in this article.

Many studies engaged so far to examine sorption phenomena involved analysis of batch experiments where data were sampled at even time intervals over the entire course of the process. As a result, fast changing kinetic data characteristic of the phenomena just after the onset of sorption could not be accurately depicted on an adequately short timescale. Thus, a primary objective of this article is to investigate the kinetic mechanism of metal sorption on sorbent particles, putting more emphasis on samples collected at short times after the initiation of the process, where the major part of the adsorption occurs.

To identify the most appropriate mechanism for a process, several models must be checked for suitability and consistency in a broad range of system parameters. The model selection criteria proposed by Ho et al. (5) for sorption of pollutants in aqueous systems were used herein as a general guideline. According to this, several reaction-based and diffusion-based models were tested in simulating our data. The finally chosen kinetic models are those that fit the data closely and also represent reasonable sorption mechanisms.

EQUILIBRIUM AND KINETIC MODELING

An overall mass balance of the sorbate across the sorbent surface can be written as

$$X\,q_t + C_t^s = C_o \Leftrightarrow q_t = \frac{(C_o - C_t^s)}{X} \tag{1}$$

where q is the specific metal uptake (mg of metal per g of sorbent), C is the metal bulk concentration (mg/L) and X is the sorbent feed per unit volume of solution (g/L). Subscripts o and t denote conditions at the beginning and any other instant of the process, respectively. The superscript s denotes conditions at the sorbent interface. For adequately long adsorption times: $q_t \approx q_e$ and $C_t^s \approx C_e^s \approx C_e$, where the subscript e denotes equilibrium conditions. The known Langmuir and Freundlich adsorption models have usually been employed to describe equilibrium metal uptake:

$$q_e = \frac{q_{\max} b C_e}{1 + bC_e} \tag{2a}$$

$$q_e = K_F C_e^{1/n} \tag{2b}$$

$q_{\max}$ (mg/g), b (L/mg), and K_F [mg [g (mg/L)$^{-1/n}$]$^{-1}$] and n (dimensionless) are constants.

To compare measurements from various experiments for a kinetics investigation, it is necessary to introduce a dimensionless degree of conversion. Thus, by normalizing the remaining ion concentration, C_t, with respect to some reference value, an index of sorption is defined. Taking advantage of the values of C_t before the onset of sorption, C_o, and for completed sorption, $C_\infty = C_e$, the following degree of conversion is proposed:

$$\alpha = \frac{C_o - C_t}{C_o - C_e} \tag{3}$$

Kinetic studies customarily use the basic conversion rate equation:

$$\frac{d\alpha}{dt} = k(T) f(\alpha) \tag{4}$$

where $f(\alpha)$ is a conversion-dependent function and k(T) is the reaction rate constant. Sorption in heterogeneous liquid–solid systems has usually been considered to follow Arrhenius kinetics, which means that the temperature dependence of the rate constant k may be described by the well-known expression:

$$k(T) = A \exp(-E/RT) \tag{5}$$

However, to apply the model in such systems, the morphology of the reactive surface and the final sorption capacity must not vary with temperature, conditions commonly overlooked by many workers (6).

The example used in this article deals with an investigation of the removal and depletion of cadmium and chromates from aqueous solutions by biosorption on *Aeromonas caviae* (the sorbent). This microorganism is often present in groundwater and generally in aquatic environments. Biosorption, the uptake of heavy metals by dead biomass, has gained credibility during recent years as it offers a technically feasible and economical approach (7–10); the latter author illustrated that adsorption is different from ion exchange. Below, evidence is also provided that the system examined is a complex process. More than one sorption model were often reported to describe a case study correctly (11). Attention should be paid to any possible experimental error.

The Langmuir isotherm equation was often found to fit quite satisfactorily the experimental data (12,13). Experimental adsorption isotherms of metal cations or oxyanions, obtained with different sorbent concentrations, temperatures, etc. are usually presented as figures and then, tables display the results of fitting the Langmuir and/or the Freundlich model to the data obtained; the values of the respective correlation coefficient (r^2), if high, are often used to demonstrate a satisfactory description. The assumptions involved in these models were well described, among others, by Volesky (9). The Langmuir isotherms obtained during the equilibrium study lend further support to the notion of a monolayer (14).

On the other hand, the Freundlich isotherm is heavily used in industry. This model is often accurate for higher concentrations, but is inaccurate at low concentrations due to its power-law nature. The equilibrium properties do not change proportionally with the sorbent addition, possibly due to an aggregation of solids at higher loads and hence reducing the effective adsorption area (15).

From the chemical reaction category (chemisorption), the best fit for the data sets of this study is achieved by second-order chemical reactions (16). The solution of the standard second-order reaction, based on constant stoichiometry of one metal ion per binding site, is (e.g., 4):

$$C_{\mathrm{t}} = \frac{C_{\mathrm{o}}}{1 - \dfrac{C_{\mathrm{o}}}{C_{\mathrm{e}}}\exp(-k_1 C_{\mathrm{e}} t)} \tag{6}$$

where k_1 is the reaction rate constant [L*(mg^{-1} of metal)*min^{-1}]. This adsorption model has been very effective in describing the kinetics of adsorption of gases on solids. Nevertheless, when the data were plotted, it was shown that Eq. 6 clearly fails to capture the steep concentration gradient of the early removal stage (16). This was a direct indication that adsorption on solids from a liquid phase is a process different from adsorption from a gas phase, where traditionally the remaining bulk concentration dictates the kinetics.

If the rate of sorption depends not on bulk concentration but on uptake by the sorbent, this can be described by the so-called Ritchie second-order equation according to which one metal ion occupies two binding sites (17):

$$q_{\mathrm{t}} = q_{\mathrm{e}}\left[1 - \left(\frac{1}{1 + k_2 t}\right)\right] \tag{7}$$

where k_2 is the reaction rate constant (min^{-1}). When in the above treatment, q_{e} does not dictate sorbate uptake, then a pseudo-second-order rate expression is more appropriate:

$$\frac{t}{q_{\mathrm{t}}} = \frac{1}{k_{\mathrm{m}} q_{\mathrm{m}}^2} + \frac{1}{q_{\mathrm{m}}} t \tag{8}$$

where k_{m} is the reaction rate constant [g of sorbent*(mg^{-1} of metal)*min^{-1}] and q_{m} is a numerically determined parameter which, under ideal second-order rate control, corresponds to q_{e}. In the literature (5), various other kinetic equations have been attempted: zero, first (forward or reversible)-order, Langmuir–Hinshelwood, and the Elovich-type.

Equations 7 and 8 provided quite a suitable description of data for advancing time (see Fig. 2). It is noteworthy that both models adequately capture the rapid rate of adsorption during the first minutes of the experiments. This implies that the metal uptake by the sorbent is a satisfactory rate-controlling parameter under a second-order reaction mechanism. The numerically best fit values of the rate parameters of these equations were displayed in a table (16). The predicted equilibrium sorption capacities were quite close to the experimental values for both models. Nevertheless, the rate constant of the pseudo-second-order model, k_{m}, is monotonically correlated with changes in the biomass load and in the bulk concentration, features that have been encountered in the past in biosorption (18).

In contrast, the rate constant of the Ritchie second-order equation, k_2, fluctuated, beyond any physical reasoning. In addition, Equation 8 exhibited better fitting statistics. Despite the goodness of fit for sorption at 40 and 60 °C, the reaction rate constant of both models varied randomly with temperature. Preliminary calculations,

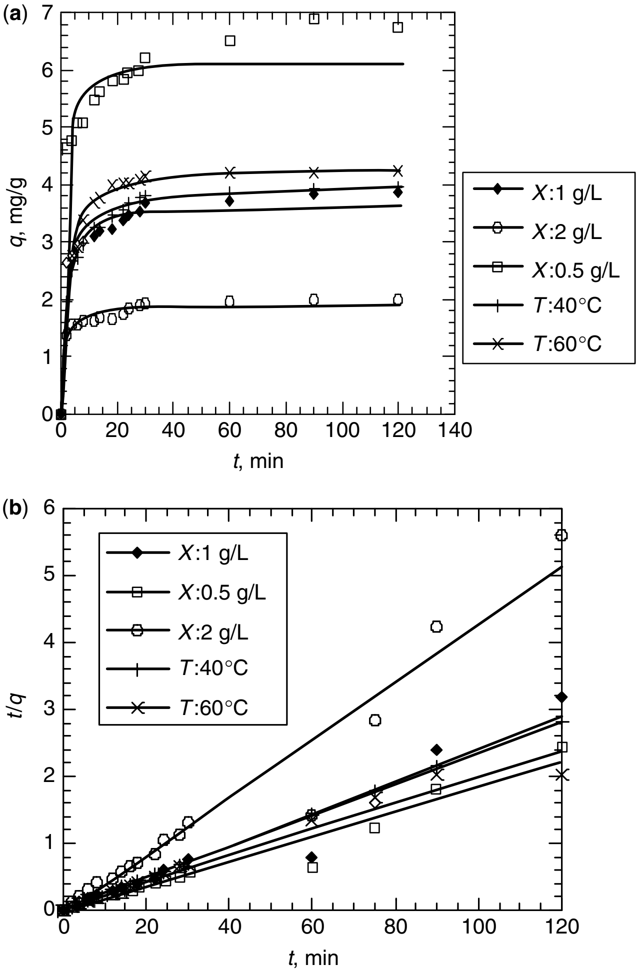

Figure 2. Comparison of experimental removal curves vs. theoretical predictions based on (**a**) the Ritchie second-order equation (at an initial cadmium concentration of 5 mg/L) and (**b**) the pseudo-second-order equation (at an initial cadmium concentration of 50 mg/L). Reprinted from Reference 16.

using the Arrhenius model between two temperatures every time, always gave activation energies below 10 kJ/mol, which is far less than expected for reaction-controlled sorption processes. The morphological changes in the biomass (sorbent) surface at different temperatures and the dependence of sorption capacity on temperature may be blamed for this irregularity.

DIFFUSION KINETICS AND MODELING

For intraparticle diffusion, Crank (19) proposed a model that takes into account the continuously decreasing bulk concentration due to sorbate removal. This gives rise to a time-dependent boundary condition for the concentration at the surface of the sorbent particle. The solution of the diffusion equation for such a boundary condition and a concentration independent diffusivity is

$$\alpha = 1 - 6\sum_{n=1}^{\infty} \frac{\exp(-\xi p_n^2 t)}{9\Lambda/(1-\Lambda) + (1-\Lambda)p_n^2} \tag{9}$$

where p_n is given by the nonzero roots of

$$\tan(p_n) = \frac{3p_n}{3 + p_n^2/(1-\Lambda)} \tag{10}$$

and $\Lambda \equiv (C_o - C_\infty)/C_o$ is the fraction of metal ultimately adsorbed by the sorbent.

Equation 9 can be solved numerically to determine ξ, the effective diffusional time constant, which for particle (also called micropore) diffusion control equals D_c/R_c^2; D_c and R_c are the intraparticle diffusion coefficient (m^2/s) and mean particle radius (m), respectively. The same expression is the solution of the diffusion equation for (macro)pore diffusion control but only where the equilibrium isotherm is linear ($q_e \approx K^* C_e$) for the concentration range under investigation. Then ξ equals $(D_p/R_p^2)/[1 + (1-\varepsilon_p)K/\varepsilon_p]$, where $\varepsilon_p D_p/[1 + (1-\varepsilon_p)K]$ is the effective macropore diffusivity, K is the equilibrium constant, ε_p is the void volume fraction, and R_p is the radius of the particle (14).

The experimental procedure in this article is based on measuring the remaining metal concentration in the bulk, so Λ is always above 0.5. For Λ greater than about 0.1, the effect of diminishing bulk concentration becomes significant, and under these conditions, the assumption of constant metal concentration at the surface of the sorbent leads to an erroneously high diffusivity. Provided that the concentration step employed in a sorption experiment corresponds to a linear section of the equilibrium isotherm, the constant K can be replaced by the local slope of the isotherm, dq_e/dC_e. This is approximately valid for the concentrations employed in this work, as can be seen from the experimental isotherms. Our interest is chiefly focused on the short time region ($\sqrt{D_{c,p}t/R_{c,p}^2} < 0.2$) where Eq. 9 converges slowly, so at least 200 terms are used in the summation to achieve satisfactory accuracy.

External mass transfer has been customarily analyzed in literature by adopting a pseudo-first-order reaction model (5,20). This approach assumes that the sorbate concentration at the sorbent surface is zero at all times. However, this is not true, particularly where a significant quantity of sorbate is adsorbed rapidly at the beginning of the process. A more realistic model should consider instead that a rapid equilibrium is established between the sorbate at the interface and that present on the sorbent surface (21), and this concept is adopted in this article.

If one combines the mass balance across the sorbent surface, Equation 1, the Langmuir adsorption isotherm, Equation 2a, and the rate of change equation in the bulk concentration, one ends up with (21)

$$\frac{dC_t}{dt} = -K_m S(C_t - C_t^s) \tag{11}$$

$$\frac{dC_t^s}{dt} = \left(\frac{K_m S}{X q_{\max} b}\right)[(C_t - C_t^s)(1 + bC_t^s)^2] \tag{12}$$

where K_m is the external mass transfer coefficient (m/s) and S is the specific surface area of the sorbent particles per unit volume of the reactor (m^2/m^3). Using the following dimensionless variables: $C^* = C_t/C_o$, $C_s^* = C_t^s/C_o^s$, and $t^* = t/\tau$, where C_o^s is determined from Eqs. 1 and 2a and τ is the total adsorption time, Eqs. 11 and 12 convert to

$$\frac{dC^*}{dt^*} = -K_m S\tau(C^* - C_s^*) \tag{13}$$

$$\frac{dC_s^*}{dt^*} = \left(\frac{K_m S\tau}{X q_{\max} b}\right)\left[(C^* - C_s^*)(1 + bC_o C_s^*)^2\right] \tag{14}$$

which is a system of two first-order ordinary differential equations that must be solved simultaneously. The initial conditions are $C^* = 1$ and $C_s^* = 0$ at $t^* = 0$. Note that normalizing C_t^s with respect to C_o^s —and not C_o, as Puranik et al. (21) did —and also t with respect to τ, improved the stability and convergence characteristics of the solution markedly due to comparable spreading of all variables over the computational domain.

Equations 9–10 and 13–14 are solved numerically to determine ξ and $K_m S$, respectively. The nonlinear numerical regression to fit experimental data to those equations is performed by the Levenberg–Marquardt method, which gradually shifts the search for the minimum of the sum of the errors squared (SSE), from the steepest descent to quadratic minimization—Gauss–Newton (22):

$$\text{SSE} = \sum_i \left[\frac{(q_{exp,i} - q_{cal,i})^2}{q_{exp,i}^2}\right] \tag{15}$$

Figure 3a presents the results of fitting Eq. 9 to biosorption data obtained with different initial concentrations, solids loads, and temperatures. It is apparent that, despite some scatter in measurements, the finite volume diffusion model can fairly well describe the entire range of data, also including the steep concentration gradient at short times. The values of Λ, ξ, and the computed values of D_c were displayed in a table (14). Such behavior has been customary as a consequence of the decreasing slope of a nonlinear equilibrium curve, example, a Langmuir isotherm, which

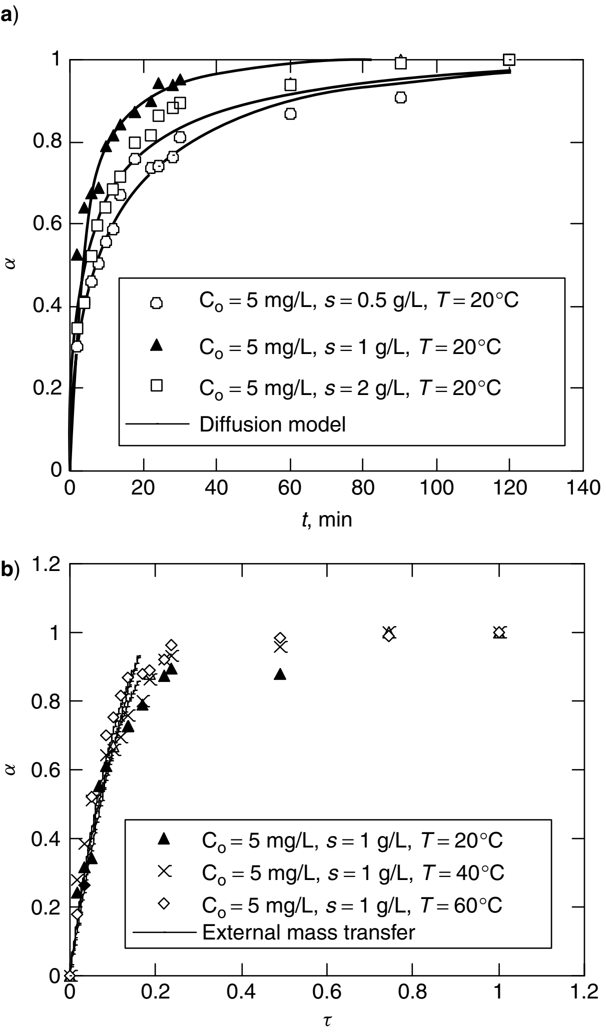

Figure 3. Experimental degree of conversion, α, vs. predictions based on the solution of (**a**) the diffusion equation, for various adsorbent loads and (**b**) the mass transfer equation, for various temperatures (both at an initial chromium concentration of 5 mg/L). Reprinted with permission from Reference 14; copyright (2004) American Chemical Society.

causes the diffusivity to increase rapidly as concentration increases (11).

A way to check further on the possibility of a pore diffusion-controlled mechanism is to perform desorption kinetic tests with adsorbent previously used for sorption. For a nonporous sorbent particle, transport of solute inside the particle may be neglected, and it can be assumed that biosorption occurs mainly at the particle surface. This idea can be effectively extended to cases of relatively large macropores, where the metal ions may have ready access to react with internal surface sites.

It was argued that, at the beginning of the process, chromium is sorbed according to quite a fast and highly favorable chemical mechanism, such as ion exchange, but soon external film diffusion comes into play (14). On this account, if one ignores the very first minute of sorption, the remaining curves were fitted pretty well by the model Eqs. 13 and 14. Figure 3b displays these results. Near the end of sorption (for α higher than 0.9), a much slower process, intraparticle diffusion, gradually becomes the rate-controlling step. Yet, this is a regime of no practical significance.

The computed mass transfer rate constant, $K_{\mathrm{m}}S$, was presented in a table (14). As regards the effect of changing the solids load, it appears that $K_{\mathrm{m}}S$ is virtually not affected. To assess the statistical significance of the determination, a derivative time series analysis has been performed where $d\alpha/dt$ is plotted versus t and then, a Hanning low-pass filter was applied to flatten out the signal undulations until $d\alpha/dt$ versus α became a reasonably smooth curve (23). Next, the integrated smoothed signal $\alpha(t)$ was processed as before and resulted in $K_{\mathrm{m}}S$ values closely comparable to the values obtained from the unfiltered data, with no preferential trend in biomass load. This is an additional positive sign that external mass transfer may be the predominant mechanism of the sorption process after the initial fast chromium removal.

As a conclusion in the example thoroughly studied, we may say that the analysis conducted has not been capable of providing strong evidence in favor of any of the examined mechanisms because several diverse kinetic models were successful in fitting the experimental data. Perhaps, the questioning of surface kinetics by Levenspiel (24) is known, concluding that it is good enough to use the simplest available correlating rate expression.

NOMENCLATURE

q:	specific metal uptake (mg/g)
C:	metal bulk concentration (mg/L)
X:	sorbent loading per unit volume of solution (g/L)
$q_{\max}$:	Langmuir constant (mg/g)
b:	Langmuir constant (L/mg)
K_{F}:	Freundlich constant {mg [g (mg/L)$^{-1/n}$]$^{-1}$}
n:	Freundlich constant (dimensionless)
E:	apparent activation energy (J/mol)
A:	preexponential factor (min^{-1})
n:	apparent reaction order
R:	gas constant (8.3136 J/mol/K)
T:	temperature (K)
t:	time (min)
k:	rate constant
k_1:	second-order chemical reaction rate constant [L*(mg^{-1} of metal)*min^{-1}]
k_2:	second-order Ritchie reaction rate constant (min^{-1})
k_{m}:	pseudo-second-order reaction rate constant [g of sorbent*(mg^{-1} of metal)*min^{-1}]
D_{c}:	intraparticle diffusion coefficient (m^2/s)
R_{c}:	mean particle radius (m)
R_{p}:	radius of a particle
K_{m}:	external mass transfer coefficient (m/s)
S:	specific surface area of sorbent particles per unit volume of the reactor (m^2/m^3)

Greek Letters

α:	degree of conversion
Λ:	fraction of metal ultimately adsorbed by the sorbent

ξ: numerically determined parameter
ε_p: void volume fraction

BIBLIOGRAPHY

1. Tien, C. (1994). *Adsorption Calculations and Modelling*. Butterworth-Heinemann, Boston.
2. Yiacoumi, S. and Tien, C. (1995). *Kinetics of Metal Ion Adsorption from Aqueous Solutions*. Kluwer Academic, Boston.
3. Xu, Y. and Schwartz, F.W. (1994). Sorption of Zn^{2+} and Cd^{2+} on hydroxyapatite surfaces. *Environ. Sci. Technol.* **28**: 1472–1480.
4. Hill, C.G. (1977). *An Introduction to Chemical Engineering Kinetics and Reactor Design*. John Wiley & Sons, New York, pp. 24–166.
5. Ho, Y.S., Ng, J.C.Y., and McKay, G. (2000). Kinetics of pollutant sorption by biosorbents: review. *Sep. Purific. Methods* **29**: 189–232.
6. Smith, J.M. (1981). *Chemical Engineering Kinetics*, 3rd Edn. McGraw-Hill, Singapore, pp. 37–92, 636–645.
7. Aksu, Z. (2001). Equilibrium and kinetic modeling of cadmium(II) biosorption by *C. vulgaris* in a batch system: Effect of temperature. *Sep. Purific. Tech.* **21**: 285–294.
8. Butter, T.J., Evison, L.M., Hancock, I.C., Holland, F.C., Matis, K.A., Philipson, A., Sheikh, A.I., and Zouboulis, A.I. (1998). The removal and recovery of cadmium from dilute aqueous solutions by biosorption and electrolysis at laboratory scale. *Water Res.* **32**: 400–406.
9. Volesky, B. (1999). Biosorption of Heavy Metals. CRC Press, Boca Raton, FL, Chap. 6. Available: http://ww2.mcgill.ca/biosorption/publication/index.htm.
10. Zouboulis, A.I., Matis, K.A., and Hancock, I.C. (1997). Biosorption of metals from dilute aqueous solutions. *Sep. Purif. Methods* **26**: 255–295.
11. Smith, E.H. (1996). Uptake of heavy metals in batch systems by a recycled iron-bearing material. *Water Res.* **30**: 2424–2434.
12. Mandjiny, S., Zouboulis, A.I., and Matis, K.A. (1995). Removal of cadmium from dilute solutions by hydroxyapatite. Part I. Sorption studies. *Sep. Sci. Tech.* **30**: 2963–2978.
13. Kefala, M.I., Zouboulis, A.I., and Matis, K.A. (1999). Biosorption of cadmium ions by *Actinomycetes* and separation by flotation. *Environ. Pollut.* **104**: 283–293.
14. Loukidou, M.X., Karapantsios, Th.D., Zouboulis, A.I., and Matis, K.A. (2004). Diffusion kinetic study of chromium(VI) biosorption by *Aeromonas caviae*. *Ind. Eng. Chem. Res.* **43**: 1748–1755.
15. Esposito, A., Pagnanelli, F., Lodi, A., Solisio, C., and Vegliò, F. (2001). Biosorption of heavy metals by *Sphaerotilus natans*: An equilibrium study at different pH and biomass concentration. *Hydrometallurgy* **60**: 129–141.
16. Loukidou, M.X., Karapantsios, Th.D., Zouboulis, A.I., and Matis, K.A. (2003). *Cadmium(II) biosorption by Aeromonas caviae: Kinetic modelling. 15th Intl. Biohydrometallurgy Symp.*, Athens.
17. Ritchie, A.G. (1977). Alternative to the Elovich equation for the kinetics of adsorption of gases on solids. *J. Chem. Soc.-Farad. Trans.* **73**: 1650–1653.
18. Cheung, C.W., Porter, J.F., and McKay, G. (2001). Sorption kinetic analysis for the removal of cadmium ions from effluents using bone char. *Water Res.* **35**: 605–612.
19. Crank, J. (1975). *Mathematics of Diffusion*. Oxford University Press, London.
20. Dzul Erosa, M.S., Saucedo, T.I., Navarro Mendoza, R., Avila Rodriquez, M., and Guibal, E. (2001). Cadmium sorption on chitosan sorbents: Kinetic and equilibrium studies. *Hydrometallurgy* **61**: 157–167.
21. Puranik, P.R., Modak, J.M., and Paknikar, K.M. (1999). A comparative study of the mass transfer kinetics of metal biosorption by microbial biomass. *Hydrometallurgy* **52**: 189–197.
22. Bates, D.M. and Watts, D.G. (1988). *Nonlinear Regression Analysis and its Applications*. John Wiley & Sons, New York.
23. Bendar, J.S. and Piersol, A.G. (1986). *Random Data: Analysis and Measurement Procedures*. John Wiley & Sons, New York.
24. Levenspiel, O. (1999). *Chemical Reaction Engineering*, 3rd Edn. John Wiley & Sons, New York, p. 380.

SOUND IN WATER

D.L. Marrin
Hanalei, Hawaii

As sound travels through an elastic medium such as water, a wave is created that imparts energy to the individual molecules—causing them to compress and then to relax along the path of the wave. Because the medium is elastic, the distance separating individual water molecules is permitted to vary just slightly, so that the propagation of sound may be defined as a periodic variation in pressure that is transmitted via adjacent molecules. Sound emanating from a point source in water usually (but not always) results in a three-dimensional waveform that spreads spherically through the medium (1). Similarly, a planar source of sound in water can be considered multiple point sources that create a wave consisting of all the individual waves.

Sound waves are *longitudinal*, meaning that molecules (e.g., water or air) comprising the elastic medium move in a direction parallel to the propagating wave. Because water has a significantly higher density than air, water transmits sound more rapidly and efficiently than air. The water molecules oscillate, or move back and forth over a very short longitudinal distance; hence, sound waves are classified as *mechanical*. By contrast, light waves and microwaves are classified as *electromagnetic* and do not involve the movement of water molecules for propagation. The speed of sound in water is about 1500 meters per second or 3300 miles per hour (i.e., approximately five times greater than that in air), which varies slightly with temperature, pressure, and salinity.

FREQUENCY AND AMPLITUDE

The frequency of a sound wave is the number of waves that pass by a given point during a specified period of time. A single wave consists of a cycle that includes both compression (squeezing together) and rarefaction (rebounding apart); the distance over which a cycle is continually repeated is known as the wavelength. The period of time over which waves are normally counted is one second; therefore, sound frequencies are expressed as cycles per second or *hertz*. Whereas high-frequency sounds possess short wavelengths, low-frequency sounds possess long wavelengths. Frequency ranges and estimated intensity levels for common underwater sounds are listed in Table 1. Frequency-dependent differences in underwater sound attenuation result in low-pitched sounds that travel farther than high-pitched sounds generated at similar intensities.

The amplitude of a sound wave is the maximum longitudinal displacement (i.e., during one cycle of compression and rarefaction) of water molecules relative to their resting, or equilibrium, position. The intensity of a sound wave is proportional to its amplitude and is often quantified in units of *decibels*. The decibel scale is based on a ratio between the intensities (expressed as a power) of a measured sound and a reference sound that corresponds to a standard pressure. Many of the man-made sounds listed in Table 1 are actually more intense than naturally produced sounds, contributing, at least in part, to the controversy regarding the possible effects of anthropogenic noise on marine organisms. Sound intensities are commonly presented for a distance of only 1 meter from the source and, as such, are rarely indicative of ambient underwater noise (e.g., the cumulative intensity of sounds measured at greater distances from their respective sources).

FATE OF OCEANIC SOUND

The propagation of sounds in the ocean is a rather complex matter owing to differences in water density as a function of depth, latitude, seafloor topography, and many other factors. Changes in seawater temperature, salinity, and pressure contribute to changes in density that, in turn, affect sound waves traveling through the medium. Generally, sound is attenuated by *spreading* (proportional to distance), *reflection* (due to solid structures or boundaries between water of different densities), *scattering* (due to rough surfaces), *absorption* (conversion of acoustic to thermal energy), and *refraction* (deflection of sound waves from a straight path) (5). Ambient noise in a shipping channel can exceed that outside a channel by as much as 45 decibels (re 1 μPa), but noise in a channel is attenuated about 100 decibels from that measured directly beneath ships (2).

Due to the vertical stratification of oceans, sound behaves differently in turbulent shallow waters from the way it does in either more isothermal middepth waters or waters immediately overlying the bottom (see Table 2). Sound transmission at the surface is highly dependent on local conditions (e.g., wind, precipitation, swells, bubbles) and is usually restricted to low-frequency sounds. However, most of the ocean lies between shallow and bottom waters, creating a relatively thick layer that transmits a spectrum of sound waves that is far less attenuated than that in either shallower or deeper waters. Both humans and whales have taken advantage of this *sofar channel* to broadcast signals and songs, respectively, over vast oceanic distances (6). Sounds generated within the sofar channel itself or at shallower depths alongside seamounts are entrained and transmitted within this deep isothermal layer as a result of many contributing physical processes.

UNDERWATER DETECTION

Detecting and identifying an underwater sound from a particular source is often difficult due to the aforementioned effects of the environment and to the myriad of background noises from which any single sound must be distinguished. Sea sounds are detected by hydrophones (i.e., the underwater equivalent of a microphone) that are often positioned in specific arrays, thus permitting the relative position of a sound source to be determined. This type of *passive* detection may be used to locate oceanic precipitation, breaking surface waves, underwater explosions, undersea volcanoes, shipping traffic, and marine organisms that emit audible sounds (e.g., crustaceans, fishes, mammals) (7).

Active detection techniques, such as sonar, are routinely used to locate underwater objects because the objects themselves need not emit sound and because the results

Table 1. Approximate Ranges for the Frequencies and Intensity Levels of Common Underwater Sounds. Reported Ranges were Complied from Data Presented in Various Sources[a]

Sound Source	Frequency, Hertz	Intensity Level, Decibels re 1 μPa at 1 m[b]
Ship engines and propellers	10–5000	160–190
Navigation and profiling sonars	100–3000	180–230
Explosive devices/air guns	1000–17,000	190–260
Military surveillance sonars	1000–10,000	190–235
Icebreaking/drilling operations	20–1000	100–150
Whale songs and moans	10–8000	120–190
Dolphin clicks and whistles	500–25,000	100–180
Cetacean echolocation	10,000–150,000	130–230
Snapping shrimp colony	2000–15,000	180–190
Lightning strikes/undersea volcanoes/earthquakes	0.1–20,000	Up to 260

[a] Reference 2–4.

[b] Intensity levels are presented in units of decibels relative to a reference pressure of 1 micropascal at a distance of 1 meter from the sound source. Sound level attenuation is a function of distance from the source, frequency range, and various environmental factors.

Table 2. Approximate Depths and Representative Sound Speeds for Various Water Layers in the Ocean. Reported Ranges were Complied from Data Presented in Various Sources[a]

Vertical Zone	Depth, Meters	Speed, m/s	Comments
Overlying air	NA[b]	330	Significantly lower acoustic impedance than water; minimal transfer of sound energy
Surface water layer	0 to 50	1515	Influenced by weather (e.g., wind, rain, temp.) and turbulence that affect sound waves
Seasonal thermal layer	50 to 800	1485–1505	Sound channel (mostly for low frequencies); limited vertical thickness; refracts sound waves originating shallower and deeper
Deep isothermal layer	800 to 4000+[c]	1485–1525	Optimal sound channel (for all frequencies); extensive vertical thickness and horizontal extent; minimal sound attenuation
Bottom water layer	4000+ to seafloor[c]	1525	Influenced by seafloor topographic features that reflect and scatter sound waves
Underlying rock	NA[b]	2000–6000	Higher acoustic impedance than water; moderate transfer of sound energy

[a] References 2, 6, and 7.

[b] Not applicable.

[c] Indicates that the transition between deep isothermal layers and bottom water layers may exceed 4000 meters, depending on the exact location within an ocean basin.

are generally more interpretable. The most basic type of sonar is echolocation, whereby a sound pulse (acoustic energy) is introduced into the water. Depending on the sonic wavelength and the size of the objects, a portion of that acoustic energy is reflected back to the source (e.g., a ship) where it is detected, transduced, and analyzed. The distance, movement, size, and even shape of some objects may be determined using sophisticated sonars (e.g., Doppler, multibeam) and data analysis techniques (8). Several marine organisms—most notably cetaceans—are believed to use echolocation as a means of finding or assessing prey, predators, conspecifics, seabed features, and environmental conditions (2).

ULTRASOUND AND INFRASOUND

Most of the underwater sounds discussed thus far lie within the *sonic* range of humans (i.e., frequencies of 20 to 20,000 hertz), but water also transmits mechanical waves of *ultrasonic* (>20,000 hertz) and *infrasonic* (<20 hertz) frequencies. Dolphins emit whistles at frequencies as high as 30,000 hertz and complex echolocation clicks at frequencies exceeding 300,000 hertz (8). Obviously, the upper limit of hearing of dolphins far exceeds that of humans. In addition to marine organisms, ambient oceanic noise associated with bubbles (i.e., trapped gases), surface winds, waves, and sea spray can reach ultrasonic frequencies. Man-made ultrasounds include everything from complex signals for communication to guided waves for inspecting underwater pipelines (9).

The moans of baleen whales are the most common source of biologically produced infrasound, which has been documented down to about 10 hertz (8). By contrast, a number of geologic and meteorologic events (e.g., volcanoes, earthquakes, sea ice cracking, hurricanes) produce infrasonic waves that propagate through the ocean at frequencies of approximately 0.1 to 10 hertz. Even the breaking of large surf, either onto the shore or on the water itself, creates infrasound at frequencies of 1 to 5 hertz (10). Anthropogenic infrasound in the ocean is created primarily via shipping traffic, low-frequency active sonar (LFAS), and acoustic thermography. A thermography technique known as ATOC generates tones down to about 1 hertz, whereas LFAS sweeps multiple tones of 100 to 500 hertz across each other and produces resulting tones as low as 0.1 hertz (2). Both ATOC and LFAS generate peak intensity sounds that are greater than those of ships.

BIBLIOGRAPHY

1. Leighton, T.G. (1994). *The Acoustic Bubble*. Academic, San Diego, Chap. 1.
2. Stocker, M. (2003). Ocean bio-acoustics and noise pollution. *J. Acoust. Ecol.* **4**: 16–29.
3. Science of Sound in the Sea. (2004). *Office of Marine Programs*. University of Rhode Island, HTML format.
4. Greene, C.R. and Moore, S.E. (1995). Man-made noise. In: *Marine Mammals and Noise*. W.J. Richardson et al. (Eds.). Academic, San Diego, Chap. 6.
5. Haines, G. (1974). *Sound Underwater*. Crane Russak, New York, pp. 22–25.
6. Payne, R. (1995). *Among Whales*. Scribner, New York, Appendix.
7. Medwin, H. and Clay, C.S. (1998). *Fundamentals of Acoustical Oceanography*. Academic, San Diego, Chap. 1 and 10.
8. Thompson, D.H. and Richardson, W.J. (1995). Marine mammal sounds. In: *Marine Mammals and Noise*. W.J. Richardson et al. (Eds.). Academic, San Diego, Chap. 7.
9. Na, W.B. and Kundu, T. (2002). Underwater pipeline inspection using guided waves. *J. Pressure Vessel Technol.* **124**: 196–200.
10. Garcés, M., Hetzer, C., Marrifield, M., Willis, M., and Aucan, J. (2003). Observations of surf infrasound in Hawai'i. *Geophys. Res. Lett.* **30**: 2264–2266.

WATER ON THE SPACE STATION

Patrick L. Barry
Tony Phillips
NASA

Rationing and recycling will be an essential part of life on the International Space Station. In this article, Science@NASA explores where the crew will get their water and how they will (re)use it.

November 2, 2000—Future astronauts poised to blast off for an extended stay on the International Space Station (ISS) might first consider dashing to the restroom for a quick splash at the lavatory, or better yet, a luxurious hot shower. Once on board the ISS, spacefarers are in for a steady diet of sponge baths using water distilled from—among other places—their crewmates breath!

If you're squeamish, read no farther, because the crew will eventually include lab rodents—and they'll be breathing, too. All of the denizens of the space station lose water when they exhale or sweat. Such vapors add to the ambient cabin humidity, which is eventually condensed and returned to the general water supply.

Sometimes it's better *not* to think about where your next glass of water is coming from!

Rationing and recycling will be an essential part of daily life on the ISS. In orbit, where Earth's natural life support system is missing, the Space Station itself has to provide abundant power, clean water, and breathable air at the right temperature and humidity—24 hours a day, 7 days a week, indefinitely. Nothing can go to waste.

In this article, the first of a series about the practical challenges of living in space, *Science@NASA* will examine how the Space Station's **E**nvironmental **C**ontrol and **L**ife **S**upport **S**ystem (ECLSS), under continuing development at the Marshall Space Flight Center, will help astronauts use *and re-use* their precious supplies of water. Future installments will explore air management, thermal control and fire suppression—in short, all of the things that will make the Space Station comfortable and safe.

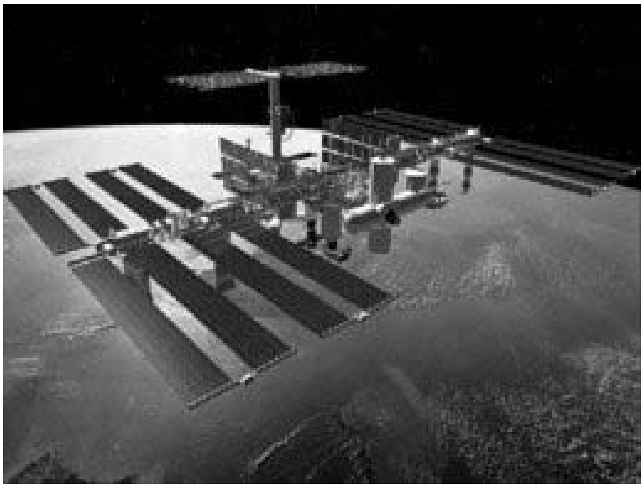

MAKING A SPLASH IN SPACE

Before recycling can begin, there has to be some water to start with.

"We have plenty of water on the Space Station now," says Jim Reuter, leader of the ECLSS group at the Marshall Space Flight Center. "The Russian module Zarya is packed with contingency water containers (CWCs) that were carried over from the Space Shuttle during assembly missions earlier this year. They look like duffle bags and each one holds about 90 lbs."

"But it's expensive to ferry water from Earth," he added. "We have to recycle. There's already a Russian-built water processor in orbit that collects humidity from the air. Here at Marshall we're building a regenerative system that will be able to recycle almost every drop of water on the station and support a crew of seven with minimal resupplies."

The ECLSS Water Recycling System (WRS), developed at the MSFC, will reclaim waste waters from the Space Shuttle's fuel cells, from urine, from oral hygiene and hand washing, and by condensing humidity from the air. Without such careful recycling 40,000 pounds per year of water from Earth would be required to resupply a minimum of four crewmembers for the life of the station.

Not even research animals are excused from the program.

"Lab animals on the ISS breath and urinate, too, and we plan to reclaim their waste products along with the

Shuttle pilot Terry Wilcutt with 7 contingency water containers destined for the space station Mir

---

crew's. A full complement of 72 rats would equal about one human in terms of water reclamation," says Layne Carter, a water-processing specialist at the MSFC.

It might sound disgusting, but water leaving the space station's purification machines will be cleaner than what most of us drink on Earth.

"The water that we generate is much cleaner than anything you'll ever get out of any tap in the United States," says Carter. "We certainly do a much more aggressive treatment process (than municipal waste water treatment plants). We have practically ultra-pure water by the time our water's finished."

MIMICKING MOTHER EARTH

On Earth, water that passes through animals' bodies is made fresh again by natural processes. Microbes in the soil break down urea and convert it to a form that plants can absorb and use to build new plant tissue. The granular soil also acts as a physical filter. Bits of clay cling to nutrients in urine electrostatically, purifying the water and providing nutrients for plants.

Water excreted by animals also evaporates into the atmosphere and rains back down to the Earth as fresh water—a natural form of distillation.

Water purification machines on the ISS partly mimic these processes, but they do not rely on microbes or any other living things.

"While you try to mimic what's happening on Earth—which is so complicated if you really think about it—we have to use systems that we can control 100 percent," said Monsi Roman, chief microbiologist for the ECLSS project at MSFC. ECLSS depends on machines—not microbes—because, "if a machine breaks, you can fix it."

The water purification machines on the ISS will cleanse wastewater in a three-step process.

The first step is a filter that removes particles and debris. Then the water passes through the "multi-filtration beds," which contain substances that remove organic and inorganic impurities. And finally, the "catalytic oxidation reactor" removes volatile organic compounds and kills bacteria and viruses.

EVERY DROP COUNTS

Once the water is purified, astronauts will do everything possible to use it efficiently. "On the ground, people flick on the faucet and they probably waste a couple of liters of water just because it's free and the water pressure is high," notes Carter.

"On the ISS, the water pressure will be about half what you might experience in a typical household," Carter said. "We don't use faucets on the ISS, we use a wash cloth. It's much more efficient. If you're an astronaut, you'll wet the wash cloth with a spray nozzle and then use the cloth to wash your hands."

On the space station, people will wash their hands with less than one-tenth the water that people typically use on Earth. Instead of consuming 50 liters to take a shower, which is typical on Earth, denizens of the ISS will use less than 4 liters to bathe.

Even with intense conservation and recycling efforts, the Space Station will gradually lose water because of inefficiencies in the life support system.

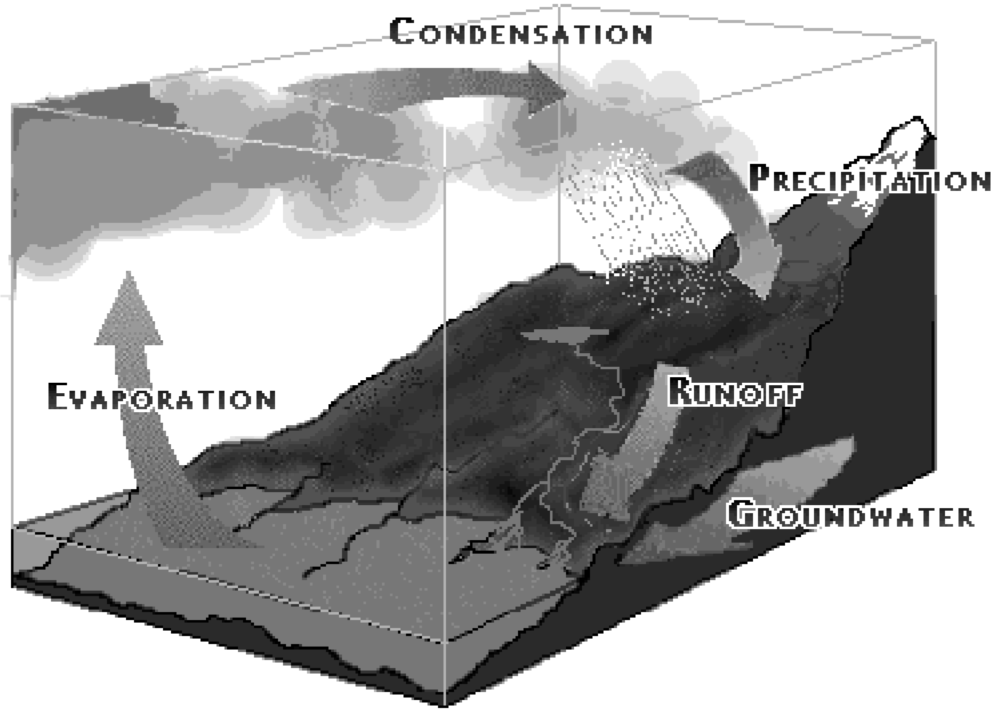

When water evaporates from the ocean and surface waters, it leaves behind impurities. In the absence of air pollution, nearly pure water falls back to the ground as precipitation

"We will always need resupply, because none of the water reprocessing technology that is available right now for space flight ... is 100 percent efficient. So there's always some minimal loss," said Marybeth Edeen, deputy assistant manager of environmental control and life support at NASA's Johnson Space Center.

Water is lost by the Space Station in several ways: the water recycling systems produce a small amount of unusable brine; the oxygen-generating system consumes water; air that's lost in the air locks takes humidity with it; and the CO2 removal systems leach some water out of the air, to name a few.

Lost water will be replaced by carrying it over from the Shuttle or from the Russian Progress rocket. The Shuttle produces water as its fuel cells combine hydrogen and oxygen to create electricity, and the Progress rocket can be outfitted to carry large containers of water.

NASA scientists will continue to look for ways to improve the life support systems of the Space Station, reducing water losses and finding ways to reuse other waste products. If the water recycling systems can be improved to an efficiency of greater than about 95 percent, then the water contained in the Station's food supply would be enough to replace the lost water, Edeen said.

One of the "nodes" that will become a part of the Space Station. The ECLSS life support equipment will be housed in Node 3, which is scheduled to be attached to the station in October 2005

Sign up for our EXPRESS SCIENCE NEWS delivery

"It takes processes that are slightly more efficient than we have developed for the space station to do that," Edeen said. "Those are the next generation water processing systems. Those are being developed now, but they're not ready for space flight yet."

The ECLSS life support system will join the Space Station as part of Node 3, which is scheduled to launch in October 2005. Until then, the environment inside the ISS will be maintained primarily by life support systems on the Russian Zvezda Service Module.

WEB LINKS

International Space Station—NASA's Web page for the International Space Station

Wheels in the Sky—Science@NASA article about humanity's dreams of a space station from the science fiction fantasies of the Nineteenth Century to Wernher von Braun's catalytic vision in the 1950s

Advanced Life Support Web Page—from the Johnson Space Flight Center

Environmental Control and Life Support Systems—describes the life support systems being developed at Marshall Space Flight Center

STRONTIUM ISOTOPES IN WATER AND ROCK

Gholam A. Kazemi
Shahrood University of
Technology
Shahrood, Iran

INTRODUCTION

A brief summary of strontium and strontium isotopes, their concentrations in natural materials, and their evolution in the geologic system is given in the following paragraph. These introductory notes have been taken from a benchmark paper by Capo et al. (1). Anyone intending to deal with strontium isotopes should consult that paper as well as the paper by Clow et al., (2). Textbooks such as those by Faure (3) and Faure (4) provide excellent coverage of strontium isotopes in rocks. Being an encyclopedia of water, the emphasis here is more on strontium isotopes in water.

Strontium has four naturally occurring isotopes, whose approximate abundances are ^{84}Sr, 0.56%; ^{86}Sr, 9.87%; ^{87}Sr, 7.04%; and ^{88}Sr, 82.53%. All four isotopes are stable, although the fraction of ^{87}Sr varies due to radioactive decay of ^{87}Rb. The ionic radius of strontium (1.18 Å) is similar to that of calcium (1.00 Å), and, hence, Sr substitutes for Ca in minerals, including plagioclase feldspar, apatite, sulfates such as gypsum and anhydrite, and carbonates (calcite, dolomite, and especially aragonite). The average concentration of Sr in soils is 240 ppm but may fall below 10 ppm or exceed 1000 ppm in some cases. Weathering of bed rock or sediments can be a significant source of strontium in soil. Soil parent materials often have distinct Sr isotopic signatures. The average concentration of Sr in crustal rocks is 370 ppm but can vary from 1 ppm in some ultramafic rocks to several percent in some aragonitic corals. If Rb and Sr are incorporated into a mineral or rock at its formation and the system remains closed with respect to those elements, then the amount of ^{87}Sr increases over time as radioactive ^{87}Rb decays; the amounts of ^{84}Sr, ^{86}Sr, and ^{88}Sr remain constant. Therefore, older rocks in general have higher ^{87}Sr/^{86}Sr ratios than younger ones that have the same initial Rb/Sr ratio. Over geologic time, rocks of a given age, composed of minerals that have a high Rb/Sr ratio (e.g., granites in the continental crust), develop a higher ^{87}Sr/^{86}Sr than rocks that have a lower Rb/Sr ratio (e.g., oceanic basalts). Thus, ^{87}Sr/^{86}Sr ratios in geologic materials are indicators of both age and geochemical origin. The strontium in rocks is released by weathering, cycled through vegetation and animals, and eventually enters the oceans, primarily via rivers. Minerals within a crystalline rock or sediment generally have variable Rb/Sr ratios and therefore a range of ^{87}Sr/^{86}Sr ratios. River water has a Sr concentration that varies from ~6 to 800 ppb and averages ~60 ppb. The Sr concentration in seawater is about 8 ppm. Precipitation near ocean basins generally has a strontium isotopic composition similar to that of seawater (0.709). However, the concentration of Sr in rainfall is generally lower than that in seawater by several orders of magnitude (usually <1 ppb). Strontium leaves the oceans, the largest reservoir of dissolved Sr, primarily by deposition in marine carbonate. A small amount of Sr is also transferred directly from the oceans to the atmosphere and hence transferred to the continents in precipitation.

SR ISOTOPES IN WATER AND ROCK: SOURCES AND IMPORTANCE

The Sr isotopic ratio in natural water reflects the contribution made by various minerals as they dissolve in water. Rivers are the major provider of strontium to the oceans and define, to a great extent, the marine ^{87}Sr/^{86}Sr ratio. Strontium in rivers is mainly derived from two end members: limestone and evaporites as one end member (high Sr and low ^{87}Sr/^{86}Sr ratio of 0.706–0.709) and silicate rocks as the other end member that are low in Sr and have a higher ^{87}Sr/^{86}Sr ratio. Processes such as atmospheric deposition, cation exchange, and biological cycling, too, influence the Sr isotope ratio in water. However, unless there have been recent anthropogenic or natural disturbances, annual gains or losses of Sr from cation exchange and biological pool should be minimal.

Sr isotopes in water samples are used (1) as a chemical tracer to determine the sources of atmospheric contributions to soil cation reservoirs over time (1); (2) to determine the contribution of various components to the water balance of a lake (5); (3) to distinguish weathering reactions from cation exchange processes (6); (4) to investigate the origin of brine, brine migration, and the relationship between brine and diagenetic phases (7); (5) to characterize weathering reactions and identify sources of dissolved Ca in stream water (2); (6) as a tracer of cation sources in stream water (8); (7) to correlate and date marine sequences of late Neogene age (9); (8) to identify the origin of salts in groundwater (10); (9) to investigate the effect of carbonate dissolution on the chemical composition of groundwater (11); and (10) to indicate the influence of seawater on fresh groundwater (12).

Analysis of Sr Isotopes in a water sample

The strontium isotopes of a water sample are measured under clean laboratory conditions using a mass spectrometer, such as a Sector 54 type (13), a Finnigan Mat 261 multicollector model (2), or a VG354 thermal ionization brand (9). To analyze water for a Sr isotope ratio using a Sector 54 Type Mass Spectrometer, an initial aliquot of the sample is placed in a microwave oven to be evaporated to dryness. For water that has a total Sr content of more than 0.5 ppm, the aliquot is 2 mL, and evaporation time is about 1 hour; for water that has low concentrations such as rainwater, 500 mL is required which takes much longer to evaporate to dryness. Then 3 mL of 6 molar HCl is added to the residue to dissolve it and left to stabilize for a few hours. The following steps are then taken one after another (13):

1. Add 1 mL of 2.5 molar HCl, and evaporate to dryness by heating slowly.
2. Add 1 mL of 6 molar HCl, and leave to dissolve.
3. Transfer to centrifuge vessels, and centrifuge for around 5 minutes.
4. Pour 1 mL of the sample into the column (the column detail is AG W50-X 8).
5. Add 1 mL of 2.5 molar HCl to the column.
6. Add 1 mL of 2.5 molar HCl to the column.
7. add 1 +30 mL of 2.5 molar HCl to the column in two stages, respectively.
8. Add 10 mL of 2.5 molar HCl to the column.
9. Collect Sr by placing a container beneath the column and letting it drain.
10. Add 1 mL of HNO_3, and slowly heat to dryness.
11. Add 1 mL of HNO_3, and slowly heat to dryness.
12. Add 1 mL of Milli-Q water to the sample.
13. Place a small volume of the sample in the holding filament.
14. Heat sample slowly.
15. Add a small volume of phosphoric acid to the sample to facilitate ionization.
16. Heat the sample rapidly to a very high temperature (to burn) to remove any Rb.
17. Place the sample in the sample holder.
18. Transfer sample/s to the mass spectrometer.
19. Start running sample with mass spectrometer.

Analysis of Sr Isotopes in a rock sample

Rocks Other Than Limestone. To analyze a rock sample for Sr isotopes ratio using a Sector 54 Type Mass Spectrometer under clean laboratory condition, first it has to be ground to a 1 micron size. After grinding, for all rock types but limestone, the following steps should be undertaken one after another:

1. 0.1 gram of sample to accuracy of 0.0001 gram is selected.
2. 2 mL of HF acid and 10 drops of HNO_3 are added to each sample.
3. Samples are heated to dryness.
4. 2 mL of HF is again added to each sample.
5. 10 drops of $HClO_4$ are added to each sample, and sample is placed on a hot plate at 150 °C.
6. 5 drops of $HClO_4$ are added to each sample.
7. 5 more drops of $HClO_4$ are added to each sample.
8. Samples are placed on a hot plate at 180 °C.
9. Samples are then placed on a hot plate at 130 °C on the next day, and 1 mL 2.5 M HCl is added to each sample.
10. Samples are transferred to centrifuge tube to be centrifuged for 5 minutes.
11. From this step forward, all the steps are exactly the same as those for water samples from Step 5 onward.

Limestone. Due to its high solubility compared to other rock types, limestone is prepared for analysis using different techniques and the following steps:

1. 0.1 mg of sample is selected.
2. 1 mL of 1 M acetic acid is added to the sample.
3. 2 mL of 1 M acetic acid is added.
4. After reaction subsides, sample is centrifuged, and the residue is decanted to a clean beaker and heated to dryness at 150 °C.
5. 3 mL of 6.25 M HCl is added and evaporated to dryness, and 1 mL more is added again and heated to dryness.
6. 1 mL of 2.5 M HCl is added.
7. The sample is then loaded into columns (like other rock types).
8. The Sr is then collected and dried.
9. The columns are flushed and cleaned.
10. The dry sample (Sr) is again dissolved with 1 mL of 2.5 M HCl and loaded into the columns again. In contrast to the other samples, carbonates have a higher calcium content. Therefore, they need to be cleaned up twice before they are loaded into the mass spectrometer.
11. The rest of the procedure is like that for other rock types.

REPORTING THE RESULTS OF SR ISOTOPIC ANALYSIS

Before reporting the results of the strontium isotopic analysis, they should first be corrected for mass fractionation to $^{88}Sr/^{86}Sr = 8.37521$. Also, the value of isotopic ratio ($^{88}Sr/^{86}Sr$) of the NBS (National Bureau of Standards) 987 Sr metal used as a reference by the spectrometer should be included in the report. Measured strontium isotopic values are expressed as both ratios (ratio of ^{87}Sr to ^{86}Sr) and standard delta notation ($\delta^{87}Sr$) relative to NBS standard 987. To calculate $\delta^{87}Sr$, the following formula is used:

$$\delta^{87}\mathrm{Sr} = \left[\frac{(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})\mathrm{sample}}{(^{87}\mathrm{Sr}/^{86}\mathrm{Sr})\mathrm{seawater}} - 1\right] \times 10^5$$

where $(^{87}Sr/^{86}Sr)_{sample}$ is the measured ratio of $^{87}Sr/^{86}Sr$ in a sample that has been normalized to $^{86}Sr/^{88}Sr = 0.1194$ and $(^{87}Sr/^{86}Sr)_{seawater}$ is the ratio of $^{87}Sr/^{86}Sr$ in modern seawater which is about 0.709172 (9). In most analysis, the precision of the $^{87}Sr/^{86}Sr$ ratios is ± 0.0001 or better, and higher precision is of little value because seasonal variations in $^{87}Sr/^{86}Sr$ ratios in rivers are commonly in the third figure (14).

In a majority of cases, the Sr isotopic composition of a water or rock sample is a mixture of two or more end members. To calculate the relative proportion of each end member of a two component mixture, the following equation applies:

$$(^{87}\mathrm{Sr}/^{\mathrm{total}}\mathrm{Sr})_{\mathrm{m}} = x(^{87}\mathrm{Sr}/^{\mathrm{total}}\mathrm{Sr})_{\mathrm{a}} + (1 - x)(^{87}\mathrm{Sr}/^{\mathrm{total}}\mathrm{Sr})_{\mathrm{b}} \quad (1)$$

where the subscript "m" is mixture, "a" is the first end member, "b" is the second end member, x is the proportion of end member a, and $(1-x)$ is the proportion of end member b. To solve this equation, $^{87}Sr/^{total}Sr$ can be calculated from the equation given by (Reference 2) as follows:

$$^{87}\mathrm{Sr}/^{\mathrm{total}}\mathrm{Sr} = (^{87}\mathrm{Sr}/^{86}\mathrm{Sr})/(3 +^{87}\mathrm{Sr}/^{86}\mathrm{Sr}) \quad (2)$$

Example (from Reference 13). In Table 1, the isotopic composition of groundwater (GW) is considered a mixture of soil and rainwater isotopic compositions. The relative proportion of Sr from rainfall and soil to groundwater in Piezometer 40, P40, is computed here. In Equation 1, "m" = groundwater; "a" = rainfall; and "b" = soil; x = proportion of end member a, rainfall; and $1-x$ = proportion of end member b, soil. To calculate the proportion of each end member, first the $^{87}Sr/^{total}Sr$ ratio (Eq. 2) must be computed:

$$\begin{aligned}
^{87}\mathrm{Sr}/^{\mathrm{total}}\mathrm{Sr}\text{ for rainfall} &= (0.7089)/(3 + 0.7089)\\
&= 0.191135\\
^{87}\mathrm{Sr}/^{\mathrm{total}}\mathrm{Sr}\text{ for soil} &= (0.72068)/(3 + 0.72068)\\
&= 0.19368\\
^{87}\mathrm{Sr}/^{\mathrm{total}}\mathrm{Sr}\text{ for groundwater} &= (0.7105)/(3 + 0.7105)\\
&= 0.191091
\end{aligned}$$

Substituting these values in Eq. 1, the relative contribution of rainfall to the Sr budget of groundwater in P40 is calculated as 86%, and only 14% of the Sr is from soils.

Case Studies

Kazemi (13) and Kazemi and Milne-Home (15), through a strontium isotope study of rainwater, surface water, groundwater (shallow and deep), soil, and rocks of a salinized watershed in Eastern Australia (Buckinbah Creek watershed) found that the ratio of $^{87}Sr/^{86}Sr$ of shallow groundwater was similar to that of soil and

Table 1. The ^{87}Sr/^{86}Sr of Various Materials in the Buckinbah Creek Watershed in Eastern Australia[a]

GW (Shallow)		GW (Deep)		Surface Water		Rocks[b]		Soils	
B7	0.71197	B4	0.7084	S6	0.7099	Kabadah Fm[c]	0.70555	P45-Top	0.709
P19	0.7058	B3	0.708	S2	0.7087	Wansey Fm[d]	0.70813	P41	0.7373
P37	0.71	B2	0.7083	S5	0.7076	CV	0.72547	P42-Top	0.7135
P45	0.7078	B1	0.7084			Kabadah Fm[e]	0.7051	P40	0.72068
P40	0.7105					Burrawong Ls	0.70812	P40B	0.72298
B5	0.7099						0.71047		0.720692
P41	0.7169			Rain	0.7089				
P39	0.7092			Leaf	0.7115				
P42	0.7125								
P15	0.71023								

[a] *Source*: References 13.

[b] CV: Canowindra Volcanics, Burrawong Ls: Burrawong limestone, Fm: Formation.

[c] Rock type- shale.

[d] Rock type- volcanoclastic sandstone.

[e] Rock type- volcanoclastic.

rainwater (Table 1). They showed that between 20 to 90% of the Sr budget of shallow groundwater and surface waters is derived from rainwater. They further concluded that the deep groundwater is the source of baseflow at the watershed outlet (due to identical ^{87}Sr/^{86}Sr ratios of deep groundwater and baseflow).

Hunt et al. (6) used water and strontium isotopes to investigate the hydrology of a natural and a constructed wetland. They showed that the porewater from the natural wetland had a ^{87}Sr content ranging from 0.00 to 0.86 per mil (^{87}Sr/^{86} Sr $= 0.71026$ to 0.71087) and the ^{87}Sr of pore water from the constructed wetland ranged from -2.35 to $+1.24$ per mil (^{87}Sr/^{86} Sr $= 0.70859$ to 0.71114). They attributed this difference to the presence of a thick peat layer in the natural wetland.

The ^{87}Sr/^{86}Sr ratio of 32 brine samples from the Upper Jurassic Smackover Formation in southern Arkansas was measured at 0.7071–0.7101, more radiogenic than the Sr in Late Jurassic seawater because of the contribution of Sr from detrital sources (7).

The linear negative correlation between the Sr and ^{87}Sr/^{86}Sr ratio (when plotted against each other) has been used by Faure (3), Dogramaci et al. (16), and Singh et al. (17) as proof of the mixing of two waterbodies. Similarly, the negative correlation between the molar ratio of Mg/Ca and Sr isotopes is an indication of the incongruent dissolution of calcitic minerals in the aquifer (11). Incongruent dissolution of carbonates occurs progressively through two mechanisms:

1. Dissolution of calcitic minerals, causing the addition of Sr to groundwater and modification of the ^{87}Sr/^{86}Sr ratio; and
2. Reprecipitation of calcitic minerals, thereby removing some of the Sr from the groundwater without changing the Sr isotope ratio, approaching eventually the ratio in calcitic minerals (11).

A large data set of Sr isotopic ratios of the rivers of the Ganges, Orinoco, and Amazon basins is found in Palmer and Edmond (14).

The study by Lyons et al. (18) focuses on the strontium isotopic signature of some streams and lakes in Antarctica. They found that the ratios of ^{87}Sr/^{86}Sr of Lake Fryxell, Lake Hoare, and Lake Bonney were 0.70895, 0.71057, and 0.71187–0.71204, respectively. They also found that in Lake Fryxell and Lake Hoare, the ^{87}Sr/^{86}Sr became less radiogenic with depth, whereas in Lake Bonney, the ratios increase slightly with depth.

Sr isotopes together with isotopes of Nd and Pb of Cenozoic basalts in eastern Australia were studied by Zhang et al. (19) to show that the isotopic signatures of the Pacific Ocean midoceanic ridge basalts (MORB) type characterize the lava-field basalts in southeastern Australia, whereas the mantle isotopic signatures of Indian Ocean MORB type characterize younger basalts from northeastern Australia.

BIBLIOGRAPHY

1. Capo, R.C., Stewart, B.W., and Chadwick, O.A. (1998). Strontium isotopes as tracers of ecosystem processes: Theory and methods. *Geoderma* **82**: 197–225.
2. Clow, D.W., Alisa Mast, M., Bullen, T.D., and Turk, J.T. (1997). Strontium 87/strontium 86 as a tracer of mineral weathering reactions and calcium sources in an alpine/subalpine watershed, Loch Vale, Colorado. *Water Resour. Res.* **33**(6): 1335–1351.
3. Faure, G. (1986). *Principles of Isotope Geology*. John Wiley & Sons, New York.
4. Faure, G. (2001). *Origin of Igneous Rocks: The Isotopic Evidence*. Springer, Heidelberg.
5. Clauer, N., Chauduhuri, S., Toulkeridis, T., and Blanc, G. (2000). Fluctuations of caspian sea level: Beyond climatic variations. *Geology* **28**(11): 1015–1018.
6. Hunt, R.J., Bullen, T.D., Krabbenhoft, D.P., and Kendall, C. (1998). Using stable isotopes of water and strontium to investigate the hydrology of a natural and a constructed wetland. *Groundwater* **36**(3): 434–443.
7. Stueber, A.M., Pushkar, P., and Hetherington, E.A. (1984). A strontium isotopic study of Smackover brines and associated solids, southern Arkansas. *Geochim. Cosmochim. Acta* **48**: 1637–1649.
8. Miller, E.K., Blum, J.D., and Friedland, A.J. (1993). Determining of soil exchangeable-cation loss and weathering rates using Sr isotopes. *Nature* **362**: 438–441.
9. Hodell, D.A., Mead, G.A., and Mueller, P.A. (1990). Variation in the strontium isotopic composition of seawater (8 Ma to

present): Implications for chemical weathering rates and dissolved fluxes to the ocean. *Chem. Geol.* **80**: 291–307.

10. Morgan, K. and Jankowski, J. (2003). Origin of salts in a fractured bedrock aquifer in the Central West Region, NSW, Australia. *Proc. Int. Conf. Groundwater Fractured Rocks*, Prague, Krasny-Hrkal-Bruthans (Eds.). IHP-VI, series on groundwater no. 7, pp. 221–222.
11. Dogramaci, S.S., Herczeg, A.L., and Bone, Y. (1998). ^{87}Sr/^{86}Sr in Groundwater as Indicators of Carbonate Dissolution. *Proc. 9th Int. Symp.* on *water–rock Interaction*, Taupo, New Zealand, March–April, 1998, G.B. Arehart and J.R. Hulston (Eds.). A.A. Balkema, Rotterdam, pp. 211–214.
12. Jorgensen, N.O. and Heinemeier, J. (2003). Isotope studies of saline groundwater in a sandstone aquifer on Bornholm, Denmark. *Proc. Int. Conf. Groundwater Fractured Rocks*, Prague, Krasny-Hrkal-Bruthans (Eds.). IHP-VI, series on groundwater no. 7, pp. 205–206.
13. Kazemi, G.A. (1999). *Groundwater factors in the management of dryland and stream salinity in the upper Macquarie Valley, New South Wales, Australia*. Unpublished Ph.D. Thesis, University of Technology, Sydney.
14. Palmer, M.R. and Edmond, J.M. (1992). Controls over the strontium isotope composition of river water. *Geochim. Cosmochim. Acta* **56**: 2099–2111.
15. Kazemi, G.A. and Milne-Home, W.A. (1999). The application of strontium, oxygen 18 and deuterium isotopes in the estimation of the solute budget and hydrology of a small watershed in Eastern Australia. *Eos Trans. AGU* **80**(17) *Spring Meet*. Suppl. S139, 1999.
16. Dogramaci, S.S., Herczeg, A.L., and Bone, Y. (1998). *Chemical evolution of groundwater in the Murray Basin: Geochemistry and mass transfer modelling*. Abstract No. 49, Geol. Soc. Aust., 14th Aust. Geol. Convention.
17. Singh, K.S. et al. (1998). Chemical and strontium, oxygen, and carbon isotopic compositions of carbonates from the Lesser Himalaya: Implications to the strontium isotope composition of the source waters of the Ganga, Ghaghara, and the Indus rivers. *Geochim. Cosmochim. Acta* **62**: 743–755.
18. Lyons, W.B. et al. (2002). Strontium isotopic signatures of the streams and lakes of Taylor Valley, Southern Victoria Land, Antarctica: Chemical weathering in a polar climate. *Aquatic Geochem.* **8**(2): 75–95.
19. Zhang, M., O'Reilly, S.Y., and Chen, D. (1999). Location of Pacific and Indian mid-ocean ridge-type mantle in two time slices: Evidence from Pb, Sr, and Nd isotopes for Cenozoic Australian basalts. *Geology* **27**(1): 39–42.

TECHNETIUM IN WATER

Jay Clausen
AMEC Earth and
Environmental, Inc.
Westford, Massachusetts

INTRODUCTION

Technetium, element 43, was predicted on the basis of its location in the periodic table. Technetium was erroneously reported as having been discovered in 1925 by Noddack and Tacke (1), who called it masurium. However, Element 43 was actually discovered in Italy (2). It was found in a sample of molybdenum, which had been bombarded with deuterons in the University of California—Berkeley cyclotron by E. Lawrence and then sent to Perrier and Segre. Technetium was the first element to be produced artificially.

Twenty-two isotopes of technetium whose masses range from 90 to 111 are known, and all are radioactive (Table 1). Technetium has three long-lived radioactive isotopes: technetium-97 [^{97}Tc] (half-life $\{T_{1/2}\} = 2.6 \times 10^6$ years), ^{98}Tc ($T_{1/2} = 4.2 \times 10^6$ years), and ^{99}Tc ($T_{1/2} = 2.13 \times 10^5$ years). An important isotope is ^{95m}Tc ("m" stands for metastable state) [$T_{1/2} = 61$ days], which is used in tracer work. However, the most useful isotope of technetium is ^{99m}Tc ($T_{1/2} = 6.01$ hours), which is used in many medical radioisotope tests because of its short half-life, easy detection of the gamma ray it emits, and the ability of technetium to bind chemically to many biologically active molecules.

Technetium-99 is a weak, beta-emitting radionuclide, 0.292 million-electron volts (MeV), produced during nuclear fission, which decays to the stable isotope ruthenium-99. It has a specific activity of 3.78×10^{10} transmutations per minute per gram, which corresponds to 17 microcuries per milligram (μCi/mg) (4). Technetium metal is a silver-gray color, tarnishes slowly in the air, and has properties similar to those of rhenium. However, as mentioned earlier, technetium does not occur naturally.

However, ^{99}Tc in terrestrial material has been evaluated since its discovery. In 1961, ^{99}Tc was isolated and identified in African pitchblende (a uranium-rich ore) in extremely minute quantities (5). It was theorized that ^{99}Tc was produced as a result of the spontaneous fission of uranium-238 (^{238}U).

Technetium-99 is produced primarily through anthropogenic processes such as detonation of nuclear weapons, reprocessing of spent nuclear reactor fuel rods containing uranium, and disposal of nuclear waste. Reprocessing involves dissolving the spent nuclear fuel in acid to separate the potentially reusable components of the spent fuel

Table 1. Isotopes, Half-Lives, and Decay Processes for Technetium[a]

Nuclide	Half-Life[b]	Type of Decay	Nuclide	Half-Life	Type of Decay
^{91}Tc	3.2 m	β^+, γ	^{99}Tc	2.13×10^5 y	B^-
^{92}Tc	4.4 m	β^+, γ	^{100}Tc	15.8 s	B^-, γ
^{93m}Tc	43.5 m	ε, γ	^{101}Tc	14 m	B^-, γ
^{93}Tc	2.7 h	$\varepsilon, \beta^+, \gamma$	^{102m}Tc	4.3 m	B^-, γ
^{94}Tc	53 m	β^+, γ	^{102}Tc	6.3 m	B^-, γ
^{94}Tc	4.9 h	$\varepsilon, \beta^+, \gamma$	^{103}Tc	50 s	B^-, γ
^{95m}Tc	60 d	$\varepsilon, \beta^+, \gamma$	^{104}Tc	18.0 m	B^-, γ
^{95}Tc	20 h	ε, γ	^{105}Tc	7.6 m	B^-, γ
^{96m}Tc	52 m	ε, γ	^{106}Tc	36 s	B^-, γ
^{96}Tc	4.3 d	ε, γ	^{107}Tc	21 s	B^-, γ
^{97m}Tc	91 d	γ	^{108}Tc	5 s	B^-, γ
^{97}Tc	2.6×10^6 y	ε	^{109}Tc	1 s	B^-
^{98}Tc	4.2×10^6 y	B^-, γ	^{110}Tc	0.83 s	B^-, γ
^{99m}Tc	6 h	γ, B^-	^{111}Tc	?	?

[a] Reference 3.

[b] d: days; h: hours; m: minutes; s: seconds; y: years.

[c] β^+: positron; β^-: beta particle; ε: internal electron conversion; γ: gamma ray.

(the uranium and plutonium) from the waste component (the fission products). The fission products, which are of no use, must be disposed of properly. In a 1000-megawatt nuclear reactor (thermal fission yield of approximately 6%), approximately 28 g of ^{99}Tc is produced daily from the slow neutron-induced fission of ^{235}U or plutonium-239 [^{239} Pu] (6). The long half-life and mobility of ^{99}Tc makes it an important element to the nuclear fuel industry in terms of its potential environmental impact. This is a particular concern for designing underground nuclear waste storage facilities because ^{99}Tc is the most mobile of all radionuclides present.

There are no known uses of technetium other than ^{99m}Tc for medical purposes. Technetium has superconducting properties and is a good corrosion preventer (7). However, its unstable nature and production of radioactive particles negate its practical use.

ANALYSIS

Analysis for ^{99}Tc typically consists of methyl ethyl ketone (MEK) extraction followed by liquid scintillation counting using United States Environmental Protection Agency (USEPA) Method 900.0. The typical quantification limit is 1 becquerel per liter (Bq/L), which is equivalent to 25 picocuries per liter (pCi/L).

Chemistry

The oxidation states most commonly found for ^{99}Tc are Tc^0, Tc^{+4}, and Tc^{+7}. The most stable chemical species of ^{99}Tc in aqueous solutions is the pertechnetate anion, TcO_4^-, in the Tc^{+7} valence state, which is highly mobile and nonreactive (8–12). However, ^{99}Tc is very sensitive to changes in redox conditions (13). Under reducing conditions, ^{99}Tc can be precipitated by various cations to form anhydrous technetium dioxide, TcO_2, (as Tc^{+4}) or technetium heptoxide, Tc_2O_7, or other insoluble oxides and hydroxides such as Tc_2S_7 or TcS_2 or incorporated into minerals (14–17):

$$TcO_4^- + 4H^+ + 3e^- = TcO_2 + 2H_2O;\ E_0 = +0.738 \text{ volts} \tag{1}$$

In the presence of oxygen, the Tc^{+4} oxide or oxyhydroxide can be reoxidized (14,18):

$$[TcO_{2\,(s)} \text{ or } TcO(OH_{2\,(s)}] \rightarrow TcO_4^- \tag{2}$$

The pertechnetate anion also forms complexes with organic compounds containing –C –O and –COO functional groups, especially under reducing conditions. The thermodynamic data for technetium (Table 2) can be used to develop a phase diagram to identify which species are likely to be present at a given set of E_h and pH conditions (Fig. 1).

The pertechnetate ion is mobile under oxidizing conditions (Fig. 1). Bidoglio et al. (20,21) found in column experiments that the oxidized species persists, even under slightly reducing or anoxic conditions. Technetium can undergo a number of surface reactions on various minerals under reduced conditions.

Table 2. Thermodynamic Data for Technetium[a]

Species[b]	ΔG^0_f kcal/gfw[c]
$TcOH_{(s)}$	−56.1
$Tc(OH)_{2\,(s)}$	−110.2
$Tc_3O_{4\,(s)}$	−206.3
$TcO_{4\,(s)}^-$	−149.1
$TcS_{2\,(s)}$	−51.0
$TcO_{2\,(s)}$	−88.3

[a] Reference 19.
[b] s: solid.
[c] kcal/gfw: kilocalories per gram formula weight

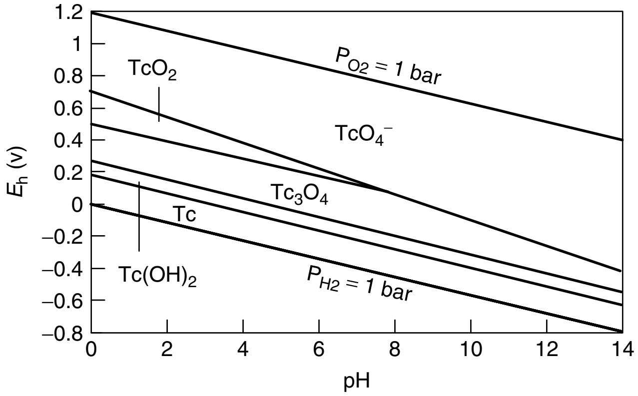

Figure 1. E_h-pH diagram for the Tc−O−H system at a concentration of 90 μg/L under standard temperature and pressure.

The solubility of sodium pertechnetate is reportedly 11 moles per liter [mol/L] (13). The solubility of ^{99}Tc in a reducing environment varies from 1.32×10^{-8} to 2.56×10^{-9} mol/L and is controlled by the Tc species present (14). Redox thus is a key variable affecting solubility (8). The solubility of technetium under reducing conditions in carbonate groundwater is pH dependent. At pH levels between 5 and 7.5, the solubility of Tc^{+4} is less than 10^{-8} mol/L. Pilkington (22) suggests that the solubility of TcO_2 is closer to 10^{-7} mol/L under reducing conditions. At higher pHs of 8 to 11, solubility increases, and the dominant ionic complex is $TcO(OH)_3CO^-_3$.

Equally important are the fate-and-transport properties of ^{99}Tc. Under aerobic soil conditions, ^{99}Tc behaves as an inert tracer and readily passes through soil unless there is a high content of organic matter or metal oxide (10,23–30). As Gu and Dowlen (31) report, this is not surprising because soil mineral surfaces are negatively charged as is TcO_4^- and therefore the two repel each other. Once introduced to surface water or groundwater, it can be transported significant distances due to its recalcitrant nature and lack of sorption. Given the concern over ^{99}Tc mobility, extensive studies have been conducted in various geologic media to assess suitable nuclear waste disposal storage facilities. Adsorption coefficients for various minerals are presented in Sheppard et al. (23), Wildung et al. (27), Palmer and Meyer (32), and

Strickert et al. (33). Sheppard et al. (23) indicate that soil redox conditions and organic matter content are the key variables influencing ^{99}Tc mobility in soil.

REGULATORY STANDARDS

The USEPA has set an effective dose equivalent of 4 millirem (mrem)/year for exposure, which they incorrectly calculated as 900 pCi/L (36 Bq/L). The correct activity equivalent to 4 mrem/year is 3790 pCi/L [152 Bq/L] (34). The U. S. Department of Energy (DOE) Order 5400.5 sets a permissible ingested-water limit of 100 mrem/year or 100,000 pCi/L (4000 Bq/L), whereas the U. S. Nuclear Regulatory Agency (NRC) has set a value of 50 mrem/year or 60,000 pCi/L (2400 Bq/L). The NRC has set the maximum activity of ^{99}Tc that can be released to a sanitary sewer at 500 mrem/year or 600,000 pCi/L (24,000 Bq/L). The proposed action level for groundwater cleanup at the Paducah Gaseous Diffusion Plant (PGDP) site is 25 pCi/L (1 Bq/L), which is the quantification limit.

Rainwater

Technetium has been released to the atmosphere in the form of Tc_2O_7 and $HTcO_4$ as a consequence of aboveground detonation of nuclear weapons through the fission of ^{235}U and ^{239}Pu. The nuclear industry also releases TcF_6 in addition to Tc_2O_7 and $HTcO_4$ during uranium enrichment. During a nuclear detonation, elemental ^{99}Tc burns to form Tc_2O_7, which upon cooling and contact with water vapor forms $HTcO_4$. The activity of ^{99}Tc measured in rainwater reportedly ranges from 8.2×10^{-8} to 1.0×10^{-6} μg/L (35,36). Work by Leon (37) and Attrep et al. (36) suggest the concentration of ^{99}Tc in rainwater has increased from the early 1960s through the 1980s. Both papers suggest that the source of this increase in ^{99}Tc in rainwater is from the reprocessing of spent nuclear fuel rather than nuclear weapons testing.

Surface Water

Technetium-99 is a surface water and groundwater contaminant at a number of DOE sites such as the PGDP in Paducah, Kentucky; the Portsmouth Gaseous Diffusion Plant in Piketon, Ohio; Oak Ridge National Laboratory Y-12 Plant, Savannah River Laboratory in Aiken, South Carolina; the Pacific Northwest National Laboratory 200 west area; and the Nevada Test Site (30,38–40).

Currently, the largest annual release of ^{99}Tc to the environment occurs at the two nuclear fuel-reprocessing facilities in Sellafield, United Kingdom. Technetium-99 laden effluent water from these facilities is discharged directly into the Irish Sea.

Contamination has been detected as far west as Svalbard and along the Norwegian coastline, distances of several hundred miles. The total discharge from these two plants amounts to between 70 and 90 tera (TBq)/year of ^{99}Tc (41).

Techentium-99 is reported in the Savannah River as a result of discharges from the U. S. Department of Energy Savannah River Plant located near Aiken, South Carolina (42,43). Technetium-99 is present at Oak Ridge National Laboratory where radioactive liquids generated from routine operations were disposed of in shallow waste pits (44). Nearby seeps have ^{99}Tc activity levels of up to 1185 pCi/L [3,200 Bq/L].

Blaylock and Frank (45) conducted a study in which $TcO_4{}^-$ was added to a freshwater pond. Their study indicated that 84% of the $TcO_4{}^-$ added remained in solution; very little was taken up by biota or by settling out into sediments. The vast majority of the ^{99}Tc was removed through drainage, indicating that the ultimate reservoir for technetium released into the environment is the ocean. Similarly, in a study of the Rhone River, it was found that the ^{99}Tc activity varied from 2.5 to 53×10^{-3} pCi/L (0.1 to 2.1×10^{-3} Bq/L) in seven samples collected in 1985 and 1986 (46).

Finally, significant quantities of ^{99}Tc laden effluent water were released to a settling lagoon as well as to several small streams surrounding the PGDP. Activity levels as high as 100,000 pCi/L (4000 Bq/L) were routinely released to surface water. Ferrous iron sulfate was used to precipitate metals in the lagoon before discharge into a nearby stream. The addition of this material shifted the redox state of the lagoon such that $TcO_4{}^-$ was precipitated from solution. The likely form of the technetium precipitate was Tc_2S_7. In addition, the iron-rich clay and organic-rich stream sediments retained much of the ^{99}Tc, limiting its downgradient migration from the lagoon. Therefore, the ^{99}Tc migrated only several thousand feet from the point of release. However, downgradient of the site, groundwater discharges to the surface water streams and ^{99}Tc-rich water has been measured in several seeps (47).

Ground Water

The most extensive groundwater distribution of ^{99}Tc is at the PGDP where a 2-$\frac{1}{2}$ mile long plume of ^{99}Tc contaminated groundwater is commingled with the solvent trichloroethene. The ^{99}Tc was introduced to the site during the 1970s when fissioned uranium was reprocessed. ^{99}Tc became entrained in processing equipment at the PGDP, which was later cleaned with organic solvents and released to the environment. Gu et al. (48) determined conclusively that the species of ^{99}Tc mobile in groundwater at the PGDP was the pertechnetate anion, $TcO_4{}^-$. This later was supported through the development of phase diagrams using site-specific E_h and pH groundwater data (38,39). In addition, Gu and Dowlen (31) and Gu et al. (48) conclusively ruled out colloidal and natural organic matter facilitated transport as viable mechanisms of ^{99}Tc mobilization. Previous studies of PGDP, other sites, and laboratory studies assumed that $TcO_4{}^-$ was the form of ^{99}Tc in most environmental settings. The estimated partitioning coefficient of ^{99}Tc exceeded 27,000 milliliter per gram [mL/g] (48). Technetium-99 activities as high as 179 Bq/L (4800 pCi/L) were observed in the groundwater plume (38,39). The estimated mass of ^{99}Tc dissolved in the groundwater plume was 4 to 6 lb, which equates to an activity of 1.0 to 1.6×10^{12} Bq/L (28 to 42×10^{12} pCi) of ^{99}Tc.

At the Hanford Site in Washington, ^{99}Tc has been found in groundwater in the Hanford 200 West Area where spent nuclear fuel was reprocessed and then the waste material was stored in underground tanks (30). Technetium-99 levels between 900 and 27,000 pCi/L (36 and 1080 Bq/L) have been found in groundwater.

Technetium-99 activity in shallow groundwater at the Portsmouth Gaseous Diffusion Plant has been observed at the 1000 pCi/L (40 Bq/L) level (49). The causes of contamination are similar to those discussed for the Paducah site. However, because the Portsmouth site received feed material from the Paducah site, the quantity of ^{99}Tc introduced into the site was much lower.

A unique study was conducted by the Lawrence Livermore National Laboratory at the Nevada Test site to assess the distribution of radionuclides in caverns produced by underground nuclear testing. The work by Buddemeier et al. (50) and Schroeder et al. (40) found that ^{99}Tc was present in groundwater associated with two separate nuclear weapons tests. The ^{99}Tc appeared to be associated and transported with colloidal material, but its mobility was reduced upon contact with zones of high levels of iron oxyhydroxides.

There are several known nongovernmental sites containing ^{99}Tc. One such facility is the Fields Brook National Priority Listed (NPL) site in Ashtabula, Ohio. The site was formerly a storage site for liquid waste, which contained ^{99}Tc as well as other radionuclides and solvents. Technetium-99 activities as high as 100,000 pCi/L (4000 Bq/L) have been detected in groundwater. A second site is the uranium fuel plant in Hematite, Montana. Recycled uranium, containing ^{99}Tc, was used to fabricate metallic fuel rods. Low-levels of ^{99}Tc have been observed in groundwater.

REMEDIATION

The technologies explored for remediating ^{99}Tc in water include ion exchange, liquid–liquid extraction, precipitation with various forms of iron, redox manipulation, the FORAGER sponge, granular activated carbon (GAC), and natural attenuation (10,16,18,31–33,48,51–65). The report by Bostick et al. (53) provides an excellent reference list of the many studies conducted on ^{99}Tc removal from aqueous waste streams. At the Portsmouth and Paducah Gaseous Diffusion plants, ion exchange is being used for treating ^{99}Tc-contaminated groundwater. Although, Gu et al. (31) found that GAC was superior to ion exchange due to its low cost, efficiency, and sorption capacity, GAC has not been employed at these sites.

CONCLUSION

Technetium is an element not typically found in nature and has environmental consequence at facilities that enrich or reprocess uranium fissioned material. These facilities are typically government operated. The high solubility of TcO_4^- and its stability across a wide spectrum of redox conditions enables the ion to be extremely mobile once introduced into water. As a consequence, at the few sites where ^{99}Tc has been introduced into the natural environment, it has produced very large plumes of contaminated surface and groundwater.

BIBLIOGRAPHY

1. Noddack, W. and Tacke, I. (1925). Die Ekamangane. *Naturwissenschaften* **13**: 567.
2. Perrier, C. and Segre, E. (1937). *Nature* (London). **140**: 193.
3. Seelmann-Eggebert, W., Pfennig, G., and Munzel, H. (1974). *Chart of the Nuclides*. Gersbach Verlag, Munich, Germany.
4. Schwochau, K. (1981). The analytical chemistry of technetium. *Top. Curr. Chem.* **96**: 109–147.
5. Kenna, B.T. and Kuroda, P.K. (1961). Technetium 99. *J. Inorg. Nucl. Chem. Lett.* **23**: 142.
6. De Alleluia, T.B. (1983). Tc: Technetium, System No. 69. In: *Gmelin Handbook of Chemistry*. Supplement Vol. 1, 8th Edn. Springer-Verlag, Berlin.
7. Cartledge, G.H. (1973). Twenty-year inhibition of corrosion by the pertechnetate ion. *Corrosion* **29**: 361–362.
8. Eriksen, T.E., Ndalamba, P., Bruno, J., and Caceci, M. (1992). The solubility of $TcO_2 \cdot nH_2O$ in neutral to alkaline solutions under constant p_{CO_2}. *Radiochimica Acta* **58/59**: 67–70.
9. Bruhl, H., Bock, W., Trapp, C., and Winkler, A. (1989). *Ausbreitungsverhalten von Technetium (Selen) in geochemisch unterschiedlichen Grundwasserleitern*. Institut fur Angeandte Geologie Freie Unversitat Berlin. Final report 02 U 5695 A9.
10. Vandergraaf, T.T., Ticknor, K.V., and George, I.M. (1984). *Reactions Between Technetium in Solution and Iron-Containing Minerals Under Oxic and Anoxic Conditions*. Atomic Energy of Canada. Limited Report AECL-1957. Pinawa, Manitoba, Canada, pp. 25–43.
11. Paquette, J., Reid, J.A.K., and Rosinger, E.L.J. (1980). *Review of Technetium Behavior in Relation to Nuclear Waste Disposal*. Whiteshell Nuclear Research Establishment. Technical Record TR-25. Pinawa, Manitoba, Canada.
12. Allard, B., Kigatsi, H., and Torstenfelt, B. (1979). Technetium: Reduction and sorption in granitic bedrock. *Radiochem. Radioanalytical Lett.* **37**: 223–230.
13. Gu, B. and Schulz, R.K. (1991). *Anion Retention in Soil: Possible Application to Reduce Migration of Buried Technetium and Iodine*. U.S. Nuclear Regulatory Commission, Washington, DC.
14. Meyer, R.E., Arnold, W.D., and Case, F.I. (1987). *The Solubility of Electrodeposited Tc(IV) Oxides*. Oak Ridge National Laboratory, Oak Ridge, TN.
15. Lieser, K.H. and Bauscher, C.H. (1987). Technetium in the hydrosphere and in the geosphere 1. Chemistry of technetium and iron in natural waters and influence of the redox potential on the sorption of technetium. *Radiochimica Acta* **42**: 205–213.
16. Lee, S.K. and Bondietti, E.A. (1983). Technetium behavior in sulfide and ferrous iron solutions. *Mater. Res. Soc. Symp. Proc.* **15**: 315–321.
17. Wildung, R.E., Routson, R.C., Serne, R.J., and Garland, T.R. (1974). *Pertechnetate, Iodide, and Methyl Iodide Retention by Surface Soils. In Pacific Northwest Laboratory Annual Report for 1974 to the USAEC Division of Biomedical and Environmental Research*. B.E. Vaughan (Ed.). Pacific Northwest National Laboratory, Richland, WA, pp. 37–40.
18. Meyer, R.E., Arnold, W.D., and Case, F.I. (1984). *Valence Effects on the Sorption of Nuclide on Rocks and Minerals*. Oak Ridge National Laboratory, Oak Ridge, TN.
19. Rard, J.A. (1983). *Critical Review of the Chemistry and Thermodynamics of Technetium and Some of its Inorganic*

Compounds and Aqueous Species. Lawrence Livermore National Laboratory, Livermore, CA.

20. Bidoglio, G., Avogadro, A., and De Plano, A. (1985). Influence of redox environments on the geochemical behavior of radionuclides. *Mater. Res. Symp. Proc.* **50**: 709–716.
21. Bidoglio, G., Chatt, A., De Plano, A., and Zorn, F. (1983). Investigation on technetium species in ground water and migration through clay columns. *J. Radioanalytical Chem.* **79**: 153–164.
22. Pilkington, N.J. (1990). The solubility of technetium in the near-field environment of a radioactive waste repository. *J. Less-Common Met.* **161**: 203–212.
23. Sheppard, S.C., Sheppard, M.I., and Evenden, W.G. (1990). A novel method used to examine variation in Tc sorption among 34 soils, aerated and anoxic. *J. Environ. Radioactivity* **11**: 215–233.
24. Henrot, J., Garten, C.T., Jr., and Blaylock, B.G. (1989). *Behavior of Technetium in Soil: Sorption-Desorption Processes*. Oak Ridge National Laboratory, Oak Ridge, TN.
25. Schulte, E.H. and Scoppa, P. (1987). Sources and behavior of technetium in the environment. *Sci. Total Environ.* **64**: 163–179.
26. Wolfrum, C. and Bunzl, K. (1986). Sorption and desorption of technetium by humic substances under oxic and anoxic conditions. *J. Radioanal. Nucl. Chem.* **99**: 315–323.
27. Wildung, R.E., Garland, T.R., McFadden, K.M., and Cowan, C.E. (1986). Technetium sorption in surface soils. In: *Technetium in the Environment*. G. Desmet and C. Myttenaere (Eds.). Elsevier Applied Science, London, pp. 115–129.
28. Wildung, R.E., Garland, T.R., and Cataldo, D.A. (1977). Accumulation of technetium by plants. *Health Phys.* **32**: 314–317.
29. Bondietti, E.A. and Francis, C.W. (1979). Geologic migration potentials of technetium-99 and neptunium-237. *Science* **203**: 1337–1340.
30. Brown, D.J. (1967). Migration Characteristics of Radionuclides Through Sediments Underlying the Hanford Reservation. In: *Disposal of Radioactive Wastes in to the Ground*. International Atomic Energy Agency Symposium. Vienna, Austria.
31. Gu, B., Dowlen, K.E., and Clausen, J.L. (1996). Efficient separation and recovery of technetium-99 from contaminated groundwater. *Efficient Sep.* **6**: 123–132.
32. Palmer, D.A. and Meyer, R.E. (1981). Adsorption of technetium on selected inorganic ion-exchange materials and on a range of naturally occurring minerals under oxic conditions. *J. Inorg. Nucl. Chem.* **43**: 2979–2984.
33. Strickert, R.G., Friedman, A.M., and Fried, S. (1980). The sorption of technetium and iodine radioisotopes by various minerals. *Nucl. Technol.* **49**: 253–266.
34. Davenport, D. (1994). *Interoffice Memorandum to L. Stinton*. Martine Marietta Energy Systems, Piketon, OH.
35. Erhardt, K.C. and Attrep, M., Jr. (1978). Technetium-99 in the atmosphere. *Environ. Sci. Technol.* **12**: 55–57.
36. Attrep, M., Jr., Enochs, J.A., and Broz, L.D. (1971). Atmospheric technetium-99. *Environ. Sci. Technol.* **5**: 344–345.
37. Leon, M. (1989). Determination and levels of ^{99}Tc in environmental and biological samples. In: *Technetium in the Environment*. G. Desmet and C. Myttenaere (Eds.). Elsevier Applied Science, London, pp. 171–178.
38. Clausen, J.L., Zutman, J.L., Pickering, D.A., and Farrow, N.D. (1995). *Final Report on Drive-Point Profiling of the Northwest Plume and Analysis of Related Data*. Lockheed Martin Energy Systems, Inc., Kevil, KY.
39. Clausen, J.L., Richards, W.L., Korte, N.E., and Liang, L. (1995). ORNL/MMES research into remedial applications of zero-valence metals, 3: Removal of TCE, cis-1,2-DCE, vinyl chloride and technetium. *Am. Chem. Soc. Symp. Reductive Dehalogenation Proc.*, April 2–7, Anaheim, CA, pp. 755–758.
40. Schroeder, N.S., Morgan, D., Rokop, D.J., and Fabryka-Martin, J. (1992). Migration of Technetium-99 in the Alluvial Aquifer at the Nevada Test Site. Los Alamos National Laboratory, LA-UR-92-463, Los Alamos, NM.
41. Brown, J., Kolstad, A.K., Lind, B., Rudjord, A.L., and Strand, P. (1998). *Technetium-99 contamination in the North Sea and in Norwegian Coastal areas 1996 and 1997*. Stralevern Report. 1998:3. Oslo, Norway.
42. Turcotte, M.S. (1982). *Environmental Behavior of Technetium-99*. Savannah River Plant and Laboratory Report DP-1644. Aiken, SC.
43. Anderson, T.J. (1979). Development of chemical isolation and concentration techniques for Tc-99 analysis by resin—bead mass spectrometry. In: *Radioelement Analysis*. Savannah River Laboratory, Aiken, SC.
44. Olsen, C.R. et al. (1986). Geochemical and environmental processes affecting radionuclide migration from a formerly used seepage trench. *Geochim. Cosmochim. Acta* **50**: 593–607.
45. Blaylock, B.G. and Frank, M.L. (1981). Bioaccumulation and distribution of Tc-95m in an experimental freshwater pond. In: *Int. Symp. Migration Terrestrial Environ. Long-Lived Radionuclides Nucl. Fuel Cycle.* Savannah River Laboratory, Aiken, SC.
46. Ballestra, S. et al. (1987). *J. Radiochem.* **115**: 51.
47. Fryar, A. and Lasage, D.M. (2001). Technetium-99 as an in-situ tracer of ground-water/stream interactions and organic co-contaminant fate. *Geol. Soc. Am. Annu. Meet.*, November 5–8, Boston, MA.
48. Gu, B., Clausen, J.L., McDonald, J.A., and McCarthy, J.F. (1994). *Assessment of the Influences of Groundwater Colloids on the Migration of ^{99}Tc at the PGDP Site, Paducah, KY*. Oak Ridge National Laboratory, Oak Ridge, TN.
49. Muck, M.T., Korte, N.E., Liang, L., Marsh, D., Zutman, J.L., and Clausen, J.L. (1995). *Technology Screening Report for Tc Removal from Groundwater Portsmouth Uranium Enrichment Plant, Piketon, Ohio*. Martin Marietta Energy Systems, Inc., Piketon, OH.
50. Buddemeier, R.W., Finkel, R.C., Marsh, K.V., Ruggieri, M.R., Rego, J.H., and Silva, R.J. (1992). Radionuclide migration at the Cheshire and Cambric event sites. In: *Hydrology and Radionuclide Migration Program 1989 Progress Report*. K.V. Marsh (Ed.). Lawrence Livermore National Laboratory, Albuquerque, NM, pp. 5–20.
51. Clausen, J.L., Sturchio, N.C., Heraty, L., Abrajano, T., and Huang, L. (1997). *Evaluation of Natural Attenuation Processes for Trichloroethene and Technetium-99 in the Northwest and Northeast Plumes at the Paducah Gaseous Diffusion Plant Paducah, Kentucky*. Lockheed Martin Energy Systems, Inc., Kevil, KY.
52. Clausen, J.L. and Richards, W.L. (1994). *Treatability Studies Using Iron Filings to Remediate Trichloroethene and Technetium*. Martin Marietta Energy Systems. KY/ER-51. Paducah, KY.
53. Bostick, W.D., Osberne, P.E., Del Cul, G.D., and Simmons, D.W. (1995). *Treatment of Aqueous Solutions Contaminated with Technetium-99 Originating from Uranium Enrichment Activities: Final Report*. Martin Marietta Energy Systems, Inc., Oak Ridge, TN.

54. Ishanullah. (1994). Optimization of various factors for the separation of technetium using anion-exchange resins. *Sep. Sci. Technol.* **29**: 239–247.
55. Del Cul, G.D., Bostick, W.D., Trotter, D.R., and Osberne, P.E. (1993). Technetium-99 removal from process solutions and contaminated groundwater. *Sep. Sci. Technol.* **28**: 551–564.
56. Kawasaki, M., Amori, T., and Hasegawa, K. (1993). Adsorption of pertechnetate on an anion exchange resin. *Radiochimica Acta* **63**: 53–56.
57. Takeuchi, M., Tanaka, S., and Yamawaki, M. (1993). Solvent extraction of Tc(vii) and Re(vii) from nitric acid solution with the mixture of CMPO and TBP. *Radiochimica Acta* **63**: 97–100.
58. Yamagishi, I. and Kubota, M. (1993). Recovery of technetium with active carbon column in partitioning process of high-level liquid waste. *J. Nucl. Sci. Technol.* **30**: 717–719.
59. Karalova, Z.K., Lavrinovich, E.A., and Myasoedov, B.F. (1991). Separation of technetium from accompanying elements in basic solutions by extraction with triphenyltetrazolium chloride. *Sov. Radiochem.* **33**: 266–269.
60. Chen, Q.J., Dahlgaard, H., Hansen, H.J.M., and Aarkrog, A. (1990). Determination of Tc-99 in environmental samples by anion exchange and liquid extraction at controlled valency. *Anal. Chim. Acta* **228**: 163–167.
61. Emery, M.F. and Lim, C.K. (1989). Separation of cationic technetium-99m amine complexes on porous graphiti carbon. *J. Chromatogr.* **479**: 212–215.
62. Bostick, W.D. and Evans-Brown, B.S. (1988). *Sorptive Removal of Technetium from Heavy Metals Sludge Filtrate Containing Nitrate Ion*. Oak Ridge National Laboratory, Oak Ridge, TN.
63. Yamagishi, I. and Kubota, M. (1989). Separation of technetium with active carbon. *J. Nucl. Sci. Technol.* **26**: 1038–1044.
64. Walton, F.B., Paquette, J., Ross, J.P.M., and Lawrence, W.E. (1986). Tc(IV) and Tc(VII) interactions with oxyhydroxides. *Nucl. Chem. Waste Manage.* **6**: 121–126.
65. Fried, S., Friedman, A.M., Cohen, D., Hines, J.J., and Strickert, R.G. (1978). *The Migration of Radioactive Processing Wastes in Selected Rocks*. Argonne National Laboratory, Argonne, IL.

WATER—NATURE'S MAGICIAN

Environment Canada

Water is the solvent, the medium and the participant in most of the chemical reactions occurring in our environment.

Water is at once simple and complex. A water molecule itself is simple, made up of three atoms: two hydrogen and one oxygen, H_20. The configuration of these building blocks produces a molecule with almost magical properties.

WATER—VISIBLE AND INVISIBLE

On earth, water is found as a liquid, as a solid (ice) or as a gas (water vapour).

- Canada has about 7% of the world's renewable freshwater supply, compared with 18% for Brazil, 9% for China, and 8% for the United States.
- The ice we skate on in winter is water in its *solid* form. Unlike most substances, which are densest in their solid state, ice is less dense than water and thus floats. If this were not the case, lakes and rivers would freeze from the bottom up. Fish could not survive, and it is unlikely that rivers and lakes in northern countries would ever completely thaw.
- Water vapour forms a kind of global "blanket" which helps to keep the earth warm. Heat radiated from the sun-warmed surface of the earth is absorbed and held by the vapour.

WATER'S MAGICAL PROPERTIES

Water molecules are attracted to each other, creating *hydrogen bonds*. These strong bonds determine almost every physical property of water and many of its chemical properties too.

Boiling and Freezing

Pure water at sea level boils at 100 °C and freezes at 0 °C. At higher elevations (lower atmospheric pressures) water's boiling temperature decreases. This is why it takes longer to boil an egg at higher altitudes. The temperature does not get high enough to cook the egg properly. If a substance is dissolved in water, then the freezing point is lowered. That is why we spread salt on streets in winter to prevent ice formation.

Thermal Properties

Water absorbs or releases more heat than many substances for each degree of temperature increase or

---

This article is produced by Environment Canada and is reproduced with permission.

decrease. Because of this, it is widely used for cooling and for transferring heat in thermal and chemical processes.

Differences in temperature between lakes and rivers and the surrounding air may have a variety of effects. For example, local fog or mist is likely to occur if a lake cools the surrounding air enough to cause saturation; consequently small water droplets are suspended in the air

Large bodies of water, such as the oceans or the Great Lakes, have a profound influence on climate. They are the world's great heat reservoirs and heat exchangers and the source of much of the moisture that falls as rain and snow over adjacent land masses.

When water is colder than the air, precipitation is curbed, winds are reduced, and fog banks are formed.

Surface Tension

Surface tension is a measure of the strength of the water's surface film. The attraction between the water molecules creates a strong film, which among other common liquids is only surpassed by that of mercury.

This surface tension permits water to hold up substances heavier and denser than itself. A steel needle carefully placed on the surface of a glass of water will float. Some aquatic insects such as the water strider rely on surface tension to walk on water.

Surface tension is essential for the transfer of energy from wind to water to create waves. Waves are necessary for rapid oxygen diffusion in lakes and seas.

Molecules in Motion

Water molecules as well as binding to each other, bind to many other substances such as glass, cotton, plant tissues, and soils. This is called adhesion. For example, in a thin glass tube, when the molecules at the edge reach for and adhere to the molecules of glass just above them, they at the same time tow other water molecules along with them. The water surface, in turn, pulls the entire body of water to a new level until the downward force of gravity is too great to be overcome. This process is called *capillary action*.

Thus water readily wets many materials. Capillary action allows a paper towel or a sponge to be used to soak up spilled water. Without this property, the nutrients needed by plants and trees would remain in the soil.

The Universal Solvent

An extraordinary property of water is its ability to dissolve other substances. There is hardly a substance known which has not been identified in solution in the earth's waters. Were it not for the solvent property of water, life could not exist because water transfers nutrients vital to life in animals and plants.

A drop of rain water falling through the air dissolves atmospheric gases. When rain reaches the earth, it affects the quality of the land, lakes and rivers.

Did You Know

- Raindrops are not tear-shaped. Scientists, using high-speed cameras, have discovered that raindrops resemble the shape of a small hamburger bun.
- About 70% of the human body is water.
- Life on earth probably originated in water.
- More than half of the world's animal and plant species live in the water.
- Almost 75% of the earth is covered in water.
- The human body needs 2 liters of water a day in our climate; we can last only a few days without water.
- Most of our food is water: tomatoes (95%), spinach (91%), milk (90%), apples (85%), potatoes (80%), beef (61%), hot dogs (56%).

THE SUN-POWERED CYCLE

The endless circulation of water from the atmosphere to the earth and its return to the atmosphere through condensation, precipitation, evaporation and transpiration is called the hydrologic cycle.

Heating of the ocean water by the sun is the key process that keeps the hydrologic cycle in motion. Water evaporates, then falls as precipitation in the form of rain, hail, snow, sleet, drizzle or fog. On its way to earth some precipitation may evaporate or, when it falls over land, be intercepted by vegetation before reaching the ground. The cycle continues in three different ways:

- *Evaporation / Transpiration*— On average, as much as 40% of precipitation in Canada is evaporated or transpired.
- *Percolation Into the Ground*— Water moves downward through cracks and pores in soil and rocks to the water table. Water can move back up by capillary action or it can move vertically or horizontally under the earth's surface until it re-enters a surface water system.
- *Surface Runoff*— Water runs overland into nearby streams and lakes; the steeper the land and the less porous the soil, the greater the runoff. Overland flow is particularly visible in urban areas. Rivers join each other and eventually form one major river that carries all of the subbasins' runoff into the ocean.

Although the hydrologic cycle balances what goes up with what comes down, one phase of the cycle is “frozen” in the colder regions during the winter season. During the Canadian winter, for example, most of the precipitation is simply stored as snow or ice on the ground. Later, during the spring melt, huge quantities of water are released quickly, which results in heavy spring runoff and flooding.

FRESHWATER SERIES A-1

Note: A resource guide, entitled Let's Not Take Water For Granted, is available to help classroom teachers of grades 5–7 use the information from the Water Fact Sheets.

FREEZING AND SUPERCOOLING OF WATER

Karl Erik Zachariassen
Norwegian University of
Science and Technology
Trondheim, Norway

Water in the crystalline solid state is known as ice, and for pure water at normal atmospheric pressure, ice is the physically stable form at temperatures below 0 °C.

Ice has a specific gravity of 0.92 g/ml, which is about 8% below the corresponding value of water, reflecting the fact that water expands by about 8% when it enters the solid state. The specific heat capacity of ice is 0.502 cal/g.deg, which is about half the corresponding value of water. The amount of heat liberated during the solidification (heat of fusion) is 79.71 cal/g.

The transition of water from the liquid to the solid state is known as freezing, and pure water has a so-called freezing point of 0 °C, which does not imply that pure water will freeze when cooled to this temperature. A small sample of pure water can be cooled to as much as −40 °C before it freezes spontaneously. For this reason, the term freezing point is misleading, and it is often replaced by the more precise terms *melting point* or *equilibrium freezing point*.

A substance that exists in the liquid state at a temperature below the equilibrium freezing point is said to be *supercooled* or *undercooled*, and the temperature where freezing is initiated is referred to as the *supercooling point* or the *temperature of crystallization*.

When a sample of supercooled water freezes, the released heat of fusion will immediately heat the system to the melting point. As a result of the high heat of fusion of water, the melting point temperature will be reached following the freezing of only a small fraction of the sample. From this point, the freezing can proceed only as fast as the released heat of fusion can be removed from the system. When pure water freezes, the temperature will remain at 0 °C until all water is frozen.

Initiation of freezing requires that the water sample contains an ice-like structure, which can act as a nucleus for ice formation, from which ice can grow. Liquid water is assumed to contain aggregates of water molecules organized in an ice-like pattern, but which usually are too unstable to act as nucleators. As the temperature drops, these aggregates become more stable and increase in size. On sufficient cooling, one of these aggregates will eventually reach the size critical for ice nucleation, and freezing will be initiated from this aggregate. When nucleation is initiated in this manner from the water molecules themselves, it is known as *homogeneous* nucleation.

Water samples often contain other substances that organize water molecules in an ice-like pattern and that cause freezing to be initiated at a higher temperature than would otherwise be the case. Such substances are known as *ice nucleators,* and freezing catalyzed by an ice nucleator is referred to as *heterogeneous freezing*. Ice nucleator activity is seen in various types of crystals and dust particles, and biological organisms may synthesize nucleating lipoproteins, which are physiologically important as ice nucleators, and which may have a high nucleating activity. Certain bacteria produce ice nucleators that can nucleate ice formations at above −1 °C.

FACTORS AFFECTING FREEZING

The equilibrium freezing point is influenced by factors such as solute concentration (Fig. 1) and hydrostatic pressure (Fig. 2). The osmolal melting point depression is 1.86 °C/Osm, whereas pressure depresses melting points by about 1 °C/100 bar.

The effect of these parameters on supercooling points depends on whether nucleation is homogeneous or heterogeneous. Homogeneous nucleation temperature is depressed 1.5–3 times more than melting points by increased osmolality and pressure, whereas, at least for the more active nucleators, nucleation temperature depression seems to be strictly equivalent to the corresponding melting point depression.

Nucleation temperature of purified water samples (no potent ice nucleators) also depends on sample volume as shown by Bigg in 1953 (1) (Fig. 3). When potent ice nucleators are present, the nucleation temperature depends only on the activity of the most active nucleator, and it does not matter if this nucleator occurs in a large or small water volume.

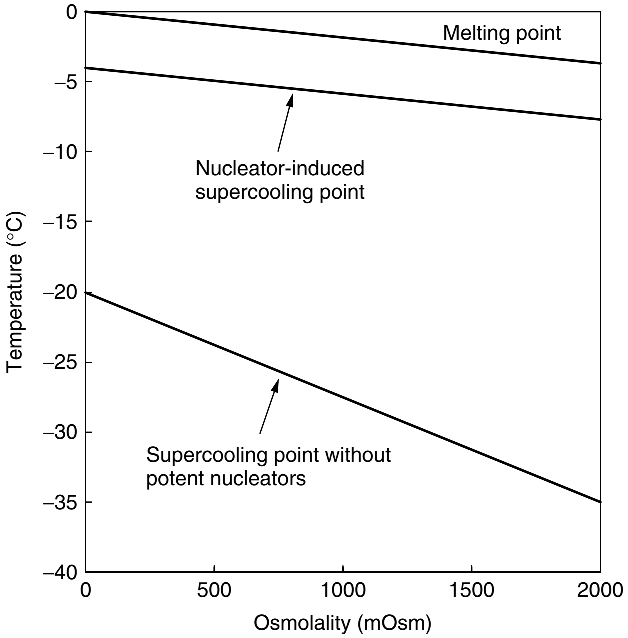

Figure 1.

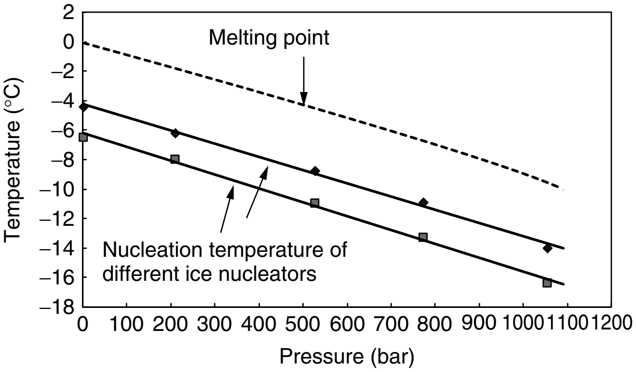

Figure 2.

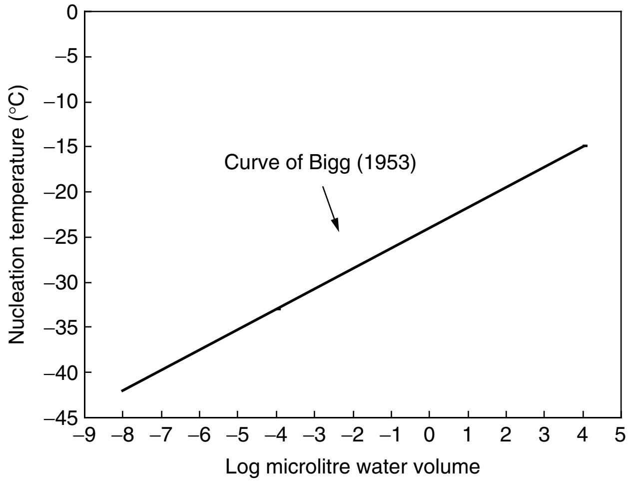

Figure 3.

FREEZING OF SOLUTIONS

When a solution freezes, the solutes are not incorporated in the ice structure, but remain in the liquid fraction surrounding the ice. Here, they become concentrated precisely until the solution reaches a melting point equal to the actual temperature. In this situation, vapor pressure equilibrium will exist between the solution and the ice, and no further freezing will occur until the released heat of fusion is removed from the system. As the heat is removed and the freezing proceeds, a gradual reduction of the temperature of the system will occur as the fluid fraction becomes more concentrated and the melting temperature of the fluid fraction drops. The solute concentration will increase until the eutectic point is reached. At the eutectic point, water and solutes freeze together.

BIBLIOGRAPHY

1. Bigg, E.K. (1953). The supercooling of water. *Proc. Roy. Soc. London. B* **66**: 688–694.

READING LIST

Zachariassen, K.E. and Kristiansen, E. (2000). Ice nucleation and antinucleation in nature. A review. *Cryobiology* **41**: 257–279.

CHEMICAL PRECIPITATION

Bradley A. Striebig
Gonzaga University
Spokane, Washington

Chemical precipitation is the removal of dissolved and suspended solids from aqueous solutions through chemical means. Chemical precipitation includes the removal of dissolved solids by altering the aqueous chemistry to make the dissolved solids less soluble. It also includes removing suspended solids by increasing their size in a process referred to as chemical flocculation. The larger the particle size of the suspended solids, the easier it is to remove them, usually through gravimetric settling in a process called clarification.

USES OF CHEMICAL PRECIPITATION

Chemical precipitation processes are utilized in water and wastewater treatment due to their effectiveness for pollutant removal with relatively low capital investment. The chemical reaction rates are typically very fast compared to biological reactions, so residence times in treatment processes or the reactor size may be significantly smaller than in other treatment techniques. Chemical precipitation processes can also process very toxic wastewaters. However, chemical precipitation processes often require the on-going addition of chemicals to control the pH and polymers to increase the size of the particles. The residuals collected from chemical precipitation processes may also be quite toxic and expensive to dispose of as sludge. Chemical precipitation is an effective tool for water treatment, and the most common uses are briefly described.

Industrial wastewaters often require pretreatment before being released to the sewer or receiving waters. High strength industrial wastewaters, such as those used in electroplating, contain high concentrations of dissolved metals. These metals can be removed from the wastewater by altering the pH of the water to reduce the solubility of the metal ions. The insoluble fractions of the metal ions form solid particles or precipitates, which are removed through filtration or gravimetric clarification. This allows industrial wastewaters to significantly reduce their metals concentrations in order to meet RCRA standards and also reduce the toxicity of the wastewater.

Municipal wastewaters may contain dissolved ions and a large amount of cell mass from biological treatment processes. The focus of chemical precipitation on municipal wastewaters is to remove the cell mass or suspended solids prior to and after biological treatment. The suspended solids are removed by adding chemical polymers, which attach to the existing suspended solids and help combine the individual particles into larger particles called flocs. This process, called coagulation, is accomplished by the chemical destabilization of particles and the subsequent aggregation. The process of gentle mixing and increasing the size of the suspended solids particles is called flocculation. The large floc particles are removed through gravimetric settling in a process called clarification. Thus a large fraction of the suspended solids and also the remaining BOD (in the form of cell mass) can be removed with the aid of chemical precipitation. Chemical precipitation is also commonly used when phosphorus removal is required at municipal wastewater treatment plants.

Industrial process water and recycled wastewater use chemical precipitation to remove hardness and dissolved salts from water. This allows a higher quality of water needed for specific industrial processes. The limits of hardness and salt removal are governed by the solubility of the dissolved ions in solution. Therefore, chemical precipitation has limited uses in industries that require highly purified water. However, chemical precipitation can be used as an inexpensive pretreatment step prior to ion exchange or reverse osmosis processes, which produce very high-quality water.

DISSOLVED SOLIDS

Dissolved solids are defined as the solid residuals in water that pass through a filter with a nominal pore size of 2.0 μm. These solids include dissolved ions and small colloidal solids. The dissolved ions in solution can be predicted based on the solubility product constant for the dissolved species. The solubility product constant, or solubility product, is an equilibrium constant that describes the reaction by which ions dissolve and solids precipitate:

$$\mathrm{A}_z\mathrm{B}_{y(s)} \rightleftharpoons z\mathrm{A}^{y+} + y\mathrm{B}^{z-}$$

For ions A and B, the generic equilibrium constant is

$$K_s = \frac{\{\mathrm{A}^{y+}\}^z\{\mathrm{B}^{z-}\}^y}{\{\mathrm{A}_z\mathrm{B}_y\}_{(s)}}$$

Many salts will form and precipitate as either hydroxides, with the hydroxide ion (OH^-), or carbonates, with the carbonate ion ($\mathrm{CO_3}^{2-}$). The concentration of these ions can often be predicted through an equilibrium analysis with the aid of electron or pH diagrams. Since the activity of the solid substance is defined as 1, and the solubility product is a constant value, the activity of the anion of interest (A) can be predicted in the aqueous solution of interest. Solubility constants for a wide array of dissolved ions of interest in water and wastewater are found in the referenced literature (1–3).

In a dilute solution under ideal conditions the activity of the ions of interest can be closely approximated by their concentration:

$$\{\mathrm{A}\} \approx [\mathrm{A}]$$

However, many wastewaters have high concentrations of dissolved ions and do not behave ideally. An activity coefficient, γ_A, must be used to establish an accurate relationship between the activity of a compound in solution and the concentration:

$$\{\mathrm{A}\} \approx \gamma_\mathrm{A}[\mathrm{A}]$$

The activity coefficient can be estimated from the ionic strength of the solution. The ionic strength is also commonly estimated from measurements of the total dissolved solids (TDS) and the electrical conductivity:

Langelier approximation	$I = 2.5 \times 10^{-5}$ (TDS)[mg/L]
Russell approximation	$I = 1.6 \times 10^{-5}$(conductivity) [μ mho]

One estimate of the activity coefficient is the Debye–Hückel limiting law for solutions with ionic strength less than 0.005 mol/L ($I < 0.005\,M$):

$$\gamma_\mathrm{A} \cong -Az^2 I^{1/2}$$

where A is a function of the dielectric constant of the solution. For water at 25 °C, $A = 0.51$. The term z is the ionic charge of the ion of interest. For more concentrated solutions ($I > 0.005\,M$), other estimates for the activity coefficient, such as the extended Debye–Hückel model and the Davies approximation can be used.

The potential to precipitate multiple ions from solution can be approximated from the activity of the ions in solution and the solubility product. Often, the pH of the solution is used as a "master variable." Solids lines on a graph of this type indicate the solids equilibrium concentration. For most hydroxide species, a minimum soluble concentration appears between a pH of 8 and 12. Similar graphical representations can be made for sulfide and carbonate species. This minimum solubility can serve as a starting point for estimating and optimizing chemical precipitation of various ions in solution.

The removal of hexavalent chromium (Cr^{6+}) from solution is one illustration of the potential of chemical precipitation processes to reduce the toxicity of wastewater. Hexavalent chromium can be converted to the less toxic trivalent form by reaction at low pH with a reducing agent such as ferrous bisulfite. Figure 1 shows a solution containing hexavalent chromium, which when dissolved in water results in a yellow color. Figure 2 shows the change in color associated with the change in the oxidation state of the chromium from the 6+ state to the 3+ state. When the pH of the solution is raised to 8 or 9, the trivalent chrome becomes much less soluble in solution than the hexavalent form of chromium as shown in Fig. 3. The resulting solids precipitate out of solution to form a concentrated sludge as shown in Fig. 4.

SUSPENDED SOLIDS

Suspended solids are composed of floating matter, settleable matter, and large colloidal matter. This matter

Figure 1. Hexavalent chromium dissolved in water.

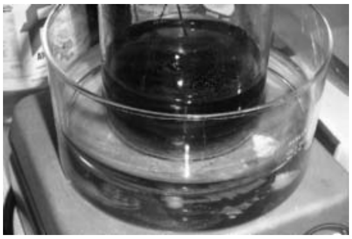

Figure 2. Reduction of hexavalent chromium to trivalent chromium.

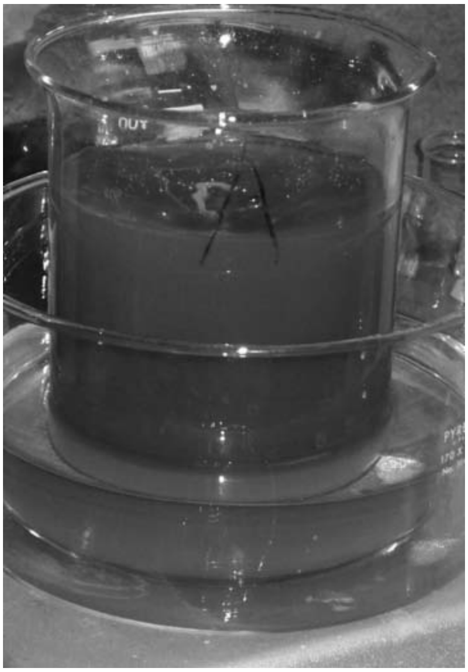

Figure 3. Formation of insoluble chromium oxide salts.

Figure 4. Precipitation of solid insoluble chromium into a sludge blanket.

may be organic or inorganic in nature. In municipal wastewater treatment plants, a significant fraction of the suspended solids consist of biological cell mass. The amount of suspended solids in water is an important parameter in determining water quality. Suspended solids are typically removed through a process of coagulation, flocculation, and gravimetric settling (or clarification). Chemical precipitation or coagulation of the suspended particles involves the destabilization of colloidal particles. Coagulation is usually followed by flocculation, which is the addition of a chemical, typically an organic polymer, to increase particle aggregation.

Suspended solids in water often have a charge that acts at the surface of the particle. Oil droplets, gas bubbles, and other inert substances may have a negative charge due to the preferential absorption of anions. Cellular materials such as microorganisms and proteins develop a charge through the ionization of carboxyl and amino acids functional groups. These charges create a force or zeta potential around the particles, which prevent them from increasing in size through aggregation as illustrated in Fig. 5. Coagulants are added to overcome these surface forces by the addition of potential determining ions or electrolytes. Once these charges have been neutralized, flocculants can be added to increase the speed of aggregation (Fig. 6).

Flocculant polymers that are anionic and nonionic become attached to the particles. These long particles increase the rate of aggregation and become intertwined

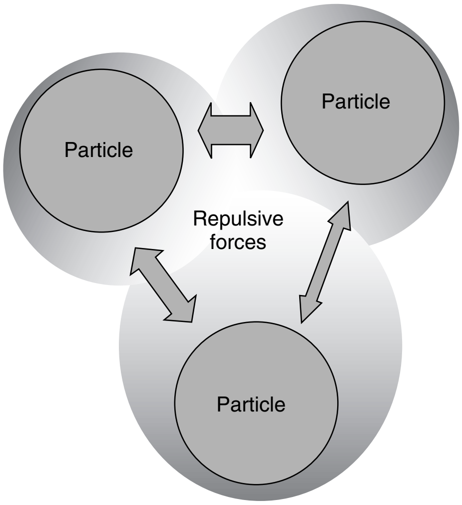

Figure 5. Repulsive forces acting on the surface of suspended particles.

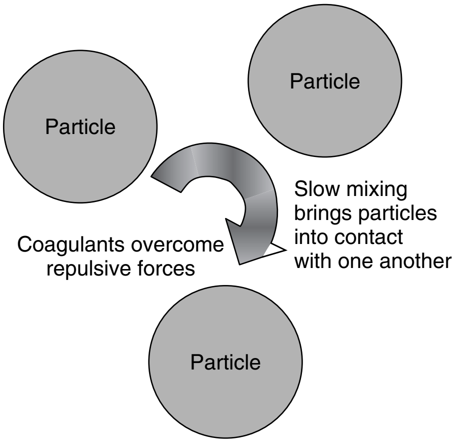

Figure 6. Chemical coagulants overcome repulsive forces and the solution is mixed.

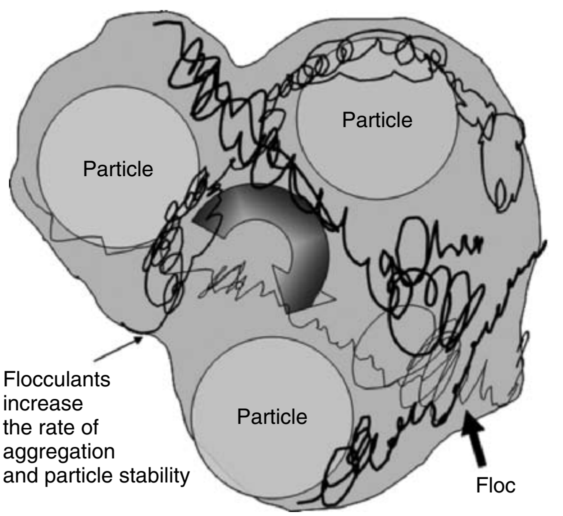

Figure 7. Flocculants increase the rate of aggregation to form large flocs.

during flocculation. This results in a significant increase in the particle size and mass (Fig. 7). The large particles called flocs are stabilized by the polymers, which prevent them from breaking up during the mixing process. The large flocs can then be removed with gravimetric settling in the clarification process. The resultant solids collected in the clarifier are the wastewater solids, or sludge.

Chemical precipitation by coagulation and flocculation has been used in the past to enhance removal of suspended solids and BOD. These processes are widely used to improve the performance of primary clarifiers, in basic physical–chemical processing of industrial wastewaters, for phosphorus removal, and for heavy metal removal.

BIBLIOGRAPHY

1. Tchobanoglous, G., Burton, F.L., and Stensel, H.D. (2003). *Wastewater Engineering Treatment and Reuse*, 4th Edn. Metcalf and Eddy, Inc/McGraw Hill, Boston, pp. 42–52, 475–517.
2. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry Chemical Equilibria and Rates in Natural Waters*, 3rd Edn. John Wiley & Sons, Hoboken, NJ, pp. 252–398.
3. Jensen, J.N. (2003). *A Problem-Solving Approach to Aquatic Chemistry*. John Wiley & Sons, Hoboken, NJ, pp. 418–450.

ANTIMONY IN AQUATIC SYSTEMS

Montserrat Filella
University of Geneva
Geneva, Switzerland

INTRODUCTION

Antimony is a naturally occurring element. It belongs to group 15 of the periodic table of the elements. Antimony can exist in a variety of oxidation states (−III, 0, III, V) but it is mainly found in two oxidation states (III and V) in environmental, biological, and geochemical samples. Although antimony was already known to the ancients, it is still often overlooked, both as an element of environmental concern and as a subject for study, probably because of its lower abundance and the relative insolubility of most of its compounds. This is reflected in the poor standard of existing data on the behavior of antimony in natural waters. However, interest in the study of this element seems to be growing and an increasing number of papers on the subject are being published. A useful series of comprehensive reviews on antimony in the environment has recently been published (1,2).

Antimony has no known biological function and, like arsenic, is toxic. Trivalent species are reported to be more toxic than pentavalent forms (3). The U.S. Environmental Protection Agency (EPA) (4) and the Commission of the European Communities (5) both consider antimony and its compounds to be pollutants of prime interest. The U.S. EPA drinking water standards are: MCLG (maximum contaminant level goal) and MCL (maximum contaminant level), both 6 μg/L (6). The European Union's maximum admissible concentration of antimony in drinking water is 5 μg/L (7). Antimony is on the list of hazardous substances under the Basel convention on the restriction of the transfer of hazardous waste across borders (8).

ABUNDANCE

The relative abundance of antimony in different systems is shown in Table 1. According to Goldschmidt's classification, antimony is a strong chalcophile element of which the most common natural form is the sulfide mineral stibnite, Sb_2S_3. Antimony is occasionally found as a native metal or as an intermediate compound with arsenic. More than 113 different sulfide, antimonide, arsenide, and telluride minerals have been identified. Sulfosalts represent the most common antimony-bearing minerals, usually in some combination with silver, copper, arsenic, mercury, and lead. More than 35 oxide, hydroxide, and oxyhalide mineral species are known. Antimony is also found in seven silicate minerals, all of which are rare (9). Antimony is also a common component of coal and petroleum.

Antimony is present in the aquatic environment as a result of rock weathering, soil runoff, and anthropogenic activities. Typically, concentrations of dissolved antimony in unpolluted waters are less than 1 μg/L. The mean antimony concentration in surface marine waters is 184 ± 45 ng/L (1). However, in the proximity of anthropogenic sources, concentrations can reach up to 100 times natural levels.

Airborne supply to aquatic and terrestrial systems is important for the environmental fate of antimony in systems located far from direct sources of pollution. Atmospheric emission values for antimony are shown in Table 2. The values in this table should be viewed with some caution because the data used comes from relatively old studies. For instance, antimony production was considered to be only 80,000 metric tons/yr (142,000 metric tons in 2003) and, more importantly, the uses of antimony have changed significantly over the past few years. Pacyna and Pacyna (15) estimated natural and anthropogenic emissions into the atmosphere to be 2400 tons/yr and

Table 1. Antimony Geochemical Abundance

Material	Sb Abundance[a]
Cosmic abundance	0.246[b]
Chondrites	0.1
Mean crustal average	0.2
Upper continental crust	0.31
Lower continental crust	0.30
Terrestrial abundance	0.7
Basic rocks (basalt)	0.15
Granitic rocks	0.2
Sedimentary rocks	
Shales	1–2
Carbonates	0.2
Deep sea clays	1

[a] All values in ppm (mg/g), except for the cosmic abundance.

[b] Atomic abundance relative to Si (=1,000,000) deduced from the following mean composition of meteorites: silicate = 100, sulfide = 7, metal = 10.6 parts per weight.

Source: Adapted from Reference 1.

1600 tons/yr, respectively. Very recently, Shotyk and co-workers (16) dismissed all previous values of natural antimony releases as being grossly overestimated. Based on antimony concentrations in ombrotrophic peat from Switzerland and Scotland, these authors suggested that 90–154 metric tons/yr would be more appropriate values.

USES

The first recorded use of antimony as a pigment to make mascara is found in an Egyptian papyrus dating from 1600 B.C. According to Emsley (17), the first uses of antimony date back even further; part of a 5000-year-old vase now in the Louvre is made of almost pure antimony. A peat core from an ombrotrophic Swiss bog revealed significant enrichments of antimony dating back to Roman times, which indicates that the anthropogenic fluxes of this element have exceeded natural fluxes for more than 2000 years. The present day enrichment factor (relative to the element/Sc ratios of typical crust rocks) is on the order of 70 times (as compared to 20 for As and 130 for Pb) (18). The chronology and intensity of the antimony enrichments in peat are similar in many ways to those of lead, which indicates that the extent of the human impact on the geochemical cycle of antimony is comparable to that of lead (16).

The production and use of antimony have steadily increased. World reserves of antimony are in excess of 2 million tons and are located principally in China, South Africa, Bolivia, Russia, and Tajikistan (19). The world's mines produced 142,000 tons of antimony in 2003 (19). Table 3 shows the main uses of primary antimony in the United States in 2003, which can be seen as representative of the use of antimony across the developed world. Uses of antimony have changed over the years. Traditionally, it was used in lead–antimony alloys. Bulk-secondary antimony could be recovered as antimonial lead, most of which was regenerated and then consumed by the battery industry. Today, the main use of antimony is as Sb_2O_3 in flame-retardants, which have applications in, among others, children's clothing and toys and aircraft and automobile seat covers. Antimony trioxide cannot be recycled and is released into the environment.

Table 2. Global Emissions[a] of Antimony into the Atmosphere, 10^3 metric tons/yr

Natural Sources	Sb Emission	Anthropogenic Sources	Sb Emission
Windborne dust (includes industrial sources of dust)	0.8	Energy production (coal, oil, gas)	1.3
Sea salt spray	0.6	Mining	0.1
Volcanic activity	0.7	Smelting and refining	1.4
Forest fires	0.2	Waste incineration	0.7
Biogenic sources	0.3		
Total	*2.6*	*Total*	*3.5*

[a] Adapted from Reference 1. Original estimations come from References 10–13. A completely different estimation, which is often cited in the literature, was published by Lantzy and Mackenzie in 1979 (14). They estimated natural and anthropogenic emissions of antimony into the atmosphere to be 0.98×10^3 metric tons/yr and 38×10^3 metric tons/yr, respectively.

Table 3. Reported Industrial Consumption (20) of Primary Antimony in the United States in 2002 (Metric Tons of Antimony Content)[a]

Uses	Sb Amount[b]
Metal products	
Antimonial lead	910
Bearing metal and bearings	43
Solder	85
Ammunition, cable covering, castings, sheet and pipe, and type metal	1,370
Total	*2,410*
Nonmetal products	
Ammunition primers	W
Ceramics and glass	487
Pigments	597
Plastics	532
Fireworks and rubber products	488
Total	*2,100*
Flameretardants	
Adhesives	W
Plastics	3,680
Rubber	126
Textiles	150
Paper and pigments	767
Total	*4,720*
Grand Total	*9,230*

[a] Secondary antimonial lead was 5600 metric tons in 2003.

[b] W = Withheld to avoid disclosing company proprietary data.

ANTIMONY SOLUTION CHEMISTRY

Both Sb(III) and Sb(V) ions hydrolyze easily in aqueous solution, thus making it difficult to keep antimony ions stable in solution except in highly acidic media. Since most of the techniques used for the study of solution equilibria (solubility measurements, potentiometry, voltammetry, spectroscopy) necessitate the use of macroamounts of reactants, relatively little information is available on the chemical species of antimony that exist in aqueous solution, and on their stability constants in particular. Moreover, most of the published studies are relatively old. Only the system antimony–sulfur has been the object of recent studies, mainly in relation to the understanding of hydrothermal systems.

The following main points can be made regarding the antimony inorganic chemistry in solution (2): (1) According to thermodynamic data, antimony should exist as Sb(V) in oxic systems and as Sb(III) in anoxic ones. (2) At environmental concentration levels, at the pH range commonly found in natural waters and in the absence of sulfur, antimony exists as a soluble species regardless of its oxidation state. Antimony(V) is present as a negative species in solution [$Sb(OH)_6^-$, also represented as SbO_3^-] and Sb(III) as a neutral species [$Sb(OH)_3$, also often quoted as $HSbO_2$]. (3) It is not clear whether Sb(V) and Sb(III) form any chloride-containing species in seawater, but hydroxychloro species would be the most probable entities. (4) Antimony oxides exhibit low water solubility, which may affect the fate of antimony in soils and water, particularly since today antimony is mostly used as Sb_2O_3(s) in flame-retardant applications. Moreover, it is thought that, when released into the atmosphere as an aerosol, antimony is oxidized to antimony trioxide through reactions with atmospheric oxidants. (5) In anoxic systems, and in the presence of sulfur, antimony forms insoluble stibnite, Sb_2S_3(s), and soluble SbS_2^{2-}, depending on the pH. The formation of polymerized species at natural concentration levels (e.g., $Sb_2S_4^{2-}$, $Sb_2S_5^{4-}$, $Sb_4Sb_7^{2-}$) is not proved. The formation of an Sb(V)-containing species (SbS_4^{3-}) has been observed by EXAFS.

Eh–pH diagrams can illustrate the redox behavior of elements like antimony. Figure 1 shows one such diagram for the system Sb–S–H_2O at environmentally realistic concentrations of antimony and dissolved sulfur.

A comprehensive, thermodynamically consistent antimony database of equilibrium constants of antimony with inorganic ligands has recently been published (21).

Barely any data have been reported on Sb(V) binding by low molecular weight organic ligands. More extensive studies have been carried out on the binding of Sb(III), which behaves like a borderline metal. The extent of antimony binding by natural organic matter (NOM) is a point of some debate. A few studies report significant Sb–NOM interactions, but most do not, and observed environmental behavior seems to suggest that Sb–NOM interactions are not relevant in most aquatic systems.

Hydrous oxides of manganese and iron have been reported to significantly sorb Sb(III) but very few systematic studies have been carried out on antimony sorption by pure metal oxides. These oxides will also oxidize the trivalent antimony to pentavalent antimony. Several studies have found a correlation between the adsorption of antimony in soil and oxic sediments and the presence of manganese and iron oxides (22,23).

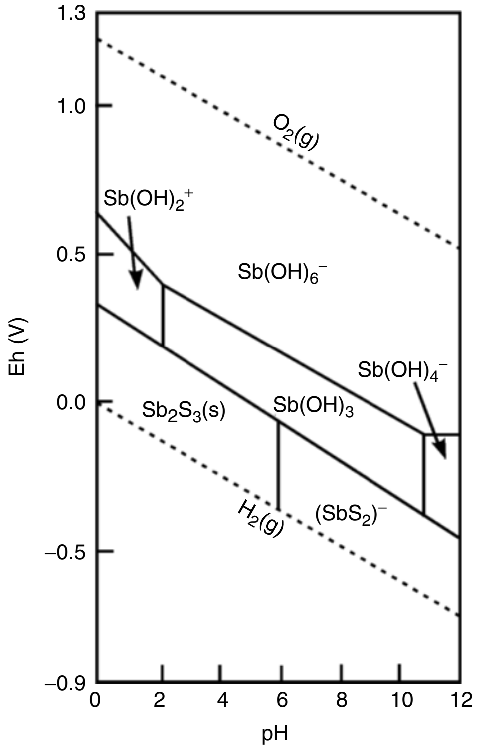

Figure 1. Eh–pH diagram of antimony in the Sb–S–H_2O system at a dissolved antimony concentration of 10^{-8} mol/L and a dissolved sulfur concentration of 10^{-3} mol/L. Adapted from Reference 2.

ANALYTICAL METHODS

The total concentrations of antimony in aqueous and solid environmental samples can readily be determined using current analytical techniques, primarily the result of the development of microwave digestion techniques for solid samples and NAA and plasma-based analyte detection systems. Both ICP-AES and ICP-MS techniques are in widespread use, although ICP-MS has less spectral interference and a lower detection limit (ng/L). An excellent critical review of the methodologies currently available for the environmental analysis of antimony has recently been published (24). However, determining antimony speciation still presents a significant analytical challenge (25). Recent data are almost exclusively based on hydride generation methods and often only oxidation state data are obtained. A lack of stability in the antimony species being investigated, particularly in trivalent forms, throughout the entire analytical process may result in an overestimation of pentavalent species concentrations.

Determining methylated antimony species is hampered by severe molecular rearrangements originating from the demethylation of trimethylantimony during the analysis. This problem has been the subject of extensive literary discussion but a solution has yet to be found (26).

ANTIMONY CONCENTRATIONS IN THE DIFFERENT AQUATIC COMPARTMENTS

The distribution and behavior of antimony in freshwater systems have not been studied extensively. Published concentrations range from a few nanograms/liter to a few micrograms/liter depending on location. They reflect the wide range of physical and chemical conditions existing in freshwater bodies and are very sensitive to the proximity of sources of pollution.

The concentration of antimony in ocean surface waters is about 200 ng/L. The behavior of antimony in the open seas is not considered to be highly reactive. According to some authors, it behaves as a conservative element and to others as a mildly scavenged element with surface (atmospheric) input. The comparison between antimony contents in deep waters of different oceans indicates that no antimony accumulation is produced during deep-water oceanic circulation (27). Antimony concentrations seem to reflect regional geochemical characteristics in coastal environments. Antimony behavior in estuaries varies somewhat according to the characteristics of the estuary. In some cases, it behaves conservatively and in others it shows a midestuary maximum. Behavior has even been seen to vary in the same estuary from one year to another.

Antimony concentrations in sediments are on the order of a few micrograms/gram of sediment and rarely exceed 1 mg/g. Higher concentrations are directly related to anthropogenic sources, mainly proximity to smelting plants and outfalls of sewage and fertilizer facilities.

ANTIMONY SPECIATION IN NATURAL WATERS

Today it is widely accepted that an understanding of biogeochemical processes requires knowledge of the chemical forms, or species, that are present in the natural environment. Despite this, the speciation of antimony in natural aquatic systems has yet to form the subject of many in-depth studies.

Dissolved Versus Particulate Fractions

Weak interaction between anionic species and suspended particles negatively charged in natural waters is to be expected and it is generally accepted that antimony is present as "dissolved" in aquatic systems. Experimental data (2) exist to support this. Fractionation of the so-called dissolved phase into several classes by ultrafiltration showed that most of the dissolved antimony is present in the smallest molecular weight region (2,28).

Redox

Antimony occurs in two oxidation states in natural waters and, thus, its behavior and toxicity can be affected by changes in the redox status of the aquatic environment. Antimony speciation studies have focused mainly on the separation and identification of Sb(III) and Sb(V) species.

Antimony (V) is the predominant species found in oxygenated systems but thermodynamically unstable Sb(III) has also been detected in different oxic marine water, fresh waters, groundwaters, and rain waters. This contrasts with thermodynamic equilibrium predictions, which suggest that Sb(V) concentration should exceed Sb(III) by several orders of magnitude. For thermodynamically unstable species to be present in water, mechanisms for their production and slow rates of interconversion must exist. The presence of Sb(III) has often been attributed to phytoplankton activity. However, although possible, a detailed analysis of published data has so far failed to produce conclusive proof to validate this hypothesis. Recently, the photochemical reduction of Sb(V) to Sb(III) has been suggested as an alternative to the phytoplankton source.

The true speciation of antimony under anoxic conditions remains unclear. According to thermodynamic calculations, in the absence of oxygen, antimony should only be present in the trivalent form. However, the occurrence of pentavalent forms of antimony has been reported in different anoxic systems. A variety of mechanisms have been suggested to explain these results, including delivery of Sb(V) on sinking detritus from oxic waters, formation of thiocomplexes by the pentavalent element, and advection of surface waters containing high concentrations of antimonate. All these mechanisms must be accompanied by relatively slow rates of reduction.

Methylated Species

Methylated species have been detected in surface seawaters, in estuaries, and in some rivers. They usually account for 10% or less of the total dissolved antimony. The monomethyl species is more abundant than the dimethyl form and close to the surface values tend to be higher. Monomethyl, dimethyl, trimethyl, and triethyl antimony derivatives have been detected in sediments from rivers and harbors. However, the identification of methylated antimony species has been questioned because some of these compounds undergo rearrangement reactions when subjected to the hydride generation procedures typically used for aquatic speciation.

The organoantimony compounds detected by Dodd et al. (26) in a freshwater plant (*Potamogetan pectinatus*) and a stibnolipid isolated from a marine diatom by Benson (29) are, as yet, the only organoantimony compounds with unequivocal biological origin to be detected in aquatic environments.

KINETIC ISSUES

As mentioned, the biogeochemical cycle of antimony is more complicated than equilibrium thermodynamic calculations predict (1). Several authors have cited the kinetic stabilization of thermodynamically unstable species to explain their presence in natural systems. Several processes may influence oxidation and reduction rates, including chemical and photochemical transformations and biological mediated reactions. It is well known that, when a redox process is accompanied by hydrolysis reactions, as is likely to be the case for antimony, the overall process can be much slower than expected. At present, kinetic information for redox reactions of antimony in natural waters is limited and rate constants are, for the most part, unknown. Cutter estimated an overall Sb(III) oxidation rate of 0.008 d^{-1} in the 65–85 m layer of the Black Sea from depth profiles (30). However, antimony residence times in seawater are probably longer because this rate included all forms of removal; that is, Sb(III) oxidation and also Sb(III) scavenging by sedimenting particles. Moreover, it was calculated at the oxic–suboxic interface of the Black Sea, where the presence of manganese and iron oxides is likely to increase the Sb(III) oxidation rate (31). Indeed, Belzile and co-workers (32), in a study on the oxidation of Sb(III) in the presence of natural and synthetic iron and manganese oxyhydroxides, showed that Sb(III) can be rapidly oxidized to Sb(V) by both compounds. Iron oxides also proved to play an important role in antimony redox chemistry in lake sediments (23). However, iron and manganese oxides are not likely to be present in oxic waters in significant quantities and thus are unlikely to influence antimony chemistry in oxic systems. It has recently been shown that Sb(III) oxidation by hydrogen peroxide and iodate is strongly pH dependent: the neutral $Sb(OH)_3$ species is unreactive to hydrogen peroxide and the presence of $Sb(OH)_4^-$ is needed for the oxidation to occur (33,34).

The addition of organic substances such as tartaric, lactic, citric, and ascorbic acids to natural or synthetic solutions has been shown to have a stabilizing effect on Sb(III). Solutions of Sb(III) prepared from potassium antimonyl tartrate have been reported to remain stable for long periods. The organic matter present in natural waters probably has a similar stabilizing effect. Ligand stabilization has been used to explain the apparent presence of Fe(II) in oxic solutions that contain relatively high levels of organic matter.

Little information is available on Sb(V) reduction kinetics in anoxic media. The reduction rate for Sb(V) in water at the bottom of the Black Sea was estimated to be 1.1×10^{-6} d^{-1} (35).

ANTIMONY AND AQUATIC BIOTA

Reported concentrations for antimony in fresh water and marine algae range from 0.02 to 1 μg/g dry weight (1). Although algae bioaccumulation and detoxifying mechanisms may play a key role in antimony redox speciation in natural waters, few studies have been carried out. Kantin (36) collected three marine algae near San Diego Bay in the United States and Sb(V) was the dominant species in all of them. Only *Sargassum* sp. contained up to 30% of Sb(III), thus demonstrating an ability to form reduced antimony species. Andreae and Froelich (37) analyzed the chemical speciation of antimony in phytoplankton (largely diatoms) from the eastern North Pacific and found that the trivalent species formed a significant part of the inorganic antimony. Both studies have since been used to justify the frequently cited biological origin of the Sb(III) found in oxic waters. Although no detailed studies exist, antimonate is expected to behave differently in bioavailability than arsenate and phosphate because of the weaker Lewis acidity and larger ionic radius of the antimonate ion in solution. More recently, the algae *Chlorella vulgaris*, isolated from an arsenic-polluted environment, excreted 40% Sb(V) and 60% Sb(III) on exposure to Sb(III), suggesting that a change in oxidation state is used by this alga as a detoxifying mechanism (38).

BIBLIOGRAPHY

1. Filella, M., Belzile, N., and Chen, Y-W. (2002). Antimony in the environment: a review focused on natural waters. I. Occurrence. *Earth Sci. Rev.* **57**: 125–168.
2. Filella, M., Belzile, N., and Chen, Y-W. (2002). Antimony in the environment: a review focused on natural waters. II. Relevant solution chemistry. *Earth Sci. Rev.* **59**: 265–285.
3. Bencze, K. (1994). Antimony. In: *Handbook on Metals in Clinical and Analytical Chemistry*. H.G. Seiler, A. Sigel, and H. Sigel (Eds.). Marcel Dekker, New York, pp. 227–236.
4. United States Environmental Protection Agency. (1979). *Water Related Fate of the 129 Priority Pollutants*. Vol. 1, EP-440/4-79-029A, U.S. EPA, Washington, DC.
5. Council of the European Communities. (1976). Council Directive 76/464/EEC of 4 May 1976 on pollution caused by certain dangerous substances discharged into the aquatic environment of the Community. Official Journal L 129, 18/05/1976, pp. 23–29.
6. United States Environmental Protection Agency. (1999). *National Primary Drinking Water Standards*. Doc. 810-F-94-001, U.S. EPA, Washington, DC.
7. Council of the European Union. (1998). Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. Official Journal L 330, 05/12/1998, pp. 32–54.
8. United Nations Environmental Program. (1999). *Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and Their Disposal* (with amended Annex I and two additional Annexes VIII and IX, adopted at the fourth meeting of the Conference of the Parties in 1998), SBC No. 99/001. UNEP, Geneva.
9. Lueth, V.W. (1999). Antimony: element and geochemistry. In: *Encyclopedia of Geochemistry*. C.P. Marshall and R.W. Fairbridge (Eds.). Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 15–16.
10. Nriagu, J.O. and Pacyna, J.M. (1988). Quantitative assessment of worldwide contamination of air, water and soils by trace metals. *Nature* **333**: 134–139.
11. Nriagu, J.O. (1989). A global assessment of natural sources of atmospheric trace metals. *Nature* **338**: 47–49.
12. Nriagu, J.O. (1990). Global metal pollution. Poisoning the biosphere? *Environment* **32**: 7–11, 28–33.
13. Maeda, S. (1994). Safety and environmental effects. In: *The Chemistry of Organic Arsenic, Antimony and Bismuth Compounds*. S. Patai (Ed.). Wiley, Hoboken, NJ, pp. 725–759.
14. Lantzy, R.J. and Mackenzie, F.T. (1979). Atmospheric trace metals: global cycles and assessment of man's impact. *Geochim. Cosmochim. Acta* **43**: 511–525.
15. Pacyna, J.M. and Pacyna, E.G. (2001). An assessment of global and regional emissions of trace metals to the atmosphere from anthropogenic sources worldwide. *Environ. Rev.* **9**: 169–298.
16. Shotyk, W. et al. (1996). Two thousand years of atmospheric arsenic, antimony, and lead deposition recorded in an ombrotrophic peat bog profile, Jura Mountains, Switzerland. *Earth Planet. Sci. Lett.* **145**: E1–E7.
17. Emsley, J. (2001). *Nature's Building Blocks. An A–Z Guide to the Elements*. Oxford University Press, New York.

18. Shotyk, W., Krachler, M., and Chen, B. (2004). Antimony in recent, ombrotrophic peat from Switzerland and Scotland: comparison with natural background values (5,320 to 8,020^{14}C yr BP) and implications for the global Sb cycle. *Global Biogeochem. Cycles* **18**: GB1016.
19. Carlin, J.F., Jr. (2004). Antimony. In: *U.S. Geological Survey Mineral Commodity Summaries*.
20. Carlin, J.F. (2003). *U.S. Geological Survey Minerals Yearbook*.
21. Filella, M. and May, P.M. (2003). Computer simulation of the low-molecular-weight inorganic species distribution of antimony(III) and antimony(V) in natural waters. *Geochim. Cosmochim. Acta* **67**: 4013–4031.
22. Müller, B. et al. (2002). As, Sb, Mo, and other elements in sedimentary Fe/Mn layers of Lake Baikal. *Environ. Sci. Technol.* **36**: 411–420.
23. Chen, Y-W., Deng, T-L., Filella, M., and Belzile, N. (2003). Distribution and early diagenesis of antimony species in sediments and porewaters of freshwater lakes. *Environ. Sci. Technol.* **37**: 1163–1168.
24. Nash, M.J., Maskall, J.E., and Hill, S.J. (2000). Methodologies for determination of antimony in terrestrial environmental samples. *J. Environ. Monit.* **2**: 97–109.
25. Krachler, M., Emons, H., and Zheng, J. (2001). Speciation of antimony for the 21st century: promises and pitfalls. *Trends Anal. Chem.* **20**: 79–90.
26. Dodd, X. et al. (1996). Antimony speciation in freshwater plant extracts by using hydride generation–gas chromatography–mass spectrometry. *Analyst.* **121**: 223–228.
27. Cutter, G.A. and Cutter, L.S. (1998). Metalloids in the high latitude North Atlantic Ocean: sources and internal cycling. *Mar. Chem.* **61**: 25–36.
28. Stolpe, X., Hassellöv, M., Andersson, K., and Turner, D. (2004). High resolution ICPMS as an on-line detector for flow-field-flow fractionation; multi-element determination of colloidal size distributions in a natural water sample. *Anal. Chim. Acta.* in press.
29. Benson, A. (1988). Antimony metabolites in marine algae. In: *The Biological Alkylation of Heavy Metals*. P.J. Craig and F. Glocking (Eds.). *Special Publication*, Vol. 66, Royal Society of Chemistry, London, pp. 135–137.
30. Cutter, G.A. (1992). Kinetic controls on metalloid speciation in seawater. *Mar. Chem.* **40**: 65–80.
31. Cutter, L.S., Cutter, G.A., and San Diego-McGlone, M.L.C. (1991). Simultaneous determination of inorganic arsenic and antimony species in natural waters using selective hydride generation with gas chromatography/photoionization detection. *Anal. Chem.* **63**: 1138–1142.
32. Belzile, N., Chen, Y-W., and Wang, Z. (2001). Oxidation of antimony(III) by amorphous iron and manganese oxyhydroxides. *Chem. Geol.* **174**: 379–387.
33. Quentel, F., Filella, M., Elleouet, C., and Madec, C-L. (2004). Kinetic studies on Sb(III) oxidation by hydrogen peroxide in natural waters. *Environ. Sci. Technol.* **38**: 2843–2848.
34. Quentel, F., Filella, M., Elleouet, C., and Madec, C-L. (2004). Sb(III) oxidation by iodate in seawater: a cautionary tale. (Submitted).
35. Cutter, G.A. (1991). Dissolved arsenic and antimony in the Black Sea. *Deep-Sea Res.* **38**: S825–S843.
36. Kantin, R. (1983). Chemical speciation of antimony in marine algae. *Limnol. Oceanogr.* **28**: 165–168.
37. Andreae, M.O. and Froelich, P.N. (1984). Arsenic, antimony, and germanium biogeochemistry in the Baltic Sea. *Tellus* **36**: 101–117.
38. Maeda, S. et al. (1997). Bioaccumulation of antimony by *Chlorella vulgaris* and the association mode of antimony in the cell. *Appl. Organomet. Chem.* **11**: 393–396.

WATER LAW AND ECONOMICS

THE CLEAN WATER ACT

Timothy J. Ryan
Ohio University
Athens, Ohio

U.S. water legislation dates back to the late nineteenth century, when the 49th Congress enacted the Rivers and Harbors Act of 1886, which ultimately became the Rivers and Harbors Act of 1899. This law was designed to minimize obstructions in any navigable waterway in the United States, and it prohibited the deposit of any material on banks of navigable waters that would lead to the impeding of navigation. The law basically allowed for fines of $500 to $2,500 as well as imprisonment. Recognizing the continuing public health threat posed by water contamination, in 1948 Congress enacted the Federal Water Pollution Control Act (FWPCA) in order to "enhance the quality and value of our water resources and to establish a national policy for the prevention, control and abatement of water pollution." The original FWPCA aimed to establish programs that would improve sanitary conditions of both surface and groundwater while minimizing or eliminating the pollution of interstate waterways and tributaries. These goals thus differed from the original Rivers and Harbors Act's goals by adding pollution regulation in hopes of improving the public's water supply, habitats of aquatic life, and water for recreational, agricultural, and industrial uses. FWPCA and several subsequent amendments set out the basic legal authority for federal regulation of water quality to the present day.

The original FWPCA, or simply the "Act," was subjected to incremental but limited improvements that eventually resulted in a somewhat disparate collection of water-related requirements. Eleven reorganizations and restructurings of federal agency responsibilities compounded the difficulty of effectively implementing the law. For this reason, the 1972 amendments to the FWPCA restructured the authority for water pollution control and consolidated authority in the Administrator of the then newly formed United States Environmental Protection Agency (EPA). Formed in 1970 by President Nixon in response to the growing public demand for an overall healthier environment, the EPA was the first federal authority to implement pollution control programs for industry. Growing public awareness and concern for controlling water pollution throughout the mid-1970s led to amendments of the Act in 1977. As amended, this law became commonly known as the Clean Water Act (CWA).

The CWA, and it's associated rules found in Title 33 of the Code of Federal Regulations, established the basic priorities and structure for regulating pollution releases into the waters of the United States. The Act's major subchapters include six Titles: (I) Research and Related Programs, (II) Grants for Construction of Treatment Works, (III) Standards and Enforcement, (IV) Permits and Licenses, (V) General Provisions, and (VI) State Water Pollution Control Revolving Funds. The Act continued requirements to set water quality standards for all contaminants in surface waters and made it unlawful to discharge any pollutant from a single, or point, source into navigable waters, unless a permit was obtained under the Act. It also funded the construction of sewage treatment plants under the construction grants program and recognized the need for planning to address the critical problems posed by diffuse, or nonpoint, sources of pollution.

NATIONAL POLLUTANT DISCHARGE ELIMINATION SYSTEM

One of the most significant features of the Clean Water Act is the creation of a National Pollutant Discharge Elimination System (NPDES) under Section 402 of Title IV, Permits and Licensing. The NPDES prohibited industrial sources and publicly owned treatment works from discharging pollutants into navigable waters without a process review and permit. Under NPDES, all point source discharges of pollution into the waters of the United States are required to obtain a permit. Point source discharges come from "any discernible, confined, and discrete conveyance including but not limited to any pipe, ditch, channel, tunnel, conduit, well, discrete fissure, container, rolling stock, concentrated animal feeding operation, landfill leachate collection system, vessel or other floating craft from which pollutants are or may be discharged." The EPA Administrator may issue a permit for discharge on the condition that the discharge meets applicable requirements, which are detailed in the Act. Permitted discharges are only allowed insofar as federal effluent limitations and state water quality standards can continue to be met. The Act also provides that, given EPA approval, a state may administer its own permit program in lieu of the federal program so long as the state effort is deemed at least as effective as the federal controls.

NONPOINT SOURCE POLLUTION CONTROL

Since its use in 1972, the term "nonpoint source pollution" has been used in a variety of sometimes confusing contexts. Most simply, a nonpoint source is any source of pollution that is not otherwise legally considered as a point source. In 1987, Congress enacted Section 319 of Title III, Standards and Enforcement, of the Clean Water Act to establish a national program to control nonpoint sources of pollution. In its guidance for the program, the EPA rigorously defines nonpoint source pollution as that "... caused by rainfall or snowmelt moving over and through the ground and carrying natural

Table 1. Toxic Pollutants Designated Pursuant to Section 307(a)(1) of the Clean Water Act

1. Acenaphthene
2. Acrolein
3. Acrylonitrile
4. Aldrin/Dieldrin
5. Antimony and compounds
6. Arsenic and compounds
7. Asbestos
8. Benzene
9. Benzidine
10. Beryllium and compounds
11. Cadmium and compounds
12. Carbon tetrachloride
13. Chlordane (technical mixture and metabolites)
14. Chlorinated benzenes (other than di-chlorobenzenes)
15. Chlorinated ethanes (including 1,2-di-chloroethane, 1,1,1 trichloroethane, and hexachloroethane)-
16. Chloroalkyl ethers (chloroethyl and mixed ethers)
17. Chlorinated naphthalene
18. Chlorinated phenols (other than those listed elsewhere; includes trichlorophenols and chlorinated cresols)
19. Chloroform
20. 2-chlorophenol
21. Chromium and compounds
22. Copper and compounds
23. Cyanides
24. DDT and metabolites
25. Dichlorobenzenes (1,2-, 1,3-, and 1,4-di-chlorobenzenes)
26. Dichlorobenzidine
27. Dichloroethylenes (1,1-, and 1,2-dichloroethylene)
28. 2,4-dichlorophenol
29. Dichloropropane and dichloropropene
30. 2,4-dimethylphenol
31. Dinitrotoluene
32. Diphenylhydrazine
33. Endosulfan and metabolites
34. Endrin and metabolites
35. Ethylbenzene
36. Fluoranthene
37. Haloethers (other than those listed elsewhere; includes chlorophenylphenyl ethers, bromophenylphenyl ether, bis(dichloroisopropyl) ether, bis-(chloroethoxy) methane and polychlorinated diphenyl ethers)
38. Halomethanes (other than those listed elsewhere; includes methylene chloride, methylchloride, methylbromide, bromoform, dichlorobromomethane
39. Heptachlor and metabolites
40. Hexachlorobutadiene
41. Hexachlorocyclohexane
42. Hexachlorocyclopentadiene
43. Isophorone
44. Lead and compounds
45. Mercury and compounds
46. Naphthalene
47. Nickel and compounds
48. Nitrobenzene
49. Nitrophenols (including 2,4-dinitrophenol, dinitrocresol)
50. Nitrosamines
51. Pentachlorophenol
52. Phenol
53. Phthalate esters
54. Polychlorinated biphenyls (PCBs)
55. Polynuclear aromatic hydrocarbons (including benzanthracenes, benzopyrenes, benzofluoranthene, chrysenes, dibenz-anthracenes, and indenopyrenes)
56. Selenium and compounds
57. Silver and compounds
58. 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD)
59. Tetrachloroethylene
60. Thallium and compounds
61. Toluene
62. Toxaphene
63. Trichloroethylene
64. Vinyl chloride
65. Zinc and compounds

and human-made pollutants into lakes, rivers, streams, wetlands, estuaries, other coastal waters, and ground water. Atmospheric deposition and hydrologic modification are also sources of nonpoint pollution." A major source of typically nonpoint source pollution was recognized in urban stormwater. In response, and to address the specific concerns of contaminated stormwater runoff, the National Stormwater Program was established as part of the 1987 Clean Water Act amendments. Although stormwater is considered nonpoint source pollution, the hazardous constituents in the water may come from various points sources such as industry, sewers, or construction. Thus, the National Stormwater Program is controlled through NPDES, previously discussed.

EFFLUENT LIMITATIONS AND WATER QUALITY STANDARDS

In Section 401 of Title IV, Permits and Licensing, the Act prohibits the discharge of pollutants in excess of effluent limitations, and makes it unlawful to discharge radiological, chemical, or biological warfare agents; high-level radioactive waste; or any medical waste into navigable waters. Effluent limitations from point sources other than publicly owned treatment works (POTWs) must be treated using best practicable control technology, or BACT. The effluent limitations provide standards for a variety of pollutants and pollution sources, including but not limited to sources such as coal mining, landfills, construction, iron and steel manufacturing, meat products, pharmaceutical, paper products, and leather and tanning industry. Each source type has a different set of EPA effluent limitation standards to meet. Toxic pollutants defined in the Act (Table 1) require treatment using the best available technology that is economically achievable. If the source discharges into a POTW (e.g., via a connection to a sanitary sewer), it must comply with applicable pretreatment requirements as well as POTW expectations.

WATERSHED RESTORATION PLANS

Section 303(d) of Title III, Standards and Enforcement, of the Act requires that states and the EPA identify rivers, lakes, and coastal waters threatened or already impaired as a result of one or more specific pollutants. The states or EPA must develop "Total Maximum Daily Loads" (TDML) for every water body that has been placed on the EPA-designated threatened or impaired waters list. A TDML is thus a pollution cap or ceiling for the water body, and it has the functional effect of a cleanup plan. An associated plan for the affected water body then determines exactly how to attain that cap. The TMDL must specify problems, determine pollution reductions needed to solve the problems, identify pollution sources, and assign responsibilities for needed action. Also under Title III, various other specific and significant sources of pollution are regulated. Thermal discharges, marine sanitation devices, and oil and hazardous substance discharge are a few aspects of water pollution so regulated. As all of these sources have the potential to be hazardous polluters, the Act sets standards to which each must abide in different manners to ensure the quality of the water.

PLANS, GRANTS, AND FUNDING

The CWA both requires and assists in the development and implementation of waste treatment management plans and practices to achieve the water quality goals of the Act. Plans and practices must provide for treatment of waste prior to any discharge of pollutants into receiving waters, as well as the secure disposal of other pollution sources to avoid migration that could ultimately cause water or other environmental pollution. To accomplish these goals, Title II of the Act outlines a program of grants to states, municipalities, or other agencies for the construction of governmental water treatment facilities, or POTWs. Revisions of the Act in 1981 streamlined the municipal construction grants process, improving the capabilities of treatment plants built under the program. Additional changes in 1987 phased out the construction grants program but replaced it with the State Water Pollution Control Revolving Fund (found in Title VI of the Act), known within professional circles as the Clean Water State Revolving Fund. That funding strategy addressed water quality needs by building on EPA-State partnerships to identify and meet state and regional water quality issues.

KEY AMENDMENTS AND ASSOCIATED WATER PROTECTION LAWS

Safe Drinking Water Act

Over the years, many other environmentally related laws have changed or otherwise impacted parts of the CWA. The Safe Drinking Water Act (SDWA) was originally passed by Congress in 1974 to specifically protect the nation's public drinking water supply. The SDWA authorized the EPA to set national health-based standards for drinking water to protect against both naturally occurring as well as man-made contaminants that may be found in drinking water. These standards are divided into primary and secondary standards. Primary standards are set to protect the public's health, are legally enforceable, and apply to public water systems. Secondary standards are nonenforceable guidelines set to regulate contaminants in the drinking water that may cause negative cosmetic or aesthetic effects, such as tooth discoloration or bad odor. States may choose to adopt these secondary standards as enforceable regulations. The law was amended in 1986 and 1996 and requires various actions to protect drinking water sources, including rivers, lakes, reservoirs, springs, and groundwater.

Oil Pollution Act

As a response to the Exxon Valdez disaster of 1989, when a crude oil tanker ran into a reef on the coast of Alaska and spilled 10.8 million gallons of the oil into the Prince William Sound, Congress passed the Oil Pollution Act in 1990. This act amended the CWA to significantly raise the limits of liability to would-be polluters, and the amount of compensation that must me made to injured parties resulting from an oil spill.

Great Lakes Critical Programs Act

As a result of their immense size and value as a freshwater supply, state, federal, and international steps have been taken to protect the water quality of the Great Lakes. Title I of the Great Lakes Critical Programs Act of 1990, for example, put into place parts of the U.S.-Canadian Great Lakes Water Quality Agreement of 1978, wherein the two nations agreed to jointly reduce certain toxic pollutants entering these resources. That law required the EPA to establish water quality criteria for the Great Lakes for 29 toxic pollutants with a particular focus on persistent bioaccumulative toxics (such agents tend to occur in higher concentration in aquatic biota than in open waters). The criteria are intended to protect human health, aquatic life, and wildlife and include the first-ever EPA wildlife criteria to protect birds and mammals from long-term exposure to mercury, DDT, PCBs, and dioxin. It also required the EPA to help the states implement the criteria on a specific schedule.

With a growing world population and intense industrial demands on a finite water supply, regulations concerning U.S. water resources will continue to change and evolve with time. As the most prominent water quality law presently in effect, the CWA will continue to play a significant roll in the maintenance of water quality for future generations in the face of technological changes and demands.

READING LIST

EPA Laws and Regulations—Clean Water Act History. Available: http://www.epa.gov/region5/water/cwa.htm.

EPA Water History Website. Available: http://www.epa.gov/history/topics/fwpca/05.htm.

Digest of Federal Resource Laws of Interest to the U.S. Fish and Wildlife Service. Federal Water Pollution Control Act (Clean Water Act). Available: http://laws.fws.gov/lawsdigest/fwatrpo.html.

River Network Website—Understanding the Clean Water Act. Available: http://cleanwateract.org/pages/c1.htm.

Knudson, S. Quick Reference: Clean Water Act (CWA). Center for Sustainable Systems, Ann Arbor, MI. Available: http://css.snre.umich.edu/css_doc/CWA.pdf.

Copeland, C. (2002, 2004). CRS Report for Congress—Clean Water Act: Summary of the Law.

Clean Water Act. Available: http://www.boozman.house.gov/UploadedFiles/H20%20%20Clean%20Water%20Act%20A%20 Summary%20 of%20 the%20Law.pdf.

Environmental Literacy Council—Clean Water Act. Available: http://www.enviroliteracy.org/article.php/372.html.

EPA Electronic "Snapshot" of the Clean Water Act (the Act in PDF format). Available: http://www.epa.gov/region5/water/cwa.htm#ECWA.

NOAA Legislative Summaries—Federal Water Pollution Control Act. Available: http://www.csc.noaa.gov/opis/html/summary/cwa.htm.

CLEAN WATER ACT, WATER QUALITY CRITERIA/STANDARDS, TMDLS, AND WEIGHT-OF-EVIDENCE APPROACH FOR REGULATING WATER QUALITY

G. Fred Lee
Anne Jones-Lee
G. Fred Lee & Associates
El Macero, California

INTRODUCTION

The 1972 amendments to the Federal Water Pollution Control Act (Clean Water Act, CWA) established a regulatory framework to maintain and enhance water quality in waterbodies of the United States. The CWA specified that waterbodies were to be classified with respect to their beneficial uses, such as domestic water supply, propagation of fish and aquatic life, and recreation, and that the U.S. EPA was to develop national water quality criteria for chemical and other agents that would protect the designated beneficial uses of waterbodies. The numeric chemical-concentration-based national water quality criteria that it eventually developed were designed to be protective in any water under worst-case exposure conditions. The U.S. EPA criteria, in turn, were to be used by the states as the basis for water quality standards, the benchmark for control of beneficial use-impairment of waterbodies.

The way in which these CWA provisions have been implemented has been to require that a waterbody in which a water quality standard was exceeded be listed as a Clean Water Act section 303(d)-"impaired" waterbody. Such a listing requires that a program be initiated to control the sources of the constituents responsible for the water quality standard violation, which is done through a total maximum daily load (TMDL) assessment through which is determined a load that would not result in exceeding the standard; that load is then apportioned among those discharging the constituent.

Although this TMDL approach was originally adopted in 1972, it was not until the late 1990s that it began to be implemented to any significant extent. This implementation developed from environmental groups filing suit against the U.S. EPA for having failed to implement the TMDL provisions of the Clean Water Act. As part of settling the lawsuit, the U.S. EPA agreed to see to the implementation of TMDLs by the states. Considerable controversy, exists however, about the appropriateness of the TMDL process as it is typically being implemented for regulating water quality caused largely by the use of numeric U.S. EPA national water quality criteria and state standards equivalent to them as the benchmark for evaluation of water quality. Following is a synopsis of how this approach came to be established and issues that remain inadequately addressed in the pursuit of the goals of the CWA. Also described is a "best professional judgment" triad weight-of-evidence approach that provides a technically valid, cost-effective framework for assessing water quality/beneficial-use impairments and their remediation.

EARLY REGULATORY APPROACH—ISSUES STILL NEEDING ATTENTION

Lee (1) discussed several significant technical shortcomings inherent in the TMDL approach adopted by the U.S. EPA for regulating chemical constituents for the protection of water quality. For example, in the 1960s, when toxicity tests were starting to evaluate the toxicity of wastewater discharges, it was often found that a portion of the heavy metals in some discharges was in forms that did not cause toxicity. This finding was in accord with what would be expected based on the aquatic chemistry of heavy metals. It was recognized that heavy metals exist in a variety of chemical forms; given the chemical nature of the forms, it could be expected that not all of them would be equally available/toxic to aquatic life, although this distinction could not be made with chemical analytical techniques. This situation was sufficiently well known that by the early 1970s, the National Academies of Science and Engineering concluded, as part of their development of the *Blue Book of Water Quality Criteria* (2), that heavy metals in wastewater discharges could not be reliably regulated based on chemical concentrations. Rather, because of the numerous unquantifiable factors that control the manifestation of a chemical's toxicity, they recommended a toxicity test approach to determine the availability of the heavy metals, either alone or in combination with other metals or other substances, in a particular water. The National Academies of Science and Engineering Blue Book Criteria were adopted by the U.S. EPA (3) in its 1976 *Red Book Criteria*, which were the first official water quality criteria developed pursuant to the Clean Water Act.

In the early 1980s, however, the U.S. EPA abandoned the toxicity testing approach recommended by the National Academies of Science and Engineering. In its place, it adopted a policy of applying the numeric, worst-case, national water quality criteria (generally, chronic exposure, safe concentrations) to the concentrations of total recoverable metals (i.e., those forms that are measurable after strong acid digestion of the sample), rather than to the available forms, in water quality evaluation and regulation. That approach lead to overregulation of heavy metals because in some situations, substantial parts of the "total recoverable" heavy metals are nontoxic/unavailable.

Some relief from the overregulation of heavy metals was provided when application of the criteria was shifted from "total recoverable" forms to "ambient-water dissolved" forms of metals (4). The shift in focus at that time was not based on the finding of any new information because it had been well-established in the 1960s and 1970s that particulate forms of heavy metals in the water column were nontoxic. Focusing on dissolved forms of heavy metals corrected a long-standing problem in the implementation of the national water quality criteria into state water quality standards. That notwithstanding, the Agency has still not addressed its inappropriate application of water quality criteria for many other constituents, such as organics, to total concentrations rather than properly addressing contaminant availability.

Although deficiencies in the conventional application of the water quality criteria were generally recognized, they were not addressed by the U.S. EPA primarily because the regulations that were developed were not being enforced by either the U.S. EPA or many of the states. As a result, the promulgation took place of the National Toxics Rule in subsequent revisions of the Clean Water Act, through which Congress mandated that states either adopt the U.S. EPA criteria for toxics or have them imposed on them by the U.S. EPA. By the early 1990s, all states had adopted U.S. EPA criteria for "toxics." California's regulations adopting the U.S. EPA criteria as state standards, however, were soon judged invalid through court action because California state law also requires consideration of economic impact of water pollution control regulations. As the California State Water Resources Control Board did not comply with those state requirements, the courts determined that the regulations must be voided. Thus, for many years, California did not have water quality criteria/objectives for "toxics." Finally, in 2000, the U.S. EPA Region 9 imposed what became known as the California Toxics Rule criteria (5). They are the U.S. EPA criteria for "toxics" that were originally adopted in the mid-1980s, or subsequent updates, such as the U.S. EPA (4,6–8). The most recent update of national water quality criteria occurred in 2002, when the U.S. EPA (9) developed its currently recommended national water quality criteria. The U.S. EPA requires that as states update their criteria, they incorporate the 2002 criteria as the state standards.

Lee et al. (10), Lee and Jones (11), and Lee and Jones-Lee (12) discussed alternative approaches for assessing and controlling the impact of contaminants on water quality that took better account of contaminant availability to affect water quality, thus directing the financial resources available first toward defining those constituents that adversely impact the beneficial uses of a waterbody and then toward controlling those constituents to the extent necessary to protect those uses.

PRIORITY POLLUTANT LIST

The 1972 Clean Water Act also mandated that the U.S. EPA develop a list of "priority pollutants" that was to include those chemicals found in water that should receive the highest attention for regulatory action. National water quality criteria were then to be developed for each of those chemicals to protect fish and aquatic life in all waters. Congress, however, did not fund the U.S. EPA adequately to carry out this mandate. Finally, when the U.S. EPA could not develop the list within the timeframe allowed, an environmental group filed suit to force the U.S. EPA to do so. In response to that suit, in the mid-1970s, the U.S. EPA's attorneys and environmental group attorneys, with limited technical input and without public peer review, promulgated what is now known as the "Priority Pollutant" list.

Despite the intention to focus on water quality problems, in actuality, the "Priority Pollutant" list has proven to be detrimental to properly defining the constituents that are significantly adverse to the beneficial uses of waterbodies. The list was not properly peer-reviewed for its inclusion and prioritization of pollutants of real water quality significance. Instead, the list focused primarily on what are known as rodent carcinogens, i.e., those constituents that, at high concentrations over extended periods of exposure, cause cancer in rats, but do not necessarily have a great significance to water quality. Large amounts of public resources have been devoted to analyzing for and then developing control programs for many of the rodent carcinogens, especially the chlorinated solvents, while neglecting many pollutants of greater water quality significance.

In looking at this situation, Dr. Christian Daughton, Chief, Environmental Chemistry Branch, U.S. EPA National Exposure Research Laboratory, indicated that there are more than 22 million organic and inorganic substances, with nearly 6 million commercially available. The current water quality regulatory approach addresses fewer than 200 of those chemicals as potential water pollutants, i.e., the "priority pollutants." He stated, "Regulated pollutants compose but a very small piece of the universe of chemical stressors to which organisms can be exposed on a continual basis" (13).

IMPLEMENTATION OF TMDLS

The establishment of TMDLs focuses on achieving water quality standards in receiving water. This focus led to a review of the TMDL-based water pollution control program by the National Research Council (NRC). The NRC's review (14) discussed technical deficiencies in the U.S. EPA's TMDL program. Waterbodies have been placed inappropriately on the Clean Water Act section 303(d) list of impaired waterbodies on the basis of unreliable evaluation. TMDL goals of achieving worst-case water quality standards for total concentrations are often inappropriate goals for solving real, significant, water quality/use-impairment problems in a technically valid, cost-effective manner. Most importantly, inadequate time and inadequate funding are available to support the development of TMDLs as they are being administered through the U.S. EPA and state regulatory agencies.

The first step toward establishing a more appropriate TMDL process should be an assessment of the appropriateness of the water quality standards that established the 303(d) listing and the standards that serve as TMDL goals to correct water quality impairment. As a considerable part of the TMDL program is directed toward sources such as agricultural runoff and urban runoff that frequently contain constituents that are, in substantial proportion, nontoxic, and unavailable, it is important that the U.S. EPA and the states refocus TMDL programs on controlling toxic, available forms, as opposed to total concentrations of constituents.

BPJ TRIAD WEIGHT-OF-EVIDENCE

Because of the technical inappropriateness and unreliability of using overly simplistic, even though administratively expedient, indicators of impact, increasing attention is being given to the use of a triad "weight-of-evidence" approach as a regulatory tool for water quality impact assessment and management. Although this approach has been configured in several ways with varying degrees of technical validity, a well-accepted, technically appropriate format is a "best professional judgment" (BPJ) evaluation of a triad of key parameters: aquatic life toxicity/bioaccumulation of the contaminant(s); aquatic organism assemblage information in the aquatic system of concern; and chemical kinetic/thermodynamic information pertaining to the contaminant(s) of interest and the aquatic system of concern. These components are described below.

Toxicity/Bioaccumulation

The availability of most chemical contaminants to aquatic life cannot be determined through chemical analysis; the availability can be affected by the character of the ambient water, the nature of the exposure conditions, and the sensitivity of the organisms of concern. Furthermore, routine chemical analysis does not cover all potential toxicants that may be present. Therefore, a key component of a BPJ weight-of-evidence evaluation for water quality impact assessment and management is aquatic life toxicity testing and/or bioaccumulation testing of aquatic organism tissue for potentially hazardous chemicals that are a threat to human health or higher trophic-level organisms that use aquatic life as food. The importance of using the BJP triad framework for the interpretation of the results of this testing is illustrated by the fact that, as discussed by Lee and Jones-Lee (15), finding aquatic life toxicity in laboratory tests of ambient water or sediment cannot be assumed to mean that that toxicity represents a significant impairment of the beneficial uses of the waterbody that is of concern to the public. It is not necessarily possible to equate laboratory-based water column or sediment toxicity with water quality impairment. For example, many sediments have natural toxicity because of low dissolved oxygen, ammonia, and hydrogen sulfide; yet the waterbodies associated with them have excellent fisheries and high water quality. The other aspects of the triad must be used with this information to make a best professional judgment regarding the significance of the toxicity and bioaccumulation information.

Organism Assemblage

Organism assemblage information includes description of the numbers, types, and characteristics of aquatic life and, as appropriate, terrestrial organisms such as fish-eating birds, present at a potentially impacted site. Insight into whether water quality has been adversely affected can be gained from understanding how the organism assemblage at the location compares with that which might be expected to be supported there. However, a variety of physical (flow, temperature, sunlight, sediment, and habitat alteration, etc.), non-potential-pollutant chemical (TDS, nutrients, organic constituents, hardness, alkalinity, etc.), and biological (reproductive cycles, disease, predation, etc.) factors other than chemical potential pollutants can affect the numbers, types, and characteristics of aquatic life in a waterbody's water column or sediments. Therefore, it is critical that this information be considered in the context of a BPJ evaluation in assessing the water quality impacts of chemicals on the beneficial uses of a waterbody.

Chemical Information

Chemical concentration has been one of the most convenient, yet most misunderstood and misused, pieces of information in water quality evaluation and management. Unfortunately, chemical information and "chemistry" have often been considered to be a laundry list of total concentrations of a few regulated constituents having water quality standards, which are mechanically applied to discharges, ambient water, or sediment. This result occurs despite that for half a century, it has been known that the total concentration of a potentially toxic constituent in the water column and/or sediments is an unreliable basis for estimating the impacts of that contaminant on the beneficial uses of a waterbody. Thus, the administratively expedient application of such values directly can distort the significance of that chemical contaminant to water quality.

The reason the total concentration of a selected chemical is unreliable for assessing water quality/use-impairments is that many chemical constituents exist in aquatic systems in a variety of chemical forms, only some of which are toxic or available to affect aquatic life. The aquatic "chemistry," i.e., the kinetics (rates) and thermodynamics (positions of equilibrium) of reactions that a chemical can undergo in a natural water system, is illustrated in the aquatic chemistry "wheel" presented in Fig. 1.

As illustrated, chemicals can undergo oxidation/reduction, volatilization, photochemical, complexation, sorption/desorption, precipitation, hydrolysis, and acid/base reactions in natural water systems. Many reactions alter the availability of the chemical to affect aquatic life. Although measurement of the total concentration of a chemical includes essentially all of these forms, the amount of the "available" forms of the chemical present depends in part on the nature of the chemical and the types and amounts of materials in the water and/or sediment that act to "detoxify" the chemical, i.e., render it nontoxic or unavailable to affect aquatic life through these reactions. These types of materials include organic carbon, sulfides, carbonates, hydrous oxides, clay minerals, and others. The amount of available forms also depends on the rates of reaction, the extent to which these reactions occur, and the comparative availability among the forms.

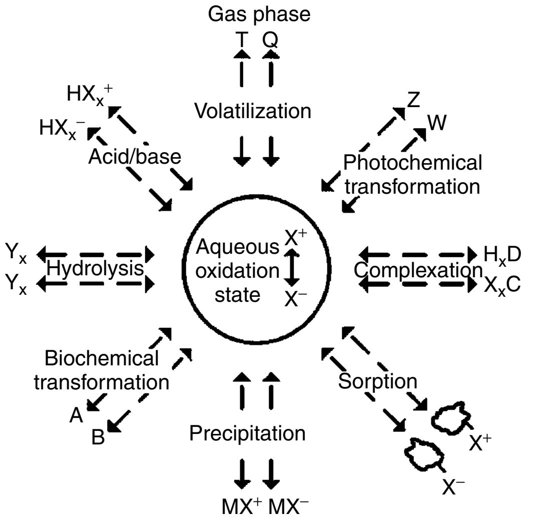

- Distribution among species depends on kinetics & thermodynamics of reactions in the particular aquatic system
- Each chemical species has its own toxicity characteristics many forms are non-toxic

Figure 1. Aquatic chemistry of chemical constituents.

No simplistic method exists by which to reliably quantitatively account for these interactions by mathematical manipulation or chemical analysis. It is for this reason that chemical concentration information, especially total concentrations of chemicals, should not be used alone in assessing water quality or impairment of water quality. The disregard for the aquatic chemistry of contaminants and for the nature and duration of organism exposure in aquatic systems makes the application of worst-case, numeric water quality criteria typically overly restrictive for the protection of beneficial uses of waterbodies. Although chemical concentration information can raise issues to consider and sources that may be worthy of further investigation, it is unreliable for reaching a conclusion about "water quality" or beneficial use-impairment. It is for this reason that the aquatic toxicity/bioaccumulation and organism assemblage information are also critical parts of the triad evaluation, and the need for objective, technically informed, best professional judgment is clear.

CO-OCCURRENCE-BASED SEDIMENT QUALITY GUIDELINES: CAUTION

The issue of sediment quality criteria and guidelines warrants special attention in this discussion. Reliable evaluation and regulation of the impacts of contaminants in sediment systems is substantially more complex and more ill-defined than in ambient waters. Although it has been long understood that no relationship exists between the total concentrations of chemical contaminants in sediments and toxicity, bioaccumulation, or adverse impact on beneficial uses of a water, chemical concentration is simple and numeric, and lends itself to easy decision-making. In an attempt to make simple chemical analysis relevant, some have developed and advanced the use of co-occurrence-based "sediment quality guidelines" (16–18). A group of sediments was evaluated for having some type of biological impact. Then the concentrations of a few selected contaminants in those sediments that exhibit an "impact" were examined. Basically, then, for each contaminant, the lowest concentration associated with an "impacted" sediment was said to be adverse, a cause for concern in any sediment, and the basis for a "sediment quality guideline." No consideration was given to the actual cause of the "impact" reported, to the fact that total concentration is not related to impact, to several chemical and conditions that are well-understood to cause sediment toxicity, or to sediments having that or higher concentrations of the contaminant without exhibiting adverse impacts. The only basis for the so-called "guideline" concentration was the "co-occurrence" of the contaminant in the sediment with some biological impact attributed to that sediment. It is entirely expected that chemical constituents derived from urban industrial areas, although having biological effects, often occur in the presence of other constituents, which, although in nontoxic, unavailable forms, are present in elevated concentrations.

Co-occurrence-based approaches exemplify inappropriate use of chemical information in a water quality assessment but nonetheless are popularized because of their ease of application. As discussed by Lee and Jones-Lee (19–21), no cause-and-effect relationship is established in the co-occurrence-based values. Although so-called "correlations" exist between toxicity and a sediment quality guideline is exceeded, this is coincidental and unreliable for an assessment of the cause of the biological impacts. The fact that a chemical constituent exceeds a particular "sediment quality guideline" does not mean that that constituent is in any way related to biological effects, such as toxicity, bioaccumulation, and/or changes in organism assemblages. The actual cause of the biological response can readily be either a constituent(s) that is not measured or not considered in the scheme or a combination of constituents that, although measured, do not, individually or summed, exceed the "sediment quality guidelines." Thus, no expectation can exist that funds spent to achieve "sediment quality guideline" values will result in any improvement in sediment/water quality or that sediments targeted by the exceedance of guideline values are, in fact, of the greatest concern.

Some try to skirt the fundamental technical flaws in the approach by limiting its use to "screening" sediments. However, using a patently unreliable "screening" tool can do nothing but provide patently unreliable results, which will serve to misdirect concern, responsibility, and funds for remediation, and leave real problem areas unaddressed. Although sediments that exceed one or more "sediment quality guidelines" may, in fact, merit further investigation or remediation, the guideline values are meaningless for making that assessment; under no circumstances should anyone assume that the exceedance of a guideline value represents a cause-and-effect relationship that can determine the likely cause of a biological response. As discussed by O'Connor (22–24), O'Connor and Paul (25), O'Connor et al. (26), Engler (27), DiToro (28), Chapman (29), Burton (30), and Lee and Jones-Lee (15,19–21), the co-occurrence approaches are obviously technically invalid and unreliable for assessing cause-and-effect. A reliable evaluation and regulatory program must be based on reliable assessments of the cause of the adverse effect.

RECOMMENDED APPROACH FOR INCORPORATION OF CHEMICAL INFORMATION INTO A BPJ TRIAD WEIGHT-OF-EVIDENCE WATER QUALITY EVALUATION

The recommended approach for the use of chemical concentration information in a BPJ triad weight-of-evidence evaluation of the cause of a water quality impairment begins with the reliable definition of the water quality/use-impairment that is of concern. The existence of a contaminant in total concentrations greater than a worst-case-based standard/guideline is not, in itself, an adverse impact unless that contaminant is causing adverse impacts to the beneficial uses of the waterbody. Thus, for example, before measuring the concentrations of copper, lead, zinc, and cadmium that typically occur in street and highway stormwater runoff at concentrations above U.S. EPA worst-case-based water quality criteria and state water quality standards based on those criteria, the chemical impact evaluation approach determines whether the water or sediment of concern is toxic. Jones-Lee and Lee (31) describe an evaluation monitoring approach that has been developed to focus on chemical impacts rather than on chemical concentrations.

If toxicity is found in laboratory tests of a water or sediment, then an assessment should be made as to whether that toxicity translates to an adverse impact on the waterbody's beneficial uses, such as fisheries, survival or reproduction of desired aquatic life, and so on. It should not be assumed that toxicity measured in a standard toxicity test necessarily translates to a toxicity that is significantly altering the numbers, types, and characteristics of desirable forms of aquatic life in a waterbody. This fact is especially true for situations such as urban area and highway stormwater runoff, where there can be short-term pulses of toxicity associated with a runoff event that are not of sufficient magnitude or duration to exceed the critical magnitude/duration coupling needed to be adverse to important forms of aquatic life in a waterbody.

Figure 2 shows a typical relationship between duration of exposure and toxicity; as illustrated, the manifestation of a toxicity response to available forms of a contaminant depends on the duration of exposure. Elevated concentrations of a toxicant can be tolerated by aquatic life, provided that the duration of exposure of this toxicity is shorter than the critical, or threshold, toxicity/duration of exposure coupling for that contaminant and organism in the ambient water. Exposure durations of aquatic life can be affected by the intermittent nature of the introduction of the contaminant, the rate and nature of dilution in the ambient water, avoidance or attraction behavior of the organism, and so on.

If measured toxicity is determined to potentially affect the beneficial uses of the waterbody, then a toxicity identification evaluation (TIE) should be conducted to evaluate the cause of the toxicity (32–34). The TIE process defines the availability of contaminants present and can elucidate the potential availability of the toxicity-causing contaminant(s) from its various sources. This process, then, can direct control toward those sources that are contributing significantly to the toxicity. Not only is this approach more technically sound, but it also addresses the issue of potential impacts of unrecognized, unmeasured, and/or unregulated pollutants that has been the subject of increasing importance in water quality management.

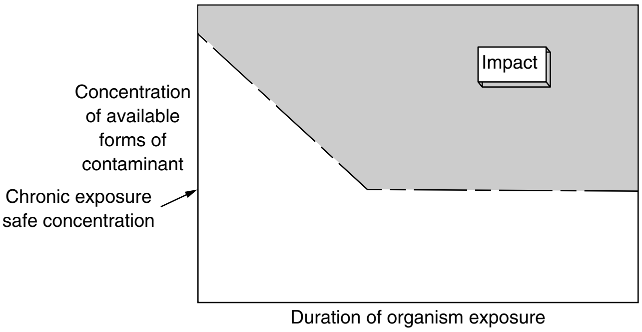

Figure 2. Critical concentration—duration of exposure relationship.

SUMMARY AND CONCLUSION

The federal Clean Water Act and amendments have directed the development of water quality criteria and standards and a framework for identifying and addressing beneficial use-impairments in the waters of the country. In the quest for administratively simple means by which to evaluate and regulate water quality and sediment quality, worst-case-based criteria and standards for a select group of chemicals have been developed. These are applied to the concentrations of contaminants in ambient waters as a means of judging water quality. This approach is known to be unreliable for evaluating or regulating water quality because it does not properly account for the beneficial uses of the water; the nature, behavior, and forms of chemical contaminants; or the unrecognized and unregulated pollutants.

The BPJ triad weight-of-evidence approach described herein provides a technically sound foundation and approach for assessing water quality impacts, causes of the impacts, and directions for controlling the impact. It is not as administratively expedient as simple comparison of chemical concentration data; however, that simplistic approach is not reliable for determining impairments or effecting control of real water quality problems. The BPJ triad weight-of-evidence approach requires the allocation of sufficient funds to determine the characteristics of the constituents/conditions of concern, with particular emphasis on properly defining toxicity and water quality cause-and-effect relationships. It also requires that individuals knowledgeable in aquatic chemistry, aquatic toxicology, and water quality provide guidance on, and appropriate interpretation of the results of, the kinds of chemical and toxicity studies that are needed to appropriately incorporate chemical information into assessing the water quality significance of chemical constituents in impacting the beneficial uses of a waterbody.

Numeric weight-of-evidence approaches in which arbitrary scale factors are assigned to each of the three components of the triad are technically invalid, because the scaling factors do not represent the relationship between a chemical constituent in a water or sediment and its impact on the water quality-beneficial uses of a waterbody.

Significant problems can occur with the use of the BPJ approach in incorporating chemical information into the triad. The use of total concentrations of constituents and/or the exceedance of a co-occurrence-based so-called sediment quality guideline is technically invalid. Such an approach can distort the triad water/sediment quality evaluation because it incorporates information into the triad that is not related to the impact of the chemicals on aquatic-life-related beneficial uses.

The BPJ weight-of-evidence approach should be based on the consensus of a panel of experts who, in a public, interactive, peer-review process, consider the information available, define what additional information is needed, and then render an opinion as to the integrated assessment of the information available on the significance of a particular chemical constituent in impacting the beneficial uses of a waterbody. The characteristics of the components of a BPJ weight-of-evidence approach that focuses on the appropriate use of chemical information are discussed further by Lee and Jones-Lee (35).

BIBLIOGRAPHY

1. Lee, G.F. (2001). Comments on US EPA TMDL Program. U.S. Environmental Protection Agency, Office of Water, Washington DC, December. Available: http://www.gfredlee.com/TMDLcomments.pdf.
2. NAS/NAE. (1973). *Water Quality Criteria of 1972*. National Academies of Science and Engineering, EPA/R3-73-033, U.S. Environmental Protection Agency, Washington, DC.
3. U.S. EPA. (1976). *Quality Criteria for Water*. U.S. Environmental Protection Agency, Washington, DC.
4. U.S. EPA. (1995). *Stay of Federal Water Quality Criteria for Metals; Water Quality Standards; Establishment of Numeric Criteria for Priority Toxic Pollutants; States' Compliance Revision of Metals Criteria; Final Rules*. *Federal Register* **60**(86): 22228–22237, U.S. Environmental Protection Agency, Washington, DC.
5. U.S. EPA. (2000). *Water Quality Standards; Establishment of Numeric Criteria for Priority Toxic Pollutants for the State of California; Rule*. U.S. Environmental Protection Agency, Region 9, Federal Register 40 CFR Part 131, Vol. 65, No. 97, [FRL-6587-9], RIN 2040-AC44, San Francisco, CA.
6. U.S. EPA. (1987). *Quality Criteria for Water 1986*. EPA 440/5086-001, U.S. Environmental Protection Agency, Office of Water Regulations and Standards, Washington, DC.
7. U.S. EPA. (1996). *1995 Updates: Water Quality Criteria Documents for the Protection of Aquatic Life in Ambient Water*. EPA-820-B-96-001, U.S. Environmental Protection Agency, Office of Water, Washington, DC.
8. U.S. EPA. (1999). *National Recommended Water Quality Criteria—Correction*. US Environmental Protection Agency, Office of Water, EPA 822-Z-99-001, Washington, DC.
9. U.S. EPA. (2002). *National Recommended Water Quality Criteria: 2002*. EPA-822-R-02-047, U.S. Environmental Protection Agency, Washington, DC.
10. Lee, G.F., Jones, R.A., and Newbry, B.W. (1982). Alternative approach to assessing water quality impact of wastewater effluents. *J. Water Pollut. Control Fed.* **54**: 165–174.
11. Lee, G.F. and Jones, R.A. (1987). Assessment of the Degree of Treatment Required for Toxic Wastewater Effluents. *Proc. Int. Conf. on Innovative Biological Treatment of Toxic Wastewaters*, U.S. Army Construction Engineering Research Laboratory, Champaign, IL, pp. 652–677.
12. Lee, G.F. and Jones-Lee, A. (1995). Appropriate use of numeric chemical water quality criteria. *Health and Ecological Risk Assessment* **1**: 5–11; Letter to the Editor, Supplemental Discussion, **2**: 233–234 (1996). Available: http://www.gfredlee.com/chemcri.htm.
13. Daughton, C.G. (2004). Pharmaceuticals and Personal Care Products (PPCPs) as Environmental Pollutants: Pollution from Personal Actions. Presentation at the California Bay-Delta Authority Contaminant Stressor Workshop, February. Additional information on PPCPs and other unregulated chemicals is available at www.epa.gov/nerlesd1/chemistry/pharma/index.htm.
14. NRC. (2001). *Assessing the TMDL Approach to Water Quality Management*. National Research Council, National Academy Press, Washington, DC.
15. Lee, G.F. and Jones-Lee, A. (1996). Evaluation of the Water Quality Significance of the Chemical Constituents in Aquatic Sediments: Coupling Sediment Quality Evaluation Results to Significant Water Quality Impacts. WEFTEC '96, Surface Water Quality and Ecology I & II, Proc. Water Environ. Fed. Annual Conference **4**: 317–328. Available: http://www.gfredlee.com/wef_seda.htm.
16. Long, E.R. and Morgan L.G. (1991). The potential for biological effects of sediment-sorbed contaminants tested in the national status and trends program. Tech. Memo. NOS OMA52, National Oceanic and Atmospheric Administration, Seattle, WA.
17. Long, E.R., MacDonald, D.D., Smith, S.L., and Calder, F.D. (1995). Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. *Environ. Manage.* **19**: 81–97.
18. Long, E.R. and MacDonald, D.D. (1998). Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems. *Human and Ecological Risk Assessment* **4**(5): 1019–1039.
19. Lee, G.F. and Jones-Lee, A. (1993). Sediment Quality Criteria: Numeric Chemical vs. Biological Effects-Based Approaches. Proc. Water Environment Federation National Conference, Surface Water Quality & Ecology, pp. 389–400. Available: http://www.gfredlee.com/sedqualcri.html.
20. Lee, G.F. and Jones-Lee, A. (1996). 'Co-Occurrence' in Sediment Quality Assessment. Report of G. Fred Lee & Associates, El Macero, CA, February. Available: http://www.members.aol.com/apple27298/COOCCUR2PAP.pdf.
21. Lee, G.F. and Jones-Lee, A. (2004). Unreliability of Co-Occurrence-Based Sediment Quality 'Guidelines' for Evaluating the Water Quality Impacts of Sediment-Associated Chemicals. Report of G. Fred Lee & Associates, El Macero, CA, September. Available: http://www.gfredlee.com/psedqual2.htm.

22. O'Connor, T.R. (1999). Sediment quality guidelines do not guide. *Environ. Toxicol. Chem. SETAC News* **19**: 28–29.
23. O'Connor, T.R. (1999). Sediment Quality Guidelines Reply-to-Reply. NOAA National Status and Trends Program, May.
24. O'Connor, T.P. (2002). Empirical and theoretical short-comings of sediment-quality guidelines. In: *Handbook on Sediment Quality*. Whittemore et al. (Eds.). Special Publication, Water Environment Federation, Alexandria, VA, pp. 317–325.
25. O'Connor, T.P. and Paul, J.F. (2000). Misfit between Sediment Toxicity and Chemistry. *Mar. Poll. Bull.* **40**: 59–64.
26. O'Connor, T.P. et al. (1998). Comparisons of measured sediment toxicity with predictions based on chemical guidelines. *Environ. Toxicol. Chem.* **17**: 468–471.
27. Engler, R. (2004). U.S. Army Corps of Engineer Waterways Experiment Station, Vicksburg, MS, Personal Communication.
28. DiToro, D. (2002). Sediment Toxicity Prediction—What is Known and Unknown. Presented at Sediment Quality Assessment (SQA5), Aquatic Ecosystem Health and Management Society, Chicago, IL.
29. Chapman, P.M. (2004). 'Paracelsus' Dictum for Sediment Quality (and other) Assessments. Presented at the 2002 Fifth International Symposium on Sediment Quality Assessment (SQA5), In: *Aquatic Ecosystem Health and Management*. M. Munawar, (Ed.). Vol. 7, pp. 369–374.
30. Burton, G.A., Jr. (2002). Summary of Pellston Workshop on Use of Sediment Quality Guidelines (SQGs) and Related Tools for the Assessment of Contaminated Sediments. Presented at the 2002 Fifth International Symposium on Sediment Quality Assessment (SQA5), Chicago, IL.
31. Jones-Lee, A. and Lee, G.F. (1998). Evaluation Monitoring as an Alternative to Conventional Water Quality Monitoring for Water Quality Characterization/Management. Proc. of the NWQMC National Conference "Monitoring: Critical Foundations to Protect Our Waters." U.S. Environmental Protection Agency, Washington, DC, pp. 499–512. Available: http://www.gfredlee.com/wqchar_man.html.
32. U.S. EPA. (1991). *Methods for Aquatic Toxicity Identification Evaluations, Phase I Toxicity Characterization Procedures*, 2nd Edn. EPA/600/6-91/003, U.S. EPA Office of Research & Development, Washington, DC.
33. U.S. EPA. (1993). *Methods for Aquatic Toxicity Identification Evaluations, Phase II Toxicity Identification Procedures for Samples Exhibiting Acute and Chronic Toxicity*. EPA/600/R-92/080, U.S. EPA Office of Research & Development, Washington, DC.
34. Lee, G.F. and Jones-Lee, A. (2002). Issues in Developing a Water Quality Monitoring Program for Evaluation of the Water Quality/Beneficial-Use Impacts of Stormwater Runoff and Irrigation Water Discharges from Irrigated Agriculture in the Central Valley, CA. California Water Institute Report TP 02–07 to the California Water Resources Control Board/Central Valley Regional Water Quality Control Board, California State University Fresno, Fresno, CA, December. Available: http://www.gfredlee.com/Agwaivermonitoring-dec.pdf.
35. Lee, G.F. and Jones-Lee, A. (2004). Appropriate Incorporation of Chemical Information in a Best Professional Judgment 'Triad' Weight of Evidence Evaluation of Sediment Quality. Presented at the 2002 Fifth International Symposium on Sediment Quality Assessment (SQA5), In: *Aquatic Ecosystem Health and Management*. M. Munawar, (Ed.). Vol. 7, pp. 351–356. Available: http://www.gfredlee.com/BPJWOEpaper-pdf.

THE CONSTITUTION AND EARLY ATTEMPTS AT RATIONAL WATER PLANNING

Martin Reuss, Ph.D
Office of History Headquarters
U.S. Army Corps of Engineers

In the first decades of the American Republic, river and harbor bills met overwhelming resistance from the executive branch, mainly because of concerns over constitutionality, one more issue that reverberates through much of America's history. The specific points usually centered on presumed lack of constitutional authority to construct works of mainly local and even private benefit or works on rivers that were not clearly navigable. Pre-Civil War Presidents either practiced an inconsistent policy towards public works or became adamant opponents. Thomas Jefferson objected to federal involvement because it would empower the government at the expense of the common man, burden the taxpayer, lead to projects benefiting one location at the cost of another, and would enrich men at public expense. James Madison, who had been an early internal improvements advocate, on the evening before his departure from the White House vetoed the so-called Bonus Bill that would have provided funds for public works, declaring that the Constitution did not give Congress the power to appropriate money for or to construct public works without an "inadmissible latitude of construction." James Monroe at first thought that Congress could appropriate funds for public works but agreed with Madison that the federal government had no authority to construct the projects. Later, he determined that Congress might construct public works after all but only for those projects that were "national not state, general not local," a clarification that left the proverbial barn door open for defining local and general.

Andrew Jackson saw himself as a friend of internal improvements, but he feared the extension of federal power, sought a clarifying amendment on the appropriate national role in internal improvements, and admitted, like Monroe, that, while the federal government could appropriate money for truly national projects, it could not actually construct the projects itself. Strict constitutional constructionist James K. Polk vetoed every rivers and harbors bill sent to him. He even went to his office on the last day of his administration armed with a veto should Congress try to pass an internal improvements bill at the last moment. Abraham Lincoln, a young Whig congressman from Illinois, succinctly captured the problem in his denunciation of Polk's veto of the 1848 rivers and harbors act. "The just conclusion from all this is that, if the nation refuses to make improvements of the more general kind because their benefits might be somewhat local, a state may, for the same reason, refuse to make an improvement of a local kind because its benefits may be somewhat general. A state may well say to the nation, 'If you will do nothing for me I will

---

do nothing for you.'" In the following decade, Franklin Pierce vetoed five rivers and harbors bills on the grounds of unconstitutionality. As Civil War erupted in the land, the political and philosophical jousting over water projects remained as short of resolution as ever.

Given the fears, hopes, and questions facing the early American republic, it is little wonder that it saw no successful implementation of coordinated public works administration. Perhaps more surprising is that this became a permanent condition in the United States. Funding issues, sectional friction, and constitutional questions invariably posed insurmountable barriers. Prior to the Civil War, the federal government attempted twice to develop and implement a national program of public works. The first was the well-known Gallatin Plan. At the request of Congress, Secretary of Treasury Albert Gallatin proposed in 1808, an ambitious network of roads and canals connecting the Eastern seaboard with the interior and a coastwise water route to shorten distances between major Atlantic seaboard cities. Gallatin argued that the federal government should construct internal improvements that provide "annual additional income to the nation" but are beyond the capacity of private entrepreneurs to build. His formulation harkens back to Adam Smith's The Wealth of Nations and anticipates the term "National Economic Development " that appeared in twentieth century economic jargon. However, his effort fell victim to lack of funds (both private and public), New England opposition to the Jefferson Administration, and, finally, growing preoccupation with real and apparent British threats to the United States, which eventually resulted in war. It is also of more than passing interest that Gallatin himself agreed with President Jefferson that his plan could never be efficiently realized without a constitutional amendment.

The next great attempt occurred in 1824. President's Monroe's vacillation, the growing clout of new states interested in waterborne commerce, and a favorable Supreme Court ruling (Gibbons v. Ogden) that sanctioned federal control over interstate commerce, including rivers, based on the Commerce Clause of the Constitution, allowed passage of the General Survey Act in 1824, after weeks of acrimonious debate. The act carried largely because of support from the Middle Atlantic states (except Delaware) and the new states west of the Appalachians. It authorized the President to use the Army Engineers to survey (not build) roads and canals (not rivers) that may be deemed "of national importance in a commercial or military point of view, or necessary to the transportation of the public mail."

Once the bill passed, Secretary of War John C. Calhoun organized a Board of Engineers for Internal Improvements to determine which routes should be surveyed among the scores suggested. Like Gallatin's plan, this program could have become the beginning of a great nationwide, coordinated system of internal improvements. Instead, once projects were surveyed, they became subject to the same parochialism in Congress that had doomed earlier, similar ideas, and congressmen continued to introduce pet projects for funding despite a contrary recommendation from the army engineers or the absence of a survey altogether. Its planning role severely diminished, the Board of Engineers languished, and a reorganization of the Corps of Engineers in 1830 provided an excuse for its abolishment. Six years later, Congress repealed the General Survey Act, partially a response to the legislature's own abuse of the act, using Army Engineers to survey potential projects of clear local and even private interest. Thus began a contest between rational administration and congressional politics that has remained unresolved and contentious to the present.

This article is based on Martin Reuss, "The Development of American Water Resources, Planners, Politicians, and Constitutional Interpretation," in Paul Slack and Julie Trottier, eds., Managing Water Resources, Past and Present: The Twelfth Annual Linacre Lectures. Oxford University Press, 2004. Forthcoming. Used by permission of Oxford University Press.

The Development of American Water Resources: Planners, Politicians, and Constitutional Interpretation.

The Constitution and Early Attempts at Rational Water Planning.

The Expansion of Federal Water Projects.

ECONOMIC VALUE OF WATER: ESTIMATION

Jorge Ramírez-Vallejo
Harvard University
Cambridge, Massachusetts

> "Water has an economic value in all its competing uses and should be recognized as an economic good." The Dublin Statement, 1993.

MOTIVATION FOR VALUING WATER

The global awareness of the vital role water plays in sustainable national development is increasing. Water is vital for sustaining life on the earth. It is crucial for economic and social development, including energy production, agriculture, and domestic and industrial water supplies. Therefore, each unit of water should be used efficiently, equitably, and soundly. The economic value that water generates should be given due attention, when apportioning scarce water resources among competing uses.

Water is often regarded as a naturally occurring, freely available commodity. However, its supply, especially in an unpolluted state, is often limited compared to the demands placed upon it. Hence, water is both a constraint on and an essential input of economic activity, and decisions are required to allocate it among competing uses. Considerable resources are committed to delivering required standards of water services in adequate quantities and qualities. In addition, water-related environmental qualities, such as those associated with wetland regimes, are recognized as valuable to society. These characteristics of water imply that it can be regarded as an economic commodity—an economic good—whose use can be guided by economic principles (1).

Water, as other resources, has value whenever users are willing to pay a price for it rather than do without. The economic value and economic benefit of water can be considered synonymous. The benefits from this resource are derived from various uses. First, we find benefits of water from personal drinking, cooking, and sanitation, and from that those contribute to productive activities on farms and in businesses and industries. These uses are called *consumptive* uses. Second, there are *in-stream* water benefits from other uses such as hydroelectric power generation and waterways transportation. These uses are called *nonconsumptive* uses. The third type of value of water comes from the economic benefits that water has in recreation, aesthetics, and fish and wildlife habitat. As income and leisure time grow, these benefits are becoming more and more important. Fourth, there are the economic benefits of waste disposal using water as a sink for carrying away a wide range of residuals from the processes of human production and consumption. Water is used then to dilute them and in some instances to transform them into less undesirable form. Fifth, we find the *dis-values* of water that come from floods and negative water quality impacts. Finally, we have the *nonuse* value of water that comes from people who are willing to pay for environmental services that might neither be used nor experienced in the future.

Economic valuation is a process of expressing preferences in a money metric. But, why bother valuing water anyway? This is the obvious question that comes to mind when approaching this subject. The key answer to this question is simple but fundamental: for the future well-being of humanity. The longer we ignore or distort water's value, the more overused and misallocated the resource becomes. Without price signals or other indicators of value to help guide policy, we tend to manage the resource blindly.

Estimates of the economic value are key for management decisions on allocation. Economic valuation is also important because it aids in determining if people want proposed projects and helps in estimating the degree to which they are willing to pay for benefits. When public budgets are constrained, this information becomes important for making sound investment decisions. For policy decisions, water values are also important to set the levels that the various users should pay for this resource.

Estimating the economic value of water helps in decisions for optimal allocation of the resource among various uses. Limited supply and increasing demand for water are common around the world. Under these circumstances, the economic value of water for each of its competing uses helps the decision maker decide how much water to allocate to each sector or use, and also helps determine the use of water in time, how much to use now and how much to save for the future.

In 1992, experts from 100 countries, intergovernmental organizations, and NGOs gathered in Dublin for the International Conference on Water and the Environment. The conference influenced preparations for the UN Conference on Environment and Development, held later that year in Rio de Janeiro.[1] Its brief concluding statement remains an authoritative affirmation of the water policy agenda; it contained four "guiding principles." The most important principle from this statement that has generated a great deal of controversy has to do with the consideration that water has an economic value in all its competing uses and should be recognized as an economic good. However, the statement recognizes the basic right of all human beings to have access to clean water and sanitation at an affordable price (see Appendix 1).

ECONOMIC CONCEPTS

To understand better the methodologies for estimating the economic value of water described below, it is important to review two key economic concepts (2).

The Demand Curve and the Price Elasticity of Water Demand

When we talk about willingness to pay for value, we mean the maximum amount that a individual is willing to pay for a product. This is the sum of the product price (P) and the consumer surplus. If we wish to estimate the maximum that the individual is willing to pay for a proposed change, we should estimate the demand curve.

A consumer's willingness to pay for an increment of supply is the corresponding area under the demand curve (price vs. quantity), although the amount the consumer actually pays for the increment is the water price times the quantity. The difference between these two values is what it is called the consumer surplus (the benefit for the consumer that he/she did not pay for it).

Assuming that the price elasticity of demand (ε) is known, is constant over the relevant range, and is not equal to one, and second, that the relevant initial price P_1 and quantity Q_1 can be specified, then the area under the curve (i.e., the desired measure of value—denoted V) for a change in quantity Q_1 to Q_2 is given by

$$V = [(P_1{Q_1}^{1/\varepsilon})/(1 - 1/\varepsilon)][(Q_1/Q_1^{1/\varepsilon}) - (Q_2/Q_2^{1/\varepsilon})]$$

This expression represents the value of an increment of quantity Q to the final user. This incremental value will provide a measure of benefits to compare with costs in an investment project evaluation.

---

[1] Awareness of the seriousness of water problems has evolved gradually over the years, and important recommendations have been formulated since the United Nations Water Conference in Mar del Plata (Argentina) in 1977. Other events that have focused on water issues include the 1992 International Conference on Water and the Environment, convened by the World Meteorological Organization (WMO) in Dublin, Ireland, and the 1992 United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro. The Ministerial Conference on Drinking Water and Environmental Sanitation held in Noordwijk in March 1994 adopted an action program. In 1994, the United Nations Commission on Sustainable Development requested a comprehensive assessment of the world's freshwater resources, including projections of future needs the results of the study were considered by the 1997 special session of the UN General Assembly reviewing the implementation of Agenda 21.

Estimates of price elasticities of water demand have shown that water use is consistently price-inelastic when there is no substitute, whereas residual use (i.e., watering the garden) is more elastic and differs significantly across geographical regions and between climate seasons (i.e., winter–summer, rainy–dry) (3).

The price responsiveness of water demand reflects the heterogeneity of demand components. A small portion of the demand is for uses for which there are no substitutes and that are of great necessity to the consumer; other uses are of lesser value and have ready substitutes.

Treatment of Time

The rate of conversion of value between time periods is called the rate of time of preference. Defining the best long-term management of the resource requires balancing the needs of the present and those of the future. Using a discount rate that adequately considers the value of goods or services for future generations creates this balance. A high discount rate implicitly places a low value on water's value to future generations. A low discount rate implies the opposite.

Another issue related to time is the difference between long-run and short-run values of water. The distinction between these two relates to the fixity of certain inputs necessary to make water available for consumption, such as supply infrastructure. Other things remaining constant, we would expect that values estimated for short-run contexts would be larger than values for the long run. Price elasticity of demand is less in the short run than in the long run when adjustments to shortages and infrastructure can be made. Accordingly, willingness to pay in the short run is usually higher than in the long run.

ESTIMATION METHODOLOGIES

The principles for valuing natural assets such as energy and mineral deposits, forests, and aquifers were set forth more than 60 years ago, but the principles for measuring consumptive value of water for household use were set forth a century ago.

Water services in many situations are difficult to value because much of the information needed for valuation is not readily available. When markets exist, they automatically generate useful information, but unfortunately, most of the services provided by water are not traded on markets. Therefore, one has to rely on techniques to value nonmarket goods.

Economists have developed indirect methods (e.g., hedonic price models, travel cost method, averting behavior), which infer values from other behaviors associated with the good. The strengths of indirect methods come from the fact that they rely on observed behaviors of producers and consumers. Examples are the market price as result of human behavior (demand and supply) as a function of the quality of the water nearby. Recent summaries of indirect approaches can be found in Braden and Kolstad (4), Mendelsohn and Markstrom (5), Smith (6,7), and Freeman (8).

The valuation method must be well matched to the context and the water function or service of interest. It is hard to generalize in the abstract about the validity and reliability of specific valuation approaches. The validity of the methodology used depends on the valuation context and the type of service that are of interest. Different approaches are needed to value different services. A valuation framework takes into account how time, institutions, water quality and quantity, hydrologic factors, and services interact to affect the resource's value.

One of the major challenges in valuing water resources is how to integrate the hydrologic and physical components into a valuation scheme. Valuation of the extractive and onsite services of water requires understanding the hydrology and ecology of the water source. Hydrologic information includes variables such as rainfall, runoff, infiltration, water balance, depth to groundwater, aquifer characteristics, and water quality.

DIRECT VALUATION APPROACHES

Market Prices

When market prices exist, the equilibrium price represents the willingness of potential buyers to pay for the good or service of water. In other words, when water has a market and it can be bought and sold, then the economic value of water is its price. However, in some cases due to government intervention such as subsidies or market failures, these prices are distorted and need to be adjusted to obtain the real economic value of the resource.

Transactions between a buyer and a seller are increasingly observed when water rights are defined and other conditions exist that make the water market function. Easter and Hearne (9) and Rosegrant and Binswanger (10) describe examples of water rights markets. If there is a market for perpetual water rights, the value of water can be estimated by using the perpetuity formula $V = A/r$ where r is the interest rate (discount rate), V the value of the perpetual water right, and A the value of water.

Contingent Valuation Method

Economic value is not a fixed, inherent attribute of a good or service but rather depends on time, circumstances, and individual preferences. One technique that attempts to measure total value is the contingent valuation method (CVM). CVM values are elicited directly from individuals via interviews or questionnaires in the form of statements of maximum willingness to pay (WTP) or minimum willingness to accept compensation (WTA) for hypothetical changes in environmental goods, such as water quantity or quality. However, there are numerous methodological controversies associated with the application of CVM, including the way the hypothetical water change (in quantity or quality) is to be specified, the elicitation format for asking valuation questions, the appropriate measure to be elicited, and various types of response biases. CVM provides reliable estimates of value when an individual has a close connection to the resource being valued.

This methodology is used when no value measures can be derived from observing individual choices through a market, such as water quality issues. This approach can potentially measure the economic benefits (or damages) of a wide assortment of beneficial (or adverse) effects in a way that is consistent with economic theory (11).

The CVM method is an ex-ante technique that can be used to measure the value of water without the need for observable behavior (data). In addition, this is the only method (direct or indirect) that measures existent or bequest values (use and nonuse values). It is an expensive method because of the need to carry out a survey, and there are concerns about its use for nonuse value applications.

Using this approach, respondents of surveys are offered conditions simulating a hypothetical market in which they are asked to express willingness to pay for existing or potential environmental conditions not registered on any market. The contingent valuation method (CVM) derives estimates for social benefits attributable to water, when considered a public or nonmarket good. This method is used frequently for projects or programs that affect water quality. The procedural steps for applying the CVM are as follows (12):

The first component describes the setting under which the respondent is to imagine herself/himself. Second are the choice questions, which will be used to infer values of the amenity or policy change. The third element asks questions about the respondent to identify demand shift variables such as age, education, income, and gender. For the forms of questions used and the way to design the survey instrument in CVM studies, see Mitchell and Carson (11) and Cummings et al. (12).

The CVM approach, given its hypothetical attributes, is susceptible to various sources of errors or biases. Some of them are (1) the questionnaire scenario can encourage strategic behavior, (2) the compliance bias or a tendency to fit the responses to the perceived preferences of the interviewer, (3) the possibility that the scenario provides implied values cues to the respondent, and (4) scenario misspecification when the respondent fails to understand the scenario intended by the researcher. These biases demand a careful questionnaire design.

The willingness to pay is estimated by using the following expression:

$$\mathrm{WTP} = f(\Delta Q_j, \Delta P_j, \Delta \mathrm{SP}_j / I_j, H_j)$$

where ΔQ_j defines a specific change in water quality, ΔP_j represents the change in price of water from primary supplies, ΔSP_j represents the change in price of substitute sources of water, I_j represents household income for individual j net of taxes, and Hj represents various nonincome characteristics of individual j's household.

INDIRECT VALUATION APPROACHES

Derived Demand

Water demand (households, farms, etc.) tends to be very site-specific and is influenced by a range of natural and socioeconomic factors. The demand relationship is represented algebraically by the familiar function,

$$Q_w = Q_w(P_w, P_a, P; Y; \mathbf{Z})$$

where Q_w refers to the individual's level of water consumption in a specified time period, P_w refers to the price of water, P_a denotes the price of an alternative water source, P refers to an average price index representing all other goods and services, Y is the consumer's income, and $\mathbf{Z}$ is a vector representing other factors such as climate and consumer preferences (13). This demand curve is usually estimated by using cross-sectional observations (data taken for the same time period) from a sample of water service agencies which exhibit a range of water rates and often variation in other factors influencing demand, such as income, rainfall, and temperature. The general approach is to apply regression analysis to estimate the parameters of the demand equation, parameters that allow us to estimate the price elasticity of demand.

The gross value of water is derived by integrating (area under the curve) the inverse of the econometrically derived demand function, $P(Q_w)$, over the appropriate increment in quantity. If the demand function was estimated for tap water and one is looking for the consumer demand function for raw water for intersectional allocation, then an additional step is necessary. The costs of capturing, transporting, treating, and storing water must be deducted from the estimated willingness to pay for tap water. When tap water (domestic) is priced to recover fully the costs of supplying it, the average revenue can be subtracted from the value of water estimated before. See Booker and Young (14) for an application of this point.

The derived demand approach is usually used to estimate the value of water in the municipal environment. Its main advantages are that it is based on observable data from firms/households using water and is relatively inexpensive. However, using this approach, it is not possible to measure the nonuse value of water and, sometimes, understates the willingness to pay for the resource (2).

The approach described above does not capture the benefit of loss in value due to the reliability dimension (the probability of having an interrupted service). Some other methodologies such as the contingent valuation method have been used in the past to estimate the reliability attribute of water supply. See, for example, Carson (15), Altaf et al. (16), and Howe and Smith (17).

Valuing Water as an Intermediate Good

Water is commonly used for producing goods; it is not a final product itself. Agricultural irrigation, industry, and hydroelectric or thermal power generation are examples of these uses. For intermediate goods, the theory of producer's demand for inputs provides the conceptual basis for valuing the economic value of water (18). The producer's demand for an input is its value as a marginal product (VMP); therefore measuring the economic value of water requires an approximation of the VMP. Three methods of valuing water as an intermediate good are in common use: the residual approach, the hedonic approach, and the alternative cost approach.

The Residual Approach. The "residual" method is the most frequently used approach for estimating the price of intermediate goods, particularly of irrigation water. Under this method, the incremental contribution of each input in the production process is determined. If appropriate prices can be assigned by market forces to all inputs but one, water, for example, the remainder of total value of product is imputed to the remaining residual input. In algebraic terms,

$$P_w = [\text{TVP}y - [(P_k Q_k) + (P_l Q_l) + (P_R Q_R)]]/Q_w$$

where P_w and Q_w are the price and quantity of water, respectively, TVP_y is the total value product, and the parenthesis terms are the price multiplied by the quantity of each of the other inputs in the production process (K, L, R).

This technique is extremely sensitive to small variations in the assumptions about the nature of the production function and about prices. Biases can be introduced by any of several types of errors and omissions. Therefore, this approach is most suitable when water contributes significantly to the value of the output (agricultural production, hydropower generation).

Sometimes, mathematical programming (optimization) models are used to estimate the value of water using the residual approach. These models are built to obtain the optimum allocation of water and other resources (inputs) so as to maximize profits, subject to constraints on resource availability and institutional capacity (19). Mathematical programming is advantageous when a wide range of technological options is to be studied.

The Alternative Cost Approach. This method is also appropriate for measuring the value of water when water is an intermediate good. The technique is attractive under the assumption that, if a given project of specified output costs less than the next best public or private project which can achieve the same output, then the cost of the next best project can be assigned as the benefit to the public project in consideration. This approach has been employed for valuing of water in municipal, industrial, hydroelectric, and thermal electric power projects. Steiner (20) remains the primary authority on this approach. See also Herfindahl and Kneese (21).

Averting Behavior Model

Using this method, water value is estimated by the actions taken to avoid or reduce damage from exposure to water contaminants, such as changing daily routines to avoid exposure. However, averting expenditures and true reduction of pollution differ because such expenditures do not measure all the costs related to pollution that affect the consumer's utility. Therefore, the ability of this valuation approach to provide a lower bound to willingness to pay depends on various assumptions. A list of these conditions is explained in detailed in the literature (22).

The averting behavior technique is based on observable data from actual behavior and choices and is relatively inexpensive. It provides a lower bound of the willingness to pay if certain assumptions are met. However, its estimates do not capture full losses from environmental degradation, and several key assumptions must be met to obtain reliable estimates. It is usually used for expost analysis and is limited to current situations; as in the derived-demand approach, it does not estimate nonuse values.

Hedonic Price Model

This pricing method is based on the premise that people value a good because of the attributes of that good rather than the good itself. These models encompass both land (housing) price models and wage models that account for variations due to environmental attributes. Therefore, by regressing housing prices to the attributes of a house and site, including environmental characteristics, one obtains the marginal value of the characteristic q. In the estimating equation,

$$P = F(\mathbf{S}, \mathbf{N}, \mathbf{Q}) + \varepsilon$$

P is the sale price of the asset, $\mathbf{S}$ is a vector of structural characteristics, $\mathbf{N}$ is a vector of neighborhood characteristics, $\mathbf{Q}$ represents the environmental attribute(s) of interest, F is a function to be estimated by multiple-regression methods, and ε is an error term. $\partial P/\partial Q$ is the marginal value of the characteristic q or the amount of value gained or lost due to a defined change in the environment.

Estimating the economic value of environmental resources by hedonic methods is quite difficult in practice, and the technique is subject to serious limitations (8). First, a large enough sample might be difficult to obtain. Second, if water resources are already in public ownership, market transactions may not be available. Third, sometimes the effect of an environmental attribute or characteristic of price may be small and hard to detect statistically or to disentangle from the effects of all other variables. See a detailed description and examples in Palmquist (23) and McConnell (24).

Travel Cost Method

The travel cost is the most widely used example of the observed indirect methods. It involves two steps: the first is to estimate the individual recreationist's demand for the resource, and the second is to derive statistically the relevant aggregate resource demand curve. Travel cost methods encompass a variety of models, ranging from the simple single-site travel cost model to regional and generalized models that incorporate quality indexes and account for substitution across sites. The basic premise behind these models is that the travel cost incurred in traveling to a site can be regarded as the price of access to the site. Therefore, under some assumptions, it is possible to derive the demand for visits to a site as a function of the price of admission, which is set equivalent to the cost of travel to the site.

This method is based on observable data from actual behavior and choices and is relatively inexpensive. However, sometimes is not easy to observe consumer behavior regarding the choice of recreation sites, and it is limited to the water uses that include travel such as recreation. It does not measure nonuse values.

The major drawback is that sometimes it is difficult to isolate the contribution of water to the attractiveness of a site. In other words, there are some other attributes that contribute to the total site value and are difficult to isolate. Another problem in using this approach is to decide what to include when estimating the cost of travel (cost of gas, oil, tires, etc), and whether or not the opportunity cost of the recreationist's time should be included in the estimate. A detailed explanation of this approach and real world examples are in Walsh (25), McConnell (24), Freeman (8), and Fletcher et al. (26).

Benefit Transfer and Meta-Analysis

The process of benefit transfer consists of adapting estimates relating to sites that have been studied to sites lacking such studies. It is an approach that consists of pooling data from already completed studies and then applying multiple regression analysis using explanatory variables that are site specific and methodology specific. For a detailed description, see Walsh et al. (27). A similar approach but using research findings is called meta-analysis; it has been used for estimating the benefits of natural resources. See, for example, Smith and Kaoru (28).

AN APPLICATION OF VALUATION

Quality of Water

Valuing the quality of water usually requires establishing its relationship to the health conditions of its users. In the past, various methodologies have been used to measure the economic value of changes in health benefits or risks from a change in water quality levels. For this purpose, they have used various tools or techniques such as water demand functions, averting or defensive expenditures, changes in production costs, the hedonic price method, or stated preference methods such as the contingent valuation method or conjoint analysis (29).

Another service supported by water quality is the provision of drinking water for livestock and irrigation. A change in the quality of water used for these purposes may result in a change in the value of livestock or crops, a change in agricultural production costs, or a change in human health or health risks by increasing or decreasing the trace chemicals found on plants ingested by people. The change in the value of crops can be estimated using a market demand function. The change in production costs can be estimated using cost functions. For changes in human health and health risks, one could use the damage avoided approach and stated preference methods. Similar methods should be used for the eventual changes in the value of crops from a commercial fish harvest, commercial hunting harvest, or commercial plant harvest.

A change in surface or groundwater quality can also have an impact on the quantity or quality of recreational services such as fishing, hunting, or just enjoying the scenery. Potential techniques for valuing these effects include the travel cost method and stated preference methods.

On the other hand, the effects of changes in nonuse or passive use services are some of the most difficult effect to value monetarily. A change in nonuse or passive use services results in a direct change in personal utility or satisfaction that may be reflected in the decision on consumption of water. The most viable techniques for measuring these effects are stated preference methods (11,15).

A comprehensive, rigorous and useful discussion of the theory and practice of valuing environmental resources is given in Freeman (8). See also Anderson and Bishop (30), Bentkover et al (31), Johansson (32), Braden and Kolstad (4), and Pearce and Warford (33).

BIBLIOGRAPHY

1. Young, R.A. and Gray, S.L. (1972). *Economic Value of Water: Concepts and Empirical Estimates*. National Water Commission, Washington, DC.
2. National Research Council. (1997). *Committee on Valuing Ground Water, Valuing Ground Water: Economic Concepts and Approaches*. National Academy Press, Washington, DC.
3. Gibbons, D. (1986). *The Economic Value of Water*. Resources for the Future, Washington, DC.
4. Braden, J.B. and Kolstad, C.D. (1991). *Measuring the Demand for Environmental Quality*. Elsevier Science, Amsterdam.
5. Mendelsohn, R. and Markstrom, D.C. (1988). The use of travel cost and hedonic methods in assessing environmental benefits. In: *Amenity Resource Valuation: Integrating Economics with Other Disciplines*. G.L. Peterson, B.L. Driver, and R. Gregory (Eds.). Venture Publishing, State College, PA.
6. Smith, V.K. (1989). Taking stock of progress with travel cost recreation demand methods: theory and implementation. *Mar. Resour. Econ.* **6**: 279–310.
7. Smith, V.K. (1993). Nonmarket valuation of environmental resources: an interpretive appraisal. *Land Econ.* **69**: 1–26.
8. Freeman, A.M., III. (1993). *The Measurement of Environmental and Resource Values: Theory and Methods*. Resources for the Future Press, Washington, DC.
9. Easter, K.W. and Hearne, R. (1995). Water markets and decentralized water resource management: international problems and opportunities. *Water Resour. Bull.* **31**(1): 9–20.
10. Rosegrant, M.W. and Binswanger, H.P. (1995). Markets in tradeable water rights: potential for efficiency gains in developing countries. *World Dev.* **22**(11): 1613–1621.
11. Mitchell, R.C. and Carson, R.T. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Johns Hopkins University Press for Resources for the Future, Baltimore, MD.
12. Cummings, R.G., Brookshire, D.S., and Schulze, W.D. (1986). *Valuing Environmental Goods: A State of the Art Assessment of the Contingent Valuation Method*. Rowman and Allanheld, Totawa, NJ.
13. Munasinghe, M. (1992). *Water Supply and Environmental Management: Developing World Applications*. Westview Press, Boulder, CO.
14. Booker, J.F. and Young, R.A. (1985). Modelling Intrastate and Interestate Markets for Colorado River Water Resources. *Environ. Econ. Manage.* **26**(1): 66–87.
15. Carson, R.T. (1991). Constructed markets. In: *Measuring the Demand for Environmental Quality*. J.B. Braden and C.D. Kolstad (Eds.). Elsevier Science, Amsterdam.
16. Altaf, M.A., Whittington, D., Jamal, H., Liu, J.L., and Smith, V.K. (1993). Rethinking rural water supply policy in the Punjab Pakistan. *Water Resour. Res.* **29**(7): 1943–1954.

17. Howe, C.W. and Smith, M.G. (1994). The value of water supply reliability in urban water systems. *Environ. Econ. Manage.* **26**(1): 19–30.
18. Just, R.E., Hueth, D.L., and Schmitz, A. (1982). *Applied Welfare Economics and Public Policy*. Prentice-Hall, Englewood Cliffs, NJ.
19. Williams, H.P. (1990). *Model Building in Mathematical Programming*, 3rd Edn. John Wiley & Sons, New York.
20. Steiner, P.O. (1965). The role of alternative cost in project design and selection. *Q. J. Econ.* **79**(3) August.
21. Herfindahl, O.C. and Kneese, A.V. (1974). *Economic Theory of Natural Resources*. Merrill, Columbus, OH.
22. Abdala, C.W., Roack, B.A., and Epp, D.J. (1992). Valuing environmental quality changes using averting expenditures: An application to ground water contamination. *Land Econ.* **68**(2): 163–169.
23. Palmquist, R.B. (1991). Hedonic methods. In: *Measuring the Demand for Environmental Quality*. J.B. Braden and C.D. Kolstad (Eds.). Elsevier Science Publishers, Amsterdam.
24. McConnell, K.E. (1993). Indirect Methods of Natural Resource Damage Assessment. In: *Valuing Natural Assets: The Economics of Natural Resource Damage Assessment*. R.J. Kopp and V.K. Smith (Eds.). Resources for the Future, Washington, DC.
25. Walsh, R.G. (1986). *Recreation Economic Decisions: Comparing Benefits and Costs*. Vbenture Publishing, State College, PA.
26. Fletcher, J.J., Adamovicz, V.L., and Graham-Tomasi, T. (1990). The travel cost model of recreational demand: theoretical and empirical issues. *Leisure Sci.* **12**: 119–147.
27. Walsh, R.G., Johnson, D.M., and McKean, J.R. (1992). *Benefit Transfer of Outdoor Recreation Demand Studies*. 1968–1988. Water Resources Research Vol. 28.
28. Smith, V.K. and Kaoru, Y. (1990). Signals or noise? Explaining the variation in recreational benefit estimates. *Am. J. Agric. Econ.* **71**(2): 419–433.
29. Bergwstrom, J.C., Boyle, K.J., and Poe, G.L. (2001). *The Economic Value of Water Quality, New Horizons in Environmental Economics*. Edward Elgar, Northampton, MA.
30. Anderson, G.D. and Bishop, R.C. (1986). The valuation problem. In: *Natural Resources Economics: Policy Problems and Contemporary Analysis*. D.W. Bromley (Ed.). Kluwer Nijhoff, Boston, pp. 89–137.
31. Bentkover, J.D., Covello, V.T., and Munpower, J. (Eds.). (1999). *Benefit Assessment: The State of the Art*. Reidel, Dordrecht, the Netherlands.
32. Johansson, P-O. (1987). *The Economic Theory and Measurement of Environmental Benefits*. Cambridge University Press, Cambridge, UK.
33. Pearce, D.W. and Warford, J.J. (1993). *World Without End: Economics, Environment and Sustainable Development*. Oxford University Press (for the World Bank), New York.
34. Young, R.A. (1996). *Measuring Economic Benefits for Water Investment and Policies*. World Bank Technical Report No 33B, Washington, DC.

APPENDIX 1. THE DUBLIN STATEMENT

In commending this Dublin Statement to the world leaders assembled at the United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro in June 1992, the Conference participants urge all governments to study carefully the specific activities and means of implementation recommended in the Conference Report, and to translate those recommendations into urgent action programmes for water and sustainable development.

Guiding Principles

Concerted action is needed to reverse the present trends of overconsumption, pollution, and rising threats from drought and floods. The Conference Report sets out recommendations for action at local, national and international levels, based on four guiding principles.

Principle No. 1. Fresh water is a finite and vulnerable resource, essential to sustain life, development and the environment

Since water sustains life, effective management of water resources demands a holistic approach, linking social and economic development with protection of natural ecosystems. Effective management links land and water uses across the whole of a catchment area or groundwater aquifer.

Principle No. 2. Water development and management should be based on a participatory approach, involving users, planners and policy-makers at all levels

The participatory approach involves raising awareness of the importance of water among policy-makers and the general public. It means that decisions are taken at the lowest appropriate level, with full public consultation and involvement of users in the planning and implementation of water projects.

Principle No. 3. Women play a central part in the provision, management and safeguarding of water

This pivotal role of women as providers and users of water and guardians of the living environment has seldom been reflected in institutional arrangements for the development and management of water resources. Acceptance and implementation of this principle requires positive policies to address women's specific needs and to equip and empower women to participate at all levels in water resources programmes, including decision-making and implementation, in ways defined by them.

Principle No. 4. Water has an economic value in all its competing uses and should be recognized as an economic good

Within this principle, it is vital to recognize first the basic right of all human beings to have access to clean water and sanitation at an affordable price. Past failure to recognize the economic value of water has led to wasteful and environmentally damaging uses of the resource. Managing water as an economic good is an important way of achieving efficient and equitable use, and of encouraging conservation and protection of water resources.

Acknowledgments
It must be acknowledged that this chapter owes a great deal in both substance and form to the paper written by Robert A Young

in 1996 for the World Bank based on the pioneering work by him and S. Lee Gray in 1972 (1,34).

WATER SUPPLY PLANNING—FEDERAL

David L. Feldman
University of Tennessee
Knoxville, Tennessee

The U.S. government has been engaged in water supply planning for domestic, agricultural, commercial, and industrial needs from its inception. Planning was initially undertaken for inland navigation and flood control. Later, primacy was given to irrigation, municipal needs, and hydroelectricity. The Army Corps of Engineers, the largest and most important federal water supply planning entity, was initially given responsibility for all these efforts. In 1902, the Bureau of Reclamation (BOR) was established and given principal responsibility for water supply projects in the West.

GENERAL PLANNING ISSUES

Federal water supply planning has been characterized by recurring policy debates on four issues:

- the value of *structural* measures (e.g., dams, impoundments) versus *nonstructural* measures such as conservation and demand-side management;
- the primacy to be accorded the *philosophies and missions* of competing agencies (e.g., the Corps, BOR) in the planning process;
- the advantages of *comprehensive, conjoint coordination* by a single entity (e.g., Interior Department, Water Resources Council), facilitated by state cost-sharing; and
- the appropriate role of *public participation*, particularly inclusion of environmental NGOs in plan formulation.

These issues have frequently arisen in discussions over specific plans. In the 1940s, for example, the Pick–Sloan Missouri River project, one of the largest federal water supply planning efforts, was conceived as a compromise between the Corps and BOR, which shared responsibility for water supply in the upper Missouri basin though they had different priorities. The BOR envisioned irrigation as the primary objective of main-stem dam construction, and the Corps gave primacy to flood control and hydropower. When these agencies failed to reconcile priorities, Congress opted to allow each to develop separate plans which would later be integrated by a series of review boards charged with managing and overseeing reservoir operations (1).

HISTORY OF FEDERAL WATER SUPPLY PLANNING EFFORTS

For much of the nineteenth century, water supply planning meant the development by states and communities of engineering-based solutions to meet specific stakeholder objectives such as flood control, municipal supply, or water treatment (2). During the twentieth century planning embraced "multiple objectives" undertaken by *all* levels of government. It is now generally conceived as a comprehensive effort to harmonize environmental, economic, financial, and other concerns (3).

Five distinct periods in the evolution of federal water supply planning are discernible: (1) a formative period consisting of federal intervention in water supply in response to specific regional needs; (2) a federal–state cooperative planning era, following the Civil War, whose focus was mostly navigational improvements; (3) a "multiprogram" era, beginning in the 1930s, characterized by establishment of national priorities; (4) an effort to "federalize" water supply planning through greater presidential-level coordination; and (5) a new federal–state partnership coupled with greater sensitivity to environmental protection and restoration needs.

Formative Developments

The earliest federal water supply planning efforts coincided with exploratory surveys of the intermountain West. Led in many cases by Army engineers, one of these expeditions was intended to identify and trace the origins and suitability for development of water supply sources. The journeys of Lewis and Clark (1804–05) and Zebulon Pike (1806–07) in the upper Missouri and Arkansas River basins, respectively, are the best known of these; Stephen H. Long's Central Rocky Mountain-Yellowstone expedition (1819–20) also surveyed regional water resources.

In 1824, the Corps was officially designated as the steward of the nation's water resources. Between this time and the Civil War, the Corps identified mill sites and needs for navigational improvements and engaged in channel widening and other flood hazard mitigation activities (4,5). These efforts generally produced site-specific solutions to local problems. Despite fierce engineering debates, particularly over proposals to harness the Mississippi River for navigation and flood prevention (6), they failed to set broad, consistent development priorities.

The Advent of Federal–State Cooperative Planning

The origins of a more comprehensive approach to water supply planning may be traced to 1879, when explorer and geographer John Wesley Powell decried the system of Western water law, charging that by encouraging the first users of a river to lay claim to its entire flow (the law of prior appropriation), harm is imposed on other users and the in-stream environment. As first director of the U.S. Geological Survey, Powell proposed dividing the West's arid lands into irrigation districts where each farmer was allotted 80 acres. These districts would be drawn to conform to watershed topography, and drainage basins would serve as state boundaries (7,8).

Powell's proposals discouraged Western migration and drew the ire of Congress and Western land development interests. Ironically, this adverse reaction partly hastened federal efforts to initiate large-scale reclamation projects designed, principally, to support Western agriculture.

These projects would be justified (and subsidized) by their flood control and hydroelectric power benefits—a pattern which would characterize federal water supply planning until the late twentieth century. In exchange for broad political support, a reciprocal system of large-scale, publically financed water projects was eventually provided to Eastern states, principally for navigation, power, and flood abatement (9–13).

In 1879 the Mississippi River Commission (MRC) was also established, one of the first efforts to coordinate federal–state water supply planning efforts in a single basin. MRC was formed to deepen and widen the river's navigable channel and provide flood abatement. Limited efforts to accelerate intergovernmental coordination efforts elsewhere occurred under the presidency of Theodore Roosevelt who "was interested in securing state cooperation and assistance in his program" to develop comprehensive management of the nation's rivers (14). However, these efforts were largely unidirectional in their scope: federal construction agencies that had greater fiscal and technical resources continued to make most decisions. Another goal of Teddy Roosevelt—reducing the power of private interests to exploit the development of water resources for their own gain—was not seriously tackled until the 1930s.

Multiprogram Planning and the Great Depression

The first multipurpose federal water supply planning efforts began when the Tennessee Valley Authority (TVA) was established in 1933. The TVA was granted broad power over a seven-state area to engage in regional water supply planning, flood control, hydropower production, navigation, and economic development. In President Franklin Roosevelt's words, it was to be a "... corporation clothed with the power of government but possessed with the flexibility and initiative of private enterprise" (15). Local entities were incorporated into its economic development programs, and long-term political alliances were built to ensure programmatic and funding stability. Most importantly, TVA "supplanted" other federal water resource agencies in its operating region, and the absence of large private power companies also helped it achieve comprehensive breadth (16).

The political conditions that permitted TVA to do all these things were uniquely characteristic of American politics during the 1930s. Some contend that these conditions may no longer exist, exemplified by the fact that later federal efforts to develop TVA-like entities in the Missouri and Columbia basins failed to gain support in Congress and among local elites (13,15). In short, no TVA-like entity has been replicated anywhere else in the United States.

In other regions, FDR's Interior Secretary Harold Ickes formed basin study groups with formal staffs, planning funds, and some powers of program review during this period. These groups had little formal access to Congress and lacked the ability to initiate new projects or overrule decisions of established water agencies regarding what they often deemed as uneconomical or environmentally damaging policies. They also had no control over planning objectives, which continued to be left to the Corps and other "construction agencies" (17).

A parallel, noteworthy development in this period was the federal interstate compact. The Colorado River Compact, one of the oldest of these, allocates both claimed and unused waters in a seven-state basin based on future needs and water-use priorities. The compact entered into force in 1922, but Arizona, due to resentment over what it considered to be an inequitable division of its share of the lower Colorado, refused to ratify it until 1944. When it did so, it tried to use its ratification to leverage federal funds to build the long-coveted Central Arizona Project, designed to divert Colorado River water to Phoenix, Tucson, and surrounding rural communities via aqueduct (18–20).

Although interstate compacts vary in intent, scope of activity, and structure, all derive their legal authority from the compact clause of the U.S. Constitution (Art. I, Sec. 10, clause 3) which mandates Congress's consent when states enter into binding agreements with one another. This endows them with rights to embrace public as well as private facilities and projects within their jurisdiction. It also binds the actions of signatory states *and* federal agencies and authorizes allocation of interstate waters in accordance with the doctrine of equitable apportionment (i.e., in ways that do not disturb or impair rights awarded to parties under Supreme Court decree without their consent). In effect, their decisions have the force of national law.

During the 1960s, in response to deteriorating water quality and droughts in the Northeast, compact commissions were established for the Delaware (DRBC), Susquehanna (SRBC), and Potomac River (PRBC) basins. These were empowered to allocate interstate waters (including groundwater and interbasin diversions), regulate water quality, and manage interstate bridges and ports. The DRBC, as a forerunner of this model, included numerous federal partners such as Interior Department and Corps of Engineers officials (21,22).

Since 1997, most compacts have continued to function without federal appropriations as a result of Congressional adoption of Heritage Foundation recommendations which urged that Congress "defund" programs whose benefits were viewed as more regional than national in scope. The DRBC and SRBC have sustained themselves without federal appropriations, and Congress has not withdrawn from any compact. Also in 1997, the DRBC's Executive Director advised officials of Alabama, Florida, and Georgia during formulation of two new compacts in that region (the ACF-ACT compacts), and the DRBC and SRBC were explicitly adopted as models by these compacts' framers (21).

"Federalizing" Water Supply Planning: The Water Resources Council and Other Efforts

In the early 1960s, Congress embarked on an effort to provide greater coordination of water supply planning. The Water Resources Council (WRC) was established in 1961 to hasten a more holistic approach to planning by encouraging interagency cooperation and integrating water quality and quantity concerns. It was initially composed of representatives from the Departments of Agriculture, Army, Commerce, and Interior.

In 1965, WRC was empowered by the Water Resources Planning Act (PL 89-80) to engage in comprehensive water supply planning by "maintain(ing) a continuing study of the nation's water and related land resources and to prepare periodic assessments to determine the adequacy of these resources to meet present and future requirements" (23). The Act encouraged states to establish "Title II" river basin commissions charged with integrating and harmonizing state assessments of water supply problems and encouraging regional supply planning. These were staffed by state *and* federal representatives.

The WRC employed competitive planning grants to states and river basin commissions for gathering basic data on water; identifying problems; and developing long-term ("level B") plans for "comprehensive, coordinated management of water resources." The commissions also tried to coordinate the efforts of the Corps of Engineers, BOR, and other federal agencies to provide assistance to states and communities in drawing up these plans. The WRC conducted two national assessments in 1968 and 1978. Both comprehensively depicted the nation's current and projected water needs; analyzed water-related problems from national, regional, and major river basin vantage points; and provided a wealth of data on water quality, quantity, and land use trends and problems (23,24).

As long as the WRC's efforts remained confined to preparing broad sets of recommendations for others to weigh and adopt, they generated little controversy and even less public attention. In the mid-1970s, however, WRC was charged by the Office of Management and Budget with developing "principles and standards" to be applied to evaluation of water projects, as well as cost-share guidance. The WRC was also granted power to discourage projects of dubious economic value or potentially adverse environmental impact. Controversy came to a head during the presidency of Jimmy Carter (1977–1981) who actively sought to diminish Congress's tradition of "pork barrel" water projects, charging that they were wasteful of public expenditures, economically unjustifiable, and in many cases, ecologically harmful. Criticism grew in Congress that WRC was exercising an "anti-dam" agenda through budget-cutting and impoundments to eliminate locally popular, necessary projects (25).

As a result of these criticisms and its eventual demise in 1981 under President Reagan, WRC is regarded as a failure. Critics charge that WRC did not ensure meaningful citizen participation at the inception of federal agency planning; overcome disagreements among states, communities, and other traditional beneficiaries of "pork-barrel" projects; or overcome the perception that its planning efforts were perceived as imposed from above. Moreover, it failed to find "an appropriate ideological basis for policy" (26,27).

Mention should also be made of the U.S. National Water Commission (NWC, 1967–1973), which grew out of controversies in the 1960s surrounding proposals to dam the lower Colorado River. Its chief legacy was the attempt to place fiscal discipline upon federal water supply efforts and to identify ways to better coordinate federal agency water supply planning and pollution control activities (28,29). NWC did succeed in placing species diversity, riverine protection, and other environmental issues on a policy footing equal to traditional economic development drivers of federal water supply planning. Moreover, the NWC's final report provided a "penetrating critique of water resources decision making" (29, pp. 4–24). Its recommendations to improve groundwater management, liberalize in-stream flow protection, urge greater wastewater reuse, and increase agricultural water efficiency remain part of many recent water planning efforts.

The 'New' Paradigm: Federal–State Partnership and Environmental Sensitivity

Since the 1980s, a "new" paradigm for federal water supply planning has evolved that is characterized by greater stakeholder participation in formulating policy alternatives; environmentally sound, socially just water resources management; reliance upon drainage basins as planning units; and greater state–federal cost sharing (30). This paradigm, advocates believe, promotes a management approach more accountable to *varied* interests, amenable to "regionally and locally tailored solutions" (29), and protective of fish and wildlife. Eroding federal interest in water supply, exemplified by Congress' 1997 'zeroing-out' of funding for river basin commissions and a general decline in support for new water projects, has helped legitimize this new paradigm (31–33).

In 1980, Congress passed the *Pacific Northwest Electric Power Planning and Conservation Act* [16 U.S.C. § 839b(h)]. The Act's purpose was to restore salmon spawning runs on the Columbia and Snake Rivers by making salmon a "coequal partner" with hydroelectric power in the operation of the Columbia Basin's 150 + dams (34). The Act also established the Northwest Power Planning Council (NPPC), a multistate, multiagency partnership comprised of the Bonneville Power Authority (BPA), the Pacific Northwest Electric Power and Conservation Council, and governor-appointed representatives from Washington, Oregon, Idaho, and Montana. NPPC has encouraged cooperation among utility systems, conservation and efficient power use, greater involvement of state and local governments in regional power planning, and oversight of electric power system planning and regional fish and wildlife recovery. NPPC is developing a long-range plan to ensure low-cost power for the region while protecting and rebuilding fishery and wildlife populations depleted by hydropower development.

Because NPPC represents a bold step toward adopting this "new" paradigm, its efforts have been heralded as an innovative means of bringing stakeholders together. Its activities have even led to creation of a regional "water budget" designed both to protect fisheries and provide for other services (more than 4 million acre-feet for salmon, seasonally allocated on the Columbia and Snake Rivers), and the set aside of 44,000 river miles of prime fish and wildlife habitat from new hydropower development.

Although it had an enormous effect on regional electric policies and regional approaches to fish and wildlife protection and restoration, NPPC has not prevented several species of Snake River salmon from being listed under the *Endangered Species Act*. By 1990, annual fish runs had been reduced from an estimated 10–16 million to fewer than 3 million, despite more than $100 million in expenditures on environmental improvements at dams and decreased hydropower production. By 1995, many West Coast salmonid species had been driven to extinction, and others remain at risk of extinction. Adding to its problems, in 1994, the Ninth Circuit Court of Appeals rejected NPPC's initial Salmon Strategy and criticized the council's "sacrificing the Act's fish and wildlife goals" (35, p. 190). NPPC has promulgated a new plan to provide for endangered salmon recovery as well as to ensure protection of other fish species.

Although not formally a "compact," NPPC does have compact-like powers to develop plans and meld the operations of state and federal water supply entities. The various water budgets and operational and engineering changes developed on the riverine system are products, in part, of data management tools refined in the light of experience. Eleven state and federal agencies, 13 Native American tribes, eight utilities, and numerous fish, forest, and environmental groups are among the stakeholders involved in restoring the Columbia River system.

Despite its mixed record, NPPC has forced environmental issues onto the agenda of water and power agencies in the Pacific Northwest. As one observer has stated, ecological imperatives, aesthetics, and demands for equity "have given rise to a greater consciousness of the ethical implications of water use" (36). Moreover, unlike previous federal efforts, NPPC has encouraged extensive public involvement and planning for long-term threats by incorporating local community and tribal concerns (35,36). Clearly, NPPC is a major departure from previous federal water supply planning efforts. It may also be a window into the future.

CONCLUSIONS: WHITHER FEDERAL WATER SUPPLY PLANNING?

Federal water supply planning efforts began as attempts to promote support for regional power, reclamation, or flood control projects; or to thwart the influence of private interests who wanted to develop such projects. For most of their history, these efforts limited public involvement to a "consultative" role; failed to provide adequate interagency coordination; and, sometimes fell short of being truly comprehensive.

We now know that effective federal water supply planning requires the ability to overcome agency "turf wars" by pursuing objectives defined by regional stakeholders; supplant authority vested in several agencies (e.g., TVA), and comprehensively manage several water supply needs simultaneously (e.g., DRBC and SRBC have mandates over water quality *and* supply). The record of federal planning further suggests that effective efforts are those perceived as economical and efficient, and thus able to generate public support for projects and programs (37). Finally, they must encourage collaboration among those stakeholders most directly affected by water supply problems. Collaboration overcomes resistance to change, facilitates new opportunities for funding, and stimulates resilient policy ideas (38). Today, public involvement is an essential part of all water supply planning efforts in the United States. As we enter a new millennium, questions that remain to be answered include, how do we determine the objectives of future plans, and what should the federal role be in their implementation?

BIBLIOGRAPHY

1. Feldman, D.L. (1995). *Water Resources Management: In Search of an Environmental Ethic*. Johns Hopkins University Press, Baltimore, MD.
2. Kundell, J.E., DeMeo, T.A., and Myszewski, M. (2001). *Developing a Comprehensive State Water Management Plan—A Framework for Managing Georgia's Water Resources*. Sponsored by Home Depot and Georgia Power. Research Atlanta, Inc./Georgia State University, Andrew Young School of Policy Studies, Atlanta, GA.
3. Grigg, N.S. (1996). *Water Resources Management: Principles, Regulations, and Cases*. McGraw-Hill, New York.
4. Merritt, R.H. (1979). *Creativity, Conflict, and Controversy: A History of the St. Paul District, U.S. Army Corps of Engineers*. Government Printing Office, Washington, DC.
5. Holmes, B.H. (1979). *A History of Federal Water Resources Programs and Policies, 1961–1970*. Economics, Statistics, and Cooperatives Service, U.S. Department of Agriculture, Washington, DC.
6. Barry, J.M. (1998). *A Rising Tide: The Great Mississippi Flood of 1927 and How it Changed America*. Touchstone Books, New York.
7. Powell, J.W. (1879). *Lands of the Arid Region of the United States*. GPO, Washington, DC.
8. McNamee, G. (1994). *Gila: The Life and Death of an American River*. University of New Mexico Press, Albuquerque, NM.
9. Foster, C.H. and Rogers, P. (1988). *Federal Water Policy: Toward an Agenda for Action*. Energy and Environmental Policy Center, JFK School of Government, Harvard University, Cambridge, MA.
10. Wilkinson, C.F. (1992). *Crossing the Next Meridian: Water, Power, and the Future of the American West*. Island Press, Washington, DC.
11. Mann, D.E. (1993). Political science: the past and future of water resources policy and management. In: *Water Resources Administration in the United States: Policy, Practice, and Emerging Issues*. M. Reuss (Ed.). Michigan State University Press, E. Lansing, MI, pp. 55–65.
12. Harden, B. (1994). *A River Lost: The Life and Death of the Columbia*. Macmillan, New York.
13. Pitzer, P.A. (1995). *Grand Coulee: Harnessing a Dream*. Washington State Univ. Press, Pullman, WA.
14. Caldwell, L.K. (1947). Interstate cooperation in river basin development. *Iowa Law Rev.* **32**: 232–243.
15. Conkin, P. (1986). Intellectual and political roots. In: *TVA: Fifty Years of Grass-roots Bureaucracy*. E.C. Hargrove and P.C. Conkin (Eds.). University of Tennessee Press, Knoxville, TN, pp. 3–34.
16. Selznick, P.W. (1966). *TVA and the Grassroots*. Harper and Row, New York.

17. Derthick, M. (1974). *Between State and Nation*. Brookings Institution Press, Washington, DC.
18. Sheridan, T. (1998). The big canal: the political ecology of the central arizona project. In: *Water, Culture, and Power: Local Struggles in a Global Context*. J.M. Donahue and B. R. Johnston (Eds.). Island Press, Washington, DC, pp. 163–186.
19. Grant, D.L. (1998). Introduction to interstate allocation problems. In: *Waters and Water Rights*. R.E. Beck (Editor-in-Chief). 1997 Replacement volumes and Supplemental Pocket Parts. Michie, Charlottesville, VA.
20. *Arizona v. California*. (1963). 373 U.S. 546, 565.
21. *Delaware River Basin Commission Annual Report*. (1997). 1998. DRBC, West Trenton, NJ.
22. Weston, T.R. (1999). *Interstate Water Rights and River Basin Compacts in the East*. Unpublished manuscript, Kirkpatrick and Lockhart, LLP, Harrisburg, PA.
23. (U.S. WRC) U.S. Water Resources Council. (1968). *First National Assessment: The Nation's Water Resources*. U.S. Government Printing Office, Washington, DC.
24. (U.S. WRC) U.S. Water Resources Council. (1978). *The Nation's Water Resources, 1975–2000, Volume 1: Summary, 2nd National Water Assessment*. Water Resources Council, December, Washington, DC.
25. Fradkin, P.L. (1996). *A River No More: The Colorado River and the West*. University of California Press, Berkeley, CA.
26. Robie, R.B. (1980). The impact of federal policy on state planning: a cautionary example. *Ame. Water Works Assoc. J.* **72**(February): 70–73.
27. Goslin, I.V. (1978). Colorado river development. In: *Values and Choices in the Development of the Colorado River Basin*. D.F. Peterson and A.B. Crawford (Eds.). University of Arizona Press, Tucson, AZ.
28. U.S. National Water Commission. (1972). *Water Resources Planning. Consulting Panel on Water Resources Planning*. NWC, Washington, DC.
29. *Water in the West: Challenge for the Next Century*. (1998). Report of the Western Water Policy Review Advisory Commission, Springfield, VA, NTIS, June.
30. Cortner, H.A. and Moote, M.A. (1994). Setting the political agenda: paradigmatic shifts in land and water policy. In: *Environmental Policy and Biodiversity*. R.E. Grumbine (Ed.). Island Press, Washington, DC, pp. 365–377.
31. McGinnis, M.V. (1995). On the verge of collapse: the Columbia River system, wild salmon, and the northwest power planning council. *Natural Resources Journal* **35**: 63–92.
32. Miller, K., Rhodes, S.L., and MacDonnell, L.J. (1996). Global change in microcosm: the case of U.S. Water Institutions. *Policy Sci.* **29**: 271–290.
33. Cody, B.A. (1999). *Western Water Resource Issues*. A Congressional Research Service Brief for Congress. Congressional Research Service, Washington, DC, March 18.
34. Blumm, M. (1998). Columbia river basin. In: *Waters and Water Rights*. Vol. 6. R.E. Beck (Ed.). Michie Publishing, Charlottesville, VA.
35. Getches, D.H. (1998). Water in the Columbia river basin: from a source of performence to an instrument for economic growth. In: *Traditional and Modern Approaches to the Environment on the Pacific Rim: Tensions and Values*. H. Coward (Ed.). State University of New York Press, Albany, NY, pp. 177–193.
36. Volkman, J.M. (1996). *A River in Common: The Columbia River, the Salmon Ecosystem, and Water Quality*. Western Water Policy Review Advisory Commission, Denver, CO.
37. American Society of Civil Engineers. (1984). *Environmental Objectives in Water Resources Planning and Management*. Water Resources Planning and Management Division. ASCE, New York.
38. Chrislip, D.D. and Larson, C.E. (1994). *Collaborative Leadership: How Citizens and Civic Leaders Can Make A Difference*. Jossey-Bass, San Francisco, CA.

FLOOD CONTROL ACT OF 1944

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

The Flood Control Act of 1944, enacted on December 22 of that year, established federal policy regarding jurisdiction over rivers and other bodies of water, specifically watersheds. The act states that all water-related construction and navigation projects will be approved only if they substantially improve navigational capacity and do not excessively interfere with standard water usage. The plan was proposed to prevent flooding and to allow more irrigation, easier navigation, increased river-related recreation, and the production of hydroelectric power. The Army Corp of Engineers was designated to oversee these tasks.

A major component of the Flood Control Act of 1944 is known as the Pick–Sloan plan. Named after its two creators, Glenn Sloan and Lewis Pick, members of the Army Corps of Engineers, this plan authorized the construction of six dams in three states on the Missouri River. The plan created the largest reservoir in North America. Prior to the 1940s, national water navigation and flood control projects were run by four different agencies that had little or no communication and planning among them: the Army Corps of Engineers, the Federal Power Commission, the Reclamation Bureau, and the Soil Conservation Service. As a result of massive and catastrophic flooding that occurred in 1936, all this changed. In March of that year, flooding killed 107 people and caused $270,000,000 of damage. The United States had never before had very many flood control projects; after these floods, it became clear that this needed to change. The federal government responded by passing the 1936 Flood Control Act and 8 years later, the Flood Control Act of 1944.

The Flood Control Act has 15 sections. The first section is composed of rules for water-related planning submissions to Congress, specifically, that they must not interfere with established water uses. These include domestic and municipal use, as well as irrigation and industrial uses. It also asserts the importance of state involvement in projects and plans that will affect them, as determined by the Chief Engineer of the Army Corps.

The second section explains that flood control includes channel and drainage improvements. This section also states that river improvement for flood control is under the War Department (now the Department of Defense) as directed by the Secretary of War (now the Secretary of Defense), and supervised by the Chief Engineer of the Army Corps. Watershed alteration, erosion prevention, and water flow control projects, on the other hand, are stated to be under the control of the Department of Agriculture, as directed by the Secretary of Agriculture. These jurisdictions are subject to change only by Acts of Congress.

The third section deals with the 1936 Flood Control Act. It says that section 3 of that Act still applies to all projects. It also allows a 5-year period in which local interests working on water-related projects can secure whatever local cooperation the Secretary of War deems necessary.

Section 4 authorizes the Chief Engineer of the Army Corps to operate and maintain reservoir facilities in the United States. The Secretary of War is authorized to lease reservoir land but must give preference to federal, state, and local government. The Secretary must also make reservoir areas available to the general public for recreational purposes, as adjudged consistent with the public interest.

Section 5 instructs the Secretary of War to deliver to the Secretary of the Interior all energy produced at reservoirs that is not required for their operation. The Secretary of the Interior is instructed to dispense this energy to the public at low rates and is allowed to build or acquire transmission lines and other facilities with money from Congress only if necessary.

The sixth section of the Act authorizes the Secretary of War to sell surplus reservoir water for reasonable prices, so long as this does not negatively affect standard water uses.

Section 7 instructs the Secretary of War to regulate the usage of reservoir storage intended for navigation and flood control. This section does not generally apply to the Tennessee Valley Authority.

Section 8 states that the Secretary of War may, on advisement of the Secretary of the Interior, allow the use of water for irrigation. The Secretary of the Interior is then authorized to construct and maintain whatever structures are necessary for irrigation. However, the Secretary is allowed to undertake such projects only after receiving Congressional authorization. This section does not apply to dams and reservoirs built for irrigation prior to 1944 by the Army Corps of Engineers.

Section 9 of the Flood Control Act of 1944 refers to specific flood control projects and allocates $200,000,000 for those projects; specifically, the Missouri River project.

The tenth section details many other projects which are authorized by the Act and affect the following areas: the Lake Champlain Basin, the Blackstone River Basin, the Connecticut River Basin, the Thames River Basin, the Housatonic River Basin, the Susquehanna River Basin, the Roanoke River Basin, the Edisto River Basin, the Savannah River Basin, the Mobile River Basin, the Alabama-Coosa River Basin, the Lower Mississippi River, the Red-Ouachita River Basin, the Arkansas River Basin, the White River Basin, the Upper Mississippi River Basin, the Red River of the North Basin, the Missouri River Basin, the Ohio River Basin, the Great Lakes Basin, the Colorado River Basin (Texas), the Brazos River Basin, the Rio Grande Basin, the Great Salt Basin, the Colorado River Basin, the San Diego River Basin, the Ventura River Basin, the Santa Ana River Basin, the Los Angeles-San Gabriel Basin and Ballona Creek, the Pajaro River Basin, the Sacramento-San Joaquin River Basin, the San Joaquin River, the Napa River Basin, the Coquille River Basin, the Nehalem River Basin, the Willamette River Basin, the Columbia River Basin, the Willapa River Basin, the Chehalis River Basin, and the Territory of Hawaii.

Section 11 lists many more areas for which the Secretary of War is instructed to oversee preliminary studies of flood control and other plans. These plans are not to be undertaken until authorized by law.

Section 12 appropriates $950,000,000 for all of these projects. It also appropriates $10,000,000 to the Departments of Agriculture and War for any preliminary studies of flood control improvements, $1,500,000 to the Federal Power Commission for these studies, and a $500,000 emergency fund for emergency flood prevention.

Section 13 authorized various projects to reduce soil erosion and to improve water flow and runoff in these areas: the Los Angeles River Basin, the Santa Ynez River Basin, the Trinity River Basin (Texas), the Little Tallahatchie River Watershed, the Yazoo River Watershed, the Coosa River Watershed, the Little Sioux River Watershed, the Potomac River Watershed, the Buffalo Creek Watershed, the Colorado River Watershed (Texas), and the Washita River Watershed.

Section 14 of the Act deals with the balance of some $10,000,000 appropriated by an earlier act. It is to be used for the work detailed in section 13 of the Flood Control Act of 1944, and no more than 20% is to be spent on any one project.

The 15th and final section of the Act amends an earlier act approved in 1938. It allows the Secretary of War to undertake emergency measures to avert dangerous, destructive, and/or fatal flooding, based on codicil that he may not spend more than $100,000 in any fiscal year.

This Act permanently altered the manner in which water issues are handled in the United States and has saved billions of dollars in property damage by averting flooding.

READING LIST

Flood Control Act of 1944.

http://lwas.fws.gov/lawsdigest/flood.html.

http://www.usace.army.mil/inet/usace-docs/eng-pamphlets/ep870-1-29/c-6.pdf.

GREAT LAKES GOVERNORS' AGREEMENT

Frederic Lasserre
Laval University
Quebec City, Quebec, Canada

Several institutions govern the use of the Great Lakes water and try to coordinate the actions undertaken by national governments (the United States and Canada) as well as state and provincial actions. Some of these institutions are old; some are much more recent and stem from political concerns regarding the quality of the water in the Great Lakes, as well as potential diversions that would have taken large quantities out of the basin.

THE GREAT LAKES GOVERNANCE

The International Joint Commission (IJC) prevents and resolves disputes between the United States of America and Canada under the 1909 Boundary Waters Treaty and "pursues the common good" of both countries as an "independent and objective advisor" to the two governments (1). Transportation was one of the main concerns of water management at the time: there was no water quantity problem, nor was there much talk of large variations in Great Lakes levels that could hamper human or industrial consumption. The negotiators wanted to ensure that water would always be sufficient to allow for the safe passage of ships through locks. Article III of the Treaty states that "no further or other uses or obstructions or diversions, whether temporary or permanent, of boundary waters on either side of the line, affecting the natural level or flow of boundary waters on the other side of the line shall be made, except by authority of the United States or the Dominion of Canada within their respective jurisdictions, and with the approval, as hereinafter provided, of a joint commission, to be known as the International Joint Commission." The IJC is, therefore, as its name underlines, an international institution between two sovereign States.

In particular, the Commission rules upon applications for approval of projects affecting boundary or transboundary waters and may regulate the operation of these projects; it assists the two countries in protecting the transboundary environment, including the implementation of the Great Lakes Water Quality Agreement (1972, renewed in 1978) and improving transboundary air quality, and it alerts the governments to emerging issues along the boundary that may give rise to bilateral disputes. The IJC, in particular, in its *Final Report on Protection of the Waters of the Great Lakes* (2000), recommended that Canadian and U.S. federal, provincial, and state governments should not permit diversion of water from the Great Lakes Basin unless the proponent can demonstrate that the removal would not endanger the integrity of the Great Lakes ecosystem (2).

The Great Lakes Commission (GLC) is the only regional organization whose statutory mandate is to represent the eight Great Lakes states on a variety of environmental and economic issues. It was created in 1955 and was granted Congressional consent in 1968. The Great Lakes Commission is now a binational agency: its members include the eight Great Lakes states—Illinois, Indiana, Michigan, Minnesota, New York, Ohio, Pennsylvania, and Wisconsin; the Canadian provinces of Ontario and Québec have had associate member status since 1999.

The Commission is dedicated to managing the waterbody of the Great Lakes: it promotes the orderly, integrated and comprehensive development, use, and conservation of the water and related natural resources of the Great Lakes Basin and St. Lawrence River. Each jurisdiction appoints a delegation of three to five members comprised of senior agency officials, legislators, or appointees of the governor or premier.

The Commission works actively at the member states' civil servant level. It is responsible for sharing information among members, coordinating state and provincial positions on issues of regional concern, and advocating those positions on which the members agree (3).

The GLC was set up by the member states as the St. Lawrence Seaway construction was well under way. The concern of the states was, at the time, to coordinate their views on ways to manage the economic, ecological, and social impacts of the increasing commercial traffic that would result from opening the Seaway in 1959.

The Council of Great Lakes Governors (CGLG), created in 1983, is a partnership of the Governors of the eight Great Lakes states. It has no legal basis. It is a pressure group that coordinates actions among the member states. It is, therefore, an institution separate from the Great Lakes Commission. Besides, the GLC is more technical, the CGLG more political in its aims and functioning. In recent years, the Canadian premiers of Ontario and Quebec have joined with the Council of Governors in sharing their views on the Great Lakes region.

The CGLG is not limited to Great Lakes water management. For instance, the CGLG coordinated negotiations among member states. The resulting 1988 Economic Development Agreement provides for coordinating the commercial promotion and trade missions of the member states abroad.

The central issue of common concern to all of the Governors was protecting and managing the waters of the Great Lakes. As the primary source of freshwater for the region as well as a trade route that connects the region to the world through the St. Lawrence Seaway, proper management of the lakes is crucial to the health of the residents of the Great Lakes region as well as to the continued growth of the region's economy. Consequently, coordinating Great Lakes regional water policy was the original impetus for creating the Council (4).

A MAJOR IMPETUS: PREVENTING DIVERSION FROM THE GREAT LAKES

The Advent of the Great Lakes Charter

Water pollution was one concern, and the Governors tackled the issue of improving Great Lakes water quality. In 1986, the Governors signed the Toxic Substances Control Agreement (TSCA) that provides for developing programs for companies to adopt pollution-prevention devices.

Diversion control and basin water integrity were another major concern. Governmental archives from the early 1980s attest to western United States lobbying for diversion of Great Lakes water to quench their lack of water. Great Lakes states wanted to resist these projects, both for environmental reasons as well as for political reasons: why would the Great Lakes states give to California an added value at a time when so many firms were leaving the area and moved away to the West Coast (5). The Great Lakes Charter stemmed from this growing concern that Great Lakes water could be diverted to water-scarce regions of the United States, the Midwest, and the West especially. States and provinces from the Great Lakes Basin had discussed their fears of diversion schemes at Mackinac Island (Michigan) in 1982 and at Indianapolis (Indiana) in 1983. The Great Lakes Charter, signed in 1985 by the eight Great Lakes Governors and the premiers of Ontario and Quebec, is the final resolution of a process set up after the Indianapolis meeting. The Charter created a notice and consultation process for Great Lakes diversions. The signatories agreed that no Great Lakes state or province would proceed with any new or increased diversion or consumptive use of Great Lakes water of more than five million gallons per day without notifying, consulting, and seeking the consent of all affected Great Lakes states and provinces. It also calls for coordinating water policies among member states and exchanging hydrologic data.

The Great Lakes states also lobbied the federal government that was in the process of designing a new law governing water use in the United States. The subsequent Water Resources Development Act of 1986 (WRDA) prohibits "any diversion of Great Lakes water by any State, federal agency, or private entity for use outside the Great Lakes basin unless such diversion is approved by the Governor of each of the Great Lakes States." It also prohibits any Federal agency from studying the transfer of Great Lakes water for use outside the Great Lakes Basin, unless done under the auspices of the IJC. To ensure compliance with a 1967 Supreme Court (modified in 1980) consent decree, the Act also appropriates federal resources to monitor and measure Lake Michigan's water flow into the Chicago River.

The Council coordinates the authority granted to the Governors under the Water Resources Development Act of 1986 (WRDA). The Act requires the Governors' unanimous approval on any proposed out-of-basin diversion of water from the Great Lakes Basin.

In 1997, the Great Lakes states entered into a Memorandum of Understanding on the Lake Michigan diversion, concluding a lengthy mediation process on the matter of Illinois' diversion of Lake Michigan water at Chicago. Under the U.S. Supreme Court decrees (1967 and 1980), Illinois was limited to 3,200 cfs each year. Illinois had exceeded that limit by nearly 15%. Illinois will further reduce its annual diversion during the following 14 years to restore to Lake Michigan the excess amount of water it has withdrawn since 1980 and construct new lakefront structures that do not allow leakage. All eight Great Lakes states and the U.S. federal government participated in the discussions. The Province of Ontario, as well as the City of Chicago and the Army Corps of Engineers, were also involved as observers or by providing technical support.

Diversions Are Still Feared: Banning Water Exports?

Diversion projects are far from being outmoded, despite the conclusion set forth by many observers that bulk water is, for now, too expensive to move over large distances. Large and small diversion schemes are still pushed by pressure groups. For instance, in the spring of 1998, the Province of Ontario approved a permit for the Nova Group of Sault Saint Marie Ontario that would have allowed the Nova Group to remove 60 million gallons per year from Lake Superior that would then be sold in Asia. In direct response to concerns expressed by the Great Lakes Governors via the Council, the Ontario government rescinded the permit. The accompanying public uproar led to the previously mentioned review of the issue of bulk removals by the International Joint Commission (IJC) in 2000. The Governors, through the Council, were represented on the study team that advised the IJC on these issues. Again in November 2002, Milwaukee mayor, John Norquist, denounced a diversion project to bring water to far western suburbs, outside the Great Lakes Basin.

So as to address a growing public concern that large water diversion schemes could someday be implemented, despite the Charter and the WRDA, the Council of Great Lakes Governors convened in 1999 and agreed there was a political need to reinforce the provisions of the Charter by

- forging a new, more binding agreement to manage the Great Lakes waters;
- creating a new standard requiring an improvement to the water and water-dependent natural resources of the Great Lakes before allowing new water uses;
- pledging to obtain better information so that the water is managed rationally; and
- lowering the trigger level for diversions that need to be approved under the Charter.

The Great Lakes Governors and the premiers of Ontario and Quebec signed the Great Lakes Charter Annex on June 18, 2001 in Niagara Falls, New York. The Annex, an amendment to the Great Lakes Charter of 1985, provides a blueprint for creating a new set of binding agreements among the states and provinces on the effective management of Great Lakes water resources. It also outlines a series of principles for reviewing water withdrawals from the Great Lakes Basin that is grounded in protecting, conserving, restoring, and improving the Great Lakes ecosystem. The new agreement must be set up by the member States before 2004 and then ratified at the two national levels of Canada and the United States to prevent any attempt to divert water from the Great Lakes Basin. A particular emphasis was placed in the Annex on the need to close loopholes that legal experts claim have arisen for water exports under the North American Free Trade Agreement and the General Agreement on Trade and Tariffs.

The history of the Great Lakes agreements attests to the evolution of the concerns regarding the largest body of freshwater. From a very transportation-oriented perspective, as embodied in the 1909 Boundary Waters Treaty, the focus evolved to pondering the possible impacts of the opening of the St. Lawrence Seaway in 1955, by the Great Lakes Commission, to emphasis placed on water quality and quantity, by the Great Lakes Charter of 1985. The advocacy of a major water diversion by western states pressure groups and recent strong variations in the Great Lakes levels stirred an important public controversy that led the Great Lakes Basin state and province leaders to sign the Annex in 2001 to ban all large water diversion out of the basin, thus implementing recommendations already made by the IJC in 2000.

Acknowledgment
This research was made possible thanks to a grant from the FQRSC.

BIBLIOGRAPHY

1. IJC Status. Available: http://www.ijc.org.
2. Lasserre, F. (2002). La guerre de l'eau aura-t-elle lieu? In: *Espaces et enjeux: Méthodes d'une Géopolitique Critique*. F. Lasserre and E. Gonon (Eds.). Collection Chaire Raoul Dandurand/UQÀM, L'Harmattan, Montréal and Paris, p. 251; Lasserre, F. (2002). L'Amérique a soif. Les États-Unis obligeront-ils Ottawa à céder l'eau du Canada? In: *Eaux et Territoires: Tensions, Coopérations et Géopolitique de l'eau*. F. Lasserre and L. Descroix (Eds.). Presses de l'Université du Québec/L'Harmattan, Quebec City/Paris, p. 315.
3. Great Lakes Commission. Available: http://www.glc.org/about/.
4. Council of Great Lakes Governors. Available: www.cglg.com.
5. Lasserre, F. (2001). L'Amérique a soif. Les besoins en eau de l'Ouest des États-Unis conduiront-ils Ottawa à céder l'eau du Canada? *Revue internationale d'Études canadiennes* **24**: 196–214.

U.S./CANADIAN BOUNDARY WATERS TREATY AND THE GREAT LAKES WATER QUALITY AGREEMENT

David W. Yoskowitz
Texas A&M University—
Corpus Christi
Laredo, Texas

The United States and Canada established the International Joint Commission (IJC) to take on certain responsibilities for transboundary bodies of water. The United States also has a similar arrangement with Mexico known as the International Boundary Water Commission (IBWC). Note that the task of the IBWC by comparison is not as involved because it has only two rivers in its jurisdiction, whereas the IJC has more than 200 transborder rivers (1), plus lakes and common aquifers.

FORM

The IJC was created as a result of the Boundary Waters Treaty of 1909 between the United States and Great Britain because it still controlled Canadian foreign policy. The pressure for a treaty came from within as populations of communities on both sides of the border began to grow and demands on water increased along with it. Initially, obstacles had to be overcome in the legal wording of the document because the two countries look at boundary waters differently. Americans tend to look at boundary water resources from the point of view of equity and equitable utilization, in which rights and obligations of the individual water user are recognized. Canadians, on the other hand, see water issues in terms of equality and the rights and responsibilities of the two states.

Institutionally, it was the early irrigation congresses and conferences of the 1890s that ultimately promoted movement toward considering bilateral agreements and treaties and, as a result, the creation of the IJC (2).

The treaty begins with assurance that navigation along common bodies of water shall remain open, and no country may impose excessive tools on the other for use of canals, locks, etc. Until power production and irrigation became important uses of water near the turn of the century, international water law had focused on navigation rights. A strong impetus for a treaty was to secure navigation rights, although it was less important during negotiation than other boundary water issues (3).

The first time the IJC is mentioned in the Treaty Relating to Boundary Water Between the United States and Canada is in Article III with regard to the authority to approve/disapprove obstructions or diversions of boundary water that affect the natural level or flow of the water on the other side of the line, but it is not until Article VII that establishes the IJC with six commissioners, three from each side, appointed by their respective governments. Its jurisdiction covers all cases involving use, obstruction, or diversion for which approval is required as mentioned in the treaty (3). A majority is required to make a decision, and, if evenly split, separate reports will be made to the respective governments and settled through diplomatic channels (4).

Each country was assured equal and similar rights in the use of the water, but there was an order of precedence:

1. uses for domestic and sanitary purposes; then
2. uses for navigation, including the service of canals for navigation; then
3. uses for power and irrigation.

These reflect the interests of the time and thus are dated because they do not reflect more current interests in fisheries, recreation, wildlife, and the environment in general (3). What makes the IJC important are the investigative powers that were accorded to them in Article IX, which states that, "The IJC is authorized in each case so referred to examine into and report upon the facts and circumstances of the particular questions referred" (4). Reports issued by the IJC are not to be considered decisions, as was made very clear in the treaty, and therefore the governments could ultimately ignore the IJC as they saw fit.

FUNCTION

The Boundary Water Treaty put the IJC into action using some broad guidelines but has changed since its inception at the beginning of the century. Along the northern border, the questions have moved from quantity to quality, and the IJC has a leading role in the answers. The governments have delegated new responsibilities to the IJC based on its track record and reputation for impartiality (3). An existing problem is that the commission has no follow-up implementation authority, which gives it largely an advisory role.

Several categories of management issues/problems dominate the work that the IJC becomes involved in, including upstream–downstream conflicts, common pool resources, and integrated river basin developments (5). Work on these issues can take the form of either an approval or reference. Approval is required if an application is made that falls under the scope of either Article III or IV of the Treaty (3), and usually consists of constructing dams, hydropower projects, or diverting flows. The IJC appoints a technical advisory board of members from each country to conduct the investigative work on whether to approve/disapprove a project. References, on the other hand, take much more time and involve constant monitoring of whatever issue has been referred to the commission. Advisory boards are set up in much the same way as they are in the approval category (3).

The strength of the commission lies in several areas where it could become an instrument for boundary environmental protection and management: (a) The IJC performs administrative and quasi-judicial tasks handling a range of issues; (b) There is experience and legitimacy in dealing with boundary issues on an impartial basis; (c) The IJC is an arbitrator of fact; (d) It is a mediator of policy; (e) Top water managers in each country meet regularly under the IJC umbrella, which has created an informal network of contact among government officials and experts (6).

GREAT LAKES WATER QUALITY AGREEMENT AND THE IJC

The Agreement, first signed in 1972 and renewed in 1978, expresses the commitment of each country to restore and maintain the chemical, physical, and biological integrity of the Great Lakes Basin ecosystem and includes a number of objectives and guidelines to achieve these goals. It reaffirms the rights and obligation of Canada and the United States under the Boundary Waters Treaty.

In 1987, a Protocol was signed amending the 1978 Agreement. The amendments aim to strengthen the programs, practices, and technology described in the 1978 Agreement and to increase accountability for implementing them. Timetables are set for implementing specific programs.

The parties will meet biennially to discuss progress and report periodically to the Commission. New annexes address atmospheric deposition of toxic pollutants, contaminated sediments, groundwater, and nonpoint sources of pollution. Annexes are also added to incorporate the development and implementation of remedial action plans for areas of concern and lakewide management plans to control critical pollutants.

The IJC monitors and assesses progress under the Agreement and advises each government on matters related to the quality of the boundary waters of the Great Lakes system. The Agreement also calls on the IJC to assist the United States and Canada with joint programs under the Agreement and provides for two binational boards—the Great Lakes Water Quality Board and the Great Lakes Science Advisory Board—to advise the Commission.

The general objectives of the agreement, expressed in Article III of the Agreement (7), are to keep the waters:

(a) free from substances that directly or indirectly enter the waters as a result of human activity and that will settle to form putrescent or otherwise objectionable sludge deposits, or that will adversely affect aquatic life or waterfowl;
(b) free from floating materials such as debris, oil, scum, and other immiscible substances resulting from human activities in amounts that are unsightly or deleterious;
(c) free from materials and heat directly or indirectly entering the water as a result of human activity that alone, or in combination with other materials, will produce color, odor, taste, or other conditions in such a degree as to interfere with beneficial uses;
(d) free from materials and heat directly or indirectly entering the water as a result of human activity that alone, or in combination with other materials, will produce conditions that are toxic or harmful to human, animal, or aquatic life; and
(e) free from nutrients directly or indirectly entering the waters as a result of human activity in amounts that create growths of aquatic life that interfere with beneficial uses.

BIBLIOGRAPHY

1. Allee, D.J. (1993). Subnational governance and the international joint—commission: local management of United States and Canadian boundary water. *Natl. Resour. J.* **33**: 133–151.
2. Carroll, J.E. (1986). Water resources management as an issues in environmental diplomacy. *Natl. Resour. J.* **26**: 207–242.
3. Lemarquand, D. (1993). The international joint commission and changing Canada—United States boundary relations. *Natl. Resour. J.* **33**: 59–91.
4. *Treaty Relating to Boundary Water Between the United States and Canada, January 11, 1909*. U.S.—Great Britain, 36 Stat, 2449.
5. Sadler, B. (1993). Shared resources, common future, sustainable management of Canada–United States border waters. *Natl. Resour. J.* **33**: 375–396.
6. Lemarquand, D. (1986). Preconditions to cooperation in Canada–United States boundary waters. *Natl. Resour. J.* **26**: 361–380.
7. *Great Lakes Water Quality Agreement of 1978 Agreement, with Annexes and Terms of Reference, between the United States and Canada signed at Ottawa November 22, 1978 and Phosphorus Load Reduction Supplement signed October 16*, 1983 as amended by Protocol signed November 18, 1987.

GREAT LAKES WATER QUALITY INITIATIVE

ELSIE F. MILLANO
ERM, Inc.
Vernon Hills, Illinois

GENERAL DESCRIPTION

The Great Lakes Water Quality Initiative (the “Initiative”) is a program to improve the quality of the waters in the

Great Lakes system. The Great Lakes consist of (1) Lake Ontario, Lake Erie, Lake Huron (including Lake Saint Clair), Lake Michigan, and Lake Superior; and (2) their connecting channels, including the Saint Mary's River, the Saint Clair River, the Detroit River, the Niagara River, and the Saint Lawrence River to the Canadian border. The Great Lakes system includes all streams, rivers, lakes, and other bodies of water within the drainage basin of the Great Lakes within the United States. The effort was undertaken by the U.S. Environmental Protection Agency (U.S. EPA) and the eight Great Lakes U.S. states (Illinois, Indiana, Michigan, Minnesota, Ohio, Pennsylvania, New York, and Wisconsin), supported by the U.S. Fish and Wildlife Service, the National Park Service, environmental groups, municipalities, industry, and academia.

The main objective of the Initiative is to reduce the discharge of bioaccumulative compounds of concern (BCCs) into the Great Lakes system. The first phase of the Initiative consisted of the development of a consistent set of water quality criteria to be applied throughout the Great Lakes system by the Great Lakes States and by any Indian Tribes that have jurisdiction over portions of the Great Lakes system.

The Initiative participants collected available scientific information on persistent toxic chemicals in the Great Lakes system, reviewed and evaluated the data and the different regulations and procedures to establish water quality criteria of the participant states and developed the water quality criteria and procedures issued in 1995 as the Water Quality Guidance for the Great Lakes System (the "Guidance").

The second phase of the Initiative, called the Great Lakes Toxic Reduction Effort (GLTRE), is under development. The GLTRE focuses on reducing BCCs contributed to the Great Lakes system by nonpoint sources.

HISTORY

In the spring of 1989, the Council of Great Lakes Governors unanimously agreed to participate in the Initiative with the U.S. EPA, to establish consistent water quality criteria throughout the Great Lakes system. The Initiative was a response by the United States to continuous contamination of the Great Lakes system by persistent and bioaccumulative toxic (PBT) pollutants (such as polychlorinated biphenyls, dioxins, DDT, DDE, and mercury), despite efforts to bring the Great Lakes system back into health that date back to 1972, when Canada and the United States signed the Great Lakes Water Quality Agreement (GLWQA).

At the time of the 1972 GLWQA, algae growth and reduced dissolved oxygen concentration were reducing the fish population in the Great Lakes system. The 1972 GLWQA specified (1) objectives to reduce the discharge of pollutants that caused nuisance conditions (color, odor, floating oil and scum, aquatic weeds, and algae) and of substances toxic to human, animal, or aquatic life; and (2) numerical targets for reducing loadings of phosphorus into the lower Great Lakes (i.e., Lake Erie and Lake Ontario). The 1972 GLWQA also contained measures to reduce the loading of phosphorus, including effluent limits on all major municipal sewage treatment facilities, limitations on the phosphorus content in household detergents, and reduction in nonpoint source runoff loadings.

During the 1970s, the passage of the Clean Water Act in the United States and similar efforts in Canada reduced the discharge of untreated municipal and industrial wastewaters and provided funding for further evaluating the effects of pollutants on the Great Lakes system. The additional research indicated that PBT substances were responsible for the declining population and accumulation of pollutants in several species of animals, including bald eagles, minks, herring gulls, double-crested cormorants, common and Caspian terns, ospreys, and lake trout among others.

As a result, the 1978 revisions to the GLWQA established a policy calling for virtual elimination of the discharge of PBT substances. In addition, the 1978 GLWQA specified modified phosphorus target loadings for the lower Great Lakes and new phosphorus target loadings for the upper Great Lakes.

In the 1980s, scientists detected effects of PBT substances on humans, consisting of neurological effects on infants whose mothers had eaten Lake Michigan fish contaminated with PBT substances during their pregnancies. In 1987, Canada and the United States amended the GLWQA by Protocol to require the development of (1) Lakewide Management Plans (LaMPs) to control and reduce PBT substances and (2) Remedial Action Plans (RAPs) for 43 recognized local areas of concern in the lakes (now 42, one of them was delisted as such an area in 1994) affected by PBT substances. The 1987 amendments called for greater coordination among different jurisdictional authorities to allow the success of the LaMPs and RAPs, plus more focused attention on PBT substance pollution from land runoff, contaminated sediment, airborne toxic substances, and contaminated groundwater.

As previously indicated, one of the objectives of the Initiative was to develop water quality guidance for BCCs in the Great Lakes system to ensure consistency in requirements among the Great Lakes states. In 1990, the U.S. Congress approved an amendment to the Clean Water Act that required the U.S. EPA to publish water quality guidance for the Great Lakes system in accordance with the 1987 GLWQA and to perform other key activities in the 1987 GLWQA, imposed statutory deadlines for implementing the water quality guidance once issued, and provided funding for programs to meet the 1987 GLWQA requirements. The Water Quality Guidance for the Great Lakes System (the "Guidance") was issued by the U.S. EPA in March 1995 and supplemented in November 2000. The Guidance uses the term BCC to refer to PBT substances. The GLTRE is currently focusing on the other aspects of the 1987 GLWQA.

WATER QUALITY GUIDANCE FOR THE GREAT LAKES SYSTEM

Under United States regulations, each state or Indian tribe has the right to set water quality criteria for the waters under their jurisdiction. These criteria are, in turn, used by each state or Indian tribe to develop water quality standard-based discharge limits for wastewater effluents under the National Pollutant Discharge Elimination System (NPDES). The issuance of the Guidance ensures that all point source discharges to the Great Lakes system will meet a set of consistently developed criteria.

General Description

The proposed regulation was issued in 1993. The final Guidance, which was issued in March 1995 and then supplemented in November 2000, contains

- water quality criteria for a set of 29 chemical constituents to protect aquatic life, human health, and/or wildlife;
- procedures to develop water quality criteria for additional pollutants;
- an antidegradation policy to be adopted, at a minimum, for BCCs;
- a ban on mixing zones for BCCs; and
- procedures to implement the water quality criteria guidance.

Mixing zones, which are established by each state on a point-source basis, allow discharge of compounds at concentrations higher than the surface waterbody standard as long as the concentration in the water near or outside the boundary of the mixing zone meets the water quality criteria.

The March 1995 Guidance contained all five elements listed above, but the ban on mixing zones was vacated in June 1997 by the U.S. Court of Appeals in the District of Columbia and reinstated in November 2000 with modifications. The final Guidance is codified at Part 132 of Title 40 of the Code of Federal Regulations, and includes the six tables and the six appendixes shown in Table 1. The implementation procedures included in Appendix F are listed in Table 2.

Table 1. Attachments to the Water Quality Initiative Regulation

Attachment	Description
Table 1	Acute Water Quality Criteria for Protection of Aquatic Life in Ambient Water
Table 2	Chronic Water Quality Criteria for Protection of Aquatic Life in Ambient Water
Table 3	Water Quality Criteria for Protection of Human Health
Table 4	Water Quality Criteria for Protection of Wildlife
Table 5	Pollutants Subject to Federal, State, and Tribal Requirements
Table 6	Pollutants of Initial Focus in the Great Lakes Water Quality Initiative
Appendix A	Methodologies for Development of Aquatic Life Criteria and Values
Appendix B	Methodology for Deriving Bioaccumulation Factors
Appendix C	Methodologies for Development of Human Health Criteria and Values
Appendix D	Methodology for the Development of Wildlife Criteria
Appendix E	Antidegradation Policy
Appendix F	Implementation Procedures

The Guidance required the Great Lakes states to adopt and submit to the U.S. EPA for review, within two years of publication of the final Guidance (i.e., March 23, 1997), minimum water quality standards, antidegradation policies, and implementation procedure regulations in compliance with the corresponding Guidance procedures. To date, all states have submitted their proposed regulations and obtained either approval or partial approval of their submittals. When an element of a state's submittal was inconsistent with the Guidance, the U.S. EPA disapproved that element of the submittal and established that the specific state has to meet the Guidance's requirements for that specific element.

Water Quality Criteria

Table 3 presents the 29 chemical constituents for which water quality criteria were issued, along with an indication of whether each compound is a BCC and what population is covered by the water quality criteria. The Guidance includes water quality criteria for

- protecting aquatic life on an acute and chronic basis;
- protecting human health from drinking water and eating fish, or from recreationally using the Great Lakes waters; and
- protecting wildlife against ingestion of aquatic life containing the indicated chemical constituents.

Tables 4 through 7 reproduce the Guidance's water quality criteria, and Table 8 presents the pollutants of initial focus in the Guidance. Only chemical constituents for which sufficient information was available to determine a water quality standard based on requirements included in the Guidance appear in Tables 4 through 7; those chemical constituents include only nine of the 22 BCCs shown in Section A of Table 8.

Table 2. List of Implementation Procedures

Procedure #	Description
1	Site-Specific Modification to Criteria and Values
2	Variances from Water Quality Standards for Point Sources
3	Total Maximum Daily Loads, Wasteload Allocations for Point Sources, Load Allocations for Nonpoint Sources, Wasteload Allocations in the Absence of a TMDL, and Preliminary Wasteload Allocations for Purposes of Determining the Need for Water Quality Based Effluent Limits
4	Additivity
5	Reasonable Potential to Exceed Water Quality Standards
6	Whole Effluent Toxicity Requirements
7	Loading Limits
8	Water Quality-Based Effluent Limitations Below the Quantification Level
9	Compliance Schedules

Table 3. Summary of the Parameters for which a Water Quality Criterion/Standard is Presented in Part 132

Parameter	BCC?	Acute Aquatic Life	Chronic Aquatic Life	Human Health	Wildlife
Arsenic (III)		XX	XX		
Benzene				XX	
Cadmium		XX	XX		
Chlordane	Yes			XX	
Chlorobenzene				XX	
Chromium (III)		XX	XX		
Chromium (VI)		XX	XX		
Copper		XX	XX		
Cyanide		XX	XX		
Cyanides				XX	
DDT				XX	
DDT and metabolites	Yes				XX
Dieldrin	Yes	XX	XX	XX	
2,4-Dimethylphenol				XX	
2,4-Dinitrophenol				XX	
Endrin		XX	XX		
Hexachlorobenzene	Yes			XX	
Hexachloroethane				XX	
Lindane	Yes	XX		XX	
Mercury (II)	Mercury	XX	XX		
Mercury (including methylmercury)	Mercury			XX	XX
Methylene chloride				XX	
Nickel		XX	XX		
Parathion		XX	XX		
PCBs (class)	Yes			XX	XX
Pentachlorophenol		XX	XX		
Selenium		XX	XX		
2,3,7,8-TCDD	Yes			XX	XX
Toluene				XX	
Toxaphene	Yes			XX	
Trichloroethylene				XX	
Zinc		XX	XX		

Key:

BCC = Bioaccumulative chemical of concern

DDD = 1, 1-Dichloro-2,2-bis(p-chlorophenyl)ethane

DDE = 1, 1-Dichloro-2,2-bis(p-chlorophenyl)ethylene

DDT = 1, 1, 1-Trichloro-2,2-bis(p-chlorophenyl)ethane

TCDD = Tetrachlorodibenzo-p-dioxin

Notes:

The concentration of 2,3,7,8-TCDD in the wastewater discharge to be compared with the effluent standards for the protection of human health must be calculated by determining the equivalent toxicity of all dioxins and furans.

DDT's metabolites include DDD and DDE.

The Guidance establishes that water quality criteria for protecting aquatic life, human health, and wildlife for pollutants not included in Tables 4 through 7 must be developed by each state using the procedures in Appendixes A, C, and D of the Guidance. In addition, water quality criteria for the chemical constituents or parameters listed in Table 9 can be established on a state-specific basis, which may or may not be in accordance with the procedures in the Guidance.

Antidegradation Policy

The antidegradation policy consists of the following provisions:

- applies, at a minimum, to BCCs;
- includes an antidegradation standard that requires maintaining the water quality necessary to protect existing uses;
- allows the identification of high-quality waters for individual pollutants;
- provides procedures for implementing the antidegradation standard;
- describes the information required for demonstrations to lower water quality; and
- presents the procedures to be followed to evaluate demonstrations to lower water quality and issue decisions on lowering water quality; these demonstrations are subject to public participation and intergovernmental coordination.

Mixing Zones Ban

With respect to the ban on mixing zones for BCCs, the Guidance establishes the following requirements:

- elimination of the use of mixing zones for the BCCs shown in Section A of Table 8 for new dischargers;

- phasing out of mixing zones for BCCs discharged by existing facilities by November 15, 2010; and
- allowing an exception to the mixing zone ban for dischargers that (1) have discharges at higher concentrations of BCCs than specified in the Guidance as a result of water conservation measures that produce an overall reduction of the mass of BCCs discharged, or (2) treat their discharge using the best possible technology and would suffer unreasonable economic effects if further control measures were taken.

GREAT LAKES TOXIC REDUCTION EFFORT

The GLTRE is evaluating nonpoint sources, wet-weather point sources, and atmospheric deposition to determine the need for additional controls on the discharge of BCCs from these sources. Some of the issues being evaluated or to be evaluated include

- additional regulation of BCCs under the Clean Air Act;
- spill prevention planning;
- reduction of BCC discharge from hazardous waste sites being remediated under the Resource Conservation and Recovery Act and the Comprehensive Environmental Response, Compensation, and Liability Act;
- improved pesticide registration procedures;
- additional regulation of combined sewer overflows and stormwater outfalls (wet-weather sources);
- public education on the dangers of BCCs;
- improved reporting of BCCs under the Toxic Release Inventory; and
- procedures to apply the National Contaminated Sediment Management Strategy in the Great Lakes Basin.

Table 4. Acute Water Quality Criteria for Protection of Aquatic Life in Ambient Water[1]

Part	Chemical	CMC (μg/L)	Conversion Factor (CF)
(a)	Arsenic (III)	339.8[a,b]	1.000
	Chromium (VI)	16.02[a,b]	0.982
	Cyanide	22[c]	n/a
	Dieldrin	0.24[d]	n/a
	Endrin	0.086[d]	n/a
	Lindane	0.95[d]	n/a
	Mercury (II)	1.694[a,b]	0.85
	Parathion	0.065[d]	n/a
	Selenium	19.34[a,b]	0.922

Part	Chemical	m(A)	b(A)	Conversion Factor (CF)
(b)	Cadmium[e,f]	1.128	−3.6867	0.85
	Chromium (III)[e,f]	0.819	+3.7256	0.316
	Copper[e,f]	0.9422	−1.700	0.960
	Nickel[e,f]	0.846	+2.255	0.998
	Pentachlorophenol[g]	1.005	−4.869	n/a
	Zinc[e,f]	0.8473	+0.884	0.978

Key:
The term “n/a” means not applicable.
The term “ln” represents the natural logarithm function.[2]
The term “exp” represents the base e exponential function.
CMC is Criterion Maximum Concentration.
CMC(tr) is the CMC expressed as total recoverable.
CMC(d) is the CMC expressed as a dissolved concentration.
CMC(t) is the CMC expressed as a total concentration.
Notes:
[a] CMC = CMC(tr)
[b] CMC(d) = [CMC(tr)]*CF. The CMC(d) shall be rounded to two significant digits.
[c] CMC should be considered free cyanide as CN.
[d] CMC = CMC(t)
[e] CMC(tr) = exp{m(A)*[ln (hardness)] + b(A)}
[f] CMC(d) = [CMC(tr)]*CF. The CMC(d) shall be rounded to two significant digits.
[g] CMC(t) = exp m(A)*[(pH) + b(A)]. The CMC(t) shall be rounded to two significant digits.

Table 5. Chronic Water Quality Criteria for Protection of Aquatic Life in Ambient Water[3]

Part	Chemical	CCC (μg/L)	Conversion Factor (CF)
(a)	Arsenic (III)	147.9[a,b]	1.000
	Chromium (VI)	10.98[a,b]	0.962
	Cyanide	5.2[c]	n/a
	Dieldrin	0.056[d]	n/a
	Endrin	0.036[d]	n/a
	Mercury (II)	0.9081[a,b]	0.85
	Parathion	0.013[d]	n/a
	Selenium	5[a,b]	0.922

Part	Chemical	m(C)	b(C)	Conversion Factor (CF)
(b)	Cadmium[e,f]	0.7852	−2.715	0.85
	Chromium (III)[e,f]	0.819	+0.6848	0.860
	Copper[e,f]	0.8545	−1.702	0.960
	Nickel[e,f]	0.846	+0.0584	0.997
	Pentachlorophenol[g]	1.005	−5.134	n/a
	Zinc[e,f]	0.8473	+0.884	0.986

Key:
The term “n/a’” means not applicable.
The term “exp” represents the base e exponential function.
The term “ln” represents the natural logarithm function.[4]
CCC is Criterion Continuous Concentration.
CCC(tr) is the CCC expressed as total recoverable.
CCC(d) is the CCC expressed as a dissolved concentration.
CCC(t) is the CCC expressed as a total concentration.
Notes:
[a] CCC = CCC(tr).
[b] CCC(d) = [CCC(tr)]*CF. The CCC(d) shall be rounded to two significant digits.
[c] CCC should be considered free cyanide as CN.
[d] CCC = CCC(t).
[e] CCC(tr) = exp{m(C)*[ln (hardness)] + b(C)}.
[f] CCC(d) = [CCC(tr)]*(CF). The CCC(d) shall be rounded to two significant digits.
[g] CMC(t) = exp{m(A)*(pH) + b(A)}. The CMC(t) shall be rounded to two significant digits.

[1] EPA recommends that metals criteria be expressed as dissolved concentrations.
[2] Not included in the regulation.
[3] EPA recommends that metals criteria be expressed as dissolved concentrations.
[4] Not included in the regulation.

Table 6. Water Quality Criteria for Protection of Human Health

Chemical	HNV (μg/L) Drinking	HNV (μg/L) Non-Drinking	HCV (μg/L) Drinking	HCV (μg/L) Non-Drinking
Benzene	1.9E1	5.1E2	1.2E1	3.1E2
Chlordane	1.4E−3	1.4E−3	2.5E−4	2.5E−4
Chlorobenzene	4.7E2	3.2E3		
Cyanides	6.0E2	4.8E4		
DDT	2.0E−3	2.0E−3	1.5E−4	1.5E−4
Dieldrin	4.1E−4	4.1E−4	6.5E−6	6.5E−6
2,4-Dimethylphenol	4.5E2	8.7E3		
2,4-Dinitrophenol	5.5E1	2.8E3		
Hexachlorobenzene	4.6E−2	4.6E−2	4.5E−4	4.5E−4
Hexachloroethane	6.0	7.6	5.3	6.7
Lindane	4.7E−1	5.0E−1		
Mercury (including methylmercury)	1.8E−3	1.8E−3		
Methylene chloride	1.6E3	9.0E4	4.7E1	2.6E3
PCBs (class)			3.9E−6	3.9E−6
2,3,7,8-TCDD	6.7E−8	6.7E−8	8.6E−9	8.6E−9
Toluene	5.6E3	5.1E4		
Toxaphene			6.8E−5	6.8E−5
Trichloroethylene			2.9E1	3.7E2

Key:
DDT = 1, 1, 1-Trichloro-2,2-bis(p-chlorophenyl)ethane
HNV = Human non-cancer value
HCV = Human cancer value
PCBs = Polychlorinated biphenyls
TCDD = Tetrachlorodibenzo-p-dioxin.

Table 7. Water Quality Criteria for Protection of Wildlife

Chemical	Criteria (μg/L)
DDT and metabolites	1.1E−5
Mercury (including methylmercury)	1.3E−3
PCBs (class)	7.4E−5
2,3,7,8-TCDD	3.1E−9

Key:
DDD = 1, 1-Dichloro-2,2-bis(p-chlorophenyl)ethane
DDE = 1, 1-Dichloro-2,2-bis(p-chlorophenyl)ethylene
DDT = 1, 1, 1-Trichloro-2,2-bis(p-chlorophenyl)ethane
PCBs = Polychlorinated biphenyls
TCDD = Tetrachlorodibenzo-p-dioxin
Note:
DDT and metabolites include DDD and DDE.

In addition to the GLTRE efforts to control BCCs, each of the Great Lakes states has prepared a Lake Management Plan (LaMP) under the binational GLWQA, which will evaluate current sources of BCCs and recommend enhancements to any media-specific program that would result in improved quality of the Great Lakes.

READING LIST

Additional information can be found in the web pages listed below. Information presented in this entry is current as of December 2001.

Great Lakes Initiative: http://www.epa.gov/waterscience/GLI/.

Table 8. Pollutants of Initial Focus in the Great Lakes Water Quality Initiative

Section A. Pollutants that Are Bioaccumulative Chemicals of Concern

Chlordane
4,4′-DDE; p,p′-DDE
Dieldrin
Hexachlorobutadiene; hexachloro-1, 3-butadiene
alpha-Hexachlorocyclohexane; alpha-BHC
delta-Hexachlorocyclohexane; delta-BHC
Mercury
Octachlorostyrene
Pentachlorobenzene
2,3,7,8-TCDD; dioxin
1,2,4,5-Tetrachlorobenzene
4,4′-DDD; p,p′-DDD; 4,4′-TDE; p,p′-TDE
4,4′-DDT; p,p′-DDT
Hexachlorobenzene
Hexachlorocyclohexanes; BHCs
beta-Hexachlorocyclohexane; beta-BHC
Lindane; gamma-hexachlorocyclohexane; gamma-BHC
Mirex
PCBs; polychlorinated biphenyls
Photomirex
1,2,3,4-Tetrachlorobenzene
Toxaphene

Section B. Pollutants that Are Not Bioaccumulative Chemicals of Concern

Acenaphthene
Acrolein; 2-propenal
Aldrin
Anthracene
Arsenic
1,2-Benzanthracene; benz[a]anthracene
Benzidine
3,4-Benzofluoranthene; benzo[b]fluoranthene
1,12-Benzoperylene; benzo[ghi]perylene
bis(2-Chloroethoxy) methane
bis(2-Chloroisopropyl) ether
4-Bromophenyl phenyl ether
Cadmium
Chlorobenzene
Chlorodibromomethane
2-Chloroethyl vinyl ether
2-Chloronaphthalene
4-Chlorophenyl phenyl ether
Chromium
Acenaphthylene
Acrylonitrile
Aluminum
Antimony
Asbestos
Benzene
Benzo[a]pyrene; 3,4-benzopyrene
11,12-Benzofluoranthene; benzo[k]fluoranthene
Beryllium
Bis(2-chloroethyl) ether
Bromoform; tribromomethane
Butyl benzyl phthalate
Carbon tetrachloride; tetrachloromethane
p-Chloro-m-cresol; 4-chloro-3-methylphenol
Chloroethane
Chloroform; trichloromethane
2-Chlorophenol
Chlorpyrifos
Chrysene

Table 8. (*continued*)

Copper	Cyanide
2,4-D; 2,4-Dichlorophenoxyacetic acid	DEHP; di(2-ethylhexyl) phthalate
Diazinon	1,2:5,6-Dibenzanthracene; dibenz[a,h]anthracene
Dibutyl phthalate; di-n-butyl phthalate	1,2-Dichlorobenzene
1,3-Dichlorobenzene	1,4-Dichlorobenzene
3,3′-Dichlorobenzidine	Dichlorobromomethane; bromodichloromethane
1,1-Dichloroethane	1,2-Dichloroethane
1,1-Dichloroethylene; vinylidene chloride	1,2-trans-Dichloroethylene
2,4-Dichlorophenol	1,2-Dichloropropane
1,3-Dichloropropene; 1,3-dichloropropylene	Diethyl phthalate
2,4-Dimethylphenol; 2,4-xylenol	Dimethyl phthalate
4,6-Dinitro-o-cresol; 2-methyl-4,6-dinitrophenol	2,4-Dinitrophenol
2,4-Dinitrotoluene	2,6-Dinitrotoluene
Dioctyl phthalate; di-n-octyl phthalate	1,2-Diphenylhydrazine
Endosulfan; thiodan	alpha-Endosulfan
beta-Endosulfan	Endosulfan sulfate
Endrin	Endrin aldehyde
Ethylbenzene	Fluoranthene
Fluorene; 9H-fluorene	Fluoride
Guthion	Heptachlor
Heptachlor epoxide	Hexachlorocyclopentadiene
Hexachloroethane	Indeno[1,2,3-cd]pyrene; 2,3-o-phenylene pyrene
Isophorone	Lead
Malathion	Methoxychlor
Methyl bromide; bromomethane	Methyl chloride; chloromethane
Methylene chloride; dichloromethane	Naphthalene
Nickel	Nitrobenzene
2-Nitrophenol	4-Nitrophenol
n-Nitrosodimethylamine	n-Nitrosodiphenylamine
n-Nitrosodipropylamine; n-nitrosodi-n-propylamine	Parathion
Pentachlorophenol	Phenanthrene
Phenol	Iron
Pyrene	Selenium
Silver	1,1,2,2-Tetrachloroethane
Tetrachloroethylene	Thallium
Toluene; methylbenzene	1,2,4-Trichlorobenzene
1,1,1-Trichloroethane	1,1,2-Trichloroethane
Trichloroethylene; trichloroethene	2,4,6-Trichlorophenol
Vinyl chloride; chloroethylene; chloroethene	Zinc

Key:
DDD = 1, 1-Dichloro-2,2-bis(p-chlorophenyl)ethane
DDE = 1, 1-Dichloro-2,2-bis(p-chlorophenyl)ethylene
DDT = 1, 1, 1-Trichloro-2,2-bis(p-chlorophenyl)ethane
TCDD = Tetrachlorodibenzo-p-dioxin
TDE = DDD

Table 9. Pollutants Subject to Federal, State, and Tribal Requirements

Alkalinity	Ammonia
Bacteria	Biochemical oxygen demand (BOD)
Chlorine	Color
Dissolved oxygen	Dissolved solids
pH	Phosphorus
Salinity	Temperature
Total and suspended solids	Turbidity

Federal Register—Environmental Documents: http://www.epa.gov/fedrgstr/.

Code of Federal Regulations: http://www.access.gpo.gov/nara/cfr/index.html.

Great Lakes Water Quality Agreement: http://www.on.ec.gc.ca/glwqa.

Great Lakes National Program Office: http://www.epa.gov/glnpo/.

QUANTITATIVE GROUNDWATER LAW

MARSHALL LAWSON
Land Conservation Legal Services, LLC
Columbia, South Carolina

INTRODUCTION

From ancient times, the law looked upon surface water and groundwater as separate resources, requiring separate rules for ownership and use. Whereas surface water was considered a public good to be shared among competing users, groundwater was equated with private property to be appropriated by the landowner at will.[1]

---

[1] Compare Dig. 43.12.3 (Paulus, Ad Sabinum 16) ("Rivers which flow [perennially] are public, and their banks are also public.");

The separate classification and treatment of groundwater can be attributed, in part, to a primitive understanding of hydrogeology; a mathematical description of groundwater storage and movement was not formulated until the midnineteenth century and not widely disseminated until the late twentieth century.[2] Without an understanding of aquifer characteristics, legal authorities lacked a basis for developing sophisticated groundwater allocation rules. The occurrence and nature of groundwater was, to them, too "... uncertain ...[to] subject it to the regulations of law...".[3]

Hand in glove with a nascent understanding of hydrogeology was the propensity of nineteenth century English jurists and commentators to follow the perceived equities of Roman water law as "indirect authority" for English water law.[4] The English Reasonable Use Rule, for example, embodies the Roman concept of limited property rights in public surface waters.[5] Likewise, the Roman classification of groundwater as private property free from any equitable rules of sharing appeared well over a millennium latter in English and American common law as Absolute Dominion or the Rule of Capture.[6]

The most important reason, however, for the development of separate legal rules for surface and groundwater centered not upon science or legal history but technology. Prior to the invention of the deep-well turbine pump in the last decade of the nineteenth century, groundwater could not commonly be exploited at rates sufficient to encroach upon the property rights of neighboring landowners. Consequently, there was little need to impose legal restraints on groundwater use.[7] The advent of submersible pumps and power drilling equipment, however, made it possible to develop aquifers at sufficient pumping rates to cause widespread injury to the environment and other property owners in the form of diminished stream flow, saltwater intrusion, land subsidence, well interference, and groundwater mining.[8] Technology had effectively made the once functional Rule of Capture obsolete in theory, if not in practice.

As groundwater conflicts became more common, American courts attempted to mitigate the harsh Rule of Capture by adopting various equitable rules of allocation, starting with the Reasonable Use Doctrine in 1862 followed by the Correlative Rights Doctrine in 1903 and The Restatement (Second) of Torts rule in 1979.[9] In reality, the rough justice dispensed by courts under these doctrines did little to protect groundwater supplies from being diminished or extinguished by competing users, and aquifers remained theoretically susceptible to complete dewatering by overlying property owners.

In response to the failure of the common law to manage groundwater use effectively, many Eastern states enacted legislation modifying common law rights in the use of groundwater. The legislation varies widely in scope and criteria but centers upon the requirement of an administrative permit to pump groundwater above certain predetermined thresholds.[10]

In contrast to the various quasi-riparian rules governing groundwater use in states west of the Mississippi River, a distinctly American system of water use, called the Prior Appropriation Doctrine, arose in the arid West. The system allocates water rights based upon priority of beneficial use and irrespective of contiguous (or in the case of groundwater, overlying) land ownership.

Appropriation proved more suited to the climate and geology of the arid West than riparianism. The central tenets of the riparian doctrine, that water rights accrue

---

Dig. 43.20.3.1 (Pomponius, Ad Sabinum 34) ("Many may take away water from a river, but in such manner only that their neighbors are not injured, or if the stream is not large, those on the other side.") with Dig. 43.24.11 (Ulpian, Ad Edictum 71) (Labeo stating the well settled rule that groundwater was considered part of the land); Dig. 39.3.1.12 (Ulpian, Ad Edictum 53) (Marcellus holding "that to him who, digging in his own land, turns away the sources of his neighbor's spring, nothing can be done, not even an action for fraud, and surely there should be none if that which he did was for the purpose of benefiting his own field, and not with the intention of injuring his neighbor.") (The Digests or Pandects of Justinian, commenced in A.D. 530 and completed in A.D. 533, comprise part of the Corpus Juris Civilis, the great codification of the Roman Emperor Justinian. The sections of the Digest pertaining to water reveal a remarkable understanding of the equities of water use. Although written almost 1500 years ago and based upon legal principles formed over the preceding 1000 years, the Pandects provide a foundation and, in some cases, a blueprint for many of the principles of Western water law in use today); See also, Earl Finbar Murphy, Waters and Water Rights § 20.02 (Robert Beck ed., 2d ed. 1991).

[2] French Engineer Henry Darcy pioneered the basic equations of modern groundwater hydrogeology in the mid-nineteenth century by quantifying the rate and direction of groundwater flow; see C.W. Fetter, Applied Hydrogeology 76 (1988); Henri-Philbert-Gaspard Darcy, Encyclopedia Britannica, III Micropaedia 883 (1991).

[3] Murphy, supra note 1, at §§ 20.03 [citing Chatfield v. Wilson, 28 Vt. 49, 54 (1856)].

[4] See C.E. Busby, *American Water Rights Law: 5* S.C.L.Q. 106, 116 n. 57 (1952) (The language "no direct authority" in Acton v. Blundell 152 Eng. Rep. 1228 (Ex. Cham. 1843) raises the presumption the court relied on Roman law in adopting a Rule of Capture for groundwater); See also Murphy, supra note 1, at § 20.04 [citing Chasemore v. Richards 11 Eng. Rep. 140, 155 (1857)] (Lord Wensleydale quoting from the Roman sources for the proposition that groundwater lay within the domain of private property); Embrey v. Owen, 155 Eng. Rep. 579, 585 (1851) (Baron Parke adopting the Roman concept of usufructuary rights in public surface waters).

[5] Murphy, English Water Law Doctrines Before 1400, 1 Am J. Leg. Hist. 103, 105–106 (1957) (tracing the Roman concept of surface water as a public good); Embrey v. Owen, 159 Eng. Rep. 579, 585 (1851) ("[Water that] is *publici juris* [is not so] in the sense that it is a *bonum vacans,* to which the first occupant may acquire an exclusive right, but that it is public and common in this sense only: that all may reasonably use it who have a right of access to it; that none can have any property in the water itself, except in particular portion which he may choose to abstract from the [source] and take into his possession, and that during the time of possession only.").

[6] See Greenleaf v. Francis, 35 Mass. 117 (1836); Acton v. Blundell, 152 Eng. Rep. 1223 (Ex. Cham. 1843).

[7] Raphael G. Kazman, Modern Hydrogeology 272 (1988).

[8] Sax, J.L. et al. (2000). *Legal Constraints on Water Resources*, 3rd Edn. p. 345.

[9] See generally, Basset v. Salisbury Mfg. Co., 43 N. H. 569, 82 Am Dec. 179 (1862); Katz v. Walkinshaw, 74 P. 766 (Cal. 1903); Restatement (Second) of Torts § 858 (1979).

[10] Murphy, supra note 1, at § 24.05.

only to contiguous or overlying landowners and are primarily limited to on-tract use, proved impracticable in the West where the federal government owned much of the land, water resources were scarce, and mining and agriculture practices required off-stream water use. Furthermore, the subjective allocation criteria of the various riparian doctrines proved ill suited for managing scarcity.

ORIGIN AND NATURE OF PROPERTY RIGHTS IN GROUNDWATER

Property rights in water, including groundwater, have a long pedigree in European jurisprudence. Roman law is the oldest legal system to which the roots of European water law can be traced with any certainty and, as discussed above, Roman jurists considered groundwater an integral part of the land.[11] As such, the landowner enjoyed a property right in the groundwater itself (res) as well as an unfettered right of use (usufruct).[12] English and American courts firmly embraced this concept of groundwater as an absolute property right in the nineteenth century.[13] Although many U.S. courts and legislatures now define groundwater as a public resource, most, if not all, jurisdictions still acknowledge (at least tacitly) a usufructuary right of the landowner to use the groundwater underlying his property.[14]

As a form of property, regulation of groundwater use is theoretically subject to challenge as a "taking" of private property without compensation in violation of the Takings Clause of the United States Constitution.[15] At least one court has stricken modification of water rights as a violation of the Takings Clause,[16] but the decision is very much an anomaly. Other courts addressing the issue have upheld groundwater regulation in the face of takings clause challenges on one of the following grounds: (1) Permitting has not deprived a landowner of all economic use of his property right and is, therefore, not a compensable taking. (2) Landowners enjoy no constitutionally protected right in groundwater until they actually begin pumping. (3) The regulation is a valid exercise of the sovereign's police power.[17] States have broad police power under the Constitution to regulate for the public health, safety, morals, and general welfare. The sovereign's exercise of this power may not require compensation for diminished property rights resulting from regulation, as long as such regulation is not arbitrary or irrational.[18]

At least one authority would add another defense to takings clause challenges to groundwater regulation. Because the overlying landowner's right to use water under various riparian doctrines is "subject to complete destruction by a more powerful pumper, [the] ... landowner's expectation of exclusive use of groundwater [is] close to illusory."[19] Stated another way, riparian law (or quasi-riparian law for groundwater) holds that groundwater is private property but effectively treats it as common property by failing to provide a mechanism for quantifying and protecting the water right.

On the other hand, it could be argued that permitting systems for groundwater use, depending upon the terms of the permit, might actually strengthen the claim of a landowner to a property right in the use of groundwater. Consistent with this argument, a permit holder could defend his right to pump against future permit holders if such pumping reduced his own allotted yield.[20]

Regardless of the court's willingness to uphold groundwater-use regulation, it should not be assumed that courts or groundwater-use statutes have abolished private property rights in the use of groundwater. Even the Regulated Riparian Model Water Code, which proposes comprehensive regulation of water resources, makes clear that the administrative permit process does not usurp

---

[11] Ludwig A. Teclaff, Water Law in Historical Perspective 6 (1985); Dig. 43.24.11 (Ulpian, Ad Edictum 71).

[12] Dig. 39.3.1.12 (Ulpian, Ad Edictum 53); Maricopa County Mun. Water Conservation Dist. Number One v. Southwest Cotton Co., 4 P.2d 369, 372 (Ariz. 1931).

[13] See Greenleaf v. Francis, 35 Mass. 117, 121 (1836); Acton v. Blundell, 152 Eng. Rep. 1223, 1228 (Ex. Cham. 1843); Chatfield v. Wilson, 28 Vt. 49, 54 Chasemore v. Richards, 11 Eng. Rep. 140, 155 (1857); see also, John M. Gould, Law of Waters 538 (2d Ed. 1891) ("Water percolating through the ground beneath the surface, either without a definite channel, or in the courses which are unknown and unascertainable, belongs to the realty in which it is found").

[14] Murphy, supra note 1, at § 20.04 ("Notwithstanding the importance of public rights in water, both the common law and the legislatively created appropriation doctrine have made rights to water primarily the domain of private property, and this includes groundwater"); see generally, Bamford v. Upper Republican Natural Resources District 512 N.W. 2d 643 (Neb. 1994); Tequesta v. Jupiter Inlet Co., 371 So.2d 663 (Fla. 1979).

[15] U.S. Const. amend. V (The Takings Clause of the U.S. Constitution bars the taking of property by the federal or state government without compensation); see, e.g., Lucas v. South Carolina Coastal Counsel, 505 U.S. 11003 (1992) (finding, on remand, a compensable taking based upon the South Carolina Coastal Council's refusal to grant a permit to construct two homes on the petitioner's oceanfront property; see generally, Robert Meltz et al., The Takings Issue: Constitutional Limits on Land Use Control and Environmental Regulation 463-70 (1999).

[16] See American Charolaise, Ltd. V. Oklahoma Water Resources Board, 855 P.2d 568 (Okla. 1990) (The Oklahoma Supreme Court struck down legislation limiting riparian water rights, holding that riparian rights constitute private property rights and legislation permitting competing appropriators to divert the water constituted a taking of that property for public use).

[17] See Meltz, supra note 15, at 126 (citing Department of Agric. & Consumer Serv. V. Polk, 276 U.S. 272 (1928); German v. City of Philadelphia, 700 A.2d 444 (Pa. 1997); Hoeck v. City of Portland 57 F.3rd 781, 787 (9th Cir. 1995)). Compare City of Houston v. Crabb, 905 S.W.2d 669, 674-75 (Tex. Ct. App. 1995) (finding a taking where the "[c]ity failed to prove a valid exercise of its police power").

[18] SAX, supra note 8, at 396 (citing Peterson v. Dept. of Ecology, 596 P.2d 285 (1979)); see also Bamford v. Upper Republican Natural Resources Dist., 512 N.W.2d 642 (Neb. 1994); Town of Chino Valley v. City of Prescott, 638 P.2d 1324 (1981); Crookston Cattle Co. v. Minnesota Dept. of Nat'l Resources, 300 N.W.2d 769 (Minn. 1980); Village of Tequesta v. Jupiter Inlet Corp., 371 So.2d 663 (Fla. 1979); Omernik v. State, 218 N.W.2d 734 (Wis. 1974); Knight v. Grimes, 127 N.W. 2d 708 (S.D. 1964); Williams v. City of Wichita, 279 F.2d 375 (10th Cir. 1960).

[19] A. Dan Tarlock et al., Law of Water Rights and Resources § 4.09[1] (1986).

[20] Sax, supra note 8, at 396.

common law rights in water and that "the new system is a regulation rather than a taking of the older riparian rights."[21]

LEGAL CLASSIFICATIONS OF GROUNDWATER

The common law historically distinguished not only between groundwater and surface water, but also distinguished between two types of groundwater: (1) that found flowing in underground streams; and (2) so-called "percolating" groundwater, or that which did not "flow" but rather "oozed" through the void spaces within the aquifer skeleton. Courts generally assign the law of surface streams to the former, and groundwater law to the latter.[22]

As mentioned above, this dual classification arose initially because "percolating" groundwater was envisioned as too "occult in its occurrence and movement" for courts to attach formalistic rules to its use.[23] In addition, courts labored under the false assumption that a large portion of groundwater occurs as an underground analog to surface streams, and, therefore, both flow regimes should be governed by the same principles. In reality, "percolating" groundwater, like stream water, flows in an ascertainable direction and rate, albeit through an aquifer matrix and at orders of magnitude slower than surface water. Furthermore, the occurrence of groundwater as underground streams is extremely rare and confined to karstic terrains or lava flows.[24]

In acknowledgment of the hydraulic interconnection between groundwater flow and stream flow, some modern courts recognize a third legal category of groundwater, the "subflow" of surface streams. The subflow of surface streams, like water flowing in underground streams, was generally made subject to the same law that applied to surface water.[25]

Finally, statutes in a number of jurisdictions exempt certain kinds of groundwater from ordinary groundwater law. For example, groundwater pumped from a mine to allow continued mineral extraction may be exempt from regulation applied to other groundwater. Another category of groundwater often exempted from regulation as groundwater is water hot enough to have value as a geothermal resource. Water extracted from the ground as a by-product of oil and gas development may also be separately treated by statute.[26]

LEGAL RULES FOR GROUNDWATER USE

Federal Legislation

The United States has no single rule governing groundwater use. Although the federal government regulates a wide range of groundwater pollution sources, it has, thus far, left formulation of quantitative groundwater law to the states, which have in turn adopted their own unique rules and statutes governing groundwater use.[27]

State Groundwater Law

Common Law (Quasi-Riparian Doctrines). Four distinct allocation doctrines for groundwater use apply east of the Mississippi and in a handful of Western states. These doctrines are best described as quasi-riparian because the property right in the groundwater vests with ownership of overlying land. They are (1) the "English rule" of Capture or Absolute Ownership, (2) the "American rule" of Reasonable Use, (3) the "Correlative Rights Rule" of shared access, and (4) the Restatement of the Law of Torts (2d, Tent. Draft 17, § 858A).[28]

A definitive and accurate appraisal of states adhering to a particular groundwater doctrine is problematic due to the blurring of the various rules in case law; an absence of decisional law in some states; and administrative overlays on the common law that modify, in various degrees, common law rights in groundwater.

In general terms, Absolute Ownership survives in only a few jurisdictions.[29] Most Eastern states employ some variation of the Reasonable Use Rule.[30] Several states follow some variation of the Correlative Rights Rule originating in California.[31] The Restatement approach has found little favor and has been adopted in only a handful of states.[32]

Absolute Ownership (Capture). Under this doctrine, also known as the English Rule of Capture, landowners are allowed to withdraw all the water they wish for whatever

---

[21] American Society of Civil Engineers, Water Resources Planning & Management Div., the Regulated Riparian Model Water Code; Final Report of the Water Laws Committee, vii (Joseph W. Dellapenna ed,, 1997).

[22] Murphy, supra note 1, at § 20.07 (citing the American & English Encyclopedia of Law 311 (1905); Dickinson v. The Grand Junction Canal Co., 155 Eng. Rep. 953 (Ex. 1852)).

[23] See, e.g., Frazier v. Brown, 12 Ohio St. 294, 311 (1861) ("[T]he existence, origin, movement, and course of such water, and the causes which govern and direct their movements, are so secret, occult and concealed, than an attempt to administer any set of legal rules in respect to them would be involved in hopeless uncertainty, and would be, therefore, practically impossible.").

[24] Fletcher G. Driscoll, Groundwater and Wells, 672 (2d ed. 1986).

[25] Sax, supra note 8, at 360.

[26] Id. at 361.

[27] Murphy, supra note 1, at § *20.04*.

[28] SAX, supra note 8, at 364.

[29] Id. at 366; see generally, Wiggins v. Brazil Coal & Clay Corp., 452 N.E.2d 958 (Ind. 1983); City of Sherman v. Public Util. Comm'n, 643 S.W.2d 683 (Tex. 1983); St. Amand v. Lehman, 47 S.E. 949 (Ga. 1904) (but putting a restriction on the rule of capture by holding a malicious diversion or wasting of percolating underground waters would be actionable).

[30] See generally, Clinchfield Coal Corp. v. Compton, 139 S.E. 308 (1927); Rouse v. City of Kinston, 123 S.E. 482 (N.C. 1924); Nashville C. & St. L. Ry. V. Rickert, 89 S.W.2d 889 (1935), *cert. Denied* by Tennessee Supreme Court (1936); Sloss-Sheffield Steel & Iron Co. v. Wilkes, 165 So. 764 (Ala. 1938); Barclay v. Abraham, 96 N.W. 1080 (Iowa 1903); Cincinnati, N.O., & Tex. Pac. R.R. v. Gillespie, 113 S.W. 89 (Ky. 1908); Schenck v. City of Ann Arbor, 163 N.W. 109 (Mich. 1917); Pence v. Carney, 52 S.E. 702 (W. Va. 1905); Cline v. American Aggregates Corp.; 474 N.E.2d 324 (Ohio 1984); Rothrauf v. Sinking Spring Water Co., 14 A.2d 87(Pa. 1940); Wood v. Picillo, 443 A.2d 1244 (R.I. 1982); State v. Michels Pipeline Const., Inc., 217 N.W.2d 339 (Wis. 1974)

[31] Katz v. Walkinshaw, 74 P. 766 (Cal. 1903); see also Jones v. Oz-Ark-Val Poultry Co., 306 S.W.2d 111 (Ark. 1957); Prather v. Eisenmann, 261 N.W.2d 766 (Neb. 1978); Vt. Stat. Ann. Tit. 10, § 1410(a)(3) (1984).

[32] SAX, supra note 8, at 366.

purpose on or off the owner's land. The owner could even waste the water, thereby injuring a neighbor, but still face no liability.[33]

American Reasonable Use. Reasonable Use, also known as the American Rule, found favor with American courts as a modification of Absolute Dominion. In effect, Reasonable Use is equivalent to an equitable servitude upon the property right in groundwater requiring that (1) groundwater use be reasonable and (2) such use must be confined to the overlying tract of land.[34]

The term "Reasonable" varied among jurisdictions but generally meant "A waste of water or a wasteful use of water was unreasonable only if it caused harm, and any nonwasteful use of water that caused harm was, nevertheless, reasonable if it was made on or in connection with the overlying land."[35] The doctrine does not articulate a rule of allocation, but some courts imply a rule of sharing that makes the practical application of the doctrine difficult to differentiate from the Correlative Rights Doctrine.[36]

Correlative Rights ("California Rule"). This doctrine, first articulated by the California Supreme Court, requires a sharing of available water equitably among overlying landowners. They share the resource in times of scarcity proportionate either to their aerial land ownership or actual historic use.[37] Even if the water is used beneficially on the overlying land, landowners cannot infringe upon the rights of others by exceeding their proportionate shares.[38]

Off-tract uses are subordinate and are legally protected only in the event "surplus" water is available, that is, if recharge exceeds withdrawals.

If no competition for water exists, then Correlative Rights are synonymous with Reasonable Use.[39]

Restatement (Second) of Torts (Reasonable Use). Although the Restatement rule is similar to reasonable use, it draws no distinction between on-tract and off-tract use. It also differs from Reasonable Use by establishing criteria for liability. A landowner who pumps groundwater from his land and uses it for a beneficial purpose is liable for interference with other's groundwater use if

1. the withdrawal causes unreasonable harm through lowering of the water table;
2. the groundwater occurs in a distinct underground stream;
3. the withdrawal of water has a substantial effect upon a stream, river, or lake.[40]

Priority System (Prior Appropriation). In the West, a uniquely American water use rule called the Prior Appropriation Doctrine gradually supplanted riparian principles for both surface and groundwater. In contrast to riparian doctrines, the acquisition of water rights does not depend upon the ownership of contiguous or overlying land. Water rights vest by appropriating the water and putting it to beneficial use and accrue on a first in time, first in right, basis. Furthermore, water rights are quantified, can be lost through nonuse, and are transferable, as long as other water rights holders suffer no harm.

Prior appropriation developed in the mining camps of the Sierra Nevada during the California gold rush of the 1840s. As discussed before, the nature of mining operations and federal ownership of western lands were incompatible with the riparian doctrine for three reasons: (1) Under the riparian system, water rights accrued only to the owners of land. The miners, however, did not own the land they mined; it belonged to the federal government; (2) Mining operations were often conducted remotely from waterbodies requiring off-stream diversions that violated riparian prohibitions of off-tract transfers of water.[41] (3) In contrast to riparian doctrines, Appropriation provided quantifiable water rights that were enforceable and transferable. These attributes fostered predictability of planning and investment.[42]

Appropriation arose first as a common law doctrine, effectively ratifying water use practices in the mining camps. Over time, the doctrine was codified and even incorporated into some state's constitutions.[43] Most Western states have extended some version of the Prior Appropriation Doctrine, to groundwater, by statute, making it the primary doctrine for groundwater use in the West.[44]

The Appropriation rules for groundwater are essentially the same as those for surface water. In case of a dispute, priority of use is the determining factor in allocating water among competing users; junior claimants may have their supply of groundwater diminished or eliminated by senior claimants in times of scarcity.[45] The realities of hydrogeology, including the time lag between the start of pumping and deleterious effects on water levels, can make

---

[33] See Greenleaf v. Francis, 35 Mass. 117 (1836) ("Everyone has the liberty of doing in his own ground whatsoever he pleases, even although it should occasion his neighbor some other sort of inconvenience"); Acton v. Blundell, 152 Eng. Rep. 1223, 1228 (Ex. Cham. 1843) (In this case, a landowner's well was dewatered by a neighbor's excavation of a coal mine. The court held the damage to be *damnum absque injuria* [an injury without a remedy] based upon its determination that groundwater flow was inherently unpredictable: "... the person who owns the surface may dig therein; and apply all that is there found to his own purposes at his free will and pleasure; and that if, in the exercise of such right, he intercepts or drains off the water collected from underground springs in his neighbor's well, this inconvenience to his neighbor ... cannot become the ground of an action ...").

[34] Henderson v. Wade Sand & Gravel Co., 388 So.2d 900 (Ala. 1980)

[35] Id.

[36] Murphy, supra note 1, at § 22.01

[37] Joseph W. Dellapenna, Groundwater Law for Mineral Lawyers, Energy & Mineral Law Foundation (1992).

[38] Murphy, supra note 1, at § 20.07.

[39] Sax, supra note 8, at 364; Fletcher G. Driscoll, Groundwater and Wells 673 (2d ed. 1986).

[40] Restatement (Second) of Torts § 858 (1979).

[41] Sax, supra note 8, at 282

[42] Id. at 286

[43] Id. at 382

[44] Id. at 282; Murphy, supra note 1, at §§ 20.04; 20.07.

[45] Clesson Kinney, The Law of Irrigation and Water Rights and Arid Region Doctrine of Appropriation of Waters 598 (1912); Atchison v. Petersen, 87 U.S. 507 (1874); SAX, supra note 8, at 382.

application of the doctrine difficult to enforce in the case of groundwater.[46]

Active Management and Regulation of Groundwater

Permit Statutes ("Regulated Riparianism")[47]**.** Although some Eastern states enacted groundwater statutes more than a century ago, such legislation did not become common until the second half of the twentieth century. As of 1996, eighteen states had enacted regulated riparian systems for surface water sources, generally including underground water resources as well. Two other states apply a regulated riparian system only to underground water sources.[48]

The common characteristic of such legislation is the requirement of a permit to withdraw groundwater based upon reasonableness of use.[49] An administrative agency is responsible for both the decision to grant or deny a permit and, at least initially, dispute resolution.[50]

These statutes were not originally introduced as a radical revision of the water law of a particular state but emerged gradually through a process of discrete initiatives in the form of well-spacing regulations, prohibitions of wasteful pumping, or licensing of major withdrawals. In several states, it remains unclear whether the law would be better described as basically riparian rights with limited legislative alterations or as comprehensive regulatory schemes referred to as "regulated riparianism."[51]

Supply Side (Conjunctive Use) Management. The Orange County (California) water district (OCWD) takes a novel approach to groundwater allocation. Instead of imposing restrictions on pumping, the district focuses on maintaining groundwater supplies by recharging aquifers with imported surface water. No attempt is made to limit groundwater pumping directly or prevent waste. In fact, OCWD is forbidden from directly limiting the amount of groundwater an individual user may withdraw, although it may discourage use by levying a "basin equity assessment" against pumpers who withdraw more than a percentage of groundwater set by the district. In addition, the OCWD does not attempt to manage groundwater withdrawals to prevent or minimize well interference.[52]

The drawbacks to this type of management model include dependence upon imported recharge water and the cost of transmission and recharge facilities. These costs, however, are largely offset by maintaining a stock of relatively cheap local water in the form of groundwater storage.[53]

NATURE OF GROUNDWATER DISPUTES

Most groundwater disputes are the product of a race-to-pump mentality referred to as "the tragedy of the commons."[54] When many users tap the same aquifer, it becomes a common property resource. Without social or legal limits on water use from the common aquifer, there is no incentive to conserve because the conserver does not enjoy an exclusive right to the water that is saved.[55] For example, under the Reasonable Use Doctrine, two neighboring landowners are free to withdraw all the groundwater they can put to beneficial use upon their own land. The reasonableness of the use is judged solely as it relates to the purpose of such use and not its effect upon others or the sustainability of the resource.[56]

The race-to-pump has produced widespread declines in potentiometric surfaces and in severe cases, groundwater mining[57] that, in turn, led to well interference, land subsidence, saltwater intrusion,[58] and diminished stream flow. Unwise groundwater use may even threaten endangered species by desiccating springs.[59] In addition to environmental damage, these phenomena impose additional costs upon water users, ranging from a few thousand dollars incurred by individual well owners for well deepening and increased electrical costs to millions of dollars to provide alternative water supplies for municipal systems threatened by saltwater intrusion.[60]

---

[46] Kazman, supra note 7, at 272; DELLAPENNA, *supra* note 37, at 7.

[47] Dellapenna, supra note 37, at 7

[48] Model Water Code, supra note 21, at vi.

[49] Id. at viii.

[50] Dellapenna, supra note 37.

[51] Model Water Code, supra note 21, at vi.

[52] Sax, supra note 8, at 454 (citing William Bloomquist, Dividing the Waters: Governing Groundwater in Southern California, ICS Press, 1992).

[53] Id. at 455.

[54] William H. Rodgers, Jr., Environmental Law 39 (West 1994) [citing Garrett Hardin, *The Tragedy of the Commons*, 162 Science 1243 (1968)].

[55] Tom Tietenberg, Environmental and Natural Resource Economics 207 (1988)

[56] Hornsberger, Oeltjen, & Fischer, *Groundwater: From Windmills to Comprehensive Public Management*, 52 Neb. L. Rev. 179 at 205 (1973).

[57] Driscoll, supra note 24, at 64 (The potentiometric surface is an imaginary line representing the hydrostatic head throughout all or part of a confined aquifer and is the level to which water will rise in a tightly cased well. Unwise groundwater use practices, such as improper well spacing, may draw down the potentiometric surface without actually dewatering the aquifer); Kazman, supra note 7 (Groundwater mining results when the rate of withdrawal exceeds the rate of recharge to an aquifer, decreasing the volume of groundwater held in storage. This phenomenon is especially acute in arid regions where aquifers have little to no natural recharge. Without artificial recharge, it may take thousands of years for such an aquifer to regain its prepumping storage volume).

[58] J. Marshall Lawson, *Transboundary Groundwater Pollution: The Impact of Evolving Groundwater Use Laws on Salt Water Intrusion of the Floridian Aquifer along the South Carolina-Georgia Border,* 9 S. C. Env. L. J. 85 (2000).

[59] Ben F. Vaughn IV & Peter M. Emerson, Protecting the Edwards Aquifer; An Efficient and Ecological Alternative, in Water Marketing: The Next Generation 167 (Terry L. Anderson & Peter J. Hill eds., 1997) (Explosive population growth and unrestricted pumping of groundwater from the Edwards aquifer in central Texas is depleting the flow of water to the aquifer's discharge areas threatening several endangered species with extinction and threatening economic interests dependent on plentiful water supplies. This realization has prompted legislators to rethink the state's longstanding adherence to the Rule of Capture).

[60] Lawson, supra note 58, at 95 [Saltwater intrusion of the Floridian aquifer in the vicinity of Savannah, Georgia, forced

Groundwater disputes may involve adjudication of the denial for new permits to pump groundwater or limitations and conditions upon existing permits.[61] In the case of property damages from groundwater use, courts may apply groundwater law (common law or a statute on point) or even tort law (e.g., nuisance or negligence theories) to the facts of a case to determine pumping rights or fix remedies in the form of an injunction or money damages.[62]

Plaintiffs may be private citizens, corporations, or public entities such as municipalities seeking to define, expand, or defend rights to groundwater. Depending on the jurisdiction, groundwater disputes, not resolved by administrative processes, may be appealed to higher state courts.

In addition to state common law and administrative law, federal law may come into play if groundwater use affects federal water rights claims,[63] impacts a federal statute, such as The Endangered Species Act,[64] or when groundwater use has an interstate dimension such as transportation of groundwater across state lines.[65]

LEGAL REFORM

Until recently, American groundwater law has been a hodgepodge of outdated common law doctrines that lacked an effective mechanism for allocating supplies on a sustainable basis and dealt only with disputes after the fact, during litigation. The need for reform prompted proposals for statutory models that can be roughly divided between emerging market-based approaches based upon the movement of water to higher value uses,[66] on the one hand, and command and control regulatory schemes restricting water use, on the other.[67]

Regulatory proponents believe that agency experts are the best suited to allocate the limited supplies in a sustainable fashion. In their eyes, market-based approaches to natural resource management inevitably lead to competitors vying to monopolize water rights and, thus, deprive others of access to water supplies.

Market proponents, on the other hand, argue that administrative control does not end competition for water but only shifts competition from the market to the political arena thereby promoting uncertainty that tends to work against meaningful conservation.[68] They also decry the assumption long held by legal authorities that control of the production of water, like alcohol, must somehow be tied to the end use of the product and not related to the valuation of the marketplace.[69] The more economical and less environmentally harmful choice to them is a system of tradable property rights in groundwater based upon either a share system or a seniority system.[70] This type of system would allow water to seek its highest value use, benefiting both buyers and sellers and minimizing government regulation with its attendant costs.

Others would like to see a combination of the two models, market incentives with government oversight to protect outside parties and the general public from harmful side effects.[71] Such oversight could include zoning restrictions, assignment of quota rights in common pools, and the right to receive money damages for economic harm to individual water users.[72]

CONCLUSION

Although groundwater law has progressed beyond the ancient Rule of Capture in most jurisdictions, it remains far from the nimble and pristine law required for the wise management of groundwater resources. Effective groundwater law of the future must focus on sustainable use of the resource within the context of both local and regional hydrogeologic realities, a formidable task, given the complexities of hydrogeology and political impediments to meaningful legal reform.

The law should encourage efficient water use by providing quantifiable and transferable property rights in water, and government should play the role of gatekeeper in maintaining environmental integrity and access to water for newcomers. Above all, water should no longer be regarded as so plentiful that it is regarded as a free good or so precious that it can be assigned no value. Where demand exceeds thresholds the environment can safely bear, as in mining or saltwater intrusion, consumers should pay for the realistic cost of alternative water supplies. Realistic water pricing combined with sound laws incorporating the appropriate mix of market incentives and regulation would foster both conservation and the beneficial transfer of water from areas of abundance to areas of scarcity.

Regardless of the mix of regulation and economic incentives, wise use of groundwater in the future will

---

Public Service Districts on Hilton Head Island, South Carolina, to develop alternative water supplies (AWS) for the island in the form of transmission lines to the Savannah River and the construction of a deep well and desalination facility to treat heavily mineralized water from deeper aquifers. The total cost of AWS was estimated at $24,950,000.00).

[61] See, e.g., Baker v. Ore-Ida Foods, Inc., 513 P.2d 627 (Idaho 1973).

[62] See generally, Adams v. Lang, 533 So.2d 89 (Ala. 1989); Friendswood Development Co. v. Smith Southwest Industries, Inc., 576 S.W.2d 21(Tex. 1978); Finley v. Teeter Stone, Inc. 248 A.2d 106 (Md. 1968); Restatement (Second) of Torts § 858(1)(a).

[63] See, e.g., Cappaert v. United States, 426 U.S. 128 (1976).

[64] 16 U.S.C.A. §§ 1531–1534 (1982).

[65] Vaughn & Emerson, supra note 59, at 177; Sporhase v. Nebraska, ex rel. Douglas, 458 U.S. 941 (1982) (The U.S. Supreme Court holding that a Nebraska statutory restriction on the interstate transportation of its groundwater violated the Commerce Clause of the U.S. Constitution); See also U.S. Const. art. I, § 10, cl. 3; art. III, § 2. cl. 1,2.

[66] Vaughn & Emerson, supra note 59, at 167; Terry L. Anderson & Pamela Snyder, Water Markets: Priming the Invisible Pump 161; Tom Tietenberg, Environmental and Natural Resource Economics 203 (1988); James E. Kundell & Diana Tetens, Carl Vinson Inst. of Gov't, Whose Water Is It?: Major Water Allocation Issues Facing Georgia 9 (1988).

[67] See generally, Model Water Code, supra note 21.

[68] Kazman, supra note 7, at 278 (Citing J. Hirshleifer et al., Water Supply Economics, Technology, and Policy 35 (1960)).

[69] Id. at 283.

[70] Vaughn & Emerson, supra note 59, at 180.

[71] Kazman, supra note 7, at 367 (Citing J. Hirshleifer et al., Water Supply Economics, Technology, and Policy 364 (1960)).

[72] Id. at 337.

require the education and cooperation of stakeholders, and also integrating the law with sound scientific, economic, and technological principles.

ISLAMIC WATER LAW

A. Zaharuddin
S. Sabri
University of Malaya
Petaling Jaya Selangor,
Malaysia

ISLAMIC WATER LAWS

Islam covers all aspects of human life. It regulates the relationships among God (Allah), man, and nature. It is based on the recognition of the Creator and of human submission to His will. Muslims believe that everything originates from the One God and everyone is responsible to Him. According to Islam, Allah created nature for the benefit of humans. The relationship between humans and nature is based on harmony; humans are urged to explore and use natural resources in a sustainable manner.

Water is considered one of the most profound elements in Islam. Water is also necessary for cleaning one's home and personal effects, as well as for general hygiene. It is indispensable to agriculture and industry: *He it is who sends down water from the sky. With it, we bring forth plants of every kind* (6:99). It cannot be denied that all human beings rely on water for life and good health, but for Muslims, water is first and foremost, a social good in Islam. It is also regarded as a blessing from Allah that gives and sustains all life in this world. Furthermore, the word "water" appears sixty-three times in the holy book of Muslims, Al-Quran. This indicator shows the importance of water in our daily life.

Before the coming of the Prophet Muhammad, there were no standard water regulations in the Arab world. The only guide concerning water laws was through customs practiced by tribes. There are four fundamentals in Islamic law, the first three can be considered the roots of Muslim dogma. The first is the Holy *Quran*, the divine revelation from Allah to humans through the Prophet Muhammad. It consists of 114 suras (chapters) and 6666 words. It also considered the basis for all legislation.

Second are the *Sunna* and the *Hadith*, which represent the sum total of habits or general pattern of life of the Prophet Mohammad, inclusive of his word, acts, and tacit approval. Third, is the *Ijma*, which means the universal consensus of opinion of the nations or Muslim community. And finally, *Qiyas* or deduction by legal analogy. It consists of the following precedents in deciding new cases, which have not been discussed or found before.

WATER RIGHTS IN ISLAM

In Islam, water is considered a gift from God, so no individual literally owns it. Humans act as the stewards of water as well as other resources that belong to the community. The environment is protected from humans by specifics injunctions against upsetting its natural order through pollution or other activities. This is to ensure that Allah's gifts are available to all living things in the present and future.

Although humans are the most favored creation, we are also responsible for ensuring that water is available to all living things. Allah commands believers to "*make not mischief on earth*." (2:11); this means that we cannot spoil or degrade natural resources. The prophet Mohammad, very sensibly, forbade urination into stagnant water and advised guarding against this practice. This shows how Islam has underlined that water sources should be guarded from any contempt that can pollute it. This opens the door for penalties or punishment in modern legislation for those who disobey. Apart from that, waterside co-owners who enjoy the right to use water should care for and maintain water sources. They have to bear the burden of maintenance and cleaning operations. The expenses can be taken from taxes or special funds for that purpose.

Muslims believe that ensuring social justice is the cornerstone of Islam. The recognition of water as a vital resource, to which everyone has the right to a fair share, is emphasized by the following hadith, which effectively makes water a community resource to which all, rich or poor, have a right: "*Muslims have a common share in three things, pasture, water, and fire*", as narrated by Abu Daud.

Apart from that, animals cannot be allowed to die of thirst, and the water that remains after humans quench their thirst must be given to them. The immense value of giving water to any creature is reflected by the following hadith: "*A prostitute was forgiven by Allah because she took off her shoe and tying it with her head cover and drew out some water from a well to be given to a panting dog passing by that was about to die of thirst. So Allah forgave her because of that*," as narrated by Al-Bukhari.

The Al-Quran notes that the gift of water is for flora as well: "*vegetation of all kinds*" (6:99) is nourished by rainwater that Allah sends down. This verse supports the statement that water is made available by Allah so that all life should receive support according to its needs, including humans, animals, and plants. As in Islam, humankind has the first right to resources that Allah has provided for its creation. It is well accepted by Islamic scholars (1) that the priority rights of the use of water are first the right of humans to quench their thirst; second, the right of cattle and household animals; and third the right of irrigation.

Nevertheless most Islamic scholars have concluded that individuals or groups have the right to use, sell, and recover the value added costs of water. These judgments are based totally on two hadiths: the first, about Othman, the third caliph, who purchased a well at Ruma (a place near Medina), proves that wells can be owned and traded (2) and secondly, as Prophet Muhammad said, "*It is better to go to the wood and cut and sell lumber to feed himself than to beg people for help*," narrated by Muslims. This implies that common property resources such as wood and water can be sold and traded (3).

In his time, the Prophet Muhammad discouraged the selling of water and even forbade the sale of excess water. As noted, he also encouraged Othman to buy the well at Ruma and give away its water free. These examples reflect the prophet's desire for the poor and weak to have access to water. Apart from that, he encouraged people to work for a living, even by taking wood (a free natural resource) and selling it; a number of Islamic scholars believe it should also be applied to water, where we have the right to use, trade, and sell.

As narrated by Abu Hurairah, Allah's Apostle, said, *"No doubt, you had better gather a bundle of wood and carry it on your back (and earn your living thereby) rather than ask somebody who may give you or not*." Obviously, water in its natural state, such as a river or a lake, cannot be bought or sold. However, most prominent Muslim scholars agreed that if infrastructure and knowledge have been invested to withdraw it, for instance, a treatment and distribution system to convey it to people's homes, then the water utility has the right to recover its costs.

To prevent anyone new from diminishing the water flow, Islam has underlined a principle that entitles canals or wells or and any other sources of water that entail the ownership of a certain amount of adjacent land. This is known as *harim* or forbidden area. In this area, there should be no development such as land or tree cutting; it should remain as it is. Most schools of thought in Islam proclaim that no uniform size for a *harim* can be fixed and that the rules sanctioned by local customs must be adhered to, even in a region of 3 meters for a small well up to 100 meters for a spring.

Irrigation Regulations

The general principles of irrigation are that no one may alter an established irrigation system. Water may be drawn in any way, provided this does not entail any derogation of the lawful owner's rights and privileges. The assent of the waterside dwellers must be obtained, and the irrigation rights of other parties must be considered. Lower lying plots of ground are supposed to receive water from natural flows from upstream plots. The owner of the lower plot may not erect any barrier likely to obstruct the natural flow of water. Likewise, the owner of the upper plot may not do anything likely to harm the lower plot (5).

As narrated by Abdullah bin Az-Zubair,

> *An Ansari man quarreled with Az-Zubair in the presence of the Prophet about the Harra Canals which were used for irrigating date palms. The Ansari man said to Az-Zubair, "Let the water pass," but Az-Zubair refused to do so. So, the case was brought before the Prophet who said to Az-Zubair, "O Zubair! Irrigate (your land) and then let the water pass to your neighbor." On that the Ansari got angry and said to the Prophet, "Is it because he (i.e., Zubair) is your aunt's son?" On that the color of the face of Allah's Apostle changed (because of anger) and he said, "O Zubair! Irrigate (your land) and then withhold the water till it reaches the walls between the pits round the trees." Zubair said, "By Allah, I think that the following verse was revealed on this occasion": "But no, by your Lord they can have No faith until they make you judge in all disputes between them."* (4.65)

In addition Islam proclaims water conservation, and it is considered a fixed concept of Islamic teaching. A Muslim is ordered to be economical with water even if he is taking his water from a fast flowing river. The Prophet Muhammad teaches us not to overdo even in worshipping Allah; for instance, he uses only a handful of water in taking his ablution. Further more, as narrated by Anas, The Prophet used to take a bath with one Saor up to five Mudds (1 Sa' is equal to 4 Mudds) of water and used to perform an ablution with one Mudd of water. As in bathing, he also used a single pot of water. As narrated by 'Aisha (one of his wives), The Prophet and I used to take a bath from a single pot called a 'Faraq.

CONCLUSION

lslam is the religion of all times and places. It is a perfect system of temporal values. By practicing its laws, Muslims all over the world attain spirituality and also the pinnacle of all other glories. It was proven that the religious precepts of Islam do not represent an obstacle for efficient management of water resources; on the contrary, proper use of Islamic principles could significantly enhance water management practices in the countries concerned.

Before it came to mean simply law, the Arabic word *sharia* denoted the law of water (1). Therefore, it is not surprising that a detailed examination of the Quran and the Hadith shows that Islam makes a remarkable number of specific statements about water management. It is because Islam rose and developed in a desert area where water resources were extremely important and always considered sacred. It is also connected with the nature of Islam as a monotheistic religion that sought to regulate the behavior of humans according to the commands of Allah.

Actually, there is no major contradiction between what Islam says about the laws of water and the emerging international consensus on the issue. Islamic water laws are not that unique; some of the same principles could be derived by studying other faiths, their holy books, and the lives of their prophets. The major differences existed and were practiced more than 1400 years ago; they have proven adequate, applicable, and practical in our daily life. It was noted that Islam is very flexible and can accommodate all concerns regarding natural resource management, including water. Furthermore, Islamic principles for water management are completely compatible with the Dublin International Conference on Water and Environment and other accepted principles. In addition, Islam considers the protection of biological biodiversity, as well as water quantity and quality.

BIBLIOGRAPHY

1. Mallat, C. (1995). *Water in the Middle East*. I.B Tauris, New York.
2. Kitab Al-muslim (1967). Darul Nada, Kuala Lumpur.
3. Yusof Ali, A. (1977). *The Holy Quran: Text, Translation and Commentary*. American Trust Publication for The Muslim Student Association of the United States and Canada, Plainfield, IN.

4. Zouhaili, O. (1992). *Al-Fiqh wa-Dalalatuh (Islamic Jurisprudence and its proof)*. Dar El-Machariq, Damascus.
5. Querry, A. (1872). *Recueil Des Lois Concernant Les Musulmans Schytes*. Vol. 2, Imprimerie Nationale, Paris.

READING LIST

Agwan, A.R. (1997). *Islam and the Environment*. Synergy Books International, Kuala Lumpur.

Caponera, D.A. (1973). *Water Laws In Moslem Countries*. FAO Publications 20, no. 1, Organisation. Food and Agriculture Organisation, Rome

Faruqui, N.I. (2001). *Water Management in Islam*. United Nation University Press, Tokyo.

http://www.usc.edu/dept/MSA/fundamentals/hadithsunnah.

TRANSBOUNDARY WATERS IN LATIN AMERICA: CONFLICTS AND COLLABORATION

Scott Whiteford
Alicia Jimenez
Michigan State University
East Lansing, Michigan

Tensions over freshwater resources may become more frequent as pressures on water resources grow due to increasing water scarcity, degrading water quality, rapid population growth, unilateral water development, and uneven levels of economic development. As water resources become increasingly overallocated and used, the potential for conflict between nations with competing priorities and management strategies grows. Despite many obstacles, several international organizations are proposing that instead of being zones of conflict, shared water resources can provide a basis for cooperation and benefit-sharing, provided that threats to international waters are recognized and collaborative structures created. In the largest study of transnational water conflict and cooperation, it was found that cooperative interactions between riparian states over the past fifty years have outnumbered conflictive interactions by more than two to one (1).

Acknowledging the benefits of cooperative water management frameworks, policy makers have developed guiding principles and laws for international freshwater management. On a smaller scale, regional bodies and individual governments have developed protocols and treaties governing the management and protection of specific international water bodies (2).

In the Americas, there are approximately 70 international basins shared by Latin American countries (32 in North America and 38 in South America). Of those, 15 of the North American international basins and six of the South American international basins are governed by international treaties (2). These treaties refer to agreements where there is a concern with water as a scarce or consumable resource, a quantity to be managed, or an ecosystem to be improved or maintained (see Table 1).

Although treaties and agreements of cooperation exist, countries of the Americas face a great challenge in collaborating to achieve sustainable development on both sides of their borders. This is particularly important where the actions of one nation can directly or indirectly affect the well-being of its neighbors. Frequently, projects that threaten water quality are located near international borders and create winner–loser confrontations. These situations generate serious international conflicts, sometimes violent, which may extend for years. To resolve these types of prolonged conflicts, collaborative agreements are negotiated which facilitate shared water and land management. This article presents examples of conflicting situations and cooperative efforts among Latin American countries in North America, Central America, and South America. This article examines key examples of binational development that has threatened the quality of water used by two nations and created binational conflict and negotiation.

NORTH AMERICA: MEXICO–U.S. BORDER

In the U.S.–Mexico border region, water scarcity is a growing problem heightened by drought, rapid population growth, tourism, urbanization, and agriculture. The border extends almost 2000 miles and includes multiple watersheds, diverse ecological zones, political units, and a range of economic activities. The two nations have developed a set of institutions to manage water conflict which we do not examine in this article. At the end of this section, we present a list of significant binational disputes and negotiations.

The following is a brief overview of one of major conflicts between the United States and Mexico, the conflict over water quality at the lower end of the Colorado River. Because Mexico is down stream on the Colorado River, which provides much of the water for the western United States, the quality and quality of water flowing from the United States to Mexico has been the subject of dispute and international litigation. See Table 1 for the many treaties written to govern the border water resources between these countries. Although both countries have developed institutions to negotiate the disputes over water rights and responsibilities, the challenge is almost overwhelming, given that the two countries have very different water laws, national institutions, and constituencies (3).

Water conflicts over water quality in the border region arose because of the salinity from drainage water coming into the Colorado from irrigated fields in the United States and then taken from the Colorado for irrigation in the Mexicali Valley (Mexico). This valley is located on the western area of the United States- Mexico border and is one of the richest agricultural zones in North America. The valley draws water from the Colorado River and also from the Mexicali-Mesa de Andrade aquifer that transcends both regions and which is also subject to international dispute. Until recently, the Mexican government has had the dominant role in managing water for both drinking and irrigation in Mexicali Valley; nonetheless, people in Mexicali mobilized for their water rights and put pressure on their local and national governments to negotiate with the United States to protect surface water. The negotiations led to the introduction of water quality control that resolved the crisis. Nevertheless, Mexican farmers are periodically plagued by discharges of poor quality water (3).

Water of the Colorado River is governed by international agreements, but this is not the case for groundwater which was not a source of conflict until the end of the twentieth century. It was assumed to be in great supply and of high quality, but conflicts have arisen with the paving of the irrigation canals on the American side of the border that had been recharging the aquifer pumped by farms on the Mexican side of the border. The quality of the water in the aquifer is also being questioned after recent studies that showed high levels of pesticides and salts. Binational efforts to resolve the groundwater conflicts are beginning, but cooperation between the two countries over groundwater is still being negotiated. Groundwater is not the only transnational water problem being negotiated between the United States and Mexico. Other issues include conflicts over water of the Rio Bravo/Rio Grande; pollution of the Rio Bravo, Tijuana, and Nogales by urban and industrial water; and the construction of radioactive waste dumps

Table 1. Transboundary Water Treaties in the Americas

NORTH AMERICA[b]

Basin	Total Area	Date	Treaty Basin	Signatories	Treaty Name
Artibonite	8,800 km^2	02-20-1979	Frontier or shared waters	Dominican Republic; Haiti	Treaty of peace, friendship, and arbitration between both countries.
Candelaria	12,800 km^2	04-10-1987	Frontier or shared waters	Guatemala; Mexico	Agreement between both countries on the protection and improvement of the environment in the border area
Coatan achute	2,000 km^2	04-10-1987	Frontier or shared waters	Guatemala; Mexico	Agreement between both countries on the protection and improvement of the environment in the border area
Colorado	655,000 km^2	07-16-1994	Colorado	United States of America and Mexico	Minute No. 291 of the International Boundary and Water Commission (IBWC) concerning improvements to the conveying capacity of the international boundary segment of the Colorado River.
		07-18-1985	Frontier or shared waters		Agreement of cooperation regarding pollution along the inland international boundary by discharges of hazardous substances
		08-14-1983	Frontier or shared waters		Agreement on cooperation for the protection and improvement of the environment in the border area
		08-30-1973	Colorado		Agreement on the permanent and definitive solution to the salinity of the Colorado River Basin (IBWC Minute No. 242)
		04-30-1973	Colorado		Agreement extending Minute No. 241
		07-14-1972	Colorado		Agreement Minute No. 241
		08-24-1966	Colorado		Agreement concerning the loan of waters of the Colorado river for irrigation of lands in the Mexicali Valley
		11-14-1944	Colorado, Rio Bravo del Norte, Rio Grande, Tijuana		Treaty between both countries regarding the waters of those four rivers
		11-21-1900	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889
		12-22-1899	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889
		12-2-1898	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889

(*continued overleaf*)

Table 1. (*continued*)

Basin	Total Area	Date	Treaty Basin	Signatories	Treaty Name
		10-29-1897	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889
		11-6-1896	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889
		10-1-1895	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889
		03-1-1889	Colorado, Rio Grande		Convention on boundary waters
Grijalva	126,800 km^2	04-10-1987	Frontier or shared waters	Guatemala;Mexico	Agreement on the protection and improvement of the environment in the border area
Hondo	14,600 km^2	04-10-1987	Frontier or shared waters	Guatemala; Mexico	Agreement on the protection and improvement of the environment in the border area
Lempa	18,000 km^2	04-9-1938	Lempa, Paz	El Salvador; Guatemala	Treaty for the delimitation of the boundary between both countries
Massacre	800 km^2	02-20-1929	Frontier or shared waters	Dominican Republic; Haiti	Treaty of peace, friendship, and arbitration
Paz	2,200 km^2	04-9-1938	Lempa, Paz	El Salvador, Guatemala	Treaty for the delimitation of the boundary between both countries
Pedernales	400 km^2	02-20-1929	Frontier or shared waters	Dominican Republic, Haiti	Treaty of peace, friendship, and arbitration
Rio Bravo/Rio Grande	656,100 km^2	11-13-1992	Rio Grande	Mexico; United States of America	IBWC Minute No. 289—observation of the quality of the waters along the U.S. and Mexico border
		11-10-1987	Rio Bravo-Rio Grande		Boundary waters agreement
		07-18-1985	Frontier or shared waters		Agreement of cooperation regarding pollution along the inland international boundary by discharges of hazardous substances
		08-14-1983	Frontier or shared waters		Agreement on cooperation for the protection and improvement of the environment in the border area
		10-24-1960	Rio Grande		Agreement to proceed with the construction of Amistad Dam on the Rio Grande
		11-14-1944	Colorado, Rio Bravo del Norte, Rio Grande, Tijuana		Treaty between both countries regarding the waters of those four rivers
		02-1-1933	Rio Bravo/Rio Grande		Convention on the rectification of the Rio Grande (Rio Bravo del Norte)
		05-21-1906	Rio Grande		Convention on the distribution of waters of Rio Grande
		11-21-1900	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889
		12-22-1899	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889
		12-2-1898	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889
		10-29-1897	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889
		11-6-1896	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889
		10-1-1895	Colorado, Rio Grande		Boundary waters, extension of convention 03-1-1889

Table 1. (***continued***)

Basin	Total Area	Date	Treaty Basin	Signatories	Treaty Name
		03-1-1889	Colorado, Rio Grande		Convention on boundary waters
San Juan	42,200 km^2	03-22-1888	Frontier or shared waters	Costa Rica; Nicaragua	Award of the President of U.S. on the validity of the treaty of limits of 15 April 1858
Suchiate	1,600 km^2	04-10-1987	Frontier or shared waters	Guatemala; Mexico	Agreement on the protection and improvement of the environment on the border area
Tijuana	4,400 km^2	07-2-1990	Rio el Alamar, Tijuana	Mexico, United States of America	Boundary waters: San Diego/Tijuana. IBWC Minute No. 238
		07-18-1985	Frontier or shared waters		Agreement of cooperation regarding pollution of the environment along the inland international boundary by discharges of hazardous substances.
		08-14-1983	Frontier or shared waters		Agreement on cooperation for the improvement of the environment in the border area
		11-14-1944	Colorado, Rio Bravo del Norte, Rio Grande, Tijuana		Treaty between both countries regarding the waters of those four rivers
Yaqui	74,700 km^2	07-18-1985	Frontier or shared waters	Mexico; United States of America	Agreement of cooperation regarding pollution of the environment along the inland international boundary by discharges of hazardous substances
		08-14-1983	Frontier or shared waters		Agreement of cooperation for the protection and improvement of the environment in the border area

SOUTH AMERICA[a]

Basin	Total Area	Date	Treaty Basin	Signatories	Treaty Name
Amazon	5,866,100 km^2	08-2-1988	Beni, Madeira, Mamoré	Bolivia; Brazil	Exchange of notes constituting an agreement for the construction of a hydroelectric plant in Cachuela Esperanza, supplementary to the agreement on economic and technical cooperation
		02-8-1984	Beni	Bolivia, Brazil	Agreement concerning the Cachuela Esperanza hydroelectric plant
		07-3-1978	Amazon	Bolivia, Brazil, Colombia, Ecuador, Guyana, Peru, Surinam, Venezuela	Treaty for Amazonian cooperation
		05-22-1944	Amazon, Chira, Tumbes, Zarumilla	Ecuador, Peru	Declaration and exchange of notes concerning the termination of the process of demarcation of the Peruvian-Ecuadorean frontier
Chuy	200 km^2	03-11-1991	Not specified	Brazil, Uruguay	Complementary agreement to the basic scientific and technical cooperation agreement between both countries on the area of water resources

(*continued overleaf*)

Table 1. (*continued*)

Basin	Total Area	Date	Treaty Basin	Signatories	Treaty Name
		12-20-1933	Frontier or shared waters	Brazil, Uruguay	Convention regarding the determination of the legal status of Brazil and Uruguay frontier
La Plata	2,954,500 km^2	05-6-1997	Cuareim	Brazil, Uruguay	Complementary settlement to the agreement of cooperation for the use of natural resources and development of this basin
		09-2-1995	Pilcomayo	Argentina, Bolivia, Paraguay	Agreement constituting the trilateral commission for the development of this basin
		06-9-1995	Bermejo, Grande de Tarija	Argentina, Bolivia	Agreement for the multiple uses of the resources of the upper basin of the Bermejo River and Grande de Tarija River
		03-11-1991	Cuareim	Brazil, Uruguay	Agreement of cooperation for the use of natural resources and development of this basin
		03-11-1991	Not specified	Brazil, Uruguay	Complementary agreement to the basic scientific and technical cooperation agreement on water resources
		06-29-1983	Pepiri-Guazu, Uruguai	Argentina, Brazil	Decree No. 88.441, agreement for water resources exploitation within the Uruguay and Pepiri-Guaco Rivers
		05-17-1980	Pepiri-Guazu, Uruguai	Argentina, Brazil	Treaty for the development of the water resources contained in the border reaches of Uruguai and Pepiri-Guazu Rivers
		10-19-1979	Paraná	Argentina, Brazil, Paraguay	Agreement on Paraná River projects
		02-26-1975	Uruguay	Argentina, Uruguay	Statute of the Uruguay River
		04-26-1973	Iguassu, Paraná	Brazil, Paraguay	Treaty concerning the hydroelectric utilization of the Paraná River owned in condominium by the two countries
		04-7-1961	Uruguay	Argentina, Uruguay	Treaty on the boundary constituted by the Uruguay River
		01-23-1958	Paraná	Argentina, Paraguay	Agreement concerning the study of the utilization of the water power of the Apipe Falls
		01-20-1956	Acaray, Monday	Brazil, Paraguay	Agreement concerning a study on the utilization of the water power of the Acaray and Monday Rivers
		12-30-1946	Uruguay	Argentina, Uruguay	Agreement concerning the utilization of the rapids of the Uruguay River in the Salto Grande area
		06-1-1945/07-5-1939	Pilcomayo	Argentina, Paraguay	Supplementary boundary treaties
		12-20-1933	Frontier	Brazil, Uruguay	Convention for determination of the legal status of the frontier of both countries
		01-5-1910	Plate	Argentina, Uruguay	Protocol dealing with question of the jurisdiction of the River Plate

Table 1. (*continued*)

Basin	Total Area	Date	Treaty Basin	Signatories	Treaty Name
Lagoon Mirim	55,000 km^2	03-11-1991	Not specified	Brazil, Uruguay	Complementary agreement to the basic scientific and technical cooperation agreement on water resources
		07-7-1977	Juguarao, Mirim Lagoon		Treaty on cooperation for the utilization of the natural resources and the development of the Mirim Lagoon basin
		04-26-1963	Mirim Lagoon		Agreement on establishing an joint commission for the development of the Mirim Lagoon
		12-20-1933	Frontier		Convention on the determination of the legal status of the frontier between Brazil and Uruguay
Lake Titicaca-Poopo system	111,800 km^2	06-21-1993	Desaguadero, Poopó, Salar de Coipasa	Bolivia; Peru	Notas related to the creation of the Autonomous Binational Authority of the basin
		02-19-1957	Lake Titicaca		Agreement concerning a preliminary economic study of the joint utilization of Lake Titicaca waters
		07-30-1955	Lake Titicaca		Preliminary convention concerning a study of the joint utilization of Lake Titicaca
		04-20-1955	Lake Titicaca		Exchange of notes establishing a joint commission for study of the Puno−Guaqui railway line and joint use of Lake Titicaca
		07-17-1935	Lake Titicaca		Preliminary convention for exploitation of fisheries in Lake Titicaca
		06-3-1929	Arica, Tacna	Chile, Peru	Treaty between both countries for the settlement of the dispute regarding Tacna and Arica
Maroni	65,000 km^2	09-30-1915	Maroni, Marowinjine	France, Netherlands (French Guiana, Suriname, and Brazil)	Convention between France and the Netherlands to fix the boundary between Suriname and French Guiana

[a] *Source*: Reference 20.

[b] Includes Central America and the Caribbean.

in Sierra Blanca, Texas; Ruidoso, New Mexico, a toxic waste incinerator in Ciudad Juarez; and several toxic waste dumps in the border region of south Texas. All of these projects threaten the border environment. Meanwhile treatment plants have been built along the border. Major sources of pollution such as the Rio Nuevo, that flows from Mexicali to the Salton Sea in California, are now being treated and are returning a much better quality of water to the stream. Conflicts, negotiations, and cooperation will continue to be the pattern along this border for years.

CENTRAL AMERICA

Despite the importance of transnational rivers in Central America, they have not been the subject of extensive diplomacy. Dam construction projects on the Usumancinta, Salinas, and Pasión Rivers on the Mexico-Guatemala border have generated both conflicts and cooperation. The Mexican and Guatemala governments have signed a memorandum of energy cooperation which has created great concern among people who live in communities along the rivers and are afraid of being displaced by their governments by the construction of new and bigger dams (4). Community leaders have protested the lack of information about future construction plans, especially in the context of the Plan Puebla-Panama. The Plan Puebla-Panama implementation, begun in 2001, framed economic integration and sustainable development for Central America and the southern states of Mexico (5). Funding from the World Bank and the Inter-American Development Bank for the projects includes building dams on the border rivers. In response to such controversial plans, 98 nongovernmental organizations (NGOs) from 21 countries of the Americas and Europe and affected communities have created a regional movement against dams (6). The concern is focused on forced relocation of communities, the destruction of valuable archaeological sites, and loss of biodiversity (4).

Other major transboundary rivers in Central America include the Rio Hondo between Mexico and Belize and the San Juan River that divides Costa Rica and the San Juan. Heavy use of pesticides in Belizean agriculture that are killing fish and impacting coral reefs are creating concern on the Mexican side of the river. Another type of transboundary river problem, navigation rights, is reflected in a dispute between Costa Rica and Nicaragua over the San Juan River. The conflict started in 1998, when Arnoldo Aleman, a former Nicaraguan president used police to impose navigation fees on Costa Rican citizens using the river on the Nicaraguan side. It was only by the intervention of the Costa Rican government that the fees were dropped, much to the relief of people living on both sides of the river (7).

SOUTH AMERICA

Transboundary water management in South America is very challenging. The region includes one of the largest river basins in the world, the Amazon River Basin. The region, known as Amazonia, includes more than 8000 km of borders, including eight South American countries. Its enormous natural, cultural, biological, hydrological, mineral, forest, and medical resource richness has made it an object of intensive exploitation that has endangered the natural balance and health of the basin.

Another important international basin is the Orinoco, located in the south part of this basin where Venezuela and Guyana are located. The Orinoco is affected by gold mining. In 1995, more than one million liters of cyanide waste were spilled in the Omai and Esequibo Rivers in Guyana. The impact of these activities and the contamination of the zone's hydrologic network affect at least 25% of the basin's indigenous population, including the following ethnic groups: Pemon, Yanomami, Piaroa, Guahibo, Yekwana, and 17 other indigenous peoples. There are no treaties or agreements of cooperation on water quality between the countries that share this basin.

In addition to the Amazon, other rivers in South America play an important role in linking as well as dividing countries in northern South America. The Meta River that divides Colombia and Venezuela has supported transportation of produce for centuries (8). Sediments, caused by watershed degradation, create serious difficulties for transportation on the Meta River reducing its capacity for transportation. With funding from the Inter-American Development Bank, a binational dredging project of the Meta River has begun to increase trade by reducing transportation cost, but the project has been opposed by many stakeholders and conservation groups on the grounds that no impact studies were carried out and local populations had little input to the project.

Sovereignty and control over who uses rivers (sailing rights) may be a source of conflict and/or cooperation. Sailing rights in waters controlled by multiple countries are a challenge for integration. An example of this is the Orinoco-Amazonas fluvial interconnection project, which has been promoted by Venezuela's and Brazil's presidents, but integrating the Colombia has been difficult. Colombia wants free sailing rights on the Orinoco, without the approval of Venezuela and Brazil, to facilitate transportation and trade. The Orinoco–Meta integration is important for Colombia because it would provide a corridor between the Pacific and Atlantic Oceans. Both Venezuela and Brazil are reluctant to give Colombia free access and have recommended doing the fluvial interconnection gradually and in specific phases. To open the waterway to commercial vessels will require considerable dredging and construction in Colombia and cooperation from its neighbors (9).

La Plata Basin, which is shared by five countries, Brazil, Argentina, Paraguay, Bolivia, and Uruguay, is also under threat. The Pilcomayo river, part of this basin shared by Argentina, Bolivia, and Paraguay, has been contaminated periodically with arsenic, cadmium, and lead caused by spills of toxic waste containers from Bolivian mining. In the trinational zone, the Pilcomayo is an important source of water for indigenous peoples, agriculture, and its unique wildlife. There are several cooperative agreements for the water management of this river (see Table 1).

The process to build a hydro-via, a water highway, using the Parana and Paraguay Rivers is in its final discussion phase (10). The Paraguay-Parana River Hydro-Via is a plan of the five governments of the La Plata Basin to build 2100 miles of river and canal systems navigable by large barges. The project would significantly reduce the transportation costs of commodities within and beyond the region. For example, it would reduce the cost of transportation of soybeans exported to Europe from South America by half. However, the project was halted in 1997 due to the Rios Vivos Coalition's efforts to demonstrate the irreversible effects on this unique ecosystem, the Pantanal. The Pantanal is the world's largest tropical wetland ecosystem formed by the Parana-Paraguay rivers (11).

The Pantanal is an ecological wonder shared by Bolivia, Brazil, and Paraguay. The ecosystem of the region is based on seasonal flooding of the major rivers, including the transborder rivers; yet there is no trinational agreement to protect the ecosystem and the people who depend upon it. This unique region has serious problems caused by deforestation, expansion of agriculture, contamination of the water by pesticides and derivatives of combustible alcohol production, and illegal hunting of endangered species. There has been an increase in natural habitat loss and a decline in the quality of life

South America has major several shared hydroelectrical dams; over the Parana River are two: Yacyretá (Argentina and Paraguay) and the biggest dam in the world, Itaipú (Brazil and Paraguay). Another is Salto Grande, shared by Argentina and Uruguay over the Uruguay River (10). Five more international dams on the Uruguay and Parana border rivers have been under study.

The Yacyreta binational dam has had great impact on native and rural populations, forcing them to move away from their homes and farms. This mobilization has affected the health of these populations, increasing their risk of contracting malaria and other tropical diseases. The Itapu dam has also had massive social and environmental impacts. One of the most significant impacts has changed the weather of the region periodically creating semihurricane conditions. The building of dams and sharing of electricity reflect a high level of cooperation between the countries that share borders.

South America economic integration and the new infrastructure, including highways and hydroelectric dams in the southern region of South America, have also generated conflicts over transboundary waters. One criticism of the integration process is the lack of wide stakeholder participation and the lack of information to the civil society about these efforts. The projects mentioned before, the Meta River canalization, the Orinoco–Amazonas fluvial interconnection, the Venezuelan–Brazilian electrical interconnection, and the South Cone projects, among others, are usually perceived as local or isolated national efforts, without taking into account their continental scope. In addition, according to Arocha (12), most of the new infrastructural projects will produce considerable disturbances in the natural and social environment, and most likely, the indigenous population will be affected the most.

Several organizations, such as the United Nations Global Environment Facility (GEF) propose a comprehensive, ecosystem-based approach to managing transboundary waters and their drainage basins; these are often called "international water projects"(1). The goal of these organizations is to build transboundary collaboration projects to help each of the countries use the full range of technical, economic, financial, regulatory, and institutional measures needed to operationalize sustainable use strategies for transboundary waterbodies and contributing basins.

One example of a GEF international water project in Latin America is the Bermejo River Basin of Bolivia and Argentina. The GEF became interested in this basin because it has excessive levels of sediment, important transboundary biodiversity, and a binational commission for developing of the basin. One of the objectives of GEF was the involvement of stakeholder groups in the basin to determine their sustainable development future. They facilitated participatory transboundary diagnostic analysis (TDA), a scientific-based assessment of the key transboundary problems and their root causes. Originally, 16 dams were proposed for hydropower production in the basin. This was quite controversial. However, through public participation in the project, the binational commission that was created to develop hydropower transformed itself into a force to consider sustainable development options. Three dams will still be pursued, but the focus on the transboundary environment, strategic processes for participation, and practical measures consistent with sustainable development helped overcome initial transboundary concerns and NGO opposition (1).

The future of transnational rivers in Latin America and their role in enhancing sustainable development depends on binational collaboration. The watershed and basin agreements in place have generally facilitated negotiation and found solutions of mutual benefit. In other transborder water situations, mutually agreed upon conflict resolution mechanisms and management plans have yet to be developed. Until these agreements are in place, binational management will continue to generate conflict. At stake is the development of some of the most important yet vulnerable areas on the continents. Fortunately, progress is being made on almost every border.

BIBLIOGRAPHY

1. Uitto, J.I.D. and Alfred, M. (2002). Management of transboundary water resources: Lessons from international cooperation for conflict prevention. *Geogr. J.* **168**(4): 365–378.
2. UNEP, FAO, Oregon State University. (2002). *Atlas of International Freshwater Agreements*. Kenya; USA.
3. Whiteford, S. and Cortez, A. (in press). Good to the last drop: the political ecology of water and health on the border. In: *Globalization, Water and Health*. L. Whiteford and S. Whiteford (Eds.). School of American Research Press, Santa Fe, NM.
4. Mesoweb. (2002). Maya sites to be flooded by Dam, Mesoweb Reports. Available: http://www.mesoweb.com/reports.dam.html.
5. Delgado, G.C. (2002). El Corredor Biologico Mesoamericano, Tres Ecorregiones y COINBIO: esquemas del Banco Mundial integrados a la politica nacional., Biodiversidadla.org.
6. Bellinghausen, H. (2002). During the Mesoamerican forum, they demanded water, light and land for rural villages. *La Jornada*. Mexico D.F.
7. Gonzalez, A.H. and Mauricio. (2002). Acuerdo evito choque bilateral. *La Nacion*. San Jose.
8. Colmenares, G. (1997). *Tratamiento binacional de asuntos ambientales en regiones fronterizas*. La experiencia Colombo-Venezolana. CEFIR.
9. Nacional, E. (2000). Recomiendan realizar por etapas la integracion de Orinoco-Amazonas. *El Nacional*.
10. Gudynas, E. (1998). *La dimension ambiental de la integracion fisica y energetica en el Cono Sur*. Uruguay. CLAES.
11. IRN. (2001). *Stop the Hidrovia in the Pantanal*. International Rivers Network (IRN).
12. Leon-Oliveros, R. (1997). *Situacion actual y perspectivas de las relaciones fronterizas de los paises miembros del grupo Andino*. CEFIR.
13. Arocha, M. (2001). Ambiente, proyectos de integracion fisica en America Latina y desarrollo sustentable en Venezuela. Analitica.com.

UNITED STATES-MEXICO BORDER WATERS: CONVENTIONS, TREATIES, AND INSTITUTIONS

NANCY A. LOWERY
San Diego, California

SELECTED TREATIES AND CONVENTIONS

Cooperation over shared United States–Mexico waters dates back to the *Treaty of Peace, Friendship, Limits, and Settlements*, know as the Treaty of Guadalupe Hidalgo, signed in 1848. This treaty established the international boundary from the Gulf of Mexico along the Rio Grande to El Paso, Texas, then west to the Pacific Ocean just south of San Diego Harbor, a stretch of 3141 km. For 2001 km of that border, the Rio Grande serves as the boundary line. This is a long and complex boundary with topography ranging from coastal plains to mountain ranges and high desert. The climate is predominately semiarid to arid, has two major watersheds, the Rio Grande (called the Rio Bravo in Mexico) and the Colorado River, many smaller watersheds, and numerous groundwater basins. This was a treaty of intent, not of specifics. The treaty did not deal with details of the boundary line, management of water between the two countries, fluctuation of the river courses, water quality, or groundwater resources. Subsequent agreements would have to deal with these issues.

The introduction of the railroad and irrigated agriculture in the late 1800s spurred settlement, which required that questions left unanswered by the 1848 Treaty be resolved or at least addressed. Therefore several conventions were enacted to solve disputes over the boundary line, establish surveys, and fix boundary monuments to officially demarcate the border (1). These include the *Gadsden Treaty of December 30, 1853* and the *Convention of July 29, 1882* (2).

Changes in the course of the Rio Grande and Colorado rivers and the resulting fluctuations in the international boundary were dealt with by the *Convention of November 12, 1884*. This agreement established rules for maintaining the boundary line, given changes in the riverbed. The *Convention to Avoid the Difficulties Occasioned by the Reason of the Changes which Take Place in the Beds of the Rio Grande and Colorado River*, March 1, 1889, established a binational commission, the International Boundary Commission (IBC). The IBC's role was to apply the 1884 Convention rules and handle disputes related to boundary waters. The IBC had both a United States and a Mexican Section; each had a separate commissioner appointed by the federal governments of the respective countries.

None of these previously mentioned agreements dealt with the apportionment of waters between the two countries. By the end of the nineteenth century, however, allocation issues could no longer be ignored. The growth of irrigated agriculture on the upstream reaches of the Rio Grande meant that in low flow years, farmers on the Mexican side of the border did not receive sufficient water to irrigate crops. By the 1890s, Mexico was pressing for a solution to this problem, especially where agricultural pressures were at their greatest, which was in the section of the Rio Grande that forms the border between New Mexico, Texas, and Chihuahua.

To alleviate stress in that region, the May 12, 1906 *Convention Providing for the Equitable Distribution of the Waters of the Rio Grande for Irrigation Purposes* allocated the waters of the Rio Grande above Ft. Quitman, Texas, between the United States and Mexico. It guaranteed Mexico 74 Mm^3 (60,000 acre-feet) to be delivered at a point just above Ciudad Juárez, Chihuahua, in accordance with a monthly schedule. In exchange for that delivery, Mexico waived any additional claims to the waters from the upper stretch of the river to Ft. Quitman, Texas.

All water allocated under the 1906 Convention is for irrigation; none is set aside for municipal or industrial uses or for instream flows. As water demand has changed, this allocation structure has had ramifications for the region. The Convention also did not deal with water quality issues or apportion groundwater supplies. It did, however, provide for drought or catastrophic conditions. Stating that should it be impossible to deliver the quantity of water promised due to severe drought or destruction of the delivery system, Mexico's supply would be diminished in the same proportion as the supply delivered to U.S. irrigators.

In 1933, the two governments enacted the *Convention for the Rectification of the Rio Grande (Rio Bravo del Norte) in the El Paso–Juárez Valley*. This convention established the Rio Grande Rectification Project to manage the river in the increasingly developed Mesilla–El Paso–Juárez Valleys for flood control and delivery of treaty waters. It provided for joint construction, operation, and maintenance, by the IBC, of the facilities needed for these purposes, which entailed straightening a 155-mile stretch of the river and was the first joint construction treaty between the United States and Mexico.

Continued border growth in the early 1900s caused the lower Rio Grande basin and the Colorado River to experience water demand pressures similar to those that had been felt in the upper Rio Grande basin 50 years earlier, which led to the 1944 *Treaty Relating to the Utilization of Waters of the Colorado and Tijuana Rivers and of the Rio Grande*, considered the most important United States–Mexico water treaty. This treaty expanded the scope of cooperation between the two countries to include issues related to border sanitation and, more importantly, provided the legal framework for resolving water disputes between the two countries. It replaced the IBC with the International Boundary and Water Commission (IBWC), called the Comision Internacional de Limites y Aguas (CILA) in Mexico.

The new IBWC had expanded scope and authority in line with the expanded scope of the 1944 Treaty. It has the status of an international body, has two Commissioners, appointed by the presidents of each country, and permanent offices in El Paso, Texas, and Juárez, Chihuahua. It operates under the Departments of State of each country and has the authority and responsibility to negotiate and carry out agreements between the two countries regarding border water and sanitation, as outlined in the treaties. These agreements are called Minutes. Once approved by the United States and Mexican governments, Minutes are binding on those governments. As of March 2004, there are 311 Minutes (3). The Minute process has been an effective means to address serious border water issues, such as salinity, untreated effluent discharges, and drought mitigation, without having to renegotiate existing treaties or negotiate new ones.

Additionally, the 1944 Treaty (4) sets out, in Article 3, priorities for the joint use of the international waters, which would be used if the Commission were asked to make a determination of priorities. In order of preferential use, they are

- domestic and municipal uses,
- agriculture and stock raising,
- electric power,

- other industrial uses,
- navigation,
- fishing and hunting, and
- any other beneficial uses that may be determined by the Commission.

The 1944 Treaty also apportioned the shared surface waters, not already divided, between the two countries. Waters of the Rio Grande–Rio Bravo from Fort Quitman to the Gulf of Mexico are apportioned in the following manner:

Mexico has the right to:

- all of the waters reaching the main channel of the Rio Grande from the San Juan and Alamo Rivers, including the return flows from the lands irrigated from those two rivers;
- two-thirds of the flow of the main channel of the Rio Grande measured from the Conchos, San Diego, San Rodrigo, Escondido, and Salados Rivers and the Las Vacas Arroyo, subject to the U.S. right to an average 431,721,000 cubic meters (350,000 af/yr) in cycles of 5 consecutive years; and
- one-half of all other flows in the main channel of the Rio Grande downstream from Ft. Quitman not specifically allocated under this Treaty to either of the two countries.

The United States has the right to:

- all of the waters reaching the main channel of the Rio Grande from the Pecos and Devils Rivers; Goodenough Spring; and Alamito, Terlingua, San Felipe, and Pinto Creeks;
- one-third of the flow reaching the main channel of the river from the Conchos, San Diego, San Rodrigo, Escondido, and Salados Rivers and the Las Vacas Arroyo, providing that this third shall not be less than 431,721,000 cubic meters annually (350,000 af/yr) as an average amount in cycles of 5 consecutive years; and
- one-half of all other flows in the main channel of the Rio Grande downstream from Ft. Quitman not specifically allocated under this Treaty to either of the two countries.

In the case of the Colorado River, under the 1944 Treaty, Mexico is to receive 1,850,234,000 cubic meters annually (1.5 million af/yr), to be delivered in accordance with schedules provided in advance by Mexico. Mexico is also assigned any surplus waters not necessary for uses in the United States arriving at the Mexican point of diversion. The total annual amount of water delivered to Mexico is not to exceed 2,096,931,000 cubic meters (1.7 million af/yr).

The amount of Colorado River water allocated to the United States is not specifically mentioned in the 1944 Treaty. Those allocations are set out in the *Colorado River Compact of 1922*. The compact allocates 18,495,000,000 m^3 year (15 million acre-feet) to the compact States of Arizona, California, Colorado, Nevada, New Mexico, Utah, and Wyoming. The Compact, which came into effect before the 1944 Treaty, maintains in Article III that if the need arises for waters of the Colorado to be delivered to Mexico, than Mexico's share would come from excess flows. If surplus waters are unavailable, then the deficit in flows is to be shared equally by the states. Because of water variability and increasing urban and agricultural reliance on the Colorado River, supplying the allocated waters has become a problem (5). The 18.5 Bm^3 allocated under the Compact and the 1.8 Bm^3 assigned under the Treaty are higher than the annual average flow of the river, which is approximately 16.6 Bm^3 or 13.5 million acre feet (6).

In allocating Tijuana River waters, provisions for distribution are not laid out. The 1944 Treaty recommends, however, that the Commission investigate all means for equitable distribution of the waters between the two countries. This investigation and all other exploration of border surface water issues would be carried out by the newly formed IBWC.

Finally, the 1944 Treaty allows for the construction of dams, delivery structures, and flood control structures that are necessary to fulfill treaty obligations. It determined that these structures would be jointly constructed, maintained, and operated, and the costs to each country would be equal to the benefits provided. Since 1944, the IBWC has been responsible for constructing, maintaining, and operating two large dams and storage reservoirs, Falcon and Amistad, numerous diversion and delivery structures, international wastewater treatment plants, and a surface water salinity treatment plant.

On August 29, 1963 the two governments signed the *Convention for the Solution of the Problem of the Chamizal* to resolve a 100-year-old boundary and land dispute between the United States and Mexico. The Rio Grande channel in the area of El Paso, Texas, and Juárez, Chihuahua, had shifted periodically over the centuries, causing approximately 600 acres to move back and forth between the north and the south banks of the river. The Convention settled this dispute by repositioning and stabilizing 4.4 miles of the river, which transferred 437.18 acres from the United States to Mexico.

In 1970, the *Treaty to Resolve Pending Boundary Difference and Maintain the Rio Grande and the Colorado River as the International Boundary between the United States and Mexico* was enacted. This Treaty provides for the maintenance of the boundary line due to shifting river courses, erosion, or other changes in the channel. It also resolved all pending boundary disputes between the two countries. The IBWC is responsible for carrying out the provisions of the Treaty.

SELECTED BINATIONAL AGREEMENTS

There are several other agreements that were signed by the United States and Mexico that are relevant to managing water in the border region. These include the La Paz Agreement and the side agreements of the *North American Free Trade Agreement*.

The *Agreement between the United States of America and the United Mexican States on Cooperation for the*

Protection and Improvement of the Environment in the Border Area (La Paz Agreement) was signed on August 14, 1983. The La Paz Agreement is the framework for cooperation between the United States and Mexico on protecting, improving, and conserving the environment in a 100-kilometer area on either side of the border. This Agreement deals with issues of water, soil, and air pollution and provides for a system of emergency notification.

Five technical annexes were established to deal with specific projects of environmental concern, two of which are water related. Annex I, *Solution of the Border Sanitation Problem at San Diego, California and Tijuana, Baja California*, and Annex II, *Regarding Pollution of the Environment along the Inland International Boundary by Discharges of Hazardous Substances*, were both enacted in 1985. The framework set up by the La Paz Agreement paved the way for future environmental cooperation between the two countries. The 1992–1994 Integrated Border Environment Program (IBEP), the Border XXI Program, and the Border 2012 Program were developed in conformance with the *La Paz Agreement*.

The *North American Free Trade Agreement* (NAFTA) among Canada, United States, and Mexico became effective on January 1, 2004. NAFTA uniquely included side agreements on the environment. These agreements established a Commission on Environmental Cooperation (CEC) among the three signatories. CEC's role is to protect the environmental regulations of each country and to allow individuals and governments to file complaints with the Commission if they believe environmental laws are not being enforced.

The NAFTA side agreements also established the Border Environment Cooperation Commission (BECC) and the North American Development Bank (NADBank). BECC and NADBank are international organizations whose goal is to protect and improve the environment of the United States–Mexico border. BECC's mandate is to certify water, wastewater, and solid waste projects within the border zone. This area is defined using the La Paz Agreement boundaries of 100 kilometers on either side of the border. Certification is based on a set of engineering standards and economic and environmental sustainability concepts. NADBank's role as a development bank is to fund projects certified by BECC. Both of these institutions have played an integral role in water and sanitation infrastructure and management in the border region; due to the certification process, projects cannot be built without BECC approval.

CURRENT SITUATION

The complex set of Treaties, Convention, Compacts, and side agreements that laid the foundation for water allocation and management in the United States–Mexico borderlands have served the region well. Rapid growth, however, has placed stress on both the resource and the legal and institutional structures used to govern it. Most, if not all, border water resources are plagued by increased demand and diminishing supply. In 2000, the border population was approximately 11.8 million (7), primarily living in 14 sets of twin cities stretched out along the border. Based on annual growth rates of between 3% and 5%, that population is expected to reach 19.4 million by 2020 (8).

Current water management on the U.S.–Mexico border, therefore, is defined by three realities: first, seasonal and annual variability of water supply; second, overallocation of that supply; and third, increased competition for the resource due to burgeoning growth and changing priorities. Demand in the early and mid-1900s was sufficiently low that, except in times of severe drought, supply was adequate to meet demand. Growth and changing demand patterns have altered the balance and strained the institutional mechanisms that regulate and manage shared waters. Issues of how to deal with persistent drought, the allocation and joint management of groundwater basins, and environmental concerns surrounding degraded and endangered habitat and species have stretched the limits of the treaties and their institutions. The current debate centers around the question of whether the treaties in force have the flexibility to deal with emerging issues, whether they should be significantly altered, or if new treaties should be enacted (9–12).

BIBLIOGRAPHY

1. Leon Metz provides a detailed description of boundary disputes and surveys in Metz, L. C. (1989). *Border: The U.S.-Mexico Line*. Mangan Books, El Paso, TX.
2. For a complete list of treaties, see United States Department of State, O. o. t. L. A. (2003). *Treaties in Force: A List of Treaties and Other International Agreements of the United States in Force on January 1*, 2003. Washington, DC.
3. A description of the Minutes approved since 1946 can be accessed at www.ibwc.state.gov.
4. corpauUnited States Department of State. (1944). *Treaty Relating to the Utilization of Waters of the Colorado and Tijuana Rivers and of the Rio Grande*. 59 Stat. 1219; TS 994; 9 Bevans 116; 3 UNTS 313.
5. The limited ability of the 1944 Treaty to deal with drought is expertly outlined by Stephen Mumme and Ismael Aguilar Barajas in "Managing border water to the year 2020" in *The U.S.-Mexican Border Environment: Binational Water Management Planning*.
6. Dettinger, M.D. (1995). Coping with severe and sustained drought in the Southwest. *Water Resour. Bull.* October.
7. For an explanation of population and economic growth trends, see Peach, J. and Williams, J. (2000). Population and economic dynamics on the U.S.-Mexican border: Past, present and future. *The U.S.-Mexican Border Environment: A Road Map to a Sustainable 2020*. P. Ganster, San Diego State University Press, San Diego.
8. United States Environmental Protection Agency and Secretaría del Medio Ambiente y Recursos Naturales. (2003). Border 2012: U.S.-Mexico Environmental Program.
9. Lopez, M. (1997). Border tensions and the need for water: An application of equitable principles to determine water allocation from the Rio Grande to the United States and Mexico. *Georgetown Int. Environ. Law Rev.* **9**(2): 489.
10. Mumme, S. (2000). Minute 242 and beyond: Challenges and opportunities for managing transboundary groundwater on the Mexico–U.S. border. *Nat. Resour. J.* **40**(2): 341–378.

11. Brown, C., Castro, J.L. et al. (2003). Comparative analysis of transborder water management strategies: Case studies on the United States–Mexico border. *The U.S.–Mexican Border Environment: Binational Water Management Planning*. S. Michel (Ed.). San Diego State University, San Diego.
12. Mumme, S. and Aguilar Bajaras, I. (2003). Managing border water to the year 2020: The challenge of sustainable development. *The U.S.-Mexican Border Environment: Binational Water Management Planning*. S. Michel (Ed.). San Diego State University, San Diego, pp. 51–94.

NEGOTIATING BETWEEN AUTHORITY AND POLLUTERS: AN APPROACH TO MANAGING WATER QUALITY

PETR ŠAUER
ANTONÍN DVOŘÁK
University of Economics
Prague, Czech Republic

INTRODUCTION

Water pollution reduction is one of the most complicated challenges of environmental policy. Water quality engineers are able to offer many technical solutions. Some are based on “end-of-pipe” technologies; some represent “cleaner production” solutions. Solutions exist that are already economically efficient for the polluters. Many cleaner production projects are of this type. However, many projects exist that bring net control costs for the polluters, so the polluters do not have an economic interest in them. As these projects can still be efficient from the social point of view (if negative externalities would be internalized), it is important to change the institutional framework of environmental policy. The best institutions are those that not only satisfy ecological efficiency but also support economically efficient solutions, or at least cost-effective solutions, to the given problem. Our contribution is to present a relatively new approach that could lead to cost-effective solutions of water pollution reduction problems. It is based on voluntary negotiations between an authority and polluters where economic tools of environmental policy (environmental fees and subsidies) are used at the same time. The principle of the approach is explained in the and an illustration of an economic laboratory experiment follows.

PRINCIPLE OF THE APPROACH DESCRIBED

The suggested approach is based on (mostly multiple rounds) negotiations between an authority of environmental policy and economic subjects—polluters. The binding agreements are the results of the negotiations. The agreements contain: (final) commitments of the polluters to reduce pollution, commitments of the authority to contribute negotiated financial resources to pollution reduction projects, and a set of consequences in the case of noncompliance with the agreement conditions. If the water pollution charges are introduced, then there must be a commitment of the polluters to pay it, unless the charge has not been already introduced by a law, i.e., the charge already exists. NGO’s and various independent experts (like public health experts, ecologists, environmental engineers) can also take part in this procedure.

The positions and performance of the parties taking part in the negotiations are as follows [for the formalized procedure see (1)]:

The polluters have technical solutions (projects) that can reduce their pollution. They know the true information about their net costs of environmental protection but keep it confidential, which creates a so-called information asymmetry between the polluters and the authority. The fact that the polluters can keep the private information confidential and still achieve a cost-effective outcome seems to be one of the important privileges of the approach. During the negotiation process, they apply their own strategies to their benefits maximization when requesting (financial) contributions from the authority and offering water pollution reduction in the single rounds of the negotiations.

The authority has some available information about contributions of individual polluters to the pollution in the region, but it is unfamiliar with the true net costs of environmental protection of polluters and about their criteria of economic efficiency. The authority uses traditional economic tools in the process of the negotiation. The rate of payments for pollution is set. The rate of this payment can also be negotiable in more advanced models of the concept. The authority collects revenues from these payments in a fund and negotiates financial contributions from the fund with the polluters. During the process of negotiations, it maximizes various environmental effects per one unit of the financial contribution, which leads to the cost-effective solutions. Various methods and approaches for a quantification of indicators of these environmental effects can be used, such as risk assessment, multicriteria analyses, etc. Three different sets of environmental indicators could be used in the process: first, charging for the pollution; second, setting priorities for the financial support from the fund, and third, the final control of the concluded agreements. The authority can use various strategies from a spectrum defined by marginal strategies as follows: the first strategy leads to a uniform contribution from the fund per one unit of the positive environmental effect; the second strategy tries to reflect different net costs of environmental protection of the individual polluters.

LABORATORY ECONOMIC EXPERIMENT ILLUSTRATION

Let us illustrate the approach to a laboratory economic experiment. An ideal case of (politically) required water pollution reduction in a river basin was based on both environmental and economic data taken from various projects realized in practice in the Czech Republic. We considered pollution based on chemical oxygen demand (3,700 t/year), dissolved substances (790 t/year), insolved substances (1,320 t/year), phosphorous (18 t/year), and nitrogen (140 t/year). To put it more simply, the pollution was recalculated by using coefficients of harmfulness of the single pollutant. The resulted pollution was 1,454,000 units/year.

15 major polluters exist in the region. Their contribution to the pollution was transferred to the recalculated

units as follows: textile company (A; 30,760), stone manufacturing (B; 5,760), heating station (C; 7,750), pigments manufacturing (D; 4,500), heating station (E; 14), photodeveloping (F; 50), metal surface remanufacturing (G; 10,800), textile (H; 102,560), food production (I; 365,200), food production (J; 157,890), waterworks (K; 176,750), municipality I (L; 54,605), municipality II (M; 99,500), municipality III (N; 105,250), and municipality IV (O; 43,790). These polluters contribute, in total, 1,165,108 units, which is approximately 80% of total water pollution in the basin.

The environmental goal for the region was to reduce pollution from all sources at least by about one-third and to reduce pollution from the major sources by about one-half.

The laboratory experiment was conducted by Czech students of Environmental Economics and Policy Specialization at the University of Economics in Prague. The students played roles of managers of polluting entities. The teacher played the role of the authority. Charges that are paid according to Czech law (2) on pollution of surface water were used as a rate of payment for single pollutants to create revenues in the fund.

The negotiating in this experiment took three rounds. Six hypothetical agreements between the authority and the polluters were concluded at the end of the last round: I (pollution reduction 280,700 units), B (5,760), J (151,975), N (89,448), K (88,750), G (7,200). The agreed pollution reduction was 623,833 units in total, which represents about 54% of total pollution from major sources and about 43% of total pollution in the region. When repeating the experiment with teachers of economics, the agreements were concluded with a similar set of polluters C, G, I, J, B, and K. The pollution reduction was, in this case, 535,825 units [for the other experiment see (1)]. The theoretical (computed) pollution reduction satisfying the cost-effectiveness rule was 536,523 units, i.e., very close to the results of both experiments. Also, the subsidies agreed at the end of the experiments were very close to the theoretical (computed) ones: Computed subsidies were 55,799 monetary units. Students agreed with 52,650 and teachers concluded the agreements with the total subsidies at 54,780 monetary units.

Other experiments testing the described approach were conducted using ideal data. Also, in these cases, the experimental outcomes were very close to the computed ones.

In the next step, we used the same data for experimenting with the tradable pollution permit system. We were interested what the result would be if the same (water) pollution reduction target were achieved via trading the permits. In all laboratory experiments, the set of polluters who concluded the agreements with the authority and the set of polluters who sold the permits were about the same. Moreover, the price per one unit of permit was very close to the subsidy per one unit of reduced pollution.

CONCLUSIONS

It would be possible to discuss several practical applications of the approach. It could be used as a model for negotiations between a state environmental protection authority (regulator) and multiple polluters, as described above.

An environmental protection fund could be established and privately owned by the polluters that voluntarily want to solve a pollution reduction problem in a region and want to use the negotiation approach to avoid negative state authority action. The concept could be applied within a single corporation having multiple sources of pollution for solving the “principal-agents” problem where manager’s and owner’s utility functions are different.

A special application might be to organize a voluntary, experimental workshop of selected polluters and/or experts. Some information that might be useful for environmental policy decision-making, like a rough estimation of the environmental pollution charges, could be generated at the workshop (to receive information that is important).

The approach can also help with an efficient initial distribution of tradable water pollution rights. Moreover, it could contribute to resolving complicated situations (characterized by many pollutants, synergic effects, etc.) before the rights are allocated for limited number of the pollutants.

BIBLIOGRAPHY

1. Sauer, P., Dvorak, A., Lisa, A., and Fiala, P. (2003). A procedure for negotiating pollution reduction under information asymmetry. *Environ Resource Econom.* **24**: 103–119.
2. Czech National Council Act No. 58/1998 Coll. on the Charges for Pollution of Surface Water.

READING LIST

Brink, P. (Ed.). (2002). Voluntary environmental agreements. *Process, Practice and Future Use*. Greenleaf Publishing, Sheffield.

Dinar, A. (Ed.). (2000). *The Political Economy of Water Pricing Reforms*. Oxford University Press, New York.

May, C. (1998). Environmental agreements. *The Role and Effect of Environmental Agreements in Environmental Policies*.

Fiala, P. (1997). Models of cooperative decision-making, In: *Multiple Criteria Decision-Making*. G. Fandel and T. Gal (Eds.). Springer, Berlin, pp. 128–136.

Glasbergen, P. (Ed.). (1998). *Co-operative Environmental Governance: Public-Private Agreements as a Policy Strategy*. Kluwer Academic Publishers, Dordrecht, the Netherlands.

On Environmental Agreements. (1996). *Commission of the European Communities (COM 96–561 final)*. Brussels, Belgium.

OECD. (1999). *Voluntary Approaches for Environmental Policy: An Assessment*. OECD, Paris.

WATER RESOURCE ORGANIZATIONS

PETER E. BLACK
SUNY ESF
Syracuse, New York

Water management organizations evolved at local, regional, state, and federal levels. They were created, in large part, during the twentieth century, exhibit different degrees of formality and authority, and exist for different purposes, serving different groups consisting of individual qualified members and/or of the public at large. The organizations fill specific missions, roles, and meet demands of members and official segments of society. Knowing their missions and responsibilities makes it easier to seek information or how to influence their responsibilities.

INTRODUCTION

Organizations may be classified in a variety of ways. The water and related land resources agencies lend themselves to classification by mission (or responsibility), level of government, and legal status. Other water organizations may be national in scope but voluntary in membership, responsible for licensed professional activity, or open to all who subscribe to the group's philosophy. Any nongovernmental organization may be chartered under the laws of a state and may be incorporated for protection from liability actions. Some organizations may be tax-exempt, depending on their political nature and/or activity. Organizations may be temporary if their missions are limited; others may experience a shift in mission as their goals are achieved, or the membership finds them no longer relevant.

GOVERNMENTAL ORGANIZATIONS

Federal Agencies

Federal agencies derive their authority from the U.S. Constitution. State agencies, of course, derive theirs from their constitutions, which were approved by the U.S. Congress at the time of statehood, or were subsequently created by state legislatures. All the organizations' chartering legislation is subject to judicial review to ensure constitutionality. For example, the 1911 Weeks Forest Purchase Act (36 Stat 961) asserts in its opening statement of purpose of "the purchase of forest lands necessary to protection of the flow of navigable streams," the federal government may buy abandoned forest lands for inclusion in the National Forest system. Those fourteen words (in quotes) are sufficient to establish the constitutionality of the Act because the Constitution confers the right to regulate navigation on the federal government [Article I, Section 8 (3)]. The principal federal agencies charged with water and related land management are arranged by primary function(s) in Tables 1a to 1e (abstracted from Table 3.1 in Reference 1). Many agencies have other functions as well.

State Agencies

The states usually have a series of departments involved with forests, water, lands, fish and wildlife, and parks and recreation. The specific names and combinations of responsibilities are quite variable but tend to follow some typical patterns and in several different ways administer those permits and programs that federal agencies may have delegated to lower levels either because the states refused to be regulated by the federal government, they wished to have greater control over the regulations, or the federal government had neither the time nor funding to execute the mandated responsibility ("unfunded mandate").

Typically, the states create state-wide agencies such as departments of natural resources or environmental conservation, fish and wildlife, parks and recreation, agriculture and markets, and so forth. These deal with *intrastate* policies and management, regulation, granting of permits, and occasionally research and inventory. For point source pollution control permits under the control of the (federal) Environmental Protection Agency, the latter may have delegated responsibility for issuing National Pollution Discharge Elimination System (NPDES) to the state agency, in which case they are called SPDES permits. Continuing coordination between the state issuing authority and the EPA is part of these cooperative efforts. Some of the agencies may also be involved as partners with similar agencies in other states.

States may create *interstate* organizations to work with neighboring states by authority of the United States Constitution [Commerce Clause, Article I, Section 10 (3)]. These are interstate compacts that may exist for any of several reasons. Most interstate compacts administer interstate rivers and harbors. The managing organization is known as a *commission* (as opposed to a *committee*, which is usually informal. The Commissioner of the appropriate state department may serve—by appointment of the governor—as the representative of the state to an interstate commission.

Local Agencies

States may extend the constitutional right to local government to create agencies such as districts. These are extensions of state government that have the power to set boundaries, hold referenda, tax those holding real property within the district boundaries, enter into contracts, and incur indebtedness. There is a wide variety of districts; many exist for water management, including soil and water conservation districts, flood control districts, conservancy districts, drainage districts, and so on (in addition to the typical neighborhood highway districts, lighting, fire fighting, and school districts).

Typically, local organizations are formal or informal. The former include the *districts* (described above), and the latter are *associations*. Associations are not legal extensions of state government. Without formal boundaries, they may have members who agree to pay dues and may live anywhere. Usually, some common long- or short-term interest is the glue that holds the association together. Currently, there are thousands of watershed initiatives, associations that have a wide variety of collective names. They are watershed, river basin, or lake associations, councils, committees, land trusts, and so forth. They may, of course, be incorporated. Districts are often identified by other descriptive names, such as Resource Conservation Districts, Soil Conservation Districts, Conservation Districts, Natural Resource Districts, and Land Conservation Districts (as is pointed out on the NACD website http://www.nacdnet.org/). In some of the states that have strong county governments, the districts are organized on the basis of and coincident with the boundaries of the county; in others, the district is based on natural drainage.

Table 1a. Federal Land and Water Management Agencies

Agency	Date	Department	Mission, Responsibility
Bureau of Indian Affairs	1824	War, Interior	Manage forest, range and water resources; education, advocacy
Forest Service	1905	Interior, Agriculture	National forests, research, state and private forestry, and international forestry
National Park Service	1916	Interior	Preservation for science and recreation
Fish and Wildlife Service	1940	Interior	Wildlife refuge administration
Bureau of Land Management	1946	Interior	Renewable and mineral resources on public lands

Table 1b. Federal Construction and Operating Agencies

Agency	Date	Department	Mission, Responsibility
Corps of Engineers	1774	Continental Congress, War, Defense	Survey, rivers and harbors, navigation, flood control, water supply, water supply and recreation, dam safety, wetland permits, and aquatic ecosystem restoration
Bureau of Reclamation	1902	Interior	Irrigation and dams therefore
Natural Resources Conservation Service (nee Soil Conservation Service)	1933	Agriculture (Interior for a few years)	Soil and water conservation, upstream flood control, wetlands, education and information; administration of principal agricultural conservation programs; disaster response
Tennessee Valley Authority	1933	Independent	Hydropower development, flood control, economic improvement

Table 1c. Federal Regulation and Enforcement Agencies

Agency	Date	Department	Mission, Responsibility
Interstate Commerce Commission	1887	Independent	Carrier rules and regulation; sets interstate freight rates and routes
Federal Power Commission	1920	Independent	Licenses hydropower sites on navigable streams, public lands
Flood Insurance Administration	1968	Housing and Urban Development	National Flood Insurance Program
Environmental Protection Agency	1970	Independent	Pollution control permits, abatement, research/grant programs

Table 1d. Federal Research and Inventory Agencies

Agency	Date	Department	Mission, Responsibility
Geological Survey	1879	Interior	Mapping, hydrologic data collection, publication, and research; benchmark watershed program
National Weather Service (nee Weather Bureau)	1890	Signal Corps, Agriculture, Commerce, NOAA	Weather data collection, research, and publication; flood and fire weather forecasting
Agricultural Research Service	1953	Agriculture	Soil−water−vegetation research
Office of Water Research and Technology	1964	Interior, now abolished	Promotion, coordination, and funding for water resources research
National Oceanic and Atmospheric Administration	1970	Commerce	Explores, maps, and charts oceans, atmosphere, and space

Table 1e. Federal Coordination and Study Agencies

Agency	Date	Department	Mission, Responsibility
Senate Select Committee on Water Resources	1959−1961	Congress	Federal investigation of all water resource issues
Water Resources Council	1965−1981	Independent (now nonfunded)	Coordination of federal activities, flood control regulations, principles and standards for evaluation of water and related land resources projects and programs
National Water Commission	1968−1973	Independent	Nonfederal study of all water resource programs issues
Council on Environmental Quality	1970	President	Policy, research, advice to President and reports to Congress and the Public
National Study Commission	1972−1977	Independent	Charged to report on necessary adjustments to Clean Water Act halfway to 1983 and 1985 goals

PROFESSIONAL ORGANIZATIONS

There are numerous professional societies (sometimes confusingly referred to as associations) that serve the particular profession, such as architects, engineers, fishery biologists, foresters, geologists, geophysicists, hydrologists, landscape architects, limnologists, oceanographers, public administrators, range managers, soil scientists, surveyors, water works managers, well drillers, wetland managers, wildlife managers, and so forth. Many of these organizations serve, wholly or in part, important professions in water resource activities. Many have both technical and public-oriented publications that might be of interest to the lay public. Annual national and regional or state meetings are also often open to the public in whole or in part. Usually, putting the organization's initials (or full name) into the website URL line will provide the home webpage. Many water resources organizations may be found listed under links at the American Water Resources Association (http://www.awra.org), the principal interdisciplinary professional water resources organization in the United States, at the Environmental Protection Agency (http://www.epa.gov/), or at the website of the Conservation Technical Information Center (http://www.ctic.purdue.edu/CTIC/CTIC.html).

NONGOVERNMENTAL ORGANIZATIONS

Nongovernmental organizations are most often referred to collectively as "NGOs." The number continues to grow. Many of these may also be found at the websites of organizations cited above, by online search, or in the publications of the organizations.

During the past two decades (approximately), water resources management decisions have been effected by partnerships made up of mandated national, state, and local agencies in concert with business, interested individuals, and NGOs. Successful partnerships operate at all levels, with large or small numbers of participants, with and without government sanction, and over time with a shifting membership reflecting past accomplishments and initial goal achievement. Partnerships represent the finest in problem solving challenges and are truly democracy in action.

Partnerships require that all stakeholders[1] have an opportunity to take part in the process; that the participants respond to questions such as "What are our common goals?", "What can we do together better than separately?", "On what do we disagree?", and "Can we agree on an initial objective?"; that no one participant exercises fiscal, organizational, political, or social control over the process or the outcome; that participants treat one another with cordiality and respect; and that there be some identifiable constraint on the challenge, usually for water resources challenges, geographical, but also possibly economic, legal, or disaster or other crisis.

SUMMARY

Historical, legal, and cultural events during the last 200 years combine with the individual personalities of leaders and public citizens to leave a massive array of legislation, government agencies, and nongovernmental organizations all charged with or assuming responsibilities for some part of the water resources of the United States. Knowing the authority, extent of influence, nature, and relationships among the organizations is essential for sound water resources management. But, it does not guarantee it: only when people take on the responsibility of participating in the important water resources decision-making process will the vital water resources be effectively managed for a sustainable future. Partnerships are important players in the continuing policy and management decisions over water and related land resources.

BIBLIOGRAPHY

1. Black, P.E. and Fisher, B.L. (2000). *Conservation of Water and Related Land Resources*, 3rd Edn. CRC Press/Lewis, Boca Raton, FL.

READING LIST

Holmes, B.H. (1979). *History of Federal Water Resources Programs and Policies, 1961–70*. USDA, Miscellaneous Publication No. 1379, Superintendent of Documents, USGPO, Washington, DC.

Holmes, B.H. (1972). *History of Federal Water Resources Programs*. Miscellaneous Publication 12233, Department of Agriculture, Washington, DC.

Office of the Federal Register. (1998). *Government Organization Manual, 1998–1999*. National Archives and Records Administration, Washington, DC.

A BRIEF HISTORY OF THE WATER POLLUTION CONTROL ACT IN THE U.S.

David L. Russell
Global Environmental
Operations, Inc.
Lilburn, Georgia

The Rivers and Harbors Act of 1899 was the earliest piece of federal water pollution control legislation in the United States. It regulated the dumping of garbage into rivers and harbors from ships. The act was only moderately successful in its application, but it became important in the 1970s when the EPA used it to regulate discharges into navigable waters.

ORSANCO

The earliest significant start for water pollution control legislation in the United States began with the formation of the Ohio River Sanitary Commission (ORSANCO) in the 1940s. It was formed as a result of efforts by individuals

---

[1] This interesting term is currently used to identify those with a vested interest. However, that is the precise opposite of the original use of the term to identify the person in a poker game who held the bet funds—the stake—while the game was played out, in other words, it was the person who had *no* vested interest!)

from the U.S. Public Health Service (USPHS) and many of the state Departments of Health. The ORSANCO compact culminated more than 10 years of work to promote public health and reduce the incidence of waterborne disease in communities along the Ohio River. ORSANCO started as an initial effort among four states, and by 1940, the compact had six states. In 1948, the seventh and eighth states signed the compact on the same day that President Harry Truman signed the first comprehensive water pollution control legislation that directed the U.S. Public Health Service to control water pollution to promote public health.

The Public Health Service already had authority, but was reluctant to use it for water pollution control unless there was a specific prominent incident. The real power to control pollution still lay with the states.

By contrast, ORSANCO was effective because the control came from within the states as a cooperative agreement among the Departments of Health. ORSANCO was instrumental in issuing the first set of coordinated standards—the first water treatment design code manual—in 1954. ORSANCO was also effective in coordinating the development of antipollution lawsuits and still is an effective voice in the Ohio River region when it speaks for the state governments. ORSANCO has legal authority to compel individual polluters in any of the compact states to reduce their discharge of pollutants, or even to cease the discharge completely and can, if it chooses, override individual state governments.

The ORSANCO compact is still in force and has remained unchanged for 50 years. The fundamental tenet of the compact is that the discharge from one state shall not be injurious to the public health and well-being of the residents of another state.

It is important to point out that developments in water chemistry, analytical technology, and fundamental research, along with political developments, accompanied the growth and expansion of ORSANCO. Many of these developments were independent, but it was the driving interest in water pollution control and industrial wastes that created the climate for research and development to bloom. Many of the earliest names in the history of water pollution control were involved with ORSANCO as well. Harold Streeter, who is one of the formulators of the Streeter–Phelps equations (which underlie much of river modeling) was an early Commissioner of ORSANCO. State Commissioners on the ORSANCO Board of Governors (from their respective State Health Departments) include Clarence Klassen (Illinois), Blutcher Poole (Indiana), and Fred Waring (Ohio). One of many early workers in the State of Indiania was Dr. Don Bloodgood at Purdue University, who created the Purdue Industrial Waste Conference in 1946. That conference was made possible through the encouragement of some of the ORSANCO Commissioners and was instrumental as a showcase in promoting nationally prominent water pollution control research until its incorporation into the Water Environment Federation Industrial Wastes Program in 2000.

WATER POLLUTION CONTROL ACT OF 1948

In 1948, the Water Pollution Control Act (Ch. 758; P.L. 845) authorized the Public Health Service to prepare comprehensive programs for eliminating or reducing the pollution of interstate waters and tributaries and improving the sanitary condition of surface and underground waters. Under the plan, water conservation and improvement was to be emphasized. The original statute also authorized the Federal Works Administrator under the U.S. Public Health Service to assist states, municipalities, and interstate agencies in constructing treatment plants to prevent discharges of inadequately treated sewage and other wastes into interstate waters or tributaries. More information is available at http://laws.fws.gov/lawsdigest/fwatrpo.html.

Controlling pollution and industrial waste gradually became more important, and the original act was amended to authorize additional water quality programs, standards, and procedures to govern allowable discharges as well as funding for construction grants or general program funding. The authority of the original act was extended several times, through appropriations up to 1956 by Ch. 927; P.L. 579.

Progress to the Mid-1960s

The Federal Water Pollution Control Act Amendments of 1961 (P.L. 87-88) stipulated that, during the planning for any reservoir, federal agencies consider storage to regulate stream flow for water quality control, but the act was placed under section 33 U.S.C. 1252 (transportation) of the U.S. Code. This later provided impetus for the U.S. Army Corps of Engineers and the U.S. Coast Guard to have authority for “transportation related” oil spills and what later became National Pollution Control Discharge Permit issuance. Under the 1961 amendments, the Department of Health, Education, and Welfare was directed to undertake research and determine the effect of pollutants and treatment methods and assess water quality in the Great Lakes. The Secretary of HEW was authorized to take “appropriate measures to ward off pollution of interstate navigable waterways if requested by a state (33 U.S.C. 1254(f)). Few requests were issued under the legislation.

The Department of Heath, Education, and Welfare was also ordered to report to Congress by 1964 on the effects of pollution on U.S. estuaries and the economic impacts of pollution control. HEW was also told to develop recommendations for further actions that would lead to a comprehensive national policy and define the roles of federal, state, and local governments. The interests of Mexico and Canada were initially considered but only as they related to damage from pollution of boundary waters (the Great Lakes and St. Lawrence Seaway, and the Rio Grande River). Congress prohibited the discharge of oil into U.S. waterways under 33USC 466 (transportation legislation).

Congress decided to provide some form of funding to states and cities to construct sewage treatment plants. This Construction Grants Program was initially a relatively modest effort through most of the 1960s, and it was capped until late in the decade at a maximum grant of 33% of the cost of interceptor sewers and treatment works.

PL-660

Public Law 660 was the beginning of a comprehensive pollution control effort. Prior to PL-660, the Public Health Service (Commission Corps) officers were largely responsible for water pollution control. The law that was passed in the early 1960s made several important changes in the funding of water pollution control programs. A new organization, called the Federal Water Pollution Control Agency, was created under the administration of the Department of Interior. The FWPCA also had responsibility for pollution control in Alaska (P.L. 86-70) and an expanded role in construction grants (P.L. 86-624).

The Silent Spring, the Age of Aquarius, and the Vietnam War

The publication of Rachel Carson's *The Silent Spring* in 1967 and the dawning of the "Age of Aquarius,"—the coming of age of the children born after World War II—led to a major departure from the once quiet role of the federal government in pollution control. In the United States, the Vietnam War was a time of domestic turbulence and significant social and political protest. The statements of young radicals ("Don't trust anyone over the age of 30") and a growing dissatisfaction with the progress in solving the war led to social protest and dissatisfaction with other areas of government as well. "Make Love not War" and the countercultural exploration of new religions and ideas, especially those from India, quickly led to the incorporation of the concept of "Mother Earth" and a "Gaia" movement. Among the tenets of this movement were the concepts that the earth was a living sentient being, that all industry was bad and full of polluters, and that all pollution was bad and should be stopped at any cost. During this era, the first ecoterrorists became active in driving spikes in trees to be logged so that they would ruin the timber cutting and debarking machines. Other actions and rhetoric were also significant in defining and shaping the environmental movement.

The Silent Spring presented a graphic description and made a persuasive case for the ecological damage from the indiscriminate use of pesticides, principally DDT. When the book reached bestseller status and was condensed in Reader's Digest, it struck a resonant chord among disaffected youth. When combined with the societal revolution, it quickly led to the inauguration of the first "Earth Day" in 1970. Later, researchers accused Carlson of sloppy and inaccurate research and presentation of a poor case for action, but her literary talents and public opinion silenced the critics.

The Clean Water Restoration Act of 1966, P.L. 89-753, authorized the Departments of Interior and Agriculture and the Water Resources Council to study the effects of pollution, including sedimentation, on fish and wildlife, sport and commercial fishing, recreation, water supply and power, and other specified uses in the estuaries and estuarine zones of the United States.

By 1969, the United Nations entered the political fray about the environment, when the Secretary General (U Thant) predicted Armageddon from environmental causes. He was asking whether mankind could survive another 10 years without immediate and decisive actions to control pollution. Predictably, he condemned the United States as the principal culprit in creating this forthcoming environmental disaster (which it may have been) and demanded that the United States take unilateral action to control pollution. He conveniently ignored the pollution arising from the rest of the world and the benefits from their industrial activity.

In the United States, Congress created the National Environmental Policy Act in 1969 and submitted it to President Nixon for signature. The NEPA was hailed as a new direction for the federal government because, for the first time, certain specific federal actions were to become subject to environmental review and the issuance of an "environmental impact statement." According to President Nixon, this heralded, "a new decade of environmental awareness."

On December 2, President Nixon signed Reorganization Plan Number 3, which created a new cabinet-level federal agency, the Environmental Protection Agency. The plan disbanded the Federal Water Quality Administration and stripped the Department of Interior of environmental enforcement responsibilities and programs. It also demoted the role of the Public Health Service and the Corps of Engineers in managing and controlling air pollution and water pollution, respectively.

It should be noted that at that time, many of the industries located on major waterways, such as the Ohio River, were required to obtain operating discharge permits issued by the Corps of Engineers. Municipalities were, depending on size, operating with state permission and permits. The federal program was much better administered than state programs. Funding for construction of wastewater treatment facilities by municipalities was coordinated through the states, and federal monies for construction grants to municipalities were rapidly increasing to 66% of the cost of those facilities that the states had prioritized.

Other actions in 1970 under PL 91-224 included

1. the development of a National Contingency Plan for preventing discharge of oil to the seas;
2. the development of regulations for determining harmful quantities of oil discharged to U.S. waters;
3. state certification that such discharge will not violate applicable water quality standards;
4. development of performance standards for marine sanitation devices;
5. demonstration projects to control acid or other mine water pollution;
6. control of water pollution within the watersheds of the Great Lakes; and
7. required federal facilities to operate in compliance with applicable water quality standards.

PL 92-500

The passage of PL 92-500 is often considered a major landmark in defining national policy for environmental protection. The reason often cited is the stated goals of the law: to make the waters of the United States *fishable* and *swimmable* by 1985. The law also had several other important provisions in addition to dramatic increases in funding for sewage treatment works.

The financial incentive for municipalities was increased to between 85% and 95% of the applicable construction costs and was supported by a very large pool of money. The consequence of this funding was that the qualified and unqualified began designing and building wastewater treatment works at an enormous rate—so fast that the personnel required for planning, project review, and construction oversight were overwhelmed by applications. Moreover, the incentive for oversight was reduced because the municipalities had only 5% to 15% of the actual cost at risk and the administration of the federal Construction Grants Program was primarily financial, not technical. As a result, a large number of new facilities were built, many of them far in excess of the needs of the communities.

By 1982, the EPA had issued a manual citing the most common mistakes in designing facilities. This manual detailed the most common failures in designing facilities, and, by its very existence, was a ringing condemnation of the many failures of oversight of the Construction Grants Program. Many of the mistakes cited by the EPA were elementary and illustrated a degree of incompetence of many of the design professionals and failure of the regulatory community to exercise adequate planning and oversight. Planning and oversight takes qualified manpower and time and cannot be rushed.

The Construction Grants Program was a well-intentioned failure in a number of other aspects as well. First, it added a significant level of complexity and paperwork requirements to the construction program that increased engineering, administrative, and total project costs. The requirement for Davis–Bacon Act compliance, which mandates that the highest prevailing wage rate would be paid and that union labor would be used whenever and wherever possible, effectively eliminated any potential savings from the use of nonunion labor and discouraged many cost-efficient improvements.

The second item that caused the greatest percentage of failures was directed at the engineering specifications. The EPA required that the project engineer must provide three specifications for each major piece of equipment in a project and that those items would be competitively bid without regard to the quality, only the price. The result was often a direct and dramatic deterioration in the quality of equipment purchased. The equipment was often selected by the contractor who had the option to appeal to the EPA on the basis of potential savings for the project. Many projects became contractor's auctions where the final project bore only a slight resemblance to the engineer's original specifications. The result was that equipment compatibility and reliability issues often arose during construction but were often sidestepped or overlooked because of the insistence on obtaining the lowest costs.

One final impact of the Construction Grants Program under PL 92-500 was the restructuring of the U.S. wastewater treatment industry. The emphasis on equipment price caused severe competition among wastewater treatment equipment manufacturers and reduced profit margins. As a result, companies cut back on research, and that had a 10 to 20-year impact on the wastewater treatment industry. From about 1975 to 1995, new products were being developed and introduced by non-U.S. sources.

Other requirements of PL-92-500 required the Federal Power Commission not to license hydroelectric power projects unless stream flow and water quality objectives could be met; proposed a range of national objectives for water quality designed to restore and maintain the chemical, physical, and biological integrity of the national waterways (33 USC 1252); and required reissuance of permits for pollution control discharges.

This last item was the item that created the National Pollution Discharge Elimination System (NPDES) permit system. The legislation also established minimum standards and requirements and provided funding for states to assume control over their own water quality management programs, once they had demonstrated their compliance with minimum federal standards set by the EPA.

In addition, PL 92-500 significantly expanded provisions related to pollutant discharges by requiring point-source effluent limitations consistent with state water quality standards, enabled the State issuance of water quality standards, and provided initial methodology for identifying and evaluating nonpoint-source pollution. Authority was also provided for developing water quality inventories and toxic and pretreatment effluent standards.

Other provisions clarified and established liability for discharges of oil and hazardous substances, provided for federal oversight and backup responsibility for oil spill cleanup operations, and authorized the Corps of Engineers to issue permits for discharging dredged or fill material into navigable waters at specified disposal sites (Section 404-33 U.S.C. 1344). However, the EPA could block the use of a disposal site based on water quality issues related to the discharge from the site and potential deleterious effects on municipal water supplies, fisheries, wildlife, or other water quality uses.

1977 Amendments

The Clean Water Act was further amended in 1977 (P.L. 95-217) to provide additional requirements for the following: "Best Management Practices" Program as part of the statewide planning program (33 U.S.C. 1288); provision for the Fish and Wildlife Service to provide technical assistance for developing "best management practices;" authorization for the Department of the Interior to complete the National Wetlands Inventory by December 31, 1981; and additional regulations and de minimis standards for the Corps of Engineers to issue general permits for dredging and filling waters and wetlands. These exemptions included normal farming, silviculture, and ranching (33 U.S.C. 1344(f)), and general activities below certain sizes.

1987 Amendments

The Water Quality Act of 1987 (P.L. 100-4) amended the Clean Water Act to provide (1) continuation of the Chesapeake Bay Program and an office dealing with pollution of the bay; (2) establishment of a Great Lakes

National Program Office within EPA and a Great Lakes Research Office within NOAA; (3) a requirement that EPA, the Fish and Wildlife Service, and NOAA conduct research on the harmful effects of pollutants in the Great Lakes, emphasizing the bioaccumulation of these pollutants in aquatic species; (4) require states to develop strategies for cleanup of toxics in waters where the application of "best available technology" (BAT) discharge standards will not meet state Water Quality Standards; (5) an increase in the penalties for violations of Section 404 (Dredge and Fill) permits; and (6) additional state reporting on lake water quality and methods to reduce the adverse impact of excess acidity.

A $15 million funding package has been provided to demonstrate methods to reduce the acidity of lakes (often found in coal and surface mining areas), and a $400 million program has been authorized for states to implement a nonpoint pollution source (agricultural and other areas) control program with EPA oversight. Estuaries of national significance are also to be nominated under this program, and plans must be developed to protect and restore those estuaries and to protect their chemical and biological status.

THE NATIONAL POLLUTION DISCHARGE ELIMINATION SYSTEM

Rolf R. von Oppenfeld
The TESTLaw Practice Group
Phoenix, Arizona

INTRODUCTION

The Clean Water Act (CWA) authorizes the United States Environmental Protection Agency (EPA) to regulate water pollution discharges to federally regulated surface waters via a permitting process. All facilities that discharge pollutants from a point source into waters of the United States must obtain a National Pollution Discharge Elimination System (NPDES) permit from the EPA or an authorized state [42 U.S.C. §§ 1311(a), 1342; 40 C.F.R. § 122.1(b)]. Permits typically allow the permittee to discharge specified pollutants from designated outfalls. See the NPDES program homepage at http://www.epa.gov/npdes/. The NPDES program generally imposes technology-based standards and water-quality-based effluent limits on point source discharges of pollutants to surface waters, although less stringent requirements presently apply to stormwater runoff discharges.

Currently, 46 states and territories are authorized (have primacy) under Section §r402(b) of the CWA to administer the NPDES program. The EPA regions are the lead permitting authority in the other states and territories. State NPDES programs must use federal standards as a baseline; however, states may impose stricter requirements [40 C.F.R. § 123.1(I)]. State-issued permits are also subject to review by the EPA [33 U.S.C. § 1342(d)]. The EPA may make comments, objections, or recommendations to the state. If the state does not change the permit to address the EPA's concerns, the EPA may assume control of the permitting process for that facility. [33 U.S.C. § 1342(d)(4)]. The EPA may also retract program approval from states not administering the program in accordance with federal law [33 U.S.C. § 1342(c)(2)]. When the EPA region is the permitting authority in the state, the state must certify that an authorized discharge will comply with the state water quality standards [33 U.S.C. § 1341(a)].

TYPES OF PERMITS

A NPDES permit is a license to discharge a specified amount of a pollutant or pollutants into a specified receiving water under certain conditions. The NPDES permits establish effluent limitations [33 U.S.C. § 1362(11)]. However, NPDES permits also contain monitoring and reporting requirements, a duty to properly operate and maintain systems, upset and bypass provisions, recordkeeping, and inspection and entry requirements [40 C.F.R. § 122.41]. The source of much information concerning NPDES permits is contained in EPA's *Permit Writers' Manual* located online at http://cfpub1.epa.gov/npdes/writermanual.cfm.

There are two basic types of permits: the individual and general permit. NPDES permits are issued for fixed terms, not to exceed 5 years [33 U.S.C. § 1342(b)(1)(B); 40 C.F.R. § 122.46]. Individual permittees generally must apply for renewal 180 days before the permit expires. EPA has made electronic copies of NPDES permits for major facilities available online at http://www.epa.gov/NPDES/permitdocuments/.

General Permits

The permitting agency develops and issues a general permit. A general permit typically covers multiple facilities within a specified category, industry, or area. General permits are cost-effective and efficient because the agency can authorize discharges from a large number of facilities with numerous common criteria under one blanket permit. Facilities covered by a general permit "self-regulate," although the permitting agency will periodically inspect or investigate potential noncompliances. The procedures for using general permits and the general conditions that apply to all general permits can be found at 40 C.F.R. § 122.28. Most stormwater dischargers may apply for a general permit by filing a Notice of Intent (NOI) pursuant to a federal or state stormwater general permit. Many states also issue general NPDES permits for certain process wastewater discharges generally considered of minimal impact.

Individual Permits

A discharging facility must be covered by an individual permit if it does not qualify for coverage under a general permit. The authorized permitting agency prepares an individual permit for a specific facility. A permittee first provides the required information on an individual permit application.

Once the permit writer determines that the application is complete, the technical review starts. Effluent limitations are determined, and a preliminary permit containing these limits is prepared. (Id). The agency publishes public notice of the draft permit. Typically, the public notice allows for 30 days of public comment. Based on this information, the agency will issue or deny a permit (40 C.F.R. § 122.5).

DISCHARGE LIMITATIONS

Effluent limitations are established in a two-step approach consisting of technology-based and quality-based standards. First, dischargers are required to use treatments that are technologically and economically feasible for the particular industry. Additionally, dischargers may be required to use more stringent treatments to achieve certain water quality goals for the particular waterbody. Limits can be expressed as mass limits, concentration limits, visual observations, monitoring requirements, testing requirements, parameter limits, flow, pH range, or temperature limits. The limits can be imposed as maximum limitations and monthly average limitations.

Technology-Based Effluent Limitations (TBEL)

The NPDES permit system sets technology-based levels of allowable discharges, determined by a series of minimum technology-based national effluent standards for various industrial discharge categories. If no categorical technology-based standards apply, the standards are determined by the permit writer's best professional judgment (BPJ) of the appropriate technology-based reduction. See 40 C.F.R. Part 400 et seq.

TBELs are national effluent guidelines that establish limitations for all types of discharges within a specific industry, such as direct and indirect dischargers, existing and new sources, and specific discharges, including process water, cooling water, and sanitary wastewater. EPA has issued approximately 50 industrial category effluent guidelines (40 C.F.R. §§ 405–471). Guidelines are available online at http://www.epa.gov/docs/epacfr40/chapt-I.info/subch.N.htm.

The CWA created a two-step approach for dealing with the reduction of water pollution through TBELs. The first phase required industrial dischargers to meet pollution control through the best practicable control technology currently available (BPT) [33 U.S.C. § 1311(b)(1)(A)]. The second level of pollution control required using the best available technology economically achievable (BAT) [33 U.S.C. § 1311(b)(2)(A)]. BAT applies to nonconventional pollutants, whereas best conventional pollutant control technology (BCT), which is more lenient, applies to conventional pollutants, as explained below. However, BPT/BAT/BCT standards do not apply to new sources, only to existing sources. New Source dischargers must meet New Source Performance Standards (NSPS), based on the best available demonstrated control technology (BADCT) (33 U.S.C. § 1316). Also, industrial indirect dischargers that discharge to a publicly owned treatment works (POTWs) are not regulated by TBELs but are instead subject to the Pretreatment Program under 33 U.S.C. § 1317(b).

BPT Standard. The BPT standard collectively examines an industry category to determine the types of treatment facilities used in that industry. The EPA looks at the better run facilities and estimates the level of pollution control achieved through the typical technology. The EPA compares the cost of applying the technology to the entire industry with the effluent reduction benefits. See 52 Fed. Reg. 42522 (Nov. 5, 1987). This analysis yields a pollution control level demonstrating the average of the best achieved pollution reduction.

BAT Standard. BAT controls represent the maximum feasible pollutant reduction. EPA considers the technology currently used in that industry and also pilot data and technology used in other comparable industries. BAT does not require the use of a cost–benefit analysis, but the standard must be economically achievable. The BAT standard was designed to apply to toxic and nonconventional pollutants.

BCT Standard. BCT standards require industries to meet a cost-reasonableness test for the discharge of conventional pollutants, such as BOD, TSS, fecal coliform, pH, and oil and grease [51 Fed. Reg. 24973 (July 9, 1986)]. The cost-reasonableness test compares the reasonableness of the cost of the reduction in discharge levels of conventional pollutants to the cost of achieving that reduction. Conventional pollutants are subject to BCT, which are no more stringent than the BPT standard, rather than the more stringent BAT standard [33 U.S.C. § 1311(b)(2)(e)].

NSPS. New sources must comply with the stricter New Source Performance Standards, thus making the definition of a new source highly disputed. The EPA employs the substantial independence test to determine whether construction is leading to a new source or a modification of an existing facility [40 C.F.R. § 122.29(b)(1)(iii)]. This test looks at the degree to which the new unit functions independently of the existing facility, the degree of integration, and the extent to which the new unit is engaged in the same general type of activity as the existing source [49 Fed. Reg. 37998, 38043 (Sept. 26, 1984)].

NSPS are generally more stringent than those for existing sources; ideally, new sources should be built with state-of-the-art treatment technology. Requiring existing sources to retrofit under NSPS would be unreasonable. The major benefit to classification as a new source is that any new source constructed to meet NSPS may not be subjected to more stringent standards for 10 years after the date construction is complete or for the period of depreciation under the Internal Revenue Code, whichever is shorter (33 U.S.C. § 1316). However, the 10-year protection applies only to TBELs, not to WQBELs. At the end of the 10-year period, immediate compliance with the standards in effect at that time is required [33 U.S.C. § 1316(d)].

Best Professional Judgment. When TBELs have not been established for a particular industry sector or where there is a concern over a particular parameter (e.g., the permit writer thinks a receiving water is too polluted), permit writers will apply best professional judgment (BPJ) to establish permit limitations [40 C.F.R. § 125.3(e)]. The permit writer may consider data from the facility itself, from similar industries having similar waste treatment, or potentially applicable technology from an industry outside of the immediate industrial group. See 40 C.F.R. §§ 122.44(a), 125.3(c)(2).

Variances From TBELs. A qualified permittee may obtain a variance or modification from its permit's TBELs. One type of variance is a fundamentally different factors (FDF) variance [33 U.S.C. § 1311(n); 40 C.F.R. §§ 125.30–32]. The discharger must demonstrate that the factors applicable to its facility are fundamentally different from those considered for developing the effluent guidelines for its industrial category (40 C.F.R. Part 125). A variance is not issued merely because the cost of compliance would cause the facility to close. (Id). The applicant must also demonstrate that an objection was raised based on the FDF during effluent guidelines development or show why it was not. (Id).

The EPA may also modify BAT requirements or pretreatment requirements affecting nonconventional and nontoxic pollutants if the economic capability of the discharger requires less stringent limits [33 U.S.C. § 1311(g)]. The permitting agency may grant a variance from BAT for several nonconventional, nontoxic pollutants, such as ammonia, chlorine, color, iron and total phenols, where applicable BATs are unnecessarily stringent. (Id). Both the variance and the modification are granted on a permit specific basis.

Water-Quality-Based Effluent Limitations (WQBEL)

WQBELs are imposed where TBELs are insufficient to ensure water quality. WQBELs rely on the state water quality standards (WQS), and consist of two elements: (1) designated use classification and (2) criteria that are used to protect the designated uses. Surface water quality standards have become increasingly stringent during the last decade. Consequently, the WQBELs required in a NPDES permit for a particular pollutant are often more stringent than the corresponding technology-based effluent limitations.

Water Quality Standards. Water quality standards (WQS) are numeric standards developed for each stretch of a waterbody based on the designated uses. All states must classify their waters according to intended uses, for example, public drinking water, propagation of fish and wildlife, recreation, industrial, agricultural, and other uses [33 U.S.C. § 1313(c)(2)]. The WQS must also attain the CWA's goal of fishable or swimmable, if possible, and must maintain both the current and designated uses under the antidegradation policy, unless the state can demonstrate that it is unattainable or infeasible [40 C.F.R. §§ 131.10(g), 131.12].

Water quality criteria describe the physical, chemical, and biological characteristics of waters necessary to attain the designated uses. The EPA has developed numeric surface water quality criteria for over 150 pollutants [33 U.S.C. § 1314(a)(1); 63 Fed. Reg. 67547 (Dec. 7, 1998)]. Each navigable water is assigned such criteria to ensure that the water can support its designated use.

When the permitting agency determines that a discharge causes, has the reasonable potential to cause, or contributes to an in-stream excursion above the allowable ambient concentration of state numeric criteria within a state water quality standard for an individual pollutant, the NPDES permit must contain effluent limits for that pollutant [40 C.F.R. § 122.44(d)].

Toxicity-Based Limitations. Effluent limitations are sometimes based upon Whole Effluent Toxicity (WET), when it becomes difficult to set individual WQBELs [40 C.F.R. § 122.44(d)(1)(ii), (iii) and (iv)]. When the permitting agency determines that a discharge causes, has the reasonable potential to cause, or contributes to an in-stream excursion of pollutants in amounts or combinations that are *toxic* to humans, animals, plants, or other organisms, the permit must contain effluent limits for WET. WET limitations are incorporated into the NPDES permit and require the permittee to perform toxicity testing on the effluent. Toxicity tests expose various aquatic species to one or more concentrations of effluent in the laboratory to determine the short-term and long-term effects of exposure. Most permits issued since the release of EPA's WET Guidance Document require conducting toxicity tests regularly, such as monthly or quarterly [40 C.F.R. § 136, 65 Fed. Reg. 46457 (July 28, 2000)].

STORMWATER DISCHARGES

Regulation of stormwater runoff has presented challenges because stormwater runoff comes from diverse sources. EPA estimates that about 30% of known pollution to the nation's surface waters results from stormwater runoff. In 1987, Congress required EPA to establish regulations and issue permits for stormwater discharges. "Stormwater" includes stormwater runoff, snowmelt runoff, and surface runoff and drainage [40 C.F.R. § 122.26(b)(13)]. Stormwater does not include street wash water or infiltration, water that enters a sewer system below ground through defective pipes, pipe joints, connections or manholes. Discharges to retention basins that do not flow into waters of the United States or flow to POTWs also do not require stormwater permits.

The stormwater program has been implemented in two parts: Phase I and Phase II. The federal NPDES stormwater program relies on three types of permits: (1) The Construction General Permit, (2) The Multi -Sector General Permit, and (3) The Individual Permit. State agencies with primacy generally have analogous permits. More information on the federal stormwater program is available at http://cfpub.epa.gov/npdes/home.cfm?programid=6.

The Phase I Program

Phase I regulations cover separate municipal storm sewer systems (MS4s) serving populations over 100,000. Discharges into MS4s are treated as discharges to waters of the United States and require a NPDES permit (40 C.F.R. § 122.26). A MS4 is a conveyance or system of gutters, ditches, man-made channels or storm drains, owned by a state, county, municipality, or other public entity, designed or used for conveying stormwater, and is not a combined sewer or part of a POTW [40 C.F.R. § 122.26(b)(8)].

Phase I also regulates discharges from "areas associated with industrial activities" in 11 categories listed at 40 C.F.R. § 122.26(b)(14). The term "areas associated with industrial activities" covers discharges from any point source used for collecting and conveying stormwater, which is directly related to manufacturing, processing, or material storage at industrial facilities [55 Fed. Reg. 48007-15 (Nov. 16, 1990)]. The regulations break this definition down into two parts: what types of facilities have "industrial activity" based on their facility's Standard Industrial Classification (SIC) and what portions of the facilities include industrial activity. The portions of the facilities that include industrial activity include industrial plant yards, material handling sites, refuse sites, shipping and receiving areas, manufacturing buildings, raw material storage areas, etc. [40 C.F.R. § 122.26(b)]. Discharges from areas separate from the industrial activities, such as office buildings and parking lots and discharges from facilities engaged in wholesale, retail, service, or commercial activities are not included. [55 Fed. Reg. 48007 (Nov 16, 1990)]. Facilities covered by Phase I must apply for an NPDES permit, unless otherwise exempt.

Exemptions are provided until March 10, 2003 for industrial activities operated by municipalities with populations less than 100,000. See 64 Fed. Reg. 68722 (Dec. 8, 1999). Light industrial facilities (SIC codes 20–39 and part of 42) with "no exposure" of industrial activities or materials to stormwater are exempt under Phase I.

Permits for MS4s can be issued either for a particular system or on a jurisdiction-wide basis. MS4 permits must control pollution to the maximum extent practicable (MEP) [33 U.S.C. § 1342(p)(3)(B)(iii)]. The MS4 permit must also contain a requirement that prohibits nonstormwater discharges into the system's storm sewers.

Entities regulated under Phase I must obtain coverage under either an Individual Permit or a General Permit. Regulated entities must also implement stormwater pollution prevention plans (SWPPPs) or stormwater management programs; both require BMPs that effectively reduce or prevent the discharge of pollutants to receiving waters.

The Phase II Program

On December 8, 1999, the EPA published the NPDES stormwater management program Phase II Final Rule [64 Fed. Reg. 68722 (Dec. 8, 1999)]. Facility operators covered under Phase II must apply for permit coverage by March 10, 2003, except for oil and gas small construction activities which must apply by March 10, 2005 [67 Fed. Reg. 79828 (Dec. 30, 2002)].

The Phase II rule subjects two new groups of permittees to stormwater management: operators of small MS4s in urbanized areas and small construction activities that disturb 1–5 acres of land. The nonexposure exemption has been extended to include all industrial activities other than construction activities. The Phase II rule impacts about 5000 municipalities and federal systems such as military bases and prisons. Regulated operators must implement stormwater discharge best management practices (BMPs) that reduce or prevent the discharge of pollutants to receiving waters.

Stormwater Discharges Associated with Industrial Activities

Phase I contains a permitting component for industrial stormwater. The operations of industrial facilities that are within the 11 categories of "stormwater discharges associated with industrial activity" [located at 40 C.F.R. § 122.26(b)(14)(i)-(xi)] that either discharge to a MS4 or directly into waters of the United States must receive a NPDES industrial stormwater permit. Operators of industrial facilities can generally obtain either a general or individual permit to satisfy the NPDES stormwater permit requirement.

Conditional No Exposure Exclusion

"No exposure" means that all industrial materials or activities are protected by storms resistant sheltering to prevent exposure to rain, snow, snowmelt, or runoff. Certified facilities are not required to obtain a stormwater permit. Applicants must submit the "No Exposure Certification Form" every 5 years.

The no-exposure certification form is nontransferable. All new owners or operators must complete, sign, and submit a new form to claim the no-exposure exclusion. Upon submission of the new owners' form, the permitting agency may inspect the facility to verify the claim. A facility owner or operator who fails to maintain nonexposure status could be liable for an unauthorized discharge of stormwater. Under Phase II regulations, all industrial activities, except for construction, are eligible for exclusion.

The Multi-Sector General Permit

The Multi-Sector General Permit (MSGP) is designed to regulate certain stormwater discharges from nonconstruction industrial activities [63 Fed. Reg. 52430 (Sept. 30, 1998)]. The MSGP is one large permit subdivided into 29 separate sectors based on the SIC system and/or narrative descriptions. If a facility has operations that can be described by more than one sector, then it must comply with all pertinent requirements.

Facilities eligible for MSGP coverage must file a notice of intent (NOI). The NOI is essentially a promise that the applicant will comply with all MSGP conditions. The applicant must also prepare and implement a Storm Water Pollution Prevention Plan (SWPPP) prior to the submission of the NOI. The applicant can look at the MSGP for guidance on what to include in the SWPPP.

The SWPPP must create a Pollution Prevention Team comprised of individuals who develop and implement the SWPPP. The SWPPP must include a map of the facility, an assessment of the potential sources of stormwater pollution at the facility, and a list of BMPs. Generic BMPs include "good housekeeping," preventive maintenance, spill prevention and response, employee training, record keeping, nonstormwater discharge evaluation, erosion control measures, and stormwater management measures, as appropriate. Comprehensive site inspection/compliance evaluations are required.

Stormwater discharges subject to an existing NPDES permit cannot be covered under the MSGP until the existing permit expires. If the existing permit contains numeric standards that are more stringent than those required in the MSGP, the MSGP will not be available when the existing NPDES permit expires.

Stormwater Discharges Associated with Construction

Operators of both large and small construction activities must obtain a NPDES construction permit if the activity disturbs more than 1 acre or is part of a larger common plan of development or sale that will disturb more than 1 acre and the discharge is to a water of the United States or a MS4. The EPA requires that construction activities obtain a general permit. Effective on July 1, 2003, a new federal construction general permit is applicable to both large and small construction sites [68 Fed. Reg. 39087 (July 1, 2003)]. States with primacy have developed their own construction general permits.

There are two waivers available for small construction sites. The first waiver is the rainfall erosivity factor (R). The construction site operator must determine that the R in the reused universal soil loss equation is less than five. The second waiver is available operators in nonimpaired waters or who certify that stormwater controls are not needed, based on a total daily maximum allowance.

Separate Municipal Storm Sewer Systems (MS4)

Operators of large, medium, and regulated small MS4s must obtain NPDES permits to discharge pollutants and must develop stormwater management programs to prevent harmful pollutants from entering MS4s by stormwater runoff. Medium and large operators must submit permit applications under Phase I. Regulated small MS4s can choose among an individual permit, a general permit, or a modification of an existing Phase I MS4 individual permit. See Phase II Guidance, EPA 833-R-00-002 p. 4-1.

A "large MS4" is any MS4 located in an incorporated place or county with a population of 250,000 or greater, as of the 1990 Census. A "medium MS4" is any MS4 located in an incorporated place or county with a population from 100,000–249,999, as of the 1990 Census (40 C.F.R. Part 122 App. F, G, H). Many MS4s in areas of less than 100,000 population were individually brought into the Phase I program by NPDES. These regulated MS4s are not required to develop a Phase II program. See EPA Phase II Guidance, EPA 833-R-00-002, p. 4-2.

The Phase II program requires that certain small MS4s obtain a NPDES permit. A "small MS4" is any MS4 not already regulated under the Phase I definition of "small," including federally owned systems, such as military bases and veterans hospitals. See Phase II Guidance at p. 4-3. Small MS4s can be regulated if the small MS4 is wholly or partially located within an urbanized area, as defined by the 2000 U.S. Census, designated by NPDES, or designated by the permitting agency [40 C.F.R. § 122.32(a)(1), (2)]. Phase II requires that the permitting agency develop and apply designation criteria to all small MS4s located outside of an urbanized area that serves a population of at least 10,000 and a population density of at least 1,000 people per square mile.

Some MS4s receive a "physically interconnected" designation. Physically interconnected means that one MS4 is connected to a second in a way that facilitates direct discharges into the second system (Phase II Guidance, EPA 833-R-00-002 at p. 4-3-5). The permitting agency must designate most small MS4s located outside of an urbanized area that contributes substantially to pollutant loadings of a physically interconnected MS4 permitted by the permitting agency. The NPDES permitting authority may waive the coverage requirement in limited circumstances. See Phase II Guidance at p. 4-10, 11. Phase II allows operators of regulated small MS4s to choose from three permit options, the general permit, an individual permit, or a modification of an existing individual permit. The permitting agency reserves the right to determine which options are available to a regulated small MS4.

LEGAL PROTECTION FOR IN-STREAM FLOW

Melinda R. Kassen
Trout Unlimited, Colorado
Water Project
Boulder, Colorado

Water is precious. Human society needs water for drinking, other municipal uses, crops, and stock watering, as well as for industrial, aesthetic, and recreational purposes. But water left flowing through the landscape also has value, as it sustains every component of an ecosystem.

Water in rivers and lakes creates aesthetic value, as well as habitat for thousands of water-dependent species. Such species include aquatic species—fish, insects, amphibians, and plant life—as well as the land-based animals that eat them. Water flowing in a natural hydrograph helps maintain water quality. In addition, flowing rivers and streams can provide significant recreational amenities that feed local economies from fishing and boating or secondary activities such as hiking and bird-watching.

For most of recorded history, there has been enough water to supply humanity's needs without worrying about how removing water from streams for human use might affect other water-dependent species. However, because of significant human population growth, enhanced appreciation for the importance of protecting natural ecosystems, and increased concern about sustainability, societies have begun to provide legal protection to water left in a stream for the benefit of both water-dependent species and the people who use and enjoy the natural environment. These concerns have surfaced more in the western part of the United States than in the east, but concerns now arise even in less water-short areas.

Moreover, in view of a recent United Nations statement that one out of three people globally will face a water shortage by the year 2025, making water the commodity likely to cause controversy during the twenty-first century as oil did in the twentieth, understanding competing water uses and values is critical throughout the world.[1]

STATE PROTECTION OF IN-STREAM VALUES—WESTERN STATES[2]

In the United States, states, rather than the federal government, have primary authority over water allocation. Those states west of the 100th meridian are, for the most part, arid or semiarid. Since early in their settlement, they have found a need to reconcile disputes between water users; each of these states follows the doctrine of prior appropriation to allocate water among competing users.[3] Each state's system is unique, but their systems share many features. Most importantly, these systems define the value of water in terms of its use, and until late in the twentieth century, this meant use out of the stream. The prior appropriation system did not recognize any value for leaving water in a stream. Because the prior appropriation system functions as a seniority system, where the earliest user of a river retains always the first right of use, the system has created a challenge today for states and others seeking to protect remaining in-stream values, where water is "unused" except in sustaining the environment or by recreationists in a non-consumptive manner. Because water allocation is primarily a function of state law, the means to protect in-stream values must also be found primarily in the systems states use to allocate quantities of water to various uses.

All but three western states (New Mexico, North Dakota, and Oklahoma) provide some in-stream flow protection.[4] States use three different types of programs: legislatively established in-stream flows, agency actions to create or enforce in-stream flow rights, and private or market mechanisms to protect in-stream flows.

States vary in the methodology they use to quantify in-stream flow protection levels. Some, like the Tennant and R2Cross methods, look at a few simple variables, such as maximum flows, velocity, depth and wetted perimeter, and input them into a model that calculates minimum flows. States whose programs protect only minimum flows, such as Colorado and Wyoming, use these methods.[5] Others, like instream flow incremental methodology (IFIM) or physical habitat simulation system (PHABSIM), model the habitat requirements for aquatic species. Finally, there is the indicators of hydrologic alterations model developed by The Nature Conservancy; this model yields results that more closely mimic a natural hydrograph.

State Legislative Actions

In Montana, Kansas, and Idaho, the legislatures establish or approve agency recommendations for in-stream rights. Montana created a dozen such rights for blue-ribbon trout streams in a one-time effort in 1969.[6] Idaho has legislated nine stream flows or lake levels,[7] and Kansas has adopted statutory protection for 32 watercourses.[8] In addition, in 2000, Montana closed the Bitterroot entirely to new appropriations to protect the recreation and farm based economy of the area, both of which need water flowing downstream.[9]

State Agency Actions

Many Western states allow only state agencies to hold in-stream flows. Even so, neither Kansas nor Texas has recognized any in-stream flow rights; Nebraska has fewer than five. Pursuant to his statutory authority,[10] the Wyoming state engineer has approved 13 applications for new water rights from the state Game and Fish Commission to protect in-stream flows, but has only adjudicated two; adjudication is a prerequisite for enforcement. Sixty-nine applications await decision. Montana gives any state or federal agency the ability to apply for an in-stream reservation of up to 50% of a stream's average annual flow, and Montana's Department of Fish, Wildlife, and Parks now holds in-stream flow reservations on the main-stem Missouri and Yellowstone Rivers.[11]

Colorado and Oregon have the most extensive programs for in-stream protection.[12] In Oregon's program, the responsible state agency limits other diversions based on in-stream levels established for more than 1400 stream reaches.[13] On the streams where these levels exist, it is virtually impossible to obtain a new diversionary right; moreover, Oregon actively enforces its rights. In Colorado, where all water rights are a matter of court decree, the state, which has the exclusive authority to hold in-stream protection flows, has applied for and obtained more than 1350 water rights since 1973.[14] Unlike the situation in Oregon, however, Colorado's in-stream flow decrees do not automatically stop additional future development. Because Colorado also has no field personnel to monitor its rights and few gauges to measure their flows, Colorado has only once made one "call" on the river to demand that

---

[1] Statement of the UN-sponsored Third World Water Forum, Stockholm, Sweden, released August 13, 2001.

[2] Statistics in this section date from December 2001

[3] The prior appropriation system is the subject of another entry in this Encyclopedia.

[4] Gillilan and Brown, Instream Flow Protection, Island Press 1997, pp. 121-22. A 1998 New Mexico Attorney General opinion expands the definition of beneficial use there to include in-stream protection, but no court has affirmed the opinion and no entity has yet applied for an in-stream flow right.

[5] See, Wyo. Stat. Ann. §43.3.1001(c), (d); Colo. Rev. Stat. 37-92-102(3).

[6] Act of March 13, 1969, ch. 345, Mont. Laws 879-81.

[7] Idaho Code, §67-4301

[8] Kan. Stat. Ann. §82-703(c).

[9] Mont. Stat. §85-2-344.

[10] Wyo. Stat. Ann. §41-2-1001, et seq.

[11] Mont. Code Ann. §85-2-316.

[12] Colorado holds these rights pursuant to Colo. Rev. Stat. §37-92-102(3).

[13] Oregon holds these rights pursuant to Or. Rev. Stat. §§537.332 to. 360.

[14] These rights protect 8100 miles of more than 107,000 miles of state streams.

junior appropriators' diversions be curtailed to satisfy the Board's in-stream right. Instead, Colorado objects to applications for changes of water rights that might adversely affect the state's in-stream flow and negotiates court settlements that rely on others to monitor the state's rights. In both states, only the state agency owner of the right may enforce the rights with a call; members of the public must rely on the agencies.

Idaho has had a minimum stream flow statute on the books since 1937.[15] However, the program does not allow for new in-stream flow protective rights on streams that are otherwise fully appropriated—and most Idaho streams have long been overappropriated, so the statute provides little protection. Instead of obtaining water rights, the Idaho's Department of Water Resources has protected about 100 streams by adopting recommendations in its state water resources plan for minimum stream flow reservations for recreation or natural resources.[16]

The State of Washington follows a similar strategy. In 19 of its 62 water resources inventory areas, as well as the Columbia River, the state has adopted in-stream resource protection programs that establish minimum stream flows.[17] The state conditions later-issued water rights on meeting these flows. In addition, the state agency responsible for water rights has closed portions of the Yakima River Basin to further groundwater appropriations (at least pending further study) to keep some of this basin's water in-stream.[18]

Most states whose agencies have the power to protect in-stream flows acted only after 1970—in a region where traditional water rights date back to the mid-nineteenth century. Given that the prior appropriation doctrine values water rights based on their seniority and given that a senior user may dry up a stream entirely before junior users are entitled to any water, the junior nature of states' in-stream flows still leave agencies unable to protect—let alone restore—healthy stream flows. For example, Colorado has fewer than 15 rights whose priorities date earlier than 1973; in a state where most rivers were overappropriated at the beginning of the twentieth century, these rights can protect against additional future depletions but do little to restore already damaged fisheries to health. The state acquired its few, highly valuable senior rights by donation from cities seeking to protect streamfront amenities, and in one case, from a private conservation group. Wyoming also has the authority to acquire traditional existing rights by transfer or gift and change these rights to instream flow purposes,[19] but the state has never done so.

Some state agencies allowed to hold in-stream rights may not make new appropriations for such rights but may only change existing rights or acquire previously consumptive rights by donation. This is true in Utah, where the Division of Wildlife Resources holds five in-stream flow rights.[20] California and Arizona agencies may also accept such transfers.

Finally, some states can protect in-stream flows by denying other new diversionary applications. A statute directs the Utah state engineer to ensure that no new diversion will "unreasonably affect public recreation or the natural stream environment."[21] However, the state engineer has never relied exclusively, or even primarily, on this provision to deny a new or changed application for a water right. Oregon also has a "public interest" test, and though the state has never denied a new permit based on this test, Oregon has adopted rules pursuant to this provision that make it difficult to obtain a new water development permit where there are sensitive stocks of fish.[22]

Public Trust Doctrine

The public trust doctrine also plays a role in protecting in-stream water resources. Evolved from the laws of the Roman Empire, this doctrine holds that a state holds certain resources in trust for the benefit of society at large because of their public values. The ancient Romans applied the doctrine to the ownership of navigable rivers and harbors, which were held in trust by the state to protect the public's interest in commerce, navigation, and fishing. This concept was subsequently adopted by the English courts and applied to lands beneath tidal waterways.

In this country, states assumed this trust responsibility, including the ability to forbid private ownership of public trust resources. For example, the U.S. Supreme Court read the public trust to require automatic state ownership of submerged lands beneath navigable waters.[23] The Court then added that the states could not sell trust lands if the transfer would unduly burden public fishing, navigation, and commercial rights.[24] Because the public trust doctrine is primarily an aspect of state law, its application varies widely between states. Whereas the Colorado courts have never recognized the public trust doctrine as applied to water resources, California has used the doctrine aggressively to protect both lands and waters. The California Supreme Court found that its state water agency, the State Water Resources Control Board, had to consider public trust values in permitting new water withdrawals and also in reviewing existing withdrawals.[25] Thus, Los Angeles had to curtail existing diversions by one-third from tributaries of Mono Lake until the lake level recovered to a certain elevation.

Private and Market-Based Efforts

In some states, private entities can augment state programs that protect in-stream flows. The states of

---

[15] Idaho Code §42-1501 et seq.

[16] Idaho Code, sec42-1734A.

[17] Washington State Depts. of Ecology and Fish and Wildlife, "A Guide to Instream Flow Setting in Washington State." Pub. No. 03-11-007, p. 23 (March 2003).

[18] Ground Water/Surface Water Interactions to be Studied in Yakima River Basin, Department of Ecology Focus No. 99–1817, August, 1999.

[19] Wyo. Stat. Ann. §41-3-1007.

[20] Utah Code Ann. §73-3-3(11)(g)(1989 & Supp. 2001).

[21] Utah Code Ann. §73-3-8(1).

[22] OAR 690, Div. 33.

[23] *Pollard v. Hagan*, 44 U.S. 212 (1845).

[24] *Illinois Central Railroad Co. v. Illinois,* 146 U.S. 387 (1892).

[25] *National Audubon Society v. Superior Court of Alpine County*, 658 P.2d 709 (Cal. 1983).

Oregon and Washington have nonprofit water trusts that purchase mostly small quantities of water in the tributaries of targeted river basins dedicated to in-stream flow protection. In most cases, these purchases are for short periods of time. For example, although the Oregon Water Trust, has made more than one dozen permanent acquisitions (where it transfers the rights to the state), most are short-term leases, from 1–10 years. Still, the Trust has brokered more than 50 deals protecting 450 river miles. In addition, the Trust is helping irrigators take advantage of Oregon's conserved water program, which allows a water user who conserves water to use 75% of that water, provided that the user turns 25% back to the stream.[26]

Montana's program allows government agencies or nonprofit entities to lease water that is currently put to beneficial use through a diversion for in-stream flow protection for up to 30 years.[27] The state agency for fish, wildlife, and parks has entered into 10 leases under this provision. Trout Unlimited, a private conservation organization, has entered into fewer than 10, but at least one of TU's leases put 200 cfs into the Madison River, a quantity of substantial and significant magnitude. Montana also allows private water users to change a traditional diversionary right temporarily for in-stream purposes.[28] The change can last 10 years for a direct flow right and up to 30 years for water coming out of storage. Only one water diverter has sought (and been granted) a temporary change, as of 2001.

Alaska, Arizona, Nevada, and California also allow individuals to hold in-stream rights permanently. Anyone may transfer existing rights to in-stream reservations in Alaska,[29] although these reservations must undergo review every 10 years, a hurdle no other water right must jump. Two government agencies have obtained such rights, and several private entities have pending applications.

Under its statute,[30] Arizona has received 85 applications for new in-stream flow rights from individuals, environmental groups, and several federal agencies. Of these, the state has permitted 25 and perfected another 23 (after the applicants demonstrated actual use for 3 years). In addition, in 1997, the state used its Water Protection Fund to purchase unused Central Arizona Project water to sustain riparian habitat.

In California, an individual may not make a new appropriation for an in-stream flow right but may convert an existing diversionary right to in-stream protection; however, in 2001, 8 years after passage of this statutory provision, the Water Resources Control Board had yet to grant a single in-stream permit.

At least one federal agency has obtained a new in-stream flow water right in Nevada.[31] In addition, since 1988, Nevada has allowed the transfer of existing water rights to in-stream use. Both government agencies and private organizations have done so; in one case off-stream rights were transferred to protect streams on the Stillwater Wildlife Refuge.

South Dakota has no express statutory authority for in-stream flows, but the state Supreme Court has held that diversion of water was not necessary to achieve a beneficial use of water under state law.[32] As a result, the state has awarded several in-stream water rights, two new rights to a state agency, one to a federal agency and a change to a private company. In 2001, there were two additional applications pending.

Conclusion

Most states have some power to maintain, if not restore, healthy stream flows, but few have made such protection a priority. Western states water law systems simply do not make in-stream flow protection easy, and at the same time the region must accommodate significant growth. Given the limitations described above (little enforcement, seniority, or measurement), many state programs provide more of an illusion of protection than real water in-stream.

STATE PROTECTION OF IN-STREAM VALUES—EASTERN STATES

The wetter United States east of the Mississippi River follows a system of water allocation different from that of Western states. In the East, the riparian doctrine prevails, under which riparian landowners, those adjacent to a waterbody, have a right to use a portion of the flow of the stream or lake.[33] Although originally, a riparian water user could not diminish the natural flow of the stream,[34] today, holders of riparian rights can consumptively use water—even outside the watershed—if the use is "reasonable."[35] Today, urban, industrial, and agricultural demands have increased to the point that water resources are sometimes scarce in riparian jurisdictions.

To protect the public's interest in a reliable water supply, many riparian states require even landowners to obtain permits before using any water. Essentially, permits serve as a means for a state to quantify existing water uses and give the permit holder a right to use a quantity of water permanently or for a limited time. The permit requirement limits the number and size of water rights that come with riparian property ownership.

Most riparian states recognize water as a public resource and have taken steps to protect public recreational, environmental, and aesthetic values. Thus, these states have passed statutes to protect stream flows. In addition, many of these states have established minimum stream flows, lake levels, and groundwater levels or delegated the authority to set these levels to a state agency.

---

[26] Or. Rev. Stat. §537.455.

[27] Mont. Code Ann. §85-2-408(2).

[28] *Id.*

[29] Alaska Code Ann. §46.15.145.

[30] Ariz. Rev. Stat. §45–151.

[31] *State v. Morros*, 766 P.2d 263 (1988).

[32] *DeLay v. US Fish and Wildlife Service*, 524 NW2d 855 (S.D. 1994).

[33] *See* National Water Commission, A Summary-Digest of State Water Laws 32 (Richard Dewsnup & Dallin Jensen eds., 1973).

[34] *Tyler v. Wilkinson*, 24 F.Cas. 472 (C.C.R.I. 1827)

[35] *Stratton v. Mt. Hermon Boy's School*, 103 N.E. 87 (Ma. 1913).

For example, Florida has authorized its Water Management Districts to set seasonal minimum flow levels for both surface and groundwater.[36] "Minimum flow" is "...the limit at which further withdrawals would be significantly harmful to the water resources or [area] ecology."[37] Further diversions may not occur below the established minimum. When existing flows or levels do fall below or are projected to do so within 20 years, the state must implement a "recovery or prevention strategy."[38]

New Hampshire's Rivers Management and Protection Program takes a different approach.[39] There, the state develops "river corridor management plans" to maintain in-stream flows along designated rivers after someone nominates the river for protection. Once accepted, the nomination must undergo court review and legislative approval. The rules to implement this program establish a general standard of in-stream flow protection for all designated rivers.[40]

FEDERAL LAW PROTECTION OF IN-STREAM VALUES

State law is the primary vehicle for water allocation decisions, but the federal government also plays a role. This is true in part because the federal government is a major landowner, particularly in the West. In addition, the federal Bureau of Reclamation owns and operates a massive system of water projects that controls the flows on many Western rivers. Finally, the nation's web of environment and natural resources law intersects with the goal of river protection.

Federal lands reserved for specific purposes, such as national parks, monuments, forests, and refuges, almost always need water. At the beginning of the twentieth century, the U.S. Supreme Court held that in reserving lands, Congress had implicitly reserved as well the water to fulfill the purpose of the reservation and also that the water right's priority date would be the date of the initial Congressional reservation, not the later year in which the federal government sought to quantify its right.[41] In 1954, Congress passed the McCarran Amendment, a statute directing the federal agencies to obtain these reserved water rights through the individual state systems of water allocation.[42]

Only in the 1970s did federal agencies begin in earnest to enter state systems to obtain and quantify their rights. Unfortunately, given the Western states' system of prior appropriation, the time lag between the date the federal land was reserved and the year when the federal government sought to confirm its water rights has proved devastating to this federal effort. Much of the West was unsettled in the early twentieth century; since that time, large cities have grown, farming has continued to divert 80–90% of the water used, and new industrial needs for water have been created—from heap leach mining to oil shale extraction to ski area snowmaking. These intervening water users are unwilling to remain passive when the federal government seeks to jump the seniority line and gain water rights for its lands.

As a result, almost all adjudications of reserved rights have been controversial. In most cases involving national forest lands, the Forest Service has failed to obtain meaningful protection for the rivers crossing its lands, even though its mandate includes the command to provide favorable flows.[43] The National Park Service and U.S. Fish and Wildlife Service on behalf, respectively, of national parks and the wildlife refuge system have fared marginally better; yet even these agencies have failed frequently to obtain protective flows. In Colorado, the Park Service failed to obtain a right to protect the rivers within Dinosaur National Monument.[44] More recently, the Idaho Supreme Court denied the U.S. Fish and Wildlife Service a reserved right for the Deer Flat National Wildlife Refuge made up a series of islands in the Snake River.[45]

Some federal agencies also use administrative mechanisms to protect stream flows when making permitting or licensing decisions. Thus, wholly outside state water allocation systems, federal agencies impose bypass flows to fulfill requirements that derive from federal laws, such as the Federal Power Act, the Clean Water Act, and the Federal Land Management and Policy Act, as well as the agencies' authorizing statutes, for example, the Forest Service Organic Act. In part because the use of this authority has sometimes limited validly issued state water rights, agency exercise of this authority has proved highly controversial and is thus used sparingly, if at all.

The Clean Water Act provides two mechanisms for protecting in-stream values to any state willing to use them. First is the state water quality certification program.[46] Under its provisions, a state must certify that a federally permitted or licensed project will meet state water quality standards. Pursuant to this authority, the State of Washington imposed substantial flow restrictions on a hydropower facility on the Olympic peninsula; in 1994, the U.S. Supreme Court upheld these restrictions.[47] Second is the total maximum daily load (TMDL) program" which arguably provides states a mechanism to protect flows through water quality regulation.[48] Here, states must list streams whose water quality is impaired.[49] Then, states must establish TMDLs that will bring the streams back into compliance. Recently, the EPA determined that the federal government cannot impose flow-based conditions in TMDLs, yet the states, which have primary

---

[36] Fla. Stat. §373.042 (2001).

[37] *Id.* at §373.042 (1)(a).

[38] *Id.* at §373.042 2(a), 2(b).

[39] N.H. Rev. Stat §§ 483:1–483:15.

[40] Rules for the Protection of Instream Flow on Designated Rivers, Chap.Env-ws 1900(2003). This document is available online at: http://www.des.state.nh.us/rules/env-ws1900.pdf.

[41] *Winters v. US, 207 U.S. 564 (1908).*

[42] 43 U.S.C.A. §666.

[43] See, e.g., *US v. New Mexico*, 438 U.S. 696 (1978); *Potlatch Corp. v. US*, 12 P.3d 1260 (2000), *US v. City & County of Denver*, 656 P.2d 1 (Colo. 1982) (*US v. Denver).*

[44] *US v. Denver*.

[45] *US v. State*, 23 P.3d 117 (Idaho 2001).

[46] 33 USC §1341.

[47] *PUD No. 1 of Jefferson County v. Washington Department of Ecology*, ## US 700 (1994).

[48] The total maximum daily load (TMDL) program is the subject of a separate entry in this Encyclopedia.

[49] 33 USC §1333(d).

responsibility for this program, can. Montana, for example, has adopted a flow-based TMDL for an impaired stream, a tributary of the Yellowstone River.[50]

The Endangered Species Act (ESA) provides federal agencies another handle for restoring stream flows when such flows are necessary to avoid jeopardizing an endangered or threatened species' existence. Again, the use of this authority is highly controversial. In one recent case, a federal court upheld the U.S. Forest Service's imposition of a bypass flow to protect an endangered species as consistent with both the ESA and the Forest Service's own Organic Act.[51] Some water users who have received less water from federal projects as a result of actions taken under the ESA have begun to challenge these federal actions as "takings" of their property rights without compensation, in violation of the 5th amendment to the U.S. Constitution.[52] In the summer of 2001, such a curtailment occurred to benefit both endangered salmon and lake fish sacred to the Klamath tribe that has resulted in a firestorm of litigation and proposed legislation.

On the other hand, ESA mandates have forced some creative responses. The Middle Rio Grande Conservancy District, embroiled in litigation over recovery of the silvery minnow, created a water bank to allow users to lease surplus water, at least some of which may go to keep critical Rio Grande reaches wet. New Mexico and the federal government have also reached an agreement to operate federal reservoirs in the Rio Grande basin to release stored water for in-stream flows to maintain minnow habitat. An even larger effort is beginning in California. There, parties trying to address endangered species concerns in the San Francisco Bay Delta created the Environmental Mitigation Bank to trade water to benefit the ecosystem.[53]

Finally, due to the existence of major federal water storage projects throughout the West, interests have pressured the federal agencies—owners of these projects—to change their operation. Primarily to benefit endangered species, but also to enhance recreation and other values, reoperation of federally owned or licensed storage and hydropower facilities can enhance river flows in a way that restores, or at least maintains, aquatic species. This has occurred to benefit endangered species in the Rio Grande, the Klamath, and the Columbia Rivers, along with many lesser known systems where endangered species live. Through the "CalFed" project, California reoperations have benefited people as well as endangered species.

CONCLUSION

State water law systems grew out of a nineteenth century worldview that believed that all resources are endless. Therefore, it is not a surprise that governments struggle today to protect in-stream values. But, because of the pressures exerted on the system as a result of modern environmental values, decision-makers are seeking ways to ensure that at least some of the country's waterways maintain healthy flows.

WATER QUALITY

U.S. Geological Survey

APPLICABLE STATUTES

Clean Water Act (PL 92–500; 33 U.S.C. 1251 et seq.)

Safe Drinking Water Act (PL 93–523; 42 U.S.C. 300f-j-26)

CLEAN WATER ACT, AS AMENDED (PL 92–500; 33 U.S.C. 1251 ET SEQ)

Purpose

Also known as the Federal Water Pollution Control Act, it is the objective of the Clean Water Act (CWA) "to restore and maintain the chemical, physical, and biological integrity of the Nation's waters". Seven goals and policies are declared in order to meet this objective. One goal was to reach zero discharge of pollutants by 1985. Others include providing funding for the construction of publicly owned treatment works (POTWs), creating a nonpoint source pollution program, and generally, and making the waters of the U.S. "fishable and swimmable". Although not specifically stated in this declaration, the CWA contains the nation's most noteworthy wetlands legislation.

MAJOR PROVISIONS BY SECTION

201–209 (33 U.S.C. 1281–1289)—Grants for Construction of Treatment Works.

These sections originally provided federal grants for the construction of wastewater treatment plants. The program has been phased out by the 1987 amendments in favor of a revolving loan fund.

301 (33 U.S.C. 1311)—Effluent Limitations

The discharge of any pollutant into the nation's waters except for discharges in compliance with the CWA is prohibited, according to this section. Limitations are placed on existing sources which vary according to the nature of the pollutant discharged and to where the outfall is directed.

---

[50] Montana Department of Environmental Quality, Big Creek Flow Restoration Plan (TMDL), December 2000.

[51] *County of Okanogan v. Nat'l Marine Fisheries Service*, 57ERC 1321, 2003 WL 22455482 (9th Cir. 2003) (not selected for publication in Federal Register).

[52] *Tulare Lake Basin Water Storage District v. United States*, 49 Fed. Cl. 313(2001).

[53] For information about this venture, *see*, CalFed Bay-Delta Program, Programmatic Record of Decision (2000), pp 54–59. [Note: this document is available online at http:// calwater.ca.gov/Programs/Environmental Water account/adobe_pdf/ ROD8-28-00.pdf.]

302 (33 U.S.C. 1312)—Water Quality Related Effluent Limitations

Point sources which interfere with the attainment or maintenance of desired water quality are subject to the imposition of more stringent effluent limitations.

303 (33 U.S.C. 1313)—Water Quality Standards and Implementation Plans

Water quality-based regulatory controls on dischargers known as Water Quality Standards (WQS) are required by states to protect designated uses of water bodies. Technological capability is not a consideration in setting WQS.

304 (33 U.S.C. 1314)—Information and Guidelines

This section requires the EPA to develop water quality criteria and guidelines for effluent limitations, pretreatment programs, and administration of the National Pollutant Discharge Elimination System (NPDES) program.

306 (33 U.S.C. 1316)—National Standards of Performance

A list of categories of effluent sources is presented which specifies that each of the industries listed must conform to technology-based new source performance standards. The standards are to demonstrate the best demonstrated control technology.

307 (33 U.S.C. 1317)—Toxic and Pretreatment Effluent Standards

This requires that industries discharging toxic pollutants meet effluent limits that employ the best available technology economically achievable. Part (b) mandates the establishment of pretreatment standards while part {c} looks at new sources of pollutants into publicly owned treatment works (POTWs).

309 (33 U.S.C. 1319)—Enforcement

Enforcement by the states, as well as compliance orders, and administrative, civil, and criminal penalties are authorized.

311 (33 U.S.C. 1321)—Oil and Hazardous Substance Liability

The Congressional declaration of policy against discharges of oil or hazardous substances in harmful quantities into waters and adjoining shorelines is presented. A spill prevention, control, and countermeasure plan is to be developed by all facilities that handle, transport, and store oil. Any spill or discharge of a harmful quantity of oil must be reported to the National Response Center. Strict liability is assigned to owners or operators of facilities from which oil discharges occur.

319 (33 U.S.C. 1329)—Nonpoint Source Management Programs

This section requires that states identify waters that are not able to meet WQS because of nonpoint sources. The activities responsible for the pollution are to be identified and a management plan is to be created to help correct the nonpoint source problem.

401 (33 U.S.C. 1341)—Certification

Any applicant for a Federal license or permit to conduct any activity which may result in any discharge into the navigable waters, shall provide the licensing or permitting agency a certification from the State in which the discharge will originate.

402 (33 U.S.C. 1342)—National Pollutant Discharge Elimination System

One of the most critical parts of the CWA is the establishment of the National Pollutant Discharge Elimination System which translates standards into enforceable limitations. This program may be administered by the EPA or states under EPA-delegated authority. After an opportunity for public hearing, a permit may be issued for the point source discharge of any pollutant, or combination of pollutants.

404 (33 U.S.C. 1344)—Permits for Dredged or Fill Material

This is the major wetlands provision of the CWA, and largely, in environmental law. The basic gist of this section is that a permit is required from the U.S. Army Corps of Engineers for the disposal of dredged or filled materials into navigable waters, notably wetlands, with EPA concurrence, and notice and opportunity for public hearings.

505 (33 U.S.C. 1365)—Citizen Suits

Any citizen is given the right to file suit against any person in violation of an effluent standard or against EPA for failure to perform nondiscretionary duties.

PERTINENT REGULATIONS

- **33 CFR Part 320—General Regulatory Policies, Army Corps of Engineers**
  The Corps issues permits under several different acts, the CWA being only one of them. The Corps must issue permits for discharges to insure that they comply with the applicable limitations and WQS. Section 320.4 describes the general policies that the Corps will follow in reviewing all Department of the Army permits. Such permits are to consider public interest, effect on wetlands, fish and wildlife, water quality, property ownership, energy conservation, navigation, environmental benefits, and economics, among other things. The pertains particularly to dredge and fill permits covered under 404 of the CWA.
- **33 CFR Part 323, Permits for Discharges of Dredged or Fill Material into Waters of the United States, Army Corps of Engineers**
  Contains definitions and special policies, practices and procedures to be followed by the Corps in connection with review of permit applications for the discharge of dredged or fill materials under 404 of CWA. Permits issued under 404 are required

for the discharge of dredged or fill materials not exempted by 323.4 of this part or permitted by 33 CFR Part 330. Please refer to 323.4 for a detailed listing of exemptions.

- **33 CFR Part 325—Processing of Department of the Army Permits, Army Corps of Engineers**
  Contains general processing procedures for all Department of the Army permits. Attention is given to dredge and fill activities in 325.1(d)(3) & (4). Federal agencies that initiate or authorize proposed actions that include dredge or fill discharge operations must ensure that the appropriate permits are obtained. In states with approved programs, permit application is done through the appropriate State agency.
- **40 CFR Part 110—Discharge of Oil, EPA**
  This section reiterates the mandate of 311(b)(3) of the CWA. It generally prohibits the discharge of oil into navigable waters in such quantities as may be harmful.
- **40 CFR Part 112—Oil Pollution Prevention, EPA**
  This requires that owners or operators of non-transportation related onshore and offshore facilities engaged in any type of oil and gas operation prepare a Spill Prevention Control and Countermeasure Plan, and provide guidelines for preparation of the plan. Such a plan may be required of an applicant for USGS approval of any oil and gas activities.
- **40 CFR Part 122—EPA Administered Permit Programs: The National Discharge Elimination System, EPA**
  Contains definitions and basic permitting requirements for EPA-administered NPDES programs under 318, 402, and 405 of the CWA. Permit applications and special NPDES program requirements are discussed in Subpart B. This is followed by permit conditions in subpart C. Subpart D covers the transfer, modification, revocation and reissuance, and termination of permits. Federal agencies that initiate or authorize proposed actions that include point source operations must ensure that the appropriate permits are obtained. In states with approved programs, permit application is done through the appropriate State agency.
- **40 CFR Part 123—State Program Requirements, EPA**
  Describes the general requirements and additional requirements for states and the EPA to obtain and give approval, revision, and withdrawal of state NPDES programs. State program information shall be made available to EPA upon request. EPA also has the right to review proposed general permits for 90 days.
- **40 CFR Part 125—Criteria and Standards for the National Pollutant Discharge Elimination System, EPA**
  This section prescribes criteria and standards for various requirements imposed as conditions for NPDES permit approval. Some of the criteria expanded upon are those presented for the imposition of technology-based treatment requirements in permits as given under 301(b) and 402(a)(1), those for modifying secondary treatment requirements under 301(h), those for Best Management Practices authorized under 304(e), and those applying to ocean dumping, in Subparts A, G, K, & M, respectively.
- **40 CFR Part 129—Toxic Pollutant Effluent Standards, EPA**
  This section designates toxic pollutant effluent standards and applies to owners or operators of specified facilities discharging into navigable waters. Section 129.4 listed the pollutants to be regulated. Each owner/operator is given 60 days to notify the Regional Administrator of any listed pollutant that is discharged. Much of the regulation gives specific information on each of the toxic pollutants.
- **40 CFR Part 130—Water Quality Planning and Management, EPA**
  Section 303 of the CWA gives the authority for promulgation of water quality standards (WQS) by states. Here, policies are established for water quality planning, management, and implementations of 303. The Water Quality Management process from the CWA provides the authority for a "consistent national approach for maintaining, improving and protecting water quality while allowing States to implement the most effective individual programs". After WQS are set by each state, implementation of the standards may be achieved by issuing permits, building publicly-owned treatment works, or instituting Best Management Practices (BMP) through a water quality management plan. Total maximum daily loads (TMDL) and individual water quality-based effluent limitations are discussed in 130.7. States are required to submit water quality reports to the Regional Administrator in accordance with 305(b) of the CWA (130.8). Final points of the regulation cover state submittals to EPA and program management.
- **40 CFR Part 131—Water Quality Standards, EPA**
  While Part 130 of this title works with water quality management, this section deals with water quality standards (WQS). It is stated that a WQS defines the water quality goals of a water body by designating the uses to be made of the water and by setting criteria necessary to protect the uses. WQS should be set to provide for the protection of fish, shellfish, wildlife, water recreation, and the use and value of public water supplies, agricultural, industrial, and other purposes. The procedures are presented for developing, reviewing, revising and approving WQS by the states and EPA. In 131.12, states are ordered to develop statewide antidegradation policies. WQS shall be reviews at least every three years by the states.
- **40 CFR Part 230—Section 404 (b)(1) Guidelines for Specification of Disposal Sites for Dredged or Fill Material, EPA**

These guidelines were written to restore and maintain the chemical, physical, and biological integrity of the U.S. through the control of discharges of dredged or fill material, with the guiding principle that degradation of sites may represent an irreversible loss of valuable aquatic resources. In Subpart B, it is stated that "no discharge of dredged or fill material shall be permitted if there is a practicable alternative to the proposed discharge which would have less adverse impact on the aquatic ecosystem". The permitting authority must determine the potential effects of dredging or filling activities on the components of the aquatic environment.

- **40 CFR Part 231—Section 404(c)Procedures, EPA**
  Contains procedures for EPA in exercising its authority to veto the specification by the Army Corps of Engineers or by a State of a disposal site for a 404 permit.
- **40 CFR Parts 401-471—Effluent Guidelines and Standards, EPA**
  Prescribes effluent limitations and pretreatment and performance standards, categorized by industries, that must be complied with as conditions for NPDES permit approval. Part 401 gives general provisions and Part 403 covers general pretreatment regulations for existing and new pollution sources. The remaining parts are industry-specific. Of particular interest to the USGS are:

  40 CFR Part 434—Coal Mining Point Source Category
  40 CFR Part 435—Offshore Segment of the Oil and Gas Extraction Point Source Category
  40 CFR Part 436—Mineral Mining and Processing Point Source Category
  40 CFR Part 440—Ore Mining and Dressing Point Source Category

- **518 DM 1—Comprehensive Waste Management, Department of the Interior**
  This chapter defines waste to include solid and hazardous waste, hazardous materials, and hazardous substances. Departmental policies, responsibilities, and functions regarding waste management are presented, with an aim towards prevention of hazardous waste generation. The use of sound waste management practices is mandated.

SAFE DRINKING WATER ACT (PL 93–523; 42 U.S.C. 300F—J-10)

Purpose

In 1974, the Safe Drinking Water Act (SDWA) was enacted with the general intent to protect the quality of drinking water the public receives from public water systems. To accomplish this, the SDWA focuses on two approaches. The first is to assure the quality of drinking water coming from the tap. The other approach is to prevent the contamination of groundwater that may be a source for drinking water.

MAJOR PROVISIONS BY SECTION

§1412 (42 U.S.C. 300g-1)—National Drinking Water Regulations

The EPA is required to promulgate national drinking water regulations, known as Maximum Contaminant Level Goals (MCLGs) and Maximum Contaminant Levels (MCLs) for public water systems. These standards specify minimum levels of drinking water quality and are to be issued for any contaminant that is known or anticipated to have any adverse effect on health.

§1413 (42 U.S.C. 300g-2)—State Primary Enforcement Responsibility; Regulations; Notice of Hearing; Publication in Federal Register; Applications

States are given the authorization to assume primary enforcement of the Act.

§1415–1416 (42 U.S.C. 300g-4 & 5)—Variances—Exemptions

The conditions are offered in which states may issue variances and exemptions.

§1417 (42 U.S.C. 300g-6)—Prohibition on Use of Lead Pipes, Solder, or Flux

Any use of lead in pipes, solder, or flux in public water systems where the water is intended for human consumption is prohibited.

§1422 (42 U.S.C. 300h-1)—State Primary Enforcement Responsibility

This section requires the establishment of state underground injection control programs so as to protect current and potential underground sources of drinking water from contamination.

§1424 (42 U.S.C. 300h-3)—Interim Regulation of Underground Injections

This declares that Federal agencies shall make no commitment of Federal assistance to actions that will contaminate an aquifer designated as the sole or principal drinking-water source for an area which, if contaminated, would create a significant hazard to the public health.

§1428 (42 U.S.C. 300h-7)—State Programs to Establish Wellhead Protection Area

States are given the authority to develop wellhead protection programs that address the kinds of activities that might be conducted in proximity to wells, particularly seeking to limit activities that might pose threats to the well fields.

§1431 (42 U.S.C. 300i)—Emergency Powers

If there is an imminent and substantial endangerment to public health through drinking water, EPA is given emergency powers to act against contamination.

§1445 (42 U.S.C. 300j-4)—Records and Inspection

The EPA is to promulgate regulations requiring drinking water monitoring. This section also mandates that EPA establish record keeping requirements.

§1448 (42 U.S.C. 300j-7)—Judicial Review

The D.C. Circuit is given the jurisdiction of judicial review of national primary drinking water regulations. The U.C. Courts of Appeal are given jurisdiction for any other EPA action when the petitioner resides or transacts business in that particular region.

§1449 (42 U.S.C. 300j-8)—Citizen's Civil Action

This section gives citizens the right to file suit to enforce any mandatory provisions of the SDWA.

Part F—Additional Requirements to Regulate the Safety of Drinking Water

§1461–1465 (42 U.S.C. 300j-21 to 26)

These particular sections codify the Lead Contamination Act of 1988. This requires that any lead-lined tanks of drinking water coolers be recalled.

PERTINENT REGULATIONS

- **40 CFR Part 141—National Primary Drinking Water Regulations, EPA**
  One of the most significant provisions of the SDWA is the establishment of National Primary Drinking Water Regulations. The regulations are given life here, with the establishment of maximum contaminant levels (MCLs) for inorganic and organic constituents, Subpart B and maximum contaminant level goals (MCLGs), Subpart F. Much attention is given to the monitoring and analytical requirements for the regulated water quality parameters, Subpart C. Filtration and disinfection are given much the same kind of thorough treatment. Also included are sections covering control of lead and copper, muse of non-centralized treatment devices and treatment techniques.
- **40 CFR Part 142—National Primary Drinking Water Regulations Implementation, EPA**
  With only a few exceptions, this part applies to the public water system in each state. States are given primary enforcement responsibility, given that the state has an EPA-approved program. If a state wishes to revise its program, it may do so if it follows the requirements given in 142.12. All states with enforcement responsibility are required to submit to the EPA Administrator a report containing various components dealing with national primary drinking water regulations. Variances and exemptions to the primary regulations may be granted by either the states of the Administrator. Another section (142.60) discusses best available technologies for a list of contaminants as they pertain to national primary drinking water regulations.
- **40 CFR Part 143—National Secondary Drinking Water Regulations, EPA**
  While Part 141 of this title is concerned with primary drinking water regulations, this part deals with national secondary drinking water regulations. Secondary regulations control contaminants that primarily affect the aesthetic qualities of drinking water. Levels are given for selected contaminants and it is explained that states may establish higher or lower levels depending on special conditions with that state, given that public health and welfare are not adversely affected. Section 143.4 gives monitoring requirements, stating that monitoring should occur no less frequently than the schedule used for the National Interim Primary Drinking Water Regulations.
- **40 CFR Part 144—Underground Injection Control Program, EPA**
  This regulation stems from Part C of the SDWA which allows for the establishment of an Underground Injection Control (UIC) program in each state. General program requirements are given in Subpart B. These requirements address prohibition of unauthorized injection, prohibition of movement of fluid into underground sources of drinking water, prohibition of Class IV wells and requirements for wells injecting hazardous waste. In 144.16 of this subpart, direction is given for cases when injection does not occur into, through, or above an underground source of drinking water. Subpart C takes a look at authorization of underground injection by rule where each class of wells is given particular authorization requirements. Authorization is also available by permit, and this is the topic of Subpart D. Going hand-in-hand with Subpart D is Subpart E, which discusses permit conditions. Finally, in Subpart F, financial responsibility for hazardous waste injection wells is described.
- **40 CFR Part 146—Underground Injection Control Program: Criteria and Standards, EPA**
  By working closely with Part 144 of this title, this part sets forth the technical criteria and standards for the Underground Injection Control Programs. Criteria are given for exempted aquifers, which include aquifers used as an "underground source of drinking water". Injection wells are then classified as Class I–V, based on the material injected. The following subparts (B–G) then outline, in detail, the criteria for each class of injection wells.
- **40 CFR Part 149—Sole Source Aquifers, EPA**
  Pursuant to 1427 of the SDWA, this regulation was written to provide criteria for identifying critical aquifer protection areas. Section 149.3 refers to a Critical Aquifer Protection Area as one which was designated as a sole or principal source aquifer prior to June 19, 1986 for which an area wide ground-water

quality protection plan was approved. The second definition entails major recharge areas of a sole or principal source aquifer designated before June 19, 1988. Edwards Underground Reservoir, of the San Antonio area, is given attention in Subpart B.

- **Environmental Statement Memorandum No. ESM94-5—Environmental Impacts on Groundwater, Department of Interior**
  Implements CEQ memorandum issued on November 19, 1976, concerning 1424(e) of the Safe Drinking Water Act. EISs will analyze impacts on the quantity and quality of ground water with specific emphasis on drinking-water sources. For proposed actions which may affect ground water, early consultation is to be initiated with the appropriate District Hydrologist, Water Resources Division (WRD). If it is possible that the proposed action may affect an aquifer which has been designated or is being petitioned for designation for special protection, early consultation should be initiated with the Regional EPA Administrator.
- **516 DM 2 Appendix 2 (2.2), Department of Interior**
  Environmental documents (EA, EIS, FONSI) must be prepared for actions which may adversely affect such unique geographic characteristics as sole or principal source drinking-water aquifers.
- **518 DM 1—Comprehensive Waste Management, Department of the Interior**
  This chapter defines waste to include solid and hazardous waste, hazardous materials, and hazardous substances. Departmental policies, responsibilities, and functions regarding waste management are presented, with an aim towards prevention of hazardous waste generation. The use of sound waste management practices is mandated.

DEFINITIONS

Best Management Practices (BMP)

Schedules of activities, prohibitions of practices, maintenance procedures, and other management practices to prevent or reduce the pollution of “waters of the United States”. (40 CFR Part 122.2)

Contaminant

Any physical, chemical, biological, or radiological substance or matter in water.(SDWA, 1401(6))

Discharge of a Pollutant

Any addition of any pollutant to navigable waters from any point source. (CWA, 502(12)(A))

Exemption

A document for water systems having technical and financial difficulty meeting national primary drinking water regulations effective for one year granted by EPA “due to compelling factors”.

Maximum Contaminant Level (MCL)

The maximum permissible level of a contaminant in water which is delivered to any user of a public water system. (SDWA, 1401(3))

Maximum Contaminant Level Goal (MCLG)

The level at which no known or anticipated adverse effects on the health of persons occur and which allows an adequate margin of safety.

National Pollutant Discharge Elimination System (NPDES)

The national program for issuing, modifying, revoking and reissuing, terminating, monitoring and enforcing permits, and imposing and enforcing pretreatment requirements, under sections 307, 402, 318, and 405 of CWA. (40 CFR Part 122.2)

Navigable Waters

Waters of the United States, including the territorial seas. (CWA, 502(7))

Person

An individual, corporation, partnership, association, state, municipality, commission, or political subdivision of a State, or any interstate body. (CWA, 502(5))

Point Source

Any discernible, confined and discrete conveyance, including but not limited to any pipe, ditch, channel, tunnel, conduit, well, discrete fissure, container, rolling stock, concentrated animal feeding operation, or vessel, or other floating craft, from which pollutants are or may be discharged. This term does not include agricultural stormwater discharges and return flows from irrigated agriculture. (CWA, 502(14))

Pollutant

Dredged spoil, solid waste, incinerator residue, filter backwash, sewage, garbage, sewage sludge, munitions, chemical wastes, biological materials, radioactive materials (except those regulated under the Atomic Energy Act of 1954), heat, wrecked or discarded equipment, rock, sand, cellar dirt and industrial, municipal, and agricultural waste discharged into water. It does not mean: (a) sewage from vessels; or (b) water, gas, or other material which is injected into a well to facilitate production of oil or gas, or water derived in association with oil and gas production and disposed of in a well, if the well used either to facilitate production or for disposal purposes is approved by authority of the State in which the well is located, and if the State determines that the injection or disposal will not result in the degradation of ground or surface water sources. (CWA, 502(6))

Public Water System

A system for the provision to the public of piped water for human consumption, if such system has at least fifteen service connections or regularly serves at least twenty-five individuals. (SDWA, 1401(4))

Publicly Owned Treatment Works (POTW)

Any device or system used in the treatment of municipal sewage or industrial wastes of a liquid nature which is owned by a "State" or "municipality". This definition includes sewer, pipes, or other conveyances only if they convey wastewater to a POTW providing treatment. (40 CFR Part 122.2)

Recharge Zone

The area through which water enters a sole or principal source aquifer.

Significant Hazard to Public Health

Any level of contaminant which causes or may cause the aquifer to exceed any maximum contaminant level set forth in any promulgated National Primary Drinking Water Standard at any point where the water may be used for drinking purposes or which may otherwise adversely affect the health of persons, or which may require a public water system to install additional treatment to prevent such adverse effect.

Sole or Principal Source Aquifer

An aquifer which supplies 50 percent or more of the drinking water for an area.

Streamflow Source Zone

The upstream headwaters area which drains into an aquifer recharge zone.

Toxic Pollutants

Those pollutants…which after discharge and upon exposure, ingestion, inhalation or assimilation into any organism…,will, on the basis of the information available to the Administrator, cause death, disease, behavioral abnormalities, cancer, genetic mutations, physiological malfunctions or physical deformations, in such organisms or their offspring. (CWA, 502(13))

Variance

A document for water systems having technical and financial difficulty meeting national primary drinking water regulations which postpones compliance when the issuing of which "will not result in an unreasonable risk to health".

Waters of the United States

a) all waters which are currently used, were used in the past, or may be susceptible to use in interstate or foreign commerce, including all waters which are subject to the ebb and flow of the tide; b) all interstate waters, including interstate "wetlands"; c) all other waters such as interstate lakes, rivers, streams…, mudflats, sandflats, "wetlands", sloughs, prarie potholes, wet meadows, playa lakes, or natural ponds the use, degradation, or destruction of which would affect…interstate or foreign commerce…(40 CFR Part 122.2)

Wetlands

Those areas that are inundated or saturated by surface or groundwater at a frequency and duration sufficient to support…a prevalence of vegetation typically adapted for life in saturated soil conditions. Wetlands generally include swamps, marshes, bogs, and similar areas. (40 CFR Part 122.2)

APPLICABLE ACTIONS

Clean Water Act

NPDES permits are required for the discharge of pollutants from any point source into waters of the U.S. [40 CFR 122.51(c)]

Section 404 permits are required for the discharge of dredged or fill material into waters of the U.S. (33 CFR 323.3)

Safe Drinking Water Act

Any action which may result in degradation of groundwater quality and/or pose a hazard to public health.

SUMMARY OF IMPLEMENTATION PROCEDURES

Clean Water Act

The procedures for determination of potential water quality impact from USGS activities and the need for various permits are extremely complicated and require close coordination with the Environmental Affairs Office, USGS, with regional offices of the EPA and the Army Corps of Engineers, and with state water pollution control agencies.

Safe Drinking Water Act (Consult ESM 94-5 for details)

1. In all cases where a project may involve ground water, one or more aquifers, or a recharge area, consultation with the appropriate District Hydrologist, WRD, should be initiated early in the planning stages.
2. Where a project may have any possibility of direct or indirect contamination of a sole or principal source aquifer, a recharge zone for such an aquifer, or a related streamflow source zone that has been designated or for which a petition for designation is being processed, early consultation should be initiated with the Regional Administrator, EPA.
3. If a project-related ground water impact is possible, an environmental assessment should be prepared. If an EIS is to be prepared, a copy of the Notice of Intent should be sent to the appropriate EPA Regional Administrator and he/she should review the draft EIS.

4. Final plans for projects having a potential for adverse impacts should incorporate appropriate mitigation measures.

INTERFACE BETWEEN FEDERAL WATER QUALITY REGULATION AND STATE ALLOCATION OF WATER QUANTITY

Melinda R. Kassen
Trout Unlimited, Colorado
Water Project
Boulder, Colorado

Healthy rivers and lakes require both clean and flowing water. In the United States, however, there are two separate and distinct systems for regulating water quality and water allocation, i.e., the amount of water taken out of a waterbody for beneficial use and, as a result, the amount of water left in the waterbody for environmental, aesthetic, or recreational purposes. States have the primary authority to allocate quantities of water for various users. Thus, state allocation decisions affect how much water flows in rivers and streams. At the same time, federal law—the Clean Water Act—regulates water quality, although the Act gives the states certain responsibilities and allows the Federal Government to delegate other responsibilities to the states, albeit in both cases with federal oversight. Because a river has aspects of both flow and quality; however, actions taken pursuant to either quality or quantity regulation can affect regulatory actions in the other system. Such instances of overlap create substantial friction both between levels of government and between the governments and those using or enjoying the nation's waters.

WATER QUANTITY

The 50 states each have developed their own systems of water allocation. However, the states in the western half of the country generally follow the system of prior appropriation, whereas the states in the wetter eastern half generally follow the riparian system of water allocation. Even federal agencies must procure their water rights through the state allocation systems.

In the temperate east of the country, states follow the riparian doctrine for water allocation. This common-law doctrine, imported from England with the first European settlers to North America, allows landowners bordering a river to take what water they need for use on their land—reasonable use—so long as such use does not impair others' use on the same river. Increasingly, riparian states are refining their doctrines and adopted statutory frameworks, including water use permits. Both common-law riparianism and the statutory variants provide that one owner's use should not adversely affect the water quality to the point where another riparian user cannot make reasonable use of the river. If such a situation develops, the injured riparian user can seek to enjoin the offending action and/or compensation for damages.

West of the 100th Meridian where natural precipitation is half or less than what is available further east (as well as in Alaska and Hawaii), states follow the doctrine of prior appropriation. Prior appropriation allows anyone to divert water from a river for beneficial use in priority. The priority system assigns water users seniority based on when they first started using the water; the "first in time, first in right" precept means that the most senior appropriator gets to fulfill his or her entire right before the next appropriator can take any water out of the stream.

Many western states do not have any public interest, environmental, or other balancing tests that prevent appropriators from drying up streams entirely. Although most western states do have some type of "instream flow" protection programs, these programs were typically added into the system within the past few decades and do not disturb existing diversionary rights.

WATER QUALITY

Before 1965, states had exclusive authority to regulate water quality except to the extent that the U.S. Army Corps of Engineers may have affected water quality through its regulation of navigable waters under the 1899 Rivers and Harbor Act. Even the 1965 federal legislation did not create federal legislative authority. So, it was not until 1972 that Congress passed the Federal Water Pollution Control Act, which, for the first time, created a federal floor for water quality protection and two mandatory permit programs. The goals of the Act were no less than to "restore and maintain the chemical, physical and biological integrity of the Nation's waters" and eliminate all discharges of pollutants into the nation's waters by 1985. [33 U.S.C. §1251 (2002)].

The Act, later amended and retitled the Clean Water Act, retained the preexisting scheme, albeit substantially strengthened, for states to adopt their own water quality standards—composed of the uses and chemical quality for all their water bodies—however, now subject to approval by the federal Environmental Protection Agency (EPA). In addition, federal agencies, or delegated state agencies, issued discharge permits to ensure compliance with the water quality standards. Finally, where a waterbody does not meet the water quality standards that the state has adopted for it, the Clean Water Act requires the state to list this waterbody as impaired and then to establish a "Total Maximum Daily Load" (TMDL) of pollutants that, once achieved, will result in the waterbody meeting its standards. This latter program is similar to the standards-setting program in that the states have primary authority for listing and setting TMDLs, but the EPA must ultimately approve the state's actions, and if the state's actions are insufficient to meet federal law, then EPA establishes the TMDLs.

The Clean Water Act creates two different permitting programs. The first, the National Pollutant Discharge Elimination System (NPDES), requires entities, including many industries but also municipal wastewater treatment facilities, discharging pollutants to a waterbody through a discrete conveyance or point source, often a pipe, to obtain a permit. This permit imposes "end of the pipe" effluent limits to ensure compliance with two distinct sets of requirements—the state-adopted water quality standards for the receiving water and a federally determined set of best-available technologies, established on an industry-by-industry basis. The second permit program regulates those who discharge dredged or fill material into waters of the United States, including wetlands tributary to rivers, lakes, and estuaries. Such permits must again protect the receiving water's water quality standards and must ensure that the discharge of dredged or fill materials does not have an "unacceptable adverse effect" on either municipal water supplies or the environment. Although Congress gave federal agencies the initial authority to administer both permit programs, the Act also encouraged states to seek delegation of the programs. Most states administer their own NPDES permit programs, rather than the EPA; only a handful of states have chosen to seek delegation of the permit program for dredged and fill materials from the U.S. Army Corps of Engineers. With regard to both types of permits, states must certify that the discharge, as ultimately regulated, will meet the state's water quality standards. In fact, states must certify compliance with water quality standards for all federal permits that affect water quality, even those not issued pursuant to the Clean Water Act, for example, hydropower licenses that the Federal Energy Regulatory Commission issues.

The Clean Water Act creates many additional programs and includes significant sources of federal funding to enable the states to meet their regulatory responsibility, and to enable municipalities to achieve the necessary levels of sewage treatment. In addition, 30 years after passage of the Act, most waters that remain impaired are so as a result of "non-point sources" of pollutants, for which the Clean Water Act does not require permits. These sources include runoff from farms, construction sites, and abandoned mines. Although the Clean Water Act did not mandate controls for these sources of contamination, it includes programs to encourage their voluntary cleanup.

INTEGRATION

In water rights adjudications and permit proceedings, in most western states, the decision maker has only a limited ability to consider water quality. At one extreme, California has fully integrated water quality protection into water rights decisions, such that the state agency responsible for granting water rights has in a few instances required a diverter to leave water instream for water quality protection. In addition, since 1983, California implements its common-law public trust doctrine to require water quality protection in the context of water diversions. Thus, the California Supreme Court found that the California Water Resources Board had an affirmative duty to protect Mono Lake's water quality and environment from adverse effects caused by diversions by the City of Los Angeles, although the court still recognized a need to balance water rights with environmental protection and did not limit existing diversions [*National Audubon Society v. Superior Court*, 658 P.2d 709 (Cal. 1983)]. Other states such as Colorado and Idaho have explicitly rejected this approach, whereas still others, e.g., Oklahoma and Nebraska, simply have not articulated a position in the decades since the Mono Lake ruling.

In fact, in most other states, the water quality protections embedded in water rights and permit proceedings are limited. Only 7 of the 17 western states have either statutory or public interest requirements to protect a senior diverter's water quality from injury as a result of changes to existing water rights.

Most states require a new diverter to protect water quality at least to the point that the quality is sufficient for a senior diverter's purposes. However, in Colorado, a senior water user cannot seek protection from a lessening of water quality because of new diversions, or changes to existing diversions, although that user is entitled to have water quality maintained when there is an exchange of water on the river, i.e., when the senior is receiving water from a different source because of changes elsewhere in the system. Thus, senior diverters on the Arkansas River were denied relief as a result of a change to water rights that resulted in water formerly used for irrigation on the plains being diverted in the headwaters out of the basin entirely, even though the change substantially reduced the water quality of the flows remaining the basin.

Although 1970s era slogan proclaimed that, "Dilution is not the solution to pollution," entities discharging pollutants to a stream pursuant to a Clean Water Act permit have also sought protection from changes to water rights that would remove dilution flows. In Colorado, at least, this effort has not been successful. Thus, the courts denied relief to a discharging industry that objected to a change of water rights that would have removed stream flow such that the discharger would have faced significant additional water treatment costs to be able to meet its permit limits. The court remarked on the subservient nature of water quality regulation to the water rights allocation system, which did not allow consideration of water quality, or injury to a discharger, as opposed to another water user, in determining whether to grant a water right [*City of Thornton v. Bijou Irrigation Co.*, 926 P.2d 1, 90–94 (Colo. 1996)].

On the water quality front, even as it adopted a federal regulatory scheme for water quality protection, Congress was mindful of the states' role in water allocation. Thus, the 1972 law includes a section that its provisions not "be construed as impairing or in any manner affecting any right or jurisdiction of the States with respect to the waters . . . of such States" [33 U.S.C. §1370(2) (2002)]. Yet, by 1977, after 5 years of Clean Water Act implementation, some western water user interests had become alarmed as to the potential for Clean Water Act regulation to "interfere" with the exercise of water rights. Congress added an additional provision to the Act, as part of a broad package of amendments, that no regulatory actions it authorized should "supercede, abrogate or impair" the exercise of state water rights [33 U.S.C. §1251(g)(2002)]. However, for several reasons, including that this language is a goal and that legitimate water quality regulation may affect water rights, the Courts have upheld water quality regulation to occur that appears at least to affect individual water users. For example, a farmer could not discharge dredged and fill material without a permit, even though he was doing so to better use his water rights [*United States v. Akers*, 785 F.2d 814 (9th Cir. 1986)]. And, a state could impose conditions in a water certification for a federal permit that had the effect of limiting the amount of water the permittee could use [*PUD No. 1 of Jefferson County v. Wash. Dep't of Ecology*, 114 S.Ct. 1900 (1994); *see also Public Util. Dist. No.1 of Pend Oreille County v. Wash. Dep't of Ecology*, 51 P.3d 744 (Wash. 2002)]. Finally, an entity seeking to build a new dam must obtain a Clean Water Act dredge and fill permit even if that entity already has a state water right to build the dam [*Riverside Irrigation Dist. v. Andrews*, 758 F.2d 508 (10th Cir. 1985)]. In making these findings, the courts repeatedly have held that legitimate regulation under the Clean Water Act may proceed even if there are indirect effects on the exercise of a water right.

In contrast to the courts' rulings, state legislatures have often taken action expressly to exempt water rights holders from regulation under state Clean Water Act programs. Thus, Colorado adopted the federal Act's language regarding not to "supercede, abrogate or impair," but not in the goals section of the Act; rather the provision is part of the state Water Quality Control Act's implementation [Colo. Rev. Stat. §25-8-104 (2004)]. Similarly, after the U.S. and state Supreme Court rulings in Washington confirming the state water quality agency authority to impose minimum stream flows that might affect the yield of water rights through water quality certifications, the state legislature adopted statutory amendments forbidding the state agency from adversely affecting water rights in water quality regulation [Wash. Rev. Code §90-48-422 (2004)].

A recent set of cases that illustrate the potential tension between water quality regulation and state water allocation laws involve artificial transfers of water, i.e., the man-induced movement of water from one waterbody to another. In each case, an entity was moving water from a contaminated watershed to a relatively cleaner waterbody through a conveyance—via a snowmaking system at a ski area, a municipal drinking water collection system, and a flood control system. Three U.S. courts of Appeal ruled that the discharger had to obtain an NPDES permit, because there was a discharge of pollutants from a point source to waters of the United States. The U.S. Supreme Court considered one of these cases in 2004 but did not reach the ultimate issue, although the Court did hold that the transferring water from one distinct water to another through a pipe did constitute a point source subject to permitting, even though the discharger was not, itself, "adding" pollutants to the water being transported [*South Florida Water Mgmt. Dist. v. Miccosukee Tribe of Indians*, 124 S.Ct. 1537, 1543 (2004)]. In other words, it was enough that the water being discharged was polluted already and was being discharged to a cleaner system.

As all three of the cases developed east of the Mississippi, the dischargers did not argue that they had water rights that allowed them to operate their systems free from Clean Water Act regulation, but many western states and water users weighed in to make that point. In fact, diverters have argued that requiring all transbasin diversions of water to operate only under an NPDES permit would overwhelm the states' permitting systems and adversely affect the use of state water rights. Those in support of requiring such permits have suggested that many of these diversions with discharges could be accomplished under states' general permitting programs in a way that would not be onerous. Only one state, Pennsylvania, has historically required permits for these kinds of discharges.

There are other places in the Clean Water Act program where water rights issues collide with water quality regulation. The Clean Water Act requires states to establish water quality standards, including use designations and numeric or narrative criteria for individual pollutants (33 U.S.C. §1313). To the extent that the exercise of water rights has affected what uses a waterbody can attain, because of the loss of dilution flow or fundamental changes in the natural system because of a dam, diversion, or artificial increase in the quantity of water coursing through the channel, states wrestle with how to accommodate human activity while encouraging operations that will maintain the Clean Water Act's "fishable/swimmable" goal for water quality.

In the program that requires states to list waters that remain impaired after all permitted discharges have permits [33 U.S.C. 1313(d)] EPA has interpreted the Clean Water Act not to require listing for impairment because of pollution, the definition of which includes hydrologic modifications such as dams and diversions that allow parties to exercise their water rights [64 Fed. Reg. 46011 (August 23, 1999); 65 Fed. Reg. 43586, 43592 (July 13, 2000)]. The problem is that the exercise of water rights that result in changes to the flow regime can lead both to impairment and to certain pollutant levels above the adopted standards. Examples include excessive temperatures, sediment, and nutrient levels. As temperature, sediment, and nutrients are pollutants, there have been listings, for example, in the Yakima River Basin in Washington State, where the underlying issue is most likely a change in the flow regime as a result of the exercise of water rights. So as not to challenge the exercise of water rights, other states, including California, New Mexico, and Washington, try to solve high temperatures with requirements to plant shade trees along a river's banks. See, e.g., http://www.ecy.wa.gov/programs/wq/tmdl/-watershed/index.html. However, EPA has rejected state proposals to solve impairment with flow enhancement programs. See http://www.-deq.state.mt.us/wqinfo/TMDL/approvalLtrs/Approval_BigCreekYelBasin.pdf.

Thus, state resistance to integrating water quality regulation, a federal mandate, with state water allocation systems has hampered protection of the environment as well as protection of the water quality necessary to sustain competing water uses. However, fears regarding the costs associated with water users having to protect the environment have kept most state legislatures from adopting statutes that integrate water quality protection into their water allocation systems in a way that would impose additional burdens on water development.

REGULATORY ISSUES AND REMEDIATION: RISK, COSTS, AND BENEFITS

David B. Vance
ARCADIS
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Major regulatory changes have occurred over the past decade regarding enforcement and cleanup standards. These changes have occurred in part, because of the realization that enormous financial and human resources were being spent on the remediation of properties with low concentrations of residual fuel hydrocarbon (primarily gasoline and diesel) contamination with low impact to human receptors. A major report from a national laboratory in 1995 changed the regulatory approach to focus attention on the risk benefits and exposure pathways associated with subsurface fuel contamination. Risked based decision making takes into account the risk, costs, and benefits when evaluating regulatory approaches to remediation

INTRODUCTION

A Lawerence Livermore National Laboratory (LLNL) report (1) found that once fuel leak sources have been removed, petroleum hydrocarbon contamination generally does not spread far from the leak sites. Given time, naturally occurring indigenous microbes in the soil and water usually break down most of the contamination before drinking water sources are impacted. The report noted that only 136 sites of the 28,051 evaluated (<0.5 of 1%) affected drinking water supply wells. The LLNL study concluded that at many low-risk sites, active cleanup could be stopped and natural attenuation by naturally occurring biological and geochemical methods would reduce the contamination over time (1).

The LLNL study was important at the time. About 1200 out of 20,000 leaking underground fuel sites were examined in detail. It appeared that in most cases, the length of the hydrocarbon plumes was limited to several hundred feet and less than 50 feet vertically. Natural bioattenuation, abiotic degradation, and diffusion processes stabilized or attenuated these plumes over time (2). LLNL concluded that the threat to human health and long-term effects on the environment caused by petroleum hydrocarbon contamination may have been vastly overestimated. Water supplies can generally be purchased for about $50 to $200 per acre foot, or even $1000 per acre foot at the most expensive price. Cleanups of petroleum contamination of groundwater were costing up to $600,000 per acre feet to reclaim groundwater and to meet cleanup goals (2).

Critics to the LLNL report noted that LLNL had huge environmental cleanups to contend with, so their recommendations may have been politically motivated, none of the actual data discussed in the LLNL report were included in the report for examination, the report relies on written and verbal communications and papers supplied by the petroleum industry, and no public health professionals from CAL EPA, the Department of Health Services (DHS), or other agencies were solicited for their recommendations or review (3).

Even with the controversy, the LLNL report did focus public policy discussion on how clean is clean and what financial and human resources are worth using to clean up hydrocarbon contamination. These discussions led to consideration of risked based cleanup standards.

Changing from enforcement driven by arbitrary cleanup standards to action (or no action) driven by health-based risk assessments in conjunction with a cost/benefit analysis. The current debate includes the application of this concept at the highest level for the promulgation of new regulations or modification of existing regulations. However, the more powerful application of the new regulatory paradigm is on a site-specific basis. The purpose of risked based regulatory decision making is to examine the components of the site-specific risk/benefit/cost process. The following discussion is predicated on two basic assumptions: First, if a release of contaminants has occurred at a site, the source of that release will be abated. Second, free product hydrocarbons are not allowed to remain on or near the water table.

These actions quickly limit the potential size of a contaminant plume, minimize further migration of the core of that plume, and recover contaminants as free product versus dissolved phase. All of which offer significant cost savings in the remediation effort. A health-based risk assessment determines what concentration of contaminant represents an acceptable risk level. The cost side of the process is driven by the effort required to achieve that risk-determined contaminant concentration. Health-based risk assessment can be a complex process. However, it has two key technical components: first, is the dose/response relationship of a chemical and second, is the pathway through which that chemical can ultimately enter a human body. A third nontechnical issue is how accepting the public is of the process, particularly those exposed to the risk. Dose/response relationships are categorized into two broad areas: noncarcinogenic and carcinogenic.

Noncarcinogenic responses are generally characterized by a threshold dose. The threshold represents a level of exposure to a contaminant that can be tolerated by a person with no adverse effects until the protective mechanisms of the person are overwhelmed. As individual levels of tolerance vary, threshold determination can be controversial.

Carcinogenic responses are assumed to have no threshold. This assumption means that some finite cancer risk exists no matter how small the dose.

The primary routes of exposure to groundwater are ingestion, inhalation of vapors during bathing or showering, and dermal exposure during bathing or showering. Direct ingestion is straightforward, and inhalation and dermal adsorption are dependent on physical properties of the contaminant such as vapor pressure and partition coefficients.

The public perception of risk is contradictory. Significant risk to which a person will voluntarily expose themselves, such as driving, improper diet, or smoking are of concern, but deemed acceptable. In contrast, much smaller risks associated with environmental impairment that is imposed, and not voluntarily engaged, often generate alarm at any level above zero risk. This issue can only be resolved by education to each person's aggregate risk during their lifetime from many sources with a view toward the benefits each receives from our industrial society. This should be an important part of the current risk versus cost and benefit debate.

In a practical sense, the most effective method of controlling the risk from contaminated groundwater is simply to not use it, which in many situations may be the most viable alternative. However, our groundwater resources are finite and are diminishing in quality with time. One solution that is likely to become more common is surface treatment to remove contaminants. The most cost-effective goals are to maintain as much of our groundwater as possible in pristine condition and the timely restoration of contaminated aquifers to acceptable conditions.

Cost effectiveness enters from establishing a health-based contaminant concentration that is considered acceptable and expending the minimum effort sufficient to ensure that level is maintained. Whenever possible, natural attenuation should be used for the final clean up of the core and the dilute distal portions of the plume.

Groundwater recovery or *in situ* treatment should be designed with three goals in mind:

1. Control and treat the core of the plume in which the dissolved concentrations exceed the capacity of the aquifer to naturally attenuate in a timely fashion.
2. If required, recover groundwater at a rate just sufficient to achieve the above goal.
3. Design a system that takes into account the heterogeneity of the aquifer matrix.

A fine-grained matrix will in general take a significant amount of time to remediate because the overall mass transport rates the matrix will support are low and the amount of adsorbed contaminant is high. However, groundwater recovery rates or *in situ* remediation systems need only be sufficient to meet the low mass transfer requirements. On the other hand, the migration distances are likely to be minimal.

The worst case exists in aquifers that contain strata of fine-grained material mixed with highly permeable continuous layers. Pumping rates must be high to obtain hydraulic control, but treatment durations are extended because of the slow leaching of adsorbed contaminants from fine-grained material. In aquifers of this type with pronounced hydraulic gradients and subsequent high groundwater flow rates, it may be necessary to allow much of the contamination to spread through the aquifer such that concentrations will be at levels amenable to natural attenuation. A health-based risk assessment will be key in this scenario, which in essence uses an extensive portion of the aquifer matrix as a chemical/biological reactor for contaminant destruction.

For permeable heterogenetic aquifers with small hydraulic gradients and low groundwater flow rates, it will be more economical to run a groundwater recovery system or an *in situ* treatment system for short durations on a periodic basis. The core of the process is the diffusion of contaminants from fine-grained units into the more highly permeable portions of the aquifer. This contaminated pore volume is then treated to restore the contaminant concentrations to low levels providing a concentration gradient that drives the diffusion process. Initial capital costs must be met, but operation and maintenance costs will be significantly decreased over the life of the project.

If natural attenuation processes stop on a site because of a lack of a suitable electron acceptor, such as oxygen for aerobic bioremediation, other methods are available to enhance site conditions and to continue with the natural attenuation approach. The consumption of any available oxygen within the core of a hydrocarbon plume is quite common. In these cases, even the nitrate and sulfate concentrations may be low within the core of the plume because of the microbial consumption of these alternative electron acceptors. For these projects that are too high in concentration to close, but too low for active and aggressive remediation methods, enhanced bioremediation holds promise of lowering concentrations to regulatory acceptable levels, wherein natural attenuation can be used. More detailed information about natural and enhanced bioremediation methods can be found in Suthersan (4).

Risk-based decision making and natural attenuation has undergone changes over the past several years. Risk-based corrective action (RBCA) is a U.S. Environmental Protection Agency (EPA) cleanup process designed to protect human health and the environment while allowing for practical and cost-effective site-specific remedial measures. RBCA has an expanded use of natural attenuation as a cleanup option.

Some regulatory agencies are moving toward the use of risk-based screening levels and decision making. This approach uses lookup tables of conservative, risk-based screening levels (RBSLs) for over 100 chemicals commonly found in impacted soil and groundwater at sites where releases of hazardous substances have occurred. The approach is intended to help expedite the preparation of environmental risk assessments at sites where impacted soil and groundwater has been identified. As an alternative to preparing a formal risk assessment, soil and groundwater data collected at a site can be directly compared with the RBSLs and the need for additional work can be evaluated. It is anticipated that documents like these will be especially beneficial for use at small- to medium-sized sites, where the preparation of a more formal risk assessment may not be warranted or feasible because of time and cost constraints (5).

This new regulatory approach will rationalize the process of environmental restoration. Regulators, industry, and the public will benefit. For it to succeed, we all must understand and accept reasonable levels of risk. Cost effectiveness will hinge on the understanding and exploitation of site-specific aquifer characteristics, and acceptance of cleanup time frames that allow for the use of natural attenuation.

BIBLIOGRAPHY

1. Rice, D.W., Gose, R.D., Michaelsen, J.C., Dooher, B.P., MacQueen, D.H., Cullen, S.J., Kastenberg, W.E., Everett, L.G., and Marino, M.A. (1995). *California Leaking Underground Fuel Tank (LUFT) Historical Case Analyses*. Submitted to the California State Water Resources Control Board, Underground Storage Tank Program and the Senate Bill 1764 Leaking Underground Fuel Tank Advisory Committee. Livermore, CA, November 16.
2. Weiss, R.B. (1996). State Regulatory Policy Takes A Giant Step to Close Low Risk Sites. *California Environmental Business Council News*.
3. Elliott, J. (1995). State Loosens Gasoline Spill Cleanup Rules. *Albion Monitor*. December 21.
4. Suthersan, S. (2002). *Natural and Enhanced Remediation Systems*. CRC Press, Boca Raton, FL.
5. Brewer, R. (2003). Screening For Environmental Concerns at Sites With Contaminated Soil and Groundwater; regional Water Quality Control Board. San Francisco Region. California. Available: http://www.swrcb.ca.gov/rwqcb2/esl.htm.

THE SAFE DRINKING WATER ACT

Rolf R. von Oppenfeld
The TESTLaw Practice Group
Phoenix, Arizona

OVERVIEW OF THE SAFE DRINKING WATER ACT

The Safe Drinking Water Act (SDWA) was enacted in 1974 and subsequently amended in 1986 and 1996 to protect the public from risks posed by drinking water contamination. The Act required the Environmental Protection Agency (EPA) to establish standards or treatment techniques for contaminants that could adversely affect human health, establish requirements for monitoring the quality of drinking water supplies and for ensuring the proper operation and maintenance of water systems, and protect groundwater from subsurface injection of polluting liquids. EPA posted a general guide to the SDWA provisions at http://www.epa.gov/OGWDW/SDWAsumm.html and posted its general approach for implementing the provisions at http://www.OGWDW/sdwa/sdwa.html.

The 1996 Amendments are important because Congress gave the EPA greater flexibility under the SDWA and emphasized the cost/benefit analytical approach to regulation. For an overview of the program, see the Office of Groundwater and Drinkingwater at http://www.epa.gov/safewater.html. The SDWA regulations can be found at 40 C.F.R. Parts 141–149 and may be accessed on the internet at http://www.water.access.gpo.gov/nara/cfr or http://epa.gov/OGWDW.regs. html. Inquiries about the regulations can also be directed to the Safe Drinking Water Hotline at 1-800-426-4791.

The EPA has delegated primary compliance and enforcement responsibility for the SDWA program to qualifying states. A detailed description of state primacy requirements can be found at http://www.epa.gov/safewater/pws/primacy.html. If states fail to ensure proper standards, the EPA may bring independent enforcement against violators through the Act's federal enforcement procedures. With respect to the regulated community, the SWDA focuses on operators of public water systems (PWSs) and on persons engaged in underground injections of fluids into wells.

Regulation of Public Water Systems

The SDWA regulates public water systems (PWSs) that provide water for human consumption through pipes or other constructed conveyances and regularly provide service to more than 25 persons or have 15 or more service connections [42 U.S.C. § 300f(4)(A)]. The EPA separates PWSs into the following categories:

1. community water systems and
2. noncommunity water systems that may be either
    a. nontransient, or
    b. transient.

Community water systems (CWSs) supply at least 15 service connections used by year-round residents or regularly serve at least 25 year-round residents [42 U.S.C. § 300f(15)]. Of approximately 180,000 water systems in the United States, only 55,000 are considered CWSs. They range from large municipal systems that serve millions of people to small systems that serve fewer than 100 people. CWSs can be publicly owned systems, including systems owned and operated by municipalities, townships, counties, water districts, and water authorities; privately owned systems that may be owned and operated by groups ranging from investor-owned water companies to homeowners associations; or ancillary systems that provide water as an ancillary function of their principal business or enterprise. Ancillary systems primarily service mobile home parks and a variety of institutional water providers.

Noncommunity PWSs are those systems that are not community water systems [42 U.S.C. § 300f(16)]. Examples of noncommunity PWSs include small systems in recreational areas and transient situations. Noncommunity systems may not have year-round residents but would have at least 15 service connections used by travelers or intermittent users for at least 60 days a year or serve an average of 25 individuals for at least 60 days a year. Specific examples of noncommunity systems include restaurants, campgrounds, motels, schools, and factories with their own wells or other water supply.

Nontransient, noncommunity water systems (NTNCWS) regularly serve at least 25 of the same people for 6 months of the year (Id). Schools and factories that serve water to 25 or more of the same people for 6 or more months of the year are nontransient, noncommunity water systems. All remaining water systems are classified as transient noncommunity systems (TNCWS). TNCWSs do not serve the same people on a day-to-day basis. In fiscal year 2000, there were more than 54,000 CWSs, more than 20,000 NTNCWS, and more than 93,000 TNCWS.

Certain PWSs are excluded from coverage and need not comply with SDWA regulations (42 U.S.C. § 300g). Excluded PWSs

1. consist only of distribution and storage facilities (have no collection and treatment facilities);
2. obtain all their water from a regulated PWS (but are not owned or operated by that PWS);
3. do not sell water; and
4. are not interstate carriers of passengers (Id).

This exclusion means that facilities such as hotels, factories, schools, and other businesses are not subject to SDWA regulation merely because they operate a storage tank and act as conduits from a PWS to consumers. However, the exclusion does not apply to private facilities that maintain their own wells or water supply. The EPA provides guidance for its definition of PWSs at http://www.epa/gov/OGWDW/guide/define.html.

SDWA DRINKING WATER STANDARDS AND GOALS

The SDWA authorizes the EPA to set drinking water standards called National Primary Drinking Water Regulations (NPDWRs). The EPA's NPDWRs are located at 40 C.F.R. Parts 141 and 142. NPDWRs are designed to protect consumers against the adverse effects caused by contaminants in drinking water. The SDWA requires that the EPA publish NPDWRs for contaminants that may have an adverse affect on human health, occur, or are likely to occur in PWSs at harmful levels and where the health risk can be reduced by regulation.

The 1996 Amendments made dramatic changes in the SDWA implementation process. Congress required the EPA to consider the following criteria when setting future drinking water standards and goals:

1. use of best available, peer-reviewed science [42 U.S.C. § 300g-1(b)(3)(A)(i)];
2. the assessment of health effects of populations at greater risk [42 U.S.C. § 300g-1(b)(3)(C)(i)(V)];
3. a list of treatment technologies for small systems, considering the quality of source water [42 U.S.C. § 300g-1(b)(4)(E)(ii)];
4. an assessment of the incremental costs and benefits of each alternative MCL considered [42 U.S.C. § 300g-1(b)(3)(C)(i)(IV)].

Maximum Contaminant Levels (MCLs)

A NPDWR must specify either a maximum contaminant level (MCL) or a treatment technique [42 U.S.C. §§ 300f(1), 300g-1(a)(3)]. The EPA is authorized to promulgate a NPDWR "that requires the use of a treatment technique in lieu of establishing a MCL," if the EPA finds that "it is not economically or technologically feasible to ascertain the level of contaminant" [42 U.S.C. § 300f(1)(C)]. Two significant NPDWRs that rely on treatment techniques are the Surface Water Treatment Rule and the Lead and Copper Rule.

MCLs are enforceable standards and represent the maximum permissible level of a contaminant in water delivered to any use of a PWS (40 CFR § 141.2). MCLs must be set as close to the maximum containment level goals as feasible. "Feasible" means using the best technology, treatment techniques, or other means which the administrator finds available after considering costs [42 U.S.C. § 300g-1(b)(4)(D)]. There are limits, however, to the use of cost considerations. First, the costs considered must be limited to those likely to be incurred [42 U.S.C. § 300g-1(b)(3)(C)]. Second, no consideration of costs is allowed to establish a MCL for "contaminants that are disinfectants or disinfection byproducts, or to establish a maximum contaminant level or treatment technique requirement for controlling *Cryptosporidium*" [42 U.S.C. § 300g-1(b)(6)(C)].

Under certain conditions, MCLs may be set at a level different from the feasible level. A variation is allowed where the means employed to achieve the feasible level of one contaminant would increase concentrations of other contaminants that pose a threat to human health. A level other than the feasible level may also be set when the means of reaching the feasible level for one contaminant interferes with the treatment processes necessary to comply with other NPDWRs. In these situations, the MCL or treatment process specified must "minimize the overall risk of adverse health effects" [42 U.S.C. § 300g-1(b)(5)(B)(i)].

The EPA must publish findings based on a scientific risk assessment of the probable health benefits and probable costs of each alternative MCL under consideration [42 U.S.C. § 300g-1(b)(3)(C)]. The published findings should be based on all of the following:

1. The EPA must use the best available, peer-reviewed science and studies.
2. Data must be collected by accepted methods.
3. Information on health effects must be presented by the EPA in an objective, comprehensive, and understandable manner.
4. Analysis for MCLs must show the quantifiable and nonquantifiable health risk reduction benefits, costs, and any increased health risk of compliance with the MCL.
5. For treatment techniques, analysis must show the health risk benefits and costs likely to result from compliance and alternative treatment techniques.
6. Valid approaches for measuring benefits must be identified.

If the EPA determines that the costs of a promulgated MCL outweigh the benefits, the EPA may promulgate a new MCL [42 U.S.C. § 300g-1(b)(6)]. The choice of treatment technique is left to the PWS, but the EPA must list available technology and methods to meet the MCL [42 U.S.C. § 300g-1(b)(4)(E)]. In addition, this list must include affordable methods for small

PWSs (Id). The EPA maintains a list of MCLs at http://www.epa.gov/safewater/mcl.html.

Judicial Review of MCLs

Section 1448(a) of the SDWA provides for judicial review of EPA's MCLGs and MCLs [42 U.S.C. § 300j-7(a)]. Only the U.S. Court of Appeals for the District of Columbia may review MCL determinations. A petition for review must be filed within 45 days, starting on the date of the promulgation of the regulation, unless solely based on grounds occurring after expiration of the 45-day period. There is no judicial review of a regulation in any enforcement proceeding (Id). Judicial review of MCLs is conducted under a narrow standard of review. The Court of Appeals will reverse only if an EPA action is arbitrary, capricious, an abuse of discretion, or not in accordance with the law. The EPA is given great deference because of the level of technical expertise it uses to evaluate complex scientific data.

The Drinking Water Contaminant Candidate List

In 1996, Congress set its most recent schedule for setting standards for drinking water contaminants. See 42 U.S.C. § 300g-1(b)(1)(B). First, the EPA was required to publish a list of contaminants that are not subject to any proposed or promulgated NPDWR, that are known or anticipated to occur in PWSs, and that *may* require regulation under the SDWA. The EPA was required to publish a Drinking Water Contaminant Candidate List (CCL) in February 1998 and every 5 years thereafter. The list must be published only after consultation with the scientific community and public comment. Following publication of the list, the EPA must determine which of the listed contaminants to regulate. The determination must be completed every 5 years for at least five listed contaminants.

Unregulated Contaminant Monitoring Regulations

The 1996 amendments also required EPA to develop a program to monitor unregulated contaminants and publish a list of monitored contaminants [42 U.S.C. § 300j-4(a)(2)]. In response, the EPA promulgated the Unregulated Contaminant Monitoring Regulations (UCMR) [64 Fed. Reg. 50556 (Sept. 17, 1999)]. The UCMR works in conjunction with the CCL. The EPA can use the data generated through the UCMR to prioritize contaminants on the CCL and help determine which contaminants to regulate. Several states may also petition to have certain contaminants monitored under the UCMR [42 U.S.C. § 300j-4(a)(2)(B)]. The UCMR covers all CWS and NTNCWS that serve more than 10,000 customers. An additional, 800 systems serving under 10,000 customers were randomly selected to test for unregulated contaminants.

The UCMR created three lists of contaminants for monitoring. Substances were placed on lists depending on their known environmental occurrence and the ability to test analytically for them. List 1, the "Assessment Monitoring" list, contains those substances known to occur in more than two states or more than 10 PWSs, and for which monitoring methods were available. See 40 C.F.R. § 141.40 Table 1. EPA has little monitoring data for List 2, the "Screening Survey," and List 3, the "Pre-Screen Testing List."

Chemical Monitoring Reform (CMR)

To reduce unnecessary monitoring and reporting requirements, the 1996 Amendments contained a provision for modifying NPDWR monitoring requirements. This CMR provision streamlined drinking water monitoring requirements for 64 chemicals that may occur in the *source water* of PWSs. For a system to qualify for alternative monitoring, it must demonstrate that a contaminant is not present in its water supply or if it is present, that it is reliably and consistently below the MCL. If the contaminant is later detected at or above the MCL, then the system must eliminate the contamination, or it must monitor as required under the NPDWR.

Secondary Regulations

Section 1412(c) of the SDWA also authorizes the EPA to promulgate National Secondary Drinking Water Regulations (NSDWRs) [42 U.S.C. § 300g-1(c)]. A NSDWR is "a regulation which applies to PWSs and which specifies the maximum contaminant levels which, in the judgment of the Administrator, are requisite to protect the public welfare" [42 U.S.C. § 300f(2)]. The NSDWR "may apply to any contaminant in drinking water which may adversely affect the odor or appearance of such water and consequently may cause a substantial number of persons served by the PWSs providing such water to discontinue its use, or which may otherwise adversely affect the public welfare" (Id). NSDWRs are not federally enforceable but offer additional guidance based on odor, aesthetics, and appearance to water systems and states. NSDWRs may vary according to geography and other circumstances. To date, NSDWRs have been established for several contaminants, including copper, fluoride, pH, and total dissolved solids. These regulations are in 40 C.F.R. Part 143.

Consumer Information

A major objective of the SDWA is to increase communication between PWSs and their customers. Consequently, Congress required an annual "Consumer Confidence Report" for PWSs. EPA regulations require each CWS to mail to each customer, at least once annually, a report on the level of contaminants in the drinking water purveyed by that system. These reports also indicate the source of the water supplied to the consumer and notify customers that a source water assessment is available.

VARIANCES AND EXEMPTIONS

General Variance

A state that has primacy over the SDWA may issue a variance if it determines that a system cannot comply with a MCL, despite application of the best available technology (BAT). The EPA determines the BAT and treatment method for each NPDWR after examination for efficacy under field conditions, rather than solely under laboratory conditions, and after considering cost. (40 C.F.R. § 141.2). A state may not issue a variance where an unreasonable risk to health exists or where alternative sources of water are reasonably available to the system. Under Section 1415, EPA must propose and promulgate its BAT finding at the same time it proposes and promulgates a MCL for each contaminant (42 U.S.C. § 300g-4). BATs may vary among systems, depending upon the number of persons served by the system or for other physical conditions related to engineering feasibility and costs of complying with MCLs.

The 1996 amendments allow the EPA to review and object to any variance granted by a state [42 U.S.C. § 300g-4(e)(10)]. If the state does not resolve these objections, the EPA may overturn the variance when it violates SDWA requirements. Consumers served by a PWS that has been granted a variance may petition EPA to begin the objection and review process.

A variance will be granted for a maximum of 5 years with the possibility of a 3-year extension. At the time a variance is granted, it will contain a schedule of steps necessary to achieve ultimate compliance with the particular NPDWR applicable in the variance. The variance may also contain additional control measures for protecting public health during the period that the variance is in effect. The EPA's final rule, issued in August 1998, states that 5 years is a reasonable time frame for a PWS to come into compliance; however, the state has the flexibility to set a feasible schedule if it determines that more than 5 years are required [63 Fed. Reg. 43841 (Aug. 14, 1998)].

Small System Variance

Congress responded to concerns that NPDWR compliance costs unfairly burden small PWSs. The 1996 Amendments added a special variance applicable only to small systems. See 42 U.S.C. § 300g-4(e). A small system variance may be available to systems that serve fewer than 3,300 persons (or up to 10,000 with special EPA approval). The variance should be of special interest to small community systems as well as the schools, factories, and other institutions that qualify as noncommunity water systems under the Act. The small system variance is available only where EPA has identified a less expensive version of the best available technology, called "variance technology," and the water system cannot afford to comply with the regular NPDWR. All variances must continue to provide adequate protection of human health. The proposed maximum period for a small system variance is 3 years with the possibility of a 2-year extension to complete improvements, obtain an alternate water source, or obtain federal or state financial assistance.

Exemptions

Under Section 1416(a), a state may exempt a PWS from any MCL or treatment technique requirement if it finds that

1. due to compelling factors (which may include economic factors), the system is unable to comply;
2. the system was in operation on the effective date of the MCL or treatment technique, or, for a newer system, that no reasonable alternative source of drinking water is available to that system;
3. the exemption will not result in an unreasonable risk to health; and
4. restructuring and management changes cannot reasonably be made that will result in compliance or otherwise improve the quality of the water (42 U.S.C. § 300g-5).

Under section 1418(b), a state is to prescribe a compliance schedule and a schedule for implementing any required interim control measures at the same time it grants an exemption [42 U.S.C. § 300g-7(b)].

In a major change in the exemption provisions, the 1996 Amendments deleted provisions which limited an exemption to 12 months, subject to a 3-year extension [42 U.S.C. § 300g-5(b)(2)(A)]. The new provisions require that the schedule for an exemption include compliance with each contaminant level and treatment technique for which the exemption was granted, as soon as practicable, but not later than 3 years after the otherwise applicable compliance date. The only exception to this time period is for small systems that serve less than 3300 persons, under certain specified conditions, for which extensions may be renewed for one or more additional 2-year periods, but not to exceed a total of 6 years [42 U.S.C. § 300g-5(b)(2)(C)].

The amendments now require that the state also consider whether the PWS is a "disadvantaged community" and whether management or restructuring changes can be made that will result in compliance or, if compliance cannot be achieved, would improve the quality of the drinking water [42 U.S.C. § 300g-5(b)(4)]. Measures to develop an alternative source of water supply must also be considered.

The 1996 Amendments also gave states more flexibility to adopt alternative protective strategies. For example, a state exercising primary enforcement responsibility for a PWS may, case-by-case and after notice and opportunity for public comment, establish watershed management programs as an alternative to filtration for systems that have uninhabited, undeveloped watersheds. See 42 U.S.C. § 300g-1(b)(7)(C)(v).

UNDERGROUND INJECTION CONTROL (UIC) UNDER THE SDWA

Scope of the UIC Program

SDWA established the UIC program to protect existing and future underground sources of drinking water from contamination caused by the subsurface disposal of waste fluids into wells. See 42 U.S.C. §§ 300h-300h-5. Improper operation of an injection well that allows movement of contaminants into an underground source of drinking water can lead to substantial civil penalties. The SDWA allows states to administer their own UIC programs if the state regulations and statutes meet the EPA's minimum requirements specified in 40 C.F.R. Parts 144–146. See http://www.epa.gov/safewater/states.html.

The overriding standard applicable to all underground injection activities is that

> No owner or operator shall construct, operate, maintain, convert, plug, abandon, or conduct any other injection activity in a manner that allows the movement of fluid containing any contaminant into underground sources of drinking water, if the presence of that contaminant may cause a violation of any primary drinking water regulation under 40 C.F.R. Part 142 or may otherwise adversely affect the health of persons . . . [40 C.F.R. § 144.12(a)].

Underground injection is defined as well injection, or the subsurface emplacement of fluids through a bored, drilled or driven well, or through a dug well where the depth of the dug well is greater than the largest surface dimension" (40 C.F.R. § 144.3). Fluid is broadly defined as "any material or substance which flows or moves whether in a semisolid, liquid, sludge, gas, or any other former state" (Id).

EPA has defined five classes of underground injection systems:

Class I: Wells injecting hazardous or nonhazardous waste below the lower most underground source of drinking water

Class II: Wells injecting fluids associated with oil and gas production

Class III: Wells injecting fluids for mineral extraction

Class IV: Wells injecting hazardous waste into or above an underground source of drinking water

Class V: All injection wells not covered by Classes I–IV (40 C.F.R. § 144.6).

All Class IV wells are banned by SDWA (40 C.F.R. § 144.13). Class V wells are "authorized by rule," and no individual permit is required. There are currently few regulatory requirements for Class V wells other than reporting under 40 C.F.R. § 144.26. Class V wells are the systems most commonly used by industry and small commercial operations. Common Class V wells include air conditioning return flow wells, cesspools, and cooling water return flow wells. See EPA's Class V websites at http://www.epa.gov.sefewater/uic/classv.html and http://www.epa.gov/safewater/uic/c5fin-fs.html. The EPA has announced that it is discontinuing its permission to operate Class V wells for motor vehicle waste disposal and large capacity cesspools. [64 Fed.Reg. 68545 (Dec. 7, 1999)].

Certain Class I, II, and III wells are not required to obtain individual permits (40 C.F.R. § 144.21). These wells may be authorized by rule if the operator injects within 1 year of a new UIC rule or performs an inventory under 40 C.F.R. § 144.26 within 1 year. All other Class I, II, and III wells require an UIC permit. An application for a UIC permit for a new injection well must be submitted by a "reasonable time before construction is expected to begin" [40 C.F.R. § 144.31(c)(2)]. If a facility or activity is owned by one person but is operated by another person, it is the operator's duty to obtain an UIC permit [40 C.F.R. § 144.31(b)]. Applicants for a UIC permit must use the application form provided by the EPA [40 C.F.R. § 144.31(e)]. The EPA may issue a permit for an area, rather than for each well individually, if the wells are within the same well field, facility site, reservoir, project, or similar unit in the same state; the wells are operated by a single owner or operator; and if the wells are used to inject something other than hazardous waste [40 C.F.R. § 144.33(a)].

If cause exists, the Administrator may modify or revoke and reissue a UIC permit. The Administrator may terminate a UIC permit during its term, or deny a permit renewal application if there is noncompliance by the permittee, failure to disclose relevant facts, or the permitted activity endangers human health or the environment (40 CFR § 144.40). A detailed description of the UIC program is available at http://www.epa.gov/safewater/uic.html.

OTHER SDWA PROVISIONS

The SDWA gives the EPA Administrator the power to classify an aquifer as a "sole source aquifer" (SSA) if the aquifer is the sole source of drinking water for a community. The SSA designation gives the EPA greater ability to scrutinize federal projects in the designated area to ensure that the aquifer is not contaminated [42 U.S.C. § 300h-3(e)]. The SSA regulations are located at 40 C.F.R. Part 149. The SDWA also requires that states with primacy to implement a Source Water Assessment Program to protect source waters for PWSs better from contamination (42 U.S.C. § 300j-13). SDWA also provides federal funding under the Drinking Water State Revolving Loan Fund to states for drinking water infrastructure projects (42 U.S.C. § 300j-12).

REPRESENTING GEOPOLITICS OF (HYDRO) BORDERS IN SOUTH ASIA

Eva Saroch
Panjab University
Chandigarh, India

INTRODUCTION

In the tradition of political geography, the concept of boundary/border has been of major importance where researchers have understood it mainly as an empirical manifestation of political processes and particularly of the sovereignty of the state. But borders not only represent a 'fixed line' in space or time, rather, they need to be understood as being produced and reproduced in the *socio-spatial consciousness* of society (a broader form of consciousness that reflects the ideological and hegemonic structures). Therefore, it is argued here that boundaries, apart from being located on terra-firma as a physical line, also exist metaphorically, symbolically in many ways as a manifestation of socio-spatial consciousness and as 'plays of power' that generate and exert various codes, rules and procedures of 'distinction' in the process of Othering such as us/them, sacred/profane, inclusive/exclusive, to impose specific understanding through the construction of knowledge in sociocultural practices and discourses in politics, administration, economics, or the organization of ethnic relations on various scales.

In the context of traditional geopolitics of water, 'mastering space' by deploying strategies of spatial domination and control and acquiring and guarding access to freshwater resources has always been a highly contested geopolitical issue both among and within sovereign states. A critical geopolitical approach, focusing the intricate but intimate interplay between space and its natural endowments such as water, and the socio-spatial consciousness of its various stakeholders, seems to suggest that rivers, for example, are not simply 'located' physically on ground as water flows or as static lines on maps but are situated/implicated in the very process of 'Othering' or the construction of symbolic/cultural boundaries between 'us' and 'them' (see Campbell, 1992, in Ref. 1, p. 3). Attention is thus drawn to how growing disagreements and rivalries over who gets, when, how and how much transforms seemingly innocent water flows into highly resilient hydro-borders. A critical geopolitical perspective motivates us to look for the sites where hydroborders are produced and reproduced, both in the form of various discourses (administrative, political, economic, cultural, and religious) and practices on various—spatial scales local, provincial, and national. At yet another analytical level, it is worth exploring how the 'criticality' of water resources is inscribed on the mental maps of various stakeholders (at local, national, and regional levels) as symbols of socio-cultural–religious identification or differentiation, identity formation, and power-political pursuit of primacy and security.

Even though the states and societies of South Asia are hydrographically united—they share the Indus River Basin (India and Pakistan), the Ganges–Brahmaputra–Meghna Basin (India and Bangladesh), and the Ganges–Mahakali Basin (India and Nepal), which the brief profile of the South Asian waterways suggests—geopolitically they stand divided. Hydro-borders in the form of mental borders or barriers continue to appear and reappear in different guises within an existing asymmetry of wealth, knowledge, and information, and as part of conflict and struggle among various actors for spatial control or mastering space (2, p. 115).

Thus, it is argued here that in the context of traditional, state-centric geopolitics of water resources, which seems to persist in post-colonial, post-partition South Asia, the 'hydroborders' in South Asia continue to be constructed, reinforced, and legitimized through competing notions of domestic stability/instability, security/insecurity, and legitimacy/illegitimacy. Constructed through 'nationalist' metaphors, state-sponsored security narratives and even politically motivated religious speeches, these mental borders, deceptively hydrographic in argument as well as in appearance, remain integral to the politics of place-making and group differentiation in the region. Drawing upon insights from the ongoing tensions/conflicts between India–Pakistan, India–Bangladesh, and India–Nepal (as well as within these countries) over the so-called 'sharing' of water resources, this article critically examines the interplay between the boundary-defying/unifying nature of waterways and the bounding/excluding logic of geopolitical discourses, imaginations, and representations.

A PREFACE TO SOUTH ASIAN WATERWAYS

Ecologically speaking, a river knows no boundaries. Flowing through farms, cities, and nations, rivers join rather than divide places. Any river that flows through two or more nations is supported by an ecosystem that cuts across political boundaries and joins the basin states. In similar perspective, South Asian waterways are the epitomes of sociocultural synergies, binding the South Asian landscape into one composite whole. The ancient 'Sapt Sindu Valley' and the 'Gangetic Basin,' together known as the Indo–Gangetic Basin, encompass South Asian waterways. The River Indus and the River Ganga are the principal rivers of two separate systems, but their separation is overridden by the unity of the Indo–Gangetic Plain, which spreads between the Himalayas and the Deccan. The Great Plains consist of two river basins, those of the 'River Indus' and the 'Ganga–Brahmaputra,' and the network of these rivers constitutes the 'Himalayan river system,' which has not only maintained the geographical contiguity on the subcontinent but has also been a fulcrum of different cultural streams. Having sustained inhabitants, the aquatic communities, and the ecosystem over centuries, these vast river systems have been the unifying–binding force for the South Asian geo-economic and geo-cultural landscape.

PARTITION, OTHERNESS, AND HYDRO-BORDERS

A critical geopolitical approach to the study of partition suggests that the partition of 'British India' was the convergence of two different phenomena—ethno-politics and geopolitics. According to Denis Retaille, ethnopolitics takes identity not territory as its starting point (3, p. 34), whereas geopolitics is a question of rivalries between powers for land and for those who live there or rivalries for territory (4, p. 122). Each complements the other in producing the other's 'excess' or excessive geopolitics. Excessive geopolitics "What", as aptly put forth by Chaturvedi, "does" is transform ideological–ethnopolitical conflicts relating to the boundaries of identity groups into geopolitical rivalries over 'our' territory versus 'their' territory. Consequently, the cold-blooded logic of 'territory for the sake of territory,' or Lebensraum, subsumes the more humane logics of common habitat or shared spaces" (5, p. 150). What sustains excessive geopolitics is the 'reflexive Otherness.' The image of the 'Other' is significant in producing borders that are both physical and symbolic. The permanent presence of 'Other' in discourse over water conflicts in South Asia has produced a manifold character of hydro-borders that are manifest in various foreign policy pronouncements, in defining national identity, and in so-called development projects. Even the language of dispute or the dispute over language suggests the deadly hold of 'nationalist language' not only in conceptualizing or thematizing the problem but also strategizing the resolution of the dispute.

Sir Cyril Radcliffe, a British lawyer who was given just six weeks to draw boundaries in Punjab in the west and Bengal in the east, was aware that the network of canals, River Indus, and its five tributaries were a lifeline for united Punjab and also the continued integration of the irrigation system, at least on a functional basis, was fundamental for the basin as a whole. Such a joint venture run by both countries would not only safeguard the interests of both peoples but also form a basis of cooperation which might prove fruitful in the years to come (as cited in Ref. 6, p. 164). But Radcliffe's hope of joint management of the basin for the optimal benefit of the people and ecology was rebuffed by both Jinnah and Nehru. According to Jinnah, "he would rather have Pakistan deserts than fertile fields watered by courtesy of Hindus." Nehru curtly informed that "what India did with the Indian rivers was India's affair" (6). This is a classic example that shows how a common past is repudiated as a time of oppression in which 'we' were simply the victims of oppressive 'others.' It also shows how partition has always been a disciplining event that shuts out other possibilities of resolving the conundrum whose one form is partitioning country, homelands, or the network of rivers, the arteries of the basin.

Hence, the boundary line in Punjab was drawn between Lahore and Amritsar, cutting across the well-integrated canal system that draws its water from the Indus and its five tributaries, Jehlum, Chenab, Ravi, Sutluj, and Beas. The enforced geography of partition was such that the headwaters of the rivers of the Indus Basin remained in India, thereby giving India the status of upper riparian. Pakistan felt that its livelihood was threatened by the prospect of unilateral Indian control over the tributaries that fed water into the Pakistani portion of the basin. Thus, Pakistan apprehended that a 'hostile' India might, if it so desired, using tunnels, divert water within Kashmir from the Western to the Eastern rivers, affecting a transfer before the water even reached Pakistan or the fringes of 'Azad Kashmir' (which India chooses to call Pakistan-occupied Kashmir) (7, p. 234). Whereas India certainly had its own ambitions for the 'profitable development' of the basin, Pakistani media and the politicians launched a large-scale national campaign over a canal water dispute. Pakistan even used the Indus water dispute as a political tool in the battle over Kashmir being debated at the United Nations. Former Foreign Secretary of India, J.N. Dixit, while addressing a plenary session in a seminar, is reported to have said that, "it was well accepted in our strategic studies circles, that much of the passion on Kashmir was generated because India controls the headwaters of the Indus" (8, p. 161).

A greater insight into the deliberations of the pre-Indus Waters Treaty period reveals that the Treaty signed between India and Pakistan in 1960 was dictated and driven by the logic of 'Othering' that jettisoned and abandoned the more humane logic and vision of joint development and administration of the integrated Indus Basin for the optimal benefit of the people of the two countries. The Treaty led to the fissuring/partitioning of the Indus and its five tributaries (the three eastern rivers went to India and the three western rivers to Pakistan) and of the 'homelands' that dwelled in the canal colonies, but the Treaty also carried, within itself, seeds of bypassed, unresolved, old, and unanticipated new water conflicts in the basin.

Even after partition became an accomplished reality, dialogues to overcome the bitter divides created in the wake of partition became prey to the partitioned times. Nothing can be more illustrative in this context than the water disputes at provincial and local levels. There are intrastate water disputes, for example, between Punjab and Sindh (Pakistan), and Punjab, Haryana, and Rajasthan (India), among others. Within India, in the wake of the partition, the issue that agitated the Sikhs even more than the territorial dispute was the unjust and illegal allocation of Punjab's river waters and hydroelectric power generated from them to non-riparian neighbor states. More recently, Mr. Praksah Singh Badal, former Chief Minister of Punjab in India, reportedly said that every drop of water that flows through Punjab belongs to Punjab. In Pakistan, the water conflicts between Punjab and Sindh have gained momentum. According to Sindh, the upper riparian province has been progressively increasing its illegal share of the waters of the Indus River since 1889 (9). Apart from this, in 1960, the military regime in Pakistan surrendered, under the Indus Waters Treaty, the three eastern rivers to India along with Sindh's acknowledged share without prior consent or compensation to Sindh, leading to water scarcity in the province.

It is possible to argue that what lies at the very core of the matter is the entrenched notion of 'us' versus 'them', as well as a nagging sense of insecurity on the part of Islamabad that a 'permanently hostile' India, which also happens to be the upper riparian state, might use water resources as a politico-strategic weapon to paralyze Pakistan's economy completely. Such imaginations and suspicions continue to mar India–Pakistani relations, and they also keep alive the bitter memories of the bloody historical circumstances, in which the two 'Midnight Children' were born. Needless to say, the persistence of such hydro-borders makes any solutions of relatively fewer complex issues such as the Tulbul Navigation Project/Wular Barrage virtually impossible.

Similarly, the Ganges water dispute between India and Pakistan (and after 1971 between India and Bangladesh) also had its genesis in partition. The Farakka dispute dates back to 1951 when India decided to construct a barrage across the Ganges at Farakka to divert water to Bhagirathi to maintain its navigability, which was being hampered due to siltation. However, a deeper analysis suggests that Farakka is not only a technical problem involving the issue of water sharing but, as put forth by Begum Khurshida, a 'techno-political tangle' (10, p. 196). Technical debates have intermingled with political rhetoric, narratives, and nationalist myths; Farakka is not only a symbol of Indian domination (10, p. 202) but also projects an idiosyncratic 'upper riparian psyche' that intimidates lower riparian Bangladesh and jeopardizes its water security. The entire history of the dispute and the tortuous course of negotiations over Farakka suggests that what was being negotiated was not the flowing waters, but the strategic resource for nationalist politics between India and Bangladesh.

The Mahakali Treaty of 1996, signed between India and Nepal (concerning the integrated development of the Mahakali river), is yet another example where the ethics of water sharing could not bypass the deeply entrenched issues of sovereignty and (in)security or for that matter ideological contestations. In Nepal, the treaty is considered another sellout of water resources. The treaty is said to be 'the blatant manifestation of domination of Indian expansionism, on one hand, and total surrender to the Indian big brother on the other.' Thus, so far, the dominant trend among the two nation-states has been to perceive water issues both as a major problem area and as a valuable resource to enhance their respective 'national' power and defense (8, p. 167). In the process, the water issues have been (mis)used to sustain conventional geopolitical considerations.

It can also be argued that behind these self-locked positions are the deeply seated hydro-borders, created through various kind of geopolitical perceptions, metaphors/rhetorics/narratives by various protagonists both within and among the sovereign states involved at various levels in the dispute over water issues. As a consequence, water has become a crucial item of resource politics, nationalism, provincialism, and sub-regionalism (10, p. 196). All such linkages have shaped and are central to the elite politics of water and the so-called 'sustainable development' of South Asian waterways.

MAPPING OF RIVERS IN THE 'CARTOGRAPHIC ANXIETY' OF NATION

South Asian states continue to suffer from, what Sankaran Krishna (11, p. 82) calls, 'cartographic anxiety' surrounding the questions of national identity and survival. As meaning is invested in all aspects of cartography, so defined cartography becomes nothing less than the 'social and political production of nationality' itself. And these cartographic representations not only become alive during the election campaign, in the everyday life along the 'border' (see 12), but interestingly enough, different imaginations/mapping of the river that produce hydro-borders are invoked in the construction/representation of 'imagined landscape' or 'national identity' or securing 'territorial unity.'

Pakistan, argues Ian Talbot, has been a state searching for its national identity, and due to persistent failure to find such an identity, ethnic and tribal identities remain strong. Even five decades after its creation, the question remains whether Pakistan is a land for the Muslims or a nation of Muslims moving towards its destiny as a nation? (as cited in Ref. 12, p. 153). In an attempt to define 'national identity,' sometimes 'pasts' are invoked and 'histories are constructed' (12). For instance, in a major restatement (rather 'rewriting of history of the Indian subcontinent') Pakistani historian Aitzaz Ahsan, argues in his work *The Indus Saga and the Making of Pakistan,* that the Indus region (Pakistan) has a primordial existence outside India and centrifugal forces of its own (13). The essential purpose of this work is, in the words of its author, "to discover and define 'Indus person,' the Pakistani citizen." According to Ahsan, history states that Pakistan is said to have existed "for almost five and half thousand of the last six thousand years" and emphasizes that the "Indus has seldom been part of India."

Subsequently, his work contends that the Indian subcontinent has always been at least two different worlds. The divide between India and Indus is civilizational and cultural. There has existed a 'primordial divide' between the two: 'a palpable divide between two lands, two peoples and two civilisations: Indus and India. In fact the twain has seldom, if ever, truly met. Arguing that these differences are deep rooted, primordial, and many, he observes that

> The sub-continent is made up of two civilisations, Indus and Indic or Gangetic and that Indus has been one large, independent, politico-economic zone for the past countless centuries. The truly 'Indic region' comprising the Gangetic plain and peninsular India, on the one hand, and the Indus basin, consisting of Indus river and its tributaries (i.e., Pakistan) on the other. Further, more than the giant mountain range that separate the Indus region from the Central Asia, that discernible hump between Indus and India along a line roughly connecting Gurdaspur, in Eastern Punjab, to Kaithawar on the Arabian sea, has been the critical dividing line of history and culture. It approximates the division between India and Pakistan, thus giving to that border the sanction and strength of history (13, p. 289).

This kind of rewriting of history is an attempt to evade the conspicuous hydrographic oneness/unity of the Indus–Gangetic basin. By ingeniously selecting the facts that serve his 'partitioning and Othering' analyses and fitting his mapping of national identity based on the re-presentation and apt distortion of the 'Indus Basin' hydrographically, historically, culturally, and civilizationally, he is able to suitably construct a 'Indus Person' or Pakistani identity. Further, in reconstructing a narrative of a distinct cultural heritage, Ahsan has traced the civilizational roots of a nation-state that came into being only in 1947, into the distant past (14, p. 19). According to David Ludden, such an attempt which constructs beleaguered historical narratives around partition and sought to inscribe the past with a new meaning, is a feature of the politics of the cultural anxieties that have surfaced lately across contemporary South Asia (14).

In the traditional sense, river has been used as a boundary to define the territorial limits of a state, but to engage in rituals to purify 'frontier rivers' is not only to revive, the sacredness of the river but also to deploy new meaning(s) to the process of territorialization. The two-day Sindhu Darshan festival (the brainchild of Home Minister L.K. Advani), in Ladakh (on the Indian side of Kashmir), where the river Sindhu or Sindh flows before entering Pakistan territory, was at one level a welcome gesture toward integrating far-flung parts of the country into the mainland, whereas, on the other hand, a symbolic attempt to revive the link between 'Sindhu and Hindu', or a ceremony affirming the sacredness of the River Indus. What has followed 'Sindhu Darshan' is the so-called 'Brahmaputra Darshan,' a similar kind of festival in the northeast part of India, 50 km from the IndoChina border (15).

It is important to note that in imagination about India's national identity, there has always been a conceptual space for Hindu nationalism. Further, Hindu nationalism has two simultaneous impulses: a commitment to the territorial integrity of India as well as a political commitment to Hinduism. And in the geopolitical imagination of Hindu nationalists, 'national unity' and 'identity' are based on territory and religion (16, p. 235). Hence, it is argued that such festivals/practices, that are held at the banks of frontier rivers, are an attempt to engage in the process of sacralization and re-territorialization of the 'frontiers' of the nation-state and also to promote and consolidate the notion of 'Akhand Bharat' (undivided, greater India). These examples also reveal how different geopolitical imaginations and representations of the river have provided meaning(s) and a rationale to different cartographic anxieties, a facet of a larger post-colonial 'anxiety syndrome' or as Samaddar says, 'insecurity syndrome' (17, p. 32), surrounding the question of territorial unity and national identity. Needless to say, through representation of a river in a certain way, the state generates a representation of self and in the process the boundary-defying river also becomes a boundary-producing entity.

CONCLUSION

The ethics of sharing water resources posits that water, as a common natural heritage—especially transboundary watercourses, including aquifers—must be managed in an ecologically sustainable manner. But this ethical imperative is undermined as and when domestic geo-political rivalries, conflicting national ideologies, and geopolitical discourses conspire together to produce hydro-borders that are deeply entrenched and frequently (mis)used by the states as instruments or media of social distinction, domination, or control (18, p. 65). In their quest for domestic stability, political legitimacy, regime survival, and identity building, in both religious-communal and secular-national terms, South Asian states continue to articulate and defend a 'national vision,' that has little space to accommodate the concerns of other state or non-state actors. The process of Otherization that continues to make its place, that is, production and re-production of hydroborders in the bilateral or multilateral dialogue between India/Pakistan/or India/Nepal/Bangladesh over the water disputes does not reflect a dialogue of accommodation/agreement or cooperation but a dialogue of disagreement, separation, and partition itself. This also suggests that a new kind of resource geopolitics, in which different power structures are sustained through the discourse and practices, is in the making in South Asia. As long as the logic of Othering, as manifested in the form of hydroborders, continues to make its place in the South Asian geopolitics of water, the prospects for an ecologically sustainable, gender-sensitive, culturally appropriate, and socially just development and management of South Asian waterways is difficult to achieve. The prospects for peace in South Asian region also require that the construction and reconstruction of hydroborders be undermined new, normative and humane geopolitics of sharing and sustaining a common habitat in place of state-centric, power-political geopolitics of mastering space and resources.

BIBLIOGRAPHY

1. Paasi, A. (2000). *Space, Boundaries and The Territorial Construction of Identities*. Presented at the BRIT-IV, Conference *Geopolitics, Identities and Sustainability*, Feburary 20–26, Chandigarh.
2. Ahmed, I., Dixit, A., and Nandy, A. (1999). Water, power and people: A South Asian manifesto on the politics and knowledge of water. *Water Nepal* **7**: 113–121.
3. Retaille, D. (2000). Geopolitics in history. *Geopolitics* **5**(2), Autumn: 35–51.
4. Lacoste, Y. (2000). Rivalries for territories. *Geopolitics* **5**(2), Autumn: 120–158.
5. Chaturvedi, S. (2003). Towards a critical geography of partition(s): Some reflections on and from South Asia. A guest editorial in *Environ. Plann.: Soc. Space* **21**(2): 148–154.
6. Michel, A.A. (1967). *The Indus Rivers: A Study of the Effects of Partition*. Yale University Press, New Haven, CT.
7. Chapman, G.P. (2000). *The Geopolitics of South Asia: From Early Empires to India, Pakistan and Bangladesh*. Ashgate, Aldershot.
8. Gyawali, D. (2001). Water accords in South Asia: Pluralist politics under monistic design. In: *Peace as a Process: Reconciliation and Conflict Resolution in South Asia*. R. Samaddar and H. Reifeld (Eds.). Manohar, New Delhi.
9. Tarren, I. (2001). Sindhi Association of North America: Letters and petition. March 21. Available: http://www.sanalist.org/letters.htm.
10. Samaddar, R. (1997). Flowing waters and the nationalist metaphors. *Stud. Conflict Terrorism* **20**: 195–206.
11. Krishna, S. (1997). Cartographic anxiety: Mapping the body politic in India. In: *Political Geography: A Reader*. J. Agnew (Ed.). Arnold, London.
12. Chaturvedi, S. (2002). Process of othering in case of India and Pakistan. *J. Econ. Soc. Geogr.* **93**(2): 149–159.
13. Ahsan, A. (1996). *The Indus Saga and the Making of Pakistan*. Oxford University Press, Karachi.
14. Tan, T.Y. and Kudaisya, G. (2000). *The Aftermath of Partition in South Asia*. Routledge, London.
15. The India Express. (2001). After Sindhu, it's Brahmaputra Darshan in Arunachal. February 18.
16. Varshney, A. (1993). Contested meanings: India's national identity, Hindu nationalism, and politics of anxiety. *Daedelus* **122**(June): 227–261.
17. Samaddar, R. (2000–2001). The last hurrah that continues. *Transeuropeannes* **19/20**: 31. 48.
18. Chaturvedi, S. (2001). Mental borders and the prospects of peace in South Asia. In: *Peace as a Process: Reconciliation and Conflict Resolution in South Asia*. R. Samaddar and H. Reifeld (Eds.). Manohar, New Delhi, pp. 61–80.

WATER TRANSFERS

David L. Feldman
University of Tennessee
Knoxville, Tennessee

Water transfers are the diversion, reallocation, or appropriation of surface or groundwater from one watershed, drainage basin, or aquifer to another. All transfers entail water exchange between catchments. They may also move water from one *political jurisdiction* to another, exacerbating their impacts and complicating efforts to resolve stakeholder conflicts that they generate. In addition, transfers may be endeavored through various *means* including pipelines, tunnels, aqueducts, pumping facilities, or even seaborne tankers—an idea contemplated under proposals to transfer water from the Great Lakes to the Middle East.

MOTIVES AND MAJOR ISSUES

Water transfers are nearly as old as civilization. The first recorded large-scale transfers occurred in the ancient Near East. About 4000 years ago, the Sumerians discovered how to divert the Euphrates River (in what is present-day Iraq) to irrigate crops and reduce reliance on the vagaries of seasonal rainfall. Other nearby societies followed suit and built administrative infrastructures to support transfers, leading to speculation that the need to maintain such infrastructures contributed to the growth of political authoritarianism (1,2).

One motive for transfers is to provide sufficient water to satisfy agricultural, industrial, or other needs to regions perceived to have inadequate water supplies. A secondary reason is to promote regional development. This was a major motive for building diversion projects in the southwest United States in the nineteenth century, for example. For the exporting region, the benefits of transfers include economic compensation and better relations with neighbors. However, exporting regions are not always willing partners in transfers, in part because they may have unintended environmental consequences or may be undertaken without adequate consultation (3). Regardless of motive, *five* major issues tend to be associated with transfers:

- *Adequacy of Return Flow.* Most water transfers involve some return flow to the basin of origin. Issues raised by this return flow include how much and when water will be returned. If the return flow is relatively prompt and abundant, harm to ecological resources from low river levels, changes in hydrology, and/or losses to the local economy due to foregone benefits are likely to be less than if return flows are small in volume and/or late.
- *Loss of Water Rights.* Transfers impose impacts on those who have access to water through legal guarantee. These may be comparable from place-to-place, but remedies available under common law differ among regions. In states that have prior appropriation systems, there has tended to be less recourse to deterring transfers thought to be harmful than in states that have riparian rights systems—although water rights issues arise in *both* systems.
- *Adverse Ecological Impacts.* In general, water transfers generate two distinct forms of ecological impact: to in-stream flow (the volume, flow, and quality of water naturally found in rivers and streams prior to the building of impoundments or diversion projects) and biota transfer (the consignment, usually unintended, of exotic fish or other aquatic species from one basin to another, that adversely affect native species' welfare).
- *Water Agency Objectives.* Agencies that manage natural or cultural resources are affected by water transfer proposals in several ways. Often, transfer decisions are undertaken by an agency to meet obligations to promote regional development through water delivery (e.g., the Bureau of Reclamation). Conversely, an agency charged with protecting natural/cultural areas (e.g., National Park Service) may be reluctant to permit transfers.
- *Adverse Social Impacts.* Although difficult to measure, transfers generate significant, often adverse social impacts comparable to those engendered by other resource allocation decisions involving unequal power and authority, representation, and resource access. Specific impacts include changes in lifestyle, reductions in regional income, fewer opportunities for economic development, and abridgment of recreational amenities.

ADEQUATE RETURN FLOW

Water transfers, especially if large, can have profound effects on the quality of water in the basin from whence water is taken, *and* on the quality of water in the receiving basin. Here, we consider the first of these. Many water transfers involve what amount to temporary shifts of water between basins. Used, often untreated water is rerouted back to the basin of origin. Such return flows are often negotiated in agreements to contract transfers and, in riparian rights systems, may be *required* to ensure that natural stream flow is unimpeded (in general, this is one reason why riparian rights systems discourage transfers).

When transfers are permitted in riparian rights systems, return flows are supposed to occur in the vicinity of the same riparian landholding from which water was originally diverted, and its quality is supposed to be *unimpaired*. Historically, such return flows are rarely monitored in many states, do not occur with regularity, are usually easier to ensure when the transfer is between two adjoining watersheds *in* the same state, and are difficult to measure precisely when they occur. Moreover, if they do not occur during periods of low flow, when most needed, the ecological resources of the exporting basin may be harmed.

WATER LAW AND WATER TRANSFERS

All water law systems predicate certain principles regarding the judicious transfer of water from one basin to another. At the outset, two characteristics of American water law, significant for water transfers, must be noted. First, far more attention has been paid to regulating *surface*, as opposed to *groundwater,* transfers because it is difficult to determine exactly what constitutes a groundwater transfer. Underground aquifers are connected to one another and to surface waters in enigmatic ways often hard to prove in courts of law. Negotiating allocation formulas for surface waters usually entails measuring the amount of water to be apportioned at gauging stations at different seasons of the year and, then, limiting how much water an "upper" state or user can draw, *or* guaranteeing a certain amount of water to the "lower" state or user. For aquifers, however, there is no upper and lower user in the same sense as for surface waters (5).

Second, states may not discriminate in the way they manage out-of-state, as opposed to in-state, transfers. Article 1, Section 9 of the Constitution prohibits states from giving preference, in their regulation of enterprises, to businesses operating within their borders over those based in other states. Originally applied to regulation of ports and vessels traversing interstate waters, the clause has been construed by federal courts to mean that Congress has broad plenary power to regulate interstate commerce, manage and use water for navigation, and regulate nonnavigable streams. It has also been construed to mean that water is a commodity in interstate commerce and states cannot discriminate against one another's ability to engage in such commerce. This is an important consideration in efforts by states to regulate interbasin transfers (see "Conclusions").

Riparian Doctrine and Water Transfers

Under riparian law, the predominant legal system in the eastern United States, a landowner has the right to "reasonable" use of the water flowing past her/his property, subject to the rights of other downstream riparian landowners. Water used but not consumed must be returned to the watercourse unimpaired inquality. In general, diverting water from the drainage area of a watercourse *without* return flow is deemed "unreasonable" and, if a downstream riparian complains, can be expressly forbidden (6,7).

In practice, courts have ruled that "reasonable" diversions are allowable under riparian rights systems and typically define such transfers as those that are socially beneficial and impose no harm. Included under this definition are transfers undertaken for the use and benefit of public supply systems. In practice, definitions of "reasonableness" vary from case-to-case and court-to-court, and riparian doctrine and court precedents are unclear on the conditions for mitigating their adverse impacts. How to ensure adequately timed return flow is also uncertain (8,9). Moreover, a "reasonable use" transfer right can be superseded if a subsequent user can demonstrate a need that is more reasonable (10).

Transfers Under Prior Appropriation Doctrine

Appropriation law, the predominant water rights doctrine in 19 western and southwestern states lying west of the 100th meridian, quantifies and prioritizes water rights according to the "first in time, first in right" principle: the first to put water to a beneficial and reasonable use acquires a right superior to later appropriators. This principle grants use of a specified amount of water for a stated beneficial use each year, subject only to the rights of other appropriators (11,12). Land ownership is not the basis for this right: one can acquire it through purchase or other means. Moreover, water rights exist only when water is appropriated (i.e., diverted and used) for a *beneficial* purpose. Historically, this encouraged transfers.

Traditionally, off-stream uses (e.g., irrigation) were considered beneficial, whereas in-stream uses by fish and wildlife were treated as inherently wasteful (7). Senior appropriators could take their full allotment even if insufficient water remained for junior appropriators' needs. This practice served to bias the concept of "beneficial" to mean off-stream, consumptive uses. More important, appropriated waters could be used virtually anywhere, regardless of the distance from a stream. Most appropriation states now define beneficial use by statute. Preferred uses are explicitly defined, and reforms intended to reduce environmental damage through unregulated transfers, such as prescription of minimum stream flows, have been adopted (13,14).

ECOLOGICAL IMPACTS OF TRANSFERS

Dams, impoundments, pipelines, and other "hydrologic alterations" that enable water transfers may lead to severe changes in in-stream flow volumes, velocity, and quality in both exporting *and* destination watersheds, as well as fragmentation and loss of aquatic and riparian habitat (15,16). An important, but often overlooked impact is *biota transfer*, which can also result in both basins of origin and destination.

The Garrison Diversion Unit (GDU), a large transfer project in the Upper Missouri basin authorized by Congress in 1965, exemplifies these problems. The project has been criticized by Canadians due to concerns that return flows from its irrigation subunits could have negative impacts on fisheries and flood stage levels in the Red River flowing north into Manitoba. Canada claims that Missouri River biota would be transferred into Canadian waters in violation of the 1909 *Boundary Waters Treaty* designed to prevent contamination, diversion, or degradation of the transboundary waters shared by both nations. The project's size and scope have been significantly curtailed since the 1990s, but the biota transfer issue, especially, exemplifies two general problems of water transfers applicable elsewhere: (1) the benefits and costs of transfers may differ across affected political jurisdictions, and (2) uncertainty regarding impacts may be exacerbated by international disagreement. Canadian scientists were much more adamant about the risks of biota transfer than U.S. researchers whose work was sponsored by the Bureau of Reclamation (3).

AGENCY MANAGEMENT IMPACTS

Proposals for transferring water from one basin to another may challenge the responsibilities of an agency charged with protecting environmental quality or cultural amenities (e.g., historically significant sites) located within lands embraced by its mandate. By contrast, the provision of water supply may be viewed as so central to the mandated responsibility of another resource-development agency as to justify transfers and even override other considerations.

An example of how these divergent views may play out is offered by the experience of the National Park Service (NPS) and Bureau of Reclamation (BOR), agencies which have, on occasion, seen their interests conflict over water transfers. Under the National Park Service Act of 1916, NPS is responsible for protecting the natural, unimpaired flow of rivers flowing through its parks and other protected areas. By contrast, since 1902, the BOR has been responsible for developing irrigation and other water development projects in the West.

Generally, the NPS tries to balance its claims to needed in-stream flow against the reasonable desire of states (both eastern and western) to protect appropriative and riparian rights. To achieve this balance, the NPS asserts rights “for quantities determined to be the minimum amounts needed to protect the primary purposes” of a given park (12). According to one Interior Department solicitor, “appropriation for authorized federal purposes cannot be strictly limited by what state water law says is a ‘diversion’ of water or a ‘beneficial use’ for which water can be appropriated” (17).

In practice, when agencies cannot agree on conditions for transfers or when the economic and political stakes are high, Congress may intervene as it did in 1947 in requiring the NPS to allow the BOR to construct a water diversion project (the Colorado-Big Thompson Project) through a portion of Rocky Mountain National Park. This project’s purpose was transferring water from the uppermost reaches of the Colorado River basin, which rises on the state’s western slope, to the Platte River Basin northeast of Denver. The project’s water storage capacity exceeds 710,000 acre-feet, and it now irrigates 1,000,000 acres (18). To compensate for potential harm, storage reservoirs were constructed for the project, in part, to ensure adequate in-stream flow to Upper Colorado River users and to permit reliable transfers to the eastern slope.

ADVERSE SOCIAL IMPACTS

All water transfers generate social impacts. In the western United States, where water marketing has gained widespread acceptance as a means of diverting water from low- to high-value uses, these impacts can be considerable. Studies show that water marketing in conjunction with transfer can be unpopular, especially in an importing region. In the San Joaquin Valley of California and the Grand Valley of western Colorado, two potential water-exporting areas, residents it was found, were more likely to oppose water transfers than residents of water-importing areas. Concerns included potential impacts to residents’ economic welfare and way of life and whether allocations between exporting and importing regions provide for reciprocal benefits such as protecting future needs and preventing adverse environmental impacts (19). Generally, agricultural interests are more likely than other groups to approve of water transfers and to favor marketing as a tool for enabling such transfers to occur.

Despite such concerns, other studies suggest that water transfers may serve as an effective means of diverting water from low- to higher valued needs and even as a means of *restoring* in-stream flow (20). In parts of the American West, for example, including California, Montana, Nevada, Oregon, Texas, and Washington, special markets negotiated by environmental groups have had some success in reducing volumes of water diverted off-stream for irrigation and maintaining stream flows for fish and other environmental and recreational benefits (20: 14). In at least one instance, governmental entities have actually purchased the rights to in-stream water, establishing, in effect, a “reverse” transfer (21: pp. 5–11).

Potential barriers to this practice include the lack of a physical infrastructure for moving water from place-to-place; ambiguities in water law permitting markets, especially under riparian systems; and ‘aligning’ the price of water between buyer and seller. An underlying investment assumption in such deals is that, over time, transfer of water for restoring in-stream flow will yield a greater benefit than diversion for off-stream uses (22).

CONCLUSIONS—WATER TRANSFER DISPUTES AND THEIR IMPLICATIONS

Since the 1980s, disputes over water supply and allocation have arisen in many parts of the southeast United States, underscoring the fact that *no* region is immune from conflicts over competing demands for water. Sources of these disputes include population growth, urbanization, and competing demands, especially for agricultural irrigation, power, and public uses, as well as debate over how to balance off-stream and in-stream needs (23). One result of these disputes has been plans, real or contemplated, to transfer water from one basin to another.

One of the most contentious of such plans began in 1983 when the city of Virginia Beach, Virginia, indicated its intent to divert water from the Roanoke River to satisfy its growing needs. This led North Carolina to protest the project’s potentially adverse impacts on ecological resources and the local economy of that part of the Roanoke Basin in the latter. After a nearly 15-year battle costing upward of $12 million, the project was approved after court challenges failed to demonstrate convincingly adverse effects from plans to divert approximately 60 MGD. A continuing legacy of this dispute, however, is the precedent which it may have established for *additional* diversions, as well as *how* to measure adverse impacts: as present-day losses or as potential future ones (24,25)?

Recently, conflict between Alabama, Florida, and Georgia over water allocation in the Apalachicola–Chattahoochee–Flint basin led some Atlanta-area officials—contemplating an eventual ceiling on withdrawals from a tri-state water compact—to look elsewhere for new supply. Among options explored, beginning in 1998, was the possibility of diverting water from the Tennessee River in Southeast Tennessee via pipeline. The mere consideration of this alternative partly convinced officials in Tennessee of that state's vulnerability to outside demands and led the legislature to pass an *Interbasin Water Transfer Act* in 2000 to regulate transfers of water from a basin originating in Tennessee to any other basin—in or outside the state (26–28).

These and other disputes over water transfers are instructive for one major reason: they underscore that decisions to undertake, or to deny, transfers, though partly based on perceived needs and environmental concerns, often result from the distribution of power and authority. Worldwide, a significant impediment to different jurisdictions reaching common accord on water allocation, and coming into conflict over transfers, is the perception by protagonists that they lack adequate influence over the outcome of negotiations or the ability to leverage decisions.

Oftentimes, this lack of influence is manifested in a lack of legal or economic tools (e.g., equitable water markets, effective regulatory frameworks, infrastructure to allocate water and manage competing demands) to resolve disputes. Just as inequalities of power and resources have been recognized as inhibiting durable solutions to water conflicts between, for example, Israel and Jordan, and Israel and the Palestinian West Bank (especially when compounded by secrecy and fundamental value differences), they may also inhibit compromise and accord among states within a nation (29–32). Likewise, when countries—like states within a country—find the means to cooperate, it is often the result of developing conjoint, equitable cooperative management schemes that are adequately funded, as in the Mekong River basin states of Thailand, Cambodia, and Vietnam, for example (33). The challenges facing transfer decisions are formidable, but opportunities for working toward common solutions are plentiful.

BIBLIOGRAPHY

1. Fradkin, P.L. (1996). *A River No More: The Colorado River and the West*. University of California Press, Berkeley, CA.
2. Bright, J. (2000). *A History of Israel*. John Knox Press, Westminster.
3. Feldman, D.L. (1991). The great plains garrison diversion unit and the search for an environmental ethic. *Policy Sci.* **24**(1): 41–64.
4. Jones, J.D., Sikora, V.A., and Woodward, J. (1983). *Study of Tennessee Water Resources Law: Legal Considerations for Effective Water Management Under Conditions of Shortage*. Research Report No. 97, Water Resources Research Center, The University of Tennessee, Knoxville, TN.
5. Grant, D.L. (1998). Introduction to interstate allocation problems. *Waters and Water Rights*. R.E. Beck (Ed.). Michie, Charlottesville, VA.
6. Laitos, J.G. and Tomain, J.P. (1992). *Energy and Natural Resources Law in a Nutshell*. West Publishing, St. Paul, MN.
7. Butler, L.L. (1990). Environmental water rights: an evolving concept of public property. *Va Environ. Law J.* **9**: 323–327.
8. Tarlock, A.D. (1997 and 1998 update). *Law of Water Rights and Resources*. Clark, Boardman and Callaghan, Deerfield, IL.
9. Wright, K.R. (Ed.). (1998). *Water Rights of the Eastern United States*. American Water Works Association, Denver, CO.
10. Samson, S.A. and Bacchus, S. (2000). Point—water marketing: the other side of the coin. *Water Resour. Impact* **2**(6) November: 15–16.
11. Pearson, R.P. (2000). Managing water scarcity—southwestern style. *Troubled Waters—Managing a Vital Resource—Global Issues Electron J.* March.
12. Gillilan, D.M. and Brown, T.C. (1997). *Instream Flow Protection: Seeking a Balance in Western Water Use*. Island Press, Washington, DC.
13. Tarlock, A.D., Corbridge, J.N., Jr., and Getches, D.H. (1993). *Water Resource Management: A Casebook in Law and Public Policy*, 4th Edn. The Foundation Press, Westbury, NY.
14. Kenney, D.S., McAllister, S.T., Caile, W.H., and Peckham, J.S. (2000). *The New Watershed Source Book—A Directory and Review of Watershed Initiatives in the Western United States*. Natural Resources Law Center, University of Colorado School of Law, Boulder, CO, April.
15. Pringle, C.M. (2000). Threats to U.S. public lands from cumulative hydrologic alterations outside of their boundaries. *Ecol. Appl.* **10**(4) August: 971–989.
16. Rosenberg, D.M., McCully, P., and Pringle, C.M. (2000). Global-scale environmental effects of hydrological alterations: Introduction. *BioScience* **50**(9) September: 746–751.
17. *Decisions of the Department of the Interior—Federal Water Rights of the National Park Service, Fish and Wildlife Service, Bureau of Reclamation and the Bureau of Land Management*. (1979). Office of the Solicitor, U.S. Department of the Interior, Washington, DC, June 25.
18. U.S. Bureau of Reclamation. (1957). *Colorado-Big Thompson Project,* Vol. 1. *Planning, Legislation, and General Description*. U.S. Department of the Interior, Denver, CO.
19. Keenan, S.P., Krannich, R.S., and Walker, M.S. (1999). Public perceptions of water transfers and markets: describing differences in water use communities. *Soc. Nat. Resour.* **12**: 279–292.
20. Landry, C.J. (2000). Agriculture and water markets in the new millennium. *Water Resour. Impact.* **2**(3) May: 13–15.
21. *Water in the West: Challenge for the Next Century*. (1998). Report of the Western Water Policy Review Advisory Commission. Published by NTIS: Springfield, VA, June.
22. *Restoring the Waters*. (1997). Natural Resources Law Center, the University of Colorado School of Law, Boulder, CO, May.
23. Arrandale, T. (1999). The eastern water wars *Governing* (August): 30–34.
24. Leahy, T. (1998). Lake Gaston project resolves water supply shortage. *Public Works: Eng. Constr. Maint.* **129**(7) June: 44–48.
25. (1995). Hopes dimming for pact to settle Gaston dispute. *Waterweek* June 19: 7.
26. (1997). *Atlanta Regional Water Supply Plan Update*. Adopted December 3, Atlanta Regional Commission—unpublished report.
27. Feldman, D.L. and Elmendorf, J. (2000). *Water Supply Challenges Facing Tennessee: Case Study Analyses and*

the Need for Long-Term Planning. Prepared for the Environmental Policy Office, Tennessee Department of Environment and Conservation, by the Energy, Environment and Resources Center, University of Tennessee, Nashville, TN, June.

28. (2000). *Inter-Basin Water Transfer Act*. H.B. 3002 & S.B. 3074. General Assembly of Tennessee.
29. Hassoun, R. (1998). Water between Arabs and Israelis: reaching twice-promised resources. In: *Water, Culture, and Power: Local Struggles in a Global Context*. J.M. Donahue and B.R. Johnston (Eds.). Island Press, Washington, DC, pp. 313–338.
30. Elmusa, S.S. (1995). Dividing common water resources according to international water law: the case of the Palestinian Israeli waters. *Nat. Resour. J.* **35**(Spring): 223.
31. Morris, M.E. (1993). *Dividing the Waters: Reaching Equitable Water Solutions in the Middle East*. RAND/P-7840. RAND Corporation, Santa Monica, CA.
32. Morris, M.E. (1992). *Poisoned Wells: The Politics of Water in the Middle East. Reprinted from Middle East Magazine*, RAND/RP-139. RAND Corporation, Santa Monica, CA.
33. Browder, G. and Ortolano, L. (2000). The evolution of an international water resources management regime in the Mekong River Basin. *Nat. Resour. J.* **40**(3): 499–532.

RESERVED WATER RIGHTS FOR INDIAN AND FEDERAL LANDS

EVE WOODS
Denver, Colorado

Most rights to the use of water are created under state law. In general, states east of the Missouri River follow riparian law and the nineteen contiguous western states follow the prior appropriation doctrine or a hybrid of the two. The federal government has a strong policy of deference to state water law. However, some federal lands and tribal lands have a unique type of water right that occurs under federal law. These "reserved water rights" attach to federal lands that the Government kept or set aside for particular purposes and to Indian reservation lands that are held for the benefit of a federally recognized American Indian tribe.[1]

Historically, the United States set aside land from the public domain for federal purposes such as parks and national forest without taking the actions necessary to establish a water right under state law. It also negotiated treaties with American Indian Tribes that failed to address the question of a water supply for the lands set aside for the tribe. The doctrine of reserved water rights is based on the contract law theory that the federal government intended to reserve adequate water for itself when it set aside land for federal purposes, even though it did not say so expressly in the document setting aside the land (1). Likewise, when an Indian tribe reserved some of its land for itself while ceding the rest to the United States by treaty, this doctrine recognizes that the tribe must have intended to reserve some waters for itself, to make the land it reserved useable.

The United States Supreme Court first recognized a reserved water right in the 1908 case of "Winters v. United States." The Court held that the Fort Belknap Tribe in Montana and the United States had impliedly reserved water for the Fort Belknap Reservation via their treaty of 1888. The Tribe was awarded 5000 acre feet of water from the Milk River for irrigation use on the Reservation, with a priority date of 1888, which gave the tribal water right priority over the rights of the upstream settlers who had begun farming the area in 1898. Reserved water rights are also called "Winters rights" after the name of that case (2).

From 1908 to 1963, the reserved rights doctrine was applied only to tribal water rights, and rarely at that. In 1963, in *Arizona v. California*, the reserved rights doctrine was extended to provide water rights for nontribal federal lands as well.

Reserved water rights share some characteristics with both riparian water rights and prior appropriation water rights. Reserved rights are riparian in the sense that they extend to the waters adjacent to or beneath the lands that they serve (3). They are riparian because they exist independent of whether the water has been diverted or applied to beneficial use. The full quantity of the reserved right develops in lieu of use and cannot be lost by nonuse. However, unlike riparian rights, a reserved right has a priority date. It is administered within the prior appropriation system—"first in time is first in right"—and so has the power to call out junior users in times of water shortage.

QUANTITY OF THE RIGHT

The quantity of a reserved water right is "that amount necessary to fulfill the purposes of the reservation" (4). It encompasses both the current and the future needs for water on the reserved lands (5).

This amorphous quantification standard has been difficult to implement. For federal lands set aside under an Act of Congress, the Congressional purpose stated or implied in that Act is the "purpose of the reservation" used for measuring the water right (6). The court uses expert testimony to determine the amount of water necessary to preserve wildlife or habitat, or to meet whatever purpose applies. If the land was set aside for more than one purpose, there can be only one reserved water right, and it will only be enough water to meet that primary purpose of the reservation (7).

Until recently, quantification of tribal reserved water rights was usually based on the PIA standard, i.e., enough water to grow crops on the "practicably irrigable acreage" of the Indian reservation, which was because many treaties stated that the purpose of the reservation was to develop a viable agricultural economy, intending that the tribal members would assimilate to the lifestyle of the nonIndian farmers who were then settling the West. To determine the number of practicably irrigable acres, courts look to economic and engineering experts to determine for how many acres the economic benefits of growing crops

---

[1] A list of federally-recognized tribes is published annually by the Department of the Interior in the Federal Register.

will at least equal the costs of transporting and applying the water to the land (8).

More recently, an Arizona court rejected the PIA standard in favor of a "homeland" standard for quantifying tribal reserved rights (9). The purpose of the reservation in that case was not agriculture, but was to provide a permanent homeland for the tribe. The court held that determining the minimal amount of water necessary to accomplish the homeland purpose—both now and in the future—would require analysis of six diverse factors concerning the tribe's cultural practices and the potential for economic development on the reservation. That new test has not yet been applied by any court.

The quantities of reserved rights vary greatly. They range from one thousand acre feet to hundreds of thousands of acre feet. Many tribal and federal reserved rights have yet to be quantified. The eventual use of those reserved rights will have a potentially enormous effect on the amount of water available to junior users on those streams.

McCARREN AMENDMENT

Although they are created under federal law, reserved water rights can be adjudicated by state courts (43 U.S.C. § 666). In the McCarran Act, the United States Congress waived federal sovereign immunity from suit to allow state courts to join federal claims into a comprehensive adjudication of a river system or other source. This 1952 Act prompted many western states to revise their statutes to allow for general stream adjudications, thereby forcing the quantification of reserved rights claims. Such adjudications are complex, requiring inter alia, a field analysis of each water claim in the basin to determine the quantity and priority date of each right.

PRIORITY DATE

The priority date of a reserved right is the date of the federal act or tribal treaty reserving the land, which is usually a relatively early right that will have priority over all water users who settled the area after that date. Under certain circumstances, a tribe may have a "time immemorial" priority date for its reserved right. This water right is, by definition, the earliest right on the stream (10).

The Pueblos of New Mexico also have time immemorial water rights. However, Pueblo water rights are not reserved rights. The quantity of a Pueblo water right is limited to past beneficial use and cannot be expanded to meet current or future Pueblo needs (11). Like reserved rights, Pueblo rights cannot be lost for nonuse.

BIBLIOGRAPHY

1. *United States v. Winans*. (1905). 198 U.S. 371.
2. *Winters v. United States*. (1908). 207 U.S. 564.
3. *In re General Adjudication of Gila River*. (1999). 989 P.2d 739 (Ariz).
4. *In re Big Horn*. (1988). 753 P.2d 76 (Wyo), *aff'd sub nom, Wyoming v. U.S.* (1989). 492 U.S. 406.
5. *Arizona v. California*. (1963). 373 U.S. 546.
6. *United States v. New Mexico*. (1978). 438 U.S. 696.
7. *Cappaert v. United States*. (1976). 426 U.S. 128.
8. *Arizona v. California*. (1963). 373 U.S. 546.
9. *In Re General Adjudication of Gila River*. (2001). 35 P.2d 68 (Ariz.).
10. *United States v. Adair*. (1984). 723 F.2d 1394 (9th Cir. 1983), *cert. denied*, 467 U.S. 1252.
11. *New Mexico v. Aamodt*. (1976). 537 F.2d 1102 (10th Cir.).

READING LIST

Tarlock, Corbridge, and Getches. (2002). *Water Resource Management: A Casebook in Law and Public Policy*, 5th Edn. Foundation Press, New York, pp. 839–909.

Colby, B. (2004). *Tribal Reserved Water Rights*.

WETLANDS POLICY IN THE UNITED STATES: FROM DRAINAGE TO RESTORATION

Andrea K. Gerlak
Columbia University
New York, New York

Historically, swamps have been regarded as wastelands, evoking reactions of "disgust at their sight and smell, fear of malaria and yellow fever, and unease about rich resources running to waste within them." George Washington's leadership in the Great Dismal Swamp Company characterized early notions of swamps. The Company, designed to drain Virginia's swamp for timber production, promoted the notion that swamps were useless lands (1). The Swamp Land Acts of the 1850s are a further illustration as the federal government encouraged states to "reclaim" nearly 6.5 million acres of wetlands.

By the beginning of the twentieth century, perceptions began to change. Sportsmen and hunters called attention to declining number of waterfowl and initial federal efforts regarding wetlands focused almost exclusively on waterfowl habitat. In 1903, President Theodore Roosevelt established the first National Wildlife Refuge by executive order to protect endangered species on Pelican Island, Florida. With the Migratory Bird Hunting Stamp Act of 1934, Congress established a special fund to finance wetland acquisitions for duck habitat.

Despite such efforts, however, water resources continued to be managed largely for economic development purposes. The U.S. Army Corps of Engineers (Corps) became a key instrument in "modernization" during America's era of industrialization (2), and exerted broad discretion by undertaking improvements in the navigability of the nation's waterways under the 1880 Rivers and Harbors Act and the 1899 Refuse Act. Many federal programs provided incentives for wetland conversion, including subsidies for agriculture, grazing and forestry, reservoir construction, and flood relief and insurance.

By the 1960s, environmental values emerged to challenge the old paradigm. Public alarm over declines in many animal species, including the symbolic bald eagle, led to endangered species legislation in the 1960s and 1970s. The term "wetland" replaced "swamp." The Army Corps inched toward reform as well. In response to an emerging concern for water pollution, the Corps began to revise its permit regulations to take into account "public interest review," beyond pure economic impacts. Based on *Zabel* v. *Tabb* [430 F.2d 199 (5th Cir. 1970), cert. denied, 401 U.S. 910 (1971)] in 1971, the first federal court case over wetlands, the Corps was granted "ecological authority" to weigh noneconomic factors in reviewing and denying permit applications under its original 1899 authority. Real change occurred from passage of the Federal Water Pollution Control Act of 1972. Known as the Clean Water Act (CWA), this legislation marked the beginning of broad-based wetlands protection in the United States. Wetlands came to be valued as biologically diverse ecosystems providing great public benefits such as improved water quality and protection from floods. This marked the beginning of federal efforts at wetlands protection. During the next 30 years, wetlands management would become highly politicized (3).

WETLANDS, LEGISLATION, AND LITIGATION

By the Federal Water Pollution Control Act amendments of 1972, Congress established a comprehensive program of regulations and permits to control water pollution. Section 404 of the CWA marked the first federal policy designed to preserve wetlands. Section 404(b)(1) states that "the guiding principle should be that degradation or destruction of special sites may represent an irreversible loss of valuable aquatic resources." In less than two pages, Congress established a shared authority between the Army Corps and the newly created U.S. Environmental Protection Agency (EPA). As a political compromise, the Corps was designated as the lead agency to administer the permit system for discharging dredged and fill materials into navigable waters. The Corps' preliminary permit decisions would then be reviewed by the EPA that had authority to veto the Corps' decision. This compromise, although appearing innocuous, would complicate and burden Section 404 implementation.

A historic tradition in water resource management, the Corps embodied the older distributive/allocative model of policy-making, referred to as the "pork barrel." Such an approach allowed a high degree of local and state agenda-setting autonomy. The newly created EPA, meanwhile, embraced the new, centralized regulatory model, referred to as "command and control," and focused primarily on water quality. Given the distinct organizational culture of the agencies (4) implementation of Section 404 immediately was characterized by conflict between the EPA and Corps. Generally, the Corps preferred a limited approach, whereas the EPA pushed for expansive authority to regulate all wetlands in the United States.

The federal judiciary would play a powerful role as environmental groups challenged the Corps' implementation and regulatory practices. By *Natural Resources Defense Council* v. *Calloway* (392 F. Supp. 685) in 1975, the U.S. Supreme Court pushed the Corps towards broader wetlands protection and forced the Corps to revise its regulations to include significant intrastate waters, in addition to interstate waterbodies. This case sparked debate in Congress and by the 1977 amendments to the CWA, Congress refined and expanded its definition of wetlands demonstrating a clear congressional commitment to wetlands protection. In response to industry concerns, however, the amendments allowed exemptions for ongoing normal ranching, farming, and silviculture.

President Carter attempted to create a more unified wetland policy by issuing Executive Order 11990 (42 Fed. Reg. 26 (1977). In calling on all federal agencies to minimize wetland loss on land managed by the federal government, this action represents the highest administrative action ever taken to protect wetlands. Two year later, Carter's Attorney General Benjamin Civiletti gave greater weight to the EPA in wetlands implementation [43 Op. Attorney General 15 (1979)]. In attempting to resolve interagency disputes, he determined that the EPA enjoyed ultimate authority to decide the scope of wetlands regulation because that agency carried the responsibility for implementing the other sections of the 1972 CWA.

The courts pushed wetlands protection even further with *U.S.* v. *Riverside Bayview Homes* (474 U.S. 121) in 1985. Here, the U.S. Supreme Court issued a landmark decision in wetlands policy. By defining the waters covered broadly, the Court ruled that the Corps and EPA could regulate wetlands even when no surface water or other apparent hydrologic connection exists between the wetlands and an adjacent waterbody.

THE POLITICS OF WETLAND MANAGEMENT

At about the same time that it enacted the CWA, Congress passed legislation that brought still more federal agencies and departments into the wetlands policy arena. A 1970 statute, The Water Bank Act, created a federal program that paid farmers to preserve wetlands habitat for waterfowl. Administered by the Department of Agriculture's Soil Conservation Service (now the Natural Resources and Conservation Service), this program has grown significantly during the past 30 years. In 1973, still another dimension was added when Congress passed the Endangered Species Act (ESA). Administered by the U.S. Fish and Wildlife Service (FWS), this highly controversial statute added another federal player to the list of agencies implementing the wetlands protection policy. Consultation with the FWS became the norm because so many wetlands are home to endangered and threatened species. In some situations, the FWS has been given the authority to manage sensitive wetland areas. For example, in 1973, more than 107,000 acres of the Great Dismal Swamp was deeded to the Department of the Interior for the creation of the Great Dismal Swamp National Wildlife Refuge. How far we have come from the days of George Washington's Great Dismal Swamp Company. Thus, during the last 30 years, the FWS has become an integral component of national wetlands policy.

By the end of the "environmental decade" of the 1970s, at least four federal agencies had major responsibilities for protecting wetlands. A number of other federal agencies, including the Forest Service, the Bureau of Reclamation, and the Bureau of Land Management, had ancillary responsibilities. Table 1 provides a list of the primary federal agencies involved in wetlands protection today.

Other legislation added several more bureaucratic participants to an already crowded landscape. For example, the Farm Service Agency in the Department of Agriculture came to administer and enforce the Swampbuster provision contained in the 1985 Farm Bill. This disincentive program denies federal farm program benefits to producers who plant an agricultural commodity on converted wetlands. Officials of the Corps believe that Swampbuster has significantly reduced agriculture-related Section 404 permit applications (6). Environmentalists, however, have been critical of the program and the Corps' inadequate enforcement of it (7). Nonetheless, beginning in the 1980s, falling agricultural prices coupled with this shift in federal wetland policies helped reduce agriculture-related wetland loss. According to the U.S. Department of Agriculture analysis, wetland conversions on agricultural lands nationwide slowed to less than 30,000 acres per year between 1987 and 1991, down from 51,000 acres per year between 1982 and 1987 (8, p. 234). However, the building boom of the 1980s increased urban wetlands conversion to 89,000 acres per year, or 57% of the total gross conversion (9, p. 21).

Table 1. Primary Federal Agencies Involved in Wetland Protection[a]

Federal Agency	Role in Wetland Protection
Department of Defense, Army Corps of Engineers	Administers wetland permitting and enforcement under Section 404 of the Clean Water Act
Environmental Protection Agency	Performs oversight over Corps' permitting process and enforcement
Department of the Interior, Fish and Wildlife Service	Reviews Section 404 permit applications, makes recommendations to the Corps, and promotes voluntary restoration programs
Department of Agriculture, Farm Service Agency	Manages voluntary programs to help protect and restore wetlands associated with farmland
Department of Agriculture, Natural Resources Conservation Service	Delineates wetlands under farm bill legislation and provides technical and financial assistance to landowners to restore and enhance wetlands
Department of Commerce, National Oceanic and Atmospheric Administration	Initiates wetland restoration activities associated with marine ecosystems

[a] Reference 5.

In the modern era, implementation of Section 404 has been characterized by heightened polarization of competing interests and politicization of science. The polarization of interests can be traced to the wetlands manual debate of the late 1980s and early 1990s. In 1987, agencies moved to adopt a single manual to delineate wetlands during the permit process. Economic interests, led by the Farm Bureau and the National Wetlands Coalition (NWC), vehemently opposed such efforts. Mobilized under the NWC, this coalition of agriculture, oil, mining, utilities, and real estate developers, strove to "work with Congress and the Administration for legislative reform of regulatory improvements to the federal wetlands permitting program" (10). By drafting legislation and lobbying members of Congress, NWC seeks to protect private property rights and streamline the regulatory process to offer greater regulatory flexibility.

The White House responded with a proposal that restricted the Corps' and EPA's jurisdiction by narrowing the definition of a wetland. Environmental critics charged that the Bush Administration's 1991 manual would write off some 20 to 40 million acres of wetlands (11, p. 5). Field tests of the manual indicated that 30–80% of land delineated as wetlands in the 1989 manual were excluded by the 1991 manual (12). The debate over a politically acceptable definition of wetlands was never resolved. Implementing agencies returned to the 1987 wetlands manual for the wetlands delineation process. Congress eventually moved the wetland delineation issue out of the political arena and into the scientific one. In 1992, Congress authorized a study of wetlands by the National Research Council. Released in 1995, the National Research Council's report, titled "Wetlands: Characteristics and Boundaries," contained a scientifically grounded wetland definition and more than 80 recommendations for improving federal wetlands regulation (13). The report still awaits serious attention and implementation.

Most recently, the debate has focused on the 2001 Supreme Court decision, *Solid Waste Agency of Northern Cook County v. U.S. Army Corps of Engineers* [121 S.Cf. 675]. In this case, a county solid waste agency sought permission to develop a disposal site on 17 acres of land that contained isolated, seasonal ponds used by migratory birds. By a 5–4 decision, the Court denied CWA protection to isolated, nonnavigable ponds and wetlands contained within a single state under the Migratory Bird Rule. In doing so, the Court "reversed elements of more than 15 years of federal jurisdiction when it ruled that the federal government overstepped its authority in regulating certain isolated wetlands" (14, p. 2). Regional offices of the Corps and EPA were then instructed to withhold clean water protection from those types of isolated waters and to seek guidance from headquarters in determining whether to protect other small intrastate streams and waterways that enjoy federal protection (15). Both environmental and industry representatives agree that there is a need for congressional guidance following this ruling. Of course, both sides have radically different notions of what that response should entail.

THE MODERN DEBATE: WETLANDS CONSERVATION AND RESTORATION

Today, the wetlands debate has been framed by a "no-net-loss" goal. Adopted from the National Wetlands Policy Forum, a task force convened by the Conservation Foundation at the request of the EPA and headed by New Jersey Governor Thomas H. Kean in 1988, "no-net-loss" means conservation of wetlands whenever possible and that acres converted should be offset through restoration and creation, thereby, maintaining or increasing the wetland resource base (16). Every president since George H. Bush has adopted this approach, and in doing so, the debate has been framed around acreage. As such, it is the numbers that matter. Wetland values and functions become less important. Recently, the National Academy of Science's (NAS) Committee on Mitigating Wetlands Loss reported that it is "not convinced that the goal of no net loss for permitted wetlands is being met for wetland functions" (17, p. 3). According to the NAS, insufficient data exists to determine if the goal was even being met.

What has been determined is that wetland conversion rates have dropped during from 458,000 acres per year during the 1950s–1970s to 58,500 acres per year in the 1980s–1990s. This amounts to an 80% reduction in average annual loss from the prior decade (18, p. 9). Figure 1 displays wetlands loss from the 1950s to the 1990s. Conservation programs have helped stem wetland loss nationally. Recent reauthorization of the North American Wetlands Conservation Act where the federal government matches donations for state wildlife agencies, conservationists, landowners, and sportsmen demonstrates this. So too does reauthorization of the Wetlands Reserve Program, which provides financial incentives to farmers to restore and protect wetlands through long-term easements. Congress has elected the conservation approach as opposed to amending the Section 404 permit process. Employing incentive programs like these, as opposed to regulation, are understandably more popular with the public than the Section 404 permit process.

By focusing on wetlands acreage framed by the "no-net-loss" goal, restoration strategies have emerged as the preferred policy solution, attempting to reconcile economic growth and wetland protection. Restoration efforts are becoming increasingly popular as a strategy to mitigate wetland losses and counter past development. Restoration of the Florida Everglades, a joint federal–state effort at an estimated cost of $8 billion over the next two decades, represents the largest ecological restoration project ever attempted. Despite restorative successes of coastal areas (19), however, some scientists are increasingly concerned that hard to create wetlands like bogs and fens are being replaced with common, easy to create wetland types such as cattail marshes (20). Moreover, the National Research Council reports that restored wetlands often cannot re-create natural levels of biodiversity and sustainability (21). In the coming years, the debate over wetlands protection is likely to focus as much on the *quality* of wetlands as it does now on the *quantity* of acres restored.

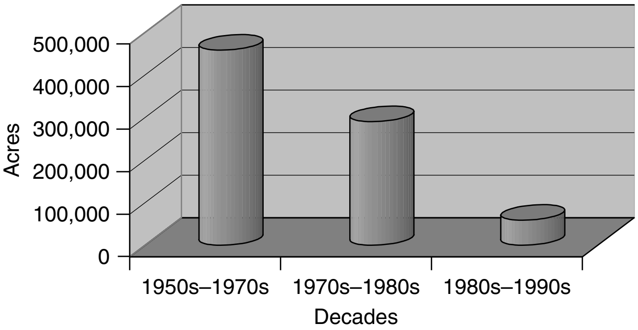

Figure 1. Wetlands loss figures for the mid-1950s to mid-1970s, mid-1970s to mid-1980s, and 1986–1997 (18).

CONCLUSION

In summary, by the 1960s and 1970s, environmental values challenged wetland drainage and destruction practices. Passage of the Clean Water Act marked the first significant federal efforts at wetland protection. Wetland management during the past 30 years has been characterized by litigation, an expansion of federal agency involvement, polarization of interest, and a politicization of science. We have not yet achieved the "no-net-loss" goal, but conservation programs and restoration efforts have helped to stem wetland loss nationally. Will they be enough to protect the wetlands that remain in the United States? Only time will tell.

BIBLIOGRAPHY

1. O'Hagan, K. (1991). Pumping with the intent to kill: Evading wetlands jurisdiction under Section 404 of the clean water act through draining. *DePaul Law Rev.* **40**: 1059–1080.
2. Feldman, D.L. (1991). *Water Resources Management: In Search of an Environmental Ethic.* Johns Hopkins University Press, Baltimore, MD.
3. Gerlak, A.K. (1996). Federal wetland policy: Is this policy all wet? In: *At the Nexus: Science Policy*. D.L. Soden (Ed.). Nova Science, Comack, NY.
4. Clarke, J.N. and McCool, D.C. (1996). *Staking Out the Terrain: Power and Performance Among Natural Resource Agencies*. State University of New York Press, Albany, NY.
5. U.S. General Accounting Office. (1998). *Wetlands Overview: Problems with Acreage Data Persist*. General Accounting Office, Washington, DC.
6. Zinn, J. and Copeland, C. (1996). *Agricultural Wetlands: Current Program and Legislative Proposals*. CRS Report for Congress 96-35 ENR, The Library of Congress, Congressional Research Service, Washington, DC.
7. Environmental Working Group. (1995). *"Taxing" the Taxpayer: Wetlands and Farm Subsidies in Perspective*. Environmental Working Group Policy Memorandum, Washington, DC.
8. Heimlich, R.E. (1994). Costs of an agricultural wetland reserve. *Land Econ.* **70**(2): May.
9. Heimlich, R.E., Wiebe, K.D., Claassen, R., Gadsby, D., and House, R.M. (1998). *Wetlands and Agriculture: Private Interests and Public Benefits*. United States Department of Agriculture, Economic Research Service, Washington, DC.

10. National Wetlands Coalition. *Mission Statement* at http://www.thenwc.org/home.htm, accessed September 5, 2003.
11. Keefer, M. (1991). *Wetlands: Going, Going, Going*. National Wildlife Federation Leader, Reston, VA.
12. Environmental Defense Fund and World Wildlife Fund. (1992). *How Wet is a Wetland? The Impacts of the Proposed Revisions to the Federal Wetlands Delineation Manual*. Environmental Defense Fund and World Wildlife Fund, New York and Washington, DC.
13. National Research Council. (1995). *Wetlands: Characteristics and Boundaries*. National Academy Press, Washington, DC.
14. National Wetlands Newsletter. (2001). *The SWANCC Decision*. Environmental Law Institute, Washington, DC, January–February.
15. U.S. Environmental Protection Agency. (2003). *Administration to Reaffirm Commitment to No Net Loss of Wetlands and Address Approach for Protecting Isolated Waters in Light of Supreme Court Ruling on Jurisdictional Issues*. Environmental Protection Agency, Washington, DC.
16. The Conservation Foundation. (1988). *Protecting America's Wetlands: An Action Agenda—The Final Report of the National Wetlands Policy Forum*. Conservation Foundation, Washington, DC.
17. National Research Council. (2001). *Compensating for Wetland Losses Under the Clean Water Act*. National Academy Press, Washington, DC.
18. Dahl, T.E. (2000). *Status and Trends of Wetlands in the Conterminous United States, 1986–1997*. United States Department of the Interior, Fish and Wildlife Service, Washington, DC.
19. Kukoy, S.J. and Canter, L.W. (1995). *Mitigation Banking as a Tool for Improving Wetland Preservation via Section 404 of the Clean Water Act*. The Environmental Professional 17, Gainesville, FL.
20. Hunt, R.J. (1996). *Do Created Wetlands Replace the Wetlands that are Destroyed*. U.S. Department of the Interior, U.S. Geological Survey, Madison, WI.
21. National Research Council. (2001). *Compensating for Wetland Losses under the Clean Water Act*. National Academy Press, Washington, DC.

WATER HISTORY, ART, AND CULTURE

CURIOUS USES OF AGRICULTURAL WATER IN THE WORLD

Eric P. Mollard
Institut de Recherche pour le Développement
France

The agricultural use of regularly flooded zones along large eastern rivers, from the Nile to the Yangtze, goes back to the births of cities and almost to the birth of agriculture. Elsewhere, wetland farming has been heterogeneous. Initially, marshes, flood plains, and coastal areas were sparsely reclaimed, whereas nearby identical wetlands were forsaken and are still unchanged today. At a later time, public intervention based on resource and technology resulting from the Industrial Revolution undertook a lot of total draining. The territorial gain for modern agriculture was accompanied by a dramatic loss of specific advantages such as incomparable aquatic and land biodiversity which can also serve agriculture.

Wetlands show two contradictory characteristics: they discourage and they attract. They are inhospitable because of the fear of marsh fevers, wizards, and legendary animals. More prosaically, great land fertility and moisture are beneficial at the cost of hard labor, not in the direction of general draining, but in that of the work for horticulture. The promise of such lands that have high productivity and are healthier when water is under control, could overcome the fears, real or imaginary, of the wilderness.

The physical duality of high productivity and unhealthiness was translated over time by a social duality, especially between natives and foreigners. Natives supplemented their livelihood by fishing and collecting reeds in a common resource land. The cases of fever were undoubtedly used to keep young people or free riders from defying the elders' authority and to discourage foreigners from seeking free lands. Due to a spooky atmosphere on foggy days, the local populations ended up believing in evil spells launched by their ancestors. Where the foreigners succeeded in settling, one often noted a dramatic social reversal. Indeed, through intense labor and the great fertility of the marsh soils, the last ones to settle became the first. The Aztecs in old Mexico epitomized such history. In the twelfth century A.D., a tribe, originating from the arid north and vassals of the kingdoms occupying the high plateau, could settle only on the immense lake that occupied most of the Valley of Mexico and had a few small islands full of snakes. Tenaciously, the Aztec people created a lake-based agriculture, the chinampas, that enabled them to build one of the largest cities in the world and to subjugate the neighboring kingdoms later.

The relation between wetlands, foreigners, a power that can become an empire, old legends, and current speeches from many stakeholders confers an integral social history on this environment. However, it would be erroneous to believe that the curious uses of water result today only in some anecdotes and relics. Two examples suffice to show how this history is not finished. The mulberry tree dikes on the Pearl River in China are a contemporary and extremely long-lived peculiar use of water. The second example is even more dynamic. The delta of the Chao Phraya River, which used to be a rice bowl in Thailand, is becoming a second Holland; it has horticulture for export, thanks to the technique of raised gardens.

The peculiar uses of water in agriculture prove to be invaluable elements in understanding the relation between water and societies, as well as in designing alternatives to commonplace agriculture instead of the common thought of large polder and irrigation schemes. In this article, we distinguish two types of uses: "raised" style and "hollow" style. The first refer to peasant reclamation, on an individual basis but as a large extension, whereas the second deal with some types of irrigation. The most curious feature of the story is that a local peculiarity is not so peculiar on a planetary level. Inventive men have found similar solutions on every continent.

THE AGRICULTURE OF HOLLOWS

The agriculture of hollows is not a simple adaptation to marshes and lowlands. It embraces the entire range of intensification from floating gardens to cultivated ponds.

Floating Agriculture

Nowadays, vegetables are still grown on natural, floating vegetation rafts or rafts man-made of reed bundles. They are used sometimes as seedbeds before firm ground transplanting. Floating gardens are located on some large lakes in Asia, Latin America, and Africa. Those of Lake Inlé were rediscovered when the border opened in Myamar, whereas those of Lake Dahl, in Kashmir, are less known. Several Chinese lakes, such as Lake Tien close to Kunming in Yunnan, would have witnessed this type of water farming. On Lake Titicaca in the Andes, the rafts of reed are platforms with fishing holes, as in an ice barrier, several thatched houses, some animals, and a small garden. It is not exactly a floating garden because agricultural activities take place on firm ground. On Lake Chad, the natural floating islands are used by fishermen who must continually move under penalty of sinking. In Yap, in Micronesia, the taro is sometimes cultivated on floating gardens cut in the marsh.

Former Chinese authors distinguished two types of rafts: wooden platforms and rafts made of reeds. Both were used for growing rice on the marshy edges of rivers, where a dike would have been difficult to build. Nowadays, the wooden rafts have disappeared, and the floating gardens

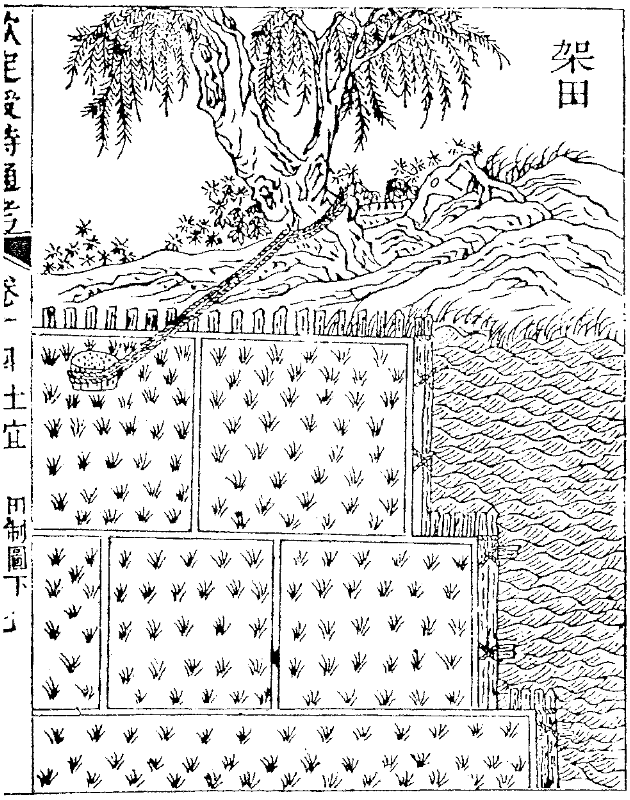

Figure 1. Floating fields.

on reeds are used for more lucrative market gardening (see Fig. 1).

Floating rice growing is a type of flood farming. When the water depth does not exceed 1 meter, one speaks of deep water rice growing with stand bearing and fast lengthening varieties; some of them survive short immersion. Beyond 1 meter, rice varieties have floating stems that lengthen in more than 5 meters of water and lie down when the water recedes. Such farming still exits in the extensive basins along the unprotected rivers in Asia (Chao Phraya, Ganges, Irrawaddy, and Mekong) and in Africa (the inland delta of Niger in Mali). If the water lasts, the harvest takes place in boats. The light work required during the season enabled a family to cultivate up to 20 ha. The great surplus fed the clergy, the warriors, and the rulers of the Angkor Empire in Kampuchea.

The Qochas

In the Quechua language, qocha means lake and generally, any artificial or natural basin. The qochas form connected sets of basins. The stretches of land that separate them are dedicated to animal breeding. Qochas are found in a few areas of the Bolivian and Peruvian altiplano at an altitude of 3850 meters, in particular in the department of Puno. Qochas have three basic forms: round, oblong, and rectangular (see Figs. 2 and 3). Most frequent are round ones that measure between 30 and 200 meters in diameter and have an average surface of 6,000 m^2. The rectangular ones are 150 m by 30 m. Whatever the form, the base is horizontal and the depth varies from 1 to 6 meters (on average, 2 m). Its diameter is crossed by a deep furrow through which water arrives and leaves; the base is surrounded by one or two circular furrows. Perpendicular furrows join the soft sloped edges. The qocha is used as a cultivated field or water tank. Its size is generally expressed in *masas*, the surface plowed by a team of two men and one woman in 1 day. Each family has six or seven *masas*. The crops are mainly potato, qaniwa, and kinuwa, rotating with oats, barley, corn, or a grazed fallow of variable duration.

The Diked Fields

Bordo and caja, dike and basin field, respectively, are words indiscriminately used for many techniques of counterfields or diked fields in Mexico. In tune with farmers, it is preferable to use a neutral terminology because, behind the appearance of a similar technique, the functions and practices vary according to site: flooded banks of a lake or river, rivers of a torrent, or mountain valley. Among the agricultural effects of the use of the technique are irrigation, protection against flooding, fertilization, control of weeds and pests, and water storage.

The Mexican midwest is a string of basins partitioned by volcanic cones of all sizes; many basins do not have natural discharge systems. Before the reclamation of many lakes, flooding was inescapable, and not very deep lakes retracted in dry season. Individual or organized diked fields were both used. Earth embankments measured between 60 centimeters and 2 meters high. They enclosed a surface ranging from several to more than 50 ha. A master diked field was filled by tapping a river or a source. In rainy years, the secondary diked fields were filled,

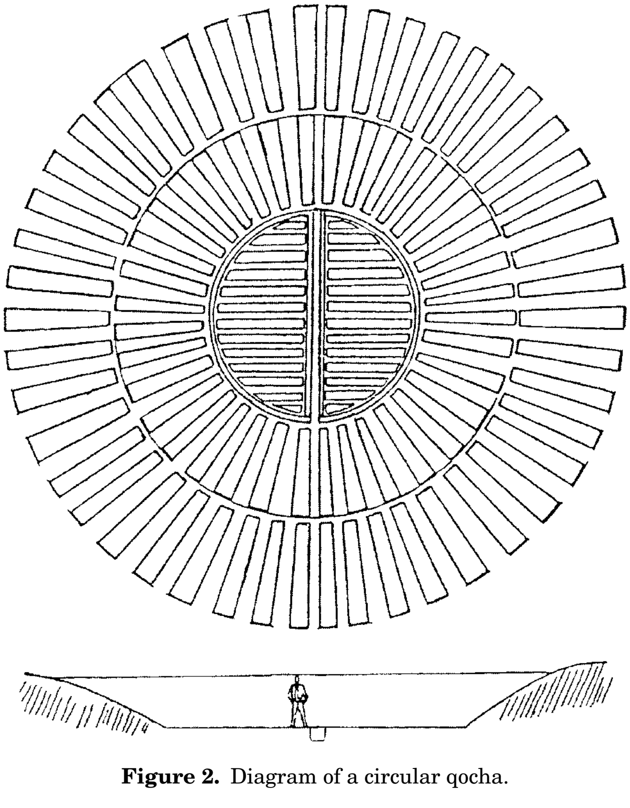

Figure 2. Diagram of a circular qocha.

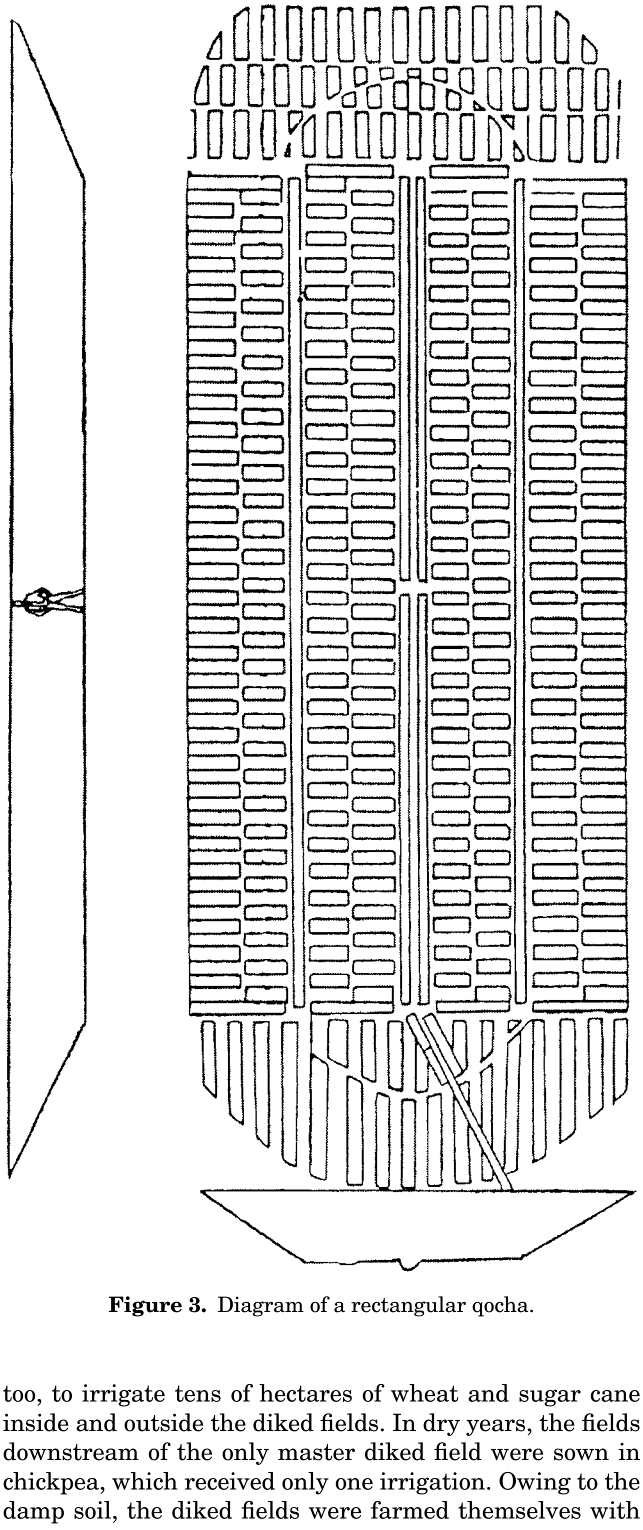

Figure 3. Diagram of a rectangular qocha.

too, to irrigate tens of hectares of wheat and sugar cane inside and outside the diked fields. In dry years, the fields downstream of the only master diked field were sown in chickpea, which received only one irrigation. Owing to the damp soil, the diked fields were farmed themselves with the same chickpea. The dry years provided a considerable surplus of quality fodder which made it possible to feed the cattle of the haciendas and also large pigsties. This region of Mexico is still the greatest industrial pig-breeding area in spite of the disappearance of most bordos during the twentieth century.

The haciendas built the systems of diked fields. After the agrarian reform, many of them were shared with many peasants who were unable to maintain such extensive systems. However, it is possible nowadays to find some diked fields used for lentil and chickpea farming.

Cultivated Ponds: Dombes and Pearl River

Dombes constitutes a undulating basin northeast of Lyons, France. The singularity of the area rests on 1 year of the crop, the assec (or dry pond), which alternates with two years of ponding, the évolage (the root eve comes from the Latin aqua). Agrifishing activity is closely connected with rights on water and land. These rights are contradictory to the French law that does not allow a land to be flooded by a dam built by a neighbor. Only a combination of circumstances and powerful events have perennialized this agriculture.

Dombes is a relatively wet area where clay soils are difficult to work and prevent agriculture from developing. As an independent principality until 1762, it was used as refuge against invaders. The proximity of Lyons, a gastronomical center requiring fresh fish of high quality, and the political independence of the close Dombes, which made it possible to maintain a special law there, are the causes of this agriculture. Indeed, each pond is a particular case combining technical adjustments and contractual agreements between landowners, the owners of the water and dam, and the sharecroppers growing assec, as well as the stakeholders upstream and downstream who can benefit or be affected by runoff, flooding, or water scarcity.

The technique of mulberry tree dikes associated with a fishing pond draws its peculiarity from the narrow association of farming and fishing. These joint activities are more productive than as separate activities. If integrated agriculture is well represented in China and in Vietnam, it seems that the delta of the Pearl River has world exclusivity in the association of the mulberry tree and fishing ponds. In the northern deltas, the frosts damage the mulberry trees and, in the south, heat and extensive flooding limit fish farming. Other situations in the world share this context, but they lack population and markets, which govern integrated agriculture, and the skills required in silkworm farming that the mulberry leaves supplied.

THE AGRICULTURE OF RAISED STYLE

The Gardens of the War

In the 1970s, the remote valleys in the center of New Guinea revealed a Stone Age society. The agriculture of the Dani tribes there is curious. Tuber crops are grown on raised fields built in marshes with spatula-ended sticks and worked by hand. A few-year fallow is characteristic of semi-intensive agriculture. Furthermore, they are "gardens of war" insofar as watchers prevent the raids of enemy tribes that seek to steal the working women. The Dani supplement their production with tubers produced in slash and burn farming. In such a society, where production is not for marketing, unlimited labor is first a social activity, as well as war.

Chinampas and Hortillonnages

Hortillonnages are raised beds from 0.80–1.50 m high. They are separated by ditches and right-angle channels;

both are full of water all year. The drainage makes the roots of the plants safe and gives them the possibility of benefiting from the soil fertility and permanent irrigation. The variety of raised fields in the world provides many local adaptations. The term "Chinese beds" must be reserved for raised beds constructed each season in paddy fields. As for the denominations of raised beds, drained fields, ridged agriculture, poldered fields, campos elevados, camellones, and chinampas, they apply to simple banked-up beds, more or less permanent traditional beds, furrow irrigation, and the various market gardening types. The French language has the specific term of hortillonnages. An unverified anecdote credits Julius Caesar with the use of the term hortilani (perhaps for garden) when he crossed the Somme river near Amiens. Dani from New Guinea show that this type of wetland use is accessible by technology from the Stone Age and that it can be old.

Classification must distinguish seasonal and permanent uses. Seasonal cropping is related to the drying up of channels, flooding, or winter. Permanent cropping is possible in tropical countries; irrigation is ensured, and flooding is controlled. In this case, one distinguishes individual flood control by poldered raised beds on the Thai pattern or by a collective dike on the lakes as was the case at the time of the Aztecs in Mexico. The first pattern includes two alternatives: on the one hand, those crafted for vegetable farming on lower beds (80 cm) that have bent surfaces and, on the other hand, orchard beds $1\frac{1}{2}$ meters high and flat surfaces. For vegetables, construction is by hand to avoid compacted soil, whereas the second option is mechanized. In Thailand, one also finds raised beds for coconut groves where water freely circulates from tides and flooding. In Mexico, Mayas built channeled market gardens where the ditches are dug on the edges of marshes; they are lengthened by traditional embanked raised beds.

In Europe, raised gardens are nowadays located under the cathedrals at the centers of cities, such as in Amiens and Bourges. Some are found in the marshes near Saint-Omer to produce cauliflower. The Atlantic marshes have preserved their raised meadows, but horticulture appeared only between the two wars; the fields were used mainly by cows, which are taken there in small boats. In Latin America, Mexico City nowadays preserves the chinampas in the district of Xochimilco wedged into the urban fabric. In Asia, they do not seem to be more than three centuries old, even in China. In Sumatra and Thailand they are promising a extension for export horticulture. In Burma and Vietnam, their development is still limited. In Java, at the beginning of the century, the surjan technique was developed, which associates rice growing in channels and a mixed crop on raised beds.

Finally, the many archaeological ruins gathered under the term camellon all over the American continent, which are vast undulations, do not necessarily proceed from raised bed or chinampa patterns. They would be closer to the semi-intensive agriculture practiced nowadays in Africa and in the valley bottoms of Rwanda, designed for producing sweet potatoes. Intensive, permanent horticulture requires a market and, as long as no close cities have been identified, camellon ruins bear witness to a more seasonal, low-input horticulture.

READING LIST

Altieri, M.A. Traditional farming in Latin America. *The Ecologist* **21**(2): 93–96.

Armillas, P. (1971). Gardens on swamps. Archeological research verifies historical data on Aztec land reclamation in the valley of Mexico. *Science* **174**(4010): 653–661.

Bérard, L. (1989). *Structures foncieres et organization sociale*. La Dombes au début du XXe siecle, Monde Alpin et Rhodanien (Le), No. 1–2, pp. 39–56.

Billaud, J-P. (1984). *Marais poitevin: rencontres de la terre et de l'eau*. L'Harmattan, Paris, p. 265.

Bruneau, M. and Bernot, L. (1972). Une population lacustre: les Intha du lac Inlé. *Jatba* **XIX**(10–11).

Chabert, J.P. (1984). Terres et eaux en Dombes. In: *Les hommes et l'eau*. Etudes Rurales Paris, No 93–94, pp. 87–89.

Coe, M.D. (1964). The chinampas of Mexico. *Sci. Am.* **211**(1): 90–98.

Dubois, P. (1913). La vallée de la Somme. Les hortillonnages, les pêcheries, les tourbières. *Bull. Soc. Géo. Lille* **59**: 324–338.

Jeffery, D. (1974). Burma's Leg Rowers and Floating Farms. *National Geography*: June 826–840.

Jimenez Osornio, J. (1990). Famous floating gardens threatened. *Ileia Newsletter*: October 26.

Jimenez-Osornio, J.J. and del Amo, R.S. (1988). *Intensive Mexican Traditional Agroecosystem: The Chinampa*. In: Allen and Van Dusen (Eds.). pp. 451–452.

Lamballe, P. (1993). Floating rice systems of South Prey Veng. *Gret* **20**.

Lambert, J.D.H., Siemens, A.H., and Arnason, J.T. (1984). Ancient Maya drained field agriculture: its possible application today in the new river floodplain, Belize. *Agric. Ecosystems Environ.* **11**: 67–84.

Leng, T.S. (1987). *Cultivation of deepwater and floating rice in Kampuchea*. Proc. Int. Deepwater Rice Workshop, IRRI, pp. 159–167.

Limouzin, P. (1994). Une agriculture périurbaine en voie de banalisation: l'exemple amiénois. *Bull. Assoc. Géogr. Franç.* **2**: 156–159.

Micha, J-C. (1987). L'aménagement des marais tropicaux en écosystème agro-piscicole. *Bull. Séanc. Acad. R. Sci.* Outre-Mer **31**(1985-3): 371–389.

Moriarty, J.R. (1968). Floating gardens (chinampas) agriculture in the old lakes of Mexico. *America Indigena 28(2) et Katunob* **7**(1) 1969: 59–75.

Perceveaux, P. (1972). Les étangs dombistes dans la première moitié du XIX siècle. *Bull. Soc. Naturalistes Archéologues Ain.* **86**: 83–107.

Salas, H. (1988). *Ecological reclamation of the chinampa area of Xochimilco, Mexico City*. In: A. Patricia and D. Van Dusen (Eds.). pp. 469–474.

Sceau, R. (1980). Les étangs de la Dombes: Fondements socio-économiques d'un système hydraulique. *Revue de géographie de Lyon* **55**(2): 129–159.

Serpenti, L.M. (1965). *Cultivators in the Swamps*. Amsterdam van Gorcum, Assen, p. 308.

Siemens, A.H. *Reducing the risk: some indications regarding pre-hispanic wetland agricultural intensification from contemporary use of a wetland/terra firma boundary zone in*

Central Veracruz, University of British Columbia, dep. geography/Vancouver, Ecological Studies.

Siemens, A.H. and Puleston, D.E. (1972). Ridged fields and associated features in Southern Campeche: new perspectives of the lowland Maya. *Am. Antiquity* **37**: 228–239.

Sluyter, A. (1994). Intensive wetland agriculture in Mesoamerica: space, time, and form. *Annals. Assoc. Am. Geographers* **84**(4): 557–584.

Song, X.F. (1997). Rice cultivation on floating-beds in different natural waters. *Chin. Rice Res. Newsl.* **5**(2): 4–5.

Thom, D.J. and Wells, J.C. (1987). Farming systems in the Niger inland delta, Mali. *Geographical Rev.* **77**(3): 328–342.

Warren, D.M. Slikkerveer, L.J., and Brokensha, D. (1995). *The Cultural Dimension of Development. Indigenous Knowledge Systems*. Intermediate Technology Publications, London.

Wilken, G.C. (1983). *A Note on Buoyency and Other Dubious Characteristics of the "Floating" Chinampas of Mexico*. In: Prehistoric Intensive agriculture in the tropics. Farrington (Ed.). British archeological Reports, Int. series 232, Oxford, pp. 31–48.

Wilken, G.C. (1997). Food-producing systems available to the ancient Maya. *Amer. Antiquity* **36**: 432–448.

You, X.L. (1991). Growing rice on the floating fields in ancient China. *Chin. J. Rice Sci.* **5**(2): 79–82.

WATER BETWEEN ARABS AND ISRAELIS: RESEARCHING TWICE-PROMISED RESOURCES

Clive D. Lipchin
Ann Arbor, Michigan

INTRODUCTION

On November 29, 1947, the General Assembly of the United Nations passed Resolution 181 in favor of partitioning Palestine into a Jewish state and an Arab state. The Zionist movement, which had been preparing itself for statehood accepted the plan, but the Palestine Arabs rejected it. On May 14, 1948, the state of Israel was proclaimed. The borders of the newly proclaimed state of Israel were a strategic nightmare representing two fighting serpents, one the new Jewish state, the other the West Bank under the autonomy of Jordan and the Gaza Strip under Egypt (1). The problems with the borders were many: lack of military strategic depth in the center of the country, an arid south, and a fractured north. No less problematic were access to water resources.

The majority of Israel's water resources are transboundary and are thus, in a sense, shared with her neighbors. The most contentious are the underground aquifers of the West Bank shared between Israel and the Palestinians and the tributaries to the Jordan River that originate in the occupied Golan Heights, which belonged to Syria prior to the 1967 Six-Day War. Since 1948, Israel has embarked on an ambitious development program for the country. The drive behind this program was the Zionist ethos of developing the land to create an attachment between the Jew and the soil. Massive agricultural and water projects ensued in the mid-1950s and 1960s in realizing the Zionist dream. The most significant was the National Water Carrier completed in 1964. The National Water Carrier is a complex system of pipes, pumps, and canals that transfers water from the Sea of Galilee in the north to the arid Negev Desert in the south for irrigation. The repercussions of this interbasin transfer have been the decline of the flow of the Jordan River that flows south from the Sea of Galilee to the Dead Sea. Today, the Jordan River is a trickle of its former self, and the Dead Sea is at its lowest level ever.

Since the mid-1980s, agricultural expansion has declined in Israel and has been surpassed by rapid industrial and urban development. Although agriculture is still the primary consumer of water in the country, the growing industrial and urban sectors require more and more water. In 1948, the Jewish population was approximately 500,000; today, it is approximately 5 million. At least 30% of Israel's water budget now comes from the aquifer systems that underlie the West Bank. These aquifers are recharged on the eastern slopes of the West Bank within autonomous Palestinian regions but flow westward to the Mediterranean across the 1967 "Green Line" into Israel. Both Palestinians and Israelis claim sole sovereignty over these resources due to this hydrologic reality.

Water has featured prominently in Israel's continuous peace negotiations with her Arab neighbors. Two significant opportunities have occurred in equitably managing these transboundary waters. The first was the shuttle diplomacy of U.S. special ambassador Eric Johnston who was sent to the region by the Eisenhower administration in 1953 and 1954. The second were the Oslo Interim Accords signed between Israel and the Palestinians in 1993 and the Israel–Jordan Peace Treaty of 1994.

THE FIRST PROMISE

During the early 1950s and mid-1960s, Israel embarked on two major public works projects to harness water resources for agricultural development. Both of these projects, the draining of the Hula Swamps north of the Sea of Galilee and the creation of the National Water Carrier on the northwestern shore of the Sea of Galilee, entailed diversion of waters from the Jordan River watershed without any formal agreement between the other riparians of the time, Syria and Jordan. Syria's objections to the projects were focused on the confiscation of Arab-owned land in the region of the swamps and Israel's activities in the demilitarized zones set up after the 1948 War of Independence (2). The rattling of sabers over the Jordan at this time prompted the Eisenhower administration to send special ambassador Eric Johnston to try to resolve the water dispute.

Eric Johnston came to the region with a proposal for a joint water development project to provide hydroelectricity and an irrigation network for the benefit of all the states in the region and, in addition, to create fertile land capable of supporting up to 900,000 Palestinian refugees on the West Bank of the Jordan (1). The plan was based on the Tennessee Valley Authority (TVA), which provides hydroelectric power to seven states in the Tennessee Valley. The TVA is the largest public utility in the United States.

Johnston visited the region in 1953 and 1954. The final agreement concerned allocation of the waters of the Jordan River. The final allocations were 45% to Israel and 55% to the Arabs (1). The Johnston Accords, as the agreement became known, was the first serious attempt at cooperation between Israel and her neighbors. It also vindicated third-party mediation (this time, on behalf of the United States) to bring the parties together. The role of external mediators in the Arab–Israeli conflict is now a requisite condition expressed by all parties. Nonetheless, the Arab leaders refused to sign the agreement because, as a matter of principle, they were opposed to the formal recognition of Israel (1). But both Jordan and Israel informed the Americans that they would treat the agreement as if it had been signed, and it served as the basis for Israel's diversion of water via the National Water Carrier and Jordan's diversion of the Yarmouk River via the East Ghor canal (2).

Israel's relationship with Syria, however, increased in belligerency. After completion of the National Water Carrier in 1964, Syria sought to deprive Israel of the headwaters of the Jordan by diverting the waters around Israel and began construction of a canal in 1965 for this purpose. According to Sherman (3), Syria had little to gain from the canal either economically or strategically, and Israel responded to the canal's construction as a personal affront to her national aspirations. In a series of military reprisals, Israel destroyed the bulldozers being used to dig the canal and, in effect, caused the Syrians to abandon the project. It is widely believed that these skirmishes contributed to the atmosphere of confrontation that finally led to the Six-Day War in 1967.

In the interim period between the Johnston Accords of 1954 and the signing of the Oslo Declaration of Principles in 1993, Israel and her neighbors undertook an independent approach to water development. Increasing populations and growing demands for more freshwater in the region coupled with a lack of a joint vision for shared management have created a serious deficit in the water budget of all the countries. The urgency of solving the water allocation problem, something that all the parties agree to, is now seen by many as the elusive catalyst for an overall final peace agreement.

THE SECOND PROMISE

At the conclusion of the Six-Day War in 1967, the hydrostrategic map of the region changed in favor of Israel. By the occupation of the West Bank and Gaza, Israel gained access to the mountain and coastal aquifers and to the headwaters of the Jordan in the Golan Heights. Since 1967, Israel's rapid economic growth has deepened her reliance on the aquifer systems of the West Bank. As the quality of life among Israel's citizens has risen, the gap between Israeli and Palestinian water consumption has grown. Before the peace process began, average annual Palestinian domestic consumption was less than 35 cubic meters, whereas in Israel, per capita consumption was about 100 cubic meters (2).

The Oslo process that began in the mid-1990s was considered the means to redress this balance as well as to lead to an overall peace settlement between Israel and the Palestinians. The principal protocol for implementing an interim accord between Israel and the Palestinians on water allocation is Article 40, together with its appendices, signed in Washington by the two parties in 1995 (4). This agreement is commonly known as Oslo Two.

The agreement created a Joint Water Committee (JWC) to deal with all water- and sewage-related issues during the interim period and led to the creation of the Palestinian Water Authority, which gradually took over control of some of the wells that were initially operated and maintained by the Israeli water company, Mekorot. The most significant aspect of Article 40 is Israel's recognition of Palestinian water rights that are to be determined in the Permanent Status Agreement. This agreement was to be completed within five years of the Oslo process.

The Oslo process and open negotiation between Israel and the Palestinians have suffered an irrevocable setback due to the current violence. Many consider the Oslo process now dead. Criticism among the Palestinians for the failure of Oslo is a lack of visible movement on the ground. Israeli settlements in the Occupied Territories continue to expand, the transfer of water resources to the Palestinian Authority by Israel has been delayed, and the freedom to transfer water and expand water and sewage networks among Palestinian urban centers has been curtailed. Part of this problem is the continued effort in Israeli political and military spheres to link water to security concerns. Nonetheless, even at the height of the recent violence, some channels remain open between Israeli and Palestinian scientists for cooperative water management. It remains to be seen if this cooperation can include the political and military sphere. Whether or not the Oslo process is dead, autonomous Palestinian institutions do exist to manage water resources for the majority of its citizens. Coordinated action, however, still remains a necessity for an equitable solution between Israel and the future Palestinian state.

The multilateral activities between Israel and her neighbors that began in Madrid in 1991 paved the way for the Israel–Jordan Treaty of Peace signed in 1994. Out of all of Israel's neighbors, Jordan has always been the most open to negotiations with the Jewish state. Tacit agreements between the two parties on strategic and water issues existed prior to the signing of the peace treaty. Having a formal peace treaty in place, both parties looked forward to peaceful cooperation and exchange in economics, culture, and security.

The Israel–Jordan peace treaty legally spells out mutually recognized water allocations for the Yarmouk and Jordan Rivers, as well as groundwater in the Arava/Araba valley, and calls for joint efforts to prevent water pollution (5). The treaty also acknowledges the water deficit facing both countries and calls for joint efforts to alleviate water shortages through cooperative efforts on both a regional and international scale (5). In the treaty, Israel agrees to provide storage services of 20 million cubic meters of water to Jordan during the winter in the Sea of Galilee and to work with Jordan to find an additional 50 million cubic meters (2). In return, Israel can continue to use wells in the Arava/Araba valley for agricultural purposes despite the fact that, according to the treaty, these wells now lie on the Jordanian side of the border. The region of the Arava/Araba valley is one of the few regions where open cooperation between Israel and an Arab state continues despite the current violence. Israeli agricultural expertise is being used to develop agricultural projects on the Jordanian side of the valley, and cooperation on combating and preventing oil spills in the Gulf of Aqaba continues between the two cities on the gulf, Eilat in Israel and Aqaba in Jordan.

RENEGOTIATING PROMISES MADE

Israel and her neighbors are inching toward a system of integrated watershed management. Today, agreements exist among three riparians of the Jordan River watershed: Israel, the Palestinian Authority, and Jordan. A fourth riparian has so far been excluded from the process. Movement on the Israeli–Syrian track is stalled in the starting gates. Syria insists on the precondition that Israel withdraw from all of the Golan Heights before any negotiations can begin. Israel rejects any form of preconditions as a basis for negotiation. The hydrostrategic importance of the Golan Heights as the headwaters of the Jordan are considered vital to Israel's water needs. For Syria, the issue of the Golan is not water, but pride in regaining the Golan and its position of influence in the Arab world; most of its water comes from the Tigris–Euphrates.

Accommodation with Syria will most likely require a "land for peace" deal. At the moment, the parties are far from agreement on what such a deal will constitute, but whatever deal is finally agreed on will affect the current agreements between Israel and the Palestinian Authority and Jordan. Both of these agreements discuss water allocations independently of Syria, the fourth riparian. Therefore, the allocation systems currently in place will have to be renegotiated once all the riparians are included ultimately which may seem impossible in today's climate, but it cannot be ignored.

Opportunities for agreement now exist to augment any allocation agreement that may be signed. Desalination is soon to become a reality in the region. Israel, as the dominant riparian, has a lot to gain politically and economically in cooperating with her neighbors in setting up and managing such plants. Trading in desalinated water may be one path toward economic cooperation among the riparians to improve the quality of life for all citizens. Desalination may also help to relieve the pressure on the existing freshwater in the region. Demand management policies also need to be considered, such as a move from agricultural hegemony over water, especially in Israel, toward domestic and industrial uses.

It is likely that the need and the recognition of a sustainable water supply in the region will push the parties together in negotiating an equitable water allocation and management system. This system can then provide the foundation for political agreement by the parties, an agreement that so far has proven to be an elusive promise.

BIBLIOGRAPHY

1. Shlaim, A. (2000). *The Iron Wall: Israel and the Arab World*. Penguin Books, New York.
2. Feitelson, E. (2000). The Ebb and flow of Arab-Israeli water conflicts: are past confrontations likely to resurface? *Water Policy* **2**: 343–363.
3. Sherman, M. (1999). *The Politics of Water in the Middle East: An Israeli Perspective on the Hydro-Political Aspects of the Conflict*. St. Martin's Press, New York.
4. Scarpa, D.J. (2001). *Oslo Interim Accords: With Special Reference to Article 40 (1995)*. Unpublished Manuscript.
5. Wolf, A.T. (2001). *Transboundary Waters: Sharing Benefits, Lessons Learned*. Thematic Background Paper. International Conference on Freshwater, Bonn, Germany, 2001.

THE MYTH OF BAD CHOLESTEROL: WHY WATER IS A BETTER CHOLESTEROL-LOWERING MEDICATION

F. Batmanghelidj
Global Health Solutions
Falls Church, Virginia

We in the medical profession, totally oblivious to the vital roles of cholesterol in the body, have been duped into thinking that this substance causes arterial disease of the heart and brain. The pharmaceutical industry has capitalized on the slogan of "bad cholesterol" and produced toxic chemicals that minimally lower the level of cholesterol in the body in the process, cause liver damage to thousands of people, some of whom die as a result of using these medications.

It is surprising that none of the frequently quoted and media-popularized doctors has reflected on the fact that cholesterol levels are measured in blood taken from the veins, yet nowhere in medical literature is there a single case of cholesterol having caused obstruction of the veins. Venous blood moves far move slowly than arterial blood and thus would be more inclined to have cholesterol deposits if the assumption of "bad cholesterol" were accurate. This mistake by us in the medical community and its exploitation by the pharmaceutical industry has caused unnecessary mass hysteria and a multibillion-dollar annual drain of financial resources by those who can least afford it in these times of serious economic crisis.

In truth, the so-called bad cholesterol is actually far more beneficial than is appreciated. The reason for its rise in the body is complications caused by chronic unintentional dehydration and insufficient urine production. Dehydration produces concentrated, acidic blood that becomes even more dehydrated during its passage through the lungs before reaching the heart—because of evaporation of water in the lungs during breathing. The membranes of the blood vessels of the heart and main arteries going up to the brain become vulnerable to the shearing pressure produced by the concentrated, acidic blood.

These membranes begin to lose water to the concentrated blood and thus lose resistance and structural integrity. The shearing force of thick, toxic blood tends to produce abrasions and minute tears in the lining of the arteries. Sections of these membranes could easily peel off and become emboli reaching the brain, kidneys, and other organs if some protective measure were not in place. To prevent the damaged blood vessel walls from getting peeled off by blood working its way under the torn lip of the membrane, low-density (so-called bad) cholesterol immediately coats and covers up the abrasions and tears and protects the underlying tissue like a waterproof bandage until the tissue heals. The advantage of “greasy” cholesterol is that blood components could roll over the patch without sticking to it and produce thrombosis in these vital arteries.

Thus, the vital, lifesaving role of low-density cholesterol proves that this substance is of utmost importance. It saves the lives of those who do not adequately hydrate their bodies by enabling their blood to flow easily through the blood vessels without causing damage.

Cholesterol has infinite vital roles. If the physiology of the body places a greater demand on some aspects of cholesterol functions, the liver manufactures more of it. Not knowing any of these functions and interfering with its actions is foolish—yet it is the present state of medicine, as far as cholesterol is concerned. Here are some other of these actions.

Cholesterol is an element from which many of our hormones are made. The body also makes vitamin D from cholesterol, which is deposited in skin that is exposed to sunlight. Cholesterol is used in the insulating membranes that cover our nerve systems. Cholesterol is used in dehydrated cells to bond the cell membrane structure together. There is no such thing as bad cholesterol. If all the primary ingredients are available for its normal functions, the human body does not engage in making things that are bad for its survival. Until now, we did not know that water is a vital nutrient that the body needs at all times—and in sufficient quantity.

By a convoluted process, the rise in the cholesterol concentration of blood is also an indicator of developing osteoporosis. Briefly, water manufactures hydroelectric energy for many of the energy-dependent physiological functions of the body; neurotransmission is on top of the list. In dehydration and low hydroelectric energy output, ATP energy trapped in calcium bonds becomes an alternative source of energy supply. Each atom of calcium that bonds to another calcium traps one unit of ATP energy.

Thus, the bones of the body, the repositories of much stored energy, begin to “decalcify” to contribute to the emergency energy needs of the body. As an integrated biochemical process in the body and to offset the catastrophic outcome of severely decalcified and softened bones, the manufacture of more vitamin D to increase the rate of bone formation becomes a dominant drive. To achieve this, the body has no alternative but to manufacture more cholesterol to be converted to vitamin D by sunlight on exposed skin. This process was used at the turn of last century to cure rickets in children, and it should be used for adults suffering from osteoporosis. They also need to be better hydrated for their bones to become firm and better calcified.

Water itself—not caffeine-containing beverages that further dehydrate—is a better cholesterol-lowering medication than any chemical on the market. It is absolutely safe, unlike the dangerous medications so often used. Please share this information with those you care for.

For more information about my medical breakthrough on the topic of chronic unintentional dehydration and the diseases it causes, read as many of the following titles as you can. They are the products of more than 24 years of focused research on the topic of deep dehydration inside the cells of the body: *Your Body’s Many Cries for Water*; *ABC of Asthma, Allergies and Lupus; Water for Health, for Healing, for Life; How to Deal with Back Pain and Rheumatoid Joint Pain; Water Cures: Drugs Kill*; and my 2004 book, *Obesity, Cancer, and Depression: The Deadly Diseases of Dehydration*. For more information on dehydration, visit www.watercure.com and www.nafhim.org.

EARLY CLOCKS

Physics Laboratory, National
Institute of Standards and
Technology

Not until somewhat recently (that is, in terms of human history) did people find a need for knowing the time of day. As best we know, 5000 to 6000 years ago great civilizations in the Middle East and North Africa began to make clocks to augment their calendars. With their attendant

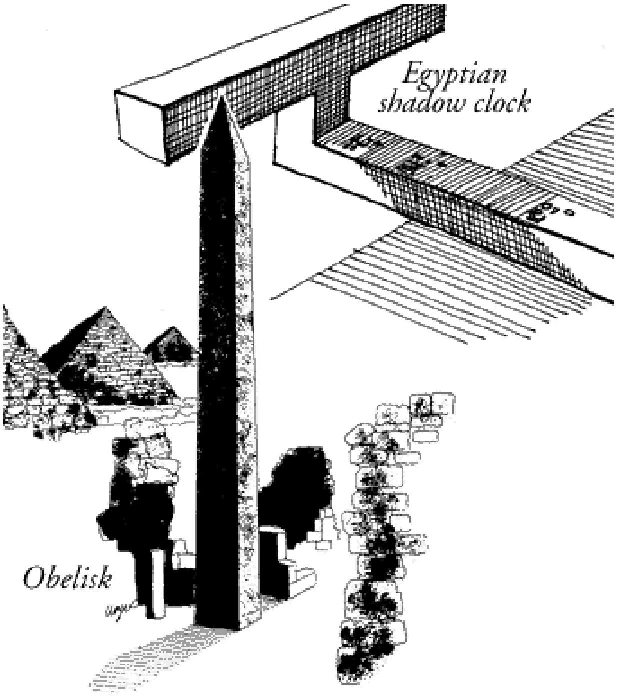

---

bureaucracies, formal religions, and other burgeoning societal activities, these cultures apparently found a need to organize their time more efficiently.

SUN CLOCKS

The Sumerian culture was lost without passing on its knowledge, but the Egyptians were apparently the next to formally divide their day into parts something like our hours. Obelisks (slender, tapering, four-sided monuments) were built as early as 3500 BCE. Their moving shadows formed a kind of sundial, enabling people to partition the day into morning and afternoon. Obelisks also showed the year's longest and shortest days when the shadow at noon was the shortest or longest of the year. Later, additional markers around the base of the monument would indicate further subdivisions of time.

Another Egyptian shadow clock or sundial, possibly the first portable timepiece, came into use around 1500 BCE. This device divided a sunlit day into 10 parts plus two "twilight hours" in the morning and evening. When the long stem with 5 variably spaced marks was oriented east and west in the morning, an elevated crossbar on the east end cast a moving shadow over the marks. At noon, the device was turned in the opposite direction to measure the afternoon "hours."

The *merkhet*, the oldest known astronomical tool, was an Egyptian development of around 600 BCE. A pair of merkhets was used to establish a north-south line (or meridian) by aligning them with the Pole Star. They could then be used to mark off nighttime hours by determining when certain other stars crossed the meridian.

In the quest for better year-round accuracy, sundials evolved from flat horizontal or vertical plates to more elaborate forms. One version was the hemispherical dial, a bowl-shaped depression cut into a block of stone, carrying a central vertical gnomon (pointer) and scribed with sets of hour lines for different seasons. The hemicycle, said to have been invented about 300 BCE, removed the useless half of the hemisphere to give an appearance of a half-bowl cut into the edge of a squared block. By 30 BCE, Vitruvius could describe 13 different sundial styles in use in Greece, Asia Minor, and Italy.

ELEMENTS OF A CLOCK

Before we continue describing the evolution of ways to mark the passage of time, perhaps we should broadly define what constitutes a clock. All clocks must have two basic components:

- *A regular, constant or repetitive process or action to mark off equal increments of time.* Early examples of such processes included the movement of the sun across the sky, candles marked in increments, oil lamps with marked reservoirs, sand glasses (hourglasses), and in the Orient, knotted cords and small stone or metal mazes filled with incense that would burn at a certain pace. Modern clocks use a balance wheel, pendulum, vibrating crystal, or electromagnetic waves associated with the internal workings of atoms as their regulators.
- *A means of keeping track of the increments of time and displaying the result.* Our ways of keeping track of the passage of time include the position of clock hands and digital time displays.

The history of timekeeping is the story of the search for ever more consistent actions or processes to regulate the rate of a clock.

WATER CLOCKS

Water clocks were among the earliest timekeepers that didn't depend on the observation of celestial bodies. One of the oldest was found in the tomb of the Egyptian pharaoh Amenhotep I, buried around 1500 BCE. Later named *clepsydras* ("water thieves") by the Greeks, who began using them about 325 BCE, these were stone vessels with sloping sides that allowed water to drip at a nearly constant rate from a small hole near the bottom. Other clepsydras were cylindrical or bowl-shaped containers designed to slowly fill with water coming in at a constant rate. Markings on the inside surfaces measured the passage of "hours" as the water level reached them. These clocks were used to determine hours at night, but may have been used in daylight as well. Another version consisted of a metal bowl with a hole in the bottom; when placed in a container of water the bowl would fill and sink in a certain time. These were still in use in North Africa in the 20th century.

More elaborate and impressive mechanized water clocks were developed between 100 BCE and 500 CE by Greek and Roman horologists and astronomers. The added complexity was aimed at making the flow more constant by regulating the pressure, and at providing fancier displays of the passage of time. Some water clocks rang bells and gongs; others opened doors and windows to show little figures of people, or moved pointers, dials, and astrological models of the universe.

A Macedonian astronomer, Andronikos, supervised the construction of his *Horologion*, known today as the Tower of the Winds, in the Athens marketplace in the first half

Early water clock

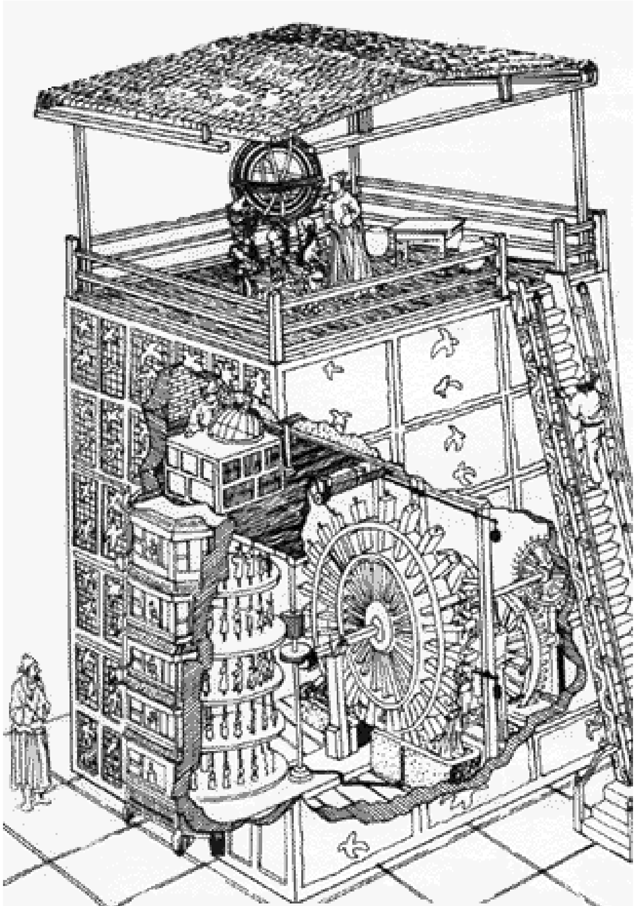

Su Sung water clock tower

of the first century BCE. This octagonal structure showed scholars and shoppers both sundials and mechanical hour indicators. It featured a 24 hour mechanized clepsydra and indicators for the eight winds from which the tower got its name, and it displayed the seasons of the year and astrological dates and periods. The Romans also developed mechanized clepsydras, though their complexity accomplished little improvement over simpler methods for determining the passage of time.

In the Far East, mechanized astronomical/astrological clock making developed from 200 to 1300 CE. Third-century Chinese clepsydras drove various mechanisms that illustrated astronomical phenomena. One of the most elaborate clock towers was built by Su Sung and his associates in 1088 CE. Su Sung's mechanism incorporated a water-driven escapement invented about 725 CE. The Su Sung clock tower, over 30 feet tall, possessed a bronze power-driven armillary sphere for observations, an automatically rotating celestial globe, and five front panels with doors that permitted the viewing of changing manikins which rang bells or gongs, and held tablets indicating the hour or other special times of the day.

Since the rate of flow of water is very difficult to control accurately, a clock based on that flow could never achieve excellent accuracy. People were naturally led to other approaches.

WATER CLOCKS

William E. Marks
Water Consciousness, Inc.
Martha's Vineyard,
Massachusetts

The science or art of measuring time is known as "horology," and, for the greater part of human history, time was calculated through the use of water clocks.

Ever since the dawn of humanity, we have enjoyed a relationship with water. It was this relationship with water that gave birth to the concept of measuring time.

It is believed that the first human-made timekeeping device was the water clock invented in Egypt around 1400 B.C. This was a very rudimentary timekeeping device based on the flow of water out of or into a bowl-type container. This filling or emptying of a measurable vessel with water provided a basis for telling time. The energy source that powered the first water clocks, and all others that followed, was the fluid properties of water responding to the influence of gravity.

The name given to early water clocks in Greece was "klepsydra" (pronounced KLEP-suh-druh). This name meant "thief of water" and was derived from two Greek words meaning "thief" and "water." These Greek root words can be traced to such present-day words as "kleptomaniac," for one who steals, and water words such as hydrant and hydrology.

The klepsydra migrated from Egypt to Greece, and then later into Rome. The Romans Latinized the Greek *klepsydra* to *clepsydra*, or the *horologium ex aqua*. In early history, there was probably one clepsydra for each village or town. Each clepsydra was set up in a central location and was watched over by responsible representatives, who took turns keeping watch over the device. This "watching" of the clepsydra was necessary to protect the time device and to have an attentive person standing by to fill the water basin constantly.

Some locations used a horn to sound the passage of a certain amount of time; others used a drum or the human voice. In this fashion, everyone within earshot had an idea of the time of day. This ancient system of sounding the passage of time remains in use today. Today, we find the sounding of bells at designated times in centrally located bell towers throughout the world, as well as the blowing of a horn. In some places, there are bell-keepers who ring the tower bell or bells manually at designated times. In other locations, the ringing of bells or blowing of a horn is mechanized and occurs on an automated basis.

In early England and later New England in the United States, there were so-called "walking clocks." A "walking clock" was a person who would use a loud voice to announce the time while walking the streets.

In Muslim countries, where Islam is widely practiced, prayer is timed at five different intervals during the day. In some Islamic areas, there are towers where a human occupant will loudly voice prayers to Allah at the five designated times. While praying aloud to Allah, the prayer-giver always faces in the direction of Mecca.

It is due to this keeping “watch” over the flow of water in early clocks and the watching of time in bell towers and by town criers that we have the etymology for the modern timepiece known today as the “watch.”

THE FIRST CLOCK WAS A WATER CLOCK

Without doubt, the very first so-called “water clock” was humanity observing how water was connected to the timing of natural and cosmic events. The timing of moon tides, the timed event of spring floods when the earth turns on its axis and causes snow and icemelt in the mountains, and the calendar of humans watching celestial events to anticipate spring rains or monsoons are three examples.

To survive, humans require a steady source of freshwater. Therefore, the earliest humans naturally learned to observe, understand, and live within the water cycle of their respective regions. Each water cycle was connected to the sourcing of drinking water, the seasonal growth of various wild plant foods, and the migratory patterns of fish and animals. In this fashion, early humans became aware of a natural water rhythm that was connected to the timing cycle of rainfall, which in turn determined the seasons of life. Because of this intimate relationship of the cycle of water to the cycle of life, humans became aware of a water clock that was inherent in the scheme of nature.

For these reasons, early humans evolved in regions where freshwater was readily available constantly. This is why modern archaeologists find evidence of early civilizations evolving along river basins such as the Tigris–Euphrates Valley and along the Nile River. Besides using the river water for consumption and bathing, the wildlife within and around the rivers provided a steady and plentiful source of food. The rivers also provided a convenient way for evolving humans to travel.

As humans gained knowledge of their regional water cycles, they eventually figured out a methodology for raising livestock and growing crops. In other words, early humans became aware of the “big” water clock of nature, a water clock that had a timing sequence that coincided with certain celestial events. Consequently, we see the construction of structures, such as Stonehenge in England and New Grange in Ireland. These stone structures allowed early humans to forecast, or time, future events in nature based on the water cycle.

The flooding of the Nile River, for example, from the end of June into late October, was responsible for depositing rich sediment in Egypt’s fertile Nile Valley. After the flooding subsided, various crops were planted from late October into late February. These crops were harvested from late February until the end of June, when the cycle would begin again.

The ancient Egyptian calendar was known as a “nilometer” because it was based on a vertical scale that recorded the annual flood levels of the Nile River. In essence, this was a “water clock” of huge dimension. Eventually, the Egyptians fine-tuned their nilometer water clock by breaking it down into twelve months of thirty days each; five days were added at the end. For the early Egyptians, this water calendar worked well with the seasonal flows of the Nile, which is why it became known as the “Nile year.” The Nile year calendar, used by the Egyptians more than 6200 years ago, predated the invention of writing.

The use of the number “12” in creating the number of months is often related to the Egyptians observation of the moon cycle. There are about 12 moon cycles in a year. The division of the year into 12 months may also have provided a basis for the Egyptians and Babylonians to divide the day and night into 12 hour segments.

Originally, the Egyptians and Babylonians divided sunrise to sunset into 12 parts they called “hours,” and sunset to sunrise into 12 parts. The lengths of hours for the day and night are not always the same through the year, so water clocks had to be adjusted every night and morning. Eventually they figured out that it was easier to keep time when they created the concept of the 24-hour day. The 24-hour day encompassed the passage of both day and night hours and therefore required little adjustment of water clocks.

Egyptian water clocks were among the earliest timekeeping devices that did not depend on the observation of celestial bodies. One of the oldest was found in the tomb of the Egyptian pharaoh, Amenhotep I, buried around 1500 B.C.E.

As explained by the National Institute of Standards and Technology Physics Laboratory, early water clocks were named clepsydras (“water thieves”) by the Greeks. The Greeks began using clepsydras about 325 B.C.E. Clepsydras were stone vessels with sloping sides that allowed water to drip at a nearly constant rate from a small hole near the bottom. The drip-drip sound of water dropping from a water clock was an ancient precursor to the ticktock sound emitted by certain mechanical clocks invented in later times. Water clocks were replaced by the invention of mechanical clocks about 700 years ago.

Other clepsydras were cylindrical or bowl-shaped containers designed to fill slowly with water flowing in at a constant rate. Markings on the inside surfaces measured the passage of “hours” as the water level reached them. These clocks were especially used to determine the passing hours at night but may have been used in daylight as well.

Another version is known as the “sinking bowl” (see Fig. 1). This consisted of a metal or clay bowl with a hole in the bottom center. The larger the bowl, the longer it would take to fill with water and sink below the surface. When placed in a water container or calm body of water, the bowl would slowly fill and then sink after a certain amount of time. As soon as one bowl sank, the next bowl would be set on top of the water. Counting the number of bowls that sank during the daytime, or nighttime, created a crude basis for telling time. This method was often used to keep track of the “dark hours” through the night. The amount of time a bowl took to sink was sometimes measured during daylight so it could be calibrated with a sundial. This method helped ancient humans to compare the passage of time between daylight hours and dark hours, which allowed for knowledge of the coming change of seasons. As daylight hours grew shorter and nighttime hours longer, the earth naturally began to cool.

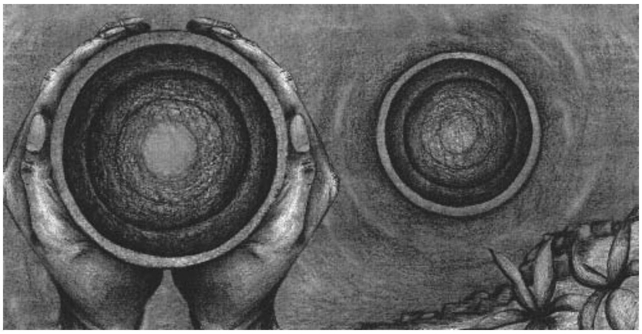

Figure 1. Sinking bowl water clock circa 1400 B.C. to present day. Drawing by: Sarah Adams.

The sinking bowl water clock was so low tech and functional that it remained in use for thousands of years in the Middle East, as well as in parts of North Africa into the twentieth century. The sinking bowl water clock remains in use in some parts of the world today.

More elaborate and impressive mechanized water clocks were developed between 100 B.C.E. and 500 C.E. by Greek and Roman horologists and astronomers. The added complexity was aimed at making the flow more constant by regulating the pressure and at providing fancier displays of the passage of time. Some water clocks rang bells and gongs; others opened doors and windows to show little figures of people, or moved pointers, dials, and astrological models of the universe.

The Greeks as well as the Romans used the water clock as a mechanism for regulating speaking time in courts. In Greece, the clepsydra was used to limit the time for pleading one's case in Athenian courts. Based on water clocks surviving from that time, the water would flow only for about 6 minutes. There are records of legal representatives asking for the flow of the water clock to be stopped so that certain laws or evidence may be presented to the court. On a more humorous level, there are stories where comments about how the orator should quench his thirst by drinking water from the water clock and thereby abbreviate his speaking time.

On a similar note, Romans also used the clepsydra to limit speaking time in Roman legal courts. This limitation of time on legal arguments by Roman orators was instituted by the command of Pompey and was alluded to by Cicero when he stated, "latrare ad clepsydram." When the clepsydra's water ran out, an orator's time ran out, and he was cut off from speaking further. Compared to the 6-minute klepsydras used by Greek courts, Roman courts used a 20-minute clepsydra.

During legal arguments, opposing parties were supposed to enjoy equal amounts of water. However, there are surviving records indicating that lawyers would sometimes request additional amounts of time by asking for a certain quantity of "clepsydrae." If a lawyer spoke too long, the opposing lawyers would argue that his water should be taken away by the court.

Martial, the Latin poet of Rome (c. 40–c. 102), once penned these words about a tiresome speaker who constantly moistened his throat by drinking from a glass of water:

> Seven water-clocks allowance you asked for in loud tones, Caecilianus, and the judge unwillingly granted them. But you speak much and long, and with backtilted head, swill tepid water out of glass flasks. That you may once for all sate your oratory and your thirst, we beg you, Caecilianus, now drink out of the water clock.

This legal limitation of time for presenting legal arguments remains in use in certain governing bodies throughout the world today. Depending on the governing rules of the land, an elected official or legal representative is allowed to hold the floor for only a designated amount of time.

It is also from the Roman legal system that we derived our modern-day A.M. and P.M. designations, representing before midday (*ante meridiem*, A.M.), and after midday (*post meridiem*, P.M.). This A.M. and P.M. split of the day was of legal import to the Romans because lawyers were required to appear in courts before noon.

On another note, the private ownership of a water clock in a home was considered a great status symbol during Roman times. Also, a person who gave the appearance of wasting time in life was referred to as "one who wastes water." Thus, we find an intimate connection between water and life, as perceived in the minds of ancient Greeks and Romans.

In Greece, a Macedonian astronomer, named Andronikos, supervised the construction of a complex Horologion. This mechanical device, known today as the Tower of the Winds, was located in the Athens marketplace during the first half of the first century B.C.E. This octagonal structure showed scholars and shoppers both sundials and mechanical hour indicators. It featured a 24-hour mechanized clepsydra and indicators for the eight winds from which the tower got its name, and it displayed the seasons of the year and astrological dates and periods.

The Romans also developed mechanized clepsydras (water clocks), though their complexity accomplished little improvement over simpler methods for determining the passage of time.

The Greeks divided the year into 12 parts they called "months." They then divided each month into 30 parts they called "days." The Greek year had a total of 360 days, which was equal to 12 times 30. The earth circles the Sun in 1 year, so the Greeks developed the mathematical concept of dividing the circle into 360°.

The division of the hour into 60 minutes and each minute into 60 seconds, probably finds its source in the ancient Sumerian sexagesimal system. This Sumerian system is based on the number 60 and was developed approximately 4000 years ago.

In the Far East, mechanized astronomical/astrological clock making developed from 200–1300 C.E. Third-century Chinese clepsydras drove various mechanisms that illustrated astronomical phenomena. One of the most elaborate clock towers was built by Su Sung and his associates in 1088 C.E. Su Sung's mechanism incorporated a water-driven escapement invented about 725 C.E. The Su Sung clock tower, over 30 feet tall, possessed a bronze, waterpower-driven armillary sphere for observations, an automatically rotating celestial globe, and five front panels with doors that permitted viewing changing manikins which rang bells or gongs and held tablets indicating the hour or other special times of the day. This and other mechanical efforts to use water for keeping time had their limitations. These limitations are inherent in the nature of water.

Water is a sensitive entity that is subject to linear and nonlinear influences. These influences include but are not limited to gravity; temperature; atmospheric pressure (weight of air above at any given moment); wind; alignment of the Moon and other planets; viscosity; sunspots (cause of ionic excitement in atmospheric water and related creation of aurora borealis and aurora australis); speed of the earth's rotation; the "Heisenberg principle" based on quantum mechanics and how that which is observed responds to being observed; sound (travels four times faster through water than through air); light; projection of intention, thought, and emotion as expressed in Masaru Emoto's pioneering water research; chemical and electrical composition of contact materials; suspended solids; the rate of molecular activity within water; the presence of vortical influences; and the spectrum of various wave influences.

Due to water's sensitivity and all of the variables that influence its character at any given second, the rate of flow of water is very difficult to control accurately. Therefore, a clock based on water flow has certain limitations and challenges for achieving high accuracy.

However, throughout time, there have been claims of water clocks that rivaled the accuracy of mechanical clocks. Once such example is an hour-glass type water clock, which had to be inverted as soon as the water from the upper glass chamber emptied into the lower glass chamber. One account of another accurate type of clepsydra, which was invented by a man named, Amontons, was published in the pamphlet, *Remarques & Experiences Physiques fur la Construction d'une nouvelle Clepsydre*, &c. Jombert, Paris, 1695.

Today, no matter what philosophical perspective a person may have about time, the basic fact is that the mechanics of keeping accurate time is a concept created for use as a linear tool.

Regardless of the variety and accuracy of human-made time devices, humankind remains subject to the vicissitudes and rhythm of flowing water, as well as its inherent cycle expressed in nature.

Understanding the rhythm and cycles of water is vital to the survival of life as well as civilization, so we cannot substitute mechanical devices for the water clock presented to us by our natural environment.

Our near future interaction with the earth's water cycle may well determine how much time our species has to survive. In other words, the available water in the earth's clepsydra is rapidly approaching empty at this time, so it behooves us to do more than just talk about altering the way we interact with water.

READING LIST

Beckman, J. (1814). *History of inventions and discoveries* [from old catalog]. Translated from the German by William Johnston, 2nd Edn. Printed for J. Walker [and others], London.

Hill, J.R. *Arabic Water-Clocks*. University of Aleppo, Institute for the History of Arabic Science, Syria, p. 1981.

Pollio, M.V. (1982). *De Architecture: The Como 1521 Edition in Exact Photographic Reproduction.* Ayer Co. Pub., North Stratford, NH.

Rees, A. *Ree's clocks, watches and chronometers, a selection from The Cyclopedia; or universal dictionary of the arts, sciences and literature.* pp. 1819–1820.

Rohr, R.R.J. (1996). *Sundials: History, Theory, and Practice.* Dover, Mineola, NY.

Waugh, A.E. (1993). *Sundials, Their Theory and Construction.* Dover, Mineola, NY.

National Institute of Standards and Technology Physics Laboratory.

www.ClocksWorld.com. *Water Clock*.

www.jim-diana-hart.home.att.net/clock.html. *Medieval Clocks*.

www.zetnet.co.uk/sea.jnp/earth.4/time.htm. *Time and the History of its Measurement, Reproduced from the Science and Technology Equipment News No. 6 by kind permission of the Scottish Schools Equipment Research Centre.*

www.inventors.about.com/library/weekly/aa071401a.htm. *The Invention of Clocks*.

www.cyrstalinks.com/clocks.html. *Clocks*.

www.rit.edu/-dansma/GoodenowOrrRossWaterClocks. *Mathematical Models of Water Clocks*.

www.sts.sunyit.edu/timetech/chinese.html. *Water Clocks*.

www.library.thinkquest.org/CO118228/water_clocks.htm. *Water Clocks*.

www.emie.cummings.net/clock.htm. *Clock - a History - Timekeepers*.

OUR EVOLVING WATER CONSCIOUSNESS

William E. Marks
Water Consciousness, Inc.
Martha's Vineyard,
Massachusetts

Since the dawn of our species, being conscious of water has been vital to our survival, health, and social behavior.

Hundreds of thousands of years before the written word, humans thought and communicated about water. Besides being essential for survival, water remains a mysterious fluid that science continues to research and study.

Prior to the invention of writing, communication about water found expression in human sign language and the spoken word. In early human writings, such as the Egyptian hieroglyphic system, the symbol for water is a ripple pattern that looks like this (𓈖). This early hieroglyphic symbol resembles almost exactly a purported water symbol carved into bone by a Neanderthal during the Paleolithic period.

Another ancient water symbol is the Chinese character "shui" (水). This character is of such antiquity that it predates Chinese recorded history and is of unknown origin. In *The Book of Changes* (I Ching), which is more than 3000 years old, there is the water trigram "Kan" (☵), whose symbol is obviously an evolution out of the more ancient water character of "shui."

The Chinese character for "water," it was recently discovered, exists as a natural water form. In Masaru Emoto's book, *Messages from Water*, he shows a photograph of crystallized water on page 12 that contains what appears to be a Chinese character. This photo is accompanied by the following text:

> "When water freezes, it becomes crystallized. At the moment right before it returns to its water form (with a rise in temperature, between −5 °C and 0 °C) it creates a crystal shape that is identical to the Chinese character for water. Did people in ancient times know this and make the Chinese character for "water" based on this information?"

Simply stated, the *Chinese Book of Changes* helps mankind deal with the idea of change. Or, as reported in the *Analects* (a book of conversations of Confucius, one of the authors of the I Ching), Confucius states, while standing by a river, "Everything flows on and on like this river, without pause, day and night."

THE TALE OF GILGAMESH

On another note, we find the incredible story of water expressed in the epic poem of Gilgamesh, a poem written in Sumer, an ancient civilization located in the lower Euphrates River Valley. Sumer is where writing is believed to have been invented. These first writings were written by Sumerians on clay tablets in the cuneiform script. Cuneiform script is composed of wedge-shaped characters imprinted into wet clay before it is heated (fired) for permanence.

The Sumerians believed that in the beginning only a vast swirling sea existed and that all things came from this sea. Because of this, it is no surprise to learn the Sumerian word mar meant "sea" as well as "womb." Archaeological evidence indicates that Sumer was the home of some of the earliest humans. The fact that humans lived continuously in Sumer since prehistoric times supports the theory of the way the region's ancient oral history found expression in the first writings of the Sumerians.

As can be expected, interpretations of the ancient texts found in Sumer vary. However, one thing stands without question—"water" served as the source of creation. The god Ea-Enki, the principal divinity of the liquid element—especially freshwater, is expressed as the Creator. Ea-Enki is presented as coming from the sea at the dawn of time and is said to be of supreme intelligence, skillful, ingenious, and the One who knows all things.

The name "Ea," which means "House of the Water," makes up the first two letters of "earth." Ea's domain was the Apsu—all the waters that surrounded the planet and kept it afloat, as well as the springs, rivers, and lakes. The freshwaters of the earth were considered the source of happiness, knowledge, and wisdom.

When writing was invented in Sumer, oral stories of old were rendered in clay. From these first writings came the Gilgamesh poem. However, the discovery of the Gilgamesh poem created concern in the minds of some religious leaders.

In 1872, members of the Society of Biblical Archaeology in London were surprised by the implications drawn by archaeologist, George Smith. Smith had recently deciphered the Gilgamesh poem. A poem that predated the Bible by thousands of years expressed a story about the "great flood," and included information about building a ship to save people and animals.

The discovery and translation of the Gilgamesh poem gave humanity a glimpse into water-related information and rituals passed down by oral history long before the invention of writing. Support for this can be found in the John Gardner and John Maier translation of Gilgamesh, where we are informed that the Gilgamesh "Creation Epic" was recited in Babylon each year during the "New Year's Festival," long before the creation of writing.

METAMORPHOSIS

From Gods to a Unifying Water Philosophy

More than 2500 years ago, in the Greek territory of Ionia, lived a man named Thales. Today, in philosophical circles, Thales is often described as "the world's first philosopher" and, according to Aristotle, "the founder of Ionian natural philosophy." For it is in Ionia where we find history's first naturalists recorded.

Thales, born in the rural Ionian town of Miletus, is renowned for his transformation of many of humankind's belief systems. The profound philosophical concept that Thales delivered to humankind was simply this—water is the first cause of all things seen and unseen.

Before Thales advanced his water philosophy, human thought was captured by the belief in polytheism. Under the concept of polytheism, people believed that all natural and imagined phenomena were created by a plethora of gods.

Besides contributing his water philosophy to Greek thought, Thales introduced other unique concepts to Greek thought. In fact, Thales' contributions were so profound and respected that during his time he was canonized as the wisest member of the Seven Sages of Greece.

Most of what we know about Thales flows from the writings of Plato, Aristotle, Plutarch, and other Greek philosophers. Thales was an original thinker who possessed a keen understanding of the myths and oral histories from hundreds of thousands of years in the past. At the same time, he was a scientist, astronomer, mathematician, teacher, author, politician, and a "seer" who could forecast future events.

For instance, Thales forecast a solar eclipse in the year 585 B.C. Modern scientists have checked on this event and confirm that a solar eclipse did take place on May 28, 585 B.C. and that it was visible in Ionia. How Thales was able to predict this event remains a mystery. According to information available about Thales' time in history, no mathematical formula existed for predicting an eclipse, and his feat is all the more impressive because Thales had to take into consideration the geographical latitude of Ionia.

Thales was born into a world filled with the so-called ancient mythologies and gods that were recorded in the first writings of humankind. By closely observing water and the natural world around him, Thales was able to synthesize his new philosophy into a form of natural science. This natural science helped to define the creation of life and humanity's connection to life and all creation through the entity of water.

Desiring to grasp the "first cause" of all things, Thales searched for a unifying hypothesis for the creation and sustenance of things both visible and invisible. This led him to the startling revelation that all things come from water, that the earth floats on water, that water is the constituent of all things, and that water is the cause of earthquakes. According to Plutarch's Miscellanies, Vol. III,

> Thales the Milesian doth affirm that water is the principle whence all things in the universe spring.... He pronounced, that all things had their original from water, and into water all things are resolved. His first reason was, that whatsoever was the prolific seed in all animals was a principle, and that is moist; so it is probable that all things receive their original from humidity. His second reason was, that all plants are nourished and fructified by that thing which is moist, of which being deprived they wither away. Thirdly, that the fire of which the sun and stars are made is nourished by water exhalations,—yea, and the world itself.

Today, we know that continental plates float and crash into each other and that there are chemical and other changes in groundwater prior to an earthquake. We also know from NASA planetary scientists that water was present in our part of the Universe before the creation of our solar system.

As far as Thales' theory "... that the fire of which the sun and stars are made is nourished by water exhalations,—yea, and the world itself," modern science has corroborated this fact. Recently, scientists discovered a direct relationship between the intensity and frequency of sunspot activity and weather patterns on the earth. Scientists also learned that sunspots contain water. As reported by Reuters on May 26, 1995,

> In research being released today in the journal *Science*, the experts report finding water not in the gaseous heart of the sun itself, but on somewhat cooler sunspots.
>
> "It's really sort of interesting," said Peter Bernath, a chemist who was part of the research team from the National Optical Astronomy Observatories in Tucson, Arizona, and Canada's University of Waterloo. "One is surprised that water can exist on the sun because it is so hot."
>
> Hydrogen and oxygen molecules both exist on the Sun, which is about 10,200 °F. They apparently combine to form water in sunspots, which are about 5200 °F. Because of the heat, the water molecules form vapor or steam, instead of pools or puddles.

Further scientific support of Thales' water theory was revealed by research data provided by a specially designed satellite called the Submillimeter-Wave Astronomy Satellite (SWAS). This satellite was launched in December 1998 to detect radio waves emitted by water molecules in space.

In an interview with the author of this article, Gary Melnick of the Harvard–Smithsonian Astrophysical Observatory in Cambridge provided several insights as to how modern science is providing support for Thales' water philosophy. Melnick is the leading scientist for the SWAS project. "We're seeing water everywhere in the universe. Every region we've looked at so far contains water," Melnick said.

Melnick explained that water found in interstellar space is crucial to the formation of stars. According to Melnick, when a big gas cloud begins to coalesce (form into a mass) under its own gravity, the gases begin to heat up. However, as the gases heat up, they begin to expand naturally. If water was not present to mitigate the heat generated by the condensing gas mixture, it would be difficult for the star to form. "It is because of this phenomenon," said Melnick, "that we see water helping to give birth to stars." Melnick also explained that one of the strangest discoveries of the SWAS project was finding a giant, water producing gas cloud in the heart of the Orion nebula. Based on SWAS data, Orion forms enough water every 24 minutes to fill all of the earth's oceans.

It is from our evolving scientific information, such as that expressed above, that we find corroboration of Thales' statements from more than 2500 years ago. However, along with our evolving water science and philosophies, there is also the wide practice of using water for ceremonial purposes. The use of water for the baptismal ceremony is very ancient and touches on the spiritual dimension of water.

Baptism by water

One of the most ancient and still widely used water rituals is that of baptism. From the second century forward, the value of the baptismal act was widely recognized and practiced in Christianity. Even though there are many meanings and beliefs relative to baptism, as practiced by Catholics, Protestants, and many others, water is always respected for its power to cleanse the body and regenerate the spirit.

For instance, the relationship of Moses to water harks back to his earliest beginnings. Besides his sprinkling of the covenant upon the people, we learn that the name Moses means "the water drawer." Also, not to be overlooked is Moses' origin from the Nile River. Exodus 2:1–10 tells us that as a baby, Moses was discovered floating in a basket by the Pharaoh's daughter. The word "tebah," used in the Old Testament for the basket carrying Moses, is also the same word used to describe Noah's Ark.

Besides Moses, there are other historic personages whose origins spring from water. The god Akki, "the water carrier," rescued Sargon, the founder of Babylon, from water. Karna, the icon of an early Indian epic, was found floating in the Ganges. And, just as there are many renowned people whose origins are linked with water, others find their way through various forms of water baptism.

The early "mystery" religions (Eleusinian, Orphic, Mithraic, Egyptian, Syrian) required baptism and other water rituals to wash away evil in preparation for the rites of communion with the god whose mastery of death assured immortality. These baptisms sometimes symbolized dying of the old life and rising again to the new. A similar idea underlies the baptism by immersion that is part of the initiation of proselytes to Judaism.

Today, many ancient water rituals can be found in almost all the world's religions.

Leonardo da Vinci (1452–1519)

About 500 years ago, Leonardo da Vinci was recording his thoughts about vortical energy as a result of his experiments with water. The subjects of water, hydrology, and hydraulics played a major part in Leonardo's lifetime studies. This fact was established when it was discovered that the largest number of Leonardo's notes were devoted to the subject of water than to any other subject.

The existence of water vortices intrigued Leonardo and prompted him to speculate about the existence and behavior of vortices in the air and in the cosmos. The power and meaning of vortices in water also led Leonardo to study closely the behavior of water under different conditions. In one of his experiments, Leonardo created elongated boxes with glass sidewalls so he could observe the motion of flowing water beneath the surface.

Based on his water experiments, Leonardo invented a double-gated dam. This dam used the pressure of water to make a watertight seal when the two gates met in the center with a projecting bias toward the high water side. This same principle is in evidence today at the Panama Canal and other places around the world.

Besides engineering and scientific inquiries into water, Leonardo's fascination with water dynamics and vortices is apparent in his art. The flowing motions of water vortices are expressed in his brush strokes and are obvious in his paintings, especially when depicting the hair of his subjects. Examples of this can be seen in his paintings "Leda, with the Swan;" the "Mona Lisa," and his "Self-Portrait" (see Fig. 1).

In the Mona Lisa, the hair emerges from the top of her head and flows downward like a fountain of water flowing from a spring, the breast area of her garment appears like a plunging waterfall, the flowing cloth over her left shoulder demonstrates the twisting vortex of water, and the dress sleeves on her arms appear as cascading ripples of darkness and reflected light, much like that of water overflowing from the base of a waterfall.

In Leonardo's "Self-Portrait," again we have the flow of hair springing from the head area. However, in this painting, the wavy flow of his head hair merges in dramatic fashion with his beard.

The existence of bubbles in water was also of interest to Leonardo; he noted from close observations that bubbles rise through water in a spiral motion. His written notes, as expressed in Leonardo's Codex Leicester Folio 23V (now owned by Bill Gates), depict how Leonardo documented that the "... motions of waters always move in a circle from surface to bottom."

Leonardo spent many years in his makeshift laboratory and in the field observing the movements of water and air. To see the fluid dynamics of water at work, Leonardo did experiments using glass so he could watch the motion of flowing water under various scenarios. During his field research, he maintained detailed notes and drawings about his experiences and observations. At times, Leonardo's mind would seemingly take off into other dimensions as his observations triggered ideas. Often, he would jot down or sketch these thoughts along the margins of his papers.

One such series of notes in the upper right hand corner of one of his papers was the outline for his proposed treatise on water. This treatise was divided into fifteen books, each book dealing with a different aspect of water. The classification of these books, as recorded by Leonardo, is as follows:

1. Of Water in Itself
2. Of The Sea
3. Of the Veins
4. Of Rivers

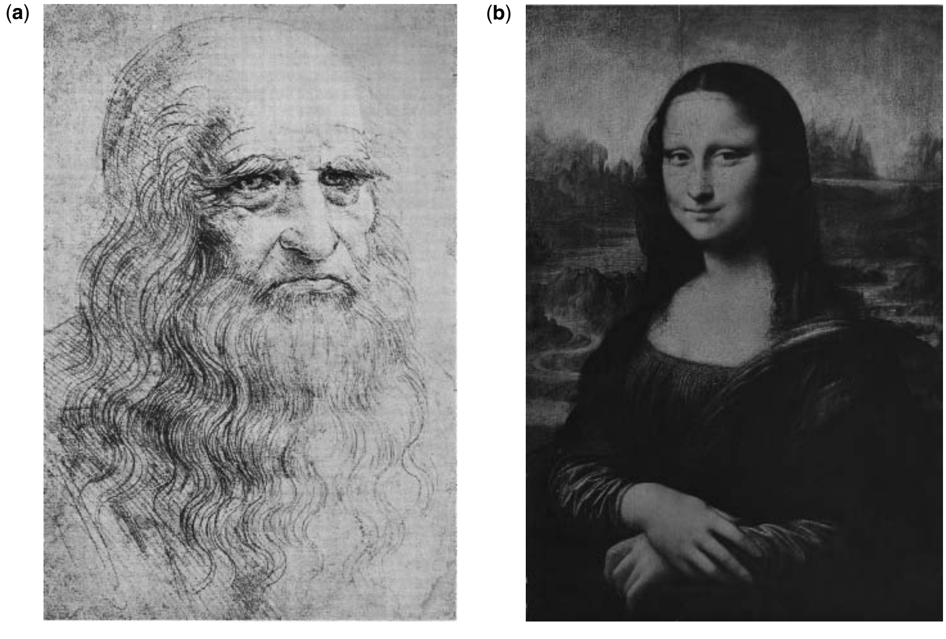

Figure 1. (**a**) Leonardo's "Self-Portrait," demonstrates water-like flow of hair springing from his head and merging in unifying waves with his beard. (**b**) Portrait of the Mona Lisa—depicts flow dynamics of water as her hair flows from her head, then cascades like a waterfall over her breasts down onto her arms, then continues flowing over her arms. Since da Vinci was careful in creating his compositions, it is important to note the mountains and "headwaters" in the background behind Mona Lisa's head, and the curving river meanders in the valleys behind both her shoulders that flow from left to right—in the same direction as the focus of Mona Lisa's eyes.

5. Of the Nature of Bottoms
6. Of Objects
7. Of Various Kinds of Gravel
8. Of the Surface of Water
9. Of Things Moving in It
10. Of River Repairs
11. Of Conduits
12. Of Canals
13. Of Machines Turned by Water
14. Of Raising Water
15. Of Things Worn Away by Water

Given the fullness of the life which Leonardo lived, he never found the time to complete the above books on water. His writings, especially those found in the Codex Leicester, contain many references and brief notes to be included in these books. An example of this can be found in Leonardo's notes dealing with precipitation:

> Write how clouds are formed and how they dissolve, and what it is that causes vapour to rise from the water of the earth into the air, and the cause of mists and of the air becoming thickened, and why it appears more blue or less blue at one time than another. Write in the same way of the regions of the air and the cause of snow and hail, and how water contracts and becomes hard in the form of ice, and of the new shapes that the snow forms in the air . . .

Another of Leonardo's field observations dealt with the physical expression of the vortex principle on a grand scale. His written observations about a water spout he saw along the seashore were recorded on a page entitled "Of wind twists and eddies involving water." In Leonardo's words,

> It often happens that, when one wind meets another at an obtuse angle, these two winds circle around together and twine themselves into the shape of a huge column, and becoming thus condensed, the air acquires weight. I once saw such winds, raging around together, produce a hollow in the sand of the seashore as deep as the height of a man, removing from it stones of considerable size, and carrying sand and seaweed though the air for the space of a mile and dropping them in the water, whirling them around and transforming them into a dense column, which formed dark thick clouds as its upper extremity. . ..

In spite of Leonardo's research and interest into various water phenomena, he was never able to provide a viable theory about the hydrologic cycle. The prevailing thinking in Leonardo's time was that the weight of the oceans waters forced water into the center of the earth. From the center of the earth, the water then flowed underground to where it eventually surfaced on land to appear in lakes, ponds, rivers, and swamps.

DISCOVERING WATER

Hydrologic Cycle

An important turning point in the correct human grasp of the hydrologic cycle came in 1580 with the publication of Bernard Palissy's book, *Discours admirables*. Palissy, a potter by trade, was a keen observer of nature and the workings of water. In *Discours admirables*, Palissy presented a concept of the hydrologic cycle that is accurate according to today's definition. Palissy's book also provided other new hydrologic concepts such as the origin of springs and rivers, how wells are recharged from nearby rivers, the hydrodynamics of artesian wells, the use of forestation for controlling erosion, and plans for constructing fountains for domestic water supply.

It was not until the late 1600s, however, that Edmund Halley, of Halley's Comet fame, actually proved that Palissy's theory was correct. Using his sense of observation and mathematical skills, Halley roughly figured the quantity of water draining from all the rivers of North Africa and Europe into the Mediterranean. From this, Halley determined that the combined flow was about equal to the water falling as rain or snow on the region's drainage area.

$H_2 + O$ = Water

In the late 1700s, Henry Cavendish, a chemist and physicist, experimented with sending electricity through water inside a sealed glass tube. To his surprise, the water disappeared. Many experiments later, Cavendish determined that water is made of two gases. Setting a match to different combinations of these two gases, Cavendish discovered that when he combined two parts of the "inflamable air" with one part of the "vital air," there was an explosion that left a residue of water drops inside the wall of the container. Because of these experiments, humans became aware of how the molecular structure of known matter could be split apart and recombined. Ironically, this experiment opened the door leading to the creation of the atom bomb.

Building upon the work of Cavendish, Antoine Lavoisier renamed "inflammable air" hydrogen—"the water producer"—and "vital air" as oxygen.

However, it was not until 1805 that Louis Joseph Gay-Lussac and Alexander von Humboldt proved that water was made up of two parts hydrogen and one part oxygen. This breakthrough proved that water is neither elemental (in the ancient use of the term), nor an element of chemistry, but a true compound unto itself. This new designation of water set it apart from air and earth (which are mixtures), and fire, which is a chemical process.

FIRST THERMOMETER

Before the invention of the thermometer, there was no way for people to qualify or quantify changes in temperature. The only way people could relate to changes in temperature was to comment how hot or cold something was. Comments like, "It's really a hot day today," or, "Brrrrr, the water is so cold, it gives me chills," were the only means of measuring temperature before the invention of the thermometer.

Galileo, along with others before his time, made the common sense observation that substances expanded when exposed to heat and contracted when chilled. In 1603, Galileo invented what is considered the world's first "thermometer" (Greek—"heat measure"). He did this by placing a heated glass tube upside down in a bowl of room-temperature water. As the hot air in the glass tube contracted, it pulled a column of water up into the tube. Through observation, Galileo made notes about how changes in room temperature caused the column of water in the tube to go up and down. This was one of the very first scientific instruments using glass, other than Leonardo da Vinci's earlier glass-sided boxes that were used for studying the vortical energy of flowing water.

In the mid-1600s, Ferdinand II, Grand Duke of Tuscany, discovered that placing water inside a round bulb with a narrow tube attached that was open at the top, provided a more sensitive way for observing temperature change. This allowed more accurate observations of water's response to temperature under different conditions, changes that, compared with Galileo's first thermometer, were more accurate because there was less influence from atmospheric air pressure.

Eventually, mercury replaced water or alcohol as the liquid of choice in thermometers. Mercury proved more stable because it does not freeze or evaporate. However, it was not until 1714, more than 100 years following Galileo's first thermometer, that Gabriel Daniel Fahrenheit invented a sealed mercury thermometer in combination with a graduated scale. History remains unclear how Fahrenheit selected the scale with which we are now familiar, a scale where water freezes at 32 °F, and boils at 212 °F. As is obvious, the study of water's response to temperature change determined the points on the thermometer. Determining the point on the scale where water turned to ice was considered a notable temperature event, as was the boiling point where water changed into a gas.

In 1742, 28 years after Fahrenheit's creation of a scaled thermometer, Anders Celsius of Sweden, fixed the freezing point of water at 0° and the boiling point at 100°. At first, this was recognized as the centigrade scale because of the 100-degree range where water remained in liquid form. However, in 1948 the centigrade scale became officially recognized by scientists as the "Celsius scale." This was to pay homage to the inventor and to be consistent with the naming of the Fahrenheit scale after its inventor.

The invention of the thermometer opened a whole new era of scientific and medical observation and inquiry. For instance, it was determined that the average human body temperature is 98.6 °F. A degree or two above or below this reading indicated of potential health problems.

Of course, the invention of the thermometer also gave humanity something more to talk about in its everyday life. Instead of just saying, "It's a really a hot day today," we can now say, "My goodness, it's over 90° today!"

SEAWATER BLOOD

About 100 years ago, the French scientist Rene' Quinton established the chemical similarities between seawater and blood plasma. This revelation created considerable media attention and opened the minds of many people to the possible origins of humankind and other life forms.

Today, science and medicine inform us that about 90% of blood plasma is water. This seawater-like blood water distributes necessary nutrients and oxygen to every cell in the human body through a network of approximately 60,000 miles of arteries and veins.

Because water is a universal solvent, seawater contains all elements known to humankind. However, scientists have determined that seawater contains only 11 elements in concentrations greater than 1/100%. Not surprisingly, these 11 elements are present in the human body at the same concentrations found in seawater.

OUR EVOLVING WATER CONSCIOUSNESS

In 1530, Nicolas Copernicus published his pivotal work, *De Revolutionibus*. This work theorized that the earth rotated on its axis every 24 hours and traveled around the Sun every 12 months, certainly, a startling theory for its time.

Today, there is an evolving water theory that may be as startling as Copernicus' discovery. This theory states that water is a conscious entity that responds to human thought and thereby creates our living reality. By doing so, water manifests the world within and around us in direct response to our thoughts and emotions. Think positive and loving thoughts—water responds by creating a healthy environment within and without. Think negative and destructive thoughts—water responds by creating a negative and destructive environment.

In her book, *The Entangled Universe*, Mae Wan Ho, Ph.D., establishes the recent discovery that we humans, as well as "... all living organisms are liquid crystalline." Because of our liquid crystalline makeup, we and other life forms are sensitive to various forms of energy, including the energy waves from sound, thoughts, and emotions. In this same book, Mae Wan Ho also writes:

> The visible body just happens to be where wave function of the organism is most dense. Invisible quantum waves are spreading out from each of us and permeating into all other organisms. At the same time, each of us has the waves of every other organism entangled within our own makeup.... We are participants in the creation drama that is constantly unfolding. We are constantly co-creating and re-creating ourselves and other organisms in the universe, shaping our common futures, making dreams come true, and realizing our potentials and ideals.

In other words, we as liquid crystalline beings, are the creators of our living reality, the same reality that responds to our thoughts and emotions and, in return, influences our very own existence.

Crystals in Water

The fact that "all living organisms are liquid crystalline" and all living organisms contain water is consistent with the fact that water naturally contains crystals. The presence of crystals in water becomes obvious when water is frozen. Snowflakes are another form of frozen water crystals. And, just as no two human faces are exactly alike, the same is true for snowflakes and other forms of frozen water crystals.

In the book, *Messages from Water*, Masaru Emoto provides scientific proof from his research results and photographs, as to how the crystalline structures of various waters respond to environmental factors, pictures, thoughts, words, music, and other influences. Emoto's research shows how water is a sensitive, lifelike entity, that is responsive to a spectrum of human and other energies. For example, Emoto would take a sample of distilled water and freeze it to obtain a baseline of what the frozen crystals looked like. Emoto would then expose samples of the same distilled water to different influences, such as certain thoughts, music, and written words. He would then freeze the water from the samples to see what changes, if any, occurred to the water's crystals. For instance, the projection of loving and caring thoughts into a water sample resulted in fully formed hexagonal crystals, whereas the projection of negative and violent thoughts into a sample of the same water produced incomplete or malformed crystals.

A FUTURE BASED ON WATER

Due to present water shortages and dire forecasts, the subject of water will soon permeate the minds of all levels of people around the world. The building of dams, such as the Three Gorges project in China, the large-scale desalination of seawater in the Middle East, the overpumping of the world's groundwater aquifers, the proposed transoceanic transport of Canadian water via huge floating water bags, and the competition amongst the world's most powerful transnationals to control sources of water are examples of the way humans are attempting to manipulate and distribute water. Simply stated, whoever controls water has political and financial power, as well as power over which life forms live or die.

However, beyond the physical management of water for profit and/or power, there are hints of other dimensions of water that may soon lead humanity to higher levels of consciousness. The fact that water responds to human thoughts and emotions is opening new possibilities to human evolution. In the near future, we may find that the projection of positive and loving human energies into water will deliver us into a healthier world filled with abundant food. If so, such a conscious progression in our evolution will give humanity the opportunity to help restore life to a planet that has been most supportive of our earlier stages of evolution.

This evolution of our water intelligence also includes our gaining a better understanding of the DNA molecule. For we are now discovering that the configuration of DNA is based on hydrogen bonding—the same as water. DNA is the basis for passing along genetic codes, and the H-bond of DNA is intimately connected with surrounding water clusters, so we may find a new basis for understanding genetic intelligence.

Perhaps, in the final analysis, the evolutionary progression of humanity will read as follows: Stone Age; Bronze Age; Iron Age; Industrial Age; Space Age; Computer Age, and then finally, Age of Water.

Whether or not humanity survives the Age of Water, in my opinion, will be determined by humanity's ability to work consciously with the living values and energies existing within water.

EFFECTIVE WATER EDUCATION STRATEGIES IN A NONTRADITIONAL SETTING

Sharon L. Lien
The Groundwater Foundation
Lincoln, Nebraska

INTRODUCTION

The Groundwater Foundation (TGF) is a nonprofit organization whose mission is educating people to care about and for groundwater and encouraging groundwater stewardship in all citizens. TGF's belief, which has come out of 20 years of experience, is that once citizens understand the importance of groundwater, they act to protect it.

Because of this, TGF puts particular emphasis on education and programming for the general public, especially youth. However, these educational programs have not been in the formal classroom setting, but in nontraditional settings, such as a festival environment, in competitions, or in the field.

This article discusses three specific examples of effective water education in nontraditional settings—Children's Groundwater Festival, Groundwater University, and Awesome Aquifers for Science Olympiad—and the impact these efforts have had on educating youth on groundwater protection and conservation.

Children's Groundwater Festival

From its inception in 1989, the Children's Groundwater Festival has been a success. Formulated around the concept of making learning fun, the Festival is staged in a fair-like atmosphere that is stimulating and exciting for participants, very unlike traditional classroom education. Whether witnessing the water cycle or building a well in a cup, students learn about every aspect of water and the environment in a fun way that helps them retain concepts and change their behaviors.

The Festival was designed as an annual 1-day event for fourth- and fifth-grade students of Nebraska. It is a dynamic, energetic day filled with hands-on education. Students from all over the state travel to Grand Island to learn about groundwater and related resources through interactive presentations given by groundwater and natural resources professionals who represent government agencies, environmental organizations, higher education, private business, and interested citizens. At the Festival, students have access to many water-related, experiential activities. There is nowhere else that children can have access to more environmental experts in such a fun and energizing atmosphere. This is a field trip like no other!

What makes the festival concept so effective is the unique and innovative methods used to teach the students about water. They learn what an aquifer is by making one with ice cream, discover how water is delivered to their home by becoming a human water distribution system, and gain an appreciation of the value of water by operating a working model of a treatment plant. Students don't just listen; they experiment, create, participate, react and learn, empowering them to make a difference.

The Children's Groundwater Festival has become an international model of hands-on environmental education. Over the years, water educational festivals have sprung up across the United States and around the world. To meet the demand for specialized training for festival organizers, TGF has published several books and provides workshops on organizing a water festival.

Groundwater University

Groundwater University (GU) was created in 1994 to provide intensive, experience-based groundwater education for youth and to encourage them to pursue science and water-related studies or careers. This program, directed toward junior and senior high school students, provides 3 days of hands-on activities, field trips, and presentations from top water professionals. GU offers participants the opportunity to study groundwater in a summer camp setting. Student activities include

- constructing mini groundwater models
- hiking the Platte River and digging down to the water table
- Participating in field trips to monitor wells, streams, and lakes to collect and test water samples
- investigating the relationship between microorganisms and water quality under a microscope
- building leadership and group skills through individual and team projects on topics of interest

As part of this program, students have the opportunity to share information about many groundwater issues that are unique to their communities. Because GU faculty includes professional hydrologists, chemists, irrigation experts, well drillers, and geologists from government agencies, private water testing laboratories, universities and colleges, and municipal water systems, students gain insight into water and science careers.

All activities at GU are hands-on and require participation. Some activities are done by teams, others are individual. Students take a posttest, and results are compared with those on their pretests to measure the change in knowledge.

Several GU graduates have continued their groundwater-related activities. Through GU, TGF educates students about groundwater helps them understand the connection between safe water supplies and public health, and ignites in them an interest in pursuing science and water-related studies and careers.

Awesome Aquifers for Science Olympiad

Science Olympiad (SO) is an international program devoted to improving the quality of science education, increasing student interest in science, and providing recognition for outstanding achievement in science education by both students and teachers. Student teams in participating middle and high schools all across the United States learn about different scientific topics; build models for some events; and compete at the local, state, and national levels in dozens of different SO events. The program emphasizes science learning, participation, interaction, building team spirit, and having fun. More than 2 million students are involved annually in SO.

Because an event on groundwater did not exist in the SO program, TGF created Awesome Aquifers. This event is designed to increase student understanding of groundwater concepts. Prior to the tournament, students conduct groundwater research and design and build a groundwater flow model according to SO rules that are standardized among participating schools. Students are required to demonstrate certain concepts, but there are almost no restrictions on how they build their models, so the event encourages student creativity.

During competition, students use their models to demonstrate an understanding of the hydrologic cycle, the physical makeup of an aquifer, and changes in the groundwater system. Students can earn bonus points for demonstrating possible contamination remediation techniques, such as using an intervening well to intercept a contaminant plume, chemical or physical treatment, or pump and aeration treatment.

Awesome Aquifers provides an opportunity to teach some of the nation's brightest young scientists about a resource that is critical to human and environmental health at a time when earth science programs are being reduced in many schools. Other objectives for the event are to give students a chance to learn about groundwater-related careers from working scientists who serve as judges or mentors for the event, and to educate two other groups who often know little about groundwater: science teachers, who generally serve as Science Olympiad coaches, and parents, who are often quite involved with student teams.

IMPACT ON EDUCATING YOUTH

In 1994 and again in 2002, TGF invested the time and resources to determine the effectiveness of the Children's Groundwater Festival and Groundwater University in impacting students' knowledge and behaviors positively toward groundwater. Results from this third-party behavior impact study indicated:

- The formats of these programs are especially well geared to the desired learning styles and approaches of students, and students have few suggestions for ways to improve either the Festival or Groundwater University. Students who have participated in these programs are likely to remember them for years, and up to 6 years later are able to relate what they learned during the event.
- The Festival did have a definable impact on students both in terms of gain in knowledge and change in behavior. After students attended the Festival, they exhibited an increase in knowledge as demonstrated through pretest and posttest scores. They also demonstrated a willingness to change their

behaviors, and in some cases actually adopted water conservation and protective actions.

- For Groundwater University students, learning goes far beyond groundwater education; it encompasses the "whole child." Students stated that they attended to learn about groundwater, but equally important for them was the opportunity to learn socialization skills, which they considered just as valuable. Almost all students noted that they appreciated learning about the cultural differences among themselves and their peers from other areas of the country. Many stated that they appreciated opportunities simply to interact with "experts" in the field (noted speakers, college professors, groundwater professionals, etc.) who had been invited to lead parts of the program.
- Students see the connections between groundwater and their own academic and professional interests. Though much of the information presented at Groundwater University is science-related, students can quickly perceive the connection of groundwater to other areas of interest and study, including writing, art, education, etc. Most importantly, the study determined that Groundwater University translates into behavioral change for most of the students (1).

Evaluations conducted by TGF have shown that Awesome Aquifers is also an effective activity to educate students on groundwater concepts. Based on information provided by students who have participated in the event,

- 97 percent of the students who participated liked the event.
- 89 percent indicated that they would participate again.
- 87 percent indicated that the event influenced them to conserve or protect groundwater.

Additionally, students overwhelmingly stated that they learned about the importance of groundwater, the complexity of an aquifer, and that groundwater is an important water source. Comments from students about what they learned included

- "This event helped me learn about the connection between pollution and the environment."
- "I learned that it is very hard to remove pollutants from an aquifer."
- "I learned more about the water cycle and how pollution affects it."
- "What an aquifer is, how it works, and what can we do to help."
- "A lot more about groundwater and the wonderful resource it is!"
- "I learned more than I thought possible and loved every bit of it."
- "We learned a plethora of information about conservation and protection."
- "Human activities affect the groundwater system hugely."
- "I learned how complex the underground water system is."
- "How waste can affect the water."
- "I learned a lot about the groundwater system and had an experience in actually seeing how a real aquifer works."
- "All about wells and the art of well construction. It also taught me how to protect and conserve groundwater."
- "We learned about groundwater and land water systems, and their relationships."

Real changes usually emerge from educational strategies that give the student a sense of involvement and ownership. Hands-on activities are a reliable approach for this as is student-directed learning(2). Thinking out of the box, or in this case out of the classroom, can allow water educators to explore less traditional educational strategies. By building creative and diverse programs that engage and empower youth and measuring the results of these efforts, the goal of protecting our groundwater supplies now and for future generations can be achieved.

BIBLIOGRAPHY

1. Marsters, M. (2002). Follow-up Behavioral Impact Study on Groundwater Foundation Student Education Programs.
2. Coyle, K. (2004). *Understanding Environmental Literacy in America: And Making it a Reality*. National Environmental Education and Training Foundation, Washington, DC.

EVOLUTION

David Haley
Picacadilly
Ulverston, Cumbria, United
Kingdom

> "Those stars are the fleshed forebears
> Of these dark hills, bowed like labourers,
> And of my blood." (1)

In the cosmic sense of evolution, water may be viewed as part of the design of a supernatural "creator," or a quirk of fate—that a certain condition occurred for a combination of atoms (H_2O) to coincide in a particular atmosphere and remain for a sustained period of time within the atmosphere of a planet. And as Theodore Schwenk writes in his analysis of flow forms, *Sensitive Chaos*: "It is not possible to speak of the beginning or end of a circulatory system; everything is inwardly connected and reciprocally related. Water is essentially the element of circulatory systems" (2). He continues: "The great rhythms of the tides are a response to forces that work in the interplay of earth and cosmos…and for which, through its great impressionability, the element of water is a receptive 'sense organ'" (3). Governed by the Sun and Moon, this earth-bound phenomenon is claimed by creation myths and scientific conjecture and is the source of inspiration for poets and artists.

> Fallen from heaven, lies across
> The lap of his mother, broken by world.
>
> But water will go on
> Issuing from heaven
>
> In dumbness uttering spirit brightness
> Through its broken mouth.
>
> Scattered in a million pieces and buried
> Its dry tombs will split, at a sign in the sky,
>
> At a rendering of veils.
> It will rise, in a time after times,
>
> After swallowing death and the pit
> It will return stainless
>
> For the delivery of this world.
> So the river is a god
>
> Knee-deep among reeds, watching men,
> Or hung by the heels down the door of a dam
>
> It is a god, and inviolable.
> Immortal. And will wash itself of all deaths (4).

Some cultures and communities believe in forms of divine creation with regard to evolution, but the understanding of much Western science continues to be based on a synthesis of Darwin's natural variation and selection ideas and Mendel's model of genetic inheritance. Central to this theory, known as the Neo-Darwinian synthesis, is the idea that natural selection is based on the transfer of genes through reproduction.

New knowledge of genetics and other evolutionary factors continue to fuel scientific debate, providing different interpretations with regard to the timescales and population patterns of evolutionary change and development. However, what seems to be emerging is that "... it may in general be inappropriate to use a theory of any kind as an ultimate framework for life as a whole" (5).

Lakoff and Johnson explore these concepts in their book, *Philosophy in the Flesh*: "In evolutionary theory, survival is keyed to the ability to fit ecological niches. Fitting a niche and hence surviving can occur for many reasons: coloration that hides one from predators, a large number of offspring, the availability of food ..." (6), and to these we may add the availability of freshwater and the ability to adapt to climate change. They then examine Western culture's understanding of evolution: "There is a common folk theory of evolution, that evolution is a competitive struggle to survive and reproduce. The folk theory has normative implications: Competitive struggle to survive and reproduce is natural. Moreover, it is good, because it got us where we are" (7).

In *The Web of Life: A New Synthesis of Matter and Mind*, Fritjof Capra uses Maturana and Varela's theory of autopoiesis and Lovelock and Margulis' Gaia theory to explore the unfolding of life differently from Darwinian theories. Capra sees evolution as a process of biological and environmental co-evolution—"an ongoing dance that proceeds through a subtle interplay of competition and cooperation, creation, and mutual adaptation" (8). Other developmental strategies are noted in the recombination of DNA among bacteria; Lynn Margulis' theory of "symbiogenesis" recognizes the vital importance of cooperation in evolutionary processes.

Another pattern of development to consider is known as convergence, when different organisms meet similar challenges by evolving in similar ways (e.g., the development of eyes in diverse species). Convergence has also come into play during periods of intense growth and innovation following planetary catastrophes. "Thus the disastrous depletion of hydrogen in the earth's atmosphere over two billion years ago led to one of the greatest evolutionary innovations, the use of water in photosynthesis" (9). However, this led to a catastrophic accumulation of oxygen, which in turn led to the evolution of oxygen-breathing bacteria and that recombined with the hydrogen that was escaping the atmosphere to form water, "...thus keeping the planet moist and preventing its oceans from escaping" (10). Other interactive living processes, like fermentation, also contribute to the balancing of the earth's systems, known as Gaia: "Gaia is just symbiosis as seen from space: all organisms are touching because all are bathed in the same air and the same flowing water" (11).

"The Russian mineralogist Vladimir Vernadsky (1863–1945) recognized life as a great geologic force. Anticipating Hypersea, he called living matter "animated water." Animated water is an excellent description of life" (12).

What is certain is that there is no life without water. Animals evolved from aquatic microorganisms about 700 million years ago and plants followed around 200 million years later. Then, between 400 and 450 million years ago, plants followed by animals started to leave water, to live on land. This major habitat transition called for major evolutionary innovations to cope with the absence of water. These included respiratory systems adapted to oxygen, skin that could cope with increased radiation from the Sun, and muscular and skeletal development to deal with increased gravity. To this day, however, all animals maintain their origins in the form of the

> ... womb that simulates the wetness, buoyancy and salinity of the ancient maritime environment. Moreover, the salt concentrations in the mammal blood and other bodily fluids are remarkably similar to those of the oceans. We came out of the ocean more than 400 million years ago, but we never completely left the seawater behind. We still find it in our blood, sweat and tears (13).

Capra brings the evolutionary story up to the present day:

> Far away from their African homeland, the early humans had to endure extremely harsh climatic conditions that had a strong impact on their further evolution. The entire evolutionary history of the human species, from the emergence of *Homo habilis* to the agricultural revolution almost two million years later, coincided with the famous ice ages (14).

CLIMATE CHANGE

It would seem that extreme and sudden changes in climatic conditions have given rise to great evolutionary innovations and further developments occur over relatively long periods of adaptation to prevailing conditions. However, as the dinosaurs found, not all species or groups of species survive such changes, and these concerns are among the important issues facing the Intergovernmental Panel on Climate Change (IPCC).

In recent history, an Irish scientist, John Tyndall (1820–1893), discovered the importance of water vapor in maintaining the earth's life-sustaining atmosphere.

> Tyndall's experiments also showed that molecules of water vapour, carbon dioxide and ozone are the best absorbers of heat radiation and that even in small quantities these gases absorb much more strongly than the atmosphere itself, a phenomenon of great meteorological importance. He concluded that among the constituents of the atmosphere, water vapour is the strongest absorber of radiant heat and is therefore the most important gas controlling the Earth's surface air temperature. He said that without water vapour the Earth's surface would be "held fast in the iron grip of frost." He later speculated how changes in water vapor and carbon dioxide could be related to climate change (15).

From the nineteenth century, Tyndall's grasp of the potential climatic problems facing the twentyfirst century are now the source of much speculation. In the publication Climate Change 2001: Synthesis Report, Summary for Policymakers, the IPCC was clear about certain consequences:

> Changes in sea level, snow cover, ice extent, and precipitation are consistent with a warming climate near the Earth's surface. Examples of these include a more active hydrological cycle with more heavy precipitation events and shifts in precipitation, widespread retreat of non-polar glaciers, increases in sea level and ocean-heat content, and decreases in snow cover and sea-ice extent and thickness. For instance, it is very likely that the twentieth century warming has contributed significantly to the observed sea-level rise, through thermal expansion of seawater and widespread loss of land ice (16).

A CONFLUENCE OF IDEAS

Uncertainty about how to model predictions of the future effects of global warming and climate change are still contentious issues among the world's top scientists. Different methods of selecting and presenting information produce different scenarios that favor opposing political, social, and commercial lobbying groups. Professor Peter Sammonds of University College London believes that "(t)here is a serious problem that the starting point for discussion on the environment has tended to not be considered scientific knowledge but alarmist interpretations of future climate set in a highly moralistic framework" (17). He ascribes most of the effects of climate change to naturally occurring phenomena, challenging the validity of available data and its misinterpretation. Sammonds criticizes the insurance industry for "(t)his talking up of the threat of global warming..." to meet their own ends and asserts that "... nothing should be done which would curtail global economic growth through misplaced strategies to curb carbon dioxide emission. Otherwise, we would be damaging our very ability to adapt to environmental changes."

However, the IPCC's 2001 Summary concludes,

> Significant progress has been made in ... many aspects of the knowledge required to understand climate change and the human response to it. However, there remain important areas where further work is required, in particular:

- The detection and attribution of climate change
- The understanding and prediction of regional changes in climate and climate extremes
- The quantification of climate change impacts at the global, regional, and local levels
- The analysis of adaptation and mitigation activities
- The integration of all aspects of the climate change issue into strategies for sustainable development
- Comprehensive and integrated investigations to support the judgement as to what constitutes dangerous anthropogenic interference with the climate system (18).

Writing from an explicitly "green" perspective, in her book, *Water Wars*, Vandana Shiva expresses her views on the relationship between water and climate change:

> Water is life, but too much or too little of it can become a threat to life. The stories of Noah and Vishnu Purana are tales of mythic floods that whipped out life on the planet. While floods and droughts have always occurred, they have become more intense and more frequent. These climatic extremes are linked to climate change, which is in turn linked to pollution of the atmosphere by fossil fuels ... The impact of climate crisis on all forms of life is mediated through water in the form of floods, cyclones, heat waves, and droughts (19).

This echoes the Renaissance concern for apocalyptic visions, and Leonardo da Vinci's (1452–1519) late drawings of the Deluge (Fig. 1) seem to express this Dantesque preoccupation—water becomes a frightening phenomenon in violent climatic conditions. But Leonardo's aesthetic was to envision these apparently catastrophic events as the spectacularly exciting stuff of life. This description illustrates his imagination and love of the dramatic, as much as his lifetime study of water that could equally find a lyrical likeness between a vortex in water and the curls of a young girl's hair:

> And the ruins of a mountain fall into the depth of a valley, creating a shore for the swollen waters of its river, which having already burst its banks, will rush onwards with monstrous waves, the greatest of which will strike and destroy the walls of the cities and farmhouses in the valley. And the ruins of the high buildings in these cities will throw up immense dust that rises like smoke or wreathed clouds against the falling rain. But the engorged waters will cascade around the pool that contains them, striking in eddying whirlpools

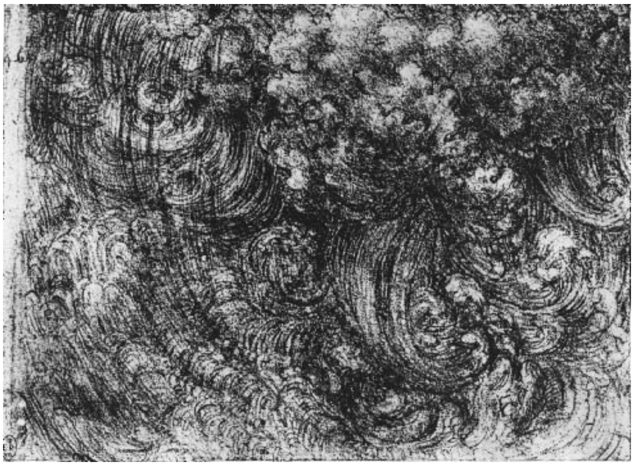

Figure 1. Leonardo da Vinci Deluge drawing No. 46 (146)—12383 Windsor collection.

> against different obstacles, throwing up muddy foam as they leap into the air before falling back and then again being dashed into the air. And the whirling waves that flow from the site of percussion are impelled across other eddies going in opposite direction, and afterward their recoil will be tossed up into the air, without dashing off from the surface. And at the outlet where the water flows from the pool, the spent waves will be seen spreading out toward the outlet, and afterward, falling or pouring through the air, and gaining weight and impetus movement, [this water] will hit the water below, pierce through it in a furious rush to reach bottom, and then being thrown back, the [water] returns to the surface of the lake, carrying with it the air that was submerged within it, and it, remains at the outlet, its foam mingled with logs of wood and other things lighter than water . . .The crests of the waves of the sea tumble down to their bases, whirling and falling with friction above the bubbles on their side (20).

Our evolutionary destiny and those of many other species depend on our ability to maintain adequate supplies of freshwater and adapt to the accelerating effects of climate change: “We must always bear in mind that the water we are using today is borrowed from future generations, who will require it for their survival” (21). Despite pollution, species depletion, and the uncertainties of the effects of global warming, the oceans still hold the greatest diversity of life on the planet and the potential for the continuum of evolution. As E. C. Pielou notes: “Living things depend on water, but water does not depend on living things. It has a life of its own” (22).

BIBLIOGRAPHY

1. Hughes, T. (1995). Fire-Eater. *Ted Hughes: New Selected Poems 1957–1994*. Faber and Faber, London, p. 30.
2. Schwenk, T. (1996) *Sensitive Chaos: The Creation of Flowing Forms in Water and Air*. Rudolf Steiner Press, London, p. 13.
3. Schwenk, T. (1996). Ibid.
4. Hughes, T. (1995). Ibid., *The River* p. 243.
5. Steen, F. (2001). *Evolutionary Theory: A Theory of Changes in Organic Design through Controlled Random Mutations and Contingent Selection*. Available: http://cogweb.ucla.edu/ep/Evolution.html.
6. Lakoff, G. and Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. Basic Books, New York, p. 560.
7. Lakoff, G. and Johnson, M. (1999). Ibid., p. 557.
8. Capra, F. (1996). *The Web of Life: A New Synthesis of Mind and Matter*. HarperCollins, London, p. 222.
9. Capra, F. (1996). Ibid., p. 227.
10. Capra, F. (1996). Ibid., p. 232.
11. Margulis, L. (1998). *The Symbiotic Planet: A New Look at Evolution*. Weidenfrld & Nicholson, London, p. 2.
12. Margulis, L., (1998). Ibid.
13. Capra, F. (1996). Ibid., p. 244.
14. Capra, F. (1996). Ibid., p. 253.
15. The Tyndall Centre web site, *A brief biography of John Tyndall, FRS, DCL, LLD*. Available: www.tyn@uea.ac.uk.
16. The International Panel on Climate Change. (2001). *Climate Change 2001: Synthesis Report, Summary for Policymakers—IPCC Third Assessment Report*, p. 6.
17. Sammonds, P. (2002). Extreme climates. In: *Nature's Revenge? Hurricanes, Floods and Climate Change for Institute of Ideas*. Hodder & Stoughton, London, p. 21.
18. Ibid. (2001). *The International Panel on Climate Change*. p. 33.
19. Shiva, V. (2002). *Water Wars: Privatization, Pollution and Profit*. Pluto Press, London, p. 40.
20. Bambach, C.C. (Ed.). (2003). BAMBACH, C. C. quoting DA VINCI, L., Winsor RL 12665r; Richter 1970, Vol. 1, pp. 355–356, no. 609, *Leonardo Da Vinci Master Draftsman*. The Metropolitan Museum of Art, New York, p. 630.
21. Kyoto Excutive Committee for the 3rd World Water Forum. (2003). *Kyoto Water Declaration*.
22. Pielou, E.C. (1998). *Fresh Water*. University of Chicago Press, Chicago, IL.

HISTORY OF POND FISHERIES IN POLAND

Janusz Guziur
University of Warmia and
Mazury in Olsztyn
Olsztyn, Poland

The origins of carp breeding in Europe and Poland are sunk in the obscurity of bygone times. In Europe, the first historic record of the carp as a cure for ailments is in the writings of St. Hildegard (1098–1170), the prioress of a Benedictine convent. Some time later, carp breeding in ponds was mentioned by the monk, Albertus Magnus (1230–1310), in the archives of the Kladbrubkski monastery in Bohemia (1115), and in an investiture act by King Przemyslaw Otakar II granted in 1227 in Moravia.

Several hypotheses have been suggested concerning the introduction of carp breeding in Poland. It is believed that two routes were taken to bring the carp from its native Black Sea watershed to Poland, but whichever way it traveled, the fish was brought to Poland by the clergy (bishoprics). The oldest monastic order that first engaged in farming and fish breeding was the Order of the Cistercians, which was brought to Poland in twelfth century from the province of Brabant (on the border of Belgium and France).

The domestic mirror carp is known to have been bred in Poland since the mid-thirteenth century. This breed is not found in the wild in Europe or Asia, unlike the wild river carp (sazan), which has local forms that live free, for example, in the Danube. Because of its inferior body exterior and growth rate, wild sazan differs distinctly from domesticated mirror carp and fine-scale carp.

In the 1960s and 1970s, attempts were made to 'refresh' the blood of farm carp by breeding Danube sazan in fish ponds or crossing it with breeding carp (1–3).

According to Balon (3), mirror carp, also known as royal carp (*rex cyprinorium*—Johnston 1650) originated from mutants of the Danube sazan, which Roman legionnaires brought to Rome, where it made its home and from where it was later taken to monasteries around west Europe (Fig. 1). The monasteries propagated the breeding of the cultured (domestic) carp, also in Poland; this was encouraged by frequent and lengthy religious fasts (up to 180 days a year) (3,4).

The first known written document mentioning carp in Poland, dating back to the year 1466, was penned by the famous chronicler of the Battle of Grunwald, Jan Długosz (1415–1480). Długosz (5) describes three carp in the coat of arms of the Silesian nobles Korbczok (*tres pisces, qui carpiones vocantur*), which was later confirmed by Paprocki's Armorial (1584), or two trouts (*duo pisces, qui trutae appelantur*) in the coat of arms of the family Musznik de Dombrova. The carp also appears in the accounting books of the Order of Teutonic Knights (Deutscherdrensritter) of 1399 (6). This evidence proves that the carp was widespread in Poland as early as the thirteenth and fourteenth centuries.

Based on books by several authors (4,6–12), four periods of carp breeding development in Poland can be distinguished:

First Period from the Mid-Thirteenth to the Mid-Seventeenth Centuries (To the End of the Swedish Wars)

This was the time of dynamic development and flourishing of pond fish farming (Fig. 2). The first printed books on fish farming came out. The first pond fish farming textbook in Europe was written by Jan Dubrawski (*Dubravius*) from Ołomuniec (13), who also described fish farms in Polish lands (Cieszyn Silesia). Soon to follow was the first Polish textbook on fish ponds by Olbrycht Strumieński (14) "O sprawie, sypaniu, wymierzaniu i rybieniu stawów," and then a book by Stroynowski (15), reprinted in 1863. In 1568, the Polish historian Marcin Kromer included some information on carp farming in his description of the Polish state (4).

Figure 1. Fishing in a monastery pond (3).

Pond fish farming in Poland, and especially in Upper Silesia, the Polish cradle of pond management, was considered the most modern and strongest in Europe, on a par with fish breeding in Bohemia and Moravia, in terms of modern technologies of constructing fish ponds and management methods (including restocking) as well as the volume of production. Restocking, which involved moving growing fish to other ponds, prepared and filled with water in advance, developed in Poland in the second half of the sixteenth century and became strongest in the following centers: Oświęcim-Zatorze, estates of the archbishops of Gniezno and Krakow, and in the valley of the Barycz river (Milicz). Noteworthy is the fact that the first information about restocking can be found earlier in the works of Strumieński (14) and partly Dubravius (13). In the early eighteenth century, when pond fish breeding in Europe was in decline, this method was forsaken and remained forgotten until the end of the nineteenth century (4,10,16).

The most extensive information on fish management in the fifteenth–seventeenth centuries is available from the region of Silesia, especially Opawa-Cieszyn, Pszczyna, and Racibórz Silesia, described by the Czech historians of carp breeding, including Dubravius (13) and Hurt (17). The area of Silesia had favorable climatic (warm summers), soil (clays with enclaves of loess), and especially water conditions (the sources of the large Polish rivers, the Oder,

Figure 2. Pond fisherman from Milicz region, seventeenth century (9).

Vistula, Olza, and Soła). Politically, Silesia was loosely tied to the Czech Kingdom, which allowed the region to enjoy great economic independence. The total estimated area of fish ponds in upper Silesia in the fifteenth–sixteenth centuries was about 25–30 thousand ha, which contained all categories of ponds used presently (12). Those ponds were used for breeding other fish species apart from the carp: the so-called black fish, including all fish that have a dark skin and scale color (crucian carp, burbot, tench, mud loach, wels) and white fish (common bream, roach, perch, ide, chub, ruffe, bleak, sunbleak) as well as other less valuable species (9). Both groups of fish equaled 15–20% of the total amount of carp, pike, and walleye fished from ponds. From the sixteenth century on, brown trout were farmed in the mountainous parts of Silesia (Cieszyn, Opawa). The fish was carried by highlanders (traveling on foot) over the mountains and then transported to the court in Vienna.

In those days in Poland, as in neighboring countries (Bohemia, Moravia, Hungary, and Bavaria), some nobles and dukes commissioned construction of huge fish ponds built as 'monuments to glory,' such as the ponds *Stary* (700 ha) and *Grabownica* (500 ha) near Wrocław and Milicz, *Kaliski* (625 ha) near Opole, *Bieruński* (500 ha) near Katowice, and *Sigmund August's* (480 ha) near Knyszyn (Białystok). The latter, also called '*Czechowizna*,' was built in 1560 by Czechs brought to Poland by Queen Bona (12,18). Many ponds and whole fish farms have survived and are still in use, for example, **Gołysz, Ochaby**, and **Landek** near Cieszyn; **Zator, Grojec, Osiek**, and **Czernichów** near Kraków; **Łyszkowice** near Łowicz; **Knyszyn** near Białystok; and **Przygodzice** and **Milicz** in the Barycz river valley, together with several smaller ponds in the valleys of the upper Oder and Vistula rivers. The high status that fish and fish pond farming attained then is confirmed by the emblems of many Polish towns that bear images of carp, pike, trout, and other fish.

Second Period from the Mid-Seventeenth Century to the 1860s

This was the time when fish pond breeding declined in all of Europe. Carp breeding became unprofitable and was nearly discontinued. In Poland, more than half of the existing ponds were liquidated (dried) and used for more profitable production of cereals, predominantly wheat (9,17). It was then, however, that Revised Krzysztof Kluk published his excellent work on farming (19), which remained popular until the mid-nineteenth century. The third volume of this book was devoted to fish and fish ponds (9,20,21).

Third Period from 1868 to 1939

This was another period of dynamic development, especially in modern methods of fish pond management. The impulse for this renewed development came from the area of Cieszyn Silesia, a fish pond 'reservoir,' located in the upper Vistula river watershed. It was initiated by the precursor of modern carp breeding methods, Tomas Dubisz (*Dubisch*). Tomasz Dubisz (1813–1888), a Slovak who was born near the Danube and lived in Vienna, was an illiterate fish breeder brought to Ruda Rożaniecka (1868) by Baron Wattman. Later, he worked as a master fisherman in Landek located on the land estates belonging to Cieszyn Manors (property of Archduke Albrecht Habsburg), which today forms Poland's largest experimental fish farm of the Polish Academy of Sciences in Gołysz, Cieszyn Silesia. The year 1869 marked an important event in carp pond breeding, when Tomasz Dubish, master fisherman at a complex of ponds in Iłownica (PAN Gołysz) introduced a new (or actually forgotten) carp breeding method of restocking. That is the system of moving 1-year-old fingerlings to a succession of ponds, prepared earlier and filled with water, reducing the density of stocking at the same time (10,12,22). The creation and implementation of this system made it possible to select the breeding material, shorten the breeding cycle from 4–5 to 2–3 years, and achieve dynamic growth in annual profit (Dubisz himself attained an increase in profit from 3000 to 40,000 zlotys per pond). Soon this system of fish farming was adopted all over Europe (16,18,21) and, with small modifications, is practically used until the present (11,23).

The breakthrough in carp breeding achieved by Tomasz Dubish was continued by his pupils and successors, including Adolf Gash (1839–1915) from Czechowice, a creator of a new breed of Polish carp (known then as the Galician carp) characterized by a strong back; Michał Naimski (1842–1915); Paweł Morcinek (1961–1938) from Skoczów, a known fish breeder active in Silesia and Moravia and also the first biographer of Tomasz Dubisz; Andrzej Gostkowski (1839–1889) from Wadowice; and later the fish breeder Oskar Rudziński (1850–1919) from Osiek near Oświęcim, together with his son, Edward Rudziński (1892–1980), the senior of Polish carp selection breeders in the years between the First and Second World Wars.

Owing to the system developed by Dubisz and the introduction of the new Polish carp breed, the productivity of ponds immediately rose by more than 30% (18). All these achievements by Polish fish farmers and breeders at the turn of the nineteenth and twentieth centuries guaranteed Polish fish farming a leading position in Europe.

The 20 Years Between World Wars I and II

This was the time of a dynamic increase in the area of ponds in Poland. In 1919, the Parliament of the Polish Republic passed an act on leasing unused farmland properties, and a year later (1920), on their liquidation and lease and management free of charge. The two legal acts made it possible to diminish the area of fallow lands from 4.6 million ha (1918) to 0.37 million ha (1923). The total area of fish ponds increased from 37.6 thousand ha in 1923 to 88.8 thousand ha in 1938 (Table 1), mainly in the Lakelands; fewer ponds were in the regions of Pemarania, Wilno, Wołyń, and Podlasie. New fish ponds appeared on numerous water and unused land areas (9,10).

In the years prior to the outbreak of World War I, Poland produced more than 13,000 tons of fish (Table 2) and held first position as a European fish producer in the volume of output (30% of the European fish production) and the amount of fish produced per capita (0.37 kg). The average yield of Polish fish ponds was 163 kg/ha, with

Table 1. Fish Culture in Poland (1923–1938)[a]

	Year		
Feature	1923	1928	1938
Total area of ponds, ha	37,550	58,750	88,775
(%)	100	156	236
Number of pond farms, n	1,350	2,925	7,562
[%]	100	217	560
Pond production, tons	6,000	9,000	13,000
[%]	100	150	217

[a] References 4 and 10.

strong regional variations. The highest yields were still produced by fish breeders of Silesia and the areas near Kraków (to 800 kg/ha) and in Masovia (300–500 kg/ha), the lowest output was recorded in the provinces of Wilno, Polesie, and in Pomerania (to 123 kg/ha).

The first Polish Committee of Carp Selection was established in those years, headed by Prof. Franciszek Staff (1889–1966) and Edward Rudziński. After World War II, the Committee was presided over by Prof. Kaziemierz Stegman (1911–1978) of the SGGW in Warszawa.

Higher yields were obtained from fish ponds by widespread use of lupine to nourish fish (76% of the pond area) and mechanical and agronomic cultivation of the pond bottom, including fertilization. Pond operation was also aided by high unemployment and inexpensive labor in the countryside. There were also more than 50 centers of trout breeding (mountain and upland trout) of the area of 56 ha.

The FAO publications (Sarig 1966, cited by Ref. 11) indicate that carp from Poland and Germany were transported to Yugoslavia and further to Israel and the countries of the Far East, Indonesia (Java), North America, and Australia.

Postwar Period

After World War II and the subsequent change in the borders of the Polish state, the area of ponds in Poland diminished by more than 22 thousand ha and totaled 66,525 ha in 1945. Although about 9000 ha of fertile and well-managed ponds were taken over in the new lands in the west at the expense of 31.4 thousand ha of less fertile ponds in the former eastern territories of Poland (Table 3), the fish pond output attained at that time was comparable to that of the year 1938.

At present, the largest areas of carp ponds are in south and central Poland (three-quarters of the total area). In north Poland, in the territories of natural lakes (total area of 320 thousand ha) or in the mountainous regions, there are few ponds. The largest area of carp ponds is in the provinces of Silesia (Katowice)—11.4 thousand ha; Lower Silesia (Wrocław)—9.3 thousand ha; Masovia (Warsaw)—6.8 thousand ha; Pokarpackie (Rzeszów)—6.1

Table 3. Changes in the Structure of Inland Waters in Poland after World War II

	Area of Waters Ha		
Specification	Ponds	Lakes	Total
West Poland (new territories)[a]	9,150	202,850	212,000
—taken over by the USSR	31,400	103,536	134,936
—increase in total	—	+99,314	+77,064
Difference (−less)	−22,250	—	—

[a] Opole, Wrocław, Zielona Góra, Gorzów Wlkp., Szczecin.

Table 4. Fish Production in Ponds in Poland (1993–2001)[a]

Production (Thousand Tons)							
	Eating Fish[b]		Stocking Fish[c]		Total Production	Mean Price of 1 kg of Carp	
Year	CARP (K_H)	Others	K_1	K_2	[100.0%]	Zł	$ US
1993	20.4–61.4%	1.5	2.2	9.1	33.2	3.6	—
1995	19.6–64.5%	0.7	3.2	6.9	30.4	4.4	—
1996	21.4–63.5%	1.4	2.5	8.4	33.7	6.8	1.76
1999	21.4–64.1%	1.1	3.1	7.8	33.4	6.8	1.55
2000	22.7–60.2%	1.5	2.7	10.8	37.7	5.9	1.44
2001	20.3–68.1%	1.4	3.4	4.7	29.8	5.5	1.31

[a] Reference 24.

[b] K_h commercial, eating carp (2- and 3-year-old).

[c] K_1: 1-year-old carp (fingerling); K_2; 2-year-old carp.

Table 2. Total Area of Fish Ponds and Fish Production in Europe Between World Wars I and II[a]

		Area of Fish Ponds, ha			
Country	[year]	Total	Pond Area per 100,000 ha of Geographic Area	Carp Production, t	Fish Consumption, kg/person
Soviet Union[b]	[1935]	89,500	14.9	12,550	0.09
POLAND	[1938]	88,775	228.6	13,000	0.37
Germany	[1929]	75,000	131.6	7,500	0.11
Czechoslovakia	[1937]	52,000	359.4	2,700	0.16
Hungary	[1929]	9,800	93.3	2,600	0.25
Yugoslavia	[1931]	5,300	21.4	2,000	0.13
Latvia	[1930]	5,000	74.2	325	0.16
Lithuania	[1930]	3,000	35.7	300	0.08
Austria	[1932]	2,500	35.4	300	0.04
France	[1934]	1,200	21.5	.	.
Sweden	[1938]	1,600	3.5	.	0.03

[a] Reference 10.

[b] European part.

thousand ha; Wielkopolskie (Poznań)—5.8 thousand ha; and Lubelskie (Lublin)—5.7 thousand ha. In some parts of Poland, the total area of fish ponds reaches 6% of the geographical area of the locality (Pszczyna, Skoczów, Oświęcim) or up to 10% (Milicz near Wrocław). The total registered area of carp and trout ponds in Poland is 71 thousand ha (72% of which is the area used). This figure does not include several thousand home ponds built in the last 10 years by private individual farmers.

In the postwar years, inland fisheries in Poland, including pond fish farming, has been reorganized several times. As a result, until 1990 more than three-fourths of the pond area was managed by the State Fish Farms. The remaining area was exploited by fishermen's cooperatives, the Polish Angling Association, and, to a much smaller extent, by farmers and individual private fish farmers [about 16 thousand ha of ponds (23)].

Large and medium size trout farms are now located mainly in north Poland, in the area rich in clean, cold water. Fewer trout ponds are found in south Poland (the Karpaty and Beskidy mountains), predominantly smaller and older pounds (some dating back to prewar times), for example, łopuszna on the Dunajec River. The dynamic growth of this branch of the economy began in the 1960s and 1970s.

Currently there are more than 180 trout pounds in Poland, including very large fish farms whose output is 200–1,000 tons annually and produce mainly rainbow trout, for instance, Mylof (1150 t), Zabrodzie and Tarnowo (400–700 t). The output of trout in Poland was 10–12 thousand tons in the years 2000–2002 and of carp, more than 20,000 tons (Table 4).

Apart from carp farms, since the year 2001, the Polish Angling Association has been managing 16 trout breeding farms (producing mainly salmon fingerlings). These farms are located in the mountainous regions and in the lowlands. They produce fingerlings of salmonids, mainly salmon, and some species of river fish (idus, asp, vimba bream, barbel., chub, beaked carp undermouth, burbot).

The political and social transformation in Poland led to the liquidation of the State Fish Farms. A number of fish farms, some well functioning (both financially and technologically), have been divided into smaller units, not always organizationally efficient. Some of them, subject to restructuring, were taken over by the Agricultural Property Agency of the State Treasury (APA) as one-man companies of the State Treasury until their complete privatization. A considerable number of inland water reservoirs (lakes) has been leased and carp and trout ponds have been sold to private fish farmers and producers.

BIBLIOGRAPHY

1. Rudziński, E. (1961). Vergleichende Untesuschungen über den Wildkarpfen der Donau den Teichkarpfen. *Z. Fischer* **10**: 105–135.
2. Leszczyńska, W. and Biniakowski, L. (1967). Breeding of the wild carp in the sport fisheries stocking centre Otorowo. (in Polish). *Gospod. Rybna* **19**(2): 19–21.
3. Balon, E.K. (1974). *Domestication of the Carp (Cyprinus Carpio L).* R. Ont. Museum Life Sci. Misc. Publ, Ontario.
4. Włodek, J.M. (1980). Development of pond breeding of carp in south Poland. *Post. Nauk Roln.* **2/80**: 107–130.
5. Długosz, J. (1863–1887). *Chorografia Regni Poloniae. Historia Poloniae*. Opera omnia. Cracov.
6. Rudziński, E. (1962). Fragments of the history of the European domestic carp. (In Polish). *Gosp. Rybna* **9**: 3–5.
7. Rudziński, E. (1963). Carp in Poland. (In Polish). *Gosp. Rybna.* **1**: 6–8.
8. Górzyński, S. (1964). *Outline of the History of Fishing in Ancient Poland*. (In Polish). PWRiL, Warszawa.
9. Nyrek, A. (1966). *Fishery in Upper Silesia From Mid 16th to Mid 19th Century*. (In Polish, with French summaries). Travaux de la Societe des Scienc.et des Lett.de Wrocław, Vol. A-111.
10. Szczygielski, W. (1967). *A Brief History of Inland Fishery in Poland*. (In Polish, with Russian and English summaries). PWRiL, Warszawa.
11. Wolny, P. (1976). *Carp*. PWRiL, Warszawa, wyd. II.
12. Guziur, J. (1991). *Fisheries in Small Inland Reservoirs*. (In Polish). PWRiL, Warszawa.
13. Dubravius, J. (1543). *On Fish Ponds and Fish which Grow in Them* (Polish translation). Cracow, about 1600.
14. Strumieński, O. (1573). *On Management, Construction, Measuring and Stocking of Ponds*. (In ancient Polish). F. Kucharzewski, 1897, Cracov.
15. Stroynowski, S. (1609). *Description of Pond Order & Farm Management*. (In Old Polish). In: Gawarecki Z. and Kohn A. (1860). *Polskie stawowe gospodarstwo*. Warszawa.
16. Burda, V. (1935). *Grunlagen der Kapfenzucht*. Neudamm, Germany.
17. Hurt, R. (1960). *History of Pond Fisheries in Moravia and Silesia*. Vol. I–II. (In Czech). Brno, Opava.
18. Maylert, J. (1927). *Construction and Management of Polish Ponds*. (In Polish). Warszawa.
19. Kluk, K. (1780). *Natural History, Origins and Management of Domestic and Wild Animals* .... (In ancient Polish). Vol. III, Warszawa.
20. Leśniewski, P. (1837). *Polish Fishery*. (In Polish). Warszawa.
21. Morcinek, P. (1909). *Geschichte des Dubischverfahrens und des von Dubisch Herangezuchteten Schlesischen Karpfen*. Lands- und Fortswirtschafliche Gessellschaft, Troppau.
22. Guziur, J. (2001). *Polish and Czech Contacts in the Development of Pond Fishery in Cieszyn Silesia*. (In Czech). *Int. Fish Conf.* VURH-Ju Vodňany 2–5.05.
23. Guziur, J. and Turkowski, K. (1995). *Agriculture—Aquaculture Integrated Farming in North-East Poland—Present State and Prospects*. *Int. Aquacult. Conf. "Aquaculture in Eastern European Countries"*. St. Zagora (Bulg.), July 26–28, p. 15.
24. Krueger, A. (2000). *Pond Fisheries in Poland in 2001.* (In Polish). Int. Fish. Conf. PTR, Poznan, February 7-8, p. 8.

WATER: THE KEY TO NATURAL HEALTH AND HEALING

F. Batmanghelidj
Global Health Solutions
Falls Church, Virginia

It is true that the micromechanical side of medicine has advanced at dizzying rates, but our scientific understanding of the way the human body works is still in the Dark Ages. The inventive, resourceful, but commerce-motivated employees in science have managed to walk a tightrope between true science and illogical nonsense-science of medicine, to the point that the human body is now viewed as a badly designed product of nature in constant need of their interpretation and meddlesome interferences.

They have sold this false bill of goods to the gullible public to the point that a vast percentage of America's GNP is allocated to cope with their thought in our society. This trend started at the turn of the last century, when scientists were flexing their muscles to gain acceptance for their understanding of the human body. Some of these scientists did produce miracles—but a few did unimaginable damage.

Walter Bradford Cannon, for instance, a prominent physician of his time, did untold harm to the progress of medicine. Cannon asserted that thirst in the human body reveals itself by the presence of a dry mouth, brushing aside the earlier view that thirst is more of a general sensation that needs to be studied and defined. The simplistic but false dogma of "dry mouth portending thirst" gained a foothold in medical science and has been with us ever since.

This seemingly insignificant assertion changed the course of scientific inquiry into the way the human body works. It pushed science in the direction of the chemical composition of the body and how these solid components of the body interact with one another. It gave way to the dominance of "solutes" and application of the science of particle chemistry to the physiological functions of the body—an integrated and complex water-dependent multisystem system. Even now, the hydrolytic role of water in chemical reactions is not given any consideration when evaluating the outcome of chemical actions in the body.

This direction of totally erroneous thought has thrown medical science into ignorance about the complex, water-dependent natural design of the human body. Solute chemists took over and made a mess of medical science. Nowhere is this mistake more apparent than in AIDS research, when every ill that befalls a person with this diagnosis is blamed on a minute particle called a retrovirus—in complete disregard of the obvious.

To survive through the millennia, the human body has overcome deadly viruses such as polio, smallpox, measles, flu, and more, and has developed the ability to mount a very effective immune response within 9 days. In other words, if a virus has not been able to kill the person in the first 9 days of its attack, it will itself be killed, and the immune system will be all the wiser for the encounter. Surely if the body, generation after generation, has survived fast viruses, getting over slow viruses must be a cinch. With this logic installed within the human body, how could we blame a devastating "disease process" on a retrovirus?

Surely the presence of the retrovirus in people with AIDS has other, less commercial explanations that are being ignored for the obvious reason that a natural solution would emerge. If you read my book *Your Body's Many Cries for Water*, you will discover that the AIDS virus is a byproduct of a larger metabolic disturbance that has a drastic disruption of the amino acid pool composition of the body at the foundation of the problem. Naturally, from this direction of thought will emerge a noncommercial natural treatment protocol that will emancipate society from the burden of costly AIDS research.

THE ILLOGICAL FOUNDATIONS OF MEDICAL SCIENCE IN THE TWENTIETH CENTURY

In all, four errors in scientific assumptions have made a mess of our view of the human body—in itself an efficient, self-repairing, self-replicating design:

1. A dry mouth is the only signal indicating the body's need for water.
2. Water has no direct initiative in chemical reactions of the body and is only a subservient and passive element in the actions of the solid matter that it dissolves and circulates in the body.
3. The human body is efficient in regulating its water intake throughout the life span from childhood to old age.
4. Any fluid can satisfy the long-term water needs of the body.

WHY ARE THESE ASSUMPTIONS FALSE?

Dry mouth is not a reliable indicator of thirst because the body produces ample saliva even though there may be established and symptom-producing intracellular drought in different regions of the body. To facilitate the act of chewing and swallowing food, producing saliva, even in a dehydrated state, is an overriding primary function in the body. The salivary glands produce about 2 liters of saliva a day to supply the upper parts of the intestinal tract with water and some digestive enzymes, especially when the body is dehydrated. Oversalivation begins when the body is short of sodium, a contributing component of dehydration.

To think that water has no direct and dominant role in the chemical and physiological functions of the body reflects an earlier basic ignorance in molecular chemistry that has been applied to human physiology. It is now almost impossible to expunge this error from the way medical science is evolving. This thoughtless assumption is one of the main reasons we have a "sick-care system," not a healthcare system based on preventing disease.

Water itself must break down first before any chemical reaction can take place. It is the hydrolytic property of water that drives all chemical reactions. A chunk of meat has no energy value unless water acts on it and performs its hydrolytic act of breaking it down to small, energized particles that can then, and only then, be assimilated within the metabolic systems of the body.

Advising people to rely on the sharpness of their thirst sensation to hydrate their bodies optimally is yet another mistake of modern medical education. As we grow older, we gradually lose our thirst perception, to the point that the elderly can become obviously dehydrated yet not have the urge to reach for a glass of water sitting right next to them.

- Does this mean that the body can function efficiently in a comparatively drier state?
- Does this mean that fully hydrated plumlike cells can become prunelike without impacting most of the physiological events of the body that depend on their efficiency of function?

The answer to both questions is an emphatic no. In their earliest prunelike state, even minimally dehydrated cells begin to produce symptoms. Modern medicine has classified these manifestations as "diseases of unknown etiology." The dis-eases of dehydration have led to the creation of various fields of specialty medicine. Larger and larger hospitals and nursing homes have been formed to serve the medical profession and maintain a firm grip on a society that still does not know that water is a disease-preventing medication. It is essential for survival, and it is the master key to vibrant well-being of the human body.

The fourth mistake is the view that anything fluid can supply the long-term water needs of the body—hence the emergence of the soft drink industry, wines and beers included. Suffice it to say that both caffeine and alcohol are addictive and have cumulative adverse chemical effects on the body. Alcohol, even in small doses, negatively affects brain function by shutting down an important water-delivery system to the brain cells. In larger quantities, this effect becomes pronounced to the point of causing hangover headaches, indicators of brain thirst.

Caffeine is a chemical much like morphine and cocaine that some plants manufacture to protect themselves against animals that would feast on their nourishing leaves, seeds, and exotic poppies, the instruments of their survival as species on Mother Earth. Caffeine is a dehydrating substance. It also suppresses the brain tissue enzymes that are vital to the retentive powers of the nervous system. When the plant made caffeine in its process of development, it did so to make its feasting foes stupid enough to become easier prey to their own predators. The plant had no idea that humans would become its protectors and harvest its caffeine in greater and greater quantities. For the plant, caffeine is a chemical weapon.

If a child or a teenager is allowed to drink caffeinated sodas in place of water, attention deficit disorder and low scholastic results should be the early expected outcome. And all we need do is visit a few old age homes to see the state of the mind of people parked there by a society that does not understand the important difference between simple, tasteless natural water and taste-enhanced caffeine-laced water. The introduction of artificial sweeteners has been an added burden on our society. You can get a clear understanding of this problem by referring to my book, *Your Body's Many Cries for Water*.

About 100 years of fast-track travel on the train of these medical misconceptions has produced the commerce-driven quagmire of a sick-care system now known to be based on ignorance, if not fraud. The drug industry's hijacking of medical education has been so successful that even the shocking recent revelation that prescription medications—taken correctly and according to a physician's instructions—make over 2 million people sicker and kill more than 106,000 a year does not raise any eyebrows in official quarters. It is also known that another 98,000 die from mistaken prescriptions. These are only the *reported* cases; God knows how many die silently. These figures together—204,000 a year—put the use of prescription medications as the third most frequent cause of death in America, only after heart disease and cancer. Doctors are now fully aware that the chemicals they prescribe have been manufactured as palliatives.

WHY WATER IS A BETTER MEDICATION THAN ALL THE DRUGS ON THE MARKET

Water is a better medication than whatever is manufactured by the drug industry for any of the presently recognized painful degenerative diseases. These diseases are the products of an inefficient, water-deprived, prunelike state of cells that cannot keep up with the demands placed on them. Each year, more than 700,000 people die of heart disease, 500,000 die of cancer, and arterial disease of the brain kills over 160,000. All of these deaths might have been prevented if water was given a primary position in the daily diet.

THE NATURAL ELOQUENCE OF PAIN: IT SIGNIFIES THIRST

It took me many years of clinical and scientific research to uncover the significance of the pain mechanism as a primary indicator of water shortage in the body.

The human body is a satellite of water-dependent units of life—about a hundred trillion of them housed together in one capsule on their journey of life. The cells of the body, each endowed with the same genetic secrets of life, have accepted a division of labor for an orderly conduct of business in their collective land-based shelter—"the body." It has an intricate water-dependent design—a mobile chemical refinery—that has established a foothold on "dry land." The medical pundits have misunderstood the various ways the body deals with its inner cleansing processes. Nonetheless, they use commerce-suited treatment protocols for different pain sensations, without the slightest understanding why the body, at times, manifests diverse localized or general pains.

It is now clear that there is a direct link between overproduction of histamine, as the neurotransmitter that regulates the water balance of the body, and the induced pain-producing inflammatory effect of histamine that is its mechanism of directing increased circulation to a drought-stricken area. If the blood is concentrated because of water shortage, the pain threshold of the brain is automatically lowered and pain is felt until the drought-stricken area receives water. Many systems and mechanisms are engaged in this process, all of them regulated by the water level of the body.

The trigger mechanism is the balance between the acidity and the alkalinity inside cells. The interior of the cells of the body must be alkaline—a pH of about 7.4. Water and the element sodium from salt flow in and out of the cells and wash out the acid that is produced as a result of the metabolism of working cells. In dehydration, the rate of acid clearance is diminished, and the interiors of the cells in that area become more and more acidic. At a certain level of acidity—a level that obviously eats into and hurts the gene pool—an acid-sensitive substance in the nerve tissue called kininogen is converted to kinin. Kinin itself is a pain-producing substance. Thus, the nerve endings in the area report the local chemical environmental change to the brain.

The brain translates this information into a form of pain for the conscious mind and takes away the ability to use the area so that extra acid production is stopped until water comes and washes the excess acid away. This process causes a localized inflammatory reaction. The circulation opens up, and more blood is brought to cleanse the area and activate the repair processes to "undo" the damage. This damage is often in the joints of the hands and the legs; it could also be in the spinal column, in the heart muscle (causing angina), or in the intestinal tract (causing colitis).

There are about 10 frequently produced pains that should denote dehydration in the area in question: heartburn, joint pain, back pain, migraine headaches, angina, colitis pain, fibromyalgia, false appendicitis pain, peptic ulcer pains, and the eye pains of glaucoma. It is said that 110 million Americans suffer one or more of these pains at one time or another. The cost of pain management for these people is estimated at more than $100 billion a year.

RESOURCE MANAGEMENT: NOT DISEASE

Water is the universal solvent. It is also the solvent of what you might have been told is disease. When the body is short of water, the ingredients that it dissolves and carries around are also missing. Like any well-organized industrial complex, the body has a resource allocation program—a strict rationing plan. Water and the substances it distributes become subject to preplanned priority distribution protocols. Vital cells in vital organs are served first, even at the expense of the less critical parts of the body.

Water conservation, the emergency water-distribution and resource-management programs of the body, have been called allergies, asthma, hypertension, diabetes, and the autoimmune group of diseases.

Allergies

The body deals with foreign agents by producing antibodies that attach to the offending agents and chemically neutralize them, be they pollen, bacteria, or viruses. Histamine has the dual role of being the water- and resource-management regulator, as well as the immune system monitor and activator.

The design of the immune system is such that in dehydration and overactivity of histamine as the resource manager, the activity of the immune system is inhibited at—of all places—the very center of activity, the bone marrow. There are twice as many inhibitor cells responsive to histamine in the bone marrow as there are killer cells within the immune system. Thus, in dehydration and the added presence of histamine, the immune system becomes less and less active and leaves the body vulnerable to pollen, infections, and development of cancer cells. The relationship of allergies and the broader implications of the diminished immune system activity initiated by unintentional dehydration should be taken very seriously. Allergic reactions of the body in childhood translate to a variety of health problems in later years.

Asthma

Air that is exhaled in breathing carries with it lots of water—the winter steam in front of your face. In 24 hours, we lose about a liter of water through breathing. In its capacity as the body's primary water-management regulator, histamine brings about bronchial constrictions to restrict the free flow of air from the lungs, along with the water that would otherwise go with it. The restrictive breathing in asthma will totally disappear if the body becomes fully hydrated.

Seventeen million people in America, mostly children, suffer from asthma—never realizing that water is their lifesaving medication and is freely available in their own kitchens. I predict that in the near future, it will be revealed that the drug industry knew about the relationship of histamine and dehydration, but concealed this from doctors and the public to profit from innocent children who live in constant fear of suffocation and frequently die from asthma.

Hypertension

Sixty million people in America alone suffer from hypertension—yet these people are actually experiencing internal drought. The extra pressure is needed to filter water from an already concentrated blood and inject it through clusters of holes—as wide as only one water molecule—in the membranes of the cells in vital organs. This is how organs such as the brain, the liver, the lungs, the endocrine glands, the heart, and more, receive some water when the body is dehydrated. When the blood becomes concentrated, it tends to pull water osmotically out of the cells of the body. The process of injecting water into vital organs is a compensation mechanism. If adequate water and some minerals needed by the cell interior are taken daily to the point of making the blood more dilute, blood pressure levels will drop back to normal. For more information about hypertension, read *Your Body's Many Cries for Water*—and take a look at another book, *Water Cures: Drugs Kill*, to see how easily others have gotten over hypertension by adjusting their daily water intake.

Type II Diabetes

When food enters the stomach and intestines and is processed, insulin is produced to help feed the cells of the body. It facilitates the entry of circulating sugar and amino acids into the cell through its cell membrane receptor channels. Sugar and amino acids also pull the water that is osmotically attached to them into the cell through these channels. In a dehydrated body, this would be counterproductive for the needs of the brain. The brain is 85% water and does not need insulin to pick up its food components from blood circulation. To avoid the catastrophic outcome of delivering the vital needs of the brain to less vital organs, the process that regulates blood flow to the pancreas, when the body is suffering from internal drought, also inhibits insulin release from the pancreatic cells that produce it. The extra sugar in blood circulation also osmotically compensates for the salt that is missing in dehydration. Increased water intake and supplementation of minerals, with salt on top of the list, will correct the physiological imbalance in this condition, and the diabetes will clear up.

For more information about chronic unintentional dehydration and the diseases it causes, read as many of the following titles as you can. They are the products of more than 24 years of focused research on the topic of deep dehydration inside the cells of the body: *Your Body's Many Cries for Water; ABC of Asthma, Allergies and Lupus; Water for Health, for Healing, for Life; How to Deal with Back Pain and Rheumatoid Joint Pain; Water Cures: Drugs Kill*; and my 2004 book, *Obesity, Cancer, and Depression: The Deadly Diseases of Dehydration*. These books will leave no doubt in your mind that your body is truly a marvel of creation, designed to give your soul a perfect habitat for you to achieve the anxiety-free greatness you are capable of. When your brain is not preoccupied with the aches and pains of your body, you will have the freedom to soar through the limitless horizon of ideas and ideals—how the curative power of water came into light.

For more information on dehydration, visit www.watercure.com and www.nafhim.org.

WATER IN HISTORY

Geoff Petts
University of Birmingham
Birmingham, United Kingdom

The story of water in history is about human ambition to achieve food security, economic growth, and political power. It is both a social and a technological history. The history of civilization is about social organization and coordination, and for six millennia, successful water resource management was the basis of civilizations. Early cultures were built on the food surpluses produced by managing the floodwaters of large rivers for irrigation. In humid areas dominated by rain-fed agriculture, water was influential in other ways by enabling trade along inland waterways. Water also powered mills that fueled industrial and economic growth. But the successful use of riparian and floodplain sites was frequently interrupted by tragedies caused by flood and drought. The notorious Huang He was popularly known as 'China's Sorrow' because of its devastating floods and dramatic changes in its course, but it is also prone to severe droughts. In recent history, major floods in 1855 and 1938 took millions of lives and in destroying crops, brought widespread famine and disease. In the great drought of 1877–79, 13 million people died of starvation. Similarly, early Sumerian and Babylonian civilizations had to survive flood disasters. The Biblical famine described in Genesis (47: 13–21) may relate to a period of persistent drought when the Nile flood failed. In more recent times, the great Mississippi flood of 1927 was a disaster for millions of people, and it also changed the course of the American nation.

Throughout our history, the popular pioneering vision has seen the human conquest of nature as the route to greater wealth and security (1). In all long-settled regions of the world, modernization involved large-scale water projects to reclaim 'wastelands,' 'conserve' water, 'control' floods, and 'harvest' energy. But water is also fundamental to sustaining the earth's ecosystems upon which human societies depend. Two key events in history supported the establishment of the first civilizations: the advent of farming and the control of water. From that point on, humans gradually and then with increasing power, exerted control on the ecosystems in which they lived. During the past 6000 years, but especially since the Industrial Revolution, the human desire to control waters and to reclaim wetlands has led to widespread ecological degradation (2,3). From the mid-1960s and more widely during the past two decades, the role of water in history has changed to become a symbol for a new social philosophy (4), that focuses on *sustainable* environmental management, the protection of biodiversity, the restoration of river corridors and lakes, and the protection of the seas and oceans. The history of this relationship between humans and water is summarized in Table 1. This highlights four phases in the human conquest of water: (1) overcoming the threat of drought, (2) generating power, (3) protecting public health, and (4) protecting (and restoring) water environments.

WATER IN THE ANCIENT WORLD

The foundation of the earliest civilizations involved the efficient and equitable distribution of water through advancement of both technology and management systems. Goubert (5) considered our preoccupation with water one of the subdivisions of the religion of progress. In Judeo-Christian literature, water is a symbol of salvation demanding human vigilance and the imposition of order (6). However, the first revolution in water technology was possible only following the introduction of an indispensable new tool, script. This was necessary for organizing people, water, and soil in quantities that were too vast to be handled efficiently by the unrecorded memorizing of oral arrangements and instructions (7, p. 51). Thus, the Sumerian civilization depended for its survival on effective control and also on effective administration of the waters of the lower Tigris-Euphrates (7). Early, highly structured 'hydraulic civilizations' (8) developed throughout the Old World, in northern India, Southeast Asia, China, and MesoAmerica. These civilizations were founded on centralized state authorities. Water yielded

Table 1. Water in History

Years Ago	Date	
0	2000	
	1990	Emerging science of restoration ecology supports practical restoration schemes across Europe, North America, and Australia.
	1960	Concerns for public health and advances in water treatment begin to clean up water quality; and
		social concern for the environment prompts political action to control point-source pollution, protect wild areas, and assess environmental impacts of water developments.
50	1950	
	1930	Era of the megaproject begins with Tennessee Valley Authority (1933) and the Hoover Dam and Lake Mead on the Colorado River, U.S. (1935).
		Major rivers across Europe and in the U.S. become severely polluted.
100	1900	
	1890	Demonstration of the link between disease (e.g., cholera) and water.
	1880	First hydroelectric power schemes.
	1840	Extensive floodplain/wetland reclamation across Europe and the U.S.
		Urban growth, water pollution, and epidemics widespread.
200	1800	
		Water power drives the Industrial Revolution and supports growth of urban and industrial centres in upland valleys.
500	1500	Water power still chief provider of mechanical power for corn milling, textile manufacture, and iron works.
	1200	Earliest records of water pollution problems in Europe.
1000	1000	Water mills widespread.
	300	Romans built major, water-powered mills and engineered channels for inland transport.
2000	0	
	100 B.C.	The city of Rome has extensive and complex water-supply system.
		Large-scale land reclamation and flood control works in Egypt and China.
		Large-scale irrigation possible from introduction of Archimedes' screw pump and water wheel.
2500	500 B.C.	
	1500 B.C.	Introduction of water-lifting technology (shadouf) enabled localized perennial irrigation.
		1800–1550 B.C. famine in Egypt as Nile flood fails = ? Biblical famine (Genesis 47: 13–21)
		Regular Nile floods create stable and prosperous Egypt.
4000	2000 B.C.	
		Dry period (2200–2000 B.C.) caused Dark Age for many early civilizations from Greece to Mesopotamia.
		Ancient hydraulic civilizations appeared.
6000	4000 B.C.	

privileges for a nobility that generated considerable revenues from irrigated agriculture, mills, fisheries, and navigation. However, political power and social stability depended on advances in water management technology to secure food surpluses and water supplies for urban centers.

In these early civilizations, advances in water management were inspired by the desire to conquer drought and flood. Diversion weirs, storage reservoirs, and irrigation systems were developed along valley bottoms. In China, the Great Yu undertook complex drainage and flood control works on the Huang He about 4200 years ago (9). In mountain valleys, terracing was developed to conserve water, and soil conservation was an important by-product. However, two inventions had a particularly dramatic impact on the relationship between human civilizations and water. First, the invention of the arch by the Sumerians around 6000 years ago (10) overcame obstacles to transport and water conveyance. The bridge has been described as a symbol of the human conquest of nature, and the river crossing is prominent in social, economic, and military history (10). Worldwide, folklore indicates that bridges had an important place in the lives of early communities and a strong ritualistic character. River gods were made angry by the building of bridges that deprived them of their regular toll of drowning people and animals, and sacrifices were deliberately made during construction to appease the angry spirits (10). Second, the invention of the water wheel around 2300 years ago (11) made conservation and management of limited or fluctuating water supplies possible.

The Egyptian Nile spawned a great civilization some 5000 years ago which depended entirely on the river and its annual flood (11). The oldest agricultural sites in Egypt date from around 7200–6000 years ago. By the end of this period, the ancient Egyptian hydraulic civilization had evolved into strong state, kingship, and political capitals with temples, palaces, and royal cemeteries. Levees, cross dikes, dams, and canals were built to capture flood waters for cultivating a single winter crop. This farming method has been traditionally credited to King Menes, the founder of the 1st Dynasty in 5100 y B.P. However, a 200-year dry period between about 4200 and 4000 years ago had an adverse effect on the economic base of many early civilizations from Greece to Mesopotamia that entered a Dark Age dominated by social disorder, crop failures, famine, and political collapse. This pattern of economic prosperity and political stability during years of a reliable flood and periods of dry years of political impotence, famine, and civil war is repeated in the ancient history of these hydraulic civilizations.

Early basin agriculture was no more than an advanced stage of subsistence agriculture. The invention of lift irrigation about 3500 years ago sustained perennial irrigation. New technology (the shadouf or pale and bucket lever) and deep wells allowed the development of a second growing season during the summer months. This led to a new era of wealth. But this technology was effective only for small plots. Large-scale irrigation became possible only in Ptolemaic times, from about 2300 years ago, when the Archimedes screw pump and the water wheel were introduced. They could raise water to higher elevations and at a faster rate than the shadouf. Ptolemaic engineers were supported by a strong administration to supervise irrigation, land surveying, and dike, weir, and canal maintenance. Their efforts doubled the area of cultivable land and reclaimed some 325,000 acres of fertile land for agriculture by lowering the lake level in the Fayum basin. However, this was also the first era of major human impact upon river corridor ecosystems. Conversion of the floodplain to intensive agriculture destroyed the virgin pasture, wooded thickets, and pools.

ROME TO THE MIDDLE AGES

The Romans were the greatest engineers in ancient history. They built large water-powered mills and complex water-supply systems with long aqueducts. By about 2100 years ago, Rome had a large and complicated water distribution system capable of supplying an estimated 680,000–900,000 m^3 of water per day. The Roman Empire was in no sense a hydraulic civilization, but the monumental masonry aqueducts that led water to their cities endure as symbols of the power and glory of the empire.

The Romans channelized rivers to aid navigation, and important towns, such as York in the United Kingdom, became established at the head of navigable waterways (12). However, the most important legacy of the Roman Empire was the water mill, which Marx saw as a key force in the making of feudal civilizations. In England alone, 5000 water mills were listed in the Doomsday Survey of 1086 (13). Water power remained the chief provider of mechanical power for corn milling, textile manufacture and iron works until the nineteenth century.

However, another revolution in water engineering technology followed the change from feudal to early capitalist systems of land management. By the fifteenth century, the core areas of European market capitalism, upper Italy and Flanders, witnessed early innovation in the mechanical control of water flows and river regulation (6). From the late sixteenth century, the pace and scale of water engineering accelerated, underwritten legislatively and financially by the new centralized states and a capitalist land market. A primary driver was to reclaim marshlands and floodplains to increase the area for productive agriculture and the area that could contribute taxes, and to contribute to national pride in conquering nature (14). Thus, new engineering techniques were advanced to achieve more efficient land drainage. The Dutch developed dredging technology and designs for floodgates, retaining walls, and groins. In Italy, the problem of *La bonifica* (land reclamation in its broadest sense) stimulated advances in the systematic control of rivers. The ‘fens’ of England were drained between 1630 and 1653 by the Dutch engineer Cornelius Vermuyden using techniques previously employed in Holland and Italy. Once again, the management of the drainage required the transfer of administration from localized, decentralized, self-governing bodies to a more centralized authority, in this case, Parliament. But now there had evolved a new driver of progress, a scientific approach to water management. By the end of the seventeenth century the ‘art’ of water management had been replaced by a scientific methodology (15), not least as a result of the work of Castelli (1577–1643), the founder of modern hydraulics.

Prestigious projects were conceived. The Canal d’Entreroches, built between 1638 and 1664 and originally designed to link Lake Neuenberg and Lake Geneva, crossed the great European water divide between the catchments of the Rhine and Rhone rivers (16). This caught the public imagination and received funding from the Dutch and the citizens of Bern and Geneva. The scheme failed, but one section was completed and remained open for 180 years, transporting wine and salt. Other notable schemes include, in France, Vauban’s great military works for Louis XIV in the Canal du Midi (6). Thus began an era of belief in the domination of nature by science, freeing societies from scarcity, want, and natural disasters. This belief grew into the nineteenth century when Europeans exhibited an extraordinary faith in the power of science to control, manipulate, and enhance the natural world. Scientific innovation and technical skill drove dramatic expansion of European commercial and industrial power. Heffernan (17) considers Roudaire’s *la mer interieure*, a proposal to create a vast sea deep within the Saharan desert, a classical illustration of European imperial ambition.

The Postindustrial Revolution Era

In the Industrial Revolution, water assumed an ever-widening role for power, navigation, water supply, and waste disposal. Urban and economic growth brought wetland drainage, canalization of natural channels for boat traffic, and embanking and dredging of rivers for flood control. Freshwater springs dried up as groundwater abstraction increased to supply rapidly growing populations. Overcrowding resulted from the desire to house as many workers as possible in a small space near the industrial workplace. Contaminated ditches and streams became open sewers and breeding grounds for disease. Small streams and brooks were often lost, buried beneath expanding urban developments. Large lowland rivers stagnated, and important fisheries died. In London, chronic conditions had been reached on streams and rivers by the late eighteenth century, but even the tidal river was severely polluted by the 1840s (18). However, it was not until the link between water and health was established at the end of the nineteenth century that a revolution in water and wastewater management occurred (5,19).

Water and Health

Great epidemics of cholera and typhoid in nineteenth century Europe led to the establishment of Public Health Committees composed of public-spirited people and led to systematic sampling of water for laboratory analysis. From the 1830s, the notion of public health placed the issue of water supply in a wider context, including housing, cleanliness of towns, poverty, domestic habits, and disease. Cholera arrived in Britain in 1831 and in London, the most severe outbreak, in 1848–9, claimed 14,137 lives. Popular belief, known as the 'miasmic orthodoxy,' was that illness was spread by an infected atmosphere. The first connection between water and health was made by a London doctor, John Snow, who in 1849, suggested that water polluted by sewage might be the vehicle by which cholera was transmitted. In 1854, he traced the cause of a cholera epidemic in Soho to a single source, a polluted well. However, until the Pasteurian revolution (1880–1890), 'clean' water was synonymous with running water and 'dirty' water with stagnant water. Running waters were seen as self-purifying, capable of diluting wastes and destroying organic pollutants by oxidation. Public health management focused on draining ditches and marshes where mosquitoes bred, separating clean water supplies from sewage, and providing running water via complex systems of pipes and pumps to supply city dwellers. Paradoxically, public health problems were made worse by attempts to exploit the 'self-cleaning' capacity of running water by creating complex sewage collection and conveyance systems. The introduction of the water closet into common use in the 1840s (and the dramatic growth of London's population) coincided with a severe cholera outbreak in 1849 and appears to have promoted epidemics because the sewer systems to which they were connected could not support the volume of liquid waste. In London, the River Thames was so bad in the summer of 1858 that it became known as the 'Year of the Great Stink.'

Pasteur focused attention on the role of microbes in water pollution and established a new scientifically based definition of 'pure' water. In 1883, the German bacteriologist, Robert Koch, isolated the cholera bacillus and confirmed that cholera is usually spread by water contaminated by the feces of an infected person. Clean, running water became an indispensable element of the urban landscape; domestic space was restructured, the toilet became a separate room, personal hygiene became a private and solitary activity, and fountains gained utilitarian as well as aesthetic and religious significance.

The Technological Revolution in Water Storage and Power Generation

The twentieth century realized the control of water, both through the advancement of large-dam building technology (2) and of basinwide management programs. America's New Deal and the electrification of the Soviet Union were both as much social and ideological as they were symbols of national pride. Developments accelerated at the end of the nineteenth century following new advances in earth-moving and concrete technology. New schemes such as the Vyrnwy Dam in the River Severn Basin, United Kingdom, represented a new phase in both dam building (the dam was more than 40 m high and impounded a reservoir of 59.7 million m^3 (the largest in Europe at the time), the beginning of large-scale interbasin transfers across 100 km, and raised issues of multipurpose river regulation for water-supply, navigation, and fisheries (15).

Hydroelectric power was first demonstrated in 1882. The invention of electricity had a dramatic impact because it can be readily and exactly applied to a wide variety of functions and also because it can be easily delivered to a precise and remote point where it is wanted. The earliest operating combined water turbine and electricity generator was built in Northumberland, United Kingdom in 1879–82, and the first commercial hydroelectric power station was operated at Salmon Leap Falls on the River Bush in Northern Ireland in 1883 (20). However, the control of Niagara Falls between 1890 and 1910 inspired the utopian imagination through the capacity to control nature. The giant turbines allowed generating electricity on a new scale, and the scheme also demonstrated the potential for long-distance power transmission, on lines to New York (21).

During the 1930s in the United States, the wholesale transformation of river basins in the Appalachian valleys of the Tennessee Valley scheme and the construction of massive dams and reservoirs in the Colorado basin (22) opened the megaproject era. The 221-m high Hoover Dam and its 37,000-million m^3 reservoir, Lake Mead, dwarfed earlier schemes (2). In the Soviet Union, Lenin's vision of communism as socialism plus electrification of the whole country evolved in the great HEP schemes along the Volga, Dnieper and Don, and later along the Siberian River. All regions of the world experienced further acceleration in the rate of dam construction after 1950 to secure water supplies and generate hydroelectric power.

Harvey (23, p. 13) sees the twentieth century as one characterized by a new 'lust to dominate nature' and a fierce and uncompromising patriotism fed by reckless personal ambition. He suggests that the mammoth river regulation and HEP schemes built in Africa in the immediate postcolonial era, such as Kariba (1959) and Cabora Bassa (1974) on the Zambezi (1959) and Akosombo on the Volta (1965), are examples. By 1990, there were more than 40,000 large dams over 15 m high, and 400 megaprojects had been constructed or were planned. These megaprojects have dams over 150 m high, a dam volume of over 25 million m^3, a reservoir volume exceeding 25 billion m^3, and an installed capacity of more than 1000 megawatts (24). Irrigated agriculture produced 40% of the world's food but consumed some 2500 km^3 of water. By the mid-1970s, the great multipurpose dam, once a symbol of social advancement and technological prowess, was being attacked as environmentally and socially destructive (25). By 1984, there was a wealth of scientific research to demonstrate the ecological damage caused by dams and river regulation (2).

The most recent chapter in the history of the Nile valley (11) began in the nineteenth century when the desire to modernize Egypt focused on expanding the cultivation of cotton, a summer crop, to generate foreign exchange. Later rice and sugar cane, both large water users, became desirable commodities. The construction of dams sluices and canals began in 1861. The Aswan Dam was first completed in 1902 and then enlarged in 1912 and 1993. Finally, it created a storage reservoir of 5.2 billion cubic meters and supports an irrigated area five times that under early basin agriculture and a cropping intensity of nearly 170%. But this phase of development was not without its environmental consequences. The raising of soil moisture levels created favorable habitat for plant and animal parasites and caused waterlogged soils and salinization. Pesticide use exceeded 16,000 metric tons in the mid-1980s when the fertilizer application was nearly 150 kg per 1.038 acre. A high count of undesirable algae developed in the regulated flows causing odor and taste problems. Thirty of the 47 commercial fish species disappeared, and the sardine catch in the eastern Mediterranean was decimated. This stage in the regulation of the Nile changed the face of Egypt (11).

In 1990, Barbara Cummings published a book entitled *Dam the Rivers, Damn the People*, that focused attention on the wide environmental and social impacts of large hydroelectric schemes in Amazonia (26). In 1996, McCully in *Silenced Rivers. The Ecology and Politics of Large Dams*, wrote of 'pharaonic works' that mortgage an entire society to benefit the elite and to enhance national prestige, but which bring little long-term benefit to society at large (27). Yet megaprojects continued to be built, and new projects, such as the Three Gorges in China and Alqueva in Portugal, continue to fuel the sustainable development debate.

In the developed economies, problems of river regulation were subordinate to those of pollution. A typical industrial river in Europe, the Rhine (28), shows the growing severity of fecal and organic pollution until 1972. Since then, improved sewage and wastewater treatment have decreased the input of labile organics to rivers. However, following the impacts of fecal and organic pollution, other major water quality problems appeared (29). Three common problems have been accelerated eutrophication (30), resulting in nutrient input rates much higher those of natural environments; acidification (31) caused by 'acid rain' from long-range cross-frontier transport of air pollutants, especially sulphur dioxide; and salinization of rivers, lakes and soils in dry land areas (32). Pesticides, organic micropollutants, and metals have caused additional problems that often have slow, long-term effects where problems are magnified through food chains to impact higher animals severely, especially the top predators (33).

The change in the industrial base during the second half of the twentieth century led, finally, to the death of the waterfront. Shipping that had provided the transport system for the Industrial Revolution was replaced by road, rail and air transport, and a technological revolution caused the virtual disappearance of labor-intensive heavy industry. In the United Kingdom, the catastrophic decline of the powerful shipbuilding industry of the great deepwater cities, London, Liverpool, Newcastle, and Glasgow of the Thames, Mersey, Tyne, and Clyde, respectively, led to mass unemployment and decay of the urban fabric. The social response to urban living was one where people became increasingly alienated from, and disinterested in, the local natural environment. Nature and wildness became associated with remote areas, to be visited when desired, for recreation and leisure. Yet water has retained its value for societies, for supply and power production and also for its important amenity value. In aesthetic terms, water is strongly prospect oriented; the riverbank and lakeshore provide the classic compromise offering the advantage of good visibility and security at the same time. Indeed, the introduction of a reservoir or canal into an artificial landscape can greatly enhance landscape quality, and the waterfront has become an attractive and valuable asset for business and residential developments.

The New Age of Environmental Responsibility

The last 50 years of our history—since the quantitative revolution in science of the 1960s—has witnessed scientific advances that have demonstrated the full nature of human impacts on waters. Early classic works include *Man's Role in Changing the Face of the Earth* edited by William Thomas in 1956 and *Water, Earth and Man* edited by Richard Chorley in 1969 (34,35). However, from a political perspective, the dramatic impact of Rachel Carson's book, *Silent Spring* (36), aroused public concern about contamination of water and air and the threat to wildlife and human health and provided the catalyst for change in policy and practice.

By the last decade of the twentieth century, a new approach to water management emerged (4) that has five facets. The first focuses on wastewater treatment and pollution control. The second is concerned with allocating water to protect aquatic ecosystems from the impacts of abstraction and river flow regulation. The third is concerned with the 'nature-like' channel designs and environmentally sensitive civil engineering works. Fourth, there is increasing concern for establishing protected areas and habitat refuges isolated from human interference, including fishing and hunting as well as habitat disturbance. However, the fifth is particularly important: collaborative decision-making involving informed stakeholders and public participation, supported by public educational strategies promoting community ownership of waters and their associated ecosystems.

Advancements in the scientific understanding of large waterbodies (37) and complex water transfer schemes (38) have provided the context for these advances. New scientifically based tools for water management have evolved (39,40). However, legislation and public participation are key elements of the new approach. Legislation for managing commercial and recreational fisheries has a long history (41). In the late twentieth century, some governments passed legislation limiting the uses of waterbodies for a wide range of reasons, including nature conservation (42). At the Earth Summit of 1992, *The Rio Declaration* and *Agenda 21: Programme for Action for Sustainable Development* (43) placed biodiversity, ecosystems, and popular participation at the head of the 'development' agenda. We now need to close the debate about unsustainable exploitation or scientifically based stewardship.

BIBLIOGRAPHY

1. Cosgrove, D. and Petts, G. (Eds.). (1990). *Water, Engineering and Landscape*. Belhaven Press, London.
2. Petts, G.E. (1984). *Impounded Rivers*. John Wiley, Chichester, UK.
3. Brooks, A. (1988). Channelized Rivers: perspectives for environmental management. John Wiley and Sons, Chichester, UK.
4. Boon, P.J., Davies, B.R., and Petts, G.E. (Eds.). (2000). *Global Perspectives on River Conservation: Science, Policy and Practice*. John Wiley, Chichester, UK.
5. Goubert, J-P. (1986). *The Conquest of Water. The Advent of Health in the Industrial Age*. Princeton University Press, Princeton, NJ.
6. Cosgrove, D. (1990). *An Elemental Division: Water Control and Engineered Landscape*. In: Cosgrove and Petts (Eds.). *op cit.*, pp. 1–11.
7. Toynbee, A. (1976). *Mankind and Mother Earth*. Oxford University Press, Oxford.
8. Wittfogel, K.A. (1956). *The Hydraulic Civilisations*. In: Thomas, *op cit.*, pp. 152–164.
9. Changning, L. (1989). Problems in management of the Yellow River basin. *Regulated Rivers* **3**: 361–370.
10. Mare, E. (1954). *The Bridges of Britain*. B.T. Batsford, London.
11. Said, R. (1993). *The River Nile. Geology, Hydrology and Utilization*. Pergamon Press, Oxford.
12. Petts, G.E. (1990). Historical analysis of fluvial hydro systems. In: *Historical Change of Large Alluvial Rivers: Western Europe*. G.E. Petts, H. Moller, and A.L. Roux (Eds.). John Wiley, Chichester, UK, pp. 1–18.
13. Sheail, J. (1988). River regulation in the United Kingdom: An historical perspective. *Regulated Rivers* **2**: 221–232.
14. Butlin, R. (1990). *Drainage and Land-use in the Fenlands and Fen-Edge of Northeast Cambridgeshire in the Seventeenth and Eighteenth Centuries*. In: Cosgrove and Petts (Eds.). *op cit.*, pp. 54–76.
15. Petts, G.E. (1988). River regulation in the United Kingdom. *Regulated Rivers* **2**: 201–220.
16. Vischer, D. (1989). Impact of 18th and 19th century river training works: Three case studies from Switzerland. In: *Historical Change of Large Alluvial Rivers: Western Europe*. G.E. Petts, H. Moller, and A.L. Roux (Eds.). John Wiley, Chichester, UK, pp. 19–40.
17. Heffernan, M.J. (1990). *Bringing the Desert to Bloom: French Ambitions in the Sahara Desert During the Late Nineteenth Century—the Strange Case of 'La Mer Interieure'*. In: Cosgrove and Petts (Eds.). *op cit.*, pp. 94–114.
18. Wheeler, A. (1979). *The Tidal Thames: The History of a River and its Fishes*. Routledge and Kegan Paul, London.
19. Halliday, S. (1999). *The Great Stink of London. Sir Joseph Bazalgette and the Cleansing of the Victorian Capital*. Sutton, Stroud.
20. Gibbon, H.D. (1969). *The History of Water in Ulster*. Augustus M. Kelley, New York.
21. McKinsey, E. (1985). *Niagara Falls: Icon of the American Sublime*. Cambridge University Press, Cambridge.
22. Allen, F. (1952). *The Big Change: America Transforms Itself, 1900–1950*. Harper and Boss, New York.
23. Harvey, D. (1989). *The Condition of Postmodernity*. Blackwell, Oxford.
24. Gleick, P.H. (1998). *The World's Water. The Biennial Report on Freshwater Resources*. Island Press, Washington, DC.
25. White, G.F. (1977). *Environmental Effects of Complex River Development*. Westview, Boulder, CO.
26. Cummings, B.J. (1990). *Dam the Rivers, Damn the People*. Earthscan, London.
27. McCully, P. (1996). *Silenced Rivers: the Ecology and Politics of Large Dams*. Zed Books, London.
28. Kempe, S., Pettine, M., and Cauwet, G. (1991). Biogeochemistry of European rivers. In: *Biogeochemistry of Major World Rivers*. E.T. Degenes, S. Kempe, and J.E. Richey (Eds.). SCOPE 42, John Wiley, Chichester, UK, pp. 169–212.
29. Meybeck, M. and Helmer, R. (1989). The quality of rivers: from pristine stage to global pollution. *Palaeogeography, Palaeolimnology, Palaeoecology (Global and Planetary Change Section)* **75**: 283–309.
30. Harper, D.M. (1992). *Eutrophication of Freshwaters: Principles, Problems and Restoration*. Chapman and Hall, London.
31. Batterbee, R.W. (1994). Surface water acidification. In: *The Changing Environment*. N. Roberts (Ed.). Blackwell, Oxford, pp. 213–240.
32. Grainger, A. (1990). *The Threatening Desert*. Earlscar, London.
33. Carson, R. (1965). Silent Spring. Penguin, Bristol, UK.
34. Thomas, W.L., Jr. (Ed.). (1956). *Man's Role in Changing the Face of the Earth*. University of Chicago Press, Chicago, IL.
35. Chorley, R.J. (Ed.). (1969). *Water, Earth and Man*. Methuen, London.
36. Carson, R. (1966). *Silent Spring*. Penguin, Harmondsworth.
37. Lubinski, K., Wiener, J., and Bhowmik, N. (Eds.). (1995). Sustaining the ecological integrity of large floodplain rivers. *Regulated Rivers* **11**: 1–388.
38. Davies, B.R., Thoms, M.C., and Meador, M. (1992). The ecological impacts of inter-basin water transfers, and their threats to river basin integrity and conservation. *Aquatic Conservation: Mar. Freshwater Ecosystems* **2**: 325–349.
39. Calow, P. and Petts, G.E. (Eds.). (1992). *The Rivers Handbook*. Vol. 1. Blackwell Scientific, Oxford.
40. Calow, P. and Petts, G.E. (Eds.). (1994). *The Rivers Handbook*. Vol. 2. Blackwell Scientific, Oxford.
41. Welcomme, R.L. (1985). *River Fisheries*. Fish technical Paper 262. FAO, Rome.
42. Palmer, C.G., Peckham, B., and Soltan, F. (2000). *The Role of Legislation in River Conservation*. In: Boon, Davies, and Petts (Eds.). *op cit.*, pp. 475–492.
43. United Nations Conference on Environment and Development (UNCED). (1993). Agenda 21: Programme of Action for Sustainable Development. United Nations Department of Public Information, United Nations, New York.

READING LIST

Barry, J.M. (1997). *Rising Tide: The Great Mississippi Flood of 1927 and How It Changed America*. Simon and Schuster, New York.

Biswas, A.K. (1970). *History of Hydrology*. North Holland, Amsterdam, London.

Harris, L.E. (1957). Land drainage and reclamation. In: *A History of Technology*. C. Singer, E.J. Holmyard, A.R. Hall, and T.I. Williams (Eds.). Volume III From the *Renaissance to the Industrial Revolution c. 1500–1750*, Oxford.

Lambert, W.G. and Millard, A.R. (1969). *Atra-hasis—The Babylonian Story of the Flood*. Oxford University Press, London.

Showers, K.B. (2000). *Popular Participation in River Conservation*. In: Boon, Davies and Petts (Eds.). *op cit.*, pp. 459–474.

HOOVER DAM HISTORY

Bureau of Reclamation—U.S.
Department of the Interior

SPILLWAYS

The following article was put together by George, a former guide at Hoover Dam.

Many people who take the tour here at Hoover Dam want to know when they will get to see the water go over the top of the dam. Well, the water has never gone over the top of the dam and probably never will. We don't want the water to go over the top of the dam for a couple of reasons. First, the power house is located at the foot of the dam. The power house contains 17 large generators, each producing enough electricity to service about 100,000 people. All that water would be bad for the electrical generators (water and electricity don't play well together). Second, there are about 18,200 vehicles a day going across the top of the dam, and we don't want those vehicles to get swept away. The road across the top of the dam is a federal highway, and it is the shortest way to get from Las Vegas to points east.

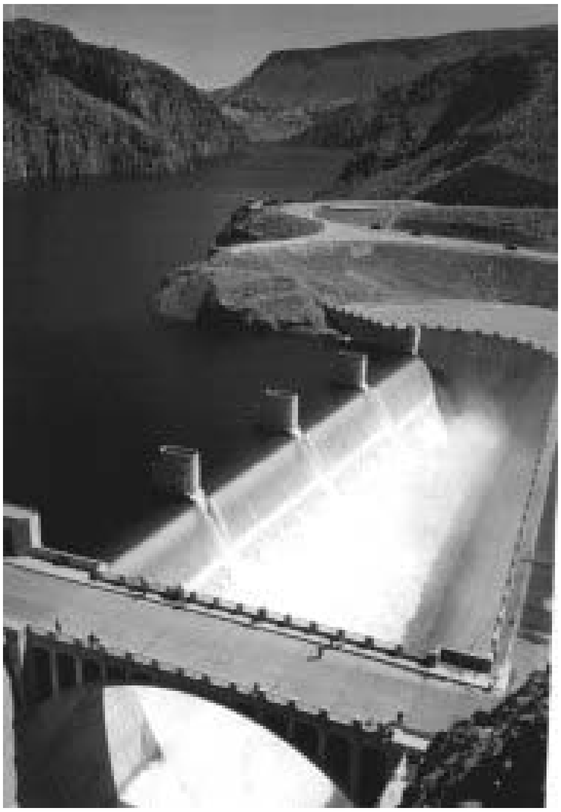

Water will probably never go over the top of the dam due to the spillways. The spillways work just like the overflow hole in your bathtub or sink at home (if you don't remember seeing that hole, go look for it right now). If the water ever gets up that high, it will go in the hole and down the drain, not over the top and onto the bathroom

---

floor (unless, you have children and they plugged up the hole). The spillways are located 27 feet below the top of the dam, one on each side of the dam. Any water getting up that high will go into the spillways then into tunnels 50 feet in diameter, and 600 feet long which are inclined at a steep angle and connect to two of the original diversion tunnels. Each spillway can handle 200,000 cubic feet per second (cfs) of water. The flow at Niagara Falls is about 200,000 cfs, so there is the potential for two Niagara Falls here.

Each spillway has four steel drum gates, each 100 feet long and 16 feet high. These gates can't stop the water going into the spillway, but they do allow an additional 16 feet of water to be stored in the reservoir. Each gate weighs approximately 5,000,000 pounds. Automatic control with optional manual operation is provided for raising and lowering the gates. When in raised position a gate may be held continuously in that position by the pressure of water against its bottom, until the water surface of the reservoir rises above a fixed point, when by action of a float the gate is automatically lowered. As the flood peak decreases, the gate can be operated manually so as to gradually empty the flood control portion of the reservoir without creation of flood conditions down stream. The spillways have been used twice. The first time, in 1941, was a test of the system. The second time, in 1983, was for a flood.

The Arizona spillway was placed in operation for the first time on August 6, 1941, soon after the reservoir level had reached a maximum elevation of 1220.44. The drum gates were raised for several hours on August 14, 1941, and a hurried inspection revealed that the tunnel lining was intact, and the inclined portion showed little or no signs of erosion at that time. Operations were then continued without interruption until the reservoir level had been lowered to elevation 1205.60 on December 1, 1942. The average discharge flow through the Arizona spillway during this period was approximately 13,500 cfs with a maximum flow of 38,000 cfs on October 28, 1941, when one of the drum gates dropped without warning.

That much water falling 600 feet down a very steep tunnel caused erosion of the tunnel lining. The eroded area was approximately 115 feet long and 30 feet wide, with a maximum depth of approximately 45 feet. The original volume of the cavity was 1069.6 cubic yards. Repair work was started almost immediately, but because it was believed that ordinary concrete was not suitable it was decided to utilize the Prepack and Intrusion process of concrete repair developed by the Dur-ite Company of Chicago, IL. After repair, the tunnel was polished smooth to help prevent future erosion.

During 1983, record flows into Lake Mead were recorded. The record surface elevation was recorded on July 24, with more than two feet of water spilling over the raised spillway gates of Nevada and Arizona. The record flows through the spillway tunnels again caused erosion in the concrete base, which had to be repaired. High water was responsible for wide spread damage throughout the project.

HYDROPSYCHOLOGY

Bellie Sivakumar
University of California
Davis, California

INTRODUCTION

Population growth and environmental degradation have become primary reasons for water scarcity on Earth. Added to this, inadequate water planning and management activities have resulted in numerous water conflicts around the world, whether in developed, developing, or underdeveloped areas. Although efforts to address these problems continue on one hand, one cannot help but also witness the role of regional, social, political, economic, and even religious and linguistic factors, among others, as an increasingly troubling menace in these conflicts.

This gathering threat is further evident from the major shift in human perception and expression of the water scarcity problem. If recent reports are any indication, we are faced with a situation where any or a combination of these factors can take precedence over "humanity" and, thus, we are left with no choice but to express our unhappiness/dissatisfaction about the consequences in the "most blunt and even antagonizing" way [e.g., "What a Lovely Water War!" (1); "Water shortages foster terrorism" (2)] rather than in a "diplomatic" manner.

Under these circumstances, what we, as water scientists, planners, and managers, can and must do to alleviate the violent and dangerous consequences, in addition to what we are already doing in terms of "water science" per se, is and will be under serious public scrutiny. Asking ourselves two relevant questions (and possibly others) would help address this:

1. How do our water planning and management activities affect human behavior?
2. How does human behavior, in turn, influence our water planning and management activities?

It is clear that answering these questions is not a simple and straightforward task, particularly in light of our traditional "rule" of conducting water research, i.e., quantity and quality (although not without "exceptions" involving also social, political, economic, and other factors). It is rather more complicated, warranting inclusion of yet another important dimension, the "human side."

In view of the significance of this "human dimension," Sivakumar (3) argues that studying the transactions between humans and water-related activities (e.g., planning and management) should be an essential part of water research, and also proposes the term "hydropsychology" for this field of research. Citing an ongoing interstate water conflict in south India (the Cauvery River water dispute) as an example, Sivakumar (3) discusses the contribution of "human behavior" to this conflict, in terms of the two questions raised above. He opines that the "linguistic" factor combined with politics, rather than social and economic factors, plays the key role in this conflict, which is only worsening with the passage of time. His question, particularly directed at water researchers and managers in these two states, is how do they balance their "science and management" on one hand against their "emotional and psychological interests" on the other. Specifically discussing their responsibilities and obligations in accurately reporting the ground realities on water data and disseminating research findings, Sivakumar (3) points out an important moral dilemma they (we) face: to be (more) faithful to science or to region, community, language, etc.?

With the increasing disorder, violence, and even hatred among various human sections around the world because of social, political, economic, and other factors, it should not be difficult to foresee a gloomy world order because of the continued worsening of the water situation as well. In fact, many observers even believe that (fresh)water *will* be the source of violent conflict among communities and nation-states in the not-too-distant future [e.g., (4,5)]. As a combination of "water scarcity" and "human behavior" can have profound consequences, the traditional "water quantity and quality" research must be supplemented and complemented with the new "hydropsychology" research. The purpose of this article is to present possible ways to make progress in hydropsychology research, which, in turn, could shed some light on studying the water crisis and the associated conflicts as a whole.

PROGRESS IN HYDROPSYCHOLOGY

The field of hydropsychology is new and, therefore, research is yet to take shape. In what follows, some ideas are presented to begin research in this field.

Problem Acknowledgment

Like in any other study, the first step in studying the role of human behavior in water conflicts is the recognition that it indeed exists and the acknowledgment that it is significant. It is often possible (and even considered wise) to "overlook" this role, because studying human behavior involves some "sensitive" aspects, such as social and economic inequality, political ideological differences, and cultural, racial, religious, and linguistic differences. Any effort toward making progress in hydropsychology, therefore, requires an open and gracious admission on our part of the need to study these sensitive issues and our willingness to do so.

Problem Identification

On acknowledging the problem, the second step is to investigate the "actual" relevance of human behaviors in water conflicts and vice-versa, which is extremely important, because some of the so-called "water" conflicts may not necessarily be related to water in itself, but water may be used merely as a "scapegoat" for advancing other interests. These interests may involve one or more of the sensitive aspects mentioned above. For example, Sivakumar (3) opines that linguistic interests and political compulsions in Karnataka and Tamil Nadu play more important roles than water itself in the Cauvery River dispute, either as sources of the dispute or as obstacles to find an amicable solution. Many "water" conflicts around the world exist that have "other" facets to them than just water, e.g., the Indus River water dispute between India and Pakistan (political and religious compulsions, national pride).

Although it is clear that accurate assessment of the "correlation" between human behaviors and water conflicts is important to move forward in hydropsychology research, this task is tremendously difficult, or at least can be tricky, at times. The only way to address this issue is by identifying all the relevant parties that, we feel, may have a role to play in the conflict, in the water sector, and/or in the human behavior sector. Depending on the conflict of interest, this may involve bringing the public (e.g., community leaders) and politicians together with researchers in various disciplines, such as water science, environmental science, political science, social science, psychology, theology, etc.

Problem Solving

With the above identification of the relevant parties, the next step is to conduct extensive individual and group consultations and discussions. Of particular importance among these parties are the water researchers and managers, politicians, and community leaders. Water researchers and managers play crucial roles in water planning and management, such as assessment of water availability, storage, and distribution; the existence of any conflict of interest to us may also motivate, and even define, our specific research agenda, as highlighted by Sivakumar (3), who cites his own (regional) interests in discussing the Cauvery River water dispute. Politicians make important policy decisions in regard to water and, therefore, have the authority to approve or reject our proposals for not only water storage and distribution but also water research itself; Sivakumar (3) extensively discusses the role of local and national politics, politicians, and political compulsions in the Cauvery water dispute. Community leaders (e.g., social, cultural, religious), by their own status and privilege, have a crucial responsibility and also the means to advise and guide the people (their followers) in ways that would negate conflicts and violence; in other words, they have a responsibility to keep a "check" on the people.

Although there is no question that these aspects have been studied in water science, in one way or another, such studies have been limited only to "quantity and quality" aspects. For example, farmer participation [e.g., (6)] and socio-economic analysis (7) have been increasingly studied. A host of scientific techniques have also been employed in these studies, such as Delphi Technique, Expert System, and Decision Support System. What has not been studied, however, is the specific role of human "behavioral or psychological" aspect in water science, which, unfortunately, seems to be a more important factor in the existing water conflicts around the world, and may become even more crucial in the future ones.

As studying the human behavioral aspect transcends our traditional water science expertise, developing ideas and techniques would indeed be challenging. It is obvious that we would benefit from, and even require, concepts and tools used in behavioral science. What efforts we will make to combine the knowledge gained in water science and in behavioral science and how (much) we will succeed in supplementing and complementing the ideas and concepts remains to be seen. In this regard, it is encouraging to note that recent advances made in the study of complex systems (e.g., nonlinear dynamics and chaos) and others have proven to be useful in water science [e.g., (8)] as well as in psychology (9). The existing techniques for studying other aspects in water science, such as farmer participation and socio-economic analysis, should also help us lay out a solid groundwork in this challenging task.

Research Dissemination

The outcomes of the hydropsychology research must be disseminated not just to researchers in water and other scientific fields but also to common people. Particular emphasis must be given to dissemination through media outlets, so that it would be easily accessible by all sections of the society. Also, the communication must be simple and effective for even the layman to understand, just like he also understands "What a Lovely Water War!" and "Water shortages foster terrorism." In other words, we, as water researchers and managers, must communicate with all sections of the society (particularly with the common people) with a clear message that we understand and care about their plight and that we are willing to work with everyone to solve their water crisis and the conflicts, which, of course, is easier said than done. Not only is this task extremely challenging, it may also result in "backlashes" and "negative effects." Although we must expect these, we must also keep in mind that these negative effects are often short-term ones. We must be willing, therefore, to continue to move forward with our efforts, only for the long-term benefits of "humanity as a whole" rather than the short-term benefits of "special interest groups."

CLOSING REMARKS

Recent trends in population growth and environmental degradation have brought us to a realization that fresh water *will* be the source of violent conflicts around the world in the near future. Although lack of water is certainly part of this problem, existing water conflicts around the world reveal that other facets also exist (e.g., socio-politico-economic and even regional, religious, and linguistic factors), which are reflections of human behavior.

This article is a follow-up of a recent "hydropsychology" proposal (3) that advocated studying the transactions between humans and water-related activities to improve our water resources planning and management. Some preliminary ideas to begin research in this field were presented, which range from acknowledging the problem in the first place to its identification and solution and, finally, to dissemination of outcomes.

We, as water researchers and managers, as members of the society, and as the more fortunate members of the scientific community dealing with the most precious ingredient for our life only next to air, have a moral responsibility and obligation to find amicable solutions to existing water conflicts as well as to avoid potential future ones. It is my hope that we will start making efforts on hydropsychology research sooner rather than later, and we will strive toward planning and managing water for humanity and peace on Earth.

BIBLIOGRAPHY

1. D'Souza, D. (2002). *What a Lovely Water War!* Rediff News. October 14. Available: http://www.rediff.com.
2. Sutherland, B. (2003). *Water Shortages 'Foster Terrorism'*. BBC News Online. March 18, Kyoto, Japan.
3. Sivakumar, B. (2004). Hydropsychology: the human side of water research. *Hydrol. Sci. J.* (in review).
4. Lorey, D.E. (2003). *Global Environmental Challenges of the Twenty-First Century*. Scholarly Resources Inc., Wilmington, DE, pp. 1–312.
5. El Kody, M. (2003). *The 3rd World Water Forum*, March 16–23, Kyoto, Japan.
6. Dick, M.R. (1997). Farmer participation in irrigation: 20 years of experience and lessons for the future. *Irri. Drain. Syst.* **11**(2): 103–118.
7. New South Wales Department of Land and Water Conservation. (1998). *Water Sharing the Way Forward: Community Based Socio-Economic Assessment*. Sydney, Australia.
8. Sivakumar, B. (2000). Chaos theory in hydrology: important issues and interpretations. *J. Hydrol.* **227**(1–4): 1–20.
9. Abraham, F.D. and Gilgen, A.R. (Eds.). (1995). *Chaos Theory in Psychology*. Praeger, New Haven, CT.

JACOB'S WELL

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

Jacob's well was dug two miles southeast of the ancient city of Schechem, an area that is currently recognized as the city of Nablus in Israel. The generally accepted history of this well begins with Jacob, the original owner of the well and one of the Hebrew patriarchs. It is said that Jacob purchased the plot of land for grazing purposes "for a hundred pieces of money" (Gen. 33:18–19; Josh. 24:32; John 4:5–6). Jacob and his twelve sons were said to have dug the well together in the center of the property he purchased.

Jacob's well is famous for its religious connection to Jacob, and also to Jesus as well. It is believed that at this well, Jesus asked a Samaritan woman to retrieve water for him. Although its religious significance made it famous, there are other reasons that make this water source important. Limited rainfall and water supply plague the area. The story in the Bible (John 4:1–42) about Jesus and the Samaritan woman at Jacob's well, at a purely literal level, reveals the importance of fresh water sources to the people of this arid land.

> *"Jesus answered, 'Everyone who drinks this water will be thirsty again, but whoever drinks the water I give him will never thirst. Indeed, the water I give him will become a spring of water welling up to eternal life.*
>
> *The woman said to him, 'Sir, give me this water so that I won't get thirsty and have to keep coming here to draw water." (John 4:13–15)*

The Samaritan woman takes his offer literally as an end to her need for water, instead, Jesus attempted to draw a parallel to thing that she could relate to, her survival by means of water, his offer of salvation by drinking the "living water." Her response shows how important a water source, such as Jacob's well, was to their continued existence.

The well alone is an example how the people of this region and their livestock have sustained themselves since ancient times. The effort that was put into building a well of this magnitude illustrates the importance of this water source.

Water rights in this area have been contested for centuries. It was one of the conflicts between Samaritans and Jews in ancient times and part of the current conflict between the Israelis and the Arabs. The struggle is over a limited amount of land that is suitable for sustaining life in this area. It has become a more serious issue due to population growth over the years as a result of higher birth rates and immigration.

Jacob's well is approximately 100 feet deep and 9 feet in diameter, though it is very likely that when it was first dug, it exceeded this depth. Water gathers in the well from an underground spring. It has been called "living water" because the water comes from a moving, fresh stream. A well of these dimensions would have required a great deal of work and a large sum of money in ancient times. Today, Jacob's well is still being used; it has become a tourist attraction and religious shrine. A large, unfinished church has been erected over it. The land that surrounds the well is currently owned and tended by the Greek Orthodox Church.

READING LIST

Easton's Bible Dictionary, "Jacob's Well." www.studylight.org/dic/ebd/view.cgi?number=T1946 (July 12, 2001).

FFOZ Magazine, "Where There Water, There Is Life." www.ffoz/Magazine/Land/wateris.htm (July 12, 2001).

"Jacob's Well" www.furman.edu/mcknight/jerreg14.htm (July 12, 2001).

The Plain Truth, "The Woman at the Well," www.ptm.org/MarWomanAtWell.htm (July 12, 2001).

"Shechem," www.worldmissions.org/Clipper/HistArch/Shechem.htm (July 12, 2001).

UCU Jacob's Well, "What Is Jacob's Well" http://www.canberra.edu.au/chaplaincy/catholic/cjwell.html (July 12, 2001).

White, Harold Dr. "Sermon Storehouse" http://members.nbci.com (July 12, 2001).

JAUBERT DE PASSA: THE FIRST WORLD HISTORY OF IRRIGATION IN 1846

Eric P. Mollard
Institut de Recherche pour le Développement
France

The first world history of irrigation appeared in 1846, and it has remained the only one to the present. The four volumes of the "Recherches sur les arrosages chez les peuples anciens" (Research on Watering among Ancients) describe hydraulic enterprizes from the most modest to the most monumental, and provide the political and social history of people who mobilized and organized themselves for this purpose. François Jaubert de Passa (1785–1856) was 61 years old when he wrote The First World History of Irrigation. This Catalan landowner, trained in law and after a political career, defended the conviction of a lifetime and intended to demonstrate it by systematically writing of irrigation from antiquity to the beginning of the nineteenth century. The lessons he drew from the comparisons in various cultural and political contexts are always topical, so much so that they match contemporary international issues. How is it possible that this early work, which had the potential for becoming an important foundation for organizing irrigation, has been all but ignored for one century and a half?

THE AUTHOR, HIS TIME, HIS CONVICTIONS

The book was published at a time when France experienced an institutional shift. The first half of the nineteenth century was under a political economy, and the second was marked by the reinforcement of technocracies and state intervention.

After the Napoleonic epic and the Congress of Vienna in 1815, France resumed military operations: the Spanish expedition, naval intervention in Greece, and the seizure of Algiers. The Geography Society of Paris has existed since 1822, the date when Champollion deciphered hieroglyphics. The military interventions continued (Rio de la Plata in 1838), and commercial establishments were set up in the Ivory Coast and Madagascar. Detailed accounts of travelers animated the discussions of many scientific societies under the reign of Louis-Philippe (1830–1848). In line with Montesquieu[1] and Toqueville,[2] political systems and their advantages for economic activity were discussed.[3]

In 1845, the Great Famine in Ireland began and lasted until 1867. Agronomists theorized around Liebig's discoveries in biology and, after the physiocrats, they strove to diffuse the intensive Norfolk rotation. At the end of the first half-century, the Research is directly influenced by the agronomy of notable landowners and the political economy. Jaubert de Passa gave an account of his observations in silkworm, mulberry tree, and oak cork farming, as well as in watering in the department of Pyrénées-Orientales. As a lawyer, he wrote the rules of the Têt users association, a river which crossed Perpignan and irrigated the plantations of Rousillon. Using the Research, he wanted to convince the politicians of the

---

[1] *De l'esprit des lois* was published in 1748.

[2] *De la démocratie en Amérique* was published in 1840.

[3] The Revue d'Économie Politique was created in 1833; the review Economie Politique Chrétienne of Villeneuve-Bargemont dates from 1834; the Société d'Économie Politique was created between Augustin Thierry et Michelet's books; the Annales de la Société d'Économie Politique were published in 1846.

productivity of irrigated agriculture:[4] "How much there are not, even in France, enlightened spirits persuaded that irrigation cannot have, in the fields of Gaul, results other than to multiply the meadows and to increase the desirable production of meat." He also tried to convince people of the role of institutions in guaranteeing freedom and authority. In addition to his own writings,[5] he gathered two sources, ancient authors' and travelers' descriptions. As a comparative history of civilizations covering the Old World, except sub-Saharan Africa, this treatise incorporated the sagacity of a field man sensitive to water government and social practices. In other words, Jaubert de Passa built an irrigation history on the basis of users' behavior, the real contracts that they have, and the local relationships that mediate the interaction between a regime and productive needs.

In 1851, the reign of Napoleon III began. Interventionism, which was legitimated by Auguste Comte's positivism,[6] resulted in great reclamation operations: the Suez Canal, development of the Gascogne Moors, and damming of the Rhone river. Upon the advent of state bodies of engineers, the debates gave up the ground of the political economy and the Research losts their interest at a time that emphasized technology. One of the possible reasons for the oblivion of the social component in irrigation was that local arrangements do not make the fame of the ruler sufficiently obvious, nor does homage to "progress." It is amusing to read, one decade in advance, an explanation of the exclusivity given to monumental technology. The next choices of Napoleon III echo with those of a pharaoh: "in his turn Menes raized dams and dug canals; moreover, he rectified the Nile riverbed and created artificial lakes to multiply jobs, to fertilize new soils, and to give royalty a character of beneficial greatness,"[7] as well as an Egyptian vice-pasha at the time: "The dam (at the apex of the delta) is not a modern idea ... Mehemet-Ali wanted to resume the project, to attach his name to a colossal work worthy of his genius."[8]

The interest in great irrigation under the aegis of the state was concretized one-half century later in Algeria and by general resident Lyautey in Morocco. It is possible that the Research influenced the French law of 1865 on user associations, which granted them legal status. Conversely, it is not certain whether small peasant irrigation was ever encouraged by the ruler as he had wished.

IRRIGATION AND GOVERNANCE IN THE "RESEARCH"

The Research is composed of six geographical parts (Table 1), which detail, region by region, the history of irrigation. Jaubert de Passa's objective is not theoretical, and only the final chapter advances some principles for an institutional framework to support the installation and durability of irrigation in France. However, one can highlight five great ideas through the combinations of situations and lessons that he draws through writing.

Irrigation and Civilization

Jaubert de Passa underlined the colossal aspect of antique hydraulic works, which are not unrelated to the Pyramids, the Hanging Gardens of Babylon, and the Great Wall of China: the construction of the Imperial canal spread out over five hundred years, "we will have an overall length of 440 leagues, the longest navigable canal in the world."[9] In a more pharaonic way, "... in the area that the Ionians were the first to indicate by the name Delta (Herodotus) ... with periodic floods on a flat ground without defense, it was required there to raize hillocks to build the villages and insulated dwellings. These hillocks required immense embankments because the delta had large cities, temples of extraordinary extension ...."[10]

Would Karl Wittfogel, one century later, have written "Oriental Despotism" (2) in the same manner if he had read Jaubert de Passa? Through the study of great hydraulic civilizations of the past, Wittfogel highlighted the need for a strong bureaucracy in creating and managing hydraulic projects, a bureaucracy that gave structure to the state apparatus and eventually to the whole society. The material and ideological bases of "Oriental Despotism" would reside in great irrigation. The two authors thus agreed on the institutional and political nature that underlies the construction of collective works.[11] On the other hand, Jaubert de Passa granted more credit to governance than to institutional structures. He showed that despotism or democracy can be either a brake on or a lever for irrigation.

Table 1. Composition of the Book

1st part	Assyrian Empire	Assyria, Babylon, Armenia ...
2nd part	Hindustan and Ganges valley	Hindustan, India, Thibet, Sindh, Nepal, Assam, Burma, Siam, Java ...
3rd part	China	Waterings, hydraulic works, agriculture, Japan, Siberia
4th part	Syria, Arabia, Egypt	Damascus, Antioch, Phonicia, Arabia, Yemen, Ethiopia ...
5th part	East	Cyprus, Ancient Greece, Thessalia, Asia Minor
6th part	Ancient Sicily, Italy, Roman Africa	Property among Romans, Republic, Empire, Tuscany, Algeria, the Sahara, Morocco
Chapter IV of the 6th part	About laws and water regime in relation to agriculture	Classification of water, concession, administration, competence, and jurisdiction
Index		

---

[4] "Irrigation is a practice as miraculous in its effects as common in its means." III 131 (Part III, p. 131 from reedition of 1981, Coll. Les Introuvables. Editions d'Aujourd'hui).

[5] Jaubert de Passa published observations about Spanish irrigation in Valencia and in Elche palm grove.

[6] 1798–1857.

[7] IV325.

[8] IV394.

[9] III99.

[10] IV344.

[11] "In a hydraulic society, civil servants have privileged rents, which benefit owners in occidental society."

The Factors in the Origin of Hydraulic Works

A large hydraulic enterprize supposes the capacity to mobilize and organize an enormous labor force in a more or less coercive way. Wittfogel translates gigantism into social terms: "the delegates of power behave like tax masters of the national surplus and prestigious consumers." Unlike administrative terror and tax and legal submission in Wittfogel's book, Jaubert de Passa proposes a contract or, at least, sharing between leaders and people of social norms where religion would play an essential role, just like the confidence delegated to an administration that would animate a paternalist design of its role. Jaubert de Passa stated that tyranny within despotism is not durable.

The Research specified the factors in the origin of the canals, which do not necessarily have irrigation as their first goal. "From now on it was less a question of opening new canals than to improve, extend and especially to apply the discharge canals to the needs for agriculture."[12] The canals frequently have contradictory functions which require arbitration by the authority. The multifunctional character of the hydraulic technique applies besides to other sectors, in particular, the use of water in agriculture. Also, watering is used not only to irrigate, strictly speaking, to bring water to profit from additional production, but also to fertilize or improve the quality of production.[13] This plurality of functions, which results in distinguishing places and times, makes any attempt at generalization a delicate matter. Finally, irrigation is also related to land colonization and the possibilities of selling agricultural goods: "the multiplication of the canals accentuated the role of the Nile in transport."[14] Other factors took part in creating hydraulic works: trade, army, cities, tax department, and strengthening power,[15] and the ruler's fame. Water marketing is seldom practiced: "the public sale of derived water, in the only area of Ispahan, produced for the Shah, at the time of Chardin,[16] [...] 180,000 francs."[17] The comparison of tax pressures established by Jaubert de Passa displays the great disparity which prevailed between the countries and the times. Like ancient Greece, where the tax was at the sixteenth under Solon,[18] imperial China also had the wisdom to limit pressure: "At no time, the emperor could be unaware of the public sufferings ... China is thus the region in Asia where land is the least taxed."[19] In fact, the surpluses released by irrigated agriculture could lead to strong tax pressures. But it was not only the tax department which was to be taken into account, but all measures that encourage or discourage the farmer: "Not only the rates, but arbitrary taxes, in the East, harm agriculture."[20] Moreover, throughout his work, the author gave more emphasis to social practices and abuses than to formal government structures.

The Sharing of Social Standards

Social cement, which ensured the thousand-year-old sustainability of irrigated agriculture, rests on the standards shared by the various classes in a society: users, an attentive administration, and a fair ruler, whose richness comes from the farming community. Social submission and terror cannot function for long because farmers can unite in great masses and revolt against the establishment. Such revolts resulted in the overthrow of several Chinese dynasties by exasperated peasants. Civil disobedience, infrastructure destruction, revolts, abandonment of irrigated farming,[21] even emigration, as in the Nile delta under the Turks, are realities a government, tyrannical though it can be, authorizes only in the event of a major crisis. Consequently, the author suggested that communities placed their prerogatives and their confidence in the hands of the government to repair the infrastructure quickly, to protect rights, and to arbitrate conflicts. As the territory was extended by military invasion or land colonization, the more the state incorporated these prerogatives, the more it was able to mobilize labor and resources for increasingly monumental works. The surplus released by irrigated agriculture reinforced the class not depending on the land and the capacity to undertake increasingly colossal work, often at the request of the peasants themselves.

The contract was not always explicit because it was embedded in a total social fabric. In India, a religious reference is at the origin of irrigation: "There are even poor men who beg for several years to have a tarpa dug later in an uncultivated soil or on the edge of a commercial road. Bringing water in a deprived place is a commendable action in the eyes of Brahma;"[22] "In Egypt, in ancient times, the magistrates appointed for water supervision were taken from the caste of priests; their judgements were all the more respected because they were provided in front of the gates of the sanctuaries."[23] In an identical way, the civil water court in Valencia, Spain, always sits at the entrance to the cathedral.

Religion played a political role of countervailing power and a barrier to despotism: "As agents of this legislation, the priests always opposed a salutary brake on the sovereign authority and popularity; their

---

[12] III59.

[13] "Irrigation was mainly aimed at increasing the acreage and the return of meadows. In the rich grassland of Thessalia a horse breed adapted to warfare was reared" V57. "The Sibi district is crippled with rocky hills ... Man draws an admirable benefit from restricted resource of nature: with the streams and on poor and sandy soils, he created rich soils whose fertility seems inexhaustible" VI252.

[14] IV292.

[15] For instance, "The Persian king blocked a small valley, thereby holding the water and making a vast reservoir ... Suddenly deprived of irrigation, lowland farmers saw their millet and sesame crop perishing [...] Tired of so many pains, users came to implore at the palace gate, and they got, against high fees, that the water gates ... would be opened successively for the watering of the five small valleys" I191.

[16] 1643–1713. Author of "Voyage en Perse et aux Indes Orientales."

[17] I126.

[18] V45.

[19] III162.

[20] IV225.

[21] Except in Egypt, where rain-fed cropping is impossible.

[22] II204.

[23] VI351.

intervention imposed on the ruler circumspect functioning in conformity with the laws, at the same time as they gave to the people lessons of morals and more moderate feelings."[24] Jaubert de Passa added, "From absolute possession of the land to the control of the agricultural classes and the humiliation of the privileged caste, the distance is short and easy to cross: the priests understood it, and soon, advized or intimidated by them, Joseph returned to the owners of the lands which they had sold to the king, with the charge for them to pay annually to the royal treasure the fifth of all the products."[25] Not only is the purpose of one good administrative organization to intervene advizedly, but it must also share the social standard and be subject to control. The administration must justify its presence to the taxpayer. Tax and operational involvement of an administration are invariant and, for lack of countervailing power, it can easily justify itself in the eyes of the only ruler by an opinion which favors the monumental enterprize. This double function of the administration was particularly clear in Egypt: "Notifications transmitted by the inspector of the nilometer were going regularly to warn the governors of the provinces: the farmers and the officers of the king modeled themselves on these notifications, the ones, as Strabon says, to warn with the distribution of water and to do the necessary work on canals and dikes; others, to establish the tax, because they adjust it to the extent of the height of water."[26]

Conversely, the pragmatism that the management of water requires resulted sometimes in sacrificing the best principles. The Chinese administration thus chose the confusion of powers: "This apparent confusion in the performance of two quite distinct duties (delegate of water and judge) has as a result, in an empire where centralization and control are thorough until the last detail, to prevent conflicts between rival authorities and to facilitate the operations."[27]

The Decline of Civilizations

Nothing is ever gained, and the factors in the origin of great civilizations can be called into question. The ecological factors and the natural disasters were exploited little by Jaubert de Passa, who overlooked land salinization in Mesopotamia and the recurring silting up of tanks and canals.[28] Should we, therefore, assume that societies, at any given time, cannot mobilize themselves to rehabilitate drains, clean canals, and create new tanks?

His analysis of the decline of Egyptian irrigation under the Turkish empire led to detail governance resulting from the regime set up by the Divan.[29] First of all, he criticized the Ptolemaic kings: "Under Arcadius (395 A.D.), the tyranny of the tax department was extreme, but it respected watering, even by attacking ownership ... One of the laws stated that those who were accused of having opened a dam along the Nile were to be condemned to be burned at the stake ... Under Zenon in the year 474, the annual tribute of Egypt was suddenly raized from 50 pounds to 500 pounds of gold ... Under Phocas, the Egyptians were excluded from all public employment. Lastly, under Heraclius (year 641), the religious fights caused a civil turmoil which put an end to the domination of Greek emperors."[30] At the time of Arab domination, which was often favorable to the diffusion of irrigation, he showed that it resulted in weakening the pillars of traditional Egyptian religion. But it was especially the Turkish regime which caused the abandonment of a good part of the delta by farmers.[31]

The first problem occurred from new task sharing imposed by the Turkish government: "To reduce the burden on the treasury, the Divan maintained the old habits, and it declared the population responsible for maintenance, in spite of confiscation, which deprives them, mainly, of the benefit resulting from this work."[32] The negligence and incompetence of the high administration are added: "Since the drudgery was maintained, work from now on was exposed to the inexperience and arbitrariness of the governors: they forgot, rather generally, to carry out certain repairs in due course ... Often materials are missed, the time is short, and the unhappy fellah, curved under the whip, drives in posts, piles up corn canes."[33] The submission became physical and there remained nothing but terror to try, with dubious success, to preserve irrigated agriculture. The abuses and the privileges of big men resulted in appropriating water. Corruption increased in all levels of the administration. That will provide Bonaparte's words: "Better than Mamelucks, or let's retire."

Law and Institutions

Good legislation and freedom of undertakings are two principles on which, Jaubert de Passa claimed, irrigated agriculture needed to be developed in France. Watering must "be freed from the obstacles of the tax department, protected by the law and encouraged by the ruler."[34] He established a parallel between landed property and water rights. Private though it is, property requires an authority that is respected on the one hand; in addition, it is subject to the public interest which demands various constraints, in particular the right of easement (or land crossing). In a period when feudal survival had maintained local rights in many countries, Jaubert de Passa believed that water must belong in the public domain. In the East, such a principle is rooted in religion and is confirmed by legislation: "the laws of Zarathushtra, Manu and the

---

[24] IV331.

[25] IV400.

[26] IV300.

[27] VI306.

[28] "An earthquake would have ruined the Jordan valley, which Moises compared to the Garden of Eden" IV247.

[29] Turkish Council.

[30] IV417.

[31] "The whole Niphaiat was covered a long time ago with villages and fertile countryside ... Canals are nowadays silted; water is derived to the Damiette branch" (Champollion) IV350. Jaubert de Passa was aware of the progressive raising of the delta lands through history, but neither he nor others of his time ever made of it the explanation of its decline.

[32] IV386.

[33] IV386.

[34] II69.

followers of Buddha admit as private water only that which runs in an artificial bed."[35] In Italy, nationalization which appeared at the end of the Middle Ages authorized the reclamation of the Po Valley. In France, the phenomenon occurred later, but nationalization is essential not only due to the advantages for economic activity, but also because it trimmed feudal power down and reinforced the alliance between landowners and the crown. The state was able then to concede a right to a legal or physical entity for the benefit of the public interest. The administration must then supervise the use of this right and the application of the clauses of the concession.

JAUBERT DE PASSA BEFORE TECHNOCRACY

The "Research on Watering among Ancients" gives the primary role to local practices and governance. The Orient, where institutions favored irrigation, was at the time a source of inspiration, and this arrangement resulted in a model to be followed, even a legacy. After one and a half centuries of technical progress that came primarily from the Occident, the direction of moral debt was reversed. All the countries did not undergo colonial imposition, but all developed a qualified administration set on progress and technologies: it was gravity irrigation; it is now pressure irrigation. Jaubert de Passa had not envisaged this drift which marginalized social arrangements, nor that which, under cover of authority, led the French administration to create and manage directly the hydraulic equipment, exceeding the initial mandate of control, standardization, and sanction.

Jaubert de Passa attached less interest to the technique than to the social forms to which it owes its existence, its choice and its mode of management. In no case is the absence of irrigation ascribable to a defect in technological knowledge, and he rejects the exclusively ecological or technological explanations: "It is thus not the nature of the soil nor the defect of intelligence which deprived Sicily of the benefit of irrigation."[36] Jaubert de Passa supplemented the advantages of modern technology with a social analysis; that was the case during the debates caused by the construction of the Mehemet-Ali dam in Egypt: "Increasing the volume of water for irrigation is not only to add to the production of the land, but also to facilitate work and to assist the poor fellah; it is also to free the modest field workman from a fight, always disastrous, against the large usurpers of water; it is finally to bring back law and order in the villages, and to remove from the agents of power a means of oppression and a ceaseless cause of venality."[37]

Private initiative needs authority. It guarantees rights and secures against conflicts. However, Jaubert de Passa avoided setting up universal principles because he knew the price of flexibility when public and private interest change. His field experience led him to highlight the bad years, which can be catastrophic in irrigated agriculture: among Egyptians as among the Chinese, taxation was adjusted in bad years. Moreover, Egyptian irrigation functioned without concessions, and the great Roman landowners did their business without turning to the state for help. His demonstration applies to the French case where small holdings dominate. "Even under a successively despotic, feudal, monarchical and constitutional government one can make broad concessions with agriculture." [38] Egyptian and Chinese despotism conditioned the extraordinary epic of irrigation, whereas Turkish despotism undermined it. The Athenian and Roman democracies did not show great effectiveness in this field. It is the agreement between the type of farmers, the institutional context, and the sociocultural practices at which any reform must aim. From the remarkable diversity of situations, which is born from reading the Research, a last lesson arises: the importance of the details of arrangements for each one mirror obvious social participation. The latter cannot fit with bureaucratic monolithism, which does not lend itself either to the necessary flexibility for irrigated agriculture. This is the meaning we have to give to the paradigm of governance, which applies both in the time of Jaubert de Passa and the present.

BIBLIOGRAPHY

1. Jaubert de Passa, François-Jacques, Baron. (1846). Recherches sur les arrosages chez les peuples anciens.
2. Wittfogel, K. (1957). *Oriental Despotism. A Comparative Study of Total Power*. Yale University Press, 1967.

READING LIST

Jaubert de Paçà. (1824). Recherches historiques sur la langue catalane. Col. Fons Antic Mars 2001, p. 200.

Jaubert de Passa, François-Jacques, Baron. (1821). Mémoire sur les cours d'eau et les canaux d'arrosage des Pyrénées-Orientales.

Jaubert de Passa, François-Jacques, Baron. (1823). Notice historique sur la ville et le comté d'Empurias.

Jaubert de Passa, François-Jacques, Baron. (1823). Voyage en Espagne, dans les années 1816, 1817, 1818, 1819, ou Recherche sur les arrosages, sur les lois et coutumes qui les régissent [Extrait du Rapport sur le concours général de la pratique des arrosages, fait par M. Héricart de Thury.]

Jaubert de Passa, François-Jacques, Baron. (1833). Recherches historiques et géographiques sur la montagne de Roses et le cap de Creus.

Jaubert de Passa, François-Jacques, Baron. (1901). Journal de Voyage d'un Roussillonnais. In: Torreilles, Ph. De Barcelone à Valence en 1819, d'après le Journal de voyage d'un Roussillonnais, par Ph. Torreilles. (Extrait de la Revue des Pyrénées, tome XIII, 1901.)

Jaubert de Passa, n.d., Poesia eroticoburlesca Rossellonesa del segle XIX. Obres de Fra Pere i Poésies catalanes. Fons Antic 2001, p. 172.

Mattes, Jaubert. Notice biographique F. Jaubert de Passa.

Monglond, A. (1922). Au pays de la Vénus d'Ille. Mérimée et Jaubert de Passa. Revue d'histoire littéraire de la France.

---

[35] VI269.

[36] VI109.

[37] IV395.

[38] VI343.

BENJAMIN FRANKLIN: FROM KITE TO LIGHTNING ROD

William E. Marks
Water Consciousness, Inc.
Martha's Vineyard,
Massachusetts

Benjamin Franklin (1706–1790) is an unusual rags to riches story. At age 17, he left his home in Boston, after a falling out with his family, and found his way to Philadelphia. Arriving penniless, but rich in reading knowledge and from working as a writer and newspaper printer with his brother in Boston, he began his rise in life. He entered a common-law marriage in 1730 with Deborah Read, a Pennsylvania woman with whom he was long acquainted. They had one son who was lost to smallpox as a boy, but raised a daughter, Sally, along with Franklin's illegitimate son, William. The mother of William, with whom Franklin had an affair, has been lost to history.

Franklin was a founding father of the United States, the first Grandmaster of the Freemasons in North America, a family man, business man, a printer, inventor, author, publisher, scientist, musician, diplomat, and politician. Obviously, Franklin was many things to many people.

Franklin's relationship with water may be outlined as follows:

1. His kite flying stunt to experiment with understanding how storm clouds filled with water generated electricity. Lightning is a phenomenon that still perplexes science.
2. His invention of the lightning rod.
3. His invention of the Armonica water music instrument.
4. His study of the Gulf Stream and his epiphany for having it printed on a navigation chart.
5. His study of weather patterns.
6. His invention of swimming flippers.

FROM ELECTRICITY-PRODUCING WATER IN THE SKY TO LIGHTNING RODS

One stormy day in June of 1752, Benjamin Franklin took a walk into a field to go kite flying with his 21-year-old son, William. The field rested along the banks of the Schuylkill. Trusting in his fathers scientific know-how, William probably had no idea that he was participating in a potentially life-threatening experiment. At the time, little did William realize that Franklin had so much excitement and interest in studying storms that he could have been dubbed "the world's first official storm chaser."

As the story goes, Franklin attached a piece of metal wire to the fuselage of his homemade kite to help attract electrical charges flying among the overhead clouds. The hemp string played out through Franklin's fingers, and the kite twisted and turned in the turbulent winds as it soared higher and higher. After a considerable length of the string had played out, Franklin tied a key to the taut string. Then, with a dampened silk ribbon tied to the knuckles of his free hand, Franklin and William waited to see what would happen.

Whether it was a collection of static electricity gathered along the moist hemp string or a "gentle" lightning strike, we may never know. Regardless, Ben Franklin reported that when he placed his hand with the insulating silk ribbon near the key, he received a jolt of electricity as sparks leaped from the key. Franklin's grounded body acted as a grounding path for the electrical current built up in the electric field of the storm clouds overhead. In addition to proving that thunderclouds generate electricity, Franklin also extrapolated his negative/positive electrical research from his laboratory and theorized that the lower part of the thunderclouds was mostly negatively charged.

"The rest," as is often said, "is history." Franklin and his son, William, survived this experiment, But others who have tried to duplicate this experiment have paid dearly with their lives or their bodies were crippled.

Long before his famous kite-flying experiment, Franklin believed that lightning and static electricity shared similarities. While living in Boston in 1746, Franklin set up a laboratory to conduct his "electrical amusements." At the time, there were other scientists in Boston conducting electrical experiments. It was in 1746, during his first year of conducting electrical experiments, that Franklin suffered a not-too-amusing jolt of electricity. In a letter to a friend he expressed the shocking experience as, "...a universal blow throughout my whole body from head to foot, which seemed within as well as without; after which the first thing I took notice of was a violent quick shaking of my body..." As a result of this electrical shock, Franklin experienced numbness in his arms and the back of his neck for a few days.

As Franklin continued to conduct his electrical experiments, he wrote down his efforts and shared them in letters to Peter Collinson in London. Collinson was a friend and scientist who shared an interest in Franklin's experiments. In 1749, Franklin sent a letter to Collinson describing the concept of an electrical battery that could hold a charge for a brief period of time. In the same letter, Franklin expressed doubts as to how such an invention could be of service to humankind.

During Franklin's experiments, he observed that when glass was rubbed, it became "filled" with an electrical charge, making it "positively charged." He also observed that when other objects, such as amber, were rubbed, electricity flowed out from them, and therefore they became "negatively charged." This designation by Franklin of "positive" and "negative" electrical charges established a new base language for understanding electricity that remains with us to this day. Prior to Franklin's time the words "vitreous" and "resinous" were used to express the charges of electricity.

In his experiments, Franklin observed that when certain objects were "filled" with static electricity, they would discharge a "spark" of electricity when placed near a pointed metal needle. Franklin was not the first to speculate about a relationship between lightning and the harmless sparks created by static electricity. William

Wall, a British scholar, had suggested such a relationship in 1708.

However, Franklin was the first to propose an experiment using a lightning rod on top of a building to capture the “fire” from the clouds. The first lightning rod described by Franklin was an iron rod 8 to 10 feet long that came to a sharpened point at its uppermost tip. Figures 1–4 show a variety of lightning rods throughout the years.

Two years prior to his kite experiment, Franklin conducted experiments where he used a large sharp iron needle to draw electricity away from a charged metal sphere. Through his observations, Franklin extrapolated the results to his invention of the lightning rod to protect the tops of buildings.

In 1750 he wrote two letters to his friend, Collinson. In these letters, Franklin postulated:

> May not the knowledge of this power of points be of use to mankind, in preserving houses, churches, ships, etc., from the stroke of lightning, by directing us to fix, on the highest parts of those edifices, upright rods of iron made sharp as a needle. . .Would not these pointed rods probably draw the electrical fire silently out of a cloud before it came nigh enough to strike, and thereby secure us from that most sudden and terrible mischief!

Figure 1. Ventilator/Lightning Rod from Gay Head Lighthouse wick 1900, Martha’s Vineyard, Massachusetts (photograph courtesy of William Waterway).

Figure 2. A weathervane and lightning rod combination. Location: Trinity Methodist Church, Oak Bluffs, Massachusetts (photograph courtesy of William Waterway).

> . . .On top of some high tower or steeple, place a kind of sentry box big enough to contain a man and an electrical stand. From the middle of the stand, let an iron rod rise. . .upright 20 or 30 feet, pointed very sharp at the end. If the electrical stand be kept clean and dry, a man standing on it when such clouds are passing low might be electrified and afford sparks, the rod drawing fire to him from the cloud. If any danger to the man be apprehended (though there would be none) let him stand on the floor of his box, and now and then bring near to the rod the loop of a wire that has one end fastened to the leads; he holding it by a wax candle [i.e., insulating him from it]. So the sparks, if the rod is electrified, will strike from the rod to the wire and not affect him.

Even though Franklin’s lightning rod proposal was first published in England in 1750 by *The Gentleman’s Magazine* and then by the Royal Society of London in 1751, there was little interest in England in acting on Franklin’s invention.

However, in 1752, Franklin’s letters were translated into French. Soon after, the lightning rod concept was quickly embraced by King Louis XV and his two leading natural scientists, Comte de Buffon and Thomas-Francois D’Alibard.

On May 10, 1752 in the country village of Marly located near Paris, the French built a sentry box according to Franklin’s published recommendations. The sentry box was equipped with a 40-foot lightning rod. A

Figure 3. A cross with lightning rod attached. Location: Town of Oak Bluffs on Martha's Vineyard, Massachusetts (photograph courtesy of William Waterway).

Figure 4. Four lightning rods on top of old Whaling Church, Edgartown, Massachusetts (photograph courtesy of William Waterway).

retired soldier, Thomas Francois D'Alibard, was somehow influenced to volunteer for the experiment. As fate would deliver, a storm passed over the sentry box in midafternoon. Much to the soldier's surprise, he and others witnessed the sparks leaping between the lightning rod and the insulated wire. Soon, the same experiment was conducted and verified in other areas of France.

When the French lightning rod experiment worked according to Franklin's theory, the King of France sent a message to Franklin expressing his compliments. However, due to the slow transit of mail in those days, Franklin did not receive the French King's message until August of 1752, two months after he had already proven his theory by successfully using the dangerous kite and key rig.

In June of 1752, Franklin was in Philadelphia watching the progress of a steeple being built on top of the Christ Church. Franklin was going to use this steeple for his first lightning rod experiment.

From his research and other "electrical amusements," Franklin speculated that the placement of pointed iron "rods" on top of buildings would prevent lightning from causing fires. He figured fires could be prevented because the pointed rod would capture the lightning's "fire sparks." On this note, Franklin was half right. It did not take Franklin long to figure out that lightning rods need to be "grounded." He accomplished this by using an insulated metal wire or cable to draw the charge from the lightning rod and to guide the electric charge harmlessly down into the ground away from the structure.

In 1753, Franklin published his lightning rod research in his Poor Richard's Almanac. He also provided a detailed description of his lightning rod invention. This helped to stimulate the installation of lightning rods throughout the American colonies.

The largest lightning rod installed in Franklin's lifetime was in 1788 at Maryland's State House. This rod was installed according to Franklin's specifications and has effectively protected the State House for more than 208 years.

Before Franklin's invention, lightning destroyed or damaged many buildings. Evidence of how rampant lightning damage was in the early 1700s can be found in various records. For instance, during a period of 33 years, in Germany alone, approximately 400 towers were damaged by lightning, and 120 bell ringers were killed. On a similar note, in France between 1753 to 1786, 386 church towers were struck, and 103 bell ringers were killed.

The installation of lightning rods in France and other countries immediately began to reduce the number of buildings struck by lightning and also reduced the deaths of bell ringers and other people in buildings with tall steeples. As the use of lightning rods slowly began to reduce the number of fires and deaths, the governments and the public began to appreciate the practical applications of Franklin's research into electricity and weather, especially, when one considers the antiquated fire fighting apparatus available in the 1700s. In Franklin's time, entire towns were known to burn down in fires ignited by lightning strikes.

Buildings, such as those with tall steeples (churches, tall mansions, and commercial buildings) were all susceptible to lightning strikes. Once outfitted with lightning rods, these tall structures were less likely to be struck by lightning. The tangible results of Franklin's invention in saving homes and buildings soon made him an international celebrity and a respected scientific genius.

However, when it came to installing lightning rods on churches or on public buildings where the church held influence, there was considerable resistance from many Catholic and Protestant church leaders. Relatively speaking, it was as though history was repeating itself; these same institutions also condemned the scientific theories of Copernicus and Galileo.

Both of these faiths taught and preached that being struck by lightning was a "sign" of God's punishment. When Franklin's experiments indicated that being struck by lightning was the unfortunate result of a random act in nature, he was immediately attacked by various church leaders.

Prior to Franklin's lightning rod, there was the age-old belief that the ringing of "consecrated" bells in a tower would help repel violent storms and lightning strikes. Therefore, as a thunder-and-lightning storm approached a town, the local bell ringer would do his religious duty to save the community. Little did he know that grabbing hold of the ringing rope and ringing the bells to ward off the evil storm was playing roulette with lightning. During thunderstorms, it was the accepted practice to ring the bells as violently as possible. This was done on the theory that the sound waves from the consecrated bells would break up the lightning in the overhead clouds. It is for this reason we find the words "Fulgaro Frango" inscribed on many medieval bells. "Fulgaro Frango" means "I break up the lightning flashes."

As people awakened to the relationship of lightning and the death of bell ringers, there soon became a struggle between the religious belief of ringing bells to ward off evil and the reality of being struck by lightning. Eventually, in 1786, 34 years after Franklin's lightning rod was first successfully tested in France, the Parliament of Paris signed an edict. The passing of this edict was "to make the custom of ringing church bells during storms illegal on account of the many deaths it caused to those pulling the ropes."

Due to this 1786 edict in France, we find one of the first examples of the forthcoming separation between church and state. A separation that was more or less prompted by Franklin's founding of the "scientific method," a method of inquiry into the nature of our world that caused some people to question the teachings of religion. This also helped to form the basis of fear from a religious perspective.

A well-documented attack against Franklin's lightning rod came in 1755 from a church leader in Massachusetts. The Reverend Thomas Prince, pastor of the Old South Church, published a sermon on the subject of lightning rods. In that sermon, he expressed the opinion that the recent big earthquake in Massachusetts was related to the installation of "iron points invented by the sagacious Mr. Franklin." His sermon also mentioned that when it came to installing lightning rods—"in Boston, are more erected than anywhere else in New England, and Boston seems to be more dreadfully shaken. Oh! there is no getting out of the mighty hand of God."

Due to the strong influence of religion in England, the first lightning rod installation on a church spire did not take place until 1762—10 years after the French began their installations. In the interim, London's Saint Bride's church was damaged by lightning several times.

In Austria's mountains of Carinthia, the Catholic church of Rosenberg was struck so frequently by lightning that peasants attending services were killed on several occasions. As a result, many of the peasants stopped attending services during stormy weather. The church steeple suffered lightning damage so severe, that it was completely rebuilt three times. When a lightning rod was installed in 1778, all lightning damage ceased.

In Italy, the tower of St. Mark's in Venice suffered lightning strike after strike. Records from 1417 onward show that the wooden spire was struck by lightning and consumed by fire on several occasions and suffered severe damage from many other lightning strikes. Due to the influence of the church, St. Mark's spire went unprotected. This is in spite of the fact that Beccaria, the respected Italian physicist, had introduced Franklin's invention to his country many years earlier. Even after St. Mark's spire was hit by lightning in 1761 and 1762, the church refused to install the "heretical rod." In time, as overwhelming evidence of the rod's effectiveness grew in neighboring countries, the church acquiesced, and a lightning rod was finally installed in 1766. Afterward, St. Mark's tower was never again damaged by lightning.

However, lessons such as this came too late to save the Church of St. Nazaire in Brescia, where the Republic of Venice had stored 100 tons of gunpowder. In those times, church vaults were often used for this purpose because they were dry. In 1767, a lightning strike hit the church spire and traveled down to the stored gunpowder. Over 3,000 people lost their lives, and a large portion of the city was destroyed.

Because of this common use of churches for storing gunpowder, there are similar stories well into the 1800s.

For example, in 1856, on the island of Rhodes, a lightning strike exploded the munition vaults and killed more than 4000 people.

Franklin's invention of the lightning rod proved to people the connection of lightning to storm clouds and thunder. This was the first proof that lightning was an electrical force generated in nature by the action of water molecules in the clouds. The mystery of how lightning is created by the motion of water molecules in storms was something that intrigued Franklin. However, he could only theorize as to the distribution of negative and positive charges in the clouds.

Franklin believed that his creation of the lightning rod was his greatest and most useful invention for the benefit of humankind. In spite of Franklin's apparent habit of keeping records and sharing information about his experiments through letters, it is of historic note that he never documented his kite experiment or wrote letters about the event. The only witness to the kite experiment was his son, William. And, the only written account of the kite and key experiment was by another person some 15 years after the event took place.

Franklin's lightning rod design was of an iron rod honed to a sharp point on its skyward tip. This design probably was developed as an extension of his laboratory experiments where he used an oversized needle with a sharpened point.

England was one of the last countries to begin installing lightning rods. The first lightning rod installation in England took place in 1760. The location was the Eddystone Lighthouse, an old wooden structure that had a history of damage by lightning. This lighthouse was vital to marine navigation safety and the commercial shipping interests of England.

As fate would deliver, nature further helped spur the use of Franklin's pointed lightning rod in England. In 1764, a bolt of lightning struck England's Saint Bride's steeple, already struck many times in the past. Saint Bride's church, a historic building in England, has a spire that stabs skyward prominently. The steeple's damage from lightning in 1764 prompted the Royal Society to recommend the use of Franklin's pointed lightning rods to protect the rest of the country's buildings.

As Franklin's pointed lightning rods were installed throughout Europe and America, King George III decided on a different approach. He decided to protect his palace with blunt-knobbed lightning rods, which he believed were superior to Franklin's. At the time, this also represented a political statement because King George viewed Franklin as a rebel. The use of blunt-knobbed lightning rods therefore represented a symbol of loyalty.

Regardless of King George III's, motivation to use blunt-tipped lightning rods in 1752, varying opinions and arguments have existed since that time, as to which lightning rod design is superior.

The controversy over which lightning rod was most effective was recently put to rest. In experiments conducted by the American Geophysical Union during the past decade, it was determined that the blunt lightning rod is more effective. Even though Franklin's design did not win this competition, he still deserves credit for bringing the lightning rod invention to the service of humankind. Besides the pointed and blunt versions of the lightning rod, there are many variations on the theme. Lightning rods have almost evolved into a folk art form, and many have considerable value to collectors.

On a closing note, as in the case of his inventions of the glass Armonica, bifocals, and the Franklin wood stove, Franklin never made money from his inventions, including the lightning rod. To quote Franklin, "As we enjoy great Advantages from the Inventions of others we should be glad of an Opportunity to serve others by any Invention of ours, and this we should do freely and generously."

This philosophy of Franklin's, about serving humanity, was also reflected in the words he wrote in 1790, the year of his death. From the tone of these words, it is evident that Franklin approached his earthly departure with philosophical humor:

> "I believe in one God, Creator of the universe....That the most acceptable service we render Him is doing good to His other children....As to Jesus...I have...some doubts as to his divinity; though it is a question I do not dogmatize upon, having never studied it, and think it needless to busy myself with it now, when I expect soon an opportunity of knowing the truth with less trouble."

WATER AND THE HISTORY OF MAN

Ann Azadpour-Keeley
National Risk Management
Research Laboratory, ORD,
U.S.EPA, Ada, Oklahoma

Jack W. Keeley
Environmental Engineer
Ada, Oklahoma

> And Isaac digged again the wells of water, which they had digged in the days of Abraham his father; for the Philistines had stopped them after the death of Abraham: and he called their names after the names by which his father had called them. (Genesis 26:18)

Before the earth was created, hydrogen, the oldest and most abundant material in the universe, and oxygen, created in the womb of stars, combined to form H_2O. Since the beginning of life on this planet and probably other planets scattered throughout infinite space, water has played a critical role in all facets of human existence, including the creation of life itself. We are made of water, and water produces the food we eat, from fish to fowl to fiber. It is the only substance on the earth that exists naturally in three phases, liquid, solid, and gas. As it serves us today in industry, agriculture, medicine, transportation, and the production of energy, it has served humans and all of the earth's creatures since the beginning of time.

In the days of Abraham, following family feuds and other conflicts, the victor often stopped wells by filling them with earth and stones. It was the equivalent of burning the village. Wells of water were so important that their owners gave them names, and communities

Previous Page

For example, in 1856, on the island of Rhodes, a lightning strike exploded the munition vaults and killed more than 4000 people.

Franklin's invention of the lightning rod proved to people the connection of lightning to storm clouds and thunder. This was the first proof that lightning was an electrical force generated in nature by the action of water molecules in the clouds. The mystery of how lightning is created by the motion of water molecules in storms was something that intrigued Franklin. However, he could only theorize as to the distribution of negative and positive charges in the clouds.

Franklin believed that his creation of the lightning rod was his greatest and most useful invention for the benefit of humankind. In spite of Franklin's apparent habit of keeping records and sharing information about his experiments through letters, it is of historic note that he never documented his kite experiment or wrote letters about the event. The only witness to the kite experiment was his son, William. And, the only written account of the kite and key experiment was by another person some 15 years after the event took place.

Franklin's lightning rod design was of an iron rod honed to a sharp point on its skyward tip. This design probably was developed as an extension of his laboratory experiments where he used an oversized needle with a sharpened point.

England was one of the last countries to begin installing lightning rods. The first lightning rod installation in England took place in 1760. The location was the Eddystone Lighthouse, an old wooden structure that had a history of damage by lightning. This lighthouse was vital to marine navigation safety and the commercial shipping interests of England.

As fate would deliver, nature further helped spur the use of Franklin's pointed lightning rod in England. In 1764, a bolt of lightning struck England's Saint Bride's steeple, already struck many times in the past. Saint Bride's church, a historic building in England, has a spire that stabs skyward prominently. The steeple's damage from lightning in 1764 prompted the Royal Society to recommend the use of Franklin's pointed lightning rods to protect the rest of the country's buildings.

As Franklin's pointed lightning rods were installed throughout Europe and America, King George III decided on a different approach. He decided to protect his palace with blunt-knobbed lightning rods, which he believed were superior to Franklin's. At the time, this also represented a political statement because King George viewed Franklin as a rebel. The use of blunt-knobbed lightning rods therefore represented a symbol of loyalty.

Regardless of King George III's, motivation to use blunt-tipped lightning rods in 1752, varying opinions and arguments have existed since that time, as to which lightning rod design is superior.

The controversy over which lightning rod was most effective was recently put to rest. In experiments conducted by the American Geophysical Union during the past decade, it was determined that the blunt lightning rod is more effective. Even though Franklin's design did not win this competition, he still deserves credit for bringing the lightning rod invention to the service of humankind. Besides the pointed and blunt versions of the lightning rod, there are many variations on the theme. Lightning rods have almost evolved into a folk art form, and many have considerable value to collectors.

On a closing note, as in the case of his inventions of the glass Armonica, bifocals, and the Franklin wood stove, Franklin never made money from his inventions, including the lightning rod. To quote Franklin, "As we enjoy great Advantages from the Inventions of others we should be glad of an Opportunity to serve others by any Invention of ours, and this we should do freely and generously."

This philosophy of Franklin's, about serving humanity, was also reflected in the words he wrote in 1790, the year of his death. From the tone of these words, it is evident that Franklin approached his earthly departure with philosophical humor:

> "I believe in one God, Creator of the universe....That the most acceptable service we render Him is doing good to His other children....As to Jesus...I have...some doubts as to his divinity; though it is a question I do not dogmatize upon, having never studied it, and think it needless to busy myself with it now, when I expect soon an opportunity of knowing the truth with less trouble."

WATER AND THE HISTORY OF MAN

ANN AZADPOUR-KEELEY
National Risk Management
Research Laboratory, ORD,
U.S.EPA, Ada, Oklahoma

JACK W. KEELEY
Environmental Engineer
Ada, Oklahoma

> And Isaac digged again the wells of water, which they had digged in the days of Abraham his father; for the Philistines had stopped them after the death of Abraham: and he called their names after the names by which his father had called them. (Genesis 26:18)

Before the earth was created, hydrogen, the oldest and most abundant material in the universe, and oxygen, created in the womb of stars, combined to form H_2O. Since the beginning of life on this planet and probably other planets scattered throughout infinite space, water has played a critical role in all facets of human existence, including the creation of life itself. We are made of water, and water produces the food we eat, from fish to fowl to fiber. It is the only substance on the earth that exists naturally in three phases, liquid, solid, and gas. As it serves us today in industry, agriculture, medicine, transportation, and the production of energy, it has served humans and all of the earth's creatures since the beginning of time.

In the days of Abraham, following family feuds and other conflicts, the victor often stopped wells by filling them with earth and stones. It was the equivalent of burning the village. Wells of water were so important that their owners gave them names, and communities which took root in the area assumed the names of the wells. In John 4:6, Jesus meets a woman from Samaria at Jacob's well and later reveals that He is the Messiah. The Christian Bible is replete with references to the importance of water to the fabric of life. It speaks of irrigation, of aqueducts, its creation, its purification, and even of it being turned into wine.

The Koran, the sacred book of the Moslems, was revealed to Mohammed by the angel Gabriel between A.D. 610 and 632. It mentions *water* 95 times in reference to such issues as springs pouring from rocks, hygiene, irrigation, pollution, prayer, thirst, and law. It even suggests that humans were made from water.

Hammurabi, the father of law, discussed the legal ramifications of water management in his Codes which were scripted 1700 years B.C. The Greeks constructed a 4200 foot tunnel or aqueduct to bring water to the city of Athens 2500 years ago, and the early Romans built 10 aqueducts which brought 220 million gallons of water to the city daily from mountain springs. The Aqua Claudia still stands in part around Rome, and the Pont du Gard in southern France was designed 20 centuries ago by the engineer Marcus Agrippa. The greatest water purveyor of all was the Zaghouan aqueduct in North Africa which crossed 130 miles of desert to deliver water to ancient Carthage. Even as the wheel was invented about 2000 B.C., the water clock (Clepsydra), fashioned by Plato, was first used to record time about 400 B.C., and the water mill came to be in the first century B.C.

Early civilizations took root near sources of abundant water: the Tigris–Euphrates Valley of Mesopotamia, the Nile Valley in Egypt, the Indus Valley in India, and the Hwang Ho Valley in China. These were the cradles of civilization which prospered, in great measure, because of the availability of water and the construction of extensive irrigation systems which made the land productive. In some cases, however, when the water systems failed, the civilization collapsed. It is believed, for example, that the Sumerian civilization of Mesopotamia failed because of poor irrigation practices which caused increased salinity in the soil and the eventual collapse of agriculture.

To an extent, the problems of managing a water system are mirrored today as major cities are forced to obtain water from greater and greater distances from the city. Often the search for available water as well as water rights litigation spans decades. The fall of groundwater levels increases pumping costs until irrigation is no longer cost-effective, and the effects of pollution increase both the cost of palatable water as well as the threat to human health.

Water has played an important as well as diverse role in many aspects of human history. It was the focus of religious rites, science, art, music, mythology, transportation, power, healing and architecture, to mention only the obvious. For example, the immersion into or the sprinkling of water have been forms of baptism long before the birth of Christianity. Ea or Enki had the distinction of being both the god of water and wisdom in several ancient religious peoples, including the Hittites and Babylonians and was often considered the creator of humanity. And, of course, the early scientists and philosophers such as Empedocles and Aristotle believed that everything was composed of only the four elements, earth, air, fire, and water. We all remember the story of the Greek mathematician Archimedes who formulated his fluid displacement law after stepping into a public bath and observed the overflowing water. He was so excited that he ran home naked crying Eureka! Eureka! He later used this water displacement knowledge to demonstrate that the king's goldsmith had adulterated a gold crown with less expensive silver.

For centuries, people around the world have bathed in mineral water to relax their souls and cure ailments. There are a number of religious shrines where the sick have been cured by holy water. One example is spring water near the site where Bernadette, a French saint, had a vision of the Virgin Mary in 1858. In music, we find Handel's Water Music (1717). Claude Monet produced a series of water-lily murals between 1900 and 1909, and in architecture Frank Lloyd Wright created the stunning Falling Water structure near Pittsburgh in 1936.

Even as water has given birth to all of the creatures of the earth and served them so well in many ways, so has it emulated the Apocalyptic horsemen of pestilence and death. Its overabundance has caused floods since the days of Noah, and its absence has resulted in drought and starvation for countless millions. Water from hurricanes and typhoons causes death and property damage of herculean proportions each year in all corners of the globe. Waterborne diseases (diarrhea is the leading cause of death in the world) cause 80% of the illness in Third World countries, and malaria from mosquitoes born in stagnant water affects 400 million people worldwide and results in 5 million deaths. The Industrial Revolution made sewers of streams and rivers.

The history of human association with water has been a union as blessed as it has been complicated, and on many occasions as wicked as Pandora's box. Perhaps it can best been capsulated in the words of John Dyer:

> And see the rivers how they run through woods and meads, in shade and sun, sometimes swift, sometimes slow, wave succeeding wave, they go a various journey to the deep like human life to endless sleep!

WATER, BACTERIA, LIFE ON MARS, AND MICROBIAL DIVERSITY

David B. Vance
ARCADIS
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Microbial life is extremely diverse with regard to ambient conditions, metabolic pathways, and geographic dispersion at depth in soils and sedimentary formations on Earth. Original bacterial life on Earth started in anaerobic conditions, without free oxygen in the atmosphere. Water chemistry in the primitive oceans would have been devoid of significant dissolved oxygen. The primitive atmosphere was chiefly carbon monoxide (CO), hydrogen sulfide (H_2S), nitrogen (N_2), hydrogen (H_2), and water. Some of the more volatile trace elements such as chromium, manganese, and sodium may also have been present (1). Since that time, almost every niche on Earth has been colonized by some type of microbial life. Given the robust nature of microbial life on Earth, it is possible that microbial life existed on other planets having similar elements and compounds. As the very early history of Mars may have been somewhat similar to that of Earth, Mars may have had the conditions for microbial life. In the oceans, the gases in the primitive atmosphere were present as dissolved constituents. In addition, given the reducing conditions, soluble forms of reduced metals were also present, with ferrous iron likely being a dominant species.

INTRODUCTION

From 3.5 to 1.5 billion years ago, the atmosphere of Earth was anoxic (primarily consisting of nitrogen and methane) and the dominant life forms were methanogenic, sulfate-reducing and iron-reducing anaerobic microbes. Near the end off this period, around 2 billion years ago, bacteria that could use nitrate to convert ferrous iron to ferric literally began to precipitate the soluble ferrous iron out of the ocean, forming the great Precambrian banded iron deposits and the later Paleozoic reworked redbeds.

Slowly starting about 3 billion years ago microbes that used the sun's light as an energy source became the dominant life form on the planet. These photosynthetic species began to become active and excrete oxygen as a waste product of their metabolism. Between 1.5 and 1.0 billion years ago, the "first" great environmental disaster occurred when photosynthetic life became abundant enough to elevate atmospheric oxygen to near current levels.

The type of bacteria that are indicated to have existed on Mars appear to be similar to those on Earth termed archaebacteria. These are bacteria with an ancient history, first appearing on Earth 3.5 billion years ago under the atmospheric conditions described above. Recent advances in genetic analysis indicate that these bacteria are a third form of life different from regular bacteria (Eubacteria) and all multicellular life (Eukaryotes). These archaebacteria survive through exotic metabolic processes that rely on a wide variety of chemical reactions to provide energy.

The archaebacteria fall into four major classes:

- Acidophiles that are able to grow in the pH range of 0.0 to 5.0.
- Thermophiles that can grow in a temperature range of +30 °C to +250 °C.
- Halophiles that survive in 10% to 27% sodium chloride solutions.
- Methanogens that are the best known archaebacteria and form methane in sediments and the intestinal tracts of animals.

MARTIAN ARCHAEACTERIA?

Also, modern (relatively speaking) Eubacteria are capable of adaptation to extreme environments, including:

- Alkalophiles that can grow in the pH range of 8.5 to 11.5.
- Barotolerant bacteria can exist under hydrostatic pressures up to 1100 atmospheres.
- Other chemoautotrophic bacteria such as denitrifiers, nitrate reducers, ferric iron reducers, sulfate reducers, hydrogen formers, sulfide oxidizers, and ferrous iron oxidizers.

The biochemical diversity is astounding. Just a few specifics are as follows:

- Bacteria have been found in aquifers at depths of 1.5 miles.
- They have been found in oil production formations at depths of almost 4 miles.
- Within the last year, bacteria have been found that survive using only surrounding rock and that generate hydrogen gas in the process; these are near 1 mile deep in basalt.
- Sulfate reduction is extremely common place in groundwater. Well water with the odor of hydrogen sulfide (rotten eggs) is a manifestation of bacteria-driven sulfate reduction.

Additional aquifer conditions under which it is common to find sulfate-reducing bacterial species are as follows:

- Eh conditions less than −200 mV.
- The presence of greater than 20 millimolar sulfate from geologic origins.

In the western United States, many aquifers are anaerobic and contain sulfate. Anthranilic acid, benzyl alcohol, and P-cresol have all been demonstrated to be biodegraded in the presence of sulfate-reducing bacteria. Research is currently under way evaluating the use of sulfate reducers for the degradation of chlorinated solvents and PCBs.

Two other dominant and effective anaerobic processes are denitrification and methanogenesis. In general, most aromatic hydrocarbons can be degraded anaerobically under methanogenic or denitrifying conditions.

Some thought should always be put into the exploitation of these alternative bacterial systems. The exact stoichiometrics should always be worked out with a complete assessment of all byproducts. For example, it is important to evaluate the overall electron flux in nitrate reduction systems:

For each mole of nitrate consumed, one mole of KOH or NaOH is generated, which in turn can lower aquifer pH and form limestone or other precipitates, that plug the infiltration or treatment zones. An acid must be added concurrent with the nitrate source in the appropriate amount.

Bacteria that produce hydrogen can sequentially dehalogenate chlorinated hydrocarbons to innocuous ethane or ethene gas.

Of course, all various metabolic systems are not equal with regard to energy yielded to the organism for each reaction that takes place.

On a scale of 1 to 10, with 10 having the highest energy yield:

- Aerobic oxygen based metabolism is 10.
- Nitrate reduction is 9.
- Iron reduction is 5
- Sulfate reduction is 4.
- Methanogenesis is 2.

Thus, the source of the dominance of aerobic life was precipitated by events 1 billion years ago (as described above). Given surface conditions, it is the most cost-effective metabolic system available. However, all other metabolic options still take precedent under the appropriate conditions. It has proven particularly difficult to drive hydrocarbon contaminated aquifers into oxic conditions mainly because of the poor transport properties of oxygen in water (the reason hemoglobin exists in our blood cells is to address that issue within our bodies). The power of many of the other metabolic systems is that they are driven by chemical species that are highly soluble in water (sulfate and nitrate, for example). The less-efficient energy yield is compensated by the dramatically improved transport properties.

In addition, these alternative metabolic systems are showing to have some capacity to degrade xenobiotic hydrocarbons that prove to be recalcitrant to aerobic attack.

The power of bioremediation lies in harnessing the vitality of bacterial life in the subsurface. The ubiquitous nature of bacterial life is constantly being revised upward. Archaebacteria exist under the extremes of physical/chemical conditions. Bacteria are found at such depths that it is almost certain they were laid down with the surrounding sediments and have been able to survive since. The potential presence of microbial life on Mars is an even more extreme example of the concept. *In situ* bioremediation is a process that appears to be viable almost under any subsurface conditions.

The history of Mars exploration has focused in a large part on the potential for past or current life on the planet. The first close-up picture of Mars was obtained in 1965. Since that time, a variety of satellites have traveled by Mars and a few spacecraft have landed on the surface of Mars. In 1976, the National Aeronautics and Space Adminstration's (NASA) Viking Mission to Mars had two spacecraft, *Viking 1* and *Viking 2*. Each craft consisted of an orbiter and a lander. NASA's Mars Global Surveyor began to orbit Mars in 1997 for the purpose of high-resolution imaging of the surface of Mars, studies of the topography and gravity, and the role of water and dust on the surface and in the atmosphere. This project also evaluated the weather and climate of Mars, the composition of the surface and atmosphere, and the existence and evolution of the Martian magnetic field (2). The Mars Pathfinder with the Sojourner Rover toured the surface of Mars in 1997. In 2001, the NASA Mars Surveyor 2001 lander was renamed the 2001 Mars Odyssey; the mission was to gather data to help determine whether the environment on Mars was ever conducive to life, to characterize the climate and geology of Mars, and to study potential radiation hazards on Mars.

The most recent information about Mars has come from NASA's two Rover Opportunity and Rover Spirit landing crafts, which inspected Mars during 2004. These remote studies indicate that the elements on Mars are the same elements found on Earth. On Earth, life started in water, and there is ample evidence of large amounts of water on Mars in the past (3).

Based on the similarities with Earth, the geologic conditions on Mars are still tantalizing scientists to examine the shallow Martian soil for any sign of past or current microbial life forms.

BIBLIOGRAPHY

1. Smith, D.G. (1981). *The Cambridge Encyclopedia of Earth Sciences*. Crown Publishers Inc./Cambridge University Press, New York.
2. National Aeronautics and Space Adminstration (NASA). (2004). Mars Global Surveyor. Available: http://nssdc.gsfc.nasa.gov/planetary/marsuv.html.
3. Baker, V. and Nummedal, D. (1978). *The Channeled Scabland: A Guide to the Geomorphology of the Columbia Basin, Washington*. Prepared for the Comparative Planetary Geology Field Conference held in the Columbia Basin, June 5–8, 1978. Sponsored by the Planetary Geology Program, Office of Space Science, National Aeronautics and Space Administration.

CANALS IN THE MEKONG DELTA: A HISTORICAL OVERVIEW FROM 200 C.E. TO THE PRESENT

David Biggs
University of Washington
Seattle, Washington

The Mekong Delta is one of the most productive agricultural regions in the world. It supports more than 20 million people living in a dense network of canals and creeks that today irrigate more than 2 million hectares of rice paddy. This landscape of dense settlement and intensive irrigated agriculture is relatively recent, although the earliest canal projects date to "Fu Nan" or Oc Eo culture circa 200 C.E. Today's water regime is not so much a single water system as it is a landscape built of multiple layers of canals in various stages of development and degradation. Some waterways continue to serve as primary transportation and irrigation works: others lie abandoned, their traces sometimes only visible through analysis of aerial photography or historical reports. This overview describes these historical layers of canalization and the context for some of the major projects in each period.

GEOLOGIC OVERVIEW

The Mekong Delta of today comprises roughly 6 million hectares (5 million in Vietnam and 1 million in Cambodia) of alluvial plains. It is bordered by the Gulf of Thailand in the west and the South China Sea in the east. The main channel of the Mekong River is believed to have formed sometime between 2 and 50 million years ago in the late Tertiary period. The older, higher terraces of its delta may have formed as early as 1 million years ago. Corings of sediment in the lower delta alluvium show alternating bands affected by freshwater and saltwater, suggesting that as recently as 6000 years ago, mean sea level may have reached a post-Ice Age high water point before declining to recent historical levels.

The slope of the delta is so flat that tidal surges affect water flow in canals and creeks up to 100 kilometers upstream. Two separate tidal regimes along the western and eastern coasts cause complicated periodic variations in water currents, depending on the proximity to both coasts. When opposing water currents meet, either at an intersection of waterways or along the path of a canal, they deposit sediment into bars that often block boat traffic. Vietnamese settlers, and probably earlier inhabitants, formed markets at these halfway points, convenient places for exchange. Choices for early modern settlements here were largely responses to water conditions and associated soil deposits.

From a soils perspective, the delta can be divided into subregions connected to geologic and ecological processes: a freshwater alluvial zone, the Dông Tháp floodplain, the Long Xuyên Quadrangle, the eastern coastal zone, and the Cà Mau Peninsula (Fig. 1). The Mekong River deposits fresh sediment along the banks of the channel, creating alluvial banks that rise 2–3 meters above mean sea level. This freshwater alluvium area fosters the development of the richest natural paddy in the delta, allowing two and three rice crops per year. The combination of high banks and good drainage allows farmers to irrigate easily and also to raise fruit trees without fear of flooding. Ancient alluvium is characterized by higher soils associated with a postulated high seas episode that occurred roughly 6000 years before the present. In the present era, this higher land forms a natural barrier to the Dông Tháp floodplain, a depressed area characterized by high soil acidity and annual floods. The other large flood zone is the Long Xuyên Quadrangle south of the main river channel. The Quadrangle is remarkable for granite and limestone

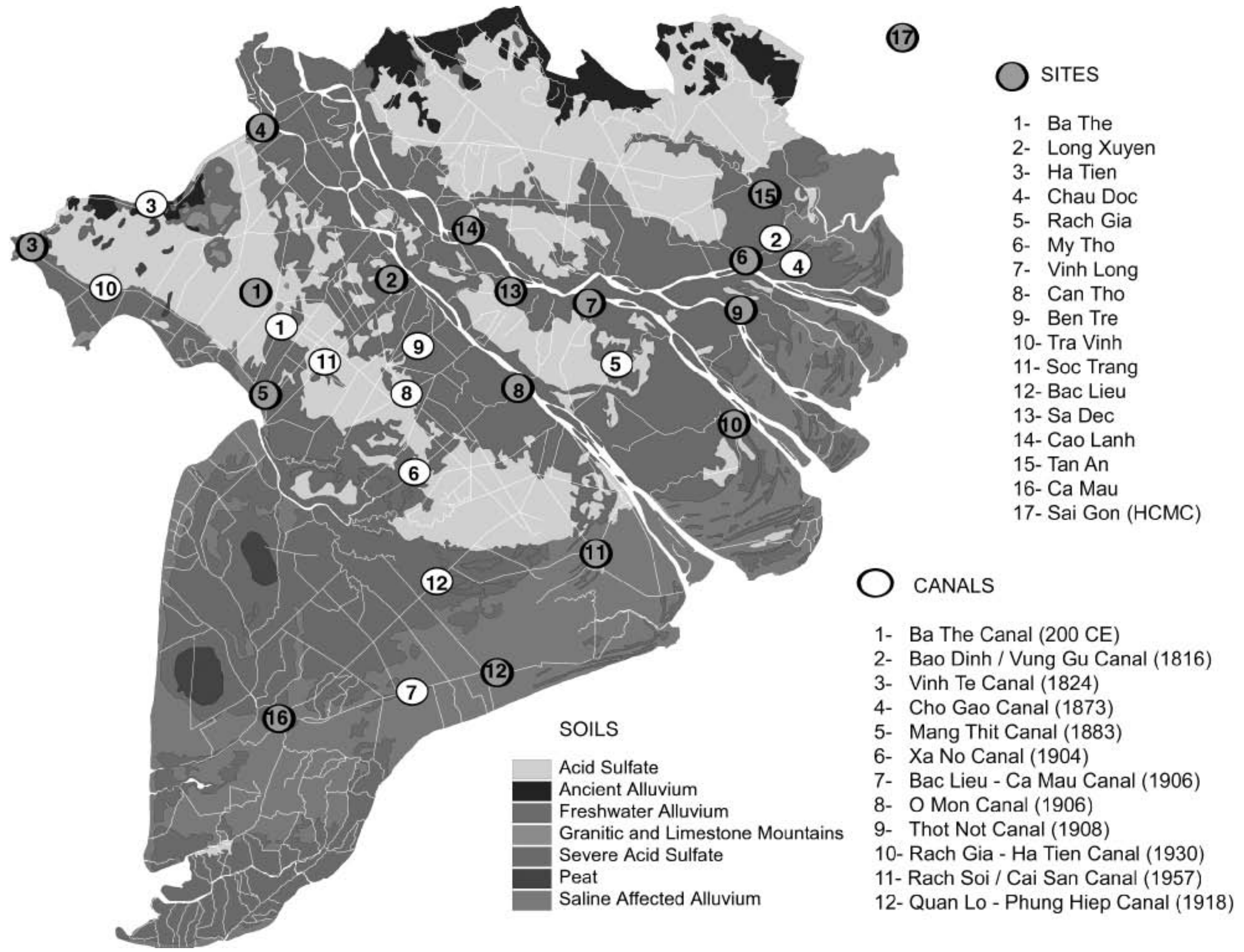

Figure 1. Mekong Delta divided into subregions.

mountains that rise up dramatically from the plains some 200–700 meters to form islands on the land. At the base of Ba Thê mountain in this Quadrangle lie some of the most substantial sites for what George Coedès described as the “Fu Nan Kingdom”: the first example of an Indianized kingdom in Indochina.

Both of these two depressed floodplains, whose average altitudes are less than 1 meter above mean sea level, form two flood sinks that form virtual, inland seas each September as the Mekong River reaches its annual flood stage in the wet season monsoon. The annual flooding of the Mekong from August to October is important for far-reaching effects on almost all activity in the region. The Tonle Sap (Great Lake) in Cambodia grows to three times its dry season size in June, July, and August. As the lake reaches its limit, water flow reverses in the Tonle Sap River, flows downstream, and recombines with the mainstream rivers at Phnom Penh. The river then floods the Long Xuyên Quadrangle and the Dông Tháp Plain.

ANCIENT SETTLEMENT

The earliest traces of organized human settlement in the Mekong Delta date to the “Fu Nan” period (100 B.C.E.–550 C.E.). The types of evidence found in these settlements can be grouped roughly into wooden piers and structures, jewelry and statuary, and vestiges of infrastructure—brick foundations, brick towers, and canal traces. The largest brick foundations lie at the Ba Thê site in the Long Xuyên Quadrangle. Using aerial photography in 1931, an amateur archeologist, Pierre Paris, discovered traces of canals crossing the western delta, from Rạch Gía on the coast to Angkor Borei about 100 km to the northwest. What had previously been noted through ground surveys as straight creeks were actually the remaining traces of human-constructed waterways. Louis Malleret, a former director of the Ecole Francaise de l’Extrême-Orient, led digs in the 1940s and 1950s at canal sites discovered through aerial and ground surveys. Malleret’s work and that of successive Vietnamese and international researchers has proven through comparison of pottery shards and other means that these canals dated to roughly 200–400 C.E. Styles in the pottery indicate containers for oils and other goods that resemble similar containers from eastern India in that time. The ancient canals also extended deep into the Cambodian portion of the Mekong Delta, connecting the gulf coast ports with inland river ports at Tak Ev and Angkor Borei.

In the ancient world, the Mekong Delta was an important halfway point for sailing ships travelling with the monsoons between India and China. As a possible warehousing center developed along these ancient canals, the Oc Eo site allowed ships from both destinations a useful trade interchange and provided them with important naval stores and valuable supplies brought downriver from the forest interior: rhinoceros horn, elephants, feathers, aromatic woods, and wild spices. Chinese records from both trade missions and accounts of Fu Nan monks visiting China note major settlements along the main course of the Mekong River below Phnom Penh, including the sites at Oc Eo/Ba Thê, Angkor Borei, and a site in Dông Tháp Muòi. Inscriptions described Indian merchant colonies at Southeast Asian ports, including those in the Mekong Delta. The situation of Indian merchants here enriched local rulers, allowing them to raise funds and craftsmen to build ships of their own.

Environmental factors may have played a role in the decline of pottery and evidence of population in the lower delta. Soil cores at the Oc Eo site reveal high concentrations of Rhizophera pollen spores in deeper layers, suggesting that saltwater mangrove trees populated a floodplain that is entirely freshwater today. Every year, the sea coast may have receded further from the Oc Eo site, and the canals may have slowly filled with sediment from the rivers, turning them into shallow creeks and ponds. This would in turn change the wetlands ecology to one favorable for freshwater tree species. This freshwater plant community would have also attracted mosquitoes that carry endemic malaria and parasites, thus putting dense populations at greater risk of disease outbreaks.

VIETNAMESE SETTLEMENT

There is no further evidence of major canal projects or dense settlement in the Mekong Delta for more than 1000 years after the decline of “Fu Nan.” From the ninth to fourteenth centuries, the upstream kingdom of Angkor controlled the entire region and conducted its major trade routes by highways that stretched westward into present-day Thailand and to the west coast of the Isthmus of Kra. Archeological evidence near present-day coastal towns of Trà Vinh and Sóc Trang suggest that river ports here conducted small-scale local trade with Southeast Asian and Chinese merchant ships. One possible explanation for the lag in development of the Mekong Delta after Fu Nan could be that the very same waterways and flat landscape that facilitated easy transportation also allowed easy attack by enemy fleets. Steles at Angkor Thom (Siem Reap) depict naval battles with ships from Champa that attacked the Khmer capital by way of the Mekong River. The numerous islands in the Gulf of Thailand also provided easy harbor for all manner of pirates—Malay, Siamese, Chinese, European. Until Vietnamese settlers began migrating southward to the delta in the seventeenth century, island-like Khmer trading ports and administrative outposts of the Khmer kingdom characterized the area. Populations of Khmer villagers lived in some of the richest, naturally irrigated alluvial regions in areas that are today still ethnically Khmer: Sóc Trang, Trà Vinh, Trà Nóc, and Sa Déc.

One of the most important factors in Vietnamese settlement and early modern canalization was the increasing importance of rice as a trade commodity. Until the seventeenth century, rice was rarely traded by ship and in cities as were gold, spices, ivory, and aromatic woods. It was a local subsistence commodity and generally in greater abundance than the human labor necessary to cultivate it. In the seventeenth century, increasing population pressures in the Red River Delta brought on widespread famines and social unrest; to the present day, this population density in the north and the "southward progression" have played important roles in spurring development of waterways and infrastructure in the Mekong Delta. The upset of the ruling Minh Dynasty in 1644 also led to a mass exodus of Minh loyalists to Vietnam and other kingdoms that granted them asylum. In 1679, 3000 Fukien troops arrived in the Vietnamese kingdom. The Vietnamese ruler, anxious not to have them so close to his capital, sent them to the south where they formed the commercial towns of Biên Hòa, Sài Gòn, and Mỹ Tho. In 1681, another Chinese emigré, Mặc Cúu, immigrated to Cambodia and succeeded in winning a concession from the Khmer king to develop a trading port at Hà Tiên. From 1690 to 1780, Hà Tiên became a virtual city-state, trading most of the forest products of Cambodia by a Chinese-controlled shipping network centered at this Gulf Coast port. The more powerful kingdom of Siam now controlled the Isthmus of Kra, so Hà Tiên became an important trading port much in the same way as Oc Eo before it.

The Vietnamese kingdom established military and administrative posts in the Mekong Delta in the eighteenth century. In 1732, military campaigns led to the establishment of camps at Mesar (Mỹ Tho), Kampong Reussey (Bên Trè), Koh Gong (Gò Cong) and Peam Ba-rach (Long Xuyên). The court sent troops and new migrants to establish military plantations [dôn diên] in the new territories. These military plantations functioned as both garrisons and farming communities. At the same time, the Chinese developed mercantile posts and arranged the marketing of rice. The Tây Son civil war (1789–1802) and repeated Siamese invasions at Hà Tiên (1705, 1717, and 1771) led to the eventual abandonment of Hà Tiên and the establishment of Vietnamese control of the delta in 1802. For 60 years, the Court enticed settlers to establish garrisons and villages. In 1818, still worried about the threat of Siamese invasion, King Gia Long ordered his military governor in the delta, Thoại Ngọc Hâu, to assemble an army of conscript laborers to dig a canal 67 kilometers from a new river citadel at Châu Dôc to Hà Tiên. More than 50,000 laborers worked for 5 years and through several Khmer-led rebellions to complete the waterway. This strategic project was intended to allow greater Vietnamese control of the Gulf Coast and direct water routes to river garrisons. Two other major projects in this period connected the increasingly dense Sài Gòn watershed with the intermediate Vàm Co watershed and the Mekong river. The Bạo Dình Canal in 1816 allowed more direct access to the Khmer territory east of Mỹ Tho and solidified trading routes and administrative communication in the eastern delta. The Thoại Hà (Long Xuyên–Rach Giá) Canal was also completed in 1816. The rapidity of this project (3 months) and its location in the heart of the former "Fu Nan" kingdom suggests that this waterway may have dated to a much earlier period and only been "cleaned" in 1816. Historical records in the nineteenth century do not specify this difference, and they often alternately describe canals as rivers.

Midcentury violence led to the abandonment of coastal outposts for periods of years, and the neglect of these new waterways led to their being choked off with silt. The natural processes of sediment accretion and growth of aquatic plants played an important role in canal projects, especially in the costs of their upkeep. The interplay of tidal pulses and opposing river currents led to formation of transverse sediment bars across the channel of a canal. In a matter of months, if not continually cleaned by hand, these bars would reduce the navigable channel to a depth of half a meter, even at high tide in the dry season. Besides this hydraulic challenge, the water hyacinth or luc bình could quickly reproduce and choke off an entire waterway with a thick, floating mat of roots on the water's surface. Thus, without the continuous application of human labor in large numbers, most artificial waterways could not remain navigable for long. With each social uprising, canals quickly deteriorated so that after a few years, the costs of redredging would approach the original cost of construction. Another theme of canal degradation that also continued through the colonial and postcolonial eras was its role in aiding rebel groups.

With the French colonial conquest beginning in 1858 and lasting until 1867, rebel groups could only survive in the deep swamps far away from deeper rivers and streams. Battles at the Plain of Reeds in 1873 proved that control of waterways was vital for the French colonial regime. For two decades, the French colony's ruling admirals focused their efforts on linking strategic areas with the colonial metropole at Sài Gòn. In 1875, Admiral Duperré organized a permanent committee to improve naval communications between Sài Gòn and the Mekong Delta. This commission surveyed existing and proposed routes until 1879. During this time, the colonial government mobilized armies of corvée [conscript] laborers into government service. These first canals included Trà On (1876), Chò Gạo (1877), Set Say (1878), Phụ Tục (1878), Mirador (1879), and Saintard (1879). Of these, the Chò Gao Canal was one of the most vital. Via several large waterways, it linked Sài Gòn with Mỹ Tho, the largest river port at the time. A gunboat or cannonière could now make the trip to the delta in just 8 hours.

The same problems of sediment accretion and controlling vegetation continued to dog French engineers as it had their Vietnamese predecessors. In just a few years, Chò Gạo was silted again, preventing passage of the larger boats and forcing them to make the much longer journey to the river port by entering the river from the sea. In 1883, the government paid the Eiffel company to construct steel railway bridges across the intermediate Vàm Co River to connect Sài Gòn with Mỹ Tho by train. As the colonial government grew stronger in the 1880s and the French succeeded in using mechanical dredges at Suez, the regime introduced mechanical dredges. The shallow and high-clay sediments here, however, forced the government to rely on labor crews until new dredges were built in the 1890s.

The period of colonial dredging contracts with a mechanical dredging firm lasted from 1894–1945. During the peak of colonial expansion from 1894–1930, the colonial regime dredged some of the major "transverse" waterways connecting the lower branch of the Mekong or Hậu Giang with the Gulf of Thailand. This lower part of the delta was the least inhabited, making it the easiest space for French colons to build plantations several thousand hectares in size. Major canals such as Xà No (1904), Lai Hiêu (1906), Thôt Nôt (1908), Ô

Môn (1906), Vinh Tê (redredged,1914), Quạn Lô–Phụng Hiêp (1918), Ba Thê (redredged,1923), and Tri Tôn (1928) were developed. They also, dredged important waterways into Dông Tháp (1890–1904), from Bac Liêu to Cà Mau (1906), and from Rạch Giá to Hà Tiên (1930). This massive effort in mechanical dredging was accompanied by the immigration of laborers from the north; the population increased from 500,000 to almost 4 million in 1930.

Some areas, especially the fields near Cân Tho, flourished, but many of the waterways proved to be both economic and ecological disasters. The Lai Hiêu Canal brought on higher flooding, and its embankments cut off existing water circulation that led both to the spread of cholera in 1908 and the failure of many preexisting fields. A series of typhoons from 1904–1908 caused abnormally high flooding in the Plain of Reeds and the abandonment of plantations and fields after several years of lost crops. Canals also led to drying of soils in floodplains that were high in sulfate content. This drying of soils produced acid sulfate that then leached into the waterways and in turn quickly reduced crop productivity. With the onset of the world depression in 1931 and the collapse in rice markets, the masses of migrant laborers employed as tenant farmers on the plantations began to protest. Hydraulic engineers also began to debate the future of projects in the area. One camp argued that it was necessary first to improve existing yields and to let some of the failed projects lie fallow. Another camp argued for more sophisticated mechanical approaches to drainage and irrigation, suggesting the redevelopment of casier or grids of irrigation canals where farmers could be resettled. Except for small settlement projects near Ba Thê, the colonial government did not develop any other large projects before it was ousted from power in 1945 and subsequently entered into a 9-year war with the Viêt Minh.

During the Second Indochina War (1945–1954), the Government of South Vietnam attempted to develop several of these casier styled resettlement projects in a fashion similar to those proposed by the French before them. Of these, the single permanent project was the Cái San Project, a series of "transverse" canals in the lower delta and a smaller network of intersecting canals numbered from 1 to 17 and letters A to H. From 1955–1962, the Government resettled approximately 50,000 settlers, mostly Catholic, from their original villages in North Vietnam to 3-hectare tracts at Cái Săn. During the 1960s, American military and private contractors further built up the area with a highway paralleling the Rạch Sọi Canal (Highway 80) and an airport at the southern end of the area (today Rạch Giá Airport). For the most part, both wars prevented steady extension of waterways due to continued guerilla attacks. Like preceding antistate rebels, preventing easy circulation of government troops on the waterways and (increasingly) roads was vital to survival. Most of the familiar scenes of combat during the "American War" were encounters in the swampy interiors of swamps where previous canals had either deteriorated or been abandoned because of their low agricultural utility.

Since the conclusion of the war in 1975, and especially with the market orientation stimulated by Dôi Mói Reform in 1986, the Government today has begun a series of new projects. Like projects enacted in the 1950s, many of these contemporary development projects were continuations of earlier plans developed in the 1960s and early 1970s. For example, in 1996, the Government began work on the Mỹ Thuận Bridge, a project first surveyed by the Nippon Koei Corporation in 1963. This bridge, completed in 2001, became the longest span bridge in Southeast Asia and for the first time, allowed continuous highway transport across the upper branch of the Mekong River. A second span is due to be completed near C`ân Tho in 2005. Now at peace, government leaders have refocused their attention on Dông Tháp and former guerilla strongholds for future agricultural development. The Vietnamese delta population now is well over 17 million persons, however, so decisions to build canals and flood dikes are becoming increasingly difficult. Flood control in one area often results in diversion of the floodwaters to another zone or province downstream, thus giving rise to interprovincial disputes over access to clean water. A related issue is the construction of dams on the upper and middle Mekong Basin. The Mekong Committee is an international advisory body made up of representatives of the six riparian countries. With the reunification of the region and all member states into both the Mekong Committee and ASEAN, downstream countries have begun to lobby against diversionary projects upstream.

Another future concern in the Mekong Delta is the impact of global warming. As sea level rises just a few feet, thousands of hectares along the coast would be permanently inundated with salt water. The Mekong Delta is one of the flattest deltas in the world, so fluctuations in both river levels and sea levels could require the construction of unprecedented sea dikes to protect areas of coastal rice paddy.

The apparent "system" of canals and water projects in the Mekong Delta reflects nearly 2000 years of experimentation. Some of those trials lasted into the present, and some others are barely visible now. The economic or environmental health of this system is largely dependent on regional social factors and the intensity of human activity here. Currently, the challenge for engineers and planners in this region is to reconcile these factors to distribute "clean" water equitably in a more predictable fashion.

READING LIST

Biggs, D. (2003). Problematic progress: Reading environmental and social change in the Mekong Delta. *J. Southeast Asian Stud* **34** (1): 77–96.

Coedes, G. (1967). *The Indianized States of Southeast Asia*. Translated by Susan Brown Cowing. W.F. Vella (Ed.). East-West Center Press, Honolulu, HI.

Paris, P. (1931). Anciens canaux reconnus sur photographies aériennes dans les provinces de Ta Kev et de Châu Dôc. *Bulletin de l'Ecole Française de l'Extrême-Orient* **XXXI**: 221–225.

Xuân, V.T. et al. (1998). *Development of Farming Systems in the Mekong Delta*. Hô Chí Minh City Publishing House, Hô Chí Minh City.

CONFLICT AND WATER USE IN THE MIDDLE EAST

CLIVE D. LIPCHIN
Ann Arbor, Michigan

INTRODUCTION

Freshwater is unevenly and patchily distributed across the surface of the earth, some regions have either ample supply or unlimited access, and other regions suffer from unpredictable rainfall patterns and long periods of drought. As human populations grow and their activities continue to alter environmental patterns, the demand for freshwater will increase as the availability and quality of freshwater decreases. These conditions make water-related conflicts more likely between nation states that are forced to share a common water resource under the strain of locally increasing populations and increasing regional demands for water. In the semiarid and arid regions of the globe, especially the Middle East, conflict over water seems a more and more likely prospect (Table 1).

Where water is scarce, competition for limited supplies can lead nations to seek access to water as a matter of national security. As recently as the mid-1980s, U.S. government intelligence services estimated that there were at least 10 places in the world where war could break out over dwindling shared water resources; the majority of these locations are in the Middle East. Jordan, Israel, Cyprus, Malta, and the countries of the Arabian Peninsula are sliding into the perilous zone where all available fresh surface and groundwater supplies will be fully used. In Israel, the Palestinian Authority, and Jordan, a water deficit already exists, and all available water resources are overexploited.

The conflicts in the Middle East are probably unique in the world and are many faceted and complex. As the birthplace of modern Western civilization and three of the world's largest religions, conflicts, whether political, racial, or religious, have been a part of the system since civilization first began between the twin rivers of the Tigris–Euphrates. The ongoing conflicts taking place on today's political stage can be traced back to Biblical times and beyond. In the face of such an historic record of conflict, solving disputes over shared water resources automatically involves cultural, economic, and political factors.

Table 1. Percapita Water Availability in 1990 and 2025[a]

Country	1990	2025, Cubic Meters Per Person Per Year
Kuwait	75	57
Saudia Arabia	306	113
United Arab Emirates	308	176
Jordan	327	121
Yemen	445	152
Israel	461	264
Qatar	1171	684
Oman	1266	410
Lebanon	1818	1113
Iran	2025	816
Syria	2914	1021
Iraq	5531	2162

[a] Reference 1.

The emergence of modern independent states since the decline of colonialism in the region has strengthened nationalistic ideals and exacerbated common property issues that transcend their borders. Given the importance of the region as the world's main supplier of fossil fuels, the international community has been quick to realize the delicate power balance in the region and how precarious and volatile it can be. In the current global political environment, it is in the world's best interest to foster cooperation among the states in equitably managing their shared waters.

Across the globe, many rivers, lakes, and groundwater aquifers transcend arbitrary political borders. This geographical fact has led to the geopolitical reality of disputes over shared waters; the situation in the Middle East (especially the Nile, Jordan, and Tigris–Euphrates) is just one such example. Other regions where water disputes have or are likely to develop are the Indus, Ganges, and Brahmaputra in southern Asia; and the Colorado, Rio Grande, and Parana in the Americas.

Water is a fundamental resource vital for survival, and it also provides a source of economic and political strength. Under these conditions, ensuring access to water provides a justification for going to war, and therefore water supply systems can become a goal of military conquest (2). Gleik (3) outlined the following characteristics that make water likely to be a source of strategic rivalry:

1. the degree of scarcity,
2. the extent to which the water supply is shared by more than one region or state,
3. the relative powers of the basin states, and
4. the ease of access to alternative freshwater sources.

The Middle East and its many ideological, religious, and geographical disputes fits within this framework.

SETTING

The current water crisis plaguing the region can be traced back to the arbitrary political division of the region into competing states by the withdrawal of the post-World War I colonial powers (4). The subsequent map of the region disregarded the issue of water in forming possible natural political boundaries. Except for Lebanon, none of the states that the colonial powers created was provided with independent water resources, and no mechanisms were established for coordinating the use of internationally shared water resources (4). The Tigris–Euphrates Basin was sectioned irregularly and placed in the domains of three competing states (Turkey, Syria, and Iraq); the Jordan River was divided up among four states (Israel, Jordan, the Palestinian Authority, and Syria) (4). The Nile River Basin is shared by no less than nine states, and the balance of power is heavily skewed toward the two

Table 2. International River Basins in the Middle East[a]

River Basin	Total Area of Basin, Square Kilometers	Countries in Basin	Area, Square Kilometers	Percentage of Total Area
Tigris	378,850	Iran	220,000	58
		Iraq	110,000	29
		Turkey	48,000	13
		Syria	850	<1
Euphrates	444,000	Iraq	177,000	40
		Turkey	125,000	28
		Syria	76,000	17
		Saudi Arabia	66,000	15
Jordan	19,580	Jordan	7,650	39
		Syria	7,150	36
		Israel	4,100	21
		Lebanon	950	5
Nile	3,031,000	Sudan	1,900,000	63
		Ethiopia	368,000	12
		Egypt	300,000	10
		Uganda	233,000	8
		Tanzania	116,000	4
		Kenya	55,000	2
		Zaire	23,000	1
		Rwanda	21,500	1
		Burundi	14,500	<1

[a] Adapted from Reference 1.

economically and militarily powerful downstream states, Egypt and the Sudan (Table 2).

An example to illustrate the tensions inherent in water resources development in the region is the conflict that took place between Israel and Syria. In the 1950s, Syria tried to prevent Israel from building its National Water Carrier, an extensive canal and piping system to provide water to the extremely arid areas of southern Israel. When Syria attempted to divert the headwaters of the Jordan in the mid-1960s, Israel used force, including air strikes against the diversion facilities, to prevent their construction and operation (5). These military actions, it is thought, contributed to the outbreak of the 1967 Six-Day War.

Having more mouths to feed, the allocation of freshwater becomes a pressing problem. More mouths means more food, and so, more water needs to be allocated to agricultural production. This issue is especially true for Turkey, Syria, and Iraq, which have all embarked independently on ambitious irrigation projects on the Tigris–Euphrates river system to increase their agricultural capacity (6). Water allocation, for the riparian states along the Jordan River, has become a balancing act between agriculture and domestic use as more of the population becomes urbanized. The allocation of water can lead immediately to issues of conflict as upstream states control the flow of a river course and inadvertently disadvantage those downstream that depend on the same flow.

The Tigris–Euphrates headwaters originate in Turkey but empty into the sea at the Persian Gulf in Iraq, flowing through Syria on the way to the sea. Egypt is entirely dependent on the Nile River for practically all of its water requirements but is at the mercy of the upstream states (especially Ethiopia) that contain the headwaters of the Nile, the Blue and White Nile, within their borders. Israel, Jordan, and the Palestinian Authority are in an even worse position. They possess few internal renewable water resources and all depend on one river system, the Jordan and associated aquifers, which has its headwaters originating outside of their territories in Lebanon and Syria.

Wolf (7) has provided a framework for analyzing transboundary waters and the potential for conflict. Within this framework, the problem, background, attempts at conflict management, and the potential for a positive outcome where conflict can be averted is provided. Wolf's framework is applied here in analyzing the three primary water systems in the Middle East: the Tigris–Euphrates, the Jordan, and the Nile.

THE TIGRIS–EUPHRATES RIVER BASIN

The Problem

In 1975, unilateral water developments came close to provoking warfare among the riparians of the Tigris–Euphrates (Turkey, Syria, and Iraq). Since the 1960s, population pressures drove the three states to pursue water development projects independently, primarily for agricultural expansion. Specifically, these were projects in southern Anatolia in Turkey, the Keban Dam (1965–1973), and in Syria, the Tabqa Dam (1968–1973) (7).

Background

From source to sea, the Euphrates is the longest river in the whole region of West Asia (8). It traverses a distance of 2,700 kilometers, of which some 40% is in the modern state of Turkey, 25% in Syria, and 35% in Iraq. Its twin, the Tigris, has a total length of 1,900 kilometers, of which about 20% lies in Turkey, 78% in Iraq, and only 2% along the pointed northeastern corner of Syria, known as the “Duck's Beak” (4).

Probably the most ambitious development project carried out in the region is Turkey's South East Anatolia Project known by the Turkish acronym GAP (for Guneydogu Anadolu Projesi, in Turkish). The project aims to develop the regions bordering Syria and Iraq that encompass the headwaters of the Tigris and Euphrates. The area is sparsely inhabited by a population of around 6 million; the majority consists of ethnic Kurds, who claim the region as part of their national homeland. By transforming this semiarid region into the country's breadbasket, Turkey hopes to encourage and promote the development of agriculture and industry and also to offset the Kurdish majority by attracting Turks into the area and so diluting Kurdish nationalistic claims (6).

The Turkish government has promoted the GAP project for its great eventual benefits for all, but recently, more than 250,000 people were displaced from their homes, which were to be inundated by dams. A conflict broke out between the Turkish government and the Kurdish Workers Party (PKK), which has repeatedly opposed GAP and regarded it as a Turkish theft of Kurdish waters (4).

The GAP plan calls for constructing 80 dams, 66 hydroelectric power stations that have a total capacity of 7700 megawatts, and 68 irrigation projects covering up to 2 million hectares. Among the principal dams are the Keban, the Karakaya, and the Ataturk, which is the linchpin of the entire GAP complex (4).

Despite its great agricultural potential, Iraq has, in the last decade, become a net importer of grain to feed its population of 26 million. Iraq's main irrigated lands are the regions of lower Mesopotamia. However, waterlogging and salinization plague this region and therefore an imperative need has developed for a coordinated scheme of rehabilitation and sustainable water management, especially to provide regional drainage.

In 1953, Iraq began work on a regional canal to provide comprehensive drainage. This canal has been dubbed "The Third River" because it flows midway between the Tigris and Euphrates. During the Gulf War, bombing caused great damage to Iraq's hydraulic works, but it managed to complete the project in December 1992.

The project was designed to drain 1.5 million hectares of land, allowing Iraq to increase its domestic food output significantly and thus help achieve its aims of economic independence. However, like Turkey's GAP project, Iraq's third river has displaced the indigenous Marsh Arabs, and the UN has criticized the canal as an environmental crime that threatens to destroy an entire ecosystem (6). An ulterior motive for draining the region has been the possible exposure of oil deposits that it is thought underlie the marshes and are considered as extensive as those in Kuwait.

More than Turkey and Iraq, Syria depends nearly entirely on the flow of the Euphrates for the development of its economy. Apart from the Euphrates, Syria has access to a few underground aquifers from which, through the sinking of thousands of wells, has already caused overdrawing and salinization of the groundwater. In 1974, Syria began its greatest engineering feat, the damming of the Euphrates and the creation of Lake Assad. The dam, built with Soviet engineering and financial aid, was supposed to irrigate some 400,000 hectares, generate electricity, and make the region prosperous (4). In this project, as for the projects described, 70,000 indigenous Bedouin were displaced. In addition, the project has turned out to be a disappointment; the imported Soviet water generators proved faulty, and the land originally designated for irrigation proved unsuitable because the soil contains large amounts of gypsum (4).

Attempts at Conflict Management

According to Shultz (8), the real problems that the three states currently face are management, apportionment, and development planning, which have led to disagreements among them. This joint dependency on the river waters clearly indicates that the national security of the three states are linked. Turkey, Syria, and Iraq may, therefore, be considered to form a hydropolitical security complex. A hydrosecurity complex is defined as those states that are geographically part owners and technically users of the rivers and, as a consequence, consider the rivers a major security issue (8).

Disputes over the Tigris–Euphrates waters consist of conflicts between upstream and downstream neighbors, but also between the state and ethnoreligious groups (Kurds, Marsh Arabs, and Bedouins). Both Syria and Iraq depend on Turkey for the continual flow of the Euphrates. The intensification in using the Euphrates for Turkey's GAP project has caused sporadic decreases in flow in both Syria and Iraq. Such a decrease in flow is especially important for Syria, as, unlike Iraq, it cannot rely on the flow of the Tigris. The issues of an amicable apportionment of the river's flow could probably be easily met because there is enough water for all three states (8). Surprisingly enough, however, no tripartite agreement has yet been signed by the three riparians on flow regulation, dams, and sustainable water management.

The issue in contention then is how to weigh historical rights against proportionate contributions to flow, taking into consideration the real needs of each country. These are the availability of energy, the need for hydroelectricity, the feasibility of developing economic alternatives to irrigation farming, the efficiency of water use, and the size of each country's population. In 1980, a Joint Economic Committee was established between Turkey and Iraq, which included meetings related to water resources. Syria began participating in 1983, but meetings have been sporadic at best (7).

Outcome

Turkish and Syrian leaders had agreed to resolve the issues of water allocation by the end of 1993, but no agreement has yet been negotiated. Thus far, it seems that the levels of mistrust among the countries makes any serious settlement unlikely in the near future.

THE JORDAN RIVER WATERSHED

The Problem

Ever since the mid-1950s, the issue facing the riparians of the Jordan River watershed has been one of recognizing rights and equitably apportioning surface and groundwaters. An agreement came close to being signed by the riparians in the early 1950s, which was based on the Johnson accords, named after U.S. special envoy Eric Johnson. Although Israel agreed to sign the accords, the Arab states officially rejected the agreement but have unofficially used the accords as a basis for ongoing negotiation. Since 1991, beginning with the Arab–Israeli peace accords, especially those between Israel and Jordan and between Israel and the Palestinian Authority, progress has been made on rights, use, and apportionment. The Oslo accords and the Israel–Jordan Peace Treaty of 1994 both included means for joint water resource management. Today, the ongoing Palestinian–Israeli violence has made the agreements signed at Oslo exceedingly difficult to implement.

Background

The Jordan Basin is an elongated valley in the central Middle East. Draining some 18,300 square kilometers, it extends from Mount Hermon in the north to the Dead Sea in the south and lies within the pre-June 1967 boundaries of Israel, Jordan, Lebanon, and Syria. Its waters, which originate in rainfall and in rivers and streams of Lebanon, Syria, and the Golan Heights, drain the lands to the east and to the west of the Jordan Valley. Precipitation in the basin ranges from more than 1000 mm/y in the north to less than 50 mm/y in the south but averages less than 200 mm/y on both sides of the Jordan River (9). Much of the basin is arid or semiarid and requires irrigation for agricultural development.

Critical to the region's water supply is access to the water resources of the Jordan and also to the underground aquifers of the region. Such access is especially crucial for Israel because 60% of its total potential water supplies are in the region's two largest aquifers, the coastal plain aquifer and the mountain aquifer (6).

As early as the mid-1970s, Israel has been overabstracting water from the coastal aquifer, causing saltwater intrusion. The Gaza Strip also depends on this aquifer, from which it is estimated that 65 MCM/y (million cubic meters per year) is available. Rapid population growth, however, has severely taxed the system and caused lowering of the water table and severe intrusion of seawater into wells. Consequently, most of the local drinking water in Gaza exceeds the salinity level of 500 ppm, considered the upper threshold for safe drinking water; in some areas, the salinity level has reached 1500 ppm and made drinking water next to intolerable (4).

The mountain aquifer or Yarkon–Taninim aquifer extends beyond the Green Line (Israel's pre-1967 borders) into the mountainous regions of the West Bank. The aquifer extends about 150 kilometers from north to south. The aquifer's sustainable yield of fresh water is estimated at about 300 MCM/y. It is tapped by more than 300 wells, whose total annual pumpage averages 375 MCM. Like the coastal aquifer, the mountain aquifer is suffering from overpumping and a decline in water quality.

Since the 1950s, Israel and Jordan have pursued unilateral plans for developing the Jordan River Basin. Israel drained the Huleh swamps to create arable land and constructed the National Water Carrier that pumps water from the Sea of Galilee to the central and southern regions of the country. Jordan created the King Abdullah canal on the eastern shores of the Jordan for irrigation and drinking purposes. Both of these projects have, in effect, dried up the natural flow of the Jordan to its terminus at the Dead Sea.

Attempts at Conflict Management

Although never signed, the Johnson Plan of 1955 offered a regional approach to the management of the Jordan, although groundwater was not considered. Today, groundwater is a key point in negotiations between Israelis and Palestinians because groundwater development will be integral to creating a Palestinian State.

Outcome

Since the Madrid talks of October 30, 1991, Israel and the Arab states agreed on the need for multilateral negotiations on water resources. This multilateral track runs in parallel with the bilateral negotiations on a political settlement between Israel and the Palestinian Authority (now stalled) and Israel and Syria (also stalled). The multilateral track on water serves as a venue for discussing problems of water supply and demand, and joint institutions have been created between Israelis and Palestinians and between Israel and Jordan on water issues. The idea of multilateral working groups is to create opportunity for progress on the bilateral track (7). In other words, agreement on water issues can become a catalyst for bilateral political agreements. The current Working Group on Water Resources has thus been more about fact-finding and sharing information than on the more contentious issues of water rights and allocations (7). It can be said, however, that Israel and Jordan have, to a large degree, agreed on these issues by signing the Israel–Jordan Peace Treaty in 1994. Nonetheless, a true accounting of rights and allocations in the region will need to include the Palestinian Authority as a fellow riparian to the Jordan, as well as Syria and Lebanon, that contain the river's headwaters.

THE NILE RIVER BASIN

The Problem

Management of the Nile waters is a case of historical rights (essentially claimed by Egypt) to the river and the political and economical power of the riparians. Egypt, as the power broker of the region, wields the greatest influence over the Nile, even though all of the Nile's water flow originates outside of Egypt. To date, only Egypt and the Sudan have negotiated on means for managing the Nile. However, nine sovereign states share the waters of the Nile. The core question is thus where best to control the river for the benefit of all riparians—upstream or down (7).

Background

The Nile is one of the great natural wonders of the world: It is the longest river and flows 6,825 km south to north across 35° latitude. The Nile's catchment basin covers approximately one-tenth of the African continent, an area of 3,007,000 square kilometers. Because of its size and variety of climates and topographies, the Nile Basin constitutes one of the most complex of all major river basins. Throughout the northern reaches of the Nile, human civilizations have depended on the river for their very survival; without the river, Egyptian civilization would not exist.

The Nile has two major tributaries, the White Nile and the Blue Nile. These two principal headwaters are very different from each other; they arise in contrasting climatic and physiographical areas and are characterized by disparate hydrologic regimes. The White Nile flows out of the tropical rainbelt of Central Africa and has relatively little interseasonal variation. The seasonal rainfall fluctuations in the Blue Nile headwater region of Ethiopia are much more marked than those prevailing at the source of the White Nile. In sum, approximately 86% of the Nile's flow originates in Ethiopia: The Blue

Nile contributes 73% and the Atbara 13%. The other 14% is contributed by the White Nile (10). The total volume of water available in any given year is subject to substantial hydrologic variability.

The Nile River Basin, along with all the water it carries, traverses the territories of nine separate nations: Burundi, Egypt, Ethiopia, Kenya, Rwanda, the Sudan, Tanzania, Uganda, and Zaire. Although these countries share one of the greatest rivers in the world, many inequities over this common resource exist and are especially difficult to redress. The Nile presents a classic riparian problem: A river does not recognize boundaries between states—it flows continuously. The ways in which it is diverted, collected, used, and returned to the stream determine the quality and quantity of water that each successive downstream nation will receive.

The Nile region is unique in economics and development. The pattern of water demand and use in the Nile Basin contrasts sharply with the pattern of supply: The upper valley states, in particular Ethiopia, are best placed geographically, but the lower valley states, especially Egypt, have a vital resource interest and can exercise decisive military and economic power. The paradox is that the countries that contribute the most water use the least, but the countries that use the most are those that have political and economic influence.

Egypt and the Sudan account for more than 90% of the water drawn from the Nile. However, there are seven other basin states, and their demand for water will inevitably grow. The upper riparian states, Tanzania, Uganda, Rwanda, Kenya, Burundi, and Zaire, use 0.05 km^3/yr collectively; Ethiopia uses only 0.6 km^3/yr. This is ironic because combined, these states contribute 72 km^3 of water per year, or about 86% of the Nile's flow (10).

Attempts at Conflict Management

Optimal development of the basin's water resources depends on agreement and cooperation among individual countries that share several common difficulties. Among them are population growth, insufficient agricultural production, and evaporative losses (10). These all directly impact the available water supply for all basin countries. An integrated approach to resolving these issues is most logical given that all countries have a stake in the outcome. In addition, these dilemmas transcend international boundaries, so any action taken by one nation will almost assuredly impact other nations in the region.

In 1929, the Nile Waters Agreement between Egypt and the Sudan was signed. This agreement guaranteed allocations to Sudan, but the entire seasonal flow of the river was reserved for Egypt (7). The agreement was heavily weighted in Egypt's favor by assurances that no works would be developed on the river that would threaten Egypt's interests. Intermittent negotiations occurred between Egypt and the Sudan on the building of the Aswan High Dam in Egypt in the late 1950s. No agreement on joint management of the dam was reached, in 1959, Sudan unilaterally built the Sennar Dam and by so doing, repudiated the 1929 agreement (7). However, when the military regime gained power in the Sudan in 1958, a Nile Waters Treaty was signed in 1959 called the Agreement for the Full Utilization of the Nile Waters.

The upstream states' reaction to the treaty was that the downstream states had simply divided the recognized average flow of the Nile between themselves and left nothing for any other countries (10).

Outcome

Egypt and the Sudan have agreed that any claims to Nile waters by the other riparians will be met with one unified Egyptian–Sudanese position. No other state riparian to the Nile has ever exercised a legal claim to the waters allocated in the 1959 Treaty (7). Ethiopia, which has an estimated 75–85% of the river's annual flow, has not been a major player in Nile hydropolitics. So far, it has pursued unilateral projects to meet its growing need for irrigation. Recently, however, a new Nile Initiative created dialogue among all the riparians on joint planning for the basin, and a new treaty, inclusive of all the riparians, may be in the offing.

CLOSING

Wars over water have not occurred. The last 50 years have seen only 37 acute disputes that involved violence. During the same period, 157 treaties have been negotiated and signed (7). But the growing lack of clean freshwater coupled with the rise in demand has led to intense political instability, and acute violence has occasionally resulted (7). Many of the world's water conflicts are also in the developing world, where political instability is coupled with socioeconomic ills such as poverty. Within the Middle East, the legacy of colonialism, mistrust between riparians, disparities in economic, infrastructural, and social capacity, and political orientation further complicate an integrated and regional approach to water resources development.

Middle Eastern water development has followed a pattern of unilateral development first due to perceived claims of sovereignty or to avoid political intricacies. At some point, a regional power may implement a project that will impact other riparians, such as the Aswan High Dam on the Nile and the GAP project in Turkey. The absence of relations or institutions for resolving these developments can be a flashpoint for violence. General legal principles for transboundary water management are, however, currently being defined and improved on, such as the Convention on the Non-Navigational Uses of International Watercourse ratified by the UN Assembly in 1997 (7). These principles are beginning to provide a framework to which countries can turn in formulating cooperative water regimes.

As described before, although the issues are highly contentious and complex, all-out war over water has not occurred in the Middle East. Positive outcomes to conflict resolution have occurred in some cases, for example, the Israel–Jordan peace treaty. In other conflicts, the countries are moving slowly toward resolution, as in the case of the Nile riparians.

BIBLIOGRAPHY

1. Gleik, P.H. (1994). Water, war and peace in the middle east. *Environment* **36**: 6–15.
2. Flakenmark, M. (1986). *Fresh Waters as a Factor in Strategic Policy and Action*. In: Westing (Ed.). Global Resources and International Conflict, pp. 85–113.
3. Gleik, P.H. (1993). Water and conflict. fresh water resources and international security. *International Security* **18**(1): 79–112.
4. Hillel, D. (1994). *Rivers of Eden. The Struggle for Water and the Quest for Peace in the Middle East*. Oxford University Press, New York.
5. Lowi, M.R. (1993). Bridging the divide. Transboundary resource disputes and the case of West Bank Water. *Int. Security* **18**(1): 113–138.
6. Lipchin, C. (1997). Water scarcity, international security and resource disputes-the case of the tigris-euphrates and jordan basin river systems. In: *Population-Environment Dynamics: Ten Case Studies*. S.L. Arlinghaus, W.D. Drake, and D. Jacobs (Eds.). University of Michigan, Ann Arbor, MI. Available: http://www-personal.umich.edu/wddrake/lipchin.html.
7. Wolf, A.T. (2001). *Transboundary Waters: Sharing Benefits, Lessons Learned*. Thematic Background Paper, International Conference on Freshwater, Bonn, Germany.
8. Shultz, M. (1995). Turkey, Syria and Iraq: A Hydropolitical Complex. In: *Hydropolitics. Conflict over Water as a Development Constraint*. L. Ohlson (Ed.). University Press, Dhaka.
9. Inbar, M. and Maos, J. (1984). Water Resource Planning and Development in the Northern Jordan Valley. *Water Int.* **9**(1): 19.
10. Smith, J. (1997). Nine Nations, One Nile. In: *Population-Environment Dynamics: Ten Case Studies*. S.L. Arlinghaus, W.D. Drake, and D. Jacobs (Eds.). University of Michigan, Ann Arbor, MI. Available: http://www-personal.umich.edu/wddrake/smith.html.

READING LIST

Starr, J.R. (1991). Water Wars. Financial Policy.

BENJAMIN FRANKLIN'S ARMONICA: A WATER MUSIC INSTRUMENT

William E. Marks
Water Consciousness, Inc.
Martha's Vineyard,
Massachusetts

The art of creating music from water-filled glass receptacles is probably as ancient as glassmaking. Some music scholars date the phenomenon of water music back to 2500 B.C. Because this time predates recorded history, all we have as documentation are references in early folklore and tales.

In the ancient days of Pythagoras, a variety of music was composed by musicians using water-filled bowls that held different levels of water. This use of creating water music according to a mathematical scale, resonated with Pythagorean philosophy, which believed in transmigration of the soul and a related belief in a numerical scale.

In Persia, where some methods of their glassblowing techniques remain a mystery to this day, we find accounts, as early as the fourteenth century, of water music being played from glasses. Largely because glass music was often played by amateur performers who created their own folk music, glass music has a history that is obscure and mysterious.

For instance, in Europe, the earliest reference to musical glasses dates back to 1492. Shortly thereafter, we find the visionary Nostradamus (1503–1566) using a water-filled bowl to see into the affairs of humankind across the span of time. This method of using a water bowl for divining the future is known as "scrying."

As expressed in the book, *The Holy Order of Water*, Nostradamus would sit alone at night on a "tripod stool of brass" and look deeply into a large bowl of water. Using a special branch that was wetted with water on both ends, he would rub one end of the wet branch in a circular motion around the rim of the water-filled bowl. This would produce tones, or ethereal voices, that according to Nostradamus, gave him information and insight about the future.

It is curious to note that there is little if any elaboration on the source of the water used by Nostradamus to see into other dimensions, a fact that may have been a hidden secret—just as it was hidden in writings concerning the creation of the "Philosopher's Stone," which is often referred to in ancient lore. Certainly, with what modern science is revealing at this time, we know that the quality and source of waters used for healing and for other "higher" purposes has to be of unique quality and handled with care.

When Galileo (1564–1642) was writing about pendulums, he wrote the following about water music being played with wine glasses filled with water:

> If one bows the bass string on a viola rather smartly and brings near it a goblet of fine, thin glass having the same tone [tuono] as that of the string, this goblet will vibrate and audibly resound. That the undulations of the medium are widely dispersed about the sounding body is evinced by the fact that a glass of water may be made to emit a tone merely by the friction of the fingertip upon the rim of the glass; for in this water is produced a series of regular waves. The same phenomenon is observed to better advantage by fixing the base of the goblet upon the bottom of a rather large vessel of water filled nearly to the edge of the goblet; for if, as before, we sound the glass by friction of the finger, we shall see ripples spreading with the utmost regularity and with high speed to large distances about the glass. I have often remarked, in thus sounding a rather large glass nearly full of water, that at first the waves are spaced with great uniformity, and when, as sometimes happens, the tone of the glass jumps an octave higher I have noted that at this moment each of the aforesaid waves divides into two; a phenomenon which shows clearly that the ratio involved in the octave [forma dell' ottava] is two (1,2).

For whatever reason, this ancient connection of the soul to water glass music plays itself again and again in various countries throughout history, including the life and times of Benjamin Franklin. Water is the source of life, and life depends on an ongoing healthy source of water, so the connection of creation and the soul to water cannot be taken for granted. As expressed above by Galileo, the water in glasses near the vibrating glass of water will vibrate as well. It takes only a small leap of imagination to grasp the concept that the cellular water within the bodies of people near a musician playing water glasses will also vibrate.

Throughout the seventeenth century, amateur performances are documented that featured sets of wine glasses arranged on a table, played with moistened fingers and tuned by the addition of small amounts of water. This "glass music" became popular throughout Europe and was played in churches and parlor circles (known as 'Gay Wine Music' in 1667 London). The unique qualities of this water glass music captured the imagination of many composers and was well entrenched in Europe during the height of Romanticism in the 1700s.

During 1743, Handel's Water Music was played on an 'Angelic Organ' (a set of tuned wine glasses) by the Irish musician Richard Pockridge. Franklin may have been exposed to this rendition of Handel's Water Music on water glasses because he attended many music concerts while in Europe. In 1759, Franklin wrote the following in his diary after watching Handel's last concert, "... the sublime old Handel was led to the organ and conducted the Messiah for the last time eight days before his death."

Also of musical record, we have the composer Gluck, who in 1746 impressed audiences throughout Europe with his "Verrillon." In doing so, Gluck used a set of wine glasses to produce his water music ('verres' is the French word for 'glass').

In 1757, the Pennsylvania assembly directed Franklin "to go home to England" and negotiate a fairer system of taxation for the colonies. That same year, while living in London he heard E. H. Delaval of the Royal Society play music on water-filled glasses. Afterward, Franklin wrote about the Delaval water music concert he attended:

> He collected a number of glasses of different sizes, fixed them near each other on a table, and tuned them by putting into them water, more or less as each note required. The tones were brought out by passing his fingers round their brims. He was unfortunately burnt here, with his instrument, in a fire which consumed the house he lived in (1759). Mr. E. (Edmund Hussey) Delaval, a most ingenious member of our Royal Society, made one in imitation of it, with a better choice and form of glasses, which was the first I saw or heard (3, p. 297).

Upon hearing the water glasses played, Franklin was moved to refer to it as, "That charming science." Having inherited an interest in music from his father, Franklin already had a strong foundation in the history, theory, and artistry of playing and composing music. Wanting to hear more harmonies with his melody, he set about inventing a way to mimic the sound of water glass music.

> Being charmed by the sweetness of its tones and the music he produced from it, I wished to see the glasses disposed in a more convenient form, and brought together in a narrower compass, so as to admit a greater number of tunes and all within reach of hand to a person sitting before the instrument (3, p. 298).

Enlisting the assistance of a London glassblower named Charles James, Franklin had 37 glass bowls (hemispheres) blown and ground to different sizes and thicknesses to produce 37 distinct tones. Each of these was then painted to help identify the note it produced. In the center of each glass bowl was a hole into which Franklin inserted a piece of cork. He then lined up the bowls chromatically, from lowest tone to highest, on a horizontal spindle (iron rod) according to size. The largest glass bowl was mounted to the left and then each subsequent bowl installed to the right was progressively smaller. The larger the bowl, the lower the tone, and the smaller the bowl, the higher the tone.

The iron spindle was attached to an apparatus similar to a spinning wheel, which was then turned by a foot treadle, a configuration common in those days on foot-operated sewing machines.

As the foot treadle turned the wheel, the spindle rotated the glass bowls. This allowed Franklin to sit at the Armonica, use his feet to turn the bowls, and by wetting his fingers from a bowl of water, produce music that rivaled that of the musical glasses containing water.

To make it easier to identify the notes of the bowls, Franklin painted the bowls using a combination of the seven colors of the rainbow along with the color white. Franklin's "Armonica" included two of the three octaves of the musical scale, giving musicians a greater creative range for their compositions (See Fig. 1).

The vibrations produced from the spinning glasses in the Armonica produced sounds with different frequencies when touched with moistened fingers. Therefore, a variety of musical tones could be created. The handmade glass used in Franklin's time was mostly soda-lime glass (containing sodium) or lead glass (containing lead).

When playing the Armonica, a moistened finger held to the rim of the glass as it rotates, produces a pure and constant tone. The pitch was altered by changing the pressure of the finger touching the rim of the glass, by adding more water to one's fingers, or by decreasing the size of the glass being touched. Therefore, Franklin's use of glasses of varying sizes arranged on the Armonica allowed a musician to create a range of melodies on the instrument. Because Franklin's Armonica instrument included two to three octaves of the musical scale, it permitted a composer greater personal freedom in creating musical arrangements.

When and how Franklin decided to coin the word "Armonica" for his musical invention is not known. What we do know is that on July 13, 1762, Franklin sent a letter to Padre Giambastista Beccaria of Italy, describing the attributes of this new musical invention:

> The advantages of this instrument are that its tones are incomparable sweet beyond those of any other; that they can be swelled and softened at pleasure by stronger or weaker pressures on the fingers, and continued at any length, and that the instrument, being once well tuned, never again wants tuning (3, p. 298).

Figure 1. 1920s painting of Franklin playing the Armonica. "Of all my inventions, the glass armonica has given me the greatest personal satisfaction"—Benjamin Franklin. (Photograph courtesy William Waterway).

In the same letter, Franklin writes, "In honour of your musical language, I have borrowed from it the name of this instrument, calling it the Armonica."

Considering this letter, it appears appropriate that Franklin chose "Armonica" as the name his musical instrument invention because "armonica" is the Italian word for "harmony."

Shortly after its introduction in England and Europe, the Armonica became widely accepted as a new musical instrument. As greater numbers of musicians began playing the Armonica, it soon became a fad. To meet the demand, Armonica factories were built. One of them reportedly employed more than 100 workers. It is estimated that more than 6000 Armonicas were manufactured during a period of about 40 years.

It is of historic note that many of the musicians who played the Armonica were women, which was unusual for that time. The French queen Marie Antoinette took lessons from a famed European Armonica player named Marianne Davies.

Always the charming statesman, Franklin would often take his Armonica to social gatherings, where he would set about entertaining the social elite. In short order, famous composers such as Schmittbauer, Richard Straus, J. A. Schultz, Mozart, Haydn, Galuppi, Martini, Beethoven, Donizetti, J. G. Naumann, Jommelli, J. F. Reichardt, Saint-Saens, and Leopold Rollig began composing music to fit Franklin's Armonica. In 1787, Rollig wrote, "...the sensation the harmonica [Armonica] produced after its first appearance and the unanimous applause of all who heard it, make the instrument...the most satisfying and most beautiful mankind has ever possessed."

Rollig's statement seems to echo what Franklin wrote 24 years earlier on December 11, 1763,

> I play some of the softest Tunes on my armonica, with which Entertainment our people here are quite charmed and conceive the Scottish Tunes to be the finest in the World. And indeed, there is so much simple Beauty in many of them, that in my Opinion they will never die, but in all ages find a Number of Admirers.

To say the least, Franklin was already renowned in England and Europe for his invention of the lightning rod. With the invention of the Armonica, Franklin's overseas popularity made him one of America's first international superstars.

Even though Franklin invented the Armonica, it is historical fact that Germany's society of musicians and fine glassmakers embraced this instrument more than musicians in any other nation. Most of the periods' literature on the Glass Armonica was available only in the German language. The only method book for playing the Armonica was authored by J. C. Müller in 1788 and published in Leipzig, Germany.

ARMONICA AS A HEALING WATER INSTRUMENT

Today, as we have evolved to understand the relationship of music to water for healing, it was something that Franklin seemed to understand in his day as well. We have the following account of Franklin using his Armonica for healing Princess Izabella Czartoryska of Poland from "melancholia."

As expressed in her own writings of 1772,

> I was ill, in a state of melancholia, and writing my testament and farewell letters. Wishing to distract me, my husband explained to me who Franklin was and to what he owed his fame...Franklin had a noble face with an expression of engaging kindness. Surprised by my immobility, he took my hands and gazed at me saying: pauvre jeune femme ["poor young lady"]. He then opened an armonica, sat down and played long. The music made a strong impression on me and tears began flowing from my eyes. Then Franklin sat by my side and looking with compassion said, "Madam, you are cured." Indeed that moment was a reaction to my melancholia. Franklin offered to teach me how to play the armonica—I accepted without hesitation, hence he gave me twelve lessons (4).

Another documented case of the Armonica used for healing involves Franz Mesmer (1734–1815). Mesmer was a Viennese psychiatrist who is known for his theory on "animal magnetism." Today, many people believe that Mesmer invented hypnosis. This misconception is due to the use of his name in the word "mesmerize." However, Mesmer did not invent or use hypnosis. One of Mesmer's water healing methods was to use a wooden tub filled with magnetized water. Besides using tubs with treated water, Mesmer also used the Armonica in his treatments.

The following is an account given by a Dr. Le Roux sometime between 1778 and 1779. Dr. Le Roux had taken an army surgeon to Mesmer's clinic for treatment of gout:

> After several turns around the room, Mr. Mesmer unbuttoned the patient's shirt and, moving back somewhat, placed his finger against the part affected. My friend felt a tickling pain. Mr. Mesmer then moved his finger perpendicularly across his abdomen and chest, and the pain followed the finger exactly. He then asked the patient to extend his index finger and pointed his own finger toward it at a distance of three or four steps, whereupon my friend felt an electric tingling at the tip of his finger, which penetrated the whole finger toward the palm. Mr. Mesmer then seated him near the harmonica [Armonica]; he had hardly begun to play when my friend was affected emotionally, trembled, lost his breath, changed color, and felt pulled toward the floor (5).

Being ahead of his time, Mesmer was using a water treatment method that is now becoming quite popular in Europe. Water, light, and sound are now used to stimulate healing while a patient is submerged in a tub of mineral water.

Mesmer, Mozart, and Franklin were Freemasons and were thus connected in their appreciation and use of the Armonica for helping humankind.

As noted in history, all three were highly regarded when it came to playing the Armonica. In 1773, Leopold Mozart met Marianne Davies in Vienna and started playing the Armonica shortly afterward. In a letter to his wife, Mozart wrote,

> Do you know that Herr von Mesmer plays Miss Davies' armonica unusually well? He is the only person in Vienna who has learned it, and he possesses a much finer instrument than Miss Davies does. Wolfgang too has played upon it. How I should love to have one!

Mesmer believed so much in the power of the Armonica that his dying request was to have a priest play the Armonica for him.

Besides the use of the Armonica for healing in the time of Benjamin Franklin, the harp was also used for healing. The ethereal connection of healing harp and Armonica music can be found in the adage, "If the harp is the instrument of the Angels, the Armonica is the voice of the Angels."

The use of the harp and Armonica for healing is based on the vibrational tones generated by these instruments. In theory, these tones cause the intracellular and intercellular water of the human body to be energized in a way that is conducive to healing. Testimony to this can be found in one of today's musical healing societies known as "Healing Harps (www.healingharps.org).

On another note, just as Franklin's lightning rod invention quickly found its way from America to all of Europe, in reverse fashion, his Armonica quickly found its way from Europe to all of America.

Needless to say, when Franklin returned home to America, he set up a room specifically for music and entertaining. This became known as the Franklins' "Blue Room." Even though this room was complemented with an assortment of musical instruments, the Armonica was always Franklin's favorite. In testimony to Franklin's appreciation for his musical invention, he leaves us the following quote, "Of all my inventions, the glass Armonica has given me the greatest personal satisfaction."

Ben Franklin collected no money from his glass Armonica invention. He refused to patent any of his inventions, saying: "As we enjoy great Advantages from the Inventions of others we should be glad of an Opportunity to serve others by any Invention of ours, and this we should do freely and generously."

THE DEATH AND RESURRECTION OF THE ARMONICA

After about 40 years as a popular musical instrument, the Armonica suddenly began to fall into disfavor during the early 1900s. There are various theories as to why this occurred. Some theorists say people were fearful of contracting lead poisoning from the bowls which were made of lead crystal. Others say that the Armonica was responsible for causing nervous disorders and possibly insanity and that the toning of the Armonica opens doorways to ethereal realms that caused maladies and possibly death. In Germany, where the Armonica was most popular, there was the unfortunate coincidence of a young child dying while an Armonica was being played in a performance hall. Rumors became so rampant that the Armonica became banned in some towns.

As controversy raged throughout Europe over the theorized negative health impacts of the Armonica, Franklin ignored these rumors and continued playing the Armonica his entire life without suffering any illness. However, the Armonica's popularity never really returned to what it was during the height of its popularity.

From a more practical perspective, it was believed that the growing popularity and versatility of the piano may have contributed to the Armonica's downfall. In about 1720, Christofori finished his real pianoforte. Christofori constructed a case superior to those used for harpsichords, which allowed the piano to withstand the increased strain of heavier strings. Christofori also added other improvements to the piano, which allowed the performer to produce a delicate pianissimo as well as a strong fortissimo, not possible on either the clavichord or harpsichord. These improvements in the piano also produced tones of music which could fill a concert hall and were easily heard by the entire audience.

Regardless of what caused the demise of the Armonica, there is now a renewed interest in its use. In the 1980s, American glassmaker, Gerhard Finkenbeiner of Waltham, Massachusetts, researched and resurrected Franklin's Armonica from obscurity. His glassmaking company, G. Finkenbeiner, Inc., makes a quality Quartz Glass Harmonica (Armonica) that contains no lead. According to Finkenbeiner, the use of pure crystal glass also gives a superior sound to the reinvented Armonica.

Finkenbeiner's version of the Armonica is finding interest and popularity as a musical instrument. It has been used at the Metropolitan Opera; in radio commercials, and in various movie tracks.

Today, a handful of musicians play the Armonica. Noted musician, William Wilde Zeitler, has composed several new musical compositions for the Armonica, which are available as CDs.

Another musician, Cecilia Brauer, is world renowned for her Armonica performances. Ms. Brauer gives lecture/demonstration programs called “Ben Franklin and the Armonica” at museums, historical sites, libraries, and social institutions such as the Carnegie Museum; the Franklin Institute, and the National Portrait Gallery in Washington, DC. Besides composing original music for several CDs, Ms. Brauer is also known for playing the Armonica at the Metropolitan Opera House, for the sound track of the movie “Interview with the Vampire,” and for several PBS specials.

On an ethereal note, Gerhard Finkenbeiner, the glassmaker from Massachusetts responsible for resurrecting interest in the Armonica, disappeared under strange circumstances in the late 1990s. The mystery of his sudden disappearance has yet to be solved. However, company employees have rallied to keep the company as an ongoing enterprise. In 2003, the employees of G. Finkenbeiner, Inc., issued the following statement:

> The Quartz Glass Harmonica, Gerhard’s invention and musical triumph, will also be continued faithfully by G. Finkenbeiner Inc., as Gerhard would have wanted it that way. This unique musical apparatus and its wonderfully ethereal tones will continue to delight listeners around the world.
>
> If you have any questions or comments, we would love to hear from you. Your satisfaction is our priority! Your continued support has been an invaluable asset during these somber times, and we thank you sincerely for your patronage.

THE ARMONICA’S FUTURE

To have a grasp of the Armonica’s future, the author of this article contacted two the world’s most prominent players, William Wilde Zeitler and Cecilia Brauer. Their responses speak for themselves:

William Wilde Zeitler (info@glassarmonica.com)

> In spite of the publicity the Armonica has received of late on the History Channel, NPR, etc., I think its future is quite precarious. At present I’m the ONLY full-time composer/performer for the instrument; I know of one other full-time player, and maybe a half-dozen part time players. That definitely puts it on “the endangered musical instrument list.

Cecilia Brauer (brauerarmonica@webtv.net)

> It is very difficult to predict the Armonica’s future. So far as I know, the only manufacturer of the Armonica in the U.S. was Gerhard Finkenbeiner in Waltham, Massachusetts. He “disappeared” in a private airplane flight almost five years ago. And, unless his descendants decide to keep the Armonica segment of his company, G. Finkenbeiner Inc., as a going concern, I do not know what the prognosis is. I would hate to see it disappear again as it is such a beautiful, unique sounding instrument and also such an important part of Franklin’s history. Franklin felt that of all the things he accomplished, the Armonica gave him his greatest personal joy.
>
> There is a man, Sacha Reckert in Germany, who makes them, but they are somewhat different than Gerhard’s. Perhaps the people at Finkenbeiner’s might be able to help. The contact is DHession@finkenbeiner.com.

Based on the opinions of William Wilde Zeitler and Cecilia Brauer, it certainly does sound as though the future of the Armonica is somewhat dubious. However, this is one of the few glass water instruments in the world, and due to its history in the cultures of Europe and America, the Armonica will hopefully continue to provide enjoyment for many future generations.

BIBLIOGRAPHY

1. Galileo, G. (1638). *Two New Sciences*.
2. Galileo, G. (1638). *Dialogues Concerning Two New Sciences*. H. Crew and A. de Salvio (Eds.). Dover Publications, New York, ISBN 0-486-60099-8.
3. Van Doren, C. (1991). *Benjamin Franklin*, New York.
4. Translated in Lipowski, Z.J. (1984). Benjamin Franklin as a psychotherapist: A forerunner of brief psychotherapy. *Perspect. Biol. Med.* **27**: 361–366.
5. Pattie, F.A. (1994). *Mesmer and Animal Magnetism*. Edmonsten Publishing Co., Hamilton, NY, p. 73.

WATER TRANSPORTATION OCCUPATIONS

Bureau of Labor Statistics, U.S.
Department of Labor

SIGNIFICANT POINTS

- Many jobs in water transportation occupations require a merchant mariner’s document or a license from the U.S. Coast Guard.
- Merchant mariners on oceangoing ships are hired for periods ranging from a single voyage to several continuous voyages and may be away from home continuously for months.
- Jobs aboard oceangoing vessels have high pay, but competition for them remains keen, and merchant mariners might have to wait months between work opportunities.

NATURE OF THE WORK

The movement of huge amounts of cargo, as well as passengers, between nations and within our Nation depends on workers in water transportation occupations, also known on commercial ships as merchant mariners. They operate and maintain deep-sea merchant ships,

---

tugboats, towboats, ferries, dredges, excursion vessels, and other waterborne craft on the oceans, the Great Lakes, rivers, canals, and other waterways, as well as in harbors. (Workers who operate watercraft used in commercial fishing are described in the section on **fishers and fishing vessel operators** elsewhere in the *Handbook*.)

Captains, mates, and pilots of water vessels command or supervise the operations of ships and water vessels, both within domestic waterways and on the deep sea. *Captains* or *masters* are in overall command of the operation of a vessel, and they supervise the work of all other officers and crew. They determine the course and speed of the vessel, maneuver to avoid hazards, and continuously monitor the vessel's position with charts and navigational aides. Captains either direct or oversee crew members who steer the vessel, determine its location, operate engines, communicate with other vessels, perform maintenance, handle lines, or operate equipment on the vessel. Captains and their department heads ensure that proper procedures and safety practices are followed, check to make sure that machinery and equipment are in good working order, and oversee the loading and discharging of cargo or passengers. They also maintain logs and other records tracking the ships' movements, efforts at controlling pollution, and cargo and passengers carried.

Deck officers or *mates* direct the routine operation of the vessel for the captain during the shifts when they are on watch. All mates stand watch for specified periods, usually 4 hours on and 8 hours off. However, on smaller vessels, there may be only one mate (called a *pilot* on some inland towing vessels), who alternates watches with the captain. The mate would assume command of the ship if the captain became incapacitated. When more than one mate is necessary aboard a ship, they typically are designated chief mate or first mate, second mate, third mate, etc. Mates also supervise and coordinate activities of the crew aboard the ship. They inspect the cargo holds during loading to ensure that the load is stowed according to specifications and regulations. Mates supervise crew members engaged in maintenance and the primary upkeep of the vessel.

Pilots guide ships in and out of harbors, through straits, and on rivers and other confined waterways where a familiarity with local water depths, winds, tides, currents, and hazards such as reefs and shoals are of prime importance. Pilots on river and canal vessels usually are regular crew members, like mates. Harbor pilots are generally independent contractors who accompany vessels while they enter or leave port. Harbor pilots may pilot many ships in a single day. *Motorboat operators* operate small, motor-driven boats that carry six of fewer passengers on fishing charters. They also take depth soundings in turning basins and serve as liaisons between ships, between ship and shore, between harbors and beaches, or on area patrol.

Ship engineers operate, maintain, and repair propulsion engines, boilers, generators, pumps, and other machinery. Merchant marine vessels usually have four engineering officers: A chief engineer and a first, second, and third assistant engineer. Assistant engineers stand periodic watches, overseeing the safe operation of engines and machinery.

Marine oilers and more experienced *qualified members of the engine department*, or QMEDs, maintain the vessel in proper running order in the engine spaces below decks, under the direction of the ship's engineering officers. These workers lubricate gears, shafts, bearings, and other moving parts of engines and motors; read pressure and temperature gauges; record data; and sometimes assist with repairs and adjust machinery.

Sailors operate the vessel and its deck equipment under the direction of the ship's officers and keep the nonengineering areas in good condition. They stand watch, looking out for other vessels and obstructions in the ship's path, as well as for navigational aids such as buoys and lighthouses. They also steer the ship, measure water depth in shallow water, and maintain and operate deck equipment such as lifeboats, anchors, and cargo-handling gear. On vessels handling liquid cargo, mariners designated as *pumpmen* hook up hoses, operate pumps, and clean tanks; on tugboats or tow vessels, they tie barges together into tow units, inspect them periodically, and disconnect them when the destination is reached. When docking or departing, they handle lines. They also perform routine maintenance chores, such as repairing lines, chipping rust, and painting and cleaning decks or other areas. Experienced sailors are designated *able seamen* on oceangoing vessels, but may be called simply *deckhands* on inland waters; larger vessels usually have a *boatswain*, or head seaman.

A typical deep-sea merchant ship has a captain, three deck officers or mates, a chief engineer and three assistant engineers, a radio operator, plus six or more unlicensed seamen, such as able seamen, oilers, QMEDs, and cooks or food handlers. The size and service of the ship determine the number of crewmembers for a particular voyage. Small vessels operating in harbors, on rivers, or along the coast may have a crew comprising only a captain and one deckhand. The cooking responsibilities usually fall under the deckhands' duties.

On larger coastal ships, the crew may include a captain, a mate or pilot, an engineer, and seven or eight seamen. Some ships may have special unlicensed positions for entry level apprentice trainees. Unlicensed positions on a large ship may include a full-time cook, an electrician, and machinery mechanics. On cruise ships, *bedroom stewards* keep passengers' quarters clean and comfortable.

WORKING CONDITIONS

Merchant mariners spend extended periods at sea. Most deep-sea mariners are hired for one or more voyages that last for several months; there is no job security after that. The length of time between voyages varies depending on job availability and personal preference.

The rate of unionization for these workers is about 24 percent, much higher than the average for all occupations. Consequently, merchant marine officers and seamen, both veterans and beginners, are hired for voyages through union hiring halls or directly by shipping companies. Hiring halls rank the candidates by the length of time the person has been out of work and fill open slots accordingly. Hiring halls typically are found in major seaports.

At sea, these workers usually stand watch for 4 hours and are off for 8 hours, 7 days a week. Those employed on Great Lakes ships work 60 days and have 30 days off, but do not work in the winter when the lakes are frozen. Workers on rivers, on canals, and in harbors are more likely to have year-round work. Some work 8- or 12-hour shifts and go home every day. Others work steadily for a week or a month and then have an extended period off. When working, they usually are on duty for 6 or 12 hours and off for 6 or 12 hours. Those on smaller vessels are normally assigned to one vessel and have steady employment.

People in water transportation occupations work in all weather conditions. Although merchant mariners try to avoid severe storms while at sea, working in damp and cold conditions often is inevitable. While it is uncommon nowadays for vessels to suffer disasters such as fire, explosion, or a sinking, workers face the possibility that they may have to abandon their craft on short notice if it collides with other vessels or runs aground. They also risk injury or death from falling overboard and hazards associated with working with machinery, heavy loads, and dangerous cargo. However, modern safety management procedures, advanced emergency communications, and effective international rescue systems place modern mariners in a much safer position.

Most newer vessels are air conditioned, soundproofed from noisy machinery, and equipped with comfortable living quarters. For some mariners, these amenities have helped ease the sometimes difficult circumstances of long periods away from home. Also, modern communications, especially email, link modern mariners to their families. Nevertheless, some mariners dislike the long periods away from home and the confinement aboard ship and consequently leave the occupation.

EMPLOYMENT

Water transportation workers held about 68,000 jobs in 2002. The total number that worked at some point in the year was perhaps twice as large because many merchant marine officers and seamen worked only part of the year. The following tabulation shows employment in the occupations that make up this group:

Sailors and marine oilers	27,000
Captains, mates, and pilots of water vessels	25,000
Ship engineers	8,200
Motorboat operators	4,100
All other water transportation workers	3,600

About 30 percent of all workers were employed in water transportation services. About one half worked in inland water transportation—primarily the Mississippi River system—while the other half were employed in water transportation on the deep seas, along the coasts, and on the Great Lakes. About another 28 percent worked in establishments related to port and harbor operations, marine cargo handling, or navigational services to shipping. The Federal government employed approximately 5 percent of all water transportation workers, most of whom worked on supply ships and are Civilian Mariners of the Department Navy's Military Sealift Command.

TRAINING, OTHER QUALIFICATIONS, AND ADVANCEMENT

Entry, training, and educational requirements for most water transportation occupations are established and regulated by the U.S. Coast Guard, an agency of the U.S. Department of Homeland Security. All officers and operators of commercially operated vessels must be licensed by the Coast Guard, which offers various kinds of licenses, depending on the position and type of vessel.

There are two ways to qualify for a deck or engineering officer's license: applicants either must accumulate sea time and meet regulatory requirements or must graduate from the U.S. Merchant Marine Academy or one of the six State maritime academies. In both cases, applicants must pass a written examination. Federal regulations also require that an applicant pass a physical examination, a drug screening, and a National Driver Register Check before being considered. Persons without formal training can be licensed if they pass the written exam and possess sea service appropriate to the license for which they are applying. However, it is difficult to pass the examination without substantial formal schooling or independent study. Also, because seamen may work 6 or fewer months a year, it can take 5 to 8 years to accumulate the necessary experience. The academies offer a 4-year academic program leading to a bachelor-of-science degree, a license (issued only by the Coast Guard) as a third mate (deck officer) or third assistant engineer (engineering officer), and, if the person is qualified, a commission as ensign in the U.S. Naval Reserve, Merchant Marine Reserve, or Coast Guard Reserve. With experience and additional training, third officers may qualify for higher rank.

Sailors and unlicensed engineers working on U.S. flagged deep-sea and Great Lakes vessels must hold a Coast Guard-issued document. In addition, they must hold certification when working aboard liquid-carrying vessels. Able seamen also must hold government-issued certification. For employment in the merchant marine as an unlicensed seaman, a merchant mariner's document issued by the Coast Guard is needed. Most of the jobs must be filled by U.S. citizens; however, a small percentage of applicants for merchant mariner documents do not need to be U.S. citizens, but must at least be aliens legally admitted into the United States and holding a green card. A medical certificate of excellent health attesting to vision, color perception, and general physical condition is required for higher level deckhands and unlicensed engineers. While no experience or formal schooling is required, training at a union-operated school is the best source. Beginners are classified as ordinary seamen and may be assigned to any of the three unlicensed departments: Deck, engine, or steward. With experience at sea and perhaps union-sponsored training, an ordinary seaman can pass the able-seaman exam and move up with 3 years of service.

No special training or experience is needed to become a seaman or deckhand on vessels operating in harbors or on rivers or other waterways. Newly hired workers generally are given a short introductory course and then learn skills on the job. After sufficient experience, they are eligible to take a Coast Guard exam to qualify as a mate, pilot, or captain. Substantial knowledge gained through experience, courses taught at approved schools, and independent study is needed to pass the exam.

Harbor pilot training usually consists of an extended apprenticeship with a towing company or a pilots' association. Entrants may be able seamen or licensed officers.

JOB OUTLOOK

Keen competition is expected to continue for jobs in water transportation occupations. Overall, employment in these occupations is projected to **grow more slowly than the average** for all occupations through the year 2012. Opportunities will vary by sector, and some of the best opportunities will be in scenic transportation and sightseeing and deep sea, coastal, and Great Lakes transportation.

Employment in deep-sea shipping for American mariners is expected to stabilize after several years of decline. New international regulations have raised shipping standards with respect to safety, training, and working conditions. Consequently, competition from ships that sail under foreign *flags of convenience* should lessen as insurance rates rise for ships that do not meet the new standards. Insuring ships under industrialized countries' flags, including that of the United States, should become less expensive, increasing the amount of international cargo carried by U.S. ships. A fleet of deep-sea U.S.-flagged ships is considered to be vital to the Nation's defense, so some receive Federal support through a maritime security subsidy and other provisions in laws that limit certain Federal cargoes to ships that fly the U.S. flag. Possible future developments include "fast ships"—oceangoing cargo vessels that use jet propulsion—which would decrease ocean-crossing times significantly. If such plans are successful, the industry will benefit in terms of increased business and employment.

Vessels on rivers and canals and on the Great Lakes carry mostly bulk products, such as coal, iron ore, petroleum, sand and gravel, grain, and chemicals. Although shipments of these products are expected to grow through the year 2012, current imports of steel are dampening employment on the Lakes, but actually leading to greater chances for overall employment for transport up the Mississippi River system. Employment in water transportation services is likely to rise, and efforts are underway at the Federal level that could lead to significantly greater use of ferries to handle commuter traffic around major metropolitan areas.

Employment growth also is expected in passenger cruise ships within U.S. waters. Vessels that operate between U.S. ports are required by law to be U.S.-flagged vessels. The building and staffing of several new cruise ships that will travel around the Hawaiian Islands will create new opportunities for employment at sea in the cruise line industry, which is composed mostly of foreign-flagged ships.

Openings within the traditional water transportation sector for mariners, although expanding only slightly, should be quite numerous because of the sizable need to replace those leaving the occupation Some experienced merchant mariners may continue to go without work for varying periods. However, this situation appears to be changing, with demand for licensed and unlicensed personnel rising. Maritime academy graduates who have not found licensed shipboard jobs in the U.S. merchant marine find jobs in related industries. Because they are commissioned as ensigns in the Naval or Coast Guard Reserve, some are selected for active duty in those branches of the Service. Some find jobs as seamen on U.S.-flagged or foreign-flagged vessels, tugboats, and other watercraft or enter civilian jobs with the U.S. Navy or Coast Guard. Some take land-based jobs with shipping companies, marine insurance companies, manufacturers of boilers or related machinery, or other related jobs.

EARNINGS

Earnings vary widely with the particular water transportation position and the worker's experience, ranging from the minimum wage for some beginning seamen or mate positions to more than $37.37 an hour for some experienced ship engineers. Median hourly earnings of water transportation occupations in 2002 were as follows:

Ship engineers	$24.61
Captains, mates, and pilots of water vessels	23.97
All other water transportation workers	14.67
Sailors and marine oilers	13.64
Motorboat operators	12.71

Annual pay for captains of larger vessels, such as container ships, oil tankers, or passenger ships, may exceed $100,000, but only after many years of experience. Similarly, captains of tugboats often earn more than the median reported here, with earnings dependent on the port and the nature of the cargo.

RELATED OCCUPATIONS

Workers in other occupations who make their living on the seas and coastal waters include **fishers and fishing vessel operators** and some **members of branches of the Armed Forces**.

SOURCES OF ADDITIONAL INFORMATION

Information on a program called "Careers Afloat", which includes a substantial listing of training and employment descriptive information and contacts in the U.S., may be obtained through:

- Maritime Administration, U.S. Department of Transportation, 400 7th St. SW., Room 7302, Washington, DC 20590. Internet: http://www.marad.dot.gov

Information on merchant marine careers, training, and licensing requirements is available from any of the following organizations:

- Military Sealift Command, APMC, PO Box 120, Camp Pendleton, Virginia Beach, VA 23548-0120.
- Seafarers' International Union, 5201 Auth Way, Camp Springs, MD 20746.
- Paul Hall Center for Maritime Training and Education, P.O. Box 75, Piney Point, MD 20674-0075. Internet: http://www.seafarers.org/phc
- International Organization of Masters, Mates, and Pilots, 700 Maritime Boulevard, Linthicum Heights, MD 21090-1941.
- U.S. Coast Guard National Maritime Center, 4200 Wilson Boulevard, Suite 630, Arlington, VA 22203-1804. Internet: http://www.uscg.mil/stcw/index.htm

OOH ONET CODES

53-5011.01, 53-5011.02, 53-5021.01, 53-5021.02, 53-5021.03, 53-5022.00, 53-5031.00

Suggested citation: Bureau of Labor Statistics, U.S. Department of Labor, *Occupational Outlook Handbook, 2004-05 Edition*, Water Transportation Occupations, on the Internet at http://www.bls.gov/oco/ocos247.htm (visited August 14, 2004).

A CONCISE GLIMPSE OF WATER IN THE HISTORY OF PHOTOGRAPHY

Basia Irland
University of New Mexico
Albuquerque, New Mexico

> "#125. When you photograph water you strip it of its form: of its restless, liquid reality."
> "#126. When you photograph water you give it an image that in essential ways bears no likeness to water."
> "#127. When you photograph water you speak in graphic terms about something that has no graphic aspect."
> Rebecca Horn, *Dictionary of Water* (1)

Water is universal and indispensable to life. Therefore, it has always been a central metaphor, both sprightly and exacting, for artists. In this brief article, we examine how a few photographers have looked to water for inspiration and also used it as the very substance of their work. There have, of course, been hundreds of photographers who have tested water as their subject matter or used it in their work in various ways at some stage of their career; some of the more notable and interesting are presented here. A few examples have been chosen, beginning with the earliest photographers of the nineteenth century.

Not just images of water caught on photographic paper, but the intricate process of developing and fixing these images depended on water. We tend to take the process for granted today as we have moved into an era of digital images, but it took many fumbling experiments before it was possible to capture the world "out there" and put it onto an object we could hold in our hands and gaze upon for generations (or at least as long as the chemical process would allow). The English scientist, Sir John Herschel (born 1792 and buried in 1871 next to Sir Isaac Newton at Westminster Abbey), explained in his letters to William Henry Fox Talbot (1800–1877), one of the earliest photographic artists, about the important role that clean water played in rinsing images. Although Herschel is best remembered for his discovery of hyposulfate as a fixative, he also discovered that one could fix photographs with plain water. After several unsuccessful attempts, it was found that the water must be as pure as possible and the rinsing process has to be repeated three times. Melting snow worked, as did distilled water. Salt traces in well water and the many forms of pollution finding their way into England's water supply would ruin a print. In March of 1839, Herschel communicated to his scientific colleagues at the Royal Society about the chemical compounds used in his new discoveries to aid ways of developing a print and how the photograph needed to be immersed two or three times in clean spring water and dried between blotting papers.

Most of Talbot's photographs of water were incidental fluid flowing under the more important human-made constructions of bridges (*Lattice Iron Bridge*, 1841; *The Hungerford Suspension Bridge*, 1845; and *The Suspension Bridge at Rouen*, 1843), or doubling the impact of architecture (*Lacock Abbey, Reflected in the Avon*, 1840–41). But in 1844, he took a series, *Sun Pictures of Scotland*, that leave the city and focus on the water itself with such titles as, *A Mountain Rivulet which flows at the foot of Doune Castle, Scenery of Loch Katrine*, and *A Waterfall*.

Colonel Stuart Wortley (1832–1890) presents an almost operatic view of the sea, complete with reflected moonlight, dark suspended cumulus clouds, and titles borrowed from poems such as "How calm, how beautiful comes on / The stilly hour when storms are gone" (1869) and "All the air was white with moonlight / All the water black with shadow" (1869). Wortley, along with Gustave Le Gray, was among the first photographers to capture a sense of stopped movement in an enlarged image. *A Wave Rolling In* (1861) shows a small, single breaking wave caught on film along with two tiny black figures in a rowboat off in the distance.

Gustave Le Gray (born in France, 1820), spent a 5-year period from 1855–1860 primarily involved in photographing sailing ships and ports, but mostly ocean waves. With a taste for the dramatic similar to Wortley's, Le Gray's seascapes (albumen prints from wet collodion on glass negatives) show dark, brooding bodies of water barely lit from the sun's dim rays hidden behind black ominous clouds. Most of these ocean views were taken on the Normandy coast and often included breaking surf in the lower right-hand corner. Le Gray was experimenting with ways to capture the movement of water and the atmospheric effects of light and cloud by using two overlaid negatives. Often, he would use the same seascape negative in combination with a different sky negative. *Grande vague—Sete* (1856/59) predates Wortley's wave photograph by 2 years. It shows a somber jetty jutting halfway across the print with a not so large wave breaking against shadowed rocks in the foreground. On October 27, 1999 this particular print broke the record for the sale of a single photograph at Sotheby's London auction, selling for $760,150. Le Gray spent the last 20 years of his life primarily photographing antiquities in Cairo, Egypt and died in 1882.

Anna Atkins (1799–1871), one of the earliest photographers, incorporated Herschel's cyanotype method, which was inexpensive and easily worked, to produce camera-less photograms using flowing pieces of sea algae. The first part of her series of these photograms was produced in 1843. *Photographs of British Algae; Cyanotype Impressions*, a handmade portfolio of images, was the first use of photography for scientific illustrations. Photograms, which use neither camera nor negative, are produced when objects, placed on a photosensitive surface and exposed to light briefly to create a shadow, are then developed. "Dispensing with lens, aperture, depth of field, and perspective, the photogram assumes the guise of a real trace, a visual footprint on the scale of the original, a memory of the essential contour devoid of redundant information" (2).

The ocean-blue tint of the cyanotype lent itself to the aquatic pictures of algae. Used until recently in the form of the architectural blueprint, cyanotypes involve only two chemical compounds, ferric ammonium citrate and potassium ferricyanide, which are applied to paper using a sponge or brush. After the paper is dried in the dark, an object such as Atkin's algae is placed on the paper and exposed to the sun. Rinsing the exposed cyanotype in water makes the print permanent and brings out the blue color. Halstead Place, where Atkins worked, was known for the quality of its well water.

Atkins was already an accomplished scientific illustrator, especially of seashells. In the handwritten cyanotype introduction to Part I of her book, Atkins explains that the need to make accurate drawings of such intricate algae led her to use the photographic process developed by Sir John Herschel to explore her studies of botany and make them available to a wider audience through publication. Atkins made numerous trips to the ocean to collect specimens of algae, or her friends would provide her with examples. Seaweed would be washed in water and pressed between blotting paper to dry within a few days. The silhouettes of algae hover on her photographic paper as if still in the sea. *Bangia fusco-purpurea* is as delicate as winter breath. *Polysiphonia violacea* is a soft, dendritic wisp. *Himanthalia lorea* begins with two bulbs and spreads out across the paper as if painted with a brush, exiting the confines of the page. In addition to her 10-year period of work with algae (Fig. 1), Atkins made prints of feathers (peacock, parrot, emu, partridge), grasses, lace, flowering plants, and an extensive study of ferns.

Two contemporary British photographers who have also created photographic images without using a camera are Susan Durges (born 1955) and Adam Fuss (1961). A baby of only a few weeks old floats in a liquid universe outside the mother's womb. A snake slithers silently through turquoise fluid leaving a trail of chain-link ripples. Perfectly formed concentric circles radiate outward from a still center. Adam Fuss fills a shallow glass tray with several inches of water over photosensitive paper and then places the baby or snake directly into this warm bath. As Eugenia Parry writes in her essay for the book, *Adam Fuss*, "He wasn't guided by Talbot or Atkins. He'd barely heard of them. But he shared their desires: To let what is mute speak; to capture and keep an object's mysterious self-inscription" (3).

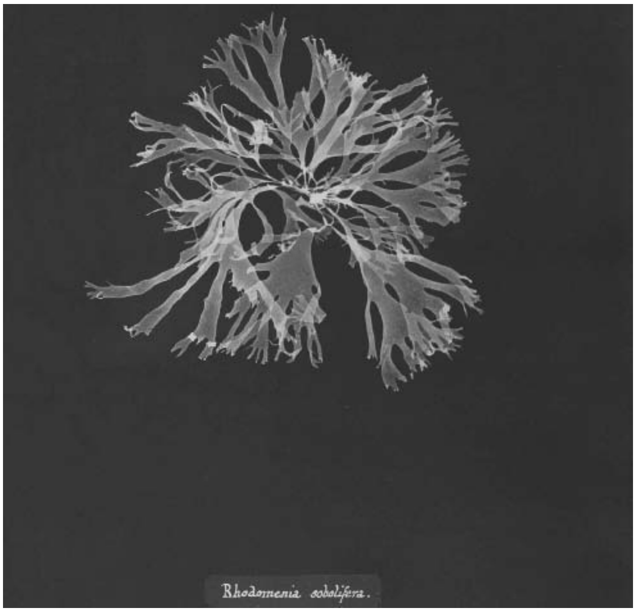

Figure 1. Artist: Anna Atkins. Title: Rhodomenia Sobolifera. Cyanotype negative, bound into a copy of Volume 1 of British Algae, 1850. 26.4 × 20.4 cm. Courtesy: Hans P. Kraus, Jr., Inc., New York.

Whereas most photographers have created photograms within the confines of the studio, Susan Derges (resides in Devon, England) in her *River Taw* series (1997–99), submerges photographic paper directly into a river at night. Enormous sheets of photosensitive paper (66″ × 24″) resting in an aluminum slide are placed just beneath the water surface and exposed to a bright, diffuse flash that captures images of water patterns sometimes overlaid with tree shadows. These are then processed in the darkroom. "Ambient light in the sky adds a colour cast to the Cibachrome (colour positive) images, which ranges from deep blue at full moon to dark green at new moon," writes Martin Kemp in an essay about the artist, in *Susan Derges Liquid Form 1985–99* (2).

Her *Waterfall* series (1997–98, 7′ × 3 1/2′) was created by holding the light-sensitive paper at the top of a water-mill chute using the same nighttime exposures. To create the *Shoreline* series, 1997–99, Derges put the photographic paper on the shoreline at night, waiting for the exact moment a wave would break and exposing it to a microsecond of flash. The ambient light from the moon or a nearby town affected the color of the print, often turning it magenta or dusky pink. Small bits of sand deposited by the wave created vortices and spiral forms.

Another British photographer who eschews preconceived notions of making a photograph is Steven Pippin (born 1960, Surrey, England). In 1993, Pippin began his performance/project, *The Continued Saga of an Amateur Photographer*, which took place in the bathroom on a British train. Ingrid Schaffner, in her essay, "Dirty Hole: Steven Pippin's Obscure Routes," describes how he "...goes about his business converting the toilet into a camera, taking a picture, and developing it. All the while the train rattles and hurls over the tracks" (4). The semicircular prints fit neatly into the commode bowl.

A toilet is not the only object Pippin converted into a camera. He also used washing machines in 1997 for *Laundromat-Locomotion.* In this project, he trip-wired the washers to take photographs and then developed them within the same machine. The wash and spin cycles ripped, scratched, and cracked the emulsion, creating prints that appear Victorian in age.

Across the ocean in the United States and back in the 1930s, we find Dr. Harold Edgerton, Center for Advanced Visual Studies, Massachusetts Institute of Technology, explaining in his book entitled *flash! Seeing the Unseen by Ultra High-Speed Photography* (1939), the process for taking photos that are now part of our iconic language of modernist images (5). These include a bullet passing through an apple, a child popping a balloon, and a drop of milk forming a crown-shape. In one photograph, using the stroboscope, a device that uses a flash to illuminate a moving object intermittently, what appears to be ice or translucent plastic is actually tap water coming out of a faucet caught in Edgerton's camera at 1/50,000 of a second. He also stopped the fluid action of water used to cut pulp and showed the figure eight emitting from a lawn sprinkler. He recorded spiral patterns revolving around a ship's propeller. These studies of motion are reminiscent of Leonardo da Vinci's investigative drawings in pen, ink, and red chalk done in the early 1500s. In *Studies of the Formation of Water in Motion* and *Studies of Water Formation* we see the same vortical shapes that Edgerton caught with his stop-action method 500 years later. The text written on da Vinci's 1510 *Studies of an Old Man and of Swirling Water* reads: "...thus the water forms eddying whirlpools, one part of which is due to the impetus of the principal current and the other to the incidental motion and return flow."

Theodor Schwenk was obsessed with these same vortices formed by fluids and in *Sensitive Chaos*, 1965, he compared liquid spiral forms with other natural phenomenon in his photographs of the bark of a cypress tree, descending hail clouds, unfurling fern leaves, and a drawing of horns of the African kudu antelope.

A few examples of Schwenk's descriptions in his index to the photographs of various vortices are "25. Trains of vortices also arise if a solid object is drawn in a straight line through stationary liquid. 26. The vortices push into the surrounding liquid like the ball part of a joint into its socket. A suitable length of exposure reveals a delicate structure. 33. A suitable length of exposure has revealed something of the structure of vortices. A meandering stream winds its way between the separate vortices. 41. A photograph of a vortex taken under water reveals the spiraling surface between the water and the air which is being sucked in" (6).

In *River Sutra*, Gita Metha describes how every river should have its own minstrel—someone to sing the praises of that body of flowing water (7). Perhaps every river, stream, brook and creek should also have its own photographer. Even though Roni Horn (8) describes water as too elusive an element to hold within the confines of a photograph's rectilinear frame, that certainly has not stopped artists throughout history from trying to capture something of the nature of a particular river. The 1875-mile length of the Rio Grande/Rio Bravo flows from its source in the San Juan Mountains of Southern Colorado, through New Mexico, and provides the border between Texas and Mexico. Photographer Laura Gilpin's (9) documentation in the 1940s of this muddy liquid ribbon (published in 1949 as *The Rio Grande, River of Destiny*) began in 1945 at the end of a war that left a scarcity of photographic materials and difficulties in traveling by car. She also experienced 2 years of drought in her three expeditions along the river. The book's simple chapter headings, "The Source," "Midstream," and "The Border," lead the viewer along a journey witnessing the flow of the river and also the surrounding towns and people living along the banks.

Photographer Jim Bones, in *Rio Grande, Mountains to the Sea*, however, focuses only on the plants and moving water within the riparian zone. "I fell in love with the Rio Grande because of a dream that repeatedly told me if I wanted to know the secrets of my own being, I must go first to the mountains and then to the sea. For the past twenty years, as I walked, rode, and floated throughout the Rio Grande's home country, I discovered that rivers cut to the very heart of things, of rock and life and even dreams" (10). Bones understands that the "end" of a river, as it empties into the sea, is also a beginning because it is here that the hydrologic cycle of evaporation, cloud formation, and rain continue the process of creating rivers all over again.

One of the chapters in Hans Silvester's book, H_2O *The Beauty and Mystery of Water* (11) is dedicated to "The Water Cycle, Disrupted" and mirrors back to the viewer the devastation humans can cause to bodies of water. It is a wake-up call to change our ways. A few of the disturbing images include five children playing in untreated sewage from Mexico City in the Tula River; a dead egret with a fly on its head floating in the polluted Coto Donana, Spain; and an irrigation canal, Bouches-du-Rhone, France, near Marseille's waste discharge where the trash, particularly plastic bags, is piled so high and wide that it is difficult to see the water.

The photographs of the River Thames for Roni Horn's book, *Another Water*, were taken in the Central London area in January and May of 1999 (8). Filling the page, with no reference points or shoreline, the water images could have been taken anywhere. Running continuously along the bottom of each page are 832 comments about the Thames or water in general. "#193. When you go down to the river you're killing two birds with one stone: you stand there and you go places." "#142. Water is sexy. I want to swim in it—to move, song-like, through it. I want to swim in it—push, kick, thrash, slide myself through it. I want to swim in it. I want to be held by it."

In among the references to water are brief descriptions of those who have chosen to take their life by drowning. "#465. An old man was found in the river last Christmas Eve. He was wearing so many layers of clothing (including two wool coats and a jacket) police couldn't lift him out of the water. (They towed it to the pier.)." "#586. A young Parisian woman came to London recently to drown herself in the river. It's curious how the Thames attracts people from far away. I've never heard of any other river doing this. I mean people don't travel from Canada to kill themselves in the Hudson—or even from Ohio." "#446. A young man was found in the river yesterday with a pair of Walkman headphones wrapped around his neck." And every so often there will be a whole white page with an autopsy or police report of someone who was found in the Thames, including details about their lives, whether they wore dentures and exactly where and how their body was located.

Another Water was followed the next year in 2001 by a second book, *Dictionary of Water,* using similar photographs of the Thames taken during the same months. In fact, Horn did not even take the photographs. Dusseldorf artist Nic Tenwiggenhorn made these images assisted by Uwe Schmidt and Helena Blaker. There is no mention within either of Horn's books about the Thames of the photographs being changed by computer, but on a web site providing information about *Dictionary of Water*, it states that this is a collection of digitally manipulated photographic images.

Industrial photographer Wim Cox (born Netherlands, 1938, moved to Germany, 1961) created likenesses of water to those similar by Horn, but they are accompanied by a very different tone of writing done by his brother, filmmaker Paul Cox (born 1940, Netherlands, resides, Melbourne, Australia) (12). The brief text is a poem—a kind of love poem—to water. It begins, "I am. I am water." And so, throughout history, water has pulled photographers to its side, embraced them, and helped in creating memorable images.

BIBLIOGRAPHY

1. Horn, R. (2001). *Dictionary of Water*, Edition 7L, Steidl Publishers, Göttingen, Germany.
2. Kemp, M. (1999). *Susan Derges Liquid Form 1985–99*. Michael Hue Williams Fine Art, London.
3. Parry, E. (1997). *Adam Fuss*. Arena Editions, Santa Fe, NM.
4. Hirsch, F. "International Auction Review," and Schaffner, I. "Dirty Hole: Steven Pippin's Obscure Routes," *Art on Paper*. **4**: 3, (New York, January–February 2000), Fanning New York, 96, pp. 34–39
5. Edgerton, H.E. and Killian, J.R., Jr. (1939). *Flash! Seeing the Unseen by Ultra High-Speed Photography*. Hale, Cushman and Flint, Boston, MA.
6. Schwenk, T. (1965). *Sensitive Chaos*. Rudolf Steiner Press, Stuttgart, Germany.
7. Mehta, G. (1993). *A River Sutra.* Doubleday, Bantam, New York.
8. Horn, R. (2000). *Another Water*. Scalo, Zurich.
9. Gilpin, L. (1949). *The Rio Grande, River of Destiny*. Duell, Sloan and Pearce, New York.
10. Bones, J. (1985). *Rio Grande, Mountains to the Sea*. Texas Monthly Press, Austin, TX.
11. Silvester, H. (2001). *H_2O—The Beauty and Mystery of Water*. Harry N. Abrams, New York.
12. Cox, W. and Cox, P. (1997). *I Am.* Schuffelen, Cologne.

READING LIST

Buckland, G. (1980). *Fox Talbot and the Invention of Photography*. David R. Godine, Boston, MA.

Janis, E.P. (1987). *The Photography of Gustave Le Gray*. The Art Institute of Chicago and the University of Chicago Press, Chicago, IL.

Popham, A.E. (1945). (Deputy Keeper of Prints and Drawings, British Museum). *The Drawings of Leonardo da Vinci*. Reynal and Hitchcock, New York.

Schaaf, L.J. (1985). *Sun Gardens*. Victorian Photograms by Anna Atkins. An Aperture Book, Boston, MA.

Schaaf, L.J. (1992). *Out of the Shadows; Herschel, Talbot and the Invention of Photography*. Yale University Press, New Haven, CT.

WATER AS A HUMAN RIGHT

Laurel Phoenix
Green Bay, Wisconsin

"The human right to drinking water is fundamental to life and health. Sufficient and safe drinking water is a precondition for the realization of human rights... the human right to water entitles everyone to sufficient; affordable; physically accessible; safe and acceptable water for personal and domestic uses" (1).

The human right to physical and economic access to water has recently been declared by nongovernmental institutions and civil society groups worldwide. The movement to establish this right has been driven by two related issues. First, global populations are currently experiencing a water crisis predicted to become much worse. Second, the growing global trend of water privatization has created additional concerns regarding access to potable water.

GLOBAL WATER CRISIS

The United Nations, Earth Policy Institute, Pacific Institute, and other organizations have been following concurrent global trends of rising population and decreasing access to potable water. Currently, there are over 1 billion people lacking access to safe water and by 2010 this number will grow to 2.5 billion (2). Associated with this problem is lack of sanitation. There are currently 2.4 billion people lacking sanitation facilities. As this sanitation problem is exacerbated by growing populations, even more water becomes too polluted to safely drink.

PRIVATIZATION

Privatization of water supplies in developing countries has been met with ambivalence. Rather than increasing the number of people served with potable water, the higher water bills attendant with privatization have made poorer citizens lose access. In several instances, the very poor cannot afford to pay the higher rates without forfeiting some other necessity, like heating or food (3). Privatized facilities in many cases have shut off the water to those who cannot pay, forcing them to use polluted supplies, resulting in sickness or death. For example, Nkobongo, South Africa privatized their water system and rates were increased. As water was shut off to those who could not pay, people turned to polluted lakes and streams, creating a spectacular cholera epidemic killing almost 300 people and sickening over 250,000 over the course of 300 miles and several years (4). This scenario was repeated in many townships, causing the government to spend enormous amounts of money to deal with the health crisis. Moreover, when utilities are privatized in developing countries, there is little oversight regarding how often or to what degree rates are raised. Consequently, water shut-offs or the use of prepaid water meter cards can externalize the problems of poverty in one region to a wider area, while private water companies are insulated from the public health crisis that ensues.

Opponents of privatization demand that the minimum amount of water needed for basic needs ought to be provided for free, and then fees for water used per household can be charged for water used in addition to this subsistence requirement. This minimum amount varies from 20 liters per person per day for drinking and cooking to 40 liters per person per day to include water for hygiene. Opponents also contend that once a profit motive is introduced into provision of water, companies may demonstrate more concern for their profit margin than for the quality of the water, either through delaying needed maintenance, reducing quality control, or hiking prices even further and cutting off more citizens from a potable supply.

DECLARATIONS OF WATER AS A HUMAN RIGHT

In response to the global water crisis and privatization concerns, numerous groups have issued statements delineating the right to water. The National Forum on Water Privatization meeting in Ghana declared in May of 2001 that “water is a fundamental human right, essential to human life to which every person, rich or poor, man or woman, child or adult is entitled.” The Blue Planet Project created a Treaty Initiative to Share and Protect the Global Water Commons, where they stated: “The global fresh water supply is a shared legacy, a public trust and a fundamental human right and, therefore, a collective responsibility” (5). By “collective” they mean every nation. The public trust theme is repeated in the 3rd World Water Forum: “Water, as a public trust and an inalienable human right, must be controlled by the peoples and communities that rely on it for their lives and livelihoods” (6).

IS THIS A NEW HUMAN RIGHT?

Human rights are not established in a vacuum. Citizens agitate for rights only when clear injustices or inequities occur. New rights are incorporated into law when those currently holding those rights agree to share those rights with the disenfranchised. Now that so many people do not have access to clean water globally, the call for human rights to water focuses world governments’ attention on the distribution and access to water worldwide.

Human rights to critical natural resources spring from ancient laws regarding use of the commons. The Roman Emperor Justinian declared in 533–534 B.C.E. that running water, seas, air, and wildlife were things that could not be owned but could be used under regulation to ensure vital common resources were not exploited (7). Perhaps because of its ubiquity, governments never thought to declare a right to something so obvious. Even so, a right to water naturally comes before other rights for equality, freedom, and so on, because there would be no need to establish these other rights if you could not live to enjoy them. Scanlon et al. (8) surveyed documents written by international organizations such as the UN Charter or the Universal Declaration of Human Rights that support a “right to life” or “right to higher standards of living,” noting they implied a right to water but did not state it. The Geneva Conventions and Protocols declare a right to drinking water only. Two human rights treaties, the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW) and the Convention on the Rights of the Child, both specifically mention a right to water. For women, they declare a right to a water supply, and clean drinking water is mandated for children. Although mention of rights to water can be found in international law documents, they are not clearly defined, and recognized nation-states still need to incorporate them into national law.

RELATED ISSUES

Other issues interrelated with human rights to water are international influences [e.g., globalization and intergovernmental organizations (IGOs)] and governmental obligations to protect marginalized populations and manage water wisely.

Globalization

Globalization has hastened and consolidated the ability of multinational water corporations to search the world for new contracts. Globalization has also injected a new level of haste in developing countries wanting to industrialize, but first needing adequate water infrastructure to attract foreign investment. Countries wanting improved infrastructure but not having the means to pay for it often will enter into agreements for privatization that are not in their citizens’ best interests. Mechanisms for institutionalizing globalization, such as the General Agreement on Tariffs and Trade (GATT), North American Free Trade Agreement (NAFTA), and many bilateral investment treaties (BITs), all assist in establishing free trade, deregulation at the state level, and the promotion of privatization. The negative effects on countries and their citizens with regard to democratic control over their water supplies or their ability to disagree with corporate objectives are analyzed in detail in *Thirst for Control* (9) and *Blue Gold: The Fight to Stop the Corporate Theft of the World's Water* (10).

IGOs

Related to globalization is the influence that IGOs and related multilateral financial institutions such as the World Bank and International Monetary Fund (IMF) have. Loans to developing countries are often preconditioned on the privatization of the country's water and/or sewer facilities (11). At the same time, guarantees regarding percentage of profit or length of contract are given to corporations. Consequently, countries can be assuming risks while corporations receive security from risk. The terms of these agreements can put citizens at risk and increase the likelihood that the poorest in cities will lose access to potable water.

Governmental Obligations

The human right to water is related to the public trust, a doctrine that recognizes the government as trustee for its citizens regarding common resources. Essentially, the government will hold ownership over common properties like water bodies, wildlife, and air; this then gives government the ability to regulate use or access to these common resources so that current and future populations will be able to benefit from them. This doctrine also presumes that a government will give equal access and treatment to all of its citizens, rich or poor, and uphold their right to access. Consequently, the human right to water is dependent on a government that takes its public trust doctrine seriously to declare a human right to safe water, and then to follow through with laws and policy that ensures each person will be able to physically and economically access that water.

Water for the Poor

The rural poor comprise 80% of the world population without access to safe water. The other 20% are the poor in city slums and squatter settlements on the city periphery. Rural people must walk hours to access dirty and open water sources contaminated by animal and human waste. The poor in the city either walk to an open stream or pay water vendors 6 to 12 times the cost of piped water (12). Money spent on expensive water will be money unavailable for food or housing or education.

Women's and Children's Rights

Declaring a human right to water benefits most women and children around the world, because it is commonly their job to search out and carry water back to their homes (13). Rural African women spend 26% of their day getting water from open sources such as rivers, muddy holes, or small lakes. Walking to distant water not only exposes women and children to rape, other physical attacks, schistosomiasis, or more mosquito bites but also takes time away from growing food, earning income, and consistent attendance at school. Children exposed to unsafe water become vulnerable to disease and death. Chemically polluted water endangers children more than adults because children consume more liters of water per kilogram of body weight. A study in Karachi, India concluded that people without water and sanitation spent six times more on medical bills than people linked to those services (14). Health is also compromised if there is not enough water to grow or cook a variety of food. As unsafe water translates into poor health, little money for other necessities, and little time to become educated, it predisposes the poor to a life of weak constitutions and unskilled, part-time labor. They remain unable to move up out of poverty.

Indigenous People

These people have often not enjoyed the same civil rights as other citizens of their country, so it is common for their waters to be diverted for use elsewhere (e.g., a large dam project) or to be polluted from mineral mining or drilling for oil. Civil governments rarely come to their aid and sometimes are the instigators of the water theft or pollution.

Intergenerational Water Resources

Because governments are assumed to be entities existing far beyond any human lifespan, the public trust implies that a government manage its natural resources for current as well as future citizens. In countries with growing populations, this requires a government to plan for future increased demand while managing current resources, which necessitates sustainable use of current resources as well as conservation measures to "produce" more water for future demand. Future generations are a marginalized population because their needs are barely acknowledged in current decision-making and highly discounted in cost/benefit calculations.

Irrigation and Water Management

Agriculture uses over 75% of the available fresh water in the world; therefore increases in efficiency would free up more water available for domestic use. Irrigation that "wastes" water has different driving forces in poor and rich countries. Irrigation in low or no technology countries tends to be the simplest forms, such as flooding the fields, diverting water into irrigation channels for too long, or using unlined irrigation ditches. These farmers generally cannot afford to install high-tech, pinpointed delivery systems, so much water is lost to evaporation or percolation.

In wealthier countries, subsidies for particular crops (e.g., cotton and rice) encourage the growth of water-intensive crops rather than other crops that the region's rainfall can support. If the farms rely primarily on groundwater, choosing a more water-intensive crop essentially exports additional water out of the area. This combination of agricultural subsidies and outdated water laws result in no monetary or legal incentives to conserve water.

CONCLUSION

The global water crisis and growing privatization have increased calls for human rights to water. Organizations interested in the establishment of this right argue that individual governments must protect their citizens from the negative effects of global forces, and that in securing and working toward full establishment of this right, they will fulfill their public trust obligations to their disenfranchised citizens. Creating "new" water by increasing global irrigation efficiency will require either increased expenditures on the part of developing countries, or a transfer of technology and financial assistance to them from the developed countries.

BIBLIOGRAPHY

1. United Nations. (2002). *General Comment on the Right to Water*. Adopted by the United Nations Committee on Economic, Social and Cultural Rights, November 27.
2. Intergovernmental Council of the IHP. (2004). *IHP Contribution to the World Water Assessment Programme (WWAP)*. Document # SC.2004/CONF.203/CLD.32; IHP/IC-XVI/INF.19.
3. Public Citizen's Critical Mass Energy and Environment Program. (2001). *Water Privatization Case Study: Cochabamba, Bolivia*. Public Citizen, Washington, DC.
4. International Consortium of Investigative Journalists. (2003). Metered to death: how a water experiment caused riots and a cholera epidemic. Available: http://cw.groupstone.net/Scripts/WebObj. . .jectid=D1D1364E000000F33188ACAE95C0CBF8.
5. This treaty can be downloaded at: http://www.blueplanetproject.net/english/treaty/.
6. This report can be downloaded at: http://www.unesco.org/water/wwap/wwdr/index.shtml. The most recent version of this report is: UN/WWAP (United Nations/World Water Assessment Programme). (2003). *UN World Water Development Report: Water for People, Water for Life*. UNESCO (United Nations Educational, Scientific and Cultural Organization) and Berghahn Books, Paris, New York and Oxford.
7. Getches, D.H. (1997). *Water Law in a Nutshell*, 3rd Edn. West Publishing, St. Paul, MN.
8. Scanlon, J., Cassar, A., and Nemes, N. (2004). *Water as a Human Right?* IUCN Environmental Policy and Law Paper # 51. International Union for Conservation of Nature and Natural Resources (IUCN).
9. Available: http://www.blueplanetproject.net/english/resources/publications/.
10. Barlow, M. and Clarke, T. (2002). *Blue Gold: The Fight to Stop the Corporate Theft of the World's Water*. The New Press, New York.
11. Public Citizen. (2004). Inter-American Development Bank water and sanitation sector loans promoting privatization. Available: http://www.citizen.org/print_article.cfm?ID=11376; Grusky, Sara. June 29, 2004. IMF forces water privatization on poor countries. Globalization challenge initiative. Available: http://www.waterobservatory.org/library.cfm?refID=33640; Eaton, J.M. July 20, 2002. Critical Perspectives on Water Privatization: Referenced Index and Digest of Major Studies, Reports, and Analyses. corporate accountability committee, Sierra Club. Available: http://www.waterobservatory.org/library.cfm?refID=33775; Grusksy, S. May, 2004. The IMF, the World Bank and the global water companies: a shared agenda. International Water Working Group; Grusky, S. and M. Fiil-Flynn. April 2004. Will the World Bank back down?: water privatization in a climate of global protest. Public Citizen. Available: http://www.worldbankwatch.org.
12. World Health Organization (WHO). (2003). *The Right to Water*. Health and Human Rights Publication Series #3. WHO, Paris, pp. 16, 23.
13. World Health Organization (WHO). (2003). *The Right to Water*. Health and Human Rights Publication Series #3. WHO, Paris.
14. WHO. (2003). *The Right to Water*. Health and Human Rights Publication Series #3. WHO, Paris, p. 24.

ANCIENT WATER AND SOIL CONSERVATION ECOSYSTEMS OF SRI LANKA

LUCIUS O. MENDIS
Colombo, Sri Lanka

INTRODUCTION

Sri Lanka, an island of approximately 65,000 sq km, in the Indian ocean, lying between 6° and 8° North latitude, at the southern tip of the Indian subcontinent, experiences two monsoons, the northeast between October and March and the southwest between April and September, with occasional intermonsoon rains. Northeast monsoon rainfall is spread over much of the island, but the southwest monsoon is largely intercepted by a south-central hill massif, and a dry southwest wind blows over the north and east. On account of topography and rainfall, 103 rivers rising in the central highlands flow in a radial pattern to the sea (Fig. 1). Perennial rivers are called *ganga*, and nonperennial rivers and streams are called *oya, ara, dola*, or *ela* in Sinhala, and *aru* in Tamil. (Sinhalese constitute 74% of the population, Tamils about 12%.) A wet zone and dry zone are defined today; the criterion is 75 inches of annual rainfall. Humid-tropical conditions exist, ideal for year-round cultivation of rice, a water-intensive crop.

ANCIENT DEVELOPMENT

There were three political regions, Rajarata (King's country), Ruhunurata, and Mayarata in ancient Sri Lanka whose written history goes back to the mid-first millennium B.C. Actual writing began early in the Christian era on ola (palm) leaf manuscripts, many still preserved in temples and archives. The Mahavamsa, or great chronicle, and Culavamsa, little chronicle, translated by modern scholars, conform with contemporaneous rock inscriptions.

A famous king, Parakrama Bahu I, said *"Let not a drop of rain water flow to the sea without being made useful to man!"* (Mahavamsa, quoted in Reference 1). Human intervention in nature's hydrologic cycle is seen in the plains of ancient Rajarata and Ruhunurata in the form of river diversion systems and small, medium,

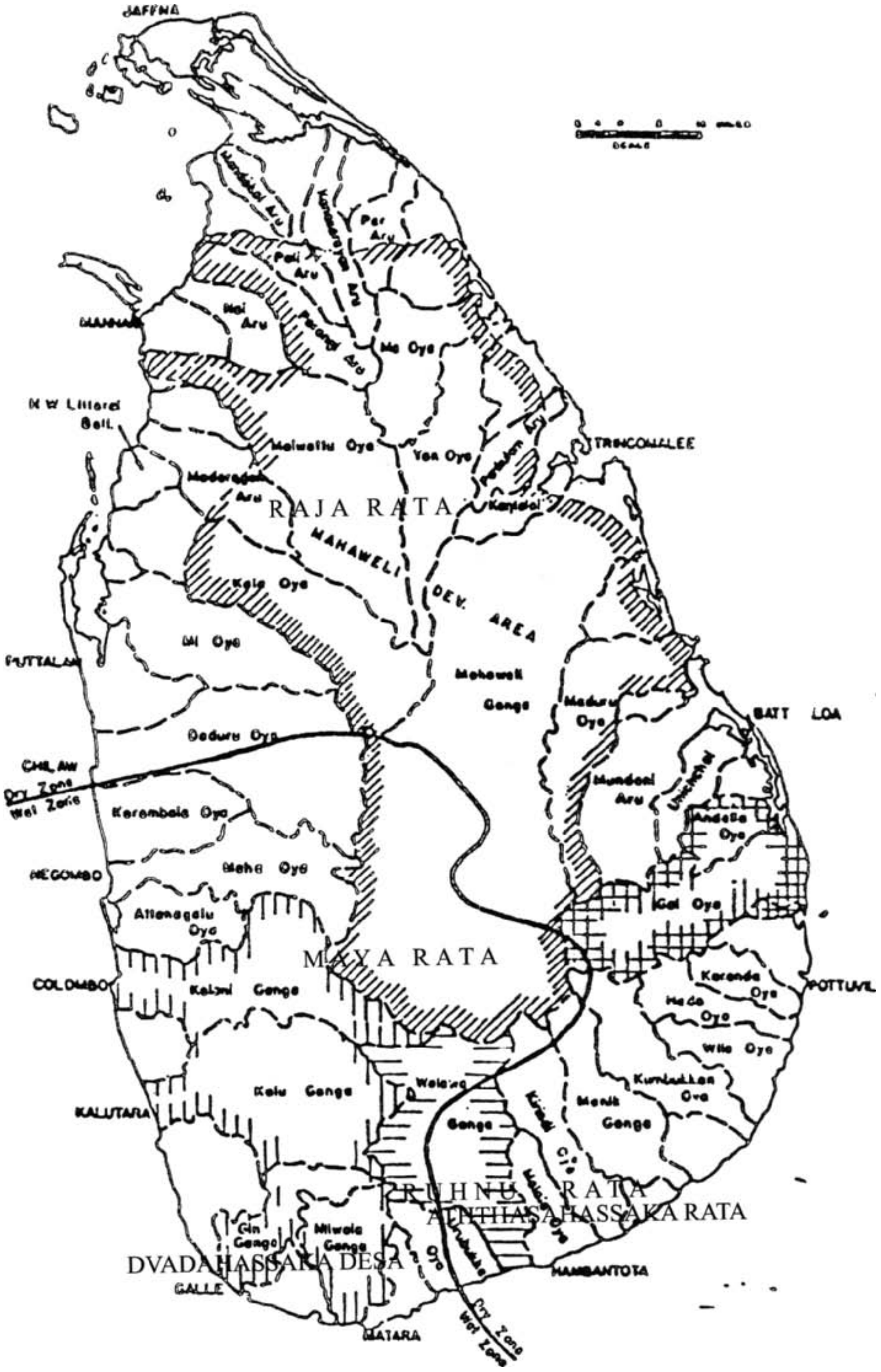

Figure 1. Rivers of Sri Lanka.

and large reservoirs. In wet zone uplands, much later, terraced rice fields saw reuse of irrigation water under gravity, benefiting from forest cover induced fertility. In colonial times, destruction of forests for planting coffee, and later tea, led to adverse impacts (2). Ancient diversion systems include stone anicuts (derived from the Tamil language) or weirs, called *amuna* in Sinhala and *tekkam* in Tamil, channels diverting streams for both irrigated agriculture, and storage in reservoirs. Large reservoirs with massive earth dams were built after the invention of the sluice with its *bisokotuwa* or access tower, forerunner of the valve tower (3, p. 379). Small tanks (Portuguese *tanque*) were synonymous with villages. Small earth dams without sluices, locally known as *vetiyas,* are incorrectly described on topographical survey sheets as *abandoned small tanks*. New metric-scale topo sheets have omitted many of these vetiyas, which checked surface runoff in nonperennial streams for storage in the earth itself, a water harvesting technique understood in India today (4) but not in Sri Lanka. Conservation of fauna began when

Table 1. Contrasting Perspectives on Ancient and Modern Projects

	Hydraulic Engineering Perspective—Hard Technology Design of Modern Projects	Ecosystems Perspective—Soft Technology Design seen in Ancient Projects
1. Water	Inanimate, active	animate, passive
2. Small tank	Inefficient, early stage in evolution and development of irrigation systems, to be submerged by a large reservoir built later	microwater and soil conservation ecosystem in downstream development area below a large reservoir; name of tank often identical to name of village
3. Large reservoir	Efficient irrigation system in combination with downstream channel distribution irrigation system without any small tanks in its command area	Main item in a macrowater and soil conservation ecosystem with microwater and soil conservation ecosystems in its command area
4. Diversion structure, dam or weir, and diversion channel	Built to augment a large reservoir—last stage in development of ancient irrigation systems in Sri Lanka (5,6)	Earliest stage in irrigated agriculture and evolution and development of ancient water and soil conservation ecosystems in Sri Lanka (7)
5. Vetiya, breached small earth embankment or bund, found in dry zone jungles	Embankment or bund of abandoned small tank	Deflection structure—microwater and soil conservation ecosystem; maintains water table
6. Downstream development area in modern projects	Cleared of all vegetation to lay out channel irrigation systems	Should be designed as a series of microwater and soil conservation ecosystems, including forest areas, as in ancient systems
7. Forest areas	Limited to catchment areas in modern projects	Not only in catchment areas in ancient projects, but interspersed with fields in downstream development areas for better nutrient flow

King Devanampiya Tissa, on a hunting trip about 223 B.C., was converted to Buddhism. Arahant Mahinda, son of Emperor Asoka of India, preached a sermon at Mihintale, a mount near the capital city Anuradhapura: *"The birds of the air and the beasts of the jungle have as much right to life as thou, O great King!"* (4). Hence, the ancient systems involved conservation of water and *interactions* of water, soil, flora, and fauna, with *functions* including irrigation, drainage, flood control, and domestic water supply. They should be called *water and soil conservation ecosystems*, as described in Table 1. Three aspects merit study:

- —their evolution and development beginning in about the mid-first millennium B.C.;
- —their stability and sustainability for about 17 centuries until about the thirteenth century A.D.; and
- —their apparently irreversible decline after the Parakrama Bahu I (1153–1186) era.

These are being addressed in a multidisciplinary project, *Science and Civilisation in Sri Lanka,* started recently following Joseph Needham's *Science and Civilisation in China* (8).

ANCIENT RAJARATA

Anuradhapura in Rajarata was the capital city of Sri Lanka from about the fourth century B.C. to about the tenth century. Polonnaruwa, further east, became an alternative capital in about the eighth century on account of invasions from southern India and remained the capital until the thirteenth century, after which drift to the southwest began (9). Water and soil conservation ecosystems in Rajarata were highly developed in the Anuradhapura period (5,10–13). Only restoration was necessary in modern times, until the advent of new projects, Gal oya (1947–1949) and the aid funded Mahaweli (1978–1986). However, the Sri Lanka National Atlas does not have a chapter on the ancient systems.

In western Rajarata, Anuradhapura city tanks were built in stages: Basawakkulama (fourth century B.C.), Nuwara weva (reservoir) (first century), Tissa weva (third century) and Nachchaduwa weva (fifth century). The $54\frac{1}{2}$-mile-long channel, Kalaweva Jayaganga (River of Victory), starts from the right bank sluice of the great twin reservoirs Kalaweva–Balaluweva built by King Dhatusena (459–477), feeding Tissa weva directly, with a branch channel to Nachchaduwa. Brohier (10) documented the *Interrelation of Large Reservoirs and Channels* in western and eastern Rajarata. Nicholas and Paranavitana (12) further expanded this (Fig. 2).

The drought-resisting capability of this interconnected system augmented *cascades of small tanks* whose names were synonymous with village names [Fig. 3 (14)] [Fig. 4 (15)].

In eastern Rajarata, ancient ecosystems were developed through more than 11 centuries from the first century to the twelfth century (5,13,16,17). The Elahera stone diversion anicut and the first 24 miles of the Elahera canal were built by King Vasabha (65–109); King Mahasena (276–303) built Minneriya weva at the tail end of this canal and started the extension of the

Elahera–Minneriya canal beyond Minneriya weva. King Agga Bodhi I (575–608) completed this extension, and King Agga Bodhi II (608–618) built Gantala weva, the modern Kantalai reservoir, at the tail end of this canal, and an extension to the sea near Trincomalee. Finally, King Parakrama Bahu I (1153–1186) improved the whole system and constructed three large reservoirs, called Seas of Parakrama: Parakrama Samudra, restored in modern times; Parakrama Talaka still unidentified; and Parakrama Sagara, created by raising the Elahera anicut and strengthening the earthen canal embankment between Elahera and Konduruweva. This formed a series of reservoirs because there is a fall of about 60 feet in this reach (5,15) (Fig. 5). Parakrama Sagara was also called *Koththabadhdhanijjara*, translated as *the weir furnished with a reservoir* and also as *the reservoir whose flood escape was walled up* (18,19).

ANCIENT RUHUNURATA

In ancient Ruhunurata, remains of thousands of ancient irrigation works are found in the southeastern area, once described as Aththasahasaka rata, the region of 8000 villages, and the southwestern area called Dvadahassaka desa, the region of 12,000 villages (Fig. 1). But they have not been studied and understood as those in ancient Rajarata, and modern development has impacted the ancient projects adversely.

JAFFNA PENINSULA

The climate and topography of the northern Jaffna peninsula are quite distinctive. Rainfall is comparatively sparse, and underlying limestone makes groundwater the

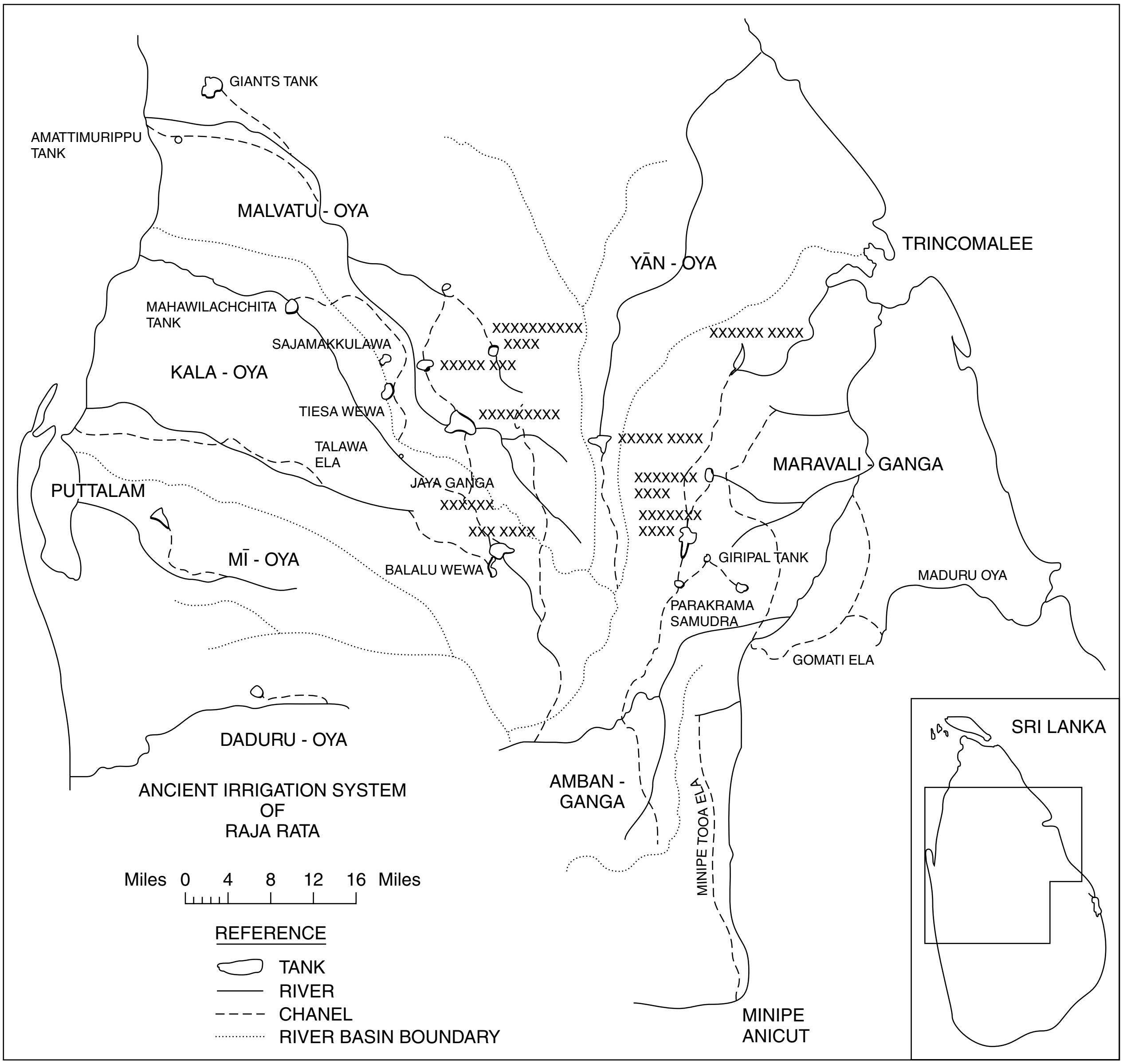

Figure 2. Interrelation of large reservoirs and channels.

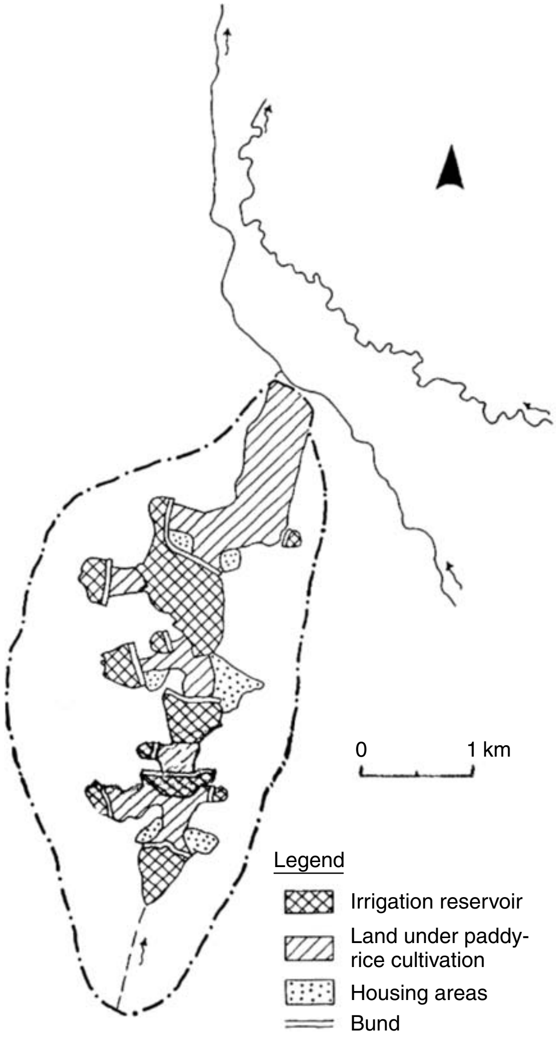

Figure 3. General layout of the cascade of four villages located in a local catchment.

principal source of water, extracted from dug wells, using the well sweep from ancient times (20).

MODERN DEVELOPMENT

Irrigation engineers today use a hydraulic engineering approach to design new projects, almost diametrically opposite to an ecosystems approach. In the hydraulic engineering or hard technology perspective, water is treated as inanimate and active, in the ecosystems or soft technology perspective, water is animate and passive. Consequences are shown in Table 1.

A four-stage hypothesis for the evolution and development of ancient irrigation systems follows the hydraulic engineering perspective (6,21):

—rainwater tanks from which water was bailed

—small storage reservoirs, called village tanks

—large storage reservoirs that submerged a number of small tanks

—augmentation of large reservoirs by diverting a river

Here, the evolution and development of storage reservoirs, in terms of storage capacity is followed by river diversion as a later development. But river diversion irrigation systems or water management in space existed before storage reservoirs or water management in time (which depended on the invention of the sluice) in all ancient civilizations.

Therefore, an alternative seven-stage hypothesis was proposed (7):

1. Rain-fed agriculture
2. Temporary or seasonal river diversion and inundation irrigation on river banks
3. Permanent river diversion and development of canal irrigation systems
4. Development of weirs and spillways on contour channels
5. Invention of the sorowwa (sluice) with its bisokotuwa (access tower)
6. Construction of small, medium, and large storage reservoirs
7. Damming a perennial river using a sluice for temporary river diversion, or the twin-tank technique, as in Kalaweva–Balaluweva in the fifth century

Meanwhile, a map entitled the Water Resources Development Plan (1959) showed proposed new large reservoirs, submerging ancient small village tanks, following the third stage of the four-stage hypothesis. Of ancient water and soil conservation ecosystems, only very large reservoirs and channels were shown.

In the southern area, location of the gigantic new Uda Walawe and Lunuganvehera reservoirs, follows the third stage in the four-stage hypothesis. Both Uda Walawe (1965–67) (Fig. 6) and Lunuganvehera (1978–1986) were built without considering alternative upstream sites. Environmental problems have made both these new projects unstable and unsustainable, contributing to insurrections against the government in 1971 and 1987–1989 that were suppressed with much loss of life and damage to property (22–25).

A similar possibility exists if Moragahakande, selected from the Water Resources Development Plan, 1959, for cost feasibility studies, is built (Fig. 5), compounded by the ethnic discord between separatist Tamils and the government. This large new reservoir will also destroy any possibility of restoring the potentially stable and sustainable ancient Parakrama Sagara.

A RIVER FOR JAFFNA—THE ARUMUGAM PLAN

A proposal called A River for Jaffna, prepared 50 years ago and partly implemented, will avoid all potential problems of Moragahakande. It involves enhancing both the quality and quantity of Jaffna groundwater by infinite dilution over the years by floodwater from the mainland Iranamadu reservoir (Fig. 7).

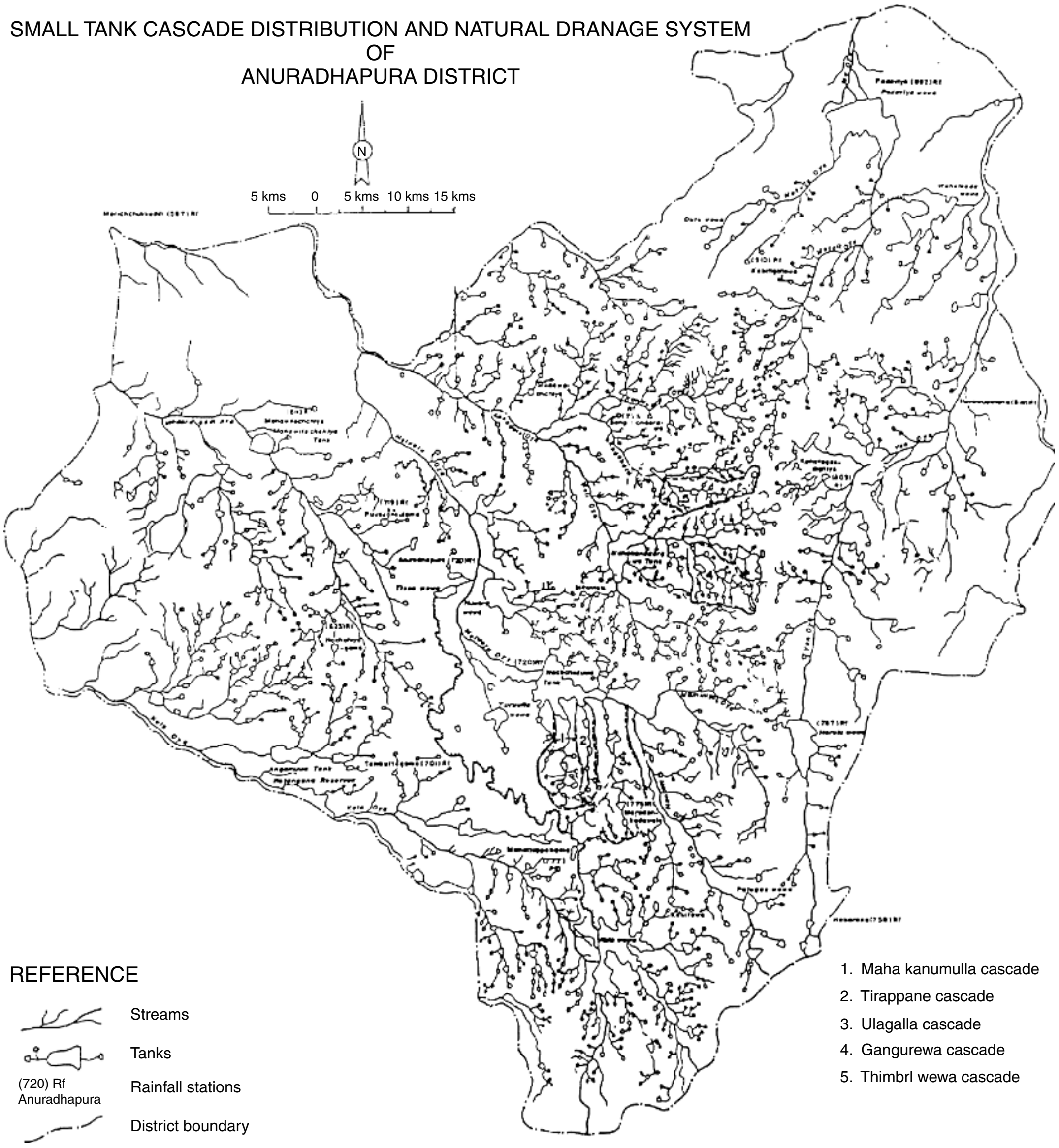

Figure 4. Cascades of small tanks.

BIBLIOGRAPHY

1. Weeramantry, C.G. (2000). *Separate Opinion* in the Gabcikovo Nagymaros case in the World Court, 1997. Republished for the Institution of Engineers, Sri Lanka. Vishwa Lekha.
2. Senanayake, D.S. (1934). *Agriculture and Patriotism*. Government Press, pp. 39–40.
3. Parker, H.W. (1909). *Ancient Ceylon*. Lusacs, London.
4. Center for Science and Environment. (1997). *Dying Wisdom*. State of India's Environment 4, p. 354.
5. Brohier, R.L. (1934). *Ancient Irrigation Works in Ceylon*. Government Press, Colombo.
6. Needham, J. (1971). *Science and Civilization in China*. Vol. 4, Part 3, Cambridge, p. 369.
7. Mendis, D.L.O. (1986). Evolution and development of irrigation ecosystems and social formations in ancient Sri Lanka. *Trans. Inst. Eng. Sri Lanka* 13–29.
8. Mendis, D.L.O. and Kamaladasa, B. (Eds.). (2003). *Water for People and Nature*. Arumugam Commemoration volume. Sri Lanka Water Heritage. Sarvodaya Vishwa Lekha, Ratmalana.
9. Indrapala, K. (1971). *Collapse of the Rajarata Civilisation and the Drift to the Southwest*. Ceylon Studies Seminar, University of Peradeniya.
10. Brohier, R.L. (1937). Inter-relation of groups of large reservoirs and channels in Rajarata. *J. R. Asiatic Soc. Ceylon* **34**(90): 70–89.
11. Brohier. (1991).
12. Nicholas, C.W. and Paranavitana, S. (1960). *Concise History of Ceylon*. Gunasena.
13. Arumugam, S. (1969). *Water Resources of Ceylon*. Water Resources Board, Colombo.
14. Sandell, K. (1988). Ecosystems in theory and practice. Farmers perspectives on water, nutrients and sustainability in low-resource agriculture in the dry zone of Sri Lanka. *Linkoping*, p. 94.

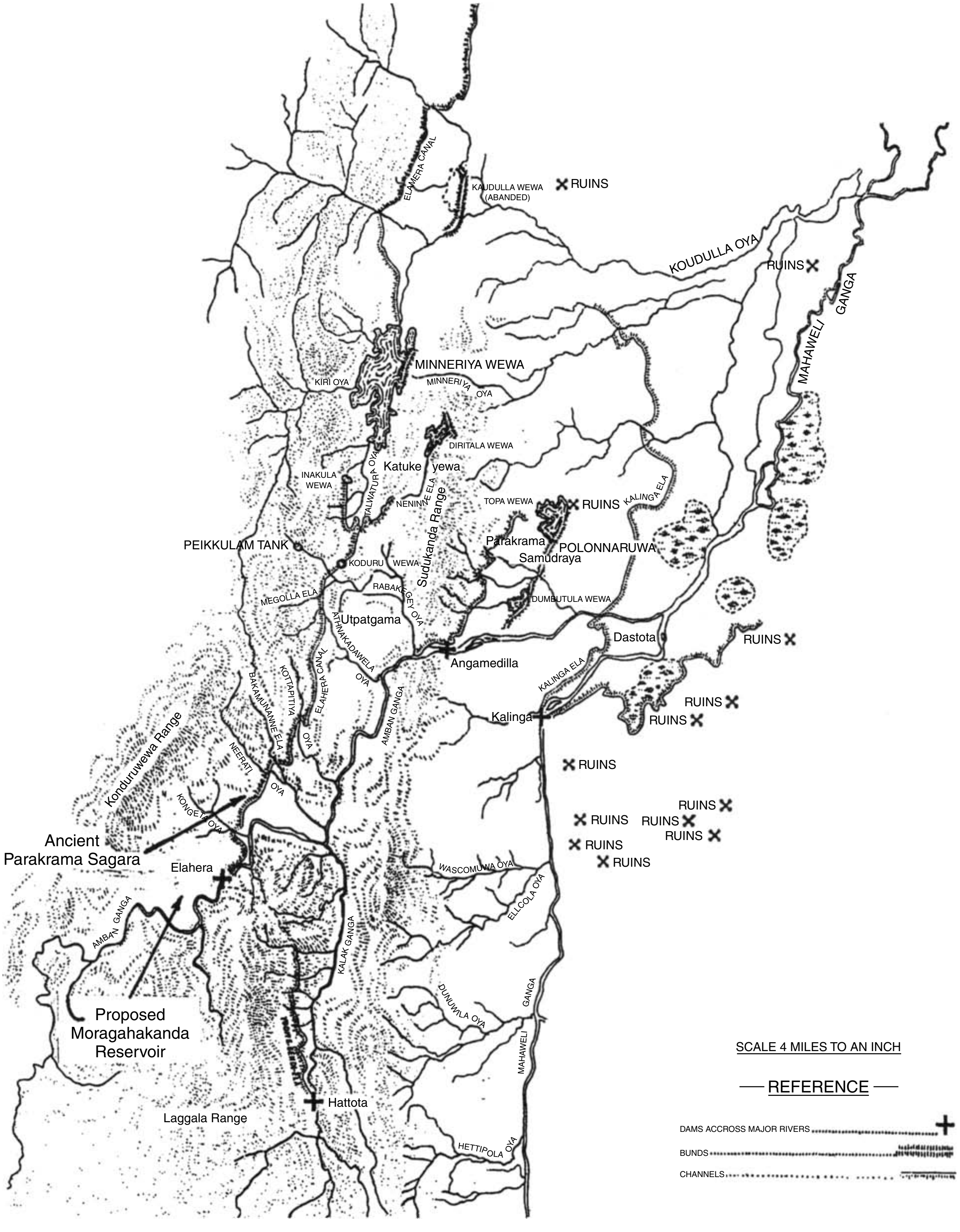

Reference : The Tamankaduwa District and the Elahera - Minneriya Canal R L Brohier, 1941
Diagram No.2 facing page 4, (Ancient Parakrama Sagara and proposed Moragahakanda reservoir, added)

Figure 5. Diagram showing the hydrography of Tamankaduwa and the ancient irrigation schemes.

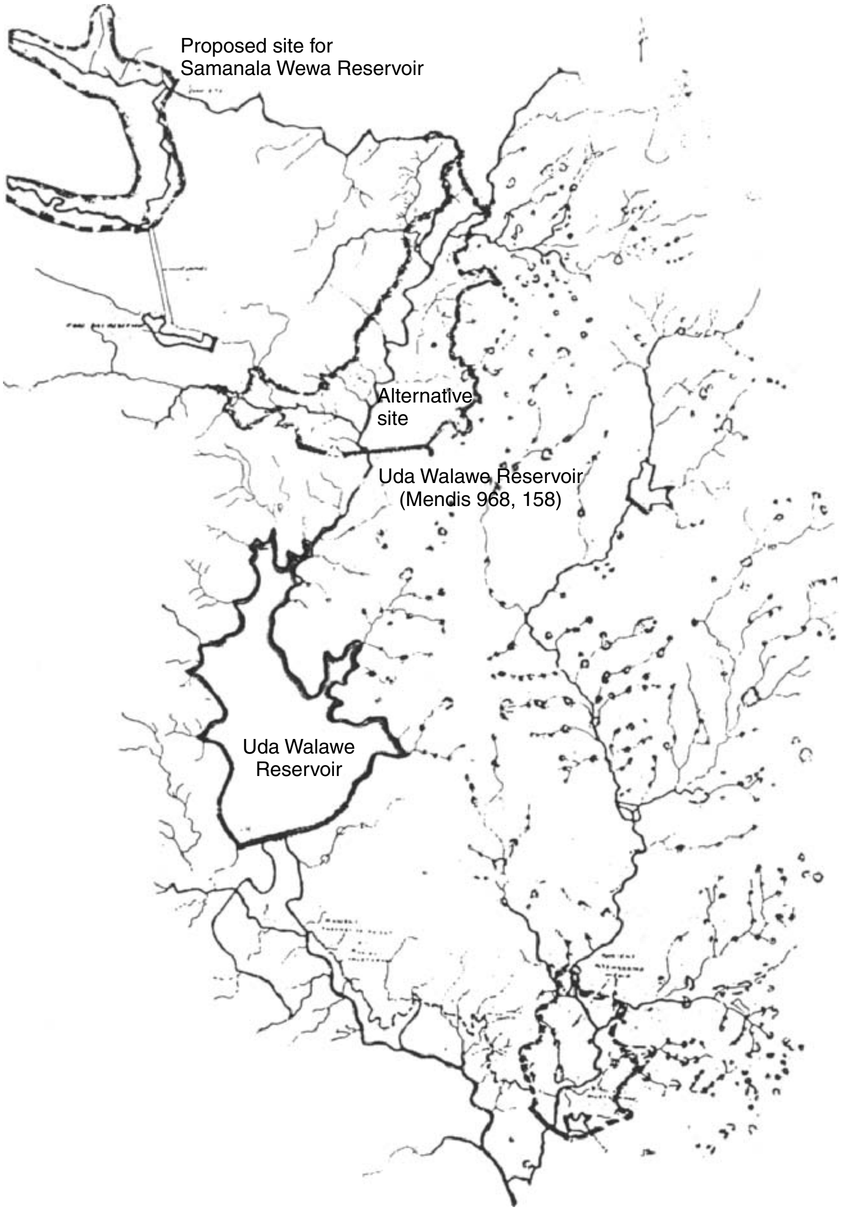

Figure 6. Walawe basin development ancient dispersed small-scale system vs modern centralized large-scale system.

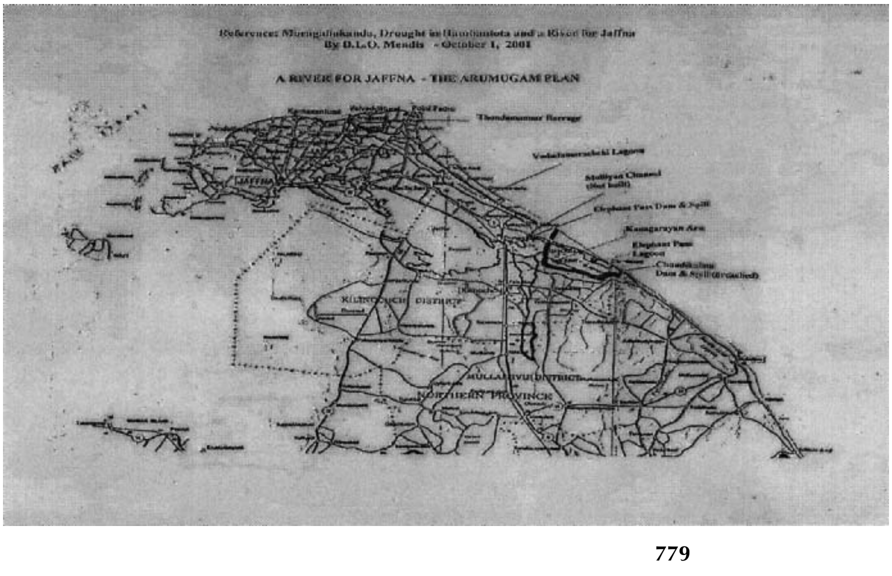

Figure 7. The Arumugam plan.

15. Panabokke, C.R. *Small Tank Cascades in the Rajarata*. In: Arumugam Commemoration, pp. 123–132.
16. Ward, SirH. (1857). Minutes of inspection. *Ceylon Almanac*. quoted in Brohier, 1934, Vol. 1, pp. 28–33.
17. Mendis, D.L.O. (2002). *Water Heritage of Sri Lanka*. History of Water Conservation series. Sarvodaya Vishwa Lekha, Ratmalana.
18. Geiger, W. (1953). *Mahavamsa (Translation)*. Government Press, Colombo.
19. Geiger, W. (1960). *Culture of Ceylon in Medieval Times*. Otto-Harrosowitz, Weisbaaden.
20. Tennant, J.E. (1859). *Ceylon*. Longmans, London. Reprinted Asian Educational Services, 1996, Vol. II, p. 552.
21. Brohier. (1956).
22. Mendis, D.L.O. (1989). *Hydraulic Civilisations, Irrigation Ecosystems and the Modern State*. Second Professor E O E Pereira Commemoration Lecture, Institution of Engineers, Sri Lanka.
23. Mendis, D.L.O. (1989). Development of underdevelopment in southern Sri Lanka: Destabilisation of ancient irrigation ecosystems by the impact of hydraulic engineering. *Trans. Inst. Eng. Sri Lanka:* 31–50.
24. Mendis, D.L.O. (1991). Non-ethnic causes of violent conflict in Sri Lanka: Environmental degradation on account of destruction of ancient irrigation ecosystems by the impact of hydraulic engineering. *Proc. 41st Int. Pugwash Conf. Sci. World Affairs*, Continental Press, Singapore, pp. 698–709.
25. Mendis, D.L.O. (1995). *Past, present and future conflict over water in Sri Lanka*. *Proc. 45th Int. Pugwash* Conf., Hiroshima. World Scientific, Singapore, pp. 791–800.

BEN FRANKLIN'S GULF STREAM WEATHER AND SWIM FINS

William E. Marks
Water Consciousness, Inc.
Martha's Vineyard,
Massachusetts

THE DISCOVERY OF THE GULF STREAM

In 1950, renowned nature writer Rachel Carson's book, *The Sea Around Us*, was published by Oxford University Press (1). In this book, Carson discusses at length, Benjamin Franklin's chart of the Gulf Stream (Fig. 1). A copy of Ben Franklin's chart is reproduced in Carson's book as an endpaper. In her writings, Carson credits the accuracy of Franklin's chart that was first created in 1769 and uses it and other information to help explain the flow of ocean currents.

In all fairness to the historic record, Franklin did not discover the Gulf Stream. The pre-Columbian Florida Seminole Indians, who traveled local waters, probably were aware of the Gulf Stream's current and its influence on their local sea travel and fisheries. As well, in 1492, Christopher Columbus pioneered Gulf Stream observations while sailing the equatorial current near the West Indies. While taking soundings, Columbus observed that the warmer surface waters were moving faster and in a different direction than the colder deeper waters (2).

In addition to the above, and thanks to research of Midshipman Pamela Phillips of the U.S. Naval Academy and Richard Gasparovic of the Applied Physics laboratory of John Hopkins University, we have the following time line:

1513: Ponce de Leon describes the Florida Current.
1519: Existence of the Gulf Stream is widely known to sailors.
1575: Andre Thevet associates the Gulf Stream with the rivers of the Gulf of Mexico.
1590: John White becomes the first person to mention the counter-currents on the coastward side of the Gulf Stream.
1606: Lescabot describes the meeting of the Labrador and Gulf Stream currents.
1660: Varenius classifies the Gulf Stream as a "perpetual special" motion of the sea and describes it as a giant stream.
1665: Kircher publishes showing the Gulf Stream.
1685: Happelius publishes second chart showing the Gulf Stream.
1769: Benjamin Franklin and Timothy Folger design and publish their chart of the Gulf Stream with notes on how to avoid it.

As expressed above, the first actual record of the Gulf Stream was on April 22, 1513, as related in the Chronicle of Ponce de Leon's Expedition, which eventually led to the so-called discovery of Florida. This 1513 document states:

> A current such that, although they had great wind, they could not proceed forward, but backward and it seemed that they were proceeding well; and at the end it was known that in such wise current was more powerful than the wind (3).

Further maritime observations of the Gulf Stream were recorded on an ongoing basis after its initial discovery. However, these observations remained the "secret" of the ship's captains and pilots. This fact was noted in an article by Franklin Bache, published by the American Philosophical Society in 1936 (4):

> Apparently the Spaniards soon learned that this northerly flowing current was not merely a local current but one of wide extent; for six years later [after 1513] when Antonio de Alaminos set out for Spain from Vera Cruz, he sailed northward with the Gulf Stream for a number of days before turning east toward Europe. This same Aliminos was pilot of Ponce de Leon's expedition of 1513, when the Gulf stream was first noted. It is, therefore, quite proper to credit the discovery of the Gulf Stream to Alaminos (Marmer H. A. *The Gulf Stream and its Problems*).
>
> For two and half centuries following its discovery the growth of knowledge regarding the Gulf Stream was slow. The story is told in detail by Kohl and more briefly by Pillsbury. During this period, to be sure, the mariner and more especially the whaler, became acquainted with the Gulf Stream throughout the greater part of its course. Much of this information, however, was kept as a professional secret,. . . (Marmer, 1929).

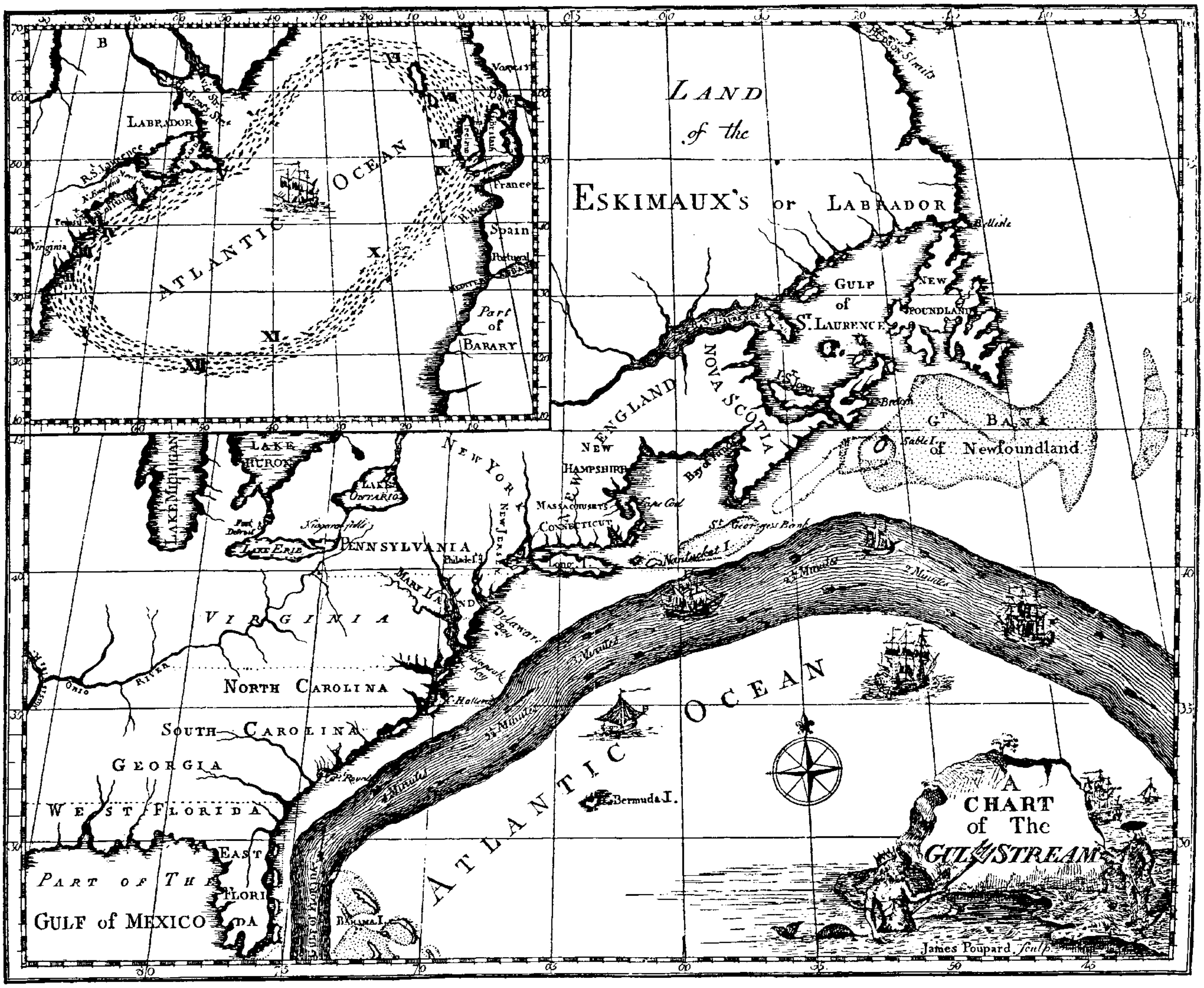

Figure 1. A chart of the Gulf Stream.

In 1513, after Ponce de Leon landed in Florida in the area known today as St. Augustine, he attempted to return to Puerto Rico after failing to locate the Fountain of Youth. However, as documented above, the current was so strong that his ships floated northward, even though they were sailing with a fair wind that should have taken them southward.

Six years later, Anton de Alaminos, found himself guiding Cortes's treasure ships from Mexico. With his prior knowledge as a pilot for Ponce de Leon, Alaminos used the northerly current of the Gulf Stream to his advantage.

Even though the discovery of the Gulf Stream by Alaminos was lost to history, there was another sailor who documented the current in 1735. Captain Walter Haxton, was a shipowner who also had a financial interest in understanding and using the currents of the ocean to his advantage. Besides being a captain, Haxton also owned several tobacco farms in the Maryland area and often shipped his produce to buyers in London.

In 1735, Haxton created the first detailed chart of the Chesapeake Bay area, which included bottom soundings, prevailing winds, and water current velocities and direction. As an addendum to this chart, Haxton plotted the latitude and longitude of the Gulf Stream (Northeast Current), as well as its direction and velocity.

Some 64 years before Franklin's chart in 1769, Haxton penned the following:

> It is generally known by those who trade to the northern parts of America that the current which comes out of the Gulph of Florida runs constantly along the coast of Carolina and Virginia and considerably further to the Northward, varying its course as it is obstructed by shores. Now if said current always runs nearly in the same part or space of the ocean (as from a great number of Tryals and observations which I have made in 23 voyages to Maryland, I have reason to think it does), the knowledge of its Limits Course and Strength may be very useful to those who have occasion to sail in it.

For some reason, Haxton's chart and related information about the Gulf Stream never received wide distribution or application. It was due to such oversight, as well as centuries of secrecy by various captains over the existence of the Gulf Stream, that the stage was set for the Franklin–Folger chart to make a splash in the annals of oceanic history. So much so, that most people today still harbor the belief that Benjamin Franklin discovered the Gulf Stream.

THE FRANKLIN/FOLGER GULF STREAM CHART

When Franklin had his and Folger's version of the Gulf Stream printed on a navigational chart in 1769, there was no such chart available. Franklin could only create the chart after receiving the benefit of Timothy Folger's drawings and advice. At the time, Folger was an active Nantucket whaling captain. Folger is credited for depicting the width and breadth of the Gulf Stream's current by drawing it on an Atlantic Ocean chart. Folger's chart included notes on how the Gulf Stream's current influenced the speed of ships crossing the Atlantic between America and Europe. Folger's drawing of the shape and speed of the Gulf Stream's current were extremely accurate for his time and in general, hold true to this day. Folger drew the Gulf Stream on a chart in answer to Franklin's inquiry about the forces of nature that may be impacting the speed of ships crossing the Atlantic Ocean. On another note, Folger was Franklin's cousin, twice removed (Nantucket Historical Society).

Today, it is generally recognized that the Franklin–Folger chart represents an accurate summary of the average width, length, and direction of the Gulf Stream, as well as the average speed of its current.

The name "Gulph Stream" was first suggested in 1762 by Benjamin Franklin because of the stream's sourcing from the Gulf of Mexico. Prior to 1762 it was generally known as the "Northeast Current."

The initial concept of creating Franklin's famous chart of the Gulf Stream was prompted in 1768 when Franklin was in London serving as Deputy Postmaster General for North America (Van Doren, 1938). While in London, Franklin received complaints from America as to how westbound mail packets [English ships] took up to two weeks longer than American merchant ships (5) *Am. Philos. Soc.* lxxvi (5). (1936). At about the same time, the English postal authorities had written to Franklin asking similar questions. These inquiries from Americans and the English about the speed of ships crossing the Atlantic prompted Franklin to do some research.

While researching the answers to these complaints in 1768, Franklin made inquiries of Timothy Folger, who happened to be in London at the time. Folger informed Franklin that based on experience, Nantucket whalers had developed a working knowledge of the Gulf Stream between Europe and America. This knowledge came from the whalers following the migration patterns of whales, which were consistently found north or south beyond the edges of the Gulf Stream, but never inside the Gulf Stream. Whaling captains from Nantucket and Martha's Vineyard knew that whales could be found feeding along the Gulf Streams plankton-rich boundaries. They also had learned that sailing back to New England whaling ports could be achieved in less time by staying north of the Gulf Stream and thereby avoiding its current.

Over time, Nantucket whalers learned to use the current of the Gulf Stream to their advantage when sailing in a northerly direction to England, but avoiding it when sailing in a southerly direction back to America. Folger told Franklin that American captains often told British captains about avoiding the current, but their hubris caused them to ignore the Americans advice. "They were too wise to be counselled by simple American fishermen," Folger told Franklin.

A letter from Franklin dated October 29, 1768, tells of Folger's role in the creation of the Gulf Stream chart:

> Discoursing with Captain Folger a very intelligent Mariner of the Island of Nantucket in New England concerning the long passages made by some Ships bound from England to New York I received from him the following information. Vizt
> That the Island in which he lives is Inhabited Chiefly by people concerned in the Whale Fishery, in which they employ near 150 Sail Vessels, that the whales are found generally near the Edges of the Gulph Stream, a strong current so called which comes out of the Gulph of Florida, passing Northeasterly along the Coast of America, and the turning off most Easterly running at the rate of 4, 3 1/2, 3 and 2 1/2 Miles an Hour; that the Whaling Business leading these people to Cruise along the Edges of the Stream in quest of Whales,. . ."

When the chart was created in 1769, Franklin wrote the following notes based in large part on the advice he received from Nantucket's Timothy Folger:

> By observing these directions and keeping between the stream and the shoals [that lie to the south of Nantucket and St. George's Banks], the passage from the Banks of Newfoundland to New York, Delaware, or Virginia, may be considerably shortened; for so you will have the advantage of the eddy current, which moves contrary to the Gulph Stream. Whereas if to avoid the shoals you keep too far to the southward, and get into that stream, you will be retarded by it at the rate of 60 or 70 miles a day.
> The Nantucket whale-men being extremely well acquainted with the Gulph Stream, its course, strength and extent, by their constant practice of whaling on the edges of it, from their island quite down to the Bahamas, this draft of that stream was obtained from one of them, Capt. Folger, and caused to be engraved on the old chart in London, for the benefit of navigators
> . . . Note, the Nantucket captains who are acquainted with this stream, make their voyages from England to Boston in as short a time generally as others take in going from Boston to England, viz. from 20 to 30 days.
> A stranger may know when he is in the Gulph Stream, by the warmth of the water, which is much greater than that of the water on each side of it. If then he is bound to the westward, he should cross the stream to get out of it as soon as possible (6).

Besides the above two sources referring to Timothy Folger's role in creating the Gulf Stream chart, we also have the following quote from Franklin's *Maritime Observations*. These observations were written at sea in August, 1785, and expressed in his letter to David Le Roy:

> I then observed it was a pity no notice was taken of this current upon the charts, and requested him [Timothy Folger]

> to mark it out for me, which he readily complied with, adding directions for avoiding it in sailing from Europe to North America. I procured it to be engraved by order from the general post-office, on the old chart of the Atlantic, at Mount and Page's, Tower Hill; and copies were sent down to Falmouth for the [English] captains of the packets [ships], who slighted it however; but is since printed in France, of which edition I annex a copy.

It is in Franklin's *Maritime Observations* that we find mention of his two charts of the Gulf Stream. Both of which he published—the first chart in England in 1769, and the second in France circa 1778. It was also during this 1785 voyage that Franklin expressed his theory about how the Trade Winds create the Gulf Stream. Franklin's theory was that the winds created the Gulf Stream current by the friction of the moving air on the surface of the water. This theory, in general has remained correct to this day. As expressed in Franklin's words:

> This Stream is probably generated by the great accumulation of water on the eastern coast of America between the tropics, by the Trade Winds which constantly blow there. It is known that a large piece of water ten miles broad and generally only three feet deep, has by a strong wind had its waters driven to one side and sustained so as to become six feet deep, while the windward side is laid dry.
> ... The power of the wind to raise water above its common level in the sea is known to us in America, by the high tides occasioned in all our seaports when a strong northeaster blows against the Gulf Stream.

After having the chart of the Gulf Stream printed, Franklin made copies available to British captains. However, Franklin noted that the British captains refused to believe in information about the Gulf Stream proffered by Nantucket whalers. On similar note, the Franklin–Folger chart, and its related information about the Gulf Stream were also rejected by English captains. For this reason, as well as the revolt of the colonies against England in the early 1770 s, Franklin's 1769 chart of the Gulf Stream became extremely rare.

It is believed by some historians that Franklin and other Americans in England may have suppressed the distribution of this 1769 chart. The reason that Franklin suppressed his 1769 chart of the Gulf Stream may have been to prevent the English Navy from using it to its advantage against the colonies. This has merit because it was in 1786 and after the Revolutionary War, that Franklin again published Folger's sketch of the Gulf Stream as part of an article in the *Proceedings of the American Philosophical Society of Philadelphia* (7).

When Franklin was returning to America from London in March 1775, he took temperature measurements of the Gulf Stream four times each day. During this voyage he discovered that the Gulf Stream was warmer than surrounding waters and that it had a color different from the colder waters on either side. These observations were consistent with those of Nantucket whalers.

In 1776, when Franklin was sent by the colonies as envoy to Paris to negotiate a treaty, he again took temperature measurements of the Gulf Stream and kept a record of his observations. These observations rekindled Franklin's interest in understanding the Gulf Stream. Upon arrival in France, he had the original Folger chart of the Gulf Stream copied and printed by Le Rouge. Soon after, the Le Rouge chart was made available to French captains sailing to America with arms and supplies. Even though the Le Rouge version of the Gulf Stream chart is undated, the copy in the Bibliotheque National has the following note inscribed on its underside, "communique' par M. de Franklin en 1778."

In 1785, at age 79, and five years before his death, Franklin made his last transatlantic voyage when he sailed home from France on the London packet. Continuing with his research in to the Gulf Stream, Franklin directed his grandnephew, Jonathan Williams, Jr., when and how to take temperature readings of the Gulf Stream. It was during this final voyage that Franklin wrote *Maritime Observations*. This paper was published in 1786 and includes the best known versions of the making of the Franklin–Folger chart of the Gulf Stream, as well as its related genesis.

The original Franklin–Folger chart of 1769 was lost to history for close to 200 years and was assumed forever lost. During that time, the only surviving record of the original 1769 chart was expressed in the writings of Franklin and others. Concern for finding a copy of the 1769 chart surfaced from time to time. For example, on April 24, 1936, The American Philosophical Society printed an article entitled, "WHERE IS FRANKLIN'S FIRST CHART OF THE GULF STREAM?." The article was authored by Franklin Bache, and was an international plea for people to search for a copy of the chart. It was not until 52 years later that this plea was answered. And, most appropriately, by a scientist who had interest in studying ocean currents.

In 1978, Dr. Philip L. Richardson of Woods Hole Oceanographic Institute found the original 1769 Franklin–Folger chart while doing research in France. As expressed by Richardson (8) *Science* (1980). **207**; 8.

> In September 1978, I found two prints of the Franklin–Folger chart in the Bibliotheque Nationale in Paris (Fig. 1) (7). It occurred to me that a copy of the chart might have been saved by the French because Franklin was envoy to France from 1776 to 1785 and both Franklin and his ideas were highly regarded by the French.
> Richardson would later publish his in-depth Gulf Stream research in *The Geographical Magazine*, London, LII (8), pgs. 575–58, in an article entitled "Progress on the Gulf Stream (9)."

When the author of this article contacted Dr. Richardson by email at Woods Hole Oceanographic Institute, he answered:

> I discovered the 1769 chart while I was working on Gulf Stream Rings, which are eddies that pinch off from Gulf Stream meanders. We were tracking them, measuring their paths in the ocean and their decay. I was spending a year in Paris analyzing some data and writing papers.
> I was thrilled by my discovery and spent a lot of time in further research of the Gulf Stream and Ben Franklin, and other early charts of the Gulf Stream. Based on that, I published the 1980 paper in Science vol 207, 643–645 and one in the book

Oceanography: *The Past*, (1980). Springer Verlag, New York. I also worked on derelict ship trajectories in the Stream, and compared them to drifting buoy trajectories and with Walter, Haxton's 1735 description of the Stream (10) (1982, *J. Mar. Res. Supp.,* **40**: 597–603).

Today, Dr. Philip L. Richardson has the status of Scientist Emeritus at Woods Hole Oceanographic Institute, where he continues to pursue his passion of oceanographic research.

FRANKLIN AND WEATHER

"Some are weatherwise, some are otherwise," wrote Franklin under his pen name of "Poor Richard" in 1735.

It was also in Poor Richard's Almanac that Franklin dared to be so weatherwise that he published some of the first recorded weather forecasts. Poor Richard's Almanac was first published in 1732 under Franklin's nom de plume of Richard Saunders.

Franklin recorded some of his first theories on weather phenomena while in Philadelphia in 1743. On the evening of October 21, he had plans to observe a lunar eclipse which was scheduled to occur at 8:30 P.M. As Franklin waited with anticipation to watch the eclipse, his plans were suddenly interrupted when a fast moving violent storm blocked his view of the moon.

As the following weeks passed, Franklin read in various newspapers that the violent storm tracked northeasterly, causing considerable damage from Virginia all the way to Boston. However, much to Franklin's surprise, the newspapers from Boston reported accounts of the lunar eclipse, and that the storm arrived in Boston long after the eclipse had come and gone.

This series of events intrigued Franklin; he became curious why Boston, which was located hundreds of miles northeast of Philadelphia, would experience the delayed arrival of the storm. Another factor that intrigued Franklin was that the winds and the direction of the storm seemed contrary. After much inquiry and observation of other storms, Franklin told his friend Jared Eliot that his research had led him to the "very singular opinion that, though the course of the wind is from the northeast to the southwest, yet the course of the storm is from the southwest to the northeast." This observation by Franklins established him as the first to observe that storm winds can rotate counterclockwise, even though the entire storm is traveling in a clockwise direction.

In his attempt to decipher how such weather patterns could occur, Franklin theorized that heated air rising skyward in the south created a vacuum effect (low pressure), which in turn attracted winds from the north (high pressure). This theory of low and high pressure systems was the first known explanation of how storms moved in the Northern Hemisphere.

Today, weather forecasters recognize and respect the accuracy of Franklin's theory as it contributes to our understanding of the hydrologic cycle and forecasting the movement of hurricanes. "With this began the science of weather prediction," said the renowned scholar, William Morris Davis, 150 years later. Perhaps this distinction qualifies Franklin as America's first meteorological scientist.

Franklin's penchant for weather phenomena surfaced again in 1755 when he had the occasion to see a dust devil, or whirlwind funnel, bouncing across the landscape, while he was on horseback in Maryland. Franklin was so excited that he chased the dust devil and rode beside it for a considerable time before it disappeared in the forest. He wrote of this experience at length, on August 25, to his longtime friend and correspondent, Peter Collinson.

This lifelong interest in weather surfaced again 6 years before his death when he published *Meteorological Imaginations and Conjectures*. In this publication, Franklin explores such weather phenomena as heat radiation, insulating clouds, the shifting direction of winds, and creation of fog. In his own fashion, Franklin correctly defines circulation patterns of the hydrologic cycle. As well, he is correct in his attempt to deduce the rare manifestation of hail and ice during the summer months. He correctly figures out that warm moist air reaches high into the atmosphere where it cools and freezes into ice. This heavy ice then falls to earth before it is melted by the summer sun.

Also, toward the twilight of his life, Franklin again made some brilliant deductions. Before he died in 1790, he researched the effects that volcanic eruptions had on weather patterns and cloud electrification. In doing so, he correctly hypothesized that the severe winter of 1783–84 was exacerbated by the summer 1783 volcanic eruption in Iceland. Franklin theorized that dust from volcanic eruptions was instrumental in reducing the amount of solar energy reaching the earth's surface.

SWIM FINS

Since boyhood, Franklin enjoyed the experience of swimming and had a reputation for his strength and agility at plying the waters. His innate ingenuity expressed itself early when, still a boy, Franklin invented swimming fins, which he secured to his hands to help him swim faster and with less fatigue. According to Franklin, he developed the swim fins as a way to reduce what he referred to as "a laborious and fatiguing operation."

No matter where he traveled, Franklin always found time to enjoy his favorite exercise of swimming. When in Boston, he swam in the Atlantic Ocean and the Charles River. And when in England, he swam the Thames River. Benjamin Franklin "was considered to be an expert swimmer and while in England attracted attention by swimming the two to three miles from Chelsea to Blackfriars" (Franklin Bache, 1936).

On his first trip to London in 1724, Franklin was known for entertaining observers with his "ornamental" maneuvers while splashing around in the Thames. Franklin enjoyed swimming so much that he once considered taking a full-time job as a swimming instructor while in England.

Out of respect for Franklin's lifelong interest in swimming, he was inducted posthumously into the International Swimming Hall of Fame.

On a closing note, it should be recognized that even though Benjamin Franklin was affectionately referred to and often addressed as Dr. Franklin, Benjamin Franklin never received a Ph.D. from any academic institution. However, this in no way detracts from the historic contributions Franklin made to the worlds of science, politics, music, invention, social studies, publishing, printing, meteorology, navigation, optometry, swimming, oceanography, and God knows what else.

In Franklin, we find a human being who had an unsatiable curiosity about the world and a strong belief that he could make a difference through his contributions.

BIBLIOGRAPHY

1. Carson, R.L. (1950). *The Sea Around Us*. Oxford University Press, London.
2. Franklin, B. (1793). *The Private Life of the Late Benjamin Franklin, LL.D.* (Originally written by Himself, and now translated from the French). London, Printed For J. Parsons, No. 21.
3. Scisco, L.D. (1913), The track of Ponce de Leon in 1513. *Bull. Am. Geogr. Soc.* **45**: 721–735.
4. Tucker, T. (2003). *Bolt of Fate: Benjamin Franklin and His Electric Kite Hoax*. Public Affairs Press, Springfield, IL.
5. *The Old Print Shop Portfolio*. (1952). Vol. XI, January, Number 5.
6. Brown, L.A. (1951). The River in the Ocean. In: *Essays Honoring Lawrence C. Wroth*. The Anthoensen Press, Portland, OR.
7. Clarence, S. B., Frederick, R. G., Curt, F. B., Goldschmidt, E.P., William, A. J., Rollo, G. S., Henry, R. W. et al. (1951). Bibliographers—including an introduction by Wilmarth Lewis. *Essays Honoring Lawrence C. Wroth*. The Anthoensen Press, Portland, OR.
8. Bache, F. (1936). Where is Franklin's First Chart of the Gulf Stream?, *Proceedings of the American Philosophical Society* **LXXVI**: 731–741.
9. Pillsbury, J. (1891). *The Gulf Stream: Methods of the Investigation and Results of the Research*. (Appendix No. 10—Report for 1890) United States Coast and Geodetic Survey (Washington 1891), pp. 461–620.
10. Cohen, I.B. (1972). Benjamin Franklin. In: *Dictionary of Scientific Biography*. C.C. Gillispie (Ed.). Charles Scribner's Sons, New York.
11. Translated in Lipowski, Z.J. (1984). Benjamin Franklin as a Psychotherapist: A Forerunner of brief Psychotherapy. *Perspectives in Biology and Medicine*, **27**: 361–366.
12. Tucker, T. (2003). *Bolt of Fate: Benjamin Franklin and His Electric Kite Hoax*. Public Affairs Press, Springfield, IL.
13. TIME Graphic by Gabel, E., and Dykman, J. (2003). Ben Franklin the Great Inventor. *Time Magazine* (July 7), pp. 60–61.
14. Golden, F. (2003). When Sparks Flew. *Time Magazine* (July 7), pp. 55, 58.
15. Hall, A.J. (1975). Philosopher of Dissent Benj. Franklin. *National Geographic* **148**(1): 93–122.
16. Marmer, H.A. (1929). The Gulf Stream and its Problems. Disbound copy published in 1929 Annual Report by Smithsonian Institution, Washington, DC.
17. Available: http://www.colonialhall.com/franklin.asp
18. Available: http://www.earlyamerica.com/lives/franklin
19. Available: http://www.educationalsynthesis.org/famamer/Franklin.html
20. Available: http://www.elywelyn.net/docs/quotes/franklin.html
21. Available: http://www.english.udel.edu/lemay/franklin
22. Available: http://www.famouspeople.com/famamer/Franklin.html
23. Available: http://www.geocities.com/peterroberts.geo/BFrWr.html
24. Available: http://www.ideafinder.com/history/inventors/franklin.htm
25. Available: http://www.inventors/about.com/library/inventors/blfranklin.htm
26. Available: http://www.pbs.org/franklin
27. Available: http://www.quotationspage.com/quotes/Benjamin_Franklin

WATER SYMBOLISM

D.L. Marrin
Hanalei, Hawaii

Since the beginning of recorded history, water has held a special place in the spiritual and physical lives of people who represent a wide range of cultures and civilizations. Though rituals, languages, and beliefs among ancient cultures differed greatly, water was routinely used to symbolize a myriad of places and events in both the observable and unobservable realms.

MESOPOTAMIA

The first generally recognized written history comes from the Sumerians, who inhabited the Mesopotamian region more than 5000 years ago. In the Sumerian pantheon, various gods personified heaven, earth, air, and water. *Enki*, the primary water god, represented wisdom and evolved from a great underworld ocean that also gave rise to heaven (1). The underworld was sometimes referred to as a river and apparently existed beneath the surface of the earth as a freshwater (not saline) body of water. Enki (later known as *Ea* by the Babylonians) ruled over this watery underworld and symbolized the primordial waters, which were associated with the "chaos" that existed before creation. Myths from around the world recognize that creation was preceded by an original state of chaos, which defined a state of formlessness and was often identified as a watery abyss or primordial sea (2). Such a sea apparently reflected the fact that ancient people recognized creation as the emergence of form from formlessness.

EGYPT

The Egyptians, similar to the Sumerians and Babylonians, believed that the heavens and the earth (including all the inhabitants of the earth) were created from the primordial or celestial waters. *Ra*, the great Egyptian sun god, was intimately involved in creating the manifested world as his Sun disk appeared above the waters to signal the beginning of time (3). Ra made his symbolic journey over the water in a boat because the Sun was made of fire and, therefore, could not have risen out of the waters of chaos on its own. According to the Egyptian pantheon, everything is created from the waters of chaos by a being that, similar to God in the Bible's Book of Genesis, appears above or within the primordial waters. The primordial waters or seas often symbolized the formless chaos and undifferentiated matter of the underworld waters that comprised the cosmos before its division into the realms of heaven and earth. Hence, heaven, earth, and all life emerged from "the waters."

CHINA

The Chinese culture was one of only a few that did not associate the primordial chaos with water in the form of a sea. Instead, the ancient Chinese apparently portrayed chaos as a misty vapor that embodied the cosmic energy governing space, time, and matter (4). The formless misty vapor, from which the duality of *yin* and *yang* were born, gave rise to all earthly life forms. Some historians believe that the ancient Chinese people divided even the spirit world into a Ministry of Waters, whereby seawater is overseen by four dragon-kings representing the cardinal directions and freshwater is ruled by four dragon-kings representing each of the major rivers of China.

OCEANIA

The ancient Maori culture of New Zealand referred to water as *wai*, which also means remembrance or the recollection of something that has been. In the Maori culture, water symbolizes the memory of all that has ever been and will be (5). Like so many ancient and indigenous cultures, the Maori consider that water's memory serves as the spark of life and is intimately associated with the process of creation. How water is able to retain or access memory is not a component of ancient or indigenous myths; however, the almost universal insight that such memory is constrained neither by space nor time is one that has stirred considerable controversy in the modern world.

GREECE

Born from the primordial chaos (water) was the Greek goddess *Gaia*, who represents the manifested world and the first pantheistic being to inhabit our planet after it was formed. Gaia (as the earth mother) and Uranus (as the sky) gave birth to *Oceanus*, who is generally considered the Greek personification of water. The god known as Oceanus is, of course, at the root of our English word "ocean," where most of the water on the earth is contained. The ancient Greeks believed that water was everywhere and also constituted everything, such that all things manifest in this world represented different transformations of water.

In addition to their pantheistic characters, the ancient Greeks were one of the first cultures to begin incorporating their intuitive or mythical views of water into a more philosophical and, from a modern perspective, comprehensible format. Though such early philosophies often echoed the message of ancient myths, they began to explain water's relationship to other aspects of the physical world. Losing little of its mystique as a symbol of the primordial chaos, water was associated with geometries, numbers, vibrations, and physical structures.

ELEMENTS AND ENERGIES

A sixth century B.C. philosopher, named Thales, hypothesized that water is the primary substance of all being, out of which everything is created and to which everything returns (6). He was perhaps the first to recognize water's unusual physical properties and its apparent universality, both as a solvent and as a so-called creational medium. During the fifth century B.C., a Hermetic philosopher, named Empedocles, expanded on Thales' notion by proposing that all matter in the universe is composed of differing combinations of four original substances and two moving forces. Empedocles referred to these four substances as the *elements* of fire, air, water, and earth, and he identified love and strife as the *moving forces*. The moving forces energized the combining or dissociating of elements, such that matter could not be created without love nor uncreated (destroyed) without its opposite moving force of strife.

In addition to the four elements, there is a fifth substance that, according to Greek philosopher Plato, was essential to creating the universe (7). This mysterious fifth substance was referred to as *aether*, a word coined by the Greeks to identify the substance filling all space beyond the earth. Similar to the four elements, aether is an integral component of most ancient beliefs, where it is generally associated with a life energy or vital force. The ancient Hindus also recognized the four elements and aether, or *akasha*, as fundamental components of the natural world.

THE PLATONIC SOLIDS

The four elements and aether were symbolized by three-dimensional geometries known as Platonic solids (see Fig. 1). These angular solids were believed to connect the etheric and worldly realms to the primordial waters of chaos (usually represented by a sphere). The solid that most often symbolizes water is the icosahedron, which represents the creative principle connecting the etheric and worldly realms. The reciprocal geometric relationship between the icosahedron (representing water) and the dodecahedron (representing aether) symbolized the intimate relationship between the two. Because matter or the material world (represented by a cube) was supposedly manifested from the aether via the water, it follows that the icosahedron and dodecahedron are mathematically related to the cube according to the golden ratio (as ϕ and $1/\phi$, respectively). The golden ratio ($\phi = 1.618\ldots$) was believed by some ancient cultures to represent the mathematical code underlying the creation of the manifested world.

BIBLICAL PERSPECTIVES

Water is frequently referenced in the Bible in a variety of contexts designed to communicate very different kinds of messages. Biblical scholars believe that the significance of water was heightened in a desert region (i.e., the Holy Land) where water was scarce and drought was a constant threat to life. Beyond water's more practical

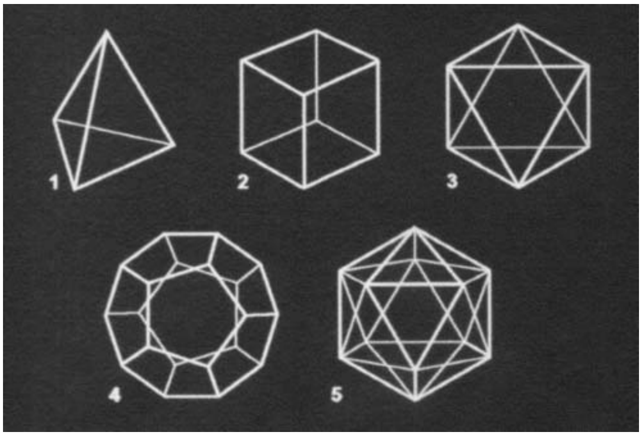

Figure 1. The five regular Platonic solids include (1) the tetrahedron symbolizing fire, (2) the cube symbolizing the earth, (3) the octahedron symbolizing air, (4) the dodecahedron symbolizing aether, and (5) the icosahedron symbolizing water.

considerations, the Bible portrays the spiritual aspects of water under several commonly recognized labels (8).

- *Cosmic water* describes a primordial force that only God can govern. The cosmic sea symbolizes the undifferentiated realm from which God created heaven and the earth.
- *Living water* symbolizes a source of life and a gift from God. Living water was considered distinct from ordinary water and symbolized both wisdom and knowledge.
- *Ceremonial water* symbolizes both the cleansing process and the passage between life and death. Ceremonial water is portrayed as an agent for removing or dissolving impurities.

If the biblical account of creation is consistent with that of other ancient accounts, heaven and earth are manifested from the symbolic, rather than literal, waters of which God is one. In this context, cosmic water is analogous to the waters of chaos rather than to literal water present in interstellar space (i.e., the modern scientific reference to cosmic water). Living water appears to be a designation for a special type of literal water, such that there is a recognizable difference between living and ordinary water. Living water is very rare, satisfies one's thirst, and also imparts life to the world. Most of the water used for religious ceremonies is ordinary water (sometimes symbolizing living water) that is blessed by an appropriate person.

ART AND NATURE

If one studies water in its natural state, it is difficult not to notice the variety of *flow forms* (e.g., whirlpools, eddies, ripples) that constantly appear and disappear. Many ancient cultures placed considerable import on these motions and believed that they held great powers for both creation and destruction. Renaissance artist Leonardo da Vinci is best known for his paintings and sculptures, but he was also a great student of water and its flow forms. He apparently believed that the vortical motion of fluids (particularly water) was a key to understanding and using the power of the universe. Two of the twentieth century's most renowned students of water flow forms were the European naturalists Theodor Schwenk and Viktor Schauberger, who expanded on ancient understandings and symbolisms of water.

According to Schauberger, water is the earth's blood and the essence of the *first substance*, whose boundaries and banks are the capillaries that guide it and in which it circulates (9). His view of water is similar to ancient perceptions inasmuch as water serves as the primordial substance of the universe and the fundamental component of both an organism's circulatory system and the planet's hydrologic cycle. Schwenk hypothesized that water's relationship to time was inherent in its rhythmical movement, which spanned cycles that ranged from seconds to years (10). He referred to different wave patterns and flow forms within a body of water as possessing distinct frequencies, overtones, and resonances—not unlike a musical instrument. Others have referred to this rhythm as "water's song," which often symbolizes the sound-shapes or music contained within the ripples and waves of water bodies.

POSTMODERN METAPHORS

The term "living water" is used today as a label or symbol for various aqueous concoctions in which ordinary water is treated with a combination of electrolytes, crystals, metals, colloids, and polymers, or exposed to various fields. The objective is to produce an optimal fluid that supports biological processes and, hence, retards both aging and disease. The exact mechanisms underlying a wide range of health claims have yet to be elucidated, but the modern practice of *water structuring* is based on the ancient understanding that water is able to retain or access a type of information that is colloquially referred to as "memory."

Perhaps the most recent water-related metaphor is that of *virtual water*, which was originally used to describe and, in some instances, quantify the influence of policy decisions on water resources in arid regions (11). Virtual water is a term used primarily in international trade for estimating the amount of water that is implicitly exchanged in goods and services (e.g., a ton of wheat requires 1000 tons of water) and the net gain for nations that import such goods and services. Although no actual water is traded during such transactions, the virtual water metaphor illustrates the extent to which real water is currently perceived as an economic commodity and is routinely transported across both geographic and political boundaries.

BIBLIOGRAPHY

1. Abraham, R. (1994). *Chaos, Gaia, Eros.* HarperCollins, San Francisco, CA, p. 126.
2. Andrews, T. (1998). *A Dictionary of Nature Myths.* Oxford University Press, New York, p. 36.

3. Mercatante, A. (1978). *Who's Who in Egyptian Mythology*. Clarkson N. Potter, New York, pp. 129–130.
4. McLeish, K. (1996). *Myth*. Facts on File, New York.
5. Brailsford, B. (1997). *Song of the Whale*. StonePrint Press, Hamilton, New Zealand, p. 174.
6. Lugt, H. (1983). *Vortex Flow in Nature and Technology*. John Wiley & Sons, New York, p. 6.
7. Lawlor, R. (1982). *Sacred Geometry*. Thames & Hudson, London, p. 96.
8. Ryken, L., Wilhoit, J., and Longman, T. (Eds.). (1998). *Dictionary of Biblical Imagery*. InterVarsity Press, Downers Grove, IL, pp. 929–932.
9. Coats, C. (1996). *Living Energies*. Gateway Books, Bath, UK, pp. 107–117.
10. Schwenk, T. (1965). *Sensitive Chaos*. Rudolf Steiner Press, London, pp. 18–20.
11. Wichelns, D. (2004). The policy relevance of virtual water can be enhanced by considering comparative advantages. *Agric. Water Manage.* **66**: 49–63.

GORDON AND FRANKLIN RIVERS AND THE TASMANIAN WILDERNESS WORLD HERITAGE AREA

Arthur M. Holst
Philadelphia Water Department
Philadelphia, Pennsylvania

The Tasmanian Wilderness World Heritage Area (TWWHA) is one of the largest temperate wildernesses in the world. Covering more than 8500 square miles, or more than 20% of the island of Tasmania, it contains three state reserves, four national parks, four conservation areas, two protected archaeological sites, and three forest reserves. Mountain ranges, rivers, Australia's deepest caves, waterfalls, and many endangered species are also present, as well as three lakes: Lake Gordon, Lake Peddler, and Lake St. Clair. The TWWHA was adjudged worthy of environmental protection in 1982 by the World Heritage Committee.

There are many reasons that the Tasmanian Wilderness was chosen as a World Heritage Area. It contains the largest peatlands in the Southern Hemisphere. It has great geological significance; it contains rocks from all but one geological period, and its oldest rocks are from the Precambrian period, more than a billion years ago. Also, sandstones and siltstones contain fossils of early marine life, fossilized when Tasmania was still underwater (Fig. 1). Much of the Wilderness' mountain landscape is the result of three glacial periods that occurred during the Pleistocene Epoch.

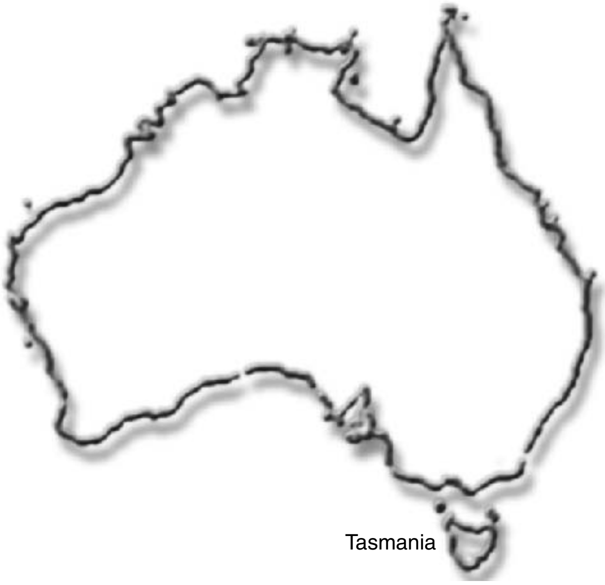

Figure 1. The island of Tasmania.

One of TWWHA's attractions is the Cradle Mountain–Lake St. Clair National Park. Lake St. Clair, the source of the Derwent River, is Australia's deepest lake and is enclosed by mountains. The Park contains myrtle, pine, and beech trees. Another attraction is the Central Plateau and Walls of Jerusalem National Park. This plateau was formed by glaciers during the last Ice Age; the area's lakes and tarns were created when the Ice Age's glaciers melted.

Many people visit the Tasmanian Wilderness World Heritage Area for its various activities. One of the most popular is known as bushwalking. This is basically hiking. There are advanced bushwalkers, and there are amateur bushwalkers, who take walks ranging from 15 minutes to an entire day. Advanced bushwalkers make longer and more arduous journeys over paths known as tracks. Tasmania's best-known track is called the Overland Track. About 80 km long, this track takes bushwalkers from 5–10 days to traverse. It features an alpine backdrop and passes the highest peak in Tasmania, Mount Ossa. This track is not exceptionally difficult; it is rated T1 by the Walking Track Management Strategy, the lowest rating, based on track hardening and track facilities.

Another favorite track of experienced bushwalkers is Frenchman's Cap. This track leads to the peak of the white quartzite Frenchman's Cap, the most prominent mountain in Franklin–Gordon Wild Rivers National Park. This track is considerably more difficult than the Overland Track; it has been assigned a T2 rating by the Walking Track Management Strategy. It usually takes bushwalkers 4 or 5 days to traverse this track.

A third popular track is the South Coast Track. It also has a T2 rating. This track has magnificent coastline views and spectacular beaches. It is located in Southwest National Park, and the main reason it is T2 is that the track is rather remote and far from any possible assistance. It is more than 80 km long and takes from 6–10 days to complete. It is often combined with the Port Davey Track to form a T3-grade track that takes from 10–16 days.

Another popular TWWHA activity is camping. Huts are provided along many common paths for general use, although tents are often used when the huts are full or where huts are unavailable.

The second most popular recreational activity in the TWWHA, after walking, is fishing. Local lakes, of which there are many, are often stocked with trout. These lakes, as well as rivers, are also sometimes used for boating, rafting, and kayaking. Some people opt to climb the peak of Frenchman's Cap. And the long, deep caves can be explored by experienced spelunkers. Marakoopa Cave is open to the general public.

Franklin–Gordon Rivers Wild Park (FGRWP), instated as a national park on May 13, 1981, is located in the southwest corner of Tasmania, between Southwest National Park, Cradle Mountain–Lake St. Clair National Park, and Walls of Jerusalem National Park. Not only the Franklin and Gordon Rivers, but also Jane River and Denison River are in FGRWP.

The Tasmanian World Wilderness Heritage Area is run by the Convention for the Protection of the World's Cultural and Natural Heritage, or World Heritage Convention. It was formed by UNESCO, the United Nations Educational, Scientific, and Cultural Organization, and it is the most extensive environmental convention in the world today. The area of the TWWHA is also world-renowned for its animal population. An unusually large number of species are endemic to Tasmania due to its remoteness from the rest of the world. Among these are the duck-billed platypus, the orange-bellied parrot, and the Tasmanian devil.

Along with all of these natural factors, the TWWHA possesses cultural significance as well. Thousands of years ago, Tasmania was detached from mainland Australia by Bass Strait, and the local Aborigines remained unaffected by outside influences until "discovered" by Abel Tasman in 1642. Many cave paintings have been found in Southwest Tasmania, including probably the southernmost cave paintings in the world.

The TWWHA's pristine quality is threatened by many factors. A major one is the soil erosion caused by bushwalkers hiking through the area. Many regulations are in place to attempt to lessen its impact on the environment. Another factor is cave degradation. Tasmania's many caves can be damaged easily by excessive visitors, so guides are available in many instances. Also, nonnative species have been introduced to the area, and they can have negative effects on local ecology. Fire can also devastate the environment; much of the TWWHA is designated fuel stove use only; the illegal use of campfires carries heavy fines.

There are many World Heritage Areas, but few can compare to the TWWHA in natural and cultural significance. It is one of only 13 World Heritage properties to satisfy all natural criteria. To this day, only one other World Heritage site has fulfilled as many criteria as the TWWHA.

READING LIST

Available: http://www.tasmaniatravel.com.au/misctext/hilite4.html.

Available: http://svc004.bne151p.server-web.com/spotlight/isncribed/taswilder.htm.

Available: http://www.dpiwe.tas.gov.au/inter.nsf/WebPages/UTAR-52SAD6?open.

Available: http://www.dpiwe.tas.gov.au/inter.nsf/Webpages/UTAR-52V8EK?open.

Available: http://www.dpiwe.tas.gov.au/inter.nsf/WebPages/UTAR-52W7QQ?open.

THE MISTAKE OF WAITING TO GET THIRSTY

F. Batmanghelidj
Global Health Solutions
Falls Church, Virginia

At the start of the twenty-first century, medicine is still muddled about the way the human body deals with a water shortage in its complex water-dependent inner environment. It is not yet clear even to physiologists of long standing in the field how local or general water shortages in the body manifest themselves. The following example is particularly dangerous: A respected scientist made an announcement that was quickly and widely publicized—unacceptable to me because its adoption will undoubtedly hurt people. As a means of airing the reasons why we need to prevent thirst rather that wait to get thirsty, I am responding here.

Heinz Valtin, M.D., an emeritus professor at Dartmouth Medical School, has ventured the opinion that there is no scientific merit in drinking eight 8-ounce glasses of water a day and not waiting to get thirsty before correcting dehydration. This view—published in the *American Journal of Physiology*, August 2002, and reiterated in the 2003 report of the Food and Nutrition Panel of the National Academy of Sciences—is the very foundation of all that is wrong with modern medicine, which is costing this nation $1.7 trillion a year (a figure rising at the rate of 12% annually). Dr. Valtin's view is as absurd as waiting for the final stages of a killer infection before giving a patient the appropriate antibiotics. His views are based on the erroneous inherited assumption that a dry mouth is an accurate sign of dehydration.

Dr. Valtin does not seem to be aware of an important paradigmatic shift in medicine. In the past, it was wrongly assumed that solutes in the body regulate all functions and the solvent has no direct role in any physiological functions. In medical schools, it is taught that water is only a solvent, a packing material and a means of transport and has no metabolic function of its own. I have come across this limited view at another Ivy League medical school from another eminent professor of physiology who, like Dr. Valtin, researched and taught the water-regulatory mechanisms of the kidney to medical students and doctors. Only when I asked him what "hydrolysis" is did the penny drop: He admitted that water is a nutrient and does indeed possess a dominant metabolic role in all physiological functions of the body.

Dr. Valtin's emphasis on the water-regulatory role of the kidneys limits his knowledge to the body's mechanisms of water deficit management. He seems to base his views of thirst management on the vital roles of vasopressin, the antidiuretic hormone, and the renin-angiotensin system, the elements that are engaged in the drought-management programs of the body when the body has already become dehydrated. He defines *dehydration* as a state of the body when it loses 5% of its water content. Yet he believes that we should wait until, at some such level of water loss, the urge to drink will correct the water deficit in the body. This view might have seemed plausible 25 years ago. Today, it exposes the tragic limitations of knowledge of human physiology available at a prestigious medical school in America.

Dr. Valtin does not take into consideration the fact that water is a nutrient. Its vital hydrolytic role would be lost to all the physiological functions affected by its shortage in its osmotically "free state." Another oversight is the fact that it is the interior of the cells of the body that would become drastically dehydrated. In dehydration, 66% of the water loss is from the interior of the cells, 26% of the loss is from extracellular fluid volume, and only 8% is borne by the blood tissue in the vascular system, which constricts within its network of capillaries and maintains the integrity of the circulation system.

Philippa M. Wiggin has shown that the mechanism that controls or brings about the effective function of cation pumps uses the energy-transforming property of water, the solvent: "The source of energy for cation transport or ATP synthesis lies in increases in chemical potentials with increasing hydration of small cations and polyphosphate anions in the highly structured interfacial aqueous phase of the two phosphorylated intermediates" (1). Waiting to get thirsty, when body fluids become concentrated, we lose the energy-generating properties of water in the dehydrated cells of the body. This is a major reason why we should prevent dehydration, rather than wait to correct it. This new understanding of the role of water in cation exchange is enough justification to let the body engage in prudent surplus water management rather than forcing it into drought and deficit water management, which is what Dr. Valtin is recommending.

In his research on "conformational change in biological macromolecules," Ephraim Katchalski-Katzir of the Weizmann Institute of Science has shown that the "proteins and enzymes of the body function more efficiently in solutions of lower viscosity"(2). Thus water loss from the interior of the cells would adversely affect their efficiency of function. This finding alone negates Dr. Valtin's view that we should let dehydration get established before drinking water. It is desirable that all cells of the body function efficiently within their physiological roles, so it would be more prudent to hydrate the body optimally rather than wait for the drought-management programs of the body to induce thirst. Furthermore, it is much easier for the body to deal with a slight surplus of water than to suffer from its shortfall and have to ration and allocate water to vital organs at the expense of less vital body functions. The circulation of concentrated blood in the vascular system is truly an invitation to catastrophe.

The tragedy of waiting to get thirsty hits home when it is realized that the sharpness of thirst perception is gradually lost as we get older. Phillips and associates have shown that even after 24 hours of water deprivation, the elderly still do not recognize that they are thirsty: "The important finding is that despite their obvious physiologic need, the elderly subjects were not markedly thirsty"(3). Bruce and associates have shown that, between the ages of 20 and 70, the ratio of water inside the cells to the amount of water outside the cells changes drastically—from 1.1 to 0.8.(4) Undoubtedly, this marked change in the intracellular water balance would not take place if the osmotic push and pull of life could favor water diffusion through the cell membranes everywhere in the body at the rate of 10^{-3} centimeters per second. Only by relying on the reverse osmotic process of expanding the extracellular water content of the body to filter and inject "load-free" water into vital cells by the actions of vasopressin and the renin-angiotensin-aldosterone systems—when the body physiology is constantly forced to rely on its drought-management programs—could such a drastic change in the water balance of the body result.

Two other scientific discoveries are disregarded when Dr. Valtin recommends that people should wait until they get thirsty before they drink water. One, the initiation of the thirst mechanisms is not triggered by vasopressin and the renin-angiotensin systems—these systems are involved only in water conservation and forced hydration of the cells. Thirst is initiated when the Na^+-K^+-ATPase pump is inadequately hydrated. It is water that generates voltage gradient by adequately hydrating the pump proteins in the neurotransmission systems of the body(1). This is the reason that brain tissue is 85% water(5) and cannot endure the level of "thirst-inducing" dehydration that is considered safe in the article published by Dr. Valtin.

Two, the missing piece of the scientific puzzle in the water-regulatory mechanisms of the body—which has been exposed since 1987, and Dr. Valtin and his colleagues need to know about it—is the coupled activity of the neurotransmitter histamine to the efficiency of the cation exchange, its role in initiating the drought-management programs, and its role in catabolic processes when the body is becoming more and more dehydrated(5). Based on the primary water-regulatory functions of histamine and the active role of water in all physiological and metabolic functions of the body—as the hydrolytic initiator of all solute functions—the symptoms of thirst are those produced by excess histamine activity and its subordinate mechanisms, which get engaged in the drought-management programs of the body. They include asthma, allergies, and the major pains of the body, such as heartburn, colitis pain, rheumatoid joint pain, back pain, migraine headaches, fibromyalgic pains, and even anginal pain(5,6). And because vasopressin and the renin-angiotensin-aldosterone activity in the body are subordinate to the activation of histamine, their role in raising blood pressure is a part of the drought-management programs of the body(6). The forced delivery of water into vital cells demands a greater injection pressure to counteract the direction of osmotic pull of water from inside the cells of the body, when it is dehydrated.

From the new perspective of my clinical and scientific research into the molecular physiology of dehydration and the peer-reviewed introduction of a paradigmatic shift in medical science recognizing histamine as a neurotransmitter in charge of the water regulation of the body, I can safely say that the 60 million Americans with hypertension, the 110 million with chronic pain, the

15 million with diabetes, the 17 million with asthma, the 50 million with allergies, and more, all did exactly as Dr. Valtin recommends. They all waited to get thirsty. Had they realized that water is a natural antihistamine(5,7,8) and a more effective diuretic, these people could have been saved the agony of their health problems.

For more information about the medical breakthrough on chronic unintentional dehydration and the diseases it causes, read as many of the following titles as you can. They are the products of more than 24 years of focused research on the topic of deep dehydration inside the cells of the body: *Your Body's Many Cries for Water; ABC of Asthma, Allergies and Lupus; Water for Health, for Healing, for Life; How to Deal with Back Pain and Rheumatoid Joint Pain; Water Cures: Drugs Kill*; and my 2004 book, *Obesity, Cancer, and Depression: The Deadly Diseases of Dehydration.* For more information on dehydration, visit www.watercure.com and www.nafhim.org.

BIBLIOGRAPHY

1. Wiggins, P.M. (1982). A Mechanism of ATP-driven cation pumps. In: *Biophysics of Water*. F. Franks and S.F. Mathis (Eds.). John Wiley & Sons, New York, pp. 266–269.
2. Katchalski-Katzir, E. (1984). Conformational changes in biological macromolecules. *Biorheology* **21**: 57–74.
3. Phillips, P.A., Rolls, B.J., Ledingham, J.G.G., Forsling, M.L., Morton, J.J., Crowe, M.J., and Wollner, L. (1985). Reduced thirst after water deprivation in healthy elderly men. *N. Engl. J. Med.* **311**(12): 753–759.
4. Bruce, A., Anderson, M., Arvidsson, B., and Isacksson, B. (1980). Body composition, predictions of normal body potassium, body water and body fat in adults on the basis of body height, body weight and age. *Scand. J. Clin. Lab. Invest.* **40** 461–473.
5. Batmanghelidj, F.M.D. (1987). Pain: A need for paradigm change. *Anticancer Res.* **7** (5B): 971–990. Available: http://www.watercure.com.
6. Batmanghelidj, F.M.D. *Your Body's Many Cries for Water*. Global Health Solutions.
7. Batmanghelidj, F.M.D. (1989). Neurotransmitter histamine: An alternative view. *3rd Interscience World Conf. Inflammation Analgesics Immunomodulators*, Monte Carlo, p. 37. Abstract and full article Available: http://www.watercure.com.
8. Batmanghelidj, F.M.D. *ABC of Asthma, Allergies and Lupus*. Global Health Solutions.

WATER AND WELL-BEING

David Haley
Picacadilly
Ulverston, Cumbria, United
Kingdom

> "Connected in water
> Where the seas touch the sky
> Tears bring forth oceans." (1)

If well-being is defined as "... the maximisation of the pleasure available to us" (2) and is related to health, safety, welfare, comfort, security, happiness, goodness, and contentment, then water must play an important part in maintaining this desirable state. For without water, none of these qualities of life can be achieved. Water connects all life and water connects all aspects of well-being, including social, economic, environmental, health, and aesthetic well-being, as a basic human right.

Article 25, Universal Declaration of Human Rights (1948) states: "Everyone has the right to a standard of living adequate for the health and well-being of himself and his family" (3). Kofi Annan, United Nations Secretary-General confirmed that "(a)ccess to safe water is a fundamental human need and therefore, a basic human right.

Contaminated water jeopardizes both the physical and social health of all people. It is an affront to human dignity" (4). But Dr. Gro Harlem Brundtland, Director-General World Health Organisation recognized that the right to water and well-being is not available to all: "Yes we have entered the new millennium with one of our most fundamental conditions of human development unmet: universal access to water. Of the world's 6 billion people, at least 1.1 billion lack access to safe drinking water. The lives of these people who are among the poorest on our planet are often devastated by this deprivation, which impedes the enjoyment of health and other human rights such as the right to food and adequate housing" (5).

NEEDS FOR WELL-BEING

Apart from the supply of safe drinking water and irrigation for food production, water is also essential to food preparation and personal hygiene. Environmental hygiene through adequate sanitation, safe housing, and water-based cultural practices are just as important to psychological well-being as they are to physical health. And water for the survival of ecosystems is critical to supporting biodiversity and also to maintaining aesthetic beauty. This, again is confirmed by the World Health Organisation: "Human livelihoods are intimately connected to natural life on the planet, and that, in turn, depends on water for survival. Natural ecosystems are often found in water or are heavily dependent on it. Without water, soil would parch, forests would wither, and species would die out" (6).

However, uncertainty—fear of the unknown—undermines well-being, particularly when it is related to an essential life-sustaining resource. As seen through history to the present day, uncertain access to adequate, safe water, provokes hostility and even war between communities competing for this vital element of survival. Likewise, the threat of flood has great physical and psychological impacts, threatening people's ability to house themselves and attachment to place.

CULTURAL DIVERSITY

> "By means of water we give life to everything." (7)—Koran 21:30

Every culture has a particular and very important relationship with water. In Jewish, Christian, and Muslim religions, water plays a fundamental role in creation, sacred dwelling, blessings, affirmation of faith, and divine retribution. Vandana Shiva points out that "… the root of the Urdu word *abadi*, or human settlement, is *ab*, or water, reflecting the foundation of human settlements and civilization along water sources" (8).

From China, the Taoist yin/yang symbol depicts the dynamic balance and unity between opposite forces, representing our potential well-being with water and symbolizing the flow of life—water is by its very nature a dynamic element. However, too much or too little brings disaster, so our ability to adapt to change and the flow of water is a measure of our well-being. The ancient Chinese art of "geomancy," Feng Shui, translates into English as wind and water and pays particular attention to achieving well-being through the relationship with water. This important cultural concern influences the harmonious use of water within the home, office, restaurant, and other places of business because the water is believed to draw in money. Inappropriate use or placement of water could bring misfortune. Likewise, the beneficial design and positioning of buildings in relation to water is taken very seriously—supposedly, the tender for the construction of the Headquarters of the Hong Kong and Shanghi Bank was won by the architect, Sir Norman Foster, because his design incorporated Feng Shui principles of the building's relationship to the sea.

An old Chinese saying, "he who controls water, governs the empire," (9) refers to the emperor's ability to govern a contented population by controlling floods, rivers, and canals; a factor still understood, if contentiously applied in modern China, as plans for vast dam developments threaten the well-being of communities and ecosystems in return for hydroelectric power and cash crop irrigation. Like many societies, now, water is becoming valued as a commercial commodity, a finite resource that may be controlled and sold for private, industrial, agricultural, or leisure use. Reverence for water as a heavenly gift, spiritual source, or common resource seems to be diminishing.

In her book *Water Wars*, Vandana Shiva strongly maintains that "(w)ater is a common because it is the ecological basis of all life and because its sustainability and equitable allocation depend on cooperation among community members" (10). And she asserts that "(c)ommunity control avoids ecological breakdown and prevents social conflict. Over the centuries, indigenous water management systems have relied on ancient knowledge and evolved into complex systems that ensure the equitable distribution of water" (11). Water was, and still is recognized as a prime factor in well-being. This condition is beautifully expressed in the ancient Rig Veda hymn, *Water of Life:*

> Waters, you are the ones that bring us the life force.
> Help us to find nourishment,
> So that we may look upon great joy.
> Let us share in the most delicious sap that you have,
> As if you are loving mothers,
> Let us go straight to the house of the one,
> For whom you waters give us life and give us birth.
> For our well-being, let the goddesses be an aid to us,
> The waters be for us to drink.
> Let them cause well-being and health to flow over us.
> Mistresses of all the things that are chosen,
> Rulers over all peoples,
> The waters are the ones I beg for a cure.
> Waters—yield your cure as an armour for my body,
> So that I may see the sun for a long time.
> Waters—carry away all of this that has gone bad in me.
> Either that I have done in malicious deceit,
> Or whatever lie I have sworn to,
> I here sort the waters today.
> We have joined with their sap,
> Oh Agni, full of moisture,
> Come and flood me with splendour! (12)

THE FLOW OF LIFE—WATER AND WELL-BEING

Let us consider water and well-being as a flowing life system. "Water in one form or another circulates around us, all the time and everywhere" (13). Blockages, floods, and droughts may be thought of as disruptions to this flow. Evolution may be considered a process of morphing forms through structural diversity within an aquatic planetary system, rather than a linear progression. And climate change is one form of that evolutionary process. We may then see water as a dynamic major determinant in the structure, pattern, and process of evolution, life, and well-being. "The waters on the surface of the earth flow together wherever they can, as for example in the ocean, where all rivers come together" (14).

In his book, *Sensitive Chaos*: *The Creation of Flowing Forms in Water and Air*, Theodore Schwenk develops this idea:

> It is in the nature of water that it possesses those characteristics fundamental to all organisms for the maintenance of life; yet all is purely functional and without fixed form. These living qualities are nevertheless so clearly present and so evidently at work that the picture of the human organism is conjured up before us in three-fold nature—metabolic, nerve-sense and rhythmic organisations—revealing in harmony the basic characteristics of all living organisms.

And this notion is reinforced by architect philosopher, Christopher Alexander: "We came from the water; our bodies are largely water; and water plays a fundamental role in our psychology. We need constant access to water, all around us; and we cannot have it without reverence for water in all its forms." (15) Alexander continues: "Jung and the Jungian analysts take great bodies of water as representing the dreamer's unconscious. We even speculate, in light of the psychoanalytic evidence, that going into the water may bring a person closer to the unconscious processes in his life" (16).

And finally, marine biologist Commodore Jacques Cousteau expresses his love for water and the profound sense of well-being it can evokes:

> My body floated weightlessly through space, the water took possession of my skin, the clear outlines of marine creatures had something almost provocative, and economy of movement acquired moral significance. Gravity—I saw it in a flash—was

the original sin, committed by the first living beings who left the sea. Redemption could come only when we returned to the ocean as already the sea mammals have done. (17)

BIBLIOGRAPHY

1. Haley, D (2000). *Seeking an Unacceptable Profile: Species Nova [to look anew].* Performed at Between Nat. Int. Conf. Lancaster University, UK.
2. Lakoff, G. and Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books, New York.
3. Article 25, Universal Declaration of Human Rights. (1948). *The Right to Water*. World Health Organization, France, p. 7.
4. Annan, K. (2003). *The Right to Water*. World Health Organization, France, p. 6.
5. Brutland, G.H. (2003). *The Right to Water*. World Health Organization, France, p. 3.
6. *The Right To Water*. (2003). World Health Organization, France, p. 19.
7. Koran 21:30. (2003). *The Right to Water*. World Health Organization, France, p. 3.
8. Shiva, V. (2002). *Water Wars: Privatization, Pollution and Profit*. Pluto Press, London, p. 20.
9. Rossbach, S. (1995). Feng Shui: An ancient theory of village siting. In: *The Village as Solar Ecology*. (East Falmouth, MA.: The New Alchemy Institute, 1980), *Feng Shui: Ancient Chinese wisdom on arranging a harmonious living environment*. Rider, London. p. 8.
10. Ibid., Shiva, V. (2002). p. 24.
11. Ibid., Shiva, V. (2002). p. 127.
12. Ibid., Shiva, V. (2002). p. VII.
13. Pielou, E.C. (1998). *Fresh Water*. University of Chicago Press, Chicago, IL, p. 246.
14. Trans. Wilhelm, R. (Eng. Trans. Baynes, C.F.). (1980). *The I Ching or Book of Changes*. Routledge & Kegan Paul, London, p. 35.
15. Alexander, C., Ishikawa, S., Silverstine, M., et al. (1977). *A Pattern Language: Towns, Buildings, Construction*. Oxford University Press, New York.
16. Ibid., Alexander, C., Ishikawa, S., Silverstine, M., et al. (1977). p. 324.
17. Schwenk, T. (1996). *Preface—Sensitive Chaos: The Creation of Flowing Forms in Water and Air*. J.Y. Cousteau (Ed.). Rudolf Steiner Press, London, p. 7.

FREE FLOWING WATER: A SOURCE OF WISDOM

MAGGIE LEE
Santa Fe, New Mexico

"Indeed it is because life is wisdom and water wisdom's element, that there can be such a thing as the water of life" (1).

In early August, while visiting gardens in the Boulder area of the Rocky Mountains, I discovered El Dorado Canyon, 10 minutes south. A brochure on the area relates…the tilted layers carry groundwater from the Rockies, down and eastward to a depth of 8,000′, before it is pushed to the surface as an artesian spring. Nearby, this tenacious spring feeds a 20 × 80′ pool. Swimming in this cool, turquoise, silky water, I moved with noticeable ease and buoyancy…refreshed in beneficent living water.

This article's title and beginning sentence, so eloquently expressed by Theodor Schwenk, continues with his experience of living water. "Water constantly pouring out of a spring in the middle of a flower-filled mountain meadow, sparkling in the pure cold air, with light reflected from the shimmering snowy peaks around." Generous, effervescent, living water.

Clearly, the beauty in the looking and the beauty in the seen, connected.

Sensing a similar mystery, David Abrams writes (when our perceptive practices reveal)…"a reciprocal phenomenon organized as much by the surrounding worlds as by (the perceiver)" (2).

Recently I discovered a passage by Rudolf Steiner saying that springs are the eyes of the earth.

In his extraordinary, rich, and comprehensive lectures and writings of the early 1960s and into the 1970s, Theodor Schwenk (1910–1986) proposed a new 'water consciousness,' a transformative way of thinking about this universal element. He described and illustrated fluid circulation, movement, and rhythms of the earth's water cycle and resulting organic forms. His experiments demonstrated the rhythmical, metamorphic potential hidden within fluid processes. Included here are a few of his insights and discoveries.

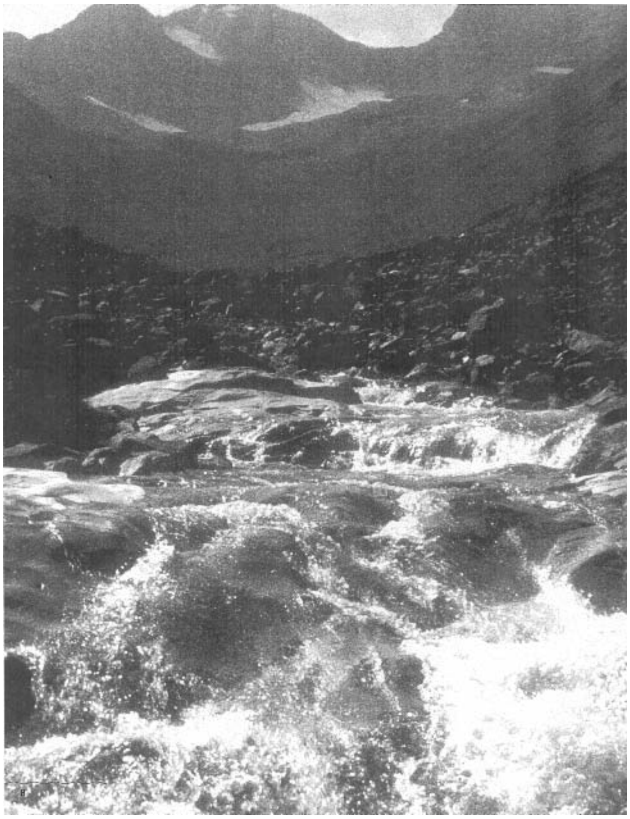

Figure 1. Waterscapes p. 8 edited by Herbert Dreiseitl, Dieter Grau, and Karl H.C. Luduig, 2001, Birkhauser, Switzerland.

In *Water, The Element of Life*, Schwenk develops a picture.

> Water as mediator between the centric (earthly) and the universal forces, setting them in balance and interweaving them with its very substance, reveals the activity of the heavens on earth. It is the port of entry through which cosmic-peripheral forces pass into the earth realm, directing the shaping, 'formative forces', organizing and patterning organic forms (Fig. 1).
>
> Renouncing any form of its own, it becomes the creative matrix for form in everything else.
>
> Contemplating this vast living organism, Earth, one's attention is drawn again and again to the layer-structure of its great enveloping mantles and to the rhythms that play in and through them.
>
> A glance at the surface configuration of the earth reminds us that 71% of it is covered by water. This watery surface in its immense extension provides a plane of contact with the atmosphere. . .where it works as the great regulator in matters of climate, meteorological processes and their rhythms.

Water's ability to regulate the climate is due to its thermal properties, in its tremendous capacity for storing warmth. . .evident too, in the origin of ocean currents and biological processes.

> Water's ability to absorb gases allows life on continents to come together with deep seas and thus these seas become great regulators of the respiratory processes of earth. In autumn, the plant world withers away and no longer inhales carbon dioxide and in the breaking down processes of plant substance, in the disintegrating foliage, much carbon dioxide is released into the air and is absorbed by the seas. When, however, in spring, the plant world grows anew and forms the substance for its foliage out of the carbon dioxide in the air, the waters of the earth once again release some of the necessary carbon dioxide. By absorbing carbon dioxide, it increases its ability to dissolve solid substance.
>
> Almost every property of water shows itself, in its relationship to warmth, as maximally suited to support the life of earth and its inhabitants. In this vein, water is unequaled, as L.J. Henderson has shown; the temperature anomaly, as a result of which water reaches its greatest density at +4 °C and grows lighter again at the freezing point, is the reason solid ice does not sink, rather floats, thus keeping earth from becoming a totally lifeless block of ice (3).

"Water is an element that brings a state of balance everywhere. Rhythm is its life element and the more it can be active rhythmically, the more it remains alive in its innermost nature" (4).

Linking similarities between water and thinking, Schwenk offers,

> The capacity of water in the realm of substances to dissolve and bind together, reappears in thinking as a spiritual activity.
>
> Like water, thought can create forms, can unite and relate the forms to one another as ideas. We speak of a capacity to think fluently when someone is skillfully able to carry out creation of form in thought, harmoniously coordinating the stream of thoughts and progressing from one idea to another.
>
> In thinking there prevails the etheric life of the water forces; through water flows the wisdom of the universe. Is it not this wisdom itself which has created the element of water, a tool for its own activity.
>
> Everything in Nature forms one indivisible fabric woven of living interchange. An all encompassing world of life comes into being from the interplay of cosmic-peripheral forces, meteorological forces, forces of the elements, the earth and all its living organisms (5).

Viktor Schauberger was a pioneering Austrian forester and brilliant inventor (1885–1958). His inventions were linked to water supply, natural river regulation, agriculture, propulsion and energy generation. They were all based on his observations of an underlying principle governing all natural processes, implosion.

> Implosion is an inward, curving (centripetal) suctional motion which follows a spiraling, whirling path as a vortex. Modern technology is primarily based on explosion, that is the outward moving, centrifugal motion of heating, burning, pushing or exploding. Inefficient, most of the energy is lost during an explosive process through frictional resistance, which produces useless waste heat. A characteristic feature of the life-generating vortex, is that the outside moves slowly, while the inside moves rapidly. As water is imploded in a vortex, suspended particles, which are denser than water, are sucked into the center of flow, so the frictional resistance is reduced and the speed of flow increased (6).

In 1961, Theodor Schwenk joined with George Adams, a mathematician, who for decades, with the assistance of Olive Whicher, had been researching projective geometry connected with organic forms, to found the Institute of Flow Sciences in the Black Forest in southern Germany. Just then his classic book on water, *Sensitive Chaos*, was published. Describing his experiences in seas worldwide, Jacques Cousteau wrote in the Preface,. . . "all around us there arose from the living sea a hymn to the 'sensitive chaos.' These memories have now taken for me a new meaning suggested by the book of Theodor Schwenk" (Fig. 2).

John Wilkes, a sculptor, had also joined the staff and was responsible for technical production of geometric models. In 1970, Wilkes, now at Emerson College in Sussex, England, began developing the Flowform Method and founded the Flow Design Research Group, currently based at the Virbela Rhythm Research Institute at Emerson College.

Concerned with enhancing water's quality, he posed his initial question, "can we create a form for water that will enable it to manifest its potential for metamorphosis and order?"

> This was inspired by observing controlled experiments revealing paths of vortices in contrast to the complexity of water movements in outer nature. Streaming water generates asymmetrical forms, but studying the results of living processes, it became clear that these forms tend strongly toward symmetry. Such forms have to do with thrust processes. Wilkes created simple symmetrical channels with mirrored meandering walls and changing proportions. He was looking at symmetry in this moment as of a higher order, with the idea of lifting water towards the living. The fortuitous event

Figure 2. Tentacles of the sea curl inland, eroding the Dutch coastline along the Waddenzee. In this view from 14,000′, a very low tide has revealed pale plains of mud and sand, giving the tidelands the look of a river delta. The darker veins are channels cut deep by the land-stealing sea; when the tide is high, they serve fisherman as navigation lanes (5).

> of proportion variation led to pulsing process in one particular place. This opened up for him a wide spectrum of investigation which led to the Flowform Method.
>
> Water's capacity to support life is actually related to its ability to move within itself. . .to generate membranes (surfaces) within its own volume. These form sensitive 'organs' through which water mediates between individual organisms and its total surroundings, carrying information. Movement, especially vortical movement, sensitizes water to this task (Fig. 3).
>
> The Flowform is a vessel, generally ovoid in shape, with a narrow entrance and exit, which create a resistance, inducing an oscillating figure of eight (lemniscate) movement. Flowing into alternate sides, whorls or vortices form, folding the water upon itself, oxygenating, enlivening the inner mobility of its many surfaces, the basis of its vitality, its freshness. Oxygenation fosters biological processes necessary to effect purification. The frequency of the rhythm varies with the size of Flowform, while the character of movement depends on the shape (7).

Special restorative value derives from water features designed in water's own terms. . .shaped surfaces caressed intimately by flowing, pulsing water.

Since 1973, in England and in collaboration with associates worldwide, well over 1000 Flowform installations have spanned more than 30 countries.

Also at that time, Arne Klingborg, director of the Rudolf Steiner Seminariet in Jarna, Sweden, invited Wilkes to test the beautifying and ecological possibilities of Flowforms in a lagoon purification system. Three Flowform cascades are operating through seven ponds with progressive wetland ecosystems from an algae pool to reed bed. Wilkes tells us, "not only is the coliform bacterial concentration in the treated water significantly below official levels permissible for public swimming, but the entire site serves as a community park and bird sanctuary" (Fig. 4).

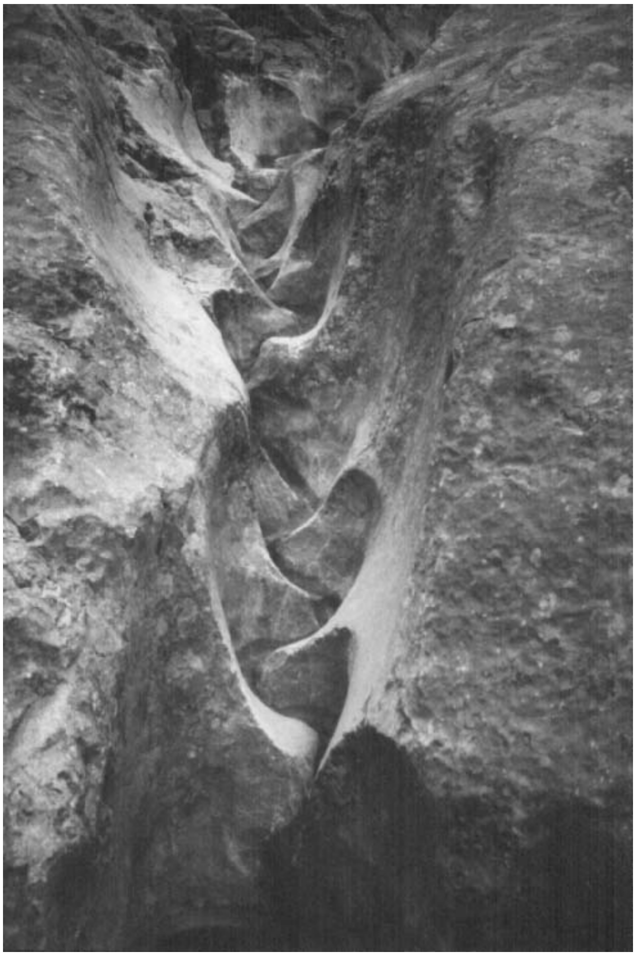

Figure 3. Such meander channels can be found below glaciers where streams have caused erosion over thousands of years. Photo supplied by John Wilkes.

Other applications of Flowforms include water purification, desalinated water treatment, stored-water treatment, swimming pools; air-conditioning and humidity; agricultural research centers, farms, homes, public gardens, parks and offices; aquaculture systems such as fish ladders for seasonally migrating species; pharmaceutics; and therapeutic and medicinal application.

In 1997, Flowforms America was founded. Its CEO, Sven Schunemann, worked with John Wilkes for a short period before this. Sven relates. . ."When water's movement is optimized for circularity, the liquid itself is able to be more effective as a natural element. Naturally flowing water is better aerated as more surface area is exposed."

John Wilkes added,

Figure 4. Malmö Flowform installed by a mountain stream at Sundet, Mösvatn, Norway. Water from the stream higher up is directed through the vessels which generate an ordered rhythmic swinging movement, like a living pulse, in contrast to the multitude of rhythmic movements in the streambed. If mounted below a hydroelectric turbine at a much larger scale, cast in situ, these more life-related rhythms, would improve the vital qualities of the water for irrigation, in a shorter distance! (11).

> It is necessary for Flowform creators to learn, in working with water's circular movements, to find the very special proportions which will finally generate pulsing processes.
> "Naturally flowing water is also more bioavailable…by allowing the nutrients it carries to be accessed more easily by plants. A plant, in turn, plays its part by taking nutrients out of such circulating water; thus fulfilling two roles, that of cleansing the water and of feeding itself" (8).

Wolfram Schwenk, a hydrobiologist and leading member of the Institute of Flow Sciences, writes about water and the creative process.

> If a creative artist wants to bring about cosmic and organic Flowforms in water, this cannot be done in the same way as a sculpture would be created. They can be evoked by handling water as an instrument. The creative process then takes place in the water itself. The artist learns how water twists and turns as it flows, eddies, trickles and spurts, surges, rests and reflects; enhancing and emphasizing by creative design, the theme develops as a joint work of art. Inspired, others may discover and experience special features of water. I know of no better, more attractive and sustainable ecological study course than this (9).

From his website, Herbert Dreiseitl informs, when referring to technology and infrastructure, of the lack of respect for the intrinsic aesthetics of water. He is actively participating in solving this dilemma.

During some months of 1978, he came to study the Flowform process with John Wilkes. In 1980, he established Atelier Dreiseitl in Uberlingen on Lake Constance in southern Germany. In his recently published book, *Waterscapes* (10), he writes, "to do justice to water, we have to go into the waterworld ourselves and experiment with it and learn to think in an integrated and interdisciplinary manner, about its flow and flexibility."

With regard to integrated planning, he states,

> water always creates a relationship between the detail and the whole. Each individual drop contributes to the balance of the earth's climate. Water projects become valuable when they help this process and can show that the place is being addressed and how it is connected with the world around it (10).

Further, he brings attention " to the special inner flexibility necessary to work with water successfully. The designs improve gradually and start to carry the signature of the water itself" (10).

Over the years, Atelier Dreiseitl has had opportunity through the scope of its commissions, which range from open-space planning in parks, water playgrounds, town squares to housing developments and urban hydrology, of linking water art to other complex themes, such as rainwater and surface water management, restoration and creation of waterbodies, incorporating Flowforms, storage and circulation cisterns, light and sound technologies, inlaid waterworks, and air-conditioning systems. Herbert relates, "demonstrating sustainable environmental technologies succeeds when everyone involved in the planning process, really does use interdisciplinary practices" (10) (Fig. 5).

Again, from his beautifully presented website, "water resources can only become sustainable when social values grant water the necessary priority. Water deserves everyones awareness" (10).

John Wilkes' new book, *Flowforms*, *The Rhythmic Power of Water*, is now available through Steiner Books (11).

Through our inventive, progressive and collaborative efforts, demonstrated and practiced with care, in our art and technology, may we allow water to be vigorous and free to the depths of its innermost nature.

> Choose to be changed. Oh be infused with the rapture of a fire that shows not what's changing in it as it burns for the

Figure 5. A water playground by the river—here everyone can join in. Water can be raised, played with and experienced! Water playground in Pforzheim, Germany (from 10).

directing spirit of the earth entire loves the figure of flight, not so much as the point where it turns

Sonnets to Orpheus II, 12 by Rilke
John Lash, translation after Walter Kaufman (12).
Maggie Lee
Santa Fe, New Mexico
January 28, 2004

BIBLIOGRAPHY

1. Schwenk, T. and Schwenk, W. (1989). *Water, The Element of Life*. Anthroposophic Press, New York.
2. Abrams, D. (1985). The perceptual implications of Gaia. *The Ecologist* **15**(3).
3. Henderson, L.J. (1914). *Die Umwelt des Lebens, Wiesbaden*. Cited in Water, The Element of Life, Schwenk.
4. Schwenk, Water, The Element of Life.
5. Schwenk, T. (1965). *Sensitive Chaos*. Rudolf Steiner Press. London, UK.
6. Coates, C. (1966). *Living Energies*. Gateway Books, UK Also see Alexandersson, O. (1996). *Living Water*. Gateway Books, UK, Bath.
7. Wilkes, J. Virbela Rhythm Research Institute, jwilkes 30@aol.com.
8. Schunemann, S. Available: http://www.flowformsamerica.com
9. Schwenk, W. (2001). Water as an open space. *Waterscapes*. H. Dreiseitl (Ed.). Birkhauser, Basel, Switzerland, pp. 106–107.
10. Dreiseitl, H. (2001). *Waterscapes*. Birkhauser, pp. 40–41.
11. Available: http://www.steinerbooks.org.
12. Website for John Lash is www.metahistory.org.
13. Baumgartner, W. (1989). *The Artesian Spring*. Energy Unlimited, Magdalena, NM.
14. Knight, D. and Stromber, J. (1998). *Nature's Miracles*. Centre for Implosion Research. Plymouth, UK www.ImplosionResearch.com. . ..as advised by Dr. Hal Williams, Anthroposophic Physician, October 2002.

Sidebars

1. Why does water rise to such great heights in mountain springs or artesian wells?
   Artesian springs come bubbling up vertically out of the ground. The source of the water comes welling up in pulses from the center of a column. . .this movement is similar to a telescopic extension.
   We can observe that in every stream, the water at the edges flow more slowly than in the middle, where the water is colder. . .faster layers flow past slower layers. . .building layer upon layer and creating extensive inner surfaces. These layers have different temperatures and densities. The inside layer is most dense, coolest and has the highest electrical potential, resulting in increased centripetal, spiraling motion. The outside layers warm up and expand in that warming up.
   Within a closed channel, i.e., underground, the outer layers exert pressure on the inside layers. . .tension builds from the temperature difference of the layers, disturbing the equilibrium inviting water to move! Just as a plum seed is shot forward from its fruit if you squeeze it. . .the denser body shoots forward as relief from pressure. Water always seeks the path of least resistance (adapted from Reference 13).
2. The vortex with its different speeds, is closely akin to the great movements of the planetary system. A given planet circles round the Sun as though in a vortex, in as much as it moves faster when near the Sun and slowly when further away. The vortex in its law of movement is thus a miniature image of the solar system and its planets.
   Also, if a very small floating object with a fixed pointer is allowed to circle in a vortex, the pointer always points in the direction in which it was originally placed, that is, it always remains parallel to itself!. . .it is always directed to the same point at infinity. This illustrates how a vortex is oriented, as though by invisible threads, with respect to the entire firmament of fixed stars (5, pp. 44–45).
3. Water as an information carrier.
   Various researchers such as Professor Benveniste, Dr. W. Ludwig (Treven and Talkenhammer, ed., Umweltmedizin Move-Verlag-Idstein, 1991), and Professor David Schweitzer have provided clear proof that water acts as a liquid tape recorder and is able to receive, store, and transmit electromagnetic vibrations. Because water molecules have positive and negative poles, they behave like little magnets. They attach themselves to their neighboring molecules and form clusters of several hundred molecules. These clusters are very sensitive structures and are impressionable by vibrational influences. This is what give water the ability to store information.
   This is closely linked to homeopathy, where a substance is diluted so many times, that eventually there is no molecule of the original substance left. . .yet it still has an effect. Homeopathy works because of the cluster ability to store vibrational imprints.
   Every substance and element has its own individual vibrational pattern, a bit like an energy blueprint. If you carry out the homeopathic process of diluting and succussing, the vibrational pattern of the remedy material becomes locked into the cluster structure of the carrier water. When you take this homeopathically prepared remedy, the cluster structure is transferred into the body and you react to the vibrational pattern of the original substance from which the remedy was prepared (14).

THE MEDICINAL PROPERTIES OF THE WATERS OF SARATOGA SPRINGS

A. Trent Millet
Newark, Vermont

My personal experience with American waters led me to investigate the widespread use of water for medical purposes in Europe, the United States, and other countries using Western medical systems similar or identical to our own.

It is important to note initially, however, that a number of Europe's great medical universities were established on or around waters that for centuries have been known to have medicinal properties. The European health systems and insurance companies have a long-standing practice of sending their clientele to spa for three weeks a year, both prescriptively as a treatment and as a preventive medium. The most commonly stated reason, throughout countries such as France, Poland, Germany, Czechoslovakia, the former USSR, and Hungary, is an expected ten to fifteen year increase in vital life expectancy: "vital" meaning not only the possible increase of number of years alive, but also, most importantly, the number of years individuals continue to be active and contributing members of society.

Czechoslovakia, Hungary, and Germany as of the late 1950s and early 1960s all had over 300 state-run spas. The former USSR had over 3500 (1). At that time in Europe alone, over 3,700,000 people per year were attending spas.

Ninety percent of patients sent to Bad Neuenahr are directed there by their own doctors (Ref. 1, p. 506) 50% of patients there suffering from diabetes. Health insurance companies, physicians representing those companies, social security agencies, unions, and corporations and municipal governments refer 50% of those sent to Bad Hall (Ref. 1, p. 507). "Thus, the state insurance fund of Hungary declared that its computed outlay for therapy and compensation in the rheumatoid group of diseases diminished 23% percent by the use of spa therapy" (Ref. 1, p. 507).

Of 2 million people sent to German spas by insurance companies, 25% suffered from rheumatism, 21% from diseases of the heart and blood vessels, 21% from respiratory diseases, 12% from digestive track disorders, and 8% from conditions causing exhaustion (Ref. 1, p. 507). Similar statistics continue in present-day Europe (A. Trent Miller, personal observation).

I began investigating the reported healing properties and medical use of American waters in 1998. A large number of American waters were used medicinally up until the 1960s. I concentrated my research around the Saratoga, New York area due to the well cataloged medicinal and historic use of Saratoga's many waters.

The anecdotal evidence, personal interviews, and personal experience lie outside the scope of this work. We are concerned here with patient statistics, prescriptive use, and the qualitative medical properties of Saratoga's spa waters.

People have been coming to use the waters for health since the 1700s. The 300-year use of Saratoga's waters for healing reached its pinnacle with the building of Sartoga's Roosevelt Spa by President Franklin D. Roosevelt in the 1930s. The famous Lincoln and Washington Baths had already been operating for almost half a century with 107,299 treatments given by the two combined baths in 1927 (2). By 1930, Lincoln Bath had the ability to give 4500 treatments a day.

Roosevelt Spa and the surrounding Saratoga State Park contained bathhouses, a drinking hall, inhalation rooms, and a bottling plant. Saratoga had always been known for its multitude of drinking halls, several bottling plants, and large number of bathhouses. Due to the fact that Roosevelt Spa was under the direction of three eminent doctors, Simon Baruch, MD, Walter S. McClellan, MD, and Oskar Baudisch, Ph.D., who directed the Research Lab, a great deal of the medical evidence, studies, and resulting statistics exist today. Hence my focus on Roosevelt Spa. By the 1930s only 18 of Saratoga's original 145 waters remained available for medicinal and public use. (About 127 springs were shut down because of the exhaustive pumping of carbon dioxide for use in the increasingly popular soda fountains, lowering the water level of the entire area.) The spa waters were commonly used in conjunction with other treatments, such as electrotherapy, ultraviolet and interred therapy, massage, and standard allopathic medicines. Because of this, statistics pertaining to the results of water treatments alone are difficult to find. Of the comprehensive number of waters, we will be looking at the treatment rationale of several. Table 1 defines the specific, subtle, and very individual ingredients contained in 17 of the 70 plus waters once available (3).

In consideration of this, the following is an example of how doctors used the waters prescriptively. "The choice of springs should be guided by the nature of the local manifestations, and the preponderance of constitutional plethora or anemia. For the uncomplicated dyscrasia, notably alkaline waters, with comparatively slight saline admixture—such as the Kissengen and Vichy—may answer well. With abdominal torpidity, a greater proportion of chloride of sodium is indicated, and then the Patterson, Hathorn, Congress or Excelsior may be recommended; or, if plethoric engorgement, is still more conspicuous, the powerfully saline Carlsbad, Champion or Geyser. In atonic cases, with well marked debility, often resulting from repeated attacks, or too prolonged depressing treatment, the saline-chalybeates may prove most beneficial."

Of the springs mentioned above, only Hathorn, Congress, and Geyser still flow. The spring called "Old Red," which is an iron water, has been used for all kinds of eruptive skin diseases and inflammation of the eyes. Dr. Oskar Baudisch, an authority on magnetochemistry, worked on the iron molecule. "In the course of his research he investigated the iron contained in Saratoga mineral water. He demonstrated that the iron contained in the water in spite of the relatively small amount in solution exceeded much larger amounts in therapeutic value as a remedy for anemia. This effect he showed to be due to the fact that the peculiar internal structure of the specific iron molecule contained in the water was similar to that of the iron molecule contained in the hemoglobin of blood. For this reason the iron in the water can be readily assimilated in the body and is, therefore, more readily available to the body than are other forms of iron" (Ref. 2, p. 37).

In the case of Old Red mentioned above, the water also contains two human cell salts (2). It is explained by Dr. Grace Swagger in a video interview contained in "Spa City, The History of Saratoga Spas," that the iron and cell salts were originally dissolved from the flesh of dinosaurs, animals, and plants. Kenneth Cohen, in his book *Honoring the Medicine* (4), backs this up with a statement that "the

Table 1. Analyses of the Waters of the Saratoga Spa Ions and Radicals as Presented in Solution

Element or Radical	Geyser, ppm	Polaris, ppm	Hayes, ppm	Orenda, ppm	Hathorn # 3, ppm	Lincoln, ppm	State Seal, ppm
Sodium	850	560	3025	2420	3,820	1,150	2.0
Potassium	83	80	333	266	340	219	0.15
Lithium	2.1	5	11.9	8	10	6.3	0.01
Calcium	375	370	724	672	872	348	32.2
Magnesium	75	95	277	224	353	171	4.4
Barium	12.5	2.1	12	15.6	25	8.1	0.038
Strontium	1.0	—	10.5	10	12	9.9	0.11
Ammonium	12.0	1.2	13	13	14	6	Trace
Iron	4.0	1.1	1.8	1.7	2.2	8.2	0.40
Aluminum	4.1	6	9.0	8	4.0	3	0.034
Manganese	Trace	—	0.3	0.3	0.4	None	0.002
Bicarbonate	2,562	2130	4,550	3,600	4,850	2,608	104
Chloride	820	1000	4,500	3,800	6,030	1,538	4.8
Bromide	16	5	51	46	53	34	None
Iodide	0.9	—	2.1	2	2.9	1.3	None
Metaborate	6.0	0.9	7.0	7.5	7.0	4.1	0.25
Silica	12	17	11	11	12	51	5.0
Sulfate	None	None	None	None	None	None	22
Total solids dried at 110 °C	4,836	3,260	13,539	11,105	16,407	6,166	175
Radium 226, PCi/L		102	284	232	430	48	

same water that flowed through the bodies of the dinosaurs has been recycled for millions of years to flow through our own bodies today."

Hathorn Spring, a saline–alkaline carbon dioxide charged water (which has a taste similar to the commercial product Alka Seltzer) has been used for stomach and digestive problems for over a century. Hathorn water was listed as one of the four most important medicinal waters bottled in Saratoga. Apparently, the very specific type of minerals and their very refined particle size and frequency allow absorption into the system regardless of the patient's age. Hence, having the proper mineralization to interact with vitamins allows continued enzyme production throughout a person's life. The common belief that people approaching a half century or more cease to make digestive enzymes is not the case when ingesting this type of water.

One of the waters considered the most important is the carbon dioxide water. The CO_2 water of the Lincoln group of wells at Saratoga is considered to have the highest content of carbon dioxide in the world. Its most common use is in the treatment of heart-related ailments. "The natural CO_2 is considered to have special merit in the treatment of the cardiac patient because of its unique ability to cause vasodilatation of the cutaneous blood vessels without any increase in the work of the heart, such as would occur in a warm bath causing an equal vasodilatation. On the contrary, the CO_2 bath reduces the work of the heart. Though administered at a neutral temperature, the CO_2 bath gives the sensation of warmth because of the layer of insulating gas bubbles next to the skin. It has been found that the carbonic acid dissolved in a natural CO_2 bath diffuses through the skin at a rate of about 30 mL per minute. The hyperemia is due to released histamine and acetylcholine. As a result, the pressure in the afferent capillary loop is increased about four times; the venous capillary pressure increases; the deep venous plexuses of the skin are emptied; and the increased vascularity of the skin drains the splanchnic pool. The diastolic pressure and the heart rate are lowered. In addition, the hydrostatic pressure exerted by the bath improves the venous return flow through compression of the lymph channels and veins (5).

Although Saratoga utilized the CO_2 waters for heart ailments, it was largely in conjunction with other medical treatments. There are no records detailing patient results from just the CO_2 water treatments alone. However, Tables 2–7 show the use of the water by disease classification and the patient results, illustrating the effectiveness in diseases other than the heart.

Table 3 shows the distribution of all the patients in this group on the basis of the usual classification of rheumatic diseases. The largest group was diagnosed as osteoarthritis in its various forms and included 438 of the patients treated. Rheumatoid arthritis was second on the list with a total of 331 patients, nearly all of whom were classified as chronic. The remainder of the patients represented smaller groups as indicated in the table. From the study of this table it is clear that the two large groups of patients with osteoarthritis and rheumatoid arthritis accounted for a large majority of the patients.

The results of treatment recorded by the hospital physicians on discharge have been tabulated for each of the groups of rheumatic disease in Table 4. They show that 262 patients or 26.3% were slightly improved, 568 patients or 57.2% were moderately improved, and 86 patients or 8.7% were markedly improved. This gives a total improvement of 92.2%. Only 77 patients or 7.8% were unimproved or were transferred for other treatment. No deaths occurred in the series.

It is obvious immediately that any statistical analysis dealing with the number of patients treated in groups 3 to 8 would have little value. Because of the larger number of

Table 2. Total Number of Patients by Disease Classification

Primary Condition	Percent	July 1932 to June 1942	July 1942 to March 1952	Total
Heart and circulatory disorders including variations of blood pressure	30.8%	38,500	47,000	85,500
Rheumatic conditions including arthritis, myositis, fibrositis, and neuritis	23.7%	29,625	36,166	65,791
Gastrointestinal ailments including liver and gallbladder	17.5%	22,000	26,858	48,858
Nervous conditions including both functional and organic disorders	8.4%	10,500	12,818	23,318
Metabolic diseases including diabetes, obesity, and glandular disorders	4.1%	5,125	6,257	11,382
Skin disorders (noninfectious)	2.1%	2,625	3,255	5,880
Miscellaneous	3.2%	4,000	4,883	8,883
No disease including general debility	10.1%	12,625	15,428	28,053
Total	*100%*	*125,000*	*152,600*	*277,600*

Compiled by Dr. McClellan.

Table 3. Results of Treatment of 331 Patients with Rheumatoid Arthritis at the Saratoga Springs Veteran's Hospital

Number of Baths	Sight Improvement	Percent	Moderate Improvement	Percent	Marked Improvement	Percent	Unimproved	Percent
1–10	8	44.6	3	16.6	0	1	7	38.6
11–21	22	27.9	53	63.4	3	3.8	4	4.9
22–42	40	23	117	67.3	14	8.1	3	1.6
Over 42	11	19.3	40	70.2	0	0	6	10.5
Totals[a]	*81*	*24.5*	*213*	*64.3*	*17*	*5.1*	*20*	*6.1*

[a] Total percent of improvement is 93.9%.

Table 4. Results of Treatment

Type of Arthritis	Total Number	Slight Improvement	Percent	Moderate Improvement	Percent	Marked Improvement	Percent	Unimproved	Percent
1. Osteoarthritis	438	101	23.4	260	59.3	38	8.6	39	8.7
2. Rheumatoid arthritis	331	81	24.5	213	64.3	17	5.1	20	6.1
3. Traumatic arthritis	94	38	40.4	31	33.0	19	20.2	6	6.4
4. Mixed (rheumatoid and osteoarthritis)	41	12	29.0	21	51.2	4	9.9	4	9.9
5. Rheumatoid spondylitis	36	10	27.8	2.0	55.6	4	11.1	2	5.5
6. Specific infectious (Gc, Tb, etc.)	31	10	32.3	17	54.8	1	3.2	3	9.7
7. Gout	15	7	46.7	4	26.6	3	20.0	1	6.7
8. Unknown	7	3	42.8	2	28.6	0	0	2	28.6
Total	*993*	*262*	*26.3*	*568*	*57.2*	*86*	*8.7*	*77*	*7.8*

Compiled by Dr. McClellen and Dr. Comstock for Lincoln Water (CO_2), Saratoga, New York, July 1932 to March 1952.

Table 5. Osteoarthritis Results for a Total of 438 Patients

Number of Baths	Slight Improvement	Percent	Moderate Improvement	Percent	Marked Improvement	Percent	Unimproved	Percent
1–10	4	12.5	16	50	4	12.5	8	25.0
11–21	35	27.2	65	52	12	10.2	12	10.2
22–42	48	20.6	149	63	19	8.1	17	7.6
Over 42	14	23.5	30	61	3	6.1	2	4.2
Total[a]	*101*	*23.4*	*260*	*59*	*38*	*8.6*	*39*	*8.7*

[a] Total percent of improvement is 93.1%.

patients with osteoarthritis and with rheumatoid arthritis, a more detailed study was made.

In another series of observations, McClellan and Comstock found that radon added to the bath was also absorbed through the skin, both from baths of plain water and, to a greater degree, from baths in the naturally carbonated mineral water. Both carbon dioxide and radon are important substances in many spas about the world. The demonstration that they are actually absorbed from the water of the bath supports the conclusion that the absorbed carbon dioxide or radon plays a role in the physiologic changes observed.

Table 5 presents the results of treatment for patients with osteoarthritis analyzed in relation to the total amount of treatment received. The total amount of treatment was gauged by the number of mineral water baths, because for practically all patients it represented the major feature of the program. Even with 10 baths or less, some patients showed improvement, but the proportion of those who were not improved was much higher than in the other groups. When the patients received 11–21 baths, there was some increase in the percentage of patients who obtained benefit. When the baths were continued longer, over 21 in number, there was considerable increase in those who showed moderate improvement. When the treatment program extended beyond what is usually considered two average courses of treatment, little additional benefit was evident.

It is interesting to note that the variation of improvement was essentially in keeping with the figures, noted in the larger groups of patients. With the variety of conditions in the remaining 20% of the patients, it was impossible to apply this detailed analysis because of the relatively small numbers.

Moving southwest 45 miles west of Albany is Sharon Springs, New York. Although now open only a few weeks per year, at one time the sulfur magnesium and calcium waters were waters utilized in over 82,000 treatments a year during the 12-week season. After WWII and up to the present day, many of the victims of the German concentration camps, predominantly Hasidic Jews, and currently the Russian victims of similar camps use the water of Sharon Springs for recovery and rejuvenation. Most are paid for by the medical services of their countries of origin. The sulfur water is particularly beneficial for rheumatoid and osteoarthritis, skin problems, and general revitalization of the body.

Table 6. Rheumatoid Spondylitis

Number of Baths	Number of Patients	Improvement			
		Slight	Moderate	Marked	None
1–10	2	1	1	0	0
11–21	13	2	9	1	1
22–42	12	4	6	2	0
Over 42	9	3	4	1	1
Total	*36*	*10*	*20*	*4*	*2*

Table 7. Gout

Number of Baths	Number of Patients	Improvement			
		Slight	Moderate	Marked	None
1–10	0	0	0	0	0
11–21	2	1	1	0	0
22–42	7	2	2	3	0
Over 42	1	1	0	0	0
Total	*10*	*4*	*3*	*3*	*0*
No bath	5	3	1	0	1
Total	*15*				

It is common to find calcium and magnesium waters in close proximity to the sulfur springs. The calcium–magnesium balance is one of the most important balances in the body, along with the sodium chloride and acid alkalinity (pH, also adjustable with the use of water). The type of magnesium and calcium found in the water is ideally suited to the human body. Several qualified sources (6) have stated that the use of natural calcium and magnesium water maintains that ratio at ideal levels.

In closing, although this entry is certainly not a complete and comprehensive discussion of all the medicinal uses and statistics, it is an example of water's inherent natural medicinal qualities.

BIBLIOGRAPHY

1. American Congress of Physical Medicine and Rehabilitation. (1960). *Archives of Physical Medicine and Rehabilitation*. Vol. 41, Nov., pp. 504–513.
2. Swanner, G.M. (1988). *Saratoga Queen of Spas*. North Country Books.
3. Irwin, J.A. (1892). *Hydrotherapy at Saratoga*. Cassell Publishing, New York.
4. Cohen, K. (2003). *Honoring the Medicine*. Ballentine Books, New York.
5. European spas, the medicinal properties of the water. (1960). *Arch. Phys. Med. Rehab.* **41**: 9510.
6. Lithium Report.

A HISTORY OF HAWAIIAN FRESHWATER RESOURCES

M. Eric Benbow
Mollie D. McIntosh
Michigan State University
East Lansing, Michigan

Albert J. Burky
University of Dayton
Dayton, Ohio

INTRODUCTION

The present economy of the Hawaiian Islands is primarily dependent on tourism. Visitors spend millions of dollars at hotels, in shopping malls and shops, at restaurants, and travel to and from islands. This money provides the economic foundation for many local Hawaiian families, and it serves to drive the state's economy. However, the economy has not always been based on tourism, but rather on the agricultural strength of sugar cane from the late nineteenth century through most of the twentieth century. Historically, the rise of agriculture was based on freshwater, the resource that drives tourism and expanding commercial development today. However, the surface waters and aquifers are limited, and a balanced concern exists for the effects of continued freshwater withdrawal from these sources. In particular, the streams of Hawaii are home to many endemic organisms (found only in Hawaii) that are threatened by continued surface water withdrawal. Many of the aquifers are being overpumped, increasing the salinity of groundwater resources and sometimes influencing surface flows.

For this topic, we provide a brief historical description of the geological origin and predominate climate patterns of the Hawaiian Islands, an introduction to endemism and the freshwater ecosystems, early Hawaiian cultural use and contemporary freshwater management, the expansion of sugar cane as a catalyst for freshwater exploitation, the cultural consequences of sugar cane expansion, and existing political struggles. We conclude with a short discussion about the future of Hawaiian freshwater resources, one of delicate balance and immediate concern.

GEOLOGICAL ORIGIN

The Hawaiian Islands have grown from a permanent "hot spot" located at an unknown depth beneath the Pacific Ocean seafloor. As it has been difficult to determine detailed information on this hot spot, the geological formation of Hawaii remains partly theory (1). The hot spot is a huge lake of melted rock under tremendous pressure in the Earth's mantle where molten rock erupts, accumulates on the ocean floor, and finally emerges from the ocean surface after millions of years (2). The hot spot is believed to remain stationary, but the outer crustal plate has been shifting (1). The shifting plates, timing of eruptions, and sea level changes have resulted in a chain of submerged and remnant volcanic islands to the northwest (Emperor Seamount Chain) of Midway Atoll, and the existing main Hawaiian Islands to the southeast (2).

The largest Hawaiian Islands are the most recently formed. The two oldest are Kauai and Niihau, with Oahu and the islands of Maui Nui (Maui, Molokai, Lanai, and Kahoolawe) sequentially younger moving southeast, and the appropriately named Big Island (or Hawaii) that is volcanically active (Fig. 1) (3,4). A new island about 1000 m (called Loihi) below the ocean surface is developing off the southern coast of the Big Island (5). As one travels from the northwest to the southeast, the existing islands become younger, larger, taller, and less eroded.

CLIMATE PATTERNS

Remotely located in the Pacific Ocean, the climate and precipitation patterns of Hawaii are dependent on predominant wind patterns and ocean currents. The Trade Winds (Prevailing Westerly Winds) blow across the Hawaiian Islands in a northeasterly to southwesterly direction, with stronger intensity occurring during the drier summer months (Fig. 1) (6). As these winds move across the ocean collecting water vapor, they move up, around, and sometimes over the mountains. As the air moves to higher altitudes, it cools, clouds form, and it rains, causing a rain shadow effect responsible for perennially wet northeastern shores (windward) opposite dry southwestern shores (leeward) (6). High rainfall, lush tropical rainforest vegetation, rushing streams, and abundant cascades and waterfalls characterize the windward sides, whereas the leeward sides have little rainfall and desert lands of dry-tolerant shrubs and cacti. The spatial distribution of precipitation combined with steep elevational gradients and ocean currents create climate patterns conducive for biome and ecosystem variety in a relatively small geographical area, ecosystems where many organisms evolved and are found nowhere else on Earth.

HAWAIIAN ECOSYSTEMS AND ENDEMISM

The variety of ecosystems and biomes in Hawaii area a result of five general characteristics: (a) geographic

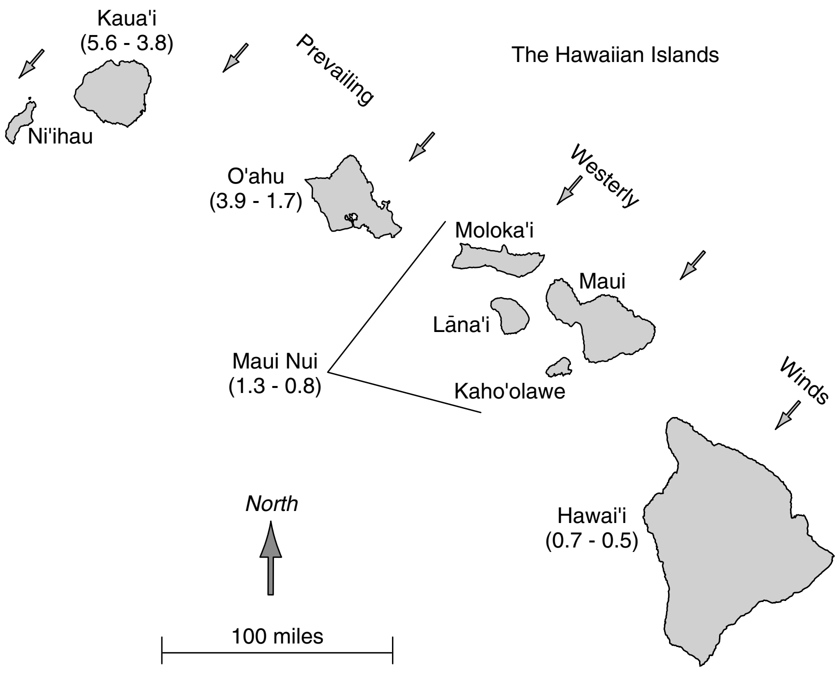

Figure 1. A map of the main Hawaiian Islands. Approximate island (or island group) ages are given in parentheses below the island name. Island ages are in millions of years and were estimated from two sources (3,4). Maui, Moloka'i, Lana'i, and Kaho'olawe were once a single island called Maui Nui until about 0.4–0.3 million years ago.

isolation, (b) small relative size and area, (c) recent age, (d) unique evolution, and (e) climate patterns (7). The Hawaiian Islands are some of the most isolated land masses in the world, which kept original immigrant species dispersal minimal (8). Relatively few species became established, but those that did underwent rapid speciation in the absence of many variables typical of mainland ecosystems (7). For example, some birds and many insects are flightless in response to evolving in predator absence (8). Many species never evolved effective defenses against predators or aggressive alien competitors, making the endemic flora and fauna easy prey, or out competed by introduced species that evolved with aggressive predators and competitors. Hawaii has approximately 10,000–15,000 endemic and indigenous species with over 30% of the total threatened or endangered species in the United States (4,7). In light of this, appropriate conservation, preservation, and management are needed to protect these rare habitats and endemic communities. The freshwater communities are threatened by continued and expanding riparian (land adjacent to a stream or river) habitat destruction, surface water removal, and introduced and invasive species.

HAWAIIAN STREAM ANIMAL COMMUNITIES

Hawaiian streams respond to unpredictable rainfall events that determine flow volume changes over short- and long-term time periods. The geology is commonly porous and precipitation quickly saturates riparian soils, creating rapid and high-volume surface runoff. Flash floods are common throughout the year. The local name for a flood is “freshet”, indicative of the historical description, and contemporary observations of scoured and rearranged stream channels moved by tremendous hydrodynamic forces. It is thought that many Hawaiian stream organisms have evolved specialized life cycles (e.g., amphidromy, see below) to accommodate these unpredictable and frequently extreme events (9–11).

AMPHIDROMY

The native Hawaiian fish, shrimp, prawn, and snail species have a diadromous life cycle, meaning they migrate between freshwater and seawater (1). These native species are amphidromous, a subcategory of diadromy (Fig. 2) (12,13). Adults reproduce in streams, eggs hatch, and embryos (now called larvae) drift to the ocean for growth and development into postlarvae (from 10–30 mm body length) over several months (14–16), and then migrate back into the streams. In amphidromy, the purpose of migration is not for immediate breeding as in diadromy (e.g., salmon); instead, postlarvae, not adults, migrate into freshwater to feed, grow, and remain for the rest of their existence through multiple reproductive events (12,17,18).

There are four endemic and one indigenous amphidromous fish species in Hawaiian streams (Table 1) (16,19,20). The pelvic fins of the gobies are fused and form a ventral sucker used to cling to rocks and climb waterfalls as postlarvae and early juveniles (16,21). One fish species has been reported to climb single waterfalls >100 m high and combined waterfalls totaling a height of 300 m (19). The native shrimp, snails, and prawns (Table 1) also have the ability to climb waterfalls (20). Postlarval migration has been found to be directly related to stream discharge (22,23).

OTHER ENDEMIC STREAM ANIMALS

The native stream organisms have closely related marine relatives, and it is evident that the freshwater species

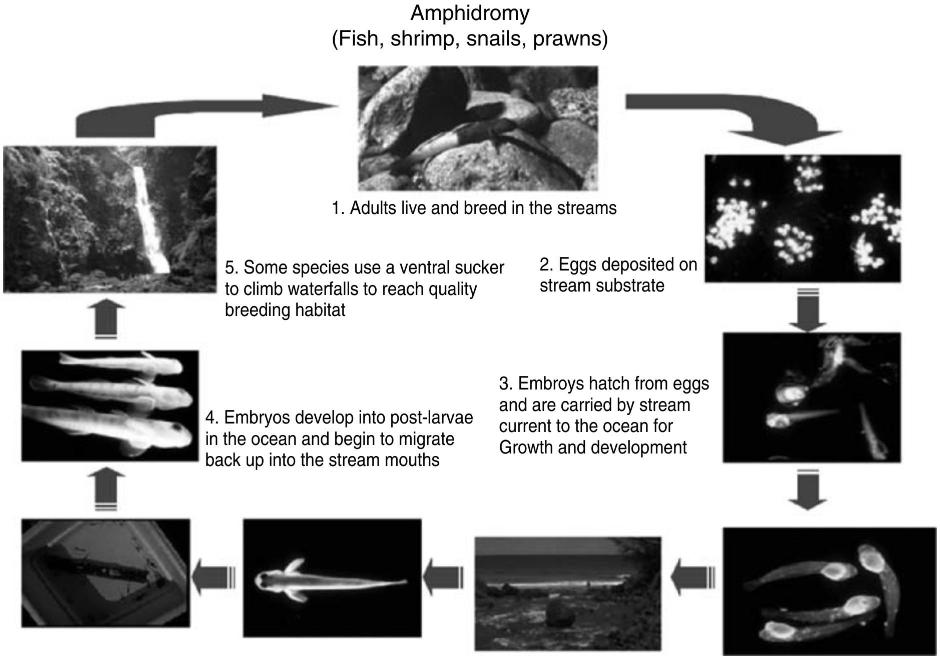

Figure 2. A schematic depicting the amphidromous life cycle of the native fish, snails, shrimp, and prawns found in Hawaiian streams. The life cycle requires postlarval development in the ocean and a migration back into the streams that is not immediately for breeding.

Table 1. The Freshwater Animals of Hawaiian Streams Including the Common Name, Scientific Name, Hawaiian Name, and Geographic Range

Taxa Type	Common Name	Scientific Name	Hawaiian Name	Range
Fish	general term for freshwater fishes		‘o’opu	
	freshwater amphidromous gobies	*Lentipes concolor*	‘o’opu ’alamo’o ‘o’opu hi’u kole ‘o’opu hi‘u’ula ‘o’opu nu’ukole	endemic
		Sicyopterus stimpsoni	‘o’opu nopili	endemic
		Awaous guamensis	‘o’opu nakea	indigenous
		Stenogobius hawaiiensis	‘o’opu naniha	endemic
	freshwater amphidromous eleotrid	*Eleotris sandwicensis*	‘o’opu ’akupa	endemic
	euryhaline-Hawaiian flag-tail fish	*Kuhlia sandvicensis*	aholehole	endemic
	euryhaline-striped or gray mullet	*Mugil cephalus*	‘ama’ama	worldwide
Snails	freshwater amphidromous snail	*Neritina granosa*	hihiwai wi	endemic
	estuarine amphidromous snail	*Neritina vespertina*	hapawai hapakai	endemic
	brackish/marine neritid snail	*Theodoxus cariosus*	pipiwai	endemic
	lymnaeid snails	*Erinna newcombi*	*	endemic
Sponge	freshwater sponge	*Heteromyenia baileyi*	*	indigenous
Crustaceans	general term for freshwater shrimp		’opae	
	freshwater, mountain srhimp	*Atyoida bisulcata*	’opaekala’ole ’opae kuahiwi ’opae kolo	endemic
	freshwater prawn	*Macrobrachium grandimanus*	’opae ’oeha’a	endemic
	anchialine pond shrimp	*Halocaridina rubra*	’opae ’ula	endemic
Insects	adult dragonfly	*Anas* spp.	pinao	endemic
	adult damselfly	*Megalagrion* spp.	pinao ’ula	endemic
	immature dragonfly or damselfly	*Megalagrion & Anax* spp.	lohelohe	endemic
	immature dragonfly or damselfly	*Megalagrion & Anax* spp.	lohaloha	endemic
	immature dragonfly or damselfly	*Megalagrion & Anax* spp.	pua’alohehole	endemic
	immature dragonfly or damselfly	*Megalagrion & Anax* spp.	’olopelope	endemic
	many other endemic insects	various species	*	endemic

*Indicates that a Hawaiian name is not available.

evolved from marine habitats (Table 1) (10,16,24,25). Some examples of this can be found in stream insect communities consisting of midges, *Telmatogeton* (Chironomidae); beach flies, *Procanace* (Canacidae); and shore flies, *Scatella* (Ephydridae), which are thought to have radiated into streams from ancestral intertidal habitats (26–28). The Hawaiian *Telmatogeton* complex is unique with all five endemic species restricted to freshwater, the only representatives of an exclusively marine intertidal subfamily worldwide (26–29).

Another important insect group is the endemic damselflies of the *Megalagrion* complex. This genus has shown rapid speciation in concert with specific freshwater and terrestrial habitats (30–32). Contrary to many other aquatic insects of Hawaii, the endemic damselflies are thought to have evolved from one freshwater ancestor to Hawaii (31,32). As a result of recent extinctions, numerous research initiatives have been carried out to understand the evolution, biology, and conservation of this unique group of insects (31,32). Other freshwater organisms are reported in Table 1.

SURFACE WATER REMOVAL

The streams of Hawaii are habitat to relatively recently evolved communities with close marine ancestors (e.g., fish, snails, shrimp, and some insects) and/or isolated populations that have undergone rapid speciation (e.g., the damselflies). For millennia, these systems connected the mountains to the ocean along what is called in Hawaiian the mauka—makai continuum. The physical connection was the flowing streams supporting the biological connections of the amphidromous species. However, many streams were heavily diverted in the late 1800s for irrigating an expanding sugar cane industry. Stream flow removal and riparian degradation has had the following effects: (a) destroying and eliminating breeding habitat for all stream animals, (b) preventing amphidromous egg and larval drift to the ocean, (c) obstructing postlarval recruitment back into the streams, and (d) facilitating invasive species establishment (9,22,33–35). Water is a critical and limited resource, not only to humans but also to the native stream organisms. Balancing these will take compromise and scientifically based freshwater resource management.

FRESHWATER RESOURCE USE AND MANAGEMENT

Historic Management

The ahupua’a system was the first freshwater management system in Hawaii. Developed by the ali’i (chiefs) of the first Hawaiians 1500–1600 years ago, it maintained adequate freshwater resources for the entire population of each island. The ahupua'a was what now is considered a watershed and extended from the headwater springs into the ocean. Along this mauka—makai continuum, the land along the stream was divided into pie-like slices called 'ili kuponos, that were designated to maka'ainana (commoners) and overseen by a chief-appointed konohiki. The commoners did not own the land or water, for they only maintained them for the chiefs. Each konohiki was appointed to a single ahupua'a, and each 'ili kupono was divided into kalo lo'i (taro patches). Stream water was diverted by pani wai (diversions) and carried through small 'auwai (ditches) to the lo'i. The lo'i were connected in a stair-step manner by additional ditches maintained by the responsible commoners. If the ditches were not maintained upstream, all downstream lo'i would be affected. Contrary to contemporary freshwater management, the stream water was always returned to the stream before it entered the ocean, and was never carried out of the watershed. Therefore, the water continually connected the highest mountain reaches to the lowland settlements near the sea, maintaining the mauka—makai continuum.

The early Hawaiians understood the connectivity of the land, freshwater, and sea. The ahupua'a system maintained adequate freshwater resources for human use while balancing ecological function. Hawaiians also understood the power and importance of water to their culture. For example, the Hawaiian word for water, wai, is the root for wealth (waiwai) and law (kanawai) (36). Over the next 60 years, native Hawaiians would find just how important water could be to foreigners, or haoles—those without breathe, or mana (spirit). The sugar cane industry completely changed freshwater resource management: water would be removed from the ahupua'a, along with much of the native Hawaiian culture.

Contemporary Management

The influence of western missionaries and merchants shifted the Hawaiian perception of water from a shared resource to a commodity that provided maximum use and greatest reward to plantation owners (37). After the arrival of Captain James Cook in 1778, freshwater management changed little until the The Great Mahele of 1848, enacted by King Kamehameha III under heavy influence by the decedents of original missionaries, the new owners of the sugar cane industry (and pineapple later). The Great Mahele was the first act to privatize land and water in Hawaii. For nearly 25 years the kings of Hawaii attempted to negotiate reciprocity treaties with the United States In 1876, King Kalakaua finally established a reciprocity treaty assuring duty-free exchange between Hawaii and the United States, giving sugar companies a competitive edge in the world market and giving Pearl Harbor to the United States, which was the same year of the first water licenses, the creation of the first private water company, and the construction of the diversion-ditch systems, which relocated millions of gallons of water from the windward watersheds to leeward sugar cane fields. According to Wilcox (36), one pound of sugar takes approximately 4000 pounds, or 500 gallons, of water.

The growth of sugar cane and pineapple took over and "transformed Hawaii from a traditional, insular, agrarian, and debt-ridden society into a multicultural, cosmopolitan, and prosperous one" (36). During this time, the number of sugar plantations rose from 5 in 1857 to 90 in 1884 (36). From the late 1870s to the mid-1900s, the Hawaiian sugar cane industry continued to grow, in part because of political access to surface and groundwater resources. By the early 1900s, almost every stream on the windward sides of each island was at least partially diverted, and in 1920, the sugar cane industry was diverting >800 mgd and pumping an additional 400 mgd from aquifers (36). Almost 20 years later, the entire city of Boston was only consuming 80 mdg (36). The privately developed surface water of Hawaii was quite different from publicly controlled surface water development of the western United States, which put control of water into the hands of western businessmen, those who would eventually contribute to the overthrow of the last Hawaiian Monarchy. (Several books are available that describe the sequence of events that led to the overthrow (38–40).)

With the developing agricultural industry, the population and ethnic diversity expanded as well. A larger labor force was needed to construct the ditches and diversions and work the agricultural fields. For cheap labor, the agricultural companies looked to Japan, China, and other Asian nations. This immigration resulted in several ethnic populations living and working on plantation property, creating mixed nonHawaiian communities that are the roots of the ethnic diversity found in Hawaii today. As the population and successful industries increased, the popularity of the islands grew in the mid-1900s, whereas the sugar cane industry met fierce competition from other nations, and has resulted in the slow demise of the former economic strength. With the buildup and maintained military presence of Pearl Harber from World War II, the economic focus of Hawaii saw a change from agriculture to military establishment to tourism that is now the economic base of the state. Along the southwestern coast of each island, numerous hotels, shopping malls, golf courses, and restaurants have been developed because most people prefer to vacation on the sunny, dry side of the island. However, similar to the sugar cane scenario, water is not readily available on the leeward side. Therefore, water remains diverted from streams, or pumped from dwindling freshwater aquifers, in order to accommodate tourists, water the green lawns, and fill hotel swimming pools and spas.

A BRIEF INTRODUCTION TO HAWAIIAN WATER LAW

Water Rights

The expansion of the sugar industry resulted in many water disputes between sugar companies, various landowners, and native Hawaiians. These disputes helped recognize the following water rights: (a) appurtenant, (b) riparian, (c) prescriptive, and (d) surplus (41). Appurtenant rights developed from ancient rights and were officially declared in the earliest Hawaiian water case, Peck vs. Baily (1867). This case determined that if land was entitled by a landowner to cultivate taro, than some stream water was also. Riparian rights permit water use to those who own land next to a stream; however, use must be reasonable and not harmful to other landowners (42). In riparian law, the focus is on maintaining stream integrity while allowing landowner priority uses (e.g., domestic drinking and washing) over nonpriority uses (e.g., irrigation, mining) (41). Both appurtenant and riparian water rights consider water as a common good. Prescriptive and surplus water rights address private ownership of water sources.

Prescriptive water rights are granted if a party has proven using water "belonging" to another party for an extended and continuous period of time (e.g., 20 years) (37). Surplus water is defined as the amount of water in a stream not covered by any prior water right (37). In court, cases of surplus water were heard independently, which precluded the establishment of defined standards for future disputes. For example, in the case of Hawaiian Commercial & Sugar Co. v. Wailuku Sugar Co. (1904), the court ruled that surplus water traveling through an ahupua'a belonged to the private land owner, but in Carter v. Territory (1917), the court concluded that surplus water was to be used reasonably according to riparian principles, thus shared among landowners. The inconsistent rulings and absence of an established in-stream flow standard left the status of surplus waters questionable until the state code in 1987.

Summary of Case Law

From the 1850s to the 1980s, the water rights of Hawaii were evolving according to case law, yet no standard criteria for the allocation, transport, and quantity of freshwater was resolved. There were several political, economical, and social reasons for no established criteria. First, the rise of the sugar cane industry boosted the economy, and with it came political power to the foreign owners. Second, during this period, the Hawaiian Islands shifted from a monarchy to a republic to a U.S. Territory, and finally became a U.S. state, which created an atmosphere of economic and political change that left many native Hawaiians in a state of social and political confusion (36), where they did not resist changes in water law, as it was not in their culture to protest actions of the King. As a result of this, there was difficulty in establishing a cause-effect relationship, and many of the villages most affected were too small to fight the wealthy plantation owners (36,37).

In 1978, the Hawaii State Constitution mandated that a water code and commission be established to manage water resources. And in the late 1980s, a Hawaii State Water Code and the Commission on Water Resource Management was passed by the Hawaii State Legislature (37). The State Water Code became a regulatory mechanism for state agencies and counties and required the Commission to monitor and guide water allocation by establishing minimum in-stream flow standards that were to enhance, protect, and re-establish beneficial in-stream uses. Current trends have been developing during the 1990s that are testing the utility of the Hawaii State Water Code. The first of several cases, and a landmark case for future petitions, was the Waiahole Combined Contested Case Hearing. In this case, the Hawaii Supreme Court ruled for the establishment of minimum in-stream flow standards and than any additional flow secondarily allocated for other uses. In-stream flow standards are currently being investigated. The Waiahole Ditch Case was the first test of the State Water Code, yet many similar cases are being petitioned at the time of this writing. The diminishing sugar industry is opening a new chapter in Hawaiian case law, a chapter that must not only consider economics but also "traditional and customary Hawaiian rights, protection and procreation of fish and wildlife, ecological balance, scenic beauty, public recreation, beneficial in-stream uses, and public interest" (36).

SUMMARY AND SYNTHESIS

The history of Hawaiian water resource issues is complex. The system of surface water management went through a dramatic change with the explosion of agriculture and the associated political power of missionary descendents; however, this explosion was inherently dependent on privately developed surface waters—a feedback proliferation of each other. The surface water development from the late 1800s through the mid-1900s drained watersheds of both water and native Hawaiian culture. The efficient removal of water high in the watersheds left a trickle downstream for native taro farmers. With the construction of massive diversion-ditch systems, native Hawaiians emigrated from the ahupua'a to build ditches and work sugar cane fields. At the same time, thousands of immigrants from China, Japan, and the Philippines melted into the sugar cane plantation farms, changing the demographics of Hawaii. The push by sugar businessmen to get an extended reciprocity treaty secured rights of Pearl Harbor to the United States; it also led to the coup of the Hawaiian Monarchy. At the same time, streams were being impacted by water removal and rapid riparian destruction, negatively affecting the native biological communities of the mauka–makai continuum.

The degradation of stream habitats in Hawaii has become an increasingly important issue to many native Hawaiians, scientists, private organizations, and individuals. Water resources have been, and will continue to be, a highly debated political issue. Water is necessary for agricultural, industrial, and municipal uses throughout Hawaii; however, in order to maintain the unique biodiversity of the streams, a compromise must be made between continued development and conservation. Thus, the future of the Hawaii's freshwater ecosystems and economic security is dependent on wise freshwater management and allocation, which should not be made in a scientific vacuum. Objective and sustainable answers can only be achieved with a solid scientific base, one free from political bias that has been the history of Hawaiian freshwater resources.

BIBLIOGRAPHY

1. Culliney, J.L. (1988). *Islands in a Far Sea: Nature and Man in Hawaii*. University of Hawaii Press, Honolulu, HI.
2. Hazlett, R.W. and Hyndman, D.W. (1996). *Roadside Geology of Hawaii*. Mountain Press, Missoula, MT.
3. Stearns, H.T. (1985). *Geology of the State of Hawaii*, 2nd Edn. Pacific books, Palo Alto, CA.
4. Juvik, S.P. and Juvik, J.O. (1998). *Atlas of Hawaii*, 3rd Edn. University of Hawaii Press, Honolulu, HI.
5. U.S.G.S. (1997). *U.S. Geological Survey Programs in Hawaii and the Pacific*. Department of Interior, United States Geological Survey.
6. Carlquist, S. (1980). *Hawaii: A Natural History*. Pacific Tropical Botanical Gardens, Honolulu, HI.
7. Mueller-Dombois, D., Bridges, K.W., and Carson, H.L. (1981). *Island Ecosystems: Biological Organization in Selected Hawaiian Communities*. Hutchinson Ross, Stroudsburg, PA.
8. Kay, E.A. (Ed.). (1994). *A Natural History of the Hawaiian Islands, Selected Readings II*. University of Hawaii Press, Honolulu, HI.
9. Way, C.M., Burky, A.J., Harding, J.M., Hau, S., and Puleloa, W.K.L.C. (1998). Reproductive biology of the endemic goby, *Lentipes concolor*, from Makamaka'ole Stream, Maui and Waikolu Stream, Moloka'i. *Environ. Biol. Fish.* **51**: 53–65.
10. Benbow, M.E. (1999). *Natural History and Bioenergetics of an Endemic Hawaiian Chironomid: Fluctuating Stability in a Stochastic Environment.* Ph.D. Dissertation. Department of Biology, University of Dayton, Dayton, OH.
11. Benbow, M.E., Burky, A.J., and Way, C.M. (1997). Larval habitat preference of the endemic Hawaiian midge, *Telmatogeton torrenticola* Terry (Telmatogetoninae). *Hydrobiologia* **346**: 129–136.
12. McDowall, R.M. (1996). Is there such a thing as amphidromy? *Micronesica* **30**(1): 3–14.
13. McDowall, R.M. (1997). The evolution of diadromy in fishes (revisited) and its place in phylogenetic analysis. *Rev. Fish Biol. Fish* **7**(4): 443–462.
14. Radtke, R.L. and Kinzie, R.A. (1996). Evidence of a marine larval stage in endemic Hawaiian stream gobies from isolated high-elevation locations. *T. Am. Fish Soc.* **125**: 613–621.
15. Radtke, R.L., Kinzie, R.A., and Folsom, S.D. (1988). Age at recruitment of Hawaiian freshwater gobies. *Environ. Biol. Fish* **23**(3): 205–213.
16. Benbow, M.E., Burky, A.J., and Way, C.M. (2004). Morphological characteristics and species separation of Hawaiian postlarval amphidromous fishes. *Micronesica* **37**(1): 127–143.
17. Lindstrom, D.P. (1998). *Reproduction, Early Development and Larval Transport Dynamics of Amphidromous Hawaiian Gobioids.* Ph.D. Dissertation, University of Hawaii, Honolulu, HI.
18. McDowall, R.M. (1992). Diadromy: origins and definitions of terminology. *Copeia* **1**: 248–251.
19. Kinzie, R.A., III (1990). *Species Profiles: Life Histories and Environmental Requirements of Coastal Vertebrates and Invertebrates Pacific Ocean Region. Amphidromous Macrofauna of Hawaiian Island Streams.* US Army Engineer Waterways Experiment Station, Vicksburg, MS.
20. Kinzie, R.A., III and Ford, J.I. (1982). *Population Biology in Small Hawaiian Streams*. Water Resources Research Center, University of Hawaii at Manoa, Honolulu, HI, p. 147.
21. Benbow, M.E., Orzetti, L.L., McIntosh, M.D., and Burky, A.J. (2002). A note on cascade climbing of migrating goby and shrimp postlarvae in two Maui streams. *Micronesica* **34**(2): 243–248.
22. Benbow, M.E., Burky, A.J. and Way, C.M. (2004). The use of two modified Breder traps to quantitatively study amphidromous upstream migration. *Hydrobiologia* **527**: 139–151.
23. Burky, A.J., Benbow, M.E., and Way, C.M. (2004). *Two modified Breder traps for quantitative studies of postlarval amphidromy, Hydrobiologia.*
24. Kinzie, R.A.I. (1997). Evolution and life history patterns in freshwater gobies. *Micronesica* **30**(1): 27–40.
25. McDowall, R.M. (2003). Hawaiian biogeography and the islands' freshwater fish fauna. *J. Biogeogr.* **30**(5): 703–710.
26. Hardy, D.E. and Delfinado, M.D. (1980). *Diptera: Cyclorrhapha III, Series Schizophora, Section Acalypterae, Exclusive of Family Drosophilidae.* Vol. 13. The University Press of Hawaii, Honolulu, HI.
27. Newman, L.J. (1988). Evolutionary relationships of the Hawaiian and North American *Telmatogeton* (Insecta; Diptera: Chironomidae). *Pacific. Sci.* **42**(1–2): 56–64.
28. Newman, L.J. (1977). Chromosomal evolution of the Hawaiian *Telmatogeton* (Chironomidae, Diptera). *Chromosoma* **64**: 349–369.
29. Wirth, W.W. (1947). A review of the genus *Telmatogeton* Schiner, with descriptions of three new Hawaiian species (Diptera: Tendipedidae). *Proceedings of the Hawaiian Entomological Society* **13**: 143–191.
30. Jordan, S., Simon, C., and Polhemus, D.A. (2003). Molecular systematics and adaptive radiation of Hawaii's endemic damselfly genus *Megalagrion* (Odonata: Coenagrionidae). *Syst. Biol.* **52**(1): 89–109.
31. Polhemus, D.A. (1997). Phylogenetic analysis of the Hawaiian damselfly genus *Megalagrion* (Odonata: Coenagrionidae): implications for biogeography, ecology, and conservation biology. *Pacific Science* **51**(4): 395–412.
32. Polhemus, D.A. and Asquith, A. (1996). *Hawaiian Damselflies: A Field Identification Guide*. Bishop Museum Press, Honolulu, HI.
33. Englund, R.A. and Filbert, R.B. (1999). Flow restoration and persistence of introduced species in Waikele Stream, O'ahu. *Micronesica* **32**(1): 143–154.
34. Benbow, M.E., Burky, A.J., and Way, C.M. (2003). Life cycle of a torrenticolous Hawaiian chironomid (*Telmatogeton torrenticola*): stream flow and microhabitat effects. *Ann. Limnol. Int. J. Lim.* **39**(2): 103–114.
35. McIntosh, M.D., Benbow, M.E., and Burky, A.J. (2002). Effects of stream diversion on riffle macroinvertebrate communities in a Maui, Hawaii, stream. *River Res. Applic.* **18**: 569–581.
36. Wilcox, C. (1996). *Sugar Water: Hawaii's Plantation Ditches*. University of Hawaii Press, Honolulu, HI.
37. Mackenzie, M.K. (1991). *Native Hawaiian Rights Handbook*. Native Hawaiian Legal Corporation, Honolulu, HI.
38. Liliuokalani. (1990). *Hawaii's Story by Hawaii's Queen*. Mutual Publishing, Honolulu, HI.
39. Daws, G. (1968). *Shoal of Time: A History of the Hawaiian Islands*. University of Hawaii Press, Honolulu, HI.
40. Cooper, G. and Daws, G. (1990). *Land and Power in Hawaii*. University of Hawaii Press, Honolulu, HI.
41. MacKenzie, M.K. (1991). Water rights, In: *Native Hawaiian Rights Handbook*. M.K. MacKenzie (Ed.). Native Hawaiian Legal Corporation, Honolulu, HI, pp. 149–172.
42. Schoenbaum, T.J. and Rosenburg, R.H. (1996). *Environmental Policy Law: Problems, Cases and Readings*. University Casebook Series, Westbury, NY.

GROUND WATER

ACID MINE DRAINAGE: SOURCES AND TREATMENT IN THE UNITED STATES

Jeffrey G. Skousen
West Virginia University
Morgantown, West Virginia

Acid mine drainage (AMD) occurs when metal sulfides are exposed to oxidizing conditions. Leaching of reaction products into surface waters pollute over 20,000 km of streams in the United States alone. Mining companies must predict the potential of creating AMD by using overburden analyses. Where a potential exists, special handling of overburden materials and quick coverage of acid-producing materials in the backfill should be practiced. The addition of acid-neutralizing materials can reduce or eliminate AMD problems. Placing acid-producing materials under dry barriers can isolate these materials from air and water. Other AMD control technologies being researched include injection of alkaline materials (ashes and limestone) into abandoned underground mines and into buried acid material in mine backfills, remining of abandoned areas, and installation of alkaline recharge trenches. Chemicals used for treating AMD are $Ca(OH)_2$, CaO, NaOH, Na_2CO_3, and NH_3, with each having advantages under certain conditions. Under low-flow situations, all chemicals except $Ca(OH)_2$ are cost effective, whereas at high flow, $Ca(OH)_2$ and CaO are clearly the most cost effective. Floc, the metal hydroxide material collected after treatment, is disposed of in abandoned deep mines, refuse piles, or left in collection ponds. Wetlands remove metals from AMD through formation of oxyhydroxides and sulfides, exchange and organic complexation reactions, and direct plant uptake. Aerobic wetlands are used when water contains enough alkalinity to promote metal precipitation, and anaerobic wetlands are used when alkalinity must be generated by microbial sulfate reduction and limestone dissolution. Anoxic limestone drains are buried trenches of limestone that intercept AMD underground to generate alkalinity. Under anoxia, limestone should not be coated with Fe^{+3} hydroxides in the drain, which decreases the likelihood of clogging. Vertical flow wetlands pretreat oxygenated AMD with organic matter to remove oxygen and Fe^{+3}, and then the water is introduced into limestone underneath the organic matter. Open limestone channels use limestone in aerobic environments to treat AMD. Coating of limestone occurs, and the reduced limestone dissolution is designed into the treatment system. Alkaline leach beds, containing either limestone or slag, add alkalinity to acid water. At present, most passive systems offer short-term treatment and are more practical for installation on abandoned sites or watershed restoration projects where effluent limits do not apply and where some removal of acid and metals will benefit a stream.

Acid mine drainage (AMD) forms when sulfide minerals deep in the earth are exposed during coal and metal mining, highway construction, and other large-scale excavations. Upon exposure to water and oxygen, sulfide minerals oxidize to form acidic products, which then can be dissolved in water. The water containing these dissolved products often has a low pH, high amounts of dissolved metals such as iron (Fe) and aluminum (Al), and sulfate.

The metal concentrations in AMD depend on the type and quantity of sulfide minerals present, and the overall water quality from disturbed areas depends on the acid-producing (sulfide) and acid-neutralizing (carbonate) minerals contained in the disturbed rock. The carbonate content of overburden determines whether there is enough neutralization potential or base to counteract the acid produced from pyrite oxidation. Of the many types of acid-neutralizing compounds present in rocks, only carbonates (and some clays) occur in sufficient quantity to effectively neutralize acid-producing rocks. A balance between the acid-producing potential and neutralizing capacity of the disturbed overburden will indicate the ultimate acidity or alkalinity that might be expected in the material upon complete weathering.

Approximately 20,000 km of streams and rivers in the United States are degraded by AMD, but sulfide minerals occur throughout the world causing similar problems. About 90% of the AMD reaching streams originates in abandoned surface and deep mines. No company or individual claims responsibility for reclaiming abandoned mine lands and contaminated water flowing from these sites is not treated.

Control of AMD before land disturbance requires an understanding of three important factors: (1) overburden geochemistry, (2) method and precision of overburden handling and placement in the backfill during reclamation, and (3) the postmining hydrology of the site.

OVERBURDEN ANALYSES, HANDLING, AND PLACEMENT

Premining analysis of soils and overburden are required by law (1). Identifying the acid-producing or acid-neutralizing status of rock layers before disturbance aids in developing overburden handling and placement plans. Acid-base accounting provides a simple, relatively inexpensive, and consistent procedure to evaluate overburden chemistry. It balances potential acidity (based on total or pyritic sulfur content) against total neutralizers. Samples containing more acid-producing than acid-neutralizing materials are “deficient” and can cause AMD, whereas those rock samples with the reverse situation have “excess” neutralizing materials and will not cause AMD. Rock layers with equal proportions of each type of material should be subjected to leaching or weathering analyses (2). Kinetic tests such as humidity cells and leach columns are important because they examine the rate of acid-producing and neutralization reactions. This information from kinetic tests can supplement information given by acid-base accounting and help regulators in permitting decisions (3).

The prevailing approach to control AMD is to keep water away from pyritic material. Once overburden materials have been classified, an overburden handling and placement plan for the site can be designed.

Segregating and placing acid-producing materials above the water table is generally recommended (2,4). Where alkaline materials overwhelm acid-producing materials, no special handling is necessary. Where acid-producing materials cannot be neutralized by onsite alkaline materials, it is necessary to import a sufficient amount to neutralize the potential acidity or the disturbance activity may not be allowed.

Postmining Hydrology

The hydrology of a backfill and its effect on AMD are complex. Generally, the porosity and hydraulic conductivity of the materials in a backfill are greater than those of the consolidated rock overburden that existed before mining, and changes in flow patterns and rates should be expected after mining (5). Water does not move uniformly through the backfill by a consistent wetting front. As water moves into coarse materials in the backfill, it follows the path of least resistance and continues downward through voids or conduits until it encounters a barrier or other compacted layer. Therefore, the chemistry of the water from a backfill will reflect only the rock types encountered in the water flow path, and not the entire geochemistry of the total overburden (6).

Diverting surface water above the site to decrease the amount of water entering the mined area is highly recommended. If it cannot be diverted, incoming water can be treated with limestone to improve water quality. Under certain conditions, pyritic material can be placed where it will be rapidly and permanently inundated, thereby preventing oxidation. Inundation is only suggested where a water table may be reestablished, such as below drainage deep mines (see WET COVERS).

CONTROL OF AMD

Acid mine drainage control can be undertaken where AMD exists or is anticipated. Control methods treat the acid-producing rock directly and stop or retard the production of acidity. Treatment methods add chemicals directly to acidic water exiting the rock mass. Companies disturbing land in acid-producing areas must often treat AMD, and they face the prospect of long-term water treatment and its liabilities and expense. Cost-effective methods, which prevent the formation of AMD at its source, are preferable. Some control methods are most suited for abandoned mines, and others are only practical on active operations. Other methods can be used in either setting.

Land Reclamation

Backfilling (regrading the land back to contour) and revegetation together are effective methods of reducing acid loads from disturbed lands (7). Water flow from seeps can be reduced by diversion and reclamation, and on some sites where flow may not be reduced, water quality can change from acid to alkaline by proper handling of overburden. Diverting surface water or channeling surface waters to control volume, direction, and contact time can minimize the effects of AMD on receiving streams. Surface diversion involves construction of drainage ditches to move surface water quickly off the site before infiltration or by providing impervious channels to convey water across the disturbed area.

Alkaline Amendment to Active Disturbances

Certain alkaline amendments can control AMD from acid-producing materials (8–11). All alkaline amendment schemes rely on acid-base accounting or kinetic tests to identify the required alkalinity for neutralization of acidic materials. Special handling of overburden seeks to blend acid-producing and acid-neutralizing rocks in the disturbance/reclamation process to develop a neutral rock mass. The pit floor or material under coal is often rich in pyrite, so isolating it from groundwater may be necessary by building highwall drains (which move incoming groundwater away from the pit floor) or placing impermeable barriers on the pit floor. Acid-forming material can be compacted or capped within the spoil (12).

If insufficient alkalinity is available in the spoil, then external sources of alkalinity must be imported (13,14). Limestone is often the least expensive and most readily available source of alkalinity. It has a neutralization potential of between 75% and 100%, and it is safe and easy to handle. On the other hand, it has no cementing properties and cannot be used as a barrier. Fluidized bed combustion ashes generally have neutralizing amounts of between 20% and 40%, and they tend to harden into cement after wetting (15). Other power-generation ashes, like flue gas desulfurization products and scrubber sludges, may also have significant neutralization potential, which make them suitable alkaline amendment materials (16). Other materials, like kiln dust, produced by lime and cement kilns, or lime muds, grit, and dregs from pulp and paper industries contain neutralization products (10). Steel slags, when fresh, have neutralizing amounts from 45% to 90%. Slags are produced by several processes, so care is needed to ensure that candidate slags are not prone to leaching metal ions like Cr, Mn, and Ni. Phosphate rock has been used in some studies to control AMD. It may react with Fe released during pyrite oxidation to form insoluble coatings (17), but phosphate usually costs much more than other calcium-based amendments and is needed in about the same amounts (18).

Alkaline Recharge Trenches

Alkaline recharge trenches (19) are surface ditches or cells filled with alkaline material, which can minimize or eliminate acid seeps through an alkaline-loading process with infiltrating water. Alkaline recharge trenches were constructed on top of an 8-ha, acid-producing coal refuse disposal site, and after 3 years, the drainage water showed 25% to 90% acidity reductions with 70% to 95% reductions in Fe and sulfate (20). Pumping water into alkaline trenches greatly accelerates the movement of alkalinity into the backfill and can cause acid seeps to turn alkaline (21).

Dry Barriers

Dry barriers retard the movement of water and oxygen into areas containing acid-producing rock. These “water control” technologies (4) include impervious membranes, dry seals, hydraulic mine seals, and grout curtains/walls. Surface barriers can achieve substantial reductions in water flow through piles, but generally they do not control AMD completely. Grouts can separate acid-producing rock and groundwater. Injection of grout barriers or curtains may significantly reduce the volume of groundwater moving through backfills. Gabr et al. (22) found that a 1.5-m-thick grout wall (installed by pumping a mixture of Class F fly ash and Portland cement grout into vertical boreholes near the highwall) reduced groundwater inflow from the highwall to the backfill by 80%, which results in some seeps drying up and others being substantially reduced in flow. At the Heath Steele Metal Mine in New Brunswick, a soil cover was designed to exclude oxygen and water from a tailings pile (23). It consisted of a 10-cm gravel layer for erosion control, 30-cm gravel/sand layer as an evaporation barrier, 60-cm compacted till (conductivity of 10^{-6} cm/sec), 30-cm sand, and pyritic waste rock. This barrier excluded 98% of precipitation, and oxygen concentrations in the waste rock dropped from 20% initially to around 1%. At the Upshur Mining Complex in West Virginia, Meek (12) reported covering a 20-ha spoil pile with a 39-mil PVC liner, and this treatment reduced acid loads by 70%.

Wet Covers

Disposal of sulfide tailings under a water cover, such as in a lake or fjord, is another way to prevent acid generation by excluding oxygen from sulfides. Wet covers also include flooding of aboveground tailings in ponds. Fraser and Robertson (24) studied four freshwater lakes used for subaqueous tailings disposal and found that the reactivity of tailings under water was low and that there were low concentrations of dissolved metals, thereby allowing biological communities to exist.

Alkaline Amendment to Abandoned Mines

Abandoned surface mines comprise huge volumes of spoil of unknown composition and hydrology. Abandoned underground mines are problematic because they are often partially caved and flooded, cannot be accessed, and have unreliable or nonexistent mine maps. Re-handling and mixing alkalinity into an already reclaimed backfill is generally prohibitively expensive.

Filling abandoned underground mine voids with nonpermeable materials is one of the best methods to prevent AMD. Underground mine voids are extensive (a 60-ha mine with a coal bed height of 1.5 m and a recovery rate of 65% would contain about 600,000 m^3 of voids), so fill material and the placement method must be cheap. Mixtures of Class F fly ash and 3–5% Portland cement control subsidence in mined-under residential areas and these slurries are generally injected through vertical boreholes at between 8- and 16-m centers. Pneumatic (air pressure) and slurry injection for placing fly ash in abandoned underground mines can extend the borehole spacing to about 30 m (25). On reclaimed surface mines still producing AMD, researchers in Pennsylvania saw small improvements in water quality after injecting coal combustion residues into buried pods of pyritic materials.

Remining and Reclamation

“Remining” means returning to abandoned surface or underground mines for further coal removal. Where AMD occurs, remining reduces acid loads by (1) decreasing infiltration rates, (2) covering acid-producing materials, and (3) removing the remaining coal, which is the source of most of the pyrite. Hawkins (26) found contaminant loads of 57 discharges from remined sites in Pennsylvania to be reduced after remining and reclamation. Short-term loads were sometimes increased during the first six months after remining and reclamation, but reduction in loads after six months resulted from decreased flow rather than large changes in concentrations. Ten remining sites in Pennsylvania and West Virginia were reclaimed to current standards (which included eliminating highwalls, covering refuse, and revegetating the entire area), and all sites had improved water quality (15).

CHEMICAL TREATMENT OF AMD

If AMD problems develop during mining or after reclamation, a plan to treat the discharge must be developed. A water treatment system consists of an inflow pipe or ditch, a storage tank or bin holding the treatment chemical, a valve to control its application rate, a settling pond to capture precipitated metal oxyhydroxides, and a discharge point. At the discharge point, water samples are analyzed to monitor whether specified parameters are being attained. Water discharge permits (NPDES) on surface mines usually require monitoring of pH, total suspended solids, and Fe and Mn concentrations. The type and size of a chemical treatment system is based on flow rate, pH, oxidation status, and concentrations of metals in the AMD. The receiving stream’s designated use and seasonal fluctuations in flow rate are also important. After evaluating these variables over a period of time, the operator can consider the economics of different chemicals.

Six chemicals treat AMD (Table 1). Each is more or less appropriate for a specific condition. The best choice depends on both technical (acidity levels, flow, and the types and concentrations of metals) and economic factors (chemical prices, labor, machinery and equipment, treatment duration, and interest rates). Enough alkalinity must be added to raise pH to between 6 and 9 so insoluble metal hydroxides will form and settle out. Treatment of AMD with high Fe (ferric) concentrations often affords coprecipitation of other metals with the Fe hydroxide, thereby removing them from AMD at a lower pH. Limestone has been used for decades to raise pH and precipitate metals in AMD. It has the lowest material cost and is the safest and easiest to handle of the AMD chemicals. Unfortunately, it is limited because of its low solubility and tendency to develop an external coating, or armor, of $Fe(OH)_3$ when added to AMD. Fine-ground limestone may be dumped in streams directly or the limestone may be pulverized by water-powered rotating drums and metered into the stream. Limestone has also treated AMD in anaerobic (anoxic limestone drains) and aerobic environments (open limestone channels).

Table 1. Chemical Compounds Used in AMD Treatment

Common Name	Chemical Name	Formula	Conversion Factor[a]	Neutralization Efficiency[b]	2000 Cost[c] $ per Mg or L Bulk	<Bulk
Limestone	Calcium carbonate	$CaCO_3$	1.00	30%	$11	$16
Hydrated Lime	Calcium hydroxide	$Ca(OH)_2$	0.74	90%	$66	$110
Pebble Quicklime	Calcium oxide	CaO	0.56	90%	$88	$264
Soda Ash	Sodium carbonate	Na_2CO_3	1.06	60%	$220	$350
Caustic Soda (solid)	Sodium hydroxide	NaOH	0.80	100%	$750	$970
20% Liquid Caustic	Sodium hydroxide	NaOH	784	100%	$0.06	$0.16
50% Liquid Caustic	Sodium hydroxide	NaOH	256	100%	$0.29	$0.33
Ammonia	Anhydrous ammonia	NH_3	0.34	100%	$330	$750

[a] The conversion factor may be multiplied by the estimated milligrams acid/yr to get milligrams of chemical needed for neutralization per year. For liquid caustic, the conversion factor gives liters needed for neutralization.

[b] Neutralization efficiency estimates the relative effectiveness of the chemical in neutralizing AMD acidity. For example, if 100 Mg of acid/yr was the amount of acid to be neutralized, then it can be estimated that 82 Mg of hydrated lime would be needed to neutralize the acidity in the water (100(0.74)/0.90).

[c] Price of chemical depends on the quantity being delivered. Bulk means delivery of chemical in a large truck, whereas < Bulk means purchased in small quantities. Liquid caustic prices are for liters. Others in milligrams.

Lime

Hydrated lime is common for treating AMD. As a powder, it tends to be hydrophobic, and extensive mechanical mixing is required for dissolution. Hydrated lime is particularly useful and cost effective in large-flow, high-acidity situations where a lime treatment plant with a mixer/aerator is constructed to help dispense and mix the chemical with the water (27). Hydrated lime has limited effectiveness if a very high pH (>9) is required to remove ions such as Mn. Unfortunately, increasing the lime rate increases the volume of unreacted lime that enters the floc-settling pond.

Pebble quicklime (CaO) is used with the Aquafix Water Treatment System using a water wheel concept (28). A water wheel is turned based on water flow, which causes a screw feeder to dispense the chemical. This system was initially used for small and/or periodic flows of high acidity because CaO is very reactive, but water wheels have been attached to large silos for high-flow/high-acidity situations. Tests show an average of 75% cost savings over NaOH systems and about 20% to 40% savings over NH_3 systems.

Soda Ash

Soda ash (Na_2CO_3) generally treats AMD in remote areas with low flow and low amounts of acidity and metals. This choice is usually based on convenience rather than on chemical cost. Soda ash comes as solid briquettes and is gravity fed into water through bins. The number of briquettes used per day is determined by the rate of flow and quality of the water. One problem is that the briquettes absorb moisture, expand, and stick to the corners of the bin and will not drop into the stream. For short-term treatment, some operators use a much simpler system that employs a wooden box or barrel with holes that allows water inflow and outflow. The operator simply fills the barrel with briquettes on a regular basis and places the barrel in the flowing water. This system offers less control of the amount of chemical used.

Caustic Soda

Caustic soda (i.e., lye, NaOH) is often used in remote low-flow, high-acidity situations, or if Mn concentrations in the AMD are high. The system can be gravity fed by dripping liquid NaOH directly into the AMD. Caustic is very soluble, disperses rapidly, and raises the pH quickly. Caustic should be applied at the surface of ponds because the chemical is denser than water. The major drawbacks of using liquid NaOH for AMD treatment are high cost and dangers in handling.

Ammonia

Ammonia compounds (NH_3 or NH_4OH) are extremely hazardous. NH_3 is compressed and stored as a liquid but returns to the gaseous state when released. Ammonia is extremely soluble, reacts rapidly, and can raise the pH of receiving water to 9.2. At pH 9.2, it buffers the solution to further pH increases, and therefore very high amounts of NH_3 must be added to go beyond 9.2. Injection of NH_3 into AMD is one of the quickest ways to raise water pH, and it should be injected near the bottom of the pond or water inlet because NH_3 is less dense than water. NH_3 is cheap, and a cost reduction of 50% to 70% is usually realized when NH_3 is substituted for NaOH (29). Major disadvantages of using NH_3 include (1) the hazards; (2) uncertainty concerning nitrification, denitrification, and acidification downstream; and (3) consequences of excessive application rates, which cause toxic conditions to aquatic life.

Costs of Treating AMD

Costs were estimated for five treatment chemicals under four sets of flow and acid concentration conditions [Table 1 from Skousen et al. (30)]. Na_2CO_3 had the highest labor requirements (10 hours per week) because the dispensers must be filled by hand and inspected frequently. Caustic had the highest reagent cost per mole of acid-neutralizing capacity, and Na_2CO_3 had the second highest. Hydrated lime treatment systems had the highest installation costs of the five chemicals because of the need to construct a lime treatment plant and install a pond aerator. However, the cost of $Ca(OH)_2$ was very low, and the combination of high installation costs and low reagent cost made $Ca(OH)_2$ systems particularly appropriate for long-term treatment of high-flow/high-acidity conditions.

For a 5-year treatment, NH_3 had the lowest annual cost for the low-flow/low-acid situation. Pebble quicklime had about the same cost as the NH_3 system, but slightly higher installation costs. Caustic was third because of its high labor and reagent costs, and Na_2CO_3 was fourth because of high labor costs. Hydrated lime was the most expensive because of its high installation costs. At high-flow/high-acidity, the $Ca(OH)_2$ and CaO systems were clearly the cheapest treatment systems (annual costs of about $250,000 less than NH_3, the next best alternative).

After chemical treatment, the treated water flows into sedimentation ponds so metals in the water can precipitate. All AMD treatment chemicals cause the formation of metal hydroxide sludge or floc. Sufficient residence time of the water (dictated by pond size and depth) is important for adequate metal precipitation. The amount of metal floc generated depends on water quality and quantity, which in turn determines how often the ponds must be cleaned. Knowing the chemical and AMD being treated will provide an estimate of the stability of metal compounds in the floc. Floc disposal options include (1) leaving it submerged indefinitely, (2) pumping or hauling it to abandoned deep mines or to pits dug on surface mines, and (3) dumping it into refuse piles. Pumping flocs onto land and letting them age and dry is a good strategy for disposal, because they become crystalline and behave like soil material.

Each AMD is unique, requiring site-specific treatment. Each AMD source should be tested with various chemicals by titration tests to evaluate the most effective chemical for precipitation of the metals. The costs of each AMD treatment system based on neutralization (in terms of the reagent cost, capital investment, and maintenance of the dispensing system) and floc disposal should be evaluated to determine the most cost-effective system.

PASSIVE TREATMENT OF AMD

Active chemical treatment of AMD is often an expensive, long-term proposition. Passive treatment systems have been developed that do not require continuous chemical inputs and that take advantage of natural chemical and biological processes to cleanse contaminated mine waters. Passive technologies include constructed wetlands, anoxic limestone drains, vertical flow wetlands (also known as SAPS), open limestone channels, and alkaline leach beds (Fig. 1). In low-flow and low-acidity situations, passive systems can be reliably implemented as a single permanent solution for many AMD problems.

Constructed Wetlands

Wetlands are of two basic types: aerobic and anaerobic. Metals are retained within wetlands by (1) formation of metal oxides and oxyhydroxides, (2) formation of metal sulfides, (3) organic complexation reactions, (4) exchange with other cations on negatively charged sites, and (5) direct uptake by living plants. Other beneficial reactions in wetlands include generation of alkalinity caused by microbial mineralization of dead organic matter, microbial dissimilatory reduction of Fe oxyhydroxides and SO_4, and dissolution of carbonates.

Aerobic wetlands consist of relatively shallow ponds (<30 cm) with wetland vegetation. Aerobic wetlands promote metal oxidation and hydrolysis, thereby causing precipitation and physical retention of Fe, Al, and Mn oxyhydroxides. Successful metal removal depends on dissolved metal concentrations, dissolved oxygen content, pH and net acidity of the mine water, the presence of active microbial biomass, and detention time of the water in the wetland. The pH and net acidity/alkalinity of the water are particularly important because pH influences both the solubility of metal hydroxide precipitates and the kinetics of metal oxidation and hydrolysis. Therefore, aerobic wetlands are best used in conjunction with water that contains net alkalinity to neutralize metal acidity.

Anaerobic wetlands consisting of deep ponds (>30 cm) with substrates of soil, peat moss, spent mushroom compost, sawdust, straw/manure, hay bales, or other organic mixtures, often underlain or admixed with limestone. Anaerobic wetlands are most successful when used to treat small flows of acidic water. Anaerobic wetlands use chemical and microbial reduction reactions to precipitate metals and neutralize acidity. The water infiltrates through a thick permeable organic subsurface that becomes anaerobic because of high biological oxygen demand. Other chemical mechanisms that occur *in situ* include metal exchanges, formation and precipitation of metal sulfides, microbial-generated alkalinity, and formation of carbonate alkalinity (because of limestone dissolution). As anaerobic wetlands produce alkalinity, they can be used in net acidic and high dissolved oxygen (>2 mg/L) AMD. Microbial mechanisms of alkalinity production are critical to long-term AMD treatment. Under high acid loads (>300 mg/L), pH-sensitive microbial activities are eventually overwhelmed. At present, the sizing value for Fe removal in these wetlands is 10 gs per day per meter squared (31).

Sorption onto organic materials (such as peat and sawdust) can initially remove 50% to 80% of the metals in AMD (32), but the exchange capacity declines with time. Over the long term, metal hydroxide precipitation is the predominant form of metal retention in a wetland. Wieder (33) reported up to 70% of the Fe in a wetland to be composed of Fe^{+3} oxyhydroxides, whereas the other 30% is reduced and combined with sulfides (34).

Sulfate reducing bacteria (SRB) reactors have been used to generate alkalinity by optimizing anaerobic conditions. Good success has been noted for several systems receiving high and low flows (35,36).

Anoxic Limestone Drains

Anoxic limestone drains are buried cells or trenches of limestone into which anoxic water is introduced. The limestone raises pH and adds alkalinity. Under anoxic conditions, the limestone does not coat or armor with Fe hydroxides because Fe^{+2} does not precipitate as $Fe(OH)_2$ at pH 6.0. Faulkner and Skousen (37) reported both successes and failures among 11 anoxic drains in WV. Failures resulted when ferric iron and Al precipitate as hydroxides in the limestone causing plugging and coating.

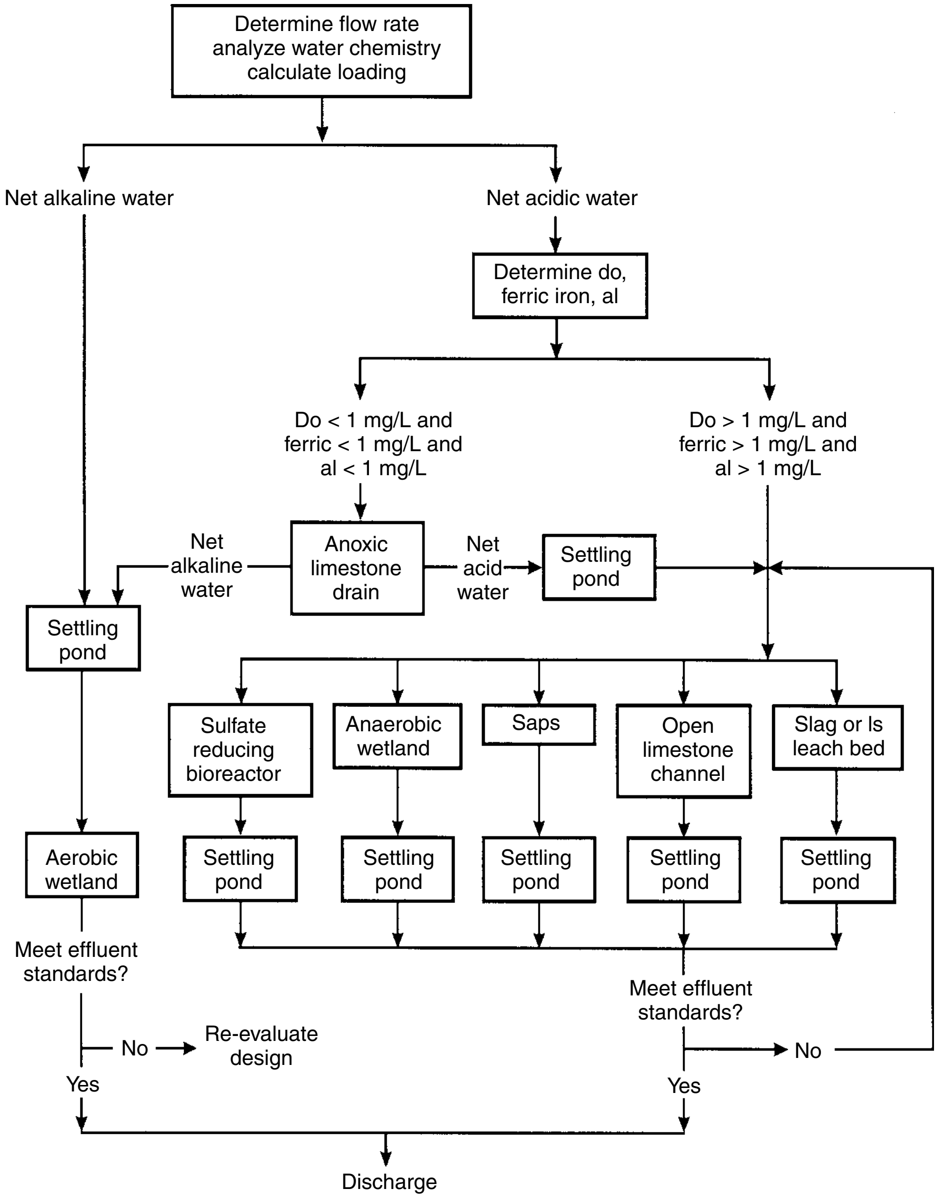

Figure 1. Diagram of possible passive treatment systems to treat mine water based on water flow and chemistry.

Longevity of treatment is a major concern for anoxic drains, especially in terms of water flow through the limestone. Selection of the appropriate water and environmental conditions is critical for long-term alkalinity generation in an anoxic drain. Eventual clogging of the limestone pore spaces with precipitated Al and Fe hydroxides, and gypsum is predicted (38). For optimum performance, no Fe^{+3}, dissolved oxygen, or Al should be present in the AMD. Like wetlands, anoxic limestone drains may be a solution for AMD treatment for specific water conditions or for a finite period after which the system must be replenished or replaced.

Vertical Flow Wetlands

In these modified wetlands [called SAPS by Kepler and McCleary (39)], 1 to 3 m of acid water is ponded over an organic compost of 0.2 to 0.3 m, underlain by 0.5 to 1 m of limestone. Below the limestone are drainage pipes that convey the water into an aerobic pond where metals are precipitated. The hydraulic head drives ponded water through the anaerobic organic compost, where oxygen stripping as well as Fe and sulfate reduction can occur before water entry into the limestone. Water with high metal loads can be successively cycled through additional wetlands. Iron and Al clogging of limestone and pipes can be removed by flushing the system (40). Much work is being done on these wetlands presently, and refinements are being made for better water treatment.

Open Limestone Channels

Open limestone channels are another means of introducing alkalinity to acid water (41). We usually assume that armored limestone ceases to dissolve, but Ziemkiewicz et al. (42) found armored limestone to be 50% to 90% effective in neutralizing acid compared with unarmored limestone. Seven open channels in the field reduced acidity in AMD by 4% to 62% compared with a 2% acid reduction in a sandstone channel. Open limestone channels show promise for neutralizing AMD in watershed restoration projects and AML reclamation projects where there can be only a one-time installation cost, little to no maintenance is required, and water exiting the system does not have to meet water quality standards. Long channels of limestone can convey acid water to a stream or other discharge point. Cross sections of channels can be designed with calculated amounts of limestone (which will become armored) to treat the water. Open limestone channels work best on steep slopes (>20%), where flow velocities keep metal hydroxides in suspension, thereby limiting plugging. If constructed correctly, open limestone channels should be maintenance free and provide AMD treatment for decades.

Alkaline Leach Beds

Limestone, when placed in an open pond or leach bed, will dissolve slowly over time and continually add alkalinity to water unless the limestone gets coated with metal hydroxides, thereby reducing its dissolution rate (41). Therefore, limestone treatment in aerobic systems works best in low-pH, metal-free water, and can add alkalinity to streams before encountering acid water downstream (42). As limestone generally reacts relatively slowly under field conditions, steel slag, a byproduct of steel making and composed of hydrated amorphous silica and calcium compounds, can be used as an alkaline material to add alkalinity to water. Steel slags have high neutralization potentials (from about 50–70%), can generate exceptionally high levels of alkalinity in water, and do not armor (43). Steel slag fines can be used in leach beds. Effluents from slag leach beds attain high pH (>10) and have alkalinity concentrations in the thousands of milligrams/liter. Slag leach beds may receive AMD directly, or effluent from "fresh water" beds may be combined with an AMD source downstream to treat acid indirectly.

SUMMARY

Acid mine drainage occurs when metal sulfides are oxidized. Leaching of reaction products into surface waters pollute over 20,000 km of streams in the United States alone. Companies must predict AMD before mining by using overburden analyses. On sites where a potential exists, special handling of overburden materials and quick coverage of acid-producing materials in the backfill should be practiced. Alkaline addition with materials such as kiln dust and FBC ash can reduce or completely eliminate AMD problems. Other control techniques include dry barriers, wet barriers, injection of alkaline materials into underground mines, remining of abandoned areas, and alkaline recharge trenches. Five chemicals typically treat AMD, and each has characteristics that make it suitable for specific applications. Companies must select a chemical that treats the water adequately and cost-effectively. Passive systems are low maintenance systems that are implemented on abandoned mine land and stream restoration projects. Certain systems are more suited to specific water quality and show good success where the acid levels do not overwhelm the system.

BIBLIOGRAPHY

1. Sobek, A., Skousen, J., and Fisher, S. (2000). Chemical and physical properties of overburdens and minesoils. In: *Reclamation of Drastically Disturbed Lands*, 2nd Edn. American Society of Agronomy, Madison, WI.
2. Skousen, J.G., Sencindiver, J.C., and Smith, R.M. (1987). *A Review of Procedures for Surface Mining and Reclamation in Areas with Acid-Producing Materials*. EWRC 871, West Virginia University, Morgantown, WV.
3. Geidel, G., Caruccio, F.T., Hornberger, R., and Brady, K. (2000). Guidelines and recommendations for use of kinetic tests for coal mining (AMD) prediction in the eastern U.S. In: *Prediction of Water Quality at Surface Coal Mines*. National Mine Land Reclamation Center, West Virginia University, Morgantown, WV.
4. Skousen, J., Rose, A., Geidel, G., Foreman, J., Evans, R., and Hellier, W. (1998). *Handbook of Technologies for Avoidance and Remediation of Acid Mine Drainage*. National Mine Land Reclamation Center, West Virginia University, Morgantown, WV.
5. Caruccio, F.T. and Geidel, G. (1989). Water management strategies in abating acid mine drainage—Is water diversion really beneficial? In: *1989 Multinational Conference on Mine Planning and Design*. 16–17 Sept. 1989, University of Kentucky, Lexington, KY.
6. Ziemkiewicz, P.F. and Skousen, J.G. (1992). Prevention of acid mine drainage by alkaline addition. *Green Lands* **22**(2): 42–51.
7. Faulkner, B.B. and Skousen, J.G. (1995). Effects of land reclamation and passive treatment systems on improving water quality. *Green Lands* **25**: 34–40.
8. Brady, K., Smith, M.W., Beam, R.L., and Cravotta, C.A. (1990). Effectiveness of the use of alkaline materials at surface coal mines in preventing or abating acid mine drainage: Part 2. Mine site case studies. In: *Proceedings, 1990 Mining and Reclamation Conference*. J. Skousen et al. (Eds.). 23–26 April 1990, West Virginia University, Morgantown, WV.
9. Perry, E.F. and Brady, K.B. (1995). Influence of neutralization potential on surface mine drainage quality in Pennsylvania. In: *Proceedings, Sixteenth Annual Surface Mine Drainage Task Force Symposium*. 4–5 April 1995, West Virginia University, Morgantown, WV.
10. Rich, D.H. and Hutchison, K.R. (1994). Coal refuse disposal using engineering design and lime chemistry. In: *International Land Reclamation and Mine Drainage Conference*. 24–29 April 1994, USDI, Bureau of Mines SP 06A-94, Pittsburgh, PA.
11. Rose, A.W., Phelps, L.B., Parizek, R.R., and Evans, D.R. (1995). Effectiveness of lime kiln flue dust in preventing acid mine drainage at the Kauffman surface coal mine, Clearfield County, Pennsylvania. In: *Proceedings, 1995 National Meeting of the American Society for Surface Mining and Reclamation*. 3–8 June 1995, Gillette, WY.
12. Meek, F.A. (1994). Evaluation of acid prevention techniques used in surface mining. In: *International Land Reclamation and Mine Drainage Conference*. 24–29 April 1994, USDI, Bureau of Mines SP 06B-94, Pittsburgh, PA.
13. Skousen, J. and Larew, G. (1994). Alkaline overburden addition to acid-producing materials to prevent acid mine drainage. In: *International Land Reclamation and Mine*

Drainage Conference. 24–29 April 1994, USDI, Bureau of Mines SP 06B-94, Pittsburgh, PA.

14. Wiram, V.P. and Naumann, H.E. (1995). Alkaline additions to the backfill: A key mining/reclamation component to acid mine drainage prevention. In: *Proceedings, Sixteenth Annual Surface Mine Drainage Task Force Symposium*. 4–5 April 1995, West Virginia University, Morgantown, WV.
15. Skousen, J., Bhumbla, D., Gorman, J., and Sencindiver, J. (1997). Hydraulic conductivity of ash mixtures and metal release upon leaching. In: *1997 National Meeting of the American Society for Surface Mining and Reclamation*. 10–15 May 1997, Austin, TX.
16. Stehouwer, R., Sutton, P., Fowler, R., and Dick, W. (1995). Minespoil amendment with dry flue gas desulfurization by-products: element solubility and mobility. *J. Environ. Qual.* **24**: 165–174.
17. Evangelou, V.P. (1995). *Pyrite Oxidation and its Control*. CRC Press, Boca Raton, FL.
18. Ziemkiewicz, P.F. and Meek, F.A. (1994). Long term behavior of acid forming rock: results of 11-year field studies. In: *International Land Reclamation and Mine Drainage Conference*. 24–29 April 1994, USDI, Bureau of Mines SP 06B-94, Pittsburgh, PA.
19. Caruccio, F.T., Geidel, G., and Williams, R. (1984). Induced alkaline recharge zones to mitigate acid seeps. In: *Proceedings, National Symposium on Surface Mining, Hydrology, Sedimentology and Reclamation*. 7–10 Dec. 1984, Univ. of Kentucky, Lexington, KY.
20. Nawrot, J.R., Conley, P.S., and Sandusky, J.E. (1994). Concentrated alkaline recharge pools for acid seep abatement: principles, design, construction and performance. In: *International Land Reclamation and Mine Drainage Conference*. 24–29 April 1994, USDI, Bureau of Mines SP 06A-94, Pittsburgh, PA.
21. Ziemkiewicz, P.F., Donovan, J., Frazier, J., Daly, M., Black, C., and Werner, E. (2000). Experimental injection of alkaline lime slurry for *in situ* remediation of an acidic surface mine aquifer. In: *Proceedings, Twenty-first West Virginia Surface Mine Drainage Task Force Symp.* April 4–5, 2000, Morgantown, WV.
22. Gabr, M.A., Bowders, J.J., and Runner, M.S. (1994). Assessment of acid mine drainage remediation schemes on groundwater flow regimes at a reclaimed mine site. In: *International Land Reclamation and Mine Drainage Conference*. 24–29 April 1994, USDI, Bureau of Mines SP 06B-94, Pittsburgh, PA.
23. Bell, A.V., Riley, M.D., and Yanful, E.G. (1994). Evaluation of a composite soil cover to control acid waste rock pile drainage. In: *International Land Reclamation and Mine Drainage Conference*. 24–29 April 1994, USDI, Bureau of Mines SP 06B-94, Pittsburgh, PA.
24. Fraser, W.W. and Robertson, J.D. (1994). Subaqueous disposal of reactive mine waste: an overview and update of case studies-MEND/Canada. In: *International Land Reclamation and Mine Drainage Conference*. 24–29 April 1994, USDI, Bureau of Mines SP 06A-94, Pittsburgh, PA.
25. Burnett, J.M., Burnett, M., Ziemkiewicz, P., and Black, D.C. (1995). Pneumatic backfilling of coal combustion residues in underground mines. In: *Proceedings, Sixteenth Annual Surface Mine Drainage Task Force Symposium*. 4–5 April 1995, West Virginia University, Morgantown, WV.
26. Hawkins, J.W. (1994). Assessment of contaminant load changes caused by remining abandoned coal mines. In: *International Land Reclamation and Mine Drainage Conference*. 24–29 April 1994, USDI, Bureau of Mines SP 06A-94, Pittsburgh, PA.
27. Skousen, J. and Ziemkiewicz, P. (1996). *Acid Mine Drainage Control and Treatment*, 2nd Edn. National Research Center for Coal and Energy, National Mine Land Reclamation Center, West Virginia University, Morgantown, WV.
28. Jenkins, M. and Skousen, J. (2001). Acid mine drainage treatment costs with calcium oxide and the Aquafix Machine. *Green Lands* **31**: 46–51.
29. Skousen, J., Politan, K., Hilton, T., and Meek, A. (1990). Acid mine drainage treatment systems: chemicals and costs. *Green Lands* **20**(4): 31–37.
30. Skousen, J.G., Sexstone, A., and Ziemkiewicz, P. (2000). Acid mine drainage control and treatment. In: *Reclamation of Drastically Disturbed Lands*, 2nd Edn. American Society of Agronomy, Madison, WI.
31. Hedin, R.S. and Nairn, R.W. (1992). Passive treatment of coal mine drainage. *Course Notes for Workshop*. U.S. Bureau of Mines, Pittsburgh, PA.
32. Brodie, G.A., Hammer, D.A., and Tomljanovich, D.A. (1988). An evaluation of substrate types in constructed wetlands acid drainage treatment systems. In: *Mine Drainage and Surface Mine Reclamation*. 19–21 April 1988, Vol. 1, Info. Circular 9183, U.S. Bureau of Mines, Pittsburgh, PA.
33. Wieder, R.K. (1993). Ion input/output budgets for wetlands constructed for acid coal mine drainage treatment. *Water, Air, and Soil Pollution* **71**: 231–270.
34. Wieder, R.K. (1992). The Kentucky wetlands project: A field study to evaluate man-made wetlands for acid coal mine drainage treatment. Final Report to the U.S. Office of Surface Mining, Villanova Univ., Villanova, PA.
35. Canty, M. (2000). Innovative *in situ* treatment of acid mine drainage using sulfate-reducing bacteria. In: *Proceedings, Fifth International Conference on Acid Rock Drainage*. Society for Mining, Metallurgy, and Exploration, Inc., Denver, CO.
36. Gusek, J., Mann, C., Wildeman, T., and Murphy, D. (2000). Operational results of a 1200 gpm passive bioreactor for metal mine drainage, Missouri. In: *Proceedings, Fifth International Conference on Acid Rock Drainage*. Society for Mining, Metallurgy, and Exploration, Inc., Denver, CO.
37. Faulkner, B.B. and Skousen, J.G. (1994). *Treatment of Acid Mine Drainage by Passive Treatment Systems*. In: *International Land Reclamation and Mine Drainage Conference*. 24–29 April 1994, USDI, Bureau of Mines SP 06A-94, Pittsburgh, PA.
38. Nairn, R.W., Hedin, R.S., and Watzlaf, G.R. (1991). A preliminary review of the use of anoxic limestone drains in the passive treatment of acid mine drainage. In: *Proceedings, Twelfth Annual West Virginia Surface Mine Drainage Task Force Symposium*. 3–4 April 1991, West Virginia University, Morgantown, WV.
39. Kepler, D.A. and McCleary, E. (1997). Passive aluminum treatment successes. In: *Proceedings, Eighteenth Annual West Virginia Surface Mine Drainage Task Force Symposium*. 15–16 April 1997, West Virginia University, Morgantown, WV.
40. Kepler, D.A. and McCleary, E. (1994). Successive alkalinity-producing systems (SAPS) for the treatment of acidic mine drainage. In: *International Land Reclamation and Mine Drainage Conference*. 24–29 April 1994, USDI, Bureau of Mines SP 06A-94, Pittsburgh, PA.
41. Ziemkiewicz, P.F., Skousen, J., and Lovett, R. (1994). Open limestone channels for treating acid mine drainage: a new look at an old idea. *Green Lands* **24**(4): 36–41.

42. Ziemkiewicz, P.F. et al. (1997). Acid mine drainage treatment with armored limestone in open limestone channels. *J. Environ. Qual.* **26**: 718–726.
43. Ziemkiewicz, P.F., Skousen, J., and Simmons, J. (2001). Cost benefit analysis of passive treatment systems. In: *Proceedings, 22nd West Virginia Surface Mine Drainage Task Force Symposium*. 3–4 April, Morgantown, WV.
44. Ziemkiewicz, P.F. and Skousen, J. (1998). The use of steel slag in acid mine drainage treatment and control. In: *Proceedings, 19th West Virginia Surface Mine Drainage Task Force Symposium*. 7–8 April, Morgantown, WV.

READING LIST

Rich, D.H. and Hutchison, K.R. (1990). Neutralization and stabilization of combined refuse using lime kiln dust at High Power Mountain. In: *Proceedings, 1990 Mining and Reclamation Conference*. 23–26 April 1990, West Virginia University, Morgantown, WV.

Skousen, J.G., Hedin, R., and Faulkner, B.B. (1997). Water quality changes and costs of remining in Pennsylvania and West Virginia. In: *1997 National Meeting of the American Society for Surface Mining and Reclamation*. 10–15 May 1997, Austin, TX.

AQUIFERS

THOMAS HARTER
University of California
Davis, California

GENERAL DEFINITION

An aquifer is a geologic formation or geologic unit from which significant amounts of groundwater can be pumped for domestic, municipal, or agricultural uses. The four major types of rock formations that serve as aquifers are unconsolidated sand and gravel, sandstone, carbonate rocks, and fractured volcanic rocks. Aquifers may also occur in other geologic formations, particularly in fractured zones of igneous, metamorphic, or sedimentary rocks.

ORIGIN OF THE WORD

The word aquifer was probably adopted around the early twentieth century from the French word aquifère, which originates from the two Latin words aqua (water) and ferre (to carry, to bear). Hence, literally translated from Latin, aquifer means 'that which carries water.'

FURTHER DEFINITIONS

There is no strict definition of the hydrogeologic attributes or volumetric extent necessary to make a geologic formation or geologic unit an aquifer. Rather, the term aquifer is used for local formations that have relatively higher permeability than surrounding formations. Geologic units that form an aquifer in one setting may therefore not be considered to constitute an aquifer in other settings. For example, a small intermontane valley aquifer that consists of very shallow, sandy loam deposits a few feet thick may supply enough water to maintain a pumping rate of 0.5–2 gallons per minute, enough for domestic water supply wells and some stock supply wells. This may be significantly more water than would be provided by the hardrock formations underlying and bounding such a shallow aquifer. In contrast, the same sandy loam deposits of the same thickness would not be considered an aquifer if they were part of an unconsolidated sedimentary sequence in a larger alluvial basin, where gravel and sand aquifers yield from 50 to more than 1,000 gallons per minute.

ROLE OF AN AQUIFER IN THE HYDROLOGIC CYCLE

Aquifers are part of the hydrologic cycle. They receive water through

- recharge from precipitation,
- recharge from irrigation return water,
- seepage from rivers and streams,
- lateral transfer of water from neighboring aquifer basins, and
- leakage from aquifer formations situated either above or below the aquifer.

Water that collects in aquifers from those sources over periods of years, decades, centuries, and even millennia is discharged back to the surface through (Fig. 1)

- springs,
- subsurface discharge into rivers and streams,
- lateral outflow to downgradient aquifers,
- vertical leakage to overlying or underlying aquifers, and
- man-made wells.

LIMITATION TO FORMATIONS WITH APPROPRIATE WATER QUALITY

The term aquifer is applied to formations that produce low to moderate salinity water (appropriate for domestic, municipal, or agricultural uses). Geologic formations containing exclusively brackish or saline groundwater (even if they are made of highly permeable material) are typically not referred to as aquifers unless the salinity was induced through human activity (e.g., in seawater intrusion, which is the advance of saline seawater into an overpumped aquifer).

AQUIFER SIZE

Aquifers can be vastly different in size: a small local aquifer in a mountainous setting may be only a few feet thick and extend over an area of a few acres to tens of acres. Other aquifers span entire regions. For example, the Ogallala aquifer in the western-central United States underlies most of the High Plains region, which extends

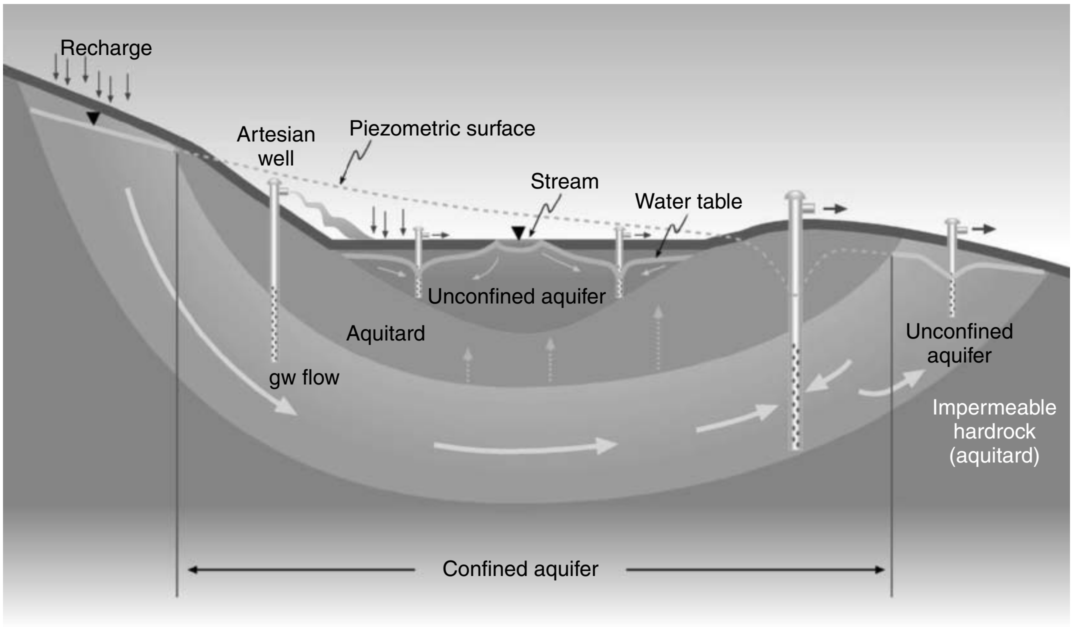

Figure 1. Schematic representation of uncontinued aquifers, confined aquifers, aquitards, and aquicludes. Blade vertical arrows indicate recharge. Black horizontal arrows indicate pumping. Light colored arrows indicate the direction of groundwater movement.

eastward from the Rocky Mountains through parts of Texas, Oklahoma, Colorado, Kansas, and Nebraska. The aquifer consists of alluvial sediments, predominantly sands and gravel. It is an important production aquifer.

An aquifer is characterized by its geologic extent (regional extent and thickness), the type of geologic formations that makes up the aquifer, the hydraulic conductivity, the transmissivity (which is defined as the product of hydraulic conductivity and aquifer thickness), the specific yield (the drainable porosity), the specific storage (the amount of water and rock compressed by hydrostatic pressure in a confined aquifer, see below), and the specific capacity (specific capacity is the amount of water pumped from a well per foot of water level drawdown created by pumping). The hydraulic conductivity of aquifers typically ranges from 1 m/day to more than 100 m/day. The specific capacity of wells located in aquifers may range from less than 0.1 gpm/ft (small, low-yielding aquifers suitable for domestic water supplies) to more than 100 gpm/ft (large production aquifers suitable for municipal and irrigation pumping).

AQUIFER CHARACTERIZATION

The amount of water that can be pumped from an aquifer depends primarily on four parameters: the hydraulic conductivity (also called the permeability) of the aquifer, the thickness of the aquifer, the specific yield or specific storage of the aquifer (related mostly to its porosity), and the amount of competition for water between wells. All four of these may change from location to location. The amount of pumping wells and the rate at which wells are pumping may be different from area to area; the thickness of the aquifer naturally changes with the thickness of the geologic formation. With respect to hydraulic conductivity, porosity, and specific yield or specific storage, hydrogeologists have found that small variations that occur in the geologic composition of aquifer formations often result in large localized changes in hydraulic conductivity. This latter phenomenon is referred to as “natural aquifer heterogeneity.” As a result of all this variability, each well within the same aquifer will have a different specific capacity. Sometimes, the specific capacity of wells can vary quite significantly from well to well, especially in fractured rock aquifers, but also in unconsolidated aquifers with sand and gravel.

Hydraulic conductivity, thickness, and specific yield or specific storage of an aquifer are determined indirectly by using literature values available for specific geologic formations, by using computer models in conjunction with local observations of groundwater fluxes or groundwater table fluctuations, or directly by performing an aquifer test (pumping test).

AQUIFERS, AQUITARDS, AND AQUICLUDES

Aquifers are the major hydrogeologic units within the hydrogeologic framework of a region from which groundwater is or can be extracted. The description of local or regional hydrogeology centers around the description of aquifers, that is, of those geologic formations with the highest significance—locally or regionally—with respect to (potential) groundwater production. Geologic formations that bound aquifers are referred to as aquicludes or aquitards. Aquicludes are, for all practical purposes, impermeable. Important aquicludes are thick, continuous clay formations and unfractured igneous rocks. Aquitards are geologic formations that have a lower hydraulic conductivity than adjacent (aquifer) formations and therefore act as a partial barrier to groundwater flow between overlying aquifers. Aquitards can consist of material similar to aquifers, but either the amount of fine sediments is much larger (in unconsolidated formations) relative to the aquifer formation, or the degree of fracturing and size of fractures is smaller than that in the aquifer formation (in hardrock formations).

CONFINED AND UNCONFINED AQUIFERS

Aquifers can be either unconfined or confined, depending on the existence of an overlying aquitard or aquiclude. In an unconfined aquifer, there is no overlying aquitard or aquiclude. Recharge to the aquifer from the land surface or from and to streams is not restricted. The water table moves freely up and down, depending on the water stored, added to, or removed from the unconfined aquifer. The water level in a borehole drilled into an unconfined aquifer will be the same as the water level in the aquifer (if we ignore the effects of the capillary fringe).

In a confined aquifer, on the other hand, water is "sandwiched" between two aquitards or between an aquitard and an aquiclude above and below the aquifer. Water in a confined aquifer is under hydrostatic pressure created by the weight of the overlying geologic formations and the water pressure created by the higher water levels in the usually remote recharge area of a confined aquifer. Due to the pressure in a confined aquifer, the water level in a borehole drilled into a confined aquifer will rise significantly above the top of the aquifer. An artesian well occurs where the pressure is so large that the water level in a well drilled into the confined aquifer rises above the land surface. A confined aquifer does not have a water table—it is completely filled with groundwater. The water level in wells drilled into a confined aquifer, instead, corresponds to the hydrostatic pressure head or potentiometric surface of the aquifer, which is located higher than the upper boundary of the aquifer itself. If the hydrostatic pressure head falls below the top of the confined aquifer, it becomes unconfined.

An aquifer that is confined by an aquitard rather than an aquiclude is referred to as a "leaky aquifer" or a "semiconfined aquifer." The aquitard is not always a contiguous layer of less permeable material. Local accumulations of multiple, smaller clay lenses and other clay-rich or otherwise impermeable layers dispersed within a more permeable formation may render the entire formation an aquitard. The actual low permeable lenses are not contiguous, but the overall effect of their presence within such a heterogeneous formation on the regional aquifer below is identical to that of a continuous aquitard formation.

PERCHED WATER TABLE

Occasionally, water collects above an impermeable or low permeability layer within the unsaturated (aerated) zone and forms a "perched" water table. By definition, a "perched" water table is a saturated groundwater zone separated from the aquifer below by a zone that is unsaturated or aerated. This should not be confused with an unconfined shallow aquifer that is separated from a deeper confined aquifer through thick but saturated layers of clay.

ARTIFICIAL RECHARGE OF UNCONFINED AQUIFER

S.N. Rai
National Geophysical Research
Institute
Hyderabad, India

V.P. Singh
Louisiana State University
Baton Rouge, Louisiana

INTRODUCTION

Groundwater plays a major role in augmenting water supply to meet the ever-increasing domestic, agricultural, and industrial demands. Increasing dependence of water supply on groundwater resources is resulting in increasing use of aquifers as a source of fresh water supply and subsurface reservoir for storing excess surface water. Aquifers are the geological formations that can store water as well as allow the flow of significant amount of water through their pores under ordinary field conditions. If the aquifer is bounded by two impermeable formations from top and bottom, it is called a confined aquifer. If the upper boundary of the aquifer is the water table, it is called an unconfined aquifer. The advantage of unconfined aquifers over confined aquifers to serve as a subsurface reservoir is that the storage of groundwater in large quantity is possible only in unconfined aquifer, which is because the storativity of the unconfined aquifer is linked to the porosity and not to the elastic properties of the water and solid matrix, as in case of the confined aquifer (1). Also, the vast surface area of the unconfined aquifer above the water table is available to receive the surface applied recharge, whereas in case of the confined aquifer, only a small open area exposed near to the ground surface or leaky portion of the aquifer boundary is available to receive the recharge (Fig. 1). This article deals with the artificial recharging of unconfined aquifer and related problems.

Natural replenishment of aquifers occurs very slowly. Therefore, withdrawal of groundwater at a rate greater than the natural replenishment rate causes declining of groundwater level, which may lead to decreased water supply, contamination of fresh water by intrusion of pollutant water from nearby sources, seawater intrusion into the aquifer of coastal areas, etc. To increase the natural replenishment, artificial recharging of the aquifer is becoming increasingly important in groundwater management. The artificial recharge may be defined as an augmentation of surface water into aquifers by some artificially planned operations. The source of water for recharge may be direct precipitation, imported water, or reclaimed wastewater. The purpose of artificial recharging

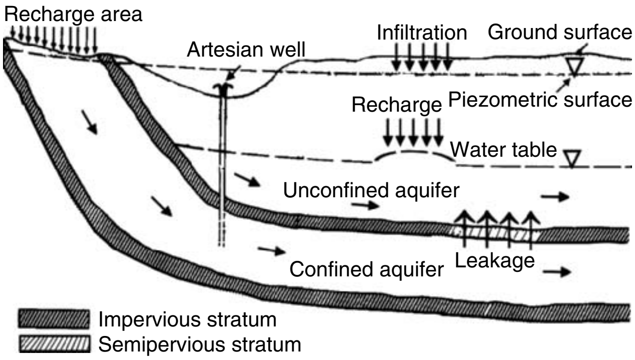

Figure 1. Aquifer types.

of groundwater systems has been to reduce, stop, or even reverse the declining trend of groundwater level; to protect fresh groundwater in coastal aquifers against saline water intrusion from the ocean; and store surface water, including flood or other surplus water, imported water, and reclaimed wastewater for future use.

RECHARGE METHODS

A variety of direct surface, direct subsurface, and indirect recharge techniques have been developed to recharge groundwater systems. The choice of a technique depends on the source of water, quality of the water, the type of aquifer, topographical condition, etc. The most widely practiced methods are direct surface techniques, which include surface flooding in basins, ponds, lakes, ditches, trenches, and furrow systems; stream and channel modification; and bunds (2–5). Trenches are constructed mostly in foothill regions to arrest the runoff and put it into the aquifers for storage. Stream channel modification involves alteration in the course of stream flow to detain stream flow and increasing the stream bed area for recharging purposes. Construction of check dams across the stream flow is one technique of stream channel modification. It enhances artificial recharge in two ways. Above the dam, impoundments enhance recharge by increasing the recharge area and detaining water for a longer period by reducing the rate of water flow. Below the dam, recharge is enhanced through exposure of a larger area than the original area of stream channel flow. Bunds, which are small earthen barriers, are constructed in agricultural lands with slopes to facilitate impounding of runoff for a longer duration, thereby increasing recharge. In indirect subsurface recharge techniques, water is injected directly into an aquifer through (a) natural openings in the aquifers, (b) pits or shafts, and (c) wells. In contrast to the direct surface techniques, groundwater recharge by indirect subsurface techniques is practiced mostly for recharging the confined aquifer and where the topography or existing land use, such as in urban areas, makes recharge by surface flooding impractical. Indirect recharge techniques involve special cases in which potable water supply is provided by river bank or sand dune filtration of generally polluted river water (6,7).

In many cases, excess recharging leads to the growth of water table near the ground surface and causes several types of environmental problems, such as water logging, soil salinity, etc. In these situations, proper management of groundwater resources is needed to overcome the shortage of water supply on one hand and to prevent the environmental problems on the other hand. In order to address the management problem, one must be able to predict the response of the aquifer system to any proposed operational policy of groundwater resources development such as artificial recharging. Such problems are referred to as forecasting problems. Their solution will provide the new state of the groundwater system. Once the new state is known, one can check whether the related recharge scheme is feasible. Then one can compare responses of different proposed recharge schemes in order to select the best scheme that can meet the preset objectives of groundwater resources development without disturbing the regional water balance and without creating any kind of environmental problems. The forecasting problems are effectively tackled by application of modeling techniques. A model is the simplified representation of a complex real physical system and the processes taking place in it. It can be physical (for example, a laboratory sand pack model), electrical analog, or mathematical. Development and applications of mathematical models are much easier than the other two types of models. Therefore, mathematical models are mostly in use today for solving groundwater management problems.

MATHEMATICAL MODELING

Modeling begins with a conceptual understanding of the physical problem (in this case, groundwater flow in the unconfined aquifer). The next step is translating the physical problem into a mathematical framework resulting in equation forms that describe the groundwater flow. Mathematical models may be deterministic, statistical, or some combination of the two. Deterministic models retain a good measure of physical insight while permitting any number of problems of the same class to be tackled with the same model. Here, discussion will be confined to deterministic models.

Formulations of groundwater flow equations are based on the conservation principles dealing with mass and momentum. These principles require that the net quantity of mass (or momentum) entering or leaving a specified volume of aquifer during a given time interval be equal to the change in the amount of mass (or moment) stored in the volume. Groundwater flow equations are formulated by combining the equation of motion in the form of Darcy's law, which follows principle of conservation of momentum with the mass balance equations, also known as continuity equations, which follows the principle of conservation of mass. Some mathematical models commonly used for solving the forecasting problem in the presence of recharge are discussed below:

2-D groundwater flow in an inhomogeneous anisotropic unconfined aquifer with a horizontal base is described by the following equation (1,8):

$$\frac{\partial}{\partial x}\left(K_x h\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y h\frac{\partial h}{\partial y}\right)+N(x,y,t)=S_y\frac{\partial h}{\partial t} \qquad (1)$$

in which h is the variable water table height measured from the horizontal base of the aquifer, K_x and K_y are the hydraulic conductivities in x and y directions, respectively, S_y is the specific yield, t is time of observation, and $N(x, y, t)$ is the sum of all recharge rates from distributed sources (recharge basins, ponds, streams, etc.) and withdrawal rates from distributed sinks (wells, leakage boundaries, etc.) and is represented by

$$N(x,y,t) = \begin{cases} \sum_{i=1}^{n} N_i(t) & \text{for } x_{i1} \le x \le x_{i2}, y_{i1} \le y \le y_{i2} \\ o & \text{elsewhere} \end{cases} \tag{2}$$

where n is the total number of basins, $N_i(t)$ is the time-varying recharge (or pumping) rate for the ith basin (or well, respectively), and x_{i1}, x_{i2}, y_{i1}, and y_{i2} are the coordinates of ith basin (or well). $N_i(t)$ is positive for recharge to the aquifer and negative for pumping.

For an inhomogeneous isotropic aquifer, Eq. (1) becomes

$$\frac{\partial}{\partial x}\left(K\,h\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K\,h\frac{\partial h}{\partial y}\right) + N(x,y,t) = S_y\frac{\partial h}{\partial t} \tag{3}$$

Equations (1 and 3) are nonlinear second-order partial differential equation. The nonlinearity is because of the presence of h as a coefficient of partial derivatives on the left-hand side. Solving these equations because of nonlinearity is possible only by numerical methods, such as finite difference, finite element, and boundary elements (1,9,10). These equations need to be linearized for their analytical solution.

For homogenous isotropic aquifers ($K = \text{constant}$), Eq. (3) can be written in the following two forms:

$$K\left[\frac{\partial}{\partial x}\left(h\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(h\frac{\partial h}{\partial y}\right)\right] + N(x,y,t) = S_y\frac{\partial h}{\partial t} \tag{4}$$

$$\frac{\partial^2 h^2}{\partial x^2} + \frac{\partial^2 h^2}{\partial y^2} + \frac{2N(x,y,t)}{K} = \frac{S_y}{Kh}\frac{\partial h^2}{\partial t} \tag{5}$$

Two procedures of lineraization are commonly used to facilitate analytical solutions of Eqs. (4 and 5). According to the first procedure, i.e., the Baumann procedure of lineraization, if the variation in h is much less than the initial height of the water table h_0, then the coefficient h appearing on the left-hand side of Eq. (4) can be replaced by h_0 (11). Then Eq. (4) can be rewritten as

$$T\left(\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right) + N(x,y,t) = S_y\frac{\partial h}{\partial t} \tag{6}$$

where $T = Kh_0$ (known as transmissivity). Now, Eq. (6) is linear in h. Sometimes the mean depth of saturation is also used in place of h_0.

In the second procedure, i.e., the Hantush procedure of linearization, h appearing in the denominator on the right-hand side of Eq. (5) is replaced by the weighted mean of the depth of saturation $\overline{h}$, a constant of linearization that is approximated by $0.5[h_0 + h(t_e)]$, and t_e is the period at the end of which $\overline{h}$ is to be approximated (12). Then, Eq. (5) becomes

$$\frac{\partial^2 h^2}{\partial x^2} + \frac{\partial^2 h^2}{\partial y^2} + \frac{2N(x,y,t)}{K} = \frac{S_y}{K\overline{h}}\frac{\partial h^2}{\partial t} \tag{7}$$

Now, Eq. (7) becomes linear in h^2. By substituting a new variable H, defined as $H = h^2 - {h_0}^2$, into Eq. (7) gives

$$\frac{\partial^2 H}{\partial x^2} + \frac{\partial^2 H}{\partial y^2} + \frac{2N(x,y,t)}{K} = \frac{S_y}{K\overline{h}}\frac{\partial H}{\partial t} \tag{8}$$

GROUNDWATER FLOW EQUATIONS FOR SLOPING AQUIFER

2-D groundwater flow in a sloping unconfined aquifer is described by (13,14)

$$\frac{\partial^2 s}{\partial x^2} + \frac{\partial^2 s}{\partial y^2} - 2a\frac{\partial s}{\partial x} + \frac{2N(t)}{K} = \frac{1}{\Re}\frac{\partial s}{\partial t} \tag{9}$$

where $s = h^2$, $a = \theta/2D$, $\theta = $ slope of the base, $D = $ the mean depth of saturation, $\Re = KD/S_y$, $x, y = $ space coordinates, $t = $ time of observation, and $N(t) = $ time varying rate of recharge.

One-dimensional groundwater flow equations can be obtained by substituting zero for the derivative of y in the above equations. These equations are used to predict the water table fluctuations in response to artificial recharge from strip basins, canals, channels, etc.

GROUNDWATER FLOW EQUATIONS IN CYLINDRICAL COORDINATES

These types of equations are used to describe groundwater flow induced by recharging through circular-shaped recharge basins/wells and is given by (15,16).

$$S\frac{\partial h}{\partial t} = -\frac{1}{r}\frac{\partial}{\partial r}(rq) + N(r) \tag{10}$$

where q is defined by Darcy's law as

$$q = -Kh\left(\frac{\partial h}{\partial r}\right) \tag{11}$$

The groundwater flow equations presented here are in the form of partial differential equations having an infinite numbers of solutions. To obtain a unique solution for a particular problem, more information about the aquifer's parameters, geometry of the flow domain and recharge basins, duration and rate of recharge rate, and initial and boundary conditions are needed. These values can be deduced from field as well as experimental methods (1,17,18).

Initial Conditions

Initial conditions describe distribution of h at all points of the flow domain at the beginning of the investigation, that is, at $t = 0$, which is expressed as

$$h = \psi(x,y,0) \tag{12}$$

where ψ is a known value of h for all points of the flow domain.

Boundary Conditions

These conditions describe the nature of interaction of the flow system with its surroundings. Three types of boundary conditions are generally encountered in groundwater flow problems.

- Dirichlet boundary condition—In this case, h is prescribed for all points of the boundary for the entire period of investigation, which is expressed as

  $$h = \psi\ (x, y, t) \tag{13}$$

  where $\psi\ (x, y, t)$ are known values of h at all points on the boundary.
- Neumann boundary condition—This type of boundary condition prescribes the flux across the boundary of the flow system and can be expressed as

  $$q = \psi_1\ (x, y, t) \tag{14}$$

  where $\psi_1\ (x, y, t)$ are known values at the boundary. A special case of this boundary condition is the no flow boundary condition in which flux is zero. This condition occurs at impermeable surface or at the groundwater divide, a surface across which no flow takes place.
- Cauchy boundary condition—This boundary condition is encountered at the semipervious boundary layer between the aquifer and an open water body such as a river. As a result of the resistance to the flow offered by the semipervious boundary that lies between the aquifer and the river, the water level in the river differs from that in the aquifer on the other side of the semipervious boundary. In this case, the flux is defined by

  $$q = K' \frac{h - h_0}{b} \tag{15}$$

  where h is the head at $x = 0$, h_0 is the water level in the river, and b and K' are the thickness and hydraulic conductivity of the semipervious boundary, respectively.

The purpose of solving a groundwater flow equation is to obtain the values of $h(x, y, t)$. Generally, two types of methods, namely analytical methods and numerical methods, are used for this purpose. Numerical methods are used to solve the nonlinear groundwater flow equation to tackle the real field problems, and analytical methods are used to solve the linearized form of groundwater flow equations. Analytical methods commonly used for the solution of groundwater problems include Fourier transforms, Laplace transforms, integral balance methods, method of separation of variables, approximate analytic methods, etc. Details about these methods can be found in many books (19–25). Most of the analytical solutions developed earlier for this purpose were based on the assumption of constant recharge. Warner et al. (26) have reviewed the performance of some such analytical solutions (27–31). However, the rate of recharge largely depends on the infiltration rate, which initially decreases because of swelling and dispersion of soil particles. After some time, the infiltration rate increases because of the release of air entrapped into soil pores and reaches to a maximum value. Then, it starts decreasing because of clogging of soil pores beneath the bottom of the basin. Recharge rate also follows a more or less similar pattern of variation with some time lag and less intensity. When it falls below a prescribed low value, the recharge operation is discontinued for some time. After drying and, if necessary, scrapping of the silted base of the basin, a high recharge rate closer to its initial value is rejuvenated in the next phase of recharge operation (1,32–34). Zomorodi (35) has also shown that the analytical solution of Dagan (36), which is based on the assumption of constant recharge rate, fails to predict the recession of the water table caused by decrease in the recharge rate. Therefore, it would be more appropriate to consider recharge rate as time-dependent to simulate the actual field conditions. Some solutions have been developed for the time-varying recharge cases in which the decreasing rate of recharge has been represented by two linear elements (37–39) or by exponential function (14,16,40–44). However, approximation of time-varying recharge by two linear elements or exponential function is possible only for one recharge cycle. However, recharge is applied intermittently for more than one cycle separated by dry periods. Manglik et al. (45), Rai et al. (46), Manglik and Rai (48), and Rai and Manglik (49) used a general scheme of recharge approximation for any number of recharge cycles. In this scheme, time-varying recharge is approximated by a number of linear elements of different lengths and slopes depending on the nature of variation of recharge rate. Later on, this scheme was modified to represent rates of recharge from any number of basins. In mathematical form, this scheme can be represented by

$$N_i(t) = \begin{cases} r_{ij}t + c_{ij} & t_j \le t \le t_{j+1} [j = 1, 2, \cdots, k-1] \\ r_{ik}t + c_{ik} & t \ge t_k \qquad [j = k] \end{cases} \tag{16}$$

where r_{ij} and c_{ij} are the slope and intercept of the jth line element of the recharge rate for the ith basin and k is the number of elements. The advantage of this scheme of recharge approximation is that any type of variation in the recharge rate for any number of recharge cycles from any number of recharge basins of different dimensions located anywhere within the flow domain can be approximated with the help of the required number of linear elements of different lengths and slopes depending on the nature of variation of recharge rate. By using this recharge scheme, several analytical solutions to describe water table fluctuation in different flow systems representing different physical conditions have been developed (50–53). The following analytical solution given by Manglik and Rai (50) is considered as an example to demonstrate the application of these solutions in prediction of water table fluctuation in the presence of time-varying recharge and pumping,

$$h^2(x,y,t) = h_0^2 + \frac{8a}{K\pi^2}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{mn}\sin\left(\frac{m\pi\, x}{A}\right)\sin\left(\frac{n\pi\, y}{B}\right)$$

$$\sum_{i=1}^{\mathrm{N}}\left[\cos\frac{m\pi\, x_{i2}}{A} - \cos\frac{m\pi\, x_{i1}}{A}\right] \tag{17}$$

$$\left[\cos\frac{n\pi\, y_{i2}}{B} - \sin\frac{n\pi\, y_{i1}}{B}\right]\left[\sum_{j=1}^{k-1}R_{ij} + R_{ik}\right]$$

in which A and B are the length and width of the aquifer, $a = K\bar{h}/S_y$, and m and n are integers representing number of Fourier coefficients. $\lambda = \frac{a\,\pi^2}{4}\left[\frac{m^2}{A^2} + \frac{n^2}{B^2}\right]$

$$R_{ij} = \frac{r_{ij}}{\lambda}[t_{j+1}\exp\{-\lambda(t-t_{j+1})\} - t_j\exp\{-\lambda(t-t_j)\}]$$
$$-\left(\frac{r_{ij}}{\lambda^2} - \frac{c_{ij}}{\lambda}\right)[\exp\{-\lambda(t-t_{j+1})\} - \exp\{-\lambda(t-t_j)\}]$$
$$R_{ik} = \frac{r_{ik}}{\lambda}[t - t_k\exp\{-\lambda(t-t_k)\}]$$
$$-\left(\frac{r_{ik}}{\lambda^2} - \frac{c_{ik}}{\lambda}\right)[1 - \exp\{-\lambda(t-t_k)\}]$$

This solution is obtained by solving Eq. (8) with recharge/pumping rates defined by Eq. (16) and subjected to the horizontal water table as an initial condition and Dirichlet boundary condition. In order to demonstrate the application of Eq. (17) in the prediction of water table fluctuation, we consider an example in which an unconfined aquifer of 10 × 10 km^2 dimension is having two recharge basins of dimension 60 × 40 m^2 and 50 × 50 m^2 centered at (4470 m, 4500 m) and (5875 m, 5530 m), respectively, and two wells each of 10 × 10 cm^2 dimension centered at (5000 m, 4500 m) and (5000 m, 5500 m), respectively. The pattern of time-varying recharge rate and pumping rate are shown in Fig. 2. The recharge operation for both the basins consists of two wet periods and one dry period, each of 20 days duration. During the first wet period, the rate of recharge decreases from its initial value of 0.8 m d^{-1} to a lower value of 0.7 m d^{-1} after 2 days. It again increases and attains maximum value of 0.9 m d^{-1} on the fourth day. After that, it starts decreasing and reduces to zero on twentieth day. The second cycle of recharge operation begins on the fortieth day and continues until the sixtieth day. The nature of variation of recharge rate for the second cycle is considered similar to the first cycle. Pumping of groundwater at a rate of 10^5 m d^{-1} from each well is considered for two periods. The first period is from the tenth to the twentieth day, and the second period is from the fortieth to the fiftieth day after a gap of 20 days. Numerical values of other controlling parameters are $h_0 = 20$ m, $K = 8$ m d^{-1}, and $S_y = 0.20$. Two water table profiles computed for $t = 45$ days along a line parallel to the x-axis at $y = 4500$ m are shown in Fig. 3. These profiles pass through the center of one recharge basin and one well centered at (4470 m, 4500m) and (5000 m, 4500 m). The profile represented by the dotted curve is in response to recharge only. Hence, it shows only growth of groundwater mound. The profile represented by continuous curve shows growth as well as depression of the water table at the respective site of recharging and pumping. This example demonstrates the capabilities of prediction of water table variations in response to time-varying recharge and withdrawal.

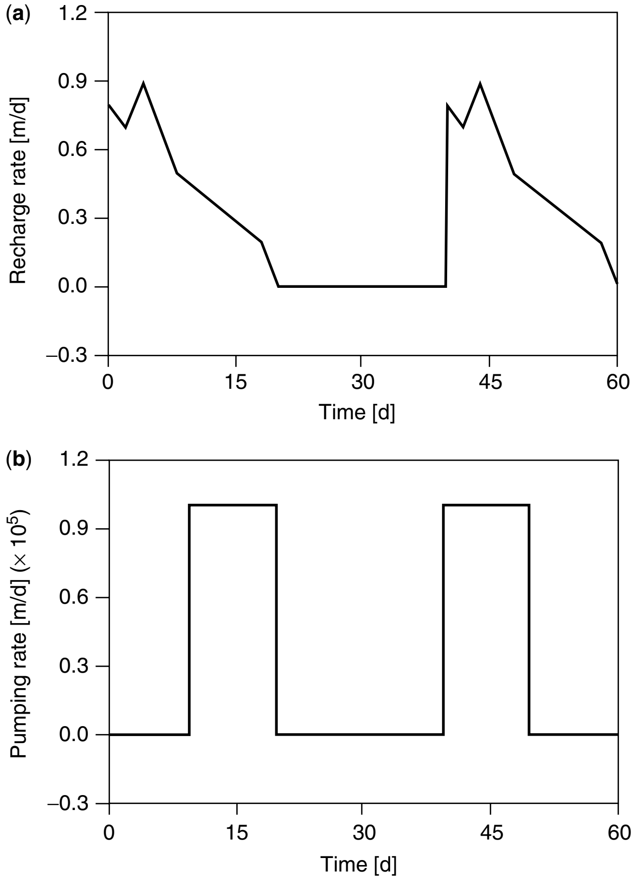

Figure 2. Nature of time-varying (**a**) recharge, and (**b**) pumping rates.

Accurate estimation of the varying recharge rate is a major problem in groundwater resources management. If the time history of water table variation at a site of an observation well is known, then analytical

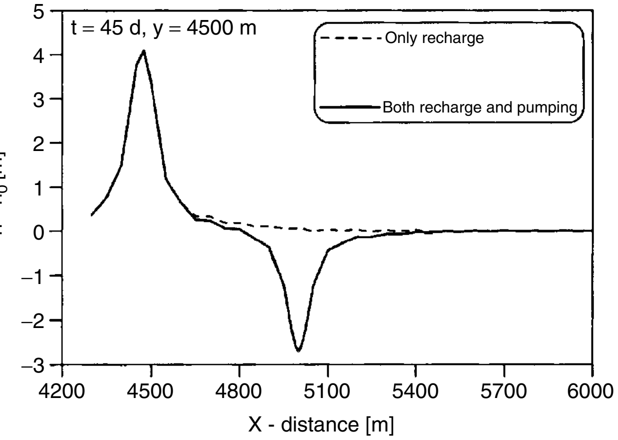

Figure 3. Water table profiles in the presence of only recharge (dashed curve), and both recharge and pumping (solid curve).

solutions can be used for the estimation of varying recharge rate by making a judicious selection of recharge rate using trial and error method, such that the computed water table variation matches well with the observed one. Although the application of analytical solutions is restricted to the relatively homogeneous isotropic aquifer system having boundaries of simple geometrical shapes, their application is fast and simple in comparison with that of the numerical methods. Analytical solutions are also useful for other purposes, such as analysis of the effects of various controlling parameters, such as aquifers properties, initial and boundary conditions, intensity and duration of recharge rate, shape, size, and location of a recharge basin, etc., on the response of the aquifer system. Such information is very essential for the judicious selection of a suitable recharge scheme out of many proposed schemes to achieve the preset objectives of groundwater resource management.

Acknowledgment

We wish to thank Dr. S. Thiagarajan for his help in preparation of this work. SNR wish to thank Dr. V.P. Dimri, Director, NGRI for according permission to publish the work.

BIBLIOGRAPHY

1. Bear, J. (1979). *Hydraulics of Groundwater*. McGraw-Hill, New York, p. 567.
2. Asano, T. (1985). Overview: Artificial recharge of groundwater. In: *Artificial Recharge of Groundwater*. T. Asano (Ed.). Butterworth Publ., New York, pp. 3–19.
3. Oaksford, E.T. (1985). Artificial recharge: methods, hydraulics and monitoring. In: *Artificial Recharge of Groundwater*. T. Asano (Ed.). Butterworth Publ., New York, pp. 69–127.
4. Singh, R. (2002). Building community leaders for groundwater resources management—working on people's priorities. In: *Proceedings of Intl. Conference on Sustainable Development and Management of Groundwater Resources in Semi-Arid Region with Special Reference to Hard Rocks*. M. Thanagarajan, S. N. Rai, and V. S. Singh (Eds.). Oxford & IBH Publ. Co., New Delhi, India, pp. 526–533.
5. Katyal, J.C., Singh, R.P., Sharma, S., Das, S.K., Padmanabhan, M.V., and Mishra, P.K. (1995). *Field Manual on Watershed Management*. Central Research Institute for Dryland Agricultural, Hyderabad, India, pp. 1–87.
6. Wilderer, P.A., Forstner, U., and Kuntschik, O.R. (1985). The role of river bank filtration along the Rhine river for municipal and industrial water supply. In: *Artificial Recharge of Groundwater*. T. Asano (Ed.). Butterworth Publ., New York, pp. 509–528.
7. Piet, G.J. and Zoedeman, B.C.J. (1985). Bank and dune infiltration of surface water in the Netherlands. In: *Artificial Recharge of Groundwater*. T. Asano (Ed.). Butterworth Publ., New York, pp. 529–540.
8. Rai, S.N. (2002). Groundwater flow modeling. In: *Dynamics of Earth's Fluid Systems*. S.N. Rai, D.V. Ramana, and A. Manglik (Eds.). A. A. Balkema Publ., Holland, pp. 49–62.
9. Pinder, G.F. and Gray, W.G. (1977). *Finite Element Simulation in Surface and Subsurface Hydrology*. Academic Press, New York, p. 295.
10. Cabral, J.J.S.P., Wrobel, L.C., and Brebbia, C. (1990). Boundary element analysis of unconfined flow in porous media using B-splines. In: *Proceedings of the 8th Intl. Conference on 'Computational methods in water resources, Part B*. Venice, pp. 405–411.
11. Rao, N.H. and Sarma, P.B.S. (1980). Growth of groundwater mound in response to recharge. *Groundwater*. **18**: 587–595.
12. Marino, M.A. (1967). Hele-Shaw model study of the growth and decay of groundwater ridges. *J. Geophys. Res.* **72**: 1195–1205.
13. Baumann, P. (1965). Technical development in groundwater recharge. In: *Advances in Hydroscience*. Vol. 2. V.T. Chow (Ed.). Academic Press, New York, pp. 209–279.
14. Ramana, D.V., Rai, S.N., and Singh, R.N. (1995). Water table fluctuation due to transient recharge in a 2-D aquifer system with inclined base. *Water. Resour. Manag.* **9**: 127–138.
15. Mercer, J.W. and Faust, C.R. (1980). Groundwater modeling; Mathematical models. *Groundwater*. **16**: 212–227.
16. Rai, S.N., Ramana, D.V., and Singh, R.N. (1998). On the prediction of ground water mound formation in response to transient recharge from a circular basin. *Water. Resour. Manag.* **12**: 271–284.
17. Todd, D.K. (1980). *Groundwater Hydrology*, 2nd Edn. John Wiley & Sons, New York, p. 535.
18. Rushtun, K.R. (2003). *Groundwater Hydrology: Conceptual and Computational Models*. John Wiley & Sons, New York, p. 407.
19. Carslaw, M.S. and Jaeger, J.C. (1959). *Conduction of Heat in Solids*. Oxford Univ. Press, Oxford, UK, p. 510.
20. Polubarinova-Kochina, P.Ya. (1962). *Theory of Groundwater Movement*. Princeton University Press, Princeton, NJ, p. 613.
21. Abramowidz, M. and Stegun, I.A. (1970). *Handbook of Mathematical Functions*. Dover, New York, p. 1046.
22. Ozisik, M.N. (1980). *Heat Conduction*. John Wiley & Sons, New York, p. 687.
23. Sneddon, I.N. (1974). *The Use of Integral Transforms*. Tata McGraw Hill, New Delhi, India.
24. Lee, T-C. (1999). *Applied Mathematics in Hydrogeology*. Lewis, Boca Raton, FL, p. 382.
25. Bruggeman, G.A. (1999). *Analytical Solution of Geohydrological Problems*. Elsevier, New York, p. 959.
26. Warner, J.W., Molden, D., Chehata, M., and Sunada, D.K. (1989). Mathematical analysis of artificial recharge from basins. *Water Resour. Bull.* **25**(2): 401–411.
27. Baumann, P. (1952). Groundwater movement controlled through spreading. *Amer. Soc. Civ. Eng. Trans.* **117**: 1024–1074.
28. Hantush, M.S. (1967). Growth and decay of groundwater mounds in response to uniform percolation. *Water Resour. Res.* **3**(1): 227–234.
29. Glover, R.E. (1961). *Mathematical Derivations as Pertain to Groundwater Recharge*. Agricultural Research Service, USDA, Ft. Collins, CO, p. 81.
30. Hunt, B.W. (1971). Vertical recharge of unconfined aquifers. *J. Hydraul. Div. ASCE*. **97**(HY7): 1017–1030.
31. Rao, N.H. and Sarma, P.B.S. (1981). Groundwater recharge from rectangular areas. *Groundwater*. **19**: 271–274.
32. Singh, V.P. (1989). *Hydrologic Systems—Watershed Modelling*. Vol. 2. Prentice Hall, Englewood Cliffs, NJ, p. 320.
33. Detay, M. (1995). Rational groundwater reservoir management, the role of artificial recharge. In: *Artificial Recharge of Groundwater II*. A.I. Johnson and R.D.G. Pyne (Eds.). ASCE, New York, pp. 231–240.

34. Mousavi, S.F. and Rezai, V. (1999). Evaluation of scraping treatment to restore infiltration capacity of three artificial recharge projects in central Iran. *Hydrogeol. J.* **7**: 490–500.
35. Zomorodi, K. (1991). Evaluation of response of a water table to a variable recharge rate. *Hydrol. Sci. J.* **36**: 67–78.
36. Dagan, G. (1966). *Linearized Solutions of Free-Surface Groundwater Flow with Uniform Recharge*. Technion Publication No. 84, Technion, Israel Institute of Technology, Tel Aviv, Israel.
37. Rai, S.N. and Singh, R.N. (1979). Variation of water table induced by time varying recharge. *Geophys. Res. Bull.* **17**(2): 97–109.
38. Rai, S.N. and Singh, R.N. (1980). Dynamic Response of an unconfined aquifer system subjected to transient recharge. *Geophy. Res. Bull.* **18**(2): 49–56.
39. Rai, S.N. and Singh, R.N. (1981). A mathematical model of water table fluctuations in a semi-infinite aquifer induced by localised transient recharge. *Water Resour. Res.* **17**(4): 1028–1032.
40. Abdulrazzak, M.J. and Morel-Seytoux, H.J. (1983). Recharge from an ephemeral stream following wetting front arrival to water table. *Water Resour. Res.* **19**: 194–200.
41. Rai, S.N., Manglik, A., and Singh, R.N. (1994). Water table fluctuation in response to transient recharge from a rectangular basin. *Water Resour. Manag.* **8**(1): 1–10.
42. Rai, S.N. and Singh, R.N. (1995). An analytical solution for water table fluctuation in a finite aquifer due to transient recharge from a strip basin. *Water Resour. Manag.* **9**(1): 27–37.
43. Rai, S.N. and Singh, R.N. (1996). On the prediction of ground water mound formation due to transient recharge from a rectangular area. *Water. Resour. Manag.* **10**: 189–198.
44. Rai, S.N. and Singh, R.N. (1998). Evolution of the water table in a finite aquifer due to transient recharge from two parallel strip basins. *Water. Resour. Manag.* **12**: 199–208.
45. Manglik, A., Rai, S.N., and Singh, R.N. (1997). Response of an unconfined aquifer induced by time varying recharge from a rectangular basin. *Water Resour. Manag.* **11**(3): 185–196.
46. Rai, S.N., Ramana, D.V., and Manglik, A. (1997). Modeling of water table fluctuation in finite aquifer system in response to transient recharge. In: *Proceeding of Intl. Symposium of Emerging Trend in Hydrology*. D. C. Singhal et al. (Eds.). pp. 243–250.
47. Rai, S.N., Ramana, D.V., Thiagarajan, S., and Manglik, A. (2001). Modelling of groundwater mound formation due to transient recharge. *Hydrological Processes* **15**(8): 1507–1514.
48. Manglik, A. and Rai, S.N. (1998). Two-dimensional modelling of water table fluctuations due to time-varying recharge from rectangular basin. *Water Resour. Manag.* **12**: 467–475.
49. Rai, S.N. and Manglik, A. (1999). Modelling of water table variation in response to time varying recharge from multiple basins using the linearized Boussinesq equation. *J. Hydrol.* **220**: 141–148.
50. Manglik, A. and Rai, S.N. (2000). Modelling of water table fluctuation in response to time varying recharge and withdrawal. *Water Resour. Manag.* **14**(5): 339–347.
51. Rai, S.N. and Manglik, A. (2000). Water table variation due to time varying recharge and withdrawal. In: *Groundwater: Past Achievement and Future Challenges*. Oliver Silio et al. (Eds.). Proc. of XXX IAH Congress, Minneapolis, MN, pp. 259–262.
52. Rai, S.N. and Manglik, A. (2001). Modelling of water table fluctuations due to time-varying recharge from canal seepage. In: *New Approaches Characterizing Groundwater Flow*. K.-P. Seiler and S. Wohnlich (Eds.). Balkema Pub., The Hague, The Netherlands, pp. 775–778.
53. Manglik, A., Rai, S.N., and Singh, V.S. (2004). Modeling of aquifer response to time varying recharge and pumping from multiple basins and wells. *J. Hydrol.* **292**: 23–29.

GROUNDWATER AND ARSENIC: CHEMICAL BEHAVIOR AND TREATMENT

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

INTRODUCTION

Arsenic is not an abundant element in the earth's crust; the average crustal concentration of arsenic is 1.8 mg/Kg (1) to 5 mg/Kg (2). Arsenic ranks as the 52nd element in abundance, between tin and molybdenum (1). However, through geogenic processing of crustal materials, arsenic can be concentrated in soils to a typical range of 2 to 20 mg/Kg (3) to 1 to 50 mg/Kg (2), with concentrations as high as 70 mg/Kg being unremarkable (3). Human activity generates anthropogenic arsenic, which makes it the third most common regulated inorganic contaminant found at U.S. Superfund sites.

Arsenical copper was in use by 4000 BC, and the toxic effects of arsenic were documented by early Greek writers. More recently, arsenic has been linked to skin, bladder, and other cancers (4). The U.S. Environmental Protection Agency (USEPA) lowered the arsenic standard in drinking water from 50 μg/L to 10 μg/L, effective January 23, 2006.

Modern usage of arsenic includes formulation of pesticides and herbicides, decolorization of glass, paint manufacturing, the production of semiconductors, and the treatment/preservation of wood. Pressure treated lumber was commonly treated for decades using copper-chromium-arsenate (CCA). This product, also called "green wood," has been used for foundation lumber and more recently as wood for outdoor children's play structures and picnic tables. The CCA wood is being phased out for toxicity concerns and environmental reasons. Many of the pressure treatment lumber facilities have significant soil and groundwater contaminated with arsenic as well as chromium.

CHEMICAL CHARACTER

Although arsenic occurs in more than 20 minerals, only a few are commonly found in ore deposits (5). Arsenic may occur as a semimetallic element (As^0), arsenate (As^{5+}), arsenite (As^{3+}), or arsine (As^{3-}). The biogeochemistry of arsenic involves adsorption, biotransformation, REDOX reactions, and precipitation-dissolution processes (6,7).

The chemical character of arsenic is labile and readily changes oxidation state or chemical form through chemical or biological reactions that are common in the environment. Therefore, rather than solubility equilibrium controlling the mobility of arsenic, it is usually controlled by REDOX conditions, pH, biological activity, and adsorption/desorption reactions. Arsenic in groundwater most often occurs from geogenic sources, although anthropogenic arsenic pollution does occur. Geogenic arsenic is almost exclusively an arsenite or arsenate. The most oxidized pentavalent form, arsenic, forms oxyanions ($H_3ASO_4{}^-$, $H_2ASO_4{}^-$, H2ASO4-, $HASO_4{}^{2-}$, $ASO_3{}^-$). These arsenic oxyanions are isomorphous with oxyanions of phosphorous, substituting for phosphate in both marine organisms and phosphate deposits (4). Arsenite is the trivalent form that also forms a series of oxyanions that change specific configuration and charge with pH. Of critical importance with regard to the controls of the mobility of arsenite is the fact that at a pH of 9.5 or lower, the arsenite oxyanion is not charged. This result obviates all ionic interactions of the species.

Common arsenic minerals are arsenopyrite (FeAsS), enargite (Cu,AsS), proustite (Ag,AsS), and lollingite ($FeAs_2$). Late-stage magmatic crystallization (pyrometasomatic and hydrothermal stages) contributes to arsenic-rich sulfides. In sedimentary rocks, arsenic is commonly found adsorbed onto fine-grained sedimentary rocks, such as iron and manganese oxides (4). According to the U.S. Geological Survey, arsenic concentrations in sedimentary iron-ores range from 65 to 650 mg/Kg (8). Arsenic is also associated with sedimentary pyrite at concentrations of 100 to 77,000 mg/Kg (6).

Anthropogenic arsenic may have any form including organic arsine species. Groundwater in acidic to intermediate volcanic rocks, or in sediments derived from those rocks, will often have arsenic concentrations exceeding 50 μg/L.

Figure 1 illustrates the difference in molecular structure between arsenate and arsenite. The double bond oxygen in the arsenate molecule influences its ability to become ionized through the of hydrogen ions. The process is termed dissociation. A negative charge develops on the arsenate molecule when dissociation occurs. The double bond oxygen increases the capacity to delocalize that charge, which cases the loss of hydrogen ions. The propensity for ionization is expressed by the constant of dissociation, pKa. The pKa value, which is a negative log, shows a greater degree of dissociation with a smaller value.

For arsenate and arsenite, pKa values are as follows:

$$\text{Arsenate- } H_3ASO_4 \quad pK_1 = 2.19 \quad pK_2 = 6.94$$

$$pK_3 = 11.5$$

$$\text{Arsenite- } H_3ASO_3 \quad pK_1 = 9.20 \quad pK_2 = 14.22^*$$

$$pK_3 = 19.22^*$$

The pH at which these ionization steps occur is significantly different between arsenate and arsenite, as illustrated in Fig. 2 (10,11). Figure 2 also shows the control of REDOX potential (Eh) on the arsenate/arsenite transition. This Eh/pH relationship is key in understanding arsenic mobility in groundwater and the effectiveness of arsenic water treatment systems.

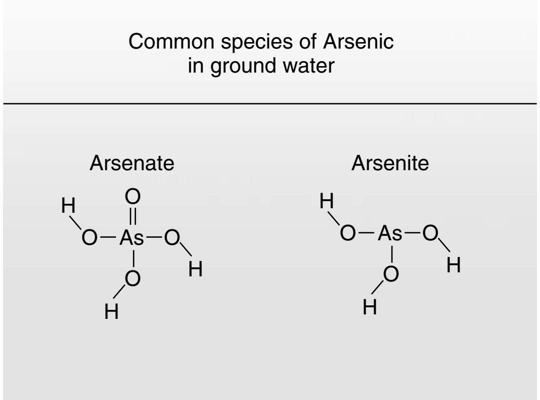

Figure 1. Difference in molecular configuration of arsenate and arsenite.

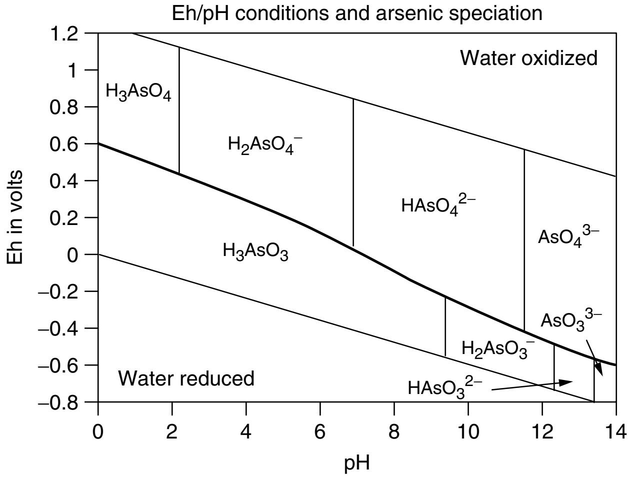

Figure 2. Control of arsenic speciation by Eh and pH conditions.

Arsenic Immobilization

The previous section described the conditions under which arsenic can become an ionized species. The most commonly recognized adsorption reactions are based on ion exchange between charged adsorption sites and charged soluble ions. However, London Van der Waals bonding is another mechanism that is also responsible for adsorption. This type of bonding is the result of complex interactions among the electron clouds of molecules, molecular polarity, and attractive forces of an atomic nucleus for electrons beyond its own electron cloud. Consequently, some degree of immobilization can occur with soluble species that are not ionized. Arsenic immobilization through ionic adsorption can be controlled within normal oxidizing Eh/pH conditions. London Van der Waals bonding is complex to the point of unpredictability except for arsenic

---

*These pKa values are extrapolated from the strength of oxygen acid rules (9).

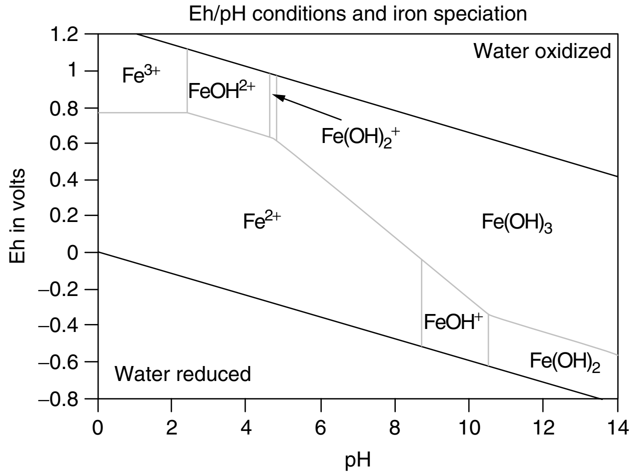

Figure 3. Iron speciation as controlled by Eh/pH conditions.

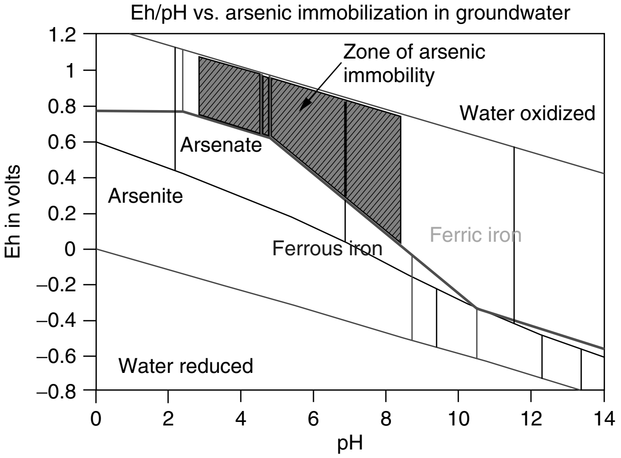

Figure 4. Arsenic mobility in groundwater as controlled by the effect of Eh/pH conditions on the speciation of arsenic and iron.

mobility at extreme Eh/pH conditions that can be obtained in industrial settings, but not in groundwater.

Components of soil that participates in both types of adsorptive reactions include clays, carbonaceous material, and oxides of iron, aluminum, and manganese. In the most shallow soils, the organic fractions typically dominate, whereas at greater depths, iron oxyhydroxides play the principal adsorptive role.

The typical iron content of soil ranges from 0.5% and 5%. Not only is iron common, but as with arsenic, it is also labile and readily reflects changes in surrounding Eh/pH conditions. This relationship for iron is illustrated in Fig. 3 (12).

Ferric hydroxide acts as an amphoteric ion exchanger. Depending on pH conditions, the ferric hydroxide has the capacity for cation or anion exchange. Given the average iron concentration in soil and soluble arsenic concentrations in groundwater at 50 μg/L, ferric hydroxides in sediment can potentially adsorb 0.5 to 5 pounds of arsenic per cubic yard of aquifer matrix, which may then act as a significant potential reservoir for arsenic release under changing Eh/pH conditions.

ARSENIC IN GROUNDWATER

Arsenic concentrations up to 12,000 μg/L have been reported for the St. Peter aquifer in eastern Wisconsin (4). In this case, the oxidation of arsenic sulfides in a sulfide cement horizon (SCH) within the aquifer is a source of the high arsenic concentrations.

Figure 4 superimposes the Eh/pH relationship for the arsenic and iron systems; it illustrates the conditions under which arsenic will be immobilized in a groundwater system. Of equal importance, it illustrates how arsenic adsorbed to ferric hydroxides in sediment can be released at exposure to groundwater that is chemically reducing. Two effects would be at work: Arsenate is reduced to arsenite that will not remain ionically bound to the geologic substrate, and ferric iron is reduced to ferrous, which is soluble under normal pH conditions. Outside the immobilized zone, arsenic mobility is variable. London Van der Waals bonding of arsenite is in effect, but it is not sufficient to assure complete immobilization.

WATER TREATMENT SYSTEMS

Introduction

Following is a brief review of various technologies used for the removal of arsenic from drinking water and industrial wastewater. Table 1 summarizes the effectiveness of each and gives the source for the information.

Arsenite Oxidation

As previously described, the ionization chemistry of arsenic in groundwater precludes the removal of arsenite by ion exchange within normal pH ranges. Other technologies including coprecipitation, electrodialysis, and reverse osmosis are also affected by arsenite's dissociation profile. One solution to this problem is an oxidation step to form arsenate. Figure 5 illustrates the Eh/pH range required for this process. Oxidation, particularly of drinking water, may be problematic. Chemical residues of the oxidant, byproducts from oxidation of other organic or inorganic species, reagent costs, and operational issues are all factors. Oxygen would be ideal, as it is thermodynamically capable of this oxidizing step. The kinetics for oxidizing arsenic compounds in groundwater are exceedingly slow (22). It is possible to use gas diffusion technologies that slowly release dissolved oxygen into aquifers and that have demonstrated the capacity to convert anaerobic groundwater systems into aerobic systems within 3 to 6 months (23). This process will convert soluble ferrous iron to insoluble ferric iron oxides capable of attracting arsenate to their surfaces. The oxidation of arsenite is complex and may take additional time or the presence of other abiotic or biological (24) stimulants. In Bangladesh, *in situ* concentrations of arsenic less than 0.1 mg/L were readily removed by the oxygenation of groundwater; concentrations greater than that had only 50% removal (25).

Other chemicals can affect arsenite oxidation including free chlorine, hypochlorite, ozone, permanganate, and hydrogen peroxide with ferrous iron.

Table 1. Effectiveness of Arsenic Water Treatment Methods

Treatmenttechnology	Initial Arsenic Concentration	Final Arsenic Concentration*	Reference
Iron Coprecipitation	Oxidized Arsenic 56 mg/L	10 μg/L	(13)
	Arsenate 350 μg/L	6 μg/L	(14)
	Arsenite 350 μg/L	140 μg/L	(14)
	Arsenate 560 μg/L	10 μg/L	(15)
	Arsenate 300 μg/L	6 μg/L	(16)
	Arsenite 300 μg/L	138 μg/L	(16)
Alum Coprecipitation	Arsenate 350 μg/L	74 μg/L	(14)
	Arsenite 350 μg/L	263 μg/L	(14)
	Arsenate 300 μg/L	30 μg/L	(16)
	Arsenite 300 μg/L	249 μg/L	(16)
Lime Precipitation	Arsenate 500 mg/L	4 mg/L	(17)
	Arsenite 500 mg/L	2 mg/L	(17)
	Arsenate 2 mg/L	20 μg/L	(17)
	Arsenite 2 mg/L	160 μg/L	(17)
Activated Alumina	Arsenate 100 μg/L	4 μg/L	(18)
	Arsenite 100 μg/L	Ineffective	(18)
	Arsenate 57 μg/L	ND at pH 6.0	(19)
	Arsenite 31 μg/L	Ineffective	(19)
Ion Exchange	Arsenate 100 mg/L	<800 μg/L	Vance
	Arsenate 68 mg/L	12.2 to 0 mg/L	(20)
Reverse Osmosis	Arsenite 37 μg/L	7 μg/L	(19)
	Arsenate 51 μg/L	1.5 μg/L	(19)
Electrodialysis	Arsenite 188 μg/L	136 μg/L	(19)
Sulfide Precipitation	Arsenate 132 mg/L	26 mg/L	(21)
Activated Carbon	Arsenite 500 μg/L	300 μg/L	(21)

*The current limit for drinking water is 50 μg/L.

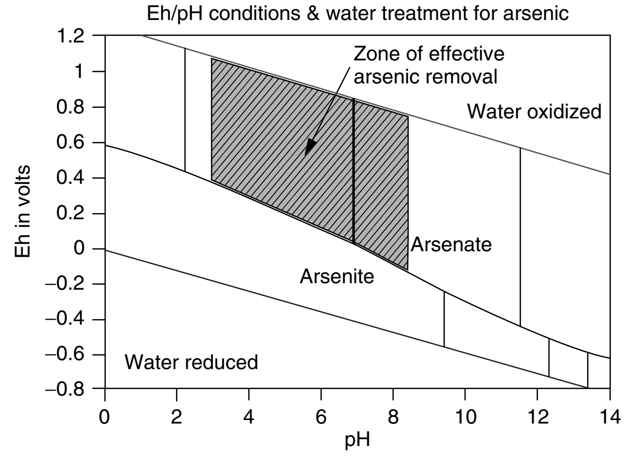

Figure 5. Eh/pH range required for effective treatment of soluble arsenic.

Iron Coprecipitation

Coprecipitation of arsenate with ferric iron is recognized as overall the most effective and practical existing method of arsenic removal. Ferric iron coprecipitation is particularly useful in the mining industry, where large amounts of ferric iron and arsenic can be byproducts of production or refining. Adding ferric iron salts for the treatment of drinking water is usually necessary. Figure 4 shows that arsenate is readily removed by iron coprecipitation and the Eh/pH conditions that must be maintained to effect that removal. Because of London Van der Waals bonding, ferric iron coprecipitation of arsenite is also moderately effective, with 50% removal at a pH of 7.0 (26).

The use of iron hydroxides for the coprecipitation of arsenic in industrial wastewater (in which arsenic is in the mg/L range) requires iron dosage four to eight times higher than that of the soluble arsenic; a greater iron dosage yields no further benefit (27).

Alum Coprecipitation

Suspended aluminum salts (alum) can remove arsenate via mechanisms similar to those for ferric hydroxides. However, it is less effective over a narrower pH range for arsenate removal and is ineffective for removal of arsenite (16).

Lime Precipitation

In testing by Nishimura and Tozawa (17), removal of arsenic with lime precipitation was feasible, but not necessarily practical. High concentrations of arsenic were removed to concentrations of 2 and 4 mg/L for arsenate and arsenite, respectively. Removal to lower levels required a second treatment step in which initial arsenic concentrations of 2 mg/L were lowered to 20 μg/L for arsenate and 160 μg/L for arsenite. However, to achieve these removal efficiencies, the lime dosage was between 5 and 15 g/L.

Activated Alumina

Activated alumina can be effective for the removal of arsenate under moderately acidic pH conditions

(5.5 to 6.0). High levels of competing anions such as sulfate significantly reduce the effectiveness of activated alumina (18). The use of activated alumina for complete arsenite removal is ineffective because of the nonionic character of arsenite in that pH range (see Fig. 1). Some initial arsenite removal is observed because of London Van der Waals bonding, but compared with arsenate adsorption, this capacity is rapidly exhausted.

Ion Exchange

Ion exchange has the potential for soluble arsenic removal. Anion exchange resins are available in two basic forms, weak base and strong base. Many weak-base anion exchangers are capable of significant adsorption because of London Van der Waals bonding in addition to ion exchange, which gives them a higher level of adsorptive capacity for nonionic arsenite. The author has evaluated anion exchange resins for arsenate removal. The most effective activation was in the hydroxyl form. Chloride and acetate were also tested. Weak base resins had higher loading capacities than did strong base (6% vs. 4.8%), but they did not have adequate removal efficiencies (75% vs. 99+% for the strong base resin). Anion exchange resins are also prone to chromatographimg because of the presence of competing anions in the treated water. However, ion exchange is an area of intense research where the development of anionic chelating exchange resins or ion exchange polymers may dramatically improve the technology for arsenic treatment.

Reverse Osmosis

Reverse osmosis (RO) has been shown to have a removal efficiency greater than 97%. Electrodialysis was only 73% effective. When used to treat 100% arsenite, removal was only 28% (19).

Other

Rosehart (21) evaluated a series of removal technologies including activated carbon and sulfide precipitation. Neither performed at a level adequate for use in the treatment of drinking water.

CONCLUSIONS

The behavior of arsenic in groundwater and industrial wastewater is dominated by REDOX and pH conditions. Under a limited range of specific Eh/pH conditions, the ability to predict total immobility of arsenic in groundwater and in water treatment systems exists (see Figs. 4 and 5). Except for those conditions, arsenic will be partially mobile, the magnitude of which is difficult to predict.

Implications of this behavior include:

- Arsenic treatment without control of Eh/pH is likely to be ineffective.
- The injection of water in a reduced oxidation state into sediments with adsorbed arsenic may cause a dramatic mobilization of arsenic into the groundwater.
- If ferric iron sludges used for the coprecipitation of arsenic are disposed under improper Eh/pH conditions, arsenic will remobilize. *In situ* remediation via recovery or stabilization of arsenic contaminated groundwater should be focused on Eh control through chemical or biological methods

Site-specific arsenic chemistry including source, mobility, migration, fate, and transport is complex. Any *in situ* treatment of groundwater containing arsenic should be evaluated carefully with laboratory bench tests, computer groundwater modeling, and field pilot tests before full-scale remediation is attempted.

BIBLIOGRAPHY

1. Demayo, A. (1985). Elements in the Earth's crust. In: *CRC Handbook of Chemistry and Physics*, 66th Edn. R.C. Weast (Ed.). CRC Press Inc., Boca Raton, FL, p. F145.
2. Lindsay, W.L. (1979). *Chemical Equilibria in Soils*. John Wiley & Sons, New York.
3. Yan Chu, H. (1994). Arsenic distribution in soils. In: *Arsenic in the Environment, Part I: Cycling and Characterization*. J.O. Nriagu (Ed.). John Wiley & Sons, New York, pp. 17–50.
4. National Research Council (NRC). (1999). *Arsenic in Drinking Water*. National Academy Press, Washington, DC.
5. Johnson, J.A. and Schreiber, M. (2004). *Arsenic: Perspectives Regarding a New Environmental Concern, The Professional Geologist*. American Institute of Professional Geologists, Westminster, CO, pp. 41–44.
6. Smedley, P.L. and Kinniburgh, D.G. (2002). A review of the source, behaviour and distribution of arsenic in natural waters. *Appl. Geochem.* **17**(5): 517–568.
7. Welch, A.H., Westjohn, D.B., Helsel, D.R., and Wanty, R.B. (2000). Arsenic in ground water of the United States: Occurrence and geochemistry. *Ground Water* **38**(4): 589–604.
8. Gualtieri, J.L. (1973). United States Mineral Resources. In: *Geological Survey Professional Paper #820*. U.S. Geological Survey, Reston, VA, pp. 51–61.
9. Pauling, L. (1970). *General Chemistry*. Dover Publications, Mineola, NY.
10. Robins, R.G. (1985). The aqueous chemistry of arsenic in relation to hydrometallurgical processes, impurtity control and disposal. *Proceedings of CIM Metallurgical Society, 15th Annual Hydrometallurgical Meeting*. pp. 1–1–1-26.
11. Welch, A.A., Lico, M.S., and Hughes, J.L. (1988). Arsenic in groundwater of the Western United States. *Groundwater* **26**(3): 333–347.
12. Hem, J.D. (1961). Stability field diagrams as aids in iron chemistry studies. *J. Am. Water Works Assoc.* **53**(2): 211–232.
13. Brewster, M.D. (1992). Removing arsenic from contaminated wastewater. *Water Treatment Technol.* **November**: 54–57.
14. Logsdon, G.S., Sorg, T.J., and Symons, J.M. (1974). Removal of Heavy Metals by Conventional Treatment. *Proceedings of the 16th Water Quality Conference*. University of Illinois, Urbana, IL, pp. 11–133.
15. Shen, Y.S. (1973). Study of arsenic removal from drinking water. *J. Am. Water Works Assoc.* **651**: 543–548.

16. Sorg, T.J. and Logsdon, G.S. (1978). Treatment technology to meet the interim primary drinking water regulations for inorganics, Part 2. *J. Am.Water Works Assoc.* **70**: 379–393.
17. Nishimura, T. and Tozawa, K. (1985). Removal of arsenic from wastewater by addition of calcium hydroxide and stabilization of arsenic-bearing precipitates by calcination. *Proceedings of CIM Metallurgical Society, 15th Annual Hydrometallurgical Meeting*. pp. 3–1–3-18.
18. Frank, P. and Clifford, D. (1986). *Arsenic (III) Oxidation and Removal from Drinking Water*. U.S. EPA, EPA-600-52-86/021.
19. Clifford, D. and Lin, C-C. (1991). *Arsenic(III) and Arsenic(V) Removal from Drinking Water in San Ysidro, New Mexico*. U.S. EPA Proj. Sum., EPA/600/S2-91/011.
20. Calmon, C. (1973). Comment. *J. Am. Water Works Assoc.* **651**: 568–569.
21. Rosehart, R. and Lee, J. (1972). Effective methods of arsenic removal from gold mine wastes. *Can. Min. J.* **June**: 53–57.
22. Clifford, D., Ceber, L., and Chow, S. (1983). Arsenic(III)/Arsenic(V) separation by chloride- form ion-exchange resins. XI. *Am. Water Works Assoc. Water Qual. Tech. Conf.*, Norfolk, VA.
23. Kolhatkar, R. (2002). Stable isotope analyses to demonstrate MTBE biodegradation in ground water. *Groundwater Resources Association of California MTBE Symposium*, San Jose, CA, Oct. 17, Abstracts/presentation.
24. Phillips, S.E. and Taylor, M.L. (1976). Oxidation of arsenite to arsenate by *Alcaligenes* faecalis. *App. Env. Microbiol.* **32**(3): 392–399.
25. Sarkar, A.R. and Rahman, O.T. (2001). In-situ removal of arsenci—experiences of DPHE-Danida Pilot Project. In: *Technologic for Arsenci Removal from Drinking Water*. M. Amhed, A. Feroze, M. Ashraf, and Z. Adeel (Eds.). pp. 201–206.
26. Pierce, M.L. and Moore, C.B. (1982). Adsorption of arsenite and arsenate on amorphous iron hydroxide. *Water Res.* **16**: 1247–1253.
27. Krause, E. and Ettel, V.A. (1985). Ferric arsenate compounds: Are they environmentally safe? Solubilities of basic ferric arsenates. *Proceedings of CIM Metallurgical Society, 15th Annual Hydrometallurgical Meeting*. pp. 5–1–5-20.

TREATMENT OF ARSENIC, CHROMIUM, AND BIOFOULING IN WATER SUPPLY WELLS

WALTER W. LOO
Environmental & Technology Services
Oakland, California

JASON J. WEN
City of Downey
Downey, California

INTRODUCTION

Arsenic

Arsenic has been used as a component of pesticides and thus may enter streams or groundwater through waste disposal or agricultural drainage. Arsenic is also present in volcanic gases and is a common constituent of geothermal or spring water.

As small amounts of arsenic can be toxic to humans, it is considered a highly undesirable impurity in water supplies; an upper concentration limit of 50 μg/L, established in 1976 by the U.S. EPA, has now been lowered to 10 μg/L.

Arsenic may form metal arsenides in which its oxidation state is negative (1). Arsenic may also form sulfides and can be present as an accessory element in sulfide ore deposits. In solution in water, the stable forms are arsenate (As^{5+}) or arsenite (As^{3+}) oxyanions. A pH–E_h diagram showing fields of dominance of aqueous arsenates (1) indicates that the monovalent arsenate anion $H_2AsO_4^-$ is expected to predominate between pH 3 and 7 and the divalent species $HAsO_4^{2-}$ takes over from pH 7 to 11. Mildly reducing conditions favor the uncharged arsenite ion $HAsO_2$ (aq).

Chromium

Chromium is an inorganic element that is used in electroplating, leather tanning, wood treatment, pigment manufacture, and cooling tower treatment for corrosion control. Chromium can contaminate drinking water sources through discharges from industries, leaching from hazardous waste sites, or it may occur naturally from the erosion of natural deposits. Two forms of chromium can occur in water sources: chromium (III) and chromium (VI). Chromium (III) is an essential nutrient at trace concentrations. Chromium (VI) is toxic and is the basis for setting the chromium drinking water standard. The ratio of the two forms can vary quite a bit in natural waters. Evidence also exists to suggest that chromium (VI) may be converted to chromium (III) in the human body, particularly in the acidic digestive system.

The U.S. Environmental Protection Agency (EPA) classified chromium (VI) as a human carcinogen by inhalation. In 1991, it reviewed the existing standard for total chromium and the most recent scientific research available. As a result, the EPA actually raised the maximum contaminant level from 50 to 100 parts per billion as total chromium, based on its conclusion that chromium (VI) is not carcinogenic by ingestion.

In California, the drinking water standard is 50 parts per billion for total chromium [the sum of chromium (III) and chromium (VI)]. A public health goal is a "risk assessment" of the concentration of a contaminant in drinking water that poses no significant health risk to the consumer. The recommended public health goal determined by the Cal/EPA Office of Environmental Health Hazard Assessment (state health experts) for total chromium is 2.5 parts per billion.

Well Biofouling

Iron and manganese biofouling (usually associated with "iron bacteria") are common in water supply wells and attached appurtenances. Although biofouling sometimes has no overt symptoms, it may cause clogging, corrosion, and water quality degradation. These problems pose a considerable challenge to water utilities and well owners in North America and around the world. Precise estimates of increased operational and mitigation costs resulting from damage and loss of efficiency from well biofouling are not available. Nevertheless, it is believed to be widespread, costly, underreported, and not adequately remedied in actual practice, resulting in frequent operating problems. Although preventive maintenance would be more cost-effective, the most common approach to iron biofouling problems in wells is crisis management. Preventive maintenance is seldom employed, however, because suitable maintenance monitoring methods and practical protocols for detecting iron and manganese biofouling problems before they severely affect well production and water quality have not been available (2).

Understanding the cause of water well deterioration and developing ways to sustain water well environments is important in maintaining and improving the quality of life in rural areas. The deterioration of well yield and water quality is a concern to individuals, small communities, and industries that rely on water wells as their principal source of water. Currently, when the quality or quantity of water produced declines dramatically, wells are often abandoned, or treatments are applied with little understanding of the cause of these problems. The cost of replacing these wells can have a significant economic impact on well owners. Correctly identifying the cause of water well problems offers the possibility of effective treatment and maintenance instead of well abandonment. Losses in water well production and water quality have traditionally been attributed to the chemical and physical properties of the water well environment. Many of these problems can be solved by well-established diagnostic and rehabilitative techniques. However, less recognized is that groundwater contains microorganisms such as bacteria, and the activities of these microorganisms also cause significant water well problems. Water well deterioration caused by microbiological activity is termed biofouling. Installing and pumping a well increases the level of oxygen and nutrients in the well and in the surrounding aquifer, encouraging bacterial cells, which are naturally present in groundwater, to anchor themselves to surfaces in the well and around the well intake. Once attached, these bacteria quickly multiply and colonize these surfaces. The bacterial colonies form a gel-like slime or biofilm that captures chemicals, minerals, and other deposits, such as clays and silts, moves to the well during pumping, and forms biomasses.

Some of the byproducts of bacterial growth, such as oxidized iron and manganese, also accumulate in these secretions, which leads to the production of the red or black slimes often found in toilet tanks or observed on pumps and discharge lines when they are pulled from a well. Biofouling of a water well occurs when biofilms accumulate a sufficient amount of debris to interfere with water flow and affect water quality. If uncontrolled, well biofouling can affect well performance in various ways. Biofilms and the debris they collect can quickly coat, harden, and plug the well screen, the sand pack, the surrounding aquifer material, and may even plug water lines and affect the performance of household treatment systems. In addition, the bacteria living within the biofilm can increase the rate of iron oxidation and iron buildup in the well and distribution pipes, which leads to occasional discoloration of well water. Biofouling can also result in the production of odors such as rotten egg or fishy smells, changes in taste, and corrosion of steel and iron casings and pipes. Once developed, a biomass can protect the bacterial cells from environmental changes such as changes in pH, temperature, and fluid velocities, making treatment chemicals less effective and removal of plugging material more difficult, which emphasizes the importance of regular well maintenance. A number of field and laboratory tests exist that can be used to monitor water quality and biological activity in groundwater. If performed regularly 1 month after the well is installed and then once every 6 months, these tests indicate when water quality is changing or when biological activity is increasing. Changes in water quality and increased levels of biological activity indicate that well maintenance is required. Ideally, appropriate well maintenance chemicals should be applied before well performance is significantly affected. Establishing a monitoring schedule, where pumping water levels and well pumping rates are recorded, is also an effective way to identify when preventive maintenance measures are no longer effective and well rehabilitation is required.

Extracellular slimes are composed largely of polysaccharides and, in general, are the major component of the biofouling mass. Biomass associated with viable active cells is a relatively minor component. Figure 1 illustrates how the maximum level of biological activity commonly occurs under redox conditions that are at the periphery of oxidation in the E_h range of -50 to $+150$ mV. Biofouling can be complex and caused by a variety of bacteria. However, the appearance and odor of bacterial slimes are diagnostic.

ARSENIC TREATMENT

Of the 14 treatment technologies that the EPA reviewed, five are relevant technologies for small systems—ion exchange, activated alumina, and membrane technologies (reverse osmosis, nanofiltration, and electrodialysis reversal). Seven alternative technologies are categorized as still emerging (iron oxide-coated sand, granular ferric hydroxide, iron filings, sulfur-modified iron, greensand filtration, iron addition with microfiltration, and conventional iron/manganese removal). The last two technologies—coagulation/filtration and lime softening—are used primarily in larger systems and are not expected to be installed solely for arsenic removal.

Ion Exchange

Ion exchange, in particular, will probably be a very common technology used to comply with the arsenic regulation. It is recommended for systems with low sulfate (<120 mg/L) and total dissolved solids (TDS). The effect of competing ions drives the regeneration frequency and, in turn, the cost. Ions that compete with arsenic are sulfate (the most significant competitor), fluoride, selenium, and nitrate. Systems that have high levels of these contaminants may need a pretreatment phase as well. The EPA has data on co-occurrence but would like to hear how much competition is occurring in the field and how

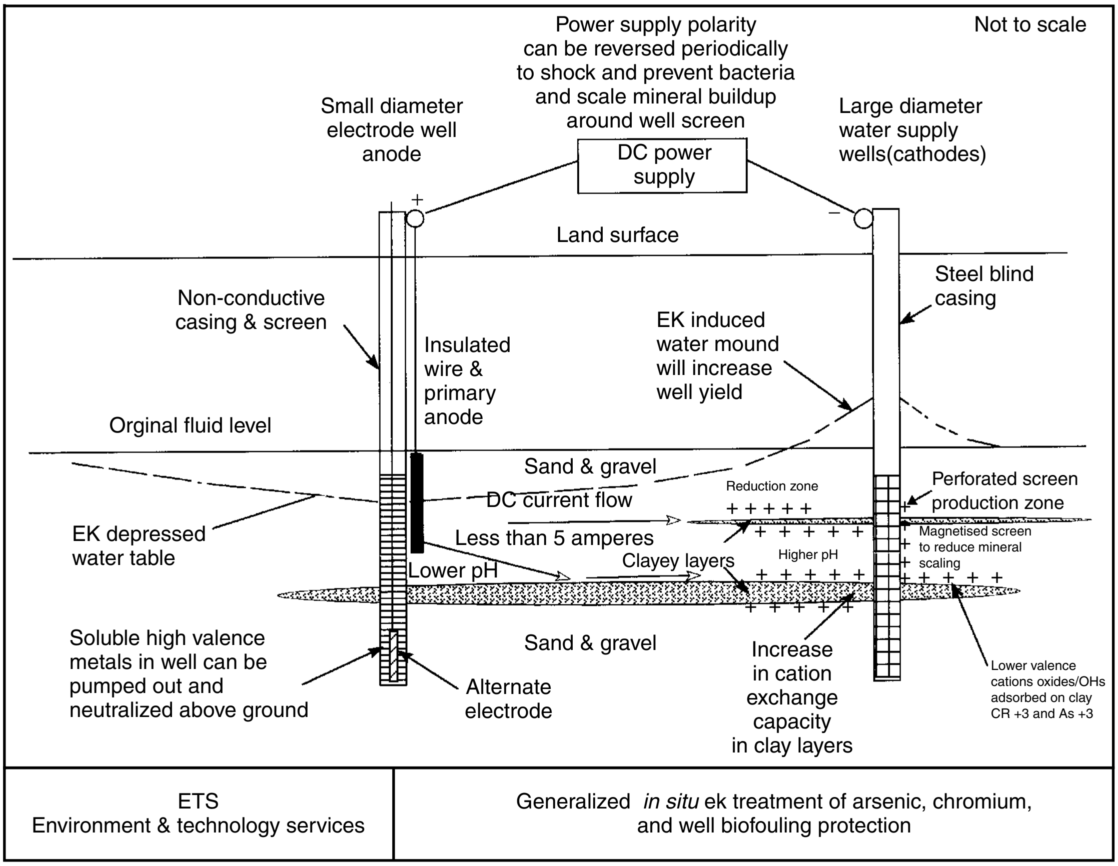

Figure 1. Generalized *in situ* EK treatment of arsenic, chromium, and well biofouling protection.

viable people think this technology would be. The waste stream, or brine, can be reused, which reduces the volume of waste and increases its concentration. The following are frequently asked questions about ion exchange:

1. How often is regeneration feasible? Will it vary by size category?
2. Can the EPA cap sulfate at 120 mg/L? Where is the TDS cap?
3. How often will an ion exchange system need regeneration? How long does it take to regenerate? Can you afford to have two systems to switch back and forth during regeneration? How much back-up water supply is necessary to provide water during regeneration? Do you have that much storage capacity available? How much regeneration is feasible for small systems? Daily, once a week, once a month? Can we expect variation by size?
4. Should we calculate costs for different removal percentages assuming full stream treatment or blending for ion exchange? Can all sources be treated at one location? Do different wells/streams serve different parts of the system?

Coagulation/filtration and lime softening are intended for larger systems. If small systems install these technologies, the EPA expects it would be in package plant form to reduce costs. A package plant is one bought "off-the-shelf" versus one custom designed for a site. A well-trained operator is needed to run these technologies; an off-site operator could utilize remote telemetry. The EPA does not expect these technologies to be installed only for arsenic removal, but if there is another contaminant in the water, it may be practical. Sludge disposal needs to be considered and may be an issue for small systems. Again, the EPA would like information from those experienced with these types of technologies.

All the technologies looked at and discussed so far work best when the arsenic is in the form of arsenic (V). **Pretreatment** converts arsenic (III) to arsenic (V). Surface water tends to favor arsenic (V), but groundwater tends to contain arsenic (III). Data shows that chlorine and potassium permanganate are effective in oxidizing arsenic (III) to arsenic (V). Possible problems include the existence of chlorine, which increases the potential to create disinfection byproducts, and membrane fouling of subsequent treatments such as RO. The EPA's Office of Research and Development is researching other preoxidants (including ozone and hydrogen peroxide, which are expected to be effective) to provide more data in the next few months.

Point of use (POU)/Point of entry (POE) devices may be appropriate for small systems serving 10,000 or fewer and are new elements of the SDWA. POE is whole-house treatment, whereas POU treats water at the tap. The EPA is looking at these devices as possible compliance technologies to meet the new MCL for small systems. Some of the nontransient, noncommunity systems with large flows of water might use POE and POU devices to treat the minor part of the flow provided for potable use. The POE/POU technologies that are available for arsenic removal are smaller versions of reverse osmosis, activated alumina, and ion exchange. Note that POU/POE technologies must be maintained by the public water system. Therefore, the need exists for substantial recordkeeping. It also increases the responsibility on the part of customers, as it requires them to facilitate entry into their homes by the utility for maintenance. The water utility is ultimately responsible for ensuring that these devices are maintained properly. The EPA is trying to determine the system size cutoff where centralized treatment would be more affordable than POU and POE devices.

Waste Disposal will be an important issue for both large and small plants. If a plant is located inland and uses membrane technologies, operators may have to pretreat prior to discharge. If the plant is discharging to a sanitary sewer because of the membranes, there may be very high salinity in the discharge as well as high levels of arsenic that might be above local sewer regulations. Ion exchange, reverse osmosis, and activated alumina treatment brines will be even more concentrated (on the order of 30,000 TDS) and more than likely will require pretreatment prior to discharge to either a receiving body of water or the sanitary sewer.

CHROMIUM TREATMENT

Reverse osmosis (RO) systems can often improve the quality of water. Reverse osmosis water treatment has been used extensively to convert brackish water or seawater to drinking water, to clean up wastewater, and to recover dissolved salts from industrial processes. It is becoming more popular in the home market as homeowners become increasingly concerned about contaminants that affect their health, as well as about nonhazardous chemicals that affect the taste, odor, or color of their drinking water. People considering the installation of a water treatment system to reduce toxic chemicals should first have their water tested to determine how much, if any, hazardous compounds are in the water. Public water supplies are routinely monitored and treated as required under the federal Safe Drinking Water Act and state regulations. Private water systems should be tested at the owner's initiative based on knowledge of land use and contamination incidents in the area (3).

Reducing Contaminants Through RO

Reverse osmosis reduces the concentration of dissolved solids, including a variety of ions and metals and very fine suspended particles such as asbestos, that may be found in water. An RO device may be installed following a water softener to reduce the concentration of sodium ions exchanged for hardness ions. RO also removes arsenic (As), chromium (Cr), nitrate (NO_3^-), certain organic contaminants, some detergents, and specific pesticides.

The Reverse Osmosis Process

In reverse osmosis, a cellophane-like membrane separates purified water from contaminated water. An understanding of osmosis is needed before further describing RO. Osmosis occurs when two solutions containing different quantities of dissolved chemicals are separated by a semipermeable membrane that allows only some compounds to pass through. The osmotic pressure of the dissolved chemical causes pure water to pass through the membrane from the dilute to the more concentrated solution. In reverse osmosis, water pressure applied to the concentrated side forces the process of osmosis into reverse. Under enough pressure, pure water is "squeezed" through the membrane from the concentrated to the dilute side. Salts dissolved in water as charged ions are repelled by the RO membrane. Treated water is collected in a storage container. The rejected impurities on the concentrated side of the membrane are washed away in a stream of wastewater, not accumulated as on a traditional filter. The RO membrane also functions as an ultrafiltration device, screening out particles, including microorganisms, that are physically too large to pass through the membrane's pores. RO membranes can remove compounds in the 0.0001 to 0.1 micron size range (thousands of times smaller than a human hair).

Design of an RO System

Although reverse osmosis is simple, a complete water treatment system is often complex, depending on the quality of the incoming water before treatment and the consumer's needs. Most home RO systems are point-of-use (POU) units placed beneath the kitchen sink to treat water for cooking and drinking. Point-of-entry (POE) systems that treat all water entering the household are more expensive to purchase and operate than POU systems.

A typical home reverse osmosis system consists of pretreatment and posttreatment filters as well as an RO membrane, flow regulator, storage container for the treated water, and dispensing faucet. The pressure for RO is usually supplied by the feed line pressure of the water system in the home, but a booster pump may be needed to produce an adequate volume of treated water. A sediment prefilter is essential for removing relatively large sand grains and silt that may tear or clog the RO membrane or clog a pump or flow regulator. Water softeners are used in advance of the RO system when household water is excessively hard. If the water is chlorinated or contains other oxidizing chemicals such as bromine, an activated carbon prefilter is needed to protect membranes sensitive to these chemicals.

To remove certain pesticides and organic solvents, activated carbon (AC) posttreatment must be included in the system. A standard AC filter positioned after the storage tank removes compounds that cause unpleasant taste and odors, including those from the tank or plastic tubing, just before water is dispensed. To remove high levels of organic chemicals such as trihalomethanes, volatile organic chemicals, and chloramines, an additional prolonged contact carbon filter (PCCF) is placed between the RO membrane and the storage tank. Combining an activated carbon filter with RO expands the range of chemicals the system can remove. Furthermore, AC treatment is improved because RO removes compounds that adversely affect AC adsorption.

The storage tank, tubing, and dispensing faucet should be made of plastic, stainless steel, or other nontoxic materials. The low pH and mineral content of RO-treated water may corrode copper pipes and allow lead to leach into the drinking water from brass components.

RO Membrane Materials

The most common RO membrane materials are polyamide thin film composites (TFC) or cellulosic types [cellulose acetate (CA), cellulose triacetate (CTA), or blends]. Very thin membranes are made from these synthetic fibers. Membrane material can be spiral-wound around a tube, or hollow fibers can be bundled together, providing a tremendous surface area for water treatment inside a compact cylindrical element. Hollow fiber membranes have greater surface area (and therefore greater capacity) but are more easily clogged than the spiral-wound membranes commonly used in home RO systems.

The flux, or capacity, of the RO membrane indicates how much treated water it can produce per day. Typically, RO membranes for home systems are rated in the range of 10 to 35 gallons per day. Thus, under standard operating conditions, it could take from 2 to 6 hours to fill a two and-a-half-gallon storage tank. CA/CTA membranes have adequate capacity for most households, but TFC membranes should be used if large volumes of treated water are needed.

RO membranes are rated for their ability to reject compounds from contaminated water. A rejection rate (% rejection) is calculated for each specific ion or contaminant as well as for reduction of total dissolved solids (TDS). It is important that consumers know their specific requirements for water quality when buying a system. For example, high rejection rates are essential when high nitrates or lead concentrations in the water must be brought below the EPA maximum contaminant or action levels.

Efficiency of RO Systems

The performance of an RO system depends on membrane type, flow control, feed water quality (e.g., turbidity, TDS, and pH), temperature, and pressure. The standard at which manufacturers rate RO system performance is 77 °F, 60 pounds per square inch (psi), and TDS at 500 parts per million (ppm). Only part of the water that flows into an RO system comes out as treated water. Part of the water fed into the system is used to wash away the rejected compounds and goes down the drain as waste. The recovery rate, or efficiency, of the system is calculated by dividing the volume of treated water produced by the volume of water fed into the system. If not properly designed, RO systems can use large quantities of water to produce relatively little treated water. Most home RO systems are designed for 20% to 30% recovery (i.e., 2–3 gallons of treated water are produced for every 10 gallons put into the system). Home RO systems can operate at higher recovery rates but doing so may shorten membrane life. The flow regulator on the reject stream must be properly adjusted. If the flow is slow, the recovery rate is high, but RO membranes are easily fouled if concentrated impurities are not washed away quickly enough. If the flow is too fast, the recovery rate is low and too much water goes down the drain. Overall water quality affects the efficiency of an RO system and its ability to remove specific contaminants. The higher the TDS, the lower the recovery rate of treated water. The amount of treated water produced decreases by 1% to 2% for every degree below the standard temperature of 77 °F. An RO system supplied with well water at a temperature of 60 °F produces only three-quarters of the volume it would produce at 77 °F. For an RO system to function properly, there must be enough water pressure. Although most home RO systems are rated at 60 pounds per square inch, the incoming feed line pressure of many private water systems is less than 40 psi. The RO system must work against back pressure created in the storage tank as it fills with water and compresses the air in the tank. The RO device must also overcome osmotic pressure, bonding between water molecules, and dissolved impurities; the higher the TDS level, the greater the osmotic pressure. The net water pressure at the RO membrane can be calculated by subtracting back pressure and osmotic pressure from feed line pressure. If the net water pressure at the membrane is lower than 15 psi, treated water production is less efficient and contaminant rejection rates are lower. Auxiliary pumps can be added to the treatment system to boost pressure and improve the quality and quantity of water produced. High-quality RO systems have valves that shut off the flow whenever storage tank pressure reaches two-thirds of the feed pressure; at that point, low net water pressure can result in low rejection rates. In some systems, once the storage tank is filled, surplus treated water is discarded; water loss from such units is frequently excessive. A system that automatically shuts off when the pressure on the tank reaches a given level saves water.

Maintenance of an RO System

An RO system must be well maintained to ensure reliable performance. Clogged RO membranes, filters, or flow controls decrease water flow and systems performance. If fouling is detected in early stages, the membrane can often be cleaned and regenerated. The cleaning procedure varies depending on the type of membrane and fouling. Completely clogged or torn RO membranes must be replaced. In addition, pre- or postfilters must be replaced once a year or more often, depending on the volume of water fed through the system and the quality of the feed water. Damage to RO membranes cannot be easily seen. The treated water must be analyzed periodically to determine whether the membrane is intact and doing its job. Many systems now have a built-in continuous monitor that indicates a high TDS level, a sign that the system is not operating properly. It may also be necessary to test regularly for specific health-related contaminants such as nitrates or lead. Microorganisms, dead or alive, can clog RO membranes. To prevent biofouling, RO units must be disinfected periodically with chlorine or other biocides provided by the manufacturer. Continuous chlorination can be used with cellulosic membranes to protect the system from biofouling and eliminate the particle-trapping slime that worsens other forms of fouling such as scaling. Chlorine and other oxidizing disinfectants are harmful to thin film composite membranes. If the feed water is chlorinated, an activated carbon unit must be used to remove the oxidizing chemicals before they reach the TFC membrane. Activated carbon (AC) prefilters should not be used on nonchlorinated water supplies because they provide a place for microorganisms to multiply and lead to increased biofouling of the RO membrane surface. It is important to replace AC filters periodically following the manufacturer's instructions, especially after an extended shutdown period during which microorganisms can flourish.

Choosing an RO System

Homeowners who are thinking about buying reverse osmosis systems should determine their initial water quality and their goals in adding water treatment systems. RO removes many inorganic impurities from drinking water, especially nitrate. Its effectiveness depends not only on the type of membrane but on feed water quality, temperature, pressure, and flow control, as well as the type and concentration of specific contaminants to be removed. A typical RO system consists of a sediment filter, pump, reverse osmosis membrane, flow regulator, storage tank, final activated carbon filter (for taste and odors), and dispensing faucet. An AC prefilter is sometimes needed for dechlorination. RO is commonly used to treat only the water used for drinking and cooking at the point of use rather than at the point of entry for all household use. RO membrane types vary in their ability to reject contaminants and differ in capacity (the volume of treated water produced per day). Water pressure is an important factor in determining the RO system's rejection rate, capacity, and recovery rate (amount of treated water produced per amount of feed water used). Maintenance of an RO system is essential for reliable performance. High levels of TDS and microorganisms in the system are commonly the cause of fouled membranes. Treated water should be monitored for TDS and the level of any specific contaminants that may affect health.

A list of home water treatment devices certified by various Department of Health Services can be found on their websites.

BIOFOULING TREATMENT

Historically, there have been three approaches to the declining operation of a water well of any common type. One is to simply abandon the well and install a new well of similar or greater capacity to replace the abandoned well. A second involves attempting to change the operating techniques (e.g., pump times, volumes, sequences of up- and downtimes, and control flow by drawdown limitations) or change some components in the well (e.g., pump, screen) in the hope that the well will recover from the observed problem. The third is to attempt an analytical approach to the problem by determining

—first, the cause;

—second, confirm that the effects witnessed can be related to the cause identified;

—and third, determine and apply a treatment strategy that counteracts the cause and allows the well to function as designed.

Increasing economic and environmental costs and concerns are now restricting the ability of a well user simply to replace a failing well. Economic concerns relate to the increasing costs involved in well replacement and the growing sensitivity for maximizing the use of each well installation by extending its useful life (i.e., environmental sustainability). Environmental concerns are being brought to the fore because groundwaters are no longer seen as an infinite resource. In some areas, aquifers are now being heavily depleted by the demand, and there is little flexibility to provide additional capacity. Another major environmental concern is the impact of various forms of pollution on well fields. In the past decades, general attitudes may be summarized by an "out of sight, out of mind" approach in which groundwater was given a lower status than surface waters. Various chemical leakages from industry, agriculture, and various service industries were not considered as important as those in surface waters. When a pollutant impacted surface water, the effects could often be relatively quickly appreciated through radical eutrophication, deteriorating water quality, and water unacceptable to users. One major difference between surface and groundwaters is the fact that the former flows as large unconfined masses whereas the latter moves as a confined mass within porous media. This difference is very critical to the current understanding of groundwater flow and quality.

It is not easy to appreciate the complex interactions that occur between flowing groundwater and the media through which it is passing as it moves to a well, a spring, or interfaces with another aquifer. For the last century, it has been popularly believed that groundwater is essentially sterile (devoid of biological activity) and that all activities within an aquifer may be explained almost exclusively by a combination of physical and chemical processes. Today, the hydrology of groundwater systems still leans heavily on this assumption. Through the science of subsurface microbiology (the study of microorganisms in the crust of the planet), it is now becoming increasingly evident that groundwater movement and quality are affected by microbiological interactions. In the past decades, these have been ignored, and one of the major consequences has been that the effects of these microorganisms as biological filters (interface) have been ignored. Pollutants within a groundwater system may become entrapped (and possibly degraded) within these biological filters and so not appear in the groundwater resurfacing through a well. Environmental monitoring of the product (postdiluvial, after the "event") water from a well may not necessarily give an "accurate" picture of the chemical loading in the transient (causal) water itself. There has been a tendency for groundwater users to rely on product ("biofiltered") water for environmental assessment, and yet this water may not accurately allow a risk assessment for that well (due to bioentrapment of some chemicals of concern).

In the next two decades, the realization of the nature of the biological interfaces within and around water wells may cause much tighter environmental constraints to be placed on new well installations, which would mean that greater attention would be paid to extending the service life of existing wells through preventive maintenance and effective rehabilitative programs. The mindset that a water well is a physical object set within a chemical and physical world has to change. This mindset has generated a "traditional" attitude that a dysfunctional well is simply a result of chemically driven corrosion, encrustation, clogging processes, or the physical collapse of the system (through such events as "silting up" and "collapsed" aquifer and well structures).

Acidization has commonly been applied as a remediation technique to dissolve and disperse the clogs and encrustations, and various disinfectants (such as different formulations of chlorine) were used to control any coliform and other bacteria that may be growing down the borehole (and presents a potential health risk). Slime formations were considered by many to be simply physical-chemical accumulates that may result in clogging, encrustation, and corrosion. Even today, camera logging a water well is considered sufficient to view all biological and much of the chemical deposits (e.g., silts and salts) that can be causing problems around a well. Please see the section on the "Preliminary Diagnosis of Biological Fouling of Water Wells Using TV Camera Logging Methods."

Combinations of disinfectants, selected acids and even, in more recent times, dispersants (a.k.a. wetting agents) have become part of the arsenal of weapons used to rehabilitate problems in a well. One of the findings from these actions has been that "no one size fits all" and that each well should be treated as unique and requires customization of the treatment parameters to optimize maintenance practices. This approach stems from observations that each well can be characterized as different from other wells in the same field. Many instances exist where two wells of the same construction and characterization placed within feet (meters) of each other in supposedly the same aquifer formation bear very different characteristics. An unfortunate result of this is that a treatment may be successfully applied to one well in a field but that same treatment may fail on a neighboring well of exactly the same characteristics in construction, operation, and mode of failure (4).

INNOVATIVE *IN SITU* TREATMENT

In situ electrokinetic treatment of chromium, arsenic, and biofouling may provide a cost-effective solution (5). The electrokinetic treatment process involves applying direct current (dc) in a medium (soil and water). The flow of electrons from anode to cathode creates a migration of cations in the medium toward the cathode. The electrolysis of water creates a higher pH and oxidizes the metal or reduces its valence, which renders the metal into a nontoxic form near the cathode. Clayey material may show a dramatic increase in cation exchange capacity under a high pH near the cathode.(6–8) The proposed electrokinetic process can be applied both *in situ* and *ex situ*. Figure 1 presents a more realistic distribution of the various valence states of metals under the proposed *in situ* electrokinetic influence.

Besides the treatment of metals, the beneficial side effects of the *in situ* EK treatment are as follows:

- The well itself is set up as a cathode, so iron bacteria will not live on the surface of the well casing and perforation because of the high pH.
- The electrokinetically induced water migration toward the cathode (well) may induce an increase in hydraulic head, thus increasing well yield.

No such treatment has been tried on chromium, arsenic, and manganese in a saturated medium to date. There are many successes of the proposed EK process in the laboratory and in soil. However, we have successfully demonstrated the electrokinetic control of selenium and boron in clayey saturated media at two sites in the Panoche Irrigation District, Central Valley, California. We believe that the same EK control can be applied to chromium, arsenic, and biofouling because of the similarity of the multiple valence forms to those of selenium and boron. Due to the simplicity of the cathode and anode setup, we believe that the proposed EK processes will prove cost-effective. The proposed *in situ* EK treatment (once set up) is permanent. It has a one-time capital cost and minimum long-term maintenance costs. The continuous operating electricity demand will not exceed 50 amperes at 30 to 100 Vdc or 1500 to 5,000 watts per site.

BIBLIOGRAPHY

1. Hem, J.D. (1979). Reaction of metal ions at surfaces of hydrous iron oxide. *Geochim. Cosmochim. Acta* **41**: 527–538.
2. Smith, S.A. (1992). *Methods for Monitoring Iron and Manganese Biofouling in Water Wells*. AWWA Research Foundation Publication Number 90617.
3. Kneen, B., Lemley, A., and Wagenet, L. (1995). *Reverse Osmosis Treatment of Drinking Water, Fact Sheet 4*. Cornell Cooperative Extension, New York State College of Human Ecology.
4. Cullimore. (1997).
5. Lehr, J.H. (Ed.). (2001). *Standard Handbook of Environmental Health, Science and Technology*. Chapters 14.4 & 14.6 on Bioremediation & Electrokinetic Treatment of Hazardous Wastes by Walter W. Loo, McGraw-Hill, New York.
6. Driscoll, F.G., (1986). *Groundwater and Wells—Well Failure & Iron Bacteria*. Johnson Filtration Systems, Inc., St. Paul, MN, pp. 646–655.
7. Hem, J.D. (1989). *Study and Interpretation of the Chemical Characteristics of Natural Water*. U.S. Geological Survey Water-Supply Paper 2254.
8. Walter, D.A. (1997), *Geochemistry and Microbiology of Iron-Related Well-Screen Encrustation and Aquifer Biofouling in Suffolk County, Long Island, New York*. U.S. Geological Survey Water-Resources Investigations Report 97-4032.

ARTESIAN WATER

David P. Commander
Water and Rivers Commission
East Perth, Australia

Artesian water is groundwater that flows freely at the surface from an artesian bore or well. The name artesian comes from the flowing wells constructed in the twelfth century at the Carthusian monastery near Lilliers in the northeastern French province of Artois.

The term "artesian" has also been extended to refer to artesian aquifers, in which the groundwater is under a sufficient pressure head to flow at the surface, and to artesian basins, sedimentary basins that may contain artesian aquifers. The term subartesian refers to the situation where the water level rises above the top of the aquifer but is below ground surface.

The artesian condition is, however, an accident of topography, not necessarily a property of the water, the bore, or the aquifer. For this reason, it is preferable now to use the term "confined" to refer to the aquifer. "Confined" implies the presence of a low permeability confining bed above the confined aquifer, which allows a positive (above ground) hydraulic head to be maintained. Rarely, flowing water can be obtained from an unconfined aquifer without a confining bed, provided there is a sufficient upward head gradient.

Artesian water is common in sedimentary basins, where there are extensive aquifers overlain by confining beds. The classic artesian basin is shaped like a saucer; the aquifer crops out around the elevated margins of the basin, where it can be recharged by rainfall or stream flow (see Fig. 1). The cities of London and Paris lie at the centers of their basins and were among the first cities to exploit artesian water on a large scale. The Great Artesian Basin in eastern Australia is one of the largest examples; it is 1500 km across. However, artesian conditions can occur in other geological formations.

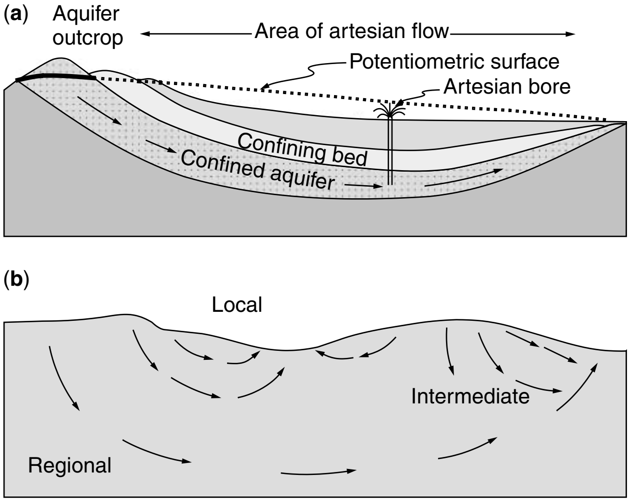

Figure 1. Groundwater flow in a classical artesian basin (**a**) contrasting with cross-formational flow in local, intermediate, and regional flow systems in a nonclassical basin (**b**).

The phenomenon of free-flowing groundwater was first documented sometime in the first millennium B.C. at Dakhla Oasis in the western desert of Egypt, during the course of progressively deepening hand-dug wells. The first to explain the correct nature of artesian wells were Mahammad Karaji (died 1016) and al Biruni (973–1048), but in Europe, the explanation had to await an understanding of the hydrologic cycle. In 1691, Bernadino Ramazzini (1633–1714) described the artesian wells that were commonplace in the Italian city of Modena. He drew a geologic section, although it appears he envisaged underground rivers rather than porous strata as the water source. In 1715, Antonio Vallisnieri (1661–1730), President of the University of Padua, reasoned that rainfall on the mountains was the source of the artesian water at Modena and illustrated his work with geologic sketches of the Alps made by Johann Scheuchzer.

The existence of artesian water began to be widely appreciated in Europe in the early 1800s, after French engineers introduced Chinese drilling technology. The success of the Grenelle artesian bore near Paris, in which water rose to a height of 33 meters and flowed at the rate of 4000 cubic meters per day stimulated the development of artesian water. In the absence of suitable pumps, the availability of free-flowing water was an important economic factor in opening up new lands for agriculture and pastoral use in the Americas and Australia. The existence of the Great Artesian Basin in Eastern Australia was postulated in 1881. Proved by drilling in 1887, it rapidly led to pastoral development. The recognition of waterborne diseases and the contamination of shallow water supplies in cities was also an incentive to developing artesian supplies of clean drinking water in the Eastern United States, Europe, and Australia by the turn of the century.

Such was the impact of free-flowing water on the public that drilling for artesian water was often carried out in geologically unfavorable places (and despite professional advice); the perception was widely held that artesian water would be encountered, providing the boring was deep enough.

Geologic textbooks of the mid- to late 1800s illustrated the structure of classical artesian basins such as the Paris and London Basins, although it was recognized by the end of the century that flowing wells could also occur under different conditions. A series of papers by U.S. Geological Survey authors provided detailed explanations of artesian water. Thomas Chamberlin's classic paper (1), which is recognized as the beginning of the science of hydrogeology in North America, systematically laid down the conditions under which artesian flows would be obtained: the presence of a pervious stratum, impermeable beds above and below, an inclination of the beds so that the edge is higher than the surface at the well, and adequate rainfall.

Chamberlin's observations can now be qualified somewhat. Few, if any, confining beds are now considered completely impermeable, and in many places, the confining beds yield significant amounts of water by leakage to adjacent artesian aquifers. Tóth (2) has drawn together ideas on regional hydraulic continuity, recognizing cross-formational flow in which it is not necessary to have inclined beds to produce artesian conditions. In many artesian basins, the water is of great antiquity, a million and a half years in the case of the discharge from the Australian Great Artesian Basin, although current rainfall is necessary to maintain the hydraulic head in recharge areas.

In 1928, Oscar Meinzer's paper on the elasticity of artesian aquifers drew together much of the work from the intervening period, demonstrating the concept of elastic storage, whereby flowing water is released by the reduction in hydraulic head and compressibility of the aquifer matrix (3).

However, there were many who denied the conventional theory (held in the main by Geological Surveys) that the origin of the water was meteoric and that the pressure was related to the elevation of the water table in outcrop areas of artesian aquifers. In Australia, distinguished academic geologists such as J. W. Gregory (1906) and Alexander du Toit (1917) invoked juvenile and connate water, rock pressure, and gas bubbling to explain the origin of water and pressure. These ideas were supported by American geologists whose oil-field experience led to the concept of closed and hydraulically isolated systems.

There is still debate surrounding the explanation for the origin of artesian pressure, stimulated by studies of nonclassical basins such as the Hungarian Basin, and this has led to a better understanding of basin hydrodynamics.

Studies of the Great Artesian Basin clearly show the hydraulic continuity of aquifers over large distances and the observed potentiometric heads that decline gently away from the recharge areas. Furthermore, isotopic age dating using chlorine-36 indicates a transit time across the basin of 1.5 million years that correlates with hydrodynamic modeling.

Artesian water is released from elastic storage, and as the pressure is reduced, the aquifer matrix bears more of the hydrostatic pressure. Thus, subsidence occurs in most situations where the compressibility of the aquifer is much greater than that of water itself. Artesian flows may also lead to irreversible consolidation of aquifers, particularly in geologically recent environments that are not fully consolidated. In large artesian systems such as the Great Artesian Basin, free-flowing water is effectively mined from elastic storage, as there is little potential to induce more recharge on the timescale of abstraction. Maintenance of artesian head is the priority for management, given the economic consequences once bores cease to flow. In the Great Artesian Basin, a program is underway to cap and control the flow of artesian bores to maintain the aboveground pressure and to prevent the waste of water, which had been allowed to flow many kilometers in open "bore drains" to water livestock.

Owing to the great depths from which it may come, artesian water is often much warmer than the shallow groundwater and may have to be cooled before use. A feature of artesian bores is that the head of water varies according to the temperature. A bore may cease to flow after the water is shut in and the water column cools. Artesian flows may also cease after degassing in a shut-in bore. Artesian flows are also affected by barometric pressure changes.

Artesian water may also be relatively highly mineralized, owing to the long travel time, and contain gases and dissolved iron, which may precipitate at the surface. The words may, therefore, have a bad connotation in terms of drinking water.

BIBLIOGRAPHY

1. Chamberlin, T.C. (1885). *Requisite and Qualifying Conditions of Artesian Wells*. U.S.G.S. Annual Report No. 5, pp. 131–141.
2. Tóth, J. (1963). A theoretical analysis of groundwater flow in small drainage basins. *J. Geophys. Res.* **68**(16): 4795–4812.
3. Meinzer, O.E. (1928). Compressibility and elasticity of artesian aquifers. *Econ. Geol.* **23**: 263–291.

MODELING CONTAMINANT TRANSPORT AND BIODEGRADATION IN GROUNDWATER

Sascha E. Oswald
UFZ Centre for Environmental Research
Leipzig-Halle, Germany

INTRODUCTION

What Is a Groundwater Contaminant?

When dealing with contaminant transport modeling, we have to be clear from the beginning that almost any substance could be a contaminant. It all depends on the nature of the object which may be at risk and in which concentrations the substance is present. This object often is the human being, and sometimes a special one, like a baby, but sometimes also other living organisms, or even a habitat or ecosystem that may be endangered. The concentration, or more precisely the intake, can determine whether a substance is harmless or even an essential nutrient, or harmful, for example, in case of chloride or zinc. That is the toxicological perspective. From a chemical perspective, there are organic and inorganic contaminants; from a physical perspective, there are dissolved and particulate or free-phase contaminants, and biologically, there are a variety of animals, microorganisms and viruses, or nonliving substances, that could be harmful and thus could be contaminants. In the following, we focus only on the more familiar type of contaminant in groundwater, that is an organic, dissolved, non living substance, and refer the reader otherwise to more comprehensive descriptions of groundwater contaminants.

Why Model Contaminant Transport?

Contaminants released into the subsurface may be spread by transport in groundwater to locations further away. That is why assessing the transport of a contaminant may be useful. A typical objective may be to protect the groundwater resource itself from contamination or to protect somebody who could be affected by the contaminated groundwater. The latter case is the basis for the source—pathway—receptor approach to risk assessment, which judges the impact of a contaminant by its pathway from a source to a particular receptor that will be potentially harmed by the contaminant. Modeling is only one way to assess contaminant transport, but can describe a complex situation in detail and thus may allow predictions.

Why Model Biodegradation?

Biodegradation is an important factor in the fate of organic contaminants and may reduce the concentration and also the total mass of a contaminant. Although several processes can contribute to the natural attenuation of contaminants, biodegradation is often the main process by which mass may be destroyed; other processes such as sorption merely retain the contaminant mass for later release. Therefore, modeling monitored as well as enhanced natural attenuation, usually has to include biodegradation modeling.

Processes

The key processes to be modeled are advection, dispersion, sorption, and degradation by microbes. However, further processes could also be relevant, for example, ion exchange, precipitation and dissolution of minerals, speciation, and surface-complexation.

Modeling Basis and Related Modeling (Flow, Transport, Hydrogeochemical)

The basis for transport modeling is flow modeling; transport modeling itself is the basis for modeling contaminant transport and biodegradation (see Fig. 1). Often in more general terms this is called reactive transport modeling, which includes transport plus biodegradation and/or hydrogeochemical modeling. To model a field scale problem of reactive transport appropriately, one should plan this as a set of modeling tasks building up to this challenging undertaking.

Scale and Type of Transport

The scale of the problem is an important factor when building a model because it influences the modeling concept and the modeling tools to be chosen. Scales may range from pore scale, for pore blocking by bacterial conglomerates, to regional scale, for agricultural contaminants. Here, we focus on an intermediate to subregional scale, such as a contaminant plume originating from a point source or a remediation operation on a site scale. The processes included in a contaminant transport model may be special transport processes such as fracture networks, multiple porosity behavior, or preferential flow paths, but we will limit the discussion to the standard, equivalent porous medium approach.

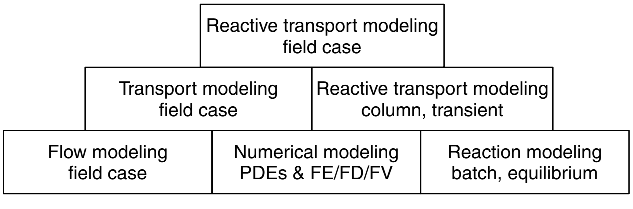

Figure 1. Key pyramid for achieving reactive transport modeling on a field scale.

THE BASIC CONCEPTS

A modeling project starts with the definition of its aims and purpose, and subsequently the modeling concept is built up. Conceptualization of the problem will probably require consideration of several aspects, especially in reactive transport modeling.

Chemistry

The chemical nature of the contaminant, the groundwater, and the aquifer may determine the fate of a contaminant to a high degree. The presence of electron acceptors that are usually needed for the degradation of organic contaminants, electron donors, enables certain reactions to occur in principle. The groundwater pH and redox condition determine which reactions actually can occur at a certain location. All of these may vary spatially and temporally. Also, the occurrence of other contaminants may affect the fate of a particular contaminant substantially. Furthermore, the presence of organic matter, charged surfaces, and minerals in the aquifer constitute pools for sorption of solute species. Altogether, this should be taken into account when defining the set of chemical species and the possible reactions.

It should also be mentioned that chemical reactions are treated as either equilibrium or kinetic reaction, depending on how their rates compare to advection and diffusion timescales and rates; equilibrium is faster than transport, and kinetic is slower. These two reaction types imply different resulting behavior, and also different types of equations, parameters, and numerical solution techniques; more details may be found in the literature on hydro(geo)chemical modeling.

Microbiology

Biodegradation implies that there is a biological activity by which degradation occurs, as opposed to chemical degradation. Populations of bacteria in aquifers constitute biota able to perform degradation in groundwater. Usually, there is a microbial community in an aquifer that has some diversity but is not really flourishing compared to soil or the biosphere. Both features are likely to be affected by biodegradation. The bacteria will try to biodegrade the contaminant when they can gain energy or carbon by this. The way they achieve the biodegradation is via enzymes, which they release or at least make available inside the cell. The kinetics of these enzymes can be described by the Michaelis–Menten formula, which is the starting point for equations used to describe biodegradation in groundwater.

In a constant microbial population feeding on a contaminant that is present in abundance as well as all other substances needed for the reaction, the rate of biodegradation is set by the enzymes available and is a constant in this quasi-steady-state scenario. This means the enzymes are saturated by the contaminant as a substrate, and this biodegradation kinetics is called *zero-order degradation*. Here, the degradation rate, that is, the change of the contaminant concentration C with time due to biodegradation, is a constant:

$$\frac{\partial C}{\partial t} = const. \tag{1}$$

However, often the contaminant, as the substrate for the bacteria, is not present in abundance, but its availability for the reaction is limited, similar to the concentration of the enzymes. Due to the mass action law, the reaction rate is proportional to the contaminant concentration. This results in a *first-order degradation* with first-order rate constant λ, where

$$\frac{\partial C}{\partial t} = \lambda C \tag{2}$$

Note that a radionuclide decays according to the same type of equation. In both cases, a half-life can be defined based on λ: $T_{1/2} = \ln 2/\lambda$.

The concentration of substrate may decrease during biodegradation, or perhaps the contaminant is not present in large quantities. Hence, a more realistic approach is to account for both former cases. A way of doing so is an empirical form of the Michaelis–Menten kinetics, called Monod degradation kinetics

$$\frac{\partial C}{\partial t} = \mu_{\max}\frac{C}{k + C} \tag{3}$$

where k is the half saturation constant and $\mu_{\max} = \lambda \cdot k$ is the maximum possible degradation rate constant. For small C, $C << k$, the degradation rate is proportional to C, whereas for large C, $C >> k$, it levels off to the constant rate $\mu_{\max}$, as in zero-order degradation.

An even more realistic concept takes into account that bacterial enzymes do not perform the degradation by themselves but merely facilitate a reaction between contaminant and another substance, usually an electron acceptor with concentration EA. The latter has to be present at the same place and time as the contaminant to allow the reaction to proceed. Then the degradation is described by the double Monod formulation:

$$\frac{\partial C}{\partial t} = \mu_{\max}\frac{C}{k_C + C} \cdot \frac{EA}{k_{EA} + EA} \tag{4}$$

And finally, to add some microbiology to it, bacterial populations can grow and die. Furthermore, they can be attached to grain surfaces and even create biofilms, or be transported with the water flow, or a mixture of both. This may affect the biodegradation very much, because the degradation rate is basically proportional to the amount of active bacteria in a unit volume, called X. This is taken into account when using the following equation, which constitutes *higher order degradation* behavior:

$$\frac{\partial C}{\partial t} = \mu_{\max}X\frac{C}{k_C + C} \cdot \frac{EA}{k_{EA} + EA}$$

and

$$\frac{\partial X}{\partial t} = const \cdot \frac{\partial C}{\partial t} \tag{5}$$

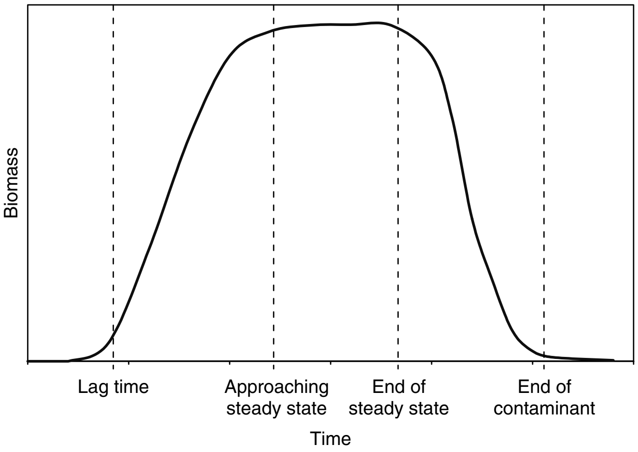

Figure 2. Example of the development of a bacterial population that is degrading a contaminant.

A typical case is that a certain type of bacteria is present in small quantities (see Fig. 2). This population starts to grow, when environmental conditions are favorable for them: there is a contaminant they can feed on and an electron acceptor they can use for the degradation reaction. It takes a particular time, until the size of the population reaches a level that makes a real difference in the contaminant concentration. Then its growth slows down, for example, when one of the substances for the degradation reactions becomes limiting. This could lead to a steady population of microbes, especially because bacteria grow but also die. Finally, when the environmental conditions change again, for example, by final consumption of the contaminant, the decay of bacteria is larger then their growth, and the population declines.

Sometimes the degradation product of a contaminant is a contaminant itself, and so on. Then a whole chain of degradation has to be taken into account, for example, in the biodegradation of PCE (1). Such a chain may be modeled on the basis of first-order degradation of each member of the chain. In this case, it is mathematically a process similar to chains of radioactive decay. However, it could also be modeled as a set of components, which are linked by the degradation-production term, and each could follow any of the mentioned degradation kinetics.

A further possible process is inhibition of a particular biodegradation process by the presence of relevant amounts of a certain substance, for example, oxygen, or of the contaminant itself, which would be called self-inhibition. To account for inhibition, an additional term can be added to the biodegradation equation, which reduces the degradation rate for this case.

A range of processes could also be included in modeling biodegradation, if they are relevant for the specific problem: (1) transport of dissolved species into a biofilm of bacteria; (2) bacterial attachment and detachment to the solid matrix; (3) sorbed contaminants may not be degraded, as long as they are bound to surfaces or organic matter; and (4) bacteria may even feed on a contaminant in a solid phase, on the other hand.

Geochemical and microbial processes are closely linked in groundwater. Microbial activity is based on the availability of substrates, electron acceptors, redox and pH conditions, yet influences all these conditions, and thus may initiate additional geochemical processes, such as precipitation of minerals. Therefore, a realistic and detailed model of biodegradation reactions in groundwater must also include a set of geochemical reactions.

Dispersion

Hydrodynamic dispersion may seem a minor issue relative to those discussed before. Nevertheless, dispersion at the plume fringe can be a key factor in biodegradation, even more than for a nonreactive contaminant. The velocity-based mechanical dispersion creates mixing between species necessary for reaction on a scale depending on longitudinal and transverse dispersivities. Molecular diffusion does the same on smaller scales down to the pore scale and may even dominate hydrodynamic transverse dispersion, especially in the vertical direction. By its nature, biodegradation at plume fringes tends to create, or at least maintain, high concentration gradients of the consumed species, as opposed to dispersion that smooths them.

Here, the relevance of dispersion is stressed, but for modeling approaches, we refer to transport modeling literature. However, it should be clear that only part of biodegradation happens at the plume fringes and part in the plume core where electron acceptors still may be stored as minerals.

MATHEMATICAL MODELING APPROACHES

Analytical Modeling

The solution of contaminant transport and biodegradation equations by analytical methods is possible for only a limited number of problems that are also simple in terms of boundary conditions, geometry, and biodegradation concepts. Nevertheless, they are very useful in answering basic questions, and they are quick, accurate, and straightforward to implement; Domenico and Schwartz (2) give a good collection of such solutions. Hence, it should always be checked first, if an analytical approach can address the particular modeling objective. The U.S. Environmental Protection Agency (EPA) provides the spreadsheet model *Bioscreen* based on analytical solutions, which, for example, allows obtaining quickly an impression of the effects of parameter values, different reaction models, and source zone extent (3). Finally, these types of approaches are used widely for risk assessment, but rather due to their simplicity and swiftness, than for their general applicability.

Numerical Modeling

Numerical modeling is based on techniques such as finite difference, finite element, or relatives of them. Only a few aspects are presented here that are specific to modeling reactive transport and other coupled equations.

- Reactive processes can be fast or slow compared with timescales of transport processes in aquifers or columns. Fast reactions are treated as an equilibrium process and may be described mathematically by thermodynamic mass action laws; kinetic reactions might be formulated as biodegradation kinetics, as before. Incorporating both imposes additional numerical challenges.
- The slow part of the reaction system is described via partial differential equations, and the fast part via nondifferential equations. There are different ways of combining them (4).
- Transport and reactions can be treated as sequential numerical steps or in one single step. The former is the less accurate method, even if the sequential steps are iterated, and may also require small time steps, whereas the latter is slower and needs more memory (5).
- Biodegradation can be handled as a single step, but alternatively also as a first step providing intermediate products followed by one or several processes of electron acceptor consumption. In the latter case, the first step is usually treated as slow, and the second step as fast. For a current literature survey of one- versus two-step, approaches, see Brun and Engesgaard (6).

NUMERICAL MODELING TOOLS

Purpose

A broad range of codes, packages, and modules deals with transport and biodegradation modeling. They are applied usually to modeling problems that are studied in greater detail, and more complex processes have to be included in the modeling concept. The selection of a modeling tool for a particular study should be governed by the aim of the study, but in practice also depends on the availability of programs and the familiarity of the modeler with particular codes.

Types

One-dimensional geochemical codes can calculate transport and speciation together with a comprehensive list of geochemical reactions, including biodegradation. One of these codes, PHREEQC (7), is probably the most frequently used program with these specifications. It is a thoroughly tested, well maintained and growing program, offers a lot of features including graphical user interface and inverse modeling, and is taught frequently in courses around the world.

For two-dimensional applications, existing numerical programs typically come from an engineering background and include flow and transport modeling with added capabilities for reactive transport. The most famous and most widely used program package in this category is the MODFLOW family (8). Here, several packages for reactive transport simulations have been added to the flow simulation packages: MT3D (9), RT3D (10), and PHT3D (11) and may be applied via one of the commercial MODFLOW packages with a graphical user-interface. Alternatively, BIOPLUME III is a noncommercial code to simulate natural attenuation of hydrocarbons in two dimensions (12).

Furthermore, there are other programs, which are designed especially to deal with complex, three-dimensional reactive transport simulations incorporating a lot of real complexity. These are often research codes rather than commercially spread programs, and thus they may be less well documented, less user-friendly, less easily available, and restrictions may apply for their use. But they represent the state of the art with respect to a comprehensive biochemical interpretation of field scale biodegradation modeling, and their use is often free of charge. Some examples of codes falling in this category are presented in Table 1.

An important note: The programs named so far represent a snapshot of today. However, these programs are a type of software that may be further developed in the future, but their development may also be discontinued or follow a different route. Therefore they should be seen as examples rather than unchanging recommendations. However, there are organizations that offer up-to-date programs in contaminant transport and biodegradation modeling, and even offer some of them free on-line. Two of them are the U.S. Geological Survey (USGS) (http://water.usgs.gov/software/ground_water.html) and the U.S. Environmental Protection Agency (EPA) (http://www.epa.gov/ada/csmos.html).

Handling

One should be aware that the hard work only starts after selecting the software fitting the purpose and the budget and learning how to run it. At that time, one has to specify actual parameter values, and this issue is even more important for reactive transport modeling than for other types of modeling in groundwater. There are more parameters involved, and often their values are difficult to obtain. This is one of the big and crucial problems in reactive transport modeling, especially for field cases: How to get the parameter values for biodegradation and the large number of components involved? Field tests are expensive and potentially give ambiguous results due to the dynamics and heterogeneity of biodegradation and plumes, but lab tests represent the field site only to a degree. For example, degradation rates found in the lab are typically higher, even by one or two orders of magnitude, than those fitted to field degradation. Nevertheless, laboratory tests can yield valuable information: (1) The biodegradation concept and relative importance of processes seem to be a valid starting point for field case modeling. (2) There are parameters that are more transferable to the field than rate constants, which are, for, example half saturation constants and inhibition coefficients. (3) Modeling of a lab experiments forces the modeler to get a complete and consistent picture of processes involved and parameters to be quantified.

Table 1. Examples of Numerical Programs as Tools for Sophisticated, Three-Dimensional Reactive Transport Simulation in Groundwater. This is not a Complete List of Such Codes

Name	Place or Organization of Development	Reference
BIO3D & BIONAPL3D	University of Waterloo, Canada	(13,14)
CRUNCH	Lawrence Livermore National Laboratory/Pacific Northwest National Laboratory, USA	(15)
Hydrobiogeochem 123D	OAK Ridge National Laboratory/Penn State University, USA	(16)
MIN3P	University of Waterloo and UBC Vancouver, Canada	(4)
TBC	University of Heidelberg, Germany	(17)
UG	IWR, University of Heidelberg, Germany	(18)

There will almost always be some parameters in reactive transport modeling that have to be fitted, preferably by inverse modeling, but maybe also by adjusting them by hand. Both ways work better when more data are available on all aspects of the problem. In any case, the sensitivity of the model to the choice of the parameter values has to be assessed thoroughly, and the fitted values have to be checked afterward for their plausibility, at least if no boundary limits have been used for parameter fitting. The sensitivity of parameters and the range of sensible parameter values determine the inherent uncertainty of the model. However, stochastic approaches are required for these complex, coupled, and nonlinear systems to get a real handle on the range of possible outcomes of certain model properties. Unfortunately, this adds heavily on top of the numerical load required anyway, and thus often a stochastic assessment has to be based on simplified conceptual models.

This leads us to the fact that the hardware for reactive transport modeling is still a limitation in dealing with complex problems, besides the problem of appropriate parameter values. On a standard PC of the current generation, even one-dimensional simulations may take hours to days to complete. This illustrates the need to enhance calculation power by using parallel processing capabilities for these simulations, especially for two-dimensional and three-dimensional problems. However, besides access to such hardware, this also requires software that can run on multiple processors at the same time.

Finally, modeling also means that the user has to deal with modeling results and has to interpret them. One aspect is that numerical modeling results should be frequently checked for consistency and accuracy, which is true even more for reactive transport modeling than for pure transport modeling. Another aspect is that often a large amount of data is produced, which leads to the necessity to visualize intelligently and efficiently the usually time-dependent data to obtain the best interpretation, especially for two- and three-dimensional data sets.

CLOSING REMARKS

This article is only an introduction to the field and touches on several subjects. For an excellent and clear book to get into this topic, we recommend Zheng and Bennett (19). Furthermore, the IAHS red book series provides a cross section of contemporary scientific contributions, for example, Thornton and Oswald (20) for natural and enhanced attenuation of contaminants or Stauffer et al. (21) and Kovar and Hrkal (22) for a broad range of issues on groundwater flow and contaminant transport modeling.

Acknowledgments
Dr. Ian Watson of the Department of Civil and Structural Engineering, University of Sheffield, UK, helped in reviewing and improving this entry.

BIBLIOGRAPHY

1. Wiedemeier, T.H., Rifai, H.S., Newell, Ch.J., and Wilson, J.T. (1999). *Natural Attenuation of Fuels and Chlorinated Solvents in the Subsurface*. John Wiley & Sons, New York.
2. Domenico, P.A. and Schwartz, F. (1997). *Physical and Chemical Hydrogeology*. John Wiley & Sons, New York.
3. Newell, C.J., McLeod, R.K., and Gonzales, J.R. (1996). *BIOSCREEN—Natural Attenuation Decision Support System*. User's Manual, Version 1.3, EPA/600/R-96/087.
4. Mayer, K.U., Frind, E., and Blowes, D.W. (2002). Multicomponent reactive transport modeling in variably saturated porous media using a generalized formulation for kinetically controlled reactions. *Water Resour. Res.* **38**(9): doi:10.1029/2001WR000862.
5. Saaltink, M.W., Carrera, J., and Ayora, C. (2000). A comparison of two approaches for reactive transport modelling. *J. Geochem. Exploration* **69**: 97–101.
6. Brun, A. and Engesgaard, P. (2002). Modelling of transport and biogeochemical processes in pollution plumes: Literature review and model development. *J. Hydrol.* **256**(3–4): 211–227.
7. Parkhurst, D.L. and Appelo, C.A.J. (1999). *User's guide to PHREEQC (Version 2)—A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations*. U.S. Geological Survey, Denver, CO, Water-Resources Investigations Report 99-4259, p. 310. Available: http://wwwbrr.cr.usgs.gov/projects/GWC_coupled/phreeqc/.
8. USGS. (1997). *Modeling Ground-Water Flow with Modeling Ground-Water Flow with MODFLOW and Related Programs*, USGS Fact Sheet FS-121-97. Available: http://water.usgs.gov/nrp/gwsoftware/modflow.html
9. Zheng, C. and Wang, P.P. (1999). *MT3DMS: A Modular Three-Dimensional Multispecies Transport Model for Simulation of Advection, Dispersion and Chemical Reactions of Contaminants in Groundwater Systems*, Documentation and User's Guide, Contract Report SERDP-99-1. U.S. Army Engineer Research and Development Center, Vicksburg, MS.
10. Clement, T.P. (1997). *A modular computer model for simulating reactive multi-species transport in three-dimensional groundwater systems*, Rep. PNNL-SA-28967. Pacific Northwest Natl. Lab., Richland, WA.
11. Prommer, H., Barry, D.A., and Zheng, C. (2003). MODFLOW/MT3DMS-Based Reactive Multicomponent Transport Modeling. *Ground Water* **41**(2): 247–257.
12. Rafai, H.S., Newell, C.J., Gonzales, J.R., Dendrou, S., Kennedy, L., and Wilson, J.T. (1998). *BIOPLUME III—Natural Attenuation Decision Support System*. User's Manual, Version 1.0, EPA/600/R-98/010.
13. Schirmer, M., Molson, J.W., Frind, E.O., and Barker, J.F. (2000). Biodegradation modelling of a dissolved gasoline plume applying independent laboratory and field parameters. *J. Contam. Hydrol.* **46**: 339–374.
14. Molson, J.W., Barker, J.F., Frind, E.O., and Schirmer, M. (2002). Modeling the impact of ethanol on the persistence of benzene in gasoline-contaminated groundwater. *Water Resour. Res.* **38**(1): doi:10.1029/2001WR000589.
15. Steefel, C.I. (2001). *GIMRT, version 1.2: Software for Modeling Multicomponent, Multidimensional Reactive Transport*. Users Guide, Technical Report UCRL-MA-143182. Lawrence Livermore National Laboratory, Livermore, CA.
16. Yeh, G-T., Salvage, K.M., Gwo, J.P., Zachara, J.M., and Szecsody, J.E. (1998). *HydroBioGeoChem: A Coupled Model of Hydrologic Transport and Mixed Biogeochemical Kinetic/Equilibrium Reactions in Saturated-Unsaturated Media*. Report ORNL/TM-13668. Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, TN.
17. Schäfer, D., Schäfer, W., and Kinzelbach, W. (1998). Simulation of reactive processes related to biodegradation in aquifers—1. Structure of the three-dimensional reactive transport model. *J. Contam. Hydrol.* **31**(1–2): 167–186.
18. Bastian, P., Birken, K., Lang, S., Johannsen K., Neuß, N., Rentz-Reichert, H., and Wieners, C. (1997). UG: A flexible software toolbox for solving partial differential equations. *Computing and Visualization in Science* **1**: 27–40.
19. Zheng, C. and Bennett, G. (2002). *Applied Contaminant Transport Modeling*. John Wiley & Sons, New York.
20. Thornton, S.F. and Oswald, S.E. (Eds.). (2002). *Natural and Enhanced Restoration of Groundwater Pollution*. Red book, IAHS Publication no. 275. IAHS Press, Wallingford, UK.
21. Stauffer, F., Kinzelbach, W., Kovar, K., and Hoehn, E. (Eds.). (2000). *Calibration and Reliability in Groundwater Modelling: Coping with Uncertainty*. Red book, IAHS Publication no.265. IAHS Press, Wallingford, UK.
22. Kovar, K. and Hrkal, Z. (Eds.). (2003). *Calibration and Reliability in Groundwater Modelling: A Few Steps Closer to Reality*. Red book, IAHS Publication no. 277. IAHS Press, Wallingford, UK.

BIOFOULING IN WATER WELLS

Stuart A. Smith
Smith-Comeskey Groundwater
Science LLC
Upper Sandusky, Ohio

BIOFOULING EFFECTS ON PRODUCTION EFFICIENCY AND WATER QUALITY

Biofouling as a term has its origins in pre-World War II studies of marine fouling and generally refers to coating, encrustation, and corrosion associated with the attachment of organisms to surfaces, including human-engineered structures. The phenomenon arises from the tendency of life across all known kingdoms to cling to and use surfaces. Biofouling takes many forms, ranging from bacterial-viral films plaguing urinary tracts to fouling of water intakes and other maritime structures (and even baleen whales) by films that also include macroorganisms such as barnacles and zebra mussels.

Biofouling of water wells and associated downstream systems tends to result from the formation of biofilms by bacteria. It is now widely recognized that biofouling is the first or second most costly deteriorating factor for groundwater systems in North America (1–3), and the problem has been recognized as important worldwide (4–6) for many years (7), although not universally accepted.

The composition of biofilms in any particular water well most likely echoes the populations of microflora in the associated aquifer and earth materials through which recharge passes. Those that thrive or persist are adapted to the environmental conditions in the well and contributing aquifer and vary by aquifer conditions (3,8). Where conditions permit, biofilm communities may closely resemble those in surficial soils, or they may be distinct.

In transmissive media systems such as aquifers and associated water wells, an important function of biofilms is to manipulate oxidation-reduction gradients and to accumulate and recycle resources (nutrients, metals, and organic carbon). One result of this activity is the deposition of reactive metal oxides and sulfides and rapid transformation of other redox-sensitive species, particularly in the nitrogen and sulfur series. The oxides are bulky solids and, along with slimes, contribute to blocking pore spaces and various engineered apertures.

The best known form of well water biofouling, iron and manganese biofouling, develops from both direct microbial function and indirect or passive interactions among biofilm, aqueous solutions, and surfaces. Ferrous iron compounds [Fe(II)] or ions (Fe^{2+}) can also be oxidized to the ferric [Fe(III)] state by nonbiological (abiotic) oxidants such as chlorine or oxygen, so the precipitation of iron by bacterial action is often hard to distinguish from abiotic processes, and both mechanisms are considered to operate in wells (8). Autooxidation of Mn in ambient ground waters is not favored due to the high E_h of the Mn(II)–Mn(IV) transformation (approx. +600 mV at pH 7.0), thus the transformation is presumptively mediated by the presence of biofilms. The formation of metal oxides is synergistic: as reactive iron oxihydroxides such as ferrihydrite are formed, more metal oxides deposit abiotically. Bacterial extracellular polymer (ECP) slimes that sequester metal ions and metal oxide colloids are involved in these interactions in a process analogous to iron sequestration treatments used to control iron precipitation.

The formation of iron sulfides and sulfur-oxidizing biofilms are two other significant water well biofouling phenomena where sulfide is present in groundwater. Sulfide is formed by microbial sulfate-reducing respiration and secondarily in some systems by putrefaction. Iron sulfides (which form dense black deposits) form spontaneously when aqueous sulfide ions come in contact with Fe^{2+} in solution. Sulfide oxidizers commonly form biofilms in the water columns of wells accessing sulfide groundwater. These biofilms can be highly viscous and have high integrity, easily blocking apertures. Chlorine and chlorine-based disinfectants can react with such components of the biofilm to produce halogenated organics.

Sulfides and other particles, such as calcite redeposited from solution (often also mediated by microbial activity), enhance the clogging potential of these biofilms. In some systems, mineral FeS_x clogs occur at the lower redox potentials that favor them (typically in the aquifer), and sulfur-oxidizer slimes form at higher E_h and pH (typically in the well water column).

Other less commonly described forms of metal oxidizing biofouling have been described, including aluminum biofouling (9). Aluminum hydroxide is coprecipitated with sulfate in high-organic-carbon white deposits. The composition varies according to variations in aeration. A variation is aluminum complexed in biofilms, often with iron deposited. Similar deposits have been described for samples from North Carolina in the United States.

Biofouling becomes a matter of public health concern when biofilms harbor or permit the proliferation of pathogens or indicator bacteria. If biofilms harbor and enhance survival of pathogens, they may result in illness. If the biofilms resist disinfection treatments, pathogens may avoid destruction. A recognized practical problem is positive results in total coliform tests that can be attributed to adaptive biofilm microflora that grow in commonly used growth media that is intended to be selective for the total coliform group of bacteria.

BIOFOULING EFFECTS ON WELL WATER SYSTEMS AND WATER QUALITY

Biofouling affects well water quality and well and water system metallic components in a variety of ways. Often the earliest manifestation of biofouling is the water quality degradation that accompanies mass bacterial growth and associated iron, manganese, and sulfur transformations. Typical symptoms are

- "red" or "black" water, increased turbidity (often far above typical groundwater), transiently high chlorine demand, and taste and odor problems that can result in a lack of confidence in the quality of the water supply.
- alteration of water quality in pumped samples. Biofouling can have a significant impact on the representativeness of groundwater samples from wells. Fluctuating Fe and Mn levels occur as Mn and Fe complexed with ECP may occur in suspension and can be detected in high levels in analytical results from unfiltered samples. Iron biofilms efficiently remove Fe and Mn, as well as other metals, from solution through chelation and deposition on reactive Fe(III) and Mn(IV) oxide mineral surfaces (6,8).

Chemical digestion of Fe- and Mn-biofilm samples indicate the presence of metals (e.g., Pb, As, Se) not detected in associated water samples (8). Thus, sample water pumped from biofouled wells can have metal levels lower than ambient ground water outside of the influence of the well (2,10–12).

- Work summarized in Smith (3) and subsequent work in alluvial aquifers and engineered systems (13) shows that (1) low total Fe (at or below the 300-μg/L standard) and (2) total Fe:Mn ratios of samples near 1:1 or upset (i.e., Mn:Fe >1:1) are typical of pumped water samples from biofouled water wells in microbially active alluvial aquifers.
- Elevated iron and manganese concentrations in pumped groundwater are also typically the result of bacterial activities in the aquifer, including respiratory Fe(III) and Mn(IV) reduction in the presence of abundant organic carbon and also corrosion of metal equipment. If unfiltered samples containing biofilm colloids are digested, very high Fe and Mn levels may be recorded (13).
- Biofilm influences have significant impacts on oxidation-reduction potential (3,8,13) and explain discrepancies between E_h values calculated from different couples and between calculations and measured values (8).
- Corrosion is enhanced by biofilm action and may occur where corrosion-incrustation index calculations predict that it would not occur. Such microbially influenced corrosion (MIC; 14) can also affect materials not expected to be subject to corrosion, accelerates well aging, and contributes corrosion-product constituents to water samples (15).
- Intermittent total coliform positive test results, mostly caused by bacteria not necessarily part of the total coliform group, but possessing the galactosidase enzyme.

It is important to note that corroding, encrusting, and nonencrusting biofouling effects, as well as test results that pose health concerns, often occur simultaneously.

BIOFOULING EFFECTS ON SYSTEM HYDRAULIC PERFORMANCE

In a well, biofouling phenomena may encrust or loosely plug well borehole intake areas and screens, pumps, and other downstream equipment. The initial process is the formation of a biofilm on surfaces in the well (casing, screen, pump) and in the aquifer in the vicinity of the well. The time course of this process to result in water quality or pumping problems may vary considerably, depending on site-specific conditions:

- Clogging (both formation/well and pump/discharge systems). Well clogging is usually expressed as reduced specific capacity (yield in volume/time unit versus drawdown during pumping).
- Corrosion: Cathodic depolarization of steel surfaces, erosion by organic acids, and intergranular cracking corrosion of stainless steels.
- Biofilms on interior pipe walls become increasingly hard or thicker over time. This tuberculation (the entire complex of corrosion/encrustation) can reduce the hydraulic capacity of a system. Tubercles increase hydraulic resistance by reducing diameter and by increasing the roughness of the interior surface of the pipe, dramatically increasing the energy cost to pump.
- Moreover, encrustations pack around lineshafts and bearings inside column pipe and reduce shaft life due to increased friction and reduced cooling. These developments may be aggravated by well operating choices (8).
- Biofouling shares top troublemaking honors with abrasion and clogging by silt and sand in groundwater systems for water supply, long-term dewatering, and on-site treatment to remove contamination (1,16,17). In fact, the two problems are frequently interactive (1,18–20), and biofilms may immobilize solid-phase silica (8), contributing to clogging.

With regard to well hydraulic performance,

- Reductions in specific capacity due to biofouling vary locally, but annual reductions are on the order of 1 to 3.4 on Long Island, NY, ranging up to 47% annually (21), for example.
- Considerable biofouling buildup can occur before well specific capacity is impacted.
- Wells accessing aquifers with low hydraulic conductivity (K) show faster and more enhanced performance decline compared to wells in formations with higher K.
- Higher Fe and Mn levels, particularly when oxidizing conditions prevail, can accelerate performance decline, although rapid biofouling clogging can occur at low levels. Higher total P, sulfate and organic carbon are also associated with enhanced biofouling development and clogging (6,8,22).

See related items, EVALUATION OF MICROBIAL COMPONENTS OF BIOFOULING, and WELL MAINTENANCE, for diagnostic methods and their applications where biofouling occurs in wells.

BIBLIOGRAPHY

1. Borch, M.A., Smith, S.A., and Noble, L.N. (1993). *Evaluation and Restoration of Water Supply Wells*. AWWA Research Foundation, Denver, CO.
2. Chapelle, F.H. (1993). *Ground-Water Microbiology and Geochemistry*. John Wiley & Sons, New York.
3. Smith, S.A. (1992). *Methods for Monitoring Iron and Manganese Biofouling in Water Supply Wells*. AWWA Research Foundation, Denver, CO.
4. Alcalde, R.E. and Gariboglio, M.A. (1990). Biofouling in Sierra Colorado water supply: A case study. *Microbiol. Civ.*

Eng., FEMS Symposium No. 59. P. Howsam (Ed.). E.&F.N. Spon, London, pp. 183–191.

5. Barbic, F., Krajcic, O., and Savic, I. (1990). Complexity of causes of well decrease. *Microbiol. Civ. Eng., FEMS Symposium No. 59*. P. Howsam (Ed.). E.&F.N. Spon, London, pp. 198–208.
6. Houben, G.J. (2003). Iron oxide incrustations in wells. Part 1: Genesis, mineralogy, and geochemistry. *Appl. Geochem.* **18**: 927–939.
7. Harder, E.C. (1919). *Iron-Depositing Bacteria and Their Geologic Relations*. U.S. Geological Survey Water Supply Paper 113, U.S. Geological Survey, Washington, DC.
8. Walter, D.A. (1997). *Geochemistry and Microbiology of Iron-Related Well-Screen Encrustation and Aquifer Biofouling in Suffolk County, Long Island, New York*. Water Resources Investigation Report 97-4032. U.S Geological Survey, Coram, NY.
9. McLaughlan, R.G., Knight, M.J., and Steutz, R.M. (1993). *Fouling and Corrosion of Groundwater Wells: A Research Study*. National Centre for Groundwater Management, University of Technology, Sydney, Australia.
10. Vuorinen, A. and Carlson, L. (1985). Special of heavy metals in Finnish lake ores; selective extraction analysis. *Int. J. Environ. Anal. Chem.* **20**: 179–186.
11. Mouchet, P. (1982). From conventional to biological iron removal in France. *JAWWA* **84**: 158–167.
12. Van Riemsdijk, W.H. and Hiemstra, T. (1993). Adsorption to heterogeneous surfaces. *Metals in Groundwater*. H.E. Allen, E.M. Perdue, and D.S. Brown (Eds.). CRC Lewis, Boca Raton, FL, pp. 1–36.
13. Smith, S. and Hosler, D.M. (2001). *Report of Investigations with Recommendations, Biological Fouling of the Pressure Relief Drainage System, Pablo Canyon Dam, Montana*. Bureau of Reclamations, Denver, CO.
14. Videla, H.A. (1996). *Manual of Biocorrosion*. CRC Press Lewis, Boca Raton, FL.
15. Oakley, D. and Korte, N.E. (1996). Nickel and chromium in ground water samples as influenced by well construction and sampling methods. *Ground Water Monitoring and Remediation* **16** (1): 93–99.
16. Howsam, P. and Tyrrel, S., (1989). Diagnosis and monitoring of biofouling in enclosed flow systems—experience in groundwater systems. *Biofouling* **1**: 343–351.
17. Smith, S.A. (1990). Well maintenance and rehabilitation in North America: An overview. In: *Water Wells Monitoring, Maintenance, and Rehabilitation*. P. Howsam (Ed.). E.&F.N. Spon, London, pp. 8–16.
18. Howsam, P., Missteârs, B., and Jones, C. (1995). *Monitoring, Maintenance and Rehabilitation of Water Supply Boreholes*. Report 137, Construction Industry Research and Information Association, London, UK.
19. Smith, S.A. (1995). *Monitoring and Remediation Wells: Problems Prevention and Cures*. CRC Lewis, Boca Raton, FL.
20. Alford, G. and Cullimore, D.R. (1999). *The Application of Heat and Chemicals in the Control of Biofouling Events in Wells*. CRC Press Lewis, Boca Raton, FL.
21. Walter, D.A. (1997). *Effects and Distribution of Iron-Related Well-Screen Encrustation and Aquifer Biofouling in Suffolk County, Long Island, New York*. Water Resources Investigation Report 96-4217. U.S Geological Survey, Coram, NY.
22. Cullimore, D.R. (1993). *Practical Manual of Microbiology*. CRC Press Lewis, Boca Raton, FL.

IN SITU BIOREMEDIATION OF CONTAMINATED GROUNDWATER

Jim C. Philp
Napier University
Edinburgh, Scotland, United Kingdom

Colin C. Cunningham
The University of Edinburgh
Edinburgh, Scotland, United Kingdom

Maria S. Kuyukina
Irena B. Ivshina
Institute of Ecology and Genetics of Microorganisms of the RAS
Perm, Russia

INTRODUCTION

Globally, the use of groundwater for various purposes is enormous. It is used as potable, industrial (especially cooling), and irrigation water on a huge scale. At least twelve megacities (population over 10 million) could not function without groundwater, and typically at least 25% of the water for these cities comes from aquifers (1). China alone has over 500 cities, and two-thirds of the water for them comes from aquifers.

Despite this importance, the number of instances of groundwater contamination due to accidental spills or unsatisfactory disposal is beyond counting. Up to a certain point, contamination can be attenuated by natural processes, especially biodegradation. In this regard, the biologically active zone is the vadose (unsaturated) zone where attenuation rates are highest. Contaminant removal continues in the saturated zone but usually at much lower rates, and migration of contaminants to the saturated zone can disperse the contaminants. Although it brings about dilution, this latter process often cannot be relied upon for complete decontamination.

Beyond a threshold, this natural attenuation cannot continue, and a decision has to be made whether or not to intervene with cleanup technology. This decision is now based on risk assessment. Risk assessment usually uses source–pathway–receptor analysis, and its outcome also determines, which treatment technology is to be used if treatment is necessary.

RISK ASSESSMENT

The objective of source–pathway–receptor analysis is the identification of the linkage(s) between them. Risk-based remedial design has as an objective the selection of the strategy to break the linkages and thus remove the risk. If, for example, there is a source of pollution and potential receptors (usually human) but no pathway to link them, or the pathway(s) can be blocked, then the risk is removed. If the receptor is human, then it is more likely that the risk will be removed by treating the source.

WHERE IS THE BOUNDARY BETWEEN GROUNDWATER AND SOIL TREATMENT?

Necessarily, if groundwater is to be treated by in-situ technologies, then the boundary between soil and water treatment becomes blurred, especially for vadose zone treatment.

SOME STATISTICS

The U.S. EPA has a recorded history of the treatment technologies applied to contaminated soil and groundwater at Superfund and other priority sites, which is updated through annual status reports (ASRs). As this history spans several decades, it is instructive to examine the record to identify the main technologies and trends in usage. Figure 1 shows that ex situ pump and treat methods have dominated groundwater treatment. This involves pumping the groundwater to the surface and selecting a treatment. Monitored natural attenuation (MNA) is not regarded as an engineered technology (2), but it is briefly described.

Specifically at Superfund sites (Fig. 2), *in situ* groundwater treatment remedies have been chosen 169 times at 135 sites (3). Air sparging has been chosen most frequently, followed by bioremediation. However, a comparison of those projects at the predesign, design, or installation stages (preoperational or future) and the operational and complete (current and historical) shows that the popularity of bioremediation has increased and that of air sparging has dropped dramatically.

Groundwater remedy type	Number of sites
Pump and treat	713
In situ treatment	135
Monitored natural attenuation	201
Other	822

Figure 1. Total number of sites with a groundwater remedy (3). Sites may be included in more than one category.

IN SITU BIOREMEDIATION TECHNOLOGIES

The treatment of groundwater in-situ has several advantages and disadvantages compared to pump and treat. The advantages include reduced site disturbance, which is crucial when a site is still being used; remediation around or under buildings without disturbance; and reduced worker exposure to volatile compounds. The quoted disadvantages include that site assessment must fully describe the hydrogeology and contaminant distribution, leading to increased cost; difficulty in control of reaction conditions; difficulty in accurately predicting end points; and therefore, the need for careful monitoring of the process. As it is difficult to take samples during *in situ* treatment, then it is difficult to assess if contamination hot spots have been remediated.

Additionally, in bioremediation, several other factors must be considered. Regularly quoted advantages of bioremediation are

- It is often less expensive than other techniques.
- Because the contaminants are mineralized, they are permanently eliminated.
- Long-term liability is thus eliminated.
- It has "green" credentials, representing a sustainable option.

Likewise, there are quoted disadvantages:

- Bioremediation cannot treat all contaminant types; most notably heavy metals cannot be biodegraded.
- Low permeability soil makes in-situ bioremediation of groundwater untenable.
- It may be too slow in certain circumstances.

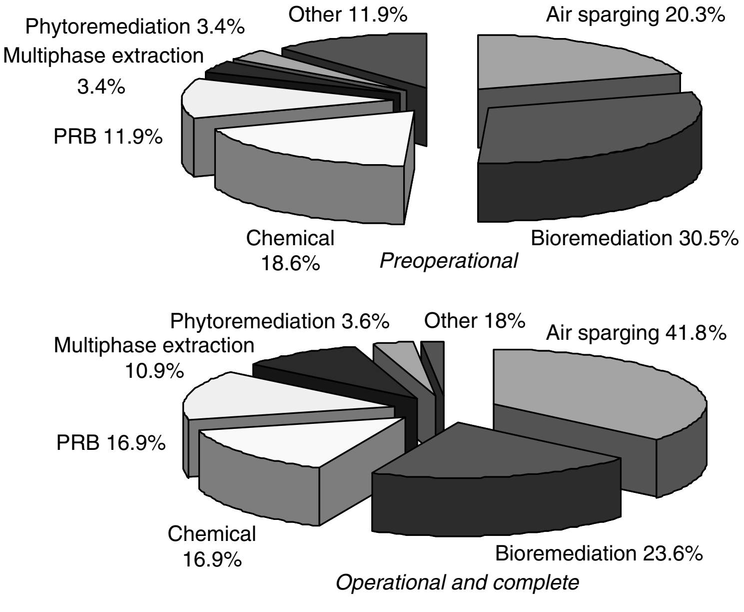

Figure 2. *In situ* groundwater treatment projects by technology, fiscal years 1982–2002 (adapted from Reference 3).

- It is often regarded as "incomplete" in that the contaminants may not be removed completely.
- Physicochemical conditions that are not ideal for microbial growth can slow the process down considerably.
- There is a worry that metabolites may be toxic.

As a result, the suitability of bioremediation has to be assessed case by case, and small-scale treatability studies are required (4), which adds to the time and expense of a project.

Aerobic or Anaerobic?

More than any other factor, it is likely that oxygen limitation occurs in groundwaters, and the most rapid biodegradation processes normally involve oxygenase enzymes (5). The supply of oxygen for *in situ* bioremediation adds considerably to the expense and the technical difficulty. This supply involves the use of either blowers or vacuum pumps to draw air through the system. A measure of the pivotal role of oxygen is in the use of hydrogen peroxide to enhance aeration. That its use has been considered at all is remarkable: concentrations of H_2O_2 above 100 to 200 mg/L are toxic to microorganisms; it can be consumed very quickly, limiting treatment to the regions near the injection well; a groundwater circulation system must be created; and good soil permeability is essential.

Recent advances in the knowledge of anaerobic biodegradation microbiology have opened up the prospect of greater acceptance of anaerobic bioremediation where maintaining aerobic conditions is not feasible. The existence of microorganisms capable of coupling the anaerobic reduction of Fe^{3+} to the oxidation of organic compounds shows promise (6). Virtually our entire knowledge of anaerobic metabolism of hydrocarbons has been gained since around 1990 (for reviews, see References 7 and 8). Several alkylbenzenes, alkanes or alkenes, are anaerobically biodegraded by denitrifying, ferric iron-reducing or sulfate-reducing bacteria. Another group of anaerobic hydrocarbon-degrading bacteria are "proton reducers" that rely on syntrophic associations with methanogens (8). As some of the most significant groundwater pollutants are chlorinated solvents, then a role for microbial reductive dehalogenation is feasible. Although slow, even the reductive dechlorination of dioxins is possible (9). A recent discovery has shown that even benzene can be oxidized completely under anaerobic conditions by pure cultures, using nitrate as the electron acceptor (10). Benzene is a particular problem in groundwater because it is relatively soluble and mobile.

Monitored Natural Attenuation

This technique involves monitoring the natural physical, chemical, and biological processes in soil and groundwater that are used to destroy a pollutant or limit its spread or migration (11). In nearly all situations, however, microbial reactions are the dominant processes driving natural attenuation (12), so it can be considered a long-term *in situ* bioremediation process.

Monitored natural attenuation may be useful where natural processes within the polluted area are effective in stabilizing or reducing the size of a contaminated groundwater plume, but there cannot be unacceptable impacts or risks to receptors. Monitoring and modeling are used to enable predicting the rate of attenuation and the rate of migration of the pollutant with some degree of confidence. The overall cost-effectiveness of MNA must take into consideration the long-term expenses involved in monitoring compared to the cost of a more intrusive approach, which is likely to be more costly during operational remediation, but requires no long-term monitoring.

Bioventing

This method uses indigenous microbes to biodegrade organic contaminants in the unsaturated zone above the water table. Technically, then, it is an *in situ* source treatment for contaminated soil (13), but it can be considered here as it inevitably results in soil water remediation. Bioventing is, after all, designed to remove contamination from the vadose zone to prevent future contamination of groundwater.

It combines supplying extra oxygen with vapor extraction to induce forced airflow through the contaminated area and enhance natural biodegradation (Fig. 3). Air is blown into the center of the area of contaminated soil above the water table and sucked out through peripheral boreholes to off-gas treatment prior to emission to the atmosphere. In its engineering manifestations, it closely resembles soil vapor extraction (SVE), but the two technologies have fundamentally different goals. SVE endeavors to maximize volatilization of low molecular weight contaminants, with some incidental bioremediation. Bioventing, however, endeavors to maximize biodegradation of the contaminant(s) regardless of molecular weight, with some incidental volatilization (14).

To enhance the biological process, nutrients can be supplied to the contaminated area. The injection of oxygen stimulates the microbes in the contaminated soil to

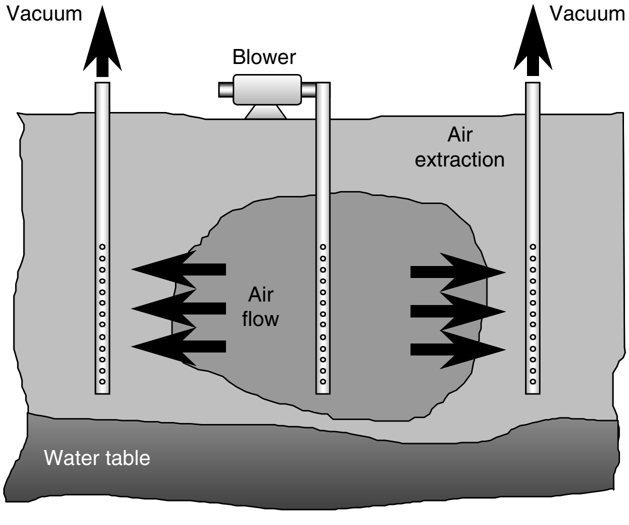

Figure 3. Bioventing.

degrade the organic contaminants to CO_2 and water. The process also mobilizes volatile compounds (either present in the soil or produced during biodegradation) to move toward boreholes, making extraction simpler.

Bioventing has now been used successfully at over 1000 sites and seems best suited to the bioremediation of middle distillate fuels such as diesel and jet fuel (15), but also nonchlorinated solvents, some pesticides, wood preservatives, and other organic chemicals have been removed. It is not successful in limited permeability soils, and application to the saturated zone relies on the water level being reduced (16). However, it can be successfully combined with biosparging.

Biosparging

Biosparging is similar to bioventing, in that air is injected, but in this case it is introduced below the water table (in the saturated zone) (Fig. 4) to increase the dissolved oxygen concentration in the groundwater and thus stimulate the activity of the indigenous microorganisms, thereby stimulating aerobic bioremediation. It can also have a similar stimulatory effect in the unsaturated zone. Like bioventing, nutrients may be added to enhance the biological process.

The success of the technique depends on adequate diffusion of the injected air away from the boreholes into the surrounding groundwater and soil. The location and number of boreholes depend primarily on the subsurface soil structure and permeability, and like bioventing, biosparging is best suited to permeable soils (16). Like bioventing, biosparging uses equipment that is readily available and easy to install.

Pump and Treat

The contaminated aqueous phase from the saturated zone is pumped, via a recovery well, to a treatment tank on the surface. In the treatment tank, nutrients, oxygen, and other electron acceptors (e.g., sulfate and nitrate) are added before the groundwater is pumped back into the ground via an injection well and recirculated through the contaminated zone (Fig. 5). The oxygen and nutrients in the injected groundwater stimulate the microbes in the contaminated zone to biodegrade contaminants dissolved in the groundwater and present in the soil. Groundwater extraction and injection continue until monitoring data show that the remedial objectives have been achieved. Technically, then, when pump and treat is operated as a bioremediation technology, it combines features of ex situ (water is pumped to the surface) and *in situ* (biological treatment continues underground).

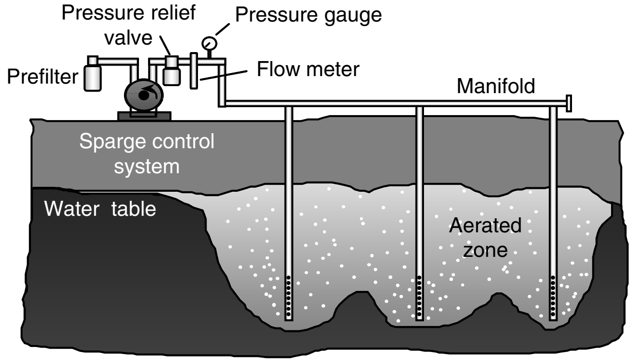

Figure 4. Biosparging.

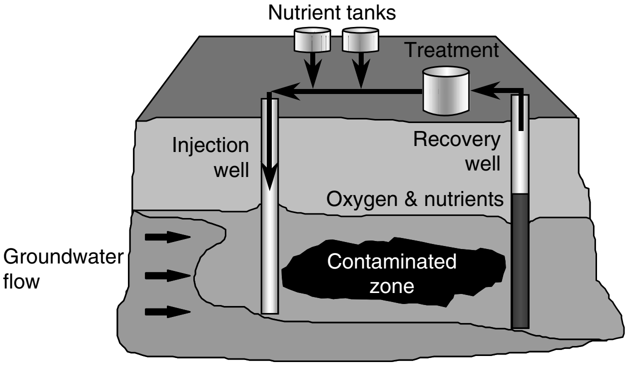

Figure 5. Pump and treat.

Pump and treat is a technology of great flexibility because the surface treatment need not be biological. In fact, a treatment train employing a variety of different technologies is possible at the surface. A much higher degree of process control is possible at the surface, for example, temperature, pH. For these reasons, pump and treat has historically been a very popular treatment for contaminated groundwater. However, there have been frequent questions over the efficacy of pump and treat, probably because, as the technology has become accepted, expectations have become too high (17).

Permeable Reactive Barriers

A PRBs consists of a reactive material that is placed in the path of flowing groundwater (Fig. 6), and due to the permeability of the chosen reactive material, it removes contaminants from the flow as the groundwater passes through. These barriers allow the passage of water while prohibiting the movement of contaminants. Agents within the barrier are materials such as zero-valence metals, for example, Fe^0, chelators, sorbents, and microbes (18). The contaminants are either degraded or retained in concentrated form by the barrier material, which may need to be replaced periodically.

PRBs have considerable flexibility in that combined chemical and biological treatments are possible within

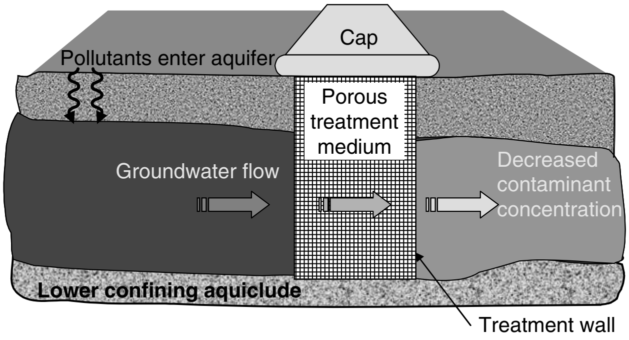

Figure 6. Permeable reactive barrier.

the same barrier, or barriers can be sequenced. For example, a contaminated aquifer beneath a petrochemical site may typically contain both chlorinated solvents and petroleum hydrocarbons. The chlorinated compounds may be removed by zero-valence iron in the permeable reactive barrier. With time, the permeable barrier will become colonized by bacteria, and bioremediation of petroleum hydrocarbons will occur, given the correct conditions. Alternatively, hydrocarbon-oxidizing bacteria may be deliberately added (the practice of bioaugmentation). Another aspect of the flexibility of the PRB is that in many cases it can be installed at working sites with minimal disruption to normal operations.

COSTS

The costs of all remediation technologies are constantly shifting and vary greatly from country to country. So quoted figures per tonne or per cubic meter of contaminated groundwater or soil would soon be out of date. Rather, it should be possible to identify economies of scale if confidence is gained in a technology and it is used increasingly at full scale. This confidence in bioremediation has been lacking in the past, and in many parts of the world, it is still not accepted at full scale.

If a technology is following a typical economies of scale pattern, then the more it is used, the more costs will decrease. The US EPA (19) examined the correlation between unit costs and quantity treated for six different soil and groundwater remediation technologies: bioremediation, thermal desorption, soil vapor extraction, on-site incineration, pump and treat, and permeable reactive barriers. Some important findings were reported:

1. Four of the six technologies (bioventing, thermal desorption, soil vapor extraction and pump and treat) evidenced a correlation between unit cost and quantity treated, thus exhibiting an economies of scale pattern.
2. Bioventing had the best correlation of these four technologies.
3. No other bioremediation technology exhibited any such correlation.
4. Pump and treat groundwater remediation systems showed a correlation for both unit capital costs and unit average operating costs.

The lack of correlation for the other bioremediation technologies may simply result from lack of data from a sufficiently large number of full-scale projects, but the signs are encouraging that bioventing has proven to be a market success.

BIBLIOGRAPHY

1. Morris, B.L., Lawrence, A.R.L., Chilton, P.J.C., Adams, B., Calow, R.C., and Klinck, B.A. (2003). *Groundwater and Its Susceptibility to Degradation: A Global Assessment of the Problem and Options for Management*. Early Warning and Assessment Report Series RS.03-3. United Nations Environment Programme, Nairobi, Kenya.
2. U.S. EPA. (1999). *Use of Monitored Natural Attenuation at Superfund, RCRA Corrective Action and Underground Storage Tank Sites*. OSWER Directive 9200.4-17P.
3. U.S. EPA. (2004). *Treatment Technologies for Site Cleanup: Annual Status Report*. 11th Edn. EPA-542-R-03-009.
4. Boopathy, R. (2000). Factors limiting bioremediation technologies. *Bioresource Technol.* **74**: 63–67.
5. Dagley, S. (1975). A biochemical approach to some problems of environmental pollution. *Essays Biochem.* **11**: 81–138.
6. Crawford, R.L. (1991). Bioremediation of groundwater pollution. *Curr. Opin. Biotechnol.* **2**: 436–439.
7. Holliger, C. and Zehnder, A.J.B. (1996). Anaerobic biodegradation of hydrocarbons. *Curr. Opin. Biotechnol.* **7**: 326–330.
8. Heider, J., Spormann, A.M., Beller, H.R., and Widdel, F. (1999). Anaerobic bacterial metabolism of hydrocarbons. *FEMS Microbiol. Rev.* **22**: 459–473.
9. Adriaens, P., Fu, Q., and Grbic-Galic, D. (1995). Bioavailability and transformation of highly chlorinated dibenzo-*p*-dioxins and dibenzofurans in anaerobic soils and sediments. *Environ. Sci. Technol.* **29**: 2252–2260.
10. Coates, J.D., Chackraborty, R., Lack, J.G., O'Connor, S.M., Cole, K.A., Bender, K.S., and Achenbach, L.A. (2001). Anaerobic benzene oxidation coupled to nitrate reduction in pure culture of two strains of *Dechloromonas*. *Nature* **411**: 1039–1043.
11. BioWise. (2000). *Contaminated Land Remediation: A Review of Biological Technology*. DTI/BW/18/5000/12/2000/NP.
12. Smets, B.F. and Pritchard, P.H. (2003). Elucidating the microbial component of natural attenuation. *Curr. Opin. Biotechnol.* **14**: 283–288.
13. U.S. EPA. (2001). *Use of Bioremediation at Superfund Sites*. EPA 542-R-01-019.
14. U.S. EPA. (1995). *Bioventing Principles and Practice*. Volume 1: Bioventing principles. EPA/540/R-95/534a.
15. U.S. EPA. (1995). *Soil Vapor Extraction (SVE) Enhancement Technology Resource Guide*. EPA/542-B-95-003.
16. Martin, I. and Bardos, P. (1995). *A Review of Full Scale Treatment Technologies for the Remediation of Contaminated Soil*. EPP, Richmond, VA.
17. U.S. EPA. (1996). *Pump-and-Treat Ground-Water Remediation. A Guide for Decision Makers and Practitioners*. EPA/625/R-95/005.
18. U.S. EPA. (2001). *Treatment Technologies for Site Cleanup: Annual Status Report (Tenth Edition)*. EPA-542-R01-004.
19. U.S. EPA. (2001). *Remediation Technology Cost Compendium—Year 2000*. EPA-542-R-01-009.

PROCESS LIMITATIONS OF *IN SITU* BIOREMEDIATION OF GROUNDWATER

DAVID B. VANCE
ARCADIS G&M, Inc.
Midland, Texas

JAMES A. JACOBS
Environmental Bio-Systems, Inc.
Mill Valley, California

Based on the origin and susceptibility to microbial interaction, hydrocarbons may be divided into two broad classes: petroleum hydrocarbons and complex chlorinated hydrocarbons.

Petroleum hydrocarbons are largely associated with the production, storage, or use of fuels, lubricants, and chemical feedstocks. Petroleum hydrocarbons have been demonstrated to be biodegradable by numerous species of bacteria. Over a period of 3.5 billion years, bacteria have been able to evolve genetic resources that allow some of them to potentially use petroleum hydrocarbons as a source of food.

Complex industrial hydrocarbons include chlorinated aliphatic and aromatic hydrocarbons, MTBE, pesticides and herbicides, and polymers. These new synthetic compounds have been manufactured for about 100 years. Bacteria have not had time to evolve the genetic information required to utilize them as a source of food. Due to recalcitrance to microbial attack, these complex industrial chemicals are termed xenobiotic.

Successful *in situ* bioremediation of groundwater has been demonstrated at sites impacted with petroleum hydrocarbons. The most successful method of application has been through stimulation of indigenous microbial species. Microbial stimulation is the process of ensuring that environmental conditions, nutrient availability, and requirements for an electron acceptor are adequate in the contaminated portions of the aquifer.

The most common cause for failure of saturated zone *in situ* bioremediation is the lack of adequate mass transport of the electron acceptor (usually oxygen). In this regard, the physical setting of the site is critical. Overall permeability and the scale and degree of heterogeneity are the factors governing the advective and diffusional transport rates of contaminants and remediation reagents in the subsurface. If mass transport rates are too low, saturated zone *in situ* bioremediation is not a viable option.

Given adequate mass transport properties, site-specific microbiological conditions can also impact the process. Unfortunately, the presence of indigenous microbes and efficient mass transport may still prove insufficient for effective bioremediation. Specific reasons for the poor performance of *in situ* bioremediation systems relate to unoptimized subsurface conditions.

There is uncertainty with regard to the effect of hydrocarbon availability on the effectiveness of biodegradation. Can bacteria degrade hydrocarbons adsorbed to surfaces or degrade hydrocarbons with low levels of solubility? Or must the hydrocarbon be solubilized before it can be biodegraded? Contradictory laboratory evidence and field evidence have been published for both scenarios. With the predominance of evidence indicating that solubilization must take place, degradation reactions with extracellular bacterial exudates are much less likely.

The answer is likely consortia specific and dependent on the ability of the bacteria to synthesize appropriate biosurfactants. This ability may be absent in some instances.

Although petroleum hydrocarbons are amenable to primarily aerobic biodegradation, for it to occur the indigenous bacteria must have the appropriate genetic information. This genetic information is specific and precise. The presence of a specific hydrocarbon will stimulate the synthesis of an oxygenase enzyme that is expressly configured to react with that stimulating hydrocarbon.

For remediation, indigenous microbes generally possess the genetic information required for appropriate enzyme production and the contaminant will stimulate the production of those enzymes. General microbial stimulation has the potential to produce a large amount of biomass that may not take part in the biodegradation process and actually cause harm through biofouling and plugging of injection wells, galleries, or surrounding formations. There is potential to lose critical subsurface mass transport capabilities.

BACTERIAL TRANSPORT

There are practical limits to the degree of cleanup obtainable using bioremediation. Hydrocarbons at the low parts per million (ppm) level may not be capable of supporting significant levels of microbial activity even under stimulation. Sites with relatively high levels of hydrocarbon impact may actually be better candidates for bioremediation than those lightly impacted at levels slightly above regulatory action levels. Toxicity related to the presence of heavy metals, such as chromium, arsenic, or lead, or low temperature of the groundwater have been observed to inhibit bacterial growth in a variety of settings. Stimulating electron acceptors must also be available at sufficient concentrations; native sulfate concentrations less than about 20 mg/L do not stimulate sulfate reducing bacteria even in the presence of usable carbon substrates.

Xenobiotic industrial compounds are often recalcitrant to direct aerobic microbial attack. However, over the last 20 years a biodegradation process termed co-oxidation (or cometabolism) has been successfully demonstrated by researchers. For example, the aerobic degradation of trichloroethylene (TCE) has been accomplished using monooxygenase and dioxygenase enzymes produced through the use of petroleum hydrocarbons as a metabolizable substrate (food source) and stimulus for enzyme production. This general process is termed co-oxidation and the hydrocarbon substrate used as a food source is the cometabolite. Many different hydrocarbon substrates have been observed to stimulate the generation of cooxidation enzymes. The currently known cometabolic substrates fall into two broad classes:

1. Analog substrates, which are hydrocarbons that have a geometry similar to the targeted xenobiotic compound.
2. Methanotrophic (which is different than methanogenic) microbial systems have proved particularly effective at generating xenobiotic active enzymes.

Enzymes with co-oxidizing potential have a strong natural affinity for the hydrocarbon that originally stimulated its generation. The enzyme is genetically tailored to the compound used as a food source. Over 300 mol of methane are required to biodegrade 1 mol of TCE via co-oxidation. The efficiency of the co-oxidation process is extremely poor. Under field conditions where mass transport is a critical success factor, a 300-fold decrease in the effectiveness of the reactants in the contaminated zone often can be impractical.

Accurate assessment of potential limiting factors such aerobic terminal electron acceptor (dissolved oxygen), geochemical conditions (pH, temperature, conductivity), and macronutrients (orthophosphate and ammonia as nitrogen) should be documented as part of the bioremediation evaluation process. Bioremediation is a dynamic process requiring monitoring of the hydrological, geochemical, and biological conditions over the life of a project.

NATURAL ATTENUATION: GROUNDWATER REMEDIATION BY NO ACTION

In a time of reappraisal for the allocation of financial resources to environmental action, a question of ever increasing importance is the consequence of no action concerning the release of petroleum or chlorinated hydrocarbons into groundwater. An important portion of that answer comes from the application of site-specific health based risk assessments. However, in instances where human consumption or exposure is not an issue, no action may be a reasonable alternative, even at elevated dissolved contaminant concentrations. The issue then becomes the determination of the consequence of no action under conditions where the sole process for remediation is natural attenuation.

The physical, chemical, geological, and biological processes that take place in a contaminated aquifer are complex. In most instances, a “native” aquifer is in a long-standing state of chemical equilibrium between the groundwater and the geologic matrix through which it flows. The release of anthropogenic hydrocarbons into an aquifer upsets that equilibrium. The dissolved concentration of the contaminant as it migrates through the aquifer is controlled by adsorption, dispersion, volatilization, and degradation. Adsorption affects the overall residence time of the release and dispersion affects the downgradient shape and dissolved concentration in the plume. Only volatilization and degradation contribute to the removal of contaminant from the aquifer, and at low concentrations degradation is the dominant mechanism for attenuation.

The mechanisms for attenuation through degradation can be broadly divided into two categories, biological and abiotic chemical action. This discussion is predicated on relatively “normal” groundwater conditions under which biological action proceeds at a rate orders of magnitude greater than abiotic processes. Extremes of pH, redox conditions, ionic strength, or temperature may make an exception to that generalization. Transformation can be chemically complex, dependent on the environmental conditions described above and affected by aquifer heterogeneity related to granular or fracture variability.

The factor controlling the rate of aerobic degradation is the availability of oxygen and the rate at which it can be introduced into the groundwater (through the groundwater–table interface) or the rate at which oxygen-rich groundwater can pass through the zones of adsorbed contamination. Each pound of petroleum hydrocarbons requires about 3.08 pounds of oxygen for complete degradation (1).

Typical *in situ* aerobic decay rates for groundwater are in the range of 35 μg/L·d (equivalent to about 0.5 oz/d per cubic yard of aquifer matrix).

Natural attenuation occurs both in the source zone and in the dissolved phase plume. In the source zone, oxygen will be rapidly consumed and portions of the aquifer will then host anaerobic degradation. Anaerobic degradation is limited by the availability of appropriate anaerobic electron acceptors such as nitrate, sulfate, or iron. When their availability is limited, degradation will stop after the production of aliphatic and aromatic organic acids; similarly, at low levels of dissolved oxygen (DO), aerobic degradation may also stop with the production of organic acids. The intrinsic biodegradation process and the alternative terminal electron acceptors are shown in Fig. 1 (2).

Optimum aerobic biodegradation occurs with the dissolved oxygen above 2 mg/L. Below that, the aerobic degradation rate of aromatic hydrocarbons will decrease dramatically. Conversely, under complete anaerobic conditions, nitrate reducing, iron reducing, and suflate reducing bacteria can effectively degrade hydrocarbons. However, at DO concentrations as low as 0.1–0.4 mg/L, anaerobic degradation rates will be reduced to just a few percent of optimum.

Because of all the mechanisms described above, hydrocarbon plumes tend to achieve a stable shape and size even when there is a continuous source of free phase hydrocarbon release. Steady state is achieved when the area of the plume edge is great enough to provide for a natural degradation rate equivalent to the rate of hydrocarbon infiltration. The edges of the dissolved plume do not have enough DO to support optimum rates of aerobic degradation but have too much DO to allow for optimum anaerobic degradation. The interior of a plume will support anaerobic natural attenuation, which is typically limited by the availability of iron in the mineral matrix and sulfate in the native groundwater and to a lesser extent the mineral matrix as well. However, once the source of hydrocarbon has been removed, a dissolved plume will narrow and dissipate from the edges inward, due to the availability of DO from groundwater along those edges.

The selection of a no action natural attenuation option should be based on an appropriate analysis of data gathered during the assessment of the site. First-order decay rates are appropriate for the evaluation of degradation kinetics at low concentrations, less than 1 ppm (an appropriate level to assume at the periphery of a plume). Given first-order decay rates, the analysis has a focus that is twofold—the effect of attenuation over time and the effect over distance.

Attenuation over time is measured at the edges of a plume using concentration measurements gathered repeatedly from specific monitor wells. The minimum recommended time is one year, with quarterly sampling from the selected monitor wells. The data for each well is then semilog plotted as log concentration against time. The slope of the line is the first-order decay constant in percent per day.

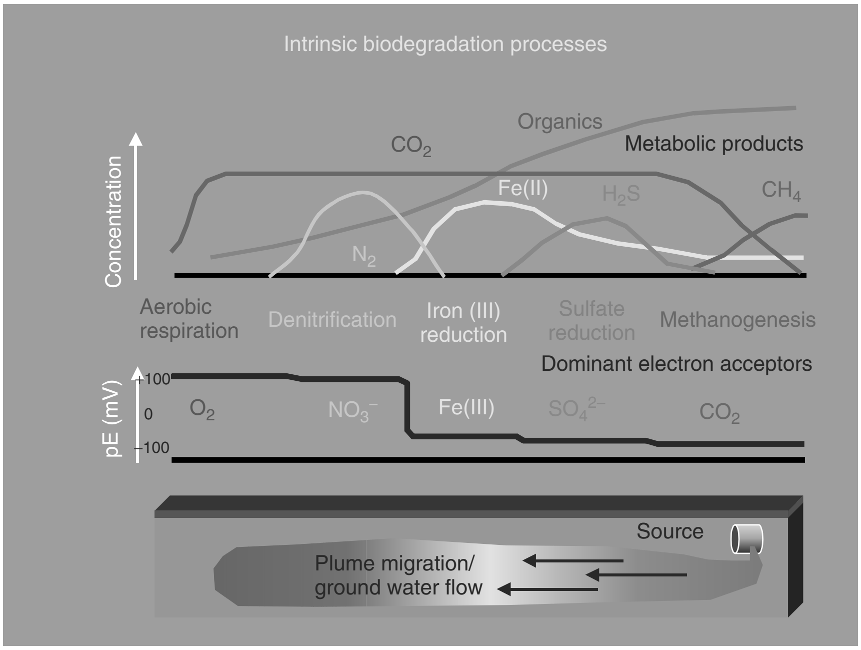

Figure 1. Intrinsic biodegration process (2). Petroleum hydrocarbons are referred to as organics in the diagram.

Attenuation with distance more accurately incorporates the effects of aquifer heterogeneity. Data for this analysis is obtained from a minimum of three monitoring wells, preferably along the long axis of the plume. This data is semilog plotted as log concentration versus distance. The slope of this line is equal to the decay constant divided by the groundwater velocity.

With this data, decisions can be made based on the site-specific contaminant dynamics under no action natural attenuation. This, in conjunction with a health based risk assessment, can allow for sound decision-making by the business and regulatory community.

In summary, the adoption of a no action alternative is most applicable to the dissolved phase plume only. Except for volumetrically small releases, it will still be necessary to remove or remediate the source zone of an impacted aquifer, after which natural attenuation may be a reasonable approach to the residual dissolved phases. Also implicit in this approach is that no action does not preclude the performance of requisite assessment activity, which can represent a significant long-term liability in some cases. Nonetheless, after proper source abatement, assessment, and analysis, the reliance on natural attenuation mechanisms for the final stages of cleanup is a cost effective and, if properly managed, environmentally sound resolution to aerially extensive dissolved phase hydrocarbon contamination.

NATURAL ATTENUATION: THE EFFECT OF PUMP AND TREAT REMEDIATION

In situ groundwater remediation has matured over the past 25 years, particularly with regard to understanding the dynamics of the interactions between contaminants, the impacted saturated soil matrix, and microbiological activity. Recent interest in the phenomenon of natural attenuation has served to illustrate the variety of microbial ecosystems that are present in a contaminant plume, each system determined by redox conditions and availability of electron acceptors. While natural attenuation is an attractive alternative to those responsible for groundwater contamination, the regulating communities are more skeptical. The need for proactive groundwater pumping remediation and the efficacy of natural attenuation pose a potentially complex balance that is governed by the subsurface conditions of each individual site. There is no universal applicable rule for the resolution of that balance. It is the responsibility of remediation designers to make those site-specific determinations and to provide the regulating community with information sufficient to support the proactive and the natural attenuation portions of each individual cleanup. Our purpose here is to point out some of the most significant factors impacting that balance.

The first and most dominant control is the nature of the saturated soil matrix. Several factors must be evaluated for remedial design:

1. The degree and scale of sediment heterogeneity, which determines how much and what portion of a contaminated aquifer can be affected with advective groundwater flow. Low permeability regions must rely on diffusional transport, which will dominate the overall remediation rate in the treatment zone.
2. The time of exposure to the contaminant is a direct function of the impact of heterogeneity described above. The contaminant will diffuse into the nonadvective portions of the aquifer soil matrix. At a minimum, remediation will take as long as the initial exposure. Due to the adsorptive retardation reactions, remediation is likely to take longer than the exposure time.

3. The geochemical composition of the soil matrix is also a factor in remedial design. Carbonaceous material and clays have a much higher propensity for the adsorption of organic contaminants. Iron oxides, in turn, have high adsorptive capacity for metal contaminants. Iron and sulfur minerals may be sources of electron acceptors as redox conditions are modified through the interaction of indigenous microbial populations and the contaminant.
4. The background geochemical makeup of the groundwater as well as that in the contaminant plume is an important remedial design factor. Dissolved oxygen, sulfate, nitrate, and iron can all potentially serve as alternate terminal electron acceptors to aid in the degradation of organic contaminants. Ferrous iron, hydrogen sulfide, and carbon dioxide are indicative end products of those reactions.
5. The distribution of the organic contaminant is another factor affecting remedial design. Free phase hydrocarbons should be recovered proactively with an extractive technology (pump and treat) or in the case of CVOCs possibly an *in situ* chemical oxidation injection technology. In cases where impact is shallow, excavation and disposal is still an extremely viable option.

The treatment of dissolved and adsorbed hydrocarbons is the point at which the balance between proactive remediation and natural attenuation must be determined.

One of the most important contributions that a pump and treat system makes to the *in situ* remediation of contaminated groundwater is plume capture and hydraulic control in the source zone and the core of the dissolved and adsorbed plumes. Background groundwater that is drawn through the plume perpendicular to the natural groundwater flow direction must also be evaluated.

In the past, the focus of pump and treat remediation has been on how it acts to flush and remove the contaminant. The contributions made by recent developments on the mechanisms of natural attenuation reside in the role of electron acceptors present at background concentrations within the aquifer. From the exterior to the interior of a plume, the specific electron acceptor zones are aerobic, denitrification, sulfate/iron reducing, and methanogenic.

The boundary between each specific redox zone and the active electron acceptor is controlled by the kinetics of the degradation process in each zone and the advective transport rates of groundwater through that zone. In most instances, the dominant effect is the groundwater transport rate.

The natural concentrations of these electron acceptors cover a wide range. Natural oxygen levels commonly range from 2 to 8 mg/L. Groundwater sulfate concentrations in soils derived from sedimentary rocks are typically in the 25-mg/L range, with higher values of several hundred mg/L not uncommon. Ferric oxides are commonly present in soils in the range of 0.5–5%; the ability of indigenous iron reducing bacteria to access that iron will vary from location to location. Given adequate permeability and the presence of appropriate electron acceptors, natural enhancement of pump and treat systems is possible and worth the relatively inexpensive analyses (some of which can be done with field kits) required to evaluate.

NATURAL ATTENUATION: TRANSVERSE DISPERSION AS THE NATURAL DRIVING FORCE

Dispersion is the process by which the interface of contaminated groundwater with native groundwater does not remain abrupt. The leading edge of a contaminant plume will arrive at a given point more rapidly than it would if advection alone were the acting driving force. The mean transport velocity of the contaminant mass remains the same, but concentration gradients are setup. This occurs simultaneously with the phenomenon of the contaminant occupying, with time, an increasing volume of groundwater. There are two fundamental types of dispersion—longitudinal and transverse. Both are accentuated by the inhomogeneous and anisotropic physical configuration of the permeable matrix within a groundwater system. There is also a contribution to both from diffusional transport as well.

Longtitudinal dispersion is caused by differences in groundwater velocity through pore spaces that vary in width or tortuosity. The result is dispersion that occurs along the direction of groundwater flow. Transverse dispersion is driven by groundwater flowing around individual particles in the aquifer matrix; the effect occurs perpendicular to the groundwater flow direction. Transverse dispersion is effective in mixing contaminated groundwater with native groundwater at the edges of an elongated plume and occurs only when there is a point source of contamination. In the context of natural attenuation, longitudinal dispersion is purely a dilution phenomena; transverse dispersion provides an influx of electron acceptors.

Both types of dispersion are dependent on horizontal and vertical variations in permeability. Increasing anisotropy and heterogeneity increase the magnitude of dispersion. Groundwater velocity also plays a role: At low velocities the effects of diffusion may equal those of dispersion.

The ratio of longitudinal to transverse dispersivity ranges from 1 to 24; most commonly, horizontal transverse dispersivity is 20% to 10% of the longitudinal dispersivity, and vertical transverse dispersivity is 2% to 1% of the longitudinal dispersivity. Instead of "football" shaped plumes, this difference in vertical and horizontal dispersivity tends to generate plumes that in three dimensions are "surfboard" shaped. The determination of the specifics of the effect of dispersivity on a contaminant concentration at a given time and location in an aquifer is an extremely complex process that requires detailed knowledge of the physical configuration of the aquifer matrix and solution to partial differential equations for final values. In many instances, a purely empirical approach is the only practical means of assessment rather than modeling.

The physical scale that is examined also has an impact: values of dispersivity change as one examines an aquifer on the scale of inches, feet, or thousands of feet.

Transverse dispersion is what serves to physically mix groundwater containing contaminants with adjacent groundwater that is unimpacted and contains natural electron acceptors. The phenomenon is probably best understood in terms of angle of divergence, that is, the angle between the two edges of the plume as it migrates from a point source. In granular materials that angle can be as low as 2° and in groundwater flowing through fracture systems as high as 20°. The most accurate determination of the angle of divergence is from information gathered as close as possible to the point representing the source of contamination.

The picture we normally have of a contaminant plume as it migrates away from a point source is one that is (in the two horizontal dimensions) tear drop shaped. The plume initially spreads downgradient and cross gradient until some point is reached at which the distal edges of the plume first travel parallel to the advective groundwater flow direction, and then begin to turn inward to the plume axis and close.

The distal portion of the plume is the dominant area where longitudinal dispersion (as well as transverse dispersion) is in effect. Adsorption and other attenuation reactions also take place, but for the purposes of this discussion those effects are ignored. Under pure hydrodynamic effects, a plume should dilute itself taking the shape of a cone at a constant angle of dispersion. This assumes that the nature of the geologic matrix and groundwater velocity stay constant, which is unlikely, especially in the vertical component. The most common result is that the downgradient shape of the plume is a consequence of attenuation caused by the migration of electron acceptors into the margins of the plume through the force of transverse dispersion—not through purely hydrodynamic flow effects.

NATURAL ATTENUATION: METHANOGENIC SYSTEMS

The ultimate no action alternative may be the use of methanogenic microbial systems. In this case, the electron acceptor is the contaminating hydrocarbon that may contribute to the degradation of petroleum and xenobiotic hydrocarbons There is a difference between methanotrophic bacteria and methanogenic bacteria. Methanotrophs use oxygen to oxidize methane into carbon dioxide (CO_2). Methanotrophic bacterial systems have received a great deal of attention over the last 15 years since it has been found that methane monooxygenase (the enzyme generated by methanotrophs to react with methane) can degrade a wide variety of chlorinated hydrocarbons. The process is known as cometabolism and is definitely an aerobic process.

Methanogenesis is the process of degrading hydrocarbons with the end product being methane (CH_4) gas and carbon dioxide. The general reaction is as follows:

$$2\text{C (organic contaminant)} + 2H_2O \rightarrow CO_2 + CH_4$$

This is a strictly anaerobic process; methanogenic bacteria are poisoned by the presence of oxygen at levels as low as 0.18 mg/L of soluble oxygen (as O_2).

The redox conditions under which these two different microbial systems operate are literally at opposite ends of the spectrum: methanotrophic reactions occur at the Eh range of +250 mV while methanogenic reactions occur at the Eh range of −200 mV.

As an aside, methanogenic bacteria are one of the three classes of bacteria termed Archaebacteria, which are representative of organisms that first appeared on Earth some 3.5 billion years ago. Although their activity is inhibited by oxygen, these bacteria are robust enough to appear in a wide variety of natural locations such as the intestinal tracts of ruminant mammals like cows, sewage digesters, groundwater, and soil.

The concentrations of methane and carbon dioxide are expressed as mole percent. This data is somewhat remarkable since this should be a difficult reaction to initiate. Oxygen, nitrate, and sulfate are all toxic or inhibitory to methanogenic activity. That means that at any site at which evidence of methanogenesis is present (methane gas), there has been a series of biodegradation reactions that have consumed the alternative electron acceptors. Migrating from the exterior to the interior of a plume, the type of redox conditions (Eh) will change from + 250 mV to − 200 mV, with oxygen, nitrate, ferric iron, and sulfate progressively being consumed.

The degradation of chlorinated xenobiotic compounds under methanogenic conditions is particularly enhanced via reductive dehalogenation reactions that involve substitution of hydrogen in the CVOC carbon chain. This can take place from the direct actions of bacteria and from the presence of molecular hydrogen that is produced from water during the methanogenic process.

Conceptually, methanogenesis might be considered "the ultimate of no action alternatives." The physical/chemical requirement for the removal of all other potential terminal electron acceptors infers that the hydrodynamics of such a system are relatively quiescent. The transport of alternate electron acceptors into the core of the contamination plume must be at a rate slow enough to allow for the consumption of all alternate electron acceptors before the methanogenic core zone is reached.

If this situation occurs, it can be a positive argument for absolutely no action involving pump and treat systems. That would increase groundwater velocities, introducing inhibitory electron acceptors into the active methanogenic zone. Groundwater systems in fine grained soils, where transport properties are poor, would be ideal for the exploitation of methanogenic degradation.

Of course, one problem with methanogenic degradation is the kinetics of the process. Methanogenic degradation occurs at rates that are orders of magnitude slower than the rates seen with other electron acceptors. However, at some sites with poor transport conditions or inaccessibility (i.e., a deep groundwater table), this may still be a reasonable alternative.

The number and types of hydrocarbons degraded under methanogenic conditions are very limited. Some laboratory studies have demonstrated the methanogenic degradation of toluene and *o*-xylene with no degradation of *m*-xylene, *p*-xylene, ethylbenzene, or benzene. Others have found evidence of degradation of benzene with recalcitrance toward other compounds.

The methanogenic process appears to be extremely selective and not capable of the complete degradation of all contaminants present in a typical hydrocarbon plume on the order of years. In addition, there is likely a great deal of microbial heterogeneity with specific degradation capacity varying from location to location. In instances where a specific compound has been released (such as toluene used as a solvent), methanogenesis may be a viable natural attenuation process. It also has value in the natural dehalogenation of chlorinated compounds.

In instances where an aquifer has good transport qualities and a supply of natural electron acceptors, natural attenuation alone can be a viable process and acceptable to the regulatory community. However, responsible parties must be prepared for the long-term monitoring that will often be required in support of a natural attenuation program—more accurately termed monitored natural attenuation by the regulating community. In other cases, supplementation of electron acceptors or improvement of groundwater dynamics through pump and treat may provide an adequate minimal approach. But, in most instances, methanogenesis as the "ultimate of no action alternatives" is not going to be practically applicable.

BIBLIOGRAPHY

1. Kun, J. (2004). *Practical Design Calculations for Groundwater and Soil Remediation*. CRC Press, Boca Raton, FL, p. 219.
2. Sloan, R. (2004). Slide: intrinsic biodegradation processes. In: *Aerobic and Anaerobic Bioremediation and Monitored Natural Attenuation of VOCs Short Course*. National Ground Water Association (NGWA) Conference on MTBE and Perchlorate, Costa Mesa, CA, June 3–4, p. 102.

BLACK MESA MONITORING PROGRAM

Bureau of Indian Affairs and
Arizona Department of Water
Resources—U.S. Geological
Survey

PROBLEM

The N aquifer is the major source of water for industrial and municipal users in the 5,400 square-mile Black Mesa area of northeastern Arizona. The aquifer consists of three rock formations-the Navajo Sandstone, Kayenta Formation, and Wingate Sandstone, which are hydraulically connected and function as a single aquifer.

Annual withdrawals from the N aquifer for industrial and municipal use have increased from about 70 acre-ft in 1965, to 4,300 acre-ft in 1972, to 7,700 acre-ft in 2000. The Navajo Nation and Hopi Tribe live in the Black Mesa area, and they depend on ground water from the N aquifer to

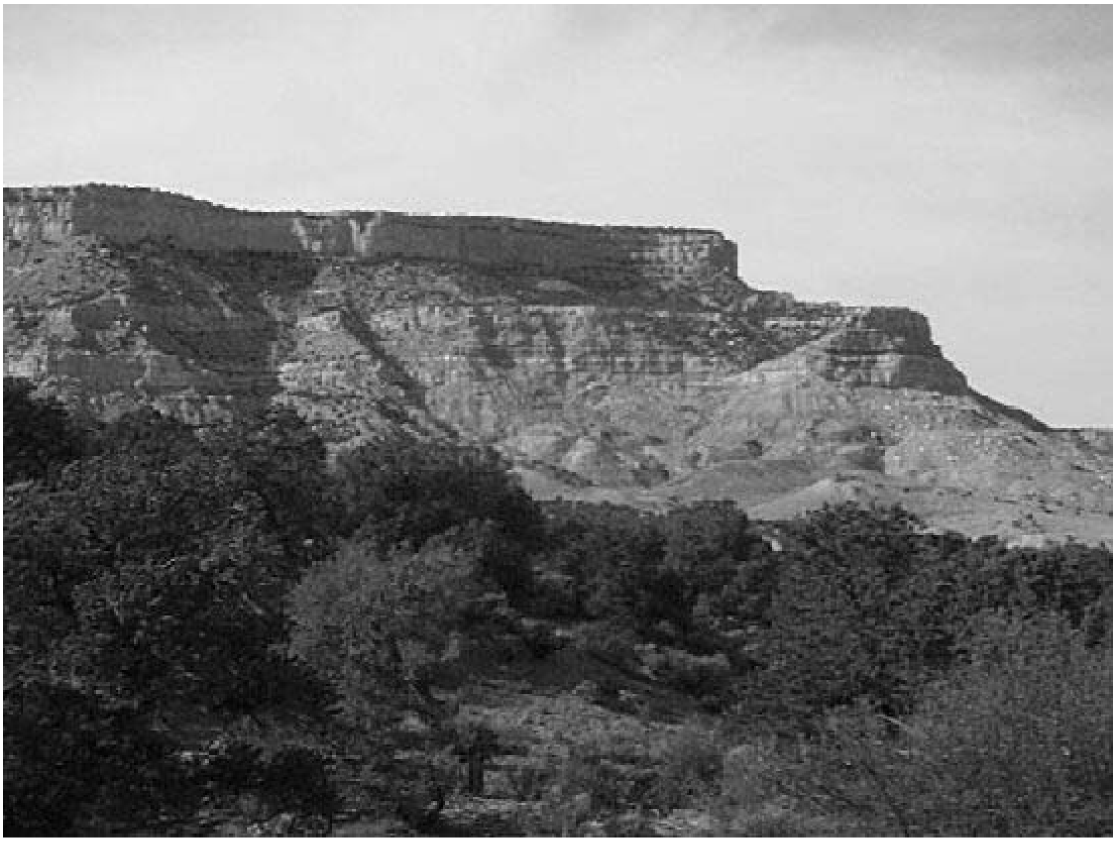

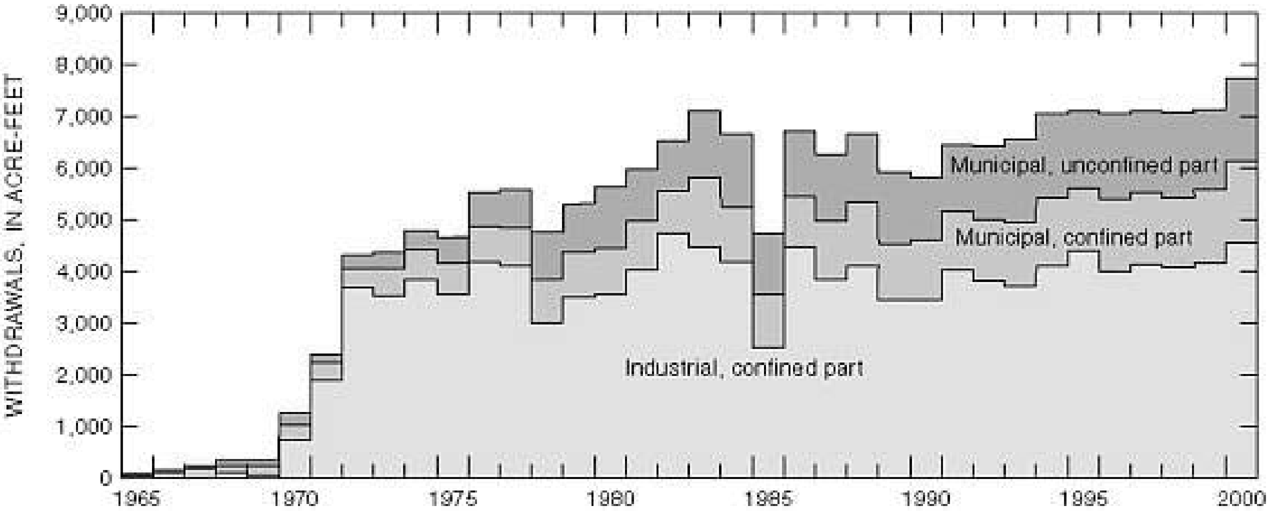

Withdrawals from the N aquifer, Black Mesa area, Arizona, 1965–2000.

HYDROGRAPHS FOR WELLS IN UNCONFINED AREA

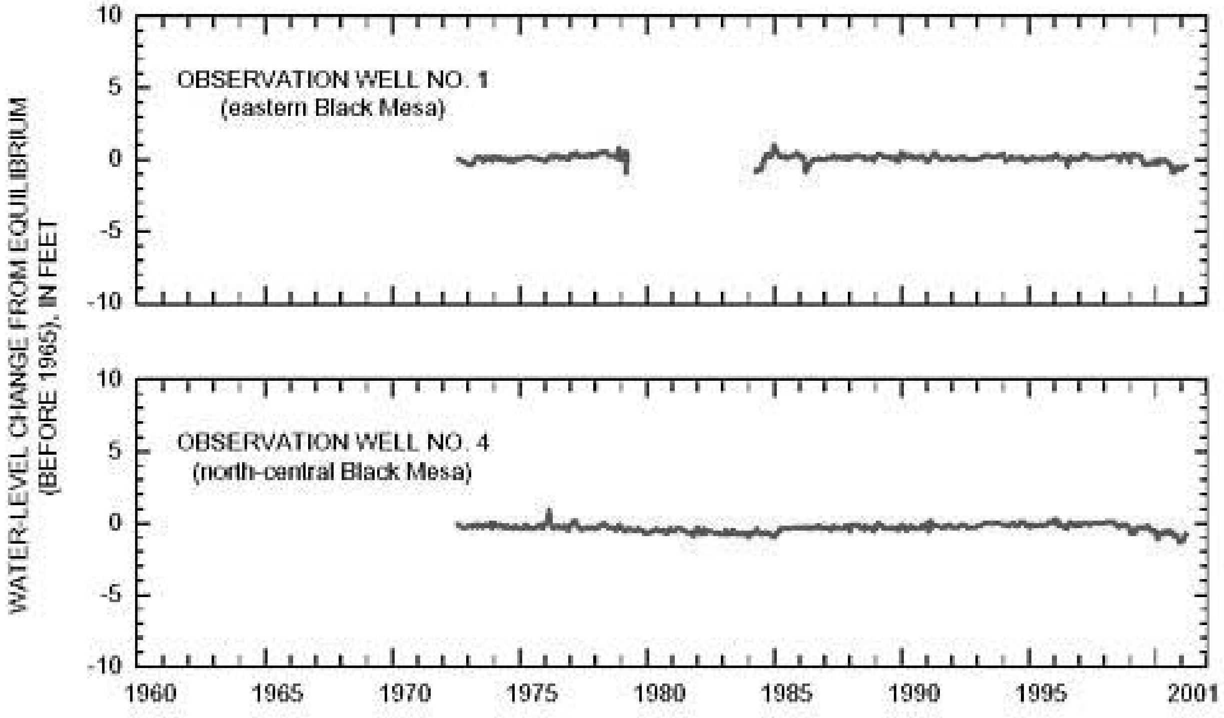

meet municipal, domestic, livestock, and irrigation needs. In addition, the springs and streams fed by groundwater discharge are an important part of their culture. Peabody Western Coal Company began operating a strip mine in the northern part of Black Mesa in 1968. The company withdraws water from the N aquifer for a slurry pipeline to transport coal to a powerplant in Laughlin, Nevada. Withdrawals by the company accounted for about 70–80 percent of the total withdrawals in the early and mid 1970's and have been about 60 percent of the total withdrawals from 1978 to 2000.

The Navajo Nation and Hopi Tribe have been concerned about the effect of this increasing pumpage on long-term water supply, discharge in streams and springs,

HYDROGRAPHS FOR WELLS IN CONFINED AREA

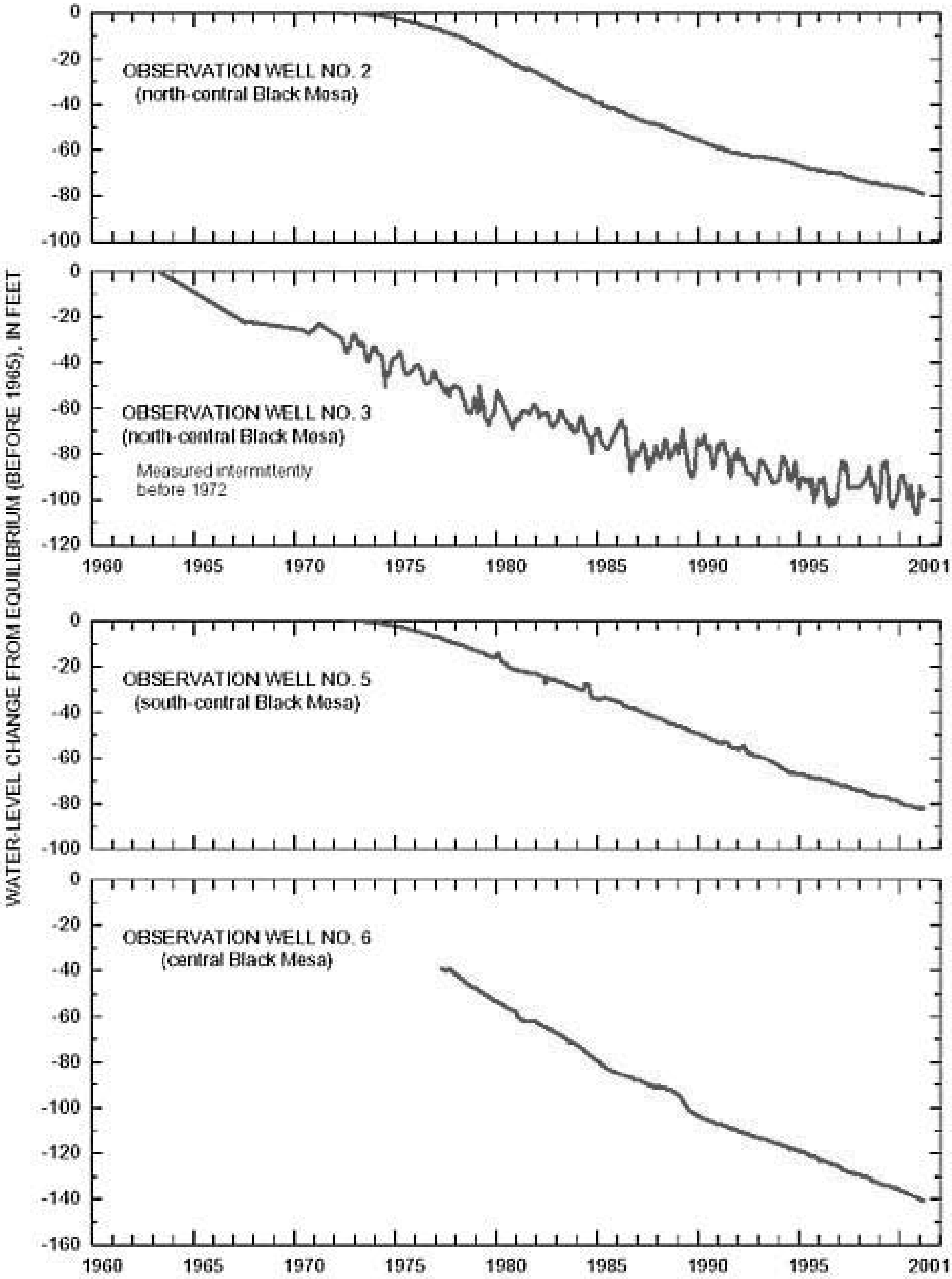

and quality of ground water. This concern led to the establishment of a long-term program in 1971 to monitor ground-water levels, ground-water discharge, ground-water quality, and surface-water discharge. From the 1960's to 2001, water levels in the N aquifer declined by over 50 feet in 9 of 19 long-term monitoring wells in confined areas. No large changes have been detected in ground-water levels in unconfined areas, ground-water discharge, ground-water quality, or surface-water discharge.

OBJECTIVE

To collect hydrologic data in a monitoring network that is designed to determine the long-term effects of industrial and municipal ground-water withdrawals on the N aquifer.

APPROACH

A long-term monitoring program has been established to collect hydrologic data in the Black Mesa area. Data are collected that describe the ground-water system, surface-water flow, and ground-water quality. Continuous measurements of ground-water levels have been made in six wells since 1972. Continuous data describing streamflow have been collected for 23 years in Moenkopi Wash, and for 3 to 6 years in three other streams. Once a year, ground-water levels are measured in about 26 wells, discharge is measured from 4 springs, and water-quality data are collected from 12 wells and 4 springs. Annual ground-water withdrawal data are collected from about 35 municipal well systems. Peabody Western Coal Company provides annual ground-water withdrawal data for industrial use.

These hydrologic data are entered into a computer data base. A report is prepared each year of the program. The report contains the data collected each year, and it shows comparisons of annual and long-term changes in ground-water levels, ground-water discharge, surface-water flow, and water quality. Long-term water-level changes in the six continuous-observation wells show that water levels in those confined areas have declined 80 to 140 feet and water levels in those unconfined areas have not changed.

RELEVANCE AND BENEFITS

The long-term available supply of water in the Black Mesa area is critical to many parties. The Hopi Tribe and Navajo Nation use water to meet their needs for public supply, irrigation, and livestock. In addition, sustained springflow and streamflow are important to their culture. Peabody Western Coal Company uses water to transport coal in a slurry pipeline to a powerplant. The hydrologic data collected in this monitoring program are needed to understand the available water supply and the effects of industrial and municipal ground-water withdrawals.

The U.S. Geological Survey has a commitment to provide data and information to Indian Tribes and other Interior Department Agencies. The Office of Surface Mining of the Interior Department uses the data from this program to facilitate their oversight and regulation of the coal mining operations of Peabody Western Coal Company.

The long-term ground-water, surface-water, and water-quality data collected for this program provides an important opportunity to investigate and gain a better understanding of a hydrologic system of bedrock geology in an arid climate in which there are many competing water-use interests.

BRINE DEPOSITS

Stephen M. Testa
Mokelumne Hill, California

INTRODUCTION

Brines are warm to hot, saturated to nearly saturated, highly saline, ocean and lake waters, containing Ca, Na, K, and Cl and minor amounts of other elements, typical of pore fluids in restricted basins. From an oceanographic perspective, brine is seawater that, due to high evaporation rates or freezing, contains more than the usual amount of dissolved salts, typically about 35%. A brine lake is essentially a salt lake. Brine lakes, like the Dead Sea, the Great Salt Lake in Utah, the Mono Lake region in northern California, and the Salton Sea region in southern California, develop as a result of high evaporation rates in an arid desert that lacks an outlet to the ocean (Fig. 1). The salt is derived from either minerals washed out of the surrounding watershed or from a geological deposit or area previously connected to the ocean.

A brine cell or pocket is a small inclusion usually the shape of an elongated tube about 0.05 mm in diameter that contains residual liquid more saline than seawater. A brine pit is a salt well or opening at the mouth of a salt

Figure 1. Mono Lake, a terminal brine lake situated in northeastern California. Tufa occurs at saline lakes worldwide, but not in the abundance and variety in shape as at Mono Lake. Mono brine shrimp are unique to this lake.

spring, from which water is taken to be evaporated for making salt.

Brines also occur in the subsurface notably as subsurface oil-field waters and geothermal mineralizing fluids. Subsurface brines in sandstone and other porous rocks are largely regarded as connate or buried seawater. Some brines also form locally by solution of rock salt beds.

SEAWATER BRINES

Seawater is relatively complex, but also relatively concentrated (1). Its principal ions are Na^+, Mg^{2+}, Ca^{2+}, K^+, Cl^- and $SO_4{}^{2-}$. Minor constituents include the elements Br, Sr, F, Ar, Li, Rb, I, Al, Fe, Zn, and Mo. Variable minor constituents include C, N, O, Si, P, and Ba. Seawater brines become more concentrated by evaporation, and in this process, deviation from ideal chemical behavior is greater (2). Microscopic cavities in such minerals as halite contain trapped brines representative of ancient seawater. Due to the protection of the brine provided by the halite crystal, such water has not undergone any significant chemical changes. The ionic ratios in these trapped brines are similar in composition to evaporating modern seawater.

Three properties of seawater and its brines predominant in evaporate production are concentration, density, and vapor pressure. Total salts are commonly stated in salinity, reflecting the resulting concentration from evaporation. The evaporation rate, however, indicates how much brine has undergone concentration via evaporation. Brine density is most important in its control of the circulation pattern in evaporate basins. The density of brines derived from seawater rises sharply from the 1.03 g/cm^3 of normal seawater to 1.29 g/cm^3 at the beginning of halite deposition. Vapor pressure is also important in its control of the rate of evaporation. Other properties include heat capacity, light transparency, and viscosity. Brines have substantially lower heat capacity, slightly lower thermal conductivity, and are less transparent to light than seawater.

BRINE OCCURRENCES

Geologically, brines are generated in a variety of environments and under varying sedimentary processes (2). Two main types of evaporate deposits are recognized: capillary evaporites and open-water evaporites (3). Capillary evaporites precipitate from interstitial brines that have a preexisting sedimentary layer. They contain remains of the host sediment, so capillary evaporites are always more or less dirty. Open-water evaporites precipitate from exposed water bodies. The detrital fraction, which is often eolian derived, makes up a small portion of these salt layers, so open-water evaporites are generally very clean.

The primary depositional settings are (1) coastal intertidal and supratidal zones called sabkhas, (2) small lagoons on coasts and atolls, (3) large deep-water marine basins, (4) sub-sea-level basins with inflow of marine water, and (5) nonmarine interior basins. Tectonically, these deposits occur along continental margins and shelves, interior shallow and deep cratonic basins, and rifted continental margins.

Brine deposits are typically extremely localized and reflect their unique conditions of formation. Brine deposits including salt are derived from four primary sources: seawater, sedimentary bedded deposits, surface playa deposits, and salt domes. Such areas and conditions include deserts of the southwestern United States, High Andean salares, Middle Eastern sabkhas, the East African Rift Valley, and subsurface Permo-Triassic deposits.

Recent studies have discovered brine pools on the seafloor within the Gulf of Mexico. These deposits are characterized by a distinct surface and shoreline. The pools are derived from the Luann Salt Layer and form in crater-like depressions that contain very concentrated brines and methane. These pools also support dense peripheral mussel beds.

Extensive evidence also exists that brines commonly form deep along midocean ridge hydrothermal systems. Such occurrences result from heating of sea-water-derived hydrothermal fluids at supercritical conditions, although it is not clearly understood how these brines are stored in the crust. One widely accepted model is the development of dense brines formed by formation of a two-layer system, whereas a recirculating brine layer underlies a single-phase seawater cell. If the temperature of the seawater cell remains high enough for substantial supercritical two-phase separation, the brine layer will grow. Should temperatures drop into the single-range region, the brine layer erodes.

Brines have also been recently discovered on other planets within our solar system. A growing body of evidence indicates that an ocean exists but is hidden beneath the icy crust of Jupiter's moon, Europa. Hydrated salts are indicated by spectral evidence and thermal evolution models of Europa's interior and laboratory studies of meteorites. The hydrated mineral deposits may reflect exposure of salty ocean water to the surface. On Mars, evaporate deposits may represent significant sinks of mobile anions and cations among the materials composing the Martian surface and upper crust. The nature of evaporate-precursor brines formed under Martian conditions is poorly understood at this time. Salts depress the freezing point of water significantly, so the presence of salts in the Martian soil could explain why water might be flowing in very high latitudes.

EVAPORITE MINERALOGY AND THE CONCEPT OF FACIES

Brines fall into three main types based on the predominating acid radical: chloride brines, sulfate brines, and alkali or volcanic brines. They include bitterns (natural brines) with other chlorides, bromides, iodides, and sulfates. The metallic ions in greatest abundance are sodium, magnesium, calcium, and potassium. Chloride type brines include many terrestrial brines as well as seawater. Terrestrial brines that contain chloride in excess of sodium are comparatively rare. Chloride type brines contain a higher proportion of sulfate than seawater. Seawater is an impure solution of sodium chloride and contains enough chloride to combine with all the sodium (the most abundant metallic ion) and part of the magnesium (the second most abundant metallic ion). The transition from chloride to sulfate type brines is gradational, whereby sulfate is the predominate acid radical. Alkali brines contain carbonate, sulfate, and borate; chloride is present in subordinate amounts.

Based on provenance, brines can be divided into marine brines that are concentrated from ocean water, continental brines concentrated from groundwater, and formation waters circulating in deeper bedrock horizons.

Evaporites are sedimentary deposits composed of minerals that have precipitated from brines concentrated during evaporation. These deposits are products of waters of inland desert basins, as well as interstitial waters of sediments located along ocean margins. Evaporites also occur from replacement of rocks that are not evaporites by evaporate minerals (4). Sedimentary processes vary significantly. Source fluids for evaporate precipitation can be classified by the degree of solute concentration of undersaturated brines. The three types distinguished include waters of low salinity or hypohaline waters, intermediate salinity or mesohaline waters, and high salinity or hypersaline brines.

Evaporites have a distinct and characteristic mineralogy. In the study of their geochemistry, origin, and geologic setting, mineralogy is the most important characteristic (2). Critical minerals derived from the evaporation of seawater are summarized in Table 1, along with the general facies sequence for marine evaporites. Facies in geological terms characterizes the critical variations or stages in the mineralogy of evaporite rocks during brine evaporation. Facies in this context essentially represent subdivisions of the evaporate process, recognizable through the appearance of new albeit critical minerals, regardless of proportion. For example, the mere presence of halite would characterize the assemblage as within the halite facies. The appearance of a potash-magnesia, halite, calcium sulfate, or calcium carbonate mineral is part of natural crystallization which occurs as water is removed during evaporation. This process is dependent on a balance between the proportions of various components in seawater and the corresponding mineral solubility in the brine. An extensive listing of the most common evaporate minerals is presented by Braitsch (5) and Sonnenfeld and Perthuisot (3).

Structurally, evaporates occur in two widely different structural forms: beds or lenses, and as salt structures which includes such features as bosses, plugs, ridges, and domes. Lenses and bed type deposits, such as the great Permian Basin, extend throughout parts of Kansas, Colorado, Oklahoma, Texas, and New Mexico. Nonbedded structures are common along or near the Gulf of Mexico. Bedded salt deposits vary in thickness but are typically of the order of 15 to 45 feet. The thickness of salt domes or diapirs remains uncertain.

BRINE RESOURCES

Known resources of evaporate and brine resources in the United States range from 100 years for potassium compound and iodine production, to unlimited, reflecting seawater as a source. The birthplace of subsurface brine production was in the Michigan Basin, an intracratonic basin that hosts thick successions of carbonates and evaporate-bearing strata.

Natural subsurface brines are the feedstock for a wide variety of industrial minerals and chemicals. For example, bromide is produced in Arkansas from the Upper Jurassic Smackover Formation, magnesium chloride and bromide are produced from the Devonian Detroit River Group in Michigan; lithium carbonate is produced in Chile and California from aquifers beneath playas and salares; iodide is produced from the Lower Pennsylvanian Morrow Formation in Oklahoma and coproduced in Japanese natural gas wells; and calcium chloride is produced near Slave Lake, Drumheller, and Brooks, Alberta.

Table 1. Summary of Evaporite Facies[a]

Facies	Subfacies	Mineralogy	Salinity, wt. %	Density, g/cm^3	Fraction Evaporated, wt. % H_2O	Concentration × Seawater, wt H_2O	Concentration × Seawater, Vol. of Brines
	Bittern subfacies	Bischoffite, tachyhydrite			99.2	120	
Potash-magnesia (supersaline)	Potash subfacies	Carnallite, sylvite, kainite, kieserite, halite, anhydrite or polyhalite	380	1.31	98.7	75	78.0
	$MgSO_4$ subfacies	Epsomite, bloedite, halite, polyhalite or anhydrite	375	1.29	98.4	65	68.0
Halite facies		Halite, anhydrite or gypsum	300	1.20	91.0	11.5	12.2
$CaSO_4$ facies		Gypsum or anhydrite, dolomite	150	1.10	72.0	3.5	3.6
Dolomite facies		Dolomite, calcite	35	1.02	0	1.0	1.0
$CaCO_3$ facies							
	(normal marine)	Calcite, aragonite	10	1.01	—		
	(brackish)		1	1.00	—		
	(terrestrial)		0	1.00	—		

[a] Modified after Reference 2.

BRINE PRODUCTION

Brines are extracted via several means, including underground mining, solar evaporation, and solution mining and mechanical evaporation. Underground halite deposits are conventionally mined by drilling and blasting. Vertical shafts are about 20 feet in diameter and extend to depths of 500 to more than 2000 feet below ground into the salt deposit. In the United States, 70% of the salt produced is extracted from natural or synthetic brines or seawater; the remaining is mined as a solid.

Products developed from brines include specialized industrial minerals used in the chemical industry for the manufacture of glass, fertilizers, pharmaceuticals, and batteries (6). Specialized industrial minerals produced include soda ash, potash, borax and boric acid, potassium chloride and sulfate, lithium, and nitrate and iodine. Salt produced from brine, and a small amount of dry salt, is used to produce chemicals such as chlorine gas and caustic soda (sodium hydroxide), among others. Deep unused brine-bearing aquifers are also used for to reduce greenhouse gas emissions via subsurface injection.

BIBLIOGRAPHY

1. Goldberg. (1965).
2. Holzer, W.T. (1979). Mineralogy of evaporites. In: *Mineralogical Society of America Short Course Notes*, *Marine Minerals*. Vol. 6, pp. 211–294.
3. Sonnenfeld, P. and Perthuisot, J.P. (1989). *Brines and Evaporites*. American Geophysical Union, Short Course in Geology, Vol. 3.
4. Friedman, G.M. (1978). Depositional environments of evaporite deposits. In: *Marine Evaporites*. Society of Economic Paleontologists and Mineralogists (SEPM) Short Course No. 4, W.E. Dean and B.C. Schreiber (Eds.). pp. 177–184.
5. Braitsch, O. (1971). *Salt Deposits Their Origin and Composition*. Springer-Verlag, New York.
6. Jensen, M.L. and Bateman, A.M. (1979). *Economic Mineral Deposits*, 3rd Edn. John Wiley & Sons, New York, pp. 553–562.

CONNATE WATER

Richard C. Brody
UC Berkeley
Berkeley, California

CHEMISTRY

Because of its long contact with rock material, connate water can change in chemical composition throughout the history of the rock and become highly mineralized (see Table 1). Connate water can be dense and saline compared with seawater; connate water salinities can range from 20 to more than 300 grams per liter. For comparison, seawater salinity is approximately 35 grams per liter total dissolved solids (1).

AGE

Some of the oldest connate water may stay beneath the land surface for millions of years, which is in contrast to the age of groundwater, which ranges from a few years or less to tens of thousands of years. Because of its long underground residence, connate water may move long distances, even though its velocity may be very low (3).

Table 1. Geochemical Composition of Connate Water Compared to Aquifer Water[a]

Chemical Component	Connate, mg/1	Aquifer, mg/1
Na^+	65,000	2,600
K^+	500	64
Mg^{2+}	28,000	210
Ca^{2+}	165,000	430
Ba^{2+}	0.1	0.1
Sr^{2+}	1,300	6.5
Fe^{2+}	0.1	0.1
Cl^-	138,360	5,100
Br^-	10	10
$SO_4{}^{2-}$	260	1,900
HCO_3	100	150
Organics	55	5
pH	5.8	7.65

[a] Ref. 2.

OIL RESERVOIR MONITORING

Crude oil is always produced with connate water, and the water-to-oil ratio is often greater than 10 to 1; therefore, the oil and gas exploration industry extensively monitors connate water (4). Connate water tends to rise toward the surface during oil and gas extraction, so resistivity logs track connate water movement to recognize and mitigate coning problems associated with excessive production yields (5). Distinguishing the difference between connate water and other types of underground water (e.g., meteoric water) through water analyses can aid in both oil and gas exploration. Variations in fluid composition can help to delimit reservoir boundaries and reveal the connectivity of different strata. Saline waters are generally more favorable for locating petroleum reservoirs. If the trapped fluid is saline, then there is less chance that any associated petroleum has been degraded by contact with meteoric water or flushed from the reservoir (1).

BIBLIOGRAPHY

1. Robin, R. (2001). *Oilfield Water Notes.* University of Saskatchewan. Available: http://www.usask.ca/geology/classes/geol463/46304.pdf.
2. Kane, R.D. (2000). InterCorr International, Inc. *Use of Portable Monitoring Units to Assess Microbial Activity, Corrosion and Souring in Water Handling and Injection Systems.* Available: http://www.corrosionsource.com/CS2000/session01/paper0107/paper0107.htm.
3. Bouwer, H. (1978). *Groundwater Hydrology*. McGraw-Hill, New York, p. 7.
4. IPEC. (2001). Integrated Petroleum Environmental Consortium. *Locating Oil-Water Interfaces In Process Vessels.* Available: http://ipec.utulsa.edu/Ipec/lopresti_abs.html.

5. JPT. (2002). Journal of Petroleum Hydrology. *Technology Update*. Available: http://www.spe.org/spe/cda/views/jpt/jptMaster/0,1513,1648_2300_5376_3,00.html.

CONSOLIDATED WATER BEARING ROCKS

Nitish Priyadarshi
Ranchi University
Ranchi, Jharkhand, India

INTRODUCTION

Consolidated rocks (otherwise known as bedrock) consist of rock and mineral particles of different sizes and shapes that have been welded together by heat and pressure or chemical reaction into a rock mass. Aquifers of this type are commonly composed of one or more of the following rocks: sandstone, limestone, granite, or lava. Water flows through these rocks through fractures, gas pores, and other openings in the rock.

CONSOLIDATED ROCKS

Consolidated rocks are usually classified according to their origin. There are three types: sedimentary, igneous, and metamorphic.

SEDIMENTARY ROCKS

Sedimentary rock is formed by the deposition of material suspended in water; the material may have been material weathered from older rocks, plant or animal remains, or precipitated chemicals. Important types of sedimentary rock aquifers include sandstone, carbonate rock, and conglomerate.

Sandstone

Sandstone is a cemented form of sand and gravel; the sand grains are cemented together when dissolved silica and calcium contained in the pore fluid precipitate. Sandstone shows great variation in water yielding capacity which is controlled chiefly by its texture and the nature of the cementing materials. Thus, whereas coarse-grained sandstone with rather imperfect cement may prove excellent aquifers, fine-grained varieties, especially those that are thoroughly cemented, may be the worst types from which no yield of water is possible (1).

Carbonate Rocks

Carbonate formations include limestone ($CaCO_3$) and dolomite (a mixture of $CaCO_3$ and $MgCO_3$). These deposits exhibit mostly secondary porosity due to fracturing and dissolution openings because $CaCO_3$ is soluble in rainwater. Limestones vary widely in density, porosity, and permeability, depending on the degree of consolidation and development of permeable zones after deposition. Openings in limestone may range from microscopic original pores to large solution caverns forming subterranean channels sufficiently large to carry the entire flow of a stream (2). Large springs are frequently found in limestone areas.

The dissolution of calcium carbonate by water causes prevailingly hard groundwater in limestone aquifers; by dissolving the rock, water, also tends to increase the pore space and permeability with time. Solution development of limestone forms a karst terrane, characterized by solution channels, closed depressions, subterranean drainage through sinkholes, and caves. Such regions normally contain large quantities of groundwater (3). Major limestone aquifers occur in the southeastern United States and in the Mediterranean area (4,5).

Conglomerates

Conglomerates like sandstone are cemented forms of sand gravel. As such, their porosity and yield have been reduced by the cement. Conglomerates have limited distribution and are unimportant as aquifers (6).

IGNEOUS ROCKS

Igneous rocks are formed from the cooling and solidification of molten magma originating in the earth's core. Important consolidated rock aquifers include basalts.

The important points in this context are as follows:

1. Basalts rocks are often vesicular. Vesicles, in them are of considerable size and number, and if they are interconnected (by fracture), they can serve as aquifers.
2. Contraction joints (like columnar joints) and other fractures, if present, also contribute to the porosity and permeability of basalts.
3. Basalts originate as lava flows, so sometimes, they may overlie buried valleys that offer good groundwater potential.

The Columbia River Plateau covering eastern Washington and Oregon, and Idaho, averages about 500 m in thickness and is one of the largest basalt deposits in the world. Basalt aquifers are critically important water sources for the Hawaiian Islands. Most of the largest springs in the United States are associated with basalt deposits.

Rhyolites are less permeable than basalt. Granites have very low porosity and permeability; what little exists is primarily due to fracturing. Although not important as aquifers, these materials are candidates for the host rock for high-level radioactive waste.

METAMORPHIC ROCKS

These are sedimentary or igneous rocks that have been altered by heat or pressure. They generally form poor aquifers.

BIBLIOGRAPHY

1. Singh, P. (1978). *Engineering and General Geology*. Katson, Ludhiana.

2. Brucker, R.W. et al. (1972). Role of vertical shafts in the movement of groundwater in carbonate aquifers. *Ground Water* **10**(6): 5–13.
3. LeBrand, H.E. and Stringfield, V.T. (1973). Karst hydrology—a review. *J. Hydrology* **20**: 97–120.
4. Burdon, D.J. and Papakis, N. (1963). *Handbook of Karst Hydrogeology with Special Reference to the Carbonate Aquifers of the Mediterranean Region*. Inst. Geology and Subsfc. Research, United Nations Spec. Fund, Karst Groundwater Inv., Athens.
5. LaMoreaux, P.E. et al. (1970). *Hydrology of Limestone Terranes, Annotated Bibliography of Carbonate Rocks*. Geological Survey of Alabama Bull. 94(A), Alabama.
6. Todd, D.K. (1995). *Ground Water Hydrology*. Wiley, Toronto, Canada, p. 51.

SENSITIVITY OF GROUNDWATER TO CONTAMINATION

Michael D. Trojan
Minnesota Pollution Control Agency
St. Paul, Minnesota

For this discussion, groundwater sensitivity is defined as *the likelihood for contaminants to reach a specified position in the groundwater system after introduction at some location above the specified position* (1,2). Sensitivity is a widely applied concept that has several uses. Sensitivity analyses often lead to the development of maps showing the relative sensitivity of groundwater across a geographic area. These maps are used as a screening tool for land-use decisions and have increasingly been used in source water protection. Sensitivity analyses are also used in program development, such as identifying areas where groundwater monitoring will occur. The concept of groundwater sensitivity also has educational value because hydrogeologic concepts are easily expressed and the audience can visualize relationships between land use and aquifer protection.

WHAT MAKES AN AQUIFER SENSITIVE TO CONTAMINATION?

Groundwater sensitivity applies to a variety of situations; each requires a different level of analysis. Sensitivity, vulnerability, and susceptibility, which are often defined differently, are used interchangeably in this discussion. The simplest scenario involves transport of a conservative contaminant in water from the land surface to groundwater. This scenario can be modified to consider land use, account for attenuation or retardation of contaminants, consider a point below the top of the saturated zone, or consider risk to human or ecological receptors at some point within an aquifer.

Under the scenario where a conservative contaminant is transported from the land surface to groundwater, factors that affect sensitivity include the permeability of geologic materials, the thickness of the unsaturated zone, and recharge. Highly permeable geologic materials transmit water quickly and thus increase sensitivity (see Hydraulic Conductivity/Transmissibility). In heterogeneous deposits, the geologic material that has the lowest permeability typically has the greatest effect on the rate of water movement, provided the geologic material is sufficiently thick, laterally continuous, and does not have extensive macroporosity. As the thickness of the unsaturated zone increases, sensitivity decreases because water can be stored and has a farther distance to travel. Sensitivity increases as the quantity of recharge increases because contaminants are transported by recharge water. Recharge is often computed from a water budget by subtracting evaporation from precipitation and assuming no runoff or soil storage.

Even in sensitive hydrogeologic settings, groundwater will not become contaminated unless there is a source of contamination. We can generalize about how groundwater quality will be affected by land use (see Land Use Impacts on Groundwater Quality). Agriculture can contribute significant quantities of nitrates and pesticides to an aquifer, nonsewered residential areas can contribute large quantities of nitrates and chlorides (3), and industrial and older sewered residential areas can contribute large quantities of volatile organic chemicals (VOCs), chlorides, and some metals such as manganese. Within each of these land uses, specific management practices further affect contaminant volumes. Several studies show direct, although not necessarily linear relationships, between the quantity of chemical applied at or near the land surface and concentrations in sensitive aquifers. Thus, for example, groundwater is more sensitive in areas that use greater fertilizer application rates.

Nearly all contaminants attenuate to some extent within the unsaturated zone. The extent and rate of attenuation depend on the properties of the contaminant and on the chemical, physical, and biological properties of the geologic materials (Table 1). These properties are too numerous to discuss individually, but adsorption and degradation are the two primary attenuation processes that affect contaminants. Contaminants that are not readily adsorbed percolate to groundwater and result in increased sensitivity. Except in highly weathered, acidic soils, these are chemicals that have a negative charge, such as chloride, nitrate, and acid herbicides (2,4-D, Dicamba). Contaminants that are quickly degraded biologically or chemically represent a lower risk to groundwater than more persistent contaminants. In addition to the properties of a chemical, degradation is affected by factors such as temperature, the presence of oxygen, and pH. For example, most halogenated contaminants are persistent except under reducing conditions, nitrate is persistent in the presence of oxygen, and benzene is persistent in the absence of oxygen.

We are often interested in the sensitivity at some point below the top of the saturated zone, such as at a well. In addition to having a greater distance to travel, a contaminant may be affected by geochemical changes within the saturated zone. Trojan et al. (4), for example, showed that in the glacial aquifers of Minnesota, nitrate was quickly attenuated in the upper few meters of groundwater, even though these aquifers were mapped

Table 1. Characteristics of Different Classes of Contaminants. This Information is Generalized and Varies within Each Contaminant Class and Between Different Geologic Deposits and Hydrologic Environments

Contaminant Class	Examples	Adsorption	Persistence	Toxicity
Halogenated organics with one or more benzene rings	PCBs, dioxins, many pesticides, chlorobenzenes	High, increases with greater extent of halogenation	High, increases with greater extent of halogenation	High, increases with greater extent of halogenation
Polyhalogenated aliphatics	Industrial solvents (TCE, PCE), some pesticides	Low, increases with greater extent of halogenation	Moderate, increases with greater extent of halogenation	Moderate, increases with greater extent of halogenation
Nonhalogenated polynuclear aromatics	Pyrene, benzo(a)pyrene	Moderate but increases rapidly with increasing molecular weight	Moderate but increases rapidly with increasing molecular weight	Moderate but increases rapidly with increasing molecular weight
Other nonhalogenated aromatics	Benzene, toluene	Low	Low to high, depending on geochemical environment	Low except for some chemicals, such as benzene
Nonhalogenated aliphatics	Oil, alkanes,	Moderate to high	Low to high, depending on geochemical environment	Low
Metals	Lead, copper	Moderate to high	High	Moderate to high
Nonmetals	Arsenic, boron	Moderate but varies widely	High	Moderate to high
Anions	Chloride, nitrate	Low	High	Low
Radionuclides	Cesium-137, radon-222	High	High	High

as sensitive to contamination. In this case, denitrification was the most likely cause for disappearance of the nitrate. In other scenarios, contaminants may be adsorbed within an aquifer.

Few sensitivity methods consider the risk to receptors. Risk analysis does not consider whether a chemical will reach groundwater but whether it will pose a risk to receptors. Risk is a function of dose and toxicity. Thus, consumption of 1,1,2-trichloroethene (TCE) at a concentration of 5 μg/L poses a greater risk than consumption at 1 μg/L; consumption of TCE at 5 μg/L [maximum contaminant level (MCL) = 5 μg/L] poses a greater risk than consumption of xylene at 5000 μg/L (MCL = 10,000 μg/L). Because we have to consider specific exposure points (e.g., a well), the quantity of contaminant being transported, and the contaminant toxicity, estimates of sensitivity based on risk can be very complicated.

Table 2 provides a summary of factors that affect groundwater sensitivity to contamination. Figure 1 provides a schematic showing how different factors affect sensitivity. Figure 2 shows how sensitivity varies when different receptor points or contaminants are considered.

METHODS FOR ASSESSING SENSITIVITY

The Commission on Geosciences, Environment, and Resources (2) provides an excellent discussion of methods for assessing sensitivity. The most commonly employed methods are overlay and index methods. These involve combining various physical properties of the hydrogeologic system. Each property is assigned a score or other sensitivity value based on perceived sensitivity. Overlay and index methods use many of the factors in Table 2. DRASTIC (5) is a widely employed index and overlay method that uses seven factors in the sensitivity assessment. Overlay and index methods can be relatively simple and may require small amounts of information. Sensitivity analyses using these methods often result in plan view maps depicting relative groundwater sensitivity, usually through a color-coded scheme. These maps are useful interpretive and screening tools. The scale of these maps is usually not appropriate for site-specific decisions, although they have been used for this purpose.

Analytic or numeric methods predict the time it takes contaminants to reach groundwater. These methods typically consist of mathematical models. Attenuation of chemicals is often considered. Examples include PRZM (6), GLEAMS (7), and LEACHM (8). An advantage of analytic and numeric methods is that they quantify the processes that affect the movement of water and contaminant. The accuracy of these methods depends on the quality of data used in the model.

Statistical methods use statistical techniques, such as regression analysis, to predict the likelihood of contamination. These methods require data on contaminant concentrations and may require additional information to derive sensitivity estimates. For example, concentrations of nitrate may be correlated with depth to water and sand content in the vadose zone. Because they use actual data, statistical methods can provide accurate estimates of sensitivity, information on variability in a sensitivity analysis, and estimates of certainty. They are, however, data intensive.

LIMITATIONS OF THE SENSITIVITY CONCEPT

In 1993, the Commission on Geosciences, Environment, and Resources prepared a report on ground water vulnerability assessments (2). The Commission, “in struggling with the manifold technical and practical difficulties

Table 2. Summary of Factors that Affect Groundwater Sensitivity

Property	Effect on Sensitivity	Low Sensitivity	High Sensitivity
	Hydrologic Factors		
Aquifer material	Rate at which water moves within an aquifer	Shale, most hard rocks, clay, silt	Limestone, sandstone, sand, gravel
Depth to bedrock	Distance that water must travel to reach groundwater	Large distances (>50 feet)	Short distances (<10 feet)
Depth to water	Distance that water must travel to reach groundwater	Large distances (>50 feet)	Short distances (<10 feet)
Recharge	Amount of water reaching groundwater	Small quantities (<1 inch)	Large quantities (>10 inches)
Soil material	Rate at which water moves to or within groundwater	Clay, silt, organic soils	Sand, gravel
Thickness of confining layers	Rate at which water moves to or within groundwater	Large thickness (>20 feet)	Small thickness (<5 feet)
Topography	Amount of water reaching groundwater	Upland areas where water runs off	Lowland areas where water accumulates
Type of bedrock	Rate at which water moves to or within groundwater	Unfractured shale and hard rocks	Limestone, sandstone, fractured rock
	Other Factors		
Soil/aquifer material	Attenuation of chemical	High organic or clay content	Low organic or clay content
Contaminant	Persistence, mobility, toxicity	Rapidly degraded, low mobility, low toxicity	Persistent, highly mobile, highly toxic
Land use	Amount and type of chemical released	Low chemical inputs (forest, grassland)	Large chemical input (row crop agriculture, nonsewered communities)

affecting the performance of vulnerability assessments today, nearly concluded that their limitations are so great as to be of no use in management decision-making." Though the concept of sensitivity is applicable to land-use decisions, program development, or education, there are two important concerns over their use.

First, the analyses and resulting products, usually maps, are only as good as the data that go into them. Sensitivity analyses generally do not result in collection of new data and rely instead on existing information. Geologic maps, soil maps, and climate data are often available, but values for depth to water and hydraulic conductivity may be difficult to find. Data become more limiting when factors such as specific contaminants, land use, and risk are considered.

Second, maps or other products developed from sensitivity analyses can be misused. Many people who make land-use decisions rely on these maps. They often assume that sensitive areas will become contaminated and less sensitive areas will not, and they are often not aware of the way the maps were developed. Examples of misuse include

- using a 1:100,000 sensitivity map to site a landfill;
- assuming that all areas mapped as sensitive have equal sensitivity;
- assuming that the same data were available throughout a mapped area; and
- using a sensitivity analysis based on a conservative tracer to predict sensitivity to contamination with pesticides.

Researchers have increasingly modified existing methods to provide better tools to local users of sensitivity maps. Rupert (9) improved the effectiveness of a modified DRASTIC ground water vulnerability map by calibrating the point rating schemes to actual groundwater quality data by using nonparametric statistical techniques and a geographic information system. Snyder et al. (10) used a particle tracking model in conjunction with DRASTIC to evaluate groundwater vulnerability. Erwin and Tesoriero (11) related the occurrence of elevated nitrate concentrations in samples from public supply wells to natural factors to assess aquifer susceptibility using logistic regression. Holtschlag and Luukkonen (12) used a vulnerability model that incorporated several contaminant properties to predict atrazine leaching. Although researchers have generally focused on improving the accuracy of sensitivity estimates, the problems discussed before can also be overcome by improving data accessibility, providing information about the reliability

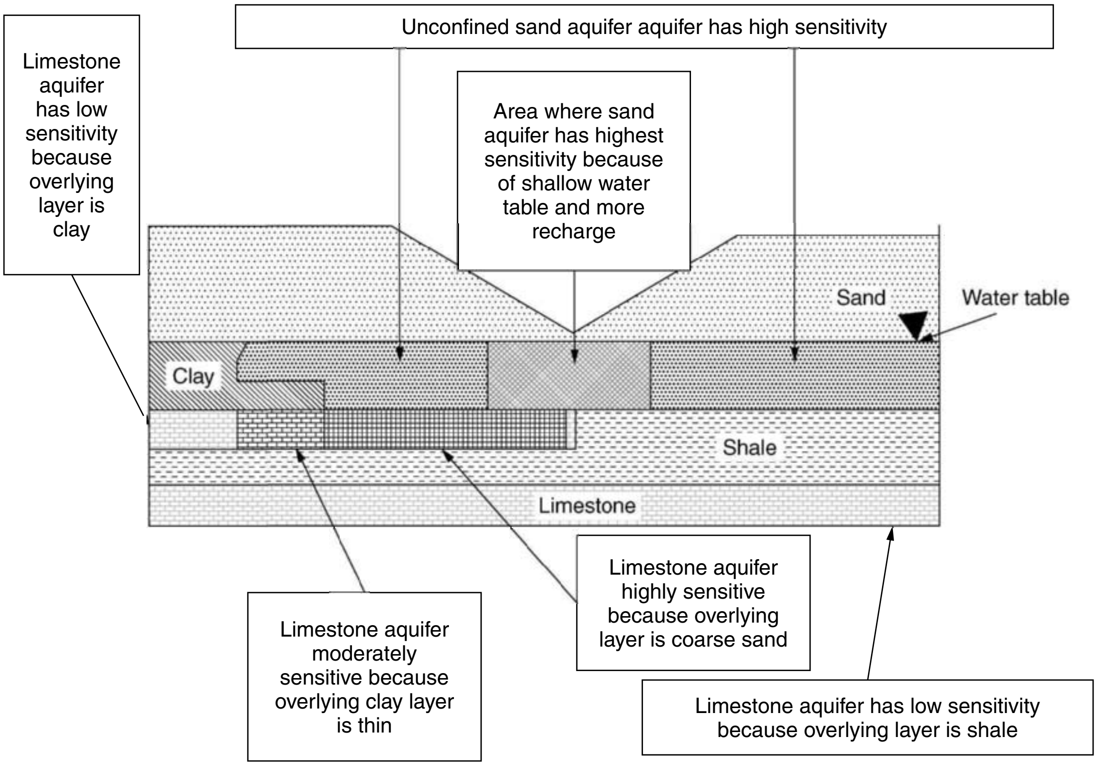

Figure 1. Schematic representation of geologic factors that affect groundwater sensitivity.

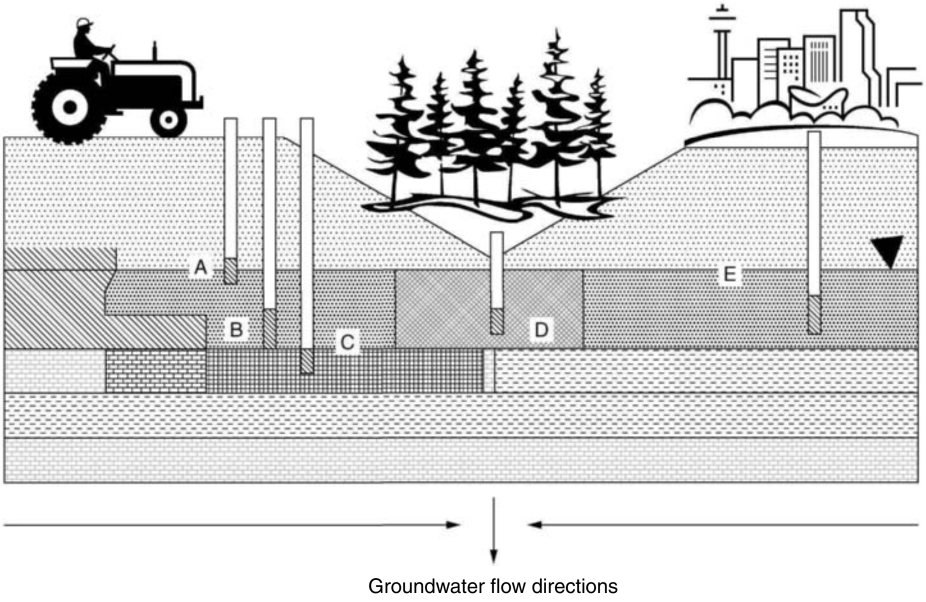

Figure 2. Schematic illustrating how a variety of factors affect groundwater sensitivity. Considering transport of a conservative chemical, the groundwater is considered sensitive at each of five well locations (A, B, C, D, E). At location A, the aquifer is most sensitive because the well is screened at the water table. The chemicals of concern are nitrate and pesticides from agriculture. At location B, the well may be less sensitive than at A because contaminants can be degraded or adsorbed within the aquifer. The chemicals of concern are nitrate and pesticides from agriculture. At location C, the well may be less sensitive than at B if the contaminant is degraded or adsorbed, but if the aquifer is fractured, sensitivity may be greater than at B. The chemicals of concern are nitrate and pesticides from agriculture. At location D, the well may be less sensitive than at all other well locations because the overlying land use does not contribute contaminants. Contaminants introduced at other locations may be degraded or adsorbed before reaching the well. At location E, sensitivity is similar to that at B but the contaminants of concern are VOCs and trace elements from commercial, industrial, and residential land use.

of data, providing estimates of variability or uncertainty, and training users.

SUMMARY

Understanding groundwater sensitivity and identifying areas where groundwater is sensitive to contamination is a potentially useful tool for managers, planners, and educators. Sensitivity analyses often lead to the production of maps showing relative sensitivity across a geographic area. Despite a potential for misuse, researchers have increasingly focused on improving the accuracy of these maps, including calibration, for local applications.

BIBLIOGRAPHY

1. United States Environmental Protection Agency. (1993). *A Review of Methods for Assessing Aquifer Sensitivity and Ground Water Vulnerability to Pesticide Contamination*. US EPA, Washington, DC.
2. Commission on Geosciences, Environment and Resources. (1993). *Ground Water Vulnerability Assessment: Predicting Relative Contamination Potential Under Conditions of Uncertainty*. National Academy Press, Washington, DC.
3. Trojan, M.D., Maloney, J.S., Stockinger, J.M., Eid, E.P., and Lahtinen, M.J. (2003). Effects of land use on ground water quality in the anoka sand plain aquifer of minnesota. *Ground Water*. **41**(4): 482–492.
4. Trojan, M.D., Campion, M.E., Maloney, J.S., Stockinger, J.M., and Eid, E.P. (2002). Estimating aquifer sensitivity to nitrate contamination using geochemical information. *Ground Water Monitoring and Remediation* **22**(4): 100–108.
5. Aller, L., Bennett, T., Lehr, J.H., Petty, R.J., and Hackett, G. (1987). *Drastic: A Standardized System for Evaluating Ground Water Pollution Potential Using Hydrogeologic Settings*. EPA-600/2-87-035. US EPA, Ada, OK.
6. Carsel, R.F., Smith, C.N., Mulkey, L.A., Dean, J.D., and Jowise, P. (1984). *User's Manual for the Pesticide Root Zone Model (PRZM): Release 1*. US EPA, Environmental Research Laboratory, Athens, GA.
7. Knisel, W.G., Leonard, R.A., and Davis, F.M. (1989). *GLEAMS User Manual. Southeast Watershed Research Laboratory*. Tifton, GA.
8. Wagenet, R.J. and Hutson, J.L. (1987). *LEACHM: A Finite-Difference Model for Simulating Water, Salt, and Pesticide Movement in the Plant Root Zone, Continuum 2*. New York State Resources Council, Cornell University, Ithaca, NY.
9. Rupert, M.G. (2001). Calibration of the drastic ground water vulnerability mapping method. *Ground Water* **39**(4): 625–630.
10. Snyder, D.T., Wilkinson, J.M., and Orzol, L.L. (1997). Use of A Ground-Water Flow Model With Particle Tracking to Evaluate Ground-Water Vulnerability, Clark County, Washington. USGS Water-Supply Paper 2488.
11. Erwin, M.L. and Tesoriero, A.J. (1997). *Predicting Ground-Water Vulnerability to Nitrate in the Puget Sound Basin*. USGS Fact Sheet FS-061-97.
12. Holtschlag, D.J. and Luukkonen, C.L. (1998). *Assessment of Ground-Water Vulnerability to Atrazine Leaching in Kent County, Michigan: Review, Comparison to Results of Other Studies, and Verification*. USGS Water-Resources Investigations Report 98–4006.

WATER CONTAMINATION BY LOW LEVEL ORGANIC WASTE COMPOUNDS IN THE HYDROLOGIC SYSTEM

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Timothy K. Parker
Groundwater Resources of
California
Sacramento, California

Laboratory methods have improved dramatically during the past decade with much lower detection levels enabling some interesting results of human impact on the condition of the hydrologic system. Low levels of organic contaminants are present in surface and groundwaters throughout the nation. Methyl tert butyl ether (MtBE) and some chlorinated solvents have been measured at low levels in public supply wells near rivers. Low levels of chloroform and other trihalomethanes produced as a result of the water disinfection process (also called disinfection byproducts) appear nearly ubiquitous in the urban environment. The source of disinfection byproducts is from the use and application of treated water in the urban environment for landscape irrigation, transport of household wastewater through leaky sewer systems, and also from leaky water pipe systems emanating form water treatment plants. An additional constituent which has low risk-based level, N-nitrosodimethylamine (NDMA), has also been found associated with water treatment plant disinfection byproducts.

A well-known nation wide study was conducted by the U.S. Geological Survey (USGS) in 1999 to 2000. The USGS collected surface water samples from 139 streams in 30 states across the lower 48 states (Fig. 1). The streams tended to be downstream of intense urban or livestock areas and consequently were suspected of contamination. The surface water samples were analyzed for organic wastewater contaminants (OWCs). The results of the surface water sample evaluation were published in Environmental Science & Technology (1) and in a USGS Open-File Report (2).

The USGS found that 80% of the streams sampled generally contained low concentrations of OWCs in the <1.0 μg/L range, impacted primarily by residential, industrial, and agricultural products. Seven chemical groups (steroids, nonprescription drugs, insect repellent, detergent metabolites, disinfectants, plasticizers, and fire retardants) were found in more than 60% of the stream samples, and three groups (detergent metabolites, steroids, and plasticizers) contributed to almost 80% of the total measured concentration. More than one-third (33 of 95 target OWCs) of the chemicals detected are known or are suspected of exhibiting at least weak hormonal activity and have the potential to disrupt normal endocrine function (Fig. 2).

One source of the chemicals identified is products and materials typically used in farming and industry. Additionally, a significant contribution was from residential

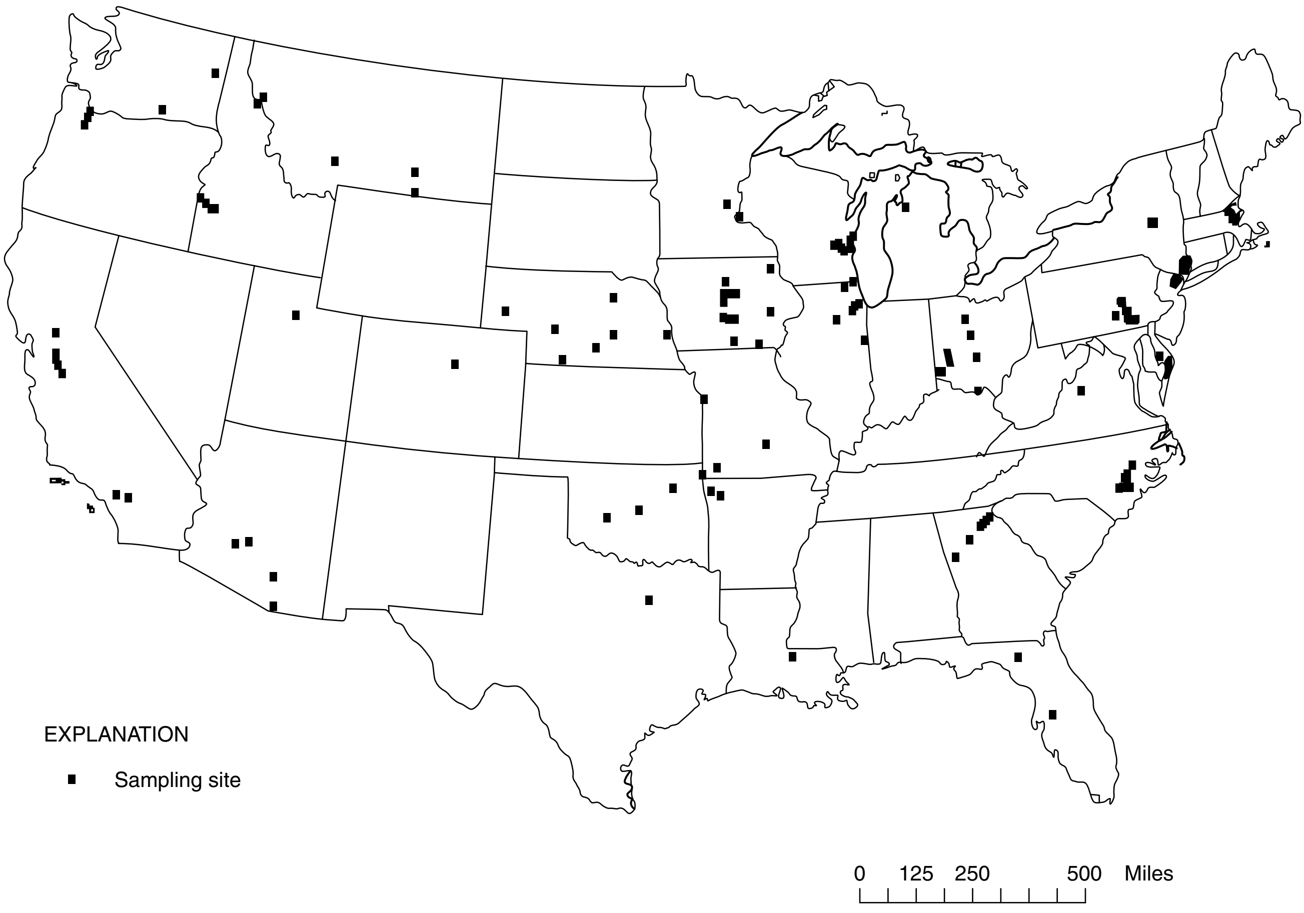

Figure 1. Sampling locations of USGS study (courtesy of USGS, Ref. 1).

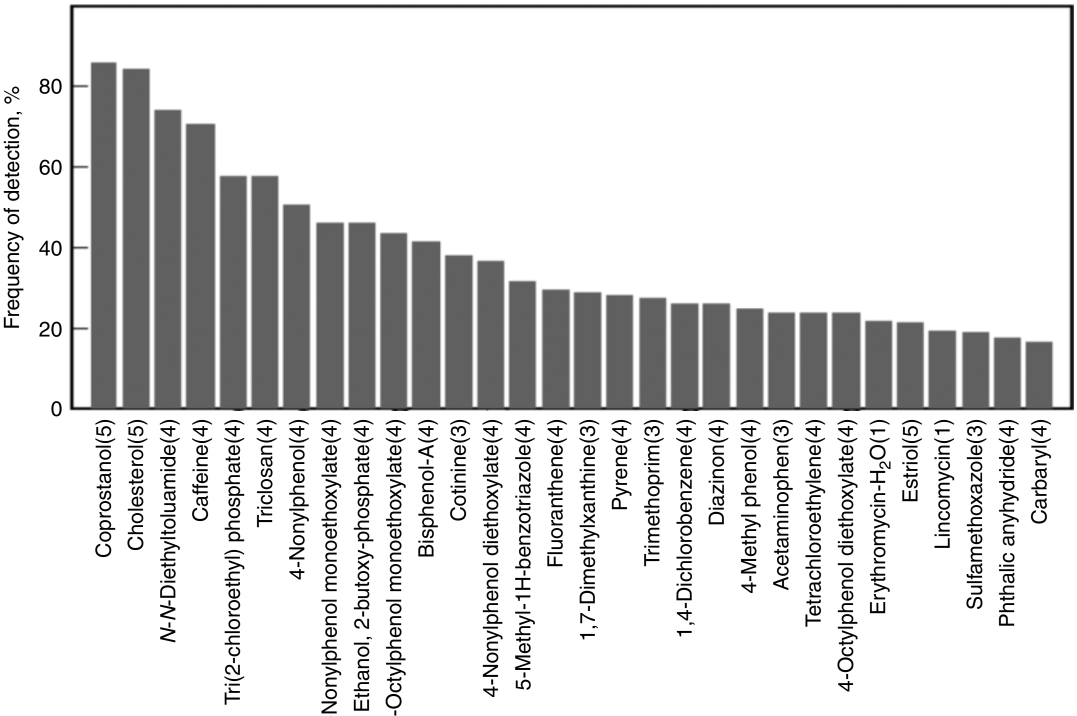

Figure 2. Most frequently detected compounds. The analytical method number is provided (in parentheses) at the end of each compound name (1).

use, drugs typically found in the medicine cabinet, and disposed of as is or as processed human waste in toilets and sinks. Some commonly detected constituents include caffeine in 70.6% of the samples at a maximum concentration of 5.7 μg/L, insect repellent *N*, *N*-diethyltoluamide in 74.1% of the samples at a maximum concentration of 1.1 μg/L, plant/animal steroid cholesterol in 84.3% of the samples at a maximum concentration of 60 μg/L, and fecal steroid coprostanol in 85.7% of the samples at a maximum concentration of 150 μg/L. The USGS study shows that many of the chemicals used in modern life survive human waste processing, current wastewater treatment plant processing, as well as biodegradation in the natural environment, and are detected as contaminants in potential water supplies.

A long list of contaminants have been documented in surface waters throughout the country partly due to the low laboratory detection levels. Considering the interconnection of surface water and groundwater, and the ability of these contaminants to move in the subsurface environment, our shallow groundwater resources are vulnerable to and likely impacted by contamination from these surface water sources. In California, the Ambient Groundwater Monitoring and Assessment Program (GAMA) implemented by the State Water Resources Control Board (SWRCB) is a program to assess the water quality and relative susceptibility of groundwater resources to contamination throughout the state. This assessment of relative contamination vulnerability is made based on the results of two types of analyses that are not routinely carried out at public water supply wells: ultra low-level measurement (part per trillion) of volatile organic compounds (VOCs), and groundwater age dating (using the tritium-helium-3 method). In addition, stable isotopes of oxygen are measured to help determine recharge water source location. Interpreted together with existing water quality and hydrogeologic data, these parameters assist in assessing the flow field of a groundwater basin, and indicate the degree of interconnection between near-surface sources (or potential sources) of contamination, and deeper groundwater pumped at high capacity production wells.

In GAMA studies that were conducted by the USGS and Lawrence Livermore National Laboratory under contract to SWRCB, the two most frequently detected volatile organic compounds were disinfection by products (trihalomethanes) dominantly chloroform, and methyl *tert* butyl ether (MtBE), which were found more frequently and at higher concentration in samples from rivers than in samples from nearby shallow public supply wells, suggesting the rivers may be the source of the groundwater occurrences (4,5). Additionally, chloroform was found associated with wellhead disinfection systems and may be present due to back-flushing of treated water (6). In one study of the Los Angeles Basin area, it was determined that the overall distribution of low level VOCs in groundwater was related to the hydrological and engineered recharge facilities (4). In a study of the Sacramento area, the frequent detection of one VOC (Perchloroethylene PCE) was believed to be due to mobilization of numerous point sources from spills or leaks at dry cleaners or other businesses, or possibly, from leaky sewer lines that hold small amounts of PCE (5).

Although these common chemicals are detected in very low or trace amounts, the possible cumulative or synergistic effects of these chemicals on plants, animals, and humans is currently unknown. Additional data from the USGS is forthcoming. There are no potential adverse health effects of these chemicals at such low documented concentrations.

If treatment or remediation is deemed appropriate, advanced oxidation, enhanced bioremediation, or improved water polishing/filtering could be added to existing wastewater treatment plants to reduce contaminants of these types from entering the environment. More limited pretreatment for septic systems could also help to reduce these chemicals in the environment. For individual water users, small point-of-use treatment systems could be used to remove trace amounts of contaminants (3).

This is a chemical society: Chemicals are widely used on a daily basis at a personal level in the household (personal care products, gasoline, solvents, pesticides, etc), and for a better quality of life in industry and agriculture (wide variety of chemicals). Anthropogenic impacts on the environment, which were largely uncontrolled until the last thirty years, are still unavoidable even with the current best standards of practice. Further, society is still continuing to understand the environmental impacts of past chemical handling and disposal practices as laboratory and investigative analytical tools improve. With population increasing at the current rate, projections are for increased demands on water supply, increased stress of the existing water supply and hydrologic systems, and shortages in dry years, unless steps are taken to increase water supplies and reliability through strategies including water conservation, recycling, desalination, conjunctive use, and artificial recharge. No doubt there will continue to be more “new” chemicals found to be present in our hydrologic system in the future that we did not know about before, but may have been present for some time. This will require planned adaptive management approaches and practices to assess the potential affects on humans and ecosystems, to develop alternative management strategies and remediation if appropriate, and to have on-going educational programs beginning at the elementary school level to edify society as to the information and risks so that society can make informed decisions as to how these challenges should be managed in the future.

BIBLIOGRAPHY

1. Kolpin, D.W. et al. (2002). Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999–2000: A national reconnaissance, USGS. *Environ. Sci. Technol.* **36**(6): 1202–1211.
2. Barnes, K.K., Kolpin, D.W., Meyer, M.T., Thurman, E.M., Furlong, E.T., Zaugg, S.D., and Barber, L.B. (2002). *Water-Quality Data for Pharmaceuticals, hormones, and Other Organic Waste Water Contaminants in U.S. Streams, 1999–2000*. USGS Open File Report 0-2-94, Iowa City, IA.
3. Jacobs, J.A. and Parker, T. (2002). *Low Concentrations of Organic Contaminants in the Hydrologic System*. *Groundwater*

Resourc. Assoc. 11th Annu. Meet., September 18, 2002, Newport Beach, CA.

4. Moran, J.E., Hudson, G.B., Eaton, G.F., and Leif, R. (2004). *Ambient Groundwater Monitoring and Assessment—A Contamination Vulnerability Assessment for the Sacramento Area Groundwater*. Lawrence Livermore National Laboratory UCLR-TR-203258.
5. Shelton, J.L., Burow, K.R., Belitz, K., Dubrovsky, N.M., Land, M., and Gronberg, J.A. (2000). *Low-Level Volatile Organic Compounds in Active Public Supply Wells as Groundwater—Water Tracers in the Los Angeles Physiographic Basin, California*. U.S. Geological Survey Water—Resources Investigations Report 01-4188.
6. Moran, J.E., Hudson, G.B., Eaton, G.F., and Leif, R. (2002). *A Contamination Vulnerability Assessment of the Livermore—Amador and Niles Cone Groundwater Basins*. Lawrence Livermore National Laboratory UCLR-TR-148831.

DARCY'S LAW

Joseph Holden
University of Leeds
Leeds, United Kingdom

Darcy's law is a mathematical relationship originally determined by Henry Darcy in 1856 that permits quantification of fluid flow through porous media. Many sophisticated theoretical and practical derivations have been based on Darcy's law (1), and it has become the foundation for quantitative groundwater flow science and soil physics (2). In its most useful form, the formula can be used to calculate the amount of water (or other fluid) flowing through a given cross-sectional area of a porous substance and equates volumetric discharge per unit time (q) to the product of area (A), hydraulic gradient (I = change in head per unit length) and a coefficient (K = saturated hydraulic conductivity). In its simplest terms, the relationship may be expressed as $q = KIA$.

Figure 1 provides a schematic diagram for establishing the relationship in a laboratory soil column test for one-dimensional flow. Note that $\Delta h/L = I$. For a given porous media in a tube of cross-sectional area A, the hydraulic conductivity, K, can be determined as the only unknown in the equation. By changing the hydraulic head, it is also possible to confirm that the relationship between head and discharge is linear for most materials. It should also be possible to determine the head or discharge conditions for which the relationship deviates from the linear form and therefore the threshold beyond which Darcy's law is no longer applicable.

Since it was determined, Darcy's law has been found valid for any Newtonian fluid. It is important to note, however, that the relationship holds only for laminar (nonturbulent) flow of fluids in homogeneous porous media and that it was established under saturated flow conditions for one-dimensional flow. The Darcy equation is strictly applicable only when the cross section being considered is much greater than the dimensions of its microstructure, so that it can reasonably be regarded as uniform. Most groundwater movement takes place in small interstices, so that an aquifer provides considerable resistance to water movement, allowing laminar flow.

The determination of laminar flow depends upon the magnitude of the Reynold's number, a dimensionless ratio of inertial forces to viscous forces. At low Reynold's numbers, viscous forces dominate, and Darcy's law is valid. There follows a transition zone in which inertial forces become more important; Darcy's law cannot be accurately applied to the nonlinear laminar flow in this zone (3). Flow in the turbulent zone is both nonlinear and nonlaminar, and deviations from Darcy's Law can become very large. In many aquifer materials, however, the assumption of laminar water flow may not cause any inaccuracies.

Flow analysis predictions based upon Darcy's law in the presence of massively fissured rocks such as karstic limestones or highly fractured crystalline rocks can lead to large errors. Flow in such cases cannot be described adequately by a linear relationship such as Darcy's law, and a more detailed analysis is required. In very large interstices such as those found in many limestone and volcanic areas, groundwater flow is almost identical to the turbulent flow of surface water (1). Others have claimed that in clays that have very small pores and low hydraulic conductivities the very low flow rates are less than proportional to the hydraulic gradient (4,5). A possible explanation is that much of the water in such material is strongly held by absorptive forces and may be more rigid and less mobile than ordinary water (6). Similar processes may explain the large store of water held in peat soils that have very low hydraulic conductivities (7,8).

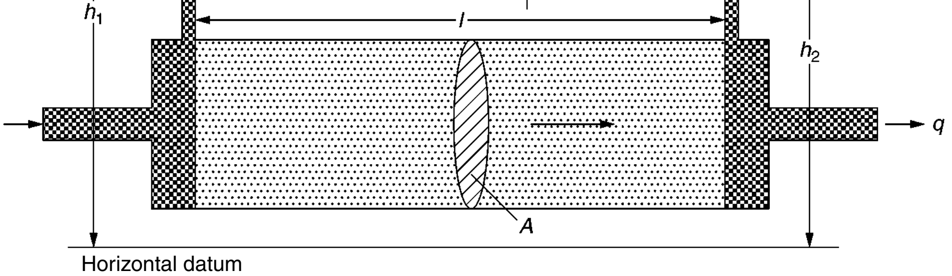

Figure 1. Schematic diagram of a soil column test for establishing components of Darcy's law.

As Hofmann and Hofmann (3) discuss, Darcy himself was careful not attribute too much to the specific magnitudes of individual data readings. Rather, he emphasized two aspects that his data supported, the proportionality of flow rate to total hydraulic head and the inverse proportionality of flow rate to column length (9). The limited experimental accuracy and scope were recognized by Darcy and have remained apparent to hydrologists ever since. By itself, Darcy's law suffices to describe only steady flow conditions, so that for most field applications it must be combined with mass conservation (or the continuity law) to obtain the general flow equation or, for saturated conditions, the Laplace equation. A direct solution of the latter for groundwater flow is generally not possible, and hence there is a need for various approximations or indirect methods of analysis. An important extension of Darcy's law to groundwater flow is its application in three-dimensions; see Childs (10) and Shaw (11) for more detail.

Fundamental to the application of Darcy's law is knowledge of the hydraulic conductivity of the saturated medium. The factors affecting hydraulic conductivity include those associated with the fluid and those associated with the aquifer, including temperature, salinity, pore space geometry, and soil or rock surface roughness (1). Hydraulic conductivity is often measured on a variety of scales ranging from a laboratory soil column, as shown in Fig. 1, to field-scale borehole tests and timed movement of tracers. As the scale of approach increases, often the estimation of hydraulic conductivity increases because of the incorporation of ever larger and more extensive fracture systems. A practical difficulty is that virtually any matrix considered, even on the laboratory scale, will be heterogeneous to some extent. Such heterogeneities do not theoretically prohibit the use of Darcy's law, but they lend uncertainty to generalizations on a large scale. It is for this reason that statistical models employing probabilistic parameter distributions are sometimes used to model the spatial variability of hydraulic conductivity (3).

However, particular problems emerge in assessing spatial variability in hydraulic conductivity and applying Darcy's law to groundwater flow when media are anisotropic, which is the case for highly layered soils such as peats where hydraulic conductivity may differ in the vertical direction by several orders of magnitude across just a few centimeters (8). Furthermore, full saturation of soils may be prevented by gases occupying pore spaces after entrapment upon rapid infiltration or gas production as a result of biogenic decomposition. This can lead to a reduction of hydraulic conductivity (12). Many wetland soils are compressible and so the field worker must be careful not to misapply rigid soil theories to these soils. Instead, compressible soil calculations have been developed but are not yet widely applied (8,13,14).

Despite concern about its strict validity, Darcy's law constitutes an adequate description of groundwater flow and fluid flow through porous media. It can be successfully applied to virtually all normal cases of groundwater flow and is equally applicable to confined and unconfined conditions (1).

BIBLIOGRAPHY

1. Ward, R.C. and Robinson, M. (2000). *Principles of Hydrology*. McGraw-Hill, London.
2. Freeze, R.A. (1994). Henry Darcy and the Fountains of Dijon. *Groundwater* **32**: 23–30.
3. Hofmann, J.R. and Hofmann, P.A. (1992). Darcy's law and structural explanation in hydrology. In: *PSA 1992: Proceedings of the 1992 Biennial Meeting of the Philosophy of Science Association*, Volume 1. D. Hull, M. Forbes, and K. Okruhlik (Eds.). Philosophy of Science Association, East Lansing, MI, pp. 23–35.
4. Millar, R.J. and Low, P.F. (1963). Threshold gradient for water flow in clay systems. *Proc. Soil Sci. Soc. Am.* **27**: 605–609.
5. Swartzendruber, D. (1962). Non Darcy behaviour in liquid saturated porous media. *J. Geophy. Res.* **67**: 5205–5213.
6. Hillel, D. (1982). *Introduction to Soil Physics*. Academic Press, New York.
7. Ingram, H.A.P. (1983). Hydrology. In: *Ecosystems of the World 4A, Mires: Swamp, Bog, Fen and Moor*. A.J.P. Gore (Ed.). Elsevier, Oxford, pp. 67–158.
8. Holden, J. and Burt, T.P. (2003). Hydraulic conductivity in upland blanket peat; measurement and variability. *Hydrological Proc.* **17**: 1227–1237.
9. Darcy, H. (1856). *Les Fontaines Publiques de la Ville de Dijon*. Victor Dalmont, Paris.
10. Childs, E.C. (1969). *Introduction to the Physical Basis of Soil Water Phenomenon*. John Wiley, London.
11. Shaw, E.M. (1994). *Hydrology in Practice*. Stanley Thornes, Cheltenham.
12. Beckwith, C.W. and Baird, A.J. (2001). The effects of biogenic gas bubbles on water flow in poorly decomposed blanket peat. *Water Resources Research* **37**: 551–558.
13. Premchitt, J. and Brand, E.W. (1981). Pore pressure equalisation of piezometers in compressible soils. *Geotechnique* **31**: 105–123.
14. Baird, A.J. and Gaffney, S.W. (1994). Cylindrical piezometer responses in a humified fen peat. *Nordic Hydrology* **25**: 167–182.

GROUNDWATER DATING WITH RADIOCARBON

Nitish Priyadarshi
Ranchi University
Ranchi, Jharkhand, India

INTRODUCTION

Groundwater age, based on measurement of the concentrations of chemical and isotopic substances in water, refers to the time elapsed since the introduction of the chemical or isotopic substance into the water, or to be more precise, the time elapsed since the chemical or isotopic substance was recharged and isolated from the modern reservoir. Although we often refer to dating of groundwater, we are actually dating a chemical substance that is dissolved in the groundwater, not the water itself. Unless we recognize and account for all the physical and chemical processes that affect the concentrations of tracers in an aquifer, the tracer-based age is not necessarily equal to the transit time of the water carrying the substance. Rather than referring to groundwater dating, it would probably be more correct to say we are interpreting temporal aspects of chemical and isotopic substances in groundwater.

For many applications in hydrology, we are interested in the travel time of water through an aquifer. The accuracy of the age determined depends in part on how perfectly the chemical and isotopic substances we are using as dating tools are transported in the aqueous phase. The concentrations of all tracers are, to some degree, affected by transport processes, and in the case of some tracers, such as chlorofluorocarbons, their concentrations can also be affected by chemical processes such as degradation and sorption during transit.

Various environmental isotopes and tracers are used to determine the age of groundwater. Carbon-14 is used to date ground waters older than 1000 years. Chloroflurocarbons (Freon) and tritium techniques are used to date groundwater that is less than 50 years old (Table 1).

The residence time of water underground has always been a topic of considerable speculation. But with the advent of radioisotopes, determination of the age of groundwater has become possible (1).

Carbon-14 (^{14}C), commonly referred to as radiocarbon, is continuously being produced in the atmosphere. This results from cosmic ray bombardment of nitrogen nuclei. Atmospheric testing of thermonuclear weapons doubled the ^{14}C concentration in the atmosphere. Carbon-14 is expressed in terms of percent of modern carbon (pmC). ^{14}C is useful when dating groundwaters that are between 1000 and 30,000 years old.

This isotope is present in groundwater as dissolved bicarbonate originating from the biologically active layers of the soil where CO_2 is generated by root respiration and the decay of humus (2).

Table 1. List of Environmental Tracers and Isotopes Used for Age Determination

Environmental Isotope/Tracer	Age Range, Years
Chloroflurocarbons (CFC-11, CFC-12, and CFC-113)	0 to 50
Tritium	0 to 50
Tritium/Helium-3	0 to 30
Carbon-14	1000 to 30,000

^{14}C

The ^{14}C generated in the atmosphere is carried down to the earth's surface by precipitation and becomes incorporated into the biomass or transported into waterbodies such as lakes, oceans, and groundwater. ^{14}C undergoes radioactive decay (to ^{14}N), so that once isolated from the atmosphere, the amount of ^{14}C decreases with time according to the equation

$$(^{14}\mathrm{C})_t = (^{14}\mathrm{C})_o e^{-kt}$$

where $(^{14}\mathrm{C})_t$ is the amount present at time t, $(^{14}\mathrm{C})_o$ is the amount present at $t = 0$, and k is the decay constant, which is related to the half-life $t_{1/2}$ by the equation

$$t_{1/2} = \ln 2/k$$

To determine the time since water lost contact with the atmosphere, it is necessary to know $(^{14}\mathrm{C})_o$. This is determined by tree rings for the most recent 7000 years; there is no accurate way to determine it prior to 7000 years ago, so it is generally assumed arbitrarily to have been constant. This gives rise to a timescale in "^{14}C years," which may be different from astronomical years (3).

There are some complications in the behavior of ^{14}C during recharge, so that the "absolute" age of a groundwater cannot be determined reliably. However, if the ^{14}C concentration is measured at several points along a flow line within an aquifer, the differences in age between the points and hence the flow velocity can be determined. One complication is that dissolution of carbonate minerals or oxidation of organic matter within the aquifer may add "old" or "dead" (no detectable ^{14}C) carbon to the water and give an erroneously old age. The contribution of carbon from these sources can sometimes be estimated from ^{13}C/^{12}C measurements and chemical arguments (4), so that corrections can be made. Another complication is mixing. A low ^{14}C concentration may mean that we are looking at relatively "old" water, or it may mean that we are looking at a mixture of relatively "young" water and "dead" water. ^{14}C measurements can be interpreted as ages only when mixing is insignificant.

Measurements of water samples taken from deep wells in deserts of the United Arab Republic and Saudi Arabia indicate ages of 20,000 to 30,000 years (5). This period is compatible with the Wisconsin Ice Age, when these desert areas last had high rainfall capable of recharging the underlying major aquifers.

BIBLIOGRAPHY

1. Todd, D.K. (1995). *Ground Water Hydrology*. Wiley, Toronto, Canada, pp. 24–25.
2. Wigley, T.M.L. (1975). Carbon-14 dating of groundwater from closed and open systems. *Water Resour. Res.* **11**: 324–328.
3. Drever, J.I. (1982). *The Geochemistry of Natural Waters*. Prentice-Hall, Englewood Cliffs, NJ.
4. Wigley, T.M.L. (1976). Effect of mineral precipitation on isotopic composition and ^{14}C dating of ground water. *Nature* **263**: 219–221.
5. Thatcher, L. et al. (1961). Dating desert ground water. *Science* **134**(3472): 105–106.

GROUNDWATER DATING WITH H–HE

Craig E. Divine
John D. Humphrey
Colorado School of Mines
Golden, Colorado

PRODUCTION AND DECAY BEHAVIOR OF ATMOSPHERIC TRITIUM

Tritium (^{3}H) is a relatively short-lived radioactive isotope of hydrogen that is produced naturally in the atmosphere and enters the hydrologic cycle through meteoric precipitation. Natural atmospheric ^{3}H results from the interaction of nitrogen-14 (^{14}N) and cosmic-ray neutrons (n) by

$$^{14}\mathrm{N} + n \longrightarrow {}^3\mathrm{H} + {}^{12}\mathrm{C} \tag{1}$$

Tritium is then rapidly incorporated into water molecules. Tritium decays to helium-3 (^{3}He) by beta decay:

$$^3\mathrm{H} \longrightarrow {}^3\ \mathrm{He} + \beta^- \tag{2}$$

Tritium concentrations in water are typically reported as "tritium units" (TU), where 1 TU is equivalent to one ^{3}H atom per 10^{18} hydrogen atoms and 0.118 Bq kg^{-1} (1 Bq = 1 disintegration per second). Decay of 1 TU yields approximately 0.402 pcm^3 kg^{-1} of ^{3}He. The natural atmospheric production rate of ^{3}H is approximately 0.5 ± 0.3 atoms of ^{3}H cm^{-2} s^{-1} (1). The background ^{3}H concentration in meteoric precipitation before 1951 is estimated in the range between 0.5 and 20 TU, and most measurements are less than 10 TU (2) and (3, Fig. 7.2). The rate of ^{3}H decay and ^{3}He production follow these first-order rate laws:

$$^3\mathrm{H}_\mathrm{t} = {}^3\mathrm{H}_0 e^{-\lambda t} \tag{3a}$$

$$^3\mathrm{He}_\mathrm{t} = {}^3\mathrm{H}_0 e^{\lambda t} \tag{3b}$$

where $^3\mathrm{H}_\mathrm{t}$ and $^3\mathrm{He}_\mathrm{t}$ are the number of tritium and helium-3 atoms at time t, $^3\mathrm{H}_0$ is the initial number of tritium atoms, and λ is the tritium decay constant (0.05626 yr^{-1}). The tritium "half-life" ($t_{1/2}$), or time it takes for half of the starting tritium to decay to ^{3}He, is best estimated at 12.32 yr (4500 ± 8 day) (4). Figure 1 graphically shows the decay and generation behavior of ^{3}H and ^{3}He through time. The half-life is related to the decay constant by

$$\lambda = \frac{\ln 2}{t_{1/2}} \tag{4}$$

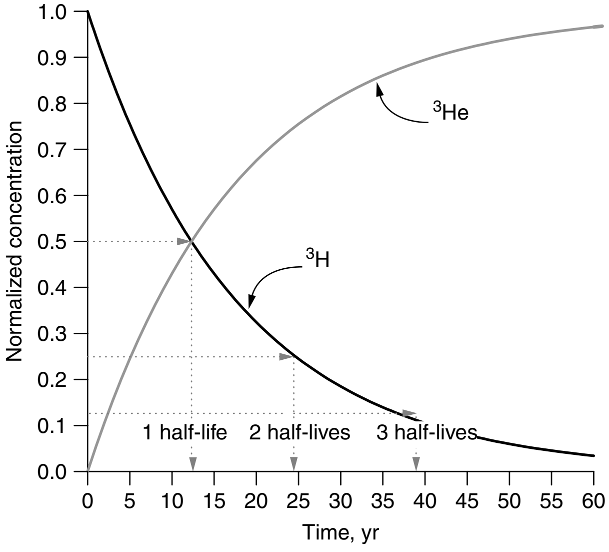

Figure 1. Tritium (^{3}H) decay and ^{3}He production through time. Values are normalized to the initial ^{3}H concentration ($^3\mathrm{H}_0$).

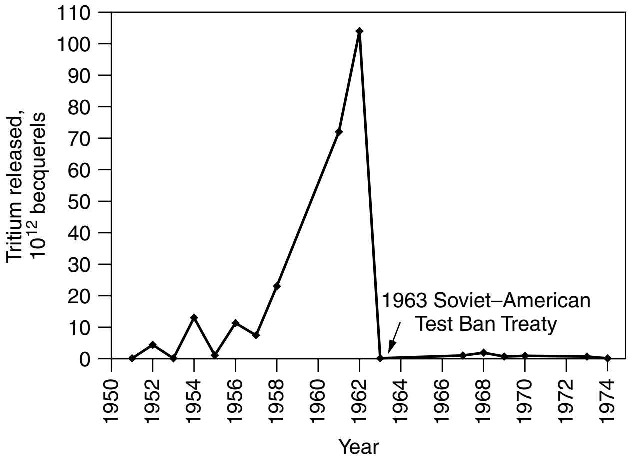

Figure 2. Anthropogenic ^{3}H released from atmospheric detonation of thermonuclear weapons testing (developed from data summarized by Reference 3).

During 1951, it was demonstrated that significant energy is released during hydrogen fusion to provide the basis for developing thermonuclear superbombs. Atmospheric testing (detonation) of these weapons during the next decade resulted in a dramatic increase in atmospheric ^{3}H concentration, until implementation of a Soviet–American atmospheric test ban treaty in 1963 (Fig. 2). Atmospheric thermonuclear weapons testing released vast quantities of high-energy neutrons that interacted with ^{14}N and led to the production of ^{3}H (Eq. 1). The incredible amount of ^{3}H produced in the atmosphere resulted in a distinct precipitation input signal of anthropogenic ^{3}H (Fig. 3), and a notable peak in 1963.

PRINCIPLES OF THE ^{3}H–^{3}HE DATING METHOD

The mid-1960s ^{3}H peak has been widely used as a direct time marker to determine groundwater ages. However, because of the relatively short half-life of ^{3}H, this peak has become difficult to differentiate in many systems. This is particularly true in the Southern Hemisphere where the anthropogenic ^{3}H in precipitation was notably lower than in the Northern Hemisphere. However, measurement of both parent ^{3}H and increases in its decay product, ^{3}He, can dramatically increase the sensitivity of the method. The "in-grown" ^{3}He (or ^{3}He produced by ^{3}H decay) can be equated to the initial $^3\mathrm{H}_0$ and $^3\mathrm{H}_\mathrm{t}$ by

$$^3\mathrm{He}_\mathrm{t} = {}^3\mathrm{H}_0(1 - e^{-\lambda t}) = {}^3\mathrm{H}_\mathrm{t}(e^{\lambda t} - 1) \tag{5}$$

Assuming that other ^{3}He sources are insignificant or can be accounted for and that there have been no systematic losses of tritiogenic ^{3}He, the relationships in Equation 5 can be rearranged, and the apparent age of the water (τ) can then be calculated from (Fig. 4)

$$\tau = \frac{t_{1/2}}{\ln 2} \ln\left(1 + \frac{^3\mathrm{He}_\mathrm{t}}{^3\mathrm{H}_\mathrm{t}}\right) \tag{6}$$

The ^{3}H–^{3}He method offers two significant advantages: (a) the method does not rely on accurate knowledge of the

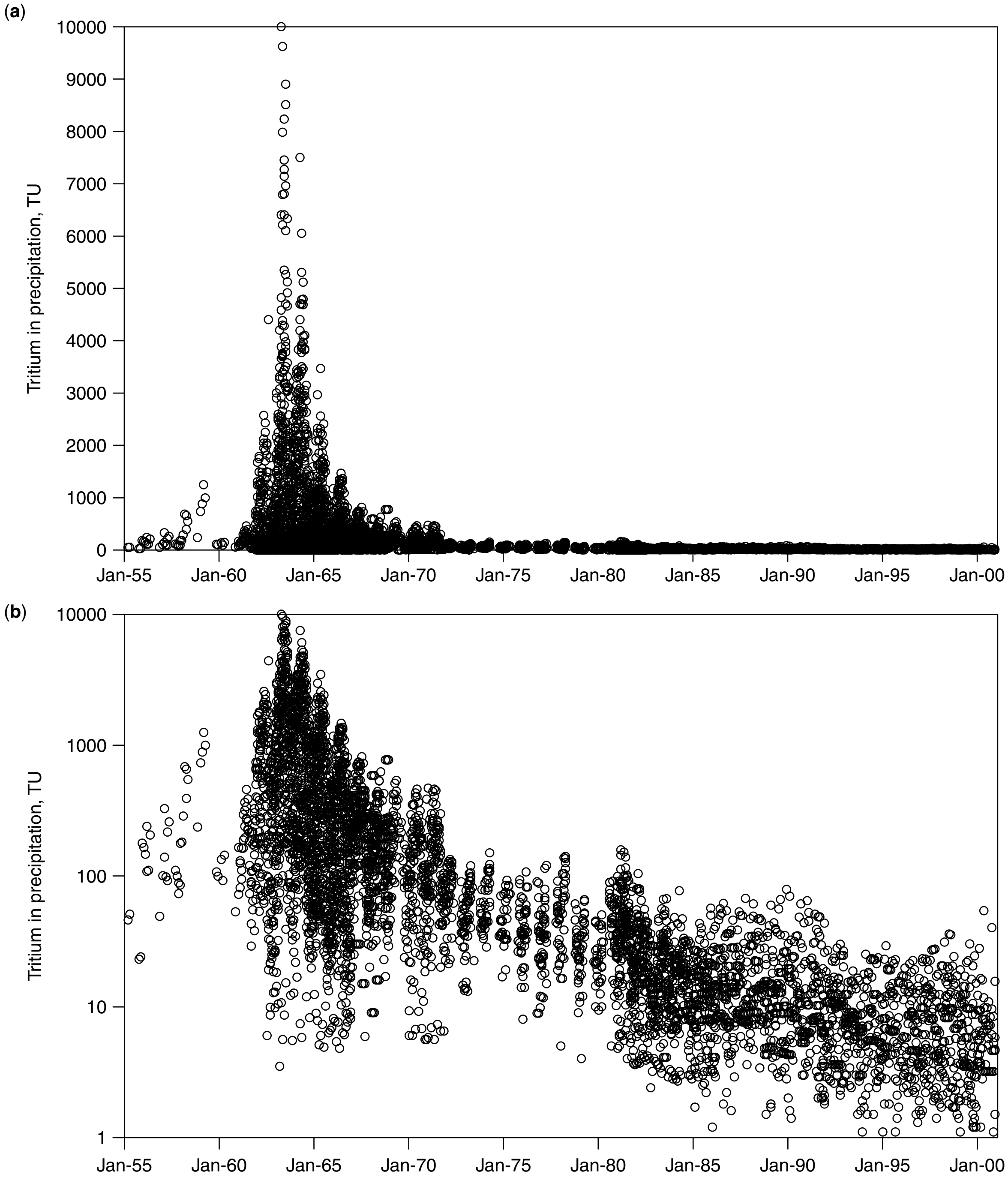

Figure 3. Tritium in precipitation (reported in tritium units, TU) measured at more than 50 locations in North America between 1955 and 2001 (top: arithmetic scale, bottom: logarithmic scale) (data obtained from Reference 5).

original ^{3}H input or identification of the mid-1960s peak, and (b) the method is generally less sensitive to analytical measurement errors (2).

The “clock” for the ^{3}H–^{3}He method begins at the seasonal water table low, where ^{3}He produced is no longer lost to the atmosphere and begins to accumulate in the groundwater. If the unsaturated zone is thick, the apparent ^{3}H–^{3}He age at the “bomb peak” signal may be significantly less than the actual time since the peak ^{3}H fallout. The difference in apparent ages can be used to estimate the travel time through the unsaturated zone. For example, Solomon et al. (6) determined an apparent ^{3}H–^{3}He age of 16 years for groundwater samples near the ^{3}H peak signal in a study at Cape Cod, Massachusetts. However, the samples were collected 30 years after the actual peak ^{3}H fallout in 1963. The difference in these age values resulted in an unsaturated zone travel time estimate of 14 years.

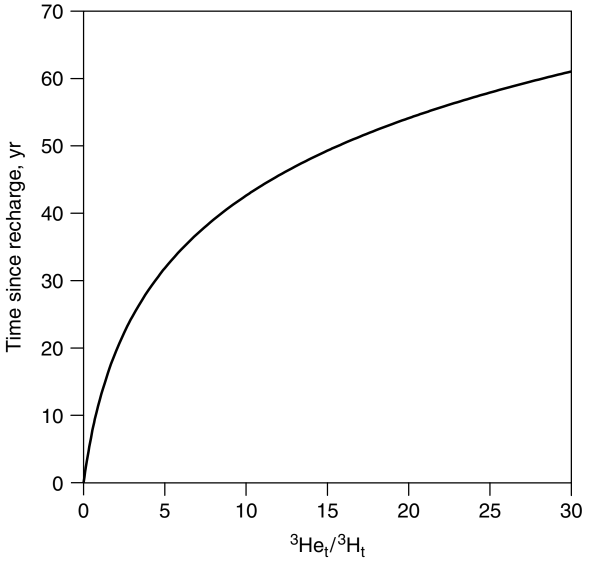

Figure 4. Calculated time since recharge from measured $^3He_t/^3H_t$ ratio (Eq. 6).

There are additional sources of ^{3}He in groundwater, and the measured ^{3}He value must be corrected for these sources before applying Eq. 6. Tritium-generated ^{3}He ($^3He_{trit}$) is determined from total measured tritium ($^3He_{total}$) by

$$^3He_{trit} = {}^3He_{total} - {}^3He_{atm} - {}^3He_{excess} - {}^3He_{crust} - {}^3He_{mantle} \quad (7)$$

where $^3He_{atm}$ is the ^{3}He concentration in solubility equilibrium with the atmosphere, $^3He_{excess}$ is the ^{3}He resulting from excess air, $^3He_{crust}$ is the ^{3}He produced in the crust, and $^3He_{mantle}$ is the ^{3}He produced in the mantle (also called primordial ^{3}He). Atmospheric ^{3}He dissolving in groundwater at the time of recharge is typically the most significant process resulting in excess ^{3}He. This process can be accounted for by considering the following (3):

- The atmospheric ^{4}He concentration is 5.24 parts per million (by volume), and the atmospheric $^3He/^4He$ ratio is 1.3×10^{-6}.
- The temperature-dependent aqueous solubility of helium (at 10 °C, the value is 4.75×10^{-8} cm^3 STP/cm^3 H_2O).
- ^{4}He is more soluble than ^{3}He in water and has a fractionation factor ($\alpha_{water-air}$) of 0.983.

Although the $^3H-^3He$ method is fairly insensitive to recharge temperatures, it is very sensitive to excess air, particularly for young groundwater (7). In some systems, ^{3}He from excess trapped air caused by a transient wetting front may be significant and can be identified by supersaturation of other gases, such as Ne and N_2. For most systems, subsurface generated ^{3}He ($^3He_{crust}$ and $^3He_{mantle}$) is insignificant; however, correction methods for these cases are discussed by Schlosser (8,9), Solomon et al. (2), and Cook and Solomon (7). Other potential errors in the $^3H-^3He$ method may be caused by diffusive loss and dispersion/mixing effects. ^{3}He diffusive loss to the atmosphere was greatest during the mid-1960 peak (due to the high concentration gradient), so this method may incorrectly calculate younger ages for groundwater below the mid-1960 peak (2). Because helium has a relatively high Henry's law constant (105.2 [dimensionless] at 25 °C) (10), care must be taken during sampling to prevent ^{3}He loss due to gas stripping.

ANALYTICAL MEASUREMENT METHODS

Tritium is typically measured by either low-level disintegration counting or by the more sensitive ^{3}He in growth method using mass spectrometry. Depending on the specific instruments and analytical procedures, ^{3}H detection limits of ~0.05–0.8 TU and analytical precision of ±2.5–5% can be achieved (9). For ^{3}He measurement, water samples are commonly collected in pinched-off copper tubes or gas-filled diffusion samplers. The helium is then isolated from all other dissolved gases and analyzed by a helium isotope mass spectrometer to resolve the $^3He/^4He$ isotopic ratio. Achievable measurement precision for the $^3He/^4He$ ratio is approximately 0.2% (9). Although approximately 40 years have passed since the mid-1960 ^{3}H "bomb peak," these analytical capabilities should permit identifying the peak signal in favorable hydrogeologic settings for several more decades (8). In most cases, analytical uncertainties result in calculated age uncertainties of less than approximately 10% (7).

CASE STUDY: COMPARISON OF THE $^3H-^3HE$ METHOD TO THE CFC AND ^{85}KR DATING METHODS

Ekwurzel et al. (11) compared calculated groundwater ages for three different dating methods, including the $^3H-^3He$ method, in a study on the Delmarva Peninsula (approximately 15,690 km^2) located on the east coast of the United States. The Delmarva Peninsula is characterized by low topographical relief and little urban development. The hydrogeology of the area is relatively simple, consisting of low hydraulic gradients, highly permeable surficial materials, and shallow water tables. Discrete groundwater samples were collected from approximately 30 wells throughout the Delmarva Peninsula (most wells were screened across less than 1 m of the aquifer). Additionally, age dating methods were compared for a single flow system on the peninsula at Locust Grove.

Most age values calculated from the three methods agreed within approximately 2 years, indicating conservative behavior of the various tracers and reliability of the dating methods. A few wells that produced large apparent age discrepancies were located in areas of significant mixing. Dissolved N_2 concentrations from several wells with large tracer age variations (>10 years) indicated that gas stripping may have occurred in these samples. A one-dimensional advection–dispersion model, applied to estimate the effect of dispersion on the age correlation, determined that hydrodynamic dispersion effects were negligible. As stated earlier, fairly close agreement was observed among the different dating methods for this study; however, the locations that produced significantly different tracer ages demonstrate that use of multiple methods for groundwater age dating is important because tracers are affected differently by diffusion, dispersion, sorption, degradation, gas entrapment, excess air, recharge temperature, and sampling errors. For additional case studies on the $^3H-^3He$ groundwater dating method, see Schlosser (9), Solomon et al. (6), Cook et al. (12), Cook and Solomon (7), and Vengosh et al. (13).

BIBLIOGRAPHY

1. Craig, H. and Lal, D. (1961). The production rate of natural tritium. *Tellus* **13**: 85–105.
2. Solomon, D.K., Schiff, S.L., Poreda, R.J., and Clarke, W.B. (1993). A validation of the $^3H/^3He$ method for determining groundwater recharge. *Water Resour. Res.* **29**: 2951–2962.
3. Clarke, I.D. and Fritz, P. (1997). *Environmental Isotopes in Hydrogeology*. CRC Press, New York, pp. 20, 174–188.
4. Lucas, L.L. and Unterwerger, M.P. (2000). Comprehensive review and critical evaluation of the half-life of tritium. *J. Res. Natl. Inst. Stand. Technol.* **105**: 514–549.
5. International Atomic Energy Agency. (2001). Isotope hydrology information system. The ISOHIS Database. Available: http://isohis.iaea.org.
6. Solomon, D.K, Poreda, R.J, Cook, P.G., and Hunt, A. (1995). Site characterization using $^3H/^3He$ ground-water ages, Cape Cod, MA. *Ground Water* **33**: 988–996.
7. Cook, P.G. and Solomon, D.K. (1997). Recent advances in dating young groundwater: Chlorofluorocarbons, $^3H/^3He$ and ^{85}Kr. *J. Hydrol.* **191**: 245–265.
8. Schlosser, P., Stute, M., Sonntag, C., and Munnich, K.O. (1989). Tritiogenic ^3He in shallow groundwater. *Earth Planetary Sci. Lett.* **94**: 245–256.
9. Schlosser, P. (1992). Tritium/^3He dating of waters in natural systems. In: *Isotopes of Noble Gases as Tracers in Environmental Studies*. IAEA Proceedings, Vienna, Austria, pp. 123–145.
10. Yaws, C.L. (2001). *Matheson Gas Data Book*, 7th Edn. McGraw-Hill, New York.
11. Ekwurzel, B. et al. (1994). Dating of shallow groundwater: Comparison of the transient tracers $^3H/^3He$, chlorofluorocarbons, and ^{85}Kr. *Water Resour. Res.* **30**: 1693–1707.
12. Cook, P.G. et al. (1996). Inferring shallow groundwater flow in saprolite and fractured rock using environmental tracers. *Water Resour. Res.* **32**: 1501–1509.
13. Vengosh, A., Gill, J., Davidsson, M.L., and Hudson, B. (2002). A multi-isotope (B, Sr, O, H, and C) and age dating (^3H-^3He and ^{14}C) study of groundwater from Salinas Valley, California: Hydrochemistry, dynamics, and contamination processes. *Water Resour. Res.* **38**: 1–17.

DATING GROUNDWATERS WITH TRITIUM

STEPHEN M. TESTA
Mokelumne Hill, California

INTRODUCTION

Tritium (^3H) is one of the most common radioisotopes used to identify the presence of modern groundwater recharge and estimate apparent groundwater ages (1). Tritium, a short-lived radioactive isotope of the element hydrogen, has one proton and two neutrons. Tritium emits only a very weak beta particle and has a half-life of 12.43 years. Although tritium can occur as a gas, its most common form is in water because, like nonradioactive hydrogen, radioactive tritium readily reacts with oxygen to form water, which allows groundwater to be dated. As tritium replaces one of the stable hydrogens in the water molecule, (H_2), it is referred to as the molecule HTO, or tritiated water (T_20). Similarly to water, tritium is colorless and odorless.

Tritium is produced naturally in both the atmosphere and hydrosphere of the earth when cosmic rays strike air molecules, notably by the interaction of ^{14}N with cosmic ray neutrons in accordance with the following reactions:

$$^{14}_{7}N + ^{1}_{0}n \rightarrow ^{3}_{1}H + ^{12}_{6}C$$

and

$$^2H + ^2H \rightarrow ^3H + ^1H$$

Cosmic rays interact with nitrogen (^{14}N) or with deuterium (^2H) and form tritium and carbon (^{12}C). The primary interactions occur within the upper atmosphere, and tritium falls to earth via meteoric precipitation. Atmospheric tritium is thus rapidly incorporated into water molecules allowing it to be removed from the atmosphere as rain.

The concentration of tritium in meteoric water is measured in "tritium units" (TU), where 1 TU is equal to an abundance of one atom of ^3H per 10^{18} atoms of hydrogen. This is equivalent to 7.1 disintegrations of ^3H per minute per liter of water.

ANTHROPOGENIC TRITIUM

The occurrence of tritium in the environment has been significantly changed over the years by anthropogenic influences. Tritium has several important uses; the most significant is its use as a component in the triggering mechanism in thermonuclear (fusion) weapons; very large quantities are required to maintain nuclear weapons capabilities. On May 9, 1951, the world's first thermonuclear flame reaction was ignited in a laboratory by a ^2H-^3H-^{235}U device code-named George. Following the birth of this concept, atmospheric detonation of these devices resulted in anthropogenic production of ^3H which raised concentrations in the stratosphere by several orders of magnitude. The proliferation of atmosphere testing continued throughout the mid-1950s, early 1960s, and into the 1970s. Tritium is also produced in accelerator cooling water systems or reactors, which are subjected to large proton or neutron fluxes via spallation reactions with oxygen, nitrogen, and carbon nuclei in the water systems and to a much smaller degree by radiative capture of neutrons by the deuterium nuclei in water. The tritium atom then combines with a hydrogen and oxygen atom to form tritiated water. Anthropogenic tritium has thus been formed by nuclear weapon explosions in the

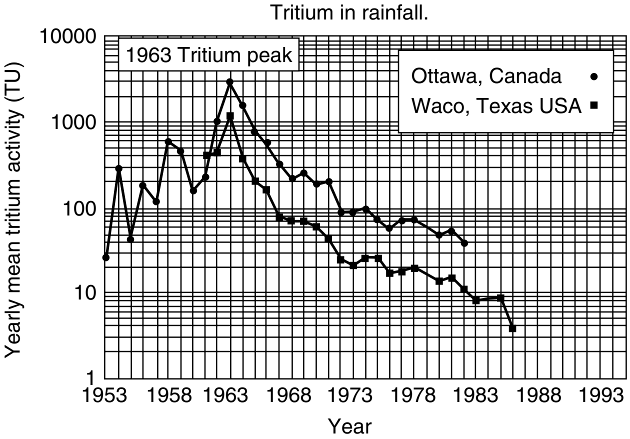

Figure 1. Tritium chart for Ottawa, Canada, and Waco, Texas, (2).

atmosphere and the operation of nuclear reactors and particle acceleration.

The longest record of atmospheric tritium measurements was initiated for precipitation in Ottawa, Canada, beginning in 1953. Atmospheric testing of nuclear devices commenced in 1952. After 1953–1954, the amount of tritium increased and reached a maximum in 1963–1964. Concentrations of ^{3}H in precipitation have decreased since the mid-1960 bomb peak, except for some small secondary peaks by China and France in 1973 and 1975. Tritium input to groundwater has subsequently occurred in a series of spikes following these periods of atmospheric testing (Fig. 1). The majority of tritium produced between the early 1950s through mid-1970s in the atmosphere has since been removed.

Other uses include various self-luminescent devices such as aircraft dials, gun sights, luminous paints, exit signs, and watches. Tritium is also used in life science research and in studies investigating the safety of new drugs.

DATING GROUNDWATER USING TRITIUM

The presence of tritium in the atmosphere from weapons testing has proved useful from several perspectives. The residual presence of nuclear-produced tritium has served as a tracer for atmospheric studies and the study of hydrologic and oceanographic processes. Tritium is considered a dating element for groundwater. By knowing the tritium concentration in precipitation and its distribution in groundwater, an apparent date for the age of the groundwater can be determined, which also allows evaluating aquifer water average transit time (T). This parameter is important because it can be used to address issues about evaluating, exploiting, and protecting water resources.

Dating groundwater can be used in ecologically sensitive coastal settings to understand the hydrogeologic setting and determine whether water is derived from surficial and/or subsurface groundwater. Oceanographers used the sharp spike in surface tritium levels over the years to measure the rate at which lower and upper ocean levels mixed. Tritium is a good marker for tracking ocean water.

Several factors affect tritium concentration in meteoric water. In the lower stratosphere, the residence time for tritium in water is between 1 and 10 years. When tritium water reaches the lower troposphere, residence time is 5 to 20 days (3). It is known that tritium concentration as meteoric precipitation in the Northern Hemisphere has fluctuated widely from less than 25 TU, prior to 1953, to more than 2200 TU in 1964 following extensive testing in the atmosphere (3). In the Southern Hemisphere, the level of bomb-produced tritium is much lower. Other factors to consider include

1. the natural production rate in the stratosphere which is estimated at 0.5 ± 0.3 atoms of ^{3}H/cm^2/s (4);
2. the decay of tritium by beta emission to stable $^{3/2}$He with a half-life of 12.26 y;
3. the seasonal injection of tritium from the stratosphere into the troposphere;
4. the presence of bomb-produced tritium; and
5. the presence of tritium produced locally by nuclear reactors and particle accelerators.

METHODOLOGY

Groundwater is dated by retrieving water samples from surface water bodies, ice or precipitation, or wells in the case of groundwater. By observing the spike in tritium levels, likely the result of atmospheric bomb testing, an age can be estimated. For example, if the spike in groundwater is determined to be 20 miles from its recharge or source area, then it could be deduced that the spike has moved 20 miles during the past 40 years, or an average rate of 0.5 miles per year.

Apparent groundwater age can be estimated directly from tritium precipitation records corrected for first-order decay. Groundwater is age dated by determining the early and late ratio of tritium in accordance with the following equation:

$$^{3}\mathrm{H}_{\text{early sample}} \times e^{\lambda\Delta t}/^{3}\mathrm{H}_{\text{later sample}}$$

where $\lambda = 0.0565\ \mathrm{yr}^{-t}$ (tritium decay constant)
$\Delta t =$ the time between samples in years

A ratio greater than 1 indicates that the tritium peak has passed, or in other words, the water is older 1963. Conversely, a ratio less than 1 indicates that the peak has not yet arrived, or the water is younger than 1963.

Some amount of hydrodynamic dispersion, however, is likely to occur, which explains why natural tritium levels in precipitation are very low and reflect a secular equilibrium between natural production and the combination of decay in the atmosphere in addition to losses in the hydrosphere and ocean. Advective dispersion and diffusion are typically addressed using a mixing model that more accurately estimates groundwater ages by accounting for attenuation of tritium levels due to these processes (5). Mixing models approximate dispersion and diffusion processes by assuming a Gaussian distribution of tritium activities.

Three kinds of models are commonly used:

1. the well-mixed or exponential model which assumes that water introduced into a system is completely and uniformly mixed with aquifer water. Thus, the water output from the system is actually representative of the aquifer waters.
2. the piston-flow model which assumes that the incoming water traverses the aquifer at constant velocity.
3. the dispersion model which assumes a dispersion phenomenon reflecting the heterogeneity of the aquifer material.

Different age waters can be distinguished. For example, tritium contents less than 2 TU typically characterize old groundwater, 2 to 10 TU may characterize mixing between old and recent water, 10 to 40 TU may characterize recent groundwater, and greater than 40 TU may characterize groundwater in close proximity to a nuclear station.

The main application of tritium is to differentiate pre-1952 water from younger water. Knowing the concentration of tritium in precipitation (the source) and its distribution in groundwater, an apparent date can be estimated.

The use of tritium to study the movement of water in the subsurface, evaluating source areas, and dating the last time subsurface water was last exposed to the atmosphere is hindered by several factors. The most important limitations are the continuing decay of nuclear-produced tritium, its unequal global distribution, and the local injection of anthropogenic tritium. For example, a tritium concentration of 30 TU may reflect late 1950s water that has decayed through three half-lives, or a 1970s water originally with 75 TU that has decayed through one half-life (6).

In many scenarios, locating the position of the mid-1960 peak is difficult due to the relatively high density of vertical sampling required, which can be expensive and impractical. Furthermore, although tritium can be used to locate the depth of the mid-1960s peak, many samples are required due to radioactive decay to locate its position. Furthermore, in waters younger than the mid-1960s, the bomb peak is not present.

DATING YOUNG GROUNDWATER USING TRITIUM/HELIUM RATIOS

Tritium is not useful for the dating relatively younger water. A way around this is the use of ^{3}H/^{3}He ratios. Radioactive decay of ^{3}H produces the noble gas helium-3 (^{3}He). First introduced by Tolstikhin and Kamenskiy (7), measurements of ^{3}H and tritiogenic ^{3}He define a quasi-stable tracer of initial ^{3}H input to groundwater and may be used to determine the position of the mid-1960s bomb peak in recharge areas and recharge rates (8–11).

Sources of helium are primarily atmospheric. Other sources include terrigenic He (R_{terr}), water in rocks that are enriched in certain elements such as thorium (Th) or Uranium (U), typically of crustal and mantle material, and samples representative of the mixing of relatively young water with old water containing R_{terr}.

Using tritium and helium concentrations, the tritium/helium age can be calculated as follows (12):

$$T_{\text{groundwater}} = t^{1/2}/\text{In } 2 \times \text{In } [1 + (^3\text{He}/^3\text{H})]$$

where $T = 12.3$ years (the half-life of tritium)
$^3\text{He}/^3\text{H} =$ helium/tritium concentration ratio (TU)

To convert the concentration units required to evaluate this equation:

$$1 \text{ cm}^3 \text{ STP g}^{-1} = 4.019 \times 10^{14} \text{ tritium units (for freshwater)}$$

This conversion can be used in conjunction with a correlation to the tritium peak method for verification.

Conditions for using this technique include the following:

1. Detectable tritium in the water sample needs to be greater than approximately 0.5 TU.
2. If R_{terr} terrigenic is present, then Ne data are also needed to define $^3\text{He}_{\text{tri}}$.
3. The ^{3}He/^{4}He ratio of R_{terr} must be known
4. If the amount of R_{terr} is small (i.e., <5% of the dissolved ^{4}He), the ^{3}H/^{3}He age may be insensitive to large uncertainties in R_{terr}.
5. R_{terr} must be known for samples with a large fraction of R_{terr} (i.e., within approximately 1% or better).
6. If R_{terr} cannot be defined with sufficient precision for age determination, a range in R_{terr} can be evaluated.

^{3}He/^{3}H allows dating groundwater of very young ages (approximately 1 year with an uncertainty in age of approximately 0.5 years due to analytic constraints) from a single water sample. This makes ^{3}He/^{3}H very attractive and complements other techniques for dating groundwater. ^{3}He/^{3}H can be applied to dating waters that have been recharged in the past approximately 30 years, so this technique can be used for interpreting ages of water from existing domestic, industrial, and municipal production supply wells and recharge operations. Water source(s) identification, flow paths, and flow rates for a variety of purposes and objectives can also be evaluated.

From an environmental perspective, it can also be used to monitor the use of certain chemicals released into the environment such as chlorofluorocarbons (CFCs) and sulfur hexafluoride. For example, for CFCs, the ^{3}H/^{3}He age is defined as the time elapsed since the water was isolated from the atmosphere following recharge.

Groundwater in aquifers is generally older than expected on the basis of flow velocity. Recent studies suggest that the rate of mass exchange between aquifers and aquitards or confining layers is assumed to be small (13). The effect on age, however, does not depend on mixing rates but rather on the ratio of fluid volume in aquitards to aquifers. This is important in understanding the relationship between groundwater flow and radiometric age distribution. It means that flow rates along an aquifer are higher than predicted by age measurements. Aquitard water is generally very old; thus, at high mixing rates, less-old water is supplied to the aquifer, and less-old water is being supplied to the aquifer because younger water is moving into the aquitard. Thus, mixing increases the age in an aquifer and also has a counterbalancing effect of decreasing the age in aquitards.

BIBLIOGRAPHY

1. Clark, I. and Fritz, P. (1997). *Environmental Isotopes in Hydrogeology*. Lewis, Boca Raton, FL.
2. Morrison, R.D. (2000). *Environmental Forensics, Principles & Applications*. CRC Press, New York.
3. Gat, J.R. (1980). The isotopes of hydrogen and oxygen in precipitation. In: *Handbook of Environmental Isotope Geochemistry*. P. Fritz and J.Ch. Fontes (Eds.). Elsevier, Amsterdam, Vol. 1A, pp. 21–47.
4. Graig, H. and Lal, D. (1961). The production rate of natural tritium. *Tellus* **13**: 85–105.
5. Wade, S., Leap, D.I., and Fritz, S.J. (2001). Estimating Indiana groundwater age from tritium activities. *Geol. Soc. Am. Annu. Meet.* Abstract, November 5–8, 2001.
6. Dominico, P.A. and Schwartz, F.W. (1990). *Physical and Chemical Hydrogeology*. John Wiley & Sons, New York, pp. 526–527.
7. Tolstikhin, I.N. and Kamenskiy, I.L. (1969). Determination of groundwater ages by the T-He method. *Geochem. Int.* **6**: 810–811.
8. Schlosser, P. et al. (1988). Tritium/^{3}He dating of shallow groundwater. *Earth Planetary Sci. Lett.* **89**: 353–362.
9. Schlosser, P. et al. (1989). Tritiogenic ^{3}He in shallow groundwater. *Earth Planetary Sci. Lett.* **94**: 245–256.
10. Solomon, D.K. and Sudicky, E.A. (1991). Tritium and helium-3 isotope ratios for direct estimation of spatial variations in groundwater recharge. *Water Resour. Res.* **27**(9): 2309–2319.
11. Solomon, D.K. et al. (1993). A validation of the ^{3}H/^{3}He method for determining groundwater recharge. *Water Resour. Res.* **29**(9): 2951–2962.
12. Aeschbach-Herrig, W. et al. (1998). A 3H/3He study of ground water flow in a fractured bedrock aquifer. *Ground Water* **36**(4): 661–670.
13. Bethke, C.M. and Johnson, T.M. (2002). Paradox of groundwater age. *Geology* **30**(2): 107–110.

RECHARGE IN DESERT REGIONS AROUND THE WORLD

Kristine Uhlman
University of Arizona
Tucson, Arizona

INTRODUCTION

Arid regions are characterized by dryness, variously defined as rainfall insufficient for crops without irrigation, evaporation/evapotranspiration rates greater than precipitation, or less than 25 cm of annual rainfall. Most of the true deserts are found between 15 and 50° latitude north or south of the equator. Dry regions in the Arctic and Antarctic are not normally classified as deserts, although parts of northern Greenland and the McMurdo Sound area of Antarctica have closed basins and saline lakes and, except for temperature, would be hard to distinguish from deserts at lower latitudes (1). Middle latitude deserts in the basin interiors of the continents are characterized by scant rainfall and high summer temperatures. The trade-wind deserts, notably the Sahara of North Africa, exhibit negligible precipitation and large daily temperature ranges. Recharge rates in arid regions typically range between 3 to 4% of precipitation (2); however, a rate as low as 0.066 mm/yr or 0.033% has been measured in the Rio Grande Basin, New Mexico (3).

General

Approximately one-tenth of the land surface of the planet is arid, and an equal or slightly greater area is classified as semiarid. Arid climatic conditions profoundly affect landscape topography and depositional setting, which controls aquifer type in nonindurated deposits. Erosion and transport by wind and water result in fluvial and aeolian deposits ranging from sand dunes to valley fill deposits of stratified sand, gravel, silts, and clays. Some desert basins lack thick alluvial deposits, and groundwater is found in fractured consolidated bedrock. Carbonate aquifers exhibiting karstic features are found in some deserts and are likely vestiges of tropical climates predating current arid conditions. Regardless of the fabric of the aquifer material, it is generally true that average annual rainfall less than 12 to 25 cm may be insufficient to allow recharge to an arid region aquifer. For this reason, the sustainability of groundwater resources in arid regions is of serious concern and the focus of numerous investigations.

Historical Background

Several recharge investigations have been conducted in the southwestern United States related to water resource management and artificial recharge of arid region aquifers using waters imported from distant sources. With the construction of the California Aqueduct, the Columbia River Project (Eastern Washington State), and the Central Arizona Project, large tracts of land in the arid American west have been studied to design and engineer enhanced recharge and aquifer storage and recovery (water banking) projects (4–9).

Following environmental regulatory compliance actions in the mid 1980s, aquifer vulnerability to contamination due to agricultural and industrial practices has been the focus of several studies (3,10–12). The design of the low-level radioactive waste disposal facility in Nevada has been supported by the Amargosa Desert Research Site under the United States Geological Survey (USGS) Low-Level Radioactive Waste Program arid recharge investigations (10,13,14).

In response to recognized water resource issues related to drought, the National Science Foundation Sustainability of Semi-Arid Hydrology and Riparian Areas (SAHRA) program was initiated in 2000 to facilitate the transfer of knowledge to support the development of water resources policy. This multi-institutional/multidisciplinary program has conducted several studies on the Rio Grande and San Pedro/Gila River Basins (Arizona), including the Semi-Arid Land Surface Atmosphere (SALSA) project to investigate the watershed-basin-scale hydrologic water balance using remote sensing tools (15).

FUNDAMENTALS OF SUBSURFACE RECHARGE

Following Potter's (10) classification, arid and semiarid regions are those whose precipitation to evaporation (P/E) ratios are smaller than 0.5 and between 0.5 and 1.0, respectively. In addition to a lower net precipitation rate, rainfalls are typically infrequent and of short duration; high-intensity seasonal storms contribute a major portion of the annual rainfall over short periods of time. Under these conditions, the shallow soil surface layer infiltration capacity may be exceeded and excess overland flow propagates rapidly to flash flooding. Even low-intensity rainfalls, can lead to significant surface runoff when desert soil crusts have formed (16).

Water focused and concentrated by topography into ephemeral washes, wadis, sabkhas, and playas provides locations of enhanced recharge in arid areas. A significant component of recharge to arid basin aquifers occurs along mountain fronts. Important aspects along the mountain front include the partitioning of rainfall and snowmelt into surface runoff, deep infiltration along fractures and faults, and vegetation-controlled evapotranspiration (17). Focused flow along mountain stream channels into the major washes in the Tucson Basin (Arizona) visibly infiltrates flood-plain sediment, and surface water commonly does not leave the watershed (18). Half the annual recharge of 9 to 10 mm/year to the Ogallala Aquifer in the southern High Plains (Oklahoma and Texas) occurs through playa floors that cover only 6% of the aquifer land surface area (19).

Arid recharge requires the downward vertical flow of water in the unsaturated zone and an excess availability of water above the evapotranspiration demand. Once infiltrated, loss mechanisms in the unsaturated zone (uptake by plant roots, evaporation, and vertical vapor flow) increase the possibility that water never reaches an underlying deep aquifer (20). The measured ratio of groundwater recharge to precipitation ranges between 2 to 3% in most arid regions. An exception has been found, however, in the coastal sabkhas of Abu Dhabi, where measured recharge was approximately 90% of the mean annual precipitation of 70 mm. This high ratio is believed to be the result of several factors: the most important is the high water table, which allows infiltration to reach the saturated zone before evaporation within the vadose zone can occur (21). Fluid pathways in the unsaturated zone are inherently complex; transport distance and time of travel are significant factors that control arid region recharge rates. Because of this, unsaturated zone techniques are generally required to estimate recharge in arid regions (11).

Vadose Zone Capillary Flow

Capillary forces, together with gravity, determine fluid flow processes in unsaturated alluvium; finer grained materials respond like sponges so that there is no free drainage, even at high water content. Capillary rise causes water to move upward against the pull of gravity and takes place in small interstitial pores under negative soil water pressures that prevent free drainage. For example, it takes 1000 days for a water particle to travel though a 10 m deep vadose zone of fine sand whose volumetric soil water content is 0.15 and whose unsaturated hydraulic conductivity is 0.15 cm/day (the hydraulic gradient below the root zone equals unity, so water flux becomes equal to the hydraulic conductivity) (19). Capillary recharge through unsaturated fine-grained alluvium is unlikely in an arid setting.

In a study of recharge in the intermittent washes in the western Mojave Desert near Victorville, California, it was found that infiltration did not occur at control sites beyond the wash where water did not accumulate. Water did not pass the root zone, and the δ^{18}O and δ deuterium isotopic composition of water retained in the shallow coarse-grained soils were consistent with removal of water by vapor transport (9). Results from Southern Nevada research indicate that soil water may be as old as 120,000 years near the bottom of the deep (240–400 m) vadose zones, which shows there has been little or no infiltration of recent precipitation (14).

Vadose Zone Macropore Flow

"Macropore" is a common name for a wide range of large pores whose diameters are generally greater than 3 mm, such as cracks in clay soils, rock fractures, fissures in sediments, polygonal pressure cracks, pipes in the indurated calcic horizon, worm holes, and old root channels. Following precipitation that produces runoff and ponding, macropores that start at the soil surface quickly become conduits for recharge. The velocity of a water particle in a macropore in the soils from the previous example is approximately 30 cm/s, and it takes only 33 s for a water particle to travel 10 m from a ponded soil surface to the water table. Macropore transport reduces evaporation losses and results in a much larger recharge rate (19).

Techniques for quantifying macropore recharge were developed in the American southern high plains semiarid alluvial fan and dune aeolian deposits, where calculated recharge flux through macropores was between 60 and 80% of the total volume of recharge that reached the water table. Using a storm model to estimate rainfall-runoff with isotope geochemistry supporting the model simulations, 95% of the total recharge was found focused into 5% of the land area where water was focused into depressions and valleys (22).

In the southern high plains of Texas and New Mexico, recharge flux averages 9 mm/year regionally or approximately 2% of precipitation. Use of stable isotopes and tritium data indicate that approximately half of the recharge occurred in the floor of playa lakes that occupy 6% of the land area; much of the remaining recharge is believed to be from macropore flow (23).

MEASUREMENT TECHNIQUES

Approaches to quantify recharge vary depending on the climate, topography, and geology of the aquifer;

processes that control the spatial distribution and volume of infiltration that recharges arid aquifers are poorly understood. For these reasons, numerous techniques are typically used, and results are compared to calibrate the analyses.

Water Balance/Modeling

Rainfall/runoff modeling is widely used to estimate recharge where recharge is calculated as a residual of the water budget. This water budget approach assumes that deep drainage below the root zone will eventually reach the water table and become recharge, sometimes neglecting the potential for deep vapor fluxes to remove water from the system. When estimated as a residual in water balance models, recharge may be in error by an order of magnitude. In a study by Gee and Hillel (24), it was recommended that quantification of drainage by lysimetry and tracer tests should be used to calibrate qualitative model estimates. Modeling efforts that ignore seasonality or transient conditions cannot simulate recharge when average annual evapotranspiration exceeds precipitation (25,26). Arid site recharge studies found significant variability under similar climatic and soil conditions, but using different plant cover and topography, estimates of recharge based on fixed fractions of annual precipitation do not reflect the seasonal plant and soil factors that control recharge.

Chloride mass-balance methods that integrate temporal and aerial distribution of groundwater recharge have been found well suited to areas of large temporal and spatial variation in recharge. For small alluvial aquifers in the wadi systems of the A'Asir and A'Hijaz mountains of western Saudi Arabia, Bazuhair (2) found that recharge varied between 3% and 4% of precipitation using this technique.

Physical Measurement

Stream gaging, seepage meters, and lysimeters provide point estimates of recharge and discharge but are not representative of the entire system (24). Streamflow and groundwater data collected in the Rillito River in the Tucson Basin was compared to microgravity measurements that showed an increase in groundwater storage within the area of the floodplain. Coupled with neutron moisture readings and environmental tracer measurements, it was found that infiltration is controlled by geology and structure, where initial depths to groundwater were greatest (27–29).

Environmental Tracers—Recharge Sources

Many of the alluvial aquifers in the Basin and Range region of the American southwest were hydrologically closed during development, resulting in readily soluble evaporate material, such as gypsum and halite, interbedded with alluvial fill. The basin fill materials are typically of late Cenozoic age and consist of thousands of feet of coarse gravel to silt and clay derived primarily from igneous and metamorphic rocks eroded from the surrounding mountains. During precipitation that is intense enough to produce runoff in ephemeral streams, recharge to the aquifer occurs near the points where stream channels leave the mountains and flow across the alluvium. Recharge from ephemeral streams generally has low dissolved solids concentrations, but as water moves downward toward the zone of saturation and down-gradient, there is extensive opportunity for interaction between water and rock. This results in a regional geochemical pattern of groundwater chemistry where more dilute water occurs near recharge areas (4). High total dissolved solids (TDS), chloride, or boron concentrations indicate low water flux and concentration by evaporation within the vadose zone or long-term residence within the saturated zone (30). Mapping of geochemical tracers within the Middle Rio Grande Basin identified spatial patterns of recharge, and focused recharge was found beneath arroyos (13).

Other tracer methods include age-dating of groundwater with historical tracers such as bomb tritium, ^{36}Cl, carbon-14, and oxygen isotope/deuterium ratios (20). On the basis of carbon-14 data, most groundwater in the regional aquifer beneath Oro Grande Wash in the Mojave Desert was recharged between 5000 and 20,000 years ago (9). In the intensely arid region of the Western Desert of Egypt, deuterium data show that modern water cannot be the source of recharge for groundwater, and it is speculated that the waters derived either from distant recharge in Chad or the Sudan or are very old Pleistocene waters from a cooler glacial time (31). Oxygen isotope ratios can determine temperature and coincidental elevation of recharge waters, allowing for differentiation between winter storm/summer storm precipitation, mountain-front recharge, and snowmelt waters (18).

In the Trans-Pecos Texas, Chihuahuan Desert, ^{87}Sr/^{86}Sr ratios (in combination with other chemical and isotopic data) were used to infer flow paths and residence times of recharged water. Concentrated recharge was found within regional fracture trends in carbonate rocks in the highlands, losing streams on proximal portions of alluvial fans, and irrigation return flow along well casings (32).

Environmental Tracers—Recharge Rates

Mapping of environmental tracers allows the investigator to find recharge sources, but it is often more difficult to quantify recharge rates. Artificial tracer tests of soil moisture through unsaturated, poorly sorted, horizontally stratified alluvial deposits found that transport takes place along preferential flow paths. Bromide was used as a tracer to measure the velocity of soil moisture, which ranged from approximately 55 to 275 cm per day on a test plot where water was artificially applied to natural soils (6).

Because recharged water from desert washes also spreads laterally away from the wash as depth increases, the flux of water decreases with increasing depth. In addition, arid soils are less permeable beyond the channel due to deposition of soluble salts (9); therefore, quantifying recharge rates in these settings is difficult.

Studies from the Amargosa Desert in southern Nevada, the Middle Rio Grande Basin in north-central New Mexico, and the Palouse Catena of the Columbia River regional aquifer in southeastern Washington were used to compare tracer-based recharge estimates with water-balance models. Mean annual precipitation at the three study areas was approximately 0.1, 0.25, and 0.5 m per year, respectively. Estimates of groundwater recharge were derived from vadose zone profiling of chloride, stable isotopes, and tritium and compared. The geochemical chloride profiles suggested that recharge rates are in approximate equilibrium with current climatic conditions. This was corroborated by isotopic and hydraulic data; however, chloride profiles at the driest sites suggested features that could be interpreted as relicts of past wetter conditions (13,25,26,33).

Mapped geochemical and isotopic data reflect recharge from basin margins and from the Rio Grande River in the Middle Rio Grande Basin, New Mexico. Isotopically depleted waters were some of the oldest waters in the basin and represent water recharged in the area of the Jemez Mountains during the last glacial period about 20,000 years ago. Younger waters, about 7,000 years old, were also mapped along the basin margins. Radiocarbon ages were used to estimate modern and paleorecharge rates, and recharge rates during the last glacial period were estimated to be at least six times greater than the modern radiocarbon-based recharge (34).

INCIDENTAL RECHARGE

Arid land recharge may include incidental recharge due to environmental pollutant releases, irrigation return flow, canal seepage, sewage effluent seepage in dry washes, and seepage from surface impoundments such as stormwater retention ponds. The evidence for irrigation and industrial-related recharge includes the presence of pesticides, caffeine, pharmaceuticals, and other chemical pollutants in groundwater (20,35), although it should be recognized that some contaminant chemicals may be transported through the subsurface by nonaqueous means. Treated municipal wastewater released into ephemeral, effluent-dependent stream channels is an important source of recharge in some regions (36). The Santa Cruz River in south-central Arizona contributes 6.0 acre-feet/day/mile to the subsurface, or approximately 80% of the flow volume results in recharge (37).

SUMMARY

Hydrologic processes in arid or semiarid lands are distinctly different from those in humid climates and exhibit wide variation in magnitudes and type of recharge. Precipitation that results in focused flow and ponding allows macropore transport of recharge, whereas loss mechanisms in the unsaturated zone allow unsaturated water or vapor flux to occur in either vertical direction, depending on surface conditions. Recharge studies use a combination of approaches, such as thermal and geophysical techniques, geochemical tracers, radioisotopes, and unsaturated zone modeling calibrated against streamflow and groundwater level information to characterize this complex hydrologic process (28). Extremes in temperature and precipitation in arid regions necessitate focused consideration of the amount and distribution of recharge to the aquifers so as to manage limited water resources adequately.

BIBLIOGRAPHY

1. Davis, S. and DeWiest, R.J.M. (1966). *Hydrogeology*. John Wiley & Sons, New York.
2. Bazuhair, A.S. and Wood, W.W. (1996). Chloride mass-balance method for estimating ground water recharge in arid areas; examples from western Saudi Arabia. *J. Hydrol.* **186**(1–4): 153–159.
3. Daniel B. Stephens & Associates, Inc. (1996). *Estimation of Infiltration and Recharge for Environmental Site Assessment*. American Petroleum Institute, Albuquerque, NM, Health and Environmental Sciences Department, Publication Number 4643.
4. Hem, J.D. and Robertson, F.N. (1987). Hydrogeochemistry of ground water recharge in alluvial aquifers, Southern Arizona. *Symp. Proc., 3rd Symp. Artificial Recharge Groundwater in Arizona*. May 20–21, 1987. Salt River Project, Phoenix, AZ, pp. 30–51.
5. Drost, B.W., Whitman, K.J., and Gonthier, J.B. (1990). *The Geologic Framework of the Columbia Plateau Regional Aquifer System, Washington, Oregon, and Idaho*. U.S. Geological Survey, Water-Resources Investigation Report 87–4238.
6. Graham, D.D. (1991). Results and significance of an unsaturated-zone tracer test at an artificial-recharge basin, Tucson, Arizona. *Symp. Proc., 5th Bienn. Symp. Artificial Recharge Groundwater*, May 29–31, 1991. Salt River Project, Phoenix, AZ, pp. 30–51.
7. Rockhold, M.L., Fayer, M.J., Kincaid, C.T., and Gee, G.W. (1995). *Estimation of Ground Water Recharge for the Performance Assessment of a Low-Level Waste Disposal Facility at the Hanford Site*. Battelle Pacific Northwest Laboratory, Richland, WA.
8. Watson, M. (1999). *Final Technical Memorandum, Identification of Candidate Recharge Areas and Pilot Test Sites, Farmington Groundwater Recharge/Wetlands Feasibility Study*. U.S. Army Corps of Engineers.
9. Izbicki, J.A., Michel, R.L., and Martin, P. (2000). Chloride and tritium concentrations in a thick unsaturated zone underlying an intermittent stream in the Mojave Desert, Southern California. *Conf. Proc. Nat. Recharge Groundwater Symp.*, June 2, 2000, Tempe, AZ, pp. 1–8.
10. Potter, L.D. (1992). Desert characteristics as related to waste disposal. In: *Deserts as Dumps? The Disposal of Hazardous Materials in Arid Ecosystems*. C.C. Rwith and B.M. Thomson (Eds.). Univ. of New Mexico Press, Albuquerque, NM, pp. 21–56.
11. Scanlan, B.R., Tyler, S.W., and Wierenga, P.J. (1997). Hydrologic issues in arid, unsaturated systems and implications for contaminant transport. *Rev. Geophys.* **35**(4): 461–490.
12. Desconnets, J.C., Taupin, J.D., Lebel, T., and Leduc, C. (1997). Hydrology of the HAPEX-Sahel Central Super-Site: Surface water drainage and aquifer recharge through the pool system. *J. Hydrol.* **188–189**(1–4): 155–178.
13. Stonestrom, D.A. (1999). What do environmental tracers say about patterns of ground water recharge, past and present? Examples from the Amargosa Desert, Rio Grande Basin, and Palouse Catena. *Geological Society of America, 1999 Annu. Meet., Abstracts with Programs—Geological Society of America*, Oct. 25–28, 1999, p. 87.
14. Tyler, S.W. (1999). Ground water recharge in arid regions: Questions about today and the past? (Abstract) Ground water supply issues in the next century. *NGWA National Convention and Exposition*, Dec. 3–6, 1999, p. 7.

15. Goodrich, D.C., Chehbouni, A., Goff, B., MacNish, B., Maddock, T., Moran, S., Williams, D.G., and Watts, C. (1998). *An Overview of the 1997 Activities of the Semi-Arid Land-Atmosphere (SALSA) Program*. American Meteorological Society. Special Symposium on Hydrology. Phoenix, AZ.
16. Abu-Awwad, A.M. and Sharanawi, M.R. (1997). Water harvesting and infiltration in arid areas affected by surface crust: Examples from Jordan. *J. Arid Environ.* **37**: 443–452.
17. Wilson, J.L. and Guan, H. (2004). *Mountain-Block Hydrology and Mountain-Front Recharge*. F.M. Phillips, J.F. Hogan, and B. Scanlon (Eds.). Water Resources Monograph Series. American Geophysical Union, Washington, DC.
18. Eastoe, C.J. (2003). Locating recharge zones with isotopes. *Southwest Hydrol.* **January/February**: 22–23.
19. Hendrickx, J.M.H. and Walker, G.R. (1997). Recharge from precipitation. In: *Recharge of Phreatic Aquifers in (Semi-) Arid Areas*. I. Simmers (Ed.). Vol. 19, International Association of Hydrogeologists, Balkema, Rotterdam, the Netherlands, pp. 19–114.
20. Scanlon, B.R. (in press). *Evaluation of Methods for Estimating Basin Scale Recharge in Semiarid and Arid Regions. Recharge and Vadose Zone Processes: Alluvial Basins in the Southwestern United States*. F.M. Phillips, J.F. Hogan, and B. Scanlon (Eds.). Water Resources Monograph Series. American Geophysical Union, Washington, DC.
21. Sanford, W.E. and Wood, W.W. (1999). A large recharge/precipitation ratio on the coastal sabkhas of Abu Dhabi. *Geological Society of America, 1999 Annu. Meet.*, *Abstracts with Programs—Geological Society of America*, Oct. 25–28, 1999, pp. 214–215.
22. Wood, W.W., Rainwater, K.A., and Thompson, D.B. (1997). Quantifying macropore recharge: Examples from a semi-arid area. *Ground Water* **35**(6): 1097–1106.
23. Wood, W.W. and Sanford, W.E. (1995). Chemical and isotopic methods for quantifying ground water recharge in a regional, semiarid environment. *Ground Water* **33**(3): 458–468.
24. Gee, G.W. and Hillel, D. (1988). Groundwater recharge in arid regions: Review and critique of estimation methods. *Hydrological Processes* **2**(3): 255–266.
25. Walvoord, M., Phillips, F.M., Tyler, S.W., and Hartsough, P.C. (2002). Deep arid system hydrodynamics, Part 2: Application to paleohydrologic reconstruction using vadose-zone profiles from the Northern Mojave Desert. *Water Resour. Res.* **38**(12): 1291. doi: 10.1029/2001 WR000825.
26. Walvoord, M. and Phillips, F.M. (in press). Deep arid system hydrodynamics, Part 3: Vegetation control on desert-floor recharge. *J. Hydrol.*
27. Pool, D.R. and Wellman, J. (1998). Streamflow infiltration and ground-water recharge along Rillito Creek in response to the El Nino-related precipitation in 1992–93 and in 1998. *11th Annu. Symp. Arizona Hydrological Society*, September 23–26, 1998, Tucson, AZ.
28. Hoffman, J.P. (1999). Recharge processes studied in Rillito Creek, Tucson, Arizona. *12th Annu. Symp.* Arizona Hydrological Society, Pinetop-Lakeside, AZ, September 9–11, 1999.
29. Bailey, M.A., Hoffmann, J.P., and Ferre, P.A. (2000). *Investigation of Subsurface Temperature Variations to Quantify Infiltration in an Ephemeral Stream, Tucson, Arizona*. Geological Society of America, *Abstracts with Programs—Geological Society of America*, Nov. 9–18, 2000, p. 185.
30. Hoffman, J.P. and Steinfampf, W.C. (1999). A study of recharge processes in an ephemeral stream in Tucson, Arizona. *Proc. 9th Bienn. Symp. Artificial Recharge of Groundwater*, Tempe, AZ, June 10–12, 1999, pp. 49–55.
31. Landis, G.P. et al. (1997). Reconnaissance hydrogen isotope geochemistry of Egyptian ground waters. *Ann. Geol. Surv. of Egypt* **20**(Part II): 655–673.
32. Sharp, J.M. and Uliana, M.M. (1998). Regional groundwater flow in a semi-arid region—controls of fracture systems in carbonate units. *Geological Society of America Symp.*, Toronto, Canada. Presentation and Abstract.
33. Walvoord, M., Plummer, M.A., Phillips, F.M., and Wolfsberg, A.V. (2002). Deep arid system hydrodynamic, Part 1: Equilibrium states and response times in thick desert vadose zones. *Water Resour. Res.* **38**(12): 1308. doi: 10.1029/2001 WR000824.
34. Plummer, L.N. (2002). Environmental tracers and how they are used to understand the aquifer. In: *Ground Water Resources of the Middle Rio Grande Basin*. J.R. Baroline and J.C. Cole (Eds.). U.S. Geological Survey Circular 1222, pp. 82–83.
35. Seiler, R.L., Zaugg, S.D., Thomas, J.M., and Howcroft, D.L. (1999). Caffeine and pharmaceuticals as indicators of waste water contamination in wells. *Ground Water* **37**(3): pp. 405–410.
36. Sinclair, K. (1999). *City of Moses Lake, Dunes Municipal Waste Water Treatment Plant, Class II Ground Water Inspection*. Washington State Department of Ecology, Environmental Assessment Program. Water Body No. WA-41-9289GW. Olympia, WA.
37. Wilson, L.G., DeCook, K.J., and Neuman, S.P. (1980). *Regional Recharge Research for Southwest Alluvial Basins*. Water Resources Research Center, University of Arizona, Tucson, AZ.

HYDROLOGIC FEASIBILITY ASSESSMENT AND DESIGN IN PHYTOREMEDIATION

J.W. Weaver
S.C. Mccutcheon
S.L. Hutchinson
(from *Phytoremediation: Transformation and Control of Contaminants*, Wiley 2003)

SUMMARY OF PRACTICAL IMPLICATIONS

Several design considerations for phytoremediation are based on the capability of trees to remove contaminated water from aquifers. These capabilities depend upon the amount of water that can be withdrawn, the vertical distribution of contaminants, and flow patterns in the aquifer. For each of these, an example field site is given where methods were applied to quantify the characteristics. These include a model of forest evaporation for estimating the balance between precipitation and evapotranspiration, a simple model of recharge-driven plume diving, and the use of groundwater flow modeling for assessing various features of the groundwater flow.

ASSESSMENT AND CHARACTERIZATION OF PHYTOCONTAINMENT

To clean up a contaminated site, characterizing the type and extent of pollutants at the site and developing a conceptual remediation plan are essential. A proficient plan covers the expected effectiveness of the remedial technology, projection of the time required to achieve cleanup, and cost. The best way of assessing these factors is to use models to determine the capability of the existing or projected vegetation to contribute to the control and cleanup of the site. No other method has proven applicable and useful for projecting future scenarios, especially for selection and design.

Phytohydraulic containment (1) of contaminated groundwater depends on the evapotranspiration rate exceeding the flow rate of groundwater from the water table down to the lowest extent of the contaminated plume. Phytodegradation and phytotransformation require that the contaminants be in contact with the plants, rhizosphere microorganisms, and any associated exogenous enzymes, all of which can degrade or transform pollutants. Because roots grow from the surface downward to the water table and remain in the capillary zone in most cases, the main factor influencing the effectiveness of phytoremediation is the ability of the plants to draw contaminated water to the roots. Two critical elements are the rate and amount of water withdrawn directly from the aquifer (*versus* soil water from precipitation) and the vertical distribution of contaminants. In other cases to be illustrated in this chapter, plant control of infiltration will limit leaching from contaminated soil. Thus, the local water balance is the vital basis for the assessment and design of phytoremediation to control and remove soil and shallow groundwater contaminants.

The physiologies of vegetation at a site, local hydrology and geology, and the chemical characteristics of the contaminants control the leaching and uptake of contaminated water. The capabilities to transpire large amounts of water, to put roots into the capillary or phreatic zones, and to tolerate stress from other plants and wildlife, contaminants, disease, and occasional lack of nutrients and water are the vital plant attributes for phytocontainment and treatment. The local hydrology involves meteorology (precipitation, wind speed and direction, and air temperature); interception; surface conditions that control runoff *versus* infiltration; percolation through the vadose zone; quantity of flow and direction; and change in storage in the surficial aquifer. The local soil stratigraphy, porosity and hydraulic conductivity of different layers, and hydraulic gradient are the important hydrogeologic factors. The dominant chemical characteristic is the extent of sorption and uptake.

The art of using vadose and groundwater models and measurements to assess and design phytohydraulic control and treatment is progressing rapidly. Hong et al. (1) used modeling to guide the field design of a methyl *tertiary*-butyl ether (MTBE) phytoremediation application in Houston, Texas. They simulated the saturated and unsaturated zones and established that deep-planted (approximately 5 meters or 15 feet) hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN-34' 'Imperial Carolina') trees could control this shallow methyl *tertiary*-butyl ether plume after full canopy development. Field studies are underway to confirm the projections. Quinn et al. (2) estimated the effects of hybrid poplars (*Populus trichocarpa* × *Populus deltoides*) planted to a depth greater than 8 meters (25 feet) into a confined aquifer at Argonne National Laboratory near Chicago, Illinois. Simulations with the U.S. Geological Survey standard groundwater model MODFLOW and the associated particle tracking code indicate that a high degree of containment should be achieved after tree canopy closure. Extensive parts of the trichloroethylene contaminated confined aquifer are expected to be dewatered, or nearly so, on a seasonal basis. Eberts et al. (this book) and Hirsch et al. (this book) along with this chapter, complete a summary of the pioneering work with vadose and phreatic zone modeling to assess and design phytocontainment and treatment systems based on evapotranspiration. Landmeyer (3) reviews some of the measurement techniques that may be useful in phytocontainment and treatment assessments. Ferro et al. (4) reviews measurement and assessment techniques (not involving modeling) applied at a petroleum hydrocarbon contaminated site in Ogden, Utah.

This chapter presents three field examples that illustrate important design considerations for the application of phytoremediation to shallow aquifer contamination. The first case presents a method for estimating the effective evapotranspiration rate for a site in Joliet, Illinois. The second case illustrates the need for vertical characterization of plumes in general by introducing a clear-cut example of a landscape feature that causes plume diving. Hydrogeologic features that cause plumes to migrate deeper into aquifers work against phytoremediation because of the increased evapotranspiration required to control the plume. The second case study also reveals that conventional monitoring approaches that use long-screened wells have a great potential for missing plumes or underestimating the length. The third case explores a proposed phytoremediation system at a site in Florida where a contaminant plume dived upon emerging from under a paved area. Methods to control the vertical location of the plume and potential application of phytoremediation are assessed using the U.S. Geological Survey modular groundwater flow model, MODFLOW (5).

AVERAGED TRANSPORT THROUGH THE VADOSE ZONE

Joliet Army Ammunition Plant

Trinitrotoluene (TNT) and hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) plumes were found in the shallow aquifer at the Joliet Army Ammunition Plant. Figure 1 shows the area in a 1:7200 scale aerial photograph of the site. Of primary interest was a forested area along the banks of Prairie Creek. In this area of the site, there is a possibility either of naturally occurring phytoremediation by existing trees or of enhancing phytoremediation by planting new trees. Either option depends upon the capability of the trees to remove contaminants from the shallow aquifer. The objective of the work at the Joliet plant was to assess the potential for net withdrawal of water from the aquifer.

The photograph reveals some of the irregular nature of the forest cover in this area. A significant fraction of the forested area consists of a lower density of forest cover where lower evapotranspiration rates would be expected.

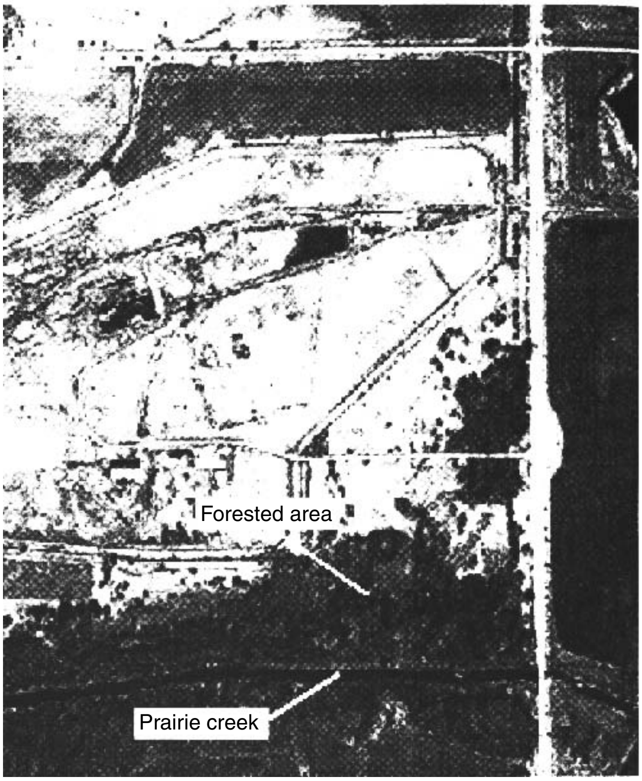

Figure 1. Aerial photograph (1:7200) of the Joliet Army Ammunition Plant site showing Prairie Creek and the forested area overlying the contaminated aquifer. Note that the side of the photograph is oriented north and south.

A vegetation survey conducted at the site showed that the lower density forests contained a number of dead or dying trees. Thus, the role of the existing vegetation in phytoremediation of the contaminant plume depends on the influence of the forest cover and the capability to continue extracting water in the future over this section of the groundwater flow path. Because of the relatively low quality of the existing vegetation, replanting with young trees [hybrid poplars (*Populus* spp.) or willows (*Salix* spp.)] or native prairie grasses might provide a more sustainable treatment system.

Evapotranspiration Estimates

Several equations have been developed for estimating evapotranspiration. The most applicable for this analysis is the Penman-Monteith equation (6). This equation is based on a site energy budget. The budget is based on (1) calculation of all the short- and long-wave energy available for evaporation at the plant canopy, and (2) water vapor diffusing out of the leaves, controlled by the surface or stomatal resistance, and then out into the atmosphere, controlled by aerodynamic resistance (7). The equation for evapotranspiration ET in millimeters per day is as follows

$$ET = \frac{1}{\lambda}\left[\frac{\Delta A + \rho_a c_p \dfrac{D}{r_a}}{\Delta + \gamma\left(1 + \dfrac{r_s}{r_a}\right)}\right] \tag{1}$$

where λ is the latent heat of vaporization equal to $2.501 + 0.002361\ T_w$ in million joules per kilogram, in which T_w is the water temperature at the interface with the air; Δ is the gradient of the saturated vapor pressure in kilopascals per °C; A is the available energy in million joules per square meter per day; ρ_a is the density of moist air in kilograms per cubic meter equal to $3.486P_a/(275 + T_a)$, in which P_a is atmospheric pressure in kilopascals and T_a is air temperature in °C; c_p is the specific heat of moist air equal to 1.013 kilojoules per kilogram per °C; D is the vapor pressure deficit in kilopascals at a reference height; r_a is the aerodynamic resistance in seconds per meter; γ is the psychrometric constant in kilopascals per °C equal to $c_pP_a/(1000\varepsilon\lambda)$, in which ε is the ratio of the molecular weight of water vapor to that of dry air or 0.622; and r_s is the surface resistance of the land cover in seconds per meter that also includes the effect of vegetation. [See Shuttleworth (7) for a description of the reference grass or crop method to estimate relative resistance of a different vegetative land cover.] Most of these variables are strongly related to the meteorological conditions of the site, affirming that the feasibility and success of phytocontainment is very site specific.

A practical approach (7) for estimating evapotranspiration, based on the energy balance and aerodynamic effects, uses generalized long-term climatic data to approximate evapotranspiration. For the evaluation at the Joliet site, existing climatic data were used to estimate evapotranspiration. In the U.S., the Solar and Meteorological Observation Network, SAMSON (8), provides hourly climate observations recorded for a 30-year period (1961 to 1990) from 239 stations located around the country. The data from SAMSON include (1) hourly precipitation (rainfall and snowfall) amounts, (2) air temperature, (3) percent cloud cover, (4) solar radiation, and (5) relative humidity. These data provide the means to estimate averaged hourly, daily, monthly, or annual evaporation amounts for locations in all parts of the U.S. For Joliet, the nearest SAMSON station is located at Chicago. The complete hourly SAMSON file for this station consists of approximately 38 megabytes of data. Practical use of these data in Equation (1) requires post processing to generate, in this case, monthly rainfall and climatic estimates from the 28-year-long record for the Chicago station.

Tables 1 and 2 give a summary of the processed SAMSON data for the Chicago station. These values were used in Equation (1) to generate the forest evapotranspiration estimates given in Table 3. The net extraction, defined here as the potential evapotranspiration less the rainfall, gives an estimate of the amount of water that the forested area near Prairie Creek can remove from the aquifer. For the months of December through March, the net extraction is assumed to be zero because average temperatures were less than 4.4 °C (40 °F). Thus, in the warmer months of the year, the estimated evapotranspiration exceeds the amount of water available from rainfall. Net extraction ranges from 222 to 348 millimeters per month. These correspond to 2.4×10^5 to 3.7×10^5 gallons per acre per month.

Table 1. Monthly Storm Statistics Determined from the SAMSON Database for the Chicago Station

Month	Average Interarrival Time (days)	Average Depth (Millimeters)	Average Storm Duration (hours)	Average Number of Storms
January	1.70	2	0.133	18.0
February	1.59	2	0.123	16.3
March	1.55	3	0.140	19.3
April	1.39	4	0.128	20.3
May	1.69	5	0.113	17.1
June	3.07	6	0.099	15.9
July	2.21	7	0.100	14.2
August	2.13	8	0.105	14.5
September	1.99	6	0.106	14.9
October	2.07	4	0.125	14.5
November	1.72	4	0.143	17.7
December	1.55	3	0.144	19.3

Table 2. Monthly Temperature and Relative Humidity Statistics Determined from the SAMSON Database for the Chicago Station

Month	Average Temperature (°C)	Minimum Temperature (°C)	Maximum Temperature (°C)	Average Dew Point Temperature (°C)	Average Relative Humidity
January	−5.83	−32.2	17.8	−10.19	0.72
February	−3.33	−25.0	21.7	−7.86	0.71
March	2.82	−22.2	30.6	−2.49	0.70
April	9.21	−13.3	32.8	2.17	0.65
May	15.16	−4.4	33.9	7.70	0.64
June	20.47	3.3	39.4	13.10	0.66
July	23.12	5.6	38.9	16.53	0.68
August	22.14	5.0	37.8	16.17	0.71
September	18.01	−1.1	36.1	12.10	0.71
October	11.59	−7.2	32.2	5.35	0.69
November	4.67	−17.2	24.4	−0.19	0.73
December	−2.66	−31.7	21.7	−6.74	0.76

Table 3. Monthly Estimated Evapotranspiration, Rainfall, and Net Extraction Estimates Determined from the SAMSON Database for the Chicago Station and the Forest Evapotranspiration Equation (1)

Month	Estimated Evapotranspiration (Millimeters Per Day)	Estimated Evapotranspiration (Millimeters)	Monthly Precipitation (Millimeters)	Net Extraction (Millimeters)	Net Extraction (Gallons Per Acre)
January	4.87	151	42.7		
February	6.21	174	34.7		
March	9.88	306	66.6		
April	11.34	352	92.3	259	2.8×10^5
May	12.22	379	86.4	293	3.1×10^5
June	14.80	444	95.7	348	3.7×10^5
July	13.85	429	99.0	330	3.5×10^5
August	12.27	380	118	262	2.8×10^5
September	10.42	313	90.6	222	2.4×10^5
October	8.62	267	63.1	204	2.1×10^5
November	5.75	173	77.7	95	0.4×10^5
December	5.50	171	65.2		

The estimated amount of water evaporated and transpired in excess of the rainfall amount was assumed to be drawn from the aquifer. This analysis assumed, however, that the trees can draw water from the aquifer without hydraulic limitation. The depth to the aquifer and the low hydraulic conductivity of the aquifer materials limit the actual rate of withdrawal. Another limitation was the variable density of vegetation in the forest. Dead or dying trees reduce the amount of water extracted from the aquifer. This analysis suggested that in the absence of site-specific evapotranspiration measurements, which would quantify how the existing vegetation actually draws water from the aquifer, there is the potential for the trees to create a net withdrawal of water from the Joliet surficial aquifer during the summer months. Nevertheless, additional site investigation and performance monitoring of any phytoremediation remedy is necessary. These feasibility and application procedures require collection of site-specific climatic and hydrogeologic data to verify and adjust the forecasts from Equation (1). Sap-flow measurements and sampling of tissues from trees are necessary to provide verification of remediation efficacy.

VERTICAL CHARACTERIZATION OF CONTAMINANT DISTRIBUTIONS

Typical Characteristics of Surficial Aquifers

With a positive potential for withdrawal of water from an aquifer, there are at least two important additional considerations. First is the vertical location of contaminants in the aquifer. Before being able to assess the prospects for removal, accurate delineation of contamination is required. The following example describes a case where accurate vertical delineation was used to reveal, in addition to the distribution of contaminants, landscape features that were responsible for diving of the plume (9). Because the contaminants move with the water, the location of aquifer recharge is also an important consideration for phytoremediation to be successful. This case indicates clearly that phytoremediation, which is dependent upon the capability of trees to remove water from the top of an aquifer, requires an understanding of the vertical distribution of the contaminants. The second consideration is the depth to which water is removed from the surficial aquifer. At each site, these recharge and discharge zones are important determinants of the groundwater flow. The importance of characterizing recharge, evapotranspiration, and discharge is illustrated in a second example where phytoremediation was considered to remediate a chlorinated solvent plume. See the section “Feasibility Investigation for Phytohydraulic Containment and Treatment.”

Vertical dimensions that are small relative to the length of the flow system characterize many aquifers. In order to better visualize the stratigraphy or vertical distribution of contaminants, however, commonly drawings are vertically exaggerated, which visually exaggerates the importance of vertical variations. Aquifers and some contaminant plumes can extend over distances of kilometers at least, whereas the thickness is of the order of meters. This results in ratios of length to thickness on the order of 100:1. For analysis of the flow at water tables or in surficial aquifers, the approach taken by Dupuit and Forchheimer (10) recognizes the two orders of magnitude difference in scale. The approach also uses the assumptions that the head in the aquifer is depth-independent (vertical potentials) and that the discharge is proportional to the slope of the water table (Darcy’s equation).

Two more observations are significant. First, because the hydraulic conductivity of a confining layer is orders-of-magnitude lower than that of an aquifer, flow in the confining layer tends to be almost vertical (11). Thus, because of the typical scale, the flow in the aquifer is nearly horizontal as a consequence of vertical potentials and nearly vertical in confining units as a consequence of permeability distributions. Conceptually, the three-dimensional character of flow has been reduced to aquifers with planar flow and confining units with vertical flow. The second important observation concerns the apparent dilution of contaminant plumes that is commonly characterized by aquifer dispersivity. In tabulating dispersivity, Gelhar et al. (12) stated that vertical dispersivities in aquifers were not often measured, but that available data suggested that the values of vertical dispersivity were two orders-of-magnitude lower than longitudinal dispersivity. Therefore, the dilution of plumes over the thickness of the aquifer is likely to be minimal when compared to longitudinal dispersion. The conventional explanation for the difference in dispersivities is that sediments resulting from depositional systems have a preferential orientation—that of the water that deposited the sediments. The practical implication is that transport of contaminants perpendicular to main direction of deposition is minimized.

The conclusions drawn from these observations are that water entering the aquifer vertically is turned so that flow is nearly horizontal. This follows from the Dupuit observations. As recharge accretes along the length of the flow path, more and more water is added to the aquifer. Because of continuity and limited vertical dispersivity, this water is added to the top of the aquifer and contaminant plumes appear to “dive” or move deeper into the aquifer. Likewise, when removed from a surficial aquifer due to evapotranspiration, water is removed from the top and a contaminant plume would appear to rise back to the water table. Such flow characteristics are similar in concept to the classic work of Toth (13) but consistent with the scaling observations of Dupuit and Forchheimer.

At some sites, unlined drainage ditches, leaking water mains and sewer pipes, irrigation, and the flow pattern in the aquifer can determine the vertical distribution of contaminants. Where recharge is the likely cause of diving, the amount of water that infiltrates into the area above the plume, and the amount that this recharge contributes to flow in the aquifer determine where and how much diving will take place.

Plume Diving: East Patchogue, New York

A gasoline release at an East Patchogue, New York, underground storage tank (*i.e.*, UST) facility created large benzene, toluene, ethylbenzene and xylene (BTEX), and methyl *tertiary*-butyl ether (MTBE) plumes. The plumes were detected because the contaminants showed up in a private water supply well, located 1200 meters (4000 feet) down gradient from the source. The well screen was about 15 meters (50 feet) below the water table, where much of the methyl *tertiary*-butyl ether mass was located. The site investigation started at this point and proceeded up gradient to identify the source.

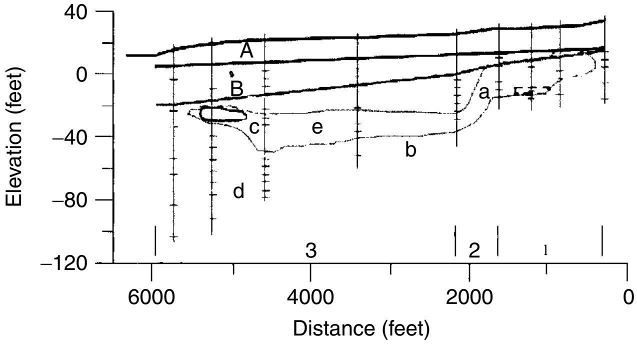

Figure 2. Vertical section through the xylene, benzene, and methyl *tertiary*-butyl ether plumes at East Patchogue, New York. A leaking tank was located at the right-hand edge of the drawing and the flow is to the left. The aquifer was divided into three recharge zones labeled 1, 2, and 3 below the plumes. In zones 1 and 3, the recharge was 559 millimeters per year (22 inches per year), whereas zone 2 recharge dominated by a gravel pit was 1120 millimeters per year (44 inches per year). The water table is marked "A" and the predicted top of the plumes by "B" (see www.epa.gov/athens/onsite for how the top of the plumes was derived). The plume labels are xylene: a—concentrations of 1000 micrograms per liter; benzene: b—100 micrograms per liter and c—1000 micrograms per liter; and methyl *tertiary*-butyl ether: d—100 micrograms per liter and e—1000 micrograms per liter. Note that 1 foot is equal to 0.3048 meter.

Because of the importance of the aquifer for drinking water supply, the State of New York undertook an extensive investigation of the site that included vertical characterization of the plumes. Multilevel samplers with 15-centimeter (6-inch) screens at 1.5-meter (5-foot) intervals were used. A resulting vertical section through the plume showed that BTEX and methyl *tertiary*-butyl ether tended to dive into the aquifer with distance from the source (Fig. 2). Furthermore, a significant amount of diving occurred as the benzene and xylene plumes passed under a gravel pit. At East Patchogue, the maximum depth of the hydrocarbon plumes reached 7.62 meters (25 feet) below the water table.

By studying the well logs and performing a detailed hydraulic characterization of the aquifer with a borehole flowmeter, vertical migration controlled by stratigraphy was ruled out because the hydraulic conductivities varied by less than a factor of two over the aquifer. Weaver et al. (14) present an Internet calculator for plume diving in simple aquifer systems (www.epa.gov/athens/learn2model) that is based on a mass balance and Dupuit-Forchheimer flows in aquifers (15). The calculator reproduces the upper bound on the contaminant distribution by accounting for enhanced recharge from the gravel pit (Fig. 2). Because of the lack of stratigraphic evidence for preferential flow paths and the ability of the model to reproduce the observed plume diving, recharge-driven plume diving is the likely reason for the observed vertical distribution of the contaminants at East Patchague (9).

This example sheds light not only on how recharge pushes the plume downward, but also what happens when water discharges from aquifers. Where water comes up at discharge points, so will the contaminants: along wetlands, streams, rivers, lakes, or the ocean. The ocean is the expected destination of the methyl *tertiary*-butyl ether plume at East Patchogue, where the groundwater discharges into Great South Bay, adjacent to the southern shore of Long Island. The groundwater and contaminants move upward as the plume approaches the discharge point at the bottom of the Bay.

Despite the low vertical dispersivity, contaminant distributions still, of course, vary over the vertical. This was seen in the East Patchogue plume where the maximum concentration at each sample location was higher than the average taken over the entire plume thickness. With conventional sampling techniques, there is effective averaging over the length of the well screens used. With no knowledge of the underlying vertical contaminant distribution, the well screen might be placed anywhere relative to the contaminant distribution. This could easily lead to well screens that sample part of the contaminant distribution and part clean water. The resulting concentrations will be between undetectable and the maximum for the location [(14); see the average borehole concentration calculator that is a part of the OnSite system at www.epa.gov/athens/onsite for a specific example]. Therefore, the same hydrologic characterization of a plume determines the feasibility for phytohydraulic containment and treatment, and also cost effectively dictates when and where the cleanup should be monitored.

FEASIBILITY INVESTIGATION FOR PHYTOHYDRAULIC CONTAINMENT AND TREATMENT

The Orlando Naval Training Center is one of thousands of U.S. Department of Defense facilities contaminated with chlorinated solvents. The plume is shallow and within the rooting depth of proven phytoremediation approaches using tree planting. The Training Center has been closed and the Navy is cleaning up and preparing the site to be turned over to the City of Orlando. As such the Naval Southern Command in Charleston, South Carolina has pilot tested a number of innovative remediation methods. One of the methods investigated was the use of phytoremediation to cost effectively clean up the plume. For this purpose, the EPA National Exposure Research Laboratory in Athens, Georgia was requested by the Navy to establish guidelines for determining the feasibility of applying phytocontainment and treatment and to apply those guidelines at the Orlando Naval Training Center.

The initial phase involved (1) designating appropriate characterization data for an assessment, (2) selecting the appropriate flow and transport models from a multimedia context, (3) vegetation surveys for indigenous phytoremediation potential (conducted by Navy forester), (4) pioneering sap-flow measurements of existing trees and extrapolation to stand-level estimates of existing evapotranspiration for assessments, (5) projecting improved evapotranspiration from replanting, and (6) assessing chlorinated solvent degradation rates in various media by vegetation and microorganisms. This section reports on the selection and application of appropriate models for assessment of the feasibility of phytocontainment and treatment. Existing stand assessments of evapotranspiration and projections were conducted by the U.S. Forest Service Coweeta Hydrologic Laboratory (16) based on sap-flow measurements and extrapolation and projection methods as described in Vose et al. (this book). Nzengung (17) collected vegetation from the site, soil and aquifer cores, and wetland and lake sediments to determine the level of ongoing attenuation. These laboratory incubations provide expected rates and kinetics of degradation and transformation of tetrachloroethylene and trichloroethylene that contaminate the site.

This section covers all other important components of a feasibility investigation—selecting and setting up a groundwater model, and parametrizing that model to simulate the effects of existing and projected vegetation, and runoff and infiltration control. Depth of the plume capture is used as the criteria for projecting the likelihood of using vegetation to capture and treat the plume. As a result, these projections are conservative because the transformation and degradation resulting from contaminant reactions in lake and wetland sediments and from plant enzymatic processes have not been fully assessed to date.

The Orlando Naval Training Center in Florida has a plume of chlorinated solvents that has resulted in detection of significant amounts of vinyl chloride in adjacent Lake Druid shown in Fig. 3. Lake Druid is also bordered by a number of residences and represents a valuable esthetic and recreational asset to adjoining residents.

Site Assessment and Apparent Plume Diving

A groundwater plume of chlorinated ethenes originated at the site of a laundry and dry cleaning facility. Prior to the closing in 1994, hazardous materials and water contaminated with chlorinated solvents were released to the groundwater from the surge tank beside Building 1100, and from spills of tetrachloroethylene and trichloroethylene in the facility. The contaminated groundwater plume flowed from the site of the former laundry underneath and into the nearby Lake Druid. The site plan shows that the source area is contained within an area of very high impervious cover. This consists of the buildings, paved driveways, and parking lots. Immediately down gradient from the parking lot is an unlined drainage ditch. The ditch that receives runoff from the parking lot and buildings on the site runs perpendicular to and crosses the plume that emerges from beneath the pavement. Beyond a narrow grassy area lies a densely wooded area that borders Lake Druid, where the plume discharges. Located nearly over the top of the plume is a creek that feeds the lake. Much of the wooded area is a seepage wetland.

Figure 4 shows the distribution of the chlorinated ethene contaminants taken over a vertical section of the

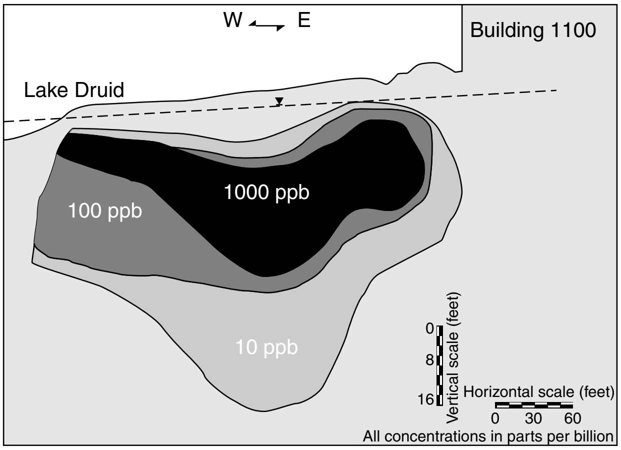

All concentrations in parts per billion

Figure 4. Vertical section of chlorinated ethene plume at Orlando Naval Training Center. Lake Druid is on the left and the contaminant source on the right. Note that 1 foot = 0.3048 meter and ppb is the concentration in micrograms per liter.

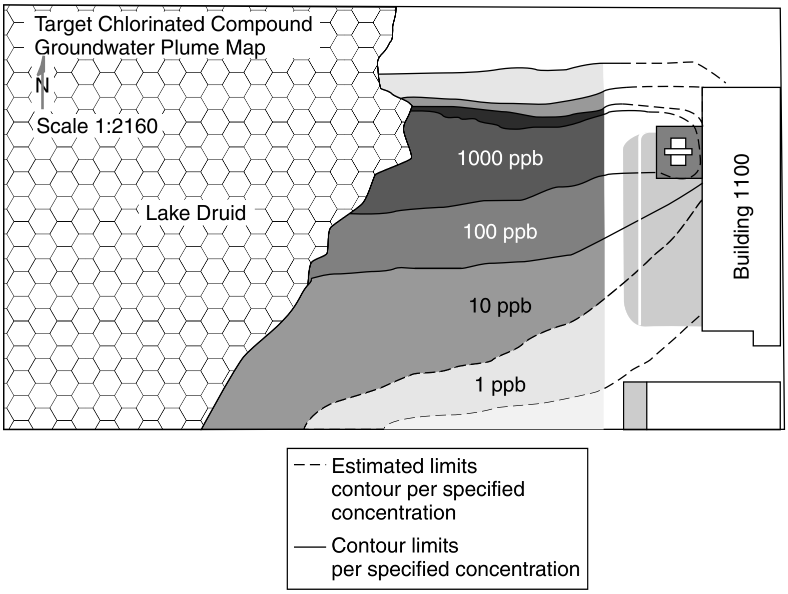

Figure 3. Site map of the Orlando Naval Training Center showing the Building 1100 source area and the shoreline of Lake Druid. The light shading surrounding Building 1100 delineates impervious parking lots and roads. Note that ppb is the concentration in micrograms per liter.

aquifer spanning the distance from the former laundry to the lakeshore. The data were obtained using a direct push probe operated on a 0.6-meter (2-foot) interval. These results show that the concentrations are low in the upper part of the aquifer, once the plume is down gradient of the ditch. In these locations the depth of uncontaminated water having undetectable amounts of chlorinated ethenes ranges up to 2.4 meters (7 feet) below the water table. The horizontal distance spanned by the diving portion of the plume is approximately 67 meters (220 feet). Thus the plume moves deeper into the aquifer upon emergence from the paved area and upon passing under the unlined ditch.

Model Application for Orlando Naval Training Center

The effects of recharge and discharge on the groundwater flow at Building 1100 were simulated with two models. First, the simple screening model for plume diving was applied (14). The model assumes that flow occurs between two points with fixed heads, the aquifer is a single uniform layer, and the flow system can be described by a set of one-dimensional segments. These assumptions are satisfied for a screening analysis of the Orlando Naval Training Center plume, although other details of the flow system require the use of a numerical model. This consideration led to the application of MODFLOW with particle tracking for a more detailed evaluation.

The screening-level plume diving calculation allows determination of the vertical displacement of the plume generated by the supply and withdrawal of water from the aquifer. The simple model was used to assess the amount of diving that would occur as the plume passed under the unlined ditch. With an estimated average annual rainfall of 1400 milliliters (55 inches) at Orlando, the amount of water that could run off into the ditch is very large. Estimates based on the area of the pavement surrounding the source ranged up to the equivalent of 38000 millimeters per year (1500 inches per year) concentrated in the ditch. The actual amount of infiltration through the ditch was not determined for this study. However, the leakage through the ditch would be a function of the transient depth of water, soil properties, soil antecedent moisture content, and the vegetation in the ditch. The online calculator (www.epa.gov/athens/onsite) was used to determine how much water would be required to infiltrate through the ditch to cause the observed diving of the plume. The purpose of this exercise was to determine if the observed plume diving could be attributed to reasonable recharge amounts ranging from the rainfall rate of 1400 milliliters (55 inches) to the estimated runoff amount of 38000 millimeters per year (1500 inches per year). Figure 5 shows the results from infiltration amounts of 1400 milliliters per year (55 inches per year) and 2800 millimeters per year (110 inches per year). Thus, if the ditch recharges double the average annual precipitation to the aquifer from the impervious area, this amount can account for the observed plume diving. This amount of water 2800 millimeters per year (110 inches per year) is less than 10 percent of the maximum estimated flow in the ditch and thus forms a plausible explanation for the vertical distribution of the contamination observed at the site. In the upper portion of the aquifer, there is no evidence for stratigraphic preferential flow paths and density-driven plume diving is not evident because the plume dives only after emerging from below the pavement.

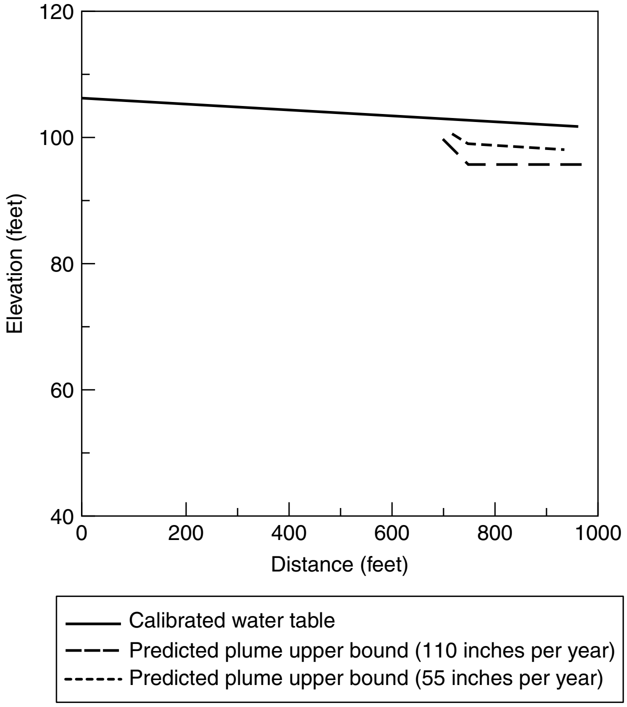

Figure 5. Predicted plume diving using the OnSite plume diving calculator (www.epa.gov/athens/onsite) and two estimates of infiltration through the unlined ditch. Note that 1 foot is 0.3048 meter and 1 inch is 25.4 millimeters.

Numerical Modeling

The U.S. Geological Survey modular groundwater flow (MODFLOW) code with particle tracking was used to perform more detailed simulations of the site. These simulations provided a more precise determination of the vertical and horizontal distribution of flow and a tool with which to evaluate the various design alternatives. Some of the parameters used in the model were taken from a previous modeling study (18) that was not focused on the plume and Lake Druid. Halford (18) calibrated the model to regional flow conditions. Calibrated parameters for hydraulic conductivity, porosity, and aquifer and layer thicknesses from Halford (18) were used in the current model. Boundary conditions were taken to match observed water levels collected around the site and consistent with the large-scale flows simulated by Halford (18). The need for site-scale water levels constrained the modeled flow domain to the area immediately surrounding Lake Druid. For an adequate assessment of the potential effects of vegetation, the area around Lake Druid and Building 1100 was modeled in more detail, both horizontally and vertically. To provide sufficient vertical delineation of the plume, the model simulations were based on 12 layers to adequately resolve the lake bathymetry. The refinements provided sufficient resolution to explore the factors that control the vertical movement of the contaminants moving from Building 1100 to Lake Druid.

The amount of transpiration from either the existing vegetation or by any proposed plantings was difficult to estimate. A generalized estimate of 1300 millimeters per year (51 inches per year) was available from the Florida Agricultural Extension Service for central Florida. An estimate of 500 millimeters per year (20 inches per year) was determined for nearby orange groves (19). Finally, a site-specific estimate of 200 to 300 millimeters per year (8 to 12 inches per year) was made by Vose and Elliot (16) for the existing mature trees over the down gradient end of the plume. The estimate was made by measuring the sapflow rate in nine trees in spring, summer, and fall. The estimate included only mature trees and did not account for the understory vegetation. Vose and Elliot (16) believed that this might not be too severe a limitation because the understory may be cycling shallow soil moisture, and they cite studies that show that understory transpiration can offset reduced overstory transpiration in open stands. The potential evapotranspiration at full stocking of the stand was 600 to 900 millimeters per year (24 to 35 inches per year), a higher value because of the relatively sparse distribution of the existing mature trees (also see Vose et al. this book). With the estimated precipitation of 1400 millimeters per year (55 inches per year) in central Florida, the estimated evapotranspiration ranges from 14 to 93 percent of that value. The lower values reflected the site-specific measurements.

These values suggested that in any case there was no strong driving force for removal of water from the aquifer by the vegetation. The generalized or regional estimate of evapotranspiration of 1300 millimeters per year (51 inches per year) came close to balancing precipitation. With the site-specific estimate of 200 to 300 millimeters per year (8 to 12 inches per year) accounting for some of the precipitation, the remainder would evaporate, contribute to soil moisture, or run off.

The MODFLOW results reproduced the presumed flow pattern in the cross section of the aquifer (Fig. 6). The simulated downward flow just down gradient of the paved area accurately reflects the observed plume (Fig. 4). The 12-layer representation of the aquifer allows adequate representation of the bathymetry of the lake in order to properly resolve the pattern of groundwater flow to the bottom of the lake. The discharge pattern was consistent with measurements of the attenuated plume emerging into the lake (data not shown in this chapter). The significance of these simulations was that the recharge and discharge patterns of the aquifer alone, can explain the vertical distribution of the contaminants (Fig. 4). Furthermore, the use of 12 thin layers to represent the surficial aquifer is unusual [see for example (18)] but necessary to represent a plume in this hydrogeologic setting.

These observations lead to the following strategies for phytoremediation at the site. First, the prospects for the trees transpiring sufficient water to pull the contaminant plume up to the root zone for treatment is too unlikely for phytoremediation alone to be considered as an adequate

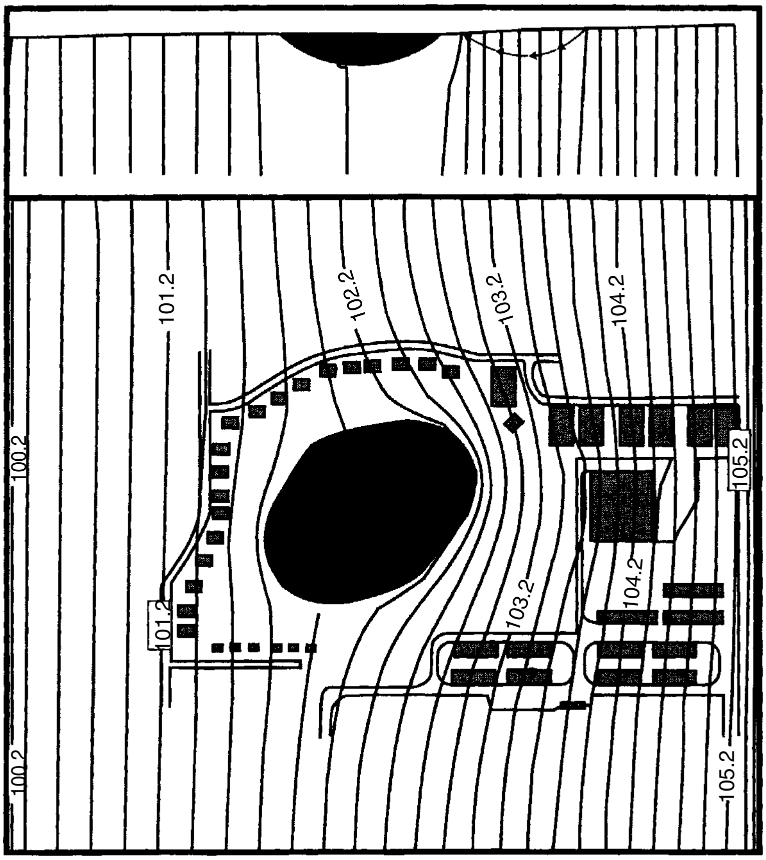

Figure 6. MODFLOW simulation of part of the Orlando, Florida U.S. Naval Training Center showing the effects of infiltration through the unlined ditch and upwelling at the shoreline. The upper panel represents a vertical section through the source and Lake Druid. The nearly vertical lines represent constant head contours, which correspond to the contours (labeled in feet above mean sea level; note that 1 foot is 0.3048 meter) on the plan view shown in the lower panel. The arrows in the upper panel represent a flow path from the unlined ditch to a discharge point along the shoreline. The scale is approximately 1:8280. The sides of the lower panel are vertically oriented north (top) to south (bottom).

cleanup remedy. This follows from the fact that the existing vegetation, extracting at least of 200 millimeters per year (8 inches per year), has not contained the existing contaminant plume.

Because the transpiration of trees alone cannot contain the plume, clearly the combined effects of replanted trees and natural attenuation in the seepage wetland and lake sediments needed to be investigated next. Nzengung (17) found significant potential for microbial and enzymatic degradation of chlorinated solvents in the wetland and lake sediments, thus requiring a refined hydraulic analysis. The refined analysis established that enhanced evapotranspiration would pull some of the plume upward into the seepage wetland and probably achieve better attenuation. This next investigation established more contact with a wider range of shallow rooted aquatic and wetland plants at the lakeshore without threatening the health of the existing wetland (Medina et al. this book). Nevertheless, the quantification of attenuation rates (17) was deemed too uncertain to forecast whether this enhanced natural attenuation would be sufficient without a field pilot investigation. These pilot investigations are currently (2003) underway. Because forecasting was not possible for the enhanced natural attenuation option, other supplemental design scenarios were also investigated. These included diversion of the surface runoff (from the impervious parking lots and roofs) from the unlined ditch to alleviate some plume diving, and planting more trees through the parking lot cover and over the source to contain the plume. Because of the ecological sensitivity of the seepage wetlands, an investigation of alternative planting was useful to avoid any temporary disruption of the wetland with supplemental planting if possible, especially to avoid planting alien species that are of regulatory concern to the State of Florida.

Simulations of eliminating the localized recharge to the aquifer at the edge of the parking lot established that most but not all of the downward movement of plume below the water table would be eliminated. However, some infiltration through the down gradient tree and shrub cover would continue because rainfall in this humid area exceeds potential evapotranspiration (Vose et al. this book). In addition, this sandy aquifer has high conductivity that limits contact of the plume with roots, wetland soils, and lake sediments, even under the best of conditions.

Finally, if trees were planted over the source zone, the simulations demonstrate that transpiration could remove water from the upper part of the aquifer despite the hydraulic conductivity. With the proper selection of trees that transform and degrade chlorinated solvents, some of the contaminated water could be treated, leaving less for natural attenuation. This could have the effect of removing contaminants from the aquifer, before the plume enters wetland soil and lake sediments. The MODFLOW results show that upward gradients would be established through the upper 7 meters (23 feet) of the aquifer if the evapotranspiration rate was 800 millimeters per year and through the upper 10 meters (30 feet) of the aquifer if the evapotranspiration was 1300 millimeters per year. These amounts exceed the site specific measured values of 200 to 300 millimeters per year (16) because the observed values were net values that integrated the effect of the trees working against infiltration. The higher values used in the model were selected because recharge was assumed to be zero due to the presence of the pavement. Forcing the trees to draw water from the aquifer exclusively over the source area maximized the impact on the contaminated portion of the aquifer by maximizing the withdrawal of water from the source zone. Overall, runoff diversion, tree planting over the source, and enhanced natural attenuation in the seepage wetland and nearer the lakeshore have the potential to manage contamination at this site but the uncertainty in enhanced attenuation requires a field pilot test to be sure this is the best remedy for this particular plume.

FEASIBILITY AND DESIGN IMPLICATIONS

Design of phytoremediation systems requires delineation of contaminant plumes and the parameters governing flow and transport in shallow aquifers. Because trees remove water and contaminants from the aquifer, estimates of evapotranspiration are needed to design phytocontainment systems. For the Joliet Army Ammunition Plant, the estimates were based on a forest evaporation model (7) using the Penman-Monteith equation (6) with climate data from the U.S. SAMSON database. From these readily available data, the feasibility of phytoremediation can be assessed.

Vertical delineation of contaminant plumes is best accomplished with short-screened or point sampling that eliminates (1) mixing of clean and contaminated water caused by screen placement that is only partly contained within the plume, and (2) averaging of concentration over long well screens. Both of these sampling effects or artifacts cause contaminant concentrations to be underestimated. Further, if the horizontal and vertical boundaries of the plume are not accurately defined, then transport processes are not correctly determined and phytoremediation potential can not be accurately assessed.

Sampling at the Orlando Naval Training Center revealed that the chlorinated solvent plume was depressed once the plume emerged from below the paved parking lot located adjacent to the source. Because of the presence of an unlined ditch adjacent to the parking lot, the plume dived due to localized recharge from the bottom of the unlined ditch. This plume characteristic was quantified with two models. These models included a simplified analytical solution of flow and a three dimensional numerical model of the site (MODFLOW) to simulate groundwater flow. Like the field data, the numerical model needed to be sufficiently resolved in the vertical to simulate the vertical distribution of flow and the bathymetry of the lake. Thus, a greater-than-usual number of layers were required for this phytocontainment assessment.

Both of these models showed that with reasonable rates of recharge and evapotranspiration and using some regional scale MODFLOW parameters, the observed plume diving could be accounted for just from the hydraulics of flow in the aquifer. Thus, the application of phytoremediation to this site was established as critically dependent upon the localized pattern of flow in the aquifer.

The analysis of the Naval Training Center site shows that the localized plume characteristics must be known to plan the locations for tree planting or design other remedial options for the site.

At the Naval Training Center, observations and simulations show that the contaminated water discharges to Lake Druid. Thus, the shoreline vegetation has the potential to intercept and treat the contaminated aquifer water. This interception of the groundwater flow is dependent upon evapotranspiration and all other characteristics of flow in the aquifer. Other elements of a remedy at this site include plume control by eliminating runoff from the parking lot to the unlined ditch. This would have the effect of preventing some of the downward migration of the contaminants. The depth of contamination at the source, however, is approximately 10 meters (34 feet). The depth of contamination is due to the location of the sump and the denser-than-water nature of the contaminants. If trees were planted and successfully grown at the source, then some of the contaminated water would migrate upward. The effectiveness of this source control depends upon the evapotranspiration rate that could be achieved by the trees and by the continued prevention of recharge in the parking lot.

Overall, runoff diversion, source phytocontainment, and phyto-enhancement of natural attenuation may be sufficient to remediate this site over a long period (until the nonaqueous phase is completely dissolved). However, due to the uncertainty in extrapolating treatment and attinuation kinetics, field pilot studies are necessary. These pilot investigations are underway at the Orlando site and other source removal remedies have been tested.

Acknowledgments

Dan Reible and John Quinn provided insightful and through peer reviews used in the evaluation of this chapter. Jerry Schnoor accepted a revised version of the chapter for this book. Support for the work came from the U.S. Environmental Protection Agency Office of Research and Development Contaminated Sites Program and a cooperative interagency research agreement with the U.S. Navy Southern Command. Cliff Casey of the U.S. Navy was instrumental in conceiving of and pursuing the work at the Orlando Naval Training Center. This chapter has been reviewed in accordance with the U.S. Environmental Protection Agency peer and administrative review policies and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

BIBLIOGRAPHY

1. Hong, M.S. et al. (2001). Phytoremediation of MTBE from a groundwater plume. *Environ. Sci. Technol.* **35**(6): 1231–1239.
2. Quinn, J., M.C. Negri, R.R. Hinchman, L.P. Moos, and J.B. Wozinak (2001). Predicting the effect of deep-rooted hybrid poplars on the groundwater flow system at a large-scale phytoremediation site. *Int. J. Phytoremed.* **3**(1): 41–60.
3. Landmeyer, J.E. (2001). Monitoring the effect of poplar trees on petroleum-hydrocarbon and chlorinated-solvent contaminated ground water. *Int. J. Phytoremed.* **3**(1): 61–85.
4. Ferro, A., J. Chard, R. Kjelgren, B. Chard, D. Turner, and T. Montague (2001). Groundwater capture using hybrid poplar trees: evaluation of a system in Ogden, Utah. *Inter. J. Phytoremed* **3**(1): 87–104.
5. McDonald, M.G. and A.W. Harbaugh (1988). *A modular three-dimensional finite-difference ground-water flow model*. U.S. Geological Survey Techniques of Water Resources Investigations Book 6, Chapter A1.
6. Jensen, M.E., R.D. Burman, R.G. Allen, eds. (1990). *Evapotranspiration and Irrigation Water Requirements*. Manual No. 70. American Society of Civil Engineers, New York.
7. Shuttleworth, J. (1993). Evaporation. In: *Handbook of Hydrology*. D.R. Maidment (Ed.). McGraw-Hill, New York, pp. 4.1–4.53.
8. U.S. Department of Commerce (1993). *Solar and Meterological Surface Observation Network 1961–1990*. National Climatic Data Center (http://www.ncdc.noaa.gov).
9. Weaver, J.W., J.T. Haas, and J.T. Wilson (1996). Analysis of the gasoline spill at East Patchogue, New York. In: *Proceedings of Non-Aqueous Phase Liquids (NAPLs) in Subsurface Environment: Assessment and Remediation*. L. Reddi (Ed.). American Society of Civil Engineers, Washington, DC, November 12–14, pp. 707–718.
10. Bear, J. (1972). *Dynamics of Fluids in Porous Media*. Dover, New York.
11. Charbeneau, R.J. (1999). *Groundwater Hydraulics and Pollutant Transport*. Prentice Hall, Upper Saddle River, New Jersey, pp. 593.
12. Gelhar, L.W., C. Welty, and K.R. Rehfeldt (1992). A critical review of data on field scale dispersion in aquifers. *Water Resour. Res.* **28**(7): 1955–1974.
13. Toth, J., (1963). A theoretical analysis of groundwater flow in small drainage basins. *J. Geophys. Res.* **68**: 4375–4387.
14. Weaver, J.W., C.B. Sosik, and J.W. Washington (2001). *OnSite: the online site assessment tool*. U.S. Environmental Protection Agency National Exposure Research Laboratory. Athens, Georgia. (www.epa.gov/athens/onsite) (In review).
15. Strack (1984). Three-dimensional streamlines in Dupuit-Forcheimer flow. *Water Resour. Res.* **20**(7): 812–822.
16. Vose, J.M. and K.J. Elliot (1999). *Preliminary estimates of transpiration at the Naval Training Center, Orlando, Florida: results from phase I.* Report by the U.S. Department of Agriculture Forest Service Coweta Hydrologic Laboratory, Otto, North Carolina.
17. Nzengung, V.A. (2000). *Data on phytotransformation and natural attenuation processes for the degradation of perchloroethene (PCE) and trichloroethene (TCE): data specific for the Orlando Naval Training Center (NTC), Orlando, Florida, task 1 to 4 (separate reports)*. Contract report. Contract #: UKC019 (QT-RT-99-002122), U.S. Environmental Protection Agency National Exposure Research Laboratory, Athens, Georgia.
18. Halford, K.J. (1998). *Assessment of the potential effects of phytoremediation on ground-water flow around area C at Orlando Naval Training Center, Florida*. U.S. Geological Survey Water Investigations Report 98–4110.
19. Summer, D.M. (1996). *Evapotranspiration from successional vegetation in a deforested area of the Lake Wales Ridge*, Florida, U.S. Geological Survey Water-Resources Investigations Report 96-4244, 37 pp.

READING LIST

Pennington, J.C., D.W. Harrelson, M. Zakikhani, D. Gunnison, J. Clarke, C.J. McGrath, H. Fredrickson, J.H. May, E. Perkins, C.A. Hayes, and D. Ringelberg (1998). *Feasibility of using natural attenuation as a remedial alternative for explosives-contaminated groundwater at site L1*. Joliet Army Ammunition

Plant, Joliet, Illinois. U.S. Army Corps of Engineers, Waterways Experiment Station Technical Report EL-98-8, Vicksburg, Mississippi.

Vogel, T., K. Huang, R. Zhang, M.Th. van Genuchten (1996). *The HYDRUS code for simulating one-dimensional water flow, solute transport, and heat movement in variably-saturated media*. U.S. Salinity Laboratory, Agricultural Research Service, U.S. Department of Agriculture. Research Report 140, Riverside, California.

WELL DESIGN AND CONSTRUCTION

Nikolay Voutchkov
Poseidon Resources Corporation
Stamford, Connecticut

INTRODUCTION

Wells are used to collect groundwater from alluvial fresh water aquifers, saline water from brackish water aquifers, and seawater from beach shore aquifers. Wells are relatively simple to build, and the source water they collect is pretreated via slow filtration through the subsurface granular geologic formations in the area of source water extraction. Therefore, source water from wells is usually of better quality in terms of solids, silt, oil and grease, natural organic contamination, and pathogen content compared to source water collected from open surface water (i.e., river, lake, or ocean) intakes.

Key well performance characteristics are yield, static and pumping water levels, and cone of depression. Well yield indicates how much water can be withdrawn from a given well for a preset period of time. Well yield is typically measured in cubic meters per second (m^3/s) or cubic feet per second (cfs) for large wells and in liters per second (lps) or gallons per hour (gph) for small wells. Pumping and static water levels are the groundwater levels in the well when pumping from the well is on and off, respectively. When a well is operational, the surface of the groundwater level in the aquifer takes an inverted cone shape due to directional water flow toward the well. This inverted shape is called a cone of depression (Fig. 1).

Well type, size, and capacity are determined by a hydrogeologic investigation which typically includes the following key steps:

1. Complete a preliminary geologic survey to identify if the selected site is generally suitable for constructing of subsurface water intake.
2. Drill-test to collect samples of the aquifer formation deposits for visual classification and grain-size distribution analysis.
3. Install one or more test wells and observation wells, and conduct a pumping test to determine the site-specific hydraulic characteristics of the aquifer for subsurface system design and determination of the intake system yield.
4. Collect an adequate amount of samples of the source water and analyze sample water quality, with special emphasis on the content of iron, manganese, barium, strontium, silica, radon, carbon dioxide, arsenic, and hydrogen sulfide in the source water. If the aquifer water quality is under the influence of a surface water source (i.e., river, lake, ocean) whose quality and quantity vary seasonally, then complete year-round intake water quality sampling to determine seasonal fluctuations of source water quality.
5. If the subsurface intake system requires multiple collection facilities (wells, or infiltration galleries/river bank filtration facilities), then complete a computer model analysis to establish the response of the production aquifer to pumping and the potential impact of groundwater collection on adjacent fresh or saline water aquifers, which could interact with the water supply aquifer.

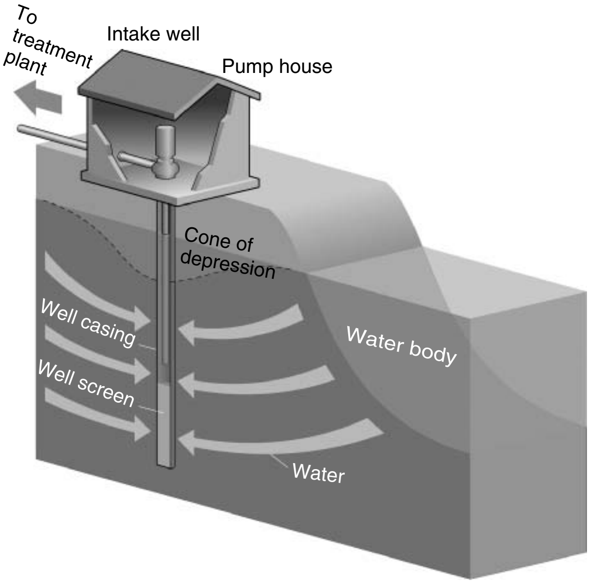

Figure 1. Vertical intake well.

The key factors that determine if the use of subsurface intake is practical or/and economical are the type of the water source aquifer (confined vs. unconfined); the aquifer permeability (hydraulic conductivity), which is a measure of the velocity of water movement through the ground (typically measured in m/s); the average specific yield (productivity) of the aquifer (in m^3/day per linear meter of riverbank or seashore along which the collector wells are located); the thickness of the production aquifer deposits; and the existence of nearby fresh or brackish water aquifers, which could be negatively impacted by the intake well operations or may have a measurable effect on intake well water quality. A confined aquifer (also often referred to as an artesian aquifer) is a water saturated geologic formation between two layers of low permeability (i.e., bedrock) that restrict the vertical movement of the groundwater in or out of the aquifer. Confined aquifers are often contained by the surrounding geologic formations and are pressurized therefore, collecting water from such aquifers may not require pumping. Unconfined aquifers are groundwater saturated formations whose upper surface (commonly referred to as the water table) fluctuates over time. Aquifer water table fluctuation is driven by recharge from surface runoff (rain or snowmelt)

Previous Page

Plant, Joliet, Illinois. U.S. Army Corps of Engineers, Waterways Experiment Station Technical Report EL-98-8, Vicksburg, Mississippi.

Vogel, T., K. Huang, R. Zhang, M.Th. van Genuchten (1996). *The HYDRUS code for simulating one-dimensional water flow, solute transport, and heat movement in variably-saturated media*. U.S. Salinity Laboratory, Agricultural Research Service, U.S. Department of Agriculture. Research Report 140, Riverside, California.

WELL DESIGN AND CONSTRUCTION

Nikolay Voutchkov
Poseidon Resources Corporation
Stamford, Connecticut

INTRODUCTION

Wells are used to collect groundwater from alluvial fresh water aquifers, saline water from brackish water aquifers, and seawater from beach shore aquifers. Wells are relatively simple to build, and the source water they collect is pretreated via slow filtration through the subsurface granular geologic formations in the area of source water extraction. Therefore, source water from wells is usually of better quality in terms of solids, silt, oil and grease, natural organic contamination, and pathogen content compared to source water collected from open surface water (i.e., river, lake, or ocean) intakes.

Key well performance characteristics are yield, static and pumping water levels, and cone of depression. Well yield indicates how much water can be withdrawn from a given well for a preset period of time. Well yield is typically measured in cubic meters per second (m^3/s) or cubic feet per second (cfs) for large wells and in liters per second (lps) or gallons per hour (gph) for small wells. Pumping and static water levels are the groundwater levels in the well when pumping from the well is on and off, respectively. When a well is operational, the surface of the groundwater level in the aquifer takes an inverted cone shape due to directional water flow toward the well. This inverted shape is called a cone of depression (Fig. 1).

Well type, size, and capacity are determined by a hydrogeologic investigation which typically includes the following key steps:

1. Complete a preliminary geologic survey to identify if the selected site is generally suitable for constructing of subsurface water intake.
2. Drill-test to collect samples of the aquifer formation deposits for visual classification and grain-size distribution analysis.
3. Install one or more test wells and observation wells, and conduct a pumping test to determine the site-specific hydraulic characteristics of the aquifer for subsurface system design and determination of the intake system yield.
4. Collect an adequate amount of samples of the source water and analyze sample water quality, with special emphasis on the content of iron, manganese, barium, strontium, silica, radon, carbon dioxide, arsenic, and hydrogen sulfide in the source water. If the aquifer water quality is under the influence of a surface water source (i.e., river, lake, ocean) whose quality and quantity vary seasonally, then complete year-round intake water quality sampling to determine seasonal fluctuations of source water quality.

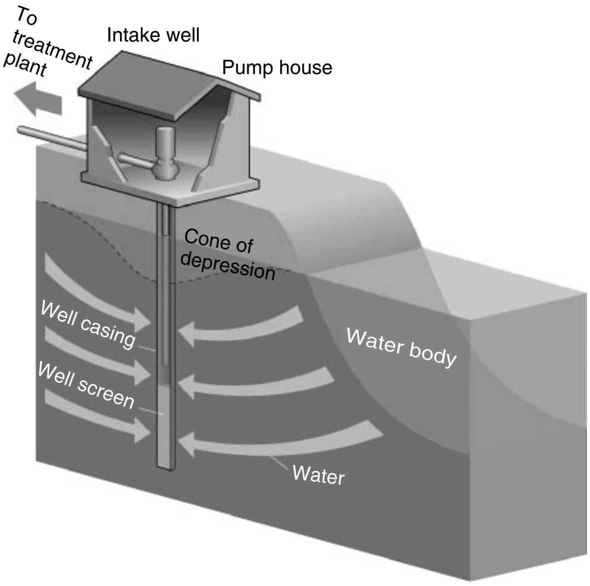

Figure 1. Vertical intake well.

5. If the subsurface intake system requires multiple collection facilities (wells, or infiltration galleries/river bank filtration facilities), then complete a computer model analysis to establish the response of the production aquifer to pumping and the potential impact of groundwater collection on adjacent fresh or saline water aquifers, which could interact with the water supply aquifer.

The key factors that determine if the use of subsurface intake is practical or/and economical are the type of the water source aquifer (confined vs. unconfined); the aquifer permeability (hydraulic conductivity), which is a measure of the velocity of water movement through the ground (typically measured in m/s); the average specific yield (productivity) of the aquifer (in m^3/day per linear meter of riverbank or seashore along which the collector wells are located); the thickness of the production aquifer deposits; and the existence of nearby fresh or brackish water aquifers, which could be negatively impacted by the intake well operations or may have a measurable effect on intake well water quality. A confined aquifer (also often referred to as an artesian aquifer) is a water saturated geologic formation between two layers of low permeability (i.e., bedrock) that restrict the vertical movement of the groundwater in or out of the aquifer. Confined aquifers are often contained by the surrounding geologic formations and are pressurized therefore, collecting water from such aquifers may not require pumping. Unconfined aquifers are groundwater saturated formations whose upper surface (commonly referred to as the water table) fluctuates over time. Aquifer water table fluctuation is driven by recharge from surface runoff (rain or snowmelt) or fluctuation of the water table of the surface waterbody (river, lake, etc.) connected to the aquifer.

Unconfined alluvial aquifers located along riverbanks are commonly used for freshwater supplies. Confined or semiconfined aquifers yield the most suitable source of water for brackish water desalination systems (1). Aquifers that favor the construction of wells are coarse-grained, porous, and highly permeable geologic formations (i.e., sand, gravel, and limestone) connected to a riverbed (for freshwater intakes) or the ocean floor (for seawater intakes) whose specific yield (transmissivity) exceeds 1,500 m^3/day/m and whose water carrying zone is at least 6 meters (20 feet) deep (2). The higher the aquifer permeability, transmissivity, and thickness, the larger the well yield the aquifer can support.

WELL SITE SELECTION

Once an aquifer of suitable yield has been identified, the location of the actual field is to be selected. Whenever possible, the well field should be located perpendicularly to the main direction of groundwater flow (i.e., along the banks of a nearby river). The most suitable location for collecting a large amount of river water is on a river island or within a meander. If significant clogging of the riverbed is expected, then the wells should be located closer to the riverbanks.

In addition, a potable water well field has to be sited away and uphill from potential contamination sources such as septic tanks, landfills, industrial plants, and underground fuel storage tanks. Most states regulate the location of private potable water wells and require locating wells at least 15 to 30 meters (50 to 100 feet) away from septic systems.

INTAKE WELL TYPES AND KEY DESIGN CRITERIA

Intake wells are typically vertical or horizontal water collectors drilled into the source water aquifer. Vertical intake wells are usually less costly than horizontal wells, but their yield is relatively small, typically, 0.004 to 0.044 m^3/s (0.1 to 1.0 MGD). Therefore, vertical wells are typically used for supplying relatively small quantities of water usually less than 20,000 m^3/day (5 MGD).

Vertical Intake Wells

Vertical intake wells consist of the following key elements: a casing, well screen, filter pack, well seal, and surface seal. Many wells have submersible or vertical turbine pumps installed inside the well casing (Fig. 1). A well casing is a steel or nonmetallic (typically, fiberglass) pipe which lines the well borehole to protect the well from caving in. Most state standards require steel casings for freshwater supply wells and specify the thickness of the casing. The diameter of the casing has to house the well intake pump and provide ample room for pump service. The diameter of the well casing is determined mainly by the well screen size and yield, but the well borehole diameter has to be at least 0.1 meters (4 in.) larger than the well casing to accommodate the installation of the well seal. Usually, the well casing diameter is between 200 and 1200 mm (8 and 48 in.), and the well depth is typically less than 75 meters (250 feet).

The well screen is the intake portion of the well and is a sieve-like structure that has slotted or perforated openings. The well screen is located at a depth corresponding to the water carrying zone of the aquifer. Screen depth, size of openings, diameter, and length are key well performance design criteria. These well parameters are selected to maximize a well's safe yield; control well entrance velocity; and to avoid excessive entrance of sand and other particulates, which have a negative impact on well useful life and water quality. The performance of the well screen is enhanced by a gravel (filter) pack, which consists of clean, uniform, and well rounded gravel and sand placed between the borehole wall and the well screen to prefilter the groundwater entering the well. Typically, the gravel pack depth extends at least 1 meter (3 feet) above the well screen. A well seal is installed above the filter pack to prevent soil and contaminants from entering the well screen area. The well seal is a cylindrical layer of cement, bentonite, or clay placed in the annulus of the well between the well casting and the borehole. Typically, the well seal extends at least 0.6 meters (2 feet) above the top of the gravel pack and usually through the elevation of the soil frost zone. The aboveground portion of the well is finished with a concrete surface seal. The surface and well seals protect the well from surface runoff contamination and support the casing.

A detailed methodology for sizing well screen diameter, length, openings and capacity, and for selecting the grain size of the material for the gravel pack is presented elsewhere (3). This source also provides information on determining the thickness of the gravel pack around the well screen, the permissible approach velocity of the water stream converging to the well, and ultimately the design capacity of a new well or the maximum capacity of an existing well.

Once constructed, vertical well performance has to be monitored frequently to secure long-term performance and identify early signs of potential malfunction and failure. The most common causes of well failure are borehole collapse, corrosion of the casing, improper or defective construction techniques, growth of organisms within the well borehole, and formation of mineral concentrations or crusts in the open-hole or screened section of the well borehole.

Horizontal Collector Wells

Horizontal (Ranney) collector wells consist of a caisson that extends below the ground surface and has water well collector screens (laterals) projected horizontally from inside the caisson into the surrounding aquifer (Fig. 2). The well screens in the collector wells are placed horizontally, but a higher rate of source water collection is possible than that of vertical wells, which allows, collecting the same intake water quantity with fewer wells.

Individual horizontal intake wells are typically designed to collect from 0.0044 to 1.75 m^3/s (0.1 to 40.0 MGD) of source water per well. The largest 1.75 m^3/s (40 MGD) horizontal collector wells in the United States were installed on the banks of the Platte River in 1990

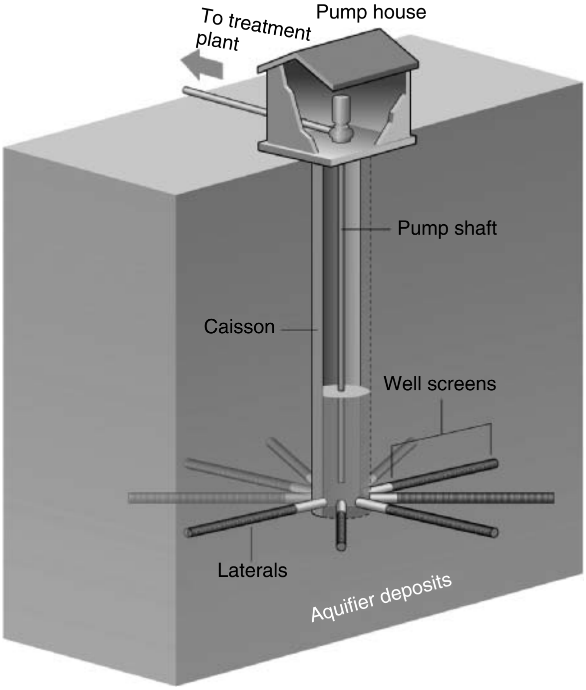

Figure 2. Horizontal collector well.

for water supply to the City of Lincoln, Nebraska. In the United States, utilities use horizontal collector wells whose laterals are located directly under the riverbed.

The caisson of the horizontal collector well is constructed of reinforced concrete of 2.7 to 6.0 m (9 feet to 20 feet) inside diameter and a wall thickness of approximately 0.5 to 1.0 meters (1.5 to 3.0 feet). The caisson depth varies according to site-specific geologic conditions, ranging from approximately 10 meters to over 45 meters (30 to 150 feet).

The number, length, and location of the horizontal laterals are determined by a detailed hydrogeologic investigation. Typically, the diameter of the laterals ranges from 0.2 to 0.3 meters (8 to 12 inches) and their length extends up to 60 meters (200 feet). The size of the lateral screens is selected to accommodate the grain size of the underground soil formation. If necessary, an artificial gravel-pack filter can be installed around the screens to suit finer grained deposits. Usually, one well has 2 to 14 laterals.

When horizontal wells are used for fresh or brackish groundwater intake, the caisson is extended above the floodplain elevation for this location to protect the pumping equipment, electrical, and instrumentation and control equipment from flooding. When used as seawater beach wells, especially for smaller size applications, Ranney-type wells can be constructed watertight at or below grade to minimize their visual impact on the shoreline. In large intake capacity applications, horizontal beach wells are typically coupled so that the intake pump station is installed above the well caisson. The well intake pump station can be designed with submersible pumps to minimize noise. However, for medium and larger size wells, most frequently vertical turbine pumps are used because these pumps usually have higher energy efficiency and require less power. A more detailed description of horizontal collector well design and construction is presented elsewhere (4).

Infiltration Galleries

Infiltration galleries are riverbank or seashore filtration systems that are typically implemented when conventional horizontal or vertical intake wells cannot be used because of unfavorable hydrogeologic conditions. For example, they are suitable for intakes where the permeability of the underground soil formation is relatively low, or for river or seashore filtration, where the thickness of the beach or the onshore sediments is insufficient to develop conventional intake wells.

Infiltration galleries consist of an excavation trench which is filled with filtration media of size and depth similar to that of the granular media filters used for conventional water treatment plants. Vertical or horizontal collector wells are installed at equal distances (usually 30 to 60 meters) inside the filter media. Typically the capacity of a single collection well is 0.009 to 0.09 m^3/s (0.2 to 2.0 MGD).

A common type of infiltration gallery is a horizontal well collection system that has a single trench (Fig. 3). The media in the wells are configured in three distinctive layers: a bottom layer of sand of approximately 2 to 3.5 meters (6 to 10 feet), followed by a 1.2 to 2 meter (4 to 6 feet) layer of graded gravel pack surrounding the horizontal well collector screens, topped by a (6 to 9 meter) (20- to 30-foot) layer of sand. Horizontal well collector screens are typically designed for an inflow velocity of 3 cm/s (0.1 ft/s) or less.

Infiltration galleries can be designed either similarly to conventional rapid sand filters (if the natural source water movement, such as river flow or ocean water wave motion, can provide adequate flushing of the infiltration gallery media contact surface with the waterbody) or can be constructed as slow sand filtration systems, which have at least a 9-meter (30-feet) layer of sand overlying the collection well screens. Infiltration galleries are usually 15 to 20% more costly to construct than conventional vertical or horizontal intake wells, and therefore, their use is warranted only when the hyrogeologic conditions of the intake site are not suitable for conventional intake wells.

Riverbed/Seabed Filtration Intake Systems

These intake systems consist of a submerged, slow sand medium filtration system (filtration bed) located in a riverbed or in the near-shore surf zone of the ocean floor (Fig. 4). The filtration bed is connected to a series of intake wells located on the shore via tunnels or horizontal collector pipes.

The filtration bed is sized and configured using design criteria similar to those for slow sand water treatment plant filters. The filter bed depth is typically between 0.9 m and 1.2 m (3 to 4 ft). The gravel supporting the filter bed's sand is 0.15 to 0.5 m (0.5 to 1.5 ft) deep. The design surface loading rate of the filter medium is typically between 0.05 and 0.2 m^3/m^2h (0.2 and 0.8 gpm/sq ft). Similar to slow sand filters, a mat of natural organic materials is formed on the surface of the filtration bed. This mat enhances removal of organics and fine particles from the source

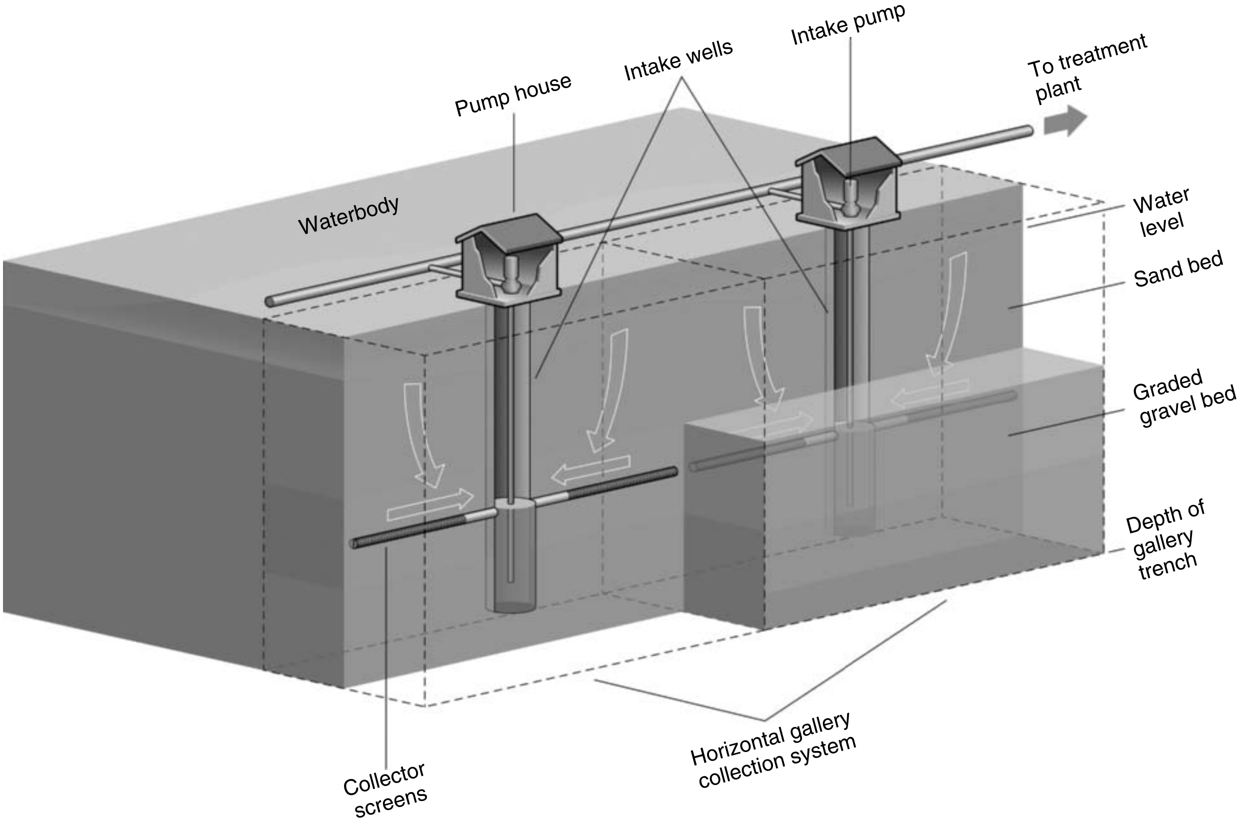

Figure 3. Infiltration gallery.

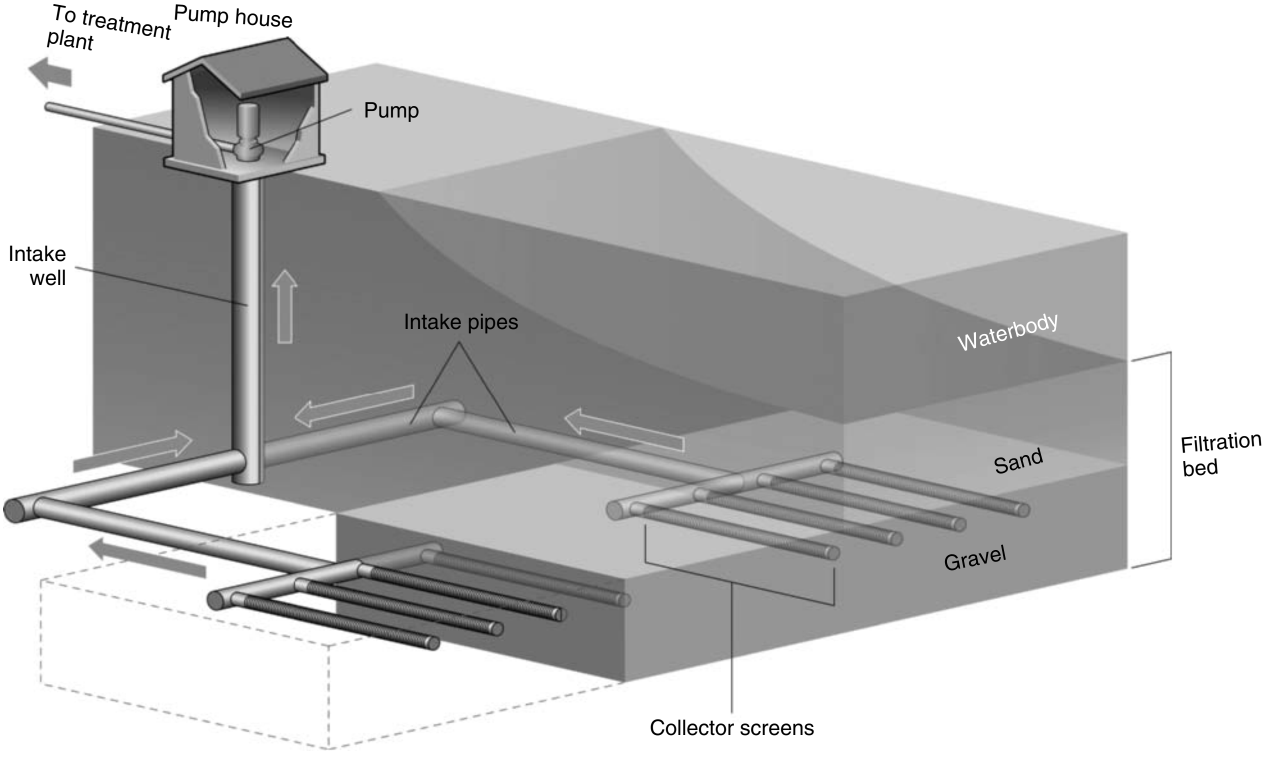

Figure 4. Riverbed/seabed filtration intake system.

water. The surface filtration mat is often removed from the surface of the filtration bed by naturally occurring seasonal scouring, such as river flooding or storms. When the mat is removed and some of the filtration bed sand is lost over time, the sand medium has to be replaced to its original depth to maintain filtration bed efficiency.

An alternative riverbed filtration system, which consists of a soft-soil tunnel constructed in the sand and gravel aquifer, whose well screens extend horizontally from the tunnel under the river, has been designed on the Ohio River for the Louisville Water Company in Kentucky (2). The riverbed filtration system consists of twenty-three 61-meter (200-ft) long laterals located at an equidistance of 79.3 meters (260 ft) in the riverbed, connected to the collection tunnel of a total length of 1830 meters (6000 ft). The laterals are 0.3 m in diameter, and the riverbed serves as a natural slow sand filtration system. The experience gained at Louisville indicates that the maximum water yield of this system is 0.22 to 0.31 m^3/s (5 to 7 MGD) per 305 m (1000 feet) of riverbank length.

The largest seawater desalination plant with a seabed intake system currently under construction is the 0.58 m^3/s (13.2 MGD) Fukuoka District RO facility in Japan. This plant is planned to be operational in late 2005. The Fukuoka seawater desalination plant seabed intake area is 30,000 m^2 (312,000 sq ft).

Typically, riverbed and seabed filtration intakes are the costliest well intake systems. Their construction costs are approximately 1.2 to 2.3 times higher than those of conventional intake wells (5).

BIBLIOGRAPHY

1. Missimer, T.M. (1999). Raw water quality—the critical design factor for brackish water reverse osmosis treatment facilities. *Desalination and Water Reuse* **9**(1): 41–47.
2. Chittaranjan, R., Melin, G., and Linsky, R. (Eds.). (2002). *Riverbank Filtration—Improving Source Water Quality*. Kluwer Academic, Dordrecht, the Netherlands, p. 24.
3. Williams, E.B. (1981). Fundamental concepts of well design. *Ground Water* **19**(5): 527–542.
4. Hunt, H.C. (1996). Filtered seawater supplies—naturally. *Desalination and Water Reuse* **6**(2): 32–37.
5. Wright, R.R. and Missimer, T.M. (1997). Alternative intake systems for seawater membrane water treatment plants. *Proc. Int. Desalination Assoc. Conf.*, Madrid, Vol. 3, pp. 407–422.

PHYSICAL PROPERTIES OF DNAPLs AND GROUNDWATER CONTAMINATION

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

The physical behavior of dense nonaqueous phase liquids (DNAPLs) can broadly be divided between two classes based on melting point. Polynuclear aromatic hydrocarbons (PAHs) have melting points that are at the lowest, 80 °C (for naphthalene). Consequently, PAHs are solids at the temperatures typical in the subsurface; thus, the potential for migration through the soil column or groundwater systems is limited. Cosolvents can make PAHs soluble and mobile; however, the PAH solutions (when dissolved by petroleum hydrocarbons) are typically lighter than water and do not behave as DNAPLs. The second class of DNAPLs are chlorinated solvents. These compounds are liquids at normal subsurface temperatures, making them extremely mobile in unsaturated and saturated zone soils. Here we discuss those and other properties that make contamination by chlorinated solvents one of the most ubiquitous and recalcitrant of groundwater problems. A good overview of chlorinated solvent characteristics and remediation is included in Pankow and Cherry (1).

The introduction of chlorinated solvents was perceived as a great innovation for safety and an improvement in the quality of the degreasers. Prior to that introduction, distilled mineral spirits were the only available solvents. Explosions and fires from mineral spirits were a constant threat and occurrence causing death and injury. In today's climate where risk is evaluated in terms of one event in a million cancer risk, there are numerous lives saved through the use of chlorinated solvents. One of the most common chlorinated solvents is tetrachloroethene, also called perchloroethene (PCE) or "perc." This solvent is mostly associated with dry cleaner leaks. Another commonly used industrial degreaser, trichloroethene (TCE), has been found at many industrial facilities.

The properties that make chlorinated solvents so problematic with regard to groundwater contamination include density, solubility, viscosity, surface tension, and dehydration.

Chlorinated solvents are significantly denser than water. The density values in g/cm^3 (water is 1.00) are:

1,1-Dichloroethane (1,1-DCA)	1.18
1,1-Dichloroethene (1,1-DCE)	1.22
Trichloroethane (TCA)	1.34
Trichloroethene (TCE)	1.46
Tetrachloroethene or perchloroethylene (PCE)	1.62

Upon release into the environment, DNAPLs sink through the vadose zone, through the water table, and, over the long-term, through aquitards. A release of chlorinated solvents will travel downward or laterally until it has been exposed to a volume of soil sufficient to retain the released DNAPLs in pore spaces within the soil matrix.

Due to high density, chlorinated solvents can flow upgradient to the groundwater flow along the down dip surface of an aquitard. The most problematic phenomenon caused by high density chlorinated solvents is the dimple effect. Upon release, chlorinated solvents will travel to an aquitard and seek the lowest point on the surface of that aquitard. It is unlikely that any given stratigraphic unit acting as an aquitard is perfectly planar; it will have small to large dips, depressions, and dimples. Each of these can act as a reservoir for DNAPLs. These reservoirs will be outside an advective flow regime set up by a "pump and treat" groundwater system. The only means of transport from these reservoirs is through diffusion, which is effective enough to cause significant dissolved phase contamination, but not effective enough to offer timely remediation.

Chlorinated solvents are soluble in water. Values in mg/L are:

1,1-DCA	5100
1,1-DCE	400
TCA	700
TCE	1100
PCE	200

These levels of solubility represent significant potential for groundwater transport of released chlorinated solvents in the dissolved phase.

Most chlorinated solvents are less viscous than water. Values in centiposes (water is 1.00) are:

1,1-DCA	0.51
1,1-DCE	0.36
TCA	0.90
TCE	0.57
PCE	0.93

The low viscosity makes chlorinated solvents extremely mobile in the vadose zone; they can readily flow through flow channels that may be as small as a human hair.

The surface tension of chlorinated solvents with respect to water has significant impact on the manner in which chlorinated solvents penetrate the capillary fringe zone and then migrate through the saturated zone. Water has a high surface tension of 73 dyn/cm; chlorinated solvents have surface tensions in the range of 20–40 dyn/cm. When free phase chlorinated solvents migrate to the capillary fringe, they will be held up until a head is built up sufficient to overcome the capillary retention of the water. The capillary retention is directly proportional to the interfacial tension and inversely proportional to the pore throat radius and liquid density. The smaller the pore size, the greater the head required to displace entrained water. This can result in significant lateral spreading of a DNAPLs release above the capillary fringe in fine grained soils.

Another effect of surface tension forces involves how surfaces in larger pore spaces are wetted. The tendency of one fluid to replace another on a surface is termed wettability.

As broad rules, the following apply:

In soil, water is the wettting fluid with respect to solvents or air.

Solvents are wetting fluids in air, but not in the presence of water.

With respect to carbonaceous soil components, solvents are wetting in the presence of air or water.

Free phase chlorinated solvents have the capacity to dehydrate clays, causing cracking and further migration through what at first analysis appears to be impermeable layers. Bentonite pellets used to seal wells in areas contaminated with DNAPLs may not work, due the failure of the bentonite to swell in those conditions. DNAPLs can also migrate through previously installed and sealed wells via the same mechanism.

All of the above phenomena impact the fate and transport of DNAPLs to and through groundwater systems. Those physical properties also manifest themselves by the nature of a given release. The release of a large volume of DNAPLs in a short period of time causes rapid migration laterally as well as downward through the vadose and saturated zone that overcomes any small permeability differences. This leaves a significant volume of residual contamination entrained in the zone of passage. Conversely, DNAPLs release that occurs slowly over a long period of time will travel through narrow channels representing the most permeable path through the subsurface. Small differences in permeability will be exploited under these conditions. Under slow release conditions, less of the DNAPLs is entrained within the soil matrix: the overall potential for vertical migration is greater, more DNAPLs will penetrate to a greater depth, and the potential to impact groundwater is also greater.

Subsurface “dimples”—slight depressions or irregularities at the top of an underlying aquitard where DNAPLs may collect—have been observed. These features are difficult to locate.

The overall effect of the physical properties described above, particularly where the “dimple” effect is manifest, may make timely remediation of groundwater contaminated with DNAPLs impossible. In those instances, the most economic method of protecting health will lie in municipal or point of use treatment systems for the recovered groundwater.

BIBLIOGRAPHY

1. Pankow, J.F. and Cherry, J.A. (1996). Dense chlorinated solvents, and other DNAPLs. In: *Groundwater: History, Behavior, and Remediation*. Waterloo Press, Guelph, Ontario, Canada, p. 522.

WATER DOWSING (WITCHING)

David P. Commander
Water and Rivers Commission
East Perth, Australia

‘Water dowsing’ is the supposed ability to locate underground water using a simple handheld device such as a forked stick, bent wires, or a plumb bob. ‘Dowsing’ is the general term applied to ‘divining’ for objects or materials concealed beneath the ground. Divining, or divination, implies a supernatural intervention and has been practiced for at least 2000 years for a variety of purposes. Water dowsing is also known as water witching.

To dowse for water, the ‘diviner’ walks along the ground holding the device, which then ‘responds’ to the presence of an ‘underground stream.’ The traditional European divining rod is a forked hazel twig, which is grasped in both hands and bent under tension; the ends held horizontal and point, forward. Only a slight involuntary movement on the part of the operator is then needed to make the twig jerk downward. Other such methods in use are two pieces of bent wire, one in each hand, which then cross to indicate the 'underground stream,' or a plumb bob, which begins to rotate.

The concept relies on the hypothesis that groundwater occurs in streams, similar to surface water. A diviner can then trace the path of an 'underground stream' from the surface. He may pick up several streams and indicate where they cross. A successful bore or well must be sunk exactly on the stream, or failure will be justified by having missed the 'stream.' Some diviners also claim to be able to detect the salinity of the underground water, or the depth of the water table, water-bearing strata or fracture, and the likely well yield.

There is vast literature on dowsing, and modern texts can be found in 'New Age' bookshops. Water witching is carried out in all hydrogeologic environments, ranging from extensive shallow aquifers, where success is assured, to fractured rock terrains in which success rates are poor. However, scientifically constructed tests always produce statistically random results, often to the chagrin of the diviner, who may have a real belief in his ability to dowse for underground water. Water witchers often have experience and local knowledge that aids them in selecting of favorable areas before actually applying the technique to locate an exact drilling site.

Dowsing with a forked twig was first mentioned in 1556 by Georgius Agricola in his descriptions of mining practices in the Harz Mountains of central Germany where the diving rod was being used supposedly to locate mineral veins (Fig. 1). Agricola mentions that the material for the twig was selected for different minerals, for instance, hazel for silver and ash for copper. Agricola is scathing about the practice, referring to dowsers as cunning manipulators and to the faith among common miners 1.

The practice of water dowsing spread through Europe in the seventeenth century, and then to North America and Australasia, where it is a widely held belief among farmers. The first mention of divining applied to water was apparently to locate a well for a convent in Spain (from a life of Saint Theresa, in Ref. 2, p. 171) in 1568. However, Arthur J. Ellis (3), who comprehensively reviewed the history of water witching, suggested that the practice gained wider credence from the activities of Baron and Baroness de Beausoleil in the early 1600s. The Beausoleils traveled through mining areas, even visiting

Figure 1. Use of divining rods in prospecting for ore 1. (From Agricola, De re metallica, German edition, 1580).

America, and the Baroness apparently used divining to discover the Chateau Thierry mineral water in France. After this, the divining rod seems to have been exclusively applied to locating groundwater, no doubt because of the interest in mineral water at the time and the high success rate of water wells in northern Europe. Connection with witchcraft led to the name water witching, which is used exclusively in North America.

BIBLIOGRAPHY

1. Agricola, G. (1912). *1556, De re metallica libri XII*, 1st Ed. Frobenius, Basel, p. 538 (English translation by H.C. Hoover and L.H. Hoover, The Mining Magazine, London), p. 640.
2. Barrett, W.F. (1911). *The So-Called Divining or Dowsing Rod. Psychical Research*. Home Library series, ch 12, pp. 167–186.
3. Ellis, A.J. (1917). *The Divining Rod—A History of Water Witching*. United States Geological Survey, Water-Supply Paper 416.

SUBSURFACE DRAINAGE

Harendra S. Chauhan
G.B. Pant University of
Agriculture and Technology
Uttar Pradesh, India

The basic reason for agricultural lands to be affected by waterlogging and salinity is the inadequacy of the natural drainage system to handle the water that reaches the land either by natural or artificial means. Under such situations, an artificial drainage system has to be provided. The meaning of land drainage varies in different regions and contexts. A geographer or a surface hydrologist may use it to mean the pattern of natural watercourses in a hilly area. A pedologist may think in terms of the permeability of a poorly drained or well-drained soil. In Peru, engineers interpret drainage as reconstruction of natural waterways rather than removal of water. In Canada, it may mean reclamation of marshland for agricultural or urban development. In Holland, a Dutch farmer may use it for the installation of drainpipe in the soil. Thus, in different areas, this term implies different meanings and objectives. If the land use is primarily agricultural then land drainage could be defined as the establishment and operation of a system by which the flow of water from the soil is enhanced, so that agriculture can benefit from the subsequently reduced or controlled water level in the soil. Agricultural land drainage aims at reclaiming and conserving land for agriculture, to increase crop yields, to permit cultivating more than one crop in an area, and to reduce the cost of production

INTRODUCTION

There are no hard and fast lines between the objectives of different kinds of agricultural drainage, but broadly there are two types of drainage, surface drainage and subsurface drainage. Surface drainage is the removal of excess water from the land surface to create more favorable conditions for plant growth. Subsurface drainage is the removal or control of groundwater and/or removal or control of salts using water as the vehicle. The excess water may be from precipitation, water applied in irrigation, losses from conveyance channels and storage systems, and water that has seeped from groundwater in upper reaches. It is the drainage of excess water from the land surface at a rate that will prevent long periods of ponding without excessive erosion, so that crops would have favorable moisture conditions. Any drain or well-designed system to control or lower the groundwater is considered subsurface drainage. It can be broadly classified in two categories as (1) horizontal drainage and (2) vertical drainage. The other category of drainage frequently referred to nowadays is biodrainage. The objective of this article is to discuss various aspects of subsurface drainage. The main focus will be on horizontal drainage hereafter referred to as subsurface drainage. Vertical drainage and biodrainage will also be discussed briefly later.

HISTORICAL PERSPECTIVE

Agricultural drainage to improve crop production is an age-old concept that has been used with profitable results for quite some time in the past with different orientations and variations. Even though agricultural drainage is discussed in ancient Greek history, its development in different parts of the world is not very well known. In Greek civilization, some areas were reclaimed by systems of ditches. Drainage of humid and semi-humid areas started in the tenth century and extended to different parts of Europe in the sixteenth and seventeenth centuries. Herodotus has referred to the oldest instance of drainage in 400 B.C. by the Egyptians in the Nile Valley. In the United Kingdom, the earliest work is traced to the thirteenth century when Romney marsh was reclaimed, but the major work was carried out about 250 years before. In the United States, the first tile drains were laid by John Johnson of New York for a total length of 56 km. Due to the import of tiling machines in 1848 more than 100 tile factories existed in Illinois, Indiana, and Ohio. By 1882, about 43,230 km of tile drains were installed in Indiana alone. The Bureau of Reclamation had constructed about 4990 km of drains in 1931. By this time, subsurface drains had been installed extensively in most of the states in the United States.

In India, the problem of waterlogging and soil salinity were observed in village Munak in the former Punjab as early as 1855. Due to the construction of large canal systems, salinity was observed in several areas. In the Ganges canal, salinity was observed in 1876, and in lower Chenab canals, it was observed after these canals opened in 1892. The first subsurface drainage experiment was laid out in India in 1873, although it was not considered successful. In Punjab, a Waterlogging Board was constituted in 1925. In 1928, Inglis and Gokhale established that subsurface drainage is effective in lowering the water table at the Baramati farm.

The earliest drainage method consisted of making drains from stones. With the use of stones, the old method of drainage by open trenches was abandoned. Another method of draining permanent pastures was through turf drains the remains of which can still be seen in many fields in the mid lands in the United Kingdom. Some of these even had useful lives of 100 years. Besides turf drains, other varieties of drains were plug drains principally used in arable lands, straw drains used extensively in the heavy land districts, and brushwood drains made from fagots or bundles of brushwood, which were placed at in the bottom of deep trenches. Then, came brick drains followed by U-tiles or horseshoe drains. The last to come were cylindrical brick tiles or clay tiles, which have now been replaced by PVC perforated/corrugated pipes.

HORIZONTAL SUBSURFACE DRAINAGE

Buried pipes, pipeless drains, or deep open ditches are commonly used to drain lands in this kind of drainage. Pipe drains consist of a system of tiles or pipes such that excess water enters in lateral lines through the joints between two tiles or perforations in the pipes and flows towards collector/main drains. Pipeless drains, called mole drains, are cylindrical channels artificially made in the subsoil by a mole plow. When the cost of regular pipe drains becomes too high due to close spacing, mole drains are used. The principles and hydraulics of mole drains are similar to those of pipe drains. It is a temporary method of drainage because moles deteriorate in a few years and have to be remade frequently. Deep open ditches, if constructed close enough, can work similarly to pipe drains or mole drains to lower the water table and remove excess water. Practically, open ditches are difficult to maintain and cause a lot of inconvenience to men, animals, and farm equipment that cross them.

SUBSURFACE DRAINAGE THEORIES

A number of mathematical solutions have been proposed to describe water table behavior and to evolve drainage criteria for appropriate design of drainage systems, using steady-state or transient flow conditions in horizontal or sloping aquifers. The solutions have been obtained by employing experimental, analytical, or numerical approaches. These solutions help to evolve drainage criteria that help in determining the drainage rate and spacing for a preselected suitable depth of drains, depending on soil characteristics and water table height midway between drains. The basic objective of evolving drainage criteria is to attain and maintain an optimum or a reasonably permissible value of average water table depth and root zone that is either free from salts or has electrical conductivity below the tolerance limit of the plants to get the maximum net benefit. Some of the theories are presented here.

Steady State

In this situation, recharge due to rainfall or irrigation is assumed constant and equal to the discharge from drains, such that water table does not change with time. This criterion describes the quantity of water to be removed in a given period of time under conditions of given water table height above the drain level. This criterion is generally applicable in flat areas of a humid region, where low intensity, long duration rainfall occurs in winter and spring. By selecting appropriate soil water hydrologic parameters, it can also be used for arid salt-affected lands. Based on this theory, Colding (1), Rothe (2), Kozeny (3), and Hooghoudt (4) obtained the same relationship independently. Hooghoudt (5), based on the Dupuit–Forchheimer assumptions, proposed an equation for computing of drain spacing as

$$L^2 = \frac{4Kh_0(2d + h_0)}{q} \tag{1}$$

where L is the drain spacing, K is the hydraulic conductivity of the soil, d is the depth to an impermeable layer below the drain axis, q is the rate of recharge, and h_0 is the height of the water table above the drain axis at the midpoint between the drains. It is more appropriately applicable only when the depth to impermeable layer d is relatively small compared to h. For comparatively large values of d, the flow to drains will tend to be radial. To account for the extra resistance caused by radial flow, Hooghoudt introduced the concept of equivalent depth (d_e), such that d is replaced by d_e. In such cases, Equation 1 may be written as

$$L^2 = \frac{8K_b d_e h_0 + 4K_a h_0^2}{q} \tag{2}$$

In this equation, K_a and K_b are the hydraulic conductivities of the soil above and below the drain, levels. The first term pertains to flow below the drains and the second term indicates flow above the drains.

Ernst (6) extended Hooghoudt's approach to make it applicable to stratified soils that have a slowly permeable top layer. He assumed horizontal and radial flow resistance in his analysis. His solution is simple, but the presence of the geometry factor in the solution makes it slightly cumbersome for practical applicability. Nomographs developed by Van Beers (7) made Ernst's solution simpler for practical purposes.

Kirkham (8) obtained an analytical solution for a problem similar to Hooghoudt's, using potential theory and ignoring the flow above the drains. Later on Kirkham (9) reported that, if vertical flow in the upper region is also taken into account, the solution becomes

$$h = \frac{qL}{K}\left(\frac{1}{1 - q/K}\right)F_k \tag{3}$$

where

$$F_k = \frac{(L - d\sqrt{2})^2}{8dL} + \frac{1}{\pi}\ln\left(\frac{d}{r\sqrt{2}}\right) + f(d, L)$$

r is the radius of the subsurface drain, and $f(d, L)$ is a function of d and L, which was evaluated from an infinite series, which introduced a reduction in flow depth. The new reduced depth, d_e, was called equivalent depth and was defined as $L/8F_k$.

Wessling (10) analyzed the solution given by Hooghoudt (5) and Kirkham (9) and concluded that both solutions gave almost identical results within a 5% variation. Lovell and Youngs (11) calculated water table heights above drains using a number of steady-state equations and found that it could be computed with sufficient accuracy for smaller as well as for larger depths of the impermeable barrier. At intermediate depths, the solutions led to large uncertainties in the estimates. It was also noticed that among all drainage equations, Hooghoudt's equivalent depth equation, when used with an optimum drain radius obtained by a hodographic analysis for finite soil depths, gave reasonable values of lateral drain spacing which were within the known limits of uncertainty.

Unsteady State

This theory is generally used in areas that have a high intensity of rainfall or in arid lands where the rise in the water table due to frequent irrigation needs to be lowered by drainage. It describes the fall of the water table required within a certain period after the water table has risen near the land surface due to instantaneous recharge. Most of the unsteady-state theories relate to the study of groundwater flow in a homogeneous, isotropic, unconfined aquifer lying over a horizontal impermeable layer and are characterized by Dupuit–Forchheimer assumptions based on the Boussinesq (12) partial differential equation, which is

$$h\frac{\partial^2 h}{\partial x^2} + \left(\frac{\partial h}{\partial x}\right)^2 = \frac{f}{K}\frac{\partial h}{\partial t} \tag{4}$$

where K is the hydraulic conductivity of the aquifer, h is the height of the phreatic surface above the impermeable layer, and f is the drainable porosity (dimensionless). Equation 4 is a nonlinear second-order partial differential equation. Solutions of nonlinear or linearized forms of this equation for appropriate boundary conditions representing different physical flow conditions have been obtained in the past by a number of researchers. There are three physically identified commonly occurring flow situations: (1) a falling water table between two level drains, (2) a falling water table between two bilevel drains, and (3) a water table between two drains/canals fluctuating due to recharge or evapotranspiration.

Falling Water Table between Two Level Drains. Boussinesq (13) obtained an exact analytical solution of Eq. 4 with initial conditions $h(L/2, 0) = h_0$ and boundary conditions $h(0, t) = h(L, t) = 0$, assuming drains located at the impervious layer. His solution, which describes the temporal and spatial variation of a falling water table in nondimensional form, can be written as:

$$\frac{h}{h_0} = \frac{1}{\left(\dfrac{9Kh_0 t}{fL^2} - 1\right)} \tag{5}$$

where h_0 is the initial midpoint of the water table height, h is the water table height after lowering it from an initial value of h_0 in time t, and L is the spacing between two drains.

Glover (14) also obtained the same solution for Eq. 4 as given by Boussinesq (13). Dumm (14) assumed an initially horizontal groundwater table at a certain height above the drain level because of instantaneous recharge due to rainfall or irrigation. Equation 4 was linearized by assuming that the free water surface has a small slope enabling h, associated with $(\partial^2 h/\partial x^2)$, to be replaced by D, average depth of flow, and neglecting $(\partial h/\partial x)^2$ by considering that it is very small. Equation 4 is thus reduced to

$$\frac{\partial^2 h}{\partial x^2} = \frac{f}{KD}\frac{\partial h}{\partial t} \tag{6}$$

Assuming that the initial water table is horizontal and zero end conditions, a Fourier series solution for falling water table heights midway between the drains was obtained as

$$h\left(\frac{L}{2}, t\right) = \frac{4h_0}{\pi}\sum_{n=1,-3,5}^{\infty}\frac{1}{n}e^{-\frac{n^2\pi^2 KDt}{fL^2}} \tag{7}$$

Dumm (15) assumed that the initial water table was a fourth-degree parabola and obtained the solution for the boundary conditions used by Glover (14). The only difference between Eq. 8 and his solution was the shape factor which was replaced by 1.16 instead of a value of $4/\pi$.

Brooks (16) presented a solution to the nonlinear Boussinesq equation describing unsteady flow toward equally spaced drains above a horizontal impermeable barrier in the form of a perturbation series. Van Schilfgaarde (17, 18) derived an exact solution to Boussinesq's Equation 4 for the boundary conditions when drains did not reach the impermeable layer. He also applied a convergence correction to the solution of the drainage problem and finally obtained an equation for drain spacing as

$$L = 3A\left[\frac{K(d_e + h)(d_e + h_m)t}{2f(h_m - h)}\right]^{\frac{1}{2}} \tag{8}$$

where

$$A = \left[1 - \left(\frac{d_e}{d_e + h_m}\right)^2\right]^{\frac{1}{2}}$$

d_e is the equivalent depth of an impervious layer below the drain, and h_m is the maximum water table height above the drain level. The results of this solution compared with those obtained from the Glover–Dumm (14) solution differed by the order of 10 to 40%.

Role of Evapotranspiration in Drainage Design

Work was also carried out by Brooks (16), Moody (19), Pandey and Gupta (20), Singh et al. (21, 22), and Nikam et al. (23) to include the contribution of other processes, mainly evapotranspiration, in lowering the water table. Both steady and nonsteady approaches have been used.

Falling Water Table Between Two Bilevel Drains. De Boer and Chu (24) developed an unsteady-state theory for a bilevel drainage system using the Bouwer and Van Schilfgaarde (25) approach. They suggested using a bilevel drainage system because by having deep and shallow drain lines on an alternate basis, the excavation and digging cost would be reduced substantially and thus a bilevel drainage system would be economical compared to a conventional level drainage system. Studies were also carried out by Chu and De Boer (26), Sabti (27), Verma et al. (28), and Upadhyaya and Chauhan (29).

Fluctuating Water Table. Kraijenhoff van de Leur (30) studied the effect of constant steady-state recharge during the drainage process instead of an instantaneous recharge, as assumed by Dumm. The governing equation and initial and boundary conditions are written as

$$\frac{\partial^2 h}{\partial x^2} + \frac{R}{Kd} = \frac{f}{KD}\frac{\partial h}{\partial t}$$

$$h(x,0) = 0, \qquad 0 < x < L$$

$$h(0,t) = h(L,t) = 0, \quad t > 0$$

where R is the recharge rate per unit area. The height of the water table midway between parallel drains at any time t is given as

$$h(L/2,t) = h_t = \frac{4}{\pi}\frac{R}{f}J\sum_{n=1,3,5}^{\infty}\frac{1}{n^3}(1 - e^{-\frac{n^2 t}{J}})$$

where $J = fL^2/\pi^2 KD$.

This equation is not used for routine drain spacing computations; however, it is useful for hydrologic studies such as estimating water table fluctuation or drain discharge variation with time for a given recharge pattern.

Maasland (31) also analyzed the problem of water table fluctuations in response to constant recharge, intermittent constant recharge, and intermittent instantaneous recharge independently and almost at the same time as Kraijenhoff van de Leur (1958) using similar and a few different boundary conditions.

NUMERICAL SOLUTIONS

Dass and Morel-Seytoux (32) solved the one-dimensional nonlinear Boussinesq equation by Galerkin's finite element technique for three initial water table conditions. They observed that Galerkin's method was applicable to any initial water table profile. Studies were also carried out by Marino (33), Singh and Jacob (34), Singh and Rai (35), Skaggs (36), Pandey et al. (37,38), and Sewa Ram et al. (39).

Drainage of Sloping Lands

When wetness is prolonged, subsurface drainage and hillside seeps are an acute problem in many parts of the world such as the pre-Alps in Switzerland and adjacent countries and the tea gardens of Assam, India, where drainage of sloping lands is considered quite an important problem.

Because of saturation of the soil, the growth of plantations and trees is retarded, and there is difficulty in new afforestations and plantations. In the southern Appalachian Mountains, such problems were experienced for which elaborate experimental studies were carried out to study storm flow response on sloping forested watersheds. The theoretical analysis of this problem has also been difficult and is a classical hydrologic issue. Some notable works may be mentioned as those of Werner (40), Schmid and Luthin (41), Wooding and Chapman (42), Luthin and Guitjens (43), Chauhan et al. (44), Childs (45), Jaiswal and Chauhan (46), Sewa Ram and Chauhan (47), and Upadhyaya and Chauhan (48).

DESIGN CRITERIA FOR HORIZONTAL DRAINS

Steady-State Criteria

This criterion is generally applicable in flat areas of humid regions, where low intensity, long duration rainfall occurs in winter and spring. In this situation, recharge due to rainfall or irrigation is assumed constant and equal to the discharge from drains such that the water table height does not change with time. This criterion describes the quantity of water to be removed in a given period of time under conditions of given water table heights above the drain level. In this criterion, drain discharge is matched to provide a clear root-zone depth above the midpoint of the water table. Such a criterion can be expressed as for example, a drain discharge of 7 mm per day when the water table depth is at the midpoint of two drains at 50 cm below the ground level. Hooghoudt's equation has generally been recommended for designing drain spacing in homogeneous soils because of its simplicity and ease of use. For layered soil or for soil that has a slowly permeable top layer, the Ernst equation (6) has been recommended because it takes radial flow into account in the total flow.

Unsteady-State Criteria

This criterion is generally used in areas of high intensity storms that occur quite commonly or in irrigated lands which requires establishing a rate of fall of the midpoint of the water table between two horizontal drains to match the sensitivity of a selected crop that can sustain a high water table for a limited period. The Glover–Dumm equation is recommended for use by the United States Bureau of Reclamation. According to this criterion, the water table should drop from the ground surface to at least 30 cm below in 24 hrs and to 50 cm below in 48 hrs (49). When the water table rises to some 15 cm below the ground surface, it should drop to 35–40 cm below in 1 day (50).

Equivalence Between Steady-state and Nonsteady State Criteria. Steady-state criteria are comparable to the falling water table criteria. For the drainage criterion consisting of a discharge rate of 0.007 m/day, when the water table depth is 0.50 m, the hydraulic head (h) will be 50 cm if the drain depth is 1 m. If the water table rises to the surface, both the discharge rates and hydraulic head will increase by two times and become 0.014 m/day and 100 cm, respectively. For a clay soil whose drainable porosity is 4 to 5%, this would result in a drop of the water table from the ground surface to about 30 cm below in 1 day which corresponds to the U.S. standard.

Vertical Drainage

Vertical drainage and conjunctive use of canals and groundwater is another good way of lowering the water table and providing subsurface drainage. With good quality groundwater, it is an excellent practice. On one hand, it lowers the water table and, on the other hand, through the same process, it also provides irrigation for increasing the production of agricultural crops. However, it is important to understand the different issues involved. Sometimes, the objective of vertical drainage is steady-state lowering of water tables in the root zone for growing crops due to overlapping of drawdown caused by pumping closely spaced wells. In other words, it is similar to the well-known well point system. Such a system is used for building foundations or other structures in high water table areas where rapid lowering of water tables is required for short periods. Such a system would not only be structurally costly but would also involve high recurring energy costs. Such a system does not seem to be practicable for growing agricultural crops. Thus, lowering of water tables should be considered in terms of total withdrawal on a long-term basis instead of transient lowering in a short period for growing crops that require lowering of 0.3–0.5 m in 1 to 2 days, as for maize or soybean.

De Ridder (51) and Attia and Twinhof (52) recommend the technical feasibility of tube well drainage only based on the geometry of the aquifer, hydraulic parameters, the thickness and hydraulic resistance of the clay cap, the groundwater quality, and the rate of recharge. According to them, well drainage enables lowering the groundwater to a much greater depth than gravity drainage. Where deeper layers of substrata are more pervious than layers near the surface, pumping from these layers may reduce the artesian pressure that is often present, and create a vertical downward flow through the upper layers. If the pervious strata are found at a depth of 5 m or more, well drainage can be applied successfully, if the transmissivity of the aquifer is fairly high. If the aquifer is semiconfined, an additional criterion would be the value of the hydraulic resistance of the upper clay layer. The energy requirement is another consideration that should be looked into for well drainage. Unlike gravity drainage, well drainage is not economically feasible in small areas because too large a portion of the water drained consists of foreign water. Well drainage may also not be feasible in areas where artesian pressure is too high.

In a case study of SCARPS in Pakistan by Awan (53), tube wells were installed to lower the water table and provide more water for intensive agriculture and to reduce salinity. Large capacity tube wells were installed near the outlets to feed one or two watercourses. A large number of problems were faced from vertical drainage. The operational and maintenance costs were quite heavy. There were problems of spares making tube wells remain idle. Eighty percent of the total budget went to electricity bills leaving little money for maintenance. Several disputes cropped up among the farmers who were supplied water from the same watercourse. The authorities were forced to shift the policy in the sweet water zone from large public tube wells to the private sector. In a later study, Smedema and Zimmer (54) also favored private investments in the freshwater zone, which seemed quite successful. It was also observed that providing horizontal drainage instead of tube well drainage could have reduced the salinity of the drainage effluent.

Even in the Australian Wakool drainage project of about 57,000 ha, affected lands had water tables within 2 to 3 m of the surface. An attempt to lower the water table by using 52 tube wells faced similar problems, as in SCARPS in Pakistan. The high capital and growing operating and maintenance costs affected the economic value of the scheme, and it seemed unlikely that a scheme of such a magnitude could ever be built again for Australian agriculture.

Biodrainage

Biodrainage is being propagated as a method of subsurface drainage. It is an economical method of drainage, it improves the ecology of the area, and it is environmentally friendly. It provides costly wood and biomass useful for many purposes. It can transpire water from the groundwater table in good amounts. A large range of plants tolerant to salinity can be grown in waterlogged, salt-affected lands. However, it is important to identify appropriate situations where this method can be effective and useful. It is unfair to assign it an objective, which it cannot achieve, and then misinterpret the approach. The issue to be examined is whether meeting the shortage of food production is a pressing need now or in the future in developing countries and, if so, how can biodrainage be used to reclaim waterlogged and salt-affected lands for food production. If there is no immediate necessity to increase food production, then it could be a good alternative to put all such areas under trees. However, if the priority favors food production, then it has to be seen whether forest or tree crops can lower water tables at a fast enough rate to grow maize/soybean or other crops without the ill effects of waterlogging. Even if plants can lower the water table at a fast rate, one would have to look for a proper cropping system. It needs to be seen whether one can grow food crops simultaneously with tree crops and maintain a deep enough water table. Alternatively, one will be required to grow tree crops for a few years, and, after the water table has been lowered, then one can grow grain crops; this has to be properly understood and practically planned. Similarly, it is understandable that many agricultural, horticultural, and other tree crops can tolerate salinity of different levels. But, whether it is practicable to reclaim a saline soil whose electrical conductivity is more than 50 dS/m, through tree crops, so as to make it suitable for growing agricultural crops, needs to be explored.

According to Smedema (55) there are three situations for biodrainage: waterlogged landscape depressions, areas along a canal to intercept canal seepage, and growing plants to act as parallel field drainage. In the first, instead of installing a pumped well for drawing down the water table, the drainage improvement could be done by planting trees, shrubs, or other vegetation. In the second, in place of an interceptor drain practised in the past at the toe of the embankment of a canal, drainage conditions could be improved by planting a strip of suitable vegetation. In the third, biological line sinks (strips of planted trees/shrubs) can be used instead of conventional parallel field drainage systems such as ditches or pipe drains. Such a system could also work as a typical windbreak for a more favorable microclimatic environment. According to Smedema (55) this is generally not a realistic option, when common crops are facing severely waterlogged conditions. However, biodrainage under certain conditions has merit according to which it may be confined to growing tolerant species under not too extreme waterlogging conditions. The scope of biodrainage seems to be more favorable in an arid zone where drainage surpluses are small in relation to evapotranspiration rates (1 to 2 mm/day vs. 10 mm/day). The claims of using biodrainage to cope with salinity do not seem to be supported by experimental evidence from Australia and California. Biodrainage without any physical drainage is more likely in the long run to lead to a harmful buildup of root-zone salinity.

PRESENT STATUS AND SCOPE OF DRAINAGE

Land drainage in the past has played an important role in the modernization of rain-fed and irrigated agriculture throughout the world. Drainage of European agricultural lands, where improved drainage was needed in as much as 25% of agricultural land, helped Europe to attain self-sufficiency in food production. Nearly 70% of this land has been put under subsurface drainage during the last 30 years. Drainage development in the United States, helped by other technologies, created one of the most efficient production systems. As half of wetlands were converted into cropped land, drainage developments since 1960 are being influenced by changing environmental ethics because there is more emphasis on protecting wetlands. In spite of this, even in the United States, about 10 to 20 mha of cropland still need drainage. Out of a total agricultural area of 18.65 mha in the United Kingdom, besides presently drained lands, a further 3.5 mha still require drainage. Considerable experience in design and construction of drainage systems has been gained in Russia. The total area of drained land exceeds 1 mha. More than half of the land is under subsurface drainage.

According to Smedema, the world's total cropped land area stands at 1450–1500 mha of which 250 mha (70%) is provided with irrigation facilities. Of this cropland, 150–200 mha (10–15%) have drainage facilities. Of this about a 50 mha area is irrigated, and the remaining 100–150 mha are rain fed. All of the latter are in Europe and North America. It is estimated that 250–300 mha of rain-fed cropland still needs drainage of which 25–30 mha would be a reasonable medium-term program to cover the humid tropic lands of Southeast Asia. The current rate of drainage development is estimated at 0.5–1.09 mha/year.

Drainage, it is predicted, will increase food production in the irrigated area by 1–1.5% and in the rain-fed area by 0.5–1%. The global average will be about 1%. The need for drainage of irrigated lands consists of surface drainage and subsurface drainage to combat waterlogging and salinity. A medium-term (25 years) drainage plan would cover 10–15 mha surface drainage and 2–3 mha of subsurface drainage. All of this is located in developing countries.

BIBLIOGRAPHY

1. Colding, L.A. (1872). Om louene for vandets be voegelse I jorden. Danske vidensk. Selsk. Skr. 5 *Raekke Natur Vog math. Afd.* **9B**: 563–621 (Copenhagen).
2. Rothe, J. (1924). Diestrangent fernung bei Dranungen Landw. *Jahrb*. **59**: 453–490.
3. Kozeny, J. (1932). Hydrologische grandlagaen des Dranuer suches. *Trans. 6th Comm. International Society of Soil Science.* **A**: 42–67.
4. Hooghoudt, S.B. (1936). Bepaling van den doorlaatfactor van den grond met behulp van pompproeven (z.g. boorgatenmethode). *Verslag Landbouwk. Onderzolk*. **42**: 449–541.
5. Hooghoudt, S.B. (1940). Bijdragen tet de Bennis van eenige natuurheundige grootheden van den grond, 7, Algemeene beschsuwing van het probleem. Van de detail ontwatering en de infiltratic door middle van parallel loopende drains, greppels slooten, en kanalen. vaersl. *Landiouwk. Ond*. **46**: 515–707.
6. Ernst, L.F. (1956). Calculation of steady flow of ground water in vertical cross section. *Netherlands Journal of Agriculture Science* **9**: 126–131.
7. Van Beers, W.F.J. (1965). Some nomographs for the calculation of drain spacings. *Int. Inst. Land Reclaim. Improve. Bull.* No. 8, Wageningen, p. 45.
8. Kirkham, D. (1958). Seepage of steady rainfall through soil into drains. *Transactions American Geographical Union* **39**(5): 892–908.
9. Kirkham, D. (1960). An upper limit for the height of the water table in drainage design formula. *Trans. 7th International Congress for Science* **I**: 486–492.
10. Wessling, J. (1964). A comparison of the steady-state drain spacing formulas of Hooghoudt and Kirkham in connection with design practices. *Int. Land Water Manage. Res. Tech. Bull.* 34, Wageningen.
11. Lovell, G.J. and Youngs, E.G. (1984). A comparison of steady state land drainage equation. *Agricultural Water Management* **9**: 1–21.
12. Boussinesq, J. (1877). Du movement non permanent des eaux souterraines. *Essai Sur la theorie des eaux courantes*, Memoires presents par divers savants a l'. Academie des Sciences de institut de France, 23, 1, Paris, France, p. 680 (Note en Basde 252–260).
13. Boussinesq, J. (1904). Recherches theoriques sur l'ecoulement des nappes d'eau infiltree's dans le sol et sur le debit des sources. *Journal de Math. Pures et Appl.* Series 5, Tome X (in French) pp. 5–78.
14. Dumm, L.D. (1954). Drain spacing formula: New formula for determining depth and spacing of subsurface drains in irrigated lands. *Agricultural Engineering* **35**(10): 726–730.
15. Dumm, L.D. (1960). Validity and use of the transient concept in subsurface drainage. Paper presented at American Society of Agricultural Engineers meeting held at Memphis, pp. 4–7.
16. Brooks, R.H. (1961). Unsteady flow of groundwater into drain tile. *Amer. Soc. Civil Eng., Proc.* **87**(IR2): 27–37.
17. Schilfgaarde, J. (1963). Design of tile drainage for falling water tables. *Amer. Soc. Civil Eng. Proc.* **89**(IR2): 1–17.

18. Schilfgaarde, J. (1965). Transient design of drainage systems. *Amer. Soc. Civil Eng. Proc.* **91**(IR3): 9–22.
19. Moody, W.T. (1966). Nonlinear differential equation of drain spacing. *Journal of Irrigation and Drainage Division, ASCE* **92**(IR2): 1–9.
20. Pandey, R.S. and Gupta, S.K. (1990). Drainage design equation with simultaneous evaporation from soil surface. *ICID Bulletin* **39**(1): 19–25.
21. Singh, K.M., Singh, O.P., Ram, S., and Chauhan, H.S. (1992). Modified steady state drainage equations for transient conditions in subsurface drainage. *Agricultural Water Management* **20**(4): 329–339.
22. Singh, K.M., Singh, O.P., Chauhan, H.S., and Ram, S. (1996). Comparison of subsurface drainage theories for drainage of waterlogged saline soils of Haryana state, India. *Applied Engineering in Agriculture* **8**(5): 653–657.
23. Nikam, P.J., Chauhan, H.S., Gupta, S.K., and Ram, S. (1992). Water table behaviour in drained lands: effect of evapotranspiration from the water table. *Agricultural Water Management* **20**: 313–328.
24. De Boer, D.W. and Chu, S.T. (1975). Bi-level subsurface drainage theory. *Transactions of the American Society of Agricultural Engineers* **18**(4): 664–667.
25. Bouwer, H. and Van Schilfgaarde, J. (1963). Simplified method of predicting fall of water table in drained lands. *Transactions of the American Society of Agricultural Engineers* **6**(4): 228–291,296.
26. Chu, S.T. and DeBoer, D.W. (1976). Field and laboratory evaluation of bi-level drainage theory. *Transactions of the American Society of Agricultural Engineers* **19**(3): 478–481.
27. Sabti, N.A. (1989). Linear and nonlinear solution of the Boussinesq equation for the bi-level drainage problem. *Agricultural Water Management* **16**: 269–278.
28. Verma, A.K., Gupta, S.K., Singh, K.K., and Chauhan, H.S. (1998). An analytical solution for design of bi-level drainage systems. *Agricultural Water Management* **37**: 75–92.
29. Upadhyaya, A. and Chauhan, H.S. (2000). An analytical solution for bi-level drainage design in the presence of evapotranspiration. *Agricultural Water Management* **45**: 167–184.
30. Kraijenhoff Van de Leur, D.A. (1958). A study of non-steady ground water flow with special reference to a reservoir coefficient. *De Ingenieur* **70**: 87–94.
31. Maasland, M. (1959). Water table fluctuations induced by intermittent recharge. *Journal of Geophysical Research* **64**(5): 549–559.
32. Dass, P. and Morel-Seytoux, H.J. (1974). Subsurface drainage solutions by Galerkin's method. *Journal of Irrigation and Drainage Division* **100**(IR1): 1–15.
33. Marino, M.A. (1974). Water table fluctuation in response to recharge. *Journal of Irrigation Drainage Division, ASCE* **100**: 117–125.
34. Singh, S.R. and Jacob, C.M. (1976). Numerical solution of Boussinesq equation. *Journal of the Engineering Mechanics Division, Proceedings of ASCE* **102**(EM5): 807–823.
35. Singh, R.N. and Rai, S.N. (1989). A solution for the nonlinear Boussinesq equation for phreatic flow using an integral balance approach. *Journal of Hydrology* **109**(3/4): 319–323.
36. Skaggs, R.W. (1991). Modelling water table response to subirrigation and drainage. *Transactions of the American Society of Agricultural Engineers* **34**(1): 169–175.
37. Pandey, R.S., Bhattacharya, A.K., Singh, O.P., and Gupta, S.K. (1992). Drawdown solutions with variable drainable porosity. *Journal of Irrigation and Drainage Engineering* **118**(3): 382–396.
38. Pandey, R.S., Bhattacharya, A.K., Singh, O.P., and Gupta, S.K. (1997). Water table drawdown during drainage with evaporation/evapotranspiration. *Agricultural Water Management* **35**: 61–73.
39. Ram, S., Jaiswal, C.S., and Chauhan, H.S. (1994). Transient water table rise with canal seepage and recharge. *Journal of Hydrology* **163**: 197–202.
40. Werner, P.W. (1957). Some problems in non-artesian ground water flow. *Transactions American Geophysical Union* **38**(4): 511–518.
41. Schmid, P.A. and Luthin, J.N. (1964). The drainage of sloping lands. *Journal of Geophysical Research* **69**: 1525–1529.
42. Wooding, R.A. and Chapman, T.G. (1966). Groundwater flow over a sloping impermeable layer, 1. Application of the Dupuit-Forchheimer assumptions. *J. Geophys. Res.* **71**: 2895–2902.
43. Luthin, J.N. and Guitjens, J.C. (1967). Transient solutions for drainage of sloping land. *Journal of the Irrigation and Drainage Division, Proceedings of American Society of Civil Engineers* **93**(IR3): 43–51.
44. Chauhan, H.S., Schwab, G.O., and Hamdy, M.Y. (1968). Analytical and computer solution of transient water table for drainage of sloping land. *Water Resources Research* **4**(3): 673–679.
45. Childs, E.C. (1971). Drainage of ground water resting on sloping bed. *Water Resources Research* **7**(3): 1256–1263.
46. Jaiswal, C.S. and Chauhan, H.S. (1975). A Hele-Shaw model study of steady state flow in an unconfined aquifers resting on a sloping bed. *Water Resources Research* **II**(4): 596–600.
47. Ram, S. and Chauhan, H.S. (1987). Analytical and experimental solutions for drainage of sloping lands with time varying recharge. *Water Resources Research. American Geophysical Union* **23**(6): 1090–1096.
48. Upadhyaya, A. and Chauhan, H.S. (2001). Falling water tables in a horizontal/sloping aquifer. *Journal of Irrigation and Drainage Engg. ASCE* **127**(6): 376–384.
49. Kidder, E.H. and Lytle, W.F. (1994). Drainage investigation in plastic till soils of Northeastern Illinois. *Agr. Eng.* **39**: 384–386.
50. Neal, J.H. (1934). Proper spacing and depth of tile drains determined by the physical properties of the soil. *Minnesota Agr. Exp. Sta. Tech. Bull.* 101.
51. De Ridder, N.A. (1973). Drainage by means of pumping from wells pp. 223–237; *Theories of Field Drainage and Watershed Runoff*. Vol. II. *Drainage Principles and Applications; ILRI*. Wageningen, the Netherlands.
52. Attia, F.A.R. and Twinhof, A. (1989). Feasibility of tube well drainage in nile valley. pp. 303–326 Chapter 13 *Land Drainage in Egypt*. M.H. Amer and de Ridder (Eds.). Cairo, p. 377.
53. Awan, N.M. (1991). *Salinity control and reclamation project, a case study in 'Approaches to Integrated Water Resource Management in Humid Tropics Arid and Semi Arid Zones in Developing Countries' compiled by Maynara*. M. Hufschmidt, Janusz, Kindler.
54. Smedema, B. and Zimmer, D. (1994). *Vertical drainage and conjunctive use. GRID ISSN. 1021-268X*, pp. 7–8.
55. Smedema, B. (Dec. 1997). Biological drainage myth or opportunity? *GRID Iptrid Network Magazine*, p. 9.
56. Smedema, L.K. (2000). Global needs and challenges, the role of Drainage in today's world. *Proc. 8th ICID International Drainage Workshop*, 31st Jan–4th Feb, New Delhi, Vol. 1–1.

DRAWDOWN

John E. Moore
USGS (Retired)
Denver, Colorado

According to Wilson and Moore (1), drawdown is the lowering of the water level in a well as a result of withdrawal. Drawdown data are collected to evaluate the performance of the well and to determine the aquifer's hydraulic character (transmissivity). Drawdown measurements also provide information on the well efficiency and performance. For example, measurements can be used with well discharge data to detect deterioration of the well screen.

The drawdown in a pumping well is the sum of head loss factors (2). Some of the factors are natural processes, head loss in the aquifer, loss in the damage zone (well drilling), well development, and turbulent loss in the filter zone. The damage zone consists of drilling debris, filter cake, and drilling fluid.

When water is withdrawn from a well, the initial discharge is from casing storage and from aquifer storage near the well bore. As withdrawal continues, water is withdrawn at greater distances from the well. The cone depression or drawdown expands and deepens more slowly with time because an increasing amount of stored water is available from each foot of expansion of the cone (3). The cone of depression will continue to enlarge until the following conditions are met (4):

1. It intercepts a body of water (stream, lake, or wetland) that supplies enough water to equal the pumping rate.
2. Enough recharge from precipitation occurs within the cone of depression's influence to equal the pumping rate.
3. Leakage occurs from overlying or underlying aquifers to equal the pumping rate.

The cone of depression around a well is controlled by the transmissivity of the aquifer. In aquifers of a low transmissivity (10,000 gpd/ft_2), the cone is deep and has steep sides and a small radius. In aquifers of high transmissivity (100,000 gpd/ft_2), the cone is shallow and has a large radius

Slichter in 1899, Thiem in 1906, and Theis in 1935 (5) introduced the equations to predict drawdown. The transmissivity of the aquifer determines the shape of the drawdown curve.

The following is a list of key terms associated with well drawdown (6):

Static level— water level of a well that is not affected by the withdrawal of ground water.

Pumping level— the level of water in the well during pumping.

Well yield— the discharge of a well or rate at which a well yields water either by pumping or free flow.

Specific capacity— an expression of the productivity of a well obtained by dividing the rate of discharge by the drawdown in the well. It should be described on the basis of the number of hours of pumping prior to the time the measurement is made because it generally decreases with time as drawdown increases.

Well efficiency (well hydraulics)— defined as the ratio of actual specific capacity to theoretical specific capacity. Actual specific capacity is related to drawdown in the well and theoretical specific capacity. Well efficiency is determined by a step drawdown test.

Step drawdown test— In this test, the well is pumped at several (three or more) successively higher pumping rates, and the drawdown at each rate is recorded. The test is usually conducted during 1 day. The discharge is kept constant through each step.

Getting access to the well to take drawdown measurements is sometimes difficult. Ideally, you can gain access through a pipe intended for measurements. Measurements of drawdown in wells can be obtained by one of the following methods (7):

1. wetted steel tape
2. air-line submergence method
3. electrical tape
4. pressure transducer

Wetted Steel Tape

Before the 1960s, most water level measurements were made with a steel tape (most likely a 100- or 200-foot Lufkin tape). It is typically used for depths up to 90 feet. Electrical methods and pressure transducers in part have replaced tapes. However, the steel tape still has applications today, for example, to calibrate pressure transducers. To use the wetted tape method, the hydrologist will need a steel measuring tape, a weight, and carpenter's chalk. Before the tape is lowered down the well, the lower 1 to 2 feet or so of the tape is coated with carpenter's chalk. Problems that may be encountered using this method are moisture on the well casing, cascading water, and oil (leaking from pump lubrication). In some cases, this method can be very inaccurate. For example, if the water level is below the well screen, water flowing into the well can wash the chalk off the tape.

Electrical Tape

Most electrical tapes are marked every one-tenth foot. An electrical probe is lowered into the water, which completes an electrical circuit, and this sounds a buzzer or light.

Pressure Transducer

The transducer probe measures the column of water above the probe. It is very useful in aquifer tests where the water level changes rapidly with time.

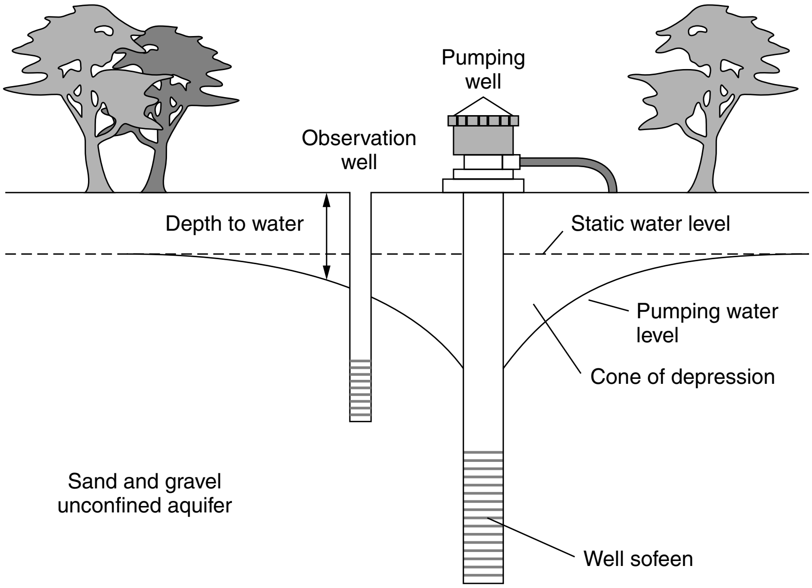

Figure 1. Diagram of a well (6).

Specific capacity tests can be used to determine the amount of water a well will yield. In this test, the pumping rate and drawdown (water-level changes) are monitored for a set period of time. The first step is to measure the initial water level in the well. Commonly, a well is pumped at several successively increasing rates for uniform periods (typically 1 hour) to establish a rate that can be maintained for long-term pumping. The well is then pumped at a steady rate and the water-level changes are monitored at the pumped well. Water levels should also be monitored in at least one observation well 2 to 20 meters (6 to 65 feet) from the pumped well. The water level will decline quickly at first as water is removed from the well, then more slowly as the rate of flow into the well approaches the pumping rate. The ratio of the discharge rate (Q) to the water-level change (drawdown, dd) gives the well's specific capacity, or $\mathrm{Sc} = Q/\mathrm{dd}$. For example, if the discharge rate is 6 liters per second (L/s) [100 gallons per minute] and drawdown is 3 meters (10 feet), the specific capacity of the well is 2 L/s per meter (10 gpm/ft) of drawdown. Once the specific capacity and the available amount of drawdown are known, the yield of the well can be determined from the formula $Q = \mathrm{Sc} \times \mathrm{dd}$. An estimate of the transmissibility in an unconfined aquifer can be obtained by multiplying the specific capacity by 2000.

A step-drawdown test evaluates the performance of the well. Well performance can be affected by resistance to flow in the aquifer itself; partial penetration of the well screen, incomplete removal of mud from the gravel envelope, or invasion of fines into the envelope; and blockage of part of the screen area. The well should be developed prior to the test by using a surge block and/or pumping until the well discharge is clear. In this test, the well is pumped at several (three or more) successively higher pumping rates, and the drawdown for each rate is recorded. The test is usually conducted during 1 day. The discharge is kept constant through each step. The test measures the change in specific capacity. The data provide a basis to choose the pump size and rate for the aquifer test and for long-term production.

BIBLIOGRAPHY

1. Wilson, W. and Moore, J. (1998). *Glossary of Hydrology*. American Geological Institute, Alexandria, VA, p. 248.
2. Roscoe Moss Company. (1990). *Handbook of Ground Water Development*. Wiley-Interscience, New York, p. 493.
3. Theis, C.V. (1938). The significance and nature of the cone of depression in ground water bodies. *Econ. Geol.* **35**:889–902.
4. Driscoll, F.G. (1986). *Groundwater and Wells*. Johnson Division, UOP Inc.
5. Theis, C.V. (1935). The lowering of the piezometer surface and the rate and discharge of a well using ground water storage. *Trans. AGU* **16**: 519–524.
6. Moore, J.E. et al. (1995) *Groundwater-a Primer*. American Geological Institute, Alexandria, VA, p. 53.
7. Moore, J.E. (2002). *Field Hydrogeology, a Guide for Site Investigations and Report Preparation*. Lewis, Boca Raton, FL, p. 195.

WATER LEVEL DRAWDOWN

Mark D. Steele
MDC Systems, Inc.
Berwyn, Pennsylvania

INTRODUCTION AND DEFINITIONS

Contrary to the popular image of caverns and underground rivers, ground water is the water that fills or saturates the soil in a specific location. *Water level drawdown* describes the effect on the actual or potential water surface of removing water from a well, trench, or sump. The water table is the interface between this saturated water/soil zone (or aquifer) and an unsaturated water/soil zone. The concept of an aquifer can be easily seen by filling a jar with rocks. Water can then be poured into the jar to fill up the empty spaces between the rocks. The level of the water in the jar is analogous to the subsurface water table, and the volume saturated with water is the aquifer. If one then placed a straw in the middle of the jar and began to pull out water, the water table would become depressed in a cone shape around the straw similar to the depression in the water level of a sink as the water goes down the drain. This depression in the water table is known as the cone of depression. The vertical difference, at any given point, between the original (or static) water level and the water surface formed by the cone of depression is called the *drawdown*.

The goal of this article is to provide an understanding of the key terms and concepts associated with drawdown and an introduction to the uses of drawdown in groundwater hydrology. The meaning of drawdown varies slightly between unconfined and confined (artesian) aquifers.

DRAWDOWN IN UNCONFINED AND CONFINED AQUIFERS

Depending upon whether an aquifer is unconfined or confined, drawdown is discussed in terms of either the water table or the piezometric surface, respectively (Fig. 1).

Unconfined aquifers have no impermeable layer between the water table and the ground surface. They can also be called “free” or “water table” aquifers. The water table in these aquifers is in direct contact with the atmosphere through the pores of the unsaturated soil above, and the water table is in balance with the air pressure. These aquifers are influenced by the infiltration of storm water and are also more easily contaminated through the infiltration of chemicals from surface spills.

For unconfined aquifers, the static water level is the height of the water in the well before pumping (Fig. 2). This is often calculated by measuring the depth to the water

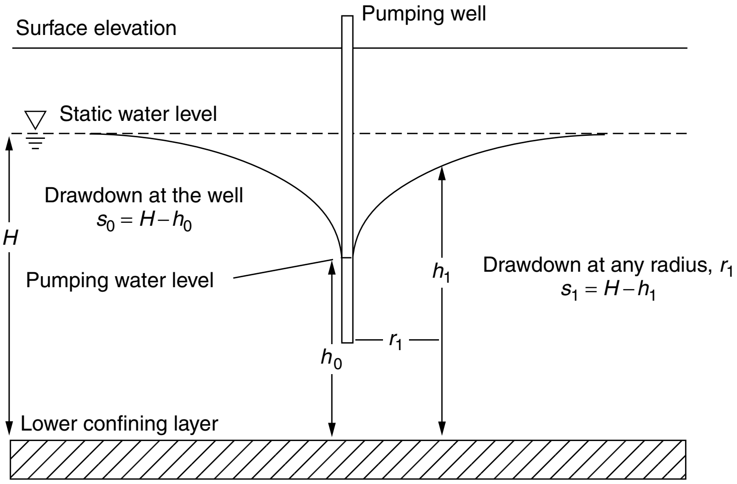

Figure 1. Drawdown in an unconfined aquifer.

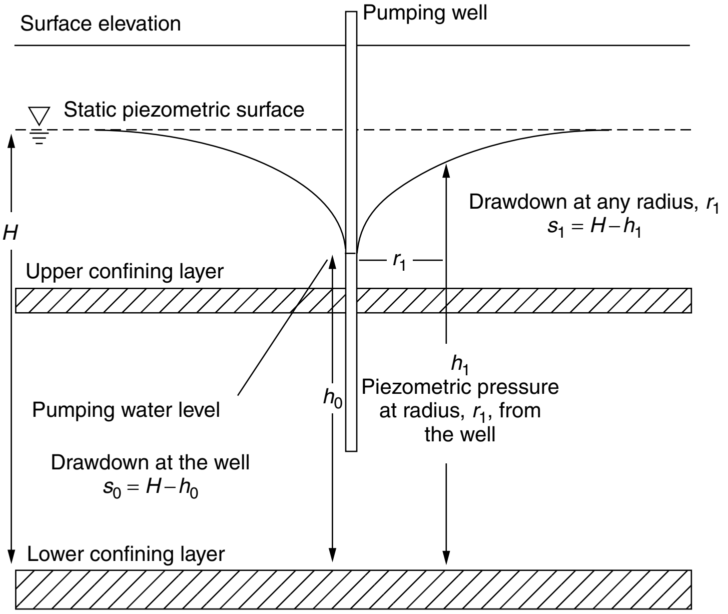

Figure 2. Drawdown in a confined (artesian) aquifer.

from the surface elevation. It can be given as "depth-to-water" or, sometimes, can be given as an elevation compared to mean sea level (MSL). In groundwater equations, it is expressed by a capital "H" and is measured, not from the surface down, but from the lower confining layer up. It is a measurement of pressure or head. So, if the depth to water is 26 feet and the surface elevation is 686 feet above MSL, the static water level is at 660 feet above MSL. If boring logs show that the lower confining level is at an elevation of 585 feet, then $H = 75$ feet ($660 - 585 = 75$).

Confined (or artesian) aquifers have a confining layer, an impermeable bed, between the aquifer and the ground surface. As a result, the water in confined aquifers is often under pressure. Artesian springs are one result of this phenomenon. If a well is drilled into a confined aquifer, the water level in the well will rise to the height of the pressure in the aquifer. Artesian springs occur when the water pressure is such that the water rises to the surface through a fracture in the confining layer. This height, or head, is the *piezometric* level at the well. A combination of all piezometric or *potentiometric* heights throughout a confined aquifer is the *piezometric surface*. The piezometric surface is what the water table would be if there were no upper confining layer.

In confined aquifers, the static piezometric surface is the piezometric surface measured before pumping begins. It is expressed in the same terms as the static water level and is also measured as a pressure, or head, from the lower confining layer.

The pumping water level in a well is a water level in the well during pumping. Like the static water level or static piezometric surface, the pumping water level is measured from the lower confining bed and is a pressure or head in feet or meters. It is usually designated by a lower case "h."

The difference between the static level and pumping level at the well is the *drawdown* at the well. It is usually designated by s_0. Observation wells can be placed at various distances or radii from the pumping well to observe the drawdown at these locations. The difference between the static level and the pumping level at any radius, r_1, from the well is the drawdown, s_1, and it is the difference between H and the pumping water level, h_1, at that radius.

If all the pumping levels could be viewed in cross section, they would form a curve, the *drawdown curve* similar to these shown in the two figures. Drawdown in three dimensions is an inverted cone known as the *cone of depression*. The cone of depression defines the distance at which drawdown occurs as a result of the pumping of the well in question. This distance is known as the *radius of influence* of the well, and the area circumscribed by this radius is known as the *area of influence*.

IMPORTANCE AND USE OF DRAWDOWN INFORMATION

The measurement of drawdown is very useful to the study of the hydraulic characteristics of an aquifer. To understand this importance, it is first necessary to review the basic equilibrium equations for both unconfined and confined aquifers. The equation for an unconfined aquifer is

$$Q = \frac{\pi K(H^2 - h^2)}{\ln\left(\frac{r_i}{r_w}\right)} \tag{1}$$

where

Q = pumping rate

K = coefficient of permeability

H = saturated thickness at the radius of influence (static water level)

h = saturated thickness at the well

r_i = radius of influence

r_w = radius of the well

The equation for a confined aquifer is

$$Q = \frac{2\pi KM(H - h)}{\ln\left(\frac{r_i}{r_w}\right)} \tag{2}$$

where

M = thickness of confined aquifer

H = piezometric pressure at the radius of influence (static piezometric)

h = piezometric pressure at the well

r_i = radius of influence

r_w = radius of the well

When these equations are adjusted to solve for K at any two radii, r_1 and r_2, using drawdown at those two radii, s_1 and s_2, then the equations become
for an unconfined aquifer

$$K = \frac{Q \ln\left(\frac{r_2}{r_1}\right)}{\pi(s_1 - s_2)} \tag{3}$$

and for a confined aquifer

$$K = \frac{Q \ln\left(\frac{r_2}{r_1}\right)}{2\pi M({h_2}^2 - {h_1}^2)} \tag{4}$$

These equations can be used as part of a pump test to determine the permeability of an aquifer. The distances to the various confining layers and the thickness of the aquifers (as well as the static levels) should have been determined by soil borings. So, H and M are known. A pumping well is drilled along with one or more observation wells set at known radii from the pumping well. The well is then pumped at a known rate, Q, and time–drawdown data are recorded for each of the wells. From this information, the coefficient of permeability for the aquifer can be developed. This coefficient can then be applied to design other wells in this aquifer or for modeling the aquifer.

READING LIST

Cedergran, H. (1989). *Seepage, Drainage and Flow Nets*, 3rd Edn. John Wiley & Sons, New York.

Corbitt, R.A. (1999). *Standard Handbook of Environmental Engineering*, 2nd Edn. McGraw-Hill, New York, pp. 5.32–5.56.

Driscoll, F.G. (1987). *Groundwater and Wells*, 2nd Edn. Johnson Division, St. Paul, MN.

Ground Water Manual, A Water Resources Technical Publication. (1985). U.S. Department of the Interior, Bureau of Reclamation.

Parmley, R.O. (1992). *Hydraulics Field Manual*. McGraw-Hill, New York.

Powers, J.P. (1992). *Construction Dewatering*, 2nd Edn. John Wiley & Sons, New York.

WATER WELL DRILLING TECHNIQUES

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Water well installation, for groundwater monitoring projects or for water supply wells for irrigation or human consumption involves a variety of drilling techniques. Rig selection is related to accessibility, time and cost of project, sediment type (consolidated rock or unconsolidated soils), sample type (undisturbed vs. disturbed), and sample integrity. For environmental projects, designing wells and evaluating groundwater also involves assessing the soil, sediments, or rocks above and below the water table.

The vadose zone includes the zone immediately above the water table to the surface. The lower vadose zone includes the capillary fringe, the zone where water is drawn upward by capillary force. The water table is the top of the saturated or phreatic zone within the aquifer.

More detail regarding water wells, aquifers, and drilling techniques is provided in Driscoll (1), Testa (2), Sisk (3), and DWR (4,5). There are numerous differences in drilling techniques, methods and equipment, many regional variations, and specific local requirements. A generalized summary of selected water well drilling techniques follows.

ROTARY DRILLING

Rotary drilling techniques include direct mud rotary, air rotary, air rotary with a casing driver, and dual-wall reverse circulation. Rotary drilling techniques are commonly used for deep water supply wells. Direct mud rotary drilling uses fluid, which is pumped down through the bit at the end of the drill rods and is circulated up the annular space back to the surface. The fluid at the surface is routed via a pipe or ditch to a sedimentation tank or pit, then to a suction pit where the fluid is recirculated back through the drill rods. Air rotary drilling uses air as a circulation medium instead of water. In unconsolidated deposits, direct mud or air rotary can be used, providing that a casing is driven as the drill bit is advanced. In dual-wall reverse circulation, the circulating medium (mud or air) is pumped downward between the outer casing and inner drill pipe, out through the drill bit, then up the inside of the drill pipe.

Figure 1. Rotary drilling rig used to assess shallow spillage from train derailment area at mile 262 (Photo P045; EPA).

Rotary drilling techniques are commonly limited to consolidated deposits of rocks (Fig. 1). In mud rotary, a mud filter cake develops along the borehole wall, potentially reducing aquifer permeability. Where a resistant layer such as overlying basalt flows or conglomerate strata exists at shallow levels within the vadose zone and above the target depth, an air rotary rig can be used to drill to a predetermined depth followed by another more suitable drilling method (2).

CABLE-TOOL DRILLING

Cable-tool drilling is the oldest drilling technique available and is for installing water supply wells in selected locations. It is not used often in the environmental field, as the technique is slow, noisy, and dusty. The exception is the use of cable-tool drilling in glacial environments containing large cobbles in the Pacific Northwest portion of the United States or in young volcanics such as in Hawaii. Cable-tool rigs, called percussion or spudder rigs, operate by repeatedly lifting and dropping the heavy string of drilling tools in the borehole, crushing larger cobbles and rocks into smaller fragments. During cable-tool drilling, the hole is continuously cased by an unperforated steel casing with a drive shoe. The casing is attached on top by a rope socket to a cable that is suspended through a pulley from the mast of the drill rig. The process of driving the casing downward a few feet is followed by periodically bailing the borehole of the broken rocks and accumulated soils from the bottom of the borehole. Formation water or added water is used to create a slurry at the bottom of the borehole.

WIRE LINE CORING

Coring is the drilling method that produces cylindrically shaped cores. A rotary rig is used in conjunction with water, drilling mud, or air. Cutting is accomplished by drill bits located at the end of the rotating barrel or tube. The barrel gradually slides down into the annular opening. The core is then separated from the rest of the formation mass, and the barrel containing the core is retrieved.

HOLLOW-STEM AUGER DRILLING

Hollow-stem continuous flight auger drilling techniques are commonly used for subsurface environmental projects (Fig. 2). Hollow stems consist of a series of continuous, interconnected hollow auger flights. The hollow-stem flight augers are hydraulically pressed downward and rotated to start drilling. Soil cuttings are rotated up the outside of the continuous flighting in the borehole annulus. A center rod with plug and pilot bit are mounted at the bottom. The plug is designed to keep soil from entering the mouth of the lead auger while drilling. Upon reaching the sampling depth, the center rod string with plug and pilot bit attached is removed from the mouth of the auger and replaced by a soil sampler.

The soil sampler is lowered into the borehole through the hollow stem of the auger (the center tube), and sampling is started. Samples can be continuously retrieved but are typically collected at 5-foot intervals, at hydrologic or lithologic changes, or at intervals of obvious contamination.

HORIZONTAL DRILLING

Horizontal or lateral radial wells are used more in subsurface environmental remediation although they have been used by the oil industry for decades. Based on the configuration in map view, horizontal water wells emanate from a center hub well. The most obvious application is where the area of concern, such as a contaminant plume, is inaccessible due to aboveground structures, tankage, roads, or subsurface structures such as landfills, lagoons, pits, pipelines, or wells.

Figure 2. Close-up of a hollow-stem auger rig (Courtesy of Joe Ryan, University of Colorado).

Figure 3. Direct push technology rig (Courtesy FAST-TEK).

DIRECT PUSH DRILLING

Direct push technology (DPT) rigs are used almost exclusively in the environmental field (Fig. 3). DPT rigs rely on the static weight of the vehicle, typically a pick-up truck or van, combined with the percussion of the onboard industrial jackhammer as the energy for advancing the soil and groundwater samplers. DPT rigs are designed for easier accessibility than larger more conventional drilling rigs. DPT rigs can install small-diameter driven well points or piezometers. Unlike rotary auger rigs, DPT rigs do not generate large amounts of drilling derived wastes. DPT rigs, including cone penetrometer technology (CPT) rigs, can also drive various sensors tools to obtain data about subsurface conditions (6).

WELL INSTALLATION

Well materials for water supply wells and monitoring wells must be chemically compatible with potential contaminants as well as water geochemistry, such as pH, iron content, turbidity, alkalinity, and other parameters (Fig. 4). Casing, both blank and screen sections, can be constructed of fiberglass-reinforced plastic, stainless steel, concrete, or thermoplastic which include polyvinyl chloride (PVC), acrylonitrile butadiene styrene (ABS), and styrene rubber (SR). Based on cost, availability, and chemical compatibility, the most common casings and screens used for shallow drilling projects in the environmental field are made of PVC. Deeper wells are typically constructed of either thicker gauge PVC or steel. Well screen openings are commonly slotted, lowered, or created through a wire wrapping process. Slot sizes can be determined by previous well installations in the area or by a grain size sieve analysis. Centering devices may be fastened to the casing to assure even distribution of filter material and grout within the borehole annulus.

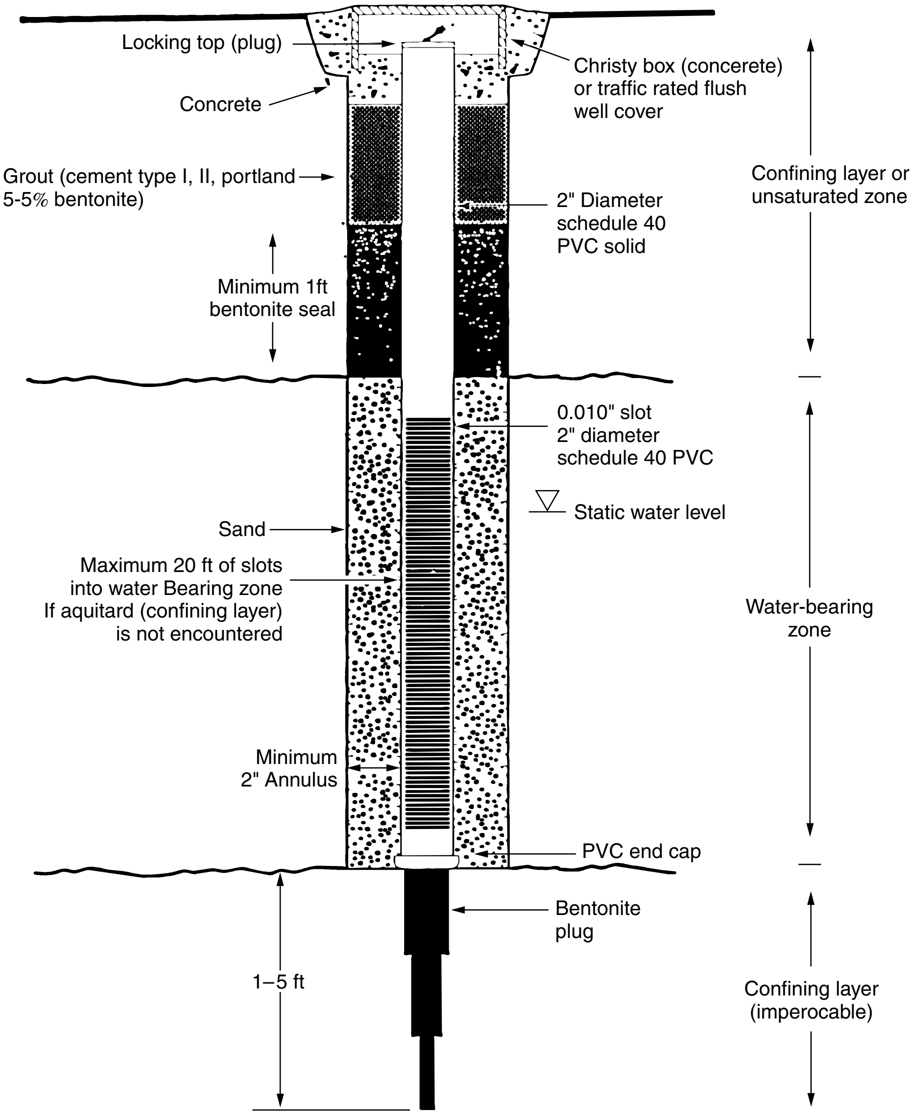

Figure 4. Groundwater monitoring well construction diagram.

BIBLIOGRAPHY

1. Driscoll, F.G. (1986). *Groundwater and Wells*, 2nd Edn. Johnson Filtration Systems, St. Paul, MN, pp. 268–339.
2. Testa, S.M. (1994). *Geologic Aspects of Hazardous Waste Management*. CRC, Boca Raton, FL, pp. 145–187.
3. Sisk, S.J. (1981). *NEIC Manual for Groundwater/Subsurface Investigations at Hazardous Waste Sites*. EPA-330/9-81-002, U.S. Environmental Protection Agency, Denver, CO.
4. California Department of Water Resources. (1981). California Well Standards, Bulletin 74–81, Sacramento, CA.
5. California Department of Water Resources. (1990). California Well Standards, Bulletin 74–90, Sacramento, CA.
6. Jacobs, J. (2000). Monitoring well construction and sampling techniques. In: *Standard Handbook of Environmental Science, Health, and Technology*. J. Lehr (Ed.). McGraw-Hill, New York, pp. 11.46–11.68.

READING LIST

American Society for Testing Materials. (1984). *Standard Practice for Description and Identification of Soils (Visual-Manual Procedure), Method D 2488-84*. Philadelphia, PA.

California Regional Water Quality Control Board. (1989). *Leaking Underground Fuel Tank (LUFT) Field Manual: Guidelines for Site Assessment, Cleanup, and Underground Storage Tank Closure*. Sacramento, CA.

California Regional Water Quality Control Board. (1990). *Tri-Regional Board Staff Recommendations for Preliminary Investigation and Evaluation of Underground Tank Sites*. Sacramento, CA.

Munsell Color. (1988). *Munsell Soil Color Charts*. Munsell Color, Baltimore, MD.

GROUNDWATER DYE TRACING IN KARST

David M. Bednar, Jr.
Michael Baker Jr., Inc.
Shreveport, Louisiana

Groundwater dye tracing has proven to be an effective tool for aquifer characterization, protection, and to provide remediation strategies in karst areas. Tracing groundwater flow routes with fluorescent dyes is highly successful because they are water soluble and inexpensive. Fluorescent dyes are usually the best tracers to use in karst because they are easily detected in concentrations that are one to three orders of magnitude less than those at which nonfluorescent dyes can be measured spectrometrically (1). Additionally, they are safe to use, work effectively in different hydrogeologic settings, and can be used effectively to trace water from subsurface to surface water bodies. Dye tracing in karst has been used successfully to (1) delineate spring recharge areas and subsurface basins, (2) determine site-specific hydrology, (3) estimate groundwater flow velocities, (4) trace groundwater flow from areas of recharge to discharge, (5) delineate wellhead protection areas, (6) map and characterize conduit flow routes, (7) detect leakage from residential sewage disposal systems, (8) identify sources of potential pollution from hazardous waste sites, (9) detect leakage from dam sites, and (10) characterize groundwater flow routes to important springs and cave streams along highway corridors.

WHAT IS KARST?

To understand why groundwater dye tracing is an effective tool for characterizing karst groundwater resources, one must have a basic understanding of karst and the unique hydrologic characteristics of its landscape. Karst refers to lands primarily underlain by limestone and dolomite where surface water is integrally connected to the groundwater system through preferential flow routes, which results in the formation of distinctive surface landforms and hydrologic features as well as subsurface features. Lack of these landforms and features does not mean that karst is not present.

A characteristic feature of karst areas is the sometimes rapid interconnection of surface water with the groundwater flow system. Even in the absence of surface streams, a karst region is a zone of drainage into the aquifer; the entire area can be a recharge zone (21). In most karst areas, two general types of recharge have been recognized: discrete and diffuse recharge. Discrete recharge, also known as concentrated recharge, is characterized by relatively rapid movement of water through localized areas (such as through sinkholes, losing streams, or other areas) toward the groundwater flow system. Areas of discrete recharge transport much of the water through preferential flow routes that commonly transport water at rates several orders of magnitude greater than those encountered on nonkarst groundwater systems, support turbulent flow, are too large to provide effective filtration for most pathogens, and provide minimal adsorption or other natural cleansing processes (4). Substantially greater quantities of water per unit area enter the groundwater system through discrete recharge areas than through diffuse recharge (5). For this reason, the groundwater system is highly vulnerable to contamination from accidental spills and poor land use practices. Groundwater flow velocities in many karst areas vary as much as 10 to 1500 ft per hour between the same two points—the latter, in response to storms—and are tens of thousands to several million times faster than those characteristic of many granular aquifers (1).

Diffuse recharge refers to the general and relatively slow seepage and percolation of recharge toward the groundwater system.

THE EPIKARST

The epikarst is the uppermost portion of the bedrock that consists of fissures and cavities formed by dissolution. The dissolution features in the epikarstic zone are organized to move infiltrating water laterally to downgradient seeps and springs or to collector structures such as shafts that conduct the water farther into the subsurface (6). The epikarst can vary from essentially zero to 30 meters or more (4) and is controlled by such factors as climate, depth of groundwater circulation, and bedrock structure.

Aley (7) summarized general dye recovery results for karst aquifers where dye was introduced directly into locations of discrete recharge that ultimately discharged to springs that passed through the epikarst. Aley recognized three hydrologically distinctive epikarstic zones in karst areas; rapid draining epikarsts, seasonally saturated epikarsts, and perennially saturated epikarsts. Rapid draining epikarsts commonly occur in areas of high topographic relief in bedrock of high solubility and negligible sediment infiltration. Rapidly draining epikarsts are saturated with water for short periods of time, especially after storms, and have little water storage or detainment. Seasonally saturated epikarsts occur in areas of moderate relief where the solubility of the bedrock has resulted in the development of soil and residuum thickness and at elevations greater than local perennial streams. Water is typically stored seasonally and after major storms lasting periods of weeks to months. Perennially saturated epikarsts occur in areas of low to moderate relief along perennial streams and are mostly saturated with water.

Groundwater dye tracing in the epikarst is more complex than that reported in most karst groundwater tracing reports. Aley (7) stressed that dye tracing in the epikarst requires extensive sampling and quantitative analysis, more detailed and quantitative analysis of background fluorescence characteristics, and simultaneous use of multiple dyes with dye quantities and analytical approaches selected to minimize the chance that small dye recoveries are obscured by another dye.

Fluorescent Dyes

Many fluorescent dyes have been used in groundwater dye tracing to characterize flow routes in karst areas. For general problem solving in karst areas, eosine, fluorescein, pyranine, rhodamine WT, and sulforhodamine B are the most useful fluorescent dyes (8). The characteristics of these dyes are well documented in Aley (9), Smart and Laidlaw (10), and Kass (11). The selection of dyes to be used, the location of dye introduction points, the manner in which the dyes are introduced, the sampling strategy employed, and the analytical approach used must be tailored to the hydrogeologic setting, the issues of concern, and the quality and credibility of the data needed for the study (12).

Sampling Locations

In karst landscapes, water that goes down into the subsurface eventually comes up at a spring or a series of springs. Springs represent the final terminus or discharge points of groundwater flow in karst areas and serve as excellent sampling locations to monitor for dye introduced into the groundwater system (surface streams, pumping wells, and monitoring wells can also be used).

It is equally important to know where the dye does not go as where the dye does go. All possible discharge points should be sampled. All springs within a radius of perhaps 5 to 15 miles from a facility, especially those within ±90° of the vector of the hydraulic gradient from it, should be monitored during dye traces (13). For maximum results, dye tracing should be conducted during low and high flow conditions. Groundwater movement during periods of high flow may be diverted to higher preferential flow routes that may discharge at springs in adjacent groundwater basins. For this reason, it may be necessary to monitor 10 to 40 sites during a single trace to ensure that dye is detected. However, due to time and budget constraints, dye trace tests are often designed to maximize current hydrologic conditions.

GROUNDWATER TRACING METHODS

Activated carbon samplers are often used to recover dyes. Activated carbon samplers (also called charcoal samplers or passive detectors) consist of a few grains of activated coconut charcoal placed in heat-sealed packets of fiberglass screening. They are placed at sampling locations to adsorb continuously and thus accumulate specific fluorescent dyes (Figs. 1 and 2). As an illustration, a charcoal sampler in place in flowing water containing fluorescein or rhodamine WT for a week will typically contain about 400 times more dye upon analysis than the mean dye concentration in the water being sampled (14).

The sampling interval is based on site-specific conditions and the questions that need to be answered by the study. Weekly sampling intervals are appropriate for most studies. Once dye introduction points and sampling locations have been identified, Aley (7) recommends one or more rounds of sampling at most (and preferably all) of the sampling locations to characterize background fluorescence before the final selection of dye types and quantities are determined. The amount of dye used for tracing studies is typically based on professional experience, as there is no credible standard equation for estimating dye quantities needed for groundwater tracing work (8).

Figure 1. Activated charcoal packet placed at Big Spring, Hardy County, West Virginia. Spring discharges near red house located in the upper left corner of photo.

Figure 2. Activated charcoal packet attached to rock prior to placement in Waites Run, Hardy County, West Virginia.

SELECTION OF DYE INTRODUCTION POINTS

In karst areas, sinking streams (Fig. 3) and sinkholes (Fig. 4) are commonly used as dye introduction points. However, the selection of appropriate dye introduction points depends on the type of study being performed. On some occasions, it is often difficult to encounter overland flow at ideal introduction points. This problem can be resolved by using “dry sets.” A dry set involves the placement of a dye so that it will be flushed into a surface drainageway, sinkhole, culvert, or roadside ditch by the first storm flow (8).

Additional successful techniques used for introducing dye, especially at hazardous waste sites, include epikarstic dye introduction points (EDIPs) and dye introduction trenches (DITs). EDIPs are vertical boreholes that enter the top of the bedrock. DITs are constructed with a backhoe, typically 17 to 33 feet long, and extend into the epikarst. The construction and use of EDIPs and DITs, in addition to determining water quantities used to flush dye into the subsurface, are described in Aley (7,8).

Groundwater monitoring wells are typically poorly suited for use as dye introduction points primarily because they often fail to intersect the preferential flow openings that transmit water flow.

ANALYTICAL STRATEGIES

Quinlan (13) identified three analytical strategies used during dye trace studies. Qualitative tracing involves introducing dye at some discrete recharge point or zone and either visually observing the dye plume or

Figure 3. Fluorescein dye introduced into a sinking stream that flows across recharge units of the Capon Warm Springs Complex in Hampshire County, West Virginia.

visually observing dye removed from passive detectors. This method can be somewhat problematic especially when the dye arrives at a monitoring station at night. Qualitative dye tracing can be used to delineate boundaries of groundwater basins and identify point-to-point connections between input and recovery points, allowing special attention to be given to those points that have the highest potential for contaminating the groundwater system (15). Time of travel and flow velocity may generally be determined through qualitative dye tracing. Semiquantitative tracing uses passive detectors at sampling points followed by analysis with an instrument such as a scanning spectrofluorometer. Quantitative tracing includes instrumental analysis of water samples collected at sampling stations. Passive detectors are often used in concert with collection of water samples. The result of quantitative dye tracing is a set of measured dye concentrations, each sampled at a specific place and time. This method makes it possible to determine, with greater precision, groundwater flow velocities within preferential flow routes from points of recharge to discharge. Quantitative tracing provides a measure for determining the effectiveness of the monitoring system by estimating the tracer loss involved. Inadequate tracer recoveries are an indication that losses other than sorption or decay may be significant (16).

DYE TRACING AT WASTE DISPOSAL SITES

Figure 4. Fluorescein dye introduced into a perennial stream that flows into a sinkhole near Greenland Gap, Grant County, West Virginia.

Groundwater dye tracing in karst is used for a wide variety of purposes, however, one of the most important uses of dye tracing is for investigations at waste disposal sites. Investigations at waste sites in mature karst should include, locating all springs, streams in sinkhole bottoms, and major streams in caves; dye tracing to establish connections between the disposal site, springs, and underground streams; monitoring connected points; monitoring at least one spring shown by dye tracing not connected to the site, as a control; and, dye tracing to delineate boundaries of groundwater basins (17).

Groundwater monitoring at facilities regulated under the Resource Conservation and Recovery Act (RCRA) and the Comprehensive Environmental Response, Conservation, and Liability Act (CERCLA, more commonly known as Superfund sites) require the installation of groundwater wells to monitor the potential leakage of contaminants from the site. The problem with using groundwater monitoring wells in karst areas is that they often fail to intercept the conduits through which pollutants move. In many karst settings, most groundwater flow occurs in discrete conduits that are dendritic or trellised and terminates at springs whose water quality is representative of the mean of the ground water basin (17). Monitoring wells located on the basis of fracture trace and lineament analysis often fail to intercept major conduits in karst areas (18).

BIBLIOGRAPHY

1. Quinlan, J.F. et al. (1992). Rationale for the Design of Cost-Effective Groundwater Monitoring Systems in Limestone and Dolomite Terranes: Cost-Effective as Conceived is Not Cost-Effective as Built if the System Design and Sampling Frequency Inadequately Considers Site Hydrogeology. In: *Proceedings of the 8th Waste Testing and Quality Assurance Symposium*. U.S. Environmental Protection Agency, pp. 552–570.
2. Veni, G. et al. (2001). *Living With Karst: A Fragile Foundation*. American Geological Institute.
3. Aley, T. (1978). Ozark hydrology: A predictive model. *Journal of the Missouri Speleological Survey* **18**: 185.
4. Aley, T. (2000). Karst systems. In: *The Standard Handbook of Environmental Science, Health, and Technology*. H. Jay and J.K. Lehr (Eds.). McGraw-Hill, New York, pp. 19.1–19.10.
5. Aley, T. and Halterman, D. (1982). A Conceptual Characterization of the Subsurface Movement of Toxic Chemicals in Soluble Rock Lands. In: *Proceedings of the Fifth National Cave Management Symposium*, pp. 77–80.
6. Huntoon, P.W. (1995). Is it appropriate to apply porous media ground water circulation models to karstic aquifers?. In: *Groundwater Models for Resources Analysis and Management*. Lewis Publishers, Boca Raton, FL, pp. 339–358.
7. Aley, T. (1997). Groundwater tracing in the epikarst. In: *The Engineering Geology and Hydrogeology of Karst Terranes*. Beck and Stephenson (Eds.). Balkema, Rotterdam, pp. 207–211.
8. Aley, T. (1999). *The Ozark Underground Laboratory's Groundwater Tracing Handbook*. The Ozark Underground Laboratory, p. 35.
9. Aley, T. (1995). *Recharge Area Delineation for Bear Creek Springs and Smokehouse Spring*. Contract Report to the Arkansas Highway and Transportation Department.
10. Smart, P.L. and Laidlaw, I.M.S. (1977). An evaluation of some fluorescent dyes for water tracing. *Water Resources Research*, **13**(1): 15–33.
11. Kass, W. (1998). *Tracing Technique in Geohydrology*. A.A. Balkema, Rotterdam.
12. Bednar, D., Jr. and Aley, T. (2001). Groundwater dye tracing: An effective tool to use during the highway development process to avoid or minimize impacts to karst groundwater resources. In: *Geotechnical and Environmental Applications of Karst Geology and Hydrology—Proceedings of the 8th Multidisciplinary Conference on Sinkholes and The Engineering and Environmental Impacts of Karst*. B.F. Beck and J.G. Herring (Eds.). A.A. Balkema Publishers, Louisville, KY, April 1–4, pp. 201–207.
13. Quinlan, J.F. (1989). *Ground-Water Monitoring in Karst Terranes: Recommended Protocols and Implicit Assumptions*. U.S. Environmental Protection Agency.
14. Aley, T. (1994). *Final Report: Tracing Study to Assess Groundwater Flow Routes for Selected Springs Near Appalachian Corridor H*. Contract report prepared for Michael Baker Jr. Inc. for the benefit of the West Virginia Department of Transportation.
15. Mull, D.S., Liebermann, T.D., Smoot, J.L., and Woosley, Jr., L.H. (1988). *Application of Dye Tracing Techniques for Determining Solute-Transport Characteristics of Ground Water in Karst Terranes*. EPA 904/6-88-001, United States Environmental Protection Agency.
16. Field, M.S. (1999). *The QTRACER Program for Tracer-Breakthrough Curve Analysis for Karst and Fractured-Rock Aquifers*. U.S. Environmental Protection Agency.
17. Quinlan, J.F. and Ewers, R.O. (1985). Ground Water Flow in Limestone Terranes: Strategy Rationale and Procedure for Reliable, Efficient Monitoring of Ground Water Quality in Karst Areas. In: *Proceedings of the National Symposium and Exposition on Aquifer Restoration and Groundwater Monitoring*. National Groundwater Association, pp. 197–234.
18. American Society For Testing and Materials (ASTM). (1995). *Standard Guide for Design of Ground-water Monitoring Systems in Karst and Fractured-Rock Aquifers*. American Society For Testing and Materials.

EARTHQUAKES—RATTLING THE EARTH'S PLUMBING SYSTEM

Michelle Sneed
Devin L. Galloway
William L. Cunningham
U.S. Geological Survey
Denver, Colorado

Hydrogeologic responses to earthquakes have been known for decades, and have occurred both close to, and thousands of miles from earthquake epicenters. Water wells have become turbid, dry or begun flowing, discharge of springs and ground water to streams has increased and new springs have formed, and well and surface-water quality have become degraded as a result of earthquakes. Earthquakes affect our Earth's intricate plumbing system—whether you live near the notoriously active San Andreas Fault in California, or far from active faults in Florida, an earthquake near or far can affect you and the water resources you depend on.

RESPONSES IN WATER WELLS

Water-level fluctuations caused by the 1964 magnitude(M) 8.5 Alaska earthquake were recorded in 716 wells in the United States (1); the earthquake also was registered on water-level recorders in many other countries. Responses of water levels in wells to earthquakes are influenced by such factors as the magnitude and depth of the earthquake, distance from the epicenter, and the hydrogeologic environment, which may be consolidated rock, unconsolidated sediment, karst, or hydrothermally altered. The depth of the well, whether the aquifer is confined or unconfined, and well construction also influence the degree of water-level fluctuations in wells in response to seismic waves. Some aquifers may act as resonators, which may amplify the response.

The most common type of observed ground-water response is an instantaneous water-level offset, or step, which may be either an increase or a decrease and may occur near or far from the epicenter. Recovery to the pre-earthquake water level can be so rapid that no change will be detected if the water level is measured

---

Photo ©1994, Stephen P. Mulqueen

This petroleum seep in the Ojai Valley, California developed as a direct result of the 1994 Northridge earthquake in southern California. Notice the oil has not yet spread far and grass blades can be seen standing through the crude oil. The photo was taken a few days after the earthquake.

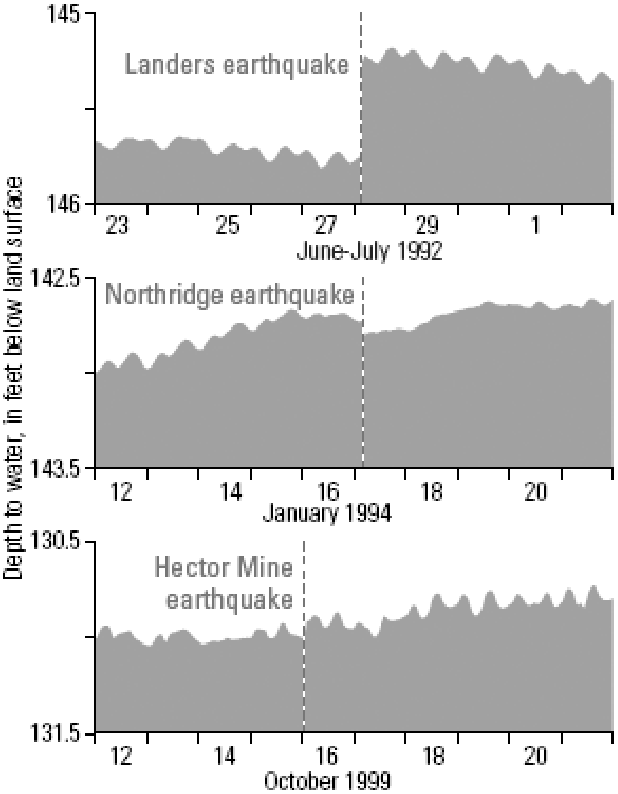

Well MO-18/02W/29-0017 in Wisconsin responded to the distant Denali Fault earthquake in Alaska.

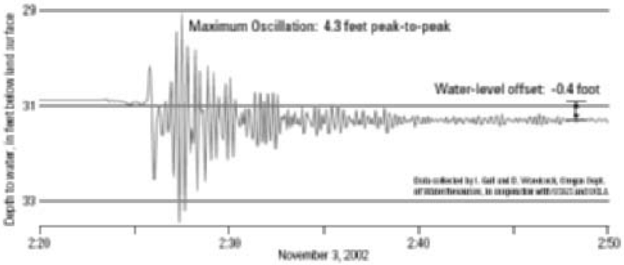

Hydroseismogram recorded in a well near Grants Pass, Oregon, shows water-level oscillations and a small offset relative to pre-earthquake water level in response to the Denali Fault earthquake.

infrequently, or it may take as long as days or months. Steps can be large enough to make a well flow at land surface, or render it dry. The 1998 M5.2 Pymatuning earthquake in northwestern Pennsylvania caused about 120 local household-supply wells to go dry within 3 months after the earthquake (2). The 2002 M7.9 Denali Fault earthquake in Alaska caused a 2-foot water-level rise in a well in Wisconsin, more than a thousand miles from the epicenter.

The other type of ground-water response is a water-level oscillation, which occurs more often, but is less commonly recorded. In the few cases where oscillations have been recorded, they resemble long-period seismograms, known as hydroseismograms. A well in Grants Pass, Oregon, is instrumented to record water levels at 1-second intervals, and the record from the Denali Fault earthquake shows peak-to-peak seismic oscillations of more than 4 feet and a permanent offset of 0.4 feet.

RESPONSES IN STREAMS, SPRINGS, SEEPS, AND LAKES

Well 8N/10W-1Q1 in the western Mojave Desert, California responded to several southern California earthquakes.

Surface-water responses to earthquakes include changes in chemistry, seiches (wave oscillations) in lakes and other open water bodies, increases in stream, spring, and seep discharge, some instances of springs going dry or the appearance of new springs, and a very few examples of decreases in stream discharge. After the Denali Fault earthquake, many eyewitnesses throughout the contiguous United States reported water "sloshing" back and forth in lakes, bayous, ponds, and pools. Seiches lasted as long as half an hour, resulting in broken moorings in Lake Pontchartrain, Louisiana (Seth Moran, U.S. Geological Survey, written commun., 2002).

Streamflow may continue to increase for a few days and then gradually decline toward the pre-earthquake baseline condition for several months. Water flowing in a stream can come from overland flow or from ground water discharging into the streambed (baseflow). An abrupt increase in streamflow without precipitation or upstream reservoir release indicates increased contributions from the ground-water system. A coseismic (at the time

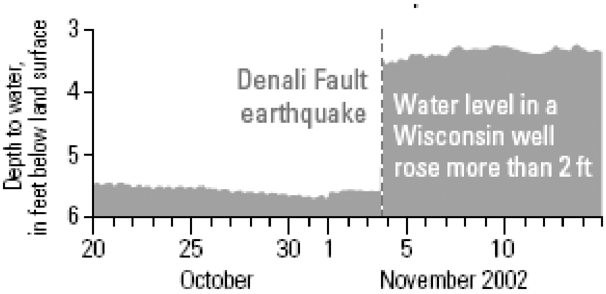

the seismic wave train arrives) discharge increase of approximately 40 percent was recorded in Santa Paula Creek, California, 35 miles WNW., of the 1994 M6.7 Northridge, California earthquake epicenter (3), and increased discharges were measured at numerous sites in the California Coast Range in response to the 1989 M7.1 Loma Prieta, California earthquake (4). The appearance of new seeps and springs and turbidity in streams also are common. Following the Northridge earthquake, a new oil seep began flowing in the Northern Ojai Valley (Stephen P. Mulqueen, Department of Conservation, Division of Oil, Gas, and Geothermal Resources, oral commun., 2003).

PRECURSORY RESPONSES

Many preseismic or precursory hydrogeologic responses have been reported, but in most cases, documentation has not been sufficient to convince the scientific community of their predictive value or to rule out explanations unrelated to earthquakes. Some investigators believe earthquakes cannot be predicted (5), whereas others believe that while earthquake prediction is difficult because of the youthfulness of seismic research and the limited number of case studies to investigate, prediction will not forever be impossible (6). Roeloffs (7) states "The documentation of a few of these reports [of precursory hydrogeologic changes] is now approaching levels that require them to be given scientific credibility." Water levels and chemical data seem to be the best documented precursory signals to date, and are monitored in several countries as potential indicators of an impending earthquake.

Japan has some of the best documented occurrences of precursory hydrogeological changes, preceding both the 1978 M6.8 Izu-Oshima earthquake (8,9), and the more recent, disastrous 1995 M7.2 Kobe earthquake (10). Before either or both of these events, water-level changes occurred in many wells, and there appeared to have been precursory changes in ground-water temperature, radon, chloride, and sulfate concentrations, as well as dissolved gas ratios in mineral springs (11,12). The preseismic changes in chloride and sulfate content of the Kobe event were reconstructed by analyzing dated, bottled ground water that is distributed in the domestic market.

The well-documented and similar hydrogeologic precursors at Izu-Oshima and Kobe have motivated the Japanese government to continue water-level monitoring at more than 30 wells in seismically active areas (13). The suggestion of precursory hydrogeologic changes is tantalizing, because the ultimate goal of most earthquake research is to reduce the hazard posed by major events, and providing some warning of impending activity would be very useful in that regard.

HAZARDS RELATED TO HYDROGEOLOGIC RESPONSES

Hydrogeologic responses to earthquakes can create a variety of hazards. Water supply may be disrupted if wells go dry or become too turbid to pump, and infrastructure damage may result from ground motion. After the 1992 M7.3 Landers, California, earthquake, bubbles of gas (carbon dioxide-enriched air) appeared in water from several San Bernardino, California, supply wells, clogging and disabling the filtration system (14). Hydrogeologic responses that cause wells to flow or streamflow to increase may cause flooding, landslides, erosion, and other related problems. Water quality also may be adversely affected

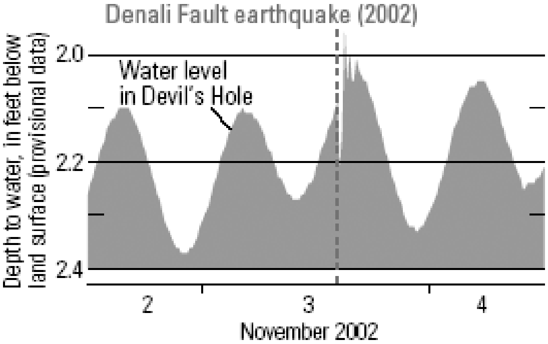

This pristine spring in the unique environment of Ash Meadows in Nevada is hydrologically connected to Devil's Hole, which is a nearly vertical cavern below the water table in southern Nevada. An endangered species, Devil's Hole pupfish, has lived in Devil's Hole pool for thousands of years. The water level in Devil's Hole oscillated in response to the Denali Fault earthquake, which disrupted the spawning areas of the pupfish near the surface of the pool.

by increased discharge, turbidity, or hydrogeochemical changes. Following the Landers earthquake, a pre-existing oil and natural gas seep in Tapo Canyon became more active, eventually flowing into and polluting the Santa Clara River (14).

Water-quantity and quality changes following an earthquake can adversely affect sensitive ecosystems, harming the plants and animals that live there. The endangered Devil's Hole pupfish (Cyprinodon diabolis) has lived for thousands of years in only one pool in Ash Meadows, Nevada, spawning on a single 6 × 13 foot shelf just below the water surface. In 1978, an earthquake in Mexico caused such violent water-level oscillations in Devil's Hole that algae was scoured from the shallow shelf (15), and the 2002 Denali Fault earthquake also apparently caused oscillations of 5–6 feet (Jim Roche, National Park Service, written commun., 2002). A persistent drop of water level below the shelf could extinguish the pupfish, whose life span is only a few years. Although court orders and careful management have led to population stability, the existence of the Devil's Hole pupfish remains precarious.

Ground-water systems are mechanically coupled to the rocks and sediments in which they exist. In addition to hydrogeologic responses to earthquakes, hydrogeologic changes may cause earthquakes or volcanic events. Earthquakes can be induced by the filling of surface reservoirs, or by annual or shorter-term fluctuations in reservoir levels, as is the case in most shallow earthquakes in the Aswan, Egypt area (16). Earthquakes also can be induced by the injection or withdrawal of fluids through wells, as was illustrated by the earthquakes caused by injection of waste fluid from munitions production at the Rocky Mountain Arsenal in the 1960s (17). More speculatively, the observation that large earthquakes can cause distant hydrogeologic changes may help explain how the 1992 M7.3 Landers, California earthquake triggered microearthquakes at a number of locations many hundreds of miles from the epicenter (18).

MECHANISMS FOR HYDROGEOLOGIC RESPONSES TO EARTHQUAKES

Known hydrogeologic responses most often occur as the seismic wave train arrives (coseismic) or sometime thereafter (postseismic); evidence of precursory (preseismic) hydrogeologic changes is becoming more compelling, but additional scientific investigation is needed to explain these phenomena.

Water levels in wells respond to the seismic-wave induced expansion and contraction of the aquifer tapped by the well, in turn causing step or oscillatory fluid-pressure changes. If flow in the well-aquifer system is fast enough, then these pressure changes cause flow into and out of the well, which sets up resonant motion of the water column, and the result is the seldom-recorded water-level oscillation. The more commonly recorded water-level offsets have been measured both near and far from earthquake epicenters. Offsets are expected in the 'near field', or static strain field, of an earthquake because the earthquake subjects the earth's crust, including its aquifer systems, to stress and permanent strain (deformation). This deformation process results in altered fluid pressure within the aquifer systems, and consequently, a steplike change in water level would be expected (3,14). The earthquake-induced water-level decline in well 9N/10 W-36J1 is the expected response for an extensionally strained region (14). Providing that the water-level response to strain is known, this relation allows scientists to use the earthquake response of water levels in favorably located and designed wells to measure strain, and also to calculate otherwise difficult-to-measure hydrologic properties of the aquifer. At distances of hundreds or thousands of miles, the offset can often be directly related to the amount of deformation produced by movement on the earthquake fault (19).

Possible mechanisms for responses are variable. For upward steps in shallow wells, compaction of overlying alluvium such as occurs during liquefaction may be the mechanism producing the offset (19). Fluid-pressure declines also can be caused by the escape of small amounts of exsolved gas from pore spaces in response to seismic waves. In a fractured rock environment such as northwestern Pennsylvania where the Pymatuning earthquake occurred, permeability may be altered due to the unclogging, widening, or narrowing of a fracture, or the creation of new fractures (2). Similarly, an increase in co- and post-seismic ground-water discharge at the surface through a spring, seep, or stream, can be caused by an increase in vent or fracture conductance (20), or an increase in the subsurface fluid pressure or permeability in the formation (4,14). The earthquake-induced increased discharge in Santa Paula spring is the expected response for a contractionally strained region (14). For a brief review of streamflow and water well responses to earthquakes and discussion of potential mechanisms see Montgomery and Manga (21).

RECORDING ACCURATE HYDROGEOLOGIC RESPONSES

Measurement techniques often do not adequately reflect hydrogeologic responses to earthquakes. In the recent past, water levels and discharge were recorded continuously using a drum recorder, similar to the method in which earthquakes are recorded on a seismograph. More recently, scientists have migrated to more cost-effective digital measurements that are recorded on an electronic datalogger. Ground-water levels typically change slowly, and dataloggers have limited storage capacity, such that data are recorded at 15-minute intervals, or more often, at 30- or 60-minute intervals. With a recording interval on this order, a water-level response from an earthquake may not be adequately defined by the periodic data; the recovery to the pre-earthquake water level can occur so

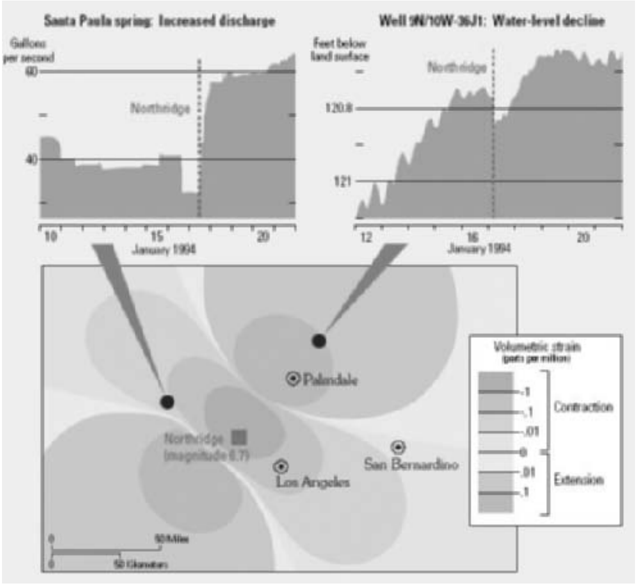

quickly that no change will be detected. In some cases, the water-level response is not in the expected direction because the initial response was not recorded, or because the response is affected by the particular well construction (the well-bore storage effect). Although stream or spring discharge may change as a result of a change in the ground-water contribution to the baseflow of the stream, surface-water discharge measurements often are not accurate enough to reliably detect small changes in flow. Additionally, flow is often controlled artificially, either partially or completely, preventing the detection of earthquake-induced changes.

CONCLUSIONS

Hydrogeologic responses to large distant earthquakes have important scientific implications with regard to our earth's intricate plumbing system. The exact mechanism linking hydrogeologic changes and earthquakes is not fully understood, but monitoring these changes improves our insights into the responsible mechanisms, and may improve our frustratingly imprecise ability to forecast the timing, magnitude, and impact of earthquakes.

BIBLIOGRAPHY

1. Vorhis, Robert C., 1966, Hydrologic effects of the earthquake of March 27, 1964 outside Alaska: U.S. Geological Survey Professional Paper 544-C, 54 p.
2. Fleeger, Gary M., Goode, D. J., Buckwalter, T.F., and Risser, D.W., 1999, Hydrologic effects of the Pymatuning Earthquake of September 25, 1998, in Northwestern Pennsylvania: U.S. Geological Survey Water-Resources Investigations Report 99–4170, 8 p.
3. Quilty, E.G., Farrar, C.D., Galloway, D.L., Hamlin, S.N., Laczniak, R.J., Roeloffs, E.A., Sorey, M.L., and Woodcock, D.E., 1995, Hydrologic effects associated with the January 17, 1994 Northridge, California, earthquake: U.S. Geological Survey Open-File Report 95–813, 47 p.
4. Rojstaczer, S., and Wolf, S., 1992, Permeability changes associated with large earthquakes: An example from Loma Prieta, California, 10/17/89: Geology, 20, p. 211–214.
5. Geller, R.J., Jackson, D.D., Kagan, Y.Y., and Mulargia, F., 1997, Earthquakes cannot be predicted: Science, 275, p. 1616–1617.
6. Wyss, Max, 2001, Why is earthquake prediction research not progressing faster?: Tectonophysics, 338, p. 217–223.
7. Roeloffs, E., 1996, Poroelastic techniques in the study of earthquake-related hydrologic phenomena: Advances in Geophysics, 37, p. 135–195.
8. Oki, Y., and Hiraga, S., 1979, Water level anomalies prior to the Izu-Oshima earthquake, January 14, 1978, Translated from the Japanese by E. Silberman, in Technical Report 79–2, Lamar-Merifield Geologists, Special Technical Report for the U.S. Geological Survey.
9. Groundwater-level monitoring for prediction of earthquakes: Pure and Applied Geophysics 126, 1987.
10. Tsunogai, U., and Wakita, H., 1995, Precursory chemical changes in ground water: Kobe earthquake, Japan: Science, 269, p. 61–63.
11. Igarashi, G., Saeki, S., Takahata, N., Sumikawa, K., Tasaka, S., Sasaki, Y., Takahashi, M., and Sano, Y., 1995, Groundwater radon anomaly before the Kobe earthquake in Japan: Science, 269, p. 60–61.
12. Sugisaki, R., Ito, T., Nagamine, K., and Kawabe, I., 1996, Gas geochemical changes at mineral springs associated with the 1995 southern Hyogo earthquake (M=7.2), Japan: Earth and Planetary Science Letters, 139, p. 239–249.
13. Tsukuda, E., Koizumi, N., and Yasuto, K., 2000, Integrated groundwater monitoring studies in the intensified observation area and near major active faults, Bulletin of the Geological Survey of Japan, v. 51, no. 9, p. 435–445.
14. Roeloffs, E., Danskin, W.R., Farrar, C.D., Galloway, D.L., Hamlin, S.N., Quilty, E.G., Quinn, H.M., Schaefer, D.H., Sorey, M.L., and Woodcock, D.E., 1995, Hydrologic effects associated with the June 28, 1992 Landers, California, earthquake sequence: U.S. Geological Survey Open-File Report 95–42, 68 p.
15. Death-Valley.us, Devil's hole pupfish—*Cyprinodon diabolis*: Death-Valley.us Forums, article 107, posted July 18, 2002, accessed January 29, 2003, at http://www.death-valley.us/article107.html
16. Awad, M., and Mizoue, M., 1995, Earthquake activity in the Aswan region: Pure and Applied Geophysics, 145, p. 69–86.
17. Healy, J.H., Rubey, W.W., Griggs, D.T., and Raleigh, C.B., 1968, The Denver earthquakes: Science 161, p. 1,301–1,310.
18. Hill, D.P., and others, 1993, Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake: Science, v. 260, no. 5114, p. 1617–1623.
19. Roeloffs, E., 1998, Persistent water level changes in a well near Parkfield, California, due to local and distant earthquakes: Journal of Geophysical Research, v. 103 (B1), p. 869–889.
20. Muir-Wood, R., and King, G.C.P., 1993, Hydrological signatures of earthquake strain: Journal of Geophysical Research, 98, p. 22,035–22,068.
21. Montgomery, D.R., and Manga, M., 2003, Streamflow and water well responses to earthquakes: Science, v. 300, p. 2047–2049.
22. Roeloffs, E., 1988, Hydrologic precursors to earthquakes: A review: Pure and Applied Geophysics 126, p. 177–209.

IN SITU ELECTROKINETIC TREATMENT OF MtBE, BENZENE, AND CHLORINATED SOLVENTS

Walter W. Loo
Environmental & Technology Services
Oakland, California

INTRODUCTION

The practice of soil and groundwater cleanup should follow the intent of the U.S. EPA laws and regulations such as the Safe Drinking Water Act (SDWA), the Resources Conservation and Recovery Act (RCRA) and the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA). Any environmental treatment process must be able to, or must be engineered to protect and conserve our resources (air, water, soil, and human) and not create undesirable side effects. Many of the conventional waste treatment technologies cannot meet these objectives and simply involve transferring the waste from one location or phase to another. This is true for waste excavation, landfill disposal, incineration, pump and treat, thermal desorption, and soil vapor extraction. Sound waste treatment technologies must be able to transform toxic and hazardous chemicals into harmless chemicals or render them harmless with the least energy consumption to achieve a good economic balance. To the extent possible, all treated soil and groundwater should be recycled for beneficial use. Most organic chemicals can be oxidized to harmless carbon dioxide, water, and inorganic chloride with existing and available innovative waste treatment technologies. The most difficult media for most waste treatment processes are relatively impermeable clayey media and fracture media. These media often absorb/adsorb the toxic and hazardous wastes spilled below the ground surface. Often, more than 90% of the waste spilled is attracted to these media. It is most fortunate that electrokinetic treatment processes can overcome the difficulties of relatively impermeable media and achieve or enhance the oxidation of many toxic and hazardous wastes at relatively low energy consumption.

The following sections discuss the *in situ* treatment processes, tests and design parameters, achievable cleanup levels, and case histories of successful treatment.

ENGINEERED ELECTROBIOCHEMICAL PROCESSES

Most petroleum hydrocarbons, including gasoline, diesel, jet fuel, kerosene, aviation fuel, motor oil, benzene, toluene, ethylbenzene, and xylenes are biodegradable. There has been some concern over the bioremediation of MtBE. However, electrokinetically enhanced cometabolic biotreatment of MtBE has been demonstrated in the field with repeated success (1).

Most chlorinated solvents are biodegradable through the application of the cometabolic biotreatment processes. However, because of the earlier publication of laboratory successes by the U.S. EPA in the 1980s, many scientists and engineers began to apply flammable methane and butane blindly into the vadose zone without questioning the side effects for health and safety (2–4). The same applies to the addition of toxic phenol, dichlorobenzene, and toluene into the groundwater as cosubstrates for bioremediation, which violates the underground injection control regulation established by the U.S. EPA. However, electrokinetically enhanced cometabolic biotreatment of chlorinated solvents using glucose sugar has been demonstrated in the field with repeated success (1).

Biochemical Processes

Bioremediation can be defined as the use of naturally occurring microorganisms to detoxify hazardous wastes (1), which is particularly true for hazardous organic wastes. The basic mechanism of bioremediation is promotion of the growth of microorganisms that are adapted to the organic wastes spilled in soil and groundwater. Stimulation of the growth of the microorganisms is controlled by proper temperature, oxygen, moisture, nutrients, and distribution of such in the impacted media. During growth, the microorganism will secrete enzymes and biosurfactants to break down the hazardous organic molecules (detoxification) and make them available for cellar absorption. Subsequently, the microorganism will grow and multiply, and the process will repeat itself. The end result is mineralization of the carbon source (from spilled hazardous organic wastes) aerobically into carbon dioxide, water, and biomass without undesirable side effects.

The application of bioremediation in hazardous waste management began in the early 1980s. The basic principles of biotreatment in the environment are described by Gaudy and Gaudy (5), who emphasized the enhancement factors, or engineering, in applying the bioremediation technology in the field. The biotreatment of excavated soil in engineered biopiles began to gain recognition in the late 1980s. In the 1990s, engineered bioremediation took such forms as soil bioventing, *in situ* aqueous phase biotreatment of soil and groundwater, and bioreactors to treat water, sludge, and organic vapor. In the mid-1990s, the evaluation of intrinsic or natural attenuative bioremediation prior to an engineering feasibility study became popular. The application of cometabolic bioremediaion treatment processes also flourished for chlorinated solvents to petroleum hydrocarbons.

The bioremediation professional must follow the basic fundamental science and engineering to complete a bioremediation project successfully. The following sections cover the science and engineering of bioremediation processes, microbial ecology, bioenhancement, soil microbes and pathogens, the aerobic process, the anaerobic processes, and the cometabolic process.

To carry out a bioremediation process properly, the living environment favorable for the microorganism must be understood. Microorganisms used in bioremediation are generally classified as mesophiles or the warm temperature type. The optimum range of temperature in which they thrive is from 80–100 °F. When the temperature of the bioremediation system drops below 40 °F, the microorganism goes into hibernation. Microorganisms must also have optimal moisture to grow and to survive. Dry conditions are not favorable for bacterial growth. All microorganisms require an available carbon source and nutrients for growth. An adapted indigenous microorganism in a hazardous waste environment relies on the carbons in organic chemicals as a food source (substrate). The nutrient requirements include sources of phosphorus, nitrogen, and iron. The microorganism also relies on the availability of oxygen in the environment to mineralize the hazardous chemicals into carbon dioxide and water under aerobic conditions.

Bioenhancement Methods. It is evident that when hazardous wastes spill into soil and groundwater, it is not an optimized microbial ecology for bioremediation. Therefore, we must engineer a favorable microbial ecology for optimal bioremediation to take place. The following sections cover an optimal microbial ecology that can be engineered or enhanced for all situations.

Oxygen Enhancement. If oxygen is not available in soil (vadose soil) or in groundwater (aquifer), this anaerobic condition is prone to create hydrogen sulfide (toxic gas) and methane gases (flammable gas). The quantity of oxygen required must satisfy the chemical oxygen demand (COD) of the soil and groundwater media. It is important to point out that in fully saturated media, both the COD of the soil and the groundwater must be satisfied. The objective is to supply enough oxygen to satisfy the COD of the impacted media. Oxygen can be enhanced by injecting air or oxygen into the pore space of the vadose soil. Dissolved oxygen in groundwater can be enhanced by adding oxidants, by *in situ* electrolytic (electric current) generation of water in the impacted media, or by creating of hydroxyl radicals (a strong oxidant) at the initial stage of the electrolysis of water. Another source of oxygen is through introducing carbohydrates, which will biodegrade and provide enough oxygen to mineralize the hazardous organic compound into carbon dioxide, water, and biomass.

Heat Enhancement. To achieve an optimal growth temperature (80–100 °F) for mesophilic microorganisms, the vadose zone soil can be heated by circulating warm air into the pore space in the soil. It may not be necessary to heat up the entire solid matrix of the mass because most of the hazardous waste tends to adhere to the surface of the solid matrix. Passive solar panel air heating can provide an economical means of heat delivery. Recirculation of heated air in the subsurface will also save energy cost because the soil mass itself acts like a heat sink that has very good insulating value. The same applies to heating groundwater by a passive solar water heater. The heating requirements (in BTUs) must be estimated based on the weight of the impacted media. The weight of the impacted media can be calculated as the volume of the soil or groundwater plumes. A swimming pool passive solar heating panel can supply more than adequate heating value in less than 30 days for most bioremediation projects.

Nutrient Enhancement. Sometimes, there are insufficient nutrients in the form of nitrogen, phosphorus, and iron in the soil and groundwater plumes. Nutrient enhancement applies only when there is a deficiency. For proper bioremediation, a carbon (C) to nitrogen (N) to phosphorus (P) weight ratio of 100:10:1 should be followed. The weight of carbon in the plume is determined by the total organic carbon (TOC) concentration times the weight of the plume that requires cleanup.

Moisture Enhancement. If warm air (80° to 100 °F) is being recirculated within the soil plume for soil remediation, the moisture in the warm air (humidity) should be adequate enhancement. No need exists for moisture enhancement in saturated media.

Nutrient Distribution. The most difficult media for hazardous waste treatment processes to work in are a relatively low-permeability clayey medium and a fracture medium. These media often absorb/adsorb the toxic and hazardous waste spilled below the ground surface. Often, more than 90% of the waste spilled is attracted to these media. It is most fortunate that the electrokinetic treatment process (electro-osmosis and electromigration of water) can overcome the difficulties of relatively low-permeability media, achieve or enhance the aerobic biochemical oxidation of the human toxic and hazardous wastes, and consume relatively little energy (see ELECTROKINETIC PROCESSES).

CoMetabolic Processes With Glucose and Sugar. The basic mechanism of bioremediation is triggering the growth of microorganisms that are adapted to organic wastes spilled in the soil and groundwater. Stimulation of the growth of the microorganism is controlled by proper temperature, oxygen, moisture, nutrients, and cosubstrates, and distribution of such in the impacted media. During growth, the microorganism will secrete enzymes and biosurfactants to break down the hazardous organic molecules (detoxification) and to make them available through cellular absorption. Subsequently, the microorganism will grow and multiply, and the process repeats until the food source is depleted. The end result is mineralization of the carbon source (from spilled hazardous organic wastes) aerobically into carbon dioxide, water, and biomass without undesirable side effects. The cometabolic processes for bioremediation of chlorinated solvents and petroleum hydrocarbons are the most promising trouble-free biotreatment processes and are very efficient. The cometabolic process can be defined as the introduction of an easily biodegradable substrate into the environment, which triggers the secretion of enzymes from microorganisms that are adapted to the spilled hazardous organic waste.

The biotreatment of chlorinated solvents is a highly sought after solution to widespread soil and groundwater contamination problems. However, most of the knowledge of biotreatment of chlorinated solvents is found only in research laboratories. A successful laboratory demonstration of the cometabolic biotreatment of trichloroethene (TCE) by methanotrophic bacteria columns was achieved by the EPA Ada Laboratory in 1985 (6). In 1987, the EPA Gulf Breeze Laboratory successfully demonstrated the cometabolic biotreatment of TCE by *Pseudomonas putida* through an aromatic pathway (7). In 1989, Loo, (8) successfully demonstrated the first field closure of the cometabolic biotreatment of TCE and trichloroethane (TCA) together with toluene in soil through heat and nutrient enhancement by the growth of indigenous bacteria *Bacilli* and *Pseudomonas fluorescens*. In 1991, Stanford University demonstrated partial success of the cometabolic biotreatment of chlorinated solvents in groundwater by methanotrophic bacteria at Moffet Field, California. Loo et al. (9) demonstrated the bio-treatment of TCE in the laboratory, field pilot, and in field application using glucose, which is nontoxic and nonhazardous, as a co-substrate. Loo et al. (9) developed and successfully applied the cometabolic biotreatment of petroleum hydrocarbons using glucose and sucrose as a co-substrate in both soil and groundwater Subsequently, many other site closures were achieved through the application of the cometabolic bioremediation of petroleum hydrocarbons.

Electrokinetic Processes

Direct current (DC) electricity is a widely used technology in industrial, commercial, and military applications. Direct current from solar panels is also used in space age technology for the electrolysis of water to produce hydrogen as fuel and oxygen for breathing. For environmental protection, direct current is used in the cathodic protection of steel underground storage tanks and pipelines from corrosion and leaks. Electrokinetic (EK) processes involve the application of DC electricity to soil and groundwater in the subsurface (1). EK processes originated or based their work on the geotechnical dewatering of clayey material by Casagrande (10–12). Other significant pioneering EK applications in oil and gas recovery were developed by Chilingar et al. (13–16).

The application of EK technology in the environmental field was first reported in the literature in the 1980s (17), most of the applications were for the isolation or recovery of metals. Van Doren and Bruell (18) first reported that the EK process destroyed benzene in wet clay in a laboratory bench scale test. Loo (8) and Loo and Wang (33) first reported the successful commercial application of the EK process as the primary enhancement process for removing or desorbing chlorinated solvents from a thick clayey soil layer at a defense contractor site closure located in Anaheim, California. The U.S. Army Corp of Engineers (19) listed EK treatment as a viable remedial process for treating hazardous wastes. In 1995 at the HAZMACON Conference, Santa Clara, California, the best paper award was given to the paper entitled “Elecrokinetic Treatment of Hazardous Wastes in Soil and Groundwater” (20). In 1995, the U.S. EPA (21) summarized the application and development of electrokinetic (EK) treatment processes by various private companies, the U.S. Department of Defense, the U.S. Department of Energy, and various universities as a potential cost-effective treatment for hazardous wastes. In 1997, the U.S. EPA (17) summarized and updated various EK applications in research, development, and commercial treatment of hazardous and radioactive wastes.

The EK processes works effectively for both permeable and relatively impermeable media. For relatively impermeable porous media like clay and silt, direct current conducts well in the media. The electrical conductivity of clay is often 1000 times greater than that of sand, sandstone, limestone, and igneous and metamorphic rocks. The same applies to fractured media where the fractures are often filled with highly conductive clayey and fine grain minerals. For porous media like sand, the flow of electricity is most likely through water that has naturally occurring electrolytic ions such as chloride, bicarbonate, nitrate, potassium, and sodium. The water molecules and positively charged ions (cations) will also be “dragged” by the electron flow from the anode toward the cathode and create a hydraulic mound (high) around the cathode by the induced electrokinetic gradient (Fig. 1). The EK processes works for both the vadose zone and the saturated zone in the subsurface, as long as there is adequate moisture in the medium. A dry medium does not conduct electricity very well.

As mentioned in the previous section, electrolysis of the water molecule will produce dissolved hydrogen around the anode and dissolved oxygen around the cathode. Therefore, higher dissolved oxygen around the cathode will help and encourage aerobic biotreatment of organic compounds. Electromigration will also help to bring the desorbed organic compounds to places enriched with oxygen and nutrients.

Regulatory And Permitting Requirements. No regulation or permitting requirement apply directly to the use of direct current in the subsurface. When the EK process is applied as an enhancement process for bioventing in the vadose zone, the extracted volatile organic vapor can be recirculated back into the center of the soil plume to avoid an air emission permit.

The EK process can make the groundwater migrate back and forth within the groundwater (thorough *in situ* mixing), so there is no need to pump and treat and thus no discharge permit. Also, when electrolytes and nutrients are added into the subsurface for EK processes, no underground injection control permit is required as long as all electrolytes and nutrients added are nontoxic and nonhazardous.

EK application in the subsurface is similar to cathodic protection with an impressed current and to the function of electrical grounding devices, which require no permit to construct and operate.

IN SITU FIELD FEASIBILITY CHARACTERIZATION

Before implementing a full-scale bioremediation project, one should conduct a field pilot confirmation test for mineralization products. The purpose of the field pilot test is to determine if the indigenous microorganism responds to the various enhancement elements. A simple oxygen slug injection into a vadose zone well while monitoring the respiration of the metabolic function of the indigenous microorganism will suffice. The simple oxygen enhancement field test should show a favorable increase in carbon dioxide and a reduction in oxygen in the well.

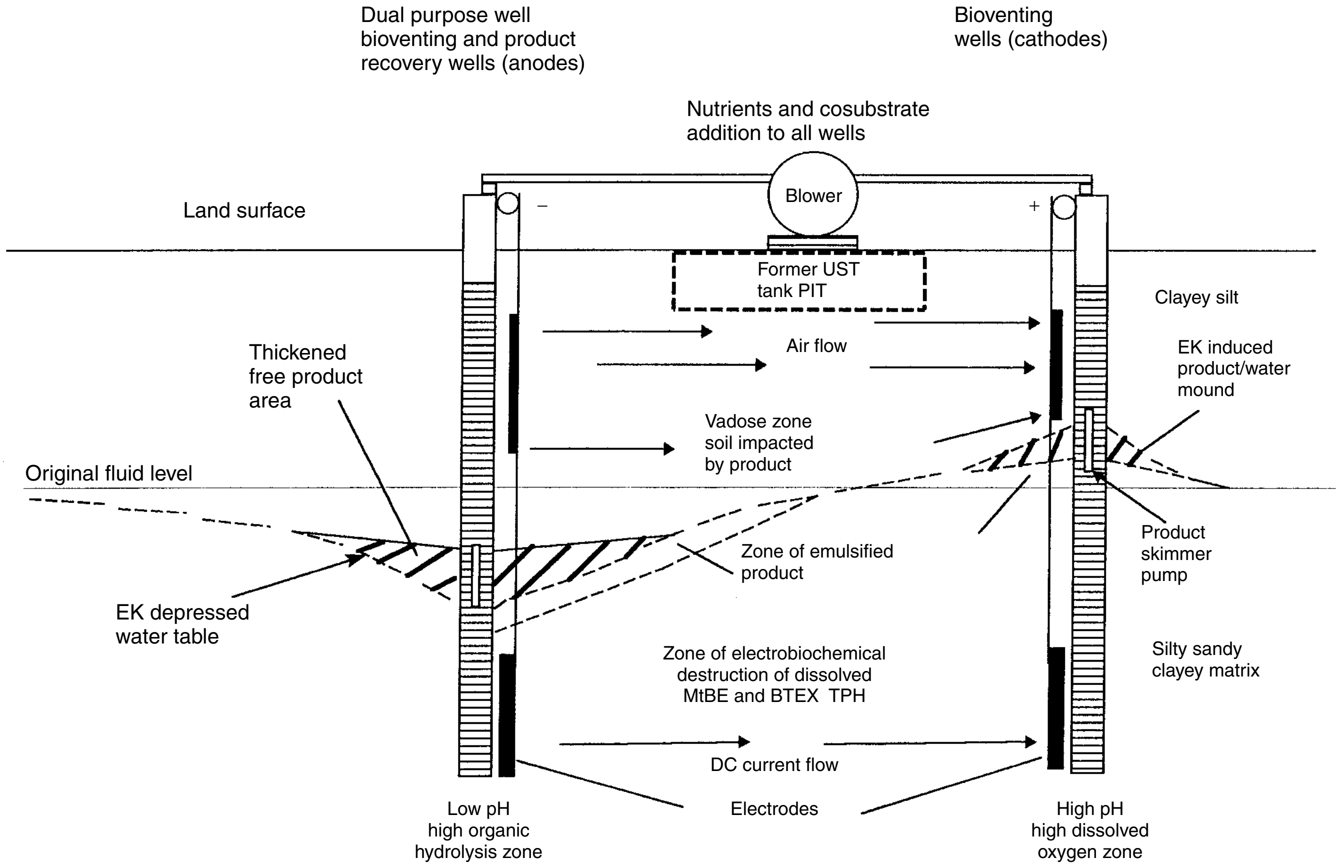

Figure 1. Generalized *in situ* electrobiochemical remedial cross section.

This whole field test involves only one vadose well in the middle of the impacted soil plume, one compressed oxygen bottle, and an oxygen/carbon dioxide field.

Biochemical Field Survey

The initial distribution of the organic hazardous waste in the soil and groundwater is the starting point for the design of a bioremediation project. For the baseline condition of a soil plume, a statistically significant soil sampling plan will be required because of the heterogeneous nature of the distribution of organic hazardous wastes in the soil. The plan should follow the requirements of the EPA SW-846 sampling procedure for the lateral extent of a soil plume to ensure statistical significance. For the vertical extent, one should add randomness by staggering depths instead of using a regular depth interval to ensure even distribution and to eliminate any significant lithologic layering effects. It is fairly easy to establish a baseline for dissolved constituents in the groundwater because of the slow dilution resulting from the influx of clean groundwater from upgradient. The only requirement is an adequate number of monitoring wells. The key sample analysis parameters required for bioremediation design and engineering are as follows:

weight of the soil and groundwater plumes
concentration of wastes in soil and water
total heterotroph plate count
chemical oxygen demand (COD)
total organic carbons (TOC)
total Kjeldahl nitrogen (TKN)
nitrate
total phosphorus
iron
temperature within the soil and groundwater plume
dissolved oxygen and pH in groundwater
moisture content in soil
concentration of oxygen and carbon dioxide in soil pores

When the soil and groundwater plumes are fully characterized in weight, chemical requirement, engineering, installation, equipment, delivering lines, monitoring, and analysis, operating and maintenance costs can then be estimated for economic feasibility. The ultimate cost should be presented in dollars per ton of soil or groundwater plume that requires biotreatment.

Electrokinetic Field Surveys

The main purpose of the *in situ* direct current survey is to detect any anisotropic (uneven and bias) flow directions. The survey also may provide clues to unidentified buried metallic lines and structures that may require cathodic protection. Field DC flow should be conducted with combinations in pairs of all available wells located inside and outside of the plume. For example, for four wells there should be six survey measurements. For six wells, there should be 15 survey measurements. The survey should be conducted at fixed voltage (electrical potential) and measure the direct current flow in amperes. The DC flow survey requires 110 volt AC power with a DC converter rated for 50 to 100 DC volts (range) and a maximum of 10 DC amperes. Wires must be insulated to prevent short circuits.

IN SITU ELECTROBIOCHEMICAL TREATMENT DESIGN

The economical alternative to digging, hauling, and disposing or *ex situ* treatment of contaminated soil is bioventing in the subsurface. The economical alternative to pumping, treating, and discharging of contaminated water is *in situ* groundwater biotreatment.

Soil Bioventing

Bioventing is a method of treating petroleum hydrocarbon contaminated soil in the vadose zone by circulating air through the pores of the soil matrix. It is particularly effective when the moisture content and nutrient level in the soil can be amended to prime enhancement conditions for bio treatment. Bioventing can treat both volatile and nonvolatile organic hydrocarbons. Hinchee et al. (22,23) and Miller et al. (24) developed the bioventing concept for bioremediating petroleum hydrocarbons at many U.S. Air Force sites. However, only oxygen enhancement was adopted for soil cleanup at many of these sites. The process involved injecting air at one well into the contaminated soil plume, with little evaluation of the engineering feasibility. Little engineering consideration was given to the migration of organic vapor that was pushed out of the soil plume and formed condensates, which further spreads the groundwater contamination. This may be the most economic means of treating petroleum hydrocarbons in soil for the U.S. Air Force, but it may face the side effects of uncontrolled spreading of the dissolved groundwater plume.

The first fully enhanced bioventing system for treating gasoline was applied in the clayey marine conglomerate at the PepBoys site, San Diego, California (25). The electroosmosis process desorbed or squeezed out the gasoline in the clayey soil and made the gasoline available for bioventing. The extracted air was heated and reinjected into the vadose zone soil for heat enhanced biotreatment. The heated air also carried the water squeezed from the clay layer as humidity for moisture enhancement. Pure oxygen gas was injected into the middle of the soil plume to replenish the oxygen used during aerobic bioventing. Nitrogen-based nutrient gas was also injected into the middle of the soil plume for nutrient enhancement. Nitrogen-based nutrient gas injection into soil is widely used in large-scale agricultural farming practices.

Groundwater Biotreatment

The aqueous phase biotreatment discussed in this section does not involve pumping out any water from the ground. The most challenging aspect of *in situ* aqueous biotreatment is the distribution or penetration of nutrient, oxygen, and heat into the clayey or less permeable areas within the saturated media. Electrokinetic processes can overcome these conditions.

Passive Biotreatment. Passive biotreatment involves oxygen, nutrient, heat, and electrokinetic enhancement to biodegrade toxic organic chemicals into harmless carbon dioxide, water, and biomass. Oxygen can be enhanced by creating dissolved oxygen through electrolysis of groundwater or by using the oxygen in food additive cosubstrates dissolved in the groundwater. Nutrients can be enhanced by adding ammonium-based food additive chemicals into the groundwater. Heat can be enhanced economically by infiltrating solar heated water into the groundwater through an infiltration gallery or through injection into a remedial well. Electrokinetic enhancement or mixing can be achieved by electromigration of nutrient/electrolyte through permeable and impermeable water saturated media.

Loo (26) demonstrated that benzene can be biotreated to a nondetectable level in groundwater. Loo (27) also demonstrated that MtBE can be electrobiochemically treated to a nondetectable level in groundwater. Finally, Loo et al. (28) and ETS showed that chlorinated solvents can be biotreated to nondetectable levels in groundwater.

The following sections provide the requirements for EK design and implementation. The basic electrokinetic feasibility test parameters are as follows:

field pH, conductivity, and dissolved oxygen of water
laboratory pH and conductivity of soil
chemical oxygen demand (COD)
contaminant concentration in soil above and below the water table
contaminant concentration in groundwater
selective anion and cation analysis

Representative samples must be obtained at the center of the plume where the concentration of the contamination is the highest. Representative samples must also be obtained at the peripheral monitoring wells upgradient, downgradient and cross gradient.

Electrokinetic System Components

Electrokinetic system components may include but are not limited to electrodes, electrolytes, wiring, and power supply. The description and selection of these components are described in the following sections.

An electrode can be made of any good conductor. All electrodes for environmental cleanup must be relatively inert and should not contain hazardous or toxic chemicals. Typically, copper and stainless steel electrodes should be avoided. The most common electrodes are made of iron, black steel, carbon, and graphite. More exotic electrodes used are made of gold, platinum, and titanium. The electrodes can be ordered in any shape and form to fit the particular application. The most common form is a round solid rod or a tube from 1 to 10 feet long.

The use of food additive chemicals as an electrolyte is preferred to increase the flow of direct current in the ground. Note that the electrolyte should be applied on an as needed basis. Soil and groundwater may contain enough electrolyte minerals to provide adequate direct current flow. The application of electrolyte sometimes is necessary for maintenance of electrodes to prevent corrosion at the anode and mineral scaling at the cathode. Potassium and sodium chlorides are not considered environmentally friendly electrolytes electrolysis of potassium and sodium chlorides will yield toxic chlorine gas, which may create unsafe and unhealthy conditions.

All electrical wiring must be insulated and rated for at least 30 amperes. The length of wires to cathodes and anodes need not be equal, but multiple wires to either cathode or anode must be equal in length and size to prevent short circuits. It is often very common for each circuit wire to carry less than 2 amperes of DC current.

Direct current power supplies often draw from 110 Volt AC or 220 Volt AC power sources. Same of the common names for the DC power converter are DC testing power supply, DC welding units, an automobile battery with a proper recharging source, and electrical solar panels. It is advised that no more than 10 DC amperes be drawn from any converters to prevent overheating.

ACHIEVABLE CLEANUP LEVELS

EK processes can achieve the very low levels of concentrations of various hazardous wastes in soil and groundwater listed here as examples:

	In Soil	In groundwater
Petroleum hydrocarbon (gasoline)	Less than 50 ppm	Nondetectable (less than 1 ppm)
Petroleum hydrocarbon (diesel kerosene, jet fuel, motor oil, etc.)	Less than 100 ppm	Nondetectable (less than 1 ppm)
Benzene	Less than 100 ppb	Nondetectable (less than 1 ppb)
Methyl tertiary-butyl ether (MtBE)	Less than 1 ppm	Nondetectable (less than 5 ppb)
Chlorinated solvents (PCE, TCE)	Less than 100 ppb	Nondetectable (less than 5 ppb)

ECONOMICS OF *INSITU* ELECTROBIOCHEMICAL TREATMENT

For a typical small-scale project, assuming less than 3,500 tons of contaminated soil and groundwater, the following is a breakdown of the in-situ electrobiochemical treatment project cost estimate in 2002 US dollars:

Planning and preliminary design	$3,000
Field feasibility testing	$20,000
Detailed engineering design	$5,000
Remedial system installation (six remedial wells)	$40,000
Startup remedial testing	$13,000
Operation and maintenance (6 months)	$18,000
Remedial monitoring	$6,000
Closure confirmation sampling	$5,000

The unit price remedial cost estimate is about $30 per ton. For medium size remedial projects of more than 10,000 tons, it is not unusual to complete the project at less than $20 per ton. For larger size remedial projects of 50,000 tons or more, it is not unusual to complete the project at less than $15 per ton.

The economics of power consumption can be measured in kilowatt hours (kwh) and each kwh costs US $0.10 to $0.20 or less than $2 per ton depending on the cost of power at the location. Each watt of energy is equal to the product of voltage (1 volt) and amperage (1 ampere). Normally, the cost of electricity is insignificant compared to the overall remedial project cost.

CASE HISTORIES

There are many documented examples of field pilot-scale EK remediation of toxic/radioactive metals and hazardous organic chemicals. The individual project descriptions are included in the publicly available reference documents listed below:

U.S. ARMY CORP OF ENGINEERS 1992
Installation Restoration and Hazardous Control Technologies

U.S. EPA 1995 PUBLICATION 542-K-94-007
In-Situ Remediation Technology: Electrokinetics

U.S. EPA 1997 PUBLICATION 402-R-97-006
Electrokinetic Laboratory and Field Processes Applicable to Radioactive and Hazardous Mixed Waste in Soil and Groundwater

Electrokinetic Enhanced Soil Bioventing

The conventional application of soil vapor extraction of volatile organic chemicals (VOCs) cannot overcome low permeability soil media to render the successful completion of a project. The electrokinetic process can enhance the desorption of water and contaminants from clay by breaking down the clay layer's electrically charged bonds. This will make the VOCs available for contact and treatment. The major drawback of soil vapor extraction is air emission which is a physical phase transfer process instead of destruction. The EK enhanced closed loop bioventing process was applied to destroy the VOCs without air emission at sites with clayey soil. The following is a list of sites where EK enhanced bioventing processes were successfully implemented:

PepBoys Site, San Diego, California (25)
Bioventing of gasoline in clayey soil

Bat Rentals Site, Las Vegas, Nevada (29)
Bioventing of gasoline, diesel and kerosene in clayey caliche soil

Evergreen Site, Los Angeles, California (30)
Bioventing of gasoline in clayey soil

DBM Oil Site, Long Beach, California (30)
Bioventing of diesel and waste oil in clayey soil

Cadillac Site, Northridge, California (27)
Bioventing of gasoline in clayey soil

Former Texaco Site, Long Beach, California (27)
Bioventing of gasoline in silty/clayey soil

Former ARCO Site, Monterey Park, California (27)
Bioventing of gasoline in clayey soil

Electrochemical Oxidation

It is difficult to distinguish which electrokinetic process is responsible for destroying petroleum hydrocarbons. The bench test on a soil sample from the PepBoys Site indicated that electrochemical oxidation is responsible for destroying petroleum hydrocarbons.

For all the bioventing projects described in the previous section, the electrochemical oxidation process is responsible for a portion of the destruction of the petroleum hydrocarbons rather than bioventing alone.

Electrobiochemical Oxidation

Electrobiochemical oxidation applies to the aqueous phase biotreatment of petroleum hydrocarbons in soil and groundwater. Electrokinetic processes are responsible for distributing nutrients and oxidants in impacted soil and groundwater. It can be used together with bioventing in the vadose zone. The following is a list of sites where EK enhanced aqueous phase biotreatment processes were successfully implemented:

Westland Site, Hayward, California (26)
Electrobiochemical of gasoline and diesel in Bay Mud clay and water

Bat Rentals Site, Las Vegas, Nevada (29)
Bioventing(wet) of gasoline, diesel and kerosene in clayey caliche soil and shallow groundwater

Cadillac Site, Northridge, California (27)
Bioventing(wet) of gasoline in clayey soil and groundwater

Former Texaco Site, Long Beach, California (27)
Bioventing(wet) of gasoline in silty/clayey soil and groundwater

Former ARCO Site, Monterey Park, California (27)
Bioventing(wet) of gasoline in clayey soil

At all these sites, food additive nutrients and oxidants were introduced as electrolytes by electrokinetically induced migration through wells (horizontal migration) and infiltration galleries (vertical migration) to penetrate into the clayey matrix of soil and groundwater. Additions of nontoxic and nonhazardous food additives, nutrients, and oxidants into the subsurface are not regulated by underground injection control regulations. Petroleum hydrocarbons like BTEX in soil and groundwater were treated to nondetectable levels at most sites.

Electrolysis of MtBE and Benzene

The two most resistant dissolved petroleum hydrocarbons in the environment are MtBE and benzene due to their higher solubility in water and low cleanup levels which most often cannot be achieved by conventional remedial treatment technologies.

The gasoline additive MtBE is highly soluble in water and usually migrates furthest away downgradient from the spill location. Air stripping of MtBE is not an effective treatment option. There are reports that MtBE in the air were solublized into air-stripper blowdown water. Plus, MtBE is not readily biodegradable in the subsurface. The following is a list of sites where EK induced electrolysis processes were successfully implemented:

Cadillac Site, Northridge, California (27)
Bioventing(wet) of gasoline in clayey soil and groundwater

Former Texaco Site, Long Beach, California (27)
Bioventing(wet) of gasoline in silty/clayey soil and groundwater

MtBE was first treated in groundwater, accidentally (not by design), in 1997 at the Cadillac Site, Northridge, California, while treating dissolved BTEX using the electrobiochemical aqueous phase treatment. It was discovered that MtBE soon “disappeared” after the start-up of the electrokinetic enhanced biotreatment. The MtBE in groundwater was treated to nondetectable levels in less than 3 months of treatment. It also appeared that benzene concentration was decreasing at a slower rate at various monitoring wells. The electrolysis of MtBE and benzene cannot actually be confirmed. It was a surprise to document the MtBE disappearance at the Cadillac Site while at the time the oil company sponsored numerous seminars that pointed to no effective remedy for MtBE in the environment.

The same electrobiochemical treatment technique was applied to MtBE and BTEX at the former Texaco Site in Long Beach, California, in 1997. The dissolved benzene and MtBE in groundwater were treated to nondetectable levels in two monitoring wells outside of the zone of influence of the electrobiochemical treatment area which was focused on the former underground storage tank pit area. This indicated that the only thing that can influence these peripheral monitoring wells is the DC flow field. This is the first confirmation of the electrolytic breakdown of dissolved MtBE and benzene in groundwater. The nondetectable MtBE and benzene performance in the vadose zone soil in the tank pit area is also a first.

Chlorinated Solvents

The conventional remedial treatment technology of chlorinated solvent in soil is soil vapor extraction. The newer and innovative remedial treatment of chlorinated solvents in groundwater is cometabolic biotreatment. None of the above will be effective when treating clayey material. Electrokinetic processes excel in desorbing the contaminants from clay and distributing nutrient, oxidants and cosubstrates in the soil and groundwater.

The following is a list of sites where the EK enhanced chlorinated solvent treatment processes were successfully implemented:

Northrop ESD Site, Anaheim, California (31)
EK enhanced soil vapor extraction of chlorinated solvents in clayey soil

The Good Guys Site, Emeryville, California (8,20,28, 30,32) EK enhanced cometabolic biotreatment of chlorinated solvents in clayey soil and groundwater

BIBLIOGRAPHY

1. Loo, W.W. (2001). Bioremediation & electrokinetic treatment of hazardous wastes. *Standard Handbook of Environmental Health, Science and Technology*. J.H. Lehr (Ed.). McGraw-Hill, New York, Chapters 14.4 & 14.6.
2. Roberts, P. et al. (1990). Biostimulation of Methanotrophic Bacteria to Transform Halogenated Alkene for Aquifer Restoration. *Conference Proceedings of Petroleum Hydrocarbons and Organic Chemicals in Groundwater*, Houston, TX, pp. 203–217.
3. Garnier, P., Auria, R., Magana, M., and Revah, S. (1999). Cometabolic Bio-treatment of MTBE by a Soil Consortium. *Proceedings of The Fifth International In situ and On-Site Bio-remediation Symposium*, Battelle Press, San Diego, CA, Vol. 5(3), pp. 31–35.
4. Battelle Press. (1999). Engineered Approaches for *In situ* Bio-remediation of Chlorinated Solvents. *Proceedings of The Fifth International In Situ and On-Site Bio-remediation Symposium*, San Diego, CA, Vol. 5(2).
5. Gaudy, A.F. and Gaudy, E.T. (1980). *Microbiology for Environmental Scientists and Engineers*. McGraw-Hill, New York.
6. Wilson, J.T. and Wilson, B.H. (1985). Biotransformation of trichloroethylene in soil. *Applied and Environmental Microbiology* **49**(1): 242–243.
7. Nelson, M.J.K. et al. (1987). Bio-treatment of trichloroethylene and involvement of an Aromatic Biodegradative pathway. *Applied and Environmental Microbiology* **53**(5): 949–954.
8. Loo, W. W. (1991). Heat Enhanced Bio-remediation of Chlorinated Solvents and Toluene in Soil. *Conference Proceedings of R&D 1991*, Anaheim, CA, pp. 133–136.
9. Loo, W.W. et al. (1993). Field Bio-treatment of Chlorinated Solvents: A Co-Metabolic Process Utilizing Glucose as Co-Substrate. *Proceedings of HAZMACON 1993*, Association of Bay Area Government(ABAG), Santa Clara, CA.
10. Casagrande, L. (1947). *The Application of Electro-Osmosis to Practical Problems, in Foundations and Earthworks*. Building Technical Paper No. 30, H. M. Stationary Office, London.
11. Casagrande, L. (1948). Electro-Osmosis. *Proc. 2nd Int. Conf. on Soil Mech. and Found. Eng.*, vol. 1, pp. 218–222.
12. Casagrande, L. (1952). Electro-Osmotic Stabilization of Soils. *Boston Soc. Civ. Eng.* **39**(1): 51.
13. Chilingar, G.V. et al. (1963). Possible use of electric current for increasing volumetric rate of flow of oil and water during primary and secondary recovery. *Chem. Chron.* **28**(1): 1–4.
14. Chilingar, G.V. et al. (1964). Use of direct electrical current for increasing the flow rate of reservoir fluids during petroleum recovery. *J. Canadian Pet. Tech.* **3**(1): 8–14.
15. Chilingar, G.V. et al. (1968). Possible use of direct current for augmenting reservoir energy during petroleum production. *Compass of Sigma Gamma Epsilon* **45**(4): 272–285.
16. Chilingar, G.V. et al. (1970). Effect of direct electrical current on permeability of sandstone cores. *J. Pet. Tech.* **22**(7): 830–836.
17. US EPA. (1997). *Electrokinetic Laboratory and Field Processes Applicable to Radioactive and Hazardous Mixed Waste in Soil and Groundwater*. US EPA publication 402-R-97-006, Center for Remediation Technology and Tools, Washington, DC.
18. Van Doren, E.P. and Bruell, C.J. (1987). Electro-Osmotic Removal of Benzene from a Water Saturated Clay. *Proc. of Petroleum Hydrocarbons and Organic Chemicals in Ground Water: Prevention, Detection and Restoration*, Nov. 17–19, 1987, Houston, TX, pp. 107–126.
19. U.S. Army Corps of Engineers. (1992). *Installation Restoration and Hazardous Control Technologies*. Prepared for US Air Force, US Navy, US Army, and US EPA, pp. 45–46.
20. Loo, W.W. (1995). Electrokinetic Treatment of Hazardous Wastes in Soil and Groundwater, HAZMACON'95 Award Winning Paper. *Proceedings of HAZMACON'95*, sponsored by the Association of Bay Area Governments, San Jose Convention Center, San Jose, CA, pp. 147–158.
21. US EPA. (1995). *In-Situ Remediation Technology: Electrokinetics*. US EPA publication 542-K-94-007, Office of Solid Waste and Emergency Response, Office of Technology Innovation Office, Washington, DC.
22. Hinchee, R.E. et al. (1989). Enhancing Bio-treatment of Petroleum Hydrocarbon Fuels Through Soil Venting. *Conference Proceedings of Petroleum Hydrocarbons and Organic Chemicals in Groundwater*, Houston, TX, pp. 235–248.
23. Hinchee, R.E. et al. (1990). Enhanced Bioreclamation Soil Venting and Groundwater Extraction: A Cost-Effectiveness and Feasibility Comparison. *Conference Proceedings of Petroleum Hydrocarbons and Organic Chemicals in Groundwater*, Houston, TX, pp. 147–164.
24. Miller, R.N. et al. (1990). A Field Investigation of Enhanced Petroleum Hydrocarbon Bio-treatment in the Vadose Zone at Tyndall AFB, Florida. *Conference Proceedings of Petroleum Hydrocarbons and Organic Chemicals in Groundwater*, Houston, TX, pp. 339–351.
25. Loo, W.W., Wang, I.S., and Fan, K.T. (1994). Electrokinetic Enhanced Bioventing of Gasoline in Clayey Soil: A Case

History. *Proceedings of the SUPERFUND XV Conference*, Washington, DC.

26. Loo, W.W. (1994). Electrokinetic Enhanced Passive In-Situ Co-Metabolic Bio-treatment of Gasoline and Diesel in Clayey Soil and Aquifer: A Case History. *Proceedings of the SUPERFUND XV Conference*, Washington, DC.
27. Loo, W.W. (1998). Groundwater Case Closures of MtBE, Benzene & Chlorinated Solvents to Non-Detect By In-Situ Electrokinetic Processes. *Presented at the HAZMACON Conference*, Santa Clara Convention Center, Santa Clara, CA, April 1998, sponsored by the Association of Bay Area Governments.
28. Loo, W.W. (1993b). Bio-treatment of Chlorinated Solvents in Soil and Groundwater Utilizing Glucose as the Co-Substrate. *Proceedings of the 1993 HAZMACON Conference*, San Jose, CA.
29. Loo, W.W. (1994). Passive In-Situ Bio-treatment of Gasoline, Diesel and Kerosene in Soil and Groundwater. *Proceedings of the 1994 HAZMACON Conference*, San Jose, CA.
30. Loo, W.W. (1996). Experience of Soil and Groundwater Project Closures in California and Nevada. *Proceedings of SUPERFUND WEST'96, Las Vegas, Nevada*, Sponsored by Hazardous Waste Management Association, Washington, DC.
31. Loo, W.W. et al. (1991). *Soil Closure Report, submitted to Santa Ana Regional Water Quality Control Board*. Electrokinetic Enhanced Soil Remediation with Chlorinated Solvents at the Northrop ESD Site, Anaheim, CA.
32. Loo, W.W. (1993). Engineered Soil Vapor Extraction: Two Case Histories. *Proceedings of the 1993 HAZMACON Conference*, San Jose, CA.
33. Loo, W.W. and Wang, I.S. (1991). Remediation of Groundwater Aquifer by In-Situ Electrolysis and Electro-Osmosis. *Proceedings of National Research and Development Conference on the Control of Hazardous Materials*, February 20–22, 1991, Anaheim, CA, p. 163.

FIELD CAPACITY

Joseph Skopp
University of Nebraska
Lincoln, Nebraska

Storage of soil water is important to soil flora and fauna, to hydrologic processes, and to a variety of soil physical properties. One element of soil water storage is an upper limit on water that is available for plant use, referred to as "field capacity." This entry provides a history of the concept, its relation to soil water holding properties, predictive techniques for field capacity, applications, and failures of the field capacity concept.

HISTORY OF THE CONCEPT

Plant Available Water

Early researchers of plant water relations quickly realized that in water limited regions, the ultimate crop yield depended not only on the amount of rainfall received but on what portion of that rainfall was usable. Three key factors could reduce the usefulness of rainfall: runoff, drainage below the root zone, and evaporation. The net effect of rainfall, runoff, and drainage results in retention of water in the root zone. Evaporation and transpiration reduce this retained water. This retention was studied by Veihmeyer and Hendrickson (1). The concepts developed by Veihmeyer and Hendrickson provide a means to evaluate whether a soil retains sufficient water to support a particular crop. Veihmeyer and Hendrickson (2) redefined their original definition of field capacity to "the amount of water held in soil after excess water has drained away and the rate of downward movement has materially decreased, which usually takes place within 2–3 days after a rainfall event or irrigation in pervious soils of uniform structure and texture."

Field studies focused on measuring the soil water content following a heavy rainfall (that penetrated deeply into the soil). It was observed that the water content would change after the rain stopped. This change gradually slowed such that after 3 days it was difficult to observe additional changes. The inaccurate concept took hold: that after 1–3 days, drainage stopped, and the water retained in the soil at that time would then be available for transpiration. The choice of 2–3 days was partially the result of limited means to measure soil water content. However, the concept took hold that the soil stored water and an upper limit on stored water (available for plant use) could be defined. This upper limit is called "field capacity."

Correspondingly, there is a lower limit to the soil's ability to store water that plants can use. This lower limit is called the "wilting point" or "permanent wilting point." This also has a historical basis, usually limited to a particular plant but more typically taken as the volumetric water content at a matric potential of −1.5 MPa (−15 bars). The difference between field capacity and the permanent wilting point is called the "available water capacity" or the potential "plant available water." Figure 1 shows how the soil porosity (or void space) might be divided into these components. For dryland agriculture, the "available water capacity" approximates the soil's water storage ability. For irrigated agriculture, the field capacity is combined with a lower limit of water content usually set at a point where irrigation is started.

A related interpretation of the field capacity is the water held against gravity. Early attempts to transfer these ideas to the lab depended on a centrifuge to remove

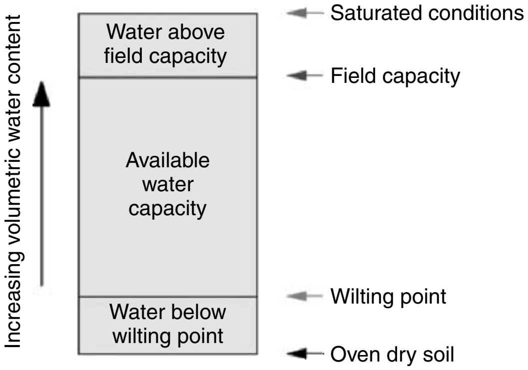

Figure 1. Naive concept of water filled soil porosity.

soil water (3–5). Some workers attempted to relate field capacity to these soil water contents (measured under specific centrifuge conditions, e.g., moisture equivalent) with limited success. Many of these early researchers were limited by their understanding of water movement.

There are several concerns in regard to the definition of field capacity proposed by Veihmeyer and Hendrickson. These include identifying when water movement ceases and the rarity of "soils of uniform structure and texture." The following section addresses the first issue, and the second is postponed to a later section.

Does Drainage Cease?

Later research into the process of drainage showed variation across soil types. Typically, sandier soils drain faster before a noticeable drop-off in water movement. More finely textured soils show a much longer decay process for changes in water content. Thus, at least sands appeared to conform to the Veihmeyer and Hendrickson concept. However, Gardner et al. (6) and Hillel et al. (7) clearly showed that water content does not approach equilibrium up to 30 days after infiltration even on a uniform sandy soil. Figure 2 shows an example of this drainage process for a sandy soil. Thus, for many soil environments, the concept of field capacity is an oversimplification and invalid (i.e., there is no well-defined drainage-induced upper limit to soil water content).

Infiltration Storage Capacity and Runoff

As indicated earlier, not all rainfall enters the soil. Depending on the infiltration characteristics of the soil and the rainfall (intensity and duration), water may pond and runoff. The ability of soils to accept rainfall (i.e., infiltration) depends on a number of soil hydraulic properties (most directly hydraulic conductivity and the soil water characteristic) as well as surface roughness. These infiltration properties may in turn correlate with properties such as field capacity. Thus, some hydrologic runoff models may require soil hydraulic properties as input parameters, which provides an indirect linkage of soil properties such as field capacity to runoff. It also creates a demand for estimates of field capacity, no matter how poorly defined it may be.

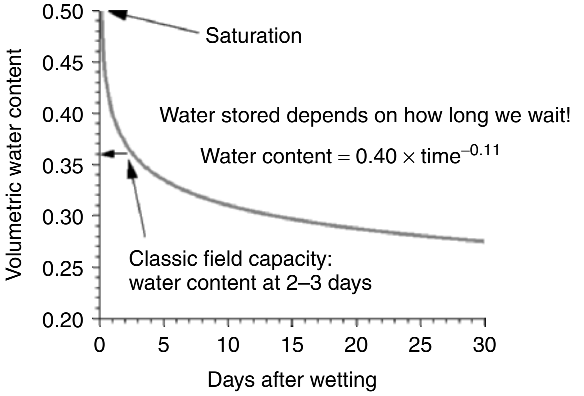

Figure 2. Example of a drainage curve for a sandy soil.

Erosion

When runoff occurs, the potential for water erosion and transport of eroded material also arises. Ponding and runoff contribute separately to the detachment of soil material (the initial process in soil erosion). Subsequent to detachment, runoff can remove the suspended material and transport it off-field or downstream. Runoff also contributes with greater velocities resulting in greater detachment. Finally, the ability of runoff to transport eroded material is also proportional to its velocity.

Field capacity influences the water erosion process through its indirect effect on infiltration. Typically, a larger field capacity results in greater infiltration rates and lower runoff velocities compared to a similar soil with smaller field capacity. These effects may be masked if the storm is very small or very large.

Wind erosion is also influenced indirectly by field capacity. The ability of wind to detach and suspend particles depends on the water content and vegetation (or plant residue) near the soil surface. The linkage between field capacity and some of these variables is weak, but some modeling efforts still rely on inputs of field capacity in their prediction algorithms.

RELATION TO WATER HOLDING PROPERTIES

The original experiments by Veihmeyer and Hendrickson (1,2) were time-consuming, labor-intensive, and not always readily adaptable to all soils and soil environments. A number of researchers looked for alternative means of estimating soil water content (or another property) that would relate to the concept of field capacity (i.e., the water content of a soil following a heavy rain and after 2–3 days of drainage). Several methods have been proposed; the most widely adopted is the use of volumetric water content at a specified matric potential. The water content is usually reported with units of $m^3\ m^{-3}$ to remove confusion with gravimetric water content. We describe three approaches to determining this water content, based on matric potential, unsaturated hydraulic conductivity, or soil water flux.

Matric Potential: The 1/10 and 1/3 Bar Criteria

A variety of workers (8,9) have examined the behavior of soil water content from field experiments meant to simulate field capacity. A broad pattern emerged suggesting that soil matric potential (a measure of the energy with which water is held by soil) provides an indicator independent of soil texture. An initial examination suggests that a range of values from −5 to −50 kPa might be reasonable; values closer to zero are used for coarse textured soils, and the more negative values are for more finely textured soils. Some researchers suggested that −10 kPa (or 1/10 bar in older units) was appropriate for sands and −33 kPa (or 1/3 bar) for all other soils. Thus, the volumetric water content at either −10 kPa or −33 kPa is commonly reported as the field capacity. The relative simplicity of this concept resulted in the adoption of these values as standard measurements by the USDA Natural Resources Conservation Service and by most introductory textbooks. However, not all researchers have adopted these values, and for individual soils, a different matric potential may be more appropriate. Below (under the section titled Deficiencies), we mention some of the profound limitations and errors that can be introduced when using the concept of field capacity as the volumetric water content at 1/10 or 1/3 bar.

Unsaturated Hydraulic Conductivity Concept

A second approach to field capacity looks for a soil property that describes water movement. In particular, a property is desired that characterizes the point where water movement slows to an imperceptible level. Unsaturated water movement is described by a Buckingham–Darcy law where the “unsaturated hydraulic conductivity function” controls water movement. In this approach, the appropriate water content (or matric potential) is the value where the unsaturated hydraulic conductivity is reduced to some fraction of its saturated value. A variety of fractions have been proposed (0.1 or 1/e). This method is rarely used because of the difficulty currently associated with measuring unsaturated hydraulic conductivity.

Flux Concept

A third concept looks directly at water movement. The Buckingham–Darcy law can be used to predict the flux of water. The concept of field capacity is interpreted to mean the point where the flux is close to zero (10). Like the hydraulic conductivity method, just described, it relies on basic concepts of water movement, and it may set up a criterion relative to saturated flow. Unlike the hydraulic conductivity method, it uses a directly measured water movement (the flux), rather than an indirect physical property, and may provide an absolute criterion for zero flux, rather than a relative one. Meyer and Gee (10) suggest that a drainage flux of 3×10^{-7} cm/s represents field capacity.

PREDICTIONS OF WATER CONTENT AT 1/10 AND 1/3 BAR

The lack of a valid laboratory method combined with the time and expense of field methods has not reduced the desire for information about field capacity. Instead, many researchers have adopted the most convenient measurement (i.e., the volumetric water content at 1/10 bar or 1/3 bar). A number of laboratories will perform this analysis. However, water content measurements at these matric potentials is extremely sensitive to sample disturbance. This requires collecting and transporting undisturbed samples to laboratories for measurement. Handling of undisturbed samples is costly and requires special training that many laboratories do not have. The coefficient of variation for undisturbed samples has been reported as between 20 and 50% (11). Thus, laboratory measurements have clear advantages over field methods, but they are not without their own set of difficulties.

One solution to the problems just described is to develop regression or physical models to predict the water content at a specified matric potential. A variety of regression models have been proposed. Some of them are described by Acutis and Donatelli (12), Baumer and Brasher (13), Cassel et al. (14), Ritchie et al. (15), and Rawls et al. (16). More recently, Gijsman et al. (17) compared a variety of predictive techniques and found a great deal of variation among the models they tested. These authors recommend the method described by Saxton et al. (18) based on their independently obtained data set. Note that use of any of these relations does not remove the concerns expressed later in this article.

APPLICATIONS

The use of field capacity (usually as 1/10 or 1/3 bar water content) has grown beyond the original goal of understanding crop yields. In this section, we briefly look at current applications for crop yield prediction as well as a surrogate for other soil hydraulic properties that are more difficult to measure.

Agronomic/Ecological Models

Advances in computational power allow researchers to describe water movement in soils after rainfall and during plant growth by using physical transport principles. One common application of such models is the prediction of plant growth and crop yield (17,19). These models typically require the input of soil physical properties, possibly including field capacity. Input to these models is the laboratory measured water content at 1/10 or 1/3 bar as one point on a soil water retention curve that is used to predict other hydraulic properties. In other words, these models tend to be dynamic and allow soil water content to change continuously beyond the 2- or 3-day period despite the input of a so-called field capacity. In addition to crop yield/plant growth, these models may be used to develop irrigation timing strategies.

Hydrologic Models

Another area where field capacity estimates are used is in watershed modeling (20,21). These hydrologic models provide a means to describe the response of a watershed to a rainfall. A key component of these models is the infiltration process. In some cases, empirical infiltration models are used. Otherwise, physical soil water transport models (similar to agronomic models) are used. In either case, parameter inputs are typically needed either as an explicit parameter or to calculate the needed parameter (e.g., to estimate unsaturated hydraulic conductivity).

Other Applications

Field capacity (or at least the volumetric water content at 1/10 or 1/3 bar) can frequently act as a surrogate for other (less readily measured) soil properties, which, combined with its ready availability in a variety of databases, has led to its appearance in a variety of publications that are not directly tied to water movement or plant growth. In this case the use of the phrase “field capacity” refers only to a specific laboratory measured water content and has lost any “field” significance.

Examples of this use of field capacity include observations of microbial occurrence and activity (22), solute transport (23), soil strength (24,25), and soil erodibility (26).

DEFICIENCIES IN MEASUREMENTS

A variety of criticisms have been directed toward the field capacity concept by Richards (27), Baver et al. (28), Hillel (7) and Cassel and Nielsen (29). In addition to questions regarding the time when drainage ceases (mentioned earlier), there are three basic areas of concern with field capacity as a widely used concept describing water movement in field soils. The importance of changes in soil structure, soil layering, and initial conditions. We briefly discuss these issues as presented by the previously cited researchers.

Dependence on Structure and Disturbance

Field capacity measurements in the field and laboratory measurements (for matric potentials between 0 and 100 kP) are sensitive to changes in soil structure or aggregation. Changes in soil structure due to tillage, root growth, faunal activity, or freeze/thaw cycles are expected to alter soil hydraulic properties and the movement of water. Water quality (e.g., sodium or total salts) is also expected to influence aggregation and ultimately water movement.

Dependence on Soil Morphology and Internal Drainage

Soils are rarely uniform. Layers that have different hydraulic properties are typical. These layers may result in perched water tables or hindered drainage, which can be expected to extend drainage times. Equally important, drainage will not be uniform throughout the soil profile, so there will not be a single water content at a point of restricted water movement.

This existence of soil layers plays particular havoc with the use of laboratory determined values based on a specific matric potential. It may be possible to measure each soil layer independently, but the sequence of layers may be more important in controlling drainage patterns.

Dependence on Initial Conditions

The soil conditions prior to wetting (required in the basic definition of field capacity) are assumed to be unimportant. In fact, initial soil water content can influence the uniformity of wetting (30). It is not possible to measure field capacity under all possible scenarios of initial water content distribution or wetting history. Also of importance, the existence and position of the water table will influence both wetting and the drainage process. The presence of a shallow water table may act to reduce drainage. The use of laboratory methods to estimate field capacity cannot account for these conditions.

SUMMARY

The field capacity concept has been used to approximate the ability of soils to store water and to calculate an "available water capacity." However, soil water is not a static reservoir, and such a picture oversimplifies soil water movement and availability. Field representations of soil water storage require field measurements. Laboratory measurements (particularly water content at 1/10 and 1/3 bar) are popular because of their availability and seeming simplicity. Laboratory measurements can be useful for describing soils or estimating other soil hydraulic properties; however, their use as a reflection of water storage may be misleading and should be used with extreme caution. Any representation of soil water storage must be site specific and must take antecedent conditions into account.

BIBLIOGRAPHY

1. Veihmeyer, F.J. and Hendrickson, A.H. (1931). The moisture equivalent as a measure of the field capacity of soils. *Soil Sci.* **32**: 181–193.
2. Veihmeyer, F.J. and Hendrickson, A.H. (1949). Methods of measuring field capacity and wilting percentages of soils. *Soil Sci.* **68**: 75–94.
3. Briggs, L.J. and Mclane, J.W. (1907). *The Moisture Equivalent of Soils*. U.S.D.A. Bureau of Soils Bull. 45.
4. Briggs, L.J. and Shantz, H.L. (1912). *The Wilting Coefficient for Different Plants and its Indirect Determination*. U.S.D.A. Bureau Plant Indus. Bull. 230.
5. Russel, J.C. and Burr, W.W. (1925). Studies on the moisture equivalent of soils. *Soil Sci.* **19**: 251–266.
6. Gardner, W.R., Hillel, D., and Benyamini. (1970). Post irrigation movement of soil water: I. Redistribution. *Water Resour. Res.* **6**: 851–861.
7. Hillel, D., Krentos, V.D., and Stylianou, Y. (1972). Procedure and test of an internal drainage method for measuring soil hydraulic charcteristics in situ. *Soil Sci.* **114**: 395–400.
8. Coleman, E.A. (1947). A laboratory procedure for determining the field capacity of soils. *Soil Sci.* **63**: 277–283.
9. Jamison, V.C. and Kroth, E.M. (1958). Available moisture storage capacity in relation to textural composition and organic matter content of several Missouri soils. *Soil Sci. Soc. Am. Proc.* **22**: 189–192.
10. Meyer, P.D. and Gee, G.W. (1999). Flux-based estimation of field capacity. *J. Geotechnical Geoenvironmental Eng.* **125**(7): 595–599.
11. Warrick, A.W. and Nielsen, D.R. (1980). Spatial variability of soil physical properties in the field. In: *Applications of Soil Physics*. D. Hillel (Ed.). Academic Press, New York, Chap. 13.
12. Acutis, M. and Donatelli, M. (2003). *SOIL*PAR 2.00: software to estimate soil hydrological parameters and functions. *Eur. J. Agron.* **18**: 373–377.
13. Baumer, O.W. and Brasher, B.R. (1982). *Prediction of Water Content at Selection Suctions*. Am. Soc. Agric. Eng. Paper 82–2590.
14. Cassel, D.K., Ratliff, L.F., and Ritchie, J.T. (1983). Models for estimating in situ potential extractable water using soil physical and chemical properties. *Soil Sci. Soc. Am. J.* **47**: 764–769.
15. Ritchie, J.T., Gerakis, A., and Suleiman, A. (1999). Simple model to estimate field-measured soil water limits. *Trans. ASAE* **42**: 1609–1614.
16. Rawls, W.J., Gish, T.J., and Brakensiek, D.L. (1991). Estimating soil water retention from soil physical properties and characteristics. *Adv. Agron.* **16**: 213–234.

17. Gijsman, A.J., Jagtap, S.S., and Jones, J.W. (2002). Wading through a swamp of complete confusion: How to choose a method for estimating soil water retention parameters for crop models. *Eur. J. Agron.* **18**: 75–105.
18. Saxton, K.E., Rawls, W.J., Romberger, J.S., and Papendick, R.I. (1986). Estimating generalized soil-water characteristics from texture. *Soil Sci. Soc. Am. J.* **50**: 1031–1036.
19. Ittersum, M.K. et al. (2003). On approaches and applications of the Wageningen crop models. *Eur. J. Agron.* **18**: 201–234.
20. Descroix, L. et al. (2002). An experimental analysis of hydrodynamic behaviour on soils and hillslopes in a subtropical mountainous environment (Western Sierra Madre, Mexico). *J. Hydrol.* **266**: 1–14.
21. Devonec, E. and Barros, A.P. (2002). Exploring the transferability of a land-surface hydrology model. *J. Hydrol.* **265**: 258–282.
22. Nunez-Requeira, L. et al. (2002). The influence of some physicochemical parameters on the microbial growth in soils. *Thermochimica Acta* **394**: 123–131.
23. Dutt, G.R. and Tanji, K.K. (1962). Predicting concentrations of solutes in water percolated through a column of soil. *J. Geophys. Res.* **67**: 3437–3439.
24. Becher, H.H. (1998). Resistances to penetration of aggregates from loess-derived topsoils at different soil water tensions. *Soil Tillage Res.* **47**: 73–81.
25. Towner, G.D. and Childs, E.C. (1972). The mechanical strength of unsaturated porous granular material. *J. Soil Sci.* **23**: 481–498.
26. USDA. (1995). Available: http://topsoil.nserl.purdue.edu/nserlweb/weppmain/jhtml/appx1.html.
27. Baver, L.D., Gardner, W.H., and Gardner, W.R. (1972). *Soil Physics*, 4th Edn. John Wiley & Sons, New York.
28. Richards, L.A. (1960). Advances in Soil Physics. *Trans. Int. Cong. Soil Sci.* **7th Madison I**: 67–69.
29. Cassel, D.K. and Nielsen, D.R. (1986). Field capacity and available water capacity. In: *Methods of Soil Analysis*. Part 1, *Physical and Mineralogical Methods*. Agronomy Monograph #9, 2nd Edn. American Society of Agronomy, Madison, WI, Chap. 36.
30. Rickard, D.S. (1973). Plant and soil water: Information required in agronomic trials. In: *Soil Plant Water Symp. 1973*. Palmerston North, N.Z. New Zealand Dept. of Scientific and Industrial Research, Wellington, N.Z., pp. 107–112.

GROUNDWATER FLOW PROPERTIES

Mukand Singh Babel
Asian Institute of Technology
Pathumthani, Thailand

Groundwater constitutes a large component of the freshwater cycle. It moves more or less continuously from areas of recharge to areas of discharge through interconnected void spaces in porous media, fractured media, and large passages. Groundwater flow results mainly from precipitation that infiltrates into the soil; it also originates from influent streams, seepage from reservoirs, artificial recharge, seepage from oceans, etc. It often augments streamflow, eventually moving through the soil to the stream channel, also referred to as baseflow, or dry-weather flow.

The driving force for groundwater movement is measured in terms of hydraulic head, which represents the fluid pressure potential and elevation potential. It is also referred to as the piezometric or potentiometric head. Contour maps of hydraulic head (piezometric or potentiometric maps), which can be constructed from water-table elevations, can be used to infer directions of groundwater flow.

The hydraulic and physical properties of a porous medium influence groundwater flow. They are porosity, hydraulic conductivity and permeability, specific storage, transmissivity and storage coefficient. A hydrologic unit (aquifer) is homogenous if its hydraulic properties are the same at every location. However, owing to the spatial and temporal variability in the geologic processes that create and modify rocks and sediments, no unit is truly homogenous.

HYDRAULIC CONDUCTIVITY

Hydraulic conductivity (K) is a measure of the ability of a fluid to move through interconnected void spaces in a sediment or rock. It is a function of both the medium and the fluid properties. A medium has a unit hydraulic conductivity if it transmits in unit time a unit volume of groundwater at the prevailing kinematic viscosity through a unit cross-sectional area, measured at right angle to the direction of flow, under a unit hydraulic gradient. The SI unit is

$$K = -\frac{q}{dh/dl} = -\frac{\text{m/day}}{\text{m/m}} = \text{m/day} \tag{1}$$

Several factors affect hydraulic conductivity—those that pertain to the water-bearing material itself and those related to the groundwater as a fluid. The geometry of the pore spaces through which groundwater movement occurs, the grain size distribution, the surface roughness of the particles, and other secondary geologic processes such as faulting and folding, secondary deposition, and secondary solution in rock (such as limestone) are the major factors pertaining to the porous medium. The fluid characteristics that affect hydraulic conductivity are density, viscosity, and salinity. Temperature, by inversely affecting viscosity, has a direct influence on groundwater flow. The effect of salinity, on the other hand, is brought about by changes in the ionic concentrations of the groundwater and also due to the chemical effect of the saline water on the aquifer material. However, the effects of fluid characteristics on hydraulic conductivity tend to be rather less important than the effects of the aquifer characteristics.

PERMEABILITY

The term permeability (k), used synonymously with intrinsic permeability, of a porous medium defines its ability to transmit a fluid, which is a property only of the medium and is independent of fluid properties. Permeability (k) is defined as

$$K = \frac{k\rho g}{\mu} \tag{2}$$

Table 1. Representative Values of Hydraulic Conductivity and Permeability[a]

Sediment or Rock Type	Hydraulic Conductivity (m/day)	Permeability (m^2)
Clays	10^{-7}–10^{-3}	10^{-19}–10^{-15}
Silts	10^{-4}–10^{0}	10^{-16}–10^{-12}
Fine to coarse sands	10^{-2}–10^{+3}	10^{-14}–10^{-9}
Gravels	10^{+2}–10^{+5}	10^{-10}–10^{-7}
Shales (matrix)	10^{-8}–10^{-4}	10^{-20}–10^{-16}
Shales (fractured and weathered)	10^{-4}–10^{0}	10^{-16}–10^{-12}
Sandstones (well-cemented)	10^{-5}–10^{-2}	10^{-17}–10^{-14}
Sandstones (friable)	10^{-3}–10^{0}	10^{-15}–10^{-12}
Salt	10^{-10}–10^{-8}	10^{-22}–10^{-20}
Anhydrite	10^{-7}–10^{-6}	10^{-19}–10^{-18}
Unfractured igneous and metamorphic rocks	10^{-9}–10^{-5}	10^{-21}–10^{-17}
Fractured igneous and metamorphic rocks	10^{-5}–10^{-1}	10^{-17}–10^{-13}

[a] Reference 1.

where K = hydraulic conductivity, k = permeability, ρ = fluid density, g = acceleration due to gravity, and μ = dynamic viscosity of the fluid. Table 1 presents representative values of hydraulic conductivity and permeability for a variety of sediment and rock types.

Owing to the spatial and temporal variability in the geologic processes that create and modify rocks and sediments, heterogeneity occurs in a range of spatial scales. In a porous multilayer medium, the variation in hydraulic conductivity among the depositional units can amount to five or six orders of magnitude. The groundwater analyst has to decide whether or not the homogeneity approximations are adequate, and if not, at what scale the heterogeneity should be specified. Some averaging rules can be applied to average a set of measurements to a single value of effective hydraulic conductivity (K_e) when using a homogenous medium approximation.

The effective hydraulic conductivity (K_e) depends upon the properties of the porous medium and the pattern of fluid flow. For steady-state flow, with a spatially uniform hydraulic gradient, the following averaging rules apply:

1. For a perfectly stratified medium, n layers of layer thickness d_i, and hydraulic conductivity K_i:
   (a) When flow is parallel to the layering, K_e is given by applying the arithmetic mean as

$$K_e = K_A = \sum_{i=1}^{n} \frac{d_i K_i}{\sum_{i=1}^{n} d_i} \qquad (3)$$

   (b) When flow is perpendicular to the layering, K_e is given by applying the harmonic mean as

$$K_e = K_H = \sum_{i=1}^{n} \frac{d_i}{\sum_{i=1}^{n} \frac{d_i}{K_i}} \qquad (4)$$

2. For a heterogeneous medium, nonstratified, and m measurements
   (a) For two-dimensional models, K_e is given by applying the geometric mean as

$$K_e = K_G = (K_1 K_2 \cdots\cdots K_m)^{1/m} \qquad (5)$$

   (b) For three-dimensional models, K_e is given as,

$$K_e = (1 + \sigma^2{}_y/6) \qquad (6)$$

where $\sigma^2{}_y$ is the variance of the natural logarithms of the hydraulic conductivity measurements.

TRANSMISSIVITY

Transmissivity (T) is defined as the amount of water that can be transmitted horizontally through a unit width by a fully saturated thickness of an aquifer under a unit hydraulic gradient. The transmissivity of an aquifer is equal to the hydraulic conductivity of the aquifer multiplied by the saturated thickness of the aquifer. Thus,

$$T = Kb \qquad (7)$$

where T = transmissivity, K = hydraulic conductivity, and b = saturated thickness of the aquifer.

STORAGE COEFFICIENT

Specific storage (S_s) is defined as the volume of water that a unit volume of aquifer releases from storage under a unit decline in hydraulic head. Storage coefficient or storativity (S) is the product of specific storage (S_s) and aquifer thickness. It is defined as the volume of water that a permeable unit will absorb or expel from storage per unit surface area per unit change in head. It is a dimensionless quantity, and this property becomes important when an aquifer loses or gains water. Storativity is due to porosity in the aquifer and is different for confined and unconfined aquifers. For a unit decline in hydraulic head, the volume of water released from a confined aquifer is substantially less than that for an unconfined aquifer, indicating a comparative low value of the storage coefficient for a confined aquifer.

SPECIFIC YIELD (S_Y)

The storativity of an unconfined aquifer termed specific yield, is defined as the volume of water that an unconfined aquifer releases from storage per unit surface area of aquifer per unit decline in the water table, as shown in Fig. 1. Table 2 gives representative values of the specific yield for different aquifer materials.

STORATIVITY IN CONFINED AND UNCONFINED AQUIFERS

The storativity in unconfined aquifers is higher than that in confined aquifers. The higher values of storativity for unconfined aquifers (specific yield), as shown in Table 2, reflect the fact that releases from storage in unconfined aquifers represent an actual dewatering of the soil pores, whereas releases from storage in confined aquifers

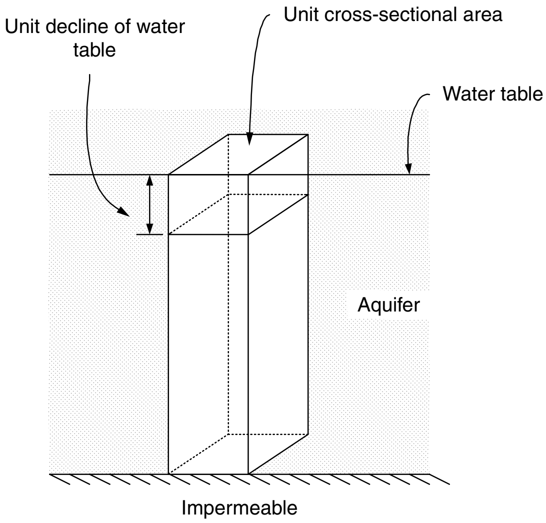

Figure 1. Schematic representation of storativity in unconfined aquifers.

Table 2. Representative Values of Specific Yield[a]

Material	Specific Yield, %
Clays	1–5
Silts	10–20
Fine sands	10–30
Sands and gravels	20–30
Sandstone	5–20
Shale	0.5–5
Limestone	0.5–20

[a] Reference 1.

represent only the secondary effects of water expansion and aquifer compaction brought about by changes in fluid pressure. In most confined aquifers, the storativity values fall in the range 0.00005 to 0.005 (2), indicating that large pressure changes over extensive areas are required to produce substantial water yields. Thus, the same yield can be realized with smaller head changes over less extensive areas in unconfined aquifers compared to confined aquifers, making them more efficient for exploitation by wells.

BIBLIOGRAPHY

1. Maidment, D.R. (1993). *Handbook of Hydrology*. McGraw-Hill, New York.
2. Todd, D.K. (1980). *Groundwater Hydrology*. John Wiley & Sons, New York.

FLUORIDE CONTAMINATION IN GROUND WATER

Chakresh K. Jain
National Institute of Hydrology
Roorkee, India

Fluoride occurrence in ground water is a natural phenomenon, influenced by the local and regional geological setting and hydrogeological conditions of the region. Fluoride content much higher than the maximum permissible limit has been found in many regions throughout the world. The problem assumes more significance in India and other developing countries because a major portion of the population that lives in rural as well as urban areas have to depend on available groundwater resources for their day-to-day requirements. An estimated 62 million people in 17 out of 32 states are affected with dental, skeletal, and/or nonskeletal fluorosis in India. The extent of fluoride contamination varies from 1.0 to 48.0 mg/L.

INTRODUCTION

One of the serious health problems facing India today is the prevalence of a disease known as fluorosis. It involves millions of people and arises primarily due to an excess of fluoride in drinking waters. Fluorosis was first detected in India among cattle (bullocks used for ploughing land) by the farmers of Andhra Pradesh during early 1930s. The farmers noticed the inability of the bullocks to walk due to painful and stiff joints. During later years, the same disease was detected in human beings. Shortt et al. (1) published the first report on endemic fluorosis in India, when the disease was prevalent in four states, Andhra Pradesh, Tamil Nadu, Punjab, and Uttar Pradesh (2). During the period 1960–1986, nine more states have been identified as endemic for fluorosis. During 1990–1992, two more states, Kerala and Jammu and Kashmir have also been identified as endemic for the disease. As per the report published by Rajiv Gandhi National Drinking Water Mission (3), fourteen states and Delhi have been identified as endemic for fluorosis. Datta (4) has also reported fluoride contamination in the ground water of the National Capital Territory of Delhi. Recently, elevated levels of fluoride have been reported from the Birhum district of West Bengal and Karbi-Anglong and Nagaon districts of Assam (5–8). Thus, the total number of states endemic for fluorosis at present is seventeen: Andhra Pradesh, Assam, Bihar, Delhi, Gujarat, Haryana, Jammu and Kashmir, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, and West Bengal.

Fluoride is essential for normal mineralization of bones and formation of dental enamel. About 96% of the fluoride in the body is found in bones and teeth. Calcium-rich constituents of teeth, enamel and dentine, have a strong affinity for fluoride during the formation of teeth. Fluoride combines with calcium to form calcium fluoroapatite crystals during the mineralization of teeth. The principal sources of fluoride for human physiology are water and food. At the same time, fluoride also has considerable physiological importance for humans and animals. If one is exposed to fluoride for long durations, it may pose serious health hazards. The major health problems caused by fluoride are dental caries, teeth mottling, skeletal fluorosis, and deformation of bones in children as well as in adults (9).

The Geological Survey of India has brought out considerable data that reveal that fluorite, topaz, apatite, rock phosphate, phosphatic nodules, and phosphorites are wide-spread in India and contain high percentages of fluoride (10). As a result of rich mineral content, fluoride leaches out and contaminates water and soil, thereby contaminating agricultural crops as well. Several natural fluoride-bearing minerals identified in India include fluorides, fluorites, and cryolite; phosphates, fluorapatite, and wagnerite; silicates and topaz; and the mica group, magnesium mica and lithium mica.

The permissible limits for fluoride in drinking water have been defined by various agencies. In view of the health problems, the Bureau of Indian Standards, formerly known as Indian Standards Institution (ISI), has laid down the Indian standard of 1.0 mg/L as the desirable limit (11). The Indian Council of Medical Research (ICMR) has recommended a desirable level of 1.0 mg/L and the maximum permissible limit of 1.5 mg/L, which means that the body may tolerate fluoride up to the limit of 1.5 mg/L depending on the nutritional standards and body physiology. The World Health Organization (12) has recommended 0.5 mg/L as the maximum limit for fluoride.

In India, more than 76% of the population lives in rural areas. The problem of endemic fluorosis occurs with varying intensity in different parts of the country. An extensive survey of community water supplies has shown that around 25 million people in rural areas consume water whose fluoride content is more than the maximum permissible concentration of 1.5 mg F/L (13). According to a Sub-mission set up by the Government of India under a Technical Mission, about 25 million people are suffering from fluorosis at various stages, and another 25 million are reported to be prone to it (14,15). Jain et al. (16) have reviewed the status of fluoride contamination in groundwater and Susheela (7) reported the overall status of fluoride contamination and the suffering of the people in India. It is stated that an estimated 62 million people in 17 states out of 32 states are affected with dental, skeletal, and/or nonskeletal fluorosis.

Rajasthan is one of the states where fluorides at high levels in all 31 districts and the problem of dental and skeletal fluorosis is widespread (17). The studies carried out by the Rajasthan Voluntary Health Association (RVHA) noticed that the total number of problem villages that have high fluoride content in groundwater is 2702. About three million people are consuming water with excess fluoride, out of which one-fifth of the affected population are children below the age of 7 years, and about 35,000 people are consuming water that has more than 10 mg/L of fluoride.

As much as 90 mg/L of fluoride was found in groundwater in western Rajasthan (18). High fluoride content is attributed to the circulation of groundwater through fluorite-bearing granites and acid volcanics and basic dykes (that act as barriers and result in ample contact time for circulating water to dissolve fluorite). High fluoride content is always associated with a lower Ca/F ratio. It is an established fact that waters high in calcium are low in fluoride content, and therefore the low fluoride concentration in the eastern region can possibly be explained by the presence of high calcium in the water due to the movement of groundwater through carbonate rocks and a higher recharge from surface water to the Sukri River that possibly dilutes the fluoride content further.

The fluoride concentration in the Nagpur district of Maharashtra state varies from 0 to 44 mg/L (19). The fluoride content increases with an increase in salinity. High bicarbonate and sodium content were observed in groundwater rich in fluoride. High values of phosphorous were also observed in association with high fluoride content indicating its possible occurrence due to mineral weathering and local pollution.

Wodeyar and Sreenivasan (20) reported the occurrence of fluoride in the Bellary district in the state of Karnataka. Deeper aquifers have shown higher fluoride concentrations than shallow aquifers. A high degree of weathering and easy accessibility of circulating water to the weathered rocks due to intensive and long-term irrigation are responsible for leaching of fluoride from parent minerals in soil and rocks. Further concentration has been brought about by the arid climate of the region and the long residence time of groundwater in the aquifer. The influence of local lithology and soil, aided by other factors like very low freshwater exchange due to the arid climate of the region, is also responsible for higher concentrations of fluoride in groundwater.

Shukla et al. (21) studied the fluoride level in human teeth obtained from two different regions, Bhilai and Shillong in the state of Madhya Pradesh. They reported that the Bhilai region is fluoridated and have an average fluoride level of 1.3 ppm, whereas the Shillong region is nonfluoridated and has an average fluoride level of 0.02 ppm. The fluoride content of the teeth, it was found, increases with age, and therefore, the duration of exposure to fluoride has a significant effect on its accumulation in the dental structure. The biologically active age group of 10–30 years incorporates fluoride in teeth more rapidly compared to the passive age group of 50–70 years in both regions.

Apparao and Karthikeyan (22) reported a fluoride content of less than 1 mg/L in the drinking water of seven fluorosis-affected villages of the Nilakottai and Dindigul blocks in the state of Tamil Nadu. The poor socioeconomic groups of these villages are affected much by dental as well as skeletal fluorosis. The general nutritional status of their diet is poor, and specifically, the calcium and vitamin C intake are very low. It is surprising that people exposed to 0.51 mg/L F water are affected, and such reports are scanty in India. However, analysis of food grains grown in these areas indicated 4.0 to 4.6 mg/kg of fluoride which is a significant amount of the total fluoride intake and hence the prevalence of fluorosis in these areas. Other fluoride-affected districts include Vellore, Dharmapuri, Trichy, Karur, Salem, Namakkal, Erode, Coimbatore, and Virudhunagar (23–28). Agricultural practices and the depth to the groundwater table, it has been found, have an important impact on fluoride concentration.

Malhotra et al. (29,30) studied fluoride contamination in the groundwater of the Allahabad, Pratapgarh, Gonda, Ballia, Bareilly, and Fatehpur districts in the state of Uttar Pradesh and reported apparently high levels of fluoride. The maximum fluoride content of 6.5 mg/L was found in rural areas of Phulpur in the Allahabad district. The extent of fluoride increases in summer months, apparently because of a drop in the water level. The evidence suggests that due to a gradual drop in the water level below the earth's crust a greater concentration of fluoride is found in the water because of greater dissolution of fluoride from rocks and soil.

Gupta et al. (31,32) and Shrivastav and Choudhary (33) reported fluoride contamination in the Agra district the highest concentration of 22 mg/L was in Bainkhera village. A significant positive correlation of fluoride with pH and negative correlation with hardness and magnesium was observed. Dental fluorosis was observed among school children in most of the villages, even where water had a fluoride concentration well within permissible limits. These observations clearly indicate the need for redefining the permissible levels of fluoride in drinking water and an exhaustive assessment of fluoride levels in foods.

Swamy (13) presented an interdisciplinary approach to understanding the problem of fluorosis against the background of present day development. Sudarshan and Reddy (2) described the pollution of fluoride in groundwater and its impact on the environment and on the socioeconomic status of the people. Mariappan et al. (28) described the occurrence and method of removing fluoride.

SOURCES OF FLUORIDE CONTAMINATION

Geological Sources

Fluoride occurrence in groundwater is a natural phenomenon, influenced by the local and regional geologic setting and hydrogeological conditions of the region. Fluoride occurs abundantly in the earth's crust as a component of rocks and minerals. The dissolution of fluoride from geologic formations occurs through the rainwater and repeated irrigation of agricultural lands. As rain water percolates through the soil, it comes in contact with the rocks and minerals in the aquifer materials. Due to the acid in the soil, dissolution of fluoride from the country rocks occurs.

The main sources of fluoride in natural water are fluorite (CaF_2), fluorapatite [$3Ca_3(PO_4)_2{-}CaF_2$], cryolite (Na_3AlF_6), magnesium fluoride (MgF_2), and as a replacement of, ions in the crystal lattice of mica and many other minerals (34,35). The other important fluoride-bearing minerals are biotite, muscovite, lepidolite, tourmaline, hornblende series minerals, glucophane riebeckite, asbestos (chrysolite, actinolite, anthophyllite), sphene, apophyllite, zinnwaldite, etc.

Dissolution of fluoride in groundwater itself may also contribute to the fluoride contamination phenomenon. The fluoride content in groundwater becomes higher in summer season due to a drop in the water level. The higher concentration of fluoride in the summer may be further attributed to the higher dissolution of fluoride, which may be due to the presence of air in the minerals/rocks cavities, which indicates that oxygen in the cavities of the geologic formation catalyzes the fluoride dissolution process. The evidence suggests that, as the ground water level gradually drops below the earth's crust, a greater concentration of fluoride is found in the water because of greater dissolution of fluoride from rocks and soil.

Industrial Sources

Industries are also a very important source of groundwater contamination. The effluents and other by-products of industries often constitute sources of groundwater pollution. Industries that release significant quantities of fluorides in process waste streams include fluorosilicone acid and fluorocarbon polymer synthesis, gasoline production; and manufacture's of coke, ceramic, cement, enamel, fiberglass, electronics, pesticides and fertilizers, surface heating operations, metal etching (with hydrofluoric acid), and wood preservatives (36).

Fluoride discharged from fertilizer manufacturing processes is typically in the form of silicon tetrafluoride, as a result of processing of phosphate rock. Power plant boiler cleaning wastes may contain fluoride due to its presence in the cleaning formulation. Both fluorspar and limestone are among the basic fluxing materials used in steelmaking. Air pollution control scrubbers water from steel manufacturing in basic oxygen, open hearth, and electric arc furnaces as well as in the sintering plant is the principal source of fluoride-containing wastewater from this industry. The primary aluminum processing industry uses the fluoride compound cryolite as a catalyst in bauxite ore reduction, The gaseous fluorides resulting from this process are discharged directly into the atmosphere or the fluorides into aqueous waste streams from wet scrubbing of the process fumes (37). Average fluoride values for aluminum reduction plants are reported as 107–145 ppm in wastewater streams. It has been observed that the number of echinocytes increases, depending on the duration of fluoride exposure among the workers in aluminum smelters (38). Fluoride concentrations ranging from 1,000–3,000 ppm have been reported for a glass manufacturing process. Glass and plating wastes typically contain fluoride in the form of hydrogen fluoride or fluoride ion, depending on the pH of the waste.

A high amount of fluoride has been reported in the groundwater in the vicinity of the aluminum industry (34,39). The main sources of fluoride in brick works are local clay and burning of coal. Malhotra et al. (30) have reported the fluoride content in groundwater in the vicinities of brickworks in Pratapgarh, Mirapur, Kechavpur, Mander, and Ahmedpur. Clarke et al. (40) have also reported fluoride contamination in groundwater due to by-products of brick kilns. Fluoride has also been detected in groundwater and air near Agra due to the 16 brick industries in the area (36).

Nonpoint Sources

Nonpoint sources of groundwater pollution also contribute fluoride to ground water to some extent. Modern agricultural practices that involve the application of fertilizers coupled with pesticides which contain about 1–3% fluoride also contribute fluoride to groundwater (41). The use of phosphatic fertilizers might also be one of the factors contributing to high fluoride concentration, which is being leached down to the groundwater by irrigation return flow (2,36). Irrigation with water containing small amounts of fluoride ion would tend to concentrate these ions in the soil. If the calcium ions are removed as calcite, fluoride ions will either be absorbed or coprecipitated with calcite. If the carbonate concentration is high enough to precipitate calcium as calcite without concomitant removal of fluoride ions, these ions may move down to the water front and ultimately join the main groundwater body (42). The U.S. Environmental Protection Agency (43) has recommended 1 mg/L of fluoride in irrigation water for continuous use but up to 15 mg/L of fluoride for short term use (20 years) on fine soils.

The production of fly ash is another important nonpoint source of fluoride contamination. More than 150 million tonnes of fly ash one produced annually worldwide from the combustion of coal in power plants (44). At least half of this amount is disposed of by landfill, thus contributing to environmental pollution due to leaching of fluoride. In addition to this, smoke particles from the aluminum industry, brickworks, and volcanoes also contain traces of fluroride (37,39,40). Fluoride in the atmosphere has been detected near Agra, India, due to the presence of about 16 brickworks in this area (36). A possibility of groundwater contamination may exist due to traces of fluoride in the atmosphere during the rainy season. The Geological Survey of India has also brought out considerable data which reveal that fluorite, topaz, apatite, rock phosphates, phosphatic nodules and phosphorites are widespread in India and contain high percentages of fluoride. As a result of their rich mineral content, fluoride leaches out and contaminates not only the groundwater but the soil as well, thereby contaminating agricultural crops also.

Other Miscellaneous Sources

Besides water, fluoride is also found in agricultural crops and other edible items. It is well recognized that fluoride enters the human body through various foods, though the data available in India are rather scanty. The information that is available is based on the analysis of food or agricultural crops in certain regions. Thus, data, from a specific location may not be applicable to other regions due to the varying concentration of fluoride in water and soil. Chary et al. (45) and Lakdawala and Punekar (46) reported the fluoride content in some commonly used foods in the Prakasam district and Bombay (India). Nanda and Kapoor (47) found the fluoride content in pine and betel. Sengupta and Pal (48) also reported the fluoride contents of different foodstuffs. Gradien (49) reported that fluoride content generally decreased in the order: tea> cereals> vegetables> pulses> fish> meat> fruits.

In addition to foodstuffs, fluoride has also been reported in cosmetics and drugs. The use of drugs containing sodium fluoride for osteoporosis, otosteosclerosis, and dental caries is very common. Different brands of toothpaste contain excessive amounts of fluoride. Fluoride enters into the circulation directly from the oral cavity through the fine blood vessels of the mouth. Fluoride is a persistent bioaccumulator; even small amounts that enter through fluoride toothpaste one guaranteed entry in children as well as adults, and the cumulative effects of fluoride are a matter of serious concern.

CONTROLLING FACTORS AND MECHANISM

Controlling Factors

Fluoride contamination in ground water is controlled by a number of parameters. The most important are

- distribution of easily weathered fluoride-bearing minerals
- the accessibility of circulating water to these minerals
- pH of the percolating water
- calcium content of the leaching water
- temperature of the percolating water and the soil
- exchangeable ions in the percolating water
- extent of fresh water exchange in an aquifer
- evaporation and evapotranspiration
- complexing of fluoride ions with other ions
- presence of CO_2 and other chemicals in draining water
- residence time of the percolating water in soil

The pH of circulating water is an important factor that controls leaching of fluoride from fluoride-bearing minerals. Wodeyar and Sreenivasan (20) have indicated that higher alkalinity of waters promotes leaching of fluoride and thus affects the concentration of fluoride in groundwater. Alkaline water dissolves fluoride-bearing minerals under simultaneous precipitation of calcium carbonate (50). It has been reported that the degree of weathering and easy accessibility of circulating waters from weathered rocks, due to intensive and long term irrigation, are responsible for leaching fluoride from the parent minerals in soil and rocks. A high content of fluoride has been reported in black cotton soil due to excessive canal irrigation (20). Further, concentration of fluoride has been brought about by the arid climate of the region and the long residence time of groundwater in the aquifer (20,50). Very low fresh water exchange due to the arid climate of the region is also responsible for higher concentration of fluoride in groundwater (20).

The fluoride concentration in groundwater, it was found, is positively correlated with calcium content but with a very low degree of validity (14,15). This observation is similar to that reported by Somani et al. (51), but there is no agreement to that report by Gupta et al. (31,32). This departure from the normal trend may be due to an irregular distribution of fluoride-bearing minerals in the soil, their solubilization characteristics, the nature of the product with soil, and other environmental conditions (51). Generally, waters of high calcium contents are low in fluoride content (52). The presence of carbon dioxide also affects the fluoride dissolution process in rocks.

Fluoride in soil and ground water is also concentrated by evaporation and evapotranspiration due to arid and semiarid atmospheric conditions following scanty rainfall. The overwithdrawal of groundwater may also have favorable effects on fluoride concentration in groundwater. The combined effect of evapotranspiration and long-term contact of the water in the aquifer (due to low hydraulic conductivity of the weathered zone) activates the process of dissolution (50). Ion exchange of the different elements in the soil and circulating water during the weathering process help in dissolving of fluoride from rocks and minerals (17,50). Besides, the formation of ion pairs such as $CaSO_4$, $CaHCO_3$ and complexing of fluoride with aluminium, beryllium, ferric ions, and series of mixed fluoride hydroxide complexes with boron affect fluoride contamination (35). The temperature also plays a crucial role in the fluoride content of groundwater because the amount of fluoride ingested by living beings is influenced primarily by air temperature (20). Besides, the dissolution of fluoride from minerals/rocks is a physicochemical process that is also controlled by the temperature of the water and the soil itself. Chand (36) presented a correlation of ambient temperature with fluoride content.

Mechanism

During the process of chemical weathering, dissolution of fluoride species in natural water is controlled by calcium ions and governed by thermodynamic principles. The $CaCO_3$ equilibrium in groundwater plays an important role in this process. The equilibrium constant of calcite can be evaluated from the following reactions (53).

$$\mathrm{CaCO_3} + \mathrm{H^-} \Longleftrightarrow \mathrm{Ca^{2+}} + \mathrm{HCO_3}^-$$

$$K_{\mathrm{CaCO3}} = [\mathrm{Ca^{2+}}][\mathrm{HCO_3}^-]/[\mathrm{H^+}] = 97 \text{at } 25\,^\circ\mathrm{C} \qquad (1)$$

$(CaCO_3$ is constant)

The fluoride equilibrium is given by (54):

$$\mathrm{CaF_2} \Longleftrightarrow \mathrm{Ca^{2+}} + 2\mathrm{F^-}$$

$$K_{\mathrm{CaF2}} = [\mathrm{Ca^{2+}}][\mathrm{F^-}]^2 = 10^{-10.58} \text{ at } 25\,^\circ\mathrm{C} \qquad (2)$$

(CaF is constant)

Dividing the first equation by the second, the solubility of calcite and fluorite can be represented by a third constant K:

$$K = \frac{[\mathrm{HCO_3}^-]}{[\mathrm{H^+}][\mathrm{F^-}]^2}$$

$$\text{or} \quad [\mathrm{F^-}]^2 = \mathrm{K'}\frac{[\mathrm{HCO_3}^-]}{[\mathrm{H^+}]} \text{ (where } \mathrm{K'} = 1/\mathrm{K})$$

$$\text{or} \quad [\mathrm{F^-}] \propto [\mathrm{HCO_3}^-]/[\mathrm{H^-}]$$

It is evident that the activity of fluoride is directly proportional to the bicarbonate ion at constant pH. Thus, according to the principle of ionic product, if the concentration of calcium and fluoride in water exceeds the solubility product of fluorite ($10^{-10.58}$ at 25 °C), CaF_2 precipitates. Before reaching saturation the calcium ion has a positive correlation with fluoride ion, and after this stage, there will be a negative correlation between calcium and fluoride ions. In fact, the total concentration of fluoride in a solution will be somewhat greater due to the presence of other electrolytes (ionic strength and complexing effects). But it appears that a high fluoride concentration is more likely to occur in water of low calcium concentration. Handa (55), Das (56), and Gupta et al. (57) summed up the characteristics of ground water whose fluoride content is from dissolved minerals, as follows:

- negative correlation of calcium and fluoride ions
- positive correlation of bicarbonate and fluoride ions
- close of saturation with respect to calcium fluoride
- saturated with calcium carbonate

Thus, the dissolution mechanism of fluoride from its minerals and rocks can be explained on the basis of the solubility product. The dissolution mechanism is physicochemical and, therefore, is governed by thermodynamic parameters, too. The important parameters are temperature, pH, ionic strength, and pressure.

REMEDIAL MEASURES

There are two types of remedial measures to control fluoride contamination. The first is to control fluoride contamination in groundwater, and the second includes removal of fluoride from fluoride-containing water. The control of fluoride contamination in groundwater is very difficult because the contamination of fluoride in groundwater is controlled by a number of hydrogeologic and physicochemical parameters. However, various artificial recharge techniques, including the aquifer storage recovery (ASR) technique may be applied to improve the quality of water by dilution.

The ASR technique is being followed in many parts of the world. In this technique, water is stored underground in wells when it is available, and this water is recovered from the same wells when needed to meet peak, long-term, and emergency water needs. The technique is being applied in United States, Canada, England, Australia, Israel, and other countries. The technique has proved to be a viable, cost-effective option for storing large volumes of fresh water not only in fresh, but also in brackish and other nonpotable aquifers at depths as low as 900 m. Most ASR sites store drinking water in confined aquifers containing water that is brackish or contains constituents such as nitrates, fluorides, iron, manganese, and hydrogen sulfide, all unsuitable for drinking purposes except following treatment. Mixing between the drinking water and the native water in the aquifer can be controlled in most situations by the proper design and operation of ASR wells, so that recovered water has acceptable quality. The operation includes development of a buffer zone surrounding the ASR well to contain the stored water and development of a target storage volume for each well so that recovered water will meet flow, volume, and water quality criteria with acceptability. This technique, however, still remains to be tried in India.

Excess withdrawal of groundwater should be avoided to the extent possible. Aquifers should be recharged periodically so that air cannot enter the aquifer. In addition, only those types of raw materials should be used in industries, which do not release fluoride into the environment. In addition to this, contamination through nonpoint sources should also be minimized by checking man-made activities and the use of fluoride-containing fertilizers and pesticides.

The second type of remedial measure includes removal of fluoride from fluoride-containing water (28,58). There are several methods that have been advocated for defluoridation of drinking water. These methods can be broadly divided into two categories, those based upon the addition of some chemical to the water during the softening or coagulation processes and those based upon ion exchange or adsorption processes. Adsorption or ion exchange processes are recommended for treating low concentrations. These processes are performed by using lime and alum, bone char and synthetic bone, activated carbon and bauxite, ion exchange, activated alumina, and reverse osmosis.

CONCLUSION

The problem of high fluoride concentration in groundwater resources has now become one of the most important health-related geoenvironmental issues in India influenced by the regional and local geological and hydrological conditions of the region. It is high time that an affordable solution is found to minimize fluoride contamination to maintain the health of the large population of the country. There is an immediate need to defluoride the water system either by community or by domestic defluoridation techniques. Demonstration-cum-awareness camps for the purpose should be arranged in fluorosis endemic areas. There is a need to carry out detailed fluoride mapping, hydrological studies for existing water sources to show flow lines, and hydrogeochemical surveys where fluorosis is endemic. In the affected areas, the government should apply firm guidelines for using groundwater, so that tube wells and/or hand pumps in high fluoride zones can be discouraged. Short-term solutions to minimize the fluoride level in drinking water could be using domestic defluoridation equipment or filters.

BIBLIOGRAPHY

1. Shortt, H.E. (1937). Endemic fluorosis in the Madras presidency. *Ind. J. Med. Res.* **25**: 553–561.
2. Sudarshan, V. and Reddy, B.R. (1991). Pollution of fluoride in groundwater and its impact on environment and socioeconomic status of the people—a case study in Sivannagudem area. *Indian J. Environ. Prot.* **11**(3): 185–192.
3. *Rajiv Gandhi National Drinking Water Mission*. (1994). New Delhi, Vol. 1, p. 20
4. Datta, P.S. (1996). Fluoride in ground water in Delhi area. *J. Contam. Hydrol.* **24**(1): 85–96.
5. Sengupta, A. (1999). *Proc. Workshop on Ground Water Pollution and Protection with Special Reference to Arsenic Contamination*. Central Ground Water Board, Science City, Calcutta, pp. 69–74.
6. Chakraborti, D. and Bhatta, A. (1999). *The Statesman*. 22 August 1999, pp. 8–9.
7. Susheela, A.K. (1999). Fluorosis management programme in India. *Curr. Sci.* **77**(10): 1250–1256.
8. Chakraborti, D. et al. (2000). Fluorosis in Assam, India. *Current Science* **78**(12): 1421–1423.
9. Susheela, A.K., Kumar, A., Bhatnagar, M., and Bahadur, R. (1993). Prevalence of endemic fluorosis with gastro-intestinal manifestation in people living in some north Indian village. *Fluoride* **26**(2): 94–104.
10. Vasavada, B.J. (1998). Excessive fluoride in Gujarat—a perspective plan of solution. *J. Indian Water Works Assoc.* **30**(3): 191–198.
11. BIS. (1991). Drinking Water—Specifications. Bureau of Indian Standards, New Delhi.
12. WHO. (1996). Guidelines for Drinking Water Quality, Vol. 2. World Health Organization, Geneva.
13. Swamy, N.K. (1990). Control of fluorosis—An interdisciplinary approach. *Indian J. Environ. Prot.* **10**(3): 181–185.
14. Garg, V.K., Dahiya, S., Chaudhary, A., and Shikha, D. (1998). Fluoride distribution in under ground water of Jind district, Haryana, India. *Ecol. Env. & Cons.* **4**(1–2): 19–23.
15. Garg, V.K., Sharma, I.S., and Bishnoi, M.S. (1998). Fluoride in ground water of Ulkana town, district Hisar. Haryana. *Poll. Res.* **17**(2): 149–152.
16. Jain, C.K., Ali, I., and Sharma, M.K. (1999). Fluoride contamination in ground water—Indian scenario. *Indian J. Environ. Prot.* **19**(4): 260–266.
17. Agrawal, V., Vaish, A.K., and Vaish, P. (1997). Groundwater quality: Focus on fluoride and fluorosis in Rajasthan. *Curr. Sci.* **73**(9): 743–746.
18. Gupta, S.C. and Vijeya, R.P. (1988). Some inherent problems of groundwater quality in western Rajasthan. *Trans. Isdt.* 97–103.
19. Gupta, S.C. (1991). Chemical characteristics of ground waters in Nagpur district, Rajasthan. *Indian J. Environ. Hlth.* **33**(3): 341–349.
20. Wodeyar, B.K. and Sreenivasan, G. (1996). Occurrence of fluoride in the ground water and its impact in Peddavankahalla basin, Bellary district, Karnataka—a preliminary study. *Curr. Sci.* **70**(1): 71–73.
21. Shukla, N., Moitra, J.K., and Trivedi, R.C. (1995). Fluoride level in exfoliated human teeth from 2 contrast environmental regions in India. *Indian J. Environ. Prot.* **15**(12): 903–905.
22. Apparao, B.V. and Karthikeyan, G. (1986). Permissible limits of fluoride in drinking water in Indian rural environment. *Indian J. Environ. Prot.* **6**(3): 172–175.
23. Mariappan, P., Vasudeven, T., and Yegnaraman, V. (1998). Surveillance of ground water in Thiruppathur block of Sivagangai district. *Indian J. Environ. Prot.* **19**(4): 250–254.
24. Mariappan, P., Vasudeven, T., and Yegnaraman, V. (1999). Fluoride distribution in ground waters of Salem district. *Indian J. Environ. Prot.* **19**(9): 681–683.
25. Mariappan, P., Yegnaraman, V., and Vasudeven, T. (1999). Correlation between fluoride and alkalinity in ground waters of fluorosis endemic—Salem district. *Indian J. Environ. Prot.* **20**(3): 182–187.
26. Mariappan, P., Yegnaraman, V., and Vasudeven, T. (1999). Control of fluoride content in the ground water—Scope for an interdisciplinary approach. *Poll. Res.* **20**(6): 421–426.
27. Mariappan, P., Yegnaraman, V., and Vasudeven, T. (2000). Correlation between water table and fluoride content in the ground waters of Salem district, Tamil Nadu. *Poll. Res.* **20**(3): 182–187.
28. Mariappan, P., Yegnaraman, V., and Vasudeven, T. (2000). Occurrence and removal of fluoride in ground water. *Poll. Res.* **19**(2): 165–177.
29. Malhotra, S.K., Verma, M., Verma, A.N., and Kulshreshtha, K. (1997). Fluoride and fluorosis in India—3. Osteo-dental fluorosis in school going children in Allahabad city (U.P.). *Chemistry & Biology of Herbal Medicine*: 203–210.
30. Malhotra, S.K., Verma, A.N., and Kulshreshtha, K. (1998). Fluoride and fluorosis in India—I. Drinking water source contamination in Utter Pradesh. *Poll. Res.* **17**(4): 311–315.

31. Gupta, M.K., Singh, V., Rajwanshi, P., Srivastava, S., and Dass, S. (1994). Ground water fluoride levels in a rural area of district Agra. *Indian J. Environ. Prot.* **14**(5): 370–372.
32. Gupta, M.K., Singh, V., Rajwanshi, P., Srivastava, S., and Dass, S. (1994). Fluoride in ground water at Agra. *Indian J. Environ. Hlth.* **36**(1): 43–46.
33. Shrivastav, R. and Choudhary, B. (1997). Drinking water quality in an average Indian city: A case study of Agra (U.P.). *Poll. Res.* **16**(1): 63–65.
34. Das, S., Mehta, B.C., Das, P.K., Srivastava, S.K., and Samanta, S.K. (1998). Source of high fluoride in ground water around Angul, Dhenkenal district, Orissa. *Poll. Res.* **17**(4): 385–392.
35. Das, S., Mehta, B.C., Das, P.K., Srivastava, S.K., and Samanta, S.K. (1999). Sources of high fluoride in ground water around Angul, Dhenkenal district, Orissa. *Poll. Res.* **18**(1): 21–28.
36. Chand, D. (1999). Fluoride and human health—cause for concern. *Indian J. Environ. Prot.* **19**(2): 81–89.
37. Samal, U.N. (1988). Dental fluorosis in school children in the vicinity of aluminium factory in India. *Fluoride* **21**(3): 137–141.
38. Susheela, A.K. (1991). *Prevention and Control of Fluorosis. Technical Information for Training cum Awareness Camp for Doctors, Public Health Engineers and other Officers*. Published by National Technology Mission of Drinking Water, New Delhi.
39. Sun, Z., Cheng, Y., Zhou, J., and Wei, R. (1998). Research on the effect of fluoride pollution in atmosphere near an aluminium electrolysis plant on regional fall wheat growth. *Proc. Annu. Meet., Air Waste Manag.*, Assoc. 91st TPE 09/P1-TPE 09P/7.
40. Clarke, M.L., Harvey, D.G., and Humphrevs, D.J. (1981). *Veterinary Toxicology*, 2nd Edn. Bailliere Tindal & Cassle Ltd., London, pp. 48–54.
41. Suresh, T. (1996). Fluoride in ground water of Chiknaya Kanahalli Taluk, Tumkur district, Karnataka. *Proc. Workshop on Challenges in Groundwater Development*, Madras, pp. 185–187.
42. Handa, B.K. (1979). Effect of return irrigation flows from irrigated land on the chemical composition of ground water from shallow unconfined aquifers. *Prog. Water Tech.* **11**: 337–349.
43. EPA Report. (1976). *National Interim Primary Drinking Water Regulation*. EPA Publication No. EPA-570/9-76-003.
44. Piekos, R. and Paslawska, S. (1998). Leaching characteristics of fluoride from coal fly ash. *Fluoride* **31**(4): 188–192.
45. Chary, V., Rao, R.J., and Naidu, M.G.C. (1975). Fluoride content of some raw vegetable foods available at Podile, Prakasam district, Andhra Pradesh. *Proc. Symp. Fluorosis*, Hyderabad, India, pp. 144–150.
46. Lakdawala, D.R. and Punekar, B.D. (1973). Fluoride content of water and commonly consumed foods in Bombay and supply of dietary intake. *Ind. J. Med. Res.* **16**: 1679–1687.
47. Nanda, R.S. and Kapoor, K (1971). Fluoride content of pine and betel and its constituents. *Ind. J. Med. Res.* **59**: 1966–1968.
48. Sengupta, S.R. and Pal, B. (1971). Iodine and fluoride contents of foodstuffs. *Ind. J. Nutr. Dicter.* **8**: 66–71.
49. Gradien, P. (1992). Cancer incident and mortality in workers exposed to fluoride. *J. Mat. Cancer Inst.* **184**: 1903–1909.
50. Rao, N.S., Rao, J.P., Rao, B.N., Babu, P.N., Reddy, P.M., and Devadas, D.J. (1998). A preliminary report on fluoride content in ground water of Guntur area, Andhra Pradesh, India. *Curr. Sci.* **75**(9): 887–888.
51. Somani, I.C., Gandhi, A.P., and Palwal, K.V. (1972). Note on toxicity of fluorine in well waters of Nagpur and Jaipur districts of Western Rajasthan. *Ind. J. Agric. Sci.* **48**(2): 752–754.
52. Maithani, P.B., Gurjar, R., Banerjee, R., Balaji, K.B., Ramachandran, S., and Singh, R. (1998). Anomalous fluoride in ground water from western part of Sirohi district, Rajasthan and its crippling effects on human health. *Curr. Sci.* **74**(9): 773–776.
53. Hem, J.D. (1970). *Study and Interpretation of the Chemical Characteristics of Natural Waters*. USGS Water Supply, Paper 1475, pp. 363–369.
54. Brown, D.W. and Roberson, C.E. (1977). Solubility of natural fluorite at 25 °C. *USGS J. Res.* **5**(4): 506–517.
55. Handa, B.K. (1975). Geochemistry and genesis of fluoride containing ground water in India. *Ground Water* **13**(3): 275–281.
56. Das, D.K. (1985). Incident of high fluoride in deep ground water in Betwa basin, Madhya Pradesh, Central India. *G.S.I. Res.* **116**(2): 23–30.
57. Gupta, S.C., Doshi, C.S., and Paliwal, B.L. (1986). Occurrence and chemistry of high fluoride ground water in Jalore district of Western Rajasthan. *Annal of Arid Zone* **25**(4): 255–265.
58. Jain, A.K. (1997). *Akalanka's Water Supply and Treatment: Central Public Health and Environmental Engineering Organizations: Manual on Water Supply and Treatment*. Akalank Publications, New Delhi, India, pp. 203–208.

ROCK FRACTURE

Nitish Priyadarshi
Ranchi University
Ranchi, Jharkhand, India

INTRODUCTION

From their origin, rocks are more or less disturbed by forces acting within the lithosphere. When a mass of rock is not strong enough to resist forces that are tending either to compress it or to stretch it, the rock is deformed. The change of form is brought about by flow in the deeper parts of the lithosphere and fractures in the upper parts (1). Fractures in rocks are either joints or faults.

JOINTS

The term joint is most commonly used in reference to relatively continuous and through-going fractures that are reasonably planar and along which there has been imperceptible movement. Joints may form as a result of either diastrophism or contraction.

Classification of Joints

Theoretically, joints may be classified according to whether they have been formed by compression or tension. Joints due to compression are (1) diagonal joints, (2) irregular cracks induced by the expansion of rocks consequent upon chemical alteration, and (3) probably a majority of the regular joint system in stratified rocks. Tension joints include (1) irregular cracks formed in the shrinkage accompanying certain kinds of rock alteration; (2) cross-joints in igneous rocks; (3) hexagonal columnar structures and associated fractures due to cooling; (4) small local fractures; (5) fractures clearly associated with tension faulting; and (6) cracks due to drying of muds, clays, and argillaceous limestones.

Description of Joints

1. A series of parallel joints is called a "joint set."
2. Two or more joints intersecting each other produce a "joint system."
3. A persistent joint or set that may be horizontal or vertical is called a "master joint."

Probably all consolidated rocks and a good share of unconsolidated deposits contain joints. Although not well recognized by most individuals involved in groundwater problems, joints exert a major control on water movement and chemical quantity. Characteristically, joints are open and serve as major conduits or pipes. Water can move through them quickly, perhaps carrying contaminants, and, being open, the filtration effect is lost. It is a good possibility that the outbreak of many waterborne diseases that can be traced to groundwater supplies result from the transmission of infectious agents through fractures to wells and springs.

Most joints, at least initially, are tight fractures, but because of weathering, the joint may be enlarged into an open fissure; this is especially common in limestone regions.

Effect of Joints

Knowledge of joints is important in many kinds of geologic studies. Quarry operations, especially those involved in obtaining blocks of certain dimensions and sizes, are obviously greatly influenced by joints. Closely spaced horizontal joints are obviously of great concern in tunneling. A large joint dipping into a highway cut is the site of a potential landslide. Wells drilled in granites for water supply are more productive in highly jointed rocks than in less jointed rocks (2).

If joints are too numerous (i.e., more sets), closely spaced, and of great magnitude, then such a fractured site will be physically too weak to withstand the stresses of dams and bridges. Saturation with water along with the accompanying decay of rocks will make the site more unsuitable for foundations.

Value of Joints

Although joints are often difficult to interpret, they are nonetheless very important structures. For ages, quarry workers have taken advantage of joint-controlled planes of weakness in removing building blocks of granite and limestones from bedrock. These fracture weaknesses exert profound control on weathering and erosion and, thus, on fashioning landscape. Many scenic attractions owe much of their uniqueness to weathering and erosion of horizontal layers of systematically broken up, steeply dipping joints.

Beyond their scenic value, joints constitute structures, of indisputable geologic and economic significance. Joints invite circulation of fluids, including rain and groundwater, hydrothermal mineralizing solutions, and oil and gas. As cracks in rocks, joints can be thought of as structures that significantly contribute to the bulk porosity and permeability of rocks.

Explorers appreciate the benefits of the circulation of fluids through jointed rocks. Petroleum geologists evaluate the nature and degree of development of joints as one guide to the reservoir quality of sedimentary formations. To increase the yield of reservoir rocks in oil and gas fields where production is waning, it is common practice to "crack" the rocks artificially, either by explosives or by high-pressure pumping of fluids into wells.

Joints can serve as sites of deposition of metallic and nonmetallic minerals. In almost all hydrothermal deposits, a part of the mineralization is localized in and around joints. The minerals are deposited either through open-space filling of joints or through selective replacement of chemically favorable rocks adjacent to the joint surfaces along which hydrothermal fluids once circulated.

FAULTS

A fault may be defined as a fracture along which there has been slipping of contiguous masses against one another. Points formerly in contact have been dislocated or displaced along the fracture. Solid rocks or unconsolidated sands, gravel, etc. may be dislocated in this way. Faulting may result from compression, tension, or torsion. Some faults in loose or weakly consolidated clays, sands, and gravels are produced by the removal of a support.

Faults are most common in the deformed rocks of mountain ranges, suggesting either lengthening or shortening of the crust. Movement along a fault may be horizontal, vertical, or a combination. The most common types of faults are called normal, reverse, and lateral (Fig.1). A normal fault, which indicates stretching of the crust, is one in which the upper or hanging wall has moved down relative to the lower or foot wall. The Red Sea, Dead Sea, and the large lake basins in the East African highlands, among many others, lie in a graben, which is a block bounded by normal faults. A reverse or thrust fault implies compression and shortening of the crust. It is distinguished by the fact that the hanging wall has moved up relative to the foot wall. A lateral fault is one in which the movement has been largely horizontal. The San Andreas Fault, extending some 600 miles from San Francisco Bay to the Gulf of California, is the most notable lateral fault in the United States. Movement along this fault produced the 1906 San Francisco earthquakes.

Recognition of Faults in the Field

To recognize faults in the field, a number of criteria are used. The faults may be directly seen in the field, particularly in artificial exposures such as river cuttings and road cuttings. In most of cases, faults are recognized by stratigraphic and physiographic evidence such as (2)

1. repetition or omission of strata,
2. discontinuity of structures,
3. features characteristic of fault planes,
4. silicification and mineralization,

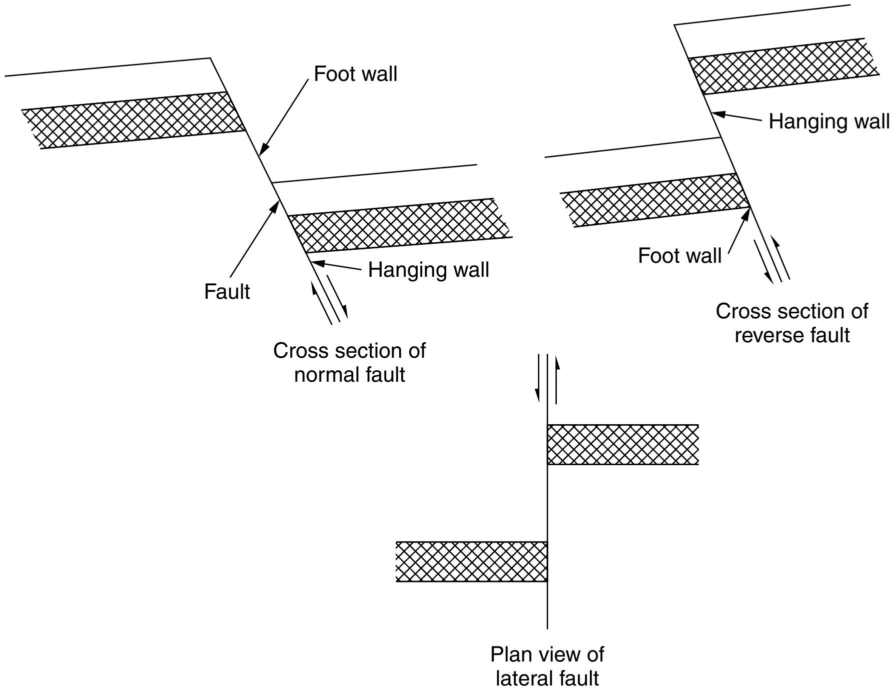

Figure 1. Cross sections of normal, reverse, and lateral faults.

5. sudden changes in sedimentary facies, and
6. physiographic data.

Effects of Faulting

Faulted areas are neither safe nor stable for the foundation of civil engineering works because of the various harmful effects produced by faults. Some important effects of faulting follow (3):

1. Faults cause considerable fracturing and shattering of rocks along fault zones, which means that they are not compact, massive, or strong. Such places are reduced to physically very weak grounds and hence are unfit as foundation sites to withstand heavy loads of structures such as dams.
2. When such porous and fractured zones are saturated with water, their strength is reduced further.
3. The same fractures act as channels for movement of groundwater, which may cause severe groundwater problems in tunnels and leakage problems in reservoirs.
4. The most dangerous features of faulting are its possible recurrence at the same place, which means that the faulted ground is unstable as long as faulting remains active there. Vertical, lateral, or rotational movements are likely to take place at the time of renewed faulting. Naturally, under such conditions, any civil engineering structure cannot have a safe or stable foundation. And if constructions are made, they are likely to collapse when renewed ground movements occur.
5. The fault plane itself is a very prominent fracture plane in the fault zone and therefore may act as a severe source of water leakage. When such percolated water reaches underground, it decomposes the shear zone or fault zones. Such weathering further reduces the competence of rocks.

BIBLIOGRAPHY

1. Lahee, F.H. (1987). *Field Geology*. CBS, New Delhi, p. 222.
2. Billings, M.P. (1984). *Structural Geology*. Prentice-Hall of India, New Delhi, pp. 140–200.
3. Kesavulu, N.C. (1997). *Textbook of Engineering Geology*. Macmillan India, New Delhi, pp. 241–242.

GEOCHEMICAL MODELS

Mário Abel Gonçalves
Faculdade de Ciências da
Universidade de Lisoba
Lisbon, Portugal

INTRODUCTION

A geochemical model, like any other model, is an abstract representation of a given reality, normally reduced to a set of master variables and described by mathematical equations, which aim at representing natural processes that occur in a system. The output data of these models are a quantitative representation of an outcome that can be observable in the natural system or subject to experimental validation. This operational definition of a geochemical model suggests that a model is nothing more than a set of mathematical equations, which is not strictly true. Any set of equations representing a process must be bounded within limits imposed by nature; otherwise the outcome of these models may be totally unrealistic. Thus, the modeler is called to set proper initial and boundary conditions such that the natural system is correctly represented, and to feed the model with the correct parameters, most of them previously determined experimentally. The quality and self-consistency of the thermodynamic data used, as well as other parameters, such as kinetic ones, are of fundamental importance for the outcome of geochemical models. This issue is considered one of the most critical in any geochemical model. When using any of the available computer programs for geochemical modeling, the choice of the thermodynamic database is left to the modeler. Several compiled thermodynamic databases are available, but this does not mean that all data is internally self-consistent. It is, however, possible to find some databases that are self-consistent relative to some set of chemical species.

Geochemical models have been extensively reviewed in the literature, such as Yeh and Tripathi (1), Mangold and Tsang (2), Appelo and Postma (3), Nordstrom and Munoz (4), and Nordstrom (5). Some textbooks on aqueous geochemistry or geochemical modeling also discuss and include several examples and case studies where specific geochemical computer models have been used (6–8). Of relevance is also the book of Albarède (9), which provides a wealth of mathematical methods and extensive examples on how to build and develop geochemical models in a truly wide range of applications.

COMPONENTS OF GEOCHEMICAL MODELS

Geochemical models may have several components that can be combined in different configurations. These components may be coupled within the model or may imply certain feedback loops.

An essential component of these models is chemical reactions, and these determine, for example, the chemical speciation in solution or the saturation states relative to solid phases. Within these reactions, biological processes may be involved, which take active part in certain reactions, boosting their kinetics (catalysts), hindering the formation of certain compounds (inhibitors), or just transforming chemical compounds (such as the biodegradation of organic pollutants).

Chemical species and compounds in solution are carried away with the water by advection and disperse through the medium by molecular diffusion. These are the components of transport of chemical elements in solution, which determine their spreading in the system.

Transport of chemical elements are thus a function of water velocity and, consequently, of fluid flow in the system. Physical parameters of the system, such as porosity and permeability, determine the patterns of fluid flow and velocity variations in space and time. In the coupled reactive transport models, precipitation/dissolution of mineral phases can reduce/increase the porosity and permeability of the medium and change fluid flow patterns.

In certain systems, such as large-scale sedimentary basins, it is necessary to consider heat transport. Heat can increase the kinetics of several chemical reactions and induce fluid flow along thermal gradients.

TYPES OF GEOCHEMICAL MODELS

The description of the different types of geochemical models is not extensive and outlines only their main characteristics, as presented by Zhu and Anderson (8). Geochemical models are generally grouped according to their level of complexity. The simplest ones are the speciation-solubility models. These models are meant to compute the thermodynamic equilibrium of species in a system at a given temperature and pressure. Therefore, the output comprises the concentration and activity of the various ionic and molecular species in a solution. It also includes the saturation state of the solution relative to several minerals and the distribution of stable species on surfaces or ion-exchange sites in equilibrium with the aqueous solution.

Reaction-path models calculate the sequence of equilibrium states of a system in response to incremental additions (or subtractions) of mass to the system, change in temperature and/or pressure, and mass transfer between phases in the system. The configuration of these models can be diverse and includes the addition of a reactant (such as a titration), fixation of the activity of a chemical species modeling a buffered system, incremental feeding of a reactant solution (as in a continuous stirred tank reactor), and kinetic controls of heterogeneous reactions.

Another group of models corresponds to inverse mass balance models. These specialized models derive the initial composition of a water solution from its actual final composition, which takes into account the reactions and mass transfer between water and solid and/or gas phases, in agreement with the available data of a system. Thus, the initial composition of the water is determined by subtracting the amount of dissolved species caused by reaction with minerals and other phases in the system from its final composition. Inverse mass balance calculations may also involve the determination of the fractions of different waters that have, at some given time, mixed completely.

Finally, coupled reaction-transport models are the most complex. In these models, both the partial differential equations describing the advection-dispersion transport and the set of algebraic equations describing the chemical equilibrium are solved. These models can also include heat transfer and fluid flow, thus increasing the number of equations to be solved. The level of complexity depends also on the details of chemical reactions considered. These details can include multicomponent reactive transport, which accounts for the kinetics of mineral dissolution and precipitation; adsorption onto mineral surfaces; and radioactive decay, to name but a few.

FINAL REMARKS: MODEL VALIDATION AND USEFULNESS

The outcome of geochemical models can be either observable in nature or subject to experimental testing. Both processes are fundamental for model validation, and they are surely the ultimate test that a model must face. However, the process is not as simple as it may seem. Usually, geochemical models may adequately describe several processes and mechanisms in nature, but nature's inherent complexity puts a limit to model precision and accuracy, which limits considerably its proper validation. Normally, the number of variables assumed within a model is limited and corresponds to a fraction of the ones found in nature. In such complex models, slight variations in parameters may induce diverse outcomes, such as in climate modeling. Thus, models are not only helpful tools to gain insight into the workings of nature, but they also must have some sort of predictive power. A model outcome may not be accurate enough to make a prediction relative to contaminant dispersion in a groundwater system, for example. However, it may give enough confidence to help make decisions on regulatory issues.

BIBLIOGRAPHY

1. Yeh, G.T. and Tripathi, V.S. (1989). A critical evaluation of recent developments in hydrogeochemical transport models of reactive multichemical components. *Water Resour. Res.* **25**: 93–108.
2. Mangold, D.C. and Tsang, C.-F. (1991). A summary of subsurface hydrological and hydrochemical models. *Rev. Geophys.* **29**: 51–79.
3. Appelo, C.A.J. and Postma, D. (1993). *Geochemistry, Groundwater and Pollution*. A.A. Balkema, Rotterdam.
4. Nordstrom, D.K. and Munoz, J.L. (1994). *Geochemical Thermodynamics*, 2nd Edn. Blackwell, Boston, MA.
5. Nordstrom, D.K. (2004). Modeling low-temperature geochemical processes. In: *Treatise on Geochemistry—Surface and Ground Water, Weathering, and Soils*. J.I. Drever (Ed.). Vol. 5, pp. 37–72.
6. Langmuir, D. (1996). *Aqueous Environmental Geochemistry*. Prentice-Hall, Upper Saddle River, NJ.
7. Drever, J.I. (1997). *The Geochemistry of Natural Waters*, 3rd Edn. Prentice-Hall, Upper Saddle River, NJ.
8. Zhu, C. and Anderson, G. (2002). *Environmental Applications of Geochemical Modeling*. Cambridge University Press, Cambridge, UK.
9. Albarède, F. (1996). *Introduction to Geochemical Modeling*. Cambridge University Press, Cambridge, UK.

GEOCHEMICAL MODELING-COMPUTER CODES

Mário Abel Gonçalves
Faculdade de Ciências da
Universidade de Lisoba
Lisoba, Portugal

An important share of geochemical studies increasingly relies on the use of computer programs[1] to model diverse geochemical systems. Most of these programs are freely available to the general public, or at a symbolic cost for educational and research purposes. Although developing open source and/or precompiled codes is markedly important, their regular maintenance should not be dismissed. Programming languages evolve, as well as operating systems and computer hardware, which means that codes without regular revision become outdated and eventually useless as their compiled versions may stop working properly under new operating systems.

All geochemical models rely very much on the availability of good quality, self-consistent thermodynamic data. This data is stored in database files that are accessed by the program while it is executed, making it one of its core elements. In the absence of specific data in the database for the problem to be modeled, some programs allow the incorporation of the data in simulations or, alternatively, the database can be modified by incorporation of new data. As new and improved experimental thermodynamic measurements are continuously being made, thermodynamic databases should also be regularly revised and updated.

The description that follows is meant to address mainly those computer programs most readily accessible and does not pretend to be an exhaustive list of all available programs. Although presenting the address of websites where these programs and codes are stored and may be obtained, one should be aware that this information will potentially become out of date rather quickly.

USGS CODES

The USGS supports various projects for developing software, including aqueous geochemistry computer programs, which include the chemical speciation program WATEQ4F (1), well suited for processing large numbers of water analyses. The most recent upgrades include the revision of the thermodynamic data on uranium and arsenic species. The most complete computer programs available are the ones from the PHREEQC (2) family (http://wwwbrr.cr.usgs.gov/projects/GWC_coupled/).
PHREEQC is a program that performs chemical speciation calculations, reaction-path modeling, one-dimensional transport, and inverse geochemical calculations. Currently, it contains a basic interpreter allowing a very flexible use of the program, meeting each one's needs, especially for modeling kinetic data. The latest revisions (February–April 2003) include isotope fractionation modeling (3). PHREEQC uses its own thermodynamic database, and also the LLNL and WATEQ4F databases, which are still updated and corrected regularly. Two graphical user interfaces (GUI) were developed: PHREEQCI by USGS and PHREEQC for Windows by Vincent Post from the Vrije Universiteit Amsterdam (http://www.geo.vu.nl/users/posv/phreeqc). The latter allows the graphical display of the output, which is unavailable in the original program, which is rather achieved by using the PHRQCGRF program. PHAST is a three-dimensional multicomponent reaction-transport model that simulates transient groundwater flow, that may or may not include geochemical reactions. PHAST combines the HST3D simulator (4) for the transport calculations with PHREEQC for geochemical calculations. PHRQPITZ is specially designed to use with brines, as it implements Pitzer's equation for the calculation of activity coefficients.

---

[1] We will use the term "computer code" as a set of written instructions aiming at solving a set of specific problems, and computer program to a compiled code to be executed as a standalone application under a given operating system.

Other computer programs include OTIS (5), used for the geochemical modeling solute transport in streams and rivers. Recently, Bowser and Jones (6) presented a Microsoft® Excel spreadsheet for a mineral-solute mass-balance model in order to study and understand the mineralogical controls on water composition in surface and groundwater systems dominated by silicate lithologies. All of these programs, and others, are available from USGS webpages at http://water.usgs.gov/software/geochemical.html.

USEPA CODES

The USEPA has a series of supported computer codes, the most popular of which is the MINTEQA2/PRODEFA2 (last release in 1999 is version 4.0), widely used in environmental geochemistry problems (7), which is a chemical equilibrium computer model that is able to calculate chemical speciation, solubility equilibrium, titration, and surface complexation modeling. It also includes the Gaussian model for the interaction of dissolved organic matter (DOM) with cations. However, it lacks database maintenance. Gustafsson (8) has been developing VisualMINTEQ, a GUI version of this program that also presents other improvements, such as the NIST database, adsorption with five surface complexation models, ion-exchange, and inclusion of both the Stockholm Humic Model and the NICA-Donnan model for metal-DOM complexation to name only a few.

The program BIOPLUME III (9) is a 2-D finite difference model that accounts for advection, diffusion, adsorption, and biodegradation in groundwater systems to model natural attenuation of organic contaminants. BIOCHLOR (10) and BIOSCREEN (11) are both Microsoft® Excel spreadsheet-based codes that model natural attenuation of chlorinated solvents and petroleum-derived hydrocarbons in water systems, respectively.

CHEMFLO-2000 (12) is a model that simulates water flow and chemical transport and fate in the vadose zone. CHEMFLO-2000 is a program that is written in Java®, which makes it platform-independent. All of these programs can be obtained from http://www.epa.gov/ada/csmos/models.

OTHER CODES

The set of computer codes known as EQ3/6 (13) supported by Lawrence Livermore National Laboratory (LLNL) was originally developed to model water-rock interactions in hydrothermal systems. It is currently one of the most complete programs applied to several problems, including municipal and industrial waste situations, and has been used to assess natural and engineered remediation processes. Unlike the programs presented until now, it must be purchased from LLNL. Closely related but mostly used for a range of high temperature and pressure is SUPCRT92 (14). This program has been discontinued, but still available on request to the authors.

The Geochemist Workbench® (15) is a commercial software with a range of capabilities similar to EQ3/6 and PHREEQC. It is available for Windows® only, but taking advantage of this environment makes it user friendly, with graphical capabilities included that are normally absent from most of these programs.

The Windermere Humic Aqueous Model (WHAM) version 6 (16) models the ion–humic substances interaction in surface waters using surface complexation. It also incorporates cation exchange on clays. However, precipitation and dissolution of solids as well as oxidation-reduction reactions cannot be simulated. This program must also be purchased for use. Alternatively, WinHumicV is a freely available GUI version of WHAM with model V implemented (17).

Steefel and Yabusaki (18) developed the GIMRT/OS3D codes for 2-D and 3-D multicomponent coupled reactive-transport modeling for flow in porous media. Both of these programs were superseded by the program CRUNCH (http://www.csteefel.com/CrunchPublic/WebCrunch.html), which can be obtained from the developer (C. I. Steefel) on request.

ORCHESTRA (19) (http://www.meeussen.nl/orchestra/) represents a new class of computer programs for use in geochemical reactive-transport modeling. This program is actually a framework where chemical speciation models can be implemented by the user and combine them with kinetic and transport processes. It is written in Java® and takes advantage of object-oriented programming. In the same class of programs is MEDIA (http://www.nioo.knaw.nl/homepages/meysman/), to simulate the biogeochemistry of marine and estuarine sediments.

BIBLIOGRAPHY

1. Ball, J.W. and Nordstrom, D.K. (1991). *User's Manual for WATEQ4F, with Revised Thermodynamic Data Base and Test Cases for Calculating Speciation of Major, Trace, and Redox Elements in Natural Waters*. U.S. Geological Survey, Open-File Report 91-183, Washington, DC.
2. Parkhurst, D.L. and Appelo, C.A.J. (1999). *User's Guide to PHREEQC (Version 2)—A Computer Program for Speciation, Batch-Reaction, One-Dimensional Transport, and Inverse Geochemical Calculations*. U.S. Geological Survey Water-Resources Investigations Report 99-4259. Washington, DC, p. 312.
3. Thorstenson, D.C. and Parkhurst, D.L. (2002). *Calculation of Individual Isotope Equilibrium Constants for Implementation in Geochemical Models*. U.S. Geological Survey Water-Resources Investigation Report 02-4172. Washington, DC, p. 129.
4. Kipp, K.L. (1987). *HST3D—A Computer Code for Simulation of Heat and Solute Transport in Three-dimensional Ground-water Flow Systems*. U.S. Geological Survey Water-Resources Investigations Report 86-4095. Washington, DC, p. 517.
5. Runkel, R.L. (1998). *One Dimensional Transport with Inflow and Storage (OTIS): A Solute Transport Model for Streams and Rivers*. U.S. Geological Survey Water-Resources Investigation Report 98-4018. Washington, DC, p. 73.
6. Bowser, C.J. and Jones, B.F. (2002). Mineralogic controls on the composition of natural waters dominated by silicate hydrolysis. *Amer. J. Sci.* **302**: 582–662.
7. Allison, J.D., Brown, D.S., and Novo-Gradac, K.J. (1991). *MINTEQA2/PRODEFA2, A Geochemical Assessment Model for Environmental Systems: Version 3.0 User's Manual*. U. S. Environmental Protection Agency, Washington, DC, p. 107.

8. Gustafsson, J.P. (2004). Visual MINTEQ ver. 2.30. Available: http://www.lwr.kth.se/english/OurSoftware/Vminteq/index.htm.
9. Rafai, H.S., Newel, C.J., Gonzales, J.R., Dendrou, S., Kennedy, L., and Wilson, J.T. (1998). *BIOPLUME III—Natural Attenuation Decision Support System, User's Manual Version 1.0. EPA/600/R-98/010*. USEPA, Washington, DC, p. 282.
10. Aziz, C.E., Newel, C.J., Gonzales, J.R., Haas, P., Clement, T.P., and Sun, Y. (2000). *BIOCHLOR—Natural Attenuation Decision Support System*, User's Manual, Version 1.0. EPA/600/R-00/008. USEPA, Washington, DC, p. 46.
11. Newel, C.J., McLoed, R.K., and Gonzales, J.R. (1996). *BIOSCREEN—Natural Attenuation Decision Support System*, User's Manual, Version 1.3. EPA/600/R-96/087. USEPA, Washington, DC, p. 65.
12. Nofziger, D.L. and Wu, J. (2003). *CHEMFLO-2000—Interactive Software for Simulating Water and Chemical Movement in Unsatured Soils*. EPA/600/R-03/008. USEPA, Washington, DC, p. 69.
13. Wolery, T.J. (1992). *EQ3/6, A Software Package for Geochemical Modeling of Aqueous Systems: Package Overview and Installation Guide (Version 7.0)*. (UCRL-MA-110662 PT I).
14. Johnson, J.W., Oelkers, E.H., and Helgeson, H.C. (1992). SUPCRT92: A software package for calculating the standard molal thermodynamic properties of minerals, gases, aqueous species, and reactions from 1 to 5000 bar and 0 to 1000 °C. *Comput. Geosci.* **18**: 899–947.
15. Bethke, C.M. (1996). *Geochemical Reaction Modeling, Concepts and Applications*. Oxford University Press, New York.
16. Tipping, E. (1998). Humic ion-binding model VI: an improved description of the interactions of protons and metal ions with humic substances. *Aqua. Geochem.* **4**: 3–48.
17. Gustafsson, J.P. (1999). WinHumicV. http://www.lwr.kth.se/english/OurSoftWare/WinHumicV/.
18. Steefel, C.I. and Yabusaki, S.B. (1996). *OS3D/GIMRT, Software for Multicomponent—Multidimensional Reactive Transport*. User manual and programmer's guide. PNL-11166. Battelle, Richland, WA.
19. Meeussen, J.C.L. (2003). ORCHESTRA: An object-oriented framework for implementing chemical equilibrium models. *Environment. Sci. Technol.* **37**: 1175–1182.

READING LIST

Plummer, L.N., Parkhurst, D.L., Fleming, G.W., and Dunkle, S.A. (1988). *A Computer Program Incorporating Pitzer's Equations for Calculation of Geochemical Reactions in Brines: U.S. Geological Survey Water-Resources Investigations Report 88-4153*. Washington, DC, p. 310.

GEOCHEMICAL MODELING—COMPUTER CODE CONCEPTS

Geoffrey Thyne
Colorado School of Mines
Golden, Colorado

INTRODUCTION

This article is focused on geochemical models in aqueous systems. The applications for these models have grown over the last 20 years as the capabilities and flexibility of the codes has increased in conjunction with the increased speed of personal computers. Many earlier computer models were designed for specific questions related to aqueous speciation (1,2). Many earlier models have been discontinued or superseded by newer programs that incorporate new features and capabilities, increased flexibility, and improved input and output options. This change has been driven by the wider use of geochemical modeling and the increase in modeling as a component of environmental studies. Some more important recent applications include modeling high-level radioactive waste disposal, environmental issues associated with mining, landfill leachate, injection of hazardous wastes into deep wells, water resources issues, and artificial recharge to aquifers (3).

All available models use the same basic approach, that of calculating the thermodynamic equilibrium state of a specified system that can include water, solutes, surfaces, and solid and gas phases. These models comprise four major components. They are as follows:

1. Input: specific information that defines the system of interest such as concentrations of solutes, temperature, partial pressure of gases, and composition of solid phases.
2. Equations that are solved by the model.
3. Equilibrium and kinetic formulations between solutes of interest.
4. Output: in tabular or graphic form.

CAPABILITIES AND METHODS

The computer codes require initial input constraints that generally consist of water chemistry analyses, units of the measurement, temperature, dissolved gas content, pH, and redox potential (Eh). The models work by converting the chemical concentrations, usually reported in wt./wt. or wt./volume terms such as mg/kg or mg/L, to moles, and then solving a series of simultaneous nonlinear algebraic equations (chemical reaction, charge balance, and mass balance equations) to determine the activity-concentration relationship for all chemical species in the specified system. The models usually require electrical balance and will force charge balance with one of the components (can be designated), as they solve the matrix of nonlinear equations. The capabilities of modern codes include calculation of pH and Eh, speciation of aqueous species, equilibrium with gases and minerals, oxidation and reduction reactions (redox), kinetic reactions, and reactions with surfaces.

The nonlinear algebraic equations are solved with an iterative approach by the Newton–Raphson method (4). The equations to be solved are drawn from a database that contains equations in the standard chemical mass action form. In theory, any reaction such as sorption of solute to surface that can be represented in this form can be incorporated into the model. These equations represent chemical interactions with reactants on the right and products on the left. Reactions are assumed to reach equilibrium (the point of lowest free energy in the system) when there is no change in concentration on either side.

$$\mathrm{CaCO_3} \leftrightarrow \mathrm{Ca^{2+}} + \mathrm{CO_3{}^{2-}}$$

Note that the arrow in the calcite dissolution example above goes in both ways; that is, the reaction as written is reversible. Once a mineral reaches equilibrium with a solution, adding more mineral will not increase the dissolved concentration because we have already saturated the solution. But removing ions from the right side (e.g., lowering the concentration by dilution with distilled water) will cause more solid to dissolve.

We express this in a mathematical form where, at equilibrium, the ratio of the concentration of reactants (on the bottom) and products (on the top) is equal to K, known as the proportionality constant or distribution coefficient or equilibrium constant.

$$\mathrm{K_{calcite}} = \frac{[\mathrm{Ca^{2+}}][\mathrm{CO_3{}^{2-}}]}{[\mathrm{CaCO_3}]}$$

Kinetic reactions, those involving time, are included by assuming that the chemical reaction will proceed to equilibrium, but at a specified rate. The available kinetic reactions include mineral dissolution and precipitation, redox reactions and microbial growth, and metabolism of solutes. The rate laws used in the codes vary, but all codes with kinetic capabilities include simple first-order rate laws, and they may include more complex rate formulations such as cross-affinity, Michaelis–Menten, and Monod formulations (5).

THERMODYNAMIC BACKGROUND

The formalism that allows us to relate mass actions equations (balanced chemical reactions) to actual solutions is called chemical thermodynamics or, more precisely, equilibrium thermodynamics. The basic idea is that elements, molecules, and compounds all contain some internal energy and that all systems try to reach a state where that energy is minimized (equilibrium). Natural systems, particularly low-temperature systems, do not always reach equilibrium, but they do move in that direction. The internal energy of an element, molecule, or compound is expressed as its enthalpy or internal heat. The free energy of an element is an element, molecule or compound is the sum of its internal heat and its internal tendency toward disorder (entropy).

$$G = H - TS$$

Assuming constant T and P, we get

$$\Delta G = \Delta H - T\Delta S$$

The total free energy of a component in the system is dependent on this inherent energy of an element, molecule, or compound and the amount (concentration). When two or more elements, molecules, and compounds are combined, the result is a reaction that minimizes the energy of the new system, lowering the ΔG. The free energy of a reaction is calculated by

$$\Delta G^0 \text{ rxn.} = \sum \Delta G^0 \text{ products} - \sum \Delta G^0 \text{ reactants}$$

The ΔG^0 values (standard free energy) for many compounds can be found in the back of textbooks such as Drever (6). This way we can calculate a ΔG value for any reaction that for which we can write a balanced chemical equation. The minimum energy state (equilibrium) between the reactants and products is related to the ΔG value by

$$\log K_{\text{rxn.}} = \frac{-\Delta G^0 \text{ rxn}}{2.303\, RT}$$

where R is the universal gas constant in kJ/mole; T is the temperature in Kelvin; and K is the equilibrium constant for the reaction.

COMPUTER MODELS

The computer models are divided into two basic types, speciation models and reaction-path models. In both cases, the models are fundamentally static; that is, no explicit transport function exists; however, some forms of transport can be simulated by manipulation of the models. More complex reaction-transport models that explicitly incorporate transport are briefly described below.

The equilibrium models are speciation models in that they can calculate the speciation (distribution) of aqueous species for any element or compound included in the database. Speciation models calculate activities (chemically reactive concentration), species distribution for elements in the database, saturation indices, and ion ratios at the specified conditions of pH and redox potential (ORP or Eh). Most models allow selection of method of activity calculation (Davies, Debye-Huckel, extended Debye-Huckel, Pitzer). Some models incorporate adsorption, solid phase solutions, and kinetics. Only one model, PHREEQC, has the inverse modeling option. This features uses mass balance constraints to calculate the mass transfer of minerals and gases that would produce an ending water composition given a specified starting water composition (7). This method does not model mass transport; it only calculates and provides statistical measures of fit for possible solutions to the mass balance between starting and ending water compositions.

The next step in complexity is the reaction path (mass transfer) models. The reaction path models use speciation calculation as a starting point and then make forward predictions of changes along the specified reaction path (specified change in T, P, pH, addition of new reactants such as another fluid or solid). The program makes small incremental steps with stepwise addition or removal of mass (dissolution or precipitation), and it can include changes in temperature or pressure along the reaction path. Typical questions posed by modelers include:

- If I change a variable (pH, pe, P_{CO2}), how does system change?
- What happens if I mix water A with water B?
- What is concentration of A^+ in water saturated with mineral AB?

- What happens to the water chemistry *A* if I react the water with mineral *B*?

Limitations exist with any model. The field input data may be corrupt with bad analysis, missing parameter, or electrical imbalance. Speciation models assume equilibrium conditions, which may not be the case. The databases are also a source of uncertainty. They do not always contain all elements or species of interest; the data have some uncertainty, and some data may be inaccurate data (6). Some available codes try and minimize this problem by including several of the most popular databases such as the MINTEQ database (EPA-approved database specializing in metals), WATEQ (USGS database specializing in minerals), and the LLNL database (the most complete database available that is compiled and maintained by Lawrence Livermore National Laboratory). For environmental applications, the limited data for organic compounds remain a concern.

Other limitations include the redox reactions that are of particular importance in metal transport. These reactions are difficult to model correctly because redox reactions may have different rates producing natural systems that are not in redox equilibrium (8). This problem can be addressed by modeling redox reactions as rate-limited (kinetic) formulations if the data are available.

The most complex models explicitly incorporate transport and reaction. The codes couple and solve both the partial differential equations of flow using the advective-dispersion equation and the nonlinear algebraic equations of chemical equilibrium. The general approach is to solve for the reaction term in each cell with the chemical module of the code, and then separately solve for the effects of transport (split-operator approach). The effects of adsorption are solved in the transport module with the retardation portion of the equation (3,8,10). These models are much more complex than are the reaction path models. Presently, only three commonly used models have transport capabilities, HYDROGEOCHEM2, PHREEQC (1D) and the related PHAST (3D) code, and Geochemist's Workbench.

Table 1 lists the most common programs, their sources, and some of the most useful capabilities. The list is not meant to be exhaustive; rather, it offers an overview to serve as a starting point for further investigation. Details of the capabilities of each program can be found on the listed websites or in the manuals. More detailed comparison of these and other models are available in related publications (1,11,12).

BIBLIOGRAPHY

1. Nordstrom, D.K. et al. (1979). Comparison of computerized chemical models for equilibrium calculations in aqueous systems; chemical modeling in aqueous systems; Speciation, sorption, solubility, and kinetics. *ACS Symposium Series* **93**: 857–892.
2. Allison, J.D., Brown, D.S., and Novac-Gradac, K.J. (1991). *MINTEQA2/PRODEFA2, A Geochemical Assessment Model for Environmental Systems: Version 3.0 User's Manual*. U.S. EPA, Athens, GA.
3. Zhu, C. and Anderson, G. (2002). *Environmental Applications of Geochemical Modeling*. Cambridge University Press, Cambridge, UK.
4. Bethke. (1996).
5. Bethke. (2002).
6. Drever, J.I. (1997). *The Geochemistry of Natural Waters*, 3rd Edn. Prentice-Hall, Englewood Cliffs, NJ.
7. Garrels, R.M. and Mackenzie, F.T. (1967). *Origin of the Chemical Compositions of Some Springs and Lakes in Equilibrium Concepts in Natural Waters*. American Cancer Society, Washington, DC, pp. 222–242.
8. Lindberg, R.D. and Runnels, D.D. (1984). Ground water redox reactions: An analysis of equilibrium state applied to Eh measurements and geochemical modeling. *Science* **225**: 925–927.
9. Appelo, C.A.J. and Postma, D. (1993) *Geochemistry, Groundwater and Pollution*. A. A. Balkema, Rotterdam.
10. Parkhurst, D. and Appelo, C.A.J. (1999). *USER'S Guide to PHREEQC (VERSION 2)—A Computer Program for Speciation, Batch-Reaction, One-dimensional Transport and Inverse Geochemical Calculations*. Water-Resources Investigations Report 99-4259.
11. Engesgaard, P. and Christensen, T.H. (1988). A review of chemical solute transport models. *Nordic Hydrol.* **19**(3): 183–216.
12. Mangold, D.C. and Tsang, C.-F. (1991). A summary of subsurface hydrological and hydrochemical models. *Rev. Geophys.* **29**(1): 51–79.

Table 1. Comparison of Selected Codes' Capabilities and Features

Program	Source	Speciation	Reaction Path	Tabular Output	Graphic Output	Surface Rxns.	Kinetics	Inverse	Transport	Multiple Databases
EQ3/6	LLNL	yes	yes	Yes	no	no	yes	no	no	no
GWB	Rockware*	yes	yes	Yes	yes	yes	yes	no	yes	yes
HYDROGEOCHEM2	SSG*	yes	yes	Yes	no	yes	yes	no	yes	no
MINTEQ	EPA	yes	no	yes	no	yes	no	no	no	no
MINEQL+	ERS*	yes	no	yes	some	no	no	no	no	no
PHREEQC	USGS	yes	yes	yes	no	yes	yes	yes	yes	yes

*Commercial programs; others are freeware.

EQ3/6—http://www.earthsci.unibe.ch/tutorial/eq36.htm

GWB—Rockware—http://www.rockware.com

HYDROGEOCHEM—http://www.scisoftware.com/environmental_software/software.php

MINEQL+—http://www.mineql.com/

MINTEQ—http://soils.stanford.edu/classes/GES166_266items%5Cminteq.htm

PHREEQC—http://wwwbrr.cr.usgs.gov/projects/GWC_coupled/phreeqc/

READING LIST

Glynn, P.D., Engesgaard, P., Kipp, K.L., Mallard, G.E., and Aronson, D.A. (1991). *Two Geochemical Mass Transport Codes; PHREEQM and MST1D, Their Use and Limitations at the Pinal Creek Toxic-Waste Site*. U.S. geological survey toxic substances hydrology program. Abstracts of the technical meeting, Monterey, CA, March 11–15, 1991. Open-File Report—U.S. Geological Survey OF 91-0088.

Grove, D.B. and Stollenwerk, K.G. (1987). Chemical reactions simulated by ground-water-quality models. *Water Resour. Bull.* **23**(4): 601–615.

GEOLOGICAL OCCURRENCE OF GROUNDWATER

JOHN E. MOORE
USGS (Retired)
Denver, Colorado

The occurrence, movement, and storage of groundwater are controlled by geology. The geologic factors that control groundwater are petrography, stratigraphy, structure, geomorphology, lithology, and thickness (1). The petrography of a given rock type controls the porosity and permeability. Porosity defines the storage capacity of an aquifer. There are two types of porosity: primary and secondary. Primary porosity, such as pores between sand grains, is created when rocks are formed. The shape, sorting, and packing of grains determine primary porosity. Sedimentary rocks are poorly sorted when the grains are not the same size creating spaces between the larger grains that are filled by smaller grains. Secondary porosity such as joints, fractures, and solution opening, is formed after the rock has been deposited. The number and arrangement of fracture openings and the degree to which they are filled with fine-grained material control secondary porosity.

Aquifers (1) are classified as unconfined and confined. An unconfined aquifer has the water table as its upper boundary (2). Recharge to unconfined aquifers is primarily by downward seepage through the unsaturated zone. The water table in an unconfined aquifer rises or declines in response to rainfall and changes in stream stage. When a well that tap an unconfined aquifer is pumped, the water level is lowered, gravity causes water to flow to the well, and sediments near the well are dewatered. Unconfined aquifers are usually the uppermost aquifers and, therefore, are more susceptible to contamination from activities at the land surface.

A confined (artesian aquifer) contains water under pressure greater than that of atmospheric. Rocks of permeability lower than the aquifer overlie a confined aquifer. The low-permeability layer that adjoins a confined aquifer is called a confining bed. A confining bed has very low permeability that restricts the movement of groundwater either into or out of the aquifer. Confining beds are thus poorly transmissive to groundwater flow. Because the water is under pressure, water levels in wells rise above the base of the confining bed. If the water level in a well that taps a confined aquifer stands above the land surface, the well is called a flowing artesian well. In some cases, a fault (fractures along which rocks have moved) will allow the passage of water from a confined aquifer to the surface, resulting in a spring.

Sand and gravel aquifers are the source of most of the groundwater pumped in many parts of the world, including North America, The Netherlands, France, Spain, and China. Sand and gravel aquifers are common near large to moderately sized streams. Rivers or the meltwater from glaciers formed these aquifers.

Limestone aquifers are the sources of some of the largest well and spring yields because limestone is soluble in water. Openings that existed when the rocks were formed are frequently enlarged by solution (dissolved by water), providing highly permeable flow paths for groundwater (3).

Basalt and other volcanic rocks are some of the most productive aquifers. Basalt aquifers contain water-bearing spaces in the form of shrinkage cracks, joints, and lava caves. Lava tubes are formed when tunneling lava ceases to flow and drains out, leaving a long, cavernous formation.

Fractured igneous and metamorphic rock aquifers are the principal sources of groundwater for people who live in mountainous areas. Where fractures are numerous and interconnected, these rocks can supply water to wells and can be classified as aquifers. Wells are commonly 50 to 100 feet deep (15 to 30 meters). Granite and metamorphic rocks have not been extensively developed as aquifers. Groundwater movement in these rocks is irregular, which makes exploration for a water supply difficult (4).

Sandstone aquifers are formed by the cementation of sand. Their porosity ranges from 5–30%. Their permeability is largely a function of the amount of cement (clay, calcite, and quartz). Sandstone is an important source of groundwater in Libya, Egypt (Nubian), Britain (the Permo-Triassic sandstones), the north central United States (St. Peter-Mount Simon Sandstone), and in the west central United States (the Dakota Sandstone).

BIBLIOGRAPHY

1. Meinzer, O.E. (1923). *The occurrence of ground water in the United States, with a discussion of principles*. U.S. Geol. Survey Water Supply Paper 489.
2. Lohman, S.W. et al. (1972). *Definitions of selected groundwater terms-revisions and conceptual refinements*. U.S. Geol. Survey Water Supply Paper 1988.
3. Heath, R.C. (1983). *Ground water regions of the United States*. U.S. Geol. Survey Water Supply Paper 2220.
4. Moore, J. E. et al. (1995). *Groundwater, a Primer*. American Geological Institute, Alexandria, VA, p. 53.

GEOPHYSICS AND REMOTE SENSING

JOHN R. JANSEN
Aquifer Science & Technology
Waukesha, Wisconsin

When data cannot be gathered by direct observation, such as observing an outcrop or taking a physical sample, information can often be obtained by using remote sensing techniques. Remote sensing methods measure physical properties of materials by measuring changes in the flux of various forms of energy using natural or transmitted fields. The science of applying these methods to earth materials is called geophysics

Geophysics is commonly used to obtain subsurface information for a variety of resource development and engineering applications. A partial list of applications for which geophysical methods can be used includes mapping aquifers, mapping water quality, mapping geologic structure, measuring flowing fluids, finding buried objects for engineering and archeological purposes, and measuring *in situ* soil and rock properties. The major energy fields used include magnetic fields (magnetometry), electrical fields (electrical resistivity and spontaneous potential), electromagnetic fields (electromagnetic induction and ground-penetrating radar), propagation of seismic waves (seismic reflection, seismic refraction, passive acoustic emission monitoring, and spectral analysis of surface waves), the gravitational field (gravimetry), gamma-ray radiation (gamma-ray spectroscopy), and heat transfer (geothermal). Other physical properties can be measured by bombarding the material with gamma rays (electron density) or high-energy neutrons (hydrogen content). Less commonly, properties such as thermal conductivity, electrical chargeability, or magnetic resonance are also measured.

The methods can be broadly classified in several ways. Most commonly, methods are classified by their mode of deployment. Surface methods are commonly applied from the ground surface. Borehole methods are used within boreholes or wells. Several techniques can also be applied using airborne or marine systems. Methods can also be classified as active or passive measurement. Methods that measure variations in natural fields, such as gravimetry or magnetometry, are called passive methods because they use natural fields propagating through the earth. Other methods are called active methods because the measurement is based on the response of earth material to some form of transmitted energy. Active methods include most seismic methods and most electrical methods.

There are many types of geophysical methods; each has specific advantages and limitations. The choice of the proper method requires understanding the target body, the matrix material, and the environment in which the survey is to be conducted. The following is a brief outline of the most common geophysical methods with a summary of the characteristics of each method. The list is by no means complete.

SURFACE METHODS

Electrical Methods

Electrical Resistivity. The electrical resistivity method uses two electrodes planted at the surface to pass an electric current through the ground. A second pair of electrodes is used to measure the potential difference between two points. The measured potential and the applied current are used to calculate the electrical resistance per unit length (resistivity) of the subsurface. The depth of measurement is changed by changing the relative position and spacing of the electrodes. Several newer systems use cables with multiple electrodes connected to a switching system to select the electrode pairs. The field data produce an apparent resistivity value, which is a function of all materials penetrated by the current. The data must be modeled to obtain the unique resistivity and thickness of each layer for a nonuniform subsurface.

Applications

Can distinguish saturated from unsaturated materials.
Can distinguish sandy material from clay-rich material.
Can distinguish high-conductivity groundwater (inorganic contaminants) from low-conductivity groundwater.
Less susceptible to cultural interference than electromagnetic (EM) methods.
Better vertical resolution for resistive targets than EM methods.

Limitations

Requires electrical coupling with the surface, which can be difficult in dry soils.
Generally requires three to five times the surface array length to the depth of investigation.
Field acquisition generally slower than EM methods.
Poorer lateral resolution than EM methods at depths of more than a few meters.
Relatively insensitive to changes in the resistivity of highly conductive bodies.
Highly resistive near-surface material can mask conductive bodies at depth.
Can be affected by strong EM fields such as from high-voltage power lines.
Interpretation of field data produces nonunique solutions (i.e., many possible solutions can fit data nearly as well as actual conditions).

Spontaneous Potential. The SP method uses two nonpolarizing electrodes and a high-impedance voltmeter to measure natural voltages between two points generated by several natural processes. Anomalies can be due to variations in the geochemical properties of soils, fluids, or moving groundwater. Although its origins are in the mining industry for finding sulfide ore bodies, the method is principally a profiling technique to find seepage pathways such as for a dam investigation.

Applications

Relatively fast and simple data acquisition and interpretation.
Can detect zones of high groundwater flow.
Can detect oxidizing metal bodies.

Limitations

Requires special nonpolarizing electrodes for accurate voltage measurements.

Susceptible to interference from near-surface soil conditions.
Interpretation is qualitative and can be ambiguous.

Induced Polarization. The IP method measures a property known as chargeability, which is the ability of subsurface materials to hold a charge after an electrical current is shut off. It is basically a modification of the electrical resistivity method that measures induced electrical polarization between two electrodes caused by an applied low-frequency alternating current. Measurements are typically made by measuring the decay of voltage with time after the applied current is shut off, called the time domain method, or by measuring the suppression of voltage relative to the frequency of the applied external current, called the complex resistivity method. IP surveys have been used to determine the clay content of the subsurface or the geochemical properties of several minerals in the subsurface. The surveys can be conducted as profiles or as depth soundings.

Applications

Can be used to estimate the clay content of the subsurface and thereby infer hydraulic conductivity.
Can be acquired simultaneously with electrical resistivity data.
Can be used to detect organic contamination of soils under certain conditions, primarily sites that have 6% to 12% clay content.
Can be used to measure the *in situ* oxidation state of some inorganic compounds.

Limitations

Requires surface electrodes and better electrode coupling than standard resistivity surveys.
Requires higher currents than electrical resistivity—potentially dangerous.
Ability to convert IP soundings to depth sections is limited.
IP response is relatively insensitive to increases in clay content above 7% to 10%.
Interpretation is qualitative and can be ambiguous.

Azimuthal Resistivity. The azimuthal resistivity method is a modification of the standard electrical resistivity method that measures electrical resistivity versus azimuthal orientation at one location. The method is used to identify areas of fractured till or bedrock and determine the orientation of the principal fracture set. Several specialized electrode arrays have been developed to reduce the number of readings required, including the tripotential array and the square array.

Applications

Can determine the fracture orientation and fracture porosity from the surface.
Uses standard resistivity equipment.

Limitations

Data acquisition can be relatively slow and require a relatively large area of open land.
Limited ability to determine location of individual fractures.
Steeply dipping or irregular bedrock surfaces can skew data.
Cannot distinguish clay-filled from fluid-filled fractures.
Interpretation is quantitative for uniformly fractured materials with no overburden; quantitative interpretation is more limited for more common cases such as discrete fracture zones with overburden.

Seismic Methods

Seismic Refraction. The seismic refraction method uses the principle of refraction (bending) of sound energy across a boundary between two materials that have different sound velocities. The travel time of sound energy from a seismic source, such as a hammer blow or explosive, to a line of ground motion detectors, called geophones, is measured by a device called a seismograph. The travel time versus distance between the source and the geophones is used to calculate the thickness and depth of the geologic layers present. The method can resolve only layered systems in which each layer has a higher sound velocity than the layer above it. A layer that is too thin or whose velocity is lower than the layer above it cannot be detected and will introduce an error in the depth calculation of deeper layers. The shape of dipping or undulating surfaces can be resolved if several source locations are used in different positions relative to the geophone string.

Applications

Relatively simple and fast.
Requires less interpretation than other seismic methods.
Directly measures the seismic velocity of the materials present.
Measured velocities can be used to determine the type of bedrock or soil.
Can provide information on depth to bedrock to within about 10%.
Can determine the slope and shape of a bedrock surface.
Can map dense till sheets beneath softer sediments.
Can identify larger faults under the right conditions.

Limitations

Typically requires three to five times the surface array length to the depth of penetration.
Practical depth limit approximately 150 to 200 feet without a strong seismic source such as explosives.
Cannot detect low velocity layers beneath high velocity layers (i.e., sand beneath clay).
Limited ability to measure changes in velocity along a layer.
Cannot resolve steep dips.
Water table mapping only possible in coarse, sandy soils.
Sensitive to vibrations from cultural features such as highways.

Seismic Reflection. The reflection methods uses a string of geophones connected to a seismograph to measure direct reflections of sound energy from geologic boundaries between layers with contrasting acoustic impedance values (density times sound velocity). After significant processing, the data are typically plotted as a time section that shows the two-way travel time to major reflectors that is analogous to a geologic cross section, except that the depth axis is presented in travel time, which is not linear with depth. The method is typically used for deeper, more detailed studies than the refraction method, and the limitation of increasing velocity with depth does not apply. Faults, geologic structure, and dipping or truncating beds are commonly mapped by reflection.

Applications

The depth of penetration can be several times the surface array length.

More sophisticated interpretation and processing techniques are available.

Produces a continuous image of the subsurface that provides a time cross section that can be converted into a depth cross section.

Lateral resolution is much better than refraction.

Can tolerate steeper dips than refraction.

Can accurately map small faults.

Can be used onshore or offshore.

Can handle low velocity layers beneath high velocity layers.

Limitations

Requires more sophisticated field equipment, processing, and field procedures.

Data acquisition and processing are more intensive than refraction.

Requires sophisticated analysis and accurate seismic velocity data from an independent source to convert from time to depth cross section.

Difficult to apply to shallow exploration targets (i.e., above approximately 50 to 100 feet).

Interpretation of data subjective and requires experienced interpreter.

Sensitive to vibrations from cultural sources such as highways.

Gravitational Methods

Gravimeter Surveys. The gravitational method uses a sensitive balance, called a gravimeter, to measure variations in the force of gravity at the surface caused by variations in the density of the subsurface. A common modification of the method is called the microgravity method. Microgravity uses a more sensitive gravimeter to detect small near-surface features, such as voids or cavities.

Applications

Can be used to detect buried bodies of contrasting density.

Can be used to detect bedrock valleys, faults, cavities, and other geologic structures.

Data acquisition relatively simple.

Forward modeling can easily be used to predict the success of the method for a given target.

Limitations

Requires a series of complex data corrections requiring precise locations and elevations at points of measurement and surrounding topography.

Materials of similar density do not produce a measurable anomaly (e.g., granite and quartzite)

Large anomalies at depth can mask shallow, smaller anomalies.

Interpretation of data produces a nonunique solution.

Can be difficult to collect accurate data around buildings or in areas of highly irregular topography.

Magnetic Methods

The magnetic method measures variations in the intensity of the earth's magnetic field caused by material of high magnetic susceptibility such as ferrous iron. Measurements can be made with single sensor units, so-called total field magnetometers, that measure the total magnitude of the magnetic field at a given point. Other types of magnetometers make simultaneous measurements at two elevations and compute the vertical gradient of the magnetic field, which increases the sensitivity to shallow targets and decreases the interference from adjacent objects. Some magnetometers measure only the vertical component of the magnetic field, which also reduces interference from adjacent objects. Three types of magnetometers are currently available: (1) Proton precession instruments measure the total field and commonly use two sensors to measure the vertical gradient. Proton precession magnetometers are commonly available and relatively inexpensive, but they are sensitive to steep magnetic gradients from cultural sources, such as power lines and large metal bodies, and do not provide reliable data in many highly developed sites. (2) Flux gate magnetometers measure only a single (usually vertical) component of the magnetic field. The readings are slower, and the sensor must be properly oriented, making data acquisition slower; but the sensor functions well in the presence of steep vertical gradients and is less sensitive to horizontal fields from adjacent objects. The units are less common and more expensive, but they work on sites where proton procession units do not. (3) Cesium magnetometers make fast, very accurate total field measurements, and two sensors are commonly used to make vertical gradient measurements. The sensors can handle steep magnetic gradients and operate at a higher data acquisition rate than other magnetometers. The sensor has an axial blind spot of about 30° that may produce erroneously low readings if the sensor is not properly aligned with the target bodies. The units are more common than flux gate systems but more expensive than proton systems.

Applications

Field acquisition is fast and relatively easy.

Can readily detect buried ferrous objects (e.g., buried drums, tanks, and pipelines).

Can be used to detect geologic structures where materials of contrasting magnetic susceptibility are present (e.g., igneous vs. sedimentary rock).

Limitations

Large masses of ferrous objects can have coalescing anomalies that prevent precise location of individual targets.

Susceptible to interference from surface metallic objects such as fences or strong electromagnetic fields from power lines.

Does not detect material of low magnetic susceptibility (e.g., nonferrous objects such as aluminum or fiber barrels).

Interpretation of data produces nonunique results.

Electromagnetic Methods

Frequency Domain Electromagnetic Induction. Frequency domain EM systems use a two-coil system to transmit an electromagnetic field to induce "eddy" currents in the subsurface and measure the resultant magnetic field, which is a product of the primary field and the induced fields. In some systems, the receiver coil is coincident with the transmitter coil. Systems typically operate at a single frequency that can be selected by the operator. Some systems use a sweep across a range of frequencies. The depth of investigation is controlled by changing the separation between the coils or the frequency of the transmitted field. Two components of the magnetic field are measured, the in-phase or quadrature component and the out of phase component. Some units are designed to operate in low-conductivity soils and use the quadrature term to set the coil spacing. These units are called low induction number systems and generally provide a direct reading of subsurface conductivity. They are easier to use in the field but do not make accurate measurements in soils above a few hundred millimhos per meter. Other systems require measuring the coil spacing independently. These systems make an accurate measurement of subsurface conductivity over a broader range of conditions, but the data must be processed to provide a conductivity measurement. The in-phase component of these units can be used as a direct indicator of high induction number material, such as metal, in the subsurface. Both types of systems produce an apparent conductivity measurement that must be modeled to obtain individual layer conductivity and thickness for a nonuniform subsurface.

Applications

No electrodes required; can be used on surfaces where electrode plants would be impossible.

Data acquisition is fast and efficient.

Can be conducted through freshwater.

Different coil orientations and intercoil spacing can be used to accommodate different depths of investigation.

Lateral resolution much better than electrical resistivity techniques.

Can be used to detect buried metal.

Detects nonferrous metal (e.g., copper, aluminum, brass, bronze).

Sensitive to inclined conductive sheets such as faults or fracture zones.

Can be used for azimuthal surveys to map fracture density and orientation.

Simple profiling instruments can be used by relatively inexperienced operator (e.g., Geonics EM34 or EM31).

Multiple-frequency systems can make vertical depth soundings.

Limitations

Depth interpretation of sounding data requires an experienced interpreter.

Vertical resolution generally poorer than with electrical resistivity.

More sensitive to cultural interference (pipelines, metal fences, power lines, etc.) than electrical resistivity.

Relatively insensitive to changes in the conductivity of highly resistive targets.

Highly conductive surface material limits depth of penetration.

Cannot be used through reinforced concrete.

Problem of nonunique interpretation greater than with resistivity method.

Very Low Frequency Induction. The VLF method uses low-frequency radio waves from one of several military transmitters as a plane wave source. The instrument measures the magnetic field generated by induced eddy currents. Some instruments use a pair of electrodes to measure the phase shift of the electric field. The depth of penetration of the system is limited by the frequency of the military transmitter and the conductivity of the subsurface. Penetration may be limited in areas of clay-rich soils, and the method is generally incapable of detecting features that are aligned perpendicularly to the direction of propagation of the plane wave. The method is generally used for profiling to detect fractures in bedrock, although it does have some limited depth sounding capabilities.

Applications

Simple instrumentation and operation.

Very rapid and efficient.

Sensitive to inclined conductive sheets such as faults or fracture zones.

Can map lateral variations in conductivity such as conductive plumes, changes in soil type, or landfill boundaries.

Limitations

Limited range in transmitted frequencies available that limits choices of the depth of investigation.

Certain transmitters are out of service periodically.

Relatively insensitive to changes in conductivity in highly resistive targets.

Highly conductive surface materials limit the depth of penetration.

Location of transmitters makes it difficult to detect linear conductive bodies oriented approximately perpendicularly to the direction of propagation of the plane wave.
Long, narrow, near-surface conductors (such as pipelines, fences, or groundwater-filled bedrock depressions) produce strong anomalies that can mimic or mask deeper anomalies.
Depth of investigation is totally dependent on subsurface conductivity for a given transmitter frequency.
Interpretation is largely qualitative and nonunique.

Magnetotellurics. The magnetotelluric (MT) method uses plane waves from natural electromagnetic fields generated by oscillations in the earth's ionosphere as transmitter signals or from transmitters positioned remote from the area of investigation. The magnetic and electric fields are measured in two horizontal components using two orthogonal electrode dipoles and two orthogonal magnetometers. The intensity and phase of electric and magnetic fields are used to determine the conductivity of the subsurface. The method is commonly used to map conductive targets at depths of hundreds or thousands of meters, where other EM methods are impractical. The magnitude and frequency range of the natural EM signals are variable, and data collection often is limited to the later afternoon or evenings. Deep studies require very low source frequencies that can require hours of data collection to record a single sounding. Shallower studies frequently require a remote transmitter to provide enough EM energy in the audio-frequency range, known as a controlled source audio-frequency magnetotelluric (CSAMT) survey. Interference from power lines and other utilities is common in developed areas. Modern MT systems generally can collect both natural source MT and CSAMT data in a single sounding to provide a broader spectrum of frequencies.

Applications

Can find conductive anomalies at depths of a few hundred to a few thousand meters.
A large area can be surveyed with a single transmitter setup (CSAMT) or with no transmitter (MT).

Limitations

The natural source is often biased toward low frequency, which seriously limits the sensitivity of the method to shallow features without a CSAMT transmitter.
Relatively insensitive to small, near-surface targets.
Relatively insensitive to highly resistive targets.
Sensitive to interference from cultural features such as power lines and buried utilities.
Interpretation is nonunique

Time Domain Electromagnetic Induction. The TEM (or TDEM) method uses a loop of wire to pass a current of several amps. The current is shut off almost instantaneously to create a broad-frequency-pulsed EM source. Typical transmitter loops are square or rectangular and a few meters to a few hundred meters on a side. The receiver uses a small magnetic coil to measure the magnetic field over a series of time windows following the shutoff of the current in the transmitter coil. The magnetic field measured at the receiver is a function of the induced eddy currents, which are a function of the conductivity of the subsurface. The data produce a sounding that can be modeled to determine the change in electrical conductivity versus depth. Several systems are available that cover a range in exploration depths of a few tens of meters to a few hundred meters.

Applications

Well suited to map conductive bodies at depths of a few tens of meters to a few hundred meters.
Often used to map zones of saline groundwater.
Vertical resolution significantly better than most multiple-frequency FDEM methods.
Can be used to find resistive targets under favorable circumstances.

Limitations

Practical upper limit of method is approximately 10 to 20 meters, depending on near-surface conductivity.
Interpretation produces nonunique results.
Sensitive to interference from linear conductors such as power lines, fences, or pipelines within distances of one or two times the depth of investigation from the transmitter loop.

Metal Detectors. Metal detectors are essentially specialized FDEM or TEM devices designed to find buried metal objects. The FDEM instruments use the high response of metal objects to distinguish them from natural materials. TEM instruments take advantage of the relatively long duration of the eddy currents induced in metal bodies. In most instruments, the depth of penetration is related to the coil size, transmitter frequency, and the surface area and shape of the metal target. The depth of penetration is generally unaffected by soil type over the conductivity range of typical soils. The instruments have high lateral resolution, but the vertical resolution of the depth of the target is generally limited.

Applications

Ease of use.
Very portable.
High lateral resolution.

Limitations

Limited depth of investigation.
Response of target related to the surface area of the target rather than the volume of metal.

Ground Penetrating Radar. GPR uses a high-frequency pulse of EM energy to probe the subsurface. The EM energy is reflected off boundaries of contrasting dielectric constant (the ability of a material to separate an electric charge), so the instrument can find nonmetallic or metallic targets. The output is a time versus amplitude plot that is generally displayed as a series of plots along a transect line. The field plots are analogous to cross sections, except that the vertical access is in two-way travel time, not depth. The data can be processed similarly to seismic reflection data to produce depth sections. The instruments are configured with single transmitter/receiver antennas or with separate transmitter and receiver antennas that can be offset to collect data at different separations to increase the data processing that can be conducted. Under ideal conditions, the GPR method provides very high-resolution images of buried bodies, stratigraphic structure, or areas of disrupted soils.

The penetration of the method is determined by the conductivity of the subsurface and the degree of scattering from rubble or other small targets. Penetration is severely limited in highly conductive soils or in areas with lots of buried rubble or other discontinuous reflectors. The depth of penetration can be increased to some degree by using a lower frequency antenna but at the cost of resolution of the target.

Applications

Resolution can be of the order of a few centimeters.
Provides a cross section of subsurface in the field.
Can detect buried, ferrous or nonferrous targets.
Can detect disturbed soil zones.
Sophisticated data processing techniques available (similar to seismic reflection processing).

Limitations

Equipment is somewhat cumbersome; usually requires a relatively flat surface.
Depth of penetration is seriously limited by conductive material such as clay or water. Penetration through moist clay can be less than 1 foot.
Decreasing transmitter frequency to increase penetration decreases resolution.
Penetration and resolution limited by scattering effects at sites with buried cobbles or rubble.

Geothermal Techniques

The geothermal method uses soil temperature measurements to detect zones of anomalous flow at depth. Soil temperature is measured by using a shallow probe driven a few feet below the surface or a dedicated monitoring probe installed in a shallow borehole. Surface temperature variations diminish with depth below the surface. Daily temperature variations penetrate less than 1 or 2 meters, whereas seasonal temperature variations generally penetrate only about 10 meters. Below a depth of a few meters, soil temperature variations are largely a function of heat flow from within the earth. On a local scale, variations in heat flux often indicate zones of anomalous groundwater flow. Geothermal methods have been used to find permeable zones at depths of a few meters to over 100 meters using probes at depths of less than 1 meter to approximately 15 meters. A principal advantage of the method is that it detects anomalous flow instead of some other less directly related property, such as electrical resistivity. The method is susceptible to thermal interference by infiltration from shallow sources or from small features near the probes.

Applications

Identifies areas of high groundwater flow such as bedrock fractures or clean sand and gravel deposits.
One of the few methods that can directly detect moving groundwater from the surface.
Can pinpoint location of maximum permeability.
Can detect permeable features at depths of more than 100 feet beneath probes.

Limitations

Requires relatively uniform soils above target interval.
Small sand lenses near probes can create false anomalies or mask deeper anomalies.
Interpretation is qualitative.
Probe installation is relatively labor-intensive.

GROUNDWATER GEOPHYSICS—BOREHOLE METHODS

Most of the geophysical measurements made at the surface can also be made in a borehole using sensors lowered down the hole on a cable connected to up-hole recording equipment. Borehole measurements tend to be more accurate and precise because the borehole environment tends to be more predictable and less prone to noise. The field of measurement is typically only a few inches to a few feet around the borehole. Due to the small scale of the measurement, borehole methods tend to produce excellent vertical resolution. The combination of high resolution and greater accuracy make most borehole methods suitable for quantitative measurements and small-scale correlation of stratigraphic properties.

Borehole measurements are commonly used to determine the physical properties of the formation and formation fluids. The measurements can be used to measure directly or interpret several properties such as lithology, porosity, water quality, borehole diameter or alignment, borehole flow, formation or fluid temperature, mineralogy, and other properties of interest.

A variety of logging tools are available. Each tool has specific requirements for the borehole environment. Some tools can measure through casing; most cannot. Some tools require a fluid-filled hole; others operate in air, water, or drilling mud.

Electrical Logs

Spontaneous Potential Log. The SP log measures the electrical potential between a surface reference electrode and an exposed electrode that is raised up or down the borehole. The SP response is caused by a difference in resistivity between the borehole fluid and the formation fluid.

Applications

Quick and inexpensive.
Can be a good sand versus shale indicator.
Can be used to estimate total dissolved solids (TDS) of formation fluid.

Limitations

If borehole fluid and formation fluid have similar resistivity (typical for a water well), there is no SP response.
Works only in uncased fluid-filled boreholes.

Single Point Resistance Log. The SPR log measures the resistance between a grounded surface electrode and an electrode that is moved up or down the borehole. The tool measures the resistance of the total thickness of the formation between two electrodes. As a result, the measurement is in units of resistance, which is not a material property. The log is used primarily to detect changes in the formation but is not appropriate for quantitative analysis of the physical properties of the formation or the formation fluid.

Applications

Quick and inexpensive.
Good indicator of formation contacts.

Limitations

Measures resistance, not resistivity.
Works only in uncased fluid-filled boreholes.

Resistivity Log. Resistivity logs use two or more potential electrodes in the borehole to measure the change in potential difference along the borehole across a fixed electrode separation. Both current electrodes can be located on the down-hole sonde, or more commonly, one current electrode is on the sonde, and the other is planted at the surface. The depth of penetration of the current can be controlled by using different electrode spacings. Typically, a short electrode spacing is used to provide a high-resolution measurement of formation changes, and one or more longer electrode spacings are used for more accurate measurements of the bulk formation resistivity. Some sondes use additional current electrodes to focus the electrical field into the formation. The data can be corrected for borehole effects to provide a quantitative measurement of the bulk formation resistivity which is a function of the resistivity of the formation and formation fluid.

Applications

Measures bulk resistivity of formation.
Can be used to determine the formation lithology, porosity, and fluid resistivity.

Limitations

Works only in uncased fluid-filled boreholes.

Induction Log. Induction logs use a frequency domain electromagnetic (EM) sonde to measure the conductivity of the formation around the borehole. Some tools include a short spacing electrode tool to measure the conductivity of the borehole fluid independently.

Applications

Can be run in fluid-filled or air-filled borehole.
Can be run in open hole or through PVC casing.
Can be run with shallow or deep depths of investigation (dual induction) to determine the invasive effects of borehole fluids.

Limitations

More expensive than electrode logs.
Response of highly resistive materials (freshwater aquifers) is small, so the sensitivity of the tool is reduced.
Cannot be run in highly conductive borehole fluids.

Passive Radioactivity Log

Natural Gamma Log. Natural gamma logs use a sensor that measures the gamma-ray energy emitted by the formation. In most formations, gamma-ray emissions are almost entirely produced by shale minerals. As a result, the gamma-ray response is a good indicator of the concentration of clay in the formation.

Applications

Good clay indicator
Can be run in cased or uncased, fluid-filled or air-filled boreholes.

Limitations

Gamma-ray units (API units or counts per second) are arbitrary units and do not relate directly to clay mineral content.
Raw data must be corrected for borehole diameter effects before making direct comparisons.

Spectral Gamma Log. A spectral gamma log separates gamma-ray energy by frequency to measure the spectrum of gamma rays emitted by the formation. The frequency spectrum of gamma rays emitted by the formation is diagnostic of the radioactive isotopes in the formation.

The gamma-ray spectrum is commonly used to determine the concentrations of potassium, uranium, and thorium in the formation (KUT logs), but full spectrum tools are available that allow identifying additional elements.

Applications

Can be used to perform “complex formation analysis,” where the mineral content and groundwater flow history of formation are estimated.
Can be used to detect uranium or radium-enriched zones of sandstone aquifers.
May have application to radioactive contamination problems.

Limitations

Equipment is significantly more expensive than natural gamma tools.

Porosity Log

Several logging tools have been developed to measure the porosity of a formation. Porosity is a useful property to measure, but it does not always relate directly to permeability, which is generally the property of greatest interest. If more than one porosity tool is used, the data can be cross-plotted to identify formation type directly and correct the porosity measurement for lithologic effects. Most porosity tools require the use of a radioactive source in the tool. The use of radioactive sources requires special licensing and is prohibited in water wells in some states. The following sections describe the most common types of porosity tools.

Neutron Density Log. Neutron density logs use a radioactive source to bombard a formation with high-energy neutrons and records gamma-ray energy "bounced back" from the formation. The tool measures the hydrogen content of a formation, which is equivalent to the water content of the formation if no hydrocarbons are present. The water content of the formation is equivalent to the porosity below the water table and the soil moisture content above the water table. Some tools do not compensate for changes in borehole diameter and produce erroneously high estimates of porosity in washouts in the borehole. Compensated tools can accommodate some degree of variation in borehole diameter, but erroneous data can be recorded in zones with rough borehole walls. Bound water in shale zones produces erroneously high values of measured porosity.

Applications

Directly measures water content of formation. Can be related to porosity of nonshale units.
Measures porosity below water table and moisture content above water table.
Can be run in a cased or uncased hole.

Limitations

Sensitive to variations in borehole diameter.
Will see "bound water" in shale zones as porosity and give unrealistically high values.
Presence of hydrocarbons (pure product) erroneously suppresses calculated porosity.
Stringent licensing requirements to make a log due to the radioactive source.
Application to cased holes is limited.

Gamma Density Log. Gamma density tools use a radioactive source to bombard a formation with high-energy gamma rays. The tool has a sensor that records gamma rays "scattered" back to the detector. The magnitude of the backscattered gamma rays is a measurement of the electron density of formation that is related directly to bulk density. Given a bulk density measurement, the formation porosity is calculated for an assumed rock matrix (sandstone or limestone).

Applications

Will record accurate porosity values in shale zones.
Can be run in cased or uncased boreholes.

Limitations

Sensitive to variations in borehole diameter.
Presence of hydrocarbons causes erroneously high porosity values.
Stringent licensing requirements to make a log due to the radioactive source.

Sonic Log. Sonic logs use a source of high-frequency sound and one or more detectors on the down-hole sonde to measure the transit time of sound across a known distance of the formation. By comparing the measured velocity of sound in the formation to the known sound velocity for the rock type, the percentage of void space (porosity) can be calculated. Sonic tools do not require a radioactive source. The measurement is sensitive to irregular hole conditions. Specialized versions of the tool record the full waveform of the sound energy transmitted and can be used to detect poor bonds between the casing and grout or find voids behind the casing.

Applications

Porosity can be calculated for a given matrix (sandstone or limestone).
Modified sonic log can be used to inspect the bond of the cement grout to the casing (cement bond log).
Full waveform sonic logs can be run to measure the engineering properties of the formation.

Limitations

Sensitive to borehole diameter and roughness of borehole wall.
Works only in a fluid-filled uncased hole (except the cement bond log).

Borehole Condition Log

Several types of logs have been developed to measure the physical condition of the borehole, rather than measuring the properties of the surrounding formation. These tools can be used to determine the integrity of a borehole, determine if the hole is straight, or determine if flow is occurring between different zones. The following section describes several of the more common logs.

Down-Hole Televising Log. Down-hole televising logs use a down-hole video camera to inspect a borehole visually. Most cameras provide a fish-eye, forward-looking view with a depth reading on the image. Some cameras offer a side-looking view that looks directly at the borehole wall and can be rotated 360°.

Applications

Produces a visual record of a well bore on VCR tape or DVD.
Can be used to inspect borehole irregularities or obstructions.
Can be used to inspect the casing.
Can be used to detect fractures or solution cavities.
Shows small-scale stratigraphic layering with more resolution than other logs.

Valuable tool in assessing options to recover lost logging tools or drilling equipment.

Limitations

Requires clear borehole fluid. May have to add clear water to flush the hole if the water is cloudy.
Produces a qualitative record. Does not measure any physical property in calibrated units.
Most tools do not provide an oriented image.

Caliper Log. Caliper tools use spring-loaded arms that scrape along the borehole wall to record the variation in diameter. Simple tools that use two arms tend to turn so that the tool always measures the maximum hole diameter. Three-arm tools average the diameter at three points and provide better measurements of the average hole diameter. Four-arm caliper tools are available that provide oriented independent measurements in two coordinates in the horizontal plane that provide additional information on the shape of the hole.

Applications

Can identify ledges or washouts that can identify incompetent zones or fractures.
Can be used to calculate volumes of grout or backfill.
Can be measured in x and y planes (four-arm caliper) to estimate the shape of the borehole.

Limitations

Two-arm caliper measures only maximum hole diameter. Can be misleading in elliptical boreholes.
Does not measure alignment of borehole.

Alignment Log. Alignment tools measure the deviation of the center of the borehole from vertical. Simple alignment tools use a weighted plumb bob suspended from a tripod. The deflection of the plumb bob is measured by crosshairs at the surface. More sophisticated alignment tools use cameras with targets mounted on a system of gimbals or gyroscopes and magnetometers to measure the deviation of the borehole.

Applications

Can identify deviations in a borehole that can limit the size of the casing or pump that can be installed.
Can be used to calculate actual well depth and actual lateral position of a borehole.

Limitations

Must be combined with a caliper log to determine the shape of the hole accurately.
Large washouts can cause erroneous alignment readings unless the diameter of the hole is taken into account.
Sharp doglegs or extreme deviations can cause a surface line to “hold up” on the side of a borehole for simple tripod and crosshair systems and make measurements from below impossible.
Alignment tools that use magnetometers do not work inside a metal casing.

Acoustic Televiewer. Produces a high-resolution image of a borehole wall using a high-frequency sonic tool that is rotated to scan the hole. The log produces an image that shows relief on the borehole wall as shading or color patterns. Standard log formats show the borehole wall from 0 to 360° on the horizontal axis and depth on the vertical axis. The tools are generally oriented with respect to true directions. Televiewer logs are often used to map fractures and plot their strike and dip. The tool requires a fluid-filled hole, but water can be added to raise the water level artificially for logging.

Applications

Can measure the aperture, strike, and dip of fractures.
Can operate in either mud- or water-filled boreholes.

Limitations

Does not operate in air-filled holes.
Measures only the surface of the borehole; does not penetrate scale or casing.
The resolution of the tool is limited to several millimeters by the wavelength of the sonic source.
Scrapes or washouts from the drilling process can obscure true fractures.

Optical Televiewer. Produces a high-resolution image of the borehole wall using an optical scanning tool that is rotated to scan the hole. The log produces an image of the borehole wall in true color. Standard log formats show the borehole wall from 0 to 360° on the horizontal axis and depth on the vertical axis. The tools are generally oriented with respect to true directions. Optical televiewer logs are often used to map fractures, plot their strike and dip, and image small-scale stratigraphic features. The tool requires a clear borehole, either air or clear water.

Applications

Can measure the aperture, strike, and dip of fractures.
Can image individual grains and other fine stratigraphic features.

Limitations

Does not operate in mud-filled holes, cloudy water, or dusty air-filled holes.
Measures only the surface of the borehole; does not penetrate scale or casing.
Scrapes or washouts from the drilling process can obscure true fractures.

Flowmeter Log. Flowmeter tools are used to measure the flow within a borehole that develops from head differences between different zones or while pumping. Spinner tools are simple flowmeters that use small impellers to measure the flow past the tube. The tools are usually trolled up or down the borehole to overcome the frictional resistance of the impeller and its bearings.

For stronger flows, the tool can be held stationary to measure the absolute flow velocity past the tool without correcting for the speed of the tool in the borehole. Heat pulse tools can be used to measure lower flow rates. Heat pulse tools use a heating element to create a small pulse of heated water, and temperature sensors are placed at known distances above and below the heating element. The tool measures the time required for the pulse of heated water to move upward or downward to the sensor and calculates the flow velocity and direction. A modification of the heat pulse flowmeter has been used to measure horizontal flow in boreholes. The diameter of the hole must be known before the flow velocity can be converted into a volumetric flow rate. Specialized tools have been developed to detect flow behind a casing.

Applications

Can measure the flow up or down a borehole.
Can be used to estimate the head difference and transmissivity of different zones of a borehole.

Limitations

The friction of the bearings and inertial mass of the impeller impose a lower limit of flow that can be measured by a spinner tool.
Heat pulse tools have an upper limit of flow rate that can be accurately measured.

Temperature Logs

Down-Hole Temperature Log. Temperature logs use down-hole thermistors or thermocouples to measure the variation in the temperature of the borehole fluid with depth. Temperature variations in the hole can be used as indicators of flow. The data can be plotted as absolute temperature or as a differential temperature that removes the effect of the geothermal gradient from the plot.

Applications

Provides direct measurement of fluid temperature.
Can be used to detect zones of increased fluid flow within a borehole or between two boreholes.
Directly measures the geothermal gradient.

Limitations

Geothermal gradient causes a continuous drift in borehole temperatures that can mask thermal variations within the borehole.
Requires temperature and head variations between units to create a thermal anomaly to detect permeable zones within a borehole.
Flow within a borehole during or after drilling can mask geothermal gradient measurements.

GEOPHYSICAL SURVEY DESIGN

When designing a geophysical survey, several factors must be considered, including the physical properties of the target, the desired resolution, the physical properties of the matrix around the target, the ambient noise in the survey area, and the physical access at the site. To be detected, the feature you are trying to find must have some physical property that has sufficient contrast with the surrounding matrix. A variety of physical properties can be remotely sensed by geophysical methods. Some targets, such as buried pipes in sand, present obvious contrasts of a number of properties. Some targets, such as excavations backfilled with native soils, present less obvious contrasts or can be identified only by contrasts of related features, such as disruption of macropore drainage structures causing perched soil moisture over old graves.

Geophysical methods each have characteristic resolution limits that are a function of the wavelength of the fields used for measurement and the limitations of the equipment. The resolution of the various geophysical methods varies from several tens of feet for deep studies using electromagnetic induction methods to fractions of an inch using ground penetrating radar with high-frequency antennas. The resolution of the geophysical method used must match the physical dimensions of the target body if the target is to be detected. The magnitude and size of the anomaly must be predicted with reasonable accuracy, typically by forward modeling, to determine the minimum grid size that can be used for a reasonable probability of detection.

Site conditions also play an important role in determining the ultimate success of the survey. Surface and subsurface conditions can create a noisy environment that can easily obscure the signal from the target body. Noise sources are different for each method. For instance, ground vibrations from traffic or machinery can be a major source of noise for seismic or gravity methods but do not affect other methods such as electromagnetic induction or magnetics.

Physical conditions at a site also affect the utility of geophysical methods. Major obstructions such as buildings or developed land often limit the area where data can be gathered and create inherent limitations on the geophysical survey. Buildings, parked cars, dumpsters, overhead power lines, buried utilities, fences, buried rubble, and other cultural features are not often considered when designing a geophysical survey. In addition, subsurface conditions such as soil type, uniformity, and other factors affect the propagation of energy and can limit the performance of most methods. Surface conditions, such as paved surfaces, heavy vegetation, frozen ground, rough topography, and surface debris, must be considered when designing a survey.

Assuming that the site conditions are acceptable, it is still necessary to select a method that can detect the target body. The following is a partial list of typical survey objectives with some suggestions of methods to consider.

1. Locating optimal well locations (bedrock fractures or sand and gravel deposits)
    a. EM
    b. Electrical resistivity

   - c. Azimuthal EM
   - d. Azimuthal resistivity
   - e. Geothermal
2. Determining depth to bedrock and mapping bedrock surface
   - a. Seismic refraction
   - b. Gravity
   - c. Seismic reflection
   - d. GPR
3. Locating preferred pathways for groundwater flow (bedrock fractures, voids, seepage paths, or sand and gravel lenses)
   - a. EM
   - b. Electrical resistivity
   - c. Azimuthal EM
   - d. Azimuthal resistivity
   - e. Geothermal
   - f. SP
   - g. GPR
4. Finding buried metal bodies (tanks, drums, Buicks, etc.)
   - a. Magnetometer
   - b. Electromagnetic induction (EM)
   - c. Ground penetrating radar (GPR)
   - d. Metal detectors
5. Mapping conductive ground water plumes (inorganics in the range of hundreds of ppm)
   - a. EM
   - b. Electrical resistivity
   - c. GPR
6. Mapping continuity of till sheets or clay layers
   - a. Seismic refraction
   - b. EM
   - c. Electrical resistivity
   - d. Seismic reflection
7. Determining the lateral extent or thickness of refuse
   - a. EM
   - b. Electrical resistivity
   - c. Magnetometer
8. Detecting fractures in tills
   - a. Azimuthal EM
   - b. Azimuthal resistivity
9. Finding abandoned well casings
   - a. Magnetometer
   - b. EM
10. Detecting voids
    - a. Gravity
    - b. GPR
    - c. Resistivity
11. Direct detection of hydrocarbons
    - a. GPR (LNAPLS, DNAPLS, or biodegrading plumes)
    - b. Electrical resistivity (biodegrading plumes)
    - c. Induced polarization (bound only to clay)

GEOTHERMAL WATER

Nitish Priyadarshi
Ranchi University
Ranchi, Jharkhand, India

Geothermal water has a temperature appreciably higher than that of the local average annual air temperature. However, in general, a spring is considered hot when its temperature is about 12.2 °C higher than mean annual ambient temperature (1). The relative terms geothermal water, warm springs, and hot springs are common.

From earliest times, hot geothermal water or geothermal springs have been used for bathing and cooking. The first technological success in using geothermal energy was in Italy in 1904 where the worlds first geothermally driven electrical plant was opened and operated. Today, applications of low- and moderate-temperature (100 to 300 °F) geothermal waters have expanded enormously to include heating large tracts of homes and buildings (district heating), heating greenhouses for growing vegetables and flowers, fish farming (aquaculture), drying foods and lumber, and many other uses. In the United States alone, there are 17 district heating systems, 38 greenhouse complexes, 28 fish farms, 12 industrial plants, and 218 spas that use geothermal waters to provide heat. The district heating system in Boise, Idaho, has been operating since the 1890s and continues to provide heat today. In Iceland, most of the homes and other buildings are connected to geothermal district heating systems, and in the Paris basin in France, many homes are heated by bringing geothermal water to the surface. Geothermal greenhouses are prominent in Italy and in the western United States.

Worldwide, there are about 12,000 thermal megawatts of installed direct uses of geothermal fluids in nearly 30 countries, replacing the combustion of fossil fuels equivalent to burning 830 million gallons of oil or 4.4 million tons of coal per year. This illustrates the trend of countries switching to alternative energy sources such as geothermal energy.

Geothermal water discharges from numerous springs located mostly in mountainous or plateau areas. The springs are connected by faults to deeply buried reservoirs that contain geothermal water, which moves upward along the fault zones to discharge at the land surface. Much geothermal water discharges as hot springs that flow steadily instead of erupting at intervals.

Hydrothermal phenomena involving the release of water and steam are nearly always associated with volcanic rocks and tend to be concentrated in regions where large geothermal gradients occur. By implication, aquifers must also be present that permit water to percolate to great depths—often 1500 to 3000 m (2). One theory used to explain how geothermal water becomes heated in areas that are underlain by complex geologic structures is that when precipitation falls in highland areas, it recharges the aquifer system. Some of the water moves downward along faults and fracture zones to great depths. As the water descends, it becomes heated because of the geothermal gradient. At some depth, the heated water becomes lighter than the overlying water and then moves upward along faults to discharge as spring flow.

Deeply circulating groundwater can also become heated by cooling magma (molten igneous rock) at great depths in the crust of the earth. The water warms as it descends, possibly along fault zones that overlie the magma chamber, until it absorbs enough heat to become lighter than overlying water. The warm water then rises to the surface. The mechanism for circulating the water is the same, regardless of whether the water becomes heated by the geothermal gradient or by the buried cooling magma. It can also be heated through mantle decay of radioactive elements such as U, Th, and Ra, tectonic activity, metamorphic processes, and exothermic reactions of minerals.

More than 10,000 individual thermal features, including geysers, hot springs, mud pots, and fumaroles (steam vents), have been identified. Geysers are the best known and certainly the most spectacular features. They are a type of hot spring that periodically emits sudden, violent eruptions of stream and hot water. Water from surface sources and/or shallow aquifers drains downward into a deep vertical tube where it is heated above the boiling point. As the pressure increases, the steam pushes upward; this releases some water at the surface, which reduces the hydrostatic pressure and causes the deeper superheated water to accelerate upward and to flash into stream. The geyser then surges into full eruption for a short interval until the pressure is dissipated; thereafter, the filling begins, and the cycle is repeated (2). The explosive release of pressure can cause a column of steam and hot water to rise 200 feet or more into air. The period between eruptions depends on several factors, including the volume of steam and water that is ejected and how rapidly ground water refills the tubes and chambers. A mud pot, results when only a limited supply of water is available. Here, water mixes with clay and undissolved particles brought to the surface, forming a muddy suspension from the small amount of water and steam that continues to bubble to the surface. A fumarole, meaning smoke, is an opening through which only steam and other gases such as hydrogen sulfide and carbon dioxide discharge. These features are normally found on hillsides above the level of flowing thermal springs; water can often be heard boiling underground (3).

Excessive concentration of certain dissolved minerals in geothermal water poses water quality problems. The most common of these minerals are dissolved fluoride, arsenic, and iron. Concentration of dissolved fluoride in excess of 4 milligrams per liter can cause mottling of teeth, especially children's, and can cause bones to become brittle. Dissolved iron is not detrimental to human health, but concentrations in excess of 200 micrograms per liter can cause staining of kitchen and bathroom fixtures and can cause clogging of well-screen openings and pumps. Concentrations of dissolved arsenic in excess of 30 micrograms per liter are toxic to humans. Sodium, calcium, magnesium, and their salts are nontoxic to human beings. Chloride salts of these metals in higher concentrations may be toxic to plants because of their chloride ions rather than the metals. A chloride ion concentration of >300 mg/L is in general toxic to vegetation and results in defoliation, chlorosis, bronzing, and burning of plants (4). Boron in trace quantities is necessary for plants, but it is toxic to certain plants in concentrations as low as 1 to 2 mg/L.

Natural steam that always contains some noncondensable gases, such as hydrogen sulphide, and ammonia, presents the greatest potential hazard. Ammonia affects primarily the upper respiratory tract and causes coughing, vomiting, and redness of the mucus membranes of the mouth, nose, lips, and pharynx (4). In geothermal operations, ammonia is not likely to present a direct toxic hazard except possibly in the immediate vicinity of the power plant. Hydrogen sulphide, which is a noticeable geothermal effluent, is much more toxic than commonly realized. A concentration of more than 600 ppm hydrogen sulfide can cause death within 1 hour. It can also paralyze the respiratory center.

High values of total dissolved solids in geothermal water sometimes interfere with the usefulness of water as a source of energy. Rain, the prime source of meteoric water, which feeds the geothermal reservoir, is practically devoid of chemical ingredients. But after entering the atmosphere, it picks up gases such as CO_2, O_2, SO_2, and dust particles. As it falls to the ground, its concentration of chemical constituents may be up to 40 mg/L (5). Other important factors are climatic conditions and the state of weathering of rocks and their mineralogical compositions. Thus, even before entering the geothermal reservoir, water is considerably rich, chemically. Groundwaters whose salt content is around 1000 mg/L are not uncommon. The major cations in groundwater are Ca, Mg, Na, and the anions are HCO_3, SO_4 and Cl. Geothermal water commonly contains large concentrations of silica (if the water has moved through limestone or other calcite-rich rocks).

Along with these radicals, gases dissolved in meteoric water such as CO_2, O_2, He, and Ar also enter the reservoir. The chemistry of meteoric water is modified in the geothermal reservoir due to rock–water interaction under the influence of temperature, contributions from magmatic sources, mixing of different types of waters, and decay of U, Th, and Ra. Closer to the surface, within the zone of oxidation, air can also enter thermal waters. In that case, the nitrogen and oxygen content will be high, and their mutual ratio may also match that of the atmosphere. Some geothermal waters may be brines, for example, the Salton Sea area, California, whose salinity is 30% (6); others may be as pure as distilled water, for example, the Rajgir thermal spring of Bihar, India (7).

Microorganisms can thrive in geothermal waters, even at boiling temperature (8). Extremely thermophilic bacteria that survive at 350 °C have also been found (9). Fortunately, these microorganisms do not survive in animal systems. Pathogenic organisms can enter thermal waters through soil and air and also through contamination by animal/man in thermal water pools. Thermal springs of various kinds are found throughout the world. Notable areas exist in the United States, Iceland, New Zealand, the Kamchatka Peninsula of the former Soviet Union, Brazil, Argentina, Ethiopia, Zambia, China, Tibet, India, Thailand, Taiwan, and Japan. Yellowstone National Park in Wyoming, United States, that contains literally thousands of hydrothermal features possesses the greatest concentration of thermal springs in the world (10). This area marks the site of an enormous volcanic eruption 600,000 years ago. Today, a temperature of 240 °C exists only 300 m below the ground surface.

As economies expand and population grows, energy demand worldwide is increasing rapidly. If we are to increase our energy consumption while simultaneously reducing environmental pollution, we must change our fuel mix, which today relies heavily on fossil and nuclear fuels. The use of geothermal energy has enormous environmental advantages over the use of fossil or nuclear fuels. Among these advantages are far fewer and more easily controlled atmospheric emissions, maintenance of groundwater quality, and much smaller dedicated land requirements. The small quantities of gases emitted from geothermal power plants are naturally occurring and result from geologic processes. Because the earth is porous and permeable, these gases would eventually find their way to the surface, even in the absence of geothermal power development.

BIBLIOGRAPHY

1. Alfro, C. and Wallace, M. (1994). Origin and classification of springs and historical review with current applications. *Environ. Geol.* **24**(2): 112–124.
2. Todd, D.K. (1995). *Ground Water Hydrology*. Wiley, Toronto, Canada, p. 51.
3. Poland, J.F. (1972). *Glossary of Selected Terms Useful in Studies of the Mechanics of Aquifer Systems and Land Subsidence due to Fluid Withdrawl*. U.S. Geological Survey Water-Supply Paper 2025, p. 9.
4. Pandey, S.N. and Srivastava, G.C. (1996). In: *Environmental Hazards of Indian Geothermal Fields*. U.L. Pitale and R.N. Padhi (Eds.). *Geol. Surv. India*, Spl.Vol. 45, pp. 375–378.
5. Hawkes, H.E. and Webb, J.S. (1962). *Geochemistry in Mineral Exploration*. Harper and Row, New York.
6. Werner, H.H. (1975). Contribution to mineral extraction from superheated geothermal brines, Salton Sea Area, California. *Geothermics* **2**: 1651–1657.
7. Geological Survey of India Special Publication. (1991). Geothermal Atlas of India, p. 106.
8. Yun, Z. (1986). Thermophilic microorganisms in the hot springs of Tengencheng geothermal area, West Yunan, China. *Geothermics* **15**: 347–358.
9. Baross, J.A. and Dening, J.W. (1983). Growth of "black smoker" bacteria at a temperature of at least 250 °C. *Nature* **303**: 423–426.
10. Keefer, W.R. (1971). *The Geologic Story of Yellowstone National Park*. U.S. Geological Survey Bull. Publ. No. 1347, p. 92.

GHIJBEN–HERZBERG EQUILIBRIUM

Ekkehard Holzbecher
Humboldt Universität Berlin
Berlin, Germany

THE ORIGINAL PAPERS

In 1889 a note was published in The Hague (The Netherlands), which dealt with a suitable location for a new borehole for the pumping of drinking water for the city of Amsterdam. The authors, Drabbe and Badon Ghijben (1), outline in length the different geological formations, their extension and their connection with each other, and especially their connection with various saline surface water bodies in that region. In the final part, salinity becomes the subject of their consideration. The authors describe that saltwater is found in the subsurface, when there is a hydraulic gradient from the surface water into the aquifer and contrast this to the situation in the dunes at the North Sea coast. There, a freshwater lens can be observed on top of the salt water. In one sentence, the topic of this contribution can be summarized: If the height of the groundwater table above the mean sea level is denoted by a, the equilibrium between saltwater and freshwater is reached at a depth of $a/0.0238 = 42a$, when the seawater weight exceeds the freshwater weight by a factor of 1.0238. A simple calculation example follows using the given formula and a reference to an even earlier communication of Conrad in 1881, who is cited to have used factors between 40 and 50 in analogy to the proposed 42.

Twelve years later, the other note (2) appeared, in which the Ghijben–Herzberg equilibrium (GHE) is named. Herzberg, an engineer from Berlin (Germany), in a talk in Vienna on behalf of the annual meeting of the German Association of Gas- and Water-Specialists, was concerned with the water supply of villages on the German North Sea coast. More specifically, he was interested in the water supply on the North Sea islands, which, due to a rise in tourism, had an increasing demand for drinking water in the summer season. Compared with his Dutch counterpart, Herzberg provided a much more detailed derivation of the equilibrium rule. In order to illustrate his arguments, Herzberg presented a figure, which is reproduced in Fig. 1.

In Herzberg's version, t denotes the elevation of the groundwater table above mean sea level and is thus identical to Badon Ghijben's a; h is the depth of the freshwater boundary below mean sea level. A saltwater column of length h has the same weight as a freshwater column of length $H = t + h$, when the condition

$$h \cdot 1.027 = h + t \tag{1}$$

is fulfilled, and when it is taken into account that the specific weight of saltwater exceeds that of freshwater by a factor of 1.027. Solving Eq. 1 for h delivers

$$h = 37t \tag{2}$$

In both studies it is shown that a wide freshwater lens is to be expected in coastal regions. The extension of this length is roughly 40 times the elevation of the groundwater table above the seawater level. For practical purposes—the authors' main concern was water supply—one conclusion is that quite a large amount of fresh groundwater is available, as the lens extends much deeper into the subsurface than above the seawater horizon. Another conclusion is that the occurrence of freshwater at greater depth does not rely on the existence of an additional freshwater source in deeper horizons.

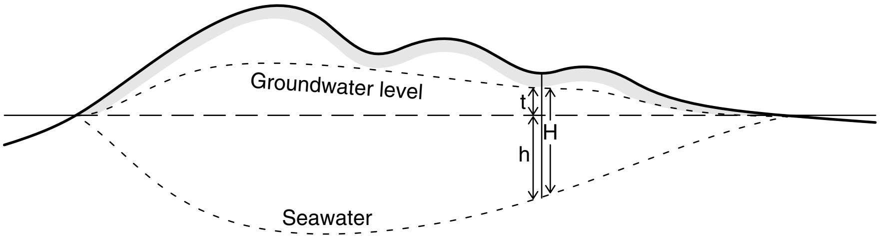

Figure 1. Illustration concerning the freshwater–saltwater equilibrium. Adapted from Reference 2.

Following the Ghijben note, a strong dispute took place in the beginning of the twentieth century in The Netherlands with respect to the maximum amount of pumping by the water works, as reported by de Vries (3). Proponents of the GHE argued in favor of such a regulated maximum.

In both publications the density difference is recognized as the major factor concerning the conditions of subsurface water near the sea coast. Freshwater is lighter than saltwater, and in order to reach an equilibrium (i.e., equal pressure), the freshwater column must extend further into the subsurface. The general situation with freshwater salinity ρ_f and water of elevated salinity ρ_s in the deeper subsurface, in current formulation, would be written as

$$h_f = -\frac{\rho_s - \rho_f}{\rho_f} z \tag{3}$$

where h_f denotes the elevation of freshwater above sea level, while z measures the depth of the interface below that level. As both extend in different directions from the zero sea level, they have different signs.

The equilibrium formulas given above can be taken as rule of a thumb. In most textbooks the formula is used with a factor of 40. The deviation of the factor in the formulation of Badon Ghijben (42) from the one given by Herzberg (37) stems from different assumptions concerning the relative weight of saltwater in relation to the freshwater. According to current knowledge, the density of seawater ρ_s lies between 1.02 and 1.029 g/m^3, depending on salinity and temperature. For a mean salinity of 35 ppt and a temperature of 10 °C (typical for subsurface water), the density is 1.027 g/m^3 (4), which is the value used by Herzberg. In real situations, it has to be taken into account that, due to mixing, dissolution, or precipitation, the subsurface saltwater may have a different salinity than seawater and that freshwater may have a slightly elevated density too.

CONDITIONS

The Ghijben–Herzberg equilibrium (GHE) has proved to be valid in several coastal systems all around the world. In both original publications, correspondence with observations is stressed. Herzberg describes that he already used the equilibrium formula for predictions concerning the depth of the saltwater in the sandy subsurface of several North Sea islands and that he obtained a good confirmation of the rule by measurements in boreholes. De Vries (3) lists field measurements and laboratory experiments, performed by Pennink in The Netherlands at the beginning of the twentieth century. However, the formula is linked to several conditions.

In the derivation it is assumed that the transition from the fresh to saline water appears quite rapidly. When the scale of the transition zone is very small in relation to the scale of the depth of the aquifer, one speaks of a *sharp interface*, which is a common term in the current literature; see, for example, Reference 5. One condition for a sharp interface is small dispersion, more specifically small transverse dispersion, as the flow direction coincides more or less with the interface (see below). But the tides and seasonal fluctuations in groundwater recharge and/or in the amount of groundwater being withdrawn also have an influence on the width of the transition zone. Extended descriptions have been derived nowadays, which are valid for situations with transition zones (see below).

Herzberg demanded caution concerning the transfer of his results obtained for the islands to cities near the coast, pinpointing the different scale, especially concerning the water supply. In a remark to Herzberg in Vienna, Halbertsma, another Dutchman, stated that the situation along the Dutch coast is quite different from one place to the other, but in some parts it is very similar to the situation described by Herzberg. As a counterexample he reports a situation where a confining impermeable clay layer can be found near the surface, which separates the freshwater on top from the saline water.

Custodio (6) illustrates several different situations, concerning the position of one or more clay layers, which disturb the ideal situation assumed for the GHE, showing that the GHE may over- or underestimate the position of the interface depending on local circumstances. In atoll islands a permeable layer below the upper strata of low hydraulic conductivity also results in a different situation, and the GHE can surely not be applied on coastlines of volcanic rock.

In the very vicinity of the coastline the GHE is mostly not valid. The GHE predicts that both fresh and saltwater heads (see below) should approach zero. But at many locations submarine groundwater discharge of freshwater has been observed (see Fig. 2). Kooi and Groen (7) provide an overview of the phenomenon and interpretations using different sharp-interface approaches.

The original formulation is given for phreatic aquifers and for the hydrostatic situation, that is, where flow in the freshwater zone as well as in the saltwater zone is neglected. Various different extensions have been made since the first formulations. Thus it is possible to use the generalizations of the Ghijben–Herzberg relation in situations for which the simple formula was not stated originally. In that sense, it is possible to use GHE as a

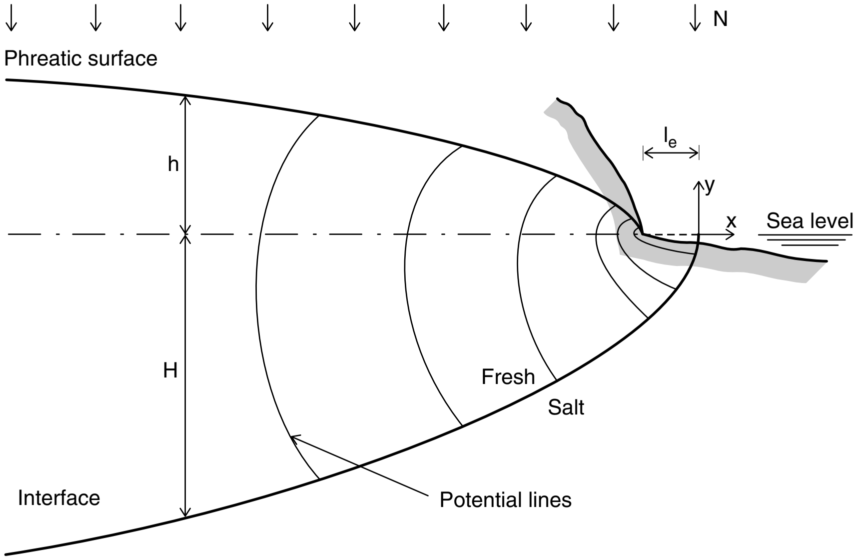

Figure 2. Schematic view of flow in an unconfined aquifer near the coast, describing submarine groundwater discharge. Adapted from Reference 12.

principle for confined aquifers, for dynamic situations, and for a dispersed transition zone, as described in the following.

When the amount of water withdrawal becomes a significant part of the entire water budget of a coastal system, the Ghijben–Herzberg formula should be questioned. Hydrostatics then has to be replaced by a quantification of the hydrological balance of the system.

GENERALIZATIONS

The original derivation assumes hydrostatic conditions, while the real situation is hydrodynamic in all cases. The authors were aware of that: in the report Badon Ghijben describes that under the North Sea dunes there is a relatively strong freshwater flow, recharged by infiltrating rainwater in the upper part, directed toward the sea, while there is a very weak flow of saltwater in the opposite direction in the lower part. In-between there is a transition from inflow to outflow.

However, the hydrostatic state can be taken as a limiting case from which the dynamic situation deviates only slightly as long as the groundwater velocities are small. The GHE idea of an equilibrium between freshwater and saltwater can then be transferred to dynamic situations. Hubbert (8) introduced the variable of *freshwater head*, defined as

$$h_f = \frac{\rho}{\rho_f}h - \frac{\rho - \rho_f}{\rho_f}z \tag{4}$$

where ρ and h denote density and head, measured at depth z. As outlined in the overview papers by Reilly and Goodman (9) and Cheng and Quazar (10), the freshwater head was used also by Muskat (11) in a description of the interface between oil and water in the subsurface. For $h = 0$ (stagnant saltwater) and $\rho = \rho_s$, the Ghijben–Herzberg formula (Equation 3) results for the interface position at depth z.

With the help of the freshwater head h_f, descriptions of situations with an interface can be simplified. A solution for freshwater head distribution is determined first; the location of the interface is calculated on the basis of the GHE afterward. This procedure is followed in many publications concerning different situations with fluids of different densities.

van der Veer (12) presents a simple formula for the one-dimensional flow in a coastal aquifer and a complex solution for flow in a vertical cross section, both based on the GHE. It is shown that distant from the shore both solutions coincide; that is, the use of the simpler solution is justified. But when the coastline is approached, the differences between both solutions increase. With the two-dimensional solution, it is possible to describe the submarine groundwater discharge into the sea.

Strack (13,14) uses the GHE and freshwater head to construct analytical solutions for shallow interface flow in confined and unconfined situations. In the case of the confined situation, the piezometric head h is the relevant variable, not the water table in the confined layer. Using these solutions the regional flow in coastal aquifers toward wells can be calculated in a two-dimensional horizontal domain.

The cross-sectional view, given in Fig. 3, depicting schematically the flow in the confined situation, is adapted from Strack (14). Due to pumping of freshwater in the upper part of the aquifer, the interface between saltwater and freshwater rises. Such an upconing may lead to severe problems, when the interface comes close to the well filter, as the salinity of the pumped water may increase above limits, which makes the water unsuitable for the drinking water supply, irrigation, or other purposes.

According to the static GHE, the interface rise exceeds the water table decline by a factor of 40 in coastal aquifers. Recent experience in Sweden shows that the GHE overestimates the rise of the interface (15). Dynamic approaches using well-known formulas for single wells were combined with the GHE by Wang (16) and by Motz (17). In regard to saltwater upconing, the GHE today

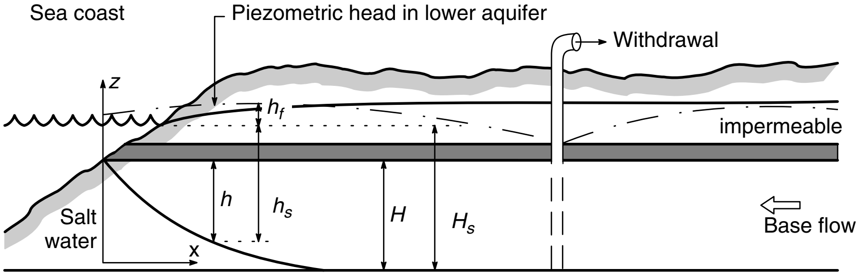

Figure 3. Schematic view of flow in a confined coastal aquifer with groundwater withdrawal. Adapted from Reference 14.

is often applied in combination with finite differences or finite elements (18,19).

The mentioned analytical solutions are valid for horizontal flow in an aquifer with a sharp interface, but vertical fluxes cannot be treated correctly by freshwater heads. As a further generalization of the GHE, Lusczynski (20) showed that for the general three-dimensional case in variable density flow, the gradients of freshwater head are responsible for horizontal groundwater flow, while for flow in the vertical direction the so-called environmental water head h_e is responsible. h_e is a generalization of Eq. 3:

$$h_e = \frac{\rho}{\rho_f}h - \frac{\rho - \rho_a}{\rho_f}z \qquad (5)$$

where ρ_a denotes the average density in the column between position z and sea level.

Note that for the latter extensions of the GHE the condition of a sharp interface has been dropped. But with ρ_a another parameter is introduced, which cannot be estimated easily without any knowledge of the mean position of the interface: remember that the principal purpose of the Ghijben–Herzberg formula was to deliver an estimate for that position.

In fact, saltwater intrusion is a problem of density-driven flow (21), in which flow and transport processes are coupled and cannot be calculated independently. Naturally, the flow in the aquifer determines the salt distribution, but on the other hand the salinity influences the flow via density and possibly via viscosity.

In order to treat such complex interactions, computer models are often used today. An overview on numerical approaches, codes, and applications, concerning saltwater intrusion, is given by Sorek and Pinder (22). Henry (23) presented a widely discussed solution for the miscible displacement in a confined aquifer, which has been used by numerous modelers as a test case. Figure 4 represents the solution in a vertical cross section, calculated with the new FEMLAB code (24) for multiphysics simulations. Depicted are streamlines extending from the left freshwater boundary to the seawater boundary on the right-hand side. On the seaside, streamlines enter in the lower part and return to the seaside after a u-turn. Along the paths salinity changes due to dispersion and diffusion. The grey pattern visualizes salinity, which is high in the lower right corner.

The Henry solution and with it the penetration of the saltwater wedge and its dispersion depends on various flow and transport parameters: the ambient freshwater flow and the dispersivities—dependencies that cannot be predicted by the GHE, which is based on water densities only.

CONCLUSION

The history of the GHE offers some more peculiarities. According to Carlston (25), the static equilibrium condition

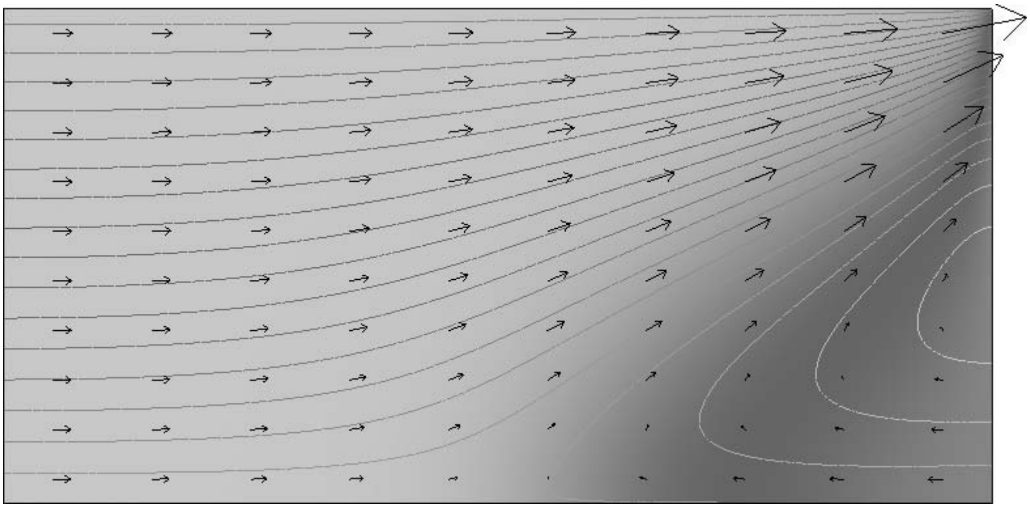

Figure 4. The solution for saltwater intrusion presented by Henry (23).

for saltwater and freshwater was formulated long before the referenced authors. Du Commun (26) already envisaged a U-tube that is filled with fluids of different density on both sides. The equilibrium in that tube was characterized by different fluid levels in both ends of the tube, and the height difference could be calculated from the density difference between the fluids. Du Commun did not provide a formula (like Badon Ghijben) but illustrated the calculation by an example. The background for his derivations was a practical problem: on the ground of a distillery in New Brunswick (New Jersey) a water well had been drilled. At that time, where not much was known about groundwater, the water level in the well and its fluctuations with the tide became the subject of a newspaper dispute (25).

Acknowledgment
The author is grateful B. Kobisch for reprinting original figures.

BIBLIOGRAPHY

1. Drabbe, J. and Badon Ghijben, W. (1889). *Nota in verband met de voorgenomen putboring nabij Amsterdam*. Tijdschrift van het Koninklijk Institut voor Ingenieurs, The Hague, 8–22.
2. Herzberg, A. (1901). Die Wasserversorgung einiger Nordseebäder. *J. Gasbeleuchtung Wasser.* **44**: 815–819, 842–844.
3. De Vries, J.J. (2004). From speculation to science: the founding of groundwater hydrology in the Netherlands. In: *Dutch Pioneers of the Earth Sciences*. R.P.W. Visser and J.L.R. Touret (Eds.). Koninklijke Nederlandse Akademie van Wetenschappen, Amsterdam, pp. 139–164.
4. Fofonoff, P. and Millard, R.C., Jr. (1983). Algorithms for computation of fundamental properties of seawater. *Unesco Tech. Pap. Mar. Sci.* **44**: 53.
5. Bear, J., Cheng, A.H.-D., Sorek, S., Quazar, D., and Herrera, I. (Eds.). (1999). *Seawater Intrusion in Coastal Aquifers—Concepts, Methods and Practices*. Kluwer Academic Publ., Dordrecht, the Netherlands, pp. 163–191.
6. Custodio, E. (2002). Coastal aquifers as important hydrogeological structures. In: *Groundwater and Human Development*. E. Bocanegra, D. Martines, and H. Massone (Eds.). 1905–1918.
7. Kooi, H. and Groen, J. (2001). Offshore continuation of coastal groundwater systems; predictions using sharp-interface approximations and variable-density flow modelling. *J. Hydrol.* **246**: 19–35.
8. Hubbert, M.K. (1940). The theory of ground-water motion. *J. Geol.* **XLVIII**(8): 785–938.
9. Reilly, T.E. and Goodman, A.S. (1985). Quantitative analysis of saltwater–freshwater relationships in groundwater systems—a historical perspective. *J. Hydrol.* **80**: 125–160.
10. Cheng, A.H.-D. and Quazar, D. (1999). Analytical solutions. In: *Seawater Intrusion in Coastal Aquifers—Concepts, Methods and Practices*. J. Bear, A.H.-D. Cheng, S. Sorek, D. Quazar, and I. Herrera (Eds.). Kluwer Academic Publ., Dordrecht, the Netherlands, pp. 163–191.
11. Muskat, M. (1937). *The Flow of Homogeneous Fluids Through Porous Media*. McGraw-Hill, New York.
12. van der Veer, P. (1977). Analytical solution for steady interface flow in a coastal aquifer involving a phreatic surface with precipitation. *J. Hydrol.* **34**: 1–11.
13. Strack, O.D.L. (1976). A single-potential solution for regional interface problems in coastal aquifers. *Water Resour. Res.* **12**(6): 1165–1174.
14. Strack, O.D.L. (1989). *Groundwater Mechanics*. Prentice-Hall, Englewood Cliffs, NJ.
15. Bäckblom, G. (2002). *Experience on Grouting to Limit Inflow to Tunnels*. Research and development and case studies from Sweden, Working Report 2002-18, Posiva Oy, Helsinki, Finland.
16. Wang, F.C. (1965). Approximate theory for skimming well formation in the Indus plain of West Pakistan. *J. Geophys. Res.* **70**(20): 5055–5063.
17. Motz, L.H. (1992). Salt-water upconing in an aquifer overlain by a leaky confining bed. *Groundwater* **30**(2): 192–198.
18. Wirojanagud, P. and Charbeneau, R.J. (1985). Saltwater upconing in unconfined aquifers. *J. Hydrol. Eng.* **111**(3): 417–434.
19. Reilly, T.E., Frimpter, M.H., LeBlanc, D.R., and Goodman, A.S. (1987). Analysis of steady-state salt-water upconing with application at Truro well field. *Groundwater* **25**(2).
20. Lusczynski, N.J. (1961). Head and flow of ground water of variable density. *J. Geophys. Res.* **66**(12): 4247–4256.
21. Holzbecher, E. (1998). *Modeling Density-Driven Flow in Porous Media*. Springer, Heidelberg, Germany.
22. Sorek, S. and Pinder, G.F. (1999). Numerical modelling issues. In: *Seawater Intrusion in Coastal Aquifers—Concepts, Methods and Practices*. J. Bear, A.H.-D. Cheng, S. Sorek, D. Quazar, and I. Herrera (Eds.). Kluwer Academic Publ., Dordrecht, the Netherlands, pp. 399–461.
23. Henry, H.R. (1964). Effects of dispersion on salt water enroachment in coastal aquifers. Geol. Surv. Water-Supply Paper 1613-C.
24. FEMLAB. (2004). Version 3.1, COMSOL AB, Tegnérgatan 23, SE-111 40, Stockholm, Sweden.
25. Carlston, C.W. (1963). An early American statement of the Badon Ghyben–Herzberg principle of static fresh-water–salt-water balance. *Am. J. Sci.* **261**: 88–91.
26. DuCommun, J. (1828). On the causes of fresh water springs, fountains, etc. *Am. J. Sci.* **14**: 174–176.

GROUNDWATER BALANCE

C.P. Kumar
National Institute of Hydrology
Roorkee, India

INTRODUCTION

Rapid industrial development, urbanization, and increase in agricultural production have led to freshwater shortages in many parts of the world. In view of increasing demand of water for various purposes like agricultural, domestic, and industrial, a greater emphasis is being placed on a planned and optimal utilization of water resources. The water resources of the basins remain almost constant while the demand for water continues to increase.

As a result of uneven distribution of rainfall both in time and space, the surface water resources are unevenly distributed. Also, increasing intensities of irrigation from surface water alone may result in an alarming rise of the water table, creating problems of waterlogging and salinization, affecting crop growth adversely and rendering large areas unproductive, which has resulted in increased emphasis on development of groundwater resources. The simultaneous development of groundwater, especially through dug wells and shallow tubewells, will lower the water table and provide vertical drainage, which can prevent waterlogging and salinization. Areas that are already waterlogged can also be reclaimed.

On the other hand, continually increased withdrawals from a groundwater reservoir in excess of replenishable recharge may result in regular lowering of the water table. In such a situation, a serious problem is created resulting in drying of shallow wells and increasing pumping head for deeper wells and tubewells, which has led to an emphasis on planned and optimal development of water resources. An appropriate strategy will be to develop water resources with planning based on conjunctive use of surface water and groundwater.

For a sustainable development of water resources, it is imperative to make a quantitative estimation of the available water resources. For this, the first task would be to make a realistic assessment of the surface water and groundwater resources and then plan their use in such a way that full crop water requirements are met and neither waterlogging nor excessive lowering of the groundwater table occurs. It is necessary to maintain the groundwater reservoir in a state of dynamic equilibrium over a period of a time, and the water level fluctuations have to be kept within a particular range over the monsoon and nonmonsoon seasons.

A complexity of factors, hydrogeological, hydrological, and climatological, control the groundwater occurrence and movement. The precise assessment of recharge and discharge is rather difficult, as no techniques are currently available for their direct measurements. Hence, the methods employed for groundwater resource estimation are all indirect. Groundwater, being a dynamic and replenishable resource, is generally estimated based on the component of annual recharge, which could be subjected to development by means of suitable groundwater structures.

For quantification of groundwater resources, proper understanding of the behavior and characteristics of the water-bearing rock formation, known as an aquifer, is essential. An aquifer has two main functions: (a) to transit water (conduit function) and (b) to store water (storage function). The groundwater resources in unconfined aquifers can be classified as static and dynamic. The static resources can be defined as the amount of groundwater available in the permeable portion of the aquifer below the zone of water level fluctuation. The dynamic resources can be defined as the amount of groundwater available in the zone of water level fluctuation. The replenishable groundwater resource is essentially a dynamic resource that is replenished annually or periodically by precipitation, irrigation return flow, canal seepage, tank seepage, influent seepage, etc.

The methodologies adopted for computing groundwater resources are generally based on the hydrologic budget techniques. The hydrologic equation for groundwater regime is a specialized form of water balance equation that requires quantification of the components of inflow to and outflow from a groundwater reservoir, as well as changes in storage therein. Some of these are directly measurable, few may be determined by differences between measured volumes or rates of flow of surface water, and some require indirect methods of estimation.

Water balance techniques have been extensively used to make quantitative estimates of water resources and the impact of man's activities on the hydrological cycle. The study of water balance requires the systematic presentation of data on the water supply and its use within a given study area for a specific period. The water balance of an area is defined by the hydrologic equation, which is basically a statement of the law of conservation of mass as applied to the hydrological cycle. With the water balance approach, it is possible to quantitatively evaluate individual contribution of sources of water in the system over different time periods and to establish the degree of variation in water regime because of changes in components of the system.

A basinwise approach yields the best results where the groundwater basin can be characterized by prominent drainages. A thorough study of the topography, geology, and aquifer conditions should be employed. The limit of the groundwater basin is controlled not only by topography but also by the disposition, structure and permeability of rocks and the configuration of the water table. Generally, in igneous and metamorphic rocks, the surface water and groundwater basins are coincident for all practical purposes, but marked differences may be encountered in stratified sedimentary formations. Therefore, the study area for groundwater balance study is preferably taken as a doab, which is bounded on two sides by two streams and on the other two sides by other aquifers or extension of the same aquifer. Once the study area is identified, comprehensive studies can be undertaken to estimate for selected period of time, the input and output of water, and change in storage to draw up the water balance of the basin.

The estimation of groundwater balance of a region requires quantification of all individual inflows to or outflows from a groundwater system and change in groundwater storage over a given time period. The basic concept of water balance is:

Input to the system − outflow from the system = change in storage of the system (over a period of time)

The general methodology of computing groundwater balance consists of the following:

- Identification of significant components,
- Evaluating and quantifying individual components, and
- Presentation in the form of water balance equation.

The groundwater balance study of an area may serve the following purposes:

- As a check on whether all flow components involved in the system have been quantitatively accounted for, and what components have the greatest bearing on the problem under study.

- To calculate one unknown component of the groundwater balance equation, provided all other components are quantitatively known with sufficient accuracy.
- As a model of the hydrological processes under study, which can be used to predict the effect that changes imposed on certain components will have on the other components of groundwater system.

GROUNDWATER BALANCE EQUATION

Considering the various inflow and outflow components in a given study area, the groundwater balance equation can be written as:

$$R_{\mathrm{r}} + R_{\mathrm{c}} + R_{\mathrm{i}} + R_{\mathrm{t}} + S_{\mathrm{i}} + I_{\mathrm{g}} = E_{\mathrm{t}} + T_{\mathrm{p}} + S_{\mathrm{e}} + O_{\mathrm{g}} + \Delta S \tag{1}$$

where R_{r} = recharge from rainfall
R_{c} = recharge from canal seepage
R_{i} = recharge from field irrigation
R_{t} = recharge from tanks
S_{i} = influent seepage from rivers
I_{g} = inflow from other basins
E_{t} = evapotranspiration from groundwater
T_{p} = draft from groundwater
S_{e} = effluent seepage to rivers
O_{g} = outflow to other basins
ΔS = change in groundwater storage

Preferably, all elements of the groundwater balance equation should be computed using independent methods. However, it is not always possible to compute all individual components of the groundwater balance equation separately. Sometimes, depending on the problem, some components can be lumped and account only for their net value in the equation.

Computations of various components usually involve errors, because of shortcomings in the estimation techniques. The groundwater balance equation, therefore, generally does not balance, even if all its components are computed by independent methods. The resultant discrepancy in groundwater balance is defined as a residual term in the balance equation, which includes errors in the quantitative determination of various components as well as values of the components that have not been accounted in the equation.

The water balance may be computed for any time interval. The complexity of the computation of the water balance tends to increase with increase in area. This is because of a related increase in the technical difficulty of accurately computing the numerous important water balance components.

DATA REQUIREMENTS FOR A GROUNDWATER BALANCE STUDY

For carrying out a groundwater balance study, the following data may be required over a given time period:

Rainfall Data

Monthly rainfall data of sufficient number of rainguage stations lying within or around the study area, along with their locations, should be available.

Land Use Data and Cropping Patterns

Land use data are required for estimating the evapotranspiration losses from the water table through forested area. Cropping pattern data are necessary for estimating the spatial and temporal distributions of groundwater withdrawals, if required. Monthly pan evaporation rates should also be available at a few locations for estimation of consumptive use requirements of different crops.

River Data

Monthly river stage and discharge data along with river cross sections are required at a few locations for estimating the river-aquifer interflows.

Canal Data

Monthly water releases into the canal and its distributories along with running days during each month are required. To account for the seepage losses through the canal system, the seepage loss test data are required in different canal reaches and distributories.

Tank Data

Monthly tank gauges and water releases should be available. In addition, depth vs. area and depth vs. capacity curves should also be available for computing the evaporation and seepage losses from tanks. Field test data are required for computing infiltration capacity to be used to evaluate the recharge from depression storage.

Water Table Data

Monthly water table data (or at least premonsoon and postmonsoon data) from a sufficient number of well-distributed observation wells along with their locations are required. The available data should comprise reduced level (RL) of water table and depth to water table.

Groundwater Draft

For estimating groundwater withdrawals, the number of each type of well operating in the area, their corresponding running hours each month, and discharge are required. If a complete inventory of wells is not available, then this can be obtained by carrying out sample surveys.

Aquifer Parameters

Data regarding the storage coefficient and transmissivity are required at a sufficient number of locations in the study area.

ESTIMATION OF GROUNDWATER BALANCE COMPONENTS

The various inflow/outflow components of the groundwater balance equation may be estimated through appropriate field experiments or other methods, as discussed below.

Recharge from Rainfall (R_r)

Rainfall is the major source of recharge to groundwater. Part of the rain water that falls on the ground is infiltrated into the soil. A part of this infiltrated water is used in filling the soil moisture deficiency, whereas the remaining portion percolates down to reach the water table, which is termed as *rainfall recharge* to the aquifer. The amount of rainfall recharge depends on various hydrometeorological and topographic factors, soil characteristics, and depth to water table. The methods for estimation of rainfall recharge involve the empirical relationships established between recharge and rainfall developed for different regions, groundwater balance approach, and soil moisture data-based methods.

Empirical Relationships. Several empirical formulae have been worked out for various regions on the basis of detailed studies. For example, Kumar and Seethapathi (1) conducted a detailed seasonal groundwater balance study in Upper Ganga Canal command area (India) for the period 1972–1973 to 1983–1984 to determine groundwater recharge from rainfall. It was observed that as the rainfall increases, the quantity of recharge also increases, but the increase is not linearly proportional. The recharge coefficient (based on the rainfall in monsoon season) was found to vary between 0.05 to 0.19 for the study area. The following empirical relationship was derived by fitting the estimated values of rainfall recharge and the corresponding values of rainfall in the monsoon season through the nonlinear regression technique.

$$R_r = 0.63\ (P - 15.28)^{0.76} \qquad (2)$$

where R_r = Groundwater recharge from rainfall in monsoon season (inch)
P = Mean rainfall in monsoon season (inch)

The relative errors (%) in the estimation of rainfall recharge computed from the proposed empirical relationship was compared with groundwater balance study. In almost every years, the relative error was found to be less than 8%. Therefore, Eq. (2) can conveniently be used for better and quick assessment of natural groundwater recharge in Upper Ganga Canal command area.

It is to be noted that the relationships, tentatively proposed for specific hydrogeological conditions, have to be examined and established or suitably altered for application to other areas. If adequate data of groundwater levels are not available, rainfall recharge may be estimated using the rainfall infiltration method. The same recharge factor may be used for both monsoon and nonmonsoon rainfall, with the condition that the recharge because of nonmonsoon rainfall may be taken as zero, if the rainfall during nonmonsoon season is less than 10% of annual rainfall. An additional 2% of rainfall recharge factor may be used in areas where watershed development with associated soil conservation measures are implemented. This additional factor is separate from contribution because of water conservation structures, such as check dams, nalla bunds, percolation tanks, etc., for which the norms are defined separately.

Groundwater Balance Approach. In this method, all components of the groundwater balance Eq. (1), except the rainfall recharge, are estimated individually. The algebraic sum of all input and output components is equated to the change in groundwater storage, as reflected by the water table fluctuation, which in turn yields the single unknown in the equation, namely, the rainfall recharge. A prerequisite for successful application of this technique is the availability of very extensive and accurate hydrological and meteorological data. The groundwater balance approach is valid for the areas where the year can be divided into monsoon and nonmonsoon seasons, with the bulk of rainfall occurring in former.

Groundwater balance study for monsoon and nonmonsoon periods is carried out separately. The former yields an estimate of recharge coefficient and the latter determines the degree of accuracy with which the components of water balance equation have been estimated. Alternatively, the average specific yield in the zone of fluctuation can be determined from a groundwater balance study for the nonmonsoon period, and using this specific yield, the recharge because of rainfall can be determined using the groundwater balance components for the monsoon period.

Soil Moisture Data-Based Methods. Soil moisture data-based methods are the lumped and distributed model and the nuclear methods. In the lumped model, the variation of soil moisture content in the vertical direction is ignored and any effective input into the soil is assumed to increase the soil moisture content uniformly. Recharge is calculated as the remainder when losses, identified in the form of runoff and evapotranspiration, have been deducted from the precipitation with proper accounting of soil moisture deficit. In the distributed model, variation of soil moisture content in the vertical direction is accounted for, and the method involves the numerical solution of partial differential equation (Richards equation) governing one-dimensional flow through unsaturated medium, with appropriate initial and boundary conditions.

Soil Water Balance Method. Water balance models were developed in the 1940s by Thornthwaite (2) and revised by Thornthwaite and Mather (3). The method is essentially a bookkeeping procedure that estimates the balance between the inflow and outflow of water. When applying this method to estimate the recharge for a catchment area, the calculation should be repeated for areas with different precipitation, evapotranspiration, crop type, and soil type. The soil water balance method is of limited practical value, because evapotranspiration is not directly measurable. Moreover, storage of moisture in the unsaturated zone and the rates of infiltration along the various possible routes to the aquifer form important and uncertain factors. Another aspect that deserves attention is the depth of the root zone, which may vary in semiarid regions between 1 and 30 meters. Results from this model are of very limited value without calibration and validation, because of the substantial uncertainty in input data.

Nuclear Methods. Nuclear techniques can be used for the determination of recharge by measuring the travel of moisture through a soil column. The technique is based on the existence of a linear relation between neutron count rate and moisture content (% by volume) for the range of moisture contents generally occurring in the unsaturated soil zone. The mixture of Beryllium (Be) and Radium (Ra) is taken as the source of neutrons. Another method is the gamma ray transmission method based on the attenuation of gamma rays in a medium through which it passes. The extent of attenuation is closely linked with moisture content of the soil medium.

Recharge from Canal Seepage (R_c)

Seepage refers to the process of water movement from a canal into and through the bed and wall material. Seepage losses from irrigation canals often constitute a significant part of the total recharge to groundwater system. Hence, it is important to properly estimate these losses for recharge assessment to groundwater system. Recharge by seepage from canals depends on the size and cross section of the canal, depth of flow, characteristics of soils in the bed and sides, and location as well as level of drains on either side of the canal. A number of empirical formulae and formulae based on theoretical considerations have been proposed to estimate the seepage losses from canals.

Recharge from canals that are in direct hydraulic connection with a phreatic aquifer, underlaid by a horizontal impermeable layer at shallow depth, can be determined by Darcy's equation, provided the flow satisfies Dupuit assumptions.

$$R_c = K\frac{h_s - h_l}{L}A \tag{3}$$

where, h_s and h_l are water-level elevations above the impermeable base, at the canal, respectively, and at distance L from it. For calculating the area of flow cross section, the average of the saturated thickness $(h_s + h_l)/2$ is taken. The crux of computation of seepage depends on correct assessment of the hydraulic conductivity, K. Knowing the percentage of sand, silt, and clay, the hydraulic conductivity of undisturbed soil can be approximately determined using the soil classification triangle showing relation of hydraulic conductivity to texture for undisturbed sample (4).

A number of investigations have been carried out to study the seepage losses from canals. United States Bureau of Reclamation (USBR) recommended the channel losses based on the channel bed material as given below:

Material	Seepage Losses (Cumec Per Million Square Meter of Wetted Area)
Clay and clay loam:	1.50
Sandy loam:	2.40
Sandy and gravely soil:	8.03
Concrete lining:	1.20

These values are valid if the water table is relatively deep. In shallow water table and waterlogged areas, the recharge from canal seepage may be suitably reduced. Specific results from case studies may be used, if available. The above norms take into consideration the type of soil in which the canal runs while computing seepage. However, the actual seepage will also be controlled by the width of canal (B), depth of flow (D), hydraulic conductivity of the bed material (K), and depth to water table. Knowing the values of B and D, the range of seepage losses (R_{c_max} and R_{c_min}) from the canal may be obtained as

$$R_{c\_max} = K(B + 2D) \quad \text{(in case of deeper water table)} \tag{4a}$$

$$R_{c\_min} = K(B - 2D) \quad \text{(in case of water table at the level of channel bed)} \tag{4b}$$

However, the various guidelines for estimating losses in the canal system are only approximate. The seepage losses may best be estimated by conducting actual tests in the field. The methods most commonly adopted are:

Inflow-Outflow Method. In this method, the water that flows into and out of the section of canal under study is measured using current meter or Parshall flume method. The difference between the quantities of water flowing into and out of the canal reach is attributed to seepage. This method is advantageous when seepage losses are to be measured in long canal reaches with few diversions.

Ponding Method. In this method, bunds are constructed in the canal at two locations, one upstream and the other downstream of the reach of canal with water filled in it. The total change in storage in the reach is measured over a period of time by measuring the rate of drop of water surface elevation in the canal reach. Alternatively, water may be added to maintain a constant water surface elevation. In this case, the volume of water added is measured along with the elapsed time to compute the rate of seepage loss. The ponding method provides an accurate means of measuring seepage losses and is especially suitable when they are small (e.g., in lined canals).

Seepage Meter Method. The seepage meter is a modified version of permeameter developed for use under water. Various types of seepage meters have been developed. The two most important are seepage meter with submerged flexible water bag and falling head seepage meter. Seepage meters are suitable for measuring local seepage rates in canals or ponds and used only in unlined or earth-lined canals. They are quickly and easily installed and give reasonably satisfactory results for the conditions at the test site, but it is difficult to obtain accurate results when seepage losses are low.

The total losses from the canal system generally consist of the evaporation losses (E_c) and the seepage losses (R_c). The evaporation losses are generally 10–15% of the total losses. Thus, the R_c value is 85–90% of the losses from the canal system.

Recharge from Field Irrigation (R_i)

Water requirements of crops are met, in parts, by rainfall, contribution of moisture from the soil profile, and applied irrigation water. A part of the water applied to irrigated field crops is lost in consumptive use and the balance infiltrates to recharge the groundwater. The process of re-entry of a part of the water used for irrigation is called return flow. Percolation from applied irrigation water, derived both from surface water and groundwater sources, constitutes one of the major components of groundwater recharge. The irrigation return flow depends on the soil type, irrigation practice, and type of crop. Therefore, irrigation return flows are site specific and will vary from one region to another.

The recharge because of irrigation return flow may be estimated, based on the source of irrigation (groundwater or surface water), the type of crop (paddy, nonpaddy), and the depth of water table below ground surface. For surface water, the recharge is to be estimated based on water released at the outlet from the canal/distribution system. For groundwater, the recharge is to be estimated based on gross draft. Where continuous supply is used instead of rotational supply, an additional recharge of 5% of application may be used. Specific results from case studies may be used, if available.

For a correct assessment of the quantum of recharge by applied irrigation, studies are required to be carried out on experimental plots under different crops in different seasonal conditions. The method of estimation comprises application of the water balance equation involving input and output of water in experimental fields.

Recharge from Tanks (R_t)

Studies have indicated that seepage from tanks varies from 9–20% of their live storage capacity. However, as data on live storage capacity of large number of tanks may not be available, seepage from the tanks may be taken as 44–60 cm per year over the total water spread, taking into account the agroclimatic conditions in the area. The seepage from percolation tanks is higher and may be taken as 50% of its gross storage. In the case of seepage from ponds and lakes, the norms as applied to tanks may be taken. Groundwater Resource Estimation Committee (5) has recommended that, based on the average area of water spread, the recharge from storage tanks and ponds may be taken as 1.4 mm/day for the period in which the tank has water. If data on the average area of water spread is not available, 60% of the maximum water spread area may be used instead of the average area of water spread.

In the case of percolation tanks, recharge may be taken as 50% of gross storage, considering the number of fillings, with half of this recharge occurring in monsoon season and the balance in nonmonsoon season. Recharge because of check dams and nala bunds may be taken as 50% of gross storage (assuming annual desilting maintenance exists), with half of this recharge occurring in the monsoon season and the balance in the nonmonsoon season.

Influent and Effluent Seepage (S_i & S_e)

The river-aquifer interaction depends on the transmissivity of the aquifer system and the gradient of the water table in respect to the river stage. Depending on the water level in the river and in the aquifer (in the vicinity of river), the river may recharge the aquifer (influent) or the aquifer may contribute to the river flow (effluent). The effluent or influent character of the river may vary from season to season and from reach to reach. The seepage from/to the river can be determined by dividing the river reach into small sub-reaches and observing the discharges at the two ends of the sub-reach along with the discharges of its tributaries and diversions, if any. The discharge at the downstream end is expressed as:

$$Q_{\mathrm{d}}.\Delta t = Q_{\mathrm{u}}.\Delta t + Q_{\mathrm{g}}.\Delta t + Q_{t}.\Delta t - Q_{\mathrm{o}}.\Delta t - E.\Delta t \pm S_{\mathrm{rb}} \quad (5)$$

where Q_{d} = discharge at the downstream section
Q_{u} = discharge at the upstream section
Q_{g} = groundwater contribution (unknown quantity; −ve computed value indicates influent conditions)
Q_{t} = discharge of tributaries
Q_{o} = discharge diverted from the river
E = rate of evaporation from river water surface and flood plain (for extensive bodies of surface water and for long time periods, evaporation from open water surfaces cannot be neglected)
S_{rb} = change in bank storage (+ for decrease and − for increase)
Δt = time period

The change in bank storage can be determined by monitoring the water table along the cross section normal to the river. Thus, using the above equation, seepage from/to the river over a certain period of time, Δt, can be computed. However, this would be the contribution from aquifers on both sides of the stream. The contribution from each side can be separated by the following method:

$$\textit{Contribution from left bank} = \frac{I_{\mathrm{L}}T_{\mathrm{L}}}{I_{\mathrm{L}}T_{\mathrm{L}} + I_{\mathrm{R}}T_{\mathrm{R}}} \cdot Q_{\mathrm{g}} \quad (6a)$$

$$\textit{Contribution from right bank} = \frac{I_{\mathrm{R}}T_{\mathrm{R}}}{I_{\mathrm{L}}T_{\mathrm{L}} + I_{\mathrm{R}}T_{\mathrm{R}}} \cdot Q_{\mathrm{g}} \quad (6b)$$

where I_{L} and T_{L} are gradient and transmissivity, respectively, on the left side and I_{R} and T_{R} are those on the right.

Inflow from and Outflow to Other Basins (I_g and O_g)

For the estimation of groundwater inflow/outflow from/to other basins, regional water table contour maps are drawn based on the observed water level data from wells located within and outside the study area. The flows into and out of a region are governed mainly by the hydraulic gradient and transmissivity of the aquifer. The gradient can be determined by taking the slope of the water table normal to water table contour. The length of the section, across which groundwater inflow/outflow occurs, is determined from contour maps, the length being measured parallel to the contour. The inflow/outflow is determined as follows:

$$I_{\mathrm{g}} \; or \; O_{\mathrm{g}} = \sum^{\mathrm{L}} T \, I \, \Delta L \quad (7)$$

where T is the transmissivity and I is the hydraulic gradient averaged over a length, ΔL, of contour line.

Evapotranspiration from Groundwater (E_t)

Evapotranspiration is the combined process of transpiration from vegetation and evaporation from both soil and free water surfaces. Potential evapotranspiration is the maximum loss of water through evapotranspiration. Evapotranspiration from groundwater occurs in waterlogged areas or in forested areas with roots extending to the water table. From the land use data, area under forests is available whereas the waterlogged areas may be demarcated from depth to water table maps. The potential evapotranspiration from such areas can be computed using standard methods.

Depth to water table maps may be prepared based on well inventory data to bring into focus the extensiveness of shallow water table areas. During well inventory, investigation should be specifically oriented toward accurately delineating water table depth for depths less than 2 meters. The evapotranspiration can be estimated based on the following equations:

$$E_t = PE_t * A \qquad \text{if } h > h_s \qquad (8a)$$

$$E_t = 0 \qquad \text{if } h < (h_s - d) \qquad (8b)$$

$$E_t = PE_t * A(h - (h_s - d))/d \qquad \text{if } (h_s - d) \le h \le h_s \qquad (8c)$$

where E_t = evapotranspiration in volume of water per unit time$[L^3T^{-1}]$
PE_t = maximum rate of evapotranspiration in volume of water per unit area per unit time $[L^3L^{-2}T^{-1}]$
A = surface area $[L^2]$
h = water table elevation $[L]$
h_s = water table elevation at which the evapotranspiration loss reaches the maximum value
d = extinction depth; when the distance between h_s and h exceeds d, evapotranspiration from groundwater ceases $[L]$

Draft from Groundwater (T_p)

Draft is the amount of water lifted from the aquifer by means of various lifting devices. To estimate groundwater draft, an inventory of wells and a sample survey of groundwater draft from various types of wells (state tubewells, private tubewells, and open wells) are required. For state tubewells, information about their number, running hours per day, discharge, and number of days of operation in a season is generally available in the concerned departments. To compute the draft from private tubewells, pumping sets, rahats, etc., sample surveys have to be conducted regarding their number, discharge, and withdrawals over the season.

In areas where wells are energized, the draft may be computed using power consumption data. By conducting tests on wells, the average draft per unit of electricity consumed can be determined for different ranges in depth to water levels. By noting the depth to water level at each distribution point and multiplying the average draft value with the number of units of electricity consumed, the draft at each point can be computed for every month.

Change in Groundwater Storage (ΔS)

To estimate the change in groundwater storage, the water levels are observed through a network of observation wells spread over the area. The water levels are highest immediately after monsoon, in the month of October or November, and lowest just before rainfall, in the month of May or June. During the monsoon season, the recharge is more than the extraction; therefore, the change in groundwater storage between the beginning and end of monsoon season indicates the total volume of water added to the groundwater reservoir, whereas the change in groundwater storage between the beginning and end of the nonmonsoon season indicates the total quantity of water withdrawn from groundwater storage. The change in storage (ΔS) is computed as follows:

$$\Delta S = \sum \Delta h\, A\, S_y \qquad (9)$$

where Δh = change in water table elevation during the given time period
A = area influenced by the well
S_y = specific yield

Groundwater Resource Estimation Committee (5) recommended that the size of the watershed as a hydrological unit could be of about 100 to 300 sq. km area, and there should be at least three spatially well-distributed observation wells in the unit, or one observation well per 100 sq. km, whichever is more. However, as per IILRI (6), the following specification may serve as a rough guide:

Size of the Area (ha)	Number of Observation Points	Number of Observation Points per 100 Hectares
100	20	20
1,000	40	4
10,000	100	1
1,00,000	300	0.3

The specific yield may be computed from pumping tests. The values of specific yield in the zone of fluctuation of water table in different parts of the basin can also be approximately determined from the soil classification triangle showing relation between particle size and specific yield (7).

ESTABLISHMENT OF RECHARGE COEFFICIENT

Groundwater balance study is a convenient way of establishing the rainfall recharge coefficient, as well as to cross check the accuracy of the various prevalent methods for the estimation of groundwater losses and recharge from other sources. The steps to be followed are:

1. Divide the year into monsoon and nonmonsoon periods.
2. Estimate all the components of the water balance equation other than rainfall recharge for monsoon period using the available hydrological and meteorological information and employing the prevalent methods for estimation.
3. Substitute these estimates in the water balance equation and thus calculate the rainfall recharge and, hence, recharge coefficient (recharge/rainfall ratio). Compare this estimate with those given by various empirical relations valid for the area of study.
4. For nonmonsoon season, estimate all the components of water balance equation including the rainfall recharge, which is calculated using recharge coefficient value obtained through the water balance of monsoon period. The rainfall recharge (R_r) will be of very small order in this case. A close balance between the left and right sides of the equation will indicate that the net recharge from all the sources of recharge and discharge has been quantified with a good degree of accuracy.

By quantifying all the inflow/outflow components of a groundwater system, one can determine which particular component has the most significant effect on the groundwater flow regime. Alternatively, a groundwater balance study may be used to compute one unknown component (e.g., the rainfall recharge) of the groundwater balance equation when all other components are known. The balance study may also serve as a model of the area under study, whereby the effect of change in one component can be used to predict the effect of changes in other components of the groundwater system. In this manner, the study of groundwater balance has a significant role in planning a rational groundwater development of a region.

CONCLUDING REMARKS

- Water balance approach, essentially a lumped model study, is a viable method of establishing the rainfall recharge coefficient and for evaluating the methods adopted for the quantification of discharge and recharge from other sources. For proper assessment of potential, present use, and additional exploitability of water resources at optimal level, a water balance study is necessary.
- Groundwater exploitation should be such that protection from depletion is provided, protection from pollution is provided, negative ecological effects are reduced to a minimum, and economic efficiency of exploitation is attained. Determination of exploitable resources should be based on hydrological investigations. These investigations logically necessitate use of a mathematical model of groundwater system for analyzing and solving the problems. The study of water balance is a prerequisite for groundwater modeling.
- A need exists for studying unsaturated and saturated flow through weathered and fractured rocks for finding the recharge components from rainfall and from percolation tanks in hard rock groundwater basins. The irrigation return flow under different soils, crops, and irrigation practices needs to be quantified. Assessment of groundwater quality in many groundwater basins is a task yet to be performed. A hydrological database for groundwater assessment should be established. Also, user-friendly software should be developed for quick assessment of regional groundwater resources.
- Nonconventional methods for utilization of water, such as through interbasin transfers, artificial recharge of groundwater, and desalination of brackish or seawater as well as traditional water conservation practices like rainwater harvesting, including rooftop rainwater harvesting, need to be practiced to further increase the usable water resources.

BIBLIOGRAPHY

1. Kumar, C.P. and Seethapathi, P.V. (2002). Assessment of natural groundwater recharge in Upper Ganga Canal command area. *J. Appl. Hydrol. Assoc. Hydrologists of India* **XV**(4): 13–20.
2. Thornthwaite, C.W. (1948). An approach towards a rational classification of climate. *Geogr. Rev.* **38**(1): 55–94.
3. Thornthwaite, C.W. and Mather, J.W. (1955). The water balance. *Publ. Climatol. Lab. Climatol. Drexel Inst. Technol.* **8**(1): 1–104.
4. Johnson, A.I. (1963). *Application of Laboratory Permeability Data, Open File Report*. USGS, Water Resources Division, Denver, CO, p. 34.
5. Groundwater Resource Estimation Methodology. (1997). *Report of the Groundwater Resource Estimation Committee*. Ministry of Water Resources, Government of India, New Delhi, India.
6. IILRI. (1974). *Drainage Principles and Applications, Survey and Investigation*. Publication 16, **III**.
7. Johnson, A.I. (1967). *Specific Yield—Compilation of Specific Yields for Various Materials, Water Supply Paper, USGS*. Denver, CO, p. 74.

HYDRAULIC HEAD

MATTHEW M. ULIANA
Texas State University—San Marcos
San Marcos, Texas

Hydraulic head is defined as the fluid energy per unit weight at a given point in a fluid system (like a pipe filled with flowing water or an aquifer). In simple, everyday terms, we can conceptualize (and measure) hydraulic head as an elevation—more specifically, the elevation to which water rises in a manometer, in a pressurized water pipe or in a piezometer.

Water elevations is a manifestation of the fluid energy in a groundwater system, and measuring a water elevation in a piezometer is really a measurement of the energy in the fluid at a certain point in a groundwater system.

WHAT DO WE MEAN BY "ENERGY"?

Everything in the universe has some amount of energy associated with it, and that energy is present in various forms. Energy drives every natural process, and the key to understanding physical processes is understanding how energy is distributed in a system.

There are two basic types of energy:

Potential Energy: energy stored in a piece of matter or at a point in a system, generally associated with position or with the thermodynamics of the system (elevation, pressure, chemical, thermal).

Kinetic: energy of motion (velocity).

At every point in an aquifer, the fluid possesses some total amount of energy. That energy is the *sum* of all the potential and kinetic energy in the fluid associated with the velocity of the fluid, pressure of the fluid, temperature of the fluid, chemical bonds in the fluid, etc.

As previously stated, the fluid energy at a point in an aquifer manifests itself as the water level in a piezometer. So we could also say that the water level, or hydraulic head, represents the total energy in the aquifer at a given point, and we can use the various energy components of the hydraulic head (elevation, pressure, velocity, etc.) to understand the driving forces behind fluid motion in the subsurface.

THE BERNOULLI EQUATION

As previously stated, fluid energies (and, subsequently, water levels) vary from one point in an aquifer to the next (Figs. 1 and 2). Now, if we recall the first and second laws of thermodynamics:

First law of thermodynamics: Energy is conserved.

$$\text{Energy added} - \text{energy subtracted}$$
$$= \text{change in total energy}$$

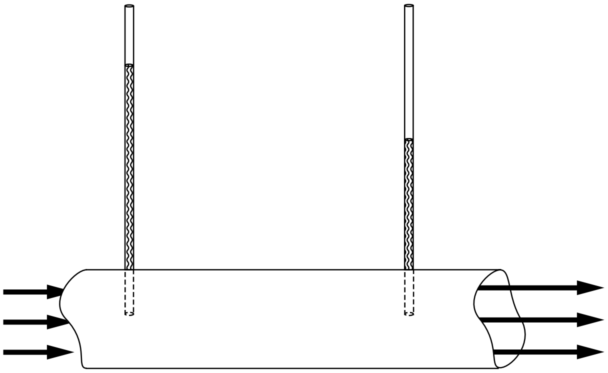

Figure 1. Pipe of flowing water with manometers showing the loss of head along the flow path.

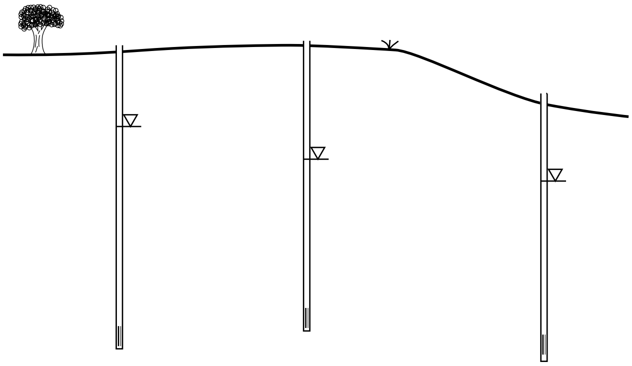

Figure 2. Cross section of aquifer showing hydraulic heads in three wells.

We could also express the first law in terms of the difference in energy at two points in a dynamic system:

$$\text{Total energy}_{(\text{atpoint1})} + \text{energy added/lost}_{(\text{betweenpoint1and2})}$$
$$= \text{total energy}_{(\text{atpoint2})}$$

Second law of thermodynamics: Closed systems move toward increasing entropy.

— In a dynamic system like an aquifer, water will move from a point of higher energy (i.e., lower entropy) to a point of lower energy (higher entropy).

— In other words, groundwater moves in the direction of decreasing hydraulic head.

The *Bernoulli equation*, which describes the total energy of a fluid at all positions along a flow path in a closed system, is basically an expression of the first and second laws of thermodynamics:

$$z_1 + \frac{p_1}{\rho_w g} + \frac{v_1^2}{2g} + H_a = z_2 + \frac{p_2}{\rho_w g} + \frac{v_2^2}{2g} + H_L + H_E \tag{1}$$

where z = elevation (L)
p = pressure ($\mathrm{M{\cdot}L^{-1}{\cdot}t^{-2}}$)
ρ_w = fluid density ($\mathrm{M{\cdot}L^{-3}}$)
g = gravitational acceleration ($\mathrm{L{\cdot}t^{-2}}$)
v = velocity ($\mathrm{L{\cdot}t^{-1}}$)
H_a = heat energy added (L)
H_L = mechanical energy lost (L)
H_E = heat energy extracted (L)

and the numerical subscripts represent two different positions along the flow path.

Let's assume that we don't add or subtract any heat energy from the system and the only change is the loss of mechanical energy from one point to the next; then we can rewrite the equation as

$$z_1 + \frac{p_1}{\rho_w g} + \frac{v_1^2}{2g} + I_1 = z_2 + \frac{p_2}{\rho_w g} + \frac{v_2^2}{2g} + I_2 \tag{2}$$

where **I** is the internal energy at each point (i.e., the rest of the potential energy not described by the other terms) and $\mathbf{I}_2\mathbf{I}_1$ is equal to $\mathbf{H}_L$ in Eq. 1.

We can understand this equation by considering that each term in the equation represents a specific component of energy with units of length. The terms of the equation represent, (respectively),

— elevation
— pressure
— kinetic (velocity)
— other internal energies (thermal, chemical)

This equation describes the change in energy from one point along a flow path to the next point. This equation expresses all the components of energy in the same units (i.e., length), so we can use it to compare the relative magnitude of the individual components. When we do that, we see that for most groundwater situations, the internal and kinetic components of the total energy are so small that we can ignore them. The result is that we can describe fluid energy in a groundwater system by only the elevation and pressure components.

HYDRAULIC HEAD AND HYDRAULIC POTENTIAL

If we assume that we can ignore velocity and internal energy components when dealing with groundwater, we can drop all that out of the equation and express the fluid energy as the sum of the elevation and pressure components. That sum is what we call *hydraulic head*; in physical terms, it is the fluid energy per unit weight, and in mathematical terms, it is

$$h = z + p/\rho_w g \tag{3}$$

where h = hydraulic head (L)
z = elevation (L)
p = pressure ($\mathrm{M{\cdot}L^{-1}{\cdot}t^{-2}}$)
ρw = fluid density ($\mathrm{M{\cdot}L^{-3}}$)
g = gravitational acceleration ($\mathrm{L{\cdot}t^{-2}}$)

If we multiply both sides of the equation by the gravitational constant, g, we get a quantity called *hydraulic potential* (Φ), which is the fluid energy per unit *mass*, or

$$\Phi = gz + p/\rho_w \tag{4}$$

such that

$$\Phi = gh \tag{5}$$

The hydraulic potential is simply a way of expressing the same fluid energy, so that it is independent of gravity.

PHYSICAL DESCRIPTION OF THE COMPONENTS OF HEAD

From Eq. 3, we can see that hydraulic head is the sum of the elevation component (or *elevation head*) and the pressure component (or *pressure head*). The physical meaning of pressure head and elevation head are defined and described in Fig. 3a–c. These figures show an idealized cross section of a piezometer that has a screened interval at the bottom. Figure 3a shows the depths and elevations measured in the field (relative to some datum, like mean sea level). Figure 3b shows the pressure, elevation, and total heads, and Fig. 3c is a combination of the other two figures.

From Fig. 3b, we see that the elevation head is the height of the screened interval above the datum. Keep in mind that the datum is arbitrarily chosen—if we wanted, we could choose the bottom of the well as the datum, and the elevation head would be zero. However, it is important to realize that the head is really only important to us when we are looking at multiple wells in the same aquifer, and we need to have a constant datum for all those wells make the comparison meaningful. Sea level is usually chosen, but it is not the necessary datum.

We also see that the pressure head is the length of the column of water in the well above the screened interval. Keep in mind that *pressure* and *pressure head* are two different things; the *pressure* at the screened interval is the force per unit area of the column of fluid above that point; the *pressure head* is the pressure divided by the product of the density of the fluid and gravitational acceleration and is manifest as the length of the column of water above the screen. If we know the density of the fluid and the length of the column of water in the well, we can calculate the pressure at the well screen.

Finally, we see that the total head is just the sum of the other two heads, or more generally, the height of the water level in the well above the datum, this brings us full circle to the concept of water levels reflecting fluid energy.

DISTRIBUTION OF HEADS (I.E., FLUID ENERGY) IN A 3-D AQUIFER

An aquifer is a dynamic system of flowing water. The fluid energy varies throughout the system and is different from one point to the next. The result is

— different water levels throughout the aquifer
— hydraulic gradients
— movement of water driven by those gradients

Head is a scalar quantity; it is measured at a point, and it has a single magnitude that doesn't vary with respect to direction. So, when we talk about heads, from a theoretical standpoint, we are talking about the energy at an infinitesimal point in an aquifer. However, that aquifer exists in three dimensions, contains an infinite number of points, and the head varies from point to point.

The trends in that variation control the directions and magnitudes of flow. We call this variation the *hydraulic gradient*. The hydraulic gradient is defined as the change in hydraulic head over the change in length, and it is directly analogous to other physical gradients (topographic slope, thermal gradient, concentration gradients, etc.)

Hydraulic heads exist in three dimensions, so hydraulic gradients are not necessarily horizontal. In reality, most groundwater flow is generally in a horizontal direction, and it is often a realistic assumption to ignore fluid movement in a vertical dimension. However, there are many situations where vertical gradients (and, subsequently, vertical flow) are significant.

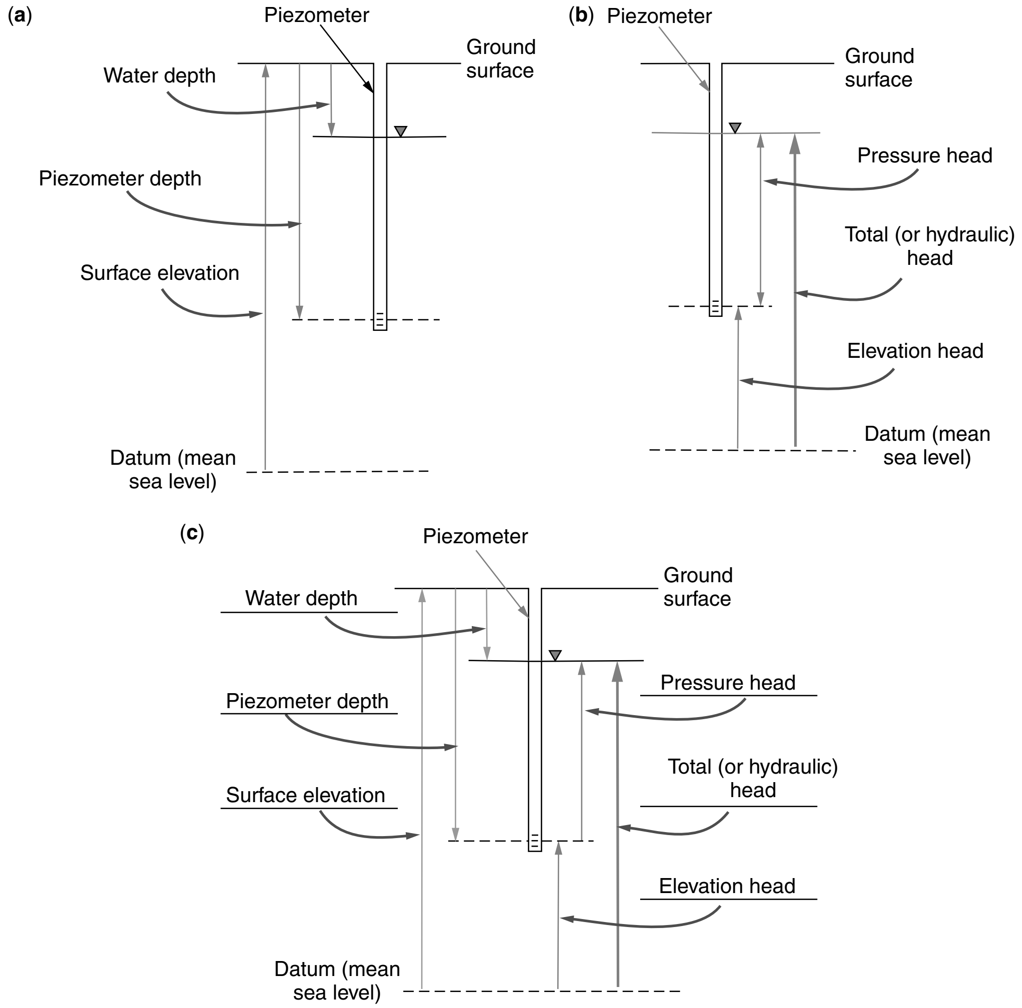

Figure 3. Hypothetical well showing depths, elevations, and heads.

THE ROLE OF HEAT IN GROUNDWATER SYSTEMS

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

The role of heat in the subsurface includes the study of the overall transport of heat in saturated groundwater systems. Once site specific dynamics of heat transport are understood the effects of temperature on physical, chemical, and biological processes that govern the transport and degradation of contaminants in an aquifer can be evaluated.

FUNDAMENTALS

Of the many parameters that are monitored in association with the assessment and remediation of contaminated groundwater, temperature is typically regarded as a background condition that must be accepted for a given site. In the northern tier of states, it is customary that in the winter months remediation processes will be slowed or stopped depending on the depth of groundwater.

DEFINITIONS

Heat is an added or external energy that causes a rise in temperature, physical expansion, evaporation or other physical change. The rise or drop of temperature in an aquifer is the result of the transport of heat into or away from the aquifer. The key measure of the relationship between heat flow and resulting temperature is heat capacity, Cp. Heat capacity is used as specific heat in English units, which is the number of BTUs required to raise one pound of aquifer matrix by one degree Fahrenheit.

There are two primary materials through which heat is transported in an aquifer, the mineral matrix and the entrained groundwater. Density of the mineral matrix and water, porosity, and heat capacity combine to govern the temperature response of an aquifer to heat flow. The specific heat Cp (as BTU/lb °F), dry density (rho, as lbs/ft^3) and percent porosity (phi) of some soils and minerals are as follows:

	Specific Heat	CpDensity (rho) lbs/ft^3	% Porosity (phi)
Water	1.00	62.4	NA
Clay	0.27	65.0	75
Sand	0.19	110.0	35
Granite	0.20	165.0	1
Limestone	0.22	155	4
Organic Fraction	0.30	55	65

The thermal energy capacity per unit volume per degree of temperature change is called the aquifer thermal capacity (q). With the above information, it is possible to calculate the thermal capacity (q) of an aquifer:

$$\mathbf{q} = (\mathrm{Cp} \times \mathrm{rho})_{\mathrm{rock}}(1 - \mathrm{phi}) + (\mathrm{Cp} \times \mathrm{rho})_{\mathrm{water}} \times \mathrm{phi}$$

Assuming a sandy aquifer with a specific heat of 0.19 for the mineral matrix, a density of 110 lb/ft^3, and a porosity of 0.35, the specific thermal capacity per cubic foot of aquifer is

$$\begin{aligned}\mathbf{q} &= (0.19 \times 110)_{\mathrm{rock}} \times (1 - 0.35) \\ &\quad + (1.0 \times 62.4)_{\mathrm{water}} \times 0.35 \\ \mathbf{q} &= 13.59_{\mathrm{rock}} + 21.84_{\mathrm{water}} \\ &= 35.4 \text{ BTU per cubic foot per 1 degree Fahrenheit}\end{aligned}$$

EXAMPLE

Given a volume in an aquifer 100 feet by 100 feet by 20 feet in depth, it would require 7,080,000 Btu to raise the temperature 1 °F. Converting Btu to kilowatt hours gives 2073 KW hours. At 6 cents per Kilowatt, that is equal to $125 in electrical cost. That estimate does not take into account the efficiency of the heat generation system, transfer of that heat into the aquifer, and effects of heat transport away from the zone undergoing heating.

Three physical processes are responsible for heat transfer in groundwater systems:

1. Conduction—heat flows from hotter regions to cooler regions through the molecular transfer of kinetic energy.
2. Convection—heat is transported along with overall mass transport of groundwater.
3. Radiation—the transference of heat through space via electromagnetic radiation.

The dominant mechanisms in groundwater systems are convection and conduction.

Thermal conductivity (k) is the amount of heat that will flow through a unit area in a unit time with a unit temperature difference. Under saturated conditions, the value of thermal conductivity is primarily governed by the texture, degree of compaction, and mineralogical composition of the aquifer matrix. Thermal conductivity decreases with the reduction of particle size. The units of thermal conductivity are Btu per hour per foot per degree Fahrenheit. Typical values are as follows:

Water	0.346
Air	0.0145
Wet Sand	0.95
Dry Sand	0.157
Wet Clay	0.87
Dry Clay	0.138

Notice that the thermal conductivity of water is about 23 times that of air. In the past, the use of heat as an augmentation to remediation has been applied to the vadose zone through the application of steam. For heat transport efficiency, that is probably not the most efficient application of a subsurface heating process.

Heat flow through a mass (qx) is calculated by the equation:

$$\mathrm{qx} = Ak(dT/dX)$$

where A is the cross-sectional area normal to the heat flow,
k is the thermal conductivity,
T is temperature in Fahrenheit,
X is the distance over which the heat must flow.

The solution to the above function requires substitution of an equation describing the system geometry for the dX term followed by integration. That is beyond the scope of this article.

Thermal diffusivity (in units of square feet per hour) is equal to the thermal conductivity (k) divided by the thermal capacity (q) of the aquifer; it measures the rate at which temperature changes occur in the soil mass. Higher values of thermal diffusivity result in more rapid changes in temperature and deeper penetration of heat into the soil.

Ratio of Thermal Velocity to Groundwater Velocity Versus Porosity

Lastly, as heated groundwater moves through an aquifer, thermal energy is transferred to the mineral matrix. As a consequence the thermal front associated with the advective flow of heated groundwater will move at a lower velocity than the groundwater. Aquifer porosity is the dominant element in this dynamic. The ratio of the velocity of the migration of the thermal front versus the groundwater velocity is affected by aquifer porosity. The smaller the ratio, the slower the thermal front velocity compared with the overall groundwater velocity.

EFFECT ON PHYSICAL PROPERTIES AND REACTION KINETICS

The two dominant physical characteristics of groundwater that change with temperature are viscosity and density. Their relationship is not linear. In the temperature range of 0 °C to 30 °C, the viscosity decreases by over 50%. This change improves the efficiency of both advective fluid flow and diffusional transport. In contrast, water density decreases significantly at temperatures over 60 °C.

Next, chemical reactions should be considered. Most salts increase in solubility as the temperature increases, the common exception to this general rule being carbonates in which solubility decreases as temperature increases. As temperature increases, the solubility of gases in groundwater decreases. The solubility (in milligrams/liter) of environmentally important gases over the temperature range of 5 °C to 35 °C is as follows:

Oxygen with air as the source	12.7 to 7.1
Carbon Dioxide	2774 to 1105
Methane	34.1 to 17.3

The relationship between temperature and the kinetics of chemical reactions is expressed by the Arrhenius equation:

$$\ln k = \ln A - E_a/RT$$

where k is the rate constant,
A is an integration constant,
E_a is the energy of activation,
R is the universal gas constant, and
T is the temperature in Kelvin.

Commonly, this relationship is graphed as ln k versus $1/T$, which results in a straight line in which k increases exponentially as T increases. Under proper circumstances, small increases in temperature can have large effects on kinetic rates.

In addition to reaction rates, temperature also affects equilibrium constants, particularly those that involve reactions with noncovalent bonds. Covalent bonds are relatively resistant to thermal perturbation. However, van der Waals interactions, hydrogen bonds, weak ionic bonds, and hydrophobic interactions are disrupted by relatively small temperature changes.

Hydrogen bonds, ionic bonds, and van der Waals interactions all form with the release of heat, and hydrophobic interactions form with heat consumption.

Increases in temperature will destabilize the first three, whereas it will stabilize the hydrophobic interactions.

As temperature increases, the solubility of hydrocarbons increases, not uncommonly two to three times in the temperature range of 5 °C to 30 °C. Through this and the bonding reactions mentioned above, the forces responsible for the retardation of hydrocarbons within the soil matrix are reduced as temperature increases, improving the flushing action of groundwater flow induced by pump and treat recovery systems.

Vapor pressure also increases with temperature, and increased vapor pressure results in increased mass transport rates in soil vapor extraction systems, which is where many of the prior applications of external heat to subsurface remediation have been used. Examples of the percent increase in vapor pressure as a result of a temperature increase from 5 °C to 35 °C are as follows:

Benzene	417%
Ethylbenzene	638%
Toluene	513%
TCE	476%
Tetrachloroethylene	545%

With regard to free product recovery, an increase in temperature will reduce the hydrocarbon viscosity and increase the recovery of free product from the surface of a water table. A temperature increase from 10 °C to 50 °C can increase the recoverable amount of middle distillates by 10%.

Molecular diffusion is driven by the presence of a chemical concentration gradient between adjacent zones in the subsurface. The presence of a temperature gradient will also set up the diffusion of dissolved components in a fluid. The phenomena is called the Soret Effect and is known generally as thermal diffusion. Mass transport rates under thermal diffusion are described by the equation:

$$N_{\mathrm{AT}} = D_{\mathrm{T}}(\mathrm{Rho})d\ln\,(T/dz)$$

where N_{AT} is the mass transport rate as number of moles/cm2/second, D_{T} is the coefficient of thermal diffusion in centimeters squared/second, T is temperature in kelvins, z is the distance in the direction of diffusion in centimeters, and Rho is the fluid density in g-moles/centimeters cubed.

The magnitude of thermal diffusion is dependent on the size and chemistry of the molecules involved as well as on the temperature. At maximum, it may reach a diffusional mass transport rate of 30% that is seen for molecular diffusion driven by chemical concentration gradients, and it is typically significantly less than that.

Lastly, a brief discussion on temperature and biological systems. Bacteria typically have a relatively narrow temperature range (about 10 °C) in which they experience maximum growth and metabolic activity. As groundwater temperatures increase, the dominant bacterial consortia will change in response. This is true with increases of temperature up to about 35 °C; temperatures of 50 °C or above will cause traumatization and partially kill many bacterial species indigenous to the subsurface. It is common for the rates of physiological and biochemical processes to undergo a two-fold increase because of a temperature change of only about 10 °C. This change is caused by the nature of the Arrhenius equation previously discussed as applied to biological systems.

A practical note with regard to bench scale studies of soil bacteria: Incubation of bacteria in the laboratory ideally should be conducted in a temperature range that is from 5 °C below the in situ groundwater temperature to 10 °C above. On average, the groundwater temperature in the United States is approximately 10 °C. Most bench scale laboratory testing is done at room temperature (about 25 °C).

In conclusion, many of the processes exploited in aid of subsurface remediation are impacted by temperature. The overall compounded physical/chemical/biological effects that can be anticipated within a relatively normal temperature range (5 °C to 35 °C) could on an additive basis make a three- to five-fold increase in remediation rates, which is not insignificant. The issue is then site specific, i.e., shallow versus deep contamination. Passive heating systems using modification of the ground surface to improve heat adsorption or low-intensity intrusive systems such as those associated with heat pump systems could offer remediation actions with possible economic benefits.

GROUNDWATER FLOW IN HETEROGENETIC SEDIMENTS AND FRACTURED ROCK SYSTEMS

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

Groundwater flow in heterogenetic sediments and fractured rock systems is highly complex. Contaminants within these types of subsurface conditions are relatively common. Similarities and differences exist in flow characteristics of heterogenetic sediments and fractured rock systems, which influence contaminant fate and transport and subsequent remediation efforts at environmentally impacted properties.

FLOW IN HETEROGENETIC SEDIMENTS

Heterogeneity in granular aquifers is common and is caused by the large variability in water flow regimes in various depositional environments. The vertical and lateral variability in granular aquifers can occur on a scale of centimeters or less.

Anisotropic advective flow, dispersion, diffusional transport, contaminant adsorption, and other physical/chemical processes taking place in these systems is complicated. Complete evaluation and understanding of these processes at any given site is impractical, and attempting to gather comprehensive data in support of that understanding is prohibitively expensive. Anisotropic advective groundwater flow dominates the heterogenetic sedimentary systems.

Heterogeneity and subsequent anisotropic groundwater flow conditions are considered normal at most contaminated sites. Near-surface granular aquifers that are impacted by the release of contaminants are usually poorly indurated and can be classified according to the depositional regime as follows:

- Fluvial that includes rivers, streams, and alluvial fans
- Glacial such as tills
- Lacustrine lake deposits
- Eolian sand dune systems

In these depositional systems (particularly fluvial and glacial), it is common to have granular deposits of high permeability contrast juxtaposed to each other, and therein lies the source of great complexity.

Granular media advective flow regimes can be classified into three broad types:

- Uniform, medium-to-coarse granular soils capable of supporting significant intergranular advective flow
- Low permeability silts, clays, and tills in which any advective flow actually occurs in secondary permeable channels such as fractures or joints
- Complex interbedded soils with zones of high permeability contrast resulting in anisotropic advective flow

Sediment source and energy of transport are the primary controllers of the characteristics of water-deposited materials. A rain storm, rising tide, or flood waters carry an initial high-energy load of coarser material (that may be mixed with finer grained silts and clays). As a given deposition event begins to subside, the denser and typically larger size grains settle first, followed by granular materials of decreasingly less dense and smaller sizes. The last stage may involve quiescent waters from which the finest clay and silt particles will settle over a significant period of time. The result of a given episode is a layer stratified by size within the overall soil mass.

Scale is important when evaluating water flow characteristics. The smallest scale heterogeneity is expressed by variations in pore size and shape. Slightly larger, bench scale heterogeneity is expressed by variance in the particle size; on the larger scale of a road cut, heterogeneity is seen as the layering of individual bedding planes; and on larger scales, as changes in size of the layers occurs, pinch outs and facies change. In layered sediments of this type, horizontal flow is dominated by the most permeable units in the sequence and vertical flow by the least permeable. Mega scale depositional heterogeneity can also be created by processes such as braided stream sediments or glacial till systems, where lenticular pods of sand may be deposited in a matrix of lower permeability silts and clays.

Material that is two or three orders of magnitude higher in permeability than the bulk soil matrix will totally govern the groundwater flow system. For example, 1 in. of sand can dominate the flow through tens of feet of silt or clay. The practical problem this presents is the determination of the spatial configuration of the physically small but hydraulically dominant units in the soil matrix. Because of this significant problem, continuous soil coring is recommended during subsurface investigations of soils with mixed sands and clay or silts, where practical.

One potentially most damaging situation includes conditions in which the permeable channels or layers are preferentially exposed to the surface. These zones are susceptible to contaminant impact followed by surface recharge, which can act as a hydraulic driver of contaminants within a matrix that overall does not support high rates of advective groundwater flow, even in the more permeable (but unexposed) portions of the coarser soil units.

The testing and analysis of these anisotropic heterogenetic flow systems is potentially extremely complex. As a matter of practicality, a significant number of assumptions must be made in the analysis of these aquifers. As readily available computing power has increased over the past decades, the number and degree of required assumptions have declined in the more sophisticated applications. Irrespective, two prime assumptions are that at some scale the heterogenetic soil can be treated as a homogenous block and that the spatial configuration of the heterogeneity has been defined. Increasing computational power typically allows for a finer mesh of blocks, but it has not addressed the issue of spatial configuration.

A significant new tool that is beginning to be put to use is the employment of fractal concepts to study the hydrology of heterogenetic soils. This approach points to a serious flaw in the value of using finer blocks in the modeling matrix. If the geometry of the heterogeneity is fractal in the soil matrix (which evidence increasingly indicates it often is), the assumption that the matrix can be represented by averaging is false. A fractal matrix will not become homogenous with averaging irrespective of block size. With further use of the concepts of fractal analysis, it is possible to model general features and hydraulic behavior and interpolate hydraulic dynamics from sparse data, an ability of potentially great value in these complex groundwater flow systems. With this approach, water level, flow, and soil data can be used from an operating groundwater recovery system to fine tune understanding of the site. Through an iterative process, the fractal model resolves hydrogeological attractors in the flow system. Known hydraulic properties model drawdown in the forward iteration, and observed drawdown and flow rates model hydraulic properties in the inverse iteration. This process has the potential to be a powerful tool for defining heterogenetic anisotropic groundwater systems.

At this juncture, computer capacity allows for primitive two-dimensional modeling in this manner. With improvements, it will soon be possible to robustly model three-dimensional systems. This technology will spatially identify which portions of the soil matrix have the greatest impact on the groundwater flow regime. It can also subtly determine areas where data gaps of high impact exist and that require additional testing wells.

FLOW IN FRACTURED ROCK SYSTEMS

Usually shallow water-bearing zones underlying a site are in shallow unconsolidated sediments. However, in some cases where bedrock outcrops or lies just meters below the surface, fractured rock hydrogeologic systems occur. In the latter case, any significant groundwater flow occurs in fracture systems within the bedrock. If soils are fine grained (i.e., tills or clays), fractures may also play a dominant role in advective flow.

The study of fluid flow through porous media was first established by Henri Darcy in 1857. Much of the science of hydrogeology was developed with and was designed for use in granular porous media because of the incentive to find and produce aquifers for large-scale groundwater consumption. It took almost 100 years before fractured flow was studied in great detail. The study of fluid flow through fractured rock was first developed by the petroleum industry in the 1950s (1). These petroleum studies resulted from observations that oil and gas production could be significantly increased by fracturing the oil-producing formations near the well bore (2). Fractured media in most instances will not produce groundwater on the same scale as homogeneous granular aquifers. Fracture flow systems are also more complex to analyze and hydraulically respond differently than do those in porous granular matrices. A good overview of contaminant transport in fractured media is presented in Schmelling and Ross (3).

To compare groundwater flow in fractured rock systems with granular media, under identical hydraulic gradients, one square meter of granular material with a hydraulic conductivity of 8.1×10^{-2} cm/s has the equivalent water-conducting capacity of one fracture in one square meter of rock with an aperture of 1 mm. In granular media, grain size, shape, and degree of sorting are the prime microscale parameters that determine hydraulic character. Fracture density, orientation, aperture, and type of rock matrix are the major parameters affecting groundwater flow in fractured media. The typical range for fracture aperture is from 0.2 to 25 mm, and fracture spacing is from 2 mm to 3 m.

Individual fractures are not infinite in extent. Therefore, where flow is supported, fracture density must be high enough to ensure connectivity through the system. The fracture density required to sustain advective flow is termed the “percolation threshold.” Below that threshold, fractures may be connected, but only in small localized regions. Above, localized regions become interconnected, and flow over significant distances can take place. With an increase in fracture density, the system becomes increasingly previous.

An expression of the parameter that determines the percolation through fracture is $N(L_F)^2$, where N is fracture density and L_F is equal to fracture length times pi/2. The percolation threshold has been found to fall around 0.3.

As an example, in an area underlain by metamorphic rocks with discrete water-producing fracture systems:

- Fractures were approximately 0.5 m in length
- Rocks with a fracture density of 50 to 200 per 0.5 m

This process gave values of NL_F^2 of 30 to 120, well above the percolation threshold of 0.3.

Given adequate connectivity, the flow an individual fracture can support is proportional to the cube of the fracture aperture. This cube rule means that a few fractures with preferentially higher apertures can dominate the flow system, and those are the ones most important to delineate. Typically, fracture aperture will decrease with depth. Usually the highest flow rates occur in the upper 9 m of a fracture system, with flow decreasing to near zero below depths of 30 m.

To some degree, all rocks or soils are reactive. In most igneous, metamorphic, or fine-grained sedimentary rocks, geochemical reactions tend to seal a given fracture over time. In carbonates, the opposite can be true, with the fracture aperture increasing with time (and the potential flow rate increasing with the cube of that aperture). The hydrodynamics of fracture flow systems will approach those in porous media in systems where fractures are randomly oriented and density is high. These systems can be analyzed with conventional granular media methodologies. To use those methods on other fracture systems is an error with potentially significant consequences.

Characterization of a fracture flow system is potentially an expensive process. Ideally, data should be gathered on fracture length, orientation, aperture, and density. Additionally, information on hydraulic head, the porosity and hydraulic characteristics of the bulk matrix, the type of contaminants, and the potential interactions between contaminants and the matrix are also important. Hydrogeologic characterization of fracture systems can be accomplished through coring, complex pumping tests, tracer tests, geophysical evaluation, or bore hole flow meters. In addition, evaluation of the hydrodynamics of a fracture system that has been defined with multiple orientations is mathematically extremely complex.

However, some level of useful knowledge can be inferred based on the structural setting of a site. The tectonic and depositional history of a given site is generally available in the geologic literature. The removal of overburden introduces stresses caused by reduction of overburden pressure, uplift of the region, and thermal stresses caused by cooling; the net stress is extensional.

Fractures that form in the tensile stress field can be placed into two classes:

- Unloading fractures, which include vertical fractures, indicating response to tensile stress in the horizontal plane and fractures horizontal or parallel to the topographic surface
- Release fractures that are fabric controlled in their orientation.

If a site has undergone even a mild degree of tectonic deformation, structural analysis can be a powerful predictive tool for the orientation of the dominant fracture sets at a site. Several points are important to use this concept:

- The deformational pattern observed in the large scale is replicated at medium, small, and microscopic scales. For example, the NE trend of the Appalachians is generally reflected as an NE-oriented fabric at all scales.
- Deformation imparts a fabric to the impacted rocks. This fabric imparts anisotropicity, which in turn will control subsequent fracture generation. For example, fractures will tend to propagate perpendicular to a strong linear fabric element.
- The stronger the degree of imprinted fabric, the greater the density of fractures in the controlled orientations.
- Depositional features such as bedding planes also impart anisotropic fabric.

Important differences exist between fractured and porous media with regard to contaminant fate and transport. Rapid transport in preferred directions can occur through rocks that normally would be thought impervious, such as bedrock fractured flow systems. Because of the general lack of organics in fractured rock systems, the flow of organic chemicals traveling through fractured rock systems is generally not retarded as much as in organic-rich granular sediments. Retardation is a function of matrix and contaminant chemistry and surface area. As an estimate, a block of granular media, such as well-sorted sand or gravel, will have a surface area 1000 to 100,000 times greater than a similar block of fractured media, such as fractured limestones or granites. Contaminant retardation will be roughly proportional. Contaminants in fractures can be mobile over long distances or until they are transported into a granular media.

Fractured rock systems create interesting challenges as well as opportunities. When drilling contaminant recovery wells in fractured terrain, it is suggested to orient the well bore such that it is perpendicular to the major water-bearing fracture set. The resulting well may not necessarily be vertical, but it will maximize the potential to intersect flow-bearing fractures, which allows for more complete contaminant removal.

BIBLIOGRAPHY

1. Gale, J.E. (1982). *Assessing the Permeability Characteristics of Fractured Rock*. Geological Society of America, Special Paper 189, pp. 163–181.
2. Duguid, J.O. and Lee, P.C.Y. (1977). Flow in fractured porous media. *Water Resources Res.* **13**: 558–566.
3. Schmelling, S.G. and Ross, R.R. (1989). *Contaminant Transport in Fractured Media: Models for Decision Makers*. United States Environmental Protection Agency Superfund Ground Water Issue, EPA/504/4-89/004, Washington, DC.

HORIZONTAL WELLS

Milovan Beljin
Cincinnati, Ohio

Horizontal wells are a new technology for solving problems in the environmental industry. Due to interest in horizontal wells for oil production, a large number of technical papers have been published regarding the reservoir engineering aspects of horizontal drilling and reservoir simulation. In the groundwater supply industry, the first theoretical analysis of groundwater flow to horizontal drains (collector wells) can be traced back to the early 1960s. In recent years, however, there has been renewed interest in horizontal wells for subsurface remediation. Horizontal wells offer significant advantages over vertical wells in environmental remediation and protection in many hydrogeologic scenarios. Horizontal well-screen orientation complements common site logistics, typical aquifer geometry, and groundwater flow patterns. A single horizontal well can replace many closely spaced vertical wells.

The logistical advantages of horizontal wells are obvious. Horizontal wells avoid the need for installing wellheads inside buildings or in the midst of complex manufacturing or process facilities. Landfills, spoil mounds, and landfill liners need not be penetrated to extract leachate or other underlying contaminants. In addition, contaminant concentrations are often highest directly beneath buildings, landfills, and other obstacles to remedial operations, so that treatment facilities are constructed tens or hundreds of feet away from the target zone of remediation.

Soils are naturally stratified and individual aquifers or water-bearing zones are much wider than they are thick. Despite the dominance of the horizontal direction in aquifer shapes and groundwater flow, the dominant tool for extracting contamination from subsurface sources is a vertical well. However, in many environmental remediation scenarios, a horizontal well offers a better match of form and function than a vertical well. The tabular geometry of many aquifer zones renders horizontal wells more productive than vertical wells. The flow characteristics of many aquifers create elongate contaminant plumes, and extracting contaminated groundwater is often more efficient using horizontal wells. A horizontal well placed through the core of a plume can recover higher concentrations of contaminants at a given flow rate than a vertical well.

Horizontal wells also offer many advantageous over vertical wells in fractured aquifers. Fractures in an aquifer are commonly vertical. Because fluid or vapor recovery from fractured zones requires penetration of numerous fractures, a horizontal well oriented normally to vertical fractures is the optimal tool for pump-and-treat or soil vapor extraction systems in vertically fractured zones. By analogy, vertical wells are efficient in highly stratified soils that have little vertical communication between strata, where fluid or vapor recovery from many thin layers through a single wellbore is required.

Injection of groundwater is part of some remediation systems, either to create a water table mound or to reinject treated effluent from a manufacturing plant into the subsurface. Water table mounds can help control flow of contaminants toward recovery wells or trenches, or they can serve as hydraulic barriers. Manufacturing plants can avoid high sewer discharge costs if their treated plant effluent can be reinjected into a nondrinking water aquifer. Reinjection can cause mounding, but the mounding can be minimized by using horizontal wells. Hydraulic barriers are most efficiently created using horizontal wells oriented perpendicularly to the groundwater flow direction.

Various drilling technologies are capable of installing horizontal wells for subsurface pollution control. Such wells are typically installed in unconsolidated soils, 10 to 200 feet deep. Selection of drilling technique depends on surface access, well placement and completion requirements, and subsurface hydrogeology. During drilling of a well, drilling mud can invade the aquifer and change its permeability in the vicinity of the well and cause formation damage or a “skin effect.” The thickness of the skin zone depends on drilling technology and also on the permeability of the aquifer. The additional drawdown due to the change in permeability and the turbulent flow around the well is called “well losses.” Because of lower flow rate per unit screen length, horizontal wells show smaller well losses due to drilling mud invasion than vertical wells.

The current high installation cost of a horizontal well compared to that of a vertical well is offset by operating and maintenance cost savings. New developments in horizontal drilling technology will further reduce the cost of installing horizontal wells, and subsurface pollution control using horizontal wells should become as common as using vertical wells.

HORIZONTAL WELLS IN GROUNDWATER REMEDIATION

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

Horizontal wells have been used in an increasing number of remediation projects. In addition to chemical and biological reactive barriers, horizontal borings can be used for groundwater extraction and control, air sparging, bioremediation, groundwater injection, for vadose zone soil vapor extraction or bioventing systems, and free product recovery. Although, in areas where there is a great deal of fluctuation in the groundwater table, use for free product recovery can be problematic.

Horizontal well technology was first used in the late 1920s by the petroleum industry to increase oil reservoir thickness per well and greatly increase oil production. In the 1970s, the technology was applied by utility companies to cross rivers and other natural or manmade barriers such as highways. The use of horizontal drilling technology for environmental applications began in the late 1980s and has escalated ever since.

Horizontal wells offer distinct advantages over trenching or vertical wells. Trenching produces massive quantities of excavation spoils and is economically limited in depth of application. In addition, horizontal wells can be installed beneath structures and other surface obstructions that would be impossible to access using trenches or vertical wells.

Horizontal wells have a unique advantage that originates in the geometry of the typical contaminated groundwater system. Horizontal permeability is on the order of ten times greater than vertical, because of the stratigraphic layering of near-surface soil horizons. The effect of increased horizontal permeability enhances the spread of contaminant plumes horizontally. Horizontal wells can be installed through contaminated zones, along their leading edges, or along a property line. All active groundwater remediation systems rely on the mass transport of water, air, or other chemicals. Mass transport is induced through the interface offered by the screened sections of a well that are exposed to the contaminated strata. Compared with vertical wells, horizontal wells can increase well screen interface by an order of magnitude or more.

Various drilling heads and cutting removal technologies have been applied to horizontal drilling. Fluid-cutting systems, mechanical cutting, augers, percussion drilling, and sonic methods are all available. The selection of which is dependent on the geology and competency of the subsurface (and the selected technology vendor). With appropriate techniques, bore holes may be extended through difficult conditions such as gravel formations, coral reefs, and even boulders or bedrock.

The removal of cuttings is usually based on mud slurry systems for installation beneath the groundwater table and air systems for installation in the vadose zone. The formulation and engineering of a mud slurry system takes great care; historically this has been one of the most common points for the failure of an installation. The mixture must be capable of holding the hole open, removing the cuttings without erosion of the walls of the well bore, and it must be capable of decomposing with time to restore the well bore to permeable operation.

Inherent with the horizontal drilling process is the ability to accurately direct the placement of the horizontal well bore. Two dominant methods of achieving directional control of the drill head are available: magnetometer/accelerometer and radio beacon. A magnetometer/accelerometer array is expensive, and the long configuration of an array makes it prohibitive for short radius bends. It is also subject to magnetic interference, which can be a serious issue around tanks or other surface or near-surface structures (i.e., buildings, piers and piles, pipelines, or utility runs). Radio beacons previously have been limited to a depth of 25 feet or less, although there have been instances where down hole wirelines have increased accessible depth. Well casings for horizontal wells must be more highly engineered than for vertical wells. They must have great tensile strength to withstand the significant forces associated with installation. Compressive strength is required to resist the overburden load as the horizontal well bore collapses with time. The screen pack is also a critical issue with regard to installation and subsequent operational efficiency. Conventional sand packs can be installed, but they require great care. Many vendors now use prepacked screens. These are nested screens, an exterior screen (typically HDPE), a packing system (sand, filter cloth, or other filter media or mixtures of media), and an internal screen typically constructed of stainless steel. These prepacked screens are stiffer and subject to larger turn radii, but they obviate the difficulties associated with installing a conventional sand pack in a horizontal boring.

Horizontal drilling is a relatively costly process to use. It is driven by necessity, or the economic advantages of scale. Necessity drivers are generally the need for the installation of a remediation system underneath a structure or facility that cannot be disturbed. Economics of scale come in to play with regard to large contaminant plumes. A single horizontal boring can replace 10 to 30 vertical wells in a plume of large aerial extent. The break-even point for the choice between horizontal or vertical wells occurs around the need for four to five vertical wells. Installations requiring more than five vertical wells along a linear trend can often be more cost effectively addressed with horizontal wells.

Two basic methods of horizontal well installation are available. The simplest involves two ends, boring downward to near the desired depth at one end, curving the boring to the horizontal, traversing the required path and distance horizontally, and then recurving to the surface at the other end. The casing is then pulled back through the boring from the distal end. A more sophisticated approach is to use a single end, through which the boring is advanced, followed by the casing being pushed through the boring, a more difficult, risky, and time-consuming process. Horizontal drilling using the two-end approach can cost as little as \$30 to \$40 per linear foot; the cost for single-ended systems start at \$80 to \$90 per linear foot.

As with many other technologies applied to the environmental industry, developers over the last decade have focused on the creation of a "just good enough" technology. Placement accuracy is sufficient, boring sizes are minimal, and construction materials are designed for the limited life of the project. Successful installation of horizontal wells depend on the selection of a vendor who is flexible enough to use an approach that is appropriate for the subsurface condition of your site; who is experienced in the use of the specialized drilling equipment and screens; who can engineer the application as well as the installation procedures; and who is experienced enough to be quick. Aside from failure caused by inadequate mud engineering, taking too long is another key source of failure; the well bores simply cave in before installation is complete.

Horizontal wells have been adapted for use in many soil and groundwater projects. The use includes groundwater extraction, air sparging, free product recovery, in situ bioremediation and bioenhancement, soil vapor extraction, in situ soil flushing, in situ radio-frequency heating, treatment walls, hydraulic and pneumatic fracturing, and leachate containment and collection (1). The overall performance of horizontal wells used for air sparging and vacuum extraction at the Savannah River Site in South Carolina (2) shows the following:

- A five-time increase in chlorinated solvent removal as compared with conventional vertical wells.
- Eight tons of chlorinated solvents were removed over a 20-week period. The equivalent would have been 11 vertical wells at a pump and treat system, each extracting at a rate of 500 gallons per minute;

The Savannah River Site project managers estimate that a 40% cost savings was predicted when compared with the use of conventional pump and treat technologies.

In the six years between 1987 and 1993, over 100 horizontal wells were drilled in the United States as part of environmental remediation projects. One quarter of them were used for groundwater extraction, one quarter of them were used for soil vapor extraction, and one half of them were used for other technologies (air injection, bioventing, and free product recovery). Eighty percent of the horizontal wells were installed at depths of 25 feet or less (2).

The advantages of horizontal wells are well known: Horizontal well screens contact a larger surface area of contaminated aquifer than do conventional vertical wells; the cost of horizontal wells, although more than vertical wells, is less expensive when comparing the fewer number of wells required for a particular remediation project; and finally because horizontal transmissivity generally exceeds vertical transmissivity in most aquifers, horizontal wells can deliver and recover more fluids, gases, and groundwater than can vertical wells (1). With these advantages and the improvements in directional drilling and global positioning systems (GPS), it is likely that horizontal wells will become more common in future remediation projects.

BIBLIOGRAPHY

1. Miller, R.R. (1996) *Horizontal Wells, Ground-Water Remediation Technologies Analysis Center (GWRTAC)*. Technology Overview Report TO-96-02.
2. Hazardous Waste Remedial Actions Program (HAZWRAP). (1995). *In situ Bioremediation Using Horizontal Wells*. Innovative Technology Summary Report, prepared for U.S. Department of Energy. Available: http://www.gnet.org/gnet/tech/reports/sbu.htm.

HEAD

TODD RASMUSSEN
The University of Georgia
Athens, Georgia

Water moves from zones of higher energy to zones of lower energy. For example, gravity causes water to move from higher to lower elevations. Water also moves from higher to lower pressures. The total energy is the sum of the all forces acting on water. This total energy is referred to as the *head*, which can determine the direction and rate of water movement.

Bernoulli's equation commonly combines the dominant forces that cause water to move:

$$H = z + \frac{P}{\gamma} + \frac{v^2}{2g} \tag{1}$$

where H is the total head, z is the elevation of the point where the head is measured, P is the fluid pressure at the point of measurement, $\gamma = \rho g$ is the fluid specific weight, ρ is the fluid density, g is the gravitational constant, and v is the fluid velocity.

The velocity component can be omitted if the flow is slow, $v \approx 0$, which is generally true in groundwater systems. When the velocity is neglected, the total head is

$$H = z + p \tag{2}$$

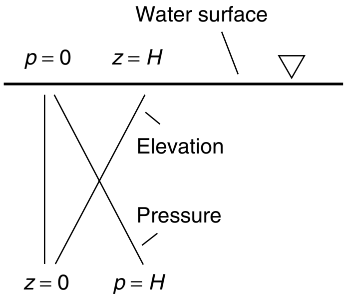

Figure 1. Total head is the sum of the elevation and pressure heads. For example, the elevation decreases and the pressure increases when moving downward from the water surface in a column of still water.

where $p = P/\gamma$ is the fluid pressure head. Figure 1 shows how the pressure head and elevation head compensate in a column of still water.

The pressure term is omitted when the water surface elevation is measured, because the fluid pressure is zero, $p = 0$, at this point. This simplification yields the equation, $H = z$, where z is the elevation of the water surface.

Important assumptions required for the use of water levels to determine the head include (1) hydrostatic conditions exist with no vertical water movement; (2) the water velocity within the aquifer is sufficiently small; (3) water within the monitoring borehole or piezometer is pure water at a standard temperature and density; and (4) the air pressure on the water surface equals the mean barometric pressure.

CAPILLARY AND OSMOTIC FORCES

Capillary and osmotic forces also affect the total head. Failure to account for these forces may result in incorrect predictions of water flow and transport.

Capillary forces develop because of the tendency of soil surfaces to attract water. Water held on soil surfaces resists the downward force of gravity, and it does not readily drain from the soil. The head must consider the negative fluid pressure that develops because of capillary forces. A measure of the force by which water is held is the *matric tension*, which can be measured with *tensiometers*.

Water may move upward above the regional water table because of capillary forces. The height of the saturated zone formed above the water table (i.e., the capillary fringe) is largely determined by the magnitude of capillary forces, which is a function of the pore surface area. Finer grained media have greater capillary forces, which result in higher capillary fringes.

Capillary forces generally increase with decreasing pore size, as shown in Fig. 2. The *capillary rise equation* relates the pore size to the height of rise:

$$\psi = \frac{2\sigma \cos \alpha}{\gamma r} \tag{3}$$

where ψ is the capillary height of rise, σ is the surface tension of water, α is the solid−liquid contact angle, and r is the pore radius (1).

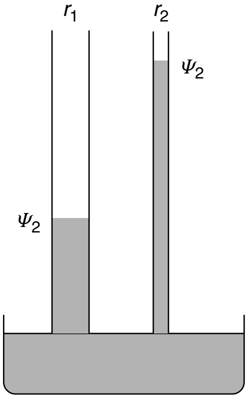

Figure 2. Height of rise of water in capillary tubes. Note that the height of rise increases as the radius of the tube decreases. Equivalently, small pores hold water to greater tensions than do large pores.

An additional force that induces fluid movement is a change in solute concentration. The *osmotic potential* decreases the head of water, which causes water to flow from areas of low solute concentrations to areas where solute concentrations are higher. In arid areas with high soil solute concentrations, salts may become concentrated as water evaporates at the soil surface, which causes an increase in the salt concentration. This increase in salt concentration induces an additional force that causes water to move upward from less saline groundwater to the surface.

Thus, evaporation can increase the height of the capillary fringe by augmenting the capillary force with osmotic forces. The osmotic height of rise in dilute solutions is

$$\phi = \frac{kTC}{\gamma} \tag{4}$$

where ϕ is the osmotic height of rise, k is the Boltzmann constant, T is the absolute temperature, and C is the solute concentration (2).

FLUID DENSITY

Water levels can be affected by the density of the fluid within the borehole. For example, water levels in a well monitoring a deep aquifer may be affected by the temperature of the water in the borehole.

Warmer water is less dense than is colder water (except below 4 °C) and so a column of warm water will display a higher water level than will a column of colder water for the same pressure at the bottom of the borehole. Because of the *geothermal gradient* (the tendency of temperature to increase with depth below the surface), deeper water within the borehole is warmer and slightly less dense than is water near the surface.

Other factors besides temperature affect the water density in a well. The dissolved solids concentration (salinity) causes the water density to increase. Suspended sediments also increase the weight of the fluid, whereas air bubbles rising to the surface lower the density of the water column. Small changes in density over a long water column cause an appreciable difference in observed water levels.

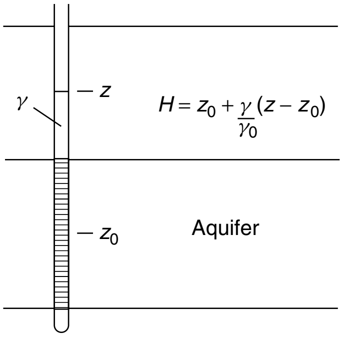

Figure 3. Total head, H, adjustment for conditions when the borehole specific weight, γ, is different from the freshwater weight, γ_o. The effect is greater for larger differences in elevation between the water level, z, and the screened zone, z_o.

To account for the variation in fluid density, observed water levels should be adjusted with a standard water density, called the *freshwater density*. The corrected head is called the *freshwater head*, which accounts for the fluid density within the borehole:

$$H = z_o + \frac{\gamma}{\gamma_o}(z - z_o) \tag{5}$$

where H is the freshwater head, z_o is the elevation of the screened interval, z is the observed water level elevation in the well, $\gamma = \rho g$ is the average fluid specific weight in the well, and $\gamma_o = \rho_o g$ is the standard freshwater specific weight. Figure 3 illustrates the geometry of this problem.

The freshwater head equation can also be written as

$$H = z_o + \varrho\,\Delta z \tag{6}$$

where $\varrho = \gamma/\gamma_o$ is the specific gravity of the fluid within the well and $\Delta z = z - z_o$ is the height of the column of water within the monitoring well above the screened zone.

These equations show that the freshwater head correction is larger for longer water columns and for water with a density that is substantially different from the freshwater density.

SPATIAL AND TEMPORAL VARIATION

Head is often observed to vary over both space and time. These variations are particularly important when mapping the regional potentiometric surface, because long-term averages may not accurately reflect the surface dynamics over time.

Temporal variations result from barometric (atmospheric pressure) influences, tidal effects, fluid density (sediment, salinity) changes, and vertical flow (nonhydrostatic conditions) within the water column (2). Spatial variations are a function of aquifer properties as well as of the regional hydrogeologic flow environment.

Heads within a specific hydrogeologic unit generally vary smoothly over space, unless some kind of intervening boundary causes a jump in the head within the formation. For example, water levels may decline smoothly, only to change abruptly when a fault displaces one side of an aquifer relative to the other, which causes a sharp change in the head across the boundary.

Also, heads in different hydrogeologic units may be different because the recharge and discharge patterns for each aquifer are different. A unit that outcrops in one area may have a different head than one that outcrops at a higher elevation. Wells that tap different aquifers are commonly observed by scientists to have different water levels, even when the wells are situated next to each other. This variation of head by aquifer is a result of the regional hydrogeologic setting, which must be considered when trying to evaluate water head data.

Spatial variations are plotted on maps with *equipotentials*, which are lines of constant head within a hydrogeologic unit. Each aquifer generally has a unique set of equipotentials that can also vary over time. In *isotropic* media (i.e., aquifers with no preferential flow direction), the direction of fluid flow is shown by placing lines, called *streamlines*, on the map that are perpendicular to the equipotentials.

Water levels in wells can also vary over time. Water levels in wells are commonly affected by barometric pressure—they fall as barometric pressure rises, and they rise as the pressure falls—because the total head in the aquifer is the sum of the water level elevation plus the atmospheric pressure on the water surface in the borehole.

Although the influence of barometric pressure is commonly neglected when monitoring water levels in wells, large changes in barometric pressure (such as when large storms pass overhead) can occasionally cause large errors in the determination of the total head.

To correct for the effects of barometric pressure, observed water levels, W, can be adjusted by the variation in pressure about the mean, $\Delta B = B - \overline{B}$:

$$H = W + \Delta B \tag{7}$$

The mean pressure, $\overline{B}$, can be taken as the average global sea-level barometric pressure, $\overline{B} = 1013.25\ hPa \approx 33.9$ ft, or, alternatively, equal to the local average barometric pressure, which varies with elevation and local weather conditions. Air pressure variations can be neglected by sealing the well, or by measuring absolute pressure instead of the gauge pressure within the monitored interval.

Barometric pressure changes cause many wells to fluctuate over short time periods (e.g., from day to day), whereas precipitation, evapotranspiration, and pumping patterns often cause longer term variations. Climatic variation over time also can have a large influence on observed head. In trying to establish a long-term average head for a well, one must consider all possible sources of short- and long-term variation.

BIBLIOGRAPHY

1. Hillel, D. (1971). *Soil and Water: Physical Principles and Processes*. Academic Press, New York.
2. Spane, F.A., Jr. and Mercer, R.B. (1985). HEADCO: A program for converting observed water levels and pressure measurements to formation pressure and standard hydraulic head. Pacific Northwest National Laboratory, Rockwell Hanford Operations, Richland, WA, Report PNL-10835.

WELL HYDRAULICS AND AQUIFER TESTS

John E. Moore
USGS (Retired)
Denver, Colorado

The objectives of an aquifer test are to identify the performance of a well (to estimate the yield capability) and to estimate aquifer properties. Accurate estimates of the hydraulic characteristics of aquifers depend on reliable aquifer test data. The tests are done on existing wells or a well drilled specifically for that purpose. An aquifer test is a controlled, field site experiment to determine hydraulic conductivity and aquifer storage. The test consists of measuring groundwater discharge and observing water level changes in the pumped well.

These are the hydrologic and geologic conditions needed for a successful aquifer test:

- Hydrogeologic conditions should not change over short distances.
- No discharging well or stream nearby.
- Discharge water should not return to the aquifer.
- The pumped well should be completed to the bottom of the aquifer and should be screened or perforated through the entire thickness of the aquifer.
- Observation wells (at least three) should be screened at the middle point in the aquifer. One observation should be located outside the area of influence of the pumping well drawdown.
- Location of observation wells should be based on the aquifer character.
- Determination of prepumping water-level trend.

The following conditions and field measurements are needed for an aquifer test:

- Accurate water-level measurements during pumping and recovery
- Pumped well developed prior to test (several hours of pumping and surging)
- Dependable power source to provide a constant pumping rate
- A flow meter that can read instantaneous and cumulative discharge
- Electrical conductance, Eh, pH, DO, and temperature measurements
- Water levels measured several hours before the test begins
- Pumping rate maintained at 5% tolerance. An optimal rate is 50% of maximum yield
- Water level is measured with an electric sounder or pressure transducer
- Remove the discharge water from the site

- Observation wells should be tested by injecting a known volume of water and measuring the recovery response
- Establish baseline trends of regional water level changes and barometric pressure changes
- Pumping well lithology and construction data

TYPES OF TESTS

Specific Capacity Test

The amount of water that a well will yield can be determined by a specific-capacity test, in which the pumping rate and water-level changes are monitored for a set period of time. The first step is to measure the initial water level in the well. Commonly, a well is pumped at several successively increasing rates for uniform periods (typically 1 hour) to establish a rate that can be maintained for long-term pumping. The well is then pumped at a steady rate and the water-level changes are monitored at the pumped well. Water levels should also be monitored in at least one observation well 2 to 20 meters (6 to 65 feet) from the pumped well. The water level will decline quickly at first, as water is removed from the well, then more slowly as the rate of flow into the well approaches the pumping rate. The ratio of the discharge rate (Q) to water-level change (drawdown, dd) gives the well's specific capacity, or $\mathrm{Sc} = Q/\mathrm{dd}$. For example, if the discharge rate is 6 liters per second (L/s) (100 gallons per minute) and the drawdown is 3 meters (10 feet), the specific capacity of the well is 2 L/s per meter (10 gpm/ft) of drawdown. Once the specific capacity and the available amount of drawdown are known, the yield of the well can be determined from the formula $Q = \mathrm{Sc} \times \mathrm{dd}$. The pump should be deep enough that the water level does not go below the pump intake. The pump depth should also be sufficient to allow drawdown caused by pumping and natural declines in water level during periods of drought

Step-Drawdown Test

The step-drawdown test evaluates the performance of a well. Well performance can be affected by resistance to flow in the aquifer itself; partial penetration of the well screen, incomplete removal of mud from the gravel envelope, or invasion of fines into the envelope; and blockage of part of the screen area. The well should be developed prior to the test using a surge block and/or pumping until the well discharge is clear In this test, the well is pumped at several (three or more) successively higher pumping rates, and the drawdown for each rate is recorded. The test is usually conducted for 1 day. The discharge is kept constant through each step. The test measures the change in specific capacity. The data provide a basis to choose the pump size and discharge rate for the aquifer test and for long-term production.

Slug Test

In this test, a small volume of water is removed from a well, or a small volume is added and the recovery of the water level is measured. The aquifer transmissivity can be determined from the time–drawdown or recovery data. The disadvantages of the test are that a data logger is needed to measure water-level changes and water removed represents only a small volume of the aquifer. Slug test data are evaluated by the Bower and Rice (1) method for unconfined aquifers and the Cooper, Bredehoeft, and Papadopulos (2) method for confined conditions. The advantages of these tests, compared to those of full aquifer tests with observation wells, are reduced cost and time. The disadvantage is that a storage coefficient is not determined and only a small volume of the aquifer is sampled. Many factors contribute to error in slug tests as follows: entrapped air, partial penetration, leaky joints, and the radius of influence of the test.

ANALYSIS OF AQUIFER TEST DATA USING THE THEIS EQUATION

The Theis equation (3) is used to determine the hydraulic characteristics of an aquifer. In this test, a well is pumped, and the rate of decline of the water level in nearby observation wells (two or more) is noted. The time drawdown is then interpreted to yield the aquifer parameters. In 1935, C.V. Theis (3) developed the first equation to include pumping time as a factor (4). The following are assumed:

1. The pumping well is screened only in the aquifer being tested.
2. The transmissibility of the aquifer is constant during the test to the limits of the cone of depression.
3. The discharging well penetrates the entire thickness of the aquifer, and its diameter is small compared to the pumping rate.

These assumptions are most nearly met by confined aquifers at sites far from the aquifer boundaries. However, if certain precautions are observed, this equation can also be used to analyze tests of unconfined aquifers (5).

BIBLIOGRAPHY

1. Bower, H. and Rice, R.C. (1976). A slug test for determining hydraulic conductivity of unconfined aquifers with completely or partially penetrating wells. *Water Resources Research* **12**: 423–438.
2. Cooper, H.H. et al. (1996). Response of a finite-diameter well to an instantaneous charge of water. *Water Resources Research* **3**: 263–269.
3. Theis, C.V. (1935). The significance of the cone of depression in groundwater bodies. *Economic Geology* **33**(8): 880–902.
4. Heath, R.C. (1989). *Basic ground-water hydrology*. USGS Water Supply Paper 2220, p. 84.
5. Lohman, S.W. (1972). *Ground-water hydraulics*. USGS Professional Paper 708, p. 70.

HYDRAULIC PROPERTIES CHARACTERIZATION

WALTER W. LOO
Environmental & Technology Services
Oakland, California

Groundwater is becoming an important issue as a source of fresh water. The groundwater that exists at the subsurface is called an aquifer. An aquifer is a geologic unit that can store and transmit water at rates fast enough to supply reasonable amounts to wells. Thus, understanding a groundwater body is an essential factor in groundwater development and protection.

The geohydrologic properties of a groundwater body consist of physical, chemical, and biological parameters of the solid matrix and liquid within its pore space. These parameters determine the quantitative, qualitative, and interpretive aspects of the groundwater body. The solid matrix may consist of various sedimentary deposits or rock types. The liquid within the pore space of the groundwater body consists of water, dissolved minerals and gases, microorganisms, and colloidal material.

INTRODUCTION

Many groundwater professionals define a groundwater body by its flow property, such as aquifer, aquitard, and aquiclude (1), which is only good in defining the rate of water yield from a groundwater body. When dealing with water quality and environmental issues, we need better definition of the other properties and characteristics, not just dissolved chemicals, hydraulic conductivity, and effective porosity.

The following parameters characterize groundwater body properties:

Geophysical parameters: electrical conductivity, hydraulic conductivity, storativity, effective porosity, temperature, media thickness, pressure or hydrostatic level, adsorption, absorption, radioactivity, electrical charge, electrokinetic flow, porous or fracture media, recharge, discharge

Biochemical parameters: pH, bacteria counts, inorganic chemicals, organic chemicals, oxidation/reduction, ion exchange potential

Most of these parameters are not required for site characterization by natural resources agencies and the U.S. EPA (2,3). It is very clear that most groundwater bodies are poorly characterized because there is no incentive or pressure to understand these parameters, as long as the groundwater is fit for use. Most textbooks define only some of the parameters but without standardized procedures. The following sections provide typical definitions of the parameters, analytical methods, use, interpretation, limitations, and the necessity for future improvement in characterization.

PARAMETER DEFINITIONS

The definitions of the following parameters are provided in Driscoll (4), Freeze and Cherry (1), Todd (5), Lohman (6), and Batu (7), Ferris et al. (8) Loo et al. (9) unless otherwise noted.

Geophysical Parameters

Electrical conductivity is the ability of a medium to conduct electricity measured in μmho/cm. Resistivity is the inverse of electrical conductivity measured in ohm-meters.

Hydraulic conductivity is the measurement of the ease of water flowing through a medium or water permeability. It is often expressed as cm/s or gpd/ft^2.

Storativity of a saturated aquifer is the volume of water release from storage per unit surface area per unit decline of hydraulic head.

Effective porosity or flow porosity is the interconnected pore space available for water to flow through (10).

Media thickness is defined as the aquifer thickness or productive water-bearing zone.

Potentiometric surface is expressed as the pressure in the water head in a confined aquifer. It is called the groundwater table when the aquifer is unconfined.

Adsorption/Absorption is the process by which molecules of dissolved chemicals in a fluid attach to solid surfaces (11–13).

Radioactivity is the radiation emitted by natural elements or radioactive wastes. Radon and tritium are commonly occurring radioisotopes of uranium or other radioactive elements (14).

Electrical charge is measured as the electrical potential difference (in millivolts) between layers of different materials. Fine-grain material such as clay and silt are slightly positively charged whereas sand and gravel are slightly negatively charged.

An **electrokinetic gradient** occurs when a flow of water is created by dc electricity flow from an anode to a cathode (1,15). It is common to create several feet of water head at the cathode when the applied direct current potential difference is 50 volts over a short distance of 40 to 50 feet between electrodes.

Recharge/Discharge Boundaries can be detected during a pump test when the drawdown curve becomes flat, recharge condition and change (steepened) slope, negative or discharge condition (16). Pumping near a surface stream can often detect a recharge condition. Pumping near a fault or lithologic discontinuity can often detect a negative boundary condition.

Biochemical Parameters

pH is an indication of the acidity or alkalinity of soil or water.

Bacteria count is the enumeration of the population of indigenous bacteria in soil or water which often is expressed in colony forming units, CFU (17).

Inorganic Chemicals are any chemicals not based on carbon such as any salts of major cations and anions and trace metals.

Organic Chemicals are carbon based compounds and can be naturally occurring or occurring as spilled chemicals.

Oxidation/Reduction Potential redox potential in short as is the state of electrochemical reaction which is measured in millivolts.

Ion exchange potential is related to the electrical charge measured in millivolts of colloidal particles which carry relatively large surface area. Colloidal particles have diameters in the range of 10^{-3} to 10^{-6} millimeters. Clayey material has very high cation exchange potential.

FIELD TESTING AND ANALYSIS OF PARAMETERS

Geophysical Parameters

Many of the following physical parameters can be defined by field pump tests:

electrical conductivity
hydraulic conductivity
storativity
effective porosity or flow porosity
temperature
radioactivity
media thickness
pressure or hydrostatic level or potentiometric surface
porous or fracture media
recharge/discharge boundaries

Pumping Test. Before a pumping test is conducted, geological and hydrological information such as the geological characteristics of the subsurface, the type of aquifer and confining bed; the thickness and lateral extent of the aquifer and confining beds, including boundary conditions, is obtained, preferably by surface resistivity and electric logs. Also, data on the groundwater flow system, including the hydraulic gradient and regional groundwater flow and existing wells in the area, should be collected. Then, the site for the well is selected considering the area representative of hydrological conditions, not near railroads and motorways, not in the vicinity of an existing discharge well, or a low water level gradient. After the well site has been chosen, the drilling operation can begin. The pumping well should be drilled to the bottom of the aquifer. Then the pump size and pump type are planned. Besides the well diameter and well depth, the location of the well screen is to be determined. The length of the well screen will largely be decided by the depth at which coarse materials are found. A general rule is to screen the well 100% of the aquifer thickness or full penetration. Partially penetrated observation wells can be used if the distance of the observation is more than 2.5 times the thickness of the aquifer tested, which is true for isotropic and anisotropic aquifer pump tests.

With regard to a pumping rate, too low or too high a pumping rate is not desirable. Too high a rate can produce water loss, and too low a rate may make too gentle a drawdown curve. The water levels measured in observation wells represent the average head at the screen of the observation wells. The measured drawdown responses to be taken during a pumping test are of two kinds, measurements of the water levels in the well and the observation wells and measurements of the discharge rate of the well.

To be practical, each pump test should last no more than one day (1440 minutes), which is enough to cover more than three time log cycles; 10, 100 and 1000 minutes, respectively. Measurements should also be made of the atmospheric pressure, temperature, the levels of nearby surface water, if present, and any precipitation. A longer duration pump test is good only for the analysis of boundary conditions (16). Drawdown data for the time period to overcome well bore storage should not be plotted for analysis.

A step drawdown test should not be conducted because it is a test for well yield and well efficiency for well drillers' use and not for the analysis of geohydrologic parameters (18).

Single Dimensional Pump Test Analysis. The data analysis for the result of the pumping test is dependent on the aquifer conditions such as isotropic or anisotropic aquifer, homogeneous or non-homogenous aquifer, non-leaky or leaky aquifer, confined or unconfined aquifer, confining aquitard, etc. Data analysis also depends on the drawdown condition such as steady state or transient conditions. Before interpreting the pumping test result, conversion of the data into appropriate units and correction of the data for external influences are necessary. The interpretation of the pumping test data is primarily a matter of identifying an unknown system. Theoretical models comprise the type of aquifer and the initial and boundary conditions, which, in a pumping test, affect the drawdown behavior of the system in their own individual ways.

Old fashioned curve matching techniques and computer fitting analysis all have inherent errors such as matching inaccuracy and oversimplified statistical assumptions. Therefore, curve matched analysis results may not be very accurate.

It is more practical to do a pump test analysis on semilog paper using the straight line plot interpretations established by Cooper and Jacob (19) for hydraulic conductivity, storativity, and radius of influence from drawdown versus time and drawdown versus distance semilog plots. The scatter of the data points is also an indication of whether anisotropic or heterogeneous conditions are encountered. An isotropic and homogeneous aquifer will be reflected by subparallel slopes on the drawdown versus time plots for observation wells, and all points on the drawdown versus distance plot will fall on a straight line (only isotropic). If the slopes of the drawdown versus time plots are not subparallel, the aquifer is heterogeneous. If the points on the drawdown versus distance do not fall or fit onto a straight line (scattered), the aquifer is anisotropic or shows a potential preferential flow direction such as channel flow, fracture media, or artificial conduits of migration (20,21).

For aquifer test analysis in unconfined aquifers, the Boulton analytical method will apply for determining hydraulic conductivity and specific yield or storativity (22). New analytical solutions for evaluating the drawdown near horizontal and slanted wells with finite length screens in water table aquifers are presented in Zhan et al. (23). These fully three-dimensional solutions consider instantaneous drainage or delayed yield and aquifer anisotropy.

For the analysis of the leakage property of the confining layers of the aquifer, Hantush (24) developed a semilog straight line analytical method, known as the 'inflection point' method.

Multidimensional Pump Test and Analysis. Two-dimensional horizontal hydraulic conductivity anisotropic tests (major and minor tensor) require three or more observation wells at different distances and orientations from the fully penetrating pumping well (25,26). The Hantush and Thomas analytical method is a hand contouring method which is not very accurate and may require many observation wells. The Papadopulos method requires drawdown data from three or more observation wells at the same elapsed time and a least squares or equivalent numerical fitting technique for an elliptical fit. The fitted ellipse orientation and axes thus define the horizontal anisotropic hydraulic conductivity tensors (20,21) Loo (27).

For the analysis of vertical hydraulic conductivity (not leakage from confining layers), the pump test requires nonoverlapping partially penetrated wells, one pumping well and one or more observation well, located within 2.5 times the aquifer's thickness from the pumping well (28). The pump test procedure for multi dimensional hydraulic conductivity tensors was first conducted in an alluvial deposit at Christensen Ranch in Powder River Basin, Wyoming, at an *in situ* uranium mining test site. Loo et al. (29) conducted a similar test at the Equity/DOE BX *in situ* oil shale project.

All pump test and analysis procedures described were documented in the manual written by Loo (20), Loo (30) and Loo (27).

Effective Porosity Testing and Analysis. No pump test can analyze effective porosity. The specific yield value from pump test analysis provides only a partial characterization because the specific retention is not defined.

Freeze and Cherry (1) described the tracer test analysis for dispersivity and effective porosity. However, the tracer test procedures have not considered retardation of the sorption properties of the solid matrix. The selection of different tracers for different media is an art more than a science because there is no perfect tracer. Therefore, a tracer test for determining effective porosity is not very practical.

The analysis of effective porosity can actually be easily determined (though seldom used) from stressing or loading an aquifer. This can be done by tidal efficiency or barometric efficiency tests or simply by sucking a vacuum on the pump test well. The effective porosity can then be calculated once the storativity is defined by the pump test analysis. Jacob (31) provides an analytical solution for a confined aquifer. Hantush (32) provides an analytical solution for an unconfined aquifer.

Radioactivity can be mapped by a natural gamma ray borehole geophysical log. If an abnormal level of radiation is mapped, then it may be necessary to test for radioisotopes of uranium, radon, and tritium and the level of their radioactivity (14).

Electrokinetic Parameter Testing and Analysis. A resistivity survey can be conducted in boreholes to distinguish the lithologic layering sequence. A surface resistivity survey can define lateral continuity or discontinuity or boundaries (33,34). These geophysical surveying methods provide large areal extent and vertical definition or three-dimensional mapping of the groundwater body.

The electrokinetic gradient was mentioned in only one paragraph in Freeze and Cherry (1). The electrokinetic gradient can be tested between monitoring wells by impressing a dc voltage across the wells (15). The response of the water level rise in the cathode well can be quite large (some times more than 10 feet) between wells 50 feet apart at a modest 50 volts potential difference and 10 amperes of dc flow. However, there is no standard test procedure at this time. This is an evolving field testing technology which may ultimately partly replace the standard pumping test. It may also provide artificially induced desorption and oxidation/reduction environments for contaminant treatment.

Elemental adsorption can be tested for cation and anion exchange capacity on soil samples in the laboratory (35). The amount of adsorbed chemical can be estimated from the Freundlich and Langmuir approximations (11). There is no standardized laboratory test for adsorption/absorption parameters of on organic compounds a fine-grain solid matrix at this time. But these are important properties because for most contaminant mass migration (mostly not very soluble), more than 90% by weight is adsorbed/absorbed by the solid matrix.

Biochemical Parameters

pH
bacteria count (17)
inorganic chemicals
organic chemicals
oxidation/reduction
ion exchange potential

All these parameters can be determined by standardized U.S. EPA laboratory analytical methods for water and wastewater. The redox potential of groundwater was required only recently for evaluating natural attenuation or intrinsic bioremediation.

However, there is very little emphasis on the test procedures for a solid matrix below the groundwater level. Most people treat soil as soil in the vadose zone. There is no requirement for characterizing soil or rock properties underneath the water table (2,3). It is important to understand that the weight of solid matrix underneath the groundwater table represents more than 90% of the total weight. The laboratory procedures can easily be changed to accommodate the characterization of these properties in a solid matrix submerged under groundwater.

USE, INTERPRETATION, AND LIMITATIONS

After the aquifer system's properties are properly characterized, then groundwater use planning and management can be implemented.

Groundwater Resources Use

Groundwater is a limited resource, and it is very easy to overdraw a groundwater system. The result is water shortages, coastal saline water intrusion, and land subsidence. The remedy for the situation is to ensure that the design of the groundwater withdrawal system is within safe yield limits that is, using groundwater at less than the natural recharge rate. Groundwater recharge usually can be expressed as 1 to 10% of natural precipitation. The rate of recharge is dependent on the evapotranspiration rate of the area.

The remedy for saline water intrusion or upwelling is either to pump less or design a water mound or recharge mound near the coastal area (5). It is most unfortunate that, once the aquifer is intruded by saline water, the aquifer will be difficult to clean up by natural dilution.

The remedy for land subsidence is to use surface water to recharge the groundwater during the wet season. This is cyclic recharge and pumping management of a groundwater basin.

Groundwater Pollution, Prevention, and Remediation

Since the U.S. EPA was formed about 30 years ago, groundwater pollution has not stopped. Nitrate pollution due to overfertilization and from feed lots, dairy farms, pig farms, and chicken farms continues unabated. This has resulted in the 'dead zone' at the estuary of the Mississippi River in the Gulf of Mexico. Inland, the Salton Sea and the Kesterson Reservoir in California here become irrigation drainage wastewater catch basins. Farmers and agricultural business are often exempted from environmental protection regulations. Most shallow groundwater in these areas is heavily polluted.

Organic solvents such as PCE and TCE exist at almost all Superfund sites, electroplating shops, dry cleaners, and electronic manufacturing sites. Fuel hydrocarbons containing benzene, MTBE, and PAHs (such as naphthalene in diesel) have reached the groundwater at many groundwater supply sources. Chromium and naturally occurring arsenic are also reaching groundwater supplies. Fortunately, there are cost-effective technologies for cleanup at the source (15).

FUTURE IMPROVEMENTS

As mentioned in the previous sections, there is much room for improvement in geohydrologic characterization. Characterization efforts need large-scale representation area. This will lead to large-scale geohydraulic tests, geophysical surveys, and more in-depth characterization of the solid matrix submerged under groundwater.

New Test Methods

Future geohydraulic tests will most likely replace the aquifer pump test by stressing the aquifer by seismic or electrokinetic methods, so that it is more cost-effective. The mapping of anisotropic flow can be done better by horizontal and vertical resistivity profiling to map these geohydrologic anomalies. Electrokinetic surveys may in the future characterize adsorption, ion exchange, electrical charge properties in situ and have a large areal representation.

Interpretive Techniques

When all is said and done, then it's time to do some real groundwater modeling using real data. There is no doubt that the theory and modeling effort have advanced much further than parameter characterization methods.

After trying for more than 20 years, leaders in geohydrology admitted in the early 1990s that groundwater modeling does not work because of the general lack of real geohydrologic data. Hopefully, groundwater modeling will work in the next decade with real data on hand.

BIBLIOGRAPHY

1. Freeze, A.R. and Cherry, J.A. (1979). *Groundwater*. Chapter 2, 9, Prentice Hall, New York, 384–462, pp. 1579.
2. RCRA Groundwater Monitoring Technical Enforcement Guidance Document. (1986). National Water Well Association.
3. Mercer, J.W. and Spalding, C.S. (1991). Site Characterization for Subsurface Remediation, EPA/625/4-91/026.
4. Driscoll, F.G. (1986). *Groundwater and Wells*. Glossary, Johnson Filtration Systems Inc., St. Paul, MN, pp. 885–892.
5. Todd, D.K. (1980). Groundwater Hydrology. In: *Groundwater and Well Hydraulics*. John Wiley & Sons, New York, pp. 111–163, Chapter 4.
6. Lohman, S.W. (1979). *Ground-water Hydraulics*. USGS Prof. Paper 708, p. 72.
7. Batu, V. (1998). Aquifer hydraulics. *A Comprehensive Guide to Hydrogeologic Data Analysis*. John Wiley & Sons, New York.
8. Ferris, Knowles, and Stallman. (1965). *Theory of Aquifer Tests*. USGS Water Supply Paper 1536-E, p. 174.
9. Loo, W.W., Wang, I.S., and McSpadden, W.R. (1984). *Effective Porosity, Basalt Waste Isolation Project (BWIP)*. Richland, WA, Report no. SD-BWI-TI-254.
10. Loo, W.W. (1984). Practical Groundwater Modelling and Effective Porosity. *Proceedings of Groundwater Modelling Conference*, National Water Well Association.
11. Devinny, J.S. et al. (1990). Subsurface migration of hazardous wastes. In: *Chemical and Physical Alteration of Wastes and Leachates*. Van Nostrand Reinhold, New York, pp. 142–168, Chapter 5.
12. Myers, D. (1991). *Principles and Applications of Surfaces, Interfaces and Colloids*. VCH Publishers, New York.
13. Rosen, M.J. (1989). *Surfactants and Interfacial Phenomena*. John Wiley & Sons, New York.

14. Murray, R.L. (1994). *Understanding Radioactive Wastes*. Battelle Press, Colombus, OH, pp. 11–22.
15. Lehr, J.H. (Ed.). (2001). *Standard Handbook of Environmental Health, Science and Technology*. Chapters 14.4 & 14.6 on Bioremediation & Electrokinetic Treatment of Hazardous Wastes by Walter W. Loo, McGraw-Hill, New York.
16. Heath, R.C. and Trainer, F.W. (1981). *Introduction to Groundwater Hydrology*. Chapter 2, 12, Water Well Journal Publishing Co., pp. 13–17, 157–170.
17. Gaudy, A.F. and Gaudy, E.T. (1980). *Microbiology for Environmental Scientists and Engineers*. McGraw-Hill, New York.
18. Walton, W.C. (1987). *Groundwater Pump Tests Design and Analysis, Step Drawdown Test Analysis*. Lewis Publishers, Boca Raton, FL, pp. 77–79.
19. Cooper, C.C. and Jacob, C.E. (1946). A generalized graphical method for evaluating formation constants and summarizing well field history. *Trans. Amer. Geophysical Union* **27**: pp. 526–534.
20. Loo, W.W. (1987). *Standard Operational Procedures of Pump Tests and Analysis*. Certificate Seminar Manual for SUPERFUND and HAZMAT Conferences.
21. Loo, W.W. (1989–1995). *Groundwater Hydrology for Beginners Certificate Seminar*. HAZMAT & SUPERFUND Conferences.
22. Boulton, N.S. (1954). *Unsteady State Radial Flow to a Pumped Well Allowing for Delayed Yield from Storage*. International Assoc. Sci. Hydrology, General Assembly, Rome, vol. 2, Publ. 37.
23. Zhan, H., Zlotnik, V.A., and Park, E. (2001). Hydraulics Of Horizontal And Slanted Wells. In: *Water Table Aquifers*, GSA Annual Meeting, November 5–8, 2001 Proceedings, Session No. 117.
24. Hantush, M.S. (1956). Analysis of Data from Pumping Test in Leaky Aquifer. *Trans. Amer. Geophysical Union* **37**: pp. 702–714.
25. Hantush and Thomas. (1966). Analysis of Data From Pumping Tests in Anisotropic Aquifers. *Journal of Geophysical Research* **71**: No. 2
26. Papadopulos, I.S. (1965). Nonsteady flow to a well in an infinite anisotropic aquifer. *Symp. Intern. Assoc. Sci. Hydrology, Dubrovinik Symposium* **1**: pp. 21–31.
27. Loo, W.W. (1989–1995). *Groundwater Plume Management Certificate Seminar*. SUPERFUND Conferences, and HAZMAT Conferences, Washington, DC.
28. Weeks, E.P. (1969). Determining the ratio of horizontal to vertical permeability by aquifer test analysis. *Water Resources Research* **5**: 1.
29. Loo, W.W., Markley, D.E., and Dougan, P. (1979). *Three Dimensional Multiple Well Testing and Reservoir Analysis of the Leached Zone of the Green River Formation*. Equity/DOE BX In-situ Oil Shale Project, Proceedings of the Ninth Annual Rocky Mountain Groundwater Conference, Reno, NV, October 22, 1979.
30. Loo, W.W. (1989–1995). *Principles of Groundwater Hydrology Certificate Seminar*. HAZMAT Conferences.
31. Jacob, C.E. (1950). Flow of groundwater. In: *Engineering Hydraulics*. John Wiley & Sons, New York, Chapter 5.
32. Hantush, M.S. (1964). Hydraulics of Well. In: *Advances in Hydroscience*. Academic Press, New York.
33. Milsom, J. (1989). Field Geophysics, Electrical Methods. *Geological Society of London Handbook Series*. Open University Press and Halsted Press, pp. 72–106.
34. Vogelsang, D. (1995). Environmental Geophysics. *Geoelectrical Methods*. Springer-Verlag, New York, pp. 9–31.
35. Dragun, J. (1988). *The Soil Chemistry of Hazardous Materials*. Hazardous Materials Control Research Institute, pp. 75–262, Chapters 3–6.

MOBILITY OF HUMIC SUBSTANCES IN GROUNDWATER

Gunnar Buckau
Institut für Nukleare
Entsorgung
Karlsruhe, Germany

INTRODUCTION

Humic substances are a relatively stable part of the global carbon inventory. They are found mainly in solid sources, in both seabed and land sediments. High concentrations are found especially in lignite and peat. The most obvious source is the organic/humus inventory in soil. Humic substances are also found in natural water. The dissolved inventory is relatively small compared to that found in sediments. Nevertheless, this inventory is a key in the mobilization of numerous trace elements and pollutants, which includes the key influence in determining the stability of and interface reactions with minerals.

Some key information on the origin, stability, and mobility of dissolved aquatic humic substances is presented here. Also described are the experimental methods required, the approach for derivation of conclusions from experimental results, and background information. The focus is on the hydrophilic part of the total humic substance inventory. The hydrophilic aquatic humic substances originate from the same source as the stationary less hydrophilic ones, including partial oxidation from this stationary inventory. The hydrophilic nature is given by oxygen-containing functional groups. Some of these, especially carboxylic acids, may change the degree of hydrophilicity depending on physicochemical conditions, that is protonation, metal ion complexation and localization of counterions, pH, metal ion concentration, and ionic strength. Consequently, the whole inventory and changes in physicochemical conditions must be kept in mind when discussing the behavior of the hydrophilic aquatic part.

Humic acid is used both as a collective term for humic substances with a sufficiently high content of hydrophilic groups to dissolve in pH neutral range or specifically for the fraction that is dissolved in pH neutral range and flocculates in the acidic range with the complementary fraction of fulvic acid, soluble also under acidic conditions. Furthermore, a fraction of organic carbon found in soil not soluble in aqueous media (humin) but with some properties common with humic and fulvic acids is also frequently considered part of the overall term. This term may also be considered for rather hydrophobic humic material or humic acid precursor material also in sediments, including peat, lignite, and mineral-bound natural organic substances, such as clay organic matter. Below, humic and fulvic acids are used for the aquatic hydrophilic part of the overall inventory unless otherwise specified.

First, some basic relevant properties of humic substances are discussed. In order to study the origin, stability, and mobility of aquatic humic substances they must be isolated and purified. In order to ensure the quality of results, the success of isolation and purification must be verified by adequate characterization of the sampled humic material. Therefore, isolation, purification, and verification methods of aquatic humic substances are described as a prerequisite for interpretation of experimental results. Also discussed are the origin, stability, and mobility of aquatic humic acid followed by a summary.

BASIC RELEVANT PROPERTIES OF HUMIC ACID

The composition, mass distribution, and functional group content vary within limits, reflecting different origins and histories. Disregarding sulfur and nitrogen (present in varying low concentrations generally on the order of up to around 1% by weight), the atomic composition is dominated by carbon, oxygen, and hydrogen. The contributions of these substances vary around $CO_{0.5}H$; that is, the formal oxidation state of carbon is around zero. With respect to aquatic humic and fulvic acids, the mass distribution centers around approximately 500 mass units. In dissolved form, the molecules are highly hydrated and possibly form associates. The size distribution is generally determined to center slightly above 1 nm diameter.

Oxygen-containing functional groups and structural entities are abundant, as seen from the elemental composition. These oxygen-containing structural elements are of key importance for the hydrophilic character of humic substances. The oxygen to carbon ratio of aquatic humic and fulvic acids is somewhat higher than for the bulk sedimentary hydrophobic humic matter. The proton exchange capacity is on the order of 7 meq/g, where protonation/deprotonation takes place from about pH 10 down into the acidic range (below pH 3 a considerable number of the groups are still ionized). There is clear spectroscopic evidence for both carboxylic and phenolic types of proton exchanging groups. Carboxylic types of groups are normally quantified with about two-thirds of the total capacity. The hydrophilic character and thus the stability in aqueous solution, of humic and fulvic acids, vary with pH, ionic strength, and metal ion complexation.

ISOLATION AND PURIFICATION METHODS

Isolation

Isolation of humic and fulvic acids from natural water, including groundwater, is based on sorption chromatography. For this purpose, the XAD-8 resin is used where humic and fulvic acids sorb at low pH. Depending on the humic substance concentrations in the water sample, a pH of typically 1 or 2 will be used, the latter especially in order to limit the total amount of acid where the sample volumes are very large. Elution of both humic and fulvic acids is achieved by desorption at high pH. Typically, HCl and NaOH are used for adjustment of pH. In order to remove salt from the original sample, the column should be washed with HCl solution (pH not higher than that of the acidified water sample) prior to alkaline humic and fulvic acid elution. Where the humic acid concentration is sufficiently high, it, or at least a large part of it, will flocculate at low pH and thus is separated prior to sorption chromatography treatment.

The alkaline elute solution should be taken to pH neutral range within a relatively short time (ideally after not more than around 10 min) in order to prevent alkaline hydrolysis and sample oxidation. If desired, the humic and fulvic acid fractions are separated from each other from this concentrate by acidification, where the humic acid fraction flocculates and may be separated by centrifugation. The so obtained humic and fulvic acid concentrates are then further processed and purified. In many cases, however, preconcentration is desired, especially directly in the field.

The application of RO to humic and fulvic acid preconcentration makes use of the retention of these substances over the RO membrane. The retentate, containing salts, particles, and the humic and fulvic acids, is recirculated as long as practical with the clean water from penetration of the membrane being discarded. With increasing retentate salt concentration, the permeation of clean water through the RO membrane is decreased and finally ceases for all practical purposes (pressure buildup). Therefore, the level of preconcentration that can be obtained by RO depends on the salt content of the original water sample. Furthermore, where Ca ions and carbonate are present, calcite will precipitate unless the sample is slightly acidified and carbonic acid released. The concentrate obtained by RO is then treated by XAD-8 chromatography as described above, resulting in further sample concentration and removal of salt.

Purification

Purification has a number of objectives, including removal of organic and inorganic contaminants such as salt, complexed metal ions, and inorganic mineral constituents. For the purpose of dissolving Si-based minerals, NaF is added and the samples are left for typically 24 h. With respect to humic acid, subsequent purification is relatively simple. Humic acid is flocculated in HCl (pH 1) and centrifuged, and the supernatant is discarded, followed by dissolution in weak NaOH. The cycle is repeated until the flocculate is finally washed with HCl until no Na (from NaOH) is found. The sample is then in its protonated form and is brought to the final product by freeze-drying.

The fulvic acid does not flocculate at low pH and thus the procedure is more tedious. The fulvic acid is sorbed/desorbed on XAD-8 in a number of cycles. Finally, the slightly alkaline solution is acidified and protonated by cation exchange chromatography. The protonated fulvic acid in its final form is then obtained by freeze-drying of this solution.

The method described has successfully been applied for isolation of about 200 mg fulvic acid from about 10 m^3 of groundwater. Subsequent analysis was successfully used for ^{14}C dating and determination of general characteristic properties (1). Details on the isolation and purification of humic and fulvic acids, including description of the RO technique, can be found in Artinger et al. (2).

Verification

Verification of the purity of humic and fulvic acids is done by characterization of the final products. It should also be borne in mind that trust in published results on the mobility of humic and fulvic acids requires evidence for successful separation and purification from the source. Characterization of humic and fulvic acids can be done by a broad spectrum of methods. For the purpose of study concerning the present topic, the quality of the samples needs to be demonstrated by a minimum of characterization methods. The elemental composition, the concentrations of inorganic constituents, and UV/Vis and IR spectra should be compared with expectation values from the literature.

Inorganic constituents will be exhibited if inorganic minerals and complexed metal ions are removed to an acceptable level. The UV/Vis absorption increases uniformly with decreasing wavelength and the logarithm of the absorption should be close to linear with the wavelength. For fulvic acids the absorption ratio of 300–400 nm is expected to fall in the range between about 5 and 8 (abs. units/g and cm), whereas the values for humic acids are lower and are expected to fall around 2.5–4. These indicators show whether the UV/Vis absorbing carbon inventory is in agreement with humic and fulvic acids.

IR spectroscopy shows a number of characteristic bands. Comparison to published spectra will reveal if there are considerable amounts of non-UV/Vis absorbing organic contaminants. Presence of such contaminants may also be identified by strong deviation from an atomic ratio around $CO_{0.5}H$.

For the purpose of determining the origin and age of humic and fulvic acids, the ^{14}C concentration is essential. In this context the ^{13}C concentration is a good indicator for possible contaminants. The ^{13}C concentration should be close to −27‰ (rel. PDB) for C-3 plant cycle origin and around −13‰ for C-4 plant cycle origin.

Again, only if a sufficient purity of the isolated substances is demonstrated can the results be considered trustworthy.

ORIGIN, STABILITY, AND MOBILITY OF AQUATIC HUMIC ACID

Determination of the origin, stability (with respect to decomposition and sorption), and mobility of aquatic humic and fulvic acids in groundwater is based on several indicators. The origin can be determined based on ^{14}C content, functional entity distribution, and general information on climatic conditions, vegetation history for soil recharge, and sedimentary and groundwater composition. Another important indicator is the cogeneration with dissolved inorganic carbon (DIC) of biogenic origin, where deviations from such a correlation show that either the DIC or the humic and fulvic acid inventory deviates from ideal tracer transport and stability behavior.

Origin

The principal origin of humic and fulvic acids is plant material. In sea sediments, marine organisms may also play a role. From the viewpoint of aquatic humic and fulvic acids in a groundwater, two principal sources may be distinguished. These are introduction from the soil zone with recharge groundwater and *in situ* generation by conversion of organic sediment material. In the former case, the source term will depend on climatic conditions and the type of vegetation, including extensive vegetation versus, for example, intense modern agriculture and land conditions such as wetland and peat deposits. In the case of *in situ* generation, an oxidizing agent and partial oxidation of the hydrophobic source material by microbial activity are required (3,4).

The different origin can be distinguished by some characteristic properties and especially the ^{14}C content (5). In aquifer systems where the groundwater residence time is sufficiently low, compared to the half-life of ^{14}C (57,300 years), the ^{14}C content of recharge humic and fulvic acids is that of the soil recharge source. Numbers deduced for fulvic acid for conditions prior to nuclear atmospheric testing are around 55 pmc (percent modern carbon) (5), basically the same as for the cogenerated dissolved inorganic carbon of biogenic origin (6). Depending on local conditions, however, this value may vary. One example is modern agriculture, where the turnover of the organic soil inventory may be higher and thus the ^{14}C concentration is higher, reflecting the lower average residence time. The ^{14}C concentration of fulvic acid and cogenerated DIC from *in situ* generation is basically zero. The relative fractions of *in situ* generated and recharge originating fulvic acid can be deduced by the overall ^{14}C concentration (5).

Another important indicator for the origin of aquatic humic and fulvic acids is the concentration varying with recharge conditions. Drainage of wetland results in a decrease in the inflow of both aquatic humic and fulvic acids and the cogenerated dissolved inorganic carbon of biogenic origin (7), which is an indicator of changes in the land use and also of changes in climatic conditions with a drastic impact on land structure and vegetation. Recharge humic and fulvic acids can only be found in water originating from areas with vegetation. For this reason, the absence of humic and fulvic acids is an indicator for vegetation free recharge conditions, also indicative of climatic changes.

In one study ^{14}C groundwater dating on the DIC indicated ages well beyond 15,000 years for the most distant part of the groundwater flow system (1). Not only did the ^{14}C concentration of fulvic acid show that these age determinations were wrong, but also the pure presence of fulvic acid showed that the groundwater at the end of the flow path could not be older than 15,000 years. The reason is that, in this area, vegetation started to develop around 15,000 years ago at the decline of the Pleistocene conditions.

Some characteristic properties of aquatic humic and fulvic acids reflect their origin. Both spectroscopic properties and basic functional entity distribution are sufficiently insensitive to the physicochemical environment to retain origin-related properties, which is especially true for the ^{14}C content, which solely depends on the concentration upon their generation and subsequent decay with a half-life of approximately 5730 years. Extraction of hydrophilic humic and fulvic acids from clay sediments revealed that the distribution of functional entities and photodynamic behavior vary with marine and terrestrial origin of the clay sediments (8,9). Other characteristic properties reflect chemical reactions of functional groups, especially redox reactions. It should thus be kept in mind that not all characteristic properties can be used as indicators for the origin of aquatic humic or fulvic acid. The content and redox state of sulfur functional entities partly reflect the redox conditions and presumably also the microbial sulfate reduction as part of the in situ generation process (5).

Stability and Mobility

In groundwater systems where the in situ generation is negligible, the fulvic acid inventory is given solely by recharge and possible follow-up geochemical reactions. Where the residence time is considerable compared to the half-life of ^{14}C, the age of the fulvic acid can be determined. A linear increase in ^{14}C age of fulvic acid with flow distance from recharge was found, with the maximum age of fulvic acid of about 15,000 years (cf. above and Ref. 1). The agreement with groundwater flow velocity calculations and the age of fulvic acid shows that the fulvic acid has an ideal tracer transport behavior over this time period, which is direct evidence for the stability and mobility with an ideal tracer behavior of aquatic fulvic acid over 15,000 years.

In the case of *in situ* generation of fulvic acid, the quantification of the recharge originating inventory by ^{14}C content and supporting spectroscopic and composition information is required. The outcome of one study is that in situ generated fulvic acid is flocculated in deeper brines, whereas the recharge originating fulvic acid remains stable in the brines. Both the conserved ^{14}C concentration and structural entity content by C-XANES show that the fulvic acid in the deeper brines (200–250 m depth) originates from recharge. The flocculation of in situ generated fulvic acid in the high ionic strength brines is shown to be the result of their less hydrophilic nature compared to the recharge fulvic acid. The similar concentrations in recharge and in the deep brines show the stability and mobility of these fulvic acids over at least several hundreds years and, furthermore, that the recharge conditions have not changed dramatically within this time period. In addition, the close correlation between the DIC of biogenic origin and humic and fulvic acids in the broad spectrum of groundwater samples shows that there is no considerable decomposition or retention of fulvic acid from either recharge or in situ generation. This correlation is also found for deep groundwater from a former wetland. This shows also that the elevated concentration from this previous source is not subject to sorption or decomposition but the dissolved humic and fulvic acids remain stable and mobile over long time periods (7).

Studies on clay organic matter show that the organic matter sorbed on the sediments are mainly of hydrophobic character and are strongly bound to mineral surfaces. Studies where humic or fulvic acid is in contact with fresh mineral surfaces frequently show strong sorption. The natural situation, however, is that active sites on sediments are coated with organic substances and subsequent introduction of hydrophilic aquatic humic and fulvic acids will behave differently and remain in solution rather than competing with sorbed hydrophobic/sediment bound organic material. Care must therefore be taken when analyzing experimental data that the real situation is reflected. Strong sorption may be expected where a recent increase in the inflow of humic acid/fulvic acid is given, such as a deposit leaking organic material into sediment previously in contact with very low organic carbon concentrations. Under stable groundwater conditions, no indication is given for retention or decomposition of humic and fulvic acids from recharge.

CONCLUSION

With respect to the origin of aquatic humic and fulvic acids, a large number of studies have shown that they are cogenerated with inorganic carbon of biogenic origin from microbial processes and that there is a strong variation in the concentration of humic and fulvic acids in natural groundwater—an indicator for recharge conditions, including climatic changes. With respect to the stability and mobility, there is no indication for decomposition of aquatic humic and fulvic acids under natural groundwater conditions for as long as 15,000 years and no indication for their retardation. An exception is the selective flocculation of less hydrophilic in situ generated fulvic acid in high ionic strength brines.

BIBLIOGRAPHY

1. Buckau, G. et al. (2000). Development of climatic and vegetation conditions and the geochemical and isotopic composition in the Franconian Albvorland aquifer system. *Appl. Geochem.* **15**(8): 1191–1201.
2. Artinger, R. et al. (2000). Characterization of groundwater humic substances: influence of sedimentary organic carbon. *Appl. Geochem.* **15**(1): 97–116.
3. Buckau, G. et al. (2000). Groundwater in-situ generation of aquatic humic and fulvic acids and the mineralization of sedimentary organic carbon. *Appl. Geochem.* **15**(6): 819–832.
4. Buckau, G. et al. (2000). Origin and mobility of humic colloids in the Gorleben aquifer system. *Appl. Geochem.* **15**(2): 171–179.
5. Schäfer, T., Buckau, G., Artinger, R., Wolf, M., Kim, J.I., Geyer, S., Bleam, W.F., Wirick, S., and Jacobsen, C. (2004). Vertical exchange of Gorleben fulvic acids of different origin, Annex 1. In: *Humic Substances in Performance Assessment of Nuclear Waste Disposal: Actinide and Iodine Migration in the Far-Field (Second Technical Progress Report)*. G. Buckau (Ed.). Report FZKA 6969, Research Center, Karlsruhe, Germany.
6. Buckau, G. et al. (2000). ^{14}C Dating of Gorleben groundwater. *Appl. Geochem.* **15**(5): 583–597.
7. Buckau, G., Wolf, M., Geyer, S., Artinger, R., and Kim, J.I. (2003). Origin and mobility of aquatic humic substances from wetland recharge in the Gorleben aquifer system. In: *Humic Substances in Performance Assessment of Nuclear Waste Disposal: Actinide and Iodine Migration in the Far-Field (First Technical Progress Report)*. G. Buckau (Ed.). Report FZKA 6800, Research Center, Karlsruhe, Germany.
8. Claret, F., Schäfer, T., Rabung, Th., Bauer, A., Wolf, M., and Buckau, G. (2004). Complexation properties of humic and

fulvic acids extracted from Callovo-Oxfordian and opalinus clay, Annex 23. In: *Humic Substances in Performance Assessment of Nuclear Waste Disposal: Actinide and Iodine Migration in the Far-Field (Second Technical Progress Report)*. G. Buckau (Ed.). Report FZKA 6969, Research Center, Karlsruhe, Germany.

9. Claret, F. et al. (2005). Differences in characteristic properties and Cm(III) complexation between terrestrial and marine origin of humic and fulvic acids from Callovo-Oxfordian and opalinus clay. *Appl. Geochem.*

ASSESSMENT OF GROUNDWATER QUALITY IN DISTRICT HARDWAR, UTTARANCHAL, INDIA

Chakresh K. Jain
National Institute of Hydrology
Roorkee, India

The groundwater quality of District Hardwar in the state of Uttaranchal (India) has been assessed to see the suitability of groundwater for domestic use. Forty-eight groundwater samples from shallow and deep aquifers were collected each during pre- and postmonsoon seasons during the year 2002. Various water quality constituents, pH, conductance, total dissolved solids, alkalinity, hardness, sodium, potassium, calcium, magnesium, chloride, sulphate, phosphate, fluoride, total coliforms, and fecal coliforms were determined. The data were analyzed with reference to BIS and WHO standards, and hydrochemical facies were determined. The concentration of total dissolved solids exceeded the desirable limit of 500 mg/L in about 25% of the samples analyzed, but the values were well within the maximum permissible limit of 2000 mg/L. The alkalinity exceeded the desirable limit of 200 mg/L in about 50% of the samples, but these were also within the maximum permissible limit of 600 mg/L. From the hardness viewpoint, about 80–85% of the samples were within the desirable limits. One sample of the study area exceeded the maximum permissible limit of 100 mg/L for nitrate. Other constituents such as chloride, sulphate, and fluoride were within the desirable limits. The bacteriological analysis of the groundwater samples indicated bacterial contamination in about 40% of the samples analyzed. Inadequate maintenance of hand pumps, improper sanitation, and unhygienic conditions around the structure may be responsible for bacterial contamination in groundwater of the region and is a cause of concern. It is recommended that the water drawn from such sources be properly disinfected before being used for drinking and other domestic purposes. The grouping of samples according to their hydrochemical facies indicates that all the samples of the study area fall under $Ca-Mg-HCO_3$ hydrochemical facies.

INTRODUCTION

Water is an essential and vital component of our life support system. In tropical regions, groundwater plays an important role in the context of fluctuating and increasing contamination of water resources. Groundwater has unique features, which render it particularly suitable for public water supply. It has excellent natural quality; is usually free from pathogens, color and turbidity; and can be consumed directly without treatment. Groundwater is widely distributed and can be frequently developed incrementally at points near water demand, thus avoiding the need for large-scale storage, treatment, and distribution systems. It is particularly important because it accounts for about 88% of safe drinking water in rural areas, where the population is widely dispersed and the infrastructure for treating and transporting surface water does not exist.

Unfortunately, the availability of groundwater is not unlimited, nor it is protected from deterioration. In most instances, extracting excessive quantities of groundwater has resulted in drying of wells, damaged ecosystems, land subsidence, saltwater intrusion, and depletion of the resource. It has been estimated that once pollution enters the subsurface environment, it may remain concealed for many years, becoming dispersed over wide areas of groundwater aquifer and rendering groundwater supplies unsuitable for consumption and other uses. The rate of depletion of groundwater levels and deterioration of groundwater quality are of immediate concern in major cities and towns of the country.

The creation of the new state of Uttaranchal has posed many challenges for planners and policy makers. Problems such as drinking water, transportation, power, housing and construction, and safety against natural hazards are very serious and require immediate attention. For sustainable development of a society, it is essential that the natural resources are used judiciously for the benefit of the existing population and also to meet the needs and aspirations of future generations. Drinking water is one such precious commodity for which a planned strategy is needed for immediate demands and also for sustainability for future needs. A large part of the state of Uttaranchal lies in the hills, where distribution of drinking water supply and its quality is a major problem that needs immediate attention. About 90% of the rural population of this region depends on natural springs for their daily water. However, due to population pressure, unplanned construction, garbage disposal, and changes in land use patterns, the water of these springs is becoming contaminated, and the discharge of these springs is declining.

A wide number of activities are associated with, the human introduction of foreign chemical and biological materials into the subsurface environment. In the long run, the most potentially hazardous of these may be the chemical fertilizers and pesticides used in agriculture. But it is possible that tremendous use of chemical fertilizers as plant nutrients may be a more significant problem, causing an increasing buildup of nutrients in some groundwaters. Bacteriological parameters are of great importance from the human point of view. It is essential to examine the presence of toxic substances and pathogenic organisms in potable water. Experience has established the significance of coliform group density as a criterion of the degree of pollution and thus of sanitary quality. The significance of the various tests and the interpretation of results are well authenticated and have been used as a basis for standards of the chemical and bacteriological quality of water supplies. In this article, the groundwater quality of District Hardwar has been assessed to see the suitability of groundwater for drinking, which will provide a proper basis for judicial management of drinking water supplies in the state.

STUDY AREA

District Hardwar, part of the Indo-Gangetic plains, lies between latitude 29°30′ and 30°20′ N and longitude 77°40′ to 78°25′ E in the state of Uttaranchal (Fig. 1). It is the largest district (in population) of Uttaranchal State and occupies an area of about 2360 km^2. Per the 2001 census, the population of District Hardwar is 14,44,213, a population density of 612 per km^2.

Physiographically, the area is generally flat except for the Siwalik Hills in the north and northeast. The area is devoid of relief features of any prominence except for deep gorges cut by gullies and rivers flowing through the area. The area is bounded by River Yamuna in the west and River Ganga in the east. The climate of the area is characterized by moderate subtropical monsoons. The average annual rainfall in the region is about 1000 mm; the major part is received during the monsoon period. The major land use is agriculture, and there is no effective forest cover. The soils of the area are loam to silty loam and are free from carbonates. The most common groundwater use is achieved by hand pumps and tube wells. Based on the lithologic logs and water table fluctuation data,

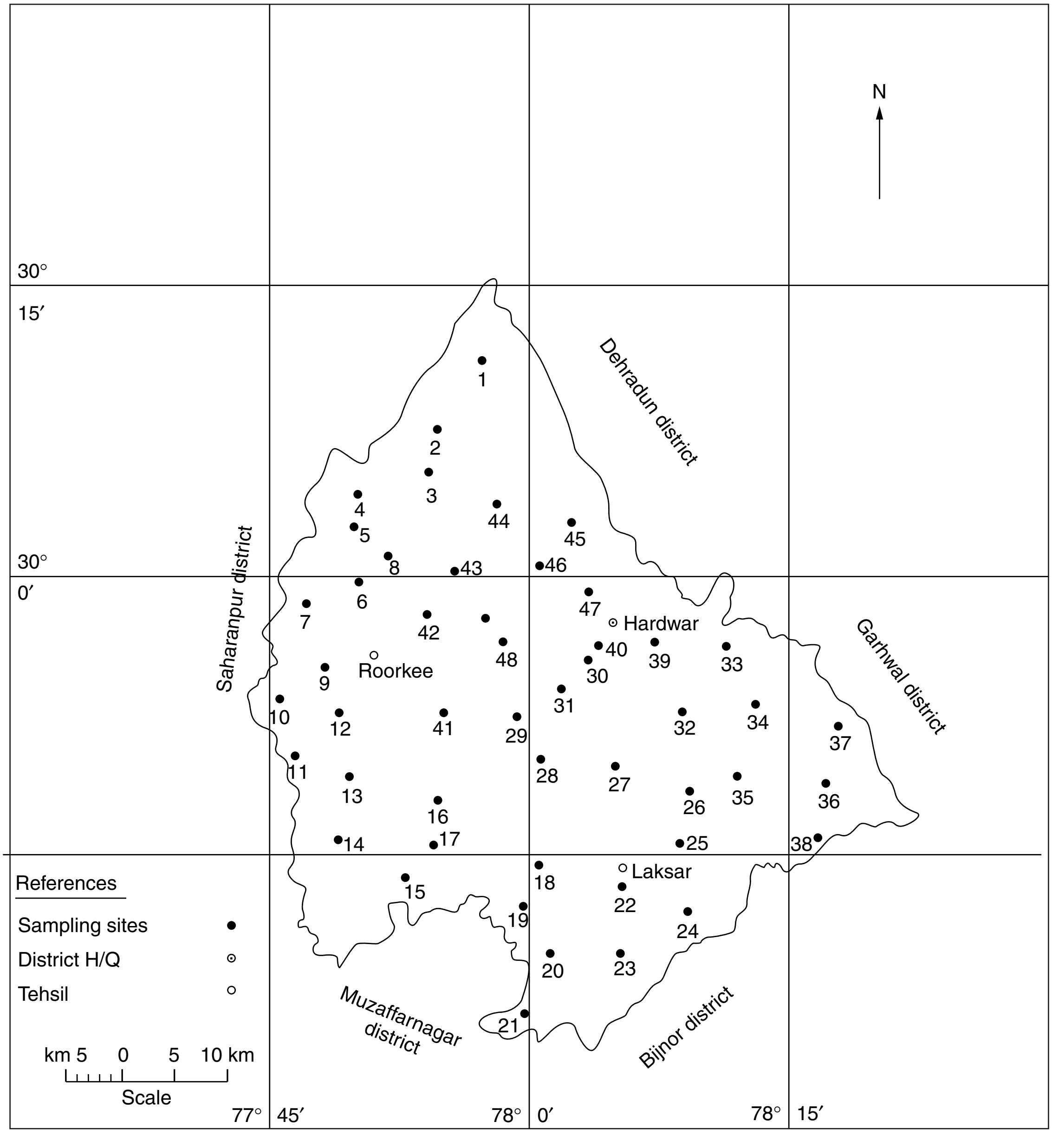

Figure 1. Study area showing location of sampling sites, District Hardwar.

two types of aquifers have been delineated in the area. The upper one is the shallow unconfined aquifer which generally extends to depths of around 25 m. The deeper one consists of semiconfined aquifers located at a depth of about 25 to 150 m below ground level separated by three to four aquifers at average depths of 25 to 55, 65 to 90, and 120 to 150 m. Water table contours in the area indicate a southward trend of groundwater flow in both unconfined and confined aquifers.

EXPERIMENTAL METHODOLOGY

Forty-eight groundwater samples from District Hardwar were collected each during pre- (June 2002) and post-monsoon (October 2002) seasons from various abstraction sources at various depths covering extensively populated area, commercial, industrial, agricultural, and residential colonies so as to obtain a good areal and vertical representation. The samples were preserved by adding an appropriate reagent (1,2). The hand pumps and tube wells were continuously pumped prior to sampling to ensure that the groundwater to be sampled was representative of the groundwater aquifer. The water samples for bacteriological analysis were collected in sterilized high-density polypropylene bottles covered with aluminum foil. All samples were stored in sampling kits maintained at 4 °C and brought to the laboratory for detailed chemical and bacteriological analysis. The details of sampling locations and source and depthwise distribution are given in Tables 1 and 2, respectively.

The physicochemical analysis was performed following standard methods (1,2). The total coliforms and fecal coliforms were determined by the multiple tube fermentation technique using MacConkey broth and EC medium, respectively.

Table 1. Description of Groundwater Sampling Locations in District Hardwar

S. No.	Location	Source	Depth, m	S. No.	Location	Source	Depth, m
1	Mohand	OW	10	25	Sultanpur	HP	15
2	Banjarewala	OW	10	26	Shahpur	HP	25
3	Buggawala	HP	30	27	Pathri	HP	15
4	Kheri	HP	38	28	Subashgarh	HP	20
5	Dadapatti	HP	38	29	Marghubpur	HP	30
6	Bahbalpur	HP	15	30	Bahadarabad	HP	20
7	Sikandarpur	HP	25	31	Alipur	HP	20
8	Bhagwanpur	HP	30	32	Katarpur	HP	10
9	Chudiala	HP	40	33	Kankhal	HP	25
10	Balswa Ganj	HP	25	34	Shyampur	HP	30
11	Manakpur	HP	40	35	Rasiya Garh	HP	40
12	Iqbalpur	HP	30	36	Gandikhatta	HP	10
13	Jharera	HP	30	37	Laldhang	HP	90
14	Sherpur	HP	30	38	Kottawali	HP	40
15	Narsen	HP	30	39	Hardwar	HP	30
16	Manglour	HP	35	40	Jwalapur	HP	30
17	Libarheri	HP	35	41	Roorkee	HP	10
18	Mahesari	HP	20	42	Gumanwala	HP	30
19	Sahipur	HP	35	43	Manubas	HP	10
20	Khanpur	HP	15	44	Bandarjud	HP	15
21	Chandpuri Kalan	HP	15	45	Beriwala	HP	5
22	Laksar	HP	15	46	Hazara	HP	10
23	Kalsiya	HP	10	47	Aurangbad	HP	10
24	Niranjanpur	HP	15	48	Daulatpur	HP	35

OW: Open well.
HP: Hand pump.

Table 2. Source and Depthwise Distribution of Sampling Sites in District Hardwar

Source Structure	Depth Range			Total Number
	<0–20 m	20–40 m	>40 m	
Hand pumps	6,18,20,21,22, 23,24,25,27,28, 30,31,32,36,41, 43,44,45,46,47	3,4,5,7,8,9, 10,11,12,13, 14,15,16,17,19, 26,29,33,34,35, 38,39,40,42,48	37	46
Tube wells	–	–	–	–
Open wells	1,2	–	–	2
Total	22	25	1	48

RESULTS AND DISCUSSION

During 1983, the Bureau of Indian Standards (BIS), earlier known as Indian Standards Institution (ISI), laid down standard specifications for drinking water which have been revised and updated from time to time. To enable the users to exercise their discretion toward water quality criteria, the maximum permissible limit was prescribed especially where no alternate source was available. The national water quality standards describe the essential and desirable characteristics that must be evaluated to assess the suitability of water for drinking (3). The hydrochemical data for the two sets of samples collected from District Hardwar during pre- and postmonsoon seasons are presented in Table 3.

General Characteristics

The pH of the groundwater of District Hardwar is mostly confined within the range 6.22 to 7.58 during the premonsoon season and 6.70 to 7.70 during the postmonsoon season. The pH values of all samples are well within the limits prescribed by the BIS (3) and the WHO (4) for various uses, including drinking and other domestic supplies.

The measurement of electrical conductivity is directly related to the concentration of ionized substances in water and may also be related to excessive hardness and/or other mineral contamination. The conductivity values in the groundwater samples of District Hardwar vary from 233–1440 μS/cm during the premonsoon season and from 221–1442 μS/cm during the postmonsoon season; about 10% of the samples had conductivity values above 1000 μS/cm during both pre- and postmonsoon seasons. The maximum conductivity of 1440 and 1442 μS/cm was observed at village Manubas (hand pump, 10 m depth) during pre- and postmonsoon season, respectively.

In natural waters, dissolved solids consists mainly of inorganic salts such as carbonates, bicarbonates, chlorides, sulfates, phosphates, and nitrates of calcium, magnesium, sodium, potassium, iron, etc. and small amounts of organic matter and dissolved gases. In the present study, the values of total dissolved solids (TDS) in the groundwater varied from 149–922 mg/L during the premonsoon season and from 141–923 mg/L during the postmonsoon season, indicating low mineralization in the area. More than 75% of the samples analyzed were within the desirable limit of 500 mg/L, and about 25% of the samples were above the desirable limit but within the maximum permissible limit of 2000 mg/L. An almost similar trend was observed during the postmonsoon season. The TDS content at deeper levels (>40 m depth) is comparatively low and lies well within the desirable limit of 500 mg/L. The TDS distribution maps for the pre- and postmonsoon seasons are shown in Fig. 2a,b. Water containing more than 500 mg/L of TDS is not considered desirable for drinking water, though more highly mineralized water is also used where better water is not available. For this reason, 500 mg/L as the desirable limit and 2000 mg/L as the maximum permissible limit have been suggested for drinking water (3). Water containing more than 500 mg/L TDS causes gastrointestinal irritation (3). No sample of District Hardwar exceeded the maximum permissible limit of 2000 mg/L.

Table 3. Hydrochemical Data for Groundwater Samples from District Hardwar[a]

Characteristics	Min	Max	Average
pH	6.22 (6.70)	7.58 (7.70)	6.87 (7.10)
Conductivity, μS/cm	233 (221)	1440 (1442)	642 (647)
TDS, mg/L	149 (141)	922 (923)	411 (414)
Alkalinity, mg/L	78 (71)	460 (482)	213 (215)
Hardness, mg/L	81 (80)	464 (427)	209 (212)
Chloride, mg/L	0.1 (1.2)	32 (35)	9.0 (9.0)
Sulphate, mg/L	0.5 (0.7)	72 (62)	22 (21)
Nitrate, mg/L	0.1 (0.1)	140 (130)	13 (13)
Phosphate, mg/L	0.01 (0.02)	0.41 (1.60)	0.03 (0.25)
Fluoride, mg/L	0.01 (0.01)	0.94 (0.88)	0.36 (0.42)
Sodium, mg/L	5.0 (4.5)	69 (71)	26 (25)
Potassium, mg/L	0.8 (0.2)	42 (32)	7.0 (7.4)
Calcium, mg/L	22 (24)	140 (135)	55 (56)
Magnesium, mg/L	5.0 (5.0)	36 (38)	18 (17)
Boron, mg/L	0.12 (0.12)	0.92 (0.87)	0.50 (0.49)

[a] Values given in parenthesis represent postmonsoon data.

Carbonates, bicarbonates, and hydroxides are the main cause of alkalinity in natural waters. Bicarbonates represent the major form because they are formed in considerable amounts by the action of carbonates upon the basic materials in the soil. The alkalinity in the groundwater varies from 78–460 mg/L during the premonsoon season and from 71–482 mg/L during the postmonsoon season. About 50% of the samples of the study area fall within the desirable limit of 200 mg/L both during the pre- and postmonsoon seasons, and the remaining 50% of the samples exceeds the desirable limit but are within the maximum permissible limit of 600 mg/L. No sample of the study area exceeded the maximum permissible limit of 600 mg/L. The high alkalinity may be due to the action of carbonates upon the basic materials in the soil.

Calcium and magnesium along with their carbonates, sulfates, and chlorides make the water hard. A limit of 300 mg/L has been recommended for potable water (3). The total hardness values in the study area range from 81–464 mg/L during the premonsoon season and from 80–427 mg/L during the postmonsoon season. About 80% of the samples of the study area fall within the desirable limit of 300 mg/L and the remaining samples exceed the desirable limit but are well within the maximum permissible limit of 600 mg/L. From the point of view of hardness all samples of District Hardwar were within the permissible limit of 600 mg/L.

The desirable limits for calcium and magnesium for drinking water are 75 and 30 mg/L, respectively (3). In the groundwater of the study area, the values for calcium and magnesium range from 22–140 mg/L and 5.0–36 mg/L, respectively, during the premonsoon season. An almost similar trend was observed during the postmonsoon season. In groundwater, the calcium content generally exceeds the magnesium content in accordance with their relative abundance in rocks. The increase in magnesium is proportionate to calcium in both seasons. All the

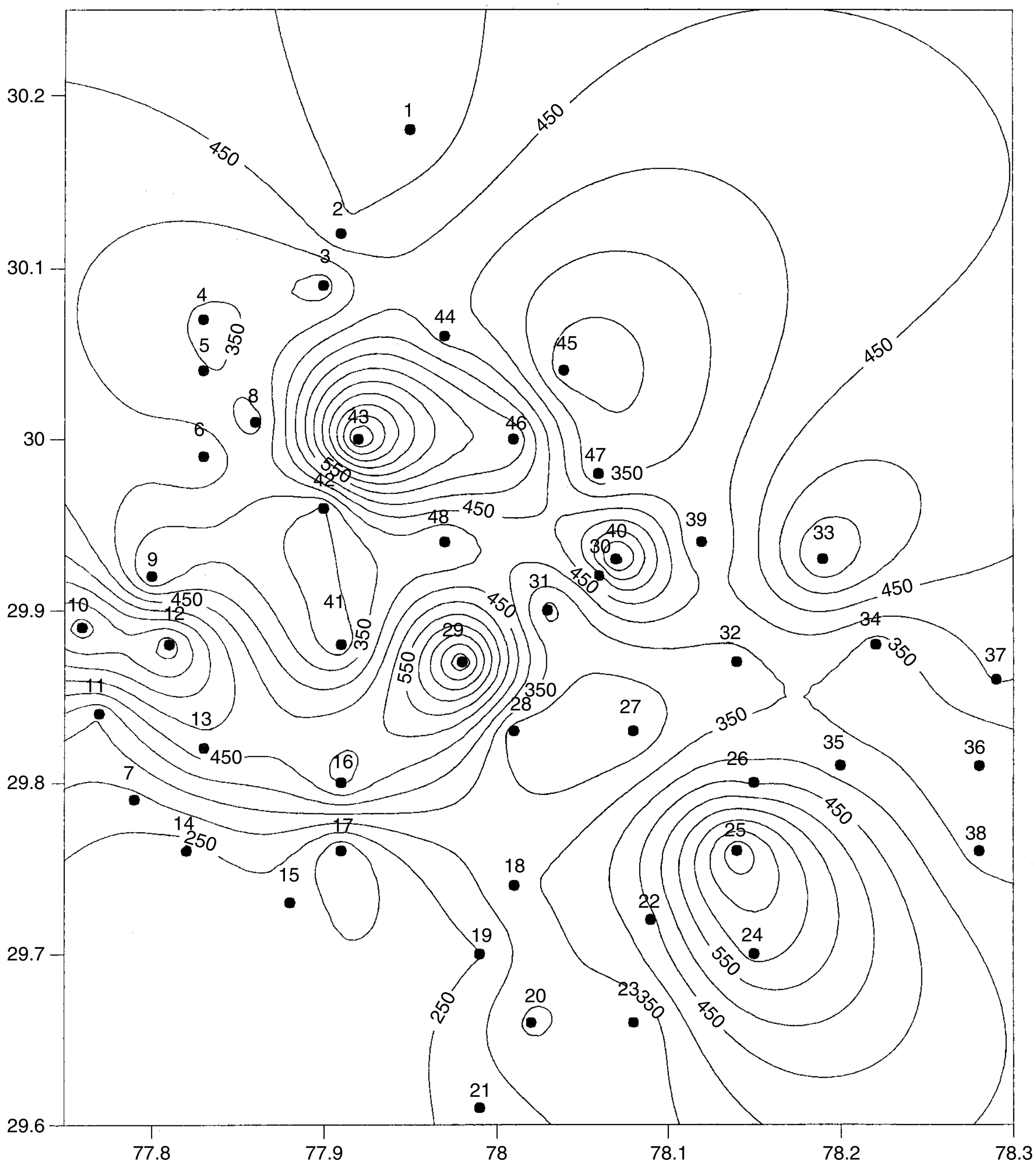

Figure 2. (**a**) Distribution of TDS in groundwater from District Hardwar (premonsoon 2002). (**b**) Distribution of TDS in groundwater from District Hardwar (postmonsoon 2002).

samples of the study area fall within the desirable limit at most places.

The concentration of sodium in the study area varied from 5.0–69 mg/L during the premonsoon season and from 4.5–71 mg/L during the postmonsoon season. The violation of BIS limits could not be ascertained for sodium as no permissible limit of sodium has been prescribed in BIS drinking water specifications. Groundwater high in sodium is not suitable for irrigation due to the sodium sensitivity of crops/plants.

The concentration of potassium in the groundwater of District Hardwar varied from 0.8–42 mg/L during the premonsoon season and from 0.2–32 mg/L during the postmonsoon season. Potassium, an essential element for humans, plants, and animals, is derived in the food chain mainly from vegetation and soil. The main sources of potassium in groundwater include rainwater, weathering of potash silicate minerals, use of potash fertilizers, and use of surface water for irrigation. It is more abundant in sedimentary rocks and commonly present in feldspar, mica, and other clay minerals. The Bureau of Indian Standards has not included potassium in drinking water standards. However, the European Economic Community has prescribed a guideline level of 10 mg/L potassium in drinking water. Per EEC criteria, about 10–15% of the samples of the study area exceeded the 10 mg/L

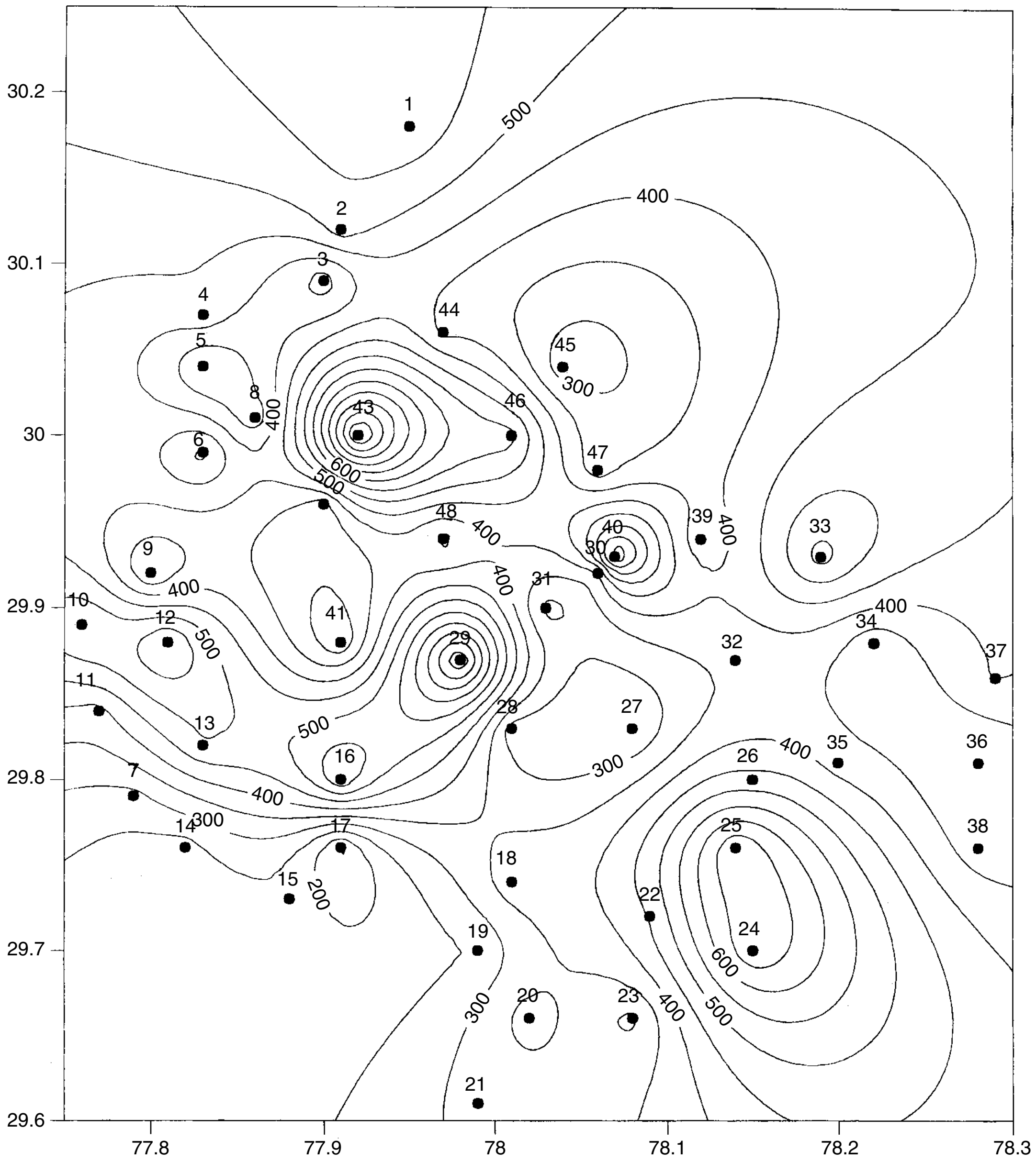

Figure 2. (*Continued*)

guideline level. Though potassium is found extensively in some igneous and sedimentary rocks, its concentration in natural waters is usually quite low because potassium minerals offer resistance to weathering and dissolution. A higher potassium content in groundwater is indicative of groundwater pollution.

The concentration of chloride in the study area is quite low and varies from 0.1–32 mg/L during the premonsoon season. An almost similar trend was observed during the postmonsoon season. The limits of chloride have been laid down primarily from taste considerations. A limit of 250 mg/L chloride has been recommended as a desirable limit for drinking water supplies (3,4). However, no adverse health effects on humans have been reported from intake of waters containing an even higher chloride content. No sample in the study area exceeded the desirable limit of 250 mg/L.

The sulfate content of groundwater generally occurs as soluble salts of calcium, magnesium, and sodium. The sulfate content changes significantly with time during infiltration of rainfall and groundwater recharge, which takes place mostly from stagnant water pools and surface runoff water collected in low-lying areas. The concentration of sulfate in the study area varied from 0.5–72 mg/L during the premonsoon season and from 0.7–62 mg/L during the postmonsoon season. It is clearly evident from the distribution maps that all the samples from District Hardwar fall within the desirable limit of 200 mg/L prescribed for drinking water supplies.

Excess nitrate content in drinking water is considered dangerous for its adverse health effects. The occurrence of high levels of nitrate in groundwater is a prominent problem in many parts of the country. The nitrate content in District Hardwar varies from 0.1–140 mg/L during the premonsoon season and from 0.1–130 mg/L during the postmonsoon season. About 95% of the samples shows nitrate content less than the desirable limit of 45 mg/L. Only one sample from Jwalapur exceeded the maximum permissible limit of 100 mg/L during both pre- and postmonsoon seasons. The nitrate distribution maps for the pre- and postmonsoon seasons are shown in Fig. 3a,b. The higher level of nitrate at Jwalapur may be attributed to improper sanitation and unhygienic conditions around the structure.

Nitrate is an effective and moderately toxic plant nutrient. A limit of 45 mg/L has been prescribed by the WHO (4) and the BIS (3) for drinking water. Its

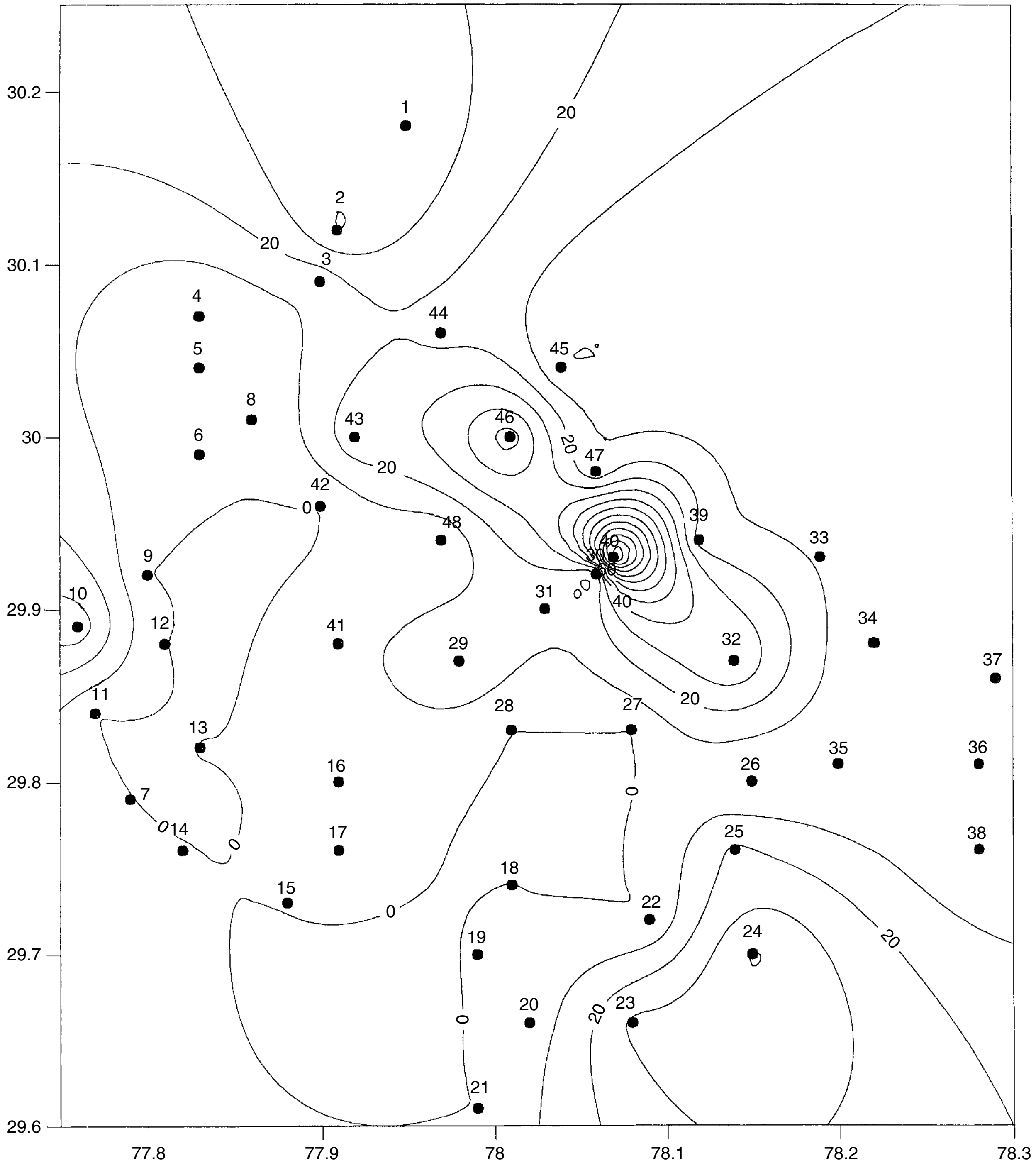

Figure 3. (**a**) Distribution of nitrate in groundwater from District Hardwar (premonsoon 2002). (**b**) Distribution of nitrate in groundwater from District Hardwar (postmonsoon 2002).

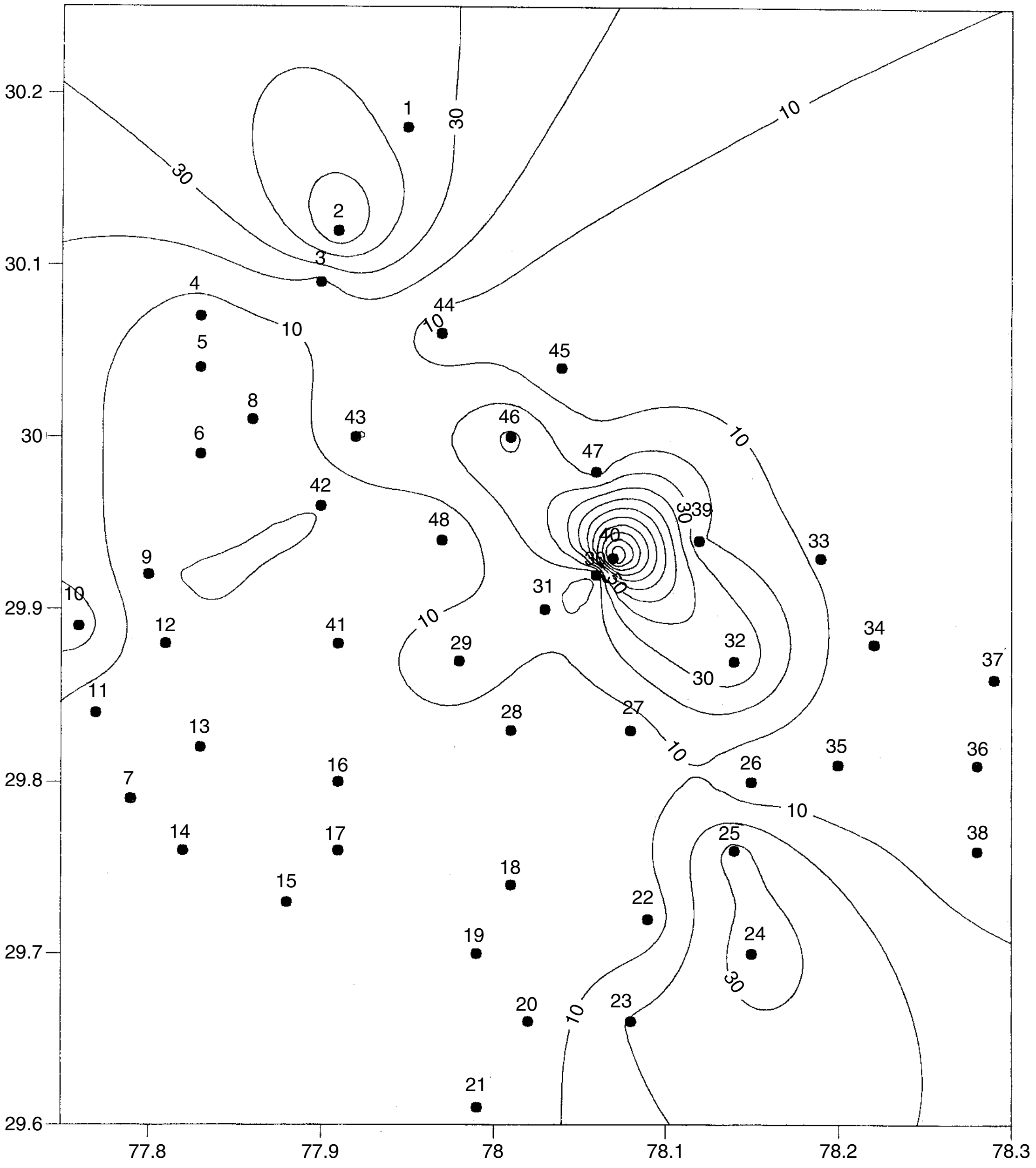

Figure 3. (*Continued*)

concentration above 45 mg/L may prove detrimented to human health. In higher concentrations, nitrate may produce a disease known as methemoglobinemia (blue babies) which generally affects bottle-fed infants. Repeated heavy doses of nitrates by ingestion may also cause cancer.

The concentration of phosphate in the study area is generally low at all locations. Phosphorous is an essential plant nutrient that is used extensively as a fertilizer. Phosphate is adsorbed or fixed as aluminium or iron phosphate in acidic soils or as calcium phosphate in alkaline or neutral soils; as a result, the concentration of phosphate in groundwater is usually low, but various chemical processes in soil strata may induce the mobility of phosphate in sub-soil and groundwater.

The fluoride content in the groundwater of District Hardwar varies from 0.01–0.94 mg/L during the premonsoon season and from 0.01–0.88 mg/L during the postmonsoon season; it lies well below the desirable limit of 1.0 mg/L in all samples. The fluoride distribution maps for pre- and postmonsoon seasons are shown in Fig. 4a,b.

The presence of fluoride in groundwater may be attributed to the localized effects of natural sources. Fluoride present in soil strata is from geological formations, such as fluorspar, and fluorapatite, and amphiboles, such as hornblende, tremolite, and mica. Weathering of igneous and sedimentary alkali silicate rocks, especially shales, contribute a major portion of fluorides to groundwaters. In addition to natural sources, considerable amounts of fluoride may be contributed by human activities. Fluoride

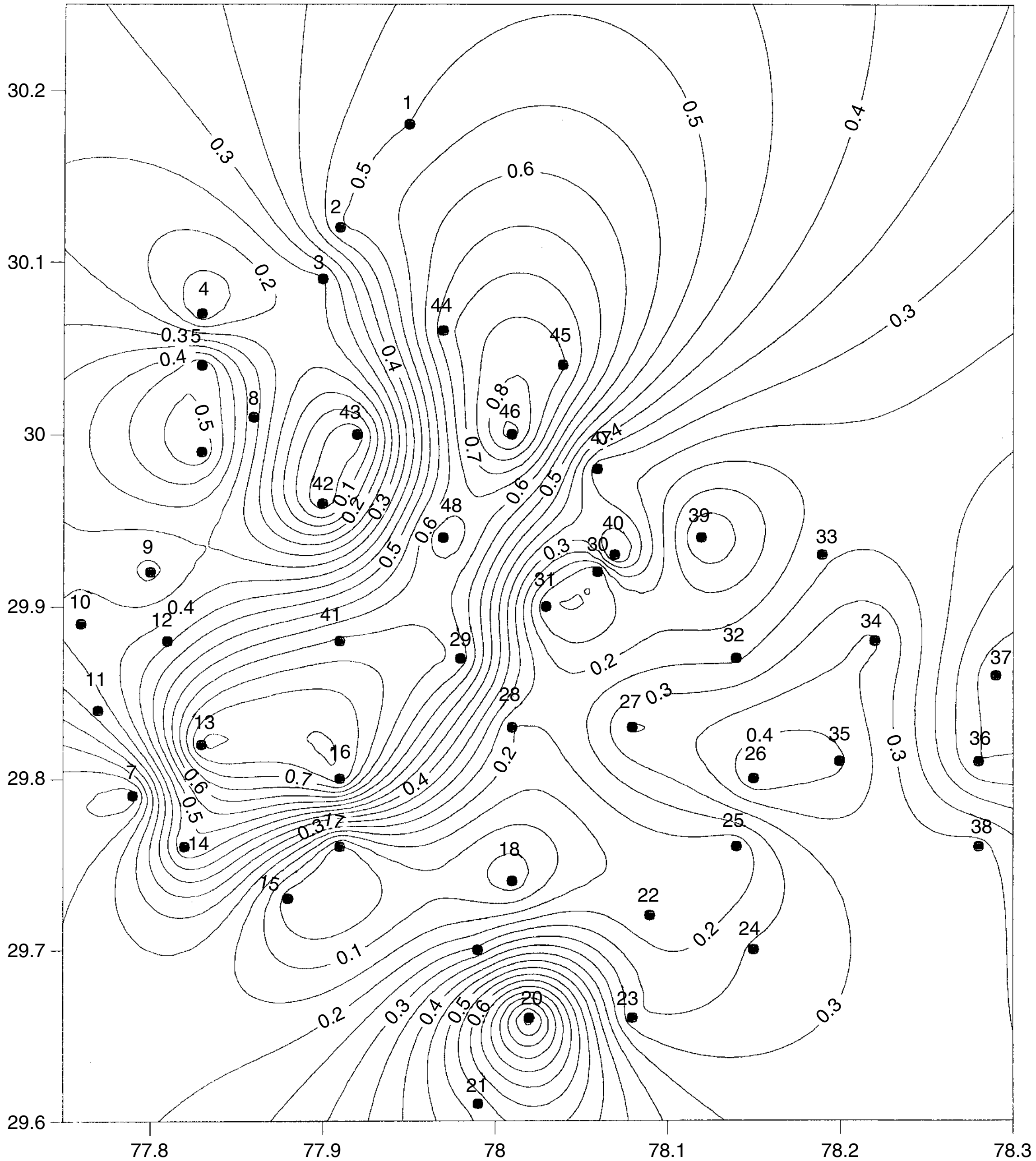

Figure 4. (**a**) Distribution of fluoride in groundwater from District Hardwar (premonsoon 2002). (**b**) Distribution of fluoride in groundwater from District Hardwar (postmonsoon 2002).

salts are commonly used in the steel, aluminum, brick, and tile industries. Fluoride containing insecticides and herbicides may be contributed through agricultural runoff. Phosphatic fertilizers, which are extensively used, often contain fluorides as impurities, and these may increase the levels of fluoride in soil. The accumulation of fluoride in soil eventually results in leaching it by percolating water, thus increasing the fluoride concentration in groundwater.

The study has clearly indicated that the concentration of total dissolved solids exceeds the desirable limit of 500 mg/L in about 25% of the samples analyzed, but the values are well within the maximum permissible limit of 2000 mg/L in all samples. The alkalinity exceeds the desirable limit of 200 mg/L in about 50% of the samples, but the levels are well within the maximum permissible limit of 600 mg/L. Total hardness exceeds the desirable limit of 300 mg/L in about 15% of the samples. The nitrate content exceeds the maximum permissible limit of 100 mg/L in only one sample (Jwalapur) of the study area. The fluoride content is well within the desirable limit in all samples analyzed.

Bacteriological Parameters

The coliform group of bacteria is the principal indicator of the suitability of water for domestic, industrial, and other uses. The density of the coliform group is the criterion for

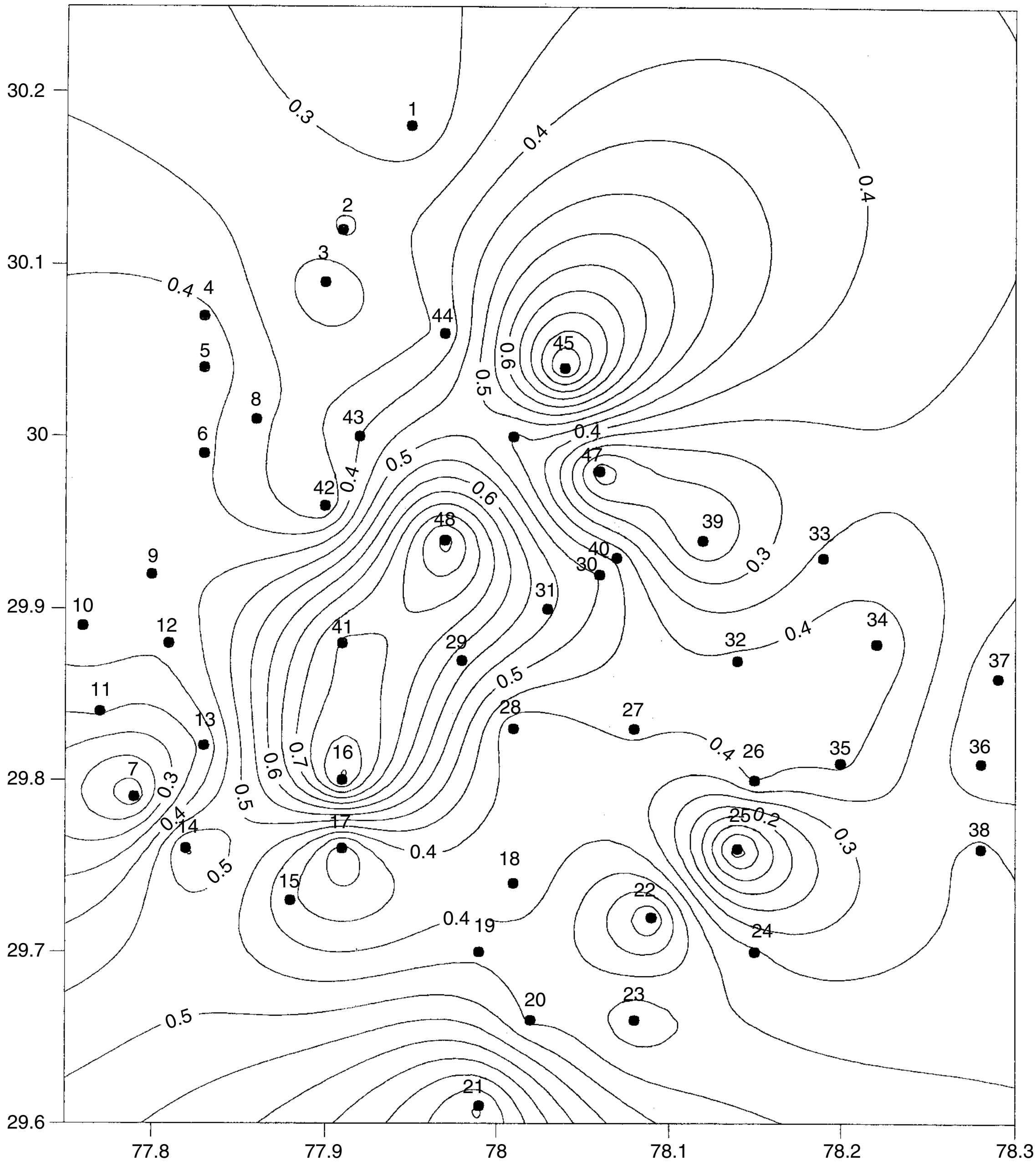

Figure 4. *(Continued)*

the degree of contamination and has been the basis for the bacteriological water quality standard. In ideal conditions, all samples taken from the distribution system should be free from coliform organisms, but in practice, this is not always attainable, and therefore, the following standard for water has been recommended (4):

— 95% of water samples should not contain any coliform organisms in 100 mL throughout the year.

— No water sample should contain *E. coli* in 100 mL water.

— No water sample should contain more than 10 coliform organisms per 100 mL.

— Coliform organisms should not be detected in 100 mL of any two consecutive water samples.

However, from bacteriological considerations, the objectives should be to reduce the coliform count to less than 10 per 100 mL, and more importantly, the absence of fecal coliform should be ensured. The presence of coliforms in water is an indicator of contamination by human or animal excrement. The presence of fecal colifirms in groundwater indicates a potential public health problem because fecal matter is a source of pathogenic bacteria and viruses. Groundwater contamination from fecal coliform bacteria is generally caused by percolation from sources of contamination (domestic sewage and septic tanks) into the aquifers and also from poor sanitation. Shallow wells are particularly susceptible to such contamination. Indiscriminate land disposal of domestic waste on the surface and improper disposal of solid waste, and leaching of wastewater from landfill areas further increase the chances of bacterial contamination in groundwater. The results of the bacteriological analysis of groundwater samples from District Hardwar are given in Table 4.

Bacteriological analysis of the groundwater samples collected from District Hardwar indicates bacterial contamination in about 40% of the samples analyzed. About 20% of the samples even exceed the permissible limit of 10 coliforms per 100 mL of sample. Inadequate maintenance of hand pumps, improper sanitation, and unhygienic conditions around structures may be responsible for bacterial contamination in the groundwater of the region and is a cause of concern. The water from such sources should be

Table 4. Bacteriological Contamination in Groundwater from District Hardwar

		Premonsoon 2002		Postmonsoon 2002	
S. No.	Location	Total Coliform per 100 mL	Fecal Coliform per 100 mL	Total Coliform per 100 mL	Fecal Coliform per 100 mL
1	Mohand	Nil	Nil	Nil	Nil
2	Banjarewala	Nil	Nil	Nil	Nil
3	Buggawala	4	Nil	Nil	Nil
4	Kheri	9	Nil	Nil	Nil
5	Dadapatti	23	Nil	Nil	Nil
6	Bahbalpur	75	Nil	4	Nil
7	Sikandarpur	7	Nil	Nil	Nil
8	Bhagwanpur	23	Nil	Nil	Nil
9	Chudiala	Nil	Nil	Nil	Nil
10	Balswa Ganj	9	Nil	Nil	Nil
11	Manakpur	14	Nil	Nil	Nil
12	Iqbalpur	28	Nil	4	Nil
13	Jharera	15	Nil	Nil	Nil
14	Sherpur	Nil	Nil	Nil	Nil
15	Narsen	4	Nil	Nil	Nil
16	Manglour	4	Nil	Nil	Nil
17	Libarheri	Nil	Nil	Nil	Nil
18	Mahesari	4	Nil	Nil	Nil
19	Sahipur	Nil	Nil	Nil	Nil
20	Khanpur	Nil	Nil	Nil	Nil
21	Chandpuri Kalan	Nil	Nil	4	Nil
22	Laksar	43	Nil	Nil	Nil
23	Kalsiya	23	Nil	15	Nil
24	Niranjanpur	4	Nil	Nil	Nil
25	Sultanpur	4	Nil	Nil	Nil
26	Shahpur	Nil	Nil	Nil	Nil
27	Pathri	Nil	Nil	Nil	Nil
28	Subashgarh	4	Nil	Nil	Nil
29	Marghubpur	Nil	Nil	Nil	Nil
30	Bahadarabad	4	Nil	Nil	Nil
31	Alipur	Nil	Nil	Nil	Nil
32	Katarpur	Nil	Nil	Nil	Nil
33	Kankhal	Nil	Nil	Nil	Nil
34	Shyampur	Nil	Nil	Nil	Nil
35	Rasiya Garh	Nil	Nil	Nil	Nil
36	Gandikhatta	Nil	Nil	Nil	Nil
37	Laldhang	Nil	Nil	Nil	Nil
38	Kottawali	Nil	Nil	Nil	Nil
39	Hardwar	Nil	Nil	Nil	Nil
40	Jwalapur	Nil	Nil	Nil	Nil
41	Roorkee	Nil	Nil	Nil	Nil
42	Gumanwala	Nil	Nil	Nil	Nil
43	Manubas	Nil	Nil	Nil	Nil
44	Bandarjud	Nil	Nil	Nil	Nil
45	Beriwala	23	Nil	Nil	Nil
46	Hazara	Nil	Nil	Nil	Nil
47	Aurangbad	Nil	Nil	Nil	Nil
48	Daulatpur	Nil	Nil	Nil	Nil

properly disinfected before being used for drinking and other domestic purposes.

Classification of Ground Water

The groundwater of District Hardwar has been classified per Chadha's diagram (5). The diagram is a somewhat modified version of the Piper trilinear diagram (6). In the Piper diagram, the milliequivalent percentages of the major cations and anions are plotted in two base triangles, and the type of water is determined on the basis of the position of the data in the respective cationic and anionic triangular fields. The plottings from the triangular fields are projected further into the central diamond field, which represents the overall character of the water. A Piper diagram allow comparisons among numerous analyses, but this type of diagram has a drawback, as all trilinear diagrams do, in that it does not portray actual ion concentration. The distribution of ions within the main field is unsystematic in hydrochemical process terms, so the diagram lacks a certain logic. This method is not very convenient when plotting a large volume of data. Nevertheless, this shortcoming does not lessen the usefulness of the Piper diagram in representing some geochemical processes.

In contrast, in Chadha's diagram, the difference in milliequivalent percentage between alkaline earths

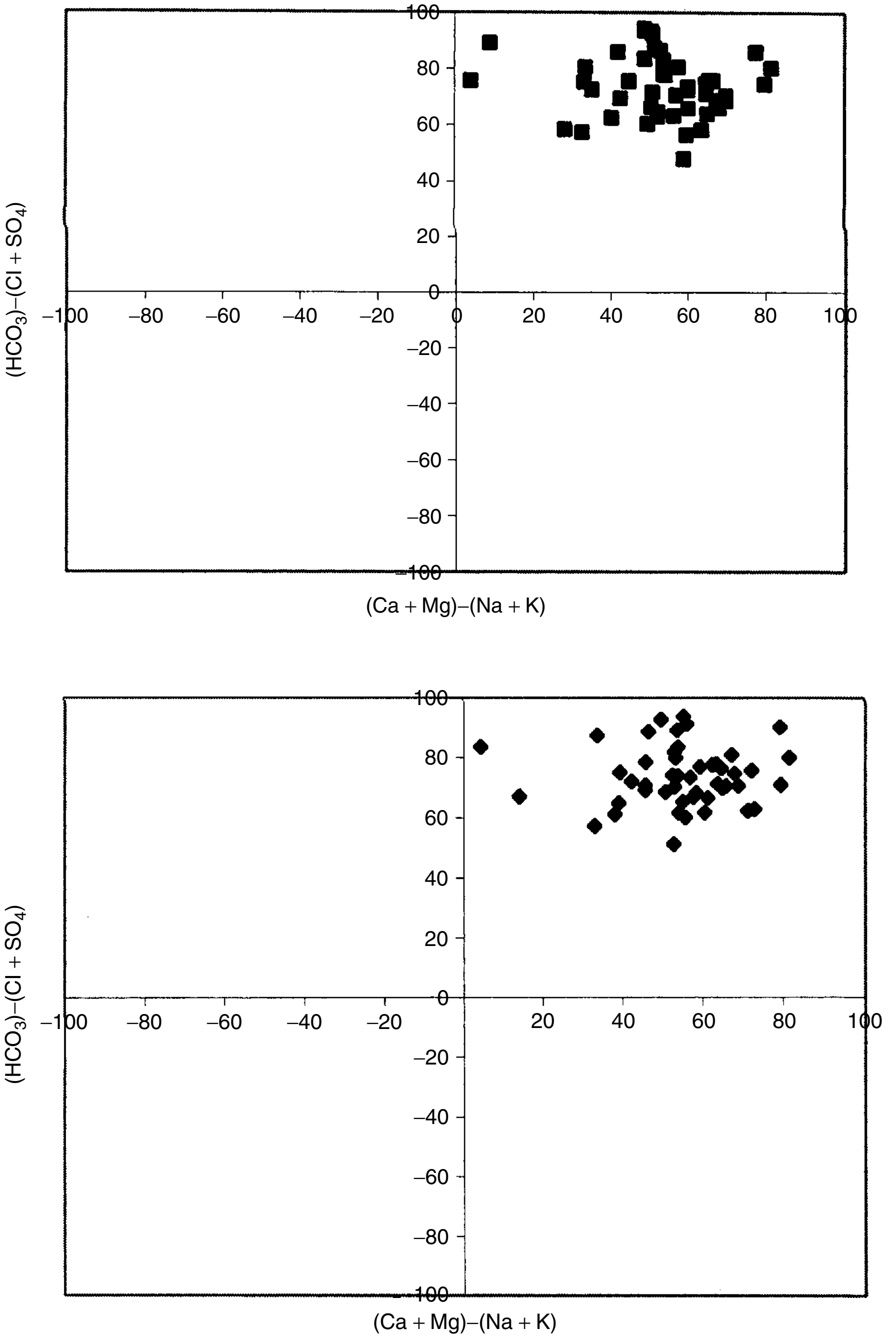

Figure 5. (**a**) Chadha's diagram showing the chemical character of groundwater in District Hardwar, Uttaranchal (premonsoon 2002). (**b**) Chadha's diagram showing the chemical character of groundwater in District Hardwar, Uttaranchal (postmonsoon 2002).

(calcium plus magnesium) and alkali metals (sodium plus potassium), expressed as a percentage of reacting values, is plotted on the x-axis, and the difference in milliequivalent percentage between weak acidic anions (carbonate plus bicarbonate) and strong acidic anions (chloride plus sulfate) is plotted on the y-axis. The resulting field of study is a square or rectangle depending upon the size of the scales chosen for the x and y coordinates. The milliequivalent percentage differences between alkaline earth and alkali metals and between weak acidic anions and strong acidic anions would plot in one of the four possible subfields of the diagram. The main advantage of this diagram is that it can be produced easily on most spreadsheet software packages.

The square or rectangular field describes the overall character of the water. The diagram has all the advantages of the diamond-shaped field of the Piper trilinear diagram and can be used to study various hydrochemical processes, such as base cation exchange, cement pollution, mixing of natural waters, sulfate reduction, saline water (end product water), and other related hydrochemical problems (5). The chemical analysis data of all samples collected from District Hardwar have been plotted on Chadha's diagram (Fig. 5a,b).

It is evident from the results that all samples of the study area fall in Group 5 ($Ca-Mg-HCO_3$ type) during both pre- and postmonsoon seasons. Chadha's diagram has all the advantages of the diamond-shaped field of the Piper trilinear diagram and can be conveniently used to study various hydrochemical processes. Another main advantage of this diagram is that it can be produced easily on most spreadsheet software packages.

CONCLUSIONS AND RECOMMENDATIONS

The groundwater quality in District Hardwar varies from place to place and with the depth of the water table. The water drawn for domestic use should be tested and analyzed to ensure the suitability of groundwater for human consumption. The groundwater abstraction sources and their surroundings should be properly maintained to ensure hygienic conditions; no sewage or polluted water should be allowed to percolate directly to a groundwater aquifer. Proper cement platforms should be constructed surrounding the ground water abstraction sources to avoid direct wellhead pollution, and the surrounding surface area should be frequently chlorinated by using bleaching powder. The hand pumps and wells, which have been identified as of suspect water quality, should be painted red to indicate and warn the public that the water drawn from the source is not fit for human consumption. In the absence of an alternate safe source of water, the water with excessive undesirable constituents must be treated by a specific treatment process before it is used for human consumption. The untreated sewage and sewerage flowing in various open drains are one of the causes of groundwater quality deterioration. A proper underground sewage system must be laid in inhabited areas, and the untreated sewage should not be allowed to flow in open drains. A proper system of collection and transportation of domestic waste should be developed. Landfill site(s) should be identified and must be scientifically designed. A mass awareness should be generated about water quality, its effect on human health, and the responsibilities of the public to safeguard water resources.

BIBLIOGRAPHY

1. Jain, C.K. and Bhatia, K.K.S. (1988). *Physico-Chemical Analysis of Water and Wastewater, User's Manual, UM-26*. National Institute of Hydrology, Roorkee, India.
2. American Public Health Association. (1992). *Standard Methods for the Examination of Water and Waste Waters*. 18th Edn. American Public Health Association, Washington, DC.
3. Bureau of Indian Standards. (1991). *Specifications for Drinking Water*. IS:10500:1991. Bureau of Indian Standards, New Delhi, India.
4. World Health Organization. (1996). *Guidelines for Drinking Water*. Vol. 2, Recommendations. World Health Organization, Geneva, Switzerland.
5. Chadha, D.K. (1999). A proposed new diagram for geochemical classification of natural waters and interpretation of chemical data. *Hydrogeol. J.* **7**(5): 431–439.
6. Piper, A.M. (1944). A graphical procedure in the geochemical interpretation of water analysis. *Trans. Am. Geophys. Union* **25**: 914–923.

IRRIGATION WATER QUALITY IN DISTRICT HARDWAR, UTTARANCHAL, INDIA

Chakresh K. Jain
National Institute of Hydrology
Roorkee, India

The groundwater quality of District Hardwar in the state of Uttaranchal (India) was assessed to determine the suitability of groundwater for irrigation. Two sets of 48 groundwater samples from shallow and deep aquifers were collected during pre- and postmonsoon seasons in the year 2002, and the suitability of the groundwater for irrigation was evaluated based on salinity, sodium adsorption ration (SAR), residual sodium carbonate (RSC), and boron content. The values of SAR ranged from 0.24–1.75 during the premonsoon season and from 0.22–1.92 during the postmonsoon season, respectively. Values of the SAR indicate that the majority of samples of the study area falls under the category of low sodium hazard, indicating no risk of sodification. In general, the groundwater of District Hardwar is safe for irrigation. According to the U.S. Salinity Laboratory classification of irrigation water, the majority of the samples (70%) falls under water type C2-S1 followed by the C3-S1 type.

INTRODUCTION

The intensive use of natural resources and the large production of wastes in modern society often pose a threat to groundwater quality and have already resulted in many incidents of groundwater contamination. Pollutants are (calcium plus magnesium) and alkali metals (sodium plus potassium), expressed as a percentage of reacting values, is plotted on the *x*-axis, and the difference in milliequivalent percentage between weak acidic anions (carbonate plus bicarbonate) and strong acidic anions (chloride plus sulfate) is plotted on the *y*-axis. The resulting field of study is a square or rectangle depending upon the size of the scales chosen for the *x* and *y* coordinates. The milliequivalent percentage differences between alkaline earth and alkali metals and between weak acidic anions and strong acidic anions would plot in one of the four possible subfields of the diagram. The main advantage of this diagram is that it can be produced easily on most spreadsheet software packages.

The square or rectangular field describes the overall character of the water. The diagram has all the advantages of the diamond-shaped field of the Piper trilinear diagram and can be used to study various hydrochemical processes, such as base cation exchange, cement pollution, mixing of natural waters, sulfate reduction, saline water (end product water), and other related hydrochemical problems (5). The chemical analysis data of all samples collected from District Hardwar have been plotted on Chadha's diagram (Fig. 5a,b).

It is evident from the results that all samples of the study area fall in Group 5 (Ca–Mg–HCO_3 type) during both pre- and postmonsoon seasons. Chadha's diagram has all the advantages of the diamond-shaped field of the Piper trilinear diagram and can be conveniently used to study various hydrochemical processes. Another main advantage of this diagram is that it can be produced easily on most spreadsheet software packages.

CONCLUSIONS AND RECOMMENDATIONS

The groundwater quality in District Hardwar varies from place to place and with the depth of the water table. The water drawn for domestic use should be tested and analyzed to ensure the suitability of groundwater for human consumption. The groundwater abstraction sources and their surroundings should be properly maintained to ensure hygienic conditions; no sewage or polluted water should be allowed to percolate directly to a groundwater aquifer. Proper cement platforms should be constructed surrounding the ground water abstraction sources to avoid direct wellhead pollution, and the surrounding surface area should be frequently chlorinated by using bleaching powder. The hand pumps and wells, which have been identified as of suspect water quality, should be painted red to indicate and warn the public that the water drawn from the source is not fit for human consumption. In the absence of an alternate safe source of water, the water with excessive undesirable constituents must be treated by a specific treatment process before it is used for human consumption. The untreated sewage and sewerage flowing in various open drains are one of the causes of groundwater quality deterioration. A proper underground sewage system must be laid in inhabited areas, and the untreated sewage should not be allowed to flow in open drains. A proper system of collection and transportation of domestic waste should be developed. Landfill site(s) should be identified and must be scientifically designed. A mass awareness should be generated about water quality, its effect on human health, and the responsibilities of the public to safeguard water resources.

BIBLIOGRAPHY

1. Jain, C.K. and Bhatia, K.K.S. (1988). *Physico-Chemical Analysis of Water and Wastewater, User's Manual, UM-26*. National Institute of Hydrology, Roorkee, India.
2. American Public Health Association. (1992). *Standard Methods for the Examination of Water and Waste Waters*. 18th Edn. American Public Health Association, Washington, DC.
3. Bureau of Indian Standards. (1991). *Specifications for Drinking Water*. IS:10500:1991. Bureau of Indian Standards, New Delhi, India.
4. World Health Organization. (1996). *Guidelines for Drinking Water*. Vol. 2, Recommendations. World Health Organization, Geneva, Switzerland.
5. Chadha, D.K. (1999). A proposed new diagram for geochemical classification of natural waters and interpretation of chemical data. *Hydrogeol. J.* **7**(5): 431–439.
6. Piper, A.M. (1944). A graphical procedure in the geochemical interpretation of water analysis. *Trans. Am. Geophys. Union* **25**: 914–923.

IRRIGATION WATER QUALITY IN DISTRICT HARDWAR, UTTARANCHAL, INDIA

Chakresh K. Jain
National Institute of Hydrology
Roorkee, India

The groundwater quality of District Hardwar in the state of Uttaranchal (India) was assessed to determine the suitability of groundwater for irrigation. Two sets of 48 groundwater samples from shallow and deep aquifers were collected during pre- and postmonsoon seasons in the year 2002, and the suitability of the groundwater for irrigation was evaluated based on salinity, sodium adsorption ration (SAR), residual sodium carbonate (RSC), and boron content. The values of SAR ranged from 0.24–1.75 during the premonsoon season and from 0.22–1.92 during the postmonsoon season, respectively. Values of the SAR indicate that the majority of samples of the study area falls under the category of low sodium hazard, indicating no risk of sodification. In general, the groundwater of District Hardwar is safe for irrigation. According to the U.S. Salinity Laboratory classification of irrigation water, the majority of the samples (70%) falls under water type C2-S1 followed by the C3-S1 type.

INTRODUCTION

The intensive use of natural resources and the large production of wastes in modern society often pose a threat to groundwater quality and have already resulted in many incidents of groundwater contamination. Pollutants are being added to groundwater systems by human activities and natural processes. Solid waste from industrial units is being dumped near factories, is subjected to reaction with percolating rainwater, and reaches the groundwater level. The percolating water picks up a large amount of dissolved constituents, reaches the aquifer system, and contaminates the groundwater.

Groundwater plays an important role in agriculture for both watering crops and for irrigating dry season crops. It is estimated that about 45% of the irrigation water requirement is met from groundwater sources. Indiscriminate use of fertilizers and pesticides in agricultural fields has become common practice in India and has already resulted in very high concentrations of these constituents in groundwater, which in the long term, may create the greatest hazard to groundwater.

Many problems originate from inefficient management of water for agricultural use, especially when it carries high salts. The problem of groundwater contamination in several parts of the country has become so acute that unless urgent steps for detailed identification and abatement are taken, extensive groundwater resources may be damaged. Keeping in view the severity of the problem, the groundwater quality of District Hardwar in the newly created state of Uttaranchal (India) was studied to examine the suitability of groundwater for irrigation.

STUDY AREA

District Hardwar is part of the Indo-Gangetic plains and lies between latitude 29°30′ to 30°20′ N and longitude 77°40′ to 78°25′ E in the state of Uttaranchal (Fig. 1). It is the largest district (populationwise) of Uttaranchal State and occupies an area of about 2,360 km^2. Per the 2001 census, the population of the District Hardwar is 14,44,213, and the population density is 612 per km^2.

Physiographically, the area is generally flat except for the Siwalik Hills in the north and north east. The area is devoid of relief features of any prominence except for deep gorges cut by drains and rivers flowing through the area. The area is bounded by River Yamuna in the west and River Ganga in the east. The climate of the area is characterized by a moderate type of subtropical monsoon. The average annual rainfall in the region is about 1000 mm; the major part is received during the monsoon period. The major land use is for agriculture, and there is no effective forest cover. The soils of the area are loam to silty loam and are free from carbonates. The most common groundwater use is achieved by hand pumps and tube wells. Based on lithologic logs and water table fluctuation data, two types of aquifers have been delineated in the area. The upper is the shallow unconfined aquifer which generally extends to depths around 25m. The deeper one is confined to semiconfined and located at

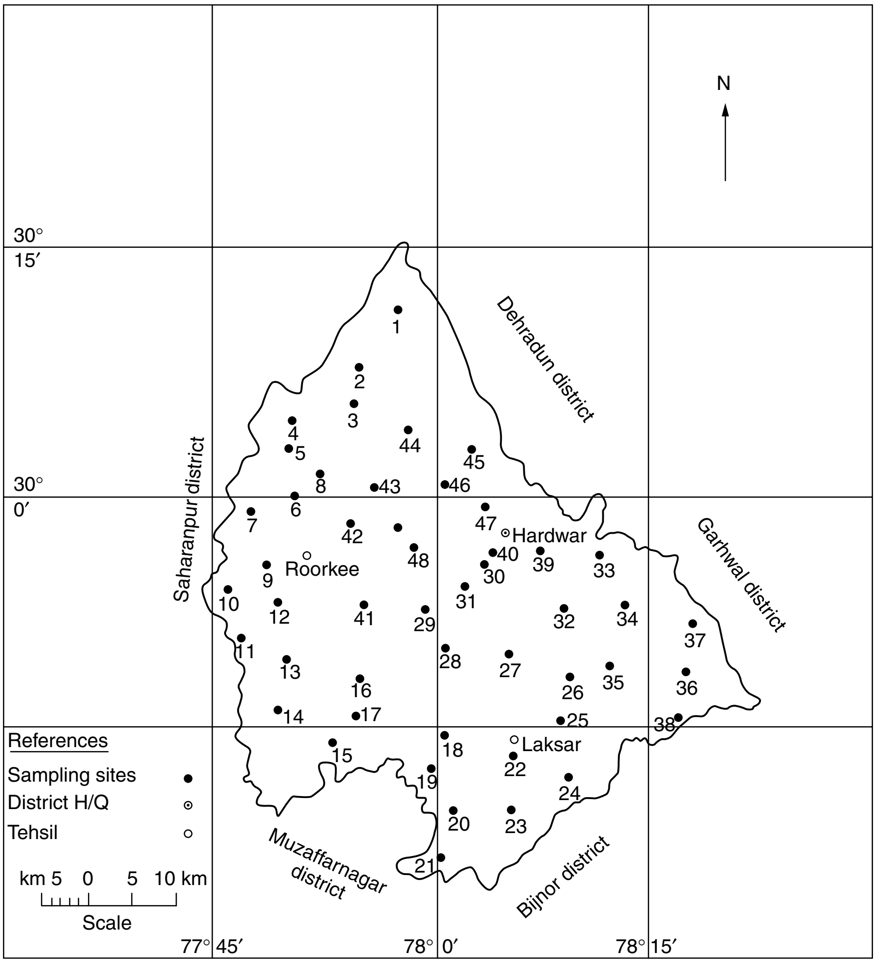

Figure 1. Study area showing location of sampling sites in District Hardwar.

a depth about 25 to 150m below ground level, separated by three to four aquifers at average depths of 25–55, 65–90, and 120–150m. Water table contours in the area indicate a southward trend of groundwater flow in both unconfined and confined aquifers.

EXPERIMENTAL METHODOLOGY

Two sets of 48 groundwater samples from District Hardwar were collected during pre- and postmonsoon seasons in the year 2002 from various abstraction sources at various depths. They were preserved by adding an appropriate reagent (1,2). The hand pumps and tube wells were continuously pumped prior to sampling to ensure that groundwater sampled was representative of the groundwater aquifer. The details of sampling locations and source and depth distribution are given in Tables 1 and 2, respectively. The physicochemical analysis was performed following standard methods (1,2).

RESULTS AND DISCUSSION

The hydrochemical data for the two sets of samples collected during pre- and postmonsoon seasons are presented in Table 3. The suitability of groundwater for domestic use was discussed in an earlier report (3). The quality of water plays an important role in irrigated agriculture. Many problems originate from inefficient management of water for agricultural use, especially when it carries high salts. The concentration and composition of dissolved constituents in water determine its quality for irrigation. The quality of water is an important consideration in any appraisal of salinity or alkaline conditions in an irrigated area. Good quality water has can provide maximum yield under good soil and water management practices. The most important characteristics of water which determine suitability of ground water for irrigation are

1. Salinity
2. Relative proportion of sodium to other cations (SAR)
3. Residual sodium carbonate (RSC)
4. Boron content

The safe limits of electrical conductivity for crops of different degrees of salt tolerance under varying soil textures and drainage conditions are given in Table 4.

Table 2. Source and Depth Distribution of Sampling Sites in District Hardwar

Source Structure	Depth Range <0–20 m	20–40 m	>40 m	Total Number
Hand pumps	6, 18, 20, 21, 22, 23, 24, 25, 27, 28, 30, 31, 32, 36, 41, 43, 44, 45, 46, 47	3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 26, 29, 33, 34, 35, 38, 39, 40, 42, 48	37	46
Tube wells	—	—	—	—
Open wells	1,2	—	—	2
Total	22	25	1	48

Table 1. Description of Groundwater Sampling Locations in District Hardwar

S.No.	Location	Source	Depth, m	S.No.	Location	Source	Depth, m
1	Mohand	OW	10	25	Sultanpur	HP	15
2	Banjarewala	OW	10	26	Shahpur	HP	25
3	Buggawala	HP	30	27	Pathri	HP	15
4	Kheri	HP	38	28	Subashgarh	HP	20
5	Dadapatti	HP	38	29	Marghubpur	HP	30
6	Bahbalpur	HP	15	30	Bahadarabad	HP	20
7	Sikandarpur	HP	25	31	Alipur	HP	20
8	Bhagwanpur	HP	30	32	Katarpur	HP	10
9	Chudiala	HP	40	33	Kankhal	HP	25
10	Balswa Ganj	HP	25	34	Shyampur	HP	30
11	Manakpur	HP	40	35	Rasiya Garh	HP	40
12	Iqbalpur	HP	30	36	Gandikhatta	HP	10
13	Jharera	HP	30	37	Laldhang	HP	90
14	Sherpur	HP	30	38	Kottawali	HP	40
15	Narsen	HP	30	39	Hardwar	HP	30
16	Manglour	HP	35	40	Jwalapur	HP	30
17	Libarheri	HP	35	41	Roorkee	HP	10
18	Mahesari	HP	20	42	Gumanwala	HP	30
19	Sahipur	HP	35	43	Manubas	HP	10
20	Khanpur	HP	15	44	Bandarjud	HP	15
21	Chandpuri Kalan	HP	15	45	Beriwala	HP	5
22	Laksar	HP	15	46	Hazara	HP	10
23	Kalsiya	HP	10	47	Aurangbad	HP	10
24	Niranjanpur	HP	15	48	Daulatpur	HP	35

OW: open well; HP: hand pump.

Table 3. Hydrochemical Data for Groundwater Samples from District Hardwar[a]

Characteristics	Min	Max	Average
pH	6.22 (6.70)	7.58 (7.70)	6.87 (7.10)
Conductivity, μS/cm	233 (221)	1440 (1442)	642 (647)
TDS, mg/L	149 (141)	922 (923)	411 (414)
Alkalinity, mg/L	78 (71)	460 (482)	213 (215)
Hardness, mg/L	81 (80)	464 (427)	209 (212)
Chloride, mg/L	0.1 (1.2)	32 (35)	9.0 (9.0)
Sulfate, mg/L	0.5 (0.7)	72 (62)	22 (21)
Nitrate, mg/L	0.1 (0.1)	140 (130)	13 (13)
Phosphate, mg/L	0.01 (0.02)	0.41 (1.60)	0.03 (0.25)
Fluoride, mg/L	0.01 (0.01)	0.94 (0.88)	0.36 (0.42)
Sodium, mg/L	5.0 (4.5)	69 (71)	26 (25)
Potassium, mg/L	0.8 (0.2)	42 (32)	7.0 (7.4)
Calcium, mg/L	22 (24)	140 (135)	55 (56)
Magnesium, mg/L	5.0 (5.0)	36 (38)	18 (17)
Boron, mg/L	0.12 (0.12)	0.92 (0.87)	0.50 (0.49)

[a] Values in parentheses represent postmonsoon data.

The quality of water is commonly expressed by classes of relative suitability for irrigation with reference to salinity levels. The recommended classification with respect to electrical conductivity, sodium content, sodium absorption ratio (SAR), and residual sodium carbonate (RSC) are given in Table 5. The values of sodium percentage (% Na), SAR, and RSC in the groundwater of District Hardwar are given in Table 6.

Salinity

Salinity is broadly related to total dissolved solids (TDS) and electrical conductivity (EC). High concentrations of TDS and electrical conductivity in irrigation water may increase the soil salinity, which affects the salt intake of a plant. The salts in the water affect the growth of plants directly and also affect the soil structure, permeability, and aeration, which indirectly affect plant growth. Soil water passes into a plant through the root zone by osmotic pressure. As the dissolved solid content of the soil water in the root zone increases, it is difficult for the plant to overcome the osmotic pressure and the plant root membranes are able to assimilate water and nutrients. Thus, the dissolved solids content of the residual water in the root zone also has to be maintained within limits by proper leaching. Negative effects are visible in plants by stunted growth, low yield, discoloration, and even leaf burns at margin or top. The electrical conductivity values for water in District Hardwar are well within the prescribed limits of 1500 μS/cm and are therefore safe for irrigation.

Relative Proportion of Sodium to Other Cations

A high salt concentration in water leads to saline soil, and high sodium leads to development of an alkaline soil. The sodium or alkaline hazard in using water for irrigation is determined by the absolute and relative concentration of cations, expressed in terms of the sodium adsorption ratio (SAR). If the proportion of sodium is high, the alkaline hazard is high, and conversely, if calcium and magnesium predominate, the hazard is less. There is a significant relationship between SAR values of irrigation water and the extent to which sodium is absorbed by the soil. If water used for irrigation is high in sodium and low in calcium, the cation-exchange complex may become saturated with sodium. This can destroy the soil structure owing to dispersion of the clay particles. A simple method of evaluating the danger of high-sodium water is the sodium adsorption ratio (SAR) (4):

$$\mathrm{SAR} = \frac{\mathrm{Na}^+}{\sqrt{(\mathrm{Ca}^{2+} + \mathrm{Mg}^{2+})/2}}$$

The sodium percentage is calculated from

$$\%\mathrm{Na} = \frac{\mathrm{Na}^+ + \mathrm{K}^+}{\mathrm{Ca}^{2+} + \mathrm{Mg}^{2+} + \mathrm{Na}^+ + \mathrm{K}^+} \times 100$$

where all ionic concentrations are expressed in milliequivalents per liter.

Table 5. Guidelines for Evaluation of Irrigation Water Quality

Water Class	Sodium (Na), %	Electrical Conductivity, μS/cm	SAR	RSC, meq/L
Excellent	<20	<250	<10	<1.25
Good	20–40	250–750	10–18	1.25–2.0
Medium	40–60	750–2250	18–26	2.0–2.5
Bad	60–80	2250–4000	>26	2.5–3.0
Very bad	>80	>4000	>26	>3.0

Table 4. Safe Limits of Electrical Conductivity for Irrigation Water

S. No.	Nature of Soil	Crop Growth	Upper Permissible Safe Limit of EC in Water, μS/cm
1.	Deep black soil and alluvial soil whose clay content is more than 30% in soils that are fairly to moderately well drained	Semitolerant	1500
		Tolerant	2000
2.	Textured soil of clay content of 20–30% in soils that are well drained internally and have a good surface drainage system	Semitolerant	2000
		Tolerant	4000
3.	Medium textured soils whose clay is 10–20% internally very well drained and has a good surface drainage system	Semitolerant	4000
		Tolerant	6000
4.	Lightly textured soils whose clay is less than 10% of soil that has an excellent internal and surface drainage system	Semitolerant	6000
		Tolerant	8000

Table 6. SAR, % Na, and RSC Values in the Groundwater of District Hardwar

		Premonsoon			Postmonsoon		
S.No.	Location	SAR	% Na	RSC	SAR	% Na	RSC
1	Mohand	1.02	27.4	0.40	0.97	27.1	−0.05
2	Banjarewala	0.81	22.7	−0.32	0.49	16.4	−0.62
3	Buggawala	0.31	11.1	−0.18	0.30	10.3	−0.05
4	Kheri	0.56	17.5	−0.17	0.67	18.3	−0.03
5	Dadapatti	0.86	25.4	0.77	0.76	25.2	0.60
6	Bahbalpur	0.89	23.4	0.61	0.96	23.5	0.47
7	Sikandarpur	0.72	24.5	0.48	0.66	23.2	0.51
8	Bhagwanpur	0.79	25.3	0.67	0.68	22.4	0.49
9	Chudiala	0.72	24.4	0.60	0.64	21.9	0.42
10	Balswa Ganj	0.90	25.0	−0.27	0.71	22.5	−0.52
11	Manakpur	0.69	23.0	0.54	0.73	23.3	0.52
12	Iqbalpur	1.04	23.7	0.48	0.97	23.4	0.55
13	Jharera	0.77	21.3	−0.10	0.76	20.8	−0.33
14	Sherpur	0.54	23.0	0.28	0.51	21.5	0.07
15	Narsen	0.50	21.0	0.14	0.56	23.0	0.26
16	Manglour	0.80	21.6	−0.38	0.90	22.1	−0.35
17	Libarheri	0.24	14.9	−0.06	0.22	13.5	−0.19
18	Mahesari	1.11	33.1	0.69	1.18	33.1	1.01
19	Sahipur	0.74	28.9	0.59	0.71	27.1	0.48
20	Khanpur	1.75	45.5	1.48	1.92	47.8	1.54
21	Chandpuri Kalan	1.60	47.9	1.11	1.51	42.9	0.93
22	Laksar	1.30	33.2	0.74	1.19	30.3	0.66
23	Kalsiya	0.71	24.4	−0.17	0.64	23.7	−0.18
24	Niranjanpur	1.28	33.5	0.01	1.39	33.5	0.26
25	Sultanpur	1.31	35.7	1.04	1.14	30.9	0.11
26	Shahpur	0.46	19.8	0.14	0.44	18.1	0.13
27	Pathri	0.25	9.9	−0.36	0.24	10.2	−0.45
28	Subashgarh	0.45	17.3	−0.37	0.42	17.1	−0.17
29	Marghubpur	1.53	29.7	0.12	1.57	30.4	0.18
30	Bahadarabad	0.54	14.9	−0.70	0.61	20.3	0.17
31	Alipur	0.49	17.0	−0.16	0.53	17.7	−0.07
32	Katarpur	0.42	17.4	−0.69	0.39	16.0	−0.77
33	Kankhal	0.63	20.0	−0.91	0.62	19.7	−0.40
34	Shyampur	0.59	19.8	0.07	0.50	17.5	−0.06
35	Rasiya Garh	0.52	16.1	−0.11	0.51	15.5	−0.17
36	Gandikhatta	0.54	18.0	−0.35	0.58	19.3	−0.23
37	Laldhang	0.54	15.7	−0.55	0.48	14.3	−0.65
38	Kottawali	0.48	16.1	−0.07	0.41	13.9	−0.31
39	Hardwar	0.54	20.3	−0.46	0.64	23.6	−0.36
40	Jwalapur	1.05	23.7	−1.76	1.07	22.9	−2.11
41	Roorkee	0.46	16.6	−0.19	0.48	18.8	0.05
42	Gumanwala	0.58	19.7	−0.15	0.72	24.6	0.31
43	Manubas	1.09	24.5	−0.10	1.37	28.8	1.09
44	Bandarjud	0.30	9.1	−0.45	0.28	9.2	−0.40
45	Beriwala	0.72	25.4	0.46	0.66	23.1	0.25
46	Hazara	1.03	28.5	0.15	0.62	21.1	−0.62
47	Aurangbad	0.99	32.2	0.32	0.75	27.1	0.24
48	Daulatpur	0.79	24.10	0.44	0.88	26.7	0.64

Calculation of SAR for a given water provides a useful index of the sodium hazard of that water for soils and crops. A low SAR (2 to 10) indicates little danger from sodium; medium hazards are between 7 and 18, high hazards between 11 and 26, and very high hazards above that. The lower the ionic strength of the solution, the greater the sodium hazards for a given SAR (4).

The values of SAR in the groundwater of District Hardwar vary from 0.24–1.75 during the premonsoon season and from 0.22–1.92 during the postmonsoon season. As evident from the SAR values, the groundwater of the study area falls under the category of low sodium hazard, which reveals that the groundwater of the study area is free from any sodium hazard. The sodium percentage in the study area varies from 9.1–47.9% during the premonsoon season and from 9.2–47.8% during the postmonsoon season. All samples are well within the permissible limit for irrigation water and are free of any sodium hazard.

Residual Sodium Carbonate

In addition to total dissolved solids, the relative abundance of sodium with respect to alkaline earths and boron and the quantity of bicarbonate and carbonate in excess of alkaline earths also influence the suitability of water for irrigation. This excess is denoted residual sodium carbonate (RSC) and is determined by the following formula:

$$\mathrm{RSC} = (\mathrm{HCO_3}^- + \mathrm{CO_3}^{2-}) - (\mathrm{Ca}^{2+} + \mathrm{Mg}^{2+})$$

where all ionic concentrations are expressed in meq/L. Groundwater containing high concentrations of carbonate and bicarbonate ions tends to precipitate calcium and magnesium as carbonate. As a result, the relative proportion of sodium increases and is fixed in the soil thereby decreasing soil permeability. If the RSC exceeds 2.5 meq/L, the water is generally unsuitable for irrigation. Excessive RSC causes the soil structure to deteriorate, as it restricts the water and air movement through soil. If the value is between 1.25 and 2.5, the water is of marginal quality; values less than 1.25 meq/L indicate that the water is safe for irrigation. During the present study, the RSC values clearly indicate that the groundwater of District Hardwar does not have any residual sodium carbonate hazard.

Boron

Boron is essential to the normal growth of all plants, but the concentration required is very small and if exceeded may cause injury. Plant species vary in boron requirement and in tolerance to excess boron, so that concentrations necessary for the growth of plants that have high boron requirements may be toxic to plants sensitive to boron. Although boron is an essential nutrient for plant growth, generally it becomes toxic beyond 2 mg/L in irrigation water for most field crops. It does not affect the physical and chemical properties of the soil, but at high concentrations, it affects the metabolic activities of the plant. During the present study, the boron content in the groundwater clearly indicates that the groundwater is safe for irrigation.

U.S. Salinity Laboratory Classification

The U.S. Salinity Laboratory classification (5) is used to study the suitability of groundwater for irrigation. In classifying irrigation waters, it is assumed that the water will be used under average conditions with respect to soil texture, infiltration rate, drainage, quantity of water used, climate, and salt tolerance of the crop. Sodium concentration is an important criterion in irrigation-water

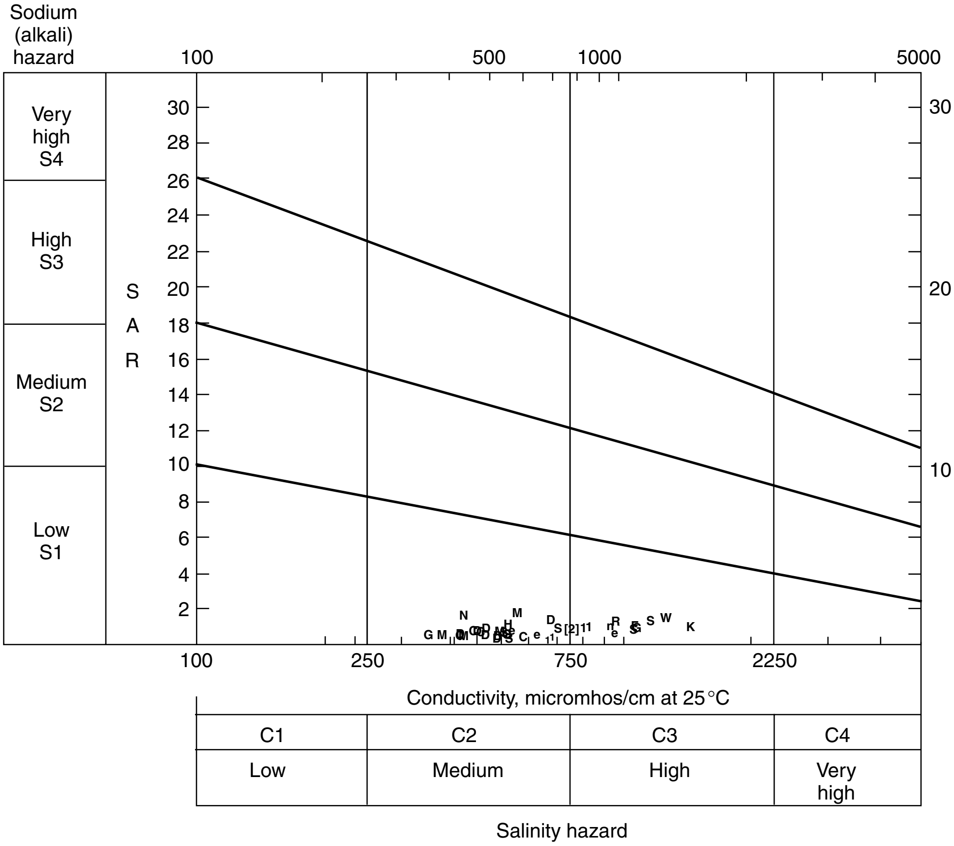

Figure 2. U.S. Salinity Laboratory classification, District Hardwar (pre-monsoon 2002).

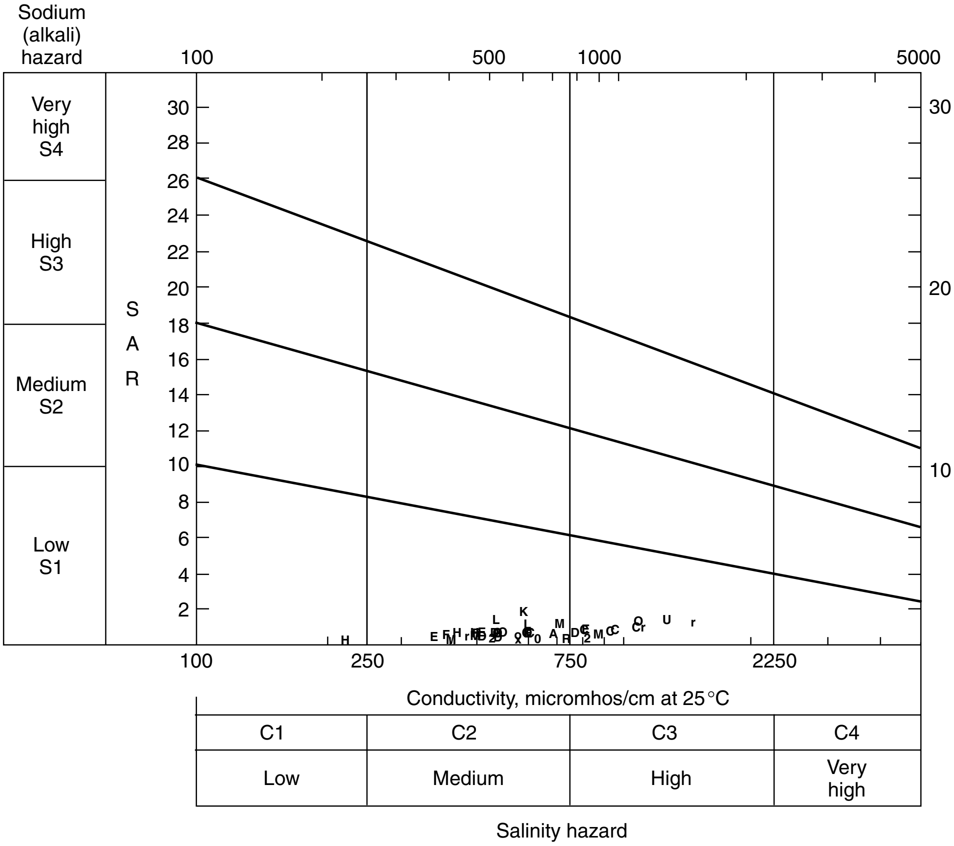

Figure 3. U.S. Salinity Laboratory classification, District Hardwar (postmonsoon 2002).

Table 7. Summarized Results of U.S. Salinity Laboratory Classification

	Sample Numbers	
Classification/Type	Premonsoon 2002	Postmonsoon 2002
C1-S1	17	17
C2-S1	3, 4, 5, 6, 7, 8, 9, 11, 14, 15, 18, 19, 20, 21, 22, 23, 26, 27, 28, 31, 32, 34, 35, 36, 37, 39, 40, 41, 42, 44, 45, 47, 48	3, 4, 5, 7, 8, 9, 11, 14, 15, 18, 19, 20, 21, 22, 23, 26, 27, 28, 30, 31, 32, 34, 35, 36, 37, 38, 39, 41, 42, 44, 45, 47, 48
C3-S1	1, 2, 10, 12, 13, 16, 24, 25, 29, 30, 33, 38, 43, 46	1, 2, 6, 10, 12, 13, 16, 24, 25, 29, 33, 40, 43, 46

classification because sodium reacts with the soil to create sodium hazards by replacing other cations. The extent of this replacement is estimated by the sodium adsorption ratio (SAR). A diagram for use in studying the suitability of groundwater for irrigation is based on the sodium adsorption ratio (SAR) and the electrical conductivity of water expressed in μS/cm. The chemical analytical data on groundwater samples from District Hardwar were processed as per the U.S. Salinity Laboratory classification for the two sets of data (Figs. 2 and 3), and the results are summarized in Table 7.

It is evident from the results that majority of the samples (about 70% of total samples) falls under water type C2-S1 (medium salinity and low SAR); such water can be used if a moderate amount of leaching occurs, and plants with moderate salt tolerance can be grown in most cases without special practices for salinity control. About 30% of the samples falls under water type C3-S1 (high salinity and low SAR); such water cannot be used on soils where drainage is restricted. Even with adequate drainage, special management for salinity control may be required, and plants with good tolerance should be selected. An almost similar trend was observed during the postmonsoon season.

CONCLUSIONS

Groundwater quality varies from place to place and with the depth of the water table. The SAR values indicate that the majority of samples from the study area falls under the category of low sodium hazard indicating no risk for sodification. In general, the groundwater of District Hardwar is safe for irrigation. According to the U.S. Salinity Laboratory classification of irrigation water, the majority of the samples (70%) falls under water type C2-S1 followed by the C3-S1 type.

BIBLIOGRAPHY

1. American Public Health Association. (1992). *Standard Methods for the Examination of Water and Waste Waters*, 18th Edn. American Public Health Association, Washington, DC.
2. Jain, C.K. and Bhatia, K.K.S. (1988). *Physico-Chemical Analysis of Water and Wastewater, User's Manual, UM-26*. National Institute of Hydrology, Roorkee, India.
3. Jain, C.K. (2003). *Ground Water Quality in District Hardwar, Uttaranchal, Technical Report*. National Institute of Hydrology, Roorkee, India.
4. Richards, L.A. (Ed.). (1954). *Diagnosis and Improvement of Saline and Alkali Soils, Agricultural Handbook 60*. U.S. Dept. Agric., Washington, DC.
5. Wilcox, L.V. (1955). *Classification and Use of Irrigation Water*. U.S. Dept. of Agric. Circular 969, Washington, DC.

INFILTRATION AND SOIL WATER PROCESSES

Joseph Holden
University of Leeds
Leeds, United Kingdom

Water at the soil surface is drawn into the soil pores under the influence of both suction and gravity. Suction (or tension/capillary) forces are largest in the small pore spaces between soil particles. Therefore, small pores fill first, and large pores empty first. As a soil becomes saturated, suction forces approach zero. As a soil dries, suction forces dominate, and there is little or no gravitational flow. For a dry soil during the early stages of wetting, the effect of soil pore suction is predominant. However, as the depth of wet soil increases, the suction gradient decreases (the same difference in suction is spread over an ever increasing depth interval) and the gravitational head gradient becomes the main driving force (1).

Infiltration processes cannot easily be separated from the processes involved in movement of water within soil because the rate of infiltration is partly controlled by the rate of soil water movement below the surface. Soil water movement continues after an infiltration, as the infiltrated water is redistributed. Infiltration and soil water processes play a key role in surface runoff, groundwater recharge, ecology, evapotranspiration, soil erosion, and transport of nutrients and other solutes in surface and subsurface waters.

Bodman and Colman (2) suggested that for a uniform soil, there would be a series of zones in the wetting part of the soil profile during an infiltration. The zone nearest the surface is a saturated zone (typically in the upper centimeter of the soil profile). As water penetrates more deeply, a zone of uniform water content, the transmission zone, develops behind a well-defined wetting front (Fig. 1).

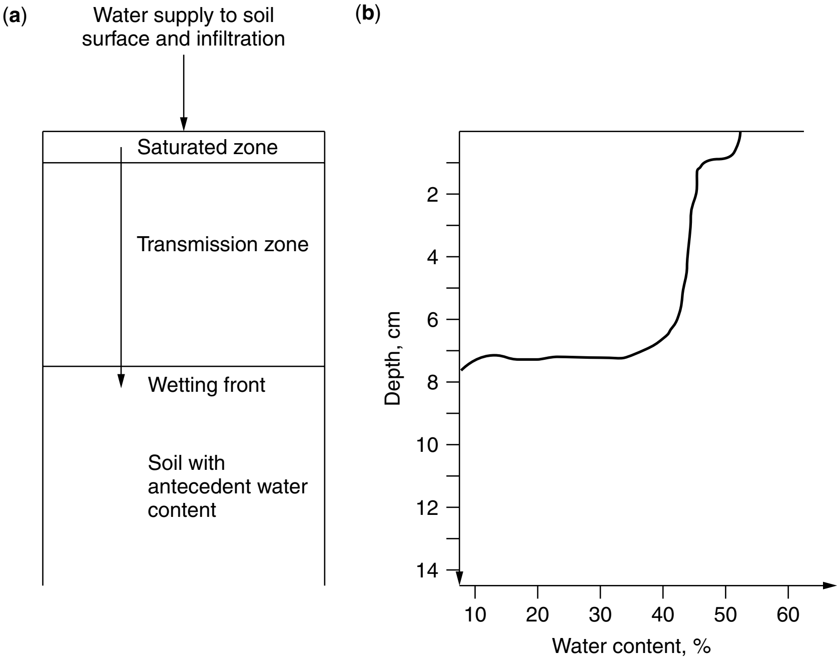

Figure 1. Soil water zones during infiltration: (**a**) theoretical zonation, (**b**) measured water content for a sandy soil 4 minutes after ponded infiltration commenced.

There is a sharp change in water content at the wetting front because the water at the boundary takes up a preferred position of minimum potential energy in the narrowest pores, for which the hydraulic conductivity is very low, and does not move at an appreciable rate until the large pores begin to fill. Figure 1 shows the water content versus depth for a sandy soil 4 minutes after ponded infiltration at the surface. The sharp transition in water content can be seen at the wetting front. Note that the soil below the wetting front still has some pre-event moisture. There is also a sharp transition in the moisture profile between the upper saturated zone and the transmission zone, sometimes called the transition zone (1).

Gravitational and suction gradients result in continued downward movement of water within the soil profile which may occur long after infiltration at the surface has ceased. As a result, the transmission zone which existed during infiltration becomes a draining zone that releases water from the wetter upper layers of the soil to deeper drier layers. As the draining zone becomes drier, the hydraulic conductivity reduces, and so the rate of soil water redistribution slows down. The suction gradient also weakens as the soil water content becomes more uniform. While the wetting front continues to advance into deeper soil layers, its movement is reduced over time until eventually the water content changes only very slowly and the soil is said to be at field capacity (1). In some soils, this may take several days (2). Theoretically, soil water redistribution following rainfall would continue until gravity and soil suction forces were in balance (4,5). At this stage and at all depths, there would be equal soil water potential. Therefore, without a hydraulic gradient, there would be no soil water movement, which is rarely observed in the field because soils are highly heterogeneous, rainfall itself is rarely an on–off uniform event (a theoretical model examining this issue was applied to a range of soil types by Ref. 6), and because evaporation from soil and plants can result in drying of surface soil layers.

Evaporation at the soil surface encourages upward movement of soil water, which therefore reduces downward movement of water over time. There is probably a two-way effect whereby drainage reduces the amount of evaporation and evaporation reduces the amount of drainage (3). During water redistribution following infiltration, the upper part of the soil profile may be drying through drainage and evaporation while the lower part is still becoming wetter, which complicates the water content–soil moisture tension (suction) relationship within the soil because it will depend on the history of wetting and drying in any given part of the soil profile (soil water hysteresis; 4,7,8).

Often, the soil water redistribution that follows an infiltration is complicated by the nonuniform nature of most soils. In addition to the effects of soil layering, many studies, for example, often assume that the water table has little impact on postinfiltration water redistribution. However, in many soils, the water table may be shallow (e.g., peats), and this may exert considerable influence on the distribution of water in a soil profile (9). Trapped soil gases may also affect water redistribution by blocking pores and preventing water entry (10). Furthermore, noncapillary macropores or voids such as decayed root channels, worm holes, and structural cracks that are open at the soil surface can capture the free water available at the surface during rainfall and overland flow. This causes the water flowing through the macropores to bypass the soil matrix. Often, water and solutes (e.g., fertilizer applications) can infiltrate into the soil via macropores but then are quickly transferred to much deeper parts of the soil profile where the water and nutrients are not readily available to plants (11,12). Bypassing flow also causes some parts of the soil profile at any given depth to wet well before other parts of the soil that are at the same depth.

BIBLIOGRAPHY

1. Ward, R.C. and Robinson, M. (2000). *Principles of Hydrology*. McGraw-Hill, London.
2. Bodman, G.B. and Colman, E.A. (1943). Moisture and energy conditions during downward entry of water into soils. *Proceedings of the Soil Science Society of America* **8**: 116–122.
3. Gardner, W.R., Hillel, D., and Benyamini, Y. (1970). Post irrigation movement of soil water, 2, Simultaneous redistribution and evaporation. *Water Resources Research* **6**: 1148–1153.
4. Childs, E.C. (1969). *An Introduction to the Physical Basis of Soil Water Phenomena*. John Wiley & Sons, London.
5. Rawls, W.J., Ahuja, L.R., Brakensiek, D.L., and Shirmohammadi, A. (1993). Infiltration and soil water movement. In: *Handbook of Hydrology*. D.R. Maidment (Ed.). McGraw-Hill, New York, pp. 5.1–5.51.
6. Corradini, C., Melone, F., and Smith, R.E. (1997). A unified model for infiltration and redistribution during complex rainfall patterns. *Journal of Hydrology* **192**: 104–124.
7. Vachaud, G. and Thony, J.L. (1971). Hysteresis during infiltration and redistribution in a soil column at different initial water contents. *Water Resources Research* **9**: 160–173.
8. Beese, F. and van der Ploeg, R.R. (1976). Influence of hysteresis on moisture flow in an undisturbed soil monolith. *Proceedings of the Soil Science Society of America* **40**: 480–484.
9. Evans, M.G., Burt, T.P., Holden, J., and Adamson, J. (1999). Runoff generation and water table fluctuations in blanket peat: evidence from UK data spanning the dry summer of 1995. *Journal of Hydrology* **221**: 141–160.
10. Beckwith, C.W. and Baird, A.J. (2001). Effect of biogenic gas bubbles on water flow through poorly decomposed blanket peat. *Water Resources Research* **37**: 551–558.
11. Beven, K.J. and Germann, P. (1982). Macropores and water flow in soils. *Water Resources Research* **18**: 1311–1325.
12. Germann, P.F. (1990). Macropores and hydrologic hillslope processes. In: *Process Studies in Hillslope Hydrology*. M.G. Anderson and T.P. Burt (Eds.). John Wiley & Sons, Chichester, pp. 327–363.

INFILTRATION/CAPACITY/RATES

JOSEPH HOLDEN
University of Leeds
Leeds, United Kingdom

Infiltration is the process of water entry into the surface of a soil. It involves surface water entry, transmission through the soil, and depletion of the soil water storage capacity. Surface water entry is influenced by surface vegetation cover, soil texture, soil porosity, soil structure (e.g., cracks, surface crusting), and compaction. Land management can influence all of these parameters and therefore influence infiltration.

The infiltration rate is the volume of water passing into the soil per unit area per unit time and therefore has dimensions of velocity (e.g., m s^{-1}; mm h^{-1}). The infiltration rate is often measured using an infiltrometer (INFILTROMETERS) and is one of the most easily and frequently measured soil water processes. Despite the units of velocity, the values of infiltration rate are actually volume fluxes expressed as volume flux density and therefore do not indicate the effective velocity of vertical water movement in the soil (1).

The maximum rate at which water soaks into or is absorbed by the soil is the infiltration capacity. Infiltration capacity may be very important in determining the proportion of incoming rainfall that runs off as infiltration-excess overland flow and the proportion that moves into the soil. The infiltration capacity is different from the infiltration rate. Infiltration may be limited by the rate of supply of rainfall (or other water supply) or the capacity of the soil to absorb water. In the former case, infiltration is limited by the water supply to the soil surface, and in the latter case, infiltration is restricted by the infiltration capacity of the soil. Use of the term infiltration rate therefore indicates that infiltration is proceeding at less than the infiltration capacity (2). If infiltration is occurring at less than the infiltration capacity, then all rain reaching the soil surface that is not held as surface storage will infiltrate into the soil. Thus, there will be a direct relationship between infiltration rate and rainfall intensity. However, when rainfall intensity exceeds the infiltration capacity of the soil surface, this relationship breaks down and may be inverted, as infiltration capacity declines through a storm as the soil wets (2). Excess rainwater that cannot infiltrate into the soil because the surface water supply is greater than the infiltration capacity will pond on the soil surface and eventually run over the surface as infiltration-excess overland flow.

The infiltration capacity of a soil generally decreases during rainfall, rapidly at first and then more slowly, until a more or less stable value has been attained (Fig. 1). Water initially crosses the soil surface interface at an initial rate (f_0), depending on existing soil moisture content. As the rainfall or other water supply to the soil surface continues, the rate of infiltration decreases (because it is limited by decreasing infiltration capacity) as the soil becomes wetter and less able to take up water. Philip (3) put forward the theory that the infiltration capacity of a soil declines over time in a manner described by the equation,

$$f = A + Bt^{-0.5}$$

where $f =$ infiltration capacity at time t, and A and B are constants. This equation may be solved using simultaneous equations, allowing constants A and B to be evaluated:

$$B = (f_1 - f_0)/(t_1^{-0.5} - t_0^{-0.5})$$

$$A = (f_0 t_1^{-0.5} - f_1 t_0^{-0.5})/(t_1^{-0.5} - t_0^{-0.5})$$

The curve of infiltration rate over time shown in Fig. 1 reduces to a constant value f_c, which is the infiltration capacity at time t_c. The Philip model is useful for short infiltration events, but other models (e.g., those based on the Richards equation) are more reliable for longer periods (4,5).

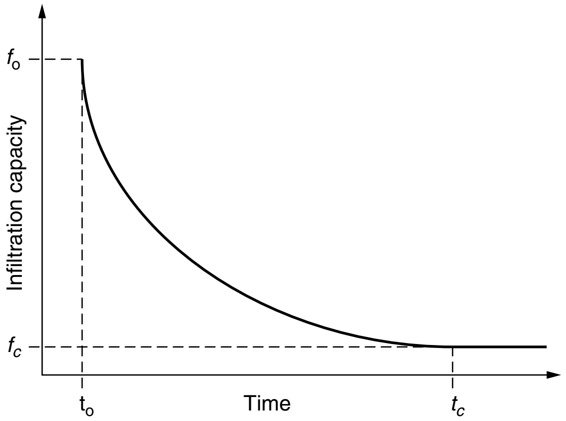

Figure 1. The decline of infiltration capacity during a rainfall.

Soil surface conditions may impose an upper limit on the rate at which water can be absorbed, despite a large available capacity of the lower soil layers to receive and to store additional infiltrating water. Often, the infiltration capacity is reduced by surface sealing resulting from compaction, washing of fine particles into surface pores, and by frost (6,7). Frozen moisture will block soil pores, and the greater the frozen moisture content, the lower the infiltration rate upon melt, as the larger pores remain blocked for longest. Holden and Burt (8) noted that infiltration capacity increases with the depth of standing water on the surface because of increased water head (9,10) and unblocking of pores due to flotation of fines (11). The number of cracks and fissures at the surface and the ground slope also influences infiltration capacity. Cultivation techniques may either increase or decrease infiltration capacity (12). Field plowing can increase soil infiltration capacity, whereas raindrop compaction subsequent to field plowing can reduce infiltration capacities when splashed silt and clay soil particles clog soil pores as the aggregates break down. Organic matter also binds aggregates together. New vegetation, however, tends to increase the infiltration capacity of a soil by retarding surface water movement, stabilizing loose particles, reducing raindrop compaction, and improving soil structure. Soils with well developed humus and litter layers tend to have high infiltration capacities. Most forest floors that have a litter layer tend, for example, to have higher infiltration than that beneath grass (2).

The infiltration capacity of a soil is dependent on surface properties and also on transmission through the soil below the surface and the storage capacity of the soil. Water cannot be absorbed by the soil surface unless the underlying soil profile can conduct the infiltrated water away. Transmission rates depend on soil texture, porosity, pore size distribution, soil stratification, antecedent soil moisture content, initial gradient of soil water potential, salinity, and biotic activity. Storage capacity depends on pore size distribution, porosity, and antecedent moisture conditions. Where the soil profile contains a relatively impermeable layer at some distance below the surface, the curve of infiltration capacity against time may display a sudden reduction of infiltration capacity because when available storage in the surface soil horizons has been filled, infiltration will be governed by the rate at which water can pass through the underlying impeding layer.

Soil properties are highly heterogeneous. Infiltration capacities and infiltration rates vary greatly even within a small area and in a short time. The effect of spatial variability on infiltration produces a difference between point infiltration values and apparent infiltration rates associated with measurements from large plots, composite areas, and estimates from watershed calculations, which may provide a reason that mean infiltration capacity across a plot increases with rainfall intensity (13,14; see also RAIN SIMULATORS). In other words, the nonuniform nature of soil across a large plot may cause in the surface of one part of the plot to have a higher infiltration capacity than the rest of the plot. Therefore, as Hawkins (13) demonstrated numerically, the mean infiltration rate across a plot will increase with rainfall intensity simply because a greater flux of water is occurring through the parts of the plot surface that have higher relative infiltration capacities. It must be acknowledged, therefore, that only a fraction of the area within a plot or hillslope need contribute to overland flow, even when the rainfall rate exceeds the mean infiltration capacity across the plot or hillslope, which also suggests that the spatial variability in infiltration parameters has more effect at lower rainfall rates, closer to the saturated hydraulic conductivity of the soil (5).

BIBLIOGRAPHY

1. White, R.E. (1997). Principles and practice of soil science. *The Soil as a Natural Resource*. Blackwell Science Ltd, Oxford.
2. Ward, R.C. and Robinson, M. (2000). *Principles of Hydrology*. McGraw-Hill, London.
3. Philip, J.R. (1957). The theory of infiltration 1. The infiltration equation and its solution. *Soil Science* **83**: 345–357.
4. Kirkby, M.J. (1985). Hillslope hydrology. In: *Hydrological Forecasting*. M.G. Anderson and T.P. Burt (Eds.). John Wiley, Chichester, pp. 37–75.
5. Rawls, W.J., Ahuja, L.R., Brakensiek, D.L., and Shirmohammadi, A. (1993). Infiltration and soil water movement. In: *Handbook of Hydrology*. D.R. Maidment (Ed.). McGraw-Hill, New York, pp. 5.1–5.51.
6. Poesen, J. (1986). Surface sealing as influenced by slope angle and position of simulated stones in the top layer of loose sediments. *Earth Surface Processes and Landforms* **11**: 1–10.
7. Romkens, M.J.M., Prasad, S.N., and Parlange, J.Y. (1990). Surface seal development in relation to rainsform intensity. *Catena Supplement* **17**: 1–11.
8. Holden, J. and Burt, T.P. (2002). Infiltration, runoff and sediment production in blanket peat catchments: Implications of field rainfall simulation experiments. *Hydrological Processes* **16**: 2537–2557.
9. Schiff, L. (1953). The effect of surface head on infiltration rates based on the performance of ring infiltrometers and ponds. *Transactions of the American Geophysical Union* **34**: 257–266.
10. Philip, J.R. (1958). The theory of infiltration 6. Effect of water depth over soil. *Soil Science* **85**: 278–286.

11. Bowyer-Bower, T.A.S. (1993). Effects of rainfall intensity and antecedent moisture on the steady-state infiltration rate in a semi-arid region. *Soil Use and Management* **9**: 69–76.
12. Imeson, A.C. and Kwaad, F.J.P.M. (1990). The response of tilled soils to wetting by rainfall and the dynamic character of soil erodibility. In: *Soil Erosion on Agricultural Land*. J. Boardman, I.D.L. Foster, and J.A. Dearing (Eds.). John Wiley, Chichester, pp. 3–14.
13. Hawkins, R.H. (1982). Interpretations of source area variability in rainfall-runoff relations. In: *Rainfall-runoff relationship. Proceedings of the International Symposium on rainfall runoff modelling held May 18–21, 1981, Mississippi State University*. Water Resources Publications, Littleton, CO, pp. 303–324.
14. Smith, R.E. (1983). Flux infiltration theory for use in watershed hydrology. In: *Advances in Infiltration*. American Society of Agricultural Engineers, St. Joseph, MI, pp. 313–323.

INFILTROMETERS

Joseph Holden
University of Leeds
Leeds, United Kingdom

Infiltrometers allow measuring of the rate of infiltration of water into a given medium. Rain simulators, a particular type of infiltrometer, are discussed elsewhere (see entry Rain Simulators). Unlike rain simulators, other infiltrometers do not simulate raindrop activity but measure the rate of infiltration from a reservoir of water at the ground surface. Their main advantage is their simplicity and ease of use. There are two main types of infiltrometer, a ponded ring infiltrometer and a disk infiltrometer.

PONDED RING INFILTROMETER

A ponded ring infiltrometer is the most common type of infiltrometer used. It is inexpensive to construct or operate and requires relatively little water for measurements. One person can set up and run several tests simultaneously. A cylindrical ring of stainless steel or plastic pipe is used, and the ring is pushed a few centimeters into the ground (Fig. 1). Care is taken to minimize disturbing the soil surface and soil structure during installation. Water is then flooded into the ring. The water inside the ring gradually infiltrates the soil, so measurements of the depth of water in the ring over time can be used to provide an infiltration rate usually in units of length per unit time (e.g., mm hr^{-1}) allowing easy comparison with rainfall intensity data. If the ring empties, it can be easily refilled.

There are two main problems in this method: (1) lateral flow below the ring; (2) changing pressure of water in the ring as the water level decreases in the ring during infiltration. To get around the lateral flow problem, it is advisable to use two concentric rings; the larger ring forms a buffer compartment around the inner ring (Fig. 2). While water is topped up in the outer ring, measurements are taken only of infiltration from the inner ring. It is preferable that a constant head of water is maintained within the rings to avoid changing pressures. The head of water should also be less than 5 cm deep. The constant head can be supplied by a Mariotte bottle placed above the rings, and readings of water usage can be made from the Mariotte bottle (Fig. 2). Thus the optimum design for a ring infiltrometer is a constant-head, double-ring device whose rings have large diameters, so that they can represent the soil surface more readily. A plot of water uptake against time should show a leveling off as the infiltration capacity is reached. It often takes several hours before a constant rate of infiltration is achieved from the infiltrometer, depending on soil type, texture, and antecedent soil moisture.

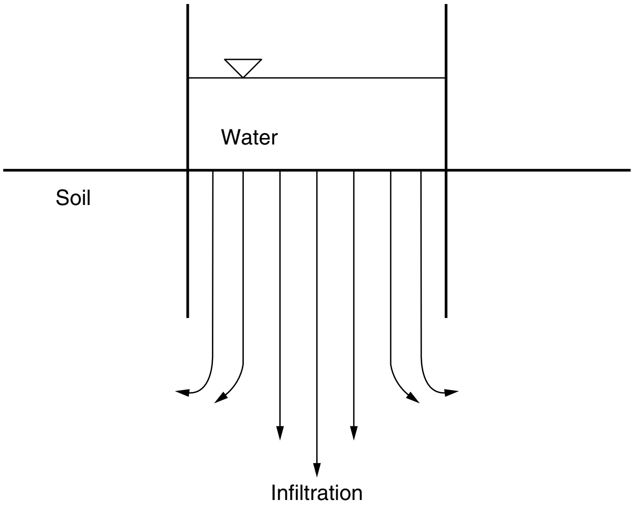

Figure 1. Infiltration from a single ring infiltrometer.

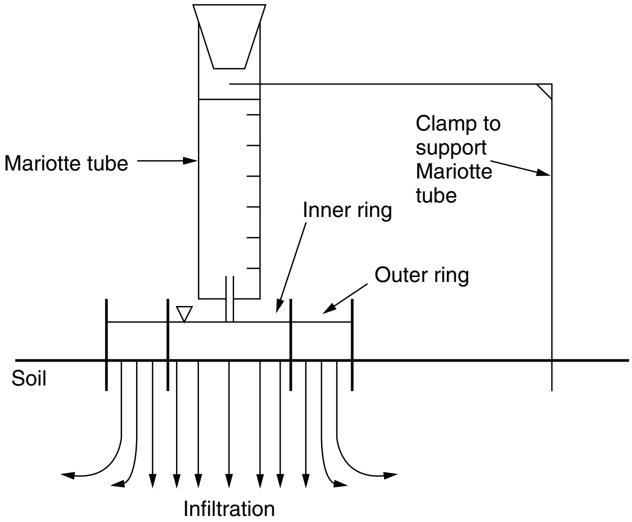

Figure 2. Double-ring infiltrometer with constant-head supply.

The method of placement is serious limitation to the use of ring infiltrometers. Knocking rings into the ground can result in destruction of the soil structure or compression. If a soil is shattered, it can disturb the contact between the ring edge and the soil, resulting in leakage and high infiltration rates. Research has demonstrated that ring infiltrometers give higher infiltration rates than rain simulators because of the effect of a pond on the soil surface, lateral seepage, and soil cracking.

DISK INFILTROMETER

Tension disk infiltrometers are a standard tool for in situ determination of saturated and near-saturated soil hydraulic properties. These infiltrometers have a porous membrane disk at their base which rests on the soil surface of interest. To assess the role of matrix and macropore flow, a tension infiltrometer allows infiltration of water into the soil matrix, while preventing flow into larger pores that may otherwise dominate the infiltration process. The infiltrometer provides a source of water at a small negative pore water pressure at the surface. The supply pressure head is controlled with a Mariotte bottle. A schematic diagram of a tension disk infiltrometer is given in Fig. 3. As for a ring infiltrometer, infiltration rates can be measured manually by observing the volume of water lost from the Mariotte bottle over time. The negative pressure prevents the larger pores that fill at greater pore water pressures from wetting and short-circuiting the flow. Hence, by subtraction, the hydrological role of larger pores during the infiltration process can be evaluated. Further details on designs for these infiltrometers are given in Ankeny et al. (1) and Zhang et al. (2). Most studies using tension disk infiltrometers have been conducted at the soil surface, although Azevedo et al. (3) looked at infiltration properties of an Iowa loamy soil at 0.15 m depth, and Logsdon et al. (4), and Messing and Jarvis (5) conducted measurements at different depths under different agricultural tillages.

Careful preparation of the soil surface is required to use a disk infiltrometer, which is to ensure even and sound contact of the disk with the soil surface. At each location, vegetation must be cut back to the soil surface, and a fine layer of moist fine sand of the same diameter as the circular base of the infiltrometer should be applied, which must be smoothed out to remove any irregularities at the soil surface and improve contact between the disk and soil surface. Moist sand is essential as air-dry sand may readily fall down into surface-vented macropores, forming ‘wicks’ (5). The infiltrometer is then placed on the sand. The weight of the infiltrometer must not be too great so as compress the soil surface significantly, as this will restrict the water flux. Therefore, it is usual to make the water supply reservoir narrow so that the total volume of water held in the infiltrometer is low, resulting in reduced weight and also aiding accurate measurements of discharge.

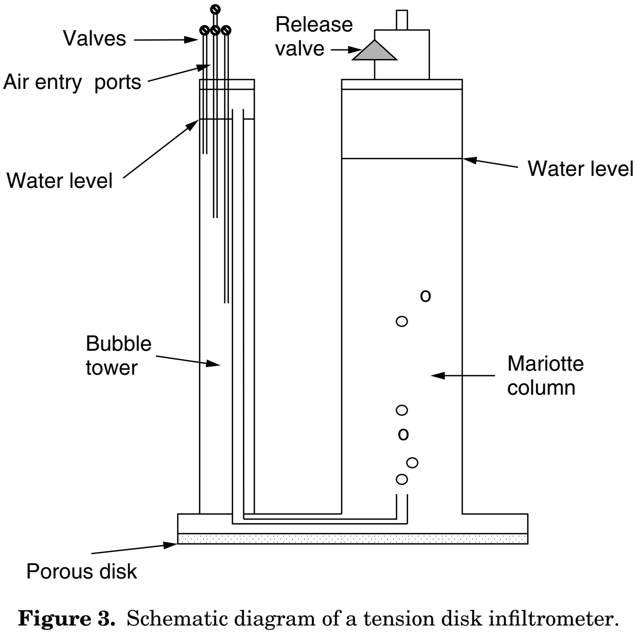

Figure 3. Schematic diagram of a tension disk infiltrometer.

If a range of supply heads is to be used, then infiltration tests are normally conducted with the lowest supply head first. Reversal of this may lead to hysteresis where drainage occurs close to the disk while wetting continues near and at the infiltration front (6). Infiltration measurements should proceed until a steady state is achieved. Users should be careful to ensure that sunlight does not heat the supply reservoir significantly: this can be reduced by shading.

Methods for analyzing the data from disk infiltrometers (e.g., hydraulic conductivity values and macropore contribution to infiltration) are given in Reynolds and Elrick (6). Typically, Wooding’s (7) solution for infiltration from a shallow pond is combined with Gardner’s (8) unsaturated hydraulic conductivity function. A range of assumptions is involved in using of these techniques, including that the hydraulic conductivity before the test is much less than that imposed under the infiltration experiment and that the soil below the tension disk is homogeneous, isotropic, and uniformly saturated, which are unlikely to be satisfied fully in most soils, and so it is necessary to evaluate the potential errors involved (9).

Water levels in supply bottles for both the ring and disk infiltrometers can be measured automatically by using electronic pressure sensors with data recorders, thus allowing an experiment to continue unattended for several hours if the infiltration rates are very low.

BIBLIOGRAPHY

1. Ankeny, M.D., Kaspar, T.C., and Horton, R. (1988). Design for an automated tension infiltrometer. *Journal of the Soil Science Society of America* **52**: 893–896.
2. Zhang, Y., Smith, R.E., Butters, G.L., and Cardon, G.E. (1999). Analysis and testing of a concentric-disk tension infiltrometer. *Soil Science Society of America Journal* **63**: 544–553.
3. Azevedo, A.S., Kanwar, R.S., and Horton, R. (1998). Effect of cultivation on hydraulic properties of an Iowa soil using tension infiltrometers. *Soil Science* **163**: 22–29.
4. Logsdon, S.D., Jordahl, J.L., and Karlen, D.L. (1993). Tillage and crop effects on ponded and tension infiltration rates. *Soil and Tillage Research* **28**: 179–189.
5. Messing, I. and Jarvis, N.J. (1993). Temporal variation in the hydraulic conductivity of a tilled clay soil as measured by tension infiltrometers. *Journal of Soil Science* **44**: 11–24.
6. Reynolds, E.D. and Elrick, D.E. (1991). Determination of hydraulic conductivity using a tension infiltrometer. *Journal of the Soil Science Society of America* **55**: 633–639.

7. Wooding, R.A. (1968). Steady infiltration from a shallow circular pond. *Water Resources Research* **4**: 1259–1273.
8. Gardner, W.R. (1958). Some steady state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. *Soil Science* **85**: 228–232.
9. Holden, J., Burt, T.P., and Cox, N.J. (2001). Macroporosity and infiltration in blanket peat: the implications of tension disc infiltrometer measurements. *Hydrological Processes* **15**: 289–303.

SUMMARY OF ISOTOPES IN CONTAMINANT HYDROGEOLOGY

Segun Adelana
University of Ilorin
Ilorin, Nigeria

INTRODUCTION

The word contaminant is a general term to describe dissolved constituents, nonaqueous phase liquids (NAPLs), or industrial solvents which, when added to water as a result of human activities, impair groundwater quality. To an important extent, the complexity of a groundwater contamination problem can be determined by whether or not NAPLs are present (1). Obviously, contaminants related to anthropogenic activities are one of the major threats facing groundwater resources today. Groundwater contamination can have a significant impact on the present and future use of groundwater resources as a source of drinking water for an increasing population living in cities and rural areas (2).

In the investigation of problems posed by contaminants in the hydrologic environment, the modern approach is to use a combination of tracers. These tracers include radioactive isotopes (such as ^{3}H, ^{13}C, ^{14}C,), stable isotopes (^{18}O, ^{2}H), and dissolved geochemical constituents (Cl^{-}, $SO_4{}^{2-}$, $NO_3{}^{-}$ and some cations). An important facet of contaminant hydrogeology is the movement of liquids that are immiscible with water through the vadose zone as well as below the water table. Such liquids are usually called nonaqueous phase liquids. There are numerous cases of groundwater contamination involving organic liquids that are either insoluble or partially soluble in water. When the densities of these liquids are greater than that of water, they are called dense nonaqueous phase liquids (DNAPLs). Those whose densities are less than that of water are light nonaqueous phase liquids (LNAPLs). They could also be partially soluble in water, such that both dissolved and nonaqueous phases may be present (3–7).

The major groups of contaminants in hydrogeology include radionuclides, trace metals, and nutrients. They could also be classified as inorganic and organic contaminants. These include synthetic organic compounds, hydrocarbons, and pathogens (7). Some types of these contaminants are inorganic chemicals that are either metallic or nonmetallic (1,8). Included in the group of organic contaminants are NAPLs, which naturally subdivide into LNAPLs and DNAPLs (7).

A wide variety of materials have been identified as contaminants found in groundwater. An extensive listing of these compounds is tabulated in (Reference 7, pp. 3–10). Many of these materials will dissolve in water to different degrees. Some of the organic compounds are only slightly soluble and will exist in both dissolved form and as an insoluble phase, which can also migrate through the ground (7). According to (2), the main relevant organic compounds of the LNAPLs are benzene, toluene, ethylbenzene, and xylenes (BTEX) from spills at gas stations and industrial facilities. The most common compounds of the group found in groundwater are perchloroethene (PCE), trichloroethene (TCE), *cis*-1,2-dichloroethene (*cis*-1,2-DCE), and dichloroethane (DCA) (9). In terms of environmental contamination, the major concern is focused on those organic compounds that are not readily degraded by bacteria, either in the soil zone or in sewage treatment facilities. The presence of these substances in the surface environment is becoming increasingly pervasive (8).

The sources that contribute to groundwater contamination are varied. Among these are land disposal of solid wastes, inorganic substances, metal and nonmetal trace elements. Other sources of groundwater contamination are from introducing of salts on roads to combat ice in winter, the infiltration of leachate from tailings in the mining industry, seepage from industrial water lagoons and septic tanks of urban waste disposal systems, wastewater discharge into rivers and lakes, agricultural wastes and fertilizers, saline or geogenic polluted water attracted by overexploitation of groundwater, and fossil fuels (1,7,8,10,11). Inorganic cations and anions occur in nature and may come from natural as well as anthropogenic sources. Some of the radionuclides are naturally occurring; others are man-made and come from the production and testing of nuclear weapons.

There are three important attributes that distinguish sources of groundwater contamination: (1) their degree of localization, (2) their loading history, and (3) the kinds of contaminants emanating from them (1). From the work of Domenico and Schwartz (1) contaminants are organized by groups according to reaction type and mode of occurrence. Six categories have been tabulated:

Category 1—Sources designed to discharge substances, for example, subsurface percolation (from septic tanks and cesspools), injection wells (hazardous and nonhazardous wastes), and land application.

Category 2—Sources designed to store, treat, and/or dispose of substances; discharge through unplanned release, for example, landfills, open dumps (including illegal dumping waste), residential or local disposal wastes, waste tailings, material stockpiles (nonwaste), graveyards, animal burial, surface and underground storage tanks, and containers.

Category 3—Sources designed to retain substances during transport or transmission, for example, pipelines and materials transport and transfer operations.

Category 4—Sources discharging substances as a consequence of other planned activities, for example, irrigation practices, pesticide applications, fertilizer applications, animal feeding operations, deicing salt applications, urban runoff, percolation of atmospheric pollutants, and mining and mining drainage.

Category 5—Sources providing conduits or inducing discharge through altered flow patterns, for example, production wells (oil/gas wells, geothermal and heat recovery wells, water supply wells), other nonwaste wells (monitoring and exploration wells), and construction excavation.

Category 6—Naturally occurring sources whose discharge is created and/or exacerbated by human activity, for example, groundwater–surface water interactions, natural leaching, saltwater intrusions and brackish water upconing, or intrusion of other poor-quality natural water.

For details of sources and background concentrations of contaminants in groundwater, refer to Forstner and Wittman (10). More information on inorganic contaminants is available in Reference 7 (Chapter 6), Reference 1 (Chapter 17); methods of detecting groundwater contamination are fully discussed in Domenico and Schwartz (1), and the sources and hydrochemical behavior of contaminants are described in Chapter 9 of Reference 8 as well as Chapter 1 of Reference 11. As a general introduction to groundwater contamination and related literature, the following are suggested for further reading: Baetsle (12), Fuhriman and Barton (13), Hall (14), Scalf et al. (15), Miller et al. (16), Summers and Spiegal (17), Cherry et al. (18), Back and Cherry (19), Fried (20), Todd and McNulty (21), Wilson et al. (22), Freeze and Cherry (8), IAEA (23, 24), Pearson et al. (25), Fetter (7), Pankow and Cherry (9), Domenico and Schwartz (1), and Seiler (11).

ENVIRONMENTAL ISOTOPES IN CONTAMINANT HYDROGEOLOGY

Introduction

For many years, environmental isotopes have been used in hydrogeology to address a range of issues such as evaluation of recharge areas, determination of groundwater residence time, and evaluation of surface water–groundwater interactions. Isotope studies applied to a wide spectrum of hydrologic problems related to both surface and groundwater resources as well as environmental studies in hydro-ecological systems are presently an established scientific discipline, called “Isotope Hydrology.” However, it is now becoming increasingly popular to use environmental isotopes as tracers to provide information about sources of contaminants and processes that affect these compounds in the groundwater flow system (2).

In essence, isotope techniques are becoming an integral part of many hydrologic investigations and sometimes a unique tool in studies related to groundwater contamination. Stable and radioactive environmental isotopes have now been used for more than four decades to study hydrologic systems, particularly in determining the hydrodynamics of aquifers, and have proved useful for understanding groundwater systems. The applications of isotopes in hydrogeology are based on the general concept of “tracing,” in which either intentionally introduced isotopes or naturally occurring (environmental) isotopes are employed.

Environmental isotopes (either radioactive or stable) have the distinct advantage over injected (artificial) tracers in that they facilitate the study of various hydrologic processes on a much larger temporal and spatial scale through their natural distribution in a hydrologic system. Thus, environmental isotope methodologies are unique in regional studies of water resources to obtain the time and space integrated characteristics of groundwater systems. The use of artificial tracers generally is effective for site-specific, local applications (26,27). Environmental isotopes can now be used to trace the pathways as well as spatial distribution and temporal changes in pollution patterns for assessing pollution migration scenarios and planning for aquifer remediation (27,28).

Recent research and the development of new analytical methods have led to fast expansion of the application of isotopes in hydrogeology. In particular is the introduction of techniques that make it possible to measure isotopic ratios of individual compounds at low concentrations (29). These have opened new and exciting possibilities for using isotopes in organic contaminant studies in groundwater. For example, nitrogen and carbon isotopes can now provide information regarding such attenuation processes as the denitrification and biodegradation of halogenated solvents, respectively (28).

Environmental isotopes have, therefore, become useful tools in differentiating between different sources of organic contaminants or in assessing biotic and abiotic transformation processes associated with natural attenuation or engineered groundwater remediation techniques (30–32). The isotope method used to evaluate the degradation of organic contaminants relies on the frequent occurrence of small differences in the degradation rate between molecules of the same contaminant that contain different isotopes. Usually, the rate is slightly faster for molecules that contain light isotopes (e.g., ^{12}C) compared to those that contain heavy isotopes (e.g., ^{13}C) because bonds between light isotopes are weaker. As a result, degrading contaminants should contain increasingly more of the heavier isotope (^{13}C) as degradation proceeds, and thus the ratio between heavy and light isotopes ($^{13}C/^{12}C$) is expected to increase (2). Oxygen-18 in dissolved oxygen also has a significant potential as a tracer to evaluate the fate of oxygen in contaminated plumes.

The sequel to these environmental isotope studies has focused on the light elements and their isotopes: hydrogen (^{1}H, ^{2}H, ^{3}H), carbon (^{12}C, ^{13}C, ^{14}C), nitrogen (^{14}N, ^{15}N), oxygen (^{16}O, ^{18}O), and sulfur (^{32}S, ^{34}S). These are the most important elements in biological systems that occur in relatively great abundance and also participate in most geochemical reactions (33). Other less important and less abundant isotopes (^{17}O, ^{33}S, ^{36}S) are not often employed in environmental studies but have significance in very specific studies. Table 1 lists the isotopes of major elements used in environmental studies and their average abundances in natural compounds.

Table 1. Isotopes of Major Elements Used in Environmental Studies[a]

Element	Isotopes	Average Terrestrial Abundance, (%)	Comments[b]
Hydrogen	^{1}H	99.984	
	^{2}H	0.015	
	^{3}H	10^{-14} to 10^{-16}	Radioactive, $t_{1/2} = 12.35$ years
Carbon	^{12}C	98.89	
	^{13}C	1.11	
	^{14}C	$\sim 10^{-10}$	Radioactive, $t_{1/2} = 5730$ years
Oxygen	^{16}O	99.76	
	$^{17}O^{c}$	0.037	
	^{18}O	0.1	
Nitrogen	^{14}N	99.34	
	^{15}N	0.366	
Sulphur	^{32}S	95.02	
	$^{33}S^{c}$	0.75	
	^{34}S	4.21	
	$^{36}S^{c}$	0.02	
Strontium	$^{84}Sr^{c}$	0.56	
	^{86}Sr	9.86	
	^{87}Sr	~7.02	
	$^{88}Sr^{c}$	82.56	
Uranium	^{234}U	~0.0056	Radioactive, $t_{1/2} = 2.47\times10^{5}$ years
	$^{235}U^{c}$	0.7205	Radioactive, $t_{1/2} = 7.13\times10^{8}$ years
	^{238}U	99.274	Radioactive, $t_{1/2} = 4.51\times10^{9}$ years

[a] Reference 33.

[b] $t_{1/2}$ = half-life.

[c] These isotopes are not presently used in environmental studies.

Uses of Isotopes in Contaminant Hydrogeology

General uses of environmental isotopes in contaminant studies are often based on evaluating the various sources of contaminants in both surface and groundwater. The isotope approach to evaluating these sources relies mostly on the distinct isotope composition, that can characterize the different sources of contaminant (2,34). Environmental isotopes have found good use as tracers both artificially and naturally (26,35–38). For general review of the use of environmental isotopes in contaminant hydrogeology and potential future use of these tracers in vulnerability studies in groundwater, see References 2,27 and 28.

Environmental isotopes have been found widely applicable, especially in identifying the origin of water that is infiltrating an aquifer (28). Obviously, there is a relationship between recharge and the vulnerability of groundwater to contamination because the transport of many contaminants to the saturated zone of the aquifer occurs in the dissolved phase as part of the recharge process (39). Generally, ^{18}O and ^{2}H, among the various environmental isotopes, are most widely used in defining groundwater recharge areas as well as their origins. Changes in the isotopic relationships between $^{18}O/^{16}O$ and $^{2}H/^{1}H$, during the evaporation and condensation of water allow defining its origin (28).

New isotope applications are being developed in organic contaminant studies of groundwater. These studies have shown that carbon isotopes allow very sensitive evaluation of the fate of precursors and by-products of the biodegradation of chlorinated compounds, which are the most common organic contaminants found in groundwater. Potentially new applications include ^{18}O to evaluate the fate of oxygen in contaminant plumes and ^{11}B in salinization and nitrate studies of groundwater (2). Table 2 shows the potential/common applications of isotopes in groundwater pollution studies.

Boron-11, a stable isotope of boron, has also been used to differentiate nitrate from sewage and fertilizers as well as to trace the sources of salinity in groundwater. This approach was used to differentiate isotopic composition of injected treated wastewater and irrigation-affected water ($\delta^{11}B > 40‰$) distinctly and to trace the effect of injection into an alluvial aquifer near El Paso, Texas (40).

It has also been possible to approach the tracing of various organic pollutants (chlorinated hydrocarbons) in groundwater by using a combination of the isotopes $^{13}C/^{12}C$, $^{2}H/^{1}H$, and $^{37}Cl/^{35}Cl$. For example, the isotopic ratios $^{37}Cl/^{35}Cl$ and $^{10}B/^{11}B$ have been used recently in pollution studies of groundwater (41,42). The strontium isotopic ratio ($^{87}Sr/^{86}Sr$) is also a valuable tracer for studies related to the mixing and origin of mineralized groundwater. It is a tracer for the origin of salinity, groundwater movement, and water–rock interactions (43,44). In the past, a combination of stable and radioactive isotopes that has dissolved geochemical constituents has been used in tracing the migration of contaminants in groundwater (45–47).

According to Hirata and Varnier (28), isotope techniques are useful in determining groundwater pollution vulnerability, to the extent that they allow for better understanding of the degree of access provided for penetration of contaminants and the attenuative capacity offered by the geologic medium (Table 3).

The most frequently used environmental isotopes in contaminant hydrogeology include nitrogen isotopes ($^{14}N/^{15}N$), chlorine - 37, the sulphur isotope (^{34}S), carbon

Table 2. Common/Potential Isotopes Used in Groundwater Pollution Studies[a]

Isotope(s)	Potential/Common Applications
^{13}C in HCO_3^- and CH_4	Identifying of methane sources Monitoring CO_2 and CH_4 production in biodegradation studies
^{34}S and ^{18}O in SO_4^{2-}	Identifying of sources of pollution Acidification and acid mine drainage
^{15}N and ^{18}O in NO_3^-, NH_4^-, N_2	Origin of nitrates Sources of pollution Microbial denitrification processes
^{11}B in B $(OH)_4^-$ and B $(OH)_3^-$	Identifying of pollution sources Characteristics of brines, source of salinity
Chlorine-37 (^{37}Cl)	Identifying pollution sources Characteristics of brines, sources of salinity
Tritium (^3H)	Indicators of recent pollution/recharge Water transport in the unsaturated zone
Strontium isotope ratio ($^{87}Sr/^{86}Sr$)	Tracer for the origin of salinity, mixing, and origin of mineralized groundwater, groundwater movement, and water–rock interactions.
Uranium isotopes (^{234}U, ^{238}U)	Sensitive indicators of radioactive pollution in natural waters, Rock–water interaction, dating of old groundwater

[a] Modified from Reference 27.

Table 3. The Use of the Isotope Technique for Groundwater Pollution Vulnerability Characterization[a]

Component of Aquifer Pollution Vulnerability	Information Related to	Isotope Technique
Hydraulic accessibility of the saturated aquifer to penetration of pollutants	Identification of recharge/discharge areas Degree of aquifer confinement Definition of period of recharge Identification of paleowaters Interconnection of surface and groundwater	^{18}O in H_2O ^2H in H_2O ^{85}Kr flow through fissured rock ^3H in H_2O (environmental and applied tracer) ^{51}Cr, ^{58}Co, ^{60}Co, ^{131}I (applied tracers)
Attenuative capacity of strata overlying the saturated zone, resulting from the physicochemical retention or reaction of pollutants	Determination of transit time of water through the saturated and unsaturated zones (advective flow)	Dating young water ^3H ^3He Dating midage water ^{39}Ar ^{14}C, correction with ^{13}C in HCO_3^- ^{81}Kr Dating old water ^{81}Kr ^{234}U
	Identification of physico chemical reactions of pollutants (degradation)	^{15}N and ^{18}O in NO_3^-, NH_4^+, N_2, microbial denitrification processes ^{13}C, biodegradation of synthetic solvents and fuel ^{34}S, oxidation/reduction of sulfur
Pollution hazard factors	*Information related to*	*Isotope technique*
Definition of pollutant origin	Identification of source of pollution	^{34}S and ^{18}O in SO_4^{2-}, acidification and acid mine drainage ^{37}Cl, ^{11}B, in $B(OH)_4^-$ and $B(OH)_3^-$ characteristics of brines; sources of salinity ^{15}N and ^{18}O in NO_3^-, NH_4^+, N_2 ^{18}O, ^2H in H_2O, salinization mechanisms, recycling of irrigation water ^{14}C and ^{13}C for origin of sources of organic compounds

[a] After Reference 28.

isotopes (^{13}C, ^{14}C), and the boron isotope (^{11}B). Those of the water molecules (^{18}O, ^2H) are also sometimes used with some of these isotopes in groundwater contamination studies. Following are brief descriptions of the isotopes commonly employed in contaminant hydrogeologic studies.

Nitrogen Isotopes ($^{14}N/^{15}N$). The most common contaminant identified in groundwater is dissolved nitrogen in the form of nitrate (NO_3^-). Other forms of inorganic nitrogen are nitrite (NO_2^-), nitrogen gas (N_2), ammonia (NH_4^+), and cyanide (CN^-). Nitrogen is also a major constituent of organic matter as amino acids (7). These contaminants are becoming widespread due to increased agricultural activities worldwide and disposal of sewage on or beneath the land surface.

Nitrogen occurs as two isotopes in nature: ^{14}N(≈99.6%) and ^{15}N(≈0.36%). ^{14}N is more abundant in the atmosphere than ^{15}N. Hence, the ^{15}N/^{14}N ratio in nitrate can be used to distinguish the origin of nitrates either from animal/human wastes or from the application of chemical fertilizers (48,49). The origin of contamination by nitrates can be determined through the relationship between ^{15}N/^{14}N and standard values (δ^{15}N). Combined isotope analyses of nitrogen and oxygen in NO_3^- leave fingerprints on natural and anthropogenic sources of nitrate; on microbial denitrification, nitrification and biological fixation processes; and on the nitrogen budget in groundwater (50–53).

Therefore, the δ^{15}N values of dissolved nitrates, ammonium, and organic nitrogen in soil water are well distinguished from one region to another. The δ^{15}N of animal waste is between +5 to 15‰, and fertilizers have a δ^{15}N of about 0‰. However, for nitrogenous fertilizers, these values lie between −4‰ and +4‰ (50,54).

For more information and related nitrogenous inputs in groundwater, refer to Foster and Crease (55), Kreitler and Jones (56), Létolle (57), Kreitler and Browning (58), White (59), Flipse et al. (60), Heaton (60), Mariotti (61), Flipse and Bonner (62), Tessendorff (63), Hübner (64), Mariotti (65), Böttcher et al. (51), Exner and Spalding (66), Aravena et al. (52), Buzek (67), Révész et al. (68), Pauwels et al. (69), Tore et al. (70), and Anwar and Ghosh (71).

Chlorine-37. Naturally, chlorine exists with two common stable isotopes: ^{35}Cl (≈75.7%) and ^{37}Cl (≈24.2%). Chloride ions are generally not reactive and as such, do not participate in redox reactions, are not sorbed onto mineral or organic surfaces, and do not form insoluble precipitates. As a result, ^{37}Cl is sometimes used as a conservative tracer in groundwater studies. Data are expressed as δ^{37}Cl with respect to the standard mean oceanic chloride (SMOC). The ratio ^{37}Cl/^{35}Cl is measured by isotope ratio mass spectrometry (IRMS). Precision must be better than ± 0.1‰ (72). δ^{37}Cl values contain information on the origin of chloride ions in fresh and polluted groundwater as well as in subsurface brines (30,41,42).

In geochemical studies generally, chlorine isotopes are not highly fractionated, but research has shown that measurable fractionations can be used to distinguish sources of groundwater mixing (73). The ^{37}Cl content of chloro contaminants in groundwater can fingerprint the sources of contaminants and salinity. It is therefore a potential tool in contaminant hydrogeology (27,43).

Sulphur Isotope (^{34}S). Sulphur has four stable isotopes: ^{32}S (95.02%), ^{33}S (0.75%), ^{34}S (4.21%), and ^{36}S (0.02%). The abundance ratio of ^{34}S and ^{32}S is generally given as a δ^{34}S value. Iron sulfide from the troilite phase of the Diablo Canyon iron meteorite (DCT) (that has a ^{32}S/^{34}S ratio of 22.220) is conventionally used as a standard (72). ^{34}S is one of the distinct isotopes that can be used to characterize different sources of sulfate contamination. For example, high sulfate in aquifers in areas close to mine tailings could be associated with the oxidation of sulfide minerals in the tailings. However, other sources of sulfate in aquifers such as gypsum of marine origin, input from volcanic sources, or sulfate from recirculated water from agriculture have been identified (2). Generally, sulfate produced from sulfide minerals is much more depleted in ^{34}S and ^{18}O than marine sulfate (74,75).

Moreover, owing to isotope fractionation during sulfate reduction, ^{34}S and ^{18}O in sulfate can also be used to evaluate the fate of sulfate in groundwater flow systems (76). The isotopic composition of sulfate can also be an indicator of anthropogenic pollution of groundwater (75).

Carbon Isotopes (^{13}C, ^{14}C). Carbon has three isotopes in nature: ^{12}C, ^{13}C, and ^{14}C. The heavy carbon isotope, ^{14}C, is unstable and decays radioactively into ^{14}N. The stable isotopes of carbon, ^{13}C and ^{12}C, have an average ratio of about 1:100 (72). The isotopic composition of carbon in the dissolved carbon constituents of groundwater is very variable. The sources of carbon dissolved in groundwater are soil CO_2, CO_2 of geogenic origin or from magmatic CO_2 (from deep crustal or mantle sources) or in fluid inclusions, living and dead organic matter in soils and rocks, methane, and carbonate minerals (53). Each of these sources has a different carbon isotopic composition and contributes to total dissolved carbon in various proportions. Therefore, the isotopic composition of dissolved inorganic carbon compounds in groundwater has a wide range of δ^{13}C values.

Carbon isotopes (^{13}C, ^{14}C) can be used to evaluate the origin of methane in landfills (2). The origin of methane is a common problem in landfills located in old quarries. Methane in this type of environment can be thermocatalytic in origin, associated with gases present in the bedrock or/and originating in the landfill itself (2). According to Schoell (77), gases of thermocatalytic origin are generally more enriched in carbon-13 ($\delta^{13}C > -50$‰) than methane of biogenic origin ($\delta^{13}C < -50$‰). In addition, Hackley et al. (78) proved that catalytic gases are devoid of carbon-14, whereas methane in landfills has a ^{14}C content higher than 100% modern carbon (pmc). Obviously, due to carbon-14 from thermonuclear explosions, any carbon in domestic waste that was deposited in a landfill after 1950 is usually characterized by a ^{14}C content higher than 100 pmc (2). Furthermore, carbon isotopes have also been employed in monitoring CO_2 and CH_4 production in biodegration studies in unsaturated areas impacted by organic contamination (79,80).

Boron Isotope (^{11}B). Natural boron has two stable isotopes, ^{11}B (≈80%) and ^{10}B (≈20%). The variation in the ratio of these two isotopes is expressed in δ^{11}B (‰) with respect to SRM-951 NBS standard (81). Generally, boron is present in groundwater mainly as trigonal undissociated boric acid, $B\ (OH)_3{}^0$, and its conjugate base, tetrahedral borate, $B\ (OH)_4{}^-$. Tetrahedrally coordinated boron is more isotopically depleted (i.e., lower in δ^{11}B) relative to trigonally coordinated boron; thus the δ^{11}B of dissolved boron in groundwater is usually higher than that in aquifer rocks (81).

Boron is used as tracer in groundwater because of its high solubility in aqueous solution, natural abundance in all waters, and the lack of effects from evaporation, volatilization, and oxidation–reduction reactions. The concentrations of boron in pristine groundwaters are generally low (<0.05 mg/L), and contaminant sources are usually enriched in boron (e.g., sea water: 4.6 mg/L; domestic wastewater: 1 mg/L; flyash leachate: 14 mg/L; landfill leachate: 6 mg/L), so the δ^{11}B of groundwater is highly sensitive to the impact of contamination (81). The large isotopic variation of the potential sources can be used to trace the origin of contamination. Vengosh (81) reported that coastal aquifers affected by seawater intrusion and marine-derived fossil brines are characterized by high δ^{11}B values (up to 60‰). In contrast, domestic wastewater that has relatively high boron concentrations and a high B/Cl ratio has δ^{11}B values of 0‰ to 12.9‰. Boron stable isotope ratios have a potential to play a role in pollution studies (82). They also have applications in the characterization of brines and geothermal waters (83).

It will suffice at this point to say that some of these environmental isotopes are in themselves a source of groundwater contamination, especially radioactive isotopes. Radioactive isotopes, also known as radionuclides, are therefore discussed in the following section in relation to groundwater contamination.

Radioactive Contaminants

It is now well known that numerous human activities have created nuclear isotopes by detonating nuclear weapons and constructing nuclear reactors. A number of radionuclides are produced in the earth's atmosphere by nuclear reactions caused by cosmic rays. Cationic radionuclides may be subjected to ion exchange and other processes that sorb the radionuclides onto mineral organic surfaces in the soil (7). Examples of transition metals that have large distribution coefficients and hence low mobilities in waters are listed in References 7 and 84. ^{90}Sr, ^{137}Cs, and ^{226}Ra are involved in cation exchange similar to other exchangeable cations, for example, Ca^{2+} and Mg^{2+} (84).

Many of these radionuclides are frequently produced as fission products from the decay of ^{235}U and ^{239} Pu and from neutron activation of stable isotopes. These processes result in the production of more than 75 radionuclides such as ^{137}Cs, ^{134}Cs, ^{58}Co, ^{51}Cr, ^{54}Mn, ^{55}Fe, ^{3}H, and ^{131}I (1). Table 4 lists the various sources of environmentally important radioactive isotopes. These are elements that undergo spontaneous decay resulting in the release of energy as well as energetic particles and consequent formation of different isotopes. They occur either naturally or are created by bombardment of the earth by cosmic radiation (7) and undergo spontaneous decay resulting in the release of energy as well as energetic particles and consequent formation of different isotopes.

The nuclear industry has been the main generator of radioactive contaminants. However, potential sources occur throughout the nuclear fuel cycle, which involves mining and milling uranium, uranium enrichment and fuel fabrication, power plant operation, fuel reprocessing, and waste disposal (1). In the examples presented by Domenico and Schwartz, the United States, Japan, France, and Germany have large numbers of light-water reactors that use enriched uranium (^{235}U) as the predominant fuel source and possibly ^{239} Pu and ^{238}U. The process of enrichment and fuel fabrication is also accompanied by generation of the following radionuclides: ^{238}U, ^{235}U, ^{137}Cs, and ^{90}Sr, which are common contaminants.

Furthermore, there are potential radionuclide contaminants that result from the processing of raw uranium ore (during mining operation) alongside other nonradioactive pollutants, which include ^{238}U, ^{230}Th, ^{226}Ra, and ^{222}Rn. Examples of locations where this type of contamination occurs are Colorado, New Mexico, Texas, Utah, Wyoming, northern Saskatchewan, and Ontario (1).

PRACTICAL APPLICATIONS OF ENVIRONMENTAL ISOTOPES IN SOLVING GROUNDWATER CONTAMINATION PROBLEMS

Many examples exist in the literature on the use of environmental isotopes in solving groundwater contamination problems. There are also case studies in publications, which have addressed the applications of isotopes in contaminant hydrogeology. In this section, the results of the application of isotopes in groundwater contamination studies, particularly in solving specific problems are summarized. The case studies cut across continents and varied hydrogeologic settings. Some of these applications are based on single isotopes, and others combine several isotopes, particularly oxygen with nitrogen and sulfur to understand the origin and migration of contaminants in groundwater. However, some case studies applied an integrated approach (using isotope techniques and conventional hydrochemistry) to interpret contaminant sources and pathways. This is not an exhaustive listing but rather selected to demonstrate the applications of environmental isotopes in solving specific hydrogeologic problems.

Table 4. Various Sources of Environmentally Important Radioactive Isotopes[a]

Source	Radionuclides
Naturally occurring	^{40}K, ^{222}Rn, ^{226}Ra, ^{230,232}Th, ^{235,238}U
Cosmic irradiation	^{3}H, ^{7}Be, ^{14}C, ^{22}Na
Nuclear weapon tests	^{3}H, ^{90}Sr, ^{137}Cs, ^{239,240} Pu
Mining wastes—uranium, phosphate, coal	^{222}Rn, ^{226}Ra, ^{230,232}Th, ^{235,238}U
Industrial wastes—nuclear power plants, weapons manufacturing, research and medical waste	^{59,63}Ni, ^{60}Co, ^{90}Sr, ^{93,99}Zr, ^{99}Tc, ^{107}Pd, ^{129}I, ^{137}Cs, ^{144}Ce, ^{151}Sm, ^{152,154} Eu, ^{237}Np, ^{239,240,242} Pu, ^{241,243}Am

[a] Reference 7.

For more information and examples, refer to the IAEA TECDOC series and other edited specialist volumes.

Fritz et al. (45) studied the possible movement of leachate from a sanitary landfill in Frankfurt am Main, Germany. Interpretation of results of both isotopic and chemical data in the pollution monitoring showed that $\delta^{2}H$ and $\delta^{18}O$ are most useful in tracing the leachate and mapping the pollution plume. Three groups were identified from $\delta^{18}O$ data: The first group (−8.6‰ to −8.0‰) is of uncontaminated groundwater; the second group (−6.7‰ to −6.3‰) is polluted groundwater, and the third group (−3.0‰ to −2.0‰) reflects recharge from lake water. The same $\delta^{18}O$ data have also been used (by these researchers) to calculate the percentage of leachate water in each well, and this was helpful in producing a map of the pollution plume at the landfill site.

Mazor et al. (85) conducted a study of groundwater contamination by a slaughterhouse at Lobatse, Botswana. The results of this study revealed that tritium and chloride were the most conservative tracers in this case. The slaughterhouse used water from a dam, and this water (when sampled and analyzed) had tritium concentrations that were significantly higher than those in the local groundwater. High tritium concentrations/contamination were observed in wells near the slaughterhouse and along the Peleng riverbed into which the effluent was released. Other wells (according to the researchers) showed low tritium and chlorine values, indicating that they were not contaminated.

Egboka (46), in the investigation of problems posed by contaminants in the hydrologic environment, used a combination of tracers: (radioisotopes, such as bomb tritium), stable isotopes (^{18}O and ^{2}H), and dissolved geochemical constituents (Cl, SO_4, NO_3 and some major cations). These isotopes, in combination with the dissolved geochemical constituents, have been used to delineate the aerial extent of spread of leachate plumes in landfills and spillage zones and around sewage lagoons. He presented two case studies where this multitechnique approach was used successfully to investigate environmental pollution and contamination. These include using bomb tritium, ^{18}O, ^{2}H, Cl^{-}, and $SO_4{}^{2-}$ to investigate the leachate migration in an aquifer beneath an abandoned landfill at CFB Borden, Ontario (Canada) and the Lake Chad drainage basin (NE Nigeria), where ^{3}H, ^{18}O, ^{2}H, ^{14}C, and ^{13}C were extensively analyzed.

Song et al. (86) used ^{34}S in 66 water samples and 30 rock (anhydrite) samples to determine the origin of sulfate in the stable isotopic composition and geochemistry of fresh water and brine in Sichuan Basin, China. The results indicate that $\delta^{34}S$ values of saline waters are similar to those in the samples collected from different sections of the same sequence and differ from different strata. These authors considered that the evaporate of the early Triassic was formed in an enclosed and reductive beach and the facies belongs to the beach salt-lake in the environment of a supratidal zone and that of the middle Triassic belongs to lagoon-salt-lakes in the intertidal environment. The consistency of $\delta^{34}S$ in water and rock implies that they have a common source of sulphur.

Wassenaar (87) carried out a representative study in Abbotsford aquifer (a regional aquifer in Canada characterized by nitrate concentrations between 80 and 160 mg/L) using the isotopes of ^{15}N and ^{18}O. In this case, it was well documented that the main sources of nitrate were manure produced in the poultry industry and inorganic fertilizers used in cultivated fields. Based on the ^{15}N data, it was shown that most of the nitrate in the contaminated groundwater has $\delta^{15}N$ values greater than +8‰. Wassenaar (87) demonstrated that the main source of this nitrate was the manure piles that were deposited on top of the aquifer.

Similarly, a nitrate study carried out by Auge and Nagy (88) in the city of La Plata (Argentina) illustrates one of the applications of nitrogen isotopes in Latin America. The isotope study, which was part of a comprehensive hydrogeologic study of this aquifer, aimed to provide information about the sources of high nitrate observed in production wells in the city of La Plata. According to these researchers, the main potential sources of nitrate in this area are fertilizers and manure from agricultural and pasture areas located upgradient of the city of La Plata and leakage from the sewage system and possibly landfills located downgradient of the center of the city. A regional nitrate study showed a plume that had high nitrate concentration below the city. From the result, the nitrogen isotope pattern showed a clear trend toward enriched $\delta^{15}N$ values correlated with an increase in nitrate concentration. The high nitrate concentration (80–120 mg/L as nitrate) and enriched $\delta^{15}N$ values (14–15‰) correspond to groundwater located below the city. The low nitrate concentration (15–40 mg/L) and depleted $\delta^{15}N$ values (7–11‰) correspond to groundwater representing the agricultural areas and wells located upgradient on the high nitrate concentration plume. The enriched $\delta^{15}N$ values indicated input of nitrate from an enriched ^{15}N source associated with domestic wastes.

Aravena and Robertson (38) used multiple isotope tracers to evaluate denitrification in groundwater and documented that septic systems are point sources of nitrate in groundwater, which was documented in a nitrate plume generated in a large communal septic system operating at a campsite. As a result of oxygen consumption by the oxidation of pyrite and organic carbon in the unsaturated zone, the groundwater became anaerobic in the shallow part of the flow system. A steady decrease in nitrate concentrations from values around 80 mg N/L at the water table to nondetectable was observed versus depth along the groundwater system. A trend toward enriched $\delta^{15}N$ values accompanied these concentration changes, reaching values as high as +58‰, demonstrated that nitrate attenuation was controlled by denitrification. A linear correlation between $\delta^{15}N$ and $\delta^{18}O$ values in nitrate also supported denitrification as the mechanism for nitrate attenuation in the aquifer.

Tandia et al. (89), in the investigation of the origin, process, and migration of nitrate compounds in the aquifers of the Dakar region, Senegal, used environmental isotopes (^{18}O, ^{2}H, ^{3}H, ^{13}C, ^{14}C) along with major inorganic constituents and fecal coliform on yearly water samples. The samples were taken from piezometers, hand-pumped wells, and dug wells in different areas of the Peninsula from 1987 to 1997. Results showed that only a limited area is affected by nitrate pollution in the confined layer. Stable isotopes, tritium and ^{14}C, gave a better understanding of nitrate contamination of groundwater in this region; tritium in groundwater indicates that nitrate concentration is recent. This has been confirmed by the presence of high activities of ^{14}C (80‰), in good agreement with recent analyses of recharge groundwater.

The distribution of the values of ^{18}O and ^{2}H (89) around the World Meteoric Water Line indicate that the recharge of the Dakar aquifer system originates from the monsoon rainfall after it has been exposed to evaporation. However, the cluster of points relating to ^{18}O and ^{2}H were also interpreted as contributions of nitrate contamination along the flow path from east to west.

Buzek et al. (67) investigated the nitrate pollution of a karstic groundwater system in Svaty Jan Pod Skalou in the Czech Republic using $\delta^{18}O$ in water and $\delta^{15}N$ in nitrate. Increasing agricultural activity in the republic after the 1960s resulted in contaminating both shallow and deep-water resources. Therefore, to identify the sources of nitrate contamination and evaluate the parameters of the karstic system, the stable isotopes $\delta^{15}N$ and $\delta^{18}O$ were used together with monitoring of the flow dynamics, water chemistry, and soil composition. The results of the isotopic investigation revealed the sources of nitrate contamination from $\delta^{15}N$ in soil solutes at sites that had different agronomic records (sequential use of organic and inorganic fertilizers). They further showed that spring-discharge nitrate is formed from an atmospheric source (an average nitrate content of 8 mg/L and $\delta^{15}N$ of −6‰) and two groundwater components that had nitrate content of 50 to 60 mg/L and $\delta^{15}N$ values from 2 to 7‰. However, the contribution of groundwater components varied in time more frequently than can be described by the model developed for flow dynamics and ^{18}O. From this case study, data from the flow dynamics, ^{18}O, and tritium gave sufficient information to estimate the volume of the karstic system and the residence times of the three contributions (direct infiltration, fast, and old components) to spring discharge. The estimated turnover time of the old component (about 22 years) could be used for calculating the spring nitrate content in the future.

Aggarwal et al. (90) initiated a study using a suite of different isotopes to fingerprint and determine the source of contaminants in groundwater systems in New Zealand. Four different isotope systems were examined: boron isotopes, oxygen and nitrogen isotopes in nitrates, oxygen isotopes in sulfates, and oxygen and sulfur isotopes in sulfates. In a case study of the semirural Manakau area as well as the South and North Islands (all in New Zealand) conducted to investigate the possibility of using nitrate isotopes to detect the source of nitrate contamination, results showed nitrogen isotope (+4 to +12‰) coupled with oxygen isotope measurements (+5 to +9‰) and demonstrated that nitrogen is not sourced from fertilizers but from some combination of septic and animal waste. However, for the case study of sulfate isotope used, sulfur and oxygen isotopic compositions of sulfate in river and lake water from seven major catchments of the North and South Islands of New Zealand were determined. The results of the isotope analyses allowed distinguishing between natural (geologic, geothermal, and volcanic) and anthropogenic (fertilizer) sulfur sources. Boron (B) isotope measurements were also carried out in the Bartholomew (Manawatu) and Petone waters in New Zealand. Results showed that the Cl/B ratio and the $\delta^{11}B$ of the water from Bartholomew may be interpreted as a consequence of mixing between saline water and freshwater. Seawater (included for comparison) showed a distinctly heavy B isotopic composition, relatively high B concentrations, and a high Cl/B ratio, whereas these two freshwaters contain a variable B isotope ratio that has a low B concentration and a much lower Cl/B ratio than that in the seawater.

Groundwater contamination in coastal aquifers of Israel, Italy, and the United States was determined by boron isotope systematics (81). This work reports isotopic results of contaminated groundwater from the coastal aquifers of the Mediterranean in Israel, the Cornia River in northwestern Italy, and the Salinas Valley in California. The interpretation of the results revealed different contamination sources; underlying saline water of a marine origin in saline plumes in the Mediterranean coastal aquifer of Israel ($\delta^{11}B = 31.7$ to 49.9‰; B/Cl ratio $\sim 1.5 \times 10^{-3}$), mixing of fresh and sea water ($\delta^{11}B = 25$ to 38‰; B/Cl ratio7×10^{-3}) in saline water associated with a saltwater intrusion into the Salinas Valley, California, and a hydrothermal contribution (high B/Cl of ~ 0.03; $\delta^{11}B = 2.4$ to 9.3‰) in groundwater from the Cornia River, Italy. Vengosh et al. (81) observed that the boron concentrations in pristine groundwaters are generally low (<0.05 mg/L), whereas contaminant sources (e.g., domestic wastewater) are enriched in boron; hence, boron isotopes can be used to evaluate the impact of anthropogenic boron on the environment. Moreover, the isotopic composition of contaminated groundwater has revealed the sources of pollution because different sources are characterized by distinguishable isotopic ratios. For example, seawater ($\delta^{11}B = 39$‰), saltwater intrusion, and marine-derived brines (40 to 60‰) sharply differ from hydrothermal fluids ($\delta^{11}B = -10$‰ to 10‰) and anthropogenic sources (sewage effluent: $\delta^{11}B = 0$‰ to 10‰; boron-fertilizer: $\delta^{11}B = -15$‰ to 7‰). This research further showed that some differences (up to 15‰) might exist between domestic wastewater and boron fertilizer.

Shivanna et al. (91) used environmental isotopes (^{18}O, ^{2}H, ^{3}H, ^{14}C, and ^{34}S) along with hydrogeology and hydrochemistry to investigate (1) the source of salinity and the origin of sulfate in the groundwater of coastal Orissa, Orissa state (India) and (2) to study the source of salinity in deep and saline groundwaters of charnockite terrain at Kokkilimedu, south of Chennai, India. The results of the first case indicate that the shallow groundwater (depth <50 m) is fresh and modern. Groundwater salinity in the intermediate aquifer (50−100 m) is due to the Flandrian transgression during the Holocene period. The saline groundwater found below the fresh deep aquifer has marine water entrapped during the late Pleistocene. The source of high sulfate in the groundwater (as identified by the researchers) is of marine origin. In most of the deep aquifers, saline samples showed enriched $\delta^{34}S$ values

(+20.8‰ to +43.8‰) compared with the seawater value, and radiocarbon ages are in the range of 12, 000 to 24, 000 years BP with negligible tritium.

However, in the second case, the results of isotope investigation showed that there are two different sources of salinity in this area: one, the salinity attributed to infiltration of seawater together with the contribution of solute from pseudotachylites observed along the fractures and shear zones in the host rock; second, salinity attributed to solutes mostly from sheared and fractured charnockite rock. The high tritium content observed in this case was interpreted as due mainly to effluent from the nuclear power plant located 0.5 km west of the site and partly due to mixing of surface water with groundwater.

Halas and Trembaczowski (92) applied stable isotopic analysis of oxygen and sulfur in sulfate to trace the sources of groundwater pollutants in the industrial area of Belchatow, central Poland, where a large brown coal deposit is exploited and the coal is burned in an electric power plant. The interpretation of isotope ratios allowed recognizing three groups of sulfates:

1. from the leaching of a Permian salt dome,
2. produced by the leaching of soluble sulfates from an ash pool, and
3. produced by the oxidation of natural sulfides in water-bearing rocks.

It was obvious from the data presented in this work that the oxygen and sulfur isotopic ratios of the dissolved sulfates are informative of their origin.

Liu et al. (93) carried out an investigation of groundwater contamination due to uranium deposits in the Zhoujiashan district of the Xiangshan uranium ore field, China. Samples of groundwater collected from five completed exploration boreholes in the area revealed the pollution source. It was traced and analyzed by using isotope as well as radiohydrochemical techniques. The researchers experimentally proved that the uranium concentration and uranium isotope ratio $^{234}U/^{238}U$ in natural waters are two sensitive indicators of radioactive pollution in natural waters. It was concluded that, under present conditions, exploration of uranium deposits might not cause serious groundwater contamination by radioactive elements (U, Ra, Rn, and Tn). Activity ratios $^{226}Ra/^{228}Ra$ and $^{230}Th/^{232}Th$ in water can be used only as supplementary indicators.

Qureshi et al. (94) used stable carbon isotope ratios as water quality indicators in the coastal areas of Karachi, Pakistan. Stable carbon isotope ratios ($\delta^{13}C$) of total dissolved inorganic carbon (TDIC), total organic and inorganic carbon in bottom sediments, as well as sea plants in polluted water sources, nonpolluted Karachi seawater, and pollution recipients were used to elaborate the pollution scenario of the shallow marine environment (off the Karachi coast). Groundwater pollution in the coastal areas of Karachi is also reflected in the stable isotope composition of carbon ($\delta^{13}C$) in the TDIC pool of organic carbon in bottom sediments, mangrove leaves, and seaweed. These results were supplemented with stable isotope analysis of nitrogen ($\delta^{15}N$) in seaweed and mangroves. Isotopic data revealed that the mangrove ecosystem and its tidal fluctuations play a key role in controlling contamination inventories in the shallow seawater off the Karachi coast, specifically the Manora Channel. The Karachi harbor zone was reportedly the most heavily polluted marine site in Manora Channel during high as well as low tide.

De Vito et al. (95) carried out a study in a 200 meter long wetland located between agricultural fields and the Boyne River in southern Ontario (Canada) using ^{15}N and ^{18}O (in nitrate) to evaluate the role of denitrification as an attenuative mechanism in nitrate. Results showed that nitrate levels of 30 and 40 mg N/L characterize the groundwater impacting the wetland. A significant change in nitrate concentration was observed by the researchers along the groundwater flow system that reached values as low as 1 mg N/L in areas close to the river. This pattern is accompanied by an isotope trend toward enriched ^{15}N and ^{18}O. The isotopic data and the dissolved organic carbon that increases along the groundwater flow system were interpreted as demonstrating that denitrification is the main process responsible for nitrate attenuation in the area.

According to Aravena (2), applications of isotopes in organic contaminants have focused on evaluating carbon and hydrogen isotope fractionation in biological and physical processes that can affect organic contaminant concentration in groundwater. The results of these studies have shown that very little isotopic fractionation occurred during physical processes (96,97). However, recent analytical developments in stable isotope research have shown that carbon isotopes are very sensitive tracers for evaluating the biodegradation of chlorinated compounds in groundwater (31,32). However, they are not sensitive enough to document biodegradation of petroleum hydrocarbons in groundwater (98,99).

BIBLIOGRAPHY

1. Domenico, P.A. and Schwartz, F.W. (1998). *Physical and Chemical Hydrogeology* (Chapter 17), 2nd Edn. John Wiley & Sons, New York, p. 506.
2. Aravena, R. (2002). Use of environmental isotopes in contaminant studies in groundwater. In: *Groundwater and Human Development*. E. Bocanegra, D. Martinez, and H. Massone (Eds.). Mar de Plata, Argentina, pp. 293–298.
3. Schwille, F. (1981). Groundwater pollution in porous media by fluids immiscible with water. *The Science of the Total Environment* **21**: 173–185.
4. Schwille, F. (1984). Migration of organic fluids immiscible with water in the unsaturated zone. In: *Pollutant in Porous Media*. B. Yanon, G. Dagen, and J. Goldshmid (Eds.). Springer-Verlag, Berlin, pp. 27–48.
5. Schwille, F. (1988). *Dense Chlorinated Solvents in Porous and Fractured Media*. Translated by Pankow, J.F., Lewis Publishers, Chelsea, MI, p. 146.
6. McWhorter, D.B. and Sunada, D.K. (1990). Exact integral solutions for two-phase flow. *Water Res. Research* **26**(3): 399–413.
7. Fetter, C.W. (1993). *Contaminant Hydrogeology*. Prentice-Hall, Englewood Cliffs, NJ, p. 458.

8. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ, p. 604.
9. Pankow, A. and Cherry, J.A. (1996). *Dense Chlorinated Solvents and other DNAPLs in Groundwater*. Waterloo Press, Waterloo, Canada.
10. Forstner, U. and Wittman, T.W. (1981). *Metal Pollution in Aquatic Environment*. Springer-Verlag, Berlin, p. 486.
11. Seiler, K.P. (2000). Man's impact on groundwater systems. In: *Environmental Isotopes in the Hydrological Cycle (Principles and Applications)*. W.G. Mook (Ed.). Vol. V, UNESCO/IAEA Publ. Series, Paris/Vienna, p. 105.
12. Baetsle, L.H. (1969). Migration of radionuclides in porous media. In: *Progress in Nuclear Energy, Series XII, Health Physics*. A.M.F. Duhamel (Ed.). Pergamon Press, New York, pp. 707–730.
13. Fuhriman, D.K. and Barton, J.R., (1971). *Ground water pollution in Arizona, California, Nevada, and Utah*. Environ. Protect. Agency. Rept. 16060 ERU.
14. Hall, E.S. (1972). Some chemical principles of groundwater pollution. In: *Groundwater Pollution in Europe*. J.A. Cole (Ed.). Water Information Centre, Port Washington, NY, pp. 269–271.
15. Scalf, M.R., Keley, J.W., and Lafevers, C.J. (1973). *Groundwater pollution in the south central States*. U.S. Environ. Protect. Agency Rept. EPA -122-73-268.
16. Miller, D.W., Deluca, F.A., and Tessier, T.L. (1974). *Groundwater contamination in the northeast States*. U.S. Environ. Protect. Agency Rept. EPA 660/2-74-056.
17. Summers, W.K. and Spiegal, Z. (1974). *Groundwater Pollution: A Bibliography*. Ann Arbor Science Publishers, Ann Arbor, MI.
18. Cherry, J.A., Gillham, R.W., and Pickens, J.F. (1975). Contaminant hydrogeology: Part I: Physical processes. *Geosci. Can.* **2**: 76–84.
19. Back, W. and Cherry, J.A. (1976). Chemical aspects of present and future hydrogeologic problems. *Proc. Symp. Adv. Groundwater Hydrol. America Water Resour. Assoc*.
20. Fried, J.J. (1976). *Ground Water Pollution*. Elsevier, Amsterdam, pp. 1–47.
21. Todd, D.K. and McNulty, D.E.O. (1976). *Polluted Groundwater*. Water Information Centre, Port Washington, NY.
22. Wilson, J.L., Lenton, R.L., and Porter, J. (Eds.). (1976). *Groundwater pollution: technology economics and management*. Dept. Civil Eng., M.I.T., Rept. TR208.
23. International Atomic Energy Agency (IAEA). (1980). *Nuclear Techniques in Groundwater Pollution Research*. IAEA, Vienna, p. 286.
24. International Atomic Energy Agency (IAEA). (1983). *Isotope Techniques in the Hydrological Assessment of Potential Sites for the Disposal of High-level Radioactive Wastes*. IAEA, Vienna, p. 164.
25. Pearson, F.J., Balderer, W., Loosli, H.H., Lehmann, B.E., Matter, A., Peters, T.J., Schmassmann, H., and Gautschi, A. (1991). *Applied Isotope Hydrogeology, a Case Study in Northern Switzerland, Studies in Environ. Sci. 43*. Elsevier, Amsterdam, p. 439.
26. Kendal, C. and McDonnell, J.J. (1998). *Isotope Tracers in Catchment Hydrology*. Elsevier, Amsterdam, Lausanne, New York, p. 839.
27. Gourcy, L. and Aggarwal, P. (2002). Strengthening the use of isotope techniques for a sustainable groundwater management. In: *Groundwater and Human Development*. E. Bocanegra, D. Martinez, and H. Massone (Eds.). Mar de Plata, Argentina, pp. 299–304.
28. Hirata, R. and Varnier, C. (2002). The use of isotopic techniques in determining groundwater pollution vulnerability—A Latin America perspective. In: *Groundwater and Human Development*. E. Bocanegra, D. Martinez, and H. Massone (Eds.). Mar de Plata, Argentina, pp. 305–313.
29. Hunkeler, D. and Aravena, R. (2000). Determination of stable carbon isotope ratios of dissolved chlorinated solvents by GC-C-IRMS: Comparison of headspace and solid-phase micro extraction. *Environ. Sci. Technol*. **34**: 2839–2844.
30. van Warmerdam, E.M. et al. (1995). Stable chlorine and carbon isotope measurements of selected chlorinated organic solvents. *Applied Geochemistry* **10**: 547–552.
31. Hunkeler, D., Aravena, R., and Butler, B. (1999). Monitoring of microbial dechlorination of tetrachloroethene (PCE) in groundwater using compound-specific stable carbon isotope ratios: Microcosm and field studies. *Environ. Sci. Technol.* **33**: 2733–2739.
32. Slater, G.F., Sherwood Lollar, B., Sleep, B., and Edwards, E. (2001). Variability in carbon isotopic fractionation during biodegradation of chlorinated ethenes: Implications for field applications. *Environ. Sci. Technol.* **35**: 901–907.
33. Fritz, P. and Fontes, J.Ch. (1980). *Handbook of Environmental Isotope Geochemistry. Vol. 1, The Terrestrial Environment*, A. p. 545.
34. Hiscock, K.M., Dennis, P.F., Feast, N.A., and Fairbaim, J.D. (1997). Experience in the use of stable isotopes to distinguish groundwater contamination from leaking sewers in urban areas. In: *Groundwater in the Urban Area*, *Proc. of the XXVI IAH Congress*. Vol. 1, Balkema, Rotterdam, Brookfield, pp. 427–432.
35. Nuti, S., Calore, C., and Noto, P. (1981). Use of environmental isotopes as natural tracers in a reinjection experiment at Larderello. In: *Proc. 7th Stanford Geothermal Reservoir Engineering Workshop*, pp. 85–89.
36. International Atomic Energy Agency (IAEA). (1991). *Use of Artificial Tracers in Hydrology*. IAEA-TECDOC-601, Vienna, p. 229.
37. International Atomic Energy Agency (IAEA). (1998). *Application of Isotope Techniques to Investigate Groundwater Pollution*. IAEA-TECDOC-1046, Vienna, p. 262.
38. Aravena, R. and Robertson, W. (1998). The use of multiple isotope tracers to evaluate denitrification in groundwater: A case study in a large septic system plume. *Ground Water* **36**: 975–982.
39. Foster, S. (1998). Groundwater recharge and pollution vulnerability of British aquifers: a critical overview. In: *Groundwater Pollution, Aquifer Recharge and Vulnerability*. N. Robins (Ed.). Special Publication. Geological Society, London, 130, pp. 7–22.
40. Vengosh, A., Heumann, K.G., Juraske, S., and Kasher, R. (1994). Boron isotope application for tracing sources of contamination on groundwater. *Environ. Sci. Technol.* **28**: 1968–1974.
41. Eggenkamp, H.G.M. (1994). The geochemistry of chlorine isotopes. *Geologica Ultraiectina, Mededelingen van de Faculteit Aardwetenschappen, Rijksunversiteit*. Utrecht, p. 116.
42. Frape, S.K., Bryant, G., Blomquist, R., and Ruskeeniemi, T. (1995). Evidence from stable chlorine isotopes for multiple sources of chlorine in groundwaters from crystalline shield environments. In: *Isotopes in Water Resources Management*. Vol. I, IAEA, Vienna, pp. 19–30.
43. Fritz, B., Clauer, N., and Sam, M. (1987). Strontium isotopic data and geochemical calculations as indicators of the origin of saline waters in crystalline rocks. In: *Saline Water and*

Gases in Crystalline Rocks. P. Fritz and S.R. Frape (Eds.). Special Paper 33, Geol. Assoc. of Canada, pp. 121–126.

44. McNutt, R.H. (1987). Sr^{87}/Sr^{86} as indicators for water-rock interactions: applications to brines found in Precambrian age rocks from Canada. In: *Saline Water and Gases in Crystalline Rocks*. P. Fritz and S.K. Frape (Eds.). Geol. Assoc. of Canada, Special Paper 33, pp. 81–88.
45. Fritz, P., Cherry, J.A., Weyer, R.V., and Sklash, M.G. (1976). Runoff analyses using environmental isotopes and major ions. In: *Interpretation of Environmental Isotope and Hydrochemical Data in Groundwater Hydrology*. IAEA, Vienna, pp. 111–130.
46. Egboka, B.C.E. (1981). Multitechnique approach using environmental tracers in hydrogeological studies. *J. Mining Geol.* **18**(1): 96–102.
47. Egboka, B.C.E., Cherry, J.A., Farvolden, R.N., and Frind, E.O. (1983). The migration of contaminants in groundwater at a landfill: a case study. *Journal of Hydrology* **63**: 51–80.
48. Aravena, R., Auge, M., Bucich, N., and Nagy, M.I. (1999). Evaluation of the origin of groundwater nitrate in the city of La Plata, Argentina, using isotope techniques. In: *Proceeding XXIX IAH Congress, Hydrogeology and Land Use Management*. Bratislava, September 6–10.
49. Flipse, W.J., Katz, B.G., Linder, J.B., and Markel, R. (1984). Sources of nitrate in groundwater in a sewered housing development, Central Long Island, New York. *Ground Water* **22**(4): 418–425.
50. Heaton, T.H.E. (1986). Isotopic study of nitrogen pollution in the hydrosphere and atmosphere: a review. *Chemical Geology (Isotope. Geosci. Sect.)* **59**: 87–102.
51. Böttcher, J., Strebel, O., Voerkelius, S., and Schmidt, H.L. (1990). Using isotope fractionation of nitrate nitrogen and nitrate oxygen for evaluation of nitrate in a sandy aquifer. *J. Hydrol.* **114**: 413–424.
52. Aravena, R., Evans, M.L., and Cherry, J.A. (1996). Stable isotopes of oxygen and nitrogen in source identification of nitrate from septic systems. *Ground Water* **31**: 180–186.
53. Geyh, M. (2000). Groundwater: Saturated and unsaturated zone. In: *Environmental Isotopes in the Hydrological Cycle (Principles and Applications)*. Vol. IV, W.G. Mook (Ed.). UNESCO/IAEA Publ. Series, Paris/Vienna, p. 196.
54. Kendall, C. and Aravena, R. (1999). Nitrate isotopes in groundwater systems. In: *Environmental Tracers in Subsurface Hydrology*. P.G. Cook and A.L. Herczeg (Eds.). pp. 261–297.
55. Foster, S.S.D. and Crease, R.I. (1972). Nitrate pollution of chalk groundwater in East Yorkshire: a hydrogeological appraisal. In: *Groundwater Pollution in Europe*. J.A. Cole (Ed.). Water Information Center, Port Washington, NY, pp. 269–271.
56. Kreitler, C.W. and Jones, D.C. (1975). Natural soil nitrate: the cause of the nitrate contamination of groundwater in runnels county, Texas. *Ground Water* **13**(1): 53–61.
57. Létolle, R.. (1980). Nitrogen-15 in the natural environment. In: *Handbook of Environmental Geochemistry 1A*. P. Fritz and J.Ch. Fontes (Eds.). Elsevier Science Publ., Amsterdam, pp. 407–434.
58. Kreitler, C.W. and Browning, L.A. (1983). Nitrogen isotope analysis of groundwater nitrate in carbonate aquifers: Natural sources versus human pollution. *J. Hydrol.* **61**: 285–301.
59. White, R.J. (1983). Nitrate in British waters. *Aqua*. **2**: 51–57.
60. Heaton, T.H.E. (1984). Sources of the nitrate in phreatic groundwater in the western Kalahari. *J. Hydrol.* **67**: 249–259.
61. Mariotti, A. (1984). Utilisation des variations naturelles d'abondance isotopique en ^{15}N pour tracer l'origine des pollutions des aquiferes par les nitrates. In: *Isotope Hydrology 1983*. IAEA Vienna, 605–633.
62. Flipse, W.J. and Bonner, F.T. (1985). Nitrogen isotope ratios of nitrate in groundwater under fertilized fields, Long Island, New York. *Ground Water* **23**(1): 59–67.
63. Tessendorff, H. (1985). Nitrates in groundwater: A European problem of growing concern. *Aqua* **4**: pp. 192–193.
64. Hübner, H. (1986). Isotope effects of nitrogen in the soil and biosphere. In: *Handbook of Environmental Geochemistry IA*. Elsevier Science Publ., Amsterdam, pp. 361–426.
65. Mariotti, A., Landreau, A., and Simon, B. (1988). ^{15}N isotope biogeochemistry and natural denitrification process in groundwater: Application to the chalk aquifer of northern France. Geochim. *Cosmochim. Acta*. **52**: 1869–1878.
66. Exner, M.E. and Spalding, R.F. (1990). *Occurrence of pesticides and nitrate in Nebraska's groundwater, Report WC1, Water Centre, Inst. Agric. & Natural Resour.,* The University of Nebraska, Lincoln, NE, pp. 3–30.
67. Buzek, F., Kadlecová, R., and Zák, K. (1998). Nitrate pollution of a karstic groundwater system. In: *Isotope Techniques in the Study of Environmental Change*. IAEA, Vienna, pp. 453–464.
68. Révész, K., Böhlke, J.K., and Yoshinari, T. (1998). $\delta^{18}O$ and $\delta^{15}N$ determination in nitrate. In: *Isotope Techniques in the Study of Environmental Change Proc.* IAEA Vienna, pp. 851–854.
69. Pauwels, H., Foucher, J., and Kloppmann, W. (2000). Denitrification and mixing in a schist aquifer: influence on water chemistry and isotopes. *Chemical Geology* **168**: 307–324.
70. Tore, C., Grilli, A., Aguirre, E., and Moya, P. (2001). Application of hydrochemistry and isotopic hydrology to identify recharge areas and characterize nitrate contamination in the Santiago aquifer. In: *Processing III Simposio Sudamericano de Geologia Isotopica*. Pucon, Chile, October, pp. 441–445.
71. Anwar, A.H.M.F. and Ghosh, P.K. (2002). Nitrate leaching to groundwater under an irrigation project. In: *Balancing the Groundwater Budget, Proc. IAH International Groundwater Conference*. Darwin, Australia, May 12–17.
72. Mook, W.G. (2000). *Environmental Isotopes in the Hydrological Cycle (Principles and Applications)*. Vol. I, UNESCO/IAEA Publ. Series, Paris/Vienna, p. 280.
73. Clark, I.D. and Fritz, P. (1997). *Environmental Isotopes in Hydrogeology*. Lewis Publishers, New York, p. 328.
74. Krouse, H.R. (1980). Sulphur isotopes in our environment. In: *Handbook of Environmental Isotope Geochemistry*. Vol. I, P. Fritz and J.C. Fontes (Eds.). *The Terrestrial Environment*. Elsevier, Amsterdam, pp. 435–471.
75. Krouse, H.R. and Grinienko, V.A. (Eds.).(1991). Stable isotopes of natural and anthropogenic sulphur in the environment. *Scope 43*. John Wiley & Sons, New York.
76. Strebel, O., Boettcher, J., and Fritz, P. (1990). Use of isotope fractionation of sulfate-sulfur and sulfate-oxygen to assess bacterial desulfurication in a sandy aquifer. *J. Hydrol.* **121**: 155–172.
77. Schoell, M. (1980). The hydrogen and carbon isotopic composition of methane from natural gases of various origins. *Geochimica Cosmochimica Acta.* **44**: 649–661.
78. Hackley, K.C., Liu, C.L., and Coleman, D.D. (1996). Environmental isotope characteristics of landfill leachates and gases. *Ground Water* **34**: 827–836.
79. Landmeyer, J.E., Vroblesky, D.A., and Chapelle, F.H. (1996). Stable isotope evidence of biodegradation zonation in a

shallow jet-fuel contaminated aquifer. *Environ. Sci. Technol.* **30**: 1120.

80. Conrad, M.E., Templeton, A.S., Daly, P.F., and Alvarez-Cohen, L. (1999). Isotopic evidence for biological controls on migration of petroleum hydrocarbons. *Organic Geochemistry* **30**: 843–859.
81. Vengosh, A., Kolodny, Y., and Spivack, A.J. (1998). Groundwater pollution determined by boron isotope systematics. In: *Application of Isotope Techniques to Investigate Groundwater Pollution IAEA-TECDOC-1046*. IAEA, Vienna, pp. 17–38.
82. Davidson, G.R. and Bassett, R.L. (1993). Application of boron isotopes for identifying contaminants such as fly ash leachate in groundwater. *Environ. Sci. Techn.* **27**: 172–176.
83. Eggenkamp, H.G.M. and Coleman, M.L. (1998). Heterogeneity of formation waters within and between oil fields by halogen isotopes. In: *Water-Rock Interaction*. G.B. Arehart and J.R. Hulston (Eds.). Balkema, Rotterdam, pp. 309–312.
84. Gee, G.W., Dhanpat, R., and Seme, R.J. (1983). Mobility of radionuclides in soils. In: *Chemical Mobility & Reactivity in Soil Systems*. pp. 208–227.
85. Mazor, E. et al. (1981). Sodium exchange in a NaCl waste disposal case (Lobatse, Botswana): implications to mineral water studies. *Environ. Geol.* **3**: 195–199.
86. Song, H., Song, Z., and Xiao, Z. (1990). The stable isotope geochemistry of fresh water and brine in Sichuan Basin. *Proc. National Seminar on Isotope Methods in Environmental Studies, Hydrology and Geochemistry*. Beijing Research. Inst. Uranium Geology, Beijing, pp. 88–96.
87. Wassenaar, L. (1995). Evaluation of the origin and fate of nitrate in the abbotsford aquifer using the isotopes of ^{15}N and ^{18}O in NO_3^-. *Applied Geochemistry* **10**: 391–405.
88. Auge, M.P. and Nagy, M.I. (1996). Origen y evolución de los nitratos en el suelo y el agua subterránea de la Plata, Provincia de Buenos Aires, Argentina. In: *Tercer Congreso Latinoamericano de Hidrología Subterránea, Noviembre*. San Luis Potosí, México, pp. 1–13.
89. Tandia, A.A., Gaye, C.B., and Faye, A. (1998). Origin, process and migration of nitrate compounds in the aquifers of Dakar region, Senegal. In: *Application of Isotope Techniques to Investigate Groundwater Pollution*. IAEA TECDOC-1046. IAEA, Vienna, pp. 67–80.
90. Aggarwal, J., Sheppard, D.S., and Robinson, B.W. (1998). Analytical developments and measurements of boron, nitrate, phosphate and sulphate and case examples of discrimination of nitrogen and sulphur sources in pollution studies. In: *Application of Isotope Techniques to Investigate Groundwater Pollution*. IAEA TECDOC-1046, IAEA, Vienna, pp. 3–15.
91. Shivanna, K., Navada, S.V., Kulkarni, K.M., Sinha, U.K., and Sharma, S. (1998). Application of isotope techniques to investigate groundwater pollution in India. In: *Application of Isotope Techniques to Investigate Groundwater Pollution. IAEA TECDOC-1046*. IAEA, Vienna, pp. 167–184.
92. Halas, S. and Trembaczowski, A. (1998). The influence of brown coal exploitation in Poland on the groundwater pollution as determined by isotopic analysis of sulphate. In: *Application of Isotope Techniques to Investigate Groundwater Pollution*. IAEA TECDOC-1046. IAEA, Vienna, pp. 81–93.
93. Liu, F., Liu, P., Zhu, W., Wu, X., and Zeng, Y. (1998). Application of isotope techniques to groundwater pollution research for Xiangshan uranium ore field, China. In: *Application of Isotope Techniques to Investigate Groundwater Pollution*. IAEA TECDOC-1046. IAEA, Vienna, pp. 107–116.
94. Qureshi, R.M., Mashiatullah, A., and Javed, T. (1998). The application of stable carbon isotope ratios as water quality indicators in coastal areas of Karachi, Pakistan. In: *Application of Isotope Techniques to Investigate Groundwater Pollution*. IAEA TECDOC-1046. IAEA, Vienna, pp. 185–207.
95. De Vito, K.J., Fitzgerald, D., Hill, A.R., and Aravena, R. (2000). Nitrate dynamics in relation to lithology and hydrologic flow path in a river riparian zone. *Journal of Environmental Quality* **29**: 1075–1084.
96. Poulsen, S.R., and Drever, J.I. (1999). Stable isotope (C, Cl, and H) fractionation during vaporization of trichloroethylene. *Environ. Sc. Technol*. **33**: 3689–3694.
97. Harrington, R.R., Poulsen, S.R., Drever, J.L., Colberg, P.J.S. and Kelley, E.F. (1999). Carbon isotope systematics of monoaromatic hydrocarbons: vaporization and adsorption Experiments. *Organic Geochemistry* **30**: 765–775.
98. Meckenstock, R.U., Morasch, B., Warthmann, R., Schink, B., Annweiler, E., Michaelis, W., and Richnow, H.H. (1999). $^{13}C/^{12}C$ isotope fractionation of aromatic hydrocarbons during microbial degradation. *Environm. Microbiol.* **1**: 409–414.
99. Ahad, J.M.E., Sherwood Lollar, B., Edwards, E.A., Slater, G.F., and Sleep, B.E. (2000). Carbon isotope fractionation during anaerobic biodegradation of toluene. Implication for intrinsic bioremediation. *Environ. Sci. Technol.* **34**: 892–896.

ENVIRONMENTAL ISOTOPES IN HYDROGEOLOGY

Segun Adelana
University of Ilorin
Ilorin, Nigeria

In the last three to four decades, environmental isotopes have contributed immensely to studies and investigations in hydrogeology, complementing physical and chemical hydrogeology. Many hydrogeologic studies use stable isotopes of the water molecule to determine groundwater quality, origin, recharge mechanism, and rock–water interaction. A good number of the applications of environmental isotopes in hydrogeology have been in the arid and semiarid areas of the world, where water scarcity is most acute and poses constraints on economic development. A substantial amount of basic data (1–3) and results of applied field investigations have already been published (4–15) on applications of different isotopes for hydrologic purposes.

Environmental isotopes provide indications of groundwater age and serve as natural tracers for groundwater provenance. Stable isotopes of carbon, boron, nitrogen, and sulphur (i.e., $^{13}C/^{12}C$, $^{11}B/^{10}B$, $^{15}N/^{14}N$, and $^{34}S/^{32}S$) can give valuable information about reactions involving these elements and can also serve as pollution tracers. On the other hand, radioactive isotopes of some of these elements decay, providing us with a measure of the circulation time and, invariably, groundwater renewability. The most common of the radioisotopes, carbon-14, is used to estimate groundwater residence time. Isotopes of the uranium series (^{234}U, ^{238}U, ^{226}Ra, and ^{222}Rn) are also useful tracers in isotope hydrogeology but are often not precise enough to establish the age of groundwater due to mineral–water interactions

HISTORICAL BACKGROUND

The history of isotopes dates back to the formation of matter by nucleosynthesis during the birth of our solar system more than 5 billion years ago. This process produced most of the stable and unstable (radioactive) isotopes that naturally exist today (16).

The term “isotope” was first used in 1913 by Soddy Frederick to describe nuclides that occupy the same position in the periodic table but differ in their nuclear properties. Thomson (1913) later showed that the element neon is made up of more than one isotope. In 1919, Francis Aston constructed a mass spectrograph that could be used in discovering virtually all elements. But it was not until 1925 that the existence of oxygen of mass 17 was first observed (32). Later on, oxygen-18 and 17 were discovered in natural material (18). In 1931, Urey discovered the hydrogen isotope of mass 2 (called deuterium) and estimated its natural concentration (19 a, b).

This achievement made it obvious that the isotopic compositions of oxygen and hydrogen were highly variable. As more precise measurements of isotopic ratios continued, it became evident that almost all D_2O and H_2O depleted in deuterium could be prepared by electrolytic decomposition of water. Much later, the concentration of pure ^{18}O was achieved in Switzerland using a thermal diffusion method (20). The work of Urey in 1946 formed the basis of isotopic fractionation (21).

Significant achievement in measuring of natural isotope abundances came with the advent of the Nier/McKinney mass spectrometer (22,23) designed for measuring small differences in isotopic abundance. The discovery of isotopes of water molecules stirred up the several earlier investigations on isotopic hydrology, as reviewed in the work of Rankama (24). Ever since then, investigations using the natural isotopic abundances of all lighter elements have continued to increase and many of them have been published in several texts, journals, and edited specialist volumes (16,25–33).

FUNDAMENTALS OF ISOTOPES IN HYDROGEOLOGY

Isotope derived its name from the Greek word, “*isos*,” meaning equal, and “*topos*,” which means place (referring to the place in the periodic table). Generally, isotopes of an element are atoms or nuclides that have the same number of protons in the nucleus and thus the same atomic number but differ in the number of neutrons and hence in their atomic mass. There are 92 naturally occurring elements comprising more than 1000 isotopes. Most of these occur in terrestrial compounds in trace amounts, but some are sufficiently abundant to be determined quantitatively through routine analysis. Hydrogen, for example, is known to have three isotopes with the following names and symbols:

1. H—common hydrogen (one proton)
2. D—deuterium, heavy and stable Hydrogen (two protons + one neutron)
3. T—tritium, radioactive hydrogen (one proton + two neutrons).

These isotopes can also be described by adding the number of particles in the nucleus of each (i.e., proton + neutron) and placing this at the upper left corner of the symbol for the element. In this way the hydrogen isotopes above may be written as ^1H, ^2H, and ^3H, respectively. For details on elements, nuclei, and valences, see Ref. 30.

There are two classes of isotopes:

(1) Stable isotopes, that do not change with time, even though their concentrations are affected by other physicochemical processes (such as evaporation or condensation).

(2) Unstable isotopes, that decay with time. The products of this decay are said to be radiogenic if they do not themselves decay. An alternative grouping of isotopes exists: (a) environmental isotopes and (b) artificial isotopes.

Environmental isotopes occur naturally, and investigators have no direct control on the variations of their concentrations: artificial isotopes are those whose variations in the environment are created by humans.

The environmental isotopes most commonly used in hydrogeology are the stable isotopes deuterium (^2H) and oxygen-18 (^{18}O) as well as the radioisotopic molecules tritium (^2H), carbon-14 (^{14}C) (see Tables 1 and 2).

The characteristics of isotopes useful to hydrogeologists are isotopic ‘signature’ and ‘decay.’ Once water reaches

Table 1. Stable Environmental Isotopes[a]

Isotope	Ratio	% Natural Abundance	Reference (abundance ratio)	Commonly Measured Phases
^2H	$^2H/^1H$	0.015	VSMOW (1.5575×10^{-4})	H_2O, CH_2O, CH_4, H_2, OH^- minerals
^3H	$^3He/^4He$	0.000138	Atmospheric He (1.3×10^{-6})	He in water or gas, crustal fluids
^6Li	$^6Li/^7Li$	7.5	L-SVEC (8.32×10^{-2})	Saline waters, rocks
^{11}B	$^{11}B/^{10}B$	80.1	NBS 951 (4.04362)	Saline waters, clays, borate, rocks
^{13}C	$^{13}C/^{12}C$	1.11	VPDB (1.1237×10^{-2})	CO_2, $CaCO_3$, DIC, $CH_{2,}$ organics
^{15}N	$^{15}N/^{14}N$	0.366	Air N_2 (3.677×10^{-3})	N_2, NH_4^+, NO_3^-, N-organics
^{18}O	$^{18}O/^{16}O$	0.204	VSMOW (2.0672×10^{-2})	H_2O, CH_2O, CO_2, NO_3^-, sulfates
			VPDB (2.0672×10^{-3})	Carbonates, silicates, OH^- minerals
^{34}S	$^{34}S/^{32}S$	4.21	CDT (4.5005×10^{-2})	Sulfates, sulfides, H_2S, S-
^{37}Cl	$^{37}Cl/^{35}Cl$	24.23	SMOC (0.324)	Saline waters, rocks, evaporites,
^{81}Br	$^{81}Br/^{79}Br$	49.31	SMOB	Developmental for saline waters
^{87}Sr	$^{87}Sr/^{86}Sr$	$^{87}Sr = 7.0$ $^{86}Sr = 9.86$	Absolute ratio measured	Water, carbonates, sulfates, feldspar

[a] Ref. 16.

Table 2. Environmental Radioisotopes[a]

Isotope	Half-life (years)	Decay mode[a]	Principal sources	Commonly Measured Phases
^{3}H	12.43	β^-	Cosmogenic, weapons testing	H_2O, CH_2O
^{14}C	5730	β^-	Cosmogenic, weapons testing, nuclear reactors	DIC, DOC, CO_2, $CaCO_3$, CH_2O
^{36}Cl	301,000	β^-	Cosmogenic and subsurface	Cl^-, surface Cl salts
^{39}Ar	269	β^-	Cosmogenic and subsurface	Ar
^{85}Kr	10.72	β^-	Nuclear fuel processing	Kr
^{81}Kr	210,000	ec	Cosmogenic and subsurface	Kr
^{129}I	1.6×10^{7}	β^-	Cosmogenic, subsurface, nuclear reactors	I^- and I in organics
^{222}Rn	3.8 days	α	Daughter of ^{226}Ra in ^{238}U series	Rn gas
^{226}Ra	1600	α	Daughter of ^{230}Th in ^{238}U series	Ra^{2+}, carbonate, clays
^{230}Th	75,400	α	Daughter of ^{234}U in ^{238}U series	Carbonates, organics
^{234}U	246,000	α	Daughter of ^{234}Pa in ^{238}U series	UO_2^{2+}, carbonate, organics
^{238}U	4.47×10^{9}	α	Primordial	UO_2^{2+}, carbonate, organics

[a] Ref. 16. β^-—beta emission; α—alpha emission, ec—electron capture.

the water table, it retains its isotopic content, as long as isotopic exchange with the reservoir at high temperature does not occur (34).

THE ISOTOPES: ^{2}H, ^{18}O, ^{3}H, ^{13}C AND ^{14}C

Stable Isotopes of the Water Molecule (^{2}H, ^{18}O)

Water is composed of hydrogen and oxygen and therefore occurs with different isotopic combinations in its molecules. As discussed in the previous section, the isotopes of hydrogen are ^{1}H, ^{2}H, ^{3}H, and those of oxygen are ^{16}O, ^{17}O, ^{18}O. Therefore, the possible stable isotopes of water species are $H_2^{16}O$, $HD^{16}O$, $D_2^{16}O$, $H_2^{17}O$, $HD^{17}O$, $D_2^{17}O$, $H_2^{18}O$, $HD^{18}O$, and $D_2^{18}O$.

Significant among these in geochemistry (based on the natural abundance of the isotopes) are $^{1}H_2^{16}O$, $HD^{16}O$, $H_2^{17}O$, and $H_2^{18}O$. The slight variations in their abundance are caused by small differences in the reactivity of the isotopes due to mass differences. Usually, $H_2^{16}O$ is about 10% lighter than $H_2^{18}O$ and therefore more reactive. Under closed conditions and without further reactions, the concentrations of isotopes of the water molecule remain stable over time.

Generally, stable environmental isotopes are measured as ratios of the two most abundant isotopes of a given element. The absolute abundance ratio of isotopes is not usually measured in natural waters because it requires a sophisticated mass spectrometric technique. Due to the low differences, the ratio of the stable isotopes R is given in a delta notation (in per mil units, i.e., parts per thousand) as a deviation from a standard. The mean isotopic composition of seawater, generally known as "SMOW" (Standard mean ocean water), is used for reference (35). However, the reference now commonly adopted for stable oxygen and hydrogen isotopic variation in natural water is V-SMOW (Vienna standard mean ocean water), which is isotopically identical to SMOW (9,36). This difference between samples and the reference standard is expressed in the following relation:

$$\text{Delta}(\delta) = (R_{\text{sample}} - R_{\text{standard}})/R_{\text{standard}} \times 1000 \quad \text{(per mil or ‰)} \tag{1}$$

(R = ratio of the heavy isotope to the light one, for example, $[^{1}H^{2}H^{16}O]/[^{1}H_2{}^{16}O]$).

By definition, the seawater standard has δ^{2}H and δ^{18}O values equal to 0‰. Negative values characterize water isotopically depleted ("lighter"), whereas positive values correspond to water samples isotopically enriched ("heavier") with respect to the standard. The measuring accuracy is 0.15‰ for δ^{18}O and 1‰ for delta δ^{2}H (37). For details of the measuring technique, refer to IAEA (9,29).

Isotopic Fractionation. Environmental isotopes of the same element can be partitioned or separated in a thermodynamic reaction due to differences in the rates of reaction of different molecular species. Fractionation is a fundamental process common to stable isotopes of H, B, C, O, N, S, and Cl and can occur under equilibrium or nonequilibrium (kinetic) conditions. Fractionation can also result from molecular diffusion. The different isotopic water molecules have various vapor pressures and freezing points. The changes in the isotopic ratio during evaporation, condensation, and chemical and biological processes—caused by these differences—are termed "isotopic fractionation."

Vapor derived from seawater is isotopically depleted compared to SMOW. One of the reasons for the depletion is the lower vapor pressure of the heavy water (e.g., $^{2}H_2^{18}O$). The fractionation or separation occurs at the transition between gaseous, liquid, or solid phases, and this is usually expressed by fractionation factor, α, which is defined as

$$\alpha = R_{\text{reactant}}/R_{\text{product}} \tag{2}$$

For example, the exchange between isotopes of the water molecule and the associated fractionation is defined as follows:

$$H_2O_{\text{liquid}} \Leftrightarrow H_2O_{\text{vapor}} \tag{3}$$

$$\alpha_{\text{liquid-vapor}} = (^{2}H/^{1}H)_{\text{liquid}}/(^{2}H/^{1}H)_{\text{vapor}} \tag{4}$$

$$\text{or } \alpha_{\text{liquid-vapor}} = (^{18}O/^{16}O)_{\text{liquid}}/(^{18}O/^{16}O)_{\text{vapor}}$$

Details of the theory and applications of stable isotopic fractionation exist in publications (16,30,34,38–40).

Of major importance in isotope fractionation is "kinetic" separation, which occurs when there is a deficit of moisture in the vapor phase. Kinetic separation results from the differing diffusion constants for the heavy (e.g., ${}^2H_2^{18}O$) and the light (e.g., ${}^1H_2^{16}O$) water molecules: during evaporation the light molecules diffuse faster than the heavy ones through the boundary layer between the water body and the atmosphere (41,42).

Isotopic fractionation is more efficient if the vapor produced in condensation is constantly removed. This leads to the concept of "Rayleigh fractionation," a process in which ${}^{18}O$ or ${}^2H$ is selectively removed from the vapor phase, so that rain becomes progressively lighter in $\delta^{18}O$ and δ^2H, as it falls farther from the ocean. Rayleight fractionation (or distillation, as it is now called) is an exponential function expressed by the relation:

$$R/R_0 = f^{(\alpha-1)} \tag{5}$$

where $R = {}^{18}O/{}^{16}O$ of residual vapor
f = fraction of residual vapor
α = fractionation factor (= 1.009)

For details on Rayleigh processes in hydrogeological applications, see Reference 16. Isotope fractionation can be used to interpret isotopic data from natural settings and is also a useful tool for elucidating and quantifying processes as well as reactions in a hydrogeologic system.

Dependencies of the Isotopic Composition. In environmental isotopic studies, ${}^{18}O$ and ${}^2H$ concentrations in precipitation provide a characteristic input signal that varies regionally and over time. The isotopic signatures encountered in precipitation depend on parameters like temperature, deficit of moisture in the air, and the isotopic ratio of the water vapor source. When these parameters are taken into consideration, the isotopic signatures give information about the origin of vapor, precipitation, and groundwater—and partly about climatic conditions during recharge processes in the past. Based on these interrelations, some generally valid dependencies of the isotopic composition of a groundwater sample can be deduced (41).

The Elevation Effect. During the rise of humid air due to orographic obstacles and successive precipitation, the concentration of heavy isotopes in the precipitation decreases with elevation. The depletion also results from the general decrease of cloud temperatures with elevation. For the elevation effect, the depletion is -1 to $-4‰$ and -0.15 to $-0.5‰$ for δ^2H and $\delta^{18}O$ per 100 m rise. The elevation correction (also known as the altitude or alpine effect) distinguishes groundwater recharged at high altitudes from that of low altitude. Therefore, it turns out to be a useful tool in hydrogeologic studies.

The Continental Effect. During the condensation of atmospheric vapor, the liquid phase (i.e., rain droplets) is isotopically enriched, and the vapor phase is isotopically depleted. However, the amount of vapor in a cloud is limited. As this process continues, the isotopic signature of the vapor and consequently, of the condensed water is continuously changing, which, invariably, leads to a situation in which both the precipitation and groundwater are found depleted with respect to heavy isotopes, as the distance away from the coast increases.

The Effect of Precipitation Rate. This is otherwise known as the amount effect that shows the dependence of the isotopic composition on the amount of rainfall: Heavier rain or greater precipitation amounts result in more negative δ^2H and $\delta^{18}O$ values. As the amount of precipitation increases, depletion in the rain can be observed. During one precipitation, of a significant difference in the isotopic signature is found due to (1) progressive condensation and (2) variations in the intensity of the rain.

The Temperature Effect. The isotopic composition of precipitation depends on the temperature at which oceanic water is evaporated into the air. Seasonal fluctuations of the isotopic composition in local precipitation are influenced by fluctuations in temperature. Rain during winter is isotopically lighter than rain during the summer. In the past, when the climate was significantly different from that of today, then the isotopic signature of groundwater formed in the past differed strongly from that of modern precipitation and modern groundwater.

Moreover, the dependence of isotopic fractionation on temperature and moisture causes an annual fluctuation (seasonal effect) and depletion with latitude (latitude effect). In arid and semiarid zones of low moisture saturation and precipitation amounts, an enrichment of the heavy isotopes in raindrops occurs while they are falling (evaporation effect). Also due to evaporation, surface water in rivers and lakes is isotopically enriched.

Meteoric Water Line. Precipitation in rivers and lakes measured worldwide (35) showed that the delta values of the stable isotopes fit along a straight line on a $\delta^2H-\delta^{18}O$ plot. This line, termed the "global meteoric water line" is characterized by the relation:

$$\delta^2H = 8\delta^{18}O + 10(‰) \tag{6}$$

Craig's global line was later refined from more than a decade of worldwide monitoring of the stable isotopic composition of precipitation [IAEA Global Network of Isotopes in Precipitation—GNIP, reported in (43)] as

$$\delta^2H = 8.13\delta^{18}O + 10.8(‰) \tag{7}$$

From Equation 6, the gradient (s) of the GMWL line is 8, and "d," the intercept on the y axis is 10‰. The value of d, first used by Dansgaard (34) to characterize the deuterium excess in global precipitation, is defined from Equation 6 as

$$d = \delta^2H - 8\delta^{18}O(‰) \tag{8}$$

Evaporation during precipitation changes in the gradient of the straight line to values <8 causes (see Table 3).

Table 3. Examples of Regional Meteoric Lines[a]

Region	Meteoric line (%)
'Global' (meteoric line)	$\delta^2H = 8\delta^{18}O + 10$
Northern Hemisphere (continental)	$\delta^2H = (8.1 \pm 1)\delta^{18}O + (11 \pm 1)$
Mediterranean (or Middle East)	$\delta^2H = 8\delta^{18}O + 22$
Maritime Alps (April 1976)	$\delta^2H = (8.0 \pm 0.1)\delta^{18}O + (12.1 \pm 1.3)$
Maritime Alps (October 1976)	$\delta^2H = (7.9 \pm 0.2)\delta^{18}O + (13.4 \pm 2.6)$
Northeastern Brazil	$\delta^2H = 6.4\delta^{18}O + 5.5$
Northern Chile	$\delta^2H = 7.9\delta^{18}O + 9.5$
Tropical Islands	$\delta^2H = (4.6 \pm 0.4)\delta^{18}O + (0.1 \pm 1.6)$

[a] Ref. 30.

Tritium (^3H)

Tritium, the radioactive isotope of hydrogen, is produced naturally in the upper atmospheric strata by the influence of cosmic radiation on nitrogen atoms:

$$^{14}N + n \Rightarrow {}^{15}N \Rightarrow {}^{12}C + {}^{3}H \tag{9}$$

Tritium, which combines in the atmosphere with oxygen to form water, may precipitate on the earth as rain and thus reach the groundwater. Tritium is usually symbolized as T or simply ^3H. Its concentration in water is expressed in tritium units (TU): 1 TU corresponds to 1 atom ^3H per 10^{18} atoms ^1H. Details of measuring techniques and procedure are fully discussed in Ref. 44.

Tritium decays with a half-life of about 12.35 years to form ^3He. The atmospheric concentration of tritium prior to 1953 was about 3–5 TU. Due to nuclear weapon tests, the concentrations in precipitation up to 1963 reached several thousand TU. Hence, since the early 1960s, the anthropogenic tritium from bombs was used as tracer to study young groundwater. In the meanwhile, however, in Europe, the concentrations have decreased to values <10 TU (45). Measurable tritium in groundwater usually signifies modern recharge. High tritium (>30 TU) indicates recharge in the 1960s, and low values (<1 TU) usually signify paleogroundwater (older groundwater) that has mixed with shallow modern groundwater.

Although, qualitative and quantitative approaches to dating groundwater are also possible for tritium (16), the direct age determination of groundwater accurate to a year is somewhat uncertain partly due to the unknown extent of mixing of each year's recharge with that of the previous year's and partly because of high local and temporal variability of input values in precipitation. However, by measuring ^3H together with its daughter ^3He, true age determination is possible by calculations not based on the complicated tritium input function.

Carbon-14 and Carbon-13

There are three isotopes of carbon in nature: common and stable carbon (^{12}C), rare and stable carbon (^{13}C), and very rare and radioactive carbon (^{14}C). ^{14}C is formed, like tritium, in the upper atmosphere from the impact of neutrons produced by cosmic radiation on nitrogen atoms:

$$^{14}N + n- \Rightarrow {}^{14}C + p \tag{10}$$

where n = neutron and p = proton.

The half-life ($t_{1/2}$) of carbon-14 is 5,730 ± 30 years. The natural ^{14}C-level in the atmosphere corresponds to a ratio $^{14}C/^{12}C$ of 1.18×10^{-12}. The ^{14}C concentration of a sample is given in pmc units (percent modern carbon) as share of the atmosphere value in 1950, which was fixed at 100 pmc. ^{14}C values >100 pmc in the atmosphere and in shallow aquifers document nuclear weapon tests during the 1950s.

The stable isotope ^{13}C is often used to determine the initial content of ^{14}C. The processes of fractionation can reflect themselves in the ratio $^{13}C/^{12}C$. The ^{13}C concentration of a sample is given as deviation in ‰ from a standard. As standard, the ^{13}C value of a marine limestone (Peedee belemnite = PDB) that has the delta value of 0‰ is used. Within the atmosphere, δ values range from −7 to −8‰. Isotopic fractionation is stronger during biological processes. Hence, biological CO_2 has values of about −25 to −15‰ depending on the predominant photosynthesis cycle.

Carbon Dating: Principles and Problems. There are two methods of sampling for ^{14}C: The first is by precipitation of approximately 60 millimoles of total carbon as barium or strontium carbonate at a pH > 9. The second method involves acidification of a water sample, gas stripping the CO_2 with nitrogen, and trapping the evolved CO_2 in a solution of $CO_3{}^{-2}$-free NaOH. ^{14}C is measured by beta counting either in gas or liquid phase or by a more recent method using accelerator mass spectrometry of a graphite target. Radiocarbon activities expressed as percent "modern" carbon (pmc) represent the activity of carbon prior to the dilution by postindustrial 'dead' fossil fuel carbonate. For example, a carbon sample that has 0 pmc is deemed dead or has an age beyond the limit detectable by radiocarbon dating. The age of a given water sample can be calculated from the relation,

$$t = 1/\lambda \ln C/C_0 \tag{11}$$

where t = age (in years), λ = decay constant of ^{14}C, C = measured ^{14}C activity, and C_0 = initial ^{14}C activity.

The age dating of groundwater that contains dissolved, inorganic radioactive carbon (^{14}C) may be used for ages up to 60,000 years, although poor preservation and subsequent contamination of old material now makes the effective dating range shorter. For organic material, the effective range is <50,000 years; for groundwater, the range is limited to 30,000 years or less. The main problems associated with ^{14}C dating are (1) determination of the initial C-14 value, (2) mixtures of waters of different ages, (3) a diffusive admixture of CO_2 from the atmosphere, and (4) dissolution of carbonates. Despite these problems, ^{14}C -age dating is widespread and yields good results in simple situations (16,46).

Correction of the Carbon-14 Age. By various interactions of recent and fossil carbon, an initial ^{14}C value between 50 and 100 pmc results. Other, partly bacterial processes can also influence the isotopic composition of groundwater. During the development of the ^{14}C method, several correction models were set up that reduce the initial activity of samples below 100 pmc. The result is a corrected age. Several correction models exist (46–51). The resulting ages are often compared with conventional ages, which are obtained with initial values of 100 pmc.

OTHER NATURALLY OCCURRING ISOTOPES

Chlorine-36 (^{36}Cl)

Chlorine-36 is a radioactive isotope of chlorine whose application to hydrology has attracted much interest in the last decade. It is naturally produced by cosmic rays interacting with atmospheric argon (^{40}Ar) and finds its way into the hydrologic cycle either as dry fallout or in precipitation. The thermonuclear bomb testing of the 1960s contributed a significant amount of ^{36}Cl, thus elevating its concentrations above the natural atmospheric abundance. ^{36}Cl behaves conservatively in most hydrologic environments, and like bomb tritium, it is useful in delineating recharge rates. However, unlike tritium, its use for dating modern groundwater is unrealistic. But its half-life of 301,000 years makes ^{36}Cl a useful tool in groundwater age determination in the range of 10^5–10^6 years.

Chlorofluorocarbons (CFCs)

Chlorofluorocarbons exist in the atmosphere as CCl_3F and CCl_2F_2 (simply referred to as CFC-11 and CFC-12, respectively). The production of Freons for use as solvents, refrigerants, and propellants has released large quantities of these compounds into the atmosphere. The source of CFCs is mainly anthropogenic, and studies have shown that their concentration in the atmosphere is steadily increasing since the production of Freons started in the mid-1940s. Hydrogeologists have found a use for CFCs, which have accumulated as contaminants in the atmosphere, as tracers like tritium. The analysis of CFC compounds is less complicated than that of tritium. This advantage, coupled with the decreasing concentration of bomb tritium (since the 1990s), is responsible for its increasing applications in hydrogeologic studies. CFC-11 and CFC-12, whose atmospheric residence times are 60 and 120 years, respectively, are not isotopes in themselves but equilibrate with water to form a dating tool for groundwater (<50 years old). They are equally applicable as tracers of groundwater and of sewage contamination in water (16,33).

Uranium Series Isotopes

The radioactive decay of uranium and thorium results in the formation of a series of isotopes, which are in themselves radiogenic. Significant among these are ^{234}U, ^{238}U, ^{226}Ra, and ^{222}Rn. The uranium series displays an enormous array of half-lives (10^5–10^9) and many geochemical distinctive characteristics that make them useful in hydrogeologic applications (16,33):

1. Evaluation of mixing between groundwater bodies can be achieved on a plot of excess ^{234}U versus ^{238}U concentration.
2. The wide range of half-lives is useful in tracing groundwater movement and investigating geochemical processes.
3. Measurement of groundwater–surface water interactions and recharge processes is made possible with ^{222}Rn (half-life $\approx$ 4 days).
4. Uranium activity ratios may provide additional information on the geochemistry of the groundwater system.

Boron Isotopes and Other Pollution Tracers

Natural boron has two stable isotopes, ^{11}B and ^{10}B. Boron is used as tracer in groundwater because of its high solubility in aqueous solution, natural abundance in all waters, and the lack of effects by evaporation, volatilization, and oxidation–reduction reactions. ^{11}B mixing curves enable identification, and in some cases, quantification of contaminants in groundwater (52).

A good number of other isotopes have been employed to trace groundwater pollution. The type of pollutant investigated governs the type of isotope used to trace groundwater pollution. For example, ^{15}N/^{14}N and ^{18}O/^{16}O are used for nitrates in groundwater, ^{34}S/^{32}S for dissolved sulfates, and ^{13}C/^{12}C for dissolved inorganic carbon. ^{11}B (with ^{18}O, ^{87}Sr/^{86}Sr, Br/Cl) is useful in tracing salinity, and ^{13}C/^{12}C, ^{2}H/^{1}H, and ^{37}Cl/^{35}Cl can be used to trace organic carbon pollutants.

SUMMARY OF APPLICATIONS OF ENVIRONMENTAL ISOTOPES

The environmental isotope technique in hydrogeology is based on the principles that stable isotopes and tritium data for groundwater correlate with local precipitation. ^{18}O and ^{2}H concentrations in precipitation provide a characteristic input signal that varies regionally and over time. The isotopic signatures encountered in precipitation depend on parameters like temperature, the deficit of moisture in the air, and the isotopic ratio of the water vapor source. When these parameters are taken into consideration, the isotopic signatures give information about the origin of vapor, precipitation, and groundwater—and partly about the climatic conditions during recharge processes in the past.

Although, qualitative and quantitative approaches to dating groundwater are possible using tritium, the accuracy of this age determination is in question due to the unknown extent of mixing of recharging water with that of the previous year and partly because of high local and temporal variability of input values in precipitation. However, by measuring ^{3}H together with its daughter ^{3}He, true age determination is possible by calculations not based on the complicated tritium input function. Age dating of groundwater with ^{14}C is also widespread in spite of the problems inherent in the method. Stable isotopes of carbon, boron, nitrogen, and sulfur (i.e., ^{13}C/^{12}C^{11}B/^{10}B, ^{15}N/^{14}N, and ^{34}S/^{32}S) can give valuable information about reactions involving these elements and can also serve as pollution tracers. On the other hand, radioactive isotopes of some of these elements decay, providing us with a measure of circulation time and invariably groundwater renewability. In all, environmental isotopes have been applied to a wide spectrum of hydrogeologic problems.

BIBLIOGRAPHY

1. IAEA. (1976). *Interpretation of Environmental Isotope and Hydrochemical Data in Groundwater Hydrology*. TECDOC-Series, IAEA, Vienna, p. 228.
2. IAEA. (1981). *Statistical Treatment of Environmental Isotope Data in Precipitation*. TECDOC-Series, IAEA, Vienna, p. 276.
3. IAEA. (1990). *World Survey of Isotope Concentration in Precipitation (1953–1987)*. Vols. 1–9, IAEA, Vienna.
4. IAEA. (1962). *Applications of Techniques in Hydrology*. TECDOC-Series, IAEA, Vienna, p. 31.
5. IAEA. (1967). *Isotopes in Hydrology*. TECDOC-Series, IAEA, Vienna, p. 740.
6. IAEA. (1978). *Isotope Hydrology—Vol. I*. TECDOC-Series, IAEA, Vienna, p. 440.
7. IAEA. (1979). *Isotope Hydrology—Vol. II*. TECDOC-Series, IAEA, Vienna, p. 540.
8. IAEA. (1980). *Arid Zone Hydrology: Investigations with Isotope Techniques*. TECDOC-Series, IAEA, Vienna, p. 265.
9. Gat, J.R. and Gonfiantini, R. (1981). *Stable Isotope Hydrology, Deuterium and Oxygen-18 in the Water Cycle*. IAEA, Tech. Rept. Series No. 210, p. 337.
10. IAEA. (1983). *Isotope Techniques in Hydrological Assessment of Potential Sites for the Disposal of High-Level Radioactive Wastes*. TECDOC-Series, IAEA, Vienna, p. 164.
11. IAEA. (1984). *Isotope Hydrology*. TECDOC-Series, IAEA, Vienna, p. 874.
12. IAEA. (1986). *Isotope Techniques in the Study of the Hydrology of Fractured and Fissured Rocks*. TECDOC-Series, IAEA, Vienna, p. 306.
13. IAEA. (1994). *Mathematical Models and Their Applications to Isotope Studies in Groundwater Hydrology*. TECDOC-777, IAEA, Vienna, p. 283.
14. IAEA. (1996). *Isotope Field Applications for Groundwater Studies in the Middle East*. TECDOC-890, IAEA, Vienna, p. 250.
15. IAEA. (1998). *Applications of Isotope Techniques to Investigate Groundwater Pollution*. TECDOC-1046, IAEA, Vienna, p. 262.
16. Clark, I.D. and Fritz, P. (1997). *Environmental Isotopes in Hydrogeology*. Lewis, Boca Raton, FL, p. 328.
17. Blackett, P.M.S. (1925). Ejection of proton from nitrogen nuclei photographed by the Wilson method. *Proc. R. Soc. (London), Ser. A* **107**: 349.
18. Giauque, W.F. and Johnston, H.L. (1929). An isotope of oxygen, mass 18. *J. Am. Chem. Soc.* **51**: 1436.
19. Urey, H.C., Brickwedde, F.G., and Murphy, G.M. (1932a). A hydrogen isotope of mass 2. *Phys. Rev.* **39**: 1645;
(1932b). Hydrogen mass 2 and its concentration. *Phys. Rev.* **40**: 145.
20. Clusius, K., Dickel, G., and Becker, E. (1943). Reindarstellung des schweren Sauerstoff-isotopes $^{18}O_2$ und des Stickstoffs $^{14}N^{15}N$. *Naturwissenschaften*. **31**: 210.
21. Urey, H.C. (1947). Thermodynamic properties of isotopic substances. *J. Chem. Soc.* **1947**: 562–581.
22. Nier, A.O. (1947). A mass spectrometry for isotope and gas analysis. *Rev. Sci. Instrum.* **18**: 398.
23. McKinney, C.R., McCrea, J.M., Epstein, S., Allen, H.A., and Urey, H.C. (1950). Improvements in mass spectrometers for the measurements of small differences in isotope abundance ratios. *Rev. Sci. Instrum.* **21**: 724.
24. Rankama, K. (1954). *Isotope Geology*. Pergamon Press, Oxford, p. 490.
25. Dostrowsky, I. and Raviv, A. (1958). Separation of the Heavy Isotopes of Oxygen by Distillation. Isotope Separation. *Proc. Int. Symp*. Amsterdam, p. 336.
26. Fidel, R. (1976). *Bibliography of Hydrogen and Oxygen Stable Isotope Hydrology*. Hebrew University, Jerusalem.
27. Fritz, P. and Fontes, J.Ch. (1980). *Handbook of Environmental Isotope Geochemistry*. Vol. 1, Elsevier, Amsterdam, p. 545.
28. Fritz, P. and Fontes, J.Ch. (1986). *Handbook of Environmental Isotope Geochemistry*. Vol. 2, Elsevier, Amsterdam, p. 557.
29. IAEA. (1995). Reference and Intercomparison Materials for Stable Isotopes of Light Elements. *Proc. Consultants meeting held in Vienna*, December 1–3, 1993. IAEA-TECDOC-825, IAEA, Vienna, p. 165.
30. Mazor, E. (1997). *Chemical and Isotopic Groundwater Hydrology*. Marcel Dekker, New York, p. 413.
31. Kendall, C. and McDonnell, J.J. (1998). *Isotope Tracers in Catchment Hydrology*. Elsevier, New York, p. 839.
32. Cook, P.G. and Herczeg, A.L. (1999). *Environmental Tracers in Subsurface Hydrology*. Kluwer, Dordrecht, the Netherlands.
33. Herczeg, A.L. and Cook, P.G. (2002). *Environmental Isotope in Groundwater*. CGS Workshop presented at the International Groundwater Conference, Darwin, Australia, May 12–17.
34. Dansgaard, W. (1964). Stable isotopes in precipitation. *Tellus*. **16**: 436–469.
35. Craig, H. (1961). Isotopic variations in meteoric waters. *Science*. **133**: 1702–1703.
36. Gonfiantini, R. (1978). Standards for stable isotope measurements in natural compounds. *Nature*. **271**: 534–536.
37. IAEA. (1998). Analytical quality in the IAEA isotope hydrology laboratory: Some recent improvements. *Water Environ. News*, IAEA, No. 3, pp. 5–7.
38. Gat, J.R. (1981). Isotopic fractionation In: *Stable Isotope Hydrology, Deuterium and Oxygen-18 in the Water Cycle*. J.R. Gat and R. Gonfiantini (Eds.). IAEA, Tech. Rept. Series No. 210.
39. Hoefs, J. (1997). *Stable Isotope Geochemistry*, 4th Edn. Springer-Verlag, Heidelberg, p. 251.
40. Kendall, C. and Caldwell, E.A. (1998). Fundamentals of isotope geochemistry. In: *Isotope Tracers in Catchment Hydrology*. C. Kendall and J.J McDonnell (Eds.). Elsevier, New York, p. 839.
41. Moser, H. and Rauert, W. (1980). *Isotope Methods in Hydrogeology* (in Deutsch). Lehrbuch Hydrogeologie, 8, Stuttgart, p. 400.
42. Jacob, H. and Sonntag, C. (1991). An 8-year record of the seasonal variation of ^2H and ^{18}O in atmospheric water vapour and precipitation at Heidelberg, Germany. *Tellus*. **43B**: 291–300.
43. Rozanski, K., Araguas-Araguas, L., and Gonfiantini, R. (1993). Isotopic Patterns in Modern Global Precipitation. In: *Continental Isotope Indicators of Climate*, American Geophysical. Union monograph 78, AGU, Washington, DC, pp. 1–36.
44. IAEA. (1981). *Low-level tritium measurement*. IAEA-TECDOC-246, p. 204.

45. Groning, M. (1994). *Noble Gases and Isotope Tracer in the Groundwater: Paleo-climatic Changes and Dynamics of Regional Groundwater Systems (in Deutsch)*. Ph.D. Thesis. University of Heidelberg, p. 136.
46. Vogel, J.C. and Ehhalt, D. (1963). The Use of Carbon Isotopes in Groundwater Studies. *Radioisotopes in Hydrology*. IAEA, Vienna, pp. 383–395.
47. Vogel, J.C. (1970). Carbon-14 Dating of Groundwater. In: *Isotope Hydrology, IAEA Symposium 129*, pp. 225–239.
48. Tamers, M.A. (1975). The validity of radiocarbon dates on groundwater. *Geophy. Surv.* **2**: 217–239.
49. Fontes, J.Ch. and Garnier, J.M. (1979). Determination of the initial ^{14}C activity of total dissolved carbon: A review of existing models and a new approach. *Water Resour. Res.* **15**: 399–413.
50. Maloszewski, P. and Zuber, A. (1991). Influence matrix diffusion and exchange reactions on radiocarbon ages in fissured carbonate aquifers. *Water Resour. Res.* **27**: 1937–1945.
51. Kalin, R. (1999). ^{14}C in Groundwater. In: P.G. Cook and A.L. Herczeg (Eds.). *Environmental Tracers in Subsurface Hydrology*. Kluwer, Dordrecht, the Netherlands.
52. Vengosh, A. (1998). Boron Isotopes and Groundwater Pollution. *Water Environment News*, IAEA, No. 3, pp. 15–16.

WATER-JETTING DRILLING TECHNOLOGIES FOR WELL INSTALLATION AND *IN SITU* REMEDIATION OF HYDROCARBONS, SOLVENTS, AND METALS

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Water-jet drilling or jetting has been used as a drilling method in both the developing world for water supply projects and in the United States for in situ remediation of hydrocarbons, chlorinated solvents, and metals.

Jetting has been documented for installing water supply wells in coarse materials such as boulders in glacial till deposits where the water pressure nozzle can move coarse gravels. This method uses a high-velocity stream of water to advance the borehole. The jet-percussion system is limited to drilling 7 to 10 cm diameter wells to depths of about 60 meters. Drilling tools consist of a chisel-shaped bit attached to the lower end of a piping string. Nozzles on each side of the bit shoot water out at moderate to high pressures. The water keeps the bit clean and lubricated and carries the pulverized drill cuttings in suspension up the annulus. The cuttings settle out by gravity in suspension pits. The water is then picked up by the suction of the mud pump and recirculated through the drill pipe to the nozzles at the bottom of the borehole (1).

As a source of clean water wells for the developing world, such as Africa, well jetting has been successfully designed for the resources, tools, and skills of even the poorest farmers. In the 1980s, George and Richard Cansdale of SWS Filtration developed jetting as a drilling technique that could be used by the developing world to obtain economical and safe supplies of water. The hand pump, called a Rower Pump, is a suction piston pump made by SWS Filtration. The Rower Pump has the entire main pump components such as a foot valve, piston valve, and pump cylinder. The operator holds the T-handle, pulls and pushes directly on the piston rod as the water pours out the top of the well. The manually operated pump pulls water from the 1 to 5 cm diameter wells at a rate of several liters per minute.

Well jetting is particularly suited to unconsolidated sediments such as sandy riverbeds or riverbanks, where water is generally found close to the surface (1 to 5 meters). The jetting technique, which generally does not penetrate more than 10 meters, can be used to drive a pipe and filter into the sandy riverbed of a dry river or wadi, so that subsurface water, close to the surface, can be pumped from below the surface. The jet of water is pushed from a tank through a pump and into a PVC or steel rod, which makes the sand fluid and allows the pipe to slide into the subsurface. On some floodplains, it is possible to jet through more than 30 meters of silt before reaching an underlying sandy aquifer. Small-diameter jetted wells, called "washbores" by local farmers, are used worldwide to irrigate crops during the dry season. Some small-bore jetted wells are used for domestic water supplies. More details are described in Power (2).

Case studies of the jetted wells have been documented in Zambia, Sudan, Nigeria, and Zaire, as well as other countries. For their efforts, the Cansdales were awarded a 1990 IBM Award for Sustainable Development for their well jetting and the muscle-powered Rower Pump. Other low-technology methods for bringing water to the surface through the jetted well include a check ball valve connected to plastic tubing.

Jetting technology can also be used to install small diameter piezometers or well points in sand. The small diameter (1 to 2.5 cm diameter) well pipe with pressure nozzles on the bottom is pushed downward as the water jets advance the borehole in front of the pipe. The well pipe is left in place at the target depth, and the well point is sealed in place at the surface with neat cement grout.

For environmental cleanup projects, jetting technology has been used successfully as a delivery method for treatment chemicals for remediating hydrocarbons, solvents, and metals. Remediation jetting using high-pressure, low to high volume injection of liquids into the subsurface through a small-diameter wand or lance driven into the subsurface has been widely used for several decades. Jetting technology uses tree root feeder systems to inject liquids into the ground. The remediation injection process (RlP®), an updated and more powerful, versatile and adaptable jetting delivery system, has been used efficiently to implement or augment a variety of environmental remediation processes, including chemical oxidation, bioremediation, pH adjustment, and metal stabilization. Handheld RIP® lances have been designed to use high-pressure liquid pumps to increase flow at the tip of the wand at pressures from 20,000–35,000 kilopascals (kPa). High-pressure injection of remediation treatment chemicals can be done in sensitive areas (Fig. 1).

At these pressures, the lances are driven downward at velocities up to 1/3 meter per second. High-pressure

Figure 1. Handheld lances used to perform remediation by jetting technology in a limited-access site (FAST-TEK).

injection points placed on close spacing, such as 0.6 to 1.5 meter centers, allow complete in situ coverage, vertically and laterally. The radius of influence around injection ports exceeds 3 meters (3). Injection port spacing includes overlapping areas (Fig. 2).

Jetting technology is used to remediate areas of limited access such as underneath slabs, railways, and buildings; around tanks, pipelines, and subsurface utilities; and into hillsides, excavation pits, and stockpiles. The flexibility and accuracy of this injection delivery system provides distinct advantages over both conventional in situ and ex situremediation systems. Hot spots can be effectively treated using this technology. As a result, jetting can provide appreciable savings in cost and time over traditional remediation technologies (4).

Jetting uses chemical oxidizers to treat soils rapidly that are contaminated by toxic and persistent organic wastes. The two most common liquid oxidizers used for jetting in soil and groundwater remediation are hydrogen peroxide and potassium permanganate.

Alkalinity, pH, and organic content must be evaluated prior to any in situ metal stabilization project. Injection ports are grouted with bentonite or neat cement. Documented jetting projects include oxidation using hydrogen peroxide and potassium permanganate for both chemical oxidation and aerobic bioremediation of a variety of hydrocarbon contaminants, including gasoline and diesel. Chlorinated solvents have been successfully remediated using chemical oxidation methods. Anaerobic biodegradation of chlorinated solvents combines jetting technology with the injection of carbon sources (molasses, cheese whey, lactic acid, or other chemicals). Metal stabilization uses chemical reductants such as sulfide-based treatments (calcium polysufide, sodium metabisulfite, or ferrous sulfate) with jetting equipment to stabilize various metals successfully, including arsenic, lead, and hexavalent chromium (5–9).

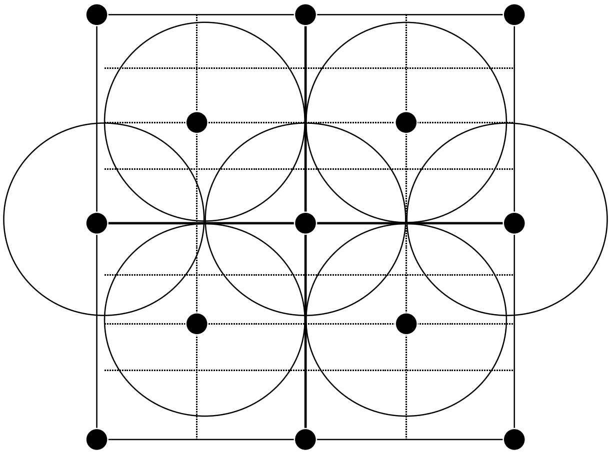

Figure 2. Plan view of overlapping injection zones, spread approximately 0.6 to 1.5 meters apart.

BIBLIOGRAPHY

1. Driscoll, F.G. (1986). *Groundwater and Wells*, 2nd Edn. Johnson Filtration Systems, St. Paul, MN, pp. 307–308.
2. Power, G. (1984). Bangladeshi Rover travels to Africa. *World Water Mag.*, June 2003. Available: http://dspace.dial.pipex.com/swsfilt/rower/rp5.htm.
3. Jacobs, J., Ruslen, D., MacLeod, S., and Taylor, T. (1999). *In-situ* high pressure remediation injection process. *Hydrovisions* Fall 1999, Groundwater Resources Association.
4. Jacobs, J. (2000). In-Situ Remediation of Hot Spots Using Jetting Technology. *52nd Annu. Natl. Ground Water Assoc. Meet.*, Las Vegas, NV, p. 24.
5. Jacobs, J. (2001). Optimization of Existing Soil Vapor Extraction and Pump and Treat Systems Using In-Situ Hydraulic Fracturing. *Groundwater Resourc. Assoc. Calif., 10th Annu. Meet.* and *23rd Biennial Groundwater Conf.*, Sacramento, CA, October 30, p. 65.
6. Jacobs, J. (2001). In-Situ Delivery Methods for Remediation of Hexavalent Chromium in Soil and Groundwater. *Natl. Joint Meet. Assoc. Eng. Geol. Am. Inst.* of *Prof. Geol.*, St. Louis, MO, October 5, p. 62.
7. Jacobs, J. (2001). Insights Into Remediation of Recalcitrant Compounds Using In-Situ Liquid Delivery Systems for Chemical Oxidation, Bioremediation, and Metals Stabilization. *Groundwater Resourc. Assoc. Calif., 2nd Symp.* in *Seri. Groundwater Contaminants: Characterization & Remediation of Emerging and Recalcitrant Contaminants*, June 14–15, 2001.
8. Jacobs, J. (2000). Applications of Jetting Technology for In-Situ Remediation. *Assoc. Eng. Geol. Groundwater Resourc. Assoc. 2000 Annu. Meet.*, San Jose, CA, p. 92.
9. Suthersan, S.S. (2002). *Natural and Enhanced Remediation Systems*. Lewis Publishers/CRC Press, Boca Raton, FL.

KARST HYDROLOGY

David M. Bednar, Jr.
Michael Baker, Jr. Inc.
Shreveport, Louisiana

The most interesting aspect of a karst landscape is its unique hydrologic system. Karst landscapes are primarily underlain by carbonate, primarily limestone and dolomite, where surface water and groundwater are interconnected to form one functional unit. The end result of this hydrologically integrated flow system is the karst spring. Springs provide substantial quantities of water for human consumption and commercial enterprises. Because karst aquifers provide effective natural cleansing, it is imperative to understand the dynamic flow systems within these aquifers to protect these vulnerable water supplies from contamination.

Karst landscapes occupy 10–20% of the earth's land surface (1). The complex hydrology of karst aquifers has been studied for more than a century, and the understanding of flow patterns has advanced significantly in the last 30 years. Investigative methods used have included spring hydrographs and chemographs, groundwater dye tracing, cave mapping, and computer modeling. However, there is still more to learn about the complex nature of karst groundwater flow systems.

This synopsis includes the terms and concepts developed throughout the last 40 years to characterize the hydrology within karst landscapes. The simultaneous efforts by many researchers throughout Europe, North America, China, and Australia in different karst settings has shaped our understanding of the hydrology of karst landscapes. Excellent sources on karst hydrology are Milanovic (2), Bonacci (3), White (4), Ford and Williams (5), Quinlan et al. (6), and White (7).

GROUNDWATER RECHARGE TO KARST SYSTEMS

Several terms have been developed throughout the years that describe where and how meteoric water enters the subsurface to replenish or recharge underlying aquifers. The first two terms refer to where the recharge waters originate. Jakucs (8) and Pitty (9) recognized allogenic and autogenic recharge. Allogenic recharge is surface water that flows from adjacent nonkarst areas onto to karst. Autogenic recharge refers to water that fall directly onto the karst landscape from precipitation. Gunn (10) proposed four recharge types for conduit flow dominated karst aquifers that combined where and how the water entered the subsurface that included diffuse allogenic, concentrated allogenic, diffuse autogenic, and concentrated autogenic recharge.

The second two terms define how the water moves into the subsurface. Aley (11) used the terms diffuse and discrete recharge. Diffuse recharge falls directly onto the karst landscape and enters the subsurface at a large number of sites where inflow is small. Water infiltrates the soil zone or rock outcrop into the groundwater flow system through joints or fissures in the uppermost layer of the bedrock. Diffuse recharge is relatively slow and can be seen in caves as drip water emerging from cracks in the roof where it deposits secondary carbonates such as dripstone and flowstone (12).

Discrete recharge, also called concentrated recharge, enters the groundwater flow system through localized areas, such as sinking streams or sinkholes. Substantially greater quantities of water per unit area enter the groundwater flow system through discrete recharge areas than through diffuse recharge areas (11). Discrete recharge and diffuse recharge represent the maximum ends of a continuum, and most karst landscapes exhibit both types. Atkinson (13) preferred to use "quickflow" to describe water entering the subsurface through sinking streams and sinkholes, and "baseflow" refers to slow percolation of water into the subsurface through nonlocalized areas.

An argument can be made that the use of any of these eight terms is appropriate and depends on the objective of the study. However, it's the authors opinion that all these terms that have been used to define recharge to karst systems simply leads to confusing terminology for those entering the field of karst science. Palmer (14), Smart and Hobbs (15), and White (12) provide more detailed discussions about recharge in karst systems.

THE EPIKARSTIC ZONE

The epikarstic zone is the uppermost part of the carbonate rock that consists of an integrated system of intersecting dissolution-widened fissures, fractures, cavities, and shafts. The epikarstic zone can vary from zero to 98 feet (16) in thickness, and its development is controlled by such factors as bedrock structure, climate, time since the last glaciation, solubility, vegetal cover, and depth of groundwater circulation. The quantity and depth of solution within the epikarst depends on the quantity of rainfall, its distribution in time, the velocity of recharge water, its composition, and the thickness of the soil and rocks (3). The percentage of bedrock removed by dissolution can range from less than 1% to more than 50% (16). Fissures are considerably widened and observation shows that they decrease rapidly with depth (17) due to the decreased amount of CO^2 in solution as the groundwater approaches equilibrium (18).

Dissolution features in the epikarstic zone are organized to move infiltrating water laterally to localized zones where collector structures, such as shafts, conduct the water farther into the subsurface (19). Water storage in the epikarst generally depends on relief. Areas of gentle topography tend to have higher storage, whereas areas of steeper topography have lower storage. Due to the higher storage capacity in many epikarstic zones, the epikarst serves as a perched aquifer.

Aley (20) divided epikarstic zones into three hydrological types: rapidly draining, seasonally saturated, and perennially saturated. These hydrological types were conceptualized on the basis of dye recovery results where dye was introduced directly into locations of discrete recharge that passed through the epikarst and ultimately discharged at springs. Generally, groundwater travel rates vary in epikarstic systems, and two or more hydrological types may exist. The primary criteria used to distinguish these zones were based on their ability to store water.

Rapidly draining epikarsts are saturated with water for short periods of time especially after major storms that provide little water storage or detainment. Infiltration of fine textured sediment is negligible. Lands commonly characterized by rapidly draining epikarsts include areas of high topographic relief, carbonate bedrock with high solubility, and areas of thin soil and residuum.

Seasonally draining epikarsts store water after major storms lasting for weeks or months or store water on a seasonal basis. Sediment infiltration in voids is greater than that in rapidly draining epikarst. Seasonally saturated epikarsts typically develop in humid areas of moderate relief, appreciable soil and residuum thickness, and at elevations greater than those of local perennial streams. Dye recovery results range from 1–10% for seasonally saturated zones.

Perennially draining epikarsts occur in humid areas of low to moderate relief along perennial streams and are mostly saturated with water. Dye recovery rates are commonly from 0.1–1% for permanently saturated epikarstic zones.

GROUNDWATER TRANSPORT IN KARST SYSTEMS

Before conceptual models of aquifer flow are discussed, one must first define the karst aquifer. The textbook definition of an aquifer does not apply to karst landscapes. Freeze and Cherry (21) define an aquifer as a saturated permeable geologic unit that can transmit significant quantities of water under ordinary hydraulic gradients, which means that most or the entire host rock in the aquifer must be saturated, which is typical in clastic rocks. However, this is not the case in karst areas. Drilling can intercept a water filled conduit that has significant quantities of water in one place, whereas another well drilled 20 feet in any direction may come up dry or provide an inadequate source of water. Furthermore, this definition considers the rock adjacent to a cave stream perched above the water table as an aquifer, even though the rock is not saturated (22), which presents a problem because the entire bedrock unit in karst is not completely saturated, as stated by the textbook definition. Therefore, a more precise definition to satisfy the unique hydrology of karst landscapes was required.

A karst aquifer is an aquifer that contains soluble rocks whose permeability structure is dominated by interconnected conduits dissolved from the host rock which are organized to facilitate the circulation of fluid in a downgradient direction wherein the permeability structure evolved as a consequence of dissolution by the fluid (19). Dissolution of the bedrock increases the permeability of karst aquifers through time. The permeability and preferential pathways created within the karst aquifer are dictated by the hydrodynamic characteristics of the flow system, not the inherited geologic fabric (23). The permeability and hydraulic gradient are the dominant forces that determine the location of conduits where driving force that governs conduit localization is the hydraulic gradient (19,24).

Several terms and analytical methods have been used to conceptualize water movement through karst aquifers. Burdon and Papakis (25) applied the terms diffuse circulation, flow through pores and fractures, and concentrated circulation, flow through larger openings to characterize aquifer flow. Ashton (26) used flood pulse analysis to study the behavior within a karst drainage system. Pitty (9) recognized that karst water exists in discrete systems of fissures. Atkinson (27) separated aquifer flow into three components, diffuse, fissure, and conduit flow. Quinlan and Ewers (28) proposed granular, diffuse, fracture, and conduit flow.

White (29) was the first to develop an oversimplified series of conceptual models for carbonate aquifer flow based on hydrogeology. He proposed two end member aquifer types: conduit flow and diffuse flow. Conduit flow systems form in areas of mature karst and have well-integrated solution channels ranging in size from small pipes to large caves (30). Response to storms is rapid, causes turbulent flow; water can reach velocities of 10^{-3} to 10^{-1} meters per second (22) and has an almost insignificant low-flow hydraulic gradient (31). However, conduit flow represents only a small portion of the aquifer porosity. It has been demonstrated from case studies that enhancement of porosity by dissolution in unconfined carbonate aquifers is relatively minor whereas enhancement of permeability is considerable because dissolution has created an efficient dendritic network of interconnecting channels that can convey 94% or more of the flow in the aquifer (32). Aquifer storage is relatively low in conduit flow systems.

Diffuse flow systems commonly occur in areas of less mature karst. Diffuse flow can occur as fracture flow in dense limestones or intergranular flow in porous limestones. These systems exhibit laminar to slightly turbulent flow through a system of small discrete pathways (widened joints or bedding planes) that are being dissolutionally enlarged extremely slowly (33). Velocities can reach 10^{-8} to 10^{-2} meters per second (22), and response to storms is less than that of conduit flow aquifers. Aquifer storage is higher in diffuse flow systems.

White (34) revised his conceptual model from 1969 to consider the effect of relief, geologic structure, and the aerial extent of the aquifer on its development. The most important conclusion from the study was that recharge was a key factor to differentiate between conduit and diffuse flow aquifers. However, White's conceptual model did not quantify the amount of diffuse and conduit flow in the aquifer. In recognition of this, Atkinson (13) studied the Cheddar Spring catchment within the carboniferous limestone of the Mendip Hills in Great Britain to determine the proportions of conduit flow and diffuse flow within an aquifer. Results indicated that the majority of true groundwater occurs within narrow fissures (diffuse flow), whereas most of the water is transported through the conduits. Recharge to the conduits is derived from quick flow as well as water draining from the diffuse portion of the aquifer. Conduit flow probably accounts for 60–80% of the water transmitted in the aquifer (13).

Quinlan et al. (6) developed a conceptual model that recognized recharge, flow, and storage as the primary components of the classification. They proposed four dominant aquifer types: hypersensitive aquifers, very sensitive aquifers, moderately sensitive aquifers, and slightly sensitive nonkarst aquifers. Each aquifer type describes the aquifer's sensitivity to contamination.

Worthington (35) reported that the threshold passage diameter for conduit flow in a karst aquifer is about 10 mm and that maximum diffuse flow velocities are $<10^{-2}\ \mathrm{m\,s^{-1}}$ and $<10^{-6}\ \mathrm{m^3\,s^{-1}}$ of diffuse flow in a fissure. Because the mean discharges of most karst springs are greater than 10^{-4} $m^3 s^{-1}$, this indicates that these springs are discharging from conduits, that the term diffuse karstic aquifer is a contradiction in terms, and that all karst springs are conduit springs. White (36) agreed with Worthington that the term "diffuse" as applied to spring flow may not be appropriate.

There is agreement among investigators today that karst aquifers consist of a combination of highly anisotropic flow routes and should be classified as triple porosity aquifers (7,37–40). Porosity exists within the rock matrix, in fractures, and in conduits. Matrix porosity refers to water that moves through the intergranular pores in the bedrock. Fracture porosity refers to water movement through fractures and bedding plane partings with the bedrock. Dissolution may or may not have taken place. Conduit or channel porosity refers to the movement of water through dissolutionally widened pathways that commonly have greater flow velocities than those of fracture porosity. Water moving through cave streams is an example of channel porosity.

Worthington et al. (32,38) provided credible support for a triple porosity model in unconfined carbonate aquifers. Case studies were examined in four contrasting carbonate aquifers in the Lower Paleozoic dolostone of Ontario, Canada; the Upper Paleozoic limestone of the Mammoth Cave area in Kentucky in the United States; the English chalk of Great Britain; and the Cenozoic limestone of the Yucatan Peninsula in Mexico. Results demonstrated that at least 96% of the storage in all four aquifers was in the matrix portion of the rock (32) and 94% or more of the flow was through the interconnected network of channels (conduits) within the aquifers.

GROUNDWATER STORAGE IN KARST SYSTEMS

The volume of water within a karst aquifer is commonly defined as storage. Because the karst aquifer is considered to have triple porosity, water in storage can be found in the rock matrix, within fractures that have undergone limited dissolution, and in larger channels where significant dissolution has occurred. Storage depends primarily on such factors as hydraulic gradient, bedrock composition, and the degree of karstification within the flow system. Just like recharge and flow, aquifer storage falls within a continuum. The following studies represent the various conceptual models proposed to characterize karst aquifer storage.

Aley (11) developed a predictive hydrologic model to evaluate the effects of land use and management on the quantity and quality of water from Ozark springs. The model was the result of a study conducted in the Hurricane Creek topographic basin in southern Missouri. Aley stressed that groundwater movement in karst areas cannot be compared to water movement in nonkarst areas. Typical groundwater movement in nonkarst areas is less that 10 feet per year and under such conditions it seems proper to view groundwater as water in storage (11). However, water movement in karst, as substantiated by numerous dye traces, can be several orders of magnitude greater than that in nonkarst areas; flow rates are up to 1 mile per day through conduits and cave streams. As a result, Aley divided groundwater movement into two components within a continuum. Water moving at rates greater than 1 foot per hour was considered "water in transit," while water moving less than 1 foot per hour was "water in storage."

Gunn (41) summarized the work of previous researchers that characterized karst aquifer storage. All studies relied on the importance of storage in the unsaturated zone and recognized four major areas of storage or stores: the soil and superficial deposits, the epikarst, water filled conduits in the phreatic zone along with associated solutionally enlarged fissures and bedding planes, and the saturated rock mass.

Smart and Hobbs (15) proposed that storage, just like recharge and flow type, fell within a continuum that has two end members: unsaturated storage and permanently saturated storage, where a seasonally saturated storage component is the central zone. The model was developed to conceptualize aquifer flow within the unsaturated and saturated zones. Two dominant types of storage were recognized in the saturated zone: dynamic storage and perennial storage. Dynamic storage occurs above base level in unconfined aquifers and flows under gravity to provide discharge at springs; storage is relatively low due to steeper hydraulic gradients. Perennial storage occurs below base level in unconfined aquifers and has a greater volume of water than dynamic storage. Due to the lower hydraulic gradients in this portion of the aquifer, storage is generally higher.

Quinlan et al. (6) modified the Smart and Hobbs model and proposed three distinct zones of storage within the karst aquifer system. The zone of limited storage, the low-end member, is confined to the soil and epikarstic zone, predominately the unsaturated zone. The second zone is identical to the zone of limited storage but a seasonally saturated zone is added. The high-end member comprises aquifers that have an unsaturated zone, a seasonally inundated saturated zone, and either perennial storage in a confined aquifer or a water table aquifer that extends well below spring level (6).

GROUNDWATER DISCHARGE FROM KARST SYSTEMS

Springs represent the pulse of a karst aquifer. For many years, investigators wrestled to find the mechanism that causes variations in spring chemistry. Analysis of spring flow and water chemistry have been the primary techniques used for karst aquifer characterization. Jakucs (8) suggested that recharge type was an important factor that influenced spring chemistry variation; springs that had high chemical variation were related to autogenic percolation recharge, whereas springs that had low chemical variation were associated with concentrated allogenic recharge. Pitty (9) stressed that the percentage of different recharge waters (allogenic recharge via sinking streams and autogenic recharge via percolation) influenced spring chemistry variation. Newson (42) recognized a linear correlation between the percentage of allogenic recharge in a spring's catchment to the coefficient of variation (COV) of hardness at springs; as the percentage of allogenic recharge increased, so did the COV of hardness at the springs. Additionally, Newson found that spring total hardness could be expressed in terms of a simple two-component mixing model (35). From the study of total hardness of springs, cave streams, and percolation water in the Mendip Hills of England, he determined the percentage of allogenic and autogenic recharge received at local springs.

Shuster and White (33) conducted a study within the folded Ordovician dolomites and limestones of central Pennsylvania using 14 springs. They suggested that flow type within the aquifer controlled the chemical character of the springs and used the COV of hardness as the principal statistical parameter. The first group of springs that discharged from fractured dolomites had a constant hardness and a COV of less than 5%, displayed a nearly constant chemical character regardless of season or storm, was recharged by diffuse infiltration, and they called them diffuse flow springs. The second group of springs discharged from karstic limestones with internal recharge from sinkholes. Hardness was very variable throughout the year and had a COV greater than 10%; they called these conduit flow springs. Jacobson and Langmuir (43) produced similar results for chemical variation at carbonate springs.

The conduit flow springs were generally undersaturated with respect to calcite. Springs that discharge from conduit flow systems are flashy based on the high ratios between maximum discharge and base flow discharge. Flow is turbulent, whereas water hardness is low at a mean of 100 ppm, but highly variable. The COV of hardness ranges from 10 to 25% or more (28,44). Groundwater flow velocities are commonly rapid, and water can travel up to a mile per day. Conduit flow springs were considered outlets from conduit flow systems and are described by pipe and channel equations (45).

The diffuse flow springs were generally saturated with respect to calcite. Turbidity is low, but water hardness may exceed 300 ppm; the COV of hardness is typically less than 5% (28). Discharge is nonflashy, groundwater flow is much slower, and water can take months to travel short distances. Diffuse flow springs were considered outlets from diffuse flow systems.

The use of hardness variation to distinguish between flow types appeared to work well for the study conducted by White for small drainage basins in temperate climates. However, larger drainage basins show less variation when springs are fed by conduits (4). It is likely that this is due to longer residence times and the contribution of other waters whose chemistry is different from that of other parts of the basin. Subsequent studies found that the binary classification developed by Shuster and White (33) was not applicable to all karst settings.

Aley (11) applied different terms to Missouri springs. As previously discussed, he divided groundwater in storage into transit water and storage water, which resulted in two spring classifications: high transit springs and high storage springs. High transit springs are similar to conduit flow springs; water is derived from areas of discrete recharge. High storage springs are similar to diffuse flow springs; water is derived from areas of diffuse recharge.

Smart and Hobbs (15) recognized that flow type alone is not responsible for the chemical variation at springs and incorporated recharge and storage components into a simple conceptual model. From this model, it is possible to map the different types of carbonate aquifers. Scanlon and Thrailkill (46) conducted a study in the relatively flat limestones and shales of the Inner Bluegrass karst region of central Kentucky. The physical and chemical characteristics of springs from the Inner Bluegrass were compared with those from the study conducted by Schuster and White (33). Results from their study showed that seasonal water chemistry variation in major low level springs and local high level springs did not correspond with the physical characteristics of the springs over time. Chemical similarities were attributed to the mode of recharge. The difference in bedrock types and structure between the two studies was considered a control factor for the relationships between the physical and chemical attributes of the springs.

Worthington et al. (47) performed a statistical analysis using data from 39 springs in six countries using temperate climates taken from the literature. Results from their study demonstrated no evidence that hardness variation was an indicator of flow conditions within an aquifer, as suggested by Schuster and White, and that greater than 75% of hardness variation was explained by recharge type. Currently, it is recognized that recharge type and storage both play an integral role in the chemical variations at karst springs.

SPRING HYDROGRAPHS AND CHEMOGRAPHS

Spring hydrograph and chemograph analysis have been widely used in concert to gain insight into the hydraulics of the karst groundwater system. A hydrograph is a time series plot of spring discharge versus time. Discharge is plotted on the vertical axis and time on the horizontal axis. The shape of the hydrograph recorded at a spring is a unique reflection of the response of the aquifer to recharge (3,5). A spring hydrograph is divided into three main responses to a discharge event: lag time, the rising limb, and the falling limb. Each of these three components is shown in Fig. 1. The lag time shows the spring discharge before the recharge reaches the spring. The rising limb represents the first arrival and increase of discharge at the spring. The peak of the graph indicates the maximum amount of spring discharge from the storm. The water that flows at the spring at this time does not reflect recharge water; it is older water in the most downstream part of the aquifer that is displaced from storage (49), which represents the pulse-through event, the arrival of energy at the spring from the storm. The falling limb, more commonly called the recession curve, shows the spring discharge as it returns to prestorm conditions.

The recession limb typically consists of a fast response segment and a slow response segment, as shown when the recession limb is plotted on semilogarithmic paper. The fast response segment can be assigned to the conduit portion of the aquifer that drains rapidly when the storm input ceases; the slow response segment can be assigned

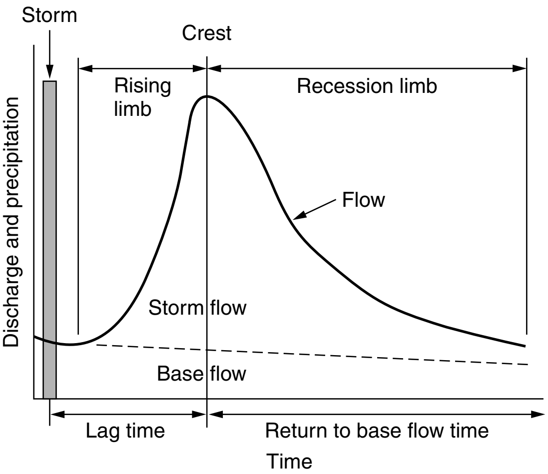

Figure 1. Example of a hypothetical hydrograph (taken from Ref. 48).

to the draining of the fracture system into the conduit after the conduit has emptied (7).

Chemographs show measured concentrations of dissolved chemical parameters in spring water over time. Although many parameters can be used to characterize the chemistry at karst springs, common parameters include turbidity, dissolved oxygen, hardness, stable isotopes, specific conductance, and temperature (5). Figure 2 shows a hydrograph and chemograph from Komlos spring in Hungary. The graph generally shows the increase in parameter concentration after the arrival of the storm pulse wave. Thus, spring hydrographs show a pattern composed of sequential and sometimes superimposed pulses of water of different quality and quantity from different stores and tributary inputs (5).

GROUNDWATER TRACING THROUGH KARST SYSTEMS

Water tracing in karst landscapes has proven to be an effective tool for aquifer characterization. A variety of tracing agents have been used throughout the years to trace and estimate groundwater flow velocities from discrete recharge areas to springs. Tracing agents used include bacteria, surfactants (wetting agents), salts, ions in solution, radioactive isotopes, environmental isotopes, club moss spores, nonfluorescent dyes, and fluorescent dyes (50). Although all agents have been widely used to trace water through karst systems, the following discussion focuses on the use of fluorescent dyes.

Tracing groundwater flow routes with fluorescent dyes is highly successful because the dyes are easily detected in concentrations that are one to three orders of magnitude less than those at which nonfluorescent dyes can be measured spectrometrically (6). Furthermore, fluorescent dyes perform well in different hydrogeologic settings and can be used effectively to trace water from subsurface to surface water bodies.

Dye tracing has been employed extensively in karst areas for a wide variety of purposes. Some of the uses of dye tracing include delineation of spring recharge areas and subsurface basins, determination of site-specific hydrology, estimation of groundwater flow velocities, detection of leakage from residential sewage disposal systems and dam sites, and identification of sources of potential pollution from hazardous waste sites. Dye tracing has also been used successfully during highway corridor studies to identify areas sensitive to accidental spills (51).

Selection of dye introduction points, the manner in which the dyes are introduced, the sampling strategy employed, and the analytical approach used must be tailored to the hydrogeologic setting and the issues of

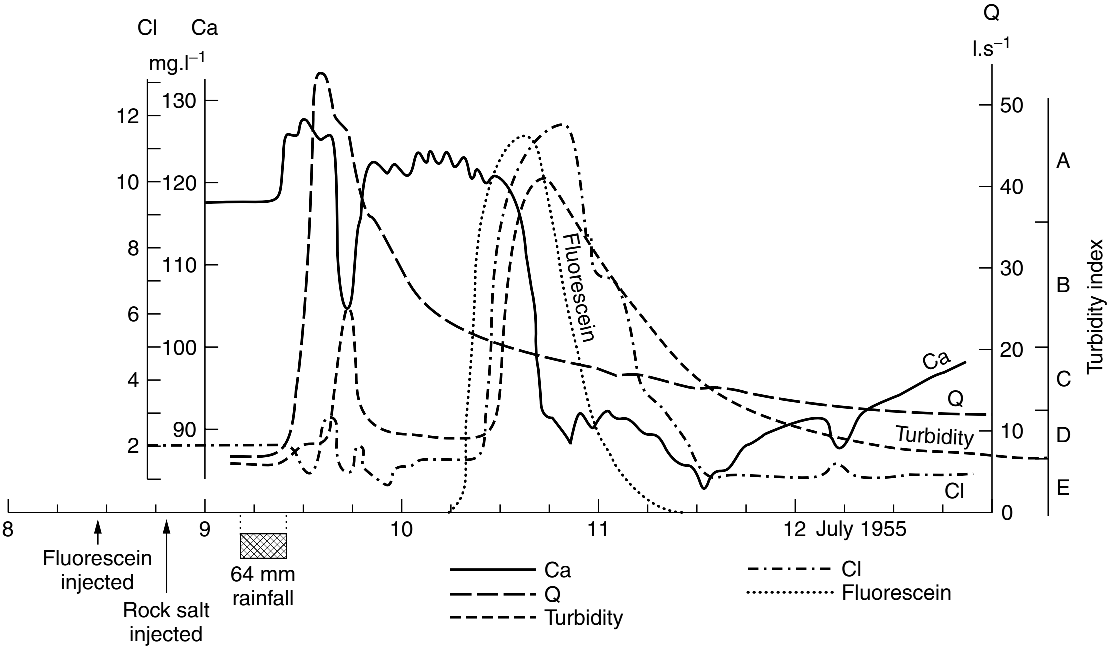

Figure 2. Hydrograph and chemograph of Komlos spring, Hungary (taken from Ref. 8).

concern (52). The amount of dye used for tracing studies is usually based on professional experience because there is no credible standard equation for estimating the dye quantities needed for groundwater tracing. One well-designed dye trace, properly done and correctly interpreted, is worth 1000 expert opinions or 100 computer simulations of groundwater flow (53). Some excellent sources for discussions of groundwater tracing in karst systems include Aley (20,52,54), Smart and Laidlaw (55), Mull et al. (56), Alexander and Quinlan (57), and Kass (50).

VULNERABILITY OF KARST LANDSCAPES

An old adage is that whatever goes up comes down. In karst areas, whatever goes down comes up through a cave, a spring or a well (58). In areas of localized recharge in karst areas, surface water is directly connected to the groundwater flow system. Karst aquifers are highly vulnerable to contamination because they commonly transport water at rates several orders of magnitude greater than those encountered in nonkarst groundwater systems, provide minimum adsorption or other natural cleansing processes, commonly support flow in the turbulent regime, and are too large to provide effective filtration for most pathogens (16). Another concern regarding contamination in karst landscapes is that contamination can enter the subsurface through one location and discharge at multiple spring locations miles apart from each other. Improved adaptation to karst conditions requires that land use and land management decisions simultaneously consider both surface and subsurface resources and conditions (59). Unfortunately, development on karst will continue, but with the implementation of proper land use strategies impacts to this vulnerable landscape can be avoided or at best minimized.

Spores can be dyed various colors and injected into different discrete recharge areas at the same time. Once collected, samples are detected with a microscope. Analysis and preparation techniques require skilled personnel and considerable care (59). An advantage to using spores rather than dyes is that they can be injected in up to six input sites and traced simultaneously, whereas only three sites can be used using during fluorescent dye traces (5).

KARST MODELING

Karst modeling involves developing of methods for reproducing karst processes in the laboratory, so that the natural phenomena can be understood better (60). Karst modeling can include interpretation of several aspects of the karst environment; some of them include groundwater geochemistry and flow patterns, distribution of porosity and permeability in the karst aquifer, fate and transport of contaminants, and groundwater monitoring and remediation strategies. Conceptual, analytical, digital, statistical, and scale models have been used to understand the nature of karst processes better.

Conceptual models are built on ideas based on field observation along with an understanding of relevant physical and chemical characteristics of the system. For a conceptual model to be sound, it must stand alone without quantitative support (60). Analytical models use mathematical equations based on physical laws to model natural phenomena. Equations of fluid flow, chemical kinetics, and conservation of mass are commonly used. When analytical models begin to approach the complexity of real systems, the only way to solve them quantitatively is to break the model into discrete space and time increments so that each increment can then be solved simultaneously (60).

Digital models come in two formats: commercial software for interpreting groundwater flow patterns and chemical equilibria and specialized numerical models designed to investigate complex physical relationships, for example, in the evolution of karst conduits (1). Statistical models are based on measured field data. A hypothesis is developed, and a statistical analysis is conducted to use the model to predict or extrapolate the occurrence of karst features or processes. Scale modeling refers to the actual construction of a model to evaluate physical processes. Sources that discuss karst modeling include White (29), Curl (61), Mangin (62,63), Gunn (64), Clemens et al. (65), and Palmer (60).

BIBLIOGRAPHY

1. Palmer, A.N. (1991). Origin and morphology of limestone caves. *Geological Society of America. Bulletin* **103**(1): 1–21.
2. Milanovic, P.T. (1981). *Karst Hydrogeology*. Water Resources Publications, p. 434.
3. Bonacci, O. (1987). *Karst Hydrology, With Specific Reference to the Dianric Karst*. Springer-Verlag, New York, p. 184.
4. White, W. (1988). *Geomorphology and Hydrology of Karst Terrains*. Oxford University Press, Oxford, UK, p. 464.
5. Ford, D. and Williams., P. (1989). *Karst Geomorphology and Hydrology*. Chapman and Hall, London, p. 701
6. Quinlan, J.F. et al. (1992). Recommended administrative/regulatory definition of karst aquifer, principles for classification of carbonate aquifers, practical evaluation of vulnerability of karst aquifers, and determination of optimal sampling frequency at springs. In: *Conference on Hydrogeology, Ecology, Monitoring, and Management of Groundwater in Karst Terranes Proceedings*. J. Quinlan and A. Stanley (Eds.). National Groundwater Association, pp. 573–635.
7. White, W. (1999). Groundwater flow in karstic aquifers. In: *The Handbook of Groundwater Engineering*. J. Delleur (Ed.). CRC Press, Boca Raton, FL.
8. Jakucs, L. (1959). Neue methoden der hohlenforchung in ungarnund ihre ergebnisse. *Die Hohle* **10**(4): 88–98.
9. Pitty, A.F. (1966). *An approach to the study of karst water, illustrated by results from Poole's Cavern, Buxton*. University of Hull, Occasional Paper in Geography No. 5.
10. Gunn, J.A. (1983). Point recharge of limestone aquifers—a model from New Zealand karst. *Journal of Hydrology* **61**: 19–29.
11. Aley, T. (1978). Ozark hydrology: a predictive model. Missouri Speleology. *Journal of the Missouri Speleological Survey* **18**: 185.
12. White, W. (1990). Surface and near-surface karst landforms. In: *Groundwater Geomorphology: The Role of Subsurface Water in Earth-Surface Processes and Landforms*.

C.G. Higgins and D.R. Coates. (Eds.). Geological Society of America Special Paper 252, pp. 157–175.

13. Atkinson, T.C. (1977). Diffuse flow and conduit flow in limestone terrain in mendip hills, somerset (Great Britain). *Journal of Hydrology* **35**: 93–110.
14. Palmer, A.N. (1984). Geomorphic interpretation of karst features. In: *Groundwater as a Geomorphic Agent*. R.G. Lafleur (Ed.). Allen and Unwin, p. 37.
15. Smart, P.L. and Hobbs, S.L. (1986). Characteristics of carbonate aquifers: a conceptual model. In: *Environmental Problems in Karst and Their Solutions Conference Proceedings*. National Water Well Association, pp. 1–14.
16. Aley, T. (2000). Karst systems. In: *The Standard Handbook of Environmental Science, Health, and Technology*. H. Jay and J.K. Lehr (Eds.). McGraw-Hill, New York, pp. 19.1–19.10.
17. Williams, P.W. (1985). *Subcutaneous Hydrology and the Development of Doline and Cockpit Karst*. Zeitschrift fur Geomorphologie, pp. 463–482.
18. Williams, P.W. (1983). The role of the subcutaneous zone in karst hydrology. *Journal of Hydrology* **61**: 45–67.
19. Huntoon, P. (1994). Is it appropriate to apply porous media groundwater circulation models to karstic aquifers? In: *Assessment of Models for Groundwater resources Analysis and Management Conference Proceedings*. Lewis, Boca Raton, FL, pp. 1–16.
20. Aley, T. (1997). Groundwater tracing in the epikarst. In: *The Engineering Geology and Hydrogeology of Karst Terranes*. Beck and Stephenson (Eds.). Balkema, Rotterdam, pp. 207–211.
21. Freeze, R.A. and J.A. Cherry. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ, p. 604.
22. Quinlan et al. (1992). *Rationale for the Design of Cost-Effective Groundwater Monitoring Systems in Limestone and Dolomite Terranes: Cost Effective as Conceived is Not Cost Effective as Built if the System Design and Sampling Frequency Inadequately Consider Site Hydrogeology*. U.S. Environmental Protection Agency, pp. 552–570.
23. Huntoon, P. (1995). Gradient controlled caves: trapper-medicine lodge area, bighorn basin, Wyoming. *Groundwater* **23**: 443–448.
24. Huntoon, P. (1999). Karstic permeability: organized flow pathways created by circulation. In: *Karst Modeling*. Karst Water Institute Special Publication 5, pp. 79–81.
25. Burdon, D.J. and Papakis, N. (1963). *Handbook of Karst Hydrogeology*. Institute for Geology and Subsurface Research/United Nations Food and Agricultural Organization, Athens, p. 276.
26. Ashton, K. (1966). The analysis of flow data from karst drainage systems. *Transactions of the Cave Research Group of Great Britain* **7**(2): 163–203.
27. Atkinson, T.C. (1985). Present and future directions in karst hydrogeology. *Annales de la Societe Geologique de Belgigue* **108**: 293–296.
28. Quinlan, J.F. and Ewers, R.O. (1985). Groundwater flow in limestone terranes: strategy rationale and procedure for reliable, efficient monitoring of groundwater quality in karst areas. *National Symposium and Exposition on Aquifer Restoration and Groundwater Monitoring, Proceedings*. National Water Well Association, pp. 197–234.
29. White, W. (1969). *Conceptual Models for Carbonate Aquifers. Groundwater* **7**, (3).
30. Crowther, J. (1983). Hydrology of autogenic percolation systems in some tropical karst outcrops, West Malaysia. *Journal of Hydrology* **60**: 227–242.
31. Field, M. (1992). Karst hydrology and chemical contamination. *Journal of Environmental Systems* **22**(1): 1–26.
32. Worthington, S. et al. (2001). Characteristics of porosity and permeability enhancement in unconfined carbonate aquifers due to the development of dissolutional channel systems. In: *Present State and Future Trends of Karst Studies Proceedings*. G. Gunay, K. Johnson, D. Ford, and A.I. Johnson (Eds.). Technical Documents in Hydrology, Vol. 1, No. 49, UNESCO, pp. 13–29.
33. Shuster, E.T. and White, W. (1971). Seasonal fluctuations in the chemistry of limestone springs: a possible means for characterizing carbonate springs. *Journal of Hydrology* **14**: 93–128.
34. White, W. (1977). Conceptual models for carbonate aquifers: revisited. In: *Hydrologic Problems in Karst Regions*. R.R. Dilamarter and S.C. Csallany (Eds.). Western Kentucky University, pp. 176–187.
35. Worthington, S. (1991). *Karst Hydrology of the Canadian Rocky Mountains*. Unpublished Ph.D. Dissertation. Department of Geography, McMaster University, Hamilton, Ontario, Canada.
36. White, W. (1999). Karst hydrology: Recent developments and open questions. In: *Hydrogeology and Engineering Geology of Sinkholes and Karst Proceedings*. Rotterdam, Balkema, pp. 3–21.
37. Worthington, S. (1999). A comprehensive strategy for understanding flow in carbonate aquifers. In: *Karst Modeling*. A.N. Palmer, M.V. Palmer, and I.D. Sasowsky (Eds.). Karst Waters Institute, Special Publication No. 5, pp. 30–37.
38. Worthington et al. (2000). Porosity and permeability enhancement in unconfined carbonate aquifers as a result of solution. In: *Speleogenesis: Evolution of Karst Aquifers*. Klimchock et al. (Eds.). National Speleological Society, Inc., pp. 463–471.
39. White, W. and White, E. (2000–2001). Conduit fragmentation, cave patterns, and the localization of karst groundwater basins: the appalachians as a test case. *Theoretical and Applied Karstology* **13–14**: 9–23.
40. Martin, J.B. and Screaton, E.J. (2001). Exchange of matrix and conduit water with examples from the floridian aquifer. In: *U.S. Geological Survey Karst Interest Group Proceedings*. Water Resources Investigations Report 01–4011, pp. 38–44.
41. Gunn, J.A. (1986). A conceptual model for conduit flow dominated karst aquifers. In: *Karst Water Resources Proceedings*. International Association of Hydrological Sciences, Publication No. 161, pp. 587–596.
42. Newson, M.D. (1971). A model of subterranean limestone erosion in the british isles based on hydrology. *Institute of British Geographers, Transactions* **54**: 51–70.
43. Jacobson, R.L. and Langmuir, D.L. (1974). Controls on the quality variations of some carbonate spring waters. *Journal of Hydrology* **23**: 247–265.
44. Quinlan, J. (1989). *Groundwater Monitoring in Karst Terranes, Recommended Protocols and Implicit Assumptions*. US Environmental Protection Agency.
45. Gale, S.J. (1984). The hydraulics of conduit flow in carbonate aquifers. *Journal of Hydrology* **70**: 309–327.
46. Scanlon, B.R. and Thrailkill, J. (1987). Chemical similarities among physically distinct spring types in a karst terrain. *Journal of Hydrology* **89**: 259–279.
47. Worthington et al. (1992). Geochemistry of springs in temperate carbonate aquifers: recharge type explains most of the variation. In: *Colloque d'Hydrologie en pays Calcaire et en Millieu Fissure (5th Neuchatel Switzerland), Proceedings*.

Annales Scientifiques de l'Universitie de Besancon, Geologie—Memoires Hors Series, No. 11, pp. 341–347.

48. Jones, W.K. (1997). *Karst Hydrology Atlas of West Virginia*. Karst Waters Institute, Special Publication 4, p. 111.
49. Huntoon, P.W. (2000). Variability of karstic permeability between unconfined and confined aquifers, Grand Canyon region, Arizona. *Environmental and Engineering Geoscience* **6**(2): 155–170.
50. Kass, W. (1998). *Tracing Technique in Geohydrology*. A.A. Balkema, Rotterdam, p. 581.
51. Bednar, D., Jr. and Aley, T., (2001). Groundwater dye tracing: an effective tool to use during the highway development process to avoid or minimize impacts to karst groundwater resources. In: *Geotechnical and Environmental Applications of Karst Geology and Hydrology, Proceedings*. B.F. Beck and J.G. Herring (Eds.). A.A. Balkema, Rotterdam, pp. 201–207.
52. Aley, T. (1999). *The Ozark Underground Laboratory's Groundwater Tracing Handbook*. The Ozark Underground Laboratory, p. 35.
53. Quinlan et al. (1988). How to Use Groundwater Tracing to "Prove" that Leakage of Harmful Materials From a Site in a Karst Terrane Will Not Occur. In: *Environmental Problems in Karst Terranes and Their Solutions Conference Proceedings*. National Water Well Association, pp. 289–301.
54. Aley, T. (1976). The water tracer's cookbook. *Journal of the Missouri Speleological Survey* **16**(6).
55. Smart, P.L. and Laidlaw, I.M.S. (1977). An evaluation of some fluorescent dyes for water tracing. *Water Resources Research* **131**: 15–33.
56. Mull, D.S., Liebermann, T.D., Smoot J.L., and Woosley, L.H., Jr. (1988). *Application of Dye Tracing Techniques for Determining Solute-Transport Characteristics of Ground Water in Karst Terranes*. EPA 904/6-88-001, United States Environmental Protection Agency.
57. Alexander, C.E. and Quinlan, J.F. (1992). *Practical Tracing of Groundwater, with Emphasis on Karst Terranes*. Geological Society of America.
58. Aley, T. (2000). Karst Groundwater. *Missouri Conservationist*, **61**, (3): 8–11.
59. Aley, T. (2000). Water and Land-Use Problems in areas of Conduit Aquifers. In: *Speleogenesis: Evolution of Karst Aquifers*. A.B. Klimchouk et al. (Eds.). National Speleological Society. pp. 481–484.
60. Palmer, A.N. (1999). Introduction. In: *Karst Modeling*. Karst Waters Institute Special Publication No. 5. A.N. Palmer, M.V. Palmer, and I.D. Sasowsky (Eds.). Karst Waters Institute, pp. 1–9.
61. Curl, R.L. (1974). Deducing flow velocity in cave conduits from scallops. *National Speleological Society Bulletin* **36**: 1–5.
62. Mangin, A. (1974–1975). Contribution a l etude hydrodynamique des aquifers kartiques. *Ann Speleol* **29**(3/4): 283–601; **30**(1): 21–124.
63. Mangin, A. (1994). Basic and applied hydrogeological research in french karstic areas. In: *COST 65 ACTION, Hydrogeological Aspects of Groundwater Protection in Karstic Area*. N. Crampon and M. Bakalowicz (Eds.). p. 11–24.
64. Gunn, J.A. (1985). Conceptual model for conduit flow dominate aquifers. In: *Karst Water Resources Proceedings. International Association of Hydrological Science*. Publication No. 161, pp. 587–596.
65. Clemens, T. et al. (1999). Simulation of the development of karst aquifers: role of the epikarst. *International Journal of Earth Science* **88**(1): 157–162.

KARST TOPOGRAPHY

DAVID M. BEDNAR, JR.
Michael Baker, Jr. Inc.
Shreveport, Louisiana

Karst is an irregularly shaped area of land where underground drainage dominates and has resulted in dissolving the bedrock to form distinctive surface and subsurface features. Karst predominately develops on lands underlain by soluble bedrock such as limestone and dolomite, but can also form in marble, gypsum, and halite. Due to differences in climate, cover, rock type, hydrology, and geologic structure, some areas show significant expression of karst development at the surface, and others show no obvious signs at all. Chemical solution of the bedrock by groundwater leads to the formation of open passageways into and through the subsurface. The degree of development of these features varies greatly from one region to another (1).

The term karst is derived from a plateau region located along the northeastern shore of the Adriatic Sea between Slovenia and Italy. Karst topography, it is estimated, covers about 20% of the earth's surface and can be found in the United States, Russia, the Baltic Republics, Great Britain, Japan, China, Vietnam, Australia, Brazil, and Europe. About 25% of the United States is karst, and it is estimated that 40% of the land east of the Mississippi River is karst. Karst topography occurs in areas of Kentucky, Tennessee, southern Indiana, and central and northern Florida.

BEDROCK DISSOLUTION

The development of karst topography begins with the first drop of precipitation from the atmosphere. Water falls through the atmosphere and reacts with carbon dioxide to produce carbonic acid. As the water infiltrates through the soil horizon, the amount of carbonic acid increases due to reactions with organic matter. Water enters the soluble rock through bedding plains and joints and slowly breaks down and transports dissolved mineral matter into the epikarst.

The epikarst or epikarstic zone (sometimes called the subcutaneous zone) is the dissolutionally weathered upper portion of the bedrock that can vary in thickness from zero to about 100 feet or more (2). Groundwater typically moves horizontally in multiple directions untill it reaches enlarged vertical fissures or shafts where the water moves into deeper portions of the karst aquifer system. Depending on topography, the epikarst can store appreciable amounts of water or transmit water quickly. Water storage is greater in gentle topography as opposed to rapid flow in steeper areas. The intensity of epikarstic development, which can be expressed as a percentage of the bedrock that has been removed by dissolution, is highly variable. This percentage can range from 1 to more than 50% (3). However, many epikarstic zones contain some degree of sedimentation. This can range anywhere from 5 to more than 95%.

Based on the results of about 1,000 positive groundwater trace, Aley (3) divided the epikarst into three hydrological types based on water storage capability. Rapid draining epikarsts are characterized by low water storage and voids that are relatively free of fine textured sediment. These areas typically occur in areas of high relief and where the soluble purity of the bedrock is high. Water that is stored seasonally and after major storms is stored in a seasonally saturated epikarst. In these areas, voids are commonly partially or completely filled with fine textured sediment, occur in humid climate or moderate relief, and the elevation of the epikarst is greater than the elevation of the local perennial streams. Perennial saturated epikarsts contain constant water storage, occur in humid climates and areas of low to moderate relief, and are located along perennial streams. However, all epikarstic zones would more than likely contain a combination of the above.

SURFACE FEATURES IN KARST

Dissolution of the bedrock underground can sometimes produce features that are visible at the surface. Sinkholes are the most common surface features in karst areas. However, the lack of sinkholes at the surface does not indicate a lack of karstification in the subsurface. Sinkholes are bowl-shaped depressions on the surface of the earth that vary in size up to 100 meters deep and 1000 meters in diameter. Subsidence sinkholes form from the gradual movement of soil into the bedrock through solutionally enlarged preferential flow routes in the bedrock (Figs. 1a, 1b, and 1c). As soil moves into the bedrock over time, a circular depression forms at the surface. Collapse sinkholes form when the roof of underlying cavities in the bedrock becomes weak from the weight of the overburden and collapses (Figs. 2a, 2b, and 2c). Caprock protected sinks occur when a sandstone unit located immediately above limestone collapses into the cavernous limestone. The rim of the depression is protected by the sandstone, so that these sinkholes tend to be very deep and steep-sided. Compound sinks form as individual sinkholes grow to form a larger sinkhole that has multiple drainage points.

Karren is another distinctive surface feature. As water flows over soluble rocks on the surface, small-scale dissolution grooves, pits, or channels form. Karren comes in many different patterns and is described in detail by Ford and Williams (4). Additionally, karren can form in the subsurface on bedrock beneath a soil cover.

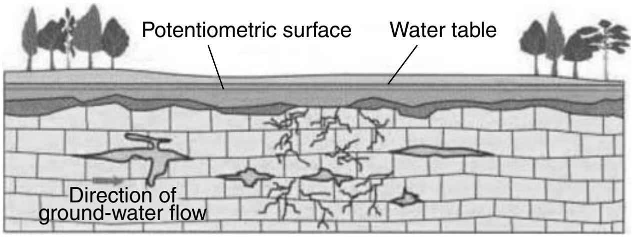

Figure 1(a). Formation of a subsidence sinkhole. From U.S. Geological Survery Report by Galloway et al. (1999).

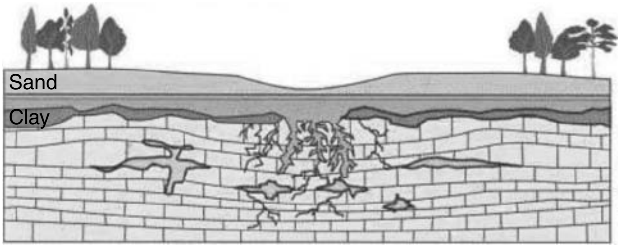

Figure 1(b).

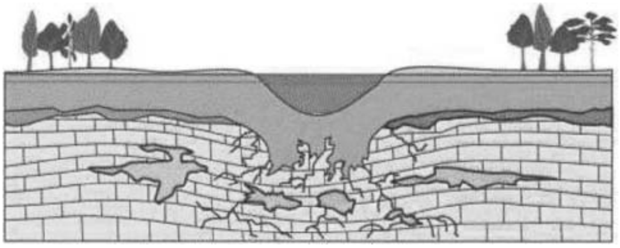

Figure 1(c).

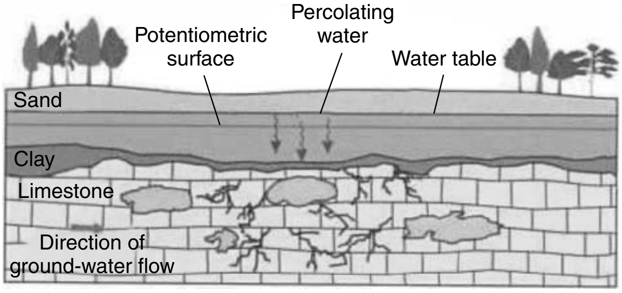

Figure 2(a). Formation of a collapse sinkhole. Form U.S. Gelogical Survery Report by Galloway et al. (1999).

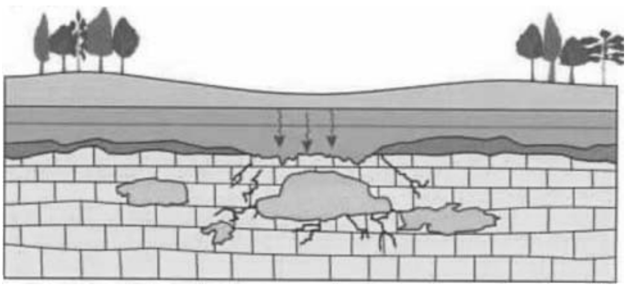

Figure 2(b).

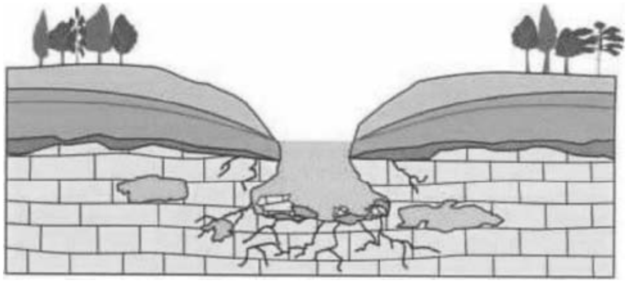

Figure 2(c).

KARST VALLEYS

Three types of distinctive valleys exist in karst areas. Some areas of karst were originally overlain by nonsoluble rocks in topographic valleys. Through time, the overlying rock erodes away by stream action to expose the underlying soluble bedrock in the stream channel. Surface water begins to recharge the groundwater flow system, and the capacity of the groundwater flow system increases to a point where it can handle the entire base flow of the stream. The valley maintains its shape, gradient, and channel, but the discharge from the basin is lost to the groundwater flow system, and it is called a dry valley. Overland flow will occur only during floods. “In some karst river valleys, rivers flow for greater distances before all water is lost to the subsurface at one location. River valleys that end abruptly in this instance are known as blind valleys. The most favorable conditions for the formation of blind valleys are hilly areas of heavy rainfall with well developed streams on impermeable rocks upstream of areas of massive limestone (5).”

Poljes represent another type of karst valley. Meaning “field” in Slovene, poljes are large closed depressions that have well-developed underground drainage systems bounded by steep-sided uplands. Runoff from the surrounding uplands discharges at springs in the polje and forms alluvial streams. These streams sink into caves or swallow holes on the opposite side during periods of low to moderate flow and discharge from them when the groundwater system is at full capacity during periods of high flow.

SPRINGS

Water that disappears into the subsurface in karst areas reappears at the surface at springs. Springs typically discharge along valley floors. Springs can discharge from the contact between two geologic strata, from a solutionally widened opening in the bedrock, or at cave entrances. Quinlan (6) stressed that springs are the pulse of karst aquifers and provide important information about flow regimes. There are two general types of springs in karst areas. Springs that exhibit turbulent flow, respond rapidly to rainfall, and have a highly variable discharge rate are known as diffuse flow springs. Springs that are less turbulent, respond slower to precipitation, and have a low variable discharge rate are called conduit flow springs. However, flow type largely depends on the type of recharge and storage in the karst aquifer as well as the size of the groundwater basin. It must be emphasized that diffuse and conduit flow springs are end members of a continuum and that most springs fall somewhere in the middle (7).

SUBSURFACE FEATURES

Caves and speleothems represent subsurface features in karst areas. A cave is a naturally occurring void in the subsurface that can be easily entered by humans. Caves are formed by the dissolution of the rock along bedding planes and fractures from the circulation of groundwater. Cave formation is affected by such parameters as water quality, tectonics, lithology, topography, geologic structure, climate, and hydrogeologic factors. Additionally, the mode of groundwater recharge helps to enhance cave development.

The shape of cave passageways is dictated by structure and stratigraphy. White (1) placed cave passageways into two broad classes, single conduit and maze passages; each has three distinct geometries. Single conduit caves are classified as those that have linear, angulate, and sinuous passageways. Linear passages are straight without any bends, and the structure of the bedrock controls the path of flow. Angulate passages are characterized by sharp bends and straight segments. Sinuous passages take the pattern of meandering streams, as seen on the surface, that have many curves and few straight segments. Cave patterns found in karst aquifers include branchwork and maze patterns. Sinkholes supply the necessary water flow to form a branchwork (single conduit caves) cave network. In this case, recharge enters the groundwater flow system through several small catchment areas into the subsurface.

Maze-type passages fall into three categories that include network, anastomatic, and spongework mazes. Network mazes consist of intersecting passageways in a uniform grid orientation that are developed by floodwater recharge. This situation is typical where runoff from large areas of insoluble rocks can flow directly into adjacent soluble rocks. The water enters the karst aquifer through a small number of sinking streams each having variable flow. This limited number of input points results in the development of mazes and blind fissures (8). Anastomatic passages look like braided streams on the surface. Spongelike passages are random, interconnected passageways that vary in size and consist of-three-dimensional patterns. The geometry of the cave passage is further dictated by the type of openings through which the water flows, including fractures, bedding plane partings, and intergranular pores.

Speleothems, formerly known as cave formations, form from carbon dioxide rich water that enters the cave atmosphere. Excess carbon dioxide in solution is lost as the groundwater enters the cave atmosphere, and carbonate minerals begin to precipitate out onto the ceiling, walls, or across the floor of the cave. Speleothems are dominantly composed of crystals of calcite and aragonite. Magnesium carbonate deposits are found to a lesser degree. Speleothems are rarely composed of dolomite.

Speleothems are divided into two broad categories, dripstone and flowstone. The common types of dripstones are called stalactites and stalagmites. Stalactites grow as water enters the cave ceiling through joints and bedding-plane partings and slowly drips to the cave floor. Drop by drop, a ring of calcite is deposited to form a hollow tube. Young stalagmites are called soda straws (Fig. 3). Excess water drops onto the floor of the cave and begins to form a mound of calcite known as a stalagmite. Stalagmites form from the cave floor, and successive layers of calcite accumulate upward with time. Sometimes, stalactites and stalagmites come together and form a column. However, this is not always the case. Figure 4 shows an example of a stalactite, stalagmite, and a column.

Figure 3. Soda straws at Mystic Caverns, north central Arkansas.

Figure 4. Stalactite, stalagmites, and a column at Mystic Caverns, north central Arkansas.

Flowstone forms as sheet flows across the cave floor or down the walls of the cave. Draperies (Fig. 5), the most spectacular forms of flowstone, are deposited as water flows over a surface on a cave wall that has an abrupt overhang and deposits in thin trails. Other speleothems within the cave environment include moonmilk, cave coral, rimstone dams, and helictites. Moonmilk is a white unconsolidated deposit of carbonate minerals that resembles cottage cheese when wet (1). Cave coral forms in cave pools when the water becomes oversaturated with calcite and the crystals drop out of solution. Rimstone dams (Fig. 6) are found along the cave floor. These features deposit along the edges of water flowing along the cave floor or at the edges of pools of water where carbon dioxide is lost. Helictites (Fig. 7) are known as eccentric speleothems that defy the law of gravity and grow in any direction from solutions fed along tiny capillaries. They are commonly very short and have a sinuous to curvilinear appearance.

Figure 5. Massive draperies at Mammoth Cave Kentucky.

Figure 6. A rimstone dam at Hidden River Cave in north central Arkansas.

Figure 7. Helictites—Mystic Caverns, north central Arkansas.

GROUNDWATER RECHARGE

Four dominant types of groundwater recharge are recognized in karst areas. Diffuse recharge enters the groundwater flow system across a broad area. Discrete recharge, also called concentrated recharge, is restricted to small areas, especially through sinkholes or losing streams. Substantially greater quantities of water per unit area enter the groundwater flow system through discrete recharge areas than through diffuse recharge (9). Usually there are several sinking points, or swallow holes, along the course of sinking streams. Surface water is continually lost to the groundwater flow system throughout the year during periods of moderate and high flow. This occurs because the groundwater flow system has reached full capacity and cannot transmit all surface water flow. However, during the summer months, overland flow decreases, and a portion of the stream becomes lost. This water loss through swallow holes or openings in the underlying bedrock eventually reappears at the surface at springs or multiple springs. Surface water that flows from adjacent non-karst areas onto karst is called allogenic recharge. Precipitation that falls directly on the karst that enters the groundwater flow system is known as autogenic recharge. The former two recharge types describe how the water enters the subsurface while the latter two types refer to the where the water originated.

TYPES OF KARST SETTINGS

Quinlan (10,11) noticed that karst settings vary greatly in appearance and classified them into two major types based on cover type. Covered karsts are areas where the bedrock is covered by some material and not exposed at the surface. Four subcategories of covered karst include mantled karst, buried karst, interstratal karst, and subaqueous karst. Mantled karst occurs when a thick layer of unconsolidated sediments partially or wholly covers the bedrock. Buried karst, also know as paleokarst, is an ancient karst landscape that was completely buried by younger rocks. In some instances, soluble rocks are buried beneath less soluble or insoluble bedrock. However, due to circulating groundwater, these rocks are eventually dissolved and are defined as interstratal karst. Karst that forms beneath bodies of water such as river, lakes, or within tidal zones is called subaqueous karst.

Exposed karsts are areas where sediment cover is absent and the rock is exposed at the surface. Four subcategories of exposed karst include naked karst, denuded karst, exhumed karst, and relict karst. Naked karst occurs primarily in alpine regions where soils are poorly developed (1). Denuded karst is interstratal karst that has been exposed at the surface. Karst that was buried by sediment and then later exposed by erosion is called exhumed karst. Relict karst refers to areas where the bedrock has been removed by erosion without any trace of topographic expression.

Additional karst settings identified include deep-seated karst, subjacent karst, and entrenched karst. Deep-seated karst is found buried beneath younger rocks. There is no evidence at the surface, and the soluble rocks are not exposed. Subjacent karst occurs when a portion of the soluble rocks is exposed and surface features are evident. Entrenched karst occurs where the entire thickness of the soluble rock is entrenched along valleys, but the insoluble cap remains over most of the interfluves (7).

VULNERABILITY OF KARST TO CONTAMINATION AND POLLUTION

Karst areas are highly vulnerable to groundwater contamination. Spills that occur on karst can enter the groundwater flow system through discrete recharge zones and travel miles within days or weeks to springs and wells used as sources of water. The high permeability of a karst aquifer provides ineffective natural cleansing of any contaminants transported through it. Contaminated soils have the potential to be washed into the groundwater flow system in diffuse recharge zones without any biological interaction with the soil horizon. Water in storage is more subject to long-term contamination than water in transit (9). Groundwater moves slower under low gradient conditions, and cleanup of contaminated aquifers can last for years.

Keeping vulnerable karst aquifers free of contaminants for water supplies is essential during proper land use planning to minimize impacts to the karst groundwater flow system. In recent years, groundwater dye tracing has proven to be a useful tool for characterizing groundwater flow systems in karst. The delineation of recharge areas for important springs is essential during land use planning.

BIBLIOGRAPHY

1. White, W. (1988). *Geomorphology and Hydrology of Karst Terrains*. Oxford University Press, New York.
2. Aley, T. (2000). Karst systems. In: *Standard Handbook of Environmental Science, Health, and Technology*. J.H. Lehr and J.K. Lehr. (Eds.). McGraw-Hill, New York, p. 1600.
3. Aley, T. (1997). Groundwater tracing in the epikarst. In: *The Engineering Geology and Hydrogeology of Karst Terranes*. Beck and Stephenson (Eds.). Balkema, Rotterdam, pp. 207–211.
4. Ford, D.C. and Williams, P.W. (1989). *Karst Geomorphology and Hydrology*. Chapman and Hall, London.
5. Sweeting, M.M. (1973). Columbia University Press.
6. Quinlan, J. et al. (1992). Recommended administrative/regulatory definition of karst aquifer, principles for classification of carbonate aquifers, practical evaluation of vulnerability of karst aquifers, and determination of optimum frequency at springs. In: *Conference on Hydrogeology, Ecology, Monitoring, and Management of Groundwater in Karst Terranes*. Quinlan and Stanley, A. (Eds.). National Groundwater Association, pp. 573–635.
7. Klimchouk et al. (2000). Speleogenesis. *Evolution of Karst Aquifers*. National Speleological Society, Inc., p. 527.
8. Palmer, A.N. (2000). Hydrogeologic control of cave patterns. In: *Speleogenesis: Evolution of Karst Aquifers*. Klimchouk et al. (Eds.) National Speleological Survey, Inc., pp. 77–90.
9. Aley, T. (1978). Ozark hydrology: a predictive model. *Journal of the Missouri Speleological Survey* **18**: 185 pages.
10. Quinlan, J. (1967). Classification of Karst Types: A review and synthesis emphasizing the North American Literature,

1941–1966. *National Speleological Society Bulletin* **29**: 107–109.

11. Quinlan, J. (1978). *Types of Karst, with emphasis on cover beds in their classification and development*. PhD Thesis, Geology, University of Texas, Austin, TX.

READING LIST

Jennings, J.N. (1971). *Karst*. MIT Press, Cambridge, MA.

DETECTING MODERN GROUNDWATERS WITH ^{85}Kr

Craig E. Divine
Geoffrey Thyne
Colorado School of Mines
Golden, Colorado

NATURAL PRODUCTION MECHANISMS AND DECAY BEHAVIOR OF ^{85}Kr

Krypton-85 (^{85}Kr) is a relatively short-lived radioactive isotope of the noble gas krypton (Kr) that is produced naturally in very minor amounts in the atmosphere by spallation reactions (n, γ) with the stable Kr isotope ^{84}Kr. ^{85}Kr decays to rubidium-85 (^{85}Rb) by beta release:

$$^{85}\mathrm{Kr} \longrightarrow {}^{85}\mathrm{Rb} + \beta^- \qquad (1)$$

The decay of ^{85}Kr follows a first-order decay rate given by

$$^{85}\mathrm{Kr_t} = {}^{85}\mathrm{Kr_0}e^{-\lambda t} \qquad (2)$$

where $^{85}\mathrm{Kr_t}$ is the ^{85}Kr concentration at time t (the elapsed time), $^{85}\mathrm{Kr_0}$ is the initial ^{85}Kr concentration, and λ is the ^{85}Kr decay constant (0.06442/yr). The ^{85}Kr half-life ($t_{1/2}$, 10.76 yr) is the time it takes for half of the initial ^{85}Kr to decay to ^{85}Rb and is related to the decay constant by

$$\lambda = \frac{\ln 2}{t_{1/2}} \qquad (3)$$

Krypton is chemically inert, and the only significant sink for atmospheric ^{85}Kr is radioactive decay.

ANTHROPOGENIC SOURCES AND ATMOSPHERIC DISTRIBUTION OF ^{85}Kr

Since the 1950s, anthropogenic contributions of ^{85}Kr to the atmosphere (primarily the result of nuclear weapons testing and nuclear fuel rod reprocessing) have completely overwhelmed natural background concentrations. Worldwide release of ^{85}Kr increased from less than 5 petabecquerels (PBq) per year in 1950 to more than 350 PBq per year in 1986 (1). Although atmospheric ^{85}Kr concentrations vary spatially with strong latitudinal correlation (higher concentrations occur in the Northern Hemisphere due to minimal ^{85}Kr releases in the Southern Hemisphere), the most significant characteristic is the nearly linear increase in concentration (Fig. 1). In contrast with the declining input signal for the tritium groundwater dating method, the increasing ^{85}Kr input signal will facilitate robust apparent age calculations well into the foreseeable future.

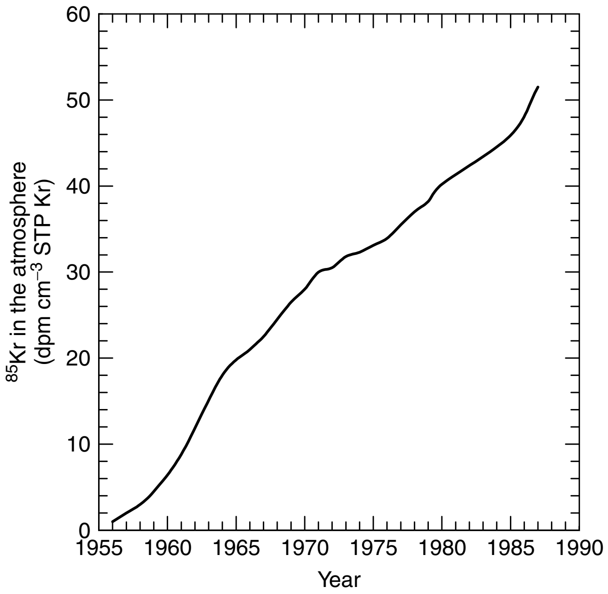

Figure 1. Activity of ^{85}Kr in the atmosphere of the Northern Hemisphere (after Reference 2).

PRINCIPLES, ADVANTAGES, AND DISADVANTAGES OF THE ^{85}Kr DATING METHOD

Anthropogenically related ^{85}Kr overwhelmingly dominates natural background activity in the atmosphere; therefore, the presence of ^{85}Kr in groundwater clearly indicates that the water is relatively young (i.e., recharged after 1950). Furthermore, the relatively simple behavior of the atmospheric concentration history permits estimating exact recharge dates to approximately 1960 by measuring the amount of ^{85}Kr activity in groundwater and accounting for radioactive decay. If dispersion is negligible, an age can be estimated by plotting the measured ^{85}Kr activity in groundwater samples back to an intercept point on the atmospheric ^{85}Kr activity curve (3). Alternatively, a decay-corrected ^{85}Kr activity curve can be plotted for the sample collection date directly relating measured activity to an apparent groundwater age (Fig. 2). For the ^{85}Kr method, the “clock” begins at the seasonal high water table because it is assumed that the ^{85}Kr concentration in the groundwater is in equilibrium with soil gas and the atmosphere (4).

A primary advantage of ^{85}Kr is that it is chemically inert in groundwater and is therefore affected only by the hydraulic properties of the aquifer and radioactive decay (5). Furthermore, age calculations do not require recharge temperature and krypton solubility information because the ^{85}Kr activity is normalized to the total ^{85}Kr present, which also implies that errors from excess air trapped in the aquifer or sample loss are negligible.

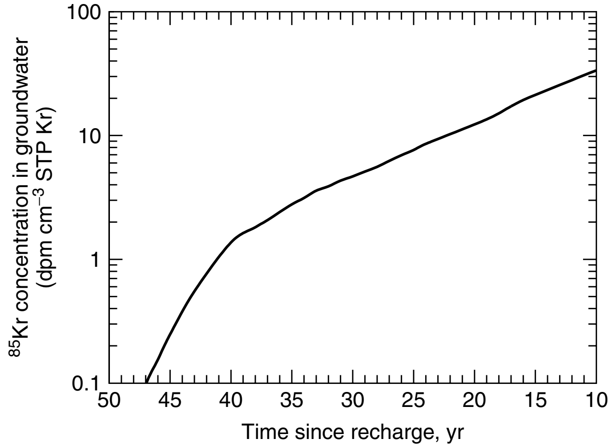

Figure 2. Relationship between apparent age (time since recharge) decay-corrected ^{85}Kr activity for samples collected in 2004 based on the activity of ^{85}Kr in the atmosphere of the Northern Hemisphere (Fig. 1).

However, because of the high current ^{85}Kr activity in the atmosphere, care must be taken to prevent sample contamination during well installation and sample collection. Note that the ^{85}Kr method may overestimate the groundwater age for deep-water tables (greater than 10 meters) due to a time lag through the unsaturated zone (3).

A significant disadvantage of the ^{85}Kr method is the large volume of groundwater required for analysis (~100 L) due to its relatively low aqueous solubility. Furthermore, analytical costs are relatively high: typically \$500–1,500 per sample (4). Water samples must be degassed in the field by vacuum extraction and isolated from the atmosphere. Sampling and analytical methods are discussed by Smethie and Mathieu (6). Ekwurzel et al. (5) report an analytical precision for ^{85}Kr of ±2.9 disintegrations per minute per cubic centimeters of total Kr (dpm cm^{-3} STP Kr).

CASE STUDY: COMPARISON OF THE ^{85}Kr METHOD TO THE ^{3}H–^{3}HE AND CFC DATING METHODS

Ekwurzel et al. (5) compared apparent groundwater ages for several different tracer dating methods, including the ^{85}Kr method, in a study on the Delmarva Peninsula located on the east coast of the United States. The hydrogeology of the area is relatively simple and consists of highly permeable surficial materials and shallow water tables with assumed vertical flow velocities near the water table interface. Discrete groundwater samples were collected in support of tritium–helium (^{3}H–^{3}He) and chlorofluorocarbon (CFC, specifically CFC-11 and CFC-12) dating methods at approximately 30 wells throughout the Delmarva Peninsula. Most wells were screened across less than 1 m of the aquifer and therefore represent relatively discrete intervals of the aquifer. Four ^{85}Kr samples were collected in addition to the other tracers at a local well-characterized flow system within the Delmarva Peninsula (Locust Grove).

Table 1. Comparison of Apparent Groundwater Ages (in years) at Locust Grove, Delmarva Peninsula for Various Dating Methods[a]

	Dating Method			
Well Name	^{85}Kr	^{3}H–^{3}He	CFC-11	CFC-12
KE Be 52	7.7	7.1	8.1	8.4
KE Be 61[b]	14.4	16.7	19.7	20.3
KE Be 62	5.4	3.1	3.2	3.6
KE Be 163	5.7	7.8	9.3	9.2

[a] Reference 5.

[b] Apparent ages are averaged for well KE Be 61 with multiple sampling for the ^{3}H–^{3}He and CFC methods.

The ^{85}Kr method yielded apparent ages that agreed closely (generally within ~30%) with the results of the other dating methods (Table 1). The effect of dispersion on apparent ages was modeled and was relatively insignificant for the ^{85}Kr method for ages younger than 25 years based on site conditions. The investigators also demonstrate that use of multiple methods for groundwater age dating may be advantageous because tracers are affected differently by potential error sources, such as diffusion, dispersion, sorption, degradation, gas entrapment, excess air, recharge temperature, and sampling errors. For additional discussion and case studies on the ^{85}Kr groundwater dating method, see Loosli et al. (7), Loosli (2), Smethie et al. (8), Cook and Solomon (3), Clark and Fritz (9), and Cook and Solomon (4).

BIBLIOGRAPHY

1. Weiss, W., Sartorius, H., and Stockburger, H. (1992). Global distribution of atmospheric ^{85}Kr. In: *Isotopes of Noble Gases as Tracers in Environmental Studies: IAEA Proc. Consultants Meet.*, May 29–June 2, 1989, Vienna, pp. 29–62.
2. Loosli, H.H. (1992). Applications of ^{37}Ar, ^{39}Ar, and ^{85}Kr in hydrology, oceanography, and atmospheric studies. In: *Isotopes of Noble Gases as Tracers in Environmental Studies: IAEA Proc. Consultants Meet.*, May 29–June 2, 1989, Vienna, pp. 29–62.
3. Cook, P.G. and Solomon, D.K. (1995). Transport of atmospheric trace gases to the water table: Implications for groundwater dating with chlorofluorocarbons and krypton 85. *Water Resour. Res.* **31**: 263–270.
4. Cook, P.G. and Solomon, D.K. (1997). Recent advances in dating young groundwater: Chlorofluorocarbons, ^{3}H/^{3}He and ^{85}Kr. *J. Hydrol.* **191**: 245–265.
5. Ekwurzel, B. et al. (1994). Dating of shallow groundwater: Comparison of the transient tracers ^{3}H/^{3}He, chlorofluorocarbons, and ^{85}Kr. *Water Resour. Res.* **30**: 1693–1707.
6. Smethie, W.M. and Mathieu, G. (1986). Measurement of krypton-85 in the ocean. *Mar. Chem.* **18**: 17–33.
7. Loosli, H.H., Lehmann, B.E., and Dappen, G. (1991). Dating by radionuclides. In: *Applied Isotope Hydrology. A Case Study in Northern Switzerland. Studies in Environmental Science*. J. Pearson et al. (Eds.). Elsevier, Amsterdam, **43**: 153–174.
8. Smethie, W.M., Solomon, D.K., Schiff, S.L., and Mathieu, G.G. (1992). Tracing groundwater flow in the Bordon aquifer using krypton-85. *J. Hydrol.* **130**: 279–297.
9. Clarke, I.D. and Fritz, P. (1997). *Environmental Isotopes in Hydrogeology*. CRC Press, New York, pp. 20, 174–188.

LAND USE IMPACTS ON GROUNDWATER QUALITY

Michael D. Trojan
Minnesota Pollution Control Agency
St. Paul, Minnesota

The manner in which humans use land has important consequences for groundwater quality. Using land for waste disposal, recreation, agriculture, residential development, and industry releases different types and quantities of chemicals to the environment. These chemicals often end up in shallow groundwater, where they may impair water quality.

Groundwater impacts from point sources, such as landfills and hazardous waste sites, are well understood. Impacts from "routine" land use, such as residential, commercial, industrial, or agricultural use, are not fully understood or readily apparent. Sources of chemicals from these land uses are called nonpoint or areawide sources. Examples include fertilized urban lawns or all feedlots in an area. This discussion focuses on water quality impacts in shallow groundwater resulting from routine land use.

To assess water quality, we compare chemical concentrations in water to a standard. Most standards are designed to protect human or ecosystem health, but standards also include ambient or natural concentrations of a chemical. For example, a nitrate concentration of 2 mg/L in groundwater under an urban subdivision is well below the drinking water standard of 10 mg/L but may be more than twice the concentration found in areas not impacted by humans.

Resource managers are increasingly concerned about impacts of human activity on groundwater quality. Between 1982 and 1997, the United States population grew about 15%. During that same period, the percentage of developed land, which includes urban land and land used for transportation, increased 34% to just over 98 million acres, which represents 6.6% of the nonfederal land in the United States (1). Urban development is only one of several land use changes that have the potential to impact groundwater quality. Others include changes in agricultural fertilizer and herbicide use, changes in the number of acres under irrigation, and changes in atmospheric deposition resulting from automobile and industrial air emissions. Groundwater impacts from land use are greatest where aquifers are sensitive to contamination. Examples include fractured bedrock or glacial aquifers that are close to the land surface and overlain by permeable geologic material.

UNDEVELOPED AREAS

Undeveloped areas include forested lands, grasslands, wetlands, and other areas not used for urban or agricultural development. These areas are minimally impacted by humans. Water quality from undeveloped areas serves as a benchmark for comparison with other land uses. Typically, pesticides and volatile organic compounds (VOCs) are absent, and concentrations of inorganic chemicals are low unless local geologic deposits are enriched in a particular element (Table 1) (2–5). Variability in chemical concentrations under undeveloped land use is low compared with that of other land uses (6).

Synthetic organic chemicals are occasionally found in undeveloped areas. Chloroform is the most commonly found VOC, atrazine and its primary metabolite, deethylatrazine, are the most common herbicides. These chemicals may be the result of atmospheric deposition (3,7). Chlorinated insecticides, such as dieldrin and DDE, are sometimes detected in groundwater under undeveloped areas and represent historic use of these persistent chemicals (3).

URBAN LAND USE

Urban areas include residential, commercial, and industrial land use. Distinctions between urban land uses are not always clear. Residential areas often include commercial development and may include small tracts of industrial or agricultural land. Residential areas served by municipal sewers (sewered) and septic systems (nonsewered) may be mixed. This intermingling of land uses makes it difficult to associate groundwater quality with a specific land use.

Studies of water quality in urban areas rarely distinguish between different types of urban land use. Consequently, Table 1 shows a wide range in chemical concentrations. Concentrations of major ions, trace elements, pesticides, and VOCs exceed concentrations found in undeveloped areas. Prometon, simazine, dicamba, bromacil, and 2,4-D are the most commonly detected herbicides (7). Prometon, bromacil, and simazine are applied for weed control in rights-of-way, and dicamba and 2,4-D are applied to urban lawns (8). These chemicals are found primarily during the growing season. The organochlorine insecticides, dieldrin and chlordane, are often observed in groundwater beneath urban areas (7,9). Atrazine and deethylatrazine are frequently observed in urban areas (10). Although used on agricultural crops, atrazine is present in precipitation in many areas of the United States (8,11,12). Concentrations of pesticides in urban ground water are typically well below drinking water standards, although standards do not exist for many of the compounds detected (13).

The widespread occurrence of VOCs is the most conspicuous difference between undeveloped and urban areas. Chlorinated hydrocarbons, particularly 1,1,2-trichloroethene (TCE), 1,1,2,2-tetrachloroethene (PCE), and 1,1,1-trichloroethane (TCA), account for the majority of VOC detections. These persistent chemicals were used as degreasers, in dry cleaning, or in septic systems (3,6). Trihalomethanes, particularly chloroform, also occur within urban areas. They are byproducts of disinfection, but other sources include atmospheric deposition and lawn watering with chlorinated water (14). Nonchlorinated hydrocarbons, such as benzene and substituted benzenes, occur in areas where underground fuel tanks have contaminated the groundwater (13).

Table 1. Summary of Median Chemical Concentration, Range in Concentration, and Number of Data Points for Three Land Uses

Chemical[a]	Median			Range			Number of Data Points		
	Und.[b]	Urb.	Agr.	Und.	Urb.	Agr.	Und.	Urb.	Agr.
Major Ions (mg/L)									
Calcium	9	89	60	5–63	20–123	16–92	3	5	5
Chloride	2.4	63	20	1.8–3.0	22–83	10–41	2	4	4
Magnesium	22	26	23	22	18.7–29	22–25	1	3	3
Nitrate	0.6	2.4	4.1	0.5–1.0	1.4–6.1	1.3–15	3	10	11
Potassium	1.1	2.6	2.0	1.1	1–3.2	1.3–5	1	3	3
Sodium	3.3	22	7.4	1–5.6	12–63	4.3–35	2	4	4
Sulfate	8	17	13	3–13	7.1–72	6.3–40	2	4	4
Synthetic Organic Chemicals (% of Samples Detected)									
Total PAHs	0	0	0	0	0–33	0	1	2	2
Total pesticide	4.4	32	86	0–35	0–70	29–100	4	9	8
Total VOCs	4	66	7	0–15	3–100	0–9	3	7	5
Trace Elements (ug/L)									
Arsenic	0.7	1.4	2.9	0.7	0.7–2	0.7–5	1	5	4
Boron	16	56	39	15–17	41–71	21–80	2	3	3
Cadmium	0.1	0.09	2.0	0.1	0.04–1.5	0.04–4	1	5	4
Chromium	1.1	3.1	3.4	1.1	1.1–6	0.59–6	1	5	4
Copper	<5.4	2	5.4	<5.4	2– < 5.4	5.4	1	6	2
Iron	10	34	4.3	10	5.6–63	3.6–4.9	1	3	2
Lead	0.04	0.06	0.027	0.04	0.05–0.07	0.027	1	5	2
Manganese	0.6	5.0	1.7	0.6	0.5–440	0.5–2.6	1	4	2
Zinc	5.9	5.5	6.6	5.9	5.1–9.1	6–7.1	1	4	2

[a] Data are compiled from many sources cited in references.

[b] Und. = undeveloped, urb. = urban, agr. = agriculture.

There are water quality differences between the three urban land uses (Table 2). Nitrate, phosphorus, and boron concentrations are higher in nonsewered areas than in sewered areas (5,15,16). In nonsewered subdivisions that are more than 10 years old and have lot sizes of 1 acre or less, nitrate concentrations often exceed the Maximum Contaminant Level (MCL) in 10% or more of sampled wells (17–20). Manganese concentrations are higher, and VOCs occur more frequently in sewered areas (5). Concentrations of TCE and PCE may exceed drinking water standards in older sewered areas where industries used these chemicals (7).

AGRICULTURAL AREAS

Agriculture includes cropland, pasture, rangeland, and areas managed for silviculture. Cropland can be divided into different cropping types and irrigated or dryland agriculture.

Nitrate concentrations and frequency of herbicide detection in agricultural areas are higher than in undeveloped and urban areas. Concentrations of most other chemicals are higher than concentrations in undeveloped areas but lower than concentrations in urban areas. VOCs are detected less frequently than in urban areas, whereas organochlorine insecticides occur at similar frequencies. DDT, DDE, dieldrin, and heptachlor epoxide are the organochlorine chemicals found most frequently.

Table 2. Summary of Median Chemical Concentrations for Different Urban Land Uses

Chemical[a]	Sewered Residential	Nonsewered Residential	Commercial-Industrial
Major Ions (mg/L)			
Calcium	105	73	74
Chloride	79	83	42
Magnesium	29	19	28
Nitrate	2.4	6.1	2.0
Potassium	2.9	1.0	3.2
Sodium	26	63	17
Sulfate	18	7.1	12
Synthetic Organic Chemicals (% of Samples Detected)			
Total PAHs	0	0	33
Total pesticide	10	0	7.5
Total VOCs	52	3.0	73
Trace Elements (ug/L)			
Arsenic	0.70	0.70	0.49
Boron	41	71	57
Cadmium	0.09	0.04	0.07
Chromium	1.2	1.1	0.81
Copper	<5.4	<5.4	<5.4
Iron	5.6	<3.1	9.3
Lead	0.07	0.05	0.06
Manganese	8.0	0.45	20
Phosphorus	0.027	0.035	0.020
Zinc	9.1	5.1	7.9

[a] Data from Ref. 5.

Carbamate insecticides occur in areas where they are applied (Table 1) (3,6).

Atrazine, deethylatrazine, and metolachlor are the most commonly observed herbicides in agricultural areas (7,10,21). The presence of other herbicides relates to regional use and includes metribuzin, EPTC, bentazon, diuron, 2,4-D, and dicamba (14,21,22). Recent studies suggest that herbicide metabolites occur at much greater frequency than parent compounds (5). Prometon, simazine, and tebuthiuron are used in road rights-of-way and occur in groundwater from agricultural areas, but at frequencies lower than those in urban areas.

Concentrations of both nitrate and herbicide are directly proportional to the amount of these chemicals applied to cropland (23). Nitrate concentrations are about 1 mg/L under noncropland, 1–3 mg/L under small grains and soybeans, 3–7 mg/L under nonirrigated row crop agriculture, and more than 10 mg/L under irrigated agriculture (5,22,24,25). Nitrate concentrations tend to be lower in areas that have extensive feedlots due to denitrification associated with organic carbon released from the feedlots. Total nitrogen concentrations in these areas, however, may exceed concentrations in areas lacking feedlots (5,24). Herbicides are generally absent under rangeland and forestland used for silviculture; concentrations under irrigated agriculture are about double those under nonirrigated agriculture. Pesticide concentrations typically are lower than drinking water standards, although standards do not exist for many chemicals.

GEOGRAPHIC AND DEPTH RELATIONSHIPS

The land use relationships discussed before are independent of geographic location, except that the chemicals differ locally (6,26). For example,

- Methyl tert-butyl ether (MtBE) is a VOC frequently found in groundwater but only in areas where the chemical is used as an additive in gasoline (7,13).
- Very densely populated urban areas have concentrations of major ions higher than those in less densely populated urban areas (6).
- Local pumping of surficial aquifers induces flow of water from different areas, resulting in a groundwater chemical signature that may not reflect the overlying land use (6).
- Specific synthetic organic compounds are often found in areas where specialty crops are grown. For example, 1,2-dichloropropane was frequently used as a fumigant for potatoes, berries, and other specialty crops. It is found in trace quantities in areas where it was applied to crops (13,22).

Impacts in shallow groundwater do not necessarily relate to water quality deeper in an aquifer. Many aquifers are protected by confining geologic layers or favorable geochemical conditions. Many aquifers, however, do not have adequate protection. Sensitive fractured bedrock aquifers, in particular, are affected by land use.

INFORMATION NEEDS

There are few studies that compare water quality beneath different land uses. Land use studies are complicated by concerns about where in an aquifer to collect samples, well spacing, defining discrete land uses, historical land use, and selection of a parameter list. Data for trace elements are generally lacking, although land use affects concentrations of arsenic, boron, lead, and manganese in ground water (3,5). Recent studies suggest that groundwater may contain chemicals not previously sampled, such as pesticide degradates; chemical additives such as tetrahydrofuran and MtBE; urban insecticides such as diazinon, carbaryl, chlorpyrifos, and malathion; pharmaceuticals; growth hormones; and antibiotics (3,5).

There is also a need for trend information. We do not understand rates of water quality change associated with land use changes. For example, nitrate loss from newly installed lawns is high the first year after establishment, particularly if the lawns are seeded. Other studies show that nitrate loss under established turfgrass is very low (5,27). Rates of change from these two conditions are not known, however. Some trend studies exist for the occurrence of nitrate in groundwater in agricultural and nonsewered residential areas (10,22,28–30). Concentrations of VOCs in urban areas should be decreasing in response to aggressive cleanup and control programs in the last 20 years, but there is limited data to support this assumption.

Finally, we need to understand the effectiveness of management strategies in protecting groundwater quality. For example, studies show that converting land to CRP improves water quality (5). There is less information on the effects of other agricultural BMPs, such as conservation tillage, crop rotation, and nutrient management. In nonsewered urban areas, increasing lot size results in lower nitrate concentrations in groundwater, but increased use of pesticides and fertilizers on larger urban lawns may offset this. In industrial areas, regulatory programs such as the Superfund have resulted in cleanup of soils and groundwater at contaminated sites, but there is limited information on the effectiveness of these cleanups in protecting water quality in aquifers.

SUMMARY

Research consistently shows that chemical concentrations in groundwater increase as the quantity used increases. The type and intensity of land use are thus the most important factors affecting groundwater quality in an area. Water resource managers need to be aware of the potential impacts of different land uses on groundwater quality.

BIBLIOGRAPHY

1. United States Department of Agriculture. (2000). *Summary Report—1997 National Resources Inventory*. Natural Resources Conservation Service, Iowa State University Statistical Laboratory, Ames, IA.

2. Clawges, R.M. and Vowinkel, E.F. (1996). Variables indicating nitrate contamination in bedrock aquifers, Newark basin, New Jersey. *Water Resources Bulletin* **32**: 1055–1066.
3. Eckhardt, D.A.V. and Stackelberg, P.E. (1995). Relation of ground-water quality to land use on long island, New York. *Ground Water* **33**(6): 1019–1033.
4. Taraba, J.L. and Dinger, J.S. (1998). Land use impacts on water quality in small karst agricultural watersheds. *University of Kentucky Waterworks Newsletter* **4**(2): 1–5.
5. Trojan, M.D. et al. (2002). Effects of land use on ground water quality in the anoka sand plain aquifer of Minnesota. Accepted for publication in Ground Water.
6. Cain, D., Helsel, D.R., and Ragone, S.E. (1989). Preliminary evaluations of regional ground-water quality in relation to land use. *Ground Water* **27**(2): 230–244.
7. Ayers, M.A., Kennen, J.G., and Stackelberg, P.E. (2000). *Water Quality in the Long Island-New Jersey Coastal Drainages*. New York and New Jersey, 1996–98. United States Geological Survey Circular 1201.
8. Andrews, W.J., Fong, A.L., Herrod, L., and Dittes, M.E. (1997). *Water-Quality Assessment of Part of the Upper Mississippi River Basin, Minnesota and Wisconsin—Ground-Water Quality in an Urban Part of the Twin Cities Metropolitan Area*. Minnesota, 1996. United States Geological Survey Water-Resources Investigations Report 97–4248.
9. Frick, E.A., Hippe, D.J., Buell, G.R., Couch, C.A., Hopkins, E.H., Wangsness, D.J., and Garrett, J.W. (1998). *Water Quality in the Appalachicola-Chattahoochee-Flint River Basin, Georgia, Alabama, and Florida, 1992–95*. U.S. Geological Survey Circular 1164.
10. U.S. Geological Survey. (1999). *The Quality of Our Nation's Waters—Nutrients and Pesticides*: U.S. Geological Survey Circular 1225.
11. Capel, P.D., Lin, M., and Wotzka, P.J. (1998). *Wet atmospheric deposition of pesticides in Minnesota. 1989–94*. U.S. Geological Survey Water Resources Investigations Report 97–4026, p. 43.
12. Ligocki, M.P., Leuenberger, C., and Pankow, J.F. (1985). Trace organic compounds in rain. II. Gas scavenging of neutral organic compounds. *Atmospheric Environment* **19**(10): 1609–1617.
13. Ebbert, J.C., Embrey, S.S., Black, R.W., Tesoriero, A.J., and Haggland, A.L. (2000). *Water Quality in the Puget Sound Basin, Washington and British Columbia, 1996–98*. United States Geological Survey Circular 1216.
14. Domagalski, J.L., Knifong, D.L., Dileanis, P.D., Brown, L.R., May, J.T., Connor, V., and Alpers, C.N. (2000). *Water Quality in the Sacramento River Basin, California, 1994–98*. United States Geological Survey Circular 1215, p. 37.
15. Hinkle, S.R. (1997). *Quality of Shallow Ground Water in Alluvial Aquifers of the Williamette Basin, 1993–95*. United States Geological Survey Water Resources Investigation Report 97–4082.
16. Lichtenberg, E. and Shapiro, L.K. (1997). Agriculture and nitrate concentrations in Maryland community water system wells. *Journal of Environmental Quality* **26**(1): 145–153.
17. Anderson, D.L., Rice, J.M., Voorhees, M.L., Kirkner, R.A., and Sherman, K.M. (1987). Ground water modeling with uncertainty analysis to assess the contamination potential from onsite sewage disposal systems (OSDS) in Florida. *Proceedings from the National Symposium on Individual and Small Community Sewage Systems*. December 14–15. Chicago, IL, p. 411.
18. Bauman, B.J. and Schafer, W.M. (1984). Estimating groundwater quality impacts from on-site sewage treatment systems. *Proceedings of the Fourth National Home Sewage Treatment Symposium*. ASAE, pp. 285–294.
19. Hantzsche, N.N. and Finnemore, E.J. (1992). Predicting ground-water nitrate-nitrogen impacts. *Ground Water* **30**: 490–499.
20. Quan, E.L., Sweet, H.R. and Illian, J.R. (1974). Subsurface Sewage Disposal and Contaminantion of Ground Water in East Portland, Oregon. *Ground Water* **12**: 356–366.
21. Anderson, H.W. (1993). *Effects of Agriculture and Residential Land Use on Ground-Water Quality, Anoka Sand Plain Aquifer, East-Central Minnesota*. United States Geological Survey Water-Investigations Report 93–4074.
22. Clark, G.M., Maret, T.R., Rupert, M.G., Maupin, M.A., Low, W.H., and Ott, D.S. (1998). *Water Quality in the Upper Snake River Basin, Idaho and Wyoming 1992–95*. U.S. Geological Survey Circular 1160.
23. Barbash, J.E., Thelin, G.P., Kolpin, D.W., and Gilliom, R.J. (1998). *Distribution of Major Herbicides in Ground Water of the United States*. Water-Resources Investigations Report 98–4245 8056–61.
24. Burkart, M.R. and Kolpin, D.W. (1993). Hydrologic and land-use factors associated with herbicides and nitrate in near-surface aquifers. *Journal of Environmental Quality* **22**(4): 646–656.
25. Frenzel, S.A., Swanson, R.B., Huntzinger, T.L., Stamer, J.K., Emmons, P.J., and Zelt, R.B. (1998). *Water Quality in the Central Nebraska Basins, Nebraska, 1992–95*. U.S. Geological Survey Circular 1163.
26. Hamilton, P.A. and Helsel, D.R. (1995). Effects of agriculture on ground-water quality in five regions of the United States. *Ground Water* **33**(2): 217–226.
27. Geron, C.A. et al. (1993). The effects of establishment methods and fertilization practices on nitrate leaching from turfgrass. *JEQ*. **22**: 119–125.
28. Keeney, D.R. and DeLuca, T.H. (1993). Des moines river nitrate in relation to watershed agricultural practices: 1945 versus 1980s. *JEQ*. **22**: 267–272.
29. Kitchen, N.R., Blanchard, P.E., Hughes, D.F., and Lerch, R.N. (1997). Impact of historical and current farming systems on groundwater nitrate in Northern Missouri. *Journal of Soil and Water Conservation* **52**(4): 272–277.
30. Miller, J.C. (1975). Nitrate contamination of the water-table aquifer by septic tank systems in the coastal plain of Delaware. In: *Water Pollution Control in Low Density Areas*. W.S. Jewell and R. Swan. (Eds.). *Proc. Rural Environ. Eng. Conf.* Hanover, NH, September, 1973. Univ. Press of New England, Hanover, NH, pp. 121–133.

GROUNDWATER CONTAMINATION FROM MUNICIPAL LANDFILLS IN THE USA

David J. Tonjes
Cashin Associates PC
Hauppauge, New York

Groundwater contamination has been a significant environmental impact associated with landfills. Landfilling can affect groundwater quality through leachate releases, landfill gas migration, and runoff. Leachate releases are the most common and serious cause of impacts. Plume chemistry often limits extensive contamination to salts, ammonia, and redox-sensitive metals. Liners and gas control systems are effective in preventing pollution at modern facilities.

INTRODUCTION

Groundwater contamination is often cited as the greatest environmental problem associated with landfills. Municipal solid waste (MSW) landfills are often said to release contaminants such as metals, volatile organic compounds (VOCs) (such as solvents and gasoline constituents), pesticides and herbicides, polychlorinated biphenyls (PCBs), and dioxins and furans to groundwater, often at concentrations associated with human health concerns (1). Additional impacts would occur where contaminated groundwater discharges to surface waters.

Landfilling proponents acknowledge that earlier operational practices at disposal sites caused impacts to ground water (2). Most are now convinced that modern landfill engineering has ended these impacts (3).

Historically, design choices have affected the degree and kind of groundwater contamination. The most important is site location. In addition, the use of active controls can limit groundwater impacts.

Landfills impact groundwater in three ways. The primary concern is leachate releases. Leachate forms within the landfill as a result of precipitation. Precipitation percolates through the fill and dissolves matter from wastes and materials such as cover soils. These liquids foster chemical and biological reactions that liberate various substances. Landfill biota are further known to alter the composition of leachate (4). All of these elements of leachate may react with each other and with landfill features, including the containment liners and associated piping. Leachates undergo alteration in the environment, react with sediments, groundwater, or vadose zone liquids, and undergo additional biological reactions (5).

Contaminants also reach groundwater through gas migration (6). Waste decomposition gases produced within the fill may migrate away, typically through the vadose zone. There, gas may dissolve in recharge to reach groundwater; gas may diffuse at the phreatic zone directly; or, if the pressures from the gas are large, it may bubble into shallow groundwater (enhancing contaminant dissolution). The principal constituents of landfill gas are methane and carbon dioxide, and they are often approximately 50% each by volume. However, chemicals that have high vapor pressures and reactive gases generated by waste reactions also are commonly found in landfill gas (7). VOCs are thus typical gas-generated groundwater contaminants; as metals are not typically found in landfill gas, landfill gas rarely causes any metal contamination of groundwater.

Runoff that was in contact with wastes is another pathway to contaminate ground water (8). Runoff differs from leachate in that liquids do not pass through the waste but have contact with surface or shallow subsurface wastes. The chemistry and biology affecting runoff are different from those generating leachate. The surfaces of landfills tend to be aerobic, whereas fills are anaerobic. Photodegradation is a surface phenomenon, and volatilization is enhanced by wind (albeit interior volatilization is promoted by elevated temperatures). Runoff flowing away from a landfill can percolate through the vadose zone to ground water and may be altered along this pathway.

HISTORICAL FACTORS REGARDING GROUNDWATER CONTAMINATION FROM LANDFILLS

Land disposal has been the most common means of waste management, other than random discards. As municipalities organized waste disposal, locations for discards burial were identified. Lands of little value (such as swamps) or areas needing improvement by fills (quarries, depressions, or shorelines) were used (9).

Early landfills may not have impacted groundwater to the extent that mid twentieth century landfills did. It is clear that cities cause deterioration of underlying groundwater resources through waste disposal. Much of Manhattan's groundwater was unpotable by 1799, for example (10). However, the primary pollutant there and for large third-world cities today appears to be septic wastes (9).

There are few reports of landfill groundwater impacts before the 1960s (11). Many early studies report plume sizes consonant with contamination initiation that coincided with the start of sanitary landfilling (12,13). However, the development of sanitary landfilling occurred when groundwater research became more widespread and as waste disposal rates greatly increased. Sanitary landfills replaced open dumps. Most dumps had small footprints and accepted low volumes of wastes; therefore, relatively small releases of contaminants would be expected. Many dumps were at groundwater discharge zones (shorelines, wetlands, depressions), minimizing impacts to regional resources. In addition, many dumps were managed by burning garbage (a source of noxious air pollution), which reduced labile carbon and formed ashes (14), which may have reduced biological and chemical activity and supported reactions with the remaining wastes or sediment matrices to minimize solubilities.

GEOGRAPHICAL FACTORS REGARDING GROUNDWATER CONTAMINATION FROM LANDFILLS

Three geographical parameters help determine if landfill groundwater contamination may occur and its seriousness. First, there must be groundwater present. Not all of the United States is underlain by aquifers. The absence of significant groundwater resources means there is no resource to impact. Secondly, the resources must be hydraulically connected to the surface. If there is an impermeable barrier (either natural or artificial) between the fill and resource or the distance to the aquifer is too far for the quantity of generated recharge to reach it, then the potential for contamination is very small (4).

Thus, local geology is paramount. In areas of unconsolidated sediments of high permeability, landfill-derived contaminants can reach and impact groundwater but may not in other geological settings (13).

Finally, landfill contamination is driven by precipitation. In desert areas, less leachate, landfill gas, and runoff are generated, compared with landfills in more humid, temperate zones. Therefore, the risk of groundwater impacts from landfills is high on the eastern coastal plain. Rainfall is plentiful, and local geology leads to high-yielding aquifers found near the surface. In the southwest, greater aridity, bedrock outcrops, and large distances to water tables tend to minimize concerns (15).

LEACHATE RELEASES

An early evaluation of landfill leachate contamination of ground water by the U.S. Geological Survey (USGS) occurred on Long Island. Plumes from two, unlined landfills were characterized as extremely salty and enriched in iron and manganese. Other trace metals (except for zinc) were not detectable. The plumes had very low dispersion constants (especially laterally) and tended to sink in the aquifer (13).

The USGS characterized landfill plume chemistry in Delaware, where the plume was delineated by redox zonation. Closest to the landfill was a methanogenic zone. Further from the fill, successive zones marked by iron reduction, manganese reduction, and nitrate reduction (for carbon oxidation by microbes) were observed—the absence of sulfate eliminated sulfate reduction. The work promoted predictions of differential chemical processes in the zones, resulting in differences in plume chemistry (16).

In 1983, a University of Waterloo landfill study was published (19) that established many basic methodologies for acquiring data and generating interpretations of leachate contamination. The discussion of plume geochemistry (18) reinforced the USGS findings and identified cation exchange and other sediment–plume interactions as major controls.

Early modeling efforts found that the leading edge of the plume moved as quickly as groundwater flowed (19), mostly because it was defined by unreactive chloride. Other plume constituents were slowed by reactions, as discussed by Nicholson et al. (18), but, generally, chlorides have defined leachate plumes.

Landfill leachate characterizations are marked by varying concentrations (Table 1), which is true across sites, and even for cells at one landfill. Two leachate samples from different cells of a Long Island landfill had chloride concentrations of 50 mg/L and 58,000 mg/L (8). These variations are generally explicable in terms of different wastes in the cells, but site-specific and time-specific leachate generation processes also cause differences. These variations make it difficult to predict impacts *a priori*, or to determine volumes of releases *post hoc*.

Commonly, it is assumed that the contaminating constituents directly represent disposed waste materials. At an extreme, this resulted in some reports stating that iron, chloride, calcium, and magnesium were dumped as discrete wastes because they were found in plumes at concentrations above drinking water standards (21). Organic compounds cause special problems in this regard. Complex molecules such as DDT or PCBs, if found in a contaminant plume, probably were disposed at the site. Benzene or toluene, however, could have been disposed of as particular solvents, released from petroleum products (22), or even generated as relatively refractory breakdown products as other organic compounds degraded. Once, chlorinated hydrocarbons found in the environment were all thought to be released as solvents. However, now it is known that particular microbes can degrade more complex chlorinated solvents to simpler compounds (23). Additionally, under certain conditions, chlorinated hydrocarbons can be generated *de novo* by environmental bacteria (25).

Table 1. Classic Definitions of Landfill Leachate Quality[a]

Parameter	Range of 20 Samples[b]	Representative Range[c]
Alkalinity	0–20,850	500–10,000
Ammonia	0–1,106	10–1,000
Biological oxygen demand	81–33,360	
Calcium	60–7,200	100–3,000
Chemical oxygen demand	40–89,250	1,000–90,000
Chloride	4.7–2,500	300–3,000
Copper	0–9.9	<10
Iron	0–2,820	1–1000
Lead	<0.1–2	<5
Magnesium	17–15,600	100–1,500
Manganese	0.06–125	0.01–100
Mercury		<0.2
Nickel		0.01–1
Nitrate		0.1–10
Organic nitrogen		10–1,000
Phosphorus as phosphate	0–130	1–100
Potassium	28–3,770	200–1,000
Sodium	0–7,700	200–1,200
Sulfate	1–1,558	10–1,000
Total dissolved solids	584–44,900	5,000–40,000
Total hardness	0–22,800	
Total organic carbon		200–30,000
Total suspended solids	10–26,500	
Zinc	0–370	0.1–100
pH	3.7–8.5	4–8

[a] (in mg/L).
[b] USEPA, cited in Ref. 14.
[c] Compiled by Ref. 20.

Efforts to avoid groundwater contamination from landfills often begin by removing hazardous substances from the wastes (1). If landfills actually generate these kinds of compounds, these waste quality improvement programs might not achieve their goal.

A contaminant plume itself can control some pollutants. Research found that iron sulfide complex generation scavenged many free metal molecules, removing the immediate threat of trace metals migrating with the plume (25). In general, microbially mediated reactions together with groundwater–sediment interactions alter a plume substantially, biodegrade and chemically neutralize organic compounds in the near vicinity of many landfills, and limit their mobility in groundwater (26). Thus, leachate plumes at distances from the landfill often are no more than the more unreactive soluble salts (e.g., chloride, sodium, calcium, but not potassium), ammonia (if not nitrified), refractory carbon compounds, and dissolved iron.

LANDFILL GAS MIGRATION

Prior to sanitary landfilling, gas does not seem to have usually been generated in large quantities at most landfills. By the 1960s, landfill gas was a major problem. The most obvious dangers were explosions, especially at structures near landfills, which caused a number of deaths and injuries (27). Health concerns from exposure to the minor constituents of landfill gas also became a concern (28).

By the 1990s, it was becoming more widely known that landfill gas could directly cause groundwater contamination. VOC contamination from gas migration was the major concern (29). Some also noted that, analogous to natural gas impacts (30), redox-sensitive metals (iron and manganese, primarily) could be released from native sediments due to increased microbial activity and associated anoxic groundwater conditions. Methane can even increase alkalinity and therefore bicarbonate concentrations by increasing aqueous carbon (31).

Gas affects groundwater in three ways. It reacts with recharge as the gas migrates through the unsaturated zone away from the landfill (6). Second, as the gas fills the vadose zone, dispersion and surface film reactions can occur at the phreatic surface (32). Finally, if the landfill is producing large amounts of gas, it may increase the pressure gradient enough to infuse gas downward into the shallowest groundwater (33) (this is probably limited to directly beneath or adjacent to the fill), which can introduce gas below the immediate surface layer of the groundwater and allow more complete dissolution of contaminants.

Methane migration thus can cause VOCs and redox-sensitive metals contamination side-gradient and upgradient to overall groundwater flows and may enrich concentrations of these contaminants downgradient in the shallow aquifer. Methane migration cannot cause increases in soluble salts or nonnative metals, as gas cannot transport them. Methane impacts also tend to be restricted to the shallowest groundwater. Thus, impacts from landfill gas can be distinguished from leachate impacts (34).

UNCONTROLLED RUNOFF IMPACTS

There has been little research on runoff impacts from landfills. Modern landfill design intends to minimize and control runoff (35). Also, impacts from these surface flows off the landfill tend to occur very close to the fill, and so may be classed as leachate release impacts. However, where recharge basins are used to collect runoff from the landfill, concentrated impacts away from leachate plumes may occur.

Runoff may encounter only aerobic conditions on the fill, and its chemistry may be determined mostly by dissolution reactions (as is not necessarily so for leachate). Runoff may include high concentrations of particular salts if distinctive materials are found on the landfill surface. One Long Island landfill had unusually high concentrations of calcium and sulfate peripheral to its leachate plume. The landfill used ground construction and demolition debris, rich in gypsum ($CaSO_4$) from wallboard, as daily cover. In addition, organic pollutants (from solvents, for example) and easily dissolved metals may also become entrained in runoff (8).

POLLUTION PREVENTION

Many regulations governing landfills intend to prevent leachate releases. Much of the remaining regulatory effort prescribes extensive groundwater monitoring. Artificial liners evolved from simple sheets of plastics or clay beds pioneered in the 1970s to complex, engineered systems described by voluminous State and Federal regulations, and are required whenever groundwater impacts are possible (2). Early, simple liners had near 100% failure rates, and so now liners made of two separate materials and redundant systems are required in many jurisdictions. Regulations also limit hydraulic head buildup as leachate collects on the liner and require leachate removal off the liner and out of the landfill entirely. New York State, which requires redundant liner systems, has records showing that these modern liner systems can universally prevent all but environmentally unimportant releases to the environment (3). There is also some research that less complex systems, if managed properly, can be effective in controlling leachate releases (36).

Landfills are subject to complex closure procedures when operations cease. The intent is to create a “dry tomb” through continued removal of leachate and installation of an artificial cap on the wastes, so that precipitation cannot percolate in. Thus, the waste mass will be isolated from the environment—swaddled in plastics and/or clays—and leachate releases will be impossible (2). The system is also subject to continued maintenance and monitoring for 30 years, according to current regulations.

Controls on landfill gas migration, primarily through gas collection with combustion, are becoming more prevalent. These systems are usually installed for odor control, greenhouse gas reductions, and explosive gas control, not groundwater quality protection (37). However, the result is to limit groundwater impacts from gas migration directly and also to reduce VOC content in leachate, lessening impacts from leachate releases.

Runoff control for groundwater protection is also not a major concern. Runoff is controlled to reduce erosion and to minimize leachate (38). However, these ends may combine to direct liquids that have long waste contact times to management facilities outside of the leachate system. Most regulations, however, state that all water that comes in contact with wastes, no matter how long or brief the time, should be treated as contaminated leachate.

Tens of thousands of landfills have accepted wastes since 1970. Most have closed, and only approximately 2150 sites were operating in 2001 (39). However, probably all but several hundred of these landfills were at least partially constructed prior to 1993, and so are not engineered to the high standard required under the Federal Resource Conservation and Recovery Act Subtitle D (40 CFR Part 258) reauthorization. If the landfill is located in an area where impacts to groundwater can occur, a rigorous analysis of water quality near Subtitle D noncompliant sites is likely to find some impacts to groundwater (although the pollution levels may not necessarily exceed water quality standards), which means that old generation landfills comprise one of the major sources of groundwater contamination in the United States.

GROUNDWATER REMEDIATION

Landfill plumes are difficult to clean up (40). They are not single constituent contaminations. Techniques appropriate for certain plume elements may be hindered by other plume constituents. Air stripping towers that work well with solvents are easily plugged by dissolved iron (which precipitates in contact with the atmosphere), for example (41). Ion exchangers do not collect nonpolar compounds such as most chlorinated organics. Many landfill remedial projects, as practical implementations, address only particular plume constituents rather than the entire range of contaminants affecting the groundwater. Source control (such as capping the landfill) remains the most common remediation for a landfill that is causing groundwater contamination.

BIOREACTORS AND THE POTENTIAL FOR GROUNDWATER IMPACTS

A popular, new concept in landfill design is the "bioreactor." It is a lined landfill where the degradation processes are enhanced through the controlled addition of liquids and/or the recirculation of leachates (42). An aim is to degrade labile carbon as quickly as possible, increasing settlement (making more air space available for disposal purposes), and causing biological stabilization of the waste mass. Recirculating leachate may also strip contaminants onto the wastes, decreasing treatment costs for the effluents (43). Bioreactor technology requires increasing the amount of liquids in the wastes and intentionally increasing gas generation (for capture and use for energy generation). Thus, the design seems to increase some of the factors that lead to landfill groundwater contamination. However, the approach relies on the documented success of modern liners to prevent leachate releases as well as active gas collection efforts to minimize environmental impacts.

SUMMARY

Landfills have historically caused extensive groundwater impacts in the United States, through leachate releases, methane migration, and uncontrolled runoff. These problems have been greatest in moist climates where aquifers have high conductivity potentials and are close to the ground surface.

Groundwater contaminated by a landfill is usually discolored and warm and smells bad. It contains high levels of salts, along with dissolved iron and often manganese. VOCs are a general concern in plumes but are not always found at significant concentrations or throughout the contamination. Many landfill plumes have much lower than expected dissolved trace metal levels because of scavenging by iron sulfides and hydroxides in low-redox, anoxic waters.

Modern engineering techniques have largely eliminated the major sources of contaminants. Modern artificial liner systems together with gas collection prevent landfill leachate and gas from carrying contaminants to groundwater. Dry tomb landfill closures isolate fills from the environment after waste disposal ceases.

However, the barriers and processes used to control access of water to wastes cannot be expected to function indefinitely. It is not clear exactly what the pollution potential of dry tomb landfills will be when the liners and caps decay and fail and the wastes are again exposed to the environment, which is an additional impetus for bioreactor research, as the material remaining in a stabilized landfill, it has been hypothesized, would be environmentally benign.

BIBLIOGRAPHY

1. Denison, R.A. (1996). Environmental life-cycle comparisons of recycling, landfilling, and incineration: a review of recent studies. *Annual Review of Energy and the Environment* **21**: 191–237.
2. Koerner, R.M. (1997). Design of Subtitle D and Subtitle C landfill containment systems. pp. 247–264. In: *Municipal Solid Wastes: Problems and Solutions*. R.E. Landruth and P.A. Rebers (Eds.). CRC Press, Lewis Publishing, New York.
3. Phaneuf, R.J. and Becker, D.A. (2001). Proper environmental containment system design and performance monitoring: can it lead to increased regulatory flexibility? *Presented at Federation of New York Solid Waste Associations Solid Waste and Recycling Conference and Trade Show*, Bolton Landing, NY, May 8.
4. Qasim, S.R. and Chiang, W. (1994). *Sanitary Landfill Leachate*. Technomic Publishing Co., Lancaster, PA.
5. Cozzarelli, I.M. et al. (2000). Geochemical and microbiological methods for evaluating anaerobic processes in an aquifer contaminated by landfill leachate. *Environmental Science and Technology* **34**(18): 4025–4033.
6. Clister, W., Brown, K., and Hibbs, S. (2001). Improvement of groundwater quality through management of landfill gas at MSW landfills. *MSW Management* **11**(3): 34–39.
7. Soltani-Ahmadi, H. (2002). A review of landfill gas components. *MSW Management* **12**(6): 120–123. (Expanded at *www.forester.net/mw_0209_nmoc.html*).
8. Tonjes, D.J. (1998). *Modified Normed Stiff Diagrams: A New Means of Interpreting Groundwater Sampling Results at Municipal Solid Landfills*. Dissertation, SUNY at Stony Brook.
9. Melosi, M.V. (2000). *The Sanitary City*. The Johns Hopkins University Press, Baltimore, MD.
10. Galusha, D. (1999). *Liquid Assets: A History of New York City's Water System*. Purple Mountain Press, Fleischmanns, NY.
11. Zanoni, A.E. (1972). Ground-water pollution and sanitary landfills: a critical review. *Ground Water* **10**(1): 3–13.
12. Apgar, M.A. and Langmuir, D. (1971). Ground-water pollution potential of a landfill above the water table. *Ground Water* **9**(6): 76–94.
13. Kimmel, G.E. and Braids, O.C. (1974). Leachate plumes in a highly permeable aquifer. *Ground Water* **12**(6): 388–392.
14. Office of Technology Assessment. (1988). *Facing America's Trash: What Next for Municipal Solid Waste?* United States Congress, Washington, DC.

15. Office of Technology Assessment. (1984). *Protecting the Nation's Groundwater from Contamination*. United States Congress, Washington, DC, Vol. 2.
16. Baedecker, M.J. and Back, W. (1979). Modern marine sediments as a natural analog to the chemically stressed environment of a landfill. *Journal of Hydrology* **43**: 393–414.
17. Cherry, J.A. (1983). Introduction. *Journal of Hydrology* **63**: vii–ix.
18. Nicholson, R.V., Cherry, J.A., and Reardon, E.J. (1983). Migration of contaminants in groundwater at a landfill: a case study, 6. Hydrogeochemistry. *Journal of Hydrology* **63**: 131–176.
19. Ostendorf, D.W., Noss, R.R., and Lederer, D.O. (1984). Landfill leachate migration through shallow unconfined aquifers. *Water Resources Research* **20**(2): 291–296.
20. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Inc., Englewood Cliffs, NJ.
21. NYSDEC. (1992). *Inactive Hazardous Wastes Disposal Sites in New York State, Annual Report: Appendix 1: List of Sites by County: Nassau, Suffolk*. New York State Department of Environmental Conservation, Albany, NY.
22. Lesage, S., Jackson, R.E., Priddle, M.W., and Riemann, P.G. (1990). Occurrence and fate of organic solvent residues in anoxic groundwater at the Gloucester Landfill, Canada. *Environmental Science and Technology* **24**(4): 559–566.
23. Leahy, J.G. and Colwell, R.R. (1990). Microbial degradation of hydrocarbons in the environment. *Microbiological Reviews* **54**(3): 305–315.
24. Myneni, S.C.B. (2002). Formation of stable chlorinated hydrocarbons in weathering plant materials. *Science* **295**: 1039–1041.
25. Kehew, A.E. and Passero, R.N. (1990). pH and redox buffering mechanisms in a glacial drift aquifer contaminated by landfill leachate. *Ground Water* **28**(5): 728–737.
26. Eganhouse, R.P., Cozzarelli, I.M., Scholl, M.A., and Matthews, L.A. (2001). Natural attenuation of volatile organic compounds (VOCs) in the leachate plume of a municipal landfill: using alkylbenzenes as process probes. *Ground Water* **39**(2): 192–202.
27. Hodgson, A.T., Garbesi, K., Sextro, R.G., and Daisey, J.M. (1992). Soil-gas contamination and entry of volatile organic compounds into a house near a landfill. *Journal of Air and Waste Management Association* **42**(3): 277–283.
28. NYSDOH. (1998). *Investigations of Cancer Incidence and Residence near 38 Landfills with Soil Gas Migration Conditions: New York State, 1980-1989*. Final Report to ATSDR. New York State Department of Health, Albany, NY.
29. Kerfoot, H.B. (1994). Landfill gas effects of groundwater samples at a municipal solid waste facility. *Journal of Air and Waste Management Association* **44**: 1293–1298.
30. Kelly, W.R., Matisoff, G., and Fisher, J.B. (1985). The effects of a gas well blow out on groundwater chemistry. *Environmental Geology Water Sciences* **7**(4): 205–213.
31. Fahrquahr, G.J. and Rovers, F.A. (1973). Gas production during refuse decomposition. *Water, Air, and Soil Pollution* **2**: 483–495.
32. O'Leary, P. and Walsh, P. (2002a). Landfill gas movement, control, and energy recovery. *Waste Age* **33**(3): 48–54.
33. Groseclose, M. et al. (2001). Hydraulic containment of landfill gas driven VOCs in groundwater using a state of the art computerized control system at the Los Reales Landfill. In: *Proceedings of the Fourth Annual Landfill Symposium*, Solid Waste Association of North America, Denver, CO, pp. 37–43.
34. Tonjes, D.J., Heil, J.H., and Black, J.A. (1995). Sliding Stiff diagrams: a sophisticated ground water analytical tool. *Ground Water Monitoring and Remediation* **15**(1): 134–139.
35. O'Leary, P. and Walsh, P. (2002b). Landfill cover and liner systems for water quality protection. *Waste Age* **33**(4): 124–129.
36. Othman, M.A., Bonaparte, R., and Gross, B.A. (1997). Preliminary results of study of composite liner field performance. In: *Proceedings of the GRI-10 Conference on Field Performance of Geosynthetics and Geosynthetic Related Systems*. R.M. Koerner, G.R. Koerner, and Y.G. Hsuan (Eds.). GSI Publications, Philadelphia, PA, pp. 115–142.
37. Guzzone, B. (2001). Landfill gas utilization: the future's so bright. *MSW Management* **11**(4): 46–51.
38. Reeves, T.S. (1999). Run-off/run-on management: observations of storm water management at landfills. In: *Proceedings of the Fourth Annual Landfill Symposium*, Solid Waste Association of North America, Denver, CO, pp. 145–154.
39. Goldstein, N. and Madtes, C. (2001). The state of garbage in America. *BioCycle* **42**(12): 42–54.
40. Christensen, T.H., Bjerg, P.L., and Kjeldsen., P. (2000). Natural attenuation: a feasible approach to remediation of ground water pollution at landfills?. *Ground Water Monitoring and Remediation* **Winter 2000**: 69–77.
41. Mackay, D.M. and Cherry, J.A. (1989). Groundwater contamination: pump-and-treat remediation. *Environmental Science and Technology* **23**(6): 630–636.
42. Campman, C. and Yates, A. (2002). Bioreactor landfills: an idea whose time has come. *MSW Management* **12**(6): 70–81.
43. Hull, P. (2000). Changing the schedule. *MSW Management* **10**(4): 40–45.

METAL ORGANIC INTERACTIONS IN SUBTITLE D LANDFILL LEACHATES AND ASSOCIATED GROUND WATERS

Mário Abel Gonçalves
Faculdade de Ciências da
Universidade de Lisoba
Lisoba, Portugal

INTRODUCTION

Organic compounds are one of the most important constituents of landfill leachate waters. Their nature and relative amount are dependent on landfill age and evolutionary stage (1). In the early stages of landfills, organic compounds present in the leachate are mostly species with molecular weight (MW) less than 1000 Da, although with a nonnegligible fraction of higher MW species. The leachate pH is typically around 5 in this stage because of the abundance of low molecular weight organic acids. This acidic stage can last for a few years or up to ten years depending on landfills. Gradually, landfill leachates evolve toward the methanogenic stage, attaining pH values around 8, where higher molecular weight organic acids develop, including fulvic- and humic-type compounds, together with the precipitation of sparingly soluble metal phases, which reduces their mobility (2). Organic compounds, with their abundant functional groups, are recognized important metal ligands in this system. Among the different types of compounds, carboxylic acids and phenols are always the most abundant ones (3).

METAL—ORGANIC COMPLEX STABILITY

Metals do not constitute an important component of landfill waters. However, because they are persistent pollutants in the environment, knowing their fate is of utmost importance. Toxicity of metals to living creatures is a well-known issue. Despite such risk, not every chemical form is equally hazardous, one such example is Cr, being an essential component to glucose metabolism, it is also extremely poisonous in the chromate form (Cr(VI)). On the contrary, Cr(III) is not particularly toxic and forms insoluble solids, which reduces its bioavailability. Depending on the speciation and complex formation, metal cations may have different migration velocities in groundwater, which may increase with metal-organic ligand association, and eventually reduce their bioavailability.

Metal-organic interaction in landfill-contaminated waters depends on the stability of complex formation and metal concentration in solution. Classification of metals into hard-sphere (type A) or soft-sphere (type B) cations (4) is helpful in grouping them according to type of complex formation and stability. Hard-sphere cations have noble gas configurations with spherical symmetric orbitals, because of the low polarizability of the electron cloud. These cations form ionic bonds mostly with ligands that have oxygen as a donor atom. Alkali and earth-alkali metals and Al^{3+} are examples in this group. Soft-sphere cations have deformable electron clouds because of their greater polarizability. These metals predominantly form covalent bonds with ligands having S or N as donor atoms. Metal cations in this group include Cu^{+}, Zn^{2+}, Cd^{2+}, Hg^{2+}, Pb^{2+}, Sn^{2+}, and Tl^{3+}. Besides, transition metal cations, such as V^{2+}, Cr^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, V^{3+}, Cr^{3+}, Mn^{3+}, Fe^{3+}, and Co^{3+}, also have similar properties (4).

For transition metal cations, a reasonably well-established rule for a sequence of complex stability indicates that the stability increases in the series (known as the Irving–Williams order) $Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$ (4). For hard-sphere cations, complex stability is proportional to the charge/radius ratio of the cation.

Complexes with monodentate ligands are usually less stable than those with multidentate complexes (chelates). Normally, monodentate complexes are easily dissociated at dilute concentrations of the metal cation, which is well illustrated in experiments with monodentate, bidentate, and tertadentate Cu(II) amine complexes. At dilute metal concentrations ($<10^{-4}$ M), where the monodentate complex is fully dissociated, the tertadentate complex remains stable in solution. Thus, even at concentration levels that are likely to be found in leachate-contaminated waters, the complexing effect of monodentate ligands is expected to be negligible. This experimental evidence may justify the verification that the highly abundant monocarboxylic acids on leachate-derived waters had little influence on the adsorption of metals such as Cu and Zn onto solid phases (5,6). Recently, Schilling and Cooper (7) identified that carboxylic groups are the most important binding sites for Cu(II) cations in soil organic matter, and that phenol groups are less important. Increase on metal sorption capacity is also observed with the addition of carboxyl and amino groups to various types of biomass (8).

METAL-ORGANIC INTERACTION IN LANDFILL-POLLUTED ENVIRONMENTS

In spite of such general knowledge about metal-organic complexation stability, it must be stressed that a great deal of experimental studies do not reflect the environment of a landfill-polluted water. Thus, the expected behavior of metals may differ substantially from what may be previously known, and several factors may induce such a view: first, most experimental work deals with manipulated and treated organic compounds; second, very few reported investigations actually dealt with metal-organic complexation in landfill lechate or leachate-polluted groundwater systems; and third, the influence that a high ionic strength environment and the presence of multiple competing ligands has on the ability to form metal-organic complexes is not known. Specific studies on metal-organic complexation associated to landfill leachate waters or leachate-polluted groundwaters are restricted to a few metal cations, most notably Pb, Cd, Ni, Cu, Cr, and Zn (e.g., 9–12).

Despite how scarce and, ultimately, complex these studies may be, it has been possible to gather important knowledge on metal-organic compound interaction. Speciation studies on a landfill leachate based on an exchange procedure with Chelex 100 resin demonstrates that the free or labile complexed fraction is the most important for all the metals studied by Majone at al. (12), increasing from 50% to 80% in the order Cd < Pb < Ni < Cu (≈80%). The slowly enchangeable fraction represents 14% for Cu, 28% for Pb, and 47% for Cd. Ni has the higher relative percentage in the inert or stable complexed fraction (17%). In the case of Cu, this metal becomes complexed either with low molecular weight compounds (<1000 Da, around 47%) or with high molecular weight compounds (>12000 Da, around 45%), these latter corresponding probably to humic acids (12). Overall, these authors concluded that the high MW organic acids have an important influence on metal speciation, and bonding to these substances increases in the rank Ni < Cd < Cu. The work of Jensen and Christensen (11) and Jensen et al. (10) showed that a significant but highly variable fraction of heavy metals have affinity for dissolved or colloidal organic matter. It must also be stressed that the distribution of metals among the different groups considered (dissolved and colloidal forms, organic- or inorganic-associated) varied between samples and between metals within the same sample (10), which demonstrates the heterogeneity of these systems and difficulty in achieving general conclusions. However, in leachates, it has been possible to show that heavy metals, such as Cu, Pb, and Cd, are predominantly associated with organic matter, either in colloidal form or dissolved. A substantial fraction of dissolved metals, such as Zn and Ni, is mostly associated with dissolved inorganic complexes, the remaining are associated with colloidal fractions (either organic or inorganic). The importance that colloidal organic matter has on the speciation and binding of Cu and Pb (10) seems to be the most plausible explanation for the differential partitioning of metals to sediments near the source of spilled landfill leachate to the surface environment (6). Besides, it has also been verified that metals such as Zn and Ni had much less affinity to organic compounds, preferentially forming insoluble sulfide phases, as in the case of Zn. Christensen et al. (9) also verified for Zn and Ni in leachate-contaminated groundwater that at the low concentrations usually observed in the natural system, complex stability decreases significantly.

BIBLIOGRAPHY

1. Christensen, T.H. et al. (2001). Biogeochemistry of landfill leachate plumes. *Appl. Geochem.* **16**: 659–718.
2. Pohland, F.G. and Kim, J.C. (2000). Microbially mediated attenuation potential of landfill bioreactor systems. *Water Sci. Technol.* **41**: 247–254.
3. Albaiges, J., Casado, F., and Ventura, F. (1986). Organic indicators of groundwater pollution by a sanitary landfill. *Water Res.* **20**: 1153–1159.
4. Stumm, W. and Morgan, J.J. (1996). *Aquatic Chemistry. Chemical Equilibria and Rates in Natural Waters*, 3rd Edn. John Wiley & Sons, New York.
5. Schroth, B.K. and Sposito, G. (1998). Effect of landfill leachate organic acids on trace metal adsorption by kaolinite. *Environment. Sci. Technol.* **32**: 1404–1408.
6. Gonçalves, M.A., Nogueira, J.M.F., Figueiras, J., Putnis, C.V., and Almeida, C. (2004). Base-metals and organic content in stream sediments in the vicinity of a landfill. *Appl. Geochem.* **19**: 137–151.
7. Schilling, M. and Cooper, W.T. (2004). Identification of copper binding sites in soil organic matter through chemical modifications and ^{13}C CP-MAS NMR Spectroscopy. *Environment. Sci. Technol.* **38**: 5059–5063.
8. Gardea-Torresdey, J.L., Tang, L., and Salvador, J.M. (1996). Copper adsorption by esterified and unesterified fractions of Sphagnum peat moss and its different humic substances. *J. Hazard. Mater.* **48**: 191–206.
9. Christensen, J.B., Jensen, D.L., and Christensen, T.H. (1996). Effect of dissolved organic carbon on the mobility of cadmium, nickel, and zinc in leachate polluted groundwater. *Water Res.* **30**: 3037–3049.
10. Jensen, D.L., Ledin, A., and Christensen, T.H. (1999). Speciation of heavy metals in landfill-leachate polluted groundwater. *Water Res.* **33**: 2642–2650.
11. Jensen, D.L. and Christensen, T.H. (1999). Colloidal and dissolved metals in leachates from four Danish landfills. *Water Res.* **33**: 2139–2147.
12. Majone, M., Papini, M.P., and Rolle, E. (1998). Influence of metal speciation in landfill leachates on kaolinite sorption. *Water Res.* **32**: 882–890.

LEACHING

David W. Kelley
University of St. Thomas
St. Paul, Minnesota

Leaching, sometimes referred to as eluviation, is the phenomenon whereby a liquid, usually water, moves through a matrix from one location to another, dissolving or suspending materials along the way. The matrix must be porous enough to allow fluid movement and is generally soil, but fractured bedrock or unconsolidated mineral or organic (humus) material can also leach liquids. Normally, this occurs under gravity and movement is downward. Occasionally, this movement is lateral when forces such as capillary action, soil matric potential, or confinement pressure draw the liquid to an area of high water tension or push the liquid from an area of high pressure. The liquid that migrates, generally referred to as leachate, contains a mix of dissolved and/or suspended ions or compounds. These materials, either present naturally in the environment or applied to the ground surface, include dissolved salts, fertilizers, soluble minerals, organic material such as humic or fulvic acids, plant nutrients, and natural or synthetic chemicals. Because water is a polar liquid, it can dissolve or transport a wide variety of substances from one location to another. Even nonpolar molecules can form micelles and be moved great distances by leaching.

Leaching is a function of several factors that affect the amount, rate, direction, and quality of the leachate: the hydraulic conductivity of the matrix, impeding layers (aquitards or aquicludes), partitioning coefficients of any chemicals present, antecedent soil water content (saturated vs. unsaturated flow), osmotic or matric potentials, the cation exchange capacity of the soil, the ionic makeup of the dissolved compounds in the percolating fluid, the water-holding capacity of the soil (largely attributable to soil texture), and the macropore flow (a function of soil porosity, structure, disturbance, burrowing animals, and root channels).

If water is applied at a rate slower than the ability of the soil surface to accept it (infiltration rate), water percolates downward once the water holding capacity of soil is exceeded. If water is applied at a faster rate, a portion of the applied water runs off the soil surface (potential erosion). See Table 1.

Most nonindustrial leaching occurs on agricultural land or under antiquated or modern landfills. Leaching rates are highest in humid regions or on land under irrigation and lowest in arid or permanently frozen regions. Leaching is governed by a few basic processes.

Table 1. Water-holding Capacity of Various Soils (Inches of Water Per Inch of Soil Depth)

Textural Class	Capacity
Sand	0.070
Loamy sand	0.065
Sandy loam	0.095
Sandy clay loam	0.150
Loam	0.165
Clay loam	0.185
Silt loam	0.20
Silt	0.23
Silty clay loam	0.25
Silty clay	0.22
Clay	0.20

Mass Flow

This is the dominant leaching process in most systems and accounts for the greatest movement of water and materials, often over large distances. When water infiltrates a dry, permeable soil, a discrete wetting front is formed that has a higher water content (lower tension) above than below. The depth of this wetting front depends on the amount of water added, the permeability of the surface (pore size, volume, and connectivity), and its initial wetness. To a lesser extent, wetting front migration is a response to the rate at which the water is applied. After water is no longer applied, the wetting front moves but at a slower rate due to reduced pressure from above.

Water infiltrating a dry soil can dissolve and transport salts or chemicals that are not strongly adsorbed on soil solids. Water can also dislodge and transport materials in suspension such as organic colloids or fine clays that are not readily soluble. This dissolved and/or suspended load is transported to depth via the percolating water, but only as far as the wetting front progresses. It can be moved more deeply by subsequent rainfall or irrigation, provided it is of sufficient volume to move percolating water past the previously established wetting front.

If a substance is loosely adsorbed onto solid surfaces in the soil, the depth to which the adsorbed chemical moves is reduced or the rate of movement slowed. This retardation factor, R, is related to the soil-water partition coefficient, K_d, of the particular substance being moved (Fig. 1):

$$R = 1 + \frac{\rho K_\mathrm{d}}{\theta} \tag{1}$$

where ρ is the dry bulk density of the soil and θ is its volumetric water content. The retardation factor varies from a low of 1 for nonadsorbed chemicals to values in excess of 100 for some highly adsorbed ones. The depth to which a substance is leached largely depends on the adsorbent relationship between a substance and the matrix through which it percolates.

The wetting front in an initially dry soil corresponds to the leading edge of the infiltrating liquid. Surfaces that are not initially dry behave similarly except in some soils, where water entering a moist soil can displace water already present and push that water ahead of the front. In such cases, a dissolved or suspended substance would move only to the depth of the infiltrating water, not all the way to the wetting front. Alternatively, because macropore flow (preferential flow paths along cracks in structured soils or fractured rock) is usually present, some water at the wetting front may represent new water. Most likely the situation is a combination of these two mechanisms of water dispersal, which leads to uncertainty in the real location or concentration of a dissolved substance and also suggests that a dissolved chemical is not likely to be concentrated in a sharp pulse but will be more diluted or dispersed.

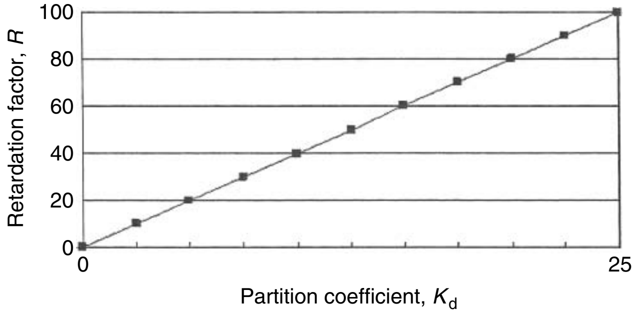

$$R = 1 + \frac{\rho K_\mathrm{d}}{\theta}$$

Figure 1.

Dispersion

Water moving through soil pores travels at different rates due to porosity, tortuosity, and connectivity differences. A pulse of chemical traveling with the water has its leading and trailing edges spread out, a phenomenon called dispersion. Under natural conditions in the field, water rarely moves through a matrix at a constant rate at all times, even if its physical properties (texture, structure, ionic exchange capacity, etc.) are extremely uniform and water is applied at a constant rate.

Diffusion

Dissolved or suspended substances also move from regions of higher to lower concentration due to osmotic potential referred to as diffusion. Diffusion is usually much less significant than mass flow; it merely spreads the chemical out in an already wetted area. Its influence is generally small and its rate slow due to adsorption (on matrices high in clay and organic matter content).

Sorption

A chemical added to a soil partitions itself between a solution phase and an adsorbed solid phase. When the cation exchange capacity (CEC) of a soil is high (e.g., soils high in smectite clays or organic matter), this bonding is strong because cations are adsorbed to the negatively charged surfaces (Table 2). In other cases (soils with low-CEC clays or lacking clays and organic matter), bonding is quite weak.

Sorption or binding of dissolved chemicals on surfaces is significant because they are not free to move with percolating water. Although sorption retards the rate of

Table 2. Normal Range of CEC Values for Common Soil Groups

Soil Groups	CEC in cmol(+)/kg
Light colored sands	3–5
Dark colored sands	10–20
Light colored loams and silt loams	10–20
Dark colored loams and silt loams	15–25
Dark colored silty clay loams and silty clays	30–40
Organic soils	50–100

chemical movement, it does not necessarily alter the rate of movement of the percolating solution itself.

Substances in a liquid–soil environment tend to establish a balance between the amount on solid surfaces and the amount in solution. Some chemicals exist primarily in the liquid phase; others are strongly adsorbed and exist primarily on solid surfaces. Molecules tend to move from one phase to another to maintain this balance. The manner in which the molecules are partitioned into solid and liquid phases depends on both matrix and chemical properties. These relationships, called sorption isotherms, can be determined experimentally. The center of the chemical pulse on a particular day is greater for substances that have smaller partition coefficients than for those that have larger ones because substances that have smaller partition coefficients are adsorbed less strongly to solid surfaces and can be dislodged and moved more readily by the percolating solution.

Degradation

Chemicals in the environment can undergo biological or chemical transformations, be absorbed by plants, volatilize into the atmosphere, be lost to surface water by surface water flow and erosion, or remain in the surface matrix and be available for leaching to groundwater. The importance of these different processes depends on the unique properties of the substance.

Many natural and synthetic chemicals break down or degrade over time from microbiological and chemical reactions. Substances with low persistence degrade quickly to other products, and highly persistent ones linger. Persistence is usually specified in terms of the degradation half-life, a measure of the time it takes for one-half of the original amount of a substance to be degraded (first-order degradation). Degradation rates are generally temperature and moisture dependent, and most biological breakdown occurs in the root zone of plants because more microbes exist there than elsewhere. Therefore, as compounds are leached to lower depths (below the root zone), they become more persistent and can leach to groundwater before surface organisms have had a chance to degrade them. Typically, root zones can vary from 1 to 6 feet below the surface, depending on the plant and its stage of development.

The chemical makeup of leachate can be altered or its volume reduced through several processes. Roots of actively photosynthesizing plants can scour out essential nutrients or absorb percolating water, denitrification and mineralization can remove or immobilize dissolved nitrogen, chemical alterations such as oxidation to less motile forms can occur (e.g., iron transformed in an oxidizing environment from a mobile, reduced ferrous Fe^{2+} state to an immobile, oxidized ferric Fe^{3+} state), cations (e.g., Na^+, $NH_4{}^+$, Ca^{2+}, Mg^{2+}, etc.) can be adsorbed onto negatively charged surfaces, and shallow subsurface agricultural tile drains can siphon off leachate before it percolates more deeply (into groundwater).

The concern with leaching is that it accounts for the bulk of nonpoint source pollution of groundwater and subsequently to any surface water to which it is hydrologically connected. Any chemicals or suspended solids that are transported to groundwater may impair water quality for aquatic life or render it nonpotable or unfit for human contact (drinking, bathing, swimming). Lawn fertilizers, agricultural chemicals (pesticides, herbicides, fungicides, miticides, etc.), road salts, petrochemical spills, atmospheric particulate fallout (Hg, Pb, oxides of sulfur or nitrogen, etc.) can all end up in groundwater via leaching.

High-risk environments that are of primary concern include

- sandy soils
- shallow-rooted plants
- high rainfall areas or excess irrigation
- shallow unprotected groundwater
- wellhead protection zones
- homesteads with open wells

Leaching affects the receiving waterbody (e.g., a change in the trophic state of a lake receiving discharge from contaminated groundwater), and also in agricultural systems, represents a two-pronged economic loss: fertilizer purchases go wasted if nutrients end up in places other than the root zone, and plant nutrients made unavailable by leaching result in lower harvests. Essential plant nutrients, naturally occurring or supplemented by fertilizers, that are lost by leaching include

- *Calcium (Ca)*: ion generally in the largest concentration in leachate
- *Magnesium (Mg), Sulfur (S), Potassium (K)*: next highest concentrations, depending on soil composition and acidity of leachate
- *Nitrogen (N)*: ranges from high (recent fertilizer application on porous soil prior to intense rainfall) to low (natural systems with no disturbance or low precipitation), depending on plant growing conditions and natural soil fertility
- *Phosphorus (P)*: limited concentration in leachate due to low solubility and high affinity for sorption onto soil particles

The U.S. Environmental Protection Agency (EPA) regulates several pollutants that routinely make their way into drinking water supplies and pose health risks for consumers in the United States. For example, the EPA currently stipulates that drinking water cannot exceed 10 mg per liter (ppm) nitrate ($NO_3{}^-$); this poses significant problems for drinking water suppliers in rural areas that rely on aquifers under agricultural fields for their water source (Fig. 2). The EPA limits lead (Pb) in drinking water to 15 μg per liter (ppb), a challenge in older urban areas that still have lead service pipes as part of their underground water systems. Recently, the EPA issued guidelines for arsenic (Ar) in drinking water; the standard was lowered to a maximum of 10 ppb beginning January 23, 2006. In addition to these specific thresholds for drinking water, the EPA also regulates waterbodies (surface, groundwater, drinking water aquifers) for dissolved salts, fecal coliform

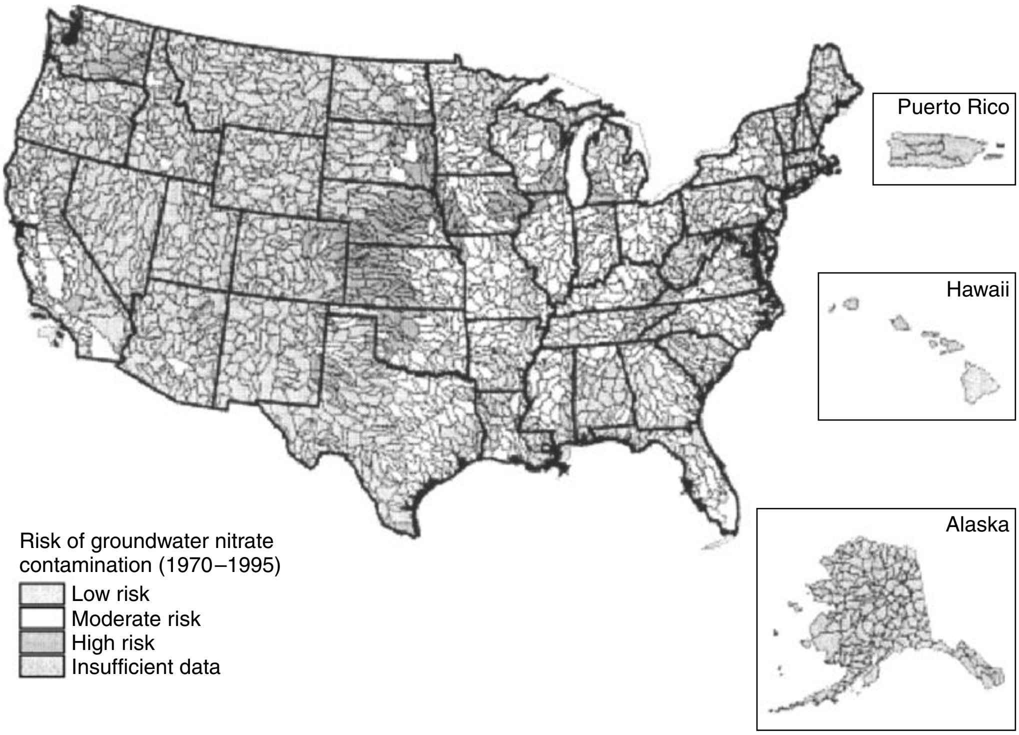

Figure 2. Risk of groundwater nitrate contamination.

bacteria, harmful chemicals, and more, some of which are attributable to leaching processes.

READING LIST

Brady, N.C. and Weil, R.R. (2002). *The Nature and Properties of Soils*, 13th Edn. Pearson Education, Upper Saddle River, NJ.

Hillel, D. (1998). *Environmental Soil Physics: Fundamentals, Applications, and Environmental Considerations*. Academic Press, New York.

Jury, W.A. and Horton, R. (2004). *Soil Physics*, 6th Edn. John Wiley & Sons, New York.

McBride, M.B. (1994). *Environmental Chemistry of Soils*. Oxford University Press, New York.

Nolan, B.T., Ruddy, B.C., Hitt, K.J., and Helsel, D.R. (1997). Risk of nitrate in groundwaters of the United States—a national perspective. *Environ. Sci. Technol.* **31**: 2229–2236.

Sawhney, B.L. and Brown, K. (Eds.). (1989). *Reactions and Movement of Organic Chemicals in Soils*. SSSA Spec. Publ. No. 22, Madison, WI.

WELL MAINTENANCE

Stuart A. Smith
Smith-Comeskey GroundWater
Science LLC
Upper Sandusky, Ohio

PREDICTIVE AND PREVENTIVE MAINTENANCE

Well maintenance consists of two overall problem-preventing scopes of work: predictive and preventive maintenance (together comprising PM). Predictive maintenance involves tasks such as monitoring water quality, power, and other indicator data to detect evidence of deterioration prior to a fault and making necessary repairs. Preventive maintenance includes tasks such as maintaining lubrication, replacing components, and preventive well treatments conducted prior to deterioration of performance. By contrast, reactive maintenance repairs deteriorated components and performance conditions.

PM GOALS

PM activities are intended to prevent or significantly slow performance deterioration and lengthen intervals between equipment replacements and service events. PM begins with design. In any engineered system, including water works and water wells, prevention of adverse water quality and performance problems begins with appropriate planning and design. These activities include

1. identifying potential problems
2. designing a system to minimize these problems
3. allowing for effective restoration.

Where industry-recommended design parameters (1,2) are followed and materials selected for resistance to deterioration due to water quality, deterioration can be slowed and prevented. These conditions are site-specific, and design is based on investigation and testing.

ANALYSES RELEVANT TO PM

PM of wells depends on analysis of water quality and well performance. The analyses performed necessarily depend on the resources available and also on the environmental difficulties posed by the well environment and the well's purpose and use. Table 1 lists relevant parameters for commercially significant wells (e.g., public and industrial water supply, irrigation).

Data collection may be greatly simplified for wells in very low-demand environments or for low-impact use (campgrounds, low-volume domestic wells). However, some data collection should be performed systematically to avoid "run to failure" typical of such wells.

RECORDS

Records are important in establishing trends that show patterns in performance, including indication of deterioration. Records should include the following:

1. Physical locations and as-built descriptions of the wells and their equipment. The geographic location of well should include reference to fixed landmarks and coordinates.
2. As-built diagram of the well's construction, with any modifications over time.
3. Any lithologic and well construction log information.
4. Records of pumping tests and geophysical, borehole flow meter, etc. tests of the completed well over time.
5. Dates of replacement of components, manufacturer and type of component, if known, and length of service, if known. Include itemized invoices with costs. Take photos or videotapes of deteriorated components for future reference and include descriptions. Include copies of product owner operation and service literature. Document any contractor service personnel.
6. Color borehole TV survey video records: Take at any zero point, before and after well rehabilitation measures, and at service intervals. Tapes and CDs may be consolidated as summary CDs of important well features over the years. Label records by well identification and date and store properly in an accessible location. Transfer to improved media as available.

The level of complexity and schedule can vary. At a minimum, for any valuable well under stress or one that is difficult to replace, routine hydrologic performance parameters, water quality, and biofouling indicators should be checked regularly (Table 2).

ADDITIONAL PM INSPECTIONS AND PRECAUTIONS

The capture zones and immediate surroundings of wells should be free of threats to the potable or monitored water quality. The area should be inspected periodically to ensure this condition. Well tops should be securely sealed to prevent contamination and secured against unauthorized entry. However, service access should not be impaired. In some wells, pumps may be pulled, inspected, cleaned, and repaired, as necessary, on a periodic or performance basis.

Table 1. Summary of Methods Relevant to Well Maintenance

Parameter	Indicator
Fe and Mn (total, Fe^{2+}/Fe^{3+}, Fe minerals, Mn^{4+}/Mn^{2+}, Mn minerals and complexes)—sometimes other metals such as Al	Indications of clogging potential, presence of biofouling, E_h shifts. Fe transformations are the most common among redox-sensitive metals in the environment. Mn is less common but locally important.
S (total, $S^{2-}/S^0/SO_4{}^{2-}$, S minerals, and complexes)	Indications of corrosion and clogging potential, presence of biofouling, E_h shifts.
E_h (redox potential):	Direct indication of probable metallic ion states, microbial activity. Usually bulk E_h, which is a composite of microenvironments.
pH	Indication of acidity/basicity and likelihood of corrosion and/or mineral encrustation. Combined with E_h to determine likely metallic mineral states present.
Conductivity	Indication of TDS content and a component of corrosivity assessment.
Major ions, hardness and alkalinity	Carbonate minerals, F, Ca, Mg, Si, Na, Cl determine the types of encrusting minerals that may be present and are used in saturation indexes. Alkalinity used in index and cleaning solution calculations.
Turbidity and particle counts	Indication of suspended particles content, suitable for assessment of relative changes indicating changes in particle pumping or biofouling.
Sand/silt content (v/v, w/v)	Indication of success of development/redevelopment, potential for abrasion and clogging.
Biofouling parameters	Select appropriate methods to permit a complete but convenient assessment of biofouling mechanisms present.
Step-drawdown and specific capacity (Q/s) testing	Flow and drawdown testing conducted periodically. Step tests to calibrate and before and after service events. Q/s testing periodically.
Power and pump performance parameters	Volts, amps, ohms in electrical power both instantaneous and trends, megohm ground leakage, vibration analysis, pressure at various Q to indicate wear.

Table 2. Recommended Data Collection Intervals (Depends on Site Conditions).[a] The Following is for a Difficult Well Situation (Valuable Well) After Its First Year

Type of Inspection	Method	Interval
Physical inspection	Borehole color video	At pump service intervals. Concentrate on screen and other stress points.
	Surface facility inspection	Quarterly
	Examination of pulled components	As needed. Wells should be equipped for easy pulling if at all possible.
Hydraulic performance	Well discharge (flow rate and pressure)	Weekly
	Drawdown	Weekly
	Conduct graphical analysis	Quarterly.
	Specific capacity test (well hydraulic performance).	Annually or at recommended shorter intervals.
	Pump performance: Conduct 5-step "pump" test, compare to "nominal" data	At least annually or at recommended shorter intervals if pump service is severe (Q/s and pump test can be a single operation).
Electrical (power)	System and motor current and load factors	Weekly—recommend installation of current monitors with alarms.
Physicochemistry	Inorganic parameters	At least quarterly using on-site instruments (calibrated) or routine monitoring (laboratory).
	Suspended particulate matter (sand, silt, clay)	Manually at well testing then quarterly.
	Turbidity or particle counting (adds colloidal component)	In-line monitors (continuous).
Biofouling microbial component	BART (or similar) analyses: Wide suite of ecological indicators	Quarterly until patterns develops then drop all but target types.
	Biofilm flow cell for microscopy	Annually on selected wells.
Treatments and service	Well hydraulic improvement and pumping systems	As testing indicates Q/s drops below 90% or pumping system degrades.
	Instrumentation calibration	In accordance with standard protocols.

[a] After Reference 3.

PM TREATMENT

PM well treatments are those applied prophylactically to remove biofouling and clogging prior to impairment of performance, which is a relatively complex and site-specific practice. However, such treatments should be targeted to the clogging conditions, as mild as possible, and capable of being easily applied.

REVIEW AND UPDATE

Methods and experience in PM change and improve over time. Experience in specific well situations also provides the basis for refining practices. Well PM plans should include provision for periodic review and revision of the PM plan and its practices. Also important is a provision for training personnel in new methods and training new employees in PM practices and purpose.

The subject of well maintenance is too complex to consider adequately in a brief article. The following are references used in the preparation of this item and additional recommended sources of information.

BIBLIOGRAPHY

1. NGWA. (1998). *Manual of Water Well Construction Practices*. National Ground Water Association, Westerville, OH.
2. Roscoe Moss Co. (1990). *Handbook of Ground Water Development*. Wiley-Interscience, New York.
3. Alford, G., Leach, R., and Smith, S.A. (2000). *Operation and Maintenance of Extraction and Injection Wells at HTRW Sites*. EP 1110-1-27. U.S. Army Corps of Engineers, St. Louis, MO.

READING LIST

ADITC. (1997). *Drilling: The Manual of Methods, Applications and Management*. CRC Lewis, Boca Raton, FL.

Borch, M.A., Smith, S.A., and Noble, L.N. (1993). *Evaluation and Restoration of Water Supply Wells*. AWWA Research Foundation, Denver, CO.

Carruthers, R. (1994). Preventative maintenance for bores. In: *Proceedings Water Industry Workshop Operations Skills*, July 25–29, 1994, Hervey Bay, Queensland Dept. of Primary Industries, Brisbane.

Gariboglio, M.A. and Smith, S.A. (1993). *Corrosión e encrustación microbiológica en sistemas de captación y conducción de agua–Aspectos teóricos y aplicados*. Consejo Federal de Inversiones, San Martin, C.F., Argentina.

Ground Water Science. (2004). Well M&R Gateway. Available: http://www.groundwatersystems.com.

Helweg, O.J., Scott, V.H., and Scalmanini, W.C. (1983). *Improving Well and Pump Efficiency*. AWWA, Denver, CO.

Howsam, P., Misstears, B., and Jones, C. (1995). *Monitoring, Maintenance and Rehabilitation of Water Supply Boreholes*. Report 137, Construction Industry Research and Information Association, London, UK, p. 113.

Smith, S.A. (1992). *Methods for Monitoring Iron and Manganese Biofouling in Water Supply Wells*. AWWA Research Foundation, Denver, CO.

Smith, S.A. (1995). *Monitoring and Remediation Wells: Problems Prevention and Cures*. CRC Lewis, Boca Raton, FL.

Smith, S.A. (in prep). *Methods for Monitoring Common Biofouling Problems of Ground Water Systems*. Ground Water Science, Upper Sandusky, OH.

Water Systems Council. (1992). *Large Submersible Pumps Manual*. Water Systems Council, Chicago, IL.

MEGAWATERSHEDS

Robert A. Bisson
Alexandria, Virginia

The term "megawatershed" describes deep-seated subsurface aquifer systems that may consist of gravel, fracture-hosted bedrock, and/or sedimentary structures, which are integrated in terms of recharge, storage, transmissivity, and containment. They may not coincide with surface topographic divides and may receive recharge from parts of several surface watersheds including massive amounts of mountain block recharge. "Megawatershed" is a conceptual model, or paradigm, that describes this new class of groundwater domain (Fig. 1). Megawatersheds are formed over geologic time by a combination of factors, including tectonically induced brittle fracturing and displacement of rock facies and strata and solutioning of parent rock facies and regolith by chemically aggressive meteoric and hydrothermal waters. The history and nature of regional tectonic stress fields and principal bounding faults, interacting with lithology and climate, play a major role in determining the geometry and extent of megawatershed boundaries.

Megawatersheds often transcend surface topographic divides and may receive recharge from parts of several surface watersheds. Similarly, as a structurally contained water resource, wells drilled into multiple aquifers in a single megawatershed may produce water from potentially dissimilar lithologies with common hydraulics related to brittle fracturing and gravity-fed contributions from adjacent and overlying unconsolidated sediments. Megawatersheds often occupy diverse host environments exhibiting both primary (continuous) and secondary (fracture) porosities with flow systems delineated by fracture systems, faulted and/or weathered lithologic contacts, and igneous intrusions (e.g., dikes), overlying or adjoining weathered rock and overlying porous unconsolidated sediments. This model supersedes the traditional watershed and aquifer models, which describe topographically controlled, functionally two-dimensional drainages and incorrectly depict most deep groundwater resources as static, poorly recharged artifacts of surface flow, confined to local discrete bedrock or alluvial units.

The megawatershed concept operates effectively and has scientific validity and economic and management utility on a scale of 10s to 1000s of square kilometers. Studies have documented megawatershed environments in Africa, Asia, the Middle East, Europe, North America, South America, and the Caribbean.

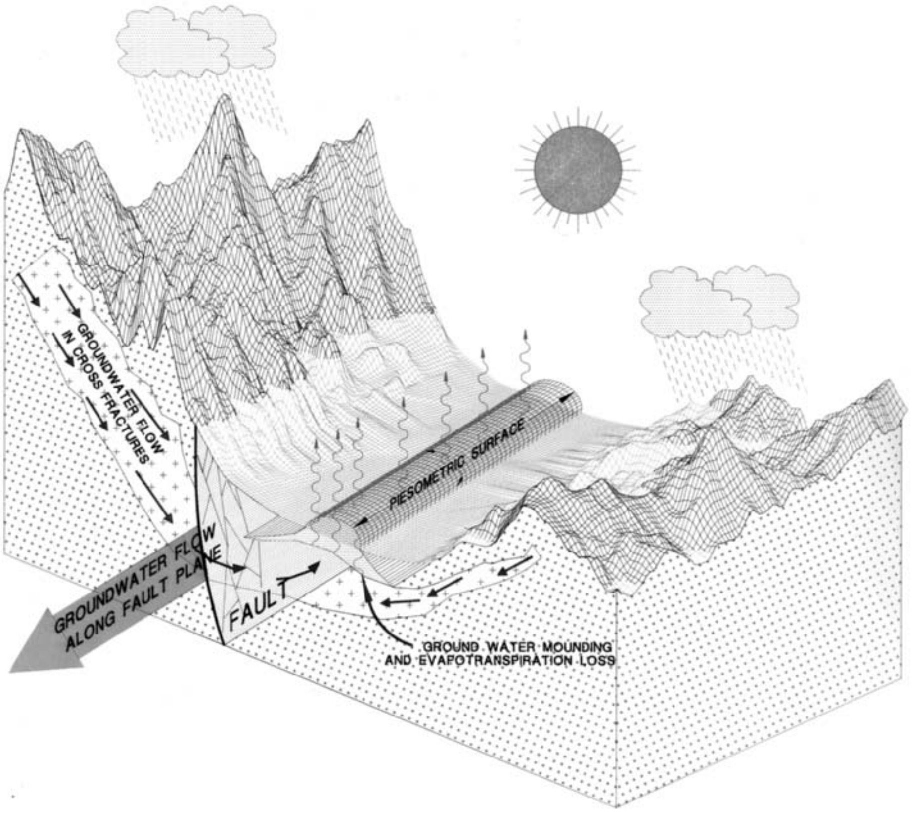

Figure 1. Megawatershed conceptual model indicating the roles of mountain block recharge and interbasin deep groundwater transfers within megawatersheds developed by Bisson Exploration Services Co. and CAD illustration by Summit Engineering, 1992.

HISTORY OF DISCOVERY

The space age has induced quantum leaps in geologic theory, and comparable advances in geophysical instrumentation and interpretive methods have permitted scientists to discover new information about the genesis, composition, and structure of the earth's crust and the internal dynamics of the many fluids, (e.g., magma, oil, water) contained within it. Equally important, the perfecting of computer-based geographic information systems (GIS) allowed skilled explorationists to combine and analyze many different types of data quickly and accurately, leading to major discoveries of oil, gas, minerals and, more recently, to breakthroughs in the understanding the nature and extent of groundwater resources, including the discovery of complex aquifer systems coined "megawatersheds" by the explorers who first documented the phenomenon, circa 1987, in the Great Rift systems of East Africa, where mountain block recharge from mountains and highlands in Ethiopia, Sudan, Kenya, and Chád remains largely untapped.

MEGAWATERSHEDS EXPLORATION PROGRAM

When the megawatershed paradigm is used as a groundwater exploration model, the result is greatly improved understanding of groundwater environments, permitting accurate, comprehensive assessments of "safe yield" from local, countrywide, or regional water resources. The megawatershed paradigm actually provides a template for measurement of all facets of the hydrologic cycle in reality-based natural hydrogeological catchments, including rainfall, evapotranspiration, surface and subsurface inflow and outflow interactions with surface water, shallow aquifers, deep fractured bedrock, and deep alluvial aquifers.

Modern explorationists employ current geologic theory, state-of-the-art exploration technologies, and interpretive methods adapted to the oil, gas, and mineral industries but rarely used in the groundwater field (Fig. 2). The "megawatersheds" paradigm of groundwater occurrence is the template the exploration program follows.

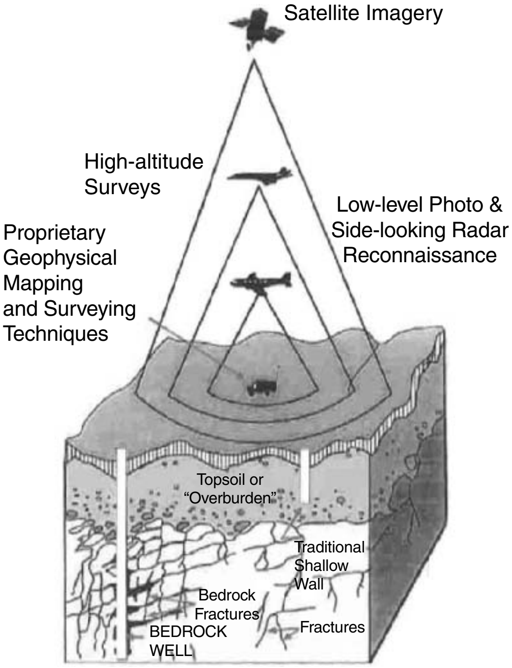

Figure 2. Illustration of megawatershed exploration program.

Case Studies

Two examples of megawatershed environments discovered using modern groundwater exploration methods are summarized here, one in the arid northwest region of Somaliland (1984–86) and the other in the humid small island developing state (SIDS) of Trinidad and Tobago, West Indies (1999–2002). These summaries are extracted from case studies included in Bisson and Lehr's book *Modern Groundwater Exploration: Discovering New Water Resources in Consolidated Rocks Using Innovative Hydrogeologic Concepts, Exploration, Drilling, Aquifer Testing and Management Methods*.

Somalia Case Study. In 1984, the Horn of Africa was experiencing a particularly brutal famine, brought about by the coincidence of periodic drought and regional and civil wars. America's ally, Somalia, was at war with the Soviet Union's ally, Ethiopia. At the same time, Somalia was mired in a bloody civil war, and the most active front for both conflicts was the remote, sparsely populated northwest part of Somalia, 1500 road-kilometers from the capital, Mogadishu, and bordered on the west and south sides by Ethiopia, to the north by the Gulf of Aden, and on the northwest, the last bastion of French colonialism in Africa, Djibouti. Several hundred thousand ethnic Somali refugees had been forced from their ancient tribal lands (annexed to Ethiopia after WWII) in the ongoing war and sought refuge in the arid Woqooyi Galbeed and Tog Dheer provinces of Somalia (Fig. 3). As a result of these circumstances, the socioeconomic fabric of the region was destabilized, cholera and other diseases were rampant in refugee camps located on the outskirts of major urban centers, and virtually no food was being produced locally.

Somalia's new patron, the United States, was working with the UNHCR and USAID to help feed and shelter those desperate people in an arid, undeveloped part of the Horn of Africa with all possible speed. Prior efforts to develop useful quantities of fresh groundwater in water-short areas had failed, and USAID's NGO contractors (New Transcentury Foundation) brought in a firm that pioneered modern groundwater exploration methods and discovered the megawatersheds phenomenon (BCI Geonetics, Inc.) to identify the groundwater resource base of the 35,000 km^2 region and develop emergency water sources for refugee camps and long-term water supplies for new refugee "cities."

Previous water development studies of northwest Somalia by other investigators focused on very limited surface and surficial water sources, but the BCI team focused on the region's bedrock groundwater potential, considering groundwater recharge and storage in unconsolidated overburden only in areas where they represented integral

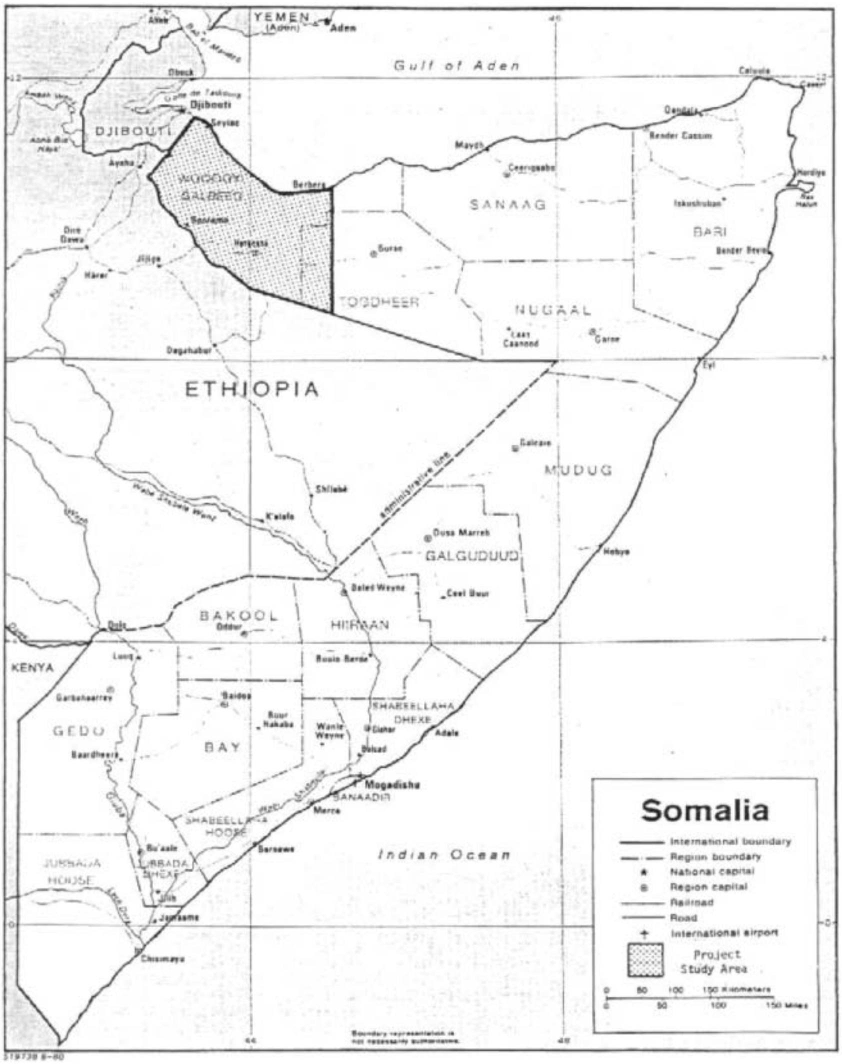

Figure 3. Somalia case study area.

hydrologic parts of deeper, underlying bedrock aquifer systems (megawatersheds).

To carry out their hydrogeologic mapping mission, the field team found only a few contemporary cultural and land-use maps and so augmented older maps with a combination of 1950s-vintage 1:125,000 scale British Ordinance Survey "spot-elevation" maps and RAF aerial photographs, plus 1980s- Tactical Pilotage Charts (TPCs), LandSat Image-maps, and high-resolution Space-Shuttle-borne Large Format Camera (LFC) photographs. The aerial and LFC photos were very useful navigational tools in certain remote areas where there were no roads leading to key bedrock outcrops and to other sites identified by BCI scientists on aerial photographs and satellite images. Thorn-bearing plants and sharp rocks punctured tires and automobile anatomies with great regularity, so undercarriage-armored Land Cruisers imported via neighboring Djibouti were outfitted with 12-ply tires for team expeditions.

The team was tasked to explore the farthest reaches of the county's lawless northwest region, and they often penetrated trackless areas where no vehicle or camel track existed and crossed unmarked borders between Ethiopia and Somalia. They were in constant danger of being captured by Ethiopian troops, strafed by MIGS during daylight, and held up at unmarked roadblocks by Somali bandits, revolutionaries, and government soldiers at night.

The region's wilderness also abounded with nonhuman predators, from hyenas to cobras, poisonous centipedes, and the nightmarish platter-sized, venomous "camel spider." The megawatersheds exploration team in Somalia encountered the most dangerous and difficult field conditions they had experienced to date, including a confrontation with Somali soldiers near the Ethiopian border that very nearly cost the lives of the team's senior scientists. BCI conducted this investigation using an advanced mineral-cum-groundwater exploration method, developed through extensive trial-and-error field-testing in the decade prior to the Somalia project. This systematic, practical approach to quantifying targeted minerals was singularly successful at identifying, drilling, and developing sustainable groundwater sources for nearby consumers in the most economical way.

The premise of the team's modern groundwater exploration program was that groundwater shares and is controlled by the same geologic environments as minerals and is most readily identified and evaluated using modern approaches proven successful in oil, gas, and economical mineral exploration. The successful application of a sophisticated groundwater exploration program required a practiced exploration team of experts in structural geology, remote sensing, surficial geology, exploration geophysics, geochemistry, hydrology, and hydrogeology. Only with an integrated team such as this could the full groundwater potential of the study area be determined.

This was a pioneering, remote-sensing project incorporating LandSat MSS imagery and high-resolution contact prints from the ITEC-built Large Format Camera (LFC), carried aboard NASA's Space Shuttle Challenger in October 1984. NASA supplied stereo-paired photographs of the extreme westernmost part of the study area, and the BCI team used the extraordinary 60,000 square kilometer "footprint", 5-meter resolution, and high planimetric accuracy of enlarged LFC prints both as "ground truth" for BCI's geological interpretations of LandSat MSS image maps and also as road maps, navigating for 2 years through thousands of square kilometers of poorly mapped, highly dangerous terrain in Somalia. The team also integrated specially processed satellite MSS image-data and British Ordinance aerial photographs into its map database for field exploration.

The exploration team's satellite image and aerial photographic analysts first studied the geological history and structural environment of the study area and then focused on identifying true surface expressions of underlying geologic structures consistent with the known tectonic fabric of the region. This work produced accurate working maps of fracture and fault systems on a scale common with other working map products. In the final phase of this part of the analysis, correlations were made between these newly (remote-sensing) interpreted "structural lineaments" and team-constructed tectonic and hydrogeologic models, verified by published geologic maps and field mapping.

In the evaluation and selection process that led to the delineation of favorable zones for further hydrologic analysis and fieldwork, the significance attached to structural lineaments increased with the strength of expression, overall length, and spatial correlation with mapped faults and fractures shown in the fault and brittle feature map—particularly those faults belonging to one of the three major structural trends described in the full report.

An investigative hydrogeologic analytical methodology was established to quantify the amount of active recharge potentially percolating through the vadose zone and available for recharge to deeper alluvial aquifers and conductive fracture zones feeding regional groundwater flow systems in northwest Somalia. The methodology developed synoptically evaluated hydrometeorological data over the study area, including identification of the dominant variables and parameters affecting recharge. Variables evaluated included topography, elevation, slope, evapotranspiration, surface roughness and permeabilities (i.e., caliche layer occurrence), soil infiltration rates and depth, wadi bed characteristics (same as surface and soils), and wadi bed surface area. These variables were investigated using data from previous studies and in the literature, plus results from new fieldwork.

Several reporting stations had documented northwest Somalia's annual precipitation, and rainfall maps of northwest Somalia had previously been published, but meteorological stations were sparse, daily precipitation was not recorded, and the methodology and assumptions on which rainfall maps were based were uncertain. The BCI team therefore sought a means of working with existing data to make it more useful by establishing correlative relationships between documented precipitation and study area topography, thereby generating a rainfall map reflecting statistical correlations of average annual rainfall with elevation. A linear regression was done with rainfall as the dependent variable and elevation as the independent variable.

Wadi beds provide an efficient pathway to the water table and, therefore, a viable mode of recharge. No information on rates of recharge through wadi beds was found. Thus, fieldwork was carried out to estimate recharge through wadi beds. Infiltration rates of the wadi bed surfaces were tested in the field to determine just how quickly storm water would be absorbed. Surface areas of the wadi beds were measured to provide relationships between infiltration rates and volume of recharge.

The slope of the land surface affects the ratio between infiltration and surface runoff. The steeper the slope, the greater the percentage of surface runoff leading to flooding conditions in the wadi beds. Steep slopes also generally have thinner soil cover and exposed bedrock, which increases the chances of groundwater recharge from the infiltration that does take place.

In spite of the early termination of the BCI project due to civil war before large production wells were drilled, the project generated the first-ever strategic groundwater development map for the region, identifying the potential for tens of millions of gallons per day of renewable groundwater available from regional fractured bedrock aquifers (megawatersheds) and solved the water crisis for many refugees and indigenous people by supplying more than 1.5 million gallons per day of fresh, potable water from six test wells. Some of the bedrock test wells reportedly have continued to supply fresh water to Borama's citizens and others for the past 18 years.

Small Island Developing States—Trinidad and Tobago Case Studies. In 1999 and again in 2000, the Water and Sewerage Authority (WASA) of Trinidad and Tobago engaged a team of exploration scientists who utilized the megawatersheds model and modern groundwater exploration technologies in an attempt to gain independence from unreliable surface water sources and to avoid the further expense and environmental impacts of additional dams and desalination plants, especially in the highly sensitive rain forest and coral reefs of Tobago. All previous attempts to find large quantities of naturally recharged and sustainable fresh groundwater using traditional hydrogeologic models and groundwater development methods had failed.

In the quest to accomplish the most comprehensive and accurate groundwater assessment ever implemented anywhere, WASA's contractor implemented data searches on the broadest possible front and applied the most advanced digital, GIS database construction, management, and computer-visualization technologies available, collecting and digitizing diverse data never before pooled; performing analyses of regional tectonic models and satellite remote sensing; digitizing over 15,000 data points from oil and gas seismic lines; carrying out hundreds of line-kilometers of geophysical surveys; interpreting 1600 hundred logs from oil, gas, and water wells; and completing intensive on-ground geologic mapping in prospective megawatersheds.

The exploration team then employed a novel "multiple convergent datasets" method of data processing and integration to produce a working groundwater map of Tobago (Fig. 4) and a state-of-the-art digital hydrogeologic map of Trinidad (Fig. 5). The new maps identified several major new megawatersheds in Tobago and increased the known, sustainable groundwater reserves almost 40-fold, from less than 0.5 million imperial gallons per day (imgd) to 40 imgd.

Subsequent drilling and long-term pumping (>36 months in several cases) confirmed the megawatersheds models and actually developed nearly 5 imgd of new deep bedrock groundwater sources for Tobago for less than US$ 10 million in 1 year, producing an amount of new water equal to a proposed US$ 80 million dam, slated for 5 years of construction in Tobago's protected rainforest.

On the island of Trinidad, modern groundwater exploration methods were successfully employed to identify megawatersheds, bedrock and sedimentary aquifers, and aquifer systems that increased the amount of known, safely extractable groundwater more than 10-fold; the models were validated by test wells and 15.8 imgd of production well long-term pumping [>24 months in several cases to date (Sept.2004)]. Digital copies of the new island-wide maps were provided to WASA for multiagency use within all government ministries.

Trinidad and Tobago is the first country and also the first small island developing state (SIDS) to employ, the megawatersheds paradigm and modern groundwater exploration methods on a national scale, in a highly successful attempt to address short-term water shortages and long-term economic development issues. Within 3 years, Trinidad and Tobago's use of the megawatershed exploration approach resulted in groundwater discoveries in the twin-island state exceeding 300 million US gallons per day, causing a true paradigm shift in the country's vision of its fresh water resources from chronic dependence on unreliable surface water sources to drought- and contamination-resistant, deep, groundwater sources.

The benefits of the Trinidad and Tobago Water Authority's innovative water development program was evident during the 2001 and 2003 dry seasons on the island of Tobago, when surface supplies were severely curtailed and the megawatersheds wells became the primary source of reliable drinking water for the island's indigenous and tourist populations.

EXOHYDROLOGY—IMPLICATIONS OF THE MEGAWATERSHEDS PARADIGM ON MARS

In the 1980s, groundwater investigators hypothesized that the striking similarity of geomorphologies between terrestrial and Martian deserts suggests a comparable geotectonic and hydrometeorological genesis, possibly indicative of paleomegawatershed formations that might retain groundwater in deep bedrock fractures and/or provide bedrock pathways for modern, tectonically induced groundwater flows (see Bisson and El-Baz, 1988).

BIBLIOGRAPHY

AFRICA

Alam, M. (1989). Water resources of the Middle East and North Africa, with particular reference to deep artesian ground water resources of the area. *International Report–World Bank and in Water Int.* **14**: 3.

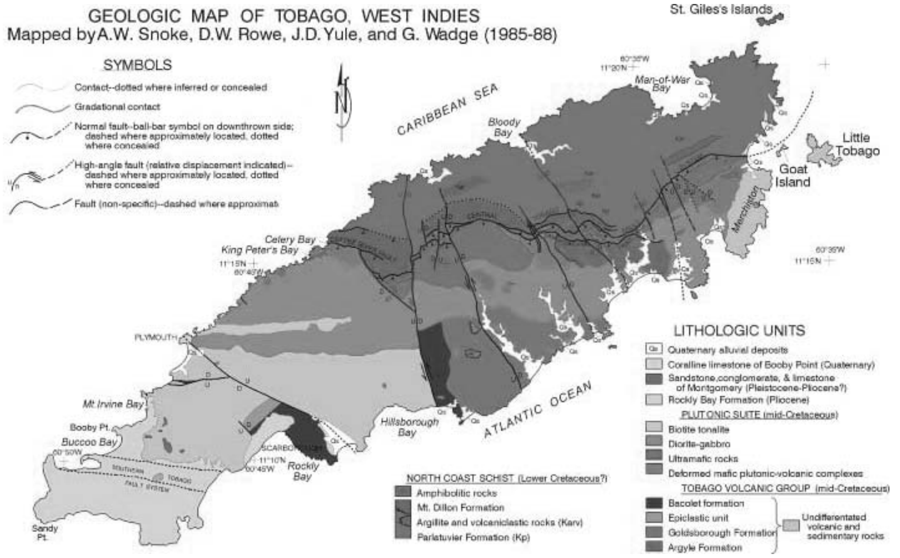

Figure 4. Tobago megawatersheds and favorable areas map.

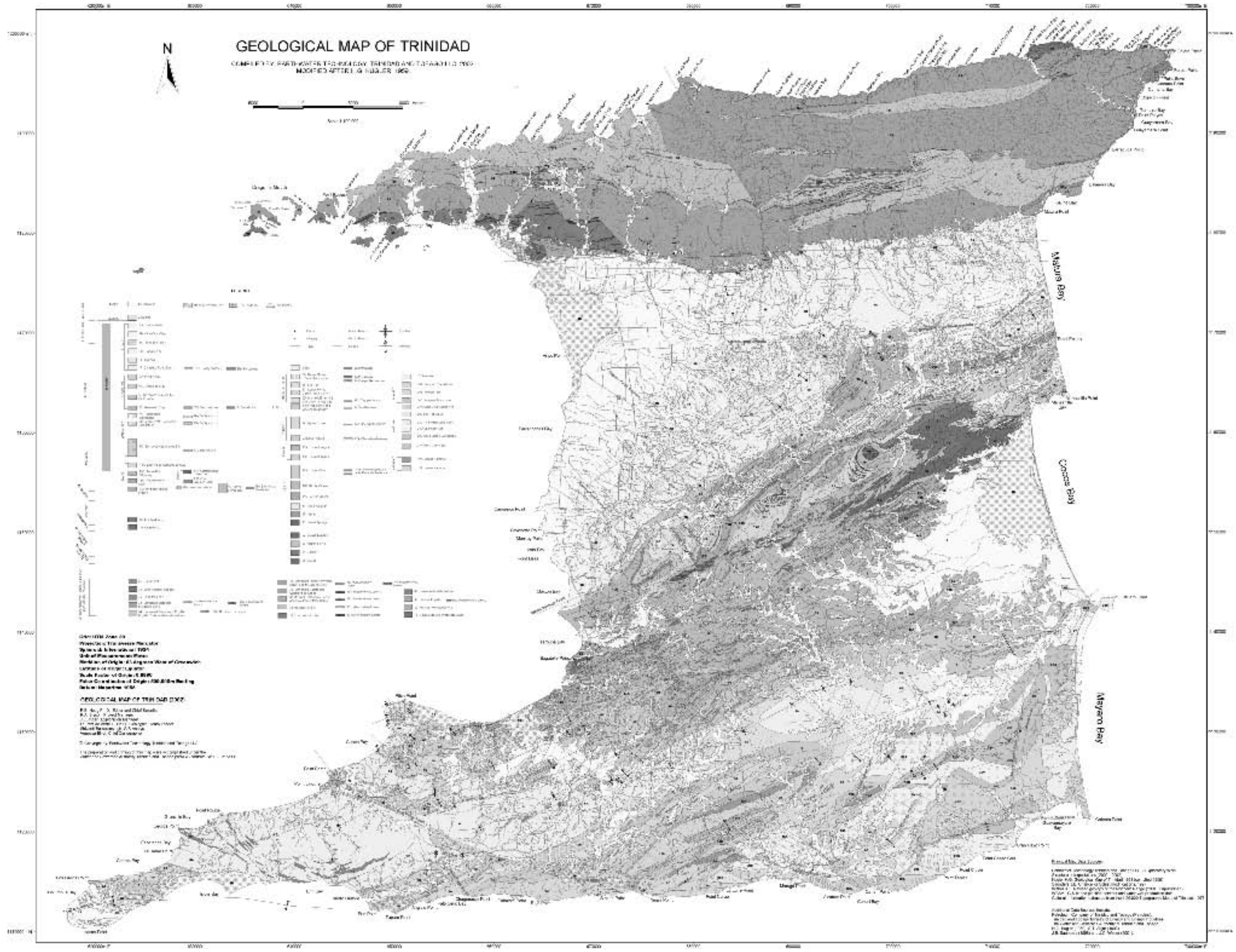

Figure 5. Hydrogeologic map of Trinidad, 2002. Copyright Earthwater Technology Trinidad & Tobago LLC and Water and Sewerage Authority of Trinidad and Tobago.

Bisson, R.A., Hoag, R.B., Jr., Ingari, J., and DeMars, R.M. (1984–1986). *BCI-Geonetics, Regional Groundwater Assessment and Development Program for Northwestern Somaliland*. Reports and Maps; for USAID.

Bisson, R.A., Bisson Exploration Services Company. (1992). *Megawatersheds Development Feasibility Study of Ethiopia*, Report and Maps, for U.S. Office of Foreign Disaster Assistance (OFDA).

Bisson, R.A. and Hoag, R.B. (1985). *BCI-Geonetics, Regional Groundwater Resources of South Africa, Namibia and Botswana*. Report and Maps, for British Petroleum Southern Africa.

Bisson, R.A. (1987). *BCI-Geonetics, Megawatersheds Development Feasibility Study of Northern Sudan*. Report and Maps; for USAID Mission Khartoum.

Bisson, R.A., Hoag, R.B., Jr., Ingari, J., Hoisington, D., and DeMars, R.M. (1990). *New Groundwater and Economic Prosperity in the Red Sea Province, Sudan*. Technical and Economic Feasibility Report, with Maps and Test and Production Well Plan, for USAID Misson Khartoum.

Bisson, R.A. (1986). *BCI-Geonetics, Inc., Megawatershed Development Feasibility Study of Eastern Zimbabwe, with Economic Analysis by Bechtel International 1983*. submitted to USAID TDA.

Ingari, J.C. (1986). *BCI-Geonetics, Inc., Sabi Basin Groundwater Geophysical Survey Report and Maps*. for USAID TDA.

ASIA

Afghanistan

Bisson, R. (2003). *Megawatersheds preferability study, Report and Maps, for Chemonics International* Earthwater Technology International.

Malaysia

Bisson, R. (1997). *Feasibility Study for Enron International*. Earthwater Technology International.

Nepal

Bisson, R. (1997). *Feasibility Study for Enron International* Earthwater Technology International.

Singapore

Bisson, R. (1997). *Feasibility Study for Enron International* Earthwater Technology International.

MIDDLE EAST

General

Bisson, R.A., Roosevelt, K., and Siski, S. (1999). New Groundwater Discoveries as a Mechanism for Regional Conflict Resolution. *Proceedings of the Conference on Water Dispute Prevention & Development*, Center for the Global South, American University, Washington, DC, October 1998.

Bisson, R.A., Sheffield, S., and Siski, S. (1995) Megawatershed Exploration: A State-of-the-Art Technique Integrating Water Resources and Environmental Management Technologies. *Proceedings of the IDA World Congress on Desalination and Water Sciences*, AbuDhabi, UAE, November 18–24, 1995.

Bisson R.A. (1994). Integrated Exploration Techniques Siscover Potential for Renewable Regional Source of Pristine Groundwater in Fractured Bedrock. *Proccedings of The Fifth International Symposium on Application of Remote Sensing and GIS Technologies to Water Resources Management*, Damascus, Syria, Oct.31–Nov.3, 1994.

Bisson R.A. (1994). Space Age Integarted Exploration and Treatment of Renewable Regional Sources of Pristine Groundwater in Fractured Rock Megawatersheds. *Desalination*, **99**: 257–273.

Bisson R.A. (1993). The Megawatersheds Discovery, The Arid World's Water Legacy of Geotectonics. *International Symposium on Water Resources in the Middle East: Policy and Institutional Aspects*, Beckman Institute, University of Illinois at Urbana, October, 1993.

Israel

Bisson, R. (1996). *Feasibility Study for Enron International* Earthwater Technology International.

Jordan

Bisson, R. (1997). *Feasibility Study for Enron International* Earthwater Technology International.

Libya

Bisson, R. and El-Baz, F. (1987). *Evaluated as part of Northern Sudan Megawatershed Study*, for USAID, BCI Geonetics, Inc.

Saudi Arabia

Western

Bisson, R.A. (1994). Integrated exploration techniques discover potential for renewable regional sources of pristine groundwater in fractured bedrock. *Proc. Fifth Int. Symp. Appl. Remote Sensing and GIS Technol. Water Resour. Manage.*, Damascus, Syria, Oct. 31–Nov. 3, 1994.

Bisson, R.A. (1994). Space Age Integrated Exploration and Treatment of Renewable Regional Sources of Pristine Groundwater in Fractured Rock Megawatersheds. *Desalination* **99**: 257–273.

Bisson, R.A. (1993). The megawatersheds discovery, the arid world's water legacy of geotectonics. *Int. Symp. Water Resour. Middle East: Policy and Institutional Aspects*, Beckman Institute, University of Illinois at Urbana, October, 1993.

Bisson, R.A. and El-Baz, F. (1990). New groundwater for Sudan - the megawatersheds exploration model. *Proc. 23rd Int. Symp.* Remote Sensing Environ., Bangkok, Thailand, April, 1990.

Bisson, R.A. and El-Baz, F. (1990). Megawatersheds exploration model for African water development. (Presented by Mr. Ghali for authors). *African Water Technol. Exhibition*, Nairobi.

Bisson, R.A. (1989). Megawatersheds exploration model. *Third World Acad. Sci. Int. Conf. Desert Environ.*, Trieste, Italy, November.

EUROPE

Ireland (Republic)

Bisson, R. and Hoag, R. (1982). *Feasibility study for private developer* BCI Geonetics, Inc.

Bisson, R.A. and El-Baz, F. (1990). Megawatershed exploration model. *Water Wastewater Int. Conf.*, Barcelona, Spain, April.

Poland–West

Bisson, R. (1991). *Feasibility study for Sargent Drilling Company* Bisson Exploration Services Company.

Sardinia

Bisson, R. (1997). *Feasibility Study for Enron International* Earthwater Technology International.

Spain, Andalucia

Bisson, R. (1997). *Feasibility Study for Enron International* Earthwater Technology International.

NORTH AMERICA

Angillilo, O.R. (1959). *Replenishing Source of Waters Flowing Through Rock Fissure Aquifers*. Unpublished manuscript, presented to State of California.

Bisson, R.A. (1989). Groundwater—The Ultimate Economic Mineral. *Proceedings, the 2nd Annual Conference on Profiting from Water*. Santa Monica, CA, May 1989.

Bisson, R.A. (1999). *Privatization Water Delivery Economics and Updating the Global Water Balance: A Win-Win Scenario* U.S. Water Report, Saringa Group, Inc., November.

Bisson, R.A. and Hofmann, P.D.(1982). No Risk/Turnkey Water Supply Contracts. *Journal of the New Hampshire Water Works Association*, Feb. 1982.

Bisson, R.A. and Hofmann, P.D.(1989). Groundwater—The Paradoxical Economic Mineral, *Water and Wastewater Magazine*. April 1989.

Bisson, R.A. and Widger, W.K.(1977). Innovative Techniques for Water Resources Studies. *Journal of the New England Water Works Assn*, June 1977.

Burbey, T.J. and Prudic, D.E. (1991). *Conceptual Evaluation of Regional Ground-Water Flow in the Carbonate-Rock Province of the Great Basin, Nevada, Utah, and Adjacent States*. Regional Aquifer System Analysis. U.S.G.S Professional Paper 1409-D.

Church, T.M. (1996). An underground route for the water cycle. *Nature* 380.

Clark, R.D. (1983). Geophysical Exploration—No Water Shortage? *Geophysics: The Leading Edge of Exploration*. **2**(7) July.

Eastoe, C.J., Gu, A., and Long, A. (2004). The origins, ages, and flow paths of groundwater in the Tucson Basin: results of a study of multiple isotope systems. In: *Groundwater Recharge in a Desert Environment: The Southwestern United States.* Hogan, J.F., Phillips, F.M., and Scanlon, B.R., (Eds.). American Geophysical Union, Washington, DC.

Hoag, R.B., Jr. and Vernon, J.H. (1993). Wellhead Protection Area Delineation in a Fractured Bedrock Aquifer in Fayetteville, Georgia. *Proceedings of the 1993 Georgia Water Resources Conference*, Athens, GA.

Long, A. and Eastoe, C. (2003) Mountain-block recharge. In: *Sustainability of Semi-Arid Hydrology And Riparian Areas (SAHARA)*, September.

Manheim, F.T. (1967). Evidence for submarine discharge of water on the Atlantic Continental Slope of the Southern United

States, and suggestions for further search. *Trans. N. Y. Acad. Sci., Sect. Geol. Sci. Ser. II* **28**: 839–853.

Manning, A.H., Canie, J.S., and Landis, G.P. (2003). Mountain-Block Recharge To The Western Denver Basin, Colorado, Paper 11-4, *Rocky Mountain—55th Annual Meeting of the Geological Society of America*, Durango, CO, May.

Person, M. and Baumgartner, L. (1995). New Evidence for long-distance fluid migration within the earth's crust. US. National Report to International Union of Geodesy and Geophysics 1991–1994, Rev. *Geophys.*, Suppl. 1083–1091, July.

Person, M., Raffensperger, J.P., Ge, S., and Garven, G. (1996). Basin-scale hydrogeologic modeling. *Rev. Geophys.* **34** (1/Feb.): 61–87.

Widger, W.K. and Bisson, R.A. (1974). News From and About the Earth Resources Technology Satellite (ERTS). *NH Society of Professional Engineers' Journal*. OBSERVER Sept. 1974.

Williams, M., Liu, F., and Gertson, J. (2004). Mountain block recharge from snowmelt runoff in the Colorado Rocky. *Conference Proceedings AGU 24th Annual Hydrology Days 2004*. Colorado State University, American Geophysical Union, March.

SOUTH AMERICA

Magaritz Aravena, M., Pena, R.H., Suzuki, O., and Grilli, A. (1990). Source of ground water in the deserts of Northern Chile: Evidence of deep circulation of ground water from the Andes. *Ground Water* **28**: 4.

CARIBBEAN

Bisson, R.A., Hoag, R.B., Ingari, J.C., Maharaj, U.S., and Jadoo, L. (2001). Megawatersheds groundwater assessment and recharge calibration on the Island of Tobago, W.I.; A comparison of results using megawatersheds versus traditional methods of groundwater assessment. *CWWA 10th Annu. Conf. & Exhibition*, Grand Cayman, October.

Bisson, R.A., and Maharaj, U., (2003). Megawatersheds of the Caribbean: New Water from Old Rocks. Presented at the *Caribbean Water and Wastewater Association (CWWA) 12th Annual Conference*, Paradise Island, Bahamas, September 29–October 3, 2003.

de Verteuil, L., Stawicki, A.W., Hoag, R.B., Bisson, R.A., Ingari, J.C., Maharaj, U.S., and Mulchansingh, K. (2001). Sedimentary megawatersheds delineation: The role of genetic sequence stratigraphy and Complex fault systems analysis in groundwater resource assessment in Trinidad. *CWWA 10th Annu. Conf. & Exhibition*, Grand Cayman, October.

Durham, B., Yoxtheimer, D., Alloway, C. and Diaz, C. (2003). Innovative water Resources for Islands. *Desalination* **156**: 155-161.

Hoag, R.B., Bisson, R.A., Ingari, J.C., Maharaj, U.S., and Sankar, R. (2000). The megawatershed concept and recharge analyses—sustainable groundwater extraction from bedrock aquifers in Tobago. *9th Annu. Conf.* Exhibition, October 2–6, Trinidad and Tobago.

Hoag, R.B., Bowman, C. and Bisson, R.A. (2001). Estimation of Groundwater Recharge in Trinidad Using Meteorological, Geographic Information Systems(GIS), and Watershed Modeling. *CWWA 10th Annual Conference & Exhibition*, Grand Cayman Island, October 2001.

Hoag, R.B., Cormier, C., Bisson, R.A., Sankar, R., Jadoo, L., and Ingari, J. (2001). Exploration and pumping test results for a high-yield crystalline bedrock well in Tobago. *CWWA 10th Annu. Conf. & Exhibition*, Grand Cayman, October.

Maharaj, U.S., Bisson, R.A., and Hoag, R.B. (2001). Novel groundwater development technology and contract structure solves long-standing water shortage problems on the Island of Tobago, West Indies. *Fourth Inter-Am.* Dialogue Water Manage., Parana, Brazil, September.

Maharaj U.S. and Tota-Maharaj, T. (2001). High-Risk Groundwater Development Option for Small Island Developing States(SIDS). *CWWA 10th Annual Conference & Exhibition*, Grand Cayman Island, October 2001.

EXOHYROLOGY-MARS

Bisson, R.A. and El-Baz, F. (1988). *Paleowatersheds on Mars: The bedrock legacy, a feasibility proposal submitted to the Lunar and Planetary Institute*, Houston Texas, by Bisson, R.A. principal investigator and El-Baz, F. co-investigator.

MASS TRANSPORT IN SATURATED MEDIA

Paul F. Hudak
University of North Texas
Denton, Texas

ADVECTION AND HYDRODYNAMIC DISPERSION

Most of the matter transported by groundwater is in dissolved form, which is referred to as solute. Principally, advection and hydrodynamic dispersion affect the movement of solutes in groundwater. Advection occurs when flowing groundwater carries along dissolved solutes. Hydrodynamic dispersion is the spreading of groundwater and its dissolved constituents by mechanical mixing and diffusion. Mechanical mixing is spreading caused by the aquifer matrix. Diffusion, which can take place in groundwater that is not flowing, is the movement of solutes from areas of high to low concentration. This latter process can be illustrated by placing a drop of dye in a beaker of stagnant water. The dye will spread even though the water is not flowing.

Longitudinal dispersion expresses mixing along flow paths. At the pore scale, mechanical causes of longitudinal dispersion include (1) fluid moving faster in the center of pores than along the edges, (2) some fluid traveling along longer pathways than other fluid, and (3) fluid in larger pores traveling faster than fluid in smaller pores (1).

Transverse dispersion occurs in all directions in a plane perpendicular to groundwater flow. However, studies of mass transport often report only the vertical and horizontal components of transverse dispersion. Bifurcating flow paths cause transverse dispersion (Fig. 1).

Beyond the pore scale, bodies of rock or sediment with different hydraulic conductivity may induce mechanical dispersion within an aquifer. For example, groundwater and dissolved solutes may diverge around a clay lense within a predominantly sandy aquifer. This process is called macrodispersion.

The magnitude of longitudinal dispersion exceeds that of transverse dispersion, as solutes spread more in the direction of groundwater flow than in any direction within a perpendicular plane. Coefficients of longitudinal (subscript L), transverse horizontal (subscript

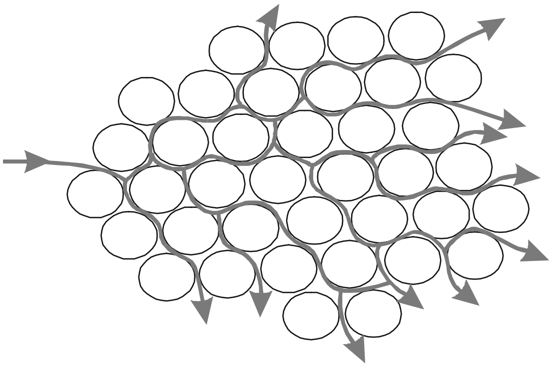

Figure 1. Pore-scale transverse dispersion in groundwater flowing between sand particles.

TH), and transverse vertical (subscript *TV*) hydrodynamic dispersion can be defined as follows:

$$D_L = \alpha_L v + D^* \quad (1)$$

$$D_{TH} = \alpha_{TH} v + D^* \quad (2)$$

$$D_{TV} = \alpha_{TV} v + D^* \quad (3)$$

where α is the dispersivity of the medium and D^* is the effective molecular diffusion coefficient. Dispersivity and the diffusion coefficient have units of length and length-squared/time, respectively. D^* ranges from 10^{-11} to 10^{-8} m^2/s for most ions in groundwater. Table 1 lists diffusion coefficients of several solutes in open water, which should be multiplied by a coefficient (less than one) to account for slower diffusion rates in tortuous, porous media. For unconsolidated media, this coefficient ranges from approximately 0.01 to 0.5 (3).

Dispersivity quantifies the ability of a medium to induce spreading by mechanical means. The magnitude of the mechanical mixing component of hydrodynamic dispersion (first term in Eqs. 1 to 3) usually exceeds molecular diffusion in flowing groundwater. Diffusion is a significant component of Eqs. 1 to 3 only for very slow-moving groundwater.

Dispersivity is a scale-dependent parameter, which attains higher values as the length of the transport domain increases. Laboratory estimates of dispersivity are routinely smaller than are field estimates. In a critical review of dispersivity observations from 59 different field sites, Gelhar et al. (4) found that, for data of high reliability, longitudinal dispersivity ranged from approximately 0.4 to 4 m. Transverse horizontal dispersivity was typically an order of magnitude less than longitudinal dispersivity, and transverse vertical dispersivity was an order of magnitude smaller than transverse horizontal dispersivity.

Table 1. Diffusion Coefficients in Water at 25 °C (2)

Solute	Coefficient (m^2/s)
Na^{+1}	1.33×10^{-9}
K^{+1}	1.96×10^{-9}
Mg^{+2}	7.05×10^{-10}
Ca^{+2}	7.93×10^{-10}
Fe^{+2}	7.19×10^{-10}
F^{-1}	1.46×10^{-9}
Cl^{-1}	2.03×10^{-9}
Br^{-1}	2.01×10^{-9}
$HCO_3{}^{-2}$	1.18×10^{-9}
$SO_4{}^{-2}$	1.07×10^{-9}
$CO_3{}^{-2}$	9.55×10^{-10}

Laboratory sand column experiments demonstrate the process of hydrodynamic dispersion. Distilled water runs through a column at a constant rate. The chloride concentration of the influent is suddenly changed from 0 to C_0 (any arbitrary concentration significantly higher than 0, for example, 1000 mg/L). Chloride concentrations, C, are monitored at the outlet, and the ratio C/C_0 is plotted against time (Fig. 2).

Early on, there are no chloride ions at the outlet; thus C and C/C_0 equal 0. Among the first group of chloride ions entering the tube, some arrive at the outlet ahead of others (because of hydrodynamic dispersion). As more chloride ions reach the outlet, C gradually increases to C_0. Had the solute been transported only by advection, the plotted curve would rise suddenly rather than gradually to C_0. The solute front "breaks through" the outlet when C/C_0 equals 0.5.

REACTIVE SOLUTES

Groundwater transports both chemically conservative and reactive solutes. Conservative solutes do not react with the aquifer medium or groundwater. Examples of conservative solutes are nitrate, chloride, and bromide. Reactive solutes have a tendency to sorb to aquifer media, moving from the liquid to solid phase, and they may go back into solution later on. Many synthetic organic chemicals are reactive, as are many cations. Negatively charged surfaces of clay particles are good sorbing media. If a solute is reactive, it will travel at a slower rate than the groundwater.

The velocity at which a reactive solute travels relative to flowing groundwater depends on both the solute and

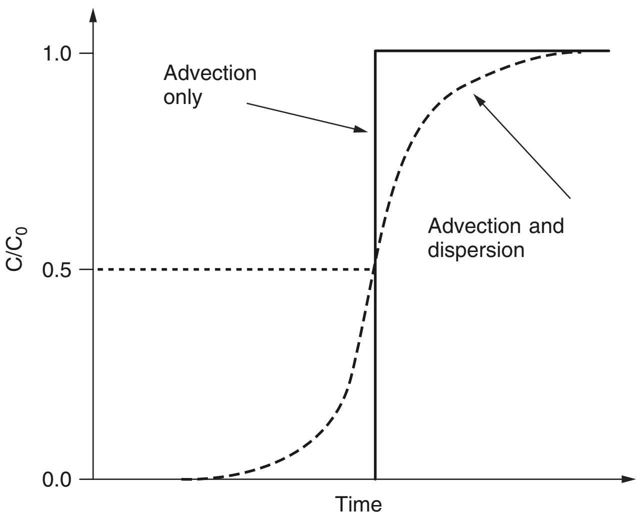

Figure 2. Solute breakthrough curve.

the aquifer material. A retardation factor, R_f, of 2 implies that the reactive solute travels at one half the velocity of groundwater, a factor of 5 implies one-fifth the velocity of groundwater, and so forth. Higher R_f values indicate slower solute transport.

Retardation factors should be carefully considered when delineating contaminant plumes. Multiple contaminants originating from the same source may occupy vastly different volumes in groundwater. Contaminants with higher retardation factors will occupy smaller volumes of aquifer. Conservative contaminants will have migrated further, and they best describe the maximum spatial impact of a leaking source. Chemical transformations and biodegradation, not quantified in R_f, also affect the spatial extent of a contaminant plume.

R_f can be computed as

$$R_f = 1 + \frac{\rho_b}{n} K_d \tag{4}$$

where K_d (commonly mL/g or L/kg) is the distribution coefficient for the solute with the aquifer. K_d can exceed 100 mL/g for reactive solutes and equals 0 for conservative solutes (5).

The distribution coefficient for a compound with a particular soil sample can be determined in a laboratory. An analyst fills jars containing soil with solutions of various contaminant concentrations. Several batches are created in this manner. The concentration of the contaminant in solution and the concentration in sediment are measured for each batch. These data are plotted on a curve (called an isotherm) fitted through the points (Fig. 3). The isotherm can take on different shapes. In the simplest case, it is linear, and K_d is the slope of the line (3). Thus, K_d has units of volume (of solution) per mass (of sediment). Langmuir (5) discusses methods for computing K_d for nonlinear isotherms.

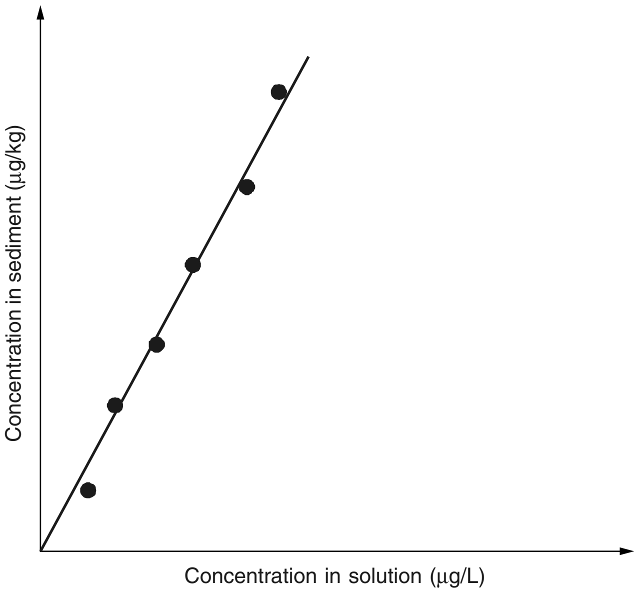

Figure 3. Data from laboratory batch experiment.

BIBLIOGRAPHY

1. Fetter, C.W. (2001). *Applied Hydrogeology*. 3rd Edn. Macmillan, New York.
2. Li, Y.H. and Gregory, S. (1974). Diffusion of ions in sea water and in deep-sea sediments. *Geochimica et Cosmochimica Acta*: 38.
3. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
4. Gelhar, L.W., Welty, C., and Rehfeldt, K.R. (1992). A critical review of data on field-scale dispersion in aquifers. *Water Resour. Res.* **28**(7): 1955–1974.
5. Langmuir, D. (1997). *Aqueous Environmental Geochemistry*. Prentice-Hall, Upper Saddle River, NJ.

SOIL AND WATER CONTAMINATION BY HEAVY METALS

Kristofor R. Brye
Kenneth F. Steele
University of Arkansas
Fayetteville, Arkansas

Soil and water contamination by heavy metals has a major adverse environmental impact. Concern has long existed over the threat of contamination of soil and water resources by heavy metals but has heightened in recent years as more knowledge has been gained about the negative and potentially serious affects that chronic exposure to heavy metals may have on the health of humans and other living organisms. Long-term exposure to lead affects human cognition; cadmium affects kidney, liver, and gastrointestinal tract functioning; and arsenic affects the skin, kidneys, and central nervous system (1). This article will address terminology, sources, mobility, risk assessment, and management and remediation alternatives for heavy metals in the environment that may contaminate soil and water resources.

TERMINOLOGY

Heavy metal is a term used to describe a subset of a large number of metals and metalloids, listed in the periodic table of the elements, which have properties of both metals and nonmetals. In general, heavy metals are elements in which the differentiating electrons are bonding electrons and typically include the top row of transition elements in the periodic table [i.e., scandium (Sc), titanium (Ti), vanadium (V), chromium (Cr), manganese (Mn), iron (Fe), cobalt (Co), nickel (Ni), copper (Cu), zinc (Zn), cadmium (Cd), and mercury (Hg)]. Arsenic (As), selenium (Se), and lead (Pb) may have electrons in outer shells (i.e., d electrons), as the differentiating electrons depending on the charge on the ion and thus are usually considered heavy metals. Heavy metals, except Fe, are also often referred to as trace metals due to their relatively small abundance in the environment.

SOURCES OF HEAVY METALS IN THE ENVIRONMENT

There are a wide variety of natural and anthropogenic sources of heavy metals in the environment that may impact or contaminate soil, surface water, and groundwater resources. Some heavy metals occur naturally in relatively small quantities as constituents of igneous and sedimentary rocks. For example, Fe composes 8.6% of basalt and only 0.38% of limestone. Titanium and Mn are also present in rocks in relatively high concentrations (0.1—1.0% range). The remainder of the heavy metals is present in rocks as trace elements ranging in concentration from 0.05–130 mg kg^{-1} (0.05 to 130 ppm) and typically are present in the 10–100 mg kg^{-1} range. Generally, heavy metals are present in higher concentrations in mafic rocks, such as basalt, and in shale. Mafic refers to igneous rocks that are largely composed of silicate minerals rich in Fe and Mg (ferromagnesium minerals), such as olivine, pyroxene, amphibole, and biotite. Physical and chemical weathering of heavy-metal-containing rocks release heavy metals into the environment. Some heavy metals contained in secondary minerals become part of the silt or clay fractions of soil; other forms of heavy metals may be retained in soil by complexing with soil clays and organic matter.

The former Kesterson Wildlife Refuge is a prime example of how naturally occurring heavy metals can adversely impact the surrounding environment. Selenium occurs naturally in the bedrock and soils of the San Joaquin Valley of central California where the former Kesterson Wildlife Refuge was located. Centuries of bedrock and soil weathering caused increased concentrations of Se in the local groundwater, which was used to irrigate crops in the region. Until 1978, much of the water flowing into the marsh was fresh water. However, the regional hydrology had changed such that by 1981 much of the influent water was from subsurface agricultural drainage water that had elevated concentrations of Se (2). The Se-contaminated marsh water was consumed and Se bioaccumulated in birds and other wildlife causing embryonic mortalities and severe deformities in aquatic birds (3) and the disappearance of most fish species (4). The marsh was ultimately drained in the mid-1980s and capped with noncontaminated soil to mitigate further loss of wildlife. Numerous additional examples exist within the United States and throughout the world of contamination and adverse environmental impacts from naturally occurring heavy metals. Some additional examples are Lake Ontario from Pb and Zn, Lake Cayuga in central New York from Cd, Co, Cu, Pb, and Zn; and Fränkische Saale in Germany from various heavy metals (5). Other natural mechanisms of heavy metal introduction into and transport within the environment include volcanic eruptions and windblown dust. The rate of human-induced heavy metal introduction into the environment has increased exponentially since the Industrial Revolution of the late 1800s.

Anthropogenic sources of heavy metals in the environment generally include atmospheric emissions, land-surface application of various heavy-metal-containing materials, and seepage from damaged or compromised containment structures. Particulate-bound heavy metals are emitted to the atmosphere during various high-temperature processes, such as fossil fuel burning for electric power, smelting of nonferrous ores, melting of ferrous ores, refuse incineration, and kiln operations (6). Due to its volatility, Hg can also be emitted in to the air as a gas and distributed over large areas. This type of transport mechanism is often cited as the source of Hg for methyl mercury contamination of aquatic ecosystems. Atmospheric emissions can be transported long distances and ultimately are deposited and accumulate on the land surface or in surface waters as dry dust particles or as part of precipitation. Heavy-metal-containing residual materials left over from high-temperature industrial processes, such as ashes, and other activities, such as municipal wastewater treatment, are also anthropogenic sources due to their disposal in and discharge to the environment.

A common practice of disposing of some types of industrial by-products is spreading them on the land surface or discharging them to surface waters. These are accepted practices as long as the heavy-metal content does not exceed regulated standards. According to the United States Environmental Protection Agency's Risk Reduction Engineering Laboratory Treatability Database (7,8), industrial sources, such as fabricated metal products except machinery and transportation equipment, chemical and allied products, and electronic and electric equipment, commonly discharge wastewaters containing Cd, Cr, Cu, Ni, and Zn in the range of 10–100 mg L^{-1}. Lead and Hg are also commonly detectable in industrial discharges. Similarly, municipal sewage sludge is often land applied as an organic amendment and fertilizer for agricultural purposes. In 1997, sewage sludge from publicly owned wastewater treatment facilities in Pennsylvania that was disposed of by land spreading contained between 5 and 800 mg kg^{-1} of As, Cd, Cr, Cu Pb, Hg, Ni, Se, and Zn (9). However, the same study also demonstrated that the concentrations of Cd, Cr, Cu, Pb, Ni, and Zn in Pennsylvania sewage sludge decreased significantly between 1978 and 1997 (9).

Aside from industrial sources and sewage sludge, another significant mode for introducing heavy metals into the environment is agriculturally related sources, including fertilizers, liming materials, animal wastes, and pesticides. Heavy metals and other trace metals are important to agriculture because of their essential roles in plant and animal nutrition. Inorganic fertilizers and liming materials are rarely pure substances and often contain other additives, fillers, or impurities that may consist of heavy-metal constituents. For example, some phosphate fertilizers contain Cd and Zn (10). Similarly, heavy metals are often contained in animal wastes (manures and litters) from confined animal feeding operations, such as for dairy, swine, and poultry, due to the use of heavy-metal-containing compounds in feed additives for growth enhancement or disease suppression. Even more so than in industrial by-products or sewage sludge, animal wastes are typically land applied as an organic amendment for agronomic benefit, thus introducing any heavy metals contained in the animal wastes into the environment. For example, poultry litter typically contains notable amounts of heavy metals. On average and depending on production and management practices, poultry litter can contain between 20 and 40 $mg\,kg^{-1}$ of As, Cd, and Pb and between 370 and 660 $mg\,kg^{-1}$ of Cu, Fe, Mn, Se, and Zn (11). Poultry litter in Alabama has also been found to contain relatively high levels of Co, Ti, and V (12). Numerous pesticides used in agronomic crops also contain heavy metals, such as As, Cu, Pb, Hg, and Zn (10).

Landfill leachate is an additional significant source of concentrated heavy metals in the environment. The wide variety of domestic refuse, especially household and unregulated small generator inputs, disposed of in landfills contain and, within the appropriate chemical conditions, may release soluble metal ions (13,14). For example, Cd, Ni, Pb, and Zn can be released from batteries; Cd and Cr can be released from paint residues; and As and Hg can be released from pesticide residues (15). These potentially toxic metals comprise the solutes of leachates produced as water percolates through landfills and discharges eventually into the surrounding soil or groundwater. Landfill leachate also contains a variety of organic compounds that may interact with toxic metal ions and, hence, influence their mobility (13).

HEAVY METAL MOBILITY

The mobility of heavy metals in the environment largely depends on the degree to which the metal is soluble. In turn, heavy metal solubility is governed by the chemical characteristics of the media in which the metal exists and the chemical nature of the metal itself.

Some heavy metals can exist in multiple valence states and in multiple forms or species in soil, such as free ions or complexes. For example, As can exist as the trivalent (As^{3+}) or pentavalent (As^{5+}) ion, Cr can exist as the trivalent (Cr^{3+}) or the hexavalent (Cr^{6+}) ion, and Fe can exist as the divalent (Fe^{2+}) or trivalent (Fe^{3+}) ion. For these metals, speciation is affected by the oxidation–reduction (redox) potential of the surrounding media. The redox potential determines the direction of a redox reaction; this is a reaction in which the complete transfer of electrons from one species to another occurs (16). Depending on the redox potential, certain metal species may be soluble or insoluble, thus affecting their mobility. For example, Fe in the oxidized state (Fe^{3+}) is relatively insoluble and typically exists as an oxide precipitant on surfaces of aggregates in soil, but Fe in the reduced state (Fe^{2+}) is highly soluble and is easily removed as water percolates through the soil. Table 1 lists the dominant chemical species in soil and water and the most toxic species of selected heavy metals.

The mobility of heavy metals in soil is affected by retention and release reactions, which include precipitation–dissolution, ion exchange, and adsorption–desorption reactions (17). Retention and release reactions in turn are affected by numerous soil properties, including pH, organic matter, bulk density, and texture (18).

Precipitation–dissolution reactions are governed by the concentrations of the reactants or solutes and the chemical characteristics of the solvent. Under the proper chemical conditions, some heavy metals may react with other solutes to form precipitates. The reverse reaction is dissolution of a fraction or all of the solid into its ionic components.

Table 1. Dominant Chemical Species in Soil and Water and Most toxic Species for Selected Heavy Metals. Adapted from Table 3 (23)

	Dominant Chemical Species		
Metal	Soil	Water	Most Toxic Species
As	$AsO_4{}^{3-}$	$AsO_4{}^{3-}$, $AsO_3{}^{3-}$	$AsO_4{}^{3-}$
Cd	Cd^{2+}	Cd^{2+}	Cd^{2+}
Co	Co^{2+}	Co^{2+}	Co^{2+}
Cr	Cr^{3+}	Cr^{3+}, Cr^{6+}	Cr^{6+}
Cu	Cu^{2+}	Cu^{2+}-fulvate	Cu^{2+}
Hg	Hg^{2+}, CH_3Hg	$Hg(OH)_2{}^0$, $HgCl_2{}^0$	CH_3Hg
Mn	Mn^{2+}, Mn^{4+}	Mn^{2+}	Mn^{2+}
Ni	Ni^{2+}	Ni^{2+}	Ni^{2+}
Pb	Pb^{2+}	$Pb(OH)^+$	Pb^{2+}
Se	$HSeO_4{}^{3-}$, $SeO_4{}^{2-}$	$SeO_4{}^{2-}$	$SeO_4{}^{2-}$
Zn	Zn^{2+}	Zn^{2+}	Zn^{2+}

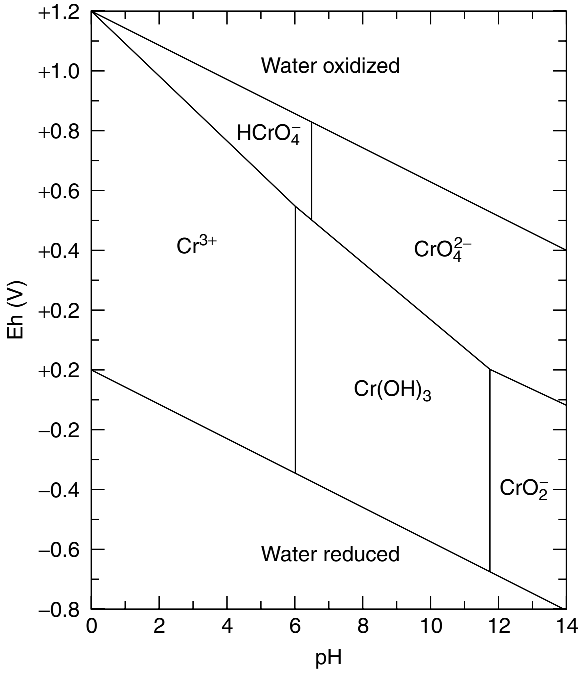

Figure 1. Eh-pH diagram for chromium (Cr) at an activity level of 10^{-6}.

Two of the most important chemical characteristics that influence precipitation–dissolution reactions and general metal speciation are pH and redox potential. The effects of pH and redox potential can best be illustrated using an Eh–pH diagram, where Eh is the reduction/oxidation potential in volts for a redox-influenced heavy metal, for example, Cr (Fig. 1). Depending on the redox conditions and pH, Cr can exist in one of five different forms, ranging from the reduced free ion (Cr^{3+}) at a $pH<6$ and Eh potential between ~ -0.3 and $+1.2$ volts to the chromate ion at a $pH>\sim 6.5$ and Eh potential between ~ -0.12 and

+0.84 volts, where Cr exists as the oxidized (Cr^{6+}) species bound with four oxygen atoms. As the Eh–pH diagram depicts for Cr activity of 10^{-6}, the pH and Eh must be high for Cr to be mobile in natural waters, where the typical pH range is 5.5 to 8.5 and the typical Eh range is 0.1 to 0.45 V (19). pH and Eh affect metal speciation and solubility, so pH and Eh also influence the ability of ionic metals to interact with the surfaces of soil particles.

Due to the presence of many clay particles that have negatively charged surfaces, most soils have some cation exchange capacity (CEC) in which cationic heavy metal species can interact by electrostatic attraction with the surface of these negatively charged clay particles. However, the type of clay minerals present is also important. For example, kaolinite, a 1:1 clay mineral, has a relatively low CEC, typically in the range of 2–5 cmol charge kg^{-1} (20). However, 2:1 clay minerals, such as those in the smectite group (e.g., montmorillonite and vermiculite), have a relatively large CEC typically in the range of 80–180 cmol charge kg^{-1} (20). Humus (organic matter) also has a relatively large CEC, typically in the range of 100–550 cmol charge kg^{-1} (20). The process of ion exchange retains metals in their ionic form in soil and reduces their mobility until the soil–solution concentration of a different cation that has a stronger affinity for the binding site increases enough to replace the cationic metal on the clay surface.

In contrast to ion exchange, adsorption is the process whereby an ionic metal species binds to the surface of a soil particle to form a complex. The degree of adsorption of a metal species depends greatly on the chemical nature of the metal species itself, but also can typically be enhanced by clay and organic matter. Organic matter sometimes acts as a chelating agent. Chelation is the binding of organic chemicals that have multiple functional groups with metals to form ring structures; the end products are organometallic complexes (20).

Many metals form inorganic and organic complexes in soil. Depending on how the metal speciates, some metal complexes, such as complexes containing As^{3+} and As^{5+}, are highly soluble in water (21). Consequently, certain metal species like inorganic As (10), organic complexes of Pb, Zn, and Cr, and oxygen complexes of Cr and Se (22) are considered relatively mobile in soil and pose a health threat when introduced into the environment.

Once in solution, either as free ions or soluble complexes, heavy metals move by mechanisms similar to those of other solutes in the soil or groundwater systems. Diffusion and advection (also called convection) are the two most common mechanisms of solute transport in soil and groundwater. Diffusion is the process by which metals move in response to a concentration gradient; therefore, diffusion occurs relatively slowly. In contrast, advection or mass flow is the process by which metal movement occurs as a result of bulk fluid (i.e., water) movement and typically occurs at a rate much faster than that of diffusion. Some heavy metals may move in the environment via air transport of metal-bound dust particles. Heavy metals may also be transported in the environment in plant tissue after root absorption and translocation to aboveground biomass.

RISK ASSESSMENT

All heavy metals are toxic to some degree when absorbed in excess of an organism's assimilative capacity (23). Once absorbed by plants, animals, or other living organisms, heavy metals can "bioaccumulate" and persist in the organism's system without dissipation. One of the most common pathways of human exposure to heavy metals is by ingestion of either contaminated drinking water or plant or animal tissue. Human exposure to heavy metals can cause a variety of adverse health effects, so acceptable limits have been set for assessing potential human health risks from exposure to heavy metals.

Maximum contaminant limits (MCLs) are standards used to define excessive exposure to metals in drinking water. An MCL is calculated to allow consuming 2 L of water per day that has a concentration up to the MCL for a period of 70 years without increasing the human risk of acquiring cancer to more than 1 in 100,000 people. In addition, the United States Environmental Protection Agency uses oral reference doses for noncarcinogenic and risk-specific doses for carcinogenic metals (24). Table 2 lists the MCLs, potential health risks to humans, common sources of drinking water contamination, and general mammalian toxicity for selected heavy metals.

MANAGEMENT AND REMEDIATION ALTERNATIVES

Prevention of environmental heavy-metal contamination is crucial because the cleanup of contaminated natural resources is difficult, time-consuming, and extremely expensive. Remediation strategies for heavy-metal contaminated soil may be employed *in situ*, where the contaminated material remains in place on-site, or *ex situ*, where the contaminated material is removed and treated off-site. Typically, the first step in the cleanup of contaminated soil is containment, which is only a temporary measure. Traditional *in situ* treatments include high temperature to produce a nonleachable material or the addition of solidifying agents to produce a cement-like material (1). More recent "advancements" in *in situ* remediation of heavy-metal-contaminated soil have included bioremediation and phytoremediation. *Ex situ* treatments most often include excavation followed by landfilling, incineration, repeated washing, incorporation into asphalt mixes, or land farming, a process by which contaminated soil is spread out and actively managed with fertilizers or plants so that degradation processes occur naturally (25).

Remediation of contaminated waters is more difficult than that for contaminated soil partly because the number of potential remediation strategies for contaminated waters is more limited. Removal of contaminated sediments from aquatic ecosystems is difficult. Dredging of contaminated sediments often exacerbates the problem by mixing the sediment with the water. A common remedial activity for groundwater is to pump contaminated water out of the saturated zone followed by treatment and disposal or reinjection into the aquifer (25). Although groundwater can be pumped out and treated, once an aquifer is contaminated, any subsequent flush of water into or through the aquifer will cause the release of more

Table 2. Limits and Potential Health Risks Associated with Selected Heavy Metals (26)

Metal	MCL[a] (mg L^{-1})	MCLG[b] (mg L^{-1})	Potential Health Effects from Exposure Above the MCL	Common Sources of Metal in Drinking Water	Mammalian Toxicity[c]
As	0.01[d]	none	Skin damage, problems with circulatory system, increased risk of contracting cancer	Erosion of natural deposits, runoff from orchards and glass and electronics-production wastes	H
Cd	0.005	0.005	Kidney damage	Corrosion of galvanized pipes, erosion of natural deposits, discharge from metal refineries, runoff from waste batteries and paints	H
Cr	0.1[e]	0.1[e]	Allergic dermatitis	Discharge from steel and pulp mills, erosion of natural deposits	H (Cr^{6+}) M (Cr^{3+})
Cu	1.3[f]	1.3	Gastrointestinal distress, liver or kidney damage	Corrosion of household plumbing systems, erosion of natural deposits	M
Hg	0.002	0.002	Kidney damage	Erosion of natural deposits, discharge from refineries and factories, runoff from landfills and croplands	H
Pb	0.015[f]	0	Delays in physical or mental development, deficits in attention span and learning abaility, kidney problems, high blood pressure	Corrosion of household plumbing systems, erosion of natural deposits	H
Se	0.05	0.05	Hair or fingernail loss, numbness in fingers or toes, circulatory problems,	Discharge from petroleum refineries, erosion of natural deposits, discharge from mines	H

[a] Maximum contaminant level (MCL) is the highest level of a contaminant that is allowed in drinking water and is an enforceable standard.

[b] Maximum contaminant level goal (MCLG) is the level of a contaminant in drinking water below which there are no known or expected health risks. MCLGs are non-enforceable public health goals.

[c] Letters represent the toxicity risk as moderate (M) or high (H); taken from table 9.1 (27).

[d] Effective as of January 23, 2006.

[e] MCL and MCLG are for total Cr.

[f] MCLs for Cu and Pb represent action levels and are based on a treatment technique to control the corrosiveness of the water.

contaminant into solution. Therefore, complete cleanup of groundwater aquifers requires a very long time.

Once the contaminated soil is remediated, several management practices may be used to help immobilize heavy metals in the soil and minimize the potential for transport and adverse environmental effects (1). The soil pH can be adjusted to 6.5 or higher, which will decrease the availability and potential mobility of cationic heavy metals. However, raising the soil pH will generally increase the availability and mobility of anionic heavy-metal complexes (1). Improving soil aeration by draining the soil will sometimes allow heavy metals to oxidize, making most heavy metals less available, except for those that form oxygen complexes (e.g., Cr) and sulfides (e.g., Hg). However, for Cr to exist as $CrO_4{}^{2-}$, requires a dissolved oxygen concentration above that which occurs naturally. Dredging often moves metal sulfides into oxidizing environments where the sulfide can potentially oxidize to sulfate, resulting in release of the metal. The addition of phosphate to a contaminated soil will help reduce the availability of cationic heavy metals due to the high binding affinity of the anionic phosphate ion for cations, but the addition of phosphate would increase the availability of anionic contaminants like As.

The use of plants to extract heavy metals from contaminated soil and water (phytoremediation) is also a very useful and cost-effective remediation strategy. There are three general mechanisms by which plants are used to stabilize or remove metals and other contaminants from soil and water. Phytoextraction is the process of cultivating plants called hyperaccumulators in metal-contaminated soil that have root systems that can absorb and translocate relatively large quantities of metals to their aboveground biomasses. After the plants have been allowed to grow for some time, the plants are harvested and incinerated or composted. There are approximately 400 hyperaccumulating plant species, but determining the appropriate species or group of plants to use depends on the heavy metal contaminant at a site (1). Rhizofiltration is used to remediate contaminated surface waters, whereby plants are initially grown hydroponically in a greenhouse using water from the contaminated site, and are then transplanted at the site. After enough time has elapsed to allow the roots to become saturated with the contaminant by a combination of adsorption and absorption, the plants, including the roots, are harvested and typically incinerated (1). Phytostabilization is a remediation process in which perennial, nonharvested, metal-tolerant plants are used to immobilize soil and water contaminants into roots and aboveground biomass by adsorption, absorption, or precipitation in the rhizosphere and reduce the mobility and bioavailability of some metals (1).

SUMMARY

Soil and water contamination by heavy metals impacts the environment negatively. Scandium, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Cd, Hg, as well as As, Se, and Pb, are generally considered heavy metals. There are numerous natural and anthropogenic sources of heavy metals in

the environment. Soil and groundwater properties, such as pH and redox potential, and the characteristics of a particular metal itself affect heavy-metal mobility in the environment. Once in the environment, excessive exposure to heavy metals poses a health risk to any living organism. A wider variety of strategies that can be used to remediate heavy-metal-contaminated soil exists than for remediating contaminated groundwater.

BIBLIOGRAPHY

1. Soil Quality Institute (SQI). (2000). *Heavy Metal Soil Contamination*. United States Department of Agriculture, Natural Resource Conservation Service, Urban Technical Note No. 3, Auburn, AL.
2. National Research Council (NRC). (1989). *Irrigation-Induced Water Quality Problems: What Can be Learned from the San Joaquin Valley Experience*. National Acad. Press, Washington, DC.
3. Ohlendorf, H.M., Hoffman, D.J., Saiki, M.K., and Aldrich, T.W. (1986). Embryonic mortality and abnormalities of aquatic birds: Apparent impacts of selenium or irrigation drainwater. *Sci. Total Environ.* **52**: 49–63.
4. Ohlendorf, H.M. (1984). The biologic system. In: *US Bureau of Reclamation and Ecological Analysis. Proceedings of a research meeting on toxicity problems at Kesterson Reservoir, California*. US Bureau of Reclamation, Mid-Pacific Region, Sacramento, CA, pp. 8–15.
5. Leland, H.V., Copenhaver, E.D., and Wilkes, D.J. (1975). Heavy metals and other trace elements. *J. Water Poll. Contr. Fed.* **47**: 1635–1656.
6. Pacyna, J.M., Münch, J., and Axenfeld, F. (1991). European inventory of trace metal emissions to the atmosphere. In: *Heavy Metals in the Environment*. J.P. Vernet (Ed.). Elsevier, New York.
7. Patterson, J.W., Minear, R.A, Gasca, E., and Petropoulou, C. (1998). Industrial discharges of metals to water. In: *Metals in Surface Waters*. H.E. Allen, A.W. Garrison, and G.W. Luther. (Eds.). Ann Arbor Press, Ann Arbor, MI.
8. United States Environmental Protection Agency (USEPA). (1993). *Risk Reduction Engineering Laboratory (RREL)*. Treatability Database, Ver. 5.0., Cincinnati, OH.
9. Stehouwer, R.C., Wolf, A.M., and Doty, W.T. (2000). Chemical monitoring of sewage sludge in Pennsylvania: Variability and application uncertainty. *J. Environ. Qual.* **29**: 1686–1695.
10. Adriano, D.C. (1986). *Trace Metals in the Terrestrial Environment*. Springer-Verlag, New York.
11. Gupta, G. and Charles, S. (1999). Trace elements in soils fertilized with poultry litter. *Poultry Sci.* **78**: 1695–1698.
12. Kpomblekou, A.K., Ankumah, R.O., and Ajwa, H.A. (2002). Trace and nontrace element contents of broiler litter. *Comm. Soil Sci. Plant Anal.* **33**: 1799–1811.
13. Gintautas, P.A., Huyck, K.A., Daniel, S.R., and Macalady, D.L. (1993). Metal-organic interactions in Subtitle D landfill leachates and associated groundwaters. In: *Metals in Groundwater*. H.E. Allen, E.M. Perdue, and D.S. Brown. (Eds.). Lewis Publishers, Chelsea, MI, pp. 275–308.
14. Pohland, F.G., Cross, W.H., and Gould, J.P. (1993). Metal speciation and mobility as influenced by landfill disposal practices. In: *Metals in Groundwater*. H.E. Allen, E.M. Perdue, and D.S. Brown (Eds.). Lewis Publishers, Chelsea, MI, pp. 411–429.
15. Jones, J.B. (2003). *Agronomic Handbook: Management of Crops, Soils, and their Fertility*. CRC Press, New York.
16. Sposito, G. (1989). *The Chemistry of Soils*. Oxford University Press, New York.
17. Amacher, M.C., Kotuby-Amacher, J., Selim, H.M., and Iskandar, I.K. (1986). Retention and release of metals by soils—evaluation of several models. *Geoderma*. **38**: 131–154.
18. Selim, H.M. and Amacher, M.C. (1997). *Reactivity and Transport of Heavy Metals in Soils*. CRC Press, New York.
19. Baas Becking, L.G.M., Kaplan, I.R., and Moore, D. (1960). Limits of the natural environment in terms of pH and oxidation-reduction potentials. *J. Geol.* **68**: 243–284.
20. Brady, N.C. and Weil, R.R. (1996). *The Nature and Properties of Soils*, 11th Edn. Prentice-Hall, Upper Saddle River, NJ.
21. Vaughn, G.T. (1993). *Investigation Report CET/LHIR148: The environmental chemistry and fate of arsenical pesticides in cattle tick dip sites and banana plantations*. CSIRO, Division of Coal and Energy Technology, Centre for Adv. Analyt. Chem., Sydney, Australia.
22. Davis, D.A., Kent, D.B., Rea, B.A., Maest, A.S., and Garabedian, S.P. (1993). Influence of redox environment and aqueous speciation on metal transport in groundwater: Preliminary results of trace injection studies. In: *Metals in Groundwater*. H.E. Allen, E.M. Perdue, and D.S. Brown. (Eds.). Lewis Publishers, Chelsea, MI, pp. 223–273.
23. Logan, T.J. and Traina, S.J. (1993). Trace metals in agricultural soils. In: *Metals in Groundwater*. H.E. Allen, E.M. Perdue, and D.S. Brown (Eds.). Lewis Publishers, Chelsea, MI, pp. 309–347.
24. Rubenstein, R. and Segal, S.A. (1993). Risk assessment of metals in groundwater. In: *Metals in Groundwater*. H.E. Allen, E.M. Perdue, and D.S. Brown. (Eds.). Lewis Publishers, Chelsea, MI, pp. 209–221.
25. National Research Council (NRC). (1997). *Innovations in Ground Water and Soil Cleanup*. National Acad. Press, Washington, DC.
26. United States Environmental Protection Agency (USEPA). (2002). *National Primary Drinking Water Sandards*. Office of Water, EPA 816-F-02-013. Washington, DC.
27. McBride, M.B. (1994). *Environmental Chemistry of Soils*. Oxford University Press, New York.

SOURCE, MOBILITY, AND REMEDIATION OF METALS

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

It is estimated that over 60% of CERCLA sites in the United States have contamination because of metals. Similarly, RCRA, general industrial, DOD, and DOE facilities also have problems with metals, which, in conjunction with a paucity of innovative approaches to deal with metal problems, means that metal contamination represents a serious environmental concern and a significant market opportunity for environmental service companies.

Metals (or other inorganics) typically become groundwater problems under the following situations:

- Activities associated with plating shops, where a wide variety of metals are present at high concentrations in forms that are soluble.
- Sites with releases of radionucleides that, because of unique health risks, can have significant impact at very low concentrations. In addition, the use of chelating and complexing agents that make these contaminants mobile in the environment is common during processing of these materials.

Metals and high levels of inorganic Total Dissolved Solids (TDS) are associated with leaks from sanitary, solid waste, and hazardous waste landfills.

High TDS impacts are also associated with salt storage areas and petroleum production activities.

Innovation in the area of hydrocarbon remediation has been significant in the last decade. However, metals may only be mobilized or immobilized, unlike hydrocarbons they cannot be degraded to less innocuous components, such as CO_2 and H_2O, for example. This limited reactivity is one of the reasons that innovative remediation technology has been slow to develop for soils or groundwater contaminated with metals, which will likely change with time, however.

Often particulate metal contamination is confined to the upper few feet of soil beneath a contaminated area. However, instances exist where metal contamination has impacted groundwater. The purpose of this article is to discuss the physical/chemical behavior of metals with a particular focus on conditions that make metals mobile and thus able to impact groundwater. Also evaluated will be remediation alternatives to metal contaminated groundwater.

The important issue with regard to metals is mobility. Specifically, under what conditions are metals mobile and under what conditions are metals immobile? Metals in the environment can take four fundamental forms: (a) raw metallic elements, (b) hydrated ionic salts, (c) covalently bonded molecules termed inorganic complexes, (d) associated with a chelating agent. Elemental metals are generally not highly soluble under normal groundwater conditions. However, naturally occurring elemental metals may cause soil samples to be deemed hazardous using a Toxic Characteristic Leaching Procedure (TCLP) test and guidelines set forth in 40 CFR 261.4. The TCLP test uses strong acids during the laboratory leaching procedure. The TCLP is an EPA SW-846 analytical method (Method 1311) that simulates sanitary landfill contaminant leaching of waste materials, which could include mobilizing naturally occurring metals in the soil sample.

The mobility of metals as hydrated ionic salts is dependent on which metallic element is participating as the positively charged ion (termed the cation) and which anion makes up the negatively charged component of the salt.

Following is a brief summary of cationic/anionic solubility relationships:

- Sodium (Na^+), Potassium (K^+), and Ammonium ($NH_4{}^+$) are cations that form salts that are all soluble.
- All metal salts of the Nitrate ($NO_3{}^-$), Nitrite ($NO_2{}^-$), Acetate ($C_2H_3O_2{}^-$), Permanganate ($MnO_4{}^-$), Perchlorate ($ClO_4{}^-$), and Chlorate ($ClO_3{}^-$) anions are soluble.
- All Chloride (Cl^-), Bromide (Br^-), and Iodide (I^-) salts are soluble except those of Lead (Pb^{2+}), Silver (Ag^+), and Mercury (Hg^{2+}).
- All Sulfate ($SO_4{}^{2-}$) salts are soluble except those of Barium (Ba2+), Strontium (Sr^{2+}), and Lead (Pb^{2+}).
- All Oxides ($O_2{}^-$), Sulfides ($S_2{}^-$), and Hydroxides (OH^-) are insoluble except those of Calcium (Ca^{2+}), Barium (Ba^{2+}), and Strontium (Sr^{2+}).
- With the exception of Sodium (Na^+), Potassium (K^+), and Ammonium ($NH_4{}^+$), all metallic salts are insoluble for anions of Carbonate ($CO_3{}^{2-}$), Phosphate ($PO_4{}^{3-}$), Sulfite ($SO_3{}^{2-}$), Borate ($BO_3{}^{3-}$), Fluoride (F^-), and Silicate ($SiO_3{}^{2-}$).

A covalently bonded inorganic molecule that contains several atoms (one of more of which are metal atoms) is termed an inorganic complex. Of particular interest and environmental concern are a class of inorganic complexes termed oxyanions, which are compounds composed of metal and oxygen atoms forming an entire molecule (rather than just an isolated metal ion) that is capable of forming a hydrated complex charged ion. These oxyanionic complexes are often soluble and, more importantly, have physical/chemical qualities that make them valuable in various industrial processes. In a modern industrial environment, metallic and metalloid oxyanions are not rare substances. The most troublesome and the most common industrial oxyanion is chromate, which contains hexavalent chromium (Cr^{+6}). The chromate molecule, in turn, forms an extremely soluble, anionically charged ion. Contributing to its high mobility in the environment is a property of the chromate ion that allows it to be soluble over the entire range of pH. From acidic to basic conditions, the chromate ion simply changes the overall negative charge it carries, all while staying soluble and not precipitating. This high degree of mobility, in conjunction with its common use, and the fact that it is a known carcinogen, makes it one of the most common problem metals found to contaminate groundwater.

The ability of chromium to form soluble oxyanionic complexes is not unique to it only. Other metals and metalloids exist that form inorganic complexes having more or less the same solubility profiles. This list includes molybdenum, vanadium, tungsten, arsenic, selenium, and tellurium. These compounds are not commonly used by our industrial society. However, arsenic and selenium can have significant groundwater impact around mining areas, irrigation complexes in the west, and, in some moderately common instances, soluble arsenic is a common component in groundwater that flows through arsenic-rich geology (such as volcanics).

Industrial chemistry often makes use of complexing agents with more than one point of attachment to a central metal atom in a complex. This type of complexing agent is termed a polydentate ligand or a chelating agent. Chelating agents form strong bonds with metals and are, in turn, extremely soluble. EDTA (ethylenediaminetetraacetic acid) is a commonly used chelating agent. Chelating agents are used in industrial chemical systems with transition metals, heavy metals, and radionucleides. If released into the groundwater, these metal-bearing chelates are extremely mobile.

It should also be remembered that, under some conditions, metals in groundwater can be mobile as colloidal-sized particles, even though the metal is in an insoluble form. Although not soluble, colloidal particles are so small they may approach within an order of magnitude the size of ionized species. As such, they are transportable in the pore spaces of some aquifer formations.

Last, it is important to understand that metals in groundwater will interact with and adsorb to components of the soil matrix. Clays, other mineral components, and carbonaceous material (especially humic and fulvic substances) can all act in this manner. However, in the case of metals, the dominant adsorptive component of the soil matrix is iron. Iron oxides (particularly of Ferric (Fe^{+3}) iron) have a very high adsorptive affinity and total capacity for metal ions and oxyanionic metal complexes. This capacity is so great that ferric hydroxide is often used in waste water treatment systems to aid in the removal of soluble metallic species such as chromate or arsenic. Knowing the total iron content of the soil matrix is paramount in understanding the fate and transport of metals in groundwater.

METALS AND GROUNDWATER PART II: *IN SITU* BIOREMEDIATION OF METALS CONTAMINATED GROUNDWATER

Significant technical developments have improved our ability to remediate groundwater contaminated with hydrocarbons. There has not been similar growth in treatment alternatives for sites contaminated with metals.

When groundwater is contaminated with metals, remediation alternatives are limited to mobilization or immobilization. Unlike hydrocarbons, biodegradation into innocuous carbon dioxide and water is not possible. Irrespective of the available of *in situ* reactions, the same metal will still ultimately be present.

The choice of mobilization or immobilization is dependent on the hydrogeologic setting of the site, the chemical properties of the metal contaminant, the geochemistry of the site soils and groundwater, and the current and future use of the contaminated groundwater. Generally, it is easier to immobilize a metal than it is to mobilize and recover it. However, an immobilized metal may still be a source of future problems if subsurface ambient conditions change.

MOBILIZATION

Metal mobilization and recovery is vulnerable to the same limitations encountered with pump and treat systems for hydrocarbons. Adsorption to components of the soil matrix can significantly increase the required duration for operation of the recovery system.

Aggressive metal mobilization (beyond that achievable with passive pump and treat) is an approach to be applied only under extreme conditions. However, given proper geochemical conditions, what may be possible is the use of complexing ligands to enhance the mobility of particularly troublesome metal contaminants (such as radionucleides). In such an application, it would be essential to maintain hydraulic control of the site. Injection wells would introduce organic complexing agents (such as gluconate) to mobilize the targeted metals. Groundwater, with the solubilized metals, could then be recovered. Once the metal removal process has been completed, conventional *in situ* bioremediation could be used to degrade the residual concentrations of the organic complexing agent. Another aggressive approach to metal mobilization is methylation. Mercury can be methylated by a number of microorganisms (methanobacterium, for example), resulting in the conversion of Hg^{+2} salts to extremely volatile methylmercury compounds. Arsenic, selenium, and tellurium can undergo similar reactions. Such volatilized methylated metals are very mobile in the subsurface. In addition, the methylated forms of these metals are extremely toxic. The unusual use of this approach would be predicated on a substantial need for total removal of the metal and the ability to exercise complete hydraulic control of groundwater and soil gas at the site.

IMMOBILIZATION

A critical component of applying metal immobilization will be the use of risk assessments to establish the safety of allowing the presence of contaminating metals in an insoluble form within the aquifer matrix.

Mechanisms for microbial mediated metal immobilization can be divided into two categories, active and passive.

Active mechanisms operate directly on the metal and include:

- precipitation;
- intracellular accumulation; and
- oxidation/reduction.

Passive mechanisms act indirectly by modifying the surrounding environment. They include:

- complexing with extracellular biological chelates;
- biosorption to cell surfaces; and
- destruction of complexing ligands.

ACTIVE IMMOBILIZATION

Precipitation

In anaerobic environments, sulfate reducing bacteria, such as Desulfovibrio and Desulfotomaculum, produce hydrogen sulfide by the reaction $SO_4{}^{2-} + 10H = H_2S + 4H_2O$. The hydrogen sulfide reacts with soluble metals forming insoluble metal sulfides. The following solubility products (in parts per billion) illustrate that point:

- Cadmium Sulfide 8.7×10^{-7};
- Copper Sulfide 8.8×10^{-14};
- Lead Sulfide 4.3×10^{-6};
- Mercury Sulfide 1.1×10^{-16}.

Intracellular Accumulation

Intracellular accumulation is a two-stage process. First, metal ions are bound passively to the surface of the bacterial cell wall by physical/chemical processes. Then, the metal ions are transferred to the interior of the cell by microbial energy systems normally associated with magnesium and potassium transport. Pseudomonas and Thiobacillus are examples of common bacterial species that are capable of this process.

Oxidation and Reduction

Many known metal redox reactions mediated by microbes exist. For example, Chromate ions (Cr^{+6}) can be reduced to insoluble trivalent chromium (Cr^{+3}) with soluble reductase enzymes generated by Escherichia coli (E. coli). Soluble uranium ions (U^{+6}) precipitate as the mineral uraninite when reduced to the U^{+4} species. This reduction reaction can be driven by reductase enzymes produced by the bacteria Desulfovibrio vulgaris.

PASSIVE IMMOBILIZATION

Extracellular Complexation

Extracellular complexation occurs from interactions of metals with extracellular polymers excreted by bacteria or from organic matter accumulated from the dead microbes.

Extracellular polymers include:

- polysaccharides,
- proteins, and
- nucleic acids.

Indigenous bacteria can be stimulated to specifically produce extracellular polysaccharides.

Another class of microbial chelating agents are siderophores, which are low-molecular-weight ligands synthesized and excreted by bacteria for capturing and supplying iron to support metabolic activity. Other metals may also complex with these ligands.

Siderophore producers include Psuedomonus, Actinomyces, and Azotobacter.

Cell Wall Binding

Three mechanisms exist for the binding of metals to bacterial cell walls:

1. Ion exchange reactions with peptidoglycan and teichoic acid (important biochemical components of the bacterial cell wall);
2. precipitation through nucleation reactions; and
3. complexation with nitrogen and oxygen ligands.

Gram positive bacteria have high adsorptive capacity, particularly Bacillus, because of the high peptidoglycan and teichoic acid content in their cell walls. In general, gram negative bacterial cell walls are low in these components and exhibit poor metal adsorption.

Destruction of Complexing Ligands

Metals associated with the plating, mining, and nuclear industries are sometimes released into the environment complexed with organic ligands such as EDTA or cyanide. These organic ligands can be biodegraded by conventional *in situ* bioremediation methods to destroy the ligand and immobilize the metal.

CONCLUSION

The use of microbiological methods for environmental remediation is not an entirely new concept. Microbial interactions with metals have long been recognized, and since the 1960s work in the field by biohydrometallurgists and biogeochemists has exploded. Given adequate hydraulic controls, this is a technology that may be easier to physically implement than *in situ* bioremediation of hydrocarbons. The success of the latter has been contingent on the ability to induce adequate mass transfer of oxygen into the impacted zone, which is often difficult. Stimulation of microbes for the manipulation of metals *in situ* is usually dependent on the mass transport of soluble salts, which is a much easier task.

METALS AND GROUNDWATER PART III: REMEDIATION OF METALS CONTAMINATED GROUNDWATER WITH ELECTROCHEMICAL PROCESSES

The remediation of groundwater contaminated with metals is constrained by the fact that metals can only be mobilized or immobilized not destroyed, as with the biodegradation of hydrocarbons. However, the use of electrochemical technology is an approach that is uniquely applicable to metals. What follows is a brief review the application of electrochemical techniques to the remediation of groundwater contaminated with metals.

Introduction to Metal Chemistry

Terminology for the technology includes electrokinetic remediation, electroreclamation, electrokinetic soil processing, and electrochemical decontamination. For purposes of this review, electrochemical remediation processes are broken into two primary areas: electromigration and electroosmosis. Electromigration is used in relatively porous formations that are capable of supporting advective groundwater flow. Electroosmosis is applied to saturated fine-grained formations such as silts or clays.

Electrochemical reactions can be used with any groundwater contaminant that has a charge such as soluble metal ions, other inorganic ions, and organic acids or phenols. Electrochemical reactions are stimulated by placing two or more electrodes in the subsurface that are connected to a direct current power source. The electrode with a positive charge is termed the cathode, and the electrode with a negative charge is formed the anode. Cations migrate toward the negative electrode and anions toward the positive electrode. Power supplies should be capable of delivering currents around one amp per square foot of cross-sectional area between the electrodes and voltage potentials of 10 to 30 volts per foot between electrodes. Estimates for the total power consumption required to treat a cubic yard of saturated soil are in the range of 50 to 250 Kilowatt Hours.

Electromigration is the process of mobilizing metal (or other) ions through an aquifer matrix without advective fluid flow. Mass transport is strictly by ion migration driven by an induced electrical field. The rate of migration for any given ion is dependent on factors that include the concentration of total ions (TDS) in the groundwater and the concentration of individual ions of concern, the charge of those ions, the electrical potential between the electrodes, the mobility of the individual ionic species, and the porosity and tortuosity of the pore spaces in the aquifer. Any groundwater will have multiple soluble ions that make up the TDS. The fraction of the total current carried by a particular ionic species in the TDS is represented by the transference number for that ion. The numerical value of the transference number is dependent on the electromagnetic gradient (the greater the gradient, the higher the mass transport rate) and concentration of the individual ionic species. The dynamics of this system are such that the application of electromigration to metal contaminated groundwater is best suited to conditions in which the metal ion makes up a major component of the TDS. Recovery efficiencies are very low if the metal ion is present at trace levels. One peripheral, but inevitable, process that occurs, and has direct impact on the physical/chemical reactions induced by electromigration, is the electrolytic decomposition of water. Water at the cathode undergoes the following reaction:

$$2H_2O + 2e^- H_2(gas) + 2OH^-$$

Water is oxidized at the anode by the following reaction:

$$2H_2O - 4e^- O_2(gas) + 4H^+$$

Hydrogen ions have an effective ionic mobility 1.8 times greater than that of the hydroxyl ion. Hydrogen ions lower the pH, and their preferential migration through a formation will contribute toward the desorption of metal species from soil particles.

The term electroosmosis refers to a technology that has been used for decades by geotechnical engineers to remove water from clays, silts, and fine sands. The effect was first described in 1809. It has been used by the construction industry to stabilize embankments, dewater foundation soils, and increase the loading capacity of pilings. Electroosmosis is primarily an electrically induced hydraulic flow rather than an ionic flow. Water flows in the same direction as electrons in the electrical field loop, from anode to cathode.

The mechanism inducing the advective flow can be described as follows:

In fine-grained soils, pore space can be visualized as a capillary tube in which the tube walls (clay soil particles) are negatively charged. Water in such a tube exists as a double layer, with an adsorbed layer tightly held directly adjacent to the pore space wall. The water in the center of the pore space and the core of the tube is less strongly held by these attractive forces. The application of an electrical gradient across the tube will cause cations in the tightly adsorbed layer adjacent to the walls of the pore space to migrate toward the cathode, taking the free water in the core of the tube (and between the adsorbed layers) along with it. In contrast, under hydraulic pumping conditions, only the free water in the core of the pore space will flow. Contaminant mass transport is induced through hydraulic flushing, in addition to some electromigration of soluble metal ions.

The finer the grain of the pore space, the greater the significance of the electroosmotic effect. In fine-grained soils, electroosmotic-induced mass transport can be two to three orders of magnitude higher than that capable of being induced hydraulically.

Potential problems in the field application of electrochemical technology include:

- High current densities and subsequent reactions at the anode can cause low pH conditions resulting in corrosion and dissolution of the electrode element.
- Precipitation of metal hydroxides or other TDS components near the cathode can cause pore space blockage.
- If contaminant and general TDS ion concentrations are not favorable, electrical costs can be unreasonably high because all ions (even undesirable ones) participate in the reactions.
- The number and spacing of electrodes can be prohibitive. Subsidence or settlement is a potential under some conditions.

Some of these issues can be addressed through process enhancements, such as using additives to depolarize the cathode and anode. Reactions induced in groundwater under an electric field are complex and will change with time. Low pH, high pH, ion mobilization, precipitation, dissolution, and desorption are all reactions that can take place within different locations in the subsurface electrical cell at different times during the operation of the system. The electrochemistry in such a system evolves with time. The successful use of this technology requires operation of the system to exploit those geochemical changes.

GENETICS OF METAL TOLERANCE AND ACCUMULATION IN HIGHER PLANTS

Rupali Datta
Dibyendu Sarkar
University of Texas at San Antonio
San Antonio, Texas

Phytoremediation of metals is rapidly developing as a cost-effective and environmentally friendly solution for cleaning up heavy-metal-contaminated sites. The most important requirements for effective phytoremediation are fast-growing, high biomass plants that uptake

and accumulate large amounts of toxic metals in their aboveground harvestable parts. A relatively small group of hyperaccumulator plants sequesters metals in their shoot tissues at high concentrations. In recent years, major scientific progress has been made in understanding the physiological mechanisms of metal uptake and transport in these plants. The majority of metal hyperaccumulators are slow growing and typically have low biomass, so bioengineering of nonaccumulators that have high biomass is essential for more effective phytoremediation. Biotechnology offers the opportunity to transfer hyperaccumulator phenotypes into fast-growing, high biomass plants that could be highly effective in phytoremediation. However, comprehensive knowledge of the genetics of metal tolerance and accumulation is essential to design transgenic plants capable of phytoremediation. After an initial focus on plants with an immediate promise of application, research is now being directed at understanding the underlying mechanisms involved in heavy metal hyperaccumulation in plants. Once the rate-limiting steps for uptake, translocation, and detoxification of metals in hyperaccumulator plants are identified, more informed construction of transgenic plants could become rather routine, which will help tremendously in more effective use of phytoremediation technology for cleaning up heavy-metal-contaminated sites.

INTRODUCTION

Mining, manufacturing, and other anthropogenic activities have contributed to extensive soil contamination during the past century, and metals are the main group of inorganic contaminants (1). Several technologies exist for remediating metal-contaminated soils (2) that can be broadly classified into five general categories: isolation, immobilization, toxicity reduction, physical separation, and extraction. Immobilization and extraction by physicochemical techniques can be expensive and are appropriate only where rapid and complete decontamination is required (3). In addition, some of these techniques such as soil washing have an adverse effect on biological activity, soil structure, and fertility. Plants that take up metals from soil offer an alternative and less expensive method of stripping metals directly from soil. The majority of plants have constitutive and adaptive mechanisms for accumulating or tolerating high contaminant concentrations in their rhizospheres. Phytoremediation, the use of such plants to clean up soils and water contaminated with pollutants, is emerging as a new tool for *in situ* metal remediation. Phytoremediation takes advantage of the fact that a living plant is a solar-driven pump that can extract and concentrate certain metals from the environment (4). This remediation method maintains the biological properties and physical structure of soil. The technique is environmentally friendly, cost-effective, visually unobtrusive, and offers the possibility of biorecovery of precious metals.

The majority of phytoremediation-related research to date has focused on hyperacccumulator plants that accumulate metals to concentrations more than 100 times higher than those in nonaccumulators (5). However, metal hyperaccumulation is a relatively rare phenomenon that occurs in only a few terrestrial plants. Only about 400 species have been identified so far as natural metal hyperaccumulators; they represent <0.2% of all angiosperms (6). Threshold values of metal concentrations are used to define metal hyperaccumulation, which varies from element to element, for example, 10,000 mg/kg dry shoot weight for Zn and Mn; 100 mg/kg for Cu, Co, Ni, As, and Se; and 100 mg/kg for Cd (7). The common traits shared by all hyperaccumulators include (1) a bioconcentration factor greater than 1, in some cases as high as 50–100; (2) a shoot-to-root ratio of metal concentration greater than 1, indicating efficient root-to-shoot transport; and (3) enhanced tolerance to metals in the medium and inside plant cells, indicating strong internal detoxification (7). A great deal of research has been conducted to elucidate the physiology and biochemistry of metal hyperaccumulation in plants, but still certain limitations must be overcome for this technology to become efficient and cost-effective on a commercial scale (8). Plants ideal for phytoremediation should be fast growing, have high biomass and extensive root systems, be easy to harvest, and tolerate and accumulate a range of metals in their harvestable parts. No single plant with all these traits has been found so far, but it is possible that high biomass nonaccumulators that are fast growing could be modified using the transgenic approach to achieve some of the properties of hyperaccumulators (9). However, this approach can be used only when the molecular mechanisms of metal uptake, tolerance, accumulation, and translocation are better understood. The primary objective of this article is to evaluate the progress made so far in understanding the physiological and molecular-genetic mechanisms that control uptake and detoxification of metals in plants.

METAL HYPERACCUMULATION IN PLANTS

Hyperaccumulation of metals by higher plants is a complex phenomenon that involves the following steps: (1) transport of metals across the plasma membrane of root cells, (2) xylem loading and translocation, and (3) detoxification and sequestration of metals at the whole plant and cellular levels (10). The term “phytoextraction” describes the use of plants to remove metals from soil by concentrating them in their harvestable parts (11). This process depends on hyperaccumulating plants that can grow naturally on soils rich in metals. The first hyperaccumulators characterized were members of the Brassicaceae and Fabaceae families. About 400 plant species have been reported so far that hyperaccumulate metals (6). Many of the Zn hyperaccumulators identified so far belong to the family Brassicaceae and are in the genus *Thlaspi.* In addition, *Cardaminopsis halleri* and *Viola calaminaria* also hyperaccumulate Zn. Several species of plants that hyperaccumulate Zn have been investigated, but only one Cd hyperaccumulator, *Thlaspi caerulescens,* has been extensively studied. The first indication of the possibility of phytoextraction of Zn was obtained from pot experiments in background and sludge-treated soils using *Thlaspi caerulescens* (12). It was estimated that a crop of *T. caerulescens* could take up 34 kg/ha of Zn and 0.16 kg/ha Cd. Recently, the Chinese brake fern *Pteris vittata* has been found able to hyperaccumulate As from soil (13). It can accumulate up to 22,630 mg As/kg in the shoot (frond) dry weight. Furthermore, the bioconcentration factor, defined as the ratio of shoot As concentration to soil As concentration, was greater than 10 for brake fern. The fern possesses three key features that are typical of metal/metalloid hyperaccumulator plants: an efficient root uptake, efficient root-to-shoot translocation, and much enhanced tolerance to As inside the plant cells (13). After the discovery of this first As hyperaccumulator, several other fern species, including *Pityrogramma calomelanos* (14), *Pteris cretica, Pteris longifolia*, and *Pteris umbrosa* (15) have recently been added to the list of As hyperaccumulators. The hyperaccumulation trait of these ferns may be exploitable in phytoremediation of As contaminated soils.

Several plant species reportedly accumulate Ni, and some of them also accumulate Co and Cu. A few species accumulate Mn and Cd (11). The major disadvantage of the continuous phytoextraction method is that majority of the plants, except for As hyperaccumulating ferns, have low biomass and slow growth rates (11).

MECHANISM OF METAL UPTAKE BY PLANTS

Transport proteins and intracellular high-affinity binding sites mediate the uptake of metals across the plasma membrane in root cells of plants. Several studies have shown that metal hyperaccumulation of Zn and Cd by *T. caerulescens* involves enhanced metal uptake by roots (15,16). Several Zn transporter genes have been cloned recently from *T. caerulescens,* which belong to the ZIP (Zn-regulated transporter/Fe-regulated transporter-like proteins) family (17). These genes, named *ZNT1* and *ZNT2*, are highly expressed in the roots of *T. caerulescens,* but their expression is not responsive to the Zn status of the plant. Through functional complementation in yeast, it was shown that *ZNT1* mediates high-affinity uptake of Zn^{2+} and low-affinity uptake of Cd^{2+} (18). Specific alterations in Zn-responsive elements, such as transcriptional activators, may play an important role in Zn hyperaccumulation in *T. caerulescens* (18). However, increased uptake of Cd by *T. caerulescens* cannot be explained by the Zn transport pathway but may be related to enhanced expression of the *IRT1* gene, which is essential for Fe uptake (10). The *IRT1* gene mediates high-affinity uptake of Cd^{2+} in *A. thaliana* (19,20).

Several classes of proteins are implicated in metal transport in plants, which include the metal (or CPx-type) ATPases that are involved in overall metal ion homeostasis and tolerance in plants, the natural resistance-associated macrophage protein (Nramp) family of proteins, and the cation diffusion facilitator (CDF) family proteins (21). CPx-type metal ATPases have been identified in a wide range of organisms and have been implicated in the transport of essential as well as potentially toxic metals, such as Cu, Zn, Cd, and Pb, across cell membranes (21). These transporters use ATP to pump a variety of charged substrates across cell membranes and are distinguished by the formation of a charged intermediate during the reaction cycle. Metal transporters have been classified as type IB and are called the CPx-ATPases because they share a common feature of a conserved intramembranous cystein-proline-cystein, cystein-proline-histidine, or cystein-proline-serine (CPx) motif, which, it is thought, functions in metal transduction. *Arabidopsis* P-type ATPase (PAA1) was the first CPx-ATPase reported in higher plants (22). Most CPx-type ATPases identified so far have been implicated in Cu or Cd transport. The physiological role of the metal ATPases in higher plants is not known. *Arabidopsis* CPx-ATPases show fairly low similarities to each other, so it is possible that they transport different substrates. They may be present in the plasma membrane and function as efflux pumps removing potentially toxic metals from the cytoplasm or may also be present at various intracellular membranes and be responsible for compartmentalization of metals, for example, sequestration in the vacuoles, Golgi, or endoplasmic reticulum. Intracellular levels of metals must be carefully controlled, so transporters represent good candidates for regulation. How they may be regulated in higher plants is not yet known, which could occur at the transcriptional level (control on initiation rates, mRNA stability, differential mRNA splicing) or at the posttranslational level (targeting, stability), which has been observed in bacteria and yeast (21).

As mentioned earlier, *Nramp* is a novel family of related proteins that has been implicated in the transport of divalent metal ions. The corresponding genes that code for these proteins are called *Nramp* genes. In higher plants, three *Nramp* homologues have been identified in rice (23). Subsequently, two *Arabidopsis* genes showing similarity to *Nramp*s were also been identified (24). Initial results suggest that *Arabidopsis Nramp* homologues encode functional metal transporters (25). Northern blot analysis indicates that the rice *Nramp* gene *OsNramp1* is expressed primarily in the roots, *OsNramp2* in the leaves, and *OsNramp3* in both tissues of rice (23). This distinct pattern of expression could mean that they are regulated differently and have distinct functions in different tissues or that they transport distinct but related ions in different parts of the plant (23).

CDF proteins have been implicated primarily in the transport of Zn, Co, and Cd in bacteria and some eukaryotes. Certain members of CDF have been implicated in Cu or Cd transport. A related Zn transporter (*ZAT1*) from *Arabidopsis* was reported by Van der Zaal et al. (26). This *ZAT1* transporter may have a role in Zn sequestration in plants. Enhanced Zn resistance was observed in transgenic plants that overexpress *ZAT1*, expressed constitutively throughout the plant. High Zn exposure of these plants led to increased Zn content in the roots. Zinc transporter (ZIP) proteins are involved in Zn and Fe uptake (27).

To enhance metal uptake, (1) the number of uptake sites could be increased, (2) the specificity of the uptake proteins could be altered, and (3) the sequestration capacity could be enhanced by increasing the number of intracellular high-affinity binding sites or the rates of transport across organelles. A comprehensive understanding of metal transport processes in plants is essential for formulating effective strategies to develop genetically engineered plants that can accumulate specific metals.

DETOXIFICATION OF METALS IN PLANTS

A general mechanism for detoxification of metals in plants is chelation of the metals by a ligand, followed by sequestration of the metal–ligand complex into the vacuole. Complexation with ligands can be extracellular, for example, the mechanism of Al tolerance by efflux of organic acids such as malate and citrate from roots. Intracellular complexation involves peptide ligands such as metallothioneins (MTs) and phytochelatins (PCs). Metallothioneins are cystein-rich polypeptides, first identified in mammalian tissues as Cd-binding peptides (11). Several MT genes and proteins have been identified in plants. MTs are gene encoded, but PCs are enzymatically synthesized. PCs have been identified in a wide variety of plant species, including angiosperms (both monocots and dicots), gymnosperms, algae (28), fungi (29,30), and marine diatoms (31), but not in animals.

PCs are a family of sulfur-rich peptides that were first identified in yeast. Most of the current knowledge of these peptides has resulted from molecular-genetic studies on yeast and *Arabidopsis* during the last few years. PCs are rapidly induced in cells and tissues exposed to a range of metal ions, such as Cd, Ni, Cu, Zn, Ag, Hg, and Pb, and anions, such as arsenate and selenite (32). PCs consist of only three amino acids, glutamine (Glu), cysteine (Cys) and glycine (Gly). They are structurally related to the tripeptide glutathione (GSH) and are enzymatically synthesized from GSH. PCs form a family of structures with increasing repetitions of the γ-Glu-Cys dipeptide followed by a terminal Gly, (γ-Glu-Cys)n-Gly, where n is generally in the range of 2–5, but could be as high as 11 (33). It was shown that GSH-deficient mutants of *Arabidopsis* are also PC-deficient and Cd-sensitive (34). GSH-dependent PC-synthase activity was identified in cultured cells of *Silene cucubalis* (35). The enzyme is active only in the presence of metal ions such as Cd, Cu, Zn, Ag, Hg, and Pb. Similar activities have been identified in *Arabidopsis* (36), tomato (37), and pea (38). Several studies have suggested that PC-synthase activity is regulated by metal ions at the level of enzyme activation. The fact that PC-synthase activity has an important role in Cd tolerance was shown in *Vigna angularis* (adzuki beans) that are hypersensitive to Cd. Cell-suspension cultures of adzuki bean did not synthesize PCs when challenged with Cd. These cells also lacked PC-synthase activity. Externally applied GSH failed to stimulate PC synthesis or confer Cd tolerance on the cells, which demonstrated the importance of PC synthesis in Cd tolerance in plants (39).

Despite the identification of PC-synthase activity more than a decade ago, the identification of a corresponding gene remained elusive until recently. PC-synthase genes were first isolated almost simultaneously by two research groups. Vatamaniuk et al. (40) identified an *Arabidopsis* cDNA, named *AtPCS1*. The expression of *AtPCS1* protein mediated an increase in Cd accumulation, pointing to a possible role in Cd chelation or sequestration. Clemens et al. (41) identified a wheat cDNA, *TaPCS1*, that increased Cd-resistance in wild-type yeast. Just like *AtPCS1*, the Cd-resistance mediated by *TaPCS1* was associated with an increase in Cd accumulation and was GSH-dependent. Both *AtPCS1* and *TaPCS1* mediated Cd tolerance are GSH-dependent and function in vacuole-deficient mutants, suggesting a cytosolic localization, and mediate *in vivo* PC biosynthesis in yeast. Further evidence of the role of PCs in metal hyperaccumulation comes from the overexpression of a bacterial glutathione synthetase, an enzyme that synthesizes GSH, by Indian mustard plants (42). These transgenic plants have increased GSH and PC concentrations and increased Cd tolerance and accumulation relative to controls. Overexpression of plant PC synthetase in transgenic yeast also increased the tolerance to and the accumulation of Cd (40). These studies show that the manipulation of GSH and PC concentrations has significant potential for increasing the accumulation of metals by plants (43).

Regulation of PC-synthase activity is likely to be the most important regulatory mechanism of the PC biosynthetic pathway. Evidence pointing to the role of posttranslational activation was obtained using plant cell cultures. PC biosynthesis reportedly occurs within minutes of exposure to Cd and is independent of de novo protein synthesis. In *Silene cucubalis* cell cultures (35), tomato (36), and *Arabidopsis* (37), PC-synthase expression is independent of metal exposure. These studies suggest that PC synthase is regulated by enzyme activation by metals. However, there are conflicting reports regarding the induction of transcript levels in response to metal exposure. PC biosynthesis probably varies among different plant species and is regulated at transcriptional and posttranscriptional levels or both, which suggests that PC-synthase activity is regulated differently in different plant species.

PC biosynthesis may also be regulated by the biosynthesis of GSH. In transgenic Indian mustard plants, increases in the expression of enzymes in GSH biosynthetic pathways led to an increase in PC biosynthesis and Cd tolerance (42,44). Wild-type Indian mustard plants respond to increased Cd exposure with increased levels of γ-glutamylcysteine synthatase (γ-*ECS*) transcript, which codes for the first enzyme in the PC biosynthetic pathway (45). Similarly, evidence also exists supporting posttranscriptional regulation of γ-ECS expression (46).

SEQUESTRATION OF METALS IN VACUOLES

Within plant cells, PC–metal complexes accumulate in the vacuole, which appears to be driven by various membrane transporters. Several classes of proteins have been implicated in metal transport, including CPx type ATPases, the Nramp family of proteins, and CDF family proteins. In yeast, PC–Cd complexes are reportedly sequestered in the vacuole. The mutant *hmt1* cannot form the HMW complex on Cd exposure. The *hmt1* gene encodes a member of the ATP-binding cassette (ABC) membrane transport protein that is located in the vacuolar membrane (47). The transport of PCs and PC–Cd complexes across the vacuolar membrane was not dependent on the proton gradient established across the membrane by vacuolar H^+-ATPase (48). When mesophyll protoplasts from tobacco plants were exposed to Cd, all Cd and PCs accumulated in the vacuole (49). Salt and Rauser (50) also identified an ATP-dependent, but proton gradient-independent activity, capable of transporting both PCs and PC–Cd complexes into tonoplast vesicles from oat roots.

GENETIC ENGINEERING OF PLANTS FOR PHYTOREMEDIATION

Effective phytoextraction of metals depends on the ability of plants to tolerate high levels of toxic metals and also on their ability to take up increased amounts of metals and translocate them to the shoots. Plants that have the ability to translocate metals from roots to shoots at high rates are more effective hyperaccumulators. In nonaccumulators, root concentrations of Zn, Cd, or Ni are ten or more times higher than shoot concentrations, but in hyperaccumulators, shoot metal concentrations can far exceed root levels (51,52). Biotechnology offers the opportunity to transfer hyperaccumulator phenotypes into fast growing, high biomass plants that could be highly effective in phytoextraction. Several recent studies have succeeded in bioengineering plants to hyperaccumulate a variety of toxic metals. Modification or overexpression of the enzymes that are involved in synthesizing glutathione and PCs might be a good approach to enhance metal tolerance and accumulation in plants. Zhu et al. (42) overexpressed the *Escherichia coli* counterparts of γ-ECS and glutathione synthetase in Indian mustard plants that accumulate more Cd than wild-type plants. Rugh et al. (53) modified yellow poplar trees with two bacterial genes, *merA* and *merB*, to detoxify methylmercury from contaminated soil. In transformed plants, *merB* catalyzes the release of Hg^{2+} from methylmercury, which is then converted to Hg^0 by *merA*. Elemental mercury is less toxic and more volatile than the mercuric ion and is released into the atmosphere. Pilon-Smits et al. (54) overexpressed the ATP-sulfurylase (*APS*) gene in Indian mustard. The transgenic plants had fourfold higher APS activity and accumulated three times more Se than wild-type plants. Recently, Dhankar et al. (55) reported a genetics-based strategy to remediate As from contaminated soils. They overexpressed two bacterial genes in *Arabidopsis*. One was the *E. coli arsC* gene encoding arsenate reductase that reduces arsenate to arsenite coupled to a light-induced soybean *rubisco* promoter. The second gene was the *E. coli* γ*-ECS* coupled to a strong constitutive actin promoter. The AsrC protein, expressed strongly in stem and leaves, catalyzes the reduction of arsenate to arsenite, whereas γ-ECS, which is the first enzyme in the PC-biosynthetic pathway, increases the pool of PCs in the plant. The transgenic plants that expressed both AsrC and γ-ECS proteins showed substantially greater As tolerance when grown on As, these plants accumulated 4- to 17-fold greater fresh shoot weight and accumulated two- to threefold more As than wild-type plants.

Most metal hyperaccumulators are slow growing and have low biomass, so bioengineering of nonaccumulators is essential for more effective phytoremediation. Conventional breeding approaches have also been proposed to improve plants for metal extraction. However, the success of this approach is doubtful due to sexual incompatibility between parent lines (56). Biotechnology has the potential of overcoming this limitation. However, a comprehensive knowledge of the genetic basis for hyperaccumulation is essential for effective use of biotechnology to design transgenic plants capable of efficient phytoremediation.

SUMMARY AND FUTURE DIRECTIONS

Phytoremediation as a cleanup process for metal contaminants is very attractive because it is effective, relatively inexpensive, and environmentally friendly. The majority of phytoremediation-related research to date has focused on hyperacccumulator plants that accumulate metals to concentrations more than 100 times higher than those in nonaccumulators. However, most of the known hyperaccumulators are slow growing and have low biomass, which are not ideal traits for optimal metal uptake. Plants ideal for phytoremediation should be fast growing, have high biomass, an extensive root system, be easy to harvest, and tolerate and accumulate a range of metals in their harvestable parts. No single plant that has all these traits have been found so far, but it is possible that nonaccumulators that are fast growing are have high biomass could be modified using the transgenic approach to achieve some of the properties of the hyperaccumulators. However, this approach can be used only when the molecular mechanisms of metal uptake, tolerance, accumulation, and translocation are better understood. Advances have been made in understanding root uptake mechanisms and the identification of transporters that sequester metals in the vacuoles and lead to hyperaccumulation of metals. Once taken up, the general mechanism for detoxification of metals in plants is chelation by a ligand, followed by sequestration of the metal–ligand complex into the vacuole. A lot of work has been done to understand metal tolerance and detoxification, but many questions regarding the genes involved in the processes controlling transport, sequestration, and detoxification of metals still remain. Once the rate-limiting steps for uptake, translocation, and detoxification of metals in hyperaccumulator plants are identified, more informed construction of transgenic plants will help tremendously in improving phytoremediation technology, paving the way for wider application.

BIBLIOGRAPHY

1. Adriano, D.C. (1992). Bioengineering of trace metals. In: *Advances in Trace Substances Research*. Lewis, Boca Raton, FL.
2. Evanko, C.R. and Dzombak, D.A. (1997). *Remediation of metals-contaminated soil and groundwater*. Technology Evaluation Report Oct. 1997, Groundwater Remediation Technology Analysis Center.
3. BIO-WISE. (2000). *Contaminated Land Remediation: A Review of Biological Technology*. D.T.I., London.
4. Raskin, I., Smith, R.D., and Salt, D.E. (1997). Phytoremediation of metals: using plants to remove pollutants from the environment. *Curr. Opin. Biotechnol.* **8**: 221–226.
5. Raskin, I. and Ensley, B.D. (2000). *Phytoremediation of Toxic Metals: Using Plants to Clean Up the Environment*. John Wiley & Sons, New York.

6. Baker, A.J.M., McGrath. S.P., Reeves, R.D., and Smith, J.A.C. (2000). Metal hyperaccumulator plants: A review of the ecology and physiology of a biochemical resource for phytoremediation of metal polluted soils. In: *Phytoremediation of Contaminated Soil and Water*. N. Terry and G. Banuelos (Eds.). Lewis, Boca Raton, FL, pp. 85–107.
7. McGrath, S.P. and Zhao, F.J. (2003). Phytoextraction of metals and metalloids from contaminated soils. *Curr. Opin. Biotechnol.* **14**: 1–6.
8. Khan, A.G., Kuek, C., Chaudhry, T.M., Khoo, C.S., and Hayes, W.J. (2000). Role of plants, mycorrhizae and phytochelators in heavy metal contaminated land remediation. *Chemosphere* **41**: 197–207.
9. Clemens, S., Palmgren, M.G., and Kramer, U. (2002). A long way ahead: Understanding and engineering plant metal accumulation. *Trends Plant Sci.* **7**: 309–315.
10. Lombi, E. et al. (2002). Influence of iron status on calcium and zinc uptake by different ecotypes of the hyperaccumulator *Thlaspi caerulescens*. *Plant Physiol.* **128**: 1359–1367.
11. Salt, D.E., Smith, R.D., and Raskin, I. (1998). Phytoremediation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**: 643–668.
12. Baker, A.J.M., Reeves, R.D., and McGrath, S.P. (1991). *In situ* decontamination of metal polluted soils using crops of metal-accumulating plants—a feasibility study. In: *In Situ Bioreclamation*. R.E. Hinchee and R.F. Olfenbuttel (Eds.). Butterworth-Heinemann, Stoneham, MA, pp. 539–544.
13. Ma, L.Q. et al. (2001). A fern that hyperaccumulates arsenic: A hardy, versatile, fast-growing plant helps to remove arsenic from contaminated soils. *Nature* **409**: 579.
14. Francesconi, K., Visoottiviseth, P., Sridokchan, W., and Goessler, W. (2002). Arsenic species in an arsenic hyperaccumulating fern, *Pityrogramma calomelanos*: A potential phytoremediator of arsenic-contaminated soils. *Sci. Total Environ.* **284**: 27–35.
15. Zhao, F.J., Hamon, R.E., Lombi, E., McLaughlin, M.J., and McGrath, S.P. (2002). Characteristics of cadmium uptake in two contrasting ecotypes of the hyperaccumulator, *Thlaspi caerulescence*. *J. Exp. Bot.* **53**: 535–543.
16. Lasat, M.M., Baker, A.J.M., and Kochian, L.V. (1996). Physiological characterization of root Zn^{2+} absorption and translocation to shoots in Zn hyperaccumulator and nonaccumulator species of *Thlaspi*. *Plant Physiol.* **112**: 1715–1722.
17. Maser, P. et al. (2001). Phylogenetic relationships within cation transporter families of *Arabidopsis*. *Plant Physiol.* **126**: 1646–1667.
18. Pence, N.S. et al. (2000). The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator, *Thlaspi caerulescence*. *Proc. Natl. Acad. Sci. USA.* **97**: 4956–4960.
19. Vert, G. et al. (2002). *IRT1*, an *Arabidopsis* transporter essential for iron uptake from the soil and for plant growth. *Plant Cell* **14**: 1223–1233.
20. Connolly, E.L., Fett, J.P., and Guerinot, M.L. (2002). Expression of the IRT1 metal transporter is controlled by metals at the levels of transcript and protein accumulation. *Plant Cell* **14**: 1347–1357.
21. Williams, L.E., Pittman, J.K., and Hall, J.L. (2000). Emerging mechanisms for heavy metal transport in plants. *Biochim. Biophys. Acta* **1465**: 104–126.
22. Tabata, K., Kashiwagi, S., Mori, H., Ueguchi, C., and Mizuno, T. (1997). Cloning of a cDNA encoding a putative metal-transporting P-type ATPase from *Arabidopsis thaliana*. *Biochim. Biophys. Acta* **1326**: 1–6.
23. Belouchi, A., Kwan, T., and Gros, P. (1997). Cloning and characterization of the OsNramp family from *Oryza sativa*, a new family of membrane proteins possibly implicated in the transport of metal ions. *Plant Mol. Biol.* **33**: 1085–1092.
24. Alonso, J.M., Hirayama, T., Roman, G., Nourizadeh, S., and Ecker, J.R. (1999). *EIN2*, a bifunctional transducer of ethylene and stress responses in *Arabidopsis*. *Science* **284**: 2148–2152.
25. Thomine, S., Clemens, S., Kim, E.J., Neumann, D., Wang, R., Crawford, N.M., and Schroeder, J.I. (1999). *Molecular Analysis of Metal Transport and Metal Tolerance in Plants*. *Plant Biol. 1999*, Meeting abstract No. 13004. Available: http://abstracts.aspb.org/aspp1999/public/cgi-bin/showabs.cgi?0958+M03.
26. Van der Zaal, B.J. et al. (1999). Over-expression of a novel *Arabidopsis* gene related to putative zinc-transporter genes from animals can lead to enhanced zinc-resistance and accumulation. *Plant Physiol.* **119**: 1047–1055.
27. Guerinot, M.L. and Eide, D. (1999). Zeroing in on zinc uptake in yeast and plants. *Curr. Opin. Plant Biol.* **2**: 244–249.
28. Gekeler, W., Grill, E., Winacker, E.-L., and Zenk, M.H. (1989). Survey of the plant kingdom for the ability to bind heavy metals through phytochelatins. *Z. Naturforsch. C.* **44**: 170–178.
29. Mehera, R.K., Tarbet, E.B., Gray, W.R., and Winge, D.R. (1988). Metal-specific synthesis of two metallothioneins and -glutamyl peptides in *Candida glabrata*. *Proc. Natl. Acad. Sci. USA* **85**: 8815–8819.
30. Kneer, R., Kutchan, T.M., Hochberger, A., and Zenk, M.H. (1992). *Saccharomyces cerevisiae* and *Neurospora crassa* contain heavy metal sequestering phytochelatin. *Arch. Microbiol.* **157**: 305–310.
31. Morelli, E. and Pratesi, E. (1997). Production of phytochelatins in the marine diatom *Phaeodactylum tricornutum* in response to copper and cadmium exposure. *Bull. Environ. Contam. Toxicol.* **59**: 657–664.
32. Rauser, W.E. (1995). Phytochelatins and related peptides. Structure, biosynthesis and function. *Plant Physiol.* **109**: 1141–1149.
33. Cobbett, C.S. (2000). Phytochelatin biosynthesis and function in metal detoxification. *Curr. Opin. Plant Biol.* **3**: 211–216.
34. Cobbett, C.S., May, M.J., Howden, R., and Rolls, B. (1998). The glutathione deficient cadmium-sensitive mutant *cad2-1* of *Arabidopsis thaliana* is deficient in γ-glutamylcysteine synthatase. *Plant J.* **16**: 73–78.
35. Grill, E., Loffler, S., Winnacker, E.L., and Zenk, M.H. (1989). Phytochelatins, the heavy-metal-binding peptides of plants, are synthesized from glutathione by a specific γ-glutamylcysteine dipeptidyl transpeptidase (phytochelatin synthase). *Proc. Natl. Acad. Sci. USA* **86**: 6838–6842.
36. Chen, J., Huang, J.W., Caster, T., and Cunningham, S.D. (1997). *Arabidopsis* as a model system for studying lead accumulation and tolerance in plants. In: *Phytoremediation of Soil and Water Contaminants*. K.L. Kruger et al. (Eds.). American Chemical Society, Washington, DC, pp. 264–273.
37. Howden, R., Goldsbrough, P.B., Anderson, C.R., and Cobbet, C.S. (1995). Cadmium-sensitive *Cad1* mutants of *Arabidopsis* are phytochelatin deficient. *Plant Physiol.* **107**: 1059–1066.
38. Klapheck, S., Schlunz, S., and Bergman, L. (1995). Synthesis of phytochelatins and homophytochelatins in *Pisum sativum* L. *Plant Physiol.* **107**: 515–521.
39. Inouhe, M. et al. (2000). Adzuki bean cells are hypersensitive to cadmium and do not synthesize phytochelatins. *Plant Physiol.* **23**: 1029–1036.

40. Vatamaniuk, O.K., Mari, S., Lu, Y.-P., and Rea, P.A. (1999). AtPCS1, a phytochelatin synthase from *Arabidopsis*: Isolation and *in vitro* reconstitution. *Proc. Natl. Acad. Sci. USA* **96**: 7110–7115.
41. Clemens, S., Kim, E.J., Neuman, D., and Schroeder, J.I. (1999). Tolerance to toxic metals by a gene family of phytochelatin synthases from plants and yeast. *EMBO J.* **18**: 3325–3333.
42. Zhu, Y.L., Pilon-Smits, E.A.H., Jouanin, L., and Terry, N. (1999). Over-expression of glutathione synthase in Indian mustard enhances cadmium accumulation and tolerance. *Plant Physiol.* **119**: 73–79.
43. Meagher, R.B. (2000). Phytoremediation of toxic elemental and organic pollutants. *Curr. Opin. Plant Biol.* **3**: 153–162.
44. Yong, L.Z. et al. (1999). Cadmium tolerance and accumulation in Indian mustard is enhanced by over-expressing γ-glutamylcysteine synthetase. *Plant Physiol.* **121**: 1169–1179.
45. Schafer, H.J., Haag-Kerwer, A., and Rausch, T. (1998). cDNA cloning and expression analysis of genes encoding GSH synthesis in roots of the heavy metal accumulator *Brassica juncea* L.: Evidence of Cd-induction of a putative mitochondrial gamma-glutamylcysteine synthetase isoform. *Plant Mol. Biol.* **37**: 87–97.
46. May, M.J., Vernoux, T., Sanchez-Fernandez, R., van-Montagu, M., and Inze, D. (1998). Evidence for post-transcriptional activation of γ-glutamyl-cysteine synthetase during plant stress responses. *Proc. Natl. Acad. Sci. USA* **95**: 12049–12054.
47. Ortiz, D.F. et al. (1992). Heavy metal tolerance in fission yeast requires an ATP-binding cassette type vacuolar membrane transporter. *EMBO J.* **11**: 3491–3499.
48. Ortiz, D.F., Ruscitti, T., McCue, K.F., and Ow, D.V. (1995). Transport of metal-binding peptides by HMT1, a fission yeast ABC-type B vacuolar membrane protein. *J. Biol. Chem.* **270**: 4721–4728.
49. Vogeli-Lange, R. and Wagner, G.J. (1990). Subcellular localization of cadmium and cadmium-binding peptides in tobacco leaves: Implications of a transport function for cadmium-binding peptides. *Plant Physiol.* **92**: 1086–1093.
50. Salt, D.E. and Rauser, W.E. (1995). MgATP-dependent transport of phytochelatins across the tonoplast of oat roots. *Plant Physiol.* **107**: 1293–1301.
51. Brown, S.L., Angle, J.S., Chaney, R.L., and Baker, A.J.M. (1995). Zinc and cadmium uptake by *Thlaspi caerulescens* and *Silene vulgaris* grown on sludge-amended soils in relation to total soil metals and soil pH. *Environ. Sci. Technol.* **29**: 1581–1585.
52. Kramer, U., Cotter-Howells, J.D., Charnock, J.M., Baker, A.J.M., and Smith, J.A.C. (1996). Free histidine as a metal chelator in plants that accumulate nickel. *Nature* **379**: 635–638.
53. Rugh, C.L., Senecoft, J.F., Meagher, R.B., and Merkle, S.A. (1998). Development of transgenic yellow poplar for mercury phytoremediation. *Nat. Biotechnol.* **33**: 616–621.
54. Pilon-Smits, E.A.H. et al. (1999). Over-expression of ATP sulfurylase in Indian mustard leads to increased selenate uptake, reduction and tolerance. *Plant Physiol.* **119**: 123–132.
55. Dhankar, O.P. and Meagher, R.B. et al. (2002). Engineering tolerance and hyperaccumulation of arsenic in plants by combining arsenate reductase and γ-glutamylcysteine synthetase expression. *Nat. Biotechnol.* **20**: 1140–1145.
56. Lasat, M.M. (2002). Phytoextraction of toxic metals: A review of biological mechanisms. *J. Environ. Qual.* **31**: 109–120.

IN SITU GROUNDWATER REMEDIATION FOR HEAVY METAL CONTAMINATION

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

THE METALS PROBLEM

Highly toxic and highly soluble heavy metals are contained in numerous waste streams, including those from power, chemical, electronics, and general manufacturing plants, and mining facilities. Heavy metals can create significant health risks to humans.

HISTORY

Arsenic, lead, copper, zinc, cadmium, molybdenum, uranium, and cyanide have been used in a variety of industrial activities such as plating and circuit board manufacturing, and occur in mining and power plant waste streams and from other processes.

Chromium is an example of one metal used by various industries for different purposes. Chromium is used in wood preservation. In that process, copper-chrome-arsenic wood preservatives are used extensively by the timber industry for pressure-treated or green wood. Chromium is also used as a nontarnishing surface coating in the metal plating industry. For making kiln bricks, chromium, due to its large capacity to withstand high temperatures, is used in the refractory industry. Finally, trivalent chromium salts are used widely as a tanning agent transforming animal hides and skins into leather. Chromium is just one heavy metal but illustrates the complexity and variability of industrial uses. With each use comes a series of associated industrial processes, waste streams, recycling programs, fate and transport of the metal in the environment, toxicity to humans, and the various oxidation states of the metal and other factors. All these factors greatly affect the cleanup process.

OVERVIEW OF METAL TREATMENT

Metal remediation has been documented. *In situ* methods using natural systems and enhanced bioremediation have shown promise (1). Various *in situ* treatment technologies, including metal remediation, are described in Nyer et al. (2). Environmental restoration of metal-contaminated soils is described in Ikandar (3). Mining and processing facilities are a source of metal contamination. Remediation of mining wastes is described in Davis et al. (4). Figure 1 shows a conceptual model of metal contaminants moving through the soil and into groundwater.

The success of *in situ* metal remediation requires adequate site characterization of both soil and groundwater zones, good understanding of the valence states of the associated metals and their likely source location, and complete chemical characterization. Chemical analyses include dissolved oxygen (DO), E_h, soil and groundwater

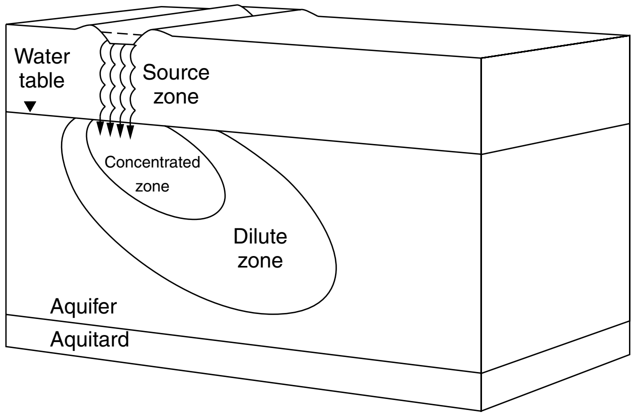

Figure 1. Conceptual model of geochemical zones in a metal contaminant plume (5,6).

pH, total organic carbon (TOC), dissolved organic carbon (DOC), particulate organic carbon (POC), cation exchange capacity (CEC), alkalinity, temperature, specific metals, and other analytes. Bench testing in the laboratory using soil and groundwater from the site allows for optimizing the chemistry in the laboratory prior to a field pilot test and an ultimate full-scale remediation program. Chemical bench tests can range from simple beaker tests where only groundwater is mixed with the reactants to soil column tests. For microbiological testing, heterotrophic count and anaerobic microbes can be evaluated. In addition, microbial bench tests can be performed with soil or groundwater, and microbial colonies can be cultured and evaluated for growth potential in various pH conditions with various bioenhancement chemicals, typically carbon sources such as molasses, cheese whey, sucrose, corn syrup, lactic acid, milk, or propane.

IN SITU TREATMENT TECHNOLOGY

Remediation of heavy metals such as lead, chromium, arsenic, and others in soil and groundwater frequently uses reduction reactions to lower the solubility of an oxidized metal in the higher of two or three valence states to a lower energy and less toxic state. Ultimately, the metal is precipitated as a sulfide or in another immobile and low toxicity form.

The solubilities of arsenic, lead, copper, zinc, cadmium, molybdenum, uranium, and cyanide are pH dependent. Recently, sulfur-based metals treatment technologies have been the focus of an increasing number of research studies and commercial applications for treating metal contamination in soil and groundwater.

Reduction Chemistry

In situ metal remediation by reduction can be either chemical or biological. For chemical reactions, sulfur compounds have been used successfully in a variety of settings. Precipitation of metals in reducing environments as sulfides, hydroxides, and carbonates has been documented as a remediation method. Even zero-valent iron has been used to reduce hexavalent chromium in groundwater through reactive permeable barrier walls, as part of a funnel and gate system. Reactive barriers for *in situ* metal treatment are documented.

Various sulfur-based reductants are available, and the more common ones include metabisulfite, ferrous sulfate, calcium polysulfide, and sodium sulfide. Selected reactions are described below:

Reductant. Metabisulfite ($S_2O_6{}^{2-}$).
In the presence of excess sulfite, Cr(VI) is reduced (7):

$$6H^+ + 2HCrO_4{}^- + 3HSO_3{}^-(\text{excess}) \Longrightarrow 2Cr^{3+} + 2SO_4{}^{2-} + S_2O_6{}^{2-} = 6H_2O$$

In the presence of excess Cr(VI), reduction to Cr(III) by sulfite is performed by this reaction (7):

$$5H^+ + 2HCrO_4{}^-(\text{excess}) + 3HSO_3{}^- \Longrightarrow 2Cr^{3+} + 3SO_4{}^{2-} + 5H_2O$$

Reductant. Ferrous sulfate ($FeSO_4$)
Dissolved Cr(VI) can be precipitated as $Cr(OH)_3$ by injecting ferrous sulfate (1).

$$\text{Acidic conditions: } Cr_2O_7{}^{2-} + 6Fe^{2+} + 14H^+ \Longrightarrow 2Cr^{3+} + 6Fe^{3+} + 7H_2O$$

Neutral or alkaline conditions:

$$CrO_4{}^{2-} + 3Fe^{2+} + 4H_2O \Longrightarrow Cr^{3+} + 3Fe^{3+} + 8OH^-$$

Both Cr(III) and Fe(III) ions are highly insoluble in natural groundwater, and these ions precipitate out as hydroxides as follows (1):

$$Fe^{3+} + 3OH^- \Longrightarrow Fe(OH)_3 \text{ (precip)}$$
$$Cr^{3+} + 3OH^- \Longrightarrow Cr(OH)_3 \text{ (precip)}$$

Reductant. Calcium polysulfide (CaS_5)

$$2CrO_4{}^{2-} + 3CaS_5 + 10H^+ \Longrightarrow 2Cr(OH)_3 + 15S^{\circ} + 3Ca^{2+} + 2H_2O$$

A reductant, such as calcium polysulfide (CaS_4), precipitates the highly soluble metal as a less soluble and nontoxic sulfide (Fig. 2). Metal hydroxides change solubility with changes in pH. Metal sulfides remain insoluble within a pH range of about 5 to 9. Calcium polysufide has a pH of 11.3 to 11.5, a specific gravity of 1.273, and is deep orange-red in solution. The excess calcium precipitates as calcium carbonate, gypsum, or calcium sulfate. Calcium polysulfide is soluble in water and comes as 29% concentration of calcium polysulfide. The extent of metals precipitation is a function of pH.

Delivery System

To deliver sulfur-based chemical reductants into the subsurface, an *in situ* delivery method uses close spacing of high-pressure injection ports, allowing the treatment

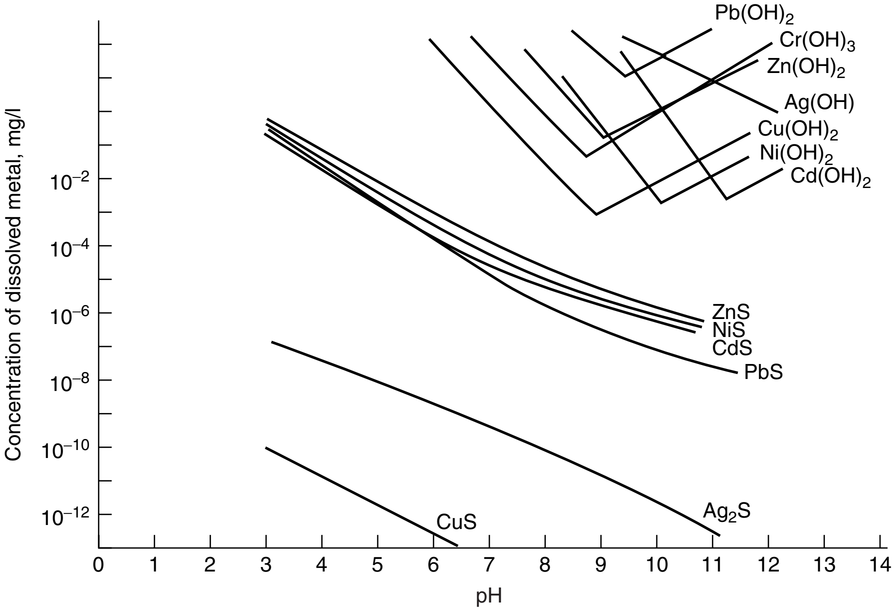

Figure 2. Solubility of metal hydroxides and sulfides as a function of pH (8).

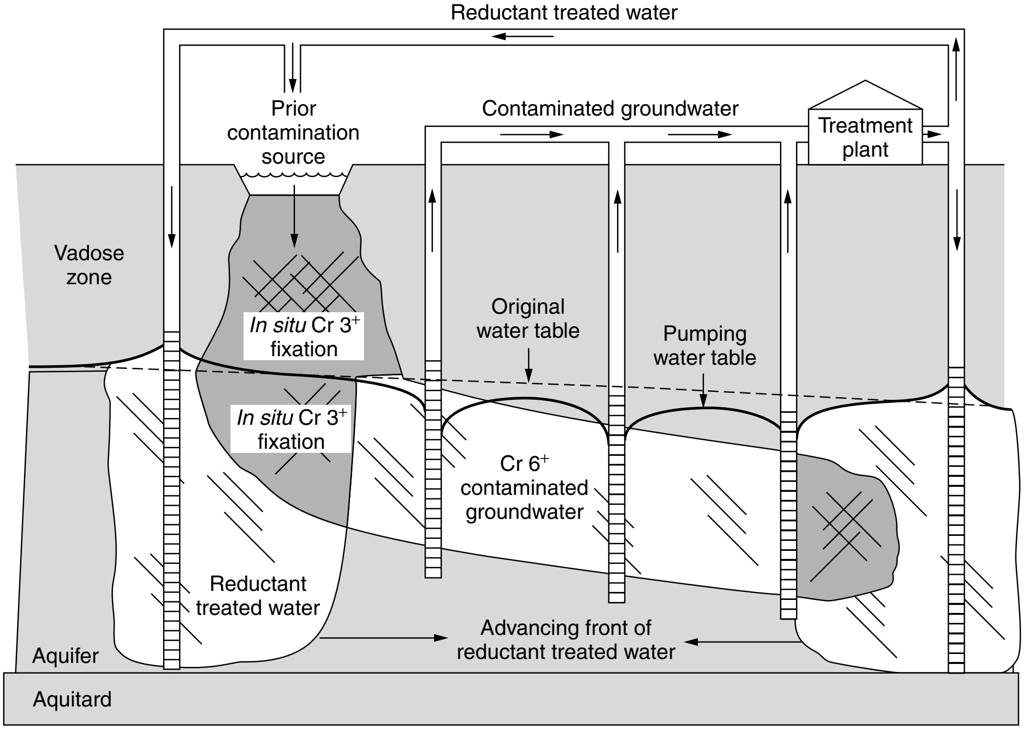

Source: Rouse, 1997

Figure 3. Schematic of an *in situ* chromium remediation process (5,9).

chemicals to contact the contaminant fully. The delivery of the reductant chemicals can be through trenches or wells. A schematic of an *in situ* remediation process for chromium is shown in Fig. 3 (5,6).

In addition, the direct push method allows placing injection ports as well as a specialized lance system to deliver treatment chemicals (10).

Besides sulfur-based reductants, lime [$Ca(OH)_2$] and other alkaline solutions such as potash (KOH) are used for chemical reduction and hydroxide precipitation (Table 1).

BIOLOGICAL METHODS OF *IN SITU* METAL REMEDIATION

Plant-based *in situ* remediation of metals, generically called phytoremediation, can be accomplished using a variety of plants that uptake the metals. An overview

Table 1. ***In situ*** **Metal Remediation**

Metal	Arsenic	Lead	Copper	Zinc
Treatment notes:	Acid medium only	Wide range (pH: 4–9)	Close to neutral (optimal pH: 5–7)	Wide range (pH: 4–9)
Metal	Cadmium	Molybdenum	Uranium	Cyanide
Treatment notes:	Wide range (pH: 4–9)	Wide range (pH: 4–9)	Wide range (pH: 4–9)	Chemical conversion produces thiocyanate[a]

[a] Thiocyanate is treated with lime, producing calcium carbonate, gypsum and ammonia.

Chromium (Cr VI) can be treated with calcium polysulfide, and Cr (VI) is reduced to Cr (III), which is then precipitated as chromium hydroxide.

of phytoremediation and other subcategories is included in Lehr (11).

Microbe-induced reduction of metals has been documented (1). Using chromium as an example, microorganisms may reduce chromium [Cr(VI)] under anaerobic conditions to trivalent chromium [Cr(III)] by exuding reducing agents or by lowering the pH. Other methods can be used to create anaerobic conditions for enhanced bioremediation and consequently reduction of metals. These enhancements use the injection of a carbon source.

PAST REMEDIATION OPTIONS

In the past, conventional remediation of soil and groundwater impacted by heavy metals has relied on soil excavation, which was expensive and disruptive. In addition, moving the soil only moved the problem, without treating the soil or reducing the long-term liability. For groundwater, pump and treat remediation relied on pumps to remove groundwater from the aquifer through a series of extraction wells or trenches. The extracted water was then treated above ground or disposed of off site. Pump and treat methods fail to address the source of the contamination in the vadose zone. Although the construction of passive permeable treatment walls containing zero-valent iron filings can reduce some metals to less toxic varieties, the passive barriers are expensive and do not treat source areas.

The chemical and biological reducing methods of remediating metals *in situ* are technically effective in significantly lowering metals concentrations in 1 to 5 years at reasonable cost. Delivery systems for the chemical reductants and bioenhancements have been developed and used at a variety of sites. The keys to success in most of these projects involve accurate subsurface assessments prior to field remediation, properly designed bench tests to optimize chemistry in the laboratory, realistic regulatory targets for site closure, and a well-designed remediation delivery system.

BIBLIOGRAPHY

1. Suthersan, S.S. (2002). *Natural and Enhanced Remediation Systems*. Lewis, Boca Raton, FL.
2. Nyer, E.K., Palmer, P.L., Carman, E.P., Boettcher, G., Bedessem, J.M., Lenzo, F., Crossman, T.L., Rorech, G.J., and Kidd, D.F. (2001). *In Situ Treatment Technology*. 2nd Edn. CRC Press, Lewis, Boca Raton, FL, p. 536.
3. Ikandar, I.K. (2001). *Environmental Restoration of Metals-Contaminated Soils*. I.K. Ikandar (Ed.). CRC Press, Lewis, Boca Raton, FL, p. 304.
4. Davis, A., Moomaw, C., Fennemore, G., and Buffington, R. (2002). Innovative strategies in remediating mining wastes. In: *Handbook of Complex Environmental Remediation Problems*. J. Lehr (Ed.). McGraw-Hill, New York, pp. 7.1–7.62.
5. Rouse, J.V., Leahy, M.C., and Brown, R.A. (1996). A geochemical way to keep metals at bay. *Environ. Eng. World* **May–June**.
6. Sabatini, D.A., Knox, R.C., Tucker, E.E., and Puls, R.W. (1997). *Innovative Measures for Subsurface Chromium Remediation: Source Zone, Concentrated Plume, and Dilute Plume*. U.S. Environmental Protection Agency, Environmental Research Brief, Washington, DC.
7. Palmer, C.D. and Wittbrodt, P.R. (1991). Processes affecting the remediation of chromium-contaminated sites. *Environ. Health Perspect.* **92**: 25–40.
8. United States Environmental Protection Agency. (1981). *US EPA Chart Comparing Sulfide and Hydroxide Solubilities*. EPA publication, Washington, DC.
9. Rouse, J.V. (1997). Natural and Enhanced Attenuation of CCA Components in Soil and Groundwater. Paper presentation for the *93rd Annu. Meet. Am. Wood-Preservers Assoc.* April pp. 27–29, Washington, DC.
10. Jacobs, J. (2001). In-Situ Liquid Delivery Systems for Chemical Oxidation, Bioremediation and Metals Stabilization, Association for Environmental Health and Sciences, *11th Annu. West Coast Conf. Contaminated Soils, Sediments Water*, March 21, 2001, San Diego, CA.
11. Lehr, J. (2000). Phytoremediation: A promising path to the elimination of heavy metals and radionuclides in groundwater and soil. In: *Standard Handbook of Environmental Science, Health and Technology*. J. Lehr (Ed.). McGraw-Hill, New York, pp. 14.26–14.29.

METHANE IN GROUNDWATER

Nitish Priyadarshi
Ranchi University
Ranchi, Jharkhand, India

Methane is a colorless, odorless, and tasteless gas, which is produced by biological decay of organic materials or by high temperatures acting on organic materials in unconsolidated and geologically young deposits. These materials include coals, organic rich shales, landfill materials, compost piles, and other accumulations of organic materials both above and underground. For example, a buried oak log, only 1 ft in diameter and 5 ft long, can generate enough methane when it decomposes underground to form an explosive mixture in 400 to 1200 m^3 of air (1).

Methane is produced by biological processes where oxygen is excluded (anaerobic conditions). Methane can be oxidized in water that contains dissolved oxygen, so the amount that reaches the surface depends on the depth of water. Its rate of production generally increases with temperature (2). It is produced only in a strongly reducing medium ($E_h < 100$ mV). Anaerobes that produce methane are called methanogens. Molecular oxygen is very toxic to them. Methane can be produced by fermenting a few simple organic compounds such as acetate, formate, methanol, or methylamines. Molecular hydrogen can also be used to reduce inorganic carbonate to methane (3). During methanogenesis, carbon dioxide is used as an electron acceptor, and methane is produced. The presence of methane in groundwater is indicative of strongly reducing conditions and occurs after oxygen, nitrate, iron, and sulfate have been depleted in the treatment zone.

Methane in water wells becomes a problem when it is allowed to build up in confined spaces. High concentrations of methane can displace oxygen or in the presence of a spark, can explode. Extensive testing for methane in water wells in southwest Colorado has been conducted during the past 10 years. As a result of this testing, many believe that methane concentrations below 1 mg/L are harmless. Methane levels up to 7 mg/L usually are not a concern but should be monitored for changes. Between 7 mg/L and 13 mg/L, additional monitoring and treatment should be considered. Also, care should be taken to ventilate confined spaces where well water is used. When methane levels are 13 mg/L or above, treatment is a must.

Theoretically, water containing as little as 1 to 2 mg/L of methane can produce an explosion in a poorly ventilated air space. Fires and explosions in well pits, basements, and water tanks have occurred from methane emitted by groundwater; safety measures include analysis to detect the presence of the gas, aeration of water before use, and adequate ventilation where the water is being used. There is a danger of people suffocating in dug wells and pump pits where high methane concentrations form (4).

BIBLIOGRAPHY

1. Anon. (1969). Gas in the ground water. *J. Am. Water Works Assoc.* **61**: 42.
2. Wild, A. (1996). *Soils and the Environment*. Cambridge University Press, New York, pp. 223–224.
3. U.S. Environmental Protection Agency. (1987). *Groundwater Hand Book*. Academic, Jaipur, India, p. 44.
4. Todd, D.K. (1995). *Groundwater Hydrology*. Wiley, Toronto, Canada, p. 307.

FOSSIL AQUIFERS

Laurel Phoenix
Green Bay, Wisconsin

Fossil aquifers are bodies of very deep aquifers that are disconnected from our contemporary hydrologic cycle. Water in these aquifers can be hundreds to thousands of years old and thousands of feet deep. As they are not recharged by today's precipitation, using them as a water supply is not sustainable. Other deep aquifers may still be connected to surface precipitation, but recharge so slowly relative to human time scales that they can practically be considered nonreplenishable fossil aquifers as well. Not only are these supplies finite, but their depth also makes them more expensive, not only for the initial drilling but also for the increased energy required to pump water from these depths.

WHY DEVELOP AN UNSUSTAINABLE AND EXPENSIVE WATER SUPPLY?

The use of fossil aquifers is a symptom of the larger world water deficit that has developed in recent history. As populations grow, and as cheap fossil fuel encourages overpumping (or mining), accessing fossil water is an easier choice than restricting current water use. Moreover, the increased import/export grain trade has globalized regional water deficits (1). For example, demand to import wheat goes up in water-stressed countries, because it takes roughly 1000 tons of water to produce one ton of wheat. Essentially, importing wheat is a surrogate for importing water, and farmers in other countries that are motivated by the increased world demand for wheat will grow wheat even if they have to pump fossil water to do it.

Fossil aquifers are often accidentally found when drilling for gas and oil, or when communities run out of surface and shallow groundwater supplies and begin exploring for deeper water. Water discovered when drilling for gas and oil has no exploration costs or risks to future water users, because the costs are already borne by the oil exploration. However, when new wells are sunk only to search for fossil water, costs are relatively high because there is no fossil fuel subsidy and some test wells come up dry.

Regions around the world have various reasons to turn to fossil aquifers for more water. Growing populations increase demand for potable water to provide at least the subsistence amounts of water that individuals need. If these populations migrate to cities in search of work, they create an increased localized water demand that may not be satisfied by current water supplies. These growing populations must also be fed, raising demand for agricultural production that requires irrigation water. As 85% of the world's fresh water is used for irrigation, inefficient irrigation methods used in developing countries waste a good deal of their available water. Countries that are recently industrializing also need more water for each new industry that comes online.

New water supplies are also needed when traditional water supplies have become polluted from sewage or industrial effluent. This polluted water essentially removes some accessible water from consideration as a potable water source. For example, in China, there are several rivers whose water is so polluted from industrial waste that the water is dangerous to touch. The famous Ganges River in India is so filled with untreated sewage that its bacterial counts are dangerously high. Some aquifers in India have become so polluted from pesticides that they can no longer be used for drinking water.

INTERNATIONAL EXAMPLES

Developing countries with dry regions, such as China, India, and Mexico, have overpumped their shallow aquifers and have sunk deeper wells to access fossil water. In China's Northern Plain, water tables have dropped 2–3 m per year in some areas (2). Water tables in the agricultural state of Guanajuato, Mexico, have dropped up to 3.3 m/yr. In various regions of India, water tables have fallen from 1 to 3 m/yr.

Countries in the arid climates of northern Africa and the Middle East have been pumping from fossil aquifers for years. Iran, Pakistan, Saudi Arabia, Yemen, Algeria, and Libya, among others, have been relying on these aquifers to provide more potable water that does not cost quite as much as desalinized water. To develop some food independence, Saudi Arabia has used fossil aquifers to support experimental wheat growing in the desert but is now dropping this program after finding it is more practical to import their wheat. Soon, the desert alfalfa fields will also be gone, which will decrease the experimental projects for raising their own milk cows. Sana'a, the capital of Yemen, relies on a fossil aquifer that will run out by 2010, forcing either a mass migration (unlikely) or constructing desalting plants to supply their two million residents. Libya has its famous $27 billion "Great Manmade River," a massive pumping and distribution project to transport fossil water from the Nubian aquifer, underlying Libya, Chad, Egypt, and Sudan, to Libyan cities on the Mediterranean coast (3). This aquifer is large enough to provide Libyans with water for over 100 years, but most of these other countries are using aquifers with a life of perhaps 10–30 years more.

Just as shared shallow aquifers have been sources of conflict between users, pumping from fossil aquifers can generate competition and conflict as well. As more countries move to access fossil aquifers, agreements between countries will have to be crafted to avoid such conflict. Aside from the Nubian aquifer (4) mentioned above, South Africa, Namibia, and Botswana share the Kalahari/Karoo aquifer, and Argentina, Brazil, Paraguay, and Uruguay share the Guarani aquifer. All are looking at how to exploit this resource.

AMERICAN EXAMPLES

In America, two good examples of fossil aquifers are the Ogallala, underlying several states, and numerous smaller aquifers in the dry Southwest. The Ogallala, which consists primarily of fossil water, is shared by eight states where water tables have dropped 100 ft in some areas. This overpumping was primarily driven by irrigation demands. Some areas have been overpumped to the point that deeper drilling has stopped once it becomes economically infeasible, which has caused a depopulation of some High Plains counties as farms go out of business and then towns lose the service jobs that supported the farmers.

The increasing use and/or hunt for fossil aquifers in pockets of the American Southwest has been driven by recent population migration to these warmer cities. Growing cities in arid regions create inevitable water crises. As demand increases and water supply decreases, these cities must either use their current sources more efficiently or continue to hunt for "new" water in competition with other cities doing the same. Reluctant to impose onerous restrictions on residents or scare off industries, these cities then turn to fossil water once surface sources and shallow groundwater sources are fully tapped, often at the expense of other communities. Las Vegas is the most infamous example, as it is now planning to pump fossil water from 28 sparsely populated basins to its west and north and transport this water through pipelines to the city (5).

Globally, more fossil water is being pumped every day. This water is important not so much as an interesting link to the past or a particular challenge to hydrogeologists, but because increasing reliance on these ancient waters in rich and poor countries alike portends an abrupt halt to regional and national economies when the pumps finally run dry.

BIBLIOGRAPHY

1. Brown, L. (2004). *Plan B: Rescuing a Planet Under Stress and a Civilization in Trouble*. W.W. Norton, New York.
2. Brown, L., Larsen, J., and Fischlowitz-Roberts, B. (2002). *The Earth Policy Reader*. W.W. Norton, New York.
3. Tyler, P. (2004). Libya's Vast Pipe Dream Taps into Desert's Ice Age Water. *New York Times*, Mar. 2.
4. http://www.heritage.umd.edu/CHRSWeb/World%20Heritage/Heritage%20Spots/Heritage%20Spots%20PH.htm. Accessed Sept. 1, 2004.
5. Ward, D.R. (2002). *Water Wars: Drought, Flood, Folly, and the Politics of Thirst*. Riverhead Books, New York.

WHAT IS A HYDROCHEMICAL MODEL?

Kung-Yao Lee
Thomas R. Fisher
Horn Point
Laboratory—UMCES
Solomons, Maryland

INTRODUCTION

A hydrochemical model is a set of mathematical algorithms that calculate water discharge ("hydro") and concentrations of sediment and chemical constituents ("chemical") in a stream or river. The computations are based on adjustable, empirical, or theoretical variables (parameters) and meteorological data for the drainage basin of the river or stream (input data). A variety of hydrochemical models have been employed by researchers, engineers, and environmental planners to predict hydrochemical responses in areas where there are no measurements or in areas where measurements are available but under different management scenarios. In general, a hydrochemical model is calibrated (adjusted to fit) with a set of observations, validated (tested for errors) with a set of reserved observations, and then applied to future or past conditions (forecasting or hindcasting, respectively) to predict watershed responses. From an environmental perspective, a hydrochemical model is a tool to investigate the effects of watershed management (e.g., changes of land use or land cover, human population, etc.) on the hydrology and chemistry of a stream or river.

Generally speaking, a hydrochemical model comprises three simulation submodels (compartments or blocks). Here we use “submodel” to represent a smaller unit or component of a hydrochemical model, and the three major submodels are generally responsible for computing (1) hydrology, (2) soil erosion, and (3) nutrient or metal geochemistry within the drainage basin.

The first submodel estimates the hydrology. This submodel partitions rainfall into physical evaporation, transpiration by plants, surface storage (puddles), overland flow during storms (runoff), and groundwater recharge because of infiltration during storms. Water storage in groundwater slowly flows to streams, creating stream flow (baseflow) between rain events, and overland flow during storms creates stormflow (see Fig. 1). Streamflow is the sum of storm and base flows and is usually integrated over a storm event or a fixed time period such as a day, month, or year.

The hydrology submodel is usually accompanied by an erosion submodel. This model component simulates the export of soil eroded by overland flow generated in the hydrology submodel. The rate of soil erosion is often based on the U.S. Department of Agriculture’s universal soil loss equation (USLE), which uses basin characteristics such as vegetation cover, slope angle and length, cultivation practices, and soil properties such as particle size and density.

Additional geochemical submodels are responsible for the simulation of chemical constituents, which are nutrients such as N, P, or Si, or metals such as Fe, Al, Pb, and so on, from both natural and anthropogenic (manmade) sources. Within the geochemical submodel, these chemical constituents are transported with sediment or in the water flows that are simulated in the first submodel. Despite important differences between hydrochemical models, it is important to realize that the modeling principles developed in a variety of hydrochemical models are similar (2); i.e., the hydrology submodel is developed first and then followed by erosional and geochemical submodels.

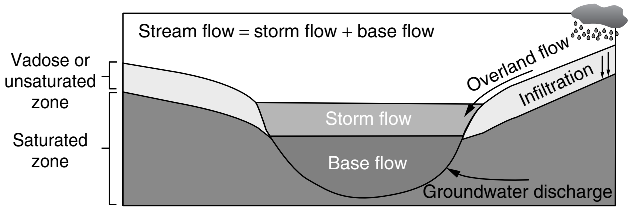

Figure 1. Simplified concept of the hydrological cycle used in hydrochemical models. Stream flow is the sum of storm flow and base flow, as used in GWLF (1) as well as in other models. However, some models may adopt different modeling approaches that differ from this illustration.

Numerous hydrochemical models are freely available. Many can be downloaded from websites or can be requested from the model developers. In Table 1, we list some commonly used hydrochemical models. Each hydrochemical model has its specific designs and limitations, such as model time step (minimum temporal resolution), number of parameters required (internal complexity), or input requirements. For detailed comparisons between these models, see Donigian and Huber (12). Because all models require at least some data input, many hydrochemical models have been ported from their original code (e.g., Fortran or Visual Basic) to more user-friendly windows applications for better visualization or to integrate with a geographic information system (GIS), which provides advanced computing power to manipulate data as subsequent input for the models (see 13–17).

PROPERTIES OF HYDROCHEMICAL MODELS

Hydrochemical models generally can be categorized as lumped-parameter and distributed models. A lumped-parameter model simplifies the watershed properties and ignores the spatial heterogeneity and distributions of these properties, using only the average watershed properties (see 18,19). Although a lumped-parameter model can efficiently generate simulation results, model users cannot spatially identify the locations where problems (e.g., nutrient or metal sources) occur. However, lumped-parameter models linked with a GIS may be applied in spatially contiguous polygons to provide some spatial resolution within a larger watershed (e.g., 15). Conversely, a distributed model can explicitly represent the spatial effects of topography or the spatial variations of properties within a watershed to locate sources; however, a distributed model requires more computer resources because of more intensive computation and more extensive data requirements.

As discussed earlier, a hydrochemical model can be designed either for long-term predictions or as an event simulation tool (or both). Time series of weather data (hourly or daily) are usually required for continuous simulation at hourly, daily, or monthly time steps. Models that predict at longer time steps ($\geq$ monthly; e.g., GWLF) do not resolve individual hydrographs (stream responses to individual storm events; see Fig. 2), whereas models used with shorter time steps (e.g., HSPF) attempt to predict both the timing and magnitude of a watershed’s response to a rain event. The storm usually occurs over hours to days, whereas the watershed response is usually days to weeks, depending on the size of the watershed. Other models (e.g., ANSWERS and AGNPS—earlier versions) have

Table 1. Hydrochemical Models Commonly Used in Watershed Ecology. Models Typically Use a Lumped Parameter Approach, Where Average Watershed Properties Such as Soil Drainage or Land Cover are Used, Without Regard to Spatial Distribution, Whereas Distributed Models Require Information on the Spatial Distributions of Watershed Properties

Acronym	Full Name	Property	Type	Model Complexity	Reference
AGNPS*	Agricultural Nonpoint Source Pollution Model	Distributed	C, S	Moderate	Young et al. (3)
ANSWERS*	Areal Nonpoint Source Watershed Environment Response Simulation	Distributed	C (2000 version), S (older version)	Moderate	Beasley and Huggins (4)
CREAMS*	Chemical, Runoff, and Erosion from Agriculture Management Systems	Distributed	C, S	High	Knisel (5)
GWLF*	Generalized Watershed Loading Functions	Lumped	C	Simple	Haith & Shoemaker (1)
HSPF	Hydrological Simulation Program—Fortran	Lumped	C, S	High	Bicknell et al. (6)
SWAT*	Soil and Water Assessment Tool	Distributed	C	Moderate	Arnold et al. (7)
SWIM*	Soil and Water Integrated Model	Distributed	C	High	Krysanova et al. (8,9)
SWMM	Storm Water Management Model	Lumped	C, S	High	Huber & Dickinson (10)
SWRRB*	Simulator for Water Resources in Rural Basins	Distributed	C	High	Williams et al. (11)

* The SCS Curve Number method calculates surface runoff.
C: Continuous; S: Single Event.

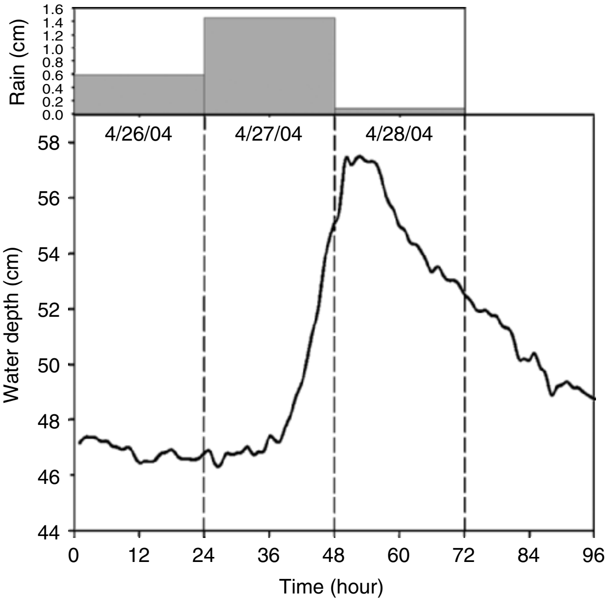

Figure 2. A hydrograph in response to a storm event that started to occur on 4/26/2004 at the Norwick Creek subwatershed of the Choptank River basin, Maryland. Note that precipitation (cm) was recorded on the daily basis (upper panel), whereas water depth (cm) at the watershed outlet was measured at a 30-min interval (lower panel). Data were compiled from Sutton (20).

been developed for the explicit purpose of predicting short-term hydrograph events such as the observations in Fig. 2.

PARAMETERIZATION OF THE HYDROCHEMICAL MODEL

A hydrochemical model generates a time series of predicted data for a time span defined by the user. To assess the accuracy and precision of the model's behavior, model output is compared on a daily, monthly, or annual basis with observed measurements. Comparisons between model output and model simulations are usually expressed in charts, bar graphs, or as tabular values to assess the model's accuracy (average ability to predict the observations without positive or negative bias, i.e., without underprediction or overprediction over a time interval) and precision (ability to reproduce the observations without scatter). These are often quantified as the cumulative model error over a given time period (model accuracy, e.g., ability to predict the observed annual stream flow for a given year) and root-mean-square (RMS) error at a given time scale (model precision, e.g., the deviations of modeled and observed stream flow at the defined time step). During calibration, model parameters are adjusted to minimize RMS or cumulative errors between the predicted and observed values, and the success of a calibration is then quantified by validation with an independent dataset not used during calibration.

The first comparison of a model's performance is called "Model Calibration." During this process, parameters are adjusted within reasonable ranges to minimize model errors. Model calibration often is a cyclical (21) and time-consuming process to adjust parameter values, although some optimization computer programs are available to automatically calibrate for the most suitable parameters. For example, HSPFEXP is an expert system that can be linked to HSPF to optimize the hydrology parameters (22). Monte Carlo methods (23) can perform statistical probability analyses to obtain the optimal values of variables. The parameter estimation software PEST (24) can also be linked to any existing model for parameter calibration. Despite different approaches to optimize the parameters, it is important to constrain the ranges of model parameters within reported values (14) during model calibration.

The second comparison of model performance is called "Model Verification" or "Model Validation." This process uses the calibrated model parameters with a different set of meteorological data not used in model calibration. In this phase, a similar set of statistical analyses is generated that set the limits of interpretation of model predictions and to investigate the model's ability to simulate natural conditions (25). For example, if validation RMS errors at the daily time scale are ±100%, then only model predictions exceeding this threshold can be considered significantly different. Model verification ensures the validity of calibrated model parameters under the current conditions, and to confirm that the model can predict the stream hydrology and chemistry under different scenarios. Model validation errors are typically large at short time scales (hours-days) and become both more accurate and more precise at longer times scales (years-decades); i.e., it is easier to predict the cumulative discharge or material flux for a given year than it is to predict the daily flow or flux.

MODELING AND REALITY

In most cases, a well-calibrated hydrochemical model can predict water flows and chemistry in situations where no data exist. It is important, however, to remember that models sometimes do not predict well in some areas or under some conditions because model parameters may be difficult to measure (26) or because the model was developed under experimental conditions or with dominant processes that may not fully represent all other cases (27). In this case, a model's predicting performances are substantially affected by these factors when applied elsewhere. For example, AGNPS and ANSWERS are limited to watershed size <200 km^2, whereas SWRRB can be applied to much larger areas (9). In other words, model parameters are often designed for particular geographical areas or watershed sizes. The alpine hydrochemical model (AHM) was, for example, originally developed for the hydrologic and biogeochemical responses of seasonally snow-covered alpine areas (28), and SWMM was mainly designed to perform predictions in urbanized areas, although both can be used elsewhere.

To simplify the natural processes, hydrochemical models often possess unique assumptions. Assumptions may be based on field observations (29) and/or empirically derived functions to simplify the response processes. For example, the Soil Conversation Service (SCS) Curve Number (30) computes direct surface runoff merely based on land use and soil information, which has empirically simplified the estimation of other processes such as interception, depression storage, and infiltration (31). Because of its simplicity, the SCS CN method has been adopted to most hydrochemical models (marked * in Table 1). However, these processes may need to be parameterized for some models, such as HSPF and SWMM.

When rain infiltrates the soil, models often incorporate other empirical assumptions and simplifications borrowed from observations. For example, the soil column above the water table can be separated into two layers—unsaturated and saturated layers (e.g., GWLF). The disadvantage of this simplified approach is reduced model flexibility when applied to other areas. Conversely, SWIM (14) divides the root zone vertically into as many as 10 layers, which increases the model flexibility when applied elsewhere. However, these models require more model variables and intensive input data.

Furthermore, it is often seen that a fixed value is used for a variable throughout all modeling tasks. For example, both GWLF and SWMM use constant values for concentrations of groundwater chemical constituents (e.g., N and P) associated with each land use because both models assume that there are no geochemical interactions for the chemical constituents in the subsurface flow. Also models often assume that the observed chemical measurements in a stream or river have promptly responded to the geochemical reactions that occur in the soil or in the groundwater during a storm event, which is usually not true because groundwater may take years to decades to become the baseflow of streams, and important geochemical reactions may occur at this time scale (29).

Hence, hydrochemical models because of model assumptions and simplifications inevitably produce prediction errors. Those errors are intensified and transferred from the hydrology submodel to the subsequent submodels, which results in decreasing reliability in predictions (32). Therefore, it is important to understand the model assumptions and limitations, and to evaluate the strength and weakness of each hydrochemical model before use in each situation (33).

BIBLIOGRAPHY

1. Haith, D.A. and Shoemaker, L.L. (1987). Generalized watershed loading functions for stream flow nutrient. AWRA. *Water Resour. Bull.* **23**(3): 471–478.
2. Krysanova, V., Meiner, A., Roosaare, J., and Vasilyew, A. (1989). Simulation modeling of the coastal waters pollution from agricultural watershed. *Ecol. Modeling* **49**: 7–29.
3. Young, R.A., Onstad, C.A., Bosch, D.D., and Anderson, W.P. (1986). AGNPS: A nonpoint source pollution model for evaluating agricultural watersheds. *J. Soil Water Conservation.* **44**(2): 168–173.
4. Beasley, D.B. and Huggins, L.F. (1981). *ANSWERS Users Manual*. EPA-905/9-82-001, U.S. EPA, Region V, Chicago, IL.
5. Knisel, W. (1980). *CREAMS: A Field-Scale Model for Chemicals, Runoff, and Erosion from Agricultural Management Systems*. U.S. Department of Agriculture. Conservation Research Report No. 26.
6. Bicknell, B.R., Imhoff, J.C., Kittle, J.L., Jr., Donigian, A.S., Jr., and Johanson, R.C. (1993). *Hydrological Simulation Program—FORTRAN, User's Manual for Release 10*. EPA/600/R-93/174, U.S. Environmental Protection Agency, Office of Research and Development, Athens, GA.
7. Arnold, J.G., Williams, J.R., Srinivasan, R., King, K.W., and Griggs, R.H. (1994). *SWAT—Soil and Water Assessment Tool*. USDA, Agricultural Research Service, Grassland, Soil and Water Research Laboratory, Temple, TX.
8. Krysanova, V., Müller-Wohlfeil, D.I., and Becker, A. (1996). *Integrated Modeling of Hydrology and Water Quality in*

Mesoscale Watersheds. PIK Report, No. 18, July 1996, Potsdam Institute for Climate Impact Research (PIK), Posdam, Germany.

9. Krysanova, V., Müller-Wohlfeil, D.I., and Becker, A. (1998). Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. *Ecol. Modeling* **106**: 261–289.
10. Huber, W.C. and Dickinson, R.E. (1988). *Storm Water Management Model*. Version 4, User's Manual, EPA/600/3-88/001a (NTIS PB88-236641/AS), U.S. EPA, Athens, GA.
11. Williams, R.R., Nicks, A.D., and Arnold, J.G. (1985). Simulator for water resources in rural basins. *ASCE Hydraulics J.*: 67–103.
12. Donigian, A.S. and Huber, W.C. (1993). *Modeling of Nonpoint Source Water Quality in Urban and Non-Urban Areas*. Environmental Research Laboratory, Office of Research and Development, U.S. EPA. EPA/600/3-91-039.
13. Needham, S.E. and Young, R.A. (1994). ANN-AGNPS: A Continuous Simulation Watershed Model. *Proceedings of the Federal Interagency Workshop on Hydrological Modeling Demands for the 90's*. U.S. Geological Survey Water-Resources Investigations Report 93–4018.
14. Krysanova, V. and Becker, A. (2000). Integrated modeling of hydrological processes and nutrient dynamics at the river basin scale. *Hydrobiologia* **410**: 131–138.
15. Lee K.-Y., Fisher, T.R., and Rochelle-Newall, E. (2001). Modeling the hydrochemistry of the Choptank River basin using GWLF and ArcInfo: 2. Model Validation and Application. *Biogeochemistry* **56**: 311–348.
16. USEPA. (2001). *Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) v. 3.0 User's Manual*. EPA-823/B01-001, Office of Water, USEPA, Washington, DC.
17. Evans, B.M. et al. (2002). A comprehensive GIS-based modeling approach for predicting nutrient loads in watersheds. *J. Spatial Hydrol.* **2**(2).
18. Lee, K.-Y., Fisher, T.R., Jordan, T.E., Correll, D.L., and Weller, D. (2000). Modeling the hydrochemistry of the Choptank River basin using GWLF and ArcInfo:1. Model calibration and validation. *Biogeochemistry* **49**: 143–172.
19. Helliwell, R.C. and Kernan, M. (2004). Modeling hydrochemical and ecological trends in acid sensitive surface waters in the Scottish Highlands. *J. Limnol.* **63**(1): 111–122.
20. Sutton, A.J. (2005). *Effectiveness of CREP: Reducing Nutrient Loads to Chesapeake Bay*. Ph.D. dissertation. University of Maryland, College Park, MD.
21. Love, J.T. and Donigian, A.S., Jr. (2002). *The Connecticut Watershed Model—Model Development, Calibration, and Validation*. WEF-Watershed 2002, February 23–27, 2002. Ft. Lauderdale, FL, WEF-2002 Specialty Conference Proceedings on CD-ROM.
22. Lumb, A.M., McCammon, R.B., and Kittle, J.L., Jr. (1994). *Users Manual for An Expert System (HSPEXP) for Calibration of the Hydrologic Simulation Program—Fortran*: U.S. Geological Survey Water-Resources Investigations Report 94–4168.
23. Niederreiter, H. and Spanier, J. (2000). Monte Carlo and Quasi-Monte Carlo Methods 1998. *Proceedings of a Conference held at the Claremont Graduate University, Claremont, California, June 22–26, 1998*, Springer-Verlag, Berlin.
24. Dorthy, J. (2004). *PEST, Model-Independent Parameter Estimation, User Manual*. 5th Edn. Watermark Numerical Computing, Brisbane, Australia.
25. Stone, A. and Seip, H.M. (1990). Are mathematical models useful for understanding water acidification? *Sci. Total Environ.* **96**: 159–174.
26. Hooper, R.P. (2001). Applying the scientific method to small catchment studies: A review of the Panola Mountain experience. *Hydrol. Processes* **15**: 2039–2050.
27. Beven, K. (1993). Prophecy reality and uncertainty in distributed hydrological modeling. *Adv. Water Resour.* **16**: 41–51.
28. Wolford, R.A. (1992). *Integrated Hydrochemical Modeling of an Alpine Watershed: Sierra Nevada, California*. Report 92-040, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ.
29. Kirchner, W.J., Hooper, R.P., Kendall, C., Neal, C., and Leavesley, G. (1996). Testing and validating environmental models. *Sci. Total Environ.* **183**: 33–47.
30. Soil Conservation Service (SCS). (1975). *Urban Hydrology for Small Watersheds*. Tech. Rel. no. 55, U.S. Department of Agriculture, U.S. Government Printing Office, Washington, DC.
31. Chow, V.T., Maidment, D.R., and Mays, L.W. (1988). *Applied Hydrology*. McGraw-Hill, New York.
32. Novotny, V. and Olem, H. (1994). *Water Quality: Prevention, Identification, and Management of Diffuse Pollution*. Van Norstrand Reinhold, New York.
33. Dunne,T. and Leopold, L.B. (1978). *Water in Environmental Planning*. W. H. Freeman and Company, San Francisco, CA.

MODELING NON-POINT SOURCE POLLUTANTS IN THE VADOSE ZONE USING GIS

Dennis L. Corwin
USDA-ARS George E. Brown,
Jr., Salinity Laboratory
Riverside, California

Keith Logue
Stanford University
Stanford, California

NON-POINT SOURCE POLLUTION: DEFINITION, SOURCES, SIGNIFICANCE, GLOBAL IMPACT, AND JUSTIFICATION FOR MODELING

Non-point source (NPS) pollutants are defined as "contaminants of [air, and] surface and subsurface soil and water resources that are diffuse in nature and cannot be traced to a point location" (1). Characteristically, NPS pollutants (1) are difficult or impossible to trace to a source; (2) enter the environment over an extensive area; (3) are related, at least in part, to certain uncontrollable meteorological events, and existing geographic and geomorphologic conditions; (4) have the potential for maintaining a relatively long active presence in the global ecosystem; and (5) may result in long-term, chronic effects on human health, and soil-aquatic degradation (2). The most common global NPS pollutants of soil and groundwater resources include biosolids and manure, persistent organic pollutants (POPs), nutrients (e.g., nitrates and phosphates), salinity, toxic heavy metals (e.g., Bi, Co, Sn, Te, Ag, Pt, Tl, Sb, Hg, As, Cd, Pb, Cr, Ni), trace elements (e.g., Se, B, Mo, Cu, Zn), and pathogens.

Often, NPS pollutants occur naturally, such as salts and trace elements in soils, or are the consequence of direct application by humans (e.g., pesticides and fertilizers), but regardless of their source, they are generally the direct consequence of human activities including agriculture, urban runoff, feedlots, hydromodification, and resource extraction (2). Specific sources of NPS pollutants include (1) excess fertilizers, herbicides, and insecticides from agricultural lands and residential areas; (2) oil, grease, and toxic chemicals from urban runoff and energy production; (3) sediment from improperly managed construction sites, crop and forest lands, and eroding stream banks; (4) naturally occurring salts and trace elements from irrigation practices; (5) acid drainage from abandoned mines; (6) pathogens (i.e., viruses and bacteria) and nutrients from livestock, and pet wastes; and (7) atmospheric deposition (2).

The significance of NPS pollutants as an environmental issue stems from their potential global impact and resultant chronic effects on human health. Because of their widespread use and often persistent and mobile nature, NPS pollutants have the capacity not only to injure the surrounding environment and ourselves, but also remote environments and their inhabitants (like the Arctic) and future generations of humans through the buildup of persistent toxic substances that mimic and disrupt human hormone systems (3). The impact of NPS pollutants on soil and water resources extends over millions of hectares of land and billions of liters of water. Throughout the world, 30% to 50% of Earth's land is believed to be affected by NPS pollutant degradation from erosion, fertilizers, pesticides, organic manures, and sewage sludge (4). Worldwide, NPS pollutants are recognized as *the* major contributors to surface and groundwater contamination (5), with agriculture as the single greatest contributor of NPS pollutants (6). Agricultural activities result in the movement of NPS pollutants from the soil surface into rivers and streams via runoff and erosion, and into subsurface soil and groundwater via leaching through the vadose zone (i.e., the portion of the soil extending from the soil surface to the groundwater table). Figure 1 reflects the worldwide extent of human-induced degradation of soil by region and by cause over the period 1945 to the late 1980s.

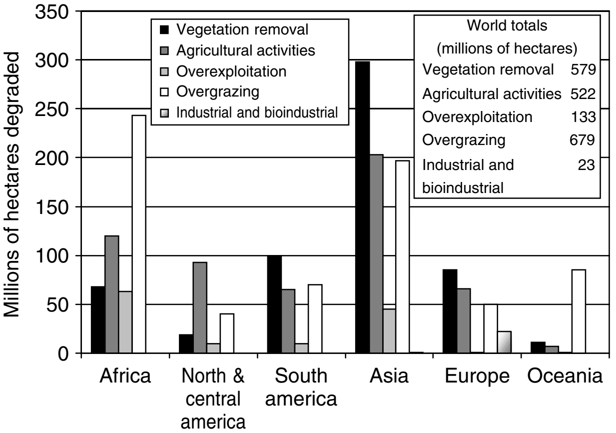

Figure 1. Human-induced soil degradation by region and by cause from 1945 to the late 1980s. *Source:* Ref. 7. With permission.

The world's population has doubled since 1950 and is expected to range from 8 to 12 billion in 2050. Barring unexpected technological breakthroughs, sustainable agriculture is viewed as the most viable means of meeting the food demands of the projected world's population. The concept of sustainable agriculture is predicated on a delicate balance of maximizing crop productivity and maintaining economic stability while minimizing the utilization of finite natural resources and the detrimental environmental impacts of associated NPS pollutants. Assessment of NPS pollutant impacts on soil-groundwater systems at local, regional, and global scales is a key component to achieving sustainable agriculture. Assessment provides the means (1) to establish the true extent of the NPS-pollution problem, (2) to evaluate mitigating management practices and regulatory policies, and (3) to predict future potential problems. The distinct advantage of prediction is that it can alter the occurrence of detrimental conditions before they occur. The reasons for modeling NPS pollutants in the vadose zone are (1) to increase the understanding of cause-and-effect relationships of spatiotemporal processes occurring in soil systems and (2) to provide a cost-effective means of synthesizing the current level of knowledge into a useable form for making environmental policy decisions (8,9).

MULTIDISCIPLINARY NATURE OF MODELING NPS POLLUTANTS IN THE VADOSE ZONE

Modeling NPS pollutants in the vadose zone is a complex environmental problem that requires a multidisciplinary, systems-based approach taken within a spatial context with an awareness of scale (10). The formidable barriers to modeling NPS pollutants are the consequence of the complexities of geographic scale and position; the complexities of the physical, chemical, and biological processes of solute transport in porous media; and the spatial complexities of the soil media's heterogeneity. The knowledge, information, and technology needed to address each of these issues crosses several subdisciplinary lines, including classic and spatial statistics, remote sensing, geographic information systems (GIS), surface and subsurface hydrology, soil science, and space science. Spatial statistics is useful in dealing with the uncertainty and variability of spatial information (11); remote sensing provides measurements of physical, chemical, and biological properties needed in environmental models (12); GIS is a means of organizing, manipulating, storing, and displaying spatial data (13); and water flow and solute transport models developed within soil science and hydrogeology are the tools for simulating future scenarios to assess potential temporal and spatial changes (2,14). Precise geographic location and areal extent are captured with the space science technology of the global positioning system (GPS).

COMPONENTS OF A NPS POLLUTANT MODEL

Modeling the fate and movement of NPS pollutants in the vadose zone is a spatial problem well suited for the integration of a deterministic solute transport model with a GIS. A GIS characteristically provides a means of representing the real world through integrated layers of constituent spatial information. To model NPS pollution within the context of a GIS, each transport parameter or variable of the deterministic transport model is represented by a three-dimensional layer of spatial information. The three-dimensional spatial distribution of each transport parameter/variable must be simulated, measured, or estimated, which creates a tremendous volume of spatial information because of the complex spatial heterogeneity exhibited by the numerous physical, chemical, and biological processes involved in solute transport through the vadose zone. GIS serves as the tool for organizing, manipulating, and visually displaying this information efficiently.

The essential components of modeling NPS pollutants consist of (1) a *model* of solute transport and/or accumulation, (2) input and parameter *dat*a for the model, and (3) a *GIS* to handle the input, manipulation, and output of spatial data. Figure 2 shows the interaction among these three basic components based on flow of information.

Because of the complex spatial heterogeneity of the vadose zone, a variety of sophisticated techniques are useful as tools to deal with the vicissitudes of soil (Fig. 2). Fuzzy logic theory provides a means of handling vague and imprecise data whether as a means to characterize map units or transitional boundaries between map units (16,17). Fractal geometry with its scale independence may offer a means of bridging a variety of gaps related to spatial variability from determining the predictability of complex spatial phenomena such as solute transport to relating difficult-to-measure soil hydraulic properties to other soil variables available from soil surveys (18). Geostatistics is useful in interpolating sparse spatial data and providing associated uncertainty (19). Hierarchical theory establishes an organizational hierarchy of pedogenetic modeling approaches and their appropriate scale of application (20,21). Wavelet analysis provides a means of determining spatial scales and the dominant processes at those scales (22,23). Neural networks and transfer functions provide a means of deriving complex hydraulic parameters from easily measured data (24). Digital terrain or digital elevation models (DEM) provide spatial geomorphologic information. Uncertainty analysis serves as a means of establishing the reliability of simulated model results based on model errors and data uncertainties (25).

Data

The effectiveness of a model to simulate a practical application is highly dependent on how well model inputs and model parameters are identified. Basically, three sources of input and parameter data for NPS pollutant models exist (15): (1) measured data, (2) estimated data, and (3) existing data. Each source of data carries distinct advantages and limitations.

A review of current measurement techniques to determine flow-related properties of subsurface porous media and soil physical properties is provided by Dane and Molz (26) and Topp et al. (27), respectively. Although direct measurement of transport parameters and variables is the most reliable means of obtaining accurate information for modeling purposes, it is also the most labor intensive and costly. A quick and easy means of obtaining these measurements is crucial to the cost-effective modeling of NPS pollutants. Remote sensing and noninvasive measurement techniques have the greatest potential for meeting the thirst for measured spatial data.

Corwin (28) provided a cursory review of some instrumental techniques developed for the remote and noninvasive measurement of variables and parameters found in transport models for the vadose zone. The review covers geophysical resistivity methods, aerial photography, x-ray tomography, ground-penetrating radar, magnetic resonance imaging, microwaves, multispectral imagery, thermal infrared imagery, and advanced very-high-resolution radiometry (AVHRR). Barnes et al. (12) provided a more

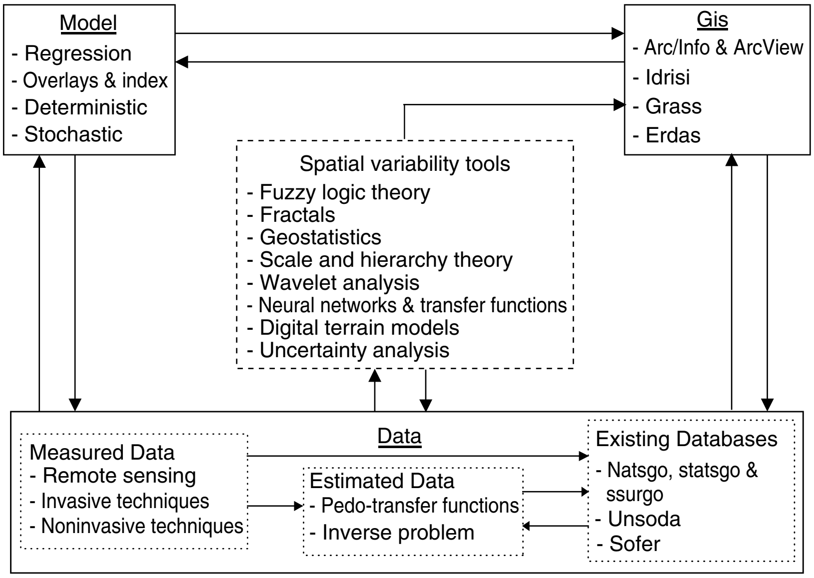

Figure 2. Integrated components of a GIS-based NPS pollutant model system. Arrows show the flow of information. Modified from Ref. 15. With permission.

recent review of remote- and ground-based sensor technology for mapping soil properties.

Even though considerable progress has been made over the past decade, the remote sensing instrumentation needed to measure *all* parameters and variables in even the simplest of transport models, for the vadose zone is not available or even on the drawing board. In most cases, remote sensing provides measurements of only the top few centimeters; consequently, it suffers from a lack *of depth information* needed in modeling the vadose zone. Noninvasive techniques such as electromagnetic induction (EMI) can provide information down to several meters in depth, but these techniques generally require measurements taken at or near the soil surface and their measurement volume is limited to tens of cubic meters or less. Nevertheless, geospatial measurements of apparent soil electrical conductivity (EC_a) with EMI is currently the most widespread and reliable means of characterizing the spatial variability of a variety of physicochemical properties in the vadose zone including salinity, texture, water content, cation exchange capacity (CEC), organic matter (OM), and bulk density (29). Furthermore, in most cases, the parameters measured by remote sensing and noninvasive techniques are often not directly applicable to solute transport models. For instance, the use of EMI to measure soil salinity is not a direct measure of salinity in the soil solution, but rather it measures EC_a, which includes the conductivity of both the solid and liquid phases, thereby requiring ground-truth soil samples for calibration.

The extreme spatiotemporal variability and the nonlinearity of many soil processes make parameterization of landscape-scale models a daunting task. The inability of remote measurement techniques to meet the demand for spatial and temporal parameter and input data has resulted in the development of transport parameter estimation techniques that estimate parameters by fitting data or are based on the formulation of transfer functions. Inverse modeling is a powerful and practical means of estimating flow and transport parameters for landscape-scale solute transport models using advanced optimization algorithms. Transfer functions relate readily-available and easy-to-measure soil properties to more complex transport variables/parameters needed for simulation.

Corwin et al. (15) provide a referenced list of the estimation methods for many common parameters in solute transport models of the vadose zone. The most common transfer function, the pedo-transfer function (PTF), uses particle-size distribution, bulk density, and soil organic-carbon content to yield soil-water retention or unsaturated hydraulic conductivity functions. Rawls et al. (30) provide a review of soil-water retention estimation methods. Reviews of methods of estimating soil hydraulic parameters for unsaturated soils have been written by van Genuchten et al. (31).

In most instances, limited resources do not permit the measurement or even estimation of needed input or parameter data. In these instances, the use of existing data is crucial. Existing soil databases for the United States include SSURGO (State Survey Geographical Database; http://www.ncg.nrcs.usda.gov/ssurgo.html), STATSGO (State Soil Geographical Database; http://www.ncg.nrcs.usda.gov/statsgo.html), and NATSGO (National Soil Geographical Database; http://www.ncg.nrcs.usda.gov/natsgo.html). SSURGO (map scale ranges from 1:12,000 to 1:63,360) is a county-level database, and it is the most detailed GIS database available from NRCS. STATSGO (map scale 1:250,000) is the state-level database designed for state, large watershed, and small river basin purposes. NATSGO (map scale 1:7,500,000) is the national soil database whose map units are defined by major land resource area (MLRA) and land resource region (LRR) boundaries.

The problem with the use of generalized rather than measured data has been associated uncertainties. Loague et al. (32) extensively reviewed the uncertainty associated with the use of an existing database for non-point source groundwater vulnerability and concluded that assessments based on this type of data are relegated to guiding data collection strategies rather than their intended purpose of groundwater vulnerability assessment. Measured input data that captures natural variability both in space and time is essential for diminishing uncertainty in simulations.

GIS

A GIS is defined by Goodchild (33) as a "general-purpose technology for handling geographic data in digital form with the following capabilities: (i) the ability to preprocess data from large stores into a form suitable for analysis (reformatting, change of projection, resampling, and generalization), (ii) direct support for analysis and modeling, and (iii) postprocessing of results (reformatting, tabulation, report generation, and mapping)." In the context of NPS pollutant modeling, a GIS is a tool that characterizes the full information content of the spatially variable data required by solute transport models. The advantages of GIS include (1) ease of data retrieval; (2) ability to discover and display information gained by testing interactions between phenomena; (3) ability to synthesize large amounts of data for spatial examination; (4) ability to make scale and projection changes, remove distortions, and perform coordinate rotation and translation; and (5) capability to discover and display spatial relationships through the application of empirical and statistical models (34). The principal benefit of coupling GIS to subsurface hydrologic models is to enable the models to deal with large volumes of spatial data that geographically anchor many environmental processes.

Model

A review of GIS-based NPS pollutant modeling in the vadose zone has been presented by Corwin et al. (15). To date, most models of NPS pollutants in the vadose zone have used deterministic models of solute transport coupled to a GIS (15,28). However, a growing recognition exists that stochastic approaches may offer the most viable means of modeling an NPS pollution (35).

The use of deterministic transport models with GIS has been justified on practical grounds based on availability, usability, widespread acceptance, and the assumption that a heterogeneous medium macroscopically behaves like a homogeneous medium with properly determined parameters and variables. The philosophy of modeling NPS pollutants in the vadose zone with a one-dimensional deterministic model of solute transport is based on the representation of physical, chemical, and biological properties influencing transport in the vadose zone with a distributed parameter structure. The validity of the assumption that a heterogeneous medium macroscopically behaves like a homogeneous medium depends on whether spatial domains can be defined and characterized that behave as stream tubes or “representative element volumes” (REV).

Three categories of deterministic models have been coupled to GIS to simulate NPS pollution in the vadose zone: regression models, overlay and index models, and transient-state solute transport models (15). Regression models have generally used multiple linear regression techniques to relate various causative factors to the presence of an NPS pollutant. These causative factors have included soil properties or conditions related to groundwater vulnerability or to the accumulation of a solute in the soil root zone. Overlay and index models refer to those models that compute an index of NPS pollutant mobility from a simple functional model of steady-state solute transport. Two types of overlay and index models have been developed: property-based and process-based. Property-based index models are established on a hydrogeologic setting (e.g., DRASTIC) or NPS pollutant properties (e.g., GUS). Process-based index models are founded on the characterization of transport processes (e.g., Rao’s attenuation factor model). Overlay and index models have been used largely to assess groundwater pollution vulnerability to pesticides and nitrates. Transient-state, process-based solute transport models include deterministic models capable of handling the movement of a pollutant in a dynamic flow system. Transient-state, process-based models describe some or all of the processes involved in solute transport in the vadose zone: water flow, solute transport, chemical reactions (adsorption-desorption, exchange, dissolution, precipitation, etc.), root growth, plant-water uptake, vapor phase flow, degradation, and dispersion/diffusion.

Jury (36) pointed out that the difficulty of constructing a three-dimensional model of chemical transport as a consequence of field variability has two significant implications: (1) Any hope of attempting to estimate a continuous spatial pattern of chemical transport must be abandoned; and (2) a possibility exists of extreme deviations from average movement so that significant concentrations of chemical may flow within relatively small fractions of the total cross-sectional area, which may be nearly impossible to detect from point measurements. The latter implication has fostered the development of stochastic solute transport models for the vadose zone as opposed to deterministic models.

Two distinct stochastic approaches are currently in use for dealing with the spatial variability encountered in modeling NPS pollutants in the vadose zone: geometric scaling and regionalized variables. Jury (36) indicates that geometric scaling uses specific “standardized variables to scale the differential equations describing transport and relates the standardized variables to some measurable or definable property of each local site of a heterogeneous field.” Once the variables are defined, the onerous task of characterizing the variability is reduced to determining the statistical and spatial distribution of these scaling parameters. In contrast, Jury (36) explains that the regionalized variable approach regards the “various parameters relevant to a field-wide description of transport as random variables characterized by a mean value and a randomly fluctuating stochastic component.”

In comparison with deterministic models, the coupling of a stochastic solute transport model to GIS is less explored. In a paper discussing the potential compatibility of stochastic transport models with GIS, Jury (35) suggested that stochastic-convective stream tube modeling seems the most compatible with GIS because it “utilizes a relatively simple local process driven by parameters that might be associated with soil morphological features, and could be integrated up to a large scale by simple arithmetic averaging over the local sites.” A stochastic stream tube model is made up of parallel, noninteracting one-dimensional soil columns whose properties are locally homogeneous, but vary from one soil column to the next. The collection of all stream tubes constitutes the field-, basin-, or regional-scale area being represented. This approach is in essence the same approach that has been undertaken in the past where deterministic piston-flow local transport models have been coupled to soil survey information; only now there is an associated stochastic component of information. Jury (35) warns that the challenge of this approach will be “to develop a reasonable local-scale model whose parameters can be related to identifiable local-scale features.”

CONSIDERATION OF SCALE WHEN MODELING NPS POLLUTANTS

Scale, as used in soil science and hydrology, refers to the “characteristic length in the spatial domain” and to the “characteristic time interval in the temporal domain” (37). So, even though space and time are continuous, only a discrete set of scales is of interest based on specific features that make them of particular use or interest (37). The existence of a hierarchy of scales has been postulated to relate to spatial or temporal features of systems of interest (see Fig. 3).

Temporal and spatial scales dictate the general type of model. The consideration of scale in model development requires observed information for the real system being modeled at the spatial and temporal scales of interest, which is to say that microscopic-scale models developed in the laboratory are not appropriate for macroscopic-scale applications, and vise versa.

Models of solute transport in the vadose zone exist at all scales. A hierarchical depiction of the scales from molecular to global showing the relationship between scale and model type is depicted in Fig. 3. An important consideration in model conceptualization is for the model to account for the predominant processes occurring at the

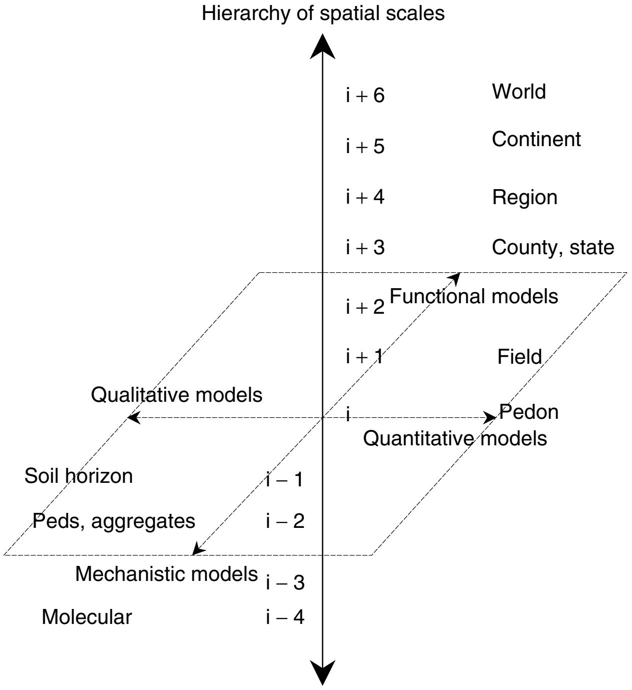

Figure 3. Organizational hierarchy of spatial scales pertinent to NPS pollutant models. *Source:* Ref. 21. With permission.

spatial and temporal scales of interest, which complies with the guideline of parsimony. Qualitatively speaking, as spatial scale increases, the complex local patterns of solute transport are attenuated and dominated by macroscale characteristics. For this reason, mechanistic models are used more frequently at the (i) to (i − 4) scales, whereas functional models are more often applied to scales ranging from (i + 1) to (i + 6). The stochastic application of deterministic models is found at the (i + 1) scale, and stochastic models generally are used at (i + 1) and (i + 2) scales. Statistical models are applied most often at the larger scales, (i + 3) to (i + 6).

The relevance of temporal domain is also a consideration not to be overlooked. Larger spatial scales appear more constant because the rapid dynamics of the lower scales are disregarded (20). For this reason, time steps of functional models can expand over days, such as the time between irrigation or precipitation events, whereas the time steps of mechanistic models characteristically extend over minutes.

The integration of solute transport models of the vadose zone into a GIS provides the ability to dynamically describe NPS pollutant transport at a range of spatial scales allowing the user to rapidly scale "up" and "down." However, this integration introduces incompatibilities between the model and data and raises basic questions regarding (1) the compatibility of the model with input and validation data, and (2) the relevance of the model to the applied spatial scale. Wagenet and Hutson (38) addressed the issue of scale dependency and proposed three scale-related factors to consider when applying GIS-based solute transport models to the simulation of NPS pollutants in soils: (1) the type of model (i.e., functional or mechanistic) must consider the scale of application, and the nature of the available data at that scale; (2) sampling and measurement of input and validation data must be spatially consistent with the model; and (3) measurement and monitoring methods must be relevant at the temporal domain being modeled.

FUTURE DIRECTION

Beven (39) asserts that the real constraint on predictability by landscape-scale environmental models is not the detail of the model structure, but rather the ability to spatially and temporally characterize the variability of model inputs and parameters. Among the various techniques for characterizing the spatial variability of model inputs (e.g., electromagnetic induction, ground penetrating radar, time domain reflectometry, remote imagery, etc.), electromagnetic induction will be the most useful in the short term (29), whereas greater future potential exists for hyperspectral imagery, which is in its infancy for soil and plant science applications.

BIBLIOGRAPHY

1. Corwin, D.L. and Wagenet, R.J. (1996). Applications of GIS to the modeling of non-point source pollutants in the vadose zone: A conference overview. *J. Environ. Qual.* **25**: 403–411.
2. Corwin, D.L. and Loague, K. (In press). Multidisciplinary approach for assessing subsurface non-point source pollution. In: *Soil-Water-Solute Processed in Environmental Systems: An Integrated Approach*. J. Alvarez-Benedi and R. Munoz-Carpena (Eds.). CRC-Press, Boca Raton, FL.
3. Duda, A. and Nawar, M. (1996). Implementing the World Bank's water resources management policy: A priority on toxic substances from nonpoint sources. *Water Sci. Technol.* **33**: 4–5, 45–51.
4. Pimental, D. (Ed.). (1993). *World Soil Erosion and Conservation*. Cambridge University Press, Cambridge, UK.
5. Duda, A.M. (1993). Addressing nonpoint sources of water pollution must become an international priority. *Water Sci. Technol.* **28**: 1–11.
6. Humenik, F.J., Smolen, M.D., and Dressing, S.A. (1987). Pollution from nonpoint sources: Where we are and where we should go. *Environ. Sci. Technol.* **21**: 8, 737–742.
7. World Resources Institute. (1998). *World Resources 1998–99: A Guide to the Global Environment*. Oxford University Press, New York.
8. Beven, K.J. (1989). Changing ideas in hydrology—the case of physically-based models. *J. Hydrol.* **105**: 157–172.
9. Grayson, R.B., Moore, I.D., and McMahon, T.A. (1992). Physically-based hydrologic modeling—II. Is the concept realistic? *Water Resour. Res.* **26**: 2659–2666.
10. Corwin, D.L., Loague, K., and Ellsworth, T.R. (1999). Advanced information technologies for assessing nonpoint source pollution in the vadose zone: Conference overview. *J. Environ. Qual.* **28**: 357–365.
11. Cressie, N.A.C. (1993). *Statistics for Spatial Data*. Wiley-Interscience, New York.
12. Barnes, E.M. et al. (2003). Remote- and ground-based sensor techniques to map soil properties. *Photogramm. Eng. Remote Sens.* **69**(6): 619–630.
13. Burrough, P.A. (1996). Opportunities and limitations of GIS-based modeling of solute transport at the regional scale. In:

Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone. D.L. Corwin and K. Loague (Eds.). SSSA Special Publication No. 48, Soil Science Society of America, Madison, WI, pp. 19–38.

14. Skaggs, T.H. and Leij, F.J. (2002). Solute transport: Theoretical background. In: *Methods of Soil Analysis: Part 4—Physical Methods*. J.H. Dane and G.C. Topp (Eds.). SSSA Book Series No. 5, Soil Science Society of America, Madison, WI, pp. 1353–1380.
15. Corwin, D.L., Vaughan, P.J., and Loague, K. (1997). Modeling nonpoint source pollutants in the vadose zone with GIS. *Environ. Sci. Technol.* **31**: 8, 2157–2175.
16. Altman, D. (1994). Fuzzy set theoretic approaches for handling imprecision in spatial analysis. *Int. J. Geogr. Inf. Syst.* **8**(3): 271–289.
17. Zhu, A.-X. (1999). Fuzzy inference of soil patterns: Implications for watershed modeling. In: *Assessment of Non-Point Source Pollutants in the Vadose Zone*. D.L. Corwin, K. Loague, and T.R. Ellsworth (Eds.). American Geophysical Union, Washington, DC, pp. 135–150.
18. Crawford, J.W., Baveye, P., Grindrod, P., and Rappoldt, C. (1999). Application of fractals to soil properties, landscape patterns, and solute transport in porous media. In: *Assessment of Non-Point Source Pollutants in the Vadose Zone*. D.L. Corwin, K. Loague, and T.R. Ellsworth (Eds.). American Geophysical Union, Washington, DC, pp. 151–164.
19. Bourgault, G. et al. (1997). Geostatistical analysis of a soil salinity data set. *Adv. Agron.* **58**: 241–292.
20. O'Neill, R.V. (1988). Hierarchy theory and global change. In: *Scales and Global Change*. Rosswall et al. (Eds.). Scientific Committee on Problems of the Environment (SCOPE), John Wiley & Sons, New York, pp. 29–45.
21. Hoosbeek, M.R. and Bryant, R.B. (1993). Towards the quantitative modeling of pedogenesis: A review. *Geoderma* **55**: 183–210.
22. Lark, R.M. and Webster, R. (1999). Analysis and eleucidation of soil variation using wavelets. *Eur. J. Soil Sci.* **50**: 185–206.
23. Lark, R.M., Kaffka, S.R., and Corwin, D.L. (2003). Multiresolution analysis of data on electrical conductivity of soil using wavelets, *J. Hydrol.* 272, 276–290, 2003.
24. Wösten, J.H.M. and Tamari, S. (1999). Application of artificial neural networks for developing pedotransfer functions of soil hydraulic parameters. In: *Assessment of Non-Point Source Pollutants in the Vadose Zone*. D.L. Corwin, K. Loague, and T.R. Ellsworth (Eds.). American Geophysical Union, Washington, DC, pp. 235–242.
25. Loague, K. and Corwin, D.L. (1996). Uncertainty in regional-scale assessments of non-point source pollutants. In: *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone*. D.L. Corwin and K. Loague (Eds.). SSSA Special Publication No. 48, Soil Science Society of America, Madison, WI, pp. 131–152.
26. Dane, J.H. and Molz, F.J. (1991). Physical measurements in subsurface hydrology. In: *Reviews of Geophysics*. Supplement, U.S. National Report to International Union of Geodesy and Geophysics 1987–1990, American Geophysical Union, Washington, DC, pp. 270–279.
27. Topp, G.C., Reynolds, W.D., and Green, R.E. (Eds.). (1992). *Advances in Measurement of Soil Physical Properties: Bringing Theory into Practice*. SSSA Special Publication No. 30, Soil Science Society of America, Madison, WI.
28. Corwin, D.L. (1996). GIS applications of deterministic solute transport models for regional-scale assessment of non-point source pollutants in the vadose zone. In: *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone*. D.L. Corwin and K. Loague (Eds.). SSSA Special Publication No. 48, Soil Science Society of America, Madison, WI, pp. 69–100.
29. Corwin, D.L. and Lesch, S.M. (In press). Apparent soil electrical conductivity measurements in agriculture. *Comput. Electron. Agric.*
30. Rawls, W.J., Gish, T.J., and Brakensiek, D.L. (1991). Estimating soil water retention from soil physical properties and characteristics. *Adv. Soil Sci.* **16**: 213–234.
31. van Genuchten, M.Th., Leij, F.J., and Lund, L.J. (Eds.). (1992). Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils. *Proc. Int'l. Workshop, Riverside*, CA, Oct. 11–13, 1989, University of California, Riverside, CA.
32. Loague, K. et al. (1996). Uncertainty of groundwater vulnerability assessments for agricultural regions in Hawaii: A review. *J. Environ. Qual.* **25**(3): 475–490.
33. Goodchild, M.F. (1993). The state of GIS for environmental problem-solving. In: *Environmental Modeling with GIS*. M.F. Goodchild, B.O. Parks, and L.T. Steyaert (Eds.). Oxford University Press, New York, pp. 8–15.
34. Walsh, S.J. (1988). Geographic information systems: An instructional tool for earth science educators. *J. Geogr.* **87**(1): 17–25.
35. Jury, W.A. (1996). Stochastic solute transport modeling trends and their potential compatibility with GIS. In: *Application of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone*. D.L. Corwin and K. Loague (Eds.). SSSA Special Publication No. 48, Soil Science Society of America, Madison, WI, pp. 57–67.
36. Jury, W.A. (1986). Spatial variability of soil properties. In: *Vadose Zone Modeling of Organic Pollutants*. S.C. Hern and S.M. Melancon (Eds.). Lewis Publishers, Chelsea, MI, pp. 245–269.
37. Baveye, P. and Boast, C.W. (1999). Physical scales and spatial predictability of transport processes in the environment. In: *Assessment of Non-Point Source Pollutants in the Vadose Zone*. D.L. Corwin, K. Loague, and T.R. Ellsworth (Eds.). *Geophysical Monograph* 108, American Geophysical Union, Washington, DC, pp. 261–280.
38. Wagenet, R.J. and Hutson, J.L. (1996). Scale-dependency of solute transport modeling/GIS applications. *J. Environ. Qual.* **25**(3): 499–510.
39. Beven, K.J. (2002). Towards an alternative blueprint for a physically-based digitally simulated hydrologic response modeling system. *Hydrol. Process.* **16**(2): 189–206.

MODELING TECHNIQUES FOR SOLUTE TRANSPORT IN GROUNDWATER

Diganta Bhusan Das
Oxford University
Oxford, United Kingdom

Navraj S. Hanspal
Loughborough University
Loughborough, United Kingdom

INTRODUCTION

Groundwater constitutes an important component of water supply systems for domestic use, industry, and agriculture. Due to increased use regularly above sustainable levels, groundwater quality often deteriorates. Contaminants in groundwater systems are introduced by hazardous industrial wastes, leachate from landfills, oil spills, agricultural activities, sites of radioactive waste repositories, and other sources. In most cases, the source of contamination is aboveground, and the contaminants pass through the unsaturated zone on their way to the saturated region.

The fundamental processes that govern solute mobility in groundwater have been studied and reasonably well understood for some time. Solute migration results from a variety of coupled physical, chemical, and biological processes. Owing to the complexities of natural soils and solute–soil–fluid interactions and frequently changing field conditions, modeling solute transport involves some degree of approximation. However, the modeling framework must retain the important aspects of real phenomena. It is also necessary for a reliable/realistic model to answer several relevant questions, such as, what is the purpose of the predictive calculations, what types of environments can be modeled, what level of precision is expected, and what are the advantages and disadvantages of the developed approach. To answer these questions while developing a model for solute movement in the subsurface, many issues must be considered. Table 1 presents some examples of such issues.

Soil is a porous medium, composed of solid and void sections (Fig. 1). The void portion that forms the *effective pore space* is made up of many interconnected channels through which fluid convection takes place. The usual way of modeling solute transport in soil has been to adopt a macroscopic approach where the flow variables are averaged over a *representative elementary volume* (REV) or *representative elementary area* (REA), depending on the dimensionality of the problem domain (1–3). The length scales of REV and REA are selected so that they represent soil heterogeneity. They are much larger than the pore scale but much smaller than the main physical domain. They are determined under the assumption that the solid matrix is distributed uniformly throughout the porous medium domain and is present in each REV or REA. Furthermore, it is assumed that the specific surface of the solid matrix is larger than that of the void portion in a REV or REA.

Convection of a solute in soil takes place due to fluid movement through its pores. As this transport occurs, in most cases, the carrier fluid is a multiphase/multifluid medium (a mixture) that can be miscible or immiscible. If two or more miscible fluids occupy the void space, even if they occupy different regions in the domain initially, they ultimately produce a multicomponent mixture due to gradients (e.g., diffusive and dispersive effects) prevailing over the fluid components. In the case of immiscible fluid flow, distinct layers are observed and interphase mass transfer among the fluid components takes place. Therefore, modeling immiscible multifluid convection of a solute requires more rigorous approaches. Models of single-phase flow can be considered a limiting case of multiphase flow phenomena. The main interest in a single-phase model lies in its simplicity and the methods developed for it which can be extended to multiphase transport mechanisms. Greenkorn (4) and Nield and Bejan (5) have given good descriptions of multicomponent, multiphase flow through porous media. Numerous other studies have also been devoted to modeling two-phase (the simplest multiphase case) or multiphase flow phenomena in soil/porous media (6–9).

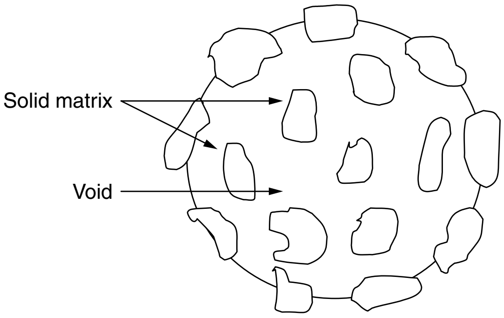

Figure 1. A representative elementary volume (REV) showing solid and void matrices in a porous medium.

Good management of water resources requires the ability to forecast the response of the managed system, for example, an aquifer, to planned operations such as pumping, recharging, and controlling conditions at aquifer boundaries. Any planning of mitigation, cleanup, or control measures requires forecasting the path and fate of contaminants in both the unsaturated and the saturated aquifer. The preferred tool for prediction is a numerical model that simulates the groundwater flow and pollutant movement and any chemical transformation. The construction of good models should be based on a thorough understanding of what happens within the modeled domain and on its boundaries, including chemical and biological processes. We need the ability to express this information as well-posed mathematical statements and, eventually, as a solvable numerical model. The model developed can be solved by employing a suitable numerical technique and appropriate computer programs. Many such programs for flow and solute transport in the subsurface are now available. This article introduces the fundamental concepts for modeling solute transport in groundwater systems. These are particularly useful for developing any remediation method for a contaminated subsurface.

SCALES OF OBSERVATION

Four spatial scales of observation have been suggested for describing solute transport phenomena in the subsurface:

1. The molecular scale, where materials are viewed as a discrete collection of molecules
2. The microscale or pore-scale, which describes solute behavior within a single pore in a porous medium
3. The macroscale or core-scale, where porous medium properties such as porosity, fluid phase saturation, and permeability can be described as continuous functions of space
4. The megascale, which is the scale of a system.

Table 1. Some Important Issues Involved in Solute Mobility in Groundwater Systems

S.N.	Issues	Details	Remarks
1.	Sources of pollutants/chemicals/solute in the ground	1. Hazardous and toxic waste disposal sites 2. Possible leaks in the compartments in ground used for storing toxic chemicals 3. Groundwater movement 4. Interphase mass transfer from air to soil 5. Dry deposition of contaminants from polluted air 6. Wet deposition on land 7. Accidental release of chemicals in the environment	Along with sources of pollutants, the amount of substance available in soil is equally important
2.	Properties of solute	1. Molecular structure 2. Vapor pressure 3. Water solubility 4. Absorption spectra 5. Particle size of particulate matter 6. Rate constants for its various reactions in the aquatic environment 7. Reactive or nonreactive 8. Life of the solute in the environment or its stability 9. Toxicological characteristics of the solute	
3.	Form of the pollutant in the environment	1. Dissolved in water 2. As suspension or colloid 3. As ions 4. As particulate matter	
4.	Direction of flow	1. One-dimensional (1-D) 2. Two-dimensional (2-D) 3. Three-dimensional (3-D)	In almost all cases, flow takes in a three-dimensional domain, but, for ease of modeling, a simpler geometry is often assumed
5.	Nature of transport mechanism	1. Saturated flow (single-phase flow) 2. Unsaturated flow (e.g., vadose zone) 3. Two-phase flow 4. Multiphase flow (more than two phases)	
6.	Soil type	1. Homogeneous soil 2. Heterogeneous soil (e.g., randomly heterogeneous, fractured media) 3. Layered soil and the no. of layers 4. Aggregated soil composed of large lumps	
7.	Soil characteristics	1. Porosity 2. Permeability 3. Dispersibility 4. Tortuosity 5. Density 6. Capillarity 7. Connectivity 8. Moisture content 9. Surface area 10. Depth 11. pH 12. Turbulence 13. Temperature 12. Absorption spectra 13. Proportion of diffusible ion	

Table 1. (*Continued*)

S.N.	Issues	Details	Remarks
8.	Transport and fate mechanisms in the soil	1. Diffusion: random motion of ions, atoms, molecules in all phases 2. Hydrodynamic dispersion: primarily from the complexities of the pore system 3. Adsorption: the passage of chemical species from one bulk phase to the surface of another where it accumulates without penetrating the structure of the second 4. Ion exchange 5. Hydrolysis 6. Photolysis 7. Volatilization 8. Microbial activity in the soil, e.g., anaerobic reaction 9. Relative velocity of a chemical with respect to moving water—sometimes the moving contaminants may be retarded with respect to moving water in soil for the following reasons: (a) ionic exchange reaction (b) precipitation and dissolution reaction (c) generation of insoluble complexes (d) formation of colloids followed by flocculants	
9.	Significance of groundwater	1. Inflow, outflow, downflow, and overflow of water from a control volume of soil 2. Infiltration and percolation of surface water 3. Level of groundwater table rises from rainfall 4. Groundwater draft from pumping of water 5. Carrying capacity of solute in groundwater	

To design soil and groundwater remediation techniques and to make reliable predictions about the efficiency of these techniques, it is necessary to identify and understand flow and transport processes on a smaller (e.g., pore scale) scale and to describe their manifestation on a larger scale (e.g., core scale). The connection of a smaller scale to a larger scale can be understood by a suitable upscaling technique. The choice of an appropriate upscaling method generally depends on the scale of observation because the relative importance of the dominant forces (e.g., capillary, viscous, and gravitational effects) may change with change in the averaging scale.

In principle, the mathematical model that describes the transport of an extensive quantity of a fluid or solid phase through a porous medium (mass balance equations for the transported quantities considered constitutive relations, initial conditions within the phase, and boundary conditions on the surface that bounds that phase) can be stated at every point within the phase considered. This description is said to be on the microscopic level, as we focus our attention on what happens at a (mathematical) point within the phase considered in the domain. Although the transport problem can be stated as a well-posed model, it is difficult to solve on this scale because the detailed geometry of the surface that bounds the phase is not known or is too complex. It is also very difficult to measure the values of variables at points within a phase to validate a flow or transport model and to determine the model parameters. As a consequence, the complete description and solution of a transport problem on a microscopic level is still unavailable.

To circumvent these difficulties, the transport problem is transformed from a microscopic to a macroscopic level, at which the problem is reformulated in terms of average values of the microscopic values. The average values are measurable quantities. We refer to this approach, employed in many branches of science, as the continuum approach.

SUBSURFACE HETEROGENEITY CHARACTERIZATION

The most typical features of groundwater systems are the spatial heterogeneity of the media and the temporal variability of the flow and transport properties. Moreover, these processes occur on multiple of scales. The mathematical descriptions of these processes on different scales are not necessarily similar. Scaling theories have been developed that attempt to quantify the links among process descriptions on various scales. However, due to spatial heterogeneity and temporal variability and the high nonlinearity of hydrologic processes, developing appropriate scaling rules is not trivial. Sometimes, new governing equations and constitutive relations to describe the flow and transport processes in groundwater have to be identified. This may yield a number of upscaled parameters in the governing equations that may have to be measured directly or indirectly. These parameters are either new, meaning they exist on only the larger scale or are averages of small-scale parameters. Questions that may arise in this regard are (1) how can parameter values measured on a small scale (e.g., laboratory experiments) be translated to effective values on a larger scale? (2) how can effective parameters of the larger scale be measured directly? (3) how can small-scale heterogeneities be accounted for in large-scale descriptions of hydrologic systems? (4) how can the large-scale behavior of a system be downscaled to its behavior on a smaller scale?

During the last 60 years, some methods have been developed to describe spatial heterogeneities in the subsurface. These methods cannot generate actual heterogeneity patterns, but they provide a theoretical basis for understanding the effects of heterogeneities and any coupled processes therein on flow and transport processes. Three types of methods have evolved:

1. descriptive methods
2. process-based methods
3. stochastic methods

Each method has its own limitations in theory and application. The descriptive method is limited mostly to large-scale solute transport problems. It cannot describe local variations in hydraulic permeability (K) and is not transferable to deterministic numerical models for flow and transport calculations. Process-based models cannot recreate the depositional faces or the experimental values of hydraulic permeability (K) at observation points. This method also requires data that are often not available, for example, sedimentation and erosion rates, and turbidity factors. Further, process-based models are computationally inefficient in computer time and storage and are highly sensitive to initial and boundary conditions. Interpretation and validation of results are therefore very subjective. Stochastic models are based on the principles of geostatistics and work well in theory. For example, they honor known data and their locations and patterns of spatial variability. Data for these models can be obtained from boreholes, soil survey maps, or any other sampling methods. However, they require an enormous amount of observational data for calculating the spatial statistics of the K distribution, and these data are generally lacking, especially in lateral directions. Therefore, one question that frequently arises in using stochastic models is, how can we estimate spatial variability in three directions from vertical data sets? Therefore, even with available methods, it is still difficult to create a reliable 3-D model that honors the important heterogeneous features at a given location.

An alternative to these approaches for characterizing heterogeneity is to use a Markov-chain approach. This is a hybrid method of deterministic and stochastic models. The earliest attempt to apply Markov chains was used to synthesize stratigraphic sequences in the vertical direction (10). However, the approach was not directed toward extension to multidimensions, and it was not used in real field situations for a long time. Later, a methodology (11) for characterizing subsurface heterogeneity based on a two-dimensional extension of the original one-dimensional Markov chain theory was developed. This model, called the coupled Markov-chain (CMC) model (11), couples two chains for 2-D problems and three chains for 3-D problems. For example, for a 2-D problem, the first chain describes the sequence in lithologies in the vertical direction, and the second chain describes the sequences of variation in the lithological structure in the horizontal direction. The two chains are coupled in the sense that a state of a cell in the domain is dependent on the state of two cells, the one on top and the other on the left of the current cell. This dependence is described in terms of transition probabilities from the two chains. The CMC is a general technique that can produce a variety of heterogeneous patterns that can be stationary or nonstationary in statistical sense.

GOVERNING MODEL EQUATIONS

The equations governing fluid flow and solute transport in groundwater are based on the principles of mass, momentum, and energy conservation. They can be expressed in two different formulations based on the dependent variables used: first, vorticity stream-function equations that are derived by incorporating the definitions of vorticity and stream-functions; second, primitive variable formulations that are expressed in terms of pressure and velocity or mass and momentum of the fluid particles. Furthermore, the governing equations of the fluid flows can be classified mathematically as elliptic, hyperbolic, or parabolic equations. Based on the physical problem under consideration, these model equations can be solved either in rectangular or curvilinear coordinate frameworks.

The use of the vorticity stream-function formulation is most commonly found in one- and two-dimensional cases. Although it is possible to extend vorticity-stream functions to 3-D domains, they give rise to many complications and are not often recommended (12).

Primitive variable formulations can be expressed in either conservative or nonconservative forms, which in turn can be dimensional or nondimensional. The conservative forms of equations derived are based on the assumption that the fluid model is fixed in space. Hence, new fluid particles interact across the domain boundaries as the fluid flows in and out. For nonconservative forms, the observer is assumed to move with the frame of the fluid particles. In other words, the equations are formulated based on a moving coordinate system. Therefore, for a nonconservative system, there is no interaction of fluid elements across boundaries. For a complete discussion of conservative and nonconservative forms of equations, readers are referred to Anderson (13).

Governing equations can also be expressed in either Eulerian or Lagrangian forms. These two formulations are similar to conservative and nonconservative equations, respectively, because they assume that the fluid model is either fixed in space or moving with the flowing fluid. However, as Abbott and Basco (14) have described, important differences exist between Eulerian and conservative forms. Similar differences are also applicable to nonconservative and Lagrangian formulations. An important characteristic of Eulerian equations is that the accelerations of fluid particles form substantial or total derivatives of each velocity component. They are expressed in terms of continuous spatial velocities and pressures. Therefore, when the flow variables are discontinuous, this formulation suffers serious drawbacks. Conservative approaches remove such limitations because they are formulated by assuming mass and momentum of the fluid particles as the dependent variables and as functions of space and time. They are represented as integral forms. This makes any discontinuity of the flow variables irrelevant in the computational domain. Conservative approaches, therefore, provide a more general way of formulating fluid flow phenomena.

The primary issue that comes up in formulating groundwater flow problems is selection of proper mathematical forms for the governing equations. Solute transport in porous domains can be represented by the convective dispersion equation (Eq. 1). The velocity profiles can be obtained by solving a suitable flow equation, for example, the Darcy or Brinkman equation. Equation 1 does not include any source/sink terms, which might become evident if there are any reactions in subsurface flow problems.

Solute Transport Equation

$$\text{Convective dispersion equation: } \frac{\partial C}{\partial t} + \overline{\nabla}.(C\mathrm{v}) = \overline{\nabla}.(D\overline{\nabla}\overline{v}) \tag{1}$$

D in Eq. 1 is the dispersion coefficient tensor and can be represented as

$$D = \begin{bmatrix} D_{xx} & 0 & 0 \\ 0 & D_{yy} & 0 \\ 0 & 0 & D_{zz} \end{bmatrix} \tag{2}$$

where D_{xx}, D_{yy}, and D_{zz} are the principal components of the dispersion coefficient tensor in the x, y, and z directions.

Flow Equations

Various linear and nonlinear relationships describe flow behavior in the subsurface (5). The two most popular flow equations are the Darcy (Eq. 3) and the Brinkman (Eq. 4) equations.

Assuming that the fluid is incompressible and neglecting any body forces, the governing momentum balance equations can be written as

$$\text{Darcy's equation:} \qquad \overline{\nabla}P = -\frac{\mu}{\overline{K}}\overline{v} \tag{3}$$

$$\text{Brinkman equation:} \qquad \overline{\nabla}P = -\frac{\mu}{\overline{K}}\overline{v} + \mu'\overline{\nabla}^2\overline{v} \tag{4}$$

K in Eqs. 3 and 4 represents anisotropic soil permeability and is a second-order tensor that can be represented as

$$K = \begin{bmatrix} K_{xx} & 0 & 0 \\ 0 & K_{yy} & 0 \\ 0 & 0 & K_{zz} \end{bmatrix} \tag{5}$$

where K_{xx}, K_{yy}, and K_{zz} are the principal components of the permeability tensor in the x, y, and z directions. Hence, the diagonal form of the anisotropic permeability is a special case which occurs when the coordinate axes coincide with the normal axes of K. In general, these components are functions of the fluid and material properties. Due to the varying range of particle sizes and shapes and complex fluid–solid interactions in field conditions, there is a lack of theoretical corelations for estimating the permeability in the subsurface.

If the soil is isotropic, then,

$$K_{xx} = K_{yy} = K_{zz}$$

Equations 3 and 4 are coupled with the continuity equation (Eq. 6), and then solved to determine the unknown field variables, velocities and pressures.

For an incompressible fluid and nondeformable porous medium, the continuity equation for conservation of mass can be written as

$$\text{Continuity equation:} \qquad \overline{\nabla}.\overline{v} = 0 \tag{6}$$

NUMERICAL TECHNIQUES FOR SOLVING GOVERNING MODEL EQUATIONS

There are mainly five different numerical approaches for modeling solute transport problems in groundwater: the finite-difference, finite-element, finite-volume, boundary-element, and spectral-expansion. They have been employed for solving many different complex fluid problems with reasonably good success (15–19). Each method was developed to handle a specific problem, to eliminate the limitations of an existing technique, or simply, as a variant form. For example, the finite-element method was developed mainly because of its ability to cope with complex geometry. Similarly, the spectral method was developed as a variant of the finite-element method (15) to enhance the accuracy of solutions. The techniques have their own advantages and disadvantages, and many similarities and dissimilarities exist among their concepts and procedures. Comparison among these methods has been made by Hirsch (16), Versteeg and Malalasekera (17), Comini et al. (18) and Garg (19), among many other authors. The choice of the technique to be used in modeling depends on all these criteria, apart from its applicability to the particular problem.

Finite-Difference Method

The procedure for deriving finite-difference equations consists of approximating the derivatives in the differential equations via a truncated Taylor series expansion. Let us consider the grid points shown in Fig. 2. For grid point

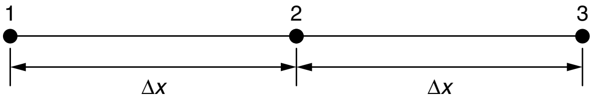

Figure 2. Computational grid for a Taylor series expansion.

2, located midway between grid points 1 and 3 such that $\Delta x = x_2 - x_1 = x_3 - x_2$, the Taylor series expansion gives

$$\varphi_1 = \varphi_2 - \Delta x\left(\frac{d\varphi}{dx}\right)_2 + \frac{1}{2}(\Delta x)^2\left(\frac{d^2\varphi}{dx^2}\right)_2 - \ldots \quad (7)$$

$$\varphi_3 = \varphi_2 + \Delta x\left(\frac{d\varphi}{dx}\right)_2 + \frac{1}{2}(\Delta x)^2\left(\frac{d^2\varphi}{dx^2}\right)_2 + \ldots \quad (8)$$

Truncating the series just after the third term and adding and subtracting the two equations, we obtain

$$\left(\frac{d\varphi}{dx}\right)_2 = \frac{\varphi_3 - \varphi_1}{2\Delta x} \quad (9)$$

$$\left(\frac{d^2\varphi}{dx^2}\right)_2 = \frac{\varphi_1 + \varphi_3 - 2\varphi_2}{(\Delta x)^2} \quad (10)$$

Substituting such expressions in the differential equations leads to a finite-difference equation, which can be cast in matrix notation and then solved using matrix solvers. For advanced literature on finite-difference methods, readers may refer to (20–23).

Finite-Element Method

This method is based on the principles of variational calculus. The differential equation to be solved can be represented as

$$L(\varphi) = 0 \quad (11)$$

Further, we assume an approximate solution for the dependent variable $\overline{\varphi}$ that contains a number of undetermined parameters, such as $a_0, a_1, a_2, \ldots, a_m$:

$$\overline{\varphi} = a_0 + a_1x + a_2x^2 + \ldots\ldots\ldots\ldots\ldots\ldots + a_mx^m \quad (12)$$

Substituting $\overline{\varphi}$ into the differential equation leaves a residual R, defined as

$$R = L(\overline{\varphi}) \quad (13)$$

We wish to make this residual small over the entire computational domain, and hence it can be proposed that

$$\int WR\,dx = 0 \quad (14)$$

where W is a weighting function and the integration is performed over the computational domain. By choosing a succession of weighting functions, we can generate as many equations as are required for evaluating the parameters. For simple regular grids, the equations derived by using the finite-difference and finite-element methods are identical. The computational meshes are generally more refined in the areas of rapid change of dependent variables. For complex problems, the finite-element method has several advantages. First, the boundary conditions are handled naturally by this method in contrast to the finite-difference method, where special formulas must be developed for the boundaries in many instances. Second, inhomogeneities and anisotropies can be taken into account easily. Moreover, the size and shape of the elements can be varied readily, and for curved geometries, the concept of isoparametric mapping can be employed to yield robust solutions. This method, developed by Zienkiewicz (24–27), was applied to structural engineering problems, and later on, was extended to subsurface hydrology.

Finite-Volume Method

The primary steps involved in the finite-volume method are grid generation to divide the domain into discrete control volumes. The governing equations are then integrated over the control volume to yield a set of discretized equation at nodal points. The integral form of the equations allows for discontinuities inside the fixed volume unlike the differential forms of governing equations, which assume that the flow properties are differentiable, hence continuous. This consideration becomes useful when dealing with flow problems that have real discontinuities encountered in subsurface flow. For a detailed text on the finite-volume method, refer to (17).

Boundary-Element Method

The boundary-element method is derived by discretizing an integral equation defined on the boundary of the domain and relates the boundary solution to the solution at points in the domain. The integral equation is commonly referred as the boundary integral equation (BIE) and the method as the boundary-element method (BEM). An integral equation formulated on the boundary can be derived only for certain classes of PDEs. Hence, the BEM is not widely applicable compared to the finite-element and finite-difference methods. The advantages of the boundary-element method are that only the boundary of the domain needs to be discretized. The dimension of the problem is effectively reduced. For example, an equation governing a three-dimensional region is transformed into one over its surface. When the domain is exterior to the boundary, as in potential flow past an obstacle, the extent of the domain is infinite, and the equation governing the infinite domain is reduced to an equation over the boundary. Thus it results in a numerical method that is easier to use and more computationally efficient than the competing methods. One can refer to (28) and (29) for detailed information on the BEM.

Spectral Expansion Method

Spectral methods are algorithms for solving certain classes of PDEs using the fast Fourier transform. Finite-difference, finite-element, and spectral methods can be distinguished on the basis of weighting functions used in the mathematical formulations. In the finite-difference method, the trial functions are polynomials of low order; in the finite-element method, they are local smooth functions (i.e., polynomials of fixed/higher order), whereas in spectral methods, they are global smooth functions such as Fourier series expansions. All spectral methods are based on the complete family of global smooth functions. Detailed discussions on the development of various spectral expansion techniques based on selecting different weighting functions can be found in (30–32).

CONCLUSIONS

The study of solute transport in groundwater systems is a multidisciplinary subject because it integrates the principles of engineering, pure science subjects, for example, applied mathematics, chemistry and physics and, other life science subjects, for example, biology. The physical, chemical, and biological mechanisms that govern solute transport in the subsurface interact in a highly complex manner, and elucidating the important or dominant control mechanisms that govern transport processes can be highly nonintuitive. On the other hand, without thorough knowledge of the governing processes and their interactions, accurate prediction and control of solute migration is severely restricted. For this reason, many experiments on both laboratory and field scales have been attempted. The knowledge gained from such experiments (benefits) is clearly evident in the scientific literature now. However, it is also pivotal to place experimental work within a consistent and coherent theoretical framework to interpret the results obtained and direct further experimental work appropriately. By examining a system in terms of suitable equations with well-defined parameters, a mathematical model can help to provide a route to such a framework.

In the text presented, the important concepts related to modeling solute migration in groundwater systems have been presented. We have discussed the various scales of observation over which solute migration may take place. The ways to characterize the ubiquitous media heterogeneities in the subsurface have been briefly mentioned. Basic equations that govern solute transport behavior in the subsurface and numerical methods to solve them have been presented. These are far from exhaustive analyses. However, it is envisaged that the text gives an overview of various basic underlying concepts for modeling solute transport in groundwater.

NOMENCLATURE

C	Solute concentration
D	Dispersion coefficient tensor
D_{xx}	x component of the dispersion coefficient tensor
D_{yy}	y component of the dispersion coefficient tensor
D_{zz}	z component of the dispersion coefficient tensor
K	Permeability tensor
K_{xx}	x component of the permeability tensor
K_{yy}	y component of the permeability tensor
K_{zz}	z component of the permeability tensor
P	Pressure
t	Dimensional time
v	Velocity vector
xx	x coordinate
yy	y coordintate
zz	z coordinate
Φ	Variable
μ	Fluid viscosity
μ'	Effective fluid viscosity

BIBLIOGRAPHY

1. Chinanzvavana, S., Grosshandler, W.L., and Davis, D.C. (1986). *Trans. ASAE* **29**(6): 1797–1801.
2. Jo, J.C., Shin, W.K., and Choi, C.Y. (1999). *Numerical Heat Transfer Part B* **36**: 95–113.
3. Kazmierczak, M., Poulikakos, D., and Pop, I. (1986). *Numerical Heat Transfer* **10**: 517–571.
4. Greenkorn, R.A. (1983). *Flow Phenomena in Porous Media*. Marcel Dekker, New York.
5. Nield, D.A. and Bejan, A. (1992). *Convection in Porous Media*. New York
6. Slattery, J.C. (1970). *AIChE J.* **16**(3): 345–352.
7. Madsen, J.H. and Jensen, K.H. (1992). *J. Hydrol.* **135**: 13–52.
8. Butts, M.B. and Jensen, K.H. (1996). *J. Hydrol.* **183**: 101–116.
9. Wang, C.Y. (1997). *Numerical Heat Transfer Part B* **31**: 85–105.
10. Krumbein, W.C. (1967). *Fortran computer programs for Markov chain experiments in geology: Computer contribution 13*. Kansas Geological Survey, Lawrence, KS.
11. Elfeki, A.E., Dekking, F.M., Bruining, J., and Kraaikamp, C. (2002). Influence of fine-scale heterogeneity patterns on large-scale behaviour of miscible transport in porous media. *Pet. Geoscience* **8**: 159–165.
12. Hoffman, K.A. and Chiang, S.T. (1993). *Computational Fluid Dynamics for Engineers*. Engineering Education System, Wichita, KS, Vols. I and II.
13. Anderson, J.D., Jr. (1995). *Computational Fluid Dynamics: The Basics with Applications*. McGraw-Hill, New York.
14. Abbott, M.B. and Basco, D.R. (1989). *Computational Fluid Dynamics—An Introduction for Engineers*. Longman Group, Harlow, UK.
15. Peyrot, R. and Taylor, T.D. (1983). *Computational Methods for Fluid Flow*. Springer, Berlin.
16. Hirsch, C. (1990). Numerical computation of internal and external flows. Vol. 2. *Computational Methods for Inviscid and Viscous Flows*. John Wiley, Chichester, UK.
17. Versteeg, H.K. and Malalasekera, W. (1995). *An Introduction to Computational Fluid Dynamics: The Finite Volume Method*. Wesley Longman, Essex, UK.
18. Comini, G., Malisan, M., and Manzan, M. (1996). *Numerical Heat Transfer Part B* **29**: 43–60.
19. Garg, V.K. (1998). *Applied Computational Fluid Dynamics*. Marcel Dekker, New York.
20. Shaw, F.S. and Southwell, R.V. (1941). *Proc. R. Soc.* **A178**: 1–17.
21. Forsythe, G.E. and Wasow, W.R. (1960). *Finite Difference Methods for Partial Differential Equations*. John Wiley & Sons, New York.
22. Richtmeyer, R.D. and Morton, K.W. (1967). *Difference Methods for Initial Value Problems*. Wiley-Interscience, New York.
23. Saulyev, V.K. (1964). *Integration of Equations of Parabolic Type by the Method of Nets*. G.J.Tee (Trans). Macmillan, New York.

24. Zienkiewicz, O.C. and Cheung, Y.K. (1967). *The Finite Element Method in Structural and Continuum Mechanics*. McGraw-Hill, London.
25. Zienkiewicz, O.C. and Taylor, R.L. (1997). *The Finite Element Method: The Basis*. Vol. 1. Butterworth Heinemann, Oxford.
26. Zienkiewicz, O.C. and Taylor, R.L. (1998). *The Finite Element Method: Solid Mechanics*. Vol. 2. Butterworth Heinemann, Oxford.
27. Zienkiewicz, O.C. and Taylor, R.L., (2000). *The Finite Element Method: Fluid Dynamics*. Vol. 3. Butterworth Heinemann, Oxford.
28. Chen, G. and Zhou, J. (1992). *Boundary Element Methods*. Academic Press, San Diego, CA.
29. Hall, W.S. (1994). *The Boundary Element Method*. Kluwer Academic, Dordrecht, the Netherlands.
30. Gottlieb, D. and Orszag, S.A. (1977). *Numerical Analysis of Spectral Methods*. Society for Industrial and Applied Mathematics, Philadelphia, PA.
31. Canuto, C., Hussaini, M.Y., Quarteroni, A., and Zang, T.A. (1988). *Spectral Methods in Fluid Dynamics*. Springer-Verlag, Berlin.
32. Mercier, B. (1989). *An Introduction to the Numerical Analysis of Spectral Methods*. Springer-Verlag, Berlin.

AMBIENT GROUNDWATER MONITORING NETWORK STRATEGIES AND DESIGN

Douglas C. Towne
Phoenix, Arizona

WHY CREATE AN AMBIENT GROUNDWATER NETWORK?

Expanding populations and associated development are creating greater water demands worldwide. Groundwater is increasingly being used to meet these needs because surface water sources are often over-appropriated, heavily polluted, or are preserved for their ecological benefits. With significant future groundwater withdrawal increases, the long-term sustainability of groundwater resources is a vital issue. Yet even where there appears to be sufficient groundwater supplies to provide for irrigation, municipal, domestic, livestock, industrial, and/or mining purposes, the quality of this resource may limit the applications for which it can be used.

Groundwater quality can be negatively impacted by a variety of physical features and cultural activities that are not always easily discerned by empirical observations. The detection and continued monitoring of groundwater quality constituents and contaminants requires expensive field sampling operations and laboratory analyses. Once deterioration of groundwater quality begins, it may be difficult to reverse, and treating the impacted water can be prohibitively expensive.

Assessing groundwater is a challenging endeavor as the resource resides beneath the land surface, available only for examination at limited sites where wells or springs are located. It is a complex, large-scale, three-dimensional resource that has a high degree of spatial variability (3). To meet these challenges are the benefits supplied by ambient groundwater monitoring, characterization, and assessment networks that address these groundwater quantity and quality concerns. The term “ambient” refers to broad-scale or basinwide conditions not associated with a specific point source, facility, or property. The term “monitoring” refers to point measurement of water level or water quality conditions of groundwater sites, and “characterization” is the description and interpretation of the hydrogeologic setting where the groundwater occurs. Finally, the term “assessment” is used to describe efforts that combine both characterization and monitoring (3).

An ambient network is a proactive strategy that uses groundwater monitoring and characterization along with scientific methods and hydrologic models to assess holistically the groundwater of a given area, usually a defined hydrologic basin. Ambient networks are designed to identify regional changes in groundwater levels and/or degradation of groundwater quality such as occurs from widespread nonpoint sources of pollution or a high density of point sources. Information gleaned from an ambient network includes determining where groundwater levels are fluctuating, locating areas of poor water quality, and identifying spatial and temporal water quality patterns. When correctly designed, the data collected in ambient monitoring programs can be statistically analyzed to make broad statements about the groundwater of a basin that have a known degree of confidence using a relatively small number of investigated sites. Several states such as Arizona and Idaho (1) have redesigned their ambient networks to reflect a statistical, rather than subjective, selection of sites. Others such as Minnesota (2) initiated new ambient networks because the previous design failed to meet objectives.

Although planning and implementing ambient networks is not inexpensive, their costs may be considered relatively minor compared to the hydrologic information they can provide when properly designed and executed. On a regional scale, such networks can determine whether aquifers are in hydrologic equilibrium at current pumping levels, which areas should not be used for municipal or domestic uses without treatment, and whether groundwater quality is being impacted over time by anthropogenic activities. On a more localized scale, networks provide valuable background groundwater level and quality information for site-specific hydrologic studies and projects that may involve remediating contaminated sites or siting production wells. Finally, networks will assist in various surface water investigations because discharges from groundwater systems often sustain river base flow and often directly influence the quality of riparian and wetland ecosystems (3).

OVERVIEW

At their zenith, ambient groundwater networks balance sometime divergent entities such as science and art, theoretical statistical considerations and pragmatic field decisions, and precise analytical laboratory tests and empirical field observations, to assess best the groundwater of a specified area. These networks provide a structured yet flexible framework from which to survey a large basin which, if desired, then can transit into smaller, localized networks to examine groundwater concerns in more specific areas.

CONCEPTUAL STRATEGIES AND DESIGNS—GENERAL OVERVIEW

When creating an ambient groundwater network, attention should focus on adhering to the general principles of probability sampling, so that inferences can then be made about the total population from data collected on the sampled population (4). Thus, statements of the network's goals and methods and which randomized site selection will be used, should be included in the planning. Four different strategies for creating an ambient network are detailed: simple random monitoring, stratified random monitoring, random monitoring within grid cells, and transect monitoring. Each strategy has various strengths and weaknesses but each method should be considered a tool to focus objectives of the study and to minimize potential sources of error rather than to dictate rigid rules blindly (4).

Simple Random monitoring

The easiest method of implementing an ambient groundwater monitoring network is simple random selection of monitoring sites within a basin. This network design is created by random selection, using either a computer or table of random numbers, in which a subset sample of sites is selected for monitoring. This design has the benefit that every potential groundwater monitoring site has an equal chance of being selected for the network. However, this design has the major drawback of being unlikely to achieve a set of monitoring sites that are spatially distributed throughout the basin. Commonly, potential measurement sites are clustered in areas of extensive groundwater development, which may lead to biases in the data.

Stratified Random Monitoring

Stratified random monitoring is a viable alternative to simple random monitoring, particularly where information is available to subdivide a basin into subunits that are likely to be more homogenous than the basin as a whole. The information used to subdivide the basin can be somewhat subjective, though stratification tends to be most effective when the subunits are very dissimilar from one another in combination with minimal variability within each subunit. Stratified random monitoring might be used when, for example, a basin has three distinct physiographic areas, and each has a different land use. These areas include an upland area primarily used for low-intensity grazing, a valley primarily used as farmland irrigated by surface water, and a mesa primarily used for farmland irrigated by groundwater. In this scenario, an important facet of the ambient groundwater monitoring network would be to distinguish water level and quality differences among the areas. Each subunit is then monitored individually by simple random selection, and monitoring sites are equally divided among the subunits. The goal of this monitoring strategy continues to be overall assessment of the basin, but an important factor is the assessment of each individual subunit. Idaho uses stratified random monitoring for its ambient network by dividing the state into 22 subareas characterized by fairly homogenous hydrogeology (1).

Random Monitoring Within Grid Cells

The most common method of implementing an ambient groundwater monitoring network is overlaying a grid on the basin (or other area of hydrologic interest), and randomly monitoring a site within each cell created by the grid. The systematic element requires that the selected sites be spatially distributed throughout the area; the random element ensures that every well within a cell has an equal chance of being monitored. The systematic grid monitoring strategy is especially effective in relatively homogenous areas as a tool to distribute sample sites adequately spatially.

More sophisticated systematic grid monitoring models aided by computer software may divide the basin into equal area cells so that each contains three random sample points (5). The random points are designated primary, secondary, and tertiary, respectively. An appropriate well or spring is then located within a specified radius of the primary sample point to monitor. If no appropriate groundwater sample site can be found near the primary point, a similar process is used for the secondary point and then the tertiary point, if necessary. If no appropriate groundwater sites can be located around any of these three random points, then any appropriate well or spring within the cell can be randomly selected for monitoring.

Random monitoring using grid cells can be adapted to more complex or comprehensive groundwater assessments. Basins that have several overlying aquifers (for example, an upper, unconfined aquifer above a lower, confined aquifer) would require selecting two random sites for monitoring within each cell to characterize each aquifer. One random site should be a deep well screened only in the lower aquifer, and the other site should be a shallow well whose depth does not exceed the upper aquifer's thickness. Wells should be avoided that have the possibility of being screened in both aquifers, as this would provide an unknown mix of groundwater. Even if only a single aquifer is spatially located within a cell, two wells pumping water from different depths (deep and shallow) could be monitored to examine for variability in groundwater levels and constituent concentrations with groundwater depth. In the perfect scenario, the shallow well could be pumping water from just below the water table to provide data on the water quality of the recharge percolating through the vadose zone.

Transect Monitoring

Transect monitoring is an especially useful design in basins that consist of steep river valleys. Most water development in these basins occurs along the flat river floodplain; minimal groundwater monitoring opportunities lie in the surrounding uplands. Transect monitoring requires selecting sites from each aquifer along, as close as possible, a straight line perpendicular to the river valley. For example, sites could be monitored at high elevation mountain hardrock, in the upland basin-fill aquifer, and finally along the floodplain aquifer. Each of these transects would allow the opportunity to examine groundwater evolution flow paths from the basin's highest recharge points down to the valley floor. A final groundwater evolution flow path is then examined along the floodplain aquifer from the most upgradient site to the most downgradient site in the basin.

Transitional Subunit Monitoring

Any of the ambient groundwater network designs described can be adjusted to examine for targeted objectives within the basin. A higher density of monitored sites could be established for preselected areas that have, as previously shown in the hydrologic literature, the potential to impact either groundwater levels or quality. Preselected sites include, but are not limited to, clusters of large capacity production wells, confined animal feedlots, mines, landfills, irrigation districts, and residential areas that have a high density of septic systems for wastewater disposal. Other activities that also have a good probability of impacting groundwater may be found in the hydrologic literature. For these targeted subunit studies, care should taken to monitor groundwater at shallow sites both upgradient and downgradient of the preselected site.

Transitional subunit monitoring may also be used to define the extent of sampling sites better in the ambient network that have either dramatic groundwater level changes and/or constituent concentration statistical outliers. This additional monitoring will assist in determining if the randomly selected site is characteristic of the area. A groundwater level outlier might indicate that the well, is tapping a shallow, perched aquifer instead of the regional aquifer. For a constituent concentration outlier, it might indicate that the well is not representative of the area because of construction flaws such as a cracked casing or an improper surface seal that could allow impacted surface flow down the well casing.

MONITORING SITE SELECTION

An important consideration when establishing an ambient network is the type of well to use for monitoring. The preferable option is to construct new wells throughout the system. This allows the major benefit of control over their location, depth, screened opening, pump type, and other characteristics. The downfalls of this option include locating suitable well sites and the major expense of drilling, developing, and the upkeep of the wells. Even if funding allows establishing the wells, finding suitable well sites might not be easy. Preferable sites for wells are public lands, including road rights-of-way to obtain unlimited access. However, it may be difficult to find public lands in all areas where ambient network wells should be established. Selecting groundwater monitoring sites on private property with the land owner's permission allows a wider range of potential well sites. However, this option has the downfall of potentially being denied access to the well at a later date by the present or future, property owner.

Using preexisting private wells also has a number of advantages and disadvantages; the major factor is significantly less initial cost to establish an ambient monitoring network. No well drilling, development, or upkeep costs are required because wells are already in place, maintained by an outside party. The disadvantages are that there is no control over well characteristics. There is also no guarantee that the well will continue in operation for monitoring in the future. Even if the current well owner is very positive about involvement in the ambient network, a future property owner could have a different position.

Preexisting wells may generally be divided into two groups. High capacity wells typically consist of those used for municipal, mining, and irrigation purposes. Low capacity wells typically consist of those used for domestic, stock, and monitoring purposes. Other potential groundwater monitoring sites include naturally flowing springs. There are many advantages and disadvantages of using each type of well in an ambient network (Table 1).

In order to minimize potential biases in well selection, the most pragmatic approach for choosing wells for the network is to use a mix of high and low capacity wells that supply groundwater for different purposes. Care should be taken to select wells that are only screened in a single aquifer and have the most complete well characteristic data available. Using both high and low capacity wells should result in a mix of deep and shallow wells that will provide for a more complete assessment of the basin. Generally, public water supply wells are monitored under state and/or federal regulations which may make them generally more amenable to additional sampling. However, other high capacity wells such as which supply irrigation or mining uses or low capacity wells may have never been previously sampled. Selecting these wells could result in obtaining important hydrologic information not previously known about the area of interest.

If some funding is available for well drilling, very shallow monitoring wells often are a good choice to construct for several reasons. Foremost, such wells would indicate what type of impacts the aquifer might be receiving from recharge through the vadose zone. Because they are screened near the top of the water table. Also, because of their shallow depth and associated lesser cost, more monitoring wells could be drilled than other types of deeper wells.

MONITORING SITE INVENTORY

A site inventory is created to collect and document information for each well or spring selected or constructed for the ambient network. The site inventory contains important hydrologic information including, but not limited to, well location, owner, water use, well depth, screened interval, pumping depth, field measurements (water level and physical parameters such as temperature, pH, and specific conductivity), sampling permission forms, prior sampling measurements, and miscellaneous field activities and observations. A preprinted field sheet should be c-cated on which much of the above information can be added should be created, so that items are not overlooked in the field. Consistent field measurement and sampling protocols must be adhered to.

Table 1. Advantages and Disadvantages of Various Types of Wells for Ambient Groundwater Networks

Well Type	Advantages	Disadvantages
	High-Capacity Production Wells	
All	– easy to purge	– might not have flow rate controls making sampling for volatile constituents difficult
	– provide a more complete measure of groundwater quality throughout depth	– because of high pumping rates may draw water from more than the aquifer the well is screened in
Irrigation	– may pump constantly during the crop's growing season	– might have only seasonal use – turbine pump oil can cause contamination – can cause contamination when used for chemigation without antisiphon devices
Mining	– may pump constantly	
Municipal	– long-term access frequently possible	– wells producing water not meeting standards might be abandoned
	– might have extensive historical water quality data	– must sample before any chlorination
	Low-Capacity Production Wells	
All	– might lack well construction information	– susceptible to local impacts such as septic systems
Domestic	– often are more shallow wells	– may have water taps after storage or pressure tanks
Monitor	– often screened near the water table to investigate aquifer recharge	– may have been originally installed to monitor some sort of contamination
Stock	– often located in areas devoid of other types of wells	– often are windmills which can be difficult to purge adequately in calm weather
	Other Groundwater Sites	
Springs	– often located in mountainous areas devoid of any wells	– may flow only seasonally or during wet periods – may not have a clearly demarcated point of emergence.

MONITORING MEASUREMENTS AND SAMPLES

To determine what measurements and samples to collect for the monitoring network, the cost of each must be weighed against how valuable the information potentially gained from it is. One viewpoint is that much of the cost of an ambient network consists of salaries, travel expenses, and sampling equipment. As such, as wide a variety of groundwater data as possible should be collected at the monitoring site, which is an especially persuasive argument during the initial monitoring, when establishing baseline conditions. However, it might not be fiscally or scientifically prudent to collect certain types of samples where there is an extremely remote possibility of detecting the contaminants. As an example, collecting samples to be analyzed for organic contamination such as pesticides or volatile organic compounds in relatively pristine areas might not be the best use of scarce resources, as these tests are relatively expensive.

Baseline data collected at ambient network sites may include, but are not limited to, the following types of samples: inorganic constituents (physical parameters, major ions, nutrients, and trace elements), isotopes (oxygen, hydrogen, tritium, and others), radiochemistry (especially in areas of granite geology or mining activity), radon gas, volatile organic compounds, and pesticides. In addition, on-site measurements of groundwater depth, temperature, pH, specific conductivity, dissolved oxygen, alkaling and redox can be determined using portable instruments. Bacteria can be sampled but it may be of limited value for long-term trend analyses because microbiological contamination in groundwater is often transient and subject to a variety of changing environmental conditions, including soil moisture and temperature.

During selection of sample types to collect for the ambient network, adequate quality assurance/quality control procedures and samples should be integrated into work plans and activities. Quality assurance data validation tests may include cation–anion balances, specific conductivity–total dissolved solids ratios, hardness (lab-calculated), pH (field-lab), specific conductivity (field-lab), and groundwater temperature–groundwater depth. Quality control samples may include, but should not necessarily be limited to, travel blanks, equipment blanks, replicate samples (both duplicates and splits), and spiked samples.

AMBIENT NETWORK SIZE

The various conceptual designs are efficient methods for characterizing groundwater because they require monitoring relatively few sites to make valid statistical statements with a known degree of confidence about the conditions of a large basin. Each design will characterize the overall basin, though the density of monitoring will depend on scientific and fiscal factors. Obviously, funding has a critical influence on the density of monitoring. When possible, a network size of 30 sites is desired as this is often large enough for a normally distributed population to be recognized as such (4).

Hydrologic factors also provide some measurement of optimal monitoring density. The hydrologic complexity of the basin and the variability of historical groundwater information are important components to consider when planning an ambient network. The hydrologic complexity of a basin is influenced by factors such as the number of aquifers, surface water features such as rivers and lakes, and the presence of high water use activities such as irrigation districts, mining and milling operations, or large urban areas. The variability of historical groundwater data can be judged qualitatively by examining the variation in groundwater levels, the types of dominant water chemistries present, and the number of trace elements detected. More extensive quantitative analyses should also be conducted which can include determining constituent concentration means and standard deviations to gauge the variability of groundwater quality in the basin.

MONITORING FREQUENCY

Once the ambient network has been established and the baseline data have been collected, it should be determined how frequently it will be monitored and what types of data will be collected. Measurement of groundwater levels should be tied to both the overall extent, and to recent changes in, the water development of the basin. Groundwater quality typically varies spatially more than temporally because of the insulating effect of the soil and vadose zone. Therefore, except in very shallow wells, it is typically unnecessary to resample wells frequently because rarely do constituent concentrations change significantly over periods of a few years. A quantitative method of determining the frequency is to examine previous groundwater quality studies in the basin covered by the ambient network. If wells or springs common to both are established, the constituent concentrations can be examined for significant statistical differences that are unlikely to occur by chance. The time period between the previous study and the baseline monitoring of the ambient network can establish the frequency needed for sampling. If new development has occurred in the basin that could potentially impact groundwater, this time period should be shortened to account for these recent influences.

In select areas of very shallow groundwater, monitoring should be seasonal. Particularly susceptible to seasonal fluctuations are perched aquifers of little storage capacity. Impacts such as high runoff from wet winters or irrigation recharge during an agricultural growing season may be reflected in high seasonal variability in both groundwater depth levels and constituent concentrations.

EFFECTIVELY USING DATA FROM AN AMBIENT NETWORK

Hydrologic data obtained from an ambient network should be used for a variety of purposes to make the investment in a program of this type as valuable and cost-effective as possible. If the ambient network is created and data collected without any subsequent assessment, the program should not be considered successful in accomplishing its goals. A common problem is adequately planning, developing, and monitoring an ambient network without the subsequent dedication to analyzing and interpreting the data and producing readable, scientifically valid reports that can be used by hydrologists, planners, and others interested in groundwater information.

Data collected from an ambient network should be used for a multitude of purposes. Most simply, efficient and accessible data storage and retrieval is a necessity for others seeking information from the ambient network, which should be as convenient as possible; potentially the data should be available electronically over the Internet. The potential audience for the ambient network data is wide and includes hydrologists, planners, real estate professionals, and anyone using groundwater.

Data from the ambient network should be examined for both spatial and temporal patterns. Empirical observations and qualitative methods are often successfully used to examine groundwater data. However, so that the hydrologist remains unbiased, statistical methods should also be used to examine the data. Groundwater data can be divided among various components of the basin to examine for statistically significant correlations. Aquifers, geologic types, watersheds, land uses, recharge zones, and fault zones are just some of the potential components used in examining for groundwater quality patterns. The ambient network becomes a powerful predictive tool when significant relationships are revealed between groundwater quality concentrations and physical or cultural features. At this point, the groundwater quality of unmonitored areas is predictable with a known degree of confidence based on statistically significant patterns revealed by the ambient network.

Finally, the data, spatial and temporal patterns and the interpretations of these patterns should be woven into a readable, scientifically valid report. Two versions could be produced: an in-depth report designed for hydrologists who require comprehensive groundwater information on a basin and a short fact sheet designed for a more casually interested audience. Both published reports should be widely disseminated among both the hydrologic community and the area of interest as well as be available on the Internet.

BIBLIOGRAPHY

1. Neely, K.W. (1994). *Idaho Statewide Ground Water Quality Monitoring Program Network Design*. Idaho Department of Water Resources, Boise, ID, p. 7.
2. Myers, G., Magdalene, S., Jakes, D., and Porcher, E. (1992). *The Redesign of the Ambient Ground Water Monitoring Program*. Minnesota Pollution Control Agency, St. Paul, MN.
3. Pitz, C.F. (2003). *Strategic Recommendations for Groundwater Assessment Efforts of the Environmental Assessment Program*. Washington State Department of Ecology, Olympia, WA.
4. Alley, W.M. (1993). *Regional Ground-Water Quality*. Van Nostrand Reinhold, New York.
5. Scott, J.C. (1990). *Computerized Stratified Random Site-Selection Approaches for Design of a Groundwater Quality Sampling Network*. U.S. Geological Survey Water-Resources Investigations Report 90–4101, Oklahoma City, OK.

MTBE

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Methyl tertiary-butyl ether (MTBE) is a synthetic compound that was developed as a solution to the problem of air pollution from vehicle emissions. MTBE was originally added to gasoline as a fuel oxygenate in approximately 1979 to reduce air emissions and make the fuel burn cleaner, allowing more complete combustion in internal combustion engines. The concept was to improve air quality by producing fuels that were more readily oxidized and decrease the resulting tropospheric ozone production. In the late 1980s, the U.S. EPA supported the use of fuel oxygenates in reformulated gasoline to achieve air quality goals.

MTBE, a byproduct of gasoline refining, was inexpensive, and readily available nationwide. Ironically, the use of this air-saving gasoline additive has taken its toll on groundwater and surface water resources throughout the nation. Life-cycle studies were not performed on MTBE and ultimately, leaks and spills of gasoline containing MTBE from underground tanks and pipelines have leaked into the groundwater. In addition, incomplete combustion of gasoline containing MTBE by two- and four-stroke engines in water skis and boats has allowed the release of MTBE into surface waters.

MTBE is an ether whose general chemical formula is $C_5H_{12}O$ and structural formula is $CH_3OC(CH_3)_3$ or $(CH3)_3COCH_3$ (1). (See Fig. 1). MTBE is highly soluble in water, more than 75 times more so than many other gasoline compounds. 250,000 MTBE-contaminated sites throughout the nation have been identified. A January 16, 2000, more than airing of the "60 Minutes" television program on MTBE highlighted the problem for a national audience, although the problem of MTBE and water resources had been known for a few years to water professionals and regulators.

PERSISTENCE IN GROUNDWATER

Sampling of groundwater in California and other states has shown that MTBE is more persistent and more mobile than many other fuel compounds.

TASTE AND ODOR

A taste and odor like turpentine has made MTBE-contaminated groundwater sources at levels exceeding the taste and odor threshold unusable for human consumption. Supplying bottled water for impacted areas and/or other large-scale treatment or water importation options will be expensive and will continue into the future. Blending clean water with MTBE-impacted water will also continue.

HEALTH RISK

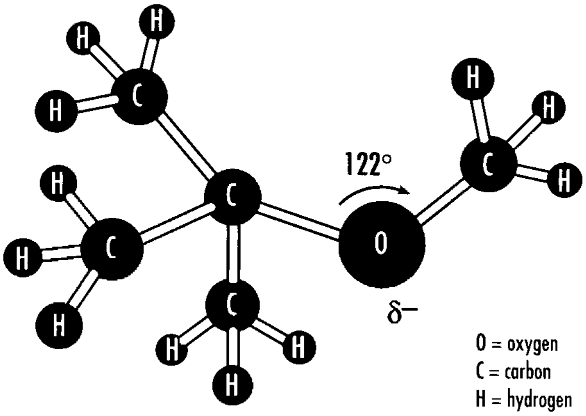

Figure 1.

Although MTBE is considered a potential health risk, there is inadequate evidence that MTBE causes cancer in humans. The strong odor and taste of MTBE at relatively low concentrations would make drinking of large amounts of high-level MTBE contaminated water unlikely, thus reducing the potential for exposure and resulting human health risks from MTBE. MTBE is not listed as a human carcinogen by the U.S. National Toxicology Panel, the California Proposition 65 Committee, or the International Agency for Research on Cancer. Limited evidence from an older Italian MTBE rodent study indicated that MTBE caused cancer in the experimental animals when administered in unrealistic doses hundreds of times greater than probable human exposure. California, a state that has the most MTBE cases, has a primary MCL of 13 μg/L and a secondary MCL (taste and odor threshold) of 5 μg/L (2).

FATE AND TRANSPORT

The MTBE issue is a chemical handling problem associated with gasoline leaking from underground storage tanks, leaking pipelines, and surface spills or from small gasoline engine equipment. Even though the majority of underground tanks in this country have been upgraded, the amount of MTBE already released with gasoline into the environment is so large that the problem of degraded drinking water sources may persist for decades. Compared to other fuel compounds, MTBE adsorbs to soils much less and dissolves in groundwater at much higher concentrations.

Due to these characteristics, MTBE migrates much faster and further in groundwater than equal amounts of other gasoline compounds. According to Wilson (2), MTBE often appears to move out ahead of a BTEX plume. The explanation is based on differences in the extent of adsorption to the aquifer materials. As the fraction of organic carbon in the aquifer sediment increases, benzene and other gasoline compounds adhere more to aquifer sediments than MTBE. Therefore, an aquifer with a high fraction of organic carbon (0.1%) would have a retardation index for MTBE and benzene, at 1.6 and 2.9, respectively (3). These differences help to explain the variations in contaminant transport rates.

GROUNDWATER INVESTIGATION

The investigation of a groundwater resource impacted with MTBE requires more detailed and more rapid sampling of soil and groundwater than other fuel-impacted sites due to the potential for faster and deeper migration of MTBE compared to other gasoline compounds. The presence of other oxygenates should also be tested at all sites containing MTBE. Some brands of gasoline contain tertiary-butyl alcohol (TBA) as an additive. As TBA is a biological breakdown product of MTBE, TBA should also be analyzed at sites containing MTBE.

REMEDIATION

MTBE is an ether compound and was considered very different from gasoline. Some investigators originally thought MTBE was a recalcitrant compound. MTBE remediation has improved after years of research in the laboratory and numerous field case studies. The same technologies that are used to treat fuel hydrocarbons have, in general, been used successfully to treat MTBE. Due to the chemical characteristics of MTBE, the cost of remediation or clean up of the contaminant is equal to or more expensive than that for other gasoline related chemicals, such as benzene, toluene, ethylbenzene, and xylenes. New technologies have improved the remediation efficiency and cost. MTBE remediation has been documented using a variety of *in situ* technologies, including, but not limited to, chemical oxidation (Fenton's chemistry and ozone) and bioremediation. Enhanced *in situ* bioremediation using various delivery systems for oxygen and peroxygens effectively reduce MTBE in groundwater. *Ex situ* technologies include pump and treat systems and soil vapor extraction systems, as well as dual-phase extraction with appropriate surface treatments and destruction. Virgin coconut shell carbon has been documented for MTBE remediation. Due to the high mobility of MTBE in the environment, rapid implementation of source control is the key to remediation of sites impacted with MTBE. More detailed MTBE information and case studies are contained in Moyer and Kostecki (4).

AIR POLLUTION

The effectiveness of MTBE in reducing air pollution in recent years is suggested, but uncertain. Reformulated gasoline containing MTBE may be related to some of the improvement in air quality, however, various conflicting air studies exist. Part of the improvement may be attributable to newer and more efficient vehicle engines with computer-controlled ignition. Replacements for MTBE must be evaluated carefully as to their fate and transport once released into the subsurface. Life-cycle studies focusing on the potential release into the air, soil, and groundwater should be performed prior to introducing any new chemicals into a product as widely used as gasoline.

BIBLIOGRAPHY

1. Motzer, W. (2001). Physical and chemical properties of MTBE. In: *MTBE: Effects on Soil and Groundwater Resources*. J. Jacobs, J. Guertin, and C. Herron (Eds.). CRC Press/Lewis, Boca Raton, FL.
2. US EPA. (March 2004). Available: http://www.epa.gov/swerust1/mtbe/mtbemnp.pdf.
3. Wilson, J. (2003). Fate and transport of MTBE and other gasoline compounds. In: *MTBE Remediation Handbook*. E. Moyer and P. Kostecki (Eds.). Amherst Scientific, Amherst, MA, pp. 19–61.
4. Moyer, E. and Kostecki, P. (2003). *MTBE Remediation Handbook*. Amherst Scientific, Amherst, MA, p. 670.

READING LIST

California Dept. of Health Services. Available: http://www.dhs.cahwnet.gov/ps/ddwem/chemicals/MtBE/MtBEindex.htm.

California State Resources Control Board's Draft Guidelines on MTBE contains technical documentation on MTBE fate and transport in the subsurface. Available: http://www.swrcb.ca.gov/cwphome/ust/MtBE_finaldraft.doc.

Diaz, A. and Dragus, D.L. (Eds.). (2002). *Oxygenates in Gasoline, Environmental Aspects*. American chemical society ACS symposium series #799, Washington, DC, p. 310.

Environmental Behavior and Fate of MTBE. Available: htttp://water.wr.usgs.gov/MtBE/fs20396/.

Jacobs, J., Guertin, J., and Herron, C. (Eds.). (2000). *MTBE: Effects on Soil and Groundwater Resources*. CRC Press/Lewis, Boca Raton, FL, p. 264.

USGS has a reference site for MTBE. Available: http://water.wr.usgs.gov/MtBE/.

LIMITING GEOCHEMICAL FACTORS IN REMEDIATION USING MONITORED NATURAL ATTENUATION AND ENHANCED BIOREMEDIATION

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California
David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

In Situ Monitored Natural Attenuation (MNA) and enhanced bioremediation are affected by a variety of limiting factors, including the availability of a terminal electron acceptor (TEA), macronutrients (nitrogen and phosphate), as well as an energy source (carbon). The TEA for aerobic degradation is limited to oxygen. For anaerobic degradation, TEA processes include denitrification using nitrate (NO_3^-), manganese reduction using manganese dioxide (MnO_2), iron reduction using bioavailable iron that includes the hydroxide form ($Fe(OH)_3$), sulfate reduction using the sulfate ion (SO_4^{2-}), and methanogenesis using carbon dioxide (CO_2) (1).

Macronutrients are measured in the laboratory as nitrogen in the form of ammonia or nitrate and orthophosphate. Nutrients, although important, are not as critical as depletion of the TEA. The energy source is the carbon in the hydrocarbons or chlorinated solvents that act as food for the microbes.

Monitored Natural Attenuation occurs when these processes are believed to be occurring on their own without additives. Natural attenuation frequently starts on many hydrocarbon sites, but the aerobic degradation process frequently stops or slows down considerably because of the limiting factors of oxygen and nutrient depletion. Enhanced bioremediation and Monitored Natural Attenuation have been described in detail in numerous volumes, including (1) and (2).

Aerated water at one atmosphere and at room temperature usually has about 8.0 mg/L dissolved oxygen. Although groundwater may have about 8 mg/L near the recharge area, oxygen depletion starts to occur as the oxygen is consumed by the chemical and biological processes that use up the oxygen. To measure these demands, laboratories measure solid or sediment oxygen demand. For dissolved oxygen demand, chemical oxygen demand and a subset, the biological oxygen demand, are typically analyzed. The biological oxygen demand is commonly evaluated as a 5-day analysis originally designed for sewer treatment plants.

To support aerobic biodegradation of petroleum hydrocarbons, aerobic processes are less than optimal when the dissolved oxygen is below 2 to 4 mg/L (3); at this oxygen level, facultative anaerobes may become active. At this stage of oxygen depletion, if a biodegradation process is to be used for the complete remediation of the site, enhanced bioremediation may be used. Enhanced bioremediation is designed wherein oxygen and sometimes nutrients are delivered into the subsurface through wells, well points, borings, or trenches. The delivery of oxygen depends on the oxygen form. Oxygen comes in many forms and can be added using a variety of techniques. Some of the more common forms of oxygen used for enhanced bioremediation include gas phase oxygen (O_2) or ozone (O_3), liquid oxygen sources such as hydrogen peroxide (H_2O_2), and solid oxygen sources including magnesium peroxide (MgO_2) and calcium peroxide (CaO_2). Some of these compounds are chemical oxidizers and, above certain threshold concentrations, can kill microbes in the pore space of the injection area during the process of releasing the contained oxygen. After the reactions occur, microbial populations can increase greatly from populations in more isolated areas of the mineral matrix.

In theory, oxygen levels of at least 3 to 3.5 times that amount of subsurface petroleum mass are required for complete mineralization of the hydrocarbon. Complete mineralization is the converting of the hydrocarbon to carbon dioxide and water:

$$\text{Benzene: } C_6H_6 + 7.5\ O_2 \rightarrow 6\ CO_2 + 3H_2O \ (3.1 \text{ oxygen requirement})$$

For anaerobic TEAs used to treat chlorinated solvents by reductive dehalogenation, one common method of TEA replenishment is to add a carbon source such as liquid molasses, lactic acid, cheese whey, or gases such as propane that also contain a large number of hydrogen atoms. In the process of the biological breakdown of the introduced chemical, hydrogen is liberated in the process. The hydrogen is used as the electron donor and the liberated carbon is used as a food source, along with the carbon-based contaminant. For petroleum hydrocarbons and some xenobiotics, such as MTBE ferric iron or sulfate, salts can be added to stimulate anaerobic biooxidation.

NUTRIENTS

Nutrient levels can be at well below reporting levels. The carbon:nitrogen:phosphorus ratios necessary to enhance biodegradation occur in the optimal range of 100:10:1 to 100:10:0.5, depending on the chemical contaminant, microbes, and other site-specific conditions (3). In some instances, anaerobic biooxidation processes may be inhibited because of the lack of trace nutrients. In these cases, native sulfate will be present in the groundwater and possibly iron. Continued use of these native inorganic TEAs can be stimulated by the addition of critical nitrogen and phosphate as trace nutrients.

SOIL MOISTURE

Soil moisture content in the unsaturated zone may be related to precipitation in the region and the retention capacity of the shallow soil. Soil moisture content of 75 to 90 percent of field capacity is considered optimal for aerobic microbial activity (4).

TEMPERATURE

Temperature has been seen as a limiting factor for biological degradation of hydrocarbons. Microbial growth rate is a function of temperature. Subsurface microbial activity has been shown to decrease significantly at temperatures below 10 °C and essentially to cease activity below 5 °C. From 10 °C to 45 °C, the rate of microbial activity typically doubles for every 10 °C increase in temperature (3). Typical soil microbial activity will slow as the temperatures reach 55 °C and will reach a steady state determined by the capacity to flush heat from the system. In cases where the load of petroleum hydrocarbons is high, the thermodynamics of the heat transport may be the rate-limiting factor rather than any biogeochemical process.

One case adjacent to the Yuba River near Donner Pass in California (7,000 feet above mean sea level) and another at South Lake Tahoe area in Nevada were evaluated for total heterotrophic plate counts and specific hydrocarbon microbial degraders. In both cases, plate counts were well below 100 colony forming units per gram of dry soil. Temperature of the water was seen as the limiting factor.

Assessment procedures must be completed that will, support the reliance on the MNA or enhanced bioremediation process by proving natural attenuation is taking place, defining the site-specific natural attenuation mechanisms, quantifying the kinetics of the process and providing an estimated time for cleanup, and insuring that the site-specific conditions will not allow unacceptable offsite migration of the contaminant plume. Given the above information, a governing regulatory body can support a decision for an MNA solution. The purpose of this article is to examine conditions under which the biological component of natural attenuation is inhibited or terminated, because of toxicity from heavy metals or certain hydrocarbons from the impacting contaminants, other compounds associated with the contaminants (at nonregulated levels), or compounds associated with the native groundwater.

The effect of toxicity on the MNA process is seen in three areas. First is a simple inhibition of activity. The microbial population continues to function, but at a rate that is slower than what would occur under normal conditions.

Second is a complete cessation of biological degradation activity. The bacteria are killed outright or, more commonly, they enter into a vegetative static state in which they are metabolically inactive. However, if the offending toxic material is removed from the system, bioactivity can resume.

Third is interference with degradational activity that is specifically focused on a contaminant of concern, which is more subtle, but particularly important to the MNA process. The overall microbial activity is not inhibited, just degradation of the contaminant. This problem is more often seen in systems where xenobiotic hydrocarbons are the targets of cometabolically generated degradation enzymes. In these cases, it is common for attenuation of the xenobiotic to take place after all of the other “natural” hydrocarbons have been consumed.

To evaluate the impact of toxicity, it is important to understand the mechanisms by which substances are toxic to microorganisms or interfere with the degradation process. The following bacteria cell components or processes are sensitive to toxic effects from the listed compounds:

- The bacterial cell wall can be totally destroyed by heavy metals, phenols, and alcohols.
- The permeability of the cell wall can be fatally increased by phenolic compounds, alcohols, detergents, and quaternary ammonium compounds.
- The alteration of interior proteins (destroying their cellular function) is caused by halogens, phenols, and alcohols.
- Interference with the action of intercellular enzymes will stop metabolic activity. Compounds toxic via this effect include cyanide, strong oxidants (chlorine, etc.), phenols, metals, and metalloids.
- Interference with nucleic acid synthesis can prevent bacterial reproduction; heavy metals can act in this fashion. Environmental conditions also play a role in the impact that toxic materials have, particularly for metals. Toxicity of metals is governed by concentration, chemical form of the metal, pH and Eh of the environment, the type of microbial system involved (i.e., aerobic or anaerobic), and potential for adaptation of the bacteria. Toxic forms of metals include soluble salts and anionic complexes. Insoluble salts, oxides, or even elemental forms are not necessarily toxic. The best method to evaluate groundwater for the presence of potentially toxic metals is to collect and test filtered water samples.

Effect of pH

Bacteria are sensitive to pH conditions. Most bacteria that are indigenous to groundwater systems and functional in the MNA process thrive in the pH range of 6.3 to 7.5. An anthropogenic pH lower than 5.0 is inhibitory or toxic from the pH effects alone. In cases where native groundwater has lower pH, the indigenous bacteria will have likely adapted to that specific condition, which is the case in lignite-rich areas of east Texas where it is common for the native groundwater pH to be as low as 4.5. Lower pH will cause an increase in soluble metal concentrations, including metals that may be native to the aquifer matrix.

REDOX Conditions

Oxidation-reduction (Redox) potential largely impacts the consortia of bacteria that will be active and the electron acceptor available to be used by the system. Almost any range of Eh conditions has the potential to offer effective biodegradation. However, some anaerobic systems are more sensitive to toxic compounds.

Toxicity Issues

Some naturally occurring metals and compounds, as well as man-made pollutants, have contributed to minimizing or stopping *in situ* biodegradation processes. As a general rule, heavy metals in excess of 2,500 mg/L are too toxic for microbial populations (5).

These toxic elements and compounds affect bacteria consortia and bacterial adaptability at the site. There are substances or compounds commonly found at contaminanted sites that have proven to be toxic to microbial populations. In instances where a concentration range is given, the low value represents inhibition of microbial activity, and the high value represents a concentration that will kill or stop microbial activity.

Copper may be the heavy metal with the highest toxicity with regard to the impact on the *in situ* biological degradational process. It is extremely toxic, inhibitory at low concentrations, and can interfere with the cometabolic degradation of chlorinated solvents at the lowest end of this range.

CONCENTRATION RANGES FOR SELECTED TOXIC SUBSTANCES

Copper 0.01 to 20 mg/L
Zinc 0.3 to 10 mg/L
Cadmium 0.1 to 20 mg/L
Chromium 25 mg/L - Inhibition
Nickel 25 mg/L - Inhibition
Lead 900 mg/L - Inhibition
Cobalt 0.3 to 10 mg/L
Mercury 0.01 to 20 mg/L
Methanol 90 mg/L - Inhibition
Isopropanol 55 mg/L - Inhibition
Acetone 75 mg/L - Inhibition
Pentachlorophenol 1 to 200 mg/L
TNT 10 to 100 mg/L
Cyanide 10 to 150 mg/L

Petroleum Hydrocarbons

The components of common petroleum hydrocarbons become inhibitory when soluble concentration ranges reach 200 to 500 mg/L. The upper range for lethal toxicity verges on 50,000 mg/L or free product, in which electron acceptor transport is more an issue than chemical toxicity. Evidence exists that methanogenesis may be inhibited by some constituents in crude oil.

Chlorinated Hydrocarbons

Most chlorinated solvents (also called chlorinated hydrocarbons) become inhibitory to microbes when soluble concentrations of these compounds are in the 100 to 500 mg/L range. They become toxic at concentrations around 1,000 mg/L, because of chemical attack on cell wall lipids. Chlorinated solvent concentrations in excess of 7,000 mg/L are inhibitory or toxic to microbial populations (5).

A baseline of total heterotrophic colony forming units and specific hydrocarbon degraders is useful prior to recommending natural or enhanced bioremediation at appropriate sites. If heterotrophic microbial population count is low (below 1,000 CFU per gram of dry soil), an assessment screen for compounds toxic to bacteria may be useful.

If toxic compounds are found, it may be necessary to perform some level of bench scale testing to quantify the toxic effect of suspect compounds. This assessment process is still a much less expensive procedure than the typical "pump and treat" active groundwater remediation projects. In some situations, pH amendments or other geochemical treatments can modify toxic conditions in the subsurface.

Monitored Natural Attenuation, enhanced aerobic bioremediation, and stimulated anaerobic biooxidation are likely to continue into the future. The first two are commonly used remedial approaches to obtain site closure for hydrocarbon-impacted properties where remediation time can be lengthy, the latter is a rapidly developing new tool for those processes. The costs of bioremediation can be significantly lower than pump and treat or other active and more aggressive remedial approaches. Assessing the limiting factors and performing geochemical studies help set baseline levels and allow for the monitoring of contaminant degradation over time.

Dissolved oxygen and nutrients are commonly depleted early on, so the use of enhancements to optimize subsurface conditions are frequently used. The power of nutrient stimulation to exploit native sulfate and iron for anaerobic biooxidation or the supplement of sulfate or iron salts for further stimulation of anaerobic biooxidation lies in the fact that the amendments are water soluble at relatively high concentrations and, in most instances, are also conservatively transported. They do not react until present with the utilizing bacteria and the impacting hydrocarbon. Sites using MNA as the sole remedial method to obtain site closure to low residual contaminant levels are not that common. Regardless of the biological method used, *in situ* bioremdiation methods are effective, but generally work on a slow time scale, with the results being observed over many months or years.

BIBLIOGRAPHY

1. Suthersan, S.S. (2002). *Natural and Enhanced Remediation Systems*. CRC Press/Lewis Publishers, Boca Raton, FL, p. 419.
2. Wiedemeier, T., Rafai, H., Newell, C., and Wilson, J., (1999). *Natural Attenuation of Fuels and Chlorinated Solvents in the Subsurface*. John Wiley & Sons, New York.
3. United States Environmental Protection Agency (USEPA). (2004). Enhanced Aerobic Bioremediation. In: *How to Evaluate Alternate Cleanup Technologies for Underground Storage Tank Sites; A Guide for Corrective Action Plan for Regulators*. EPA 510-R-04-002, Washington, DC.
4. United States Environmental Protection Agency (USEPA). (2004). Monitored Natuaral Attenuation. In: *How to Evaluate Alternate Cleanup Technologies for Underground Storage Tank Sites; A Guide for Corrective Action Plan for Regulators*. EPA 510-R-04-002, Washington, DC.
5. United States Environmental Protection Agency (USEPA). (2004). In-Situ Groundwater Bioremediation. In: *How to Evaluate Alternate Cleanup Technologies for Underground Storage Tank Sites; A Guide for Corrective Action Plan for Regulators*. EPA 510-R-04-002, Washington, DC.

NITRATE CONTAMINATION OF GROUNDWATER

Ann Azadpour-Keeley
National Risk Management
Research Laboratory
ORD, U.S. EPA
Ada, Oklahoma

Jack Keeley
Environmental Engineer
Ada, Oklahoma

Groundwater supplies drinking water for half of the people in the country and represents the total water supply for many of the western states. Numerous contaminants are anthropogenic, including petroleum products, chlorinated solvents, and pesticides, but nitrate contamination can be either natural or man-made and is one of the most common chemical contaminants of groundwater. Nitrate levels in uncontaminated groundwater are generally less than 2 mg/L, but levels have been increasing during the past five to six decades in areas where contamination has occurred.

Nitrate nitrogen is predominantly the product of the microbial decomposition of nitrogenous material, and almost all nitrogen reactions in soil and groundwater are the result of microorganisms. Through decomposition, nitrogenous matter is converted to ammonia (NH_3) which is first oxidized to nitrite (NO_2^-) followed by further oxidation to the production of nitrate (NO_3^-); both are accompanied by the production of nitrogen gas (N_2). Nitrate serves as a nutrient for the production of plant life that is consumed by living animal matter, which produces wastes products and, upon death, completes the nitrogen cycle by creating nitrogenous matter.

Sodium and calcium nitrate add nitrogen directly to the soil resulting in valuable fertilizers. A large natural deposit of sodium nitrate in Chile served as a major source of nitrogen fertilizer. Geologically deposited nitrate can also be found in the San Joaquin Valley in California, Runnels County, Texas, and shale in Montana and South Dakota.

Although the degradation of plant materials can account for small amounts of nitrate in groundwater, most results from the application of fertilizers to agricultural land, parks, lawns, and golf courses. Other sources include animal wastes, seepage from septic tanks, the application of sewage and sludge to the land, precipitation, geological deposits, and biological fixation. Canter and Knox (1), for example, estimate that one trillion gallons of septic-tank waste are released to the subsurface annually. According to Nolan et al. (2), 11.5 million tons of nitrogen as fertilizer in used in agriculture annually in the United States where the production of commercial fertilizer has increased by a factor of 20 between 1945 and 1985. Manure from farm animals accounts for an estimated 6.5 million tons of nitrate annually, and that from the atmosphere is about 3.2 million tons.

Nitrates are the cause of methemoglobinemia which is also referred to as blue baby disease and nitrate cyanosis in infants. Nitrates are reduced to nitrites in the intestinal tract and absorbed into the blood. The nitrites combine with hemoglobin which deprives organs and tissues of oxygen, resulting in cyanosis which causes a blue color of the skin. Nitrates in drinking water have also been linked to non-Hodgkin's lymphoma. Other health related problems may include impairments to the nervous system and birth defects (3). The U.S. Public Health Service established drinking water standards for nitrate in 1946 at a limit of 10 mg/L expressed as nitrogen and 45 mg/L expressed as nitrate.

Generally, nitrate contamination is pervasive at shallow depths in the vadose zone. For this reason, drinking water from rural water districts or other public water supplies is less likely to contain high nitrate concentrations because of the depth of the wells and steps taken to prevent contamination from sources near the surface such as septic tanks. When water is pumped from greater depths, it is also more likely to have a low oxygen concentration that results in possible denitrification. Rural wells, on the other hand, are usually completed at shallower or depths in areas where sources of contamination, such as septic tanks, fertilizers, or manure from feedlots or holding pens, are more likely. When obtaining water from a domestic well, care should be given to ensure that the screen is placed as deep as is practicable for that location, that the upper casing is properly sealed, and that the wellhead is elevated and protected by a concrete pad and is in a secure enclosure. In a study by Schubert et al. (4), it was found the most likely scenario for exposure to high nitrates in drinking water involved families living on farms supplied by old, shallow wells. In such cases, the water should be tested periodically for nitrates as well as bacteria.

BIBLIOGRAPHY

1. Canter, L. and Knox, R.C. (1984). *Evaluation of Septic Tank System Effects on Ground Water Quality*. EPA-600/2-84-107.
2. Nolan, B.T., Ruddy, B.C., Hitt, K.J., and Helsel, D.R. (1997). Risk of nitrate in groundwaters for the United States—A national perspective. *Environ. Sci. Technol.* **31**: 2229–2236.
3. Keeney, D. (1986). Sources of nitrate to ground water. *CRC Crit. Rev. Environ. Control* **16**: 257–304.
4. Schubert, C., Knobeloch, L., Kanarek, M.S., and Anderson, H.A. (1999). Public response to elevated nitrate in drinking water wells in Wisconsin. *Arch. Environ. Health* **54**(4): 242–247.

TREATMENT FOR NITRATES IN GROUNDWATER

Walter W. Loo
Environmental & Technology Services
Oakland, California

Jason J. Wen
City of Downey
Downey, California

INTRODUCTION

Nitrate is a naturally occurring compound that forms in the soil when nitrogen and oxygen combine. Small amounts of nitrate are normal, but excess amounts can pollute supplies of groundwater.

Drinking water high in nitrate is potentially harmful to human and animal health. Nitrate (NO_3) is a naturally occurring form of nitrogen (N) that is very mobile in water. It is essential for plant growth and is often added to soil to improve productivity. Water moving down through soil after rainfall or irrigation carries dissolved nitrate with it to groundwater. In this way, nitrate enters the water supplies of many homeowners who use wells or springs (1).

Sources of Nitrate Pollution in Water

Common sources of nitrogen in the soil are fertilizers, livestock waste, and septic systems. Excess nitrate in the soil is most often found in rural and agricultural areas. Nitrate travels easily through the soil and is carried by rain or irrigation water into groundwater supplies. Wells that tap groundwater may be affected. Shallow wells, wells in sandy soil, or wells that are improperly constructed or maintained are more likely to have nitrate contamination.

The distribution of nitrate contamination in the United State is widespread. It constitutes the largest amount of pollutant released into the environment. The following references describe the extent of nitrate pollution (1–3).

Agricultural runoff and seepage into the groundwater occurs wherever fertilizer is applied. The denitrification of nitrate effluent into the Mississippi River is what caused the 'dead zone' of the Gulf of Mexico area along the coast of Texas and Louisiana downstream from the Mississippi River estuary. Shallow groundwater in the farming states, Illinois, Ohio, Indiana, Iowa, and Minnesota, is heavily polluted with nitrate. Large cattle feedlots, dairy farms, and orchards in California, Kansas, and Illinois release large amounts of nitrate into shallow groundwater. Nitrate pollution in groundwater is a worldwide problem.

Nitrate Properties in Groundwater

The following references provide in-depth properties of nitrate physical chemistry, natural degradation processes, and migration phenomenon (4–7).

Health Effects of Nitrate/Nitrite in Water

Infants under 6 months of age are susceptible to nitrate poisoning. Bacteria that live in the digestive tracts of newborn babies convert nitrate to nitrite (NO_2). Nitrite then reacts with hemoglobin, which carries oxygen in blood, to form methemoglobin. Methemoglobin cannot carry oxygen, thus the affected baby suffers oxygen deficiency. The resulting condition is referred to as methemoglobinemia, commonly called "*blue baby syndrome*." The most noticeable symptom of nitrate poisoning is a bluish skin coloring, called cyanosis, particularly around the eyes and mouth. A baby with bluish skin should be taken to a medical facility immediately and tested for nitrate poisoning. The blood sample of an affected baby is chocolate brown instead of the normal bright red due to lack of hemoglobin. Methemoglobinemia is relatively simple to treat, and, in most reported cases, the affected baby makes a full recovery. Within several months after birth, the increasing level of hydrochloric acid in a baby's stomach kills most of the bacteria that convert nitrate to nitrite. By the age of 6 months, the digestive system is fully developed, and the risk of nitrate-induced methemoglobinemia is greatly reduced.

Water quality standards for human consumption have been set at 10 milligrams of nitrate-nitrogen per liter of water (10 mg/L NO_3-N). This level of nitrate-nitrogen is equivalent to 45 mg/L of nitrate (NO_3). When reading laboratory reports of water quality, be sure to note whether reported values are for nitrate-nitrogen or nitrate. Note that 1 mg/L equals 1 ppm (parts per million). Most reported cases of blue baby syndrome due to contaminated water have occurred when infant formula was prepared using water that contained more than 40 mg/L NO_3-N. Consumption of high-nitrate water by pregnant women and nursing mothers is not as likely to be harmful to babies as direct consumption. The health effects in these cases are not completely understood, so it is recommended that pregnant women and nursing mothers limit nitrate consumption. Possible connections between nitrate and other health problems such as nervous system disorders, cancer, and heart damage are not well documented and are currently being researched.

Ruminant animals (cattle and sheep) and infant monogastrics (baby pigs and baby chickens) are also susceptible to nitrate poisoning because of bacteria living in their digestive tracts. Horses, even though they are monogastric, are susceptible to nitrate poisoning throughout their lives. Livestock may be exposed to large quantities of nitrate in their feed as well as in contaminated water. Animals that are treated in time can recover fully from nitrate poisoning. Scientific studies indicate that water that contains more than 25 mg/L NO_3-N can be harmful to animals.

Water Use Criteria for NO_3/NO_2 in Water

The U.S. Environmental Protection Agency (EPA) has established the maximum contaminant level (MCL) for nitrate at 10 milligrams per liter (mg/L). In addition, EPA established an MCL for nitrite at 1 mg/L.

Water quality standards for human consumption have been set at 10 milligrams of nitrate-nitrogen per liter of water (10 mg/L NO_3-N). This level of nitrate-nitrogen is equivalent to 45 mg/L of nitrate (NO_3). The presence of high nitrate levels may indicate that other, more serious contaminants may also be entering your water supply. Nitrate and coliform bacteria concentrations are commonly used to indicate well water contamination by surface waters. Water with nitrite (NO_2) levels of 1 ppm or more also is not recommended for human consumption.

These guidelines will help you interpret the results of your water test. Be careful when interpreting different lab results. Nitrate can be measured as NO_3 or as NO_3-N, and the results mean very different things (Table 1).

Nitrate is colorless and odorless. The surest way to tell if you have nitrate in your water is to have it tested. If your water comes from a public water supply, such as a city supplier, it is tested regularly to ensure that it meets safe drinking water standards. If your water comes from a private well, you are responsible for testing it and ensuring that it is safe to drink.

Table 1. Guidelines for Drinking Water

Measured Level (mg/L)	Interpretation NO_3
0–44	Below the drinking water standard. The nitrate level is safe for humans and livestock.
45–176	Higher than the drinking water standard. The water should not be consumed by infants 6 months of age or younger or by pregnant or nursing women. It may be acceptable for adults and livestock, though it is not recommended.
176+	Significantly higher than the drinking water standard. The water should not be consumed. Install a water treatment system or use bottled water for drinking and cooking.
Measured Level (mg/L)	**Interpretation NO_3-N**
0–9	Below the drinking water standard. The nitrate level is safe for humans and livestock.
10–39	Higher than the drinking water standard. The water should not be consumed by infants 6 months of age or younger or by pregnant or nursing women. It may be acceptable for adults and livestock, though it is not recommended.
40+	Significantly higher than the drinking water standard. The water should not be consumed. Install a water treatment system, or use bottled water for drinking and cooking.

Water Testing Frequency

Unfortunately, nitrate is only one of several possible contaminants in well water, which is why water districts recommend that well users conduct comprehensive tests of their water supplies every 1 to 3 years, or more frequently if the taste, odor, or appearance of the water changes, or if you have specific health concerns.

Preventive Measures for Nitrate/Nitrite Pollution

Septic systems, animal waste, and fertilizer are all potential sources of nitrate contamination. Water that comes into contact with a source of nitrate can carry that contamination through the soil and into the groundwater supply. Follow these guidelines to reduce the risk of nitrate contamination:

1. Proper well location: Wells should be located uphill (up-gradient) and at least 100 feet away from septic tanks, leach fields, animal confinement areas, and fertilized areas.
2. Proper well construction: Make sure your well casing extends above the ground, and construct an earth berm around the well to divert surface runoff away from the wellhead. Also, make sure your well has at least a 50 feet deep annular seal (grouting around the outside of the well casing) and an intact concrete slab on the wellhead.
3. Operate your septic system correctly: Maintaining your septic system regularly will help avoid system failure, which can lead to water contamination, the spread of disease, and the need for costly repairs. Follow these guidelines:

Do not run heavy vehicles over the area above your septic tank, drain pipes, or leach field.

Avoid planting trees or shrubs near drain pipes or the leach field. Roots can clog the lines.

Do not dispose of chemicals or nonbiodegradable materials in your toilet or drain.

Conserve water.

Install a lint trap on your washing machine.

Hire a reputable contractor to pump out your septic tank every 2 to 3 years.

If you have two leach fields, switch them every year.

4. Reduce your use of fertilizer: Use lawn and garden fertilizers only when necessary and always according to the manufacturer's instructions. Farmers can use management practices that are both environmentally and economically sound.

BIOLOGICAL DENITRIFICATION

Biodenitrification will probably be a very common technology used to comply with nitrate/nitrite regulation (8).

The demonstration project has developed a two-stage *ex situ* anoxic biofilter biodenitrification process. The process is fixed film bioremediation, using biocarriers and specific bacteria to treat nitrate-contaminated water. Unique to the process is a patented mixed bed reactor that retains the biocarrier within the system, thus minimizing solids carryover. Fixed film treatment allows rapid and compact treatment of nitrate with minimal byproducts. Methanol is added as a source of carbon for cell growth and for metabolic processes that remove free oxygen. The resulting oxygen-deficient environment encourages the bacteria to consume nitrate. Methanol is also important to ensure that conversion of nitrate proceeds to produce nitrogen gas rather than the more toxic nitrite intermediate.

The mechanism for anoxic biodegradation of nitrate consists of two sequential denitrification reactions. Oxygen must be consumed to a dissolved oxygen concentration of <1 mg/L. In the first denitrification step, the bacteria are forced to substitute the nitrate as the electron acceptor and the nitrate is reduced to nitrite. In the second step, the nitrite is further reduced to nitrogen gas. Nitrite production is an intermediate step, and there is no a priori reason to assume that the second reaction is at least as fast and/or favored as the first reaction in the presence of a specific bacterial population. Consequently, any evaluative scheme must establish that there is no buildup of nitrite, particularly because the nitrite-nitrogen maximum contaminant level (MCL) is only 1 mg/L, one-tenth that of nitrate. High concentrations of nitrate and high nitrate/methanol ratios may also affect the concentration of residual nitrite in a particular process configuration.

A simplified process diagram of the treatment system used during the demonstration is composed of two major components: a biodenitrification system and a polishing or posttreatment system. The biodenitrification system is intended to convert nitrates in the groundwater to nitrogen, thus reducing nitrate concentrations. The posttreatment system destroys or removes intermediate compounds generated during the biological breakdown of nitrate and removes bacteria and suspended solids that are not attached to the biocarrier. The posttreatment system can also incorporate treatment for other contaminants, such as VOCs, that may be in the influent.

Biodenitrification is conducted in two reactors, R1 and R2, respectively. The major part of the oxygen removal step is conducted in R1 where aerobic bacteria reduce the dissolved oxygen levels of the influent. Methanol is metered to the tank to encourage the bacteria to begin consuming nitrate. The resulting oxygen-deficient water is pumped from the bottom of R1 to the bottom of R2, which is densely packed with biocarrier media (1 cm) that have the appearance of small foam cubes. A patented mixing apparatus within R2 directs the incoming water into a circular motion, thus ensuring intimate contact with the biocarrier. Within R2, the majority of denitrification occurs by anaerobic bacteria that are continually fed methanol and populate on the large mass of biocarrier media. After a sufficient retention time, depending on concentration and goal, denitrified water drains by gravity to an overflow tank, which allows for continuous and smooth transfer to the posttreatment system and removal of entrained bacteria and media.

Depending on the presence of other contaminants, the posttreatment system consists of a series of varying sized filters downstream of one or more contaminant-specific treatment units. For instance, ozonation may be used to oxidize any residual nitrite to nitrate and to deactivate/destroy all residual biological materials leaving the biodenitrification unit. If VOCs are present, an air stripper and/or carbon adsorption unit can be used. If state regulations require chlorination of drinking water, then chlorine can be added as a posttreatment or directly to the overflow tank immediately following denitrification.

Waste Applicability

Anoxic biodenitrification using one or more biocarriers should be applicable to industrial wastewaters and leachate from commercial, industrial, and hazardous waste sites containing various nitrate concentrations, as well as for treatment of groundwater (the medium treated during the demonstration). The presence of other contaminants could play a significant role in the effectiveness and viability of the overall treatment system. For example, if volatile chlorinated hydrocarbons are present along with nitrate, a postnitrate treatment system (e.g., carbon filters) may be necessary to remove those compounds to acceptable levels.

Demonstration Results

A site demonstration of the biodenitrification system was conducted at the location of a former public water supply well in Bendena, Kansas. This study, from May until December of 1999, was conducted in cooperation with the Kansas Department of Health and Environment (KDHE). The KDHE provided a small building and necessary utilities for the EcoMat systems. In addition, the state is analyzing water samples independently.

The demonstration focused on treating contaminated water from the Bendena Rural Water District No. 2 Public Water Supply (PWS) Well No.1. This former railroad well, constructed in the early 1900s, was at one time the sole source of water for the town of Bendena. The primary contaminant in the water is nitrate from uncertain sources ranging from 20 to 130 mg/L. Low concentrations of VOCs, particularly carbon tetrachloride (CCl_4), in the groundwater, ranging from 2 to 31 g/L, are a secondary problem.

The main goal of the study was to demonstrate that its biodenitrification system could reduce incoming nitrate-N in excess of 20 mg/L to a combined nitrate plus nitrite concentration below 10 mg/L. A second goal of the study was to demonstrate that the posttreatment system used would produce treated water that would meet applicable drinking water standards with respect to nitrate-N and nitrite-N; and that the final effluent would not have turbidity greater than 1 NTU, detectable levels of methanol (1 mg/L), increased levels of biological material or suspended solids, and a pH in the acceptable 6.5 to 8.5 range.

To evaluate both the biodenitrification system and the posttreatment system adequately, water samples were collected from four specific points along the entire process, which were (1) an influent sample point between PWS #1 and R1, (2) a partial treatment sample point between R1 and R2, (3) an intermediate effluent sample point between the biodenitrification system and the posttreatment system, and (4) a final effluent sample point downstream of the posttreatment system. To ensure a statistically adequate number of samples, an average of 30 influent and 30 effluent samples were collected for each of four separate sampling episodes. Over an approximate seven and one-half month period, EcoMat operated its system at a flow between 3 and 8 gallons per minute. Results from the biodenitrification process were encouraging when the entire system was operating at optimal performance. In those instances where the final combined nitrate-nitrite effluent concentration was above the regulatory limit, operating problems (mostly mechanical) were suspected as the primary cause.

REVERSE OSMOSIS TREATMENT

Reverse osmosis (RO) systems can often improve the quality of water. Reverse osmosis water treatment has been used extensively to convert brackish water or seawater to drinking water, to clean up wastewater, and to recover dissolved salts from industrial processes. It is becoming more popular in the home market as homeowners are increasingly concerned about contaminants that affect their health and about nonhazardous chemicals that affect the taste, odor, or color of their drinking water. People considering the installation of a water treatment system to reduce toxic chemicals should first have their water tested to determine how much, if any, hazardous compounds are in the water. Public water supplies are routinely monitored and treated as required under the federal Safe Drinking Water Act and state regulations. Private water systems should be tested at the owner's initiative based on knowledge of land use and contamination incidents in the area (9).

Reducing Contaminants through RO

Reverse osmosis reduces the concentration of **dissolved solids**, including a variety of ions and metals and very fine suspended particles such as asbestos, that may be found in water. An RO device may be installed following a water softener to reduce the concentration of sodium ions exchanged for hardness ions. RO also removes nitrate (NO_3^-), arsenic (As), chromium (Cr), certain organic contaminants, some detergents, and specific pesticides.

The Reverse Osmosis Process

In reverse osmosis, a cellophane-like membrane separates purified water from contaminated water. An understanding of osmosis is needed before further describing RO. Osmosis occurs when two solutions containing different quantities of dissolved chemicals are separated by a semipermeable membrane that allows only some compounds to pass through. Osmotic pressure of the dissolved chemical causes pure water to pass through the membrane from the dilute to the more concentrated solution. In reverse osmosis, water pressure applied to the concentrated side forces the process of osmosis into reverse. Under enough pressure, pure water is "squeezed" through the membrane from the concentrated to the dilute side. Salts dissolved in water as charged ions are repelled by the RO membrane. Treated water is collected in a storage container. The rejected impurities on the concentrated side of the membrane are washed away in a stream of wastewater, not accumulated as on a traditional filter. The RO membrane also functions as an ultrafiltration device, screening out particles, including microorganisms, that are physically too large to pass through the membrane's pores. RO membranes can remove compounds in the 0.0001 to 0.1 micron range (thousands of times smaller than a human hair).

Design of an RO System

Although the reverse osmosis process is simple, a complete water treatment system is often complex, depending on the quality of the incoming water before treatment and the consumer's needs. Most home RO systems are point-of-use (POU) units placed beneath the kitchen sink to treat water used for cooking and drinking. Point-of-entry (POE) systems that treat all the water entering the household are more expensive to purchase and operate than POU systems.

A typical home reverse osmosis system consists of pretreatment and posttreatment filters as well as an RO membrane, flow regulator, storage container for the treated water, and dispensing faucet. The pressure for RO is usually supplied by the feed line pressure of the water system in the home, but a booster pump may be needed to produce an adequate volume of treated water. A sediment prefilter is essential for removing relatively large sand grains and silt that may tear or clog the RO membrane or clog a pump or flow regulator. Water softeners are used in advance of the RO system when household water is excessively hard. If the water is chlorinated or contains other oxidizing chemicals such as bromine, an activated carbon prefilter is needed to protect membranes sensitive to these chemicals.

To remove certain pesticides and organic solvents, activated carbon (AC) posttreatment must be included in the system. A standard AC filter positioned after the storage tank removes compounds that cause unpleasant taste and odors, including those from the tank or plastic tubing, just before water is dispensed. To remove high levels of organic chemicals such as trihalomethanes, volatile organic chemicals, and chloramines, an additional prolonged contact carbon filter (PCCF) is placed between the RO membrane and the storage tank. Combining an activated carbon filter with RO expands the range of chemicals the system can remove. Furthermore, AC treatment is improved because RO removes compounds that adversely affect AC adsorption.

The storage tank, tubing, and dispensing faucet should be made of plastic, stainless steel, or other nontoxic materials. The low pH and mineral content of RO-treated water may corrode copper pipes and allow lead to leach into the drinking water from brass components.

RO Membrane Materials

The most common RO membrane materials are polyamide thin film composites (TFC) or cellulosic types [cellulose acetate (CA), cellulose triacetate (CTA), or blends]. Very thin membranes are made from these synthetic fibers. Membrane material can be spiral-wound around a tube, or hollow fibers can be bundled together, providing a tremendous surface area for water treatment inside a compact cylindrical element (Fig. 1). Hollow fiber membranes have greater surface area (and therefore greater capacity) but are more easily clogged than the spiral-wound membranes commonly used in home RO systems.

The flux, or capacity, of the RO membrane indicates how much treated water it can produce per day. Typically, RO membranes for home systems are rated in the range of 10 to 35 gallons per day. Thus, under standard operating conditions it could take from 2 to 6 hours to fill a two and-a-half-gallon storage tank. CA/CTA membranes have adequate capacity for most households, but TFC membranes should be used if large volumes of treated water are needed.

RO membranes are rated for their ability to reject compounds from contaminated water. A rejection rate (% rejection) is calculated for each specific ion or contaminant as well as for reduction of total dissolved solids (TDS). It is important that consumers know their specific requirements for water quality when buying a system. For example, high rejection rates are essential when high nitrate or lead concentrations in the water must be brought below the EPA maximum contaminant or action levels.

Efficiency of RO Systems

The performance of an RO system depends on membrane type, flow control, feed water quality (e.g., turbidity, TDS, and pH), temperature, and pressure. The standard at which manufacturers rate RO system performance is 77 °F, 60 pounds per square inch (psi), and TDS at 500 parts per million (ppm). Only part of the water that flows into an RO system comes out as treated water. Part of the water fed into the system is used to wash away the rejected compounds and goes down the drain as waste. The recovery rate, or efficiency, of the system is calculated by dividing the volume of treated water produced by the volume of water fed into the system. If not properly designed, RO systems can use large quantities of water to produce relatively little treated water. Most home RO systems are designed for 20 to 30% recovery (i.e., 2–3 gallons of treated water are produced for every 10 gallons put into the system). Home RO systems can operate at higher recovery rates, but doing so may shorten membrane life. The flow regulator on the reject stream must be properly adjusted. If the flow is slow, the recovery rate is high, but RO membranes are easily fouled if concentrated impurities are not washed away quickly enough. If the flow is too fast, the recovery rate is low and too much water goes down the drain. Overall water quality affects the efficiency of an RO system and its ability to remove specific contaminants. The higher the TDS, the lower the recovery rate of treated water. The amount of treated water produced decreases by 1 to 2% for every degree below the standard temperature of 77 °F. An RO system supplied with well water at a temperature of 60 °F produces only three-quarters of the volume it would produce at 77 °F. For an RO system to function properly, there must be enough water pressure. Although most home RO systems are rated at 60 pounds per square inch, the incoming feed line pressure of many private water systems is less than 40 psi. The RO system must work against back pressure created in the storage tank as it fills with water and compresses the air in the tank. The RO device must also overcome osmotic pressure, bonding between water molecules and dissolved impurities; the higher the TDS level, the greater the osmotic pressure. The net water pressure at the RO membrane can be calculated by subtracting back pressure and osmotic pressure from feed line pressure. If the net water pressure at the membrane is lower than 15 psi, treated water production is less efficient and contaminant rejection rates are lower. Auxiliary pumps can be added to the treatment system to boost pressure and improve the quality and quantity of water produced. High-quality RO systems have valves that shut off the flow whenever storage tank pressure reaches two-thirds of the feed pressure; at that point, low net water pressure can result in low rejection rates. In some systems, once the storage tank is filled, surplus treated water is discarded; water loss from such units is frequently excessive. A system that automatically shuts off when the pressure on the tank reaches a given level saves water.

Maintenance of an RO System

An RO system must be well maintained to ensure reliable performance. Clogged RO membranes, filters, or flow controls decrease water flow and system performance. If fouling is detected in early stages, the membrane can often be cleaned and regenerated. The cleaning procedure varies depending on the type of membrane and fouling. Completely clogged or torn RO membranes must be replaced. In addition, pre- or postfilters must be replaced once a year or more often, depending on the volume of water fed through the system and the quality of the feed water. Damage to RO membranes cannot be seen easily. The treated water must be analyzed periodically to determine whether the membrane is intact and doing its job. Many systems now have a built-in continuous monitor that indicates a high TDS level, a sign that the system is not operating properly. It may also be necessary to test regularly for specific health-related contaminants such as nitrates or lead. Microorganisms, dead or alive, can clog RO membranes. To prevent biofouling, RO units must be disinfected periodically with chlorine or other biocides provided by the manufacturer. Continuous chlorination can be used with cellulosic membranes to protect the system from biofouling and eliminate the particle-trapping slime that worsens other forms of fouling such as scaling. Chlorine and other oxidizing disinfectants are harmful to thin film composite membranes. If the feed water is chlorinated, an activated carbon unit must be used to remove the oxidizing chemicals before they reach the TFC membrane. Activated carbon (AC) prefilters should not be used on nonchlorinated water supplies because they provide a place for microorganisms to multiply and lead to increased biofouling of the RO membrane surface. It is important to replace AC filters periodically following the manufacturer's instructions, especially after an extended shutdown period during which microorganisms can flourish.

Choosing an RO System

Homeowners who are thinking about buying reverse osmosis systems should determine their initial water quality and their goals in adding water treatment systems. RO removes many inorganic impurities from drinking water, especially nitrate. Its effectiveness depends not only on the type of membrane but on feed water quality, temperature, pressure, and flow control, as well as the type and concentration of specific contaminants to be removed. A typical RO system consists of a sediment filter, pump, reverse osmosis membrane, flow regulator, storage tank, final activated carbon filter (for taste and odors), and dispensing faucet. An AC prefilter is sometimes needed for dechlorination. RO is commonly used to treat only the water used for drinking and cooking at the point of use rather than at the point of entry for all household use. RO membrane types vary in their ability to reject contaminants and differ in capacity (the volume of treated water produced per day). Water pressure is an important factor in determining the RO system's rejection rate, capacity, and recovery rate (amount of treated water produced per amount of feed water used). Maintenance of an RO system is essential for reliable performance. High levels of TDS and microorganisms in the system are commonly the cause of fouled membranes. Treated water should be monitored for TDS and the level of any specific contaminants that may affect health.

A list of home water treatment devices certified by various state Department of Health Services can be found on their web sites.

INNOVATIVE *IN SITU* TREATMENT

The following are advanced studies for *in situ* groundwater treatment containing nitrate.

Metallic Ion Treatment

Metallic iron was selected as a new reductant for chemically reductive removal of nitrate from water in this study (10). The effect of pH on the reduction rate and products of nitrate was investigated using a fixed dosage of iron powder (size: 80 mesh) of 12 mol-Fe:mol-N^{-1} under acidic ($pH = 2-5$) and aerobic conditions. The reduction of nitrate by metallic iron was a pseudo-first-order reaction under our experimental conditions. The reduction rate of nitrate increased with decreasing pH of the reaction solution, and the pseudo-first-order reaction rate constants were 0.49, 0.40, 0.05, and 0.01 h^{-1} at $pH = 2$, 3, 4, and 5, respectively. The reduction products of nitrate were ammonia and nitrogen gas, and their yields were strongly affected by pH. The yields of nitrogen gas were 30, 31, and 45% at $pH = 3$, 4, and 5, respectively, but no formation of nitrogen was found at $pH = 2$. It was also demonstrated that the indirect reduction of nitrate by hydrogen generated from the reaction between protons and metallic iron may be a major mechanism for the reduction of nitrate under the experimental conditions.

Wetland Treatment

Constructed wetlands have become popular, mainly for polishing wastewater effluent (11). In the arid southwest, where natural wetlands have all but disappeared, constructed wetlands can be a significant wildlife attractant. In our arid climate, wildlife enhancement can be considered an "intermediate" use of wastewater before it is discharged or recharged to aquifers. Despite the widespread popularity of wetlands, our understanding of the biogeochemical transformations occurring in wetlands is lacking. Furthermore, the "black box" design approach commonly used for designing treatment wetlands is archaic.

The C:N Ratio as a Design Guideline. Baker also evaluated the effect of hydraulic loading rate (HLR) and carbon addition (chopped up cattails that were "fed" to the microcosms once a week) on nitrate removal in microcosms that received water containing 30 mg NO_3-N/L. We found that the "critical" C:N ratio for denitrification was 5:1 and that the denitrification rate constant varied linearly with carbon addition.

As wetlands are inherently well suited for nitrate removal (carbon supplied by plants; anoxic mats), potential N removal rates are much higher for wetlands receiving nitrate than for wetlands receiving nonnitrified wastewater. Nitrate treatment wetlands could have broad application for treating irrigation return flow, remediating nitrate-contaminated aquifers, and removing nitrate in nitrified wastewater effluents prior to recharge.

Model of N and C Transformations in Wetlands. During 1996, nitrogen transformations were studied in a wetland receiving lagoon effluent in Kingman, Arizona. In this study, various N species (SON, PON, ammonium, nitrate) were measured for 10 months at 13 locations along the longitudinal axis of the wetland. On average, the wetland removed 75% of input N. Effluent TN levels were consistently <15 mg/L and NO_3-N was always <10 mg/L. In the same study, it was found that effluent DOC concentrations represent the balance between the degradation of DOC in wetland inflow (lagoon effluent) and DOC produced by decomposition of plants.

Low-Tech Wastewater Reuse System. Others have studied the potential for using wetlands as part of a low-tech approach to treat and recharge wastewater in small towns in the United States and Mexico that have limited water supplies and must recycle wastewater if they continue to grow. Proposed systems would treat effluent from an existing aerated lagoon in a wetland; the effluent from the wetland would then be infiltrated through the soil to recharge groundwater. The lagoon would provide most of the BOD and SS removal. Further removal of SS, BOD (and DOC), and most of the N removal would occur in the wetland. Infiltration would remove additional DOC and filter most pathogens. Finally, we developed an integrated design concept that incorporates a "plug flow" component to assure high-level treatment and more natural components (pools; riparian areas) for wildlife habitat.

In Situ Iron Filing Curtain Treatment of Nitrate Contaminated Agricultural Runoff

The common agricultural practice of overapplying slow-release nitrate-based chemical fertilizers, such as ammonium nitrate, and livestock waste disposal practices can result in substantial levels of nitrate (NO_3^-) concentrating in agricultural water runoff that enters streams, rivers, and groundwater aquifers. In many agriculturally intensive regions, the groundwater nitrate concentration exceeds human drinking water standards. These agricultural regions quite often face restricted land development opportunities as strict government nitrate approval criteria are often imposed on developers to ensure that the problem is not worsened by adding new residential and commercial sources of potential nitrate contamination. Beyond the immediate human health impacts, nitrate-laden agricultural runoff that enters rivers and lakes also represents a significant contributor to excessive growth of algae blooms that can negatively impact these ecosystems.

In recognition of these impacts, a passive reactor system for containing and treating nitrate-contaminated agricultural runoff has been developed. Nitrate-laden runoff is gravity conveyed by land drain pipes or ditches to a collection reservoir. The contained runoff is then further gravity fed through a reactor system that is kept under anaerobic conditions and uses a reactive porous carbonaceous medium to convert the nitrates to a less harmful form of nitrogen. In most practical applications of the technology, the low-cost and easily constructed reservoir and reactor system can be passively connected to the end of an existing farm field drainage tile system without any requirement for pumping or chemical addition equipment and without maintenance requirements for many years after installation.

A small-scale field test was completed during June 1991 to May 1992, the results of which were documented in a paper accepted by the *Journal Contaminant Hydrogeology* (7/93). More recently, the technology has been commercially applied at several other sites in Ontario. Canadian patent #2068283 has issued, and the technology is available for licensing in Canada. U.S Patent #5,330,651 has been issued, and the technology is licensed in the U.S to Lombardo Associates Inc. (www.lombardoassociates.com), Contact:

G.G.H. Gray, Director
gghgray@uwaterloo.ca July 2002.

In Situ Electrokinetic Treatment

Electrokinetic treatment of nitrate, chromium, arsenic, and manganese may provide a cost-effective solution (12).

The electrokinetic treatment process involves applying direct current (dc) in a medium (soil and water). The flow of electrons from anode to cathode creates a migration of cations in the medium toward the cathode. The electrolysis of water creates a higher pH and oxidizes the metal or reduces the valence of the metal which renders the metal into a nontoxic form near the cathode. Clayey material may show a dramatic increase in cation exchange capacity under a high pH near the cathode. The proposed electrokinetic process can be applied both *in situ* and *ex situ*. Figure 1 presents a more realistic distribution of the various valence states of metals under the proposed *in situ* electrokinetic influence.

Besides the treatment of the metals, the beneficial side effects of the *in situ* EK treatment are as follows:

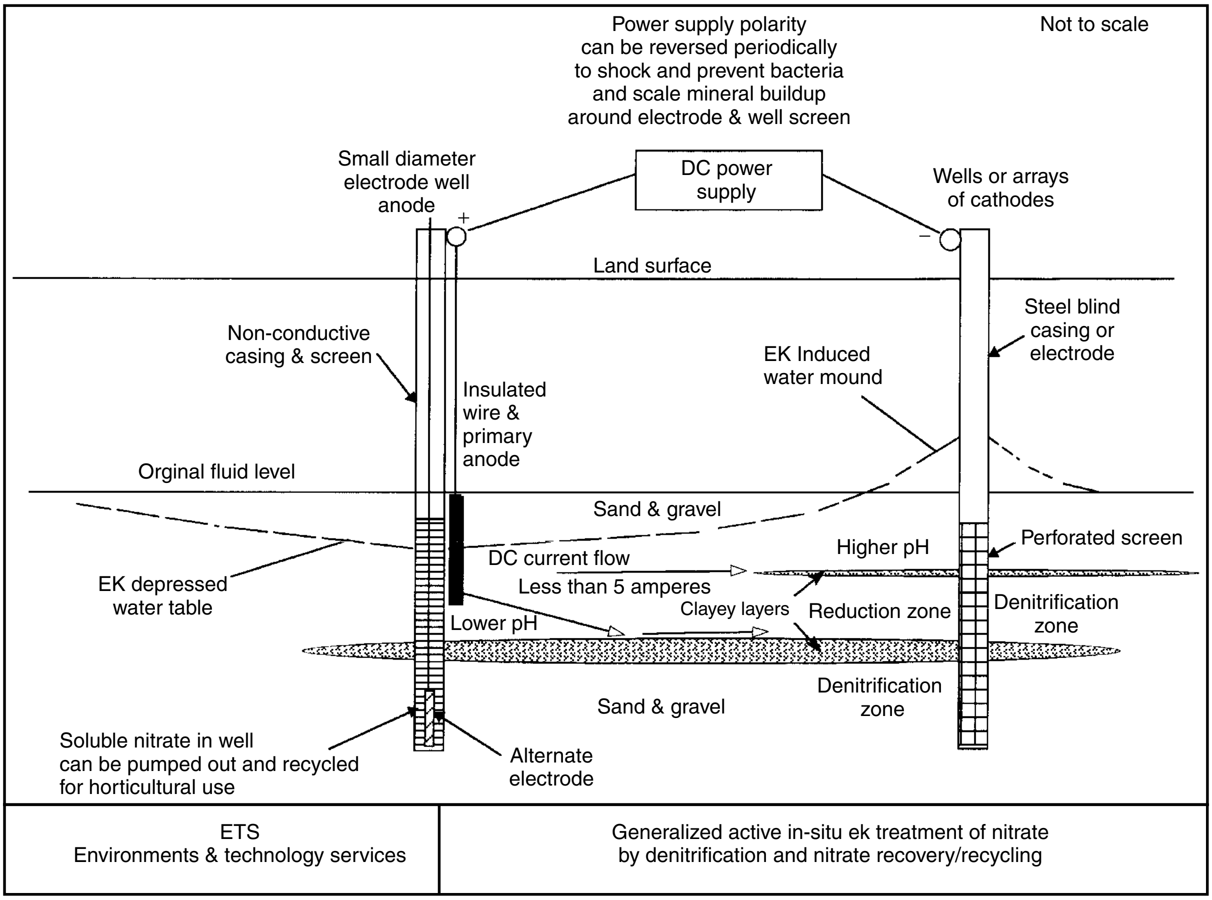

Figure 1. Generalized active *in situ* EK treatment of nitrate by denitrification and nitrate recovery/recycling.

1. The well itself is set up as cathode, so iron bacteria will not live on the surface of the well casing and perforation because of the high pH.
2. The electrokinetically induced water migration toward the cathode (well) may induce an increase in hydraulic head thus increasing well yield.

No large-scale treatment has been tried on nitrate, chromium, arsenic, and manganese in a saturated medium to date. There are many successes of the proposed EK process in the laboratory and in soil. However, we have successfully demonstrated the electrokinetic control of nitrate, selenium, and boron in clayey saturated media at two sites in the Panoche Irrigation District, Central Valley, California. We believe that the same EK control can be applied to nitrate because of the similarity of the electrochemical processes (Fig. 1).

Due to the simplicity of the cathode and anode setup, we believe that the proposed EK processes will prove very cost-effective. In particular, the proposed *in situ* EK treatment (once set up) is permanent. It has a one time capital cost and minimum long-term maintenance costs. The continuous operating electricity demand will not exceed 50 amperes at 30 to 100 Vdc or 1500 w to 5000 watts per site.

BIBLIOGRAPHY

1. Jennings, G.D. and Sneed, R.E. (1996). *Nitrate in Drinking Water*.
2. Downes, C.J. (1985). *Groundwater Quality*. Wiley, New York, Chap. 7, Fig. 7.4, p. 102.
3. Pye, V.I. and Kelley, J. (1984). *The Extent of Groundwater Contamination in the United States*. National Academy Press, Washington, DC, pp. 23–33.
4. Devinny, J.S. et al. (1990). Subsurface migration of hazardous wastes. In: *Microbiology of Subsurface Wastes*. Van Nostrand Reinhold, New York, pp. 116–141.
5. Freeze, A.R. and Cherry, J.A. (1979). *Groundwater*. Chap. 2: Physical Properties, pp. 15–79; Chap. 9.3: Hydrochemical Behavior of Contaminant–Nitrogen, pp. 413–416, Prentice-Hall, Englewood Cliffs, NJ.
6. Hem, J.D. (1989). *Study and Interpretation of the Chemical Characteristics of Natural Water—Occurrence of Nitrogen in Water*. US Geological Survey Water-Supply Paper 2254, pp. 125–126.
7. Zuane, J.D. (1997). *Handbook of Drinking Water Quality*.
8. EPN. Superfund Innovative Technology Evaluation Demonstration Bulletin Biological Denitrification Process - EcoMat, Inc. EPA/540/MR-01/501.
9. Kneen, B., Lemley, A., and Wagenet, L. (1995). *Reverse osmosis treatment of drinking water*. Fact Sheet 4. Cornell Cooperative Extension, New York State College of Human Ecology.
10. Hu, H.-Y. et al. (2001). Reductive treatment characteristics of nitrate by metallic iron in aqueous solution. *J. Chem. Eng. Jpn.* **34** (9): 1097–1102.

11. Gerke, S., Baker, L., and Xu, Y. (2001). Sequential model of nitrogen transformations in a treatment wetland receiving lagoon effluent. *Water Res.* **35** (16): 3857–3866.

12. Lehr, J.H. (Ed.). (2001). *Standard Handbook of Environmental Health, Science and Technology*. Chapters 14.4 & 14.6 on Bioremediation & Electrokinetic Treatment of Hazardous Wastes by Walter W. Loo, MacGraw-Hill, New York.

NONPOINT SOURCES

Chakresh K. Jain
National Institute of Hydrology
Roorkee, India

In recent years, it is being recognized that the nonpoint sources of water pollution have greater importance than point sources, which is due in part to the continuing efforts to reduce pollution from point sources during the past few decades, as well as recognition that nonpoint sources, such as storm water, may contain harmful contaminants. Modern agriculture has also relied, among other things, on heavy use of fertilizers and pesticides. This has largely occurred without any serious regard to the environmental consequences, especially pollution of both surface and groundwater resources. Chemicals (fertilizers and pesticides) applied in agroecosystems invariably enter the hydrologic cycle and contaminate both surface and groundwater resources. Nutrients and pesticides, particularly, are of major concern because of eutrophication and toxicity. Nutrients and pesticides are transported from cropland either by being absorbed onto eroded soil particles or dissolved in runoff water, which has increased the need to identify and quantify major sources of nutrients and pesticides deposited within river systems. In this article, some basic concepts of nonpoint source pollution, including the effects and extent of pollution in surface and groundwater bodies, pollution control measures, monitoring and modeling approaches, and management options are presented.

INTRODUCTION

The assessment of nonpoint source (NPS) pollution is a perfect example of a spatially and temporally complex, multidisciplinary environmental problem that exists on multiple scales. Nonpoint source pollution accounts for more than 50% of total water quality problems. In many areas, nonpoint source pollution, such as runoffs from cropland, urban storm water, strip mining, and construction sites are becoming major water quality problems (1). Agriculture is often considered the largest contributor to nonpoint source pollution of both surface and subsurface systems. It is mainly responsible for degrading river water quality by generating runoff from animal husbandry units, which contain predominantly organic compounds from the use of mineral fertilizers and chemical pesticides.

The problem of NPS pollution is being recognized and systematically studied in Western countries (2–8) this is not so in developing countries where agricultural pollution is usually complicated by urban and industrial pollution (9). The factors that influence nonpoint source pollution include soil erosion, sedimentation, and erosion of stream banks; washing out nutrients and organic material from livestock wastes and agricultural land; storm runoff from urban areas; and atmospheric deposition. Adsorption to the surface of sediment particles provides a mechanism for transporting many contaminants derived from agricultural fertilizers, pesticides, and industrial wastes. Deposition of sediments carrying such loads in a channel or on a flood plain can have detrimental consequences for ecology and agriculture. The sediment released into a river system can promote channel instability and cause bed degradation.

In most cases, the sources and concentrations of nonpoint source pollutants result from land use interactions with the transport system. It is a source transport problem in which the hydrologic cycle provides the transport processes to move pollutants from the source to groundwater, a stream, or a reservoir. Nonpoint sources can be urban, industrial, or agricultural pollutants that are distributed over the surface. The mode of transport is the flow of water across the soil surface and in stream channels and reservoirs or the flow of water through the soil profile.

The problem of nonpoint source pollution does not involve only the traditional pollution parameters such as suspended sediment, biochemical oxygen demand (BOD), and dissolved oxygen (DO). Some of the most serious nonpoint pollution problems do not have a parallel in the traditional point source oriented environmental pollution control area. These problems include nutrients (nitrogen and phosphorous); PCB (polychlorinated biphenyl) contamination; acid rain; and pesticide contamination of surface water, groundwater bodies, and aquatic biota.

In recent years, exponential growth in population and fragmentation of farm families have caused a reduction in land holding size, consequently forcing the farmers toward intensive cultivation (10) that leads to increased soil and nutrient loss. According to Wollast (11), the increase in the dissolved nitrogen flux in rivers represents 30% of the nitrogen fixed annually by humans in combustion processes and fertilizer production. For phosphorous, the increase is only 15% of the total phosphorous mined annually. The increased biological activity due to the entry of N and P in rivers had led in some cases to anoxia and consequent fish mortality.

To understand the nitrogen and phosphorous concentrations in surface waters, one should always take into account factors such as current and historic nutrient inputs to agricultural land from fertilization and atmospheric deposition, soil type, hydrology, historical landuse, and climatic conditions. Preferential flow also plays an important role in the subsurface losses of nutrients in many soils, especially after longer periods of drought.

Various pathways are involved in transporting N and P into a river system. Depending on the pH and temperature of soils, NH_4^+ and NO_3^- ions are produced in a watershed through ammonification and nitrification of organic matter and are mobilized into rivers through runoffs. The weathering of apatitic rocks is the major natural source of PO_4^{3-}. Rain is a source of dissolved N and S. N, P, and S from fertilizer application, sewage, and nonpoint source runoff contribute significant quantities of dissolved species of N, P, and S near human habitation.

Modern high yielding crops generally require adding fertilizers, which can be inorganic or organic. Annual uptake efficiencies by crops are relatively poor, approximately 50% and 20% for N and P, respectively (12,13). Fertilizer application rates have increased considerably during the last two–three decades (14), which has resulted in the possibility of increased NO_3-N losses from terrestrial systems into watercourses through surface runoff and/or by leaching through the soil profile (15). The relative contribution of individual sources and processes to N and P losses varies spatially and temporally; it is greatly influenced by climatic and management factors.

As per the National Water Quality Inventory Report, nonpoint agricultural pollution significantly affects water quality in 68% of the drainage basins in the United States (16). It has been reported that agricultural pollution contributes 60% of the BOD, 64% of suspended solids, and 76% of total phosphorous discharged into water resources (17).

In the Great Lakes region (18,19), the phosphorous contribution from nonpoint sources represents roughly 50% of the total load to Lakes Superior, Huron, and Erie and exceeds municipal and industrial point sources. The balance is due to atmospheric fallout and shoreline erosion. Phosphorous is the limiting nutrient for controlling eutrophication of most of the Great Lakes. More than 90% of the total loading to Lakes Superior, Michigan, Huron, and Ontario originates from nonpoint sources. In addition, the International Joint Commission research (19) has found that diffuse sources, including atmospheric inputs, account for the major loading of PCBs and organic chemicals to the lakes.

Several articles discuss the urgent need for and the problems of controlling nonpoint sources of water pollution from agriculture, and others evaluated the economic aspects of controlling NPS pollution. Wallace et al. (20) presented an example of a point-NPS trading framework developed to help a discharger comply with load allocation constraints in the Minnesota River. Hoag and Hughes-Popp (21) compared the theory of pollution credit trading and its application in the Tar-Pamlico nutrient-trading program in North Carolina, where point-NPS trading is possible. The article evaluated transaction costs, number and relative discharge of participants, abatement costs, enforcement costs, trading ratio, and loading limits and identified factors that encourage or discourage trades.

SURFACE WATER PROBLEMS

In recent years, it is being recognized that nonpoint sources of water pollution have greater importance than point sources. Several studies on the nutrient levels in surface waters from agricultural watersheds have been conducted in Pennsylvania, Maryland, Delaware, North Carolina, Colorado, Denmark, and Wisconsin (22). Increased nitrogen concentrations were correlated with a greater portion of cropland area in several Chesapeake Bay watersheds (23,24), whereas phosphorus concentrations did not correlate with land use. Sediments in agricultural drainage basins were reportedly enriched with phosphorous in Delaware's Inland Bays (25,26). Deletic and Maksimovic (27) studied the water quality of storm runoff at two experimental catchments in Belgrade, Yugoslavia, and Lund, Sweden. Rai and Sharma (28) made a comparative assessment of runoff characteristics under different land use patterns within a Himalayan watershed and reported that a large quantity of sediments leaves the watershed through its rivers.

Heathwaite and Johnes (29) studied the contribution of nitrogen and phosphorous from agricultural catchments by monitoring the chemical composition of these nutrients in agricultural runoff for both surface and subsurface flow pathways. In another publication, Heathwaite et al. (30) studied the role of nitrogen and phosphorous as key nutrients that determine the trophic status of water bodies. The paper primarily focuses on results from North America and Europe, particularly for the United Kingdom where a wide range of data exists. It is reported that, in general, N and P concentrations have increased dramatically (>20 times background concentrations) in many areas and causes vary considerably, ranging from urbanization to changes in agricultural practices. Johnes and Heathwaite (31) also modeled the impact of agriculture on the nitrogen and phosphorous loading in the surface waters of two contrasting agricultural catchments.

The concentrations of pesticides in surface waters in agricultural watersheds were monitored in Arkansas (32), Nebraska (33), Idaho and Wyoming (34), California (35), and France (36). The pesticides detected include Metolachlor, atrazine, norflurazon, cyanazine, simazine, isoproturon, and organophosphate in varying concentrations. Several studies have also been carried out to evaluate the impact of pesticides on management practices (37–40).

Laroche et al. (41) used model HSPF to study the transport of pesticides on a 78-ha watershed in the province of Quebec, Canada. Pesticide concentrations were measured at the outlet of the watershed from February to November 1993. This short period of time allowed calibrating the pesticide component but precluded verification. Variations of atrazine concentration were well simulated by the Hydrological Simulation Program FORTRAN (HSPF), simulated concentrations were in the same range as observed values, and peak concentrations occurred at the same time.

Surface water-quality impacts of urban and industrial land uses were monitored in Texas (42), Utah (43), North Carolina (44), South Africa (45), and Singapore (46). Results from these studies indicated that urban land uses contribute varying levels of metals, bacteria, sediment, nutrients, and organic chemicals to surface waters. Line et al. (44) showed that metal concentrations in runoff from industrial sites varied significantly for various categories of businesses. Regarding sources of nutrients in urban areas, Linde and Watschke (47) reported that, on average, 11% of applied phosphorus and 2% of applied nitrogen were detected in runoff from turf grass plots, whereas 14% of phosphorus and 3% of nitrogen were detected in leachate.

Bouraoui and Diallaha (48) developed a nonpoint source pollution management model (Answers-2000) to simulate long-term average annual runoff and sediment yield from agricultural watersheds. The model is based on the event-based Answers model and is intended for use without calibration. Wong et al. (49) developed a land use runoff model for Santa Monica Bay using a geographic information system (GIS) coupled with an empirical runoff model. The GIS/model has simple data requirements compared to more complex models that require routing information and is useful for predicting receiving water loading on an annual basis or for single storms.

GROUNDWATER PROBLEMS

There is wide concern in developed countries that groundwater resources are deteriorating in the long term, both in quantity and quality. Groundwater pollution is a major concern mainly because of the implications for human health when it is used for drinking water. Many of the most serious sources of health risk, such as bacterial and viral infection and toxic metals, do not have a specific agricultural origin. However, nitrates, as a cause of methaemoglobinaemia, and pesticides, which have a suspected disturbingly wide array of health effects, have largely agricultural origins and are a major concern.

Concentrations of nitrogen in groundwater are an important concern because a large number of people use groundwater as their drinking water supply. In addition, much of the nitrogen entering our streams and lakes comes from nitrates in groundwater. When urea is applied as a deicing agent during the winter, large concentrations of ammonium are observed at all soil depths; however, over time, the ammonium is oxidized to nitrate, which leaches into the groundwater (50).

Nolan et al. (51) prepared national maps of groundwater vulnerability to nitrate contamination for the United States using different input parameters (population density and the amount of nitrogen contributed by fertilizer, manure, and atmospheric sources) and aquifer vulnerability (soil drainage characteristics and the ratio of woodland acres to cropland acres in agricultural areas). In England and Wales, 125 groundwater sources supplying 1.8 million people had nitrate concentrations exceeding 50 mg/L (52).

Unfortunately, very little has been done in India to carry out well-planned primary field studies to assess the impact of nonpoint source pollution. Agarwal et al. (53) reviewed the status of the diffuse agricultural nitrate pollution of groundwater in India and reported that nitrate levels in groundwater across vast agricultural areas can be correlated with intensive irrigated agriculture, corresponding use of nitrogenous fertilizers, and groundwater development. It is further reported that diffuse agricultural pollution has already endangered the safety of potable groundwater for future generations in both rural and urban areas.

In recent years, there has been a considerable rise in the number of reported incidents of pesticide contamination in groundwater due to excessive use of pesticides in agriculture. Monitoring of pesticides in central South Dakota indicated that the movement of pesticides is controlled primarily by the geological and chemical characteristics of the medium and pesticides (54). Milde et al. (55) reported the occurrence of 22 different pesticides in 34 locations in the United States and Europe. In Britain, the most extensive survey of pesticides in surface waters and groundwaters has taken place in the Anglian region where herbicide usage on cereals has been the greatest.

Milde et al. (55) stated that the main factors influencing the contamination potential of any pesticide are molecular structure, formulation, metabolism, water solubility, adsorption by soil, the retardation factor, and resistance to chemical and biochemical degradation. The time, quantity, and frequency of application, as well as the local hydrogeologic and meteorologic conditions, also affect the quantity of pesticides that reach the water table. Kolpin et al. (56) found that 70% of all 106 municipal wells sampled were contaminated by pesticides and that the different types of aquifers affected pesticide leaching differently.

Numerous workers have studied the processes that affect the movement of pesticides into groundwater. Sorption of pesticides on different types of colloids and subsequent movement of these colloids through the soil profile were studied, and the researchers found almost no increase in the transport of atrazine (57). When fluxes of pesticides from a stream bed were measured, it was found that pesticide fluxes of atrazine were two to five orders of magnitude greater in the stream bed than fluxes of atrazine through agricultural fields (58). Guo et al. (59) found that the concentration of organic matter in the soil profile affects the transport of atrazine through the soil profile. Kolpin (60) studied the effects of land use on concentrations of agricultural chemicals and reported that the concentrations of atrazine and nitrate were positively correlated with the amount of irrigated land area and that alachlor was inversely related to the amount of highly erodible land.

Loague and Corwin (61) presented a regional-scale assessment of nonpoint source groundwater contamination, and Takarate et al. (62) studied the effect of hydrology, vegetation, and substrate on the spatial and temporal variations of nutrient concentrations in the groundwater of a flood plain. It is reported that on the regional scales associated with NPS agrochemical applications, there are staggering data management problems in assessing potential groundwater vulnerability. Geographical information system (GIS) is a timely tool that greatly facilitates the organized characterization of regional-scale variability. The strengths and weakness of using GIS in regional-scale vulnerability assessments (owing to both data and model errors) have also been discussed. A regional GIS-driven integrated assessment approach is discussed, which is based upon cost–benefit analysis and incorporates both physical and economic factors that can be used in a regulatory decision process.

POLLUTION CONTROL MEASURES

Concern for nonpoint source pollution in integrated water quality management has been growing recently. Such sources are likely to be more critical in developing countries, where agriculture and rural habitats are still dominant. The term pollution control refers to regulation of pollutants from individual outfalls or nonpoint sources, including urban developing lands, feedlots, agricultural areas, and other pollution generating land uses. The purpose of all pollution control programs is to protect the capacity of surface waters to assimilate pollution without damage or impairment of their use, to protect shellfish and wildlife, to preserve or restore the aesthetic and recreational values of surface waters, and to protect humans from adverse water quality conditions (63).

Water quality management deals with all aspects of water quality problems for all beneficial uses of water or lands from which pollution originates, whereas pollution control is mostly understood as the safe disposal of wastewater and its treatment. Water quality and pollution are determined by comparing measured physical, chemical, biological, microbiological, and radiological parameters to a set of standards and criteria.

The water quality standards used presently by water pollution engineers and scientists as well as pollution abatement authorities throughout the world are either stream standards or effluent standards. The effluent standards, which determine how much pollution can be discharged from municipal and industrial wastewater sources, are of lesser importance in nonpoint source pollution control and management. Performance standards, the equivalent of effluent standards to control pollution from lands, have been used by some local authorities to control pollution from subdivisions, construction areas, and mining. The stream standards can be related to the protection of aquatic habitat and biota and/or to intended downstream use of water.

Nutrient Control

The control of nutrients is an important issue throughout the world, both from a public health perspective and to keep natural waters free from eutrophication. The most widely used water quality standard for nitrate is the 50 mg/L limit adopted by the WHO as a precautionary level to safeguard babies from the risks of contracting methaemoglobinaemia (64). Most national authorities regard the 50 mg/L concentration as a realistic target in relation to eutrophication and, therefore, programs aimed at controlling eutrophication often use this value as an EQS. Whereas there are techniques available to remove nitrate from drinking water after it has been abstracted, the eutrophication problem is universally dependent on the control of nitrate sources.

An added and unexpected health implication related to eutrophication exists, particularly in lakes. In a number of countries in the world, the phenomenon of excessive growth of blue-green algae has caused concern where recreational pursuits take place on the lakes and also where the water is withdrawn for public supply, which is due to the recognition that such algae produce a number of toxins, which, if ingested, can cause liver damage. There are a number of well-documented cases where animals are known to have died as the result of drinking water heavily laden with these algae and, for this reason alone, nutrient control is justified.

The reduction of nitrogen and phosphorous from agriculture relies upon changes in farming practices because they give rise to diffuse sources. Plowing of grassland and other crops, particularly during autumn, leads to the release of large quantities of soil nitrogen and, therefore, a general move toward permanent pasture regimes assists in lowering nitrate leaching. When this is not possible, the use of short-term rotational crops to take up nitrogen followed by their harvesting and subsequent removal from the catchment is helpful. Animal wastes should be used carefully, avoiding overuse and direct runoff into watercourses, but wherever possible they should be used in place of synthetic fertilizers. Spreading of cattle manure and liquid sludge should be banned when aquifers are being replenished by rains. A manuring season should be prescribed for main crops depending on monsoon and crop patterns in different regions. The use of biofertilizers and green manures should be encouraged, and research needs to be undertaken to improve their quality and efficiency. Use of all types of fertilizers should be carefully controlled and matched to crop requirements.

One of the most important aspects of controlling phosphorous inputs from agriculture is the need to prevent erosion from field surfaces. Phosphates tend to bind to soil particles which, washed from fields into watercourses, become a source of phosphate in suspended form and in deposited sediments. Sediments act as a long-term source of phosphate by releasing it (i.e., by redissolution) under certain environmental conditions. Physical removal of the sediment layer to remove the bound phosphate from the catchment has been tried in a number of locations around the world. Some success has been achieved in lowering phosphate levels in the Norfolk Boards in England by a combination of diverting effluents containing phosphorous out of the area, phosphorous stripping at sewage treatment works, and by dredging 1 m of sediment. Concentrations below the target of $100\,\mu\text{g L}^{-1}$ of phosphorous were reached (65).

In addition to agricultural sources, phosphate occurs in sewage effluents as a result of its use in detergents. Control of phosphate at the source is not yet practiced, and removal relies on stripping of the orthophosphate from the sewage works effluent by chemical treatment. Reductions in the polyphosphate content of some detergents may also help in controlling phosphate.

Pesticide Control

Pesticides represent a particularly difficult area of pollution control, not only because the environmental effects in relation to aquatic flora and fauna are important but because human health issues have a very important bearing on the nature of the control applied. There are several thousand formulations of insecticides, herbicides, and fungicides in common use and, therefore, the potential for water pollution is very high. There are also very stringent limits on pesticides in water used for public supply and, consequently, the control of pollution from pesticides is quite crucial.

Many pesticides in water are derived from agricultural use, and it is therefore difficult to regulate their input to water bodies. Regulation can take place only by prevention, by indirect controls on their manufacture, storage, and use. Approval for the use of pesticides is granted in the United Kingdom by the government, following expert assessment of safety and environmental risks. The active ingredients in pesticides are approved by the European Union and placed on an approved list. This Directive allows authorization provided that the pesticide is not expected to occur in groundwater at concentrations above 0.1 μg/L. Further controls are also placed on products by Directives such as that on Classification, Packaging and Labelling of Dangerous Substances.

WATERSHED MANAGEMENT PLANNING

Watershed management essentially relates to soil and water conservation by proper land use, preventing deterioration of the soil, increasing and maintaining soil fertility, reducing soil erosion, conserving water for drinking and other farm uses, increasing the availability of basic resources, and achieving the optimum productivity of land uses (66). Rai et al. (67) described the traditional conservation practices and relationship to land use in Mamlay watershed in the Sikkim Himalaya. It was observed that the agricultural land area has considerably increased during the past 40 years (66,67). The soil without tree cover on steep upland farming systems associated with more intensive agricultural practices is vulnerable to soil erosion and reduced fertility (10).

In several countries, the process of catchment or watershed planning has been introduced. Catchment management plans are designed to cover the catchment of a river and its tributaries and any associated groundwater flows. Catchment management plans are designed to be strategic in approach and to take into account regional and national policies and all activities likely to have an impact on the watercourse. As a result of the plan, the main polluting sources may be identified and solutions agreed to which can be achieved with available resources and on a timescale agreed to by all concerned.

CONCLUSION

Nonpoint source pollution is a worldwide problem and has been responsible for contamination of both surface as well as groundwater resources. The problem is a complex function of many physical, chemical, and biological interactions in relation to soil type, management practices, and climatic conditions. As progress has been made in reducing pollutant loading from point sources, the relative importance of nonpoint sources has increased. Such sources are likely to be even more critical in developing countries, where agriculture and rural habitats are still dominant. A number of special features need to be considered to tackle the problem of nonpoint source pollution. These include (1) extremely varying rainfall and stream flow pattern; (2) still largely traditional agricultural practices, average application of fertilizers and pesticides, and significant areas under farming or only marginal irrigation; (3) a very large cattle population, where agriculture is almost always linked to animal husbandry; and (4) a culture of living close to the river with dominating in-stream uses of bathing, washing, cattle wading, waste disposal, etc. and large-scale floodplain farming. There is a need to address the problems of nonpoint source pollution and to manage it as carefully as point source pollution. Catchment management plans may be designed to cover catchments of different river systems and their tributaries and associated groundwater flows. The plans should be strategic in approach and should take into account regional and national policies and all other such activities, which are likely to have an impact on the watercourse.

BIBLIOGRAPHY

1. Alexander, G.R. (1976). Best management practices for non-point source control. In: *Proc. of the Best Management Practices for Non-point Source Pollution Control Seminar*, EPA 905/76-005, U.S. Environmental Protection Agency, Chicago, IL, p. 3–5.
2. Strebel, D., Duynisveld, W.H.M., and Bottcher, J. (1989). Nitrate pollution of ground water in western Europe. *Agri. Ecosystem and Environment* **26**: 189–214.
3. Fried, J.J. (1991). Nitrates and their control in the EEC aquatic environment, Nitrate Contamination: Exposure, Consequence, and Control. *NATO ASI Ser. G.: Ecological Science 30*. I. Bogardi and R.D. Kuzelka (Eds.). Springer Verlag, New York.
4. Frink, C.R. (1991). Estimating nutrient exports to estuaries. *J. Environ. Qual.* **20**: 717–714.
5. Goodrich, J.A., Lykims, Jr., B.W., and Clark, R.M. (1991). Drinking water from agriculturally contaminated groundwater. *J. Environ. Qual.* **20**: 707–717.
6. USEPA. (1992). *1990 National Water Quality Inventory*. Report to Congress, Washington, DC, p. 305b.
7. Cooper, C.M. (1993). Biological effects of surface water pollutants on aquatic systems—a review. *J. Environ. Qual.* **22**: 402–408.
8. Spalding, R.F. and Exner, M. (1993). Occurrence of nitrate in groundwater: A review. *J. Environ. Qual.* **22**: 392–402.
9. Duda, A.M. (1993). Addressing non-point sources of water pollution must become an international priority. *Water Sci. Tech.* **28**(3–5): 1–11.
10. Rai, S.C. and Sharma, E. (1995). Land use change and resource degradation in Sikkim Himalaya: A case study from the Mamlay watershed. *Sustainable Reconstruction of Highland and Headwater Regions*. R.B. Singh and M.J. Haigh (Eds.). Oxford & IBH Publishing, New Delhi, pp. 265–278.
11. Wollast, R. (1983). Interactions in esturies and coastal waters. In: *The Major Biogeochemical Cycles and Their Interactions*. B. Bolin and R.B. Cook (Eds.). John Wiley, Chichester, UK, pp. 385–407.
12. Nielsen, N.E., Schjorring, J.K., and Jensen, H.E. (1987). Efficiency of fertilizer uptake by spring barley. In: *Nitrogen Efficiency in Agricultural Soils*. D.S. Jenkinson and K.A. Smith. (Eds.). Elsevier Applied Science.
13. Holford, I.C.R. and Doyle, A.D. (1993). The recovery of fertilizer phosphorous by wheat, its agronomic efficiency and their relationship to soil phosphorous. *Aus. J. Agric. Res.* **44**: 1745–1756.
14. World Fertilizer Consumption Statistics. (1992). International Fertilizer Industry Association, No. 24.
15. Vinten, A.J.A., Howard, R.S., and Redman, M.H. (1991). Measurement of nitrate leaching losses from arable plots

under different nitrogen input regimes. *Soil Use Magmt.* **7**: 3–14.

16. National Water Quality Inventory. (1978). Report to Congress U.S. Environmental Protection Agency, Washington, DC.
17. Duda, A.M. and Johnson, R.J. (1985). Cost effective targeting of agricultural non-point source pollution controls. *J. Soil Water Conserv.* **40**: 108.
18. Sonzogni, W.C., Monteith, T.J., Bach, W.N., and Hughes, V.G. (1978). *United States Great Lakes Tributary Loadings*. International Joint Commission, Ontario, Canada.
19. PLUARG. (1978). *Environmental Management Strategy for the Great Lakes Systems*. International Joint Commission, Ontario, Canada.
20. Wallace, S., Micheletti, B., and Sparks, C. (1997). Nonpoint source trading creates new discharge opportunities. *Industrial Wastewater* **5**: 12.
21. Hoag, D.L. and Hughes-Popp, J.S. (1997). Theory and practice of pollution credit trading in water quality management. *Rev. Agric. Econ.* **19**: 252.
22. Owens, D.W., Corsi, S.R., and Rappold, K.F. (1997). *Evaluation of nonpoint-source contamination, Wisconsin: selected topics for water year 1995*. U.S. Geol. Surv. Open-File Rep. 96-661 A, Denver, CO.
23. Jordan, T.E., Correll, D.L., and Weller, D.E. (1997). Effects of agriculture on discharges of nutrients from coastal plain watersheds of Chesapeake Bay. *J. Environ. Qual.* **26**: 836.
24. Jordan, T.E., Correll, D.L., and Weller, D.E. (1997). Nonpoint source discharges of nutrients from Piedmont watersheds of Chesapeake Bay. *J. Am. Water Resour. Assoc.* **33**: 631.
25. Sallade, Y.E. and Sims, J.T. (1997). Phosphorus transformations in the sediments of Delaware's agricultural drainageways: I. Phosphorus forms and sorption. *J. Environ. Qual.* **26**: 1571.
26. Sallade, Y.E. and Sims, J.T. (1997). Phosphorus transformations in the sediments of Delaware's agricultural drainageways: II. Effect of reducing conditions on phosphorus release. *J. Environ. Qual.* **26**: 1579.
27. Deletic, A.B. and Maksimovic, C.T. (1998). Evaluation of water quality factors in storm runoff from paved areas. *J. Environ. Eng., ASCE* **124**(9): 869–879.
28. Rai, S.C. and Sharma, E. (1998). Comparative assessment of runoff characteristics under different land use patterns within a Himalayan watershed. *Hydrol. Process.* **12**: 2235–2248.
29. Heathwaite, A.L. and Johnes, P.J. (1996). Contribution of nitrogen species and phosphorus fractions to stream water quality in agricultural catchments. *Hydrol. Process.* **10**: 971–983.
30. Heathwaite, A.L., Johnes, P.J., and Peters, N.E. (1996). Trends in nutrients. *Hydrol. Process.* **10**: 263–293.
31. Johnes, P.J. and Heathwaite, A.L. (1997). Modelling the impact of land use change on water quality in agricultural catchments. *Hydrol. Process.* **11**: 269–286.
32. Senseman, S.A. et al. (1997). Trace level pesticide detection in Arkansas surface waters. *Environ. Sci. Technol.* **31**: 395.
33. Ma, L. and Spalding, R.F. (1997). Herbicide persistence and mobility in recharge lake watershed in York. *Nebraska, J. Environ. Qual.* **26**: 115.
34. Clark, G.M. (1997). Occurrence and flux of selected pesticides in surface water of the upper Snake River basin, Idaho and Western Wyoming. *J. Soil Water Conserv.* **52**: 381.
35. Domagalski, J.L., Dubrovsky, N.M., and Kratzer, C.R. (1997). Pesticides in the Joaquin river, california: input from dormant sprayed orchards, *J. Environ. Qual.* **26**: 454.
36. Garmouma, M., Blanchard, M., Chesterikoff, A., Ansart, P. and Chevrcuil, M. (1997). Seasonal transport of herbicides (triazines and phenylureas) in a small stream draining an agricultural basin: melarchez (France). *Klates Res.* **31**: 1489.
37. Zhang, X.C., Norton, L.D., and Hickman, M. (1997). Rain pattern and soil moisture current effects on atrazine Metolachlor losses in runoff. *J. Environ. Qual.* **26**: 1539.
38. Hong, S. and Smith, A.E. (1997). Potential movement of dithiopyr following application to golf courses. *J. Environ. Qual.* **26**: 379.
39. Lennartz, B., Louchart, X., Voltz, M., and Andrieux, P. (1997). Diuron and simazine losses to runoff water in Mediterranean vineyards. *J. Environ. Qual.* **26**: 1493.
40. Turnbull, A.B. et al. (1997). Assessment of the fate of selected adsorptive pesticides at ADAS Rosemaund. *J. Chartered Inst. Water Environ. Manage.* **11**: 24.
41. Laroche, A.M., Gallichand, J., Lagace, R., and Peasant, A. (1996). Simulating atrazine transport with HSPF in an agricultural watershed. *J. Environ. Eng., ASCE* **122**(7): 622–630.
42. Characklis, G.W. and Wiesner, M.R. (1997). Particles, metals, and water quality in runoff from large urban watershed. *J. Environ. Eng.* **123**: 753.
43. Farrell-Poe, K.L., Ranjha, A.Y., and Ramalingam, S. (1997). Bacterial contributions by rural municipalities in agricultural watersheds. *Trans. Am. Soc. Agric. Eng.* **40**: 97.
44. Line, D.E. et al. (1997). Water quality of first flush runoff from 20 industrial sites. *Water Environ. Res.* **69**: 305.
45. Jagals, P. (1997). Stormwater runoff from typical developed and developing South African urban developments: definitely not for swimming. *Water Sci. Technol. (G.B.)* **35**: 133.
46. Chui. P.C. (1997). Characteristics of stormwater quality from two urban watersheds in Singapore. *Environ. Monit. Assess.* **44**: 173.
47. Linde, D.T. and Watschke, T.L. (1997). Nutrients and sediment in runoff from creeping bentgrass and perennial ryegrass turfs. *J. Environ. Qual.* **26**: 1248.
48. Bouraoui, F. and Diallaha, T.A. (1996). Answers-2000: Runoff and sediment transport model. *J. Environ. Eng., ASCE* **122**(6): 493–502.
49. Wong, K.M., Strecker, E.W., and Stenstrom, M.K. (1997). GIS to estimate storm-water pollutant mass loading. *J. Environ. Eng., ASCE* **123**(8): 737–745.
50. Swensen, B. and Singh, B.R. (1997). Transport and transformation of urea and its derivatives through a mineral soil. *J. Environ. Qual.* **26**: 1516.
51. Nolan, B.T., Ruddy, B.C., Hitt, K.J., and Heisel, D.R. (1997). Risk of nitrate in groundwaters of the United States—a national perspective. *Environ. Sci. Technol.* **31**: 2229.
52. Department of the Environment. (1986). *Nitrates in Water—Report by the Nitrate Coordination Group*. Pollution Paper No. 26.
53. Agarwal, G.D., Lunkad, S.K., and Malkhed, T. (1999). Diffuse agricultural nitrate pollution of ground water in India. *Wat. Sci. Tech.* **39**(3): 67–75.
54. Bhatt, K. (1997). Occurrence and distribution of nitrate and pesticides in Bowdle aquifer, South Dakota. *Environ. Monit. Assess.* **47**: 223.
55. Milde, G., Pribyl, J., Kiper, M., and Friesel, P. (1982). Problems of pesticides use and the impact on groundwater. *IAH*

International Symposium on Impact of Agricultural Activities on Groundwater, Prague.

56. Kolpin, D.W. et al. (1997). Occurrence of selected herbicides and herbicide degradation products in Iowa's ground water, 1995. *Ground Water* **35**: 679.
57. Seta, A.K. and Karathanasis, A.D. (1997). Atrazine adsorption by soil colloids and co-transport through subsurface environments. *Soil Sci. Soc. Am. J.* **61**: 612.
58. Squillace, P.J., Burkart, M.R., and Simpkins, W.W. (1997). Infiltration of atrazine and metabolites from a stream to an alluvial aquifer. *J. Am. Water Resour. Assoc.* **33**: 89.
59. Guo, L., Wagenet, R.J., Hutson, J.L., and Boast, C.W. (1997). Nonequilibrium transport of reactive solutes through layered soil profiles with depth-dependent adsorption. *Environ. Sci. Technol.* **31**: 2331.
60. Kolpin, D.W. (1997). Agricultural chemicals in groundwater of the midwestern United States: Relations to land use. *J. Environ. Qual.* **26**: 1025.
61. Loague, K. and Corwin, D.L. (1998). Regional scale assessment of non-point source ground water contamination. *Hydrol. Process.*, 957–965.
62. Takarate, N., Sanchez-Perz, J.M., and Tremolieres, M. (1999). Spatial and temporal variations of nutrient concentrations in the ground water of a flood plain: effect of hydrology, vegetation and substrate. *Hydrol. Process.* **13**(10): 1511–1526.
63. Patrick, R. (1975). Some thoughts concerning correct management of water quality. In: *Urbanization and Water Quality Control*. W. Whipple, Jr. (Ed.). AWRA, Minneapolis, MN.
64. WHO. (1993). *Guidelines for Drinking Water Quality*. Vol. 1, Recommendations, 2nd Edn., World Health Organization, Geneva.
65. RCEP. (1992). *Royal Commission on Environmental Pollution, Freshwater Quality, Sixteenth Report, 1996*. Her Majesty's Stationary Office, London, pp. 65–67.
66. Sharma, E., Sundriyal, R.C., Rai, S.C., Bhatt, Y.K., Rai, L.K., Sharma, R., and Rai, Y.K. (1992). *Integrated Watershed Management*. Gyanodaya Prakashan, Nainital, p. 120.
67. Rai, S.C., Sharma, E., and Sundriyal, R.C. (1994). Conservation in the Sikkim Himalaya: Traditional knowledge and land use of the Mamlay watershed. *Env. Conserv.* **21**(1): 30–34.

ORGANIC COMPOUNDS IN GROUND WATER

Thomas B. Boving
University of Rhode Island
Kingston, Rhode Island

Virtually all natural waters contain at least some amounts of organic compounds. However, as human impact on the environment increases, the concentration of organic compounds, such as solvents or petroleum products, in the hydrosphere has increased in many places. Consequently, many aquifers experience a degradation of groundwater quality, and there is strong concern how to protect this valuable natural resource effectively from further contamination.

Organic compounds in groundwater can be of natural or anthropogenic origin. Naturally occurring organic compounds in terrestrial water include carbohydrates (sugars and cellulose), lipids (neutral fats), and proteins. The building blocks of proteins are amino acids, whereas functional proteins that act as biological catalysts are enzymes. Proteins are biopolymers that can form complexes with other organic compounds. Such compounds include tannins, which are phenolic compounds that occur naturally in bark, wood, leaves, or roots of many plants. The presence of elevated tannin levels in water results in a yellowish-brown discoloration. The parameter, dissolved organic carbon (DOC), is a commonly used measure for the concentration of organic compounds in aqueous solution. DOC is the fraction of total organic carbon (TOC is all carbon atoms covalently bonded in organic molecules) in water that passes through a 0.45 micron pore-diameter filter. Typical DOC concentrations for groundwater range from a few micrograms to several 10 mg per liter and more. DOC concentrations depend on various factors, such as climate, soil and vegetation coverage, and land use.

Many organic compounds, such as proteins, lignin, or cellulose, are suspended in water. These compounds are solid phase macromolecules. The size of the suspended matter is between that of individual small molecules and macroscopic objects. Particles of this size (ranging from approximately a few nanometer to millimeter are known as colloidal organic matter). The principal source of organic matter in water is from decaying plant material or kerogen, which is the organic matter deposited in the sediments of an aquifer. Organic matter in groundwater is in part made up of humic substances. The humic substances are further subdivided into humic acid and fulvic acids. Humic acid is the fraction of humic substances that is not soluble in water under acidic conditions ($\mathrm{pH} < 2$) but is soluble at higher pH values, whereas fulvic acid is the fraction of humic substances that is soluble in water under all pH conditions. The fraction of humic substances that is not soluble in water at any pH value is referred to as humin. Humic substances are biopolymers that vary in color from light yellow-brown (fulvic acid) to dark brown (humic acid) to black (humin). The carbon based chemical structure of humic substances—especially that of humic acids and humin compounds—is very complex and is characterized by high molecular weight ($\gg$2,000 atomic mass units). The size of the macromolecules can range from 60–500 Å (1). Under the electron microscope, humic substances take the form of rings, chains, and clusters. The polarity of these compounds is intermediate between that of many (hydrophobic or nonpolar) hydrocarbons and that of polar water. For this reason, organic matter is an attractive nonpolar environment for hydrophobic compounds in aqueous solution. The partitioning of hydrophobic organic compounds to organic matter surfaces is referred to as adsorption, whereas partitioning into the organic matter is referred to as absorption. Although sorption to the solid phase of the aquifer would essentially stop or at least retard the movement of hydrophobic compounds, both adsorption and absorption to mobile organic compounds can greatly facilitate their transport in groundwater. Facilitated transport has important implications for assessing the fate and transport of organic contaminants in aqueous solution. The concentration of humic substances in groundwater increases as the DOC concentration increases and typically ranges from about

20 to 66% of the DOC mass (2). Elevated concentrations of organic matter in natural water can cause oxygen depletion, taste and odor problems, and can interfere with the water treatment process (formation of trihalomethanes as disinfection by-products during water chlorination).

Besides naturally occurring organic compounds, many waters contain compounds of anthropogenic origin. Following the definition of Schwarzenbach et al. (3), an anthropogenic compound is one that is introduced into the environment primarily or exclusively as a consequence of human activities. Thus, synthetic compounds, such as chlorinated solvents or most pesticides, as well as naturally occurring compounds, such as petroleum components, can be found in both surface and groundwater. Given the large number of synthetic chemicals and petroleum products used for domestic, industrial, and agricultural applications (including solvents, pharmaceuticals, plastics, dyes, and detergents), a discussion of individual organic compounds in the aquatic system is impractical. Hence, chemicals that share certain structural or chemical similarities are commonly grouped together. For example, those compounds that have two or more fused benzene rings are known as polynuclear aromatic hydrocarbons or polycyclic aromatic hydrocarbons (PAH), and compounds that have several chlorine atoms attached to a central biphenyl structure are called polychlorinated biphenyls (PCB) (Fig. 1). Other grouping criteria are based on (1) a common application, for example, organic compounds that are fungicides, herbicides, or insecticides; (2) a common source of origin, such as petroleum compounds that are lumped together in a group know as total petroleum hydrocarbons (TPH); or (3) similar physical properties, such as volatility. The two most often used volatility based classifications are volatile organic compounds (VOC), which include compounds such as acetone and 1,1,1-trichloroethane, and semivolatile organic compounds (SVOC), such as naphthalene and phenol. Organic compounds that typically co-occur are often discussed together. For example, benzene, toluene, ethylbenzene, and xylenes are commonly referred to as BTEX compounds. These four chemicals are aromatic petroleum hydrocarbons typically found in contaminated groundwater near leaking underground storage tanks. It is possible that a compound can fall in more than one category, for example, naphthalene (Fig. 2) is a PAH but also a common petroleum hydrocarbon and a SVOC. Table 1 provides an overview of the most common classes of anthropogenic organic compounds in the terrestrial aquatic environment.

The physical and chemical properties of many organic compounds are well known and documented. Besides some very well maintained printed data collections (5–7), a number of reputable government and commercial web sites offer information about organic compounds (e.g., www.nedi.gov/NEDI-Catalog or esc.syrres.com/efdb/Chemfate.htm). In addition, there are digital databases, such as the Environmental Fate Database (EFDB) published by United States Environmental Protection Agency (EPA), which contain data or links to data on physical/chemical properties and the environmental fate of several thousand chemicals (16).

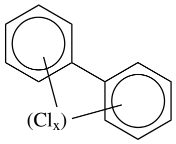

Figure 1. Polychlorinated biphenyls (PCB) are generally a complex mixture of various PCB isomers with different numbers of chlorine atoms (Cl$_x$) substituted onto the biphenyl rings.

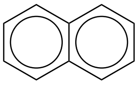

Figure 2. Naphthalene.

Many organic compounds are toxic, and even minute quantities of those chemicals make natural water unfit for human consumption. The EPA has defined threshold values for a number of organic and inorganic compounds. These so-called maximum contaminant levels (MCL) define when water is either of drinking water quality or contaminated. The MCLs and the number of entries changes with time as new scientific evidence emerges about the toxicity of listed and unlisted compounds.

New or improved analytical methods make it possible to detect compounds at lower detection limits. Better analytical resolution has also increased (and will continue to increase) the number of detectable compounds and is helping to identify possible links between the occurrence of negative health effects and the exposure to certain compounds. For example, there is evidence linking the exposure to some synthetic organic chemicals that compete with or disrupt the effects of hormones to reproductive abnormalities in wildlife, with cancer in women, and decreasing sperm counts in men. These so-called hormone or endocrine disrupters are chemicals such as some pesticides, PCBs, nonylphenols (used in surfactants), bisphenol A (a component in epoxy), phthalates (primarily used as plasticizers), and heavy metals.

Some chemicals that had been used on a large scale have been taken off the market after their detrimental biological effects and persistence in the aquatic environment became obvious. Probably the most well-known example is DDT (dichlorodiphenyltrichloroethane) (Fig. 3), which was the first chlorinated organic insecticide synthesized. DDT's effectiveness in fighting malaria and its low cost resulted in worldwide use of large quantities of that compound. In fact, DDT was credited with preventing 500 million human deaths due to malaria in little more than two decades (8). However, DDT in the environment has been linked to cancer in humans. Another organic compound that has received public attention more recently is MtBE (methyl tertiary-butyl ether) (Fig. 4). This compound has been added to gasoline and other petroleum products and has made its way into surface and groundwater. Currently, many states are trying or have already prohibited the use of MtBE as a gasoline additive after possible carcinogenic effects have been reported.

Once an organic compound has been released into the environment, bacterial activity may begin to break it down. This so-called biodegradation process works

Table 1. Important Classes of Anthropogenic Organic Compounds in Natural Waters

Class	Examples	Major uses or sources
Petroleum hydrocarbons	Benzene, toluene, ethylbenzene, xylenes, dodecane, hexane, naphthalene, pyrene	Gasoline, diesel and jet fuel, lubricating and hydraulic oils, grease, waxes
Alcohols, aldehydes, ketones, ethers	Methanol, ethanol, ethylene glycol, formaldehyde, acetone, 1,4-dioxane, tetrahydrofuran, MtBE	Solvents, chemical intermediates, antifreeze, oxygenizers (fuel additives)
Halogenated aliphatic hydrocarbons	Trichloroethene (TCE), tetrachloroethene (PCE), methylene chloride (dichloromethane), chloroform (trichloromethane), trichlorofluoromethane, methylbromide	Solvents, degreasing agents, dry cleaning fluids, refrigerants, pesticides
Polynuclear aromatic hydrocarbons (PAH)	Naphthalene, pyrene, anthracene, phenanthrene, benzo[a]pyrene, chrysene	Expect for naphthalene, there are no commercial uses. Sources: coal tar, crude oil and refined petroleum products, incomplete combustion of hydrocarbons
Polychlorinated biphenyls (PCB)	PCBs are synthetic mixtures of several congeners containing 1 to 10 chloride ions to the biphenyl. PCBs are oily liquids or waxy solids (4). Examples of commercial products containing PCB: Aroclor 1242, Clophen A30, Kenechlor 400.	Due to nonflammability, electrical insulating properties, chemical stability, and high boiling point, PCBs were used in electrical, heat transfer, and hydraulic equipment. Other uses include plasticizers in paints, plastics and rubber products; pigments, dyes, and carbonless copy paper
Substituted aromatics	Phenol, pentachlorophenol, nitrotoluene, 2,4,6-trinitrotoluene (TNT)	Pesticides, explosives, chemical intermediates
Phthalates	Di-*n*-butyl phthalate, dimethylphthlate, bis (2-ethylhexyl) phthalate. Phthalates are esters.	Plasticizers, solvents
Organosulfur and -phosphorus compounds	Malathion, Parathion, Aldicarb	Insecticides, herbicides, fungicides
Surfactants	Group of amphiphilic (partly polar, partly nonpolar) surface active compounds. Examples include alkylbenzenesulfonates (ABS), linear alkylbenzene sulfonates (LAS), fatty alcohol sulfates (FAS), quaternary ammonium chloride.	Soaps, laundry and dishwasher detergents, remediation agents

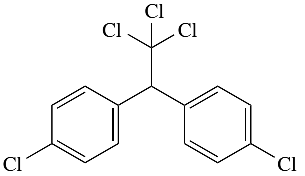

Figure 3. Dichlorodiphenyltrichloroethane (DDT).

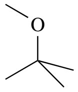

Figure 4. Methyl tertiary-butyl ether (MtBE).

best for those compounds that are naturally occurring or similar to naturally occurring compounds (crude oil and most refined petroleum products). Unfortunately, many synthetic organic chemicals are very resistant to biodegradation and can remain in the environment for a long time. In some cases, the breakdown of a chemical can result in a compound that is even more hazardous. For example, tetrachloroethene (PCE) undergoes a complex biodegradation process during which trichloroethene (TCE), dichloroethenes (DCA), chloroethane, and vinyl chloride are formed (6,7) (Fig. 5). Vinyl chloride is a known carcinogen and even as little as 2 μg/L in water is considered unsafe for human consumption. Some synthetic organic compounds have been used as molecular markers to identify waters influenced by human activity. For example, optical brighteners are synthetic organic compounds that are being added to most laundry detergents. Because these compounds pass through sewage plants or septic tanks systems almost unchanged, they can be detected in many surface and groundwaters (9,10). Other compounds, such as caffeine or the cholesterol breakdown product coprostanol, also have been used as organic tracers (11).

Water contamination from organic compounds is a global problem and is closely related to the recalcitrant nature of many organic chemicals. For example, though the solubility of many organic compounds is very low, (less than a few milligrams per liter), they exceed by far the drinking water limits in most cases.

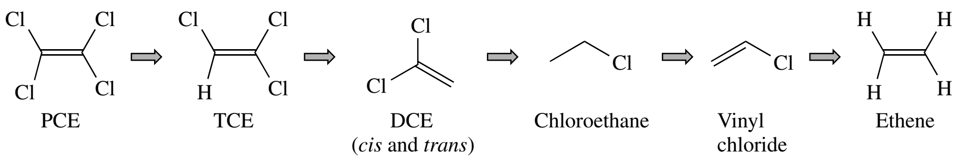

Figure 5. Simplified degradation pathway of tetrachloroethene (PCE), trichloroethene (TCE), *cis*- and *trans*-dichloroethene (DCE), chloroethane, and vinyl chloride to ethene (6).

Hence, even a relatively small spill of a toxic organic compound can contaminate a large volume of water. Many synthetic organic compounds also biodegrade very slowly, attach to organic matter or mineral surfaces, and—in the case of liquid immiscible compounds entering the subsurface—tend to have lower surface tension values than water. Thus, these compounds can enter small pore spaces or fractures from which they are hard to remove. Immiscible organic liquids are known to environmental hydrologists as nonaqueous phase liquids (NAPL). If a NAPL has a density less than that of water, it will float on the water table or the surface of a lake or a river. These compounds are called light nonaqueous phase liquids, (LNAPL). Examples of LNAPLs include gasoline, lubricating oil, and other petroleum products. Conversely, the densities of many NAPLs exceed that of water. These compounds are known as dense nonaqueous phase liquids (DNAPL). Many chlorinated solvents, such as PCE and TCE, are DNAPLs. After a NAPL, especially a DNAPL, has entered the subsurface, it is very difficult to locate and to remediate. Given the large quantities of accidentally released or illegally disposed NAPLs and other organic compounds, it will take decades or even centuries to solve the resulting environmental problems.

A number of analytical methods have been developed to analyze for organic compounds in water, but the methods approved by EPA are usually those that environmental laboratories use most (12). In many cases, the method selected depends on the desired analytical resolution and the instruments available.

Although our knowledge of organic compounds in groundwater has grown significantly during the past decades, there are still many questions to be answered. For instance, there has been a tendency to study the fate and transport of individual organic chemicals, although complex mixtures of organic compounds are much more common. Plenty of evidence now exists that interaction among multiple compounds—either synthetic or natural—results in chemical and physical behaviors distinctively different from those of simple, one-component systems. Further, physical and chemical processes in groundwater tend to be slow, and they very often exceed the length of a typical research project. Consequently, there is a slowly growing body of knowledge about the long-term (several decades) environmental fate of synthetic organic compounds in groundwater. Finally, there are a number of innovative methods under development for remediating contaminated groundwater. The expectation is that these methods will accelerate the removal of contaminants from the subsurface, therefore allowing the reuse of formerly polluted sites and, if possible, the restoration of groundwater to usable levels. Many of these methods, however, need to be tested in the field to explore their capabilities, limitations, and cost-effectiveness.

BIBLIOGRAPHY

1. Stevenson, F.J. (1985). Geochemistry of soil humic substances. In: *Humic Substances in Soil, Sediment, and Water: Geochemistry, Isolation, and Characterization*. G.R. Aiken, D.M. McKnight, R.L. Wershaw, and P. MacCarthy (Eds.). John Wiley & Sons, New York, pp. 13–52.
2. Thurman, E.M. (1985). Humic substances in groundwater. In: *Humic Substances in Soil, Sediment, and Water: Geochemistry, Isolation, and Characterization*. G.R. Aiken, D.M. McKnight, R.L. Wershaw, and P. MacCarthy (Eds.). John Wiley & Sons, New York, pp. 87–104.
3. Schwarzenbach, R.E., Gschwend, P.M., and Imoboden, D.M. (1993). *Environmental Organic Chemistry*. John Wiley & Sons, New York, p. 681.
4. United States Environmental Protection Agency. (1996). *PCBs: Cancer Dose-Response Assessment and Application to Environmental Mixtures*. EPA/600/P-96/001, September.
5. Mercer, J.W. and Cohen, R.M. (1990). A review of immiscible fluids in the subsurface. *J. Contam. Hydrol.* **17**(6): 106–163.
6. Verschueren, K. (1996). *Handbook of Environmental Data on Organic Chemicals*. 3rd Edn. John Wiley & Sons, New York, p. 2064.
7. Montgomery, J.H. (1996). *Groundwater Chemicals*. Desk Reference, 2nd Edn., Lewis Publishers, New York, p. 1345.
8. National Academy of Sciences (NAS). (1970). *The Life Sciences. Recent Progress and Application to Human Affairs, the World of Biological Research, Requirements for the Future*. National Academy of Sciences Press, Washington, DC, p. 526.
9. Stoll, J.M.A., Poiger, T.F., Lotter, A.F., Sturm, M., and Giger, W. (1997). Fluorescent whitening agents as molecular markers for domestic waste water in recent sediments of Greifensee, Switzerland. In: *Molecular Markers in Environmental Geochemistry*. R.P. Eganhouse (Ed.). ACS Symposium Series 671, American Chemical Society, Washington, DC, pp. 231–241.
10. Boving, T.B. (2002). *Fingerprinting sources of bacterial input using optical brighteners*. In preparation.
11. Sargeant, D. (1999). *Fecal Contamination Source Identification Methods in Surface Water*. Ecology Report #99–345, Dept. of Ecology, Washington State University, Pullman, WA.
12. United States Environmental Protection Agency. (1995). *Methods for the Determination of Organic Compounds in Drinking Water-Supplement III*. EPA/600/R-95-131.
13. Web-based databases: http://www.nedi.gov/NEDI-Catalog: National Environmental Data Index (NEDI) Catalog.
14. Stevenson, F.J. (1982). *Humus Chemistry; Genesis, Composition, Reactions*. John Wiley & Sons, New York, p. 443.
15. http://esc.syrres.com/efdb/Chemfate.htm: Environmental Fate Database maintained by the Syracuse Research Corporation (SRC), which is an independent, not-for-profit research and development firm chartered by the State of New York.
16. United States Environmental Protection Agency. (1995). *Environmental Fate Data Base (ENVIROFATE) Complete: XREF, CASLST, DATALOG, BIOLOG, CHEMFATE, and BIODEG (Raw Data File)*. Environmental Protection Agency, Office of Pollution, Prevention, and Toxics, Datafile-Magnetic tape, PB95-503074INC.

OVERDRAFT

JOHN E. MOORE
USGS (Retired)
Denver, Colorado

Groundwater mining (overdraft) occurs when there is a deliberate or inadvertent extraction of groundwater at a rate so in excess of replenishment that groundwater levels

Hence, even a relatively small spill of a toxic organic compound can contaminate a large volume of water. Many synthetic organic compounds also biodegrade very slowly, attach to organic matter or mineral surfaces, and—in the case of liquid immiscible compounds entering the subsurface—tend to have lower surface tension values than water. Thus, these compounds can enter small pore spaces or fractures from which they are hard to remove. Immiscible organic liquids are known to environmental hydrologists as nonaqueous phase liquids (NAPL). If a NAPL has a density less than that of water, it will float on the water table or the surface of a lake or a river. These compounds are called light nonaqueous phase liquids, (LNAPL). Examples of LNAPLs include gasoline, lubricating oil, and other petroleum products. Conversely, the densities of many NAPLs exceed that of water. These compounds are known as dense nonaqueous phase liquids (DNAPL). Many chlorinated solvents, such as PCE and TCE, are DNAPLs. After a NAPL, especially a DNAPL, has entered the subsurface, it is very difficult to locate and to remediate. Given the large quantities of accidentally released or illegally disposed NAPLs and other organic compounds, it will take decades or even centuries to solve the resulting environmental problems.

A number of analytical methods have been developed to analyze for organic compounds in water, but the methods approved by EPA are usually those that environmental laboratories use most (12). In many cases, the method selected depends on the desired analytical resolution and the instruments available.

Although our knowledge of organic compounds in groundwater has grown significantly during the past decades, there are still many questions to be answered. For instance, there has been a tendency to study the fate and transport of individual organic chemicals, although complex mixtures of organic compounds are much more common. Plenty of evidence now exists that interaction among multiple compounds—either synthetic or natural—results in chemical and physical behaviors distinctively different from those of simple, one-component systems. Further, physical and chemical processes in groundwater tend to be slow, and they very often exceed the length of a typical research project. Consequently, there is a slowly growing body of knowledge about the long-term (several decades) environmental fate of synthetic organic compounds in groundwater. Finally, there are a number of innovative methods under development for remediating contaminated groundwater. The expectation is that these methods will accelerate the removal of contaminants from the subsurface, therefore allowing the reuse of formerly polluted sites and, if possible, the restoration of groundwater to usable levels. Many of these methods, however, need to be tested in the field to explore their capabilities, limitations, and cost-effectiveness.

BIBLIOGRAPHY

1. Stevenson, F.J. (1985). Geochemistry of soil humic substances. In: *Humic Substances in Soil, Sediment, and Water: Geochemistry, Isolation, and Characterization*. G.R. Aiken, D.M. McKnight, R.L. Wershaw, and P. MacCarthy (Eds.). John Wiley & Sons, New York, pp. 13–52.
2. Thurman, E.M. (1985). Humic substances in groundwater. In: *Humic Substances in Soil, Sediment, and Water: Geochemistry, Isolation, and Characterization*. G.R. Aiken, D.M. McKnight, R.L. Wershaw, and P. MacCarthy (Eds.). John Wiley & Sons, New York, pp. 87–104.
3. Schwarzenbach, R.E., Gschwend, P.M., and Imoboden, D.M. (1993). *Environmental Organic Chemistry*. John Wiley & Sons, New York, p. 681.
4. United States Environmental Protection Agency. (1996). *PCBs: Cancer Dose-Response Assessment and Application to Environmental Mixtures*. EPA/600/P-96/001, September.
5. Mercer, J.W. and Cohen, R.M. (1990). A review of immiscible fluids in the subsurface. *J. Contam. Hydrol.* **17**(6): 106–163.
6. Verschueren, K. (1996). *Handbook of Environmental Data on Organic Chemicals*. 3rd Edn. John Wiley & Sons, New York, p. 2064.
7. Montgomery, J.H. (1996). *Groundwater Chemicals*. Desk Reference, 2nd Edn., Lewis Publishers, New York, p. 1345.
8. National Academy of Sciences (NAS). (1970). *The Life Sciences. Recent Progress and Application to Human Affairs, the World of Biological Research, Requirements for the Future*. National Academy of Sciences Press, Washington, DC, p. 526.
9. Stoll, J.M.A., Poiger, T.F., Lotter, A.F., Sturm, M., and Giger, W. (1997). Fluorescent whitening agents as molecular markers for domestic waste water in recent sediments of Greifensee, Switzerland. In: *Molecular Markers in Environmental Geochemistry*. R.P. Eganhouse (Ed.). ACS Symposium Series 671, American Chemical Society, Washington, DC, pp. 231–241.
10. Boving, T.B. (2002). *Fingerprinting sources of bacterial input using optical brighteners*. In preparation.
11. Sargeant, D. (1999). *Fecal Contamination Source Identification Methods in Surface Water*. Ecology Report #99–345, Dept. of Ecology, Washington State University, Pullman, WA.
12. United States Environmental Protection Agency. (1995). *Methods for the Determination of Organic Compounds in Drinking Water-Supplement III*. EPA/600/R-95-131.
13. Web-based databases: http://www.nedi.gov/NEDI-Catalog: National Environmental Data Index (NEDI) Catalog.
14. Stevenson, F.J. (1982). *Humus Chemistry; Genesis, Composition, Reactions*. John Wiley & Sons, New York, p. 443.
15. http://esc.syrres.com/efdb/Chemfate.htm: Environmental Fate Database maintained by the Syracuse Research Corporation (SRC), which is an independent, not-for-profit research and development firm chartered by the State of New York.
16. United States Environmental Protection Agency. (1995). *Environmental Fate Data Base (ENVIROFATE) Complete: XREF, CASLST, DATALOG, BIOLOG, CHEMFATE, and BIODEG (Raw Data File)*. Environmental Protection Agency, Office of Pollution, Prevention, and Toxics, Datafile-Magnetic tape, PB95-503074INC.

OVERDRAFT

John E. Moore
USGS (Retired)
Denver, Colorado

Groundwater mining (overdraft) occurs when there is a deliberate or inadvertent extraction of groundwater at a rate so in excess of replenishment that groundwater levels decline persistently and threaten exhaustion of the supply or at least a decline of pumping levels to uneconomic depths (1).

Under natural conditions groundwater moves from areas of recharge to areas of discharge. The water is discharged to springs, lakes, streams, wetlands, and to the ocean. In most cases, equilibrium exists, and long-term recharge of groundwater is balanced by long-term discharge. Water levels in an aquifer fluctuate in response to changes in the rate of recharge and discharge. When discharge exceeds recharge, water is released from storage, and water levels decline. Declines may amount to hundreds of feet in areas where large withdrawals from wells have caused discharge to exceed recharge for long periods of time.

Groundwater in the United States is being increasingly developed for irrigation and municipal supplies. The U.S. Geological Survey has noted that groundwater availability is a significant issue in almost every state (2). The development of groundwater has led to declining water levels in a number of areas. The water table and artesian head declines have resulted in the following:

1. stream flow depletion
2. land subsidence
3. saltwater intrusion
4. increased cost to deepen wells
5. higher costs to pump water
6. decrease in well yield
7. drying up of shallow wells and springs
8. degradation of water quality

Under natural conditions, the largest fluctuations in water levels in unconfined aquifers are seasonal. Short-term fluctuations in confined aquifers commonly occur due to such factors as changes in barometric pressure. In areas where large withdrawals from wells have caused discharge to exceed recharge for long periods of time, declines may amount to tens or even hundreds of feet. These yearly declines may stop if pumpage is reduced so that discharge is equal to or less than discharge.

In a number of areas of the United States, groundwater levels have declined 300 feet or more. Five case studies were selected where groundwater mining has had a major impact: Denver, Tucson, Chicago, California, and Las Vegas. In these areas, groundwater mining has caused land subsidence, stream flow depletion, and increased costs for pumping groundwater (3).

Denver, Colorado

The Denver Basin bedrock aquifer system supplies water to rural and suburban residents of a large area along the eastern front of the Rocky Mountains in northeastern Colorado. The geologic formations that compose the aquifer system are Cretaceous and Tertiary sandstone and conglomerate shale of the Fox Hills Sandstone, Laramie Formation, Arapahoe Formation, Denver Formation, and Dawson Arkose (Fig. 1). In 1985, 12,000 wells withdrew water from these aquifers. A large part of the rapidly growing Denver metropolitan area depends on groundwater from these aquifers. For example, about 70% of the 220,000 residents in the southern metro area obtain water from bedrock aquifers. The bedrock aquifers in the Denver Basin have supplied groundwater to the Denver area for more than 100 years. Groundwater development began in the 1880s with artesian wells in downtown Denver. The artesian pressure was used for decorative fountains, powering elevators, and church organs. The water has been used for domestic, municipal, and industrial purposes. Groundwater withdrawal from the bedrock aquifers has caused water level declines, reduced well yield, and increased pumping costs. Water levels have declined as much as 12 m (40 feet) per year in some areas. In most of the area, the artesian head will be eliminated within the next 10 years. Increased use of groundwater and water level declines has alarmed water districts, government agencies, and water users because of the increased cost and depletion of the supply.

Tucson and Phoenix, Arizona

Large water level declines have caused land subsidence and earth fissures to develop in a 3000 square mile area that includes Tucson and Phoenix. The subsidence is caused by compaction of fine-grained sediments in the basin fill. The sediments deform and compact when water-level declines subject the sediments to additional compression from the weight of the overlying deposits.

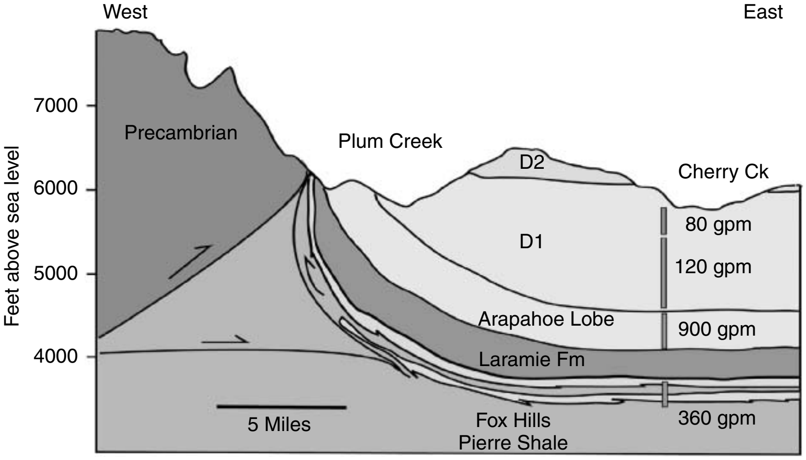

Figure 1. Denver Basin bedrock aquifers (4).

Figure 2. Pima County, Arizona. Sign warning motorists of subsidence hazard was erected after an earth fissure damaged a road in Pima County, Arizona (left photograph). Earth fissure near Picacho, Arizona (right photograph) (photographs by S.R. Anderson, U.S. Geological Survey).

Compaction increases slowly as water levels decline. Land subsidence has exceeded 15 feet since the start of groundwater development. Water levels in the Tucson area have declined 160 feet from 1945 to 1984. Compaction and land subsidence cause cracks (Fig. 2) in the land surface which can extend for hundreds to thousands of feet along the surface and can be hundreds of feet deep (5).

Central Valley, California

The Central Valley of California is one of the most intensively developed areas of irrigated agriculture. The valley is 500 miles long by 50 miles wide, and its total area is about 16,000 square miles. The valley is underlain by a large alluvial aquifer to a depth of 25,000 feet. Fresh groundwater is present to depths of as much as 4000 feet but most well water is less than 1000 feet deep. Significant development of groundwater for irrigation began in the early 1900s. As groundwater withdrawal increased to the point that groundwater withdrawal exceeded recharge, water levels started to decline. The water level declined 260 feet between 1940 and 1963. The groundwater withdrawal caused the compaction of fine-grained sediments (confining beds) which in turn resulted in subsidence of the land surface. Water level declines of 160 feet or more have resulted in land subsidence of about 30 feet in the Los Banos–Kettleman City area (Fig. 3).

Chicago, Illinois

Since 1864, large quantities of groundwater have been withdrawn from glacial drift and bedrock aquifers for municipal and industrial use. In the Chicago area, two Cambrian–Ordovician bedrock aquifers (sandstone and dolomite) supply groundwater. Water use has increased from 9.2–340 million gallons per day from 1864 to 1980. One of the first wells drilled to a depth of 711 flowed at the land surface at a rate of 400 gallons per minute. Groundwater withdrawal from the Cambrian–Ordovician bedrock aquifers (sandstone and dolomite) has caused water level (artesian head) declines of more than 800 feet (Fig. 4). No major land subsidence has been reported as a result of these large withdrawals because the rocks in the area are consolidated and resist compaction. A principal concern has been the possibility of beginning to

Figure 3. Central Valley, California: Land subsidence of 30 feet.

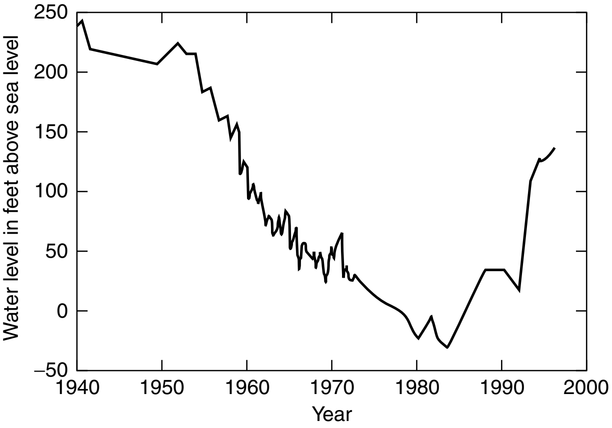

Figure 4. Chicago, hydrograph 1940–1998. (Representative trend of water levels for a deep well in Cook County, Chicago area, since 1940).

dewater the confined aquifer and thus converting it to an unconfined aquifer.

Las Vegas, Nevada

The Las Vegas Valley is the fastest growing metropolitan area in the United States. Water to support this rapid growth is being supplied by imported Colorado River water and local groundwater. Groundwater is pumped from the upper 2000 feet of unconsolidated alluvial sediments. The deeper aquifers called “principal artesian”(below 300 feet) can yield large quantities of groundwater. Overlying the principal aquifers are 100 to 300 foot thick deposits of clay. Since the 1970s, the annual groundwater withdrawal has remained between 60,000 and 90,000 acre feet. Most of the withdrawal is from the northwest area, where declines are more than 300 feet. Areas in central Las Vegas (The Strip) have experienced declines of as much as 200 feet. Since 1935, compaction of the aquifer has caused nearly 6 feet of subsidence and has led to the formation of numerous earth fissures and surface faults. Before groundwater development, the aquifer sustained the flow of many springs that discharged into Las Vegas Wash. Extensive groundwater withdrawal has caused the springs to dry up, and urban runoff has created a reservoir of poorer quality potentially contaminated water that now recharges the principal aquifer (Fig. 5).

Table 1. Subsidence and Groundwater Withdrawal

	Subsidence,		
Locality	Feet	m	Area Affected, km^2
Central Valley, CA	30	9.0	13,500
Houston-Galveston, TX	9.1	2.75	12,170
Eloy, AZ	11.8	3.6	8,700
Tokyo	15.1	4.6	2,400
Po Valley, Italy	9.8	3.0	780
London	1.1	0.35	450
Venice, Italy	0.45	0.14	400
Mexico City	28.5	8.70	225

CONCLUSION

Water level declines and subsidence have been documented in Houston, Beijing, Tokyo, Bangkok, Venice, and Mexico City (Table 1).

Water conservation and regulations should be implemented in areas where groundwater is being mined (6). Some other measures that could be used to sustain the water supply are

- artificial recharge
- water reuse
- conjunctive use of groundwater and surface water
- restrictions on lawn irrigation and other water use.

BIBLIOGRAPHY

1. Wilson, W.E. and Moore, J.E. (1998). *Glossary of Hydrology*. American Geological Institute, Arlington, VA.
2. U.S. Geological Survey. (1984). *National Water Summary 1984*. U.S. Geological Survey Water Supply Paper 2275.
3. Moore, J.E. (2002). Effects of groundwater mining on aquifers-case histories Denver, Tucson, Chicago, Central Valley, Las Vegas: *IAH Congr. Mar de Plata*.
4. Johnson, K. and Raynolds, R. (2001). *Ancient Denvers*. Denver Museum at Nature and Science, p. 34.
5. U.S. Geological Survey. (1995). Colorado, New Mexico, Nevada, and Arizona. Hydrologic Atlas 730-C.
6. Alley, W.M., Riley, T.E., and Franke, O.L. (1999). *Sustainability of Ground-Water Resources*. U.S. Geological Survey Water Supply Paper 1186.

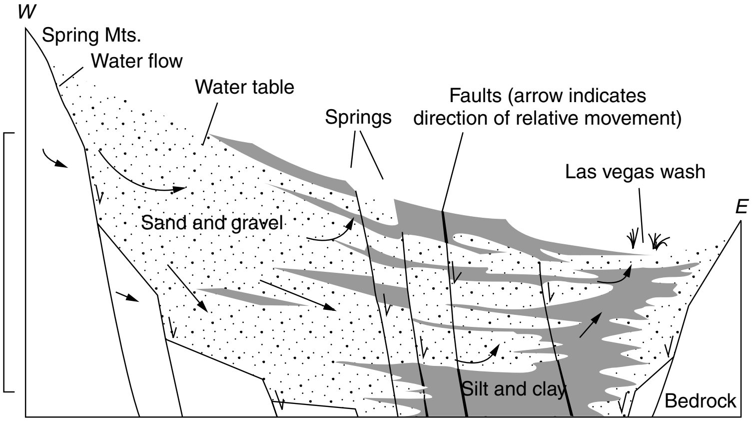

Figure 5. Las Vegas hydrogeology.

CHEMICAL OXIDATION TECHNOLOGIES FOR GROUNDWATER REMEDIATION

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

INTRODUCTION

The use of chemical oxidation for the *in situ* remediation of soils and groundwater impacted with constituents of concern (COC) is a technology that has seen significant development and application in the past decade with numerous site applications in the last eight years. Three chemicals are typically utilized as oxidants:

- Hydrogen peroxide in the form of Fenton's reagent
- Potassium permanganate or sodium permanganate
- Ozone

Each has advantages and disadvantages, as does the use of chemical oxidation technology in general. There are numerous recent reviews of field application of the technology (1,2). Our purpose here is to assess the utilization of *in situ* chemical oxidation in general and the use of these three chemical oxidants in particular. In addition, these are powerful chemicals and care should be taken with their shipping, storage, and use.

Chemical oxidation has had a long history of application in the field of wastewater treatment. Hydrogen peroxide, Fenton's reagent, potassium permanganate, ozone, and combinations of ozone with other oxidants are all chemical oxidation systems that have found application for the treatment of wastewater. One of the key design issues for the application of chemical oxidation to wastewater is the total carbon load or other inorganic constituent that can be oxidized over and beyond that represented by the COC that are targeted in that waste stream. The chemical oxidants will react with and be consumed by all constituents in the wastewater stream not just the COC. In the case of the *in situ* treatment of groundwater, there are also consuming reactions associated with the geologic mineral matrix. In many instances, this adversely impacts the cost of application to the point of being impractical.

The general value of *in situ* chemical oxidation technologies resides in two areas: first, the treatment of residual free product and second, reduction of overall remediation time frames. Specifically, *in situ* oxidation is likely to be selected for difficult applications that include:

- Low permeability soils
- Highly stratified soils
- Low solubility compounds
- High concentrations of highly soluble organics (such as ketones, alcohols, or MTBE) that would be difficult to treat with conventional surface treatment technology (air stripping or activated carbon)
- Gasoline, diesel fuel, chlorinated solvents, polychlorinated biphenyls, organic pesticides, TNT, PBX, and VX
- Target compounds with low *in situ* degradation kinetic constants
- Dense, nonaqueous phase liquids (DNAPLs).

From an economic perspective, chemical oxidants are usually only practical in limited areas, typically near or in source zones. It is important to differentiate types of source zones:

- Primary source zones refer to areas that have been exposed to free phase DNAPLs. Discrete DNAPLs pools, emulsified DNAPLs, and high levels of adsorbed COC are characteristic of such zones.
- Secondary source zones refer to areas that have been exposed to high concentrations of dissolved COC. Elevated levels of adsorbed and dissolved COC are characteristics of these zones.

Chemical oxidation systems have the potential to offer rapid (from a week to a month) removal of COC. In addition, oxidation systems can be applied to dissolved plumes that are above the levels for which natural attenuation mechanisms can be expected to work within the time frames desired for site closure. Treatment of large diffuse dissolved COC plumes is possible with chemical oxidants but often not economically practical due to the effect of side reactions that consume the oxidant reagents.

The most powerful advanced oxidation systems are based on the generation of hydroxyl radicals. The hydroxyl radical is an extremely powerful oxidation agent, second only to fluorine in power. Fenton's reagent, ozone, and sonification are different means of generating those hydroxyl radicals. Each has advantages and disadvantages that can be exploited given the specific setting of a COC impacted site. A less intense, but still applicable, oxidation method involves the use of potassium or sodium permanganate. Following is a listing of common chemical oxidants, placed in the order of their oxidizing strength:

Relative Oxidation Power ($Cl_2 = 1.0$)

Fluorine	2.23
Hydroxyl radical (OH)	2.06
Activated persulfate	1.91
Atomic oxygen (singlet)	1.78
Ozone	1.50
Persulfate	1.48
Hydrogen peroxide	1.31
Perhydroxyl radical (OOH)	1.25
Potassium permanganate	1.24
Chlorine dioxide	1.15
Chlorine	1.00
Bromine	0.80
Iodine	0.54

In addition, the geology must be considered when applying a chemical oxidant *in situ*. Although the first concern is having the liquid reagents reach all of the COC in a heterogeneous situation, the soils themselves can have a significant effect on the reactions. This technology relies on water as a carrier to deliver the chemical oxidant to the COC. The macro- and microgeologic conditions do not always allow for complete contact between the chemical oxidants and the COC present in the vadose or saturated zone.

It is important to understand that COC oxidation occurs only in the aqueous phase, involving dissolved species of both the COC and the oxidant. The solubility of a targeted COC ultimately controls the rate of possible oxidation. There is interface mass transfer and then chemical oxidation. The rate of mass removal is limited by the kinetics of the COC dissolution process rather than those of the chemical oxidation reaction.

In all cases where the final product of the oxidation product is carbon dioxide, there is potential for short-term plugging of the aquifer pore space with carbon dioxide gas bubbles. In the case of Fenton's reagent, oxygen gas will also be generated and contribute to this possible problem.

Lastly, with the exception of some permanganate applications, this technology relies on the total displacement of native groundwater from the treatment zone and replacement with the reagent laden water. In the case of permanganate, the active life is sufficient that at sites with moderate to high groundwater velocities (greater than 30 ft/yr) reagent injection can take place with subsequent reliance on advection and dispersion to mix the oxidation solutions with native groundwater during transport.

CHEMICAL OXIDATION WITH FENTON'S REAGENT

Hydrogen peroxide alone has little effectiveness for the oxidation of chlorinated solvents. Fenton's reaction was first described in 1894. During the 1930s the reaction mechanisms were fully defined. Over the last 15 years commercial reactors have been available for wastewater treatment, and in the 1990s applications for *in situ* groundwater treatment have been developed. Fenton's reagent generates hydroxyl radicals through the reaction of ferrous iron and hydrogen peroxide:

$$Fe^{2+} + H_2O_2 \rightarrow {}^{\bullet}OH + Fe^{3+} + OH^-$$

The process is self-replicating because the reaction of ferric iron with hydrogen peroxide to generate the perhydroxyl radical also occurs:

$$Fe^{3+} + H_2O_2 \rightarrow Fe^{2+} + {}^{\bullet}OOH + H^+$$

The perhydroxyl radical is a weaker oxidizer (between hydrogen peroxide and permanganate). But, more importantly, the process generates ferrous ions that in turn stimulate further reaction with hydrogen peroxide to produce more hydroxyl radicals. The hydroxyl radical can react with almost any hydrocarbon to produce carbon dioxide as a final product (and chlorides if a chlorinated hydrocarbon is treated). The key application issues are delivery of the reagents *in situ* and pH control. Optimum pH for the reaction, in the range of 3.0–4.5, is driven by the iron chemistry. It is possible to use iron chelates to get the desired reaction at higher pH values (3). Improper control of the reaction will only generate oxygen and water, not high-intensity oxidation. The technology is ideal for application in source zone areas where there are high concentrations of adsorbed or even interstitial free product. Cost is mostly dependent on the geochemistry of the geologic matrix and control of the iron chemistry. Soil buffering capacity may limit the ability to obtain optimal subsurface pH conditions.

Most of the original utilization of Fenton's reagent was in relatively low concentrations in wastewater applications. The main criteria for wastewater application was to have a sufficient ratio of hydrogen peroxide to the organic chemical that was to be destroyed. The chemical oxidation process is less efficient in the presence of soil. Due to the presence of competing organics and mineral surfaces (4) that are reactive to hydrogen peroxide, and the less than optimal environmental conditions, dosage requirements for *in situ* applications may increase 10–100-fold or more to accomplish the desired oxidation of COC. With regard to the rate of reaction and the degree of completion of the reaction, high concentrations of hydrogen peroxide (10–25%) are preferred. On the other hand, excess consumption of hydrogen peroxide by reactions with the geologic matrix can be controlled by using applications (possibly multiple) at concentrations of 5% or less.

Almost all soils have native iron bearing minerals as part of their geochemical makeup. When soils contain chemically available iron, supplemental iron salts may not be required with the hydrogen peroxide injection solution (5). However, problems may occur when too many iron minerals are present in the soil. If the iron is in a mineralized form not readily available for dissolution under low pH conditions, the natural catalytic activity of the mineralized iron decomposes the hydrogen peroxide to oxygen and water and does not create the hydroxyl radical. A second problem with soil mineralogy comes from soils rich in carbonate minerals or with high alkalinity. These soils produce groundwater with a high pH, which can require unacceptable amounts of acid to buffer the pH down to the required range. Carbonate minerals respond to low pH so rapidly that the geologic matrix itself will consume acid. Lastly, the carbonate ion preferentially scavenges hydroxyl radicals before they have a chance to react with the COC.

The heat generated by exothermic dissociation of hydrogen peroxide promotes contaminant volatilization. Rapid gas generation creates turbulence that further enhances the contact of the oxidant with the targeted COC impacted zones. Concentrations of hydrogen peroxide as low as 11% can cause groundwater to boil. The rate of hydrogen peroxide decomposition doubles with every 10 °C rise in temperature, so the energy release process, once initiated, rapidly accelerates. A pound of hydrogen peroxide can release 1200 BTU of heat energy and up to 6 ft^3 of oxygen gas, which creates volatilization and transport due to the created pressure gradients, which, when exploited in a controlled fashion, can have beneficial effects regarding contact of the oxidant with the COC entrained in the geologic matrix.

For well over a decade, the cost of 50% hydrogen peroxide has been \$3.50 a gallon. Hydrogen peroxide applications are frequently used with sulfuric acid (H_2SO_4) or acetic acid (CH_3CH_2OH) as a pretreatment to optimize the subsurface pH and solubilize the naturally occurring iron as a catalyst for Fenton's reagent. Sometimes ferrous sulfate is added to the hydrogen peroxide as a catalyst if naturally occurring iron is not available. Recently, sodium persulfate ($Na_2S_8O_2$) is being used with hydrogen peroxide for *in situ* oxidation processes. In this case, the heat of reaction from the hydrogen peroxide activates the much longer-lasting sulfate ($S_2O_8{}^{2-}$) radical.

PERMANGANATE

With the permanganate ion, the initial oxidation reactions of COC are independent of pH. For example, the oxidation rate of chlorinated alkenes by potassium permanganate is not affected by pH over a range of 3.0–11.0. However, pH does impact the type of intermediate products produced: At pH 4, formic acid is dominant; at pH 6–8, oxalic and glyoxylic acids dominate. In both cases, the organic acids are subsequently oxidized to carbon dioxide. The conversion of the intermediate oxidation products to carbon dioxide does occur more rapidly under acidic pH conditions. For the initial permanganate oxidation reactions, they can generally be characterized (7) as follows:

- Acid catalyzed reactions at $\text{pH} < 5.0$
- Uncatalyzed reactions at pH 5–9
- Base catalyzed reactions at $\text{pH} > 10$

Potassium permanganate has an affinity for organic compounds containing carbon–carbon double bonds, aldehyde groups, and hydroxyl groups. The permanganate ion borrows electron density from the pi bonds in chlorinated alkenes, which creates a bridged oxygen compound known as the cyclic hypomagnate ester (6). The intermediate ester is unstable and further reacts by a number of mechanisms including hydroxylation, hydrolysis, or cleavage. Under normal pH and temperature conditions, the primary oxidation reactions involve spontaneous cleavage of the carbon–carbon bond. Once the double bond is broken, the highly unstable carbonyl groups are immediately converted to carbon dioxide through either hydrolysis or further oxidation by the permanganate ion.

Manganese dioxide (MnO_2) is the end product of the reduction of the permanganate oxyanion under neutral to basic pH conditions by the following reaction:

$$MnO_4{}^- + 2\,H_2O + 3\,e^- \rightarrow MnO_2 + 4\,OH^-$$

The MnO_2 is insoluble and forms colloids with a typical diameter near 1 μm (8). Plugging of the effective pore space can occur due to the precipitation of MnO_2. The manganese dioxide colloids can also cause problems with surface treatment equipment used as part of oxidation programs that rely on the creation of active circulation cells. The formation of insoluble manganese dioxide reaction products can also form a coating and seal around DNAPL mass, in effect blocking further dissolution and migration of the COC into the adjacent reaction areas (9), which can be a significant problem.

At low pH conditions, the reduction of the permanganate ion proceeds to Mn(II) by the following reaction:

$$MnO_4{}^- + 8\,H^+ + 5e^- \rightarrow Mn^{2+} + 4\,H_2O$$

The Mn(II) cation is soluble in water at concentrations above the regulatory limit (50 ppb) when chloride or sulfate counteranions are present.

Examples of complete reactions of permanganate with CVOCs are as follow:

$$\text{PCE}\ C_2Cl_4 + 2\,MnO_4{}^- \rightarrow 4\,Cl^- + 2\,CO_2 + 2\,MnO_{2(s)}$$

$$\text{TCE}\ C_2HCl_3 + 2\,MnO_4{}^- \rightarrow 3\,Cl^- + 2\,CO_2 + H^+ + 2\,MnO_{2(s)}$$

$$\text{DCE}\ C_2H_2Cl_2 + 2\,MnO_4{}^- \rightarrow 2\,Cl^- + 2\,CO_2 + 2\,H^+ + 2\,MnO_{2(s)}$$

The stochiometry ratios for complete oxidation reactions of permanganate as pounds of the $KMnO_4$ salt required to oxidize 1 pound of CVOC are:

PCE	1.3: 1
TCE	2.4: 1
DCE	4.4: 1
VC	8.5: 1

(The dramatic change in consumption rate is due to the fact that this analysis is on a pound per pound basis, not a molar basis. A pound of vinyl chloride has almost seven times more moles than a pound of PCE due to the loss of three heavy chlorines from the ethene backbone.)

In contrast to Fenton's reagent, the presence of carbonate minerals in the geologic matrix has generally positive effects on permanganate oxidation (10): Mn reduction stops at Mn(IV) (insoluble) rather than proceeding to potentially soluble Mn(II). Trace minerals are coprecipitated and immobilized with the manganese oxide. Consumption of permanganate by reaction with reduced mineral such a magnetite is minimized because the rates of these reactions increase with decreasing pH.

Oxidant demand from the matrix can be attributed to natural organic matter, reduced metals, carbonates, and sulfides. Permanganate demand rates can vary from a few grams of permanganate per kilogram of soil (clean sand with dissolved COC) to as much as several hundred grams of permanganate per kilogram of soil (organic clays with 6% organic carbon content and DNAPLs).

Immobile reduced metals such as chromium, uranium, vanadium, selenium, and molybdenum (all of which are soluble as oxidized oxyanionic complexes) can be oxidized and mobilized by permanganate oxidation.

When the permanganate concentration is 10 times the concentration of the CVOCs, the half-lives of TCE and

DCE range from 24 s to 18 mins; the half-life of PCE is 257 min. Typical dosage range (by weight) is 1.5–5% permanganate in the targeted treatment zone.

The potassium salt has a solubility limit of 65 g/L (6.5%); the sodium salt is significantly more soluble. However, the sodium salt is significantly more reactive. The reactivity difference between the sodium and potassium salt is not significant with regard to COC treatment. It is significant because the increased reactivity of the sodium salt presents greater safety challenges during its shipping, storage, and application.

Potassium permanganate costs $4.75 per pound.

OZONE

Ozone has been used for the treatment of drinking water since 1903. Its oxidizing power and innocuous decay products have made it ideal for that application. Ozone can generate hydroxyl radicals via catalytic decomposition of water. The oxidative power of an ozone generated hydroxyl radical is the same as those generated by Fenton's reagent. It also has specific activity toward alkenes (including chlorinated alkenes) through the attack of the double-bond carbon in a fashion similar to permanganate oxidation. Ozone bridges the carbon–carbon double bond to form unstable ozonide intermediates, which decompose into smaller oxidation species until carbon dioxide and water, or stable refractory compounds, such as acetic acid or oxalic acid, are formed. Ozone does not react well with chlorinated alkanes, particularly chlorinated methanes. Although the fate and reaction mechanisms of ozone in porous geologic media are not completely understood, it is likely that hydroxyl radical production from ozone occurs through catalytic reactions with natural inorganic (probably iron oxides) and organic material.

Ozone is fairly stable in dry air and has a half-life of several hours in low concentration. In water, ozone half-life is several minutes. As ozone is very reactive in an aqueous environment, ozone can oxidize material between 10 and 1000 times faster (11) than most oxidants used in water treatment. As ozone has such a short half-life, it cannot be compressed and stored. Instead, it must be generated on site and used immediately. Electrical generation is the only practical and safe method for large-scale applications of ozone. In practice, ozone concentrations of 1–2% using air and 3–8% using oxygen can be obtained by corona discharge generators. At a production concentration of 1% by weight in air, 650 ft^3 of air is required to produce 1 lb of ozone. As the speed of ozone destruction by decomposition is proportional to the ozone concentration, producing higher ozone concentrations with corona discharge is not feasible.

The decomposition rate of ozone has been shown to be 25 times more rapid in geologic media than in nonreactive media such as glass beads. However, the presence of carbonate ions in the groundwater can provide stabilization and double the half-life of dissolved ozone at a pH of 8.0 (12).

Other species created from ozone include superoxide anion $O_2^{\bullet -}$, which is an intermediate to the hydroxyl radical $^{\bullet}OH$; peroxide anion O_2^{2-}; oxygen atom singlet O; and the oxygen radical anion $O^{\bullet -}$.

At atmospheric pressure, the solubility of ozone is 3.4 mg/L for 1% and 7 mg/L for 2%. Note that solubility is somewhat dependent on the concentration of the ozone in the gas phase. With decreasing temperature, solubility increases. In addition, increasing hydrostatic pressure increases ozone solubility. For example, with a 1% system at 18 psi ozone solubility is 4.2 mg/L; at 25 psi it is 6.3 mg/L; and at 32 psi it is 8 mg/L.

As the TDS of the treated water increases, the solubility decreases. For example, ozone solubility is reduced 30% in seawater (13).

In most applications 1.5–3 lb of ozone is required for each pound of COC to be treated. That consumption rate is based purely on COC degradation requirements; spontaneous decay and ancillary reactions with other groundwater or mineral matrix constituents create an ozone loading requirement over and above those baseline numbers.

The cost for the generation of ozone includes the cost of the equipment: 1 lb/d unit— $7500; 10 lb/d unit— $32,500. Forty percent of the cost is for air pretreatment. The power requirements are in the range of 8–12.6 kW·h per pound of ozone produced.

PRACTICALITY OF APPLICATION

Stoichiometric Cost Comparison

The weight ratio in terms of pounds of oxidant per pound of TCE fully oxidized are 0.8 lb/lb for Fenton's reagent and 2.4 lb/lb for potassium permanganate. The cost for Fenton's regent, including hydrogen peroxide and amendments for pH control and ferrous iron addition, is typically $1.10 per pound and potassium permanganate is $4.75 per pound. So total oxidant cost per pound of TCE destroyed would be $0.88 for Fenton's reagent and $11.40 for potassium permanganate.

Calculation of a similar cost for ozone oxidation is more complicated, involving the capital costs of the equipment, the operational life of the system, electrical efficiency, and electrical costs. Following are assumptions for those key parameters.

A 10 lb/d ozone unit is purchased for $32,000 for a project life of 3 yr, which translates to a capital cost of $2.90 per pound of ozone. It requires 10 kW·h at $ 0.10 per kW·h (or $1.00) for the electrical cost to produce 1 lb of ozone. The total per pound cost for ozone is $3.90 per pound.

It requires 2.25 lb of ozone to oxidize 1 lb of COC, translating into a total cost of $8.75 for the ozone to oxidize 1 lb of COC.

Actual Application Costs

The total costs for the use of these oxidants is not accurately reflected by just the pound per pound cost required for the stochiometric oxidation of the COC. In reality, reactions with native carbon in the geologic matrix, reactions with minerals in the geologic matrix (14), and reactions with native dissolved constituents in the groundwater can increase the consumption rate of the oxidation reagents by orders of magnitude. Fenton's reagent and ozone are particularly susceptible to these effects, potassium permanganate less so.

In practical application, these chemical oxidants are not applied based on the amount of COC present (with the exception of instances when there are significant volumes of DNAPLs present); rather, they are applied to achieve a specific reagent concentration in the treatment zone. Assuming the desired reagent concentration is 5% for Fenton's reagent and permanganate (this argument is not valid for ozone as explained below), following are the per cubic yard costs for one treatment with the respective chemical oxidant:

Fenton's reagent—$21
Potassium permanganate—$115

In practice, it is likely that two to five applications of Fenton's reagent may be required to achieve the remediation goals during a treatment campaign, raising the reagent costs to the range of $42–105. In many instances, one application of potassium permanganate will be adequate for treatment, but in cases where DNAPLs are present two or three applications may be required.

The hydrogen peroxide in Fenton's reagent can have a reaction life (under practical conditions) that ranges from 15 min to several hours. The permanganate ion can remain active in the saturated zone for months or perhaps even a year or more (10). Ozone will decompose in water within minutes, which has significant impact for *in situ* applications: Fenton's reagent can be injected directly into an aquifer, but spacing of the injection points will typically range from 1 to 7 yards. All contact with the COC occurs during advection caused by the injection process, and by mixing caused by heating and gas generation. Groundwater in the treated areas is displaced.

Permanganate salts can be injected *in situ* with the anticipation of further mixing of the reagents with the COC impacted groundwater due to natural flow conditions. Spacing of injection wells can be predicated on the impact of natural groundwater flow over a 6-month to 1-year period. Push–pull circulation systems can also be utilized to decrease the spacing of injection wells, but accommodations must be made for the removal of colloidal manganese dioxide from recovered groundwater. The reaction rate of permanganate also allows it to be applied in conditions where native groundwater flow can provide further advection for transport and dispersion for mixing. Complete displacement of native groundwater is not required.

Ozone is so reactive in water that it cannot be practically injected *in situ*. It can be used as a reactant in the bore of a circulation well or as part of a sparge system. The radius of influence of such systems is limited. As a result, the actual application of ozone as an *in situ* chemical oxidation system has been much more limited than the use of Fenton's reagent or potassium permanganate.

CONCLUSION AND CAUTIONS

A last word of caution. These are powerful chemicals. During shipment, storage, and application of chemical oxidants great care must be taken to maintain safe site conditions with regard to personnel and property.

Health and Safety Precautions for Chemical Oxidants

Hydrogen Peroxide. High strength hydrogen peroxide (greater than 5%) can cause chemical burns to the skin and eyes; lower concentrations will be irritating. In all cases, handling requires protective clothing: face shield, gloves, hard hat, rubber or PVC boots, and a rubber or PVC suit. A shower and eye wash should be available.

Decomposing hydrogen peroxide rapidly generates heat, gas, and pressure. Storage vessels and piping runs that have valve traps should have adequate ventilation and pressure relief systems.

Containers and piping should be free of all contaminants. Contaminants include copper, brass, zinc, mild steel, synthetic rubbers, polypropylene, and organic compounds (especially liquid organics). Acceptable materials for storage include aluminum, stainless steel, glass, ceramics, Teflon, polyethylene, Viton, and (for temporary storage) PVC. All metal components must be properly passivated before use.

Check valves are required to prevent backflow into hydrogen peroxide piping or tanks. Hydrogen peroxide pumps should only be constructed of stainless steel or Teflon.

Hydrogen peroxide can generate enough heat and oxygen to ignite combustible materials. Do not store hydrogen peroxide on wooden pallets or decks. Hydrogen peroxide in the presence of hydrocarbon vapors can cause vapor phase explosions. Every 10 °C increase in temperature doubles the reaction rate of hydrogen peroxide, it must be stored away from heat sources (boilers, steam lines, etc.) and should never be stored in insulated tanks.

Shipment of solutions with a concentration greater than 8% must be done as DOT hazardous.

Hydrogen peroxide is sold as a 50% solution, which is an extremely reactive material. Hydrogen peroxide vendors can readily deliver more dilute solutions. When possible, it is preferable to have the hydrogen peroxide delivered at the concentration at which it will be used for *in situ* chemical oxidization, rather than delivered at 50% and then diluted on site.

Permanganate Salts

Permanganates present a fire and explosion risk when placed in contact with organic materials. Containers of permanganate salts exposed to liquid hydrocarbons may explode. Other materials that exhibit high reactivity include metallic powders, elemental sulfur, hydrochloric acid, hydrazine, hydrogen peroxide, and metal hydrides.

Permanganate salts are irritating to the eyes, skin, respiratory system, and digestive tract from acute exposure. Chronic exposure to the skin can cause defatting and dermatitis. Chronic ingestion can cause central nervous system and kidney damage.

Employee protective equipment should include dust and mist respirator (and or area ventilation), gloves, and safety goggles.

Shipping is regulated under the Transport of Dangerous Goods Act:

TDG Classification: Class 5.1 (9.2) U.N. 1490; Packing group II
WHMIS Classification: C, E
On the DSL list

Ozone. Ozone cannot be shipped; it must be generated at the point of use. Ozone can only be generated at limited concentrations; the chief hazard is exposure to ozone inhalation.

Ozone can be detected by smell at a concentration of 0.01–0.005 ppmv. However, repeated exposure increases the detection level.

For 8 h the exposure limit is 0.1 ppmv; for 10 min the exposure limit is 0.3 ppmv. Ozone equipment and process lines should be in well ventilated areas. Exposure to aqueous ozone solutions can cause skin or eye irritation.

BIBLIOGRAPHY

1. EPA. (1998). *Field Application of In Situ Remediation Technologies: Chemical Oxidation*. EPA 542-R-98-008, September. Available: http://www.epa.gov/swertio1/download/remed/chemox.pdf.
2. ESTCP. (1999). *Technology Status Review In Situ Oxidation*. November. Available: http://www.estcp.org/documents/techdocs/ISO_Report.pdf.
3. Sun, Y. and Pignatello, J.J. (1992). *Iron (III) Chelates as Catalysts for Chemical Degradation of Pesticide Wastes by Hydrogen Peroxide*. Abstract. Division of Environmental Chemistry, AMS, San Francisco, CA, April, pp. 480–481.
4. Miller, C.M. and Valentine, R.L. (1995). Hydrogen peroxide decomposition and quinoline degradation in the presence of aquifer material. *Water Res.* **29**(10): 2353–2359.
5. Teel, A.L., Wargberg, C.R., Atkinson, D.A., and Watts, R.J. (2001). Comparison of mineral and soluble iron Fenton's catalysts for the treatment of trichloroethylene. *Water Res.* **35**(4): 977–984.
6. Yan, Y.E. and Schwartz, F.W. (2000). Kinetics and mechanisms for TCE oxidation by permanganate. *Environ. Sci. Technol.* **34**(12): 2535–2541.
7. Wilberg, K.B. and Freeman, F. (1999). Kinetics of the base-catalyzed permanganate oxidation of benzaldehyde. *J. Org. Chem.* **65**: 573–576.
8. West, O.R. et al. (1997). *A Full-Scale Demonstration of in situ Chemical Oxidation Through Recirculation at the X-701B Site Field Operations and TCE Degradation*. ORNL/TM-13556.
9. Urynowicz, M.A. and Siegrist, R.L. (2000). Chemical degradation of TCE DNAPL by permanganate. In: *Chemical Oxidation and Reactive Barriers: Remediation of Chlorinated and Recalcitrant Compounds*. The Second International Conference on Remediation of Chlorinated and Recalcitrant Compounds, Monterey, CA, May 22–25, Vol. C2-6, pp. 75–82.
10. Nelson, M.D., Parker, B.L., Al Toma, A., Cherry, J.A., and Loomer, D. (2001). Geochemical reactions resulting from *in situ* oxidation of PCE-DNAPL by $KMnO_4$ in a sandy aquifer. *Environ. Sci. Technol.* **35**(6): 1266–1275.
11. Hongine, J. and Bader, H. (1983). Rate constants of reactions of ozone with organic and inorganic compounds in water 1: Non-dissociating organic compounds. *Water Res.* **17**: 173–183.
12. Staehelin, J. and Holgne, J. (1982). Decomposition of ozone in water: rate of initiation by hydroxide ions and hydrogen peroxide. *Environ. Sci. Technol.* **16**(10): 675–681.
13. Kosak-Channing, L.F. and Heiz, G.R. (1983). Solubility of ozone in aqueous solutions of 0–0.6 M ionic strength at 5–30 °C. *Environ. Sci. Technol.* **17**(3): 145–149.
14. Barcelona, M.J. and Holm, T.R. (1991). Oxidation–reduction capacities of aquifer solids. *Environ. Sci. Technol.* **25**(9): 1565–1572.

PARTICULATE TRANSPORT IN GROUNDWATER—BACTERIA AND COLLOIDS

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Enhanced bioremediation via bioaugmentation requires understanding of groundwater flow of both colloids and bacteria.

COLLOID TRANSPORT MECHANISM

Evaluation of mass transport of entrained material in groundwater flow systems is typically focused on two primary areas, dissolved organic or inorganic species, and flowing free phase liquids. However, although generally not as dominant overall, the movement of small particles in groundwater can cause problems at the low concentrations required by regulatory limits. In addition, the proponents of *in situ* bioaugmentation, through injection of bacteria, must rely on successful particulate mass transport through the saturated soil matrix.

Colloids are a special class of material with properties that lie between that of the dissolved state and the solid (or liquid) state. Historically, the term *colloid* is applied to particles with a size range of 0.001 to 1 micron (1 micron is 0.001 millimeter or approximately 0.00004 inches). The surface area per unit mass is very high for colloids, which has a great effect on their mass transport behavior. The sources of colloids in groundwater include the following:

- Detached soil, mineral, or contaminant particles
- Colloids formed from solutes undergoing geochemical precipitation caused by changes in redox conditions from mixing with injected or percolated surface water
- Emulsions of fine droplets from free phase hydrocarbons
- Agglomerations forming micelles seeded by macromolecules such as humic acids.

— Colloids introduced directly into the groundwater from landfills or other surface sources.

On a mass basis, colloid concentrations in groundwater range from 1 to 75 mg/L. Colloid density in natural groundwater systems can range to upper limits of 10^{10} particles per liter in igneous fractures and 10^{12} particles per liter in sandy aquifers.

Contaminants can be transported as colloids resulting in unexpected mobility of low solubility material. Colloids can also act as adsorbents for contaminants that are then transported with the colloid. If contaminants have been adsorbed to colloids, it is important to remember that the transport behavior is determined by the physical/chemical properties of the colloid, not the physical/chemical properties of the contaminant. In cases where colloids are formed *in situ*, contaminants can be incorporated (or occluded) into the colloidal particle as it forms.

MECHANICAL COLLOIDAL PROCESSES

Particles larger than 2 microns in the low-flow conditions common in groundwater systems are subject to removal by sedimentation (settling under the influence of gravity). Below 0.1 microns, the effects of adsorptive process are much more pronounced. As a result, colloids and particles in the range of 0.1 to 2.0 microns are likely to be the most mobile in groundwater. Although particles at the middle of the colloid range are overall more mobile, larger particles may travel through formations more rapidly because of size exclusion. In that process, the particle travels a reduced path length through the soil matrix because it is excluded from the smaller pore spaces. Colloids or other particles can be mechanically removed via filtration by the soil matrix. The key parameter to this process is the pore entrance size, which is a function of grain size. For fine- to coarse-grained silts, pore entrance size ranges from 0.7 to 7 microns, for fine- to coarse-grained sands from 24 to 240 microns, and for fine- to coarse-grained gravels 720 to 7200 microns.

Mechanical removal of particles occurs most often by straining, a process in which particles can enter the matrix, but are caught by the smaller pore spaces as they traverse the matrix. If within the soil matrix there is groundwater flow through heterogeneity, a surface mat may form at the interface when particles are too large to enter the finer grained matrix at all. The best example of this is along the walls of fractures through fine-grained sediments.

ADSORPTIVE COLLOIDAL PROCESSES

The primary forces that influence colloids suspended in a liquid moving through a porous solid matrix include electrostatic repulsion and attraction, London–van der Walls attraction, and brownian motion. Electrostatic charged based forces are familiar. London–van der Walls attraction is a weak (but still effective) form of chemical bonding. Brownian motion is caused by molecular collisions between a particle and the surrounding fluid matrix; it becomes apparent when particle size reaches a few microns. The effect predominates colloids 0.1 microns or smaller; the smaller the size, the higher the velocity that can be imparted because of brownian motion. Adsorptive interactions of colloids may be affected by the ionic strength of the groundwater; specific ionic composition and charge; quantity, nature, and size of the suspended colloids; geologic composition of the soil matrix; and flow velocity of the groundwater.

In most instances, however, the mobility of a colloid is dependent on groundwater chemistry rather than on forces caused by advective flow. Higher mobility occurs at lower overall concentrations of total dissolved solids (TDS). Higher levels of TDS encourage deposition of colloids.

The reasons for this behavior deserves some explanation. Surfaces in an aquifer matrix at normal, near-neutral pH in general have a net negative electrostatic charge because of the predominance of silica in the minerals of the matrix. This charge on the matrix surfaces and the colloids in the groundwater system has a configuration that is described as an electric double layer. The first layer forms because of the collection of positive ions on the exposed negatively charged interfacing surfaces. Anions in solution then form the second diffuse layer around the first to counter the resulting positive surface charge.

As the ionic strength of the groundwater increases, the thickness of the double layer decreases. When a negatively charged colloid approaches a negatively charged grain within the groundwater matrix (both with double electrical layers), mutually repulsive forces increase. Conversely, if the two surfaces can approach past the repulsive maximum, attractive London–van der Waals forces will take over, overcome the repulsive forces, and the colloid is attached to the matrix surface. The high velocities imparted to colloids smaller than 0.1 microns because of brownian motion provides the mechanism for overcoming the electrostatic repulsion of the double layers. The process is delicately balanced such that the reduction of the electrical double layer thickness through increased ionic strength is also required.

The result of this colloidal behavior is beneficial with regard to typical surface contaminant sources such as landfills. The high ionic strength of leachate will serve to provide optimum conditions for the immobilization of entrained colloids. Conversely, the injection of low ionic strength water has the potential to enhance colloid mobility and even mobilize previously adsorbed material.

PARTICULATE TRANSPORT IN GROUNDWATER—BACTERIA

The transport of bacteria in groundwater systems is of concern from both ends of the mobility spectrum. Mobility is essential for the injection of bacteria using the bioaugmentation approach to saturated zone *in situ* bioremediation. Conversely, there is high concern with mobility, and a desire for immobility, with regard to pathogens from septic systems or other sources. Bacteria range in size from 0.2 to 5 microns and viruses from 0.005 to 0.1 microns, which is within the size range considered colloidal. Therefore, bacteria transport is effected by many of the same processes as those for colloids.

As with colloids in groundwater, factors affecting bacteria transport include mechanical and adsorptive processes. However, bacteria are live organisms capable of responding to environmental conditions and have other unique qualities that impact their transport properties. These qualities include surface hydrophobicity, reactive groups on the surface of the bacterial cell wall, and coatings that can be "sticky."

Mechanical Processes

Removal of bacteria can occur solely by straining within the aquifer matrix. The size of the pore space in relation to bacteria size is important enough to repeat: In fine- to coarse-grained silts, pore entrance size ranges from 0.7 to 7 microns; for fine- to coarse-grained sands from 24 to 240 microns; and for fine- to coarse-grained gravels 720 to 7200 microns. As a general rule, bacteria should be half the size of the pore entrance for adequate success in passage, and if the average bacteria size is greater than 5% of the grains (not the pore size) within the porous matrix, straining becomes an important removal mechanism.

In addition to reduction of hydraulic conductivity through the accumulation of bacterial cells, as described above, other processes unique to bacteria can mechanically reduce transport efficiency. Bacteria may excrete extracellular polymers, low solubility metabolic precipitates, or gaseous products (such as nitrogen, methane or carbon dioxide) that can potentially block pore passages.

Adsorption Processes

Increasing ionic strength of groundwater increases the capability of bacteria to adhere to soil surfaces by increasing the availability of ions to act as bridges between the surfaces of the cell and soil particles, and by decreasing the thickness of the electric double layer (a phenomena described above). Increased ionic strength also enhances the ability of bacteria to aggregate, which forms a larger overall particle more likely to be captured in pore spaces. For example, transport efficiency of bacteria through clean sands has been shown to be 2% to 3% in water with 750 ppm of total dissolved solids (TDS) versus 70% to 100% in deionized water.

Bacteria have an overall negative charge on the surface of their cell wall, primarily because of the presence of peptidoglycan, the structural backbone of the bacterial cell wall, which is rich in carboxyl and amino groups. Teichoic acids are a phosphate-rich component of bacterial cell walls, which also help contribute to the presence of a negative charge. So that conversely to the above, bacterial adsorption to positively charged surfaces (such as those presented by iron and other metal oxyhydroxides) is at its maximum under conditions of low ionic strength. In a manner similar to that for inorganic anions, the sorption of bacteria to oxyhydroxides is also pH-dependent. As a consequence, to evaluate bacteria mobility, it is important to know the specific chemical and mineralogical character of the soil matrix. In some circumstances, because of soil chemistry, high levels of bacterial adsorption will occur irrespective of the manipulation of ionic strength in the surrounding groundwater. This news is good for those concerned with septic systems, but bad for those wishing to inject bioaugmentation bacteria.

The bacterial cell wall also contains varying amounts of lipids that are responsible for hydrophobic behavior. Hydrophobic bacteria have the tendency to adsorb to surfaces in the groundwater system because of repulsion from the polar water molecule. The effect of electrostatic repulsive force decreases with increasing hydrophobicity of a bacterial species.

Biological Processes

Attachment of bacteria to a surface is a two-stage process. First, the initial adsorption caused by electrostatic or hydrophobic forces takes place. This adhesion is reversible, given adequate shearing from groundwater flow. The second stage is the irreversible binding through cellular production of exopolymers that anchor the cell to the surface. The production of these exopolymers is stimulated by the initial adsorption and has been observed to increase five-fold once a cell has attached itself to a surface.

Bacterial motility may be important under conditions of static flow or chemotaxis. Motility is the ability of bacteria to move through the use of appendages such as flagella or pili. Chemotaxis is the movement or orientation of bacteria cells along a chemical concentration gradient. By this mechanism, porous media can be penetrated by bacteria that literally grow through the pore space. Growth rates have been observed in cores at ranges up to 0.01 to 0.05 cm/hr. However, pore space is occupied by bacterial biomass, the blockage from which ultimately prevents the transport of required nutrients to the growth area.

Starvation of cells will reduce their size to less than 0.3 microns, therefore reducing overall filtration effects and allowing for penetration of a finer grained matrix. However, in response to the starvation stress, many bacteria will increase the stickiness of their cell walls to improve the chance of adhesion to a surface. In general, bacterial adsorption in uncontaminated water is driven by electrostatic forces. Bacteria adsorption rates are high in groundwater at a pH of 6 or lower. Bacteria transport mobility increases above pH 6.0, significantly so above pH 7.5. In contaminated groundwater, the situation is much more complex and is specific to the site, contaminant properties, and the bacterial species present. In the presence of contamination, bacteria metabolic processes play a much more dominant role in transport properties, overwhelming physical/chemical effects.

Coarse-grained groundwater flow systems or large aperture fracture flow systems (such as karst) have the capacity to support substantial bacterial transport. Bacteria injection is possible, and pathogen transport from septic systems is probable under such conditions. However, bacteria transport in finer grained groundwater systems is more problematic. The high TDS associated with septic systems (or other potential sources of pathogens) will enhance the tendency for immobilization. Bacterial injection for bioaugmentation may be mechanically impossible because of straining and will require careful assessment under marginal porosity conditions.

Bioaugmentation

Bioaugmentation is a technology to enhance bioremediation. One type of bioaugmentation includes culturing indigenous bacteria and reintroducing that population back into the aquifer. A second type of bioaugmentation includes culturing a specific laboratory selected strain of bacteria that has had proven success degrading a specific contaminant and injecting that bacteria into the aquifer. Based on experience, the former is more likely to be successful because the indigenous bacteria are acclimated to the site-specific conditions and food source. Injection pressures for bioaugmentation tend to be 40 psi or less. Pressures approaching 100 psi will get the bacteria further into the aquifer; however, cell walls may start to burst at higher pressures. In addition, injection pressures sufficient to fracture the sediments in an injection zone will induce flow by hydrofracturing, possibly short-circuiting the bulk of a targeted treatment zone. As a rough rule of thumb, injection pressures can increase by 1 psi for each foot of depth. Injections pressures higher than that are likely to stimulate hydrofracturing.

Groundwater flow studies must take into account both colloids and bacteria for successful enhanced bioremediation projects.

PERCHED GROUNDWATER

Leo S. Leonhart
Hargis Associates, Inc.
Tucson, Arizona

.

Perched ground water is subsurface water that forms a saturated horizon within porous media at an elevation higher than the local or regional groundwater table. This condition may result from various field conditions, but the requisite condition is a soil or rock horizon of significantly lower vertical hydraulic conductivity (K_v) (perching unit) amid contrastingly higher K_v strata. Downward movement of infiltrating water through the higher K_v materials is retarded and accumulates above and/or within the perching unit, which typically "mounds" saturated water above the perching unit, leaving an unsaturated (vadose) horizon below the perching unit. The upper limit of saturation in a perched zone is referred to as a perched water table, whereas the lower limit has been defined as an inverted water table.

HYDRAULIC PROPERTIES

K_v contrasts typically of the order of 10 to 20% of the K_v of the overlying materials, are believed to have the potential to produce perched groundwater (Fig. 1) (1,2). Stratigraphically, such contrasts are often conceptualized as fine silty or clayey lenses layered between coarser sediments such as sands or sandy gravels. Equivalent K_v contrasts can result within other stratigraphic analogues such as saprolite or ash horizons between adjacent lava flows, anisotropic fractures patterns in homogeneous crystalline rocks, or intraflow structures within flood basalts. Perching units can vary considerably in areal extent, ranging from extensive, essentially continuous strata to localized, discontinuous lenses.

Porosity within perching and overlying stratigraphic units is typical of the types of materials that comprise these strata. Notably, the porosity of clayey materials is typically higher than that of other granular materials, thereby providing high capacity for storing water but a lesser or slower rate of yield. Secondary porosity of clays, such as fractures, may facilitate downward movement, particularly as water is mounded above the perching unit.

Elastic properties are another potentially important factor in regard to perched groundwater involving clayey units. Specifically, shrinking of the clay lattice may occur upon desiccation, and swelling may occur upon

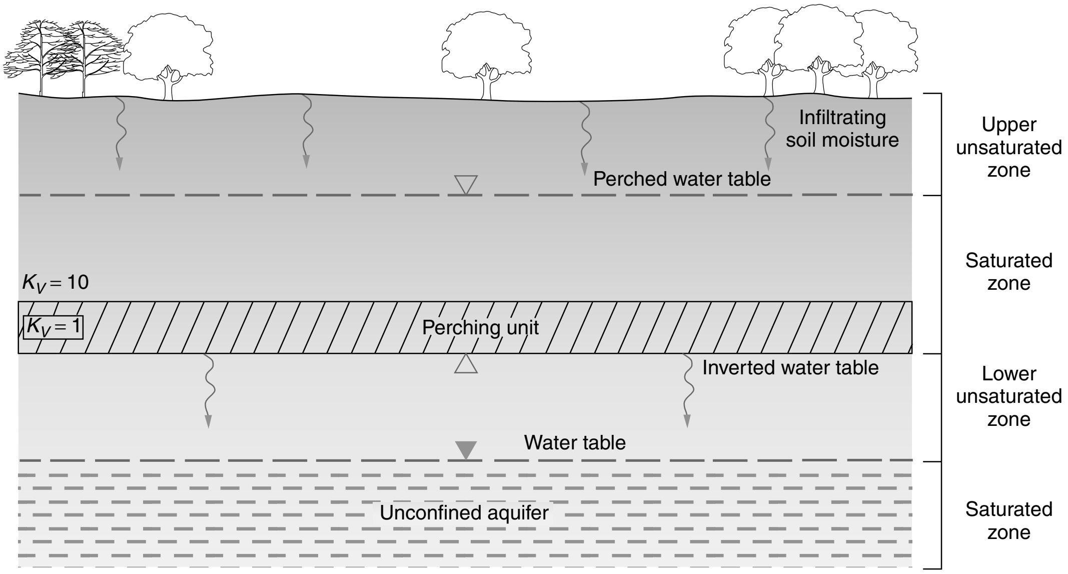

Figure 1. Perched groundwater resulting from hydraulic conductivity contrast and surface infiltration.

wetting, thereby, respectively, enabling greater or lesser rates of water movement through the perching unit. Other rarer forms of secondary porosity may result from certain geochemical interactions, such as dissolution or precipitation of minerals upon wetting.

HYDROLOGIC CONDITIONS

Hydrologic conditions that give rise to perched ground water are variable. Perched groundwater tables have resulted from various high volume natural recharge events, such as flooding or intense and long-term precipitation; artificial recharge resulting from infiltration basins, water line leaks, irrigation, disposal pits, or artificial impoundments; and transient conditions such as falling water tables. In general, all of these mechanisms involve downward movement or drainage of subsurface water. For this reason, perched groundwater is often characterized as transient or ephemeral. However, longer term recharge and discharge may give rise to near steady-state perched groundwater. Wells tapping perched groundwater can have highly variable yields, depending on the specific hydrologic conditions. Generally, however, perched groundwater is not exploited.

Mathematical equations describing the rise and fall of groundwater mounds were developed by Hantush (3) and generally relate to conditions underlying recharge basins. These equations are also applicable to perched groundwater.

Except for perched groundwater resulting from falling water tables, development of a perched groundwater zone can be conceptualized as having three phases: growth, accumulation, and drainage. The perched zone growth phase involves the initial downward movement of subsurface water and its encounter with the perching unit. The downward flux of water must be sufficient to overcome soil moisture tension and is subject to the classical principles of unsaturated flow.

Once a sufficient volume of water encounters the perching unit, accumulation or “mounding” of groundwater begins. Initially, this mounding may be limited to the overlying coarser grained units (or equivalent units in other terrains). As downward flux continues, the mound will accrete both vertically and laterally. As accumulation continues, the added hydraulic head can advance the inverted water table downward into the perching unit. Likewise, lateral movement may continue, potentially extending to the edges of the perching unit, where “spillage” into the underlying unsaturated zone can occur.

The drainage phase represents waning of the perched zone. If the rate of recharge is insufficient to maintain the areal and vertical extent of the perched zone, then it will slowly drain through or off the ends of the perching unit. Another source of discharge may result from evapotranspiration, if the perched groundwater is shallow and within reach of phreatophytes.

Examples of Perched Groundwater

Tucson Basin. The municipality of Tucson, Arizona, relied solely on its underlying aquifer for water supply until the recent delivery of Colorado River water via the Central Arizona Project. As a result of the region’s population growth and associated increases in water demand since the early 1950s, the water table declined more than 200 feet within some areas of the basin. The rapid lowering of the water table that averaged 3 to 4 feet per year from 1950 to 2000 left saturated remnants in several parts of the basin (4). The basin sediments comprise a heterogeneous sequence of unconsolidated to semiconsolidated strata ranging in textures from laterally discontinuous facies of sandy gravels to silty clays. Perched groundwater has reportedly been encountered at depths ranging from 30–85 feet below land surface and ranges in saturated thickness from less than 1 foot to more than 85 feet (5–7). In some places within the basin, the elevation of the perched water table is within the perching unit as opposed to the overlying sediments. In other areas, the perched groundwater is sustained as a result of natural recharge from runoff or artificially from irrigation.

High Volcanic Islands. Perched groundwater is common in high volcanic islands, such as Hawaii and Tahiti, owing to high precipitation upon highland areas and extreme K_v contrasts in layered shield volcanics (8–11). Perched groundwater is often classified as “high level” compared with basal groundwater, which occurs as near or below sea level lenses, “floating” on intruded seawater (Fig. 2). As also seen in the figure, high level water can also result from intrusive dikes, which act as “subsurface dams.” Perching units are generally sedimentary strata, often saprolites, dense lavas, or ash beds interbedded among lava flows. Yields from such perched sources are highly variable; however, exploitation by horizontal wells and tunnels has provided adequate local supplies. As a result of high relief, perched groundwater can often intercept the surface topography and discharge as a spring. Equivalent examples are also found in high-relief fractured crystalline terrains.

Flood Basalts. The hydrogeology of extensive flood basalts such as occur in the Columbia Plateau in parts of Washington, Oregon, and Idaho is controlled largely by the characteristics of intraflow structures. Interflow structures such as those sedimentary interbeds and saprolites described for shield volcanics are also important. In such volcanic terrains, individual lava flows are much thicker than those in shield volcanics, ranging up to 10s to 100s of meters. Intraflow structures from the base to top of an individual flow generally comprise a contact or pillow palagonite, representing a chill zone; a colonnade, consisting of prismatic, hexagonal columns with vertical cooling joints; an entablature, comprising slender, hackly, often-intersecting columns; and a flow top of vesicles and rubbly materials (Fig. 3). From flow to flow, the proportional thicknesses of these structures may vary extensively.

Interflow contacts generally represent zones of higher porosity and permeability than the dense flow interiors. Such zones of comparatively higher porosity preferentially store and transmit higher volumes of water, thereby potentially creating multiple perched horizons. Being

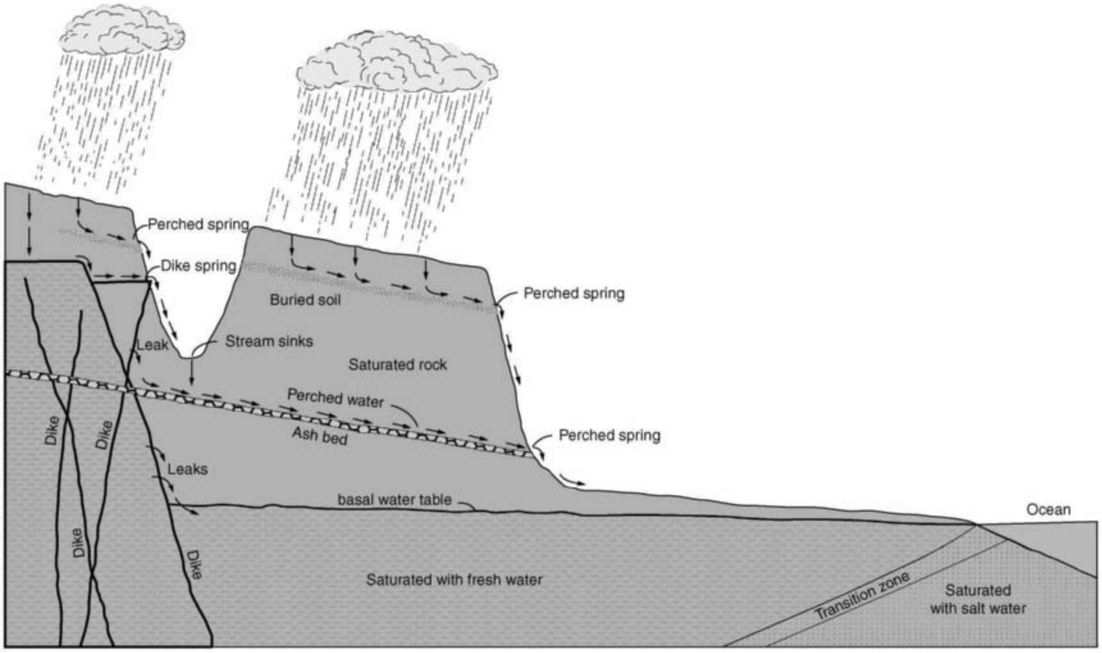

Figure 2. Conceptualization of perching on high volcanic islands (8).

a relatively competent rock, wells completed in basalt are often uncased. This construction often results in the borehole penetrating into several variably saturated horizons. As a result, cascading water from shallower perched zones is common and can be observed in borehole video logs and even heard by listening at the wellhead.

ROLE OF PERCHED GROUNDWATER IN CONTAMINANT MOVEMENT

Considering the position of perched groundwater bodies in relation to the surface, where much contaminant generation occurs, and the groundwater table, it is easy to conceptualize how perched groundwater can play a role in controlling the movement of contaminants in the subsurface. First, perched groundwater may result from artificial recharge due to the handling and disposal of liquids and liquid wastes. Leaky impoundments, such as disposal pits and sumps, underground storage tanks (USTs), pipelines, drywells, landfills, channels and ditches, sewers, and poorly constructed wells can produce reservoirs of contaminated perched groundwater. Similarly, nonpoint sources such as agricultural areas that receive fertilizer and pesticide applications, construction sites, industrial sites, and municipal areas can also constitute sources of liquid contamination.

Such reservoirs of contaminated perched groundwater can act to control the movement and fate of these contaminants within the subsurface. Perched horizons may act as continuing sources, “filters,” vapor reservoirs, or nonaqueous phase liquid (NAPL) reservoirs, depending on both the nature of the contaminant source and the site hydrogeology. At some sites, perching of contaminants may locally prevent contamination of underlying groundwater by intercepting the downward migration of contaminants. On the other hand, this condition may only postpone the inevitable contamination and may prevent earlier detection of sources and incipient problems.

Related Terms & Concepts

Semiperched Groundwater. Semiperched groundwater may be described as a zone of saturation that overlies a semiconfined aquifer. Thus, the perching unit is leaky, and a hydraulic relationship exists between the semiperched aquifer and the underlying semiconfined aquifer and often allows upward migration of water into the semiperched zone. In this case, the saturated zone effectively extends to the top of the semiperched zone, in contrast to the traditional perched zone, which is positioned over an unsaturated zone. Failure to recognize and account for such zones has been identified as a potential issue in interpreting regional groundwater level trends (13).

In another context, semiperched groundwater that occurs among cyclothems in coal regions of Pennsylvania is discussed. Again, the permeability contrasts created in a stratigraphic sequence of alternating lithologies detains or retards the vertical movement of water through the unsaturated zone creating both perched and semiperched groundwater. Depending on topography, these zones may also discharge as springs. Such terrains have been termed “hydrologic islands” and have been related to a hierarchical classification of local, intermediate, and regional groundwater flow (14).

Tension-Saturated Zone. Tension-saturated zone is another term for the capillary fringe overlying the saturated zone wherein pores of the medium are saturated, but the pressure heads are less than atmospheric (15).

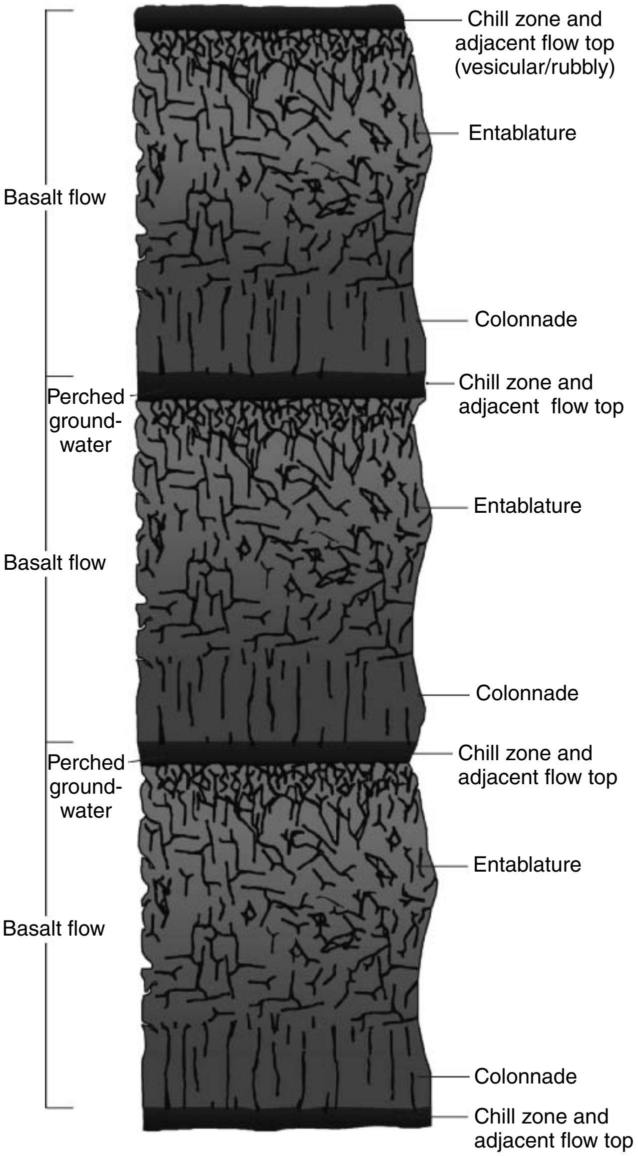

Figure 3. Conceptualized intraflow structures in flood basalts (12).

Temporary Saturated Zones. Perched groundwater may result locally from various ephemeral events as well as in response to surface disturbances. Conditions that lead to sudden recharge are favored by high antecedent moisture conditions; high-intensity rainfall; and various conditions related to storm duration, infiltration, slope, and subsurface strata. Anthropogenic activities such as timber clearing, grading, road construction, and extreme natural events such as flooding and snowmelt may create perched groundwater. In some instances, perched groundwater has been cited as a cause for slope failure (16).

BIBLIOGRAPHY

1. Bouwer, H. (1978). *Groundwater Hydrology*. McGraw-Hill, New York.
2. Freeze, R.A. (1971). Three-dimensional, transient, saturated-unsaturated flow in a groundwater basin. *Water Resour. Res.* **8**: 347–366.
3. Hantush, M.S. (1967). Growth and decay of groundwater-mounds in response to uniform percolation. *Water Resour. Res.* **3**: 227–234.
4. Anderson, S.R. (1988). *Potential for Aquifer Compaction, Land Subsidence, and Earth Fissures in the Tucson Basin, Pima County, Arizona*. U.S. Geological Survey Hydrologic Investigations Atlas HA-713.
5. Parsons. (1986). *Aviation Corridor Highway (State Route 210), Junction I-120 to Junction SR-210: Subsurface Investigations, Phase I, Geotechnical Report*. Prepared by Parsons, Brinkerhoff, Quade, Douglas, Inc. for Arizona Department of Transportation, Highways Division.
6. Leake, S.A. and Hanson, R.T. (1987). *Distribution and Movement of Trichloroethylene in Ground Water in the Tucson Area, Arizona*. U.S. Geological Survey Water-Resources Investigation Report 86–4313.
7. Nelson, D.F. and Dulaney, A.R. (1996). *A Resource Effective Management Tool for Multi-Agency Projects Through GIS: As Implemented in the Downtown Tucson Groundwater Study*. Arizona Department of Environmental Quality. ESRI Symposium Paper.
8. Stearns, H.T. and Macdonald, G.A. (1946). *Geology and Ground-Water Resources of the Island of Hawaii*. Hawaii Division of Hydrography, Bull. 9.
9. Peterson, F.L. (1972). Water development of tropic volcanic islands—types example: Hawaii. *Ground Water* **10**(5): 18–23.
10. Stearns, H.T. (1966). *Geology of the State of Hawaii*. Pacific Books, Palo Alto, CA.
11. Macdonald, G.A., Abbott, A.T., and Peterson, F.L. (1983). *Volcanoes in the Sea: The Geology of Hawaii*, 2nd Edn. University of Hawaii Press, Honolulu, HI.
12. USFS. Columbia River Basalt. Website constructed by Portland State University on behalf of U.S. Forest Service. Available: http://www.fs.fed.us/r6/columbia/basalt.htm.
13. Pederson, D.T. (2001). Effect of semi-perched ground water on monitoring of ground water levels in a developed aquifer. *Irrigation Drainage Abstr.* **27**(4): 292.
14. Poth, C.W. (1963). *Geology and Hydrology of the Mercer Quadrangle, Mercer, Lawrence, and Butler Counties, Pennsylvania*. PA Topographic and Geological Survey, Water Resource Report No. 15.
15. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
16. Chatwin, S.C., Howes, D.E., Schwab, J.W., and Swanston, D.N. (1994). A guide for management of landslide-prone terrain in the Pacific Northwest. *Land Management Handbook* No. 18. 2nd Edn. British Columbia Ministry of Forests.

PERMEABILITY

Nitish Priyadarshi
Ranchi University
Ranchi, Jharkhand, India

Permeability is a measure of the ease with which fluids pass through a porous medium. It was established experimentally by Darcy in 1856 that

$$k = (Q/A)n(dp/ds)^{-1}$$

in which k is the permeability, Q is the volume of fluid which passes in unit time through area A in the direction s of the pressure gradient dp/ds, and n is the viscosity of the fluid. The unit of permeability in most common use by geologists and members of the oil industry is the darcy (D), which corresponds to a flow of 1 $\text{cm}^3\,\text{s}^{-1}\,\text{cm}^{-2}$ at a pressure gradient of 1bar cm^{-1}. If the fluid is water, material with a permeability of 1D will pass $10^{-3}\,\text{cm}^3\,\text{s}^{-1}\,\text{cm}^{-1}$ at a pressure gradient of 1 $\text{g}\,\text{cm}^2\,\text{cm}^{-1}$.

A few typical values of permeability for specific rock types are listed in Table 1. The permeability values quoted in Table1 have been determined in experiments in which the rock specimens have usually been subjected to moderate and constant confining pressure.

Permeability is measured in darcys. One darcy is equal to the passage of 1 cm^3 of a fluid with a viscosity of 1 cP (centipoise) in 1 s under a pressure difference of 1 atm through a porous medium having a cross section of 1 cm^2 and length of 1 cm (1). The centipoise is 1/100 of a poise, which is the cgs unit of dynamic viscosity. The poise is equal to $1\,\text{dyn}\,\text{s}\,\text{cm}^{-2} = 0.1\,\text{Pa}\,\text{s}$ (because dynes/square centimeter = pressure). Permeability ranges from 1 to 5 darcy in loose sediments and from 0.05 to 1 darcy in rocks. The lowest value (0.05 darcy) applies to cherts, which are practically impermeable masses of microcrystalline quartz. A rock may have high porosity but no permeability because the holes are not interconnected (e.g., a vesicular lava) or because they are too small (e.g., clay).

Permeability is high where pore spaces are large and numerous and where they are freely intercommunicating. If the pores are large but are not intercommunicating, the rock may have low permeability even though its porosity is high. A rock with pore spaces of given size, abundance, and continuity is more permeable to fluids of low than of high viscosity. Permeability is sometimes called tightness. An oil-bearing rock is tight if it yields its oil with difficulty, and it is open if it yields its oil freely.

In the study of groundwater, the permeability of rocks is of fundamental importance. Generally speaking, highly permeable rocks form good aquifers. Sand and gravel are more permeable than silt and clay. Therefore, gravel and sand constitute a very large proportion of well-developed aquifers. Sandstones generally make good aquifers. Rocks of volcanic origin, such as basalts, make good aquifer materials only if they are fractured, vesicular, and porous. Granite, gneiss, and other crystalline and metamorphic rocks are generally impermeable and hence, do not form good aquifers (3). A geologic formation that is completely impermeable and is devoid of water is called an aquifuge.

Table 1. Typical Values of the Permeability of Some Rock Types[a]

Rock Type	Permeability, k, (D)
Sandstones and limestones	10^0–10^{-2}
	10^{-2}–10^{-5}
Argillaceous	10^{-5}
	10^{-6}–10^{-8}
Halite	$< 10^{-6}$
Marble	$< 10^{-7}$

[a] Reference 2.

When aquifers are confined by an impermeable layer or stratum, the aquifers are called confined aquifers. If the confining unit is relatively impermeable, such as clay, it is called an aquiclude. If, however, the stratum is not completely impermeable, such as sandy clay, then it is called an aquitard. Such poorly permeable strata leak water to adjacent aquifers.

The permeability of rocks depends on

1. The size and shape of the constituent grains. A uniformly coarse-grained gravel forms a highly permeable aquifer, but when the same rock consists of fractions of various sizes, there is a tendency for smaller sediments to get into and fill or block the interstices rendering it less permeable. Rounded fractions result in higher porosity and hence better permeability than angular fractions in the same grain size range (3).
2. Sorting of grains. Well-sorted materials (i.e., where grains of a particular size are in one layer followed by grains of the next size) have higher permeability than ill-sorted or unsorted grains in a rock.
3. Continuity and nature of interstices, especially when they are secondary. The mere presence of cracks or cavities or joints does not make a rock permeable; all openings must also be interconnected or mutually communicating; only then is the flow of water through the rock mass possible.

Permeability Investigation for a Reservoir Site

Many types of rocks may be traversed by a proposed reservoir. For ideal conditions, they should be impervious and impermeable. The main object of permeability investigations is, therefore, to determine the character of the rocks of the reservoir region and locate the "highly permeable" formations that can be treated as most unsuitable for the reservoir or dam site.

For reservoir sites, rocks may be divided into three groups: highly permeable, occasionally permeable, and least permeable. Highly permeable rocks include well-sorted gravel, sand, and coarse sandstone and deposits of glacial origin such as moraines and till. Occasionally permeable includes those rocks that acquire permeability due to secondary features such as cracks, fissures, solution cavities, and gas cavities. Limestone, vesicular basalt, jointed granite, and gneiss are examples of such rocks, which are otherwise impervious but are often rendered unsuitable because of prolific development of the secondary features described. Least permeable are igneous rocks, which are unjointed and thoroughly crystalline; massive limestones and sandstones can generally be considered impermeable.

Relation of Porosity and Permeability

Good correlation if pores are clean (if pore tortuosity is low); poor to no correlation if pore tortuosity is high because of clay cements, etc.

BIBLIOGRAPHY

1. Emiliani, C. (1997). *Planet Earth Cosmology, Geology and the Evolution of Life and Environment*. Cambridge University Press, New York, p. 322.
2. Fyfe, W.S., Price, N.J., and Thompson, A.B. (1978). *Fluids in the Earth's Crust*. Elsevier Scientific, New York, p. 255.
3. Singh, P. (1978). *Engineering and General Geology*. Katson, Ludhiana, p. 369.

GROUNDWATER VULNERABILITY TO PESTICIDES: AN OVERVIEW OF APPROACHES AND METHODS OF EVALUATION

Keith Loague
Stanford University
Stanford, California

Dennis L. Corwin
USDA-ARS George E. Brown,
Jr., Salinity Laboratory
Riverside, California

INTRODUCTION

Groundwater contamination is one of the most important environmental quality concerns. The subsurface migration of hazardous wastes from point sources firmly caught the public's attention in 1977 with the well-publicized Love Canal story. Characteristically, point sources are relatively easy to identify and control. Groundwater contamination from nonpoint sources (NPSs) is a greater environmental problem (1). Nonpoint sources, as a result of contaminants entering the environment over an extensive area, are difficult (at best) or impossible to trace to a specific location and have the potential for maintaining a relatively long active presence in entire ecosystems. Historically, point sources have received the greatest attention because of the conspicuous severity of their impacts. In recent years, public, political, and scientific attention have shifted toward NPSs. This shift reflects an awareness of the scope and potential impact of NPS pollution (2,3). Regulation, monitoring, litigation, and remediation have each helped to control or reduce both point and NPSs of groundwater contamination.

The word *pesticide* is a general term used to refer to the chemicals that are employed to control, destroy, mitigate, prevent, or repel pests (plant or animal). There are several classes of pesticides (e.g., algaecides, fungicides, germicides, herbicides, insecticides, miticides, molluskicides, rodenticides, and termiticides). In the United States, the Environmental Protection Agency (EPA) regulates pesticides under the authority of the 1996 amendments to the *Federal Insecticide, Fungicide, and Rodenticide Act* (FIFRA) of 1947. All pesticides must be registered and carry a *label* (providing the rules for application rates and frequencies) approved by the EPA. The distribution and trends of pesticides in the atmosphere (4), groundwater (5), surface water (6), and fluvial sediments and aquatic biota (7) have been carefully assessed. For example, there have been well over 100 monitoring studies in recent years, which quantitatively document the occurrence of pesticides in groundwater (5). The impact of widespread pesticide use, in what Freeze (8) calls *the age of awakening* (i.e., environmental perspective) or *the age of social upheaval* (i.e., new right perspective), was lucidly bought to the public's attention by Rachel Carson in her 1962 book *Silent Spring* (9), which in large part forced the banning of DDT. Barbash and Resek (5) provide a glossary of 600 pesticides that are of interest relative to groundwater impacts. Table 1 identifies 31 pesticides that have received considerable attention.

In many areas of the world, successful agriculture depends on both pesticides and irrigation. The combination of repeated pesticide applications and irrigation over a large NPS area often leads to leaching through the vadose zone and, subsequently, to groundwater contamination. For example, in the semiarid San Joaquin Valley in California, where groundwater contamination concentrations have been among the highest reported, the amount of active pesticide ingredient applied annually is on the order of 50 million kilograms. A spill of concentrated pesticide at a mixing or distribution site is an example of point source pollution. For example, in April 1977, there was an accidental spill of approximately 1900 liters of the pesticide EDB (see Table 1) within approximately 20 meters of a well that provided drinking water to the village of Kunia on the Hawaiian island of Oahu. The area around the spill eventually became a *Superfund* site.

VULNERABILITY ASSESSMENT

Assessment involves determination of the change in some constituent over time. This change can be measured in either real time or simulated by a model. Real-time measurements reflect the activities of the past, whereas simulations can provide useful glimpses into the future. Both means of assessment are valuable. The distinct advantage of simulation is that it can be used to alter the occurrence of detrimental conditions before they occur. It should be pointed out that simulation cannot replace *real* data.

By definition, a model integrates existing knowledge into a framework of rules, equations, and relationships to quantify the way a system behaves. As long as a model is applied over the range of conditions from which they it was initially developed, it serves as a useful tool for prognostication. Models can range in complexity from the simplest empirical equation to complex sets of partial differential equations that are solvable only by numerical approximation techniques.

Table 1. Thirty-One Important Pesticides[a]

Common Name	Chemical Name	Use
1,2-Dichloropropane[b,c]	1,2-Dichloropropane	Nematicide
2,4-D[b]	2,4-Dichlorophenoxy acetic acid	Herbicide
2,4-DP[c]	Butoxyethyl ester of (±) 2-(2,4-dichlorophenoxy)propanoic acid	Herbicide
Alachlor[b]	2-Chloro-*N*-(2,6-diethylphenyl)-*N*-(methoxymethyl)acetamide	Herbicide
Aldicarb[b]	2-Methyl-2-(methylthio)-propionaldehyde *O*-(methylcarbamoyl)oxime	Insecticide
Atrazine[b,c]	2-Chloro-4-ethylamino-6-isopropylamino-*S*-triazine	Herbicide
Bromacil[b]	5-Bromo-3-(*sec*-butyl)-6-methyluracil	Herbicide
Carbaryl[b]	1-Naphthyl-*N*-methylcarbamate	Insecticide
Carbofuran[b]	2,3-Dihydro-2,2-dimethyl-7-benzofuranyl-*n*-methylcarbamate	Insecticide
Carboxin[b]	5,6-Dihydro-2-methyl-1,4-oxathiin-3-carboxanilide	Fungicide
Chlorothalonil[b]	Tetrachloroisophthalonitrile	Fungicide
Cyanazine[b]	2[[4-Chloro-6(ethylamino)-*S*-triazin-2-yl]amino]-2-methylpropionitrile	Herbicide
Dalapon[b,d]	2,2 Dichloropropionic acid (sodium salt)	Herbicide
DBCP[c,d]	1,2-Dibromo-3-chloropropane	Nematicide
DCPA[b]	Dimethyl 2,3,5,6-tetrachloro-1,4-benzenedicarboxylate	Herbicide
Diazinon[b,c]	*O*, *O*-Diethyl-*O*-(2-isopropyl-4-methyl-6-pyrimidinyl)phosphorothiote	Insecticide
Dicamba[b,c]	2-Methoxy-3,6-dichlorobenzoic acid	Herbicide
Dinoseb[b,d]	2-*sec*-Butyl-4,6-dinitrophenol	Herbicide
Diphenamid[b,d]	*N*,*N*-Dimethyl-α -phenylbenzeneacetamide	Herbicide
Disulfoton[b]	*O*,*O*-Diethyl *S*-[2-(ethylthio)ethyl]phosphorodithioate	Insecticide
Diuron[b,c]	3-(3,4-Dichlorophenyl)-1,1-dimethylurea	Herbicide
EDB[c,d]	1,2-Dibromoethane	Insecticide
Methomyl[b]	*S*-Methyl-*N*-(methylcarbamoyl)oxy)-thioacetamidate	Insecticide
Metolachlor[b]	2-Chloro-*N*-(2-ethyl-6-methylphenyl)-*N*-(2-methoxy-1-methylethyl)acetamide	Insecticide
Oxamyl[b]	Methyl 2-(dimethylamino)-*N*-[[(methylamino)carbonyl]oxy]-2-oxoethanimimdothioate	Insecticide
Prometon[c]	6-Methoxy-*N*,*N*′-bis(1-methylethyl)-1,3,5-triazine-2,4-diamine	Herbicide
Prometryn [c]	*N*,*N*′-Bis(1-Methylethyl)-6-(methylthio)-1,3,5-triazine-2,4-diamine	Herbicide
Propazine[c,d]	6-Chloro-*N*,*N*′-*bis*(1-methylethyl)-1,3,5-triazine-2,4-diamine	Herbicide
Simazine[b,c]	2-Chloro-4,6-bis(ethylamino)-s-triazine	Herbicide
Tebuthiuron[b]	*N*-[5-(1,1-Dimethyl)-1,3,4-thiadizol-2-yl]-*N*,*N*′-dimethylurea	Herbicide
Trifluralin[b]	2,6-Dinitro-*N*,*N*-dipropyl-4-(trifluoromethyl)benzenemamine	Herbicide

[a] abstracted from the 341 pesticides in Hornsby et al. (10)

[b] Identified by EPA (11) as high risk for groundwater contamination.

[c] Detected in groundwater in California (12).

[d] Cancelled or severely restricted in the United States.

The use of models to aid in the assessment of groundwater vulnerability is now a well-established tool, employed both in industry and by regulatory agencies. The various mechanisms that facilitate the migration of pesticides from the surface/near surface to the water table and the methods used to simulate these processes have been the subject of considerable review (13–27). In general, subsurface flow and pesticide transport are, respectively, estimated by Richards' equation and the advection–dispersion equation, or simplifications. The processes that are important to consider in assessing groundwater vulnerability to pesticide contamination are advection, diffusion, dispersion, sorption, transformation, and volatilization (28). The increasing availability of geographic information system (GIS) software to those assessing the potential for NPS groundwater contamination from pesticides has resulted in the generation of regional-scale multicolored vulnerability maps for use in decision management (29).

Three categories of models for assessing groundwater vulnerability from pesticides are (1) regression models, (2) index models, and (3) transient-state solute transport models. Regression models generally use multiple linear regressions to relate various causative factors. Index models, which can be property or process based, compute an index of pesticide mobility. Transient-state, process-based models are capable of simulating the concentration/movement of a pesticide in a dynamic flow system.

One of the first and most promoted regional-scale groundwater vulnerability methods for assessing the impact of pesticides was the DRASTIC index (30). The early process-based simulations by Petach et al. (31) were both elegant and ambitious considering the large scale of the pesticide leaching assessments. During the last two decades, a number of process-based models have been developed for assessing the behavior of pesticides in the near surface (e.g., PRZM (32), LEACHM (33), GLEAMS (34), HYDRUS-1D (35), LEACHP (36), RZWQM (37), PRZM-2 (38), PELMO (39), MACRO (40), HYDRUS-2D (41), and PESTLA (42)). Table 2 lists selected pesticide leaching assessments for California, Hawaii, and Tenerife.

UNCERTAINTY

There can be considerable uncertainty in assessments of point and NPS pollution to groundwater from pesticides. Three sources of error inherent in modeling are (1) model error, (2) input error, and (3) parameter error. Model

Table 2. Summary of Selected Pesticide Leaching Assessments in California, Hawaii, and Tenerife

Location	Model Name	Model Category[f]	Reference
San Joaquin Valley, California	PSCLR	R	(43)
	AF[a]	I	(44,45)
	PRZM-2[b]	TSST	(46)
	TTF[c]	TSST	(47)
Oahu, Hawaii	AF	I	(48–56)
	PRZM[d]	TSST	(57–60)
Tenerife, Canary Islands, Spain	AF	I	(61,62)
	Li[e]	I	(63)
	AF, PRZM-2	I, TSST	(64)

[a] Attenuation factor (65).
[b] Pesticide root zone model - Version 2 (37).
[c] Type transfer function model (66).
[d] Pesticide root zone model (31).
[e] Leaching index (67).
[f] R: regression model; I: index model; and TSST: transient-state solute transport model.

error results in the inability of a model to simulate the given process, even with the correct input and parameter estimates. Input error is the result of errors in the source terms. Input error can arise from measurement, juxtaposition, and/or synchronization errors. Parameter error has two possible connotations. For models requiring calibration, parameter error usually is the result of model parameters that are highly interdependent and not unique. For models with physically based parameters, parameter error results from an inability to represent aerial distributions on the basis of a limited number of point measurements. The aggregation of model error, input error, and parameter error is the total (or simulation) error. For multiple-process and comprehensive models, simulation error is complicated further by the propagation of error between model components.

The methods for characterizing uncertainty can, in general, be grouped into three categories (68): (1) first-order analysis, (2) sensitivity analysis, and (3) Monte Carlo analysis. First-order analysis is a simple technique for quantifying the propagation of uncertainty from input parameter to model output. Sensitivity analysis is used to measure the impact that changing one factor has on another. Monte Carlo analysis is a stochastic technique for characterizing the uncertainty in complex hydrologic response model simulations. Loague and Corwin (68) provide examples of first-order uncertainty analysis, sensitivity analysis, and Monte Carlo simulation. Note that most of the groundwater vulnerability assessments listed in Table 2 were designed to identify the uncertainty in the estimates relative to the uncertainty in the chemical, climate, and soil databases.

FRESNO CASE STUDY, AN EXAMPLE

Between the late 1950s and the time of its statewide cancellation in August of 1977, there was widespread use of DBCP (see Table 1) throughout the San Joaquin Valley in California. More than two decades after its cancellation, DBCP-contaminated groundwater persisted as a problem in the San Joaquin Valley. The objective of the Fresno case study (46,69) was to address, from a simulation perspective, if *label recommended* NPS applications were likely to be the principal source of the DBCP groundwater contamination in Fresno County (note that DBCP was a legacy when the study was conducted). The relatively unique characteristic of the Fresno case study is that both the unsaturated near surface and the saturated subsurface were considered for a 2184 km^2 area.

The numerical model used for 1-D simulations (46) of dissolved phase DBCP concentration profiles in the unsaturated zone was PRZM-2 (37). The potential fate and transport of DBCP between the surface and the water table for multiple NPS applications were quantitatively estimated from 1172 separate 35-year simulations. The aggregate of the DBCP concentrations loaded to the water table made up the annual loading files for the 3-D saturated transient transport simulations. The numerical models used for the 3-D simulations (69) of saturated subsurface fluid flow and DBCP transport, for the same 35-year period, are MODFLOW (70) and MT3D (71), respectively. A 76,440-element mesh was used for the saturated simulations.

The simulation results from the Fresno case study lead to several conclusions (46,69): (1) the areas most likely to facilitate DBCP leaching through the entire unsaturated soil profile were targeted; (2) the first appearance of DBCP above the detectable limit at the water table was simulated as most likely to have occurred between 1961 and 1965; (3) the estimated DBCP concentrations reaching the saturated subsurface exceed the maximum contaminant level (MCL) at several locations at different times; (4) the first appearance above the MCL was between 1965 and 1970 (note that by 1990, the concentrations are below the MCL); (5) relative to the size of the study area, the extent and duration of the estimated DBCP contamination was small; (6) DBCP concentrations are a function of spatial and temporal variations in the application rates, the application frequency, the unsaturated profile thickness, the soil-hydraulic properties, and the near-surface sorption; and (7) the DBCP plume evolves (grows and retracts) with time due to the loading rates at the water table. Figure 1 shows 1971 snapshots of the simulated DBCP loading to the water table and the 3-D DBCP plume in the saturated subsurface. Note that the Fresno case study assessments were critically evaluated and extended (72,73).

EPILOGUE

The ethics, efficacy, and economics of widespread pesticide applications have all received considerable attention in recent years. Without question, it is important to assess the vulnerability of groundwater to contamination from widespread applications of pesticides. For more than 10 years, the U.S. Geological Survey's National Water-Quality Assessment (NAWQA) program has played a key role in this assessment; more than 50 major river basins and aquifers were studied (74). The overall picture emerging from the NAWQA program is that pesticides

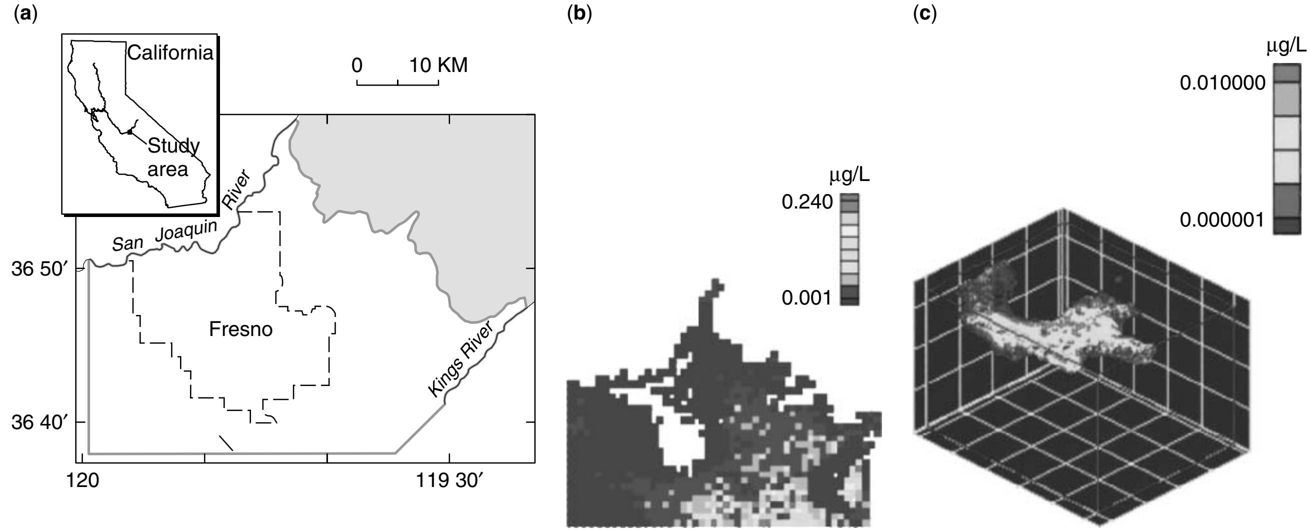

Figure 1. Results from the Fresno case study. (**a**) Location of study area. (**b**) Plan view snapshot of the simulated DBCP concentration loading at the water table for 1971 (46). (**c**) 3-D snapshot of the simulated DBCP concentration plume for 1971 (69).

typically occur at low levels, punctuated by seasonal pulses in concentration (75). Specific to groundwater, pesticides were most frequently found in shallow systems in urban and agricultural areas (74). The patterns of pesticide detection (over space and in time) identify problem areas and chemicals, warn of potential problems, and provide valuable feedback on to improve conditions (75). The use of mathematical models for assessing groundwater vulnerability from pesticides, especially after ground truth comparisons, fills in information gaps, identifies critical areas and chemicals for future monitoring, and provides the "what if" capability needed for both regulation and remediation.

BIBLIOGRAPHY

1. National Research Council. (1993). *Ground Water Vulnerability Assessment: Predicting Relative Contamination Potential under Conditions of Uncertainty*. National Academy Press, Washington, DC.
2. Corwin, D.L. and Loague, K. (Eds.). (1996). *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone*. Special Publication 48, Soil Science Society of America, Madison, WI.
3. Corwin, D.L., Loague, K., and Ellsworth, T.R. (Eds.). (1999). *Assessment of Non-Point Source Pollution in the Vadose Zone*. Geophysical Monograph 108, AGU Press, Washington, DC.
4. Majewski, M.S. and Capel, P.D. (1995). *Pesticides in the Atmosphere: Distributions, Trends, and Governing Factors*. Ann Arbor Press, Chelsea, MI.
5. Barbash, J.E. and Resek, E.A. (1996). *Pesticides in Ground Water: Distributions, Trends, and Governing Factors*. Ann Arbor Press, Chelsea, MI.
6. Larson, S.J., Capel, P.D., and Majewski, M.S. (1997). *Pesticides in Surface Water: Distributions, Trends, and Governing Factors*. Ann Arbor Press, Chelsea, MI.
7. Nowell, L.H., Capel, P.D., and Dileanis, P.D. (1999). *Pesticides in Stream Sediments and Aquatic Biota: Distributions, Trends, and Governing Factors*. CRC Press, Boca Raton, FL.
8. Freeze, R.A. (2000). *The Environmental Pendulum: A Quest for the Truth About Toxic Chemical, Human Health and Environmental Protection*. University of California Press, Berkeley, CA.
9. Carson, R. (1962). *Silent Spring*. Houghton Mifflin, Boston, MA.
10. Hornsby, A.G., Wauchope, R.D., and Herner, A.E. (1996). *Pesticides Properties in the Environment*. Springer-Verlag, New York.
11. Environmental Protection Agency. (1988). *Pesticides in Groundwater Database 1988 Interim Report*. Environmental Fate and Groundwater Branch, Office of the Pesticide and Toxic Substances, Environmental Protection Agency, Washington, DC.
12. Domagalski, J.L. and Dubrovsky, N.M. (1992). Pesticides residues in groundwater of the San Joaquin Valley, California. *J. Hydrol.* **130**: 299–338.
13. Guenzi, W.D. (Ed.). (1974). *Pesticides in Soil and Water*. Soil Science Society of America, Madison, WI.
14. Garner, W.Y., Honeycutt, R.C., and Nigg, H.N. (1986). *Evaluation of Pesticides in Ground Water*. American Chemical Society, Washington, DC.
15. Giorgini, A. and Zingales, F. (Eds.). (1986). *Agricultural Non-Point Source Pollution: Model Selection and Applications*. Elsevier, New York.
16. Saltzman, S. and Yaron, B. (Eds.). (1986). *Pesticides in Soil*. Van Nostrand Reinhold, New York.
17. Sawhney, B.L. and Brown, K. (Eds.). (1989). *Reactions and Movement of Organic Chemicals in Soils*. SSSA Special Publication No. 22, Soil Science Society of America, Madison, WI.
18. Cheng, H.H. (Ed.). (1990). *Pesticides in the Soil Environment: Processes, Impacts, and Modeling*. Soil Science Society of America, Madison, WI.
19. Hutson, D.H. and Roberts, T.R. (Eds.). (1990). *Environmental Fate of Pesticides*. John Wiley & Sons, New York.
20. Jury, W.A. and Roth, K. (1990). *Transfer Functions and Solute Movement Through Soil: Theory and Applications*. Birkhauser Verlag, Boston, MA.

21. Schnoor, J.L. (Ed.). (1992). *Fate of Pesticides and Chemical in the Environment*. John Wiley & Sons, New York.
22. Gustafson, D.I. (1993). *Pesticides in Drinking Water*. Van Nostrand Reinhold, New York.
23. Honeycutt, R.C. and Schabacker, D.J. (1994). *Mechanisms of Pesticide Movement into Ground Water*. Lewis, Boca Raton, FL.
24. Leng, M.L., Leovery, E.M.K., and Zubkoff, P.L. (Eds.). (1995). *Agrochemical Environmental Fate: State of the Art*. Lewis, Boca Raton, FL.
25. Vighi, M. and Funari, E. (Eds.). (1995). *Pesticide Risk in Groundwater*. Lewis, Boca Raton, FL.
26. Flury, M. (1996). Experimental evidence of transport of pesticides through field soils—a review. *J. Environ. Qual.* **25**: 25–45.
27. Thorsen, M., Feyen, J., and Styczen, M. (1996). Agrochemical modeling. In: *Distributed Hydrological Modelling*. Dondredit, M.B. Abbott and J.C. Refsgaard (Eds.). Kluwer Academic, The Netherlands, pp. 121–141.
28. Jury, W.A. and Horton, R. (2004). *Soil Physics*. 6th Edn. John Wiley & Sons, Hoboken, NJ.
29. Corwin, D.L., Vaughan, P.J., and Loague, K. (1997). Modeling nonpoint source pollutants in the vadose zone with GIS. *Environ. Sci. Technol.* **31**: 2157–2175.
30. Aller, L., Bennett, T., Lehr, J.H., Petty, R.J., and Hackett, G. (1987). DRASTIC: A standardized system for evaluating ground water pollution potential using hydrogeologic settings, EPA-600/2-87-035. U.S. Environmental Protection Agency, Ada, OK.
31. Petach, M.C., Wagenet, R.J., and DeGloria, S.D. (1991). Regional water flow and pesticide leaching using simulations with spatially distributed data. *Geoderma* **48**: 245–269.
32. Carsel, R.F., Smith, C.N., Mulkey, L.A., Dean, J.D., and Howise, P. (1984). User's manual for the pesticide root zone model (PRZM), Release 1. US/EPA-600/3-84-109. U.S. Environmental Protection Agency, Washington, DC.
33. Wagenet, R.J. and Huston, J.L. (1987). LEACHM: A finite-difference model for simulating water, salt, and pesticide movement in the plant root zone. Continuum 2. New York State Water Resources Institute, Cornell University, Ithaca, NY.
34. Davis, F.M. Leonard, R.A., and Knisel, W.G. (1990). Groundwater loading effects of agricultural management systems (GLEAMS). User Manual. USDA-ARS, Southeast Watershed Research Laboratory and University of Georgia, Rifton, GA.
35. Kool, J.B. and van Genuchten, M.Th. (1991). HYDRUS one-dimensional variability saturated flow and transport model, including hysteresis and root uptake. U.S. Salinity Laboratory, USDA-ARS, Riverside, CA.
36. Huston, J.L. and Wagenet, J.R. (1992). LEACHP: A process-based model of water and solute movement, transformations, plant uptake, and chemical reactions in the unsaturated zone (Version 3). Research Series 92-3, Department of Soil, Crop, and Atmospheric Sciences, New York State College of Agriculture and Life Sciences, Cornell University, Ithaca, NY.
37. DeCoursey, D.G., Ahuja, L.R., Hanson, J., Shaffer, M., Nash, R., Rojas, K.W., Hebson, C., Hodges, T., Ma, Q., Johnsen, K.E., and Ghidey, F. (1992). Root zone water quality model, Version 1.0. Technical Documentation. USDA-ARS, Great Plains Systems Research Unit, Fort Collins, CO.
38. Mullins, J.A., Carsel, R.F., Scarbough, J.E., and Ivery, A.M. (1993). PRZM-2, A model for predicting pesticide fate in the crop root and unsaturated soil zones: Users manual for release 2. EPA-600/R-93/046. USEPA Environmental Research Laboratory, Athens, GA.
39. Klein, M. (1993). PELMO—Pesticide leaching model, Version 1.5, User's Manual. Frauenhofer-Institut fur Umweltchemie and Okotoxikolgie, Schmallenberg.
40. Jarvis, N. (1994). The MACRO model (Version 3.1)—Technical description and sample simulations. Monograph, Reports, and Dissertations, 19, Department of Soil Science, Swedish University of Agricultural Science, Uppsala, Sweden.
41. Simunek, J.M, Sejna, M., and Van Genuchten, M.Th. (1996). HYDRUS-2D. Simulating water flow and solute transport in two-dimensional variability saturated media. U.S. Salinity Laboratory, USDA-ARS, Riverside, CA.
42. van den Berg, F. and Boesten, J.J.T.I. (1998). Pesticide leaching and accumulation model (PESTLA). Version 3.3. Description and user's guide, Technical Document 43, DLO-Starting Centrum, Wageningen, The Netherlands.
43. Teso, R.R., Poe, M.P., Younglove, T., and McCool, P.M. (1996). Use of logistic regression and GIS modeling to predict groundwater vulnerability to pesticides. *J. Environ. Qual.* **25**: 425–432.
44. Blanke, J.S. (1999). *Vulnerability to groundwater contamination due to agrochemical use in the San Joaquin Valley, California*. Masters Thesis, Department of Geological and Environmental Sciences, Stanford University, Stanford, CA.
45. Mills, M.B. (2004). *Groundwater vulnerability due to pesticide use in the San Joaquin Valley, California*. Masters Thesis, Earth Systems Program, Stanford University, Stanford, CA.
46. Loague, K., Lloyd, D., Nguyen, A., Davis, S.N., and Abrams, R.H. (1998). A case study simulation of DBCP groundwater contamination in Fresno County, California: 1. Leaching through the unsaturated subsurface. *J. Contam. Hydrol.* **29**: 109–136.
47. Stewart, I.T. and Loague, K. (2003). Development of type transfer functions for regional-scale non-point source groundwater vulnerability assessments. *Water Resour. Res.* **39**(12): 1359, doi:10.1029/2003WR002269.
48. Khan, M.A. and Liang, T. (1989). Mapping pesticide contamination potential. *Environ. Manage.* **13**: 233–242.
49. Loague, K.M., Yost, R.S., Green, R.E., and Liang, T.C. (1989). Uncertainty in a pesticide leaching assessment for Hawaii. *J. Contam. Hydrol.* **4**: 139–161.
50. Loague, K. et al. (1990). Impact of uncertainty in soil, climatic, and chemical information in a pesticide leaching assessment. *J. Contam. Hydrol.* **5**: 171–194; 405.
51. Loague, K. (1991). The impact of land use on estimates of pesticide leaching potential: Assessments and uncertainties. *J. Contam. Hydrol.* **8**: 157–175.
52. Loague, K. (1994). Regional scale ground water vulnerability estimates: Impact of reducing data uncertainties for assessments in Hawaii. *Ground Water* **32**: 605–616.
53. Kleveno, J.J., Loague, K., and Green, R.E. (1992). An evaluation of a pesticide mobility index: Impact of recharge variation and soil profile heterogeneity. *J. Contam. Hydrol.* **11**: 83–99.
54. Yost, R.S., Loague, K., and Green, R.E. (1993). Reducing variance in soil organic carbon estimates: Soil classification and geostatistical approaches. *Geoderma* **57**: 247–262.
55. Giambelluca, T.W., Loague, K., Green, R.E., and Nullet, M.A. (1996). Uncertainty in recharge estimation: Impact on groundwater vulnerability assessments for the Pearl Harbor Basin, Oahu, Hawaii. *J. Contam. Hydrol.* **23**: 85–112.

56. Li, Z.C., Yost, R.S., and Green, R.E. (1998). Incorporating uncertainty in a chemical leaching assessment. *J. Contam. Hydrol.* **29**: 285–299.
57. Loague, K.M., Green, R.E., Liu, C.C.K., and Liang, T.C. (1989). Simulation of organic chemical movement in Hawaii soils with PRZM: 1. Preliminary results for EDB. *Pac. Sci.* **43**: 67–95.
58. Loague, K. et al. (1989). Simulation of organic chemical movement in Hawaii soils with PRZM: 2. Predicting deep penetration of DBCP, EDB, and TCP. *Pac. Sci.* **43**: 362–383.
59. Loague, K. (1992). Simulation of organic chemical movement in Hawaii soils with PRZM: 3. Calibration. *Pac. Sci.* **46**: 353–373.
60. Loague, K. et al. (1995). Chemical leaching near the Waiawa Shaft, Oahu, Hawaii: 2. Modeling results. *Ground Water* **33**: 124–138.
61. Diaz-Diaz, R., Garcia-Hernandez, J.E., and Loague, K. (1998). Leaching potentials of four pesticides used for bananas in the Canary Islands. *J. Environ. Qual.* **27**: 562–572.
62. Diaz-Diaz, R. and Loague, K. (2000). Regional-scale leaching assessments for Tenerife: Impact of data uncertainties. *J. Environ. Qual.* **29**: 835–847.
63. Diaz-Diaz, R., Loague, K., and Notario, J.S. (1999). An assessment of agrochemical leaching potentials for Tenerife. *J. Contam. Hydrol.* **36**: 1–30.
64. Diaz- Diaz, R. and Loague, K. (2001). Assessing the potential for pesticide leaching for the pine forest areas of Tenerife. *Environ. Toxicol. Chem.* **20**: 1958–1967.
65. Rao, P.S.C., Hornsby, A.G., and Jessup, R.E. (1985). Indices for ranking the potential for pesticide contamination of groundwater. *Soil Crop Sci. Soc. Fl. Proc.* **44**: 1–8.
66. Stewart, I.T. and Loague, K. (2004). Assessing San Joaquin Valley, California, groundwater vulnerability with the type transfer function model. *J. Environ. Qual.* **33**: 1487–1498.
67. Bacci, E. (1994). *Ecotoxicology of Organic Contaminants*. Lewis, New York.
68. Loague, K. and Corwin, D.L. (1996). Uncertainty in regional-scale assessments of non-point source pollutants. In: *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone*. D.L. Corwin and K. Loague (Eds.). Special Publication 48, Soil Science Society of America, Madison, WI, pp. 131–152.
69. Loague, K. et al. (1998). A case study simulation of DBCP groundwater contamination in Fresno County, California: 2. Transport in the saturated subsurface. *J. Contam. Hydrol.* **29**: 137–163.
70. McDonald, M.G. and Harbaugh, A.W. (1988). A modular three-dimensional finite-difference groundwater flow model. Scientific Software Group, Washington, DC.
71. Zheng, C. (1992). MT3D: A modular three-dimensional model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems (Version 1.5). S.S. Papadopulos and Associates, Bethesda, MD.
72. Soutter, L.A. and Loague, K. (2000). Revisiting the Fresno DBCP case study simulations: The impact of upscaling. *J. Environ. Qual.* **29**: 1794–1805.
73. Loague, K. and Abrams, R.H. (1999). DBCP contaminated groundwater: Hot spots and non-point sources. *J. Environ. Qual.* **28**: 429–445.
74. Gilliom, R.J. (2001). Pesticides in the hydrologic system—what do we know and what's next? *Hydrological Processes* **15**: 3197–3201.
75. Graffy, E.A. (1998). Low-level detection of pesticides... so what? *J. Soil Water Conserv.* **53**: 11–12.

HIGH pH GROUNDWATER—THE EFFECT OF THE DISSOLUTION OF HARDENED CEMENT PASTES

Caijun Shi
CJS Technology, Inc.
Burlington, Ontario, Canada

Roger Spence
Oak Ridge National Laboratory
Oak Ridge, Tennessee

INTRODUCTION

Concrete is the most widely used construction material manufactured by mixing together cement, aggregates, water, air, and various admixtures. Conventional hardened concrete comprises approximately 30% by volume of cement paste phase and 18% pore voids. Many concrete structures are built above and below ground, some in contact with surface and/or groundwater. In these cases, pore voids within concrete are filled with water. Groundwater chemistry is affected when brought into contact with or close proximity to such structures since the hydration products of cement can be dissolved and leached into the groundwater.

In fact, groundwater chemistry close to aboveground structures is affected by the rainwater falling on these structures prior to soaking into the ground and becoming groundwater. The composition of groundwater varies widely, depending on the soil or rock with which the water is in equilibrium. Rainwater can be acidic ("acid rain"), which may be quickly consumed. The local soil, rocks, and minerals result in dissolution of some ions into the resulting runoff. For example, acid rain leached aluminum into lakes in the northern United States, killing the local aquatic life in those lakes. Acidic mine runoff is another source for acid groundwater. On the other hand, limestone strata result in hard water with higher pH and seawater results in higher ionic concentrations from dissolved salts. It is well known that some ions in groundwater can migrate into concrete and even deteriorate the integrity of the concrete structures.

The focus here is on how the hydration products of cements affect the pH of groundwater since they have very high equilibrium pH values. To fully evaluate this effect requires knowledge of cement chemistry, the characteristics of hydration products, and the transport of ions in hardened concrete. We first summarize cement chemistry, the effect of supplementary cementing materials on this chemistry, and the dissolution and leaching of hydration products of cement. The last section discusses the interaction with cement and its leached ions. For a given case, the known cement chemistry can be used to project what to expect for the local groundwater in question.

CHARACTERIZATION AND HYDRATION CHEMISTRY OF PORTLAND CEMENT

Portland cement is the most used cementing material in construction. It is a hydraulic cement produced by pulverizing clinker and calcium sulfate (usually gypsum—$CaSO_4 \cdot 2H_2O$) as an interground addition. Cement clinker consists mainly of tricalcium silicate ($3CaO \cdot SiO_2$—C_3S), dicalcium silicate ($2CaO \cdot SiO_2$—C_2S), tricalcium aluminate ($3CaO \cdot Al_2O_3$—C_3A), and tetracalcium aluminoferrite ($4CaO \cdot Al_2O_3 \cdot Fe_2O_3$—$C_4AF$). Different types of Portland cement are manufactured to meet various normal physical and chemical requirements for specific purposes. ASTM C150 specifies five types of common Portland cement produced by adjusting the proportions of their minerals and finenesses.

In the presence of water, C_3S and C_2S in cement hydrate to form calcium silicate hydrate gel (C-S-H gel) and $Ca(OH)_2$. In the presence of calcium sulfate, C_3A hydrates to form calcium trisulfoaluminate hydrate ($3CaO \cdot Al_2O_3 \cdot 3CaSO_4 \cdot 32H_2O$—AFt or ettringite), or calcium monosulfoaluminate hydrate ($3CaO \cdot Al_2O_3 \cdot CaSO_4 \cdot 12H_2O$—AFm or monosulfate). In the absence of calcium sulfate, C_3A reacts with water and calcium hydroxide to form tetracalcium aluminate hydrate [$3CaO \cdot Al_2O_3 \cdot Ca(OH)_2 \cdot 12H_2O$]. C_4AF reacts with water to form calcium aluminoferrite hydrates ($6CaO \cdot Al_2O_3 \cdot Fe_2O_3 \cdot 12H_2O$). These hydration reactions can be expressed as follows:

$$3(3CaO \cdot SiO_2) + 6H_2O \rightarrow 3CaO \cdot 2SiO_2 \cdot 3H_2O + 3Ca(OH)_2 \quad (1)$$

$$2(2CaO \cdot SiO_2) + 4H_2O \rightarrow 3CaO \cdot 2SiO_2 \cdot 3H_2O + Ca(OH)_2 \quad (2)$$

$$3CaO \cdot Al_2O_3 + 3CaSO_4 \cdot 2H_2O + 26H_2O \rightarrow 3CaO \cdot Al_2O_3 \cdot 3CaSO_4 \cdot 32H_2O \quad (3)$$

$$3CaO \cdot Al_2O_3 \cdot 3CaSO_4 \cdot 32H_2O + 2(3CaO \cdot Al_2O_3) + 4H_2O \rightarrow 3(3CaO \cdot Al_2O_3 \cdot CaSO_4 \cdot 12H_2O) \quad (4)$$

$$3CaO \cdot Al_2O_3 + 12H_2O + Ca(OH)_2 \rightarrow 3CaO \cdot Al_2O_3 \cdot Ca(OH)_2 \cdot 12H_2O \quad (5)$$

$$4CaO \cdot Al_2O_3 \cdot Fe_2O_3 + 10H_2O + 2Ca(OH)_2 \rightarrow 6CaO \cdot Al_2O_3 \cdot Fe_2O_3 \cdot 12H_2O \quad (6)$$

A hardened cement paste is a heterogeneous multiphase system. At room temperature, a fully hydrated Portland cement paste consists of 50–60% C-S-H gel, 20–25% $Ca(OH)_2$, 15–20% ettringite (or AFt) and AFm by volume. The minor hydration products, such as $Ca(OH)_2$, $3CaO \cdot Al_2O_3 \cdot 6H_2O$, and AFt, form in small quantities depending on the composition of the cementing material and hydration conditions.

Portland cement also contains a small amount of alkalis; but the amount is dependent on the raw materials and production process used. The presence of a small amount of alkalis does not have a significant effect on the hydration process of the cement but may have a significant effect on the pH of the pore solution. Usually, the pH value of the pore solution of a hardened Portland cement is over 12.5 to values as high as 13.5, due to the presence of alkalis.

USE OF SUPPLEMENTARY CEMENTING MATERIALS AS CEMENT REPLACEMENTS

For construction use, supplementary cementing materials, such as granulated/pelletized blast furnace slag, coal fly ash, volcanic ashes, condensed silica fume, rice husk ash, and natural pozzolans, are often used to replace Portland cement to reduce the cost or to improve the performance of concrete.

When Portland pozzolan cement contacts water, Portland cement hydrates first; the lime released from the hydration of Portland cement reacts with pozzolan to form new products depending on the composition of the pozzolan and hydration conditions. The main hydration product in lime–pozzolan mixtures is C-S-H with a C/S ratio less than 1.5, depending on the local concentration of reactants. The aluminate in the pozzolan may yield a variety of hydrates: calcium aluminate hydrate (C_4AH_{19}), gehlenite hydrate (C_2ASH_8), AFt ($C_3A \cdot 3CaSO_4 \cdot 32H_2O$), and AFm ($C_3A \cdot CaSO_4 \cdot 12H_2O$). The pozzolanic reactions can be generally expressed as follows:

$$Ca(OH)_2 + SiO_2 + (n-1)H_2O \rightarrow xCaO \cdot SiO_2 \cdot nH_2O \quad (7)$$

$$Ca(OH)_2 + Al_2O_3 + (n-1)H_2O \rightarrow xCaO \cdot Al_2O_3 \cdot nH_2O \quad (8)$$

$$Ca(OH)_2 + Al_2O_3 + SiO_2 + (n-1)H_2O \rightarrow 3CaO \cdot Al_2O_3 \cdot 2SiO_2 \cdot nH_2O \quad (9)$$

$$(1.5\text{–}2.0)CaO \cdot SiO_2 \cdot aq + SiO_2 \rightarrow (0.8\text{–}1.5)CaO \cdot SiO_2 \cdot aq \quad (10)$$

The relationship between the C/S ratio of C-S-H and equilibrium pH is shown in Fig. 1. There are two plateaus for the equilibrium pH. One corresponds to around pH 12, where the C/S ratio is greater than 1.0. The other plateau corresponds to pH 10, where the C/S ratio of C-S-H varies from 0.05 to 0.6. Typically, the C/S ratio of C-S-H from the hydration of Portland cement ranges from 1.4 to 1.7.

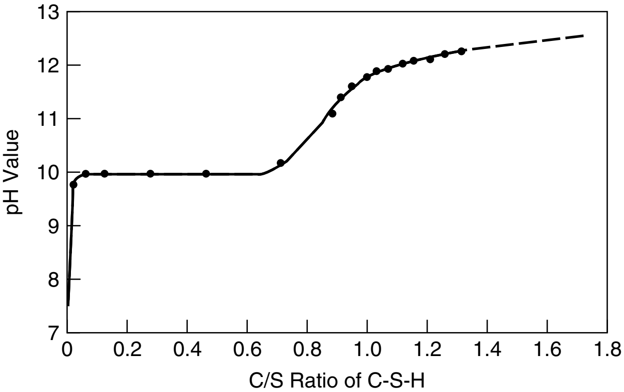

Figure 1. Relationship between C/S ratio of C-S-H and equilibrium pH. (Based on Reference 1).

DISSOLUTION AND LEACHING OF HYDRATION PRODUCTS OF CEMENTS

Table 1 lists the solubility products of common hydration products of hydraulic cements. It can be seen that $Ca(OH)_2$ has the highest solubility among all hydration products. The solubility of C-S-H decreases as its C/S ratio decreases, which means that the use of supplementary cementing materials converts the more soluble $Ca(OH)_2$ into the less soluble C-S-H.

The dissolution of $Ca(OH)_2$ and the decalcification of the C-S-H phase occur so as to maintain the chemical equilibrium of Ca^{2+} between the solid hydrates and the pore solution. Once the concentration of Ca^{2+} in the pore solution of a cement-based material is decreased by diffusion, the dissolution of $Ca(OH)_2$ starts, depleting the $Ca(OH)_2$, followed by the decalcification of the C-S-H phase. The dissolution of $Ca(OH)_2$ and the partial decalcification of C-S-H are given by the following chemical equilibrium equations:

$$Ca(OH)_2 \rightarrow Ca^{2+} + 2OH^- \qquad (11)$$

$$xCaO \cdot SiO_2 \cdot yH_2O \rightarrow (x-m)CaO \cdot (1-n)SiO_2 \cdot (y-m)H_2O + mCa(OH)_2 + nSiO_2 \qquad (12)$$

Chemical equilibrium of the solid cement hydrates and the pore solution with regard to Ca^{2+} is applied according to the model shown in Fig. 2. The initial concentration of Ca^{2+} in the pore solution is approximately 0.02 mol/L of solution and must be maintained as long as the $Ca(OH)_2$ remains in the solid. Subsequently, after all the $Ca(OH)_2$ is dissolved, the decalcification of C-S-H will ensue depending on the concentration of Ca^{2+} with the chemical equilibrium shown in the curved portion of Fig. 2.

Table 1. Solubility Product of Common Hydration Products of Cements

Compound	Solubility Product	Reference
$Ca(OH)_2$	5.5×10^{-6}	2
C-S-H[a]	$1.0 \times 10^{-24 \text{ to } -14}$	3
C_4AH_{13}	3.2×10^{-28}	2
C_2ASH_8	8.1×10^{-52}	4
AFt	1.0×10^{-40}	5
AFm	1.7×10^{-28}	5

[a] The solubility product increases as the C/S ratio of C-S-H decreases.

Table 2. pH Value and Concentrations of Ions in Immersion Water at 12 Weeks

		Ion Concentration (mmol/L)				
W/C	pH	Ca^{2+}	K^+	Na^+	$SO_4{}^{2-}$	Al^{3+}
0.4	12.8	400	110	38	1.9	2.4
0.55	13.0	710	89	33	2.4	2.5
0.7	13.0	760	38	19	2.6	2.6
0.85	13.0	840	32	12	2.7	2.8

Source: Reference 7.

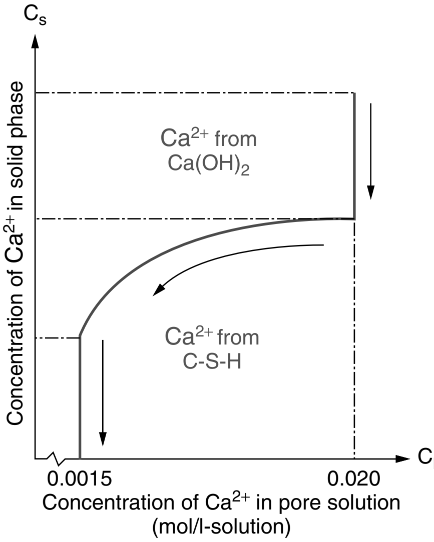

Figure 2. Relationship between Ca^{2+} concentrations in solid and liquid phases (6).

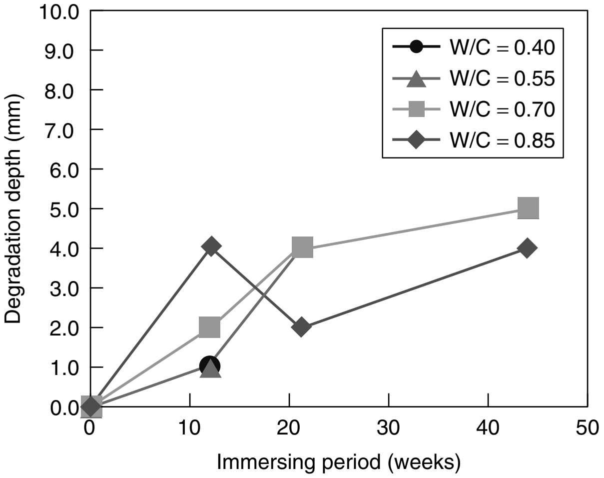

Figure 3. Degraded depths of cement pastes after immersion in water. (Based on Reference 7).

Kosuke Yokozeki et al. (7) investigated the leaching out of ingredients from hardened cement pastes made with different water to cement ratios. They immersed 20 × 20 × 20 mm hardened cement cubes in water with a water to cube (by volume) ratio of 20, then measured the degraded depths of cement paste, as shown in Fig. 3, and pH values and ion concentrations of the immersed water after 12 weeks of immersion, as shown in Table 2.

The results in Fig. 3 indicate that the degradation depth increases with immersion time, but the water to cement ratio does not show an obvious effect. The pH value of immersed water is almost 13 regardless of the water to cement ratio. The Ca^{2+}, K^+, and Na^+ concentrations are high, and the rest are low.

Several researchers (6,8–10) Sugiyama et al. confirmed that hardened cement pastes dissolve and degrade in the following five steps: (1) leaching of calcium hydroxide; (2) decalcification of C-S-H; (3) formation of C-S-H with

C/S ratio or calcium aluminate hydrates; (4) leaching of C-S-H; and (5) leaching of silica gel. As shown in Eq. 11, the initial dissolution and leaching of Ca^{2+} are always associated with hydroxyl ion. Laboratory results clearly indicate a gradual pH drop on the surface of concrete after immersion in water or a solution (11). Once it is controlled by the leaching of C-S-H, the pH of hardened concrete is around 10, as indicated in Fig. 1.

EFFECT OF LEACHING OF HYDRATION PRODUCTS OF CEMENTS ON pH OF GROUNDWATER

In general, the water chemistry is controlled by two large bulk systems: concrete and soil-rock. Both bulk systems are porous with an interfacial transitional zone. Cement chemistry dominates within the bulk of concrete structure and local groundwater chemistry dominates within the bulk of the soil-rock.

As discussed above, the pH of pore solution within concrete is over 12.5 to values as high as 13.5, due to the presence of alkalis, which is significantly higher than groundwater. Hydroxyl ions always transport from concrete to groundwater. Increased pH of groundwater can be expected around the surface of concrete but is not expected to penetrate far into the groundwater matrix, because the transport rate from the concrete is low. A transition zone in the aqueous chemistry exists at the interface between these two bulk systems, as illustrated in Fig. 4.

The width of this transitional zone is controlled mainly by the rate of flow of the groundwater around the surface of the concrete, with higher flows resulting in a narrower zone. For example, a fast-flowing river around a concrete structure makes this zone very thin, or even negligible, sweeping the surface of the concrete clean of ions leached from the concrete and turbulently mixing these ions with the river water. The width of the transitional zone can be significant for concrete sitting in quiescent perched water. Diffusion is very slow, so much so that leaching from a monolithic concrete structure can be dominated by the permeation of water through the porous concrete matrix.

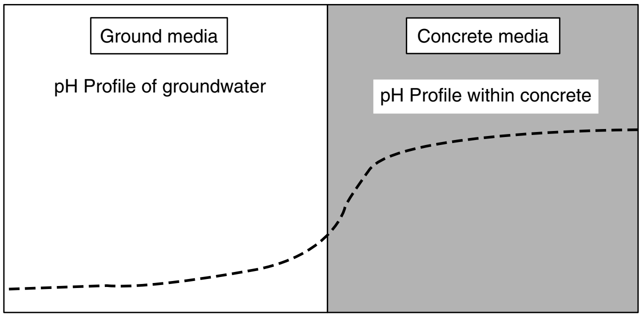

Figure 4. Schematic illustration of pH profiles between groundwater and concrete.

BIBLIOGRAPHY

1. Buil, M., Revertegat, E., and Oliver, J. (1992). A model of the attack of pure water or undersaturated lime solutions on cement. In: *Stabilization and Solidification of Hazardous, Radioactive, and Mixed Wastes*, Vol. 2, STP 1123. T.M. Gilliam and C.C. Wiles (Eds.). ASTM, Philadelphia, PH, pp. 227–241.
2. Reardon, E.J. (1990). An ion interaction model for the determination of chemical equilibria in cement/water system. *Cement Concrete Res.* **20**(2): 175–192.
3. Glasser, F.P., Lachowski, E.E., and Macphee, D.E. (1987). Compositional model for calcium silicate hydrate (C-S-H) gels, their solubilities, and free energies of formation. *J. Am. Ceramic Soc.* **70**(7): 481–485.
4. Atkins, M., Glasser, F.P., and Kindness, A. (1991). *A Thermodynamic Model for Blended Cements*. Research Report for Department of Environment, DoE/HMIP/RR/92/005.
5. Day, R.L. (1992). *The Effect of Secondary Ettringite Formation on the Durability of Concrete: A Literature Review*. Department of Civil Engineering, Research Rep. No. CE 92-2, University of Calgary.
6. Sugiyama, T., Ritthichauy, W., and Tsuji, Y. (2003). Simultaneous transport of chloride and calcium ions in hydrated cement systems. *Adv. Concrete Technol.* **1**(2): 127–138.
7. Yokozeki, K., Watanabe, K., Sakata, N., and Otsuki, N. (2003). Prediction of changes in physical properties due to leaching of hydration products from concrete. *J. Adv. Concrete Technol.* **1**(2): 161–171.
8. Furusawa, Y. (1997). Research trends on degradation of concrete caused by calcium leaching and modeling. *Concrete J. JCI* **35**(12): 29–32 (in Japanese).
9. Watanabe, K., Yokozeki, K., Furusawa, Y., Otsuki, N., and Daimon, M. (2000). Mechanisms of deterioration of cement composites due to leaching of calcium. In: *CANMET/ACI Durability of Concrete*. SP-192-70, Vol.2, pp. 1155–1172.
10. Yokozeki, K. et al. (2002). Analysis of old structures and numerical model for degradation of concrete by calcium ion leaching. *Concrete Lib. JSCE* **40**: 209–226.
11. Sergi, G., Yu, S.W., and Page, C.L. (1992). Diffusion of chloride and hydroxyl ions in cementitious materials exposed to a saline environment. *Mag. Concrete Res.* **44**(158): 63–69.

PHYTOEXTRACTION AND PHYTOSTABILIZATION: TECHNICAL, ECONOMIC AND REGULATORY CONSIDERATIONS OF THE SOIL-LEAD ISSUE

Tracey Johnston
Rupali Datta
Dibyendu Sarkar
University of Texas at San Antonio
San Antonio, Texas

Lead is a common environmental contaminant, produced from multiple anthropogenic sources. The cost of traditional remediation techniques for lead-contaminated soils is prohibitive. As a result, new technologies such as phytoextraction, the removal of contaminants from soil and water sources by plants, and phytostabilization, the immobilization of contaminants in the soil, are being examined as remediation alternatives for lead-contaminated soils. Only a few plant species have been identified so far that can tolerate and accumulate lead partly because lead is relatively immobile in soils and has low phytoavailability. However, lowering the pH or the use of soil amendments or chelating agents such as EDTA can enhance the uptake of lead by plants. Lead is already relatively immobile in most soils, so phytostabilization is a promising remediation strategy. However, phytostabilization leaves the contaminant in the soil (albeit in an insoluble form), so long-term monitoring is required to verify that the contaminant does not become soluble again. Although only limited studies on the economic viability of phytoextraction and phytostabilization have been done, the estimated costs are significantly lower than those of traditional methods.

INTRODUCTION

Lead is a bluish-gray heavy metal, which is usually found naturally in soils in levels ranging from 15 to 40 mg/kg. Galena (PbS) is the major mineral source of naturally occurring lead in soils. More than 85% of the lead mined in the United States comes from southeastern Missouri (1). It is common to find lead in soils at levels between 100 and 3000 mg/kg due to the use of lead-based paints, storage batteries, plumbing fixtures, gasoline, explosives, ammunition, and mining and smelting (2,3). According to the Toxic Release Inventory, almost 144 million pounds of lead were released to land and water in the United States between 1987 and 1993, most of which were from lead and copper smelting (1). Lead contamination in soils can lead to many environmental problems, such as lead toxicity in humans, animals, and plants; loss of vegetative cover; and ground water contamination (4). In humans, short-term exposure to lead can interfere with red blood cell activity, cause mental and physical developmental delays in babies and young children, and increase the blood pressure of adults. Long-term exposure to lead has the potential to cause cancer, kidney disease, and stroke in adults (1). Current remediation techniques for lead-contaminated soils include excavation, stabilization, and/or removal to a landfill (3). These techniques are expensive and can result in secondary problems such as human exposure and habitat destruction (3).

PHYTOREMEDIATION

Phytoremediation is the use of plants to remediate contaminated soils, sediments, sludge, ground/surface water, or wastewater (5). Phytoremediation uses a variety of plant biological processes to degrade, remove, or immobilize contaminants (3). Phytoremediation can be an attractive alternative to traditional remediation techniques because it is driven by photosynthesis and is performed on site, reducing costs and human exposure (1). Two specific methods of phytoremediation, phytoextraction and phytostabilization have been used to treat lead-contaminated soils (5,6).

Phytoextraction

Phytoextraction is the uptake of contaminants by a plant, and the toxic substance is then sequestered in the tissues (5). The plants that take up contaminants at the highest rates are called hyperaccumulators (0.1 to 3% of dry biomass) (7). *Brassica juncea* (Indian Mustard), *Helianthus annuus* (common sunflower), *Nicotiana tabacum* (tobacco), and *Zea mays* (corn) have been identified in the literature as hyperaccumulators of lead (2,8). For lead to be taken up by a plant, it must be in an aqueous phase, either in soil solution or in groundwater (9,10). The solubility of lead (in the form of Pb^{2+}) is low at neutral or alkaline pH, so only about 0.1% of the lead in lead-contaminated soils is available for uptake by plants (3). Kumar et al. (2) reported that the ratio of lead in the dry weight of plant to that in soil (called the phytoextraction coefficient) in *B. juncea* was 1.7; in comparison, Cr^{6+} and Cd^{2+} have coefficients over 50. Decreasing the pH of the soil can increase the phytoavailability of lead; however, most plants have an ideal pH range of 5–8 (3,11). Chelating agents, such as EDTA (ethylenediaminetetraacetic acid), increase lead phytoavailability and also increase the amount of lead accumulated in the aboveground biomass (12). However, by making lead more soluble, the danger exists of allowing more lead to enter the food chain or leach into groundwater (13).

Phytostabilization

In phytostabilization, vegetation is used to modify the biological, physical, and chemical properties of soil to reduce the movement of contaminants. *Brassica juncea, Festuca rubra* cv Merlin (red fescue), and *Agrostis tenuis* cv Goginan (colonial bentgrass) have been used to stabilize lead mine wastes (14). One phytostabilization study showed that lead-contaminated soil had 740 mg/kg of lead in the leachate without plant cover, compared with 22 mg/kg with *B. juncea* cover (5). Phytostabilization can be achieved through root absorption, accumulation, and adsorption. Plants may also release exudates in the root zone that could result in precipitation, complexation, or a change in the valence state of a metal contaminant (15). Roots can help precipitate lead in the form of lead phosphate (5). Some grasses, such as ryegrass, have been investigated as possible phytostabilizers due to their extensive roots and widespread growth patterns (14). In addition, plants can prevent wind and water erosion from spreading contaminants to surrounding areas. However, because contaminants are left in the soil in an insoluble form in phytostabilization, the area has to be monitored continuously to verify that they do not become soluble with time (16).

TECHNICAL CONSIDERATIONS

Vegetation

Several plants reportedly accumulate small amounts of lead, but hyperaccumulators are those plants that can take up the metal to tissue concentrations greater than 1000 mg/g (17). For effective phytoextraction, the plant should also be able to translocate lead to aboveground tissues, which allows easier harvesting and helps the next crop of plants to grow quickly due to an already existing root system (16). Depth of roots is an important characteristic of candidate plants for phytoremediation. Most hyperaccumulator plants have shallow root systems; the effective depth is about 12 inches and hence can only take up contaminants from shallow soils (6). Hyperaccumulating plants that have faster growth rates, higher biomass, and an extensive root system would be more effective for phytoremediation (18). Plants used in phytostabilization have to grow and survive long term in contaminated soils. A large root mass would also be advantageous for greater soil–root contact surface area, which would also result in the secretion of larger amounts of enzymes or exudates to immobilize lead. Prospective plants should also be hearty and tolerate the environmental stresses in a given contaminated area (16).

Soil and Lead Chemistry

Soil pH is very important in the solubility of metals such as lead. The solubility of lead is relatively low at neutral or alkaline pH (3). Decreasing the pH of the soil by using soil amendments or acidifiers can increase the solubility and hence the phytoavailability of lead. However, as the pH decreases, aluminum (Al^{3+}), which is ubiquitous in soils and toxic to plants, also becomes more phytoavailable. The optimum pH range for most plants is 5–8, so this would make the optimum pH range for lead uptake between 5 and 6.5 (11). Chelating agents, such as EDTA, have been used to increase solubility and thereby increase plant uptake (13,19). Chelating agents can also increase the translocation of lead into the aboveground biomass (19). On the other hand, by increasing the solubility of lead, there exists the danger of lead entering the food chain or leaching into underground water sources. Strategies for properly using chelating agents include using plants that rapidly take up lead and harvesting them immediately to reduce environmental risks (12).

Soil amendments can also be used to decrease solubility. Increasing the pH of soil via alkalizing agents or mineral oxides increases lead precipitation potential. Organic matter and biosolids have also been used to immobilize lead in soil (20). Phosphates can bind with lead and precipitate it from solution, usually in the form of lead carbonates, hydroxides, and phosphates (21). Therefore, when adding fertilizers to plants involved in phytoextraction, phosphate should be added typically to leaves and not to the soil itself (12). The type of soil and the amount of organic matter present can complicate the uptake of lead. Contaminated areas generally also have multiple contaminants, and there are substantial variations in the concentrations of contaminants at a single site. These scenarios can reduce the effectiveness of any phytoremediation project (16).

Other Considerations

Phytoextraction and phytostabilization use natural plant processes, so it takes longer to remediate contaminated areas using these techniques, compared with traditional methods (5,15). The length of the growing season and rainfall patterns in a contaminated area will also substantially impact the amount of time required for any phytoremediation project. Using plants that have faster growth and high biomass may reduce the time necessary for phytoextraction compared to hyperaccumulators, the majority of which are slow growing and low biomass plants (16). In phytostabilization, selecting plants that produce more exudates or enzymes can speed up the remediation process. Any phytoremediation project requires continuous monitoring of the site. However, a phytostabilization project needs long-term monitoring because of the lead that stays on site. On-site plant tissue sampling to ensure little or no lead uptake and soil and local source water sampling, to ensure that lead remains immobilized may be required for many years (16). Although some herbivorous animals avoid plants with high levels of accumulated metals, it may be necessary to restrict access to hyperaccumulating plants during phytoextraction through fencing, netting, or harvesting before the plants reach the flowering stage to avoid transfer of metals to the food chain (22).

ECONOMIC CONSIDERATIONS

To date, economic viability data for phytoextraction and phytostabilization projects are rather limited. Moreover, the available data are typically gathered from small, pilot-scale studies, which may not accurately reflect the costs and efficiency of large-scale operations (16). The costs of any potential phytoremediation project must consider the initial feasibility studies, which includes site characterization, selecting the most appropriate plants, and determining if soil amendments are necessary (16). Before remediation can begin, site start-up costs may include debris removal, fencing, irrigation systems, and drainage control measures (23). Besides the costs of planting, watering, fertilizing, and pesticide use, project costs must also include periodic monitoring for pH, nutrient levels, and contaminant levels in soils and possibly nearby surface and groundwater sources and disposal of contaminated vegetation (16). Still, estimated costs for most phytoremediation applications are usually 50–80% lower than those of traditional techniques (24). A small-scale study of lead phytoextraction in 1997 predicted that the costs of the phytoextraction project at the Magic Marker site in New Jersey would range from 50–65% of the remediation costs of traditional methods (19). In 1997, another study estimated that the costs for phytoremediating a lead-contaminated area of soil 60 cm deep would be $\$6/m^2$ compared to $\$15–\$730/m^2$ for more traditional remediation techniques (20). One of the disadvantages of using phytoremedation as a cleanup technique is that it can easily take several growing seasons to reduce the lead concentrations to acceptable levels. The cost incurred by not using the land during the period of the phytoremediation project may be a deterrent for a business expecting to make a profit from that land immediately (25). The estimated costs of a typical phytostabilization project may range from $200 to $10,000 per hectare (1 meter root depth) (7). However, the contaminants essentially stay in place in this method, so the costs of long-term monitoring (to verify that the lead has not become soluble again) should be included in any project (7). The number of cost estimates for actual phytoextraction and phytostabilization projects is limited, but there is general agreement that the costs are much lower than those of currently practiced traditional *ex situ* remediation techniques. For phytoextraction, the future recovery and eventual resale of extracted metals has been proposed; however, it is unclear if the technology and markets exist to make such an idea feasible and profitable on a large scale (16).

REGULATORY CONSIDERATIONS

Specific federal regulations for phytoremediation have not yet been developed, but several existing federal acts, statutes, and programs have implications for phytoremediation technologies (26) The Resource and Recovery Act (RCRA) of 1976, an amendment to the Solid Waste Disposal Act, deals with waste management. Two of the sections (Subtitles C - Hazardous Waste Management and D - Solid Waste Management) are most likely to impact phytoremediation projects. Remediation of treatment, storage, or disposal units will fall under the RCRA requirements. These requirements, however, are performance-based and are delegated to the states. It is up to the state authorities to assess the individual site requirements to determine if phytoremediation can be part of a site cleanup. Disposal of plants used in phytoextraction may also be governed by RCRA. The Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) was enacted in 1980 to cleanup hazardous waste contaminated sites. The trust fund created by CERCLA provides money for the cleanup of designated hazardous waste sites. Under CERCLA, a feasibility study must be completed to determine if phytoremediation is appropriate for a specific site. Any remediation plan must meet the nine criteria set forth in the National Oil and Hazardous Substances Pollution Contingency Plan (NCP) (27). States may have stricter guidelines for cleanup requirements and remaining material limitations. Any state or federal regulation concerning how much material may be left on site may have important implications for phytostabilization. In addition, the Maximum Contamination Levels (MCLs) put forth by the Safe Drinking Water Act (SDWA) may also apply. Executive Order 13112 signed in 1999 prohibits the use of invasive species unless it can be shown that the benefits of using such species outweigh the potential harm (28). Other regulations and acts that may apply to phytoremediation include the Clean Air Act, Department of Agriculture statutes, and the Toxic Substances Control Act (26).

FUTURE PROSPECTS FOR LEAD PHYTOREMEDIATION

Lead phytoextraction seems to be limited to low to medium lead soil contamination levels and to the root zone of the plants being used, so phytoextraction may emerge as a common polishing step in a multiple step process aimed at total site remediation. For some cleanup projects, traditional remediation techniques could be used to reduce the contamination to plant-tolerant levels followed by phytoextraction to reduce the contamination even further (23). Similarly, appropriate plants can be used to stabilize the remaining pollutants to reduce their bioavailability and potential for leaching as the final remediation step (20). Lead is relatively insoluble in a soil solution, so phytostabilization might be a more promising technique for lead-contaminated soils (2,23). The search for fast-growing and high biomass phytoextracting and phytostabilizing accumulator plants is ongoing (29). Selective breeding and genetic engineering of more suitable plants is a possibility. However, more information on the ecological impact of using genetically modified plants would be needed. The current regulations and the societal view of using genetically modified organisms could also hinder the use of these technologies (23).

BIBLIOGRAPHY

1. U.S. EPA. (2002). *Lead in Drinking Water Regulation: Public Education Guidance*. EPA/816/R-02/010. Office of Water, Washington, DC.
2. Kumar, N.P.B.A., Dushenkov, V., Motto, H., and Raskin, I. (1995). Phytoextraction: The use of heavy metals to remove heavy metals from soils. *Environ. Sci. Technol.* **29**: 1232–1238.
3. Huang, J.W. and Cunningham, S.D. (1996). Lead phytoextraction: Species variation in lead uptake and translocation. *New Phytol.* **134**: 75–84.
4. Todd, A.C. et al. (1996). Unraveling the toxicity of lead: An essential priority for environmental health. *Environ. Health Perspect*: 104–146.
5. Salt, D.E. et al. (1995). Phytoremediation: A novel strategy for the removal of toxic metals from the environment using plants. *Biotechnology* **13**: 468–474.
6. Blaylock, M.J., Elless, M.P., Huang, J.W., and Dushenkov, S.M. (1999). Phytoremediation of lead-contaminated soil at a New Jersey brownfield site. *Remediation* **9**: 93–101.
7. Cunningham, S.D., Berti, W.R., and Huang, J.W. (1995). Phytoremediation of contaminated soils. *TIBTECH* **13**: 393–397.
8. Brooks, R.R. (Ed.). (1998). Geobotany and hyperaccumulators. In: *Plants That Hyperaccumulate Heavy Metals*. CAB International, New York.
9. Paterson, S., Mackay, D., Tam, D., and Shiu, W.Y. (1990). Uptake of organic chemicals by plants: A review of processes, correlations and models. *Chemosphere* **21**: 297–331.
10. Shimp, J.F. et al. (1993). Beneficial effects of plants in the remediation of soil and groundwater contaminated with organic materials. *Crit. Rev. Environ. Sci. Technol.* **23**: 41–77.
11. Chlopecka, A., Bacon, J.R., Wilson, M.J., and Kay, J. (1996). Forms of cadmium, lead, and zinc in contaminated soils from southwest Poland. *J. Environ. Qual.* **25**: 69–79.
12. Huang, J.W., Chen, J., Berti, W.R., and Cunningham, S.D. (1997). Phytoremediation of lead contaminated soils: Role

of synthetic chelates in lead phytoextraction. *Environ. Sci. Technol.* **31**: 800–805.

13. Wu, J., Hsu, F.C., and Cunningham, S.D. (1999). Chelate-assisted Pb phytoextraction: Pb availability, uptake, and translocation constraints. *Environ. Sci. Technol.* **33**(11): 1898–1904.
14. Smith, R.A.H., and Bradshaw, A.D. (1979). The use of metal tolerant plant populations for the reclamation of metalliferous wastes. *J. Appl. Ecol.* **16**: 595–612.
15. Ernst, W.H.O. (1996). Availability of heavy metals and decontamination of soils by plants. *Appl. Geochem. - J. Int.* **11**: 163.
16. Pivetz, B.E., and U.S. EPA. (2001). *Ground water issue: Phytoremediation of contaminated soil and ground water at hazardous waste sites*. EPA/540/S-01/500. Office of Research and Development, Washington, DC, February 2001.
17. Brooks, R.R. (Ed.). (1998). General introduction. In: *Plants That Hyperaccumulate Heavy Metals*. CAB International, New York, pp. 1–14.
18. McGrath, S.P. (1998). Phytoextraction for soil remediation. In: *Plants That Hyperaccumulate Heavy Metals*. R.R. Brooks (Ed.). CAB International, New York, pp. 261–287.
19. Blaylock, M.J. et al. (1997). Enhanced accumulation of Pb in Indian mustard by soil-applied chelating agents. *Environ. Sci. Technol.* **31**: 860–865.
20. Berti, W.R. and Cunningham, S.D. (1997). In-place inactivation of Pb in Pb-contaminated soils. *Environ. Sci. Technol.* **31**: 1359–1364.
21. McBride, M.B. (1994). *Environmental Chemistry of Soils*. Oxford University Press, New York.
22. Pollard, A.J. (1996). *Ecology and evolution of hyperaccumulation in natural populations: Implications for phytoremediation of metals*. *Int. Phytoremediation Conf.*, May 8–10, 1996, Arlington, VA. International Business Communications, Southborough, MA.
23. U.S. EPA. (2000). *Introduction to phytoremediation*. EPA/600/R-99/107. Office of Research and Development, Washington, DC.
24. Glass, D.J. (1998). *The 1998 United States Market for Phytoremediation*. D. Glass Associates, Inc., Needham, MA.
25. Gulledge, C.S. (2004). *Personal correspondence*. Professional Performance Development Group, Inc.
26. U.S. EPA. (2000). Land Remediation & Pollution Control Remediation & Containment Phytoremediation of Hazardous Wastes: Potential Regulatory Acceptability. EPA's Office of Research and Development, Office of Solid Waste and Emergency Response, and the Office of Pollution Prevention and Toxics. Available: http://www.epa.gov/ORD/NRMRL/lrpcd/rcb/phytohaz.htm.
27. U.S. EPA. (2001). National Oil and Hazardous Substances Pollution Contingency Plan; National Priorities List, 40 CFR Part 300 [FRL-6963-3]. Available: http://www.epa.gov/fedrgstr/EPA-WASTE/2001/June/Day-12/f14470.htm.
28. U.S. EPA. (2001). Oceans, Coasts and Estuaries, Draft Report for Public Comment, U.S. Environmental Protection Agency Office of Water, Office of Wetlands, Oceans and Watersheds, Office of Wastewater Management, Washington, DC. Available: http://www.epa.gov/owow/invasive_species/ballast_report/current.html.
29. Datta, R. and Sarkar, D. (2004). Effective integrations of soil chemistry and plant molecule biology in phytoremediation of metals: An overview. *Environ. Geoscience* **11**(2): 53–63.

PHYTOEXTRACTION OF ZINC AND CADMIUM FROM SOILS USING HYPERACCUMULATOR PLANTS

Rupali Datta
Dibyendu Sarkar
University of Texas at San Antonio
San Antonio, Texas

INTRODUCTION

Mining, manufacturing and various other human activities have contributed to extensive soil contamination during the past century; heavy metals being the main group of inorganic contaminants (1). Accelerated release of metals into ecosystems has resulted in serious environmental problems and poses a threat to human health because most heavy metals are toxic at higher concentrations (2). Heavy metals can cause toxicity for several reasons: (1) by causing oxidative stress; (2) by replacing other essential metals in biomolecules, thereby disrupting their function; and (3) by reacting with thiol groups, thereby interfering with protein structure and function (3).

Zinc and Cd are ubiquitous pollutants that tend to occur together at many contaminated sites (4). Cadmium occurs in natural environments typically in association with Zn ores. Cadmium is emitted into the atmosphere from natural as well as anthropogenic sources. Metal production (drying of Zn concentrates, roasting, smelting and refining of ores) is the largest source of anthropogenic atmospheric Cd emissions, followed by waste incineration, production and disposal of nickel-cadmium batteries, fossil fuel combustion, and generation of dust by industrial processes, such as cement manufacturing (5). In the air, Cd vapor is rapidly oxidized and transferred to the soil by wet or dry deposition.

Most of the commercial production with Zn involves galvanizing iron and manufacturing of brass. Zinc oxide is also used in large quantities as an accelerant in tire vulcanizing. Worn tires are a major contributor to environmental Zn contamination (6). Zinc is also a major effluent from industries, such as soft drink flavoring, fur dressing and dyeing, fish processing, and laundry (7).

Cadmium derives its toxicological properties from its chemical similarity to Zn, which is an essential micronutrient for plants, animals, and humans. In humans, long-term exposure to Cd is associated with renal dysfunction, obstructive lung disease, and bone defects. In Japan, a bone disease called 'itai-itai' has been associated with prolonged exposure to Cd in rice and water (8). Cadmium and Zn are chemical antagonists; Zn deficiency can cause accelerated absorption of Cd (9). Zinc is an essential cofactor in a variety of cellular processes, including DNA synthesis, behavioral responses, reproduction, bone formation, growth, and wound healing. Although humans can handle relatively large quantities of Zn, very high levels of Zn can cause irritation and corrosion of the gastrointestinal tract, renal tubular necrosis, and interstitial nephritis (9).

Conventional soil cleanup methods for heavy-metal-contaminated soils include soil washing, excavation and reburial of soil, and pump and treat systems (10), which are rather expensive techniques. The use of plants for remediating metal-contaminated sites offers an attractive alternative because it is solar-driven and can be carried out *in situ*, minimizing the cost and human exposure (11). It offers a cheaper, safer, and environmentally friendly solution for the cleanup of metal-contaminated soils. The technique of phytoextraction depends on the extent to which plants can take up metals from soils and concentrate them in aboveground tissues (12). Plants that can accumulate metals to exceptionally high concentrations in their shoots are called hyperaccumulators (13).

Hyperaccumulators are an invaluable resource for successful phytoextraction of heavy metals, such as, Zn and Cd, either as phtoremediation crops themselves or as sources of genes for improving other phytoremediation crops (12). In this article, we focus on factors affecting Zn and Cd bioavailability in soils, the mechanism of Zn and Cd hyperaccumulation in plants, and future directions in improving phytoextraction through biotechnology and improved agronomic practices.

FACTORS AFFECTING ZINC AND CADMIUM BIOAVAILABILITY IN SOILS

Zinc

The amount of different Zn forms in soils varies considerably depending on the type and nature of soil constituents. Alkaline soils that are calcareous are generally high in the carbonate form of Zn, whereas soils rich in organic matter are high in organic forms (14). Water-soluble and exchangeable forms of Zn are readily bioavailable to plants; other forms of Zn are either unavailable or not as readily available to plants for uptake. In agriculture agronomic management practices can affect the distribution of metals among soil constituents. In Norwegian soils treated with farmyard manure, rock phosphate, or mixed fertilizers, the majority of the Zn ($\sim$78%) was reportedly associated with the oxides and residual matrices (15). In polluted soils from Poland, the majority of the Zn ($\sim$65%) was found in the oxide and residual fractions (16). The predicted dominant Zn species in soil solution below pH 7.7 is the divalent metal ion, Zn^{2+} (17). Above this pH, $ZnOH^+$ is more dominant (14).

Cadmium

In unpolluted soils, most of the Cd is associated with the residual fraction, which is generally not available for plant uptake (14). In contrast, the greatest amount of Cd ($\sim$37%) in heavily polluted soil in Norway was found in the exchangeable fraction; the oxide and residual fractions accounted for relatively lower concentrations, 23% and 15% of the Cd, respectively (18). Based on several studies, it has become apparent that Cd from mining and smelting activities is likely to be more bioavailable than Cd from unpolluted soil because increasing contamination results in increased amounts of Cd in the labile fractions of the soil (14). Most of the Cd in soil solutions is generally present as Cd^{2+} ions (14).

Cadmium–Zinc Interactions

The close association of Cd and Zn in geologic deposits and the chemical similarity of the two elements carry over into biological systems (14). Cadmium has no known biological function, whereas Zn is an essential nutrient. Cadmium competes with Zn in forming protein complexes, so a negative association between the two can be expected. The interaction between Zn and Cd is either antagonistic or synergistic (14). The addition of Cd-treated sewage sludge to calcareous soil decreased the concentrations of Zn in the shoots of the many plants studied (19). However, at higher levels of Cd, increasing concentrations of Zn increased Cd uptake (20). The variable results are related to the plant species, soil type, Zn/Cd status, source and level of applied Zn or Cd, and the Zn/Cd ratio (14).

PHYTOEXTRACTION OF ZINC AND CADMIUM

Phytoextraction is a term for the process by which plants remove significant quantities of substances from their substrate (21). Generally, the metal phytoextraction protocol consists of the following steps: (1) plant cultivation on a contaminated site, (2) removal of harvested metal-rich biomass, (3) postharvest treatments and subsequent disposal of the biomass as hazardous waste, and (4) eventual recovery of metals from the harvested biomass (22). Metal hyperaccumulation is a relatively rare phenomenon that occurs only in a few terrestrial plants. Only about 400 species have so far been identified as natural metal hyperaccumulators, representing $<0.2\%$ of all angiosperms (13). Threshold values of metal concentrations are used to define metal hyperaccumulation, which varies from element to element, for example, 10,000 mg/kg dry shoot weight for Zn and Mn; 100 mg/kg for Cu, Co, Ni, As, and Se; and 100 mg/kg for Cd (21). These concentrations are two to three orders of magnitude higher than in normal plant species growing on uncontaminated soils. The common traits shared by all hyperaccumulators include (1) a bioconcentration factor greater than 1, in some cases as high as 50–100; (2) a shoot-to-root ratio of metal concentration greater than 1, indicating efficient root to shoot transport; and (3) enhanced tolerance to metals in the medium and inside the plant cells, indicating strong internal detoxification (21).

Several species of plants that hyperaccumulate Zn have been investigated, but only one Cd hyperaccumulator, *Thlaspi caerulescens,* has been extensively studied. Many Zn hyperaccumulators are in the family Brassicaceae and in the genus *Thlaspi.* In addition, *Cardaminopsis halleri* and *Viola calaminaria* also hyperaccumulate Zn. The first indication of the possibility of phytoextraction of Zn was obtained from pot experiments in background and sludge-treated soils using *Thlaspi caerulescens* (23). It was estimated that a crop of *T. caerulescens* could take up 34 kg/ha of Zn and 0.16 kg/ha Cd. McGrath et al. (24) performed a field study on metal-contaminated soil from 20 years of sewage sludge application and reported that hyperaccumulator plants accumulated 10,625 mg/kg of Zn from a soil concentration of 406 mg/kg. McGrath and Dunham (25) reported that *T. caerulescens* and *C. halleri* plants, at field removal rates of 150 and 34 g/ha, could remove 10 years worth of Cd accumulation from agricultural land where P fertilizers added Cd at a rate of 3 g/ha/y. Only the bioavailable fraction of the total metal is of major ecological and human health concern, so it is acknowledged that all the metal need not be removed from the soil, a concept described as 'bioavailable element stripping' (BES). From the study of McGrath and Dunham (25), it is evident that in terms of BES, the bioavailable fraction of Cd can be removed from the soil by *T. caerulescens* in about four harvests.

Brown et al. (26) showed that *T. caerulescens* contained 18,000 mg/kg of Zn and 1,020 mg/kg of Cd in their shoots when grown on contaminated soils, there was no effect on yields, and the result was a in significant reduction in bioavailable metal concentration in the soils. McGrath and Dunham (25) used data from their field experiment to show that *T. caerulescens* and *C. halleri* could remove a maximum of 41 and 0.15 kg/ha/y of Zn and Cd, respectively. It is, however, not known if these extraction rates would decline with time and soil metal concentration.

Although Zn and Cd are chemically similar, as the above case studies illustrate, there are marked differences in their accumulation by hyperaccumulator plants. Cadmium and Zn also differ due to the difference in accumulation and also in relation to their response to concentrations in the soil. McGrath et al. (24) showed that *T. caerulescens* responded to increasing soil Zn concentrations by maintaining consistent hyperaccumulation of the metal. However, Cd accumulation depends much more on soil conditions, especially pH and extractable Cd concentrations (21), which indicates that there are differences in the mechanisms for accumulating the two metals in hyperaccumulators.

MOLECULAR BASIS OF ZINC AND CADMIUM HYPERACCUMULATION

Considerable progress has been made in understanding the molecular basis of metal hyperaccumulation, but the entire picture is far from complete (21). A genetic study using crosses between Zn hyperaccumulator *Arabidopsis halleri* and the nonaccumulator *Arabidopsis petraea* showed that Zn tolerance and hyperaccumulation are independent traits and that tolerance is controlled by a single major gene (27). Several studies have shown that metal hyperaccumulation of Zn and Cd by *T. caerulescens* involves enhanced metal uptake by the roots (4,28). Several Zn transporter genes cloned recently from *T. caerulescens* belong to the ZIP (Zn-regulated transporter/Fe-regulated transporter-like proteins) family (29). These genes, named *ZNT1* and *ZNT2*, are highly expressed in the roots of *T. caerulescens,* but their expression is not responsive to the Zn status of the plant. Through functional complementation in yeast, it was shown that *ZNT1* mediates high-affinity uptake of Zn^{2+} and low-affinity uptake of Cd^{2+} (30). Specific alterations in Zn-responsive elements, such as transcriptional activators, may play an important role in Zn hyperaccumulation in *T. caerulescens* (30). However, increased uptake of Cd by *T. caerulescens* cannot be explained by the Zn transport pathway but may be related to an enhanced expression of the *IRT1* gene, which is essential for Fe uptake (31). The *IRT1* gene, it was shown, can mediate high-affinity uptake of Cd^{2+} in *A. thaliana* (32,33).

Enhanced root-to-shoot transport is another important element of metal hyperaccumulation. Lasat et al. (34) reported that most of the Zn in the nonaccumulator species *T. arvense* was localized in the root cell vacuoles, whereas in the hyperaccumulator *T. caerulescens,* more Zn was transported to the shoot and accumulated in the vacuoles of shoot cells. This pattern of sequestration could be due to tissue-specific expression of transporters responsible for vacuolar Zn sequestration or enhanced xylem loading (12). Very little progress has been made in understanding how hyperaccumulators differ from nonaccumulators in this area (12). Increased metal tolerance is essential for the hyperaccumulation phenotype to occur in majority of natural hyperaccumulator plants (21). Hypertolerance is achieved by internal detoxification and possibly involves compartmentation and complexation. Several studies have reported that metals are sequestered in leaf vacuoles in Zn, Cd, Ni, and As hyperaccumulators (31,35,36). Metal transporter genes encoding putative vacuolar ion transport proteins have been cloned from *T. caerulescens* (*ZPT1*) (37). The *ZTP1* gene belongs to the cation diffusion facilitator (CDF) family and, it was found, is highly expressed, predominantly in *T. caerulescens* leaves. Enhanced tonoplast transport of Zn probably plays an important role in metal tolerance in hyperaccumulator plants. With regard to complexation, Zn was found coordinated with histidine in roots and with organic acids or uncomplexed in shoots of *T. caerulescens* (38).

Phytochelatins (PCs) are small metal-binding peptides whose synthesis from glutathione (39), homo-glutathione, hydroxymethyl-glutathione (40), or γ-glutamylcysteine (41) is catalyzed by a transpeptidase, named phytochelatin synthase (PCS), which is a constitutive enzyme that requires posttranslational activation by heavy metals (39,42). There is strong evidence that phytochelatins are essential for constitutive tolerance to Cd in nonhyperaccumulator plants, but recent reports indicate that in *T. caerulescens,* PCs are not involved in the tolerance of Cd (43). Phytochelatin synthase, it has been shown, is activated by a broad range of metals and metalloids, in particular Cd, Ag, Pb, Cu, Hg, Zn, Sn, Au, and As, both *in vivo* and *in vitro* (42,44). Naturally selected heavy metal hypertolerance, which is commonly found in plant populations from strongly metal-enriched soils, does not seem to be associated with enhanced PC synthesis. De Knecht et al. (45) obtained equal capacities and activation constants for Cd-induced PC synthesis in crude protein extracts prepared from the roots of Cd/Zn-hypertolerant and nonmetallicolous *Silene vulgaris* plants. The root PC concentrations measured *in vivo*, however, were much lower in the hypertolerant plants, even when compared at equal rates of Cd uptake. In the same species, Zn-induced accumulation of PCs in roots were also much higher in nonmetallicolous plants than in Zn-hypertolerant plants when compared at equal metal exposure levels and at equal rates of metal accumulation in the roots (46). Artificial overexpression of enzymes and transporters involved in the PC-based metal sequestration machinery, it has been shown, increases metal tolerance, in general, and Cd tolerance in particular (47–49). Several examples of enhanced PC synthesis in cell lines artificially selected for Cd hypertolerance have also been reported (50). However, there is no evidence of naturally selected enhanced PC synthesis in hypertolerant plant populations from Cd- and Zn-toxic environments (43).

OPTIMIZATION OF PHYTOEXTRACTION WITH AGRONOMIC PRACTICES AND GENETIC ENGINEERING

Successful phytoextraction of heavy metals such as Zn and Cd depends on a number of factors; the most important is the selection of appropriate plant species. The rate of metal removal depends on the biomass of the plant as well as the concentration of metals in the harvested parts. Hyperaccumulator plant species have the capacity to bioconcentrate metals, but their use is limited by their small size and slow growth rate. On the other hand, in nonaccumulator species, low potential for metal bioconcentration could be compensated for by production of a large biomass (51). The benefits of using fast-growing, high biomass nonaccumulators versus slow-growing, low biomass hyperaccumulators for phytoextraction have been debated extensively. Chaney et al. (52) analyzed the rate of Zn and Cd phytoextraction and concluded that nonaccumulator crops do not remove enough metal to support phytoextraction, despite their higher biomass. Moreover, metal phytotoxicity would further reduce their biomass in highly contaminated soils.

Lasat (53) argues that the success of phytoextraction would ultimately depend on agronomic practices applied at the site. Chaney et al. (52) proposed the use of ammonium sulfate as a soil additive to provide nutrients necessary for high plant yield and to acidify soil for greater metal bioavailability. Another important factor that affects biomass production is plant density (53). Other agronomic considerations include the use of crop rotation, weed and pest control and proper irrigation methods (53).

Conventional breeding approaches have also been proposed to overcome the two major disadvantages of using hyperaccumulators for phytoextraction; low biomass and slow growth rates. The success of this approach is limited by sexual incompatibility caused by anatomical differences between parent plants. Biotechnology has the potential to overcome this limitation by allowing direct gene transfer (53). No reports involving the expression of metal hyperaccumulator genes in nonaccumulator species are so far available, but an alternative approach has been used to transfer hyperaccumulation capacity to a nonaccumulating, high biomass plant species. Brewer et al. (54) used somatic hybridization to create a hybrid between *T. caerulescens* and *Brassica napus.* Some of the hybrids had high biomass combined with high metal tolerance and accumulation that make them attractive for metal phytoextraction.

SUMMARY AND FUTURE DIRECTIONS

Phytoextraction is an environmentally friendly approach for remediating metal-contaminated soils that has been widely accepted in recent years. If applied successfully, this technology has the capability to extract metals from large surface areas in a cost-effective manner. The cost of soil cleanup can be less than a quarter of the cost of excavation or *in situ* contaminant fixation (55). On the other hand, phytoextraction technology has the distinct disadvantage of taking much longer to accomplish cleanup goals compared with other conventional treatments. There have also been very few successful demonstrations of phytoextraction under field conditions, which require further optimization of the technique. Successful implementation of this technology is likely to be achieved in the near future as research progresses in the following fields: (1) improved agronomic practices to increase productivity of hyperaccumulators, (2) use of appropriate soil amendments to increase metal availability to plants, (3) development of conventional breeding or tissue culture techniques to improve metal uptake and accumulation capabilities of high biomass nonaccumulators, and (4) use of biotechnology to develop transgenic plants capable of efficient phytoextraction.

BIBLIOGRAPHY

1. Adriano, D.C. (1992). Bioengineering of trace metals. In: *Advances in Trace Substances Research*. Lewis, Boca Raton, FL.
2. Marschner, H. (1995). *Mineral Nutrition of Higher Plants*. 2nd Edn. Academic Press, London.
3. Pilon-Smits, E. and Pilon, M. (2002). Phytoremediation of metals using transgenic plants. *Crit. Rev. Plant Sci.* **21**(5): 439–456.
4. Zhao, F.J., Hamon, R.E., Lombi, E., McLaughlin, M.J., and McGrath, S.P. (2002). Characteristics of cadmium uptake in two contrasting ecotypes of the hyperaccumulator, *Thlaspi caerulescence*. *J. Exp. Bot.* **53**: 535–543.
5. Kazantzis, G. (1987). Cadmium. In: *Genotoxic and Carcinogenic Metals: Environmental and Occupational Occurrence and Exposure*. L. Fishbein et al. (Eds.). Advances in Modern Environmental Toxicology, Vol. 11, Princeton Scientific, Princeton, NJ.
6. Callender, E. and Rice, K.C. (2000). The urban environmental gradient: Anthropogenic influences on the spatial and temporal distributions of lead and zinc in sediments. *Environ. Sci. Technol.* **34**: 232–238.
7. Dinesan, K.C. and George, K.C. (1993). *Histopathological Studies on Zinc Toxicity in Milkfish Chanos Frsskal*. K. Rengarajan et al. (Eds.). CMFRI Special Publication: Mariculture Research Part 3, Cochin, India, pp. 25–27.
8. Nogawa, K., Honda, R., Kido, T., Tsuritani, I., and Yamada, Y. (1987). Limits to protect people eating cadmium in rice, based on epidemiological studies. *Trace Substances Environ. Health* **21**: 431–439.
9. Barceloux, D.G. (1999). Zinc. *J. Toxicol.: Clin. Toxicol.* **37**: 279–292.
10. Glass, D.J. (1999). *U. S. and International Markets for Phytoremediation, 1999–2000*. Report, D. Glass Assoc. Inc., Needham, MA.

11. Salt, D.E., Smith, R.D., and Raskin, I. (1998). Phytoremediation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**: 643–668.
12. Pollard, J.A., Powell, K.D., Harper, F.A., and Smith, J.A.C. (2002). The genetic basis of metal hyperaccumulation in plants. *Crit. Rev. Plant Sci.* **21**: 539–566.
13. Baker, A.J.M., McGrath, S.P., Reeves, R.D., and Smith, J.A.C. (2000). Metal hyperaccumulator plants: A review of the ecology and physiology of a biochemical resource for phytoremediation of metal polluted soils. In: *Phytoremediation of Contaminated Soil and Water*. N. Terry and G. Banuelos (Eds.). Lewis, Boca Raton, FL, pp. 85–107.
14. Adriano, D.C. (2001). *Trace Elements in Terrestrial Environments*. 2nd Edn. Springer-Verlag, New York.
15. Jeng, A.S. and Singh, B.R. (1995). Cadmium status of soils and plants from a long-term fertility experiment in southeast Norway. *Plant Soil* **175**: 67–74.
16. Chlopecka, A., Bacon, J.R., Wilson, M.J., and Kay, J. (1996). Forms of cadmium, lead, and zinc in contaminated soils from Southwest Poland. *J. Environ. Qual.* **25**: 69–79.
17. Lindsay, W.L. (1979). *Chemical Equilibria in Soils*. John Wiley & Sons, New York.
18. Hickey, M.G. and Kittrick, J.A. (1984). Chemical partitioning of cadmium, copper, nickel, and zinc in soils and sediments containing high levels of heavy metals. *J. Environ. Qual.* **13**: 372–376.
19. Bingham, F.T. and Page, A.L. (1975). *Proc. Int. Conf. Heavy Metals Environ*. University of Toronto, Ontario, Canada.
20. Gerritse, R.G., Van Driel, W., Smilde, W.W., and Van Luit, B. (1983). Uptake of heavy metals by crops in relation to their concentration in the soil solution. *Plant Soil* **75**: 393–404.
21. McGrath, S.P. and Zhao, F.J. (2003). Phytoextraction of metals and metalloids from contaminated soils. *Curr. Opin. Biotechnol.* **14**: 1–6.
22. Blaylock, M. and Huang, J. (2000). Phytoextraction of metals. In: *Phytoremediation of Toxic Metals: Using Plants to Clean Up the Environment*. I. Raskin and B. Ensley (Eds.). John Wiley & Sons, New York, pp. 53–69.
23. Baker, A.J.M., Reeves, R.D., and McGrath, S.P. (1991). *In situ* decontamination of heavy metal polluted soils using crops of metal-accumulating plants—a feasibility study. In: *In situ Bioreclamation*. R.E. Hinchee and R.F. Olfenbuttel (Eds.). Butterworth-Heinemann, Stoneham, MA, pp. 539–544.
24. McGrath, S.P., Sidoli, C.M.D., Barker, A.J.M., and Reeves, R.D. (1993). The potential for the use of metal-accumulating plants for the *in situ* decontamination of metal-polluted soils. In: *Integrated Soil and Sediment Research: A Basis for Proper Protection*. H.J.P. Eijsackers and T. Hamers (Eds.). Kluwer Academic, Dordrecht, The Netherlands, pp. 673–676.
25. McGrath, S.P. and Dunham, S.J. (1997). Potential Phytoextraction of Zinc and Cadmium from Soils using Hyperaccumulator Plants. *Proc. 4th Int. Conf. Biogeochemistry Trace Elements*. Berkeley, Ann Arbor Press, Chelsea, MI.
26. Brown, S.L., Chaney, R.L., Angle, J.S., and Barker, A.J.M. (1994). Phytoremediation potential of *Thlaspi caerulescence* and bladder campion for zinc- and cadmium-contaminated soil. *J. Environ. Qual.* **23**: 1151–1157.
27. Mcnair, M.R., Bert, C., Huitson, S.B., SaumitouLaprade, P., and Petit, D. (1999). Zinc tolerance and hyperaccumulation are genetically independent characters. *Proc. R. Soc. London Ser B* **266**: 2175–2179.
28. Lasat, M.M., Baker, A.J.M., and Kochian, L.V. (1996). Physiological characterization of root Zn^{2+} absorption and translocation to shoots in Zn hyperaccumulator and nonaccumulator species of *Thlaspi*. *Plant Physiol.* **112**: 1715–1722.
29. Maser, P. et al. (2001). Phylogenetic relationships within cation transporter families of *Arabidopsis*. *Plant Physiol.* **126**: 1646–1667.
30. Pence, N.S. et al. (2000). The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator, *Thlaspi caerulescence*. *Proc. Natl. Acad. Sci. USA* **97**: 4956–4960.
31. Lombi, E. et al. (2002). Influence of iron status on calcium and zinc uptake by different ecotypes of the hyperaccumulator *Thlaspi caerulescens*. *Plant Physiol.* **128**: 1359–1367.
32. Vert, G. et al. (2002). *IRT1*, an *Arabidopsis* transporter essential for iron uptake from the soil and for plant growth. *Plant Cell* **14**: 1223–1233.
33. Connolly, E.L., Fett, J.P., and Guerinot, M.L. (2002). Expression of the *IRT1* metal transporter is controlled by metals at the levels of transcript and protein accumulation. *Plant Cell* **14**: 1347–1357.
34. Lasat, M.M., Baker, A.J.M., and Kochian, L.V. (1998). Altered Zn compartmentation in the root symplasm and simulated Zn absorption into the leaf as mechanisms involved in Zn hyperaccumulation in *Thlaspi caerulescens*. *Plant Physiol.* **118**: 875–883.
35. Kupper, H., Zhao, F.J., and McGrath, S.P. (1999). Cellular compartmentation of zinc in leaves of the hyperaccumulator *Thlaspi caerulescense*. *Plant Physiol.* **199**: 305–312.
36. Kramer, U., Pickering, I.J., Prince, R.C., Raskin, I., and Salt, D.E. (2000). Subcellular localization and speciation of nickel in hyperaccumulator and nonaccumulator *Thlaspi* species. *Plant Physiol.* **122**: 1343–1353.
37. Assuncao, A.G.L. et al. (2001). Elevated expression of metal transporter genes in three accessions of the metal hyperaccumulator *Thlaspi caerulescence*. *Plant Cell Environ.* **24**: 217–226.
38. Salt, D.E., Prince, R.C., Baker, A.J.M., Raskin, I., and Pickering, I.J. (1999). Zinc ligands in the hyperaccumulator, *Thlaspi caerulescence* as determined using X-ray absorption spectroscopy. *Environ. Sci. Technol.* **33**: 713–717.
39. Grill, E., Loffler, S., Winnacker, E.L., and Zenk, M.H. (1989). Phytochelatins, the heavy-metal-binding peptides of plants, are synthesized from glutathione by a specific γ-glutamylcysteine dipeptidyl transpeptidase (phytochelatin synthase). *Proc. Natl. Acad. Sci. USA* **86**: 6838–6842.
40. Klapheck, S., Schlunz, S., and Bergman, L. (1995). Synthesis of phytochelatins and homo-phytochelatins in *Pisum sativum* L. *Plant Physiol.* **107**: 515–521.
41. Hayashi, Y., Nakagawa, C.W., and Mutoh, N. (1991). Two pathways in the biosynthesis of cadystins (γ-EC)nG in the cell-free system of the fission yeast. *Biochem. Cell Biol.* **69**: 115–121.
42. Chen, J., Huang, J.W., Caster, T., and Cunningham, S.D. (1997). *Arabidopsis* as a model system for studying lead accumulation and tolerance in plants. In: *Phytoremediation of Soil and Water Contaminants*. K.L. Kruger et al. (Eds.). American Chemical Society, Washington, DC, pp. 264–273.
43. Schat, H., Llugany, M., Vooijs, R., Hartley-Whitaker, J., and Bleeker, P.M. (2002). The role of phytochelatins in constitutive and adaptive metal tolerances in hyperaccumulator and non-hyperaccumulator metallophytes. *J. Exp. Bot.* **53**: 2381–2392.
44. Grill, E., Winnacker, E.L., and Zenk, M.H. (1987). Phytochelatins, a class of heavy-metal-binding peptides from

plants, are functionally analogous to metallothioneins. *Proc. Natl. Acad. Sci. USA* **84**: 439–443.

45. De Knecht, J.A. et al. (1995). Synthesis and degradation of phytochelatins in cadmium-sensitive and cadmium-tolerant *Silene vulgaris*. *Plant Sci.* **106**: 9–18.
46. Harmens, H., Cornelisse, E., Den Hartog, P.R., Ten Bookum, W.M., and Verkleij, J.A.C. (1993). Phytochelatins do not play a key role in naturally selected zinc tolerance in *Silene vulgaris* (Moench) Garcke. *Plant Physiol.* **103**: 1305–1309.
47. Ortiz, D.F., Ruscitti, T., McCue, K.F., and Ow, D.V. (1995). Transport of metal-binding peptides by HMT1, a fission yeast ABC-type B vacuolar membrane protein. *J. Biol. Chem.* **270**: 4721–4728.
48. Clemens, S., Kim, E.J., Neuman, D., and Schroeder, J.I. (1999). Tolerance to toxic metals by a gene family of phytochelatin synthases from plants and yeast. *EMBO J.* **18**: 3325–3333.
49. Zhu, Y.L., Pilon-Smits, E.A.H., Jouanin, L., and Terry, N. (1999). Overexpression of glutathione synthase in Indian mustard enhances cadmium accumulation and tolerance. *Plant Physiol.* **119**: 73–79.
50. Chen, J. and Goldsbrough, P.B. (1994). Increased activity of γ-glutamyl cystein synthetase in tomato cells selected for cadmium tolerance. *Plant Physiol.* **106**: 233–239.
51. Ebbs, D.S., Lasat, M.M., Brady, D.J., Cornish, J., Gordon, R., and Kochian, L.V. (1997). Phytoextraction of cadmium and zinc from a contaminated site. *J. Environ. Qual.* **26**: 1424–1430.
52. Chaney, R.L., Li, Y.L., Angle, J.S., Baker, A.J.M., Reeves, R.D., Brown, S.L., Homer, F.A., Malik, M., and Chin, M. (1999). Improving metal hyperaccumulator wild plants to develop commercial phytoextraction systems: Approaches and progress. In: *Phytoremediation of Contaminated Soils and Water*. N. Terry and G.S. Banuelos (Eds.). CRC Press, Boca Raton, FL.
53. Lasat, M.M. (2000). Phytoextraction of metals from contaminated soil: A review of plant/soil/metal interaction and assessment of pertinent agronomic issues. *J. Hazard Substance Res.* **2**: 1–25.
54. Brewer, E.P., Saunders, J.A., Angle, J.S., Chaney, R.L., and McIntosh, M.S. (1999). Somatic hybridization between the zinc accumulator *Thlaspi caerulescence* and *Brassica napus*. *Theor. Appl. Genet.* **99**: 761–771.
55. Black, H. (1995). Absorbing possibilities: Phytoremediation. *Environ. Health Perspect.* **103**: 1106–1108.

PHYTOREMEDIATION ENHANCEMENT OF NATURAL ATTENUATION PROCESSES

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Phytoremediation is a remediation method that uses what is in essence enhanced natural attenuation for cleanup. There are a variety of phytoremediation methods, some having multiple names: phytostabilization, rhizodegradation (phytostimulation, rhizosphere bioremediation, or plant-assisted bioremediation), rhizofiltration (contaminant uptake), phytodegradation (phytotransformation), phytovolatilization, and Phytoaccumulation (phytoextraction or hyperaccumulation). Various types of plants can be used in phytoremediation, including poplar trees, alfalfa, black locust, Indian mustard, fescue grass, crested wheatgrass, and Canada wild rye.

Figure 1 and Tables 1 and 2 summarize the overall concepts of phytoremediation and the types of processes and chemicals that can be treated.

A form of natural attenuation for indigenous plants at a site and enhanced bioremediation for cultivated plants, phytoremediation acts through two fundamental remediation processes: enhancement of saturated zone *in situ* biodegradation and phytoextraction.

The rhizosphere is the zone in the subsurface occupied by the plants' root systems. Root depths of 6–30 ft are common, with some trees and shrubs capable of root penetration to 60 ft. Phreatophytes are deep-rooted plants that draw water from beneath the water table; xerophytes are shallow-rooted and depend more directly on infiltrating rainwater. Depending on site conditions and contaminant distribution, either type of plant can be of value for phytoremediation.

Within the rhizosphere, plants contribute to enhanced *in situ* biodegradation through the supply of carbonaceous substrate and oxygen transfer. Rhizodeposition is partially

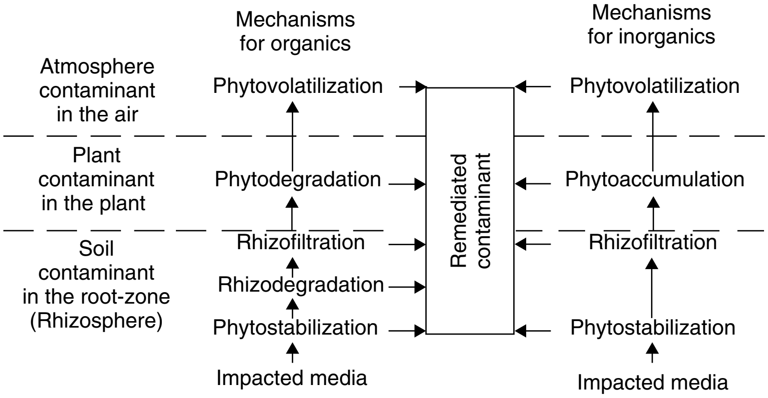

Figure 1. Overview of phytoremediation. (After Ref. 1).

Table 1. Types of Phytoremediation for Organic Constituents

Type of Phytoremediation	Process Involved	Contaminant Treated
Phytostabilization	Plants control pH, soil gases, and redox conditions in soil to immobilize contaminants. Humification of some organic compounds is expected.	Expected for phenols, chlorinated solvents (tetrachloromethane and trichloromethane), and hydrophobic organic compounds
Rhizodegradation (phytostimulation, rhizosphere bioremediation, or plant-assisted bioremediation)	Plant exudates, root necrosis, and other processes provide organic carbon and nutrients to spur soil bacteria growth by two or more orders of magnitude. Exudates stimulate degradation by mycorrhizal fungi and microbes. Live roots can pump oxygen to aerobes and dead roots may support anaerobes.	Polyaromatic hydrocarbons, BTEX, and other petroleum hydrocarbons, perchlorate, atrazine, alachlor, polychlorinated biphenyl (PCB), and other organic compounds
Rhizofiltration (contaminant uptake)	Compounds are taken up or sorbed by roots (or sorbed to algae and bacteria).	Hydrophobic organic chemicals
Phytodegradation (phytotransformation)	Aquatic and terrestrial plants take up, store, and biochemically degrade selected organic compounds to harmless by-products, products used to create new plant biomass, or by-products that are further broken down by microbes and other processes to less harmful products. Reductive and oxidative enzymes may be used in series in different parts of the plant.	Munitions (TNT, DNT, HMX, nitrobenzene, picric acid, nitrotoluene), atrazine, halogenated compounds (tetrachloromethane, trichloromethane, hexachloroethane, carbon tetrachloride, TCE, tetrachlorethane, dichloroethant), DDT and other chlorine and phosphorus based pesticides, phenols, and nitrites
Phytovolatilization	Volatile organic compounds are taken up and transpired. Some recalcitrant organic compounds are more easily degraded in the atmosphere (photodegradation).	Chlorinated solvents (trichloroethane), organic VOCs, BTEX, MTBE

Source: Adapted from Reference 1.

Table 2. Types of Phytoremediation for Inorganic Constituents

Type of Phytoremediation	Process Involved	Contaminant Treated
Phytostabilization	Plants control pH, soil gases, and redox conditions in soil to immobilize contaminants. Humification of some organic compounds is expected.	Proved for heavy metals in mine tailing ponds
Rhizofiltration (contaminant uptake)	Compounds are taken up or biosorbed by roots (or sorbed to algae and bacteria).	Heavy metals and radionuclides
Phytoaccumulation (phytoextraction or hyperaccumulation)	Metals and organic chemicals are taken up by the plant with water, or by cation pumps, sorption, and other mechanisms.	Nickel, zinc, lead, chromium, cadmium, selenium, other heavy metal radionuclides
Phytovolatilization	Volatile metals are taken up, changed in species, and transpired.	Mercury and selenium

Source: Adapted from Reference 1.

the result of the decay of dead roots and root hairs. Also important are carbonaceous root exudations, such as leakage from epidermal cells, secretions resulting from metabolic activity, mucilages from root tips (which act as lubricants for root penetration), and lysates from sloughed cells. Exudates are composed of a wide range of chemicals that include sugars, amino acids, organic acids, fatty acids, and numerous other compounds. It is estimated that 7–27% of the total plant mass is annually deposited as carbonaceous material in the rhizoshpere, amounting to 85–155 tons/acre. This carbonaceous material stimulates overall bacterial activity as well as providing substrate to support cometabolic degradation of xenobiotic hydrocarbons.

The capability to support oxygen transfer can be divided into three categories of plants: nonwetland herbaceous and woody plants with poor oxygen transfer capacity; wetland woody plants with moderate capacity; and wetland herbaceous plants with high oxygen transport capacity. Plants of the first variety will ultimately die with their roots under saturated conditions. Wetland plants adsorb oxygen through their leaves, twigs, stems, bark, and unflooded roots. This oxygen is in turn transported to the roots, where it diffuses out into the rhizosphere. Consequently, wetland plants are able to support aerobic biooxidation in the rhizosphere at rates that are significantly greater than that seen through saturated zone diffusion alone. Due to the processes described above, it is not uncommon to find bacterial population levels in the rhizoshpere elevated a magnitude or more above surrounding undeveloped soils.

The primary *in situ* remediation potential of plants for hydrocarbons lies in their capacity to enhance oxidation rates in the subsurface and provide cometabolic substrate. However, plants also have the ability to remove compounds, a process termed phytoextraction, which can be applied to organic or metal contaminants. In the case of hydrocarbons, the compound must be water soluble and have a moderate degree of lipid solubility. Lipid solubility is a function of the octanol–water partition coefficient (K_{ow}) for the compound. Compounds most readily mobilized by plants have log K_{ow} values in the range of 1–3. Compounds with values of K_{ow} in this range include BTEX hydrocarbons, chlorinated solvents, and other short-chain aliphatic hydrocarbons.

Once in the interior of the plant, the adsorbed hydrocarbons may be stored via lignification, volatilized, partially degraded through metabolization, or completely mineralized.

Compounds with values of log K_{ow} higher than 3 such as PNAs are incapable of entering the root; those with log K_{ow} values lower than 1 are rejected by the root membrane. In addition, the desorption and mass transfer of contaminant hydrocarbons from the geologic matrix may be the rate limiting step in the remediation process.

In phytoremediation of metals, the dominant active mechanism is phytoextraction and accumulation in the tissues of the plant, which is a process that has long been familiar to exploration geologists in the mining industry. Geobotany is concerned with the identification of plants or plant conditions common to metal-rich soils and biogeochemistry is concerned with actual metal concentrations in parts of plants.

The mechanisms for metal accumulation include:

- Chelation
- Precipitation
- Compartmentalization
- Translocation

These same mechanisms often contribute to the metal tolerance of the plant. To date, accumulators of lead, cadmium, chromium, nickel, cobalt, zinc, and selenium have been identified. To successfully apply this technology to a metal contaminated site, it is important that pH, organic complexes, and interfering elements be assessed and that plant species with the appropriate metal selectivity be utilized. In some instances, it may be necessary to apply soil amendments to enhance the process.

A key part of the metal extraction process is played by phytochelatins, which are low molecular weight peptides that have the capability to bind metals. Their presence in plants has likely evolved because toxicity to the plant is reduced by having the metals bound by the phytochelatins. Currently, there are several hyperaccumulating plants that have been discovered such as Indian mustard and pennycress. However, these plants are small and slow growing and thus lack enough biomass to remove significant amounts of metals. For example, there are 26 known plants that are hyperaccumulators of cobalt and exhibit desirable properties of common crops such as high growth rates and easy harvesting.

Lastly, plants have other properties with potential remediation applications. In semiarid climates, sufficient numbers of trees are capable of depressing the water table through transpiration, up to the equivalent of 3 ft of rainfall per year. Phytoremediation can be used as a form of hydraulic control. In tight soils, root penetration can improve the overall mass transport properties.

Plants can also aid in the surface stabilization of soils, preventing the windblown migration of soil with adsorbed contaminants.

Planting costs have been estimated to be in the range of $10,000 per acre, with monitoring costs parallel to those associated with other remediation technologies. Total phytoremediation costs are estimated to range from $60,000 to $100,000 per acre. The potential of phytoremediation is significant given proper design for its application. Contaminant type and distribution, soil chemistry, and climate are all important factors that must be considered.

BIBLIOGRAPHY

1. Suthersan, S. (2002). *Natural and Enhanced Remediation Systems*. Lewis Publishers/CRC Press, Boca Raton, FL, p. 419.

BACTERIA ROLE IN THE PHYTOREMEDIATION OF HEAVY METALS

Anthea Johnson
Naresh Singhal
P. Takis Elefsiniotis
University of Auckland
Auckland, New Zealand

INTRODUCTION

Rapid industrialization in the last century has led to the release of numerous contaminants, including heavy metals, to the environment. The release of heavy metals is of particular concern as they cannot be degraded, but only transformed from one form to another, and can accumulate in natural systems. Examples of heavy metals of particular significance include cadmium, copper, chromium, iron, lead, mercury, nickel, and zinc. They are released into the environment from a wide range of anthropogenic sources. Vehicles, industrial activities, building materials, and even the atmosphere contribute heavy metals to the water and soil, resulting in their occurrence at some sites in higher concentrations than is acceptable.

Heavy metals may remain dissolved in water or attach (referred to as sorption) to the mineral and organic components of soil. Although some heavy metals are required in minute quantities as micronutrients for the growth of plants and bacteria, at high concentrations they present a toxic threat to ecosystems. Bacteria and plants are susceptible to toxic effects of heavy metals, as are animals and humans who consume vegetation and crops grown in contaminated soils. Heavy metals are capable of bioaccumulation in plant and animal tissues and, therefore, can be biomagnified through the food chain.

Currently, heavy metal contaminated sites are treated by various immobilization and extraction methods. Pump-and-treat systems aim to wash heavy metals from the soil, using water alone, or metal-solubilizing compounds such as chelators and surfactants. Other methods involve thermally volatilizing or vitrifying contaminants in soil. At some sites, the soil is simply removed and dumped elsewhere. The existing treatment methods for contaminated soils are, in many cases, prohibitively expensive, which may postpone or preclude any treatment action. To gain wider acceptability and implementation, remediation technologies must be a low cost, sustainable, and effective alternative to physical and chemical treatment methods. Phytoremediation is one such emerging technology that can be used effectively for at-source contaminant mitigation.

The three basic components of phytoremediation systems are plants, bacteria, and the contaminated soil. Heavy metals are sorbed to the minerals in soil by cation exchange or to soil organic matter by complexation with organic acids. The complex interactions that have evolved between plants and bacteria in the rhizosphere (root zone) can be harnessed to remediate metal-contaminated soils. Bacteria and plants mobilize metals for uptake by excreting a vast array of organic compounds such as lipids, sterols, enzymes, and surfactants. Alternatively, they can immobilize metals via chemical transformation or precipitation of insoluble compounds. Labile metals can be sorbed to bacterial cells or taken up through plant cell membranes. Metals taken up by roots are translocated in the xylem (woody tissue) to the aboveground plant biomass, which can be harvested, thereby providing the basis for developing a decontamination technology using phytoremediation. Here we discuss the major processes by which bacteria and plants mobilize/immobilize heavy metals within the rhizosphere.

MOBILIZATION PROCESSES

The mobilization of metals from soil has a direct influence on their bioavailability to plants and hence the potential for their uptake and removal from the soil. There are a number of mechanisms by which bacteria and plants facilitate the dissolution of precipitates or the desorption of ions attached to mineral and organic fractions of soil.

Acidification/Protonation

Plant and bacterial exudates can alter the pH of soils and hence the bioavailability of contaminants and nutrients to both plants and bacteria (1). Soil acidification occurs due to proton efflux, the release of organic acids, or the addition of respiratory carbon dioxide to the soil (1). Free protons (H^+ ions) are added to the soil by a number of bacterial processes. Organic and amino acids are secreted by plant roots and over time may lower the pH of soils, increasing metal solubility. Bacterial respiration results in the addition of carbon dioxide to the soil, forming carbonic acid. The dissociation of acids in soil water causes the release of protons, which compete with metal ions for sorption or complexation sites. In this way, the protons may displace the metal cations into solution.

Plants too have a number of specialized mechanisms to increase the solubility of metals and nutrients. For example, iron is a micronutrient required for plant growth. To mobilize iron from the soil, plants may release protons to convert iron(III) to iron(II) and thus increase its solubility (2,3).

Solubilization by Biosurfactants

Biosurfactants are surface-active agents that are secreted by plants and bacteria within the rhizosphere (4). Microbially produced biosurfactants have many of the same properties as synthetic surfactants but have several advantages for *in situ* bioremediation. They are nontoxic to most plants and bacteria and are biodegradable, which prevents their accumulation in soil or groundwater. In addition, biosurfactants can be produced by bacteria localized around plant roots where they are needed.

Biosurfactants have both a hydrophobic and a hydrophilic region, which allows them to partition across interfaces of varying polarity. The hydrophobic region is often a nonpolar hydrocarbon chain, which reduces the surface tension at the interface (5). For example, it has been found that surfactin, a biosurfactant produced by *Bacillus subtilus,* reduced the surface tension of water from 72 to 27 mN/m at a concentration of only 0.005% (6). The lowering of the interfacial tension between the aqueous solution and substrate enables bacteria to access metals and degrade lipophilic substrates in the aqueous phase.

At low concentrations, surfactant molecules behave independently as monomers. On a solid surface, the surfactant molecules may form one or more layers (admicelles) with regions of similar polarity associated with each other. At concentrations above a nominal value, the surfactant molecules aggregate into micelles in the bulk phase, in addition to the interface. This value is dependent on the biosurfactant and is known as the critical micelle concentration (CMC) (7).

Biosurfactants have the capability to solubilize heavy metals that are sorbed to soil particles or soil organic matter. The surfactant molecules are adsorbed to the surface of the soil particle where they form complexes with attached metal ions (8). These complexes may then dissociate from the solid to the liquid phase, where similar surfactant molecules form micelles.

Plants produce surfactants such as lipids and sterols for root lubrication to aid root progression through the soil (9). These surfactants are also able to facilitate the desorption of contaminants and nutrients from soil particles. For example, saponins are plant-derived biosurfactants that are capable of binding to heavy metals such as cadmium, copper, lead, and zinc (10).

Chelation

Another mechanism by which bacteria and plants influence the bioavailability and toxicity of heavy metals is via the production of chelators. These are ligands that bind to metal ions forming chelates—ring-like structures that can increase the solubility of metals in solution. The production of chelates by microbes in the rhizosphere is greater than that in the bulk soil (11,12). Chelation enhances the uptake of ions with larger valence than ions with lower valence; Parker et al. (13) found that chelators reduced the uptake of divalent metal ions while increasing the uptake of trivalent ions such as aluminum. In addition, dissociated organic anions such as citrate and oxalate may also be capable of forming complexes with metal ions sorbed to soil minerals or organic matter (1).

Chemical Transformation

Bacteria are capable of altering metal speciation between soluble and insoluble forms through reduction processes. The change in oxidation state of metals influences their solubility and hence their bioavailability. Reduction under anaerobic conditions is carried out by a range of heterotrophic bacteria. During respiration, these bacteria can use a metal cation as a terminal electron acceptor, reducing to a lower valence state (14), which can cause an increase in the solubility of metals such as iron and manganese since iron(II) and manganese(II) are more soluble than iron(III) and manganese(IV) (15). Similarly, the reduction of Hg(II) to Hg(0) within bacterial cells causes the diffusion of elemental mercury out of cells (1). However, in some cases reduction can decrease the available metal in solution, such as the reduction of chromium from VI to III valence results in lower solubility and mobility (1).

IMMOBILIZATION PROCESSES

Sorption

Binding to Cells. Bacteria have a net negative charge due to the presence of hydroxyl (OH^-), carboxyl (COO^-), and phosphate (PO_4^{3-}) functional groups on the membrane surface. This causes the attraction and sorption of free metal cations to negatively charged external sites on bacterial cells. The binding capacity is dependent on the type of bacteria. For example, the binding capacity of gram-positive bacteria is greater than that of gram-negative bacteria due to the difference in the structure and polymer content of membranes and the cell wall. The binding of metal ions by bacteria increases with temperature and is greater at lower pH (16).

Heavy metal cations can form complexes, which reduces their positive charge and increases their size, which causes the cations to be less favored for transport through plant and bacterial cell membranes and instead they are retained by the surface sites on cells.

Binding to Extracellular Molecules. Bacteria and plants excrete organic substances that are capable of binding heavy metals in ionic or solid forms. Simple organic molecules include acids and alcohols, which can attach to free metal ions. Macromolecules include long-chain organic acids such as humic and fulvic acids, which are capable of binding metal cations (17). Exopolysaccharides are a major component of biofilms and other microbial assemblages and are capable of trapping particulates and ions. Bacteria can also secrete specific metal binding proteins (metallothioneins) in response to the presence of toxic metals (1).

Uptake and Intercellular Sequestration

The uptake of heavy metal cations is an energetically favorable process for plant and bacterial cells due to their negative internal charge. Negative transmembrane potentials encourage the uptake of metal cations into the cell (13). Bacteria may store metals within inclusion bodies and thus render them temporarily unavailable and hence nontoxic.

Precipitation/Crystallization/Mineralization

Bacteria are capable of reducing metals to insoluble forms, lowering their bioavailability and hence toxicity to bacteria, plants, and animals. The bacterial reduction of chromium(VI) to chromium(III) immobilizes the metal in soil (1). Researchers have identified a strain of *Xanthomonas maltophyla* that is capable of catalyzing the reduction and precipitation of chromium as well as the transformation of other toxic metal ions including lead and mercury (18).

Volatilization

Biomethylation involves the microbial transfer of methyl ($-CH_3$) groups to a metal or metal-containing compound, often forming methylated species with increased volatility. Methylation of metals such as arsenic, lead, mercury, selenium, and tin can occur in either aerobic or anaerobic environments (1). Volatile methylated species can then escape from the soil into the atmosphere. However, methylated species of metals such as mercury and arsenic can be more toxic. For example, methylmercury is a neurotoxic compound formed via the methylation of ionic mercury. Some bacteria are capable of reducing mercury from its ionic to elemental form, increasing its volatility and decreasing its toxicity (2).

INTERACTIONS BETWEEN PLANTS, BACTERIA, AND CONTAMINANTS

Plants as Energy Sources for Microbes

The growth of bacteria in the rhizosphere is carbon limited (19). Most organic carbon in the rhizosphere is derived from plant sources. Root cap mucilage, root exudates, photosynthate, and plant debris provide organic carbon for microbial cell growth. Plant roots directly or indirectly foster rhizospheric bacteria by producing exudates that bacteria and fungi may use as a carbon source. Plant root exudates consist primarily of organic acids (of low and high molecular weight) as well as sugars and amino acids. These substrates can be easily degraded and utilized as an energy source by many types of soil bacteria, encouraging a higher concentration of soil microbes in the rhizosphere than in bulk soil (20). An increase in microbial biomass usually results in higher rates of contaminant degradation or mobilization due to the increased total activity of the rhizospheric community.

Transpiration Effects on Contaminant Movement

Plant root uptake of water influences the rate and direction of contaminant migration through the soil.

Soluble metals contained within the groundwater are drawn toward the roots by the transpiration stream. In this way, the contaminants move into the region where they may be taken up by the roots and the surrounding rhizospheric community, which facilitates access to contaminants by nonmotile bacteria and improves the possibilities for microbial action resulting in transformation or degradation (21). In addition, the network of plant roots intercepts contaminants drawn downward by the infiltration of rainwater. This reduces the loss of nutrients and contaminants from the upper regions of the soil.

Secondary Plant Metabolites

Secondary plant metabolites (SPMEs) comprise a wide range of organic compounds produced by plants and exuded by their roots. Although SPMEs are not essential for basic metabolic processes, they have an important role in the nutrition and defense of plants (22).

There are many theories as to the purpose of SPME production (23). It is generally accepted that many SPMEs are effectively natural pollutant analogues (24). The resemblance of many SPMEs to organic pollutants may encourage the bacterial production of enzymes that are capable of degrading both types of compounds. Thus, the plant reduces the time needed for bacteria to develop the enzymes required for the degradation of pollutants in the rhizosphere.

SPMEs are also produced by plants for defense against herbivory as well as insect and microbial attack (25). Allelopathic SPMEs are produced to reduce competition between plant species (26). SPMEs are also produced to act as chemical signals to other plants and microbes. However, this is not always to the benefit of the plant, as this trait can also encourage the growth of parasitic plants (27).

Cometabolism of Contaminants. Secondary plant metabolites may include enzymes that are intended for the degradation of one compound, but are also capable of degrading other, perhaps xenobiotic, compounds with no additional benefit to the plant in terms of energy or nutrition. In this case, the primary compound is referred to as a cometabolite as it provides the energy source for the degradation of the secondary compound (22). As cometabolism does not require the population to produce a specific enzyme for contaminant degradation, cometabolism of contaminants is improved by increasing the microbial biomass and diversity in the rhizosphere.

Exudates as Chemical Signals. Walton et al. (28) proposed the theory for the plant production of exudates as chemical signals to bacteria. When a plant detects the presence of a toxicant, it alters the composition of the root exudates produced. In turn, the plant root exudates alter the rhizospheric community in such a way as to increase the bacterial detoxification of the soil for the benefit of the plant, which may be achieved by the release of specific enzymes, cometabolites, or energy sources to increase the number and activity of rhizospheric bacteria or to change the bacterial exudates produced (28). Associations between plants and bacteria may be both specific and nonspecific and are thought to have evolved over time for the mutualistic benefits they offer.

Selection for Metal-Resistant Bacteria. Hyperaccumulating plant species may directly or indirectly increase the selection for metal-resistant bacteria in their rhizosphere. Hyperaccumulating plants are generally defined as variants capable of accumulating metals such as nickel, cobalt, copper, zinc, cadmium, manganese, and lead within tissues to concentrations 100 times or more greater than normal plants (29). Such plants usually prefer one particular metal, although some multimetal accumulators have been encountered. More than 400 hyperaccumulating species have been identified (30). One of the better known zinc hyperaccumulating species is *Thlaspi caerulescens*, which has been well researched for its phytoremedial potential. Higher ratios of metal-resistant bacteria are seen to exist around the roots of *T. caerulescens* than the nonaccumulating plant species growing in the same zinc-contaminated soil. This behavior is attributed to the release of metals for uptake via acidification of the plant rhizosphere, resulting in selection for bacteria that can tolerate higher zinc concentrations (31).

Plant Pathogens

Plants are subject to attack from insects and microorganisms. Insects eat plant tissues, such as leaves, or suck the sugar-rich phloem from the stalks. Bacteria can produce toxins that affect plant growth and hence reduce competition for soil nutrients. These bacterial phytotoxins may have been a reason for the evolution of detoxification mechanisms by plants and the production of allelopathic compounds as a defense mechanism. Bacteria also have a role as plant pathogens, causing disease to plant tissues and affecting the ability of plants to remove metals from soil. However, bacteria may also defend plants against attack from insects, fungi, and viruses. For example, the production of toxins by endophytic bacteria may protect plants from herbivory by chewing insects.

Nutrient Requirements

In many soils, the availability of nutrients is limited, which causes competition between bacteria and plants for nutrients. Bacteria are capable of forming mutualistic or antagonistic relationships with plants with respect to nutrient availability.

The presence of rhizospheric bacteria may have beneficial effects for plant biomass production. An increased microbial population results in greater mobilization of nonlabile nutrients and increased uptake by plant roots. Limiting nutritional elements are mobilized from soils by bacterial exudates. For example, siderophores (mugineic and avenic acids) are iron-chelating ligands that solubilize iron from the soil, which assists plants in assimilating iron and increases plant resistance to iron deficiency (12).

Some species of hyperaccumulating plants require heavy metals for normal growth, in concentrations that would be toxic to most species. Bacteria may assist these plants in meeting their high demand for metals. Studies have shown that the total uptake of zinc by shoots of *Thlaspi caerulescens* was increased by a factor of 4 due to the presence of bacterial species, which increase was attributed to the twofold increase in the concentration of zinc in the shoots and a similar increase in the shoot biomass (12).

Root Area Augmentation

The presence of rhizospheric bacteria and plant mycorrhizal associations can increase the surface area of the plant root system. In addition, the presence of rhizospheric communities can also increase the biomass of plants. It has been found that bacterial seed inoculates increase plant root biomass, length, and exudate production (32). Hadas and Okon (33) demonstrated that the root growth of tomato seedlings was improved by inoculation with the bacteria *Azospirillum brasilense*. The roots had greater surface area due to the increased length and density of root hairs, with a subsequent increase in the uptake of water and minerals from the soil (33).

Soil Detoxification

Mutualistic relationships may develop between plants and bacteria living in contaminated soils. The presence of organic and inorganic contaminants may prevent the normal growth of plants and hence the supply of nutrients and energy for rhizospheric bacteria. Therefore, bacteria benefit from the nutritional gains brought about by their ability to detoxify soils. Where high concentrations of metals are present, the rhizospheric bacteria may mediate their immobilization or transformation, reducing metal bioavailability by binding metals to the negatively charged surface of bacteria, which may reduce their toxicity to plants. Additionally, rhizosphere communities may assist their host plants by mineralizing or degrading toxic organics to less hazardous forms.

CONCLUSION

At many sites around the world, heavy metals have built up to toxic concentrations in soils. High concentrations of soil metals are harmful to most forms of life including plants, bacteria, and humans. Bacteria interact with plants and heavy metal contaminants in a variety of ways to mediate the mobilization and immobilization of heavy metals within the soil. Plants may directly or indirectly stimulate these microbial processes through the release of a diverse range of root exudates. Recent research has shown that phytoremediation systems may be improved by building on the natural processes and interactions between plants and bacteria that occur in the rhizosphere. Phytoremediation may present a more sustainable solution to the worldwide problem of contaminated soils by further developing the remedial processes present in natural systems.

BIBLIOGRAPHY

1. Gadd, G.M. (2004). Microbial influence on metal mobility and application for bioremediation. *Geoderma*, in press.
2. Lasat, M.M. (2002). Phytoextraction of toxic metals: a review of biological mechanisms. *J. Environ. Qual.* **31**: 109–120.
3. Romheld, V. and Marschner, H. (1986). Mobilisation of iron in the rhizosphere of different plant species. *Adv. Plant Nutr.* **2**: 155–204.
4. Parkinson, M. (1985). Bio-surfactants. *Biotechnol. Adv.* **3**(1): 65–83.
5. Cameotra, S.S. and Bollag, J.-M. (2003). Biosurfactant-enhanced bioremediation of polycyclic aromatic hydrocarbons. *Crit. Rev. Environ. Sci. Technol.* **30**(2): 111–126.
6. Arima, K., Kakinuma, K., and Tamura, G. (1968). Surfactin, a crystalline peptidolipid surfactant produced by *Bacillus subtilis*: Isolation, characterization and its inhibition of fibrin clot formation. *Biochem. Biophys. Res. Commun.* **31**(3): 488–494.
7. Rosen, M.J. (1978). *Surfactants and Interfacial Phenomena*. John Wiley & Sons, Hoboken, NJ.
8. Mulligan, C.N., Yong, R.N., and Gibbs, B.F. (2001). Heavy metal removal from sediments by biosurfactants. *J. Hazard. Mater.* **85**(1–2): 111–125.
9. Curl, E.A. and Truelove, B. (1986). Root exudates. In: *The Rhizosphere*. E.A. Curl and B. Truelove (Eds.). Springer-Verlag, Berlin, pp. 55–90.
10. Hong, K.-J., Tokunaga, S., and Kajiuchi, T. (2002). Evaluation of remediation process with plant-derived biosurfactant for recovery of heavy metals from contaminated soils. *Chemosphere* **49**(4): 379–387.
11. White, P.J. (2001). Phytoremediation assisted by microorganisms. *Trends Plant Sci.* **6**(11): 502.
12. Whiting, S.N., De Souza, M.P., and Terry, N. (2001). Rhizosphere bacteria mobilize Zn for hyperaccumulation by *Thlaspi caerulescens*. *Environ. Sci. Technol.* **35**(15): 3144.
13. Parker, D.R., Chaney, R.L., and Norvell, W.A. (1995). Equilibrium computer models: applications to plant nutrition research. In: *Chemical Equilibrium and Reaction Models*. R.H. Leoppert, A.P. Schwab, S. Goldberg, and W.I. Madison (Eds.). Soil Science Society of America, pp. 163–200.
14. White, C. and Gadd, G.M. (1998). Reduction of metal cations and oxyanions by anaerobic and metal-resistant microorganisms: chemistry, physiology, and potential for the control and bioremediation of toxic metal pollution. In: *Microbial Life in Extreme Environments*. K. Horikoshi and W.D. Grant (Eds.). Wiley-Liss, Hoboken, NJ, pp. 233–254.
15. Lovley, D. (2000). Fe(III) and Mn(IV) reduction. In: *Environmental Microbe–Metal Interactions*. D. Lovley (Ed.). American Society of Microbiology, Washington, DC, pp. 3–30.
16. Goyal, N., Jain, S.C., and Banerjee, U.C. (2003). Comparative studies on the microbial adsorption of heavy metals. *Adv. Environ. Res.* **7**(2): 311–319.
17. Van Dijk, H. (1971). Cation binding of humic acids. *Geoderma* **5**(1): 53–67.
18. Blake, R.C. et al. (1993). Chemical transformation of toxic metals by a *Pseudomonas* strain from a toxic waste site. *Environ. Toxicol. Chem.* **12**: 1365–1376.
19. Grayston, S.J., Wang, S., Campbell, C.D., and Edwards, A.C. (1998). Selective influence of plant species on microbial diversity in the rhizosphere. *Soil Biol. Biochem.* **30**(3): 369–378.
20. Jones, D. (1998). Organic acids in the rhizosphere—a critical review. *Plant Soil* **205**: 24–44.
21. Harvey, P.J. et al. (2001). Phytoremediation of polyaromatic hydrocarbons, anilines and phenols. *Environ. Sci. Pollut. Resour.*: 1–18.

22. Singer, A.C., Crowley, D.E., and Thompson, I.P. (2003). Secondary plant metabolites in phytoremediation and biotransformation. *Trends Biotechnol.* **21**(3): 123–130.
23. Hadacek, F. (2002). Secondary metabolites as plant traits: current assessment and future perspectives. *Crit. Rev. Plant Sci.* **21**(4): 273–322.
24. Singer, A.C., Thompson, I.P., and Bailey, M.J. (2004). The tritrophic trinity: a source of pollutant-degrading enzymes and its implications for phytoremediation. *Curr. Opin. Microbiol.* **7**(3): 239–244.
25. Frey, M. et al. (1997). Analysis of a chemical plant defense mechanism in grasses. *Science* **277**: 696–699.
26. Bais, H. et al. (2003). Allelopathy and exotic plant invasion: from molecules and genes to species interactions. *Science* **301**: 1377–1380.
27. Bouwmeester, H.J., Matusova, R., Zhongkui, S., and Beale, M.H. (2003). Secondary metabolite signalling in host–parasitic plant interactions. *Curr. Opin. Plant Biol.* **6**(4): 358–364.
28. Walton, B.T., Hoylman, A.M., Perez, M.M., Anderson, T.A., Johnson, D.L., Guthrie, E.A., and Christman, R.F. (1994). Rhizosphere microbial communities as a plant defense against toxic subsances in soils. In: *Microbial Degradation of Organic Chemicals in the Rhizosphere*. T.A. Anderson and J.R. Coats (Eds.). American Chemical Society, Washington, DC, pp. 88–92.
29. Baker, A.J.M. and Brooks, R.R. (1989). Terrestrial higher plants which hyperaccumulate metallic elements—a review of their distribution, ecology and phytochemistry. *Biorecovery* **1**: 81–126.
30. Baker, A.J.M., McGrath, S.P., Reeves, R.D., and Smith, J.A.C. (2000). Metal hyperaccumulator plants: a review of the ecology and physiology of a biological resource for phytoremediation of metal-polluted soils. In: *Phytoremediation of Contaminated Soil and Water*. N. Terry and G.S. Banuelos (Eds.). Lewis Publishers, Boca Raton, FL, pp. 85–107.
31. Delorme, T.A., Gagliardi, J.V., Angle, J.S., and Chaney, R.L. (2001). Influence of the zinc hyperaccumulator *Thlaspi caerulescens J. & C. Presl.* and the nonmetal accumulator *Trifolium pratense L.* on soil microbial populations. *Can. J. Microbiol.* **47**(8): 773–776.
32. Hofflich, G., Weihe, W., and Hect-Bucholz, C. (1995). Rhizosphere colonisation of different crops with growth promoting *Pseudomonas* and *Rhizobium* bacteria. *Microbiol. Res.* **150**: 139–147.
33. Hadas, R. and Okon, Y. (1987). Effect of *Azospirillum brasilense* on root morphology and respiration in tomato seedlings. *Biol. Fertil. Soils* **5**: 241–247.

PHYTOREMEDIATION OF LEAD-CONTAMINATED SOILS

Th.A. Ioannidis
A.I. Zouboulis
Aristotle University of Thessaloniki
Thessaloniki, Greece

Several treatment technologies have been investigated for treating heavymetal-contaminated soils, such as lead. Recently, among the most promising methods is "phytoremediation," the use of certain plants to remove/accumulate or neutralize/degrade pollutants, of inorganic or organic nature. The systematic examination of plants, especially for removing lead from polluted soils has drawn serious attention, especially during the last decade. The overview of the specific technology and a brief description of the results obtained so far is briefly presented. Specific aspects that are of critical importance, such as the chemistry of soil and the interaction with the toxic metals, are also outlined.

SOURCES OF LEAD

Lead (Pb) is a bluish-gray metal that occurs naturally in minute amounts within the earth's crust. It has also been referred to as plumbum, lead metal, and pigment metal. The frequent use of this element in many industrial processes is the main reason for the widespread contamination (especially of soil) by lead. There are a variety of industrial processes that involve the production/use of lead, such as mining, smelting, manufacture of pesticides and fertilizers, dumping of municipal sewage, and burning of fossil fuels that contain a lead additive (leaded gasoline). Many commercial products and materials may also contain lead, including paints, ceramic glazes, television glass, ammunition, batteries, medical equipment, such as X-ray shields or fetal monitors, and electrical equipment. The applications of lead for roofing and for the production of ammunition have increased during the last 20 years. Lead battery recycling sites, of which 29 have been labeled as Superfund sites in the United States, and other manufacturers use more than 80% of the lead produced in the United States. On average, recycled lead products can satisfy only half of total U.S. lead requirements.

HEALTH EFFECTS

Lead has been listed as a potential carcinogen in the EPA Toxic Release Inventory (TRI). Inhalation and ingestion are the two main routes of exposure, whereas the health effects from both are the same. Pb accumulates in the body organs, such as the bones and the brain, which may lead to poisoning (plumbism), or even to death. The gastrointestinal tract, kidneys, and especially the central nervous system are also affected by the presence of lead. Children exposed to lead are at risk for impaired development, lower IQ, shortened attention span, hyperactivity and mental deterioration; children under the age of six are at a more substantial risk. Adults usually experience decreased reaction time, loss of memory, nausea, insomnia, anorexia, and weakness of the joints, when exposed to lead.

CHEMISTRY OF LEAD-CONTAMINATED SOIL

The chemistry of metal interaction with the soil matrix is fundamental to the phytoremediation concept (Fig. 1). Once introduced into the soil matrix, lead is very difficult to remove due to a series of chemical reactions that are likely

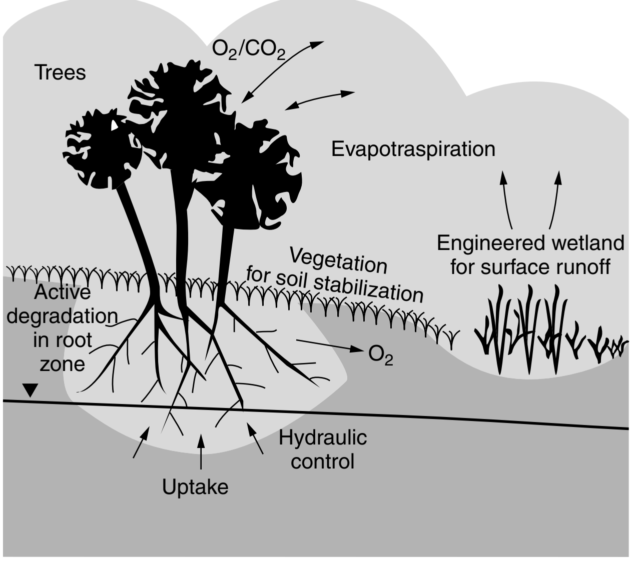

Figure 1. Plant processes that lead to environmental remediation.

to take place, such as mineralization. This transition metal resides within the upper 15–20 cm of the soil, where it is strongly bound by adsorption, ion exchange, precipitation, and complexation with existing organic matter. Lead found within the soil matrix can be classified into six general categories: ionic lead dissolved in soil water, exchangeable, carbonate, oxyhydroxide, organic, or the precipitated fraction. Sequential extractions of soil can be employed to isolate and quantify specific metallic forms, associated with different fractions (1).

The combinations of all these forms make up the total content of lead in soil. Water-soluble and exchangeable lead are the only fractions readily available for uptake by plants, whereas the oxyhydroxide, organically bound (occassionally), hydroxy carbonate, and precipitated forms of lead are the most strongly bound to the soil.

All these interactions that occur throughout the soil matrix are highly pH-dependent. The pH of soil has a significant effect on the mobility of lead from soil, as well as of other metals. The pH of soil generally ranges from 4.0–8.5. Under acidic conditions (pH values lower than 5.5), metal cations are more easily mobile, whereas anions tend to be sorbed to mineral surfaces. Metals become more available to plant roots under these conditions; however, due to the increase in toxic metal solubility, plant growth may be inhibited by phytotoxicity. The opposite occurs when the soil matrix is at an alkaline pH; anions are mobilized and cations are adsorbed to mineral surfaces or form precipitates, hence decreasing the bioavailability of metal cations and their subsequent plant uptake. The capacity of soil to bind lead increases with an increase in pH, cation exchange capacity (CEC), organic carbon content, soil/water E_h values (“mixed” redox potential), and phosphate levels (2).

Another important factor in the phytoremediation process is the presence of specific ionic species of lead, derived by respective hydrolytic reactions. Lead ions exist as Pb^{2+} at low (acidic) pH. As the pH increases (at pH values lower than 6), Pb^{2+} and $[Pb(OH)]^+$ cations are the principal species, but gradually several polynuclear species, including $[Pb_3(OH)_4]^{2+}$, $[Pb_3(OH)_5]^+$, $[Pb_4(OH)_4]^{4+}$, and $[Pb_6(OH)_8]^{4+}$, form. They become abundant at even higher (alkaline) pH levels. Nevertheless, in a recent study, the existence of polynuclear species was seriously questioned; it was concluded that mononuclear complexes $Pb(OH)_q{}^{(2-q)+}$ (aq) (where q = 1–4) are the only hydrolyzed species likely to be significant under typical environmental and biological conditions (4).

The predominant insoluble lead compounds are lead phosphates, lead carbonates and hydroxy carbonates (this form exists when the pH is above 6), and lead (hydr)oxides. Lead sulfide (PbS) is the most stable solid form within the soil matrix and can be formed under reducing conditions, when increased concentrations of sulfides are present. Under anaerobic conditions, a volatile organolead compound (tetramethyl lead) can also be formed by microbial alkylation.

PROPOSED MECHANISMS FOR THE PHYTOREMEDIATION OF LEAD IN CONTAMINATED SOILS

Depending on the nature of the specific pollutant (organic or inorganic), the biology and the properties of the plant and other factors that could act synergistically, such as soil properties, it has been proposed that different mechanisms take place during phytoremediation.

Phytoextraction, rhizofiltration, and phytostabilization are the principal mechanisms that have been proposed, when the pollutants to be removed/treated are inorganic (e.g., toxic metals or radionuclides). Rhizodegradation, phytodegradation, and phytovolatilization are the main mechanisms proposed for the removal/treatment of organic pollutants, whereas the latter mechanism (phytovolatilization) has been also proposed for the removal of certain inorganic pollutants (e.g., Se, Hg, As). Hydraulic control, vegetative cover, and riparian corridors have been suggested as methods for treating inorganic or organic pollutants in contaminated soils; these methods are considered a combination of the aforementioned mechanisms, however, with certain differences that derive mainly from the way that these technologies are engineered (5).

The basic phytoremediation mechanisms, in which lead may be involved, are phytoextraction, rhizofiltration, and phytostabilization.

Phytoextraction

Phytoextraction is the uptake of contaminants by plant roots and their translocation within the plants; the contaminants are generally removed by harvesting the plants (e.g., Fig. 2 for nickel). This concentration technology leaves a much smaller mass to be disposed of than the excavation of soil (applications of *ex situ* treatment technologies), or other media. The specific technology is most often applied to metal-contaminated soil (5).

Lead hyperaccumulation by plants has not been documented in a natural setting. However, certain plants have been identified which can take up lead.

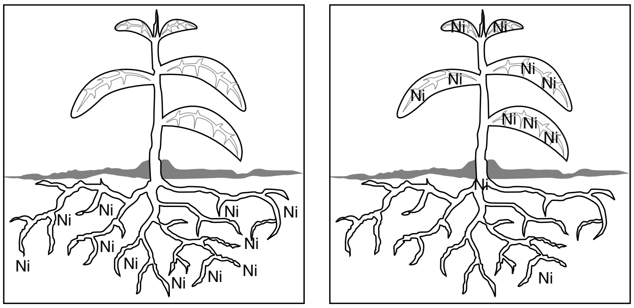

Figure 2. Phytoextraction of nickel from contaminated soil.

Many of these plants belong to the following families: *Brassicaceae, Euphorbiaceae, Asteraceae, Lamiaceae*, and *Scrophulariaceae*. *Brassica juncea*, commonly known as Indian mustard, shows especially good ability to transport lead from the roots to the shoots, which is an important characteristic in the phytoextraction of lead.

The *phytoextraction* coefficient is the ratio of the metal concentration within the biomass of the plant, compared with the metal concentration in the soil. Thus, the greater the coefficient, the greater the uptake of contaminant. Some calculations indicate that *Brassica juncea* can remove 1550 kg of lead per acre. The phytoextraction coefficient for Indian mustard was determined as 1.7; it has also been found that lead concentrations up to 500 mg/L are not phytotoxic to this Brassica species. *Thalspi rotundifolium ssp. cepaeifolium*, which is a noncrop Brassica, commonly known as pennycress, grows in soils contaminated with lead (at concentrations up to 0.82% w/w) and zinc, originating from nearby mine operations. Bench-scale studies have also shown that certain crop plants are capable of phytoextraction. Corn, alfalfa, and sorghum are effective due to their fast growth rate and large amount of biomass produced (2,5).

Rhizofiltration

Rhizofiltration is the adsorption or precipitation of contaminants, which are in the solution surrounding the root zone, onto the plant roots, by biotic or abiotic processes. Plant uptake, concentration, and translocation might occur, depending on the contaminant. Additionally, exudates from the plant roots might cause the precipitation of certain metals, depending on localized conditions. Rhizofiltration results initially in the containment of contaminants, in their immobilization or accumulation on or within the plant roots. Contaminants are then removed by physically removing the plant. Some of the experimental results are summarized in the following:

1. Pb^{2+} at a solution concentration of 2 mg/L accumulated in Indian mustard roots with a bioaccumulation coefficient of 563 after 24 h. Pb^{2+} (at solution concentrations of 35, 70, 150, 300, and 500 mg/L) accumulated in Indian mustard roots, although the root adsorption of Pb reached saturation at 92 to 114 mg Pb/g dry weight of roots. Lead was removed from the 300 and 500 mg/L solutions mainly by precipitation of lead phosphate. Lead absorption by the roots was rapid, although the amount of time required to remove 50% of the Pb from the solution increased, as the Pb concentration increased.
2. Lead accumulated in the roots of Indian mustard in (aqueous) concentrations of approximately 20 to 2,000 g/L with bioaccumulation coefficients of 500 to 2,000.
3. It has been also reported that lead at concentrations of 1 to 16 mg/L was accumulated by water milfoil (*Myriophyllum spicatum*) at a minimum residual concentration below 0.004 mg/L (5).

Phytostabilization

Phytostabilization is defined as (1) the immobilization of a contaminant in soil through absorption and accumulation by roots, adsorption onto roots, or precipitation within the root zone of plants and (2) the use of plants and plant roots to prevent contaminant migration via wind and/or water erosion, leaching, or soil dispersion.

Phytostabilization occurs through root-zone microbiology and chemistry, and/or alteration of the soil environment, or of contaminant chemistry. Soil pH may be changed by plant root exudates or through the production of CO_2. Phytostabilization can change metal solubility and mobility or impact the dissociation of organic compounds. A plant-affected oil environment can convert metals from a soluble to an insoluble oxidation state. Phytostabilization can occur through sorption, precipitation, complexation, or metal valence reduction. Plants can also be used to reduce the erosion of metal-contaminated soil. The term *"phytolignification"* refers to a specific form of phytostabilization, in which organic compounds are incorporated into plant lignin. The contaminants can be also incorporated into the humic material of soils, following a process related to phytostabilization.

It has been reported that the concentration of lead in soil leachate was 22 mg/L, when Indian mustard was present, compared to a concentration of 740 mg/L in soil leachate, without the plants. In another case, mine wastes containing lead were reportedly stabilized by grasses. Soil containing 1660 mg Pb/kg appeared to support less than 50% of plant coverage. Plants in soils containing 323 mg/kg Pb exhibited heavy *chlorosis*, whereas plants could grow in mine wastes, containing up to 4500 mg of Pb/kg of soil (5).

An overview of the correlation between the mechanisms examined and the respective media, the plants used, and the other pollutants that might be treated in parallel, when phytoremediation of lead is applied, is presented in Table 1.

THE ROLE OF SYNTHETIC CHELATES IN PHYTOREMEDIATION OF LEAD

The pH of soils, as well as the specific forms in which lead exists in the soil, might limit the potential bioavailability of lead to plants. The use of certain synthetic chelators

Table 1. Typical Plants and Mechanisms Used in Various Phytoremediation Applications Concerning Lead[a]

Mechanism	Media	Contaminants	Typical Plants
Phytostabilization	Soil, sediments	Metals (Pb, Cd, Zn, As, Cu, Cr, Se, U) Hydrophobic organics (PAHs, PCBs, dioxins, furans, pentachlorophenol, DDT, dieldrin)	Phreatophyte trees to transpire large amounts of water for hydraulic control Grasses with fibrous roots to stabilize soil erosion Dense root systems are needed to sorb/bind contaminants.
Phytoextraction	Soil, brownfields, sediments	Metals (Pb, Cd, Zn, Ni, Cu). Selenium (volatilization)	Sunflowers, Indian mustard, rapeseed plants, barley, hops, crucifers, serpentine plants, nettles, dandelions
Rhizofiltration	Groundwater, water and wastewater in lagoons or created wetlands	Metals (Pb, Cd, Zn, Ni, Cu) Radionuclides (^{137}Cs, ^{90}Sr, U) Hydrophobic organics	Aquatic Plants: — Emergents (bulrush, cattail, coontail, pondweed, arrowroot, duckweed) — Submergents (algae, stonewort, parrot, feather, Eurasian water milfoil, Hydrilla)

[a] References 5 and 6.

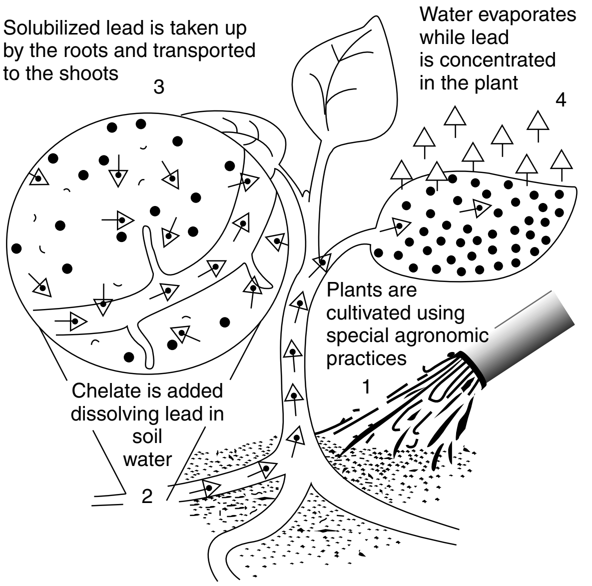

Figure 3. Enhancement of plant uptake of Pb through the use of synthetic chelates (2).

(such as EDTA) by several researchers was examined to overcome this drawback (Fig. 3).

The use of synthetic chelates in phytoremediation increases heavy metal uptake by the plants, through the increase of metal bioavailability, but also increases micronutrients availability, therefore decreasing the possibility of plant nutrient deficiencies. The goal of commercial phytoextraction is to remove or reduce the level of toxic metals within the contaminated soils (*in situ* treatment technology) to meet regulatory standards within 1 to 3 years following the application. The regulatory standard (achievement goal) for lead-contaminated soils set by the U.S. EPA is 500 mg of metal/kg of soil. Plants that can accumulate more than 1% of the target contaminant in the harvestable portion and can produce more than 20 metric tons of shoot biomass per hectare per year are usually required to achieve this goal.

Several researchers have found that this goal can be achieved by applying soil amendments and chemical chelates. Based on relevant studies, it has been shown that only 0.1% of the total amount of lead in contaminated soils exists in soluble form and, therefore, is considered bioavailable to plants for remediation. By adding synthetic chelators, the total amount of (bioavailable) lead in solution can be increased up to 100 times.

Increasing the mobility and bioavailability of lead in the soil through the application of certain chelators, organic acids, or other appropriate chemical compounds allows hyperaccumulation of metals in some plants. For lead, a number of different chelators have been tested, such as EDTA (ethylene dinitrilotetraacetic acid), CDTA (trans-1,2-cyclohexylene-dinitrilotetraacetic acid), DTPA (diethylene trinitrilopentaacetic acid), EGTA (ethylebis-[oxyethylene-trinitrilo]-tetraacetic acid), HEDTA (hydroxyethylethylene dinitrilotriacetic acid), citric acid, and malic acid. The addition of these chelates resulted in enhanced shoot lead concentrations (2). The soil additive EDTA can form hexahedral metal ion complexes with commonly found metals, such as Pb[II] or Fe[II], which can enhance their uptake by roots and translocation throughout the plant (7).

Although there are certain advantages in using synthetic chelates, environmental concerns for their impact on contaminated sites still need research. The major concern for using chelates to enhance phytoremediation and to increase the bioavailability of toxic metals, such as Pb, is the fear of lead leaching or running off into the ground or surface water, creating secondary contamination. By making the metals more soluble in the soil matrix, leaching is more probable and threatens to contaminate nearby water sources (2). The use of synthetic chelates necessitates careful mass balances to confirm that metals mobilized by chelators are not subsequently leached to groundwater.

The use of synthetic chelators (complex chemical compounds) is a questionable practice from the environmental standpoint. Natural chelators originating from plants or from microorganisms seem to be more promising. The

Table 2. Advantages and Disadvantages/Limitations of the Phytoremediation Process[a]

Advantages	Disadvantages/Limitations
Amenable to a variety of organic and inorganic compounds.	Restricted to sites with shallow contamination within rooting zone of remediative plants.
In situ/*ex situ* applications for treating contaminated soils.	May take up to several years to remediate a contaminated site.
In situ applications decrease the amount of soil disturbance, compared to conventional methods.	Restricted to sites with rather low contaminant concentrations.
Reduces the amount of waste to be landfilled (up to 95%).	Harvested plant biomass from phytoextraction may be classified as a hazardous waste.
In situ applications decrease the spread of contaminant via air and water.	Climatic conditions.
Does not require expensive equipment or highly specialized personnel.	Introduction of nonnative species may affect biodiversity.
Easy to implement and maintain.	Consumption of contaminated plant tissue is also of concern.
Low cost compared to conventional treatment methods.	
Environmentally friendly and aesthetically pleasing to the public.	

[a] Reference 2.

microbial protection of plants against phytotoxic concentrations, however, is often based on excluding heavy metals; therefore, it is uncertain whether an approach based on chemical chelators is practical for improving phytoremediation (8). Naturally secreted organic acids, such as mugenic and avenic acids, may exhibit the same mechanism (similar to EDTA) to scavenge soils for metals (7).

ADVANTAGES AND LIMITATIONS

The phytoremediation of lead involves a number of advantages and certain limitations. Both of them derive from the fact that the use of plants to treat lead-contaminated soils would affect the soil and the surrounding media; as a result, some other pollutants (e.g., other metals) might be treated as well. For this reason, the advantages and limitations of lead phytoremediation are generally regarded as similar to those found in all relevant phytoremediation applications.

The most important advantages are the low cost and the aesthetic result, which is the sight of plants. On the other hand, the most important limitations are that the biodiversity of the site might be altered and that the uptake of pollutants by the plants may change their biology (2,5,9). The specific factors that can be characterized as advantages or limitations of phytoremediation are summarized in Table 2.

CURRENT STATUS AND FUTURE TRENDS

Although the phytoremediation of lead-contaminated soils is a rather new technology, the initial results from the respective laboratory research in pilot plants as well as from field applications, can be characterized as rather promising (5). Nevertheless, a better insight that will illuminate the overall picture will be given after at least 10 years, when the field data would be adequate to help in drawing safe conclusions.

Although the research effort for already existing plants is ongoing, the use of genetic engineering for plant improvement has been proposed. However, before such plants can be used for the phytoremediation of contaminated sites, the questions that arise from these applications should be carefully answered in an appropriate time frame following extensive research.

BIBLIOGRAPHY

1. Tessier, A., Campbell, P., and Bisson, M. (1979). Sequential extraction procedure for the speciation of particulate trace metals. *Anal. Chem.* **51**: 844–850.
2. Henry, R.J. (2000). *An Overview of The Phytoremediation of Lead and Mercury*. National Network of Environmental Management Studies (NNEMS), Washinghton, DC.
3. Baes, F.C. and Mesmer, E.R. (1976). *The Hydrolysis of Cations*. John Wiley & Sons, New York.
4. Perera, W.N., Hefter, G., and Sipos, P.M. (2001). An investigation of the lead (II)-hydroxide system. *Inorg. Chem.* **40**: 3974–3978.
5. EPA. (2000). *Introduction to Phytoremediation*. EPA Report Cincinnati, OH.
6. Schnoor, L.J. (1997). *Phytoremediation*. Technology Overview Report, TO-98-01. GWRTAC, Pittsburgh, PA.
7. Meagher, B.R. (2000). Phytoremediation of toxic elemental and organic pollutants. *Curr. Opin. Plant Biol.* **3**: 153–162.
8. Van Der Lelie, D., Schwitzguebel, J.P., Glass, D.J., and Vangronsveld, B.A. (2001). Assessing phytoremediation's progress in the United States and Europe. *Environ. Sci. Technol.* **35**: 446–452.
9. Miller, R.R. (1996). *Phytoremediation*. Technology Overview Report, TO-96-03. GWRTAC, Pittsburgh, PA.

PHYTOREMEDIATION OF METHYL *TERTIARY*-BUTYL ETHER

S.K. Winnike-McMillan
Q. Zhang
L.C. Davis
L.E. Erickson
J.L. Schnoor
(from *Phytoremediation: Transformation and Control of Contaminants*, Wiley 2003)

SUMMARY OF PRACTICAL IMPLICATIONS

The capability of plants to remediate methyl *tertiary*-butyl ether (MTBE) at contaminated sites is promising to the point that at least one field test is underway and others anticipated. Both hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings and alfalfa (*Medicago sativa*) facilitate the removal of MTBE from saturated soil in the laboratory. The primary removal of MTBE is by soil volatilization facilitated by plant transpiration. Very little of the contaminant is metabolized or accumulated in laboratory studies. The MTBE that volatilizes and the small amount that diffuses from stems or transpires through leaves do not represent a risk in the atmosphere due to the rapid dispersion and degradation. Despite the different behavior of MTBE in groundwater, plants, and the atmosphere, concurrent phytoremediation of gasoline spills containing MTBE and other contaminants is expected to be very useful in the future.

INTRODUCTION

Methyl *tertiary*-butyl ether was first developed as a fuel oxygenate in the late 1970s to enhance the octane of gasoline and to improve air quality. In 1990, the U.S. Clean Air Act Amendments required fuel oxygenates to be added to gasoline in an effort to reduce atmospheric concentrations of carbon monoxide and ozone in some urban areas. The use of MTBE increased 16-fold from 1984 to 1993, to 24 billion pounds (approximately 10×10^6 metric tons). By 1998, MTBE was the fourth most common chemical product in the U.S. (Johnson et al. 2000). Oxygenates are added to over 30 percent of the gasoline used in the U.S., and MTBE is the most commonly used of these additives (Squillace et al., 1996). After formulation, MTBE constitutes up to 15 percent of gasoline by weight, giving rise to a daily usage of more than 10 million gallons (3.8×10^7 liters) in the U.S. (Johnson et al. 2000). However, continued use of MTBE is being questioned because of reports of adverse health effects, uncertainty about the efficacy of air pollution control, and the potential for contamination of groundwater (Anonymous 1997, Mehlman 1998). Several U.S. states have required that MTBE be phased out over time. A recent proposal was that the U.S. President use the authority under the Toxic Substances Control Act to ban MTBE across the U.S., but production continued in late 2000 at a rate of more than 11 million gallons or nearly 40 million liters per day.

As a part of the U.S. Geological Survey National Water Quality Assessment program from 1993 to 1994, 60 volatile organic compounds were analyzed for occurrence in groundwaters and surface waters. Of these 60 organic compounds, MTBE was the second most frequently detected chemical in samples of shallow groundwater from U.S. urban areas (Zogorski et al. 1998). Very little was detected in agricultural areas or deep groundwater wells. Surveys have reported finding MTBE in 5 to 10 percent of U.S. potable groundwater supplies (Stuckey 2000). Based on current distribution of MTBE sources and well pumping rates, Johnson et al. (2000) assessed the likelihood that more wells will become contaminated by MTBE. Although new and much more extensive databases are needed, the projected number of U.S. wells likely to experience contamination is in the thousands due to a large fraction of all community water supply wells lying within 1 kilometer of leaking underground storage tanks.

Methyl *tertiary*-butyl ether contamination of soil and groundwater can be separated into two categories: point sources, such as leaking underground storage tanks, and nonpoint sources, including infiltration and diffusion from urban air into shallow groundwater. General urban air contamination arises from fugitive emissions at gasoline stations and from gasoline-powered engines. Point sources are typically known quantifiable oxygenated gasoline spills onto the ground or leaks from underground storage tanks. Contamination by MTBE commonly occurs in conjunction with BTEX gasoline constituents (benzene, toluene, ethyl benzene, and xylenes). Gasoline from leaking underground storage tanks is the most frequent source of significant groundwater pollution with as much as one million gallons (approximately 4 million liters) of gasoline released into the ground each year (Hartley and Englande 1992), implying MTBE releases of hundreds of tons. It is estimated that out of three million underground storage tanks in use in the U.S., as many as 500,000 may be leaking petroleum into the ground (Hartley and Englande 1992). Actual leakage may be much higher than this implied level of 2 gallons per tank per year (8 liters per tank per year).

Properties of Methyl Tertiary-Butyl Ether

Methyl *tertiary*-butyl ether is an aliphatic ether with a molecular weight of 88.15 grams per mole (Fig. 1) (Squillace et al. 1997). This compound is volatile, colorless, and flammable (Caprino and Togna 1998). Because chemical structure influences many of the properties such as solubility, volatilization potential, and resistance to microbial degradation, it is useful to compare the physical and chemical properties of MTBE with those of the BTEX compounds with which it usually occurs in the environment (Table 1).

There is good evidence for microbial degradation of MTBE by several processes, as well as field data showing natural attenuation. Several processes for MTBE have been identified.

Anaerobic Microbial Transformation

Mormile et al. (1994) tested both pure and mixed cultures to anaerobically degrade MTBE. Two acetogens, *Acetobacterium woodii* and *Eubacterium limosum*, known to degrade phenyl methyl ethers, showed to MTBE transformation. However, the microbial cells consumed other gaseous substrates from the headspace of the incubation system and produced formate and acetate with no apparent toxicity caused by MTBE (Mormile et al. 1994). The capability of mixed cultures to anaerobically transform MTBE and *tertiary*-butyl alcohol was investigated under denitrifying and methanogenic conditions, with rates of transformation increasing with nutrients added (Yeh and Novak 1994). Methyl *tertiary*-butyl ether was shown to degrade in the soil taken from only one site out of three and only when nutrients were added. The degradation rates were much slower than those of the *tertiary*-butyl alcohol and complete degradation took

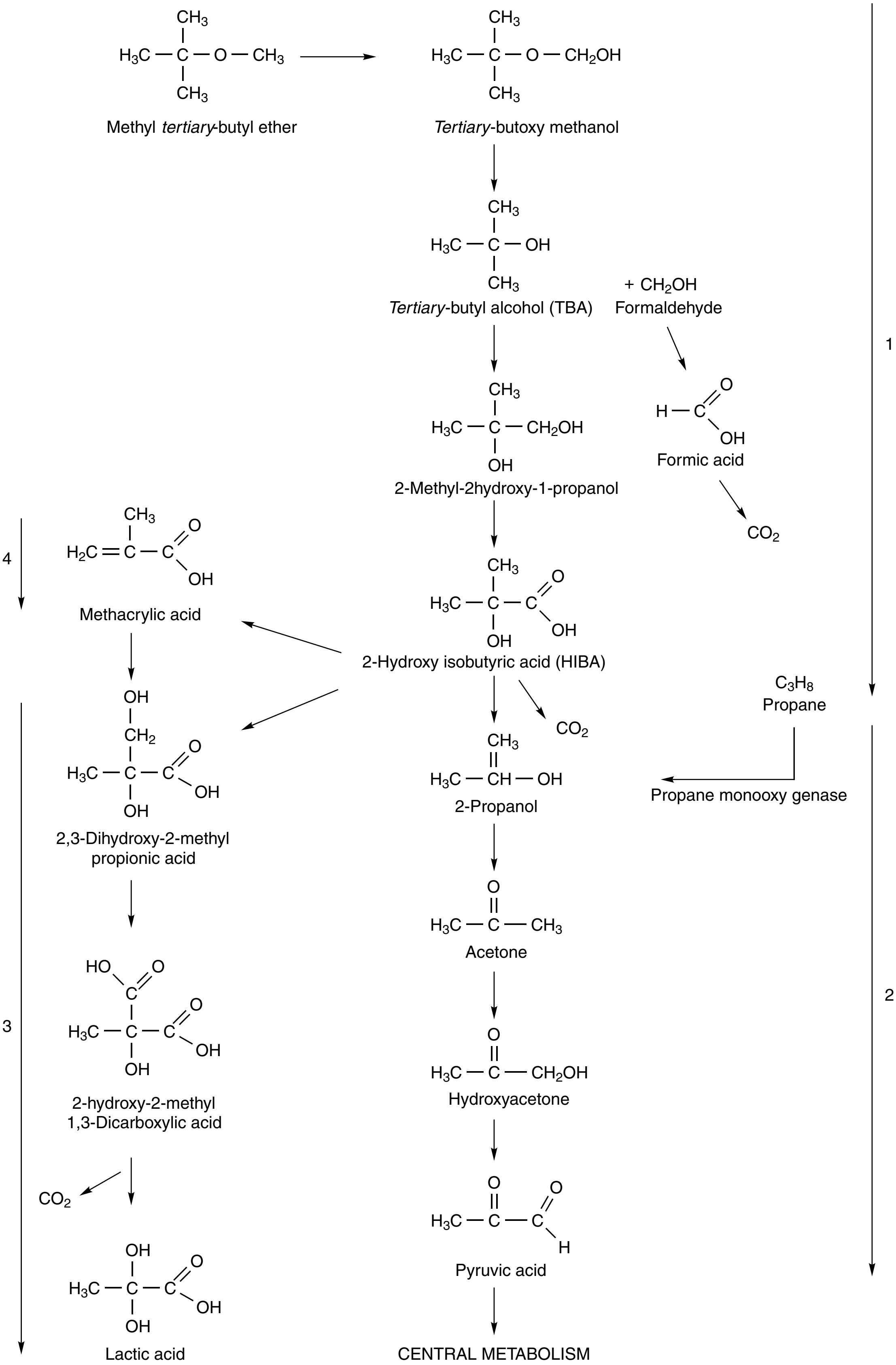

Figure 1. Methyl *tertiary*-butyl ether degradation pathway proposed by the investigation of propane oxidizing bacteria. Adapted from Steffan et al. (1997).

Table 1. Chemical and Physical Properties of Gasoline, Methyl *Tertiary*-Butyl Ether, Benzene, and Toluene

Property	Gasoline	Methyl *tertiary*-butyl ether	Benzene	Toluene
Molecular weight, grams per mole	≈100	88.15	78.11	92.13
Specific gravity	0.72	0.744	0.88	0.8669
	0.74	0.7404 to 0.7578	0.8765	
Boiling temperature, °C		53.6 to 55.2	80.1	110.6
Water solubility, milligrams per liter	100 to 200	43 000 to 54 300	1780	534.8
			1791	
Vapor pressure at 25 °C, millimeters mercury		245 to 251	95.19	28.4
Log K_{ow}		1.20	2.13	2.73
		0.94 to 1.30	1.56 to 2.15	2.11 to 2.80
Henry's law constant, atmospheres-cubic meters per gram mole		5.87×10^{-4}	5.43×10^{-3}	5.94×10^{-3}
		1.4×10^{-3}		
		3×10^{-3}		
		5.28×10^{-4}		
Dimensionless Henry's law constant (milligrams per liter of air) per (milligrams per liter of water)		0.0240	0.222	0.243
		0.0572		
		0.123		
		0.0216		
		mean: 0.0564		
Log K_{oc}		1.05	1.1 to 2.5	1.56 to 2.25
		$K_{oc} \approx 11$ cubic centimeters per gram	median = 1.9	
			$K_{oc} \approx 80$ cubic centimeters per gram	

K_{ow} is the dimensionless octanol–water partitioning coefficient.
K_{oc} is organic carbon normalized sorption coefficient.
Log is the base 10 logarithm.
Note: One atmosphere is 1.0133×10^{-5} pascals in the Henry's law constant and 1 millimeter of mercury of vapor pressure is 133.32 pascals. Adapted from Squillace et al. (1997) and references cited therein.

approximately 270 days. More recently Landmeyer et al. (1998) and Finneran and Lovley (2001) observed anaerobic degradation of MTBE under Fe(III) reducing conditions and Bradley et al. (2001) observed MTBE biodegradation under nitrate-reducing conditions

Aerobic Transformation in Mixed Cultures

Aerobic microbial transformation of MTBE was first successfully shown by Salanitro et al. (1994) using a mixed culture isolated from the sludge of an industrial chemical biotreatment operation. This culture, consisting of organisms from *Corynebacterium, Pseudomonades*, and *Achromobacters* was able to utilize MTBE at 2 milligrams per liter as a sole carbon source although none of the organisms were shown to grow alone on MTBE. Approximately 80 percent of the [^{14}C]-MTBE was mineralized to $^{14}CO_2$ or incorporated into new microbial cells after an incubation period of approximately 8 hours (Salanitro et al. 1994). *Tertiary*-butyl alcohol was observed as an intermediate. With 20 milligrams of MTBE per liter, little or no [^{14}C]-MTBE was metabolized to $^{14}CO_2$ or cellular material, suggesting that MTBE may be inhibitory to aerobes at high concentrations (Salanitro et al. 1994).

A consortium of aerobes from refinery activated sludge was also effective in transforming MTBE (Cowan and Park 1996), as was an aquifer consortium studied by Borden et al. (1997). Bradley et al. (1999) and Landmeyer et al. (2001) observed rapid metabolism of [^{14}C]-labeled MTBE using naturally occurring streambed microorganisms under mixed aerobic and anaerobic conditions in microcosms. Eweis et al. (1998) examined a compost-based biofilter microbial community and found MTBE degradation.

Aerobic Cometabolism

A fungus of the genus *Graphium* has been shown capable of cometabolizing MTBE with butane (Hardison et al. 1997), and Garnier et al. (1999) observed cometabolism with pentane by *Pseudomonas*. Steffan et al. (1997) reported indigenous propane-oxidizing bacteria isolated from contaminated and uncontaminated soils that could degrade MTBE. Isolate ENV425 that degrades MTBE was identified by fatty acid analysis to be most closely related to members of the genus *Nocardia*. All isolates were first grown with propane, then fed concentrations of MTBE in the range of 20 to 100 milligrams per liter, which was used as a sole carbon and energy source (Steffan et al. 1997). The cells were fed 20 milligrams per liter metabolized MTBE at a rate of at least 3.8 nanomoles per minute per milligram of cells without a significant lag period. Complete degradation occurred after approximately 4 hours, with *tertiary*-butyl alcohol accumulating nearly stoichiometrically, and with evolution of $^{14}CO_2$ from the methyl group. The microorganisms were capable of transforming *tertiary*-butyl alcohol at a rate of 1.8 nanomoles per minute per milligram of cells, about half that of

MTBE. Following 10 hours of transformation of *tertiary*-butyl alcohol, approximately 70 percent of the radiolabeled compound was still present in solution, suggesting the formation of water-soluble *tertiary*-butyl alcohol metabolites. Continued incubation showed 60 percent conversion to $^{14}CO_2$ after 24 hours and greater than 90 percent conversion after several days (Steffan et al. 1997). Figure 1 is a pathway for MTBE degradation proposed by Steffan et al. (1997).

Aerobic Transformation in Pure Cultures Lacking Other Substrates and Prior Induction

Four pure cultures of diverse genera have been shown capable of transforming at least a fraction of MTBE in a defined medium (Mo et al. 1997). The three strains investigated by Mo et al. (1997) represented *Methylobacterium, Rhodococcus*, and *Arthrobacter*. Also, a proteobacteria strain was derived from the biofilter examined by Eweis et al. (1998).

Potential Biotreatments for Methyl Tertiary-Butyl Ether

Salanitro et al. (2000) have successfully used bioaugmentation and oxygenation with a consortium designated MC-100 to enhance degradation of MTBE within an aquifer. Landmeyer et al. (2001) successfully used oxygenation only to enhance biodegradation of an MTBE-contaminated aquifer. Successful treatment was achieved in laboratory studies through cometabolism in an upflow fluidized-bed bioreactor (Stringfellow and Oh 2002) and in a vapor phase treatment using a trickling filter (Fortin and Deshusses 1999). In the latter case, contaminated soil from two different locations was used as an inoculum. Prince (2000) has reviewed most of the earlier literature on microbial MTBE degradation. Despite several instances of successful metabolism, MTBE appears to transform only slowly under typical field spill conditions, resulting in a need for additional cost-effective treatment methods.

A potentially useful alternative is phytoremediation. Plants have the benefit of absorbing large quantities of subsurface water by solar driven pumping (transpiration) and the roots support diverse microbial populations (Shimp et al. 1993, Davis et al. 1998a, Narayanan et al. 1999). Deep-rooted poplar (*Populus* spp.) trees have been studied in several contexts but there is little information on the capability to transpire or metabolize MTBE. Davis et al. (1998b) showed by use of Fourier transformed infrared spectroscopy that MTBE passes freely through hybrid poplars (*Populus deltoides* × *Populus nigra* ‘DN34’ ‘Imperial Carolina’) in 1- to 2-day experiments although no effort was made to determine metabolic processes that might occur within the plant. Landmeyer et al. (2000) detected MTBE in cores of trees above a gasoline-contaminated aquifer but no effort was made to track the movement through the trees.

Rubin and Ramaswami (2001) have also investigated the uptake and phytoremediation of MTBE using hybrid poplar (*Populus deltoides* × *Populus nigra* ‘DN34’, & IC) cuttings. They reported transpiration of MTBE through plants with no evidence of MTBE transformation in the poplar (*Populus* spp.) saplings (Ramaswami and Rubin 2001).

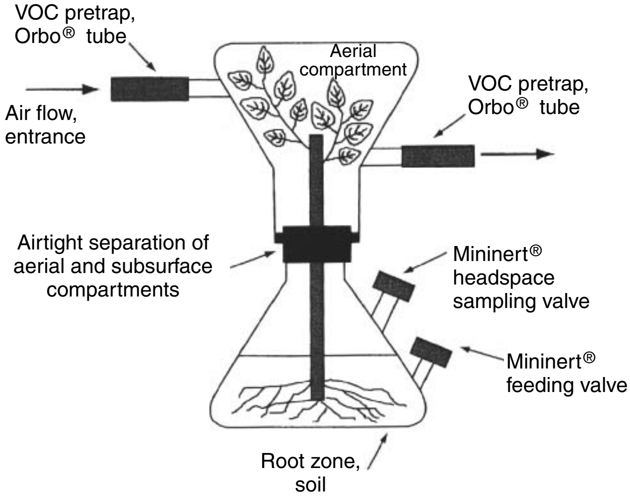

Figure 2. Schematic of the reactor design used to investigate hybrid poplar (*Populus deltoides* × *Populus nigra* ‘DN34’ ‘Imperial Carolina’) cuttings growing in soil.

UPTAKE OF METHYL TERTIARY-BUTYL ETHER BY HYBRID POPULUS IN HYDROPONIC SOLUTION

Experimental Setup

To determine the potential for hybrid poplars (*Populus deltoides* × *Populus nigra* ‘DN34’ ‘Imperial Carolina’) to take up, metabolize, and translocate MTBE, a hydroponic experiment was performed, measuring the capability of hybrid poplar (*Populus deltoides* × *Populus nigra* ‘DN34’ ‘Imperial Carolina’) cuttings to remove [^{14}C]-MTBE from aqueous solution. Uptake of [^{14}C]-radiolabeled MTBE by hybrid poplar (*Populus deltoides* × *Populus nigra* ‘DN34’ ‘Imperial Carolina’) cuttings was investigated in a manner similar to methods pioneered by Burken (1993) and Burken and Schnoor (1996). A schematic of the reactor design is provided in Fig. 2. However, in the first experiment, there was no attempt made to collect gases volatilizing from plant leaves. The study included varying concentrations of MTBE in solution to determine any toxic response. Hydroponic studies used one-quarter-strength Hoagland’s inorganic nutrient solution, fed to the root zone of batch reactors containing hybrid poplar (*Populus deltoides* × *Populus nigra* ‘DN34’ ‘Imperial Carolina’) cuttings. Eight-inch (200-millimeter) long cuttings (0.75 to 1 inches or 19 to 25 millimeters in diameter) were affixed with pre-drilled screw caps and pre-drilled Teflon®-lined septa. Teflon® tape was wrapped around the stem and acrylic caulk was used to seal the caps and septa to the cutting. Hybrid poplar (*Populus deltoides* × *Populus nigra* ‘DN34’ ‘Imperial Carolina’) cuttings were first rooted in hydroponic solution until a vascular root system appeared, then placed in 1-liter bioreactors containing 400 milliliters of nutrient solution. The bioreactors consisted of a 1-liter Erlenmeyer flask modified by attaching a sampling port to the bottom and top of each flask. The cuttings were maintained in a laboratory growth chamber at 28 °C under artificial grow lights that provided a photon flux at the leaf surface of 100 to 160 micromoles of photons per square meter per second, the full photosynthetically active range.

Controls were included in triplicate to distinguish the uptake of MTBE by hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings from losses out of the system. First, a capped control with only nutrient solution and [^{14}C]-MTBE was included to determine the contribution of bio-degradation possibly occurring inside the hydroponic solution and sorption to the glass. To determine loss of compound through leaks in the system, controls were used with solid glass rods in place of a tree cutting. Finally, two sets of planted cuttings were cut off just above the cap to create decapitated controls, which were necessary to observe the effect of an excised tree on [^{14}C]-MTBE removals. Two sizes of cuttings were used; the first was approximately the same diameter as the glass rod and the second was similar to the intact tree reactors.

After being sealed, all reactors were spiked with a combination of [^{14}C]-radiolabeled MTBE ("hot," *i.e.*, radioactive), 6 microcuries per reactor, and pure MTBE ("cold," *i.e.*, nonradioactive) to a total aqueous concentration of 10.18 milligrams per liter, (creating a cold:hot MTBE ratio of 12 037:1). Concentrations of cold MTBE were then increased to 100, 1000, and 10 000 milligrams per liter. Reactors were prepared for all concentrations in triplicate. Samples were collected daily to monitor the disappearance of [^{14}C]-MTBE from hydroponic solution. Transpiration was measured gravimetrically each day, and replacement nutrient solution was added through the lower Mininert® port. Prior to adding nutrient solution, all headspace gas surrounding the root system was removed through a series of traps to capture any MTBE that was present in the headspace above the hydroponic solution. The air was pumped out through an activated carbon trap capturing pure compound and other organic metabolites (Orbo® tube 32, large, Supelco), and a trap that allowed air to bubble through 1 molar sodium hydroxide (NaOH) to capture any carbon dioxide (CO_2). To extract the trapped MTBE, the activated carbon was placed into 4-milliliter glass vials and submerged in 2 milliliters of methanol (Fisher) for 24 to 48 hours. Samples of 100 microliters were taken and injected into 15 milliliters of Scintiverse® for counting on a liquid scintillation counter. One milliliter samples were taken from the sodium hydroxide (NaOH) in the carbon dioxide (CO_2) traps and injected into an Ultima Gold scintillation cocktail (Packard) for scintillation counting.

After 10 days, total [^{14}C]-uptake was quantified in the roots, lower stem, upper stem, leaves, and petioles of the hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings *via* oxidation to $^{14}CO_2$ in a R.J. Harvey Bio-Oxidizer, and subsequently by scintillation counting. Oxidation efficiency was determined to be 92 percent and was calibrated with pre- and post-analysis efficiency testing.

Results from Hydroponic Solutions

Based on recorded observations of visual symptoms, biomass, and transpiration rate, MTBE was found not toxic to hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings at concentrations as high as 1000 milligrams of MTBE per liter. However, the trees dosed with 10 000 milligrams of MTBE per liter died after only a few days. Transpiration rate correlated well with biomass measurements and was a good measure of growth and vitality. Hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings were healthy and grew well in all experiments at concentrations as high as 1000 milligrams of MTBE per liter.

Figure 3 displays removal of [^{14}C]-MTBE from solution by the full and excised cuttings as well as the controls. The data for the full ($n = 5$) and excised ($n = 3$) cuttings are shown as an average with the error bars representing one standard deviation. Uptake of MTBE occurred immediately in all planted reactors and continued throughout the investigation. For the five planted reactors, 66.98 ± 6.09 percent of the total MTBE was removed from hydroponic solution at the end of the 10-day experiment.

Methyl *tertiary*-butyl ether removal after 10 days for the excised cuttings was 42.97 ± 3.32 percent. The data from the toxicity portion of the experiment yielded a comparable result for the cuttings dosed with 10 000 milligrams per liter—51.52 ± 2.29 percent was removed from solution after 11 days. The removal rate was similar for the excised cuttings ($k = 0.060 \pm 0.002$ per day and the correlation coefficient $r^2 = 0.994$), and for the cuttings that died at 10 000 milligrams of MTBE per liter ($k = 0.067$ per day) as evidenced by the first order rate coefficients. The reproducibility of the removal rates and percent removed from solution are a strong indication that the cause is attributable to movement of MTBE through the dead stem as opposed to leaks from the enclosed system. Excised cuttings are apparently capable of volatilizing MTBE (and very little water) from solution to the atmosphere.

After 10 days, the experiment was terminated and the cuttings were sacrificed. The various components were bio-oxidized and the radioactivity measured with the liquid scintillation counter to determine the amount of radiolabeled compound in the plant tissue. Figure 4

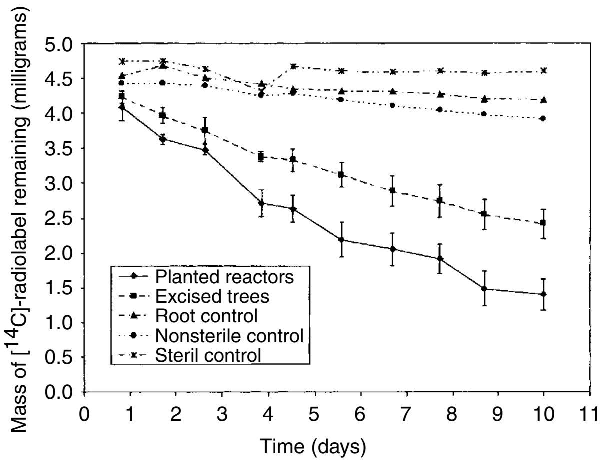

Figure 3. Hydroponic experiment with hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings: mass of methyl *tertiary*-butyl ether remaining in aqueous solution as determined by liquid scintillation counting. At time = 0, 4.3 milligrams of methyl *tertiary*-butyl ether was added for an initial concentration of 10.2 milligrams per liter in solution. Error bars on the data for full trees ($n = 5$) and excised trees ($n = 3$) represent one standard deviation.

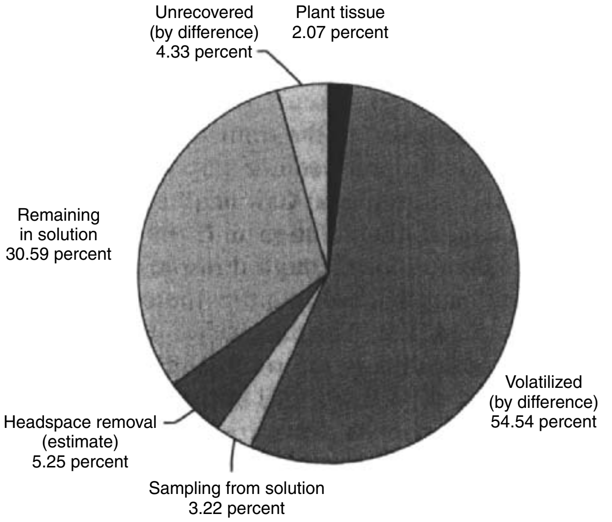

Figure 4. Hydroponic experiment: graph depicting [^{14}C]-radiolabel distribution and recovery results for the reactors containing hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings as determined by liquid scintillation counting at the termination of the 10-day investigation. Results are given as a percent of the total [^{14}C]-MTBE initial spike (6 microcuries) added to reactors at time = 0. The amount of radiolabel that was reported as volatilized was determined as the difference in the amount remaining in solution at the termination of the experiment between the unsterilized and planted reactors. The amount reported as unrecovered was the difference between the unsterilized and sterile reactors.

summarizes the radiolabel recoveries for the full cutting and displays the overall mass balance. The amount of MTBE volatilized through plants was found to be 54.5 percent. This was estimated by the difference in the amount of radiolabel remaining in solution between the planted cuttings and the sterile controls. This estimate assumes that any removal from the excised tree reactors was due to uptake and not to leaks (Winnike 1998). However, leaks of MTBE through the septums and seals can be substantial, and leaks have been shown to rival transpiration of MTBE through leaves in such experiments (Hong et al. 2001).

The average [^{14}C]-radiolabel remaining in solution was 33.02 ± 6.09 percent. This is approximately twice as much as that found in the toxicity investigations in the same reactors at the same concentration. This value was expected to be similar due to the similar experimental conditions and duration. The reason for this difference lies in the relationship between cumulative transpiration and uptake of MTBE. A comparison of average transpiration rates shows that the cuttings in the toxicity experiment transpired more (32.3 ± 11.24 milliliters per day) than the cuttings in the hydroponic experiment (26.5 ± 10.0 milliliters per day). Because the leafy portion was exposed to the atmosphere, the amount of radiolabeled MTBE transpired through the leaves was not measured directly.

UPTAKE EFFICIENCY AND TOXICITY

Methyl *tertiary*-butyl ether was taken up by the plants and translocated up the vascular system of the plant in the transpiration stream. The efficiency of uptake of MTBE from hydroponic solution by hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings was estimated by the change in MTBE mass in the solution and the volume fraction of water transpired each day. Because of the high transpiration rates of all cuttings, it was necessary to add nutrient solution to the reactors daily, diluting the initial concentration of MTBE. This resulted in a continual decrease in concentration and mass of MTBE in the root compartment. For chemicals that are readily taken up and not transformed in the roots, the efficiency of uptake is identical to the dimensionless transpiration stream concentration factor (*TSCF*). The *TSCF* is defined as the chemical concentration in the xylem water divided by the chemical concentration in the external solution (Briggs et al. 1982). The TSCF was estimated to be in the range of 0.5 to 0.8 for MTBE for hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings based on

$$U = (TSCF)\,(C)\,(T) \tag{1}$$

in which U is the rate of mass uptake (milligrams per day), C is the concentration in solution (milligrams per milliliter), and T is the rate of transpiration (milliliters per day). The concentrations used to estimate *TSCF* were based on initial experimental concentrations of 10, 100, and 1000 milligrams per liter. The *TSCF* value was not calculated for the toxic concentration of 10 000 milligrams per liter.

UPTAKE OF METHYL TERTIARY-BUTYL ETHER BY HYBRID POPULUS PLANTED IN SOIL

Experimental Setup

To assess the capability of hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings to remove MTBE from soil, an Iowa agricultural silt-loam soil (Nodeway-Ely series) was taken from an uncontaminated field site with a fraction organic matter of 0.025 (Burken 1993). The soil was air-dried, pulverized, and passed through a 2-millimeter sieve. The cuttings were allowed to root hydroponically for 2 weeks until a full vascular root system developed. Four cuttings were transplanted to individual reactors containing 500 grams of air-dried soil, and 180 milliliters of nutrient solution was added intermittently to the flask to achieve uniform soil moisture content (Fig. 2). The reactors were initially at 100 percent saturation and acclimated for 7 days. During this time, the trees transpired and the soil-water saturation level was reduced 70 to 90 percent.

After this acclimation, the leafy portion of the cutting was enclosed by attaching an inverted 1-liter Erlenmeyer flask modified with an inlet port for air flow near the top and an outlet near the bottom. The perimeter of the open end was sealed with acrylic caulk to fit snugly around the screw cap. Parafilm® tape was used to seal this joint, and a Mininert® valve was placed on the top sampling port of the bottom flask to create a fully encapsulated system. Reactors were treated with doses of 3.95 milligrams of cold MTBE and 7.1 microcuries of [^{14}C]-MTBE mixed with the feed water (resulting in a ratio of cold:hot MTBE of 10 069:1). Air was purified and pumped through the aerial portion of the reactor at 1.0 to 3.0 liters per minute. The air inlet and outlet ports of the Erlenmeyer flask were fitted with activated carbon traps to capture [^{14}C] transpired through the plants (two traps were used in series on the outlet port to ensure capture). The scrubbed air then passed through traps to capture any $^{14}CO_2$ that was mineralized by soil microorganisms or by the plant in the aerial portion of the reactor.

Results with Hybrid Populus in Soil

The primary goal of this test was to quantify volatilization of [^{14}C]-MTBE from leaf tissue. By enclosing the leafy portion of the cutting, a total mass balance was achieved. The total recovery was lower than expected (67.4 ± 6.3 percent). However, this test provided valuable information with regard to the major pathways for MTBE removal. The unrecovered portion seemed to be partially the result of leaks from the enclosed reactor vessel (Fig. 2). The four planted cuttings remained healthy throughout the test, which was terminated after 33 days because the leaves grew too large and were pressing into the sides of the aerial compartment. The longer duration of this experiment (compared to the prior experiment in hydroponic solution) was possible because of slower growth of the cuttings in the soil. Transpiration was monitored by gravimetric analysis.

The rate of transpiration varied widely throughout the test because of excess humidity in the aerial compartment. As humidity increased to high levels, transpiration decreased dramatically due to aerodynamic resistance. In addition, the average transpiration rates for all reactors were significantly lower than under hydroponic conditions as a result of the decreased water availability in the soil. Although [^{14}C]-MTBE uptake was not quantified daily, the fluctuation in transpiration rates very likely influenced the rate of MTBE uptake from the soil water. The full set of transpiration data is provided in Winnike (1998).

Radiolabel recoveries on a percent mass basis are presented in Fig. 5. Because only liquid scintillation counting was utilized, it was uncertain whether the radiolabel recovered was [^{14}C]-MTBE or a [^{14}C]-metabolite. Only a small percentage of the initial [^{14}C]-radiolabel remained in the soil (4.75 ± 1.76 percent), indicating that 33 days were adequate to remove a majority of the MTBE from the soil. In subsequent tests, the vast majority of the radiolabel exists as parent compound, MTBE, in hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') tissues.

The [^{14}C]-radiolabel that was transpired constituted the most significant flux of MTBE movement (55.55 ± 1.85 percent). Assumedly the recovered radiolabel from the activated carbon traps was [^{14}C]-MTBE. Subsequent analyses in the laboratory indicated that MTBE was predominant over other potential metabolites such as *tertiary*-butyl alcohol. The air also bubbled through a sodium hydroxide (NaOH) trap to capture $^{14}CO_2$ from the mineralization of MTBE. A small percent of the initial MTBE input was recovered as $^{14}CO_2$ (1.62 ± 0.63 percent), which indicates that either hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings or microorganisms in the soil were able to mineralize a portion of the MTBE to carbon dioxide (CO_2).

The headspace in the root compartment was sampled and analyzed for both [^{14}C] and $^{14}CO_2$ prior to watering and at the end of the test. This amount was shown to be fairly significant with recoveries of 0.63 ± 0.55 percent and 2.03 ± 1.85 percent for [^{14}C] and $^{14}CO_2$, respectively. The observed range of 0.41 to 4.50 percent demonstrates the high variability in $^{14}CO_2$ recovery. The amount of [^{14}C] remaining in plant tissue was approximately 3 percent. On the basis of total plant mass, the leaves contained 14.72 ± 4.02 milligrams of MTBE per kilogram, the petioles 11.10 ± 3.86 milligrams of MTBE per kilogram, and roots ±3.76 milligrams of MTBE per kilogram. The leaves contained the largest total percentage (34.27 ± 9.55 percent) and the roots accounted for the smallest (11.14 ± 2.12 percent).

CHANNEL TESTS WITH ALFALFA

Tests were carried out in a 6-channel system (Zhang et al. 1998). Table 2 provides information on the experimental system and the tests that were conducted. The results are described fully elsewhere (Zhang et al. 1998, Zhang 1999, Zhang et al. 1999, 2000, 2001). Table 3 and Fig. 6 show that vegetation affects the upward movement of MTBE from the saturated zone to the soil surface. Channel 4 has a much smaller fraction of MTBE and water moving to the soil surface. The recovery of MTBE was larger for Channel 4, which was barren of vegetation and unseeded with microbial cultures capable of biodegrading MTBE. As shown in Table 3, the measured loss of MTBE to the atmosphere at the soil surface was much greater than that which entered the plant and then passed into the atmosphere. The fraction of water that was lost as vapor to the atmosphere was considerably larger than the measured fraction of MTBE that entered the atmosphere (Zhang et al. 1999). The total recoveries of MTBE in the vegetated channels were less than 100 percent for all of the tests. The incomplete recovery of MTBE was probably due to enhanced rhizosphere biodegradation in the soil. Microcosm experiments with soils from the channel system were conducted, and a disappearance rate for MTBE of up to 5 milligrams per kilogram of soil per day was observed.

The water transpired daily by the plants increased from the time the alfalfa (*Medicago sativa*) was cut until the next harvest as shown in Fig. 7. The transpiration rate was on average 5 millimeters per day (0.5 cubic centimeter per square centimeter per day) at the end of July 1999. This value was in agreement with predicted transpiration rates (Davis et al. 1998a). Figure 8 provides information on the concentration of MTBE in the plant water as a function of plant height. Experiments conducted with plant stems show that MTBE diffuses through the stem surface into the atmosphere (Zhang et al. 2001). The estimated values of the diffusivity for MTBE range from 8×10^{-8} to 8×10^{-7} square centimeters per second for alfalfa (*Medicago sativa*) plant stems (Zhang et al. 2001).

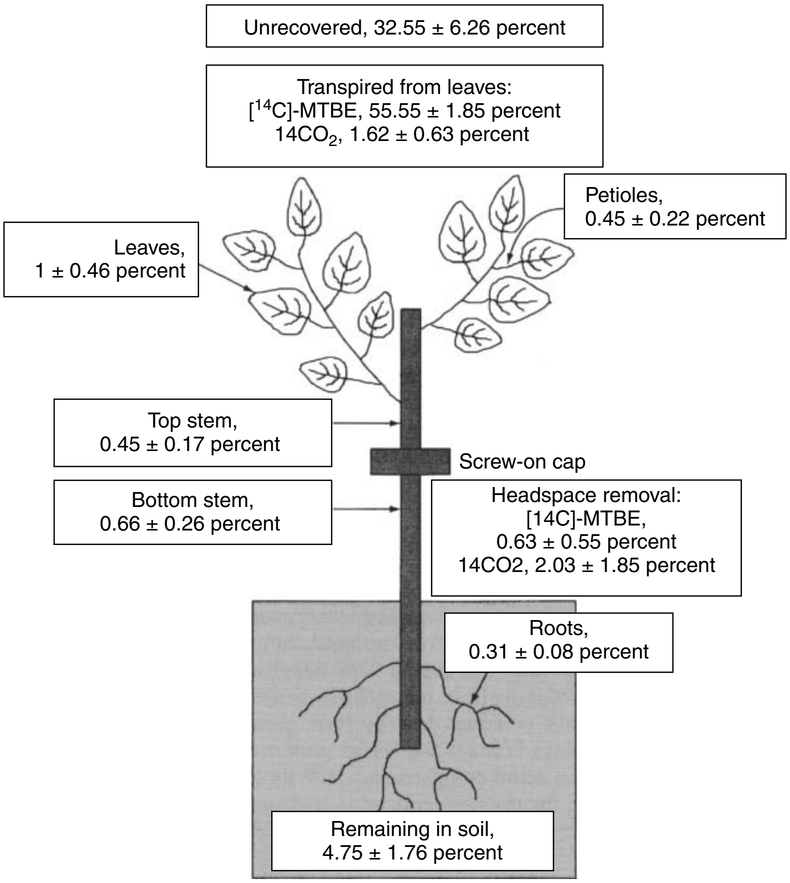

Figure 5. Schematic and mass balance of methyl *tertiary*-butyl ether uptake by hybrid poplars (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') in soil reactor.

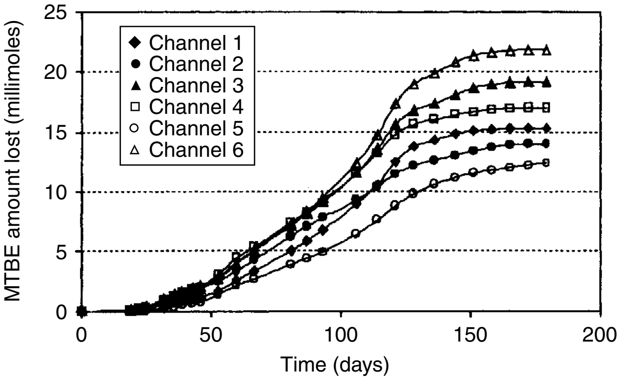

Figure 6. Cumulative loss of methyl *tertiary*-butyl ether to the atmosphere through the soil surface for each of six channels. Tests conducted from June 30 to December 25, 1998.

FATE OF METHYL TERTIARY-BUTYL ETHER IN THE ATMOSPHERE

Site-specific determinations are necessary to ensure that the amount of MTBE transferred to the atmosphere does not exceed air quality standards. This analysis takes into account the total amount of compound volatized per unit time and the mixing zones in the atmosphere. This can be compared to the large quantities of MTBE introduced into the atmosphere because of gasoline pumping at refueling stations and engine operation. The MTBE introduced to the atmosphere caused by phytoremediation is significantly less.

Table 2. Experimental Conditions and Treatments in Each Channel for the Six-Channel Experimental System

	Planted with 10 alfalfa (*Medicago*) *sativa* plants	Air sparged at 0.24 liter per channel per day	Additional bacteria (2 liters of suspension added on June 20, 1998)
Channel 1	Yes	Yes	#33 (*Rhodococcus*)
Channel 2	Yes	No	#33 (*Rhodococcus*)
Channel 3	Yes	No	No
Channel 4	No	No	No
Channel 5	Yes	No	#41 (*Arthrobacter*)
Channel 6	Yes	Yes	#41 (*Arthrobacter*)

Note: Channel dimensions were 1.1 meters long, 0.65 meter deep, and 0.10 meter wide. The soil was alluvial silty sand soil (with less than 10 percent silt) collected near a landfill in Riley County, Kansas. The system was illuminated with 12 cool white fluorescent lights (40 watts) at a height of 0.50 meter above the soil; light intensity was approximately 160 micromoles of photons or microeinsteins per square meter per second. The channels were fed with an MTBE-water solution of 0.84 millimole per liter at 1 per day from June 30, 1998 to September 20, 1998 (83-day duration). Distilled water was fed from September 21 to December 25, 1998 (96-day duration) and the channels fed with an MTBE-water solution of 0.84 millimole per liter at 1 per day from Feb. 6, 1999 until the data in Figs 7 and 8 were collected during the summer of 1999. Strain #33 (*Rhodococcus*) and Strain #41 (*Arthrobacter*) were obtained from C.F. Kulpa (Mo et al. 1997, Zhang et al. 1999).

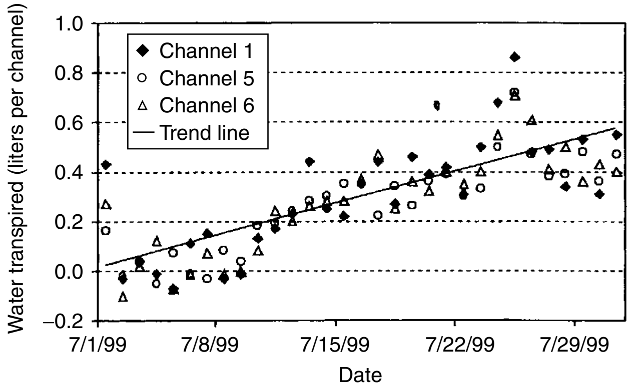

Figure 7. Water transpired daily by alfalfa (*Medicago sativa*) in three channels during July 1999; plants were cut back to 60 millimeters on July 1.

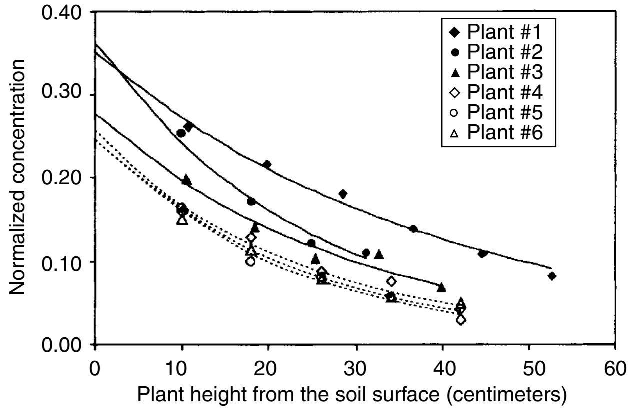

Figure 8. Methyl *tertiary*-butyl ether dimensionless concentration in plant water as a function of alfalfa (*Medicago sativa*) stem position from the soil surface. Dimensionless feed concentration of methyl *tertiary*-butyl ether is 1 which corresponds to 0.84 millimoles per liter.

The MTBE released to the atmosphere is rapidly dissipated. The published information on the degradation half-life of MTBE in the atmosphere indicates that MTBE readily degrades chemically because of the hydroxyl radicals (OH^-) present. Squillace et al. (1997) have reviewed the fate of MTBE in the atmosphere and report that the half-life of MTBE can be as short as 3 days at 25 °C.

The following method to estimate the concentration of MTBE being released through vegetation is based on mass balance considerations. If the MTBE flows upward into the unsaturated zone together with the water drawn up by transpiration, the air required to hold the associated water vapor is the critical control volume. For example, consider the case in which there is 88 milligrams of MTBE per liter in the groundwater, and the water and MTBE move into the atmosphere together at 25 °C where the vapor pressure of water is 3173 pascals (23.8 millimeters of mercury) or a mole fraction of 0.0313 at saturation in air with 1×10^5 pascals (1 atmosphere) pressure. A liter of water contains 55.6 gram moles which requires 43,434 liters of air to dissolve the water under saturation conditions at 25 °C and 1×10^5 pascals (1 atmosphere) of pressure. The 88 milligrams of MTBE per liter is 1 millimole per liter in water, but only $0.001/1776 = 0.56$ parts per million by volume in air. Thus, the concentration of MTBE in air is diluted because of the water vapor dissipation process that partially controls the rate of evapotranspiration. Because (1) MTBE concentrations in groundwater are often much smaller than 88 milligrams per liter, (2) the air is usually not saturated with water vapor, and (3) the ambient air contains some water vapor, the values of MTBE in the atmosphere as a result of phytoremediation are very small. This is especially true compared to values at a refueling station where MTBE vapors in equilibrium with the liquid fuel may be as high as 50,000 parts per million by volume.

FIELD STUDIES

The basic processes observed at the laboratory scale (*i.e.*, biodegradation, volatilization, and transpiration), also take place at a field scale. Salanitro et al. (2000) observed aerobic biodegradation in the field with and without bioaugmentation using MTBE-degrading microorganisms. Wilson et al. (2000) reported biodegradation under anaerobic conditions. Hong et al. (2001) demonstrated the beneficial contributions of vegetation with respect to plume control. Landmeyer et al. (2000) reported the uptake of MTBE into vegetation.

Phytoremediation is an effective technology for remediation of fuel-contaminated sites as shown elsewhere in this book. Based on the results of laboratory and field studies reviewed in this chapter, vegetation increases the rate of evapotranspiration and the upward movement of MTBE into the unsaturated zone and the atmosphere. The beneficial effects of vegetation for MTBE are similar to those for BTEX compounds in fuel-contaminated soils; however, MTBE tends to persist longer in the atmosphere compared to BTEX compounds because MTBE is more resistant to photochemical degradation. Furthermore, MTBE is more mobile in soil and groundwater, and MTBE is not degraded readily in the root zone. Where these different properties of BTEX and MTBE are taken into account, simultaneous phytoremediation of shallow groundwater plumes from fuel spills should be feasible at many sites (see Weaver et al. this book).

CONCLUSIONS

Methyl *tertiary*-butyl ether is a compound of concern because of possible carcinogenic effects, persistence in soil and groundwater, and potential resistance to degradation by microorganisms. The widespread occurrence of this compound in groundwater is relatively recent, and conventional treatment is either ineffective or expensive. Phytoremediation provides an inexpensive, esthetically pleasing, and effective alternative for treating numerous organic contaminants. These investigations show that hybrid poplar (*Populus deltoides* × *Populus nigra* ‘DN34’ ‘Imperial Carolina’) cuttings and alfalfa (*Medicago sativa*) are capable of cleaning up water and soil contaminated with MTBE. Specific conclusions are as follows:

- Hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings tolerate concentrations of MTBE of at least 1000 milligrams per liter. The concentrations observed in the environment near a source of MTBE contamination can be as high as 200 milligrams per liter and hybrid poplars (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') could effectively take up MTBE at this concentration without apparent toxic effects.
- Small hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings are effective at removing MTBE from hydroponic solution. Tests demonstrated that cuttings dosed with 10, 100, and 1000 milligrams per liter were able to achieve an MTBE removal of 30.6 percent remaining in solution within 10 days from 400-milliliter reactors.
- Small hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings and alfalfa (*Medicago sativa*) are effective at removing MTBE from soil. Hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') cuttings removed 95.2 ± 1.8 percent within 33 days when MTBE was initially at a soil water concentration of 22 milligrams per liter (soil concentration of 7 milligrams per kilogram). The evapotranspiration rate and amount of MTBE volatilized are considerably larger in alfalfa (*Medicago sativa*) vegetated soil compared to soil without vegetation. The evapotranspiration rate is small after alfalfa (*Medicago sativa*) plants are cut back; the rate increases nearly linearly with time as the plants increase in size over 29 days.
- The primary removal mechanism for MTBE was volatilization into the soil gas and atmosphere. When soil volatilization was not possible, significant amounts also transpired through hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') leaves and stem. Shown to move with water into the plant (*TSCF* of 0.5 to 0.8), MTBE also volatilized through the stem of excised small hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') plants and alfalfa (*Medicago sativa*) stems. Removal occurred at a slower rate than for actively transpiring hybrid poplar (*Populus deltoides* × *Populus nigra* 'DN34' 'Imperial Carolina') and alfalfa (*Medicago sativa*) plants. However, the volatilization is typically expected to be very small, especially compared to the MTBE volatilized at gasoline refueling stations. The small amount of MTBE that does volatilize into the atmosphere disperses and degrades photochemically with a half-life on the order of days.

This research provides evidence that phytoremediation could be a potential alternative for the remediation of MTBE contaminated sites.

Acknowledgments

We thank the U.S. Environmental Protection Agency Hazardous Substances Research Center, Great Plains and Rocky Mountain Region, and Equilon Enterprises LLC for financial support of this chapter. We benefited from collaboration and discussion with Marjorie Hong, Walter Farmayan, and Lee Newman.

READING LIST

Anonymous (1997). Fueling the debate. *Environ. Health Persp.* **105**(10): 1042–1043.

Borden, R.C., Daniel, R.A., LeBrun, L.E., and Davis, C.W. (1997). Intrinsic biodegradation of MTBE and BTEX in a gasoline-contaminated aquifer. *Water Resour. Res.* **33**: 1105–1115.

Bradley, P.M., Landmeyer, J.E., and Chappelle, F.H. (1999). Aerobic mineralization of MTBE and *tertiary*-butyl alcohol by streambed microorganisms. *Environ. Sci. Technol.* **33**: 1877–1879.

Bradley, P.M., Chapelle, F.H., and Landmeyer, J.E. (2001). Methyl *t*-butyl ether mineralization in surface-water sediment microcosms under denitrifying conditions. *Appl. Environ. Microbiol.* **67**(4): 1975–1978.

Briggs, G.G., Bromilow, R.H., and Evans, A.A. (1982). Relationships between lipophicity and root uptake and translocation of non-ionized chemicals by barley. *Pestic. Sci.* **13**: 495–504.

Burken, J.G. (1993). *Vegatative uptake by Populus spp. and mineralization of atrazine in variable soil types*. Masters thesis. University of Iowa, Iowa City.

Burken, J.G. and Schnoor, J.L. (1996). Phytoremediation: plant uptake of atrazine and role of root exudates. *J. Environ. Eng.* **122**(11): 958–963.

Caprino, L. and Togna, G.I. (1998). Potential health effects of gasoline and its constituents: a review of current literature (1990–1997) on toxicological data. *Environ. Health Persp.* **106**(3): 115–125.

Cowan, R.M. and Park, K. (1996). Biodegradation of the gasoline oxygenates MTBE, ETBE, TAME, TBE and TAA by aerobic mixed cultures. *Hazard. Ind. Wastes* **28**: 523–530.

Davis, L.C., Banks, M.K., Schwab, A.P., Narayanan, M., Tracy, L.E. Erickson, and J.C., (1998a). Plant-based remediation. In: *Bioremediation: Principles and Practice. Vol. 2, Biodegradation Technology Developments*. S.K. Sikdar and R.L. Irvine (Eds.). Technomic, Lancaster, Pennsylvania, pp. 183–219.

Davis, L.C., Vanderhoof, S., Dana, J., Selk, K., Smith, K., Goplen, B., and Erickson, L.E. (1998b). Movement of chlorinated solvents and other volatile organics through plants monitored by Fourier transform infrared (FT-IR) spectrometry. *J. Hazard. Sub. Res.* 4–1: 1–26.

Eweis, J.B., Schroeder, E.D., Chang, D.P.Y., and Scow, K.M. (1998). Biodegradation of MTBE in a pilot-scale biofilter. In: *Natural Attenuation: Chlorinated and Recalcitrant Compounds*. G.B. Wickramanayake and R.E. Hinchee (Eds.). Battelle Press, Columbus, Ohio, pp. 341–346.

Finneran, K.T. and Lovley, D.R. (2001). Anaerobic degradation of methyl *tert*-butyl ether (MTBE) and *tert*-butyl ether (TBE). *Environ. Sci. Technol.* **35**: 1785–1797.

Fortin, N.Y. and Deshusses, M.A. (1999). Treatment of methyl *tert*-butyl ether vapors in biotrickling filters. 1. reactor startup, steady-state performance and culture characteristics. *Environ. Sci. Technol.* **33**: 2980–2986.

Garnier, P.M., Auria, R., Auger, C., and Revah, S. (1999). Cometabolic biodegradation of methyl *t*-butyl ether by *Pseudomonas aeruginosa* grown on pentane. *Appl. Microbiol. Biotechnol.* **51**(4): 498–503.

Gomez, M.M., Abernathy, C.O., and Du, J.T. (1997). *Drinking water health advisory for methyl tertiary-butyl ether*. American Chemical Society Division of Environmental Chemistry Preprints of Papers, Washington, DC, pp. 370–372.

Hardison, L.K., Curry, S.S., Ciuffetti, L.M., and Hyman, M.R. (1997). Metabolism of diethyl ether and cometabolism of methyl *tert*-butyl ether by a filamentous fungus, a *Graphium* sp. *Appl. Environ. Microbiol.* **63**: 3059–3067.

Hartley, W.R. and Englande, A.J., Jr. (1992). Health risk assessment of the migration of unleaded gasoline—a model for petroleum products. *Water Sci. Technol.* **25**: 65–73.

Hong, M.S., Farmayan, W.F., Dortch, I.J., Chiang, C.Y., Schnoor, S.K., and McMillan, J.L. (2001). Phytoremediation of MTBE from a groundwater plume. *Environ. Sci. Technol.* **35**: 1231–1239.

Johnson, R., Pankow, J., Bender, D., Price, C., and Zogorski, J. (2000). MTBE: To what extent will past releases contaminate community water supply wells. *Environ. Sci. Technol.* **34**: A210–A217.

Landmeyer, J.E., Chapelle, F.H., Bradley, R.M., Pankow, J.F., Church, C.D., and Tratnyek, P.G., (1998). Fate of MTBE relative to benzene in a gasoline-contaminated aquifer (1993–1998). *Ground Water Mon. Rem.* **18**: 93–102.

Landmeyer, J.E., Vroblesky, D.A., and Bradley, P.M. (2000). MTBE and BTEX in trees above gasoline-contaminated ground water. In: *Case Studies in the Remediation of Chlorinated and Recalcitrant Compounds*. G.B. Wickramanayake, A.R. Gavaskar, J.T. Gibbs and J.L. Means (Eds.). Battelle Press, Columbus, Ohio, pp. 17–24.

Landmeyer, J.E., Chapelle, F.H., Herlong, H.E., and Bradley, P.M. (2001). Methyl *tert*-butyl ether biodegradation by indigenous aquifer microorganisms under natural and artificial oxic conditions. *Environ. Sci. Technol.* **35**: 1118–1126.

Mehlman, M.A. (1998). Editorial: pollution by gasoline containing hazardous methyl *tertiary*-butyl ether (MTBE). *Arch. Environ. Health* **53**(4): 245–246.

Mo, K., Lora, C.O., Wanken, A.E., Javanmardian, M., Yang, X., and Kulpa, C.F. (1997). Biodegradation of methyl *t*-butyl ether by pure bacterial cultures. *Appl. Microbiol. Biotechnol.* **47**: 67–72.

Mormile, M.R., Liu, S., and Suflita, J.M. (1994). Anaerobic biodegradation of gasoline oxygenates: extrapolation of information to multiple sites and redox conditions. *Environ. Sci. Technol.* **28**: 1727–1732.

Narayanan, M., Erickson, L.E., and Davis, L.C. (1999). Simple plant-based design strategies for volatile organic pollutants. *Environ. Progress* **18**: 231–242.

Prince, R.C. (2000). Biodegradation of methyl *tertiary*-butyl ether (MTBE) and other fuel oxygenates. *Crit. Rev. Microbiol.* **26**: 163–178.

Ramaswami, A. and Rubin, E. (2001). Measuring phytoremediation parameters for volatile organic compounds: focus on MTBE. *Pract. Periodical Hazard. Toxic Radioact. Waste Manag.* **5**: 123–129.

Rubin, E. and Ramaswami, A. (2001). The potential for phytoremediation of MTBE. *Water Res.* **35**: 1348–1353.

Salanitro, J.P., Diaz, L.A., Williams, M.P., and Wisniewski, H.L. (1994). Isolation of a bacterial culture that degrades methyl *t*-butyl ether. *Appl. Environ. Microbiol.* **60**: 2593–2596.

Salanitro, J.P., Johnson, P.C., Spinnler, G.E., Maner, P.M., Bruce, H.L. Wisniewski, and C. (2000). Field-scale demonstration through aquifer bioaugmentation and oxygenation. *Environ. Sci. Technol.* **34**: 4152–4162.

Schnoor, J.L., Licht, L.A., McCutcheon, S.C., Wolfe, N.L., and Carreira, L.H. (1995). Phytoremediation of organic and nutrient contaminants. *Environ. Sci. Technol.* **29**(7): 318–323.

Shimp, J.F., Tracy, J.C., Davis, L.C., Lee, E., Huang, W., Schnoor, L.E. Erickson, and J.L. (1993). Beneficial effects of plants in the remediation of soil and groundwater contaminated with organic materials. *Crit. Rev. Environ. Sci. Technol.* **23**(1): 41–77.

Squillace, P.J., Zogorski, J.S., Wilber, W.G., and Price, C.V. (1996). Preliminary assessment of the occurrence and possible sources of MTBE in groundwater in the United States, 1993–1994. *Environ. Sci. Technol.* **30**: 1721–1730.

Squillace, P.J., Pankow, J.F., Korte, N.E., and Zogorski, J.S. (1997). Review of the environmental behavior and fate of methyl *tert*-butyl ether. *Environ. Toxicol. Chem.* **16**(9): 1836–1844.

Steffan, R.J., McClay, K., Vainberg, S., Condee, C.W., and Zhang, D. (1997). Biodegradation of the gasoline oxygenates methyl *tert*-butyl ether, ethyl *tert*-butyl ether, and *tert*-amyl methyl ether by propane-oxidizing bacteria. *Appl. Environ. Microbiol.* **63**: 4216–4222.

Stringfellow, W.T. and Oh, K.C. (2002). Initiation of MTBE biotreatment in fluidized-bed bioreactors. *J. Environ. Eng.* **128**(9): 852–861.

Stuckey, H.T. (2000). The benefits and problems associated with MTBE. *Environ. Protect.* **11**(7): 49–52.

Wilson, J.T., Cho, J.S., Wilson, B.H., and Vardy, J.A. (2000). *Natural attenuation of MTBE in the subsurface under methanogenic conditions*. EPA/600/R-00/006. National Risk Management Research Laboratory, U.S. Environmental Protection Agency, Cincinnati, Ohio. (http://www.epa.gov/ORD/NRMRL/mtbe/MTBEreport.pdf).

Winnike, S.K. (1998). *Phytoremediation of methyl tert-butyl ether (MTBE) by hybrid poplar trees*. Masters thesis. University of Iowa, Iowa City.

Yeh, C.K. and Novak, J.T. (1994). Anaerobic biodegradation of gasoline oxygenates in soils. *Water Environ. Res.* **66**: 752–774.

Zhang, Q. (1999). *Phytoremediation of methyl tert-butyl ether (MTBE) in groundwater–experimental and modeling studies*. Ph.D. dissertation. Kansas State University, Manhattan, Kansas.

Zhang, Q., Davis, L.C., and Erickson, L.E. (1998). Using vegetation to treat methyl *tert*-butyl ether contaminated groundwater. In: *Proceedings of the 1998 Conference on Hazardous Waste Research–Bridging Gaps in Technology and Culture*. U.S. Environmental Protection Agency Hazardous Substance Research Center: Great Plains/Rocky Mountains, Kansas State University, Manhattan, Kansas. Held in Snowbird, Utah, May 18–21. pp. 262–272 (http://www.engg.ksu.edu/HSRC).

Zhang, Q., Davis, L.C., and Erickson, L.E. (1999). An experimental study of phytoremediation of methyl *tert*-butyl ether (MTBE) in groundwater. *J. Hazard. Sub. Res.* **2**: 4–1 to 4–19. (http://www.engg.ksu.edu/HSRC).

Zhang, Q., Davis, L.C., and Erickson, L.E. (2000). Plant uptake of methyl *tert*-butyl ether (MTBE) from groundwater. In: *Environmental and Pipeline Engineering 2000*. R.Y. Surampalli (Ed.). American Society of Civil Engineers, Reston, Virginia, pp. 166–173.

Zhang, Q., Davis, L.C., and Erickson, L.E. (2001). Transport of methyl *tert*-butyl ether (MTBE) through alfalfa plants. *Environ. Sci. Technol.* **35**: 725–731.

Zogorski, J.S., Delzer, G.C., Bender, D.A., Squillace, P.J., Baehr, T.J. Lopes, A.L., Stackelberg, P.E., Landmeyer, J.E., Lico, C.J. Boughton, M.S., Pankow, J.F., Johnson, R.L., and

N.R. Thomson (1998). MTBE: summary of findings and research by the U.S. Geological Survey. In: *1998 Annual Conference of the American Water Works Association*, Denver, Colorado.

PHYTOREMEDIATION OF SELENIUM-LADEN SOILS

Rupali Datta
Dibyendu Sarkar
University of Texas at San Antonio
San Antonio, Texas

Selenium (Se) is a naturally occurring metalloid that is toxic at high concentrations but is also a nutritionally essential element. In the 1980s, Se-contaminated water at the natural wildlife refuge at Kesterson Reservoir, California was responsible for widespread reproductive failure, developmental defects, and death in migratory aquatic birds and fish. The environmental and health impacts of this incident have resulted in extensive research on the behavior of Se in soils and plants. The cost of using conventional remediation technologies for cleaning up Se-contaminated soils is very high, so recent research has focused on phytoremediation of Se from contaminated soils. Selenium accumulation by plants and its subsequent volatilization are very effective in remediating Se-contaminated soil and water. Natural hyperaccumulator plants such as *Astragalus* growing in seleniferous soils accumulate high concentrations of Se in their tissues. However, the main drawback of using these plants for phytoremediation is their slow growth rate and low biomass. This has prompted researchers to investigate the possibility of developing genetically engineered plants with better Se phytoextracting capacity. Understanding the speciation and bioavailability of Se in soils and the biochemistry of Se in plants and microorganisms is the first step toward developing effective phytoremediation strategies for Se-laden soils.

INTRODUCTION

Selenium is a naturally occurring metalloid, essential for animals, humans, and microorganisms in trace amounts (1). It provides protection against diseases such as liver necrosis, muscular dystrophy, Keshan disease, infertility, and white muscle in livestock (2). At higher doses, it is toxic to animals and humans. The concentration range of Se from trace element requirement to lethality is very narrow. The minimal nutritional requirement for animals is about 0.05–0.10 mg Se/kg dry forage feed, but to 2–5 mg Se/kg dry forage is toxic (3). A chronic form of Se poisoning in livestock is 'alkali disease' caused by ingestion of plants containing high concentrations of Se. Excess Se in diet causes birth defects and sterility in animals; loss of hair, teeth, nail, fatigue; and even death in humans (4).

Selenium is found primarily in sedimentary rock formations in drier areas of the world. It is chemically very similar to sulfur (S); it has five valence states; selenide (2−), elemental selenium (0), thioselenate (2+), selenite (4+), and selenate (6+) (5). Environmental speciation of Se depends on redox conditions and pH; selenate is the major species in aerobic and neutral to alkaline environments, but selenide and elemental Se dominate under anaerobic conditions. Selenium complexes with metals and nonmetals, forming both organic and inorganic compounds (3). Interest in Se has grown during the past couple of decades, primarily due to the Kesterson Reservoir controversy in the 1980s. Selenium-contaminated water at the natural wildlife refuge at Kesterson Reservoir, California, was responsible for reproductive failure, developmental defects, and death in migratory aquatic birds (6) and fish (7). Scientists, regulators, politicians, and the general U.S. public became aware of the dangers of Se contamination as a result of this incident. Selenium toxicity is encountered in several arid and semiarid regions of the world that have seleniferous, alkaline soils derived from weathering of seleniferous rocks and shales. Se leaching from these soils is accelerated by irrigation, and it accumulates in high concentrations in drainage water, mainly as selenate. Selenium is also released into the environment by various industries such as oil refineries and electrical utilities (3).

Conventional remediation technologies for cleaning up Se-contaminated soils depend on physicochemical and electrochemical approaches, which are often prohibitively expensive. Moreover, these methods produce hazardous by-products, which must be transported to landfills, thereby adding to the remediation cost. Phytoremediation is the use of plants to extract, accumulate, and volatilize Se (8). It offers a cheaper, safer, and environmentally friendly solution for cleaning up Se-contaminated soils. This article provides an overview of the various sources of Se contamination, factors affecting Se bioavailability in soils, Se biochemistry and the mechanism of Se accumulation in plants, and future directions in the improvement of phytoremediation through biotechnology.

SOURCES OF SELENIUM CONTAMINATION

Selenium contamination in the environment arises from both natural and anthropogenic sources. Soils that are derived from shale rocks contain high concentrations of Se. As a result of irrigation, Se is leached into the subsurface drainage water and is eventually channeled into evaporation ponds. Evaporation and continual drainage into these ponds, builds up the amount of bioavailable Se to very high concentrations, which can be very damaging to wildlife, which was first demonstrated by the environmental disaster at the Kesterson Reservoir in California (9). Selenium-contaminated irrigation drainage water represents one of the most serious problems confronting agriculture in the western United States and other parts of the world whose geological and environmental conditions are similar (9).

Anthropogenic Se contamination develops from aqueous discharges of electrical power plants, coal ash leachates, oil refinery effluents, and other industrial wastewater (4). Selenium is widely used in electronics, photography, and in the production of steel, glass, ceramics, pigments, paints, pharmaceuticals, rubber, fungicides, and insecticides (10).

Of the 1430 hazardous waste sites on the United States Environmental Protection Agency's National Priorities List, at least 376 have measurable levels of Se (11).

SELENIUM IN SOILS

Most uncontaminated soils contain 0.10 to 2 ppm of total Se (12). Excessive amounts of Se is present only in the semiarid and arid regions of the world in soils derived from cretaceous shales. However, total Se in soils is not a reliable index of the actual amount available to plants. Normal soils contain <50 ppb of water-soluble Se. The chemical form of Se in soils largely determines its mobility and bioavailability to plants and animals. Selenium species in soil solution can be present primarily in three oxidation states; selenates, selenites, and selenides (12). In the oxidized alkaline soils of western San Joaquin Valley, California, selenate is the dominant soluble species constituting 98% of the soluble Se, but selenites and organic Se species are also present. The solubility and mobility of Se in soils depend on processes such as adsorption, precipitation, and transformation, which in turn depend on the chemical species of Se, pH, and redox conditions. Several studies have shown that selenite is sorbed onto soil to a much greater extent than selenate (13), which makes it less bioavailable than selenate. Selenate generally adsorbs weakly onto clays, iron oxides, and other soil phases at neutral to alkaline pH. The low solubility of Se in acid or neutral soils can be attributed to its occurrence as selenite in combination with ferric iron as ferric selenite. In soils containing reactive ferric oxides, ferric selenite–ferric hydroxide adsorption complexes are formed. Oxidation of selenite to selenate in alkaline soils is favored by a decrease in the stability of ferric hydroxide–selenite complexes (12). Oxyanions such as NO_3^- can facilitate the oxidation of Se, resulting in high levels of soluble selenate in drainage effluents from agricultural lands.

The three major mechanisms of Se transformation in soil are: oxidation–reduction, methylation–demethylation, and volatilization. Oxidation leads to mobilization, and reduction results in immobilization of Se. Methylation and volatilization lead to the loss of Se from the soil. A number of microorganisms are capable of forming volatile, alkylated Se compounds from inorganic forms of Se, which are less toxic. Dimethylselenide (DMSe) is the major metabolite of Se volatilization, although other products such as dimethyldiselenide (DMDSe), methaneselenone, methaneselenol, and dimethylselenide sulfide may also be produced (14).

SELENIUM IN PLANTS

Plant Uptake and Transport

Selenium is taken up from soils by plants primarily as selenate or selenite. Selenate is accumulated in plant cells against its likely electrochemical potential gradient through a process of active transport (3). Selenate competes directly with sulfate for uptake, indicating that both anions are taken up by the sulfate transporter protein in the root plasma membrane. Kinetic and expression studies have indicated that sulfate transporters belong to two main classes, transporters that have either high or low affinity for sulfate. The high-affinity transporter is the primary transporter involved in sulfate uptake and is expressed mainly in the roots. The low-affinity transporter is expressed both in roots and shoots and may be involved in intercellular transport of sulfate (3). Unlike selenate, there is no evidence that the uptake of selenite is mediated by membrane transporters. Selenate is taken up by the sulfate transporter, so sulfate salinity drastically inhibits plant uptake of selenate (15).

Plant Assimilation and Volatilization

After uptake, it is believed that selenate is transported to chloroplasts, where it is processed by the S assimilation pathway (1). Selenate is activated by ATP sulfurylase enzyme, forming adenosine 5′-phosphoselenate (APSe), which is a major rate-limiting step for selenate assimilation in plants. APSe is then reduced to selenite by adenosine 5′-phosphosulfate (APS) reductase, which in turn is nonenzymatically reduced to selenide by glutathione. Selenocysteine is synthesized from O-acetyl serine and selenide by cysteine synthase enzyme (1). Selenoamino acids, such as selenocysteine and selenomethionine, are nonspecifically incorporated into proteins, thereby leading to Se toxicity (3). Selenomethionine can be converted to DMSe and then volatilized. Recently, Se volatilization by plants has received attention as a possible method of phytoremediation of Se-contaminated soils.

Selenium Hyperaccumulators and Nonaccumulators

Plants differ in their ability to accumulate Se in their tissues. Certain plants can hyperaccumulate Se in their shoots when grown in seleniferous soils. Such plants, called Se hyperaccumulators, include a number of species such as *Astragalus, Stanleya, Morinda, Neptunia, Oonopsis*, and *Xylorhiza* (3). They can accumulate several hundreds to several thousands of milligrams of Se/kg dry weight in their tissues, whereas nonaccumulators contain less than 25 mg Se/kg dry weight. Secondary Se accumulators include plants such as Indian mustard and canola, which can accumulate several hundred milligrams of Se/kg tissue when grown in soils contaminated with moderate levels of Se.

The pathway for the assimilation of inorganic forms of Se to selenocysteine in Se accumulators, it is believed, is the same as for nonaccumulators (3). The main difference between Se accumulators and nonaccumulators is that accumulators metabolize selenocysteine into various nonprotein selenoamino acids. The best characterized Se accumulator is *Astragalus bisulcatus*. In this plant, selenocysteine is methylated to methylselenocysteine, which prevents its nonspecific incorporation into proteins and helps to confer Se tolerance (16).

PHYTOREMEDIATION

Phytoremediation is the use of plants to remove, contain, or render harmless certain environmental pollutants (4).

It offers a cheaper, safer, and environmentally friendly solution for cleaning up Se-contaminated areas. Plants can be used in conjunction with microbial activity to extract, accumulate, and volatilize Se (8). There are several different types of phytoremediation, which can be classified as

- Phytoextraction, which uses the ability of certain plants to take up contaminants from soil and water and accumulate them in their tissues. Such tissues can then be harvested and removed from the site. Phytoextraction is facilitated by hyperaccumulator plants that grow naturally on heavy metal and metalloid-rich soils (17). Several plant species have been evaluated for their efficiency in phytoremediation, particularly certain species of *Astragalus* (18). However, these are slow growing plants, and the Se accumulated can be leached back into the soil from the leaves by rainfall (3). Another disadvantage of phytoextraction is that the accumulated Se in plant tissues could become available to animals and birds.
- Phytostabilization, which uses plants to immobilize contaminants chemically and physically at the site, thereby preventing their movement to surrounding areas (19).
- Phytofiltration, which takes advantage of the ability of plant roots to absorb or adsorb environmental contaminants, mostly metals, from water and waste streams.
- Phytodetoxification, which involves the ability of plants to change the chemical species of the contaminant to a less toxic form. For example, plants can take up toxic hexavalent chromium and convert it to nontoxic trivalent chromium (15).
- Phytovolatilization, which uses plants and their associated microbes to metabolize certain contaminants to volatile forms, which are removed from the environment. For Se, phytovolatilization is an innovative and potentially effective remediation technology for several reasons.
    1. Volatilization minimizes the production of Se-laden plant material, thereby reducing the risk that contaminants move into the local food chain.
    2. DMSe, the major component of volatile Se resulting from phytovolatilization, is 600 times less toxic than selenate or selenite (20).

The ideal plant species for phytoremediation is one that can accumulate and volatilize large amounts of Se, grow rapidly, produce a large biomass in Se-contaminated soil, tolerate salinity and other toxic environmental conditions, and provide a safe source of forage for Se-deficient livestock (3). Indian mustard (*Brassica juncea*) has most of these preferred characteristics (21). These plants can also be genetically engineered to improve their Se-remediation capacity. Pilon-Smits et al. (22) were able to overexpress the gene encoding ATP sulfurylase in Indian mustard, which resulted in transgenic plants that accumulated two- to three-fold more Se per plant compared to wild-type plants.

GENETIC ENGINEERING TO ADVANCE SELENIUM PHYTOREMEDIATION

Phytoremediation by genetically unaltered plant species is limited by both environmental and biological factors. Hyperaccumulators can accumulate high concentrations of Se in their tissues, but phytoremediation using hyperaccumulators is often very slow due to their slow growth rate and low biomass (23). These limitations can be overcome by conventional plant breeding techniques, but developing accumulators that have suitable characteristics may take decades. Genetic engineering has the potential to produce accumulators with suitable traits for phytoremediation within a short time (4). Genetic engineering could be used to create plants that are more tolerant to contaminants, have greater efficiency in extracting pollutants, or have greater capacity for accumulating and detoxifying pollutants.

Manipulation of the Sulfate Assimilation Pathway

The most popular genetic engineering strategy is to overexpress appropriate genes in suitable plants by inserting extra copies of coding DNA under the control of constitutive or inducible promoter elements. Overexpression of genes involved in rate-limiting steps of Se assimilation and volatilization pathways should enhance such processes in transgenic plants (4), which requires thorough understanding of the mechanisms of Se uptake and metabolism in plants. Plants metabolize Se through the S assimilation pathway, so overexpressing various genes in this pathway is likely to result in increased assimilation and volatilization of Se in transgenic plants. For example, overexpression of sulfate transporter genes could greatly increase plant uptake of Se. Overexpression of ATP sulfurylase gene (coding for ATP sulfurylase protein constituting the rate-limiting step in the S and Se assimilation pathway) or overexpressing the methionine S-methyltransferase gene (coding for the first enzyme in the biochemical pathway resulting in the formation of DMSe, the major volatile form of Se produced by plants) could result in enhanced phytoremediation capacity in transgenic plants. Transformation of Indian mustard with a gene encoding ATP sulfurylase from *Arabidopsis* under the control of a constitutive promoter resulted in higher levels of ATP sulfurylase activity and increased Se accumulation in these plants (22).

Introduction of Genes from Selenium Hyperaccumulators

A second strategy is to introduce additional metabolic pathways into the plant by overexpressing unique genes from Se hyperaccumulator plants. Hyperaccumulators have a unique ability to reduce the toxic effects of Se by depleting the intracellular concentrations of selenocysteine and selenomethionine, which disrupt plant metabolism when incorporated into enzyme proteins. In hyperaccumulators, the levels of selenocysteine and selenomethionine are reduced by the presence of metabolic pathways that lead to the accumulation of Se in nonprotein amino acids (4). Neuhierl and Boeck (24) purified a methyl transferase enzyme from *A. bisulcatus*, which specifically methylated selenocysteine. When the gene encoding the *A.*

bisulcatus methyl transferase was transferred to *E. coli*, the bacteria had increased Se tolerance and reduced incorporation of Se into protein (25).

Introduction of Microbial Genes

Microbes are also likely to provide a new library of genes capable of conferring Se tolerance and phytoremediation capacity on plants. Overexpression of microbial genes has been successfully used in the past to increase heavy metal tolerance in plants (4). Identification and use of these genes to transform plants can enhance the capacity of plants for uptake, detoxification, and volatilization of Se.

Overexpressing Transcription Factors

Overexpression of ATP sulfurylase that catalyzes the rate-limiting step in the Se assimilation pathway results in increased accumulation of Se (22). Many other genes in the S/Se assimilation pathway could be similarly overexpressed, or a similar effect could be achieved by overexpressing a single transcription factor responsible for activating a number of coregulated genes (4). In the Se assimilation pathway, the expression of three genes, ATP sulfurylase, sulfate permease, and APS reductase are regulated similarly (4). It is possible that a common transcription factor can be used to modulate the expression of these genes, resulting in increased Se assimilation in plants.

SUMMARY AND FUTURE DIRECTIONS

The consequences of elevated Se accumulation at the Kesterson Reservoir National Wildlife Refuge in California fueled extensive research on the behavior of Se in soils and plants. Conventional remediation technologies for cleaning up Se-contaminated soils are often prohibitively expensive. On the other hand, Se accumulation and volatilization by plants have been effective in remediating Se-contaminated soil and water (26,27). Selenium volatilization is a particularly attractive option because toxic inorganic Se is converted to a volatile gas DMSe, which is about 600 times less toxic than inorganic Se. Although a number of Se hyperaccumulators are known, their slow growth rate and low biomass make them less than satisfactory for phytoremediation. To enhance the Se remediation potential of plants, understanding the speciation and bioavailability of Se in soils and the biochemistry of Se assimilation and volatilization in plants and microorganisms is of paramount importance.

Recent molecular studies have revealed numerous details of Se physiology and biochemistry in plants. Overexpression of genes encoding transporters and enzymes involved in the uptake, assimilation, and volatilization of sulfate has been highly successful in elucidating the role of these proteins in Se metabolism. Identification of the rate-limiting steps in the Se assimilation pathway will aid in enhancing the phytoremediation capacity of plants by genetic engineering. The role of rhizosphere bacteria in Se volatilization by plants also needs to be further elucidated, because it is well known that plants volatilize relatively low amounts of Se in the absence of bacteria. Primary Se accumulators are a potential source of genetic material that could be used to alter Se metabolism in economically important plants, but secondary accumulators, such as Indian mustard, which are fast growing and have high biomass can also be genetically engineered to improve their phytoremediation capacity.

Disposal of Se-containing vegetation is another problem that needs further study. Se is an essential trace element in humans and animals, so one solution is to use seleniferous plant materials as a forage blend in Se-deficient regions. Plants used for phytoremediation of Se may also be used to generate useful by-products such as fibers for the production of paper and building materials and energy for heat production.

BIBLIOGRAPHY

1. Ellis, D.R. and Salt, D.E. (2003). Plants, selenium and human health. *Curr. Opin. Plant Biol.* **6**: 273–279.
2. Navarro-Alarcon, M. and Lopez-Martinez, M.C. (2000). Essentiality of selenium in the human body: Relationship with different diseases. *Sci. Total Environ.* **249**: 347–371.
3. Terry, N., Zayed, A.M., de Souza, M.P., and Tarun, A.S. (2000). Selenium in plants. *Annu. Rev. Plant Physiol. Plant. Mol. Biol.* **51**: 401–432.
4. Berken, A., Mulholland, M.M., LeDuc, D.L., and Terry, N. (2002). Genetic engineering of plants to enhance selenium phytoremediation. *Crit. Rev. Plant Sci.* **21**: 567–582.
5. Frankenberger, W.T., Jr. and Engberg, R.A. (Eds.). (1998). *Environmental Chemistry of Selenium*. Marcel Dekker, New York.
6. Ohlendorf, H.M., Hoffman, D.J., Salki, M.K., and Aldrich, T.W. (1986). Embryonic mortality and abnormalities of aquatic birds: Apparent impacts of selenium from irrigation drain water. *Sci. Total Environ.* **52**: 49–63.
7. Saiki, M.K. and Lowe, T.P. (1987). Selenium in aquatic organisms from subsurface agricultural drainage water, San Joaquin Valley, California. *Arch. Environ. Contam. Toxicol.* **19**: 496–499.
8. Banuelos, G.S. (2000). Factors influencing field phytoremediation of selenium-laden soils. In: *Phytoremediation of Contaminated Soil and Water*. N. Terry and G.S. Banuelos (Eds.). Lewis, Boca Raton, FL, pp. 41–59.
9. Ohlendorf, H.M., Hothem, R.L., Aldrich, T.W., and Krynitsky, A.J. (1987). Selenium contamination of the grasslands. A major California, USA, waterfowl area. *Sci. Total Environ.* **66**: 169–184.
10. Kerr, J.A. (1998). Selenium. In: *The Electronic Encyclopedia* (TM). Academic American Encyclopedia on CD-ROM, Grolier Electronic, Danbury, CT.
11. Agency for Toxic Substances and Disease Registry. (1996). *Toxicological Profile of Selenium*. U.S. Department of Health and Human Services, Public Health Service, Atlanta, GA.
12. Adriano, D.C. (1986). *Trace Elements in Terrestrial Environments*, 2nd Edn. Springer-Verlag, New York.
13. Gissel-Nielsen, G., Gupta, U.C., Lamand, M., and Westermarck, T. (1984). Selenium in soils and plants and its importance in livestock and human nutrition. *Adv. Agron.* **37**: 397–460.
14. Thompson-Eagle, E.T. and Frankenberger, W.T. (1992). Bioremediation of soils contaminated with selenium. In: *Advanced Soil Science*. R. Lal and B.A. Stewart (Eds.).

Soil Restoration Series, 17, Springer-Verlag, New York, pp. 261–310.

15. Zayed, A.M., Lytle, C.M., and Terry, N. (1998). Accumulation and volatilization of different species of selenium by plants. *Planta* **206**: 284–292.
16. Wang, Y., Boeck, A., and Nuehierl, B. (1999). Acquisition of selenium tolerance by a selenium non-accumulating *Astragalus* species via selection. *Biofactors* **9**: 3–10.
17. Kumar, N.P.B.A., Dushenkov, V., Motto, H., and Raskin, I. (1995). Phytoextraction: The use of heavy metals to remove heavy metals from soils. *Environ. Sci. Technol.* **29**: 1232–1238.
18. Parker, D.R., Page, A.L., and Thomason, D.N. (1991). Salinity and boron tolerances of candidate plants for the removal of selenium from soils. *J. Environ. Qual.* **20**: 157–164.
19. Berti, W.R. and Cunningham, S.D. (2000). Phytostabilization of metals. In: *Phytoremediation of Toxic Metals: Using Plants to Clean Up the Environment*. I. Raskin and B.D. Ensley (Eds.). Wiley-Interscience, New York, pp. 71–88.
20. Wilbur, C.G. (1980). Toxicology of selenium: A review. *Clin. Toxicol.* **17**: 171–230.
21. Zayed, A.M., Pilon-Smits, E.A.H., deSouza, M.P., Lin, Z.Q., and Terry, N. (1999). Remediation of selenium-polluted soils and waters by phytovolatilization. In: *Phytoremediation of Metal-Contaminated Water and Soils*. N. Terry and G. Banuelos (Eds.). CRC Press, Boca Raton, FL, pp. 61–83.
22. Pilon-Smits, E.A.H. et al. (1999). Overexpression of ATP sulfurylase in Indian mustard leads to increased selenate uptake, reduction and tolerance. *Plant Physiol.* **119**: 123–132.
23. Cunningham, S.D., Shann, J.R., Crowley, D.E., and Anderson, T.A. (1997). Phytoremediation of contaminated water and soil. In: *Phytoremediation of Soil and Water Contaminants*. E.L. Kruger, T.A. Anderson, and J.R. Coats (Eds.). *ACS Symposium Series* No. 164, American Chemical Society, Washington, DC.
24. Neuhierl, B. and Boeck, A. (1996). On the mechanism of selenium tolerance in selenium accumulating plants: Purification and characterization of a specific selenocysteine methyltransferase from cultured cells of *Astragalus bisulcatus*. *Eur. J. Biochem.* **239**: 235–238.
25. Neuhierl, B., Thanbichler, M., Lottspeich, F., and Boeck, A. (1999). A family of S-methylmethionine-dependent thiol-selenol mthyltransferases: Role in selenium tolerance and evolutionary relation. *J. Biol. Chem.* **274**: 5407–5414.
26. Banuelos, G.S., Terry, N., Zayed, A., and Wu, L. (1995). Managing High Soil Selenium with Phytoremediation. *Proc. 12th, Annu. Natl. Meet. Am. Soc. Surf. Min. Reclam.*, Gillette, WY.
27. Terry, N. and Zayed, A.M. (1998). Phytoremediation of selenium. In: *Environmental Chemistry of Selenium*. W.T. Frankenberger, Jr. and R.A. Engberg (Eds.). Marcel Dekker, New York, pp. 633–656.

SOIL PIPES AND PIPE FLOW

Joseph Holden
University of Leeds
Leeds, United Kingdom

Natural soil pipes are subsurface cavities of diameter greater than 1 mm that are continuous in length such that they can transmit water, sediment, and solutes through the soil and bypass the soil matrix. Soil pipes are larger versions of soil macropores. Pipes can be up to several meters in diameter, several hundred meters long, and occur in a broad range of environments. They are most common in soils that are subject to cracking in dry periods or are highly dispersible. Soils that have a significant amounts of swelling clays also commonly contain soil pipes (e.g., montmorillonite, illite, and bentonite). Piping is common in arid and semiarid areas where shrinking and desiccation cracking are common. Piping is also common in highly organic soils in humid uplands (e.g., blanket peats). However, soil pipes are found in virtually all climates, in organic and mineral soils, and on disturbed and undisturbed land.

PIPE INITIATION

The wide range of environments in which soil pipes are found reflects the wide range of pipe initiation and development processes (Fig. 1) (1). The initiation of piping is discussed in detail by Jones (2), although there have been relatively few process measurements (3,4). The following factors are considered important in the initiation and development of piping: climate, biota, human activity, pedogeomorphic context, role of soil chemistry (influenced by sodium salts and influenced by systems exhibiting base deficiency), soil texture, erodibility, soil structure and hydraulic conductivity, clay minerals, cracking potential, and dispersivity. The dual need for water supply and desiccation means that pipes are most frequent in semiarid and temperate marine environments (1). Flow through desiccation and biotic and mass movement cracks may enhance the macropores into pipe networks. Jones (5) demonstrates that the vast majority of catchments examined that have piping in Britain, for example, face south, suggesting that desiccation cracking may be very important in the formation of piping there. Gilman and Newson (6) observed that vertical cracks in both the roof and floor of peat pipes were a common feature in mid-Wales during the drought summer of 1976 and suggested that this allowed more water to reach deeper levels and created permanent extensions of the pipe networks when rewetting took place. Faunal activity in many areas, it has been shown, results in active soil pipe formation and affects subsurface movement of water (4,7).

PIPE IDENTIFICATION

Until recently, it has been difficult to find and define soil pipe networks because pipes are often visible only at stream banks or where the pipe roof has collapsed and created a surface opening (Fig. 1b) or formed a gully. Therefore, it is difficult to map the route of a soil pipe or to know its diameter, depth, and length across a hillslope. In some soils, a change in surface vegetation may often indicate the presence of a pipe (8). Jones and Crane (9) extensively mapped 4.4 km of pipes in a drainage area of only 0.23 km^2 by a dye tracing and ground survey. Pipe locations were identified mainly by observation of collapse features, of water jets emerging from pipes, and the sound of flowing water (2). However, these techniques do not give

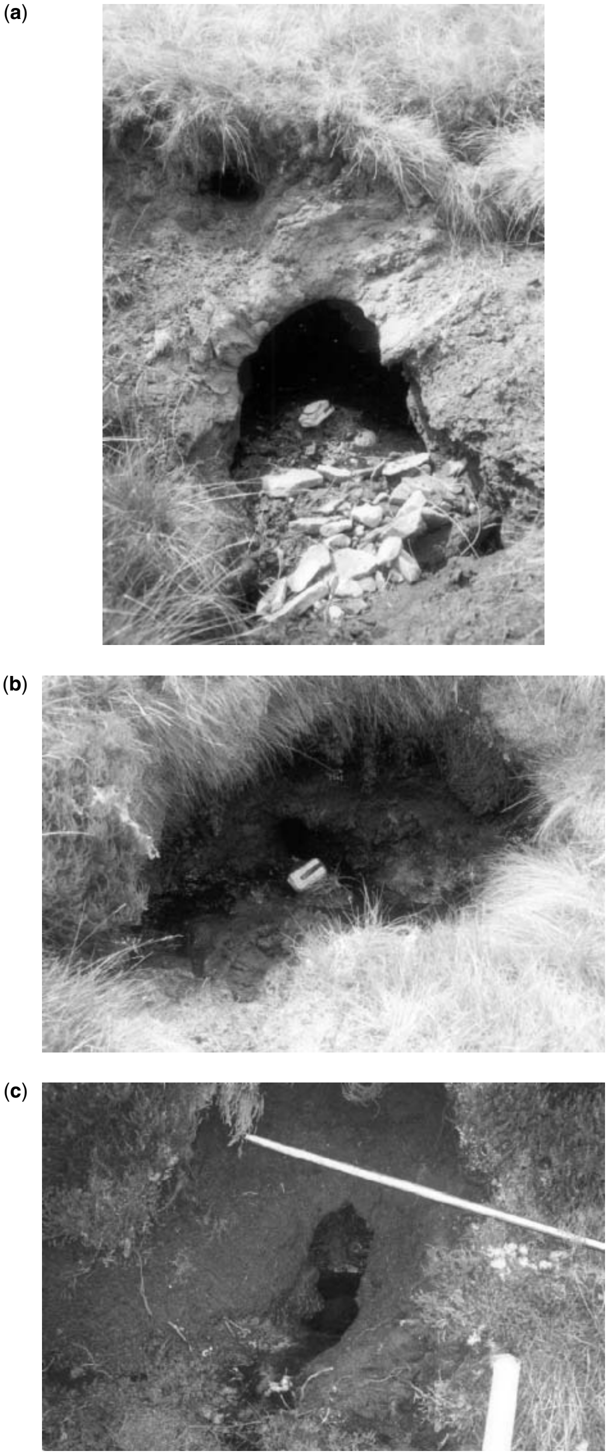

Figure 1. Natural soil pipes. (**a**) large pipe outlet with coarse sediment delivered from its base; (**b**) collapsed surface undermined by pipe erosion; (**c**) typical pipe outlet in blanket peat. Holden, J. and Burt, T.P. (2002) Piping and pipeflow in a deep peat catchment. *Catena* **48**: 163–199.

a detailed or complete picture of the subsurface network and can result in underestimates of pipe density. Often, destructive techniques (e.g., soil trenches) must be used to investigate the pipes (2), and there have been few other attempts to locate and map subsurface piping accurately.

Holden et al. (10) demonstrated that ground penetrating radar (GPR) could be used to sense soil pipes remotely (Fig. 2). The radar transmits short pulses of high-frequency (10–10,000 MHz) electromagnetic energy through the ground surface, which is reflected from boundaries between soil layers or from internal irregularities whose electrical properties are different. Thus, soil pipe cavities can be identified by radar when it crosses over them. GPR work has shown that pipe network densities are much greater than can be detected from surface observation alone. On its own, the radar can detect only subsurface cavities, but when used in conjunction with tracers (e.g., sodium chloride), it can establish the hydrologic connectivity of soil pipes (11), which will be detected by alterations in the reflection of the electromagnetic signal if the electrical conductance of the water in the soil pipe suddenly changes below the radar. Thus, it is possible to determine how pipes are connected to others within a complex subsurface drainage network.

PIPE MORPHOLOGY

The size (mean diameter) of pipes is the most recorded measure of pipe geometry often measured only at outlet locations. However, pipe size can vary significantly just a few meters upslope from the outlet. All areas have a range of pipe sizes, often within the same pipe network. The largest pipes, however, tend to develop in arid and semiarid areas (12). Some evidence suggests that small rounded pipes tend to evolve into larger flat-bedded or rectangularly shaped pipes (2). Fiberscopes have recently been developed to examine internal structures of soil pipes, although these have usually been restricted to short distances from a pipe outlet (13). The average cross-sectional area of pipe outlets per kilometer of stream bank is often taken as a good measure of piping intensity along a stream bank (14), although many reports are of pipes not directly connected to the stream. Pipes can be at any depth within a soil profile, although they are more common at the interface between two distinct soil horizons (especially at organic—mineral interfaces), where there is a sharp transition in hydraulic conductivity or erodibility. In blanket peat, however, pipes have been found that undulate

Figure 2. Ground penetrating radar being used in the field to detect soil pipes. Photo by Vincent Holden.

throughout the soil profile. Pipe floors can run counter to the surface topography such that hydraulic pressures are required to transport the water upward through those sections. An example of a typical vertical section through a soil pipe network is shown in Fig 3. Pipes thus provide rapid connectivity of water, sediment, and solutes throughout the soil profile. Soil pipes are rarely isolated conduits; they often form complex networks that have meandering and anastomosing forms (2,10). Thus, they can extend the stream channel network many hundreds of meters upslope and rapidly transmit water and sediment from great distances away from the stream channel itself.

PIPE HYDROLOGY

Laboratory work (15), modeling (16), and field measurement (17,18) have shown that piping can be a very important hydrologic phenomenon, particularly in humid temperate regions (2,6,14,19). Jones and Crane (9) reported, for example, that 49% of stream flow in the Maesnant catchment, mid-Wales, United Kingdom, was generated through the pipe network. Holden and Burt (10) found that 10% of the discharge in peat catchments moved through the pipe network but at times (depending on antecedent conditions), this could be as high as 30%. Three broad categories of pipe flow can be identified: perennial, ephemeral, and seasonal. There is typically a great variety of hydrologic response between one pipe and another, even within a small area, although some pipe outlets can simply be connections to the same pipe network as others. Jones (20) reports that piping doubles the dynamic source area for rapid runoff response in the upper Maesnant, mainly by linking points well beyond the riparian zones of seepage to the stream.

It is likely that there are several sources of pipe-flow water: overland flow entering collapse features where pipes are open to the surface; water flowing through cracks and root channels to enter the pipe network; diffuse seepage from pipe walls into the pipe system, and groundwater springs forcing water upward from below (Fig. 3). Pipes often receive water more quickly than would be expected from diffuse seepage through soil layers (8,10), and it certainly does not seem that pipe flow in all catchments is restricted by the long time it takes for rainwater to infiltrate down as far as the level of pipes. Nevertheless, Sklash et al. (21), using isotope analysis, showed that pipes in some catchments may transmit predominantly 'old' water (water from previous storms) that is richer in deuterium during storm flow. In an adjacent catchment, however, there was strong chemical evidence that storm flow in the pipes was dominated by fresh storm rainfall. In semiarid zones, pipe flow can occur while the surrounding soil or bedrock remains relatively dry (12).

PIPE EROSION AND WATER QUALITY

Piping, a form of subsurface erosion, has been associated with gully and rill development in many parts of the world. Particular associations were noted in Australia and New Zealand during 1940s soil erosion surveys. Alluvial and loess soils are particularly prone to such pipe–gully coupling. Some pipes are subject to 'episodic cut and fill' in that they can collapse and block the route for water before eventually unblocking and allowing water to flow through them again (17). Thus, pipe discharge can be highly erratic. Pipes can provide the stream system with sediment that is very different from that expected from overland flow because they often connect and erode deeper soil layers (Fig. 3). If piping is very active, then sediment losses from hillslopes can be severe and can cause problems for soil conservation workers. However, very few measurements of pipe-flow sediments have been made around the world. Hauser and Zotl (22) measured 0.5 kg of sand per week from one pipe that had an average discharge of $0.16\,\mathrm{L\,s^{-1}}$ in the Sarmatian mountains of Graz, Austria, and badland pipe sediment loads, it is thought, are an order of magnitude higher than overland flow concentrations (2). Because it is subsurface, pipe erosion can occur, even when the hillslope is densely vegetated (23).

The role of piping in major mass movement is as yet uncertain. There have been some instances where pipes have been identified as the cause of subsidence or landslides, others where pipes are thought to have developed after the mass movement, and still further cases

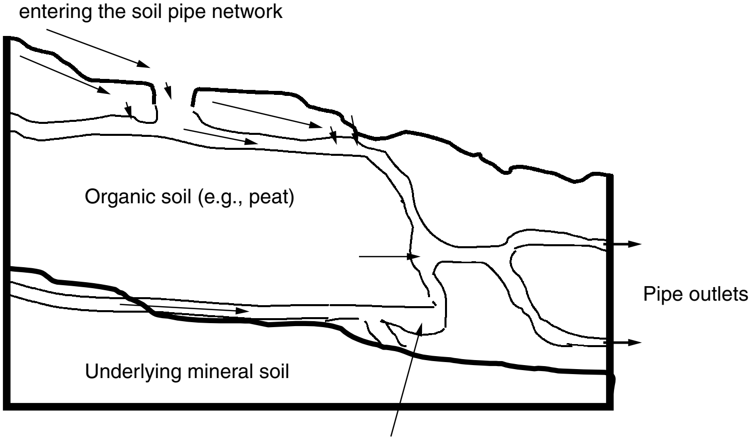

Figure 3. Schematic of a short vertical section through a soil pipe network.

where pipes it is thought, have prevented mass movement. In many circumstances, water flow is too great to be freely transmitted along these subsurface tunnels, and a pressure head builds back up the pipe. If a heavy rainstorm occurs, it is possible that restrictions in the dimensions of the pipes that act as bottlenecks will impede flow. Hence, water builds back up the pipe, and large pressures may build up within the slope. At the same time, pipes might provide a drainage path for water, prevent the buildup of gravitational forces on the slope, and thus prevent mass movements in some areas.

In some catchments, measurements of pipe flow chemistry have failed to determine any significant difference from that of overland flow, suggesting that the sources of the water were the same (24). Piping may therefore exacerbate problems of acid runoff by reducing soil buffering effects as a result of more rapid transmission, reduced soil residence times, and reduced contact between weathering mineral surfaces. Drainage and aeration of large sections of hillslope can also encourage release of sulfates and organic acids, particularly from peaty soils, which increases the acidity of the runoff. Therefore, many pipes tend to issue more acidic water during storms and contribute to acid flushes in streams that can increase acid rain impacts on fish and microfauna.

BIBLIOGRAPHY

1. Bryan, R.B. and Jones, J.A.A. (1997). The significance of soil piping processes: inventory and prospect. *Geomorphology* **20**: 209–218.
2. Jones, J.A.A. (1981). The nature of soil piping: a review of research. *British Geomorphological Research Group Monograph Series 3*, GeoBooks, Norwich, MA.
3. Torri, D. and Bryan, R. (1997). Micropiping processes and biancana evolution in Southeast Tuscany, Italy. *Geomorphology* **20**: 219–235.
4. Onda, Y. and Itakura, N. (1997). An experimental study on the burrowing activity of river crabs on subsurface water movement and piping erosion. *Geomorphology* **20**: 279–288.
5. Jones, J.A.A. (1994). Soil piping and its hydrogeomorphic function. *Cuaternario y Geomorfologia* **8**: 77–102.
6. Gilman, K. and Newson, M.D. (1980). *Soil Pipes and Pipeflow; A Hydrological Study in Upland Wales*, *British Geomorphological Research Group Research Monograph No. 1*, GeoBooks, Norwich, MA.
7. Banerjee, A.K. (1972). Morphology and genesis of pipe structure in ferralitic soils of Midnapore District, West Bengal. *Journal of the Indian Society of Soil Science* **20**: 399–402.
8. Jones, J.A.A., Wathern, P., Connelly, L.J., and Richardson, J.M. (1991). Modelling Flow in Natural Soil Pipes and its Impact on Plant Ecology in Mountain Wetlands. In: *Hydrological Basis of Ecologically Sound Management of Soil and Groundwater (Proceedings of the Vienna Symposium, August 1991)*. IAHS publication 202, pp. 131–142.
9. Jones, J.A.A. and Crane, F.G. (1984). Pipeflow and pipe erosion in the Maesnant experimental catchment. In: *Catchment Experiments in Fluvial Geomorphology*. T.P Burt and D.E. Walling (Eds.). GeoBooks, Norwich, MA, pp. 55–72.
10. Holden, J. and Burt, T.P. (2002). Piping and pipeflow in a deep peat catchment. *Catena* **48**: 163–199.
11. Holden, J. (2004). Hydrological connectivity of soil pipes determined by GPR tracer detection. *Earth Surface Processes and Landforms to be published*.
12. Bryan, R.B. and Yair, A. (Eds.). (1982). *Badland Geomorphology and Piping*. GeoBooks, Norwich, MA.
13. Terajima, T., Sakamoto, T., and Shirai, T. (2000). Morphology, structure and flow phases in soil pipes developing in forested hillslopes underlain by a Quaternary sand-gravel formation, Hokkaido, northern main island in Japan. *Hydrological Processes* **14**: 713–726.
14. Jones, J.A.A., Richardson, J.M., and Jacob, H.J. (1997). Factors controlling the distribution of piping in Britain: a reconnaissance. *Geomorphology* **20**: 289–306.
15. Sidle, R.C., Kitahara, H., Terajima, T., and Nakai, Y. (1995). Experimental studies on the effects of pipeflow on throughflow partitioning. *Journal of Hydrology* **165**: 207–219.
16. Nieber, J.L. and Warner, G.S. (1991). Soil pipe contribution to steady subsurface stormflow. *Hydrological Processes* **5**: 329–344.
17. Zhu, T.X. (1997). Deep-seated, complex tunnel systems—a hydrological study in a semiarid catchment, Loess Plateau, China. *Geomorphology* **20**: 255–267.
18. Carey, S.K. and Woo, M. (2001). Spatial variability of hillslope water balance, Wolf Creek basin, subarctic Yukon. *Hydrological Processes* **15**: 3113–3132.
19. Uchida, T., Kosugi, K., and Mizuyama, T. (1999). Runoff characteristics of pipeflow and effects of pipeflow on rainfall-runoff phenomena in a mountainous watershed. *Journal of Hydrology* **222**: 18–36.
20. Jones, J.A.A. (1987). The effects of soil piping on contributing area and erosion pattern. *Earth Surface Processes and Landforms* **12**: 229–248.
21. Sklash, M.G., Beven, K.J., Gilman, K., and Darling, W.G. (1996). Isotope studies of pipeflow at Plynlimon, Wales, UK. *Hydrological Processes* **10**: 921–944.
22. Hauser, A. and Zotl, J. (1955). Die morphologische Bedeuntung der unterirdischen Erosion durch Gesteinausspulung. *Petermans Mitt.* **99**: 18–21.
23. Jones, J.A.A. (1990). Piping effects in humid lands. In: *Groundwater Geomorphology; the Role of Subsurface Water in Earth-Surface Processes and Landforms*. C.G. Higgins and D.R. Coates (Eds.). Geology Society of America Special Paper 252, pp. 111–137.
24. Cryer, R. (1980). The chemical quality of some pipeflow waters in upland mid-Wales and its implications. *Cambria* **6**: 28–46.

LOW FLOW GROUNDWATER PURGING AND SURGING

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Groundwater monitor wells are installed primarily as tools to measure the dissolved concentrations of contaminants in groundwater in and around facilities where the release of contaminants has occurred or where the potential exists for such a release to occur (i.e., adjacent to landfills). These wells are installed with the intention of obtaining an accurate assessment of the conditions in the formation adjacent to the well. A significant body of literature and guidance documentation exists concerning the location, drilling, installation, and development of those wells to that end. However, it has become evident that the quality of the data yielded from monitor wells is dependent on screen length and location (addressed during design and installation), purging method, and purge volume. The question becomes how does one collect samples from groundwater monitoring wells that generate representative results?

The regulatory community and those involved in the design and operation of remediation systems are seeking an accurate answer to that question. However, interest is also generated from the economic pressures to reduce the cost of sampling and monitoring procedures. Standard groundwater sampling procedures involve purging three to five well volumes with a pump or a bailer, which entails expense in the form of time (often for two technicians) and disposal of the purge water. These two issues have driven recent research by the scientific community, as well as the regulatory community, which is also beginning to act by introducing new protocols for the acceptable sampling of wells.

The use of screened hydraulic probes is recognized as an accurate sampling method and as the yardstick for comparison with various monitor well sampling techniques. Their practical use is of value for initial plume delineation and assessment in aid of the design of a monitor well network; however, they are impractical as a long-term monitoring system. The groundwater monitor well is essential for that role.

For conventional purge techniques, wherein three to four well casing volumes of water were removed during purging, it was determined that water in a monitor well was “static” and in time grew “stagnant.” The purging of the well casing volumes was to remove the old and “stagnant” water, which allowed for representative aquifer water to enter the well bore for sampling.

Concerns were twofold: (1) the assumption that the mechanisms at work within the well bore that would attenuate the contaminant concentrations such as vaporization and/or biodegradation supported by oxygen diffusion; and (2) the assumption that with time, formation conditions will improve, contaminant concentrations will decrease, and samples will demonstrate that trend.

Recent studies have shown that conventional purging procedures can activate two phenomena that to give false results: vertical mass averaging within the well bore and the introduction of water lower in the formation through vertical (upward) flow paths. Vertical heterogeneity of contaminant concentrations within the well bore occurs when zones of striking hydraulic conductivity contrast are present. Contaminants preferentially flow through the more permeable units intersected by the well screen. Gross purging then simply mixes the entire well bore. Low flow discrete level sampling has been devised to address this specific phenomena. The induction of vertical uncontaminated groundwater flow from beneath a partially penetrating well is a problem in formations at the opposite end of the hydrodynamic spectrum, which are coarsely granular (sand and gravel) and relatively homogenous.

Field testing shows that low flow purging at specific vertical points within the well bore yields results that are similar to sampling via the conventional purging methods. Bailing is likely to produce the lowest concentrations, and low flow pumping slightly higher, but both tend to produce results that are significantly lower than actual concentrations within the adjacent formation. In general, it appears that the more vigorously a well is purged, the more likely it is that the samples will underrepresent contaminant concentrations within the adjacent formation. A detailed description of low flow (minimal drawdown) groundwater sampling is described by the U.S. Environmental Protection Agency.

Physical and chemical processes can potentially impact a monitor well in a static or stagnant state. The processes that could effect the concentration of dissolved components or the degree of mixing of heterogenetic water within a well bore column include:

- The groundwater flow rate through the surrounding formation and its influence via advective flow within the well
- The vertical profile of the contaminant concentration and hydraulic conductivity contrast within the adjacent formation
- Vaporization and volatilization across the air water interface at the top of the well
- Diffusion of oxygen through the air–water interface into the well bore stimulating biodegradation Diurnal and seasonal temperature fluctuations
- Diffusional transport vertically within the well bore as well as horizontally from the adjacent formation
- Barometric pumping, which under some conditions can cause water level fluctuations (and mechanical mixing) in wells located in unconfined aquifers
- Seismic activity
- Tidal forces
- Local heavy auto, truck, or train traffic

In aggregate, these forces impact all monitor wells to various degrees depending on the geologic and geographic site conditions. In some instances, “static” conditions are dynamic enough to ensure that water in a monitor well is representative of the adjacent formation.

Consequently, some regulatory bodies have begun to consider the use of non-purge or low-purge (the EPA) methods of sampling. In California, the Cal/EPA allows non-purge sampling in cases where:

- The monitor well is screened in an unconfined (water table) aquifer
- The screened interval extends above the top of the water table

- The monitor wells are located in moderate-to-high recovery aquifers consisting of sands or gravels

Purging is still required when:

- The well is new and has been sampled less than two times (this also assumes the well underwent an adequate development purge at installation)
- Chlorinated hydrocarbons, PNAs, diesel, or metals are the contaminant
- Slow recovering wells within fine-grained sediments (clays and silts) are used
- Monitor wells associated with sparging or soil vapor extraction activity are used

LOW FLOW TECHNIQUES

Low flow refers to the velocity with which water enters the pump intake and is imparted to the groundwater in the immediate area of the well screen. Low flow does not necessarily refer to the flow rate of groundwater discharged on the surface, which can be affected by flow regulators or restrictions. Flow rates from 0.1 to 0.5 L/min are used; however, on more porous formations, up to 1.0 L/min have been used (1).

Advantages of low-flow sampling (1) include:

- Groundwater samples are representative of the mobile load of contaminants present.
- Minimal disturbance occurs, which minimizes sampling artifacts.
- Less operator variability and greater operator control are provided.
- Reduced stress occurs on the aquifer when minimal drawdown sampling is performed.
- Reduced need exists for filtration and, consequently, less time sampling occurs.
- Smaller purging volume decreases waste disposal costs and sampling time.
- Better sample consistency exists.

Some of the disadvantages include higher initial capital costs, greater setup time in the field, additional training time, resistance to change by regulators or consultants, and the concern that the new data will trigger regulatory action based on new, more accurate concentration data (1). In some settings, low flow purging and sampling is appropriate simply because of the physical setting: low permeability formations where recharge rates are less than 0.1 L/min and fractured rock formations where representative samples may come from selectively oriented fracture networks.

The desire by the regulatory and regulated community to obtain accurate results, the realization of the problems associated with sampling, and the costs associated with conventional approaches are likely to drive the development and implementation of new standard protocols for the sampling of monitor wells. It also presents an opportunity for the development and application of new technology and approaches.

BIBLIOGRAPHY

1. Puls, R.W. and Barcelona, M.J. (1996). *Low-Flow (Minimal Drawdown) Groundwater Sampling Procedures*. EPA Groundwater Issue, Office of Research and Development, Office of Solid Waste and Emergency Response, EPA/540/S-95/504, April.

GROUNDWATER QUALITY

John E. Moore
USGS (Retired)
Denver, Colorado

The value of groundwater lies in its widespread occurrence and availability and also its consistent good quality. Groundwater, on the whole, is cleaner than most surface water and maintains a nearly uniform temperature.

As groundwater moves through the subsurface, physical, chemical, and biological processes interact to change its original quality. The earth materials through which groundwater moves act as natural filters to screen out some bacteria and some impurities from the water. Groundwater is usually clear and colorless and of excellent sanitary quality, making it drinkable directly as withdrawn, which does not mean, however, that groundwater is completely free of impurities. In fact, all groundwater contains some dissolved minerals.

Water is a universal solvent (1). From the time rain falls to the ground and begins to run off or pass into and through the soil and rocks, it dissolves these materials and thus picks up various mineral constituents. These dissolved minerals are rarely harmful to health and may give the water a pleasant taste.

The dissolved minerals include sodium, calcium, magnesium, and potassium, which combine with bicarbonates, sulfates, and chlorides. The character of the earth materials mainly determines the composition of groundwater. The longer the period contact of water with rocks, the more minerals are dissolved.

The most common groundwater quality problems are caused by natural hydrogen sulfide (2). A high content of dissolved mineral salts may affect the taste of drinking water and have laxative effects.

Hardness is caused by a high concentration of calcium and magnesium. It can cause scale formation in kettles, pipe systems, boilers, and tanks. The most common negative effect of hardness is its tendency to prevent soap from lathering and to form soap scum. Water hardness can be alleviated at a fairly reasonable cost by installing a water softener. It is not desirable, however, to remove all the minerals that make water hard. Water that is too soft is likely to corrode water distribution systems and tanks as well as to produce a flat, unpalatable taste and render the water less beneficial to the health of most people.

Excessive iron is another common natural problem. Iron concentrations in excess of 0.3 mg/L (milligrams per liter) do not normally cause health problems but are of concern for aesthetic and taste reasons. Water that has a high concentration of iron has a metallic taste and causes reddish stains on fixtures, cookware, and laundry.

Cascading or spraying the water into the air so that it is exposed to oxygen can reduce an excessive iron content. High levels of manganese cause problems similar to those caused by excessive iron. The stains caused by manganese are black and are harder to remove than iron stains.

Iron bacteria in groundwater do not damage health but are a nuisance because they clog the pores in the aquifer near wells and well screens, thereby reducing well yields. Bacteria also produce red-to-brown slime in toilet tanks and plumbing. Iron bacteria can be killed by pouring a small amount of disinfecting solution such as chlorine bleach into the water.

Hydrogen sulfide gas is recognizable by its “rotten egg” smell and taste (3). It is a poisonous gas that can build up in a small, enclosed area such as a well pump house. A person who works in a poorly ventilated space may become nauseated or even lose consciousness. Water containing hydrogen sulfide is corrosive to plumbing systems, a condition that can be reduced by aeration or by chlorination and filtration.

Water quality is commonly characterized as “safe” or “good.” “Safe” water typically means that the water is free from bacteria and disease-causing organisms, as well as from minerals and substances that can have adverse health effects. The term “good quality” water is a relative term whose meaning depends on the intended use of the water. Sanitary quality can be assessed by periodically analyzing water samples for coliform bacteria and nitrate. These substances do not normally occur in groundwater and their presence may indicate contamination.

Coliform bacteria are useful indicators of harmful microorganisms. According to the drinking-water standards of the United States, safe drinking water should not contain more than one coliform bacterium per 100 mL (6 cu in.) of water. If coliform bacteria occur in well water in numbers that indicate the water is unsafe, the water should be disinfected either by chlorinating or by boiling before drinking until further sample analyses indicate that the water is safe to drink. Sources of bacteria are septic tanks and leach fields and surface water contaminating the well.

High nitrate levels may indicate organic contamination from nearby sources of nitrate, such as barnyard drainage, animal waste storage, and percolation from agricultural land, fertilizers, and septic tanks (4). The U.S. national limit for nitrate-nitrogen concentration is 10 mg/L, which is equivalent to 45 mg/L of nitrate (2). The European standard is 11.4 mg/L nitrate-nitrogen (50 mL/L as nitrate. The main reason for this limit is the risk to (young) infants (blue babies) from drinking high nitrate waters.

Well water can be analyzed by either a country or state public health environmental laboratory or by a qualified private laboratory. Municipal authorities (local health departments) have sources on where water may be analyzed. Routine tests for coliform bacteria or inorganic contaminants, such as nitrate and salts, are relatively inexpensive. However, analyses for inorganic and organic industrial chemicals and pesticides can be very expensive. Those concerned about possible contaminants in their water supplies can contact their local health departments. Specialists from these departments can help determine if a cause for concern exists before extensive and potentially expensive water testing is performed. Keeping accurate records of the water tests ensures proper documentation of the water-quality history of a particular well. Even without obvious signs of contamination, it is advisable to have well water analyzed at least once a year.

If water contamination persists even after several tests, an alternative drinking-water supply should be used or a new source obtained. Leading causes of bacterial and nitrate contamination are poor well construction and poor well location. In most cases, nitrate contamination is limited to the shallow part of the water-bearing formation. When the contaminant is a volatile organic compound (VOC), it may cause problems, by drinking and also by inhalation and absorption through the skin.

BIBLIOGRAPHY

1. Hem, J.D. (1985). *Study and Interpretation of the Chemical Characteristics of Natural Water*. USGS Water-Supply Paper 2254.
2. Fetter, C.W. (2001). *Applied Hydrogeology*. Prentice-Hall, Englewood Cliffs, NJ, p. 598.
3. Moore, J.E. et al. (1995). *Groundwater-a Primer*. American Geological Institute, p. 53.
4. Waller, R.M. (1988). *Ground Water and The Rural Homeowner*. US Geological Survey, p. 36.

RADIAL WELLS

Dennis E. Williams
Geoscience Support Services
Claremont, California

Radial wells have a radial or horizontal component that comprises the well screen. Types of radial wells include collector wells, horizontal directionally drilled (HDD) wells, and slant or horizontal wells. The most common type of radial wells are collector wells, which have the basic design of a vertical shaft (or caisson) with horizontal screen laterals to collect groundwater. The vertical shaft typically extends to depths from 9 to 46 m (30 to 150 ft) and has a 5-m inside diameter. The laterals are typically tens to hundreds of meters in length and have a diameter of 305 mm (12 in.). The number of laterals designed for the well is dependent on aquifer characteristics (such as hydraulic conductivity) and water demand. Advances in radial well design have allowed for up to 500 m of horizontal well screen (total of all laterals) to be developed in formations where vertical wells can accommodate only 20 m of well screen. Radial wells take advantage of natural filtration through the aquifer materials, which reduces chemical treatment costs.

The lateral well screens are usually installed near the bottom of shallow aquifers in order to exploit the maximum saturated thickness of the aquifer. The laterals are installed by jacking the well screens through precast ports in the vertical shaft at one or more elevations (Fig. 1).

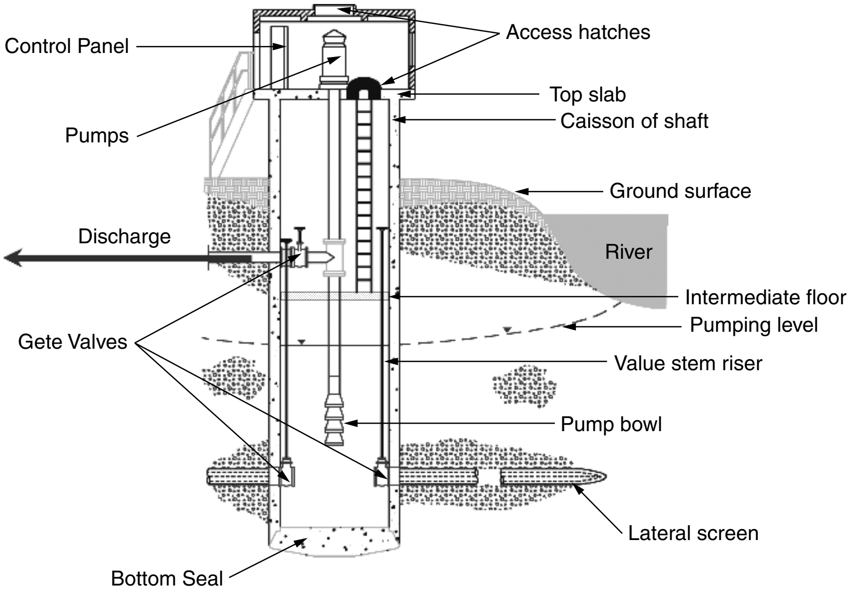

Figure 1. Lateral well schematic (after Bennett & Williams, Inc.).

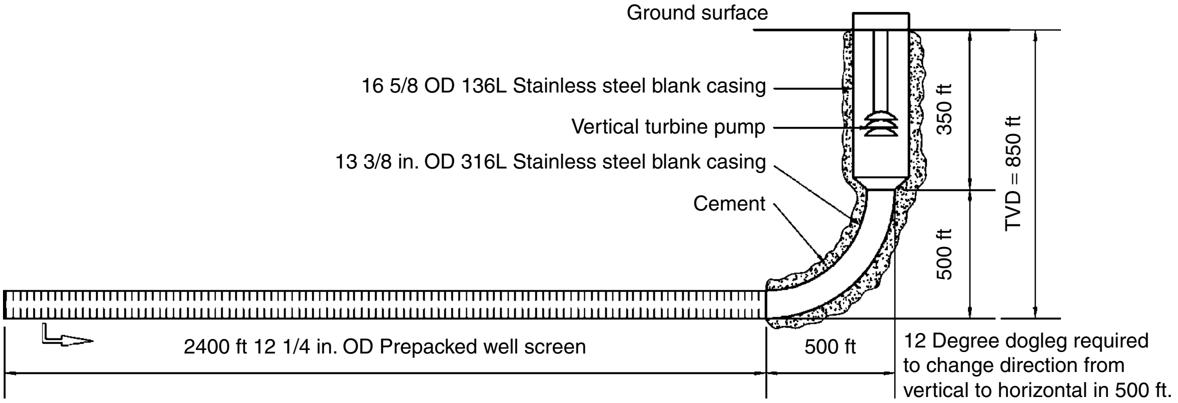

Figure 2. Horizontal directionally drilled (HDD) well.

Radial wells are reported to require less maintenance than conventional vertical wells, as the longer horizontal screens reduce water velocity entering the screen slots, thereby maximizing well efficiency and reducing well head losses and clogging. Existing radials wells can be rehabilitated with the installation of new laterals.

Radial wells are used in applications such as municipal production wells, dewatering operations, and inducing surface water infiltration for groundwater recharge.

An emerging technology in the groundwater supply industry is the use of horizontal directionally drilled (HDD) wells (Fig. 2). HDD technology is widely used in the oil and gas industry; however, HDD technology has not yet been completely transferred to the groundwater supply industry for routine or wide-scale applications. HDD wells utilize a "mud motor" technology along with sophisticated downhole navigation systems to "guide" the drill first vertically and then horizontally (or a combination of both). The application of HDD technology is in its infancy in the water well industry; however, there is potential for specialized applications such as desalination plant intake systems, or specialized recharge or enhancement of a shallow aquifer's saturated thickness. HDD wells may be constructed to extend to over 1000 m.

RECHARGE IN ARID REGIONS

Leo S. Leonhart
Hargis Associates, Inc.
Tucson, Arizona

Groundwater recharge in arid regions or deserts is a process that requires a favorable combination of hydrologic factors. Accordingly, recharge, in arid regions occurs more rarely and supplies annual quantities of water to the groundwater table lower than typical of more humid regions. Many of the processes by which recharge occurs, such as surface infiltration, stream channel loss, and seepage from impounded water bodies, however, are common to both regions.

DEFINITION OF ARID REGIONS

By definition, the aridity of a region is determined by the ratio of mean annual precipitation to mean annual potential evapotranspiration (PET). As shown in Fig. 1, arid regions in both the Northern and Southern Hemispheres are situated mostly between 10° and 35° latitude worldwide. Regions that have average annual precipitation (*P*) totaling less than 250 mm (10 in) make up an estimated 25% of the worldwide land areas outside of the polar regions (1). Arid regions are sometimes further classified as hyperarid, arid, or semiarid, according to the ranges in the *P*/PET ratio (Table 1).

VARIABILITY OF PRECIPITATION AND RECHARGE

For several reasons, estimates of mean annual precipitation are difficult to relate directly to recharge in arid regions. Largely, this difficulty develops from characteristically high temporal and spatial variability in both the distribution of precipitation and recharge in arid regions. Longer term recharge to precipitation ratios (*R*/*P*) for various arid regions in the United States generally are typically less than 5% (3–5).

Table 1. Arid Region Subclassifications

Arid Region Subclassifications	*P*/PET Ratio
Hyperarid	< 0.03
Arid	0.03 to < 0.20
Semiarid	0.20 to < 0.50

[a]After Reference 2.

High temporal variability of precipitation exists across seasonal and multiyear time frames. Average annual precipitation depths for arid regions result from large variances between extreme high and low annual totals. For example, average annual precipitation falling on the town of Arica, located in the Atacama Desert region of Chile, during a 17-year period was only 0.5 mm (0.02 in). Interestingly, this average resulted from just three measurable showers in a 17-year period (6). Such variability in precipitation infers even greater temporal variability in recharge in arid regions.

Similarly, the spatial variability of precipitation in arid regions can be high. If convective, as opposed to frontal, storms characteristically supply a significant percentage of the annual precipitation within an arid region, the areal distributions of rainfall from individual storms can be highly variable. Adjacent gauging stations can indicate significant differences in the total precipitation recorded during any given storm or during an entire year, and in the next year, a completely different pattern could be observed. Recharge, a process that requires a sufficient precipitation threshold to overcome evapotranspiration losses and soil moisture deficits in the soil, can be significantly affected by such spatial variability in precipitation.

Recharge in arid regions might be further limited due to areal variations in hydrogeologic and topographic factors. In arid regions where topographic relief is high, recharge often occurs along mountain fronts, where precipitation is greater and where there may be sustained runoff from snowpack. Other studies have shown that, in some

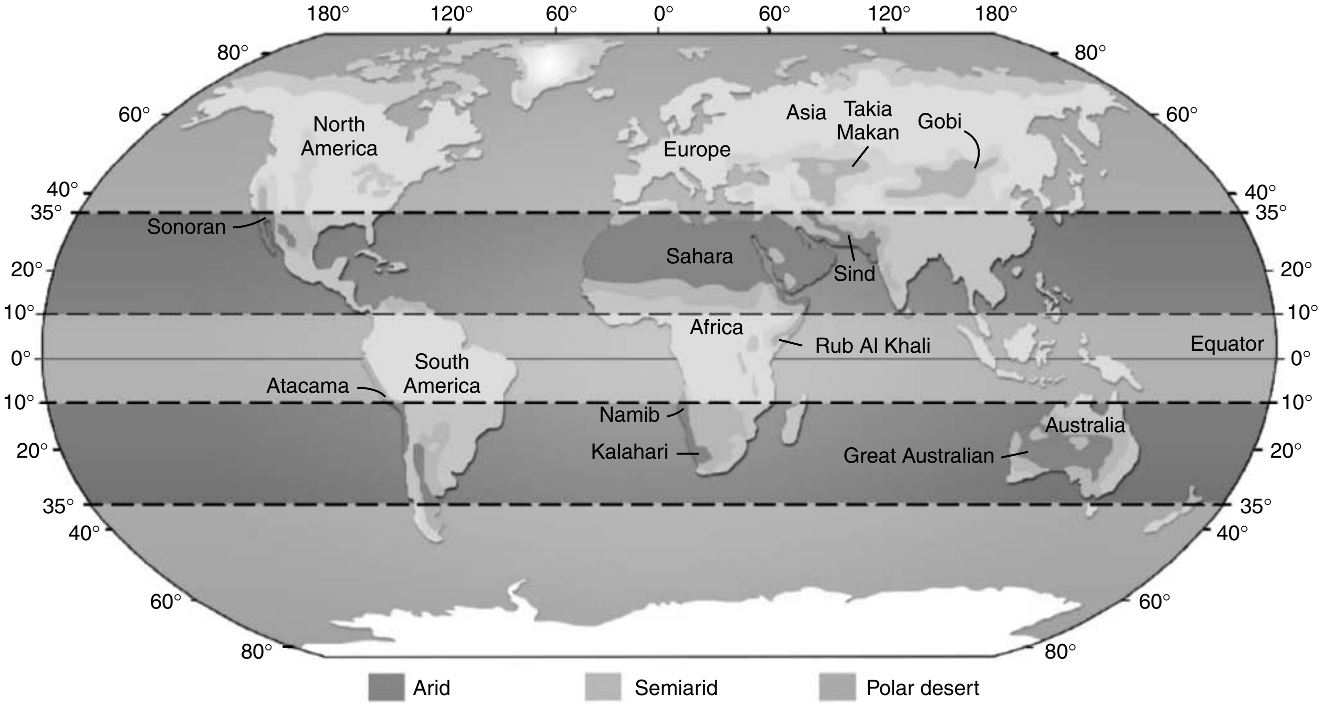

Figure 1. Worldwide pattern of arid regions (copyrighted Fig. 13–15 in Reference 1). Figure preparation/revision by Hargis and Associates, Inc., San Diego, CA.

mountainous areas, “mountain block recharge” occurs as a result of diffused infiltration through thin or absent soils overlying fractured bedrock (7). Recharge can also occur along incised alluvial channels during ephemeral or intermittent runoff. In certain low-lying areas or depressions, precipitation or runoff waters might be detained, thereby providing localized recharge through sediments or fractures. In highland areas, exposures of fractured bedrock can facilitate infiltration during precipitation or runoff, creating perched groundwater zones or deeper infiltration to the regional water table. Thus, the overlay of spatial and temporal variability of precipitation onto the areal variability of recharge areas further increases recharge variability temporally and spatially.

CONDITIONS FAVORING RECHARGE

For recharge to occur in any region, certain initial conditions must be satisfied. First, there must be sufficient surface water to infiltrate into the soil. Such conditions may result from sufficient precipitation (in terms of depth, intensity, and/or duration), runoff, or water impoundment. Second, the moisture storage potential of underlying soils must be satisfied. In other words, the void spaces in the soil matrix must become saturated. Once these conditions are met, recharge to the groundwater table becomes possible through deep percolation of soil water. Subsurface soil or bedrock conditions, however, such as fractures or heterogeneities, can sometimes locally short-circuit these overall requirements, thereby providing more direct “conduits” to the water table.

Recharge from runoff in ephemeral streams in the southwestern United States is depicted in Fig. 2. During wet seasons, sustained runoff in the alluvial channel may result in infiltration, provided there is sufficient vertical permeability in the underlying sediments. This infiltration locally recharges the groundwater table along the course of the ephemeral stream (wash or arroyo). A typical wet season recharge event is depicted in Fig. 3. During dry seasons, phreatophytes rely on the availability of shallow soil water and groundwater in the vicinity of the ephemeral stream channel, discharging groundwater by evapotranspiration. Another fraction is lost to direct surface evaporation from soils.

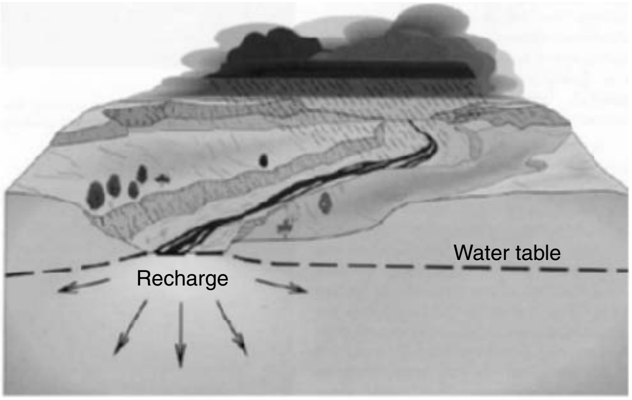

Recharge through channel bottom in wet season

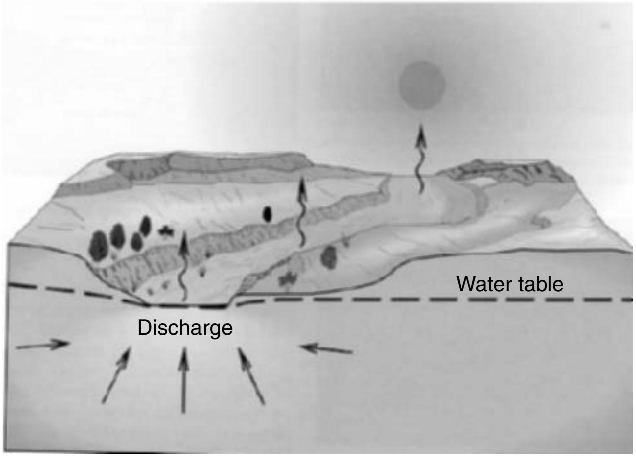

Discharge via evapotranspiration in dry season

Figure 2. Recharge from runoff in ephemeral streams.

PRECIPITATION VERSUS EVAPOTRANSPIRATION

Theoretically, soil moisture is recharged during periods when precipitation exceeds potential evapotranspiration ($P >$ PET). As discussed earlier, such conditions are rare in arid regions, which is further illustrated in the monthly average relationship between precipitation and pan evaporation (E), shown in Fig. 4. This comparison shows a site in Tucson, Arizona, elevation ~730 m (~2400 ft) above mean sea level (msl), and at Sierra Ancha, Arizona, elevation ~1555 m (5100 ft) msl. Using the average annual value of evaporation for each site as an estimate for the average annual potential evapotranspiration, the P/E ratios for Tucson and Sierra Ancha during the timeframe from 1963 through 1966 were approximately 0.15 and 0.37, respectively. The arid region subclassifications of these sites, according to Table 1 and based on the limited records evaluated, would be arid and semiarid, respectively.

From these plots, it can be seen that the recharge potential in arid regions can vary significantly with seasonal conditions. The effect of elevation differences on the P/E ratios is also evident. Both stations indicate a bimodal (double-peaked) distribution of precipitation during the year, indicating higher precipitation typically in the winter and summer months. On average, a greater amount of precipitation falls during the summer months, but the lower rate of evaporation better favors recharge potential during winter months, which is evident from the closer convergence of the P and E curves on the Tucson plot and the crossover of the two curves on the Sierra Ancha plot.

During the period of record (1963–1966), the monthly average values show that there was always a soil moisture deficit at the Tucson station, whereas a soil moisture recharge potential was present during the winter months at Sierra Ancha. These patterns help to

During runoff event

Two days later

Figure 3. Wet season recharge in Cañada del Oro Wash near Tucson, Arizona.

understand the conditions favoring recharge, but it should be understood that longer term variability might overlay these patterns. Long-term variability results from extreme conditions of drought and precipitation, earlier said to affect the calculation of average values of precipitation and evapotranspiration in arid regions. The example in Fig. 4 is based on averages calculated over a 4-year period. The averages of total annual precipitation for Tucson and Sierra Ancha during these 4 years were 340 mm (13.37 in) and 692 mm (27.25 in), respectively. Longer term records for these stations indicate lower averages of 283 mm (11.13 in) and 676 mm (26.60 in), respectively (8).

RECHARGE ESTIMATION

Recharge is a component of the hydrologic budget often calculated from a water budget equation, whereas other variables, such as precipitation, evaporation, runoff, and use quantities, usually represent field-measured parameters. Another method of estimation involves calibration of groundwater flow models against field measurements of groundwater levels. Generally, these models simulate the water budget dynamics of the study area.

Research is being performed to improve methods for quantifying recharge in arid regions. Methods under investigation include interpreting stable and radioactive isotopic signatures. The most common methods employ ratios between stable isotopes of oxygen (^{18}O/^{16}O) and hydrogen (^{2}H/^{1}H) or simply the graphical relationship between ^{2}H and ^{18}O to estimate the “isotopic age” of the water, and therefore the time since it was recharged (9). Other common methods are based on the decay constants of various radioisotopes such as tritium (^{3}H), carbon (^{14}C), and chlorine (^{36}Cl) (9,10). Recent research has examined silicon (^{32}Si) to estimate long-term recharge rates based on the “loss” of the isotope by adsorption during infiltration (11).

Other methods under current investigation for estimating recharge in arid areas include microgravity networks, temperature profiling, geochemical profiling, and various remote-sensing techniques. Microgravity networking employs finely gridded gravity instrumentation to assess seasonal and long-term changes in the mass of groundwater in storage under a basin (12). Temporal changes in the gravity profile measured across the network infer changes in the volume of groundwater stored in the underlying aquifer and hence, recharge and discharge.

Temperature profiling is based on the difference in the temperature of recharging waters versus ambient soil and groundwater. Research is focusing on both natural and artificial recharge from infiltration through stream channels as well as more diffuse recharge mechanisms. To accomplish this, multilevel temperature probes are emplaced at successive depths within soil borings drilled in known or suspected recharge areas.

Geochemical methods also require depth profiling in the soil column beneath suspected recharge areas. The concept relies on interpreting changes in soil chemistry from water movement during infiltration. Such changes result from either deposition or dissolution of minerals along the infiltration pathway. For example, using a combination of field-measured chloride data and numerical simulations, researchers have estimated long-term recharge cycles in the Badain Jaran Desert in China as far back as 1000 years (13).

Remote sensing techniques involve interpretation of tones, colors, and spectral hues observable in bare soils as well as due to vegetation densities and spectral variations. Such differences can be interpreted from photographic or scanned imagery and infer regions where shallow subsurface water is present. Infrared techniques employ either thermal- or near-infrared spectra. Thermal techniques detect temperature differences in soil moisture variations, whereas near-infrared relies on vegetation inferences.

BIBLIOGRAPHY

1. Skinner, B.J, Porter, S.C., and Botkin, D.B. (1999). *The Blue Planet, An Introduction to Earth System Science*, 2nd Edn. John Wiley & Sons, New York.
2. UNESCO. (1979). Map of the world distribution of arid regions: explanatory note. In: *Understanding Water in a Dry Environment: Hydrological Processes in Arid and SemiArid Zones*. International Association of Hydrogeologists, Balkema, Lisse, The Netherlands, Vol. 23.

(**a**) Sierra Ancha, Arizona (1963–66)

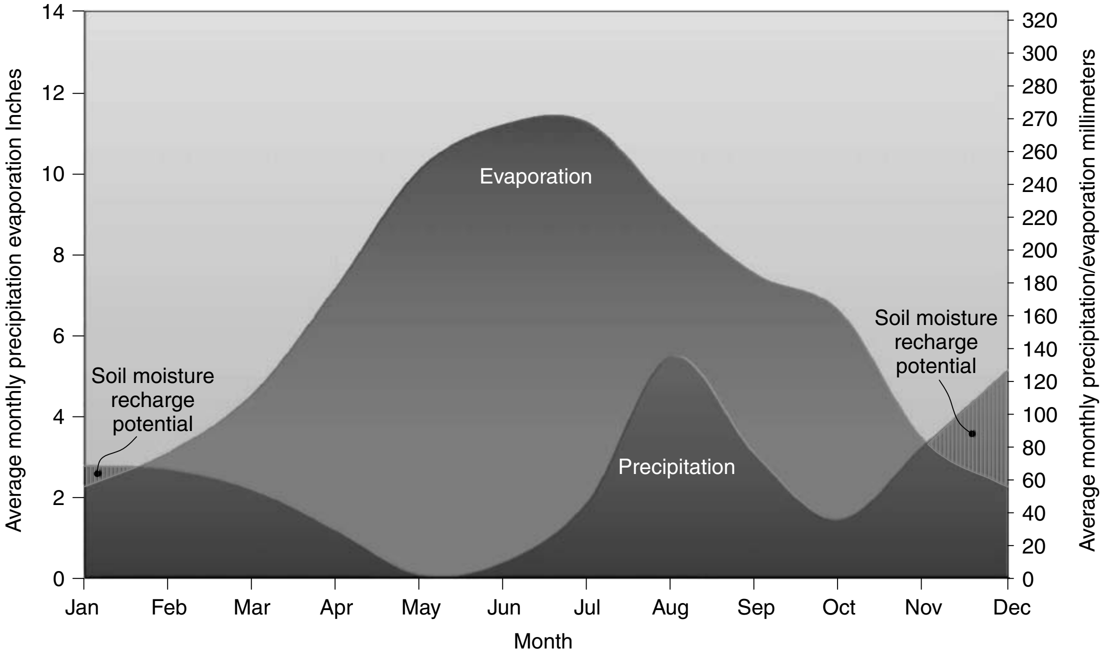

(**b**) University of Arizona, Tucson, Arizona (1963–66)

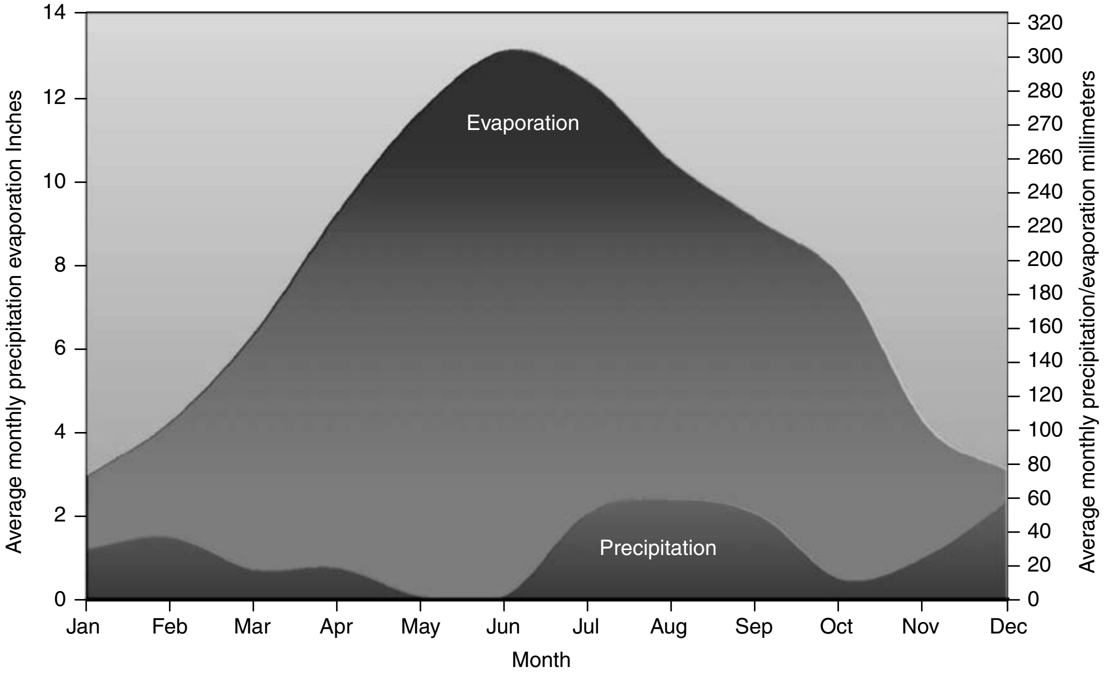

Figure 4. (**a**) Relationships between average monthly precipitation and evaporation at Tucson and Sierra Ancha, Arizona; (**b**) Comparison of soil moisture recharge potential in Tucson vs. Sierra Ancha, Arizona (data from Sellers & Hill, 1974) (Ref. 14).

3. Harshbarger, J.W. (1959). Geohydrology of Arid Lands: Arizona—A Case Study. In: *Arid Lands Colloquia, 1958–1959*. University of Arizona Press, Tucson, AZ. Available: http://alic.arid.arizona.edu/sonoran/documents/harshbarger/alc_geohydrology.html.
4. Sophocleous, M. (Ed.). (1998). *Perspectives on Sustainable Development of Water Resources in Kansas*. Kansas Geological Survey: Bulletin 239. Lawrence, KS.
5. Hevesi, J.A., Flint, A.L., and Flint, L.E. (2002). *Preliminary Estimates of Spatially Distributed Net Infiltration and Recharge for Death Valley Region, Nevada-California*. U.S. Geological Survey: Water-Resources Investigations Report 02-4010.
6. McDonald, J.E. (2002). *Climatology of Arid Lands*. Available: http://alic.arid.arizona.edu/sonoran/documents/alc_climatology_aridlands.html.

7. Long, A. and Estoe, C. (2003). *Mountain-Block Recharge*. Research abstract for work sponsored by SAHRA. Available: http://www.sahra.arizona.edu/research/TA2/2_5.html.
8. Sellers, W.D., Hill, R.H., and Sanderson-Rae, M. (1985). *Arizona Climate, One Hundred Years, 1885–1985*. University of Arizona, Tucson, AZ.
9. Davis, S.N., Campbell, D.J., Bentley, H.W., and Flynn, T.J. (1985). *Ground Water Tracers*. National Ground Water Association, Dublin, OH.
10. Faure, G. (1986). *Principles of Isotope Geology*, 2nd Edn. John Wiley & Sons, New York.
11. Ekwurzel, B. and Einloth, S. (2003). *Quantifying Groundwater Recharge Using ^{32}Si*. Research abstract for work sponsored by SAHRA. Available: http://www.sahra.arizona.edu/research/TA2/2_4.html.
12. Pool, D.R. and Schmidt, W. (1997). *Measurement of Ground-Water Storage Change and Specific Yield Using Temporal-Gravity Method Near Rillito Creek, Tucson, Arizona*. U.S. Geological Survey: Water Resources Investigation Report 97-4125.
13. Jinzhu, M., Ding, L., Jiawu, Z., and Edmunds, W.M. (2003). Groundwater recharge and climatic change during the last 1000 years from unsaturated zone of SE Badain Jaran Desert. *Chin. Sci. Bull.* **48**(14): 1469–1474.
14 Sellers, W.D. and Hill, R.H. (1974). *Arizona Climate, 1931–1972*. University of Arizona Press, Tucson, AZ.

SUB-SURFACE REDOX CHEMISTRY: A COMPARISON OF EQUILIBRIUM AND REACTION-BASED APPROACHES

Naresh Singhal
University of Auckland
Auckland, New Zealand

Jahangir Islam
City Design Limited
Auckland, New Zealand

INTRODUCTION

The study of subsurface redox chemistry has gained importance because of increased awareness of the relationships between chemical speciation and contaminant behavior in the environment. The oxidation state of a redox component can affect its solubility, sorption to soil, and toxicity. Knowledge of redox processes is therefore fundamental to understanding the fate and transport of organic and inorganic constituents in both pristine and contaminated environments.

Redox reactions involve oxidation, and accompanying reduction, of the reacting species because of transfer of electrons. Such reactions, directly or indirectly, determine the fate and transport of several biologically important elements such as carbon, nitrogen, and sulfur in the soil environment. Elements of environmental significance undergoing redox reactions include heavy metals like lead and zinc, which are multivalent and readily undergo a change in valence.

Redox reactions in natural environments can be biotic or abiotic. Most redox processes encountered in natural aquatic systems are biologically catalyzed or at least indirectly affected by living microorganisms. In subsurface systems, such as groundwater aquifers with microbial activity, the redox state is significantly influenced by microbe-mediated redox reactions. In such environments, a sequence of distinct redox zones can develop depending on the composition of aquifer materials and aqueous constituents and the relative rates of redox reactions and the transport of products or reactants.

Frequently, redox reactions in subsurface environments are evaluated with equilibrium thermodynamics. Although equilibrium thermodynamics is useful for visualizing the potential relationships between the oxidized and reduced phases of various elements, in many cases, redox equilibrium is not established because of the presence of living organisms, dependence of redox transformations on biological catalysis, and the slow kinetics of oxidation and reduction reactions (1). Time spans for attainment of equilibrium for dissolution/precipitation of secondary phases and redox reactions may range from a few hours to years. Many groundwaters are not well posed, and speciation calculations for numerous groundwaters have shown that various redox couples typically exist in a state of disequilibrium. When disequilibrium exists, geochemical model based on equilibrium thermodynamics will provide incorrect speciation results for the redox species. Reliable predictions of solid-phase dissolution/precipitation and redox reactions require knowledge of reaction kinetics. This article presents a comparison of equilibrium and reaction-based kinetic approaches for evaluating subsurface redox chemistry.

MODELING APPROACHES

The two approaches that compute the distribution of chemical species in the subsurface are the kinetic and equilibrium models. Both approaches use the concept of components and species to model systems with multiple chemical reactants. Components are a set of linearly independent chemical entities such that every species can be written as the product of a reaction involving only the components, and no component can be written in terms of components other than itself. For a given system, the set of components is not unique, but once it has been defined, the representation of species in terms of this set of components is unique. The number of components is selected such that it is the minimum number of the set of species necessary to describe the composition of all phases and species in the chemical system.

Equilibrium models focus on the final equilibrium state of a solution composed of several components in a closed system (2). Implicit in this approach is the assumption that the modeled reactions are fast and fully reversible, which results in natural systems with areas where thermodynamic equilibrium has been established, at least locally. Both of the above conditions are frequently violated, as many redox reactions are slow and irreversible. For example, consider the oxidation of organic matter, which is usually an irreversible reaction, and the diagenesis of sedimentary rock, which occurs over decades. The equilibrium composition of soil systems is generally calculated using two methods—the equilibrium constant method (3) and the free energy minimization technique (4). The equilibrium constant method uses the mass action principle to relate the redox species involved in specific reactions. The equilibrium composition is obtained by solving the mass action relations used along with the mass balance equations for different components. The free energy minimization technique involves combining the chemical potential of each species with the mass balance equation for each component. The equilibrium composition is obtained by minimizing the free energy of the system (5). Of these two methods, the equilibrium composition approach has been more commonly adopted in the models developed to date.

Kinetic formulations, typically used in reactive transport modeling, describe the transformation of redox species using rate expressions for reactions affecting specific redox species. These reactions may include sorption to soil, degradation, dissolution/precipitation, complexation, and volatilization to gaseous phase. Additionally, these reactions may be biotic, thereby requiring an additional consideration of biological growth and accompany organic matter degradation. The kinetic models can be grouped, on the basis of the logic coupling the transport and reaction equations, into two categories: two-step and one-step methods. The two-step methods involve solving the transport equations followed by the chemical reaction equations at each time step, whereas the one-step methods involve incorporating the chemical reaction equation into the transport equation and solving the resulting set of nonlinear equations explicitly or implicitly. The use of the kinetic approach to model redox reactions in groundwater systems is a relatively recent development, and much work is continuing in this area.

Equilibrium Models

The equilibrium models describe multispecies reactive chemical transport based on principles of thermodynamics, and they can describe complexation, sorption to soils, and removal via precipitation in batch and flow through systems at equilibrium. The approaches used by different investigators to model redox processes in groundwater systems can be grouped into four categories (2): effective internal approach, external approach, oxygen fugacity approach, and redox couple approach. The first three approaches assume that the system has achieved thermodynamic equilibrium, whereas the last approach handles redox processes under partial equilibrium conditions.

The effective internal approach (6,7) is based on the principle of conservation of electrons, although free electrons do not exist in the aqueous phase (2). Therefore, a redox reaction involving the loss of an electron through oxidation must be accompanied by a simultaneous reaction involving the gaining of electron through reduction. The sum of the valence states of each species in the system remains constant as chemical redox reactions progress. For redox-sensitive species, their redox states equal their valences.

The external approach (8), unlike the internal approach, considers hypothetical electron activity as an aqueous component. For computational purposes, the redox reactions are considered mathematically equivalent to complexation reactions in which the hypothetical electrons are incorporated as a component (2). Reduction half-reactions involve higher valence species reacting with the hypothetical electrons to form lower valence species. For each set of redox species, the ionic species in the highest oxidation state is chosen as master component and all other oxidation states are formed from it.

The oxygen fugacity approach uses oxygen fugacity as a redox parameter in the system (9). The mass balance equation of oxygen provides an additional equation to solve the system of chemical speciation equations involving redox reactions. The use of oxygen fugacity as a redox potential indicator can be illustrated using the following half-cell reaction for water:

$$2\mathrm{H_2O(l)} = \mathrm{O_2(g)} + 4\mathrm{H}^+ + 4e^- \tag{1}$$

Then the redox potential (Eh) and the oxygen fugacity ($\mathrm{f_{O_2}}$), using the equilibrium constant $\mathrm{K_{O_2(g)/H_2O(l)}}$ for the half-reaction, can be related by (10)

$$\log(\mathrm{f_{O_2}}) = \left(\frac{4\mathrm{F}}{2.303\mathrm{RT}}\right)\mathrm{Eh} - 4\log\{\mathrm{H}^+\} + 2\log\{\mathrm{H_2O(l)}\} + \log \mathrm{K_{O_2(g)/H_2O(l)}} \tag{2}$$

where the terms within {} represent the activities of the enclosed species.

The fourth concept, called the redox couple approach (6,9), assumes that partial equilibrium conditions prevail in the system. This approach separates each redox couple into two master species that give rise to two mass balance equations. After the complete calculation of chemical speciation, the redox potential is determined from the activities of this couple and its equilibrium constant (2). Redox couples may be preferred for modeling the redox processes if the analytical data for the couples are available. As the subsurface system may have more than one redox couple, the redox couple approach can test whether these redox couples are at equilibrium with each other. The applicability and validity of redox sequence concept can also be tested. If the aqueous solution exists at a single redox potential, the redox sequence concept requires that all redox reactions in the aqueous solution be at thermodynamic equilibrium. Using the concept of redox zones, the dominant redox couple, which controls the redox potential in a given redox zone, can determine the redox potential of the system. The use of redox zones for classifying the subsurface environment is an important concept in the numerical simulation of subsurface redox chemistry.

Kinetic Models

If the reactions are insufficiently fast, because of chemical kinetics or mass diffusion limitations, relative to the macroscopic transport processes affecting solute concentration (e.g., advection, hydrodynamic dispersion), the assumption of thermodynamic equilibrium, or even local equilibrium, may not apply. In subsurface environments, with the exception of aqueous redox reactions, it is usually valid to model most homogeneous aqueous-phase reactions (including complexation and acid-base reactions) using equilibrium thermodynamics as they are characterized by rapid rates, especially in comparison with subsurface flow. However, increasing evidence suggests that heterogeneous reaction processes (such as adsorption, biodegradation, and precipitation/dissolution) may need to be described by a kinetic model. In a kinetic model for flow-through systems, algorithms have been proposed that integrate the flow processes (advection and dispersion) with the transformation processes. A flow chart that illustrates a typical algorithm adopted for this purpose is presented in Fig. 1.

Various kinetic formulations have been proposed for adsorption, biodegradation, and precipitation of species present in soils. Some common kinetic equations used in the literature for these processes are discussed below.

Adsorption. The kinetic formulations for adsorption to soils include those based on mass transfer, first-order reaction, and nth-order reaction (11).

$$\text{Mass transfer}: \quad \frac{\partial C_s}{\partial t} = \alpha(f(C_w) - C_s) \qquad (3)$$

$$\text{First-order reaction}: \quad \frac{\partial C_s}{\partial t} = k_1 \frac{\theta}{\rho_b} C_w - k_2 C_s \qquad (4)$$

$$\text{nth-order reaction}: \quad \frac{\partial C_s}{\partial t} = k_1 \frac{\theta}{\rho_b} C_w^n - k_2 C_s \qquad (5)$$

where α is a mass transfer coefficient, θ is the soil porosity, ρ_b is the soil bulk density, $f(C_w)$ is some function of the concentration of dissolved chemical, C_w [M/L^3], k_i ($i = 1,2$) are reaction rate constants, and C_s is the mass of chemical sorbed onto the surface of intraparticle pores of the grains per unit mass of the grains (dimensionless).

Biodegradation. The primary microbial redox reactions typically observed in subsurface systems are given in Table 1. The most energetically favorable of these

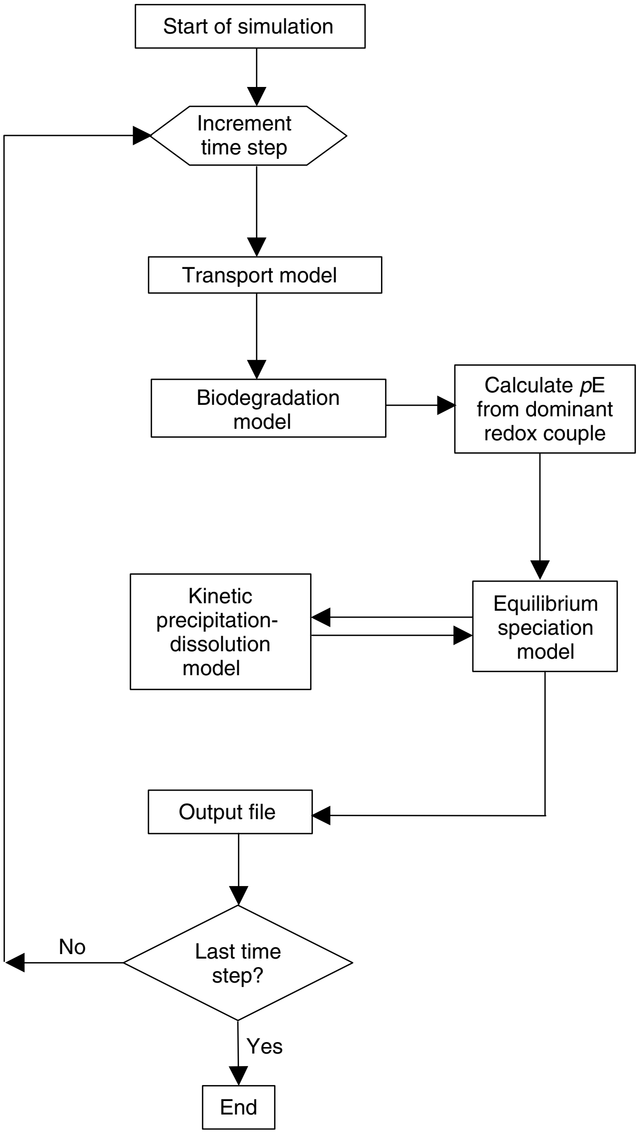

Figure 1. A typical algorithm for kinetic redox modeling in a flow through system.

Table 1. Primary Microbe Mediated Redox Reactions in Subsurface Environments

Redox Reactions	ΔG^o (kcal/mol) (pH = 7)
Aerobic respiration, oxygen reduction $CH_2O + O_2 \rightarrow CO_2 + H_2O$	−120
Denitrification $5CH_2O + 4NO_3^- + 4H^+ \rightarrow 5CO_2 + 7H_2O + 2N_2$	−114
Manganese (manganic) reduction $CH_2O + 2MnO_2 + 4H^+ \rightarrow CO_2 + 2Mn^{2+} + 3H_2O$	−81
Iron (ferric) reduction $CH_2O + 4Fe(OH)_3 + 8H^+ \rightarrow CO_2 + 4Fe^{2+} + 11H_2O$	−28
Sulfate reduction $2CH_2O + SO_4^{2-} + H^+ \rightarrow 2CO_2 + HS^- + 2H_2O$	−25
Methanogenic processes $CH_3COOH \rightarrow CH_4 + CO_2$	−22

Table 2. Kinetic Formulations for Ion Precipitation

Model	Rate Equation	Source
Davies and Jones Model	$R = sk\gamma^2\left(\{\mathrm{Ca}^{+2}\}^{1/2}\{\mathrm{CO}_3^{-2}\}^{1/2} - \frac{K_{\mathrm{sp}}^{1/2}}{\gamma}\right)^2$ k = rate constant, s = surface area of the crystal, K_{sp} = solubility product of $\mathrm{CaCO_3}$, and γ = divalent ion activity coefficient calculated using the Davies equation.	Inskeep and Bloom (13)
Empirical Model	$R_{\mathrm{P}} = sk_{\mathrm{p}}(\Omega - 1)^{\mathrm{n}}$ $R_{\mathrm{D}} = sk_{\mathrm{D}}(\Omega - 1)^{\mathrm{n}}$ Ω = disequilibrium ratio = $\{\mathrm{Ca}^{+2}\}\{\mathrm{CO}_3^{-2}\}/K_{\mathrm{sp}}$ k_{d} and k_{p} = rate constants for dissolution and precipitation, respectively, and n = reaction order.	Inskeep and Bloom (13)
House Model	$R = k((\{\mathrm{Ca}^{+2}\}\{\mathrm{CO}_3^{-2}\})^{1/2} - K_{\mathrm{sp}}^{1/2})^2$	Van Langerak et al. (15)
Nancollas and Reddy Mechanistic Model	$R = sk\gamma^2\left(\{\mathrm{Ca}^{+2}\}\{\mathrm{CO}_3^{-2}\} - \frac{K_{\mathrm{sp}}}{\gamma^2}\right)$	Inskeep and Bloom (13)
Zhang and Dawe Model	$R = k\left(\left(\frac{\left\{\mathrm{Ca}^{+2}\right\}\left\{\mathrm{CO}_3^{-2}\right\}}{K_{\mathrm{sp}}}\right)^{0.5} - 1\right)^{\mathrm{n}}$	Zhang and Dawe (16)

reactions is aerobic respiration, followed by, in order of decreasing free energy yield, denitrification, Mn(IV) reduction, Fe(III) reduction, sulfate reduction, and methanogenesis. The differences in energy yields of these reactions may lead to physical separation (spatial or temporal) of the zones in which each reaction dominates. The rates of reduction of the terminal electron acceptor and the corresponding oxidation of organic substrate are usually represented by zero-order, first-order, and multiple Monod-type expressions. Application of Monod rate expression is justified when the organic compound of interest acts as a primary energy substrate for the microbial community. Assuming an absence of diffusional resistance biodegradation of the substrate and microbial growth in soils can be represented by the following system of equations using the multiple Monod kinetics (12):

$$\frac{\partial S}{\partial t} = -\frac{\mu_{\max}}{Y}\frac{S}{K_{\mathrm{s}}+S}\frac{C_{\mathrm{w,ea}}}{K_{\mathrm{ea}}+C_{\mathrm{w,ea}}}\left(X_{\mathrm{a}} + \frac{\rho_{\mathrm{b}}X_{\mathrm{s}}}{\theta}\right) \tag{6}$$

$$\frac{\partial X_{\mathrm{w}}}{\partial t} = \mu_{\max}\frac{S}{K_{\mathrm{s}}+S}\frac{C_{\mathrm{w,ea}}}{K_{\mathrm{ea}}+C_{\mathrm{w,ea}}}X_{\mathrm{w}} - K_{\mathrm{dec}}X_{\mathrm{w}} - K_{\mathrm{att}}X_{\mathrm{w}} + \frac{\rho_{\mathrm{b}}K_{\mathrm{det}}X_{\mathrm{s}}}{\theta} \tag{7}$$

$$\frac{\partial X_{\mathrm{s}}}{\partial t} = \mu_{\max}\frac{S}{K_{\mathrm{s}}+S}\frac{C_{\mathrm{w,ea}}}{K_{\mathrm{ea}}+C_{\mathrm{w,ea}}}X_{\mathrm{s}} - K_{\mathrm{dec}}X_{\mathrm{s}} - K_{\mathrm{det}}X_{\mathrm{s}} + \frac{\theta K_{\mathrm{att}}X_{\mathrm{a}}}{\rho_{\mathrm{b}}} \tag{8}$$

where S is the substrate concentration in the bulk (mobile) fluid ($\mathrm{M/L^3}$), X_{w} is the aqueous-phase biomass concentration ($\mathrm{M/L^3}$), X_{s} is the solid-attached biomass concentration (M/M), $C_{\mathrm{w,ea}}$ is the concentration of terminal electron acceptor ($\mathrm{M/L^3}$), K_{dec} is the first-order endogenous decay coefficient (1/T), K_{att} is the biomass attachment coefficient (1/T), K_{det} is the biomass detachment coefficient ($1/T$), $\mu_{\max}$ is the maximum specific growth rate of the biomass ($1/T$), Y is the yield coefficient for the biomass (cell mass produced per mass of substrate consumed), and K_{s} and K_{ea} are the saturation constant for substrate and electron acceptor ($\mathrm{M/L^3}$), defined as the concentration for which the specific growth rate is equal to $\mu_{\max}/2$.

Precipitation and Dissolution. Precipitation and dissolution processes have often been described by chemical reactions whose rate is proportional to the deviation from equilibrium (13,14). One such kinetic formulation can be written as follows:

$$\frac{dC_{\mathrm{j}}}{dt} = -a_{\mathrm{j}}k\left(\prod_{\mathrm{j}=1}^{\mathrm{N}}\{u_{\mathrm{j}}\}^{a_{\mathrm{j}}} - K_{\mathrm{sp}}\right) \tag{9}$$

where k is the precipitation rate coefficient, N is the number of components forming the solid precipitate, C_{j} is the dissolved concentration of component j, $\{u_{\mathrm{j}}\}$ is the dissolved free ion activity of component j, a_{j} is the stoichiometric coefficient of the jth component of the solid, and K_{sp} is the equilibrium solubility product for the solid. Several similar formulations for ion precipitation, using $\mathrm{CaCO_3}$ as the precipitating mineral, are presented in Table 2.

MODEL LIMITATIONS AND APPLICABILITY

Geochemical models suffer from inadequacies in two main areas (17):

1. Violation of (local) equilibrium assumption in many environmental systems. Although such processes can be modeled using kinetics, rather than equilibrium, our current knowledge of kinetic rates and mechanisms is primitive.
2. The knowledge of adsorption of solutes on mineral surfaces, an important environmental process, is also at an early stage of understanding.

The "geochemical (and hydrological) models of natural and engineered systems at the present time are, at best, approximations to the real situation. To meet various regulatory requirements, these approximations, and predictions based on these approximations, must be made, and there is no better way to proceed than to construct models based on fundamental science. To this extent, geochemical and hydrological modeling is useful" (17).

BIBLIOGRAPHY

1. Sigg, L. (2000). Redox potential measurements in natural waters: Significance, concepts and problems. In: *Redox—Fundamentals, Processes and Applications*. J. Schüring, H.D. Schulz, W.R. Fischer, J. Böttcher, and W.H.M. Duijnisveld (Eds.). Springer-Verlag, Berlin, pp. 1–12.
2. Liu, C.W. and Narasimhan, T.N. (1989). Redox-controlled multiple-species reactive chemical transport—1. Model development. *Water Resour. Res.* **25**(5): 869–882.
3. Allison, J.D., Brown, D.S., and Novo-Gradac, K.J. (1990). *MINTEQA2/PRODEFA2, A Geochemical Assessment Model for Environmental Systems: Version 3.0 User's Manual*. Environ. Res. Lab., USEPA, Athens, GA.
4. Mattigod, S.V. (1995). Chemical equilibrium and reaction models; Applications and future trends. In: *Chemical Equilibrium and Reaction Models*. R.H. Loeppert, A.P. Schwab, and S. Goldberg (Eds.). SSSA Special Publication Number 42, Soil Science Society of America, Madison, WI, pp. 1–5.
5. Waite, T.D. (1989). Mathematical modeling of trace element speciation. In: *Trace Element Speciation: Analytical Methods and Problems*. R.H. Loeppert et al. (Eds.). CRC Press, Boca Raton, FL, pp. 117–184.
6. Parkhurst, D.L., Thorstenson, D.C., and Plummer, L.N. (1980). *PHREEQE: A Computer Program for Geochemical Calculations*. U.S. Geol. Surv. Water Resour. Invest., PB81–167801.
7. Walsh, M.P., Bryant, S.L., Schecter, S.R., and Lake, L.W. (1984). Precipitation and dissolution of solids attending flow through porous media. *AICHE J.* **30**(2): 312–328.
8. Carnahan, C.L. (1986). *Simulation of Uranium Transport with Variable Temperature and Oxidation Potential: The Computer Program THCC, Rep. 21639*. Lawrence Berkeley Lab., Berkeley, CA.
9. Wolery, T.J. (1983). *EQ3NR, A Computer Program for Geochemical Aqueous Speciation-Solubility Calculations, User's Guides and Documentation, Rep. UCRL-53414*. Lawrence Livermore National Lab., Livermore, CA, p. 191.
10. Stumm, W. and Morgan, J.J. (1981). *Aquatic Chemistry*. 2nd Edn. Wiley-Interscience, New York, p. 780.
11. Habbar, A., Kolditz, O., and Zielke, W. (2002). New geochemical simulator rockflow-RTM: development and benchmarking. In: *Geochemical Processes, Conceptual Models for Reactive Transport in Soil and Groundwater*. W-VCH Verlag Gmbh & Co. KGaA, Weinheim, Germany, pp. 95–114.
12. Islam, J. and Singhal, N. (2002). A one-dimensional reactive multi-component landfill leachate transport model. *Environ. Model. Softw.* **17**: 531–543.
13. Inskeep, W.P. and Bloom, P.R. (1985). An evaluation of rate equations for calcite precipitation kinetics at pCO_2 less than 0.01 atm and pH greater than 8. *Geochim. Cosmochim. Acta* **49**: 2165–2180.
14. Smith, S.L. and Jaffé, P.R. (1998). Modeling the transport and reaction of trace metals in water-saturated soils and sediments. *Water Resour. Res.* **34**: 3135–3147.
15. Van Langerak, E.P.A. et al. (2000). Impact of location of CaCO3 precipitation on the development of intact anaerobic sludge. *Water Res.* **34**(2): 437–446.
16. Zhang, Y. and Dawe, R. (1998). The kinetics of calcite precipitation from a high salinity water. *Appl. Geochem.* **13**(2): 177–184.
17. Zhu, C. and Anderson, G. (2002). *Environmental Applications of Geochemical Modeling*. Cambridge University Press, Cambridge, UK, p. 284.

REGIONAL FLOW SYSTEMS

Matthew M. Uliana
Texas State University—San Marcos
San Marcos, Texas

DEFINITION

A **groundwater flow system** is a three-dimensional entity that has the following components:

—a recharge area where water enters the flow system
—a discharge area where water exits the system
—boundary conditions and physical dimensions

When dealing with flow systems, we often differentiate between **local, intermediate,** and **regional** flow systems. The distinction is primarily a matter of scale; local flow systems cover a much smaller area than intermediate systems, which in turn are smaller than regional systems. **Regional groundwater flow systems** generally involve a large area (10^1–10^2 km), and usually contain several local flow systems, resulting in areas of local recharge as well as springs or other discharge features within the larger system. Often several different aquifers and geologic formations will be connected by a larger regional groundwater system.

Examples of regional flow systems are shown in Fig. 1.

The following is a discussion of each of the three components.

RECHARGE

Recharge refers to the water that is entering a groundwater system. Areas where recharge is occurring are called **recharge areas** or **recharge zones**. There are several different ways that recharge enters a flow system. Recharge can enter a flow system through direct infiltration of precipitation on the outcrop, or by infiltration through the beds of losing streams or reservoirs. In some parts of the world, especially arid regions, infiltration of storm runoff through the beds of intermittent streams is the dominant form of recharge. Water can also enter a flow system through interformational flow, which is usually in the form of flow through leaky confining layers where water is drawn in by drawdowns at wells, or where

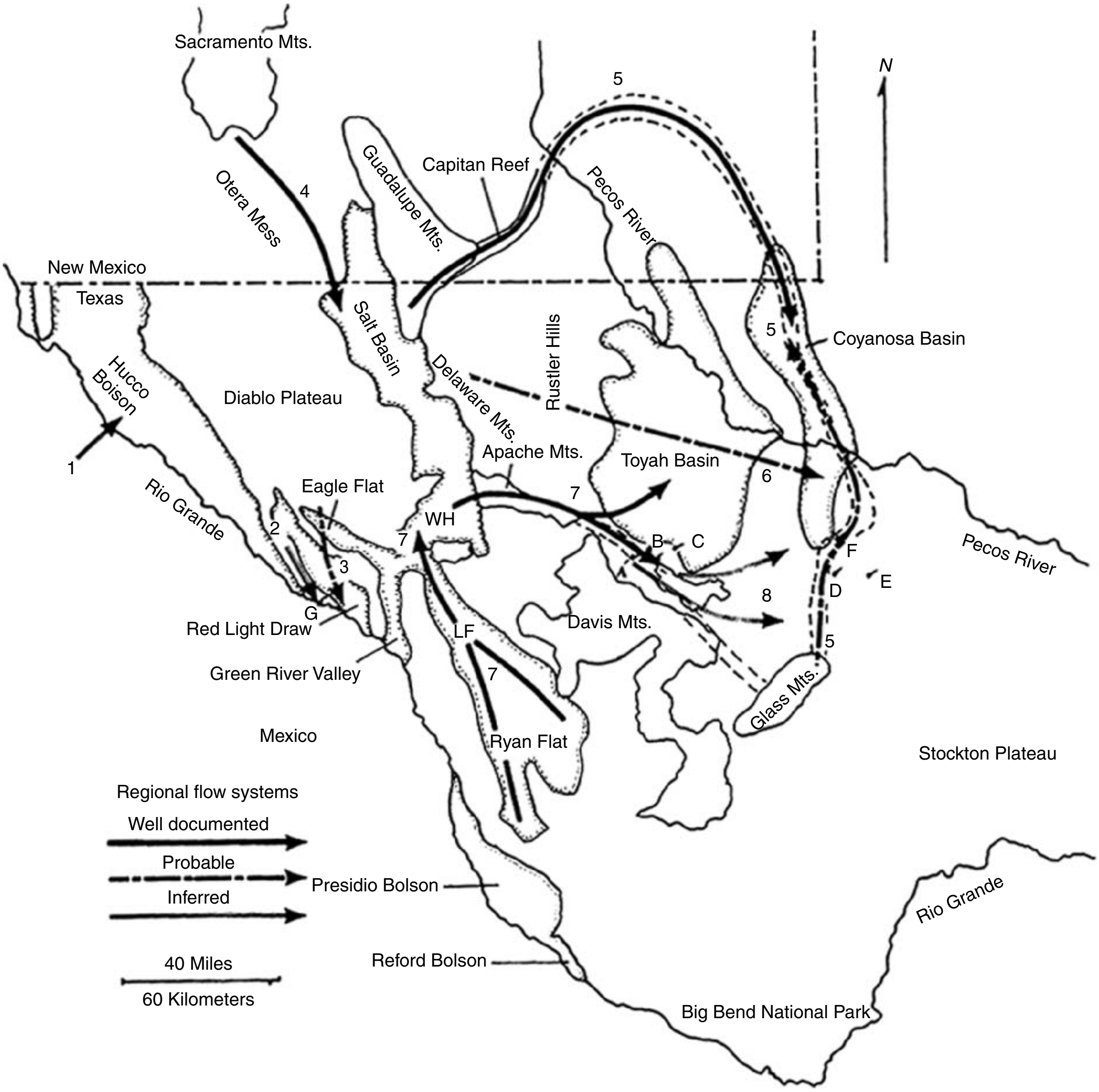

Figure 1. Regional groundwater flow systems of West Texas (after Reference 1).

underlying aquifers have significant overpressuring and are forcing water upwards.

For modeling and other research purposes, recharge is generally considered as some percentage of precipitation. Annual recharge rates in most aquifers are usually between 1–10% of annual precipitation rates. Recharge rates are difficult to quantify; many methods involve measuring precipitation and performing a water balance by quantifying all the other fluxes of water (i.e., surface runoff, evaporation, transpiration) at the surface.

DISCHARGE

There are several ways that water discharges from a flow system. These include discrete discharge to a spring or seep, discharge into a gaining stream or lake, interformational flow, or pumping from a well. In some arid regions, direct evaporation from a shallow water table is the primary mechanism for discharge from the flow system.

Discharge can also be hard to quantify, especially in areas dominated by well pumping or evaporation. Interformational flow is generally small compared to the other mechanisms. Springs and gaining streams can be gauged, and changes in flow across a certain area can be attributed to either recharge or discharge.

BOUNDARY CONDITIONS

Flow systems are three-dimensional bodies that have boundaries. We can consider two basic types of boundaries, or **boundary conditions**, that can characterize the limits of flow systems.

A **no-flow boundary** has a hydraulic gradient of zero ($h/x = 0$), and therefore no flow across the boundary. No-flow boundaries can be actual physical boundaries, such as where permeable aquifer units are in contact with low permeability bedrock at a fault (Fig. 2). A no-flow boundary also exists where flow lines are parallel, creating a **symmetry boundary** (Fig. 3). Modelers often use symmetry boundaries to constrain numerical groundwater models of aquifers. Locally high water levels can create a type of no-flow boundary called a **groundwater divide** (Fig. 4). At a groundwater divide, water flows away from the divide on either side (similar to surface runoff at a drainage divide).

A **constant-head boundary** is one characterized by a set of hydraulic heads that do not change. A nonflowing body of water, such as a lake, pond, or ocean, can create a constant-head boundary (Fig. 5). In this case, the shore of the body of water represents a single **equipotential line** (or line of constant head) in the aquifer, and flow is perpendicular to the shoreline (either into the aquifer from the surface waterbody, or vice versa). A stream can also act as a constant-head boundary (Fig. 6). In a stream boundary, the actual heads will vary along the stream gradient, but each point is considered constant. In this case, each point along the stream gradient represents a point on an equipotential.

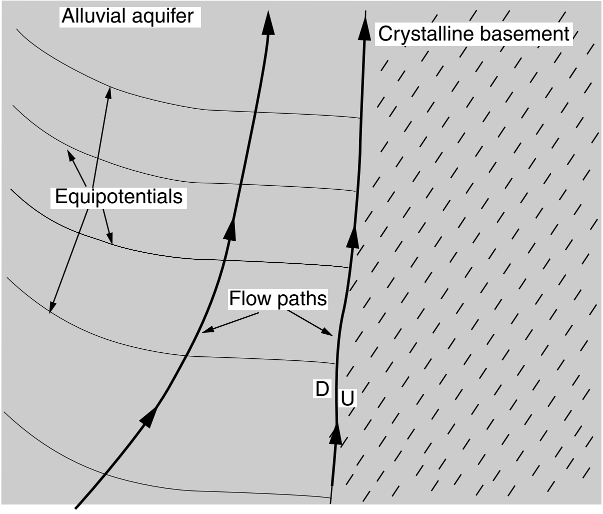

Figure 2. No-flow boundary created by juxtaposition of permeable alluvial sediments and low-permeability crystalline bedrock.

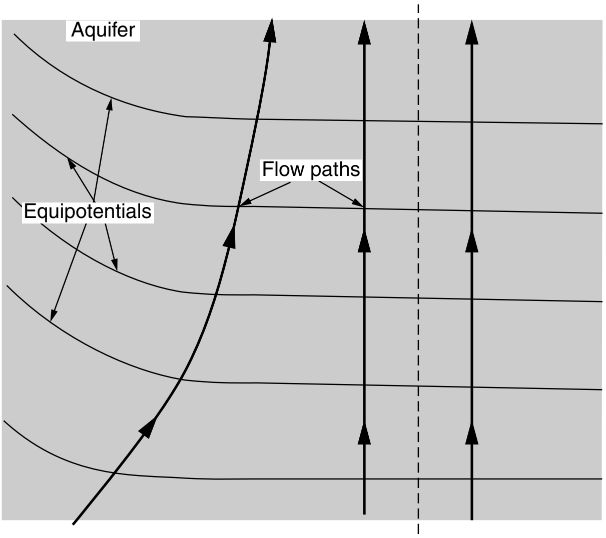

Figure 3. Parallel flow lines create a no-flow symmetry boundary.

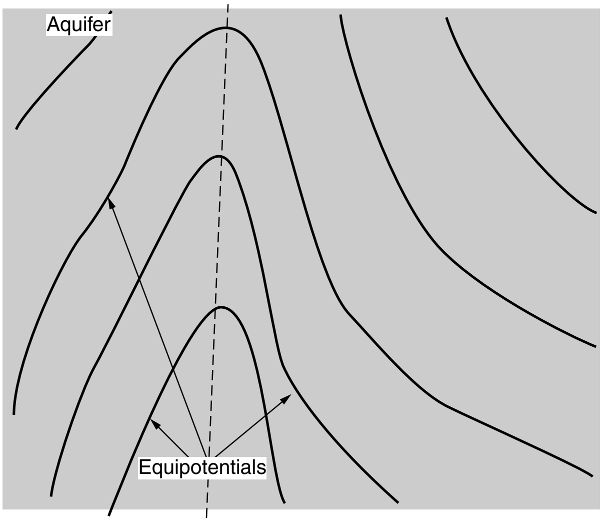

Figure 4. Groundwater divide produced by high water levels.

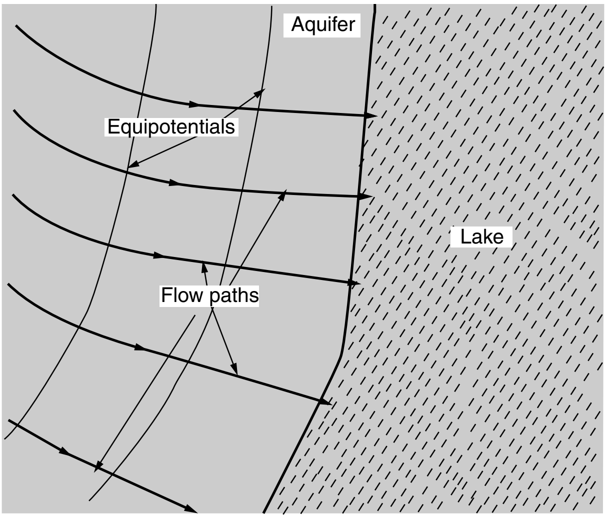

Figure 5. Constant-head boundary at a nonflowing body of water (in this case, a lake).

CHARACTERISTICS OF REGIONAL FLOW SYSTEMS

Water Chemistry

Groundwater in regional flow systems can have residence times ranging from a few decades to 10^4 years. Long residence times usually results in high concentrations of dissolved ions, and regional flow systems frequently

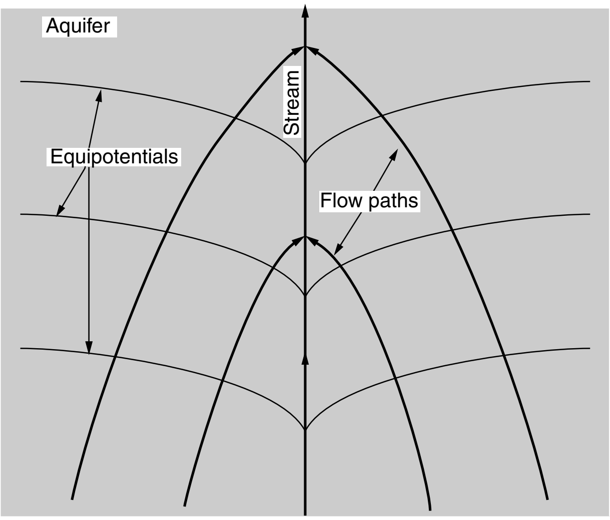

Figure 6. A stream constant-head boundary.

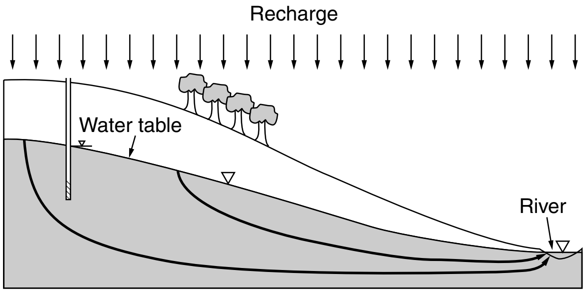

Figure 7. Cross section of water table aquifer showing the relationship between topography and the orientation of the water table.

have total dissolved solids (TDS) concentrations greater than 2000 ppm. The exact distribution of dissolved ions in the water is controlled by the minerals encountered in the subsurface; however, groundwater generally tends to evolve along flow paths toward higher concentrations of sodium and chloride, (2), and groundwater in regional flow systems tends to be dominated by NaCl type waters. Water samples are usually close to saturation with respect to common carbonate (e.g., calcite, dolomite) and sulfate (e.g., gypsum, barite) minerals. Water temperatures are generally higher than average annual air temperatures due to geothermal heating; the difference in temperature is directly related to residence time.

Spring Discharge

Springs in regional flow systems tend to exhibit relatively constant discharge, chemistry, turbidity, and temperature. These springs frequently have both local and regional flow components. These springs respond to local storms with increases in discharge and turbidity and decreases in TDS and temperature but also have a fairly constant baseflow discharge sourced from the regional systems (3). Ecosystems in springs sourced by regional flow systems often have restricted endemic organisms in or near their outlets, and these spring ecosystems often support listed endangered species.

Relationship Between Topography and Flow Systems

King (4) and Hubbert (5) noted the relationship between topography and groundwater flow patterns in unconfined aquifers. Both noticed that the water table tends to be a subdued replica of the topography (Fig. 7), and Hubbert (5) suggested that topography can control flow patterns so that high elevations are recharge areas and low elevations are discharge areas (Fig. 8).

Toth (6) and Freeze and Witherspoon (7) developed computer models that simulated the effects of topography on flow systems. Both of their models supported the conclusions of King and Hubbert. These models also indicated that sinusoidal topography can result in the formation of smaller local flow systems with local recharge and discharge areas within larger regional systems.

BIBLIOGRAPHY

1. Sharp, J.M., Jr. (2001). *Regional Groundwater Flow Systems in Trans-Pecos Texas in Aquifers of West Texas*. R.E. Mace, W.F. Mullican III, and E.S. Angle (Eds.). Texas Water Development Board, Austin, TX, Report 356, Ch. 4, pp. 41–55.
2. Back, W. (1960). Origin of hydrochemical facies of groundwater in the Atlantic Coastal Plain. *Proc. Int. Geol. Congr.*, Copenhagen, pp. 87–95.
3. Uliana, M.M. and Sharp, J.M., Jr. (2001). Tracing regional flow paths to major springs in Trans-Pecos Texas using geochemical data and geochemical models. *Chem. Geol.* **179** (1–4): 53–72.
4. King, F.H. (1899). *Principles and Conditions of the Movements of Groundwater*. U.S. Geological Survey 19th Annual Report, pt. 2, pp. 59–294.

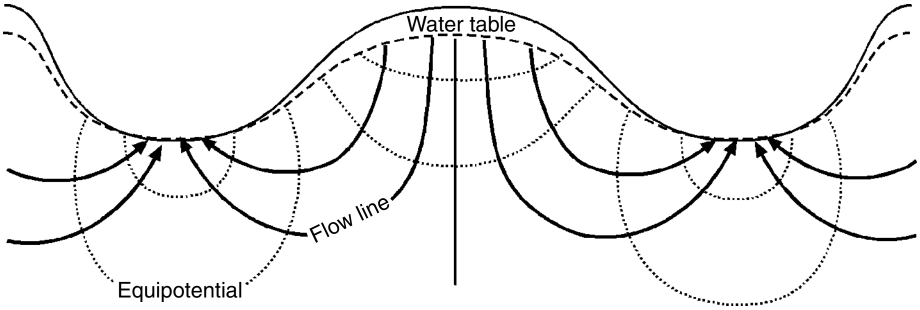

Figure 8. Flow pattern controlled by topography (after Reference 5).

5. Hubbert, M.K. (1940). The theory of groundwater motion. *J. Geol.* **48**: 785–944.
6. Toth, J. (1963). A theoretical analysis of groundwater flow in small drainage basins. *J. Geophys. Res.* **67**: 4375–4387.
7. Freeze, R.A. and Witherspoon, P.A. (1967). Theoretical analysis of regional groundwater flow. II: Effect of water table configuration and subsurface permeability variations. *Water Resour. Res.* **2**: 641–656.

GROUNDWATER REMEDIATION BY INJECTION AND PROBLEM PREVENTION

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

A variety of interesting problems can occur in the subsurface as the result of efforts to remediate groundwater. With few exceptions, groundwater remediation requires either an *in situ* injection of reactive chemicals or groundwater recovery, treatment, and disposal. The reinjection of treated water or the placement of chemicals in the subsurface used to treat contaminants can cause unintended subsurface geochemical conditions that may create new contamination issues or impact *in situ* transport properties.

Storm drains, sanitary sewers, and surface streams have been and will continue to be receptors for treated water. However, permit requirements and volume based user fees often contribute significantly to the cost of groundwater remediation systems.

In rural settings, storm or sanitary drains may not be available, and increasingly in urban settings publicly owned treatment works (POTWs) are reluctant (or unable) to accept recovered groundwater. In urban settings, water quality is not the issue; rather, it is stress on the hydraulic capacity of the POTW system. In these instances, the most viable option for water discharge is reinjection.

REINJECTION ISSUES

Treated water may meet the standards for reinjection based on the chemicals analyzed. It is becoming more apparent, however, that there could be contaminants in treated water that are not analyzed. Methyl tertiary butyl ether (MTBE) is a fuel oxygenate used in gasoline. In some pump and treat systems, the MTBE is removed and the MTBE is below the laboratory reporting level. If tertiarty butyl ether (TBA), a breakdown product of MTBE, is not analyzed, it could be present in the treated water and discharged to a stream, storm drain, or into a reinjection well, spreading an otherwise unanalyzed contaminant into the environment. Overlapping plumes of gasoline and undetected perchlorate could also create a new problem if the pumped water is treated for gasoline and then discharged to the storm drain or into the subsurface, with the undetected perchlorate. Lastly, another example is 1,4-dioxane, an extremely soluble hydrocarbon similar in transport properties to MTBE that was used as a stabilizer in percent concentrations for some chlorinated solvents.

Reinjection can be less expensive (at least to set up) than permit fees and flow based discharge fees. As such, it is sometimes implemented even when other options are available. Lastly, an injection regime may be part of a program to obtain hydraulic control of a site.

The most common source of problems for groundwater recharge systems is plugging. Any recharge system will eventually foul, and the design of the system should provide for that eventuality. The goal is to minimize the rate at which the fouling occurs. The purpose here is to examine some of the pitfalls that can occur when the recharge option is selected. Without exception, it is much easier to prevent damage from occurring than it is to remedy a damaged condition.

SUSPENDED SOLIDS

Suspended solids are a major cause of plugging. In most instances, solids will collect within the first 0.5 in. of a well pack. At a flow rate of 10 GPM and total suspended solids (TSS) at 5 mg/L, approximately 0.66 lb of solids will be introduced per day. In a year, this would total 240 lb, or about 2 ft^3 of solids, which would be capable of fouling about 100 ft^2 of injection face (30 linear feet of a 1-ft diameter well). Generally, infiltrated water must have a TSS level <2 mg/L in order to maintain adequate infiltration rates over a long period.

AIR

Entrained air can impact recharge capacity through simple physical blockage or through what is termed the “Jamin effect.” The Jamin effect results when capillary forces act on small advective channels that contain alternating air bubbles and water. These channels are then capable of responding to finite pressure gradients without allowing fluid flow, meaning that surging an impacted well will not displace the entrained air. Once in a formation, air can be extremely difficult to remove.

The problem is particularly exacerbated by allowing injected water to fall into a well (rather than piping below the groundwater surface), thus mixing with air that can be transported into the adjacent formation. Saturated zone *in situ* bioremediation systems using injected air or hydrogen peroxide can also be susceptible to air entrainment problems.

Injection of water that is cooler than the receiving aquifer will cause degassing as the two waters mix. The solubility limit of oxygen in pure water at 40 °F is 13.1 mg/L; at 50 °F it is 11.3 mg/L. Over a 24-h period, at a flow rate of 10 GPM, approximately 0.20 lb of oxygen would degas upon contact with the formation water, which represents about 2.5 ft^3 of gas entrained in the pore spaces directly adjacent to the injection zone (as depth increases the gas volume would decrease). Assuming a 1-ft diameter well, and that entrainment of air through a 1-ft section of the formation would decrease injection efficiency through the Jamin effect, a well with a 30-ft screened interval would have flow inhibited within 10 d.

MICROBIAL FOULING

To prevent microbial growth, care must be taken to ensure there is no source of carbon or nutrients (nitrogen and phosphorus) in the injected water. The carbon content of water should be evaluated through a total organic carbon (TOC) analysis, not just contaminants of concern (such as BTEX). Bacteria will exploit any carbon resource, not just those under regulatory discharge limits. TOC must be <10 mg/L to ensure that reasonable infiltration rates can be maintained over cost effective durations.

In instances where the groundwater is iron or sulfate rich, iron or sulfate reducing bacteria may also create problems.

CHEMICAL INCOMPATIBILITY

Chemical reactions between the injected and formation water are common. Typical reactions result in the formation of precipitants in the mixing zone. These reactions are complex and temperature sensitive. Differences in redox potential (usually due to dissolved oxygen levels) and carbonate chemistry are most often responsible for problems.

When injecting into clay-rich soils, ionic reactions can also come into effect. Clay particles can be dispersed (to cause plugging deeper in the formation) or swell when exposed to recharge water with ionic character different from the formation water. These reactions can be induced through exposure to water of different ionic concentration or through exposure to water with different types of cations present at any concentration (e.g., water discharged from a caustic or lime precipitation system).

Iron precipitation can be a serious problem in groundwater systems that are high in soluble ferrous iron. Iron present in recovered water may precipitate forming TSS that must be removed before reinjection. However, iron precipitation is also possible in the mix zone if the redox conditions of the recharge water are significantly different from the formation water.

OVERDEVELOPMENT

Lastly, injection efficiency has been observed to be reduced by as much as 50% from compaction of the gravel/sand pack around a well through overzealous surging during well development, pumping tests, or redevelopment performed to alleviate some of the problems described above, which is a common problem. In this instance, more is *not* better; well development should be judiciously applied based on the composition of the pack around the casing.

Impact from all of the above can be minimized in an injection system during the design stage of a project. It requires a thorough evaluation of the chemistry of the discharge that includes determination of ionic character, redox condition, chemical composition, and temperature. The same must be done for the formation water. Potential reactions and problems should be evaluated by a competent geochemist. It may be determined that a different method of water treatment would be more cost effective (i.e., eliminating air stripping to prevent precipitation reactions driven by elevated dissolved oxygen).

EH/PH CHANGES

A variety of *in situ* injection processes change Eh/pH conditions. Chemical oxidants such as ozone (O_3) and hydrogen peroxide (H_2O_2) release oxygen as part of the reaction. Enhanced bioremediation using oxygen (O_2) diffusion, air sparging, or magnesium peroxide (MgO_2) also are designed to take an anaerobic reducing environment and convert it into an aerobic environment. For example, a petroleum hydrocarbon being treated using ozone in the anaerobic core of a plume would change Eh/pH conditions in the treatment area.

These changes have the effect of changing the valence states of certain metals, such as chromium. If high concentrations of chromium are present naturally or from a release, and there are manganese containing minerals present in the mineral matrix, the formerly reduced chromium compound can be oxidized and converted from chromium (3+) [Cr(III)] to chromium (6+) [Cr(VI)], a toxic and mobile form of chromium. Normally, the oxidized chromium would move with the groundwater downgradient to naturally reducing conditions, which should restore the original Eh/pH. However, when a water supply well or vulnerable surface water is located close to the *in situ* treatment area, there may not be enough of a distance and retention time for the reconversion of Cr(VI) back to Cr(III). One option is to create an Eh/pH restoration treatment wall immediately downgradient from the initial treatment zone. The Eh/pH restoration wall should be constructed to place the liquids (molasses, corn syrup), gases (propane), or solids (nanoscale iron particles) into the aquifer through trenches, wells, ports, or other delivery means that could be supplied on a constant basis to restore the Eh/pH of the environment. After all the petroleum hydrocarbon treatment is performed in the center of the hydrocarbon plume treatment area, the Eh/pH restoration wall could be shut down after the last of the remaining oxidized water has passed the Eh/pH restoration wall. Changing Eh/pH requires redox reaction kinetics that may favor certain reactions. For example, it is much easier to convert Cr(VI) to Cr(III) than the reverse.

An oxidizing Eh/pH treatment wall would be required if originally oxidizing conditions were present and reducing chemicals (carbon sources) were used for anaerobic treatments of chlorinated solvents such as perchloroethene (PCE) or trichloroethene (TCE), for example. In this case the reduced conditions might convert the more stable and less toxic arsenic (5+) to the toxic arsenic (3+) that is soluble under normal pH conditions. A further exacerbation of potential arsenic problems occurs under reducing conditions when some portion of the ferric iron in the mineral matrix is converted to ferrous iron. The ferrous iron is soluble and the process will release any arsenic that had previously adsorbed to the iron oxides, which act as a major redardation sink for arsenic in groundwater systems. An oxidizing Eh/pH treatment wall in an area of high arsenic would supply oxygen to restore arsenic to As(V) and oxidize ferrous iron to the arsenic removing ferric oxide species, lowering the potential of creating a new subsurface problem.

Proper bench testing, chemical evaluations, and rigorous planning to prevent future treatment problems is less costly and time consuming than encountering the mistakes after developing a poorly designed project.

GROUNDWATER REMEDIATION: *IN SITU* PASSIVE METHODS

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Passive *in situ* remediation technologies (PIRT) use a variety of physical, chemical, or biological processes that act without significant human intervention in the subsurface to reduce the mass, toxicity, mobility, volume, or concentration of contaminants in soil or groundwater. Passive remediation uses the natural movement of water to disperse treatment chemicals or to move contaminated groundwater through a treatment zone.

In many early EPA Superfund site remediations, slurry walls were used. Slurry cutoff walls are a simple passive treatment method to isolate contaminants or prevent migration. Frequently, a pump and treat system is used with a slurry cutoff wall to to maintain groundwater pressure on the wall and to prevent groundwater from rising above and around the wall. Slurry walls are generally constructed of bentonite, polyethylene, or sheet piling. Project depths rarely exceed 15 meters.

PERMEABLE REACTIVE BARRIERS

Early examples of passive remediation are permeable reactive barriers (PRB) (Fig. 1). Within these passive treatment zones are included funnel and gate treatment technologies. The funnel and gate system is a hybrid of both low permeability barriers and a permeable treatment system. Originally proposed by the University of Waterloo, it is most widely used with zero-valent iron to degrade chlorinated volatile organic compounds (VOCs) such as trichloroethylene (TCE) and daughter products. These types of treatment systems are generally less than 15 meters deep.

Deeper treatment methods using drilling rigs and directional high-pressure fracturing technologies developed for depths exceeding 100 meters to inject treatment chemicals have been pioneered by GeoSierra. These deeper PRBs are generally used with zero-valent iron as the treatment material, which is injected under high pressure to form a fracture wall between closely spaced deep boreholes. The passive treatment wall is designed to be perpendicular to groundwater flow.

BIOREMEDIATION

Phytoremediation, another passive method, uses specific plants and trees to extract specific contaminants from soil and groundwater. Some of the targeted contaminants include heavy metals and radionuclides. Common plants used in phytoremediation include mustard plants as well as poplar trees. Phytoremediation has great promise and will increase in acceptance as more case studies are completed. An overview article documents some of the potential of phytoremediation (2).

For passive and semipassive methods of remediation, aerobic and anaerobic bioremediation offer options for a variety of contaminants. For low permeability sediments (clay and silts), the molecules of the dissolved gases (oxygen, propane, and others) and the microbes are much smaller than the pore throat sizes of the sediments, which allows enhanced bioremediation to be considered in geologic settings where other more active technologies such as dual-phase extraction would not be feasible.

Aerobic *in situ* enhanced bioremediation is a cost-effective way to treat residual contamination in groundwater from surface spills or leaking underground tanks of hydrocarbon fuels such as gasoline, diesel, kerosene, motor oil, jet fuels, and used oil. Hydrocarbon degrading microbes will break hydrocarbons down to carbon dioxide and water if the right amounts of oxygen and nutrients are available.

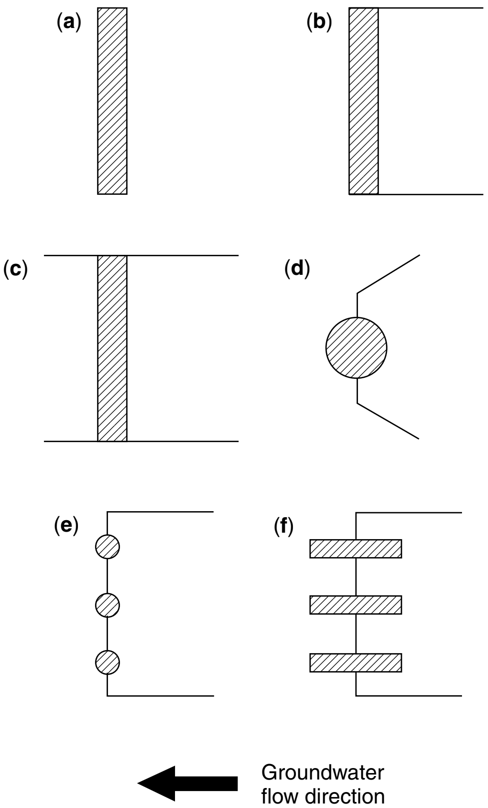

Figure 1. Possible geometries of passive treatment walls: (**a**) continuous trench, (**b**) with lateral hydraulic control, (**c**) with lateral hydraulic control and downgradient lateral barriers, (**d**) with a single treatment caisson, (**e**) with multiple treatment caissons, (**f**) with parallel alleys (1).

Typically, the aerobic microbial degradation of hydrocarbons is limited by the amount of oxygen present. Oxygen can be added in a variety of ways: from peroxygens such as magnesium peroxide, sodium peroxide, calcium peroxide, potassium permanganate, and hydrogen peroxide (H_2O_2). Oxygen sources as gases include oxygen (O_2) and ozone (O_3). Higher concentrations of oxidizers such as hydrogen peroxide and ozone would temporarily destroy microbial colonies, creating a barren zone that might last for a few weeks. Ultimately, the aerobic microbial colonies would move back into barren zones from adjacent areas and reestablish themselves with the benefit of significantly higher oxygen levels.

Anaerobic *in situ* enhanced bioremediation is used for chlorinated solvents such as tetrachloroethylene (PCE) and trichloroethylene (TCE). The same process can be used to reduce heavy metals, nitrates, and perchlorate. Nitrates are generated by agricultural processes and runoff from fertilizers. Perchlorate is an ingredient in rocket fuels and flares. For anaerobic processes, remediation requires carbon sources. Molasses, cheese whey, corn syrup, milk, glucose, as well as hydrogen, methane, and propane have been used. Reductive bioprocesses have been well documented (3).

REMEDIAL CRITERIA FOR DEVELOPING NATIONS

One of the main challenges in assessing and remediating petroleum hydrocarbon contaminated sites is developing an appropriate cost-effective remediation strategy, especially for nonindustrialized countries. For developing nations, technologies would have to match local geologic conditions, use readily available materials and resources, be performed *in situ*, require a minimal amount of equipment and little or no power, have no time constraints, and meet local and national regulatory guidelines. In addition, the technologies would require training that would allow local contractors to perform the work proficiently.

ENHANCED BIOREMEDIATION

Natural attenuation consists of unenhanced physical, chemical and biological processes that limit the migration and reduce the concentration of contaminants in the subsurface (4). The most important process in petroleum hydrocarbon degradation is aerobic bioremediation because it can destroy a large percentage of the hydrocarbon contaminant mass. Destruction occurs because of aerobic bacteria that use oxygen to oxidize hydrocarbons (reduced materials) to obtain energy. Aerobic biological degradation of hydrocarbon fuels and selected other organic compounds have been well documented (3,5–7). Enhanced bioremediation is the process of supplying an appropriate amount of oxygen, nutrients, and other amendments, if needed, to the subsurface to enhance the biodegradation process, significantly increase remediation effectiveness, and decrease treatment time.

Gasoline is a mixture of numerous compounds, including benzene, toluene, ethylbenzene, and xylenes (BTEX). Gasoline sometimes also contains additives, such as methyl tertiary-butyl ether (MTBE). Contaminant concentrations below saturation are degradable. Gasoline floating on the top of the groundwater table, also called free-product, is toxic to hydrocarbon degrading microbes.

Gasoline, MTBE, a gasoline additive, and BTEX compounds degrade aerobically in the presence of dissolved oxygen. Natural attenuation of gasoline constituents has been growing as a remediation technology since the mid 1990s. Presently, a variety of technologies are available that create low to high concentrations (10–60 ppm) of stable dissolved oxygen in groundwater. Once this elevated dissolved oxygen mixes with gasoline-contaminated ground water, natural biodegradation processes occur (3,8).

BENCH TESTS AND OXYGEN DEMAND THRESHOLD

Bioremediation enhanced by the addition of dissolved oxygen has been proven an effective technology to reduce both BTEX and MTBE. However, many groundwater environments contain numerous materials that cause competing reactions for the oxygen from solid and dissolved compounds high in ferrous iron and other metals in the reduced state, and organic and inorganic compounds that have high biological and chemical oxygen demand. These compounds, also called “oxygen sinks,” consume large volumes of injected dissolved oxygen before aerobic bacteria can use the oxygen as part of the process of consuming gasoline and associated compounds.

The oxygen demand parameters significantly affect both enhanced bioremediation and chemical oxidation technologies. Oxygen demand includes the solid and dissolved substances that compete with microbes for oxygen. The parameters include solid oxygen demand (SOD) in saturated soil and biological oxygen demand (BOD) and chemical oxygen demand (COD) in groundwater. By better understanding nonmicrobial oxygen demand, the amount of oxygen required by microbes at a specific site that has specific geologic conditions during enhanced bioremediation can be predicted. Laboratory work generally includes measuring specific variables in the soil and groundwater such as alkalinity, pH, dissolved oxygen, contaminant concentration, macronutrients, heterotrophic plate count, specific degraders, chemical oxygen demand, biological oxygen demand sulfate, nitrate, total iron, ferrous iron, total dissolved solids, total inorganic carbon, and total organic carbon. The oxygen demand threshold must be overcome before significant petroleum hydrocarbon degradation will occur under aerobic conditions.

OXYGEN DELIVERY

For aerobic bioremediation, an efficient oxygen mass transfer method is needed to place high levels of dissolved oxygen in the groundwater. There are many methods of delivering oxygen to groundwater. A spargeless delivery of supersaturated levels of dissolved oxygen into groundwater is one method to ensure an abundance of oxygen to get over the oxygen demand threshold quickly. Use of chemical oxidants such as hydrogen peroxide or ozone can be used to reduce oxygen-competing reactions rapidly.

To deliver dissolved oxygen into the aquifer, a spargeless gas emitter tool that infuses 100% oxygen into groundwater via 50.8-mm diameter (2-inch) monitor wells has been developed (8) by inVentures Technologies, Inc (Fig. 2). There are no power sources or moving parts. The oxygen is released by pressure from an industrial-grade oxygen cylinder. The proprietary structured polymer inside the oxygen diffusion tool provides a large surface area for gas transfer into the gas infusion tool, which is placed into an existing monitor well. The system can be reinstalled and reused on other sites.

The gas infusion tool is connected to a regulated supply of industrial compressed oxygen (100% concentration). Field experience has shown that dissolved oxygen levels of 30–60 mg/L have been achieved in each monitor well where a gas infusion tool is installed. The dissolved oxygen is infused into the monitoring well at a typical rate of 5 to 20 cc/minute, depending on the head pressure in the well. The effective radius of influence is typically 3 to 5 meters, even in low permeability soil such as clay (Mapleshade, New Jersey, U.S. site) or in fractured bedrock (Englewood, Colorado, U.S. site).

Field performance data for a pilot test site in Mapleshade, New Jersey, with MtBE, tert-butyl alcohol (TBA), and BTEX bioremediation illustrates the gas infusion technology (8). Depth to water is 2 to 4 meters below grade. Groundwater velocity is 0.09 to 0.12 meters/day at a gradient of 0.028. Recovery wells yield up to 4 liters per minute. The existing on-site air stripper-GAC system had iron clogging problems from ferrous iron at 30 to 80 mg/L. Dissolved oxygen (DO) concentrations in deep piezometers have increased from low baseline levels to more than 1 mg/L at a minimum of 5 meters downgradient of the infusion point. Shallow piezometers showed little DO change from baseline.

After 6 months of using gas infusion devices at the Mapleshade, New Jersey, U.S. site, an effective barrier of dissolved oxygen was established. Data indicate significant reductions in MtBE (up to 99%), TBA (50%), and benzene (85%) concentrations downgradient of the gas infusion tools.

THE IMPORTANCE OF ADEQUATE HYDROGEOLOGIC CHARACTERIZATION FOR *IN SITU* REMEDIATION

Understanding the subsurface environment is essential to developing a practical, efficient, cost-effective and timely *in situ* remediation strategy. System design is optimized when important geologic factors such as lithology, permeability, porosity, soil and groundwater chemistry, and contaminant type and concentration and their anticipated fate and transport, are fully evaluated with respect to the design and installation of oxygen injection ports or wells and rates of injection. Geologic factors control the movement, distribution, and quality of

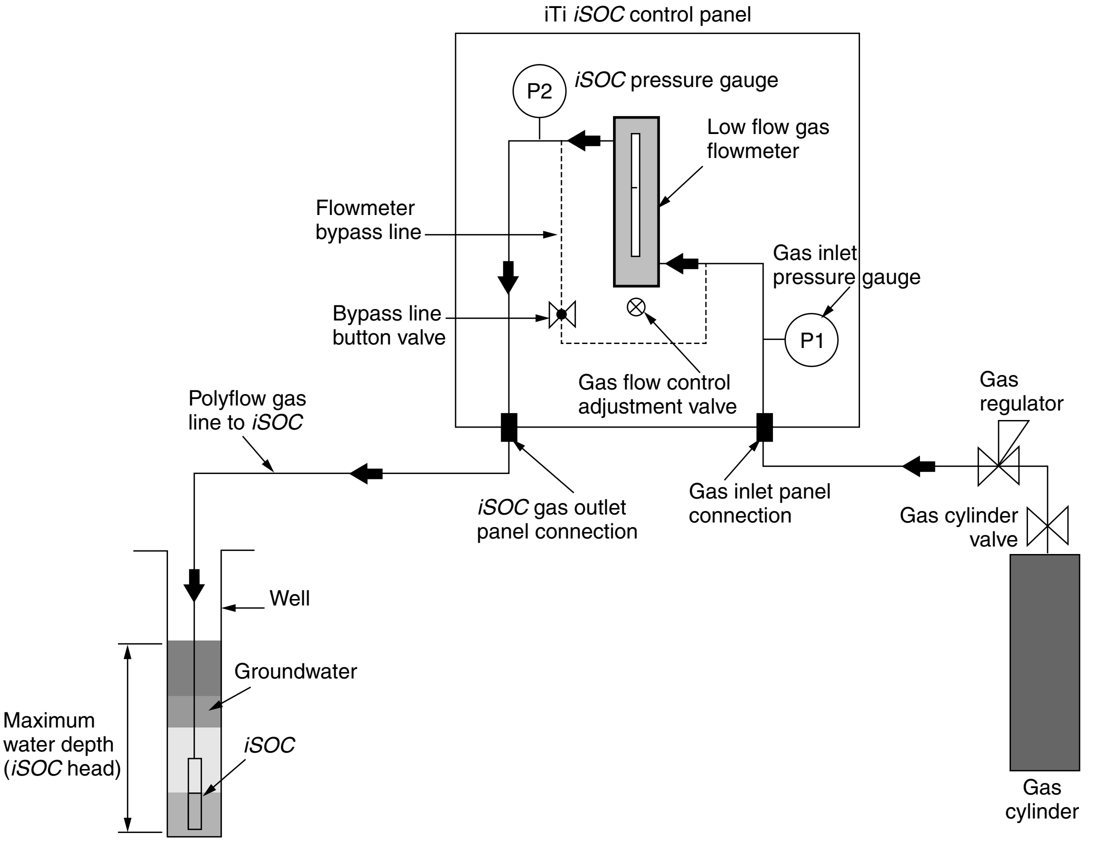

Figure 2. Diagram of semipassive gas diffusion device (iSOC).

groundwater as well as contaminants via a combination of several physical, chemical, and biological processes (9,10).

Shallow groundwater contamination by petroleum hydrocarbons results from surface or near-surface activities, including releases from underground and aboveground storage tanks and pipelines and accidental spills. The largest percentage of environmentally contaminated sites in the world lies on alluvial and coastal plains consisting of complex interstratified sediments. Lithologic contacts, some abrupt and some gradational, significantly influence permeability, porosity, and preferred flow pathways of contaminants in soil and groundwater.

Geology is the prime controlling agent for the movement of groundwater contaminants. A comprehensive understanding of the three-dimensional framework of geologic materials provides data for study of groundwater vulnerability to contamination and the lateral and vertical extent of hazardous and toxic constituents in the subsurface. Effective design and monitoring of subsurface remedial systems requires hydrogeologic characterization.

Lithology and stratigraphy are the most important factors affecting contaminant movement in soils and unconsolidated sediments. Stratigraphic features, including geometry and age relations between lenses, beds, and formations and the lithologic characteristics of sedimentary rocks, such as physical composition, soil type and chemistry, grain size, grain packing, and cementation, are among the most important factors affecting groundwater and contaminant flow in sedimentary rocks. Igneous–metamorphic rocks are geologic systems produced by deformation after deposition or crystallization. Groundwater and contaminant flow in igneous–metamorphic rocks are most affected by structural features such as cleavages, fractures, folds, and faults.

In summary, passive and semipassive remedial methods use minimal or no power, have few or no movin g parts, have few or no time constraints, and meet local and national regulatory guidelines. As for all forms of *in situ* remediation, proper assessment and hydrogeological characterization are the keys to success.

BIBLIOGRAPHY

1. Warner, S.D., Szerdy, F.S., and Yamane, C.L. (1998). *Permeable Reactive Treatment Zones: A Technology Update, Contaminated Soils*. Vol. 3. Amherst Scientific, Amherst, MA, Chap. 27, pp. 315–327.
2. Lehr, J. (2000). Phytoremediation: A promising path to the elimination of heavy metals and radionuclides in groundwater and soil. In: *Standard Handbook of Environmental Science, Health, and Technology*. J. Lehr (Ed.). McGraw-Hill, New York, pp. 14.26–14.29.
3. Suthersan, S.S. (2002). *Natural and Enhanced Remediation Systems*. CRC Press/Lewis, Boca Raton, FL.
4. Testa, S.M. (1996). *Groundwater restoration and site closure: Society of Petroleum Engineers*. Paper 35885, pp. 1–11.
5. Rice, D. et al. (1995). *Recommendations to improve the cleanup process for California's leaking underground fuel tanks*. Submitted to the California State Water Resources Control Board and the Senate Bill 1764 Leaking Fuel Tank Advisory Committee, Lawrence Livermore National Laboratory Report, Livermore, CA.
6. Mace, R.E., Fisher, R.S., Welch, D.M., and Parra, R.A. (1997). *Extent of Mass and Duration of Hydrocarbon Plumes from Leaking Petroleum Storage Sites in Texas*. Texas Bureau of Economic Geology, Geological Circular 97-1.
7. U.S. EPA. (1985). *Modeling Remedial Actions at Uncontrolled Hazardous Waste Sites*. EPA/540-2-85/001, Washington, DC.
8. Kolhatkar, R., Kuder, T., Philp, P., and Wilson, J.T. (2002). Stable Carbon Isotope Analyses to Demonstrate MTBE Biodegradation in Ground Water. *Groundwater Resour. Assoc. (GRA) MTBE Symp.*, October 17, 2002, San Jose, CA.
9. Testa, S.M. (1994). *Geological Aspects of Hazardous Waste Management*. CRC Press/Lewis, Boca Raton, FL.
10. Testa, S.M., and Winegardner, D.L. (2000). *Restoration of Contaminated Aquifers*, 2nd Edn. CRC Press/Lewis, Boca Raton, FL.

READING LIST

Jacobs, J.A. (2000). Passive *in situ* remediation technologies. In: *Handbook of Environmental Science, Health, and Technology*. J. Lehr (Ed.). McGraw-Hill, New York, pp. 14.12–14.25.

Jacobs, J. and von Wedel, R. (1997). Enhanced *in situ* biodegradation via high pressure injection of oxygenating agents and nutrients. *Am. Inst. Prof. Geol., 34th Annu. Meet.*, Houston, TX, Abstracts, p. 27.

GROUNDWATER REMEDIATION BY *IN SITU* AERATION AND VOLATILIZATION

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

INTRODUCTION

There are a variety of groundwater remediation processes that use *in situ* aeration and volatilization as a part of their technology. Some of these technologies have a bioremediation component to them as well. These technologies include soil venting, soil vapor extraction, air sparging, bioventing, and biosparging.

Soil venting is a term used for an *in situ* aeration process that is a powerful remediation technology for the treatment of soils exposed to a variety of hydrocarbons. The practice of soil venting includes the following variations in application:

- Soil vapor extraction (SVE) systems are designed to exploit hydrocarbons potential for volatilization.
- Air sparging is an aeration process in which volatilizing air is injected into the saturated zone beneath the water table. Soil venting is used to recover the vapor laden air as it exits the water table.

- Bioventing is an aeration process designed to deliver oxygen to the subsurface for use by indigenous bacteria to degrade hydrocarbons; the focus is on minimizing hydrocarbon volatilization.
- Biosparging is a variant of air sparging in which oxygen stimulated biodegradation is the aim rather than volatilization. As with air sparging, soil venting is used to recover gas discharged through the water table.

Figure 1 illustrates and compares the potential effectiveness of soil vapor extraction and bioventing.

The data illustrated in Fig. 1 was calculated using the following premises:

- The air flow rate is 10 SCFM (Standard Cubic Feet per Minute). For the vapor transport calculations, it is assumed that the 10 SCFM air stream becomes saturated with hydrocarbon vapor. Volatilization driving vapor pressures were calculated at 8 °C.
- For the biodegradation calculations, it is assumed that the oxygen provided by the 10 SCFM air flow is completely used for hydrocarbon biooxidation.
- Biodegradation of each of the hydrocarbons proceeded to carbon dioxide and water.

This data was derived from theoretical calculations predicated on the fundamental principals governing the action of each of the processes (i.e., volatilization or biooxidation). In actuality, these processes and other subsurface interactions (with soil moisture, for example) are more complex. However, the base principals do apply and are, in the overall process, upheld. Although based on a simplified system, the data serves to illustrate the fundamental principals of, and differences between, soil vapor extraction and bioventing.

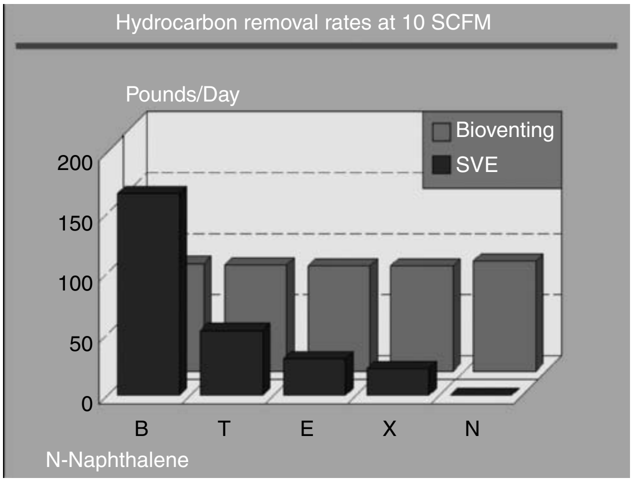

Figure 1. Hydrocarbon removal by SVE and biooxidation.

Following is a more detailed explanation of these processes and a case history that illustrates the use of biosparging technology at a site contaminated with volatile and nonvolatile hydrocarbons.

SOIL VAPOR EXTRACTION

By stimulating an advective air flow through soil venting, adsorbed hydrocarbons are exposed to air that does not carry an existing vapor load. Upon this exposure, equilibrium driven mechanisms will force the hydrocarbons to volatilize into the induced air flow. Every hydrocarbon has a specific temperature-dependent vapor pressure that determines the maximum (saturated) vapor concentration that can be obtained in the air flow. This vapor-saturated air is directed to the surface where it may be discharged to the atmosphere, or more commonly, treated to remove the hydrocarbon vapors before discharge.

The key element in this vapor extraction process is the fact that the mass transport rates are determined by a physical property (the vapor pressure) of the hydrocarbon. Not all hydrocarbons have the same vapor pressure, and the lower the vapor pressure, the lower the overall mass transport rate will be for that compound, thus the dramatic trend seen in Fig. 1 for the SVE portion of the graph. Table 1 presents the basic data used to prepare the soil vapor extraction part of Fig. 1.

Benzene is extremely volatile and offers an excellent mass transport potential of 166 pounds per day. However, naphthalene is at the other extreme; with a vapor pressure less than 1 mm Hg, it is at the lower end of hydrocarbons considered amenable to removal by soil vapor extraction. In this instance, only 0.13 pounds per day would be removed. The other petroleum hydrocarbons fall somewhere between the two.

The dependence of soil vapor extraction on the varying physical property of vapor pressure places serious constraints on the rate of SVE remediation at sites

Table 1. Vapor Pressure and Volatilization Driven Mass Transport Rate

Hydrocarbon	Vapor Pressure*	Molecular Weight	Saturated Vapor Concentration	Lbs Removed in One Day at 10 SCFM
Benzene	41.0	78.1	11.5×10^{-3} lbs/ft^3	166
Toluene	11.1	92.1	3.67×10^{-3} lbs/ft^3	52.8
Ethylbenzene	5.5	106.2	2.10×10^{-3} lbs/ft^3	30.2
Xylene	4.0	106.2	1.53×10^{-3} lbs/ft^3	22.0
Naphthalene	0.02	128.2	9.2×10^{-6} lbs/ft^3	0.13

*mm Hg at 8° Centigrade

impacted with petroleum hydrocarbons. Even gasoline has significant concentrations of hydrocarbons with relatively low vapor pressures. Heavier petroleum products, such as jet fuel, kerosene, diesel fuel, and lubricating oil, are thought to be nonresponsive to soil vapor extraction technology. Based on soil vapor extraction alone, that thought is an accurate one.

Air Sparging

Air sparging is the process of hydrocarbon volatilization stimulated by the injection of air beneath the water table. With air sparging, the volatilization process takes place under saturated conditions, emulating the action of an air stripping surface treatment system. The governing physical parameter that relates a hydrocarbon's volatilization potential from water is Henry's constant. According to Haarhoff and Cleasby (1), Henry's constant is directly related to a hydrocarbon's vapor pressure, water solubility, and temperature of the air/water/hydrocarbon system, which is a more complex physical/chemical system than volatilization of free phase or adsorbed hydrocarbons from soil in a vapor extraction system. In an air sparging system, the injected air and entrained volatilized hydrocarbons are captured above the water table with a conventional soil venting system.

$$P_a = H_a X_a$$

where P_a is the partial pressure of constituent a in air, H_a is the Henry's law constant (atm), and X_a is the solution concentration of the chemical a (mole fraction).

Some Henry's law constants for gasoline compounds include tetraethyl lead (4,700 atm), ethylbenzene (359 atm), xylenes (9266 atm), benzene (230 atm), toluene (217 atm), naphthalene (72 atm), and MTBE (27 atm) (2). Henry's law constants of greater than 100 atmospheres are generally considered volatile, and consequently more likely to be volatilized rather than biodegraded.

BIOVENTING

Bioventing is the term for aerobic biodegradation stimulated by oxygen provided to a hydrocarbon impacted subsurface zone by an air flow induced through soil venting. Many factors exist that affect the outcome and success of bioventing projects.

Intrinsic permeability (k), in cm^2/sec, and product composition are important aspects for biosparging projects. Ineffective permeability for bioventing includes k values of 10^{-16} to 10^{-12}. This range corresponds to lithologies of clay and the lower permeability ranges of glacial till. Moderate to minimal effectiveness for bioventing includes sediments with k values of 10^{-10} to 10^{-5}, including silt, loess, silty sand, and the lower permeability ranges of clean sand. The most effective permeability for bioventing includes k values of 10^{-4} to 10^{-2}, which includes the upper permeability range of clean sand and all gravels. In some generally low-permeability settings, heterogenetic prefererential pathways may exist that can offer influence beyond what would be initially expected.

Petroleum hydrocarbons are generally biodegradable, regardless of their molecular weight, as long as the indigenous microbes have an adequate supply of oxygen and nutrients. The most volatile and most soluble to least volatile and least soluble petroleum products include gasoline, kerosene, diesel, fuel oils, and lube oils. Bioventing is more effective with the more volatile components (2).

Site characteristics that must be evaluated include intrinsic permeability of the sediments, soil structure and stratification, temperature, (10 °C or higher), pH (6 to 8), microbial population density (generally greater than 1,000 colony-forming units per gram of dry soil), nutrient concentrations (ammonia as nitrogen and orthophosphate; carbon-nitrogen-phosphorous ratios are best at 100:10:1), and dissolved ferrous iron concentrations (Fe^{+2} is less than 10 mg/L). The contaminant characteristics also need to be evaluated, which include chemical structure of compounds, concentrations and toxicity, vapor pressure, product composition and boiling point, and Henry's law constant. Approximately 3 to 3.5 pounds of oxygen are needed to degrade one pound of petroleum hydrocarbons. This amount of oxygen does not take into account the oxygen demand in the sediments (sediment oxygen demand) or in dissolved in the groundwater (chemical oxygen demand with the subset of biological oxygen demand).

The power of bioventing lies in the underlying process responsible for it, the biologically mediated oxidation of hydrocarbons to carbon dioxide and water. The data calculated for the bioventing portion of Fig. 1 is based on the stoichiometric consumption of oxygen; that is, each mole of carbon present will require one mole of oxygen (O_2) to be converted to carbon dioxide (CO_2):

$$O_2 + C = CO_2$$

Each mole of hydrogen requires 1/4 mole of oxygen (as an O_2 molecule) to produce one mole of water. It can be more conveniently expressed as follows:

$$O_2 + 4\ H_2 = 2\ H_2O$$

The exact stoichiometry for each of the hydrocarbons illustrated in Fig. 1 is shown in Table 2.

A flow rate of 10 SCFM can deliver 270 pounds of oxygen to a treatment zone in a period of 24 hours. Based on the above stoichiometry, the potential biodegradation rates in pounds per day are indicated in Table 3.

This entire process is independent of any other physical property of these hydrocarbons. The prime issue with regard to *in situ* remediation by bioventing is how much

Table 2. Biooxidation Stoichiometry

Compound	Chemical Formula		Moles O_2 Required		Oxidation Products
Benzene	C_6H_6	+	7.5 O_2	=	6 CO_2 + 3 H_2O
Toluene	$C_5H_5(CH_3)$	+	9 O_2	=	7 CO_2 + 4 H_2O
Ethylbenzene	$C_6H_5(C_2H_5)$	+	10.5 O_2	=	8 CO_2 + 5 H_2O
3-Xylene	$C_6H_4(CH_3)_2$	+	10.5 O_2	=	8 CO_2 + 5 H_2O
Naphthalene	$C_{10}H_8$	+	12 O_2	=	10 CO_2 + 4 H_2O

Table 3. Biooxidation Rate

Compound	Biooxidation Potential—Lbs/Day
Benzene	88
Toluene	87
Ethylbenzene	86
3-Xylene	86
Naphthalene	90

oxygen can effectively be transported to the reaction (contaminated) zone. However, practical limits exist to the effectiveness of biodegradation and the linear effect of the stoichiometric biooxidation reaction with oxygen. The potential problem lies with compounds that are recalcitrant to biodegradation. With respect to petroleum hydrocarbons, these recalcitrant compounds are typically polynuclear aromatic (PNA) compounds having high ring counts. In many products, however, high ring count PNAs are not a significant amount of the total hydrocarbon makeup. In addition, these compounds are most often still biodegradable, but at a slower rate than that observed for the less refractory hydrocarbons.

Indigenous Bacteria

A key concern over the viability of the bioventing approach is the presence of indigenous bacteria capable of being stimulated to degrade hydrocarbons. Figure 2 illustrates soil gas data from a site impacted with petroleum hydrocarbons.

These samples were collected under static conditions, no remediation activity has taken place. This soil gas is representative of equilibrium conditions existing at the impacted site. Although volatilized hydrocarbon vapors (from the contaminant impact) are also present in the soil gas, Fig. 2 only shows the three dominant soil gases: carbon dioxide, oxygen, and methane. Under normal atmospheric conditions oxygen concentration is 21% and carbon dioxide is approximately 300 ppm (0.03%). As a result of the presence of carbonate minerals and natural organic materials, CO_2 in uncontaminated soils is typically found at concentrations from 1.5–3% (3). Methane is present in the atmosphere in trace amounts (1.5 ppm). As Fig. 2 illustrates, the soil gas concentrations at this contaminated site are significantly skewed from those levels. Carbon dioxide is elevated at 11–12%, oxygen is depressed to around 5%, and methane is elevated at 3–5%. Following is an outline of the process responsible for the generation of this soil gas blend:

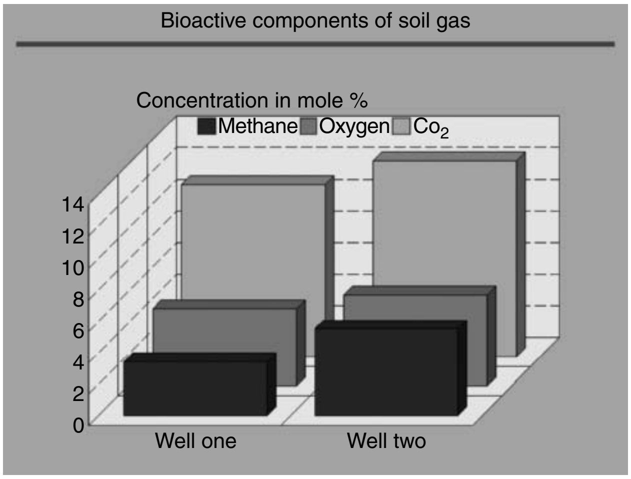

Figure 2. Bioactive soil gases in the vadose zone.

On release of hydrocarbons into the subsurface, the indigenous bacteria began an aerobically driven biooxidization process. The endproduct of this aerobic microbial degradation was carbon dioxide and water (see Table 2).

After the aerobic microbial activity had consumed oxygen in the soil gas to near the observed 5% level, facultative anaerobes became active. These bacteria have the ability to support metabolic activity under full aerobic or oxygen depressed conditions (the transition has been observed in the field and laboratories to normally occur at oxygen levels near 5%). The degradation products (seen in the gas phase) of the anaerobic activity are methane and additional carbon dioxide. The methane is a result of the fact that, in microscale isolated zones within the geologic matrix, deeply methanogenic conditions may be reached under these circumstances.

The fundamentally important point of the data illustrated in Fig. 2 is that this aerobic/anaerobic activity occurred naturally. No bacteria were added to the soil; the existing indigenous bacteria generated these gases. No nutrients were added. The indigenous bacteria were able to become active under natural subsurface conditions, using available oxygen, nitrogen, phosphorous, and trace nutrients. With depletion of the oxygen levels, the facultative anaerobic activity became dominant at the expense of the aerobic biooxidation. This anaerobic degradation occurs at a rate several orders of magnitude slower than that observed for aerobic degradation (4). If this were not the case, it would be cost effective to let the anaerobic degradation occur at its own pace with no other intervention.

What is required for timely bioremediation is the installation of a soil venting system to displace this soil gas (equilibrated to the existing chemical and microbiological conditions) with fresh, fully oxygenated air, thus restimulating the natural aerobic biodegradation of the impacting hydrocarbons, which is the essence of the bioventing process.

Biosparging

The governing processes engaged in a biosparging system are identical to those in bioventing. The biosparging stimulated biooxidation follows the same stoichiometry presented in Table 2. The injected air, biooxidation products (CO_2), and some fraction of volatilized hydrocarbons are collected above the water table with a soil venting system, just as described for air sparging. The last sentence contained a key statement, that “some fraction of volatilize hydrocarbons” are generated by a biosparging (and bioventing) system. The next section analyzes this statement in more detail.

SOIL VAPOR EXTRACTION VERSUS BIOVENTING A SIMULTANEOUS PROCESS

When a soil venting system is operated at a site impacted by petroleum hydrocarbons, both of the discussed mechanisms are engaged. The determination of which process is dominant lies in how the soil venting system is operated.

A soil venting system can be operated such that 60–90% of the impacting hydrocarbons are volatilized and 10–40% are biodegraded, i.e., soil vapor extraction. Bioventing reverses those numbers, 60–90% of the hydrocarbons are biodegraded and 10–40% are volatilized (5). Note that some degree of biodegradation is unavoidable even in a system that is designed and operated as a pure soil vapor extraction system. The site soils would have to be sterilized to prevent it. Conversely, some level of volatilization is equally unavoidable in a bioventing system.

The value of the bioventing approach lies in three areas:

First, every pound of hydrocarbon that is degraded *in situ* is a pound that will not require subsequent treatment on the surface, resulting in significant cost savings.

Second, equipment costs and operational costs (primarily energy) are less using the lower air flow rates capable of supporting the bioventing approach. Oxygen does not need to be brought to the subsurface at a rate faster than the ability of the bacteria to consume it.

Third, as Fig. 1 illustrates, soil vapor extraction does not work well with hydrocarbons that have low vapor pressures.

Remediation by soil vapor extraction is very rapid when applied to hydrocarbons with high vapor pressures. For example, a release of pure benzene could theoretically be cleaned up faster using a straight soil vapor extraction approach rather than bioventing. Although surface treatment expenses must still be taken into account in the overall project costs.

Soil vapor extraction is definitely the preferred approach when remediating biological recalcitrant compounds that have high vapor pressures such as chlorinated solvents. However, in instances where the released materials are petroleum hydrocarbons, such as fuels or lubricants, bioventing is likely to be the most cost and time effective remediation approach.

BIOSPARGING—A CASE HISTORY

Following is a field example of the biosparging process discussed above. This project is ongoing at a facility that was closing an RCRA hazardous waste drum storage area (DSA). The wastes stored in the DSA were “Spent Non-Halogenated Solvents” (F003/F005). Figure 3 shows a plan view of the site, the DSA, and the installed remediation system.

The geology beneath the site consists of:

- approximately 30 feet of sand interbedded with a peat layer 1 to 3 feet thick at a depth of 16 feet; silt and clay underlies the sand.
- depth to groundwater is 6 feet below grade (see Fig. 4).
- soil and groundwater contamination was limited to the sand/peat unit above the clay.

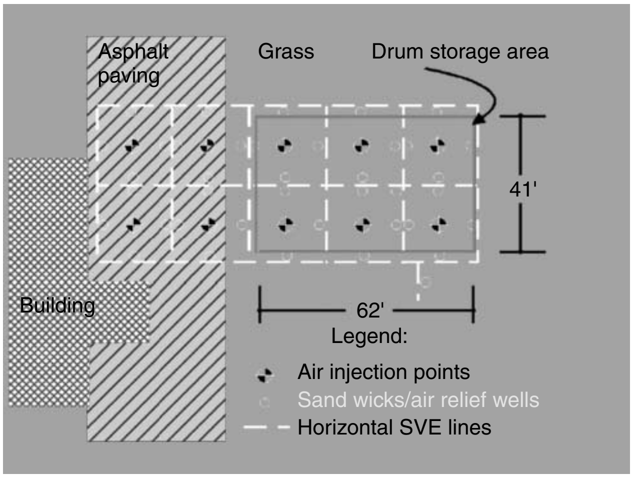

Figure 3. Plan view of biosparging site.

Soil contamination included ethylbenzene (160 mg/Kg), toluene (110 mg/Kg), total xylenes (620 mg/Kg), naphthalene (440 mg/Kg), and other polycyclic aromatics at low mg/Kg levels.

Groundwater contained 2 mg/L ethylbenzene and 16 mg/L total xylenes.

The remediation system put into place was an *in situ* saturated zone treatment using biosparging, the details (Fig. 4) of which are as follows:

- A series of sparge points were installed to a depth of 30 feet just above the surface of the lower clay layer, which placed them about 14 feet beneath the peat layer.
- A series of 4-inch borings, filled with graded sand to act as air relief wells (sand wicks), were installed to a depth of 30 feet. The sand wicks act as a conduit for injected air bubbles through the peat

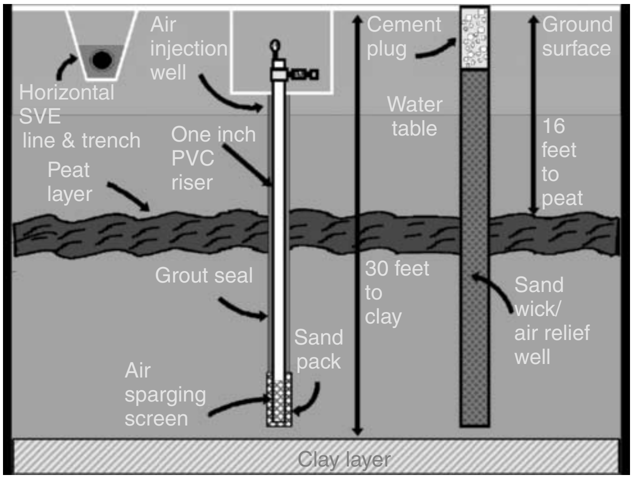

Figure 4. Cross-sectional view of biosparging site with sand wicks.

layer, preventing unacceptable horizontal migration of the injected air and any entrained vapor, which is a common and critical problem for the application of air sparging/biosparging in heterogeneous soils.

- A soil venting recovery system was installed in horizontal trenches above the water table to capture the injected air and hydrocarbon vapors generated from the biosparging system. The area was then covered with a plastic liner and soil to prevent short circuits in the soil venting system.

The discharge of the soil venting system was treated with vapor phase activated carbon.

An important point with regards to this system was the biosparging aspect of the design. Although vapor phase activated carbon was installed as a necessary treatment of coproduced hydrocarbon vapors, the intent of the remedial design was to minimize the actual volatilization and concentrate on oxygen-stimulated biodegradation of the hydrocarbons *in situ*. This approach minimized the carbon usage and also had the potential to remove the nonvolatile constituents of the impacting hydrocarbons. To this end, air injection was only at 12 SCFM, a rate estimated to match the kinetics of the microbiological biooxidation. The horizontal soil venting system was operated at 60 SCFM to ensure that all the injected air (and minimal hydrocarbon vapor load) were adequately recovered.

Although provisions were made for the addition of nutrients (nitrogen, phosphorus, and trace minerals), nutrients were not actually used in this phase of the project. Under many conditions, existing nutrient levels are often adequate to support microbial activity at an acceptable level. Under most circumstances, oxygen supplementation is the dominant factor required for microbial stimulation. Nutrient addition to the vadose and saturated zones (although possible) is a complication that should be avoided unless required. Figures 5 and 6 illustrate the results after system startup. Figure 5 illustrates the trend of the vaporized hydrocarbon and carbon dioxide observed in the recovered soil gas, and Fig. 6 illustrates the trend of dissolved oxygen (DO) in the groundwater and carbon dioxide in the soil gas.

The hydrocarbon versus carbon dioxide in soil gas above the biosparging system shows a sharp increase in the concentration in the VOCs recovered in the soil gas by day ten of system operation (Fig. 5). Initial carbon dioxide levels are quite low. In a manner very typical of microbiological systems, carbon dioxide concentrations steadily increased to a maximum level over a 30-day period.

This period is termed the "lag phase," as the indigenous microorganisms adapt to the sudden introduction of oxygen into their environment and the use of the impacting hydrocarbons as a carbon source.

In turn, the VOC concentrations continue a steady decline as degradation begins to become more dominant. The most striking feature is the continuous generation of significant amounts of carbon dioxide after day 30, with very low concurrent levels of VOC emission, which is primarily because of the biooxidation of nonvolatile hydrocarbon components (such as lighter PNAs), which are not responsive to removal through volatilization. Last, after 50 days of operation, the carbon dioxide levels precipitously decline in conjunction with VOCs becoming almost undetectable, which indicates that the hydrocarbons have been consumed and the stimulated bacteria have ran out of the hydrocarbons serving as their carbon source.

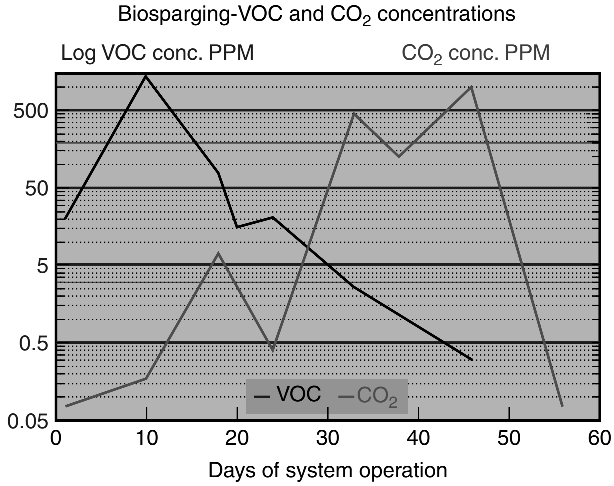

Figure 5. Hydrocarbon vs. carbon dioxide in soil gas above biosparging.

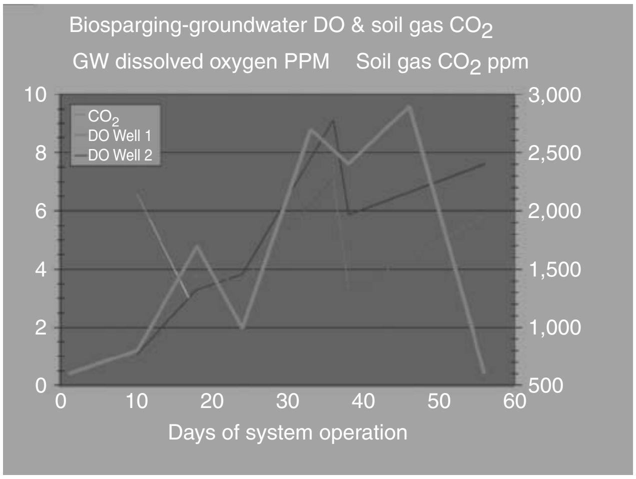

Figure 6. CO_2 in soil gas vs. dissolved oxygen in groundwater.

Figure 6 serves to further illustrate the biological activity in the saturated zone of the impacted soils. The initial low levels of the dissolved oxygen may have been because of chemical oxygen demand within the aquifer matrix (primarily dissolved iron). After meeting that COD, the DO increased in tandem with carbon dioxide production.

Of particular interest is the marked drop in DO between day 35 and day 40, which coincides with the peak biological activity. When the carbon dioxide levels began to drop, the DO increased once again. At the peak level of activity, the microbes were able to consume most of the oxygen provided by the injected air, as planned. With depletion of the food source, more oxygen is free to appear as DO. This site data has been presented to provide a topical field example of biosparging. It is from the early stages of the remediation. The hydrocarbon attenuation indicated by Figs. 5 and 6 is occurring in the advective zone of the impacted saturated zone soils and groundwater.

In conclusion, aeration is a powerful remediation tool in the vadose and saturated zones. The manner in which an aeration system is operated will determine the dominant process stimulated by that aeration: volatilization or biooxidation. Biooxidation has distinct advantages with regard to the range of nonvolatile hydrocarbons that can be remediated and offers lower potential off-gas treatment costs. The core issue is to understand the underlying mechanisms responsible for both processes and apply that knowledge accordingly.

BIBLIOGRAPHY

1. Haarhoff, J. and Cleasby, J.L. (1990). Evaluation of air stripping for the removal of organic drinking-water contaminants. *Water SA.* **16**(1): 13–22.
2. United States Environmental Protection Agency (USEPA). (2004). Biosparging. In: *How to Evaluate Alternate Cleanup Technologies for Underground Storage Tank Sites; A Guide for Corrective Action Plan for Regulators*. EPA 510-R-04-002, Washington, DC.
3. Suchomel, K.H., Kreamer, D.K., and Long, A. (1990). Production and transport of carbon dioxide in a contaminated vadose zone: a stable and radioactive carbon isotope study. *Environ. Sci. Technol.* **24**: 1824–1831.
4. Atlas, R.M. (1981). Microbial degradation of petroleum hydrocarbons: an environmental perspective. *Microbiolog. Rev.* **45**(1): 180–209.
5. Miller, R.N., Hinchee, R.E., Vogel, C.M., Dupount, R.R., and Downey, D.D. (1990). A field scale investigation of enhanced petroleum hydrocarbon biodegradation in the vadose zone at Tyndall AFB, Florida. *Proc. NWWA Petroleum Hydrocarbons and Organic Chemicals in Ground Water Conference*, Houston, TX, pp. 339–351.

READING LIST

Vance, D.B. (1993). Groundwater remediation—first principals. *National Environment. J.* **3**(2): 18–20.

Vance, D.B. (1993). Hydrophobic organic chemicals and total organic carbon—the ideal and reality of site clean-up. *National Environment. J.* **3**(3).

REMEDIATION OF CONTAMINATED SOILS

Dibyendu Sarkar
Rupali Datta
Ali H. Davani
University of Texas at San Antonio
San Antonio, Texas

Soil contamination is caused by a variety of industrial, agricultural, and domestic activities, and poses a major threat to the environment. Although soil contamination can result from a variety of organic and inorganic chemicals, this article focuses on remediation of metal-contaminated soils. Heavy metals such as Pb, Cr, Cd, Zn, Cu, and Hg can cause significant damage to environmental and human health, particularly because of their indefinite persistence in the environment. In the United States alone, 1200 sites are on the National Priority List (NPL) for remediation of contaminated soils, of which 63% are contaminated with toxic heavy metals. A plethora of remediation techniques are now available for the cleanup of such contaminated sites. The selection of the most appropriate site remediation technique depends on various factors such as site characteristics, concentration, types of pollutants to be removed, and end use of the contaminated medium. The major technologies available to date can be grouped under the following categories: isolation, immobilization, toxicity reduction, physical separation, and extraction. This article provides an overview of the various sources of heavy metal pollution in soils, the common remediation technologies available for soil cleanup, and the various factors that influence the type of remediation method that would be most appropriate for effectively remediating metal-contaminated soils.

INTRODUCTION

Environmental contamination due to human activities is an ancient problem that has been further aggravated by growing urbanization and increasing affluence. Dramatic accumulation of pollutants from chemical manufacturing, wastewater treatment, agriculture, mining and smelting, food processing, energy production, and pulp and paper production has led to deterioration of soil and water quality (1). Recent research is increasingly focused on the ecological and human health effects of the contaminants, and guidelines or standards on soil, drinking water, irrigation water, crop tissues, foodstuff, biosolids, and fertilizers are now becoming available (1). Heavy metals released by a variety of industrial, agricultural, and domestic activities are a potential risk to both human population and wildlife. In contrast to many organic pollutants, which are often degraded in the soil, the danger of heavy metals is aggravated by their indefinite persistence in the environment (2). Exposure to heavy metals has been linked to developmental retardation; various cancers; kidney damage; and development of autoimmunity leading to diseases of the joints, kidneys, circulatory and nervous systems. In addition, heavy metals can cause reduction in growth, inhibition in photosynthesis and respiration, and degradation of cell organelles in plants and microorganisms (3).

Remediation of contaminated soils involves reduction, extraction, removal, stabilization and/or containment of contaminants in the subsurface to a level acceptable for site closure. Site closure can be brought about by the reduction of risk posed by the presence of contaminants in the subsurface to an acceptable level. General approaches to remediating metal contamination include isolation, immobilization, toxicity reduction, physical separation, and extraction. Contaminated soil can be remediated by chemical, physical, or biological techniques. The available techniques can be grouped into two categories: (1) *ex situ* techniques that require removing the contaminated soil on- or off-site and (2) *in situ* methods that remediate without excavating the contaminated soil (4). The selection of the most appropriate soil and sediment remediation method depends on the site characteristics, concentration, types of pollutants to be removed, and the end use of the contaminated medium. The United States Environmental Protection Agency (USEPA) included 13 metals in its priority pollution list: Ag, As, Be, Cd, Cr, Cu, Hg, Ni, Pb, Sb, Se, Tl, and Zn (1). This review discusses the various sources of heavy metal pollution, evaluates the remediation methods available for soil cleanup, and the various factors that influence the type of remediation method that would be most appropriate for effective remediation.

SOURCES OF CONTAMINANTS

Major contributors to the metal burden in soils include discarded manufactured products in scrap heaps/landfills (As, Cr, Cu, Pb, Mn, and Zn), coal ashes (As, Cd, Pb, Mn, Hg, Mo, Ni, Se, V, and Zn), and agricultural and livestock wastes (As, Cu, and Zn) (1). Pollution sources can be broadly categorized into two sources: point sources and nonpoint sources. Point sources refer to discrete and localized contamination processes. Local accumulation of pollutants can arise from landfilling, mining, and smelter processes. Nonpoint sources are related to diffuse processes or human activities that cover large areas, for example, agricultural practices resulting in pesticide contamination (1). The major sources of soil metal contamination can be summarized as follows.

Airborne Sources

Airborne sources of metals include stack or duct emissions of air, gas, or vapor streams, and fugitive emissions such as dust from storage areas or waste piles. Some metals such as As, Cd, and Pb can also volatilize during high-temperature processing. These metals convert to oxides and condense as particulates, unless a reducing atmosphere is maintained (5).

Process Solid Waste

Process solid wastes result from a variety of industrial processes. These metal-bearing solid wastes are disposed above ground in waste piles, below ground, or under cover in landfills. Examples of process solid wastes include slag, fly ash, mold sand, abrasive waste, ion exchange resins, spent catalysts, spent activated carbon, and refractory bricks (6). Waste piles are exposed to weathering, which can lead to dispersal of the contamination to the surrounding soil, water, or air (7).

Sludge(Biosolids)

The composition of biosolids depends on the original waste stream and the process from which it was derived. Biosolids resulting from a uniform waste stream (such as wastewater treatment) are typically more homogeneous and have more uniform matrix characteristics. Sludge pits, on the other hand, often contain a mixture of wastes that have been aged and weathered, causing a variety of reactions to occur (5).

Soils

Soils can be contaminated as a result of spills or direct contact with contaminated waste streams such as airborne emissions, process solid wastes, sludges, or leachate from waste materials. The solubility of metals in soil is influenced by the chemistry of the soil and groundwater (1). Factors such as pH, pE, ion exchange capacity, and complexation/chelating with organic matter affect metal solubility directly.

CHEMICAL FATE AND MOBILITY OF METALS IN SOILS

The fate and transport of metals in soils depends significantly on the chemical form and speciation of metals (8). The term speciation means the distribution of elements among chemical forms or species. Heavy metals can occur in various forms in soils and water. For an accurate health risk assessment, not just the total metal concentration, but understanding its speciation is of paramount importance (9). In soils, speciation of metals depends on the physical and chemical characteristics of the soil, such as pH, redox, organic, carbonate, clay, and oxide content (10). A potential method for determining if heavy metals can be removed by a particular remediation technique or for predicting removal efficiencies is to determine speciation by selective extraction techniques. For example, it is believed that exchangeable carbonate and reducible oxide fractions may be amenable to soil washing techniques (11). Removal of the organically and residually bound fractions may not be economical or necessary to recover, as these fractions are generally not considered bioavailable (9).

REMEDIAL APPROACHES FOR METAL CONTAMINATED SOILS

Several technologies exist for the remediating metal contaminated soils (7), that can be broadly classified into five general categories; isolation, immobilization, toxicity reduction, physical separation, and extraction.

Isolation

Isolation technologies attempt to prevent the transport of contaminants by containing them within a designated area. Contaminated sites may also be isolated temporarily to limit transport during site assessment and site remediation (7). Isolation can be accomplished via these methods.

Capping. Capping systems provide an impermeable barrier to surface water infiltration into contaminated soil to prevent further release of contaminants into the surrounding surface water or groundwater (7). Other advantages of capping are control of odor emissions, improving aesthetics, and elimination of risks associated with human contact. Caps can be simple, single layered or complex multilayered systems (12). The selection of capping materials depends on site characteristics, remedial objectives, and risk factors associated with the site (12).

Subsurface Barriers. Subsurface barriers may be used to isolate contaminated soil and water by controlling the movement of groundwater at a contaminated site. Vertical barriers are commonly used to restrict the lateral flow of groundwater. For effective isolation of the contaminated matrix, the barrier should extend and key into a continuous, low-permeability layer, such as clay or competent bedrock, below the contaminated area (12,13). Slurry walls, sheet curtains, and grout piles are some of the widely used vertical barriers. Technologies for constructing horizontal barriers, such as grout injection by vertical boring and horizontal drilling, are also under investigation. Horizontal barriers would enable control of downward migration of contaminants by lining the site without requiring excavation of the contaminated matrix (7).

Immobilization

Immobilization technologies are designed to reduce the mobility of contaminants by changing the physical or leaching characteristics of the contaminated matrix. Mobility is usually decreased either by physically restricting contact between the contaminated matrix and surrounding groundwater or by chemically altering the contaminant to make it more stable with respect to dissolution (7). Immobilization techniques can be performed *ex situ* or *in situ* via these methods.

Solidification/Stabilization. Solidification and stabilization (S/S) immobilization technologies are the most commonly selected treatment options for metal contaminated sites (14). Solidification involves forming a solidified matrix that physically binds the contaminated material. Stabilization, also referred to as fixation, usually uses a chemical reaction to convert the waste to a less mobile form. Inorganic binders such as cement, fly ash, or blast furnace slag, or organic binders such as bitumen are used to form a barrier around the waste. The most important method used to immobilize metals is by precipitation of hydroxides within the solid matrix (15).

Vitrification. The mobility of metals can also be decreased by high-temperature treatment of the contaminated area that results in the formation of vitreous materials, usually a solid oxide. Most soils can be treated by vitrification, and a wide variety of inorganic and organic contaminants can be targeted (16). *Ex situ* vitrification includes steps such as excavation, pretreatment, mixing, feeding, and melting. The vitrified material can be recycled for use as clean fill, aggregate, or other reusable materials (5). *In situ* vitrification involves passing electric current through the soil using an array of electrodes inserted vertically into the soil (17).

Toxicity and/or Mobility Reduction

Chemical and/or biological processes can be used to alter the form of metal contaminants to decrease their toxicity and/or mobility.

Chemical Treatment. Chemical reactions can be initiated that are designed to reduce the toxicity or mobility of metal contaminants (7). The three types of reactions that can be used for this purpose are oxidation, reduction, and neutralization. Changing oxidation states of metals by oxidation or reduction can detoxify, precipitate, or solubilize metals (18). The most commonly used oxidizing agents are potassium permanganate, hydrogen peroxide, hypochlorite, and chlorine gas; some of the common reducing agents used are alkali metals, sulfur dioxide, sulfite salts, and ferrous sulfate. Chemical neutralization is used to adjust the pH of extremely acidic or basic soils. This process can be used to precipitate insoluble metal salts from contaminated water or to prepare for chemical oxidation or reduction. Chemical treatments can be performed both *in situ* or *ex situ*. Chemical agents are non specific, so they target all reactive metals, which might make them more toxic or mobile (18).

Biological Treatment. Until recently, biological techniques were more commonly used for remediating organic contaminants, but they are beginning to be applied for metal remediation. However, until now, most biological treatment applications have been at the bench and pilot scale; very few have been tested on a field scale. Biological treatment exploits natural biological processes that allow certain plants and microorganisms to aid in remediating metals through processes such as adsorption, oxidation/reduction reactions, and methylation (19). Biological treatment technology can be broadly classified.

Phytoremediation. Phytoremediation refers to the specific ability of plants to aid in metal remediation. Some plants have developed the ability to remove ions selectively from the soil to regulate uptake and distribution of metals in their tissues. Potentially useful techniques for metal remediation at contaminated sites include these (20).

Phytoextraction. Phytoextraction employs hyperaccumulating plants to remove metals from the soil by uptake into the roots and shoots of the plants. The aboveground shoots can be harvested to remove metals from the site and subsequently disposed of as hazardous waste or treated for the recovery of metals.

Phytostabilization. Phytostabilization involves using plants to limit the mobility and bioavailability of metals in soil. Phytostabilizing plants are characterized by high tolerance of metals in surrounding soils but low accumulation of metals in the plants.

Bioleaching. Bioleaching uses microorganisms to solubilize metal contaminants in the soil. This process has been adapted from the mining industry for metal remediation (7). Certain microorganisms can oxidize/reduce metal contaminants directly; others produce oxidizing/reducing agents that interact with metals to change their oxidation states, thereby decreasing their mobility and/or toxicity. Methylation by microorganisms involves attaching methyl groups to inorganic metal ions to form organometallic compounds, which can then be removed by volatilization (7).

Physical Separation

Physical separation is an ex-situ process that attempts to separate contaminated material from the rest of the soil matrix by exploiting certain physical characteristics of the metal and soil. Physical separation techniques operate on the basis of particle size, particle density, surface, and magnetic properties of the contaminated soil. These techniques are most effective when the metal is either in the form of discrete particles in the soil or if the metal is sorbed to soil particles that occur in a particular size fraction (21). Screening, classification, gravity concentration, and magnetic separation are some of the most important physical separation techniques used for soil remediation. *Screening* separates soils according to particle size when the soil passes through sieves with particular size openings. *Classification* involves separation of particles based on the velocity with which they fall through water (hydroclassification) or air (air classification) (21). *Gravity concentration* separates particles on the basis of gravity differences (21). *Magnetic separation* subjects particles to a strong magnetic field using electromagnets or magnetic filters and relies on the magnetic properties of minerals for separation (8).

Extraction

Heavy metals can be extracted from soil either by treating the soil with a solution containing chemical extractants or via electrokinetic processes.

Soil Washing. Washing can be used to remove metals from soil by chemical or physical treatment methods in aqueous suspension. Soil washing is an ex-situ process that requires soil excavation prior to treatment. Chemical treatment involves adding extraction agents that react with the contaminant and leach it from the soil (22,23). The type of extractant used depends on the contaminants present and the characteristics of the soil matrix. Many processes manipulate the acid/base chemistry of the slurry to leach contaminants from the soil. Chelating agents such as EDTA selectively bind some metals and can be used to solubilize metals from the soil matrix (22). Oxidizing and reducing agents, such as hydrogen peroxide or sodium borohydride, provide yet another option to aid in solubilizing metal because chemical oxidation/reduction can convert metals to more soluble forms (24). Surfactants can also be used to extract metals from soil (7).

Pyrometallurgical Extraction. Pyrometallurgical technologies use elevated temperature extraction and processing to remove metals from contaminated soils. Soils are treated in a high-temperature furnace to remove volatile metals from the solid phase. Subsequent treatment steps may include metal recovery or immobilization. This process usually produces a metal-bearing slag, from which the metals can be recovered for reuse (25).

In-situ Soil Flushing. This method is used to mobilize metals by leaching contaminants from soils so that they can be extracted without excavating the contaminated materials. An aqueous extracting solution is injected or sprayed onto the contaminated area to mobilize the contaminants, usually by solubilization. The extractant can be applied *via* surface flooding, sprinklers, leach fields, vertical/horizontal injection wells, basin infiltration systems, or trench infiltration systems (26). Similar extracting agents are used for in-situ soil flushing and soil washing, including acids/bases, chelating agents, and oxidizing/reducing agents (18). The liquid containing the contaminants is subjected to pump-and-treat methods. Once the water is pumped from the soil, the heavy metals are removed by sodium hydroxide or sodium sulfide precipitation, ion-exchange, activated carbon adsorption, ultrafiltration, reverse osmosis, electrolysis/electrodialysis, or biological means (9).

Electrokinetic Treatment. Electrokinetic remediation technologies apply a low density current to contaminated soil to mobilize contaminants in the form of charged species. The current is applied by inserting electrodes into the subsurface. This method relies on the natural conductivity of soils (due to the presence of water and salts) to influence movement of water, ions, and particulates through the soil profile. The contaminant metals are concentrated in the solution around the electrodes (5), and are removed by a variety of processes, including electroplating at the electrodes, precipitation/coprecipitation at the electrodes, complexation with ion exchange resins, or by pumping the water from the subsurface and treating it to recover the extracted metals (5).

SUMMARY AND FUTURE DIRECTIONS

Containment technologies, the least expensive, are widely used to control the transport of hazardous materials and to prevent the spread of contamination (9). Containment is the preferred remedial method for sites that have low levels of wastes of low toxicity and low mobility or wastes that have been pretreated to obtain these characteristics. The advantages of containment technologies include relatively simple and rapid implementation, often at a much lower cost than alternatives that require excavation, ability to treat large areas and volumes of waste, and the potential for successful containment as the final action at the site (7). Among the major disadvantages of containment technologies are the high level of uncertainties regarding the long-term effectiveness of the containment process and the need for long-term inspection because untreated contaminants remain on-site.

Metals are considered relatively immobile, so the remediation approaches generally used involve solid-phase processes such as solidification/stabilization and vitrification (9). These techniques are mostly done ex-situ, but they can be also used for in-situ remediation. They are useful at sites that are large, but the contamination is relatively shallow. The benefits of *in situ* remediation systems over conventional methods are the lower final cost for remediation, minimum cost for operations and maintenance, no moving parts that could break, and no discharge permits or waste disposal of liquids for in-situ groundwater treatment. Vitrification is expensive but can be used to remediate mixed wastes, where few technologies are currently available. Electrokinetics and soil flushing methods are also promising, and have been used successfully at a few sites (9). At present, phytoremediation and bioleaching are areas of primary focus of remediation research, because these techniques are environmentally friendly and relatively cheap. These techniques are useful for areas of low contamination, but longer treatment times are generally necessary. Some of the areas where research is still needed to use these technologies commercially are enhancing the metal accumulation capacity of plants by biotechnology, determining the correlation between soil properties and bioavailability, and developing methods to extract metals from plants (27).

BIBLIOGRAPHY

1. Adriano, D.C. (1986). *Trace Elements in Terrestrial Environments*, 2nd Edn. Springer-Verlag, New York.
2. Garbisu, C. and Alkorta, I. (2001). Phytoextraction: a cost-effective plant-based technology for the removal of metals from the environment. *Biores. Technol.* **77**: 229–236.
3. Vangronsveld, J. and Clijsters, H. (1994). Toxic effect of metals. In: *Plant and Chemical Elements*. M.G. Farago and V.C.H. Weinheim (Eds.). VCH Verlags Gesell Shaft, New York, pp. 149–177.
4. Khan, A.G. et al. (2000). Role of plants, mycorrhizae and phytochelators in heavy metal contaminated land remediation. *Chemosphere* **41**: 197–207.
5. Smith, L.A. et al. (1995). *Remedial Options for Metals Contaminated Sites*. Lewis, Boca Raton, FL.
6. Zimmerman, L. and Coles, C. (1992). Cement industry solutions to waste management-the utilization of processed waste by-products for cement manufacturing. In: *Cement Industry Solutions to Waste Management: Proc. 1st Int. Conf. Canadian Portland Cement Association*, Toronto, Canada, pp. 533–545.
7. Evanko, C.R. and Dzombak, D.A. (1997). *Remediation of Metal Contaminated Soils and Groundwater*. Technology Evaluation Report, TE-97-01. Groundwater Remediation Technologies Analysis Center, Pittsburgh, PA.
8. Allen, J.P. and Torres, I.G. (1991). Physical separation techniques for contaminated sediment. In: *Recent Developments in Separation Science*. N.N. Li (Ed.). Vol. V, CRC Press, Boca Raton, FL.
9. Mulligan, C.N., Yong, R.N., and Gibbs, B.F. (2001). Remediation technologies for metal-contaminated soils and groundwater: An evaluation. *Eng. Geol.* **60**: 193–207.
10. Kabata-Pendias, A. (1992). Trace metals in soils in Poland—occurrence and behavior. *Trace Substance Environ. Health* **25**: 53–70.
11. Li, W., Peters, R.W., Brewster, M.D., and Miller, G.A. (1995). Sequential extract evaluation of heavy metals-contaminated soils. How clean is clean? In: *Proc. Waste Manage. Assoc., 88th Annu. Meet. Exhibition*. San Antonio, TX., June 18–23.
12. Rumer, R.R. and Ryan, M.E. (1995). *Barrier Containment Technologies for Environmental Remediation Applications*. John Wiley & Sons, New York.
13. USEPA. (1985). *Handbook: Remedial Action at Waste Disposal Sites*. EPA/525/6-85/006, U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, DC.
14. Conner, J.R. (1990). *Chemical Fixation and Solidification of Hazardous Wastes*. Van Nostrand Reinhold, New York.
15. Bishop, P., Gress, D., and Olafsson, J. (1982). Cement stabilization of heavy metals: Leaching rate assessment. In: *Ind Wastes-Proc. 14th Mid-Atlantic Ind. Waste Conf.* C.A. Cole and D.A. Long (Eds.). Technomics, Lancaster, PA.
16. USEPA. (1992). *Vitrification Technologies for Treatment of Hazardous and Radioactive Waste*. Handbook EPA/625/R-92/002, U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.
17. Buelt, J.L. and Thompson, L.E. (1992). *The In-situ Vitrification Integrated Program: Focusing on an Innovative Solution on Environmental Restoration Needs*. Battelle Pacific Northwest Laboratory, Richland, WA.
18. NRC. (1994). *Alternatives for Groundwater Cleanup*. National Academy Press, Washington, DC.
19. Means, J.L. and Hinchee, R.E. (1994). *Emerging Technologies for Bioremediation of Metals*. Lewis, Boca Raton, FL.
20. Salt, D.E., Smith, R.D., and Raskin, I. (1998). Phytoremediation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**: 643–668.
21. Rosetti, P.K. (1993). *Possible Methods of Washing Fine Soil Particles Contaminated with Heavy Metals and Radionuclides*. M.S. Thesis, Carnegie Mellon University, Pittsburgh, PA.
22. Elliot, H.A. and Brown, G.A. (1989). Comparative evaluation of NTA and EDTA for extractive decontamination of Pb-polluted soils. *Water Air Soil Pollut.* **45**: 361–369.
23. Ellis, W.D. and Fogg, T. (1985). *Interim Report: Treatment of Soils Contaminated by Heavy Metals*. Hazardous Waste Engineering Research Laboratory, Office of Research and Development, USEPA, Cincinnati, OH.
24. Tuin, B.J.W. and Tels, M. (1990). Extraction kinetics of six heavy metals from contaminated clay soils. *Environ. Technol.* **11**: 541–554.
25. USEPA. (1996). *Engineering Bulletin: Technology Alternatives for the Remediation of Soils Contaminated with Arsenic, Cadmium, Chromium, Mercury and Lead*. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Cincinnati, OH.
26. USEPA. (1996). *Report: Recent Developments for In-situ Treatments of Metals-contaminated Soils*. PB96-153135, U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, DC.
27. Datta, R. and Sarkar, D. (2004). Effective integration of soil chemistry and plant molecular biology in phytoremediation of metals: an overview. *Environ. GeoSci.* (in press).

GROUNDWATER REMEDIATION PROJECT LIFE CYCLE

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

During the past decade, the groundwater remediation project life cycle has been used to characterize the status of an environmental project. If groundwater or soil contamination is suspected, a general process from assessment to closure occurs. In some cases, the project may involve many stakeholders, including landowners, responsible parties, bankers, operators/tenants, consultants, and regulators. In addition, others in the process might include environmental attorneys, realtors, developers, and interested parties such as neighbors and the community.

Although the timing and specific activities may vary from site to site, a general phased process occurs. Environmental evaluations of potential leakage or spillage of hazardous substances into the subsurface generally consist of four phases (Phase I through Phase IV). The generalized four-phase approach reflects the realistic process required to take a potentially impacted property from assessment phase (Phase I) through to the final monitoring and site closure phase (Phase IV). The phases generally proceed in order starting with Phase I. Some sites where soil or groundwater contamination is already suspected or documented might start with Phase II when a site has an underground tank or even Phase III, in the case of visually stained surface soils.

PHASE I

A Phase I Environmental Assessment is a set of noninvasive techniques to acquire information from site inspection and to obtain data supplied by others, including owners, operators, employees, and regulators. Sometimes, the Phase I Environmental Assessment is performed as part of a due diligence activity for a property transfer, as required by a financial institution on industrial, agricultural, and commercial properties. The transfer of ownership could occur during a corporate merger or acquisition or a routine sale. Other times, a Phase I Environmental Assessment is performed for financing or refinancing a loan on a commercial, industrial, or agricultural property. Other times, these activities might be required by regulatory agencies on properties when contamination is found, but no clear source of contamination is evident.

Phase I Environmental Assessments involve a site inspection, the development of the history of the property, and a review of data supplied by regulators, environmental lists and building permit databases, owners, tenants, and others.

A broad-based facility audit inspection focuses on the utilities, industrial processes, any signs of contamination, and all current and past permits and manifests (1). During the site facility inspection, the sources of all utilities (water, sewage, electrical, natural gas, electrical backup systems) are documented. Aboveground or underground storage tanks are frequently located at facilities that have backup power generators. The processes of the industrial facility from the input of the raw materials to the outflow of the finished products and associated liquid, solid and vapor wastes is reviewed. Recycling systems are documented. Visual clues of process activities such as staining, spillage, or leakage of liquids; and release of vapors or solid residues are documented and photographed. An investigator must also examine the permits for the facility, including any hazardous materials business plans, operating permits, hazardous material storage permits, building permits, waste generator permits, OSHA safety and training files, material safety data sheets (MSDS), environmental fines or compliance memos from regulatory agencies, and manifests for transportation and disposal of any hazardous materials generated on site.

Techniques include interviews of knowledgeable persons, review of historical aerial photography, and examination of published and unpublished historic maps. The maps include both historic topographic maps supplied by the USGS and Sanborn Fire Insurance Maps, as well as other local historic maps. Reverse telephone or address directories, chain-of-title reports, and regulatory databases may be reviewed. Noninvasive data may be collected on a property using handheld vapor meters. Other noninvasive techniques that would generate data include surface geophysics such as ground penetrating radar (GPR) and magnetic or induction surveys to locate buried tanks, pipes, or drums.

A passive vapor survey can be used to locate upward migrating volatile organic compounds that may lie within a few centimeters from the surface. This approach consists of placing a natural (carbon) or synthetic (Gore-sorber) adsorbent a few centimeters into the subsurface. The sorbent is usually placed in the upper end of an inverted glass or plastic container or test tube that has an open bottom where migrating vapors enter and collect. The collection device is left for an extended time, from a few days to a few weeks. Subsequently the adsorbent is removed and analyzed using a gas chromatograph. When the vapor data are evaluated, investigators can plot the various VOCs detected and concentration patterns on site maps. This exercise frequently allows an interpretation on the location of a source of detected shallow VOC contamination.

PHASE II

A Phase II Subsurface Investigation is normally invasive and is designed to evaluate the lithologic and hydrogeologic conditions by collecting soil, soil vapor, and groundwater samples. A variety of techniques and equipment are currently available for assessing the subsurface. Environmental subsurface investigation tools range in size, cost, and operating complexity from hand augers and hand-operated drive samplers to direct push technology (DPT) rigs to hollow stem auger rotary rigs. Samples are collected for field screening and physical testing. Selected soil, vapor, or groundwater samples are submitted under chain-of-custody procedures in a refrigerated environment (for volatile and semivolatile compounds) to a certified laboratory for chemical testing.

As part of Phase II, groundwater monitoring wells may be installed. Several Phase II Subsurface Investigations might be required prior to completing the Phase II process and fully characterizing the vertical and lateral extent of the soil and groundwater contamination.

PHASE III

A Phase III Corrective Action is the remediation portion of an environmental project. This phase involves designing and implementing the corrective action plan for remediation of soil and/or groundwater. A corrective action plan, evaluating remedial options and feasibility, is typically submitted to the regulator for approval prior to commencing the field work. A laboratory bench test might be designed to optimize biological or chemical conditions using soil and groundwater derived on site. This phase may also have a pilot test where the various technologies selected are field tested to verify the successes and failures prior to performing the full-scale remediation. After the bench tests and pilot scale tests, a final remedial action plan is usually prepared with the final selection of the remedial option.

Remediation might include extractive technologies (soil vapor extraction, groundwater pump and treat, dual phase extraction). These technologies remove liquids and vapors and treat the contaminants using a variety of above-ground technologies. Other technologies include bioremediation (*in situ*), chemical oxidation (*in situ*), monitored natural attenuation (*in situ*), soil excavation and removal to a landfill (dig and haul), or other technologies that are first evaluated in the corrective action plan.

Drilling techniques might be used during the remediation phase in designing and constructing vapor or groundwater extraction wells. In Phase III, it is unlikely that all contamination is removed, and therefore some residual amount of contamination will be left in place. A well-planned soil and groundwater confirmation sampling program is key to evaluating the success of a Phase III Remediation Project. As part of a Phase III Remediation Project, additional soil and groundwater sampling (a Remedial Investigation) might be planned if additional new source areas of impacted soil or groundwater are detected that were previously unknown.

Once the technology is selected and approved, the full-scale remediation is performed. Upon completing the full-scale remediation activities, equipment such as free-product pumping systems, soil vapor extraction, and aboveground air and water treatment units are demobilized and removed from the site. All drilling and remediation derived wastes are removed from the property. Site restoration may include replanting or resurfacing in disturbed areas.

PHASE IV

After remediation, the regulatory agencies generally require at least a minimum of 1 year quarterly groundwater monitoring verifying that the majority of the source of the contamination has been removed. This phase includes the proper abandonment of existing groundwater monitoring wells after site closure has been obtained. Phase IV might include a sensitive receptor survey to evaluate human health risks or sensitive environments that might be impacted if residual contamination is left in place. A risk based corrective action (RBCA) is a computer model that evaluates residual contamination as to the type of contamination, contamination levels in soil, vapor and groundwater, possible receptors, depth to the contamination, types of soils, and other factors. As part of the site closure process, deed restrictions may be required and notifications to any future contractors performing subsurface work that certain areas of the site might be impacted.

The Phase I to Phase IV project life cycle is an engineering approach that has been used on tens of thousands of sites. Every site is different, so the logical and stepwise approach is a generally accepted model for moving a project from initial discovery toward site closure.

BIBLIOGRAPHY

1. Wadge, L. (2000). Site evaluation, auditing and assessment. In: *Standard Handbook of Environmental Science, Health, and Technology*. J. Lehr (Ed.). McGraw-Hill, New York, pp. 11.1–11.6.

READING LIST

Jacobs, J. (2000). Monitoring well construction and sampling techniques. In: *Standard Handbook of Environmental Science, Health, and Technology*. J. Lehr (Ed.). McGraw-Hill, New York, pp. 11.46–11.68.

INNOVATIVE CONTAMINATED GROUNDWATER REMEDIATION TECHNOLOGIES

Thomas B. Boving
University of Rhode Island
Kingston, Rhode Island

INTRODUCTION

Contamination of groundwater is a problem of global proportions. Ever increasing demand for clean water for human, industrial, and agricultural uses, forces society to protect groundwater resources from further pollution and also to restore already polluted aquifers. Groundwater contamination originates from numerous sources, such as accidental spills, illegal dumping, leaking landfills, industrial wastewater discharge, and leakage of underground storage tanks. A major problem is that once toxic compounds have entered the subsurface, they are very difficult and costly to remove using conventional remediation strategies. Conventional remediation methods, such as flushing-out the contamination with water (known as the pump-and-treat method), generally have been proven inefficient. The National Research Council indicated that large sums of money are wasted by implementing these conventional methods (4).

The development of improved methods for remediating of contaminated aquifers has become a major environmental priority (1,2). Many innovative remediation technologies have emerged that promise fast and cost-effective cleanup of contaminated ground water. According to Cherry et al. (3), there are two general types of new technologies, those that destroy or immobilize contaminants *in situ* (e.g., chemical oxidation, chemical reductive dehalogenation, cometabolic and other bioremediation processes) and those that bring the contaminant mass to the surface for treatment or disposal (e.g., enhanced-flushing technologies, thermal treatment).

The following is a discussion of innovative remediation methods. The description of these new technologies is selective and by no means complete. It focuses on the *in situ* remediation of contaminated aquifers and cleanup technologies for vadose zone contamination but neglects *ex situ* treatment technologies (e.g., soil washing, thermal desorption). This is not an attempt to describe the theoretical and technical details involved but rather a summary of the concepts and challenges inherent in subsurface remediation. For a more detailed discussion of innovative remediation technologies, refer to a number of excellent publications (1–5) or good Internet sources (e.g., Ground-Water Remediation Technologies Analysis Center at www.gwrtac.org/html/techs.html, USEPA Technology Innovation Office at www.epa.gov/swertio1/remed.htm, or clu-in.org).

ANATOMY OF A SPILL

The area where a spill occurred and where most of the contaminant mass is concentrated is called the source zone. Sometimes the extent of the source zone is known, but more often it is not—especially in case of older spills and illegal dumping. The source zone can be limited to the vadose (i.e., unsaturated) zone or, if a sufficient volume of liquid contaminant has been spilled, can include the saturated zone. Once in the subsurface, a contaminant can partition into the vapor or aqueous phase, can remain stationary at residual concentration, or can continue to migrate as a free phase. Diffusive vapor phase transport is an important contaminant-spreading mechanism in the vadose zone (6). Dissolved contaminants in the saturated zone are transported away from the source zone by advection and dispersion. The contaminated water emanating from the source zone is called the plume. The general shape of the plume is that of an elongated teardrop; it tends to be much longer than wide. Contaminant concentrations are highest near the source zone and decrease toward the plume's outer edges. Even if the source zone is removed, the plume's center of mass (mean position of solute mass) can continue to move downgradient and spread, causing the plume to become larger. The plume continues spreading until it is intercepted by a natural or man-made hydrologic boundary, such as a river, drain, or well. In absence of a hydrologic boundary, the plume may eventually slow down or stop spreading entirely, which is the case when the contaminant degradation rate is in equilibrium with the contaminant-spreading rate. This equilibrium has been observed in some cases where petroleum products were released from underground storage tanks. Conversely, the degradation rates of many synthetic contaminants (especially those of halogenated compounds) are extremely slow. Thus, there are examples where a plume can grow across large distances, often for many miles.

INNOVATIVE REMEDIATION TECHNOLOGIES

A large variety of toxic compounds have entered the environment, including organic and inorganic compounds, heavy metals, radioactive substances, and biological contaminants. Realizing that each of these compound classes has distinct chemical and physical properties, quite different remediation approaches are required. For instance, a method that works well for removing heavy metals will most likely not work for a gasoline spill and vice versa. Thus, there are as many different remediation strategies as there are classes of toxic compounds. The hydrogeologic setting of a contaminated site may also elude or favor a specific remediation method, and, in some cases, a combination of two or more remediation technologies may be advantageous. The following is a description of innovative remediation techniques that remove or immobilize contaminants and methods that destroy contaminants *in situ*.

Removal and Immobilization Technologies

Removal and immobilization of contaminants are popular subsurface remediation strategies. Immobilization is a technique often used for heavy metal stabilization or fixation. This technology relies on injecting agents that form stable and insoluble metal precipitates with metals such as mercury or radioactive nuclides. Recently, the role of biogenic metal fixation processes has also been investigated. One of the greatest obstacles to *in situ* metal immobilization is delivery of the precipitant to the contaminant because it requires careful control of the subsurface conditions, such as pH. Therefore, immobilization has been most successfully used for treating excavated sediments, such as dredged materials.

Although applicable to many contaminant classes, contaminant removal is most often used to eliminate organic contaminants—especially nonaqueous phase liquids (NAPL). NAPLs have been identified as one of the most important groups of environmental contaminants (4,7). Examples include chlorinated solvents, gasoline, polychlorinated biphenyls, creosote, and coal tar.

NAPL's have distinct physicochemical properties and can exist in different compartments in subsurface environments, occupying the pore spaces as separate liquids or gaseous phases, dissolved in water, or attached to soil particles (8). Fluids whose densities are less than that of water are known as light nonaqueous phase liquids (LNAPL). Examples include gasoline and crude oil. LNAPLs accumulate on the water table and in the capillary fringe zone. They may also migrate downgradient away from the source zone, dissolve into the ground water, or may smear throughout the upper part of the aquifer due to water table fluctuations. In contrast, a DNAPL is a nonaqueous phase liquid that is denser than water. Examples include trichloroethene (TCE) and 1,1,1-trichloroethane. A DNAPL, therefore, can sink below the water table into deeper parts of the aquifer. Due to interfacial forces, a fraction of the contaminant may be retained in pores as immobile globules or ganglia, which tend to dissolve very slowly (9). If a DNAPL encounters less permeable geologic strata (such as fine silt or clay horizons), it may form pools. DNAPL pools, ganglia, and globules are a major problem for site remediation because they tend to be relatively small and dispersed and thus difficult to locate. With an unfavorably low surface area to mass ratio, these DNAPL bodies serve as long-term sources of groundwater contamination (3).

Conventional pump-and-treat systems may work well for plume containment, but it is widely recognized that this method is neither an effective nor economical means of complete remediation of sites contaminated with NAPLs (10). Pump-and-treat methods remove contaminants from the subsurface by forced displacement of the immiscible liquid due to artificially increased pore water velocities and by dissolution into the groundwater. Because most field sites require unrealistically high hydraulic gradients to displace residual NAPLs, dissolution is the primary removal mechanism. Thus, the inefficiency of the pump-and-treat approach is attributed to the low aqueous solubility of most NAPLs and the relatively large interfacial tensions between the groundwater and the NAPL. Overcoming low solubility and large interfacial tension is one major strategy for removing NAPLs from the subsurface. Methods that modify the contaminant solubility and interfacial tension are known as chemically enhanced flushing technologies. Based on the removal mechanism, chemically enhanced flushing technologies can be grouped loosely into two categories: (1) mobilization and removal of NAPL through immiscible displacement and (2) increase of the apparent solubility of a NAPL in the aqueous phase and removal by enhanced dissolution. Chemical agents that have been used for enhanced flushing include cosolvents (5,11), surfactants (5,12), dissolved organic matter (13), and complexing agents, such as cyclodextrin (14,15,18) (see Table 1).

Enhanced flushing of contaminated aquifers generally requires injecting an aqueous solution containing a solubility-enhancing agent. This solution is flushed through the contaminated aquifer and then extracted. Enhanced flushing works best in permeable sediments, such as sand and gravel. Conventional injection and extraction wells can be used to control the flow field of the flushing solution. This application scheme is, in principle, similar to conventional pump-and-treat systems. But due to the solubility enhancing properties of the agent in solution, mass removal rates are much higher, and consequently remediation times are much shorter. Many solubility-enhancing agents, such as most surfactants, also decrease the interfacial tension between the NAPL and the aqueous phase, which can result in mobilization of the NAPL. In some cases, mobilization is desirable because it increases the overall mass removal rate. However, if mobilized NAPL escapes hydraulic control and migrates off-site or into deeper aquifer horizons (as can be expected of DNAPLs), previously uncontaminated areas can become contaminated. Some surfactants are also sorbed to the sediment, which increases the amount of surfactant required to remove the NAPL. In addition, many cosolvents and surfactants are somewhat toxic, and there is concern about residuals left behind in the subsurface after primary remediation has been implemented. Thus, careful selection of appropriate enhancing agents and well-designed injection/extraction systems are crucial for successful application of chemically enhanced flushing technologies.

Thermal *in situ* treatment is a second approach for removing contaminants—especially volatile compounds and NAPL—from the subsurface. This method is based on heating polluted sediment and groundwater by hot air, water, steam, or energetic waves to move or mobilize contaminants toward extraction wells. The extraction wells capture the contaminants and deliver them to the surface for further treatment. In addition, heating can destroy contaminants that are unstable at elevated temperatures. Depending on the heat source, there are several thermal treatment methods available (Table 2). A special case of thermal treatment is *in situ* vitrification, which has been used to destroy or immobilize pesticides, heavy metals, and dioxins. This method requires heating the contaminated sediment by applying strong electrical currents until the contaminated sediment melts. After cooling, the vitrified material bears little resemblance to the original sediment, and it can no longer be considered an aquifer.

Air sparging and soil vapor extraction (SVE) are methods commonly used for removing volatile contaminants from groundwater and the vadose zone, respectively. Air sparging requires installing air injection and vacuum extraction wells. The injection wells are placed in the

Table 1. Chemically Enhanced Flushing Methods for Contaminant Removal from the Subsurface

Method	Removal Mechanism
Cyclodextrin flushing	Hydrophobic contaminant can form a complex within the hydrophobic interior of the cyclodextrin molecule. Because the cyclodextrin's exterior is hydrophilic, the contaminant/cyclodextrin complex has a higher solubility compared to the contaminant in aqueous solution. It can then be flushed out of the ground and treated on the surface.
Surfactant flushing	Above the critical micelle concentration, surfactant molecules cluster together to form micelles. The micelles have a hydrophobic interior into which hydrophobic contaminants can partition. Like cyclodextrin, the micelle's exterior is hydrophilic and therefore water-soluble. The contaminant bound within the micelle can then be flushed out of the ground and brought to the surface for further treatment. In addition, surfactant can lower the interfacial tension between a NAPL and the aqueous solution. This can result in NAPL mobilization.
Cosolvent flushing	The addition of cosolvents, e.g., ethyl or methyl alcohol, decreases the polarity of the aqueous phase. This enhances the dissolution of hydrophobic contaminants, which can then be flushed out of the ground. Sometimes, cosolvents and surfactants are applied together.
Dissolved organic matter flushing	Dissolved organic matter is usually applied as humic acid solution. The humic acid attracts hydrophobic contaminants and facilitates their transport. The humic acid/contaminant solution is then extracted from the ground and treated on the surface.

Table 2. Thermal Treatment Methods for Contaminant Removal from the Subsurface. Based on the Mechanism of Delivering Heat to the Contaminated Zone, Several Approaches are being Used

Method	Source of Heat and Removal Mechanism
Hot air injection	Hot air is injected through injection wells. Volatile contaminants evaporate or mobilize. The contaminant phase is captured by extraction wells. Some instable contaminants may break down when heated.
Steam injection	Similar to hot air injection, but steam is the source of heat.
Hot water injection	Injection of hot water into the source zone. Contaminants are mobilized and can be extracted from the subsurface.
Thermal conduction	Heat is transferred to the contaminated zone from heated steel wells or by using thermo blankets (shallow aquifers). Removal mechanisms are evaporation and destruction.
Radio-frequency heating	Contaminants evaporate when sediment is heated by radio waves. Wells function as antennas from which the radio waves are emitted.
Electrical resistance heating	Steel wells deliver an electrical current that heats up the sediment and groundwater. Contaminants evaporate or break down.

contaminated zone and screened below the water table. Air is pumped into the wells and forced to bubble through the contaminated aquifer. Volatile contaminants, such as gasoline and most solvents, move from the aqueous phase into the air phase. The vapors rise to the water table and into the vadose zone where they can be pulled out of the ground by vacuum extraction wells. SVE works similarly to air sparging, but vapors are removed only from the vadose zone. The vapors are pulled out of the unsaturated sediment using vacuum extraction wells. Additional air vents or air injection wells can aid SVE. In contrast to air injection wells, air vents passively allow fresh atmospheric air to enter the vadose zone. In most cases, injecting oxygen-rich atmospheric air into the subsurface enhances the bioactivity, which, in return, can result in faster breakdown of many contaminants. An extraction well can be converted into an injection well, which allows minimizing dead zones (parts of the contaminated zone that are bypassed). Depending on the size of the contaminated area, the number of wells can range from a few to more than a hundred. SVE and air sparging technologies are widely used because they are comparably inexpensive and require little engineering.

Phytoremediation is another innovative contaminant removal method. This method involves using plants to remediate contaminated water and sediments. The principal processes involved are phytostabilization, phytodegradation, phytoextraction, and phytovolatilization (Table 3). The contaminants are usually taken up in dissolved form. The efficiency of phytoremediation depends on a variety of factors, such as pH, soil properties, climate, and plant physiology. Phytoremediation, for example, has been used for remediating nitroaromatic explosives and heavy metals (19). A shortcoming of this method is that it is inefficient during winter months and it is confined to relatively shallow zones of contamination.

In Situ Destruction Technologies

The innovative remediation technologies that have been discussed so far were mostly aimed at immobilizing or removing the contaminant form the subsurface. A second strategy is based on *in situ* destruction of contaminants. Biotic and abiotic methods are available. Both approaches are discussed in the following sections.

Oxidation Processes. Many organic contaminants, such as petroleum hydrocarbons, solvents, and pesticides, can be destroyed or converted into harmless compounds when exposed to chemical oxidizers. Oxidation is a process by which an oxidant provides the chemical energy (e.g., as free oxygen radicals) to change the chemical composition of the target molecule into an oxidized and less harmful state. A simple example is the oxidation of methanol to carbon dioxide and water.

$$2CH_3OH + 3O_2 \longrightarrow 2CO_2 + 4H_2O$$

Potassium or sodium permanganate ($KMnO_4$, $NaMnO_4$) and hydrogen peroxide (H_2O_2) are commonly used oxidants for site remediation. When dissolved in water and injected into the contaminated zone, the oxidant comes in contact with the contaminant and destroys it. Because not all of the oxidant is consumed in most cases, it may be necessary to extract the remaining solution. The success of *in situ* oxidation is determined by the oxidant's ability to reach the contamination, which is relatively unproblematic in sand and gravel deposits but can be a challenge in dense silt and clay. Because permanganate converts into much less soluble manganese dioxide (MnO_2), precipitation and subsequent reduction in the aquifer's permeability may be problematic.

Another *in situ* oxidation technology is based on Fenton's reaction. Fenton's reagent is a mix of H_2O_2 and iron salts that are injected into the contaminated zone. The presence of the iron salt creates hydroxyl radicals that readily oxidize chlorinated solvents and other contaminants, such as fuel oils and BTEX. The Fenton's reagent reaction is much more efficient than H_2O_2 alone, but a disadvantage is that Fenton's reagent is toxic to microbes, making it incompatible with bioremediation.

Table 3. Phytoremediation Methods for Contaminant Removal from the Subsurface

Method	Removal and Remediation Mechanism
Phytostabilization	Contaminants are bound in the plant tissue.
Phytodegradation	Plant-assisted microbial or enzymatic breakdown of contaminants.
Phytoextraction	Bioconcentration of contaminants within harvestable zone of the plant.
Phytovolatilization	Plant takes up and volatilizes contaminants via its leaves.

The injection of ozone (O_3) has been used to remediate sites contaminated with polycyclic aromatic hydrocarbons (PAHs), fuels, chlorinated solvents, and other contaminants. Ozone is a strong oxidant and, compared to oxygen, is 12 times more soluble. Therefore, it can be delivered to the contaminant much more efficiently. However, the half-life of O_3 limits its ability to migrate through soil across large distances. Ozone is a gas and must be generated on-site, for example, by using ultraviolet light. *In situ* ultraviolet (UV) oxidation is a special case of ozone treatment. This technique exposes the contaminated groundwater to ozone and H_2O_2 generated by intense UV radiation. UV radiation is provided by UV light bulbs that are installed inside a well. The target contaminants are destroyed by direct reaction with the oxidizers and through the interaction with UV light.

Each oxidizing agent has its own special treatment application, and the selection of the oxidizing agent is based not only on the contaminant type but also on sediment and groundwater properties. Some oxidizing agents are corrosive or can cause explosions if not handled adequately. However, when designed properly, chemical oxidation is safe and has been used at hundreds of sites.

Reduction Processes. Environmental hydrogeologists have realized that zero-valent iron effectively reduces many halogenated contaminants if placed within the flow path of contaminated groundwater (16). This revelation resulted in the development of a new remediation technology, permeable reactive barriers (PRB). A PRB requires installing a permeable barrier ("wall") across the flow path of a contaminant plume. The plume moves passively through the wall, and the dissolved contaminants come in contact with reactive materials that either degrade or retain the contaminants. The earliest PRBs were iron walls that consisted of a narrow trench filled with zero-valent iron granules mixed with sand and gravel. These iron walls were first tested for passive treatment of dissolved chlorinated contaminants such as TCE. As the TCE comes in contact with the iron, its chlorine atoms are removed by one or more reductive dechlorination mechanisms. The necessary electrons are supplied by the oxidation of iron. Given sufficient residence time, the TCE degrades to harmless compounds (chloride, water, etc.), and clean groundwater flows out the other side of the PRB. Since then, other active agents such as a palladium-enhanced metal catalyst, chelators, sorbents (e.g., carbon), precipitators (e.g., limestone), and microbes were added to the PRB system to retain or destroy those compounds that the iron walls could not degrade sufficiently. By including a funnel-and-gate system the PRB's efficiency is further enhanced. The funnel-and-gate system for *in situ* treatment of contaminated plumes consists of barriers of low hydraulic conductivity (e.g., slurry walls or sheet piles). Like a funnel, the barriers channel the contaminated water to a gate that contains the *in situ* reactor (17).

PRBs are widely used to control organic contamination in ground water; but they have been applied to metals (e.g., chromium VI) and radionuclide contaminants too. Passive treatment walls are generally intended for long-term control of the migration of contaminants in groundwater. An important advantage of this technology is that no pumping and no moving parts are required. Thus, this technology may be more cost-effective than pump-and-treat methods. In the United States, several dozens PRBs have been installed, but this technology is currently limited to sites with relatively shallow plumes (less than 50 ft). The installation of a PRB requires a comparatively high initial investment and a commitment to long-term monitoring, which is because the treatment approach is limited to the contaminant plume and does not target the source zone itself.

Bioremediation. Microorganisms can use many contaminants as food and energy sources and, in the processes, break down these compounds. *In situ* groundwater bioremediation, also called enhanced biodegradation, is a remediation technology that encourages growth and reproduction of indigenous microorganisms (yeast, fungi, or bacteria) that enhance the biodegradation of toxic organic constituents. Creating a favorable environment for the microorganisms can stimulate the activity of microbes. Generally, this means providing some combination of electron acceptors (oxygen, nitrate), nutrients (nitrogen, phosphorous), moisture, and controlling the temperature and pH. In most cases, electron acceptors and nutrients are the two most critical components of a bioremediation system. These compounds are delivered by pumping air and nutrient solutions into the contaminated zone. Sometimes, microorganisms adapted to degradation of the specific contaminants or additional food sources (e.g., molasses) are added to enhance the biodegradation process.

Bioremediation can take place under anaerobic and aerobic conditions. In anaerobic conditions, no oxygen is present. The microorganisms break down chemical compounds and release the energy they need. Aerobic conditions support biological activity in which microorganisms use available atmospheric oxygen to function. Anoxic conditions exist when nitrate is the base of microorganism respiration. With sufficient oxygen, microorganisms will convert many organic contaminants to carbon dioxide and water. Degradation pathways can be very complex. Sometimes intermediate products that are less, equally, or even more toxic than the original contaminants are produced. Some contaminants are also degraded by enzymes produced by microbiological metabolism. These enzymes aid the degradation of contaminants that are difficult or impossible to treat otherwise (e.g., cometabolic biodegradation of TCE).

In situ groundwater bioremediation is a widely used technology because it can effectively degrade dissolved and adsorbed organic constituents. As a general rule of thumb, short-chain, low molecular weight, and more water-soluble compounds are biodegraded more rapidly and to lower residual levels compared to long-chain, high molecular weight, less soluble compounds. The effectiveness of bioremediation systems is controlled by how well electron acceptors and nutrients can be delivered to the contaminated zone and how bioavailable and biodegradable the target contaminants are. Aerobic degradation has been proven most effective for treating aliphatic and aromatic petroleum hydrocarbons (e.g., gasoline and diesel fuel). Anoxic, anaerobic, and cometabolic approaches are sometimes used for remediating chlorinated solvents, but remediation rates are generally slower compared to aerobic respiration of petroleum hydrocarbons.

Natural Attenuation. Dilution, dispersion, volatilization, biodegradation, adsorption, and chemical reactions with subsurface materials are processes that occur naturally and help to reduce the concentration, toxicity, and mobility of chemical or radioactive contaminants. Natural attenuation relies on these processes. Therefore, natural attenuation is not a technology as such, but rather is a cleanup strategy that uses nature's own decontamination capabilities. Nonetheless, natural attenuation is not a "do nothing" solution because it requires careful monitoring and testing of the processes that take place in the subsurface. For example, long-term contaminant degradation rates have to be sufficient to ensure public health and compliance with regulatory criteria. Therefore, monitoring of plume spreading or changes in hydrogeological, geochemical, or microbiological conditions that might reduce the effectiveness of natural attenuation is required.

Because natural attenuation does not require direct human intervention, it has become a very popular remediation strategy at many contaminated sites. However, natural attenuation may not be a viable method for removing the source of the pollution, because many of the natural attenuation mechanisms may be inactive in the presence of high contaminant concentrations. In these instances, other remediation methods must be used to remove the bulk contamination first.

CONCLUSION

Innovative remediation technologies have greatly improved the effectiveness of subsurface cleanup operations. Some technologies have already progressed into full-scale commercial applications; others remain in the developmental stage or have just emerged from research labs. However, remediation costs can be extremely high, and there is an ongoing debate about the economical cost of remediation versus the gains to public health. Against this background, improved risk assessment methods and concepts are necessary to ensure optimum remediation effectiveness—regardless of the remediation technology used.

BIBLIOGRAPHY

1. Grubb, D.G. and Sitar, N. (1994). *Evolution of technologies for in-situ cleanup of DNAPL contaminated sites*. EPA Report 600/R-94/120.
2. National Research Council (NRC) (U.S.). (1997). *Innovation in Ground Water and Soil Cleanup*. National Academy Press, Washington, DC.
3. Cherry, J.A. and Pankow, J.F. (1996). *Dense Chlorinated Solvents and other DNAPLs in Groundwater*. Waterloo Press, Portland, OR, P. 522.
4. National Research Council (NRC) (U.S.). (1994). *Alternatives for Ground Water Cleanup*. National Academy Press, Washington, DC.
5. Lowe, D.F., Oubre, C.L., and Ward, C.H. (1999). *Surfactants and Cosolvents for NAPL Remediation: A Technology Practices Manual*. Lewis Publishers, Boca Raton, FL, p. 412.
6. Conant, B.H., Gillham, R.W., and Mendoza, C.A. (1996). Vapor transport of trichloroethylene in the unsaturated zone; field and numerical modeling investigations. *Water Resources Research* **32**(1): 9–22.
7. Mercer, J.W. and Cohen, R.M. (1990). A review of immiscible fluids in the subsurface. *J. Contam. Hydrol.* **17**(6): 106–163.
8. Knox, R.C. and Sabatini, D.A. (1992). Transport and remediation of subsurface contaminants, colloidal, interfacial, and surfactant phenomena. *American Chemical Society Symposium Series* 491, American Chemical Society, Washington, DC.
9. Schwille, F. (1984). Pollutants in porous media: the unsaturated zone between soul surface and ground water. *Ecology Studies*. Springer-Verlag, New York, Vol. 47, pp. 27–48.
10. Mackay, D.M., Roberts, P.V., and Cherry, J.A. (1985). Transport of organic contaminants in ground water. *Environ. Sci. Technol.* **19**(5): 384–392.
11. Imhoff, P.T., Gleyzer, S.N., Mcbride, J.F., Vancho, L.A., Okuda, I., and Miller, C.T. (1995). Cosolvent-enhanced remediation of residual dense nonaqueous phase liquids; Experimental investigation. *Environ. Sci. Technol.* **29**(8): 1966–1976.
12. Fountain, J.C., Klimek, A., Beirkirch, M.G., and Middelton, T.M. (1991). The use of surfactants for in situ extraction of organic pollutants from a contaminated aquifer. *J. Hazardous Mater.* **28**: 295–311.
13. Johnson, W.P. and John, W.W. (1999). PCE solubilization and mobilization by commercial humic acid. *J. Contam. Hydrol.* **35**: 343–362.
14. Wang, X. and Brusseau, M.L. (1993). Solubilization of some low-polarity organic compounds by hydroxypropyl-β-cyclodextrin. *Environ. Sci. Technol.* **27**: 2821–2825.
15. Boving, T.B. and McCray, J.E. (2000). Cyclodextrin-enhanced remediation of organic and metal contaminants in porous media and ground water. *Remediation* **10**(2): 59–83.
16. Burris, D. and Cherry, J.A. (1992). Emerging plume managing technologies: In situ treatment zones. In: *Proceeding of the Air and Waste Management Association*. Annual Meeting, Kansas City, MO, June 21–26, 1992.
17. Star, R.C. and Cherry, J.A. (1994). In situ remediation of contaminated ground water: The funnel-and-gate-system. *Ground Water* **32**: 465–476.
18. Boving, T.B., Wang, X., and Brusseau, M.L. (1999). Cyclodextrin-enhanced solubilization and removal of residual chlorinated solvents from porous media. *Environmental Science and Technology* **33**(5): 764–770.
19. Rodgers, J.D. and Bunce, N.J. (2001). Treatment methods for the remediation of nitroaromatic explosives. Review paper. *Wat. Res.* **35**(9): 2101–2111.

RESISTIVITY METHODS

Hamid R. Jahani
Water Research Institute
Hakimieh, Tehran, Iran

INTRODUCTION

Many problems connected with groundwater can be investigated by geophysical methods. Of these, situation of water table, thickness of aquifer, delineation of weathered or fracture zone, fresh–salt water interfaces, groundwater quality, and groundwater pollutant flow direction are the most important.

Among all surface geophysical techniques for groundwater prospecting, geoelectrical methods are the most widely applied methods for both alluvial plains and karstic terrains all over the world because of their high capability in detecting water-bearing layers and because they are simple and inexpensive in carrying out field investigation.

The rock matrix of most geological formations is basically highly resistive and does not conduct electricity. There are, however, exceptions such as clay and shale, which are conducting minerals. These formations have low electrical resistivities when compared with other rock formations. Normally, the resistivity of a rock formation reduces only when it contains moisture. This reduction in resistivity depends on the relative quantity and quality of the water it contains.

In nature, groundwater is stored in geologically formed pore spaces. When these pores are interconnected, the groundwater moves from a high potential to a low potential zone. The rock layers at different depths have different "porosities" and, consequently, different water content. The more water in a formation, the lower the electrical resistivity, provided there is no other conducting layer such as clay or shale. If the contained water is brackish or saline, the resistivity of that layer drops even further (1).

The conductivity of an underground formation also varies as its homogeneity varies. The variation in conductivity affects the flow density, which, in turn, affects the distribution of the potential difference under the ground. The size of this effect depends on the size, shape, position, and resistivity of the heterogeneity.

If we suppose that heterogeneity is a spherical void in a karstic limestone filled with water or clay (a good example of a highly conductive material inside one of low conductivity), the concentration of the flow inside the void would be much more than that out of the void and, to the same extent, the variation in potential difference would be different (Fig. 1). This variation can be detected at the ground surface (2).

Thus, by measuring or determining the resistivities of earth layers at different depths, it is possible to infer indirectly the hydrogeologic character of a particular subsurface layer and also the quality of water it contains. The most applicable geoelectrical methods are resistivity methods in which an electrical current is passed into the ground so that the current distributes itself according to the conductivities of the geologic materials coming under the influence of the current.

RESISTIVITIES OF GEOLOGICAL FORMATIONS

The resistivity of rocks and minerals has a wide range (Fig. 2). For example, fresh igneous and metamorphic rocks have resistivities up to 10^5 to 10^6 ohm-m (ohm-meters), whereas shale and clay have resistivities less than 10 ohm-m. No other physical property found in geophysical prospecting varies so widely (3) and this is one of the best reasons for applying resistivity methods.

RESISTIVITY METHODS: THEORY AND APPLICATION

Resistivity (ρ) is a physical property of a substance. It is defined as the resistance to the flow of electric current by a unit cube of the substance when voltage is applied across the opposite faces. It is expressed in ohm-l (l is meters, feet, . . .). The resistance of a substance is a function of the resistivity, size, and shape. If the substance has a regular shape, then

$$R\rho L/A \tag{1}$$

where R = resistance, ohm;
ρ = resistivity of the substance, ohm-m;
L = length of the substance, m; and
A = cross-sectional area, m^2.

The inverse of resistivity is called conductivity.

We usually deal with heterogeneous media in geologic formations, so the term "apparent resistivity" is used instead of resistivity.

Measurement of Resistivity

In general, at least four electrodes are required for measuring the resistivities of subsurface formations. An electric current is introduced between one pair of electrodes, called current electrodes. The potential difference produced by current flow is measured by another pair of electrodes, called potential electrodes or probes. The apparent resistivity measured is (1):

$$\rho_{\mathrm{a}} = K(\Delta V)/I$$

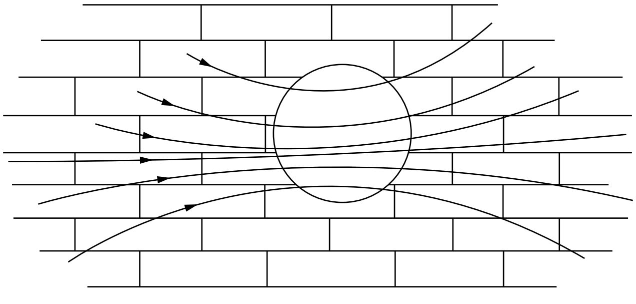

Figure 1. Concentration of electrical current in a spherical void filled with water or clay in a limestone formation (1).

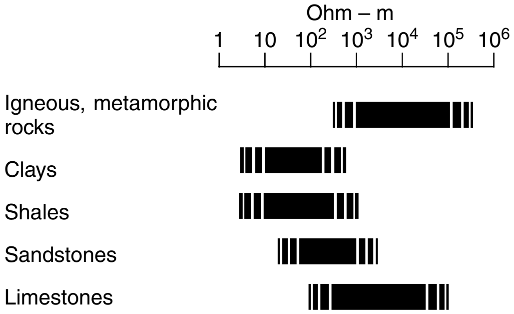

Figure 2. Ranges of resistivity in ohm-meters for some geologic formations (3).

where ρ_a = apparent resistivity, ohm-m;
K = geometric factor, which depends on the type of electrode configuration selected; and
I = electric current, (A).

Electrode Configurations. There are several electrode arrangements for geoelectrical operations (4):

Wenner Array. In this configuration, the four electrodes A, M, N, and B are placed along a straight line (Fig. 3), and $AM = MN = NB = a$.

Schlumberger Array. Schlumberger array is the most widely used method in electrical prospecting. In this configuration, the four electrodes are placed along a straight line in the same order as the Wenner array, but $AB > 5MN$ (Fig. 4).

Lee Partitioning Array. Lee partitioning array is the same as the Wenner array except that an additional potential electrode O is placed at the center of the array between the potential electrodes M and N (Fig. 5).

Dipole–Dipole Array. In this array, the distance between the current electrodes A and B (current dipole) and the distance between the potential electrodes M and N (measuring dipole) are significantly smaller than the distance r between the centers of the two dipoles (Fig. 6).

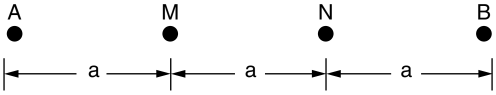

Figure 3. Wenner array (5).

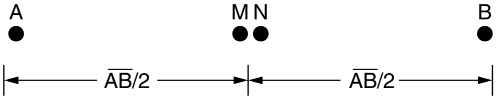

Figure 4. Schlumberger array (5).

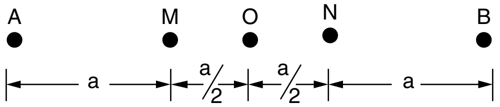

Figure 5. Lee partitioning array (5).

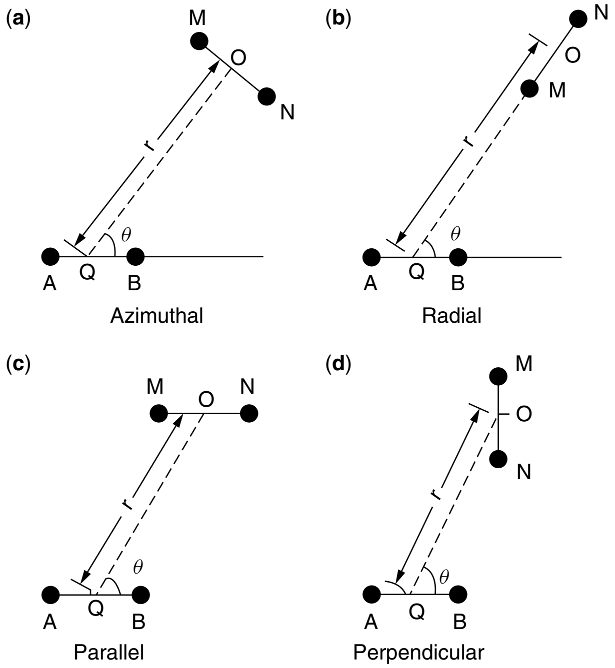

Figure 6. Dipole–dipole arrays (5).

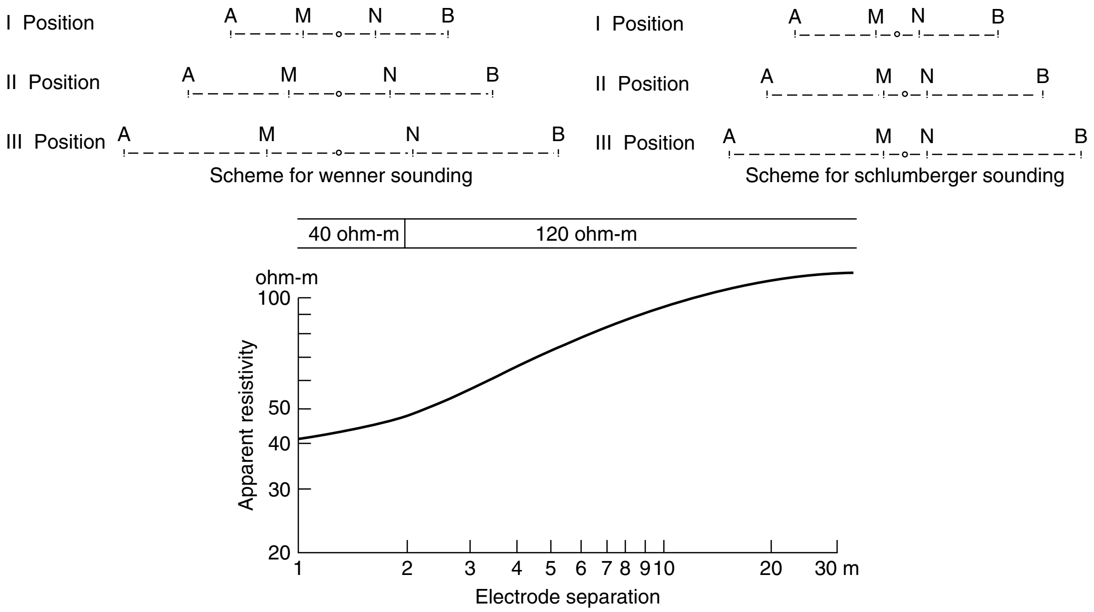

Figure 7. Scheme of array expansion for electrical sounding in Wenner and Schlumberger arrays (up). Variation of apparent resistivity with electrode separation for a two-layer earth where $\rho_{\text{layer2}} > \rho_{\text{layer1}}$ (down) (1,3).

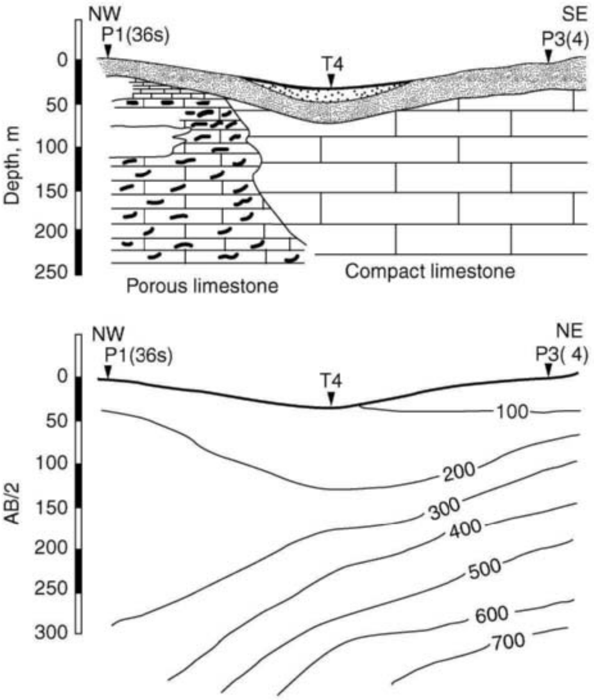

Figure 8. A geologic cross section along a karstic formation (up) and the apparent resistivity section for the same place (down) at Oshnavieh area, western Azerbaijan, Iran (6).

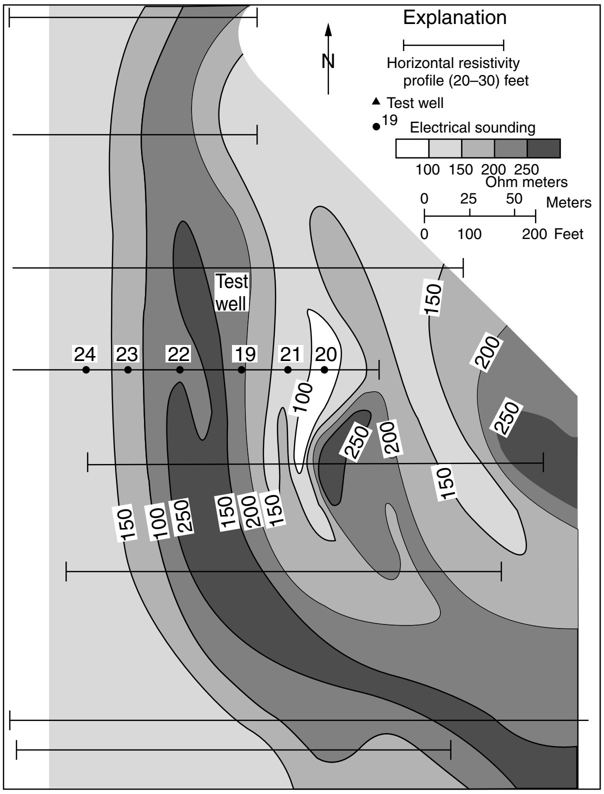

Figure 9. An apparent resistivity map for a site at Campbel, California. The low resistivity zone delineates a buried stream channel (5).

Electrical Sounding and Horizontal Profiling

Electrical sounding is a process for depth investigations, and horizontal profiling is a process for detecting lateral variations in resistivity. The basis for an electrical sounding, regardless of the electrode array selected, is that the farther away from a current source the measurement of the potential is made, the deeper the probing will be. So a curve of electrode spacing against apparent resistivity is made (Fig. 7).

In horizontal profiling, a fixed electrode spacing is chosen (preferably based on the results from a previous

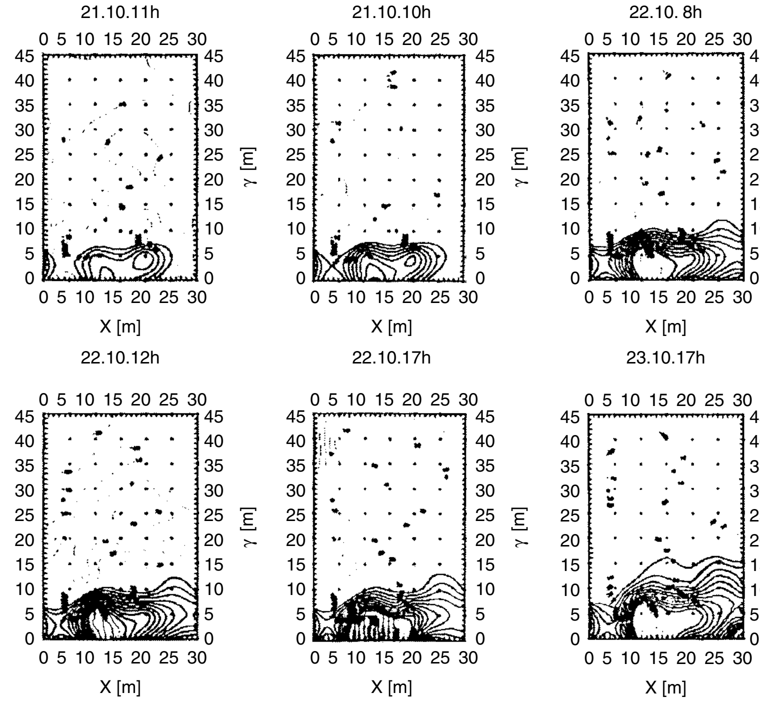

Figure 10. Determining pollution flow direction and velocity by the resistivity method (9).

electrical sounding), and the whole electrode array is moved along a profile horizontally. So a curve shows the variation of apparent resistivity for different points along the profile. Maximum apparent resistivity anomalies are obtained by orienting the profiles normal to the strike of the geologic structures.

Usually, the result of electrical sounding and horizontal profiling are apparent resistivity sections and maps. These plots indicate the variations in geoelectrical properties of the study area in vertical and planar surfaces, respectively. These plots can help in understanding the underground structure and lithology of different geoelectrical behaviors known for different geologic formations (Figs. 8 and 9).

RESISTIVITY METHODS AND GROUNDWATER POLLUTION STUDIES

Geoelectrical methods can be used easily to investigate groundwater contamination because pollution or any dissolved solid in groundwater directly affects its electrical conductivity and resistivity. So this method can be used to detect pollution plumes provided that the pollution has changed the resistivity of the groundwater enough to be detected.

Two case studies will be presented here to show the capability of the technique. In the first, a "tracer" has been intentionally introduced into the groundwater system, and the changes in groundwater resistivity have been investigated to find the velocity and direction of the tracer movement. In the second case, the resistivity method has been used to map the extent of the pollution zone in groundwater beneath a waste disposal site.

Case Study 1

This case is actually a single well technique in which a tracer is injected into the aquifer through a borehole. The tracer is NaCl, which is cheap and also has good electric properties. It is important to remember that this substance is selected to increase the conductivity of the water. If the study site is composed of saline soil and saline groundwater, injection of fresh water would again result in a suitable resistivity contrast. Before injection, geoelectrical sounding and horizontal profiling should be performed so that the base geoelectrical condition of the formation can be understood.

The study case is an unconfined shallow aquifer consisting of coarse gravel in Germany. A network for resistivity measurements is considered along the south–north direction, which, it is thought, is the direction of groundwater flow. After injecting 250 kg of salt, resistivity was measured every 5–6 hours at all points. The result is shown as isoresistivity maps in Fig. 10. The movement of the salt plume is clearly shown on the maps, and its velocity has been calculated as 7 meters per day. Contrary to what was supposed before, the direction of the movement of the plume was almost from west to east (7).

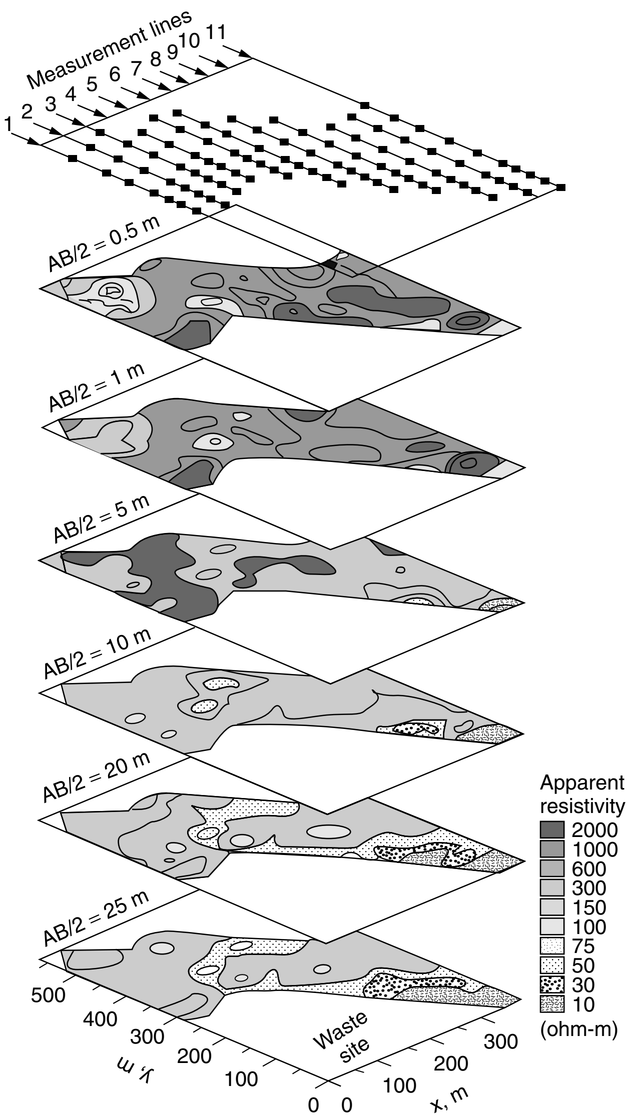

Figure 11. Geoelectrical measurement lines and apparent resistivity maps for different electrode spacings (8).

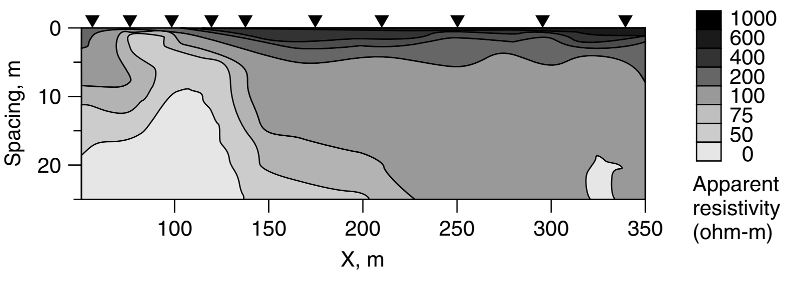

Figure 12. Resistivity pseudosection from south–north along measurement line 9 (8).

Next Page

Case Study 2

In this case, the resistivity method has been used to map the extent of contamination induced by an open waste disposal site and to determine where to locate future monitoring wells. The site is located SE of the city Isparta, Turkey, on a nearly 100 m thick Quaternary alluvial deposit of gravel, sand, and clay. The Schlumberger sounding resistivity method was used to map resistivities from 0.5 down to 25 meters in depth. Figure 11 shows the apparent resistivity maps for different electrode spacings. As expected, leachate from the waste disposal site has decreased groundwater resistivity. The deeper layers are characterized by low resistivities, and the shallower ones by high resistivities.

A resistivity pseudosection along measurement line 9 is performed, and the result is shown in Fig. 12. As can be seen, the low resistivity front from the waste disposal site base is characterized by its low resistivity and is developing northward. So, the extent of groundwater pollution can easily be manifested using geoelectrical mapping (8).

BIBLIOGRAPHY

1. Rao, T.G. (Ed.). (1992). *Groundwater Exploration Techniques*. National Geophysical Research Institute, Hyderabad, India.
2. Jahani, H.R. (2005). *Applied Geophysical Methods for Groundwater Investigations*. Iranian Hydraulic Association, Tehran, Iran.
3. Griffiths, D.H. and King, R.F. (1988). *Applied Geophysics for Geologists and Engineers*. Pergamon Press, Oxford.
4. U.S. Soil Conservation Service. (1971). *Seismic and Resistivity Methods of Geophysical Exploration*. Water Resources Publications, United States Soil Conservation Service, Washington, DC.
5. Zohdy, A.A.R. et al. (1990). *Application of Surface Geophysics to Groundwater Investigations*. U.S. Geological Survey, Denver, CO.
6. Iranian National Karst Research Center. (1997). *Resistivity Methods of Geophysical Explorations in Karstic Areas in West Azerbaijan Province*. Shiraz.
7. Hotzl, H. and Werner, A. (Eds.). (1992). *Tracer Hydrology*. A.A. Balkema, Rotterdam, The Netherlands.
8. Karlýk, G. and Ali Kaya, M. (2001). Investigation of groundwater contamination using electric and electromagnetic methods at an open waste-disposal site: A case study from Isparta, Turkey. *Environ. Geol.* **40**(6): 725–731.

RISK ANALYSIS OF BURIED WASTES FROM ELECTRICITY GENERATION

Bernard L. Cohen
University of Pittsburgh
Pittsburgh, Pennsylvania

An important aspect of generating electricity from nuclear power is the disposal of the high level radioactive waste. The most widely anticipated plan is to convert it into a rock-like material and bury it deep underground, typically 600 meters below the surface. In evaluating the safety of this procedure, a key element is to estimate the probability per year, P, for an atom of this buried waste to be dissolved by groundwater and eventually enter a human stomach. Obviously, this probability is site dependent and depends on present underground characteristics and future events in the region, many of them unpredictable, including climate changes, geological land uplifting causing rivers to change their courses, volcanoes, earthquakes, and human intrusion.

INTRODUCTION

A first estimate of P can be obtained by considering that all properties of the region are the U.S. average (lower 48 states); the purpose of this article is to obtain such an estimate. Most relevant characteristics and future changes in them are occurring somewhere in the United States and are hence taken into account with or including some estimate of their probability of occurrence. The value of P we obtain can be interpreted as the average value of P for a large number of randomly chosen sites. It seems reasonable to assume that the very elaborate site selection efforts by geological and hydrologic experts should choose a site at least as secure as a randomly chosen site.

We assume that the waste converted to a rock-like material behaves like average rock—differences between them will be considered later—and calculate P for average rock. P can be calculated as the dissolution rate, R, the probability per year for an atom of the rock to be dissolved in groundwater, times the probability, p, for an atom once dissolved into groundwater to enter a human stomach:

$$P = R \times p \tag{1}$$

Dissolution Rate for Rock, *R*

We begin by calculating R. From the rate at which rivers carry dissolved and suspended material into the oceans, it is estimated (1) that the surface of the continent is eroding at an average rate of 5×10^{-5} (hereafter, we use the notation 5 E−5) meters per year (hereafter, m/y). About 28% of this material is in solution, corresponding to 1.4 E−5 m/y removed by chemical dissolution. It is estimated (2,3) that 15% of the water flow in rivers is derived from groundwater (aquifers), and the rest comes from surface runoff. If the concentrations of dissolved materials were the same in rivers as in aquifers, this would mean that the latter dissolve $(0.15 \times 1.4\text{ E}{-5}) = 2.1\text{ E}{-6}$ m/y of rock thickness. The average concentrations of a few dissolved materials in aquifers (4) and in river water (5) are listed in Table 1. Silica and calcium are such important

Table 1. Average Concentrations in mg/kg of Dissolved Materials in Aquifers[a] and in River Water[b]

Material	Silica	Calcium	Magnesium	Potassium	Iron	Uranium
Aquifers	20	30	3	2	0.3	3 E−4
Rivers	7	15	4	2.2	0.5	4 E−4

[a] Ref. 4.
[b] Ref. 5.

Previous Page

Case Study 2

In this case, the resistivity method has been used to map the extent of contamination induced by an open waste disposal site and to determine where to locate future monitoring wells. The site is located SE of the city Isparta, Turkey, on a nearly 100 m thick Quaternary alluvial deposit of gravel, sand, and clay. The Schlumberger sounding resistivity method was used to map resistivities from 0.5 down to 25 meters in depth. Figure 11 shows the apparent resistivity maps for different electrode spacings. As expected, leachate from the waste disposal site has decreased groundwater resistivity. The deeper layers are characterized by low resistivities, and the shallower ones by high resistivities.

A resistivity pseudosection along measurement line 9 is performed, and the result is shown in Fig. 12. As can be seen, the low resistivity front from the waste disposal site base is characterized by its low resistivity and is developing northward. So, the extent of groundwater pollution can easily be manifested using geoelectrical mapping (8).

BIBLIOGRAPHY

1. Rao, T.G. (Ed.). (1992). *Groundwater Exploration Techniques*. National Geophysical Research Institute, Hyderabad, India.
2. Jahani, H.R. (2005). *Applied Geophysical Methods for Groundwater Investigations*. Iranian Hydraulic Association, Tehran, Iran.
3. Griffiths, D.H. and King, R.F. (1988). *Applied Geophysics for Geologists and Engineers*. Pergamon Press, Oxford.
4. U.S. Soil Conservation Service. (1971). *Seismic and Resistivity Methods of Geophysical Exploration*. Water Resources Publications, United States Soil Conservation Service, Washington, DC.
5. Zohdy, A.A.R. et al. (1990). *Application of Surface Geophysics to Groundwater Investigations*. U.S. Geological Survey, Denver, CO.
6. Iranian National Karst Research Center. (1997). *Resistivity Methods of Geophysical Explorations in Karstic Areas in West Azerbaijan Province*. Shiraz.
7. Hotzl, H. and Werner, A. (Eds.). (1992). *Tracer Hydrology*. A.A. Balkema, Rotterdam, The Netherlands.
8. Karlýk, G. and Ali Kaya, M. (2001). Investigation of groundwater contamination using electric and electromagnetic methods at an open waste-disposal site: A case study from Isparta, Turkey. *Environ. Geol.* **40**(6): 725–731.

RISK ANALYSIS OF BURIED WASTES FROM ELECTRICITY GENERATION

Bernard L. Cohen
University of Pittsburgh
Pittsburgh, Pennsylvania

An important aspect of generating electricity from nuclear power is the disposal of the high level radioactive waste. The most widely anticipated plan is to convert it into a rock-like material and bury it deep underground, typically 600 meters below the surface. In evaluating the safety of this procedure, a key element is to estimate the probability per year, P, for an atom of this buried waste to be dissolved by groundwater and eventually enter a human stomach. Obviously, this probability is site dependent and depends on present underground characteristics and future events in the region, many of them unpredictable, including climate changes, geological land uplifting causing rivers to change their courses, volcanoes, earthquakes, and human intrusion.

INTRODUCTION

A first estimate of P can be obtained by considering that all properties of the region are the U.S. average (lower 48 states); the purpose of this article is to obtain such an estimate. Most relevant characteristics and future changes in them are occurring somewhere in the United States and are hence taken into account with or including some estimate of their probability of occurrence. The value of P we obtain can be interpreted as the average value of P for a large number of randomly chosen sites. It seems reasonable to assume that the very elaborate site selection efforts by geological and hydrologic experts should choose a site at least as secure as a randomly chosen site.

We assume that the waste converted to a rock-like material behaves like average rock—differences between them will be considered later—and calculate P for average rock. P can be calculated as the dissolution rate, R, the probability per year for an atom of the rock to be dissolved in groundwater, times the probability, p, for an atom once dissolved into groundwater to enter a human stomach:

$$P = R \times p \tag{1}$$

Dissolution Rate for Rock, *R*

We begin by calculating R. From the rate at which rivers carry dissolved and suspended material into the oceans, it is estimated (1) that the surface of the continent is eroding at an average rate of 5×10^{-5} (hereafter, we use the notation 5 E−5) meters per year (hereafter, m/y). About 28% of this material is in solution, corresponding to 1.4 E−5 m/y removed by chemical dissolution. It is estimated (2,3) that 15% of the water flow in rivers is derived from groundwater (aquifers), and the rest comes from surface runoff. If the concentrations of dissolved materials were the same in rivers as in aquifers, this would mean that the latter dissolve $(0.15 \times 1.4\ \text{E}{-5}) = 2.1\ \text{E}{-6}$ m/y of rock thickness. The average concentrations of a few dissolved materials in aquifers (4) and in river water (5) are listed in Table 1. Silica and calcium are such important

Table 1. Average Concentrations in mg/kg of Dissolved Materials in Aquifers[a] and in River Water[b]

Material	Silica	Calcium	Magnesium	Potassium	Iron	Uranium
Aquifers	20	30	3	2	0.3	3 E−4
Rivers	7	15	4	2.2	0.5	4 E−4

[a] Ref. 4.
[b] Ref. 5.

components of rock that we interpret these data to indicate that aquifers carry about twice the concentration of dissolved material in rivers, which leads to the conclusion that $[30/(30+85)] = 26\%$ of the dissolved material in rivers, $(0.26 \times 1.4\ \mathrm{E}{-5}) = 3.6\ \mathrm{E}{-6}$ m/y of rock depth, is contributed by aquifer dissolution of rock.

The next problem is to estimate what fraction of this material, f, is removed per meter of depth at 600 m; let us say from between 599 m and 600 m. As a crude gross overestimate, we might assume that rock erosion is constant with depth down to 600 m and zero below 600 m, in which case $f = 1/600 = 16\ \mathrm{E}{-4}$. A more reasonable approach is as follows: The annual circulation of groundwater, defined as the quantity per year entering or leaving aquifers at the stated depth, has been given (2,3) as

shallow (<800 m) : 310 E9 cubic meters per year

deep (>800 m) : 6.2 E9 cubic meters per year

If we assume that the flow decreases exponentially with depth and that the rate of rock dissolution is proportional to this groundwater flow, determining the constants in the relationship from these data, we find (6) that $f = 2.6\ \mathrm{E}{-4}$, one-sixth of our crude gross overestimate.

Using this gives the quantity of rock eroded per year per meter of depth at 600 m below the surface as $(3.6\ \mathrm{E}{-6} \times 2.6\ \mathrm{E}{-4}) = 1.0\ \mathrm{E}{-9}$ m/y. If 1.0 E−9 m/y of depth is dissolved from one meter of rock depth, the probability per year for an atom of that rock to be dissolved must be 1.0 E−9. Thus,

$$R = 1.0\ \mathrm{E}{-9} \tag{2}$$

This result plays a key role in our risk analysis, so it is interesting to seek alternative approaches to calculating it. One such approach follows: A typical aquifer reaching to the waste burial depth of 600 m may be about 100 km long and have a flow velocity of 100 m/y through rock of 10% porosity, discharging into a river The water discharged from it annually per square meter of cross-sectional area is then 10% of the volume of a column of water 100 m long and one square meter in cross section, which is 10 cubic meters, or 10,000 liters (L).

Chemical analyses of groundwater (4) indicate that it typically contains 30 mg/L or 30 E−6 kg/L (milligrams per liter, or kilograms per liter) of Calcium (Ca), so it discharges (30 E−6 kg/L × 10,000L =) 0.3 kg of Ca into the river each year. This is the first entry in Column (2) of Table 2. This calcium was derived by dissolution from the rock through which the aquifer had passed, 100 km, or 1E 5 m, long, one square meter in cross section, and a specific gravity of about 3.0, which gives it a mass of $(3000 \times 1\ \mathrm{E}5) = 3\ \mathrm{E}8$ Kg. Typical rock contains 5% Ca (1), so the Ca contained in this source rock is $(0.05 \times 3\ \mathrm{E}8) = 15\ \mathrm{E}6$ kg; this is the first entry in Column (3) of Table 2. If 0.3 kg of Ca per year is dissolved from 15 E6 kg of Ca in the rock, the probability for a Ca atom to be dissolved must be $(0.3/15\ \mathrm{E}6) = 2\ \mathrm{E}{-8}$ per year; this is the first entry in Column (4) of Table 2. Table 2 includes data for other materials, obtained analogously to that for Ca.

Table 2. Calculation of Fractional Removal per year of Rock Materials by a Typical Aquifer. Columns (2) and (3) refer to a 1 Square Meter Cross Section of the Aquifer. See Discussion in Text

(1) Material	(2) Kg/y into River	(3) Kg in Rock (× 1 E6)	(4) Fraction Removed (× 1 E−8)
Ca	0.3	15	2
Mg	0.03	3	1
K	0.02	3	0.7
Fe	0.003	9	0.03
U	3 E−6	8 E−4	0.3
Silica	0.2	150	0.13
Carbonate	1.3	18	8

Of the materials listed in Table 2, the materials in the radioactive waste are most similar to Fe and U and not at all similar to carbonate ions. Thus, the numbers in Column (4) of Table 2 might give an estimate of about $R = 7\ \mathrm{E}{-9}$, seven times higher than Eq. 2. But the aquifer we have assumed in this calculation is substantially more robust than the groundwater encountered by average rock at 600 m depth, so it seems reasonable to conclude that Eq. 2 is roughly verified.

Probability of Transport from Groundwater to Human Stomachs, *p*

Next we turn to the problem of estimating the probability of transferring an atom of material dissolved in groundwater into human stomachs, p in Eq. 1. We assume that once the material is dissolved in groundwater, it moves with groundwater, eventually reaching shallow aquifers that feed into rivers. Contributions to p derive from our use of rivers and wells drilled into aquifers for potable water, from fish removed from rivers and used for food, and from use of well and river water for irrigating food crops. The potable water path is the simplest to calculate and, it turns out, the most important. The average person ingests 2 L/day of potable water which corresponds to ingestion by the U.S. population of $2 \times 365 \times 2.8\ \mathrm{E}8 = 2.0\ \mathrm{E}11$ L/y (liters per year). Of this, 45% is derived from wells, and 55% comes from rivers (7), corresponding to 9 E10 and 1.1 E11 L/y respectively entering human stomachs. The estimated water flow in U.S. rivers is 1.7 E15 L/y (1) and 1.9 E15 L/y (2,3); we use 1.8 E15 L/y. The estimated water flow in aquifers is 16% of the flow in rivers (3), or 2.9 E14 L/y The contributions to p from the use of rivers and wells for potable water are then

$$p(\text{rivers}) = (1.1\ \mathrm{E}11/1.8\ \mathrm{E}15) = 6.1\ \mathrm{E}{-5}$$

$$p(\text{well water}) = (9\ \mathrm{E}10/2.9\ \mathrm{E}14) = 3.1\ \mathrm{E}{-4}$$

Note that we have ignored removal of material by filtration processes, making these conservative estimates, that is, more likely to be high than low.

Analyses of the fish and irrigation contributions to p will be sketched below. They give results that vary considerably from element to element because of variations in bioaccumulative factors (8). For the fish pathway, $p(\text{fish})$ is rarely larger than 1 E−5, and for most important elements in the waste, it is much smaller. For the irrigation pathway, $p(\text{irrigation})$ is rarely more than 1 E−4 and somewhat less for the most important elements in the waste. We therefore take as conservative estimates

$$p(\text{fish}) = 1\ \text{E}{-5}$$

$$p(\text{irrigation}) = 1\ \text{E}{-4}$$

Summing these four contributions gives $p = 5\ \text{E}{-4}$. Using this and Eq. 2 in Eq. 1 then gives the result we are seeking for P:

$$P = (1\ \text{E}{-9} \times 5\ \text{E}{-4}) = 5\ \text{E}{-13} \tag{3}$$

Alternative Approach for Transfer from Rock to Human Stomachs

We now present an essentially independent method (9) for calculating P, based on our knowledge of quantities of elements entering human stomachs each year (10,11) and the quantities of these materials in rock (5), supplemented by an estimation of the fraction of these that derive from soil rather than from rock.

We begin with the very crude assumption that all material entering human stomachs derives from the top 600 m of rock. We know how much of each chemical element is contained in this rock (5), and we know how much of each chemical element enters human stomachs each year (10,11), so the ratio of these gives a crude estimate of P for each element. An average of these, weighting for the importance of each element in the waste, gives

$$P(\text{very crude}) = 1\ \text{E}{-11}$$

Some of the crudeness of this estimate may be reduced by dropping the implication that all depths of rock down to 600 m are equally likely to contribute and assuming that the contribution from various depths is proportional to the groundwater flow at that depth. Using the exponential decrease in this flow alluded to before, it turns out that the flow at 600 m depth is 16% of the average for the top 600 m, giving an improved, but still rough estimate:

$$P(\text{rough}) = (0.16 \times 1\ \text{E}{-11}) = 1.6\ \text{E}{-12}$$

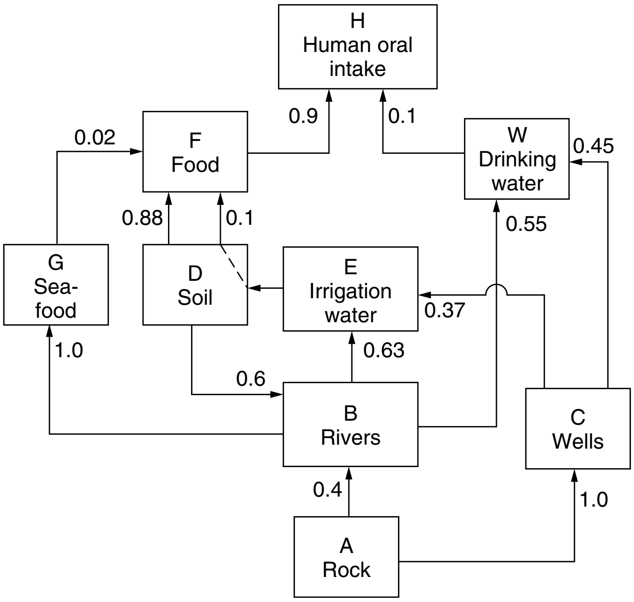

Figure 1. Pathway diagram for transmission of materials from rock and soil into food and water ingested by people. Numbers attached to each line show the fraction of the input to the target box that is derived from that line.

This result is still an overestimate because most of the material entering human stomachs derives from the top layers of soil rather than from the underlying rock. To analyze this problem, a pathway analysis was developed (9) as shown in Fig. 1. The numbers attached to the arrows are the fraction of the input into the compartment derived from that arrow. For example, 90% of human oral intake is from food, and 10% is from drinking water; 2% of the minerals in food comes from seafood (derived from rivers), 88% comes directly from soil, and 10% comes from soil via irrigation water (which is derived 63% from rivers and 37% from wells). Using Fig. 1, the various pathways from rock to human stomachs and their fractional contributions to human intake are then calculated as

$$\begin{array}{lr}
\text{ABWH} & 0.4 \times 0.55 \times 0.1 = 0.022; \\
\text{ACWH} & 1.0 \times 0.45 \times 0.1 = 0.045; \\
\text{ABEFH} & 0.4 \times 0.63 \times 0.1 \times 0.9 = 0.023; \\
\text{ACEFH} & 1.0 \times 0.37 \times 0.1 \times 0.9 = 0.033; \\
\text{ABGFH} & 0.4 \times 1.0 \times 0.02 \times 0.9 \underline{= 0.007}; \\
 & \text{TOTAL} \ = 0.13
\end{array}$$

Our final value for P is then

$$P(\text{final}) = (0.13 \times 1.6\ \text{E}{-12}) = 2\ \text{E}{-13} \tag{4}$$

This result, Eq. 4, is in substantial agreement with the estimate from Equation 3 derived from essentially independent approaches.

HIGH LEVEL WASTE VERSUS AVERAGE ROCK

This result applies to average rock, but our interest is in applying it to the rock-like material, probably a borosilicate glass, into which the waste has been converted. The principal difference is in leachability by groundwater. Leach rates by distilled water for various rocks and glasses have been measured, and the results were correlated with the Gibbs free energy of hydration (12). It was concluded that the leach rate for borosilicate glass is similar to that for basalt and about three times more rapid than the leach rate for average rock. Applying this to the results in Eqs. 3 and 4 gives the rate of transfer for an atom of buried high level waste into a human stomach as about 1.0 E−12 per year.

Concerns have been raised that the high level waste glass may be much less secure than average rock because it is not in chemical equilibrium with the local rock–ground water regime. However, chemical equilibrium is a surface phenomenon (13,14). Groundwater is saturated with dissolved silica. If a foreign silica-based material such as glass is emplaced, there is initially rapid dissolution, removing a very thin surface layer. But this brings more silica into the water which then becomes supersaturated with silica, resulting in precipitation of silica from the water onto the surface—this would be the most insoluble component of the silica originally in the water. The result is a buildup of highly insoluble silica on the surface. Waste can then be dissolved only by diffusion through this surface layer, a very slow process. Moreover, this leads to further thickening of the surface layer, which further retards the diffusion process. Consequently, the rate of dissolution decreases exponentially, only a tiny fraction of the waste becomes involved in the process, and dissolution proceeds only as for the surrounding rock.

BIBLIOGRAPHY

1. Garrels, R.M. and MacKenzie, F.T. (1971). *Evolution of Sedimentary Rocks*. W.W. Norton Co., New York.
2. Todd, D.K. (1980). *Groundwater Hydrology*. John Wiley & Sons, New York.
3. Ad Hoc Panel on Hydrology. (1962). *Scientific Hydrology, Federal Council for Science and Technology*. Washington, DC.
4. White, D.E., Hein, J.D., and Waring, G.A. (1963). *Chemical Composition of Sub-Surface Waters, Chap. F in Data of Geochemistry*, 6th Edn. M. Fleischer (Ed.). U.S. Government Printing Office, Washington, DC.
5. Bowen, H.J.M. (1979). *Environmental Chemistry of the Elements*. Academic Press, New York.
6. Cohen, B.L. (1985). A simple probabilistic risk analysis for high level waste repositories. *Nuclear Technology* **68**: 73–76.
7. Byrne, F. (1974). *Earth and Man*. Wm. C. Brown Co., Dubuque, IA.
8. Cohen, B.L. (1984). Probability for human intake of an atom randomly released into ground, rivers, oceans, and air. *Health Physics* **47**: 281–292.
9. Cohen, B.L. (1986). A generic probabilistic risk analysis for a high level waste repository. *Health Physics* **51**: 519–528.
10. International Commission on Radiological Protection. (1975). *Report of the Task Group on Reference Man*. ICRP Publication 23, Pergamon Press, New York.
11. Pennington, J.A.T. et al. (1984). Selected minerals in food survey, 1974–1982. *Jour. Am. Dietetic Assn.* **84**: 771ff.
12. Plodinek, M.J., Jantzen, C.M., and Wicks, G.C.. (1983). *Stability of Radioactive Waste Glasses Assessed from Hydration Thermodynamics*. Savannah River Laboratory Reports DP-MS-83-61, DP-MS-83-21, and DP-MS-83-66.
13. Cohen, B.L. (1980). Analysis, critique, and evaluation of high level waste water intrusion scenario studies. *Nuclear Technology* **48**: 63ff.
14. Cohen, B.L. (1985). Critique of the National Academy of Sciences study of the isolation system for geologic disposal of radioactive waste. *Nuclear Technology* **20**: 433ff.

GROUNDWATER CONTAMINATION FROM RUNOFF

NITISH PRIYADARSHI
Ranchi University
Ranchi, Jharkhand, India

INTRODUCTION

Groundwater is water that lies below the soil surface and fills the pore spaces in and around rock, sand, gravel, and other materials. Groundwater, under most conditions, is safer and more reliable for use than surface water because surface water is more readily exposed to pollutants from factories, for example, than groundwater. This by no means says that groundwater is invulnerable to contamination. Once groundwater is contaminated, it is extremely costly to remove the contaminant. Any chemicals that are easily soluble and penetrate the soil are prime candidates as groundwater pollutants.

HOW MUCH DO WE DEPEND ON GROUNDWATER?

According to 1985 U.S. figures, groundwater provides an estimated

- 22% of all freshwater withdrawals
- 53% of drinking water for the total population and 97% of drinking water for the rural population
- 40% of public water supply withdrawals
- 46% of domestic and commercial use
- 24% of industrial and mining use
- 34% of agricultural use (mostly for irrigation)

CAUSES OF URBAN RUNOFF

Urban runoff can be attributed to many things, including the amount of rainfall, the soil conditions, and the degree of urbanization. Rainwater can go in many directions once it has reached the earth's surface. Rainwater can be absorbed by the soil on the land surface; absorbed by surrounding vegetation; directly deposited into oceans, streams, and rivers; and infiltrate through the surface and subsurface soils into the groundwater.

Urban runoff can happen anytime of the year when excessive water use from irrigation, car washing, and other sources carries litter, lawn clippings, and other urban pollutants into storm drains.

The majority of people live in cities that are becoming larger and larger and are sprawling into vast suburban neighborhoods. Roofs, parking lots, streets, and other impervious surfaces of an urban environment cause rainwater to collect and be forced out through a storm drain system. If the drainage system does not connect to a wastewater treatment facility, then the rainwater and everything that is carried with it travels into groundwater, local streams, and rivers.

GROUNDWATER CONTAMINATION FROM URBAN RUNOFF

Not too long ago, urban runoff was considered an insignificant contributor to groundwater contamination. Urban runoff is now recognized as a significant source of contamination in water. Effluents from urban areas contain large concentration of oils, greases, nutrients, heavy metals, and detergents. Detergents that are soluble can pass through the soil and pollute groundwater. Raw sewage dumped in shallow soakpits and seepage from polluted lakes, ponds, and streams also pollute groundwater. Rainfall could pick up substantial contaminants from dust and air and join the aquifer below. The infiltration of liquids containing toxic pollutants may cause pollution in sandy soils and well waters.

Some of the contaminants expected in urban runoff are shown in Table 1. Groundwater moves through rocks and subsurface soil, so it has lots of opportunity to dissolve substances as it moves. For that reason, groundwater often has more dissolved substances than surface water. Even though the ground is an excellent mechanism for filtering out particulate matter, such as leaves, soil, and bugs, dissolved chemicals and gases can still occur in large enough concentrations in groundwater to cause problems. Underground water can be contaminated through urban runoff containing industrial, domestic, and agricultural chemical wastes, which includes chemicals such as pesticides and herbicides that many homeowners apply to their lawns. Contamination of groundwater by road salt is of major concern in northern areas of the United States. Salt is spread on the roads to melt ice, and salt is so soluble in water that excess sodium and chloride is transported through runoff after ice melts and finally contaminates subsurface groundwater.

In the United States today, there may be more than 20,000 known abandoned and uncontrolled hazardous waste sites, and numbers grow every year. If there is a leak, these contaminants may be carried away with runoff and finally make their way down through the soil and into the groundwater.

Landfills are another major source of contamination. Landfills are the places where our garbage is taken to be buried. Landfills are supposed to have a protective bottom layer to prevent contaminants from getting into the water. After rainfall, contaminants from the landfill (car battery acid, paint, household cleaners, etc.) are washed away with surface runoff and make their way down into the groundwater. The problem of pollution from landfills is greatest where high rainfall and shallow water tables occur. Important pollutants frequently found in leachate include BOD, COD, iron, manganese, chloride, nitrate, hardness, and trace elements. Hardness, alkalinity, and total dissolved solids are often increased.

Chemicals include products used on lawns and farm fields to fertilize plants. The primary fertilizers are compounds of nitrogen, phosphorus, and potassium. Phosphate and potassium fertilizers are readily adsorbed on soil particles and seldom constitute a pollution problem. But nitrogen in solution is only partially used by plants or adsorbed by the soils, and it is the primary fertilizer pollutant (1). When the rain comes, this chemical is washed away with the surface runoff and eventually goes into the groundwater.

Animal wastes are confined within a limited area, in beef or milk production, so large amounts of waste are deposited on the ground. Thus, for the 120 to 150 days that a beef animal remains in a feedlot, it produces more than a half ton of manure on a dry weight basis. When thousands of animals are in a single feedlot, the natural assimilative capacity of the soil becomes overtaxed. Surface runoff in contact with the manure carries highly concentrated pollutants to subsurface waters. Animal wastes may transport salts, organic loads, and bacteria into the water and soil. Nitrate-nitrogen is the most important persistent pollutant that may reach the groundwater (2).

Sanitary systems have a significant impact on shallow aquifers in urban areas. Field observations revealed that contamination of groundwater is caused by infiltration of surface runoff polluted by municipal waste and sewage and/or leakage of sanitary effluents. It has increased the salt content of groundwater, particularly nitrates

Table 1. The Contaminants that may be Expected in Runoff[a]

1. Asbestos from brake and clutch linings
2. Bacteria and viruses from animals and birds, soils, litter, livestock hauling, livestock waste hauling and on-site sewage tanks and fields
3. Bromide from auto exhaust
4. Cadmium from tire fillers and insecticides
5. Chromium from moving engine parts and brake linings
6. Copper from bearing and bushing wear, moving engine parts, brake linings, and radiator repair
7. Cyanide and chloride from deicing road salts
8. Pesticides (fungicides, herbicides, and insecticides) from roadside maintenance
9. Iron from auto bodies, moving engine parts, bridges, guardrails, overpasses, lamp standards, and other structures
10. Manganese from moving engine parts and gasoline additives
11. Nickel from diesel fuel, lubricating oil, bushing wear, brake linings, and asphalt paving
12. Nitrogen from animal wastes, on-site sewage systems, vegetative matter, and fertilizers
13. Petroleum compounds from paving, fuels spills, engine blow-by, lubricant leaks, antifreeze, and hydraulic fluids
14. Phosphorus from animal wastes, on-site sewage systems, vegetative matter, and fertilizers
15. Potassium from fertilizers
16. Sulfate from roadbeds, road salts, and fuels
17. Zinc and lead from tire fillers, motor oil additives, automotive and radiator repairs, grease, and paint manufacturing

[a] *Source*: Ministry of Water, Land and Air Protection, Government of British Columbia.

and chlorides, and has reduced the oxygen content. In addition, introduction of pathogenic bacteria and viruses into groundwater has caused many outbreaks of waterborne disease.

Solid materials are frequently stockpiled near industrial plants and construction sites. These may be the raw materials awaiting use, or they may be solid wastes placed for temporary or permanent storage. Precipitation falling on unsheltered stockpiles may transport heavy metals, salts, and other inorganic and organic constituents as pollutants to the groundwater.

Urban runoff pollutants are many and varied, depending on land uses and pollutant sources in an urban area. Typically, loadings of urban pollutants are greatest from industrial and commercial areas. Although sources of specific pollutants may vary widely in urban areas, motor vehicles are recognized as a major source of pollutants; they contribute oils, greases, hydrocarbons, and toxic metals such as lead, zinc, copper, chromium, and arsenic.

BIBLIOGRAPHY

1. Todd, D.K. (1995). *Groundwater Hydrology*. Wiley, Toronto, Canada, p. 332.
2. Gillham, R.W. and Webber, L.R. (1969). Nitrogen contamination of groundwater by barnyard leachates. *J. Water Pollut. Control Fed.* **41**(10): 1752–1762.

SALINE SEEP

JOHN E. MOORE
USGS (Refired)
Denver, Colorado

Jackson (1) defines saline seep as an intermittent or continuous saline water discharge at or near the soil surface under dryland conditions that reduces or eliminates crop growth. Human induced salinization of land and water resources due to water table rise is as old as the history of human settlement and irrigation (2). Early settlements in the valleys of the Tigris and Euphrates flourished about 4000 B.C.: however, the irrigated farming of wheat resulted in salt accumulation, and in time the area was abandoned. Historic evidence from Russia, China, India, Pakistan, South American, United States, Canada, and Australia show that saline seeps are a worldwide problem. The number of saline seeps will probably increase in the future because an increase in world population will be accompanied by increases in irrigation.

Saline seeps result from a combination of geologic, climatic, and management factors. Development of a seep starts at the recharge area. As the water moves through the aquifer, the water dissolves and accumulates salts. Saline seeps are formed by mobilization of salt stored in the soil profile. The very act of irrigation is the cause for salinity because every irrigation event adds some salt to the soil. Land clearing or irrigation can cause the groundwater to rise. The additional water moves the water table closer to the land surface. When the water table is within 2–3 feet of the surface, capillary action can lift the water to the surface where it evaporates and leaves salts behind (Figs. 1 and 2).

The salt can also move laterally to affect streams. The common soil forming minerals calcium, sodium, magnesium, and potassium form water-soluble salts in the soil.

Techniques for controlling seeps originated in Montana and Alberta (3). The two major methods to control seeps are (1) cut off the recharge water and (2) reclaim the seep area. Installing subsurface tiles to intercept the groundwater flow might cut off recharge water. The second method is to increase water use in the recharge area by changing the type of crop. Alfalfa is an excellent crop substitute because it is a deep-rooted perennial capable of

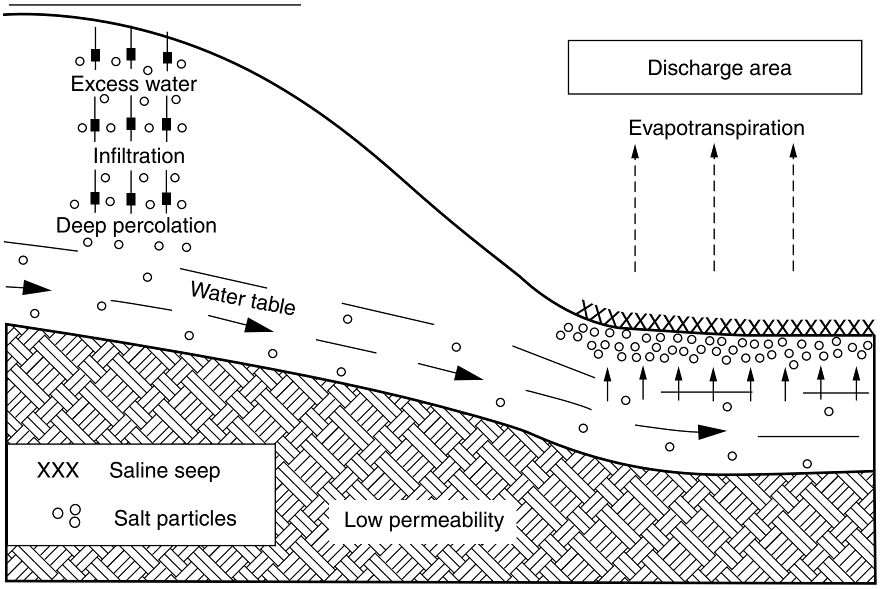

Figure 1. Water movement to seep.

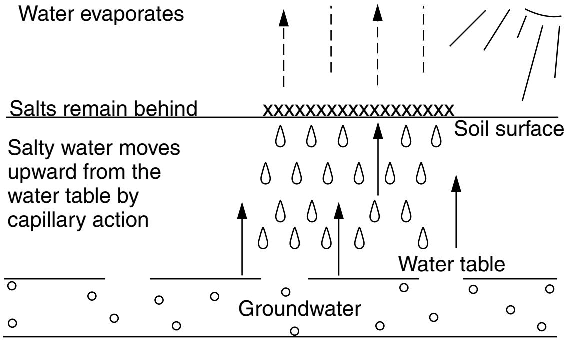

Figure 2. Evaporation and salt concentration.

using large amounts of water. Seeps have been reclaimed within 5 years by planting alfalfa in the recharge area (4).

In southeast Australia, the replacement of deep-rooted native vegetation by dryland and irrigated agriculture has resulted in a major change in surface water and groundwater systems in the Murray–Darling basin. The Murray River has a salinity of less than 25 mg/L in the headwaters and 480 mg/L downstream. Water that was previously used by vegetation for transpiration is now leaking past the vadose zone and entering the water table. The salinization process is closely linked to groundwater changes. In both irrigation and dryland farming, salts are remobilized into the production zone. The salts were originally stored in the aquifer or the vadose zone. Evaporation concentrates the salts in the near surface zone to levels where the plants are. For salinization to occur, it is necessary to have both a hydrologic change and a source of salt.

Saline soils and sodic soils are quite common in parts of western North Dakota and Montana. Most saline seeps have developed recently (postsettlement). The formation of saline seeps is closely related to the practice of summer fallow for moisture conservation (5).

Colorado River has 25 mg/L salt in the headwaters and 825 mg/L downstream at Imperial Dam. In the San Joaquin River in California, salinity increased from 330 mg/L in the 1930s to 600 mg/L in the 1970s. In the former Soviet Union, the salinity of the Syr Darya River which discharges to the Aral Sea increases from 300 mg/L to 2000 mg/L downstream.

The water table rise in agricultural areas of the world has contributed to the salinization of large areas. The development of land and water resources of the earth will be more difficult and more expensive than in the past. The control of population and the efficient use of developed resources will be essential in the future.

BIBLIOGRAPHY

1. Jackson, J. (1997). *Glossary of Geology*. American Geological Institute, p. 769.
2. Ghasemi, F. et al. (1994). *Water table rise and the expansion of land and water resources salinsation: a global view: IAH Congress Proceedings*. Water Down Under, Adelaide, Australia, pp. 607–610.
3. Brown, P.L. et al. (1983). *Saline seep diagnosis control and reclamation*. USDA-ARS conservation Research Report No. 30.
4. Lamon, R.E. and Whitney, D.A. (1992). *Management of Saline and Sodic Soils:MF-1022 Cooperative Extension Service*. Kansas State University, Manhattan, KS.
5. Wentz, D. (1997). *Dryland Saline Seeps-Types and Causes*. Alberta Agriculture.

GROUNDWATER SAMPLING TECHNIQUES FOR ENVIRONMENTAL PROJECTS

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Groundwater sampling techniques are important for accurate project results and understanding aquifers whether the project is sampling a groundwater monitoring well as part of an environmental assessment or regular sampling of a potable water supply well. Detailed groundwater sampling information can be obtained from Testa (1), Jacobs (2), the American Society for Testing Materials (3), the California Regional Water Quality Control Board (4), and Driscoll (5). Although there are numerous differences in groundwater sampling and record keeping techniques, regional variations, and specific regulatory requirements that vary from state to state, a generalized summary of selected environmental groundwater sampling techniques is given below.

PURPOSE OF GROUNDWATER SAMPLES

Groundwater samples are collected for a variety of reasons. Groundwater can be collected as a "grab" sample or from a monitoring well. Groundwater samples are used to define the vertical and aerial extent of groundwater contamination, to monitor target chemical concentrations over time, to provide a measurement for detecting a contaminant plume front, to evalvate unexpected changes in plume size or direction of flow, to determine the extent of interaquifer movement of contaminants, to determine aquifer characteristics such as permeability, transmissivity, etc.), to estimate the rate of contaminant plume movement, to develop a data base for designing remedial measures, to determine the effects of the remedial measures, to assist in performing the remedial work (provide hydraulic control and contaminant removal), to provide data base for groundwater modeling, and to evaluate the aquifer during regular sampling over the yearly hydrologic cycle (6).

GRAB GROUNDWATER SAMPLING

A "grab" groundwater sample is collected without the use of a fully installed groundwater monitoring well. Methods include using a bailer or jar to collect groundwater from an excavation or trench made with a backhoe or excavator. Other "grab" samples can be collected using a drilling rig wherein a sampler is drilled or pushed to the target depth and the sampler is pushed a bit further into the aquifer. Water is allowed to collect in the sampler. The water is removed from the sampler using a bailer or by using a peristaltic pump and tubing, check-ball valve with tubing or other similar methods. These methods, frequently called punch methods, are commonly used with hollow stem auger rigs or direct push technology rigs. Cone penetrometer (CPT) rigs can use specialized groundwater containers that are inserted into a cavity of the water sampler tip. The groundwater containers are under a vacuum. At the proper depth, the sampler is opened, the groundwater container is pierced, and water enters the container.

Temporary wells are used occasionally in environmental investigations. A temporary well consists of a well screen below a casing riser placed in an open borehole. Several well volumes can be purged prior to sampling. A bailer or other method is used to bring the water to the surface. After groundwater sampling is completed, the borehole is usually tremie grouted to seal the annulus.

WATER WELLS

Water wells are installed using a variety of rigs: hollow stem auger, mud or air-rotary, air percussion, diamond coring, cable tool, and others. After well installation, wells are developed within a day to several days of installation to remove residual drilling materials from the well bore and to improve well performance by removing any fine material in the filter pack that can pass from the native soil into the well. Well development techniques include pumping, bailing, surging, jetting, and airlifting. In most cases, surging and pumping are satisfactory. Development water is inspected for turbidity, product sheen, odors, and sediment. A minimum of several casing volumes is removed during development. The well is considered fully developed when consistent pH, temperature, and conductivity readings indicate characteristic groundwater for the aquifer. If the aquifer is slow to recharge, development will continue until the well is pumped dry. Cross contamination of wells from pumps is avoided by using dedicated equipment and proper decontamination procedures.

Due to the large volumes of purge water generated during groundwater sampling and the costs for storage and disposal of the water, some regulatory agencies are now allowing various low-flow groundwater sampling techniques for compliance groundwater monitoring. Proponents of micropurging suggest that evacuation of a specific volume of water may increase turbidity from agitation or mixing of the well water column. Larger purge volume may increase the risk of colloid mobilization. Proper well development immediately after well construction is essential when using low-flow sampling (7).

Micropurging is the evacuation of tubing, sampling devices, and other items from the well prior to sampling groundwater. There is no purging prior to sample collection with micropurging, so there is no method for determining if the well contains stagnant well water or water representative of the aquifer. For this reason, many regulatory agencies do not approve micropurging. Instead, when minimal purging is requested, low-flow sampling is preferred.

The low-flow sampling method, typically used in 2-inch diameter groundwater monitoring wells, relates to the amount of drawdown in a well during purging. The indicator parameters are collected before and during purging and allowed to stabilize prior to groundwater sample collection. Purge rates may be higher than sample rates to maximize purge efficiency. Prior to the collection of the groundwater sample, the pump rate may be lowered to lower sample turbidity and entrained air in the sample, and to mimic conditions that may exist in the natural aquifer. High turbidity in water samples being analyzed for metals requiores filtering to remove colloidaland suspended sediments that may contain metals.

Discrete sampling systems are designed to collect groundwater samples form the formation at a specific depth or multiple target depths without significant purging prior to sample collection.

GROUNDWATER SAMPLING PROTOCOL

Depth to the top of the water and free-product thickness within wells are monitored by using a water level meter or an interface probe. Water level data collected from areas that have multiple wells are used to develop groundwater contour maps for the project site. Groundwater flow is perpendicular to equipotential lines drawn on the map. If free-product is detected, a product sample is sometimes collected for source identification. In environmental projects, where several chemicals are to be analyzed for a given well, individual samples are collected in order of decreasing volatility. When the results from previous sampling are known, it is recommended to start sampling the cleanest wells first, moving to the most contaminated wells later in the sampling event.

Each chemical analysis requires specific sample preservation techniques or additives. Sealed chemical ice is placed in the coolers to maintain samples at a temperature of 4 °C. Where several types of chemicals are to be analyzed for a given well, individual samples are collected in the following order: volatile organic compounds, purgeable organic compounds, purgeable organic halogens, total organic compounds, total organic halogens, extractable organic compounds, total metals, dissolved metals, phenols, and cyanide.

QUALITY CONTROL OF GROUNDWATER SAMPLES

A QC program, which is independent of the laboratory's program, verifies sample integrity and includes the submittal of duplicates, field blanks, and travel blanks to the laboratory. The QC samples are packaged and sealed in the same manner as the other samples and are assigned independent sample numbers.

A duplicate sample is collected for about 5% of the samples or one per sampling round, whichever is greater. The duplicate sample is acquired by filling separate containers from the same well bailer as the actual sample. The contents of the bailer are evenly divided between the actual and duplicate samples to ensure duplication. The field blank is submitted to the laboratory for the same analyses as the rest of the sampling set. The field blank is acquired by dispensing deionized water from a clean sampling bailer into the containers in the same manner as groundwater samples. When sampling groundwater for analyzing volatile compounds, travel blanks are used to detect the introduction of contaminants during transportation from the field to the laboratory. The travel blank generally contains deionized water, which has been poured into the container prior to being taken into the field.

BIBLIOGRAPHY

1. Testa, S.M. (1994). *Geologic Aspects of Hazardous Waste Management*. CRC Press, Boca Raton, FL, pp. 145–187.
2. Jacobs, J. (2000). Monitoring well construction and sampling techniques. In: *Standard Handbook of Environmental Science, Health, and Technology*. J. Lehr (Ed.). McGraw-Hill, New York, pp. 11.46–11.68.
3. American Society for Testing Materials. (1984). Standard Practice for Description and Identification of Soils (Visual-Manual Procedure), Method D 2488-84. Philadelphia, PA.
4. California Regional Water Quality Control Board. (1989). Leaking Underground Fuel Tank (LUFT) Field Manual: Guidelines for Site Assessment, Cleanup, and Underground Storage Tank Closure, Sacramento, CA.
5. Driscoll, F.G. (1986). *Groundwater and Wells*, 2nd Edn. Johnson Filtration Systems, St. Paul, MN, pp. 268–339.
6. Sisk, S.J. (1981). *NEIC Manual for Groundwater/Subsurface Investigations at Hazardous Waste Sites*. EPA-330/9-81-002. U.S. Environmental Protection Agency, Denver, CO.
7. Puls, R.W. and Barcelona, M.J. (1996). Low-flow (minimal drawdown) ground water sampling procedures. EPA Groundwater Issue, April.

GROUNDWATER SAMPLING WITH PASSIVE DIFFUSION SAMPLERS

Craig E. Divine
Colorado School of Mines
Golden, Colorado

INTRODUCTION: DIFFUSION-BASED GROUNDWATER SAMPLING METHODS

Passive diffusion-based sampling methods have been used to monitor environmental contaminants, particularly trace metals and certain lipophilic compounds including polychlorinated biphenyls (PCBs), polynuclear aromatic hydrocarbons (PAHs), and pesticides, for many years. In particular, dosimeter-type and semipermeable membrane devices (SPMDs) are commonly used in limnology, oceanography, and ecotoxicology for evaluating ultralow concentrations and potential bioavailability of contaminants (1–4). These types of samplers are most frequently used for sampling sediment porewater and surface water (i.e., rivers, estuaries, harbors) and have occasionally been used for groundwater characterization in contaminant toxicology studies (5). These types of sampling devices are composed of a "trapping" phase, such as hexane, activated carbon, or triolein (a prominent fish lipid), contained within a semipermeable membrane (such as dialysis-type cellulose or a synthetic polymer). On deployment, contaminants diffuse through the membrane and are subsequently retained in the trapping phase until analysis. After the SPMD is removed, the contaminants must typically be extracted from the trapping phase before analysis. A significant disadvantage with SPMDs and dosimeter-type devices for groundwater sampling is that back-calculation to actual *in situ* aqueous concentrations may not be straightforward due to nonlinear and/or nonequilibrium sorption behavior of the trapping phase (5).

Passive diffusion (PD) samplers are similar to dosimeter-type and SPMD sampling methods in that they generally consist of a sealed container with a semipermeable membrane. The PD sampler is suspended in a monitoring well, and the solutes of interest diffuse across the membrane into the sampler. However, unlike dosimeter-type devices and SPMDs, PD samplers do not contain a trapping phase but are instead filled with water or air. Therefore, once the PD sampler is retrieved from the well (after equilibration), the concentration measured in the sampler is assumed to relate directly to *in situ* analyte concentrations at the deployment location. Passive diffusion samplers are less expensive to construct than SPMDs and easier to analyze because no extraction step is required. Furthermore, the correlation between the contaminant concentration measured in the PD sampler and the surrounding environment is straightforward: for water-filled PD samplers deployed in groundwater the concentrations are equivalent.

Passive diffusion-type sampling methods have been used in a variety of applications for many years, including dissolved gas measurement in seawater and groundwater (3,6,7). Generally, PD samplers offer several potential advantages for monitoring many common groundwater contaminants, compared to conventional sampling techniques. For sites with ongoing monitoring programs, significant cost savings may be realized due to reduced sampling-related labor and minimal sampling-related waste generation. For dissolved constituents with large Henry's Law constants, such as volatile organic compounds (VOCs) and dissolved gases, analyte loss due to volatilization can be greatly minimized. In some cases, analytical matrix interferences may be reduced if the membrane is selective against interfering constituents (i.e., the matrix in the PD sampler may be comparatively less complex). Additionally, multiple PD samplers can be deployed to provide solute concentration measurements across very discrete intervals. For example, PD samplers with dialysis-type membranes have been used to characterize ion and trace metal concentration profiles on the scale of a few centimeters in sediment (3).

MEMBRANE MATERIALS AND SPECIFIC CONSTRUCTION EXAMPLES

Because of the potentially significant costs savings and technical advantages noted above, the use of PD samplers for groundwater monitoring at contaminated sites is greatly increasing [particularly water-filled PD samplers using low-density polyethylene (LDPE) membranes].

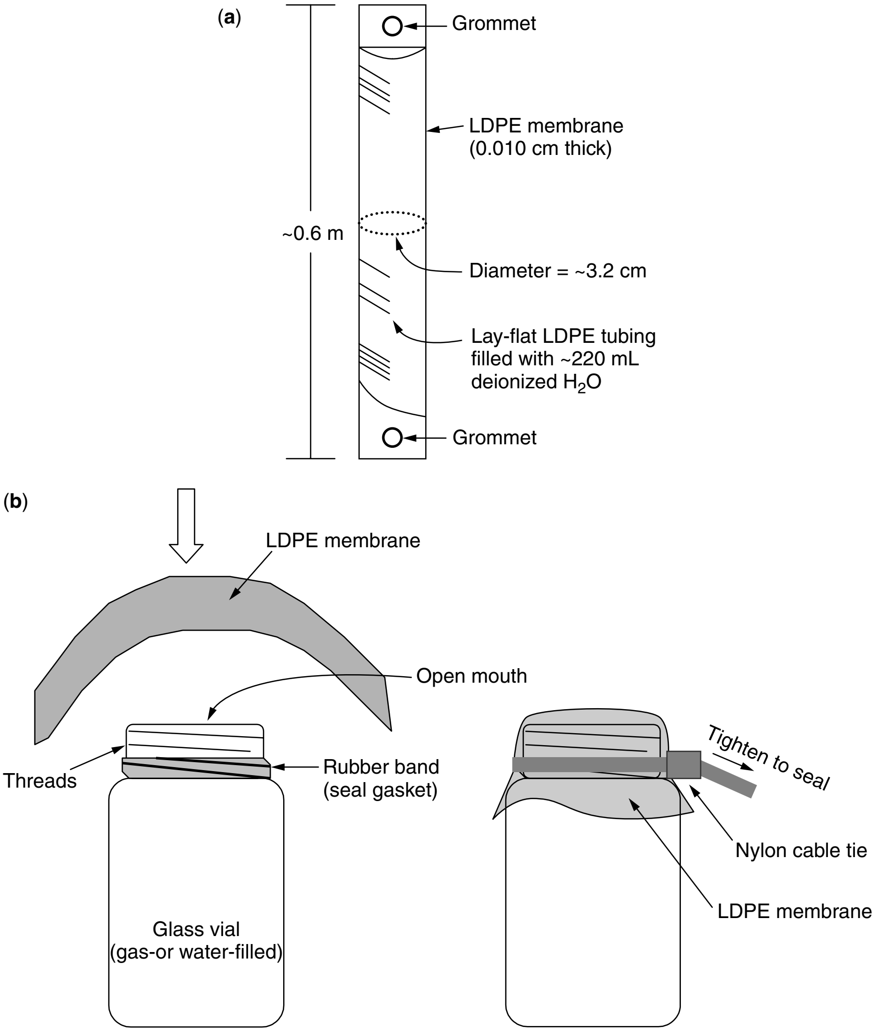

Figure 1. Schematics of passive diffusion samplers.

Laboratory and field compatibility tests indicate that LDPE membranes are suitable for most fuel and solvent VOCs; however, it has been demonstrated that the LDPE membrane is generally not suitable for measuring inorganic ions, metals, pesticides, explosives, semivolatile organics, alcohols, organic acids, and ethers [including methyl-*t*-butyl ether (MtBE)] with PD samplers (8,9). Dissolved gases, such as He, Ne, H_2, O_2, and N_2, can be measured by PD samplers with LDPE (10) or silicone (7) membranes. Passive diffusion-type samplers (also known as dialysis samplers or "peepers") that use other membrane materials, such as cellulose acetate, can be used to monitor major ions, metals, nutrients, dissolved gases, and pH (3,11).

Passive diffusion samplers can be constructed easily from inexpensive materials. As shown in Fig. 1a, the open mouth of a glass laboratory vial can be covered with a membrane and sealed with a rubber band gasket and nylon zip tie. Drop sheeting, which is widely available in the painting supply sections of hardware stores, can be used for LDPE membranes and is offered in a range of thicknesses. PD samplers filled with water can be constructed while submerged in deionized water to prevent entrapment of air bubbles. Passive diffusion samplers made of lay-flat LDPE tubing filled with deionized water are commercially available from EON Products, Inc. (Snellville, Georgia) and Columbia Analytical, Inc. (Kelso, Washington). A schematic of this type of sampler, which is also known as a PD "bag" sampler is shown in Fig. 1b. For a description of peepers and other PD samplers with dialysis-type membranes, see Davison et al. (3).

EQUILIBRATION TIME FOR PASSIVE DIFFUSION SAMPLERS

To prevent biased-low measurements, PD samplers must be deployed long enough in a well for equilibration to occur. Based on empirical laboratory and field studies, a minimum of 2–3 days is required for equilibration for most VOCs using commercially available LDPE PD bag samplers (8). Carnigan (12) found that some dialysis-type samplers require equilibration times of 3–20 days for measuring inorganics in porewater. Based on these general results, a minimum equilibration time of approximately 2 weeks is commonly recommended as a conservative approach in field studies (8,9). However, note that this is a generic rule of thumb. As equilibration is a diffusion process, the specific equilibration time depends on sampler volume, membrane surface area, membrane thickness, and the diffusion coefficient of the solute in the membrane material.

Sanford et al. (2) present the following solution that describes the concentration within the PD sampler as it equilibrates with the surrounding groundwater:

$$C_{\mathrm{s}}(t) = C_{\mathrm{r}}H\left[1 - e^{\left(\frac{-D_{\mathrm{m}}At}{\forall L_{\mathrm{m}}}\right)}\right] \tag{1}$$

where

$C_{\mathrm{s}}(t)$ concentration within the sampler at time t after PD sampler deployment [M L^{-3}],

C_{r} resident concentration at the measurement location [M L^{-3}],

D_{m} effective constituent diffusion coefficient for the membrane material [L^2 T^{-1}],

H solute dimensionless Henry's law constant [ ],

A surface area of the PD sampler membrane [L^2],

t deployment time [T],

$\forall$ internal volume of the PD sample [L^3], and

L_{m} membrane thickness [L].

This solution assumes Fickian solute diffusion across a thin membrane, a constant C_{r} concentration at the measurement location, and a well-mixed reservoir in the PD sampler. As presented, Equation 1 incorporates air–water partitioning and is applicable to gas-filled PD samplers deployed in groundwater, such as those described by Sanford et al. (7) and Divine and McCray (10). For conditions where no phase partitioning occurs (such as a water-filled PD sampler deployed in groundwater), Henry's law constant (i.e., the gas-water partition coefficient) is treated as unity.

As clearly shown in Eq. 1, the rate at which a PD sampler equilibrates relates directly to the D_{m} value of the solute, the sampler dimensions ($A, \forall$), and the thickness of the membrane (L_{m}). Therefore, PD samplers can be constructed to minimize equilibration times by using a thin membrane and increasing the $A/\forall$ ratio. The effective D_{m} value is temperature-dependent and specific to both the membrane and solute. Divine and McCray (10) estimated that a practical range is $1{-}4 \times 10^{-7}$ cm^2 s^{-1} for D_{m} values of several dissolved gases and VOCs for LDPE membranes. Sanford et al. (2) measured D_{m} values for helium and neon of 4×10^{-7} cm^2 s^{-1} for silicone membranes. Based on these D_{m} ranges, equilibration times are predicted using Eq. 1 for several PD samplers, including both constructed (Fig. 1a) and commercially available (Fig. 1b) samplers. The specific parameter values and resulting equilibration times are summarized in Table 1, and the predicted PD sampler concentration histories are shown in Fig. 2.

These results show the relative importance of the various parameters for equilibration times. For example, increasing the $A/\forall$ ratio reduces the equilibration time, and this can be accomplished for constructed PD samplers (Fig. 1a) by filling the sampler with inert sand or glass beads (i.e., compare results of Scenario 3 with Scenario 4). It is also important to note that Eq. 1 is valid only for advection-dominated systems. If the groundwater velocity is very low (diffusion-dominated systems), a concentration gradient will develop outside of the sampler, and equilibration times can be much longer (13). Equation 1 is appropriate, however, for most shallow aquifer systems.

CASE STUDY: APPLICATION OF PASSIVE DIFFUSIONS SAMPLERS AT A HETEROGENEOUS SITE

Passive diffusion samplers measure water quality directly in contact with the membrane surface; therefore, they generally measure concentrations at relatively discrete intervals within a well. Conversely, samples collected by conventional purge-and-bail methods tend to integrate water quality across the entire well screen region and therefore are most significantly influenced by water quality conditions in the highest permeable zones across the well screen interval. In other words, PD samplers tend to measure porewater concentrations over relatively small intervals, and purge-and-bail type methods tend to measure flux-weighted concentrations. For vertically uniform plume distributions and homogeneous aquifer

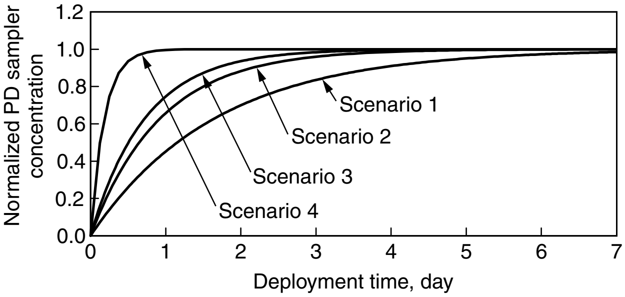

Figure 2. Passive diffusion sampler concentrations after deployment, predicted by Equation 1. The normalized concentration is defined as $C_{\mathrm{s}}/C_{\mathrm{r}}$ (at equilibrium $C_{\mathrm{s}} = C_{\mathrm{r}}$; therefore, $C_{\mathrm{s}}/C_{\mathrm{r}} = 1.0$). The parameter values used in the models are summarized in Table 1.

Table 1. Passive Diffusion Sampler Dimensions, Membrane Thickness, Membrane Diffusion Coefficients, and Predicted Equilibration Times for Various Constructed (Fig. 1a) and Commercially Available (Fig. 1b) PD Samplers. The Equilibration Time is Defined by Equation 1 when $C_{\mathrm{s}} \geq 0.95C_{\mathrm{r}}$)

	A, cm^2	$\forall$, mL	L_{m}, cm	D_{m}, (cm^2 s^{-1})	Equilibration Time, Day
Scenario 1 *Commercial "bag" sampler*	350	220	0.010	4.0×10^{-7}	0.7
Scenario 2 *Commercial "bag" sampler*	350	220	0.010	2.5×10^{-7}	2.7
Scenario 3 *Constructed sampler: sand-filled 20-mL glass vial*	3.2	12.9	0.005	2.5×10^{-7}	2.8
Scenario 4 *Constructed sampler: 20-mL glass vial*	3.2	23	0.005	2.5×10^{-7}	5.1

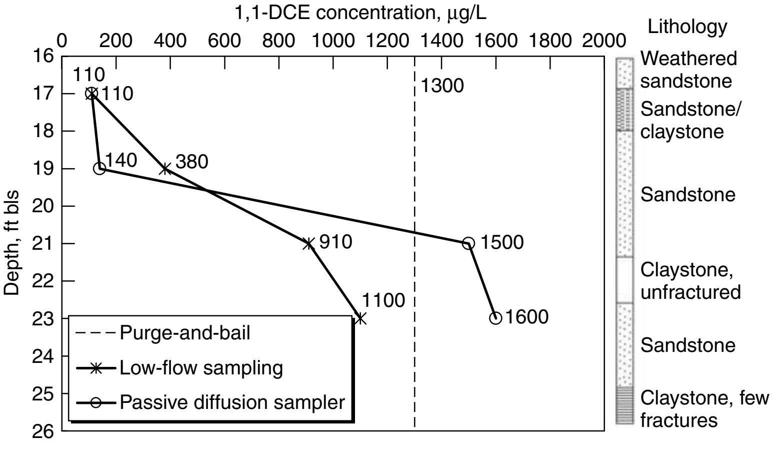

Figure 3. Comparison of sampling results for purge-and-bail, low-flow, and passive diffusion sampling methods for a single monitoring well.

conditions, results from these two sampling methods may be equivalent. However, these two methods may yield significantly different concentration values for a specific well at a heterogeneous site with plume stratification.

To illustrate this potential behavior, the concentrations of 1,1-dichloroethene (1,1-DCE) (a common VOC) measured at a monitoring well by purge-and-bail, low-flow [another discrete-type sampling technique; see Puls and Barcelona (14)]. PD sampling methods are shown in Fig. 3. The aquifer at this study site is highly heterogeneous, and this is clearly expressed in the results of the various sampling methods. Only one value (1300 μg/L) is obtained by the purge-and-bail method; however, the low-flow and PD sampling methods characterize concentrations at various discrete depth intervals and show significant contaminant stratification (approximately an order of magnitude over several vertical feet).

Figure 4 compares the results of purge-and-bail with PD sampling methods at several wells at the same heterogeneous site. The purge-and-bail measurements are generally higher than the average discrete PD sampler measurements, possibly indicating that the majority of the contaminant mass exists in highly permeable zones. These data provide additional characterization information and have potentially important implications for the site. At some sites, a similar targeted dual-method sampling strategy (i.e., use of both purge-and-bail and PD sampling methods) may be useful for evaluating contaminant distribution, remedial system design and performance, and long-term plume behavior.

Clearly, using PD samplers under heterogeneous aquifer conditions, particularly in an ongoing compliance monitoring program, requires careful consideration of ways to interpret measurement data and satisfy monitoring program objectives. For example, determining the appropriate intervals for PD sampler placement may be challenging (i.e., Fig. 3). Additionally, as indicated by the circled values in Fig. 4, the PD sampling method measured low concentrations of some nontarget VOCs at a few discrete intervals in certain wells, but these compounds were not measured above the analytical quantitation limit by the purge-and-bail method. Assuming that these differences are related to the differing averaging natures of the sampling methods, interpreting the discrete interval measurement within site objectives or regulatory requirements may be challenging. For example, an analytical water quality goal or regulatory limit may be met by one method and not by another at the same well simply due to the inherent averaging characteristics of the sampling methods. For additional discussion of the results of this study, see Divine et al. (15).

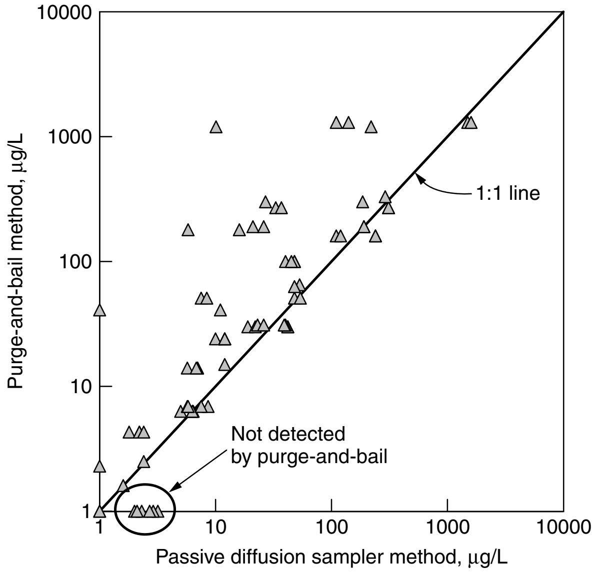

Figure 4. Comparison of sampling results (various VOCs) for purge-and-bail and passive diffusion sampling methods for a series of wells at a heterogeneous site.

BIBLIOGRAPHY

1. Sordergren, A. (1987). Solvent-filled dialysis membranes simulate uptake of pollutants by aquatic organisms. *Environ. Sci. Technol.* **21**(9): 855–859.
2. DiGiano, F.A., Elliot, D., and Leith, D. (1988). Application of passive dosimetery to the detection of trace organic contaminants in water. *Environ. Sci. Technol.* **22**(11): 1365–1367.
3. Davidson, W., Fones, G., Harper, M., Teasdale, P., and Zhang, H. (2000). Dialysis, DET, and DGT: *In situ* diffusional techniques for studying water, sediments, and soils. In: *In Situ Monitoring of Aquatic Systems*. J. Buffle and G. Horvai (Eds.). John Wiley & Sons, New York, Chapter 11.

4. Dunn, R.K., Teasdale, P.R., Warnken, J., and Schleich, R.R. (2003). Evaluation of diffusive gradient in a thin film technique for monitoring trace metal concentration in estuarine waters. *Environ. Sci. Technol.* **37**(12): 2794–2800.
5. Gustavson, K.E. and Harkine, J.M. (2000). Comparison of sampling techniques and evaluation of semipermeable membrane devices (SPMDs) for monitoring polynuclear aromatic hydrocarbons (PAHs) in groundwater. *Environ. Sci. Technol.* **34**(20): 4445–4451.
6. Dyck, W. and Da Silva, F.G. (1981). The use of ping-pong balls and latex tubing for sampling the helium content of lake sediments. *J. Geochem. Exploration* **14**: 41–48.
7. Sanford, W.E., Shropshire, R.G., and Solomon, D.K. (1996). Dissolved gas tracers in groundwater: Simplified injection, sampling, and analysis. *Water Resour. Res.* **32**(6): 1635–1642.
8. Vroblesky, D.A. and Campbell, T.R. (2001). Equilibration times, compound selectivity, and stability of diffusion samplers for collection of ground-water VOC concentrations. *Adv. Environ. Res.* **5**: 1–12.
9. Vroblesky, D.A. (2001). *User's Guide for Polyethylene-Based Passive Diffusion Bag Samplers to Obtain Volatile Organic Compound Concentrations in Wells, Part 1: Deployment, Recovery, Data Interpretation, and Quality Control and Assurance*. United States Geological Survey, Columbia, SC.
10. Divine, C.E. and McCray, J.E. (2004). Estimation of diffusion coefficients and equilibration times for LDPE passive diffusion samplers. *Environ. Sci. Technol.* **38**(6): 1849–1857.
11. Webster, I.T., Teasdale, P.R., and Grigg, N.J. (1998). Theoretical and experimental analysis of peeper equilibration dynamics. *Environ. Sci. Technol.* **32**(11): 1727–1733.
12. Carnigan, R. (1984). Interstitial water sampling by dialysis, methodological notes. *Limnol. Oceanogr.* **29**: 667–670.
13. Harrington, G.A., Cook, P.G., and Robinson, N.I. (2000). Equilibration times of gas-field diffusion samplers in slow-moving ground water systems. *Ground Water Monitoring Remediation* **20**(2): 60.
14. Puls, R.W. and Barcelona, M.J. (1996). *Low-Flow (minimal Drawdown) Ground-Water Sampling Procedures*. EPA/540/S-95/504, Ada, OK.
15. Divine, C.E., Madsen, L., Andrews, S.D., and Santangelo-Drieling, T. (in press). Passive diffusion ground water samplers at a heterogeneous site: Pilot study results. *Ground Water Monitoring and Remediation*.

SPECIFIC CAPACITY

Eric S. Wilson
E. L. Montgomery & Associates, Inc.
Tucson, Arizona

The specific capacity of a well is the pumping rate of the well divided by the drawdown at that rate:

$$S_c = \frac{Q}{\Delta h}$$

where S_c=specific capacity
Q=constant pumping rate, and
Δh=drawdown

Water well contractors often report specific capacity to characterize the performance of a well and use specific capacity to estimate the optimum depth for installing a pump to obtain the desired yield. Units typically associated with specific capacity are liters per second per meter of drawdown (l/s/m), cubic meters per day per meter of drawdown (m^3/day/m), cubic feet per day per foot of drawdown (ft^3/day/ft), or gallons per minute per foot of drawdown (gpm/ft).

When taking measurements to compute specific capacity, the pumping rate should be constant, and the period of pumping should be sufficiently long that the rate of change in drawdown is small. The specific capacity of a well tends to become smaller as the pumping rate and duration of pumping increase. At high pumping rates, nonlaminar flow losses are larger and result in an increased rate of drawdown. Water level drawdown in unconfined aquifers results in reduced aquifer transmissivity near pumping wells; larger drawdown at high pumping rates then results in progressively smaller transmissivity and smaller specific capacity. A specific capacity value remarkably smaller than expected may result when the well is not screened in the correct water-bearing zone or when well screen openings are clogged due to incomplete well development, precipitation of minerals in the screen openings, or biofouling. After many years of pumping, if the specific capacity declines to about half the original value, well rehabilitation should be considered.

Specific capacity may be used to estimate the maximum potential pumping rate of a well by multiplying the available drawdown by specific capacity. The available drawdown is the difference between the depth at which the pump is set and the depth to the nonpumping water level in the well.

Specific capacity may be used to estimate the transmissivity of aquifer media in the vicinity of the pumping well. Theis (1) suggested that specific capacity could be used to estimate the transmissivity of an aquifer using the equation,

$$T^* = \frac{Q}{\Delta h}\frac{2.3}{4\pi}\log\frac{2.25T^*t}{r^2S}$$

where $\frac{Q}{\Delta h}$=specific capacity
t=time
r=radius of pumping well
S=aquifer storativity
T^*=transmissivity; an initial guess of the value of T is made and substituted in the equation. The equation is then solved for Δh. After the initial solution, the value of T must then be adjusted until the calculated value of Δh is reasonably close to the measured value.

Driscoll (2) indicated that transmissivity could be estimated as

$$T = S_c \times 2,000 \text{ for confined aquifers}$$

or

$$T = S_c \times 1,500 \text{ for unconfined aquifers}$$

where T is in gpd/ft and
S_c is specific capacity.

Razack and Huntley (3) gave an empirical relationship between transmissivity and specific capacity as follows:

$$T = 15.3(S_c)^{0.67}$$

where T is in m^2/day
S_c in m^3/day/m,

or

$$T = 33.6(S_c)^{0.67}$$

where T is in ft^2/day
S_c in ft^3/day/ft.

Using specific capacity data to estimate transmissivity can be useful when trying to estimate potential pumping rates or estimating how much available head is needed to produce a specified pumping rate. But, as in any estimation technique, it is not always applicable or advisable. These instances may include situations where substantial drawdown occurs because the pumping duration is too short for the specific capacity test, too small a pumping rate is used for an adequate specific capacity test, or well construction is inefficient.

BIBLIOGRAPHY

1. Theis, C.V. (1963). Estimating the Transmissibility of Aquifers from the Specific Capacity of Wells. In: *Methods of Determining Permeability, Transmissibility, and Drawdown*. R. Bentall, compiler, U.S. Geological Survey Water-Supply Paper 1536-I, pp. 331–340.
2. Driscoll, F.G. (1986). *Groundwater and Wells*, 2nd Edn. Johnson Filtration Systems, Inc., St. Paul, MN.
3. Razack, M. and Huntley, D. (1991). Assessing transmissivity from specific capacity in a large and heterogeneous alluvial aquifer. *J. Ground Water* **29**(6): 856–861.

SOIL WATER

Todd Rasmussen
The University of Georgia
Athens, Georgia

Soil water is a measure of the quantity of water present within a unit volume of soil. Soil-water content can vary from fully saturated, in which case the water content equals the soil porosity, to oven-dry, in which case all free water has been evaporated.

Wet soils are characterized as having high water contents, which means that more soil pores are filled with water, whereas dry soils have more pores that are filled with air.

The volumetric water content, θ, is defined as the volume of water, V_w, per unit volume of soil, V_s:

$$\theta = \frac{V_w}{V_s} \tag{1}$$

where the range in water content is from a minimum, or residual, saturation, θ_r, to a maximum equal to the saturated porosity, θ_s, so that $\theta_r < \theta < \theta_s$.

The water content can also be defined with a gravimetric, or mass, measurement, θ_g, such that

$$\theta_g = \frac{M_w}{M_s} \tag{2}$$

where M_w is the water mass and M_s is the soil mass. We can readily convert between these two measures by using:

$$\theta = \theta_g \frac{\rho_w}{\rho_s} = \frac{\theta_g}{\rho_s} \tag{3}$$

where $\rho_w = M_w/V_w = 1$ kg/L is the water density and $\rho_s = M_s/V_s$ (kg/L) is the bulk density of the soil.

The soil water content can also be reported as a relative saturation, Θ, which is defined using

$$\Theta = \frac{\theta - \theta_r}{\theta_s - \theta_r} \tag{4}$$

where θ_r and θ_s are the residual and saturated water contents, respectively. The relative saturation varies from zero for a dry soil to a maximum of one for a fully saturated soil, $0 < \Theta < 1$.

MEASURING WATER CONTENTS

Water contents are commonly measured with time-domain reflectometry (TDR) devices, as shown in (Fig. 1). TDRs determine the water content with the electromagnetic properties of a wave pulse passing through a conducting set of rods (such as 3-mm stainless-steel welding rods) placed in the soil.

The principle of TDRs is that the velocity of electromagnetic waves along a conductor is a function of the *dielectric* coefficient of the media around the conductor. Liquid water has a dielectric constant of 80.2 at 20 °C, whereas ice is 3.2, petroleum is 1.8 to 2.2, quartz is 4.3, and air is only 1.00.

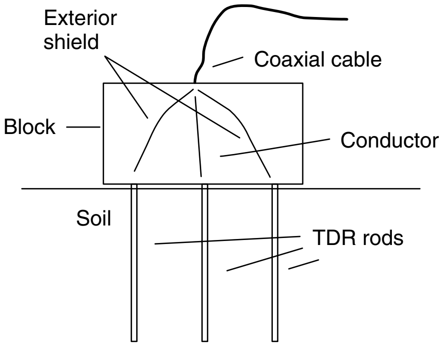

Figure 1. Time-domain reflectometer (TDR) probe that measures soil-water content. The three-rod assembly is placed within the soil. Coaxial cable connects probe to the water content tester. Center wire of coaxial cable connects to the central rod, whereas the exterior shield is connected to the two outer rods.

As the water dielectric constant is so much greater for water than for soil or air, the wave velocity is substantially retarded as the water content of a soil increases, because larger dielectric constants cause slower wave velocities and, hence, longer travel times.

Before the use of TDRs, neutron probes used the thermalization of neutrons to determine the water content (1). This older method has fallen out of favor, however, because of the rigorous licensing and inspection requirements associated with radioactive materials.

MATRIC TENSION

Water molecules in the unsaturated zone are bound by *adhesive* forces to soil surfaces and by *cohesive* forces to other water molecules, which are *surface tension* forces, and they result in fluid pressures that are less than atmospheric.

Rather than use negative pressure heads, $p < 0$, to describe the soil potential, we employ soil, or matric, tensions, $\psi > 0$. The matric tension is the negative fluid pressure head, $\psi = -p$, which increases as the pressure head becomes more negative (more tightly bound) and decreases as the pressure head becomes less negative (less tightly bound).

The total head within the soil pore fluids, h, is the difference between the elevation, z, and the matric tension:

$$H = z + p = z - \psi \tag{5}$$

Capillary theory can relate the matric tension, ψ, to the fluid potential associated with the radius of curvature, r, of fluid films within pores:

$$\psi = \frac{2\sigma \cos \alpha}{\gamma r} \tag{6}$$

where σ (Pa·m) is the surface tension and α is the contact angle at the air–water–solid interface.

The matric tension of a wet soil, $0 < \psi < 0.7$ bar, can be measured with tensiometers, as shown in (Fig. 2), which determine the negative pressure head of fluids inside a probe. The fluid within the probe is allowed to reach equilibrium with the soil fluids using a semiporous ceramic or metal filter that prevents air from entering the probe.

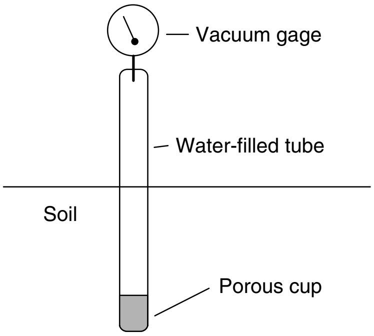

Figure 2. Soil-water tensiometer probe that monitors soil-tension. A porous cup allows water to exchange between the soil and a reservoir of water within the tube assembly. A vacuum pressure gauge monitors the water pressure inside the water reservoir.

Psychrometry is a useful technique in drier soils, $\psi > 5$ bar (2). The dewpoint temperature of the pore atmosphere is measured with a thermocouple psychrometer. The dewpoint temperature can then, along with the pore temperature, determine the relative humidity of the pore atmosphere. The matric tension can be readily determined from the relative humidity, as shown in Table 1.

MOISTURE CHARACTERISTIC CURVES

The amount of water contained within soil can be related to the fluid tension, or negative pressure, present within the water. Dry soils contain water that is held by strong adhesive forces to soil surfaces along with strong cohesive forces to other water molecules. These forces are reflected in the tension that holds the water against the force of gravity.

Each soil has a unique relationship between soil tension and water content. This relationship, termed the soil *moisture characteristic curve*, is a function that describes how water content and matric tension are related. The characteristic curve for each material varies with the pore size distribution—a soil with larger pores releases water at a lower tension than does a soil with smaller pores.

A number of formulations have been proposed that relate the relative saturation, Θ, to matric tension, ψ. The *Brooks and Corey* equation has the form:

$$\Theta = \left(\frac{\psi}{\psi_b}\right)^{-\lambda} \tag{7}$$

where ψ_b is the *bubbling pressure* and λ is a *pore-size distribution index*. This equation only holds for $\psi > \psi_o - \Theta = 1$ otherwise.

Another common formulation is the *van Genuchten* equation:

$$\Theta = \left[1 + (\alpha\psi)^n\right]^{(1-n)/n} \tag{8}$$

where α and n are empirical coefficients.

Moisture characteristic curves are commonly estimated with *pressure plates* (shown in (Fig. 3)) or *Tempe pressure cells*. In both methods, one or more soil samples are placed on top of a *porous plate* within a sealed chamber. Nitrogen gas under high pressure is forced into the chamber, and water is forced out of the soil samples through the porous

Table 1. Measures of Soil Tension

Fluid Pressure (bars)	Soil Tension (m)	Pore Radius	Relative Humidity (%)
0	0	∞	100
0.1	−1	15 μm	99.993
1	−10	1.5 μm	99.93
15	−150	100 nm	98.9
100	−1000	15 nm	93.
500	−5,000	3 nm	70.
1,000	−10,000	1.5 nm	48.
5,000	−50,000	3 Å	2.6
10,000	−100,000	1.5 Å	0.068

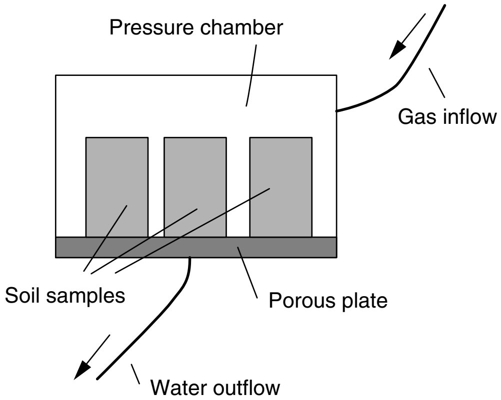

Figure 3. Pressure plate procedure for determining the soil moisture characteristic curve. A gas pressure is applied to soil samples within the chamber, and water is released through the porous plate. The water content, θ, is found for each pressure, ψ.

plate, which is permeable to water but not to air. The resulting water content of the samples can be determined once the outflow has stopped.

SAMPLING SOIL WATER

Soil water samples are more difficult to collect in the unsaturated zone than in the saturated zone because water does not freely flow into a well or piezometer. Commonly, a soil *lysimeter* collects water samples by placing a suction within a hollow tube that has a *porous cup* placed within the soil. Water flows through the porous cup into the hollow tube where it can then be collected.

Soil lysimeters are limited to conditions where the soils are wet, $\psi < 0.7$ bar. For drier conditions, soil samples can be collected and returned to the laboratory, where mechanical or gas pressure can be applied to the samples to force pore fluids out of the sample.

BIBLIOGRAPHY

1. Elder, A.N. and Rasmussen, T.C. (1994). Neutron probe calibration in unsaturated tuff. *Soil Sci. Soc. Am. J.* **58**(5): 1301–1307.
2. Rasmussen, T.C. and Rhodes, S.C. (1994). Energy-related methods: Psychrometers. In: *Vadose Zone Characterization and Monitoring; Principles, Methods and Case Studies*. L.G. Wilson, L. Everett, and S. Cullen (Eds.). Lewis Publishers, Boca Raton, FL.

SOIL AND GROUNDWATER GEOCHEMISTRY AND MICROBIOLOGY

Stuart Smith
Smith-Comeskey GroundWater Science LLC
Upper Sandusky, Ohio

DEFINITIONS

Geochemistry can be defined as the study of the chemical composition of, and of actual or possible chemical changes in, a geological matrix, including *soil*, that complex mixture of minerals, life, and organic material on the surface of the earth. *Groundwater*, of course, is water located in the saturated zone of the earth. *Microbiology* is the study of small forms of life, usually including the bacteria and archaea domains and protists and fungi of the eukaryaceae domain, and naturally also covers numerous subcategories.

The interdisciplinary studies of the interactions among microflora, earth materials, and chemical reactions are often linked under *geomicrobiology* or *biogeochemistry*.

ORIGINS

The relationship between life in soil and earth materials appears to go back to the very origins of life on earth. Bacterial and archaeans (superficially bacteria-like microflora that have entirely different genetic operations, biochemistry, and cell membranes) were interacting with iron and other minerals as part of energy transformations since within a few 100 million years of the coalescence of the planet. Many archaeans specialize in methanogenesis and adaptation to extreme environments at the present time, such as geothermal waters, undersea hydrothermal vents, and hypersaline lakes (e.g., Dead Sea, Great Salt Lake, and solar salt lagoons). Biochemical evidence suggests that they may have developed in association with such conditions, which were common early in earth history.

Bacteria and cyanobacteria (colloquially known as blue-green algae) strongly affected the earth environment from early in the Archean era (±4 billion years before present, b.y. BP). Cyanobacteria, with a photosynthetic mechanism similar to algae and plants, used carbon dioxide and expended oxygen as a waste product in shallow anoxic seas. Ferrous iron (Fe), common in water and the earth's crust, scavenged the oxygen, forming various FeIII minerals, including magnetite, hematite (recently confirmed on Mars), goethite, and less-organized mineral forms. Such iron-containing Archean sediments were deposited as banded iron formations, with iron bands in a cherty matrix, which are important iron ores worldwide where they occur (such as the Mesabi Range in Minnesota). When much of the available free oceanic FeII was depleted, oxygen began to build up in the atmosphere, with a recognizably modern atmosphere dating from 1.7 b.y. BP.

Bacteria have the most diverse biochemical and respiratory capabilities among forms of life. Although others require oxygen for respiration, bacteria can respire using sulfur, forms of nitrogen, carbon dioxide, and metals (primarily Fe and Mn) as electron acceptors. Such respiration is used to process a wide variety of organic compounds, including recalcitrant ring structures. Such versatility appears to be a very ancient capability.

In ancient shallow marine settings, cyanobacteria formed stromatolites, laminations of organic and sedimentary structures formed as sand and debris, and coprecipitated minerals were trapped in cyanobacterial masses. Although many stromatolites are found as fossil structures, a number of locations exist on the modern-day Earth where stromatolites are still forming, which are one form of the common microbial phenomenon, biofilm formation.

From there on, microbial processes and mineral-biological interactions, and the environment-manipulating "behavior" known as biofilm formation (all developed in the Archean era), have continued in action until the present.

- Bog ores (important to the development of an iron industry in the United States during the nation-forming period) and gold and copper placer ores are younger manifestations of the ancient microbial practice of depositing metals.
- Bacteria reduce oxides of Fe, Mn, N, and S to accumulate "reducing power" to oxidize ("digest") complex organic compounds. These capacities are exploited in groundwater contamination remediation. The presence of high levels of Fe and Mn in well water samples is also an indicator that organic carbon compounds are present in groundwater accessed by a well.
- Microbe-mediated biogeochemical transformations alter groundwater quality by oxidizing or reducing Fe and Mn, and mobilizing them into solution. Recently, the reduction of insoluble arsenic minerals to soluble forms and As mobilization by microbes has attracted interest as a factor in As presence in groundwater.
- Biofilm-forming microflora coat, clog, and corrode engineered structures such as wells and drains in the saturated zone, resulting in *biofouling* (further discussed in the article BIOFOULING IN WATER WELLS).

SOIL MICROBIOLOGY

Soil microbiology is a special category of geomicrobiology focused on the earth's regolith, and specifically *soil*, which is a complex of mineral, organic, and biological components. Although beyond the scope of a short article, soil microbiology has traditionally been distinguished from groundwater or geomicrobiology. This distinction has origins in academic discipline; however, studies indicate that surficial soils have species richness and total microbial biomass values that are much higher than in earth materials below the soil zone. Protists and fungi are common in soil but limited in type and numbers (but not entirely absent) in aquifers.

Soil microbiology is much more thoroughly known compared with that of less-accessible earth materials. It is also much more thoroughly influenced by human activities, which include cultivation, fertilization, construction, and application of contaminants (both deliberately and accidentally). Soil microflora and soil structure experience close interactions with animals and plants. Changes in such macroflora and fauna communities can trigger sweeping changes in soil microbiology. Clearing and cultivation of forest and prairie soils result in simplification in microbial community structure. The introduction of Eurasian earthworms into North American boreal soils (where they have been absent since the Pleistocene glaciations) has resulted in measurable shifts in the availability of organic matter and nutrients to native plants. In turn, changes in vegetation affect soil and watershed hydrology.

OVERCOMING CHALLENGES IN MICROBIOLOGICAL METHODS IN GROUNDWATER STUDIES

Background

Environmental microbiology and geomicrobiology are the oldest disciplines within the broader field of microbiology. However, industrial, public health, and clinical interests quickly overshadowed the study of the microbiology of natural environments. Until the 1970s, scientific preconceptions in many circles, based on observations of engineered filters, limited methods (and limited reading) and placed limits on groundwater microbiology. Considerations of the chemistry of groundwater assumed that chemical transformations occurred in an environment devoid of life. A notable exception was that of the microbiologists in the Soviet petroleum industry of the 1950s and 1960s, who described the microbiology of petroleum reservoirs and systems.

Analysis

Methods for the analysis of groundwater environments for microbes were very limited prior to the 1980s. The principal analytical method was culturing, which recovers only a small fraction of bacteria. It was generally assumed that water passing through the soil was purified by active microbial processes and by filtration.

Interest in the occurrence of biofouling microflora in groundwater and concerns about groundwater contamination in the 1970s and 1980s motivated innovation in methods suitable for groundwater microbiology. As groundwater contamination became more and more evident in the United States and other western nations during the 1980s, the motivation for understanding groundwater environments increased. In addition, new methods in microbiology, based on advances in molecular biology, provided microbiologists with new tools to explore this difficult-to-sample microbial habitat. Among these are phospholipid fatty acid methyl ester (PL-FAME) and nucleic acid testing, described further in EVALUATION OF MICROBIAL COMPONENTS OF BIOFOULING. Such methods provide the means to describe microbial associations in a meaningful way, as cultivation and phenotypic classification typically fail to do so.

Classification, in particular, is a problem in description. With micro-organisms, it often depends on making comparisons with libraries of known genotypes or phenotypes. Some deep subsurface communities are so isolated and specifically adapted that comparisons with described forms are weak. A general recognition exists in the geobiological community that the task of describing aquifer and other deep subsurface communities compares with the task of describing the recently discovered deep-sea hydrothermal vent ecosystems.

Sampling

Acquiring samples is perhaps the most difficult aspect of the microbiology of the deep subsurface. Such sampling is relatively expensive, at times technically challenging, and prone to sample compromise. Typically, methods for both solid and liquid sample acquisitions are adapted from those developed for other purposes (borehole and core drilling, pumped or zonal sampling). Only coring can deliver solid samples isolated from drilling contamination. A variety of water sampling techniques exist to obtain discrete samples.

However, if properly interpreted, available sampling and analytical techniques appear to be sufficient to gain an understanding of deep subsurface microbial ecology.

NOTABLE DISCOVERIES

1. Microbial community diversity: PL-FAME and rRNA analysis reveal subsurface communities that are diverse and typically adapted to coping with starvation and exploiting resource "windfalls" when they occur. Bacteria can persist in viable but nonculturable and inactive states when conditions dictate. Microbial communities can have biomass that rivals surface soils and may include fungi and protozoa where pore sizes permit.
2. Microbial communities and numbers vary significantly between hydraulically transmissive and confining formations in the same wells and boreholes.
3. Community biochemical adaptability: Bacteria that form biofilms and readily exchange genetic potential can be quite plastic in their ability to exploit resources. Early in the effort to employ bioremediation in treating contaminated groundwater and soil, it was recognized that natural communities were capable of degrading a variety of anthropogenic compounds, especially hydrocarbons.
4. Iron and sulfur biogeochemical cycles are particularly important in deeper systems, whereas nitrogen is strongly cycled in the soil and shallow saturated flow zones. In deeper systems, carbon and nitrogen are cycled locally within biofilms.
5. Groundwater contamination locally alters oxidation-reduction systems, nutrient cycling, and community structure. Microbial communities react to influxes of organic carbon by increasing biomass and tipping toward anaerobic respiration as oxygen is depleted by hydrocarbon-using or other aerobic heterotrophs, which is closely analogous to eutrophication in surface water systems. For energetic reasons, the following are favored sequentially, switching to the next when the electron acceptor is depleted: nitrate, iron (and other metal) oxides, sulfate, and carbon dioxide (reduced to methane).
6. Microbial communities may include very long-lived and ancient lineages. Studies of communities in deep, isolated systems reveal communities that may be very ancient and existing in isolation from the surface.

An article of this length is insufficient to consider the subject comprehensively (considering that entire textbooks, conferences, and hundreds of webpages are devoted to the topic). Related subjects are considered elsewhere in the *Encyclopedia of Water*.

READING LIST

Ehrlich, H.L. (2002). *Geomicrobiology*, 4th Edn. Marcel Dekker, New York.

Krumbein, W.E. (Ed.). (1983). *Microbial Geochemistry*. Blackwell Scientific Publications, Oxford, UK.

Bitton, G. and Gerba, C.P. (Eds.). (1984). *Groundwater Pollution Microbiology*. Wiley-Interscience, New York.

Sinclair, J.L. and Ghiorse, W.C. (1989). Distribution of aerobic bacteria, protozoa, algae and fungi in deep subsurface sediments. *Geomicrobiology*. **7**(1–2): 15–32.

Chapelle, F.H. (1993). *Ground-Water Microbiology and Geochemistry*. John Wiley & Sons, New York.

Chapelle, F.H. (2000). *Ground-Water Microbiology and Geochemistry*, 2nd Edn. John Wiley & Sons, New York.

Madsen, E.L. and Ghiorse, W.C. (1993). Groundwater microbiology: subsurface ecosystem processes. In: *Aquatic Microbiology*. T.E. Ford (Ed.). Blackwell Scientific Publications, Boston, MA.

LeClerc, H. (2004). Microbiology of Natural Mineral Waters. Centre Evian. Available: http://www.centre-evian.com.

Madsen, E. (2004). Bibliography of important works in environmental microbiology. Available: http://www.micro.cornell.edu/faculty/Madsen/Bibliography_of_important_works_in_Environmental_Microbiology.htm.

Sylvia, D.M., Fuhrmann, J.J., Hartel, P.G., and Zuberer, D.A. (2005). *Principles and Applications of Soil Microbiology*, 2nd Edn. Prentice-Hall, New York.

CHARACTERIZING SOIL SPATIAL VARIABILITY

Dennis L. Corwin
Scott M. Lesch
USDA-ARS George E. Brown,
Jr., Salinity Laboratory
Riverside, California

SPATIAL VARIABILITY: DEFINITION, ORIGIN, SIGNIFICANCE, AND MEASUREMENT METHODS

Spatial variability refers to the variation in soil physical, chemical, mineralogical, and biological properties both laterally across the landscape and vertically through the soil profile. The spatial heterogeneity of soils is a consequence of interacting meteorological, pedological, and biological processes and anthropogenic factors over time. Within-field spatial variability of soil properties has been a primary concern within soil science ever since the classic paper by Nielson et al. (1) concerning the variability of field-measured soil water properties. The significance of soil spatial variability stems from its complex influence on any landscape-scale soil-related issue including solute transport in the vadose zone, precision agriculture, and soil quality assessment, to mention a few.

A variety of measurement methods are available for potentially characterizing soil spatial variability including ground penetrating radar (GPR), aerial photography, multispectral and hyperspectral imagery, time-domain reflectometry (TDR), and apparent soil electrical conductivity (EC_a). However, none of these approaches has been as extensively investigated with respect to characterization of within-field soil variability as EC_a (2).

Since its early agricultural use for measuring soil salinity, the application of EC_a has evolved into a widely accepted means of establishing the spatial variability of several soil physicochemical properties that influence the EC_a measurement. Geospatial measurements of EC_a are well suited for characterizing spatial variability for several reasons. Geospatial measurements of EC_a are reliable, quick, and easy to take. The mobilization of EC_a measurement equipment is easy and can be accomplished at a reasonable cost. Finally, and most importantly, EC_a is influenced by a variety of soil properties for which the spatial variability of each could be potentially established.

NEED FOR CHARACTERIZING SOIL SPATIAL VARIABILITY

The prospect of feeding a projected additional 3 billion people over the next 30 years poses formidable, but not insurmountable, challenges. Feeding the ever-increasing world population will require a sustainable agricultural system that can keep pace with population growth. Sustainable agriculture hinges on a delicate balance of economic sustainability, crop productivity, resource utilization, and minimized detrimental environmental impacts. These factors are influenced either directly or indirectly by soil spatial variability.

Spatial variability within a field amounts to an order of magnitude or more for soils (3,4) and to a factor of 2–4 or more for crops (5). Spatial variation in crops is the result of a complex interaction of biological (e.g., pests, earthworms, microbes), edaphic (e.g., salinity, organic matter, nutrients, texture), anthropogenic (e.g., irrigation efficiency, soil compaction due to farm equipment), topographic (e.g., slope, elevation), and climatic (e.g., relative humidity, temperature, rainfall) factors. To a varying extent from one field to the next, crop patterns are influenced by soil-related properties. Conventional farming is not as efficient as it could be because it currently treats a field uniformly. To maintain uniform crop production, conventional farming overcomes the inherent variability of soil and crop conditions between and within fields with an over application of inputs (i.e., irrigation water, pesticides, fertilizers, etc.) intended to meet the needs of areas of lowest productivity. Bullock and Bullock (6) pointed out the imminent need for efficient methods to accurately measure within-field variation in soil physical and chemical properties as a key component for precision agriculture.

The effective utilization of soil as a resource depends on assessing and maintaining its quality. A fundamental component of assessing field-scale soil quality is establishing the spatial distribution of the soil properties affecting the soil's intended management goal (e.g., maximize agricultural productivity, minimize environmental impact, and/or maximize waste recycling) and its intended function (e.g., biodiversity, filtering and buffering, nutrient cycling, physical stability and structural support, resistance and resilience, and water and solute flow). It is not sufficient to take a single measurement within a field to characterize its soil quality. Rather, a sufficient number of measurements must be taken and at specific locations to representatively characterize the spatial distribution of the existing soil conditions that influence the soil's intended use. Therefore, assessing soil quality requires quantitative knowledge of each indicator property associated with a soil's quality and the spatial variability of those indicator properties.

Furthermore, spatial variability has a profound influence on environmental impacts, particularly of non-point source pollutants. This result is because of its influence on the fate and transport of agrochemical inputs (e.g., pesticides, fertilizers, etc.) in an agroecosystem. In fact, it has become clear that the real constraint on modeling solute transport is not the detail of the model structures, but defining the characteristics of individual places (7). Jury (3) provides an excellent fundamental discussion of the spatial variability of soil properties and its impact on solute transport in the vadose zone. As Jury points out, “any hope of estimating a continuous spatial pattern of chemical emissions at each point in space within a field must be abandoned due to field-scale variability of soils.”

The characterization of spatial variability is without a doubt one of the most significant areas of concern in soil science because of its broad reaching influence on all field- and landscape-scale issues. Research in the characterization of spatial variability has advanced tremendously over the past decade in large part because of georeferenced EC_a measurements.

EDAPHIC FACTORS INFLUENCING EC_a MEASUREMENTS

The earliest field applications of geophysical measurements of EC_a in soil science, which occurred in the early 1980s, involved the determination of salinity within the soil profile of arid zone soils. However, it soon became apparent that the measurement of EC_a in the field as a quantification of soil salinity was more complicated than initially anticipated because of the complexity of current flow pathways originating from the spatial heterogeneity of properties influencing current flow in soil.

Three pathways of current flow contribute to the EC_a of soil: (1) a liquid phase pathway via dissolved solids contained in the soil water occupying the large pores, (2) a solid–liquid phase pathway primarily via exchangeable cations associated with clay minerals, and (3) a solid pathway via soil particles that are in direct and continuous contact with one another (8). Of these three pathways, the solid pathway is usually negligible; hence, most soil can be approximated with a dual-pathway parallel conductance equation.

An analysis of this dual-pathway system indicates that EC_a will be influenced by a complex interaction of soil properties including salinity, soil texture, water content, ρ_b, and temperature. The ρ_b is directly influenced by clay content (soil texture) and organic matter (OM). Furthermore, the exchange surfaces on clays and OM provide a solid–liquid phase pathway primarily via exchangeable cations; consequently, the cation exchange capacity (CEC) and OM are recognized as additional factors that can influence EC_a measurements. Measurements of EC_a *must* be interpreted with these influencing factors in mind.

MOBILE EC$_a$ MEASUREMENT EQUIPMENT

An EC$_a$ survey to characterize spatial variability uses mobile geophysical equipment coupled with a global positioning system (GPS) to georeference each EC$_a$ measurement. Mobile EC$_a$ equipment has been developed by a variety of researchers (9–11). The development of mobile EC$_a$ measurement equipment has made it possible to produce EC$_a$ maps with measurements taken every few meters.

Mobile EC$_a$ measurement equipment has been developed for both electrical resistivity (ER) and electromagnetic induction (EMI) geophysical approaches. In the case of ER, four stainless-steel electrodes are inserted into the soil generally at equal distances and connected to a resistivity meter. Current is applied to the two outer electrodes with the two inner electrodes serving as the potential electrodes. Fixing the electrode array and mounting it on or pulling it from a vehicle saves considerable time for measurements. By coupling the fixed electrode array to a datalogger and GPS georeferences the EC$_a$ measurement, geospatial ECa measurements can be taken. Veris Technologies[1] (Salinas, KS; www.veristech.com) has developed a commercial mobile system for measuring EC$_a$ using the principles of ER. In a similar fashion, an EMI unit can be mobilized by pulling a sled-mounted EM-38 unit (Geonics Ltd., Mississaugua, Ontario, Canada) behind a vehicle and logging the EMI data simultaneously along with real-time GPS position information. Measurements of EC$_a$ with EMI generally involve taking readings in the horizontal coil configuration (EM$_h$), which is a shallow reading penetrating to a depth of about 0.75–1.0 m, and/or in the vertical coil configuration (EM$_v$), which is a deep reading penetrating to a depth of about 1.5 m.

EC$_a$-DIRECTED SAMPLING DESIGNS

Despite the complexities of the EC$_a$ measurement, the fact that EC$_a$ is influenced by several soil properties has broadened its utility and enabled it to play a major role in addressing the issue of characterizing spatial variability. Geospatial measurements of EC$_a$ have been successfully used for (1) identifying the soil physicochemical properties influencing crop yield patterns and soil condition, (2) establishing the spatial variation of these soil properties, and (3) characterizing the spatial distribution of soil properties influencing solute transport through the vadose zone (4,12–14).

In a typical field survey, the spatial distribution of various soil properties can often be adequately determined with conductivity survey information in conjunction with some type of soil sampling plan. Either probability (design) or prediction (model) based sampling strategies can identify the soil properties that correlate best with EC$_a$ in a particular field. Prediction-based sampling strategies objectively select sampling locations based on the observed spatial ECa pattern and thus are often referred to as “sensor-directed” sampling plans (15). Characterizing spatial variability with EC$_a$-directed soil sampling is based on the hypothesis that when EC$_a$ correlates with a soil property or properties, then spatial EC$_a$ information can identify sites that reflect the range and variability of the property or properties. In instances where EC$_a$ correlates well with a particular soil property, an EC$_a$-directed soil sampling approach can optimally establish the spatial distribution of that property while minimizing the corresponding sampling and labor costs (2,15,16). Also, if EC$_a$ is correlated with crop yield, then an EC$_a$-directed soil sampling approach can identify those soil properties that are causing the variability in crop yield (13). Details for conducting a field-scale EC$_a$ survey for the purpose of characterizing the spatial variability of soil properties influencing soil quality or crop yield variation can be found in Corwin and Lesch (16).

Figure 1 shows EM$_h$ and EM$_v$ maps of an EC$_a$ survey performed on a 6.5-ha field in Maricopa, AZ, along with the associated soil sample sites selected with a specific type of EC$_a$-directed sampling methodology known as spatial response surface sampling design (15,16). Once soil cores are taken at the selected sites and the corresponding soil physicochemical properties are analyzed, the properties that are well correlated with EC$_a$ can be identified. Statistical prediction models can then be derived that spatially map (predict) these correlated soil properties with the calibrated EC$_a$ survey data. Figure 2 shows the corresponding EC$_a$-predicted 0–0.9-m percent water content patterns for the field shown in Fig. 1.

GUIDELINES FOR CONDUCTING AN EC$_A$-DIRECTED SOIL SAMPLING SURVEY TO CHARACTERIZE SPATIAL VARIABILITY

The basic elements of a field-scale EC$_a$ survey for characterizing spatial variability include (1) EC$_a$ survey design, (2) georeferenced EC$_a$ data collection, (3) soil sample design based on georeferenced EC$_a$ data, (4) soil sample collection, (5) physicochemical analysis of pertinent soil properties, (6) spatial statistical analysis, (7) determination of the dominant soil properties influencing the EC$_a$ measurements at the study site, and (8) GIS development. The outlined steps of an EC$_a$-directed soil sampling survey are provided in Table 1.

LIMITATIONS AND PRECAUTIONS

At present, no other soil measurement provides a greater level of spatial information than that of geospatial EC$_a$ measurements. Even so, are limitations exist to the use of EC$_a$ to characterize soil spatial variability. Awareness of these weaknesses is crucial for the proper use and interpretation of geospatial EC$_a$ measurements. The complex spatial heterogeneity of the soil system has subtle influences on geospatial EC$_a$ measurements that can have significant interpretive impacts. The ability to recognize and interpret these influences can be the difference between the successful or failed application of EC$_a$ measurements for characterizing spatial variability.

First and foremost, in the absence of soil sample information, geospatial measurements of EC$_a$ by themselves do not directly quantify soil spatial variability.

---

[1] Product identification is provided solely for the benefit of the reader and does not imply the endorsement of the USDA.

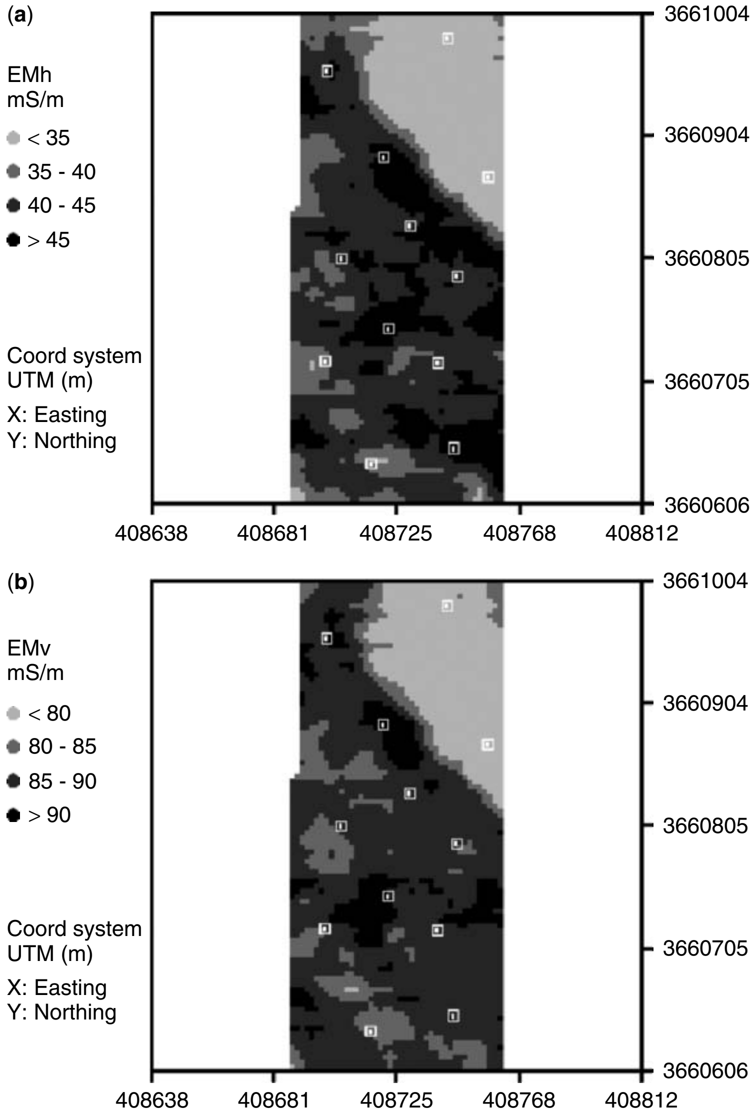

Figure 1. Maps of EC_a survey data and EC_a-directed soil sampling sites for a field in Maricopa, AZ: (**a**) EM38 horizontal and (**b**) EM38 vertical. White square symbol represents the soil sample site locations. UTM = universal transverse mercator grid system in meters.

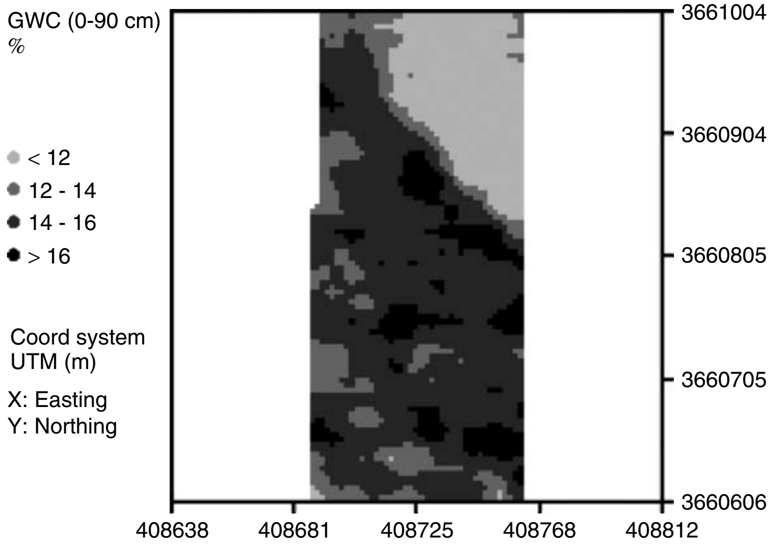

Figure 2. Map of spatial variability of percent gravimetric water content (GWC) for the 0–0.9-m sampling depth, as determined from the survey data and EC_a-directed soil sampling sites shown in Fig. 1 ($R^2 = 0.85$ for predicted and measured GWC). UTM = universal transverse mercator grid system in meters.

Table 1. Outline of Steps to Conduct an EC_a Field Survey

1. Site description and EC_a survey design
   a. record site metadata
   b. define the project's/survey's objective
   c. establish site boundaries
   d. select GPS coordinate system
   e. establish EC_a measurement intensity
2. EC_a data collection with mobile GPS-based equipment
   a. georeference site boundaries and significant physical geographic features with GPS
   b. measure georeferenced EC_a data at the predetermined spatial intensity and record associated metadata
3. Soil sample design based on georeferenced EC_a data
   a. statistically analyze EC_a data using an appropriate statistical sampling design to establish the soil sample site locations
   b. establish site locations, depth of sampling, sample depth increments, and number of cores per site
4. Soil core sampling at specified sites designated by the sample design
   a. obtain measurements of soil temperature through the profile at selected sites
   b. at randomly selected locations, obtain duplicate soil cores within a 1-m distance of one another to establish local-scale variation of soil properties
   c. record soil core observations (e.g., mottling, horizonation, textural discontinuities, etc.)
5. Laboratory analysis of appropriate soil physicochemical properties defined by project objectives
6. If needed, stochastic and/or deterministic calibration of EC_a to EC_e or to other soil properties (e.g., water content and texture)
7. Spatial statistical analysis to determine the soil properties influencing EC_a and/or crop yield
   a. soil quality assessment:
      1. perform a basic statistical analysis of physicochemical data by depth increment and by composite depth over the depth of measurement of EC_a
      2. determine the correlation between EC_a and physicochemical soil properties by composite depth over the depth of measurement of EC_a
   b. precision agriculture applications (if EC_a correlates with crop yield, then):
      1. perform a basic statistical analysis of physicochemical data by depth increment and by composite depths
      2. determine the correlation between EC_a and physicochemical soil properties by depth increment and by composite depths
      3. determine the correlation between crop yield and physicochemical soil properties by depth and by composite depths to determine depth of concern (i.e., depth with consistently highest correlation, whether positive or negative, of soil properties to yield) and the significant soil properties influencing crop yield (or crop quality)
      4. conduct an exploratory graphical analysis to determine the relationship between the significant physicochemical properties and crop yield (or crop quality)
      5. formulate a spatial linear regression (SLR) model that relates soil properties (independent variables) to crop yield or crop quality (dependent variable)
      6. adjust this model for spatial autocorrelation, if necessary, using restricted maximum likelihood or some other technique
      7. Conduct a sensitivity analysis to establish dominant soil property influencing yield or quality
8. GIS database development and graphic display of spatial distribution of soil properties

Source: Ref. 16. With permission.

Rather, EC_a-survey measurements provide the spatial information necessary to direct soil sampling. It is the combined use of both the EC_a-survey and the corresponding sample data that makes this sort of surveying technology invaluable for characterizing spatial variability. In other words, the primary strength of geospatial EC_a measurements lies in their effectiveness as a means to direct soil sampling (using a minimum number of sample sites) that best characterizes the spatial variability of those soil properties influencing EC_a in the landscape of interest.

Second, EC_a-directed soil sampling can only spatially characterize soil properties that correlate with EC_a. This correlation may be caused by a direct or indirect influence on the EC_a measurement, or the correlation may be a complete artifact. For example, salinity and water content will directly influence EC_a and CEC will indirectly influence EC_a through its influence on current flow at the surface of soil minerals. Likewise, in many instances, B and salinity distributions are similar. Consequently, a strong correlation between B and salinity (and thus EC_a) can result, even though no cause-and-effect relationship exists between B and EC_a. Thus, a solid understanding of the soil properties that directly influence and/or are indirectly measured by EC_a at a specific site is particularly essential in temporal surveying applications, because over time these correlations may or may not persist.

Third, as previously mentioned, often a dynamic and complex relationship exists between EC_a and those properties that influence EC_a. Ground-truth soil samples are obligatory if spatial measurements of EC_a are to be used for quantifying soil variability. Geospatial measurements of EC_a do not supplant the need for soil sampling, but they do minimize the number of samples necessary to characterize spatial variability. Additionally, users of EC_a survey data must exercise caution and be aware of what EC_a is actually measuring at the site of interest. Normally, the only way to establish those soil properties that influence EC_a at a specific location is by collecting representative soil samples and then determining the relationship between EC_a and the particular property or properties of interest. This process requires that every EC_a survey has an associated soil sampling design. The location and number of sites can be established (after acquiring the EC_a survey data) using model-based sample design software such as ESAP (17).

Finally, the temporal stability of EC_a measurements at a site may be of potential concern because EC_a is a product of both static and dynamic factors, which adds another dimension to the complexity of understanding and interpreting geospatial EC_a measurements. For this reason, caution must be taken to characterize spatial variability with EC_a when dynamic conditions influencing EC_a are significant. Apparent soil electrical conductivity surveys are generally conducted (1) within a set time frame to minimize the effects of dynamic properties (such as temperature and water content), (2) when the soil is at or near field capacity (to avoid current-flow problems), and (3) with regard for subtle topographic effects (e.g., bed-furrow). Protocols for conducting an EC_a survey that consider all previously discussed limitations are presented by Corwin and Lesch (16).

FUTURE DIRECTION

Because of the heterogeneous nature of soils, characterization of soil spatial variability is a fundamental component of any landscape-scale process that cannot be overlooked or superficially addressed. Significant technological advances have occurred over the past two decades, particularly in the area of sensor technology and in precisely locating a geographic position with GPS. These advances have helped the progress of research in characterizing spatial variability at field and landscape scales.

Geospatial measurement of EC_a stands out among the ground-based sensor technologies, contributing to an improved ability to characterize spatial variability. Numerous soil samples are required for representative estimates of field-scale spatial variability using traditional grid sampling, which makes grid sampling impractical because of labor and cost intensiveness. Soil sampling directed by geospatial EC_a measurements provides a viable alternative for characterizing spatial variability of a variety of soil-related physicochemical properties. Geospatial EC_a measurements provide a means of significantly reducing the number of soil samples needed to characterize spatial variability, provided that the target soil properties are well correlated with the conductivity survey data.

Directed soil sampling with geospatial EC_a measurements has its limitations in characterizing spatial variability, particularly when the soil properties of interest are not measured either directly or indirectly by the EC_a survey data. In these instances, other types of sensors should be used (if available) or a more intensive soil sampling strategy must be adopted. A robust methodology for the integrated use of multiple remote and ground-based sensors has yet to be developed. The successful integration of multispectral and hyperspectral imagery, TDR, GPR, aerial photography, and EC_a sensors is needed to provide the redundant and supplemental data necessary to unravel the spatial complexity of soil. This type of integrated methodology represents the future direction of soil spatial variability research.

BIBLIOGRAPHY

1. Nielsen, D.R., Biggar, J.W., and Erh, K.T. (1973). Spatial variability of field-measured soil-water properties. *Hilgardia* **42**(7): 215–259.
2. Corwin, D.L. and Lesch, S.M. (In press). Apparent soil electrical conductivity measurements in agriculture. *Comput. Electron. Agric.*
3. Jury, W.A. (1986). Spatial variability of soil properties. In: *Vadose Zone Modeling of Organic Pollutants*. S.C. Hern and S.M. Melancon (Eds.). Lewis Publishers, Chelsea, MI, pp. 245–269.
4. Corwin, D.L. et al. (2003). Assessment and field-scale mapping of soil quality properties of a saline-sodic soil. *Geoderma* **114**(3–4): 231–259.
5. Verhagen, A., Booltink, H.W.G., and Bouma, J. (1995). Site-specific management: Balancing production and environmental requirements at farm level. *Agric. Syst.* **49**: 369–384.
6. Bullock, D.S. and Bullock, D.G. (2000). Economic optimality of input application rates in precision farming. *Prec. Agric.* **2**: 71–101.
7. Beven, K.J. (2002). Towards an alternative blueprint for a physically-based digitally simulated hydrologic response modeling system. *Hydrol. Process.* **16**(2): 189–206.
8. Rhoades, J.D., Chanduvi, F., and Lesch, S. (1999). Soil Salinity Assessment: Methods and Interpretation of Electrical Conductivity Measurements, FAO Irrigation and Drainage Paper #57, Food and Agriculture Organization of the United Nations, Rome, Italy, pp. 1–150.
9. McNeill, J.D. (1992). Rapid, accurate mapping of soil salinity by electromagnetic ground conductivity meters. In: *Advances in Measurements of Soil Physical Properties: Bringing Theory into Practice*. G.C. Topp, W.D. Reynolds, and R.E. Green (Eds.). *SSSA Special Publication* No. 30, ASA-CSSA-SSSA, Madison, WI, pp. 201–229.
10. Rhoades, J.D. (1993). Electrical conductivity methods for measuring and mapping soil salinity. In: *Advances in*

Agronomy. D.L. Sparks (Ed.). Vol. 49, Academic Press, San Diego, CA, pp. 201–251.

11. Jaynes, D.B., Colvin, T.S., and Ambuel, J. (1993). Soil type and crop yield determinations from ground conductivity surveys. ASAE Paper No. 933552, 1993 ASAE Winter Meetings, 14–17 Dec. 1993, Chicago, IL, ASAE, St. Joseph, MI.
12. Corwin, D.L. et al. (1999). Evaluation of GIS-linked model of salt loading to groundwater. *J. Environ. Qual.* **28**: 471–480.
13. Corwin, D.L. et al. (2003). Identifying soil properties that influence cotton yield using soil sampling directed by apparent soil electrical conductivity. *Agron. J.* **95**(2): 352–364.
14. Corwin, D.L., Lesch, S.M., Oster, J.D., and Kaffka, S.R. (In press). Characterizing spatio-temporal variability with soil sampling directed by apparent soil electrical conductivity. *Geoderma*.
15. Lesch, S.M. (In press). Sensor-directed response surface sampling designs for characterizing spatial variation in soil properties. *Comput. Electron. Agric.*
16. Corwin, D.L. and Lesch, S.M. (In press). Characterizing soil spatial variability with apparent soil electrical conductivity: I. Survey protocols. *Comput. Electron. Agric.*
17. Lesch, S.M., Rhoades, J.D., and Corwin, D.L. (2000). ESAP-95 Version 2.10R: User Manual and Tutorial Guide, Research Rpt. 146, USDA-ARS George E. Brown, Jr. Salinity Laboratory, Riverside, CA.

DEEP SOIL-WATER MOVEMENT

Todd Rasmussen
The University of Georgia
Athens, Georgia

Soil water refers to subsurface water between the Earth's surface and the regional water table, called the unsaturated, or vadose zone. Figure 1 illustrates the components of flow within this zone. Deep soil-water movement usually implies the migration of water below the root zone, and is also called deep percolation.

Soil water is held by adhesive forces to soil surfaces and by cohesive forces to other water molecules. Measurements of soil-water pressures are usually negative—water does not freely drain into a well. Some soil water may exist with positive pressures, such as water that is perched in a saturated zone above a layer of lower permeability. Water under positive pressure flows freely into a well.

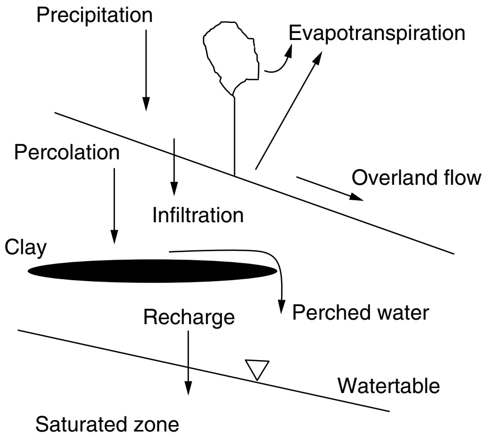

Figure 1. Flow components in the unsaturated zone.

FORCES THAT CONTROL WATER MOVEMENT

Gravity is the dominant force that controls soil-water movement. The downward force of gravity acts in conjunction with geologic media to control the magnitude and direction of fluid flow.

A second important force is the matric tension (or soil-water tension), which causes water to move from zones of lower tension to zones of higher tension. The soil tension, ψ, is the negative of the water pressure head, $\psi = -P/\gamma$, where P is the fluid pressure relative to atmospheric pressure and γ is the freshwater specific weight.

The total head, H, combines the gravitational and tension forces using $H = z - \psi$, where z is the elevation of the point of measurement.

In some circumstances, osmotic forces can also affect water movement. Osmotic forces develop when solute concentration gradients are present, causing water to move from zones of low salt concentration to zones of higher concentration. When osmotic forces are present, the total head must include the osmotic potential, ϕ, as noted in Section GW456 (head).

MATERIAL PROPERTIES

In addition to the total head, the characteristics of the geologic medium are also important for controlling the magnitude and direction of deep soil-water movement (1). The important material properties that affect soil-water movement are the unsaturated hydraulic conductivity functions and the soil-water characteristic curves.

Unsaturated Hydraulic Conductivity

The unsaturated hydraulic conductivity, $K(\psi)$, is commonly written as:

$$K(\psi) = K_r(\psi)\,K_s \tag{1}$$

where K_r is the relative unsaturated hydraulic conductivity function that varies with matric tension, ψ, and K_s is the saturated hydraulic conductivity.

As soil dries, the tension increases and the relative hydraulic conductivity decreases, whereas wetting soils show a decrease in tension and an increase in relative conductivity. The relative conductivity ranges from a theoretical minimum of zero in dry soil, $K_r \to 0$ as $\psi \to 0$, to a maximum of one in saturated soil, $K_r \to 1$ as $\psi = 0$, so that $0 < K_r \leq 1$.

Both the relative and saturated hydraulic conductivities vary with the pore-size distribution of the geologic media. Sandy materials may have a large saturated hydraulic conductivity, but a rapidly declining relative hydraulic conductivity, whereas finer materials may have a smaller saturated hydraulic conductivity but a slowly

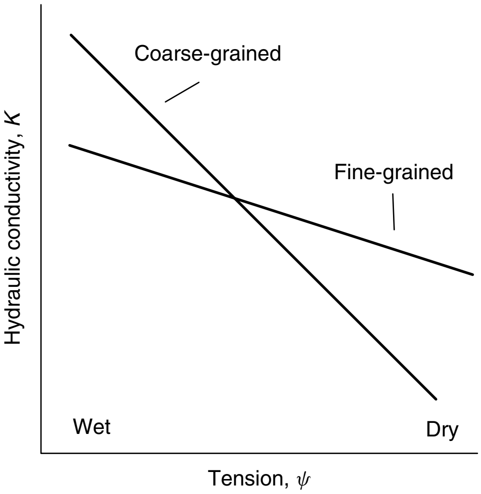

Figure 2. Unsaturated hydraulic conductivity, K, as a function of tension, ψ, for two materials. The coarser material is more conductive under wet conditions (smaller values of tension), whereas the finer material is more conductive under drier conditions.

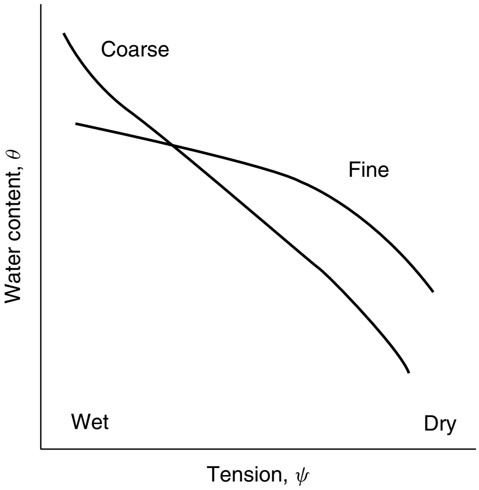

Figure 3. Soil water content, θ, as a function of tension, ψ, for two materials. The coarser material holds more water under wet conditions (smaller values of tension), whereas the finer material holds more water under drier conditions.

declining relative conductivity. This effect is shown if Fig. 2.

The crossover of the two curves shown in Fig. 2 results in the interesting phenomenon that coarser materials have a larger unsaturated hydraulic conductivity than finer material in the wet range, but a smaller conductivity in the drier range.

Thus, the flow behavior for wet conditions may be very different from the behavior under dry conditions in heterogeneous materials. Coarser materials (such as sands, fractures, and other large openings) generally dominate flow under wet conditions, whereas finer materials (such as silts, clays, and consolidated materials) can dominate flow under drier conditions.

Moisture Characteristic Curves

The ability of the unsaturated zone to hold water as a function of the soil tension is called the moisture characteristic curve. The characteristic curve for each material varies with the pore size distribution—a soil with larger pores releases water at a lower tension than a soil with smaller pores.

Figure 3 shows two soil characteristic curves, one for a coarser material and a second for a finer material. Like the unsaturated hydraulic conductivity function, the finer soil can hold more water under dry conditions, whereas the coarser soil holds more water under wetter conditions.

The water retention properties of soils also affect water movement. Changes in water flux with depth can be attributed to the uptake and release of water as the soil tension changes. For example, a wetting front may move downward with time from the surface after a large rain event, only to be absorbed by the dry soil underneath. The ability to absorb the moisture is a function of the ambient moisture tension—an already wet soil will have little ability to absorb additional water.

In some cases, water percolating down from above may not immediately reach the water table because the dry soils mediate the pulse of water. This water is retained by the soil and then slowly released over time. The initial pulse of water at the surface is transformed into a slow drainage at depth.

Only in materials with little water storage capacity (i.e., in wet soils or in materials with limited water holding capacity) will the downward movement of water be rapid. In these cases, the water that reaches the water table may or may not be the same water that was added at the surface. In some situations, the added water merely displaces the water below it, causing an equivalent volume of water to move at depth. In other situations, the added water moves quickly through the unsaturated zone.

Spatial Variability

Local subsurface conditions can also affect deep soil-water movement. Hydrogeologic units that dip in one direction, such as bedded sedimentary materials, can cause water to flow in the down-dip direction. Microscopic layering in features such as sand dunes can also affect the direction of movement.

Vertical features, such as fractures and joints, may focus water movement during wet conditions. Localized zones of preferential flow can also be found in soils with a defined structure, such as in soil with aggregates. These soils form a dual-porosity network consisting of fine pores within the aggregates and larger macropores between the aggregates.

Many explanations exist for preferential flow within soils, including bypass flow, macropore flow, fracture flow, boundary layer flow, mobile zone flow, finger flow, funnel flow, media heterogeneities, ion exclusion, and colloid transport (2).

WATER VAPOR MOVEMENT

Water vapor movement in deep systems can also be important, especially in very dry media where liquid flow is small. Water vapor is generally in equilibrium with the liquid phase and can be related to the matric tension of the pore fluids (3):

$$RH = \frac{e}{e_s} = \exp\left(\frac{g\psi}{RT}\right) \tag{2}$$

where RH is the relative humidity, e is the actual vapor pressure, e_s is the saturated vapor pressure, R is the water vapor gas constant, and T is the absolute temperature. Dissolved solids in soil water can lower the relative humidity—and increase the tension—because of osmotic forces.

The diffusion of water vapor through the subsurface can be predicted using a variety of approaches that depend on the specific conditions within the subsurface, which is because additional forces become important when considering subsurface vapor movement, including thermal effects and barometric pressure changes.

Temperature variation affects water movement because the saturated vapor pressure increases with temperature. The natural temperature increase with depth is called the geothermal gradient. Deeper water is warmer and has a higher vapor pressure than water at a shallow depth. This vapor pressure gradient causes water vapor to move upward toward the cooler zones, where it may condense.

Also, barometric fluctuations cause oscillations in the vertical movement of soil gasses, leading to the vertical advective transport of water vapor. The vertical movement may be amplified when extreme barometric pressure conditions are present, such as during a hurricane when extremely low surface barometric pressures cause an upward air pressure gradient within the unsaturated zone.

BIBLIOGRAPHY

1. Evans, D., Rasmussen, T.C., and Nicholson, T.J. (2001). Flow and transport through unsaturated fractured rock: An overview. In: *Flow and Transport Through Unsaturated Fractured Rock*, 2nd Edn. D.D. Evans, T.J. Nicholson, and T.C. Rasmussen (Eds.). *Geophysical Monograph* 42. American Geophysical Union, Washington, DC.
2. Rasmussen, T.C., Baldwin, R.H., Dowd, J.F., and Williams, A.G. (2000). Tracer vs. pressure wave velocities through unsaturated saprolite. *Soil Sci. Soc. Amer. J.* **64**(1): 75–85.
3. Hillel, D. (1971). *Soil and Water: Physical Principles and Processes*. Academic Press, New York.

SPECIFIC GRAVITY

Todd Rasmussen
The University of Georgia
Athens, Georgia

The density of water and other materials varies with temperature as well as with the amount of dissolved and suspended matter within them. The specific gravity is the observed density of a material divided by a reference density, which is usually taken to be water at its maximum density, i.e., 1 kg/L. Thus, variations in the specific gravity are a direct measure of the variation in density. The specific weight of a material equals the product of its density and the gravitational constant. The specific gravity is important because it is needed to correct for the water density in wells to determine the freshwater head. Well construction also relies on the specific gravity of drilling muds to prevent the collapse of borehole walls in unconsolidated media.

The specific gravity is a measure of the density, or weight, of a fluid or solid. Denser materials have a larger specific gravity than do less dense materials. The specific gravity, ϱ, is defined using

$$\varrho = \frac{\rho}{\rho_o} \tag{1}$$

where ρ is the density (mass per unit volume) of the material and ρ_o is the density of a reference material.

The density of water at 4 °C, $\rho_o = 1$ kg/L, is generally the reference density for liquids and solids. Table 1 provides the specific gravity for a range of elements, including very light metals, such as aluminum and titanium, to very dense metals, such as gold and platinum. Note that for liquids and solids, the specific gravity equals the density because the standard density of water is 1 kg/L. The density of hydrogen is common for gasses.

The specific gravity of a pure substance can be affected by temperature. Table 2 provides the specific gravity and specific weight for water from 0 °C to 100 °C. Note that the specific gravity decreases more rapidly as the temperature increases.

Changes in physical properties of solids, such as their crystalline structure or porosity, can also result in a change in their density. The specific gravity can also be affected by any components present within the material, including dissolved gasses, liquids, and solids, as well as suspended gasses, liquids, and solids.

For example, water with high concentrations of suspended solids (such as drilling fluids) or dissolved solids (such as brines) has higher specific gravities than does clear water. Also, a column of air bubbles rising within a water column can reduce the fluid density.

The specific gravity is also related to the specific weight, $\gamma = \rho g$, where $g = 9.807$ m/s^2 is the gravitational constant. The specific weight commonly converts measured pressures to an equivalent hydraulic head because an increase

Table 1. Specific Gravity of Selected Elements

Material	Specific Gravity
Aluminum	2.70
Titanium	4.51
Iron	7.87
Copper	8.96
Silver	10.50
Lead	11.30
Mercury	13.50
Uranium	19.10
Gold	19.30
Platinum	21.50

Table 2. Variation of Pure Water with Temperature

Temperature (°C)	Density (kg/L)	Specific Weight (Pa/m)
0	0.99984	9,805.4
10	0.99970	9,804.1
20	0.99821	9,789.4
30	0.99565	9,764.3
40	0.99222	9,730.7
50	0.98803	9,689.6
60	0.98320	9,642.2
70	0.97778	9,589.1
80	0.97182	9,530.6
90	0.96535	9,467.2
100	0.95840	9,399.0

in depth, Δz, is accompanied by a concomitant increase in pressure, ΔP, which yields $\gamma = \Delta P/\Delta z$.

MEASURING THE SPECIFIC GRAVITY

A traditional technique for measuring the fluid density uses a *hydrometer*, which compares the weight of the fluid to a standard fluid. The hydrometer is allowed to float within the fluid, and the fluid density is determined from how much higher or lower the hydrometer level is compared with pure water (Fig. 1). Heavier fluids cause the hydrometer to rise relative to clear water because of buoyancy.

An alternative method for determining the fluid density is to employ a differential pressure transducer to measure the change in fluid pressure over a known depth. The resulting ratio between the measured pressure and the depth yields an estimate of the specific weight, $\gamma = \Delta P/\Delta z$.

Pressure transducers have recently measured the concentration and particle size distribution of suspended sediments (1). The particle size distribution is determined using Stokes's law, whereas the sediment concentration is found using

$$C = \frac{\varrho_f - 1}{\varrho_s - 1} \tag{2}$$

where ϱ_f is the measured specific gravity of the fluid and ϱ_s is the specific gravity of the suspended sediments.

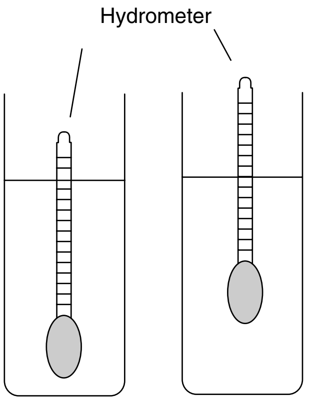

Figure 1. Hydrometer technique for measuring fluid density. Clear water causes the hydrometer to float at a constant level. Denser fluids cause the hydrometer to rise, whereas lighter fluids cause the hydrometer to sink lower in the water column.

IMPORTANCE OF SPECIFIC GRAVITY

Head Measurement

The use of water levels in wells to determine the hydraulic head within an aquifer is complicated when the density of the fluid within the borehole is not a constant (2). As noted in HEAD, changes in density affect the head according to the relationship:

$$H = z_o + \varrho \Delta z \tag{3}$$

where H is the freshwater head, z_o is the midpoint elevation of the screened zone within the aquifer, $\varrho = \gamma/\gamma_o$ is the specific gravity of the fluid within the well, $\Delta z = z - z_o$ is the height of the column of water within the monitoring well above the screened zone, and z is the water surface elevation in the well.

Drilling Fluids

The specific gravity of the drilling fluid is important during well construction. In loose materials, the process of drilling a well is complicated by the tendency of the unsupported walls to collapse.

Borehole walls can be temporarily supported until a casing is emplaced by using drilling fluids that are denser than water. The weight of the drilling fluid places an outward force on the borehole walls that helps to support them.

Heavier drilling fluids place a correspondingly heavier force on the borehole wall. One important drawback is the tendency of the drilling fluid to accumulate on the walls and within the formation, which causes a reduction in well yield if it is not removed during well development.

Lighter drilling fluids, such as foams and air, can also be used. In this case, the inflow of water from the formation into the borehole prevents the buildup of materials on the borehole wall, which eases well development. This technique is suitable in *consolidated media* where borehole collapse is not a problem.

Table 3. Specific Gravities of Selected Drilling Fluids

Drilling Fluid	Specific Gravity
Air	0.001
Mist	0.020
Stiff foam	0.100
Wet foam	0.200
Polymers	1.020
Polymers w/ NaCl	1.200
Bentonite clay	1.300
Polymers w/ CaCl	1.400
Bentonite w/ barite	2.200
Barite ($BaSO_4$)	4.200

Table 3 presents approximate specific gravities for a range of drilling fluids. Notice the wide range in specific gravities, from a low of 0.01 for air to a high of 4.2 for barite.

The values in Table 3 are approximate because the final density is a function of the relative concentrations of the liquids, solids, and gasses within the drilling fluid. Also note that the specific gravity of the air and mist are related to the density of water.

BIBLIOGRAPHY

1. Lewis, A.J. and Rasmussen, T.C. (1999). Determination of suspended sediment concentration and particle size distributions using pressure measurements. *J. Environ. Qual.* **28**: 1490–1496.
2. Driscoll, F.G. (1986). *Groundwater and Wells*, 2nd Edn. Johnson Filtration Systems, Inc., St. Paul, MN.

HOT SPRINGS

John E. Moore
USGS (Retired)
Denver, Colorado

Springs have been important in the social, cultural, and economic history of humankind. They have been an important source of water for humans. Our ancestors did not have drilling equipment, so in the absence of surface water, they had to rely on springs. Many early towns in Greece, Egypt, Mesopotamia, India, and China developed around springs. Many towns in the United States have taken their names from springs: Steamboat Springs, Colorado; Glenwood Springs, Colorado; Springfield, Missouri; and Rock Springs, Wyoming, are a few examples.

For centuries, hot springs and mineral springs were considered of medicinal or therapeutic value. Bottled “springwater” is still considered by many people of higher quality than ordinary tap water. This belief is at least partially responsible for the current high-volume sale of bottled water. Unfortunately, this association of springwater with purity is not based on fact. Springwater can contain higher concentrations of dissolved solids than local public water supplies, and springs become easily contaminated. Springs become contaminated because they are open, unprotected, and accessible to human beings and animals. Contamination can also occur from surface water flow or flooding into spring collection systems.

More than 3400 public water supply systems (1992) in the United States obtain part or all of their drinking water from springs. These systems provide drinking water for more than 7 million people. However, the average number of people served by a single public supply system using springs as a water source is small. Springs are an important component of the drinking water supply in many states. Public water supplies that use springs are more numerous in the western United States. Wyoming has 80 springs that are used for public water supply. A box used to collect water from springs is shown in Fig. 1 (1).

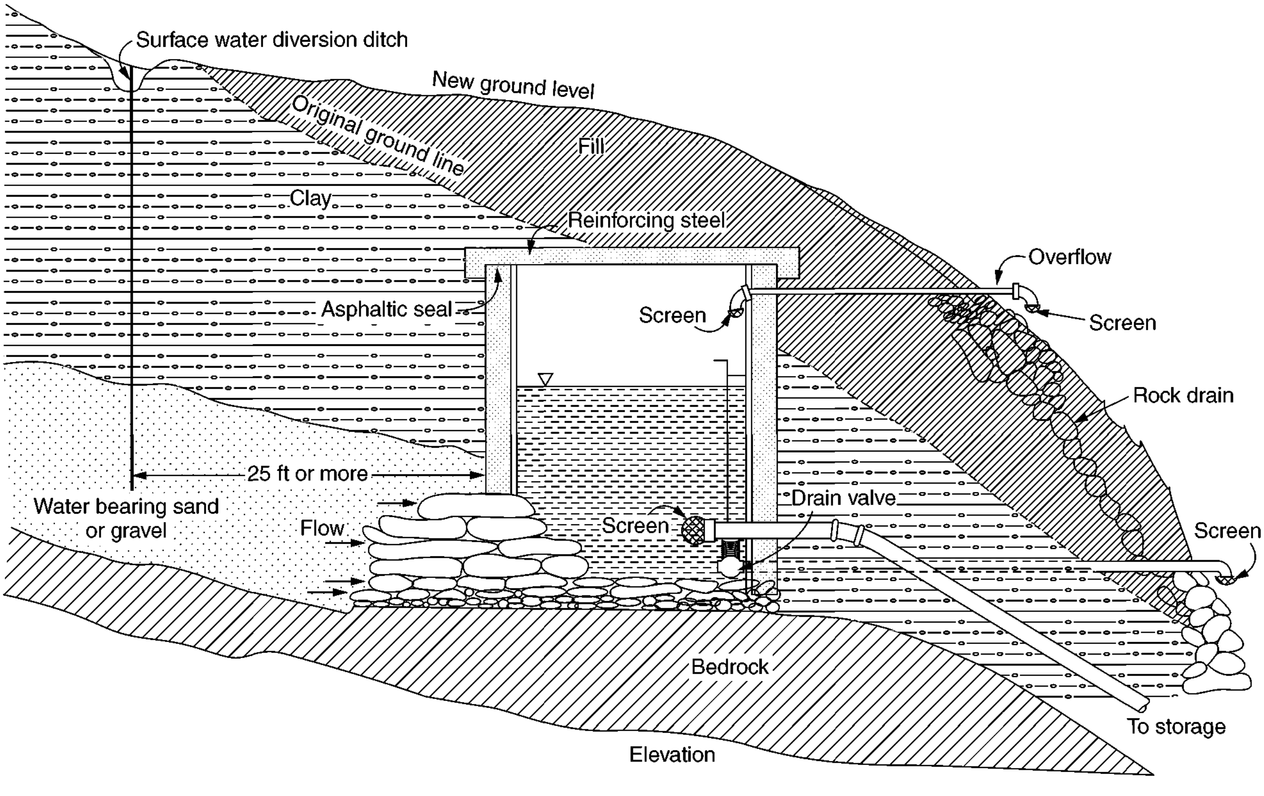

Figure 1. Diagram of a spring collection box (1).

A spring is a source of water that flows naturally from an aquifer or soil onto the land or into a body of surface water. Its occurrence depends on the nature and relationship of rocks (especially permeable and impermeable strata), on the position of the water table, and on topography.

Springs are generally classified as gravity springs and artesian springs. Thermal springs are typically considered a type of artesian spring. The vast majority of springs are gravity springs. Gravity springs are created by water that moves downgradient (downhill) and emerges at the surface. Two gravity springs, contact and fault springs are shown in Fig. 2. Artesian (nongravitational) springs occur when the potentiometric level of the groundwater system is above the land surface and water flows at the land surface under pressure either at the aquifer outcrop or from fractures or faults.

Springs are replenished by precipitation that recharges aquifers. The precipitation seeps into the soil and enters fractures, joints, bedding planes, or pore spaces in sand aquifers and sedimentary rocks. Springs occur when water flows through aquifers and discharges at the land surface through faults, fractures, or by flow along an impermeable layer. They can also occur where water flows from large orifices that result when the water dissolves carbonate rock (karst) and enlarges fractures or joints to create a passage (Fig. 2).

Springwater is sometimes forced to the surface along a fault from deep sources by thermal pressure gradients. Springs are associated with volcanism and fractures that extend to great depths in the earth's crust. Springs are a common feature in some hydrogeologic settings.

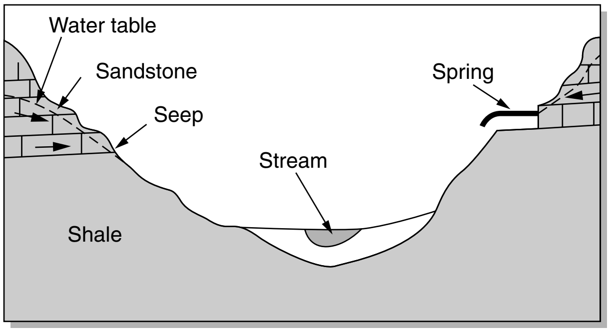

Contact spring

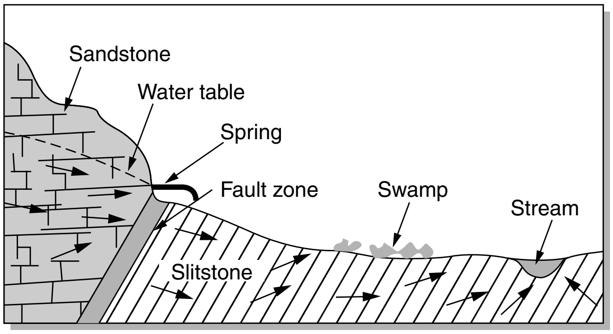

Fault spring

Figure 2. Contact and fault springs (2).

For example, springs are common in mountains and in watersheds that are underlain by fractured rock aquifers.

Springs can be regional (long flow paths) or local (short flow paths). Local springs are comparatively small, can have low flow, and are typically from shallow aquifers. The discharge from these springs often fluctuates either seasonally or in greater cycles, sometimes in response to local precipitation. Local aquifers are quickly recharged, and water movement through them is comparatively rapid, resulting in low water mineralization. Springs supported by local aquifers are more likely than regional flow springs to stop flowing periodically. Regional springs more typically have large discharges. Regional springs typically have nearly constant discharge and are more mineralized than local springs. Regional springs rarely stop flowing even during long droughts.

The discharge rate of a spring is a function of three main variables:

1. hydraulic conductivity (permeability) of the aquifer
2. area contributing recharge to the aquifer
3. quantity of recharge.

Springs are also classified based on hydrogeologic characteristics. The Water Resources Division of the U.S. Geological Survey recognizes eight principal types of springs based on hydrogeologic characteristics:

1. An artesian spring is a release of pressurized water from a confined aquifer at the aquifer outcrop or through an opening in the confining unit.
2. A depression spring is formed when the water table intersects a steeply sloping land surface. This type of spring is sensitive to seasonal fluctuations in groundwater storage and frequently disappears during dry periods.
3. A fracture spring (fault spring) is formed in fractured or jointed rocks. Water movement is through fractures, and springs form where the fractures intersect the land surface. This type of spring is particularly sensitive to seasonal fluctuations in groundwater storage and frequently disappears due to reduced flow during dry periods. Daily fluctuations of discharge of small springs are commonly the result of use of water by vegetation.
4. A contact spring occurs where a permeable water-bearing unit overlies a less permeable unit that intersects the ground surface.
5. A geyser spring is a periodic thermal spring resulting from the expansive force of superheated steam within constricted subsurface channels.
6. A perched spring occurs where infiltrating water discharges above the regional water table.
7. A seep spring discharges from numerous small openings in permeable material. These springs typically have very low discharge rates.
8. A tubular spring discharges from rounded channels (karst solution openings, lava tubes).

Springs occur in many sizes, types of discharge points, and locations with respect to topography. They occur in the highest elevations of mountainous areas and on valley floors. Many of the public water supply springs in the western United States are small; discharges are less than 5 gallons per minute. Springs vary in their physical and chemical compositions. They can be cold (50 °F) or hot (more than 10 °C above the air temperature). Shallow groundwater has a temperature within a few degrees of the ambient air temperature. Higher temperatures usually indicate deeper circulation. Thermal springs gain their temperature increases when the water comes in contact with recently emplaced igneous masses. Thermal and hot springs are due to deep-seated thermal sources and are classed as volcanic springs or fissure springs. Springs may occur singly or in groups that can include dozens of habitats in various sizes.

Springs may be highly mineralized, especially thermal springs and sometimes region springs that have a very long flow path. Thermal springs have pH values ranging from 7.2–7.6. Electrical conductance of spring water ranges from 2–10,000 milligrams per million.

A list of question that could be asked to evaluate spring conditions is shown below.

- Are there standing water and surface drainage around the spring?
- Is there deeply rooted vegetation in the spring collection area?
- Are there roots in the collection pits?
- Is there a lock-tight fence around the collection area?
- Is there a diversion channel capable of diverting surface water away from the collection area?
- Are there privies, septic systems, underground storage tanks, or barnyards near the spring?
- Is the collection chamber properly constructed (lockable)?

BIBLIOGRAPHY

1. U.S. Environmental Protection Agency. (1999). *Guidance Manual for Conducting Sanitary Surveys of Public Water Systems*.
2. Moore, J.E. et al. (1995). *Groundwater-a primer*. American Geological Institute, Arlington, VA.

READING LIST

Davis, S.N. and De Wiest, R.J.M. (1966). *Hydrogeology*. John Wiley & Sons, New York.

U.S. Environmental Protection Agency. (1997). *Investigation of Hydrogeologic Mapping to Delineate Protection Zones Around Springs-Report of Two Case Studies*.

Fetter, C.W. (2000). *Applied Hydrogeology*. Prentice-Hall, New York.

LaMoreaux, P.E. and Tanner, J.T. (2001). *Springs and Bottled Waters of the World*. Springer, New York.

Meinzer, O.E. (1923). *Outline of Ground Water Hydrology*. U.S. Geology Survey Water Supply Paper 494, p. 71.

Moore, J.E. (2002). *Field Hydrogeology*. Lewis, Boca Raton, FL.

Moore, J.E. et al. (in press). *A Field Guide to Evaluate Springs*. USEPA, Washington, DC.

Sanders, L.L. (1998). *A Manual of Field Hydrogeology*. Prentice-Hall, Upper Saddle River, NJ.

SQUEEZING WATER FROM ROCK

NASA Earth Science Enterprise
Data and Services

On December 16, 1811, an earthquake estimated at magnitude 7.0 to 8.0 on the Richter scale rocked the frontier town of New Madrid, Missouri, in the central United States. It was the first of three powerful earthquakes that would shake the central Mississippi River Valley that winter. Damages were reported as far away as Charleston, South Carolina, and Washington, DC. According to the U.S. Geological Survey, the New Madrid events were the most powerful earthquakes in recorded U.S. history.

Survivors of the New Madrid earthquakes reported not only intense ground shaking and land movement, as would be expected during an earthquake, but also an unfamiliar phenomenon: water and sand spouting up through fissures, or cracks, in the Earth's surface. In a letter published in *Lorenzo Dow's Journal*, New Madrid resident Eliza Bryan (1) wrote in 1816:

> ... the surface of hundreds of acres was, from time to time, covered over in various depths by the sand which issued from the fissures, which were made in great numbers all over this country, some of which closed up immediately after they had vomited forth their sand and water... 25 June 2003

Structures in the village of Bhuj, India, which was 80 kilometers (50 miles) from the earthquake epicenter, sustained considerable damage. (Image Copyright © James Mori, Research Center for Earthquake Prediction, Disaster Prevention Institute, Kyoto University)

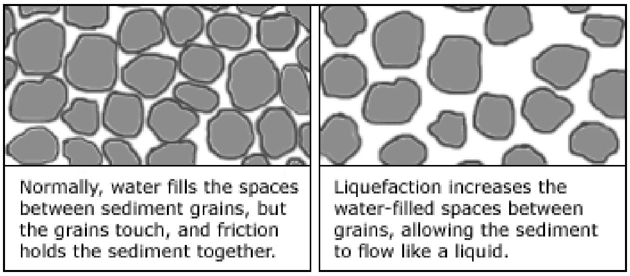

Normally, water fills the spaces between sediment grains, but the grains touch, and friction holds the sediment together.

Liquefaction increases the water-filled spaces between grains, allowing the sediment to flow like a liquid.

Modern earthquake researchers refer to this expulsion of water and sand from the ground as *earthquake dewatering*, which results from liquefaction. Liquefaction occurs when an earthquake vigorously shakes and compacts water-saturated sediments. As the compaction process displaces the water between the sediment pores, water and sand shoot upward and out of the ground.

"Imagine a cube full of sand and water. If you press it in from both sides (compressing and releasing it, then compressing it again), you build up what's called pore-water pressure," said Martitia Tuttle, geologist and consultant for the U.S. Geological Survey. "It's like shaking a coke can—when the pressure builds up and you release it, the fluid comes shooting to the surface." These fountains of water can sometimes shoot as high as 30 feet into the air, according to Tuttle (2).

Tuttle is part of a team that has been studying liquefaction in the New Madrid region in an effort to identify large prehistoric earthquakes in the geologic record. "Interpreting historic and prehistoric earthquake events involves a lot of uncertainty, due to the lack of data," Tuttle said. "We often don't know precisely where the earthquakes occurred or how big they were." What researchers needed was a modern-day event in a similar environment with which to compare the New Madrid earthquakes.

On January 26, 2001, a magnitude 7.7 earthquake near the village of Bhuj, located in the Kachchh region of northwestern India, provided scientists with just that analogue. Considered one of the two most damaging earthquakes in India's recorded history, the Bhuj earthquake killed about 20,000 people and caused an estimated $3.3 billion in damages.

The Bhuj earthquake struck the Kachchh region of northwestern India on January 26, 2001. The bottom image is a detail of the area indicated in red in the top map. (Images adapted from maps from the University of Texas at Austin General Libraries)

Earthquake Severity Table Adapted from What is Richter Magnitude? (A New Browser Window Will Open.)

Richter Magnitude	Earthquake Effects
Less than 3.5	Recorded but generally not felt
3.5 to 5.4	Often felt but rarely causes damage
Less than 6.0	Slight damage at most to well-constructed buildings; over small regions, can cause major damage to poorly constructed buildings
6.1 to 6.9	Can be destructive in populated regions up to 100 kilometers across
7.0 to 7.9	Major earthquake that can cause serious damage over areas larger than 100 kilometers across
8 or more	Great earthquake that can cause serious damage over areas several hundred kilometers across

According to Tuttle and co-authors (3) of a paper recently published in *Eos*, local residents and survey teams working in the area after the earthquake reported fountains of water and sediments surging from the ground during and immediately following the Bhuj earthquake. The water flow was significant enough to reactivate streams in previously dry river channels.

"The Bhuj earthquake provided an opportunity to study an earthquake that was very similar to the New Madrid event—it was a very large earthquake, and it occurred in an intraplate-like setting," said Tuttle. The New Madrid and Kachchh regions are both located more than 185 miles (300 kilometers) from active plate boundaries.

"Before the Bhuj earthquake, there was a lot of debate in the seismological community about the magnitude of the New Madrid earthquakes, but based on the similarity in intensity between the New Madrid and Bhuj events, there is a growing consensus among scientists that the New

Madrid earthquakes were probably larger than magnitude 7.5," said Tuttle. "The Bhuj earthquake is a modern event that helps us better understand the historic and prehistoric earthquake record in the New Madrid region."

But despite the unique study opportunity presented by the Bhuj earthquake, the area presented the researchers with some grave challenges. First, the Kachchh region of India is remote and difficult to reach. Second, it shares a politically sensitive border with Pakistan. These factors made it difficult for research teams to conduct field studies to verify liquefaction effects or the presence of water in some areas.

Bernard Pinty (4), research scientist at the Institute for Environment and Sustainability in the Joint Research Centre of the European Commission in Ispra, Italy, and colleagues found that water expelled during the earthquake could be detected in imagery from NASA's Multiangle Imaging SpectroRadiometer (MISR) satellite sensor.

The MISR instrument, which flies aboard NASA's Terra satellite, views the sunlit face of the Earth at nine widely spaced angles simultaneously. As it passes over the Earth, its nine cameras successively view each piece of the planet's surface at a spatial resolution of 300 yards (275 meters).

The change in reflection at different angles provides a means for distinguishing different types of land surface covers. Since bright soils in India's Rann of Kachchh (also known as the Rann of Kutch) reflect most of the sun's near-infrared radiation, and water bodies absorb near-infrared radiation, MISR is able to detect the contrast and indicate where dewatering occurred.

According to Pinty, these features made MISR an ideal tool for studying liquefaction in the Bhuj earthquake. "MISR's multi-angle capability was instrumental in exposing the presence of free water close to the Pakistani border, a region where it is difficult, if not impossible, to take ground measurements," he said.

These cracks in the ground surface north of Chobari, India were caused by liquefaction induced during the Bhuj earthquake. (Image Copyright © James Mori, Research Center for Earthquake Prediction, Disaster Prevention Institute, Kyoto University)

Most other satellite instruments only look straight down or towards the edge of the Earth. "It would take several more passes for another sensor to capture the entire region that the MISR imagery covers," said Tuttle. "The MISR images enabled us to see a large area both during and immediately after the earthquake, and we could see that dewatering was occurring over a huge region."

Scientists now know that the New Madrid earthquakes were not geologic flukes; similar events have occurred repeatedly in the geologic past. In fact, the central Mississippi Valley experiences more earthquakes than any other region in the United States east of the Rocky Mountains.

The New Madrid Seismic Zone, which includes parts of Tennessee, Kentucky, Missouri, Arkansas, Illinois, Indiana, and Mississippi, harbors a large liquefaction field. Scientists estimate the probability of a magnitude 6.0 or greater earthquake occurring in the New Madrid Seismic Zone within the next 50 years at 25–40 percent, according to the U.S. Geological Survey. But a similar event would have much more serious consequences today.

In 1811–1812, the central Mississippi Valley was a remote frontier. Today, the region is home to millions of people and includes the metropolitan cities of St. Louis, Missouri, and Memphis, Tennessee. In addition, most buildings in the region were not built to withstand earthquake shaking, as they often are in California and Japan.

Buildings and engineered structures are at much higher risk in areas where liquefaction occurs, due to increased ground motion. "The liquefaction process is extremely damaging to all sorts of engineered structures; it has

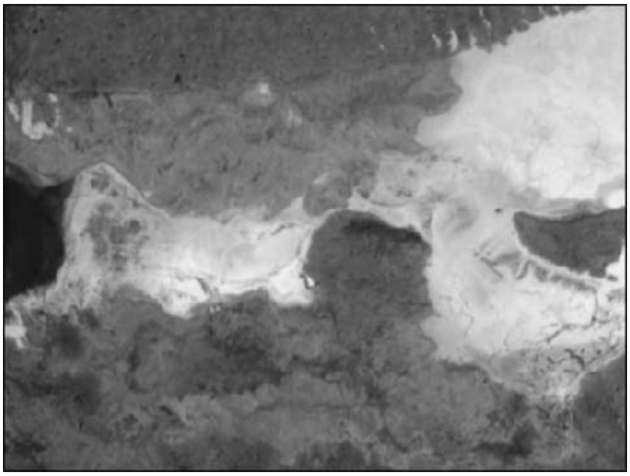

This false-color MISR image was acquired after the Bhuj earthquake, on January 31, 2001. The image shows numerous areas where groundwater flowed up to the surface, including within the Rann of Kachchh, as well as near the Indo-Pakistani border. (Image courtesy of NASA/GSFC/LaRC/JPL, MISR Team) Click here for more information.

caused major destruction in both Anchorage and San Francisco," said Pinty.

"The ground failure that resulted from liquefaction during the New Madrid earthquakes was severe. We're talking about vertical displacement of 3 to 6 feet (1 to 2 meters), and lateral displacement up to 33 feet (10 meters)," said Tuttle. "A recurrence of that type of event would have severe consequences for engineered structures."

According to Tuttle, in 1988 a magnitude 5.9 earthquake in Quebec, Canada, produced liquefaction. "Basements cracked, septic fields were disrupted, and people described water and sand shooting into their basements, out of their toilets, and into their bathtubs," she said. "And that was just a moderate-sized earthquake, not a big one."

MISR imagery, which is archived at the Langley Research Center DAAC, promises to be a boon to earthquake researchers seeking to minimize hazard risks.

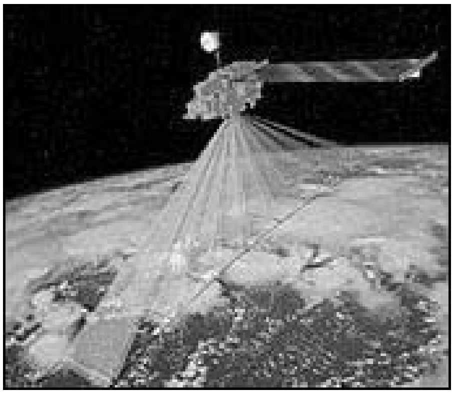

The MISR instrument views the sunlit Earth simultaneously at nine widely spaced angles, providing ongoing global coverage at high spatial detail. (Image courtesy of NASA)

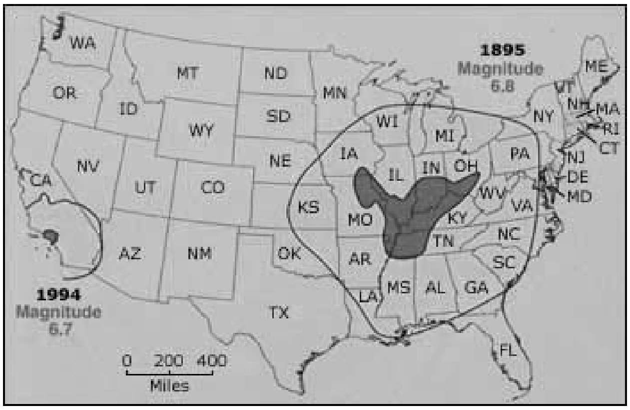

Although earthquakes in the central and eastern United States are less frequent than in the western United States, they affect much larger areas. This is shown by two areas affected by earthquakes of similar magnitude—the 1895 Charleston, Missouri, earthquake in the New Madrid Seismic Zone and the 1994 Northridge, California, earthquake. Red indicates regions of minor to major structural damage; yellow indicates areas where shaking was felt, but little or no damage to objects occurred. (Image courtesy of the U.S. Geological Survey)

"The interesting thing is that this project started as an 'after hours' sort of activity," said Pinty. "But it soon acquired its own momentum, giving us results that are of significant relevance to earthquake study applications. Our findings revealed that MISR can detect water and moist conditions on terrestrial surfaces, even under unfavorable spectral conditions."

"If we study modern earthquakes that produce liquefaction, we can better interpret the geologic record of liquefaction during past events. This helps us anticipate what is likely to happen in the future so that we can make informed decisions about reducing and mitigating hazards," said Tuttle.

For more information, visit the Langley Atmospheric Sciences Data Center DAAC. (A new browser window will open.)

"It's one of those things where people tend to think—if it hasn't happened during my lifetime, then it can't happen here," said Tuttle. "But the liquefaction field in the New Madrid region is very large. We're talking about a huge earthquake that could have a significant impact on society."

BIBLIOGRAPHY

1. Bryan, E. (1849). *New Madrid, Territory of Missouri*, March 22, 1816. In *Lorenzo Dow's Journal*. Published by Joshua Martin: 344–346. Quoted by The Virtual Times, The New Madrid Earthquake. Accessed June 5, 2003.
2. Tuttle, M.P. (2001). The use of liquefaction features in paleoseismology: Lessons learned in the New Madrid seismic zone, central United States. *Journal of Seismology*. **5**: 361–380.
3. Tuttle, M.P. et al. (2002). Observations and comparisons of liquefaction features and related effects induced by the Bhuj earthquake. *Earthquake Spectra*. **18**(Supp. A): 79–100.
4. Pinty, B. et al. (2003). Observing earthquake-related dewatering using MISR/Terra satellite data. *EOS, Transactions, American Geophysical Union*. **84**(5): 37–48.

This article contributed from *Distributed Active Archive Center (DAAC) Alliance: Supporting Earth Observing Science 2003*

For a copy, e-mail mailto:nasadaacs@nsidc.org?subject =DA 2003 annual request via EO.

STORAGE COEFFICIENT

Matthew M. Uliana
Texas State University—San Marcos
San Marcos, Texas

DEFINITIONS

A coefficient is a mathematical quantity that relates two or more variables in an equation describing some physical process. In equations describing the flow of fluids in a porous medium (e.g., groundwater in an aquifer), the **storage coefficient** is the quantity that relates the change in fluid potential (i.e., hydraulic head) in the aquifer to the change in the amount of water stored in the aquifer at a given point. In practical terms, the storage coefficient tells us the volume of water that can be removed from an aquifer for a given drop in hydraulic head.

There are three storage coefficients used in groundwater flow equations (the units of each are given in [brackets] following each definition):

Specific Storage (S_s) is the volume of water removed from a unit volume of a confined aquifer for a unit drop in hydraulic head [L^{-1}].

Storativity (S) is the volume of water removed from a unit area of an aquifer for a unit drop in hydraulic head; in confined aquifers, it is equal to the specific storage times the thickness of the aquifer, in unconfined aquifers, it is equal to the specific storage times the thickness of the aquifer plus the **specific yield** [–].

Specific Yield (S_y) is the volume of water removed from a unit volume of an unconfined aquifer for a unit drop in water level; this quantity is also defined as the volume of water that will drain by gravity from a saturated unit volume of an aquifer and is equivalent to the effective porosity [–].

Specific Storage (S_s)

The main equation that governs the flow of a fluid in a porous medium is given by

$$\frac{\partial q_x}{\partial x}+\frac{\partial q_y}{\partial y}+\frac{\partial q_z}{\partial z}=S_s\frac{\partial h}{\partial t} \tag{1}$$

where q = specific discharge (i.e., groundwater flow) in the x, y, or z direction [$L \cdot t^{-1}$]
h = hydraulic head [L]
t = time [t]
x, y, z = lengths in the principal directions in our coordinate systems [L]
S_s = specific storage [L^{-1}]

Darcy's law

$$q=K\frac{dh}{dl} \tag{2}$$

where K = hydraulic conductivity [$L \cdot t^{-1}$]
dh/dl = hydraulic gradient, or change in head with length [–]

can be substituted for q to get the expanded version of the flow equation:

$$\frac{\partial\left(K_x\frac{\partial h}{\partial x}\right)}{\partial x}+\frac{\partial\left(K_y\frac{\partial h}{\partial y}\right)}{\partial y}+\frac{\partial\left(K_z\frac{\partial h}{\partial z}\right)}{\partial z}=S_s\frac{\partial h}{\partial t} \tag{3}$$

Equations 1 and 3 are specific versions of the **continuity** (or **conservation of mass) equation**,

$$I-O=\Delta S \tag{4}$$

where I = flux into a volume of a system [$L^3 \cdot t^{-1}$]
O = flux out of a volume of a system [$L^3 \cdot t^{-1}$]
ΔS = change in the amount stored in that same volume of the system [$L^3 \cdot t^{-1}$]

The left side of Eq. 3 corresponds to the left side of the continuity equation and describes the divergence of flux at a point in the aquifer. In other words, it describes the groundwater flow vector resolved into x, y, and z components (see Fig. 1).

The right side of Eq. 3 corresponds to the right side of the continuity equation and describes the change in storage at the given point in the aquifer. The term $\partial h/\partial t$ refers to the change in hydraulic head, with respect to time, at the same point in the aquifer. The term S_s is the **specific storage**, and it relates the changes in head in the aquifer to the change in the amount of water stored in the aquifer at a certain point. In other words, it is a quantity that tells us the volume of water that can be removed from a unit volume of a confined aquifer for a unit drop in head.

CONFINED AQUIFERS—A CONCEPTUAL UNDERSTANDING

Consider a confined aquifer that consists of a layer of sand overlain by an impermeable layer of clay that acts as a confining layer. If a well is drilled into the aquifer and water is pumped from the well, the heads around the well will lower and create a cone of depression in the potentiometric surface (Fig. 2). The more water that is pumped from the well, the larger the cone of depression. In the example shown in Fig. 2, the pumping well has removed a volume of water (represented by the volume in the barrel), and in response, the water levels in

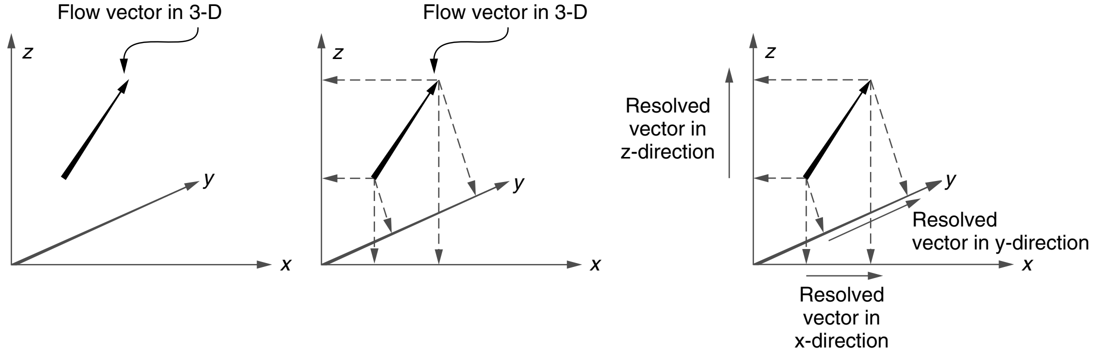

Figure 1. Flow vector resolved into component vectors in three dimensions.

the well and in the surrounding aquifer have dropped. However, the saturated thickness of the aquifer itself has not changed (i.e., the aquifer is still full to the base of the confining layer). What is the source of this water? The conceptual answer deals with elastic deformation and compressibility of the aquifer materials and the water.

This aquifer is made up of sand grains and pore spaces that are full of water. If there is no confining layer and the aquifer is simply recharged by water infiltrating from above, the pore spaces are full of water, and the pore fluid pressure is equal to atmospheric pressure plus the overlying column of water. In this case, the volume of water in the pore space is equal to the volume of the pore space.

However, if the aquifer is confined and an additional component of pressure is added to the water, that pressure will force a greater volume of water into the pores of the sand. It does this by compressing the sand grains slightly, pushing the sand grains around so that pore volume is increased, and compressing the water itself slightly. Increasing the pore fluid pressure causes the aquifer to undergo **elastic** (i.e., recoverable) **deformation**, so that a subsequent decrease in fluid pressure will cause an associated decrease in volume (Fig. 3).

This is the source of the additional volume of water. Pumping water from the confined aquifer lowers the excess fluid pressure (i.e., the pressure head above the top of the aquifer), causing the aquifer to 'relax' and yield water that is stored by the excess pressure.

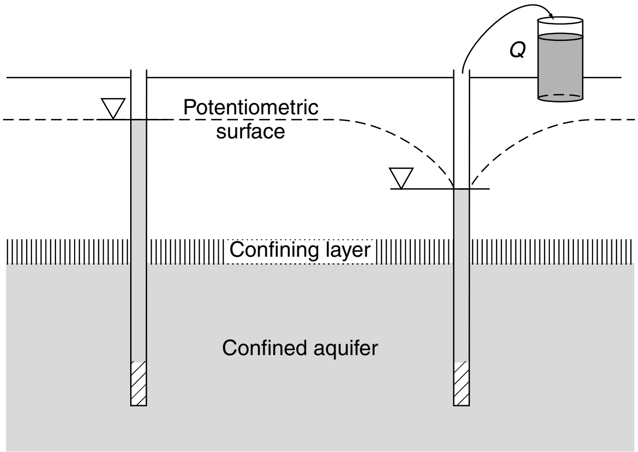

Figure 2. Pumping well in a confined aquifer.

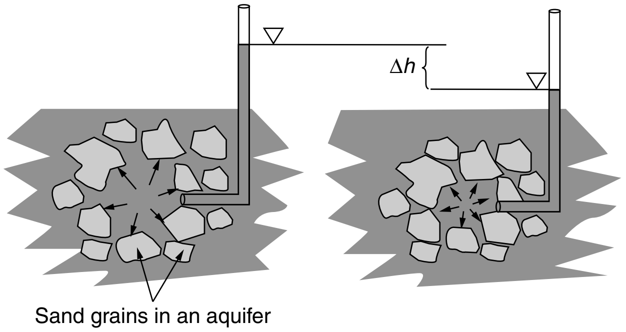

Figure 3. Elastic deformation in a sand aquifer.

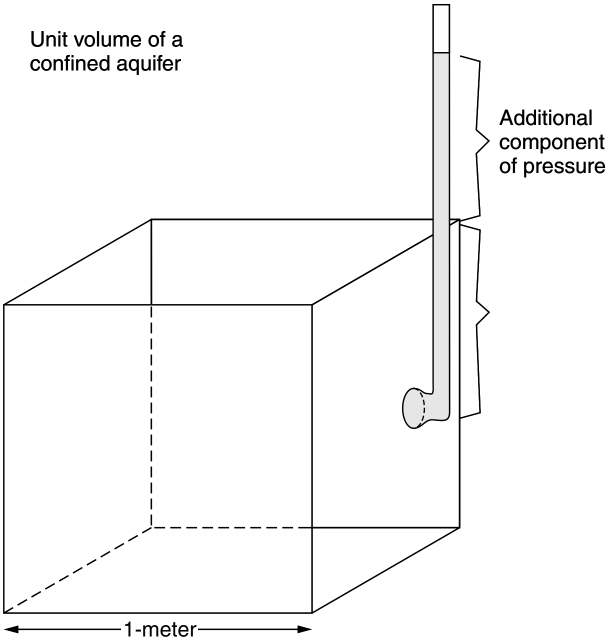

Figure 4. Unit volume of a confined aquifer.

CONFINED AQUIFERS—A MATHEMATICAL UNDERSTANDING

Consider a unit volume of a confined aquifer. This unit volume is saturated with water, and the water has an additional component of pressure that would cause the water level in a manometer to rise above the top of the cube (Fig. 4). If a volume of water is removed from the cube, it will reduce the additional component of pressure and lower the head in the cube. Removing more water will cause more head loss; less water will cause less head loss. Therefore, there is a specific volume of water that, if removed, will cause a unit drop in head. This volume of water that results in a unit drop in head in the unit volume of aquifer is the **specific storage**.

The specific storage is also called the **elastic storage coefficient**; it deals only with the amount of water removed from elastic storage. The specific storage is a function of the compressibility of the aquifer matrix (i.e., the material that makes up the aquifer) and the compressibility of the water and is strictly defined as

$$\mathrm{S}_s = \rho_w g(\alpha + n\beta) \tag{5}$$

where ρ_w = density of water [M/L^3]
g = gravitational constant [L/t^2]
α = compressibility of the aquifer matrix [1/M/Lt^2]
n = porosity [−]
β = compressibility of water [1/M/Lt^2]

STORATIVITY IN CONFINED AQUIFERS

The specific storage actually refers to an infinitesimal point in the aquifer. However, in real aquifers, the thickness of the aquifer influences how much water the aquifer will yield; a thicker aquifer will transmit and yield more water than a thinner aquifer. Multiplying the specific storage by the saturated thickness of the aquifer results in **storativity** (S):

$$S = S_s b \tag{6}$$

where b = saturated thickness of the aquifer [L].

In other words, the storativity describes the volume of water removed from a unit *area* of aquifer for a unit drop in head.

UNCONFINED AQUIFERS AND SPECIFIC YIELD

In a confined aquifer, all the water pumped out through a well is provided by the release of water from elastic storage. In an unconfined aquifer, the same thing occurs—the changes in head result in a release of water from elastic storage in the aquifer. However, in an unconfined aquifer, the pumping well is dewatering a section of the aquifer (Fig. 5), and this part of the aquifer's response adds an additional volume of water to the volume withdrawn from elastic storage.

This additional volume of water is released from the pore spaces and is due to the **specific yield** of the aquifer. Specific yield is the volume of water that drains from a rock under the force of gravity and is directly related to the effective porosity of the aquifer matrix. The volume of water released from a unit volume of aquifer due to dewatering of the aquifer is therefore the unit volume times the specific yield.

So, in unconfined aquifers, the volume of water released from a unit volume of aquifer is a function of both the specific yield (i.e., the water removed from the pore space) and the specific storage (the water released from elastic storage in the aquifer). Therefore, in an unconfined aquifer, the storativity of the aquifer is the sum of the specific yield and the thickness times the specific storage.

$$S = S_y + bS_s \tag{7}$$

However, the magnitudes of these parameters are drastically different: specific yield values in an aquifer generally range from 1% to 50%, or 0.01 to 0.5; specific storage values generally range from 1×10^{-3} to 1×10^{-6}, or 0.001 to 0.000001, which means that, in an unconfined aquifer, the volume of water provided by release from elastic storage is orders of magnitude less than that provided by dewatering the pores; so the specific storage is generally considered negligible, and unconfined storativity is a function only of the specific yield.

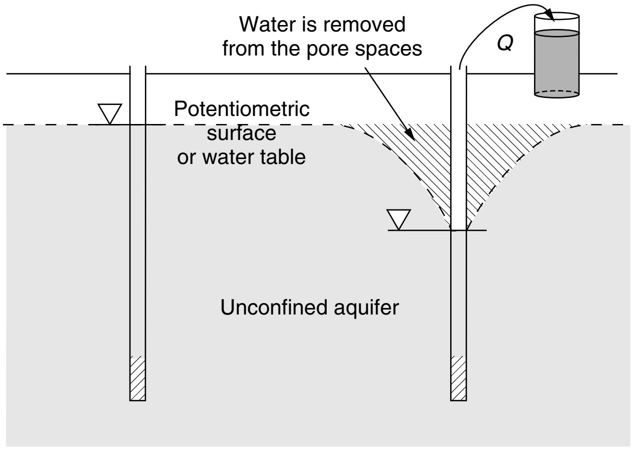

Figure 5. Pumping well in an unconfined aquifer.

QANATS: AN INGENIOUS SUSTAINABLE GROUNDWATER RESOURCE SYSTEM

Bahram Saghafian
Soil Conservation and
Watershed Management
Research Institute
Tehran, Iran

A qanat system consists of an underground tunnel that conveys water by gravity from highland aquifers to the ground surface at lower elevations. Blended with ingenuity, its purpose is simply to collect and bring underground water to the surface for use. Qanat, invented and originated in Iran more than 2700 years ago, is considered a hydraulic engineering system of an Old World era providing water to settlements since its invention. The overwhelming climate of Iran is arid and semiarid; two thirds of the country area in central, southern, and eastern parts receive an average of about 150 mm of rainfall annually. Yet, Iran is considered a farming country where diverse crops are grown. This remarkable achievement in agricultural development of arid lands has materialized partly by building qanats to tap underground water in regions where surface water is lacking or unreliable during seasons of high water demand. Numerous urban and rural qanat-watered settlements are vivid examples that not all civilizations were rooted by the sides of rivers. Qanats have helped to transform barren lands into productive assets. Thanks to its simple and efficient technique, the idea behind this sustainable water resource system was passed from one generation to the next and traversed geographic boundaries. Qanat's use is not confined to Iran and has been adopted in many regions across the world.

QANAT DESCRIPTION

Britannica characterizes a qanat as, "A qanat taps underground mountain water sources trapped in and beneath the upper reaches of alluvial fans and channels the water downhill through a series of tunnels, often several kilometres long, to the places where it is needed for irrigation and domestic use ...." The dictionary of

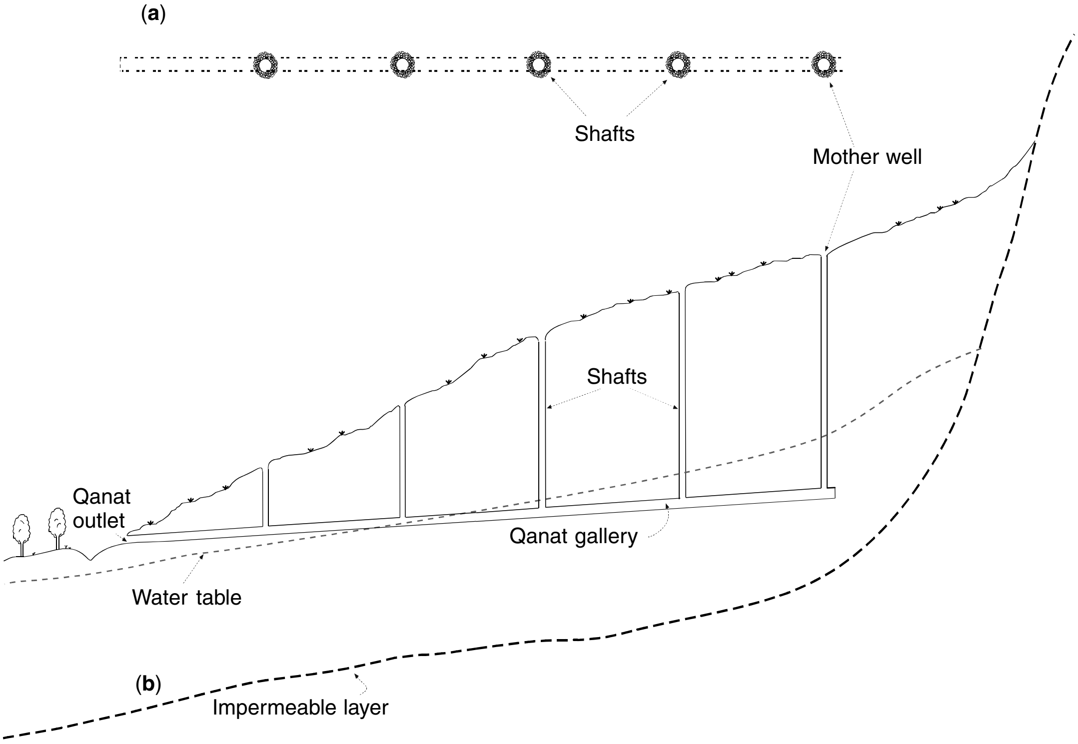

Figure 1. Schematic plan (**a**) and cross section (**b**) of a qanat system.

Islamic Architecture defines qanat as a "subterranean canal system usually used to bring water some distance from a river or mountains." More technically, Lightfoot (1) describes a qanat as "a form of subterranean aqueduct, or subsurface canal, engineered to collect groundwater and direct it through a gently sloping underground conduit to surface canals which provide water to agricultural fields."

As shown in Fig. 1, a qanat consists of a gently sloped underground tunnel, or gallery, penetrating the upland aquifer at one end and appearing in the ground surface at the other end. The gallery, is also vertically connected to the ground by several wells, or shafts, dug along its length, often at 20–30 m to a maximum of 100 m intervals, providing ventilation and access during construction and maintenance. Put differently, a qanat consists of a series of vertical shafts in sloping ground, inter-connected at the bottom by a tunnel with a gradient flatter than that of the ground. The excavated soil is piled up around the entrance of the wells to give some protection in times of storm surface runoff. The first shaft, called the mother well or main shaft, is usually sunk into the alluvial fan to a level below the groundwater table. The line of qanat wells looks like anthills in a distinctive plan view. The typical cross section of a qanat gallery is 1.5 m high and about 1 m wide, which is just sufficient to allow men and boys to do the digging and cleaning.

Qanats can be classified into hill and river-based. Hill qanats have their main shafts based in the aquifers located in hill slopes. This type of qanat is not very long. If this type of qanat taps a deep aquifer, it will enjoy a relatively constant discharge. A river-based qanat is normally recharged from a seasonal river, stretching over a long distance where the tunnel meets the ground surface.

Other terms for qanat, pronounced *kanat*, as it is called in different parts of the world, are karez (e.g., in Afghanistan and eastern Iran), foggora (e.g., in Algeria), hattara (e.g., in Morrocco), and falaj (e.g., in Oman).

HISTORY AND GEOGRAPHIC EXTENT OF QANATS

Written records and recent excavations prove that qanats first appeared in ancient Iran (then Persia). Wulff (2) describes the origin of qanat: "As early as the seventh century B.C., the Assyrian king Sargon II reported that during a campaign in Persia he had found an underground system for tapping water in operation near Lake Urmia. His son, King Sennacherib, applied the secret of using underground conduits in building an irrigation system around Nineveh, and he constructed a qanat on the Persian model to supply water for the city of Arbela." The earliest reports of qanats indicated that the water for the city of Ula northwest of Lake Urmia, located in northwest of Iran, was provided by qanats some 700 B.C. (3). English (4) reports that, based on early mining, the mountains of Kurdistan in western Iran, eastern Turkey, and northern Iraq were the birthplace of the qanat more than 2500 years ago.

The technique was then widely diffused during the expansion of Achaemenid dynasty, ruling the Persian Empire, beginning sometime in the fifth century B.C. Qanats were constructed in Mesopotamia to the shores of the Mediterranean to the west, in Egypt and the Arabian Peninsula to the south, and in Afghanistan, Central Asia, and western China to the east. The Arabs introduced the idea further to the west during the second diffusion phase in the eighth century A.D. The greatest developments of qanat technology outside the Persian core area are in southern Morocco (Marrakesh and the Tafilalet) and southern Algeria (5). Wulff (2) reported the excavation of several particularly interesting qanats constructed by Arab rulers of early medieval times in the Middle East. These include a small qanat built by the caliph of Damascus to supply water for his palace in 728 A.D. A century later, presumably with the aid of Persian engineers, the caliph Mutawakkil in Iraq constructed a qanat that brought water to his residence at Samarra from the upper Tigris River 300 miles away.

Qanats are reported to exist in Spain, Sicily, Cyprus, the Canary Islands, and in South American countries such as Mexico, Peru, and Chile. The Madrid qanat system in Spain, bringing water for urban use, still functions. The qanat systems of Tehuacan and Parras in Mexico are believed of Spanish origin, but those of Parras (Coahuila) and Pica (Chile) may be pre-Columbian (6). English (5) and Beekman et al. (7) discuss scenarios of the diffusion of qanat technology throughout the world.

CONSTRUCTING A QANAT

The laborious construction of a new qanat is a task to be engineered by an expert, who usually heads a team of diggers. The diggers, called *muqannis*, are responsible for building and repairing qanats. The expert must decide on a number of key issues prior to qanat construction: the site of the mother well and the alignment and the slope of the underground gallery. In his work, the expert uses simple instruments such as a long rope and a level. As the mother well of a qanat is usually dug in an alluvial fan at the foot of a mountain or hillside, the expert watches for signs of seepage on the ground, slight variations in the vegetation, and available groundwater. Once the site is chosen, one or sometimes more trial wells are sunk. The muqannis then start digging the trial well, normally 1 m in diameter. They use a windlass at the top and haul up the soil excavated by their teammates. They continue the digging and spoil the material around the rim of the well until the water table is struck. The expert then should determine whether the aquifer water has acceptable water yield. If so, he measures the distance from the mother well to the qanat outlet, where water is planned for use. Once the general alignment of the qanat is determined, the expert calculates the slope of the qanat gallery. This slope must allow the water to flow, yet prevent erosion of the bottom and walls of the gallery. The slope for short qanats may run from 0.5 to 2 in 1000, whereas longer qanats usually have gentler slopes. In steep land, the slope of the gallery is maintained in a looping course. The work then proceeds from the outlet of qanat by excavating the tunnel back toward the mother well. Vertical shafts at a distance of 30 m, or so, are dug to provide ventilation and access. For shafts deeper than 100 m, the muqannis set up another windlass halfway within the shaft for a two-stage haul up of spoil. The muqannis maintain a straight course by sighting along a pair of burning oil lamps. The task requires great skill and endurance. Figure 2 shows a top view of qanat during construction.

Figure 2. A top view of a qanat during construction.

Muqannis may encounter a number of dangers in their work. Soft soil in the tunnel can cause collapse of the roof. In such cases, the tunnel cross section is braced with baked clay collars to control cave-ins (Fig. 3). Another danger is air low in oxygen. Any change in the burning of the illuminating oil lamps is an indication of such danger. And last, but not least, is the hazard during the breakthrough into the mother well, which must be emptied.

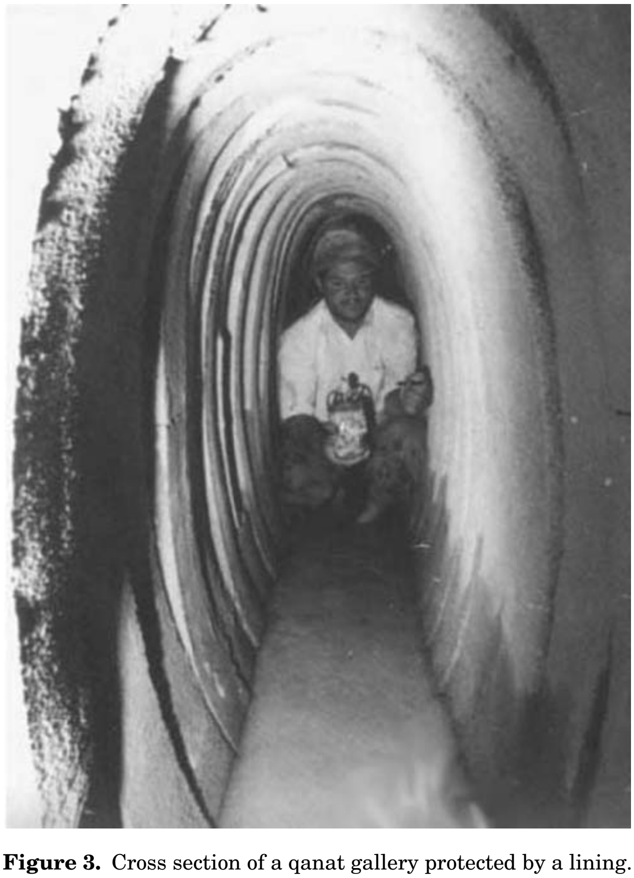

Figure 3. Cross section of a qanat gallery protected by a lining.

Muqannis are highly respected for their courage, hard work, and know-how. They are paid well for their work, and older muqannis are considered blessed or lucky.

QANATS AND WATER REGULATIONS

In Iran, the allotment of water delivered by qanats is regulated by certain laws, which are rooted in a traditional system of water rights. Normally a water master or bailiff, appointed by a qanat owner(s), supervises the water distribution according to each farmer's time share. As Wulff (2) has noted, "the bailiff is guided by an allocation system that has been fixed for hundreds of years. For instance, three hamlets in a region west of Iran, still receive the shares that were allotted to them in the 17th century by the civil engineer in the reign of Shah Abbass. The hamlets of Dastgerd and Parvar in central Iran are entitled to eight shares apiece. These allocations are built into outlets from the qanat distribution basin." English (4) provides a detailed overview of qanat ownership, the complex water distribution system, and, through some examples, how fragmentation of qanat water rights has taken place. He makes a note of strict and unforgiving adherence to communal methods of water rotation and upkeep of the water supply system serving to maintain the social cohesion of qanat-watered villages.

QANATS: COSTLY AND TIME-CONSUMING

Qanats are labor-intensive and time-consuming to build. Both the cost and time to build a new qanat varies with topography, depth of the water table, underground soil type, length of the qanat, type of ownership, availability and experience of muqannis, and local conditions. Thus, reports differ in figures provided for the cost and duration of construction of qanats.

Noting that qanats are long-lasting water resources systems if properly maintained, Wulff (2) believes that the agricultural production sustained by qanat water repays the investment in construction and maintenance by 10% to 25% depending on the size of the qanat, the water yield, and the crop type. Wulff's assessment of the cost in the 1960s showed that a 6-mile long qanat required US$13,500 to US$34,000 to build. The cost runs to about $90,000 for a qanat 10 to 15 miles long.

English (8), on the other hand, reported that it took 27 years to construct a 1-km qanat in Kerman, in central Iran, with a mother well 45 meter deep, due to three changes in ownership. Another 3-km qanat with two mother wells, 50 and 55 m deep, and a bifurcated tunnel took 17 years to build. The costs of these two qanats, respectively, were approximately $10,000 to $11,000 per km in the 1960s. As another example, the cost of constructing a 40-km long qanat near the city of Kerman with a mother well 90 m deep was roughly $213,000 when completed in 1950 (4). Such expenses adjusted for inflation can be raised only through government subsidies or large corporations in modern Iran. Due to high construction investment, qanat owners often try to increase the yield. They may extend branches from the main tunnel to reach additional aquifers or excavate the floor of the existing tunnel to tap water at a deeper level (2).

Note that, although agricultural production cannot repay the investment and maintenance costs of new qanats in some regions, existing qanats are still the second water supply system in volume providing for drinking and domestic use in Iran. Keeping the qanat cost and the enormous building effort in mind, individual farmers are unwilling and unable to fund the construction of new qanats or even maintain old ones. Thus, farmers have turned to drilled wells as the more affordable alternative.

QANAT VERSUS DRILLED WELL

People in arid regions rely on underground aquifers for their water supply. The aquifers can be tapped by a qanat, or alternatively by a drilled well, and the water is directed to an area where it is consumed. The volume of water produced by a qanat depends upon several factors: extent, type, yield, and recharge rate and area of the aquifer. Comparing a drilled well with a qanat system, there are advantages for each. The following compare the two systems:

- As history proves it, qanats are sustainable renewable water resource systems. However, they cannot meet today's increasing demand. Wells are the alternative at the expense of inducing negative water balance in the aquifers they exploit.
- Qanats need no extra energy to harvest water, whereas wells need diesel or electric power to bring water to the surface.
- Qanats can be built and maintained using low-technology tools and material locally available.
- Qanats can be used as a passive cooling system in conjunction with a wind tower (9).
- Qanats reinforce social cooperation, an important issue in traditional societies.
- Wells are dug in a short time, whereas qanat construction is slow, investment-intensive, and quite time-consuming. However, technology can be put to work to accelerate qanat construction by cutting the time and the cost.
- Only a portion of a qanat's water is used during nonirrigation season, unless conserved by other means. During periods of low water use in fall and winter, watertight gates can seal off the qanat outlet damming up the water. In spring and summer, night flow may be stored in small ponds for daytime use (10). A well pump can be switched off when water is not needed.
- A typical qanat's discharge is less than that of pumped wells. The latter are favored to meet growing demands, though, in a nonsustainable manner.

Experts in Iran now suggest an integrated approach of conjunctive use of qanats and wells. The location of a well must be optimized with respect to an existing qanat recharge area.

STATISTICS OF QANATS

No one knows how many qanats were ever built, how many are still in operation, and how much water they deliver. Available reports vary in numbers, albeit they have collected qanat statistics in different times. It is estimated that, as of the 1960s, between 40,000 to 50,000 qanats had been built in Iran during its history, of which some 27,000 running qanats were still used (11). Wulff (2) reported that some 22000 qanats in Iran, totaling 170,000 miles of underground conduits built by manual labor, had a total yield of 19,500 cubic feet of water per second, which amounts to 17 billion cubic meters of water annually. Two recent reports prepared in Iran put the total number of qanats, as of 1995, between 25,000 and 33,000, but it is unclear whether all are currently used. The reports, however, agree on the total volume of water provided by qanats at around 9 billion cubic meters. Having accepted this figure and the figure given by Wulff in 1968, an alarming decline of about 50% in water yield of qanats is inferred, which must be partly due to an excessive number of wells causing underground water tables to drop and affect qanat yield. The declining groundwater tables and increasing number of abandoned qanats have shrunk the share of water supplied by qanats to less than 10% of total use.

The discharge of individual qanats varies widely with the location and, in some cases, time of the year. Qanat discharges of a few liters to as much as 1 cubic meter per second have been reported.

There are some unique qanats in Iran described as follows:

1. Vazvan Qanat in Meimeh, Isfahan province, is 1800 m long. At a distance of 600 meter from its outlet, a stone dam has been built to store the water. A total of six gates have been installed at different elevations to control the flow of water, depending on the time and amount needed. This qanat system is essentially an underground dam.
2. The well-known and unique Moun Qanat in Ardestan, Isfahan province, is a twin-gallery qanat system, the only one of its kind reported in the world. The two galleries are independent conduits, about 3 m apart in elevation. A thin impermeable layer is believed to separate the two qanat conduits. The water in the layers is not mixed, as the upper conduit is bent around the shafts. The qanat is about 2 km long, and its total discharge waters 25 ha.
3. Gonabad region, in Khorasan province northeast of Iran, is known for some of the oldest qanats in Iran. Reports indicate that one qanat has a mother shaft with a depth of 140 meters and a length of about 70 kilometers. Another qanat is reported with a 270-m mother well.

BIBLIOGRAPHY

1. Lightfoot, D.R. (1997). QANATS in the Levant: Hydraulic technology at the periphery of early empires. *Technology and Culture* **38**(2): 451.
2. Wulff, H.E. (1968). The qanats of Iran. *Sci. Am.* **April**: 94–105.
3. Laessøe, J. (1951). The irrigation system at Ulhu, 8th century BC. *J. Cuneiform Stud.* **5**: 21–32.
4. English, P.W. (1998). Qanats and lifeworlds in Iranian plateau villages. *Yale F&ES Bull.* **103**: 187–205.
5. English, P.W. (1968). The origin and spread of qanats in the Old World. *Proc. Am. Philos. Soc.* **112**: 170–181.
6. Humlum, J. (1965). Underjordiske vandingskanaler: kareza, qanat, foggara. *Kultergeografi* **16**: 81–132.
7. Beekman, C.S., Weigand, P.S., and Pint, J.J. (1999). Old World irrigation technology in a New World context: qanats in Spanish Colonial Western Mexico. *Antiquity* **73**(279): 440–446.
8. English, P.W. (1966). *City and Village in Iran*. University of Wisconsin Press, Madison, WI.
9. Bahadori, M.N. (1978). Passive cooling systems in Iranian architecture. *Sci. Am.* **February**: 144–154.
10. Beaumont, P. (1989). The qanat: a means of water provision from groundwater sources. In: *Qanat, Kariz, and Khettara, Wishbech*. P. Beaumont, M. Bonine, and K. McLachlan (Eds.). Menas Press, pp. 13–31.
11. Goblot, H. (1979). *Les Qanats: Une Technique d 'Acquisition de l 'Eau*. Mouton, Paris.

LYSIMETERS

Kristofor R. Brye
University of Arkansas
Fayetteville, Arkansas

Lysimetry is a measurement technique that employs sampling devices called lysimeters to measure the movement of water and solutes through a porous medium. Lysimeters have most commonly been used to quantify water movement (i.e., drainage) through soil and typically consist of a porous plate surface of varying pore sizes through which water moves from the soil into a collection reservoir.

BRIEF HISTORY OF LYSIMETERS

According to Joffe (1), lysimeter studies were initiated in the late eighteenth century to address the fate of precipitation in soil. Many lysimeter studies then shifted to consider leaching of soil and precipitation-deposited constituents, such as ammonium and sulfur (1). Gradually, lysimeters were adopted to investigate fertilizer losses from agricultural and nonagricultural soil (1).

Early lysimeter research involved greatly disturbed soil profiles, unrepresentative of natural soil properties. Lysimeter tanks were dug and refilled with soil after construction was complete (1). For this reason, the validity of drainage and leaching measurements were severely questioned, as they are today. Nonetheless, many technological and methodological advancements have been made since the late 1700s, but the same fundamental question remains today as it did then: how to measure natural water movement through undisturbed soil?

DRAINAGE LYSIMETERS

Two general types of lysimeters exist for monitoring water and solute fluxes through soil, weighing and nonweighing lysimeters. Weighing lysimeters are large-scale sampling devices typically used for measuring evapotranspiration, drainage, and solute leaching. Inputs and losses of moisture are evaluated on the basis of changes in mass in a containerized column of soil within the lysimeter (2). Weighing lysimeters are expensive, labor-intensive to construct, and usually involve filling the lysimeter by reconstructing a disturbed soil profile (3,4), although large weighing lysimeters that have undisturbed soil columns have been constructed, such as the monolith lysimeters near Coshocton, Ohio. However, nonweighing lysimeters are better suited for less burdensome, replicable field measurements of *in situ* water and solute fluxes.

Nonweighing lysimeters are well adapted to monitoring drainage and leaching from soils. Nonweighing lysimeters, also known as drainage or pan lysimeters, are typically much smaller than weighing lysimeters. An advantage of a nonweighing pan lysimeter is that replicable field measurements can be obtained below undisturbed soil profiles. As a result, site-specific information can be gathered relative to the mass flux of water and leachate concentrations of soluble organic and inorganic chemical species (3). Coupling drainage measurements with solute concentrations provides a means of estimating solute mass fluxes through an undisturbed soil profile.

DRAINAGE MEASUREMENTS USING NONWEIGHING LYSIMETERS

The collection of leachate through the porous surface of a lysimeter depends on two conditions: (a) minimal discontinuities in the flow path of moving water and (b) a driving force for water movement. Maintaining a capillary connection between the soil and the porous surface of the lysimeter is imperative for high-quality results. Neglecting to minimize potential sources of error may cause ponding, preferential, and bypass flow around the lysimeter (3).

Drainage lysimeters have been designed to operate where a variety of forces act on soil water. Drainage and solute fluxes are measured and replicated in the field by some variation of a nonweighing pan lysimeter operating under tension (i.e., tension lysimeters) or tensionless (i.e., zero-tension lysimeters). In soil or any other porous material, the bulk of the water moves in response to gradients in gravitational and/or matric potentials. Zero-tension lysimeters rely on the force of gravity for water movement through the lysimeter's porous plate surface. The effect of a matric potential gradient is created by applying suction to a tension lysimeter to quantify drainage and leaching.

Both types of nonweighing lysimeters have several significant disadvantages. Nonweighing lysimeters are commonly operated after the lysimeter has been filled with soil, similar to weighing lysimeters. As a result, soil horizonation, structure, pore size distributions, and many other physical properties no longer reflect natural soil conditions. Therefore, using a disturbed profile technique for general lysimeter installation may generate biased and uninterpretable results (4).

Kung (6) described another disadvantage of nonweighing lysimeters, in which funnel flow may divert percolating water based on the angle of inclination of coarse textured soil lenses. Lenses, or soil textural discontinuities, may cause water to bypass the lysimeter or channel water directly to it. In either case, results may be erroneous.

Tension or tensionless column lysimeters that have sidewalls extending above the soil surface pose additional challenges (7–9). Impeding runoff, as the design of extended-walled column lysimeters do, unnaturally alters the water status of the contained soil column by permitting ponding during winter months and following intense rainfall. Furthermore, crop management becomes more difficult as the extended sidewalls act as obstacles to plowing, seeding, and cultivating (9).

Zero-tension lysimeters rely on the formation of a saturated soil zone above the porous plate before drainage occurs by gravitational flow, which is an unlikely natural condition in most deep, well-drained and moderately well-drained soils and alters the water content and soil water potential of the rooting zone (7). The prerequisite of saturation above these sampling devices may cause drainage water to bypass the lysimeter (3) and create unnatural matric potential gradients that cause water movement from the saturated zone to the surrounding unsaturated soil matrix (10). Zero-tension lysimeters that have upwardly extended sidewalls (7,8,11) may promote preferential flow down the side of the soil column upon drying (11). In addition, lateral flow that normally occurs under field conditions would be restricted in containerized, column lysimeters with fine textured, clayey horizons (12).

Fixed-tension lysimeters improve the flow patterns around and through the lysimeter's porous plate surface. Stored water often moves through soil under some tension. Applying tension (i.e., suction or vacuum) to a porous surface produces more natural flow patterns through a soil profile than relying on gravity to generate drainage. However, fixed-tension lysimeters typically experience vacuum depletion shortly after vacuum application unless constant vacuum is maintained. Fixed-tension lysimeters do not reproduce natural soil water potential variations that can occur during a relatively short period of time. The transient nature of soil water potentials influences unsaturated hydraulic conductivity and thus alters drainage fluxes.

RECENT USAGE OF NONWEIGHING DRAINAGE LYSIMETERS

Drainage lysimeters, both tension and tensionless, have been frequently employed in numerous studies to accomplish a variety of research goals. Zero-tension lysimeters have been used to study salt and water movement and leaching losses under conventional and no-tillage maize [*Zea mays*; (13)]; collect mobile colloids transporting organic and heavy metal compounds (11); investigate nitrate (NO_3^-) and pesticide leaching from removed, but intact, soil columns (12,14); monitor the fate and cycling of ^{15}N on revegetated mine spoils in containerized, soil mini-columns (7); determine the effect of irrigation on the nitrate leaching potential of agroecosystems (4); measure nitrate leaching from maize agroecosystems under zero, intermediate, and high fertilizer N rates (3); and elucidate the effect of various cropping systems on solute leaching and ^{15}N migration (8,9).

Jemison and Fox (10) evaluated the collection efficiency of zero-tension lysimeters. A bromide (Br^-) tracer method was used to determine an overall mean collection efficiency, based on lysimeter recovery of applied Br^-, of 52%, ranging from 13–92%, and high plot-to-plot variability of volumes were collected. Cumulative observed drainage, uncorrected for Br^- collection efficiency, measured from zero-tension lysimeters for 3 years, was consistently lower than LEACHM-predicted drainage (10). LEACHM (Leaching Estimation And Chemistry Model) is a complex computer model capable of simulating water infiltration and redistribution, solute transport and distribution in the soil, and tracer ion uptake by plants (15).

The use of fixed-tension drainage lysimeters has been somewhat more limited. Karnok and Kucharski (16) described the use of an underground rhizotron-lysimeter facility constructed to monitor root growth, evapotranspiration, and leaching under turf grass cover. Leaching losses from observation cells were collected under a continuous, fixed tension of −33.3 kpa. Dolan et al. (17) used wedge-shaped, fixed-tension microlysimeters to ascertain the effect of nitrogen placement on nitrate leaching below maize monoculture grown in sandy soil. Adams et al. (18) evaluated nitrate leaching from poultry litter and manure-amended pasture soil using pan lysimeters at a fixed tension of ∼ −10 kpa. Despite several disadvantages, fixed-tension and tensionless nonweighing lysimeters have many potential uses and applications.

EQUILIBRIUM-TENSION LYSIMETERS

Equilibrium-tension lysimetry is an innovative and improved nonweighing drainage lysimeter technique that allows applied-lysimeter suction to mimic the surrounding bulk soil matric potential fluctuations; thus it minimizes the potential for bypass flow and preserves natural drainage patterns (19). A heat dissipation sensor (19,20) is used to monitor the bulk-soil matric potential, and suction applied to the lysimeters is varied as the bulk-soil matric potential varies. Lysimeter suction was originally set manually every several days (19), but a recently developed automated vacuum system maintains equilibrium between applied-lysimeter suction and bulk-soil matric potential fluctuations within seconds (21).

The use of equilibrium-tension lysimeters is relatively new and even more limited compared to the other types of nonweighing drainage lysimeters (Fig. 1). Equilibrium-tension lysimeters have been used to quantify drainage losses for a multiyear period in a water balance study of a restored prairie and maize agroecosystems in southern Wisconsin (22). At the same study site in southern Wisconsin, equilibrium-tension lysimeters have been used to quantify multiple years of nitrogen and carbon (23,24) and soluble phosphorus fluxes (25) in drainage water from below the root zones of a restored prairie and maize agroecosystems. Equilibrium-tension lysimeters have also been used to quantify heavy metal and antibiotic leaching from the root zone of tall fescue (*Festuca arundinacea*) amended with poultry litter in northwest Arkansas (unpublished data).

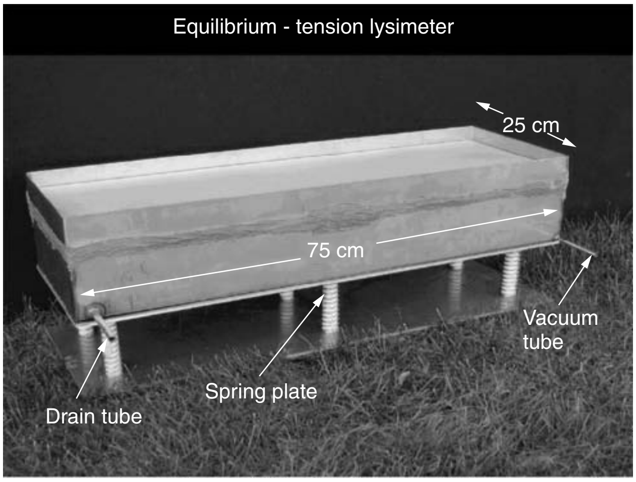

Figure 1. Stainless steel equilibrium-tension lysimeter (ETL) for measuring drainage and solute leaching below an undisturbed soil column.

Figure 2. Backhoe digging trench for lysimeter installation.

Figure 3. Rectangular cavity excavated in sidewall of the trench with nearly completed lysimeter installation.

Figure 4. Complete lysimeter installation with cavity protected from damage upon backfilling with treat plywood.

Figure 5. Soil surface following lysimeter installation and trench backfilling.

GENERAL TENSION-LYSIMETER INSTALLATION AND SAMPLING PROCEDURES

For the most representative results in a field setting, nonweighing drainage lysimeters are typically installed below an undisturbed soil column. A trench is dug by hand or with a backhoe, then a small cavity that extends into the plot or measurement area is excavated in the long sidewall of the trench (Fig. 2). The lysimeter is forced upward against the base of the undisturbed soil column. Some apparatus (e.g., a spring plate, shim, or screw jack) is installed below the lysimeter to provide the necessary upward force to maintain the capillary connection between the lysimeter's porous plate surface and the soil above (Fig. 3). Once the cavity is outfitted with the lysimeter, the trench can either be left open and accessible or completely backfilled. If the trench will be backfilled, tube extensions that extend upward out of the soil can be connected to the drain and vacuum tubes attached to the lysimeter. The lysimeter cavity itself is typically left open but protected with plywood from damage during backfilling (Fig. 4). The trench is backfilled by adding soil in the opposite order in which it was removed and is revegetated, if necessary, so that only the tube extensions and heat dissipation sensor wires if used, are visible above the soil surface (Fig. 5).

Water collected in the lysimeter's collection reservoir is typically removed under vacuum. The leachate volume is recorded, and, if necessary, a portion is saved and refrigerated for subsequent chemical analysis.

BIBLIOGRAPHY

1. Joffe, J.S. (1932). Lysimeter studies. I. Moisture percolation through soil profiles. *Soil Sci.* **34**: 123–143.
2. Soil Science Society of America (SSSA). (1996). *Glossary of Soil Science Terms*. Soil Sci. Soc. Am., Madison, WI.
3. Jemison, J.M. and Fox, R.H. (1994). Nitrate leaching from nitrogen-fertilized and manured corn measured with zero-tension pan lysimeters. *J. Environ. Qual.* **23**: 337–343.
4. Martin, E.C., Loudon, T.L., Ritchie, J.T., and Werner, A. (1994). Use of drainage lysimeters to evaluate nitrogen and irrigation management strategies to minimize nitrate leaching in maize production. *Trans. ASAE.* **37**: 79–83.
5. Owens, L.B. (1987). Nitrate leaching losses from monolith lysimeters as influenced by nitrapyrin. *J. Environ. Qual.* **16**: 34–38.
6. Kung, K.J.S. (1993). Laboratory observation of funnel flow mechanism and its influence on solute transport. *J. Environ. Qual.* **22**: 91–102.
7. Reeder, J.D. (1986). A non-weighing lysimeter design for field studies using nitrogen-15. *Soil Sci. Soc. Am. J.* **50**: 1224–1227.
8. Shipitalo, M.J. and Edwards, W.M. (1993). Seasonal patterns of water and chemical movement in tilled and no-till column lysimeters. *Soil Sci. Soc. Am. J.* **57**: 218–223.
9. Moyer, J.W., Saporito, L.S., and Janke, R.R. (1996). Design, construction, and installation of an intact soil core lysimeter. *Agron. J.* **88**: 253–256.
10. Jemison, J.M. and Fox, R.H. (1992). Estimation of zero-tension pan lysimeter collection efficiency. *Soil Sci.* **154**: 85–94.

11. Thompson, M.L. and Scharf, R.L. (1994). An improved zero-tension lysimeter to monitor colloid transport in soils. *J. Environ. Qual.* **23**: 378–383.
12. Bergstrom, L. and Johansson, R. (1991). Leaching of nitrate from monolith lysimeters of different types of agricultural soils. *J. Environ. Qual.* **20**: 801–807.
13. Tyler, D.D. and Thomas, G.W. (1977). Lysimeter measurements of nitrate and chloride losses from soil under conventional and no-tillage corn. *J. Environ. Qual.* **6**: 63–66.
14. Bergstrom, L. (1990). Leaching of chlorsulfuron and metsulfuron methyl in three Swedish soils measured in field lysimeters. *J. Environ. Qual.* **19**: 701–706.
15. Jemison, J.M., Jabro, J.D., and Fox, R.H. (1994). Evaluation of LEACHM: I. Simulation of drainage, bromide leaching, and corn bromide uptake. *Agron. J.* **86**: 843–851.
16. Karnok, K.J. and Kucharski, R.T. (1982). Design and construction of a rhizotron-lysimeter facility at the Ohio State University. *Agron. J.* **74**: 152–156.
17. Dolan, P.W., Lowery, B., Fermanich, K.J., Wollenhaupt, N.C., and McSweeney, K. (1993). Nitrogen placement and leaching in a ridge-tillage system. pp. 176–183. In: *Conference Proceedings on Agricultural Research to Protect Water Quality*. Feb. 21–24, 1993. Minneapolis, MN. Soil Water Conserv. Soc., Ankeny, IA.
18. Adams, P.L. et al. (1994). Poultry litter and manure contributions to nitrate leaching through the vadose zone. *Soil Sci. Soc. Am. J.* **58**: 1206–1211.
19. Brye, K.R., Norman, J.M., Bundy, L.G., and Gower, S.T. (1999). An equilibrium tension lysimeter for measuring drainage through soil. *Soil Sci. Soc. Am. J.* **63**: 536–542.
20. Reece, C.F. (1996). Evaluation of a line heat dissipation sensor for measuring soil matric potential. *Soil Sci. Soc. Am. J.* **60**: 1022–1028.
21. Masarik, K.C., Norman, J.M., Brye, K.R., and Baker, J.M. (2004). An improved automated approach to equilibrium tension lysimeters. *J. Environ. Qual. In Press*.
22. Brye, K.R., Norman, J.M., Bundy, L.G., and Gower, S.T. (2000). Water budget evaluation of prairie and maize ecosystems. *Soil Sci. Soc. Am. J.* **64**: 715–724.
23. Brye, K.R., Norman, J.M., Bundy, L.G., and Gower, S.T. (2001). Nitrogen and carbon leaching in agroecosystems and their role in denitrification potential. *J. Environ. Qual.* **30**: 58–70.
24. Brye, K.R., Gower, S.T., Norman, J.M., and Bundy, L.G. (2002). Carbon budgets for a prairie and agroecosystems: Effects of land use and inter-annual variability. *Ecol. Appl.* **12**: 962–979.
25. Brye, K.R. et al. (2002). Phosphorus leaching under a restored tallgrass prairie and corn agroecosystems. *J. Environ. Qual.* **31**: 769–781.

STEADY-STATE FLOW AQUIFER TESTS

Harendra S. Chauhan
G.B. Pant University of
Agriculture and Technology
Uttar Pradesh, India

Groundwater studies are conducted to evaluate the occurrence, availability, and quality of groundwater. There are numerous practical applications like optimum development of groundwater basins, safe withdrawal, groundwater recharge, seepage from canals and reservoirs, and drainage that require knowledge of the hydraulic characteristics of aquifers. Two basic aquifer properties needed to make a quantitative assessment of such applications are transmissivity and storage coefficient which represent, in a general sense, its abilities to transmit and store water, respectively. The other hydraulic properties of an aquifer are hydraulic resistance, leakage factor, and drainage factor. Performing an aquifer test is an effective way of determining these hydraulic properties of aquifers. Steady-state flow aquifer tests are conducted to determine transmissivity, hydraulic conductivity, and hydraulic resistance; more detailed unsteady-state tests are conducted if other hydraulic properties are also to be evaluated. The objective of this article is to discuss steady-state methods of determining aquifer characteristics.

AQUIFER TESTS

The procedure involves pumping water from a well for a certain time and at a certain rate and measuring its effect on the water table in the well and at a number of piezometers installed in the vicinity. When pumping is carried out in a well, the water level in its vicinity is lowered; the extent of lowering at any point is termed the, drawdown. The drawdown is maximum at the well point and reduces gradually with distance from the well, becoming zero at a certain distance termed the radius of influence. This distance varies with the well and depends on the well discharge, formation constants, and the duration of pumping. The lowering of the water level in the well during pumping creates a head difference or hydraulic gradient from the surroundings which causes the flow of water in the well. The hydraulic gradient increases as flow converges toward the well. The rate of flow is same through all the concentric cylindrical surfaces around the well, and the surface area decreases as one approaches the well. The form of the surface resembles a cone, and therefore it is known as the cone of depression, which is the characteristic feature of every pumped well. Its size and shape depend on the discharge, length of pumping, aquifer characteristic, and other factors.

Steady-State and Unsteady-State Flows

Both steady-state and unsteady-state groundwater flow conditions occur in nature and cannot be very distinctly separated. It is a matter of interpretation and choice as to which concept is used for practical utility. Although every phenomenon in nature is transient, it could be considered a series of successive steady-state events.

Steady-state groundwater movement requires that the boundary conditions of flow and aquifer characteristics remain unchanging with time. Even though the aquifer characteristics do not remain constant, they are assumed so for practical purposes. Thus, if a hydraulic disturbance like well pumping is created, even though the process is initially unsteady, eventually it becomes steady with time.

DEFINITION OF AQUIFER PROPERTIES

The transmissivity (T) of an aquifer is the rate of flow under a unit hydraulic gradient through a vertical strip of the aquifer of unit width and extending through the whole saturated thickness of the aquifer. It has the dimensions of length^2/time and, for example, may be expressed in m^2/day. T is the product of the average hydraulic conductivity (k) and the saturated thickness (b) of the aquifer.

Hydraulic conductivity (k) is the proportionality constant k in Darcy's law $v = ki$, where v is the specific discharge or Darcy's velocity (L T^{-1}) and i is the hydraulic gradient (dimensionless), which is the volume of water that will flow through a porous medium in unit time under a unit hydraulic gradient through a unit area measured at right angles to the direction of flow. k is a parameter that depends on both the properties of the porous medium and of the fluid and has units of length/time, for example, m/day.

Saturated thickness (b) is equal to the physical thickness of the aquifer between two confining layers in a confined aquifer. The saturated thickness for unconfined aquifers, equal to the difference between the free water table and the impervious layer, is not constant but varies with pumping.

Coefficient of storage (S) of an aquifer is the volume of water released from or taken into storage per unit surface area of the aquifer per unit change in head. In unconfined aquifers, S is known as the *specific yield* which represents the volume of water released from storage per unit surface area of aquifer per unit decline of the water table. Specific yield is also sometimes called effective porosity or drainable pore space. S is a dimensionless quantity.

Hydraulic resistance (c) is a property of semiconfined aquifers that characterizes the resistance of a water-bearing layer to vertical flow. It is also known as the reciprocal leakage coefficient and is given by b'/k' where b' and k' are the thickness and the hydraulic conductivity, respectively, of a semipervious layer. The dimension of the hydraulic resistance is time and is, for example, expressed in days.

Leakage factor (L) describes the spatial distribution of leakage through one or more semipervious layers into a leaky aquifer or vice versa. It is defined as $L = \sqrt{kbc}$, has the dimension of length, and it may be expressed in meters.

Of these aquifer parameters, transmissivity (T), hydraulic conductivity (k), and hydraulic resistance (c) are the properties that can be determined from steady-state aquifer tests.

STEADY-STATE METHODS OF DETERMINING AQUIFER PARAMETERS

Different well flow theories have been derived based on the consideration of a continuity equation in cylindrical or spherical coordinates, well depth, the fully or partially penetrating nature of wells; steady-state or unsteady-state flow conditions; gravity, artesian, leaky or nonleaky aquifer; constant or variable discharge conditions; and sinks or source. The methods of determining aquifer constants depend on the solution of the selected continuity equation and the specific boundary conditions. Steady-state solutions can be used to estimate transmissivity; the storage coefficient (S) cannot be determined because it does not appear in the solution. The following methods based on steady-state approaches can be used to determine the transmissivity of an aquifer.

Unconfined Aquifer Flows

Radial Flow to a Fully Penetrating Constant Discharge Well (Fig. 1). The flow is considered assuming that the aquifer is homogeneous, isotropic, of uniform thickness, and has an infinite aerial extent. The differential equation for groundwater flow in an unconfined aquifer may be written in cylindrical coordinates as

$$\frac{d}{dr}\left(\frac{r\,dh}{dr}\right) = 0 \tag{1}$$

In a well that fully penetrates the gravity aquifer, the flow will be axisymmetric. Neglecting the vertical component of the velocity as small, the equipotential surfaces may be considered coaxial cylinders. The flow across these cylinders at a distance r is given by $Q = av = 2\pi rh\left(-k\dfrac{dh}{dr}\right)$. Equation (1) may be integrated twice. The two constants may be obtained from an assumed boundary condition of a piezometric head at the well as $h(r_0) = h_0$ and the relationship of the well flow at a given distance r. The solution of Eq. 1 under the boundary conditions gives the phreatic surface of the flow as

$$h^2 = \frac{Q}{\pi k}\ln\frac{r}{r_0} + h^2 \tag{2}$$

Three methods may be developed from Eq. 2 to determine transmissivity.

Determination of T Using Two Observation Wells: Dupuit–Thiem Method. It can be seen from Eq. 2 that there is no fixed saturated thickness in gravity aquifers and only k appears in the solution describing phreatic surfaces. Thus, only k can be obtained from the solution. To obtain the transmissivity, k has to be multiplied by some average depth. Measuring heads at two distances $r_1, r_2 (r_2 \geq$

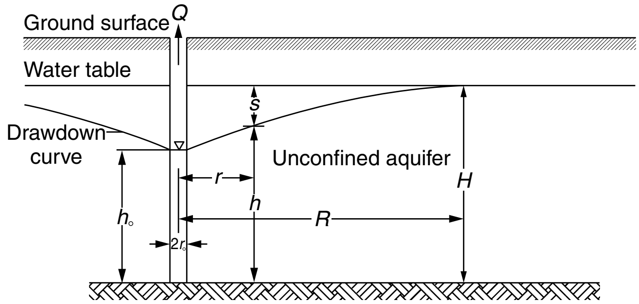

Figure 1. Flow to an unconfined aquifer.

$r_1), h(r_1) = h_1, h(r_2) = h_2$ in Eq. 2, an expression for k may be written as

$$k = \frac{Q}{\pi h(h_2^2 - h_1^2)} \ln \frac{r_2}{r_1} \tag{3}$$

Contrary to Dupuit's assumption, the vertical components of velocity are large in gravity flow and cannot be ignored. Equation 3, consequently, does not give correct results. Setting $h = H - s$, Eq. 3 can be rewritten as

$$kH = \frac{Q}{2\pi\left[\left(s_1 - \dfrac{s_1^2}{2\,H}\right) - \left(s_2 - \dfrac{s_2^2}{2H}\right)\right]} \ln \frac{r_2}{r_1} \tag{4}$$

where H is the saturated thickness of the aquifer, s_1 and s_2 are drawdowns and $s_1 - \dfrac{s_1^2}{2H}$ and $s_2 - \dfrac{s_2^2}{2H}$ are corrected drawdowns s_1' and s_2',

Transmissibility may be written in terms of corrected drawdowns or, neglecting correction terms, it may also be expressed as

$$T = kH = \frac{Q}{2\pi(s_1' - s_2')} \ln \frac{r_2}{r_1} \approx \frac{Q}{2\pi(s_1 - s_2)} \ln \frac{r_2}{r_1} \approx \frac{Q}{2\pi(h_2 - h_1)} \ln \frac{r_2}{r_1} \tag{5}$$

Determination of T Using a Single Observation Well. Due to constraints on funds for pumping tests, it may be possible to construct only one observation well. For such a situation, one may have observations only in a single well. For such conditions, setting $h(r_0) = h_0$ and $h(r_1) = h_1, (r_1 \geq r_0)$, an expression for k may be written as

$$k = \frac{Q}{\pi(h_1^2 - h_0^2)} \ln \frac{r_1}{r_0} \tag{6}$$

In the same way for such boundary conditions, T may be written in terms of corrected drawdowns. Considering that corrections have small values, they may be neglected to give approximate values of transmissivity as

$$kH = \frac{Q}{2\pi\left[\left(s_0 - \dfrac{s_0^2}{2H}\right) - \left(s_1 - \dfrac{s_1^2}{2H}\right)\right]} \ln \frac{r_1}{r_0} \approx \frac{Q}{2\pi(s_0 - s_1)} \ln \frac{r_1}{r_0} \approx \frac{Q}{2\pi(h_1 - h_0)} \ln \frac{r_1}{r_0} \tag{7}$$

Determination of T Using the Head in the Pumped Well. Due to further constraints on funds, one may not be able to make even a single observation well. For such a condition, measuring the head in the pumped well itself, $h(r_0) = h_0$, and considering an arbitrary radius of influence $h(R) = H$, the hydraulic conductivity may be obtained as

$$k = \frac{Q}{\pi(H^2 - h_0^2)} \ln \frac{R}{r_0} \tag{8}$$

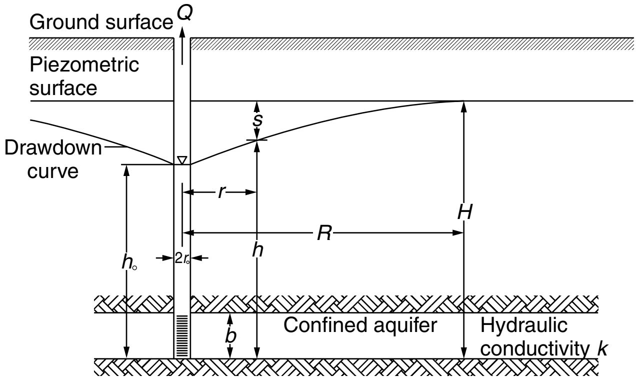

Figure 2. Flow to a confined aquifer.

Writing $h = H - s$, transmissibility may be written as

$$kH = \frac{Q}{2\pi\left(s_0 - \dfrac{s_0^2}{2H}\right)} \ln \frac{R}{r_0} \approx \frac{Q}{2\pi s_0} \ln \frac{R}{r_0} \tag{9}$$

Confined Aquifer Flows (Fig. 2)

Similar to the unconfined aquifer, flow is considered to take place assuming that the aquifer is homogeneous, isotropic, of uniform thickness, and of infinite aerial extent.

Radial Flow to a Fully Penetrating, Constant Discharge Well. Differential equations of the groundwater flow to a fully penetrating well in a confined aquifer may be written in cylindrical coordinates as

$$\frac{d}{dr}\left(r\frac{dh}{dr}\right) = 0 \tag{10}$$

In a fully penetrating aquifer of thickness b, the flow is axisymmetric, and the equipotential surfaces are coaxial circular cylinders. Flow across these cylinders in a fully penetrating artesian aquifer of thickness b at a given distance r from the well center is given as $Q = av = 2\pi rb\left(-k\dfrac{dh}{dr}\right)$. Equation 10 may be integrated twice under the assumed condition of the piezometric head at the well, $h(r_0) = h_0$, and the relationship of the well at a given distance r to give

$$h = \frac{Q}{2\pi kb} \ln \frac{r}{r_0} + h_0 \tag{11}$$

Thiem's (1906) Equilibrium Equation Using Two Observation Wells. To reduce errors because of losses in the well and through the well screen, Thiem (1) proposed measuring piezometric heads h_1, h_2 in two observation wells located at distances r_1, r_2 from a constantly discharging well Q. Thus, setting $h(r_1) = h_1, h(r_2) = h_2$ in Eq. 11 of an expression of transmissivity may be derived as

$$T = kb = \frac{Q}{2\pi(h_2 - h_1)} \ln \frac{r_2}{r_1} \tag{12}$$

Determination of T Using a Single Observation Well. Considering the head as $h(r_1) = h_1$ at a single piezometer installed at distance r_1 from the well, T may be obtained from Eq. 11 as

$$T = kb = \frac{Q}{2\pi(h_1 - h_0)} \ln \frac{r_1}{r_0} \tag{13}$$

Determination of T Using the Piezometric Head in the Pumped Well. By measuring the piezometric head in the pumped well itself and at an arbitrary radius of influence R, $h(R) = H$, T can be obtained from Eq. 11 as

$$T = kb = \frac{Q}{2\pi(H - h_0)} \ln \frac{R}{r_0} = \frac{Q}{2\pi s_0} \ln \frac{R}{r_0} \tag{14}$$

Spherical Flow to a Nonpenetrating, Constant Discharging Well in An Extensively Thick Artesian Aquifer (Fig. 3). The groundwater flow equation for such a flow system is described by a continuity equation in spherical coordinates as

$$\frac{d}{dr}\left(r^2 \frac{dh}{dr}\right) = 0 \tag{15}$$

The isopiestic surfaces to a nonpenetrating well consist of concentric hemispheres. The flow across a hemispherical surface at distance r is given by

$$Q = av = 2\pi r^2(-k dh/dr)$$

Considering the piezometric head at the well, $h(r_0) = h_0$, and the flow equation above, the piezometric surface can be obtained as

$$h = \frac{Q\left(\frac{1}{r_0} - \frac{1}{r}\right)}{4\pi k} + h_0 \tag{16}$$

Determination of k Using Two Observation Wells. It may be observed that from Eq. 16, only hydraulic conductivity can be determined as the thickness of the aquifer is not finite.

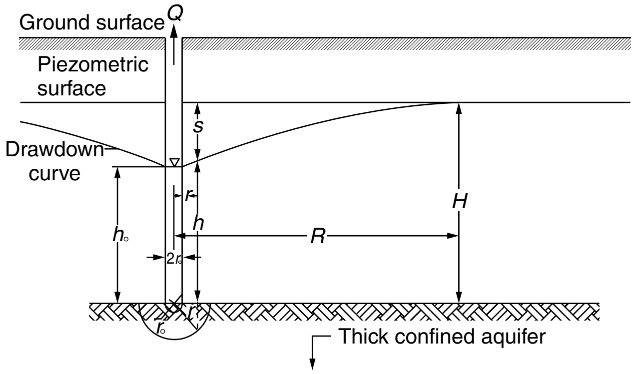

Figure 3. Spherical flow to a thick confined aquifer.

Considering heads at two observation wells located at distances r_1, r_2 as h_1, h_2, an expression for k may be obtained as

$$k = \frac{Q\left(\frac{1}{r_1} - \frac{1}{r_2}\right)}{2\pi(h_2 - h_1)} \tag{17}$$

Determination of k Considering the Head in the Well. Using the measurement of the head in the pumped well itself, $h(r_0) = h_0$, and considering an arbitrary radius of influence R, an expression for k in terms of the head and drawdown $s = H - h$ may be written as

$$k = \frac{Q\left(\frac{1}{r_0} - \frac{1}{R}\right)}{2\pi(H - h_0)} = \frac{Q\left(\frac{1}{r_0} - \frac{1}{R}\right)}{2\pi s_0} \tag{18}$$

Considering that $R > r_0$, 1/R is negligible and can be neglected; this gives

$$k \approx \frac{Q}{2\pi r_0(H - h_0)} \approx \frac{Q}{2\pi r_0 s_0} \tag{19}$$

Semiconfined Aquifer Flows

Aquifer Constants from the Steady-State Flow of a Well in a Semiconfined Aquifer: Hantush and Jacob method (2) (Fig. 4). The solution is based on the assumption that the phreatic surface remains constant, so that the leakage through the semiconfining layer takes place in proportion to the drawdown of the piezometric level. Hantush and Jacob (2) obtained a solution for $L > 3D$, as

$$s_m = \frac{Q}{2\pi kb} K_0\left(\frac{r}{L}\right) \tag{20}$$

where s_m = steady-state drawdown in meters at distance r;
Q = discharge of the pumped well in cubic meters per day;
$L = \sqrt{kbc}$, leakage factor in meters;
$c = \frac{b'}{k'}$, hydraulic resistance of a semipervious layer in meters; and
$K_0(x)$ = modified Bessel's function of the first kind and zero order.

Hantush observed that if r/L is <0.05, Eq. 20 can be approximated as

$$s_m = \frac{2.3Q}{2\pi kb}\left(\log 1.12 \frac{L}{r}\right) \tag{21}$$

A plot of s_m against r on semilog paper, with r on the logarithmic scale, will show a straight line relationship in the range where r/L is small. The slope of the straight line, the drawdown difference per log cycle of r, is expressed as

$$\Delta s_m = \frac{2.3Q}{2\pi kb}$$

$$kb = \frac{2.3Q}{2\pi \Delta s_m} \tag{22}$$

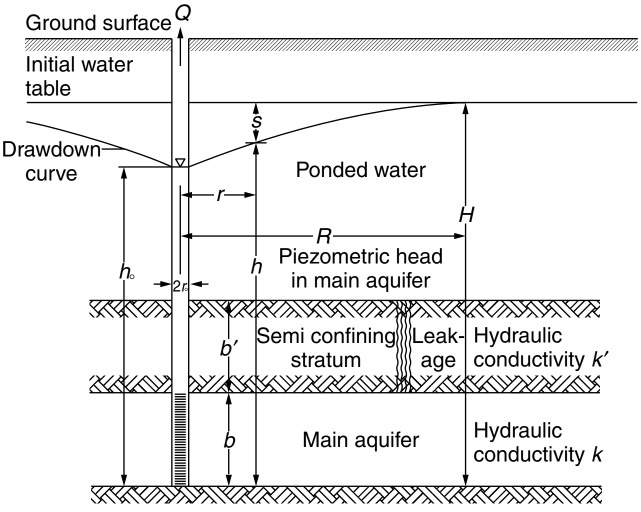

Figure 4. Flow to a semiconfined aquifer.

For any point s_m, r is selected on the straight line. Putting r in Eq. 21 gives $L.c$ can be obtained as

$$\mathrm{c} = \frac{b'}{k'} = \frac{L^2}{kb}$$

Alternatively the extended straight line portion of the curve intercepts the r axis where the drawdown is zero. At the interception point, $s_{\mathrm{m}} = 0, r = r_{\mathrm{o}}$, and Eq. 21 reduces to

$$\left(\log 1.12\frac{L}{r}\right) = 0$$

or

$$1.12\frac{L}{r_0} = 1.12\frac{\sqrt{kbc}}{r_0} = 1$$

or

$$\mathrm{c} = \frac{(r_0/1.12)^2}{kb}$$

Transmissivity from Steady Flow in a Semiconfined Aquifer: Ernst's Modification of the Thiem Method. The relationship is based on assuming that the well discharge is corrected by subtracting, from the total discharge, the amount of phreatic water obtained from the semipervious layer. The determination requires two piezometers installed in semipervious layers at distances r_1, r_2. The coaxial phreatic water from a circular area may be written approximately as $r_0 = r_1 + 1/3(r_2 - r_1)$.

The phreatic water delivered from a circular area is

$$Q' = \pi r_0^2 \Delta s' S' \times 24$$

where $\Delta s'$ = average drop of phreatic level in the semipervious layer during the last hour (m) and S' = specific yield of the semiconfined aquifer (to be estimated from well logs). The transmissivity may be found from

$$kb = \frac{(Q - Q')\ln\dfrac{r_2}{r_1}}{2\pi(s_1 - s_2)} \tag{23}$$

Approximate Methods

Logan's (3) Method for confined aquifers. Equation 14 using the observation in the pumped well itself is also known as Logan's method. According to Logan (3), the accuracy of the calculation depends on the accuracy in measuring the maximum drawdown in the well s_0 and on the accuracy of the ratio R/r_0. The ratio R/r_0 may have a large value, but the ratio of the logarithm may be small. He suggested an approximate value of 3.33 for the log ratio. Putting this value in Eq. 14 gives the transmissivity as

$$kb = \frac{2.3Q}{2\pi s_0}\log\frac{R}{r_0} \approx \frac{1.22Q}{s_0} \tag{24}$$

Logan's (3) Method for Unconfined Aquifers. Using a similar approach and considering observation in the pumped well itself, the transmissivity in an unconfined aquifer may be obtained from Eq. 9 as

$$kH = \frac{2,3Q}{2\pi s_0'}\log\frac{R}{r_0} \approx \frac{1.22Q}{s_0'} \tag{25}$$

where $s_0' \equiv s_0 - \dfrac{s_0^2}{2H}$.

Transmissivity from Other Analytical Solutions

In principle, any analytical solution of a boundary value problem of groundwater flow can be used to estimate aquifer parameters, if other variables can be measured or estimated. A few steady-state solutions for estimating transmissivity are presented.

Two Equal, Constant Discharge, Mutually Interfering Wells. The relationships between two equal and constant discharge interfering wells distance B apart may be written as

$$Q_1 = Q_2 = \frac{2\pi kb(h_a - h_0)}{\ln\dfrac{R^2}{r_0 B}} \tag{26}$$

where r_0 = well radius, h_a = average piezometric head at the external boundary, h_0 = piezometric head at the well, and R = radius of influence ($R > B$).

Thus if the values of the constant equal discharges. Q, h_0, h_a, R, and B are known, the transmissivity may be determined from the relationship

$$kb = \frac{Q\ln\dfrac{R^2}{r_0 B}}{2\pi(h_a - h_0)} \tag{27}$$

Steady-State Flow to a Partially Penetrating Well. A relationship was obtained for such a flow system by De Glee (4) as

$$h_{2b} - h_0 = \frac{Q_p}{4\pi k}\left(\frac{2}{h_s}\ln\frac{\pi h_s}{2r_0} + \frac{0.20}{b}\right)$$

where h_s is the depth of the partially penetrating well, b is the aquifer thickness, h_{2b} is the initial piezometric head at distance 2b, and h_0 is the piezometric head in the well.

$$\text{Also} \quad 1.3\,h_s \le b; \frac{h_s}{2r_w} \ge 5$$

From this relationship and measuring other parameters, the transmissivity may be found as

$$kb = \frac{Q_p b}{4\pi(h_{2b} - h_0)} \left[\frac{2}{h_s} \ln \frac{\pi h_s}{2r_0} + \frac{0.20}{b}\right] \tag{28}$$

For a partially penetrating well in an unconfined aquifer where the drawdown is also small in relation to the saturated thickness, Eq. 28 can be used by replacing b by H.

Kozeny (5) proposed a relationship, by which transmissivity may be determined from

$$T = \frac{2.3 Q_p \log \dfrac{r}{r_0}}{2\pi(h - h_0)\, G} \tag{29}$$

where

$$G = \alpha \left(1 + 7\sqrt{\frac{r_0}{2h_s}} \cos \frac{\pi\alpha}{2}\right)$$

Q_p is the discharge of a partially penetrating well, Q_f is the discharge of a fully penetrating well, $G = Q_p/Qf$ is a correction factor giving the ratio of the two discharges, α is the fractional penetration of the well screen h_s/b, and h_s is the depth of penetration of the aquifer.

PRACTICAL ASPECTS OF PUMPING TESTS (after Ref. 6)

Observation Well Size, Location, and Numbers

An observation well should be just large enough to allow accurate and rapid measurement of water levels. Small wells provide a fast response. Five-cm observation wells are usually effective when hand methods of measuring water levels are used. Larger pipes may be used with automatic recorders. Observation wells are commonly installed with 1–2 m long screens at the bottom embedded in the aquifer to measure the pressure head. Observation wells should be installed at about the same depth as the middle of the well screen in the pumped well.

Observation wells should be located at a reasonably large distance and not too close to the pumped well. Setting observation wells at distances of 30–100 m from the pumped well may work out best in most cases. When observation wells are too close to the pumped well, correct drawdown readings may be difficult. The cone of depression in a water table spreads at a slower rate than in an artesian formation, so the observation wells for a water table aquifer should be relatively closer to the pumped well.

The number of observation wells depend upon the amount of information required and the funds available for the test program. Drawdown data from the well itself or from a single observation well permit estimating the average hydraulic conductivity, transmissibility, and coefficient of storage of the aquifer. If two or more observation wells are placed at different distances, the test data can be analyzed both in terms of the time–drawdown and the distance–drawdown relationships. It is always better to have as many observation wells as conditions may allow. A minimum of three placed at varying distances and radially at equal angles around the well is desirable.

Measurement of Water Levels. The depth to water may be measured many times during the course of a pumping test. Observations should be taken at close intervals during the first 2 hours, and the time between readings should gradually increase as the test continues. The water level measurement may be recorded to the nearest 0.5 cm in the observation wells. Measurements may be made every half-minute during the first 5 minutes after starting the pump, then every 5 minutes for an hour, then every 20 minutes for about two hours. From this point on, readings at hourly intervals should be sufficient.

Water-level measurements can be taken in various ways, the wetted-tape method, mechanical sounder, electric water-level indicator, pressure gauge, or pressure logger. For detailed information on these devices, see Kruisman and De Ridder (7). Fairly accurate measurements of water levels can be made manually.

Measurements of Discharges. The other parameter that needs to be measured is the discharge from wells during pumping tests. To avoid complicated calculations later, the discharge rate should preferably be kept constant throughout the test. The discharge rate can be measured using various devices like commercial water meters, flumes, orifice weirs, containers or the jet-stream method. The procedures for installation and measurements using these devices are available in standard texts.

SUMMARY

An overview of steady-state approaches for determining the transmissivity of confined, unconfined, and semiconfined aquifers has been presented. Detailed procedures for determining transmissivity using drawdown data in a pumped well and one or two observation wells are discussed in the context of two most common methods of Thiem and Dupuit–Thiem. A few approximate methods and analytical solutions for determining aquifer constants have been briefly introduced. General guidelines for conducting pumping tests including measurement of discharge and groundwater levels and numbers and locations of observation wells have been furnished.

BIBLIOGRAPHY

1. Thiem, G. (1906). *Hydrologische Methoden*. Leipzig, p. 56.
2. Hantush, M.S. and Jacob, C.E. (1955). Non steady radial flow in an infinite leaky aquifer. *Trans. Amer. Geophysical Union*. **36**: 95–100.
3. Logan, J. (1064). *Estimating Transmissibility from Routine Production of Tests of Water Wells Ground Water* **2**(1): 35–37.

4. De Glee, G.J. (1930). *Over Grundwaterstroomingen bij Wateronttrekking Door Middle van Patten*. J. Waltman Jr., Delft, the Netherlands, p. 175.
5. Kozeny, J. (1933). *Theorie und Berech nug der Brunneu Wasser Kraft and Wasser Wartschift*. **28**: 88–92,113–116.
6. Johnson Div. UDP, Inc. (1972). *Ground Water and Wells*. 2nd Edn. Edward E. Johnson, St. Paul, MN, p. 440.
7. Kruisman, G.P. and De Ridder, N.A. (1970). *Analysis and Evaluation of Pumping Test Data, International Institute for Land Reclamation and Improvement*. Wageningen, the Netherlands.

TIDAL EFFICIENCY

Todd Rasmussen
The University of Georgia
Athens, Georgia

Ocean tides often cause water levels in coastal aquifers to fluctuate. These fluctuations result from two different causes. In confined aquifers, the additional weight of the water on the surface increases the pressure on the water at depth, which causes an increase in water levels in observation wells located near the coast. In this case, no movement of water between the ocean and the aquifer may have occurred. In unconfined aquifers, however, the loading effect does not occur. Instead, water that has moved directly through the pores within the aquifer has caused the water level response. Regardless of the mechanism, both processes can be described with the phase lag, or delay, between the tidal peak and the aquifer response peak, as well as the amplitude ratio between the magnitudes of the response and the tides

Water levels in wells located near the coast often fluctuate with a rhythm that mimics the rise and fall of the ocean. The tidal efficiency, TE, is the ratio of the change in water levels in wells, ΔW, to the change in tidal levels, ΔT:

$$TE = \frac{\Delta W}{\Delta T}$$

where both water levels are measured with the same type of units (e.g., feet, meters, psi). Tidal efficiencies can range from a maximum of one in poorly consolidated geologic media to a minimum of zero in rigid aquifers (1,2).

The tidal response in an aquifer can result from two types of phenomena, one in confined and a second in unconfined aquifers. As described in greater detail below, the confined aquifer response is caused by the increased load placed on the surface, whereas the unconfined aquifer response is caused by the movement of water in response to changing tidal water levels.

CONFINED AQUIFERS

In confined aquifers, coastal tides affect water levels because the surface load changes as ocean levels fluctuate. The weight of the ocean compresses the underlying aquifer, which causes a change in the fluid pressure. In this situation, no water moves between the ocean and the aquifer, which is similar to the effect that trains have as they pass a well in some confined aquifers—water levels rise and then fall as a train passes by (3).

The loading effect is greater in poorly consolidated aquifers where a greater proportion of the overlying load is borne by the fluid. Pore fluids in consolidated aquifers are less likely to carry the load, and so these aquifers have lower tidal efficiencies.

Coastal aquifers are not the only type of aquifer that shows a response to surface loads. The loading effect on confined aquifers can be observed in response to the placement or removal of any type of mass on the ground surface. Components of the water budget (such as precipitation, evaporatranspiration, and soil moisture) can immediately affect water levels in confined aquifers—even when no direct hydraulic communication is present (4).

An increase in water levels in a confined aquifer under a waste lagoon may be mistakenly attributed to contamination from the lagoon. The cause could be from downward leakage of water from above, or it could also be from the increased surface load causing a rapid increase in fluid pressure.

Another example of surface loading develops when barometric pressure changes cause water levels in wells to fluctuate. In these cases, the response is somewhat complicated because the barometric pressure also exerts a force on water levels in open wells. The tidal efficiency in confined aquifers is related to the barometric efficiency, BE, in that the two efficiencies must sum to unity, $TE + BE = 1$. Thus, aquifers with low tidal efficiencies have large barometric efficiencies, and vice versa. The reader is referred to Barometric Efficiency for additional discussion of this topic.

UNCONFINED AQUIFERS

A second tidal influence occurs in unconfined, surficial coastal aquifers. In these aquifers, water levels change as the hydraulic perturbation created by changing ocean levels moves horizontally through the saturated zone toward the well. Unlike the confined case in which the load is transmitted without concomitant fluid movement, water is free to move through the unconfined aquifer.

For a simple, one-dimensional flow geometry (i.e., a well placed at some distance from a linear shoreline), the response is a function of the distance of the well from the shoreline, L, and the aquifer *hydraulic diffusivity*, $D = T/S$, where T and S are the aquifer transmissivity and storativity, repsectively. Nearby wells in more diffusive aquifers respond more rapidly than do more distant wells in aquifers with lower diffusivity. The response is a function of a dimensionless variable, $\kappa = D\tau/L^2$, where τ is the response time.

This type of response is not limited to coastal areas, however. A similar type of response is found in surficial aquifers next to streams. In this case, changing water levels during floods cause a loss of water from the stream into the adjacent aquifer, with a reversal of flow later in the flood hydrograph. This type of response is also referred to as *bank storage* (5).

SINUSOIDAL OSCILLATIONS

Many types of loading phenomena display periodic variation, tides being a common example. In addition to the gravitational affects of the sun and moon on the oceans, gravitational forces also cause changes in atmospheric pressure, called *atmospheric tides*, and even in the solid earth, called *earth tides*.

These periodic variations can be characterized with the *period* of the oscillation (e.g., 24 hours for solar tides) and the *amplitude*, or *magnitude*, of the oscillation. When observing the tidal response in confined aquifers, the tidal efficiency is equal to the ratio of the magnitude of the water level fluctuation in the well to the magnitude of the ocean tides. For example, if peak daily ocean levels vary by 1 m and the peak daily groundwater levels vary by 10 cm, then the tidal efficiency is 0.10, or 10%.

Another influence is the *phase lag*, which is the delay between when the ocean tide and the aquifer water levels peak. Normally, an insignificant delay occurs between the ocean tide and a confined aquifer response; yet this delay can be large in cases in which significant borehole storage exists.

The tidal efficiency and phase lag in unconfined aquifers can determine the aquifer diffusivity. The transmissivity and storativity can be only determined, however, if the flow of water can be determined, which requires measurements of the flow rate between the ocean (or stream) and the aquifer.

BIBLIOGRAPHY

1. Bear, J. (1972). *Dynamics of Fluids in Porous Media*. Dover Publications, New York.
2. Walton, W.C. (1970). *Groundwater Resource Evaluation*. McGraw-Hill, New York.
3. Jacob, C.E. (1940). On the flow of water in an elastic artesian aquifer. *Am. Geophys. Union Trans.* **21**: 574–586.
4. vanderKamp, G. and Schmidt, R. (1997). Monitoring of total soil moisture on a scale of hectares using groundwater piezometers. *Geophys. Res. Lett.* **24**(6): 719–722.
5. Todd, D.K. (1959). *Ground Water Hydrology*. John Wiley & Sons, New York.

COMBINED FREE AND POROUS FLOW IN THE SUBSURFACE

Diganta Bhusan Das
Oxford University
Oxford, United Kingdom

Vahid Nassehi
Loughborough University
Loughborough, United Kingdom

INTRODUCTION

The presence of a fluid layer adjacent to a porous medium is common in many environmental, chemical, mechanical and petroleum engineering problems. The flow behavior in the fluid layer is determined by the properties of the fluid (e.g., viscosity and density), and surface (e.g., pressure and shear) and body (e.g., gravity and electric field) forces. However, flow in the porous medium depends additionally on the properties of the medium which include, among others, porosity and permeability.

In fluid dynamic analysis of groundwater systems, it is commonly found that water flows through combined free (nonporous) and porous pathways between surface and subsurface systems. Such coupled regimes may also be present due to construction of industrial utilities and structures in the ground. Important engineering processes that may involve combined free and porous flow domains are, for example, fluid losses/leaks from underground pipes and storage tanks in old gas works sites; dig-and-treat, pump-and-treat, and permeable reactive barrier technologies for groundwater treatment; and drilling/extraction of oil from underground reservoirs. However, in most cases, the associated transport phenomena are determined by natural hydroenvironmental conditions, which include, for example, combined surface and subsurface flow, lake–groundwater interactions, seepage through preferential flow channels and macropores, circulatory flows and rise and fall of groundwater, glaciology, and multiple karstic regions interfaced with granular porous sections. The subdomains in coupled flow systems are generally distinguished by an interfacial surface, which represents the transitional zone for contaminants/fluid mobility from free to porous sections or vice versa. To gain on understanding of the combined flow behavior, it is necessary to develop realistic methodologies for evaluating mass and momentum transfer across the free/porous flow interfaces.

The construction of well-posed mathematical formulations for combining different flow regimes can present difficulties. Two general approaches have evolved for this purpose: (1) methodologies based on treating the combined flow domain as a single continuum and (2) formulations in which the flow domain is considered as a combination of discrete sections corresponding to different flow regimes. In the former, the porous section is effectively regarded as a pseudofluid layer, and the whole domain is analyzed as a whole. As such, a suitably formulated momentum equation for the entire domain in conjunction with mass conservation equations is solved. The mathematical model identifies fluid and porous domains based on spatially varying permeability. Such a single-domain approach is usually applied where the flow transition from free to porous flow zones is not distinct and properties of the porous domain change progressively (e.g., metal solidification problems involving mushy zones). However, if the permeability of the porous medium is constant and remains above a relatively high level, then at defined interfaces between adjacent zones, transitions between free/porous regimes occur. Under these conditions, the domain should be viewed as a combination of adjacent flow fields rather than a geometric continuum. In these situations, the second approach should be used where appropriate equations of motion describing the flow in different subdomains are solved. In general, a groundwater system corresponds to this case.

GOVERNING MODEL EQUATION

The main aim of this article is to present a general methodology that can be used to simulate hydrodynamic behavior in combined free (nonporous sections) and porous flow zones in the subsurface based on the assumption of discrete flow domains, which involves selecting governing model equations for each subdomain and the matching conditions at the interface for linking them. Although fluid mobility in combined free and porous flow on the surface has been the focus of intense research for more than two decades, analyses of subsurface flow have not treated the possible interactions between free and porous flow regimes efficiently. Addressing these issues, the hydrodynamic conditions in underground zones of coupled fluid and porous domains have been analyzed on a 3-D spatial scale by Das et al. (1). This methodology is based on the principles of mass and momentum conservation and primitive variable formulations (i.e., the equations are expressed in terms of the pressure and velocity). Conservative forms of the equations have also been adopted to enhance the accuracy of the numerical calculations. The model equations are presented below.

Free Flow Region

The momentum transfer for constant density and viscosity of a fluid at a reference temperature in a free flow region can be represented by the well known Navier–Stokes (N-S) equation,

$$\nabla P_f = -\rho \frac{dv_f}{dt} + \mu \nabla^2 v_f \qquad (1)$$

where the subscript f refers to the free flowing fluid. P and v are the pressure and velocity terms, and ρ and μ are the constant density and viscosity of the fluid, respectively. In Eq. 1, the body forces are automatically incorporated in the pressure term as the resultant flow potential. The following continuity equation that describes the mass conservation for an incompressible fluid can be used for a free flow regime:

$$\nabla \cdot v_f = 0 \qquad (2)$$

Porous Flow Region

The corresponding macroscopic equations representing the flow field in the porous section are more uncertain. The momentum equations have received considerable attention in the past. The two most suitable are the Darcy (2) and Brinkman (3) equations shown below in the transient form for anisotropic porous media where the body forces have been incorporated in the pressure terms to represent the driving flow potential:

$$\text{Darcy equation: } \nabla P_p = -\rho \frac{\partial v_p}{\partial t} - \mu K^{-1} \bullet v_p \qquad (3)$$

$$\text{Brinkman equation: } \nabla P_p = -\rho \frac{\partial v_p}{\partial t} - \mu K^{-1} \bullet v_p \mu' \nabla^2 v_p \qquad (4)$$

In these equations, the subscript p refers to the porous section; P is the interstitial averaged pressure, v is the volume averaged seepage velocity; and ρ and μ are the constant density and viscosity of the fluid, respectively. μ' is the effective viscosity which depends on the fluid viscosity, μ, and the porosity. For flow of an incompressible fluid in a continuum porous medium, the equation of continuity is

$$\nabla \cdot v_p = 0 \qquad (5)$$

where v is the volume averaged seepage velocity of fluid.

Matching Boundary Condition for the Darcy Equation as a Porous Flow Equation. To couple the equations representing the two domains, it is necessary to specify the matching hydrodynamic conditions at the interfacial region separating the two subdomains. Keeping the three-dimensional (3-D) subsurface problem in mind, the matching conditions for pressure (continuous or discontinuous) and velocity components (normal and tangential) must be specified at the interface. The Darcy equation is a first-order differential equation with respect to velocity and pressure, so it is not possible to describe the velocity components fully at the free/porous domain interface, which makes it impossible to link the Darcy and the N-S equations directly because the latter requires specifying all velocity components to solve them. Special techniques have been invoked, henceforth, for this purpose. Addressing this issue, Beavers and Joseph (4) performed their well-known experiment on the measurement of mass efflux of a Poiseulle fluid flowing over a permeable block. They hypothesized the existence of a slip velocity, commonly known as the Darcy slip phenomenon, at the interface and a shear velocity in the porous boundary layer. Consequently, an ad hoc boundary condition was proposed based on the assumption that the slip velocity is proportional to the shear effects. It allowed jumps in the tangential velocities and gave a much-needed relationship between the flow within and outside porous media. The Beavers and Joseph (4) formulation was represented as

$$\left(\frac{\partial u}{\partial y}\right)_f = \frac{\gamma}{\sqrt{K}}(u_f - u_p) \qquad (6)$$

where the x direction was taken as parallel and the y direction as perpendicular to the interface. K is the direction-independent geometric permeability of the porous medium (isotropic), and γ is a slip coefficient that depends only on the properties of the porous medium and is independent of the fluid. The Beavers and Joseph condition has now been modified for cases where the flow and the interface are perpendicular rather than parallel, as in the original formulation. Jones (5) derived the following expression for the interfacial condition, assuming that the Darcy slip phenomenon is due to shear stress at the interface rather than a simple shear velocity:

$$\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)_f = \frac{\gamma}{\sqrt{K}}(u_f - u_p) \qquad (7)$$

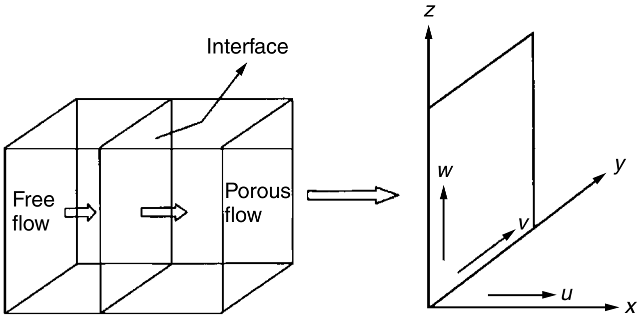

Figure 1. Model of combined free and porous flow domain for extending the Beavers and Joseph (1967) formulation to the 3-D subsurface region.

Direct extension of the above formulations to 3-D cases poses further challenges. However, by resolving the 3-D framework into two planes (x–y and x–z), it is possible to use the Beavers and Joseph formulation for three-dimensional modeling of combined flow in the subsurface (1,2). The extension can be done based on Fig. 1 where the x component of the velocity is normal to the interfacial plane and the y and z components are tangential (parallel). For the 3-D situation in Fig. 1, postulating that the Darcy slip phenomenon takes place at both x–y and x–z planes, the following relationships can be obtained heuristically based on the Jones (5) modification of Beavers and Joseph's original formulation (1,2). The velocity components that are parallel to the interface can be calculated by using these expressions:

$$\text{Velocity slip at } x\text{–}y \text{ plane: } \left(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right)_f = \frac{\gamma}{\sqrt{K_{yy}}}(v_f - v_p) \tag{8}$$

$$\text{Velocity slip at } x\text{–}z \text{ plane: } \left(\frac{\partial w}{\partial x}+\frac{\partial u}{\partial z}\right)_f = \frac{\gamma}{\sqrt{K_{zz}}}(w_f - w_p) \tag{9}$$

The slip coefficient is represented in Eqs. 8 and 9 by γ_{yz} to indicate the plane on which it is active. Equation 8 is used to simulate the jump in the transverse (y) velocity component across the interface, and Eq. 9 is used for the vertical (z) velocity component. K_{yy} and K_{zz} are the permeability in the y and z directions, respectively. The above equations, therefore, represent the slip-flow phenomena at the permeable interface with respect to directional geometric permeability. Also, the velocity component normal to the interface, the longitudinal component, can be taken as continuous across the interface:

$$u_f = u_p \tag{10}$$

However, the pressure does not necessarily remain continuous across the interface and only under static conditions, can it be said to be strictly continuous. If the pressure is to be modeled as discontinuous at the y–z interfacial plane during flow, the following relationship is used:

$$\left(-P + 2\mu\frac{\partial u}{\partial x}\right)_f = -P_p \tag{12}$$

This interfacial condition for pressure jump arises from the fact that stress in the porous side is carried by the porous materials, whereas in the free flow domain, the stress is borne by the fluid medium, which causes a discontinuity in stress, and hence fluid pressure, across the fluid/porous interface.

Matching Boundary Condition for the Brinkman Equation as a Porous Flow Equation. The use of the Brinkman equation for mathematically describing the flow in porous flow zones apparently does not cause any mathematical difficulty. The second-order derivatives of the N-S equations and the Brinkman equation with respect to velocity enable specifying all flow properties as continuous across the interface, which facilitates the task of coupling the equations of motion for the free and permeable domains. The continuity of the flow variables across the interface yields the following matching boundary conditions:

$$u_f = u_p \tag{13a}$$

$$v_f = v_p \tag{13b}$$

$$w_f = w_p \tag{13c}$$

$$P_f = P_p \tag{13d}$$

DISCUSSION AND CONCLUSIONS

Although there are superficial similarities in mathematical formulations of different coupled free and porous flow problems, Das et al. (1) clearly demonstrated that combined flow behavior in the subsurface may differ significantly from other combined flow processes encountered in chemical, petroleum and mechanical engineering applications. Therefore, models to represent specific scenarios need to be developed. In the above studies by Das et al., natural flow behavior has been defined at the exit of the domain by imposing stress-free boundary conditions. Das et al. demonstrated that at the open end (exit) of the domain, coexisting inflow and outflow might take place. In other words, fluid from the surrounding media might come inside the domain through the open end in addition to the defined inlet. For the particular domain investigated, Das et al. observed that although flow in the fluid domain is mostly unidirectional toward the free/porous interface and appears similar to pipe flow, it might become circulatory on the porous side and reverse its direction. Evidence of flow circulation, consistent with the simulated results of Das et al., can also be found in the literature dealing with various scenarios of subsurface flow (6). However, fluid circulation is a complex transient process often involving a moving center of circulation and shifts of the flow reversal front to or from the interface. After a certain time interval, the circulation may disappear, and the porous flow may become unidirectional toward the exit of the porous section. The main cause of such flow behavior has been attributed to complex/variable pressure distributions within the porous domain. The effects of a partially impermeable interface between the free and porous domains has been explored by Das and Nassehi (7). It has been demonstrated that due to the impermeable interface, there is a possibility of buildup of fluid in the free flow region. The influence of having a the multiple number of fluid/porous interfaces in a flow domain has been demonstrated by Das and Nassehi (8). This study showed that, in general, the presence of a multiple number of interfaces in the free flow domain has a minimal effect on flow behavior. However, the flow patterns on the porous side may vary depending on the direction of the free/porous interfaces. Das and Nassehi (8) also examined concentration profiles of a conservative pollutant in associated free flow and porous domains. The effects of different aspect ratios of the subdomains on the hydrodynamic behavior of free/porous regimes have been investigated by Das and Nassehi (9). This study shows the influence of the extent of free and porous domains, which govern the entire flow behavior.

The study of combined free and porous flow in the subsurface using different methodologies is still relatively new in scientific literature. Although a number of issues have been studied, future research aimed at better understanding of combined flow behavior must be continued. For example, investigations into the effects of heterogeneity in porous media and multiphase flow (e.g., oil and water) in combined domains should be carried out. A general methodology, which is not restricted to systems where free/porous interfaces are aligned to coordinate axes, also needs to be developed.

BIBLIOGRAPHY

1. Das, D.B., Nassehi, V., and Wakeman, R.J. (2002). A finite volume model for the hydrodynamics of combined free and porous flow in sub-surface regions. *Adv. Environ. Res.* **7**(1): 35–58.
2. Darcy, H.P.G. (1856). *Les Fontaines Publiques de la Ville de Dijon, Exposition et Application des Principes à Suivre et des Furmules à Employer dans les Questions de Distribution d'Eau*. Victor Dalmont, Paris.
3. Brinkman, H.C. (1947). A calculation of the viscous force exerted by a flowing fluid on a dense swarm of particles. *Appl. Sci. Res.* **A1**: 27–34.
4. Beavers, G.S. and Joseph, D.D. (1967). Boundary conditions at naturally permeable wall. *J. Fluid Mech.* **30**(1): 197–207.
5. Jones, I.P. (1973). Low Reynolds number flow past a porous spherical shell. *Proc. Cambridge Philos. Soc.* **73**: 231–238.
6. Devito, K.J., Waddington, J.M., and Branfireun, B.A. (1997). Flow reversal in peatlands influenced by local groundwater systems. *Hydrological Processes* **11**: 103–110.
7. Das, D.B. and Nassehi, V. (2001). LANDFLOW: A finite volume model of combined free and porous flow of water in contaminated land sites. *Water Sci. Technol.* **43**(7): 55–64.
8. Das, D.B. and Nassehi, V. (2003). Modelling of contaminant mobility in underground domains with multiple free/porous interfaces. *Water Resour. Res.* **39**(3), doi: 10. 1029/2002 WR001506.
9. Das, D.B. and Nassehi, V. (2002). Development of a new mathematical model for subsurface water quality management. *Water Sci. Technol.* **45**(9): 301–307.

READING LIST

Das, D.B. (2001). *Computer Modelling of Contaminant Mobility in Subsurface Regions*. Ph.D. Thesis, Chemical Engineering Department, Loughborough University, Loughborough, UK.

GROUNDWATER TRACING

Hamid Jahani R.
Water Research Institute
Hakimieh, Tehran, Iran

INTRODUCTION

Groundwater tracing is now a major field in groundwater hydrology aiming mainly at demonstrating the direction of groundwater flow and secondly, at giving as much information on the hydrologic setting as possible. It has been used for several decades to study flow conditions, for groundwater reconnaissance, and for managing water resources. During the last decades, water protection has been one of the main issues due to strong pollution of the environment. A groundwater tracer is a substance or a sort of energy conveyed by water to be detected at a sampling point and to be used in analyzing the temporal and spatial distribution of water and the pollutant. Artificial and natural tracers can help in various ways to assess and to predict contaminant transport in surface and groundwater.

Groundwater tracing can be a component of a large hydrogeological program, or it can be a hydrogeological program by itself. Nowadays, new field methods, new tracers, and new measurement techniques are being developed to make hydrogeologists more satisfied with results.

The advantage of groundwater tracing is that it provides the possibility of a direct view of a groundwater system, so the results can be relied on with assurance. In a groundwater tracing program, depending on the question and technical problem, both “natural” (those existing in the water cycle) and “artificial” (those to be introduced into groundwater system) tracers can be selected. The main group of artificial tracers, is soluble tracers, such as fluorescent dyes that are highly detectable, as well as different salts. Another group of artificial tracers includes drift bodies, such as spores, bacteria, or phages. The applicability of each tracer depends on its properties and on the condition of the flow environment. The use of certain tracers, especially for groundwater flow, is strongly restricted due to hydrogeologic conditions and aquifer properties.

For artificial tracers, the methodology comprises introducing the tracer into the groundwater system, letting it travel together with the water, sampling, and finally analyzing the water samples. For natural tracers, the natural physical, chemical, or biological composition of water is traced. For example, the natural temperature of cold water from a spring could be a guide to identifying snowmelt as the main water source and differentiating it from local groundwater.

THE HISTORICAL BACKGROUND

The groundwater tracing method is as old as the interest of humans in nature. It was used while people were curious about finding the source of undiscovered or unknown caves or springs and used chaff, wood bits, or other drifts to reveal the route of the water. The way to perform the test is now almost as simple as it was in the past, at least in most cases. But the way to analyze samples and the results evolved into complicated analytical methods. The development of new tracers and new analytical methods during the last 10 years has led to more accuracy in the results and also more ability to cover longer distances in tracing tests. While using the method during the last decades, it was aimed to find the direction of groundwater flow and its velocity; now, its major application is in studying groundwater pollution problems and protecting the environment.

GOALS OF GROUNDWATER TRACING

The most important goals to be achieved in a tracing program can be categorized into two groups:

Quantitative goals: To determine

- groundwater recharge amount
- groundwater discharge amount
- components of groundwater recharge
- age of groundwater
- elevation of recharge
- groundwater velocity
- volume of aquifer dynamic storage
- dispersion coefficient
- travel time

Quantitative goals: To determine

- origin of water
- hydraulic connection between different surface water and groundwater bodies
- hydraulic connection between different basins
- spring catchment area
- aquifer recharge areas
- aquifer discharge areas
- aquifer boundaries
- groundwater flow direction and path
- flow regime
- role of structural features in groundwater flow
- saline water–freshwater interface
- source of pollution
- pollution distribution pattern
- degree of karstification
- seepage/leakage zones in dams and reservoirs

CHARACTERISTICS OF IDEAL TRACERS

Among the characteristics for selecting a tracer as an ideal applicable tracer are the following:

The tracer

1. should be detectable in low concentrations. In most cases, the tracer material is to be diluted in a big volume of groundwater storage, so it must be easily detectable at low concentrations in water samples. The tracer concentration should be easily differentiated from background values in sampled water.
2. should be applicable as a nontoxic additive to water. It is compulsory for the tracer material not to harm flora and fauna.
3. should not be an absorptive material. The tracer substance should not be easily adsorbed to aquifer material; otherwise, it is possible that a considerable portion of the injected tracer is lost during the test, and it may be a major source of error in interpreting the results. The tracer should move in the same manner as water, and should not interact with aquifer material.
4. should be economically available. In some tracing experiments, the cost of tracer is a decisive factor.
5. should not change the hydraulic conductivity of the medium under investigation.
6. should be stable enough during the test period from injection until detection.

It is obvious that one could not find an ideal tracer that meets all of these specifications. Although there are some tracers developed specifically for groundwater tracing, selecting a suitable tracer, as well as applying a suitable amount of the tracer, is a state-of-the-art task and cannot be done without sufficient experience, favorable expertise, and knowledge about the medium under study.

GROUNDWATER TRACER TYPES

Groundwater tracers can be categorized into two main groups of artificial (to be introduced into the system) and natural (environmental or preexisting in the system) tracers (Fig. 1). Natural tracers are those present in water molecules or accompany water molecules naturally along their path, such as deuterium (^{2}H), carbon-13, or water temperature. Discharge pulses are often good tracers for water especially in the response of springs to flash rainfalls. Artificial tracers are much more diverse in type, and there are many more tracers not mentioned.

Natural (Environmental) Tracers

Natural or environmental tracers are those that already exist in the natural environment, the hydrologic cycle. They could be substances or some kinds of energy.

Environmental Isotopes. Environmental isotopes used as groundwater tracers are isotopes of hydrogen, oxygen, and carbon. They could be part of a water molecule (^{2}H or deuterium, ^{18}O, and ^{3}H or tritium), or they might enter the hydrologic cycle and move together with water molecules (^{13}C, ^{14}C, ^{14}N, ^{15}N, ^{32}S, ^{34}S, ^{36}S ...).

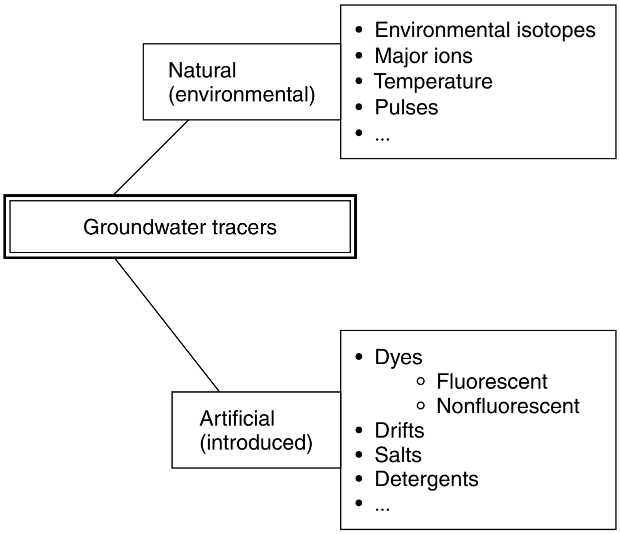

Figure 1. Groundwater tracer types.

Stable Isotopes

Hydrogen and Oxygen Isotopes. Analysis of the natural concentration of these isotopes in water has helped in understanding groundwater flow. Because of the large differences in mass between the two hydrogen isotopes, they tend to fractionate whenever evaporation or condensation takes place. Other factors being equal, waters of higher deuterium content are found near coastlines, at low elevations, in warm rains, and where partial evaporation occurs, as in saline lakes. The variations in oxygen isotope almost follow those of deuterium. The most common use of studies of ^2H and ^{18}O has been to trace the regional movement of groundwater and to locate recharge areas.

Nitrogen. The two abundant isotopes of nitrogen (^{14}N and ^{15}N) can vary significantly in nature. Ammonia escaping as vapor from decomposing animal wastes, for example, tends to remove the lighter nitrogen (^{14}N) and leaves behind a residue rich in nitrogen. In contrast, many fertilizers that have an ammonia base are isotopically light. Natural soil nitrate is somewhat between these two extremes. As a consequence, nitrogen isotopes have been useful in helping to determine the origin of unusually high amounts of nitrate in groundwater.

Most nitrogen in groundwater is in the form of nitrate anion (NO_3^-) or dissolved nitrogen gas (N_2) from the atmosphere. Locally, in zones devoid of dissolved oxygen, the chemically reduced form (NH_4^+) may predominate. In general, nitrate moves as a conservative tracer and is an important indicator of pollution. If nitrate concentrations exceed about 10 mg/L, the health of infant mammals including humans may be adversely affected. Also the presence of more than about 5 mg/L of nitrate commonly is an indirect indication of other forms of contamination, including those from fertilizers and sewage.

Sulfur. Most dissolved sulfur within shallow groundwater is bound within the sulfate ion (SO_4^{2-}). The stable sulfur isotopes (^{32}S, ^{34}S and ^{36}S) found in the sulfate ion vary quite widely and, under certain circumstances, can be useful indicators of the origin of the sulfate, which is particularly true if, for example, one wishes to distinguish sulfate originating from natural dissolution of gypsum ($CaSO_4 : 2H_2O$) from sulfate originating from an industrial spill of sulfuric acid.

Active Isotopes

Tritium. Tritium (^3H) is an unstable isotope of hydrogen whose half-life is 12.32 years. It enters the hydrologic cycle in the same manner as stable hydrogen in the H_2O molecule. This isotope is naturally produced at low levels in the atmosphere. During and after the nuclear tests in the Northern Hemisphere in the 1960s, the tritium concentration (referred to as TU: Tritium Unit) in the atmosphere increased from a level of 5–15 TU to up to 4000 TU. ^3H can be used qualitatively to date groundwater groundwater that has less than 2 to 4 TU is dated prior to 1953; if the amount is significantly greater than 10 to 20 TU, it has been in contact with the atmosphere since 1953. Tritium analysis in groundwater studies is mainly for absolute groundwater dating for ages up to 40 years.

Carbon-14. ^{14}C is the radioactive isotope of carbon. It has a half-life of 5730 years and is mainly used to date groundwaters up to 30,000 years of age. Most ^{14}C in potable groundwater is contained in the HCO_3^- ion in water. As a first approximation, the initial number of ^{14}C nuclei per total carbon nuclei, or X_0 in Eq. 1, in a water sample is considered to have been constant prior to 1950 due to the almost constant natural production of ^{14}C in the atmosphere by cosmic radiation interacting with the atmosphere.

$$X_t = X_0 e^{-\lambda t} \tag{1}$$

where λ = decay constant
t = time
X_0 = number of nuclei at zero time
X_t = number of nuclei at time t

If the only source of ^{14}C in water is originally from the active biosphere, then the ^{14}C that is measured in carbon from a water sample can be considered X_t in the equation because λ is known from experimental work, the age of the sample (t) can be determined directly. Due to some complexities, the use of ^{14}C in age determination is not so simple. The existence of carbon from carbonate formations, or formation of CH_4 gas, are factors that affect the results of age determination. But despite that, ^{14}C dating is highly useful in determining the approximate residence time of old groundwater.

Major ions. The chemical composition of groundwater is in most cases a good indicator of the environment through which the water has passed. In almost all groundwater tracing and isotopic hydrologic programs, water is analyzed for calcium, sodium, magnesium, potassium (major cations), chloride, sulfur, carbonate, bicarbonate, and nitrate (major anions).

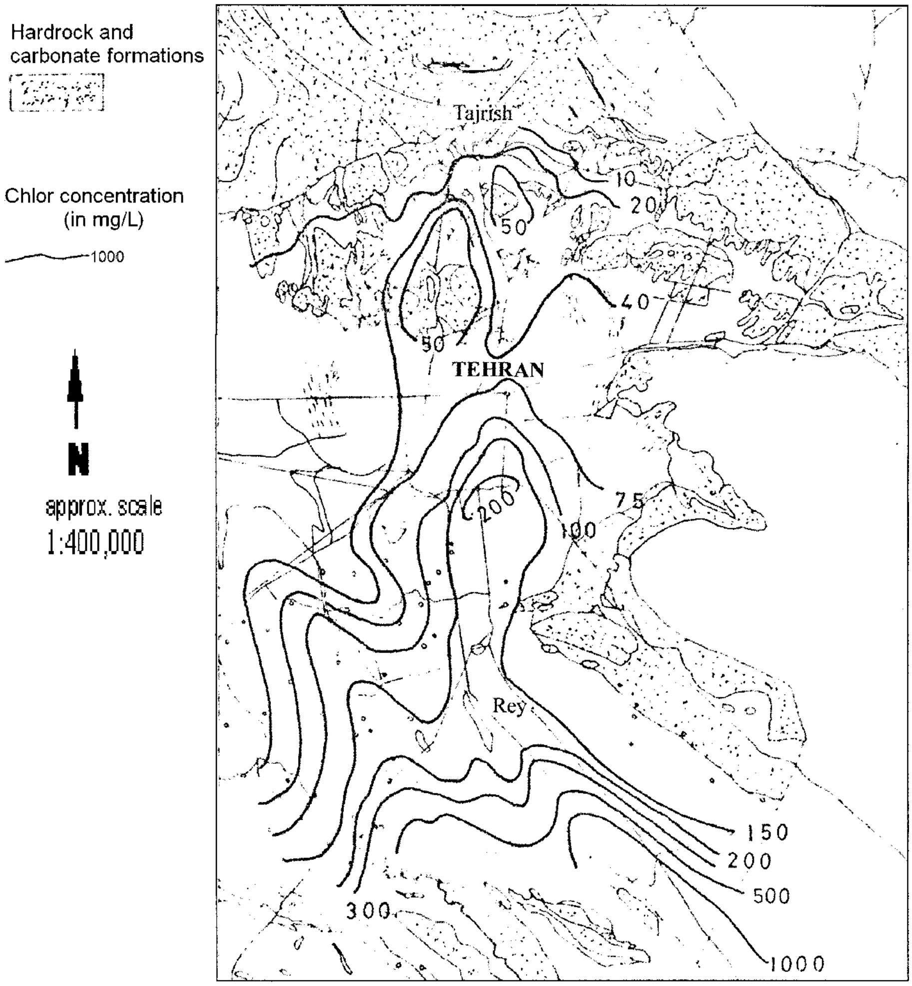

Figure 2. Isochlor map for Tehran plain, 1977, Tehran, Iran. The general groundwater flow direction is southward. As can be seen, groundwater gets richer in ions (specifically in chlor) as it flows downstream.

Typically, if we follow a recharge event and track the water down-gradient in the flow path, it can be seen that the ionic concentration in the water increases as it migrates down-gradient (Fig. 2). Generally, groundwater is the bicarbonate type in recharge areas the sulfate type midway, and the chloride type in terminal zones (e.g., basin outlet). The chemical composition of groundwater would finally be demonstrated on specific presentation charts such as Piper, Durov, Schuler, ... and different groundwater types could be easily recognized.

Temperature. Water temperature is almost a good indicator or tracer where there is an abrupt difference in the physical properties of the aquifer (e.g., a change in hydraulic conductivity, fracture properties ...) or where there is a significant temperature gradient (Fig. 3) Thermometry is a study tool for investigating changes in rock properties in a borehole column. It comprises a heat sensor that goes down through the borehole and records the water temperature along the borehole column. The record is called a temperature log.

Pulses. Pulses are sudden changes in the amount of a specific parameter that can be easily seen as a peak in a graph of the parameter versus time. It could be the graph of a spring discharge, electrical conductivity, turbidity, etc. A pulse is normally an indicator of a hydrologic event upstream. For example, a sudden rise in the amount of turbidity in a spring could be an indication of a storm in a spring recharge area. It could also be a sign of a hydraulic connection between the storm area and the spring.

Artificial Tracers

Artificial tracers are those that are introduced intentionally into groundwater. So an "injection point" is necessary in artificial tracing in addition to a sampling network.[1] While in environmental tracing, the study is performed only through a sampling network. Normally, artificial tracing and natural tracing are complementary, and sometimes both methods are applied to the same area to achieve more reliable results. Among all artificial tracers, the most applicable are fluorescent dye tracers which are discussed at the end for more emphasis.

Drifts. In areas such as karstified limestone terrains or highly faulted volcanic rocks where groundwater flow follows a turbulent pattern, drifts usually are good choices to be selected as tracers. Signal-emitting floats, yeasts,

---

[1] Sometimes the injection point and sampling point are the same. This is the case in the so-called single-well tracing method.

bacteria, viruses, and spores have reportedly been used as tracers in many cases.

Salts. Most tracers that have already been used extensively are common salts. This category of tracers includes those compounds that ionize in water, resulting in separation into charged elements possessing a positive charge (cations) or a negative charge (anions). The most common ionic tracers are chloride (Cl^-), bromide (Br^-), lithium (Li^+), and potassium (K^+). These tracers have been successfully used to determine flow paths, residence time, and aquifer properties. The most important factors that should be considered before applying such tracers are aquifer material texture and natural background concentrations in the aquifer.

The advantages of salt tracers are their availability, low cost, and stability, but they must be injected in large amounts, hundreds of kilograms or tons, and this causes problems in handling and transportation. As they are injected in huge amounts, they make water denser around the injection point, and it may affect the hydrodynamic behavior of the water.

Fluorescent Dyes. For many reasons, fluorescent dyes have become very important as tracers in hydrology and hydrogeology. They can be used in hydrology to measure transit times in streams and rivers, to measure discharge in open channels by the tracer dilution method, and to investigate dispersion processes in rivers, lakes, and seas. Of course their main application is in hydrogeologic studies for estimating hydrodynamic parameters as well as studying sources of pollution.

Application of fluorescent tracers has some great advantages and a few disadvantages:

- They are visible without the use of any instrument at higher concentrations (say more than 10 ppb). This property can be useful for qualitative experiments.
- The limit of detection for the dyes is in the sub-ppb range and therefore much better (up to a factor of 1000 to 10000) than salt tracers. So a small amount of tracer is enough.
- Instrument detection is quick and easy without time-consuming sample preparation.
- The chemical properties of the dyes are in some cases better because salts (particularly the cationic part) are influenced by ion exchange in clayey sediments.
- Fluorescent dyes do not influence the density of water near the injection point.
- Depending on the type of analytical instrumentation, the costs to determine fluorescent dyes are cheaper (using an ordinary photometer) or comparable to a flame photometer which can be used to analyze salt tracers. The detection of salts by electrical conductivity measurement is obviously cheaper but cannot distinguish between different types of salt tracers.
- A disadvantage is the relatively high cost of fluorescent dyes, but, on the other hand, up to 1000 times less dye than salt is necessary for a tracing experiment.

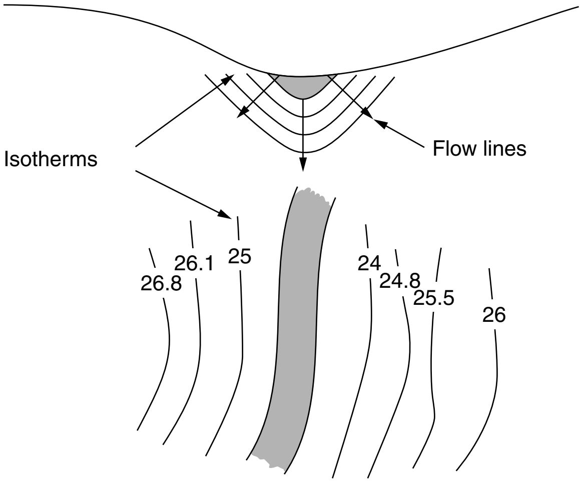

Figure 3. Sketch showing a section of a temperature gradient due to freshwater recharged from the river (up) and the plan of isotherms at both sides of the river (down).

These are the most commonly used fluorescent dyes:

Uranine. For hydrologic or hydrogeologic purposes, uranine ($C_{20}H_{10}O_5Na_2$, Color index name: Acid Yellow 73) is the most important tracer (Fig. 4). It is also the most sensitive of all. The dye is supplied as a hygroscopic orange-red powder: it is very soluble in water and has a yellowish-red color and an intense yellowish green fluorescence. In solutions at concentrations higher than 1 g/m^3 (equal to 1 ppm), the color is orange-red. The fluorescence is visible without any instrument above concentrations of 50 mg/m^3 (50 ppb) and approximately above 10 mg/m^3 (10 ppb) using a UV lamp, depending on the quality of the dye. It has no specific odor. At high concentrations, the fluorescent light is absorbed by the sample itself. The intensity of the fluorescence can be measured exactly only if the concentration is much lower (less than 100 ppb); otherwise, the solution has to be diluted.

The solubility of Uranine is 500 mg/L, and its detection limit is 0.001 mg/m^3. Its fluorescence intensity in solution is very much dependent on the pH. At pH lower than 6, the intensity tends to zero. Uranine is an anionic tracer, and it is adsorbed by clay minerals only to a minor extent. It has also good temperature stability which is important for analytics (a temperature increase in the instrument) and for injection, for example, into thermal systems. From

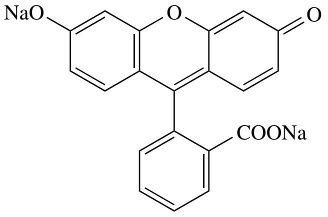

Figure 4. Molecular structure of uranine.

the view point of toxicity, it is not dangerous to humans or animals, in particular, at concentrations normally used in tracing experiments.

Eosin. Eosin ($C_{20}H_6Br_4Na_2O_5$, Acid Red 87) (Fig. 5) is a powder made of red crystals of a bluish tinge. A concentrated aqueous solution is deep brownish-red, dilute solutions are yellowish-red with a greenish fluorescence. The fluorescence is visible in more concentrated solutions (above 10 ppb).

It is very soluble in water, less in alcohol, and insoluble in ether. At high concentrations, fluorescent light is absorbed by the sample itself. For an exact analysis, the concentration should not be higher than 100 ppb, otherwise, dilution is recommended. Its solubility is 300 g/L and the detection limit is 0.01 mg/m^3. The pH effect for eosin is not as significant as that for uranine. Eosin is an anionic tracer and is generally little adsorbed by clay minerals. It can be used in karstic systems as well as in porous groundwater aquifers. It is not dangerous to humans or animals.

Among other fluorescent dyes, sulforhodamine G, Sulforhodamine B, rhodamine WT, rhodamine B, and pyranine are used most groundwater tracing.

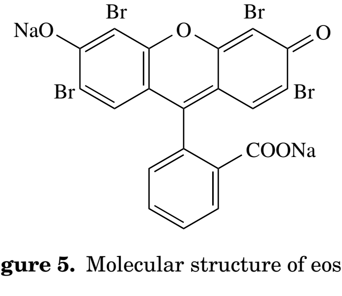

Figure 5. Molecular structure of eosin.

Tracing Experiment

While planning a tracing test, there are some factors that must be considered. First, the objectives of the study should be defined well because they affect other study components such as selecting the tracer. Then all available resources (labor, money, time,) should be considered to facilitate formatting the program. Third, all relevant information and data (maps, reports, well logs) should be available so that a good idea of the geology and hydrogeology of the site is gained.

The general direction of groundwater flow and also its velocity should be estimated before starting the test, which will allow configuring the injection and sampling network effectively.

Different Phases of the Test

Planning and General Assessment. Before starting a tracing test, the test site should be visited and examined carefully, which is to schedule the program activities and other components of the work such as selecting the injection and sampling points, selecting the tracer (which is among the most important and critical components of the work), designing the sampling program, transportation (equipment, tracers, personnel, etc.

Tracer Injection. By injecting the tracer into the aquifer, it will enter the hydrologic cycle and (it is assumed that) will travel in the same manner as water. The injection method varies with the type of tracer, study purpose, site conditions, etc.

Sampling. Sampling is done for the purpose of analyzing water to find if there is tracer in it. Actually the sampling work starts before injecting the tracer because there might be some background concentration of the tracer in the water; this is called background sampling.

Because of the very high dissemination potential of fluorescent tracers, it is very important to avoid any contact between injecting personnel/equipment and sampling. The sampling time interval for each sampling point should be scheduled according to the hydrogeologic factors and distance to the injection point.

Analysis. All samples should be carefully packed and sent to the lab to be analyzed as soon as possible. The analytical method and the instrumentation for measuring the fluorescence of the samples are discussed following. As mentioned before, concentrations higher than 10 ppb of most fluorescent dyes may be visible and can be detected without any instrument.

Detection with a Two Monochromator Spectrofluorometer. To analyze for tracers, the best method is to use a spectrofluorometer with monochromators for excitation and emission (Fig 6). The light of the excitation source is transmitted through a filter of maximum transmittance near or at the absorption maximum of the dye. The dye molecules in the solution will be excited, and the fluorescence can then be dispersed by a secondary filter or a monochromator (prism or grating). Table 1 is a summary of spectral data of some fluorescent dyes, and Table 2 is a summary of their approximate detection limits in water. The final result of an analysis is a time–concentration (or breakthrough) curve, which presents the temporal evolution of the sampling point tracer concentration (Fig. 7).

Interpretation of Results. With regard to the breakthrough curve plotted for each sampling point, one should

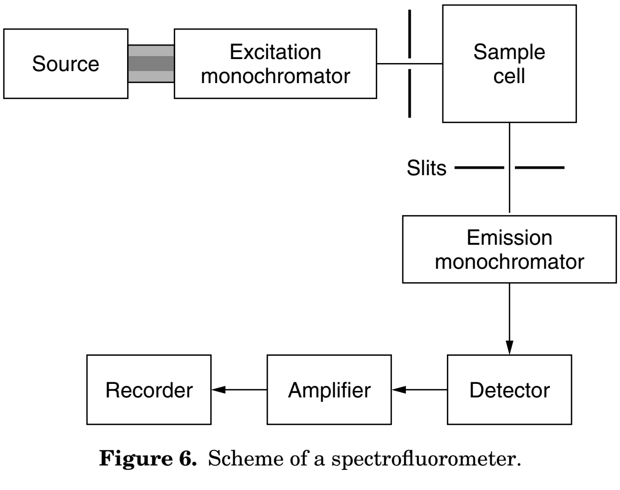

Figure 6. Scheme of a spectrofluorometer.

Table 1. Summary of Spectral Data of Some Fluorescent Dyes

Dye	Fluorescent Excitation	Fluorescent Emission	Wavelength Difference (nm)
Uranine	491.0	512.5	21.5
Eosin	515.5	536.0	20.5
RhodamineB	553.5	576.0	22.5
Pyranine	463.5	512.5	49.0

Table 2. Summary of Approximate Detection Limits for some Fluorescent Dyes

Dye	Limit of Detection (mg/m^3)
Uranine	0.001–0.025
Eosin	0.050–0.5
Rhodamine B	0.010–0.5
Pyranine	0.008–0.5

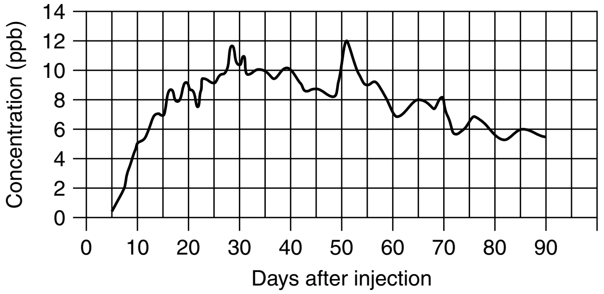

Figure 7. The breakthrough curve for a spring in the Tang-e Shemiran area, western Iran.

first of all focus on the first arrival time of the tracer. It reveals first, that there is an underground hydraulic connection between the injection point and the sampling point and, second, the maximum groundwater flow velocity in that zone. The peak concentration also leads to the average transit time. The shape of the curve is a guide to the approximate volume of water in which the tracer has been diluted. The area under the curve is equivalent to the total mass of tracer recovered from the sampling point (e.g., a spring). The summation of this figure for all sampling points should be close to the injected amount. Otherwise, there is another outlet for the aquifer which was unknown during planning.

READING LIST

Benischke, R. (2002). *Fluorescent Tracers in Hydrology*. Institute of Hydrogeology and Geothermics, Graz, Austria, p. 60.

Davis, S.N. (1985). *Groundwater Tracers*. National Groundwater Association, Westerville, OH, p. 200.

Ighanian, R. (1977). *Hydrochemical Studies of Tehran Plain*. Ministry of Energy, Tehran, Iran, p. 32.

Jahani, H.R. (2000). Groundwater Tracing in Tang-e Shemiran, *Western Iran*. Water Research Center, Tehran, Iran, p. 85.

Jahani, H.R. (1998). *Groundwater Tracing and Isotope Hydrological Studies in Rudbar-e Lorestan Area, Iran*. Water Research Center, Tehran, Iran, p. 83.

Standards Office. (2003). *Standard Guidelines on Groundwater Tracing in Karst and Hardrock Formations (Draft Edition-Farsi)*. Iranian Ministry of Energy, Tehran, Iran, p. 60.

HYDRAULIC CONDUCTIVITY/ TRANSMISSIBILITY

Steven A. Dielman
ENVIRON International
Corporation
Arlington, Virginia

Hydraulic conductivity is a fundamental parameter that governs the flow of liquids such as groundwater through aquifers and other porous media. Specifically, hydraulic conductivity is a quantitative measure of the capacity of a geologic formation or other porous media to transmit a specific fluid. It is determined by the characteristics of both the porous medium and the fluid of interest.

Transmissibility (or transmissivity) is a property closely related to hydraulic conductivity that describes the capacity of a specific water-bearing unit of a given thickness, such as an aquifer, to transmit water. Transmissibility is most simply defined as the effective hydraulic conductivity of an aquifer or other water-bearing unit multiplied by the thickness of that unit.

ORIGINS IN DARCY'S LAW

The concept of hydraulic conductivity dates back to Henry Darcy's famous 1856 paper regarding the water supply for the city of Dijon, France. In that paper, Darcy described the results of laboratory experiments that he had conducted to investigate the flow of water through sand. From the results of those experiments came the relationship now referred to as Darcy's law, it is valid for laminar flow conditions that occur in most groundwater systems and can be expressed as

$$Q = -K\frac{\partial h}{\partial l}A \tag{1}$$

where Q = volumetric flow rate (L^3/t)
K = hydraulic conductivity (L/t)
$\frac{\partial h}{\partial l}$ = hydraulic gradient (L/L)
A = Cross-sectional area perpendicular to the flow direction (L^2)

As evident from Eq. 1, the volumetric flow of a fluid through a porous medium is directly proportional to the coefficient K; Darcy recognized that it is dependent on the type of soil through which water flows. For this reason, the coefficient K has historically been referred to as Darcy's coefficient of proportionality. Today, the coefficient K is more commonly called hydraulic conductivity.

RELATIONSHIP TO PERMEABILITY

It is intuitive that the flow of a liquid through soil also depends on the characteristics of the liquid. For example, if all other factors are equal, the volumetric flow rate (Q in Darcy's equation above), and hence the hydraulic conductivity, will be greater for water than for a more viscous liquid such as a heavy oil.

The relationship of hydraulic conductivity (K) to permeability (k)[1] is useful in distinguishing the effects of the porous medium from the effects of the fluid:

$$K = \frac{k\rho g}{\mu} \tag{2}$$

where ρ = density of fluid;
μ = dynamic viscosity of fluid; and
g = gravitational acceleration constant.

Permeability k is a function only of the medium and is commonly used in petroleum studies where various fluids of differing characteristics such as oil, water, and natural gas may be present together in multiphase flow systems.

From Eq. 2, it is clear that hydraulic conductivity is a function of the characteristics of the fluid as well as the permeability of the medium (k).[2] Specifically, a soil with a given permeability will exhibit relatively greater hydraulic conductivity to fluids of lower viscosities and/or greater densities.

Rearranging Eq. 2 also allows determining the hydraulic conductivity to one fluid if the hydraulic conductivity to another is known:

$$K_1 = K_2 \frac{\mu_2}{\mu_1}\frac{\rho_1}{\rho_2} \tag{3}$$

where K_1, μ_1, and ρ_1 are properties of fluid 1 and K_2, μ_2, and ρ_2 are properties of fluid 2.

UNITS OF MEASUREMENT

The units of hydraulic conductivity are (L/t), and the units of permeability are (L^2). In the context of groundwater flow, the units of permeability are sometimes expressed as "darcys"; one darcy approximately equals 1×10^8 cm^2. Conversion factors for commonly used units of hydraulic conductivity and permeability are provided in Table 1. For most applications, the viscosity and density of water vary over a narrow range, and permeability values can be converted to hydraulic conductivity (of water), assuming standard temperature and pressure, without introducing significant error.

Because transmissibility is defined as hydraulic conductivity multiplied by aquifer thickness, its units are (L^2/t), often expressed as volume per unit width of the aquifer per unit time (e.g., gallons per day per foot or gpd/ft).

REPRESENTATIVE HYDRAULIC CONDUCTIVITY VALUES

Local hydraulic conductivity values can vary over an extremely wide range. Ranges in representative values of hydraulic conductivity for various geologic media are summarized below.

Figure 1 illustrates that hydraulic conductivity values may range over approximately 12 orders of magnitude, depending on the type of soil or rock. Even for specific types of geologic deposits, variations over several orders of magnitude are not uncommon.

In general, geologic deposits that exhibit the lowest hydraulic conductivity values are fine-grained unconsolidated materials (e.g., clays) and unfractured sedimentary and crystalline rocks. The presence of fractures in either consolidated or unconsolidated materials of otherwise low permeability can significantly increase the hydraulic conductivity of these materials.

Geologic deposits that exhibit the highest hydraulic conductivity values are typically coarse-grained unconsolidated materials (e.g., coarse sands and gravels) and rocks that are highly fractured or otherwise exhibit significant secondary permeability (e.g., solution-enhanced void space in carbonate formations or basalt shrinkage cracks).

GEOLOGIC FACTORS INFLUENCING HYDRAULIC CONDUCTIVITY

For unconsolidated materials, the factors that have the greatest influence on hydraulic conductivity are generally particle size and particle sorting. It can be theoretically shown that hydraulic conductivity is proportional to the square of the particle diameter for a uniform, porous medium, that is, doubling the particle diameter results in quadrupling hydraulic conductivity. In practice, however, geologic formations are not uniform but contain particles of varying sizes, and the definition of a representative particle diameter is not clear-cut. In general, poorly sorted deposits (i.e., those with greater variation in particle sizes) will exhibit lesser relative hydraulic conductivity compared to that of well-sorted deposits, due to infilling of pore space by finer grained particles.

In consolidated crystalline and sedimentary rocks, the hydraulic conductivity of the rock matrix tends to be low to extremely low. Transmission of water through these rocks is generally via interconnected voids such as fractures, solution channels, shrinkage cracks or metamorphic (e.g., schist) foliations. Transmission of water through such voids is said to occur via "secondary permeability."

HETEROGENEITY AND ANISOTROPY

Geologic depositional environments typically exhibit significant variability across very short distances. As a result, hydraulic conductivity values generally exhibit

---

[1] In older publications, K was sometimes referred to as the coefficient of permeability and, therefore, k become known as "specific" permeability or "intrinsic" permeability to differentiate it from K. More recently, it has become conventional to refer to K as hydraulic conductivity, allowing k to be referred to simply as permeability.

[2] It is common practice in groundwater hydrology publications to use the term hydraulic conductivity as it relates to water, unless a different fluid is specified.

Table 1. Hydraulic Conductivity and Permeability Conversion Factors[a]

	Multiply To obtain	Hydraulic Conductivity					Permeability	
		cm/s by	m/s by	ft/s by	ft/day by	US gal/day/ft^2 by	cm^2 by	Darcy by
Hydraulic Conductivity	cm/s	1	100	30.48	3.53E−04	4.72E−05	9.71E+04	9.66E−04
	m/s	0.01	1	0.305	3.52E−06	4.72E−07	9.71E−02	9.66E−06
	ft/s	3.28E−02	3.28	1	1.16E−05	1.55E−06	3.19E+03	3.17E−05
	ft/day	2.84E+03	2.84E+05	8.64E+04	1	0.134	2.75E+08	2.74
	US gal/day/ft^2	2.12E+04	2.12E+06	6.45E+05	7.48	1	2.06E+09	1.48E+01
Permeability	cm^2	1.03E−05	1.03E−03	3.14E−04	3.63E−09	2.71E−08	1	9.87E−09
	Darcy	1.04E+03	1.04E+05	3.15E+04	3.65E−01	2.72	1.01E+08	1

[a] Conversions between units of hydraulic conductivity (K) and permeability (k) assume the properties of water at approximately 15 °C.

spatial variation across several orders of magnitude within an area of hydrogeologic interest.

The terms *heterogeneity* and *anisotropy* are often employed to describe variability in hydraulic conductivity. Heterogeneity refers to variability in space, and anisotropy refers to variability in direction from a given point. Figure 2 illustrates the relationship between heterogeneity and anisotropy.

The distinction between anisotropy and heterogeneity often depends on scale and context. For example, on one scale, layered heterogeneity (see Fig. 4) may require consideration as a series of distinct geologic formations that have different representative hydraulic conductivity values. On another scale, it may be more convenient and appropriate to describe this system as one anisotropic layer.

Heterogeneity

A porous medium of uniform characteristics is said to be homogenous. In reality, no geologic system is truly homogenous with respect to hydraulic conductivity. Each system exhibits varying types and degrees of heterogeneity. An illustration of actual heterogeneity is shown in Fig. 3. This two-dimensional cross-section was created from more than 700 measurements within an area approximately 2 meters deep by 20 meters long within the intensely studied Borden aquifer, which is generally recognized as “mildly” heterogenous compared to most groundwater systems. Even in this small area of a relatively uniform aquifer, measured hydraulic conductivity values range across approximately three orders of magnitude.

In some cases, heterogeneity in hydraulic conductivity values may exhibit one or more measurable trends due to geologic conditions. Freeze and Cherry (2) identified three types of common heterogeneity patterns: trending, discontinuous, and layered.

Trending heterogeneity refers to spatial trends resulting from, for example, gradational coarsening or fining depositional sequences such as those that typically result from encroaching or receding seas. Trending heterogeneity may also result from postdepositional processes such as vertical trends in rock fracture frequency.

Discontinuous heterogeneity refers to an abrupt change in local hydraulic conductivity values, such as those across a geologic discontinuity, fault, or other sharp stratigraphic contact.

Layered heterogeneity is a commonly encountered type of heterogeneity, particularly for unconsolidated or consolidated sedimentary sequences formed in time-varying depositional environments. An idealized example of layered heterogeneity is illustrated in Fig. 4. A more realistic depiction of layered heterogeneity is illustrated in the Borden aquifer hydraulic conductivity plot depicted in Fig. 3.

In reality, distinguishing between types of heterogeneities is often difficult. For example, layered heterogeneity may be considered a series of discontinuous heterogeneities. In addition, multiple patterns in heterogeneity may be present at any given location.

Anisotropy

If hydraulic conductivity is independent of the direction of measurement, the medium is described as isotropic; if, however, the hydraulic conductivity value depends on the direction of measurement, it is considered anisotropic. Mathematically, hydraulic conductivity is represented as

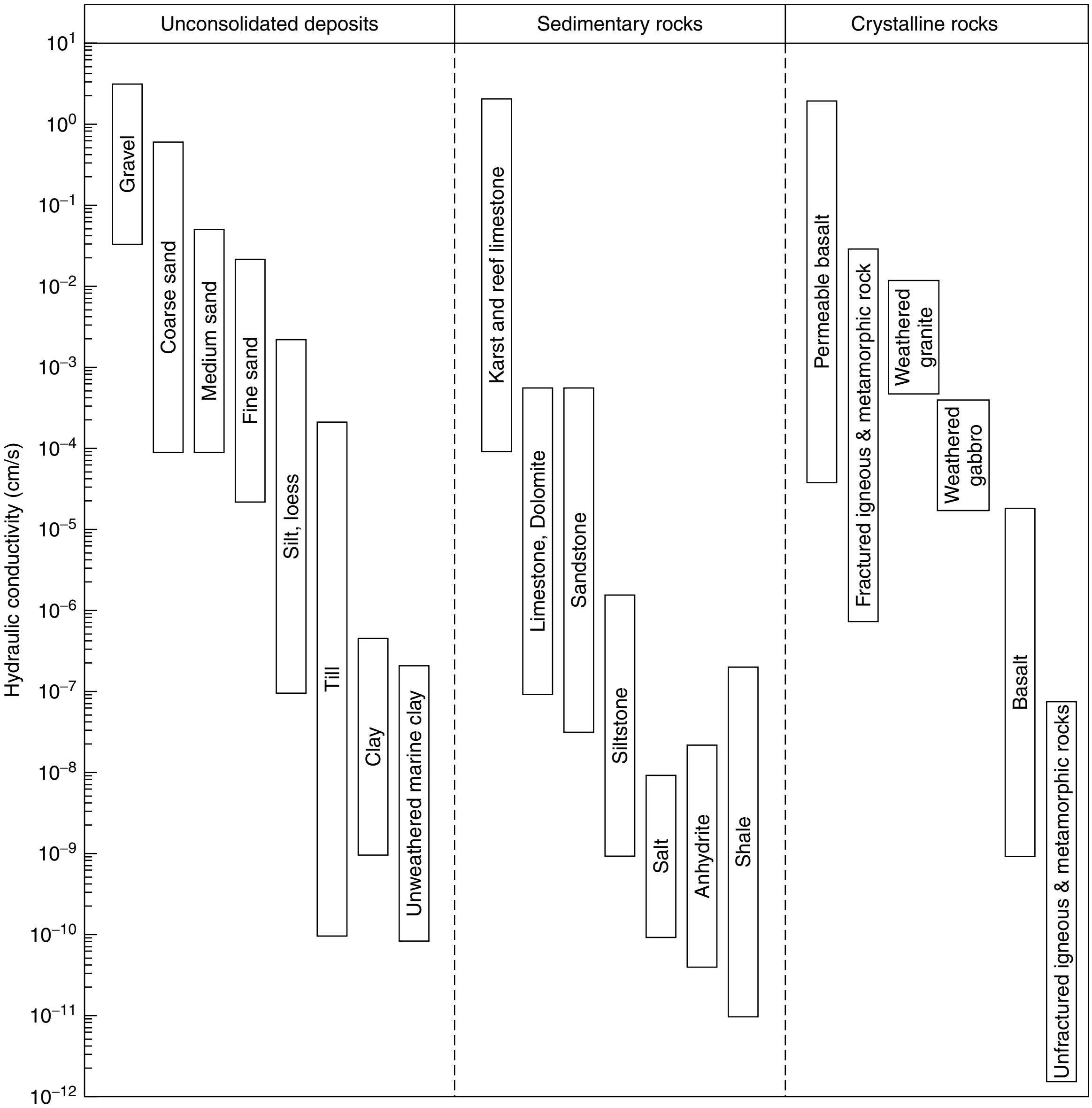

Figure 1. Representative values of hydraulic conductivity for various rock types (adapted from: Ref. 1).

a scalar (direction-independent) variable in the form of Darcy's law presented in Eq. 1 In anisotropic settings, hydraulic conductivity must be considered a tensor variable, leading to more complex forms of Darcy's law and other equations that describe the flow of groundwater and other fluids through porous media.

On a small scale, anisotropy can result from mineral grain orientation. On a larger scale, anisotropy can result from sedimentary bedding planes, metamorphic structure (e.g., shistocity) or directional fracturing of crystalline rocks resulting from postdepositional stresses.

In sedimentary deposits of horizontal or near-horizontal orientation, hydraulic conductivity is frequently greater in the lateral directions compared to the vertical direction. On a small scale, such anisotropy may be relatively modest. Freeze and Cherry (2) report ratios of horizontal hydraulic conductivity (K_{h}) to vertical hydraulic conductivity (K_{v}) between 3:1 and 10:1 for clays and shales due to mineral grain orientation. On a larger scale, however, K_{h}:K_{v} anisotropic ratios for sedimentary deposits may be extremely high due to stratified deposits of greatly different characteristics. High K_{h}:K_{v} anisotropic ratios for stratified deposits are not surprising, considering the large differences in typical hydraulic conductivities of coarse-grained and fine-grained deposits, which may differ by several orders of magnitude.

EFFECTIVE HYDRAULIC CONDUCTIVITY: STATISTICAL TREATMENT OF HYDRAULIC CONDUCTIVITY VALUES

Groundwater hydrologists have used a variety of techniques to characterize the spatial variability of hydraulic conductivity values and to assess groundwater flow in

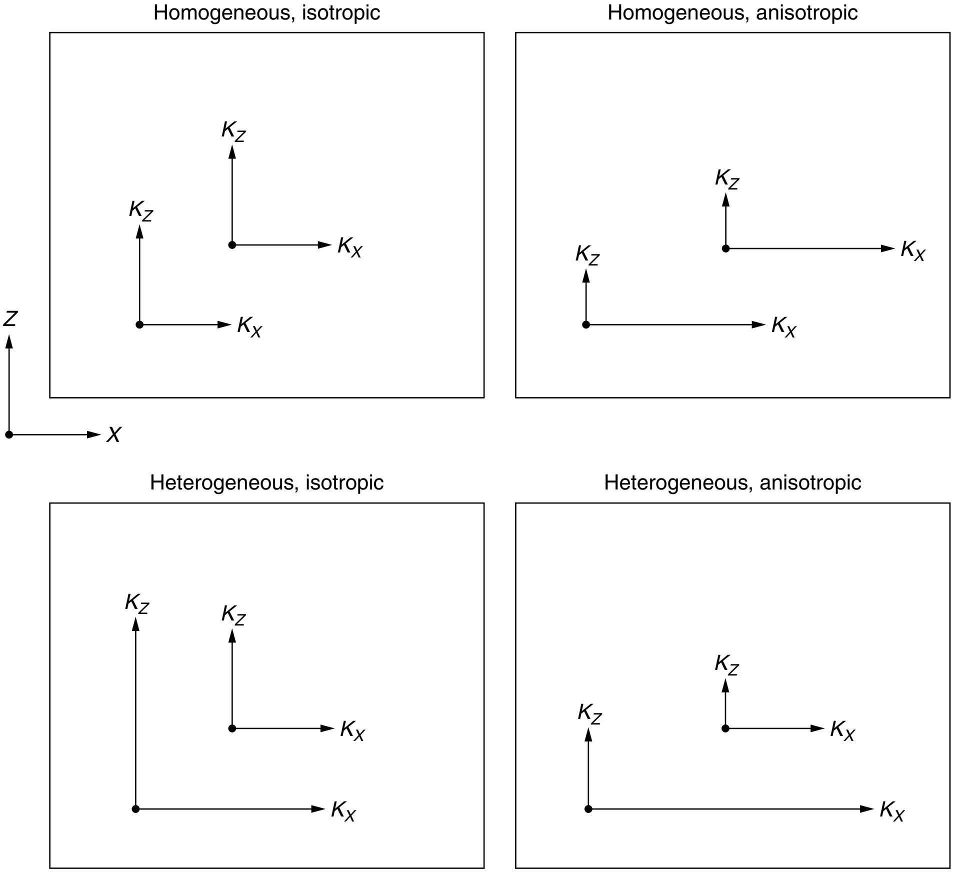

Figure 2. Four possible combinations of heterogeneity and anisotropy (adapted from Ref. 2).

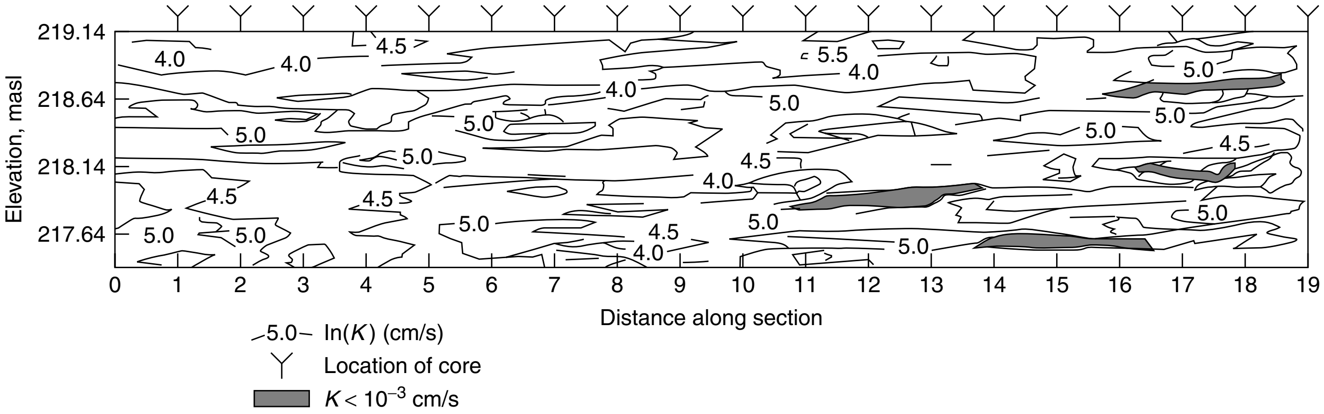

Figure 3. Hydraulic conductivity distribution at Canadian Forces Base Borden (from Ref. 3, v. 22, p. 2069–2082).

heterogenous geologic settings. One such approach is to identify an "effective" hydraulic conductivity that, applied to a heterogenous water-bearing geologic unit, will allow an accurate representation of its ability to transmit water as a whole.

Consider horizontal groundwater flow through an idealized layered geologic system such as that shown in Fig. 4 where the a number of layers equals n. The total flow through the system (Q) is equal to the sum of the flow in all layers (ΣQ_i). From Darcy's law, this can be expressed as

$$Q = \sum Q_i = \sum K_i \frac{dh}{dx} D_i W \tag{4}$$

where

K_i = hydraulic conductivity of layer i,
dh/dx = horizontal hydraulic gradient (assumed uniform in all layers),
D_i = thickness of layer i, and
W = unit width.

For combined flow through the layered system, an effective horizontal hydraulic conductivity ($K_{\text{H-EFF}}$) can be defined for groundwater flow such that

$$Q = K_{\text{H-EFF}} \frac{dh}{dx} DW \tag{5}$$

From Eqs. 4 and 5, the following relationship for $K_{\text{H-EFF}}$ is obtained:

$$K_{\text{H-EFF}} = \frac{\sum K_i D_i}{D} \tag{6}$$

This shows that, for flow parallel to the idealized layered system shown in Fig. 4, $K_{\text{H-EFF}}$ is equivalent to the arithmetic mean of K_i weighted by the thickness of each layer D_i. If all layers are of the same thickness (or all K values are uniformly distributed), $K_{\text{H-EFF}}$ is equal to the arithmetic mean of the individual K_i values.

Consider, however, vertical groundwater flow through the same system. Using a similar analysis, it can be shown that the effective horizontal conductivity in the vertical direction ($K_{\text{V-EFF}}$) is equal to

$$K_{\text{V-EFF}} = \frac{D}{\sum D_i/K_i} \tag{7}$$

If all layers are of equal thickness (or are uniformly distributed), $K_{\text{V-EFF}}$ for the idealized layered system is equivalent to the harmonic mean of K_i values, which is always lower than or equal to the arithmetic mean.

These two simplified conceptual models represent two extreme possibilities of flow uniformly parallel to or perpendicular to geologic layers, respectively. Therefore, the arithmetic and harmonic means of a population of hydraulic conductivity values are generally considered to be upper and lower bounding values, respectively, for the effective hydraulic conductivity of a heterogenous geologic system with respect to groundwater flow through that system.

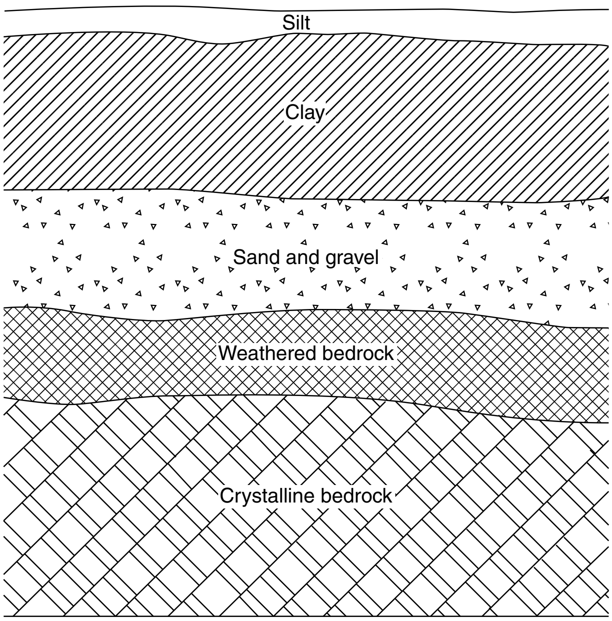

Figure 4. Example of layered heterogeneity.

Frequently, hydraulic conductivity values and well yields within a specific geologic setting are best characterized by a log-normal distribution. As such, the geometric mean is often preferred as a descriptive statistic for the central tendency (e.g., median and mode) of a population of log-normally distributed hydraulic conductivity values. Further, several studies have concluded that the geometric mean provides a better estimate of the effective hydraulic conductivity of a heterogenous groundwater system than the harmonic mean or arithmetic means, which are less than and greater than the geometric mean, respectively.

The preceding provides a simplified discussion of common statistical treatments for hydraulic conductivity values. In recent decades, much work has been conducted to apply more complex statistical models and stochastic techniques to characterize groundwater flow through heterogenous systems.

METHODS OF ESTIMATION AND MEASUREMENT

Empirical Relationships

Several efforts have been made over the years to correlate hydraulic conductivity and/or permeability empirically with characteristics of porous media through which water flows. Some of the more commonly used relationships are provided in Table 2.

Although such empirical relationships correlate hydraulic conductivity with various factors, nearly all can be simplified to the form

$$K = Cd^x \tag{8}$$

where C is a constant or shape factor that is a function of the fluid density and viscosity and may be correlated with properties of the porous medium such as porosity and d is some representative grain size (L). It has been mathematically shown (e.g., by Ref. 5) that for idealized systems of uniform grain size, $x = 2$. For all the empirical relationships shown in Table 2, x ranges between 1.5 and 2, which indicates a disproportionate increase in hydraulic conductivity as grain size increases, but suggests an exponent value of something less than 2 for many real-world situations.

Laboratory Testing

The hydraulic conductivity of geologic deposits can be measured in the laboratory by using a permeameter, a device similar to that originally devised by Darcy. Various configurations of permeameters have been used; the most common are the constant-head permeameter and the falling-head permeameter shown in Fig. 5.

Permeameters consist of a chamber to contain a sample of soil or rock and allow the application of a hydraulic gradient across the sample. Flow through the sample is measured, and hydraulic conductivity is calculated based on principles founded by Darcy. Careful consideration must be used in evaluating permeameter test data. For example, when testing unconsolidated material, was the test conducted on an intact core or on a repacked sample? If the former, the test likely represents the vertical hydraulic conductivity, which may be much lower than the horizontal hydraulic conductivity. If the latter, the disturbance of the sample during repacking may result in a value that is neither representative of the horizontal nor the vertical hydraulic conductivity.

Field Measurements—Aquifer Tests

Hydraulic conductivity and transmissibility on a larger scale are estimated by evaluating data from aquifer tests. There are many different types of aquifer tests and methods for evaluating the test results, but the most common types of tests employed in groundwater studies are pumping tests and slug tests.

Pumping tests involve pumping water from a well at a known rate and measuring the changes in water level through time at the pumping well and/or nearby observation points. Slug tests involve rapid introduction or removal of a volume of water or other material into or out of a well and measuring the changes in water level through time at the well being tested.

Table 2. Examples of Empirical and Semiempirical Relationships for Estimating Hydraulic Conductivity and Permeability[a]

Source	Equation		Parameters
Hazen (1911)	$K = Cd_{10}{}^2$		K = hydraulic conductivity (cm/s) C = constant; 100 to 150 $(\text{cm/s})^{-1}$ for loose sand d_{10} = effective grain size (cm) (10% particles are finer, 90% coarser)
Harleman et al. (1963)	$k = (6.54 \times 10^{-4})\, d_{10}{}^2$		k = permeability, cm^2 d_{10} = effective grain size, cm (10% particles are finer, 90% coarser)
Krumbein and Monk (1943)	$k = 760\, d^2 e^{-1.31\sigma}$		k = permeability (Darcys) d = geometric mean grain diameter, mm σ = log standard deviation of the size distribution
Kozeny-Carmen Bear (1972)	$K = (\rho_w g/\mu) n^3/(1-n)^2 (d_m{}^2/180)$		K = hydraulic conductivity ρ_w = fluid density μ = fluid viscosity g = gravitational constant d_m = representative grain size n = porosity
Shepard (1989)	$K = 40{,}000\, d^2$ $K = 1{,}600\, d^{1.75}$ $K = 450\, d^{1.65}$ $K = 100\, d^{1.5}$	glass spheres beach deposits channel deposits texturally immature consolidated sediments	K = hydraulic conductivity, ft/day d = mean grain size, mm

[a] Adapted from Refs. 1 and 4.

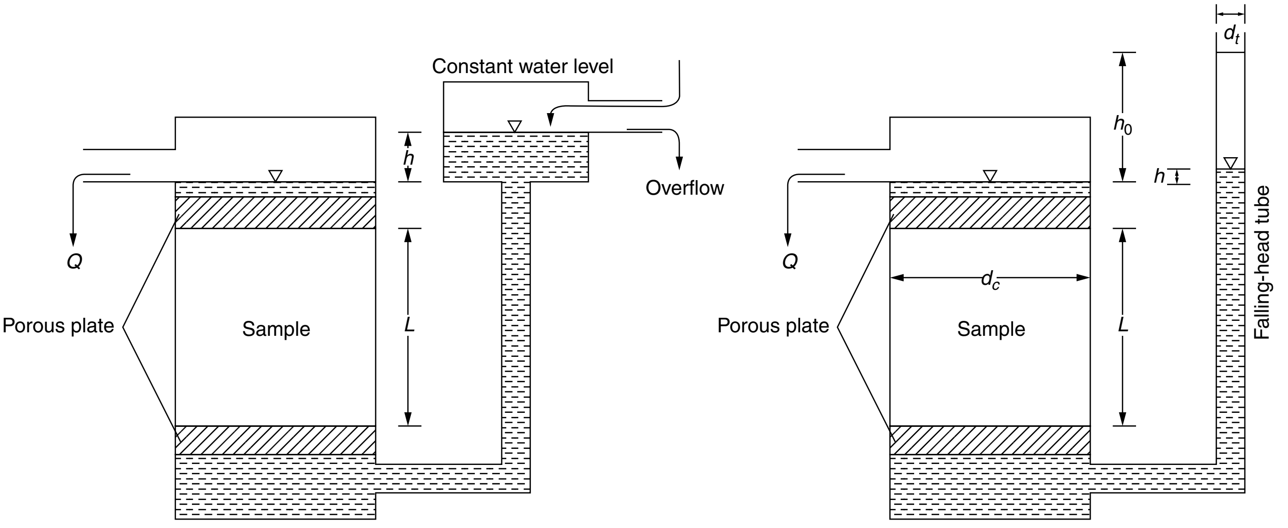

Constant-head permeameter apparatus

Falling-head permeameter apparatus

Figure 5. Permeameters (adapted from Ref. 4).

In general, pumping tests affect a significantly larger volume of the aquifer than slug tests, and the resulting hydraulic conductivity and transmissibility estimates are considered more representative of conditions across a broad area. Slug tests can be used to target specific zones and points within the aquifer but are generally considered "point estimates" of hydraulic conductivity and transmissibility.

BIBLIOGRAPHY

1. Domenico, P.A. and Schwartz, F.W. (1997). *Physical and Chemical Hydrogeology*. John Wiley & Sons, New York.
2. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
3. Sudicky, E.A. (1986). A natural gradient experiment on solute transport in a sand aquifer: spatial variability of hydraulic conductivity and its role in the dispersion process. *Water Resour. Res*. **22**: 2069–2082.
4. Fetter, C.W. (1994). *Applied Hydrogeology*. Macmillan College Publishing Company, Inc., New York.
5. Hubbert, M.K. (1940). The theory of groundwater motion. *The Journal of Geology* **48**: 785–944.

GROUNDWATER FLOW AND TRANSPORT PROCESS

Mohammad N. Almasri
An-Najah National University
Nablus, Palestine

Jagath J. Kaluarachchi
Utah State University
Logan, Utah

INTRODUCTION

Groundwater is an important source of water for agriculture and domestic drinking water. Two-thirds of the world's freshwater is found in underground aquifers. Groundwater occurs mainly in the tiny spaces between soil particles that are called voids or pores. In its natural state, groundwater is perpetually moving. This movement is governed by conventional hydraulic principles such as Darcy's law, the concept of hydraulic gradient, and the permeability of the medium (1). With the increasing water demand due to population growth, management of water resources is becoming a challenging task. Management of groundwater resources must consider both sustainable development and prevention and control of groundwater contamination. Groundwater contamination is a critical issue because of the importance of groundwater supply and quality in sustainable development. In general, contaminants found in groundwater are associated with adverse health, social, environmental, and economic impacts (2). Exposure to contaminated groundwater may occur accidentally because contaminated water might be colorless, odorless, and tasteless. It is essential to understand groundwater flow and transport processes for the best management of groundwater resources, appraisal of the sustainable yield of aquifers, reliable evaluation of the extent of groundwater contamination, assessment of the corresponding hazards and risks, and design of effective remediation systems and protective alternatives.

This chapter provides an overview of the main principles of groundwater flow and transport processes, including the basic concepts, definitions and terms, physical governing equations, and mathematical modeling.

VERTICAL DISTRIBUTION OF SUBSURFACE WATER

The subsurface occurrence of water may be divided into zones of aeration and saturation (see Fig. 1). The zone of aeration is partially occupied by water and partially by air. In the zone of saturation, all pores are filled with water. The zone of aeration may be further subdivided into the soil–water zone, the intermediate vadose zone, and the capillary zone (see Fig. 1).

The saturated zone extends from the upper surface of saturation down to underlying impermeable rock. In the absence of an overlying impermeable formation, the water table forms the upper surface of the zone of saturation, which is defined as the surface of atmospheric pressure. Water in the soil–water zone exists at less than saturation except temporarily when excessive water reaches the ground surface from rainfall or irrigation. The zone extends from the ground surface down through the major root zone. The water content and distribution within this zone depends on soil texture, vegetation, and atmospheric conditions. The intermediate vadose zone extends from the lower edge of the soil–water zone to the upper limit of the capillary zone. The capillary zone extends from the

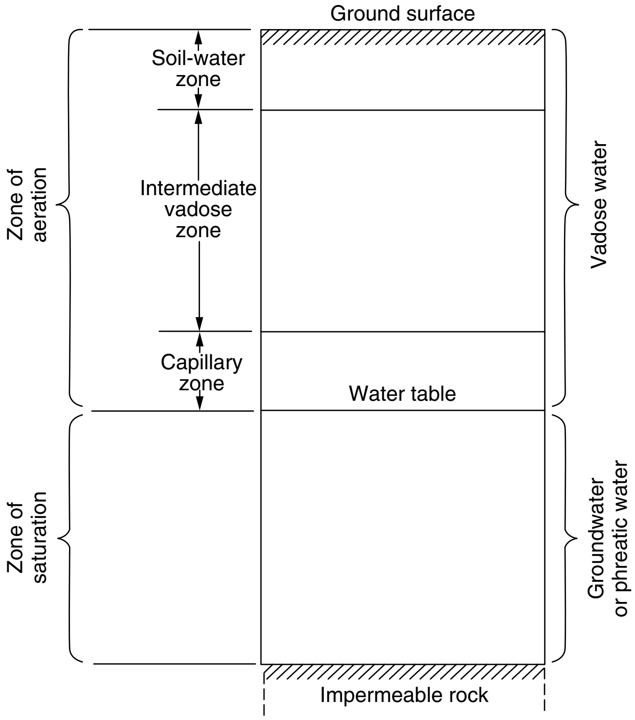

Figure 1. Illustration of saturated and unsaturated zones (1).

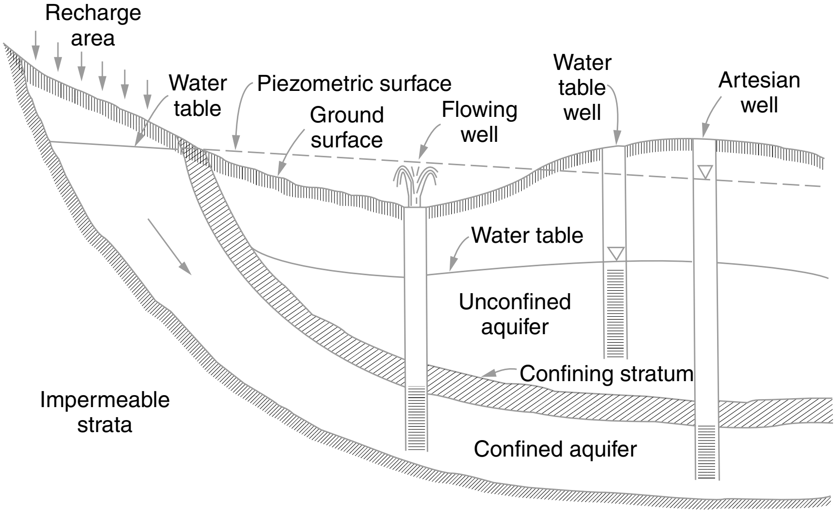

Figure 2. Schematic cross section of unconfined and confined aquifers (1).

water table up to the limit of capillary rise of water. The thickness of the capillary zone varies inversely with the pore size of a soil.

AQUIFER TYPES

An aquifer is a formation that contains sufficient saturated permeable material to yield significant quantities of water to wells and springs (3). Therefore, aquifers have the ability to store and transmit water. Aquifers might be bounded by confining beds that are relatively impermeable materials located adjacent to the aquifers. Confined aquifers occur where groundwater is confined under pressure greater than atmospheric by overlying relatively impermeable formations. In a well penetrating such an aquifer, the water level will rise above the bottom of the confining bed, as shown by the artesian and flowing wells of Fig. 2. Unconfined aquifers do not have confining beds above and the aquifer can be directly recharged by rainfall (see Fig. 2). An unconfined aquifer is one in which the water table varies in a rising and falling form.

DARCY'S LAW

In 1856, the French engineer, Henry Darcy, developed the principal equation in groundwater hydrology based on laboratory experiments he had carried out to analyze the flow of water through sand. Darcy's law can be written as

$$Q = -KA\frac{dh}{dl} \tag{1}$$

Q is the quantity of water per unit of time (L^3/T), A is the area perpendicular to flow (L^2), K is a constant of proportionality called the hydraulic conductivity (L/T), h is the hydraulic head (L), and dh/dl is the hydraulic gradient. The hydraulic conductivity expresses the ease by which water can pass through soil pores and openings and serves as a measure of the permeability of the porous medium. The hydraulic conductivity of a soil is a function of the properties of the porous medium and the fluid passing through it. In addition to hydraulic conductivity, the term transmissivity T (L^2T^{-1}) is widely employed in groundwater hydraulics. Transmissivity equals the hydraulic conductivity multiplied by the saturated thickness of the aquifer.

Darcy flux or specific discharge is given by the following relationship:

$$q = \frac{Q}{A} = -K\frac{dh}{dl} \tag{2}$$

The specific discharge q has the dimensions of velocity (L/T). Finally, the cross-sectional area of flow for a porous medium is actually much smaller than the dimensions of the aquifer. To find the velocity at which water is actually moving, the specific discharge is divided by the effective porosity to account for the actual open space available for flow. The result is the average velocity at which water moves between two points as in the following equation:

$$v = \frac{q}{n_e} = \frac{Q}{n_e A} = -\frac{Kdh}{n_e dl} \tag{3}$$

STORAGE IN CONFINED AND UNCONFINED AQUIFERS

Storage in confined aquifers is associated with the compressibility of water and the porous matrix. The change in the potentiometric head affects the arrangement of mineral grains as well as the density of water in the voids. If the water pressure increases, the mineral skeleton will expand; if it drops, the mineral skeleton will contract. Likewise, water will contract with an increase in pressure and expand if the pressure drops. The specific storage (S_s) is the amount of water per unit volume of a saturated formation that is stored or expelled from storage owing to compressibility of the mineral skeleton and the pore water per unit change in head. The value of specific storage is very small, generally 0.0001 ft^{-1} or less. In unconfined aquifers, pumpage causes drawdown in the water table itself, and the induced hydraulic gradient causes both a horizontal and vertical flow component. Thus, the drop in the water table creates a vertical hydraulic gradient which causes drainage of the porous medium (3–7).

The storativity (S) of a confined aquifer is the product of the specific storage S_s and the aquifer thickness. The value of the storativity of confined aquifers is of the order 0.005 or less. Storativity under unconfined conditions is referred to as the specific yield, defined as the ratio of the volume of water an unconfined aquifer drains by gravity due to the unit decline in the water table to the total volume of the aquifer. The coarser the material, the more closely the specific yield approaches the porosity. The volume of water drained from an aquifer as the head is lowered may be found from the formula,

$$V_w = S\,A\,\Delta h \tag{4}$$

where V_w is the volume of water drained or released (L^3), S is the storativity (L^0), A is the surface area overlying the drained aquifer (L^2), and Δh is the average decline in head (L). Note that Δh for a confined aquifer is the decline in the potentiometric head and for unconfined aquifers is the decline in the water table.

MATHEMATICAL MODEL OF GROUNDWATER FLOW

Groundwater flow is described using Darcy's law and the mass conservation equation. The governing equation, describing transient groundwater flow in a three-dimensional flow domain, is given as

$$\frac{\partial}{\partial x}\left(K_x\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y\frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z\frac{\partial h}{\partial z}\right)+W' = S_s\frac{\partial h}{\partial t} \tag{5}$$

where K_x, K_y, K_z are the hydraulic conductivity values in the x, y, and z coordinates, respectively; S_s is the specific storage; t is time; h is the hydraulic head; and W' is the source or sink strength in units of volume/time/area. To solve Eq. 5, initial and boundary conditions should be specified for the study area. Solving this equation yields the head distribution in time and space. Computer codes can be used to solve Eq. 5 numerically. MODFLOW developed by the U.S. Geological Survey (8), is one of the most commonly used groundwater softwares for solving one-, two-, or three-dimensional groundwater flow problems. MODFLOW is divided into packages that deal with single simulation aspects. For example, the Well Package simulates the effects of injection and abstraction wells and the River Package simulates the effect of rivers, etc. The WWW address for MODFLOW model and documentation is: http://water.usgs.gov/sofware/ground_water.htm. A brief description of MODFLOW can also be found in Charbeneau (7).

SOURCES OF GROUNDWATER CONTAMINATION

There are three key features that differentiate groundwater contamination sources: (1) degree of localization, (2) loading history, and (3) types of contaminants. Sources of groundwater contamination can be classified into point and nonpoint sources. Point sources include leaking underground pipelines or tanks, wastewater lagoons, septic systems, and leaking barrels of chemical wastes. These sources are identifiable, located at single locations, and have local impacts on groundwater quality. Conversely, nonpoint sources are diffuse, originating from many smaller sources, largely distributed, lack a well-defined single point of origin, and have a regional influence on groundwater quality. Examples of nonpoint sources are agricultural activities such as fertilizer and manure applications; and polluted atmospheric deposition, including dry fallout, precipitation, and runoff from roads and parking lots. The loading history defines how the concentration of a contaminant or its rate of application changed with time at the source location.

TRANSPORT PROCESSES

The physical processes that control contaminant flux into and out of a groundwater system are advection, mechanical dispersion, and molecular diffusion. In the following subsections, the basic concepts related to these physical processes are illustrated.

Advection

Transport by advection or convection entails movement and carrying contaminants with flowing groundwater (6). In advective transport, contaminants travel at a rate equal to the average linear velocity of the groundwater (5). Mathematically, advective flux is expressed by the following equation (6):

$$J = v_x n_e C \tag{6}$$

where J_x is the advective mass flux (ML^2T^{-1}) per unit area per unit time, v_x is the average linear velocity of groundwater(LT^{-1}), n_e is the effective porosity through which flow occurs, and C is the contaminant concentration (ML^{-3}). The average linear velocity can be determined from Eq. 3. The product of the average linear velocity and time since contaminants were introduced into the groundwater system defines the position of the advective front, as depicted in Fig. 3, where the vertical line represents the advective front without considering the effects of dispersion and diffusion (6).

Diffusion

The movement of contaminants from an area of high concentration to an area of lower concentration is called molecular diffusion or diffusion. This diffusive flux is

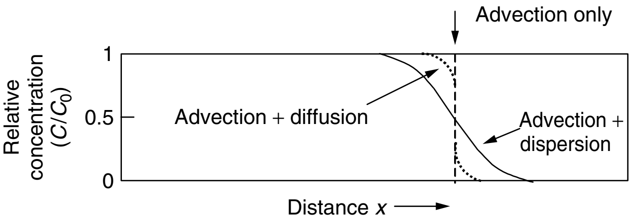

Figure 3. Advective transport and the influence of longitudinal dispersion and diffusion on the transport of a solute in one-dimensional flow.

directly proportional to the diffusion coefficient and the concentration gradient (9). This is known as Fick's first law and can be expressed in a one-dimensional system as in the following:

$$F_x = -D_{\mathrm{d}} \frac{dC}{dx} \tag{7}$$

where F_x is the diffusive mass flux ($\mathrm{ML^2T^{-1}}$), D_{d} is the diffusion coefficient for open water ($\mathrm{L^2T^{-1}}$), C is the contaminant concentration ($\mathrm{ML^{-3}}$); and dC/dx is the concentration gradient ($\mathrm{ML^{-4}}$). The negative sign signifies that the diffusive flux is in the direction of decreasing contaminant concentration.

Mechanical Dispersion

Dispersion describes the gradual spreading of contaminants beyond the region they are expected to occupy when considering advection alone. The phenomenon of dispersion is attributed mainly to the discrepancy between contaminant velocities and the average linear velocity caused by subsurface heterogeneity (5,10,11). According to Fetter (3), this discrepancy is attributed to three key factors: (1) the velocity of groundwater flow is higher in the center of pores than at the edges, (2) the contaminants can pass through different flow paths to reach the same location, and (3) groundwater moves faster in large pores. Therefore, contaminants will dilute due to mixing with noncontaminated water. Mixing, called mechanical dispersion, may occur along the flow path, is called longitudinal dispersion, and is attributed to the aforementioned key factors. The groundwater flow path may diverge at the pore scale, so the contaminant will spread in directions normal to the flow path. This mixing is called *lateral* or *transverse dispersion* (3). The difference between the straight dashed line and the solid line in Fig. 3 depicts the influence of longitudinal dispersion on the transport of a contaminant. Molecular diffusion and mechanical dispersion contribute jointly to the dispersive process of contaminant transport. They are combined to define a coefficient called *hydrodynamic dispersion* as given in the following equations (6):

$$D_{\mathrm{L}} = \alpha_{\mathrm{L}} v_{\mathrm{i}} + D^* \tag{8}$$

$$D_T = \alpha_T v_{\mathrm{i}} + D^* \tag{9}$$

where D_{L} and D_{T} are the longitudinal and transverse hydrodynamic dispersion coefficients, respectively; and α_{L} and α_{T} are the longitudinal and transverse dispersivities, respectively.

Figure 4 illustrates the process of hydrodynamic dispersion in two-dimensional groundwater flow at a constant uniform velocity, V. A tracer was instantaneously injected into the aquifer at $t = 0$. The groundwater flow carries the tracer with it. Due to the effect of dispersion, the solute spreads out such that the maximum concentration decreases with time, as shown for times t_1 and t_2 when $t_2 > t_1$. The spread of the tracer results in an elliptically shaped concentration distribution that is normally distributed in both the longitudinal and transverse directions.

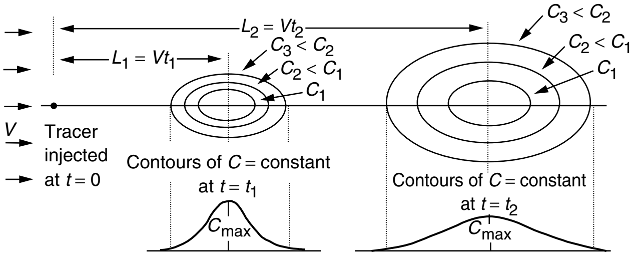

Figure 4. Longitudinal and transversal spreading of a tracer from an instantaneous point injection at $t = 0$.

As mentioned earlier, longitudinal dispersivity or more specifically, dispersivity, is a scale-dependent property that increases with the characteristic transport length distance of measurement. This phenomenon is attributed to the macroscale effects resulting from the heterogeneity of the subsurface on the field scale. Many researchers have developed relationships between longitudinal dispersivity and flow length (6,12–14). The ratio of longitudinal to transverse dispersivity ($\alpha_{\mathrm{L}}/\alpha_{\mathrm{T}}$) controls the shape of the contaminant plume. As the ratio decreases, the wider the plume will be (6). Bear (10) provided a range from 5 to 24 for $\alpha_{\mathrm{L}}/\alpha_{\mathrm{T}}$.

MATHEMATICAL MODEL OF CONTAMINANT TRANSPORT

The advection–dispersion–reaction partial differential equation that governs three-dimensional transport of a single chemical constituent in groundwater, considering advection, dispersion, fluid sinks/sources, equilibrium-controlled sorption, and first-order irreversible rate reactions, is given in the following (11):

$$R\frac{\partial C}{\partial t} = \frac{\partial}{\partial x_i}\left(D_{ij}\frac{\partial C}{\partial x_j}\right) - \frac{\partial}{\partial x_i}(v_i C) + \frac{q_{\mathrm{s}}}{n}C_{\mathrm{s}} - \lambda\left(C + \frac{\rho_{\mathrm{b}}}{n}S\right) \tag{10}$$

where the retardation factor is defined as

$$R = 1 + \frac{\rho_{\mathrm{b}}}{n}\frac{\partial S}{\partial C} \tag{11}$$

and C is the dissolved concentration ($\mathrm{ML^{-3}}$); S is the sorbed concentration ($\mathrm{MM^{-1}}$) which is a function of the dissolved concentration, C, as defined by a sorption isotherm; t is time (T); D_{ij} is the dispersion coefficient tensor ($\mathrm{L^2T^{-1}}$); v_i is the average linear groundwater velocity ($\mathrm{LT^{-1}}$); q_{s} is the volumetric flow rate per unit volume of aquifer and represents fluid sources and sinks ($\mathrm{T^{-1}}$); C_{s} is the concentration of the fluid source or sink flux ($\mathrm{ML^{-3}}$); λ is the reaction rate constant ($\mathrm{T^{-1}}$); R is the retardation factor ($\mathrm{L^0}$); ρ_{b} is the bulk density of the porous medium ($\mathrm{ML^{-3}}$); and n is the porosity ($\mathrm{L^0}$). A complete derivation of this equation is provided in Freeze and Cherry (4) and Zheng and Bennett (11).

Solving time-dependent partial differential equations governing contaminant transport requires identification of the initial and boundary conditions. The initial condition describes the distribution of the contaminant in the groundwater within the domain at the beginning of simulation. The boundary conditions can be classified into three categories: (1) specified concentration boundary, called the Dirichlet condition; (2) a specified concentration gradient boundary, called the Neuman condition; and (3) a variable flux boundary, called the Cauchy condition (5).

Solving the mathematical model yields the contaminant concentration as a function of space and time. The approach used in formulating and solving the mathematical model can be classified generally into numerical and analytical approaches. Analytical solutions of the advection–dispersion equations in one-, two-, and three-dimensional space can be found in Baetslé (14), Bear (10), Domenico (15), Domenico and Schwartz (5), and Fetter (6) among others.

Numerical methods provide an approximate solution of the mathematical model by solving a set of algebraic equations that replace the original partial differential equations. Numerical solutions are widely used and preferred because they have the capability to handle more general conditions. Numerical methods for solving the advection–dispersion–reaction equation divide the model domain into small cells, approximate the governing equation over each cell at time t, and then predict a new value at time $t + \Delta t$. This process continues forward in time for small Δt values. The most common numerical methods for solving the advection–dispersion–reaction equation include the finite-difference method, the finite-element method, and the random walk method. A detailed illustration of these methods can be found in Zheng and Bennett (11) and Sun (16). Bedient et al. (17) cited many references that provide detailed descriptions of these numerical methods. One of the common computer codes used in simulating the fate and transport of contaminants in groundwater is MT3D (18). MT3D is a three-dimensional contaminant transport model for the simulation of advection, dispersion, and chemical reactions of dissolved constituents in groundwater systems. MT3D is based on a modular structure to permit simulation of transport components independently or jointly with finite-difference groundwater flow models, such as MODFLOW. MT3D can be downloaded from http://hydro.geo.ua.edu. Further information on MT3D can be found in Zheng (18), Zheng and Bennett (11), and Zheng and Wang (19).

BIBLIOGRAPHY

1. Todd, D.K. (1980). *Groundwater Hydrology*. John Wiley & Sons, New York.
2. Office of Technology Assessment. (1984). Protecting the nation's groundwater from contamination, Office of Technology Assessment, OTA-0-233, Washington, DC.
3. Fetter, C.W. (1994). *Applied Hydrogeology*. Prentice-Hall, New York.
4. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, New York.
5. Domenico, P.A. and Schwartz, F.W. (1990). *Physical and Chemical Hydrology*. John Wiley & Sons, New York.
6. Fetter, C.W. (1999). *Contaminant Hydrogeology*. Prentice-Hall, New York.
7. Charbeneau, R.J. (2000). *Groundwater Hydraulics and Pollutant Transport*. Prentice-Hall, Upper Saddle River, NJ.
8. Harbaugh, A.W. and McDonald, M.G. (1996), User's documentation for MODFLOW-96, an update to the U.S. Geological Survey modular finite-difference ground-water flow model: U.S. Geological Survey Open-File Report 96–485.
9. Ingebritsen, S.E. and Sanford, W.E. (1998). *Groundwater in Geologic Processes*. Cambridge University Press, Cambridge, UK.
10. Bear, J. (1979). *Hydraulics of Groundwater*. McGraw-Hill, New York.
11. Zheng, C. and Bennett, G.D. (1995). *Applied Contaminant Transport Modeling: Theory and Practice*. International Thomson Publishing, Stamford, CT.
12. Neuman, S.P. (1990). Universal scaling of hydraulic conductivities and dispersivities in geologic media. *Water Resour. Res.* **26**(8): 1749–1758.
13. Xu, M. and Eckstein, Y. (1995). Use of weighted least-squares method in evaluation of the relationship between dispersivity and field scale. *Ground Water* **33**(6): 905–908.
14. Baestlé, L.H. (1999). Migration of radionuclides in porous media. *Progress in Nuclear Energy Series* XII, *Health Physics*. A.M.F. Duhamel (Ed.). Pergamon Press, Elmsford, NY, pp. 707–760.
15. Domenico, P.A. (1987). An analytical model for multidimensional transport of a decaying contaminant species. *J. Hydrolo.* **91**: 49–58.
16. Sun, N. (1996). *Mathematical Modeling of Ground Water Pollution*. Springer, New York.
17. Bedient, P.B., Rifai, H.S., and Newell, C.J. (1999). *Groundwater Contamination: Transport and Remediation*. Prentice-Hall, New York.
18. Zheng, C. (1990). MT3D, a modular three-dimensional transport model for simulation of advection, dispersion and chemical reactions of contaminants in ground water systems. Report to the US Environmental Protection Agency.
19. Zheng, C. and Wang, P.P. (1999). MT3DMS, a modular three-dimensional multi-species transport model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems; documentation and user's guide. U.S. Army Engineer Research and Development Center, Contract Report SERDP-99-1, Vicksburg, MS.

REACTIVE TRANSPORT IN THE SATURATED ZONE: CASE HISTORIES FOR PERMEABLE REACTIVE BARRIERS

LEN LI
CRAIG H. BENSON
University of
Wisconsin-Madison
Madison, Wisconsin

INTRODUCTION

The permeable reactive barrier (PRB) is an *in situ* remediation technology where contaminated groundwater is treated passively as it flows through a reactive medium. Most PRBs use granular zero valent iron (ZVI) to create redox conditions that result in degradation or immobilization of groundwater contaminants, such as chlorinated solvents and herbicides, heavy metals, and radionuclides (1–5). These redox conditions promote precipitation of secondary minerals from ions typically in groundwater as well as some groundwater contaminants (2,6,7). The most common minerals formed in PRBs are magnetite, hematite, goethite, lepidocrocite, calcite, aragonite, siderite, green rust, ferrous hydroxide, ferrous sulfide, and marcasite (8). Calcium carbonates and siderite typically are found near the entrance face of a PRB, whereas magnetite, ferrous hydroxide, green rust, and iron oxyhydroxides form throughout a PRB (6,7). Accumulation of secondary minerals causes loss of pore space and reactive surface area of the medium in PRBs, which can alter flow paths, residence times, and the treatment effectiveness of a PRB (2,7,9).

Developing a general assessment of mineral precipitation in PRBs with field and/or laboratory data is complicated by site-specific geochemical and hydrogeological conditions (7), and the relatively long period over which mineral deposition occurs (9). An alternative approach is to conduct analyses with a reactive transport simulator that predicts mineral deposition and its effects on hydraulics (10–12). Li (12) developed a reactive transport simulator specifically for this purpose. This chapter describes a comparison between ion concentrations predicted with this model and concentrations measured in the field at three PRBs.

REACTIVE TRANSPORT SIMULATOR

Equilibrium and Kinetic Reactive Transport Models

Equilibrium and kinetic models have simulated geochemical processes and mineral formation in PRBs (10,11,13,14). Equilibrium models can estimate the aqueous species and solid phases that exist in a state of thermodynamic equilibrium, but they do not account for the rate at which reactions occur and generally do not involve the movement of groundwater. In contrast, kinetic models describe the rate at which geochemical reactions occur and predict aqueous and solid phase concentrations as a function of time. In kinetic models, the geochemical reactions usually are specified by the user along with reaction rate expressions and the rate coefficients. Thus, when a kinetic model is developed, an equilibrium model often defines the most important reactions to be included in the kinetic model (e.g., see Yabusaki et al. (10)).

Equilibrium models are appropriate when the reaction rate is fast compared with the flow rate, whereas kinetic models are appropriate when flow rates are not negligible relative to reaction rates. For most PRBs, flow rates are fast enough such that reaction kinetics must be considered when evaluating mineral precipitation (10). Reactive transport codes incorporating kinetic geochemical algorithms have been developed by Mayer et al. (11) and Yabusaki et al. (10) for simulating contaminant degradation and mineral precipitation in PRBs. The code developed by Li et al. (8) is similar to those developed by Mayer et al. (11) and Yabusaki et al. (10), but it focuses solely on geochemical reactions that result in mineral precipitation.

Conceptual Model

The conceptual model used by Li et al. (8) consists of a continuous-wall PRB containing ZVI that is placed orthogonal to the groundwater flow. Groundwater flowing into the PRB is assumed to contain dissolved oxygen, Fe^{2+}, Ca^{2+}, Mg^{2+}, $SO_4{}^{2-}$, $CO_3{}^{2-}$, OH^-, $HCO_3{}^-$, and $NO_3{}^-$. These dissolved species are assumed to form $CaCO_3$, $CaMg(CO_3)_2$, $FeCO_3$, $Fe(OH)_2$ (am), FeS (am), and $Mg(OH)_2$ in response to the redox conditions induced by corrosion of the ZVI. In an actual PRB, additional ions may be present and minerals other than those included in the model may also form. However, the reactions in the model include the most common minerals observed in laboratory and field studies of PRBs (6,7,10,11). Contaminant transformation reactions were not included in the model because the volume of mineral precipitates formed as a consequence of contaminant treatment generally is negligible, even though the contaminant transformations are intrinsic to the remediation process (2,11).

Codes

The reactive transport simulator is created with the public domain codes MODFLOW and RT3D (15,16). MODFLOW simulated steady-state flow in the aquifer and PRB. Geochemical processes within the PRB were simulated with RT3D, a three-dimensional multicomponent model that describes advective–dispersive-reactive transport (16). Heads from MODLOW are input to RT3D. A detailed description of the codes that comprise the simulator can be found in Li (12) and Li et al. (8).

Li et al. (8) developed a geochemical module for RT3D that simulated the mineral-forming reactions. Geochemical reactions included in the module are summarized in Table 1. Three redox reactions are related to iron corrosion, along with six mineral precipitation reactions and one reaction for microbial sulfate reduction. A pseudo-first-order reaction rate was assumed for iron corrosion by oxygen or nitrate with the reaction rate proportional to the iron surface area and the concentration of oxygen or nitrate (11). For iron corrosion by water, the reaction rate was assumed to be proportional only to the iron surface area (10,11,17). The rate of degradation of sulfate to HS^- by sulfate-reducing bacteria was assumed to follow the Monod equation (11), and the kinetics of mineral precipitation were assumed to follow transition state theory (18). The kinetic rate expressions and solubility constants for the mineral reactions are summarized in Table 1. All geochemical reactions are assumed to occur in parallel and are solved simultaneously.

FIELD DATA SOURCES AND MODEL SETUP

Field data from three PRBs verified that the MODFLOW-RT3D simulator produces reasonable predictions of major ion concentrations. Comparisons were not made between measured and predicted solid phase concentrations at any of the sites because of a lack of field data. The selected PRBs are located at Moffett Federal Airfield in

Table 1. Geochemical Reactions and Solubility Constants

Reaction	Mineral Formed	Kinetic Rate Expression[a]	Solubility Constant Log(K_{eq})[b]
Iron Corrosion:			
$Fe^0 + H_2O + 0.5O_2(aq) \rightarrow Fe^{2+} + 2OH^-$	—	$k_{an,Fe}\ Sa_{DO}$	—
$Fe^0 + 2H_2O \rightarrow Fe^{2+} + H_2(aq) + 2OH^-$	—	$k_{an,Fe}\ S$	—
$4Fe^0 + 7H_2O + NO_3^- \rightarrow 4Fe^{2+} + 10OH^- + NH_4^+$	—	$k_{NO3}Sa_{NO_3^-}$	—
Microbial Sulfate-Reduction:			
$SO_4^{2-} + 4H_2(aq) \rightarrow HS^- + OH^- + 3H_2O$	—	$k\left(\frac{[SO_4^{2-}]}{K_{SO_4^{2-}} + [SO_4^{2-}]}\right)\left(\frac{[H_2]}{K_{H_2} + [H_2]}\right)$	—
Secondary Mineral Precipitation-Dissolution:			
$CaCO_3(s) \leftrightarrow Ca^{2+} + CO_3^{2-}$	Aragonite/Calcite	$-k_{Ca}(1 - a_{Ca^{2+}}a_{CO_3^{2-}}/K_{eq})$	−8.1
$CaMg(CO_3)_2(s) \leftrightarrow Ca^{2+} + Mg^{2+} + 2CO_3^{2-}$	Dolomite	$-k_{CaMg}(1 - a_{Ca^{2+}}a_{Mg^{2+}}a^2_{CO_3^{2-}}/K_{eq})$	−17.7
$Mg(OH)_2(s) \leftrightarrow Mg^{2+} + 2OH^-$	Brucite	$-k_{Mg2}(1 - a_{Mg^{2+}}a^2_{OH^-}/K_{eq})$	−11.2
$FeCO_3(s) \leftrightarrow Fe^{2+} + CO_3^{2-}$	Siderite	$-k_{Fe1}(1 - a_{Fe^{2+}}a_{CO_3^{2-}}/K_{eq})$	−10.5
$Fe(OH)_2(am) \leftrightarrow Fe^{2+} + 2OH^-$	Ferrous hydroxide	$-k_{Fe}(1 - a_{Fe^{2+}}a^2_{OH^-}/K_{eq})$	−15.2
$FeS(am) + H_2O \leftrightarrow Fe^{2+} + HS^- + OH^-$	Ferrous sulfide	$-k_{FeS}(1 - a_{Fe^{2+}}a_{HS^-}a_{OH^-}/K_{eq})$	−18.4

[a] *k* is rate coefficient, S is reactive surface area of ZVI, a is activity of aqueous species, and K_{eq} is solubility constant. Activity coefficients were calculated with the extended Debye—Huckel equation.

[b] At 15 °C and 101 kPa.

Mountain View, CA; Monticello, UT; and Copenhagen, Denmark. These PRBs were selected because hydrological and geochemical data were sufficiently well defined for each PRB to permit an assessment of the simulator.

For all sites, the simulations were set up assuming homogeneous hydraulic conductivity in each medium (i.e., aquifer, reactive zone, etc.). Constant-head boundaries were applied along the upstream and downstream faces of the domain to establish a hydraulic gradient orthogonal to the PRB. No flow conditions were imposed at the other boundaries. Each species in the PRBs was assumed to have an initial concentration of zero. A constant concentration boundary was applied at the upstream face of the PRBs, and other boundaries were assigned the Cauchy boundary condition with no dispersive flux. Influent groundwater concentrations entering the PRB were set at values measured in the field.

Moffett Federal Airfield

A pilot-scale funnel-and-gate PRB consisting of ZVI was installed in 1996 at Moffett Federal Airfield to remove TCE and DCE from groundwater (10). The gate is 3.0 m long (orthogonal to flow direction), 3.0 m wide (along the flow direction), and 5.5 m deep. The gate consists of a 1.8-m-thick barrier of granular ZVI sandwiched between two 0.6-m-thick layers of pea gravel (up gradient and down gradient). Groundwater samples collected within the PRB between April 1996 (installation) and October 1997 were analyzed for target contaminants, pH, major cations and anions, and indicator parameters (10). Analysis of core samples showed that magnetite, hematite, aragonite, and marcasite formed in the iron media. No direct measurements were made of siderite, FeS, $Fe(OH)_2$ (am), $Fe(OH)_3$, brucite, or green rust.

The MODFLOW-RT3D simulator simulated the nearly one-dimensional conditions in the gate at the Moffett Federal Airfield site. The input was based on data provided in Yabusaki et al. (10). MODFLOW simulated steady-state flow in the pea gravel zones and iron media zone. The initial hydraulic conductivity in the PRB was 186 m/d, and the porosity was 0.66. The hydraulic conductivity in the pea gravel zones was 86.4 m/d. A hydraulic gradient of 0.0003 was applied across the PRB to yield the Darcy velocity (0.04 m/d) reported in Yabusaki et al. (10). Five minerals reported by Yabusaki et al. (10) were included in the model: $CaCO_3$, $FeCO_3$, $Fe(OH)_2$ (am), $Mg(OH)_2$, and FeS (am). Composition of groundwater entering the PRB was based on information provided in Yabusaki et al. (10) (Table 2).

Monticello

A full-scale funnel and gate PRB containing granular iron was installed in 1999 to treat uranium in groundwater at a former uranium mill site near Monticello, UT (14). The gate is 30.0 m long, 2.4 m wide (along flow direction), and 4.5~6.0 m deep. A 1.2-m-thick reactive zone consisting of 100% granular ZVI was sandwiched between two 0.6-m-thick layers of pea gravel (up gradient and down gradient) in the gate. Samples of groundwater were collected from June 1999 (installation) to September 1999 and analyzed for target contaminants, pH, major cations and anions, and indicator parameters. Core samples showed that magnetite and $CaCO_3$ formed within the PRB.

Reactive transport in the gate at the Monticello PRB was simulated with the MODFLOW-RT3D simulator using a one-dimensional model similar to that employed for the PRB at Moffett Federal Airfield (see previous section). A one-dimensional model was selected because flow through the pea gravel zones and ZVI was nearly one-dimensional and steady state. Hydraulic conductivities of the ZVI and pea gravel zones were set at 30.9 m/d and 8.64 m/d. Based on data in Morrison et al. (14), the porosity was set at 0.50 in the ZVI and 0.30 in the pea gravel and the hydraulic gradient was set at 0.142 to yield a

Table 2. Concentrations of Aqueous Species in Influent Ground Water

Aqueous Species in Influent Ground Water	Aqueous Species Concentrations (M)		
	Moffett Federal Airfield, Mountain View, California[a]	Former Mill Site, Monticello, Utah[b]	Copenhagen Freight Yard, Copenhagen, Denmark[c]
Fe^{2+}	1.0×10^{-10}	1.0×10^{-10}	1.0×10^{-10}
Ca^{2+}	3.9×10^{-3}	8.5×10^{-3}	3.3×10^{-3}
Mg^{2+}	2.7×10^{-3}	3.6×10^{-3}	1.2×10^{-3}
OH^-	1.3×10^{-7}	5.2×10^{-8}	5.0×10^{-7}
HCO_3^-	4.7×10^{-3}	2.5×10^{-3}	1.0×10^{-2}
${CO_3}^{2-}$	1.4×10^{-6}	1.0×10^{-6}	1.0×10^{-6}
O_2(aq)	1.0×10^{-10}	1.0×10^{-10}	1.0×10^{-10}
${NO_3}^-$	3.9×10^{-5}	1.0×10^{-10}	1.0×10^{-10}
${SO_4}^{2-}$	3.6×10^{-3}	1.2×10^{-2}	1.2×10^{-3}

[a] From Yabusaki et al. (10).
[b] From Morrison et al. (14).
[c] From Killerich et al. (19).

groundwater velocity of 3 m/d. Five minerals were included in the geochemical model: $CaCO_3$, $FeCO_3$, $Fe(OH)_2$ (am), $CaMg(CO_3)_2$, and FeS. Composition of the groundwater entering the PRB was defined based on data reported in Morrison et al. (14), as summarized in Table 2.

Copenhagen

A full-scale PRB containing granular ZVI was installed in 1998 at the Copenhagen Freight Yard in Copenhagen, Denmark to treat chlorinated solvents in ground water (19). The PRB is a continuous trench 15.0 m long, 1.3 m wide (along flow direction), and 6.0 m deep and consists of a 0.9-m-thick zone of 100% granular ZVI sandwiched between 0.2-m-thick layers of pea gravel. Samples of groundwater were collected from July 1998 (installation) to September 1999 and were analyzed for target contaminants, pH, and major cations and anions (19). No direct measurements were made of minerals forming within the PRB (19).

The MODFLOW-RT3D simulator was set up to simulate one-dimensional flow and reactive transport along the centerline of the Copenhagen PRB. The model predicted steady-state flow in the surrounding aquifer, pea gravel zones, and ZVI. Hydraulic conductivity of the ZVI, pea gravel, and aquifer were set at 6.1, 2.1, and 0.6 m/d, respectively, based on site conditions (19). The porosity was set at 0.60 for the ZVI, 0.40 for the pea gravel zones, and 0.3 in the aquifer (19). A hydraulic gradient of 0.011 was applied across the aquifer based on data reported in Killerich et al. (19). Five minerals were included in the model: $CaCO_3$, $FeCO_3$, $Fe(OH)_2$(am), $CaMg(CO_3)_2$, and FeS. Composition of the groundwater entering the PRB was defined based on data reported in Killerich et al. (19), as summarized in Table 2.

COMPARISON OF PREDICTED AND MEASURED CONCENTRATIONS

Moffett Federal Airfield

Comparisons of the predicted and measured pH and concentration profiles are shown in Fig. 1 for the PRB at Moffett Field after 12 months of operation. The reaction rate coefficients initially were estimated from data in the literature, but ultimately they were adjusted by trial-and-error to reproduce the measured pH, alkalinity, and concentrations of ${SO_4}^{2-}$, Ca^{2+}, and Mg^{2+}. A similar calibration procedure was used at the other two sites. The calibrated reaction rate coefficients are summarized in Table 3 along with typical ranges based on data reported in the literature.

The predicted pH is in general agreement with the observed pH [Fig. 1(a)], including the rapid jump near the entrance face and the leveling out in the interior. The drop in alkalinity just up gradient of the PRB is not captured [Fig. 1(b)], but the alkalinity predicted by RT3D inside the PRB is similar to that measured in the field. The sulfate concentration predicted by RT3D also is in general agreement with the trends observed in the field [Fig. 1(b)], as are the Ca^{2+} and Mg^{2+} concentrations [Fig. 1(c)]. However, the predicted Ca^{2+} concentrations in the PRB are near the upper bound of the measured concentrations [Fig. 1(c)]. The field data are also too sparse to evaluate whether the simulator accurately predicts the rapid changes in geochemical conditions occurring at the entrance face.

Monticello Mill Site

Comparisons of measured and predicted pH, alkalinity, and concentrations of Ca^{2+}, Mg^{2+}, and ${SO_4}^{2-}$ for conditions at 3 months of operation are shown in Fig. 2. Calibrated rate coefficients for the PRB at Monticello are summarized in Table 3. As with the PRB at Moffett Field, insufficient data exist to determine whether the rapid changes predicted at the entrance face accurately replicate the conditions occurring in the PRB at Monticello.

The measured and predicted pH, ${HCO_3}^-$, Mg^{2+}, and ${SO_4}^{2-}$ concentrations in the Monticello PRB are in close agreement (Fig. 2). However, the predicted Ca^{2+} concentrations in the PRB are higher than are the measured concentrations [Fig. 2(b)]. One reason for this discrepancy may be that minerals other than $CaCO_3$ and $CaMg(CO_3)_2$ are precipitating (these minerals only account for 15%

Table 3. Calibrated Rate Coefficients in PRB Case Studies

Geochemical Reactions Terms	Units	Moffett Federal Airfield[c]	Former Mill Site[c]	Copenhagen Freight Yard[c]	Rate Coefficients Reported in Literature[d]
Oxygen Corrosion[a]	m^3/m^2-d	2.8×10^{-2}	2.8×10^{-2}	2.8×10^{-2}	$< 2.8 \times 10^{4}$
Water Corrosion[a]	mole/m^2-d	7.0×10^{-7}	1.5×10^{-5}	1.9×10^{-7}	$3.0 \times 10^{-8} - 5.4 \times 10^{-3}$
Nitrate Corrosion[a]	m^3/m^2-d	2.6×10^{-5}	—	2.8×10^{-5}	$1.8 \times 10^{-8} - 2.8 \times 10^{-5}$
Microbial sulfate reduction[b]	M/d	1.0×10^{-3}	1.0×10^{-3}	4.0×10^{-3}	$5.0 \times 10^{-6} - 5.0 \times 10^{-3}$
$CaCO_3$	M/d	1.0×10^{-4}	4.0×10^{-4}	1.0×10^{-5}	$2.7 \times 10^{-9} - 1.4 \times 10^{-4}$
$FeCO_3$	M/d	1.1×10^{-5}	2.1×10^{-5}	1.0×10^{-5}	$1.1 \times 10^{-5} - 2.7 \times 10^{-4}$
$Fe(OH)_2$ (am)	M/d	5.1×10^{-5}	5.1×10^{-5}	5.0×10^{-5}	$<2.2 \times 10^{-4}$
FeS (am)	M/d	1.0×10^{-6}	1.0×10^{-6}	1.0×10^{-6}	$1.1 \times 10^{-8} - 2.2 \times 10^{-5}$
$CaMg(CO_3)_2$	M/d	—	1.0×10^{-6}	1.0×10^{-9}	6.9×10^{-6}
$Mg(OH)_2$	M/d	2.0×10^{-3}	NA	NA	3.4×10^{-6}

[a] Reactive surface area of ZVI set at 3.9×10^6 m^2/m^3.

[b] Half-saturation constants for $SO_4{}^{2-}$ and H_2 (aq) were $K_{SO4} = 1.0 \times 10^{-4}$ M, and $K_{H2} = 1.0 \times 10^{-7}$ M, respectively.

[c] After calibration with field measurements.

[d] Li et al. (8).

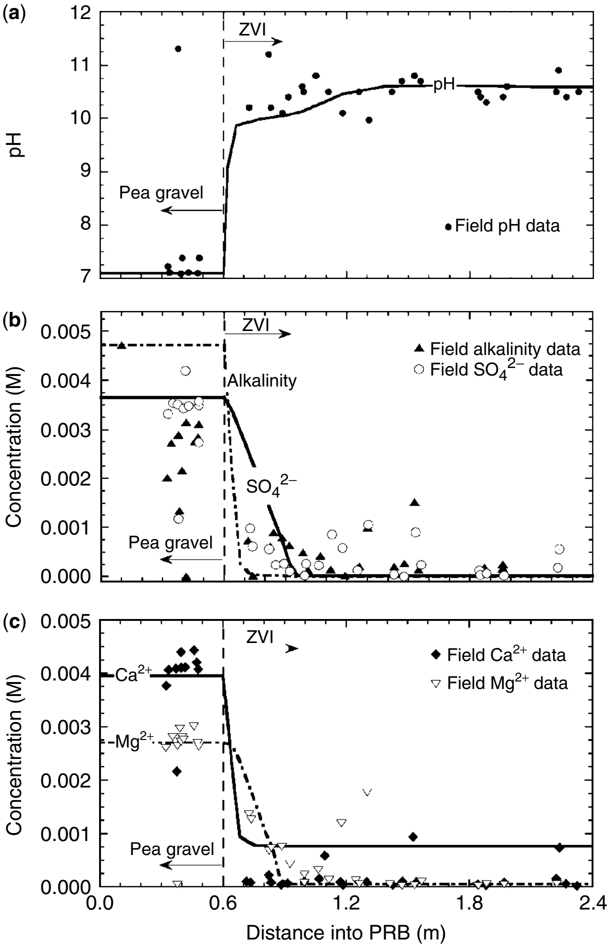

Figure 1. Measured and predicted concentrations for the PRB at Moffett Federal Airfield, Mountain View, CA, after 12 months of operation: (**a**) pH, (**b**) alkalinity and $SO_4{}^{2-}$, and (**c**) Ca^{2+} and Mg^{2+}.

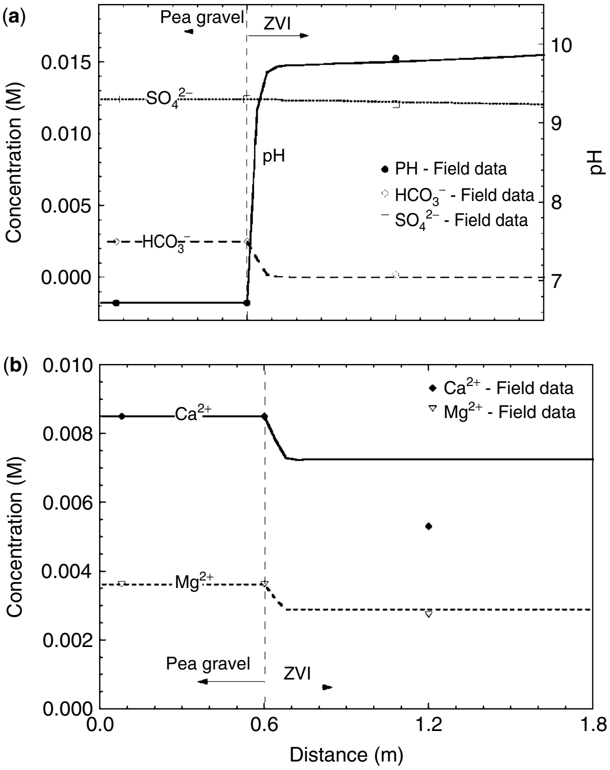

Figure 2. Measured and predicted concentrations for the PRB at the former Uranium Mill Site in Monticello, UT, after 3 months of operation: (**a**) pH, $HCO_3{}^-$, and $SO_4{}^{2-}$ and (**b**) Ca^{2+} and Mg^{2+}.

of aqueous calcium). For example, Morrison et al. (14) included iron-rich calcite ($Fe_{0.4}Ca_{0.6}CO_3$) and gypsum ($CaSO_4$) in their equilibrium model to account for calcium removal in the ZVI. However, neither of these minerals has been observed in the core samples from the PRB at Monticello. The surface complexation of calcium on the ZVI and secondary minerals surfaces, neither of which are considered in the simulator, may also be removing calcium from the solution (10).

Copenhagen Site

Comparisons of measured and predicted pH and Ca^{2+}, Mg^{2+}, $SO_4{}^{2-}$, and $HCO_3{}^-$ concentrations after 15 month of operation of the Copenhagen PRB are shown in Fig. 3. The calibrated rate coefficients corresponding to these predictions are summarized in Table 3.

Close agreement exists between the measured and predicted pH and the HCO_3^-, SO_4^{2-}, Ca^{2+}, and Mg^{2+} concentrations in the ZVI of the Copenhagen PRB [Figs. 3(a) and (b)], but the field data are too sparse to evaluate whether the trends in geochemistry predicted by the model are accurate. In contrast to the PRBs at Moffett Field and Monticello, the predicted pH climbs gradually at the Copenhagen site. This more graduate change in pH may be caused by a lower rate of iron corrosion by water (1.9×10^{-7} mole/m^2-d, Table 3), which results in less OH^- released into solution.

Comparison of Calibrated Rate Coefficients and Literature Values

The calibrated rates coefficients for the three PRB case studies are summarized in Table 3. Most of the calibrated rate coefficients fall within the range of rate coefficients from the literature (Table 3). The exceptions are the rate coefficient for $Mg(OH)_2$ at the Moffett PRB and the rate coefficient for $CaCO_3$ at the Monticello PRB. Both rate coefficients are higher than are the maximum values reported in the literature. However, the rate coefficient for $CaCO_3$ is only slightly above the upper bound from the literature, and only one rate coefficient was available in the literature for $Mg(OH)_2$. Thus, these exceptions are not considered unusual.

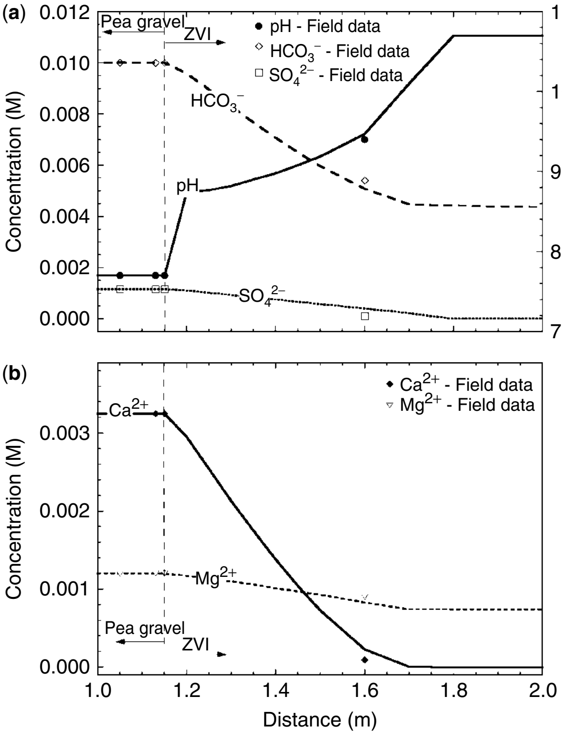

Figure 3. Measured and predicted concentrations at the centerline of the PRB at the Copenhagen Freight Yard, Copenhagen, Denmark, after 15 months of operation: (**a**) pH, HCO_3^-, and SO_4^{2-} and (**b**) Ca^{2+} and Mg^{2+}.

Although good agreement between measured and predicted concentrations has been shown in Figs. 1–3, the accuracy of the simulator when applied to other problems will depend greatly on the rate coefficients used as input. For example, the effect of varying the rate coefficient for iron corrosion by water (Table 1) within a reasonable range (1.0×10^{-7} to 1.0×10^{-5} mole/m^2-d) on pH is illustrated in Fig. 4 with the model for the PRB at Moffett Field. The influent concentrations used for the simulation are shown in Table 2, and the rate coefficients for the reactions other than iron corrosion by water are shown in Table 3. When the rate coefficient for iron corrosion by water is 10^{-7} mole/m^2-d, the predicted pH is much lower than the pH observed in the field, especially near the entrance. In contrast, when the rate coefficient for iron corrosion by water is 10^{-5} mole/m^2-d, the predicted pH increases too rapidly near the entrance and levels out approximately one pH unit higher than observed in the field.

In situ rate coefficients can vary considerably and depend on a variety of factors, and no method currently is available to predict rate coefficients accurately *a priori*. Thus, at the current time, simulators must be calibrated with field data if reliable predictions of mineral precipitation are to be obtained. Alternatively, parametric analyses can be conducted to evaluate the potential range of conditions that may be realized in the field.

SUMMARY

Comparisons between predicted and measured concentrations of common groundwater ions in PRBs have been described in this article. The simulator used for the predictions is based on the public domain codes MODFLOW and RT3D, the latter being augmented with a geochemical algorithm specifically for simulating geochemical reactions in PRBs. Field data from three PRBs were used for the comparison.

The comparisons showed that pH profiles and anion and cation concentrations in PRBs can be predicted accurately. The ability to predict solid phase concentrations has not been evaluated because of lack of field data. However,

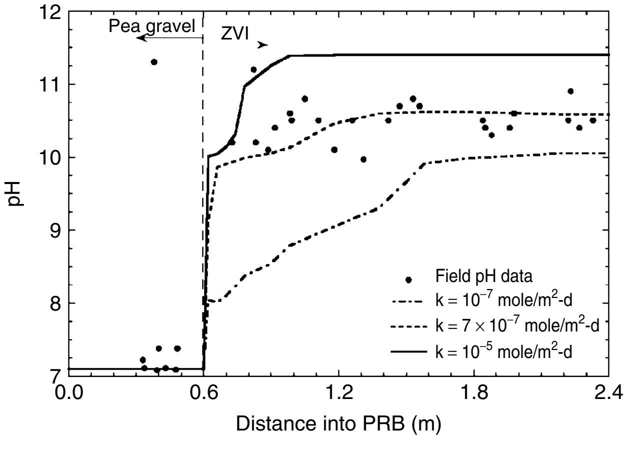

Figure 4. Measured and predicted pH for the PRB at Moffett Federal Airfield, Mountain View, CA, after 12 months of operation using various rate coefficients (k) for iron corrosion by water.

the accuracy of the predictions depends to a great extent on the reaction rate coefficients used as input. In this study, the rate coefficients were obtained by calibration because the actual *in situ* rate coefficients were unknown. Thus, predictions made *a priori* are prone to considerable uncertainty until methods are developed to accurately define reaction rate coefficients.

In many cases, the simulator predicted that the greatest changes in geochemistry generally occur at the entrance face of the PRB. However, the current field data sets contain insufficient detail to confirm that these changes in geochemistry are being predicted accurately. More detailed field observations are needed at the entrance face of the PRBs to more fully evaluate whether the simulator can accurately capture these rapid changes in geochemistry.

BIBLIOGRAPHY

1. Blowes, D.W. et al. (2000). Treatment of inorganic contaminants using permeable reactive barriers. *J. Contaminant Hydrol.* **45**: 123–137.
2. Wilkin, R.T., Puls, R., and Sewell, G.W. (2003). Long-term performance of permeable reactive barriers using zero-valent iron: geochemical and microbiological effects. *Ground Water* **41**: 493–503.
3. EPA. (1998). *Permeable Reactive Barrier Technologies for Contaminant Remediation*. United States Environmental Protection Agency, Washington, DC.
4. Lee, T. and Benson, C.H. (2004). Sorption and degradation of alachlor and metolachlor in ground water using green sands. *J. Environ. Quality* **33**: 1682–1693.
5. Lee, T., Benson, C.H., and Eykholt, G.R. (2004). Waste green sands as reactive media for groundwater contaminated with trichloroethylene (TCE). *J. Hazardous Materials* **B109**: 25–36.
6. Mackenzie, P.D., Horney, D.P., and Sivavec, T.M. (1999). Mineral precipitation and porosity losses in granular iron columns. *J. Hazardous Materials* **68**: 1–17.
7. Phillips, D.H. et al. (2003). Mineralogical characteristics and transformations during long-term operation of a zerovalent iron reactive barrier. *J. Environ. Quality* **32**: 2033–2045.
8. Li, L., Benson, C.H., and Mergener, E.A. (2004). Impact of mineral fouling on hydraulic behavior of permeable reactive barriers. *Ground Water*. In press.
9. Vikesland, P.J. et al. (2003). Longevity of granular iron in groundwater treatment processes: changes in solute transport properties over time. *J. Contaminant Hydrol.* **64**: 3–33.
10. Yabusaki, S. et al. (2001). Multicomponent reactive transport in an *in situ* zero-valent iron cell. *Environ. Sci. Technol.* **35**: 1493–1503.
11. Mayer, K.U., Blowes, D.W., and Frind, E.O. (2001). Reactive transport modeling of an *in situ* reactive barrier for the treatment of hexavalent chromium and trichloroethylene in groundwater. *Water Resour. Res.* **37**: 3091–3103.
12. Li, L. (2004). Fouling and the Long-Term Performance of Permeable Reactive Barriers. Ph.D. Dissertation, University of Wisconsin-Madison, Madison, WI.
13. Liang, L. et al. (2002). *Integration of Geochemical Data and Hydraulic Changes at Permeable Reactive Barriers Sites*. Oak Ridge National Laboratory, Oak Ridge, TN.
14. Morrison, S.J., Metzler, D.R., and Carpenter, C.E. (2001). Uranium precipitation in a permeable reactive barrier by progressive irreversible dissolution of zerovalent iron. *Environ. Sci. Technol.* **35**: 385–390.
15. McDonald, M.G. and Harbaugh, A.W. (1988). *MODFLOW: A Modular Three-Dimensional Finite-Difference Ground-Water Flow Model*. U.S. Geological Survey Techniques of Water-Resources Investigations, Washington, DC, book 6, chap. A1.
16. Clement, T.P. (1997). *RT3D: A Modular Computer Code for Simulating Reactive Multi-Species Transport in 3-Dimensional Groundwater Aquifers, PNNL-SA-11720*. Pacific Northwest National Laboratory, Richland, WA.
17. Reardon, E.J. (1995). Anaerobic corrosion of granular iron: Measurement and interpretation of hydrogen evolution rates. *Environ. Sci. Technol.* **29**: 2936–2945.
18. Lasaga, A.C. (1998). *Kinetic Theory in the Earth Sciences*. Princeton series in geochemistry. Princeton University Press, Princeton, NJ.
19. Killerich, O., Larsen, J.W., and Nielsen, C. (2000). Field results from the use of a permeable reactive wall. *Proceedings of the Second International Conference on Remediation of Chlorinated and Recalcitrant Compounds*, Monterey, CA.

TRANSPORT OF REACTIVE SOLUTE IN SOIL AND GROUNDWATER

Natalie L. Capiro
Philip B. Bedient
Rice University
Houston, Texas

INTRODUCTION

Remediation of contaminated aquifers, both vadose zone (unsaturated portion of an aquifer) soils and groundwater, has proven to be a challenging environmental endeavor. Groundwater is the source of about 60% of the United States' water usage and, with contaminant plumes impacting many states, associated human health concerns of great importance exist. Groundwater can become polluted both directly by leaking underground storage tanks and indirectly by leakage of contaminants on the soil surface that leach to the groundwater. Common pollutants include agricultural chemicals applied to crops, industrial spills, landfill leachate, disposal sites, and waste ponds. Subsurface remediation is complicated by not only the behavior of the contaminants themselves but by the heterogeneities found in typical aquifer formations. The processes that govern the fate and transport of solute in aquifer systems generally fall into the categories of hydrodynamic, abiotic, and biotic.

The universal process that impacts all forms of solute include the hydrodynamic processes, which affect contaminant transport by impacting the flow paths of groundwater in the subsurface, including both the vadose zone and the saturated aquifer. Abiotic processes affect contaminant transport physically by causing interactions between the contaminant and the stationary subsurface materials (i.e., sorption) or by affecting the form of the contaminant (i.e., through hydrolysis and redox reactions). Biotic processes are driven by microbial activities and can affect contaminant transport by metabolizing or mineralizing the contaminant (i.e., organic contaminants), or possibly by using the contaminant in the metabolic process (i.e., nitrate under denitrifying conditions).

PHYSICAL MECHANISMS OF SOLUTE TRANSPORT

The three basic hydrodynamic processes affecting contaminant transport in saturated and unsaturated porous media are advection, diffusion, and hydrodynamic dispersion. Advection represents the movement of a contaminant with the flowing groundwater according to the seepage velocity in the pore space. Advection is the most important process driving the down gradient migration of dissolved contaminants in the saturated subsurface (1). If advection was the only process acting on a tracer, a sharp interface would exist continuously as the tracer flushes out the solute-free water. In essence, the tracer would behave as a plug. A breakthrough curve, illustrating the relative effluent concentration (C/Co) vs. time in a continuous tracer input, would appear as a step function (Fig. 1a). The average velocity of advective transport is described by the amended Darcy equation:

$$v_x = -\frac{K}{n}\frac{dh}{dL} \tag{1}$$

where v_{x} is the average linear groundwater velocity [L/t], K is the hydraulic conductivity (a measure of a porous media's ability to transmit water) [L/t], n is the effective porosity [unit less], and dh/dL is the hydraulic gradient [L/L].

The average linear velocity or seepage velocity, v_x, equals the Darcy velocity divided by effective porosity, n, associated with the pore space through which water can actually flow. The average linear velocity is less than the microscopic velocities of water molecules moving along individual flowpaths, because of tortuosity. The one-dimensional (1-D) mass flux (F_x) because of advection equals the product of water flow and concentration (C) of solute, or $F_x = v_x nC$.

Darcy's law is also still valid for the macroscopic view (through the entire pore body) of unsaturated flow, except that hydraulic conductivity is now a function of moisture content. Darcy's law is then used with the unsaturated value for K and can be written as

$$v = -K(\theta)dh/dz \tag{2}$$

where v is Darcy velocity, z is depth below surface [L], $K(\theta)$ is unsaturated hydraulic conductivity [L/T], θ is volumetric moisture content, h is potential or head ($h = z + \psi$) [L/T], and ψ is tension or suction [L]. For saturated soils, the hydraulic conductivity was assumed to be constant; however, in unsaturated soil, it decreases sharply with decreasing water content, because as water drains from pores, the flow decreases (2).

Although advection is associated with bulk macroscopic groundwater movement, diffusion is a molecular-based phenomenon. The process involves spreading because of concentration gradients and random motion. Diffusion causes a solute in water to move from an area of higher concentration to an area of lower concentration. Diffusive transport can occur in the absence of velocity. Mass transport in the subsurface because of diffusion in one dimension can be described by Fick's first law of diffusion,

$$f_x = -D_d\frac{dC}{dx} \tag{3}$$

where f_x is the mass flux [M/L^2/T], D_d is the diffusion coefficient [L^2/T], and dC/dx is the concentration gradient [M/L^3/L].

Typical values of D_d are relatively constant and range from 10^{-4} to 10^{-5} cm^2/sec (3). In cases of very low velocities, such as in a tight soil or clay liner, or in the case of mass transport involving very long time periods, diffusion is the dominant hydrodynamic process. On the other hand, in areas of high groundwater velocity, molecular diffusion can actually be ignored compared with advection and other dispersion mechanisms.

Freeze and Cherry (4) defined hydrodynamic dispersion as the process in which solutes spread out and are diluted compared with simple advection alone. It is defined as the sum of molecular diffusion and mechanical dispersion, where diffusion is caused by concentration gradients alone and dispersion is generally caused by local variations in velocity around some mean velocity of flow. The net effect of the dispersive phenomena is to decrease the concentration gradient between two different solutions of the same solute. In essence, dispersion reduces the sharp interface between the two solutions. Mass transport because of dispersion can also occur normal to the direction of flow. This transverse dispersion, D_y, is caused by diverging flowpaths in the porous media and concentration gradients that cause mass to spread laterally from the main direction of flow. In most cases involving a two-dimensional plume of contamination, D_y is much less than D_x, and the shape of the plume tends to be elongated in the direction of flow (5). Many typical contaminant plumes in groundwater are represented by 2-D advective-dispersive mechanisms.

One-Dimensional Transport Equations

The governing mass transport equation is difficult to solve in field cases of practical interest because of boundary irregularities and variations in aquifer characteristics, so numerical methods must generally be employed (6). However, a limited number of relatively simple, 1-D analytical solutions exist. The simplifying assumptions include the following: (a) the tracer is ideal, with constant density and viscosity; (b) the fluid is incompressible; (c) the medium is homogeneous and isotropic; and (d) only saturated flow is considered. Initial conditions ($t = 0$) in a soil column are usually set to zero ($C(x, 0) = 0$) or to some constant background concentration. Boundary conditions must be specified at the two ends of the 1-D column. For a continuous source load at $x = 0$, the concentration is set to $C(0, t) = C_0$ for $t > 0$. The concentration at the other boundary, $x = \infty$ is set to zero $C(x, \infty) = 0$ for $t > 0$.

For an infinite 1-D column with background concentration of zero and a continuous input tracer concentration C_0 at $-\infty \leq x \leq 0$ for $t \geq 0$, Bear (7) solves the problem using the Laplace transform at $x = L$, length of the column

$$\frac{C(x,t)}{C_0} = \frac{1}{2}\left(\mathrm{erfc}\left[\frac{L - v_x t}{2\sqrt{D_x t}}\right] + \exp\left(\frac{v_x L}{D_x}\right)\mathrm{erfc}\left[\frac{L + v_x t}{2\sqrt{D_x t}}\right]\right) \tag{4}$$

where erfc is the complementary error function, $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x) = 1 - (2/\sqrt{\pi})\int_0^x e^{-u^2}du$, D_x is the coefficient of hydrodynamic dispersion, and v_x is the average seepage velocity. The center of mass $(C/C_0 = 0.5)$ of the breakthrough curve travels with the average linear velocity, v_x, and corresponds to the point where $x = v_x t$. Note that the second term on the right-hand side of Eq. (4) can generally be neglected for most practical problems. The error functions erf (β) and erfc(β) are tabulated in Domenico and Schawartz (3).

The corresponding solution can be derived for the injection of a tracer pulse (instantaneous input) at $x = 0$ with background concentration equal to zero in the 1-D column. As the slug moves downstream with v_x in the $+x$ direction, it spreads out according to

$$C(x,t) = \frac{M}{(4\pi D_x t)^{1/2}}\exp\left[-\frac{(x - v_x t)^2}{4D_x t}\right] \tag{5}$$

where M is the injected mass per unit cross-sectional area. Figure 1b shows the resulting Gaussian distribution of concentration for the instantaneous pulse source in one dimension.

Plots comparing the shapes of Eqs. (4) and (5) are shown in Figs. 1a and 1b, respectively, at $x = L$ at the end of a soil column. The differences between instantaneous (spike) source and continuous source transport problems are obvious in one dimension. The continuous source produces a response or breakthrough curve, which starts at a low concentration and eventually levels off to the initial input concentration C_0 as a function of time. The spike source produces a normal or Gaussian distribution, which continues to decrease in maximum concentration because of spreading out with time in the direction of flow. For the spike source, the amount of mass under each curve is identical if the tracer is conservative.

Additional information can be found on transport equations in several sources (3–6,8), including three-dimensional derivations and those incorporating degradation.

Techniques for Determining Hydrodynamic Parameters

Tracers and tracer tests are valuable methods for estimating dispersion and advection below the water table. In a tracer test, a chemical species is introduced to the subsurface and monitored as it is carried by groundwater through the subsurface in the direction of flow. Average seepage velocity can be determined by analyzing the resulting curves in time and space, and hydraulic conductivity (K) values can be computed from Darcy's Law (Eq. 1). One of the most successful field sites for tracer studies has been the Borden landfill site in Canada, which resulted in a series of classic papers on dispersion and adsorption processes measured during a multiyear field experiment (9,10).

Other methods of determining hydraulic conductivity exist. Permeameters may be used in a laboratory setting to determine K, whereas slug tests and pump tests are recommended for field sites. A slug test involves suddenly adding or removing a volume of water (called a slug) from a well and monitoring the time it takes for the water level to return to normal. This test is convenient, but not as accurate as a tracer test to determine the speed with which a contaminant moves through an aquifer. The pump test involves the constant removal of water from a single well and the subsequent observation of drawdown in adjacent wells. This test gives the K value for a large part of an aquifer, but may produce a large volume of water that requires proper handling and treatment (if performed in a contaminated area) (3).

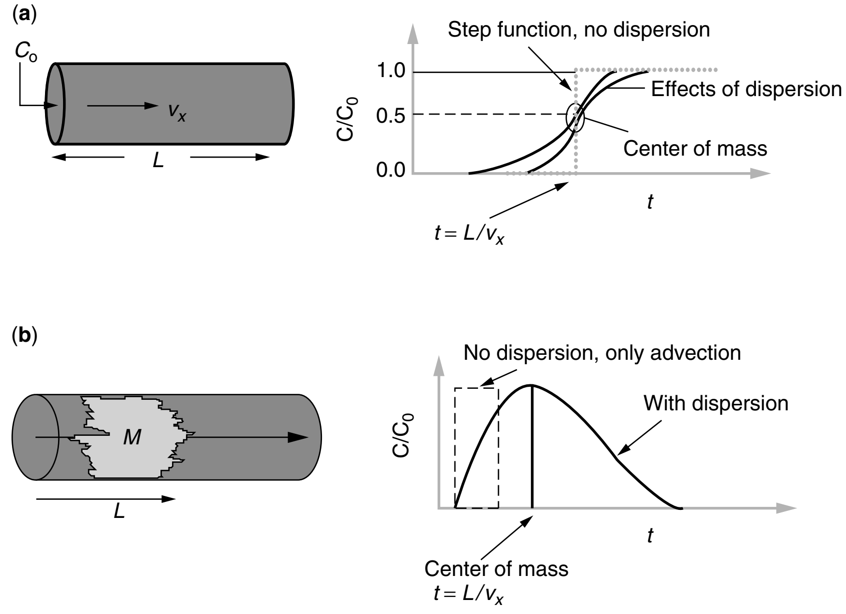

Figure 1. Breakthrough curves in one dimension for: (**a**) continuous and (**b**) instantaneous tracer inputs.

ABIOTIC SOLUTE TRANSPORT

Abiotic processes affect contaminant transport by causing interactions between the contaminant and the aquifer media or by changing the form of the contaminant that may subsequently interact with the subsurface material. The interfaces of most interest in subsurface fate and transport are liquid/solid interfaces and gas/solid interfaces. The processes that detail these interactions are discussed in this section.

Sorption

Sorption includes the phenomena of adsorption, absorption, and ion exchange. Weber (11) defines adsorption as the accumulation occurring at the interface, absorption as the partitioning between two phases (accumulation from groundwater into organic carbon), and sorption as including both adsorption and absorption. Often, the terms adsorption and sorption are interchangeable.

Many organic contaminants are removed from solution by sorption onto the aquifer matrix. Sorption of dissolved contaminants onto the aquifer matrix results in slowing (retardation) of the contaminant relative to the average groundwater seepage velocity and a reduction of dissolved organic concentrations in groundwater. In some cases, sorption is a reversible process; at a given solute concentration (steady state), some portion of the contaminant is partitioning out of solution onto the aquifer matrix and some portion is desorbing from the matrix back to solution; this process is called desorption.

The ability of chemicals to completely desorb from a solid is an area of current study (12–15). At high contaminant concentration, desorption patterns are generally predictable through equilibrium analysis. However, the phenomena of solid phase contaminant concentrations exceeding the predicted concentration, based on the surrounding aqueous phase concentrations, is observed in materials that have been contaminated for long time periods and contain low solid-phase contaminant concentrations (nominally less than 20 mg/kg). The importance of this observation is potentially enormous, because site remediation requirements include the removal of sorbed contamination. If a contaminant does not desorb as readily as it sorbed, the ability to meet cleanup levels may be seriously threatened (16).

Ion exchange is a specific category of sorption that occurs when the absorbent charge deficiency of a surface can be neutralized more efficiently by ions in solution than those ions currently absorbed (17). An example of ion exchange that is familiar is a home water softener. Home water softeners operate by accumulating the divalent cations from the groundwater onto the absorbent (ion exchange resin) while releasing monovalent ions, which reduces water hardness. In the subsurface, the resin would be natural minerals such as those found in clays. The most common ions occurring naturally are as follows (listed in order of occurrence from high to low): cations, Ca^{2+}, Mg^{2+}, Na^+, and K^+; anions, $SO_4{}^{2-}$, Cl^-, $PO_4{}^{3-}$, and $NO_3{}^-$ (18).

Cosolvation and Ionization

Cosolvation and ionization are two examples of processes that act to increase the chemical solubility, by altering the solvent phase or by altering the chemical form, respectively. Cosolvation is the process by which the mobility of one contaminant is enhanced by the presence of a solvent (19), whereas ionization refers to the gaining or losing of protons as a function of pH and thus going from a neutral form to an ionic form. These processes are solvent motivated; when a contaminant is in the presence of a solvent, it will have decreased sorption and an increased solubility leading to elevated concentrations in the surrounding solute. Based on these assumptions, it would be expected that as the fraction of cosolvent increases in solution, the level of sorption will decrease.

These attributes of cosolvents lead to increased mobility of originally slow dissolving hydrophobic (lacking affinity for water) contaminants. An example of this is the affect of ethanol on petroleum hydrocarbons. The issue concerning an ethanol releases' potential to cosolubilize benzene and other aromatic compounds in groundwater (20) has been addressed in recent groundwater modeling studies (21,22) and in laboratory studies (23,24). The reason that tremendous efforts are being employed to understand this phenomenon is that benzene, which is known to increase cancer risks, has a maximum contaminant level of only 0.005 mg/L in drinking water as set by the U.S. EPA and ethanol is now being used as a fuel additive (an oxygenate).

Oxidation and Reduction

Oxidation and reduction (or redox) refer to the transfer of electrons and contaminant species change of ions or compounds. Oxidation refers to the loss of electrons and reduction refers to the gain of electrons in a reaction. The two processes often occur in tandem; electrons given by one compound must be gained by another compound. Inorganic compounds, such as reduced metals and reduced sulfur, are possible mediators of oxidation-reduction reactions (25–27). The presence or absence of inorganic species is often dependent on aquifer conditions. A common condition is that many anaerobic environments will contain elevated levels of Fe(II) or Fe^{2+}, H_2S/HS^-, and other reduced species that are capable of donating electrons and reducing certain contaminants. For example, the two half reaction equations for the oxidation-reduction of ferrous iron (Fe^{2+}) to ferric iron (Fe^{3+}) are:

$$4\,Fe^{2+} \Leftrightarrow 4\,Fe^{3+} + 4e^-[\text{oxidation}]$$

$$O_2 + 4H^+ + 4e^- \Leftrightarrow 2H_2O[\text{reduction}]$$

Biologically mediated oxidation-reduction reactions are often more rapid than abiotic reactions for many contaminants of interest in groundwater systems, and thus abiotic oxidation-reduction are often neglected in transport calculations. In cases where they become important, oxidation-reduction processes are generally introduced as a first-order reaction (rate $= k^1X$) with respect to contaminant concentration, analogous to hydrolysis.

Hydrolysis

Hydrolysis refers to a chemical reaction between a contaminant molecule and water. Hydrolysis may function when direct biodegradation (microbially driven) of the contaminant cannot, and hydrolysis may produce a product that is susceptible to biodegradation, which can be significant in affecting fate and transport (17). With respect to organics, this process frequently results in the formation of alcohols and alkenes (28).

The rate that a chemical undergoes hydrolysis is strongly influenced by the temperature and pH of the system (27). As temperature increases, hydrolysis rates increase. As pH becomes either acidic or alkaline (e.g., greater than or less than 7.0), rates of acid-catalyzed or base-catalyzed hydrolysis can increase. Typically, hydrolysis is described with a first-order rate expression as shown in Eq. 6

$$\frac{C}{C_0} = e^{-k_{hyd}t} \tag{6}$$

where C is the concentration of the chemical species and k_{hyd} is a first-order hydrolysis rate coefficient. Knowing factors, such as pH and temperature, is important in determining the appropriate rate coefficient for a particular contaminant.

Volatilization

The transfer of a contaminant from the aqueous phase, nonaqueous phase liquid (NAPL), or sorbed phase directly to the gas phase is a process referred to as volatilization. The rate and extent to which volatilization occurs is strongly influenced by a number of parameters including the contaminant phase, the contaminant's vapor pressure, environmental factors (e.g., temperature and others), proximity in respect to the vadose zone, and other site-specific parameters (29,30). As a result of the range of factors, that play an important role in vapor transport, which include the physical transport processes of advection and dispersion, it is difficult to calculate the contribution of volatilization to the fate of chemicals in groundwater.

For the evaluation of volatilization, however, consider a closed bottle that contains water and air. If a contaminant is introduced to the bottle and allowed to equilibrate, some fraction of the contaminant added will reside in the gas phase. The distribution of a chemical between the water and the gas at equilibrium is described by the Henry's law coefficient (31,32). Mathematically, Henry's law is presented in Eq. (7), where P_c is the partial pressure of the contaminant, $[C]_{aq}$ is the aqueous phase concentration of the contaminant, and H_c is the Henry's constant.

$$H_c = \frac{P_c}{[C]_{aq}} \tag{7}$$

In groundwater systems that are isolated from the vadose zone, Henry's law can be applied to estimate the distribution of contaminants between the aqueous phase and the gas phase. If direct exchange is possible with the vadose zone, equilibrium may not be achieved and the calculation of gas phase concentrations is more difficult. Additionally, volatilization from NAPLs in the vadose zone or floating on the water table can yield high concentrations of contaminants in the gas phase (16).

BIOTIC PROCESSES AFFECTING SOLUTE TRANSPORT

Biotic processes include those biological phenomena that affect the transport and/or fate of compounds (organics, inorganics, or the microbes themselves) in the subsurface. Biodegradation of organics is one process that is of particular interest to minimize human health risks associated with the widespread contamination of chlorinated ethenes and hydrocarbons in the environment, because this process is able to achieve mineralization (endproducts CO_2,H_2O, etc.) and effectively render the contaminant innocuous (harmless). It has only been in the last two decades that the variety and magnitude of micro-organisms present in aquifers, and thus the ability of aquifers to support biodegradation, has been realized (33). Laboratory column studies (34) and improved field sampling methods are chiefly responsible for these advances.

Energy Flow and Metabolism

All living organisms require energy to survive, and the capture of usable energy through the process of catabolism is a significant part of an organism's overall metabolism. Bacteria that obtain energy from chemical forms are classified as chemotrophs. Chemotrophic organisms are further differentiated as either lithotrophs, which use inorganic sources of energy, or organotrophs, which use organic compounds. Examples of lithotrophic substrates include ammonia (NH_3), hydrogen (H_2), ferrous iron (Fe^{2+}), and sulfide (HS^-). Examples of common groundwater contaminants that serve as organotrophic substrates include fuel components and chlorinated solvents (35).

Biodegradation involves the process of extracting energy from organic chemicals via oxidation of the organic chemicals during catabolism. As a chemical is oxidized, it loses electrons. For this reason, chemotrophic substrates are often referred to as primary electron donors. Electrons lost during oxidation are coupled with the reduction of electron acceptors. Eventually, the complex sequence of oxidation-reduction reactions that occur during catabolism results in the reduction of a terminal electron acceptor. Common terminal electron acceptors in groundwater systems include oxygen (O_2), nitrate ($NO_3{}^-$), ferric iron (Fe^{3+}), sulfate ($SO_4{}^{2-}$), and carbon dioxide (CO_2). Aerobic catabolism occurs when oxygen is the terminal electron acceptor, and anaerobic conditions are in place when all other terminal electron acceptors are used. Typically, an individual bacterial strain is capable of using only one terminal electron acceptor. Facultative aerobes are the exception to this rule; these organisms can use nitrate as a terminal electron acceptor in the absence of the preferred oxygen. An important evaluation of contaminant biodegradation potential involves an analysis of the availability of primary electron donors and terminal electron acceptors, as their presence is required for organisms to obtain energy for survival and selects the types of organisms that will be present (16).

The metabolism of organic contaminants can be broadly differentiated by the ability of organisms to use the contaminant for catabolic processes. If the compound is a primary electron donor and provides the bacterium with energy for cell growth, the contaminant is referred to as a primary substrate, whereas secondary substrates are contaminants that provide the cell with energy, but are present at concentrations that are not sufficient to support the energy requirements of the organism. If a compound is metabolized fortuitously as the cell is obtaining energy from another primary electron donor, the transformation is referred to as cometabolic (36,37). However, cometabolism is generally a slow process and not all chlorinated solvents are susceptible (e.g., PCE and carbon tetrachloride). In the case of chlorinated ethenes, the contaminant serves as an anaerobic terminal electron acceptor in a process referred to as dehalorespiration, which is faster than cometabolism.

Metabolic Pathways of Common Contaminants

An overview of how bacteria are capable of metabolizing two of the primary groundwater contaminants (hydrocarbons and chlorinated solvents) is presented in this section. It is important to note that this continues to be an area of extensive research. Theoretical frameworks continue to change and it is difficult to assert definitive analysis regarding the totality of metabolic processes that may impact an individual contaminant's fate.

As a result of accidental spills and leaking underground storage tanks, hydrocarbon compounds consisting of gasoline, diesel, and jet fuels are common groundwater contaminants. Of particular concern are benzene, toluene, ethylbenzene, and xylene isomers (BTEX), which, under aerobic conditions, are rapidly biodegraded as primary substrates (38). Unfortunately, the oxygen demand resulting from the aerobic oxidation of these compounds can exceed the solubility of oxygen in water. Dissolved oxygen concentrations are rapidly depleted because of the biodegradation of these and other fuel constituents, yielding anaerobic conditions in contaminated aquifers (39).

Only recently have laboratory studies confirmed that the biodegradation of all BTEX compounds occurs under all anaerobic electron acceptor conditions (listed above). From these studies, it appears that toluene is the most degradable under all electron acceptor conditions and benzene is the least degradable. Certain modeling studies of certain well-characterized contaminated field sites infer that anaerobic benzene degradation may occur *in situ*. Only several recent laboratory studies conducted under anaerobic conditions have demonstrated that benzene may degrade in the absence of oxygen (40). Despite the possibility of anaerobic degradation of BTEX compounds, it is still generally slower than rates that are observed through aerobic processes.

Chlorinated methanes, ethanes, and ethenes comprise a group of compounds commonly referred to as chlorinated solvents. Chlorinated ethenes, particularly tetrachloroethene (PCE) and 1,1,2-trichloroethene (TCE), are commonly used as degreasers and cleaning solvents in the dry cleaning, aviation, and computer industries. Through accidental spills, improper storage, and disposal techniques, these chemicals have become major subsurface contaminants (41). The extensive impacts observed are primarily because of their widespread usage by industry, their general mobility and persistence in the environment, and the health risks associated with compound exposure.

The metabolism of chlorinated solvents is perhaps more diverse than any other group of environmental contaminants. Depending on the compound of interest, the electron acceptor condition, and the presence of inducing substrates, the metabolism of chlorinated solvents may occur through primary metabolism, secondary metabolism, cometabolism, or through terminal electron acceptor processes (16). Laboratory studies (34,36,37,42–44) have shown that, in the presence of a suitable substrate, certain microbes are capable of anaerobically metabolizing PCE by using it as an electron acceptor. In this process, known as reductive dechlorination, PCE may be dechlorinated to 1,1,2-trichloroethene (TCE), to typically cis-1,2 dichloroethene (DCE), to vinyl chloride, and eventually to ethene and ethane, an innocuous compound (45).

SUMMARY OF SOLUTE TRANSPORT PHENOMENA

Fate and transport of solutes in subsurface soils and groundwater is impacted by a combination of many processes that are difficult to quantify. No two field sites or aquifers have the exact same characteristics. Soil type, pH, temperature, groundwater flow velocity, the availability of electron donors for microbes, and the contaminant itself (among other parameters) all influence the path and lifetime of the pollution source. However, measurement techniques are improving, and understanding these processes is an area of on going research.

In summary, nonreactive (conservative) chemicals tend to move unimpeded through the subsurface and are subject to only hydrodynamic processes such as advection, dispersion, and diffusion. On the other hand, many actual contaminants (organic solvent and fuels, and inorganic metals) can be affected during groundwater transport by a number of subsurface mechanisms and chemical reactions. The processes that have the largest impact on contaminant fate and transport include advection, dispersion, sorption, cosolvation, volatilization, biological oxidation-reduction reactions, and cometabolism/mineralization. These processes are reviewed in Table 1, and Fig. 2 provides an illustration of these processes acting on a typical contaminant spill.

BIBLIOGRAPHY

1. Wiedemeier, T.H., Rifai, H.S., et al. (1999). *Natural Attenuation of Fuels and Chlorinated Solvents in the Subsurface*. John Wiley & Sons, New York.
2. Wierenga, P.J. (1995). *Water and solute transport and storage, Handbook of Vadose Zone Characterization and Monitoring*. CRC Press, Boca Raton, FL, pp. 41–60.
3. Domenico, P.A. and Schwartz, F.W. (1998). *Physical and Chemical Hydrogeology*. 2nd Edn. John Wiley & Sons, New York.

Table 1. Subsurface Processes Affecting Solute Transport

Category	Process	Contaminant (Solute) Behavior
Hydrodynamic Solute Transport	Advection	Plug flow; movement with bulk fluid
	Diffusion	Movement from high to low concentration
	Dispersion	Spreads out and dilutes
Abiotic Solute Transport	Sorption	Decreased solubility
	Ion exchange	Neutralizes the charge and sorbs
	Cosolvation	Increased solubility by solvent
	Oxidation-Reduction	Transfer of electrons
	Hydrolysis	More susceptible to biodegradation
	Volatilization	Phase transfer from aqueous or sorbed to gas
Biotic Transport Influences	Cometabolism	Metabolized as cell obtains energy
	Mineralization	Broken down to benign products
	Oxidation-Reduction	Microbial driven transfer of electrons

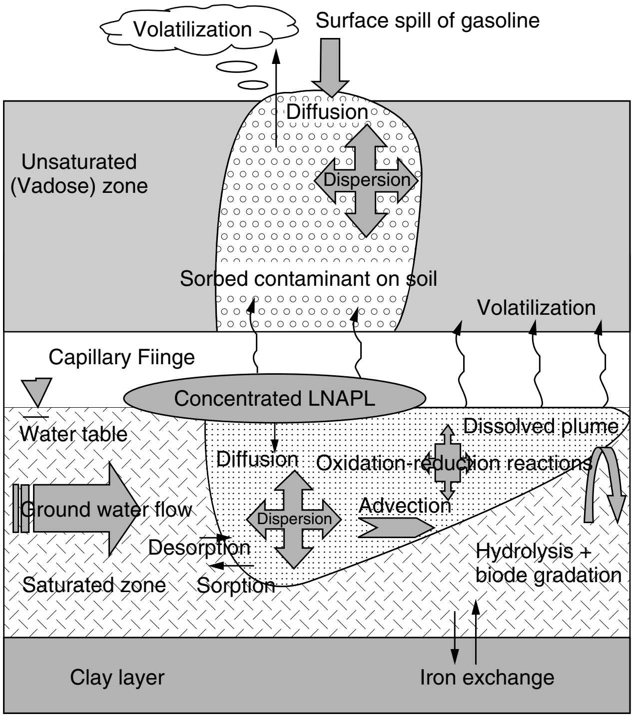

Figure 2. A typical scenario showing the processes impacting a surface gasoline spill, and its transport through the vadose zone to the groundwater.

4. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
5. Charbeneau, R.J. (2000). *Groundwater Hydraulic and Pollutant Transport*. Prentice-Hall, Upper Saddle River, NJ.
6. Bedient, P.B., Rifai, H.S., and Newell, C.J. (1999). *Ground Water Contamination, Transport and Remediation*. Prentice-Hall, Upper Saddle River, NJ.
7. Bear, J. (1979). *Hydraulics of Ground Water*. McGraw-Hill, New York.
8. Fetter, C.W. (1993). *Contaminant Hydrogeology*. Macmillian Publishing Company, New York.
9. Mackay, D.M., Freyberg, D.L., Roberts, P.V., and Cherry, J.A. (1986). *Water Resources Res.* **22**(13): 2017–2029.
10. Roberts, P.V., Goltz, M.N., and Mackay, D.M. (1986). *Water Resources Res.* **22**(13): 2047–2058.
11. Weber, W.J., Jr. (1972). *Physiochemical Process for Water Quality Control*. John Wiley & Sons, New York.
12. Fu, G., Kan, A.T., and Tomson, M.B. (1994). *Environ. Toxicol. Chem.* **13**(10): 1559–1567.
13. Kan, A.T., Fu, G., and Tomson, M.B. (1994). *Environ. Sci. Technol.* **28**: 859–867.
14. Kan, A.T., Fu, G., Hunter, M.A., and Tomson, M.B. (1997). *Environ. Sci. Technol.* **31**: 2176–2185.
15. Pignatello, J.P. and Xing, B. (1995). *Environ. Sci. Technol.* **10**: 1–11.
16. Hughes, J.B. (1999). *Ground Water Contamination, Transport and Remediation*. Prentice-Hall, Upper Saddle River, NJ, pp. 203–240.
17. Knox, R.C., Sabatini, D.A., and Canter, L.W. (1993). *Subsurface Transport and Fate Processes*. Lewis Publishers, Boca Raton, FL.
18. Mitchell, J.K. (1976). *Fundamentals of Soil Behavior*. John Wiley & Sons, New York.
19. Gorelick, S.M., Freeze, R.A., Donohue, D., and Keely, J.F. (1993). *Groundwater Contamination: Optimal Capture and Containment*. Lewis Publishers, Boca Raton, FL.
20. Rice, D., Powers, S.E., and Alvarez, P.J.J. (1999). Potential scenarios for ethanol-containing gasoline released into surface and subsurface waters. In: *Health and Environmental Assessment of the Use of Ethanol as a Fuel Oxygenate*. Report to the California Environmental Policy Council in Response to Executive Order D-5-99. Vol. 4.
21. Molsen, J.W., Barker, J.F., Frind, E.O., and Schirmer, M. (2002). *Water. Res. Research* **38**(1).
22. Powers, S.E. et al. (2001). *Crit. Rev. Environ. Sci. Technol.* **31**: 79–123.
23. Heermann, S.E. and Powers, S.E. (1998). *J. Contam. Hydrol.* **34**: 315–341.
24. Rixey, W.G. and He, X. (2001). Dissolution characteristics of ethanol from NAPL sources and the impact on BTEX groundwater concentrations. In: *Proceedings of the Conference on Petroleum Hydrocarbons and Organic Chemicals in Groundwater*. November 14–16, Houston, TX, Ground Water Publishing Company, Dublin, OH, pp. 41–52.
25. Kleeka and Gonsior. (1984). *S.J. Chemosphere* **13**: 391–402.
26. Macalady, D.L. and Wolfe, N.L. (1984). *Abiotic hydrolysis of Sorbed Pesticides*. ASC Symposium Series, No. 259, pp. 221–244.
27. Schwartzenbach, R.P., Gschwend, P.M., and Imboden, D.M. (1993). *Environmental Organic Chemistry*. John Wiley & Sons, New York.

28. Johnson, R.L., Palmer, C.D., and Fish, W. (1989). *Transport and Fate of Contaminants in the Subsurface*. U.S. Environmental Protection Agency, Ada, OK, EPA/625/4-89-019.
29. Mackay, D. and Leinonen, P.J. (1975). *Environ. Sci. Technol.* **9**: 1178–1180.
30. Mackay, D. and Wolkoff, A.W. (1973). *Environ. Sci. Technol.* **7**: 611–614.
31. Mackay, D., Shiu, W.Y., and Sutherland, R.P. (1979). *Environ. Sci. Technol.* **13**: 333–337.
32. Thibodeaux, L.T. (1996). *Environmental Chemodynamics: Movement of Chemicals in Air, Water, and Soil*. John Wiley & Sons, New York.
33. Young, L.Y. and Cerniglia, C.E. (1995). *Microbial Transformation and Degradation of Toxic Organic Chemicals*. John Wiley & Sons, New York.
34. Maymo-Gatell, X., Tandoi, V., Gossett, J.M., and Zinder, S.H. (1995). *Appl. Environ. Microbiol.* **61**: 3928–3933.
35. Rittman, B.E. and McCarty, P.L. (2001). *Environmental Biotechnology: Principles and Applications*. McGraw-Hill, New York.
36. Hollinger, C., Schraa, G., Stams, A.J.M., and Zehnder, A.B. (1993). *Appl. Environ. Microbiol.* **59**(9): 2991–2997.
37. Carr, C.S. and Hughes, J.B. (1998). *Environ. Sci. Technol.* **32**: 1817–1824.
38. Alvarez, P.J.J. and Vogel, T. (1991). *Appl. Environ. Microbiol.* **57**: 2981–2985.
39. Lee, M.D. et al. (1988). *CRC Cri. Rev. Environ. Control.* **18**: 29–89.
40. Lovely, D.R., Coates, J.D., Woodward, J.C., and Phillips, E.J.P. (1995). *Appl. Environ. Microbiol.* **61**: 953–958.
41. Pankow, J.F. and Cherry, J.A. (1996). *Dense Chlorinated Solvents and Other DNAPLs in Groundwater*. Waterloo Press, Portland, OR.
42. DiStefano, T.D., Gossett, J.M., and Zinder, S.H. (1992). *Appl. Environ. Microbiol.* **58**(11): 3622–3629.
43. Smatlak, C.R., Gossett, J.M., and Zinder, S.H. (1996). *Environ. Sci. Technol.* **30**(9): 2850–2858.
44. Yang, Y. and McCarty, P.L. (2000). *Environ. Sci. Technol.* **34**: 2979–2984.
45. Freedman, D.L. and Gossett, J.M. (1989). *Appl. Environ. Microbiol.* **55**: 2144–2151.

WATER IN THE UNSATURATED ZONE

Mukand Singh Babel
Asian Institute of Technology
Parthumthani, Thailand

The soil profile immediately below the land surface and above the water table is called the unsaturated zone, where the pores contain both water and air. Figure 1 shows a schematic diagram of the unsaturated zone. This zone differs from the saturated zone, where the pores are saturated with water. The unsaturated zone is sometimes called the vadose zone.

The thickness of an unsaturated zone can range from zero (e.g., when a lake or marsh is at the surface) to hundreds of meters, as is common in arid regions. It is a zone of natural and human-induced activity. It experiences physical phenomena such as thermodynamic interactions, transport processes of various kinds, and chemical reactions involving both natural and artificial substances. Several biological activities involving plant roots, rodents, worms, microbiota, and other organisms also take place in this zone. The zone is used for cultivating plants, constructing structures, and disposal of waste.

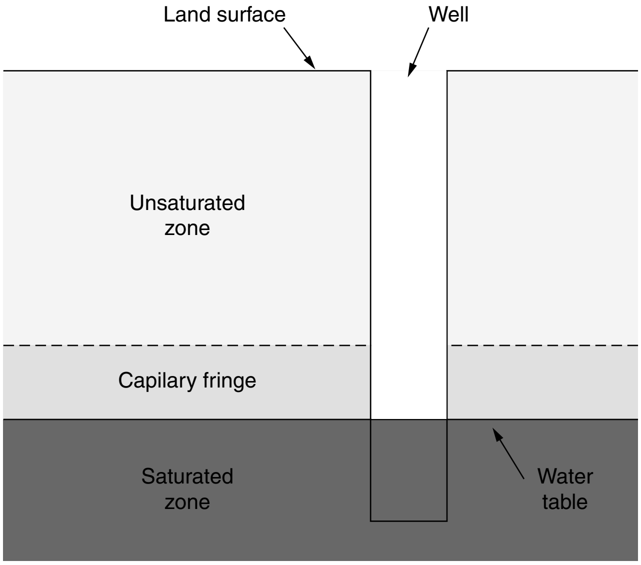

Figure 1. A schematic diagram of the unsaturated zone.

To a large degree, the unsaturated zone controls the transmission of water to aquifers (water bearing strata), as well as to the land surface, to water on the surface, and to the atmosphere. It also acts as a filter as it removes undesirable substances before they reach and affect aquifers. Figure 2 gives a pictorial representation of the unsaturated zone.

Though the unsaturated zone stores water, plant nutrients, and other substances, it is not always considered a major storage component of the hydrologic cycle because it holds only a tiny fraction of the earth's fresh water and this water is usually difficult to extract. But it is of great importance for storing water and nutrients in ways that are vital to the biosphere.

SOIL WATER

Soil water is normally considered to include both the water contained in the soil profile and the subsurface water in the unsaturated subsoil layers. Soil water may be present as gravity water in transit in larger pores, as capillary water in smaller pores, as hygroscopic water adhering

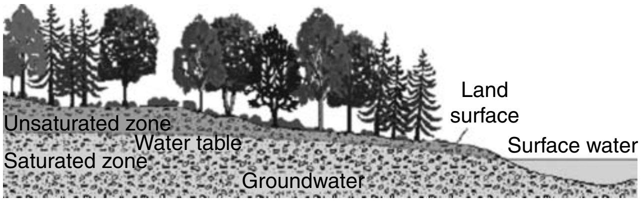

Figure 2. Unsaturated zone.

in a thin film to soil grains, and as water vapor. After a rain, gravitational water moves downward in larger pores, but it is dispersed into capillary pores or passes through the unsaturated zone to the groundwater or to a stream channel. Hygroscopic water is held very tightly to the soil surface by molecular attraction and is not normally removed from soil. Capillary water, therefore, is the important variable element of soil water.

SOIL WATER POTENTIAL

The most basic measure of the water in an unsaturated soil is water content or wetness, defined as the volume of water per bulk volume of the soil. Water is held in unsaturated soil by forces whose effect is expressed in terms of the energy state or pressure of the water. Like all matter, water can possess both kinetic and potential energy. However, in general, the flow of water in soils is too small to consider kinetic energy. Therefore, the energy state of soil water is represented by its equivalent potential energy. Potential energy is defined as the ability to do work. Work is done whenever an object is accelerated through a distance or heat is produced or transferred. Soil water potential thus measures the ability of soil water to move and/or produce heat. Water potential is important to any process in which soil water moves, such as infiltration and redistribution within the soil or removal of water from the soil by evaporation or plant uptake.

To define the soil water potential parameter, a reference body of water is defined, and the soil water potential of any other body of water is defined as the difference in potential energy concentration (potential energy per unit quantity of water) between the soil water and the reference body of water. The reference body of water is usually called the reference state, which is defined as "*pure* (no dissolved solutes), *free* (free of adsorptive forces) water at a specified temperature and *specified elevation* exposed to *atmospheric pressure*." The water potential of water in the reference state is zero.

Total soil water potential may formally be defined as the amount of work per unit quantity of pure water that must be done by external forces to transfer an infinitesimal amount of water reversibly and isothermally from the reference state to the soil at the point under consideration (1).

FACTORS AFFECTING WATER POTENTIAL

A change in the potential energy of the soil water brings about a change in the soil water potential. There are several factors that affect the potential energy of soil water, including adsorption of water onto soil particles, solutes dissolved in the soil water, the elevation of the soil water in the earth's gravitational field, and applied pressure (both positive and negative).

WATER POTENTIAL COMPONENTS

The total soil water potential can be broken into several components. A complete list of soil water potential components consists of gravitational potential, osmotic potential (also sometimes called solute potential), matric potential, vapor potential, hydrostatic pressure potential, and overburden pressure potential. The major potential components of interest in the unsaturated zone are gravitational potential, osmotic potential, and matric potential that give rise to the total potential as follows:

$$\Psi_{\mathrm{t}} = \Psi_{\mathrm{g}} + \Psi_{\mathrm{o}} + \Psi_{\mathrm{p}} + \dots \tag{1}$$

where Ψ_{t} is the total potential, Ψ_{g} is the gravitational potential, Ψ_{o} is the osmotic potential, and Ψ_{p} is the matric potential.

Gravitational potential is defined as the energy of water per unit volume that is required to move a specific amount of pure, free water from an arbitrary reference point to the soil-water elevation. It is independent of soil properties and solely dependent on the vertical distance between the arbitrary reference point and the interest point. Often, the soil surface is chosen as the point of reference, whereby the distance z (m) to each point of interest is negative by convention. Gravitational potential is given as

$$\Psi_{\mathrm{g}} = \rho_{\mathrm{l}} g z \tag{2}$$

where ρ_{l} is the density of the soil water, g is the acceleration due to gravity (m/s^2) and Ψ_{g} is the gravitational potential (J/m^3).

The osmotic potential component arises due to the salts in the soil water. It may be thought of more a suction than a pressure. In a dynamic soil–water–plant system, in the presence of a selective permeable membrane as found in the roots of plants, the pressure or energy potential (a suction) on the solution side is less than that of pure, free water. Water will pass through the membrane to the solution side until equilibrium is reached. The osmotic potential, Ψ_{o}, is given by

$$\Psi_{\mathrm{o}} = -\frac{\Pi}{\rho_{\mathrm{w}}} \tag{3}$$

where Π is the osmotic pressure and ρ_{w} is the density of water.

Soil is porous, so it acts as a semipermeable barrier to salt solutions. An individual particle has an electric field that causes a decrease in the concentration of the soil solution near the particle surface, which results in an increased cross-sectional area for water transport compared to the area available for transport of dissolved salts, causing water to move due to the osmotic difference. The amount of flow depends on the particle size of the soil and the gradient of the osmotic pressure. In coarse, saturated sand, soil pores are relatively larger than those in clays, and consequently, the thickness of fluid layer with decreased salt concentration is small compared to the total area available for transport. Thus, the semipermeability is small, and the transport of soil solution due to osmotic potential is negligible. In clays, with many more pores, the solutes in the soil solution can be completely repelled from the soil and restricted from flow. In such soils, water movement in response to an osmotic pressure gradient is significant.

The matric potential component results from the soil matrix. Water molecules can form hydrogen bonds with the surface of soil minerals that give rise to adsorption and with other water molecules that result in cohesion. In unsaturated soil, these forces of attraction between the soil mineral surfaces and the soil water exert a "pull" on the soil water reducing the free energy of the water. It is usual to deal with the combined effect of both capillary and adsorptive forces in the way in which water is held in the soil matrix. The matric potential is the potential of the water in a pore of the medium relative to the potential of the air that results primarily from the adsorptive and capillary forces of the soil matrix. When a medium is unsaturated, the water generally is at lower pressure than the air, so the matric potential is negative.

Moisture content increases as the matric potential increases. Zero matric potential is associated with high (saturated or nearly saturated) water content. As the matric potential decreases, the water content decreases. Figure 3 shows the relation between matric potential and water content, called a retention curve. It is a characteristic of a porous medium that depends on the nature of its pores. This relation strongly influences the movement of water and other substances in unsaturated media.

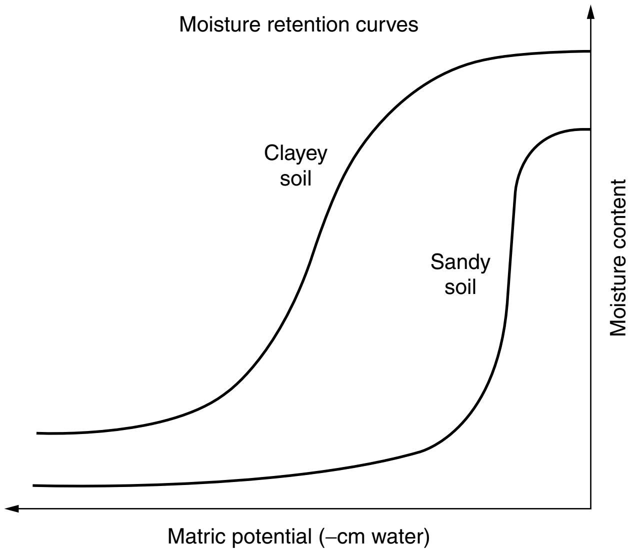

Figure 3. Moisture retention curves.

MEASUREMENT OF SOIL WATER

Soil water can be measured as either soil water content or soil water potential; conversion from one to the other uses the moisture retention curve. Their combined use provides a powerful approach to the study of soil fluxes in unsaturated soil.

The most commonly used method for measuring the water content of soil directly is the gravimetric method, which involves taking a number of soil samples of known volume and determining their weight loss when dried in an oven at 105 °C. The neutron probe is the most widely used indirect way of measuring soil water. In this method, a source of high-energy neutrons is used that is lowered into an aluminum access tube in the soil. Fast neutrons lose energy through collision with atoms of low atomic weight, such as the hydrogen in water, and are converted to slow neutrons. These slow neutrons are counted with a high count indicating high moisture content. Time-domain reflectometry (TDR) determines soil water content by measuring the dielectric constant of the soil. TDR has two significant advantages over the neutron probe method; it is nonradioactive and can be set up for an automatic, nonmanual operation. The satellite-mounted synthetic aperture radar (SAR) method has the potential for measuring soil moisture content on a drainage basin scale. It determines soil water content by monitoring changes in soil dielectric properties.

Tensiometers are the oldest and most widely used technique for measuring matric potential. They comprise a liquid-filled porous cup connected to a pressure-measuring device such as a mercury manometer or a pressure transducer. The cup is embedded in the soil, and water can flow between the soil and cup until the pressure potential inside the cup equalizes with that of the soil water. In fine-textured soils, the resistance block method may be useful. In this method, two electrodes are embedded in a porous block, which is buried in the soil. The matric potential in the block comes into equilibrium with that of the soil water, and the resistance across the electrodes varies with the resulting water content of the block. Gypsum resistance blocks can be used for measuring potentials as low as −15,000 cm (1500 kPa).

BIBLIOGRAPHY

1. Tindall, J.A. and Kunkel, J.R. (1999). *Unsaturated Zone Hydrology for Scientists and Engineers*. Prentice-Hall, Englewood, Cliffs, NJ.

GROUNDWATER AND VADOSE ZONE HYDROLOGY

Uwe Dannwolf
URS Australia Pty Ltd.
Turner, Australia

The vadose zone is the geologic media between the land surface and the regional water table.

The vadose zone

- is the foundation for most man-made structures;
- provides nutrients and water to all plant and animal life;
- protects drinking water reservoirs (aquifers);
- governs how much groundwater is replenished;
- provides the pathway for contaminants to penetrate into groundwater, and
- is one of the least understood media.

The latter point becomes evident when a definition for the vadose zone is needed. All publications include the unsaturated soil zone, but some publications include the capillary fringe; others do not. The term “vadose” is derived from the Latin noun *vadosu*s, which translates to “shallow”.

The vadose zone includes the weathered soil horizon and the unconsolidated sediments and rocks below. The capillary fringe separates the water saturated zone from the unsaturated zone. The water saturated zone is often referred to as an aquifer if it allows significant amounts of water to be transmitted. Several processes take place in the vadose zone.

PROCESSES IN THE VADOSE ZONE

The processes, which influence the vadose zone and the recharge to groundwater include

- precipitation,
- evapotranspiration,
- surface water runoff, and
- infiltration and groundwater recharge.

Precipitation

Precipitation is the water that falls on the land surface as rain, snow, and other forms such as dew, frost, and irrigation water. It is measured as volume per time period (e.g., mm/a) or as volume per area and time (e.g., L/km^2 s), which is a relatively simple parameter to measure and is often available from meteorological stations. However, care should be exercised when using this parameter for climatic water balancing because there is great variability even between neighboring stations, due to the precipitation and wind characteristics as well as the various measuring devices.

Evapotranspiration

Evapotranspiration, the reverse of precipitation, represents the transport of water from the earth back to the atmosphere through evaporation and plant transpiration. This process depends on the microclimatic conditions at a given site. This parameter depends on many factors (air temperature and humidity, solar irradiation, vegetation, wind speed, etc.). It is more difficult to measure and therefore not available from all meteorological stations. Simple (1) and complex calculations (2–5) can estimate its range. Its units are generally reported as millimeter per time period. The vegetation type and cover greatly influence this parameter and are often neglected in standard formulas, which can lead to significant variations.

Surface Water Runoff

Surface water runoff is the amount of precipitation that does not immediately enter the subsurface and contributes to surface water discharge (e.g., into local creeks). A portion of this surface water runoff appears on the surface again after a short underground passage (interflow). The amount of surface water runoff depends on the size of the catchment, vegetative cover, relief [(slope or gradient) and depressions (pond formation)], soil cover and soil type, initial soil moisture content, and rainfall intensity. For instance, if the supply rate of water to the soil surface is greater than the soil’s ability to allow the water to enter, excess water either accumulates on the soil surface or becomes runoff. Steep slopes cause greater runoff rates, allowing the water a shorter time to infiltrate. In contrast, gentle slopes have less impact on infiltration due to decreased runoff.

Surface water runoff is even more difficult to measure than any of the above parameters. Hydrological models aid in estimating it.

Effective Recharge or Infiltration

Both evapotranspiration and surface water runoff limit the amount of total precipitation that enters the subsurface, the effective recharge or infiltration into the vadose zone. The term infiltrability is generally used in soil physics and hydrology to define the maximum rate at which rain or irrigation water can be transmitted through a soil under a given condition.

As a general rule of thumb, a total of four water relevant zones exist in the vadose zone (after Reference 6):

1. a near-surface zone to shallow (<1 m) depths affected by episodic rain;
2. a zone below which shows signs of seasonal water infiltration;
3. a zone of virtually constant moisture content, in which the hydraulic gradient is unity (i.e., 1), and there is a predominantly downward flow toward the water table; and
4. a capillary zone above the water table, in which the water moves in the groundwater flow direction.

This simplified zoning is complicated by the fact that irregularities (heterogeneities) in the soil profile such as root channels, earthworm burrows, soil fissures or cracks, spatially varying sediment size therefore exhibiting variable hydraulic and physical characteristics

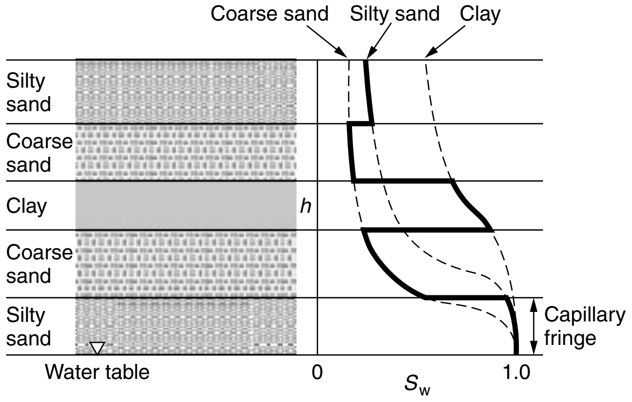

Figure 1. Saturated water content (S_w) above the water table for a layered sequence. Dashed lines represent water content for homogeneous sediments (from Reference 7, altered).

(see Fig. 1), and distinct layering allow preferential (i.e., faster) flow of water in selected areas. For instance, clay layers generally impede flow due to their lower saturated hydraulic conductivity, whereas sand layers have a lower tendency to retard the movement of infiltrating water. This process, which is also referred to as "fingering," can occur in less than a few percent and to up to 70% of the soil profile. Such layers generally retard water movement during infiltration. For the remainder of the soil profile, the entering moisture results in a semicontinuous wetting front, a zone of increased moisture content and a zone below of significantly lower moisture content. Beyond the wetting front, there is no visible penetration of water.

Infiltrating precipitation can appear as

1. retarded interflow discharging as surface runoff after a brief underground passage,
2. increased soil moisture, and
3. effective recharge.

The effective recharge can be measured by using lysimeters, but this requires a substantial effort. It is furthermore dynamic, changing constantly over time and space.

The measuring units are volume or m per time period. Infiltrating water also changes the moisture content of the soil. If rain infiltrates a previously dry soil (typical moisture content of less than 10% saturation), then the rainwater is first absorbed into the soil pores, increasing the moisture content. The soil gas in the soil profile needs to be replaced, and as a result the infiltration velocity is reduced initially. Later in the process when the soil gas dissolves gradually in the infiltrating water, the infiltration rate increases again before a renewed reduction occurs. This reduction is from increased solid particles washed into the soil profile from the rain and the growth of organisms near the soil surface. Additionally, the increase in soil moisture leads to a reduction in capillary pressure.

Snow and frost also impact the infiltration rate. In frost periods, soils of low initial water content may become porous, and the infiltration rate is increased. Soils of high initial water content may freeze and become nearly impermeable.

In detail, infiltration depends on material properties, such as

- grain size,
- porosity,
- capillary forces and soil suction,
- permeability, and
- fluid densities and viscosities.

MATERIAL DEPENDENCIES

Grain Size and Porosity

Unconsolidated materials are distinguished by their grain sizes. Various classifications exist, the grain sizes commonly decrease in the order block, cobble, gravel, sand, silt, clay, colloid, and ion. Clays are smaller than 2 μm, whereas colloids consist of particles less than 0.1 μm in size. On the other hand, gravels start at 2 mm. The distribution of soil grain size and bulk densities (often referred to as ρ) of soil and rock are commonly measured during geotechnical investigations.

With regard to water and contaminant transport, a critical parameter of vadose zone materials is the type and distribution of void spaces (pores). Loose materials contain widely distributed void spaces that typically are interconnected. These pores are usually individually very small, but numerous, and the total volume of pores usually varies from 30–50% of the total volume. However, not all of these pores are available for water circulation. A clay may exhibit a total porosity of >50%, however, only less than a few percent are available for water circulation. In contrast, in gravel sediments, almost all of the pore space (roughly 20% of the total volume) allows water transport.

As can be seen from Fig. 2, classifications exist for pore spaces. The largest pores are obviously present in the coarsest materials. Pores of the order of 60–1000 μm are present in sandy sediments. Small and smallest pores are present in clay grain size and below.

Also depicted in Fig. 2b is the approximate rate of water movement in meters per year. The coarser the sediments, the higher the rate of vertical water movement. It is, however, important to note that even in silts, the vertical water movement under unit gradient can be as high as 2 m per year.

Capillary Rise and Capillary Pressure

When water evaporates near the surface, it generates a low hydraulic potential, which is equilibrated by the capillary rise at the groundwater table. The rate at which this supply is determined is governed by the permeability (and size, respectively) of the sediments. Consequently, the height of the capillary fringe depends on the sediment's permeability, as Fig. 2a outlines. The water in the capillary fringe is held by capillary forces against gravity induced flow.

The capillary pressure (C_p, also named matrix or soil suction) for sands and silts can be estimated roughly from

$$C_p(\text{kPa}) = 0.3/\text{grain size (mm)} \quad (1)$$

The higher the soil suction, the higher the residual saturation of a sediment.

Soil–Water Characteristic Curve

The soil–water characteristic curve (also referred to as the water retention curve) for a soil is defined as the relationship between the water content and the capillary pressure. This curve is characteristic for each soil, it governs the rate at which water is taken up or released by the soil.

As can be seen from Fig. 3, when the water content decreases (i.e., the soil suction increases), the permeability of a sediment (here, a silty soil) decreases. The order of decrease depends on pore sizes. Sands have a much slighter decrease in permeability. Tight soils stay wet

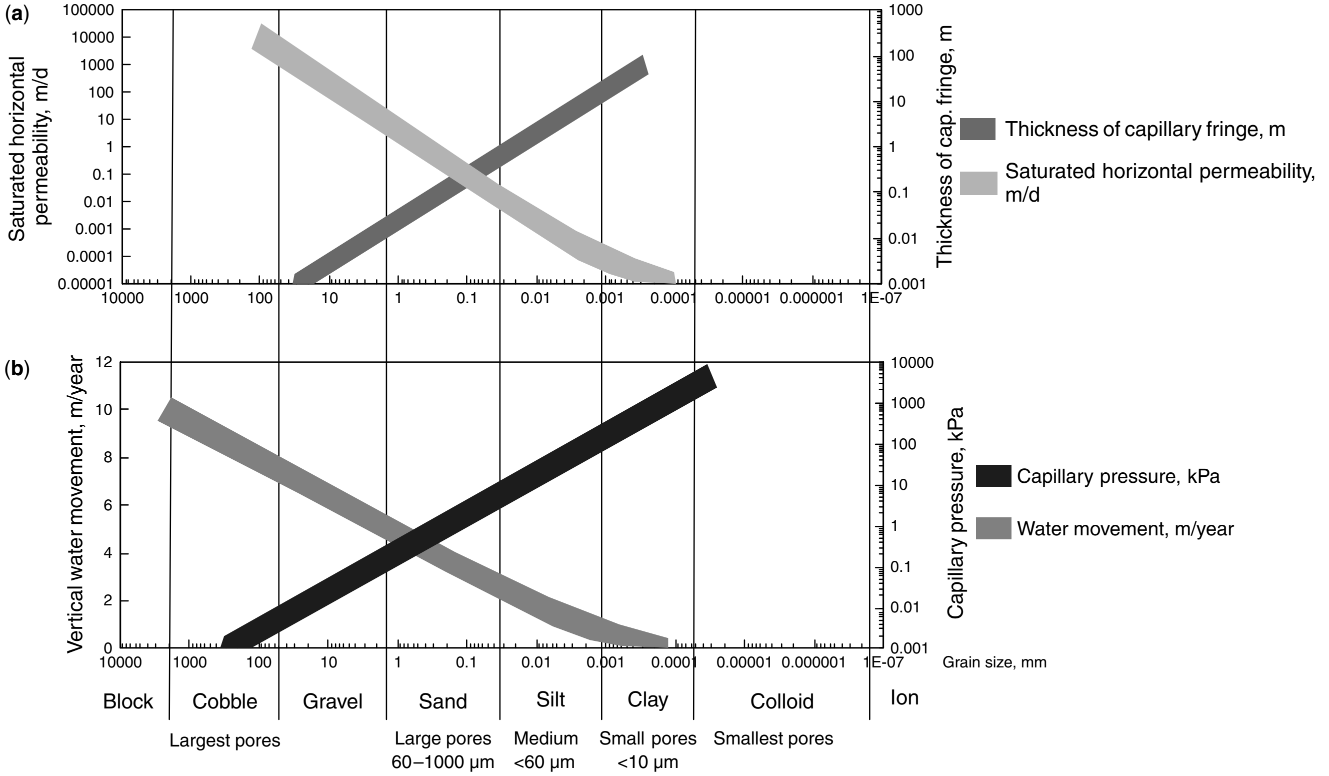

Figure 2. (**a**) Generalized diagram depicting the thickness of the capillary fringe and saturated horizontal permeability. (**b**) Generalized diagram depicting pore sizes, capillary pressure, and the vertical annual water movement rate of sediments.

for much longer, however, suffer much more from evaporative losses.

Figure 3 also introduces terms often used by soil scientists to describe the wetting and drying behavior of a soil. The term field capacity is used when the water content of a soil has reached a maximum (often referred to as Θ sat). Any further increase in soil moisture results in a water release (i.e., gravity induced flow). When the water content nearly equals its minimum, then the wilting point is reached. By definition, the wilting point equals a soil suction between 500 and 6000 kPa. A soil is referred to as "air dry" at or above a capillary pressure of 70,000 kPa. Furthermore, two terms are defined through this water retention curve: the air entry value and the residual water content. The air entry value is the meeting point of the two tangents shown in Fig. 3. The residual water content (referred to as Θ res) is defined similarly by using the lower tangent. Figure 3 further introduces a fundamental feature soils display when drying out or becoming wet (the hysteresis effect).

Hysteresis describes the effect of retardation of forces acting on a body to change (as if from viscosity or internal friction in magnetic induction and other physical phenomena). In soil science, when a soil is dried and then rewetted, the drying and wetting soil–water characteristic curves are not identical. The wetting curve is drier than the drying curve over a wide range of soil–water potentials. This hysteresis is caused by:

1. the drainage of an individual pore, which is controlled by the smallest radius. When the meniscus retreats under increasing tension, the pore's radius is decreasing. Rewetting (imbibition), on the other hand, is controlled by the pore's largest radius.
2. contact angle hysteresis. This is a well-recognized phenomenon, but it is only significant at relatively high flow rates. The slower the process, the less contact angle hysteresis. Because drainage and rewetting curves are generally measured at slow ("quasi-equilibrium") rates, this is not likely to be an important contributor.
3. Air entrapment effects. During imbibition, air can become trapped inside the larger soil pores. Given sufficient time, the entrapped air bubbles slowly dissolve in the water and cause hysteresis

Determining the Soil–Water Characteristic Curve. The most commonly used direct way to measure the soil–water potential is with a tensiometer. A body of pure water is hydraulically connected to the soil water via a permeable ceramic cup. If the soil water is under tension, then a small fraction of water in the tensiometer is pulled out, lowering the pressure inside. This negative pressure can then be measured by a gauge or pressure transducer.

Alternatively, mathematical models can help in determining the soil–water characteristics of a soil.

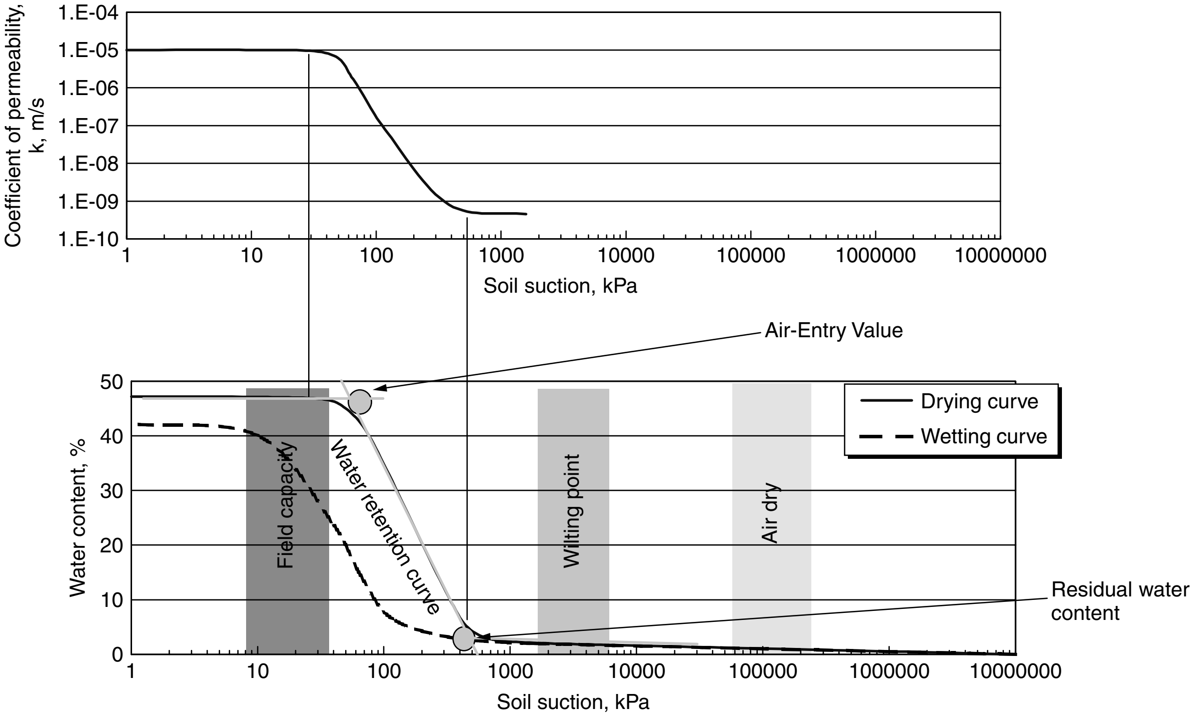

Figure 3. Visualization of the relationship between the permeability function and the soil water characteristic curve (from Reference 8, altered).

Hydraulic Conductivity

The basic formula to determine hydraulic conductivity was first developed by Henry Darcy in 1856 when he determined that the flux of water in sand columns depends on hydraulic conductivity and the hydraulic gradient. It can further be shown that hydraulic conductivity depends on both the water and the porous media properties according to

$$K = \frac{k\varphi g}{\mu} \tag{2}$$

where μ = viscosity,
ϕ = density
g = acceleration of gravity
k = intrinsic permeability of the soil (independent of fluid).

The intrinsic permeability is generally a function of the porous media, and it is proportional to the square of the grain diameter, $k \sim d^2$. Consequently, capillary pressure depends on the soil's conductivity:

$$Cp \sim 1/\sqrt{K} \tag{3}$$

This relationship is commonly used in numerical modeling studies to scale the magnitude of the capillary pressure curves in the absence of direct measurements of capillary pressure.

Measurements of Hydraulic Conductivity. The hydraulic conductivity of an aquifer with respect to water can be determined by a variety of field tests, including pumping tests from groundwater monitoring wells, slug tests, and other indirect measurements like grain-size distribution. In vadose zone studies, tensiometers are used for determining the field capacity of a soil.

FLOW PROCESSES IN THE VADOSE ZONE

Two processes control water in the vadose zone. The first is gravity, which moves water downward. The second is a capillary process that, similar to water dripped slowly on a sponge, moves water in all directions and stores and releases it. The capillary process is controlled by the nature of the sediment and rock and the pores and fractures within it. In most cases, capillary processes are dominant in fine-grained sediments (clay, silt, and the like), whereas gravity is dominant in coarse-grained sediment and large fractures. Pores and fractures can be completely filled with water in places within the vadose zone. These saturated portions of the vadose zone are typically referred to as "perched water" or "capillary fringe," depending on where they occur. At most sites, the water content is changing over time throughout the system representing a dynamic system.

Once saturation is continuous, this groundwater moves through the system from the highest potential to the lowest. Numerical modeling studies show that a range of infiltration patterns occur, which can be related to the ratio of infiltration rate to unsaturated hydraulic conductivity. A high value of this ratio reflects a prevailing hydraulic conductivity, which cannot readily redistribute the newly infiltrated moisture. Moisture accumulates in the near-surface region before advancing down through the soil as a distinct wetting front. In contrast, low values of the ratio of rainfall to unsaturated hydraulic conductivity show minimal moisture accumulation, as the relatively small volumes of infiltrating moisture are readily redistributed through the soil profile.

MODELS

Accurate *in situ* measurement of unsaturated hydraulic conductivity is cumbersome and time-consuming. One alternative to direct measurement of unsaturated hydraulic conductivity is to use theoretical methods, which predict the conductivity from more easily measured soil water retention data. Methods of this type are generally based on statistical pore-size distribution models, a large number of which have appeared in the soil science and petroleum engineering literature during the past several decades (see Reference 9 for a review). Implementation of these predictive conductivity models still requires independently measured soil–water retention data. Measured input retention data may be given either in tabular form or by closed-form analytical expressions, which contain parameters that are fitted to the observed data.

Water flow in variably saturated soils is traditionally described by the Richards equation:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left\{K(\theta)\frac{\partial h}{\partial z}\right\} + \frac{\partial K(\theta)}{\partial z} \qquad (4)$$

where h is the soil water pressure head (or fluid pressure with dimension L), θ is the moisture content [−], t is time (T), z is soil depth (L), and K is the hydraulic conductivity [LT−1]. The solution of the Richards equation requires knowledge of the unsaturated soil hydraulic functions $\theta(h)$ and $K(h)$ or $K(\theta)$.

Several equations have been proposed to describe the soil–water retention curve near saturation. Most of these functions are mathematically too complicated to be easily incorporated into predictive pore-size distribution models for hydraulic conductivity. A smooth function with attractive properties is the equation of van Genuchten (10):

$$S_{\mathrm{eff}}(h) = [1 + (ah)^{n}]^{-m} \qquad (5)$$

where a, n, and m are empirical constants affecting the shape of the retention curve (i.e., curve fitting parameters). Another mathematical solution is provided by Brooks and Corey (11).

Various computer programs exist to model unsaturated soil properties, including (but not limited to)

- Soil Vision: www.soilvision.com.
- Rosetta: U.S. Salinity Laboratory USDA; http://www.ussl.ars.usda.gov/models/rosetta/rosetta.
- RETC: Code for Quantifying the Hydraulic Functions of Unsaturated Soils, U.S. Salinity Laboratory USDA; http://www.ussl.ars.usda.gov/models/retc.
- SEEP/W: http://www.geo-slope.com/products/seepw.asp.
- Hydrus: http://www.ussl.ars.usda.gov/models/tutorials/hydrus2d/tutorial2d.htm.
- UnSat: Waterloo Hydrogeologic Inc;. http://www.flowpath.com.

BIBLIOGRAPHY

1. Linacre, E. (1993). A three-resistance model of crop and forest evaporation. *J. Theor. Appl. Climatology* **48**: 41–48.
2. Penman, H.L. (1948). Natural evaporation from open water, bare soil and grass. *Proc. R. Soc.London* **A193**: 120–146.
3. Thorntwaite, C.W. and Mather, J.R. (1955). *The water balance, Publication in Climat.* VIII, Nr.1. Drexel Inst. Laboratory of Climatology, Centerton, NJ.
4. Turc, L. (1954). Le bilan d'eau des sols, relations entre les precipitations, l'evaporation et l'ecoulement. *Trois Journée d'Hydraulique*. Alger.
5. Haude, W. (1955). Zur Bestimmung der Verdunstung auf möglichst einfache Weise. *Mitt. dt. Wetterd.* **2**: 11, Bad Kissingen.
6. Faybishenko, B. (2000). Vadose Zone Characterization and Monitoring—Current Technologies, Applications, and Future Developments. In: *Vadosc Zone Science and Technology Solutions*. Battelle Press, Columbus, OH.
7. Looney, B.B. and Falta, R.W. (Eds.). (2000). *Vadose Zone Science and Technology Solutions*. Battelle Press, Columbus, OH, pp. 1–1465.
8. Fredlund, M.D. (1988). Unsaturated seepage modelling made easy. *Geotechnical News* **June**.
9. Mualem, Y. (1986). Hydraulic conductivity of unsaturated soils: Prediction and formulas. In: *Methods of Soil Analysis*. A. Klute (Ed.), Part 1. *Physical and Mineralogical Methods*. Agron. Monogr. 9, 2nd Edn. American Society of Agronomy, Madison, WI, pp. 799–823.
10. van Genuchten, M.Th. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* **44**.
11. Brooks, R.H. and Corey, A.T. (1964). *Hydraulic Properties of Porous Media*. Hydrology Papers, No. 3, Colorado State University, Fort Collins, CO.

VADOSE ZONE MONITORING TECHNIQUES

Philip R. Rykwalder
University of Texas at San Antonio
San Antonio, Texas

William J. Blanford
Louisiana State University
Baton Rouge, Louisiana

INTRODUCTION

Monitoring fluids of the vadose zone has multiple motives, including detecting contaminants, measuring soil properties and water conditions for construction or monitoring geologic hazards, and determinations of water saturation that may be used for optimizing irrigation. Because fluids must move through the vadose zone to reach the water table, monitoring of the vadose zone allows predicting possible sources and processes of groundwater contamination. Geologic hazards such as landslides, rising water tables, or land subsidence can be partially predetermined by monitoring the vadose zone. And in agriculture, monitoring soil water content to irrigate most efficiently has a long history.

Monitoring the temporal and spatial variance in the vadose zone is vitally important to engineers, geologists, meteorologists and other environmental scientists interested in studying water and contaminants. There are many techniques that directly or indirectly assess water content, sorbed mass, gaseous and liquid contaminants, flow direction, and other parameters that may be monitored. The methods discussed here are divided into two categories, direct and indirect monitoring techniques. Direct monitoring techniques require a degree of sampling or large-scale soil disturbances. Conversely, indirect techniques are analogous to remote sensing in which no disturbance to the soil or water is required, although a well may be required in some techniques. This section discusses in short detail some of the attributes and shortfalls of most of the currently used techniques for monitoring the three phases of the vadose zone: liquid, soil, and gas. For a more in-depth understanding, readers are directed to the references listed after each entry.

INDIRECT MONITORING METHODS

Indirect methods for monitoring the vadose zone commonly have links to geophysical tools, principles, and techniques. Many vadose zone characteristics such as water content, the presence of contamination, and flow direction, among others may be determined without soil sampling. These techniques may also be used to determine the nature and extent of soil stratification for site delineation.

In practice, these methods are generally nonintrusive and do not require disturbing the soil, but some require well installation. When drilling or augering is necessitated, proper techniques must be ensured so that mud linings, drilling fluids, or other drilling by-products do not affect hydraulic conductivity, which may then affect measurements. It is also important to clean any grease or oil from drilling tools to avoid contamination and to clean tools between use.

Tension Infiltrometers

Tension infiltrometers are used to characterize hydraulic properties of saturated and partially saturated porous media, including saturated and unsaturated hydraulic conductivities. Infiltrometers are commonly used to measure hydraulic conductivity as a function of matrix potential in partially saturated porous media, sorptivity, steady-state volumetric flow rate (Q), and water content (Fig. 1).

There are variations of the disk infiltrometer and multiple methods of measuring hydraulic properties using infiltrometers, but all consist of three basic parts: the bubble tower, the reservoir, and a porous baseplate. The porous base plate lies flat on a wetted soil surface, and both the bubble tower and reservoir are connected perpendicularly to it. The purpose of the baseplate is to establish hydraulic contact between the soil and the reservoir. A thin layer of moist sand may be placed under the baseplate to ensure better hydraulic conductivity between it and the soil. The reservoir supplies water to the baseplate and the bubble tower; the latter consists of a series of tubes that relieve the vacuum in the reservoir when water infiltrates into the soil. Tension measurements correspond to the level of the water in the reservoir. An important point to note is that hydraulic conductivity must exist between the baseplate and the wet porous media for the device to operate.

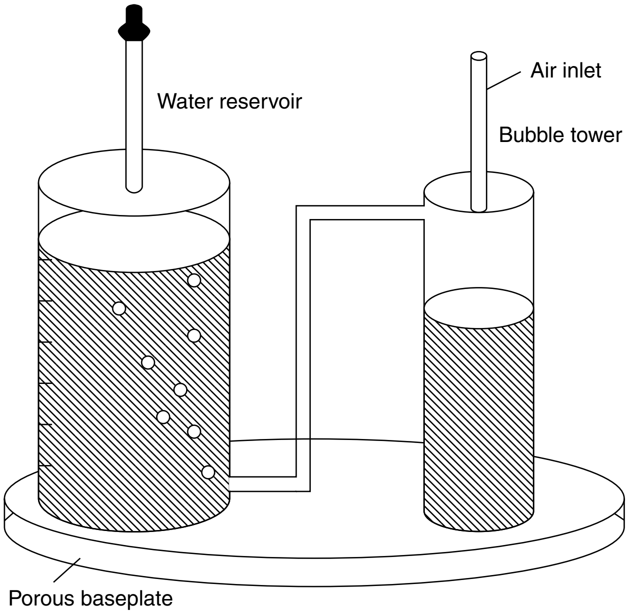

Figure 1. Tension infiltrometer. Tension measurements may be made using a handheld tensiometer on nongraduated models or on the side of graduated models (pictured).

The main advantage of disk infiltrometers is that *in situ* measurements of vadose zone hydraulic properties are rapid, reliable, and replicable and the devices are inexpensive and easy to operate. There exists, however, a limited range of tensions (generally 0 to 25 cm below the ground surface (BGS) in which the infiltrometers can operate, and measurements can be made only at the soil surface unless trenches are dug to the desired measurement level and the device used (1–3). Ideal soil properties are also assumed (soil is uniform, homogenous, and nonswelling), and pore closure may result from the weight of the device if used on freshly cultivated soils. There are other limitations that may be observed in Hussen and Warrick (4).

Neutron Moisture Probe

Initially designed for the agricultural industry, the neutron moisture probe has been adapted to environmental monitoring to obtain soil moisture levels. It is a nondestructive method of obtaining soil moisture levels when the probe is inserted in a well. The only restriction to depth of use is the length of the well used and the probe cord length.

Once lowered down a well, the probe begins to emit neutrons. The velocity at which neutrons return to the probe is correlative to the substance by which they were reflected. Neutrons reflected by heavy atoms of the soil matrix return with higher velocities compared to those of neutrons reflected by lighter hydrogen atoms. Hydrogen may be found in organic matter or in mineral lattices, but water is the greatest source of hydrogen in the subsurface. Due to this, the counts of reflected low-velocity neutrons are correlative to soil moisture content. A complete moisture stratification along the length of a well may be obtained, given a well-calibrated instrument.

Attributes of the neutron probe include ease of use and high accuracy, given a well-calibrated probe. It may be operated by one person, and soil moisture changes of the order of 1–5 volume % may be measured. Water flow pathways may also be located with the instrument. The largest drawback of the neutron moisture probes lies with calibration difficulties that may be extreme. Improper neutron probe calibration nearly renders the probe useless. Ideally, the probe is calibrated using soils collected in a drum with a known quantity of water from the field site of interest. Another drawback is that neutron probes cannot distinguish between water and other substances with similar hydrogen densities in the subsurface, such as gasoline. Problems of applying single calibrations to multiple media types also arise, which is the case in heterogeneous soils. Inaccuracies may also be result from improper well packing and the presence of certain elements, such as migrating boron or chlorine. Finally, care should be taken to mitigate exposure to the neutron source (5–8).

Psychrometers

Psychrometers represent another tool adapted from the agricultural industry for environmental monitoring. Estimation of the hydraulic gradient using the matric suction, osmotic potential, and water vapor potential are possible.

The generic class of instruments used to measure relative humidity using both wet and dry bulb thermometers are called psychrometers; they may be used for both surface (agriculture) and subsurface (environmental) applications. For the latter, they are placed in a small diameter well and surrounded by silica flour or air-filled packers. A dry bulb thermometer is employed to measure the ambient temperature, and a wet bulb thermometer is then used to measure the temperature decrease resulting from fluid evaporation in the subsurface (evaporative cooling). Liquid-filled thermometers are not accurate enough for the precision required, however, and thus more precise thermocouple thermometers are employed. The junction of two wires, a thermocouple, generates an electrical current that is a function of the junction temperature. The temperature of the thermocouple determines the activity of the water surrounding it, which is then correlated with the evaporation rate.

The direction and magnitude of water and contaminant migration may be observed easily and inexpensively with psychrometers, but regular calibration and monitoring of the thermocouple are necessary as temperature deviations of $\pm 0.001\,^\circ\mathrm{C}$ between the thermocouple and the liquid phase can introduce significant errors. Regular calibration is necessary due to corrosion of metals and the accumulation of salts on the thermocouple junction. Also, psychrometer equations assume that water is the first liquid to condense, and, at a site containing volatile liquids, this may not always be true (9,10).

Tensiometry

Tensiometers measure the force which holds water to the soil, or the soil–water potential, at shallow depths (0 to 8 m BGS). Because water moves from areas of high potential energy to low, evaluation of water's energy may be used to estimate the direction of water flow and the direction of contaminant plume migration. Rapid measurements at different locations may be obtained with tensiometers, which may then be plotted to obtain a spatial distribution of water and an estimation of the water flow direction.

A tensiometer measures matric potential using a porous ceramic cup placed in the soil, a connecting tube, and a vacuum gauge. The cup and tube are filled with water to provide hydraulic contact from the tensiometer to the water in the soil. Once placed in the ground, water flows out of the cup and into the soil until equilibrium is reached. Because the device is sealed from the atmosphere, the drop of the water level in the tube creates a vacuum which is measured by a vacuum gauge or pressure transducer from the surface.

Tensiomenters represent a low-cost, simple, and accurate measurement of matric potential (11–13). It is one of the most widely used devices for monitoring soil moisture levels in the vadose zone, though tensiometry values are limited to tension values below 0.85 bar at atmospheric pressures near 1 bar. For field applications, air coming out of solution from water or air trapped in the tube are the major shortcomings. During installation, care must be taken to ensure good hydraulic contact between the cup and the soil.

Time-Domain Reflectometry

Time-domain reflectometry (TDR) allows shallow, rapid, *in situ* measurements of water content and electrical conductivity in porous media (14–17). It is commonly used to locate underground cables. It is a relatively new development in vadose zone monitoring that was first described by Topp et al. (18). Accuracy within 1–2% of volumetric water content is feasible with TDR.

TDR uses a probe that is inserted into the soil equipped with 2 or 3 metallic rods that transmit electromagnetic pulses. The travel time of these pulses across the rods through the soil is used to compute the soil's bulk dielectric constant from which the water content is inferred, which provides an average soil water content over the depth of probe insertion.

The main advantage of TDR is that simple and rapid accumulation of shallow soil moisture content with high volumetric accuracy is possible. Calibration requirements are minimal compared to other indirect monitoring methods, and measurements for spatial and temporal analyses are easily obtained. The drawbacks associated with TDR include temperature-dependent readings and probe insertion in rocky or hard substrates because readings are limited to depth of insertion. Additionally, water molecules bound by interfacial forces in high surface area substances, for example, clays, result in lowered bulk dielectric constant compared to other soils at similar water content.

Ground-Penetrating Radar

Ground-penetrating radar (GPR) is an adaptation of seismic reflection for shallow use that images subsurface features by deploying radar waves. It may be used to image any homogeneity in the vadose zone, including contaminant plumes and soil stratification.

GPR is used to observe the electromagnetic impedance of soils using high-frequency radar waves to measure subsurface properties. It consists of an enclosed wave emitter and antenna unit that is dragged across the ground, generally in a grid pattern. Wave pulses are emitted and when an inhomogeneity is encountered, part of the incident energy is reflected back to the radar antenna within the unit. The reflected signal is amplified and transformed into a viewable image. The data are in visual form, not a quantified form, so results must be interpreted.

Any inhomogeneity in the vadose zone can be imaged with GPR. For vadose zone monitoring, contaminant plumes of both light nonaqueous phase liquids (LNAPLs) and dense nonaqueous phase liquids (DNAPLs), as well as buried utility lines and pipes may be imaged with GPR. Additionally, both underground storage tanks (UST) and leaking underground storage tanks (LUST) may be imaged and discerned. The most obvious drawback to GPR, however, is the fact that all data must be interpreted from visual images. Heterogeneities are viewed as anomalies in imaged data, which must be deciphered according to available information on subsurface conditions. Due to this fact, GPR benefits from use in combination with other vadose zone monitoring techniques to avoid misinterpretations (19,20).

Electrical Resistance Blocks

Electrical resistance blocks have been used in agriculture for more than 50 years to measure soil moisture in dry areas. In the environmental field, they are employed where other methods, including tensiometers, lysimeters, and manometers, are not operable due to low soil–water tension, generally between 0.5 and 15 bars.

The method consists of two or more electrodes that are placed in a porous block, commonly made of gypsum. The block is buried in the soil and allowed to remain until moisture equilibrium with the native soil is achieved. Changes in the electrical properties of the block, measured with the electrodes, reflect changes in the water content, which are measured with a wheatstone bridge resistance meter.

The benefits of this technique include low-cost, extensive historical use in agriculture research, and simplicity of use. It is generally useful for approximate soil moisture level changes. The disadvantages include limited historical use in the environmental industry and measurement and calibration difficulties at wet soil moisture potentials. Standard calibration curves exist, but as in many indirect techniques, soil-specific calibration of the instrument is recommended. Other limitations are associated with dissolution of the gypsum block and salinity changes in vadose zone water.

DIRECT MONITORING METHODS

Direct monitoring methods, in contrast to indirect monitoring methods, employ sampling of the vadose zone for liquid, soil, or gas. Obtaining liquid samples is fairly straightforward, using the most widely practiced method, suction lysimeters, but techniques for soil and gas sample acquisition are generally more complicated. Soil sampling commonly involves trenching or digging of some sort, which may be subject to Occupational Safety and Health Administration (OSHA) guidelines (24). Trenches deeper than 5 feet are subject to OSHA regulations and diggings deeper than 20 feet require the approval of a professional engineer. Soils are also subject to maximum allowable slope protocol, depending on their nature. In addition, certain cases may require confined space, heavy machinery, or decontamination permits. Readers are referred to OSHA literature for proper digging protocol. In gas sampling, problems stem from the nature of the phase being sampled, and, to a lesser extent, safety concerns. Vadose zone gas readily exchanges with any other gas with which it comes in contact, and thus samplers must be fully enclosed. Special safety procedures are required when considering volatile substances, such as methane.

Additionally, direct monitoring methods have a much higher risk of introducing contamination into the aquifer or exposing workers to contamination, and thus proper field procedures are absolutely necessary. In contaminated areas, all equipment must be constructed of nonreactive and low sorbing materials and be decontaminated between each use (25–27).

Direct Liquid Monitoring Techniques

Directly monitoring the liquid phase of the vadose zone involves obtaining a representative sample, whether in the form of water or NAPL. These may be important for soil–water chemistry analyses or contaminant plume delineation, among others. The primary method of collection is by using suction lysimeters placed into unconsolidated soil. As a suction device, lysimeters rely on applying a vacuum greater than the soil–water tension to draw a water sample up to the surface. A porous cup is inserted into the soil or in a shallow well, and a vacuum tube connects it to the surface. Once the vacuum is applied, water is drawn to the surface and a sample is obtained. Outside of lysimeters, other methods involve driving devices into unconsolidated sediment or lowering devices down wells above the water table to acquire liquid samples. Lysimeters have the advantage of being relatively cheap and simple to install. Virtually no maintenance is required, although sample volumes may be small.

Described by Gardner (28), wavelengths of light from gamma to X ray may be used to monitor the water content of the vadose zone by direct soil sampling. In practice, however, only gamma rays are employed. In this method, a sample is collected and a narrow beam of gamma radiation is sent through it. A recorder on the opposite side of the sample records only those rays that passed freely through the sample with no reflection along the way. By this method, the volumetric water content can be measured (29).

Direct Soil Monitoring Techniques

Due to the sorbing nature of many contaminants, soil samples are used to monitor the spatial and temporal distributions of contaminants in the soil. For certain applications, undisturbed samples are desired, in which case special precautions must be taken, and certain sampling tools and techniques must be employed. Generally, soil sampling involves inserting a device into the ground to recover a soil sample, although more intensive methods may be necessary in certain situations. Hand-operated samplers are typically used for shallow sampling and mechanically powered devices for deeper sampling. There are five basic hand-operated samplers, as described by Dorrance et al. (30), that include screw augers, barrel augers, tube samplers, hand-powered augers, and bulk samplers (shovels, etc). Each type has its own pros and cons depending on cost, ease of use, availability, size of matrix grains, degree of sediment cohesiveness, and other factors. Due to these, sampler choice is often site specific. For deeper monitoring applications, mechanically powered tools such as a drill rig must be employed.

Direct Vadose Zone Gas Monitoring

Monitoring the gas phase of the vadose zone is important because many contaminants volatilize to a significant degree. Unlike liquid and soil phases, the gaseous component of the vadose zone is problematic because atmospheric air and foreign vadose zone gases readily exchange and react upon exposure. When sampling, all gaseous vadose zone samples must avoid exposure to foreign air, and all sampling methods must be designed to avoid it. *In situ* sampling and devices with internal liners designed to seal samples from the atmosphere are preferred (31).

Monitoring through sampling the gas phase typically involves hydraulically ramming a specially designed device into the soil where it collects samples via a vacuum and pumps them to the surface for collection. Passive gas sampling by an *in situ* absorbent, such as an activated carbon trap placed in the vadose zone, may be used, but less information is provided by this method. Direct sampling of gases is more desirable, but it is more difficult and costly. An alternative is monitoring gas advection by using groundwater tracers. A tracer that partitions to the gas phase is injected, samples are taken at other sites, typically at wells, and analyzed for the tracer concentration (32–35).

BIBLIOGRAPHY

1. Ankeny, M.M., Ahed, M., Kaspar, T.C., and Horton, R. (1991). Simple method for determining unsaturated hydraulic conductivity. *Soil Sci. Soc. Am. J.* **55**: 467–470.
2. Lin, H.S. and McInnes, K.J. (1995). Water-flow in clay soil beneath a tension infiltrometer. *Soil Sci.* **159**(6): 375–382.
3. Reynolds, W.D. and Elrick, D.E. (1991). Determination of hydraulic conductivity using a tension infiltrometer. *Soil Sci. Soc. Am. J.* **55**: 633–639.
4. Hussen, A.A. and Warrick, A.W. (1993). Algebraic models for disc tension permeameters. *Water Resour. Res.* **29**: 2779–2786.
5. Amoozegar, A., Martin, K.C., and Hoover, M.T. (1989). Effect of access hole properties on soil water content determination by neutron thermalization. *Soil Sci. Soc. Am. J.* **53**: 330–335.
6. Arslan, A., Razzouk, A.K., and Alain, F. (1997). The performance and radiation exposure of some neutron probes in measuring the water content of the topsoil layer. *Aust. J. Soil Res.* **35**(6): 1397–1407.
7. Dillel, D. (1982). *Introduction to Soil Physics*. Academic Press, London.
8. Grismer, M.E., Bali, K.M., and Robinson, F.E. (1995). Field scale neutron probe calibration and variable analysis for clay soil. *J. Irrig. Drainage Eng.—ASCE* **121**(5): 354–362.
9. Comstock, J.P. (2000). Correction of thermocouple psychrometer readings for the interaction of temperature and actual water potential. *Crop Sci.* **40**(3): 709–712.
10. Rawlins, S.L. and Campbell, G.S. (1986). Water potential: Thermocouple psychrometry. In: *Methods of Soil Analysis*. A. Klute (Ed.). Part 1, 2nd Edn. pp. 597–618.
11. Nyhan, J.W. and Drennon, B.J. (1990). Tensiometer data acquisition system for hydrologic studies requiring high temporal resolution. *Soil Sci. Soc. Am. J.* **54**(1): 293–296.
12. Williams, T.H.L. (1978). Automatic scanning and recording tensiometer system. *J. Hydrol.* **39**(1–2): 175–183.
13. Whalley, W.R., Leedsharrison, P.B., Joy, P., and Hoefsloot, P. (1994). Time-domain reflectometry and tensiometry combined in an integrated soil–water monitoring system. *J. Agric. Eng. Res.* **59**(2): 141–144.
14. Jones, S.B., Wraith, J.M., and Or., D. (2002). Time domain reflectometry measurement principles and applications. *Hydrologic Processes* **16**(1): 141–153.
15. Long, D.S., Wraith, J.M., and Kegel, G. (2002). A heavy-duty time domain reflectometry soil moisture probe for use in intensive field sampling. *Soil Sci. Soc. Am. J.* **66**(2): 396–401.
16. Lee, J., Horton, R., and Jaynes, D.B. (2002). The feasibility of shallow time domain reflectometry probes to describe solute transport through undisturbed soil cores. *Soil Sci. Soc. Am. J.* **66**(1): 53–57.
17. Noborio, K. (2001). Measurement of soil water content and electrical conductivity by time domain reflectometry: A review. *Comput. Electron. Agric.* **31**: 213–237.
18. Topp, G.C., Davis, J.L., and Annan, A.P. (1980). Electromagnetic determination of soil–water content: measurement in coaxial transmission lines. *Water Resour. Res.* **16**: 574.
19. Daniels, J.J., Roberts, R., and Vendl, M. (1995). Ground penetrating radar for detection of liquid contaminants. *Appl. Geophys.* **33**: 195–207.
20. Weil, G.J., Graf, R.J., and Leann, M.F. (1994). Investigations of hazardous waste sites using thermal IR and ground penetrating radar. *Photogrammetric Eng. Remote Sensing* **60**(8): 999–1005.
21. Everett, L.G., Wilson, G.G., and Hoylman, E.W. (1984). *Vadose Zone Monitoring for Hazardous Waste Sites*. Noyes, Park Ridge, NJ.
22. Fishback, P.E. (1977). Scheduling irrigations by electrical resistance blocks. *NebGuide* G 77–340, Institute of Agriculture and Natural Resources, University of Nebraska, Lincoln, NE.

23. Strangeways, I.C. (1983). Interfacing soil moisture gypsum blocks with a modern data-logging system using a simple, low cost, dc method. *Soil Sci.* **136**: 322.
24. http://www.osha.gov/ (Occupational Safety and Health Administration).
25. Fetter, C.W. (1999). *Contaminant Hydrogeology*, 2nd Edn. Prentice-Hall, Upper Saddle River, NJ.
26. Hudak, P.F. and Loaiciga, H.A. (1999). Conjunctive vadose and saturated zone monitoring for subsurface contamination. *Environ. Monitoring Assessment* **59**(1): 15–29.
27. Mickam, J.E., Bellandi, R., and Tifft, E.C., Jr. (1989). Equipment decontamination procedures for ground water and vadose zone monitoring programs: Status and prospects. *Ground Water Monitoring Rev.* **9**(2): 100–121.
28. Gardner, W.H. (1986). Water content. In: *Methods of Soil Analysis. Part 1.* A. Klute (Ed.). Monograph 9, American Society of Agronomy, Madison, WI.
29. Wilson, L.G. (1990). Methods for sampling fluids in the vadose zone. In: *Ground Water and Vadose Zone Monitoring*. ASTM STP 1053. D.M. Nielsen and Z.I. Johnson (Eds.) American Society for Testing and Materials, Philadelphia, PA, pp. 7–24.
30. Dorrance, D.W., Wilson, L.G., Everett, L.G., and Cullen, S.J. (1994). A compendium of soil samplers for the vadose zone. In: *Handbook of Vadose Zone Characterization and Monitoring*. L.G. Wilson, G.E. Lorne, and S.J. Cullen (Eds.). CRC Press, Boca Raton, FL, pp. 401–428.
31. Lewis, T.E., Crockett, A.B., Siegrist, R.L., and Zarrabi, K. (1991). Soil sampling and analysis for volatile organic compounds. E.P.A. *Ground Water Forum Issue*, EPA/540/4-91-001.
32. Fetter, C.W. (1999). *Contaminant Hydrogeology*, 2nd Edn. Prentice-Hall, Upper Saddle River, NJ.
33. Hubbell, J.M., Wood, T.R., Higgs, B., Wylie, A.H., and McElroy, D.L. (1998). Design, installation, and uses of combination groundwater and gas sampling wells. *Ground Water Monitoring and Remediation* **18**(3): 151–157.
34. Marrin, D.L. (1988). Soil-gas sampling and misinterpretation. *Ground Water Monitoring Rev.* **8**(2): 51–54.
35. Thompson, G. and Marrin, D. (1987). Soil gas contaminant investigations: A dynamic approach. *Ground Water Monitoring Rev.* **8**(4): 88–93.

VAPOR TRANSPORT IN THE UNSATURATED ZONE

Fred D. Tillman
U.S. Environmental Protection Agency
Athens, Georgia

James A. Smith
University of Virginia
Charlottesville, Virginia

INTRODUCTION

The term "vapor" describes the gaseous phase of a substance that is normally a liquid or solid at standard atmospheric pressure and temperature. Examples of vapors that are important in environmental systems include humidity (water vapor), gasoline vapor, and solvent vapor. Soil and groundwater that are contaminated by organic solvents and fuel hydrocarbons may emit vapors that can move through the subsurface and reach the land surface. Human exposure to these harmful vapors has become an increasingly important topic in the evaluation and remediation of contaminated soil and groundwater. The potential for inhalation of vapors transported from a groundwater plume to the interior of buildings often controls the outcome of a risk based corrective action (RBCA) analysis (1). Additionally, vapors emanating from subsurface contamination contribute to the reduction of contaminant mass and may provide a beneficial mechanism of natural attenuation at sites where there are no human receptors. Subsurface processes affecting the transport of organic vapors through the soil above the water table (here defined as the unsaturated zone) may slow, reduce, or enhance the release of these vapors to potential receptors at the land surface. A careful investigation of organic vapor transport in the unsaturated zone is therefore important in assessing the potential for natural attenuation and the impact of even deep contamination on receptors at the land surface. This article describes potential sources, fate and transport, and measurement of organic vapors in the unsaturated zone. Although this article deals exclusively with organic vapors, many of the fate and transport properties discussed are also applicable to other vapors, including metals and moisture.

SOURCES OF ORGANIC VAPORS

Contaminants of concern in vapor transport in the unsaturated zone are typically volatile organic compounds (VOCs), although vapors emanating from inorganic sources such as mercury vapor may be of concern as well. A chemical is considered volatile if its Henry's law constant is 1×10^{-5} atm-m^3/mol or greater (2). Examples of VOCs that are important in impacted environmental systems include chlorinated solvents such as carbon tetrachloride, tetrachloroethylene, and trichloroethylene (and their degradation compounds), fuel hydrocarbons such as benzene, toluene, ethylbenzene, and *o*,*m*,*p*-xylenes as well as volatile pesticides such as chlordane, aldrin, and lindane. The U.S. Environmental Protection Agency lists 107 compounds whose toxicity and volatility produce a potentially unacceptable inhalation risk to receptors (2). These VOCs can be released into the subsurface environment from leaking landfill liners, improper disposal, accidental spillage, or leaking storage tanks (LSTs). Once in the subsurface, these compounds can become bound to the soil matrix, dissolved in groundwater (or soil water), and/or exist as a separate, residual phase known as a nonaqueous phase liquid (NAPL). Soil, aqueous, and NAPL-phase organics may all be sources of organic vapors in the subsurface. Therefore, organic vapor transport in the unsaturated zone requires understanding of interphase mass-transfer processes because the contaminant can be distributed among soil gas, water, soil, and NAPL phases. In modeling organic vapor transport, the equilibrium vapor concentration at the source of contamination is often estimated from the following relationships (3):

Soil contamination:

$$C_{\text{source}} = \frac{H_{\text{TS}} C_{\text{R}} \rho_{\text{b}}}{\theta_{\text{w}} + K_{\text{d}} \rho_{\text{b}} + H_{\text{TS}} \theta_{\text{a}}} \quad (4) \qquad (1)$$

Groundwater contamination:

$$C_{\text{source}} = H_{\text{TS}} C_{\text{w}} \qquad \text{Henry's Law} \qquad (2)$$

NAPL contamination:

$$C_{\text{source}} = \frac{x_i p_i \text{MW}_i}{RT} \qquad \text{Raoult's Law} \qquad (3)$$

where C_{source} is the vapor concentration at the source of contamination (ML^{-3}), H_{TS} is the dimensionless Henry's law constant at the system temperature (unitless), C_{R} is the soil concentration (MM^{-1}), ρ_{b} is the soil dry bulk density (ML^{-3}), θ_{w} is the soil water-filled porosity (L^3L^{-3}), θ_{a} is the soil air-filled porosity (L^3L^{-3}), K_{d} is the soil–water partition coefficient (L^3M^{-1}), C_{w} is concentration of the contaminant in the groundwater (ML^{-3}), x_i is the mole fraction of component i in the NAPL mixture (unitless), p_i is the vapor pressure of component i at the soil temperature ($\text{ML}^{-1}\text{T}^{-2}$), MW_i is the molecular weight of component i (M), R is the ideal gas constant ($\text{ML}^2\text{T}^{-2}\text{T}^{-1}\text{M}^{-1}$, e.g., 8.314570 $\text{J}^{\circ}\text{K}^{-1}\ \text{mol}^{-1}$), and T is the soil temperature (T). Both Henry's law constant and vapor pressure rise with temperature for organic compounds of interest, so a higher source vapor concentration is expected from higher soil and groundwater temperature.

TRANSPORT AND FATE OF ORGANIC VAPORS IN THE UNSATURATED ZONE

Organic vapors emanating from contaminated soil or groundwater or from a residual phase such as gasoline floating on the water table may move through unsaturated zone soil gas by diffusion or soil-gas advection due to pressure or density gradients or a combination of these processes.

Diffusion

Molecular diffusion is the spreading out of compounds from random collisions resulting from thermal motion of atoms. These collisions may be between molecules themselves or between molecules and their surroundings. Under most environmental conditions, molecular diffusion in natural systems proceeds from locations of higher concentration toward locations of lower concentrations. In a typical scenario, organic vapors above a contaminated water table (high concentration) diffuse toward the land surface (lower concentration). The well-known relation describing the diffusion of a compound across a unit of cross-sectional area is Fick's first law:

$$F_{\text{x}} = -D^{\text{eff}} \frac{dC}{dx} \qquad (4)$$

where F_{x} is the mass flux [$\text{ML}^{-2}\text{T}^{-1}$], D^{eff} is the effective diffusion coefficient of the compound in the gas phase [L^2T^{-1}], and dC/dx is the concentration gradient of the compound in one dimension [$\text{ML}^{-3}\text{L}^{-1}$]. From this equation, it is apparent that the rate of molecular diffusion in the gas phase depends on the concentration gradient and the effective diffusion coefficient of the compound of interest. Several relations relate the effective diffusion coefficient to the free-air diffusion coefficient of the compound and the total and air-filled porosities of the diffusing media (5–14). An increase in diffusive flux is seen in soils with greater air passageways (i.e., greater porosities and air-filled porosities). Therefore, in a layered unsaturated zone, vapor diffusion from depth to the land surface is limited by the wettest, least porous soil layer. Because free-air diffusion coefficients are compound dependent, ranging from $2.5 \times 10^{-3}\ \text{cm}^2\ \text{s}^{-1}$ for hexachloroethane to $2.71 \times 10^{-1}\ \text{cm}^2\ \text{s}^{-1}$ for chloroethane among the 107 volatile compounds of concern to the U.S.EPA (2), different chemicals diffuse at different rates under the same concentration gradients. Increased temperature also produces an increased free-air diffusion coefficient, leading to a greater rate of diffusion relative to the same system at lower temperatures. Fick's First Law may not sufficiently describe vapor diffusion in systems where pore sizes are very small (Knudsen diffusion) and when volatile species constitute a substantial fraction of the total soil-gas concentration (nonequimolar diffusion) (15–17). In these situations, it may be necessary to employ the Stefan–Maxwell equation.

Advection

The flow of soil gas (advection) in the subsurface may be caused by gas-pressure gradients or, in certain cases, gas density gradients. Pressure-driven advection is produced when differences in soil-gas pressure form, causing soil gas to flow and carrying any vapors present with it. Air pressure gradients in the subsurface of natural systems may result from several phenomena. As diurnal or weather related atmospheric pressure cycles occur at the land surface, pressure waves are transmitted into the unsaturated zone and air may flow in response—a process known as "barometric pumping." Barometric pumping may cause soil gas to flow either toward the land surface carrying soil vapor or away from the land surface bringing in fresh atmospheric air (13,15,18–25). The underpressurization of an overlying building will produce gas pressure differences in subsurface soils. This underpressurization may be caused by thermal differences between indoor and outdoor air (stack effects), wind loading on the building superstructure, and unbalanced building ventilation (26,27). Soil-gas pressure gradients may also be produced by a rapidly rising or falling water table, as in coastal zones (28), or through the buildup of gas pressure from decomposing organic matter inside a landfill (29). Finally, natural temperature differences between warmer deep and cooler shallow soil gas will cause soil gas to rise (30,31). Density-driven flow of organic vapors may occur in the vicinity of residual-phase organic compounds whose saturation gas densities are greater than that of air. As organic liquids with high vapor pressures and molecular weights volatilize, the density of the soil gas surrounding the liquid changes. In almost all cases, organic liquids have molecular weights that are greater than air, so the resulting density-driven flow is in a downward direction and is proportional to soil permeability and density differences between the vapor and air (32–34). Organic compounds for which density-driven advection may be significant include methylene chloride, 1,2-dichloroethylene, 1,1,1-trichloroethane, carbon tetrachloride, and 1,1-dichloroethane, among others (32–34).

Sorption

As organic vapors move through the unsaturated zone by diffusion and advection, they come in contact with soil moisture, infiltrating rainwater, and the soil matrix itself. Each of these interactions may affect the concentration of the contaminant in the soil gas. Depending on the compound, organic vapors may adsorb to soil grain surfaces or partition into soil organic matter (35–40). Adsorption of relatively nonpolar organic vapors is suppressed by high humidity in the subsurface because as polar water molecules can effectively out-compete organic vapors for mineral-surface adsorption sites (41,42). For these high-humidity conditions, sorption may be limited to organic vapor partition into soil organic matter (41,42). Soil moisture trapped in unsaturated-zone pore space or infiltrating rain water may also sorb organic vapors to varying degrees (43,44). Finally, gas-phase organic compounds may adsorb to the air–water interface in unsaturated zones (45,46). Each of these sorption processes may act as both a source and a sink of organic vapors in the unsaturated zone.

Biodegradation

Under favorable conditions, organic vapors in the unsaturated zone that partition into soil moisture may be biodegraded by indigenous bacteria in oxidation/reduction reactions. The aerobic biodegradation of petroleum hydrocarbons requires an abundant oxygen supply as well as necessary nutrients of nitrogen and phosphorus (47–51). When oxygen is depleted, other possible electron acceptors for biodegradation of petroleum hydrocarbons include nitrate (NO_3^-), iron oxides (e.g., $Fe(OH)_3$), sulfate (SO_4^{2-}), and carbon dioxide (CO_2) (48). Lightly chlorinated compounds (e.g., chlorobenzene, dichlorobenzene) may be biodegraded under aerobic conditions. The more highly chlorinated hydrocarbons are recalcitrant to aerobic biodegradation but may undergo direct or cometabolic anaerobic reductive dechlorination. Reductive dechlorination has been observed to be most effective under sulfate-reducing and methanogenic conditions (52). In direct reductive dechlorination, the chlorinated hydrocarbon is used as an electron acceptor and the bacteria gain energy and grow as a result of the reaction (53). In cometabolic reductive dechlorination, enzymes produced during microbial metabolism of another hydrocarbon fortuitously reduce and dehalogenate the chlorinated contaminant. For either reductive dechlorination process to be successful, a primary substrate (electron donor) such as soil organic matter, lactate, acetate, methanol, or glucose is necessary (54).

VAPOR SAMPLING AND FLUX MEASUREMENT

Organic vapors in the unsaturated zone above contaminated soil or groundwater are typically sampled to provide evidence of biodegradation (49,55), calibrate vapor flow and transport models for estimating natural attenuation (23), or to estimate the region of contaminated groundwater (42,56,57). Soil gas is often sampled through low volume, inert pipe or tubing penetrating the subsurface to the desired sampling depth (42,44,49,56,58,59). This tubing can be installed with the help of a Geoprobe®-type device or inserted in a previously drilled borehole. The sampling portion of the tubing must be sealed from the atmosphere and other sampling depths to prevent preferential flow down the borehole during sampling. Samples may be collected from these vapor probes in various ways, including direct syringe withdrawal (49), pumping of soil gas into Tedlar® bags (57,60,61), or pumping of soil gas through adsorbent traps (58,62,63). Analyses of samples collected by syringe withdrawal require on-site analysis. Samples collected in Tedlar® bags or onto adsorbent traps may be transported to an off-site laboratory for analysis. Alternatively, there are field-portable gas detectors and gas chromatographs with flame-ionization and photoionization detectors capable of measuring real time hydrocarbon concentrations (47,57,64).

Efforts to measure the flux of VOCs from the unsaturated zone to the atmosphere at the land surface have included using the aerodynamic method, which necessitates the measurement of trace concentrations and atmospheric conditions such as wind speed, temperature and pressure (65); using groundsheets sampled through a nylon sampling tube (66); and several chamber designs (58,59,67–71). Surface measurements of vapor flux are important because they provide a direct measurement of total natural flux that can be used to estimate volatilization rates for a field site. This natural attenuation process for contaminants in soil and groundwater may be significant, and, at times, approach the removal rate of active remediation systems (58). Complexities involved in the measurement of vapor transport from the subsurface to the atmosphere include possible preferential flow paths (e.g., underground utilities, macropores, etc.) and spatial and temporal variability in vapor concentrations. Spatial variability may be produced by subsurface heterogeneities that are geologic as well as constructed barriers (e.g., storm drains, basement foundations). Temporal variability in vapor concentrations is produced by changing soil moisture content, temperature, and gas-pressure gradients.

SUMMARY

Analysis of the transport of organic vapors in the unsaturated zone is fundamental in evaluating potential human exposure to harmful compounds from a contaminated subsurface as well as determining a site's potential for natural attenuation. Analysis of soil gas vapors may also aid in estimating the areal extent of a contaminated plume. Soil, aqueous, and NAPL-phase organics may all be sources of organic vapors in the subsurface, and interphase mass transfer to and from soil, soil moisture, groundwater, and soil gas all play a role in vapor transport in the unsaturated zone. Diffusion of the vapor phase contaminant proceeds from higher concentrations near the source toward lower concentrations at the land surface. Advection of soil gas results from pressure gradients in the subsurface and from density-driven flow for certain high molecular weight compounds. Under favorable conditions, organic vapors in the unsaturated zone may be biodegraded by indigenous bacteria in oxidation/reduction reactions. All biodegradation reactions of organic compounds require nutrients and electron acceptors in sufficient, usable quantities. Soil-gas concentrations may be quantified through the use of vapor probes, adsorbent traps, Tedlar® bags, direct syringe withdrawal with field-portable analytical equipment, and the use of gas chromatography. The total flux of contaminants from the subsurface to the land surface may be estimated by the aerodynamic method, through the use of ground sheets, or with flux chambers. Surface flux measurement is useful for estimating the total volatilization of contaminants from a site, but all vapor-phase measurements are complicated by spatial and temporal variability.

BIBLIOGRAPHY

1. Laubacher, R.C., Bartholomae, P., Velasco, P., and Reisinger, H.J. (1997). An evaluation of the vapor profile in the vadose zone above a gasoline plume. In: *Pet. Hydrocarbons Org. Chem. Ground Water: Prev. Detection Remediation Conf.*, Houston, TX, pp. 396–409.
2. Environmental Quality Management. (2003). *User's Guide for Evaluating Subsurface Vapor Intrusion into Buildings*. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Durham, NC.
3. U.S. Environmental Protection Agency. (2002). *Draft Guidance for Evaluating the Vapor Intrusion to Indoor Air Pathway from Groundwater And Soils*. USEPA Office of Solid Waste and Emergency Response.
4. Johnson et al. (1990).
5. Buckingham, E. (1904). *Contribution to Our Knowledge of the Aeration of Soils*. USDA Bureau of Soils Bulletin No. 25.
6. Penman, H.L. (1940). Gas and vapor movements in soil I. The diffusion of vapours through porous solids. *J. Agric. Sci.* **30**: 463–462.
7. Van Bavel, C.H.M. (1952). Gaseous diffusion and porosity in porous media. *Soil Sci.* **73**: 91–105.
8. Rust, R.H., Klute, A., and Geiseking, G.E. (1957). Diffusion-porosity measurements using a non-steady state system. *Soil Sci.* **84**: 453–465.
9. Dye, R.F. and Dallavale, J.M. (1958). Diffusion of gases in porous media. *Ind. Eng. Chem.* **50**: 1195–1200.
10. Millington, R.J. (1959). Gas diffusion in porous media. *Science* **130**: 100–102.
11. Currie, J.A. (1970). Movement of gases in soil respiration. *Sorption and Transport Processes in Soils*. SCI Monogr. Ser. **37**: 152–171.
12. Nilson, R.H., Peterson, E.W., and Lie, K.H. (1991). Atmospheric pumping: A mechanism causing vertical transport of contaminated gases through fractured permeable media. *J. Geophys. Res.* **96**: 21933–21948.
13. Bartelt-Hunt, S.L. and Smith, J.A. (2002). Measurement of effective air diffusion coefficients for trichloroethylene in heterogeneous soil. *J. Contam. Hydrol.* **56**: 193–208.
14. Rolston, D.E. and Moldrup, P. (2002). Gas diffusivity. In: *Methods of Soil Analysis: Part 4-Physical Methods*. J.H. Dane and G.C. Topp (Eds.). Soil Science Society of America, Madison, WI, pp. 1113–1158.
15. Thorstenson, D.C. and Pollock, D.W. (1989). Gas transport in unsaturated zones: Multicomponent systems and the adequacy of Fick's laws. *Water Resour. Res.* **25**: 477–507.
16. Baehr, A.L. and Bruell, C.J. (1990). Application of the Stefan-Maxwell equations to determine limitations of Fick's law when modeling organic vapor transport in sand columns. *Water Resour. Res.* **26**: 1155–1163.
17. Partridge, G.P., McLeary, K.S., Showers, T.R., Huebner, R.S., and Elliott, H.A. (2002). Single and multicomponent gas phase diffusion in porous media: Modeling and laboratory measurements. *Soil Sediment Contam.* **11**: 555–581.
18. Sleep, B.E. and Sykes, J.F. (1989). Modeling the transport of volatile organics in variably saturated media. *Water Resour. Res.* **25**: 81–92.
19. Massmann, J. and Farrier, D.F. (1992). Effects of atmospheric pressures on gas transport in the vadose zone. *Water Resour. Res.* **28**: 777–791.
20. Auer, L.H., Rosenberg, N.D., Birdsell, K.H., and Whiteny, E.M. (1996). The effects of barometric pumping on contaminant transport. *J. Contam. Hydrol.* **24**: 145–166.
21. Elberling, B., Larsen, F., Christensen, S., and Postma, D. (1998). Gas transport in a confined unsaturated zone during atmospheric pressure cycles. *Water Resour. Res.* **34**: 2855–2862.
22. Tillman, F., Choi, J.C., Katchmark, W., Smith, J.A., and Wood, H.G.I. (2001). Unsaturated-zone airflow: Implications for natural remediation of ground water by contaminant transport through the subsurface. In: *Physicochemical Groundwater Remediation*. J.A. Smith and S.E. Burns (Eds.). Kluwer Academic/Plenum, New York, pp. 307–339.
23. Choi, J.W., Tillman, F.D., and Smith, J.A. (2002). Relative importance of gas-phase diffusive and advective trichloroethylene (TCE) fluxes in the unsaturated zone under natural conditions. *Environ. Sci. Technol.* **36**: 3157–3164.
24. Neeper, D.A. (2002). Investigation of the vadose zone using barometric pressure cycles. *J. Contam. Hydrol.* **54**: 59–80.
25. Rossabi, J. and Falta, R.W. (2002). Analytical solution for subsurface gas flow to a well induced by surface pressure fluctuations. *Ground Water* **40**: 67–75.
26. Nazaroff, W.W., Lewis, S.R., Doyle, S.M., Moed, B.A., and Nero, A.V. (1987). Experiments on pollutant transport from soil into residential basements by pressure-driven airflow. *Environ. Sci. Technol.* **21**: 459–466.
27. Garbesi, K. and Sextro, R.G. (1989). Modeling and field evidence of pressure-driven entry of soil gas into a house through permeable below-grade walls. *Environ. Sci. Technol.* **23**: 1481–1487.
28. Li, L., Dong, P., and Barry, D.A. (2002). Tide-induced water table fluctuations in coastal aquifers bounded by rhythmic shorelines. *J. Hydraul. Eng.* **128**: 925–933.
29. Little, J.C., Daisey, J.M., and Nazaroff, W.W. (1992). Transport of subsurface contaminants into buildings: An exposure pathway for volatile organics. *Environ. Sci. Technol.* **26**: 2058–2066.
30. Gustin, M.S., Taylor, G.E., and Maxey, R.A. (1997). Effect of temperature and air movement on the flux of elemental mercury from substrate to the atmosphere. *J. Geophys. Res.* **102**: 3891–3898.
31. Krylov, V.V. and Ferguson, C.C. (1998). Contamination of indoor air by toxic soil vapours: The effects of subfloor

ventilation and other protective measures. *Build. Environ.* **33**: 331–347.

32. Falta, R.W., Javandel, I., Pruess, K., and Witherspoon, P.A. (1989). Density-driven flow of gas in the unsaturated zone due to the evaporation of volatile organic compounds. *Water Resour. Res.* **25**: 2159–2169.
33. Mendoza, C.A. and Frind, E.O. (1990). Advective-dispersive transport of dense organic vapors in the unsaturated zone 1. Model development. *Water Resour. Res.* **26**: 379–387.
34. Mendoza, C.A. and Frind, E.O. (1990). Advective-dispersive transport of dense organic vapors in the unsaturated zone: 2. Sensitivity analysis. *Water Resour. Res.* **26**: 387–398.
35. Goss, K.U. (1994). Adsorption of organic vapors on polar mineral surfaces and on a bulk water surface: Development of an empirical predictive model. *Environ. Sci. Technol.* **28**: 640–645.
36. Goss, K.U. (1994). Effects of temperature and relative humidity on the sorption of vapors on quartz sand. *Environ. Sci. Technol.* **26**: 2287–2294.
37. Goss, K.-U. and Eisenreich, S.J. (1996). Adsorption of VOCs from the gas phase to different minerals and a mineral mixture. *Environ. Sci. Technol.* **30**: 2135–2142.
38. Popovicova, J. and Brusseau, M.L. (1998). Contaminant mass transfer during gas-phase transport in unsaturated porous media. *Water Resour. Res.* **34**: 83–92.
39. Ruiz, J., Bilbao, R., and Murillo, M. (1998). Adsorption of different VOC onto soil minerals from gas phase: Influence of mineral, type of VOC, and air humidity. *Environ. Sci. Technol.* **32**: 1079–1084.
40. Goss, K.-U., Buschmann, J., and Schwarzenbach, R.P. (2004). Adsorption of organic vapors to air-dry soils: Model predictions and experimental validation. *Environ. Sci. Technol.* **38**: 3667–3673.
41. Chiou, C.T. and Shoup, T.D. (1985). Soil sorption of organic vapors and effects of humidity on sorptive mechanism and capacity. *Environ. Sci. Technol.* **19**: 1196–1200.
42. Smith, J.A., Chiou, C.T., Kammer, J.A., and Kile, D.E. (1990). Effect of soil moisture on the sorption of trichloroethene vapor to vadose-zone soil at Picatinny Arsenal, New Jersey. *Environ. Sci. Technol.* **24**: 676–683.
43. Cho, H.J. and Jaffé, P.R. (1990). The volatilization of organic compounds in unsaturated porous media during infiltration. *J. Contam. Hydrol.* **6**: 387–410.
44. Cho, H.J., Jaffé, P.R., and Smith, J.A. (1993). Simulating the volatilization of solvents in unsaturated soils during laboratory and field infiltration experiments. *Water Resour. Res.* **29**: 3329–3342.
45. Kim, H., Rao, P.S., and Annable, M.D. (1997). Determination of effective air-water interfacial area in partially saturated porous media using surfactant adsorption. *Water Resour. Res.* **33**: 2705–2711.
46. Kim, H., Annable, M.D., and Rao, P.S.C. (1998). Influence of air-water interfacial adsorption and gas-phase partitioning on the transport of organic chemicals in unsaturated porous media. *Environ. Sci. Technol.* **32**: 1253–1259.
47. Ostendorf, D.W. and Kampbell, D.H. (1991). Biodegradation of hydrocarbon vapors in the unsaturated zone. *Water Resour. Res.* **27**: 453–462.
48. Norris, R.D., Hinchee, R.E., Brown, R., McCarty, P.L., Wilson, J.T., Reinhard, M., Bouwer, E.J., Borden, R.C., Vogel, T.M., Thomas, J.M., and Ward, C.H. (1994). *Handbook of Bioremediation*. Lewis, Boca Raton, FL.
49. Lahvis, M.A. and Baehr, A.L. (1996). Estimation of rates of aerobic hydrocarbon biodegradation by simulation of gas transport in the unsaturated zone. *Water Resour. Res.* **32**: 2231–2249.
50. Lahvis, M.A., Baehr, A.L., and Baker, R.J. (1999). Quantification of aerobic biodegradation and volatilization rates of gasoline hydrocarbons near the water table under natural attenuation conditions. *Water Resour. Res.* **35**: 753–765.
51. Hers, I., Atwater, J., Li, L., and Zapf-Gilje, R. (2000). Evaluation of vadose zone biodegradation of BTX vapours. *J. Contam. Hydrol.* **46**: 233–264.
52. U.S. Environmental Protection Agency. (2000). *Engineered approaches to in situ bioremediation of chlorinated solvents: Fundamentals and field applications*. EPA 542-R-00-008, Washington, DC.
53. McCarty, P.L. (1997). Breathing with chlorinated solvents. *Science* **276**: 1521–1522.
54. Vogel, T.M., Criddle, C.S., and McCarty, P.L. (1987). Transformations of halogenated aliphatic compounds. *Environ. Sci. Technol.* **21**: 722–736.
55. Roggemans, S., Bruce, C.L., Johnson, P.C., and Johnson, R.L. (2001). Vadose zone natural attenuation of hydrocarbon vapors: An empirical assessment of soil gas vertical profile data. *Am. Pet. Inst. Tech. Bull.* **15**: 12.
56. Kerfoot, H.B. (1987). Soil-gas measurement for detection of groundwater contamination by volatile organic compounds. *Environ. Sci. Technol.* **21**: 1022–1024.
57. Moseley, C.L. and Meyer, M.R. (1992). Petroleum Contamination of an elementary school: A case history involving air, soil-gas and groundwater monitoring. *Environ. Sci. Technol.* **26**: 185–192.
58. Smith, J.A., Tisdale, A.K., and Cho, H.J. (1996). Quantification of natural vapor fluxes of trichloroethene in the unsaturated zone at Picatinny Arsenal, New Jersey. *Environ. Sci. Technol.* **30**: 2243–2250.
59. Tillman, F., Choi, J.-W., and Smith, J.A. (2003). A comparison of direct measurement and model simulation of total flux of volatile organic compounds from the subsurface to the atmosphere under natural field conditions. *Water Resour. Res.* **39**: 1284–1294.
60. Fischer, M.L. et al. (1996). Factors affecting indoor air concentrations of volatile organic compounds at a site of subsurface gasoline contamination. *Environ. Sci. Technol.* **30**: 2948–2957.
61. Jeng, C.Y., Kremesec, V.J., and Primack, H.S. (1996). Models of hydrocarbon vapor diffusion through soil and transport into buildings. In: *Pet. Hydrocarbons Org. Chem. Ground Water: Prev. Detection Remediation Conf.* National Ground Water Association, Houston, TX, pp. 319–338.
62. Tillman, F. (2003). *Design and testing of a chamber device to measure total flux of volatile organic compounds from the unsaturated zone under natural conditions*. Ph.D. Dissertation. University of Virginia, Charlottesville, VA.
63. Hodgson, A.T., Garbesi, K., Sextro, R.G., and Daisey, J.M. (1992). Soil-gas contamination and entry of volatile organic compounds into a house near a landfill. *J. Air Waste Manage. Assoc.* **42**: 277–283.
64. Ririe, G.T., Sweeney, R.E., Daugherty, S.J., and Peuron, P.M. (1998). A vapor transport model that is consistent with field and laboratory data. *Pet. Hydrocarbons Org. Chem. Ground Water: Prev. Detection Remediation Conf.*, Ground Water Association, Houston, TX, pp. 299–309.
65. Chen, C., Green, R.E., Thomas, D.M., and Knuteson, J.A. (1995). Modeling 1,3-dichloropropene fumigant volatilization with vapor-phase advection in the soil profile. *Environ. Sci. Technol.* **29**: 1816–1821.

66. Wyatt, D.E., Richers, D.M., and Pirkle, R.J. (1995). Barometric pumping effects on soil gas studies for geological and environmental characterization. *Environ. Geol.* **25**: 243–250.
67. Eklund, B., Balfour, W.D., and Schmidt, C.E. (1985). Measurement of fugitive volatile organic emission rates. *Environ. Prog.* **4**: 199–202.
68. Batterman, S.A., McQuown, B.C., Murthy, P.N., and McFarland, A.R. (1992). Design and evaluation of a long-term soil gas flux sampler. *Environ. Sci. Technol.* **26**: 709–714.
69. Reinhart, D.R., Cooper, D.C., and Walker, B.L. (1992). Flux chamber design and operation for the measurement of municipal solid waste landfill gas emission rates. *J. Air Waste Manage. Assoc.* **42**: 1067–1070.
70. Eklund, B., Anderson, E.P., Walker, B.L., and Burrows, D.B. (1998). Characterization of landfill gas composition at the Fresh Kills municipal solid-waste landfill. *Environ. Sci. Technol.* **32**: 2233–2237.
71. Tillman, F. and Smith, J.A. (2004). Design and laboratory testing of a chamber device to measure total flux of volatile organic compounds from the unsaturated zone under natural conditions. *J. Contam. Hydrol.* **75**: 71–90.

APPLICATIONS OF SOIL VAPOR DATA TO GROUNDWATER INVESTIGATIONS

Blayne Hartman
H&P Mobile Geochemistry
Solana Beach, California

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

This article describes the uses and interpretation of soil vapor data for a variety of common applications related primarily to groundwater. It begins with an overview of the applications of soil vapor surveys, a brief discussion of the various soil vapor methods, and then focuses on a number of specific issues relating to soil vapor surveys for groundwater investigations. The intent of this article is to guide professionals on when, why, and how soil vapor data can be useful for groundwater-related investigations and how to interpret the results.

BACKGROUND AND APPLICATIONS

Soil vapor in the pore spaces, also called soil gas, soil air, pore gas, and interstitial gas refers to the vapor that lies among the soil grains in the soil column lying between the ground surface and the water table (vadose zone). Soil vapor data are collected in a variety of ways, and the data can be used for various purposes. Soil vapor samples are most commonly analyzed for volatile organic compounds (VOCs). VOCs include simple hydrocarbons such as methane; lighter hydrocarbons such as gasoline; gasoline constituents such as benzene and methyl tert-butyl ether (MtBE); and chlorinated solvents such as trichloroethylene (TCE), trichloroethane (TCA), and tetrachlorethylene (PCE).

Soil vapor is surveyed for a variety of reasons. Historically, soil vapor surveys have been used primarily for site assessment to identify soil and groundwater contamination (1–4). More recent applications include evaluation of upward and downward vapor transport to monitor site remediation (natural and engineered) and to monitor landfills and migrating VOCs (5–7).

Soil vapor surveys are quick, relatively nondisruptive, and generally inexpensive. The surveys are frequently performed early in the site investigation, when investigation for numerous VOCs is performed. The data obtained are often used to focus subsequent soil and groundwater investigatory sampling programs.

Regulatory agencies across the country are increasingly concerned with human health risk due to the upward migration of contaminant vapors through the vadose zone into buildings. Vertical soil vapor profiles enable measuring actual soil vapor concentrations that give a more accurate calculation of the upward vapor migration than values calculated from soil or groundwater data. These data are used in risk based corrective action (RBCA) computer models (8).

Use of soil vapor data for monitoring site remediation takes many forms. Most commonly, concentrations of vapor contaminants are monitored to measure the progress of a vapor extraction system. Remediation of groundwater or soil by natural attenuation requires monitoring the concentrations of contaminants, by-products, and oxygen in the soil vapor to measure the progress of the degradation of the contaminant plume.

Soil vapor is frequently measured at the fringes of landfills to ensure that harmful compounds such as methane, chlorinated compounds (primarily vinyl chloride), and petroleum related carcinogens such as benzene are not migrating from landfills to neighboring developments.

SOIL VAPOR METHODS

Collection of Soil Vapor Samples

Three methods exist for collecting soil vapor data: (1) active, (2) passive, and (3) flux chambers. A brief overview of these methods follows. More comprehensive descriptions of soil vapor collection methods and protocols are given in documents written by Hartman (9,10), Hartman and Jacobs (5,11), EPA (1), ASTM (12), and Devitt et al. (4).

The *active* approach is quantitative and consists of the withdrawal of a volume of vapor from the subsurface, typically using a steel sampling rod driven into the ground from 0.9 to 6 meters below ground surface, and subsequent analysis of the withdrawn vapor. Originally, surveys were limited to shallow depths of 3 meters because the sampling

[1] Some of this material is excerpted from previous work, appearing in "Petroleum Hydrocarbon Contamination-Legal and Technical Issues" and "Chlorinated Solvent Contamination-Legal and Technical Issues" Copyright © 1998, 1999, Argent Communications Group. Reprinted by permission of Argent Communication Group, P.O. Box 1425, Foresthill, CA 95631. Telephone 1-800-419-2741; or sales@ argentco.com.

rods were driven and extracted using hand techniques. Currently, active surveys can routinely achieve depths exceeding 30.5 meters due to the use of direct-push probe rigs, cone penetrometer (CPT) sampling rigs, and full-size drill rigs to collect soil vapor using newly developed sampling tools.

Sample containers for active vapor sampling include syringes, gas sampling Tedlar® bags, preevacuated stainless steel canisters (e.g., SUMMA canisters), and gastight vials (e.g., 40 mL preevacuated volatile organic analysis (VOAs) vials or headspace vials). Analysis can be performed on-site or off-site, usually within 48 hours of collection. The quantitative results are typically presented in concentration units (e.g., ppmv, μg/L vapor). The active method requires actively withdrawing soil vapor and is best suited to sites that have higher vadose zone permeabilities and higher contaminant concentrations (>0.1 μg/L vapor). The active method can be used for investigating groundwater contamination, and as discussed later in this chapter, the active method is also useful for groundwater issues relating to contaminant fate and transport because it gives concentration data.

The *passive* approach consists of emplacing a natural or synthetic absorbent into the subsurface (typically 0.15 to 0.9 meters) for an extended time, and subsequently removing and analysing the absorbent. The absorbent is typically placed in the upper end of an inverted container or sleeve that has openings for migrating vapors to enter and collect. Unlike the active method, which requires one visit to the site, the passive method requires one visit to the site to deploy the absorbents and a second trip to retrieve them.

The passive method is qualitative and does not allow for real-time data. Measured values are reported as total absorbed mass (e.g., μg) or relative units because the amount of vapor meeting the absorbent is unknown. The passive method offers more sensitivity than the active method; hence it offers advantages in sites of low vadose zone permeability and lower contaminant concentrations (>0.01 μg/L vapor). Contaminants must have an appreciable vapor pressure to be detected by this method, regardless of the deployment time or contaminant concentration in the soil or water. As a rule of thumb, one should not expect compounds whose vapor pressures are less than 10^{-6} atmospheres to be detectable by passive soil vapor methodology in a reasonable time frame (<50 days of exposure) for reasonable contaminant concentrations in the water.

Flux chambers consist of an enclosed chamber that is placed on the surface for a specific period of time. Vapor concentrations are measured in the chamber through time or after a period of time. This method is also quantitative and yields both concentration data in the chamber and flux data (mass/area-time). Flux chambers are the least common soil vapor method and are seldom used for groundwater assessment. Their most useful application is for sites where volatilization of contaminants from groundwater is of concern and direct fluxes are desired Hartman (9).

Analysis of Soil Vapor Samples

The type of analytical instrumentation employed depends upon the goals of the project and the type of soil vapor method used. For screening-level active soil vapor surveys, in which the primary goal is simply to find contamination qualitatively, simple handheld or portable instruments, such as organic vapor analyzer (OVAs), organic vapor meter (OVMs), gas-monitoring meters (e.g., IR detectors), and gas detector tubes, can be used. For applications where quantitative results are desired, the analytical methodology employed is typically gas chromatography (GC) or gas chromatography-mass spectrometry (GC-MS). These latter techniques differentiate a wide variety of compounds. These latter methods are also used to analyze passive soil vapor samples and flux chamber samples.

More comprehensive treatments of the available analytical options for soil vapor samples are given by Hartman & Jacobs (5) and the EPA (1).

APPLICATIONS OF SOIL VAPOR DATA TO GROUNDWATER ISSUES

This section discusses how to use and interpret soil vapor data for applications specifically related to groundwater.

Assessing and Defining Groundwater Contamination

Soil vapor surveys have frequently been used for initial site assessment to identify groundwater contamination, and the data are used to focus subsequent groundwater investigatory sampling programs (1,3,4). The soil vapor data are typically plotted as contours or as a raised surface. More sophisticated approaches, including three-dimensional visualizations of the soil column, cross sections, normalization, and statistical analysis, are sometimes used depending on the project goals.

The applicability of a soil vapor survey to detect potential groundwater contamination depends upon a number of factors, including the local geology, contaminant properties, and the concentration of the contaminant in the groundwater (11). The maximum concentration of a contaminant in the soil vapor due to volatilization from the groundwater can be estimated from Henry's law:

$$C_{\mathrm{sg}} = HC_{\mathrm{w}} \tag{1}$$

where C_{sg} is the soil gas concentration in equilibrium with the groundwater,
C_{w} is the groundwater concentration, and
H is the dimensionless Henry constant.

As an example, consider groundwater that has a benzene concentration of 10 μg/L. The dimensionless Henry constant for benzene is 0.22, giving an equilibrium soil vapor concentration of 2.2 μg/L vapor at the groundwater interface. Hence, a soil vapor survey conducted near groundwater could at best be expected to detect only a maximum value of 2.2 μg/L in the vapor if groundwater were the only source.

Equilibrium phase partitioning is rarely achieved in the subsurface, so measured soil vapor values from groundwater volatilization will be significantly lower (generally one to two orders of magnitude) than predicted from Henry's constants. In addition, soil vapor concentrations decrease as distance increases from the groundwater due to other processes active in the vadose zone, such as biodegradation and sorption onto soils (13). For these reasons, the chances of detecting groundwater contamination from soil vapor data increases the closer the samples are collected to groundwater.

Due to its increased sensitivity for some compounds, the passive soil vapor method can offer advantages over the active method for this application. Similar to the active method, the ability of passive methods to detect groundwater contamination increases the closer the samples are collected to groundwater. However, regardless of sampling depth, the capability of passive methods to detect contaminants is still limited by the vapor pressure of the compound of concern (5).

Transport of Contaminant Vapors in the Subsurface

To understand the source of soil vapor and groundwater contamination, it is necessary to understand how contaminants move in the vapor phase. A condensed version of this topic is addressed in this section beginning with a discussion of the processes by which vapors move through the vadose zone. A more detailed presentation can be found in Hartman (6,7,14).

There are two types of physical processes by which contaminants are transported in the vapor phase. The first is advection, the bulk movement of the vapor itself. In advective transport, any contaminants in the vapor are carried along with the moving vapor. Advective transport processes can be important in soil vapor, for example, near the surface due to atmospheric pressure differences (15), near buildings that can create pressure gradients (16), and movement of vapor by density-induced flow (17).

The second type of transport process is gaseous diffusion, the motion of contaminants by molecular processes through a nonmoving vapor column, which is the primary mechanism for contaminant transport in the vapor phase through the vadose zone (18).

Contaminant transport by gaseous diffusion is described by Fick's first law as

$$\text{FLUX} = \frac{D_e\, C_{sg}}{X} \tag{2}$$

where Flux is the rate of movement of a compound per unit area,
D_e is the effective diffusion coefficient in the vadose zone,
C_{sg} is the contaminant concentration gradient in the soil vapor, and
X is depth in the vadose zone

Similar to momentum transfer (e.g., water running downhill) and heat transfer (movement from hot areas to cold areas), contaminant transfer by gaseous diffusion moves from areas of high concentration to areas of low concentration. The flux is always down the concentration gradient, regardless of the orientation of the concentration gradient with respect to the depth below the surface. Diffusional transport is a three-dimensional process, so contaminants move away from a source in all directions, similar to an expanding balloon.

Differentiating the Source of Soil Vapor/Groundwater Contamination

A common problem confronted by environmental geologists is determining the source of contamination at a site that has both soil vapor and groundwater contamination. Is the soil vapor responsible for the groundwater contamination, or is the groundwater responsible for the soil vapor concentration? Contaminant transfer by gaseous diffusion moves from areas of high concentration to areas of low concentration, so vertical profiles of the soil vapor through the vadose zone down to groundwater can provide this answer. Figure 1 shows the expected shape of soil vapor profiles in the vadose zone for three common conditions. Processes active in the vadose zone such as sorption, biodegradation, and advection may alter these ideal profiles, but the source of contamination must still be in the area of highest concentration, as dictated by Ficks law.

Vertical soil vapor is profiled by installing vapor monitoring wells at multiple depths through the vadose zone to the groundwater. For soil vapor, vapor wells are often constructed of small diameter 0.5 to 1.0 cm (1/8 in. to 1/4 in. OD) flexible tubing and "nested" in the same borehole (Fig. 2).

Repeated sampling of the vapor wells allows averaging any temporal effects. Thomson (19) describes two case histories of vertical profiling using nested vapor wells.

Contaminant Vapor Transport Rates

The distance that contaminant vapors can move by gaseous diffusion can be approximated from

$$\text{Distance} = (2 \times D_e \times t)^{1/2} \tag{3}$$

where D_e is the effective diffusivity and t is time.

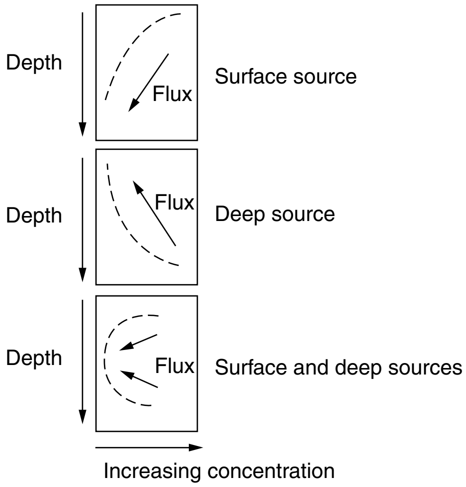

Figure 1. Common soil vapor profiles in the vadose zone for several one-dimensional situations (8).

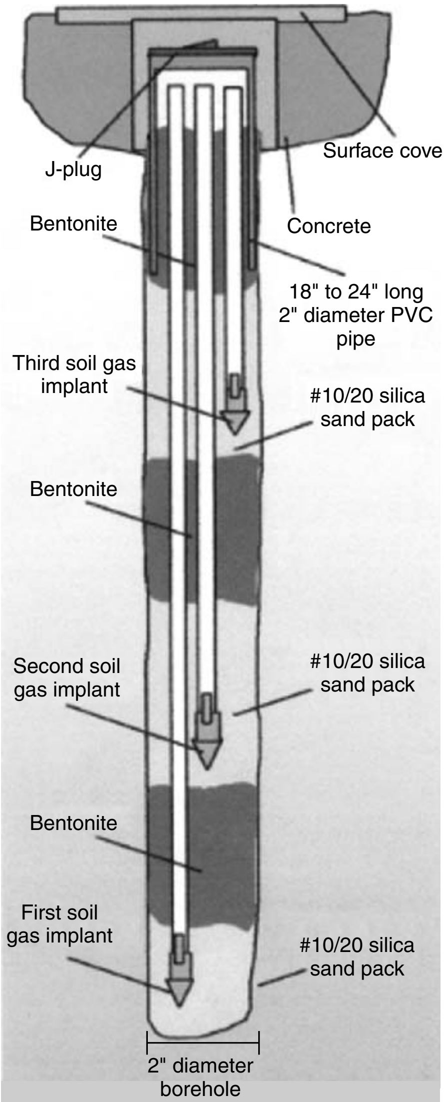

Figure 2. Schematic diagram of a multidepth, nested vapor monitoring well (19).

Through the Vadose Zone. For contaminant vapor transport through the vadose zone, the effective diffusion coefficient is the gaseous diffusion coefficient corrected for soil porosity. For many vapors, the gaseous diffusion coefficient is approximately 0.1 cm^2/s. The effect of soil porosity varies depending upon the type of soil, and several equations are available to calculate the effect of porosity on diffusivity. A general approximation is that the porosity reduces the gaseous diffusivity by a factor of 10. Thus, for vapors, D_e can be approximated as 0.01 cm^2/s.

The distance that contaminant vapors can move through the vadose zone in a year can be estimated as

$$\text{Distance} = (2 \times 0.01\ \mathrm{cm}^2/\mathrm{s} \times 31{,}536{,}000\ \mathrm{s})^{1/2}$$

$$\sim 800\ \mathrm{cm}$$

This calculation shows that contaminant vapors can move long distances through the vadoze zone in a short period of time. Within a few years, vapor contamination can move laterally underneath a neighboring room or building or downward to the groundwater surface.

Into or out of Groundwater. In contrast to movement through the vadose zone where contaminant vapors move through the soil vapor, the movement of contaminant vapors into or out of groundwater is controlled by the rate at which vapors partition into and move through the liquid. Because groundwater movement is so slow, the water interface remains relatively undisturbed (laminar flow), and vertical mixing of the water is minimal. The primary exchange process is again molecular diffusion, but in this case, the exchange rate is controlled by liquid diffusion, not gaseous diffusion. A general value for the liquid diffusion coefficient for compounds is approximately 0.00001 cm^2/s. Using the same factor of 10 reduction due to soil porosity, D_e for most liquids can be approximated as 0.000001 cm^2/s.

The distance that contaminants can move into and through groundwater in a year can be estimated as

$$\text{Distance} = (2 \times 0.000001\ \mathrm{cm}^2/\mathrm{s} \times 31{,}536{,}000\ \mathrm{s})^{1/2}$$

$$\sim 8\ \mathrm{cm}$$

These calculations show that although contaminant vapors can move through the vadose zone relatively quickly, they partition into and move through groundwater extremely slowly. The reverse situation is also true; the movement and partitioning of contaminants through groundwater and into soil vapor is also extremely slow.

Estimation of the Downward Transport of Contaminant Vapors into Groundwater

The calculations summarized in the preceding sections indicate that although contaminant vapors can move quickly down to groundwater, they do not partition into groundwater very quickly because phase partitioning across a liquid laminar interface is kinetically controlled. For this reason, phase partitioning rarely reaches equilibrium, and the use of Henry's constants to estimate groundwater concentrations from soil vapor data (or the reverse) often gives estimated values one to two orders of magnitude from measured values.

Using a modification of Fick's first law, the transfer rate of a contaminant from soil vapor into groundwater can be estimated, and in turn, reasonable estimates of the expected contaminant concentration in groundwater due to contamination in the overlying soil vapor can be estimated.

The flux is calculated as

$$\text{Flux} = K_\mathrm{L}\ (C_\mathrm{sg}/H - C_\mathrm{w}) \tag{4}$$

where K_L is the gas exchange coefficient (length/time),
C_sg is the soil gas concentration time (mass/volume),
H is the dimensionless Henry's law constant, and
C_w is the groundwater concentration (mass/volume).

The ratio C_{sg}/H represents the groundwater concentration in equilibrium with the overlying soil vapor at the interface between the soil vapor and groundwater.

The gas exchange coefficient has units of velocity and represents the distance that contaminants move through the groundwater per unit time. This coefficient is highly dependent on a compound's effective diffusivity and advective mixing of the groundwater interface. For groundwater velocities less than 30.5 meters per year and soil grain sizes less than 0.5 mm, diffusional transport dominates over dispersion, and K_L can be approximated as:

$$K_L = 1.1\,(D_e\,v/d)^{1/2} \tag{5}$$

where D_e is the effective diffusivity,
d is the horizontal distance (downgradient) over which the soil vapor plume and groundwater are in contact, and
v is the horizontal groundwater flow velocity.

For the groundwater conditions described above, the interface remains laminar (i.e., no turbulent mixing), and the gas transfer coefficient is approximately 8 cm/yr. Using this value for the gas exchange rate, estimates of the expected contaminant concentration in groundwater due to contamination in the overlying soil vapor for various soil vapor concentrations are summarized in Table 1. These calculations assume equilibrium partitioning at the soil vapor/groundwater interface, transfer into groundwater by molecular diffusion only, and uniform mixing of the contaminant into the groundwater over the well screen length.

Table 1 shows that for equilibrium concentrations (C_{sg}/H) up to 100 μg/L, the resulting ground water concentration after 5 years will be low (<10 μg/L) if liquid molecular diffusion is the only exchange process.

Table 1. Expected Contaminant Concentration in Groundwater for Various Soil Vapor Concentrations at the Groundwater Interface[a,b]

C_{sg}, μg/L	C_{sg}/H, μg/L	Flux, μg/cm^2-yr	Water conc at 1 yr, μg/L	Water conc at 5 yrs, μg/L
10	40	315	0.63	3.2
20	80	631	1.3	6.3
30	120	946	1.9	9.5
40	160	1261	2.5	13
50	200	1577	3.2	16
100	400	3154	6.3	32
500	2000	15,768	31	158
1000	4000	31,536	63	315

[a] Reference 5.

[b] Assumptions used to compute the tabulated values:

Henry's law constant:	0.25 (dimensionless)
Gas transfer coefficient:	8 cm/yr
Well screen interval:	5 m

Importance of Downward Vapor Transport for Several Common Contaminants

The potential importance of downward vapor migration for several of the more common compounds encountered by environmental consultants are given special consideration in this section.

Leaking Vapors from Chlorinated Solvents (Dry Cleaners and Vapor Degreasers). Many compounds, primarily chlorinated solvents, have relatively high vapor pressures and vapor densities three to six times greater than air. Because of these physiochemical properties, vapors may emanate from containers or pipes holding gaseous or liquid products. If air flow is restricted, such as in a closed room containing a washer unit, degreaser, or tank, leaking dense vapors can penetrate the floor and enter the upper vadose zone. Density-driven advection can continue to drive vapor downward through the vadose zone until the vapor is diluted to low enough concentrations (<1%) that density-driven advection is no longer an important factor in vapor transport and molecular diffusion takes over. Vapor clouds reaching tens of feet into the uppermost vadose zone have been documented in thousands of cases. Businesses and commercial operations that are most susceptible to this situation are those that deal with chlorinated solvents (dry cleaners, vapor degreasers, spray facilities, etc.).

The soil vapor concentration to yield an equilibrium water concentration of 100 μg/L is 35 μg/L vapor for TCA, 60 μg/L vapor for PCE, and 100 μg/L vapor for vinyl chloride. In areas of groundwater deeper than 9.1 meters, soil vapor concentrations for these compounds are generally less than these values at groundwater, unless in the presence of contaminated soil, suggesting that surface vapor leaks are not a significant contaminant pathway to groundwater in most cases. In areas of groundwater shallower than 6 meters, soil vapor concentrations exceeding these values are found, even in the presence of clean soil, suggesting that leaking dense vapors from the surface may be a significant contaminant pathway.

Leaking, Hydrocarbon Fuel Vapors. In contrast to chlorinated solvents, aboveground vapor releases are not common for petroleum hydrocarbons because they rarely are stored in confined spaces (due to their flammability). However, subsurface vapor leaks from petroleum hydrocarbons are possible from USTs and piping associated with them (vent pipes, pipe joints, and tank bungs. Gasoline vapor contains approximately 8000 μg/L vapor of benzene and lower, although still high concentrations (>1000 μg/L vapor) of toluene and xylene. Reformulated gasolines have very high concentrations of MtBE or ethanol, and their vapors contain approximately 130,000 μg/L vapor of MtBE and 11,500 μg/L vapor of ethanol. Due to these high concentrations, leaking vapors must be considered a potential source of groundwater contamination, as further described below.

Benzene and Light Aromatics (BTEX). For light aromatics, the soil vapor concentration to yield an equilibrium water concentration of 100 μg/L is approximately 25 μg/L vapor. In areas of groundwater deeper than 9.1 meters, soil vapor concentrations of aromatics are generally less than 25 μg/L vapor at the groundwater, unless in the presence of contaminated soil from a leaking UST, suggesting that vapor leaks are not a contaminant pathway in most cases.

In areas of groundwater shallower than 3 meters, soil vapor concentrations exceeding 25 μg/L vapor have been found even in the presence of clean soil, suggesting that leaking gasoline vapors from the UST are a possible contaminant pathway. In areas where the groundwater is near or above the UST, vapor leaks could be very significant in groundwater contamination.

MtBE and Ethanol. Due to extremely low Henry constants, the soil vapor concentration to yield an equilibrium water concentration of 100 μg/L is only 2 μg/L vapor for MtBE and approximately 0.02 μg/L vapor for ethanol. Hence, leaking gasoline vapors from a UST are a likely contaminant pathway to groundwater for both compounds in shallow and deep groundwater. The likelihood is even more acute in locations where the water table is near or above the UST and/or in areas that have significant amounts of rainwater infiltration.

Summary. In areas of deeper groundwater (>10 m), slow groundwater flow rates less than 30.5 meters per year and low recharge, soil vapor concentrations for most contaminants at the groundwater surface are typically well below equilibrium values (C_{sg}/H) of 100 μg/L. Hence contamination of groundwater by downward vapor transport is not likely to be significant. For compounds with very low Henry constants (MtBE, ethanol), downward vapor transport could be important, even under these groundwater conditions.

In areas where groundwater is shallower than 10 meters or higher groundwater flow rates and high recharge, soil vapor concentrations for many contaminants at the groundwater surface may approach high enough values that contamination of the groundwater by downward vapor transport may be significant. This potential will be enhanced for compounds with low Henry constants and when vapor leaks occur directly into the groundwater.

Additional discussion of the downward migration of vapors can be found in Hartman (7,14).

The Upward Migration of Vapors from Groundwater Devolatilization

Attention has been raised in the past several years to the risk to human health by the upward migration of contaminant vapors into buildings and other enclosed spaces (20–22). In cases of contaminated groundwater, it is common to assess this risk pathway before site closure is obtained.

Simplified, the upward contaminant vapor flux into a building is computed by Fick's first law, requiring the soil vapor concentration at some depth underlying the structure. Soil vapor data may be measured directly, or alternatively, in the absence of actual soil vapor data, soil vapor concentrations can be calculated from groundwater data assuming equilibrium conditions using Henry's constants (23). As discussed in the preceding sections, soil vapor values calculated from groundwater using Henry's constants are likely to be overestimated by amounts as large as 10 to 100 times; hence the resulting risk will also be overestimated. The potential error in the calculated risk introduced by the incorrect vapor concentration can be orders of magnitude greater than errors introduced using default values for other parameters, such as porosity. Thus, if a site fails the upward risk calculation using calculated soil vapor values from groundwater data, direct measurement of actual soil vapor concentrations will yield more representative soil vapor values to be used in the calculation (13).

When groundwater is extremely shallow (1.5 meters) and representative soil vapor data are unobtainable, the use of surface flux chambers may yield the best estimate of the vapor flux (24).

A more detailed discussion of the upward vapor migration risk pathway, including a recommended protocol for determining the upward vapor flux, can be found in Hartman (8).

BIBLIOGRAPHY

1. EPA. (1997). *Expedited Site Assessment Tools For Underground Storage Tank Sites*. U.S. Government Printing Office, Washington, DC.
2. Marrin, D.L. and Kerfoot, H.B. (1988). Soil gas surveying techniques. *Environ. Sci. Technol.* **22** (7): 740–745.
3. Ullom, W.L. (1995). Soil gas sampling. In: *Handbook of Vadose Zone Characterization and Monitoring*. L.G. Wilson, L.G. Everett, and S.J. Cullen (Eds.). Lewis, Boca Raton, FL, pp. 555–567.
4. Devitt, D.A., Evans, R.B., Jury, W.A., Starks, T.H., Eklund, B., and Gholson, A. (1987). *Soil gas sensing for detection and mapping of volatiles organics*, EPA/600/8-87/036. (NTIS PB87-228516).
5. Hartman, B. and Jacobs, J. (2000). Applications and interpretation of soil vapor data to volatile organic compound contamination. In: *Standard Handbook of Environmental Science, Health, and Technology*. J. Lehr and J. Lehr (Eds.). McGraw-Hill, New York.
6. Hartman, B. (1999). Applications and interpretation of soil vapor data. In: *Chlorinated Solvent Contamination: Legal and Technical Issues*. Morrison et al. (Eds.). Argent Communications Group, Foresthill, CA.
7. Hartman, B. (1998). Applications and interpretation of soil vapor data. In: *Petroleum Hydrocarbon Contamination: Legal and Technical Issues*. Morrison et al. (Eds.). Argent Communications Group, Foresthill, CA.
8. Hartman, B. (1997). The upward migration of vapors. *LUSTLine Bull.* **27**, NEIWPCC, Wilmington, MA, December 1997.
9. Hartman, B. (2003). How to collect reliable soil gas data for risk based applications. Part 2: Flux chamber method. *LustLine* # **44**, NEIWPIC, August 2003.
10. Hartman, B. (2002). How to collect reliable soil gas data for risk based applications. Part 1: Active soil gas method. *LustLine* # **42**, NEIWPIC, October 2002.
11. Hartman, B. and Jacobs, J. (2000). Soil vapor principles. In: *Standard Handbook of Environmental Science, Health, and Technology*. J. Lehr and J. Lehr (Eds.). McGraw-Hill, New York.

12. ASTM. (1993). *Standard guide for soil gas monitoring in the vadose zone*, D5314-93. *Annual Book of ASTM Standards*. Philadelphia, PA.
13. McNeel, P.J. and Dibley, V. (1997). Case study comparisons of vapor inhalation risk estimates: ASTM RBCA model predictions vs site specific soil vapor data. *Proc. 1997 Superfund XVIII Conf.*, Washington, DC, December 1997, pp. 509–516.
14. Hartman, B. (1998). The downward migration of vapors. *LUSTLine Bull.* **28**, NEIWPCC, December 1998.
15. Massman, J. and Farrier, D.F. (1992). Effects of atmospheric pressures on gas transport in the vadose zone. *Water Resourc. Res.* **28** (3): 777–791.
16. Little, J.C., Daisey, J.M., and Nazaroff, W.W. (1992). Transport of subsurface contaminants into buildings. *Environ. Sci. Technol.* **26** (11): 2058–2066.
17. Cherry, J. and Smyth, D. (1996). Non-aqueous phase liquids in the subsurface environment: Assessment & remediation. *Am. Soc. Civ. Eng., 1996 Annu. Convention & Exposition*, Washington, DC, November 10–14, 1996.
18. Thorstenson, D. and Pollock, D. (1989). Gas transport in unsaturated porous media: The adequacy of Fick's law. *Rev. Geophys.* **27**: 61–78.
19. Thomson, K.A. (1995). Soil gas sampling. In: *Handbook of Vadose Zone Characterization and Monitoring*. L.G. Wilson, L.G. Everett, and S.J. Cullen (Eds.). Lewis, Boca Raton, FL, pp. 569–588.
20. Jury, W.A., Russo, D., Streile, G., and El Abd, H. (1990). Evaluation of volatilization by organic chemicals residing below the surface. *Water Resourc. Res.* **26**: 13–20.
21. Johnson, P.C. and Ettinger, R.A. (1991). Heuristic model for predicting the intrusion rate of contaminant vapors into buildings. *Environ. Sci. Technol.* **25**: 1445–1452.
22. Jeng, C.-Y., Kremesec, V.J., and Primack, H.S. (1996). Models of hydrocarbon vapor diffusion through soil and transport into buildings. In: *Proc. 1996 Pet. Hydrocarbons & Org. Chem. Ground Water: Prevention, Detection, and Remediation*, Houston, TX, November 1996.
23. ASTM. (1995). Standard guide for risk-based corrective action (RBCA) applied at petroleum release sites, E-1739-95. *Annual Book of ASTM Standards*. Philadelphia, PA.
24. Smith, J.A., Tisdale, A.K., and Cho, H.J. (1996). Quantification of natural vapor fluxes of trichoroethen in the unsaturated zone at Picatinny Arsenal, New Jersey. *Environ. Sci. Technol.* **30**: 2243–2250.

GROUNDWATER VELOCITIES

MUKAND SINGH BABEL
Asian Institute of Technology
Pathumthani, Thailand

Groundwater flows through interconnected pore spaces in subsurface soil in response to differences in fluid pressure and elevation. Groundwater flow rates are extremely slow, compared to surface water, and they are also highly variable. For example, in clayey soils, groundwater moves as slowly as a fraction of a millimeter per day; however, in sandy soils, it moves faster, even up to more than a meter per day. Hydrology is concerned with both the speed and the direction of groundwater movement. Groundwater velocity is very important in dealing with contaminant migration and arrival time, mass flux, site characterization, reactive barrier design, and calibration of transport processes.

Hydraulic head drives flow through porous media. Two main components contributing to the hydraulic head at a given point in a groundwater flow system are pressure head and elevation head, as shown in Fig. 1. The pressure head is due to the porewater pressure at that point, which can be measured by a piezometer. The elevation head is due to gravitational energy. The higher the elevation of a fluid from sea level or some selected datum, the more gravitation potential energy it has. The third component, kinetic energy, can be ignored, as groundwater flow velocities are usually very small in magnitude. Thus, the hydraulic head (h) is the sum of the pressure head (Ψ) and the elevation head (z), expressed usually in the units of length (meters).

Henry Darcy, a French engineer who studied the movement of water through sand in 1856, found that the rate of water flow through a pipe filled with sand is proportional to the difference in the height of the water between the two ends of the tube, and inversely proportional to the length of the tube (Fig. 2). The constant of proportionality, K, is called hydraulic conductivity. This law, universally known as Darcy's law, is given as

$$Q = -KA[(h_A - h_B)/L]$$

or

$$Q = -KA(dh/dL) \tag{1}$$

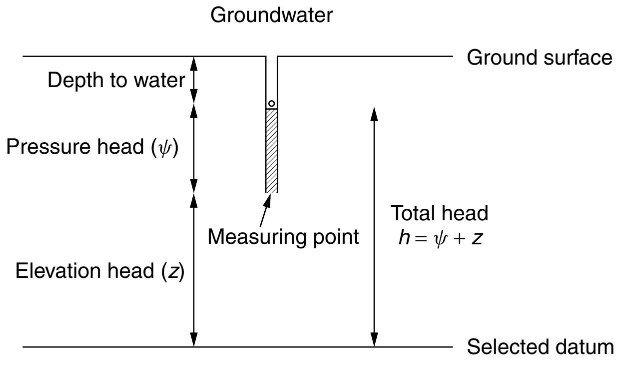

Figure 1. Elevation head, pressure head, and total head (adapted from Reference 1).

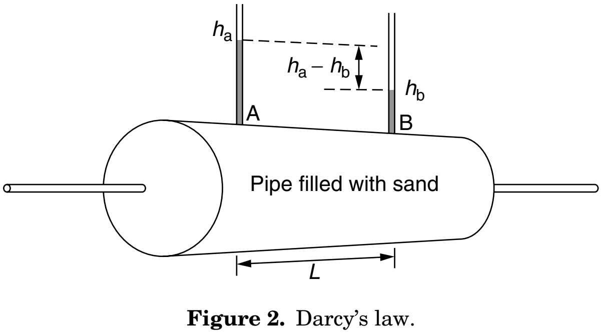

Figure 2. Darcy's law.

or

$$q = Q/A = -K\ (dh/dL) \tag{2}$$

where Q is the quantity of flow per unit of time (m^3/day); $K =$ hydraulic conductivity (m/day); $A =$ cross-sectional area (m^2); h_A and h_B are the hydraulic head at points A and B, respectively; and $dh/dL =$ hydraulic gradient. The term Q/A in Eq. 2 is also known as the specific discharge (q), which is the flow per unit area [LT^{-1}]. The specific discharge is sometimes referred to as Darcy flux, Darcy velocity, or bulk velocity.

The negative sign in Eq. 2 indicates that the fluid moves in the direction of decreasing hydraulic head. Darcy's law assumes that the flow is laminar, that is, the water follows distinct flow lines rather than mixing with other flow lines, which is the case for most groundwater flow in porous media. Most groundwater movements occur in small interstices, so that there is considerable resistance to flow imposed by the aquifer material, making the flow laminar. When the flow velocity increases, especially in media having large pores, the occurrence of turbulent eddies dissipates kinetic energy, rendering the hydraulic gradient less effective in inducing flow. In very large interstices, such as those in many limestone and volcanic areas, groundwater flow is almost identical to the turbulent flow of surface water. Darcy's law is not valid for turbulent flow, for unusually high velocity that may be found in fractures or in granular aquifers.

Equation 2 yields only an apparent velocity value, the macroscopic velocity, through the cross-sectional area of a solid matrix and interstices, which, however, is not the true velocity of water moving through the pores. Clearly, flow velocities through interstices alone will be much higher than the macroscopic value. The interstices themselves vary in shape, width, and direction, so the actual velocity in the soil or rock is highly variable. Furthermore, the flow path of a water particle around and between the grains in an aquifer is tortuous, so water particles must travel longer than the apparent distance given by the measured length of the porous medium in the direction of flow, as shown schematically in Fig. 3.

To calculate the true linear velocity, one must take into account the fact that water travels only through the pores of the sediment. The fraction of the area covered by open space is equal to the effective porosity (n_e). Thus, the effective velocity (v), also known as the porewater velocity, is obtained by dividing the macroscopic velocity by the effective porosity:

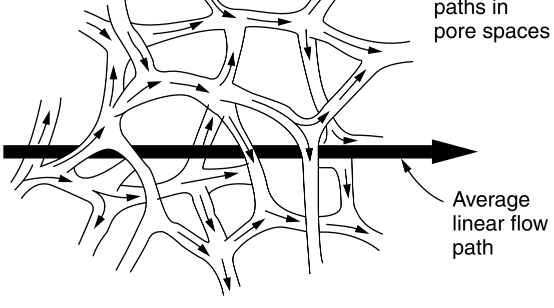

Figure 3. Concept of average linear velocity (2).

$$v = -(K/n_\mathrm{e})dh/dL \tag{3}$$

or

$$v = Q/n_\mathrm{e}/A$$

or

$$v = q/n_\mathrm{e}$$

From Eq. 3, it can be seen that, for sand with an effective porosity of 33%, the actual velocity is three times the Darcy velocity. Therefore, to define the actual velocity, one must consider the microstructure of the material. Usually, actual velocities tend to decrease with depth as porosities and hydraulic conductivities also decrease.

BIBLIOGRAPHY

1. Ward, R.C. and Robinson, M. (2000). *Principles of Hydrology*, 4th Edn. McGraw-Hill, London.
2. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.

VISCOUS FLOW

Navraj S. Hanspal
Vahid Nassehi
Loughborough University
Loughborough, United Kingdom

Diganta Bhusan Das
Oxford University
Oxford, United Kingdom

INTRODUCTION

This text describes the elementary concepts of viscous fluid flow. The principal laws that govern the motion of fluids are also discussed. From our common experience, we know that fluids deform under the action of external forces. The deforming forces and rate of deformation are related to each other through a proportionality constant known as the coefficient of viscosity. The value of this constant depends upon the chemical identity, material properties, and thermodynamic state of a given fluid. To develop a theory that shows how viscosity affects fluid dynamics, one must understand the relationships between stress and the rate of strain. The outlines given in this article provide basic concepts of viscous flow. For more detailed analysis, the reader should consult specialized texts on this topic. Some of these texts are listed in the references at the end of this article.

VISCOSITY AND NEWTON'S LAW OF VISCOSITY

The physical property by which the resistance to fluid flow can be characterized is known as viscosity. A more fundamental approach to viscosity shows that it is the property of a fluid that relates applied stress to a resulting strain rate. By experience, we come across many different types of fluids. For example, consider water and lubricating oil such as that used in automobiles. Water can flow more freely compared to lubricating oil under isothermal conditions. Hence, we say that the viscosity of water is less than that of oil.

When we deform a solid, so that it is strained, we know that the solid will exert a restoring force that opposes the strain; for small strains, the restoring force is proportional to the strain, the familiar Hooke's law. Real fluids also oppose strains; however, in a fluid it is not the amount of strain that is important but the rate at which the strain is produced.

Consider simple fluid flow between the two plates shown in Fig. 1. If the force per unit area of the top plate is F, we find that

$$\frac{F}{A} = \mu \frac{v}{d} \qquad (1)$$

where d is the separation between the plates, which defines the viscosity μ (sometimes called the absolute or dynamic viscosity). Equation 1 can be written in a more generalized form given by Eq. 2. The shear stress exerted in the x direction on a fluid surface of constant y by the fluid in the region of lesser y is designated as τ_{yx}, and the x component of the fluid velocity vector is designated as v_x:

$$\tau_{yx} = -\mu \frac{dv_x}{dy} = \mu \dot{\gamma} \qquad (2)$$

This equation is commonly known as *Newton's law of viscosity*. It states that the shear force per unit area is proportional to the negative of the local velocity gradient. The left-hand side of Eq. 2 represents the shear stress, and the right-hand side represents the rate of strain $\dot{\gamma} \cdot \tau_{yx}$ from Eq. 2 can also be interpreted as the viscous flux of the x momentum in the y direction. It can be seen that the viscous momentum flux is in the direction of the negative velocity gradient, that is, the momentum tends to go in the direction of decreasing velocity. A velocity gradient can thus be considered a driving force for momentum transport.

Fluids in which shear stress is directly proportional to the rate of deformation are termed *Newtonian fluids*. The viscosity of such fluids is independent of the rate of shear and rate of strain and always remains constant under given conditions. Examples of Newtonian fluids are air, water, and gasoline. In SI units, the units of viscosity are kg/m · s or Pa · s. In the cgs system, the unit of viscosity is g/cm · s, also called the poise (after Poiseuille) and equals 0.1 Pa · s. In fluid mechanics, the ratio of absolute viscosity to density is also termed kinematic viscosity which is often denoted by η. Generally, the viscosity of liquids decreases as temperature increases. Gases, on the other hand, have viscosities which generally increase with increasing temperature. More detailed explanations are based on the kinetic theory of gases and can be found in (1).

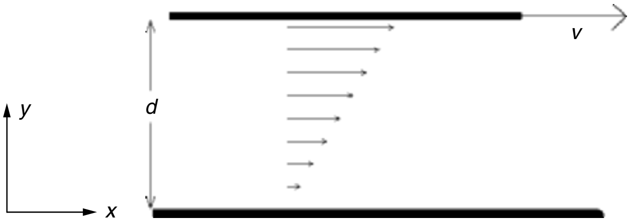

Figure 1. Viscous drag between two plates, which the bottom plate is held fixed.

NON-NEWTONIAN FLUIDS

There are fluids that do not obey the simple relationship between shear stress and shear strain rate given by Eq. 2 for a Newtonian fluid. These fluids are known as *non-Newtonian* fluids. Examples are paints, solutions of various polymers, food products such as applesauce and ketchup, emulsions and suspensions of various solids, and coal slurries. The subject of non-Newtonian flow is actually a subdivision of the science of *rheology*, which is also known as the "science of deformation and flow."

We categorize non-Newtonian fluids and discuss them under three general classifications:

1. The first is the time independent non-Newtonian fluid in which the shear rate is a unique but nonlinear function of shear stress.
2. In the second category are the time dependent non-Newtonian fluids that have more complicated shear stress and shear strain relationships. In these fluids, the shearing rate is not a single valued function of the shear stress. The shear rate depends on shearing time or on the previous shear stress rate history of that fluid.
3. In the third category are viscoelastic fluids. Unlike a truly viscous fluid in which all its energy of deformation is dissipated, some of the energy of deformation of a viscoelastic fluid may be recoverable as in the deformation of an elastic fluid.

Time-Independent Fluids

For time-independent non-Newtonian fluids,

$$\dot{\gamma} = f(\tau) \qquad (3)$$

A Newtonian fluid is a special case of this category where the function $f(\tau)$ is linear.

Time-independent non-Newtonian fluids can be commonly represented by three distinct rheograms shown in Fig. 2:

1. Bingham plastics
2. pseudoplastics
3. dilatant fluids

Bingham Plastics. Bingham plastics exhibit yield stress at zero shear rate, followed by a straight-line relationship between shear stress and shear rate. The characteristics

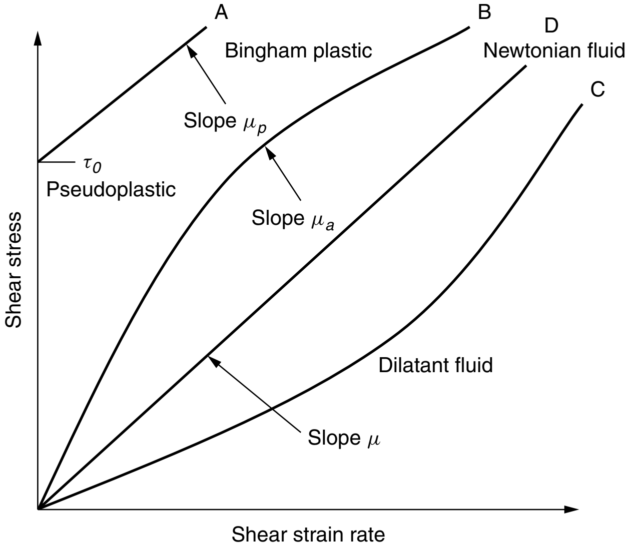

Figure 2. Shear stress–strain rate relationships for non-Newtonian fluids.

of such fluids are defined by two constants. The first is the yield stress denoted by τ_0; this is the stress that must be exceeded for the flow to begin. The second is the plastic viscosity μ_p which is the slope of the straight-line portion on curve A, Fig. 2. The model for Bingham plastics can be represented by Eq. 4:

$$\tau_{yx} = \mu_p \dot{\gamma} + \tau_0 \qquad \mathit{if}\ \tau_{yx} > \tau_0 \tag{4}$$

This model closely approximates the behavior of many real fluids such as slurries, plastics, emulsions such as paints, and suspensions of finely divided solids in a liquid. However, different constitutive models representing this type of fluids have been proposed by Herschel and Bulkley (2), Oldyroyd (3), and Casson (4).

Pseudoplastic fluids. Pseudoplastics (curve B, Fig. 2) do not have any yield stress. A pseudoplastic fluid is characterized by a progressively decreasing slope of shear stress versus strain rate. The slope has been defined as apparent viscosity,

$$\mu_a = \tau_{yx}/\dot{\gamma} \tag{5}$$

At very high shear rates, the apparent viscosity becomes constant, and the shear stress versus shear strain becomes linear. A number of empirical relations have been used to describe the behavior of pseudoplastic fluids. The simplest of these is the power law model proposed by Ostwald in 1923 [refer to (5)] and expressed in the form of Equation 6.

$$\tau_{yx} = k\dot{\gamma}$$

where $n < 1$ and k and n are constants for a particular fluid. k is known as the consistency coefficient for a given fluid, and n is the power law index. n also represents the deviation of a particular fluid from a Newtonian fluid. For a Newtonian fluid, $n = 1$, and $k = \mu$, the viscosity of the fluid. Combining Eqs. 5 and 6,

$$\mu_a = k\dot{\gamma}^{(n-1)} \tag{7}$$

The value of n is greater than unity for pseudoplastic fluids.

Dilatant Fluids. Dilatant fluids (curve C, Fig. 2) are similar to pseudoplastics in having no yield stress. They differ from pseudoplastics in a fashion such that the apparent viscosity increases as shear rate increases. The power law model given by Eq. 7 is also applicable to dilatant fluids, and so are Eqs. 6 and 7. The only difference here is that the power law index n is greater than unity. Pseudoplastic and dilatant fluids are also known as shear thinning and shear thickening fluids.

Time-Dependent fluids. Some fluids are more complex than those just described, and the apparent viscosity depends not only on the shear rate but also on the time for which the shear has been applied. Such fluids can be classified in two categories;

1. thixotropic fluids
2. rheopectic fluids

The shear stress decreases with time for a thixotropic fluid, whereas the shear stress increases with time for a rheopectic fluid. Rheopexy and thixotropy are complex phenomena resulting from transient changes of the molecular structure of time-dependent fluids under an applied shear stress. The proposed models for such fluids are hence based on many simplifying assumptions (6).

Thixotropic Fluids. The apparent viscosity of thixotropic fluids depends on the length of time of shearing as well as on the shear rate. As the fluid is sheared from the state of rest, it breaks down (on a molecular scale), but then the structural reformation increases with time. An equilibrium state is achieved when the breakdown rate is equal to the build up rate. If allowed to rest, the fluid builds up slowly and eventually regains its original consistency. As the shear is decreased, the apparent viscosity is less than that for increasing shear. For more detailed discussions, refer to (7).

Rheopectic Fluids. Molecular structure in rheopectic fluids, is formed by shear, and the behavior is opposite to that of thixotropy. A simple example of the formation of shear structure is the beating and thickening of egg whites. Many substances lose their rheopectic properties at extremely high shear rates and behave as thixotropic fluids.

Viscoelastic Fluids. A viscoelastic fluid exhibits both viscous and elastic properties. The simplest type of such a material is one which is Newtonian and obeys Hooke's law for the elastic part. The shear rate for such fluids can hence be written as

$$\dot{\gamma} = \tau_{yx}/\mu_0 + \dot{\tau}_{yx}/\lambda \tag{8}$$

where λ is the rigidity modulus. Under steady-state flow, $\dot{\gamma} = \tau_{yx}/\mu_0$ and the fluid behaves like a simple Newtonian fluid. However, if the shear is changed, an elastic effect is added. Maxwell first proposed Eq. 8 in the form

$$\tau_{yx} + \dot{\tau}_{yx}\ (\mu_0/\lambda) = \mu_0\dot{\gamma} \tag{9}$$

Liquids that obey Eq. 9 are known as Maxwell liquids. The constant $(\mu_0/\lambda)^{-1}$ is known as the relaxation time, which is the time constant for the exponential decay of stress at constant strain. If the motion is stopped, the stress relaxes at $e^{-t\lambda/\mu_0}$. Rather complex models of viscoelastic materials have been developed in which higher time derivatives of τ_{yx} and $\dot{\gamma}$ appear. A very readable text for these models is (8).

MICROSCOPIC BALANCES OF MASS, MOMENTUM, AND ENERGY

We define the control volume as a region in space within which our fluid flow system is present. Laws of conservation (or change) of mass, momentum, and energy desribe the behavior of fluid flow within this system. To yield the detailed descriptions of variables of interest, the control volume must be of infinitesimal dimensions that can shrink to zero, yielding a point volume. This approach reduces the quantities to point variables. The application of conservation principles to this infinitesimal system produces microscopic or differential conservation equations.

Shell Momentum Balances

Shell momentum balances are written to represent a fluid flow problem mathematically. The momentum balance for any fluid flow system can be written as

$$\begin{aligned}&(\text{Rate of momentum in}) - (\text{rate of momentum out})\\&\quad + (\text{sum of forces acting on the system})\\&\quad = (\text{rate of momentum accumulation})\end{aligned} \tag{10}$$

Momentum can enter the system by momentum transfer according to Newtonian (or non-Newtonian) expressions for momentum flux. Momentum can also enter from overall fluid motion, bulk transfer. The forces acting on a system in fluid flow problems are pressure forces (acting on surfaces) and gravity forces (acting on the entire volume as a whole). The general procedure for solving any viscous flow problem is

1. Write a momentum balance in form of Equation 10 for a shell of finite thickness, a microscopic shell momentum balance.
2. Let this thickness approach zero, and use the mathematical definition of the first derivative to obtain the corresponding differential equation describing the momentum flux distribution.
3. Insert the appropriate Newtonian or non-Newtonian expression for momentum flux to obtain a differential equation for the velocity distribution.
4. Integrate these differential equations which yields the momentum flux and velocity distributions for the system under consideration.

Refer to (1) to get more acquainted with shell momentum balance concepts, where this approach has been applied to simple flow problems.

EULERIAN AND LAGRANGIAN COORDINATE SYSTEMS

Conservation equations and the constitutive equation are expressed in terms of relative kinematics, velocities, gradients of velocities, strains, and rates of strain, so it is very important to quantify these variables.

Flow kinematics, the relative motion of fluid particles, can be described by using either a *Lagrangian* or a *Eulerian* description. In the Lagrangian system, the motion of an individual particle is tracked where the coordinates locate the position of a particle and are functions of time. In the Eulerian system, dependent variables such as the velocity vector and pressure are considered functions of fixed spatial coordinates and of time.

The Lagrangian framework is considered a more natural choice to represent the actual kinematics and stresses experienced by fluid particles. However, to solve complex flow problems, this approach is limited (e.g., due to the necessity of tracking particles along unknown streamlines). This system is particularly convenient in flows of viscoelastic fluids with memory that require particle tracking and calculation of deformation and stresses along streamlines. The Eulerian formulation is, in general, more convenient to use because it deals only with local or present kinematics. In most cases, all variables of interest, such as strain, rate of strain, stress, and streamlines, can be obtained from the velocity field. An additional advantage of the Eulerian system is that it involves only time as a variable in unsteady-state flow, whereas the Lagrangian system uses traveling time as a variable, even in steady-state flow. All equations described in subsequent sections have been represented in the Eulerian coordinate system.

EQUATIONS OF CHANGE

The shell balance approach was used to acquaint the beginner with the application of the principle of conservation of momentum to viscous flow problems. It is, however, not always necessary to formulate a momentum balance. It is quicker and easier to start with the equations of continuity and motion in general form and then to simplify them to fit the problem in hand. These two equations describe all problems of the viscous flow of a pure isothermal fluid.

EQUATION OF CONTINUITY

The equation of continuity is developed by applying the law of conservation of mass to a small 3-D volume element within a flowing fluid, as shown in Fig. 3.

$$\begin{aligned}&(\text{Rate of mass accumulation})\\&\quad = (\text{rate of mass in}) - (\text{Rate of mass out})\end{aligned} \tag{11}$$

We consider each of the six faces of this volume element and evaluate the mass entering and exiting the system:

$$\text{Rate of mass in} = \Delta y \Delta z \rho v_x \Big|_x + \Delta x \Delta z \rho v_y \Big|_y + \Delta x \Delta y \rho v_z \Big|_z \qquad (12)$$

$$\text{Rate of mass out} = \Delta y \Delta z \rho v_x \Big|_{x+\Delta x} + \Delta x \Delta z \rho v_y \Big|_{y+\Delta y} + \Delta x \Delta y \rho v_z \Big|_{z+\Delta z} \qquad (13)$$

$$\text{Rate of mass accumulation} = \Delta x \Delta y \Delta z \frac{\partial \rho}{\partial t} \qquad (14)$$

After inserting, Eqs. 12–14 in Eq. 11, dividing the entire equation by $\Delta x \Delta y \Delta z$, and taking the limit of dimensions approaching zero, we get the equation of continuity.

For generality, we have expressed the equation of continuity in vector form:

$$\frac{\partial \rho}{\partial t} = -(\nabla \cdot \rho v) \qquad (15)$$

Here $(\nabla \cdot \rho v)$ is called the "divergence" of $\rho \mathbf{v}$. Vector $\rho \mathbf{v}$ is the mass flux and is the net rate of mass efflux per unit volume. Equation 15 states that the rate of increase of density within a small volume element fixed in space is equal to the net rate of mass influx to the element divided by its volume. Note that the derivation can be performed for a volume element of any shape.

A very special form of the equation of continuity is Eq. 16, which is applicable to fluids of constant density, *incompressible fluids*.

$$(\nabla \cdot \mathbf{v}) = 0 \qquad (16)$$

EQUATION OF MOTION

Using the shell momentum balance approach for the 3-D volume element shown in Fig. 5, an equation of motion is developed. It should be emphasized that momentum is a vector quantity, and hence the equation of motion has components in each of the three coordinate directions x, y, and z, respectively. For simplicity, we consider only the x component here; y and z components can then be handled analogously. Momentum flows in and out of the control volume by two mechanisms; convection, which is a result of the bulk fluid flow and molecular transfer (i.e., viscous transfer), that is, by virtue of velocity gradients.

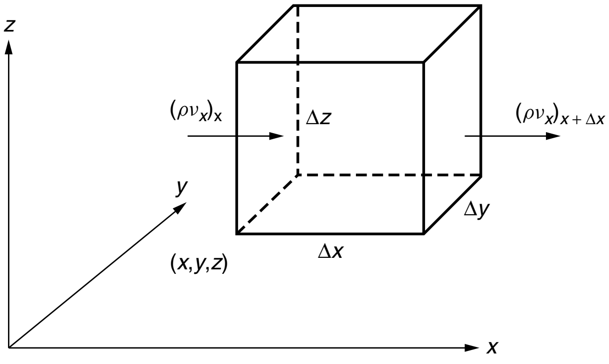

Figure 3. Region of volume in $\Delta x \Delta y \Delta z$ through which fluid is flowing.

The important forces acting on the cubic element in Fig. 4 are the fluid pressure p and the gravitational force g_x. Hence, to write Eq. 10 in mathematical form,

$$\begin{aligned} \text{Rate of momentum in} &= \text{Rate of momentum in by (convection + viscous transfer)} \\ &= \Delta y \Delta z \rho v_x v_x \Big|_x + \Delta x \Delta z \rho v_y v_x \Big|_y + \Delta x \Delta y \rho v_z v_x \Big|_z + \Delta y \Delta z \tau_{xx} \Big|_x \\ &\quad + \Delta x \Delta z \tau_{yx} \Big|_y + \Delta x \Delta y \tau_{zx} \Big|_z \end{aligned} \qquad (17)$$

$$\begin{aligned} \text{Rate of momentum out} &= \text{Rate of momentum out by (convection + viscous transfer)} \\ &= \Delta y \Delta z \rho v_x v_x \Big|_{x+\Delta x} + \Delta x \Delta z \rho v_y v_x \Big|_{y+\Delta y} + \Delta x \Delta y \rho v_z v_x \Big|_{z+\Delta z} + \Delta y \Delta z \tau_{xx} \Big|_{x+\Delta x} \\ &\quad + \Delta x \Delta z \tau_{yx} \Big|_{y+\Delta y} + \Delta x \Delta y \tau_{zx} \Big|_{z+\Delta z} \end{aligned} \qquad (18)$$

$$\begin{aligned} \text{Sum of forces acting} &= \text{body forces} + \text{pressure} \\ &= \Delta y \Delta z \left(p \Big|_x - p \Big|_{x+\Delta x} \right) + \rho g_x \Delta x \Delta y \Delta z \end{aligned} \qquad (19)$$

$$\text{Rate of accumulation of } x \text{ component of momentum} = \Delta x \Delta y \Delta z \frac{\partial \rho v_x}{\partial t} \qquad (20)$$

These stresses developing in viscous transfer can be divided into normal and shear stresses. For example, τ_{xx} is the normal stress, whereas τ_{yx} is the shear stress resulting from viscous forces. At each point in a continuous medium, whether it is solid or fluid, we need six components; each represents a component of force per unit area to define

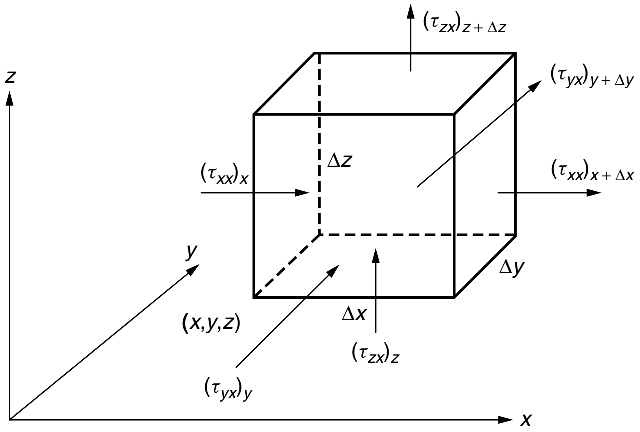

Figure 4. Direction of transfer of x component of momentum through the surfaces.

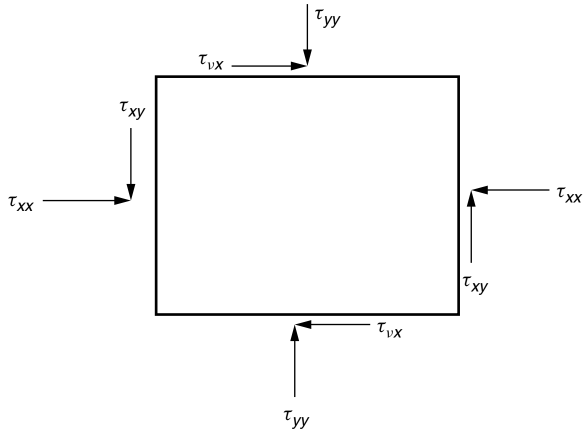

Figure 5. Local stresses acting on a cubic element in an x–y plane.

the local stress completely. The local stresses acting in the x–y plane are shown in Fig. 5.

Using Eqs. 17–20, dividing the entire equation by $\Delta x \Delta y \Delta z$, and taking the limit of the dimensions approaching zero, we get the x component of the equation of motion. Analogously, we can find the components for the y and z components. Adding the three components of equation of motion, the equation of motion can be expressed in vector notation as

$$\frac{\partial}{\partial t}\rho\mathbf{v} = [\nabla\cdot\rho\mathbf{vv}] \quad -[\nabla p] \quad -[\nabla\cdot\boldsymbol{\tau}] \quad +\rho\mathbf{g} \tag{21}$$

$\frac{\partial}{\partial t}\rho\mathbf{v}$	$[\nabla\cdot\rho\mathbf{vv}]$	$-[\nabla p]$	$-[\nabla\cdot\boldsymbol{\tau}]$	$+\rho\mathbf{g}$
Rate of increase of momentum per unit volume	Rate of momentum gain by convection per unit volume	Pressure on element per unit volume	Rate of momentum gain by viscous transfer per unit volume	Gravitational force on element per unit volume

$\boldsymbol{\tau}$ is known as the "stress tensor", $[\nabla\cdot\rho vv]$ represents the rate of loss of momentum per unit volume by fluid flow, and $(\nabla\cdot\rho v)$ represents the rate of loss of mass per unit volume. To use Eq. 21 to determine the velocity distribution, we must insert expressions for various stresses in terms of velocity gradients and fluid properties. For Newtonian fluids, these expressions can be written as

$$\tau_{xx} = -2\mu\frac{\partial v_x}{\partial x} + \frac{2}{3}\mu(\nabla\cdot\mathbf{v}) \tag{22}$$

$$\tau_{yy} = -2\mu\frac{\partial v_y}{\partial y} + \frac{2}{3}\mu(\nabla\cdot\mathbf{v}) \tag{23}$$

$$\tau_{zz} = -2\mu\frac{\partial v_z}{\partial z} + \frac{2}{3}\mu(\nabla\cdot\mathbf{v}) \tag{24}$$

$$\tau_{xy} = \tau_{yx} = -\mu\left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}\right) \tag{25}$$

$$\tau_{yz} = \tau_{zy} = -\mu\left(\frac{\partial v_y}{\partial z} + \frac{\partial v_z}{\partial y}\right) \tag{26}$$

$$\tau_{zx} = \tau_{xz} = -\mu\left(\frac{\partial v_z}{\partial x} + \frac{\partial v_x}{\partial z}\right) \tag{27}$$

Equations 22–27 have been presented with proof in (9) and can be applied to various complex situations. Equations 22–24 are normal stresses and should contain an additional term. For example, Eq. 22 should be written as

$$\tau_{xx} = -2\mu\frac{\partial v_x}{\partial x} + \left(\frac{2}{3}\mu - \kappa\right)(\nabla\cdot v) \tag{28}$$

Similarly, Eqs. 23 and 24 are also written as

$$\tau_{xx} = -2\mu\frac{\partial v_x}{\partial x} + \left(\frac{2}{3}\mu - \kappa\right)(\nabla\cdot v) \tag{29}$$

$$\tau_{xx} = -2\mu\frac{\partial v_x}{\partial x} + \left(\frac{2}{3}\mu - \kappa\right)(\nabla\cdot v) \tag{30}$$

κ is known as "bulk viscosity." The bulk viscosity is zero for low-density monoatomic gases and not too important for dense gases and liquids which are almost incompressible. Experimental measurements to estimate the value of "bulk viscosity" have been discussed by Karim and Rosenhead (10). Formulas to estimate the value of bulk density can be found in (11). When the fluid flows in the x direction between two parallel plates perpendicular to the y direction and v_x is only a function of y, then this set of six equations reduces to Eq. 2 described previously. Equations 22–27 inserted in Eq. 21 yield the equations of motion for a Newtonian fluid with varying density and viscosity.

1. For constant ρ and μ, the equation of motion can be written as

$$\rho\left(\frac{\partial v}{\partial t} + v\cdot\nabla v\right) = -\nabla p + \mu\nabla^2 v + \rho g \tag{28}$$

Equation 28 is the well-known **Navier–Stokes** equations, first developed by Navier in France in 1822.

2. When viscous effects are ignored, Eq. 28 reduces to

$$\rho\left(\frac{\partial v}{\partial t} + v\cdot\nabla v\right) = -\nabla p + \rho g \tag{29}$$

Equation 29 model laminar flow when viscous effects are negligible. These flows are termed Euler, potential, or inviscid flows. Equation 29 is the well-known **Euler** equation.

3. When the motion of a fluid is "very slow," the flow is said to be creeping or Stokes flow. Viscous forces dominate creeping flows and hence, the nonlinear inertia term in Eq. 28 is neglected. Then,

$$\rho\frac{\partial v}{\partial t} = -\nabla p + \mu\nabla^2 v + \rho g \tag{30}$$

Equation 30 is the well-known **Stokes** equations.

The equations of motion together with the equation of continuity, the equation of state $p = p(\rho)$, viscosity models illustrated in previous sections, and the intial and the boundary conditions, complete the problem specification for any viscous flow problem. These equations, when solved mathematically, determine completely the density, pressure, and velocity components in a flowing isothermal fluid.

Boundary Conditions

Boundary conditions are constraints imposed on conservation equations to describe how flow behavior conforms to the surroundings. Boundary conditions come from nature and are mathematical descriptions of the physics at the boundary. These may describe conditions along the boundary dealing with motion, external stresses, rate of mass and momentum flux, boundary values of field variables as well as the relations among them. When the solution involves the time evolution of flow fields, initial conditions are also needed. The number of boundary conditions required to solve a particular problem is determined by the nature of the governing partial differential equation. In general, elliptical equations require boundary conditions on each portion of the boundary; hyperbolic equations require boundary conditions at upstream but not at downstream boundaries; parabolic equations require initial conditions and boundary conditions everywhere except downstream. The Navier–Stokes equations are hyperbolic at high Reynolds numbers and elliptical at low Reynolds numbers. The Euler equation is the upper limit of hyperbolicity, whereas Stokes equations are the lower limit of ellipticity. Different types of boundary conditions in mathematical form have been described by Papanastasiou (12) and are applicable to a wide range of viscous flow problems.

In general, there are three types of boundary conditions:

1. First kind or *Dirichlet* boundary conditions: The value of the dependent variable along the boundary is imposed as the condition. Typical Dirichlet boundary conditions are no-slip boundary conditions for velocity and the specification of inlet/outlet velocities.
2. Second kind or *Neumann* boundary conditions: The normal derivative of the dependent variable is specified along the boundary. Examples are symmetry conditions and free surface and interface stress conditions.
3. Third kind or *Robin's* boundary condition: The dependent variable and its normal derivative are related. Slip wall conditions are typical Robin conditions.

BIBLIOGRAPHY

1. Bird, R.B., Stewart, W.E., and Lightfoot, E.N. (1994). *Transport Phenomena*. John Wiley & Sons, New York.
2. Herschel, W.H., and Bulkley, R. (1926). Flow of Non-Newtonian fluids. In: *Encyclopaedia of Fluid Mechanics*. Gulf Publishers, Houston, TX, Vol. 9, Chap. 4.
3. Oldyroyd, J.G. (1947). A rational formulation of the equations of plastic flow for a Bingham solid. *Proc. Camb. Philos. Soc.* **43**: 100–105.
4. Casson, N. (1959). *Rheology of Disperse Systems*. Pergamon Press, London.
5. Bird, R.B., Armstrong, R.C., and Hassager, O. (1977). *Dynamics of Polymeric Liquids: Fluid Mechanics*. John Wiley & Sons, New York.
6. Slibar, A. and Paslay, P.R. (1959). Retarded flow of Bingham materials. *Trans. ASME* **26**: 107–113.
7. Tanner, R.I. (1985). *Engineering Rheology*. Clarendon Press, Oxford, UK.
8. Wilikinson, W.L. (1960). *Non-Newtonian Fluids*. Pergamon Press, New York.
9. Lamb, H. (1945). *Hydrodynamics*. Dover, New York.
10. Karim, S.M. and Rosenhead, L. (1952). *Rev. Mod. Phys.* **24**: 108–116.
11. Papanastasiou, T.C., Georgiou, G.C., and Alexandrou, A.N. (2000). *Viscous Fluid Flow*. CRC Press, New York.
12. Hirschfelder, J.O., Curtiss, C.F., and Bird, R.B. (1954). *The Molecular Theory of Gases and Liquids*. John Wiley & Sons, New York.

Nomenclature

A	Area
F	Force
$\mathbf{g}$	Gravity force vector
n	Power law index
p	Pressure
$\mathbf{v}$	Velocity vector
t	Time variable
Greek Symbols	
$\dot{\gamma}$	Rate of strain
κ	Bulk viscosity
μ	Viscosity of a fluid
τ	Stress
$\dot{\tau}$	Rate of shear
λ	Elasticity modulus
ρ	Density of a fluid
Subscripts	
a	Stands for apparent
p	Stands for plastic
o	Stands for yield stress
x	Stands for x component
y	Stands for y component
z	Stands for z component

VULNERABILITY MAPPING OF GROUNDWATER RESOURCES

Thomas Harter
University of California
Davis, California

WHAT IS GROUNDWATER VULNERABILITY?

Groundwater vulnerability is a measure of the likelihood that pollution or contamination at the land surface reaches a production aquifer. It is a measure of the "degree of insulation" that natural and man-made factors provide to keep pollution away from groundwater. The term was first used in Europe in the 1960s. In defining groundwater vulnerability, a number of issues must be examined and clarified; all refer to these two questions: vulnerability *of what?* and vulnerability *to what*? Exact definitions of groundwater vulnerability have therefore varied (1).

In principle, all groundwater is vulnerable to human activity; no groundwater is completely isolated from the aboveground environment. The degree of vulnerability depends on environmental conditions, on how we define groundwater and the part of groundwater in which we are interested in (hence the question, "vulnerability *of what*?"), and on the timescale of interest. In addition, it depends on whether or not a vulnerability measure is intended to account also for the presence and type of pollutants (hence, "vulnerability *to what*?").

Vulnerability of What?

Groundwater is produced from a wide variety of hydrogeologic environments. When defining vulnerability, therefore, we need to define the portion of groundwater to be assessed. Because the shallowest groundwater zone is typically the most vulnerable, vulnerability assessments are mostly concerned with the vulnerability of the uppermost aquifer (in a multiaquifer system) or with the water table (in an unconfined aquifer system). Deep groundwater is considered less vulnerable than shallow groundwater because of the longer travel times necessary for a pollutant to reach these deeper aquifers.

Some methods of assessing vulnerability therefore account, directly or indirectly, for the travel time between the water table and a well by considering aquifer permeability or travel-time zones (e.g., 2). Timescales are important in defining vulnerability. The practical question is, should we consider groundwater vulnerable to today's activities only if the recharge time is a few years or also if recharge timescales are tens, hundreds, thousands, or tens of thousands of years? The potential travel time between contaminant source and groundwater or groundwater wells is an important consideration for protection, monitoring, and assessment. Timescales are also important when mapping vulnerability to specific pollutants that become less harmful over time and with travel distance due to radioactive decay, degradation, inactivation (in the case of pathogens), filtration, or strong sorption to soils and aquifer materials.

Vulnerability to What?

An assessment of groundwater vulnerability may or may not include consideration of the presence of pollutants. Vulnerability that is independent of whether or not contaminants (pollutants) are present and focuses primarily on a description of natural environmental conditions is often (though not always) referred to as "susceptibility," "natural vulnerability," "aquifer sensitivity," or "intrinsic vulnerability." Pollution-type-dependent vulnerability, or vulnerability to specific land uses, is sometimes referred to as "specific vulnerability" or "integrated vulnerability" (1).

WHY ASSESS GROUNDWATER VULNERABILITY?

The purpose of vulnerability assessments is to provide a decision-making tool based on the best available data and good scientific judgment. It is generally not intended to be a scientifically rigorous analysis of groundwater resources. The National Research Council (3) identified four general objectives typically achieved by groundwater vulnerability assessments: (1) to facilitate policy analysis and development at the local and regional level, (2) to provide program management, (3) to inform land use decisions, and (4) to provide general education and awareness of a region's hydrogeologic resources.

The usefulness of vulnerability maps is limited (some would say, arguable). The scientific literature lists pros and cons to vulnerability assessments (1). The facts that speak for implementing assessments are that land use decisions have to be made, land management practices need to be sensitive to the risk for groundwater contamination, and not all anthropogenic activities can be carried out isolated from groundwater. Timescales and distinctions based on travel time are important, at least with respect to some pollutants. Timescales are also important in placing groundwater monitoring wells and in scheduling and providing for planning and action–response time if pollution does occur. So a need exists to provide at least some general guidance to land use planners, decision makers, and water users. Carefully prepared vulnerability assessments allow them to make decisions that are economically sensible and at the same time hydrogeologically reasonable.

On the other hand, some may argue against the use of vulnerability maps for land use planning because groundwater flow conditions and the transport properties of the subsurface are too complex to be appropriately captured by any vulnerability tool. A second group of arguments against vulnerability zoning arises from the fact that all groundwater is vulnerable to land pollution. The only geographically differing factor is the timescale for the pollution to reach groundwater.

WHAT FACTORS DETERMINE VULNERABILITY?

We distinguish between intrinsic factors (natural conditions) and extrinsic factors (anthropogenic activities). Among intrinsic factors, the thickness and hydraulic properties of the geologic formations above the aquifer—the unsaturated zone and confining layers above the aquifer—are key factors in determining the vulnerability of an aquifer system. They are the principal natural controls that determine the recharge rate and recharge time to the aquifer. The unsaturated zone also provides key groundwater protection by

- intercepting, sorbing, and eliminating pathogenic viruses and bacteria;
- sorbing and degrading many synthetic organic chemicals; and
- attenuating heavy metals and other inorganic chemicals through sorption and complexation with mineral surfaces within the unsaturated zone and through uptake into plants and crops (e.g., fertilizer).

The potential of the soil and the unsaturated zone to sorb, degrade, or eliminate substances depends on the type of pollutant and therefore is considered only in vulnerability studies that are specific to certain land uses.

The amount of recharge at the land surface is another important factor that determines vulnerability. The higher the recharge, the higher the vulnerability of the aquifer. If climatic conditions are such that little or no recharge occurs at the land surface, downward movement of moisture through the unsaturated zone is very limited, regardless of the hydraulic properties of the unsaturated zone.

In vulnerability assessments that encompass more than just aquifer susceptibility (see above), the presence of contamination sources or possible contaminating activities is considered an important factor. The amount, spatial distribution, and chemical properties of the contaminants (possibly) released into the subsurface are highly specific to individual sources and regions.

CHARACTERIZATION OF GROUNDWATER VULNERABILITY

Numerous schemes have been developed for assessing and mapping vulnerability. These methods can be grouped into three major categories:

- index-and-overlay methods
- process-based computer simulations
- statistical analyses

Index-and-Overlay Methods

Index-and-overlay methods are based on assembling information on the most relevant factors affecting aquifer vulnerability (soil type, geologic formation type, recharge, etc.), which then is interpreted by scoring, integrating, or classifying the information to produce an index, rank, or class of "vulnerability." The scoring, ranking, and integration methods are based on expert opinion rather than processes and are inherently subjective to some degree. In the United States, the most prominent vulnerability assessment method in this category is "DRASTIC" (2). Others include the vulnerability analysis specified by California's Drinking Water Source Assessment Program (4) and the British "GOD" method (1). A brief overview of DRASTIC and GOD is given below.

The advantage of index-and-overlay methods is that they provide relatively simple algorithms or decision trees to integrate a large amount of spatial information into maps of simple vulnerability classes or indexes. The methods are designed to rely on data that are readily available from local, state, or federal agencies, such as information on soils, water level depth, precipitation, and geology. These methods are particularly suitable for use with computerized geographic information systems (GIS), which is a digital form of mapmaking, because they usually involve overlaying and aggregating multiple maps showing soil properties, depth to water table, recharge, and so on.

DRASTIC Vulnerability Mapping (2). DRASTIC is the most popular vulnerability mapping tool in the United States. Put together by a group of experts and the EPA in the mid-1980s, it has been applied to a number of groundwater basins, regions, even states.

The name stands for *d*epth to groundwater, *r*echarge rate, *a*quifer media, *s*oil media, *t*opography, *i*mpact of vadose zone media, and hydraulic *c*onductivity of the aquifer. In DRASTIC, scores of 0 to 10 are assigned to each parameter, 0 means low risk of groundwater contamination; 10 means high risk of contamination. The scores are multiplied by a parameter-specific weight. The weighted scores of all seven parameters are then added for the final DRASTIC score. Two sets of weights can be used: one for general (intrinsic) vulnerability analysis, one for specific vulnerability to pesticides.

GOD (1). GOD is a vulnerability assessment method developed in Great Britain, where most groundwater resources are in hardrock aquifers, primarily sandstone and limestone aquifers. Unconsolidated overburden or soil layers cover the fractured hardrock in many places. Like DRASTIC, GOD is an index-and-overlay method designed to map groundwater vulnerability across large regions based on a few important parameters; GOD stands for *g*roundwater occurrence, *o*verall lithology of the unsaturated zone or overlying aquitard, and *d*epth to the groundwater table. Scores are assigned to each of the three categories and then multiplied to yield a final score. In developing GOD, the method's authors have given particular consideration to the likelihood of fractures or fracture systems to develop in the soils, overburden, or overlying geologic units of the aquifer.

Process-Based Computer Simulations

Process-based computer simulations afford a great amount of realistic complexity and detail to be built into the vulnerability assessment. Computer models can account for complex physical and chemical processes on a very detailed scale. Unlike the two-dimensional maps and map layers used with other methods, computer modeling allows for complete three-dimensional resolution. Geologic and hydrogeologic variations with depth can therefore be reproduced to evaluate their effect on vulnerability. Process-based computer models focus on re-creating the flow and transport patterns within the unsaturated zone or in an actual aquifer and can be used to compute travel times or concentrations of a contaminant in the unsaturated zone or in the aquifer. Computer models do not compute vulnerability directly. Rather, vulnerability is defined as a function of what the computer models simulate. For example, high vulnerability may be defined as any region in the aquifer for which the computer model shows a travel time of less than 5 years.

The most advanced computer models also allow the analyst to compute the uncertainty that is inevitably associated with the computer model predictions due to shortcomings in the database fed into the computer and due to our limited knowledge of the "underground world". These methods use statistical methods to quantify uncertainty. However, even simple computer models can include a sensitivity analysis, that is, an investigation of the sensitivity of the model predictions to the input provided.

Computer models are not commonly used for vulnerability assessment due to their considerable data requirements and the expertise required to implement models. In other words, computer simulation models are rarely an economical alternative for vulnerability mapping. However, computer modeling is an excellent and economical tool for vulnerability mapping if

- a more localized analysis of specific vulnerability to particular land uses (particular contaminants) is required and sufficient data are available or can be collected to prepare the computer model, and
- a number of "what-if" scenarios involving complex processes need to be evaluated for making important land use planning decisions.

Four computer model codes are particularly popular for computing the fate and transport of contaminants as those contaminants travel downward through the unsaturated zone to the water table: VLEACH (EPA), PRZM (EPA), LEACH (Cornell University), and HYDRUS (U.S. Department of Agriculture/Agricultural Research Service Salinity Laboratory, Riverside, California) (for up-to-date information on these codes, see the web pages of these agencies). All four codes (listed here roughly in order of complexity and user expertise required; VLEACH is least complex and HYDRUS the most complex) compute the downward movement of water (flow) and contaminants (transport). These codes are designed specifically to determine the amounts of sorption and degradation that occur in soils and in the unsaturated zone and are typically used to compute the residual, nondegraded mass of a contaminant that arrives, over time, at the water table. For vulnerability assessments, subregions or smaller sections of land with similar unsaturated zone properties are selected and grouped. One simulation is performed for each group or section, and the results are mapped in various ways; a map of the travel time to the water table (perhaps specific to a contaminant), a map of the percent removal of a contaminant within the vadose zone, and so on.

The most common computer code used for groundwater modeling (in all three dimensions) is MODFLOW (U.S. Geological Survey). This model and other 3-D computer models can be used for highly specific site studies of groundwater contamination that are far beyond the purpose of a typical vulnerability assessment. The data requirements for running these models are tremendous and require careful data preparation and data processing, particularly in a fully three-dimensional simulation. Sometimes, groundwater models are used to compute explicitly the complex source area of the water reaching a well. In that case, the same computer model can also be used to compute specific vulnerability for particular contaminants of concern within the source area of the well. If a well-calibrated and well-documented groundwater model exists, it may be preferrable to use the groundwater model to address specific vulnerability issues rather than using a basic index or a scoring-based vulnerability assessment.

Statistical Methods

Statistical methods are used to quantify the risk of groundwater pollution by determining the statistical dependence or relationship between observed contamination, observed environmental conditions that may or may not characterize vulnerability (e.g., unsaturated zone properties, recharge), and observed land uses that are potential sources of contamination (e.g., fertilizer applications, septic tank occurrence). Once a model of this dependence or relationship has been developed from the statistical analysis, it can be used to predict—in a similar area elsewhere—the chance or risk of contamination. Such an application requires, of course, knowledge of significant environmental conditions for that area. When statistical methods are used, the risk of contamination is essentially a quantitative measure of "vulnerability." The higher the contamination risk, the higher the vulnerability.

In principle, statistical methods are not much different from index-and-overlay methods. Both establish a relationship between inherent natural conditions and groundwater vulnerability (which the statisticians refer to as groundwater contamination risk). In overlay methods, the relationship is established by a team of experts. In statistical methods, the relationship is established by statistical analysis. The advantage of the statistical method is that the statistical significance can be explicitly calculated, which provides a measure of uncertainty or certainty of the model. The disadvantage is that statistical methods are difficult to develop and, once established, can be applied only to regions that have environmental conditions similar to those of the region for which the statistical model was developed.

Few statistical methods have been developed for vulnerability assessment, primarily because they rely on large data sets from regions already contaminated. The key to the development of a statistical method is a large, high quality data set of a specific contaminant or several specific contaminants commonly found in a region. That data set is then correlated in one fashion or another to specific properties of the subsurface (depth to water, soil permeability, slope, hydraulic conductivity, etc.) to create a statistical predictive model that can be used to make a quantitative statement about the contamination risk, such as, "if this kind of unsaturated zone, aquifer, and land use properties exist in a location, then the risk for groundwater contamination is X%, with a confidence interval of plus or minus Y%." Some examples follow.

CALVUL. In California, a statistical method has been developed by the California Department of Pesticide Regulations (DPR) to determine the specific vulnerability of groundwater to pesticide residues. The method is nicknamed CALVUL (California Vulnerability approach; (5)). DPR uses the vulnerability analysis to determine, for each section of agricultural land in California, whether groundwater in that section of land is vulnerable or not. If it is vulnerable, the land section (square mile) becomes part of a groundwater protection area with restricted pesticide use. The DPR method distinguishes two types of vulnerability: (1) vulnerability to pesticide leaching in coarse soils with a shallow water table and (2) vulnerability to pesticide runoff over hardpan soils (and subsequent leaching in dry wells). Pesticide management practices depend on whether pesticide leaching or pesticide runoff is of concern.

A pesticide database for the San Joaquin Valley was used to determine the statistical relationship between pesticide residue in groundwater and a large number of soil (unsaturated zone) properties, including water holding capacity, texture, organic matter content, permeability, shrink-swell potential, slope, and infiltration. By determining which soil and geographic properties are good predictors of pesticide residue in groundwater, the statistical model can be used to identify sections of land that have similarly "vulnerable" soil and geographic conditions but no past or current pesticide residue detections (5). Presumably, the most vulnerable areas are identified and can be targeted for implementation of mitigation management practices.

Texas Case Study. In Texas, a statistical analysis similar to DPR's San Joaquin study, albeit based on a different set of statistical tools, was developed by Evans and Maidment (6). Instead of pesticide residues in groundwater, this method uses an analysis of nitrate in groundwater as the basis for delineating vulnerable groundwater areas. As in CALVUL, the assumption made by the developers of the method was that where high contamination exists, the aquifer is more vulnerable than elsewhere. Nitrate was used as the target variable because a large amount of data was available throughout Texas and because nitrate is neither sorbed nor significantly degraded in groundwater (except under anaerobic conditions). The analysis was done by using data from the entire state. From a 30-year statewide database, the probability (risk) that a well would have nitrate levels in excess of some threshold value (1, 2, 5, or 10 mg/L NO_3-N) was computed separately for each analysis unit. The exceedance probability is used as a quantifiable measure of vulnerability; the higher the chances that nitrate is found in wells within an analysis unit (area), the higher the vulnerability. The study illustrates that overall results of the statistical analysis follow the general patterns laid out by the index-and-overlay methods, albeit at considerably higher expense.

National Groundwater Nitrate Vulnerability Assessment. The U.S. Geological Survey recently published a national assessment of the probability of nitrate contamination in recently recharged groundwater (7). The assessment relied on a specialized statistical regression analysis among major factors affecting the specific vulnerability of groundwater to nitrate and the measured occurrence of elevated nitrate levels in wells. The data set used to develop the regression model included wells from 20 major groundwater basins across the United States; each contained 20–30 test wells. Nitrogen fertilizer loading and percent cropland were significant extrinsic (nitrogen use related) factors. The presence of well-drained soils, depth to a seasonally high water table, and the presence of unconsolidated sand and gravel aquifers were the key intrinsic factors that determine groundwater nitrate vulnerability.

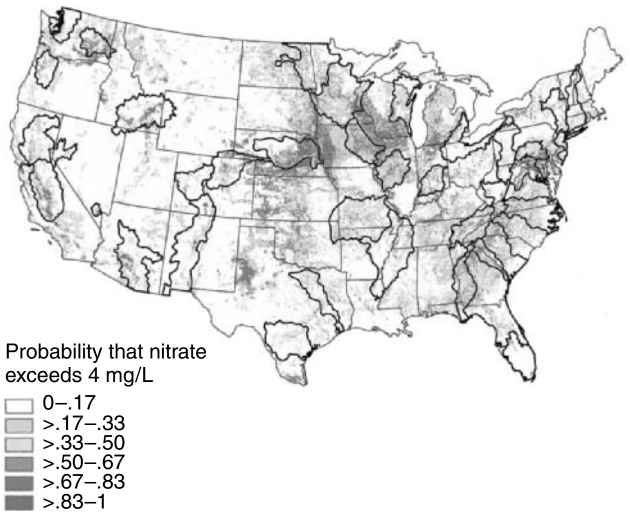

Figure 1. Groundwater vulnerability map for the United States, where vulnerability is here defined as the probability that nitrate concentration in groundwater exceeds 4 mg/l. (reprint from (7), also published at http://water.usgs.gov/nawqa/nutrients/pubs/est_v36_no10/fig3.html).

The regression model was validated (tested) on a large set of well records not used in developing the regression model and then applied to predict the probability of elevated nitrate (a measure of nitrate vulnerability) in groundwater throughout the United States at a resolution of 1 km^2 (Fig. 1, from Reference 7). The study demonstrates the advantage of the statistical method over index-and-overlay methods. The statistical assessment model uses objective, quantifiable measures to assess vulnerability. Its predictive capability has been quantified as well (through the validation test).

WHAT ARE THE LIMITATIONS OF VULNERABILITY ASSESSMENTS?

Because of the implied imprecise nature of vulnerability assessments and the inevitable subjectiveness of the underlying interpretive scheme, the National Research Council (3) issued "three rules of groundwater vulnerability." These rules, or limitations, should be spelled out explicitly with every vulnerability assessment:

- All groundwater is vulnerable to some degree.
- Uncertainty is inherent in all vulnerability assessments.
- There is a risk that the obvious may be obscured and the subtle may become indistinguishable.

The last refers to the danger, especially when using complex vulnerability assessment tools, that in light of the final vulnerability index or ranking, one may lose sight of the data used for the analysis and of the assumptions underlying vulnerability assessment schemes. The results of vulnerability assessments may mask real differences between specific sites or aquifer conditions.

AFTER THE ASSESSMENT: VERIFICATION AND POSTAUDIT

Verification refers to some independent procedure that can verify the results of the vulnerability analysis. *Postaudit* is essentially the same as verification but typically occurs years later, when additional data have been sampled that can be compared with the predictions made by the vulnerability analysis.

Verification and postaudits of vulnerability assessments can be done in many different ways. The most common approach, particularly for verifying assessments done with index-and-overlay methods, is to compare the vulnerability map with the actual occurrence of some common pollutant in groundwater. Typical pollutants used are nitrate and pesticides. However, such verification works well only where the appropriate pollution source is actually present and has been present for some time.

BIBLIOGRAPHY

1. Robins, N.S. (Ed.). (1998). *Groundwater Pollution, Aquifer Recharge and Vulnerability*. Geological Society Special Publication No. 130, Geological Society, London.
2. Aller, L., Bennett, T., Lehr, J.H., and Petty, R.J. (1985). DRASTIC—A standardized System for Evaluating Ground Water Pollution Potential Using Hydrogeologic Settings. U.S. EPA, Robert S. Kerr Environmental Research laboratory, Office of Research and Development, EPA/600/2-85-018.
3. National Research Council. (1993). *Ground Water Vulnerability Assessment, Contamination Potential under Conditions of Uncertainty*. National Academy Press, Washington, DC.
4. California Department of Health Services. (2000). Drinking Water Source Assessment and Protection Program. Available: http://www.dhs.ca.gov/ps/ddwem/dwsap/DWSAP_document.pdf.
5. Troiano, J., Spurlock, F., and Marade, J. (1999). Update of the California Vulnerability Soil Analysis for Movement of Pesticides to Ground Water. Department of Pesticide Regulation, Sacramento, CA 95814-3510, Document EH 00-05. Available: http://www.cdpr.ca.gov/docs/empm/pubs/ehapreps/eh0005.pdf.
6. Evans, T.A. and Maidment, D.R. (1995). A Spatial and Statistical Assessment of the Vulnerability of Texas Groundwater to Nitrate Contamination. Technical Report CRWR 260, Center for Research in Water Resources, Bureau of Engineering Research, The University of Texas at Austin, Austin, TX.
7. Nolan, B.T., Hitt, K.J., and Ruddy, B.C. (2002). Probability of nitrate contamination of recently recharged groundwaters in the conterminous United States. *Environ. Sci. Technol.* **36**(10): 2138–2145.

READING LIST

Blodgett, R. (1993). Drinking Water Vulnerability Assessments in Texas. *Proc.*, 1993 *Ann. Conf.*, American Water Works Association, Denver, CO.

Holtschlag, D.J. and Luukkonen, C.L. (1998). Assessment of Ground-Water Vulnerability to Atrazine Leaching in Kent County, Michigan: Review, Comparison of Results of Other Studies, and Verification, Water-Resources Investigations Report 98-4006. U.S. Geological Survey, Lansing, MI.

Rupert, M.G. (2001). Calibration of the DRASTIC ground water vulnerability mapping method. *Ground Water* **39**(4):625–630.

Snyder, D.T., Wilkinson, J.M., and Orzol, L.L. (1998). Use of a Ground-Water Flow Model with Particle Tracking to Evaluate Ground-Water Vulnerability, Clark County, Washington, Water-Supply Paper 2488. U.S. Geological Survey, Denver, CO.

Vrba, J. and Zoporozec, A. (Eds.). (1994). *Guidebook on Mapping Groundwater Vulnerability*. International Association of Hydrogeologists: International Contributions to Hydrogeology Vol. 16, Verlag Heinz Heise, Hannover, Germany.

WATER/ROCKS INTERACTION

Henry Teng
The George Washington University
Washington, DC

Water–rock interaction is a process whereby solid earth materials and aqueous phases in the Earth system exchange matter and, ultimately, energy through chemical and biochemical reactions. The extent of reactions is determined by how far the system is away from equilibrium, and the rate of reactions is controlled largely by the kinetics of mineral-water interfacial processes. In a broad sense, these reactions can be classified into two categories based on the reactants involved: (1) those between H_2O and the mineralogical components of rocks and (2) those between solutes dissolved in natural waters and the minerals. Water–rock interactions participate in many geological processes such as ore deposit formation, weathering, and pedogenesis. The outcome of these interactions plays an important role in regulating geochemical cycling of elements, the atmospheric CO_2 level, and the chemistry and nutrient availability of surface/near-surface waters at both local and global scales.

Water–rock interactions are heterogeneous reactions because of the involvement of multiple phases (that is, solid and liquid phases in most cases). Despite the complexity associated with the heterogeneity and variations in mineral assemblages, water chemistry, and environmental conditions, reactions between aqueous phase and rocks at the fundamental level can be regarded as consisting of only two types of mechanistic movement (1), chemical species attachment to and detachment from mineral surfaces. Individual or combined development of the two basic actions over time yields three primary types of water–rock interactions: adsorption (including ion exchange), dissolution, and precipitation. Following is a detailed discussion of the three types of reactions.

ADSORPTION

Chemical sorption results directly from the occurrence of undercoordinated sites and associated charges on mineral surfaces. Development of surface charge in aqueous environments can be understood as a consequence of surface complexation, by which water molecules form chemical bonds with the undercoordinated surface atoms (2). These surface-bound water molecules are then subject to a proton transfer process that shifts hydrogen ions onto neighboring surface anions. This dissociative sorption of water molecules results in one or more types of surface species, depending on the chemical composition of the minerals in question. For (hydr)oxide materials, hydroxylated surface cation centers $>X(\mathrm{OH})^0$, where $>$ symbolizes lattice bonds and X represents the cationic component, seem to be the only concern. For salt-type minerals X_mA_n, an additional, second-type nonhydroxylated surface anion site $>A\mathrm{H}^0$ has to be taken into consideration.

When present in solution environments, the surface hydroxyl groups can undergo protonation or deprotonation to form charged surface species. The relative abundance of protonated and deprotonated species thus governs the surface electric properties. Consider an oxide mineral XO. The sign and quantity of surface charges incurred in pure water will be controlled by the following reactions

$$\equiv X\mathrm{OH_2}^+ \Leftrightarrow \equiv X(\mathrm{OH})^0 + \mathrm{H}^+ \qquad K_{\mathrm{a1}}^{\mathrm{app}} \tag{1a}$$

and

$$\equiv X(\mathrm{OH})^0 \Leftrightarrow \equiv X\mathrm{O}^- + \mathrm{H}^+ \qquad K_{\mathrm{a2}}^{\mathrm{app}} \tag{1b}$$

where $K_{\mathrm{a1}}^{\mathrm{app}}$ and $K_{\mathrm{a2}}^{\mathrm{app}}$ are the apparent equilibrium constants of reactions 1a and 1b. The participation of hydrogen ions in the above reactions indicates that the relative concentrations of the protonated and deprotonated surface species, and hence the amount of surface charges, are controlled by the pH of the aqueous environment. The pH at which there are equal amounts of $>X\mathrm{OH_2}^+$ and $>X\mathrm{O}^-$ (i.e., an uncharged surface) is defined as the pH of zero net proton charge, below which the surface is positively charged for $>X\mathrm{OH_2}^+$ is more abundant than $>X\mathrm{O}^-$ and above which the surface is negatively charged because of the excess of $>X\mathrm{O}^-$.

In the presence of other sorbing ions, say $M^{(\mathrm{n}+1)+}$ and $L^{(\mathrm{n}+1)-}$, in solutions, competitive complexation becomes operative as M can exchange for H and L for OH at the $>X(\mathrm{OH})^0$ sites to result in the formation of surface complexes such as $>X\mathrm{O}M^{\mathrm{n}+}$ and $>XL^{\mathrm{n}-}$. Surface complexation under this condition becomes adsorption where both chemical bonding and electrostatic interaction take place between solute particles and surface atoms (3). Empirical approaches to adsorption employ simple equations to describe the adsorption isotherm (relationship between the solution concentration or activity and the surface concentration of an adsorbate), without concerning the specific interfacial interactions between adsorbate and adsorbent. For example, the Langmuir isotherm, one of the most common empirical models, simply treats adsorption as a free surface site $>S$ occupied by an adsorbate A to form a sorbed site $>SA$:

$$>S + A \Leftrightarrow >SA \tag{1c}$$

The adsorption constant K_{ad} given by the mass law is

$$K_{\mathrm{ad}} = \frac{[>SA]}{[>S][A]} \tag{1d}$$

Because the total amount of surface sites is $[>S]_{\mathrm{tot}} = [>S] + [>SA]$, Eq. 1d can be rewritten as

$$[>SA] = [>S]_{\mathrm{tot}} \frac{K_{\mathrm{ad}}[A]}{K_{\mathrm{ad}}[A] + 1} \tag{1e}$$

High concentrations of $[A]$ lead to $[>SA] \cong [>S]_{\mathrm{tot}}$, whereas low values of $[A]$ reduces Eq. 1e to a linear form of $[>SA] \cong K_{\mathrm{d}}[A]$. Another common empirical adsorption model, the Freundlich isotherm, uses a logarithmic relationship between $[>SA]$ and $[A]$ so that

$$[>SA] = k_f[A]^n \tag{1f}$$

where both k_f and n are constants. The Freundlich model can be thought of as the Langmuir isotherm integrated over many types of surface sites (4).

In the surface complexation model, adsorption is treated as a chemical reaction

$$>X(\mathrm{OH})^0 + M^{(\mathrm{n}+1)+} \Leftrightarrow >X\mathrm{O}M^{\mathrm{n}+} + \mathrm{H}^+ \tag{1g}$$

The equilibrium constant K^{app}

$$K^{\mathrm{app}} = \frac{[>X\mathrm{O}M^{\mathrm{n}+}][\mathrm{H}^+]}{[>X\mathrm{OH}^0][M^{(\mathrm{n}+1)+}]} = K^{\mathrm{int}}K^{\mathrm{coul}} \tag{1h}$$

now has contributions from chemical (intrinsic) as well as electrostatic (coulombic) interactions between the reactants. Quantitatively, $K^{\mathrm{coul}} = \exp(F\Delta\mathrm{Z}\Psi)$, where $\Delta\mathrm{Z}$ is the change of charges in surface species because of adsorption. The surface potential Ψ is a function of surface charge σ, and the charge-potential relationship can be derived from electrochemistry theories. A common one

$$\sigma = (8RT\varepsilon\varepsilon_0 I)^{1/2}\sinh(z\Psi F/2RT) \tag{1i}$$

is given by the Gouy–Chapmen theory (diffuse double-layer model). Here ε is the relative dielectric constant of water, ε_0 is the permittivity of vacuum, I is the ionic strength of the bulk solution, and z is the valence of the symmetrical background electrolyte. The quantity of surface charge can be computed from the concentration of charged surface species. For Eq. 1g,

$$\sigma = F[\equiv X\mathrm{O}M^{(\mathrm{n}-1)+}]/AS \tag{1j}$$

where A and S are the specific area (m^2/g) and solid concentration (g/l) of the adsorbent material, respectively. Overall, assuming the total amount of surface sites in Eq. 1g is approximated by $[>X\mathrm{O}]_{\mathrm{tot}} = [>X(\mathrm{OH})^0] + [>X\mathrm{O}M^{\mathrm{n}+}]$, one can rearrange Eq. 1h into

$$[>X\mathrm{O}M^{\mathrm{n}+}] = [>X\mathrm{O}]_{\mathrm{tot}}\left[\frac{\exp(\mathrm{nF}\Psi)K^{\mathrm{int}}[M^{(\mathrm{n}+1)+}]}{\exp(\mathrm{nF}\Psi)K^{\mathrm{int}}[M^{(\mathrm{n}+1)+}] + [\mathrm{H}^+]}\right] \tag{1k}$$

It becomes apparent that, after inspecting Eqs. 1e, 1f, and 1h, the surface complexation approach is advantageous over the empirical isotherms in that the formal provides additional insight into the dependence of adsorption upon pH and ionic strength (through Eq. 1i). Moreover, the use of K^{int} in surface complexation models makes it possible to compare results from different experiments because, unlike empirical parameters such as K_{ad} (Eq. 1d) and K_{f} (Eq. 1f), the intrinsic surface complex formation constant is not experimental condition dependent.

It is noteworthy to point out that although surface complexation reactions can occur on any mineral crystals, effective adsorption requires small grain sizes. Because of the inverse relationship between particle size and specific surface area (defined as the ratio of surface area to volume of a particles), small grains (<1 μm) tend to have a significant percentage of surface-bound atoms. Consequently, small particles have large numbers of active surface sites available to undergo complexation reactions. Micron-scale suspended particles and sub-micron-scale colloids of minerals are a common occurrence in surface waters and groundwaters (Ref. 5). It is estimated that 80% of the materials transported by rivers and discharged into the oceans is in the form of suspended particles. Although the exogenic cycle of the major cations and anions is controlled by solution process, in most cases the geochemistry of trace elements in natural waters is governed by adsorption on separate mineral grains (6).

DISSOLUTION

Minerals dissolve because the chemical potential of solute in solid and liquid phases has not reached equality, which indicates that the driving force of any dissolution reaction may be estimated by the free energy difference, ΔG, between reactants and products. As the reaction approaches equilibrium conditions where ΔG becomes zero, dissolution should cease to proceed.

Consider the dissolution of a solid compound ML,

$$ML \Leftrightarrow M^{n+} + L^{n-} \tag{2a}$$

The free energy change in this system is described by

$$\Delta G = \Delta G^{\mathrm{o}} + RT\ln(a_{\mathrm{M}}a_{\mathrm{L}}) \tag{2b}$$

where ΔG^{o} is the standard free energy change, R is the gas constant, and T is the temperature; a_{M} and a_{L} are the activity of aqueous species M^{n+} and L^{n-}, respectively. At equilibrium conditions, $\Delta G = 0$ and Eq. 2b can be rewritten as

$$a_{\mathrm{M}}a_{\mathrm{L}} = \exp(-\Delta G^{\mathrm{o}}/RT) \tag{2c}$$

Because ΔG^{o} can be computed as the difference between the standard free energy of formation for the dissolved species and the solid phase, Eq. 2c yields a numerical value defined as the standard equilibrium constant or solubility constant, K_{sp}, of the compound ML.

When the solid–fluid interface is concerned, surface energies and grain sizes have to be taken into account to define the equilibrium conditions. Consider a mineral crystal of any geometry with any number of faces that are not related by any symmetry operations. At constant temperature and pressure, the free energy change in a system containing the crystal in contact with an undersaturated solution can be described by

$$dG = -\sum_i \nu_i\mu_i dn + \sum_j \gamma_j dA_j \tag{3a}$$

where μ_i is the chemical potential of the ith species in the liquid phase; ν_i is the stoichiometric coefficient of the ith species in the solid phase; n is the number of moles of solute that are involved in the dissolution reaction; and γ_j and A_j are the surface free energy and the surface area of the jth type of faces, respectively. Using the definition of chemical potential,

$$\mu_i = \mu_i^{\mathrm{o}} + RT\ln\alpha_i \tag{3b}$$

one can rewrite Eq. 3a as

$$\frac{dG}{dn} = -\sum_i \nu_i\mu_i^{\mathrm{o}} - RT\ln\prod_i \alpha_i^{\nu_i} + \sum_j \gamma_j(dA_j/dn) \tag{3c}$$

where μ_i^{o} and α_i are the standard chemical potential and the activity of the ith species, respectively. At equilibrium conditions, $dG/dn = 0$, $\sum_i \nu_i\mu_i^{\mathrm{o}} = -RT\ln K_{\mathrm{sp}}$, it then follows that the crystal solubility $K_b = \prod_i \alpha_i^{\nu_i}$ can be expressed as

$$\ln K_b = \ln K_{\mathrm{sp}} + \frac{1}{RT}\sum_j \gamma_j(dA_j/dn) \tag{3d}$$

Equation 3d states that the solubility of a crystal is always greater than the standard solubility product of the solute. Rewrite Eq. 3d using the volume v and density ρ of the crystal and the molecular weight W of the solute, one gets

$$\ln K_b = \ln K_{\mathrm{sp}} + \frac{W/\rho}{RT}\sum_j \gamma_j(dA_j/dv) \tag{3e}$$

For isotropic crystals with any characteristic dimension l (i.e., length, width, or height), $\Sigma\gamma_j(dA_j/dv)$ can be approximated by $(2\gamma)/(3l)$ and Eq. 3e can be simplified to

$$\ln K_b = \ln K_{\mathrm{sp}} + \frac{2W/\rho}{3RT}(\gamma/l) \tag{3f}$$

Equation 3f is the statement of the classic Gibbs–Thomson relationship, which predicts that the difference between crystal solubility and K_{sp} is size-dependent, and the solubility of any crystals can only be approximated by the standard solubility product of the solute when the crystal is infinitely large. In a practical sense, the "infinite" size is usually no larger than a few micrometers. For anisotropic minerals, each crystal face has to be considered individually because the surface energy is face-specific and Eq. 3e cannot be easily simplified. Nonetheless, the conclusions drawn from Eq. 3f are just as well applicable.

Dissolution not only takes place in the form of dissociation reactions (Eq. 2a), but it can also result from reduction–oxidation reactions where the number of valence electrons changes on one or more reactant elements. Redox reactions particularly affect water–rock interactions involving minerals that contain multivalence state elements such as S and Fe. However, because of the relatively few minerals concerned, the contribution of redox chemistry to the exogenic cycle of elements is considered auxiliary (7). Most natural redox reactions are mediated by microbial communities or induced by solar radiation. In the formal scenario, oxidized mineral species are often used as electron acceptors for metabolism, whereas in the latter situation, electrons in the reactants' molecular orbital are excited photochemically to pair up with more electrons or to be released to lower the energy state of the system.

Traditionally, geochemists treat redox reactions as reactants' gaining or losing oxygens. The more modern approaches, on the other hand, choose to use electron as a basic component. In this convention, reductants are electron donors that release e^- and oxidants electron acceptors that receive e^-. The activity of electron can be expressed as a potential (Eh, in volts) or as an activity unit ($\mathrm{p}e = -\log a_{e^-}$). Quantitatively, Eh and pe can be converted to each other through

$$\mathrm{p}e = \frac{F}{2.303RT}Eh \tag{4a}$$

where F is the Faraday constant. In addition, Gibbs free energy and the activity of electrons are related by

$$\Delta G = -nFEh \tag{4b}$$

where n is the number of electrons in the reaction. If a redox dissolution reaction can be written as

$$\text{Ox (solid)} + ne^- \Leftrightarrow \text{Rd (aqueous)} \tag{4c}$$

the standard free energy expression (Eq. 2b) becomes

$$Eh = E^{\mathrm{o}} - \frac{RT}{nF}\ln a_{\mathrm{Rd}} \tag{4d}$$

or

$$\mathrm{p}e = \mathrm{p}e^{\mathrm{o}} - \frac{1}{n}\log a_{\mathrm{Rd}} = \frac{1}{n}\log K^{\mathrm{o}} - \frac{1}{n}\log a_{\mathrm{Rd}} \tag{4e}$$

where E^{o} and pe^{o} are the activity of electrons under standard conditions. Notice that the equilibrium constant of redox reaction K^{o} can still be calculated from the difference between the standard free energy of formation for the dissolved species and the solid phase (i.e., Eq. 2c).

In general, the rate R of dissolution reactions can be related to ΔG by (8)

$$R = k(1 - e^{\Delta G/RT}) \tag{5a}$$

where k is the rate constant. When the reaction is sufficiently close to equilibrium (i.e., $|\Delta G| \ll RT$), Eq. 5a reduces to

$$R = -(k/RT)\Delta G \tag{5b}$$

This expression depicts a linear relationship between the driving force and the kinetics of dissolution reactions at near-equilibrium conditions. When applying Eq. 5b to water–rock interactions, however, caution has to be used because the dissolution of many minerals in aqueous environments is initiated by the formation of etch pits on crystal surfaces. Etch pit nucleation usually requires one to overcome an energy barrier that decreases with descending saturation state (9,10). At very close to equilibrium conditions, the saturation state may be too high to bring down the barrier to the extent that allows pit formation, which indicates that dissolution kinetics may change discretely, rather than in a linear fashion, with decreasing saturation state (11,12).

A unique characteristic of mineral dissolution is that some (alimino)silicate minerals may dissolve in an incongruent or nonstoichiometric manner. This type of reaction is best exemplified by the dissolution process of Na- and K-feldspars in acidic environments. It is generally agreed that the dissolution starts with an initial fast ion exchange of hydrogen ions for alkali and alkaline earth metals, followed by a slow process that disrupt the Si-O-Al-O framework. The disruption seems to proceed by preferential release of aluminum (i.e., nonstoichiometric surface leaching) so SiO_2 accumulates *in situ*. Observations by numerous workers indeed reveal the presence of a Si-rich surface coating on feldspar crystal faces when the minerals are exposed to solutions with pH <3–5 (Ref. 13). What researchers are not in complete agreement on is to what extent the Si-rich layer is originated from surface alteration. Although results from investigations up to the end of the last century imply that most coating (up to several thousand angstroms in depth) forms through nonstoichiometric leaching of the underlying minerals (Ref. 13), recent studies using more modern experimental techniques suggest otherwise. Synchrotron radiation and high-resolution transmission electron microscopy show that the incongruent dissolution may be operative only for, at most, the first few unit cells at the mineral-water interface, and the thick surface alteration thus forms largely by interfacial dissolution–reprecipitation (14,15).

PRECIPITATION

Much of the discussion on reaction driving forces in dissolution section is directly transferable to precipitation reactions. However, precipitation is not a pure mirror image of dissolution in that the former is a combination of two separate but sequentially occurring processes, nucleation and growth. Nucleation refers to the atomic process in which solute components congregate to form a thermodynamically stable cluster, whereas the following accretion of the cluster is termed growth. Both nucleation and growth require the occurrence of supersaturation, but each of the two processes has its unique physiochemical controls.

Homogeneous nucleation occurs when a thermodynamically stable 3-D cluster is born directly from solution. On the other hand, solute particles can assemble onto an existing surface, and hence through heterogeneous nucleation, to form a stable 2-D cluster. No matter how nucleation takes place, the classic nucleation theory requires the sizes of the clusters to satisfy the following energetics for the atomic assemblage to become stable nuclei: A surface energy increase associated with the emergence of the solid phase precisely offsets the chemical potential decrease in the solution because of losing solute. Mathematically, free energy change upon formation of a spherical cluster of radius r can be expressed as

$$\Delta G = 4\pi r^2\overline{\gamma} - (4/3\pi r^3/\overline{\mathrm{v}})\Delta\mu \tag{6a}$$

where $\overline{\gamma}$ is the average surface free energy and $\overline{\mathrm{v}}$ is the specific molecular volume of the solid. The minimal size of a stable cluster

$$r_\mathrm{c} = 2\overline{\mathrm{v}}\overline{\gamma}/\Delta\mu \tag{6b}$$

can be found by solving the equation $d(\Delta G)/dr = 0$. Any cluster with $r < r_\mathrm{c}$ is unstable and will be dissolved, whereas the ones with $r > r_\mathrm{c}$ will move on to the growth stage. Inserting Eq. 6b into 6a, one has an expression that estimates the energy barrier, ΔG^*, for nucleation:

$$\Delta G^* = \frac{16\pi\overline{\gamma}^3\overline{\mathrm{v}}^2}{3\Delta\mu} \tag{6c}$$

The nucleation rate R_n, according to the conventional rate theory, is then available through

$$R_\mathrm{n} = \overline{A}\exp\Big(\frac{-\Delta G^*}{k_BT}\Big) \tag{6d}$$

where k_B is the Boltzmann constant and $\overline{A}$ is a constant related to the efficiency of particle collision (16). All implications from Eqs. 6a–6d apply to heterogeneous nucleation, except that, quantitatively, ΔG^* is usually smaller when nucleation occurs on a substrate surface. This result leads to a general conclusion that heterogeneous nucleation often occurs at a lower supersaturation than does homogeneous nucleation.

Precipitation enters the growth stage with the onset of enlargement of the minimal sized clusters. Nuclei increase their size through attachment of individual solute particles onto the cluster surfaces. During growth, crystallographic control becomes strongly operative. That is, faces that are more closely packed will prevail, whereas the ones with smaller packing density will be phased out. The more closely packed faces usually are the flat face (F-face) on the ideal Kossel crystals (17); the growth of flat surfaces is commonly recognized as following a layer-by-layer mechanism where solute particles attach to kink sites along steps (i.e., the edges of individual layers) generated either by crystal imperfection or surface nucleation. The net rate of step advance, v_s, is determined by the difference between solute fluxes attaching to and detaching from the step kinks. This relationship can be described mathematically by (18,19)

$$v_s = (d^2/\lambda_\mathrm{o})\upsilon\{\exp[(\mu_\mathrm{S} - \overline{\mathrm{E}})/k_BT] - \exp[(\mu_\mathrm{L} - \overline{\mathrm{E}})/k_BT]\} \tag{6e}$$

Here, d is the size of the attaching species in the lattice, λ_o is the interlink distance, and υ is the fundamental frequency (20) that can be approximated by k_BT/h_p where h_p is the Planck constants; μ_L and μ_S are the chemical potentials of solute in liquid (solution) and solid (crystal), respectively, and $\overline{\mathrm{E}}$ is the free energy of an intermediate activated complex.

Equation 6e can be further rearranged into

$$v_s = (d^2/\lambda_\mathrm{o})\upsilon\exp[(\mu_\mathrm{L} - \overline{\mathrm{E}})/k_BT]\{\exp[(\mu_\mathrm{S} - \mu_\mathrm{L})/k_BT] - 1\} \tag{6f}$$

Considering that the term $\exp(\mu_\mathrm{L} - \overline{\mathrm{E}})$ has an entropy factor, $\exp(\Delta\overline{\mathrm{S}}/k_B)$, and an energy factor, $\exp(-\Delta E/k_BT)$, and that $(\mu_\mathrm{S} - \mu_\mathrm{L})$ is essentially the driving force for growth and is quantitatively related to supersaturation, σ, by

$$\Delta\mu = k_BT\ln(a/a_e) = k_BT\sigma \tag{6g}$$

one can rewrite Eq. 6f as

$$v_s = \{(d^2/\lambda_\mathrm{o})\upsilon\exp(-\Delta\overline{S}/k_B)\exp(-\Delta E/k_BT)\}/ k_BT]\}\{(a - a_e)/a_e\} \tag{6h}$$

Notice that step velocity has a linear dependence on the difference between solute activity and its equilibrium value $(a - a_e)$. By convention, the quantity between the first two curly parentheses in Eq. 6h is termed the kinetic coefficient, β, of step growth. This parameter, as shown here, is characterized by the activation energy barriers, ΔE, and the corresponding entropic changes for the formation of activated intermediate complexes.

Complementary to the classic approach of atom-by-atom growth, recent findings suggest that crystals can also grow by oriented aggregation of nano-sized clusters (21,22). In this growth mode, adjacent nanocrystallites coagulate after spontaneous self-organization that situates the two parties to a common crystallographic orientation. The new, larger particles gain thermodynamic stability by forming chemical bonding between the original clusters to eliminate the interface between them. It is entirely possible that different growth mechanisms operate at different conditions. When nucleation rates are low and nuclei are far apart in the media, growth may only be accomplished by atom-by-atom attachment on growing surfaces. When sufficient nuclei are present, they can coalesce, through orientated coagulation, to form larger crystals, as long as the system is shifting toward a lower energy state.

It is important to be informed that the kinetics of interfacial reactions such as dissolution and growth can be controlled either by the pace of the actual chemical reactions or the speed at which the reactants and products are delivered to and removed from the reaction sites, or both. To date, it is generally agreed that surface reactions, rather than transport, play a predominant role in controlling the rate of mineral-water interactions. This point of view is supported by the nearly ubiquitous occurrence of etching features on both laboratory dissolved and naturally weathered grains of many minerals. The significance of this understanding is directly reflected by the intensive interests in mineral surface geochemistry that has been growing for more than two decades.

BIBLIOGRAPHY

1. Hochella, M.F., Jr. and White, A.F. (1990). *Rew. Mineral.* **23**: 1–16.
2. Schindler, P.W. and Stumm, W.W. (1987). The surface chemistry of oxides, hydroxides, and oxide minerals. In: *Aquatic Surface Chemistry*. W. Stumm (Ed.). John Wiley & Sons, New York, pp. 83–110.
3. Stumm, W. and Morgan, J.J. (1981). *Aquatic Chemistry*. 2nd Edn. John Wiley & Sons, New York, pp. 600–601.
4. Stumm, W. (1992). *Chemistry of the Solid-Water Interface*. John Wiley & Sons, New York, p. 95.
5. Langmuir, D. (1997). *Aqueous Environmental Geochemistry*. Prentice-Hall, Englewood, Cliffs, NJ, p. 344.
6. Morel, F.M.M. and Hering, J.G. (1993). *Principles and Applications of Aquatic Chemistry*. John Wiley & Sons, New York, p. 509.
7. Morel, F.M.M. and Hering, J.G. (1993). *Principles and Applications of Aquatic Chemistry*. John Wiley & Sons, New York, p. 421.
8. Lasaga, A.C. (1998). *Kinetic Theory in the Earth Sciences*. Princeton Univ. Press, Princeton, NJ, p. 188.
9. Frank, F.C. (1951). *Acta Cryst.* **4**: 497–501.
10. Cabrera, N. and Levine, M.M. (1956). *Phil. Mag.* **1**: 450–458.
11. Nagy, K.L. and Lasaga, A.C. (1992). *Geochim. Cosmochim. Acta* **56**: 3093–3111.
12. Lasaga, A.C. and Luttge, A. (2001). *Science* **291**: 2400–2404.
13. Blum, A.E. and Stillings, L. (1995). *Rev. Mineral.* **31**: 291–351.
14. Teng, H.H., Chen, L., Fenter, P.P., and Sturchio, N.C. (2001). *Geochimica et Cosmochimica Acta.* **65**: 3459–3474.
15. Hellmann, R. et al. (2003). *Phys. Chem. Minerals* **30**: 192–197.
16. Stumm, W. (1992). *Chemistry of the Solid-Water Interface*. John Wiley & Sons, New York, p. 216.
17. Kossel, W. (1927). *Nach. Ges. Wiss.*, Gottingen, p. 135.
18. Chernov, A.A. (1961). *Soviet Phys.* **4**: 116–148.
19. Chernov, A.A. and Komatsu, H. (1995). Topics in crystal growth kinetics. In: *Science and Technology of Crystal Growth*. J.P. van Eerden and O.S.L. Bruinsma (Eds.). Kluwer Academic, Amsterdam, pp. 67–80.
20. Lasaga, A.C. (1998). *Kinetic Theory in the Earth Sciences*. Princeton Univ. Press, Princeton, NJ, p. 178.
21. Penn, R.L. and Banfield, J.F. (1998). *Science* **281**: 967–971.
22. Penn, R.L. and Banfield, J.F. (1999). *Geochimica et Cosmochimica Acta.* **63**: 1549–1557.

GROUND WATER: WELLS

Roger M. Waller
U.S. Geological Survey

There's a good chance that the average Joe who had to dig a well in ancient Egypt probably did the work with his hands, a shovel, and a bucket. He would have kept digging until he reached the water table and water filled the bottom of the hole. Some wells are still dug by hand today, but more modern methods are available. It's still a dirty job, though!

Wells are extremely important to all societies. In many places wells provide a reliable and ample supply of water for home uses, irrigation, and industries. Where surface water is scarce, such as in deserts, people couldn't survive and thrive without ground water.

TYPES OF WELLS

Digging a well by hand is becoming outdated today (would YOU want to do it?). Modern wells are more often drilled by a truck-mounted drill rig. Still, there are many ways to put in a well—here are some of the common methods.

Dug Wells

Hacking at the ground with a pick and shovel is one way to dig a well. If the ground is soft and the water table is shallow, then dug wells can work. They are often lined with stones to prevent them from collapsing. They cannot be dug much deeper than the water table—just as you

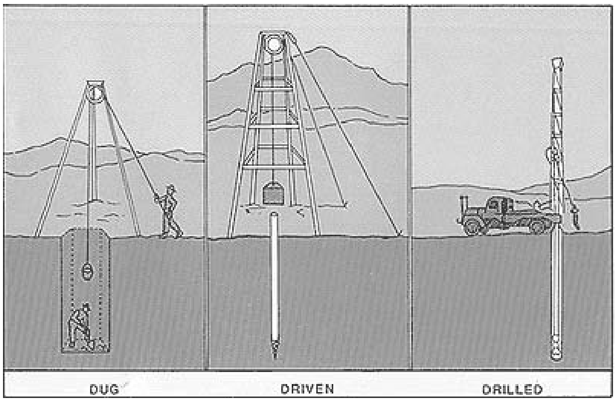

---

cannot dig a hole very deep when you are at the beach... it keeps filling up with water!

Driven Wells

Driven wells are still common today. They are built by driving a small-diameter pipe into soft earth, such as sand or gravel. A screen is usually attached to the bottom of the pipe to filter out sand and other particles. Problems? They can only tap shallow water, and because the source of the water is so close to the surface, contamination from surface pollutants can occur.

Drilled Wells

Most modern wells are drilled, which requires a fairly complicated and expensive drill rig. Drill rigs are often mounted on big trucks. They use rotary drill bits that chew away at the rock, percussion bits that smash the rock, or, if the ground is soft, large auger bits. Drilled wells can be drilled more than 1,000 feet deep. Often a pump is placed at the bottom to push water up to the surface.

WATER LEVELS IN WELLS

Ground-water users would find life easier if the water level in the aquifer that supplied their well always stayed the same. Seasonal variations in rainfall and the occasional drought affect the “height” of the underground water level. If a well is pumped at a faster rate than the aquifer around it is recharged by precipitation or other underground flow, then water levels around the well can be lowered. The water level in a well can also be lowered if other wells near it are withdrawing too much water. When water levels drop below the levels of the pump intakes, then wells will begin to pump air—they will “go dry.”

WELL SCREENS

Dennis E. Williams
Geoscience Support Services
Claremont, California

A well screen is a fairly generic term that denotes the component of a well that provides (1) an opening for water to enter the well casing from the aquifer; and (2) stabilizes material in the “near-well zone” (i.e., materials immediately adjacent to the well) and prevents it from entering the well.

Proper design of well screens therefore is paramount in successful water wells. Most modern wells are constructed using a filter pack placed between the well screen and the aquifer. In a proper well screen/filter pack design (i.e., gravel envelope well), the fundamental principle is: *The purpose of the filter pack is to stabilize the aquifer; the purpose of the well screen is to stabilize the filter pack*.

A properly designed well screen therefore allows sand-free water to flow into the well with minimum drawdown. Features of a properly designed well screen include:

1. Adequate open area (typically greater than 3–5%) to maximize well efficiency[1]
2. Strong pipe-based well screen (e.g., horizontal louver shutter screen) to provide maximum collapse strength.
3. Constant diameter throughout the entire length of the well screen.
4. Single material type (e.g., all copper-bearing or all stainless steel) to avoid galvanic corrosion from two different metals coming in contact with one another.

Well screens may be constructed with a variety of different materials and designs. The screen material can be different steels, such as carbon, copper-bearing, high strength low-alloy, and stainless, or PVC materials. The type of material used is dependent on the purpose of the well, strength requirements, groundwater quality (corrosion potential), the expected well life, and cost factor. There are many different types of well screens; however, the most commonly used are discussed below.

HORIZONTAL LOUVER SCREENS

Horizontal louver screens are a pipe-based well screen manufactured by punch-formed downward facing horizontal louver-shaped openings (Fig. 1). This type of screen is particularly suited for use in large diameter, deep, gravel envelope wells. The horizontal louver screen can be manufactured in a variety of patterns, with various percentages of open areas. Percentage open areas range from approximately 3.4% to 7.4% for Ful Flo to Super Flo patterns (2). The smooth, unobstructed interior of the horizontal louver screen allows for efficient use of highly effective development techniques such as swabbing and also facilitates employment of efficient maintenance and repair methods. Disadvantages of horizontal louver screen include

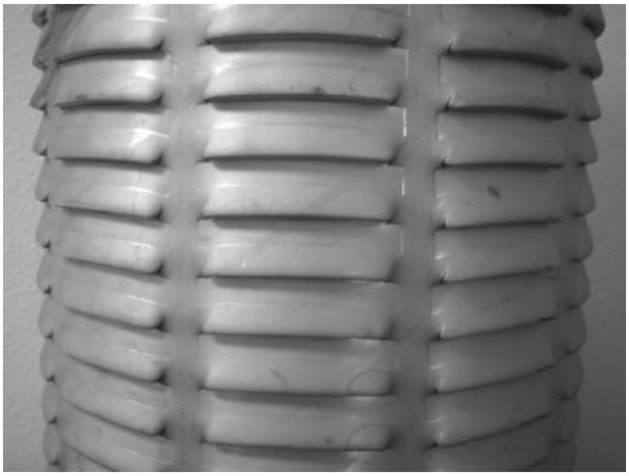

Figure 1. Horizontal louver screen.

---

[1] [It has been demonstrated (1,2) that if the percentage open area is above 3–5%, well efficiencies do not significantly increase. It has also been demonstrated that entrance velocity is not a critical design factor and only needs to be below 0.5 m/s (1.5 ft/s).]

unavailability of slot size smaller than 1 mm (0.040 in.), and difficulty in perforating though thick walled pipe [i.e., greater than 9.5 mm (0.375 in.)]. However, the superior strength capabilities and long-life characteristics have made horizontal louver screens an industry standard in deep high capacity water supply wells.

CONTINUOUS WIRE WRAP SCREEN

Wire wrap screen is manufactured by wrapping a shaped wire (steel or plastic) around an internal array of longitudinal rods (Fig. 2). The wire and rod are joined by resistance welding, which produces a cage-shaped cylindrical configuration. Continuous slot screens are very effective in naturally developed wells in uniform fine-grained sands in relatively shallow, thin aquifers since they are able to be manufactured with very small slot sizes and yet still maintain the necessary open area to minimize frictional head loss. Continuous wire wrap screens typically have open areas greater than 25%. Additionally, the design allows small particles to pass completely through the screen without becoming wedged in and causing clogging of the screen. Disadvantages of the continuous slot screen are associated with its relatively low strength, which necessitates special handling during installation and development, and also restricts the extent of future well maintenance and repair efforts. As the interior is obstructed by rods, the use of highly effective well development techniques such as swabbing is precluded. Should the screen be damaged, repair by swedging is impractical.

Figure 2. Continuous wire wrap screen.

BRIDGE SLOT SCREENS

Bridge slot screens are manufactured on a press from flat sheets or plates. The slot opening is usually vertical and provides two orifices, longitudinally aligned to the axis. The perforated steel sheets or plates are then rolled into cylinders and the seam is welded (Fig. 3). A bridge slot screen is usually installed in gravel envelope wells. Its chief advantages are reasonably high area of opening (approximately 4.8%) and minimum frictional head losses at relatively low cost. One important disadvantage is its low collapse strength due to the large number of vertically oriented slots. The manufacturing process is restricted with respect to wall thickness, which is limited to a maximum of 6.35 mm (0.250 in.).

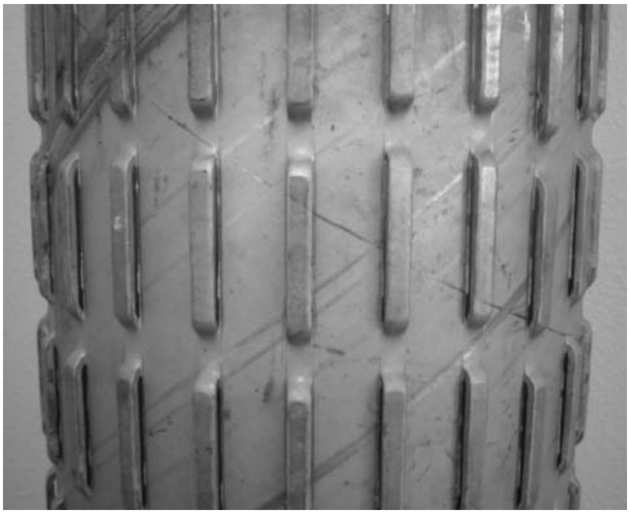

Figure 3. Bridge slot screen.

MACHINE-SLOTTED (MILLED) SCREENS

Horizontal and vertical machine-slotted screens are manufactured from casing in which openings are milled with axially oriented precision cutters (Fig. 4). Vertical slotted screens are also manufactured by prepunching flat sheets of steel prior to fabrication into cylinders. Machine-slotted screens have disadvantages that include clogging due to the parallel surfaces within the opening. As slot

Figure 4. Horizontal machine-slotted screen.

clogging is directly related to wall thickness, thicker material encourages greater plugging. Some vertical machine-slotted casings are machined with an undercut to reduce this tendency. A second drawback is the low area of openings (less than 1%), although this can be overcome at higher cost by increasing the number of slots. Collapse strength of vertically machine-slotted casings, however, is substantially reduced when the number of openings is increased. The main positive characteristics of machine-slotted screens are low cost, uniform openings, and availability in a fairly wide range of apertures and patterns.

WELL CASING PERFORATED IN PLACE

Well casing perforated in place has historically been used with cable tool wells drilled in unconsolidated alluvium. Once the casing has been driven or jacked to its total depth, a mills knife (vertical slots) or hydraulic louver-type perforator (horizontal slots) is used down the hole to perforate the casing at specific intervals. The hydraulic louver-type perforator is superior to the mills knife in producing slots due to better physical casing characteristics and more control over the aperture size; in addition, the louver-shaped openings provide greater stabilization of particles in relative fine-grained sand formations. The hydraulic louver-type perforator is limited to a slot size of 3.2–4.0 mm (0.125–0.156 in.). Typically, a well casing perforated in place would have a percentage open area of less than 1%.

The design of the screen slot or aperture size for gravel envelope wells requires the filter pack to be designed first. The screen slot size must result in 90% retention of filter pack material outside the screen.

BIBLIOGRAPHY

1. Williams, D.E. (1985). Modern techniques in well design. *J. AWWA* **77**: 9.
2. Moss, R. (1990). *Handbook of Ground Water Development*. John Wiley & Sons, Hoboken, NJ.

WELL TEST

Dennis E. Williams
Geoscience Support Services
Claremont, California

A well test is a test conducted in a completed and developed well from which data can be collected to calculate certain hydraulic parameters of the well and surrounding aquifer. Hydraulic parameters may then be used, for example, to design the maximum safe pumping rate of the well, determine well field spacing, calculate aquifer storage, or determine the rate of movement of contaminants in the groundwater. Well tests include pumping tests, recovery tests, and slug tests. Essential measurements and data collected during well tests include time, depth to water, and discharge rate. Obtaining accurate data from a pumping test involves a carefully planned test program, including consideration of continuous operation and execution of the necessary steps.

Step-drawdown tests are generally the first type of test performed on a newly completed production well following final well development. A step-drawdown test involves pumping the well at several discharge rates and measuring the change in depth to pumping level in the well as time progresses (Fig. 1). At least three "discharge steps" are required for the test with subsequent discharge rates increasing over the previous steps. Data obtained from step-drawdown tests are used to determine well production and drawdown characteristics, from which the permanent production pump can be designed. The discharge step rates are determined based on results from the final development pumping. For example, if during final development, the maximum discharge rate reached 2500 gpm with good groundwater level stabilization, then the three rates for the step-drawdown test might be 750, 1500, and 2200 gpm. The discharge rates for a step-drawdown test should be reasonably spread out over the maximum range available for the well (as determined during development).

Following the step-drawdown test, a constant-rate pumping test is usually performed. The purpose of the constant-rate test is to verify the design discharge rate estimated from the step-drawdown test and also to measure longer-term (usually greater than 24 h) drawdown effects on the pumping well and any nearby wells. Data obtained from both the pumping well and nearby wells can be used to calculate aquifer parameters, which can then be used to design spacing between wells, calculate groundwater storage volumes, or determine the rate of movement of groundwater. The discharge rate during the constant-rate test should be kept constant for the entire duration of the test and should be at the design discharge rate for the well (Fig. 2). The design discharge rate may be based on both development pumping and results from the step-drawdown test. The step-drawdown test is seldom performed for a duration sufficient to design the permanent pump and should always be verified with a longer-term constant-rate test. However, the final design pumping rate is still subject to "sound engineering

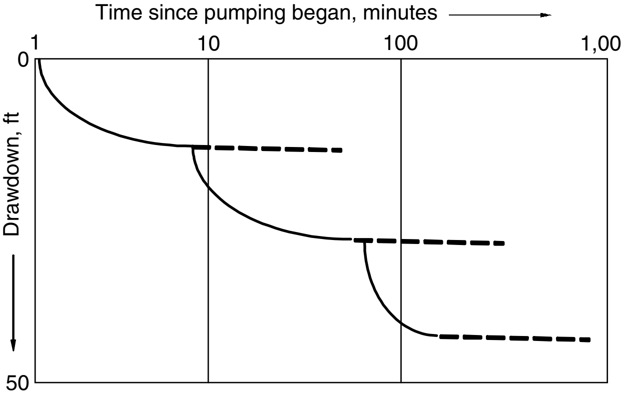

Figure 1. Step-drawdown test data.

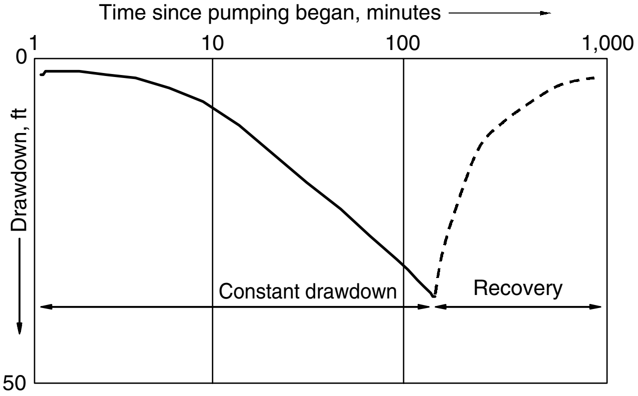

Figure 2. Constant-rate test with recovery data.

judgment" and should consider both demand requirements and pumping test results.

Recovery tests are carried out after pumping tests. The test involves measuring the rate at which the groundwater level recovers to its initial level after pumping has stopped. In general, water level data from recovery tests are more reliable due to the lack of water level fluctuations caused by discharge variations.

A slug test is a popular and inexpensive method of estimating the hydraulic properties of aquifers in small diameter wells when the hydraulic conductivity of the aquifer materials is too low to conduct a pumping test. Slug testing involves either rapidly adding or removing a "slug of water" from a well and measuring the rate at which the water level recovers (either falling-head or rising-head). Disadvantages with this method are that it can only be used to determine the aquifer's hydraulic conductivity in the immediate vicinity of the well, and results can be influenced by filter pack material adjacent to the well screen.

The type of test to be performed depends on the test purpose, available resources, and site-specific limitations. For example, in an area containing only a single well, it is not possible to calculate aquifer storage parameters from a pumping test without an observation well (i.e., nonpumping well). Similarly, observation wells not screened in the same aquifer as the pumping well cannot provide reliable interference data.

Once groundwater level data from the well tests have been collected, compiled, and graphically plotted, the data are evaluated using various methods to determine hydraulic parameters of the aquifer and well. In many cases, the calculation of parameters requires only solving simple equations or a set of equations. For each method there are basic underlying assumptions from which the method equation was derived. These assumptions need to be carefully considered with respect to the test data and specific aquifer conditions when a method is used. In addition to the basic assumptions, each of the analysis methods has other assumptions specific to the method. Many of the analytical solutions are based on the basic differential equation of groundwater flow, which can be derived from fundamental laws of physics.

SAFE YIELD OF AN AQUIFER

Eric S. Wilson
E. L. Montgomery & Associates, Inc.
Tucson, Arizona

Safe yield is a controversial term and concept that may be defined for common usage as the amount of water an aquifer system can yield over a long period without producing unacceptable effects. Unacceptable effects may include an excessive decline in groundwater level, sometimes described as groundwater mining; reduction or elimination of riparian vegetation; reduction or elimination of water in streams, lakes, springs, and wetlands; aquifer compaction and land subsidence; reduction in groundwater underflow to downgradient aquifer systems; a change in groundwater temperature; an undesirable change in the chemical quality of groundwater; and salt water intrusion into aquifer systems adjacent to oceans or saline lakes.

The definition of safe yield has evolved for nearly a century; initially, it was defined by Lee (1) as "...the limit and quantity of water which can be withdrawn regularly and permanently without dangerous depletion of the storage reserve." The definition was modified and expanded by Meinzer (2), Conkling (3), and Todd (4). Due to ambiguity and lack of agreement on the definition, many authors have suggested abandoning the term (5–7) for an alternate term such as "sustainable yield."

Incorrect concepts, regulations, or laws may indicate that safe yield is equal to recharge. This incorrect concept has been described by Bredehoeft et al. (8) as the "water-budget myth" because the concept oversimplifies the necessary information needed to understand the effects of groundwater development on an aquifer system.

It is a great deal more difficult to define safe yield of an aquifer based on analyzing the unacceptable effects of overpumping than to assume simply that safe yield is equal to recharge. When water is pumped from an aquifer, even in amounts equal to recharge, natural groundwater discharge continues. The initial result of pumping is the decline of groundwater levels near the pumping well, and with time, across larger areas of the aquifer. This water level decline increases the hydraulic gradient and assists in drawing groundwater from marginal parts of the aquifer system into the central part of the aquifer system, but eventually water levels drop sufficiently that natural groundwater discharge becomes smaller. As a result, even by pumping groundwater at rates equal to recharge, a decline in groundwater discharge may occur. This decline in discharge could possibly be defined as unacceptable, and the aquifer may not be in a state of safe yield.

When considering the pragmatic aspects of safe yield, it is important to address the principal elements of the water budget expressed by the law of conservation of mass and the water budget equation:

$$\text{Inflow} = \text{outflow} \pm \text{change in storage}$$

Pumping from an aquifer system over time changes the amount of water in storage and subsequently changes values of inflow and outflow, thus changing values in the water budget equation. Theis (9) stated that groundwater withdrawals are balanced, over time, by a loss of water somewhere. Groundwater withdrawals by pumping result in one or more of the following: (1) a decrease in the amount of groundwater in storage; (2) capture of previously rejected recharge, if available; (3) an increase in water entering the aquifer system through increased underflow, if available; (4) decreased outflows by way of underflow to downgradient aquifers; or (5) decreased outflow by reduction of discharge to streamflow or to evapotranspiration. The hydraulics of an aquifer system evolve toward, and eventually may approach, a new dynamic equilibrium under increased pumping. An exception may occur in hydraulically isolated, arid environments where little recharge is available. Under these conditions, if pumping exceeds the meager natural recharge, no subsequent equilibrium will be reached and groundwater mining is expected.

Groundwater withdrawal or development of water resources of an aquifer system occurs in stages. These stages include the predevelopment stage, partially developed stage, and fully developed stage. Under natural or predevelopment conditions, an aquifer system is in dynamic equilibrium among change in aquifer storage, groundwater underflow, and interaction between groundwater and surface water. As additional development occurs, water budget values change as a result of an extended period of groundwater pumping. A partially developed aquifer system will see moderate declines in groundwater level and decreases in groundwater in storage; an increase in recharge may occur, and a decrease in natural discharge is expected. Full development of an aquifer system results in a large magnitude of groundwater level decline, additional depletion of groundwater in storage, maximization of recharge, elimination of discharge to surface waters, and underflow to downgradient aquifers may cease.

Safe yield cannot be defined solely on the basis of hydrologic conditions. Hydrologists may describe and quantify the effect of groundwater withdrawals on an aquifer system, but safe yield is properly a determination by a team of natural resource managers, and the results must be acceptable to concerned citizens. It is often difficult to reach consensus on the true safe yield of an aquifer system. Although difficult, the term and concept are the basis of many water management programs and groundwater laws and therefore must be given careful attention.

Wood (10) states, "...scientists and engineers must avoid the trap of blurring the boundary between scientific investigations and management decisions. Neither I nor any other scientist or engineer can say whether any given volume of water or delivery rate meets the needs of anyone beyond the minimum personal requirement of two or three liters per day necessary for long-term survival. 'Meeting the needs' requires many value judgments, and their assessment lies outside the realm of science or engineering. In many ways, sustainability appears to be analogous to the management concept of aquifer 'safe yield' that has been extensively used and misused in the last 85 years."

BIBLIOGRAPHY

1. Lee, C.H. (1915). The determination of safe yield of underground reservoirs of the closed basin type. *Trans. Am. Soc. Civ. Eng.* **78**: 148–251.
2. Meinzer, O.E. (1923). *Outline of Ground-Water Hydrology with Definitions*. U.S. Geological Survey Water-Supply Paper 494.
3. Conkling, H. (1946). Utilization of ground-water storage in stream system development. *Trans. Am. Soc. Civ. Eng.* **3**: 275–305.
4. Todd, D.K. (1959). *Ground Water Hydrology*. John Wiley & Sons, New York.
5. Thomas, H.E. (1951). *The Conservation of Ground Water*. McGraw-Hill, New York.
6. Kazmann, R.G. (1956). Safe yield in ground-water development, reality or illusion? *J. Irrig. Drainage Div. Am. Soc. Civ. Eng.* **82**(IR3): 12.
7. DeWiest, R.J.M. (1965). *Geohydrology*. John Wiley & Sons, New York.
8. Bredehoeft, J.D., Papadopulos, S.S., and Cooper, H.H., Jr. (1982). Groundwater: The water-budget myth. In: *Scientific Basis of Water-Resource Management*. National Academy Press, Washington, DC, pp. 51–57.
9. Theis, C.V. (1940). The source of water derived from wells—essential factors controlling the response of an aquifer to development. *Civ. Eng.* **10**: 277–280.
10. Wood, W. (2001). Water sustainability—science or management. *Ground Water* **39**(5): 641.

SPECIFIC YIELD STORAGE EQUATION

THOMAS HARTER
University of California
Davis, California

The specific yield, S_y, is the ratio of the volume of water that can be drained by gravity from a mass of porous rock (unconsolidated or consolidated) and the volume of that mass:

$$S_y = \frac{V_W}{V_T},$$

where V_W is the volume of drainable water and V_T is the total volume of porous rock or sediments (Fig. 1). Specific yield, S_y, is typically expressed as a percentage.

Specific yield is part of the total porosity of a porous rock or sediment. The total porosity (total pore space) of a rock or sediment formation is much larger than the specific yield. Total porosity includes the fraction of pore space that is interconnected (called "effective porosity") and porosity due to isolated pores. The effective porosity itself includes specific retention and specific yield. Specific retention is that volume fraction of water that is held back by adhesion and capillary forces, when an aquifer is drained. Specific yield is the amount of water that is actually available for groundwater pumping, when sediments or rocks are drained due to lowering of the water table.

Specific yield depends on the distribution of pores, their shape, and grain size (Table 1). The larger the surface area of the rocks or sediment particles per unit volume values of inflow and outflow, thus changing values in the water budget equation. Theis (9) stated that groundwater withdrawals are balanced, over time, by a loss of water somewhere. Groundwater withdrawals by pumping result in one or more of the following: (1) a decrease in the amount of groundwater in storage; (2) capture of previously rejected recharge, if available; (3) an increase in water entering the aquifer system through increased underflow, if available; (4) decreased outflows by way of underflow to downgradient aquifers; or (5) decreased outflow by reduction of discharge to streamflow or to evapotranspiration. The hydraulics of an aquifer system evolve toward, and eventually may approach, a new dynamic equilibrium under increased pumping. An exception may occur in hydraulically isolated, arid environments where little recharge is available. Under these conditions, if pumping exceeds the meager natural recharge, no subsequent equilibrium will be reached and groundwater mining is expected.

Groundwater withdrawal or development of water resources of an aquifer system occurs in stages. These stages include the predevelopment stage, partially developed stage, and fully developed stage. Under natural or predevelopment conditions, an aquifer system is in dynamic equilibrium among change in aquifer storage, groundwater underflow, and interaction between groundwater and surface water. As additional development occurs, water budget values change as a result of an extended period of groundwater pumping. A partially developed aquifer system will see moderate declines in groundwater level and decreases in groundwater in storage; an increase in recharge may occur, and a decrease in natural discharge is expected. Full development of an aquifer system results in a large magnitude of groundwater level decline, additional depletion of groundwater in storage, maximization of recharge, elimination of discharge to surface waters, and underflow to downgradient aquifers may cease.

Safe yield cannot be defined solely on the basis of hydrologic conditions. Hydrologists may describe and quantify the effect of groundwater withdrawals on an aquifer system, but safe yield is properly a determination by a team of natural resource managers, and the results must be acceptable to concerned citizens. It is often difficult to reach consensus on the true safe yield of an aquifer system. Although difficult, the term and concept are the basis of many water management programs and groundwater laws and therefore must be given careful attention.

Wood (10) states, "...scientists and engineers must avoid the trap of blurring the boundary between scientific investigations and management decisions. Neither I nor any other scientist or engineer can say whether any given volume of water or delivery rate meets the needs of anyone beyond the minimum personal requirement of two or three liters per day necessary for long-term survival. 'Meeting the needs' requires many value judgments, and their assessment lies outside the realm of science or engineering. In many ways, sustainability appears to be analogous to the management concept of aquifer 'safe yield' that has been extensively used and misused in the last 85 years."

BIBLIOGRAPHY

1. Lee, C.H. (1915). The determination of safe yield of underground reservoirs of the closed basin type. *Trans. Am. Soc. Civ. Eng.* **78**: 148–251.
2. Meinzer, O.E. (1923). *Outline of Ground-Water Hydrology with Definitions*. U.S. Geological Survey Water-Supply Paper 494.
3. Conkling, H. (1946). Utilization of ground-water storage in stream system development. *Trans. Am. Soc. Civ. Eng.* **3**: 275–305.
4. Todd, D.K. (1959). *Ground Water Hydrology*. John Wiley & Sons, New York.
5. Thomas, H.E. (1951). *The Conservation of Ground Water*. McGraw-Hill, New York.
6. Kazmann, R.G. (1956). Safe yield in ground-water development, reality or illusion? *J. Irrig. Drainage Div. Am. Soc. Civ. Eng.* **82**(IR3): 12.
7. DeWiest, R.J.M. (1965). *Geohydrology*. John Wiley & Sons, New York.
8. Bredehoeft, J.D., Papadopulos, S.S., and Cooper, H.H., Jr. (1982). Groundwater: The water-budget myth. In: *Scientific Basis of Water-Resource Management*. National Academy Press, Washington, DC, pp. 51–57.
9. Theis, C.V. (1940). The source of water derived from wells—essential factors controlling the response of an aquifer to development. *Civ. Eng.* **10**: 277–280.
10. Wood, W. (2001). Water sustainability—science or management. *Ground Water* **39**(5): 641.

SPECIFIC YIELD STORAGE EQUATION

Thomas Harter
University of California
Davis, California

The specific yield, S_y, is the ratio of the volume of water that can be drained by gravity from a mass of porous rock (unconsolidated or consolidated) and the volume of that mass:

$$S_y = \frac{V_W}{V_T},$$

where V_W is the volume of drainable water and V_T is the total volume of porous rock or sediments (Fig. 1). Specific yield, S_y, is typically expressed as a percentage.

Specific yield is part of the total porosity of a porous rock or sediment. The total porosity (total pore space) of a rock or sediment formation is much larger than the specific yield. Total porosity includes the fraction of pore space that is interconnected (called "effective porosity") and porosity due to isolated pores. The effective porosity itself includes specific retention and specific yield. Specific retention is that volume fraction of water that is held back by adhesion and capillary forces, when an aquifer is drained. Specific yield is the amount of water that is actually available for groundwater pumping, when sediments or rocks are drained due to lowering of the water table.

Specific yield depends on the distribution of pores, their shape, and grain size (Table 1). The larger the surface area of the rocks or sediment particles per unit volume

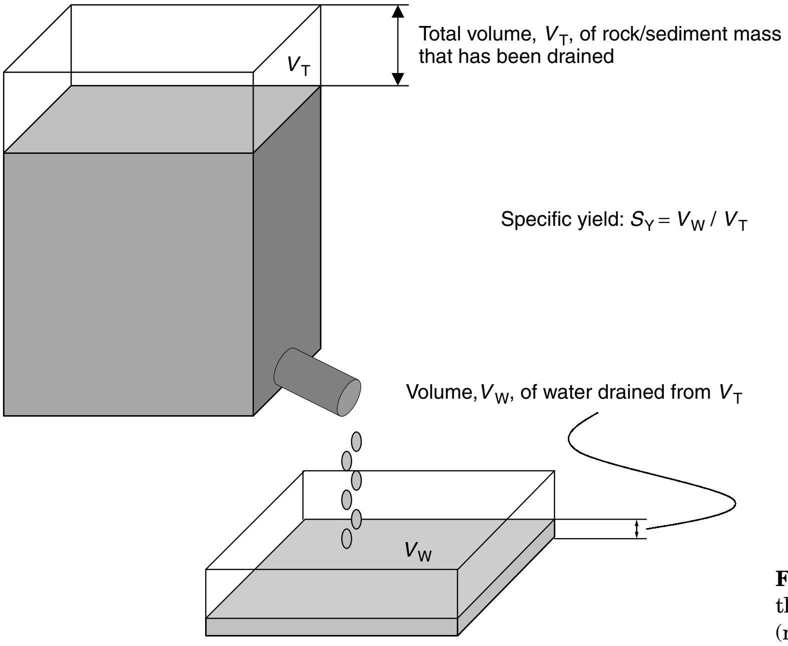

Figure 1. Illustration of the concept of specific yield. Note that the drained volume V_T contains sediment particles (not shown here). The volume V_W is much less then V_T.

("specific surface area"), the more water adheres to these surfaces. This decreases the specific yield in favor of specific retention. Fine-grained sediments have the largest specific surface area, the highest specific retention, and the lowest specific yield. Coarse-grained sediments have a very small specific surface area; hence, the specific retention is small, and the specific yield is large. Specific yield typically decreases with depth due to compaction of sediments.

Clay, silt, and fine sand have relatively high effective porosities, but their specific yields vary widely. The specific yield of clay is often 2% or less, due to its extremely small grain size. In contrast, the specific yield of fine sand can be over 20%. The highest specific yield is obtained in coarse sands and gravels (20–30%), where little water is held back by retention (large pores have only small capillary forces). This makes sand and gravel aquifers good storage reservoirs for groundwater. Peat stands out for its large effective porosity: 90% of its total volume consists of pores, half of which is drainable (specific yield: ~ 45%). These properties (high specific retention and high specific yield) make it a good amendment for garden and potting soil.

Table 1. Specific Yield of Various Sediments and Rocks

Sediment or Porous Rock	Specific Yield, in %,[a]
Gravel, coarse	10–25
Gravel, medium	15–25
Gravel, fine	20–35
Sand, gravelly	20–35
Sand, coarse	20–35
Sand, medium	15–30
Sand, fine	10–30
Silt	1–30
Clay, sandy	3–20
Clay	1–20
Sandstone	10–40
Limestone	1–25
Dune sand	30–40
Loess	15–35
Peat	30–50
Siltstone	1–35
Till	5–20
Tuff	2–35

[a] Compiled by Reference 1.

The specific yield is used to determine how much water can be produced from an unconfined aquifer per a unit (foot, meter) decline in the water table (Fig. 1). An example may illustrate this. A shallow, unconfined sand and gravel aquifer in a small montane basin has an extent of 10 km^2, an average thickness of 10 m, and a specific yield of 20%. Let us assume that there is a drought and the aquifer does not receive any recharge. How much water has been produced from the aquifer, if the water table across the entire aquifer drops by 1 m? The answer: The total volume of aquifer drained is 10 km^2 × 1 m. Then the volume of water obtained is 20% of that volume, or 10 km^2 × 0.2 m.

The same example can be turned around. Let us assume there is no pumping from the aquifer during a 6-month rainy period. During this time, recharge from a river and from precipitation amounts to a total volume of 10 km^2 × 0.1 m or an average of 100 liters/m^2. How much will this raise the water level? With specific yield at 20%, the 100 liters/m^2 recharge will fill an aquifer volume of 100 liters/0.2 = 500 liters/m^2 (0.5 m^3/m^2) in each empty cubic meter of aquifer. Hence, the water level rises by 0.5 m.

Specific yield is measured by a variety of laboratory and field methods, or by estimation from computer models (for recent reviews and applications, see references). The most common field technique is evaluation of well-pumping test data (2,3).

BIBLIOGRAPHY

1. Walton, W.C. (1991). *Principles of Groundwater Engineering*. CRC Press, Boca Raton, FL.

2. Neuman, S.P. (1987). On methods of determining specific yield. *Ground Water* **25**(6): 679–684.
3. Moench, A.F. (1994). Specific yield as determined by type-curve analysis of aquifer-test data. *Ground Water* **32**(6): 949–957.

READING LIST

Frohlich, R.K. and Kelly, W.E. (1988). Estimates of specific yield with the geoelectric resistivity method in glacial aquifers. *J. Hydrol.* **97**(1–4): 33–44.

Grimestad, G. (2002). A reassessment of ground water flow conditions and specific yield at Borden and Cape Cod. *Ground Water* **40**(1): 14–24.

Howle, J.F., Phillips, S.P., Denlinger, R.P., and Metzger, L.F. (1997). *Determination of Specific Yield and Water-Table Changes Using Temporal Microgravity Surveys Collected During the Second Injection Storage and Recovery Test at Lancaster*, Antelope Valley, California, November 1996 through April 1997. U.S. Geological Survey, Water-Resources Investigations Report 03–4019.

Johnson, A.I. (1976). *Specific Yield; Compilation of Specific Yields for Various Materials*. U.S. Geological Survey Water-Supply Paper 1662-D.

Nwankwor, G.I., Cherry, J.A., and Gillham, R.W. (1984). A comparative study of specific yield determinations for a shallow sand aquifer. *Ground Water* **22**(6): 764–772.

Pool, D.R. and Eychaner, J.H., (1995). Measurements of aquifer storage change and specific yield using gravity surveys. *Ground Water* **33**(3): 425–432.

Robson, S.G. (1993). *Techniques for Estimating Specific Yield and Specific Retention from Grain-size Data and Geophysical logs from Clastic Bedrock Aquifers*. U. S. Geological Survey Water-Resources Investigations Report, WRI 93–4198.

Tritscher, P. and Read, W.W. (2000). Specific yield for a two-dimensional flow. *Water Resour. Res.* **36**(6): 1393–1402.

MICROBIAL PROCESSES AFFECTING MONITORED NATURAL ATTENUATION OF CONTAMINANTS IN THE SUBSURFACE

Ann Azadpour-Keeley
Guy W. Sewell
National Risk Management
Research Laboratory
ORD, U.S. EPA, Ada,
Oklahoma

Hugh H. Russell
CHR$_2$ Environmental Services,
Inc.,
Oilton, Oklahoma

INTRODUCTION

The EPA Regional Ground Water Forum is a group of EPA professionals representing Regional Superfund and Resource Conservation and Recovery Act (RCRA) Offices, committed to the identification and resolution of ground-water issues impacting the remediation of Superfund and RCRA sites. Innovative technologies for subsurface remediation are being evaluated more often for specific sites, as the limitations to conventional technologies are recognized. The purpose of this Issue Paper is to provide those involved in assessing remediation technologies with some basic information regarding monitored natural attenuation (MNA) processes, specifically in determining overall contribution of microbial processes.

On April 21, 1999, the Office of Solid Waste and Emergency Response (OSWER) issued Directive 9200.4-17P, titled "Use of Monitored Natural Attenuation at Superfund, RCRA Corrective Action, and Underground Storage Tank Sites." It specifies the current EPA policy regarding the use of MNA for the remediation of contaminated soil and ground water at OSWER sites. The Directive's aim is to promote consistency in the manner in which MNA remedies are proposed, evaluated, and approved. Accordingly, "EPA does not consider MNA to be a presumptive or default remedy—it is merely one option that should be considered with other applicable remedies." Thus, during the process of selecting a site remedy, MNA may be evaluated and compared with other remedial technologies and chosen upon attainment of each relevant remedy selection criteria, including the full protection of human health and the environment, and achieving the intended site remedial objectives within a time frame that is comparable to the other remedial techniques. The Directive expects that "source control and long-term performance monitoring will be fundamental components of any MNA remedy" (U.S. EPA, 1999).

It should be emphasized that this document is not intended to be used in establishing protocol involved in natural attenuation investigations or policies leading to the interpretation of the results of those investigations. To that end, the reader is referred to the EPA Protocol for chlorinated solvents (U.S. EPA, 1998a), and ASTM for petroleum hydrocarbons (ASTM, 1998). EPA is also preparing a guidance for long term monitoring (LTM) for MNA.

For further information contact Dr. Ann Azadpour-Keeley (580-436-8890) at the Subsurface Protection and Remediation Division of the National Risk Management Research Laboratory, Ada, Oklahoma.

BACKGROUND

Interest in the natural attenuation of ground-water contaminants has increased in recent years as the result of dealing with complexities of subsurface systems and the inherent problems and costs associated with more conventional remedial technologies, such as pump-and-treat systems. There is a growing perception in the environmental community that, under favorable conditions, the selection of natural attenuation as a remedy will result in significant savings in cost over more intrusive remedial alternatives which are exacerbated by the complex geochemical, biochemical, and hydrogeological uncertainties which are dominant

---

at most sites. While this perception may or may not be true, depending upon site specific characteristics, like all relatively new technologies, actual cost and performance data will be required as the remediation alternative matures. Natural attenuation, which is also referred to as natural assimilation, intrinsic remediation, intrinsic bioremediation, natural recovery, or passive remediation, is the use of natural processes to remove contaminants from soil or ground water. While the mechanisms of chemical transformation, dispersion, dilution, sorption and volatilization have been identified, aerobic and anaerobic degradation comprise the major processes for the reduction of contaminant mass in the subsurface. Some considerations needed for the evaluation of these processes are depicted in Table 1.

There is little question that naturally occurring biodegradation processes are taking place at many sites where sediments have been contaminated (Davis et al., 1994; Lee, 1988). To those who have examined subsurface sediment, the reduced mineralogy (darkening) and unpleasant odors (anaerobic metabolic products) are indicators of microbial degradation processes. While there are limitations to natural attenuation due to factors such as complex hydrogeology, microbial toxicity of contaminants, and other physical, biological, chemical and environmental factors, many of the organic compounds introduced into the subsurface can be transformed by indigenous microorganisms. The primary challenge in evaluating natural attenuation is not in demonstrating that biodegradation is occurring. This can be a relatively easy task accomplished by determining that the production of metabolites and the loss or electron acceptors are concomitant with the distribution of contaminants. As with any other remedial technology, the appropriate evaluation of natural attenuation as a remedial alternative is to make the determination that the transformation processes are taking place at a rate that is protective of human health and the environment. The evaluation should include a reasonable expectation that these processes will continue at an acceptable rate for an acceptable period of time.

Chlorinated aliphatics are among the most widespread contaminants in ground water (Fig. 1) and soil due to their use for degreasing, dry cleaning, and as solvents (Westrick et al., 1984). At the same time, the impact of petroleum hydrocarbons among the various environmental media is ubiquitous (Tiedje and Stephens, 1988; Sharples, 1992; Kennedy and Hutchins, 1992). Because of their prominence as environmental contaminants, these two groups of chemicals will be used as case examples in this issue paper.

MONITORED NATURAL ATTENUATION MODEL

Contaminants in the subsurface partition into four phases including sorption to the soil and aquifer solids, a free phase which displaces water from the pore spaces, dissolved in the water, and vapor (Fig. 2). The degree to which contaminants partition into these phases is determined by their physical/chemical properties or notably the sorption coefficient, Henry's Law Constant (gas partition coefficient), octanol-water coefficient, and solubility (water partition coefficient). The mass in each phase will therefore depend on the preference of the contaminants of concern for each phase.

Plume movement is dependent on the same partition coefficients, in addition to the hydrology of the site itself. In most instances, chemicals will prefer to partition to organic

Table 1. Some Information Needed for Prediction of Organic Contaminant Movement and Transformation in Ground Water. (Reprinted from *Biotic and Abiotic Transformation of Halogenated Aliphatic Compounds*, T.M. Vogel, Ph.D. Thesis, Stanford University, Stanford, CA, p. 42, 1988, with permission of T.M. Vogel)

BIOLOGICAL	**Ground-Water Characteristics**	**Aquifer Characteristics**	**Contaminant Characteristics**
	ionic strength	grain size	potential products
	pH	active bacteria number	toxicity
	temperature	Monod rate-constants	concentration
	nutrients		
	substrate		
	O_2, NO_3^-, $SO_4^=$ macro (P, S, N)		
	trace		
	organism		
CHEMICAL	**Ground-Water Characteristics**	**Aquifer Characteristics**	**Contaminant Characteristics**
	ionic strength	potential catalysts metals, clays	potential products
	pH		concentration
	temperature		
	NO_3^-, $SO_4^=$, O_2		
	toxicants		
HYDRAULIC	**Contaminant Source**	**Wells**	**Hydrogeologic Environment**
	location	location	extent of aquifer and aquitard
	amount	amount	
	rate of release	depth	characteristics of aquifer
		pump rates	hydraulic gradient
			ground-water flow rate
SORPTION	**Distribution Coefficient**	**Characteristics of the Aquifer Solid**	**Contaminant Characteristics**
	characteristic of concentration	organic carbon content	octanol/water partition coefficient
		clay content	solubility

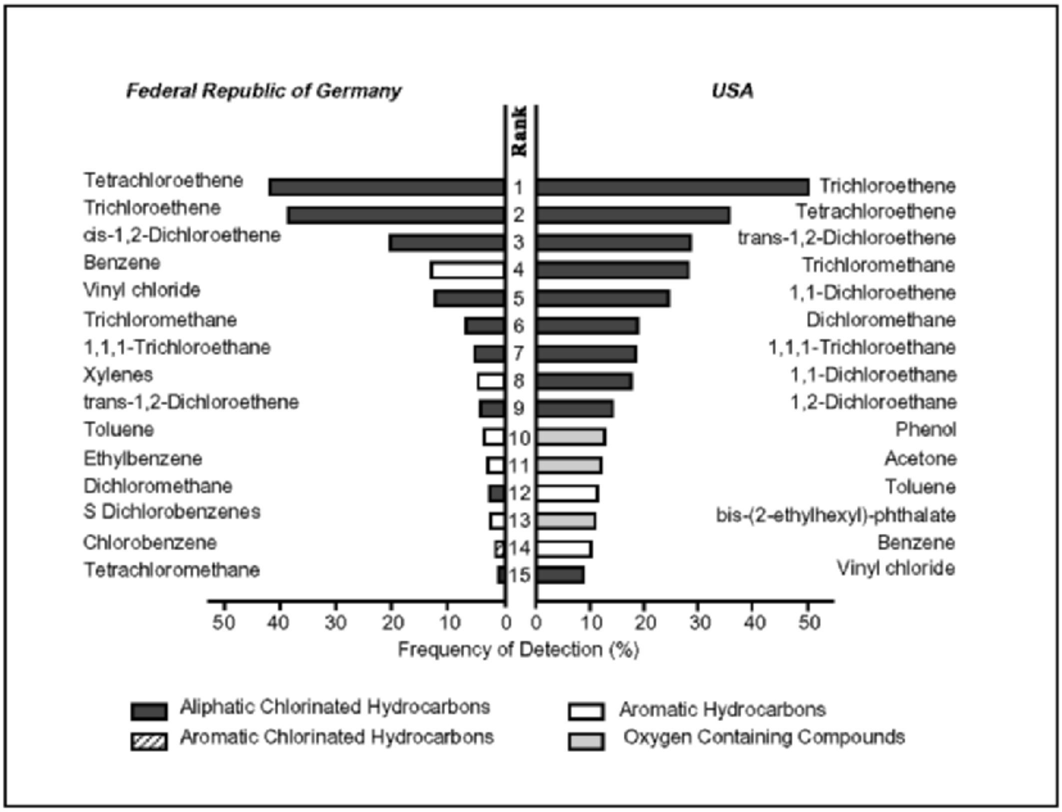

Figure 1. The 15 most frequently detected organic compounds in groundwater at waste disposal sites in Germany and the U.S. (Reprinted from *The Landfill*, Baccini, P., Ed., p. 399, 1989. Arneth, J.-D., Milde, G., Kerndorff, H., and Schleyer, R. with permission of Springer-Verlag. New York, Inc., New York, NY.).

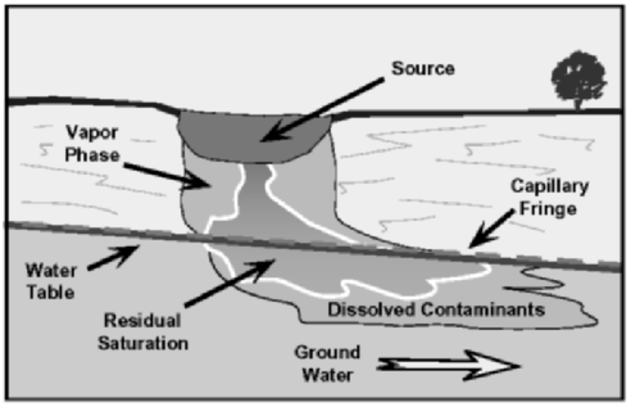

Figure 2. Distribution of contaminants in the subsurface. (Modified from *In-Situ Bioremediation of Ground Water and Geological Material: A Review of Technologies*, p. 2, 1993, by Norris et al., EPA/600/R-93/124.).

matter within sediments, to oily or free-phase material that may completely fill pore spaces, or be dissolved in water at solubility levels. The mass of the contaminants in the aqueous phase is usually lower than that in the other phases.

As water moves through pore spaces, chemicals are desorbed from sediments or dissolved from free-phase liquids. Once in solution, these chemicals move with the aqueous phase. Their movement downgradient is impeded by further sorption to sediments and biodegradation. The basis of natural attenuation is that the partition of chemicals into the aqueous phase reaches equilibrium with the processes of biological transformation at an acceptable time and distance from the source. To understand these biological processes, some knowledge of microbiology is required.

MICROBIAL PHYSIOLOGY

Bacteria, although unicellular, share characteristics with most living things. Information is encoded in DNA (deoxyribonucleic acid) and transferred through RNA (ribonucleic acid) to ribosomes to make proteins or enzymes which are used to operate systems within the organism. In regard to this discussion, enzymes are responsible for the degradation of organic carbon, which is used by the bacterial cell to produce both the building blocks of life and energy. The degradation of any organic molecule, including contaminants, requires the production and efficient utilization of enzymes. In most instances, degradation is merely a complex oxidation/reduction reaction. The electrons or reducing equivalents (hydrogen or electron-transferring molecules) produced must be transferred to a terminal electron acceptor (TEA). During the transfer process, energy is produced which is utilized by the cell.

In regard to TEAs, bacteria are generally grouped into three categories:

1. *Aerobic bacteria*—Bacteria which can only utilize molecular oxygen as a TEA. Without molecular oxygen, these bacteria are not capable of degradation.
2. *Facultative aerobes/anaerobes*—Bacteria which can utilize molecular oxygen or when oxygen concentrations are low or nonexistent, may switch to nitrate, manganese oxides or iron oxides as electron acceptors.
3. *Anaerobes*—Bacteria which cannot utilize oxygen as an electron acceptor and for which oxygen is toxic. Though members may utilize nitrate or other electron acceptors, it can be said that they generally utilize sulfate or carbon dioxide as electron acceptors.

In this discussion, three modes of contaminant degradation are addressed including aerobic, cometabolic, and anaerobic. The first is the case in which the contaminants, for example, petroleum hydrocarbons, are utilized by bacteria as a sole source of carbon. Petroleum hydrocarbons are degraded through a series of enzymatic reactions to produce needed cellular constituents. Electrons or reducing equivalents must be regenerated. If a contaminant serves as a sole source of carbon and energy, conditions must be within acceptable pH, Eh, and temperature limits and the appropriate TEA must be present. In this case, the rate of degradation will be determined by the rate of dissolution of toxic end products away from the microbial population and the rate at which the TEA is replenished. Some of the lesser chlorinated solvents, such as dichloroethene (DCE), may also serve as sole sources of carbon; however, tetrachloroethene (PCE) and trichloroethene (TCE) are not thought to serve as sole sources of carbon.

In the case of TCE (and lesser chlorinated solvents), degradation may occur through cometabolic processes. Under aerobic conditions, the enzymes necessary for the degradation, however, must be induced. Inducible enzymes are those that are not produced unless an inducer compound is present within the bacterial cell. Pertinent to this discussion are the inducers for methane monooxygenase and various mono- and dioxygenase enzymes produced by aromatic degrading bacteria.

In the presence of oxygen and methane, methanotrophic bacteria are known to produce the enzyme methane monooxygenase (Hanson and Hanson, 1996; Patel et al., 1982). The substrate for this enzyme is methane, but it has been shown to have a broad substrate specificity including chlorinated solvents (Mayer et al., 1988). Methanotrophs downgradient from a chlorinated solvent event may feed on methane produced within the anaerobic portion of the plume and cometabolically degrade some chlorinated solvents.

Numerous authors have shown that oxygenase enzymes produced by bacteria capable of degrading aromatic hydrocarbons are capable of degrading chlorinated solvents. Aromatic compounds, such as toluene and phenol, have been shown to induce the responsible enzymes. In contaminated aquifers which contain both aromatic hydrocarbons and chlorinated solvents, degradation of both may occur.

It should be noted that reports on the degradation of PCE under aerobic conditions do not exist in the peer-reviewed literature. The structure and oxidative state of PCE may preclude its aerobic degradation (Chen et al., 1996). The anaerobic process for degradation of chlorinated solvents is known as reductive dechlorination (Bouwer et al., 1981; Bouwer, 1994). Chlorinated solvents are not utilized as a carbon source, rather as an acceptor for electrons produced during the metabolism of other oxidizable carbon (electron donor). Thus, this process cannot be termed cometabolism. Electrons or reducing equivalents formed during metabolism are accepted by the chlorinated solvent. As an example, PCE will accept electrons or reducing equivalents formed during metabolism. This results in the reduction of PCE, the concomitant release of a chloride ion and the formation of TCE. While almost any degradable carbon appears capable of driving reductive dechlorination, in most instances, a low percentage of the electrons or reducing equivalents are utilized in the reductive process. This is a function of other metabolic requirements for reducing equivalents (substrate specific) and the presence or absence of more suitable TEAs.

The behavior of the chlorinated solvent plumes can be classified into three types based on their primary substrate source (U.S. EPA, 1998a).

- *Type 1 Behavior*—Primary substrate is adequate amount of anthropogenic organic carbon, solvent plume degrades
- *Type 2 Behavior*—Primary substrate is adequate amount of native organic carbon, solvent plume degrades
- *Type 3 Behavior*—Low native organic carbon concentrations or low anthropogenic organic carbon concentrations, PCE, TCE, and DCE do not degrade

For complete detoxification, the parent chlorinated solvent must be dechlorinated in a stepwise fashion to the environmentally benign ethene. This is illustrated in Fig. 3. While PCE and TCE are readily reduced as a result of their oxidative states, the more reduced daughter products (DCE and VC) are less prone to reductive processes. These intermediates tend to accumulate in anaerobic aquifers where contaminants are allowed to naturally attenuate (Lesage et al., 1990). While this may indeed be a function of the oxidative state of the lesser-chlorinated compounds, it may also be a function of the concentration of degradable organic matter within the contaminated system.

The efficacy of an anaerobic microbial population for natural degradation of chlorinated solvents is determined by the same environmental factors as for the other systems (i.e., pH, Eh, temperature, osmotic potential), the presence of a carbon source which can be readily degraded (electron donor), and a TEA other than chlorinated solvents (Vogel et al., 1987). In addition to organic carbon and a TEA, bacteria require macro- and micronutrients, most notably nitrogen and phosphorous, for the production of DNA, RNA

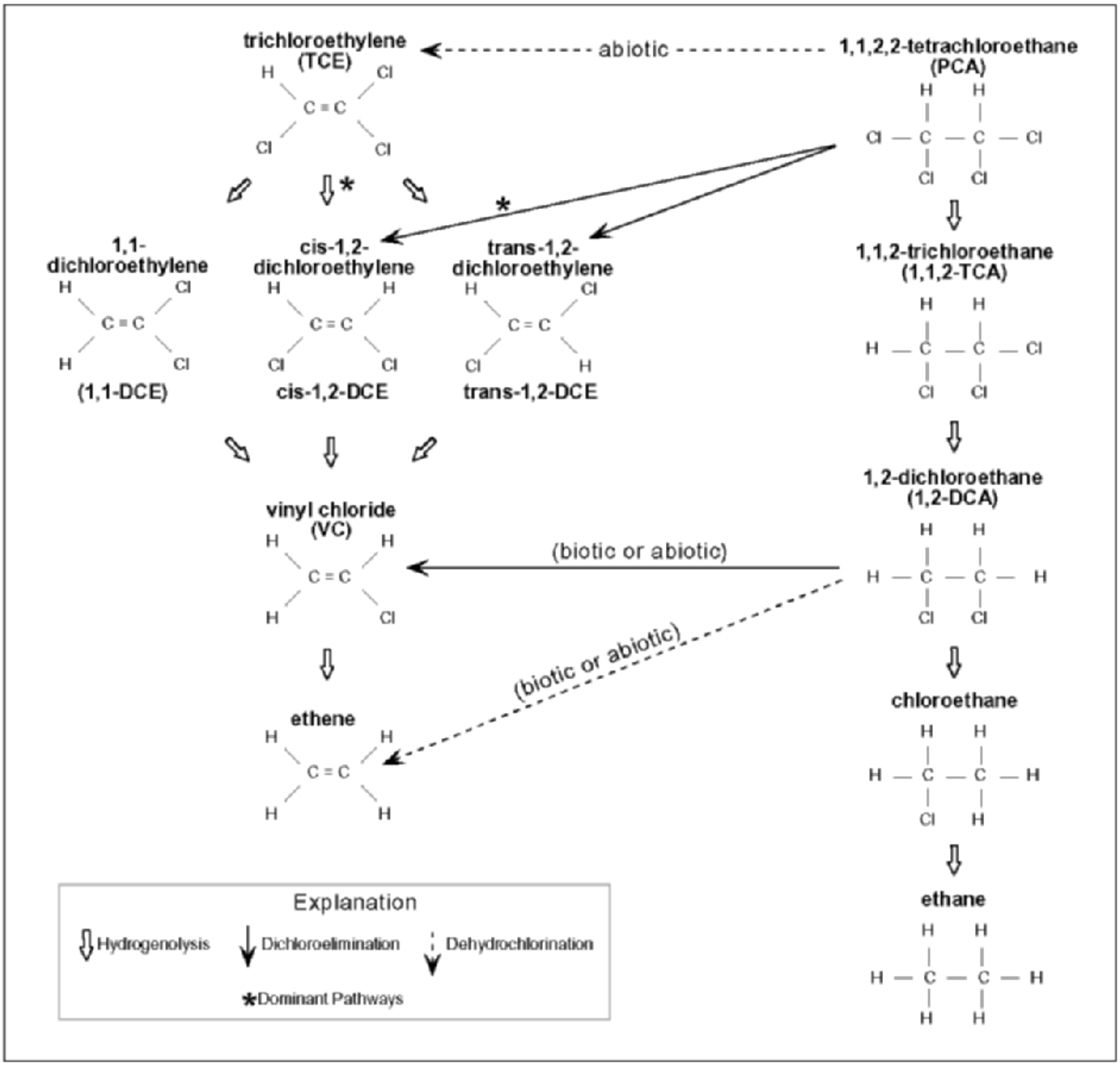

Figure 3. Anaerobic degradation pathways for trichloroethylene and 1,1,2,2-tetrachloroethane. (Modified from Chen et al., 1996; McCarty, 1996; Nyer and Duffin, 1997; and Vogel et al., 1987.).

and other needed cellular constituents. In most instances within an aquifer, sufficient macro- and micronutrients will be available for microbial processes. On the other hand, the availability of organic carbon can often be the limiting factor in the continuance of these processes. For example, during the reductive dechlorination of chlorinated solvents, the ratio of the mass of the electron donor to that of the contaminant ranges between 100/1 and 1000/1 (Bouwer, 1994).

MECHANISMS OF BIODEGRADATION

In theory, in any environment in which microbial activity occurs, there is a progression from aerobic to anaerobic (methanogenic) conditions. There is a definite sequence of electron acceptors used in this progression through distinctly different redox states (Fig. 4).

The rate, type of active microbial population, and level of activity under each of these environments are controlled by several factors. These include the concentration of the electron acceptors, substrates which can be utilized by the bacteria, and specific microbial populations leading to the progression of an aquifer from aerobic to methanogenic conditions (Salanitro, 1993). This results in a loss of organic carbon and various electron acceptors from the system as well as a progression in the types and physiological activity of the indigenous bacteria.

If microbial activity is high, the aquifer environment would be expected to progress rapidly through these conditions. The following scenario outlines a general sequence of events in which aerobic metabolism of preferential carbon sources would occur first. The carbon source may be contaminants of interest or other more readily degradable carbon which has entered

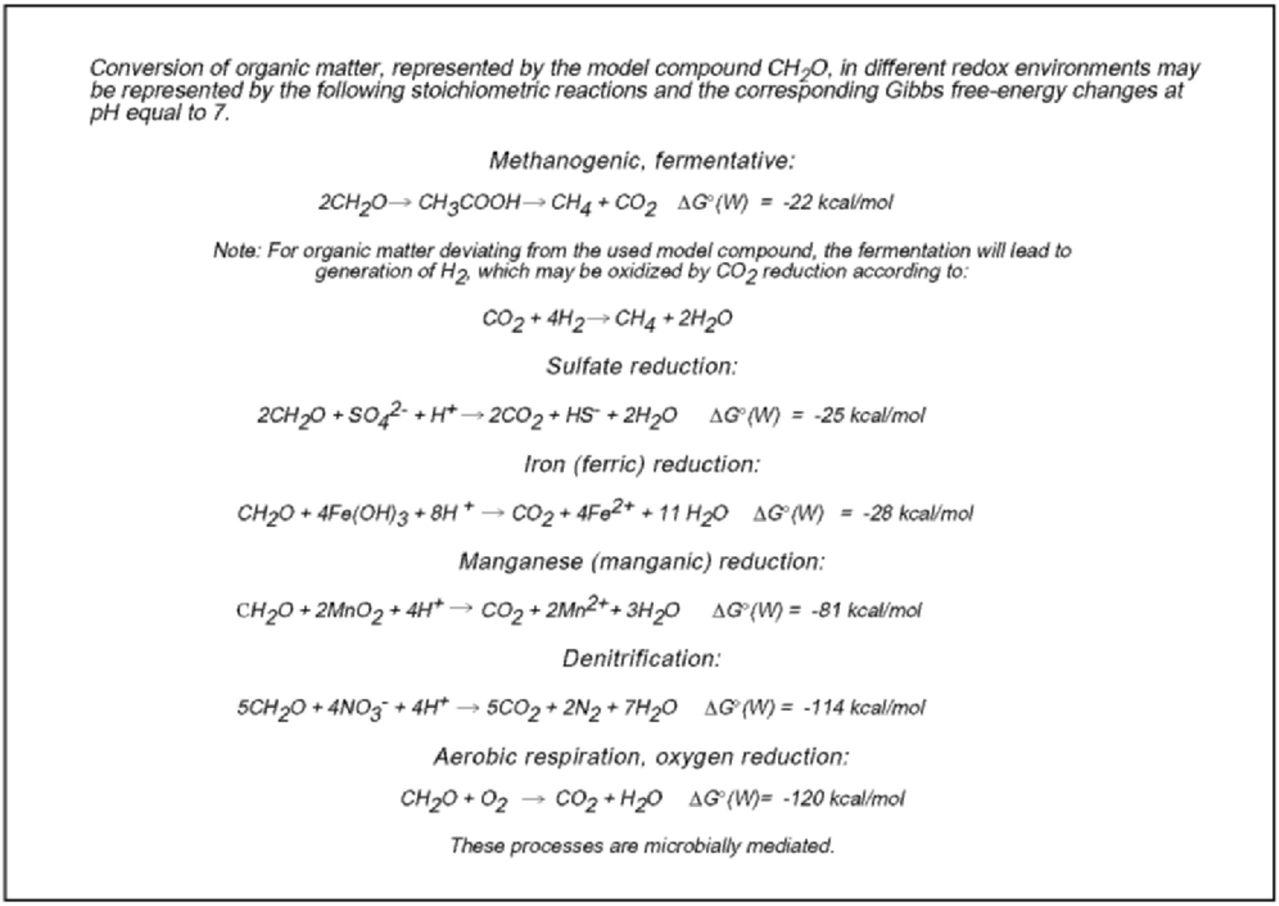

Conversion of organic matter, represented by the model compound CH_2O, in different redox environments may be represented by the following stoichiometric reactions and the corresponding Gibbs free-energy changes at pH equal to 7.

Methanogenic, fermentative:

$$2CH_2O \rightarrow CH_3COOH \rightarrow CH_4 + CO_2 \quad \Delta G^\circ(W) = -22\ \text{kcal/mol}$$

Note: For organic matter deviating from the used model compound, the fermentation will lead to generation of H_2, which may be oxidized by CO_2 reduction according to:

$$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$$

Sulfate reduction:

$$2CH_2O + SO_4^{2-} + H^+ \rightarrow 2CO_2 + HS^- + 2H_2O \quad \Delta G^\circ(W) = -25\ \text{kcal/mol}$$

Iron (ferric) reduction:

$$CH_2O + 4Fe(OH)_3 + 8H^+ \rightarrow CO_2 + 4Fe^{2+} + 11\,H_2O \quad \Delta G^\circ(W) = -28\ \text{kcal/mol}$$

Manganese (manganic) reduction:

$$CH_2O + 2MnO_2 + 4H^+ \rightarrow CO_2 + 2Mn^{2+} + 3H_2O \quad \Delta G^\circ(W) = -81\ \text{kcal/mol}$$

Denitrification:

$$5CH_2O + 4NO_3^- + 4H^+ \rightarrow 5CO_2 + 2N_2 + 7H_2O \quad \Delta G^\circ(W) = -114\ \text{kcal/mol}$$

Aerobic respiration, oxygen reduction:

$$CH_2O + O_2 \rightarrow CO_2 + H_2O \quad \Delta G^\circ(W) = -120\ \text{kcal/mol}$$

These processes are microbially mediated.

Figure 4. Examples of redox reactions involved in degradation of organic matter (expressed as the model compound CH_2O) in different redox environments. (Reprinted from *Attenuation of Landfill Leachate Pollutants in Aquifers*, box 1, p. 138, 1994, by Christensen, T.H., P. Kieldsen. H-J. Albrechtsen, G. Heron, P.H. Nielsen, P.L. Bjerg, and P.E. Holm, with permission of CRC Press, Inc., Boca Raton, FL.).

the system previously or simultaneously with the contamination event.

Oxygen-Reducing to Nitrate-Reducing Conditions

Once available oxygen is consumed, active aerobic populations begin to shift to nitrate respiration. Denitrification will continue until available nitrate is depleted, or usable carbon sources become limiting.

Nitrate-Reducing to Manganese-Reducing Conditions

Once nitrate is depleted, populations which reduce manganese may become active. Bacterial metabolism of substrates utilized by manganese-reducing populations will continue until the concentration of manganese oxide becomes limiting.

Manganese-Reducing to Iron-Reducing Conditions

When manganese oxide becomes limiting, iron reduction becomes the predominant reaction mechanism. Available evidence suggests that iron reduction does not occur until all manganese IV oxides are depleted. In addition, bacterial Mn(IV) respiration appears to be restricted to areas where sulfate is nearly or completely absent.

Iron-Reducing to Sulfate-Reducing Conditions

Iron reduction continues until substrate or carbon limitations allow sulfate-reducing bacteria to become active. Sulfate-reducing bacteria then dominate until usable carbon or sulfate limitations impede their activity.

Sulfate-Reducing to Methanogenic Conditions

Once usable carbon or sulfate limitations occur, methanogenic bacteria are able to dominate.

The ambient redox condition of the aquifer is important when determining the contribution of microbial degradation to MNA mechanisms. In the case of petroleum hydrocarbons, because of their highly reduced condition, the preferred TEA for microbial processes would be oxygen (Brown et al., 1996; Clark et al., 1997). From a thermodynamic standpoint, this is the most favorable reaction mechanism. When the soluble portion of petroleum hydrocarbons, BTEX (benzene, toluene, ethylbenzene, xylenes), are the contaminants of concern, an inverse relationship between BTEX and dissolved oxygen concentrations within a plume is indicative of the microbial metabolism of these compounds as well as other hydrocarbons in the mixture (Donaldson et al., 1992; Huesemann and Truex, 1996). Data available from various sites indicate that the natural attenuation of BTEX proceeds at higher rates under oxic conditions than normally achieved in anoxic environments (Fig. 5), with rate constants ranging from 0.3 to 1.3 percent per day when modeled as a first-order process (Chiang et al., 1989; Kemblowski et al., 1987; Salanitro, 1992; and McAllister and Chiang, 1994). Although anaerobic biodegradation of toluene and xylene under nitrate-reducing (Barbaro et al., 1992; Schocher et al., 1991), iron-reducing (Lovley et al., 1989; Lovley and Lonergan, 1990), sulfate-reducing (Beller et al., 1992a and b; Edwards and Grbić-Galić, 1992; Rabus et al., 1993), and methanogenic (Vogel and Grbić-Galić, 1986; Bouwer and McCarty, 1983; Edwards and Grbić-Galić, 1994) conditions have been extensively reported. Until recently, unequivocal biodegradation of benzene under strict anaerobiosis was not demonstrated (Edwards and Grbić-Galić, 1992; Lovley et al., 1994). According to Borden et al. (1997), even though accurate description of anaerobic biodegradation of individual BTEX constituents may not follow a simple first-order decay function, biodegradation of total BTEX seems to more closely approximate a first-order decay function.

Biodegradation of chlorinated solvents, depending on the degree of halogenation (Fig. 6), is fundamentally different from that of petroleum hydrocarbons and other oxidized chemicals (Wiedemeier et al., 1995). The preferred redox conditions for the effective degradation of these chemicals is anaerobic (exception is vinyl chloride, VC). Effective degradation of these compounds may occur only when redox conditions are below nitrate reducing.

Although under aerobic conditions, cometabolism of TCE by autotrophic bacterial populations, obtained from soil and ground water, have been demonstrated, it is generally accepted that this route of removal is limited only to low concentrations of TCE. The cometabolism of TCE proceeds in the presence of methane (Fogel et al., 1986; Henson et al., 1988; Wilson and Wilson, 1985), ammonia (Arciero et al., 1989), or toluene (Ensley, 1991; Mu and Scow, 1994; McCarty and Semprini, 1994) as cosubstrate. Due to the inherent toxicity of TCE to microorganisms responsible for degradative process (Alvarez-Cohen and McCarty, 1991a; Broholm et al., 1990), and because of the competitive inhibition between a cosubstrate and the secondary substrate for oxygenase enzymes (Alvarez-Cohen and McCarty, 1991b; Hopkins et al., 1993), special attention to concentrations of TCE and its cosubstrate is warranted. Toxic or inhibitory effects of TCE are more serious than those of 1,1,1-TCA (Broholm et al., 1990).

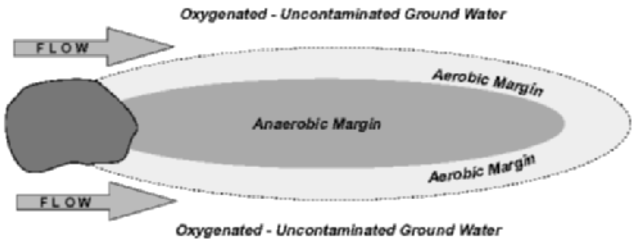

Figure 5. Plan view of a typical hydrocarbon plume undergoing natural attenuation. (Modified from *In-Situ Bioremediation of Ground Water and Geological Material: A Review of Technologies*, p. 9-8, 1993, by Norris et al., EPA/600/R-93/124.).

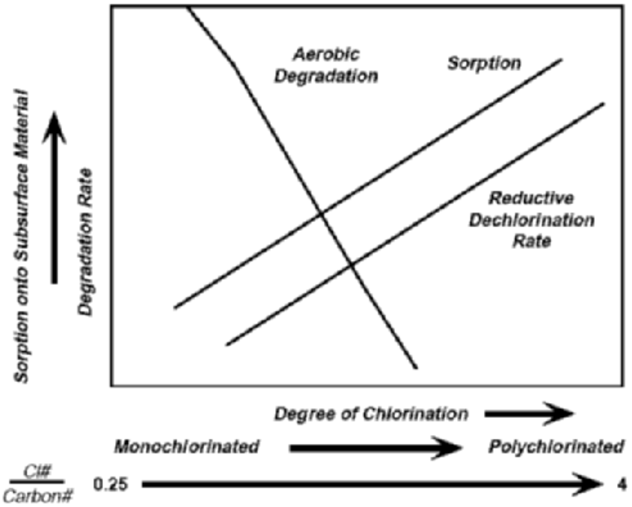

Figure 6. Relationships between degree of chlorination and anaerobic reductive dechlorination, aerobic degradation and sorption onto subsurface material. (Reprinted from *In-Situ Bioremediation of Ground Water and Geological Material: A Review of Technologies*, p. 10–19, 1993, by Norris et al., EPA/600/R-93/124.).

Microcosm studies involving anaerobic biotransformation of PCE/TCE from environmental samples including sediments (Parsons et al., 1984), ground water (Wilson et al., 1983; Sewell and Gibson, 1991), and soil (Kleopfer et al., 1985) have been documented. Also, the reductive dechlorination of PCE and other chlorinated compounds under methanogenic conditions has been reported (Vogel and McCarty, 1985 and 1987b; Vogel et al., 1987; Freedman and Gossett, 1989; McCarty, 1988). Bagley and Gossett (1990) suggested that the ability of sulfate-reducing enrichment cultures for PCE dechlorination is apparently less than that of mixed methanogenic cultures. The main by-product of anaerobic biodegradation of chlorinated ethene is VC which is more toxic than the parent compounds PCE, TCE, and DCE (Chu and Jewell, 1993). It is noted that anaerobic reduction of VC to ethylene is a slow and inefficient process (Freedman and Gossett, 1989). Gantzer and Wackett (1991) determined that dechlorination of chlorinated ethenes proceed via first-order rate constants.

The oxidation-reduction (redox) potential is a relatively simple and inexpensive indicator of the redox state of an aquifer. If the redox is positive, one can assume that dissolved oxygen is present and the system has not been stressed by biological activity (Borden et al., 1995). If the redox potential is significantly negative, it can be assumed that processes favored under aerobic conditions, such as BTEX degradation, are occurring at a substantially reduced rate. Figure 7 suggests that the redox should be -400 mV or less and the dissolved oxygen should be below

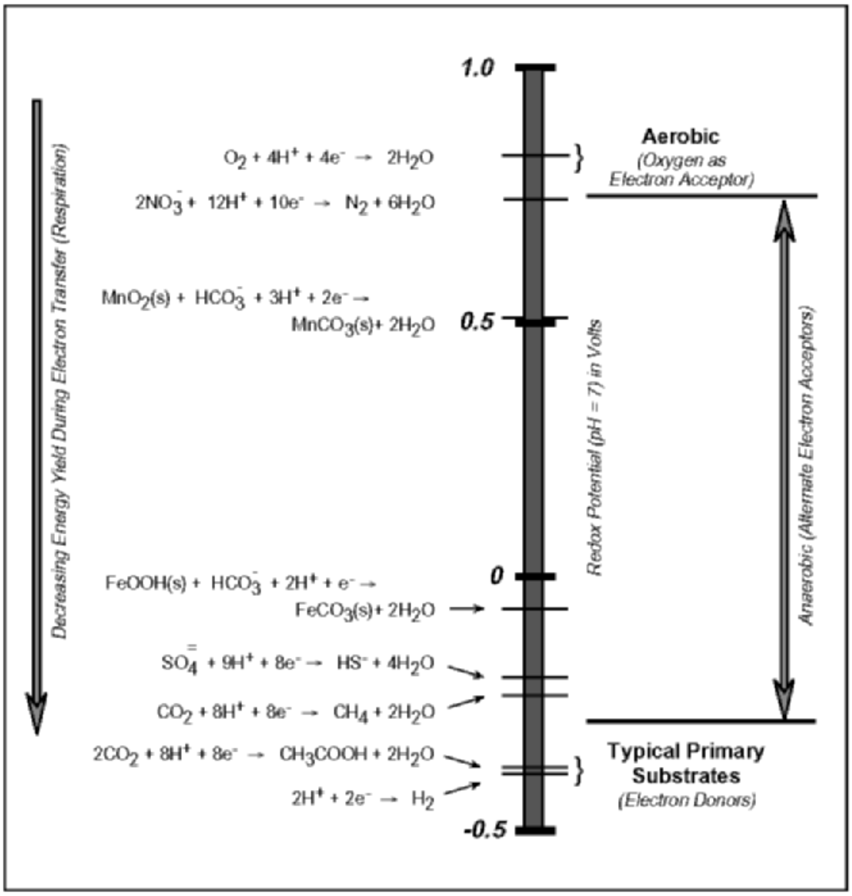

Figure 7. Important electron donors and acceptors in biotransformation processes. Redox potentials data were obtained from Stumm and Morgan (1981). (Modified from *In-Situ Bioremediation of Ground Water and Geological Material: A Review of Technologies*, p. 8-3, 1993, by Norris et al., EPA/600/R-93/124.).

0.25 mg/l before anaerobic microbial reactions could take place for the more highly chlorinated compounds (i.e., PCE, PCA). It should be stressed that one normally does not attempt to determine the actual redox conditions for comparison between different sites, rather the differences between points within a plume. Furthermore, due to the lack of internal equilibrium (Morris and Stumm, 1967) and the mixed Eh potentials of natural aqueous systems, the use of any measured master Eh as an input in equilibrium hydrogeochemical model for predicting the equilibrium chemistry of redox reactions is misleading. Instead, measuring certain sensitive species such as oxygen, Fe(II), hydrogen sulfide, or methane as qualitative guides to the redox status of the water may generate better results (Lindberg and Runnells, 1984).

In anoxic waters, where low pH and Eh exist, the reduced form of manganese, Mn(II) is favored (Stumm and Morgan, 1981). Reduction of Fe(III) and Mn(IV), due to chemical processes or microbial metabolic reactions that couple the oxidation of organic matters to the reduction of these chemical species, has a major influence on the distribution of Fe(II) and Mn(IV) in aquatic sediments, submerged soil, and ground waters (Stone and Morgan, 1984; Burdige and Nealson, 1986; Ehrlich, 1987; DiRuggiero and Gounot, 1990; Lovley, 1991). Thus, measurable Fe(II) or Mn(II) may indicate suboxic conditions in the absence of detectable oxygen concentration (Higgo et al., 1996).

In addition to establishing background conditions away from the plume, dissolved oxygen, nitrate, manganese, iron, sulfate, and sulfide should be measured along the axis of the plume, as well as transverse to it, in order to characterize biological activity with respect to the redox state at those locations. This information will allow an estimation of the current redox state at various parts of the contaminated plume, thereby defining the types of reactions that may take place.

The rate of change in the concentration of these parameters can be useful as input to predictive models. This set of data will also characterize the abundance of the principal electron acceptors, oxygen, nitrate, and sulfate, to allow an estimate of how long natural attenuation will remain a viable remedial alternative.

Another approach that may be used to indicate the terminal electron acceptor process (TEAP) predominant in the areas of contamination is the hydrogen (H_2) concentration (Lovley and Goodwin, 1988). Hydrogen concentrations for the various terminal electron acceptors are shown in Table 2 (Chapelle et al., 1995).

Parameters that investigators should analyze for petroleum hydrocarbons include dissolved oxygen, nitrate, Fe(II), sulfate, redox potential, pH, Mn(IV), dissolved methane, and total petroleum hydrocarbons (ASTM, 1998). The majority of these parameters can be determined using field measurements, Hach™ kits, and/or CHEMetrics™ test kits. Since methane is produced after other TEAs

Table 2. Range of Hydrogen Concentrations for a Given Terminal Electron-Accepting Process. (Data from Chapelle et al., 1995)

Terminal Electron Accepting Process	Hydrogen (H_2) Concentration (nanomoles per liter)
Denitrification	<0.1
Iron (III) Reduction	0.2 to 0.8
Sulfate Reduction	1 to 4
Methanogenesis	5–20

(nitrate, iron, sulfate) are depleted, dissolved methane data is superior to contaminant data (Underground Storage Tank Technology Update, 1998). The parameters for chlorinated solvents may include: temperature, redox potential, DO, sulfide, Fe(II), methane, ethane/ethene, alkalinity, pH, sulfate, nitrate, chloride, dissolved organic carbon, and hydrogen. Since Fe(III) may be dissolved from aquifer matrix, Fe(II) is measured as proxy for biodegradation due to iron reduction. Although there may be a correlation between sediment redness and the hematite content of soil, when a soil sample represents a mixture of several iron species, the color is not a useful indicator (Heron et al., 1994).

The microbial activities of a site are thus determined by the dissolved organic carbon, presence of macro- and micronutrients and the TEA (Semprini et al., 1995). The presence and concentration of each will determine not only the activity, but the predominant population.

Different levels of QA/QC may be required for those analyses determined in the field versus those performed under laboratory conditions (U.S. EPA, 1996; Klusman, 1980; Shampine et al., 1992; Koterba et al., 1996). For example, dissolved iron and oxygen, redox, and temperature must be determined on-site (Shelton, 1994; Wood, 1981) using field test kits because of the deterioration that would normally occur between the time of sample collection and that of arrival at the laboratory. On the other hand, parameters such as metals, organics (Shelton and Capel, 1994; Fishman and Friedman, 1985), and bacteria can be properly preserved by cooling, capping, or chemical fixation, and thereby subjected to a higher level of QA/QC.

FACTORS AFFECTING THE DEMONSTRATION OF NATURAL ATTENUATION

Hydrogeology

State and federal agencies are increasingly relying on risk based corrective action (RBCA) and/or MNA for cleanup of contaminated sites (Brady et al., 1998). The American Society for Testing and Materials released on RBCA protocol two years ago (ASTM, 1998) and recently finalized an MNA protocol. The U.S. EPA has recently published a directive on MNA (U.S. EPA, 1999) and a protocol for chlorinated solvents (U.S. EPA, 1998a). Since RBCA and MNA incorporate no safety factors to reduce contaminant concentrations compared to active remediation technologies, these strategies rely solely on accurate and high quality hydrogeological site characterizations (Boulding, 1993a and b; U.S. EPA, 1997a) to demonstrate adequate public protection.

Adequate monitoring is one of the most important facets of proving that natural bioremediation is occurring in the subsurface. Many factors other than natural attenuation or bioremediation can have an effect on the observed concentration of contaminants at a monitoring well (Black and Hall, 1984). The infiltrating precipitation into a system may have a profound effect on contaminant concentrations, especially if the contaminants are light nonaqueous phase liquids (LNAPLs) such as petroleum hydrocarbons (Kemblowski and Chiang, 1990; Lenhard and Parker, 1990). High rates of infiltration may lower the apparent concentration of contaminants due to dilution. Pettyjohn (1982) has shown that there are actually two time periods after a precipitation event when dilution may have an effect on monitoring. The first is a few hours after a precipitation event where flow is through macropores or "wormholes" to the water table, and the second is a few days after with flow through the vadose zone. The time required for both of these events to occur will be a function of the number and size of the macropores, overall permeability of the unsaturated zone and depth to the water table. Monitoring during these events can result in an apparent decrease in the concentration of contaminants by dilution, or if contaminants are present in the vadose zone, an increase in concentrations due to their infiltration. Seasonal variations may also occur in flow paths within an aquifer (Schmidt, 1977). These variations may cause an apparent increase or decrease in concentration of contaminants due to dilution or plume movement. Although dilution is considered to be a part of natural attenuation, this reduction in concentration should not be attributed to degradation.

A major problem with monitoring wells is that a sample from the uncontaminated portion of the aquifer may be a composite of contaminated water from the plume drawn into the cone of depression along with clean water from the aquifer. This will result in an apparent increase in contamination. The reverse is also often possible if the monitoring well is pumped for an extended period prior to sampling, the amount of clean water coming into the well in relation to contaminated water will result in an apparent decrease in the concentration of the contaminants of interest. These problems are often exacerbated by using well screens that are too long (greater than 2 meters), inconsistent screened intervals, and inappropriate sampling methods (Church and Granato, 1996).

Pumping wells other than those designed for monitoring (such as those in an interdiction field, or water supply wells for municipalities or irrigation) may influence the movement of a plume as well as flow lines within an aquifer. Depending on when these wells were designed and constructed, how they are pumped, and how they affect plume movement, apparent decreases in concentration may be observed in monitoring wells (Martin-Hayden and Robbins, 1997; Martin-Hayden et al., 1991; Robbins, 1989; Robbins and Martin-Hayden, 1991).

Monitoring programs should be designed such that these concerns are taken into account (Zeeb et al., 1999).

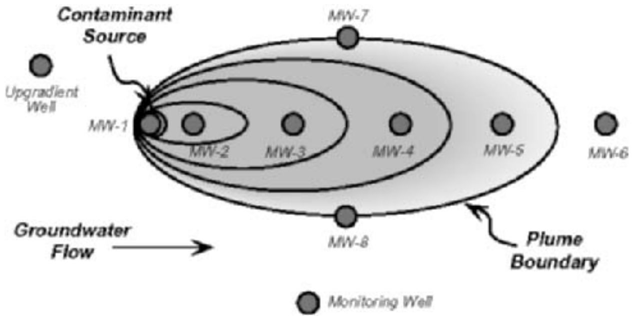

Figure 8. Recommended groundwater monitoring well network for demonstrating natural attenuation. (Reprinted from *A Practical Approach to Evaluating Natural Attenuation of Contaminants in Ground Water*, p. 166, 1994, by P.M. McAllister and C.Y. Chlang. Reprinted by permission of *Ground Water Monitoring & Remediation*, Westerville, OH, Copyright 1994. All rights reserved.).

A good monitoring program will require sampling (Puls and Barcelona, 1996) of not only monitoring wells which are completed into the plume, but also monitoring wells outside of the contaminated zone in order to establish background conditions (Fig. 8). Data from monitoring wells in the contaminated portion are then compared to background wells. The number and location of monitoring wells are not only determined by plume geometry and ground-water flow (ASTM, 1991), but by the degree of confidence required to statistically demonstrate that natural attenuation is taking place, to estimate the rate of attenuation processes, and to predict the time required to meet established remediation goals. These issues will be addressed in detail in the EPA guidance for long term monitoring (LTM) for MNA which is in preparation.

MNA Monitoring Well Network Considerations

RBCA and MNA both rely on sentinel wells for early warning signals of plume migration. The installation of monitoring wells (Aller et al., 1991) to adequately identify contaminant concentrations is of paramount importance in determining the overall contribution of biological processes to a reduction in either concentration or mass of contamination. Monitoring wells which are to be used to determine the contribution of natural bioremediation to site cleanup cannot be located until sufficient knowledge of the aquifer system is obtained (Zeeb et al., 1999). Information that must be obtained before installation include depth to water table, hydraulic conductivity (Molz et al., 1994) and gradient, direction of ground-water flow, storage coefficient or specific yield, vertical and horizontal conductivity distribution, direction of plume movement and the effects of any man-made or natural influences (i.e., lagoons or seeps) on the aquifer system. It is also important to determine if the hydraulic gradient is affected seasonally. Also, sentinel well screen depth and length are important. Often it is advantageous to use short screens to minimize averaging of vertical water quality differences (Martin-Hayden and Robbins, 1997).

The location, number (ASTM, 1995) and other pertinent data regarding monitoring wells (U.S. EPA, 1998b) for the evaluation of MNA should be determined on a site-specific basis. The design of the monitoring network will be determined by the size of the plume, site complexity, source strength, ground-water/subsurface water interactions, distance to receptors and the confidence limits each party involved wishes to place in the data obtained. By necessity, the denser the monitoring network, the greater the degree of confidence one may place in the data. The wells should be capable of monitoring singular flow paths within a plume's course and subsequent movement of contaminants along these flow paths. One way to determine natural attenuation is to determine the concentration of appropriate parameters at one location and sample the same volume of water for the same parameters at some distance downgradient. It is generally impractical to monitor flow paths within a plume with the exception of the plume axis which is the only flow line that can be located with any reasonable level of certainty.

A second way to monitor for natural attenuation assumes that the plume is in equilibrium. While less costly and less time-consuming than monitoring single flow paths for contaminants where a mass balance cannot be performed, the confidence level for data obtained is significantly less when plume equilibrium is assumed. Problems arise because plumes are never in complete equilibrium. Monitoring wells even a few feet apart differ significantly in observed concentrations of contaminants. It has been suggested that internal tracers such as trimethylbenzene (TMB), a biologically recalcitrant compound, can increase the confidence one might place in this method. One simply uses the difference between TMB concentrations at the upgradient and downgradient points to measure the observed loss which can be attributed to other factors such as dispersion. One must be cautioned, however, when using TMB as a conservative tracer due to its degradation under anaerobic conditions and the resulting production of aromatic acid intermediates (Fang et al., 1997).

To monitor the anaerobic degradation of chlorinated solvents, monitoring of singular flow paths may not be necessary, if a pseudo-mass balance can be performed. This method assumes that at each monitoring point either the parent chlorinated solvent or daughter product should be observed. Using PCE as an example at each monitoring point, samples are analyzed for PCE, TCE, DCE, VC, and ethene. The total concentration of analytes is compared on a molar basis. Changes in total molar concentrations are assumed to be losses due to flaws in the monitoring system. The pseudo-mass balance method does not account for physical loses due to dispersion, sorption, or dilution, therefore, one would also have to utilize some method, such as a conservative tracer, to account for these other processes. To place even greater confidence in data, the time frame for analysis must be sufficient to allow for differences in subsurface mobility of the various chlorinated solvents. They will not arrive at the downgradient monitoring point at the same time.

Often hazardous waste sites are monitored for natural attenuation using monitoring wells which have previously been installed, especially if plume equilibrium is assumed. This in itself presents a number of problems, especially if the wells were installed without adequate knowledge of the subsurface and plume movement. If it can be demonstrated that contamination exists between these wells, by using tracer study for example, they can be used as appropriate monitoring points for assessing natural attenuation. Reinhard and Goodman (1984) used chloride as a reference to investigate the behavior of the trace organic compounds in leachate plumes of two sanitary landfills. Kampbell et al. (1995) and Wiedemeier et al. (1996) described the use of TMB to normalize for changes in BTEX concentration due to abiotic processes of dispersion, dilution, sorption, and volatilization.

It is also necessary to determine what contribution physical processes between sequential monitoring wells have on apparent reduction of chemicals. It may also be necessary to construct new monitoring wells which are offset to present wells or to perform borings near existing wells.

Statistical Validity of Data for MNA

Adequate monitoring is critical (Reinhard and Goodman, 1984), especially when considering natural attenuation as a remedial alternative. The number of sampling points (Barcelona et al., 1994 and 1985) and sampling rounds are often insufficient to establish statistically valid trends, given the natural variability in ground water quality, the variability due to pumping and sampling, and differences between analytical laboratories. Schmidt (1977) lists a number of conditions in which large fluctuations in water quality may be noted, indicating that minor changes in water chemistry may be related to sampling procedures. In order to minimize the effects of natural seasonal variations, sampling, and subsurface heterogeneity on ground water quality, any natural attenuation monitoring program (Gibbons, 1994; Gilbert, 1987) should be based on a detailed statistical evaluation of pertinent data (Hardin and Gilbert, 1993; McDonald and Erickson, 1994; O'Brien et al., 1991; O'Brien, 1997). Although it is recognized that intensive monitoring is expensive, increasing sampling points and frequencies, along with acceptable QA/QC procedures, will give more statistically reliable information. Caution should be exercised when drawing conclusions from limited data sets, especially when attempting to model complex situations.

The number of samples required for evaluating natural attenuation is intrinsically a function of a preselected confidence error and the variance of the data. Quarterly samples collected for a year or two, for example, may not be adequate for evaluating an overall reduction in the mass of contaminants from a particular monitoring well. Sampling frequencies such as this offer small windows for viewing contaminant reduction rates. Contaminant fronts may or may not have reached the monitoring points at the time of sampling, or as stated previously, any number of processes such as infiltration, dilution, and sorption may bias results when addressing only the bioremediation component of natural attenuation.

If the reduction of contaminant mass is to be determined by temporal trends (least squares analysis), for example, the statistical confidence is based on the variance of the data and the square root of the number of samples. In a least square analysis, the correlation coefficient (r) is a function of the degrees of freedom (df) which, in this case, is the number of observations minus two (n − 2). For example, if a well were sampled quarterly for two years, the degrees of freedom would be 6 (8 − 2) which requires a correlation coefficient of about 0.8 to demonstrate statistical significance at a 95 percent level of confidence (an r of 1.0 denotes a perfect correlation). The point is that one must sample with enough frequency, over a protracted period in order to obtain a statistically meaningful correlation between the reduction of contaminant mass and time.

MNA Degradation Rate Constant Considerations

Precision and accuracy in estimation of rate constants are essential to conclude how quickly the ground-water plume will be cleaned up following the source control (McNabb and Dooher, 1998). Reviews by Vogel et al. (1987a) and Howard et al. (1991) contain a compilation of chlorinated solvents rate constants. Experimental data on the neutral and base-catalyzed abiotic hydrolysis rates of chlorinated ethanes and ethenes was determined by Jeffers et al. (1989). Chapelle et al. (1996) integrated field and laboratory data to estimate rates of petroleum hydrocarbon biodegradation.

In the derivation of the rate constants, it is of considerable significance to calculate concentrations at the point of compliance, compare rates at the site to those in the literature to determine if the site is behaving like other sites, and predict changes caused by fluctuations in flow (Weaver et al., 1996). Although microcosms are used as an effective tool to determine the biodegradation potentials, the use of the microbiological laboratory data for calculation of rate constants may be inappropriate since they are not always representative of the degradative rate(s) under the field conditions (U.S. EPA, 1997b). Results from laboratory studies may significantly over- or underestimate biodegradation rates if environmental conditions in the laboratory differ from conditions in the field (Borden, 1994). Rifai (1997) also points out that, although useful in evaluating the biodegradation potential, microcosms can disrupt the normal structure of ecosystems and prevent the direct extrapolation of microcosm-determined biodegradation rates to field scale. If rate constants for attenuation of chlorinated contaminants are to be used for exposure assessments, it is necessary to estimate the residence time of the contaminants in the aquifer as accurately as possible (Molz and Boman, 1997). One should also be cautioned to not substitute literature biodegradation rates in the place of site specific values.

Often, first-order kinetics obtained from field studies are used to approximate the degradation mechanism. Wiedemeier et al. (1996) described two methods to estimate first-order rate constants: (1) the use of a conservative tracer, a biologically recalcitrant compound found in the dissolved contaminant plume; and (2) the interpretation of a steady state contaminant plume as proposed by Buscheck and Alcantar (1995). The later method is founded on a one-dimensional steady state analytical solution to the advection-dispersion equation presented by Bear (1979).

Modeling as a Predictive Tool for MNA

MNA requires two types of models: (1) conceptual and (2) fate and transport. In order to understand ground water-flow and contaminant movement (Bear et al., 1992), the construction of a three-dimensional conceptual model of the site must be an integral part of any MNA Work Plan. A comprehensive conceptual model should be used as a clear and concise aid for the general understanding of the nature of the site, the acquired sampling results, and to indicate where additional sampling efforts should be directed. Once a conceptual model has been accepted, a period of monitoring is required to verify that the forecast of the conceptual model is adequate. Only when sufficient quantitative site characterization data are generated and the conceptual model is well developed, can an appropriate analytical or numerical fate and transport model be chosen for the site.

Care should be exercised when choosing the models to predict the fate and transport of contaminants in the subsurface. This may also include the use of screening models such as BIOSCREEN and BIOCHLOR (Newell et al., 1996 and 1998). Appropriateness of the model to the actual hydrogeologic situation, assumptions, limitations, and manner of application are all considerations (Corapcioglu and Baehr, 1987; Carey et al., 1988). The overall effectiveness of the model for predicting fate and transport of contaminants at a particular site depends on all of these factors. Caution should also be exercised when drawing conclusions from limited data sets, especially when modeling complex situations. The validity of the input data is critical in determining the accuracy of predictions made with the model.

Once a model is chosen, it may be applied using the site data (Hunt et al., 1988) and calibrated. Calibration is a process of careful modification of site hydrogeologic or contaminant transport parameters over numerous simulations to identify a set of parameters which generate simulation results which closely match field measured values of hydraulic head for the flow model and contaminant concentrations for the transport model. Calibration of the model results to observed values requires that an acceptable range of error be identified for each calibration target. This range will depend on the model purpose and also on the amount and reliability of the field data (Kresec, 1997). Once calibrated, the fate and transport model can then be used to predict the future extent and concentration of a dissolved contaminant plume by simulating the combined effects of advection, dispersion, sorption, and biodegradation (Rifai et al., 1989).

Many times during calibration, if a model does not fit observed concentrations, it is assumed that the biodegradation coefficient is the proper parameter to be adjusted. Using biodegradation to adjust a model without supporting field data should not be done until all abiotic mechanisms for reduction are explored. When using a model which incorporates a biodegradation term, care should be taken to verify that assumptions made about degradation rates and the amount and activity of biomass are valid for the site in question. Degradation rates are sensitive to a wide array of field conditions which have been discussed previously. Extrapolation of laboratory derived rates to a site can also lead to significant errors. Likewise, using models to derive degradation rates from limited field data where abiotic variables are not well defined can be misleading. Models can be useful tools in determining plume movement (Mercer, 1998) and the contribution of natural attenuation (Rifai et al., 1995) to reductions in contaminant mass provided that all model inputs are correct, particularly those associated with biodegradation rates. Kinetic constants derived from laboratory microcosms or other sites are generally not useful on a wide scale to predict overall removal rates. Site specific degradation rates should be developed and incorporated into a model.

SUMMARY

The behavior of a contamination plume, whether stable, shrinking, or expanding, is the primary evidence for the occurrence of natural attenuation. In the majority of the cases, historical data to indicate the status of a plume are not available. In these events, there are at least four basic conditions which must be present to confirm that natural attenuation processes are taking place. These include, but are probably not limited to:

1. The points of sampling must be on flow lines from the source of contamination or an upgradient point of observation. It must be demonstrated that the downgradient observations accurately reflect the abiotic and biotic processes which have occurred between the two points. Ideally, one would sample the same volume of water at the downstream point that was sampled earlier at the source or upgradient observation well. Since this is rarely practical, it must be assumed that the plume is in equilibrium with respect to natural attenuation processes between the two points of observation. If these conditions are not satisfied, any downgradient measurement of contaminant concentration must be lower than the true value, and therefore, exaggerate the effectiveness of natural attenuation. Confidence levels may be enhanced by 1) increasing the number of observation points and times at fixed frequencies, and 2) use of a conservative tracer.
2. There must be a reduction in contaminant mass or concentration. One could argue that natural attenuation results in a reduction in contaminant concentration by sorption, volatilization, or dilution, with the only loss of mass being that of volatilization to the atmosphere. Natural bioremediation, on the other hand, must result in a reduction of mass of the contaminants of concern by the eventual conversion to environmentally acceptable compounds.
3. Site geochemistry must assure that conditions are right for reduction of contaminant concentration,

such as the presence of mineral nutrients and electron acceptors, the state of redox, temperature, and pH.

4. Daughter products of contaminants must be present, perhaps with indicators of mineralization. For example, claims of intrinsic or natural bioremediation need to be supported by data including the relationship between the mass loss and the loss of oxygen. In addition to the biological utilization of oxygen, nitrate, and sulfate, these natural attenuation processes often result in the creation of by-products such as dissolved Fe(II), Mn(II), HCO_{3^-}, $CO_{3^=}$ and methane.

The use of MNA or passive remediation at contaminated ground-water sites is gaining attention both within the scientific and regulatory communities. The OSWER directive recommends that MNA be applied concurrently with or subsequently to active measures such as source control or active remedial technologies. Selection of MNA as a remedy or part of a remedy can be advantageous since it may minimize the transfer of contaminants to other media, is less intrusive to the environment, may be applied to all or part of a site, and overall remedial costs may be lower than for an active remedy. There are several factors which may limit the application of natural attenuation as a remedial alternative. They are: (1) the longer time frame that may be required to achieve remedial goals, (2) site characterization investigations that may be more extensive and costly, (3) the added responsibility for long-term monitoring and costs, (4) toxicity of byproducts, (5) potential for continued contaminant migration, and (6) required alternatives if natural attenuation fails to meet established goals.

Under proper conditions, MNA along with source removal, long-term monitoring, and land use restrictions might be selected over more expensive conventional technologies. There will be other sites where natural attenuation will not be acceptable as a remedial alternative because of regulatory constraints or the site conditions are not favorable for its application. In the end, the selection of a remedial technology at a specific site will be determined by time constraints in obtaining remediation objectives, the hydrogeology and geochemistry at the site, the contaminants of concern, regulatory constraints, and considerations of environmental exposure and cost.

DISCLAIMER

The U.S. Environmental Protection Agency through its Office of Research and Development partially funded and collaborated in the research described here under Contract No. 68-C4-0031 to Dynamac Corporation. It has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

QUALITY ASSURANCE STATEMENT

All research projects making conclusions or recommendations based on environmentally related measurements and funded by the Environmental Protection Agency are required to participate in the Agency Quality Assurance Program. This project did not involve physical measurements and as such did not require a QA plan.

READING LIST

Aller, L., T.W. Bennett, G. Hackett, R.J. Petty, J.H. Lehr, H. Sedoris, D.M. Nielsen, and J.E. Denne. 1991. *Handbook of Suggested Practices for the Design and Installation of Groundwater Monitoring Wells*. EPA/600/4-89/034. Environmental Monitoring Systems Laboratory, Las Vegas, NV. 398 pp.

Alvarez-Cohen, L., and P.L. McCarty. 1991a. Effects of toxicity, aeration, and reductant supply on trichloroethylene transformation by a mixed methanotrophic culture. *Appl. Environ. Microbiol.* **57**: 228–235.

Alvarez-Cohen, L., and P.L. McCarty. 1991b. Product toxicity and cometabolic competitive inhibition modeling of chloroform and trichloroethylene transformation by methanotrophic resting cells. *Appl. Environ. Microbiol.* **57**(4): 1031–1037.

Arciero, D., T. Vannelli, M. Logan, and A.B. Hooper. 1989. Degradation of trichloroethylene by the ammonia-oxidizing bacterium *Nitrosomonas europaea*. *Biochem. Biophys. Res. Commun.* **159**: 640–643.

Arneth. J.-D., G. Milde, H. Kerndorff, R. Schleyer. 1989. The 15 most frequently detected organic compounds in groundwater at waste disposal sites in Germany and the U.S. p. 399. In: P. Baccini (ed.), *The Landfill*. Springer-Verlag, New York, Inc., New York, NY.

ASTM. 1991. *Standard Guide for Selection of Aquifer Test Method in Determination of Hydraulic Properties by Well Techniques*. Standard D-4043-91. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1995. *Standard Practice for Design and Installation of Ground Water Monitoring Wells in Aquifers*. Standard D-5092-90. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1998. *Standard Guide for Remediation of Groundwater by Natural Attenuation at Petroleum Release Sites*. E-1943-98. West Conshohocken, Pennsylvania.

Bagley, D.M., and J.M. Gossett. 1990. Tetrachloroethene transformation to trichloroethene and cis-1,2-aichloroethene by sulfate-reducing enrichment cultures. *Appl. Environ. Microbiol.* **56**(8): 2511–2516.

Barbaro, J.R., J.F. Barker, L.A. Lemon, and C.I. Mayfield. 1992. Biotransformation of BTEX under anaerobic denitrifying conditions: Field and laboratory observations. *J. Contam. Hydrol.* **11**: 245–277.

Barcelona, M.J., J.A. Helfrich, and E.E. Garske. 1985. Sampling tubing effects on groundwater samples. *Anal. Chem.* **57**(2): 460–464.

Barcelona, M.J., H.A. Wehrmann, and M.D. Varijen. 1994. Reproducible well purging procedure and VOC stabilization criteria for ground-water sampling. *Ground Water* **32**(1): 12–22.

Bear, J. 1979. *Hydraulics of Groundwater*. McGraw-Hill, New York, NY.

Bear, J., M.S. Beljin, and R.R. Ross. 1992. *Fundamentals of Ground-Water Modeling*. EPA/540/S-92/005. ORD, Washington, DC.

Beller, H.R., D. Grbić-Galić, and M. Reinhard. 1992a. Microbial degradation of toluene under sulfate-reducing conditions and the influence of iron on the process. *Appl. Environ. Microbiol.* **58**(3): 786–793.

Beller, H.R., M. Reinhard, and Grbić-Galić. 1992b. Metabolic by products of anaerobic toluene degradation by sulfate-reducing enrichment cultures. *Appl. Environ. Microbiol.* **58**(9): 3192–3195.

Blake, S.B., and R.A. Hall. 1984. *Monitoring petroleum spills with wells–Some problems and solutions*. pp. 305–310. In: *Proceedings of the Fourth National Symposium on Aquifer Restoration and Groundwater Monitoring*. Dallas, TX, 23–25 May.

Borden, R.C. 1994. Natural Bioremediation of Hydrocarbon-Contaminated Ground Water. p. 177. In: R.D. Norris, R.E. Hinchee, R. Brown, P.L. McCarty, L. Semprini, J.T. Wilson, D.H. Kampbell, M. Reinhard, E.J. Bouwer, R.C. Borden, T.M. Vogel, J.M. Thomas, and C.H. Ward (eds.), *Handbook of Bioremediation*. Lewis Publishers, Boca Raton, FL.

Borden, R.C., C.A. Gomez, and M.T. Becker. 1995. Geochemical indicators of intrinsic bioremediation. *Ground Water* **33**(2): 180–189.

Borden, R.C., J.H. Melody, M.B. Shafer, and M.A. Barlaz. 1997. *Anaerobic biodegradation of BTEX in aquifer material*. EPA/600/S-97/003.

Boulding, J.R. 1993a. *Subsurface Characterization and Monitoring Techniques, A Desk Reference Guide, Volume I: Solids and Ground Water*. EPA/625/R-93/003a.

Boulding, J.R. 1993b. *Subsurface Characterization and Monitoring Techniques, A Desk Reference Guide, Volume II: The Vadose Zone, Field Screening and Analytical Methods*. EPA/625/R-93/003b.

Bouwer, E.J. 1994. Bioremediation of chlorinated solvents using alternate electron acceptors. p. 149. In: R.D. Norris, R.E. Hinchee, R. Brown, P.L. McCarty, L. Semprini, J.T. Wilson, D.H. Kampbell, M. Reinhard, E.J. Bouwer, R.C. Borden, T.M. Vogel, J.M. Thomas, and C.H. Ward (eds.), *Handbook of Bioremediation*. Lewis Publishers, Boca Raton, FL.

Bouwer, E.J., and P.L. McCarty. 1983. Transformation of 1- and 2-carbon halogenated aliphatic organic compounds under methanogenic conditions. *Appl. Environ. Microbiol.* **45**(4): 1286–1294.

Bouwer, E.J., B.E. Rittman, and R.L. McCarty. 1981. Anaerobic degradation of halogenated 1- and 2-carbon organic compounds. *Environ. Sci. Technol.* **15**(5): 596–599.

Brady, P.V., M.V. Brady, and D.J. Borns. 1998. *Natural Attenuation - CERCLA, RBCA's, and the Future of Environmental Remediation*. Lewis Publishers, Boca Raton, FL. 245 pp.

Broholm, K., B.J. Jensen, T.H. Christensen, and L. Olsen. 1990. Toxicity of 1,1,1-trichlororethane and trichloroethene on a mixed culture of methane-oxidizing bacteria. *Appl. Environ. Microbiol.* **56**(80): 2488–2493.

Brown, R.A., R.E. Hinchee, R.D. Norris, and J.T. Wilson. 1996. Bioremediation of petroleum hydrocarbons: A flexible, variable speed technology. *Biorem. J.* **6**(4): 95–108.

Burdige, D.J., and K.H. Nealson. 1986. Chemical and microbiological studies of sulfide-mediated manganese reduction. *Geomicrobiol. J.* **4**(4): 361–387.

Buscheck, T.E.G., and C.M. Alcantar. 1995. *Regression techniques and analytical solutions to demonstrate intrinsic bioremediation*. In: *Proceedings of the 1995 Battelle International Conference on In-Situ and On Site Bioreclamation*.

Carey, G.R., P.J. Van Geel, J.R. Murphy, E.A. McBean, and F.A. Rovers. 1998. Full-scale field application of a coupled biodegradation-redox model (BIOREDOX). pp. 213–218. In: G.B. Wickramanayake and R.H. Hinchee (eds.), *Natural Attenuation of Chlorinated Solvents*. Battelle Press, Columbus, OH.

Chapelle, F.H., P.B. McMahon, N.M. Dubrovsky, R.F. Fujii, E.T. Oaksford, and D.A. Vroblesky. 1995. Deducing the distribution of terminal electron-accepting processes in hydrologically diverse groundwater systems. *Water Resour. Res.* **31**(2): 359–371.

Chapelle, F.H., P.M. Bradley, D.R. Lovley, and D.A. Vroblesky. 1996. Measuring rates of biodegradation in a contaminated aquifer using field and laboratory methods. *Ground Water* **34**(4): 691–698.

Chen, C., J.A. Puhakka, and J.F. Ferguson. 1996. Transformation of 1,1,2,2-tetrachloroethane under methanogenic conditions. *Environ. Sci. Technol.* **30**: 542–547.

Chiang, C.Y., J.P. Salanitro, E.Y. Chai, J.D. Colthart, and C.L. Klein. 1989. Aerobic biodegradation of benzene, toluene, and xylene in a sandy aquifer: Data analysis and computer modeling. *Ground Water* **27**(6): 823–834.

Christensen, T.H., P. Kjeldsen, H-J. Albrechtsen, G. Heron, P.H. Nielsen, P.L. Bjerg, and P.E. Holm. 1994. Attenuation of landfill leachate pollutants in aquifers. *Crit. Revs. Environ. Sci. Tech.* **24**(2): 119–202.

Chu, K.H., and W.J. Jewell. 1993. Treatment of tetrachloroethylene with an anaerobic attached film process. *J. Environ. Eng. Am. Soc. Civ. Engrs.* **2**: 155–162.

Church, P.E., and G.E. Granato. 1996. Bias in ground-water data caused by well-bore flow in long-screen wells. *Ground Water* **34**(2): 262–273.

Clark, D., D. Ours, and L. Hineline. 1997. Passive remediation for groundwater impacted with hydrocarbons. *Rem. Manage.* **1**: 50–45.

Corapcioglu, M.Y., and A.L. Baehr. 1987. A compositional multiphase model for groundwater contamination by petroleum products, 1. Theoretical considerations. *Water Resour. Res.* **23**(1): 191–200.

Davis, J.W., N.J. Klier, and C.L. Carpenter. 1994. Natural biological attenuation of benzene in ground water beneath a manufacturing facility. *Ground Water* **32**(2): 215–226.

Di-Ruggiero, J., and A.M. Gounot. 1990. Microbial manganese reduction mediated by bacterial strains isolated from aquifer sediments. *Microb. Ecol.* **20**: 53–63.

Donaldson, S.G., C.M. Miller, and W.W. Miller. 1992. Remediation of gasoline-contaminated soil by passive volatilization. *J. Environ. Qual.* **21**: 94–102.

Edwards, E.A., and D. Grbić-Galić. 1992. Complete mineralization of benzene by aquifer microorganisms under strictly anaerobic conditions. *Appl. Environ. Microbiol.* **58**: 2663–2666.

Edwards, E.A., and D. Grbić-Galić. 1994. Anaerobic degradation of toluene and o-xylene by a methanogenic consortium. *Appl. Environ. Microbiol.* **60**: 313–322.

Ehrlich, H.L. 1987. Manganese oxide reduction as a form of anaerobic respiration. *Geomicrobiol. J.* **5**(3): 423–431.

Ensley, B.D. 1991. Biochemical diversity of trichloroethylene metabolism. *Ann. Rev. Microbiol.* **45**: 283–299.

Fang, J., M.J. Barcelona, and C. West. 1997. The use of aromatic acids and phospholipid ester-linked fatty acids for delineation of processes affecting an aquifer contaminated with JP-4 fuel. pp. 65–76. In: R. Eganhouse (ed.), *Molecular Markers in Environmental Geochemistry*. ACS Symposium Series 671. Amer. Chem. Soc.

Fishman, M.J., and L.C. Friedman. 1985. Methods for determination of inorganic substances in water and fluvial sediments. p. 545. In: M.J. Fishman and L.C. Friedman (eds.), *Techniques* *of Water-Resources Investigations*. U.S. Geological Survey, book 5, chap. A1.

Fogel, M.M., A.R. Taddeo, and S. Fogel. 1986. Biodegradation of chlorinated ethenes by a methane-utilizing mixed culture. *Appl. Environ. Microbiol.* **51**: 720–724.

Freedman, D.L., and J.M. Gossett. 1989. Biological reductive dechlorination of tetrachloroethylene and trichloroethylene to ethylene under methanogenic conditions. *Appl. Environ. Microbiol.* **55**: 2144–2151.

Gantzer, C.J., and L.P. Wackett. 1991. Reductive dechlorination catalyzed by bacterial transition-metal coenzymes. *Environ. Sci. Technol.* **25**: 715–722.

Gibbons, R.D. 1994. *Statistical Methods for Groundwater Monitoring*. John Wiley & Sons, New York, NY, 286 pp.

Gilbert, R.O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold Co., New York, NY, 320 pp.

Hanson, R.S., and T.E. Hanson. 1996. Methanotrophic bacteria. *Microbiol. Rev.* **60**(2): 439–471.

Hardin, J.W., and R.O. Gilbert. 1993. *Comparing statistical tests for detecting soil contaminations greater than background*. PNL-8989. Pacific Northwest Laboratory, Richland, WA.

Henson, J.M., M.V. Yates, J.W. Cochran, and D.L. Shackleford. 1988. Microbial removal of halogenated methanes, ethanes, and ethylenes in an aerobic soil exposed to methane. *FEMS Microbiol. Ecol.* **53**: 193–201.

Heron, G., C. Crouzet, Alain C.M. Bourg, and T.H. Christensen. 1994. Speciation of Fe(II) and Fe(III) in contaminated aquifer sediments using chemical extraction techniques. *Environ. Sci. Technol.* **28**: 1698–1705.

Higgo, J.J., P.H. Nielsen, and M.P. Bannon. 1996. Effect of geochemical conditions on fate of organic compounds in groundwater. *Environ. Geol.* **27**: 335–346.

Hopkins, G.D., L. Semprini, and P.L. McCarty. 1993. Microcosms and in-situ field studies of enhanced biotransformation of trichloroethylene by phenol-oxidizing microorganisms. *Appl. Environ. Microbiol.* **59**: 2277–2285.

Howard, P.H., R.S. Boethling, W.F. Jarvis, W.M. Maylan, and E.M. Michalenko. 1991. *Handbook of Environmental Degradation Rates*. Lewis Publishers, Chelsea, MI.

Huesemann, M.H., and M.J. Truex. 1996. The role of oxygen diffusion in passive bioremediation of petroleum contaminated soils. *J. Hazard. Mater.* **51**: 93–113.

Hunt, J.R., N. Sitar, and K.S. Udell. 1988. Nongaseous phase liquid transport and cleanup, 1. Analysis of mechanisms. *Water Resour. Res.* **24**(8): 1247–1258.

Jeffers, P.M., L.M. Ward, L.M. Woytowitch, and N.L. Wolfe. 1989. Homogeneous hydrolysis rate constants for selected chlorinated methanes, ethanes, ethenes, and propanes. *Environ. Sci. Technol.* **23**(8): 965–969.

Kampbell, D.H., T.H. Wiedemeier, and J.E. Hansen. 1995. Intrinsic bioremediation of fuel contamination in ground water at a field site. *J. Hazard. Mater.* **49**: 197–204.

Kemblowski, M.W., and C.Y. Chiang. 1990. Hydrocarbon thickness fluctuations in monitoring wells. *Ground Water* **28**(2): 244–252.

Kemblowski, M.W., J.P. Salanitro, G.M. Deeley, and C.C. Stanley. 1987. *Fate and transport of dissolved hydrocarbons in groundwater - a case study*. pp. 207–231. In: *Proceedings of Petroleum Hydrocarbons and Organic Chemicals in Ground Water: Prevention, Detection, and Restoration Conference*. Houston, TX, 17–19 November. NGWA: Dublin, OH.

Kennedy, L.G., and S.R. Hutchins. 1992. Applied geologic, microbiologic, and engineering constraints of in-situ BTEX bioremediation. *Remediation* **3**(1): 83–107.

Kleopfer, R.D., D.M. Easley, B.B. Hass, Jr., T.G. Deihl, D.E. Jackson, and C.J. Wurrey. 1985. Anaerobic degradation of trichloroethylene in soil. *Environ. Sci. Technol.* **19**: 277–280.

Klusman, R.W. 1980. *Sampling Designs for Biochemical Baseline Studies in the Colorado Oil Shale Region: A Manual for Practical Application*. U.S. Department of Energy. Report DOE/EV/10298-2. 180 pp.

Koterba, M., F. Wilde, and W. Lapham. 1996. *Ground-Water Data-Collection Protocols and Procedures for the National Water-Quality Assessment Program*. U.S. Geological Survey. Open-File Report 95-399. 113 pp.

Kresec, N., 1997. *Quantitative Solutions in Hydrogeology and Groundwater Modeling*. Lewis Publishers, CRC Press, Inc., Boca Raton, FL.

Lee, M.D. 1988. Biorestoration of aquifers contaminated with organic compounds. *CRC Crit. Rev. Environ. Control* **18**: 29–89.

Lenhard, R.J., and J.C. Parker. 1990. Estimation of free hydrocarbon volume from fluid levels in monitoring wells. *Ground Water* **28**(1): 57–67.

Lesage, S., R.E. Jackson, M.W. Priddle, and P.G. Riemann. 1990. Occurrence and fate of organic solvent residences in anoxic groundwater at the Gloucester landfill. *Canada Environ. Sci. Technol.* **24**(4): 559–566.

Lindberg, R.D., and D.D. Runnells. 1984. Ground water redox reactions: An analysis of equilibrium state applied to En measurements and geochemical modeling. *Science* **225**: 925–927.

Lovely, D.R. 1991. Dissimilatory Fe(III) and Mn(IV) reduction. *Microbiol. Rev.* **55**(2): 259–287.

Lovely, D.R., and D.J. Lonergan. 1990. Anaerobic oxidation of toluene, phenol, and *p*-creosol by the dissimilatory iron-reducing organism GS-15. *Appl. Environ. Microbiol.* **56**: 1858–1864.

Lovely, D.R., and S. Goodwin. 1988. Hydrogen concentrations as an indicator of the predominant terminal electron-accepting reaction in aquatic sediments. *Geochim. Cosmochim. Acta.* **52**: 2993–3003.

Lovely, D.R., J.C. Woodward, and F.H. Chapelle. 1994. Stimulated anoxic biodegradation of aromatic hydrocarbons using Fe(III) ligands. *Nature (London)* **370**: 128–131.

Lovely, D.R., M.J. Baedecker, D.J. Lonergan, I.M. Cozzarelli, E.J.P. Phillips, and D.I. Siegel. 1989. Oxidation of aromatic contaminants coupled to microbial iron reduction. *Nature (London)* **339**: 297–299.

Martin-Hayden, J.M., and G.A. Robbins. 1997. Plume distortion and apparent attenuation due to concentration averaging in monitoring wells. *Ground Water* **35**(2): 339–346.

Martin-Hayden, J.M., G.A. Robbins, and R.D. Bristol. 1991. Mass balance evaluation of monitoring well purging: Part II. Field tests at a gasoline contamination site. *J. Contam. Hydrol.* **8**: 225–241.

Mayer, K.R., D. Grbić-Galić, L. Semprini, and P.L. McCarty. 1988. Degradation of trichloroethylene by methanotrophic bacteria in a laboratory column of saturated aquifer material. *Water Sci. Tech.* **20**(11/12): 175–178.

McAllister, P.M., and C.Y. Chiang. 1994. A practical approach to evaluating natural attenuation of contaminants in ground water. *Ground Water Monit. Rem.* **14**(2): 161–173.

McCarty, P.L. 1988. Bioengineering issues related to in-situ remediation of contaminated soils and groundwater. pp. 143–162. In: G.S. Omenn (ed.), *Biotic and Abiotic Transformation of*

Chlorinated Solvents in Ground Water. Environmental Biotechnology, Plenum, New York, NY.

McCarty, P.L. 1996. *Biotic and abiotic transformations of chlorinated solvents in ground water*. pp. 5–9. In: *Proceedings of the Symposium on Natural Attenuation of Chlorinated Organics in Ground Water*. EPA/540/R-96/509. Office of Research and Development, Washington, DC.

McCarty, P.L., and L. Semprini. 1994. Ground-water treatment for chlorinated solvents. p. 87. In: R.D. Norris, R.E. Hinchee, R. Brown, P.L. McCarty, L. Semprini, J.T. Wilson, D.H. Kampbell, M. Reinhard, E.J. Bouwer, R.C. Borden, T.M. Vogel, J.M. Thomas, and C.H. Ward (eds.), *Handbook of Bioremediation*. Lewis Publishers, Boca Raton, FL.

McDonald, L.L., and W.P. Erickson. 1994. Testing for the bioequivalence in field studies: Has a distributed site been adequately reclaimed? pp. 183–198. In: D.J. Fletcher and B.F.J. Manly (eds.), *Statistics in Ecology and Environmental Monitoring*. Otago Conference Series No. 2, University of Otago Press, Dunedin, New Zealand.

McNabb, W.W. Jr., and B.P. Dooher. 1998. A critique of a steady-state method for estimating contaminant degradation rates. *Ground Water* **36**(6): 983–987.

Mercer, J.W. 1998. Transport. and remediation of chemicals in groundwater. In: R.A. Myers (ed.), *Encyclopedia of Environmental Analysis and Remediation*. John Wiley & Sons, New York, NY.

Molz, F., and G. Boman. 1997. *Site characterization tools: Using a borehole flowmeter to locate and characterize the transmissive zones of an aquifer*. In: *Proceedings of the Symposium on Natural Attenuation of Chlorinated Organics in Ground Water*. EPA/540/R-97/504. ORD, Washington, DC.

Molz, F.J., G.K. Woman, S.C. Young, W.R. Waldrop. 1994. Borehole flowmeters: Field application and data analysis *J. Hydrol.* **163**: 347–371.

Morris, J.C., and W. Stumm. 1967. p. 270. In: W. Stumm, (ed.), *Equilibrium Concept in Natural Waters Systems*. Am. Chem. Soc., Washington, DC.

Mu, D.Y., and K.M. Scow. 1994. Effect of trichloroethylene (TCE) and toluene concentrations on TCE and toluene biodegradation and the population density of TCE and toluene degraders in soil. *Appl. Environ. Microbiol.* **60**(7): 2661–2665.

Newell, C.J., R.K. McLeod, J.R. Gonzales. 1996. *BIOSCREEN Natural Attenuation Decision Support System*. User's Manual. Version 1.3. EPA/600/R-96/087. ORD, Washington, DC.

Newell, C.J., A.P. Smith, C.E. Aziz, T.A. Khan, J.R. Gonzales, and P.E. Hass. 1998. BIOCHLOR: A planning-level natural attenuation model and database for solvent sites. pp. 237–242. In: G.B. Wickramanayake and R.E. Hinchee (eds.), *Natural Attenuation of Chlorinated and Recalcitrant Compounds*. Battelle, Columbus, OH.

Norris, R.D., R.E. Hinchee, R. Brown, P.L. McCarty, L. Semprini, J.T. Wilson, D.H. Kampbell, M. Reinhard, E.J. Bouwer, R.C. Borden, T.M. Vogel, J.M. Thomas, and C.H. Ward. 1993. *In-Situ Bioremediation of Ground Water and Geological Material: A Review of Technologies*. EPA/600/R-93/124.

Nyer, E.K., and M.E. Duffin. 1997. The state of the art bioremediation. *Ground Water Monit. Rem.* **2**: 64–69.

O'Brien, R.F. 1997. Statistical issues when comparing a site to background. pp. 367–372. In: R. Hinchee (ed.), *In-Situ and On-Site Bioremediation*. Battelle Press, Columbus, OH.

O'Brien, R.F., B.K. Sinha, and W.P. Smith. 1991. A statistical procedure to evaluate cleanup standards. *J. Chemometrics* **5**(3): 249–261.

Parsons, F.Z., P.R. Wood, and J. DeMarco. 1984. Transformations of tetrachloroethylene and trichloroethylene in microcosms and groundwater. *J. AWWA.* **71**: 56–59.

Patel, R.N., C.T. Hou, A.I. Laskin, and A. Felix. 1982. Microbial oxidation of hydrocarbons: Properties of a soluble methane monooxygenase from a facultative methane-utilizing organism, *Methylobacterium* sp. strain CRL-26. *Appl. Environ. Microbiol.* **44**: 1130–1137.

Pettyjohn, W.A. 1982. Cause and effect of cyclic changes in ground-water quality. *Ground Water Monit. Rem.* **2**(1): 43–49.

Puls, R.W., and M.J. Barcelona. 1996. *Low-Flow (Minimal Drawdown) Sampling Procedures*. EPA/540/S-95/504. ORD, Washington, DC.

Rabus, R., R. Nordhaus, W. Ludwig, and F. Widdel. 1993. Complete oxidation of toluene under strictly anoxic conditions by a new sulfate-reducing bacterium. *Appl. Environ. Microbiol.* **59**: 1444–1451.

Reinhard, M., and N.L. Goodman. 1984. Occurrence and distribution of organic chemicals in two landfill leachate plumes. *Environ. Sci. Technol.* **18**: 953–961.

Rifai, H.S. 1997. Natural Aerobic Biological Attenuation. In: C.H. Ward, J.A. Cherry, and M.R. Scalf (eds.), *Subsurface Remediation*. Ann Arbor Press, Inc., Chelsea, MI.

Rifai, H.S., C.J. Newell, R. Miller, S. Taffinder, and M. Roundsaville. 1995. Simulation of natural attenuation with multiple electron acceptors. *Bioremediation* **3**: 53–58.

Rifai, H.S., P.B. Bedient, and J.T. Wilson. 1989. *BIOPLUME model for contaminant transport affected by oxygen limited biodegradation*. EPA/600/M-89/019.

Robbins, G.A. 1989. Influence of using purged and partially penetrating monitoring wells on contaminant detection, mapping and modeling. *Ground Water* **27**(2): 155–162.

Robbins, G.A., and J.M. Martin-Hayden. 1991. Mass balance evaluation of monitoring well purging: Part I. Theoretical models and implications for representative sampling. *J. Contam. Hydrol.* **8**: 203–224.

Salanitro, J.P. 1992. *Criteria for evaluating the bioremediation of aromatic hydrocarbons in aquifers*. Presented at the National Research Council (Water Science Technology Board) Committee on In-Situ Bioremediation: How Do We Know When It Works? Washington, DC, 26–29 October.

Salanitro, J.P. 1993. The role of bioattenuation in the management of aromatic hydrocarbon plumes in aquifers. *Ground Water Monit. Rem.* **13**(4): 150–161.

Schmidt, K.D. 1977. Water quality variations for pumping wells. *Ground Water* **15**(2): 130–137.

Schocher, R.J., B. Seyfried, F. Vazquez, and J. Zeyer. 1991. Anaerobic degradation of toluene by pure cultures or denitrifying bacteria. *Arch. Microbiol.* **157**: 7–12.

Semprini, L., P.K. Kitanidis, D.H. Kampbell, and J.T. Wilson. 1995. Anaerobic transformation of chlorinated aliphatic hydrocarbons in a sand aquifer based on spatial chemical distributions. *Water Resour. Res.* **13**(40): 1051–1062.

Sewell, G.W., and S.A. Gibson. 1991. Simulation of the reductive dechlorination of tetrachloroethylene in anaerobic aquifer microcosms by the addition of toluene. *Environ. Sci. Technol.* **25**: 982–984.

Shampine, W.J., L.M. Pope, and M.T. Koterba. 1992. *Integrating Quality Assurance in Project Work Plans of the U.S. Geological Survey*. U.S. Geological Survey. Open-File Report 92–162. 12 pp.

Sharples, B.P.M. 1992. Oil pollution by the offshore industry contrasted with tankers: An examination of the facts. *J. Process Mech. Eng.* **206**: 3–14.

Shelton, L.R. 1994. *Field Guide for Collection and Processing Stream-Water Samples for the National Water-Quality Assessment Program*. U.S. Geological Survey. Open-File Report 94-455, Sacramento, CA. 44 pp.

Shelton, L.R., and P.D. Capel. 1994. *Guidelines for Collecting and Processing Samples of Stream Bed Sediment for Analysis of Trace Elements and Organic Contaminants for the National Water-Quality Assessment Program*. U.S. Geological Survey. Open-File Report 94–458. 20 pp.

Stone, A.T., and J.J. Morgan. 1984. Reduction and dissolution of manganese (III) and manganese (IV) oxides by organics. 1. Reaction with hydroquinone. *Environ. Sci. Technol.* **18**(6): 450–456.

Stumm, W., and J.J. Morgan. 1981. *Aquatic Chemistry*, 2nd ed. John Wiley & Sons, New York, NY.

Tiedje, J.M., and T.O. Stevens. 1988. The ecology of anaerobic dechlorination consortium. pp. 3–14. In: G.S. Omen (ed.) *Environmental Biotechnology: Reducing Risks from Environmental Chemicals through Biotechnology*. Plenum Press, New York.

U.S. EPA. 1996. *Draft Field Methods Compendium (FMC)*. OERP #9285.2-11. Office of Emergency and Remedial Response, Washington, DC.

U.S. EPA. 1997a. *Field Analytical and Site Characterization Technologies, Summary of Applications*. EPA/542/R-97/011. OSWER, Washington, DC.

U.S. EPA. 1997b. *Draft EPA Region 4 Suggested Practices for Evaluation of a Site for Natural Attenuation (Biological Degradation) of Chlorinated Solvents*.

U.S. EPA. 1998a. *Technical Protocol for Evaluating Natural Attenuation of Chlorinated Solvents in Ground Water*. EPA/600/R-98/128. ORD, Washington, DC.

U.S. EPA. 1998b. *Field Sampling and Analysis Technologies Matrix and Reference Guide*. EPA/542/B-98/002.

U.S. EPA. 1999. *Use of Monitored Natural Attenuation at Superfund, RCRA Corrective Action, and Underground Storage Tank Sites*. OSWER Directive, 9200.4-17P. Office of Solid Waste and Emergency Response, Washington, DC.

Underground Storage Tank Technology Update. 1998. *Four Critical Considerations in Assessing Contaminated Plumes. Implications for Site-Specific MNA/RBCA Assessments at Petroleum-Contaminated Plumes*. EPA/600/J-98/228.

Vogel, T.M. 1988. *Biotic and abiotic transformations of halogenated aliphatic compounds*. Ph.D. Thesis. Stanford University. Stanford, CA.

Vogel, T.M., and D. Grbić-Galić. 1986. Incorporation of oxygen from water into toluene and benzene during anaerobic fermentative transformation. *Appl. Environ. Microbiol.* **52**: 200–202.

Vogel, T.M., and P.L. McCarty. 1985. Biotransformation of tetrachloroethylene to trichloroethylene, dichloroethylene, vinyl chloride and carbon dioxide under methanogenic conditions. *Appl. Environ. Microbiol.* **49**: 1080–1083.

Vogel, T.M., C. Criddle, and P. McCarty. 1987. Transformation of halogenated aliphatic compounds. *Environ. Sci. Technol.* **21**: 722–736.

Vogel, T.M., and P.L. McCarty. 1987a. Rate of abiotic formation of 1,1-dichloroethylene from 1,1,1-trichloroethane in groundwater. *J. Cont. Hydrol.* **1**: 299–308.

Vogel, T.M., and P.L. McCarty. 1987b. Abiotic and biotic transformation of 1,1,1-trichloroethane under methanogenic conditions. *Environ. Sci. Technol.* **21**: 1208–1213.

Weaver, J.W., J.T. Wilson, and D.H. Kampbell. 1996. *Extraction of degradation rate constants from the St. Joseph, Michigan Trichloroethene site*. p. 69. In: *Symposium on Natural Attenuation of Chlorinated Organics in Ground Water*. September 11–13, 1996, Dallas, TX. EPA/540/R-96/509.

Westrick, J.J., J.W. Mellow, and R.F. Thomas. 1984. The groundwater supply survey. *J. AWWA.* **76**: 52–59.

Wiedemeier, T.H., J.T. Wilson, D.H. Kampbell, R.N. Miller, and J.E. Hansen. 1995. *Technical Protocol for Implementing Intrinsic Remediation with Long-Term Monitoring for Natural Attenuation of Fuel Contamination Dissolved in Groundwater*. United States Air Force Center for Environmental Excellence, Technology Transfer Division, Brooks Air Force Base, San Antonio, TX.

Wiedemeier, T.H., M.A. Swanson, J.T. Wilson, D.H. Kampbell, R.N. Miller, and J.E. Hansen. 1996. Approximation of biodegradation rate constants for monoaromatic hydrocarbons (BTEX) in groundwater. *Ground Water Monit. Rem.* **16**(3): 186–194.

Wilson, J.T., J.F. McNabb, B.H. Wilson, and M.J. Noonan. 1983. Biotransformation of selected organic pollutants in groundwater. *Dev. Ind. Microbiol.* **24**: 225–233.

Wilson J.T., and B.H. Wilson. 1985. Biotransformation of trichloroethylene in soil. *Appl. Environ. Microbiol.* **49**: 242–243.

Wood, W.W. 1981. *Guidelines for Collection and Field Analysis of Ground-Water Samples for Selected Unstable Constituents*. U.S. Geological Survey Techniques of Water-Resources Investigations, Book 1, Chap. D2. 24 pp.

Zeeb, P.J., N.D. Durant, R.M. Cohen. and J.W. Mercer. 1999. *Site Investigation Approaches to Support Assessment of Natural Attenuation of Chlorinated Solvents in Groundwater: Part 1, Sandy Aquifer*. Draft of Site Characterization Manual. U.S. EPA. ORD, Washington, DC.

GROUNDWATER VULNERABILITY TO PESTICIDES: STATISTICAL APPROACHES

Jason J. Gurdak
U.S. Geological Survey
Lakewood, Colorado
and
Colorado School of Mines
Golden, Colorado

John E. McCray
Colorado School of Mines
Golden, Colorado

Groundwater vulnerability to pesticides is the potential for contamination of groundwater by pesticides from a nonpoint source or aerially distributed point sources. Due to widespread pesticide use and detection in groundwater, assessments that evaluate and estimate pesticide vulnerability have increasingly proven valuable during decisions regarding groundwater management, conservation, and remediation because of the significant costs of such actions. To predict groundwater vulnerability and understand factors that affect the transport of pesticides to groundwater resources, a variety of approaches have been developed that are broadly classified into three general categories: (1) overlay and index, (2) statistically based, and (3) process based. Often the selection of an appropriate approach for assessing groundwater vulnerability to pesticides is determined by the type and scale of the problem and by the available resources. This article provides an overview of the approaches and methods of evaluation for assessing groundwater vulnerability to pesticides.

BACKGROUND

Pesticide vulnerability assessments have become increasingly important in managing groundwater resources because of widespread pollution of groundwater by various pesticides. Pesticides are synthetic organic chemicals used to destroy insects or other organisms harmful to agriculture or animals. A comprehensive review by Barbash and Resek (1) of the published information on the distribution and behavior of pesticides in groundwater suggests that the greatest potential for unintended adverse effects from pesticides is through contamination of the earth's hydrologic system. Pesticides from every chemical class have been detected in groundwater (1), including samples collected from major aquifers of the United States, as part of the U.S. Geological Survey's National Water-Quality Assessment (NAWQA) Program. The samples collected from 1992 to 1998 by NAWQA revealed that pesticides were found most frequently in shallow groundwater in urban (54% of wells) and agricultural areas (61%), where pesticide use is greatest (2).

Knowledge of pesticide vulnerability for specific groundwater resources can be of great value for preserving and sustaining water quality. A variety of approaches have been developed to predict groundwater vulnerability and understand factors that affect the transport of pesticides to groundwater resources. Many of these approaches have been used by water-resource managers to make informed decisions regarding the potential risk of groundwater contamination by pesticides. This article provides an overview of the approaches and methods of evaluation for assessing groundwater vulnerability to pesticides.

Vulnerability is the potential for [pesticide] contamination of groundwater from a nonpoint source or aerially distributed point sources. The National Research Council (3) defined groundwater vulnerability to contamination as "the tendency or likelihood for contamination to reach a specified position in the groundwater system after introduction at some location above the uppermost aquifer." Similarly, the U.S. Environmental Protection Agency (4) defined groundwater vulnerability as "the relative ease with which a contaminant applied on or near the land surface can migrate to the aquifer of interest under a given set of agronomic management practices, pesticide characteristics, and hydrogeologic sensitivity conditions." Many definitions conceptualize two distinct components of groundwater vulnerability (5): (1) *intrinsic susceptibility*, which is defined by the inherent properties of the hydrogeologic system (hydraulic conductivity, porosity, depth to water, and soil type) and the associated sources of water and stresses for the system (rates and sources of recharge and discharge); and (2) *specific vulnerability*, or simply *vulnerability*, which is defined by the proximity of contaminant sources, fate and transport properties of the contaminant, and other factors that could increase loads of specified contaminants to the aquifer, their eventual delivery to a groundwater resource, or both. As Rupert (6) noted, although multiple definitions have been used for the term "vulnerability," they all address the same underlying question: What is the potential for groundwater contamination?

To assess the vulnerability of a groundwater resource to pesticide contamination, the physical and chemical mechanisms that control pesticide movement into groundwater must be addressed. A complete discussion of all processes and factors responsible for the behavior and fate of pesticides from land surface to groundwater is beyond the scope of this article. For additional details on pesticide fate and transport, refer to Barbash and Resek (1), a literature review of pesticide transport and leaching to groundwater by Flury (7); and Hornsby et al. (8), which is a comprehensive database that lists the properties of pesticides in the environment.

APPROACHES TO GROUNDWATER VULNERABILITY ASSESSMENTS OF PESTICIDES

A variety of approaches are presently used for assessing groundwater vulnerability to pesticides, ranging from subjective, qualitative, and opinion-based to objective, quantitative, and scientifically defensible methods. The approaches for assessing groundwater vulnerability can be classified into three general categories: (1) overlay and index, (2) statistically based, and (3) process based. Note that selection of the appropriate approach for assessing groundwater vulnerability is primarily driven by the type and scale of the problem and by the available resources.

Overlay and Index Approach

The overlay and index, one of the earliest approaches, generates vulnerability indexes based on the aggregation, or overlay, of many variables or factors deemed important in a given hydrologic setting, such as well location, geology, soil, and depth to water. This approach often is used as an initial screening tool and has an advantage over other methods because the data needed for implementation are generally available on the appropriate scale of study. The most widely used overlay and index method is DRASTIC (9), which incorporates professional judgment to weigh factors and create an aggregate point rating of groundwater vulnerability. DRASTIC is defined by seven factors: depth to water, net recharge, aquifer media, soil media, topography, impact of vadose zone media, and hydraulic conductivity of the aquifer. DRASTIC and other approaches that are based on best professional judgment are subjective and thus inherently result in less scientifically defensible groundwater vulnerability assessments. Overlay and index methods are not usually calibrated to measured pesticide concentrations in the groundwater and can result in poor correlation with actual pesticide concentration measurements (10). In addition, the seven factors considered by DRASTIC may not be the most appropriate for all groundwater systems.

Statistical Approach

Statistical approaches rely on empirical observation of measured pesticide concentration or detection in groundwater to establish relations with independent, or explanatory, variables. Explanatory variables represent the physical and chemical system of the land surface and subsurface, such as land use, soil taxonomy, lithology, depth to water, and pesticide application rates. The results of statistical models often express vulnerability as the probability of detecting pesticide concentrations within the groundwater of a specific area from which the empirical data were drawn. Although statistical models do not explicitly define underlying processes of pesticide fate and transport, important processes can be inferred based on proper selection of explanatory variables. An important assumption of statistical vulnerability models is that the area to which they are applied must be comparable to that in which they were developed (3). To address this assumption, statistical models rely on groundwater monitoring networks that are spatially representative of the area of interest.

The following sections outline two common statistical approaches used to assess groundwater vulnerability to pesticides: logistic regression and posterior probability distribution. Other statistically based approaches include cluster analysis (11), principal component analysis (11), and linear regression.

Logistic Regression. Logistic regression is a multivariate statistical approach that has been used extensively in the health sciences to predict a binary response in the dependent variable from measured explanatory variables, but recently, it has been used to estimate the probability of detecting pesticide concentrations in groundwater resources (6,12–17). Logistic regression is conceptually similar to multiple linear regressions because the relation between one dependent (response) variable and several independent (explanatory) variables is evaluated (18). However, logistic regression has an advantage over other generalized linear regression models because is does not require a normal sample distribution and censored (nondetect) values can be used. The concentrations of pesticides in groundwater generally are low; therefore, laboratory detection levels serve as a convenient threshold for establishing the binary response (i.e., detection of pesticides in groundwater samples versus nondetection of pesticides). The underlying assumption of logistic regression is that the natural logarithm of the odds ratio, which is the probability of being in a response category, is linearly related to the explanatory variables. The odds ratio is defined by the probability of exceeding the established threshold, expressed as

$$\text{Odds ratio} = \frac{p}{1-p} \tag{1}$$

where p is the probability of exceeding the threshold value.

The log of the odds ratio, or logit, transforms a variable constrained between 0 and 1 into a continuous variable that is a linear function of the explanatory variables (19). The logit transformation is

$$\log\left(\frac{p}{1-p}\right) = b_o + b_x \tag{2}$$

where b_o is the logistic regression constant and b_x is the vector of slope coefficients and explanatory variables. Predicted values of the response variable are converted back into probability units by using the logistic transformation; the logistic regression model takes the form

$$P = \frac{e^{(b_o+bX)}}{1+e^{(b_o+bX)}} \tag{3}$$

where P is the probability of the binary response event, often defined as detecting the pesticide of interest in groundwater, and e is the base of the natural logarithm.

An additional advantage of logistic regression is the user's ability for model validation using an independent set of pesticide concentration data. However, logistic regression, like other statistical approaches, has received criticism because the models do not represent the physical, chemical, and biological processes that affect groundwater vulnerability and process-based approaches. A detailed description of logistic regression analysis is presented by Hosmer and Lemeshow (18).

Posterior Probability Distribution. Using the posterior probability distribution method, Worrall and Kolpin (20) developed an approach for the direct assessment of groundwater vulnerability to pesticides based on a single observation of multiple contaminants. This method of direct assessment is independent of the compounds examined, and it has a direct probabilistic interpretation making it ideal for risk assessment. In addition, this method has the advantage that the measures of vulnerability are based on observation rather than expert opinion, indexes, or physical models (20). Furthermore, errors of the probabilistic interpretation can be easily quantified.

The posterior probability distribution method is a measure of the difference between two proportions (the proportion of detections in the region and the proportions of detections observed in a single observation well) (20). Because vulnerability is always calculated relative to the region as a whole, compounds that are either always detected or never detected add no discriminating information to the assessment. The proportion of detections of each compound is calculated for the entire data set, and the probability distribution for each of the proportions is determined based on the observed data. The posterior probability distribution can help identify differences between the proportions of detections in the well compared to that in the region and provides estimates of vulnerability of the area of the aquifer contributing to the well. Worrall and Kolpin (20) provide in-depth details regarding the posterior probability distribution methodology.

Although the posterior probability distribution method is advantageous because it can be used to determine a probabilistic measure of vulnerability based on empirical observation rather than expert opinion, this method relies on a number of significant assumptions that may not be reasonable for all groundwater systems (20).

Process-Based Approach

Process-based pesticide vulnerability models include a wide range of methods and depend primarily on the extent to which they simulate subsurface solute-transport processes (21). These deterministic models require analytical or numerical solutions of mathematical equations that represent processes governing pesticide fate and transport. Although process-based models incorporate a more complete description of the physical, chemical, and biological processes that affect groundwater vulnerability, they may not necessarily provide more reliable estimates of pesticide vulnerability because of extensive data requirements needed to define the subsurface hydrogeology (3). Often, detailed data sets of the subsurface are not available or may result in predictions that are of little value if the parameters describing the system are ill-defined or if they vary significantly in time and space (22).

Process-based pesticide vulnerability assessments include complex and simple models. Complex pesticide vulnerability models can range from analytical solutions for one-dimensional transport through the unsaturated zone to coupled, unsaturated–saturated, two- or three-dimensional models (3). The Pesticide Root Zone Model (PRZM) (23) is one of the most commonly used complex models for pesticide vulnerability modeling. PRZM is a one-dimensional, deterministic-empirical/conceptual model that simulates soil–water movement by using an empirical drainage algorithm and solute transport with the advection-dispersion equation. PRZM has two distinct advantages over simple index models for estimating pesticide leaching: (1) concentrations are estimated with reference to depth and time, and (2) heterogeneity in the soil column and variability in recharge can be accounted for (24). Other complex process-based models include the Type Transfer Function (TTF), outlined by Stewart and Loague (25); the Leaching Estimation and Chemistry Model (LEACHM) (26); and Monte-Carlo simulations of deterministic models (27).

Advantages gained by using complex pesticide vulnerability models can be minimized by extensive data requirements, which have limited their use to local or field-scale evaluations. If data are unavailable, estimates are used that may introduce additional model uncertainty. Therefore, uncertainty can be reduced when using a simple method (28). Simple index-based process models have been used more routinely for large-scale and regional pesticide vulnerability modeling. Process-based index models are first-order approximations of pesticide vulnerability that describe the relation between how fast the pesticide leaches and how fast it can be degraded. The net result of these two processes determines what fraction of soil-applied pesticides potentially reaches groundwater (29).

Index models are conceptually simplified and are valid only under steady-state flow within a uniform homogeneous ideal medium. The models represent the important transport processes for pesticides, are most appropriately used as screening tools, and are intended for evaluation and comparison of pesticide behavior under constrained and limited conditions with the advantage of requiring few input data (21). The following sections describe index process models that have routinely been used to evaluate pesticide vulnerability of groundwater.

Attenuation Factor. The attenuation factor (AF) (30) is a simple index of residence time and degradation rate of a pesticide that is equivalent to the fraction of applied pesticide mass likely to leach past a reference depth. This index is a simple solution of the one-dimensional, advection-dominated transport equation and is defined as

$$\mathrm{AF} = e\left(\frac{-0.693\ d\mathrm{RF}\theta_{\mathrm{FC}}}{\mathrm{q}t_{1/2}}\right) \tag{4}$$

where d is the distance from the land surface to groundwater; RF is the retardation factor, which accounts for pesticide sorption effects; θ_{FC} is the volumetric water content at field capacity; q is the net groundwater recharge rate; and $t_{1/2}$ is the half-life of pesticide degradation. RF is defined as

$$\mathrm{RF} = \left(1 + \frac{\rho_{\mathrm{b}} f_{\mathrm{oc}} K_{\mathrm{oc}}}{\theta_{\mathrm{FC}}} + \frac{\theta_{\mathrm{g}} K_{\mathrm{H}}}{\theta_{\mathrm{FC}}}\right) \tag{5}$$

where ρ_{b} is the soil bulk density (g/cm^3), f_{oc} is the fraction of soil organic carbon on a mass basis, K_{oc} is the organic carbon sorption coefficient (mL/g), θ_{g} is the gas constant (unitless), and K_{H} is Henry's constant for the pesticide of interest (unitless).

The AF ranges from zero to one, where zero corresponds to a minimal threat of applied pesticide leaching and a value of one corresponds to all the applied pesticide mass reaching the groundwater.

Leaching Potential Index. The Leaching Potential Index (LPI) was developed by Meeks and Dean (31) and, similar to the AF, is derived from the advection-dispersion equation. The LPI is compound specific and includes information regarding adsorption and degradation parameters. The LPI is defined as

$$\mathrm{LPI} = \frac{1000\ t_{1/2} V}{0.693\ RFd} \tag{6}$$

where V is the average linear vertical soil–water velocity (L/t).

Increases in LPI values reflect a direct relation between increases in pesticide vulnerability. It should be noted that the LPI is simply the inverse of the exponent in Equation 5 of the AF, and a factor of 1000 is arbitrarily included to increase the numerical values of the LPI (32).

Schlosser and McCray (28) conducted a sensitivity analysis by using the LPI to evaluate the relative importance of input-parameter variability on model-predicted vulnerability. Results indicated that the calculated vulnerabilities were most sensitive to organic carbon content, depth to groundwater, and vertical soil–water velocity.

Vulnerability Index. The Vulnerability Index (VI) (32) is a modified version of the LPI equation, defined as

$$\mathrm{VI} = \frac{200\ k\theta_{\mathrm{FC}}}{d\rho_{\mathrm{b}}(\mathrm{OM})}\left(\frac{t_{1/2}}{K_{\mathrm{oc}}}\right) F_{\mathrm{DGW}} \tag{7}$$

where k is the decay constant (day^{-1}); OM is the percentage of organic matter present in the soil; $(t_{1/2}/K_{\mathrm{oc}})$

is the leachability ratio, which is the measure of the pesticide's propensity to biodegrade ($t_{1/2}$) and to sorb to organic matter in the soil (K_{oc}); and F_{DGW} is a factor that accounts for the depth to groundwater.

The VI model is a practical approach to assessing many pesticides simultaneously by incorporating the leaching potential. The VI uses the leachability ratio to account for the leaching potential of multiple pesticides that have similar chemical properties. This method has a significant advantage over other index methods; by grouping pesticides according to their leachability ratio, vulnerability assessments are made for each group, rather than for each individual pesticide.

PESTICIDE VULNERABILITY MAPS

Pesticide vulnerability maps are spatial representations of models (i.e., overlay and index, statistical, or process-based) that identify areas of the aquifer that have a potential for pesticide contamination in the groundwater. Often, pesticide vulnerability maps are used to predict pesticide vulnerability of groundwater that has not been previously sampled. Many pesticide vulnerability assessments have incorporated the use of geographic information systems (GIS) for calculating pesticide vulnerability and visualizing the model over the study area. GIS is a modeling framework and tool designed for the analysis and display of spatial data. Many vulnerability studies use GIS because (1) a GIS can display and analyze the spatial variability of georeferenced variables important for the assessment, (2) a GIS can incorporate georeferenced measurements of pesticide concentration data, and (3) the digital display and data-manipulation ability of a GIS provides numerous advantages for creating maps of groundwater vulnerability to pesticide contamination.

UNCERTAINTY OF PESTICIDE VULNERABILITY ASSESSMENT

Uncertainty is unavoidable during estimation of pesticide vulnerability because of limitations in knowledge of contaminant behavior in the subsurface and significant limitations in the spatial databases used for assessments (3). Uncertainty in predictions reflect limitations due to (1) accurate spatial delineations of data and its transfer from real-world to digital data or maps; (2) spatial averaging and uncertainties of soil properties and other explanatory variables; (3) assumptions related to individual pesticide vulnerability approaches that often assume steady-state conditions or homogeneous conditions in the soil or unsaturated zone; (4) uncertainties from spatial interpolation from point estimates, if kriging or other interpolation methods are used; and (5) temporal variability of explanatory variables, such as land-use practices, may limit predictions of future vulnerability estimates for a given groundwater resource.

Increasingly, studies of pesticide vulnerability have acknowledged the inherent uncertainty in the assessment and have made attempts to quantify this uncertainty. Refer to the work of Freissinet et al. (33), Loague et al. (24), and Loague (34) for specific approaches to estimate and quantify uncertainty during pesticide vulnerability assessments.

BIBLIOGRAPHY

1. Barbash, J.E. and Resek, E.A. (1996). *Pesticides in Ground Water, Distribution, Trends, and Governing Factors*. Ann Arbor Press, Chelsea, MI.
2. Gilliom, R.J. (2001). Pesticides in the hydrologic system—what do we know and what's next? *Hydrological Processes* **15**: 3197–3201.
3. National Research Council. (1993). *Ground-Water Vulnerability Assessment: Predicting Relative Contamination Potential under Conditions of Uncertainty*. National Academy Press, Washington, DC.
4. USEPA (1993). *A Review of Methods for Assessing Aquifer Sensitivity and Ground Water Vulnerability to Pesticide Contamination*. EPA 813/R-93-002, Washington, DC.
5. Focazio, M.J., Reilly, T.E., Rupert, M.G., and Helsel, D.R. (2002). *Assessing Ground-Water Vulnerability to Contamination: Providing Scientifically Defensible Information for Decision Makers*. U.S. Geological Survey Circular 1224.
6. Rupert, M.G. (2003). *Probability of Detecting Atrazine/Desethyl-atrazine and Elevated Concentrations of Nitrate in Ground Water in Colorado*. U.S. Geological Survey Water-Resources Investigations Report 02-4269.
7. Flury, M. (1996). Experimental evidence of transport of pesticides through field soils—a review. *J. Environ. Qual.* **25**: 25–45.
8. Hornsby, A.G., Wauchope, R.D., and Herner, A.E. (1996). *Pesticide Properties in the Environment*. Springer-Verlag, New York.
9. Aller, L., Truman, B., Lehr, J.H., and Petty, R.J. (1985). *DRASTIC—A Standardized System for Evaluating Ground Water Pollution Potential Using Hydrogeologic Settings*. U.S. Environmental Protection Agency, Robert S. Kerr Environmental Research Laboratory, Office of Research and Development.
10. Rupert, M.G. (1999). *Improvements to the DRASTIC Ground-Water Vulnerability Mapping Method*. U.S. Geological Survey Fact Sheet 066-99.
11. Troiano, J., Johnson, B.R., Powell, S., and Schoenig, S. (1994). Use of cluster and principal component analyses to profile areas in California where ground water has been contaminated by pesticides. *Environ. Monitoring Assessment* **32**: 269–288.
12. Worrall, F. and Kolpin, D.W. (2004). Aquifer vulnerability to pesticide pollution—combining soil, land-use and aquifer properties with molecular descriptors. *J. Hydrol.* **293**: 191–204.
13. Rupert, M.G. (1998). *Probability of Detecting Atrazine/Desethyl-atrazine and Elevated Concentrations of Nitrate* ($NO_2 + NO_3 - N$) *in Ground Waters in the Idaho Part of the Upper Snake River Basin*. U.S. Geological Survey Water-Resources Investigations Report 98-4203.
14. Nolan, B.T., Hitt, K.J., and Ruddy, B.C. (2002). Probability of nitrate contamination of recently recharged groundwaters in the conterminous United States. *Environ. Sci. Technol.* **36**: 2138–2145.
15. Nolan, B.T. and Stoner, J.D. (2000). Nutrients in groundwaters of the conterminous United States, 1992–1995. *Environ. Sci. Technol.* **34**: 1156–1165.

16. Teso, R.R., Poe, M.P., Younglove, T., and McCool, P.M. (1996). Use of logistic regression and GIS modeling to predict groundwater vulnerability to pesticides. *J. Environ. Qual.* **25**: 425–432.
17. Eckhardt, D.A.V. and Stackelberg, P.E. (1995). Relation of ground-water quality to land use on Long Island, New York. *Ground Water* **33**: 1019–1032.
18. Hosmer, D.W. and Lemeshow, S. (1989). *Applied Logistic Regression*. John Wiley & Sons, New York.
19. Helsel, D.R. and Hirsch, R.M. (1992). *Statistical Methods in Water Resources*. Elsevier Science, New York.
20. Worrall, F. and Kolpin, D.W. (2003). Direct assessment of groundwater vulnerability from single observations of multiple contaminants. *Water Resour. Res.* **39**: 1345–1352.
21. Shukla, S., Mostaghimi, S., Shanholtz, V.O., and Collins, M.C. (1998). A GIS-based modeling approach for evaluating groundwater vulnerability to pesticides. *J. Am. Water Resour. Assoc.* **34**: 1275–1293.
22. Nolan, B.T. (1998). *Modeling Approaches for Assessing the Risk of Nonpoint-Source Contamination of Ground Water*. U.S. Geological Survey Open-file report 98-531.
23. Carsel, R.F., Smith, C.N., Mulkey, L.A., Dean, J.D., and Howise, P. (1984). *User's Manual for the Pesticide Root Zone Model (PRZM), Release 1*. USEPA, Washington, DC.
24. Loague, K., Bernknopf, R.L., Green, R.E., and Giambelluca, T.W. (1996). Uncertainty of groundwater vulnerability assessment for agricultural regions in Hawaii: Review. *J. Environ. Qual.* **25**: 475–490.
25. Stewart, I.T. and Loague, K. (2003). Development of type transfer functions for regional-scale nonpoint source groundwater vulnerability assessments. *Water Resour. Res.* **39**: 1359–1371.
26. Hutson, J.L. and Wagenet, R.J. (1992). *Leaching Estimation and Chemistry Model. Version 3, User Manual*. Res. Ser. No 92-3. Dept. of Soil, Crop and Atmospheric Sciences. Cornell University, Ithaca, NY.
27. Soutter, M. and Musy, A. (1998). Coupling 1D Monte-Carlo simulations and geostatistics to assess groundwater vulnerability to pesticide contamination on a regional scale. *J. Contam. Hydrol.* **32**: 25–39.
28. Schlosser, S.A. and McCray, J.E. (2002). Sensitivity of a pesticide leaching-potential index model to variations in hydrologic and pesticide-transport properties. *Environ. Geosciences* **9**: 66–73.
29. Rao, P.S.C. and Alley, W.M. (1993). Pesticides. In: W.M. Alley (Ed.). *Regional Ground-Water Quality*. Van Nostrand Reinhold, New York, pp. 345–377.
30. Rao, P.S.C., Hornsby, A.G., and Jessup, R.E. (1985). Indices for ranking the potential for pesticide contamination of groundwater. *Soil Crop Sci. Soc. Fl. Proc.* **44**: 1–8.
31. Meeks, Y.J. and Dean, J.D. (1990). Evaluating groundwater vulnerability to pesticides. *J. Water Resour. Plann. Manage.* **116**(5): 693–707.
32. Schlosser, S.A., McCray, J.E., Murray, K.E., and Austin, B. (2002). A subregional-scale method to assess aquifer vulnerability to pesticides. *Ground Water* **40**(4): 361–367.
33. Freissinet, C., Vauclin, M., and Erlich, M. (1999). Comparison of first-order analysis and fuzzy set approach for the evaluation of imprecision in a pesticide groundwater pollution screening model. *J. Contam. Hydrol.* **37**: 21–43.
34. Loague, K. (1991). The impact of land use on estimates of pesticide leaching potential: Assessments and uncertainties. *J. Contam. Hydrol.* **8**: 157–175.

GROUNDWATER—NATURE'S HIDDEN TREASURE

Environment Canada

Groundwater is an essential and vital resource for about a quarter of all Canadians. It is their sole source of water for drinking and washing, for farming and manufacturing, indeed, for all their daily water needs. Yet for the majority of Canadians—those who do not depend on it—groundwater is a hidden resource whose value is not well understood or appreciated.

Our image of Canada is of a land of sparkling lakes, rivers and glaciers. Groundwater, which exists everywhere under the surface of the land, is not part of this picture. Not surprisingly, therefore, concerns of Canadians about water quality focus primarily on surface waters—our lakes and rivers. The less visible, but equally important, groundwater resources have received less public attention, except in regions of Canada where people depend on them.

GROUNDWATER NEEDS PROTECTION

In recent years, however, a number of events affecting groundwater quality have contributed to a heightened public awareness and concern about the importance and vulnerability of the resource. Media reports about the contamination of wells from leaking gasoline storage tanks or dry cleaning solvents, and about the effects on groundwater of chemical leakage from landfill or industrial waste disposal sites have raised public concerns about

---

This article is produced by Environment Canada and is reproduced with permission.

groundwater quality. Early in 1990, chemical seepage, caused by the tire fire at Hagersville, Ontario, threatened to pollute groundwater supplies in that region. Newspaper reports alerted area residents and helped focus attention on the problem of groundwater pollution.

Even where we might not use it directly as a drinking water supply we must still protect groundwater, since it will carry contaminants and pollutants from the land into the lakes and rivers from which other people get a large percentage of their freshwater supply.

Groundwater and Geology

Groundwater is also important quite apart from its value as a resource or its close connection with surface water supplies. Engineers must consider groundwater when planning almost any kind of structure, either above or below the ground. Ignoring the effect of groundwater on slope stability can be both costly and dangerous. Geologists see groundwater as a major force in geological change. The fluid pressures exerted by groundwater, for example, play an important role in the occurrence of earthquakes. Geologists also know that the movement of water through underground geologic formations controls the migration and the accumulation of petroleum and the formation of some ore deposits.

WHAT IS GROUNDWATER?

It is sometimes thought that water flows through underground rivers or that it collects in underground lakes. Groundwater is not confined to only a few channels or depressions in the same way that surface water is concentrated in streams and lakes. Rather, it exists almost everywhere underground. It is found underground in the spaces between particles of rock and soil, or in crevices and cracks in rock. The water filling these openings is usually within 100 metres of the surface. Much of the earth's fresh water is found in these spaces. At greater depths, because of the weight of overlying rock, these openings are much smaller, and therefore hold considerably smaller quantities of water.

Groundwater flows slowly through water-bearing formations (aquifers) at different rates. In some places, where groundwater has dissolved limestone to form caverns and large openings, its rate of flow can be relatively fast but this is exceptional.

Many terms are used to describe the nature and extent of the groundwater resource. The level below which all the spaces are filled with water is called the *water table*. Above the water table lies the *unsaturated zone*. Here the spaces in the rock and soil contain both air and water. Water in this zone is called *soil moisture*. The entire region below the water table is called the *saturated zone*, and water in this saturated zone is called *groundwater*.

WHAT IS AN AQUIFER?

Although groundwater exists everywhere under the ground, some parts of the saturated zone contain more water than others. An *aquifer* is an underground formation of permeable rock or loose material which can produce useful quantities of water when tapped by a well. Aquifers come in all sizes. They may be small, only a few hectares in area, or very large, underlying thousands of square kilometres of the earth's surface. They may be only a few metres thick, or they may measure hundreds of metres from top to bottom.

GROUNDWATER—ALWAYS ON THE MOVE

Permeable material contains interconnected cracks or spaces that are both numerous enough and large enough to allow water to move freely. In some permeable materials groundwater may move several metres in a day; in other places, it moves only a few centimetres in a century. Groundwater moves very slowly through relatively *impermeable* materials such as clay and shale.

Groundwater scientists generally distinguish between two types of aquifers in terms of the physical attributes of the aquifer: porous media and fractured aquifers.

Porous media are those aquifers consisting of aggregates of individual particles such as sand or gravel. The groundwater occurs in and moves through the openings between the individual grains. Porous media where the grains are not connected to each other are considered *unconsolidated*. If the grains are cemented together, such aquifers are called *consolidated*. Sandstones are examples of consolidated porous media.

Fractured aquifers are rocks in which the groundwater moves through cracks, joints or fractures in otherwise solid rock. Examples of fractured aquifers include granite and basalt. Limestones are often fractured aquifers, but here the cracks and fractures may be enlarged by solution, forming large channels or even caverns. Limestone terrain where solution has been very active is termed *karst*. Porous media such as sandstone may become so highly cemented or recrystallized that all of the original space is filled. In this case, the rock is no longer a porous medium. However, if it contains cracks it can still act as a fractured aquifer.

Most of the aquifers of importance to us are unconsolidated porous media such as sand and gravel. Some very porous materials are not permeable. Clay, for instance, has many spaces between its grains, but the spaces are not large enough to permit free movement of water.

Groundwater usually flows downhill with the slope of the water table. Like surface water, groundwater flows toward, and eventually drains into streams, rivers, lakes and the oceans. Groundwater flow in the aquifers underlying surface drainage basins, however, does not always mirror the flow of water on the surface. Therefore, groundwater may move in different directions below the ground than the water flowing on the surface.

Unconfined aquifers are those that are bounded by the water table. Some aquifers, however, lie beneath layers of impermeable materials. These are called *confined aquifers*, or sometimes *artesian aquifers*. A well in such an aquifer is called an *artesian well*. The water in these wells rises higher than the top of the aquifer because of confining pressure. If the water level rises above the ground surface a *flowing artesian well* occurs. The *piezometric surface* is the level to which the water in an artesian aquifer will rise.

Groundwater As a Source of Energy

Groundwater may be used as a source of heat. Ground source heat pumps are receiving increased attention as energy efficient commercial and residential heating/cooling systems. Although initial costs are higher than air source systems—due to the additional costs of the underground installations—the much greater energy efficiency of ground source systems makes them increasingly attractive.

Research into the use of geothermal water has been carried out in a number of institutions across Canada. The City of Moose Jaw has developed a geothermal heating system for a public swimming pool and recreational facility. Carleton University in Ottawa already uses groundwater to heat and cool its buildings. The Health Centre complex in Sussex, New Brunswick, has been utilizing an aquifer for thermal energy storage since 1995.

GROUNDWATER—A MAJOR LINK IN THE HYDROLOGIC CYCLE

Groundwater circulates as part of the hydrologic cycle. As precipitation and other surface water sources recharge the groundwater it drains steadily, and sometimes very slowly, towards its discharge point.

Groundwater does not stay underground forever, and it does not lie still waiting for us to draw it from a well. The *hydrologic cycle* is the series of transformations that occur in the circulation of water from the atmosphere onto the surface and into the subsurface regions of the earth, and then back from the surface to the atmosphere. Precipitation becomes surface water, soil moisture, and groundwater. Groundwater circulates back to the surface, and from the surface all water returns to the atmosphere through evaporation and transpiration.

When precipitation falls on the land surface, part of the water runs off into the lakes and rivers. Some of the water from melting snow and from rainfall seeps into the soil and percolates into the saturated zone. This process is called recharge (Fig. 2). Places where recharge occurs are referred to as *recharge areas*.

Eventually, this water reappears above the ground. This is called *discharge*. Groundwater may flow into streams, rivers, marshes, lakes and oceans, or it may discharge in the form of *springs* and flowing wells.

Groundwater discharge can contribute significantly to surface water flow. In dry periods, the flow of some streams may be supplied entirely by groundwater. At all times of the year, in fact, the nature of underground formations has a profound effect on the volume of surface runoff. While the rate of discharge determines the volume of water moving from the saturated zone into streams, the rate of recharge determines the volume of water running over the surface. When it rains, for instance, the volume of water running into streams and rivers depends on how much rainfall the underground materials can absorb. When there is more water on the surface than can be absorbed into the groundwater zone, it runs off into streams and lakes.

The *residence time* of groundwater, i.e., the length of time water spends in the groundwater portion of the hydrologic cycle, varies enormously. Water may spend as little as days or weeks underground, or as much as 10 000 or more years. Residence times of tens, hundreds, or even thousands of years are not unusual. By comparison, the average turnover time of river water, or the time it takes the water in rivers to completely replace itself, is about two weeks.

Groundwater and Permafrost

Most of the Northwest Territories is covered by either the Canadian Shield or permafrost (permanently frozen ground), both of which inhibit the flow of groundwater. The major exceptions include the Mackenzie Mountains in the western Northwest Territories and Yukon, and the limestone terrain southwest of Great Slave Lake, where soils, fractured rock, and glacial debris provide material that can store and release groundwater.

On a local scale, the seasonal development of a thawed "active layer" above the permafrost can provide permeable pathways for the subsurface movement of water and contaminants.

ALMOST NINE MILLION CANADIANS DEPEND ON GROUNDWATER

In Canada, 8.9 million people, or 30.3% of the population, rely on groundwater for domestic use. Approximately two thirds of these users live in rural areas. In many areas, wells produce more reliable and less expensive water supplies than those obtained from nearby lakes, rivers and streams. The remaining users are located primarily in smaller municipalities where groundwater provides the primary source for their water supply systems. For instance, 100% of Prince Edward Island's population and over 60% of the population of New Brunswick rely on groundwater to meet their domestic needs.

Furthermore, the predominant use of groundwater varies by province. In Ontario, Prince Edward Island, New Brunswick, and the Yukon, the largest users of groundwater are municipalities; in Alberta, Saskatchewan, and Manitoba, the agricultural industry for livestock watering; in British Columbia, Quebec and the Northwest Territories, industry; and in Newfoundland and Nova Scotia, rural domestic use. Prince Edward Island is almost totally dependent on groundwater for all its uses.

According to some estimates, the quantity of groundwater in the earth would cover the entire surface of the globe to a depth of 120 metres. By contrast, the volume of surface water in lakes, rivers, reservoirs and swamps could be contained in a depth of about one quarter of a metre.

It is extremely difficult to estimate the volume of groundwater on the entire planet. For example, a recent review of the literature revealed estimated figures ranging from 7 000 000 to 330 000 000 cubic kilometres. However,

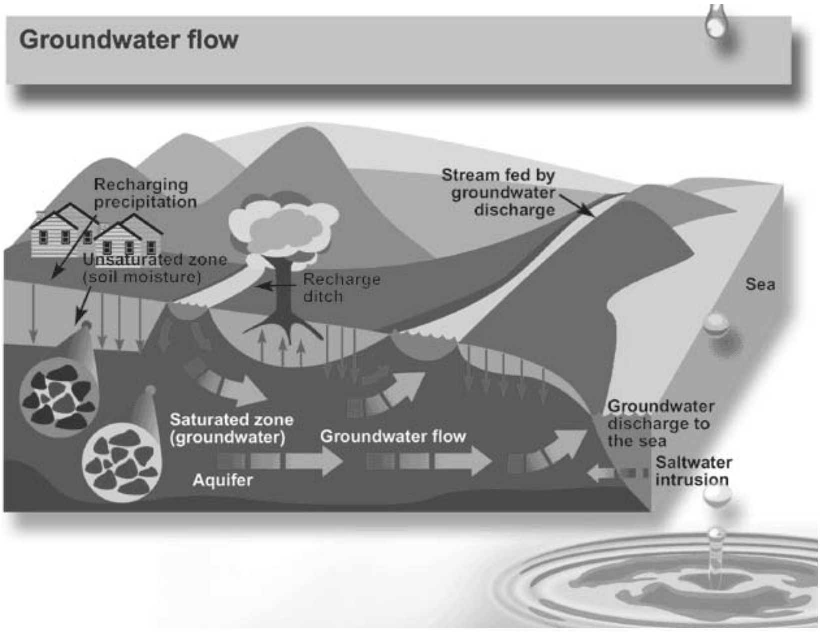

all the estimates imply that if we do not include the water frozen in ice caps, glaciers and permanent snow, groundwater makes up almost the entire volume of the earth's usable *fresh* water. Yet, this supply is often not easily accessible, and it may be difficult and expensive to develop these water supplies in some regions. The quality of the groundwater source is also a significant determining factor when identifying its use.

GROUNDWATER QUALITY

We often think of water quality as a matter of taste, clarity and odour, and in terms of other properties which determine whether water is fit for drinking. For other uses different properties may be important. Most of these properties depend on the kinds of substances that are dissolved or suspended in the water. Water for most industrial uses, for instance, must not be corrosive and must not contain dissolved solids that might precipitate on the surfaces of machinery and equipment.

Pure water is tasteless and odourless. A molecule of water contains only hydrogen and oxygen atoms. Water is never found in a pure state in nature. Both groundwater and surface water may contain many constituents, including microorganisms, gases, inorganic and organic materials.

The chemical nature of water continually evolves as it moves through the hydrologic cycle. The kinds of chemical constituents found in groundwater depend, in part, on the chemistry of the precipitation and recharge water. Near coastlines, precipitation contains higher concentrations of sodium chloride, and downwind of industrial areas, airborne sulphur and nitrogen compounds make precipitation acidic.

One of the most important natural changes in groundwater chemistry occurs in the soil. Soils contain high concentrations of carbon dioxide which dissolves in the groundwater, creating a weak acid capable of dissolving many silicate minerals. In its passage from recharge to discharge area, groundwater may dissolve substances it encounters or it may deposit some of its constituents along the way. The eventual quality of the groundwater depends on temperature and pressure conditions, on the kinds of rock and soil formations through which the groundwater flows, and possibly on the residence time. In general, faster flowing water dissolves less material. Groundwater, of course, carries with it any soluble contaminants which it encounters.

Saltwater Intrusion

Saltwater intrusion can be a problem in coastal areas where rates of groundwater pumping are high enough to cause sea water to invade freshwater aquifers (Fig. 2). The problem can be avoided by appropriate well field design and by drilling relief wells to keep the salt water away from the fresh groundwater source. Some wells pumping saltwater in Prince Edward Island, are used as convenient water supplies for shellfish farms.

Scientists assess *water quality* by measuring the amounts of the various constituents contained in the water. These amounts are often expressed as milligrams per litre (mg/L), which is equivalent to the number of grams of a substance per million grams of water.

The suitability of water for a given use depends on many factors such as hardness, salinity and pH. Acceptable values for each of these parameters for any given use depend on the use, not on the source of the water, so that the considerations important for surface water (as mentioned in Freshwater Series No. A-3, entitled "Clean Water—Life Depends on It!") are equally applicable to groundwater.

The natural quality of groundwater differs from surface water in that:

a. for any given source, its quality, temperature and other parameters are less variable over the course of time; and,
b. in nature, the range of groundwater parameters encountered is much larger than for surface water, e.g., total dissolved solids can range from 25 mg/L in some places in the Canadian Shield to 300 000 mg/L in some deep saline waters in the Interior Plains.

At any given location, groundwater tends to be harder and more saline than surface water, but this is by no means a universal rule. It is also generally the case that groundwater becomes more saline with increasing depth, but again, there are many exceptions.

Leaking Underground Storage Tanks and Piping

Leaks of petroleum products have been increasing over the last two decades because underground steel tanks installed in large numbers in the 1950s and 1960s have become corroded. Before 1980, most underground tanks were made of steel. Without adequate corrosion protection, up to half of them leak by the time they are 15 years old.

Groundwater dissolves many different compounds, and most of these substances have the potential to contaminate large quantities of water. For example, one litre of gasoline can contaminate 1 000 000 litres of groundwater. This problem is particularly severe in the Atlantic provinces where there is a high usage of groundwater. In many cases, the problem is noticed long after the aquifer is contaminated, for example, when consumers start tasting or smelling gasoline.

As groundwater flows through an aquifer it is naturally filtered. This filtering, combined with the long residence time underground, means that groundwater is usually free from disease-causing microorganisms. A source of contamination close to a well, however, can defeat these natural safeguards. Natural filtering also means that groundwater usually contains less suspended material and undissolved solids than surface water.

DNAPLs

A type of contaminant that is especially troublesome is the group of chemicals known as *dense non-aqueous phase liquids*, or *DNAPL*s. These include chemicals used in dry cleaning, wood preservation, asphalt operations, machining, and in the production and repair of automobiles, aviation equipment, munitions, and electrical equipment. They can also be generated and released in accidents, e.g., the Hagersville "tire fire." These substances are heavier than water and they sink quickly into the ground. This makes spills of DNAPLs more difficult to handle than spills of petroleum products. As with petroleum products, the problems are caused by groundwater dissolving some of the compounds in these volatile substances. These compounds can then move with the groundwater flow. Except in large cities, drinking water is rarely tested for these contaminants.

HOW WE CONTAMINATE GROUNDWATER

Any addition of undesirable substances to groundwater caused by human activities is considered to be *contamination*. It has often been assumed that contaminants left on or under the ground will stay there. This has been shown to be wishful thinking. Groundwater often spreads the effects of dumps and spills far beyond the site of the original contamination. Groundwater contamination is extremely difficult, and sometimes impossible, to clean up.

Sources of Contamination That Can Cause Groundwater Contamination

Point sources

- On-site septic systems
- Leaky tanks or pipelines containing petroleum products
- Leaks or spills of industrial chemicals at manufacturing facilities
- Underground injection wells (industrial waste)
- Municipal landfills
- Livestock wastes
- Leaky sewer lines
- Chemicals used at wood preservation facilities
- Mill tailings in mining areas
- Fly ash from coal-fired power plants
- Sludge disposal areas at petroleum refineries
- Land spreading of sewage or sewage sludge
- Graveyards
- Road salt storage areas
- Wells for disposal of liquid wastes
- Runoff of salt and other chemicals from roads and highways
- Spills related to highway or railway accidents
- Coal tar at old coal gasification sites
- Asphalt production and equipment cleaning sites

Non-point (distributed) sources

- Fertilizers on agricultural land
- Pesticides on agricultural land and forests
- Contaminants in rain, snow, and dry atmospheric fallout

Source: Adapted from Cherry, p. 395.

Groundwater contaminants come from two categories of sources: *point sources* and distributed, or *non-point sources*. Landfills, leaking gasoline storage tanks, leaking septic tanks, and accidental spills are examples of point sources. Infiltration from farm land treated with pesticides and fertilizers is an example of a non-point source.

Among the more significant point sources are municipal landfills and industrial waste disposal sites. When either of these occur in or near sand and gravel aquifers, the potential for widespread contamination is the greatest.

In Ville Mercier, Quebec, for example, the disposal of industrial wastes into lagoons in an old gravel pit over many years rendered the water supplies of thousands of residents in the region unusable. Water had to be pumped from a well 10 kilometres away to replace the area's supply.

Other point sources are individually less significant, but they occur in large numbers all across the country. Some of these dangerous and widespread sources of contamination are septic tanks, leaks and spills of petroleum products and of dense industrial organic liquids.

Septic systems are designed so that some of the sewage is degraded in the tank and some is degraded and absorbed by the surrounding sand and subsoil. Contaminants that may enter groundwater from septic systems include bacteria, viruses, detergents, and household cleaners. These can create serious contamination problems. Despite the fact that septic tanks and cesspools are known sources of contaminants, they are poorly monitored and very little studied.

Contamination can render groundwater unsuitable for use. Although the overall extent of the problem across Canada is unknown, many individual cases of contamination have been investigated such as Ville Mercier in Quebec; the highway de-icing salt problem in Nova Scotia; industrial effluents in Elmira, Ontario; various pesticides in the Prairie provinces; industrial contamination in Vancouver, British Columbia; and so on. In many cases, contamination is recognized only after groundwater users have been exposed to potential health risks. The cost of cleaning up contaminated water supplies is usually extremely high.

Contamination problems are increasing in Canada primarily because of the large and growing number of toxic compounds used in industry and agriculture. In rural Canada, scientists suspect that many household wells are contaminated by substances from such common sources as septic systems, underground tanks, used motor oil, road salt, fertilizer, pesticides, and livestock wastes. Scientists also predict that in the next few decades more contaminated aquifers will be discovered, new contaminants will be identified, and more contaminated groundwater will be discharged into wetlands, streams and lakes.

Once an aquifer is contaminated, it may be unusable for decades. The residence time, as noted earlier, can be anywhere from two weeks or 10 000 years.

Furthermore, the effects of groundwater contamination do not end with the loss of well-water supplies. Several studies have documented the migration of contaminants from disposal or spill sites to nearby lakes and rivers as this groundwater passes through the hydrologic cycle, but the processes are not as yet well understood. In Canada, pollution of surface water by groundwater is probably at least as serious as the contamination of groundwater supplies. Preventing contamination in the first place is by far the most practical solution to the problem. This can be accomplished by the adoption of effective groundwater management practices by governments, industries and all Canadians. Although progress is being made in this direction, efforts are hampered by a serious shortage of groundwater experts and a general lack of knowledge about how groundwater behaves.

GROUNDWATER AND ENGINEERING

Groundwater can also have dramatic implications for engineering and geotechnical studies. The study of groundwater is essential for engineers who construct dams, tunnels, water conveyance channels, mines, and other structures. Groundwater must be considered whenever the stability of slopes is important, whether the slope is natural or constructed. Groundwater must also be taken into account when devising measures to control flooding. In all of these situations, groundwater flow and fluid pressure can create serious *geotechnical problems*.

Groundwater, for example, may create structural weaknesses in dams, or it may flow underground right around the structure as it did at the Jerome Dam in Idaho. Water flowed so efficiently through the rock formations surrounding the reservoir that the dam would hold no water, even though it was structurally sound.

In another case, when geological exploration was being carried out in preparation for the construction of the Revelstoke Dam in British Columbia, geologists and engineers were concerned about an old landslide on the bank of the proposed reservoir. They suspected that the water held in the reservoir could increase groundwater pressures enough to make the slide unstable. The solution was to increase drainage around the slide to ensure that groundwater pressures did not increase. In 1963, these same conditions at the Vaiont reservoir in Italy caused a slide which killed 2500 people.

Other problems result from the excessive use of groundwater. *Overdrafting* occurs when people draw water out of an aquifer faster than nature can replenish it. The most obvious problem created is a shortage of water. Overdrafting, however, can also create significant geotechnical problems. Although not an issue in Canada, at many locations around the world overdrafting has caused *land subsidence*. This can produce severe engineering difficulties. Parts of Mexico City, for instance, have subsided as much as 10 metres in the past 70 years, resulting in a host of problems in its water supply and sewerage system. Land subsidence may also occur when the water table is lowered by drainage. In the early 1970s, for example, an entire residential subdivision in Ottawa subsided when a collector sewer was constructed nearby. The subsidence seriously damaged the residents' property.

Groundwater and Wetlands

Wetlands, which provide a summer home to nearly all of North America's 45 million ducks and other waterfowl, often have very close connections with the groundwater system. Some wetlands, e.g., potholes in higher ground, may serve as important groundwater recharge areas. Others, especially those in low-lying areas, may be the receptors for significant amounts of groundwater discharge. Therefore, if the underlying groundwater is contaminated, detrimental consequences will be felt by the wildlife and all other resources dependent on that wetland.

SAFEGUARDING OUR GROUNDWATER SUPPLY

Groundwater is an essential resource. It exists everywhere under the Canadian landscape and is vitally connected to our rich surface water resources.

Contamination of groundwater is a serious problem in Canada. Industrial and agricultural activities are major sources of contaminants, but Canadian households are equally important sources.

Groundwater moves so slowly that problems take a long time to appear. Because of this, and because it is so expensive to clean up a contaminated aquifer (if it can be done at all), it is preferable by far to prevent contamination from happening in the first place. For example, leaking underground storage tanks can be replaced by tanks that will not corrode; landfills can be sited in locations where leachates will not contaminate underlying groundwater; and the impacts of spills of hazardous materials reduced by restricting access to recharge areas.

All levels of government in Canada are starting to take some of the actions necessary to protect our groundwater supplies, but there is a long way to go before these measures are fully effective. At the same time, universities and government research institutes are investigating what happens to water underground and what can be done to preserve it and even improve its availability to us. Both as a society and as individuals, we must keep in mind groundwater's susceptibility to contamination.

Groundwater is just as important as the sparkling lakes and rivers of our postcard image of Canada. This national treasure may be "hidden," but it must not be forgotten.

READING LIST

American Chemical Society. *Ground Water Information Pamphlet*. American Chemical Society, Department of Public Affairs: Washington, 1983. 14 pp.

American Institute of Professional Geologists. *Ground Water: Issues and Answers*. American Institute of Professional Geologists: Arvado, Colo., 1986. 24 pp.

Bowen, Robert. *Groundwater*. Elsevier Applied Science Publishers: London, UK, 1986. 427 pp.

Cherry, John A. "Groundwater Occurrence and Contamination in Canada." In M.C. Healey and R.R. Wallace, *Canadian Aquatic Resources*, eds., Canadian Bulletin of Fisheries and Aquatic Sciences 215: 387–426. Department of Fisheries and Oceans: Ottawa, 1987.

Coon, David, and Janice Harvey. *The Groundwater Pollution Primer*. Conservation Council, Fredericton, N.B., July 1987. 44 pp.

Federal-Provincial Advisory Committee on Environmental and Occupational Health. *Guidelines for Canadian Drinking Water Quality*, 4th ed., Health and Welfare Canada: Ottawa, 1989. 25 pp.

Freeze, R.Allen, and John A. Cherry. *Groundwater*. Prentice-Hall: Englewood Cliffs N.J., 1979. 604 pp.

Hess, Paul J. *Ground-Water Use in Canada, 1981*. National Hydrology Research Institute: Ottawa, 1986. NHRI Paper No. 28, IWD Technical Bulletin No. 140. 43 pp.

Johnston, L.M., and J.A. Gilliland. *"Ground Water—Why Worry?"* Unpublished Report, National Hydrology Research Institute, Saskatoon, Sask., 1987. NHRI Contribution No. 87008.

Keating, Michael. *To the Last Drop: Canada and the World's Water Crisis*. MacMillan: Toronto, 1986. 265 pp.

Laycock, A.H. "The Amount of Canadian Water and Its Distribution." In M.C. Healey and R.R. Wallace, *Canadian Aquatic Resources*, eds., Canadian Bulletin of Fisheries and Aquatic Sciences 215: 13–42. Department of Fisheries and Oceans: Ottawa, 1987.

McNeely, R.N., V.P. Neimanis, and L. Dwyer. *Water Quality Sourcebook: A Guide to Water Quality Parameters*. Inland Waters Directorate, Water Quality Branch: Ottawa, 1979. 89 pp.

Price, Michael. *Introducing Groundwater*. George Allen & Unwin: London, UK, 1985. 195 pp.

Speidel, David H., L.C. Ruedisili, and A.F. Agnew. *Perspectives on Water, Uses and Abuses*. Oxford University Press: New York, N.Y., 1988.

Todd, David Keith. *Groundwater Hydrology*, 2nd ed. John Wiley & Sons: New York, 1980. 535 pp.

U.S. Environmental Protection Agency. *EPA Ground Water Handbook*. Government Institutes, Inc.: Rockville, Md., 1989. 212 pp.

FRESHWATER SERIES A-5

Note: A resource guide, entitled Let's Not Take Water For Granted, is available to help classroom teachers of grades 5–7 use the information from the Water Fact Sheets.

PHARMACEUTICALS, HORMONES, AND OTHER ORGANIC WASTEWATER CONTAMINANTS IN U.S. STREAMS

Herbert T. Buxton
Dana W. Kolpin
United States Geological Survey

A recent study by the Toxic Substances Hydrology Program of the U.S. Geological Survey (USGS) shows that a broad range of chemicals found in residential,

---

industrial, and agricultural wastewaters commonly occurs in mixtures at low concentrations downstream from areas of intense urbanization and animal production. The chemicals include human and veterinary drugs (including antibiotics), natural and synthetic hormones, detergent metabolites, plasticizers, insecticides, and fire retardants. One or more of these chemicals were found in 80 percent of the streams sampled. Half of the streams contained 7 or more of these chemicals, and about one-third of the streams contained 10 or more of these chemicals. This study is the first national-scale examination of these organic wastewater contaminants in streams and supports the USGS mission to assess the quantity and quality of the Nation's water resources. A more complete analysis of these and other emerging water-quality issues is ongoing.

BACKGROUND

Chemicals, used everyday in homes, industry and agriculture, can enter the environment in wastewater. These chemicals include human and veterinary drugs (including antibiotics), hormones, detergents, disinfectants, plasticizers, fire retardants, insecticides, and antioxidants. To assess whether these chemicals are entering our Nation's streams, the Toxic Substances Hydrology Program of the U.S. Geological Survey (USGS) collected and analyzed water samples from 139 streams in 30 states during 1999 and 2000. Streams were sampled that were considered susceptible to contamination from various wastewater sources, such as those downstream from intense urbanization or livestock production. Thus, the results of this study are not considered representative of all streams.

Although each of the 95 chemicals is used extensively, there is little information about the extent or occurrence of many of these compounds in the environment. Some may be indicators of certain classes of contamination sources, such as livestock or human waste, and some have human or environmental health implications. The results of this study are a starting point for investigation of the transport of a wide range of organic wastewater contaminants in the Nation's waters.

New laboratory methods were developed in several USGS research laboratories to provide the analytical capability to measure concentrations of 95 wastewater-related organic chemicals in water. Uniform sample-collection protocols and field and laboratory quality-assurance programs were followed to ensure that results are comparable and representative of actual stream conditions.

FINDINGS

One or more chemicals were detected in 80 percent of the streams sampled, and 82 of the 95 chemicals were detected at least once. Generally, these chemicals were found at very low concentrations (in most cases, less than 1 part per billion). Mixtures of the chemicals were common; 75 percent of the streams had more than one, 50 percent had 7 or more, and 34 percent had 10 or more.

The most frequently detected chemicals (found in more than half of the streams) were coprostanol (fecal steroid), cholesterol (plant and animal steroid), N-N-diethyltoluamide (insect repellent), caffeine (stimulant), triclosan (antimicrobial disinfectant), tri (2-chloroethyl) phosphate (fire retardant), and 4-nonylphenol (nonionic detergent metabolite). Steroids, nonprescription drugs, and insect repellent were the chemical groups most frequently detected. Detergent metabolites, steroids, and plasticizers generally were measured at the highest concentrations.

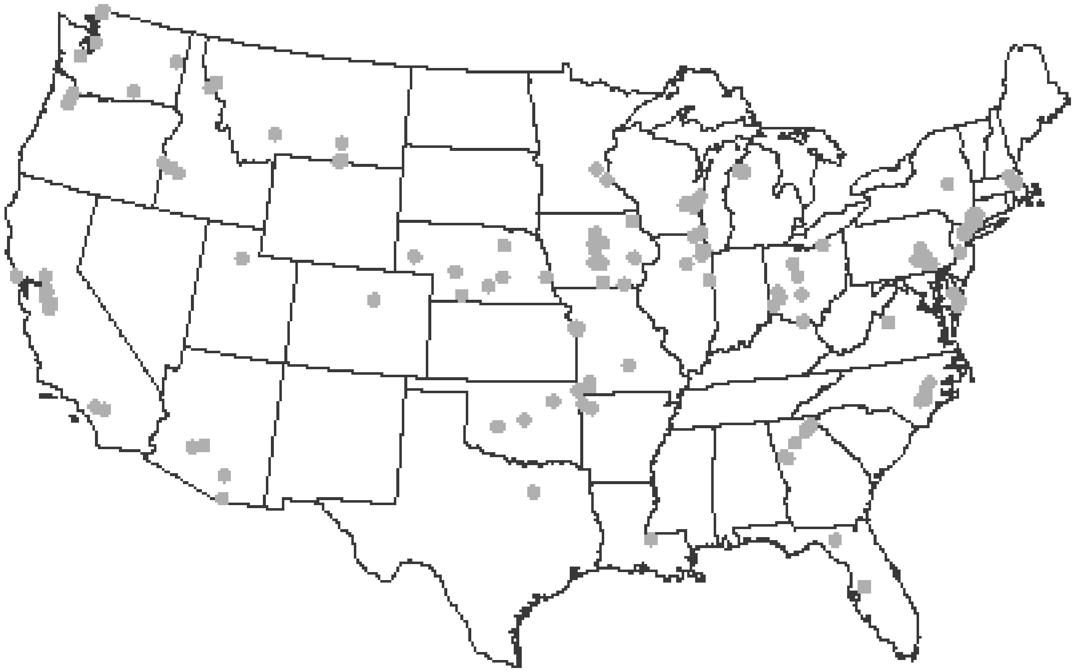

Pharmaceuticals, hormones, and other organic wastewater contaminants were measured in 139 streams during 1999 and 2000.

HUMAN AND ENVIRONMENTAL EFFECTS

Knowledge of the potential human and environmental health effects of these 95 chemicals is highly varied; drinking-water standards or other human or ecological health criteria have been established for 14. Measured concentrations rarely exceeded any of the standards or criteria. Thirty-three are known or suspected to be

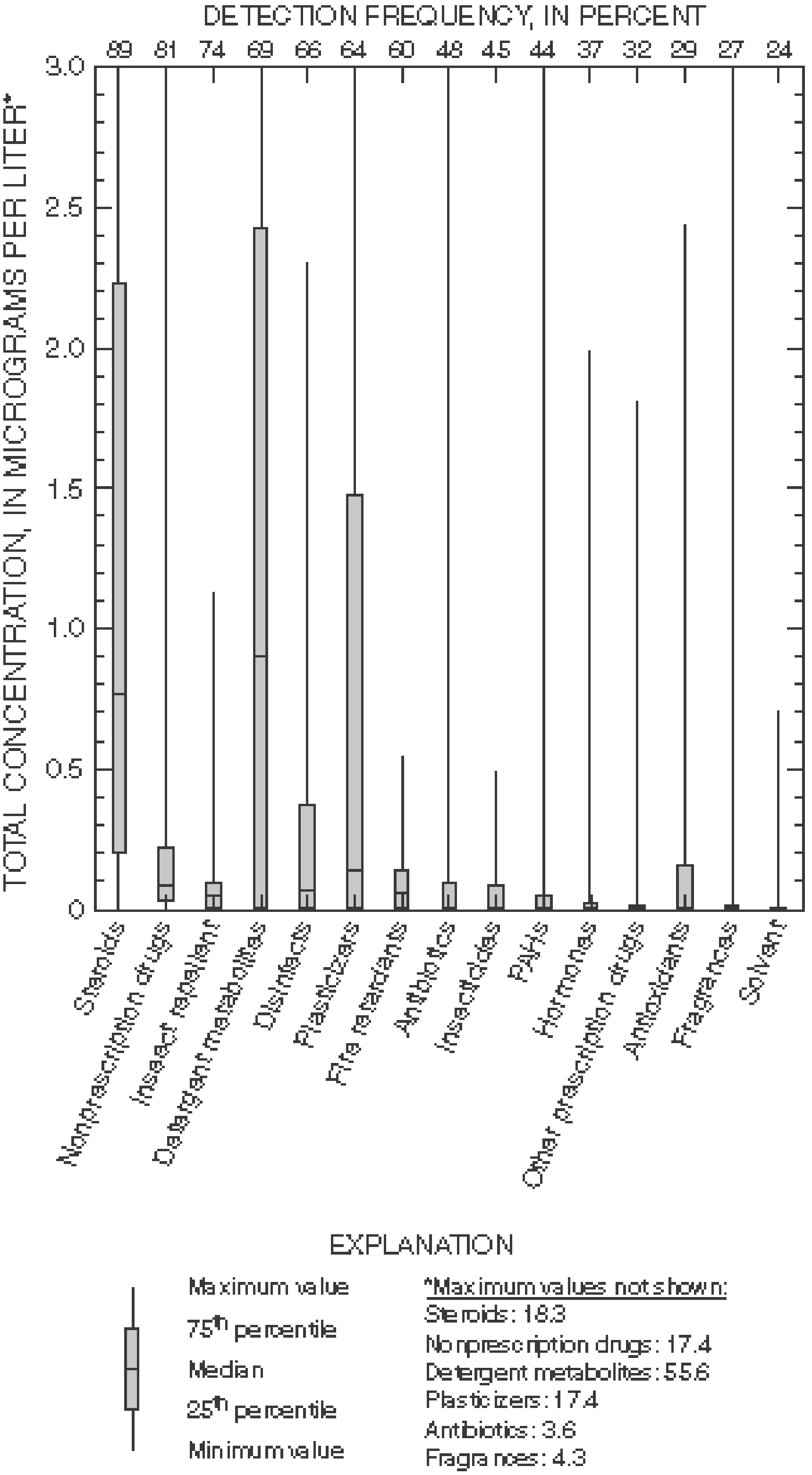

Steroids, nonprescription drugs, and an insect repellent were the three chemical groups most commonly detected in susceptible streams. Detergent metabolites, steroids, and plasticizers generally were found at the highest concentrations.

hormonally active; 46 are pharmaceutically active. Little is known about the potential health effects to humans or aquatic organisms exposed to the low levels of most of these chemicals or the mixtures commonly found in this study.

SIGNIFICANCE OF FINDINGS

This study suggests that mixtures of pharmaceuticals, hormones, and other wastewater contaminants can occur at low concentrations in streams that are susceptible to various wastewater sources. It provides methodology and guidance for future monitoring and assessment of these types of environmental contaminants, and establishes the needed foundation for setting priorities for further study of sources, pathways and effects.

Household chemicals can enter streams through wastewater discharges. A wastewater treatment facility near Atlanta, Georgia, is shown above. (Photograph by Daniel J. Hippe, U.S. Geological Survey).

Veterinary pharmaceuticals used in animal agriculture can enter streams through runoff or infiltration. A swine facility near the South Fork Iowa River, Iowa, is shown above. (Photograph by Doug Schnoebelen, U.S. Geological Survey).

FUTURE DIRECTIONS

Further analyses of these data, including relationships to specific source types, are ongoing. The Toxic Substances Hydrology Program is conducting research on the occurrence of organic wastewater chemicals in susceptible wells and drinking-water sources across the nation; assessments of antibiotics and antibiotic-resistant bacteria; the identification of wastewater indicators; and the development of new laboratory analytical capabilities, including sediment and fish tissue.

ADDITIONAL INFORMATION

These findings are based on “Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999–2000: A national reconnaissance,” an article published in the March 15, 2002 issue of Environmental Science & Technology, v. 36, no. 6, pages 1202–1211. Data are presented in a companion USGS report, “Water-quality data for pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999–2000” (USGS Open-File Report 02-94). These and other reports, data, and maps can be accessed on the Internet at http://toxics.usgs.gov.

THE ENVIRONMENTAL IMPACT OF IRON IN GROUNDWATER

David B. Vance
ARCADIS G&M, Inc.
Midland, Texas

James A. Jacobs
Environmental Bio-Systems, Inc.
Mill Valley, California

Iron is a common constituent in soils and groundwater. It readily participates in subsurface redox reactions and under some conditions can cause taste, odor, and flow problems with drinking water wells and can precipitate and plug up *in situ* groundwater remediation systems. In other geochemical conditions, iron can aid in the remediation of certain chemicals undergoing chemical transformations associated with Redox reactions. Ferrous iron is used as a terminal electron acceptor for anaerobic bioremediation processes that play a dominant role in the natural attenuation process.

BACKGROUND

The average crustal abundance of iron is 5% by weight. The iron content of soils is typically in the range of 0.5% to 5% and is dependent on the source rocks from which the soil was derived, transport mechanisms, and overall geochemical history. Not only is iron common, but it is also reactive and readily reflects changes in surrounding Eh/pH conditions. This is particularly true in soil and groundwater systems that have been environmentally impacted with hydrocarbons. In groundwater systems, iron occurs in one of two oxidation states: reduced soluble divalent ferrous iron (Fe^{+2}) or oxidized insoluble trivalent ferric iron (Fe^{+3}). The modern atmosphere has 21% oxygen, which causes most iron in shallow subsurface soils to be in the oxidized ferric state. Ferric hydroxide ($Fe(OH)_3$) is the direct result of ferrous iron oxidation and precipitation. With time, ferric hydroxide is mineralized. The principal forms of mineralized ferric iron found in soils are follows:

- Amorphous hydrous ferric oxide ($Fe_2O_3 \cdot XH_2O$)
- Maghemite (gamma-Fe_2O_3)
- Lepidocrocite (gamma-FeOOH)
- Hematite (alpha-Fe_2O_3)
- Goethite (alpha-FeOOH)

The above iron oxides are listed in order of decreasing solubility, which also reflects increasing crystallinity. Amorphous hydrous ferric oxide at neutral pH and oxidizing Eh conditions has a solubility of 0.6 μg/L, which is three orders of magnitude greater than that of goethite. However, a high concentration of iron in groundwater is not rare. Obviously, other processes are at work.

IRON COMPLEXING

Ferrous iron is soluble as a cation, and ferric iron is not. However, ferric iron can form soluble complexes with many inorganic and organic ligands, including compounds that are the byproducts of the biodegradation of aromatic petroleum hydrocarbons (catechols, which are organic acids resulting from aromatic ring cleavage). The yellow/orange coloration of groundwater at some sites contaminated with petroleum hydrocarbons is caused by these ferric-catechol complexes.

It is important to remember that complexed ions typically react more slowly than do uncomplexed ions, which impacts the behavior of ferrous iron. Ferrous iron normally can be oxidized to ferric iron in minutes (see below for more details). Complexed ferrous iron may take months to complete the same reaction, which has profound implications for the fouling of groundwater injection systems. Humic acid, fulvic acid, and tannic acid are examples of noncontaminant organic complexes. Phosphate also serves as an effective complexing agent for iron.

IRON OXIDATION

The time required for uncomplexed ferrous iron to undergo oxidation to the ferric state is dependent on many factors, the dominant ones being: pH, temperature, dissolved oxygen level, and the presence of other soluble ions. The lower the pH and temperature, the longer the time required for completion of the oxidation reaction. Increasing dissolved oxygen decreases the time required for oxidation. For example:

At pH 7.0, 90% Fe^{+2} oxidation requires 1 hour at 21 °C and 10 hours at 5 °C.

At pH 8.0, 90% Fe^{+2} oxidation occurs in 30 seconds; at pH 6.0, it requires 100 hours.

The critical dissolved oxygen concentration is 2 mg/L. Below that ferrous iron oxidation occurs slowly.

The above reactions are often responsible for the iron fouling of air strippers treating iron-rich groundwater. These reactions can also foul screens in poorly designed recovery wells pumping that groundwater.

MICROBIOLOGY AND IRON

Iron readily undergoes reduction or oxidation, depending on surrounding conditions. As a consequence, many microbiological metabolic pathways use redox couples between ferrous and ferric iron. Organic material can be biodegraded with ferric iron as the terminal electron acceptor, which results in the production of reduced soluble ferrous iron. Anaerobic conditions are required as is the presence of ferric iron in a suitable form. As the crystallinity of the ferric iron mineral increases (as discussed above), the microbiological availability decreases. Or contrarily, iron fixing bacteria can oxidize ferrous iron to ferric iron, even under oxygen poor conditions, by many methods, including the extraction of carbon dioxide (for microbial energy systems) from ferrous bicarbonate leaving insoluble ferric hydroxide and the utilization of iron bearing organic acid complexes as a carbon source, leaving precipitated ferric hydroxide. Bacteria also reduce nitrate and oxidize ferrous iron to ferric iron. It is now thought that the Precambrian banded iron formations were the consequence of the mass utilization of ferrous iron in the ancient oceans shortly before oxygenation because photosynthesis took place. These are the reactions that are typically responsible for iron fouling of well screens, piping systems, and air strippers that remediate iron-rich groundwater.

IRON OXIDES AND *IN SITU* METAL/ORGANIC ADSORPTION

Amorphous hydrous ferric oxide can be described as an amphoteric ion exchange media, which means that as pH conditions change, it has the capacity to offer hydrogen ions (H^+) or hydroxyl ions (OH^-) for cation or anion exchange, respectively. As a rough rule of thumb, amorphous ferric oxide is capable of adsorbing 0.5 millimoles of ionic material per gram. Ionic material includes oxyanions of chrome, arsenic, or selenium; heavy metal cations; or ionized organic compounds. The dominant mechanism of interaction between iron oxides and soluble organic compounds is through carboxyl and hydroxyl functional groups on the organic, which is particularly true for the organic acids that are the intermediate product of the biodegradation of petroleum hydrocarbons.

Sorption behavior is primarily related to pH (within the typical range of 5.0 to 8.0), and each ion has its optimum pH range for adsorption. Given the average concentration in soil, the iron in a cubic yard of soil is capable of adsorbing from 0.5 to 5 pounds of soluble metals as cations, anionic complexes, or a similar amount of organic biodegradation products.

SUMMARY

The reactions described above are complex. At any given location, it is difficult to determine the specific processes taking place. However, the overall effect is as follows: Impacting hydrocarbons provide a carbon source to stimulate aerobic microbial degradation, which consumes the readily available oxygen driving the system anoxic. In the anoxic state, soluble ferrous ions are generated by the action of iron reducing bactgeria and soluble ferrous and ferric complexes are formed. High levels of soluble iron is an indicative side effect of hydrocarbon contamination in groundwater.

The mobilized iron is then drawn into any installed groundwater recovery and treatment systems. There, oxygenation takes place changing redox conditions, and the elevated levels of soluble iron react and which causes fouling in the recovery and treatment systems. Adsorption and attenuation of contaminant transport is concurrently taking place on the residual solid iron oxide phases remaining in the soil matrix, which can extend the duration of a pump and treat effort. Or it can be exploited in a passive “funnel and gate” groundwater barrier remediation system.

The presence and form of naturally occurring iron in groundwater are indicators of complex reactions in the subsurface. Iron in high concentrations does affect drinking water quality. It can also be an important element in aiding or lowering the efficiency of *in situ* remediation of contaminants, depending on specific geochemical conditions. It is also a key element in the natual attenuation process. Therefore, iron data are valuable information that may greatly affect everything from groundwater quality to remedial operations.

GROUNDWATER AND COBALT: CHEMICAL BEHAVIOR AND TREATMENT

James A. Jacobs
Lise Walter
Environmental Bio-Systems, Inc.
Mill Valley, California

HISTORY OF COBALT

Cobalt has been used as a pigment since 3000 B.C. The name “Kobald” may have come from the Erzgebirge region of Saxony, a silver mining region. Kobalds are mythical creatures like gnomes who frequented the mines causing trouble. The problems were due to interference of cobalt with silver smelting and causing respiratory illnesses in the miners. Before 1914 only oxides of cobalt were used; since then the metal has been used in a variety of ways, particularly as alloys and in the electronic component industry.

MINERAL COMMODITY

Cobalt, identified by Co in the periodic table, has an atomic weight of 58.9322. There are seven radioisotopes of cobalt. It is a solid at room temperature and has been classified as a transition metal. It is magnetic, hard, and brittle and has a metallic gray color. Cobalt will also exist as cobalt(II) and cobalt(III) and will form a number of organic and inorganic salts. Cobalt is not found as a discrete metal but as a solid solution with other elements in minerals. The similarity in ionic radii allows cobalt to substitute for iron, manganese, copper, and chromium. Cobalt metal is produced as a by-product of copper, nickel, and lead production, as it is not economically feasible to isolate cobalt on its own.

Worldwide, the background concentration range of cobalt in soil is 1–40 mg/kg and the average background concentration of cobalt in soil is 8 mg/kg (1). Many mineral species contain cobalt as an integral structural component. Cobalt salts are important for biological processes such as production of vitamin B_{12} and providing the blue-green color in broccoli and other vegetation. On average, humans contain 20 parts per billion (ppb) by mass of cobalt. Cobalt is also an important commercial, industrial, and military metal. Its uses are as follows:

- Cobalt salts are used for the production of brilliant and permanent blue colors in porcelain, glass, pottery, tiles, and enamels.
- For magnet, stainless, and high speed steel production.
- For electroplating, cemented carbides, and diamond tools.
- In the petroleum and chemical industry as catalysts.
- Cobalt compounds are used as paint pigments.
- Cobalt is combined with iron, nickel, and other metals to make corrosion- and wear-resistant alloys.
- Cobalt-60 is an artificial radioactive isotope used as a tracer and radiotherapy agent for cancer treatment and as a food preservation method, also called the food irradiation process.
- Cobalt is used in batteries, car tires, recording media, semiconductor ceramics, and most importantly in superalloys in jet turbines and gas turbine generators.

SOURCES OF COBALT

Mining of cobalt occurs primarily in twelve countries, most notably Canada, Russia, Democratic Republic of Congo, Zambia, and Australia, which also refine the cobalt ore.

The main countries that only refine cobalt include Finland, Norway, and Brazil.

PHYSICAL PROPERTIES

Cobalt is a shiny, gray, brittle metal with a close packed hexagonal (CPH) crystal structure at room temperature but which changes at 421 °C to a face centered cubic form. The metal is rarely used as a structural material in the pure form but almost always as an alloy or a component of another system. The density is 8.85 g/cm^3, the melting point is 1493 °C (2719 °F), the Curie point is 1121 °C, and the atomic number for cobalt is 27. Cobalt comes in two valence states: +2 and +3. For more information on physical and chemical properties, see Reference 2.

CHEMICAL REACTIONS

Cobalt is not particularly reactive with air. However, on heating the oxide Co_3O_4 is formed. If the reaction is carried out above 900 °C, the result is cobalt(II) oxide, CoO. Cobalt does not react directly with nitrogen (N_2).

$$3Co(s) + 4O_2(g) \rightarrow 2Co_3O_4(s),$$

$$2Co(s) + O_2(g) \rightarrow 2CoO(s).$$

Water has little effect on cobalt metal. The reaction between red hot cobalt metal and steam produces cobalt(II) oxide, CoO. Simple inorganic salts of cobalt are generally very soluble in water.

$$2Co(s) + O_2(g) \rightarrow 2CoO(s).$$

Cobalt metal reacts directly with halogens, producing cobalt(II) salts.

$$Co(s) + Br_2(l) \rightarrow CoBr_2(s) \quad [\text{green}],$$

$$Co(s) + Cl_2(g) \rightarrow CoCl_2(s) \quad [\text{blue}],$$

$$Co(s) + I_2(s) \rightarrow CoI_2(s) \quad [\text{blue-black}].$$

Cobalt metal dissolves slowly in dilute sulfuric acid to form solutions containing the aqueous Co(II) ion together with hydrogen gas, H_2. In practice, Co(II) is present as the complex ion $[Co(OH_2)_6]^{2+}$ (see Reference 3).

$$Co(s) + H_2SO_4(aq) \rightarrow Co^{2+}(aq) + SO_4^{2-}(aq) + H_2(g),$$

COBALT ISOLATION

Numerous ores contain cobalt, but not many of these are of economic importance. These include the sulfides and arsenides linnaeite (Co_3S_4), cobaltite, (CoAsS), and smaltite, ($CoAs_2$). Normally the cobalt ore is "roasted" to form a mixture of metals and metal oxides. Treatment with sulfuric acid leaves metallic copper as a residue and dissolves out iron, cobalt, and nickel as the sulfates. Iron is obtained by precipitation with lime (CaO) while cobalt is produced as the hydroxide by precipitation with sodium hypochlorite (NaOCl):

$$2Co^{2+}(aq) + NaOCl(aq) + 4OH^{-}(aq) + H_2O \rightarrow 2Co(OH)_3(s) + NaCl(aq).$$

The trihydroxide $Co(OH)_3$ is heated to form the cobalt oxide and then reduced with carbon (as charcoal) to form cobalt metal (3).

$$2Co(OH)_3(\text{heat}) \rightarrow Co_2O_3 + 3H_2O,$$

$$2Co_2O_3 + 3C \rightarrow Co + 3CO_2.$$

FATE AND TRANSPORT

The geochemical behavior of cobalt generally follows the iron–manganese system. The cobalt concentration in soils is controlled by adsorption and coprecipitation reactions with manganese and iron oxide minerals. Cobalt-60 (^{60}Co) is an activation product of stable cobalt (^{59}Co). It is the isotope of primary importance to waste disposal and site remediation. ^{60}Co has a half-life of about 5.3 yr, relatively short compared to other radionuclides of environmental concern. It readily adsorbs to minerals and as such its mobility should be limited to soil and water systems near to the area in which it was released (2).

Co(II) is the stable valence state in water under most geochemical conditions. Co(III) is a strong oxidizing agent; hence it is not thermodynamically stable and will decompose under the conditions typical of natural water. Cobalt(III) will form strong complexes with organic ligands such as ethylenediamine-tetraacetic acid (EDTA) and NH_3 and exist in aqueous solution. These complexes affect cobalt's environmental mobility by increasing its solubility in aqueous solutions, decreasing its adsorption in soil, and stabilizing Co(III) in some sediment systems. Co(III) can be mobilized and reduced to Co(II) as a result of bacterial iron reduction. The solubility of Co(III) is a function of the composition of the media and buffers (see Fig. 1).

The migration and sorption of cobalt depends strongly on the pH of the medium. In acidic media, cobalt occurs in solution; an increase in pH will cause it to sorb to soil, sediments, and other precipitates such as manganese and iron oxides and clay minerals. Adsorption is minimal below a pH of 4. In the absence of organic complexants cobalt will adsorb (moderate to high) to minerals at near neutral and basic pH values. Humic acid will increase the adsorption of cobalt on all mineral sorbents by 10–60%. The highest cobalt adsorption due to humic acid will occur between pH 4.5 and 6.5. K_d for dissolved Co(II) ranges from 10^3 to 10^5 mL/g. Inorganic ligands such as cyanide in low concentrations can decrease cobalt adsorption on sediment significantly ($K_d < 2$ mL/g). The presence of other cations (K^+, Ca^{2+}, Mg^{2+}, Na^+) in solution can decrease cobalt adsorption.

Adsorption on manganese minerals results in the surface-mediated oxidation of Co(II) to Co(III) by the reduction of surface Mn(IV) sites. This will occur at pH values of 4, 6, and 7. At pH 8 and 10, Co(II) adsorption resulted in the precipitation of something similar to cobalt

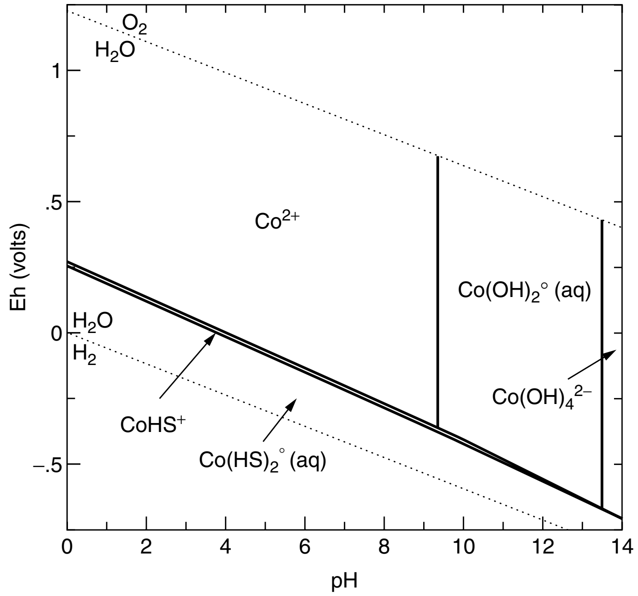

Figure 1. Eh−pH diagram showing the dominant aqueous complexes of cobalt (2). (Diagram was calculated at 25 °C and a concentration of 10−14.8 mol/L total dissolved cobalt in the presence of dissolved chloride, nitrate, sulfate, and carbonate).

hydroxide; this could explain the large range of K_d values reported in the literature (2).

HEALTH AND SAFETY

Exposure to very high levels of cobalt can cause serious health problems to the respiratory system and skin of humans and animals. Most typically, persons breathing air with high cobalt levels are at risk for lung problems, wheezing, asthma, and pneumonia. Cobalt has been found to be a sensitizer and as such dermatitis can occur on contact. Three cobalt compounds are classified as a human carcinogen on the PHS Category 1 Chemical Table maintained by the Los Alamos National Laboratory (2000). The three carcinogenic compounds include cobalt(II) arsenate, cobalt−chromium alloy (as Cr^{6+}), and beryllium−copper−cobalt alloy.

Research has shown that a high cobalt level is harmful to fetus development in pregnant animals. It is known to cause cancer in animals when placed under the skin or in the muscle. The lowest published lethal dose by oral application in rats is 1500 mg/kg, 250 mg/kg for intraperitoneal doses, and 25 mg/kg for intratracheal doses. Inhalation is the most hazardous exposure pathway of cobalt. When working with cobalt, goggles and gloves are recommended for personal protection. If cobalt powder is present, the area should be well ventilated and inert, as the powder can ignite spontaneously.

Several countries have determined different toxicity levels for cobalt.

In Ontario, the cobalt toxicity limit in soil is 40 parts per million (ppm) (4). In Sweden the limit is 30 ppm (5) and in Quebec the agricultural soil standard is 15 ppm.

PERSISTENCE IN GROUNDWATER

Cobalt will exist in groundwater as Co(II) or Co(III) depending on the pH. Under most geochemical conditions, Co(II) is the stable valence state in water. Co(III) is not thermodynamically stable and decomposes under Eh−pH conditions common for most natural waters. However, the presence of certain complexing ligands can stabilize Co(III) relative to reduction and allow it to persist in aqueous solutions. Cobalt will be retarded in groundwater due to its large K_d values (2).

COBALT CONTAMINATION

Cobalt is naturally occurring in the environment and is found in a variety of places, including plants, soil, dust, and seawater. Anthropogenic releases are by way of burning coal and oil and fuel exhaust from cars, trucks, airplanes, and many industrial processes that use any of the numerous cobalt compounds. Most sites with cobalt contamination have occurred due to mining operations. Cobalt contamination is generally greatest in the soil around mine sites due to the leaching acids in the tailings. Cobalt contamination is usually coupled with high arsenic and nickel contaminations and is rarely the contaminant of major concern due to its low toxicity. Cobalt-60 contamination does occur when storage tanks leak or disposal practices are inadequate. Contamination by ^{60}Co is generally of less concern than other radionuclides based on its short half-life. ^{60}Co is highly sorbed on sandy soils and relatively immobile. When ^{60}Co is complexed with EDTA, the cobalt mobility increases.

COBALT REMEDIATION TECHNIQUES

Remediation of cobalt-contaminated sites can be carried out via removal of the soil and groundwater. Cobalt is most hazardous to the human respiratory system. Cobalt fumes can be generated when cutting or torching cobalt-containing metals. Cobalt mists occur when groundwater or process water containing dissolved cobalt is sprayed through a nozzle, such as a process washing system or a bathroom shower.

To reduce the mobility of cobalt dust from contaminated soil, the stockpiles are usually covered with clean soil, plastic sheeting, sod, or pavement. Because cobalt is mobile in an acidic medium, the upper layers of contaminated soil can be tilled with lime to reduce the pH and prevent toxicity to plants. The soil can also be covered with a thick layer of crushed limestone. The limestone layer will raise the pH in the vadose zone and prevent further leaching of the soluble cobalt.

Phytoaccumulation is one of the biological techniques used to remediate soils close to the surface contaminated by cobalt. Phytoaccumulation is done using nonfood crops to absorb metals from the soil into the stems and leaves. There are several plants that are hyperaccumulators; they can absorb large amounts of metals in their root structures (comparative to other plants and the concentration of metal in the ground). The cobalt is transported to the leaf and stalk biomass. Ultimately, the plant containing the cobalt is harvested for disposal or reclamation of the cobalt. There are 26 known plants that are hyperaccumulators of cobalt. There are also high accumulator plants that will take metal out of the soil in larger quantities than normal plants. Cobalt will exist as an oxalate compound in leaves, and as such it is readily extracted using accumulator plants. Accumulator plants are metal specific and will contain about 100 times more of a toxic element than normal plants. Phytoaccumulation can be enhanced using EDTA, citric, or oxalic acids on the soil (1).

For conventional remediation of cobalt in the groundwater, pump and treat or natural attenuation is used. At low pH, the cobalt is mobile. The contaminated water can be removed and treated with a strong base to reduce its solubility. Treatment with sodium hypochlorite (NaOCl) will cause precipitation of cobalt trihydroxide. The movement of cobalt-containing acidic plume within an aquifer matrix causes the pH to rise due to the presence of carbonate materials, which allows cobalt to be sorbed to the soil. High dissolved oxygen content will cause manganese oxides in the soil to precipitate and increase cobalt sorption, thus reducing the amount of soluble cobalt in the aquifer. Where monitored natural attenuation is the chosen remediation method, the down-gradient surface water body will be the location at which remediation is complete as the adsorbed cobalt will be remobilized with the migration of the plume. Only when the plume has traveled to a source with adequate dissolved oxygen will the iron and manganese oxides precipitate with sorbed cobalt.

Cobalt-contaminated groundwater can be treated *in situ* with a combination of methods that will allow the pH to be raised sufficiently to immobilize the cobalt and by providing enough dissolved oxygen to allow manganese oxides to precipitate and enhance cobalt sorption. Geochemical fixation using calcium polysulfide (pH 11.3–11.5) would significantly raise the pH in an aquifer. Extraction wells can be placed down-gradient to remove the acidic plume or in the middle.

BIBLIOGRAPHY

1. Suthersan, S.S. (2002). *Natural and Enhanced Remediation Systems*, CRC Press/Lewis Publishers, Boca Raton, FL.
2. Krupka, K.M. and Serne, R.J. (2002). *Geochemical Factors Affecting the Behavior of Antimony, Cobalt, Europium, Technetium, and Uranium in Vadose Sediments*, Report PNNL-14126; Pacific Northwest National Laboratory, December, Chap. 3; pp. 3.1–3.19.
3. Web Elements (WE) (2004). Available: http://www.web-elements.com/webelements/elements/text/Co/.
4. Ontario Ministry of the Environment (OME) (2001). Fact sheet: cobalt in the environment. Available: http://www.ene.gov.on.ca/cons/.
5. Swedish Environmental Protection Agency (SEPA) (2002). Assessment of contamination level. Available: http://www.internat.naturvardsverket.se/

GROUNDWATER AND CADMIUM: CHEMICAL BEHAVIOR AND TREATMENT

James A. Jacobs
Craig Foreman
Environmental Bio-Systems, Inc.
Mill Valley, California

Cadmium occurs naturally in ores containing other elements such as oxygen (cadmium oxide), sulfur (cadmium sulfate, cadmium sulfide), and chlorine (cadmium chloride) or compounds (cadmium carbonate). Cadmium is usually extracted during the production of other metals such as zinc, lead, and copper. A primary use for cadmium is in rechargeable nickel–cadmium and solar batteries, which are sold worldwide. Cadmium does not corrode easily and is also used in metal plating and coating operations including transportation equipment and various industrial machinery as well as plastics manufacturing. It is used in numerous pigments, in baking enamels, and in photographic processes as well as in television and monitor screen phosphors. Cadmium is released into the environment through industrial waste streams and landfills. Cadmium is relatively mobile in surface and groundwater and has a United States Environmental Protection Agency (U.S. EPA) Maximum Contaminant Level Goal (MCLG) of 5 parts per billion (ppb) in drinking water (1). The World Health Organization has set the provisional tolerable daily intake of cadmium at 60–70 μg/day (2).

PHYSICAL AND CHEMICAL CHARACTERISTICS

In ancient times cadmium was known as “calamine” and was often confused with zinc carbonate. In 1817, cadmium was discovered by the German chemist Friedrich Strohmeyer from an impurity in zinc carbonate ($ZnCO_3$), also called calamine. The impurity glowed with a yellow color and was identified as a new element, cadmium. Cadmium most often occurs in small quantities associated with zinc ores, such as sphalerite (ZnS). Greenockite (CdS) is the only mineral of any consequence bearing cadmium. Cadmium is a poisonous metal. The oxidation state of cadmium is +2.

The most common forms of cadmium are Cd^{2+}, cadmium–cyanide complexes, or cadmium hydroxide solid sludge (3). Cadmium hydroxide [$Cd(OH)_2$] and cadmium carbonate ($CdCo_3$) solids are most common at high pH, whereas the cadmium ion (Cd^{2+}) and aqueous cadmium sulfate species are the dominant forms at lower pH (<8). Under reducing conditions when sulfur is present, the stable solid cadmium sulfate [CdS(s)] is formed. Cadmium will also precipitate in the presence of phosphate, arsenate, chromate, and other anions, although solubility will vary with pH and other geochemical factors (4).

Naturally occurring cadmium is composed of 6 stable isotopes; 27 radioisotopes have been characterized with the most stable being Cd-113 with a half-life of 7.7 quadrillion years, Cd-109 with a half-life of 462.6 d, and Cd-115 with a half-life of 53.46 h. All of the remaining radioactive isotopes have half-lives that are less than 2.5 h and the majority of these have half-lives that are less than 5 min. This element also has 8 metastates with the most stable being Cdm-113 ($t_{1/2}$ 14.1 y), Cdm-115 ($t_{1/2}$ 44.6 d), and Cdm-117 ($t_{1/2}$ 3.36 h). The isotopes of cadmium range in atomic weight from 96.935 amu (Cd-97) to 129.934 amu (Cd-138).

Almost all cadmium is obtained as a by-product in the treatment of zinc, copper, and lead ores. It is a soft, bluish-white metal that is easily cut with a knife and is similar in many respects to zinc. It is a component of some of the lowest melting alloys and therefore is used in bearing alloys with low coefficients of friction and great resistance to fatigue. It is also used extensively in electroplating and in many types of solder, for standard EMF cells, for Ni–Cd rechargeable batteries, and as a barrier to control nuclear fission. Cadmium compounds are used in black and white television phosphorus and in blue and green phosphorus for color television and computer monitor tubes. It forms a number of salts, of which the sulfate is most common with the sulfide being used as a yellow pigment.

SOURCES OF CADMIUM

Environmental Exposure

Numerous human activities result in the release of significant quantities of cadmium to the environment. The relative importance of individual sources varies considerably from country to country. The major sources of anthropogenic cadmium release can be divided into three categories. The first is made up of those activities involved in the mining, production, and consumption of cadmium and other nonferrous metals. The second category consists of inadvertent sources where the metal is a natural constituent of the material being processed or consumed. Sources associated with the disposal of materials that had earlier received cadmium discharges or discarded cadmium products make up the third category. As a minor impurity in a variety of phosphate fertilizers, cadmium is introduced into the environment through direct application onto agricultural lands. The amount of cadmium impurities that are contained within phosphate fertilizers varies with the mining district.

Atmospheric Cadmium

The average global total emission of the metal from human sources in 1983 was 7570 tons (5) and represented about half the total quantity of cadmium produced in that year. In comparison, the worldwide emission of lead from human activities was about 10% of the total lead produced in 1983 (5). In both the European Economic Community (EEC) and on a worldwide scale (6), about 10–15% of total airborne cadmium emissions arise from natural processes, the major source being volcanic action. Nonferrous metal production, which accounts for about 75% of the total anthropogenic cadmium emissions worldwide, accounts for only 25% in the EEC. This is primarily a result of the extensive emission controls due to legislation and regulation in developed countries.

In addition, of the two basic methods of zinc production—thermal smelting and electrolyte refining—only the former releases significant atmospheric cadmium emissions. In recent years, electrolytic refining has assumed the major share of the world's production of zinc and cadmium and has largely replaced thermal processes in Europe. The once important vertical and horizontal retort smelters, which emit large quantities of atmospheric cadmium, have been phased out in most developed countries but are still in operation in several developing countries (7).

Other industries that employ thermal processes, for example, iron production, fossil fuel combustion, and cement manufacture, all release airborne cadmium, the metal being a natural component of the raw materials. The cadmium content of these materials is generally relatively low but due to the significant consumption of the raw material, significant cadmium is released into the environment. In addition, the elevated temperatures employed result in the volatilization of cadmium and release into the atmosphere. Municipal refuse is a common cadmium waste-related source, the cadmium being derived from discarded nickel–cadmium batteries and plastics that contain cadmium pigments and stabilizers. The incineration of refuse is also a major source of atmospheric cadmium release worldwide.

Sewage sludge receives cadmium from industrial sources, particularly from the discharges of plating operations and pigment works. One disposal option, the incineration of sewage sludge, is a relatively minor source of airborne cadmium, reflecting the small quantities of sludge disposed of in this manner.

Steel production can also be considered as a waste-related source, as large quantities of cadmium-plated steel scrap are recycled by this industry, at least in developed countries. As a result, steel production is responsible for considerable emissions of atmospheric cadmium.

Aquatic Cadmium

Nonferrous metal mines represent a major source of cadmium release to the aquatic environment. Contamination can arise from mine drainage water, wastewater from the processing of ores, overflow from the tailings pond, and rainwater runoff from the general mine area. The release of these effluents to local water sources can lead to extensive contamination of aquatic environments. The cadmium content of the ore body and climatic and geographical conditions, as well as mine management policies, all influence the quantities of cadmium released from individual sites. Aquatic inputs of cadmium are not restricted to active mine sites, and abandoned mining operations are also sources of continuing contamination of aquatic environments.

Cadmium is a natural constituent of rock phosphates and deposits from some regions of the world contain markedly elevated levels of the metal. The manufacture of phosphate fertilizer results in a redistribution of the cadmium in the rock phosphate between the phosphoric acid product and the gypsum waste. In many cases, the gypsum is disposed of by dumping in coastal waters, which leads to considerable cadmium inputs. Some countries, however, recover the gypsum for use as a construction material and thus have negligible cadmium discharges.

Acidification of soils and lakes may result in enhanced mobilization of cadmium from soils and sediments and lead to increased levels in surface and groundwaters (2). The corrosion of soldered joints or zinc-galvanized plumbing by acidic waters can dissolve cadmium and produce increased levels of the metal in drinking water.

Terrestrial Cadmium

Solid wastes from a variety of human activities are disposed of in landfill sites, resulting in large cadmium inputs at the national and regional levels when expressed as a total tonnage (8,9). Of greater potential environmental significance are the solid wastes from both nonferrous metal production and from the manufacture of cadmium-containing articles, as well as the ash residues from refuse incineration. All three waste materials are characterized by elevated cadmium levels and as such require disposal to controlled sites to prevent the mobilization of the cadmium in groundwater.

Cadmium soil contamination is a characteristic feature around nonferrous metal mines and smelters, particularly in the case of those handling zinc ores. Contamination from mining is generally local but may be widespread in areas of high mineral content. Shipham, England, in the United Kingdom, is a site of extreme soil cadmium contamination. Between 1650 and 1850 the village of Shipham was the site of a major zinc mine. Once the mining stopped, the area was flattened and developed for agriculture and housing. Cadmium levels in agricultural and garden soils are some of the highest ever reported worldwide (10).

The application of municipal sewage sludge to agricultural soil as a fertilizer can also be a significant source of cadmium. Large increases in soil cadmium concentration have resulted in the past from the application of contaminated sludge in both North America and Europe. Even today, the high application rates used for sewage sludge result in relatively large cadmium inputs, a value of 80 g/ha having been estimated for the United Kingdom (9). On a national or regional basis, however, these inputs are much smaller than those from either phosphate fertilizers or atmospheric deposition.

TERRESTRIAL IMPACTS

Large cadmium concentrations in terrestrial environments are commonly a result of solid wastes disposed of in landfill sites, ashes from fossil fuel combustion, waste from cement manufacture, and the disposal of municipal refuse and sewage sludge. The solid wastes from both nonferrous metal production and the manufacture of cadmium-containing articles, as well as the ash residues from refuse incineration, are the main contributors to cadmium contamination in terrestrial environments. These three waste materials are characterized by elevated cadmium levels and as such require disposal to controlled sites to prevent the contamination of the groundwater.

The agricultural application of phosphate fertilizers represents a direct input of cadmium to arable soils. The cadmium content of phosphate fertilizers varies widely and depends on the origin of the rock phosphate that is the main component in fertilizers, particularly in third world countries. The application of municipal sewage sludge to agricultural soils as a fertilizer can also be a significant source of cadmium.

Impact on Water Bodies

Rivers contaminated with cadmium can contaminate surrounding land, either through irrigation for agricultural purposes, by the dumping of dredged sediments, or through flooding. Cadmium compounds are able to leach through soils to groundwater. When cadmium compounds do bind to the sediments of rivers, they can be more easily bioaccumulated or redissolved when sediments are disturbed, such as during flooding. Its tendency to accumulate in aquatic life is great in some species. Much of the cadmium entering fresh waters from industrial sources is rapidly adsorbed by particulate matter, where it may settle out or remain suspended, depending on local conditions. This can result in low concentrations of dissolved cadmium even in rivers that receive and transport large quantities of the metal.

Health Risks

Cadmium and solutions of its compounds are toxic. In comparing the various standards available, exposure to cadmium dust is most common in workplaces, such as battery manufacturing plants, industrial facilities and metal fabricators having metal soldering and welding fumes, and coal plants. The cadmium dust levels should not exceed 0.01 mg/m^3 (8-h time-weighted average, 40-h week). The ceiling concentration (maximum), for a period of 15 min, should not exceed 0.14 mg/m^3. Cadmium oxide fume exposure (8-h day, 40-h week) should not exceed 0.05 mg/m^3, and the maximum concentration should not exceed 0.05 mg/m^3. Cadmium stays in the body for a very long time and concentrations in the body can build from years of low exposure (1).

Effects on Humans/Mammals. The United States Department of Health and Human Services (DHHS) has determined that cadmium and cadmium compounds may reasonably be anticipated to be carcinogens (1). Apart from the gastrointestinal tract and the lungs, the organ most susceptible to chronic cadmium exposure is the kidney. Cadmium is a cumulative poison; its ready solubility in weak acids is an important precondition for its absorption in the organism: 5% of cadmium is resorbed via the gastrointestinal tract and accumulated in liver and kidneys. In Asia, “Itai-Itai” disease is caused by high cadmium concentrations in rice.

Lung cancer is one potential result of chronic inhalation of fine-particle cadmium compounds, particularly cadmium oxide, which readily dissolves in the body. Ingestion of the compounds in food results from the fact that plants and fish absorb and retain cadmium, and hazardous levels can be reached when consuming food from areas where cadmium contamination has been generated at industrial sites or by the use of phosphate fertilizers. Adverse health

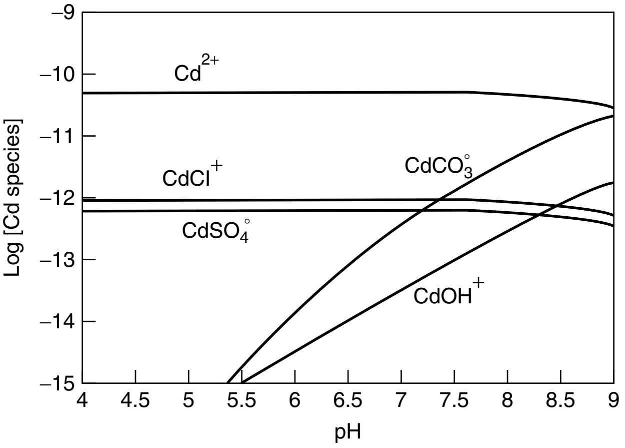

Figure 1. Inorganic cadmium speciation in soil water.

results from long-term, high level ingestion of cadmium include kidney damage and kidney stones, and debilitating effects on bones and the skeletal structure (1).

Effects on Plants. Cadmium reduces the photosynthesis and transpiration rate while increasing the respiratory rate. Small cadmium concentrations in the soil are sufficient to produce marked damage such as shortening of the shoot axis and intensive yellowing of older leaves. Absorption not only takes place via the roots but also through shoots and leaves. In addition to yield shortfalls, the main hazard results from the contamination of (accumulation in) crop plants since this is how cadmium as a cumulative poison enters the food chain.

GEOCHEMISTRY OF CADMIUM

In nature cadmium exists in the +2 oxidation state and forms a number of aqueous complexes, particularly with dissolved carbonate. The concentration of cadmium in the environment may be controlled by either adsorption or precipitation/coprecipitation processes. The concentration and extent to which cadmium is present in soils vary greatly with type of minerals, oxidation state of the system, and presence of competing cations in solution. Cadmium concentrations in uncontaminated soils are usually less than 1 mg/kg. However, higher concentrations may be present as a result of human activities or by the weathering of parent materials with high cadmium concentrations (i.e., cadmium-rich igneous rocks).

CADMIUM SPECIES AND CONDITIONS

Cadmium forms soluble complexes with inorganic and organic ligands resulting in an increase of cadmium mobility in soils. As the pH increases between 6 and 8.2, cadmium carbonate species [$CdHCO_3$ and $CdCO_3$ (aq)] assume greater significance. At pH values between 8.2 and 10, most of the cadmium in solution is expected to exist as the neutral complex $CdCO_3$(aq). The species $CdSO_4$(aq), $CdHCO_3$, $CdCl^{2+}$, and $CdOH^{+}$ are also present, but at much lower concentrations (Fig. 1).

CONDITIONS FOR THE DISSOLUTION AND PRECIPITATION OF CADMIUM

At pH values less than 7.5, most cadmium minerals are more soluble than cadmium concentrations found in oxic soils (10^{-7} M), indicating that cadmium at these concentrations is not likely to precipitate. At pH levels greater than 7.5, the solubilities of $Cd_3(PO_4)^2$ or $CdCO_3$ may control the concentrations of cadmium in soils. $CdCO_3$ precipitates in calcareous soils (pH $>$ 7.8), whereas in neutral or acidic soils, adsorption is the predominate process for removal of cadmium from solution. The activity of dissolved cadmium in cadmium-amended soils was lowest in calcareous soils. Cadmium may coprecipitate with calcium to form carbonate solid solutions, $(Ca,Cd)CO_3$. This may be an important mechanism in controlling cadmium concentrations in calcareous soils. Although cadmium itself is not sensitive to oxidation/reduction conditions, its concentration in the dissolved phase is generally very sensitive to redox state. The concentrations of dissolved cadmium greatly increase when reduced systems are oxidized, such as when dredged river sediments are land filled or rice paddies are drained.

The following mechanisms appear to be responsible for this increase in dissolved cadmium concentrations:

1. Very insoluble cadmium species in rocks or soils (e.g., greenockite) dissolve as sulfide [S(II)] that is oxidized to sulfate [S(VI)].
2. Organic materials binding cadmium are decomposed through oxidation, releasing cadmium into the environment. This mechanism appears to be important only in environments in which moderate to high organic matter concentrations are present. Dissolved cadmium concentrations greatly increased when the sediments were incubated under oxidizing conditions (Eh $>$ 350 mV).
3. Pyrites that may be present in soils or sediments get oxidized when exposed to air. The pyrite oxidizes to form $FeSO_4$, which generates high amounts of acidity when reacted with water. The decrease in the pH results in the dissolution of cadmium minerals and increase in the dissolved concentration of cadmium.

SORPTION AND DESORPTION OF CADMIUM IN SOILS

At high solution concentrations of cadmium ($>$10 mg/L), the adsorption of cadmium often correlates with the cation exchange capacity of the soil. During cation exchange, cadmium generally exchanges with adsorbed calcium and magnesium. At low solution concentrations of cadmium, surface complexation to calcite (11) and hydrous oxides of aluminum and iron (12) may be the most important adsorption mechanism. Both Cd^{2+} and possibly $CdOH^{+}$ may adsorb to aluminum and iron oxide minerals.

As with other cationic metals, cadmium adsorption exhibits pH dependency. The effect of pH on cadmium adsorption by soils, sediment, and iron oxides is influenced by the solution concentration of cadmium and the presence of competing cations or complexing ligands. In lower pH environments, these metals, based on their propensity to adsorb, would rank as follows: Pb > Cr > Zn > Cd. This order is inversely related to the pH at which hydrolysis of these metals occurs.

Competition between cations for adsorption sites strongly influences the adsorption behavior of cadmium. The presence of calcium, magnesium, and trace metal cations reduces cadmium adsorption by soils and manganese oxides and aluminum oxides. The extent of competition between cadmium and other ions depends on the relative energies of interaction between the ions and the adsorbing surface, the concentrations of the competing ions, and solution pH (12,13). The addition of copper or lead, which are more strongly adsorbed, slightly reduces cadmium adsorption by iron and aluminum oxides, suggesting that copper and lead are preferentially adsorbed by different surface sites (14). In contrast, zinc almost completely displaces cadmium.

FATE AND TRANSPORT

Cadmium can enter the environment in several ways. It can enter the air from the burning of coal and household waste and from metal mining and refining processes. It can enter water from disposal of wastewater from households or industries. Fertilizers often contain some cadmium, which results in cadmium entering the soil. Spills and leaks from hazardous waste sites can also result in excessive amounts of cadmium entering the soil or water. Cadmium may attach to small particles, which may be transported in the air as a result of an industrial process (i.e., metallurgy, zinc/lead smelting). Cadmium does not break down in the environment but may form complexes, ligands, and compounds that facilitate transportation of cadmium ions. Most cadmium stays where it enters the environment for a long time. Some of the cadmium that enters water will bind to soil but some will remain in the water. Cadmium in soil can enter water or be taken up by plants. Fish, plants, and animals take up cadmium from the environment.

Cadmium in Water

About 67–75% of the cadmium in surface water and groundwater is adsorbed on suspended matter. It can be remobilized from sediment by complexing agents. Aquatic life tends to be vulnerable to increases in cadmium concentration. Fish toxicity is dependent among other things on the calcium content of the water. Generally, high calcium content in water reduces the toxic effect of cadmium on fish. It has been noted that biological self-purification of surface water and groundwater is impaired when cadmium concentrations are in excess of 0.1 mg/L.

Cadmium in Air

Cadmium is stable in air since it forms an oxide layer.

Cadmium in the Soil

The cadmium retention capacity is a function of the number of exchangeable alkalis. There is hardly any dilution because of the adsorption on organic soil components. The accumulation horizon for cadmium is in the rhizosphere (roots). The minimum availability is at pH 6.5. Cadmium absorption by plants increases with a decreasing pH of the soil.

Cadmium in Organisms

Excessive cadmium concentrations in the environment will affect all organisms. In humans, 50% of the accumulated amount of cadmium is found in the liver and the kidneys. Cadmium is primarily excreted in urine [2 mg/d on average (0.2–3.1 mg/L)]. The biological half-life of cadmium in the human body is between 15 and 25 years (measured in the kidneys) (15). Humans consume roughly one-third of the cadmium to which they are exposed in animal foodstuffs and two-thirds in vegetable foodstuffs. Smoking significantly increases the cadmium impact.

METHODS OF REMEDIATION

There are several ways to deal with soil contaminated with toxic heavy metals such as cadmium. The following sections detail some of the more common remediation options for cadmium-contaminated soil and water.

Capping

Capping systems are used to provide an impermeable barrier to surface water infiltration to contaminated soil for prevention of further release of contaminants to the surrounding surface water or groundwater. A variety of materials are available for use in capping systems and choice of materials is site specific because local soils are often incorporated into parts of the cap. Synthetic membranes such as high density polyethylene are also available for incorporation into capping systems. Surface water controls, such as ditches and dikes, are usually included to help control drainage from the cap. Multilayered capping systems may also include a hard cover and/or a layer of topsoil to separate the underlying layers from the ground surface. Revegetation is promoted in order to reinforce the topsoil, to reduce soil erosion and runoff velocity, and to help remove water from the soil by evapotranspiration.

Subsurface Barriers: Pump and Treat

Subsurface barriers may be used to isolate contaminated soil and water by controlling the movement of groundwater at a contaminated site. These barriers are designed to reduce the movement of contaminated groundwater from the site, or to restrict the flow of uncontaminated groundwater through the contaminated site. Vertical barriers are commonly used to restrict the lateral flow of groundwater. For effective isolation of the contaminated matrix, the barrier should extend and key into a continuous, low permeability layer, such as clay or competent bedrock, below the contaminated area. If an impermeable layer is not available, a groundwater extraction system must be used to prevent transport of contaminants under the barrier and this contaminated water must be treated after extraction. Subsurface barriers are often implemented in conjunction with a capping system to control surface water infiltration. The use of circumferential barriers can prevent the escape of contamination from the site by using an infiltration barrier and collection system to create a hydraulic gradient in the inward direction.

Solidification, Stabilization, and Removal

The usual methods of dealing with soils contaminated with heavy metals are solidification, stabilization, and removal. The most commonly used method is solidification. Solidification involves a process in which the contaminated matrix is stabilized and fixated or encapsulated into a solid material by the addition of a chemical compound such as cement. The metal contaminants are chemically and/or physically bound to the matrix and are not obliterated but are contained in such a way that leaching into the environment is prevented or reduced. Removal is the process of physically removing the metal-contaminated soil from the current site and discarding the waste into a designated contaminated site. Stabilization, also referred to as fixation, usually uses a chemical reaction to convert the waste to a less mobile form. The general approach for solidification/stabilization treatment processes involves mixing or injecting treatment agents into the contaminated soils. Inorganic binders, such as cement, fly ash, or blast furnace slag, are commonly used. The dominant mechanism by which cadmium is immobilized is by precipitation of hydroxides or sulfides within the solid matrix.

ex situ Vitrification

The mobility of metal contaminants, particularly cadmium, can be decreased by high temperature treatment of the contaminated area, which results in the formation of vitreous material, usually an oxide solid. Most soils can be treated by vitrification with cadmium commonly being treated with *ex situ* vitrification. Typical stages in *ex situ* vitrification processes may include excavation, pretreatment, mixing, feeding, melting and vitrification, off-gas collection and treatment, and forming or casting of the melted product.

Bioremediation

Bioremediation is the use of microorganisms to control, precipitate, or destroy contaminants. Enhanced bioremediation of soil involves the delivery of oxygen for aerobic bioremediation or a carbon source (molasses, corn syrup, milk, cheese whey, lactic acid, propane, etc.) for anaerobic bioremediation. For cadmium, anaerobic and reducing conditions are more likely to contribute toward precipitating the cadmium as a sulfide or hydroxide than are oxidizing conditions. More information about metals stabilization and phytoremediation are described in Reference 16. For enhanced bioremediation, cadmium concentrations must be carefully evaluated so as not to be toxic to the microbial populations. Bench testing with the appropriate treatment chemicals can optimize pH, Eh, and other factors in the laboratory.

Phytoremediation

Phytoremediation is defined as the use of green plants to remove heavy metals and contaminants from the soil (16). The plants take up the toxic metals or isotopes through their roots and transport them to the stems or leaves. There are plants that tolerate excessive concentrations of metals in soils by absorbing, translocating, and storing the metals in a nontoxic way. Many of these plants have evolved on metal-rich soils. Research over the last two decades has shown that certain specialized plants have the ability to accumulate up to 3% (by dry weight) of heavy metals and up to 25% (by dry weight) in sap/latex with no apparent damage to the plant. Alpine pennycress (*Thlaspi caerulescens*) can be used for the removal of zinc and cadmium. Research has indicated that the pennycress plant typically absorbed 3–4% of the cadmium contamination from the soil (in some cases as much as 9%). Willow is another plant that proved able to absorb cadmium but never more than 1% of the soil content. Consequently, soil could advantageously be cleaned of cadmium by pennycress, possibly in combination with willow.

In Situ Electrokinetic Remediation

An emerging remediation technology, which in laboratory experiments has indicated substantial results in the removal of heavy metal contaminants particularly cadmium, is electrokinetic remediation (4). While bench tests conducted at the Naval Air Weapons Station at Pt. Mugu, California did not give statifactory results in the removal of cadmium and chromium contamination from the site, the technology holds some promise. Electrokinetic remediation is an *in situ* process in which an electrical field is created in a soil matrix by applying a low voltage direct current (DC) to electrodes placed in the soil.

The current density is generally on the order of milliamperes per square centimeter (mA/cm^2) or an electric potential difference on the order of a few volts per centimeter across electrodes placed in the ground. The electrodes can be placed in either a vertical or horizontal array. When direct current is applied to the electrodes, an electrical field develops between the anodes and cathodes. The application of the electric field has several effects on the soil, water, and contaminants. These effects include electromigration, electroosmosis, changes in pH, and electrophoresis. Electromigration refers to the movement of cations and anions under the influence of the electrical field. Cations (positively charged ions) tend to migrate toward the negatively charged cathode, and anions (negatively charged ions) migrate toward the positively charged anode. These ions concentrate in the solutions near the electrodes or may undergo reactions at the electrodes, which may plate the metals onto the electrodes or liberate gaseous compounds. The electrokinetic remediation system was designed to extract the metal contaminants via electromigration of the contaminant ions and concentration of the ions in the electrode wells. The majority of these contaminants were expected to remain in solution with only minimal plating occurring on the electrodes. Once metal contaminant concentrations had built up in the electrode well, processing of the electrolyte fluid would be conducted to extract the metal contaminants and return the processed electrolyte fluid back to the wells for continued use. This would have minimized the volume of waste generated by the process.

Precipitation/Flocculation/Sedimentation

Cadmium precipitation has been described in the literature (4). This process is the most common method of removing cadmium from soils. In precipitation reactions chemicals are added to transform dissolved cadmium constituents to form insoluble cadmium precipitates. Metals are precipitated as hydroxides, sulfides, and carbonates by adding appropriate precipitant and adjusting the pH to favor insolubility. Precipitation can be used to remove most metals and many anionic species (phosphates, sulfates, fluorides). Better removal efficiencies can be achieved with sulfite precipitation, but hydroxide precipitation, using lime or caustic, is more practiced, due to lower cost and better availability of the materials such as lime. This may change with newer chemicals such as calcium polysulfide (CaS_5). Precipitation reactions require rapid mixing, while flocculation reactions require slow, gentle mixing to encourage particle contact. After the contaminated product has reacted in the precipitation tank, it flows into the flocculation tank. In a flocculation reaction, alum, lime, ferric chloride, or polyelectrolytes are added to the inflow to reduce the repulsive forces between the precipitated particles. These particles aggregate, forming large flocs of material, which can be settled out in a sedimentation tank. For *in situ* methods, sulfur-based reducing chemicals such as calcium polysulfide, sodium metabisulfate, and ferrous sulfate can be used to precipitate cadmium and other metals. The resultant precipitate is cadmium sulfide (CdS).

CONCLUSIONS

Cadmium usage will continue to grow, especially in the area of rechargeable batteries and solar cells. The health effects of cadmium are well understood, so exposures are minimized through regulations. With recycling programs handling more spent batteries than in the past, cadmium emissions into the air and land should decline over time.

BIBLIOGRAPHY

1. ATSDR (1999). Agency for Toxic Substances and Disease Registry. *Toxological Profile for Cadmium,* U.S. Department of Health and Human Services, Public Health Service. Available: http://www.atsdr.cdc.gov/toxprofiles/tp5.html; updated July.
2. World Health Organization (WHO) (1992). *Environmental Health Criteria 134—Cadmium*, International Programme on Chemical Safety (IPCS) Monograph, Geneva, Switzerland.
3. Smith, L.A., Means, J.L., Chen, A., Alleman, B., Chapman, C.C., Tixier, J.S. Jr., Brauning, S.E., Gavaskar, A.R., and Royer, M.D. (1995). *Remedial Options for Metals–Contaminated Sites*. Lewis Publishers/CRC Press, Boca Raton, FL.
4. Evanko, C.R. and Dzombak, D.A. (2000). Remediation of metals-contaminated soil and groundwater. In: *McGraw-Hill Standard Handbook of Environmental Science, Health and Technology*, J. Lehr (Ed.). McGraw-Hill New York, pp. 14.100–14.134.
5. Nriagu, J.O. and Pacyna, J.M. (1988). Quantitative assessment of worldwide contamination of air, water and soils by trace metals. *Nature* **333**: 134–139.
6. Nriagu, J.O. (1979). Global inventory of natural and anthropogenic emissions of trace metals to the atmosphere. *Nature* **279**: 409.
7. International Lead and Zinc Study Group (ILZSG) (1988). *World Directory: Lead and Zinc Mines and Primary Metallurgical Works*. Chameleon Press, London.
8. Hutton, M., Wadge, A., and Milligan, P.J. (1988). Environmental levels of cadmium and lead in the vicinity of a major refuse incinerator. *Atmos. Environ.* **22**: 411–416.
9. Hutton, M. and Symon, C. (1986). The quantities of cadmium, lead, mercury and arsenic entering the United Kingdom environment from human activities. *Sci. Total Environ.* **57**: 129–150.
10. Thornton, I. (1988). The Shipham report. An investigation into cadmium contamination and its implications for human health. Metal content of soils and dusts. *Sci. Total Environ.* **75**: 21–39.
11. McBride, M.B. (1989). Reactions controlling heavy metal solubility in soils. *Adv. Soil Sci.* **10**: 1–56
12. Benjamin, M.M. and Leckie, J.O. (1981). Competitive adsorption of Cd, Cu, Zn, and Pb on amorphous iron oxyhydroxide. *J. Colloid Interface Sci.* **79**: 410–419.
13. Sposito, G. (1984). *The Surface Chemistry of Soils*. Oxford University Press, Oxford.
14. Benjamin, M.M. and Leckie, J.O. (1980). Adsorption of metals at oxide interface: effects of the concentrations of adsorbate and competing metals. In: *Contaminants and Sediments*. Vol. 2. R.A. Baker (Ed.). Ann Arbor Science Publishers, Ann Arbor, MI, pp. 305–322.
15. Großklaus, X. (1989).
16. Suthersan, S. (2002). *Natural and Enhanced Remediation Systems*, Lewis Publishers/CRC Press, Boca Raton, FL.

READING LIST

United States Environmental Protection Agency (USEPA), 2004, EPA Ground Water & Drinking Water Consumer Factsheet on Cadmium, http://www.epa.gov/OGWDW/index.html.

World Health Organization (WHO), (1984) Guidelines for drinking-water quality: Health criteria and other supporting information, Geneva, World Health Organization, Vol. 2, pp. 84–90.

GROUNDWATER MODELING

Ekkehard Holzbecher
Humboldt Universität Berlin
Berlin, Germany

INTRODUCTION

Since the early beginnings with applications for the oil industry (1), computer models have become an indispensable tool to study aquifers, to understand the interaction of different processes affecting the biogeochemistry and/or the heat distribution, to predict the effect of changes, and to solve practical groundwater problems. Examples are aquifer characterization, capture zone delineation, pumping and recharge well design and management, watershed simulation, groundwater pollution, potential hazards and remediation, acid mine drainage, natural attenuation, geo- and hydrothermics, and saltwater intrusion.

Computer software codes are available today for the modeler to set up groundwater models. Some codes are equipped with user-friendly interfaces and can be used for simulations of flow, as well as mass and heat transport. Some are even capable of including reactions between various chemical or biochemical species. Only a few specific programs can be mentioned here in detail. An extended description of codes is given elsewhere (2,3). Some topics have to be left out also. Multiphase flow is not considered, and modeling of sorption and biogeochemistry could only be mentioned in the context of reactive transport, just to list some topics.

DIFFERENTIAL EQUATIONS

A quantitative description of groundwater flow, transport, and reaction processes is obtained via differential equations. In three dimensions flow equations can be formulated as

$$S\frac{\partial h}{\partial t} = \left\{ \begin{array}{ll} \nabla \cdot T\,\nabla h & \text{for the confined aquifer} \\ \nabla \cdot \mathbf{K}h\,\nabla h & \text{for the unconfined aquifer} \end{array} \right\} + Q \tag{1}$$

with storativity S, transmissivity T, and the tensor of hydraulic conductivities $\mathbf{K}$. The tensor $\mathbf{K}$ allows us to take into account the direction-dependent conductivity. In the case of an isotropic porous medium, the tensor reduces to a scalar conductivity value K, which may change from one location to the other in order to consider different hydrogeological strata or inhomogeneities. In the term Q all types of sinks and sources are gathered. The value of Q is positive for sources of fluid and negative for sinks. Unknown is the variable of piezometric head h. Sometimes pressure p is used as an unknown variable, which is the more appropriate variable in cases in which water density changes (4).

The flow differential equations are derived from the principle of water mass conservation and from Darcy's law (2). The differential equations for the steady state are obtained when the storage term on the left side of differential equation 1 is omitted. More details are given elsewhere (5–7).

The general differential equation for mass transport in three-dimensional space is

$$\frac{\partial}{\partial t}(\phi c) = \nabla\phi\,\mathbf{D}\cdot\nabla c - \nabla(\mathbf{v}c) + q \tag{2}$$

with porosity ϕ, the vector of Darcy velocity $\mathbf{v}$, and the dispersion tensor $\mathbf{D}$. With the dispersion tensor $\mathbf{D}$ it is possible to account for direction-dependent dispersive processes, as they are mostly observed in porous media. The dispersion tensor depends on the parameters of molecular diffusivity D_m and longitudinal as well as transversal dispersivity, α_L and α_T. The term on the left side of the equation represents storage, the first two terms on the right side dispersion and advection. In the term q different additional processes can be taken into account, like decay, degradation, or production. Unknown is the concentration c of a chemical species. For more details see References 4, 8, and 9. In order to describe multispecies systems, the transport equations for the different species are combined by reaction terms:

$$\frac{\partial}{\partial t}(\phi\mathbf{c}) = \nabla\phi\,\mathbf{D}\cdot\nabla\mathbf{c} - \nabla(\mathbf{v}\mathbf{c}) + \mathbf{S}^T\mathbf{c} + q \tag{3}$$

$\mathbf{S}$ denotes the reaction matrix, in which all equilibrium reactions are gathered. In the given notation all unknown concentrations are gathered in a vector $\mathbf{c}$. In order to handle the system described by Equation 3, the equations are rearranged by linear operations, which makes the reaction term $\mathbf{S}^T\mathbf{c}$ vanish. In addition to the remaining transport equations, the equilibrium conditions for the reaction have to be taken into account (10,11). Sorption can be treated similarly, but instead of solving the equilibrium state equation directly, often a retardation parameter R is used as an additional coefficient on the left side of the transport equation (see References 2 and 3).

The heat transport equation is given by

$$(\rho c)^*\frac{\partial}{\partial t}T = \nabla\cdot\lambda\,\nabla T - (\rho c)_f\mathbf{v}\cdot\nabla T + H \tag{4}$$

with heat capacity $(\rho c)_f$ of the fluid as well as thermal conductivity λ and heat capacity $(\rho c)^*$ of the saturated porous medium (2,4,12). In the term H, all types of heat sources can be gathered. Sometimes direction-dependent dispersion processes are taken into account in heat transport; then a tensor has to be used, as in the mass transport equation 2. Unknown in Equation 4 is the temperature T.

Table 1 provides an overview of the types of processes and equations, their unknown variables, and parameters. While in the simplest case a model can be related uniquely to one of the three given types, some model concepts are complex. Transport models are often based on a flow model, which calculates the necessary velocity field $\mathbf{v}$. Even more complex is the coupled situation, when one of the transport variables has an influence on flow. The most prominent example is density-driven flow, in which the effect of salinity and/or temperature on the density is taken into account; see Reference 4. The complexity increases further when, for example, the effect of temperature on kinetic reaction rates in mass transport is to be considered.

BOUNDARY AND INITIAL CONDITIONS

Before the computer is engaged, the first task of the modeler is to specify the *model region*. The model region can be one, two, or three dimensional, depending on the situation to be modeled. The extent of the model region, more specifically its boundaries, has to be chosen appropriately, in order to formulate *boundary conditions*.

Table 1. Overview of Types of Models, Unknown Variables, and Parameters

Type	Unknown Variable	Parameters[a]
Flow	Head, h, or pressure, p	S, T, K, Q, ρ
Mass transport	Concentration(s), c	$\phi, D_m, \alpha_L, \alpha_T, q, R$
Heat transport	Temperature, T	$(\rho c)^*$, $(\rho c)_f$, λ, H

[a] See text for description of parameters. The list of parameters is not complete; neither are all of them necessary in all cases.

The boundary of a one-dimensional model consist of two points, the boundary of a two-dimensional area consists of closed lines, and the boundary of a three-dimensional volume is given by surfaces.

In order to complete the mathematical formulation of the problem, boundary conditions have to be formulated. At the boundaries additional information about the behavior of the system is necessary, as the variation of the unknown variable is free outside the model region, while in the interior it is fixed in all directions by the differential equation. There are three major types of boundary conditions, which are listed in Table 2.

Some typical examples for boundary conditions in field situations are given as illustration. When a river or a lake is perfectly connected to groundwater (no clogging layer), the water table of the surface water body is taken as a first-type boundary condition for flow (see Fig. 1); the concentration or temperature of the surface water can be used as first-type boundary conditions for mass or heat transport. Interfaces to nearly impermeable strata in the subsurface are also appropriate boundaries: Neumann conditions are specified for both flow and mass transport, representing no-flow for water and solute. For transport the Neumann conditions represent no dispersive flux, but in combination with a no-flow condition for water there is no advective flux across that boundary either. More complex types of boundary conditions, in which heads and fluxes are combined, are described elsewhere (7,13).

In transient simulations, that is, when the system changes with time, *initial conditions* for the unknown variable have to be specified by the modeler additionally. The task, in mathematical terms, is to solve the *initial and boundary problem* for the differential equation. For this task analytical solutions are sought or numerical models are applied.

Analytical solutions are mostly given by an explicit equation, which is valid for all locations in the model region and all time instants (6). For the evaluation and illustration of such solutions usually a computer is used. Finally, the analytical methods also depend on discretization, that is, the use of discrete time and space values instead of a continuum. But in *numerical methods* the discretization is already applied in the solution process.

Table 2. Overview of Types of Boundary Conditions

Type	Name	Known Value, to be Specified as Condition
First	Dirichlet	Unknown variable
Second	Neumann	Derivative of unknown variable normal to the boundary
Third	Cauchy or Robin	Combination of unknown variable and its derivative

DISCRETIZATION

The major step in groundwater modeling by numerical methods is discretization. The term discretization means that an aquifer or a part of it is divided into a *grid* of small pieces, which are called blocks, volumes, elements, or cells, depending on the *method of discretization*; see below. In one-dimensional models a grid size of 1000 is already a fine model; in two-dimensional cases such a grid size is still small. In three-dimensional models grids containing more than a million unknowns are not uncommon nowadays.

Grids may look very different. A grid may consist of regular rows and columns of blocks: such a regular grid is shown in Fig. 1. It may consist of irregularly shaped triangles as shown in the upper left part of Fig. 3. In fact, these are the major types of grids, connected to the methods of finite differences (FD) and finite elements (FE), but other forms are possible also: irregular rectangular grids, regular triangular grids, grids containing triangles, and other forms of geometries (14). The term grid is not only valid in two dimensions: it consists of intervals in one dimension and prisms or tetrahedrons in three dimensions.

All discretization methods are constructed to calculate approximate solutions for the initial and boundary value problem. Finer grids usually guarantee better approximations. On the other hand, finer grids require more computer space and much longer execution time on the computer. Based on available computer facilities and experience, the modeler usually decides on the grid resolution. Only advanced software tools provide the user with the ability to check the model accuracy by adaptive grids (15).

There are different numerical discretization methods, by which computer algorithms are derived from differential equations. Most important are *finite differences* (FD), *finite elements* (FE), and *finite volumes* (FV).

Using the FD method, derivatives (differential quotients) in the differential equations are approximated by difference quotients. For first and second order derivatives in the x-direction, the simplest central stencils are given by

$$\frac{\partial f}{\partial x} \approx \frac{f_{i+1} - f_{i-1}}{2\Delta x} \quad \text{and} \quad \frac{\partial^2 f}{\partial x^2} \approx \frac{f_{i+1} - 2f_i + f_{i-1}}{\Delta x^2} \tag{5}$$

where the f-values denote function values at the grid nodes: that is, f_i is the approximate value of the function at node i, f_{i-1} at the previous node, and f_{i+1} at the following node. In that way, the differential equation delivers a condition for each block or node of the grid; altogether a system of equations is obtained for the values of unknown functions at the nodes.

Using the FE technique, the solution of the differential equation is approximated by a combination of shape

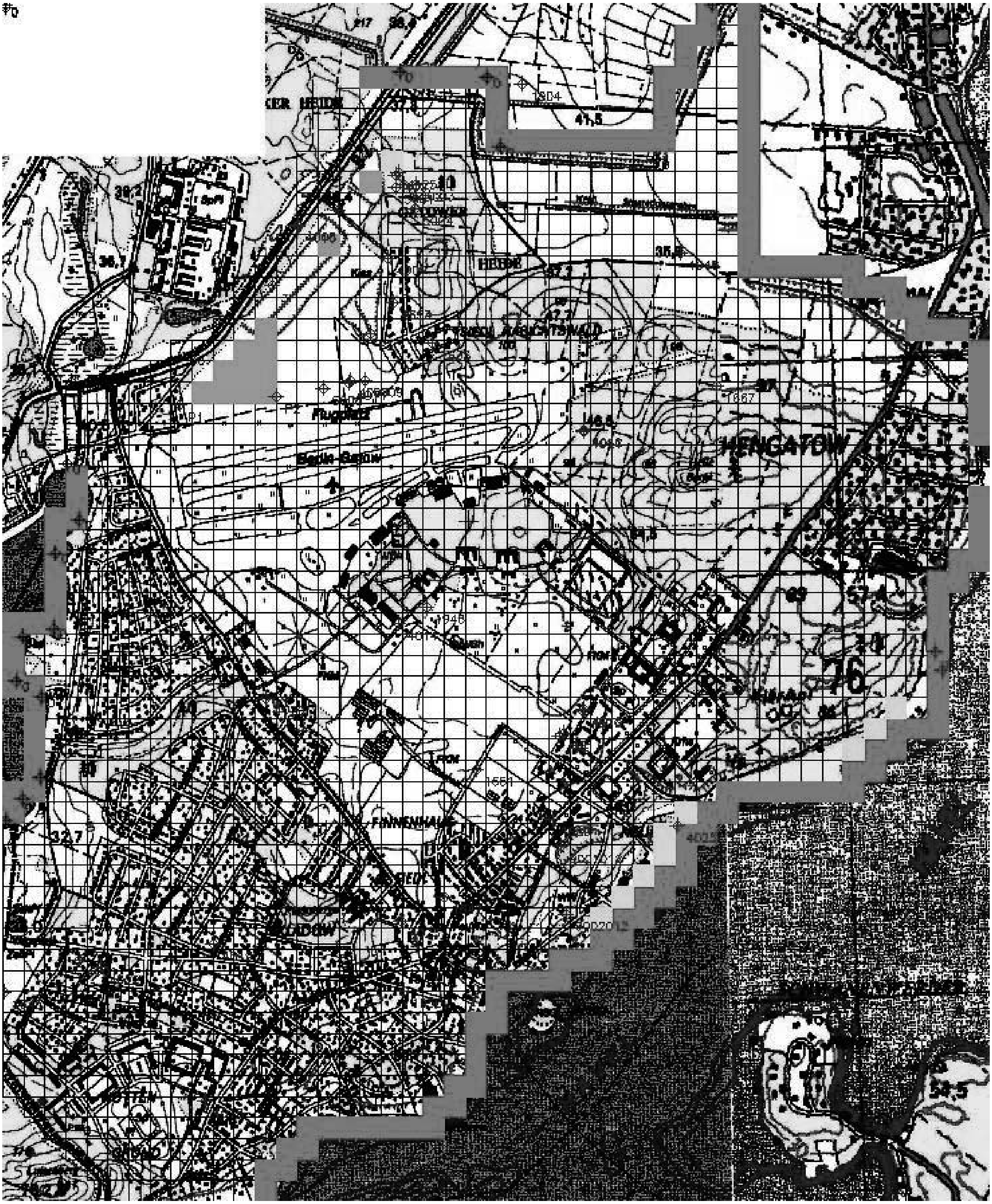

Figure 1. Regular finite volume grid for a horizontal model region. A map is depicted, underlying a regular FV grid, from which only active blocks are shown; boundaries of the model region are shorelines of lakes (west and east) and creeks (north), a water divide (northwest), and a streamline (southwest). Blocks with first-type boundary condition are dark colored pumping wells are light colored, and abandoned landfills as sources of contaminants are gray.

functions, which depend on the element type. For the most common triangular elements, the prescribed shape functions are linear within each element: in the two-dimensional situation the unknown function $f(x, y)$ is approximated by

$$f(x, y) = a_{\alpha 0} + a_{\alpha 1}x + a_{\alpha 2}y \quad \text{within element } \alpha \tag{6}$$

The so-called weak form of the differential equation delivers conditions for the coefficients $a_{\alpha j}$ for all elements in the form of a linear or nonlinear system (14). FEFLOW (16) is a popular FE groundwater code, using prisms for single layers, while in the mathematical FEMLAB code (14) more general tetrahedrons can be used.

The method of finite volumes (FV) is derived from a mass or energy balance for all volumes of the grid. In the two-dimensional rectangular volume *ij*, for example, the balance can be noted as

$$\frac{\partial V}{\partial t} = Q_{i-} + Q_{i+} + Q_{j-} + Q_{j+} + Q \tag{7}$$

where V denotes the volume or mass in the block, Q_{i-}, Q_{i+}, Q_{j-}, Q_{j+} the fluxes across the block edges, and Q other sources or sinks for water volume, mass, or energy in the volume. For further illustration of the technique, the flow equation 1 is considered: replacing the four flux terms in Equation 7 by using Darcy's law leads to a system of equations for the heads at the block centers. The popular MODFLOW code (13) is implemented following this strategy. Under certain conditions the FV technique delivers the same algorithm as the FD method.

The discretization of time is mostly implemented by a time-stepping approach; that is, starting from the initial condition the numerical simulation proceeds in time steps. For each time step, the discretized form of the left side of the differential equation is solved.

Concerning specific modeling situations, the basic techniques may deliver poor results; that is, they may produce bad approximations or results that do not resemble the solution at all, or they may even not converge. A very prominent example problem is transport in advection-dominated flow, for which standard methods deliver unrealistic oscillations or numerical diffusion (3,14). For this specific problem *operator splitting*, using different discretization methods for the different terms on the right side of the transport equation 2, has turned out to be a successful strategy to obtain accurate results even for small time steps (see References 2, 3, and 11).

Independent of the discretization method, a system of equations for the values of the unknown functions at the grid nodes is obtained. The task to find an approximation for the solution of the differential equation is thus transformed to the task of solving a linear or nonlinear system of equations. For the latter task, various solution techniques have been developed in numerical mathematics, which are included in modeling software tools.

PREPROCESSING, SOFTWARE, AND SOLUTION

The entire task of modeling can be divided into three steps: preprocessing, processing, and postprocessing. Processing means the execution of the mathematical solution algorithm for the system of equations (the *solver*), for which different alternatives are available. Taking into account several input parameters, in a model run the solver delivers values for the unknown variable in question on the grid. Depending on the discretization method, values are calculated at the nodes of the grid or at block centers.

Input parameters are prepared during preprocessing. Common software tools allow different preprocessing tasks, concerning the specification of input data. Data fields of distributed parameters can be calculated. For these calculations the user interfaces are equipped with interpolation tools, varying from simple nearest neighbor to complex krigging schemes.

The solvers are special numerical algorithms for the solution of large systems of linear or nonlinear equations. Several solvers and variants are included in most modeling codes. The default values for these solvers (often of conjugate gradient type) are usually appropriate for a wide range of model setups, but sometimes it may be necessary to change a solver parameter or switch to another solver type, in order to obtain a correct and fast solution.

User-friendly shells often not only offer the alternative to use different solvers of one processing program, but allow the user to choose different programs, for which the input data are prepared automatically. Visual MODFLOW (17), GMS (18), and PMWIN (19) are such shells, which are currently used. Operations with the values, which have been delivered by the solver, are gathered under postprocessing, which is dealt with next.

POSTPROCESSING: RESULTS AND VISUALIZATION

There are various ways to visualize large numerical data sets, delivered by a solver, in a form that provides direct insight to the hydrogeological situation. The most common output are head contours, shown as a surface plot, as a pattern of contours with or without fill pattern (see Fig. 2). When several layers of aquifers are modeled, contour plots are generally different from one layer to the other. In three-dimensional models, contour plots can be shown for layers and also for cross sections.

Very informative are flowpaths or flowlines, that is, the travelpaths of particles in the calculated flow field. Start positions have to be chosen by the modeler, and the trace of the particles is followed by a numerical algorithm (2,20). The direction can be either forward or backward in time: the path is traced forward to future positions or back to its origin. Traveltimes can be visualized as so-called isochrones, lines of equal traveltime. In transient simulations, the flowpath usually changes with time, while in steady-state flow the flowpaths are constant. In the latter case, one may speak of streamlines. In three-dimensional models it may be sufficient to depict flowpaths in one picture, in which crossovers are allowed (Fig. 2). Some modeling software delivers the very instructive three-in-one plots with one view from the top and two views in vertical cross sections, as in PMPATH (19).

With streamline plots, it is possible to visualize not only flow direction but also velocities. When streamline

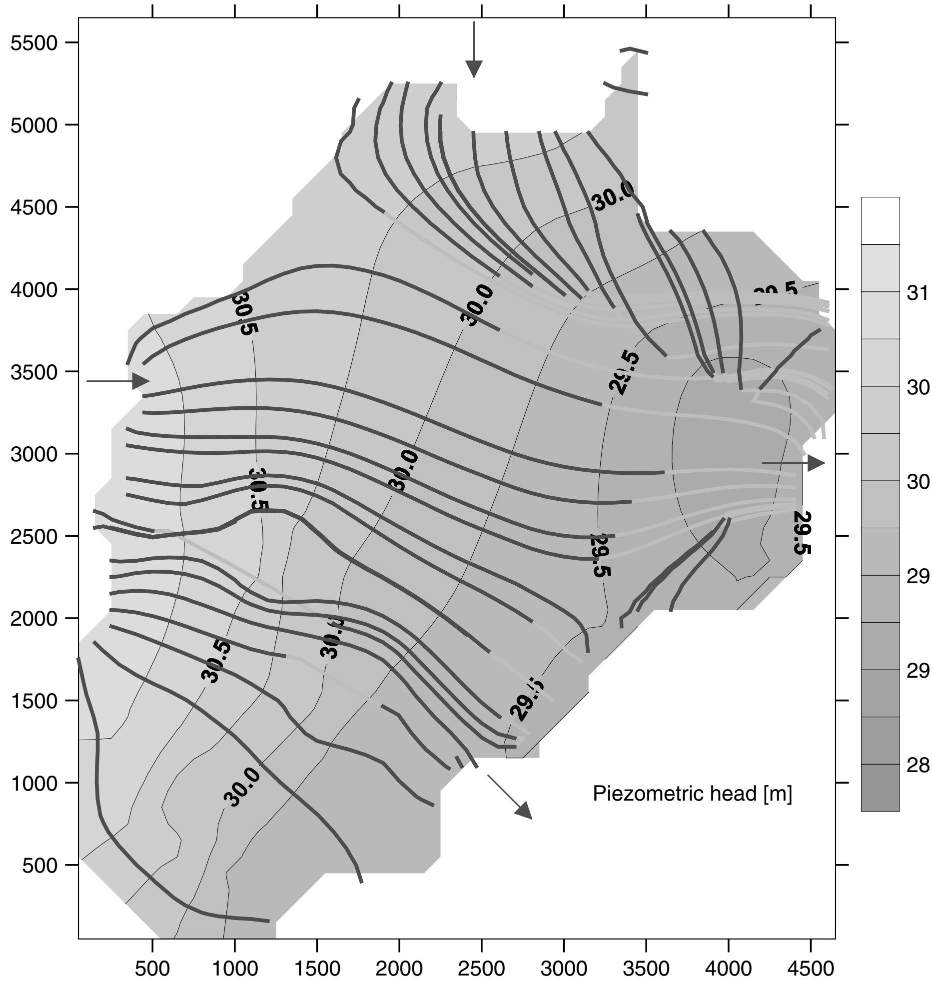

Figure 2. Head contour and flow-path plot, calculated using MODFLOW (13) and PMWIN (19).

levels are equidistant, streamlines become dense, where velocities are high; as the distance widens, the velocities are low. See the bottom left of Fig. 3 as an example.

Vector plots of velocities are also informative, as depicted in the top right of Fig. 3. Not only is the direction shown by the vectors, but the velocity can also be visualized by scaling the length of the vector with the size of the velocity. A problem with such a representation is that in most groundwater problems velocities change by several orders of magnitude within the model region; that is, some vectors become very large, while others are hard to recognize. The alternative is to scale the length of the vectors with the logarithm of the velocity.

In transport models, distributions of concentration or heat are depicted by plots of filled or unfilled contours, representing isohalines in the case of salt and isotherms in the case of temperature. An example for a steady-state concentration is given in the bottom right of Fig. 3. The transient behavior at certain locations of interest within the model region can be captured by breakthrough curves, as implemented in PMWIN (19).

Another important postprocessing tool is the calculation of a budget, for example, of fluid mass for flow models, of component mass in the case of mass transport, and of energy in the case of heat transport. Such budgets are extremely useful for the entire model region or for parts of it. Some programs allow a graphical representation of the different budget terms in a histogram (16).

PARAMETER ESTIMATION

There is an increasing demand for and use of parameter estimation in groundwater models. While the usual modeling task, as described so far, is termed *direct modeling*, parameter estimation is performed by *inverse modeling*. While in direct modeling of groundwater flow (for example) hydraulic conductivity (for example) is an input variable and piezometric head is the unknown variable, their role is exchanged in inverse modeling. However, direct modeling is part of the inverse modeling task.

After each direct modeling run, the performance of the model is evaluated. The performance measure usually is the mean deviation between measured and modeled values. Special algorithms are applied to change the input data set in a way that is likely to increase the performance. Levenberg-Marquard and Gauss-Newton are algorithms of that type, which are applied in all fields of parameter estimation. Various programs offer different types of assistance to the modeler to perform inverse

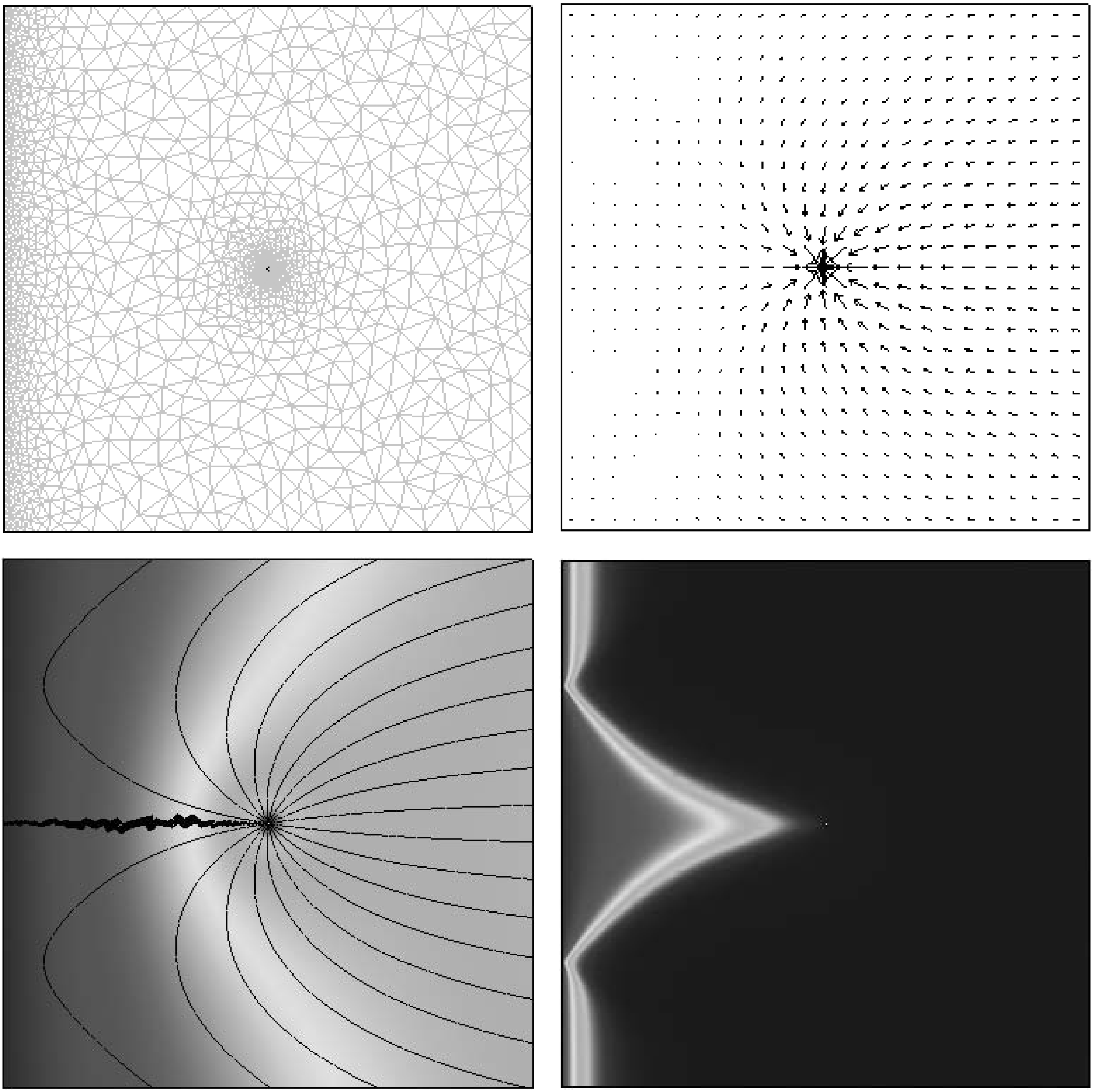

Figure 3. Finite element grid (top left), velocity vector plot (top right), head and streamline plot (bottom left), and concentration plot (bottom right) for bank filtration. The bank is located on the left boundary, the well in the middle of the depicted model region; water pumped at the well is a mixture of bank filtrate and groundwater. Simulated and visualized by FEMLAB (15).

modeling (16,19). In most cases the groundwater models are connected with general-purpose parameter estimation codes, like PEST (21) or UCODE (22).

CONCLUSION

Groundwater modeling has become a broad field in the few decades since its early beginnings. Only major features could be mentioned in a short introduction. Available software tools vary in many aspects and are in permanent development. The treated topics can thus be only a guideline for users searching for appropriate tools for specific applications. The reader may consult References 2 and 3 for more details.

Acknowledgments
The author is grateful to NASRI, an interdisciplinary project of the Kompetenzzentrum Berlin, funded by Veolia Water and Berliner Wasserbetriebe; and to KORA, an interdisciplinary project, funded by the German Ministry for Education and Science (BMBF).

BIBLIOGRAPHY

1. Peaceman, D.W. (1977). *Fundamentals of Numerical Reservoir Simulation*. Elsevier, Amsterdam.
2. Holzbecher, E. (2002). *Groundwater Modeling—Computer Simulation of Groundwater Flow and Pollution*, FiatLux Publ., Fremont. CD-ROM. Available: http://envirocomp.org.
3. Holzbecher, E. and Sorek, S. (2004). Numerical models of groundwater flow and transport. In: *Encyclopedia of Hydrological Sciences*. M. Anderson (Ed.). John Wiley & Sons, Hoboken, NJ.
4. Holzbecher, E. (1998). *Modeling Density-Driven Flow in Porous Media*. Springer, Heidelberg.
5. Bear, J. and Verruijt, A. (1987). *Modeling Groundwater Flow and Pollution*. Reidel Publ., Dordrecht.
6. Strack, O.D.L. (1989). *Groundwater Mechanics*. Prentice Hall, Englewood Cliffs, NJ.
7. Anderson, M.P. and Woessner, W.W. (1991). *Applied Groundwater Modeling*. Academic Press, San Diego, CA.
8. Zheng, C. and Bennett, G.D. (1995). *Applied Contaminant Transport Modeling*. Van Norstrand Reinhold, New York.
9. Sun N.-Z. (1995). *Mathematical Modeling of Groundwater Pollution*. Springer, New York.
10. Saaltink, M.W., Ayora, C., and Carrera, J. (1998). A mathematical formulation for reactive transport that eliminates mineral concentrations. *Water Resource Res.* **34**(7): 1649–1656.
11. Holzbecher, E. (2005). Reactive transport—concepts and numerical approaches. In: *Transport Phenomena in Porous Media III*. D.B. Ingham and I. Pop (Eds.). Elsevier, Dordrecht.
12. de Marsily, G. (1986). *Quantitative Hydrogeology*. Academic Press, San Diego, CA.
13. Harbaugh, A.W., Banta, E.R., Hill, M.C., and McDonald, M.G. (2000). *MODFLOW-2000, the U.S. Geological Survey Modular Ground-Water Model—User Guide to Modularization Concepts and the Ground-Water Flow Process*. U.S. Geological Survey Open-File Report 00-92.
14. Huyakorn, P.S. and Pinder, G. (1983). *Computational Methods in Subsurface Flow*, Academic Press, New York.
15. FEMLAB, *Version 3.1, COMSOL AB, Tegnérgatan 23*, SE-111 40 Stockholm, Sweden.
16. FEFLOW, *WASY Inst. for Res. Planning and Systems Res.*, Waltersdorfer Str. 105, D-12526 Berlin, Germany.
17. Visual MODFLOW, *Waterloo Hydrogeologic, Inc.*, 460 Phillip Street, Suite 101, Waterloo, Ontario, N2L 5J2, Canada.
18. GMS, *Environmental Modeling Systems, Inc.*, 1204 W. South Jordan Pkwy, Suite B, South Jordan, UT 84095, USA.
19. Chiang, W.-H. and Kinzelbach, W. (2003). *3D-Groundwater Modeling with PMWIN—A Simulation System for Modeling Groundwater Flow and Pollution*, p. 346.
20. Pollock, D.W. (1988). Semianalytical computation of path lines for finite difference models. *Groundwater* **26**(6): 743–750.
21. Doherty, J. (2002). *PEST—Model Independent Parameter Estimation*. Watermark Numerical Computing, Corinda, Australia.
22. Poeter, E.P. and Hill, M.C. (1998). *Documentation of UCODE, a Computer Code for Universal Inverse Modeling*. U.S. Geological Survey Water-Resources Investigations Report 98-4080.

GROUNDWATER AND BENZENE: CHEMICAL BEHAVIOR AND TREATMENT

James A. Jacobs
Keith Villiers
Environmental Bio-Systems, Inc.
Mill Valley, California

Benzene (C_6H_6) is a clear, colorless, highly flammable aromatic hydrocarbon liquid. As an important ingredient for making plastics, rubber, resins, and synthetic fabrics like nylon and polyester, benzene is one of the most commonly used industrial chemicals in the world. It is also used as a solvent in printing, paints, and dry cleaning. It is also a major component of gasoline. The wide ranging industrial applications of benzene make it widely distributed, particularly in the case of petroleum products, of which it is a primary constituent. Benzene is highly toxic to humans and the environment. Because of its severe toxic effects, both in the short and long term, benzene use has been under strict regulations for many years. Nonetheless, widespread contamination of benzene has occurred in the environment, and, therefore, significant regulatory efforts in the United States and other nations are being enacted to limit or eliminate potential impacts.

HISTORY

In 1825, Michael Faraday began a number of experiments in distillation. Among the products he worked on was crude oil, from which he extracted a colorless gas-phase compound. Faraday lit the gas compound and it burned. For decades the benefits of what Faraday called "bicarburet of hydrogen" were unknown. In 1834 Eilhard Mitscherlich isolated the same gas-phase compound by heating benzene carboxylic acid ($C_6H_5CO_2H$), also called benzoic acid, with lime (CaO). He derived the chemical formula C_6H_6 and named the compound "benzin." In 1846, A. W. Hoffman isolated the same product, later called benzene, in large quantities by distilling coal (1).

CHEMICAL STRUCTURE

The molecular structure of benzene baffled scientists for many years, although the empirical formula was known. None of the proposed molecular structures at the time supported the chemical properties observed. In 1865, Friedrich August Kelule proposed a ring structure. It is now known that benzene, like all aromatic hydrocarbons, contains a cyclohexane type of ring structure that has alternating single and double bonds.

This unique molecular structure of benzene (C_6H_6), a volatile, colorless, and flammable hydrocarbon, made it incompatible or reactive with strong oxidizers, many fluorides, perchlorate, and nitric acid. It is slightly soluble in water [580–1780 ppm (∼0.06–0.17%)] and has a boiling point of 80 °C and a freezing point of 5.5 °C. Benzene has a characteristic sweet odor, with an odor threshold of 2.7–12 ppm in air.

PRODUCTION AND USES

For a long time, benzene could only be extracted from coal; however, with the rise of crude oil use, its importance increased significantly. From the start of the twentieth century, benzene saw extraordinary growth in its applications in an abundance of industrial areas, including chemical, petroleum, agricultural, and pharmaceutical manufacturing. Benzene has been given a list of trade names by product producers, including Benzol 90, Pyrobenzol, Polystream, Coal naptha, and Phene (2).

Benzene first saw major industrial production during World War I, where it was produced specifically for the rubber industry as well as for the production of toluene, a major component of explosives. Eventually, benzene began finding uses in other industrial processes such as the manufacture of artificial leather, rubber goods, and rotogravure printing industries, and as a starting material in organic synthesis.

Benzene has been released into the air primarily from fumes and exhaust through incomplete vehicular combustion. Other air sources of benzene include the release of fumes from the production and use in manufacturing of other chemicals (2). Vapor emissions of benzene can be detected from carpet glue, textured carpet, and furniture wax (3).

Benzene is also an important component of gasoline. Benzene has been released into the subsurface through leaking fuel tanks, primarily gasoline. Discharges into water come from industrial effluents and losses during spills (2).

HAZARDOUS EFFECTS

Benzene has been recognized since 1900 as a toxic substance capable of producing acute and chronic nonmalignant effects in humans. When benzene vapors are inhaled, the benzene diffuses rapidly through the lungs and is quickly absorbed into the blood. Acute circulatory failure resulting in death within minutes often accompanies exposure to benzene concentrations as high as 20,000 ppm. Other acute effects of exposure to milder, though still high (250–500 ppm), concentrations of benzene include vertigo, nervous excitation, headache, nausea, and breathlessness. As a result of its toxicity, benzene's history has been one of regulation. In 1946, the American Conference of Governmental Industrial Hygienists recommended a threshold limit value for benzene exposure of 100 ppm. This value was reduced to 50 ppm in 1947, to 35 ppm in 1948, to 25 ppm in 1963, and to 10 ppm in 1974. The American National Standards Institute adopted a threshold limit value of 10 ppm in 1969, which OSHA adopted in 1971 without rulemaking under the authority of 29 U.S.C.A. § 655(a). This standard, codified in 29 C.F.R. § 1910.1000 Table Z-2, was based on the nonmalignant toxic effects of benzene exposure and not on any possible leukemia hazard (4).

The U.S. Congress passed the Safe Drinking Water Act (SDWA) in 1974. The law requires the U.S. EPA to determine safe levels of chemicals in drinking water, which do or do not cause health problems. These nonenforceable levels are based on possible health risks and exposures are called Maximum Contaminant Level Goals (MCLG). The MCLG for benzene is zero, as this level would not cause any harm to humans. The EPA set the enforceable Maximum Contaminant Level (MCL) for benzene at 5 μg/L or 5 parts per billion (ppb). This is the lowest level to which water systems can be reasonably required to remove benzene prior to serving the water. All community or public water supplies in the United States must abide by these regulations (2).

The carcinogenic nature of benzene was known from as early as the 1920s (5) and since then numerous medical studies have confirmed linkages between human exposure to benzene and the occurrence of leukemia. Since then, a number of U.S. health authorities [e.g., the Occupational Safety and Health Association (OSHA), the National Institute of Occupational Safety and Health (NIOSH), and the American Conference of Governmental Industrial Hygienists (ACGIH)] have set benzene exposure risk standards (at levels not exceeding 0.1 ppm per day) (6). Short-term exposure above the MCL could cause temporary nervous system disorders, immune system depression, and anemia. Long-term lifetime benzene exposure at levels above the MCL has the potential to cause chromosome aberrations and cancer (2).

CONTAMINATION AND EXPOSURE

Benzene has been found in at least 813 of the 1430 National Priorities List sites identified by the U.S. EPA. Industrial sites and gasoline service stations are the main sites of benzene contamination (7). Human exposure to benzene may result from such factors as industrial and motor vehicle emissions, tobacco smoke, industrial factories working environments, and hazardous waste disposal sites. Benzene is capable of passing into the air from soil and water sources. The largest issue is benzene contamination from leaking underground storage tanks (LUSTs) into groundwater, which may impact drinking water sources.

Remedial Measures

Environmental contamination of pure benzene is unlikely since benzene is often used as an ingredient with other compounds. Most benzene contamination occurs when other more common chemicals, such as gasoline, having benzene as a main constituent, are released into the environment. Examples of such products include petroleum fuel hydrocarbons (gasoline, diesel, and kerosene), industrial solvents, detergents, and pesticides, among others compounds (7). At present, there are a number of major contamination sites containing elevated levels of benzene and benzene-based products. In many of these cases, the contamination poses a threat to sources of groundwater utilized for human consumption.

Benzene does not readily sorb to soil and therefore may be transported large distances from its source depending on groundwater velocity and hydraulic conductivity. The potential environmental impact of subsurface benzene contamination is dependent on many factors. Under optimal subsurface conditions, benzene is biodegradable. Benzene is relatively easy to oxidize with chemical oxidants such as ozone, potassium permanganate, sodium persulfate, and hydrogen peroxide. Benzene has been extracted from the subsurface using soil vapor extraction in the vadose zone and groundwater pump and treat systems. Common aboveground treatment methods for benzene include thermal and catalytic oxidizers, granular activated carbon, biofilters, and advanced oxidation processes. Aerobic bioremediation and enhanced natural attenuation work well with benzene. Suthersan (8) describes various gasoline and benzene bioremediation methods.

Extraction technologies such as soil vapor extraction (SVE) for the vadose zone and groundwater extraction with pumps and aboveground treatment of benzene have been used for decades. A good summary of various technologies for benzene vadose zone and groundwater treatment is Reference 9.

BIBLIOGRAPHY

1. Online Lawyer Source (OLS) (2003). The history of benzene. Available: http://www.onlinelawyersource.com/benzene/history.html.
2. United States Environmental Protection Agency (U.S. EPA) (2004). Consumer factsheet on benzene. Available: http://www.epa.gov/safewater/contaminants/dw_contamfs/benzene.html.
3. National Safety Council (2004). Available: http://www.nsc.org/library/chemical/benzene.htm.
4. Environmental Law Reporter (2004). American Petroleum Institute v. Occupational Safety and Health Administration. Available: http://www.elr.info/litigation/vol8/8.20790.htm.
5. Boston-Lawyer-Referral.com (BLR) (2004). Benzene cancer cases. Available: http://www.boston-lawyer-referral.com/pgs/benzene_cancer_14.html.
6. Acute myelogenous leukemia (AML). (2004). A history of benzene standards: benzene and acute myelogenous leukemia story. Available: http://www.acute-myelogenous-leukemia.com/html/risk.html.
7. Agency for Toxic Substances and Disease Registry (ATSDR) (2004). ToxFAQs™ for benzene. Available: http://www.atsdr.cdc.gov/tfacts3.html#bookmark02.
8. Suthersan, S. (2002). *Natural and Enhanced Remediation Systems*. Lewis Publishers/CRC Press, Boca Raton, FL.
9. Testa, S.M. and Winegardner, D.L. (2000). *Restoration of Contaminated Aquifers: Petroleum Hydrocarbons and Organic Compounds*, 2nd Edn. CRC Press, Boca Raton, FL.

GROUNDWATER AND NITRATE: CHEMICAL BEHAVIOR AND TREATMENT

James A. Jacobs
Lorraine Geddes-McDonald
Environmental Bio-Systems, Inc.
Mill Valley, California

Nitrate is a naturally occurring substance, which is an essential ingredient for the growth of plants. Human activities can contribute to higher levels of nitrate in the ground and surface waters. The major source of excess nitrogen is nitrate fertilizers for agricultural, lawn, and other garden uses. Nitrates are sourced from nitrates in nitrate fertilizers, the storage or use of animal manure, leaky septic systems, leaky sanitary sewer pipelines, and improper disposal of treated wastewater containing high nitrate levels. Through storm runoff and leaching processes, nitrates have become dissolved in stormwater runoff or infiltration waters. These sources of dissolved nitrates have created significant impacts on worldwide surface water and groundwater resources.

INTRODUCTION

Besides nitrate-based fertilizers, other sources of nitrates include animal manure from feedlots, dairies, and poultry farms; septic tanks; rainwater (acid rain); and sewage sludge. Human exposure to nitrates can occur through certain foods, such as cured meats and vegetables, and through tobacco smoke (1).

HISTORY

Daniel Rutherford in Scotland discovered nitrogen in 1772. In 1905 the world's first production of nitrogen fertilizer [Nrgessalpeter (calcium nitrate)] was done at a test facility in Nottodon, Norway. Production was based on the Birkeland Eyde invention using hydroelectric power to extract nitrogen from the air.

The German physical chemist Fritz Haber invented a process in 1909 that directly combines nitrogen from the air with hydrogen from a hydrocarbon source under extremely high pressure (200–400 atm) and moderately high temperatures (400–650 °C) to make ammonia (NH_3). The process initially used coal and later natural gas as a source of hydrogen and as a source for the considerable energy required. For his efforts, Haber received the Nobel Prize in Chemistry in 1918. It was an industrial chemist, Carl Bosch, who took the Haber method and designed a large scale process using a catalyst and high pressure methods. Those efforts were rewarded with a jointly shared Nobel Prize in 1931 with Friedrich Bergius for high pressure studies. The discovery and feasibility of the Haber–Bosch process (also called the Haber Ammonia Process and the Synthetic Ammonia Process) were dependent on the newly discovered sources of energy from coal and petroleum, which made the production of ammonia economically feasible. In 1928, ammonia-based production replaced the Birkeland Eyde technology both at the Nottodon and Rjukan facilities. Calcium nitrate based on ammonia from Rjukan nitrogen–phosphorus–potassium (NPK) production using a newer nitrophosphate process, started in the mid-1930s and gained importance in the mid-1960s.

PROPERTIES OF NITRATES

Nitrates, chemical compounds containing the nitrate (NO_3) radical, are salts or esters of nitric acid (HNO_3)

formed by replacing the hydrogen with a metal (e.g., sodium or potassium) or a radical (e.g., ammonia or ethyl). Some important inorganic nitrates are potassium nitrate (KNO_3), sodium nitrate ($NaNO_3$), silver nitrate ($AgNO_3$), and ammonium nitrate (NH_4NO_3). Calcium nitrate is also used in fertilizers. The presence of nitrate in the soil is of great importance, since it is from these compounds that plants obtain the nitrogen necessary for their growth. Nitrogen-fixing bacteria are important in keeping soil supplied with nitrates.

Nitrate is present in various compounds: nitrate (NO_3), nitrate-nitrogen (NO_3-N), nitrite (NO_2), nitrate plus nitrite (NO_3+ NO_2). Each represents a different measurement of nitrogen's impact on health.

Persistence in Groundwater

Nitrate occurs naturally in the subsurface in small quantities. Due to large inputs of agricultural fertilizers and runoff from manure, the level of nitrate in groundwater has in risen significantly in some locations. Nitrate in groundwater can be reduced by denitrification, resulting in nitrogen gas (2). Commonly available activated carbon, filtration, and standard water softeners do not remove nitrate-nitrogen from water.

Taste and Odor

Nitrate is odorless and colorless, so it can be detected only through a water test.

Health Risk

The main health threat from nitrate is methemoglobinemia, or blue baby syndrome. Infants less than six months old do not have the enzyme needed to break down nitrates. Excess nitrogen in the system interferes with the blood's ability to carry oxygen, resulting in a bluish tint to the skin. This condition is rare but can be fatal.

There is no conclusive evidence that nitrates can cause cancer in adults. However, there is some evidence that high levels of nitrate consumed over a long period may be toxic. Maximum contaminant levels (MCLs) are standards that provide a minimal level of risk to health from consuming a contaminant over the course of a lifetime. Lifetime exposure to nitrate levels above MCL can cause diureses, increased starchy deposits, and hemorrhaging of the spleen (3).

A water test for nitrate is suggested for households with infants, pregnant women, nursing mothers, or elderly people. It is recommended that new water supplies should be tested to determine the baseline nitrate concentration. In addition, if a water supply has never been tested for nitrate, it should be tested. If excess nitrate is found in a water supply, then the water should be treated, or an alternative source of potable water should be secured (4).

REGULATORY ISSUES

In 1974 the U.S. congress passed the Safe Drinking Water Act. This law requires the Environmental Protection Agency (EPA) to determine safe levels of chemicals in drinking water, which do or may cause health problems. The MCL for nitrate has been set at 10 ppm and 1 ppm for nitrite (3).

FATE AND TRANSPORT

Nitrate is a highly soluble anion that readily transports in groundwater, resulting in contamination of large subsurface areas. Many sites are affected by nitrate contamination. Nitrate is easily transported in oxygen-rich water with no transformation and little or no retardation. Nitrate is not removed by filtration and the compound is stable over a considerable range of conditions. Large numbers of activities contribute to the problem: farming, fertilization, animal feedlots, dairy; manufacturing of explosives and chemicals; nuclear industry, large amounts of nitric acid, dissolved metals and actinides; mining industry.

Use of agricultural fertilizers is the major source of contamination. Nitrate contamination depends on factors such as the amount of excess nitrate not used by plants, the soil type, and the underlying geology. Areas with soils that are sandy, gravelly, or shallow over porous limestone bedrock and areas with karsts topography (cave areas) have the greatest risk of nitrate contamination of groundwater sources.

The primary inorganic nitrates that may contaminate drinking water are KNO_3 and NH_4, both of which are widely used as fertilizers. Nitrates are highly soluble in water and do not bind to soils. They have a high potential to migrate to groundwater. Because they do not evaporate, nitrates are likely to remain in H_2O until consumed by plants or other organisms (3).

Nitrate is often the dominant source of nitrogen in the soil since it generally occurs in high concentrations and is free to move to the roots by mass flow and diffusion. In addition to KNO_3 and NH_4, other nitrate-containing fertilizers including $NaNO_3$ and $Ca(NO_3)$ are quite soluble and thus very mobile in soil solution. They are quickly available to crops because of their mobility. They are susceptible to leaching under conditions of high rainfall. Nitrate supplied in commercial fertilizer is subject to leaching. The $NO_3{}^-$ form of nitrogen is completely mobile and within limits moves largely with soil water. Under conditions of excessive precipitation or irrigation, it is leached out of the vadose zone.

Immobility of nitrate occurs when large quantities of low nitrogen crop residues begin decomposing in the soil. The high amounts of carbohydrate in such residues cause the population of microflora to build up quickly. As new cells are formed, nitrogen is used to build up protoplasm. Almost invariably this leads to a decrease in levels of nitrates. Biodegradation of nitrates occurs through processes such as denitrification (5).

A study by the United States Geological Survey for point and nonpoint sources in major U.S. watersheds indicated that agricultural nitrate accounted for over 80% of the nitrogen input into the environment: (53%) fertilizers, (27%) animal manure, (14%) atmospheric, and (6%) point sources (6).

GROUNDWATER INVESTIGATION

Nitrate is highly soluble in water and poorly retained in the soil; therefore, it can move through the soil and into groundwater sources. The likelihood and seriousness of leaching is of concern as the public cost of groundwater contamination is high. Studies have shown that more public supply wells have been impacted by nitrate contamination than all other contaminants combined. Frequent sampling for nitrate contamination should be conducted.

REMEDIATION

Several approaches for actively treating nitrate contamination have been proposed. The duration, success, and cost of treatment are dependent on the method used.

Phytoremediation using shrubs and specific types of plants to remediate a nitrate-contaminated aquifer is a passive but cost effective method to remove contaminants. However, it is a slow process and could take decades to clean up nitrate contamination.

Active pump and treat remediation is also used successfully in depleting nitrate contaminant. Nitrate reduction by denitrification where nitrate is sequentially reduced to nitrite, nitric oxide, and nitrogen gas is also effective.

Nitrate can be removed in anaerobic groundwater systems, such as may occur in natural or contaminated environments with elevated concentrations of dissolved organic matter, by bacterial action through the process of denitrification (7). It is the major microbial process that destroys nitrate by reducing it into nitrogen gas (N_2):

$$\begin{array}{ccccccccc} (1) & & (2) & & (3) & & (4) & & \\ NO_3^- & \rightarrow & NO_2^- & \rightarrow & [NO] & \rightarrow & N_2O & \rightarrow & N_2 \end{array}$$

Oxidation state:

$$\begin{array}{ccccc} +5 & +3 & +2 & +1 & 0 \end{array}$$

where (1) is nitrate reductase, (2) is nitrite reductase, (3) is nitric oxide reductase, and (4) is nitrous oxide reductase.

Bacteria use nitrate as a terminal electron acceptor when oxygen is not available. Denitrification has developed over millions of years and can occur readily in an oxygen-depleted environment (anaerobic) with abundant carbon and electron acceptors.

Electron acceptors include oxygen (O_2), nitrate (NO_3^-), nitrite (NO_2^-), manganese [Mn(IV)], ferric iron [Fe(III)], sulfate (SO_4^{2-}), and carbon dioxide (CO_2) (2). Usually by the time the sulfate and carbon dioxide are used as the electron acceptor or oxidant, the nitrate is already completely reduced to nitrogen gas (N_2). More information on denitrification can be found in References 2 and 8. The half-cell reaction for denitrification produces -16.9 kcal/mole $\cdot\, e^-$ of usable energy, also called the Gibbs free energy of reaction (9):

$$5e^- + 6H^+ + NO_3^- \rightarrow 0.5N_2 + 3H_2O$$

An injection of a carbon source into wells, well points, probe rods, trenches, or other delivery systems can move a nitrate-contaminated groundwater plume into the denitrification process, which will produce nitrogen gas as the end product. Laboratory bench testing is recommended prior to pilot scale field testing to optimize the site-specific geochemistry. Adjustments of pH or additions of nutrients may be required to optimize nitrate reduction rates.

Short-term solutions to high nitrate concentrations include bottled water and blending nitrate-impacted water with clean water to deliver water that is within the safe range for human consumption. Nitrate can be removed from drinking water by several physical/chemical methods, including distillation, reverse osmosis, electrodialysis or electrodialysis reversal, and ion exchange (4).

Distillation involves heating the water to boiling and collecting and condensing the steam generated as the water boils. Merely boiling water will increase rather than decrease the nitrate concentration. In the reverse osmosis process, pressure is applied to water to force it through a semipermeable membrane. As the water passes through the membrane, dissolved salts, including nitrate, are removed. Electrodialysis uses a driving force of direct current (DC) power to transfer ionic species (e.g., nitrate, sodium) from the feed water through cation and anion transfer membranes to a concentrated stream, creating a more dilute stream as the product water. Electrodialysis reversal is a variation on the electrodialysis process, in that it uses electrode polarity reversal to automatically clean membrane surfaces. Nitrate removal by ion exchange operates on the same principle as a household water softener. In a standard water softener, calcium and magnesium ions are exchanged for sodium ions. For nitrate removal, special anion exchange resins are used that exchange chloride ions for nitrate and sulfate ions in the water as it passes through the resin (4).

CONCLUSION

Nitrate is highly soluble and is subject to leaching. Therefore, it will migrate readily to groundwater sources. High levels of rainfall or frequent irrigating make groundwater sources more susceptible to nitrate contamination from agricultural fertilizers. Enhanced natural attenuation using a carbon source (such as propane gas, molasses, corn syrup, or lactic acid) can lead to denitrification of nitrates, producing nitrogen gas as the end product.

BIBLIOGRAPHY

1. Ilvento, X. and Tarabo, X. (1997).
2. Suthersan, S. (2002). *Natural and Enhanced Remediation Systems*, Lewis Publishers/CRC Press, Boca Raton, FL, p. 419.
3. United States Environmental Protection Agency (USEPA) (2004). Consumer factsheet on nitrates/nitrites, updated November 1, 2004. Available: http://www.epa.gov/safewater/contaminants/dw_contamfs/nitrates.html.
4. Ulrick, J. (2004). Preprint of Groundwater Resources Association of California Nitrate Overview Paper, Sacramento, CA.
5. Tisdale, S.L. and Nelson, W.L. (2004). *Soil Fertility and Fertilizers: An Introduction to Nutrient Management*, Prentice-Hall, Engelwood Cliffs, NJ, p. 634.

6. Pucket, L.J., *Nonpoint and Point Sources of Nitrogen in Major Watersheds of the United States*, U.S. Geological Survey National Water Quality Assessment Program, WRI94-4001, 1994.
7. Hem, J.D. (1992). *Study and Interpretation of the Chemical Characteristics of Natural Water*, 3rd Edn. U.S. Geological Survey Water Supply Paper 2254.
8. National Resource Council (NRC) (2000). *Natural Attenuation for Groundwater Remediation*. National Academy Press, Washington, DC.
9. Wiedemeier, T.H., Rifai, H.S., Newell, C.J., and Wilson, J.T. (1999). *Natural Attenuation of Fuels and Solvents in the Subsurface*, 1st Edn. John Wiley & Sons, New York.

GROUNDWATER AND PERCHLORATE: CHEMICAL BEHAVIOR AND TREATMENT

James A. Jacobs
Shonel Dwyer
Environmental Bio-Systems, Inc.
Mill Valley, California

Large-scale perchlorate production began in the 1940s. Perchlorate leakage has been ongoing in some cases for decades. The majority of perchlorate spills are related to the use of explosive propellant as a solid rocket fuel for missiles and for the NASA space shuttle. It has only been since the late 1990s, when laboratory methods improved detection of perchlorate in the groundwater, that the full impact of the groundwater contamination became evident. A variety of point of use treatments for perchlorate are available, at varying efficiencies and costs. Of the various *in situ* technologies, anaerobic bioremediation is the most promising for widespread subsurface perchlorate treatment.

INTRODUCTION

The environmental impact/toxicity of perchlorate is an active area of research, critical to the protection of groundwater resources. Perchlorate ($ClO_4{}^-$) is the soluble anion associated with the solid salts of ammonium, magnesium, potassium, and sodium perchlorate. Perchlorate is used in the United States to make rocket fuel, fireworks, road flares, and similar products. It remains one of the key ingredients in numerous national defense weapons systems. The occurrence of perchlorate in public drinking water supplies over the years may be inextricably linked with inadequate maintenance, storage, and disposal practices followed in the past. Consequently, varying quantities of the compound leached into the soil and groundwater, especially at or near sites that developed or manufactured aerospace materials or armaments for U.S. national defense programs.

Contamination of numerous sites is highest in California, where cities like Rialto, already affected by drought and increasing population, are continuing to lose water capacity to perchlorate contamination and its spreading plume. Perchlorate is an unregulated contaminant in the United States, meaning it does not have a set maximum concentration level (MCL). A good overview of perchlorate issues is summarized by Motzer (1).

Perchlorate

Perchlorate has the following uses:

- As a solid propellant in rockets, missiles, and fireworks (primarily ammonium perchlorate)
- In road flares
- As a component of airbag inflators
- In munitions for smokeless powder
- In nuclear reactors and electronic tubes
- As additive in lubricating oils
- In tanning and finishing leather
- As a mordant for fabrics and dyes
- In electroplating, aluminum refining, and rubber manufacture
- In the production of paints and enamels
- As a component of certain imported Chilean fertilizers

About 90% of the 20 million pounds of perchlorate produced each year goes into solid rocket fuel for Air Force missiles and the NASA space shuttle (2). Although the vast majority of locations where perchlorate has been detected in the groundwater are influenced by anthropogenic activity, perchlorate has been found to occur naturally in Chilean caliche. Perchlorate is reported to be present in some caliche formations in Chile that are used to produce nitrate fertilizers (3).

Ammonium Perchlorate (ClH_4NO_4)

Ammonium perchlorate is used as an energetics booster or oxidant in solid propellant for rockets and missiles. It is therefore a national technical asset integral to the nation's strategic defense system and space exploration. Ammonium perchlorate is also used in certain fireworks, the manufacture of matches, as a component of airbag inflators, and in analytical chemistry to preserve ionic strength. Large-scale production of ammonium perchlorate began in the United States in the mid-1940s. Ammonium perchlorate has a limited shelf life and must be periodically replaced in munitions and rockets or in inventory. This has led to the disposal of large volumes of the compound since the 1940s in Nevada, California, Utah, and other states.

Potassium Perchlorate ($KClO_4$)

Potassium perchlorate until recently was used to treat hyperthyroidism resulting from Graves' disease and is still used diagnostically to test thyroid hormone production in some clinical settings (medicinal use). In addition, potassium perchlorate is used in protective breathing equipment on Air National Guard (ANG) aircraft for use in the event of depressurization, and in naval emergency escape breathing devices (4).

Sodium Perchlorate (NaClO4:H_2O)

Sodium perchlorate is mainly used in the manufacture of perchloric acid, also used in the gun-powder industry.

HISTORY OF OCCURRENCE

Several hundred drinking water wells were sampled by the California Department of Health Services (CDHS) for perchlorate beginning February 1997. Perchlorate was first detected in drinking water wells (up to 260 μg/L) near the Aerojet site in Sacramento County. Groundwater treated to remove volatile organic chemicals (such as trichloroethylene) was reinjected into the groundwater. Perchlorate, also in the contaminated shallow groundwater, has been present in the reinjected water at concentrations up to 8000 μg/L (5).

Characteristics of Perchlorate

Perchlorate is a highly soluble inorganic anion (2.09 kg/L for $NaClO_4{}^-$) that adsorbs poorly to mineral surfaces and activated carbon and is not retarded during groundwater transport.

Persistence of Perchlorate in Groundwater

Perchlorate is very stable/inert in groundwater. It does not biodegrade unless there is an absence of available, dissolved oxygen (oxygen acts as an alternate electron acceptor). Groundwater has naturally occurring dissolved oxygen, which microbes will use first before turning to other sources, so natural biodegradation of perchlorate is unlikely.

Plumes may be nearly 8 miles (which is large). Perchlorate may disperse in the aquifer as it flows downgradient in the subsurface, to the extent that it may not be detectable using available analytical methods.

Taste and Odor

There is no perceptible odor or taste that perchlorate imparts to water.

Fate and Transport

Fate and transport of the chemical essentially depend on its physical characteristics and natural attenuation. Perchlorate is highly soluble in groundwater. Volatilization is not expected to be one of the predominant pathways of transport. Perchlorate and concentrated solutions of perchlorate are denser than water (ammonium perchlorate at 1.95 g/mL), which allows it to sink. Perchlorate is kinetically stable and adsorbs poorly to mineral surfaces and activated carbon and is not retarded during groundwater transport.

HEALTH RISKS

At present, there is a large amount of information available on the health effects from exposure to perchlorate; however, there is still widespread debate on the potential toxicity of perchlorate, the ecological receptors, and the long-term effects on humans. Health studies using data from patients who were given perchlorate to treat thyroid conditions indicate that high doses can affect metabolism, growth, and development of the body. Long-term exposure to high levels of perchlorate taken as medication has been linked to a potentially serious blood disorder. However, it is impossible for the levels of perchlorate in drinking water to cause such conditions. Perchlorate slows the uptake of iodide into the thyroid. Because iodide is an essential component of thyroid hormones, at high doses perchlorate can affect how the thyroid functions. In adults, the thyroid helps to regulate metabolism; in children, the thyroid also plays a role in proper development (6).

A clinical study using human volunteers was performed to determine whether levels of perchlorate similar to the low levels that are currently being found in the groundwater are safe for people. This human study was performed in the same way that pharmaceutical studies look at the effect that soon-to-be-approved medicines have on people. At the low levels used in the study, no health effects were observed (7). In Chile, an epidemiological study was performed on pregnant women who drank water that contained naturally occurring perchlorate, at levels at least 10 times higher than what we've detected in our drinking water. There were no health effects observed (8).

The maximum contaminant level (MCL) has not been established by the U.S. Environmental Protection Agency and review and evaluation are ongoing. California has more perchlorate sites than any other state or country. Until the MCL is in place, the California Department of Health Services (CDHS) will continue to use a 6-μg/L perchlorate level as a notification level or public health goal with which to protect water consumers (5).

Perchlorate is easily absorbed into the bloodstream after ingestion but leaves the body very quickly. Infants and small children may be more sensitive to perchlorate than adults. The U.S. Environmental Protection Agency (EPA) is focusing on those most sensitive in its ongoing evaluation. Perchlorate does not appear to cause cancer at low levels that are currently being measured in drinking water. Perchlorate does not cause Graves' disease. Perchlorate does not cause changes in thyroid hormones at the levels that are currently being measured in drinking water (6).

GROUNDWATER INVESTIGATION

Widespread perchlorate contamination of groundwater only became known when a new ion chromatography (IC) method lowered the anion's detection limit in water from 400 parts per billion (ppb) to 100 ppb. After more refinement in the spring of 1997, a method was developed (EPA Method 314 using IC-MS or LC-MS-MS) that could quantify perchlorate concentrations to a limit of 4 ppb (9,10).

REMEDIATION

Treatment of perchlorate contamination in water is essentially complex. The perchlorate anion does not respond to typical water treatment techniques due to its fundamental physical and chemical nature. Chemical oxidation, granular activated carbon, and air sparging technologies have been evaluated as being relatively ineffective for perchlorate treatment. Tailored granular activated carbon impregnated with special amine coatings are being used with success for point of use applications. Fluidized bed bioreactors are being used at the Kerr McGhee, Henderson, Nevada Site. At that site, perchlorate influent is approximately 180 ppm; after 45 min of contact time, the effluent is less than 4 ppb (A. Eaton, personal communication, 2004).

The ideal treatment technology for a given perchlorate occurrence may depend on a number of factors, including perchlorate concentration, the presence and concentration of cocontaminants, pH, alkalinity, presence of natural organic matter (NOM), amount of total dissolved solids (TDS), presence of metals, and geochemical parameters (nitrate, sulfate, chloride, dissolved oxygen, redox potential, etc.). The presence of indigenous perchlorate-reducing microbes and substances inhibitory to microbial activity will also influence perchlorate treatment technology effectiveness. For *in situ* treatment of perchlorate contamination, variables related to the site hydrogeological setting, such as depth to and distribution of contaminants, soil permeability, and groundwater flow velocity, are also important.

PHYSICAL REMOVAL TECHNOLOGIES

Ion Exchange

This is a process with two similar applications of the same technology:

- Water softening removes ions from the water and replaces them with sodium (Na^+) and chloride (Cl^-) ions. This technique is employed at Aerojet, Sacramento, California.
- Deionization removes ions and replaces them with hydrogen (H^+) and hydroxyl (OH^-) ions, which combine to form water.

Ion exchange treatment has been successful in reducing perchlorate concentrations in water from 75 ppb to less than detectable levels at the San Gabriel Valley Superfund sites. Resulting brines from the ion exchange process can be problematic because the perchlorate is concentrated and not destroyed. Ion exchange is the preferred large-scale treatment method used by utilities for aboveground treatment of perchlorate.

Single Pass Adsorption Resins

This technology uses nonregenerative single-pass adsorption resins, which are tailored for perchlorate. This technology is currently being used by municipalities for large-scale water treatment.

Membrane Techniques

- *Reverse Osmosis and Nanofiltration*. This technology uses semipermeable membranes, allowing the water to pass through it, while retaining perchlorate. RO and NF can be costly due to the energy use and perchlorate disposal. They are typically used on small-scale systems.
- *Electrodialysis*. This is an electrically driven membrane separation process that separates ions from water. The process is based on the property of ion exchange membranes to selectively reject anions or cations.

In Situ Biological Treatment

In situ anaerobic biological remediation of perchlorate-contaminated waters is promising. The anaerobic conversion of chlorine in perchlorate to chloride requires the overall transfer of eight electrons (11):

$$\text{perchlorate}(ClO_4{}^-) \longrightarrow \text{chlorate}(ClO_3{}^-) \longrightarrow \text{chlorite}\ (ClO_2{}^-) \longrightarrow \text{oxygen}\ (O_2) + \text{chloride}\ (Cl^-)$$

The perchlorate and chlorate are used as electron acceptors (11). An *in situ* treatment zone or wall could be emplaced with a biodegradable carbon source, such as propane (gas infusion), molasses, corn syrup, or lactic acid.

CONCLUSION

Perchlorate plumes have been developing for decades. New technologies are being developed for aboveground and *in situ* treatment for perchlorate.

BIBLIOGRAPHY

1. Motzer, W.E. (2001). Perchlorate: problems, detection, and solutions. *Environ. Forensics* **2**(4): 301–311.
2. United States Environmental Protection Agency (U.S. EPA) (1998). *Perchlorate Environmental Contamination: Toxicological Review and Risk Characterization*. U.S. EPA, Washington, DC.
3. Urbansky, E.T., Collette, T.W., Robarge, W.P., Hall, W.L., Skillen, J.M., and Kane, P.F. (2004). United States Environmental Protection Agency (U.S. EPA), Survey of fertilizers and related materials for perchlorate; EPA/600/R-01/047; NRMRL/WSWRD. Available: http://www.epa.gov/ORD/NRMRL/Pubs/600R01047.htm; updated February 19, 2004.
4. California Environmental Protection Agency (CEPA) (2002). *Public Heath Goal for Perchlorate in Drinking Water*, Pesticide and Environmental Toxicology Section, Office of Environmental Health Hazard Assessment, December, draft, p. 100.
5. California Department of Health Services (CDHS) (2004). Perchlorate in California drinking water: overview and links. Available: http://www.dhs.ca.gov/ps/ddwem/chemicals/perchl/perchlindex.htm.
6. United States Environmental Protection Agency (U.S. EPA) (2004). Perchlorate. Available: http://www.epa.gov/safewater/ccl/perchlorate/perchlorate.html; updated August 5, 2004.
7. Greer, M.A., Goodman, G., Pleus, R.C., and Greer, S.E. (2002). Health effects assessment for environmental perchlorate contamination: the dose response for inhibition of

thyroidal radioiodine uptake in humans. *Environ. Health Perspect.* **110**(9): 927–937.

8. Crump, C. et al. (2000). Does perchlorate in drinking water affect thyroid function in newborns or school-age children? *J. Occup. Environ. Med.* **42**(6): 603–612.
9. Horst, J. (2004). *Technology Information Sheet: Emerging Contaminants: Perchlorate, ARCADIS*, Jan., p. 2.
10. Eaton, A. and Via, S. (2004). Verifying the reliability of EPA method 314 to measure perchlorate at sub ppb levels vs. new EPA method options, Abstracts, The Eleventh Symposium in GRA's Series on Groundwater Contaminants: ClO_4^-; 2004—Perchlorate in California's Groundwater, Groundwater Resources Association Meeting, Glendale, CA, August 4.
11. Suthersan, S. (2002). *Natural and Enhanced Remediation Systems*. Lewis Publishers/CRC Press, Boca Raton, FL, p. 419.

GROUNDWATER AND VINYL CHLORIDE: CHEMICAL BEHAVIOR AND TREATMENT

James A. Jacobs
Angela Munroe
Environmental Bio-Systems, Inc.
Mill Valley, California

Vinyl chloride is a colorless, flammable gas at normal temperatures with a mild, sweet odor. It exists in liquid form if it is kept under high pressure or at low temperatures (less than −13.4 °C). Vinyl chloride is also known as chloroethene, chloroethylene, and ethylene monochloride (1). All vinyl chloride is manufactured or results from the breakdown of other manufactured substances, such as trichloroethane (TCA), tetrachloroethylene (PCE), and trichloroethylene (TCE). Vinyl chloride was first produced commercially in the 1920s and is now one of the highest volume chemicals produced in the United States (2). Of the vinyl chloride produced, 96–98% is used to make poly(vinyl chloride) (PVC) with the remaining 2–4% being utilized in the manufacture of specific chlorinated hydrocarbons such as 1,1,1-TCA, 1,1,2-TCA, and vinylidene chloride (3). PVC is used to make a variety of plastic products including pipes, wire, cable coatings, and packaging materials. Other uses include furniture and automobile upholstery, wall coverings, housewares, and auto parts. Vinyl chloride has been used in the past as a refrigerant, an extraction solvent for heat-sensitive materials, and in the production of chloroacetaldehyde and methyl chloroform (4). Also, limited quantities of vinyl chloride were used as an aerosol propellant and as an ingredient in drug and cosmetic products (4). However, these practices were banned by the United States Environmental Protection Agency (U.S. EPA) in 1974.

CHARACTERISTICS

Vinyl chloride is a colorless, flammable gas with a faintly sweet odor. It is slightly soluble in water, soluble in ethanol, and very soluble in ether, carbon tetrachloride, and benzene. It is extremely flammable and is easily ignited by heat, sparks, or flames. Explosive mixtures will form with air, and it tends to self-polymerize explosively if peroxidation occurs. It reacts vigorously with oxidizers and, in the presence of moisture, corrodes iron and steel. Fires may release irritating and toxic gases including carbon dioxide, carbon monoxide, hydrogen chloride, and traces of phosgene (2). A good overview of chlorinated solvents and vinyl chloride is included in Reference 5.

PRODUCTION

Vinyl chloride was first produced commercially in the 1920s (2). In the Unites States, production of vinyl chloride steadily increased from approximately 6 to 9 billion pounds per year in the late 1970s through the 1980s to approximately 15 billion pounds per year in the mid-1990s (2). The projected demand for 2001 was 16.8 billion pounds (2). There are at least 11 current chemical suppliers of vinyl chloride in the United States. Imports reached a peak of 302 million pounds in 1989. In 2000, the United States imported more than 231 million pounds compared to 29 million pounds in 1994 and 164 million pounds in 1991 (1). Exports fluctuated from approximately 685 million pounds per year to 2.2 billion pounds per year from the late 1970s to the mid-1990s (1). In 2000, U.S. exports were approximately 1.03 billion pounds (2).

Commercial Production of Vinyl Chloride

Vinyl chloride is produced commercially by the chlorination of ethylene through one of two processes, direct chlorination or oxychlorination. Direct chlorination reacts ethylene with chlorine to produce 1,2-dichloroethane. Similarly, oxychlorination produces 1,2-dichloroethane, but this is accomplished by reacting ethylene with dry hydrogen chloride and oxygen:

$$\mathrm{HCl} + \mathrm{H_2C{=}CH_2} + 0.5\ \mathrm{O_2} \longrightarrow \mathrm{Y\ H_2C{=}CHCl} + \mathrm{H_2O}$$

After both processes, the 1,2-dichloroethane is subjected to high pressures (2.5–3.0 MPa) and temperatures (550–550 °C) (1). This causes the 1,2-dichloroethane to undergo pyrolysis, or thermal cracking, which forms the vinyl chloride monomer and hydrogen chloride. The vinyl chloride monomer is then isolated. The technical-grade product is available in 99.9% purity (1).

Production of Vinyl Chloride in Groundwater

Vinyl chloride is produced in groundwater through the biodegradation of chlorinated solvents such as PCE and TCE. Biodegradation is a process in which naturally occurring microorganisms, such as yeast, fungi, and bacteria, break down target substances, such as fuels and chlorinated solvents (6). In the absence of oxygen, microorganisms often use chlorinated compounds to aid in respiration. This is done through an electron transfer process called dehalorespiration (7). In the case where the carbon in the contaminant is the food source, the contaminant is an electron donor. In the case where food is obtained from another source, the contaminant sometimes aids this transfer by accepting electrons that are shed through respiration (6). The most common anaerobic process for degrading chlorinated compounds is an electron transfer process called reductive dechlorination. In this process, hydrogen atoms are sequentially substituted for a chlorine atom in the contaminant molecules. PCE is reductively chlorinated stepwise through TCE, DCE, vinyl chloride, and then to ethene (ethylene). However, vinyl chloride resists reductive dechlorination once it is formed. The major requirement for reductive dechlorination is the presence of other organic compounds, such as fuels, that can serve as the food source (7).

FATE AND TRANSPORT

Anthropogenic sources are responsible for all of the vinyl chloride found in the environment. Most of the vinyl chloride released to the environment is eventually transported to the atmosphere. Lesser amounts are transported to groundwater. Vinyl chloride has been detected in the ambient air near vinyl chloride and PVC manufacturing plants and hazardous waste sites. It has been found in tobacco smoke, perhaps as a result of the manufacturing process (8). When vinyl chloride is released to the ground, it is not adsorbed onto soil; it migrates readily to groundwater, where it degrades to carbon dioxide and the chloride ion. It may also remain unchanged for several months to years. Vinyl chloride has been reported in groundwater as a degradation product of chlorinated solvents TCE and PCE (9).

Effluents and emissions from vinyl chloride and PVC manufacturers are responsible for most of the vinyl chloride released to the environment. It has entered the environment at hazardous waste sites as a result of its improper disposal or leakage from storage containers or from spills. When released to the atmosphere, vinyl chloride is expected to be removed by reaction with photochemically generated hydroxyl radicals (half-life = 1–2 days). When vinyl chloride breaks down in air, it can form other harmful chemicals. Reaction products include hydrochloric acid, formaldehyde, formyl chloride, acetylene, chloroacetaldehyde, chloroacetylchloranil, and chloroethylene epoxide. In photochemical smog, the half-life of vinyl chloride is reduced to a few hours. When released to water, volatilization is expected to be the primary fate process. In waters containing photosensitizers, such as humic materials, sensitized photodegradation may also be important. When released to soil, vinyl chloride either volatilizes rapidly from soil surfaces or leaches readily through soil, ultimately entering groundwater (1).

DISPOSAL

The recommended method of disposal is total destruction by incineration. The temperature of the incinerator must be sufficient to ensure the complete combustion of the vinyl chloride in order to prevent the formation of phosgene. The recommended temperature range for incineration is 450–1600 °C, with residence times of seconds for gases and liquids, and hours for solids (1). If in solution, the vinyl chloride product may need to be adsorbed onto a combustible material prior to incineration. Recommended materials include vermiculite, sawdust, or a sand–soda ash mixture (90/10) covered with wood and paper (1). The vinyl chloride can also be dissolved in a flammable solvent prior to incineration. An acid scrubber should be used in conjunction with the incinerator in order to remove any hydrogen chloride that is produced by the combustion process (1). Alternatively, chemical destruction may be used, especially with small quantities. From 1 to 2 days is generally sufficient for complete destruction. Aqueous by-product solutions from the production of vinyl chloride are usually steam-stripped to remove volatile organic compounds, neutralized, and then treated in an activated sludge system to remove nonvolatile organic compounds remaining in the wastewater (1).

EXPOSURE

Historically, exposure to vinyl chloride has been in the industrial workplace, in particular, from inhalation of factory air. As industry has made efforts to better protect workers from vinyl chloride exposure, attention has also been drawn to other avenues of exposure, such as environmental and commercial (10).

Since vinyl chloride commonly exists in a gaseous state, one is most likely to be exposed to it by breathing it in. Vinyl chloride is not normally found in urban, suburban, or rural air in amounts that are detectable by the usual methods of analysis. However, vinyl chloride has been found in the air near plastics industries, hazardous waste sites, and landfills. The amount of vinyl chloride in the air near these places ranges from trace amounts to 0.041 mg/L of air, but may exceed 1 mg/L. Levels as high as 44 mg/L have been found at some landfills. One can also be exposed to vinyl chloride in the air through tobacco smoke from cigarettes or cigars (10). Workers who produce or use vinyl chloride and PVC are at a high risk of exposure by breathing vapors in the workplace air, or if the chemical comes into contact with their skin or eyes (4)

Exposure to vinyl chloride may also occur by drinking water from contaminated wells, but how often this occurs is not known. Most drinking water supplies do not contain vinyl chloride. In a 1982 survey, vinyl chloride was found in less than 1% of the 945 groundwater supplies tested in the United States. The concentrations found in groundwater were up to 0.008 mg/L, with a detection limit of 0.001 mg/L. Other studies have reported groundwater vinyl chloride concentrations at or below 0.38 mg/L (8). High concentrations up to 200,000 μg/L were detected in well water in the vicinity of a PVC plant in Finland ten years after leakages (9).

At one time, the flow of water through PVC pipes added very low amounts of vinyl chloride to water. For example, in one study of newly installed pipes, the drinking water had 0.001 mg/L of vinyl chloride. No current information on the amount of vinyl chloride released from PVC pipes into water is available (8).

Owing to its high volatility, vinyl chloride has rarely been detected in surface waters; the concentrations measured generally do not exceeding 10 μg/L, with a maximum of 570 μg/L from contaminated sites (9). A recent example is the finding of vinyl chloride up to 56 μg/L together with other volatile organic compounds in heavily polluted shallow rivers in Osaka, Japan (11).

Once set, poly(vinyl chloride) by itself is a hard, brittle, and inflexible material. To induce the flexibility needed for products such as food wraps, plasticizers are added. Poly(vinyl chloride) food wrappings, however, pose a potential problem as monomers and oligomers of vinyl chloride as well as some plasticizers can migrate into food from the wrappings (10). The U.S. government now regulates the amount of vinyl chloride in food packaging materials. A model of food systems shows that when levels less than 1 mg/L of vinyl chloride monomer are used in PVC packaging, “essentially zero” vinyl chloride enters food by contact with these products (8).

Estimated Total Exposure and Relative Contribution from Drinking Water

If an ambient air concentration of 0.1–2.0 μg/m^3 and a daily inhalation of 20 m^3 of air are assumed, daily intake by the inhalation route would amount to 2–40 μg. For those living and working in industrial areas, in particular, near vinyl chloride/PVC polymerization plants or near waste sites, this could be at least a factor of ten or more higher (i.e., daily intake 20–400 μg).

At a concentration of 1–2 μg/L in drinking water, daily intake would be about 2–4 μg. Daily intake from food has been estimated to be about 0.02–0.025 μg (9). It appears that inhalation is the most important route of vinyl chloride intake, although drinking water may contribute a substantial portion of daily intake where PVC piping with a high residual content of vinyl chloride monomer is used in the distribution network, in particular, in hot countries or where drinking water is from wells near sites contaminated with chlorinated hydrocarbons. Boiling drinking water reduces the risk of contamination of water with vinyl chloride (9).

HEALTH EFFECTS

A wide variety of deleterious health effects have been attributed to vinyl chloride exposure. However, prior to the 1974 discovery at the B. F. Goodrich plant in Kentucky, where the factory’s doctor linked vinyl chloride exposure to angiosarcoma, most toxicologists believed that vinyl chloride did not pose a serious health risk. In fact, prior to 1974, the “safe” exposure level was set at 200 mg/L, rather than the 5 mg/L and under levels, which are standard for most jurisdictions today. The discovery in 1974 set in motion a series of events that led government, industry, labor, and science to the realization of the hazards of vinyl chloride exposure. Subsequently, industry, labor, and government all supported research into the biological consequences of vinyl chloride exposure (10).

Short-Term Exposure

Breathing high levels of vinyl chloride leads to feelings of dizziness and sleepiness. These effects occur within 5 min at about 10,000 mg/L of vinyl chloride. One can easily smell vinyl chloride at this concentration and one can rapidly recover from these effects by breathing fresh air (8). Limited exposure to vinyl chloride may also induce a feeling of dullness, nausea, and headache; irritate the eyes, mucous membranes, and respiratory tract; and have neurotoxic and dermatological effects. Escaping compressed gas or liquid can cause frostbite or irritation of the skin and eyes (1). It is difficult to clearly define the duration of a short-term or limited exposure, and its effect will almost surely differ from person to person (10).

Long-Term Exposure

There are many possible consequences of long-term exposure to vinyl chloride. These include various forms of liver disease such as liver sarcoma and liver carcinoma, as well as diseases in other tissues such as lung cancer, digestive system cancers, cardiovascular disease, genetic defects, and even melanoma. In the case of angiosarcoma, which appears to be the signature disease of long-term vinyl chloride exposure, the average length of exposure to vinyl chloride necessary to trigger angiosarcoma development has been estimated at 18 years, though a recent study has documented workers with as little as 3.5 and 4 years of exposure contracting angiosarcoma (10). Another condition attributed to prolonged vinyl chloride exposure is a cold-induced pallor of the extremities, referred to as Raynaud’s syndrome. A potentially serious side effect of chronic exposure to vinyl chloride, even at low doses, is associated with the body’s defense mechanisms. More medical information is contained in the ATSDR Medical Management Guidelines report (12).

Many of the health effects listed above are specific to workers in VC and PVC factories. These include sclero dermiform lesions, acroosteolysis, Raynaud’s syndrome, and liver function abnormalities (10). Some men who work with vinyl chloride have complained of a lack of sex drive. Results of studies in animals show that long-term exposure may damage the sperm and testes. Some women who work with vinyl chloride have had irregular menstrual periods. Some have developed high blood pressure during pregnancy. Studies of women who live near vinyl chloride manufacturing plants could not prove that vinyl chloride causes birth defects. Studies using pregnant animals show that breathing vinyl chloride may harm their unborn offspring. Animal studies also show that vinyl chloride may cause increased numbers of miscarriages early in pregnancy. It may also cause decreased weight and delayed skeletal development in fetuses. The same very high levels of vinyl chloride that caused these fetal effects also caused adverse effects in the pregnant animals (8).

Mutation and Cancer

Vinyl chloride has been shown to be mutagenic in almost all assay systems in which it has been tested. Phylogenetically, vinyl chloride has been shown to be mutagenic in all life forms tested including bacteria, yeast, insects, plants, mammalian cells in culture, human cells, and whole animals. In fact, vinyl chloride is currently used as the positive control substance in many studies. Vinyl chloride is known to induce almost a full spectrum of mutations including several types of point mutations and chromosome aberrations. Studies have indicated that certain chromosomes elicit a greater sensitivity to vinyl chloride than most of the chromosomes, suggesting some degree of specificity in the action of vinyl chloride (10).

Results from several studies suggest that breathing air or drinking water containing low levels of vinyl chloride may increase the risk of getting cancer. The monomer of vinyl chloride is a potent human carcinogen while only a weak to moderate animal carcinogen. Unlike many carcinogens, the carcinogenic metabolite(s) of vinyl chloride is considered a complete carcinogen; that is, it does not need a cocarcinogen to produce cancer. In rodents, the mechanism by which vinyl chloride and its metabolites transform normal cells into the neoplastic state appears to initially follow the pathway involved in the development of genotoxic effects (10). The primary cancer produced by vinyl chloride is its signature disease, angiosarcoma. Sarcomas of the liver are extremely rare and have been strongly linked to vinyl chloride exposure. Evidence also exists, however, that links vinyl chloride to hepatocellular carcinoma, cancers of the digestive system, lung cancer, brain cancer, and melanoma. Many cohort and epidemiology studies have attempted to correlate exposure to vinyl chloride with the abovementioned cancers. In almost all cases, the clearest correlation is with angiosarcoma. Studies indicate that induction of angiosarcoma generally requires a fairly extensive exposure (∼18 years) to vinyl chloride. Based on at least one calculation, the latency period, the time from the initial exposure to the onset of the disease, for angiosarcoma is approximately 22.6 years. From this calculation, it has been estimated that by the year 2010 there will be another 500 cancer-related mortalities in North America and approximately 2000 more in Europe due to previous vinyl chloride exposure (10).

DETECTION OF VINYL CHLORIDE

Biological Samples

The analytical method used to analyze for the presence of vinyl chloride in biological samples is separation by gas chromatography (GC) combined with detection by mass spectrometry (MS), flame ionization detector (FID), or electron capture detector (ECD). Vinyl chloride and/or its metabolite, thiodiglycolic acid, have been detected in breath, urine, blood, and tissues. Breath samples can be concentrated by cryogenic trapping. The two methods most commonly used to prepare liquid and solid samples are concentration by a purge-and-trap technique or headspace analysis. Concentration not only increases the sensitivity but also, in certain instances, may decrease the sample separation time prior to quantization (1).

Vinyl chloride can be determined in exhaled air by concentration with a multistage cryogenic trapping system followed by thermal desorption using GC/FID, GC/ECD, and GC/MS (1). Sensitivity is in the low-ppb range. One limitation of this method is its reduced ability to detect vinyl chloride when air concentrations in the workplace are below 50 mg/L (1).

Environmental Samples

Analysis of environmental samples is similar to that of biological samples. The most common methods used to detect vinyl chloride in environmental samples are GC/MS, GC/ECD, and GC/FID. Concentration of samples is usually done by sorption on solid sorbent for air and by the purge-and-trap method for liquid and solid matrices. Alternatively, headspace above liquid and solid samples may be analyzed without preconcentration (1).

Air. The primary method of analyzing vinyl chloride in air is GC combined with MS, ECD, or FID. Air samples are usually pumped through a sample collection column with Tenax-GC, coconut activated charcoal, or spherocarb (a carbon molecular sieve material) as the most common adsorbents. It has been noted that Tenax-GC displays poor retention for vinyl chloride when the compound is present in the very low ppb range (1).

Vinyl chloride is thermally desorbed from the collection column and concentrated on a cryogenic trapping column located on the gas chromatograph. Vapors are heat released from the trapping column directly to the gas chromatograph (1). Grab samples of air can also be obtained and preconcentrated on a cryogenic trapping column. With careful technique, precision is adequate, ranging from 5% to 20% (1).

Water. Vinyl chloride can be detected in drinking water, groundwater, wastewater, and leachate from solid waste. Analysis of vinyl chloride is done by purge-and-trap or headspace GC. The primary analytical method is separation by GC combined with MS, ECD, FID, Hall's electrolytic conductivity detector (HECD), or another type of halogen-specific detector (HSD). In most methods, vinyl chloride is liberated from the liquid matrix by purging with an inert gas and concentrated by trapping on a suitable solid sorbent. Vinyl chloride is thermally desorbed and back flushed onto the column of the gas chromatograph with an inert gas. Detection of vinyl chloride is generally achieved using HECD, HSD, or MS (1). Accuracy is greater than 98% and precision ranges from 11% to 25% for GC/HECD and GC/MS (1).

REGULATIONS

In September 1974, eight months after a correlation between angiosarcoma and vinyl chloride exposure was observed, the U.S. EPA established a national emissions standard for vinyl chloride which came into effect two years later when published as part of the U.S. Clean Air Act (October 1976) (10).

The U.S. EPA requires that the amount of vinyl chloride in drinking water not exceed 0.002 mg/L. It is also required that spills or accidental releases into the environment of 1 lb or more of vinyl chloride be reported to the EPA (1). The Occupational Safety and Health Administration (OSHA) set the maximum allowable level of vinyl chloride in workroom air during an 8-h workday in a 40-h workweek at 1 mg/L (1).

The EPA has set emission standards for vinyl chloride and PVC plants. The amount of vinyl chloride allowed to be emitted varies depending on the type of production and discharge system used. To limit intake of vinyl chloride via foods to levels considered safe, the Food and Drug Administration (FDA) regulates the vinyl chloride content of various plastics. These include plastics that carry liquids and plastics that come into contact with food. The limits for vinyl chloride content vary depending on the nature of the plastics and its use (4). However, the FDA has set the allowable level for vinyl chloride at 0.002 mg/L in bottled water (2).

GROUNDWATER REMEDIATION

A variety of technologies can be used for VC removal and destruction. Selected technologies are described below.

In Situ Biological Treatment

In situ bioremediation techniques are destruction techniques directed toward stimulating the microorganisms to grow and use the contaminants as a food and energy source by creating a favorable environment for the microorganisms. Generally, this means providing some combination of oxygen (electron acceptor), nutrients, and moisture, controlling the temperature and pH for aerobic conditions, providing a carbon source as a food source and nutrients and moisture, and controlling the temperature and pH for anaerobic conditions.

Contaminants are usually destroyed and little to no residual treatment is required. Some compounds break down and the breakdown products are more toxic than the original product: for example, PCE and TCE, which break down to form the more toxic vinyl chloride. For *in situ* applications, the vinyl chloride may be mobilized in groundwater if no control techniques are used.

Enhanced Biodegradation. The rate of bioremediation of organic contaminants by microbes is enhanced by increasing the concentration of electron acceptors and nutrients in groundwater, surface water, and leachate. Oxygen is the main electron acceptor for aerobic bioremediation. Nitrate, manganese (4+), iron (3+), sulfate, and ultimately carbon dioxide serve as an alternative electron acceptor under anaerobic and reducing conditions.

Although vinyl chloride can be released into the environment as a liquid product, vinyl chloride is also a degradation by-product of the anaerobic reductive dechlorination process that starts with tetrachloroethylene (PCE) (4 chlorine atoms) and proceeds to trichloroethylene (TCE) (3 chlorine atoms) and moving through *cis*-1,2-dichloroethene (*cis*-1,2-DCE) (2 chlorine atoms) before creating vinyl chloride (VC) (1 chlorine atom) by primarily biological reactions. Other pathways to VC include limited biological reactions from 1,1-dichloroethene (1,1-DCE) and *trans*-1,2-dichroloethene (*trans*-1,2-DCE) (13). For chlorinated solvents, the carbon source for anaerobic degradation is a food source and the contaminant is the anaerobic electron acceptor.

Once at VC, the aerobic breakdown of VC into ethene and later into ethane is best done using oxygen as the electron acceptor under aerobic conditions. Oxygen enhancement for VC breakdown can be achieved by a variety of methods: air sparging, pure oxygen gas infusion using mass transfer tools in monitoring wells, and injection of peroxide compounds [hydrogen peroxide (H_2O_2), calcium peroxide (CaO_2), and, to a lesser extent, magnesium peroxide (MgO_2)]. Even ozone (O_3) can be used as an oxygen source. Vinyl chloride has been observed to degrade four times faster under aerobic conditions than under anaerobic conditions (14).

In Situ Physical/Chemical Treatment

Physical/chemical treatment uses the physical properties of the contaminants or the contaminated medium to destroy (i.e., chemically convert) or separate the contamination. Available *in situ* physical/chemical treatment technologies include chemical oxidation, air sparging, bioslurping, directional wells, dual phase extraction, thermal treatment, hydrofracturing, in-well air stripping, passive/reactive treatment walls, and soil vapor extraction (SVE).

Chemical Oxidation. The chemical oxidants most commonly employed to date include rapidly reacting hydrogen peroxide, Fenton's reagent, and ozone. Slower oxidizers include potassium permanganate and sodium persulfate. These oxidants have been able to cause the complete chemical destruction of many toxic organic chemicals including vinyl chloride. In general, the oxidants have been capable of achieving high treatment efficiencies (>90%) for chlorinated solvents. Field applications have clearly affirmed that matching the oxidant and *in situ* delivery system to the contaminants of concern and the site conditions is the key to successful implementation and achieving performance goals. Bench testing is recommended prior to performing pilot studies in the field. Oxidant delivery systems often employ vertical or horizontal injection wells and sparge points, probe rods, or lances under pressures from 40 psi up to 5000 psi. Reaction temperatures, pH, DO, and other field variables are used to help monitor progress of the subsurface reactions.

Passive Reactive Treatment Walls. Passive reactive treatment walls are barriers that allow the passage of water while causing the degradation or removal of contaminants. A permeable reaction barrier (PRB) is installed across the flow path of a contaminant plume, allowing the water portion of the plume to passively move through the wall. These barriers allow the passage of water while prohibiting the movement of contaminants by employing such agents as zero-valent metals, chelators (ligands selected for their specificity for a given metal), sorbents, and microbes. The contaminants will either be degraded or retained in a concentrated form by the barrier material. The wall could provide permanent containment for relatively benign residues or provide a decreased volume of the more toxic contaminants for subsequent treatment. An iron treatment wall consists of iron granules or other iron-bearing minerals for the treatment of chlorinated contaminants such as TCE, DCE, and vinyl chloride. As the iron is oxidized, a chlorine atom is removed from the compound by one or more reductive dechlorination mechanisms, using electrons supplied by the oxidation of iron. The iron granules are dissolved by the process, but the metal disappears so slowly that the remediation barriers can be expected to remain effective for many years, possibly even decades.

Soil Vapor Extraction. Soil vapor extraction (SVE) is an *in situ* unsaturated (vadose) zone soil remediation technology in which a vacuum is applied to the soil to induce the controlled flow of air and remove volatile and some semivolatile contaminants from the soil. The gas leaving the vadose zone may be treated to recover or destroy the contaminants, depending on local and state air discharge regulations.

For the soil surface, plastic sheeting or membranes are often placed over the soil surface to prevent short circuiting and to increase the radius of influence of the wells. Groundwater depression pumps may be used to reduce groundwater upwelling induced by the vacuum or to increase the depth of the vadose zone. Air injection is effective for facilitating extraction of deep contamination, contamination in low permeability soils, and contamination in the saturated zone. The duration of operation and maintenance for *in situ* SVE is typically medium to long term. Dual phase extraction is a variation of the SVE–groundwater extraction approach that can be used successfully with VC.

Ex Situ Physical/Chemical Treatment

Once the VC is extracted using SVE or groundwater extraction, the contaminant must be treated and destroyed on the surface.

Advanced Oxidation Processes. Advanced oxidation processes including ultraviolet (UV) radiation, ozone, and/or hydrogen peroxide are used to destroy organic contaminants as water flows into a treatment tank. UV oxidation is a destruction process that oxidizes organic and explosive constituents in wastewater by the addition of strong oxidizers and irradiation with UV light. Oxidation of target contaminants is caused by direct reaction with the oxidizers, UV photolysis, and through the synergistic action of UV light, in combination with ozone and/or hydrogen peroxide. If complete mineralization is achieved, the final products of oxidation are carbon dioxide, water, and occasionally salts. Easily oxidized organic compounds such as those with double bonds (e.g., vinyl chloride) are rapidly destroyed in UV/oxidation processes.

The main advantage of UV oxidation is that it is a destruction process, as opposed to air stripping or carbon adsorption, for which contaminants are extracted and concentrated in a separate phase. UV oxidation processes can be configured in batch or continuous flow modes, depending on the residence treatment time required. The UV oxidation process is generally done with low pressure UV lamps operating at 65 W of electricity for ozone systems and lamps operating at 15–60 kW for hydrogen peroxide systems (15).

Air Stripping. Air stripping is a full-scale technology in which volatile organics are partitioned from groundwater by greatly increasing the surface area of the contaminated water exposed to air. Henry's law constant is used to determine whether air stripping will be effective and organic compounds with constants greater than 0.01 $\text{atm}\cdot\text{m}^3/\text{mol}$ are considered amenable to stripping. Vinyl chloride has a constant of 1.2 $\text{atm}\cdot\text{m}^3/\text{mol}$. Types of aeration methods include packed towers, diffused aeration, tray aeration, and spray aeration. The typical packed tower air stripper includes a spray nozzle at the top of the tower to distribute contaminated water over the packing in the column, a fan to force air countercurrent to the water flow, and a sump at the bottom of the tower to collect decontaminated water. Packed tower air strippers are installed as permanent installations on concrete pads or as temporary units on a skid or a trailer. Aeration tanks strip volatile compounds by bubbling air into a tank through which contaminated water flows. A forced air blower and a distribution manifold are designed to ensure air–water contact without the need for any packing materials. The baffles and multiple units ensure adequate residence time for stripping to occur (15).

Containment

Containment technologies are often performed to prevent, or significantly reduce, the migration of contaminants in soils or groundwater. Containment is necessary whenever contaminated materials are to be buried or left in place at a site, such as at a landfill. Available containment technologies include physical/biological barriers.

Physical Barriers. Physical barriers (or slurry walls) are used to contain contaminated groundwater, divert contaminated groundwater from the drinking water intake, divert uncontaminated groundwater flow, and/or provide a barrier for the groundwater treatment system. Most slurry walls are constructed of a soil, bentonite, and water mixture. Unfortunately, most barriers do not provide complete containment unless active hydraulic control is maintained using a groundwater extraction system.

BIBLIOGRAPHY

1. ATSDR (1997). Agency for Toxic Substances and Disease Registry. *Toxological Profile for Vinyl Chloride*, U.S. Department of Health and Human Services, Public Health Service. Available: http://www.atsdr.cdc.gov/toxprofiles/tp20.pdf.
2. VC (2002). *Vinyl Chloride*, Tenth Report on Carcinogens. Available: http://ehp.niehs.nih.gov/roc/tenth/profiles/s186viny.pdf.
3. BMZ (1992). German Federal Ministry for Economic Cooperation and Development. *Environmental Handbook—Documentation on Monitoring and Evaluating Environmental Impacts*. Available: http://www.gtz.de/uup/publica/English/vol369.htm.
4. Val Iyioke, I. (2003). *Vinyl Chloride*. Available: http://www.iet.msu/Tox_for_Journ/Chemicals/Vinyl.htm.
5. Pankow, J.F. and Cherry, J.A. (1996). *Dense Chlorinated Solvents, and Other DNAPLs in Groundwater: History, Behavior, and Remediation*. Waterloo Press, Guelph, Ontario, Canada, p. 522.

6. NAOC (1998).
7. Doctorflood (2004). *Biodegradation Processes for Chlorinated Solvents*. Available: http://doctorflood.rice.edu/envi518/Handouts/Lisbon/04_chlor.ppt.
8. Eco-USA (2003). *Vinyl Chloride*. Available: http://www.eco-usa.net/toxics/vinylcl.shtml.
9. VC2 (2003). *Vinyl Chloride*. Available: http://www.who.int/docstore/water_sanitation_health/GDWQ/draftchemicals/vinylchloride2003.pdf.
10. Salamone, M.F. and Westlake, G. (1998). *Vinyl Chloride: An Annotated Bibliography with Emphasis on Genotoxicity and Carcinogenicity*, Ministry of the Environment, Ontario, Canada. Available: http://www.ene.gov.on.ca/envision/techdocs/3698e.pdf.
11. Yamamoto, K., Fukushima, M., Kakutani, N., and Tsurho, K. (2001). Contamination of vinyl chloride in shallow urban rivers in Osaka, Japan. *Water Res.* **35**: 561–566.
12. ATSDR (2004). Agency for Toxic Substances and Disease Registry. Medical management guidelines (MMGs) for vinyl chloride. Available: http://www.atsdr.cdc.gov/mmg.html.
13. Suthersan, S. (2002). *Natural and Enhanced Remediation Systems*. Lewis Publishers/CRC Press, Boca Raton, FL, p. 419.
14. Brady, P.V. and Brady, W.D. (2000). Monitored natural attenuation. In: *McGraw-Hill Standard Encyclopaedia of Environmental Science, Health and Technology*, J. Lehr (Ed.). McGraw-Hill, New York, pp. 14.1–14.11.
15. Chhetry, S., Liou, R., Lloyd, T., Peck, J., and Van Duren, J. (2002). *Remediation Technologies Screening Matrix and Reference Guide*, 4th Edn. Federal Remediation Technologies Roundtable. Available: http://www.frtr.gov/matrix2/top_page.html.

GROUNDWATER AND URANIUM: CHEMICAL BEHAVIOR AND TREATMENT

James A. Jacobs
Ricardo Smalling
Environmental Bio-Systems, Inc.
Mill Valley, California

The element uranium, symbol U, is a chemically reactive radioactive metallic element. It is the main fuel used in nuclear reactors. Uranium has atomic number 92, which places it in the actinide series of the periodic table. The German chemist Martin Heinrich Klaproth discovered uranium in 1789 in pitchblende. He named it after the planet Uranus. Uranium was first isolated in the metallic state in 1841and the radioactive properties were first demonstrated in 1896 by the French physicist Antoine Henri Becquerel (1).

Uranium melts at about 1132 °C and boils at about 3818 °C. It has a specific gravity of 19.05 at 25 °C and the atomic weight of the element is 238.03. Uranium has three crystalline forms, of which the one that forms at about 770 °C is malleable and ductile. Uranium is soluble in hydrochloric and nitric acids, and it is insoluble in alkalis. It displaces hydrogen from mineral acids and from the salt solutions of such metals as mercury, silver, copper, tin, platinum, and gold. When finally divided, it burns readily in air at 150–175 °C. At 1000 °C, uranium combines with nitrogen to form a yellow nitride (1).

FATE AND TRANSPORT

Uranium has multiple oxidation states of three, four, five, and six. The hexapositive compounds include uranyl trioxide, UO_3, and uranyl chloride, UO_2Cl_2. Uranium tetrachloride, UCl_4, and uranium dioxide, UO_2, are examples of the tetrapositive, or uranous, compounds. Uranous compounds are usually unstable; they revert to the hexapositive form when exposed to excess air. Uranyl salts, such as uranyl chloride, may decompose in the presence of strong light and organic matter. *In situ* soil moisture pH greatly influences the formation of uranium surface complexes (2). The carbon dioxide (CO_2) partial pressure in soil gas is a dominant factor in controlling the pH of soil moisture.

Uranium never occurs naturally in the free state but is found as an oxide or complex salt in minerals such as pitchblende and carnotite. It has an average concentration in the crust of the earth of about 2 ppm (parts per million) and, among the elements, ranks about 48th in natural abundance in crustal rocks. Pure uranium consists of more than 99% of the isotope uranium-238, 0.72% of the fissile isotope uranium-235, and a trace of uranium-234, formed by radioactive decay of uranium-238. Among the artificially produced isotopes of uranium are uranium-233, uranium-237, and uranium-239. Isotopes ranging from mass number 222 to 242 are known (3).

Uranium has a generally low solubility and will not easily reach groundwater. It is also easily adsorbed to soil particles. Uranium oxide will precipitate from uranium-bearing groundwaters as they enter a reducing environment.

Uranium is unstable and, as it decays, it releases radiation and forms decay products. Uranium-238 decay products include uranium-234, radium-226 (both of which are solids), and radon-222, a gas. Radon is soluble but does not remain in water. It can be released by simple agitation. All the decay products are radioactive as well. The half-life for uranium-238 is about 4.5 billion years; for uranium-235 it is 710 million years, and for uranium-234 it is 250,000 years.

PROCESSING OF URANIUM

There are two major methods of producing uranium. In the classical procedure for extracting uranium, pitchblende is broken up and mixed with sulfuric and nitric acids. Uranium dissolves to form uranyl sulfate, UO_2SO_4; radium and other metals in the pitchblende ore are precipitated as sulfates. With the addition of sodium hydroxide, uranium is precipitated as sodium diuranate, $Na_2U_2O_7 \cdot 6H_2O$, known also as the yellow oxide of uranium. To obtain uranium from carnotite, the ore is finely ground and treated with a hot solution of caustic soda and potash to dissolve out uranium, radium, and vanadium. After the sandy matrix is washed away, the solution is treated with sulfuric acid and barium chloride. A caustic alkali solution added to the remaining clear liquid precipitates the uranium and radium in concentrated form. These classical methods of extracting uranium from its ores have been replaced in current practice by such procedures as solvent extraction, ion exchange, and volatility methods. *In situ* leaching is another method that is used in the United States. The artificial isotope uranium-233 is produced from thorium, another actinide metal (1).

Uranium oxide precipitates from uranium-bearing groundwaters as they enter a reducing environment. It can be mobilized (redissolved) *in situ* by oxygenated leach solution.Uranium ores are widely distributed throughout the world. Deposits of pitchblende, the richest uranium ore, are found chiefly in Canada, the Democratic Republic of the Congo (DRC, formerly Zaire), and the United States. Most of the uranium mined in the United States is obtained from carnotite occurring in Colorado, Utah, New Mexico, Arizona, and Wyoming. A mineral called coffinite, discovered in 1955 in Colorado, is a high grade ore containing nearly 61% uranium. Coffinite deposits were found subsequently in Wyoming and Arizona and in several foreign countries.

USES OF URANIUM

After the discovery of nuclear fission, uranium became a strategic metal, and its uses were at first restricted mainly to the production of nuclear weapons. In 1954 the United States government relaxed controls to permit leasing of uranium enriched in the isotope uranium-235 to various private and foreign agencies for the development of nuclear power. Peacetime applications were discussed at the three International Conferences on the Peaceful Uses of Atomic Energy held in Geneva, Switzerland, in 1955, 1958, and 1964.

The potentiality of uranium as a vast source of industrial power became apparent with the launching in 1954 of the first nuclear-powered submarine, the USS *Nautilus*. By 1989, 112 nuclear power plants in the United States produced more than 101,000 megawatts electric, MW(E). In addition, there are 316 plants in 40 countries outside the United States, which produced more than 213,000 MW(E). The first such U.S. plant, which began operations at Shippingport, Pennsylvania, generates 60,000 kW and requires about 15 lb of uranium-235 per month. Conventional plants producing 60,000 kW consume about 40 million lb of coal per month. Problems of uranium scarcity, plant safety, and storage of radioactive uranium and plutonium waste products, however, have prevented the full realization of nuclear energy's potential.

Uranium-235 is used in nuclear weapons and nuclear reactors. Depleted uranium (natural uranium in which almost all of the uranium-235 has been removed) is used to make ammunition for the military, guidance devices and compasses, radiation shielding material, and X-ray targets. Uranium dioxide is used to extend the lives of incandescent lamps used for photography and motion pictures. Very small amounts of other uranium compounds are used in photography for toning, in the leather and wood industries for stains and dyes, and in the wool industry. Uranium has also been used in the past in ceramics as a coloring agent.

HEALTH EFFECTS AND TOXICITY

Natural uranium releases alpha particles and low levels of gamma rays. The alpha particles can travel only short distances and cannot penetrate human skin. Gamma radiation, however, can penetrate the body. In general, uranium outside the body is less harmful than inside the body.

Uranium can enter the body in several ways. It is inhaled or swallowed or enters through cuts in the skin. About 99% of the uranium ingested in food or water will leave a person's body in the feces; however, a small quantity may enter the bloodstream where it is eventually filtered by the kidney.

Although uranium is radioactive and can cause cancer, its primary toxic effect when consumed in water is that of a heavy metal. Heavy metals, like uranium, lead, cadmium, and arsenic, are deposited in the kidneys and can cause irreparable damage to the chief filtering mechanism of the body (4–6).

METHODS OF REMEDIATION

Several methods of remediation have been tried with sites contaminated with uranium. These include pump and treat systems for groundwater and removal of contaminated soils. In light of the high potential for negative health effects on workers, these methods are not favored. Three *in situ* remedial approaches that do not adversely affect worker health are phytoremediation, chemical precipitation, and bioremediation.

Phytoremediation

More recent methods include phytoremediation. Rhizofiltration technology uses the plant roots to uptake the contaminant. In 1996, a rhizofiltration study was performed using hydroponically grown plants in uranium-impacted groundwater. The original study was performed in a greenhouse under controlled conditions. Sunflower plants were found to have the best uptake of uranium in the laboratory studies. In the study, the uranium-impacted water was introduced into a recirculation loop with inlet concentrations of 80–350 ppb. After treatment, the outlet treatment water was reduced to less than 5 ppb, and most of the concentrations were below 0.5 ppb. The uranium contaminant was recovered from the sunflower plant roots, with no translocation of the uranium to the plant shoots. The roots were disposed of as radioactive waste (7). Additional uranium speciation and plant uptake studies are described in detail in References 8–10.

Chemical Precipitation

Long-term sequestration of uranium by *in situ* U(VI) phosphate mineral precipitation has been documented (11). In one case at Coles Hill, Virginia, uranium remains essentially immobile and mineralized during the oxidation and chemical weathering of the uranium ore. The groundwater at the site is oxic, has a pH of 6, and falls within the uranyl phosphate stability field showing that the Coles Hill U(VI) phosphates may precipitate rapidly (weeks) relative to groundwater flow. In this location, the groundwater contains 15 μg/L or parts per billion (ppb) of dissolved uranium in the groundwater. These levels are well below the EPA maximum contaminant level (MCL) of 30 μg/L (4). In similar hydrologic conditions, phosphate-based remediation of uranium could be considered. Other uranium compounds may also provide long-term stabilization of uranium-impacted groundwater.

Bioremediation

Bioremediation has been used in uranium treatment. Microbial reduction of soluble U(VI) to insoluble U(IV) can immobilize uranium from contamianted groundwater. This remedial method may be added to soil washing for concentrating uranium from contaminated soils. Laboratory studies (12) have demonstrated that U(VI) is reduced concurrently with Fe(III) reduction and molecular studies have suggested that, in most instances, metal reduction can be attributed to the activity of the bacteria Geobacteraceae. Currently, field studies are under way to evaluate *in situ* bioremediation options for uranium.

BIBLIOGRAPHY

1. United States Environmental Protection Agency (U.S. EPA) (2004). Uranium, updated Sept. 21, 2004. Available: http://www.epa.gov/radiation/radionuclides/uranium.htm.
2. Langmuir, D. (1997). *Aqueous Geochemistry*, Simon & Schuster, Upper Saddle River, NJ.
3. Seborg, G.T. (2004). Uranium. *Microsoft® Encarta® Online Encyclopedia*. Available: http://encarta.msn.com/encyclopedia_761557628/Uranium.html.
4. United States Environmental Protection Agency (U.S. EPA) (2002). Radionuclides in drinking water, updated November. Available: http://www.epa.gov/safewater/standard/pp/radnucpp.html.
5. United States Environmental Protection Agency (U.S. EPA) (2004). Use of uranium drinking water standards under 40 CFR 141 and 40 CFR 192 as remediation goals for groundwater at CERCLA sites. Available: www.epa.gov/swerffrr/documents/uranium_drinking_water_standards.htm.
6. World Health Organization (WHO) (2003). Fact sheet no. 257 on depleted uranium; revised January. Available: http://www.who.int/mediacentre/factsheets/fs257/en/.
7. Lehr, J. (2000). Phytoremediation: a Promising path to the elimination of heavy metals and radionuclides in groundwater and soil. In: *McGraw-Hill Standard Encylopedia of Environmental Science, Health and Technology*, J. Lehr (Ed.). McGraw-Hill, New York, pp. 14.26–14.29.
8. Ebbs, S.D., Norvell, W.A., and Kochian, L.V. (1998). The effect of acidification and chelating agents on the solubilization of uranium from contaminated soil. *J. Environ. Qual.* **27**: 1486–1494.
9. Ebbs S.D., Brady, D.J., and Kochian, L.V. (1998). Role of uranium speciation in the uptake and translocation of uranium by plants. *J. Exp. Bot.* **49**: 1183–1190.
10. Ebbs, S., Brady, D., Norvell, W., and Kochian, L. (2001). Uranium speciation, plant uptake, and phytoremediation. *Pract. Period. Hazard. Toxic Radioactive Waste Manage.* **5**: 130–135.
11. Jerden, J.L. Jr. and Sinha, A.K. (2002). Long-term stabilization of uranium by U(VI) phosphate mineralization: a model for uranium remediation, Geological Society of America (GSA) 2002 Denver Annual Meeting, October 27–30, Paper No. 138-8.
12. Holmes, D.E., Finneran, K.T., O'Neil, R.A., and Loverly, D.R. (2002). Enrichment of members of the family Geobacteracea associated with stimulation of dissimilatory metal reduction in uranium-contaminated aquifer sediments. *Appl. Environ. Microbiol.* May: 2300–2306.

GROUNDWATER AND MERCURY: CHEMICAL BEHAVIOR AND TREATMENT

James A. Jacobs
Silvia Johnson
Environmental Bio-Systems, Inc.
Mill Valley, California

INTRODUCTION

Mercury is rarely found as a pure metal. The chief ore of mercury is cinnabar, also called vermilion, a bright red mineral composed of mercury sulfide. Metallic mercury is recovered by heating cinnabar ore in air and condensing the mercury vapor. It was originally found primarily in Spain and Italy. Mercury's chemical symbol comes from the Greek word *hydrargyrum*, which means "liquid silver." Mercury has an atomic number of 80, atomic weight of 200.59, and a density of 13.5336 g/cm^3. It is the only metal element that is a liquid at room temperature. Mercury is not an active element, combining with other elements only with some difficulty. It is therefore a *noble* metal, like gold, silver, and the platinum group, which are all highly resistant to chemical reactions and oxidation. Mercury accumulates most efficiently in the aquatic foodweb and is highly toxic to humans (1).

There are three forms of mercury in the environment: elemental, inorganic, and organic mercury. All forms of mercury are toxic. Elemental mercury occurs naturally in three valence states: elemental (Hg^0), monovalent mercurous (Hg^{1+}), and the divalent mercuric (Hg^{2+}). Elemental mercury is the most stable form and is only slightly water soluble. Both mercuric and mercurous mercury are thermally unstable and readily decompose to elemental mercury. Vapors of elemental mercury can occur at room temperature, presenting a hazard if spills occur.

Inorganic mercury compounds contain ionic mercury usually in a salt formation (e.g., mercuric chloride).

Organic mercury compounds can be chemically synthesized or biologically converted from mercury compounds by bacteria (e.g., methylmercury). Some organic mercury compounds are water soluble and capable of transport in the aquatic food chain through the process of bioaccumulation.

Although mercury does not combine easily with other elements to form compounds, it will, however, dissolve other metals to form a solution with two or more metals. These alloys are called amalgams. Dentists have used silver–mercury amalgams, consisting of silver powder dissolved in mercury, to fill teeth. Mercury is used in mining applications to recover gold from gold ore. In this application, the mercury containing dissolved gold is heated and vaporized to leave behind the gold.

HISTORY AND USES

Spain has mercury mines that have been operating continuously for over 2000 years. Mercury was well known to the ancient Chinese. Samples of the metal have been found in Egyptian tombs dating back to 1500 B.C. Mercury was an important chemical for medieval alchemists and was suspected of being an element from which all matter is made. When a solid, mercury was thought to be gold. It was also called “quicksilver” because it is liquid and silver-white in appearance. Mercury is not usually found free in nature and is primarily obtained from the mineral cinnabar (HgS). Spain and Italy produce about half of the world’s supply of mercury. Mercury has been used in mining for hundreds of years in the separation of valuable metals, such as gold, from placer deposits (1).

Mercury has been used to produce chlorine and caustic soda; in wiring devices and switches; for electric lights; for measuring and control instruments (such as blood pressure gauges, barometers, and thermometers); and in other industrial applications.

HEALTH RISKS

All forms of mercury are highly toxic to humans and animals. Its toxicity was recognized in the nineteenth century when hat makers, who used a mercury compound in making hats, began to exhibit damage to their nervous systems. This is the origin of the phrase “mad as a hatter.” Workers in mercury-battery factories, mercury processing plants, power plant workers, and dentists are in occupations associated with mercury exposure to mercury liquid or dust. The U.S. EPA has performed a detailed assessment of exposure using various computer models (2).

Mercury concentrations in air are usually low and of little direct concern for exposure to humans (3). From the standpoint of public health, when mercury enters the water, it is the most serious problem because of the manner in which mercury accumulates in the living tissue of animals. Mercury combines chemically with enzymes in the body, causing them to lose the ability to act as catalysts for vital body functions (1).

When mercury is deposited into waterways, bacteria and algae in the water convert inorganic mercury into organic methylmercury (CH_3Hg). Fish that eat these bacteria and algae accumulate the mercury in their tissues, and the larger fish up the food chain that feed on fish with mercury in their tissues accumulate even higher levels of mercury (biomagnification) (4).

Elemental Mercury

The major health effect from elemental mercury results from inhalation of mercury vapor. Acute (short-term) inhalation exposure to high levels of elemental mercury in humans results in central nervous system effects such as hallucinations and delirium. Gastrointestinal and respiratory effects have also been noted. Studies are inconclusive regarding elemental mercury and cancer.

The central nervous system is the major organ affected by chronic (long-term) exposure to elemental mercury. Effects noted include increased excitability, irritability, insomnia, severe salivation, loss of teeth, and tremors. Chronic exposure to elemental mercury also affects the kidneys (2).

Inorganic Mercury Compounds

The major risk from inorganic mercury occurs through ingestion. Acute oral exposure to inorganic mercury compounds has been known to produce a metallic taste in the mouth, nausea, vomiting, and severe abdominal pain. The primary effect from chronic exposure to inorganic mercury is kidney damage.

Organic Mercury Compounds

The major risk posed by organic mercury occurs through ingestion. The most important organic mercury compound in terms of human exposure is methylmercury. Acute exposure to high levels of methylmercury can result in damage to the central nervous system, including blindness, deafness, intellectual deterioration, and death. No human studies are available on the carcinogenic effects of methylmercury.

Chronic methylmercury exposure occurs primarily through diet, with fish and fish products as dominant sources. The primary effect from chronic exposure to methylmercury is damage to the central nervous system, which produces malaise, numbness and tingling, blurred vision, deafness, speech difficulties, and constriction of the visual field. Infants born to women who ingest high levels of methylmercury may exhibit blindness, mental retardation, lack of coordination, and cerebral palsy. The U.S Environmental Protection Agency (EPA) currently advises pregnant women not to eat shark, swordfish, king mackerel, tilefish, and tuna.

What is known about the health effects of mercury has come from several accidental exposures, including two high dose episodes that occurred in Japan and Iraq. From 1953 to 1960, manufacturing plants in Japan discharged methylmercury into the local waterways. In 1956, the population living near Minamata Bay in southwestern Japan suffered severe neurotoxic effects from ingesting fish from the bay that had accumulated high levels of mercury because a chemical manufacturing plant had dumped tons of mercury into the water. Levels of mercury in Minamata Bay were found to be as high as 550 ppm (aquatic biota impact level for mercury is 0.000025 ppm or 0.025 μg/L). In the 1970s, over 6000 people were hospitalized in Iraq after eating imported wheat seeds that had been treated with a methylmercury fungicide. Several hundred people died, although because the incident occurred in a rural area the exact magnitude of the accident is not known (2).

FATE AND TRANSPORT

When liquid mercury spills, it forms drops that can accumulate in the smallest of spaces and then emits vapors into the air. Once in surface water, mercury enters a complex cycle in which one form can be converted to another. Mercury can easily be adsorbed onto small particles of matter. Mercury attached to particles can settle on the sediments, where it can diffuse into the water column, be resuspended, be buried by other sediments, or be methylated. Methylmercury can enter the food chain, or it can be released back into the atmosphere by volatilization (3).

The concentration of dissolved organic carbon (DOC) and pH have a strong effect on the ultimate fate of mercury in the ecosystem. Studies have shown that for the same species of fish taken from the same region, increasing the acidity of the water (decreasing pH) and/or the DOC content generally results in higher mercury levels in fish, an indicator of greater net methylation. Higher acidity and DOC levels enhance the mobility of mercury in the environment, thus making it more likely to enter the food chain. More information is given in Reference 2.

Mercury and methylmercury exposure to sunlight (specifically ultraviolet light) has an overall detoxifying effect. Sunlight can break down methylmercury to Hg(II) or Hg(0), which can leave the aquatic environment and enter the atmosphere as a gas. Since mercury is an element, it is not biodegradable. Mercury is converted among its various forms through a range of abiotic and biogeochemical transformations and during atmospheric transportation. Although its form and availability to living organisms may change over time, mercury persists in the environment.

Nearly 8 million pounds of mercury were produced in the United States in 1986. From 1987 to 1993, according to the EPA's Toxic Chemical Release Inventory, releases of mercury to land and water totaled nearly 68,000 lb, primarily from chemical and allied industries (3).

REMEDIATION

Although mercury is a globally dispersed contaminant, it is not a problem everywhere. Pollution prevention controls and equipment to contain airborne mercury and mercury spills would greatly limit future mercury releases into the environment. Aside from grossly polluted environments, mercury is normally a problem only where the rate of natural formation of methylmercury from inorganic mercury is greater than the reverse reaction. Environments that are known to favor the production of methylmercury include certain types of wetlands, dilute low pH lakes in the northeast and north central United States, parts of the Florida Everglades, newly flooded reservoirs, and coastal wetlands, particularly along the Gulf of Mexico, Atlantic Ocean, and San Francisco Bay. The highest concentration of mercury in freshwater sediments in the world is found in a section of Berry's Creek in New Jersey, near the Meadowlands Sports Complex. Several defunct industrial plants, including one mercury plant, are parts of some Superfund environmental cleanup sites (4).

In highly polluted areas where mercury has accumulated through industrial or mining activities, natural processes may bury, dilute, or erode the mercury deposits, resulting in declines in concentration. In many relatively pristine areas, however, mercury concentrations have actually increased because atmospheric deposition has increased.

For instance, concentrations of mercury in feathers of fish-eating seabirds from the northeastern Atlantic Ocean have steadily increased for more than a century. In North American sediment cores, sediments deposited since industrialization have mercury concentrations about 3–5 times those found in older sediments (2). Some sites may have become methylmercury hot spots inadvertently through human activities. Although scientists from the United States Geological Survey (USGS) and elsewhere are beginning to unravel the complex interactions between mercury and the environment, a lack of information on the sources, behavior, and effects of mercury in the environment has impeded identification of effective management responses to the nation's growing mercury problem.

Mercury contamination of shallow soil is commonly remediated by excavation and importation of clean backfill, as in the example of the mercury contamination remediation at gas pipeline sites in Kansas (5). Off-site treatment or transportation to a licensed disposal facility may be part of the remedial action.

Bioremediation is not an option for mercury, as the metal is also toxic to most microbial colonies (6). The following treatment methods have been approved by the U.S. EPA for removing mercury from drinking water: coagulation/filtration, granular activated carbon, lime softening, and reverse osmosis (3).

Remediation of mercury has been treated using capping and containment strategies wherein the mercury is surrounded by low permeability (bentonite) or nonpermeable material (steel shoring, plastic sheeting) to prevent the migration of the mercury. Other techniques include *in situ* vitrification.

Mercury is one of the most difficult wastes to stabilize. Aboveground stabilization techniques employing cement, kiln dust, fly ash, or other materials have been used to treat mercury-contaminated soils (7). Sulfide chemistry has also been used to chemically precipitate mercury into mercury sulfide. Calcium polysulfide (CaS_5) has been used in an aboveground soil washing treatment of mercury-impacted soil and refractory bricks at the Red Devil Mine in Alaska. The project converted toxic mercury and lead compounds into insoluble mercury sulfide and lead sulfide precipitates (B. Graves, personal communication). Mercury with zinc was encapsulated using an alkali-activated slag matrix. The physical encapsulation and chemical fixation mechanisms were responsible for the immobilization of the mercury ions in the alkali-activated slag matrix (8).

SPECIFIC REGULATORY GUIDELINES

The United States Environmental Protection Agency (U.S. EPA) has established a Maximum Contaminant Level of Goal for mercury in drinking water at 2 μg/L or parts per billion (ppb) (3). This is the lowest level that water systems can reasonably attain, given current treatment technologies and health information.

Efforts by manufacturers of mercury-containing products, by government programs, and by solid waste management facilities have significantly reduced mercury entering the environment from products that contain it. For example, Hennepin County programs to keep mercury out of the waste stream, together with pollution control equipment on the county's waste-to-energy plant, have brought mercury emission levels down from 67% of the amount allowed under the Minnesota Pollution Control Agency permit, to under 6% of the permitted amount (9).

The EPA has issued several major studies on mercury to the U.S. Congress, including the *Mercury Study Report to Congress* in 1997 and *Utility Hazardous Air Pollutant Report to Congress* in 1998 (2,4). The agency has also taken steps to reduce human sources of mercury, including issuing regulations for the medical waste combustion, municipal waste combustion, and chloralkali production. The EPA has proposed regulations for industrial boilers, another large source of mercury. On December 15, 2003, the EPA announced a plan to significantly reduce mercury emissions from coal power plants through a cap and trade program, a market-based approach that worked successfully to reduce acidic precipitation from sulfur dioxide. The proposal would require utilities to cut emissions by 70% over the next 15 years. The EPA has also proposed a second option that is a traditional "command and control" regulation. This alternative would require utilities to install what is known as "maximum achievable control technologies" (MACT), as required under the Clean Air Act, to reduce mercury emissions by 29% over the next three years. It is not clear which of these two approaches will become the final regulation.

CONCLUSIONS

Mercury is highly toxic to humans. Significant reduction of mercury in the environment can be attained through better pollution prevention programs and control of both mercury vapors and liquid spills. Mercury waste is difficult to remediate and several treatment methods have been developed.

Acknowledgments
The authors wish to thank Ross B. Jacobs, a student interested in environmental research and issues related to mercury, for his contributions to this entry.

BIBLIOGRAPHY

1. Stewertka, A. (2002). *A Guide to the Elements*, 2nd Edn. Oxford University Press, New York, p. 247.
2. United States Environmental Protection Agency (U.S. EPA) (1997). *Mercury Study Report to Congress*, EPA-452/R-97-003. U.S. EPA, Washington, DC.
3. United States Environmental Protection Agency (U.S. EPA) (2004). *Consumer Fact Sheet on Mercury*, U.S. EPA, Washington, DC. Available: http://www.epa.gov/safewater/contaminants/dw_contamfs/mercury.html. Updated April 27, 2004.
4. United States Environmental Protection Agency (U.S. EPA) (2003). *Mercury Study Report to Congress. Whitepaper*, U.S. EPA, Washington, DC. Available: http://www.epa.gov/oar/merwhite.html. Updated August 19, 2003.
5. Kansas Department of Health and Environment (KDHE) (1994). Mercury contamination remediation at gas pipeline sites, October, p. 8.
6. Suthersan, S. (2002). *Natural and Enhanced Remediation Systems*. Lewis Publishers/CRC Press, Boca Raton, FL.
7. United States Environmental Protection Agency (U.S. EPA) (1998). *Technologies for Immobilizing High Mercury Subcategory Wastes*, EPA Contract No. 68-W4-0005, WA No. R11032, July 17. Submitted to TechLaw, Inc., Chantilly, VA, p. 15.
8. Qian, G., Sun, D.D., and Tay, J.H. (2003). Immobilization of mercury and zinc in an alkali-activated slag matrix. *J. Hazard. Mater.* **101**(1): 65–77.
9. Minnesota Office of Environmental Assistance (MOEA) (2001). Reducing mercury in the environment; May. Available: http://www.moea.state.mn.us/berc/mercury.cfm.

GROUNDWATER AND LEAD: CHEMICAL BEHAVIOR AND TREATMENT

James A. Jacobs
Lisa Kirkland
Environmental Bio-Systems, Inc.
Mill Valley, California

Lead (Pb) occurs naturally in the earth's crust and is reported as being the most abundant among the heavy metals, with the average Pb content of the crust ranging from 13 to 16 parts per million (1). Lead has a crustal abundance of 1.4×10^1 mg/kg, with an estimated oceanic abundance of 3×10^{-5} mg/L (2). Lead makes up 0.013% of the earth's crust. There are more than 200 minerals of Pb. Natural mobilization of Pb into the environment occurs mainly from the erosion of Pb-containing rocks and via gaseous emissions during volcanic activity. Lead is mined and refined easily. On a worldwide basis, most lead is concentrated by roasting galena, also called lead sulfide (PbS), in hot air. Nearly one-third of the consumption in the United States is obtained by recycling efforts (2).

INTRODUCTION

The impact of Pb on the environment has been long and persistent. Approximately 5000 years ago Pb mining waste began the first time a curious human picked at an outcropping of lead, pulled away the dirt and rock, and discarded the unwanted material. The Romans were the first of the ancient people to use Pb on a large scale and they used Pb more than any other people until recent time. In fact, there were laws prohibiting the production of excess amounts, according to *Pliny the Elder's Natural History* (3). The name lead comes from the Anglo-Saxon word for lead. The symbol, Pb comes from the latin word for lead, *plumbus*. Plumbers were lead workers using lead pipes and lead solder.

Lead is a cumulative poison and the ancient Romans were exposed to significant levels of lead in their water system. The blood levels of Pb in the post-Roman Italian population are as high as the Roman population, which is indicative that Pb was obtained in food and drink; this may be the result of continued use of Pb water pipes and cisterns or a result of the use of pewter, poor Pb glazes, and Pb salts (e.g., Pb acetate, "sugar of lead", which was used as a sweetener). Poor wine was improved in the Middle Ages by the addition of Pb even though it was known to be a dangerous practice and outlawed and sometimes punishable by death. There are many records of Pb epidemics, which persist well into the eighteenth century (4).

New centuries brought new uses for Pb, some of which helped change the course of history. In the mid-1400s, Johannes Gutenberg created the printing press by casting letters of molten Pb and tin. Later, Pb made possible electronic circuitry, computer screens, and televisions.

As the uses of lead increased, driven by the growing number of social benefits it provided, there was more lead mining and mining waste. Until a few decades ago, there was little concern about the environmental impact of mining waste. Limited engineering solutions and lack of public concern about the issue meant that, over the years, mining companies have developed their own methods of handling and disposing of these materials. In fact, before engineers spent time deciding how to keep piles of tailings from washing or blowing away and before society understood the need for such controls, tailings were simply left in piles. As the early Pb mines were exhausted of ore and closed, the tailing piles were left behind. In some cases the piles have slumped into streams; in other cases, wind has blown away portions of the piles. In recent years, society's concern about the impact of tailings on the environment and human health has grown. That, along with engineering advancements that help stabilize tailing piles, has led to changes.

PROPERTIES

Lead is a bluish-white metal of bright luster, soft, highly malleable, ductile, and a poor conductor of electricity. Lead is a soft, malleable, and corrosion-resistant metal, belonging to group IV of the periodic table. Lead has an atomic weight of 207.19, a melting point of 327.5 °C, a boiling point of 1744 °C, and specific gravity of 11.342 g/cm^3. Lead is a solid at room temperature. The oxidation states of Pb are II and IV and there are four stable isotopes (204, 206, 207, 208) and two radioactive isotopes (210, 212). A variety of lead oxide compounds exist in nature.

USE

In ancient times the need for silver propelled Pb production. This was coupled with its properties—that is, its low melting point, ductility, malleability, and durability contributed to its use as a construction material. Currently, the construction and automobile industries are the primary users of Pb. It is estimated that 60% of Pb is consumed in the auto industry for solder in the building of cars, buses, trucks, vehicular batteries, and as gasoline additive. The use of Pb as gasoline additive has been declining since the mid-1970s as a result of legislation in many countries, responding to environmental concern about the release of Pb from auto exhausts.

Other important uses of Pb are in the manufacture of cable sheathings, pigments, pipe, foil and tube, and various alloys. Secondary use of Pb includes ammunition and type metal, varnishes, the production of inorganic compounds, protective coating, glassware, ceramics, TV tubes, shield against radioactivity, galvanizing, Pb plating, weights, and ballasts. Lead carbonate ($PbCO_3$), also called cerussite, is a white pigment used in early paints, especially in interior white paints. These paints have caused significant lead-based-paint (LBP) poisoning, especially of infants and children who might chew on the sweet-tasting lead paint chips. Currently, titanium oxide (TiO_2) has replaced the toxic lead in paints. Other lead compounds used in paints include lead sulfate or anglesite ($PbSO_4$) and lead chromate or crocoite ($PbCrO_4$). Other lead compounds include lead nitrate [$Pb(NO_3)_2$] used in fireworks and lead silicate ($PbSiO_3$) used in glass making and in the production of paints and rubber (2).

Lead arsenate [$Pb(AsO_4)_2$] was used as an insecticide and has been banned as an ingredient in pesticides (5). The high density of lead makes it useful to shield against X-ray and gamma-ray radiation in X-ray machines and other scientific instruments. It is used to absorb vibrations, as well as in the manufacture of ammunition. Most lead currently is used in lead-acid storage batteries, the type found in vehicles. Lead has been used in crystal and glass as well as solder (with tin).

The elevated levels are due:

- to the continued use of lead (Pb) pipes in drinking water transport
- chips and dust from leaded paint
- emissions from various industrial combustion processes
- deterioration of leaded paint
- historical fall-out from leaded gasoline (6).

HEALTH RISKS

The health risks associated with Pb include poisoning, with the most common symptoms being insomnia, delirium, hallucinations, tremor, colic, nausea, and vomiting. (Persons at risk would include those handling tetraethyl lead or leaded gasoline.) Symptoms of lead poisoning also include neuropathy, encephalopathy (both adult and children), renal failure, intoxication, and death. Lead anemia is seldom severe unless concomitant with iron deficiency. More importantly, low level Pb exposure can have profound, adverse physiological and cognitive effects that can have long-lasting impacts. In children this could lead to permanent damage to the central nervous system, lower performance I.Q., and impairment of fine motor function. Children are more susceptible due to their greater rates of absorption (an area that needs further examination). Lead poisoning in adults is manifested by a decrease in reaction time, lower memory capacity, weakness in fingers, wrists, or ankles, anemia, insomnia, weight loss, anorexia, malnutrition, nausea, abdominal pain, constipation, and vomiting. Lead in men may increase blood pressure and high levels may damage the male reproductive system (6,7).

Pb is incorporated in the hair, nails, and sweat, but losses from the body by these routes are trivial. Blood Pb levels tend to be higher in city dwellers than in rural dwellers, while within cities those living in areas with the greatest traffic density tend to have the highest blood levels (8). The relative contribution which the various environmental sources of Pb make to the blood Pb concentration is not known with certainty and more study needs to be done in this regard. Currently, the EPA guideline for Pb blood level is 10 Mg/dL.

FATE AND TRANSPORT

Lead released to land, surface water, and groundwater is usually elemental lead, lead oxides, lead hydroxides, and lead–metal oxyanion complexes (9).

Pb naturally is present in small quantities in practically all environmental matrices, which includes air, water, soil, and plants. The distribution of Pb is greatly exacerbated as a result of emission from the combustion of Pb-containing fuel. It must be pointed out that Pb toxicity is inversely proportional to pH in soil. Pb in water originates from a variety of sources: atmosphere, lithosphere, and, for drinking water, the possibility from plumbing fixtures as well as from mines that may percolate into underlying aquifers. In the 1970s Pb from anti-knock additives resulted in approximately 180,000 tons/yr in the United States alone. Lead is the most common metal contaminant found at U.S. Superfund sites (460 sites) (10).

Pb is rarely found in source water, but Pb mining and smelting operations may be sources of significant contamination. Eighty-eight percent of the Pb mined in the United States comes from seven mines in the New Lead Belt in southeastern Missouri. From 1987 to 1993, according to the Toxics Release Inventory, Pb compound releases to land and water totaled nearly 144 million pounds. These releases were primarily from Pb and copper smelting industries. The largest releases occurred in Missouri, Arizona, and Montana. The largest direct releases to water occurred in Ohio. Pb in the form of very small particles may be airborne and disseminated over long distances. Metal-rich mine tailings, metal smelting, electroplating, gas exhausts, energy and fuel production, downwash from power lines, intensive agriculture, and sludge dumping are the human activities that introduce the largest quantities of Pb into soils. When released to land, Pb binds to soils and does not migrate to groundwater. A total of 71,529 tons of lead were released to the land in the United States over a seven-year period. In water, it binds to sediments. A total of 485 tons were released to water in the United States over a seven-year period. The worst industries for releasing lead to water are the steelworks, lead smelting, and refining and iron foundries. The worst industries for releasing lead to land are the lead smelting and refining, copper smelting, and steelworks. Lead does not accumulate in fish but does in some shellfish, such as mussels (11).

Lead (Pb^{2+}) and lead–hydroxy complexes are the most stable forms of lead (9). Lead forms low solubility compounds by complexation with inorganic ions (Cl^-, ${CO_3}^{2-}$, ${SO_4}^{2-}$, ${PO_4}^{3-}$) and organic ligands (humic and fulvic acids, EDTA, amino acids) (12). Lead carbonate solids form above pH 6. Lead sulfide forms with high sulfide concentrations under reducing conditions. Most lead released into the environment is retained in the shallow soil. The main processes affecting the fate of lead in soil generally limits the amount of lead entering the groundwater. The processes include adsorption, ion exchange, precipitation, and complexation with sorbed organic matter in the soil (13). The amount of lead dissolved in groundwater and surface water is heavily pH dependent and to a lesser degree related to the concentration of dissolved salts and types of mineral surfaces present. In surface and groundwater, lead usually occurs as a precipitate (rather than in solution), as $PbCO_3$, PB_2O, $PB(OH)_2$, and $PbSO_4$, as sorbed ions or surface coatings on minerals, or as suspended organic matter (13).

The major pathways of exposure to lead are:

- the inhalation of lead-containing car exhausts or industrial emission
- the ingestion of lead-based paint
- the ingestion of contaminated soil or dust from hand-to-mouth activities of those living in lead-polluted environments
- the inhalation of leaded dust carried on clothing or by the wind

The specific pathways of exposure to lead for children are:

- swallowing lead paint chips
- eating contaminated food
- putting hands, toys, or objects contaminated with lead dust or soil in their mouths

GUIDELINES AND STANDARDS

In 1974, the U.S. Congress passed the Safe Drinking Water Act. This law requires the EPA to determine safe levels of chemicals in drinking water, which do or may cause health problems. These nonenforceable levels, based solely on possible health risks and exposure, are called Maximum Contaminant Level Goals (MCLG).

The MCLG for lead has been set at zero because the EPA believes this level of protection would not cause any of the potential health problems. Since lead contamination generally occurs from corrosion of household lead pipes, it cannot be directly detected or removed by the water system. Instead, the EPA is requiring water systems to control the corrosiveness of their water if the level of lead at home taps exceeds an Action Level.

The Action Level for lead has been set at 15 parts per billion (ppb) because the EPA believes, given present technology and resources, this is the lowest level to which water systems can reasonably be required to control this contaminant should it occur in drinking water at their customers' home taps.

These drinking water standards, and the regulations for ensuring these standards are met, are called National Primary Drinking Water Regulations. All U.S. public water supplies must abide by these regulations.

REMEDIATION

Various remedial methods exist for the treatment of Pb contamination such as excavation, phytoremediation, soil containment, soil washing and flushing, deep tilling, and geochemical fixation. Other remedial approaches include isolation by capping and subsurface barriers including slurry walls. Immobilization includes solidification/stabilization using cement-based binders and stabilizers, organic binders, and vitrification. Permeable treatment walls, *in situ* soil flushing, and electrokinetic extraction are also gaining popularity in the treatment of heavy metals (13). Phytoremediation, soil washing, and geochemical fixation are described below.

Phytoremediation

Phytoremediation is a relatively new technology that has been gaining commercial acceptance as research on the field scale has confirmed greenhouse and pilot scale success. While it has been used mainly in open, industrial settings, phytoremediation has been used in several urban, residential applications for Pb reduction. For example, in Massachusetts phytoremediation significantly reduced high soil lead levels in both an open lot that was revitalized into a community garden and in an urban residential backyard. In the case of the community garden, several plantings of Indian mustard lowered the Pb levels that initially exceeded 1000 ppm to concentrations less than half that amount. While phytoremediation in a full-scale residential application presents a few more challenges than that in an industrial site or an isolated lot, it can be successfully implemented over time through the use of a comprehensive monitoring and planting regime. This method is inexpensive and is particularly suited for developing countries.

Soil Washing

Soil washing is a remediation system that leaches and recovers Pb and other metals from contaminated soil, dust, sludge, or sediment. Soil washing also uses physical and mechanical sorting methods to separate the various fractions of the sediment by size. The Pb compounds are often tightly bound by fine soil constituents such as clay, manganese and iron oxides, and humus. A pretreatment, physical separation stage may involve dry screening to remove gross oversized material. The soil can be separated into oversized (gravel), sand, and fine (silt, clay, and humus) fractions. Soil, including the oversized fraction, is first washed. Most Pb contamination is typically associated with fines fraction, and this fraction is subjected to countercurrent leaching to dissolve the adsorbed Pb and other heavy metal species. The sand fraction may also contain significant Pb, especially if the contamination is due to particulate Pb, such as that found in battery recycling, ammunition burning, and scrap yard activities. In this case, the sand fraction is pretreated to remove dense metallic or magnetic materials before subjecting the sand fraction to countercurrent leaching. Sand and fines can be treated in separate parallel streams.

After dissolution of the Pb and other heavy metal contaminants, the metal ions are recovered from the aqueous leachate by a metal recovery process such as reduction, liquid ion exchange, resin ion exchange, or precipitation. The metal recovery technique depends on the metals to be recovered and the leachate employed. In most cases, a reduction process is used so that the metals are recovered in a compact form suitable for recycling. After the metals are recovered, the leachate can be reused within the soil washing system for continued leaching.

Geochemical Fixation

An *in situ* technology for Pb treatment is geochemical fixation using a chemical reductant such as calcium polysulfide, ferrous sulfate, or sodium metabisulfite. As an example, calcium polysulfide generally produces hydrogen sulfide, thiosulfate, and colloidal sulfur when placed in groundwater. The soluble Pb^{2+} in the groundwater is precipitated as an insoluble lead sulfide by thiosulfate breakdown products and the hydrogen sulfide. Lead sulfide is less soluble than lead hydroxide. The reductants tend to be injected or poured into wells, trenches, or delivery points. Some chemical reductants can remain active in the groundwater for months to years.

BIBLIOGRAPHY

1. Swaine, D.J. (1978). Lead in the environment. *J. Proc. R. Soc. New South Wales Australia*, **111**: 41–47.
2. Jefferson Lab (JLAB) (2004). It's elemental: lead. Available: http://education.jlab.org/itsselemental/ele082.html.
3. Healy, J.F. (1991). *Pliny the Elder's Natural History: A Selection*, translated by J.F. Healy, Penguin Books, London.
4. Waldron, H.A. (1980). Lead. *London School Hyg. Trop. Med.*, 155–191.
5. Adriano, D.C. (1986). *Trace Elements in the Terrestrial Environment*, Springer, New York, pp. 219–256.
6. ATSDR (1997). Agency for Toxic Substances and Disease Registry. *Toxological Profile for Vinyl Chloride,* U.S. Department of Health and Human Services, Public Health Service. Available: http://www.atsdr.cdc.gov/toxprofiles/tp20.pdf.

7. ATSDR (2004). Agency for Toxic Substances and Disease Registry. Medical management guidelines (MMGs) for vinyl chloride. Available: http://www.atsdr.cdc.gov/mmg.html.
8. World Health Organization (1977). *Inorganic Lead* (EHC 165, 1995) Health Criteria 3: *Lead*, WHO, Geneva, Switzerland.
9. Smith, L.A., Means, J.L., Chen, A., Alleman, B., Chapman, C.C., Tixier, J.S. Jr., Brauning, S.E., Gavaskar, A.R., and Royer, M.D. (1995). *Remedial Options for Metals-Contaminated Sites*, Lewis Publishers/CRC Press, Boca Raton, FL.
10. United States Environmental Protection Agency (U.S. EPA) (1996). Recent developments for *in situ* treatment of metals-contaminated soils, Office of Solid Waste and Emergency Response, Draft.
11. United States Environmental Protection Agency (U.S. EPA) (1994). Available: http:www.epa.gov/Superfund/accomp/sarc/1994/chap_.pdf.
12. Bodek, I., Lyman, W.J., Reehl, W.F., and Rosenblatt, D.H. (1988). *Environmental Inorganic Chemistry; Properties, Processes and Estimation Methods*. Pergamon Press, Elmsford, NY.
13. Evanko, C.R. and Dzombak, D.A. (2000). Remediation of metals-contaminated soil and groundwater. In: *McGraw-Hill Standard Handbook of Environmental Science, Health and Technology*. J. Lehr (Ed.). McGraw-Hill, New York, pp. 14.100–14.134.

LAMINAR FLOW

NAVRAJ S. HANSPAL
Loughborough University
Loughborough, United Kingdom

RAJ SHARMA
University of KwaZulu-Natal
Durban, South Africa

INTRODUCTION

The prime objective of this text is to describe the elementary concepts involved in laminar fluid flows. In general, any flow is characterized by the continuous deformation of fluid particles and a corresponding internal state of stress. It has been experimentally established that, under laminar situations, the rate of deformation is directly proportional to local viscous stress. Laminar flows thus fall under the category of viscous flows. The principle laws that govern the motion of fluids are also discussed. We shall take geometrically simple configurations and provide a more specific introduction of laminar fluid dynamics. The outlines given in this article provide basic concepts of laminar flows. For more detailed analysis, the reader should consult specialized texts within this topic. Some of these texts are listed in the references given at the end of this article.

VISCOSITY AND OSTWALD'S EXPERIMENT

Viscosity is a measure of how easily the fluid deforms or flows under an applied stress. Alternatively, viscosity is a measure of the fluid's internal resistance to deformation or internal friction between the fluid layers. Experimentally, under conditions of laminar flow, the force required to move a plate at constant speed against the resistance of a fluid is proportional to the area of the plate and to the velocity gradient perpendicular to the plate. The constant of proportionality is called the viscosity. Consider a simple fluid flow between the two plates shown in Fig. 1.

If the force per unit area of the top plate is measured as F, we can write that as

$$\frac{F}{A} = \mu \frac{v}{d} \tag{1}$$

with d as the separation between the plates, which essentially defines the absolute or dynamic viscosity, μ, which has the units of kg/ms. The ratio of dynamic viscosity to density of the fluid is known as kinematic viscosity and has the units of m^2/s. Experimentally, the viscosity can be measured using specially designed viscometers, whose operation is based on simple flow between one plate fixed and another moving with constant velocity v. If the gap d between the plates is small compared with the length of the plates, the flow with reasonable accuracy can be assumed to be one-dimensional. For such flows, the velocity is constant in the direction of flow but varies in the normal direction. For example, in Fig. 1, the velocity in x-direction is a function of y coordinate, i.e., $v = v(y)$. Equation (1) also represents that the viscous stresses are directly proportional to the velocity gradient, i.e.,

$$\textit{Viscous stresses} = \mu(\textit{Rate of deformation})$$

These stresses developing in the viscous transfer can be divided into normal and shear stresses. For example, τ_{xx} is the normal stress, whereas τ_{yx} is the shear stress resulting from the viscous forces. At each point in a continuous medium, whether it is solid or fluid, we need six components, each of them representing a component of force per unit area, to define the local stress completely. The local stresses acting in the x-y plane have been shown in Fig. 2.

Equation (1) can be cast in a more generalized form as given by Eq. (2). The shear stress exerted in the x-direction on a fluid surface of constant y by the fluid in the region of lesser y is designated as τ_{yx}, and the x-component of the fluid velocity vector is designated as v_x

$$\tau_{\mathrm{yx}} = -\mu \frac{du}{dy} \tag{2}$$

This law is commonly known as Newton's law of viscosity. It states that the shear force per unit area is

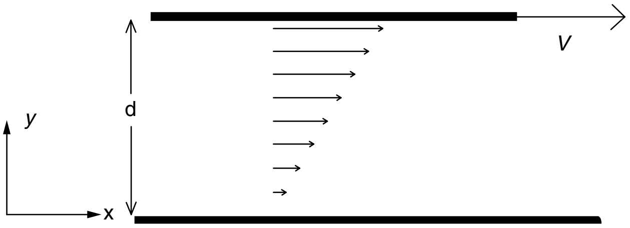

Figure 1. Viscous drag between two plates; the bottom plate is held fixed.

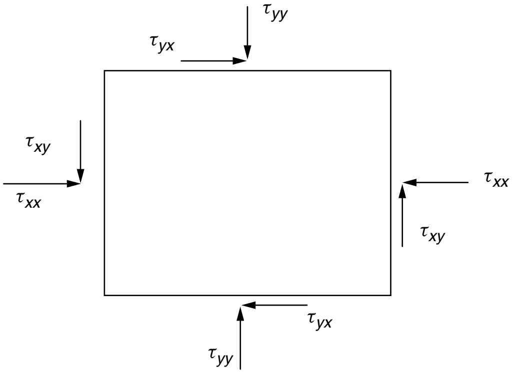

Figure 2. Local stresses acting in a *x-y* plane on a cubic element.

proportional to the negative of the local velocity gradient. In an arbitrary three-dimensional laminar flow, this relation can be generalized as

$$\begin{bmatrix} \tau_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{xy} & \tau_{yy} & \tau_{yz} \\ \tau_{xz} & \tau_{yz} & \tau_{zz} \end{bmatrix} = \mu \begin{bmatrix} \left(2\dfrac{\partial u}{\partial x}\right) & \left(\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x}\right) & \left(\dfrac{\partial u}{\partial z} + \dfrac{\partial w}{\partial x}\right) \\ \left(\dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x}\right) & \left(2\dfrac{\partial v}{\partial y}\right) & \left(\dfrac{\partial w}{\partial y} + \dfrac{\partial v}{\partial z}\right) \\ \left(\dfrac{\partial u}{\partial z} + \dfrac{\partial w}{\partial x}\right) & \left(\dfrac{\partial w}{\partial y} + \dfrac{\partial v}{\partial z}\right) & \left(2\dfrac{\partial w}{\partial z}\right) \end{bmatrix} \quad (3)$$

For detailed derivations and discussion of viscous stresses, one should refer to References 1 and 2.

Flows in which fluid viscosity is important can be of two types, namely laminar and turbulent. The basic difference between these flows was demonstrated by Osborne Reynolds in 1883. A thin stream of dye was introduced into the flow of water through the tube. At low flow rates, the dye stream was observed to follow a well-defined straight path, indicating that the fluid moves in parallel layers (*laminas*) with no macroscopic mixing across the layers, which is known as *laminar flow*. As the flow rate was increased beyond certain critical values, the dye streak broke up into an irregular motion and spread throughout the cross section of the tube, indicating the macroscopic mixing motions perpendicular to the direction of flow. Such a chaotic fluid motion is known as *turbulent flow*.

REYNOLDS NUMBER

The Reynolds Number (Re) is a quantity that engineers use to estimate if a fluid flow is laminar or turbulent. Consider the flow of an incompressible fluid through a circular pipe as shown in Fig. 3.

The variables associated with this pipe-flow configuration are the dimensions of the pipe, the rate of flow, and the physical properties of the fluid.

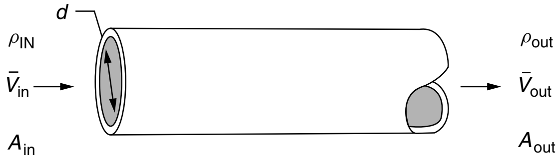

Figure 3. Flow through a pipe.

d = Internal diameter of the pipe

ρ = Density of the fluid and in an incompressible flow $\rho_{\text{in}} = \rho_{\text{out}}$ because it is constant

$\bar{v}$ = Average velocity

Q = Volumetric flow rate

A = Cross-sectional area and is constant, i.e., $(A_{\text{in}} = A_{\text{out}})$

We can then represent the Reynolds Number as a dimensionless entity as

$$\text{Re} = \frac{\rho \bar{v} d}{\mu} = \frac{\rho\left(\dfrac{Q}{A}\right)d}{\mu} = \textbf{Unitless Number} \quad (3)$$

The Reynolds Number also indicates the relative significance of the inertial forces compared with the viscous forces. Laminar flow occurs at low Reynolds Numbers, where viscous forces are dominant, and is characterized by smooth, constant fluid motion, whereas turbulent flow occurs at high Reynolds Numbers and is dominated by inertial forces, producing random eddies, vortices, and other flow fluctuations. Different ranges of Reynolds Numbers determine the nature of different flow regimes, as shown in Fig. 4.

In laminar flow, the particles in the fluid follow streamlines (see Fig. 4), and the motion of particles in the fluid is predictable. If the flow rate is very large, or if objects obstruct the flow, the fluid starts to swirl in an erratic motion. No longer can one predict the exact path a particle on the fluid will follow. This region of constantly changing flow lines (see Fig. 4) is said to consist of turbulent flow.

RANGE 1: laminar flow (see Fig. 4)

Generally, a fluid flow is laminar from $\text{Re} = 0$ to some critical value $\text{Re} = 2100$, at which transition flow begins.

RANGE 2: transition flow (see Fig. 4)

Flows in this range may fluctuate between laminar and turbulent flow. The fluid flow is on the verge of becoming turbulent. In this zone, $2100 < \text{Re} < 4000$.

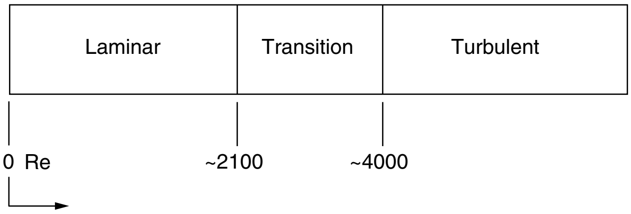

Figure 4. Different flow regimes and Reynolds Number in pipe flow.

RANGE 3: turbulent flow (see Fig. 4)

The fluid flow has become unstable. In turbulent flow, increased mixing exists that results in viscous losses, which are generally much higher than in those in laminar flow. In this zone, the values of Reynolds Number is greater than 4000.

It should be noted that the critical Reynolds Number depends on the flow type and the definition of the Reynolds Number, which varies for different geometries.

CONSERVATION LAWS FOR FLUID FLOW SYSTEMS

The behavior or dynamics of any fluid flow system can be described by laws of conservation of mass, momentum, and energy. In order to yield the detailed descriptions of variables of interest within the flow system, the control volume must be of infinitesimal dimensions that can shrink to zero, yielding a point volume. This approach reduces the quantities to point variables. The application of the conservation principles to this infinitesimal system produces microscopic or differential conservation equations. The governing momentum, energy, and mass balance differential equations for a given flow system can be derived using the shell balance concepts or directly from equations of change, which have been discussed in detail in References 2 and 3. Shell balance concepts are generally valid for simple flow systems, whereas equations of change can describe the flow behavior in complex flow real-life situations.

The governing differential equations for any flow system can be expressed in Lagrangian and Eulerian coordinate systems. In the Lagrangian approach, the fluid is assumed to be composed of particles, and the motion of individual particles is tracked. However, in the Eulerian approach, the attention is instead focused on a fixed coordinate system that defines space. We consider a prime variable of interest in fluid dynamics, that is, velocity. In the Lagrangian approach, this parameter is a function of time, whereas in Eulerian systems, it is a function of both time and spatial coordinates. Therefore, the acceleration of a particle that at time t occupies a fixed location can be calculated either by using the material or substantial derivatives as

$$a = \frac{DV}{Dt} \tag{4}$$

or by chain rule of differentiation as

$$a = \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial v}{\partial z}\frac{\partial z}{\partial t} \tag{5}$$

Equations (4) and (5) are the same; Eq. (4) represents the Lagrangian approach, whereas Eq. (5) is representative of an Eulerian system.

$$\underset{\substack{\uparrow\\ \textbf{Lagrangian Approach}}}{\frac{DV}{Dt}} = \underset{\substack{\uparrow\\ \textbf{Eulerian Approach}}}{\underline{\frac{\partial v}{\partial t} + \frac{\partial v}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial v}{\partial z}\frac{\partial z}{\partial t}}} \tag{6}$$

For advanced discussions about Lagrangian and Eulerian coordinate systems, the readers are advised to refer to References 2, 4, and 5.

EQUATIONS OF CHANGE IN LAMINAR FLOW SYSTEMS

The governing differential equations corresponding to mass and momentum conservation are solved to obtain field variables of interest in fluid dynamics (i.e., velocities and pressures), which are also known as the continuity equation and equations of change, which describe the laminar flow of a pure isothermal fluid.

EQUATION OF CONTINUITY

Equation of continuity is derived by applying the law of conservation of mass to a small 3-D volume element (see Fig. 5) within a flowing fluid. In vector notation, this equation can be represented as

$$\frac{\partial \rho}{\partial \mathbf{t}} = -(\nabla.\rho v) \tag{7}$$

Here, $(\nabla.\rho v)$ is called the "divergence" of $\rho\mathbf{v}$. This vector $\rho\mathbf{v}$ is the mass flux and is the net rate of mass efflux per unit volume. Equation (7) states that the rate of increase of the density within a small volume element fixed in space is equal to the net rate of mass influx to the element divided by its volume. A very special form of the equation of continuity is Equation (8), which is applicable for fluids of constant density, that is, incompressible fluids.

$$(\nabla.\mathbf{v}) = 0 \tag{8}$$

EQUATION OF MOTION

The equation of motion is developed by implementing the shell momentum balance concept on a 3-D volume element, as shown in Fig. 5.

It should be noted that momentum is a vector quantity, and hence the equation of motion will have components in each of the three coordinate directions x, y, and z, respectively. For detailed derivation of equation of motion, readers should refer to References 2 and 5. Adding the three components of equation of motion, the equation of

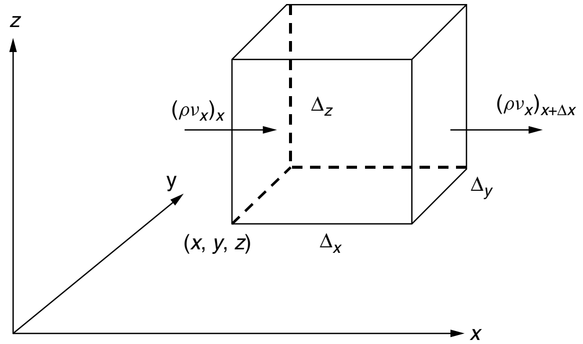

Figure 5. Region of volume in $\Delta x\Delta y\Delta z$ through which the fluid is flowing.

motion can be represented in vector notation as

$$\underset{\substack{\text{Rate of increase of}\\ \text{momentum per unit}\\ \text{volume}}}{\frac{\partial}{\partial t}\rho\mathbf{v}} = \underset{\substack{\text{Rate of momentum}\\ \text{gain by convection}\\ \text{per unit volume}}}{[\nabla.\rho\mathbf{vv}]} - \underset{\substack{\text{Pressure force on}\\ \text{element per unit}\\ \text{volume}}}{[\nabla p]} \underset{\substack{\text{Rate of momentum}\\ \text{gain by viscous}\\ \text{transfer per unit}\\ \text{volume}}}{-[\nabla\cdot\boldsymbol{\tau}]} + \underset{\substack{\text{Gravitational force}\\ \text{on element per unit}\\ \text{volume}}}{\rho\mathbf{g}} \qquad (9)$$

$\boldsymbol{\tau}$ is known as “stress tensor” and has nine components, $[\nabla.\rho vv]$ represents the rate of loss of momentum per unit volume by fluid flow, and $(\nabla.\rho v)$ represents the rate of loss of mass per unit volume. In order to use Equation (9) to determine velocity distribution, appropriate expressions for various stresses in terms of velocity gradients and fluid properties are substituted, which yields the equations of motion for a Newtonian fluid with varying density and viscosity.

For constant ρ and μ, the equation of motion can be written as

$$\rho\left(\frac{\partial v}{\partial t} + v.\nabla v\right) = -\nabla p + \mu\nabla^2 v + \rho\mathbf{g} \qquad (10)$$

Equation (10) is the well-known Navier–Stokes equation, first developed by Navier in France in 1822.

The equations of motion together with equation of continuity, equation of state $p = p(\rho)$, viscosity models for different fluids [see References 4 and 6], and the initial and the boundary conditions [see References 1, 3, 4, and 7], complete the problem specification for any laminar flow problem in hand. These equations, when solved mathematically, determine completely the density, pressure, and velocity components in an isothermal laminar flow situation.

Equation (10) can be written in nondimensional form as

$$\left(\frac{\partial v'}{\partial t} + v'.\nabla v'\right) = -\nabla p' + \frac{1}{\mathrm{Re}}\nabla^2 v' + \mathbf{g} \qquad (11)$$

where the primed quantities are dimensionless. The ratio of second term to the fourth term in Equation (11) is of order *Re*, which indicates the physical representation of Reynolds Number as

$$\mathrm{Re} = \frac{\text{inertia forces}}{\text{viscous forces}} \qquad (12)$$

Detailed derivation of nondimensional form of Navier–Stokes and Stokes equations can be found in References 2 and 8.

Concept of Dynamic Pressure

If the body of fluid is at rest, the pressure is hydrostatic and

$$\rho\mathbf{g} - \nabla p = 0 \qquad (13)$$

Subtracting Eq. (13) from Eq. (10), we get

$$\rho\left(\frac{\partial v}{\partial t} + v\cdot\nabla v\right) = -\nabla p_{\mathrm{d}} + \mu\nabla^2 v \qquad (14)$$

where $p_{\mathrm{d}} = p_{\mathrm{s}} - p$ is the pressure change because of dynamic effects. However, no accepted terminology exists for p_{d}, but we shall call it dynamic pressure. Other common terms are modified pressure (9) and excess pressure (10). For a fluid of uniform density, introduction of p_{d} eliminates gravity from the differential equations in Eq. (10). However, the process may not eliminate gravity from the problem. Gravity reappears in the problem if the boundary conditions are given in terms of total pressure p. Examples are surface gravity waves where the total pressure is fixed at the free surface. Without a free surface, however, gravity has no dynamic role, and its only effect is to add a hydrostatic contribution to the pressure field.

Equations (7)–(14) have been presented in rectangular coordinate systems. However, in real complex flow situations, depending on the geometrical configurations of the flow system, other coordinate systems such as spherical and cylindrical can be used. Hence, in these cases, Equations (7)–(14) can be cast in alternative coordinate systems using coordinate transforms. Detailed literature and methodology for performing coordinate transformations can be found in References 2 and 5.

STEADY-STATE FLOW BETWEEN PARALLEL PLATES

A laminar flow within two parallel plates is considered in Fig. 6.

The term fully developed in the figure signifies that we are considering regions beyond the developing stage near the entrance, where the velocity profile changes in the direction of flow because of the development of boundary layers from the two walls. The discussion of boundary layer theory is beyond the scope of the present text, and we shall only consider the flows under fully developed situations, i.e., far away from the entrance. For advanced discussions on boundary layer theory, readers may consult References 2 and 11.

Consider the fully developed stage of a two-dimensional flow between two infinite parallel plates, as shown in Fig. 7.

The flow is driven by an externally imposed pressure gradient, and the motion of the upper plate is maintained

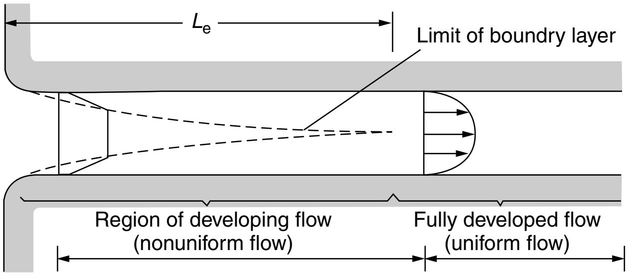

Figure 6. Developing and fully developed flows in a channel. Flow is fully developed after the boundary layers merge.

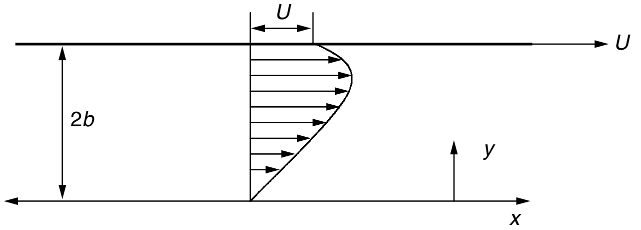

Figure 7. Steady-state flow between two parallel plates.

at velocity U. The x coordinate axis is parallel to the direction of flow. Two-dimensionality of the flow requires that $\frac{\partial}{\partial z} = 0$. Flow characteristics are also invariant in the x-direction, so that continuity requires $\frac{\partial v}{\partial z} = 0$. As $v = 0$ at $y = 0$, it follows that $v = 0$ everywhere, which implies that the flow is parallel to the walls. The momentum balance equations in x and y directions can be written as

$$0 = -\frac{1}{\rho}\frac{\partial p}{\partial x} + \mu\frac{d^2u}{dy^2} \tag{15}$$

$$0 = -\frac{1}{\rho}\frac{\partial p}{\partial y} \tag{16}$$

The y-momentum in Eq. (16) shows that p is not a function of y. In the x-momentum in Eq. (15), the first term can only be a function of x, whereas the second term can only be a function of y. The only way this can be established is that both the terms are constants. The pressure gradient is hence a constant, implying that the pressure varies linearly across the channel. Integrating Eq. (15) twice, we get

$$0 = \frac{y^2}{2}\frac{dp}{dx} + \mu u + Ay + B \tag{17}$$

where p is a function of x only. The constants of integration are determined by using the conditions imposed on lower and upper plates. For the lower boundary, $u = 0$ at $y = 0$ requires that $B = 0$. The upper boundary condition $u = u$ at $y = 2b$, requires that $A = b\frac{dp}{dx} - \mu\frac{u}{2b}$.

The velocity profile from Eq. (17) then becomes

$$u = \frac{yU}{2b} - \frac{y}{\mu}\frac{dp}{dx}\left(b - \frac{y}{2}\right) \tag{18}$$

The volumetric flow rate per unit width of the channel can be calculated using

$$Q = \int_{O}^{2b} u dy = Ub\left[1 - \frac{2b^2}{3\,\mu U}\frac{dp}{dx}\right] \tag{19}$$

The average velocity can further be written as

$$\overline{V} = \frac{Q}{2b} = \frac{u}{2}\left[1 - \frac{2b^2}{3\,\mu U}\frac{dp}{dx}\right] \tag{20}$$

Two special cases of interest have been discussed in the subsequent section,

1. *Plane Couette Flow*. The flow driven by the motion of the upper plate, without the imposition of any external pressure gradient, is called a plane Couette flow.

Equation (18) then reduces to a linear profile represented by Eq. (21) and can be seen in Fig. 8.

$$u = \frac{yU}{2b} \tag{21}$$

The magnitude of shear stress is given by

$$\tau_{\mathrm{yx}} = \mu\frac{dU}{dy} = \frac{\mu u}{2\mathrm{b}} \tag{22}$$

which is uniform across the channel.

2. *Plane Poiseuille Flow*. The flow driven by an externally imposed pressure gradient through two stationary flat walls is called a plane Poiseuille flow. In this case, Eq. (18) reduces to Eq. (23), which represents a parabolic profile, as seen in Fig. 9.

$$u = -\frac{y}{\mu}\frac{dp}{dx}\left(b - \frac{y}{2}\right) \tag{23}$$

The magnitude of shear stress is given by

$$\tau_{\mathrm{yx}} = \mu\frac{du}{dy} = (b - y)\frac{dp}{dx} \tag{24}$$

which shows that the stress distribution is linear with a magnitude of $b(dp/dx)$ at the walls (as seen in Fig. 9).

It should be noted that the constant pressure gradient and the linearity of the shear stress distribution are general results for a fully developed flow and even hold in a turbulent flow situation.

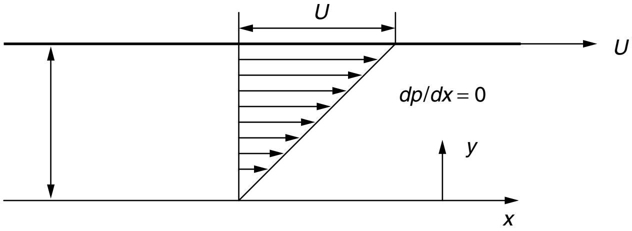

Figure 8. Plane Couette flow between two parallel plates.

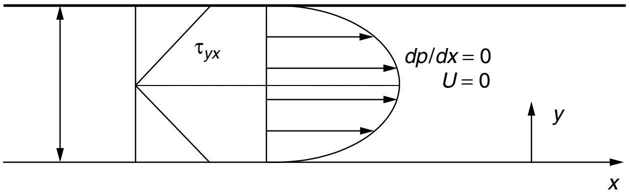

Figure 9. Plane Poiseuille flow between two parallel plates.

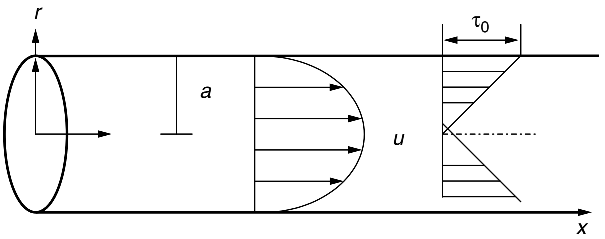

Figure 10. Laminar flow through a circular pipe.

STEADY-STATE FLOW IN A PIPE

Consider the fully developed laminar motion of a fluid through a circular pipe of radius a as shown in Fig. 10.

This type of flow is commonly known as circular Poiseuille flow. In this case, cylindrical coordinates (r, θ, x) have been used, with the x-axis coinciding with the axis of the pipe. The only nonzero component of velocity is the axial velocity u(r), and none of the variables depend on θ. The equations of motion are written in cylindrical coordinates [for detailed derivation, please refer to (2)]. The radial equation of motion gives,

$$0 = \frac{\partial p}{\partial r} \tag{25}$$

showing that p is a function of x only. The x-momentum equations gives

$$0 = \frac{dp}{dx} + \frac{\mu}{r}\frac{d}{dr}\left(r\frac{du}{dr}\right) \tag{26}$$

As the first term can only be a function of x and the second term can only be a function of r, it follows that both terms should be constants. The pressure, therefore, falls linearly along the length of the pipe. Integrating twice, we get

$$0 = \frac{r^2}{4\mu}\frac{dp}{dx} + A\ln r + B \tag{27}$$

As u must be bounded at $r = 0$, we must have $A = 0$. The wall condition, $u = 0$ at $r = a$, gives $B = -\left(\frac{a^2}{4\mu}\right)\left(\frac{dp}{dx}\right)$. The velocity profile is hence parabolic and can be represented as in Eq. (28),

$$u = \frac{r^3 - a^2}{4\mu}\frac{dp}{dx} \tag{28}$$

The shear stress at any point can be expressed using Eq. (29),

$$\tau_{\mathrm{xr}} = \mu\frac{du}{dr} = \frac{r}{2}\frac{dp}{dx} \tag{29}$$

The shear stress distribution is linear with a maximum value of $\frac{a}{2}\frac{dp}{dx}$ at the walls.

The volumetric flow rate is given by

$$Q = \int_{\mathrm{O}}^{\mathrm{a}} u 2\pi r dr = -\frac{\pi a^4}{8\mu}\frac{dp}{dx} \tag{30}$$

Equation (30) is also commonly known as the Hagen–Poiseuille law. It gives the relationship between the volumetric flow rate and the forces causing the flow, that is, the forces associated with the pressure gradient and the viscosity.

The average velocity over the circular cross section can be written as

$$\bar{v} = \frac{Q}{\pi a^2} = \frac{a^2}{8\mu}\frac{dp}{dx} \tag{31}$$

CREEPING FLOW PAST A SPHERE

When the motion of fluid is "very slow" ($\mathrm{Re} \ll 1$), the flow is said to be creeping or Stokes flow. Examples of such flows are the motion of settling of particles near the ocean bottom and the fall of moisture drops in the atmosphere. In creeping motions, viscous forces are dominant and, hence, the nonlinear inertia term present in Eq. (10) is neglected. We then have

$$\rho\frac{\partial v}{\partial t} = -\nabla p + \mu\nabla^2 v + \rho g \tag{32}$$

Equation (32) is the well-known **Stokes** equation, which is used to model creeping flows. The solution of creeping flow around a sphere was first given by Stokes in 1851. The detailed solution of velocity and pressure distributions around a sphere is quite advanced for the present text. Interested readers can refer to (2), (8), and (11) for detailed derivations. The positive pressure difference in front and negative in the rear of the sphere creates pressure drag on the sphere. Consider a fluid flowing with velocity U past a sphere of radius r. The total drag force on a sphere is obtained by calculating the total viscous stress and the normal stress and then multiplying it with the surface area of the sphere, which is found out to be

$$D = 6\pi\mu r U \tag{33}$$

One-third of the total drag comprises pressure drag, whereas two-thirds is the viscous or skin drag. It can be seen from Eq. (33) that the resistance in a creeping flow is proportional to the velocity, and is commonly known as Stokes Law.

Drag coefficient can be defined as the drag force nondimensionalized by $\rho u^2/2$ and the projected area of the sphere, which is πr^2, and stated as follows:

$$C_{\mathrm{D}} = \frac{d}{\frac{1}{2}\rho u^2 \pi r^2} = \frac{24}{\mathrm{Re}} \tag{34}$$

where $\mathrm{Re} = 2rU\rho/\mu$

Most of the problems in fluid dynamics fall under two broad categories, namely flow in channels and flow around submerged objects. Relatively simple geometrical configurations were presented in prior sections. However, real-life examples of flows in channels include pumping oil in pipes, flow of water in channels, extrusion of plastics through dyes, and fluid flows in filters. Examples of flow around submerged objects are the flow of air around an airfoil, an airplane wing, motion of fluid around particles undergoing sedimentation, etc. Flow in channels are mainly concerned with establishing a relationship between the volumetric flow rate and the pressure gradient. However, in flow past submerged objects, one is more interested in determining the relationship between the velocity of the fluid approaching the object and the drag force. As in many flow systems velocity and pressure profiles cannot be determined, the concept of friction factors/drag coefficient is found to be very useful. In other words, standard correlations are used to estimate the flow in geometrically complex systems. For an in-depth analysis, one can refer to Reference 2.

REMARKS

For understanding the physics and making generalizations in laminar flow situations, analytical solutions to simple flow configurations have been presented in the main text. With an advanced knowledge in mathematics, the readers should find it easy to solve any analytically tractable problem in laminar flow situation. Another area of utmost importance these days is Computational Fluid Dynamics, in which the equations of change are solved numerically using computers. Some of the commonly used numerical techniques for fluid flows include finite difference, finite element, and the finite volume methods. It is expected that the current text shall encourage the reader to learn numerical methods in a separate study.

BIBLIOGRAPHY

1. Alexandraou, A. (2001). *Principles of Fluid Mechanics*. Prentice-Hall, Englewood Cliffs, NJ.
2. Bird, R.B., Stewart, W.E., and Lightfoot, E.N. (1994). *Transport Phenomena*. John Wiley & Sons, New York.
3. Hanspal, N.S., Das, D.B., and Nassehi, V. (2004). Viscous flow. *Encyclopedia of Water*. John Wiley & Sons, New York.
4. Nassehi, V. (2002). *Practical Aspects of Finite Element Modelling of Polymer Processing*. John Wiley & Sons, New York.
5. Aris, R. (1997). *Vectors, Tensors and the Basic Equations of Fluid Mechanics*. Dover, New York.
6. Bird, R.B., Armstrong, R.C., and Hassager, O. (1997). *Dynamics of Polymeric Liquids: Fluid Mechanics*. John Wiley & Sons, New York.
7. Papanastasiou, T.C., Georgiou, G.C., and Alexandrou, A.N. (2000). *Viscous Fluid Flow*. CRC Press, New York.
8. Tritton, D.J. (1987). *Physical Fluid Dynamics*. Oxford Science Publications, Clarendon Press, Oxford, UK.
9. Batchelor, G.K. (1967). *An Introduction to Fluid Dynamics*. Cambridge University Press, London, UK.
10. Lighthill, M.J. (1986). *An Informal Introduction to Fluid Mechanics*. Clarendon Press, Oxford, UK.
11. White, F.M. (1991). *Viscous Fluid Flow*. McGraw-Hill, New York.

APPENDIX 1

NOMENCLATURE

A	Area
C_d	Drag coefficient
d	Diameter of the pipe
D	Drag force
F	Force
$\mathbf{g}$	Gravity force vector
p	Pressure
Q	Volumetric flow rate
r	Radius of the sphere
Re	Reynolds number
$\mathbf{v}$	Velocity vector
$\bar{v}$	Average velocity
t	Time variable
u	Velocity in x direction
v	Velocity in y direction
w	Velocity in z direction

Greek Symbols

μ	Viscosity of the fluid
τ	Stress
ρ	Density of the fluid

Subscripts

x	Stands for x component
y	Stands for y component
z	Stands for z component

FINITE ELEMENT MODELING OF COUPLED FREE AND POROUS FLOW

Navraj S. Hanspal
Atul N. Waghode
Loughborough University
Loughborough, United Kingdom

Diganta Bhusan Das
Oxford University
Oxford, United Kingdom

INTRODUCTION

The article discusses the finite element modeling (FEM) of groundwater flow in coupled free and porous domains in natural hydroenvironmental conditions. In general, free flow regime can be viewed as if the flow were in a layer of fluid or a pipe, which needs only the fluid dynamic characterization of the domain, e.g., Reynolds Number. On the other hand, porous flow regime indicates that fluid flow takes place through a porous medium, which needs the characterization for both the fluid dynamics and the medium, e.g., permeability, porosity, pore size distribution, etc. Coupled free and porous flow is observed in many hydroenvironmental conditions. As an example, conjunctive surface and subsurface flow may be cited. The flow on the surface is characterized as free flow. Some of the water from the surface seeps into the underground soil, which is a porous medium. The flow in the underground soil may be called as the subsurface or porous flow. As well as the expected media heterogeneity, e.g., presence of layers, laminations, etc., the subsurface also contains nonporous sections, e.g., conduits, fissures, buried structures that are often damaged, macropores, vugular inclusions and fractures, etc. The hydrodynamics of groundwater flow in the nonporous (e.g., vugs) and porous sections are fundamentally governed by the Navier–Stokes and Darcy equations, respectively (1–3), although simplified forms of Navier–Stokes equations are often used in practice for the nonporous domains. The overall flow behavior in the coupled nonporous and porous region can be modeled with an approach that combines Navier–Stokes and Darcy equations with suitable interfacial, boundary, and initial conditions. However, the complexity of the subsurface regions is such that it is usually not possible to predict the coupled flow behavior with confidence using the traditional analytical methods or experiments. FEM is a powerful method that can be used to validate results from experiments and/or analytical method. In other cases, where such results are not available, FEM modeling can be used independently to predict the flow field in the subsurface. The versatility of FEM lies in its ability to cope with complex and irregular geometries of the flow domain found in subsurface regimes.

In this article, the finite element method will be discussed in the context of solving coupled free and porous flow governed by Navier–Stokes and Darcy equations, respectively. Boundary conditions that may be used for simulating coupled flow problems are also discussed very briefly.

COUPLED FREE AND POROUS FLOW MODEL

The flow in the free flow regime can be mathematically expressed by the well-known Navier–Stokes equations in vector form as

$$\rho \frac{\partial \vec{v}}{\partial t} + \rho \vec{v} \cdot \nabla \vec{v} = -\nabla p \vec{\vec{\delta}} + \nabla \cdot \vec{\vec{\tau}} + \rho \vec{g} \qquad (1)$$

where ρ is the density of the incompressible fluid; $\vec{v}$, p, $\vec{\vec{\tau}}$, and $\vec{g}$ represent velocity, pressure, extra stress, and body force, respectively; and $\vec{\vec{\delta}}$ is the Kronecker delta tensor. Equation (1) expresses the conservation of momentum of free flow for an incompressible viscous fluid. The first term on the left-hand side represents the rate of change of momentum, whereas the second term represents the convective flow. On the right-hand side, the first term represents the hydrostatic pressure gradient; the second term is the deviatoric stress, whereas the last term contributes to the momentum because of the gravitational force.

The porosity of soil in the subsurface is generally low (<0.5). In such a case, the inertia forces seem to be dominated by the viscous forces, although this cannot be generalized because of its heterogeneous nature. Nevertheless, the flow in the subsurface can be mathematically expressed by the Darcy equation as

$$\rho \frac{\partial \vec{v}}{\partial t} = -\nabla p - \frac{\eta}{\vec{\vec{K}}} \cdot \vec{v} \qquad (2)$$

where $\vec{\vec{K}}$ represents the permeability tensor.

According to Newton's first law, the total mass of fluid flowing in the flow regions must be equal to the amount of fluid leaving the domain without any accumulation or consumption. This mass conservation can be represented by the continuity equation for a fluid of constant density as

$$\nabla \cdot \vec{v} = 0 \qquad (3)$$

The governing equations for the groundwater flow system are either solved analytically or numerically. Analytical models contain an analytical solution of the field equations, continuously in space or time. In numerical models, a discrete solution is obtained in both space- and time-domains by using numerical approximations of the governing partial differential equations. In general, a wide variety of numerical techniques are used in groundwater flow modeling. The most commonly used methods are the finite difference, finite element, finite volume, and analytical element methods. However, we will limit this text to discuss the use of finite element method for reasons explained earlier.

PRINCIPLES OF FINITE ELEMENT MODELING

The finite element method (FEM) is based on the principles of variational calculus. In this method, the overall geometrical domain is divided into a number of small discrete regions, termed as finite elements, with each element associated with a fixed number of nodal points. The variables within each element can be approximated by their values at each elemental node multiplied by an appropriate shape function associated with that node. This approximation is substituted in the mathematical statement of a problem, and the residual statement is weighted over the entire elemental domain. The solution is achieved by minimizing the errors introduced by the approximation.

TYPES OF FINITE ELEMENTS

Various finite element families are available, mainly based on the geometrical shapes of the elements, such as triangular or rectangular in 2-D cases or brick or prismatic elements for 3-D situations. In the context of 2-D modeling, we briefly discuss triangular and rectangular elements. Typical triangular elements are shown in Fig. 1(a) and (b).

To define the variations of the dependent variable over the elemental domain, nodal geometrical functions associated with each node, commonly known as shape functions, need to be derived. For a two-dimensional

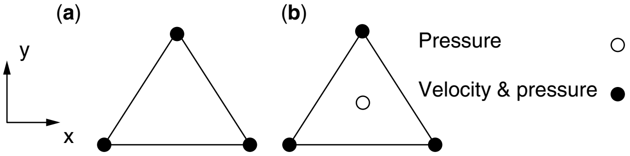

Figure 1. Triangular finite elements.

element, the approximation for the dependent variable in terms of the nodal shape functions can be written as

$$f(x,y) = \tilde{f}(x,y) = \sum_{i=1}^{p} N_i(\overline{x},\overline{y}) f_i \tag{4}$$

where the overbar indicates the values at each node of the element.

Multidimensional elements should also satisfy the following conditions:

$$\sum_{i=1}^{p} N_i(\overline{x},\overline{y}) = 1$$

$$N_i(\overline{x}_j,\overline{y}_j) = 1, \quad i = j$$

$$N_i(\overline{x}_j,\overline{y}_j) = 0, \quad i \neq j$$

In case of a three-noded element, shown in Fig. 1(a), the shape functions are

$$\left.\begin{aligned} N_i &= \frac{1}{2A}[a_i + b_i x + c_i y] \\ N_j &= \frac{1}{2A}[a_j + b_j x + c_j y] \\ N_k &= \frac{1}{2A}[a_k + b_k x + c_k y] \end{aligned}\right\} \tag{5}$$

where

$$\begin{vmatrix} 1 & X_i & Y_i \\ 1 & X_j & Y_j \\ 1 & X_k & Y_k \end{vmatrix} = 2A$$

$$\begin{aligned} a_i &= X_j Y_k - X_k Y_j, & b_i &= Y_j - Y_k & \text{and} \quad c_i &= X_k - X_j \\ a_j &= X_k Y_i - X_i Y_k, & b_j &= Y_k - Y_i & \text{and} \quad c_j &= X_i - X_k \\ a_k &= X_i Y_j - X_j Y_i, & b_k &= Y_i - Y_j & \text{and} \quad c_k &= X_j - X_i \end{aligned}$$

where (X_i, Y_i), (X_j, Y_j), and (X_k, Y_k) are the coordinates corresponding to the nodes i, j, and k, respectively. For detailed derivations, the readers can refer to Nassehi (4) and Segerlind (5). A four-noded rectangular finite element is shown in Fig. 2. The shape functions for this element can be written as

$$\left.\begin{aligned} N_i &= \frac{1}{4}\left(1 - \frac{q}{b}\right)\left(1 - \frac{r}{a}\right) \\ N_j &= \frac{1}{4}\left(1 + \frac{q}{b}\right)\left(1 - \frac{r}{a}\right) \\ N_k &= \frac{1}{4}\left(1 + \frac{q}{b}\right)\left(1 + \frac{r}{a}\right) \\ N_k &= \frac{1}{4}\left(1 - \frac{q}{b}\right)\left(1 + \frac{r}{a}\right) \end{aligned}\right\} \tag{6}$$

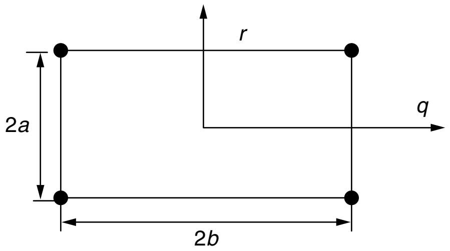

Figure 2. A linear rectangular finite element.

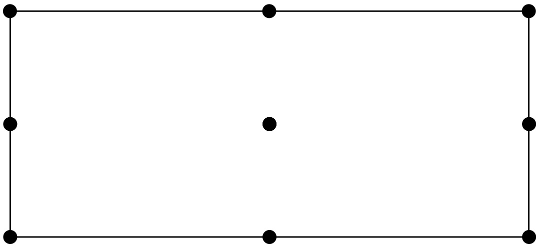

Figure 3. A biquadratic rectangular finite element.

The second member of this group is the nine-noded biquadratic rectangular element with equal order interpolation functions for velocity and pressure, as shown in Fig. 3.

Higher order elements of this family can also be readily generated using the tensor products of the higher order one-dimensional interpolation functions. In addition to all these classes of finite elements, some nonstandard elements exist, such as the Taylor–Hood elements specially designed for the solution of incompressible flow problems. In the Taylor–Hood elements, interpolation of pressure is always based on a lower order polynomial than the velocity interpolation functions.

PARAMETRIC MAPPING

Finite element method can deal with irregular or curved geometries in a straightforward and systematic way using parametric mapping (6). A regular element, known as the master element, is chosen and a local finite element approximation based on this master elemental shape function is selected. The Master element is mapped into the global coordinates to generate required curved irregular element. Two independent elements can be used in the finite element analysis, one for approximating geometry and the other for interpolation of field variable. When a typical element in global coordinate system is transformed to a master element, it is requisite to transform the global coordinate system to local coordinate system of the master element. Depending on the relationship between the degree of approximation used for coordinate transformation and for the field variable, three different types of mapping schemes can be employed.

1. *Subparametric Transformation*: Shape functions used for geometry are of lower order than that used for dependent variable.
2. *Isoparametric Transformation*: Shape functions used for mapping geometry are of the same order as for the dependent field variable.
3. *Superparametric Transformation*: Shape functions for mapping geometry are of higher order than that used for the field variable.

FINITE ELEMENT SCHEMES FOR INCOMPRESSIBLE FLOW PROBLEMS

Application of the weighted residual method to the solution of incompressible flow equations of continuity and momentum conservation is commonly based on two different schemes.

U-V-P Scheme

The name U-V-P implies U as the component of velocity in x-direction, V as a component of velocity in y-direction, and P as pressure with all these variables calculated simultaneously in the solution scheme. Zienkiewiez and Taylor (6) classified the U-V-P scheme to general category of mixed finite element methods. In this technique, both velocity and pressure are termed as primitive variables and are discretised as unknowns. In the U-V-P method, the weighted residual statements for the groundwater flow model can be formulated as

Free Flow Zone

$$\int_{\Omega} N_P \nabla\cdot\bar{v}\,d\Omega = 0$$

$$\int_{\Omega} \rho N_V\left(\frac{\partial\bar{v}}{\partial t} + \bar{v}^0\nabla\bar{v}\right)d\Omega = \int_{\Omega} N_V\left(-\nabla\bar{p}\delta + \nabla\cdot\bar{\tau} + \rho g\right)d\Omega \tag{7}$$

Porous Flow Zone

$$\left.\begin{array}{l} \int_{\Omega} N_P \nabla\cdot\bar{v}\,d\Omega = 0 \\ \int_{\Omega} \rho N_V\left(\dfrac{\partial\bar{v}}{\partial t}\right)d\Omega = \int_{\Omega} N_V\left(-\nabla\bar{p} + \dfrac{\eta}{K}\cdot\bar{v}\right)d\Omega \end{array}\right\} \tag{8}$$

where an over bar indicates interpolated variables over an element. N_v and N_P are the appropriate basis functions or weight functions.

The interpolating functions for the field variables are substituted in the formulations, and the equations are then integrated, which yields a final set of finite element working equations (7).

Penalty Scheme

The penalty method is a standard technique to reduce the number of variables in the flow calculation (4). In the penalty method, the pressure is expressed in terms of the incompressibility condition (4)

$$p = -\lambda(\nabla\cdot v) \tag{9}$$

This relation is then substituted in the momentum balance equation, after which the weighted residual finite element method is applied. λ is the penalty parameter that generally has a very large value. Elimination of pressure as a field variable leads to a compact set of working equations reducing the computational cost. The only problem with the penalty scheme lies in the selection of appropriate value of the penalty parameter to avoid ill conditioning of the solution matrices.

MODELING SCHEMES BASED ON U-V-P FORMULATION

In the modeling of incompressible flows, the Ladyzhenskaya–Babuska–Brezzi (LBB) stability condition must be satisfied (8), which poses a severe restriction on the type of approximating functions that can be used in finite element formulations. Different finite element models based on various strategies have been developed that satisfy the LBB condition. They range from the use of elements generating unequal order interpolation functions (6) for the field unknowns to a perturbed form of continuity equation (9) representing slightly compressible fluids.

Taylor–Hood Scheme

In order to satisfy the LBB condition, unequal order interpolation functions for velocities and pressure can be used. In this scheme, the incompressibility constraint can be used without any modifications (i.e., Eq. (3) remains as $\nabla\cdot\vec{v} = 0$). In addition to the requirement of LBB criterion, it is also known that because of the incompatibility of the operators in the Navier–Stokes and Darcy equations, the approximating function spaces used for the numerical solution of these equations need to be different. In essence, the Darcy equation should be treated as an elliptic Poisson equation, where the degrees of freedom should be kept as low as possible. This scheme is developed using C^0 continuous Taylor–Hood elements (10). This element is a member of the bubble element family (11), and it is expected that all other members of this family will generate stable solutions for incompressible flows with a variety of time-stepping schemes. A typical Taylor–Hood element can be seen in Fig. 4.

Perturbed Continuity Scheme

The perturbed continuity method (12) depends on the inclusion of the term $\dfrac{1}{\rho c^2}\dfrac{\partial p}{\partial t}$ in the continuity equation (i.e., Eq. 3), where ρ is the fluid density and c is the velocity of sound in the fluid. Therefore, this results in the incorporation of the temporal derivative of pressure in the continuity equation, which in physical terms amounts to considering the fluid to be slightly compressible (9). To use the time-dependent continuity equation, transient terms, such as $\dfrac{\partial\vec{v}}{\partial t}$, should also be added into the Navier–Stokes and the Darcy equations, which inherently forces the use of a time-stepping scheme in the numerical solution of the coupled flows. The perturbed continuity scheme allows the use of equal order interpolation functions for velocity and pressure. Steady-state conditions can be obtained iteratively using different time-stepping schemes.

TEMPORAL DISCRETIZATION

The finite element solution of time-dependent flow problems produces a system of linear first-order differential

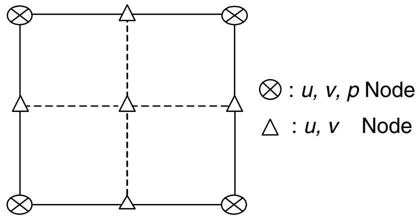

Figure 4. Taylor–Hood element.

equations in the time-domain. The method for solving time-dependent partial differential equations is the Partial Discretization Method. In this method, the space-time domain is not discretized as a whole, but the time derivatives are treated separately.

θ Time-Stepping Method

In this method, the time derivatives are kept unchanged, and the spatial discretization is carried out to form a weighted residual statement, which yields a system of ordinary differential equations in time. Consider a partial differential equation

$$\frac{\partial \phi(x,t)}{\partial t} + f[\phi(x,t)] - c = 0 \tag{10}$$

where f is a linear differential operator with respect to x.

At any time level θ, the temporal derivative is approximated by a forward difference as

$$\left.\frac{\partial \phi}{\partial t}\right|_{\theta} = \frac{\phi|_{n+\theta\Delta t} - \phi|_n}{\theta \Delta t} = \frac{\phi|_{n+1} - \phi|_n}{\Delta t} \tag{11}$$

In general, this differential system can be shown in matrix form as

$$[M]_\theta \left\{\phi^\bullet\right\}_\theta + [K]_\theta\{\phi\}_\theta = \{F_\theta\} \tag{12}$$

where $0 \leq \theta \leq 1$.

The terms, except the differential term, in the equation are approximated using a linear interpolation as

$$[K]_\theta\{\phi\}_\theta = (1-\theta)[K]_n\{\phi\}_n + \theta[K]_{n+1}\{\phi\}_{n+1} \tag{13}$$

$$\{F\}_\theta = (1-\theta)\{F\}_n + \theta\{F\}_{n+1} \tag{14}$$

Substitution of these equations and carrying out algebraic manipulation gives

$$\begin{aligned}([M]_\theta + \theta\Delta t[K]_{n+1})\{\phi\}_{n+1} = {} & ([M]_\theta - (1-\theta)\Delta t[K]_n)\{\phi\}_n \\ & + ((1-\theta)\{F\}_n + \theta\{F\}_{n+1})\Delta t \end{aligned} \tag{15}$$

On the basis of this equation, a global set is derived and solved to obtain unknowns at time level $n+1$ using the known value at time level n.

Taylor–Galerkin Method

The Taylor–Galerkin time-stepping technique is elaborated in detail in the work by Donea (13). Consider the same differential equation. Taylor series expansion of variable ϕ within the time steps n and $n+1$ gives

$$\phi^{n+1} = \phi^n + \Delta t \left.\frac{\partial \phi}{\partial t}\right|^n + \frac{1}{2}(\Delta t)^2 \left.\frac{\partial^2 \phi}{\partial t^2}\right|^n + \ldots.. \tag{16}$$

The temporal derivatives in Eq. (15) can be substituted from the parent differential Eq. (9) evaluated at that time level. Hence, we have

$$\left.\frac{\partial \phi(x,t)}{\partial t}\right|_n = -f[\phi(x,t)]|_n + c \tag{17}$$

Similarly, the second-order derivative at time step n can be replaced as

$$\left.\frac{\partial^2 \phi(x,t)}{\partial t^2}\right|_n = \frac{\partial}{\partial t}\{-f[\phi(x,t)]|_n + c\} \tag{18}$$

All the temporal derivatives in the Taylor series expansion can be substituted principally from the original differential equation. Taylor series expansion of the field variable is truncated, keeping only first few terms. Accuracy of the scheme is solely dependent on the highest order of derivative remaining in the expansion after truncation.

MATHEMATICAL MODELS FOR COUPLED FLOWS

The main challenge is to construct a sound mathematical formulation for the coupled free and porous flow dynamics (14). The selection of appropriate approximating function spaces for the numerical solution of these equations, which guarantee unified stability of both the free flow and porous flow models, is not straightforward (15). In order to circumvent this problem, most of the solution algorithms are based on two approaches.

The first approach employs the Brinkman's equation to describe the flow on a free/porous interface (16). The Brinkman's equation includes the Laplacian of the velocity field and, hence, is of the same order as the Stokes equation. Therefore, its use across a free/porous interface ensures continuity of velocity and pressure fields. The Brinkman's equation is only applicable to media of high permeability (17). In addition, the Brinkman's equation is constructed using a parameter called "effective viscosity." The second approach is based on the imposition of a slip-wall boundary condition at the free/porous interface, as suggested by Beavers and Joseph (18). Using this approach, the interfacial boundary condition is modified to match the free and porous flow conditions by an auxiliary relationship. This approach is empirical and deduced from a simple one-dimensional situation, and its extension to multidimensional cases is not well understood. The value of the slip coefficient depends on many physical parameters such as the geometrical features of the interface.

Discacciati et al. (19) developed a model for coupling of Navier–Stokes and Shallow Water equations for surface flows and Darcy equations for groundwater flows. They applied the boundary conditions similar to that of the Beavers and Joseph (18) at the free/porous interface. They developed an iterative method for the solution, which requires the solution of one problem in the fluid part and one in the porous medium.

Recently, Burman and Hansbo (15) proposed a unified approach for the solution of the coupled Stokes/Darcy problem using an edge stabilization technique in conjunction with the use of mixed P_1/P_0 finite elements. The edge stabilization, as named by these authors, involves the splitting of the viscous stress term into its normal and tangential components at the free/porous interface. The normal stress component is then equated to the pressure differential across the free/porous interface.

NODAL REPLACEMENT SCHEME FOR COUPLED FLOWS

In addition to the schemes described in the previous section, Nassehi et al. (7) have developed a nodal replacement technique based on finite element method, which can be used for modeling coupling the free and the porous flow dynamics in two-dimensional situations.

The scheme is based on the imposition of Darcy equation as a boundary condition for the Navier–Stokes equations at the free/porous interface (and vice-versa at the porous/free interface). The field variable velocities and pressures are forced to be continuous in both the free and porous regions across the interface. In the stiffness matrix corresponding to the elements in the free flow regime, which are also present on the free/porous interface, the Navier–Stokes terms corresponding to the interfacial nodes are replaced by the appropriate form of the discretised Darcy components and vice-versa. This procedure is illustrated in Fig. 5.

BOUNDARY CONDITIONS FOR COUPLED FLOW DOMAINS

For the solution of free flow equations, no-slip-wall boundary conditions (20) are generally specified on the impermeable walls of the domain. A velocity boundary condition is given at the inlet of the flow domain, whereas at the exit, stress-free boundary conditions (21) are used.

For the solution of porous flow equations, a partial slip boundary condition (4) may be specified, and at the flow exit, a pressure datum can be imposed. In the coupled flow situation, the Darcy equation is imposed as the boundary condition for the Navier–Stokes equation at the free/porous interface.

COMMERCIAL GROUNDWATER FLOW MODELS

There are many commercially available groundwater flow models based on a numerical method, which have been developed in the past. They are widely used in industries and academia for carrying out a variety of numerical studies. For example, FEMWATER is commonly used for 3-D modeling of flow and transport in groundwater systems. Salient features of some of the other frequently used groundwater flow tools have been presented in Table 1.

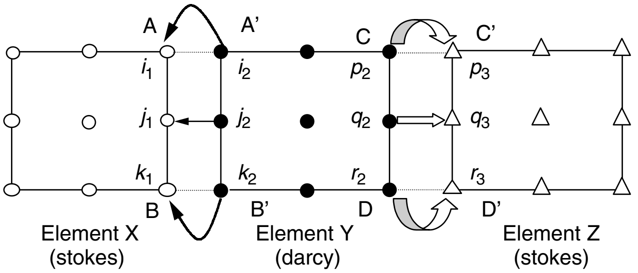

Figure 5. Schematic representation for linking of Navier–Stokes and Darcy regimes.

Table 1. Commercial Finite Element Groundwater Models

S No.	Model	Application
1	**3DFEMFAT**	3-D finite-element model of flow and transport through saturated-unsaturated media.
2	**AQUA3D**	3-D groundwater flow and contaminant transport model.
3	**ChemFlux**	Finite element mass transport model.
4	**FEFLOW**	Finite element subsurface flow system. FEFLOW is a finite-element package for simulating 3-D and 2-D fluid density coupled flow, contaminant mass (salinity), and heat transport in the subsurface.
5	**FLONET/TRANS**	2-D cross-sectional groundwater flow and contaminant transport modeling.
6	**GMS**	Groundwater Modeling Environment for MODFLOW, MODPATH, MT3D, RT3D, FEMWATER, SEAM3D, SEEP2D, PEST, UTCHEM, and UCODE.
7	**MicroFEM**	Finite-element program for multiple-aquifer steady-state and transient groundwater flow modeling.
8	**MODPATH**	3-D particle tracking program for MODFLOW.
9	**MOFAT**	Multiphase (water, oil, gas) flow and multicomponent transport model.
10	**SUTRA**	2-D saturated/unsaturated transport model.
11	**VAM2D**	2-D Variably-Saturated Groundwater Analysis Model.
12	**WinTran**	Groundwater Flow and Finite-Element Contaminant Transport Model.
13	**FEMWATER**	3-D Finite Element Model for Coupled Groundwater Flow and Contaminant Transport.

BENEFITS OF FINITE ELEMENT MODELING FOR MODELING COUPLED FLOW

The finite element method is generally found to be more robust and reliable toward obtaining stable and convergent solution, which is because it involves the minimization of errors introduced because of variable approximation. The Galerkin weighted residual finite element method is very flexible and can cope with irregular, complex, and curved geometries of the problem domain, which is the normal case in the free surface and subsurface flows. The previously developed mathematical models for the coupled free/porous problems are based on concepts like the use of the Brinkman's equation or the specification of slip-wall boundary condition at the interface. Although both of these methodologies appear to be simple, they seem to be compromising their accuracy and lacking flexibility with suitability only for specific problem domains. The novel nodal replacement technique available in FEM is a simple mathematical concept that links the physics behind the free flow and porous flow phenomena without any simplified assumptions and can be easily extended to three-dimensional flow regimes. The linking of the free and the porous flow regimes is achieved without imposition of any artificial boundary conditions at the free/porous interface. It can be concluded that the FEM schemes are highly robust and reliable for the combined flow without any mathematical problems developing from stability requirements, numerical incompatibility, or time-stepping schemes.

SUMMARY

The coupled free and porous flow phenomenon has been discussed in the context of natural hydroenvironmental conditions. Mathematical statements of free and porous flow models have been represented elaborating physical significance of each flow regime. Various finite element families have been listed and classified mainly according to the geometric shapes of the elements. Finite element formulations for the incompressible flow problems based on two common methods, namely the U-V-P and penalty schemes, have been derived. In the context of unsteady-state studies, time-stepping schemes like θ-method and Taylor–Galerkin methods have been explained for the temporal discretization of the governing equations. The currently available mathematical models for the coupled flow problems have been discussed regarding their conceptual details and related assumptions. A recently developed nodal replacement scheme for the coupled flow situation has been analyzed with the help of a pictorial representation and required boundary conditions.

Nomenclature

v	Fluid velocity
$\vec{v}$	Velocity vector
p	Fluid pressure
$\vec{g}$	Gravitational acceleration vector
$\vec{\vec{K}}$	Permeability tensor
A	Area of the finite element
N	Shape functions
c	Speed of sound in fluid

Subscripts

x	In x-direction
y	in y-direction
v	For velocity
p	For pressure

Greek Symbols

ρ	Fluid Density
η	Fluid viscosity
θ	Time-stepping parameter
Ω	Elemental domain
λ	Penalty parameter
∇	Gradient
$\vec{\vec{\delta}}$	Kronecker delta tensor
$\vec{\vec{\tau}}$	Extra stress tensor

BIBLIOGRAPHY

1. Zimmerman, R.W. and Yeo, In-W. (2000). Fluid flow in fractures: from the Navier–Stokes equations to the cubic law, In: *Dynamics of Fluids in Fractured Rock*. B. Faybishenko, P.A. Witherspoon, and S.M. Benson (Eds.). Geophysical Monograph Series, American Geophysical Union, Washington, DC, pp. 213–224.
2. Das, D.B. and Nassehi, V. (2003). *Water Resour. Res.* **39**(3): 1072.
3. Abrogast, T., Brunson, D.S., Bryant, S.L., and Jennings, Jr., J.W. (2004). A preliminary computational investigation of a macro-model for vuggy porous. *Proceedings of the Computational Methods in Water Resources 2004 International Conference*, Chapel Hill, NC, June 13–17.
4. Nassehi, V. (2002). *Practical Aspects of Finite Element Modelling of Polymer Processing*. Wiley, Oxford, UK.
5. Segerlind, L.J. (1984). *Applied Finite Element Analysis*, 2nd Edn. Wiley, New York.
6. Zienkiewicz, O.C. and Taylor, R.L. (2002). *The Finite Element Method*, Vol. 3, 5th Edn. Butterworth-Heinemann, Oxford, UK.
7. Nassehi, V., Hanspal, N.S., Waghode, A.N., and Wakeman, R.J. (2004). Finite element simulation of flow through pleated cartridge filters. *Chemical Engineering Science*.
8. Reddy, J.N. (1993). *An Introduction To Finite Element Method*, 2nd Ed. McGraw-Hill, London, UK.
9. Zienkiewicz, O.C. and Wu, J. (1991). *Int. J. Numeric. Methods Engineer.* **32**: 1189–1203.
10. Taylor, C. and Hood, P. (1973). *Computers and Fluids*. **1**: 73–101.
11. Fortin, M. and Fortin, A. (1985). *Newer Elements for Incompressible Flow: Finite Element Methods*. Vol. 6, Wiley, New York.
12. Hanspal, N.S., Waghode, A.N., Nassehi, V., and Wakeman, R.J. (2004). Numerical analysis of coupled Stokes/Darcy flows in industrial filtrations. *Transport Porous Media*.
13. Donea, J. (1984). *A Taylor-Gallerkin Method for Convective Transport Problem. Int. J. Numeric. Methods Eng.* **20**: 101–119.
14. Layton, W.J., Schieweck, F., and Yotov, I. (1993). *SIAM J. Numeric Anal.* **40**(6): 2195–2218.
15. Burman, E. and Hansbo, P. (2002). *A Unified Stabilized Method for Stokes' and Darcys' Equations*. Chalmers Finite Element Center, Chalmers University of Technology, Göteborg, Sweden.
16. Brinkman, H.C. (1947). *Appl. Scientif. Res.* **A1**: 27–34.
17. Kim, S. and Russell, W.B. (1985). *J. Fluid Mechan.* **154**: 269–286.
18. Beavers, G.S. and Joseph, D.D. (1967). *J. Fluid Mechan.* **30**(1): 197–207.
19. Discacciati, M., Miglio, E., and Quarteroni, A. (2002). *Appl. Numeric. Mathemat.* **43**: 57–74.
20. Bird, R.B., Stewart, W.E., and Lightfoot, E.N. (2000). *Transport. Phenomena*. Wiley, New York.
21. Papanastasiou et al. (1992).

UNCONFINED GROUNDWATER

Aly I. El-Kadi
University of Hawaii at Manoa
Honolulu, Hawaii

INTRODUCTION

An unconfined aquifer is a geologic formation with subterranean water that is not confined by relatively impermeable layers. Part of the precipitation and irrigation water as well as water from recharge basins, reservoirs, and ponds moving downward from the soil layer will reach the saturated groundwater zone. The top of this zone is the water table, which separates the fully saturated zone from a partially saturated zone. The latter zone is simply referred to as the unsaturated zone, but it is also known as the zone of aeration or vadose zone. Generally, the water in the unsaturated zone is under less than atmospheric pressure. The unsaturated zone includes the capillary fringe, directly above the water table, which is near saturation but still under negative pressure relative to the atmospheric pressure.

Often, groundwater flow discharges into or receives water from surface water bodies such as river channels or lakes. An effluent or gaining stream is supplied by groundwater, which is usually the case in humid climates. An influent or losing stream supplies water to the groundwater and is a characteristic of arid climates. In addition to the flow to streams from the saturated zone, there exists throughflow, which is the sporadic horizontal water flow above the water table. Such a flow normally takes place when parts of the soil are completely saturated. Flow rates are usually low, with maximum flow likely occurring on steep slopes and in pervious sediments.

The presence of a low conductivity zone or an impermeable layer (called perching bed) can lead to the formation of a perched water table above the aquifer's main water table. Springs are often formed when a perched water table intersects the ground surface. Perched groundwater can be permanent if recharge is frequent enough to maintain a saturated zone above the perching bed. Intermittent cases occur when recharge is not large or frequent enough to prevent the perched water from disappearing at times due to drainage through the perching bed or over its edges.

Water input to an unconfined aquifer mainly includes recharge due to rainfall and return irrigation, direct injection, and leakage from surface water bodies. Major outflow can include pumping of groundwater, plant usage, evaporation from the soil surface, and leakage onto surface water bodies. Water table elevation will change according to the relative magnitude of groundwater replenishment versus depletion.

When compared to confined aquifers, unconfined aquifers are more susceptible to contamination by dissolved chemicals, such as agricultural chemicals, and nonaqueous phase liquids (NAPLs), such as hydrocarbon products, due to the absence of a confining layer. Light nonaqueous phase liquids (LNAPLs), such as gasoline, can form a floating lens on top of the water table. The lens moves in the direction of groundwater flow. A dissolved phase plume can form in the saturated zone that also moves in the direction of water flow. A plume of dense NAPLs (DNAPLs) can penetrate down through the unsaturated and saturated zones and move by gravity in a direction that is independent of groundwater flow. A dissolved phase plume due to DNAPLs can also exist.

WATER FLOW

Mathematical Formulation

A description of processes involved in saturated/unsaturated flow in unconfined aquifers is covered by Bear (1) and more recently by El-Beshry (2) and Charbeneau (3). Flow in parts of the unconfined aquifer is characterized by the existence of water and air. Various mathematical approaches for flow in unconfined aquifers are described by Brutsaert and El-Kadi (4). The approaches cover various degrees of complexities. The engineering approach includes a two-equation system for the flow of water and air. If the air phase is considered as continuous and at atmospheric pressure, a single equation that describes water flow will result. The equation will include the effects of unsaturation and solid-matrix and water compressibility. If compressibility effects are ignored, the equation will reduce to the well-known Richards equation (5), which can be written in the following form:

$$\nabla \cdot (K \, \nabla h) = \frac{\partial \theta}{\partial t} \tag{1}$$

where K is the hydraulic conductivity, h is the hydraulic head, θ is the water content, and t is time. The hydraulic head is related to pressure head ψ by $h = \psi + z$, in which z is the elevation above an arbitrary datum. Equation 1 is also applicable to the saturated zone, where K equals its saturated value and θ equals porosity. Following solution of Equation 1, the water table would be identified as the surface where the pressure head value equals zero.

Needed for a complete formulation are the relationships $K(\psi)$ and $\theta(\psi)$. Popular forms of these relationships are provided by van Genuchten (6):

$$S_e = [1 + |\alpha\psi|^n]^{-m} \tag{2}$$

in which S_e is defined as

$$S_e = \frac{(\theta - \theta_r)}{(\theta_s - \theta_r)} \tag{3}$$

In Equations 2 and 3, θ_s and θ_r are saturated and residual water content, respectively, and α, n, and m are parameters with n and m related by

$$m = 1 - 1/n \tag{4}$$

The hydraulic conductivity function is given by van Genuchten (6) as

$$K = K_s S_e^{0.5}[1 - (1 - S_e^{1/m})^m]^2 \tag{5}$$

where K_s is the saturated hydraulic conductivity.

Solutions for Equation 1 are complicated by its nonlinearity due to the dependence of the conductivity and saturation on hydraulic head. Analytical solutions are available for simplified conditions, such as those for infiltration and discharges from sources in the unsaturated zone (6,7).

Numerical solutions are suitable for general problems involving more realistic flow conditions and spatially varied parameters, such as hydraulic conductivity. Examples of numerical models include VS2D (8,9), HYDRUS (10; http://www.pc-progress.cz/Fr_Hydrus.htm), and FEMWATER (11). VS2D is in the public domain and can be downloaded from the U.S. Geological Survey web site: http://www.usgs.gov/software/vs2di.html. FEMWATER is included in the commercial package GMS (http://www.ems-icom), which serves as the data pre- and postprocessor for the model. The original version of the model is in the public domain (http://www.epa.gov/ceampubl/gwater/femwater/). FEMWATER is a three-dimensional model, in contrast to HYDRUS and VS2D, which are two-dimensional models.

Care is required in choosing the size of the mesh or grid for numerical models. Small sizes, on the order of centimeters, are needed in some cases for achieving good numerical accuracies, mostly due to the nonlinear nature of Equation 1 (12). Such a restriction puts a severe limitation on the use of numerical models for large-scale three-dimensional aquifers with thick unsaturated zones.

Solutions for Equation 1 also require specification of initial and boundary conditions. The initial condition may include location of the water table and the relevant distribution of pressure head. Boundary conditions may include specified head or water flux rates. Wells are usually treated as sources (injection) or sinks (pumping).

In some applications, flow in the unsaturated zone can be overlooked and only the flow in the saturated zone is considered. In this case, a number of mathematical formulations exist, which include the use of a specific yield parameter that assumes instantaneous flow from the unsaturated zone toward the water table (1). Saturated flow models, such as MODFLOW (13), can be used in this regard. The limitations have to be recognized, however, especially for aquifers with a thick unsaturated zone compared to the saturated thickness of the aquifer.

Infiltration Models

Infiltration is the process of water entry to the subsurface en route to the saturated zone. Infiltration quantification is an important part of estimating runoff needed for watershed assessment. Numerical models can simulate infiltration. However, their use is cumbersome for watershed assessment. Empirical and physically based models are useful in this regard. Such models are reviewed by Philip (14) and El-Kadi (15). Popular models include those developed by Green and Ampt (16), Kostiakov (17), Horton (18), and Philip (19). More recent models include those by Smith and Parlange (20), Parlange et al. (21), and Parlange and Haverkamp (22).

Models for Aquifer Test Analysis

Aquifer test analysis is a popular technique for estimating hydraulic parameters that control an aquifer's capacity to transmit and store water. Both analytical and numerical solutions can be used in the process through an automatized or manual fitting process. The nature of the flow in an unconfined aquifer causes failure of the classic Theis solution (23), which is based on restrictive assumptions relevant to the nature of confined aquifer flow. Most analytical models attempt to account for the water table condition and flow in the unsaturated zone by assigning a certain boundary condition at the water table. Theoretical developments and model assumptions can be found in Freeze and Cherry (24), Kruseman and de Ridder (25,26), Driscoll (27), Dominico and Schwartz (28), Walton (29), and Fetter (30).

Popular aquifer test models include those by Boulton (31), Boulton and Streltsova (32), Boulton and Pontin (33), Dagan (34), Neuman (35–37), Moench (38,39), and Moench et al. (40). The pioneering model of Boulton (31) approximates drainage from the zone above the water table by using an exponential relation containing an empirical parameter considered as a delay index. Moench (38) substituted Boulton's (41) convolution integral for Neuman's (35,36) boundary condition for the free surface. Additional improvements were introduced by Moench (39) and Moench et al. (40).

Available commercial pump-test packages include AquiferTest (42; http://www.waterloohydrogeologic.com). The package includes the solutions of Moench (43) and Neuman (37). The package also includes a solution with a correction term that allows the use of the Theis (23) and Cooper–Jacob (44) solutions for the analysis of pumping test data for an unconfined aquifer.

CONTAMINATION

Mathematical Formulation

As discussed earlier, contamination can be caused by dissolved chemicals or NAPLs. Subsurface solute transport and fate in the unsaturated/saturated zone are described by Charbeneau (3), including degradation due to biological, chemical, and photochemical processes, volatilization, immobilization, sorption, ion exchange, advection, diffusion, dispersion, residual saturation, and preferential flow. Further discussion of these processes is provided by Hughes (45).

For conservative dissolved chemicals, that is, in the absence of chemical and biological transformations, the governing equation can have the following form (3):

$$\nabla \cdot (\theta D \cdot \nabla C) - \nabla(VC) = \frac{\partial \theta C}{\partial t} \tag{6}$$

where C is the solute concentration, D is the dispersion coefficient, and V is the seepage velocity. The seepage velocity is estimated by dividing Darcy's flux by the effective porosity, which is defined as the porosity available for water flow. Darcy's flux is estimated by solving Equation 1. Mathematical formulation for nonconservative chemical transport is covered by Freeze and Cherry (24), Fetter (46), and Charbeneau (3). Equation 6 is applicable to the saturated zone, where θ equals the effective porosity, which is constant with respect to time but can be spatially varied.

Multiphase flow in cases of NAPL contamination requires the use of various phase equations that include water flow, liquid product flow, gas flow, and dissolved phase transport and fate processes. Discussion of concepts involved, as well as practical aspects of addressing NAPL contamination, is covered by Bedient et al. (47) and Charbeneau (3).

A free interactive guide for the fate, transport, and remediation of LNAPLs is available for downloading from the website of the American Petroleum Institute: http://www.groundwater.api.org/lnaplguide/. The guide provides understanding of principles and some tools for evaluating the mobility, stability, and recoverability of LNAPLs. The guide is designed to provide an overall approach for evaluating the LNAPL at a site, assessing its potential risk, quantitatively defining mobility and recoverability, developing remedial strategies, and examining methods to enhance site-closure opportunities.

Saltwater Contamination

In addition to potential contamination by dissolved chemicals and NAPLs, saltwater can be a contamination threat for unconfined aquifers in coastal zones and islands. Classic textbooks, such as those by Bear (1), Todd (48), and Fetter (30), discuss the subject, including theoretical developments and practical aspects of the processes involved. Density effects, which can be justifiably overlooked in the cases for dissolved chemicals, dominate saltwater contamination situations.

For coastal aquifers, fresh groundwater discharges to the ocean and mixes with saltwater. The contact between saltwater and fresh water creates a zone of mixing that is characterized by brackish water with a salinity gradient. Such a transition can be sharp, with a thin mixing zone. However, under freshwater head fluctuations, excessive well pumping, and tidal fluctuations, the mixing zone can be quite large.

For unconfined aquifers, fresh groundwater moves upward to discharge near the shoreline, while cyclic water flow occurs in the salty water near the interface (49). Under ideal conditions that include a sharp interface and hydrostatic condition, the depth to the freshwater extent below sea level, Z, is expressed by the Ghyben–Herzberg equation (50,51):

$$Z = \frac{\rho_f}{\rho_s - \rho_f} H \tag{7}$$

where H is the elevation of the water table above sea level, ρ_f is the density of fresh water, and ρ_s is the density of saltwater. Such a relationship would provide a ratio of 1 to 40 for H to Z.

Equation 7 is also valid under conditions for mainly horizontal flow, which is usually valid away from the shoreline. Todd (48) and Fetter (30) provide other useful analytical expressions, such as that of the depth to the interface between fresh water and saltwater for an oceanic island with groundwater recharge. The U.S. Geological Survey has developed the numerical modes SHARP (52), which can be used in simulating freshwater and saltwater flow under the assumption of a sharp interface. The model is available from the URL site http://water.usgs.gov/software/sharp.html.

AQUIFER VARIABILITY

Variability of soil and aquifer material is common and may cause substantial difficulties in assessing water flow and contaminant transport and fate in unconfined systems. Jury et al. (53,54) summarized results from a number of unsaturated zone field studies regarding variability of hydraulic conductivity, porosity, and infiltration. Properties can be highly heterogeneous due to the nature of geological formations, including the existence of various geological units of contrasting features. Fractures due to structural or other factors can cause preferential flow and transport (55,56). Overlooking preferential flow may lead to unacceptable errors in estimating travel times and the risk of contamination (57).

Estimating parameters under variable conditions is difficult due to the costs involved in physically collecting data and the potential for affecting the integrity of the aquifer. Inverse methods involve the indirect estimation of parameters based on hydraulic head or solute concentration values. An example of available software is PEST (58), which can be used with various models in estimating their parameters. However, questions arise regarding model uniqueness.

Approaches to model systems under variable conditions include both deterministic approaches and stochastic approaches. For deterministic approaches, continuum models are adapted in which it is assumed that a representative control volume exists where aquifer or soil properties are defined. For example, a continuum model for flow in an unconfined aquifer would assume validity of the Richards equation but would use soil hydraulic properties that are modified to reflect a fracture-porous system (59). Examples of soil properties for this system are provided by Durner (60) and Mohanty et al. (61).

Other deterministic approaches include the use of discrete fracture models to account for water flow and solute transport in fractured rocks by taking into consideration the location and aperture of rock fractures (62). Charbeneau (63) and Germann (64) adopted a kinematic wave approach to account for water flow in fracture or macropore domains, which are also known as dual- or double-porosity domains. Two-domain models were developed by Duguid and Lee (65), Gerke and van Genuchten (66), Ray et al. (67), and Vogel et al. (68), while a three-domain model was developed by Gwo et al. (69). Such models simulate mobile water in soils or fractured porous media conceptualized to have two- or three-pore domains. These models are commonly referred to as dual- or multipermeability models, where advective movements of chemicals are considered in the matrix domain. Zyvoloski et al. (70, 71) developed a general purpose finite-element code for the simulation of water, heat, and multisolute species at the Yucca Mountain nuclear waste repository site.

Stochastic models assume that parameters are described as statistical variables derived from certain distributions and may include spatial correlation structures. The simplest approach involves the use of the Monte Carlo method, which is based on the repetitive solution of the governing equation with different sets of parameters derived from their statistical distributions, followed by a statistical analysis of results to quantify uncertainties (72). Other stochastic methods include the spectral approach (73) and the similar-media approach (74). For solute transport in unsaturated media, available approaches include the models of Dagan and Bresler (75), Simmons (76), and Jury et al. (77).

A CASE STUDY

Extensive studies of an unconfined aquifer at a Cape Cod site in Massachusetts have been completed. The hydrogeology of the site is described in Moench et al. (40), Hess et al. (78), LeBlanc et al. (79), Garabedian et al. (80), and Rudolf et al. (81). The unconfined aquifer is composed of unconsolidated glacial outwash sediments, which overlie crystalline bedrock at a depth of about 100 m. Clean, medium-to-course-grained, high permeability deposits cover fine-grained, relatively low permeability material at about 50 m below the water table.

Moench et al. (40) analyzed data for an aquifer test conducted at the Cape Cod site. After the aquifer was pumped at a rate of 0.02 m^3/s for 72 h, drawdown measurements were made in the pumped well and in 20 piezometers located at various distances from the pumped well and various depths below the land surface. The estimated hydraulic parameter values were consistent with estimates from prior studies (78). The results indicated that effects of heterogeneity at the site were small. The estimated parameter values were as follows: specific yield, 0.26; saturated thickness, 52 m; horizontal hydraulic conductivity, 11.68×10^{-4} m/s; vertical hydraulic conductivity, 7.11×10^{-4} m/s; and specific storage, 4.3×10^{-3} m^{-1}.

El-Kadi (82) used the same aquifer test data to examine the validity of the generalized Richards equation (GRE), which includes unsaturation and compressibility effects. The solution of the GRE closely matched the field-measured drawdowns, with some parameters close to their independently measured values, including conductivity and storage coefficient. However, the model was not fully validated due to the need for calibrated soil-hydraulic parameters that represent a finer soil than the coarse-grained highly permeable material characteristic of the site (83). Good accuracy was obtained using a physically based value for the storage coefficient when well-pore storage was considered. Saturated flow models did not provide results as accurate as those provided by the GRE model. Classifying the aquifer material as coarse can be misleading, considering that its effective soil properties are that of a finer-textured material. As such, unsaturated flow effects should not be overlooked.

Hess et al. (78) assessed the variability of hydraulic conductivity in the Cape Cod aquifer and used stochastic transport theories to estimate macrodispersivities. About 1500 estimates were obtained from borehole flowmeter tests and permeameter analyses of cores. The geometric mean for the flowmeter tests (0.11 cm/s) were larger than the mean for the permeameter tests (0.035 cm/s), possibly because of compaction of the cores. Geostatistical analyses estimated correlation scales that ranged from 2.9 to 8.0 m in the horizontal direction and from 0.18 to 0.38 m in the vertical direction. Estimates of asymptotic longitudinal dispersivity were in the range of 0.35–0.78 m.

McCobb and LeBlanc (84) studied a plume of contaminated groundwater that migrated from a former landfill on the Massachusetts Military Reservation to Red Brook Harbor, a small coastal embayment along the Cape Cod coast of Buzzards Bay. The primary contaminants were trichloroethene (TCE) and tetrachloroethene (PCE). The landfill was capped in December 1995, and a groundwater remediation system was established in 1999. The study identified the location of maximum concentrations of the contaminants. Fresh groundwater was present beneath the harbor at locations within 76 m of the shoreline. TCE was detected at several locations in fresh groundwater obtained from the sandy aquifer materials beneath the harbor.

LeBlanc et al. (79) analyzed data from a three-dimensional large-scale natural gradient experiment in the Cape Cod aquifer. Bromide (a conservative tracer) and lithium and molybdate (reactive tracers) injected in July 1985 were monitored for distances up to 280 m. The bromide cloud moved at a rate of 0.42 m/d, while the reactive chemicals moved at a reduced rate of about 50% of that for bromide. The dispersive characteristics of the aquifer were analyzed in the follow-up paper by Garabedian et al. (80). The asymptotic value for longitudinal dispersivity was 0.96 m, while the respective values for transverse-horizontal and transverse-vertical dispersivities were 1.5 cm and 1.55 mm, respectively.

BIBLIOGRAPHY

1. Bear, J. (1979). *Hydraulic of Groundwater*. McGraw-Hill, New York.
2. El-Beshry, M. (1999). Flow and transport in the unsaturated zone. In: *Groundwater Contamination, Transport and Remediation*, P.B. Bedient, H.S. Rifai, and C.J. Newell (Eds.), 2nd Edn. Prentice-Hall PTR, Upper Saddle River, NJ, pp. 289–334.
3. Charbeneau, R.J. (2000). *Groundwater Hydraulics and Pollutant Transport*. Prentice-Hall, Upper Saddle River, NJ.
4. Brutsaert, W. and El-Kadi, A.I. (1984). The relative importance of compressibility and partial saturation in unconfined groundwater flow. *Water Resour. Res.* **20**: 400–409.
5. Richards, L.A. (1931). Capillary conduction of liquids in porous mediums. *Physics* **1**: 318–333.
6. Broadbridge, P., and White, I. (1988). Constant rate rainfall infiltration: a versatile nonlinear model, I. Analytic solution. *Water Resour. Res.* **24**(1): 145–154.
6. van Genuchten, M.Th. (1978). *Calculating the unsaturated hydraulic conductivity with a new closed-form analytical model*. 78-WR-08, Water Resources Program, Department of Civil Engineering, Princeton University.
7. Warrick, A.W., Lomen, D.O., and Islas, A. (1990). An analytical solution to Richards' equation for a draining soil profile. *Water Resour. Res.* **26**(2): 253–258.
8. Healy, R.W. (1990). *Simulation of Solute Transport in Variably Saturated Porous Media with Supplemental Information on Modifications to the US Geological Survey's Computer Program VS2D*. Water-Resources Investigations Report 90–4025. U.S. Geological Survey.

9. Hsieh, P.A., Wingle, W., and Healy, R.W. (2000). *VS2DI—A Graphical Software Package for Simulating Fluid Flow and Solute or Energy Transport in Variably Saturated Porous Media*. Water-Resources Investigations Report 99–4130. U.S. Geological Survey.
10. Simunek, J., Sejna, M., and van Genuchten, M.Th. (1999). *The HYDRUS-2D Software Package for Simulating Two-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably Saturated Media. Version 2.0. IGWMC-TPS-53*. International Ground Water Modeling Center, Colorado School of Mines, Golden, CO.
11. Lin, H.-C.J., Richards, D.R., Talbot, C.A., Yeh, G.-T., Cheng, J.-R., Cheng, H.-P., and Jones, N.L. (2002). *A Three-Dimensional Finite Element Computer Model for Simulating Density-Dependent Flow and Transport in Variably Saturated Media. Version 3.0. GMS Documentation*. Environmental Modeling Systems, Inc., South Jordan, UT.
12. El-Kadi, A.I. and Ling, G. (1993). The Courant and Peclet criteria for numerical solution of the Richards equation. *Water Resour. Res.* **29**(10): 3485–3494.
13. McDonald, M.D. and Harbaugh, A.W. (1988). A modular three-dimensional finite-difference flow model. In: *Book 6, Techniques of Water Resources Investigations of the U.S. Geological Survey*.
14. Philip, J.R. (1969). Theory of infiltration. In: *Advances in Hydroscience*, Vol. 5. V.T. Chow (Ed.), Academic Press, New York, pp. 215–395.
15. El-Kadi, A.I. (1983). Modeling infiltration for water systems. GWMI 83-09, International Ground Water Modeling Center, Holcomb Research Institute, Butler University, Indianapolis, IN.
16. Green, W.A. and Ampt, G.A. (1911). Studies on soil physics, 1. The flow of air and water through soils. *J. Agric. Sci.*, 1–24.
17. Kostiakov, A.N. (1932). On the dynamics of the coefficient of water percolation in soils and on the necessity of studying it from a dynamic point of view for purpose of amelioration. *Trans. Com. Int. Soc. Sci. 6th Moscow A*, 17–21.
18. Horton, R.E. (1940). An approach towards a physical interpretation of infiltration capacity. *Soil. Sci. Soc. Am. Proc.* **5**: 399–417.
19. Philip, J.R. (1957). The theory of infiltration: 1. The infiltration equation and its solution. *Soil Sci.* **83**: 345–357.
20. Smith, R.E. and Parlange, J.Y. (1978). A parameter-efficient hydrologic infiltration model. *Water Resour. Res.* **14**: 533–538.
21. Parlange, J.Y., Lisle, I., Braddock, R.D., and Smith, R.E. (1982). The three-parameter infiltration equation. *Soil Sci.* **133**: 337–341.
22. Parlange, J.Y. and Haverkamp, R. (1989). Infiltration and ponding time. In: *Unsaturated Flow in Hydrologic Modeling: Theory and Practice*. H.J. Morel-Seytoux (Ed.), *NATO ASI Series C: Mathematical and Physical Sciences*, Vol. 275.
23. Theis, C.V. (1935). The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Am. Geophys. Union Trans.* **16**: 519–524.
24. Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
25. Kruseman, G.P. and de Ridder, N.A. (1979). *Analysis and Evaluation of Pumping Test Data*, Bull. 11. International Institute for Land Reclamation and Improvements, Wageningen, The Netherlands.
26. Kruseman, G.P. and de Ridder, N.A. (1990). *Analysis and Evaluation of Pumping Test Data*, 2nd Edn. (completely revised), *ILRI Publication* 47. International Institute for Land Reclamation and Improvements, Wageningen, The Netherlands.
27. Driscoll, F.G. (1987). *Groundwater and Wells*. Johnson Division, St. Paul, MN.
28. Dominico, P.A. and Schwartz, F.W. (1990). *Physical and Chemical Hydrogeology*. John Wiley & Sons, Hoboken, NJ.
29. Walton, W.C. (1996). *Aquifer Test Analysis with WINDOWS Software*, CRC Press, Boca Raton, FL.
30. Fetter, C.W. (2001). *Applied Hydrogeology*, 4th Edn. Prentice-Hall, Upper Saddle River, NJ.
31. Boulton, N.S. (1954). *Unsteady Radial Flow to a Pumped Well Allowing for Delayed Yield From Storage*, Publication 37. International Association of Scientific Hydrology, Rome, pp. 472–477.
32. Boulton, N.S. and Streltsova, T.D. (1975). New equations for determining the formation constants of an aquifer from pumping test data. *Water Resour. Res.* **11**: 148–153.
33. Boulton, N.S., and Pontin, J.M.A. (1971). An extended theory of delayed yield from storage applied to pumping tests in unconfined anisotropic aquifers. *J. Hydrol.* **14**: 53–65.
34. Dagan, G. (1967). A method of determining the permeability and effective porosity of unconfined anisotropic aquifers. *Water Resour. Res.* **3**: 1059–1071.
35. Neuman, S.P. (1972). Theory of flow in unconfined aquifers considering delayed response of the water table. *Water Resour. Res.* **8**: 1031–1044.
36. Neuman, S.P. (1974). Effects of partial penetration on flow in unconfined aquifers considering delayed aquifer response. *Water Resour. Res.* **10**: 303–312.
37. Neuman, S.P. (1975). Analysis of pumping test data from anisotropic unconfined aquifers considering delayed yield. *Water Resour. Res.* **11**: 329–342.
38. Moench, A.F. (1995). Combining the Neuman and Boulton models for flow to a well in an unconfined aquifer. *Ground Water* **33**: 378–384.
39. Moench, A.F. (1997). Flow to a well of finite diameter in a homogeneous, anisotropic water table aquifer. *Water Resour. Res.* **33**: 1397–1407.
40. Moench, A.F., Garabedian, S.P., and LeBlanc, D.R. (2001). Estimation of hydraulic parameters from an unconfined aquifer test conducted in a glacial outwash deposit, Cape Cod, Massachusetts. Professional Paper 1629, U.S. Geological Survey.
41. Boulton, N.S. (1963). Analysis of data from non-equilibrium pumping tests allowing for delayed yield from storage. *Proc. Inst. Civ. Eng.* **26**: 469–482.
42. Waterloo Hydrogeologic, Inc. (2000). *AquiferTest User's manual*. Waterloo, Ontario, Canada.
43. Moench, A.F. (1993). Computation of type curves for flow to partially penetrating wells in water-table aquifers. *Ground Water* **31**: 966–971.
44. Cooper, H.H. and Jacob, C.E. (1946). A generalized graphical method for evaluating formation constants and summarizing well field history. *Am. Geophys. Union Trans.* **27**: 526–534.
45. Hughes, J.B. (1999). Contaminant fate processes. In: *Groundwater Contamination, Transport and Remediation*, P.B. Bedient, H.S. Rifai, and C.J. Newell (Eds.), 2nd Edn. Prentice-Hall PTR, Upper Saddle River, NJ, pp. 203–236.
46. Fetter, C.W. (1993). *Contaminant Hydrology*. Macmillan Publishing, New York.
47. Bedient, P.B., Rifai, H.S. and Newell, C.J. (1999). *Groundwater Contamination, Transport and Remediation*, 2nd Edn. Prentice-Hall PTR, Upper Saddle River, NJ.

48. Todd, D.K. (1980). *Groundwater Hydrology*, 2nd Edn. John Wiley & Sons, Hoboken, NJ.
49. Cooper, H.H. Jr. (1959). A hypothesis concerning the dynamic balance of fresh water and salt water in coastal aquifer. *J. Geophys. Res.* **64**: 461–467.
50. Drabbe, J. and Baden-Ghyben, W. (1888–1889). *Nota in verband met de voorgenomen putboring nabij Amsterdam*. Tydyschrift van het Koninkhyk Institute van Ingenieurs, The Hague, Netherlands, pp. 8–22.
51. Herzberg, B. (1901). Die Wasserversorgung einiger Nordsubader, Munch. *J. Gasbelenchtung Wasser.* **44**: 815.
52. Essaid, H.I. (1990). *The Computer Model SHARP, a Quasi-Three-Dimensional Finite-Difference Model to Simulate Freshwater and Saltwater Flow in Layered Coastal Aquifer Systems*. Water-Resources Investigations Report 90–4130. U.S. Geological Survey.
53. Jury, W.A., Russo, D., Sposito, G., and Elabd, H. (1987). The spatial variability of water and solute transport properties in unsaturated soil: I. Analysis of property variation and spatial structure with statistical models. *Hilgardia* **55**: 1–32.
54. Jury, W.A., Russo, D., Sposito, G., and Elabd, H. (1987). The spatial variability of water and solute transport properties in unsaturated soil: II. Scaling models of water transport. *Hilgardia* **55**: 33–56.
55. Hutson, J.L. and Wagenet, R.J. (1995). A multiregion model describing water flow and solute transport in heterogeneous soils. *Soil Sci. Soc. Am. J.* **59**: 743–751.
56. Flurry, M. (1996). Experimental evidence of transport of pesticides through field soils—a review. *J. Environ. Qual.* **25**: 25–45.
57. Ray C., Vogel, T., and Dusek, J. (2004). Modeling depth-variant and domain-specific sorption and biodegradation in dual-permeability media. *J. Contam. Hydrol.* **70**: 63–87.
58. Doherty, J. (1994). *PEST—Model-Independent Parameter Estimation*, Watermark Computing, Corinda, Australia.
59. El-Kadi, A.I. and Torikai, J.D. (2001). Identifying variable saturated water-flow patterns in a steep hillslope under intermittent heavy rainfall. *Hydrogeology* **9**: 231–242.
60. Durner, W. (1994). Hydraulic conductivity estimation for soils with heterogeneous pore structure. *Water Resour. Res.* **30**(2): 211–224.
61. Mohanty, B.P., Bowman, R.S., Hendrickx, J.M.H., and van Genuchten, M.T. (1997). New piecewise-continuous hydraulic functions for modeling preferential flow in an intermittent-flood-irrigated field. *Water Resour. Res.* **33**: 2049–2064.
62. Harrison, B., Sudicky, E.A., and Cherry, J.A. (1992). Numerical analysis of solute migration through fractured deposits into underlying aquifers. *Water Resour. Res.* **28**: 515–526.
63. Charbeneau, R.J. (1984). Kinematic models for soil moisture and solute transport. *Water Resour. Res.* **20**: 699–706.
64. Germann, P.F. (1985). Kinematic wave approach to infiltration and drainage into and from soil macropores. *Trans. ASAE* **28**: 745–749.
65. Duguid, J.O. and Lee, P.C.Y. (1978). Flow in fractured porous media. *Water Resour. Res.* **13**: 558–566.
66. Gerke, H.H. and van Genuchten, M.T. (1993). A dual-porosity model for simulating the preferential movement of water and solutes in structured porous media. *Water Resour. Res.* **29**: 305–319.
67. Ray, C., Ellsworth, T.R., Valocchi, A.J., and Boast, C.W. (1997). An improved dual porosity model for chemical transport in macroporous soils. *J. Hydrol.* **193**: 270–292.
68. Vogel, T., Gerke, H.H., Zhang, R., and van Genuchten, M.Th. (2000). Modeling flow and transport in a two-dimensional dual-permeability system with spatially variable hydraulic properties. *J. Hydrol.* **238**: 78–89.
69. Gwo, J.P., Jardine, P.M., Wilson, G.V., and Yeh, G.T. (1995). A multi-pore concept to modeling mass transfer in subsurface media. *J. Hydrol.* **164**: 217–237.
70. Zyvoloski, G.A., Dash, G.V., and Kelkar, S. (1988). *FEHM: Finite Element Heat and Mass Transfer Code*. Report No. LA-11224-MS. Los Alamos National Laboratory, Los Alamos, NM.
71. Zyvoloski, G.A., Dash, G.V., and Kelkar, S. (1992). *FEHMN 1.0: Finite Element Heat and Mass Transfer Code*. Report No. LA-12062-MS, Rev. 1. Los Alamos National Laboratory, Los Alamos, NM.
72. El-Kadi, A.I. and Brutsaert, W. (1985). Applicability of effective parameters for unsteady flow in heterogeneous aquifers. *Water Resour. Res.* **21**: 183–198.
73. Bakr, A.A., Gelhar, L.W., Gutjahr, A.L., and McMillan, J.R. (1978). Stochastic analysis of spatial variability in subsurface flow, 1. Comparison of one- and three-dimensional flows. *Water Resour. Res.* **14**: 263–271.
74. Miller, E.E. and Miller, R.D. (1956). Physical theory for capillary flow phenomena. *J. Appl. Phys.* **27**: 354–362.
75. Dagan, G. and Bresler, E. (1979). Solute dispersion in unsaturated heterogeneous soil at field scale: I. Theory. *Soil Sci. Soc. Am. J.* **43**: 461–467.
76. Simmons, C.S. (1982). A stochastic-convective transport representation in one-dimensional porous media systems. *Water Resour. Res.* **18**: 1193–1214.
77. Jury, W.A., Stolzy, L.H., and Shouse, P. (1982). A field test of the transfer function model for predicting solute transport. *Water Resour. Res.* **19**: 369–375.
78. Hess, K.M., Wolf, S.H., and Celia, M.A. (1992). Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 3. Hydraulic conductivity variability and calculated macrodispersivities. *Water Resour. Res.* **28**: 2011–2027.
79. LeBlanc, D.R. et al. (1991). Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 1. Experimental design and observed tracer movement. *Water Resour. Res.* **27**: 895–910.
80. Garabedian, S.P., LeBlanc, D.R., Gelhar, L.W., and Celia, M.A. (1991). Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 2. Analysis of spatial moments for a nonreactive tracer. *Water Resour. Res.* **27**: 911–924.
81. Rudolph, D.L.R. et al. (1996). Infiltration and solute transport experiments in unsaturated sand and gravel, Cape Cod, Massachusetts: experimental design and overview of results. *Water Resour. Res.* **32**: 519–532.
82. El-Kadi, A.I. (2005). Validity of the generalized Richards equation for the analysis of three-dimensional test data for a coarse-material aquifer. *Vadose Zone J.*, in press.
83. Mace, A., Rudolph, D.L.R., and Kachanoski, R.G. (1998). Suitability of parametric models to describe the hydraulic properties of an unsaturated coarse sand and gravel. *Ground Water* **36**: 465–475.
84. McCobb, T.D. and LeBlanc, D.R. (2002). Detection of fresh ground water and a contaminant plume beneath Red Brook Harbor, Cape Cod, Massachusetts, 2000. Water-Resources Investigations Report 02–4166, U.S. Geological Survey.

MODELING OF DNAPL MIGRATION IN SATURATED POROUS MEDIA

M.S. Mohan Kumar
Indian Institute of Science
Bangalore, India

Mini Mathew
Colorado School of Mines
Golden, Colorado

INTRODUCTION

Nonaqueous phase liquids (NAPLs) such as the homogenate hydrocarbons used in dry cleaning and industrial degreasing (DNAPLs), hydrocarbon fuels and aromatic solvents (LNAPLs), and neutrally buoyant coal tars and creosotes are widespread in our environment. Because of their low solubility, NAPL sources can emit toxic contaminants to groundwater for many decades if not managed properly. The migration of dense nonaqueous phase liquids (DNAPLs), which are immiscible with water, through the saturated porous media is an important part of contaminant hydrology and in petroleum engineering. Protection and remediation of groundwater resources require an understanding of processes that affect the fate and transport of such contaminants in the subsurface environment. To implement appropriate remedial schemes in the contaminated area, it is necessary to evaluate the extent of the contaminated area. This information can be obtained by extensive field investigation, which generally is expensive and time consuming. Field investigation can be reduced or made more cost effective, if the migration pattern of the DNAPLs can be evaluated by using numerical models accurately.

The movement of DNAPLs through an initially saturated porous medium is represented as the two-phase flow of water and DNAPLs simultaneously in the porous medium (1). One of the fluids may wet the porous medium more than the other. A large number of experiments for multiphase processes have been carried out (2–4) and these have shown that Darcy's velocity for each phase in the porous medium can be described by the generalized Darcy's law. Darcy's law is extended to multiphase flow by postulating that the gradient in phase pressure of each phase is causing that fluid phase to flow in the porous medium and the effective premeability of each phase is a function of saturation of that respective phase. The continuity equations for immiscible two-phase flow of DNAPLs and water in the porous medium is given by (5,6):

$$\frac{\partial}{\partial x_i}\left[\frac{\rho_\alpha k_{ij} k_{r\alpha}(S_\alpha)}{\mu_\alpha}\left(\frac{\partial P_{ij\alpha}(S_\alpha)}{\partial x_j}+\rho_\alpha g\right)\right] \pm q_{ij\alpha} = \frac{\partial}{\partial t}(\phi \rho_\alpha S_{ij\alpha}), \alpha = \{w, nw\} \tag{1}$$

where x_i and x_j are the Cartesian coordinates in the x and z directions, t is the time, i and j are the direction indices, μ [$\mathrm{ML}^{-1}\mathrm{T}^{-1}$] is the viscosity, ρ [ML^{-3}] is the density, k is intrinsic permeability, k_α [LT^{-1}] is the relative permeability of the phase, and z [L] is the vertical distance taken as positive upward. The α represents both DNAPL and water phases, S is the phase saturation, φ is the medium porosity, q_α [LT^{-1}] is the phase source or sink term, g [LT^{-2}] is the acceleration due to gravity, and P_α is the phase pressure.

When two immiscible fluids are in contact, a discontinuity in phase pressures exists at the interface separating the fluids by the consequence of the interfacial tension, which exists between two phases in contact. The magnitude of the pressure difference depends on the interface curvature at that point, which in turn depends on the saturation of phases. The phase pressures are coupled through the capillary pressure P_C and are a function of saturation:

$$P_C(S_\alpha) = P_{nw} - P_w, \quad \alpha = \{w, nw\} \tag{2}$$

The governing equations have strong nonlinearities such as the dependence of the relative permeabilities $k_{r\alpha}$ and P_C on the respective phase saturations. Parameterization of capillary pressure saturation and the relative permeability saturation relations widely used for the two phases are given by Brooks and Corey (7) and van Genuchten (8).

STATUS OF LITERATURE

The governing equations of multiphase flow are highly nonlinear. Numerical methods of multiphase flow mainly depend on the formulation of the governing equations used for the numerical modeling. The numerical and analytical approaches of the multiphase flow are difficult due to its high nonlinearity and the coupling between the phases. Buckley and Leverett (9) were the first to analyze the multiphase fluid displacement in one dimension by a simplified mathematical model by suppressing the capillary derivative term. Fokas and Yortsos (10) developed a closed form analytical solution for one-dimensional flow with capillary drive by restricting the capillary hydraulic functions. McWhorter and Sunada (11) developed an exact quasianalytical solution using the concept of fractional flow functions known as the flux concentration relations.

A number of numerical models have been developed to simulate the migration of dense nonaqueous phase liquids (DNAPLs) in groundwater. All the models are based on one or another type of primary variables. The multiphase flow equations are highly nonlinear and it requires powerful techniques to handle the coupled equations. The choice of method for solution of multiphase flow is very important. Peaceman (12) used the explicit method and Peaceman and Rachford (13) used the alternating direction implicit (ADI) method. Results showed that the ADI method is unstable for multiphase simulations especially in heterogeneous porous media.

The simultaneous solution (SS) method was first proposed by Douglas and Rachford (14) and later extended and further analyzed by Coats et al. (15), Coats (16), and Sheffield (17) for the petroleum simulation problems. Explicit and alternating direction implicit methods are not stable for two-phase immiscible fluid problems; hence, a simultaneous implicit method was introduced (13) and was the first implicit method to be used successfully for problems in two dimensions. The first convergence analysis of the SS method was presented for two-phase flow by Douglas (18) and later by Coats (16), who analyzed the stability of both the SS and IMPES methods. The implicit pressure explicit saturation (IMPES) method was proposed by Sheldon et al. (19) and Stone and Gardner (20). The basic idea behind this method was to obtain a single pressure equation by a combination of the flow equations. Peaceman and Rachford (21) proposed the leapfrog sequential method, which gives satisfactory results in a limited number of calculations and the truncation error is larger than the simultaneous method. This method leads to an anomalous situation where one of the relative permeabilities becomes zero. Fagin and Stewart (22) developed another type of sequential method and this method is described in detail by Coats et al. (23), in which pressure of one of the phases was taken as one primary variable and was calculated implicitly and the saturation was calculated either explicitly or implicitly using the present value of pressure. This method was also subjected to saturation creep where the relative permeability of one of the phases becomes zero. Aziz and Settari (1) developed a sequential method consisting of two steps. The first step was to obtain an implicit pressure for saturation simulation. In this method, the equations of two phases were combined into one, like in the IMPES method, to obtain a single pressure equation in one phase pressure and the saturation was calculated implicitly using new phase pressures.

Kueper and Frind (5,6) carried out experiments and developed a two-dimensional finite difference model to study the simultaneous movement of DNAPLs and water in heterogeneous porous media. Numerical simulation of DNAPLs infiltration processes in the saturated heterogeneous porous media was studied by means of numerical simulations using the finite element method (24). The two-phase simulator MUFTE developed by Helmig (25) was based on a semidiscrete finite element formulation using a modified Petrov–Galerkin technique, and the nonlinearization was carried out using the Newton–Raphson method. Difference in parameterization of the constitutive relations was studied by Croise et al. (24) for homogeneous media, and the results show that the constitutive relations have significant impact on DNAPLs migration in homogeneous media. Also, the study is carried out for DNAPLs infiltration into heterogeneous media and it shows that the infiltration of DNAPLs is highly dependent on the capillary resistance at the interface between the media.

A comparative study was made between the numerical methods such as fully upwind Galerkin method, which belongs to the classical finite element method, and a mixed hybrid finite element method based on an implicit pressure explicit saturation (IMPES) approach (26). The results show that the control volume finite element approach based on coupled pressure and a saturation scheme is mass conserved locally over each control volume. Also, they observed that the IMPES method was not suited to simulate processes where high changes of capillary pressure with respect to saturation occur.

Coupling of different numerical algorithms for two-phase flow in saturated porous media was carried out using both fully implicit and sequential formulations (27). The influence of field scale heterogeneity on the infiltration and entrapment of dense nonaqueous phase liquids in saturated formulation was studied by Timothy et al. (28). Three node centered finite volume discretizations for multiphase flow in porous media have been developed by Huber and Helmig (26) to study the extent of heterogeneity at the interfaces and their associated interface conditions. Finite volume methods such as the Box method (BOX), control volume finite element method (CVFE), IFDM method, and different combinations of these methods such as BOXIFDM and CVBOXIFDM have been compared. The results show that the CVFE has difficulty generating meshes, which obey the positive transmissibility constraint. The BOX, IFDM, BOXIFDM, and CVBOXIFDM do not have any restriction on the user grid. Also, the results show that the BOX discretization may generate spurious oscillations at the corners of heterogeneities if the Brooks and Corey model is employed. Combinations of all three methods are applicable for any mesh without generating a physical saturation profile at the heterogeneities. The study of effective parameters for two-phase flow in a porous medium with periodic heterogeneities to investigate the feasibility of replacing a porous medium containing heterogeneities with an equivalent homogeneous medium was done by Ashtiani et al. (29). Selecting the proper primary variables for the efficient modeling of the highly nonlinear problems of multiphase flow in a heterogeneous porous fractured medium is carried out by Yu-Shu Wu et al. (30).

ADAPTIVE MODIFIED SEQUENTIAL METHOD

Thomas and Thurnan (31) and Forsyth (32) have described an adaptive implicit method. This involve solving implicit formulations of the governing equations in part of the domain where changes in the primary variables were large, while in the rest of the domain, which is computationally less intensive, the less robust IMPES method was employed. In most practical multiphase simulation problems, large changes in fluid pressure and saturations do not occur throughout the spatial domain at a given time step. Computational effort is therefore inefficiently spent on solving equations in areas where less activity occurs. However, the locations of active zones change with time. A study of an adaptive solution using a fully implicit method is incorporated for the two-phase simulation in the saturated porous medium by Unger et al. (33). They observed that the adaptive solution using a simultaneous method requires extensive checks for the applicability of this procedure at each node and this leads to a negligible computational benefit.

In this work, an adaptive solution using a modified sequential method is used for the modeling. The linearization of the governing equations is carried out using a modified Newton–Raphson method. The modified sequential method developed by Mini Mathew for the simulation of DNAPLs–water in the saturated zone is used in the adaptive solution. The modified sequential method consists of two steps. In the first step, the linearized equation of wetting fluid is used for the simulation of wetting fluid pressure by taking the primary variable of saturation of wetting fluid at the previous iteration on the right-hand side. The Jacobian matrix of wetting fluid pressure is a five-banded matrix and the equation is given by

$$[A_1][P_W]^{n+1/2} = [R_1]^n \tag{3}$$

where $[A_1]$ is an $N \times N$ (N is the number of nodes) Jacobian matrix, $[P_W]$ is a column vector of N unknowns of P_W, and R_1 is the column vector consisting of Eq. 4:

$$\begin{aligned} R_1 = -[[BS_{Wi(j-1)}] + [D''S_{W(i-1)j}] + [E''S_{Wij}] \\ + [F''S_{W(i+1)j}] + [H''S_{Wi(j+1)}]] \end{aligned} \tag{4}$$

where B'', D'', E'', F'', and H'' are the partial derivatives of the governing equation of wetting fluid with respect to the wetting fluid saturation. At the first time step for the wetting fluid pressure simulation, the simulation was carried out for the full domain. The active nodes were identified for the simulation of wetting fluid saturation. The nodes were categorized as either active or inactive in each iteration. For the simulation of wetting fluid pressure, the active nodes were selected in such a way that the difference in saturation of water (wetting fluid) between the present and the previous iteration at that particular node or any one of its surrounding four nodes is less than the specified tolerance level of water saturation. Other nodes, except the active nodes, are taken as inactive nodes for the pressure calculation. On inactive nodes the pressure of wetting fluid is taken from the previous iteration.

In the second step, the wetting fluid saturation is calculated using the nonwetting phase equation. The equation in matrix form is given by

$$[A_2][S_W] = [R_2]^{n+1/2} \tag{5}$$

where $[A_2]$ is an $N \times N$ (N is the number of nodes) Jacobian matrix, $[S_W]$ is a column vector of N unknowns of S_W, and R_2 is the column vector consisting of Equation 6:

$$\begin{aligned} \mathrm{R}_2 = -[[B''''P_{W_i(j-1)}] + [D''''P_{W(i-1)j}] + [E''''P_{Wij}] \\ + [F''''P_{W(i+1)j}] + [H''''P_{Wij+1}]] \end{aligned} \tag{6}$$

where B'''', D'''', E'''', F'''', and H'''' are the partial derivatives of the governing equation of nonwetting fluid with respect to the wetting fluid pressure. For the simulation of wetting fluid saturation, the active nodes were selected in such a way that the difference in water pressure between the present and the previous iteration at that particular node or at any one of its surrounding four nodes is less than the specified tolerance level of water pressure. The computation for the saturation was carried out only for the active nodes and the saturation of inactive nodes was taken from the previous iteration. This procedure was continued in each iteration until the simulation ends.

DNAPLS MIGRATION IN

In this section, an effort is made to study the effect of DNAPLs migration in randomly heterogeneous porous media. The simulation is carried out using ASMSEM model to study the handling power of the numerical method for the DNAPLs simulation on the randomly heterogeneous porous media. The properties of the fluid and medium and the domain properties are: wetting fluid density, 1000 kg·m^{-3}; nonwetting fluid density, 1460 kg·m^{-3};

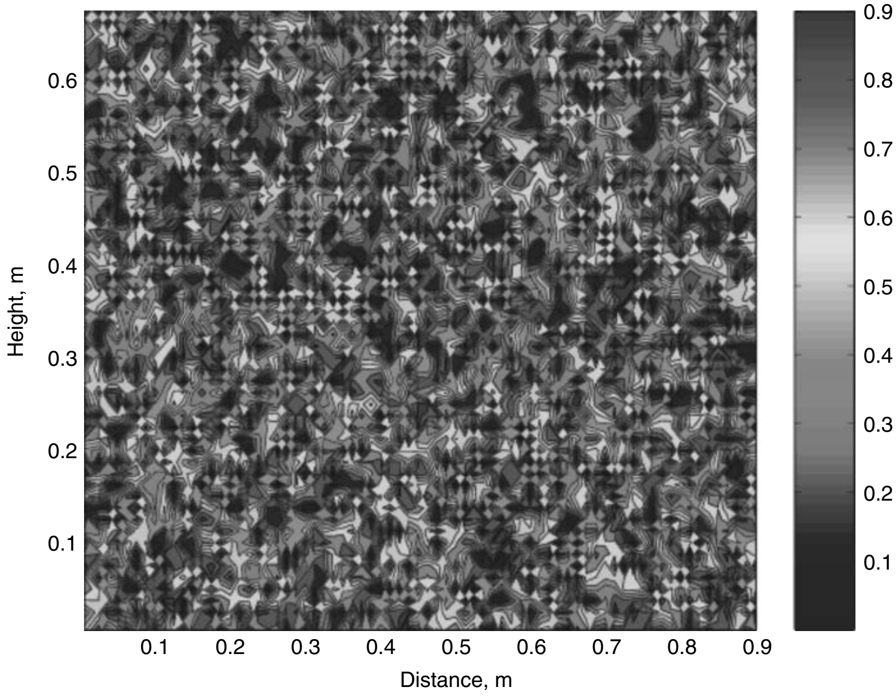

Figure 1. Random permeability field.

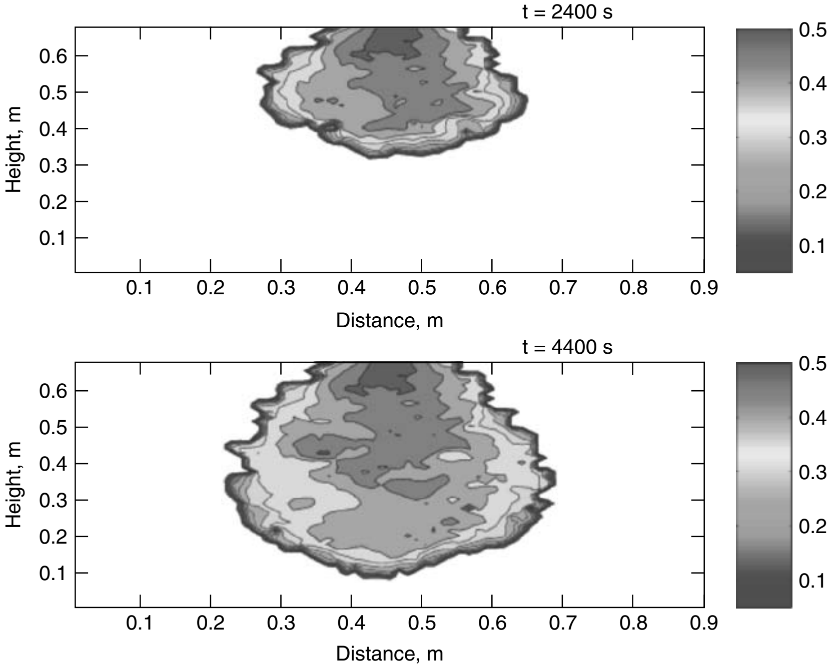

Figure 2. DNAPL migration profiles.

wetting fluid viscosity, 0.001 Pa·s; nonwetting fluid viscosity, of 0.0009 Pa·s; intrinsic permeability, 6.64E-11 m^2; porosity, 0.4; wetting phase residual saturation, 0.09; displacement pressure, 755 Pa; and pore size distribution index, 2.70. In the random porous media, the capillary pressure curve depends on the local value of the absolute permeability and is given by

$$P_C(S_W) = P_d[k^*/k]^{1/2}{S_{ew}}^{-1/\lambda}$$

where k^* is the mean value of the permeability field and the parameters P_d and λ are the Brooks and Corey parameters. The main aim of this study is to verify the handling power of the model using the adaptive modified sequential method for the DNAPLs simulation on randomly heterogeneous porous media. Figure 1 shows the permeability field, which is generated for the present study. The permeability field was generated in such a way that the mean value of the permeability field is equal to the intrinsic permeability of the medium. Figure 2 shows the DNAPLs migration profile after 2400 s and 4400 s.

CONCLUSION

The main mechanisms of DNAPLs migration in porous media have been discussed. Several works in this area, especially about modeling, have been discussed in detail. The main approaches for modeling of DNAPLs migration is presented. The advantage of the adaptive modified sequential method has been pointed out. The developed model is applied on a randomly heterogeneous porous medium to establish DNAPLs migration patterns. Such simulation models will be very useful in planning field experiments for prediction and remediation due to DNAPLs contamination.

BIBLIOGRAPHY

1. Aziz, K. and Settari, A. (1979). *Petroleum Reservoir Simulation*. Applied Science, London.
2. Wyckoff, R.D. and Botset, H.G. (1936). The flow of gas–liquid mixtures through unconsolidated sand. *Physics* **7**: 325–345.
3. Leverett, M.C. (1938). Flow of oil–water mixtures through unconsolidated sands. *Trans. Am. Inst. Min. Metall. Pet. Eng.* **132**: 149–171.
4. Scheidegger, A.E. (1974). *Physics of Flow Through Porous Media*. 3rd Edn. University of Toronto Press, Toronto, Canada.
5. Kueper, B.H. and Frind, E.O. (1991). Two phase flow in heterogeneous porous media 1. Model development. *Water Resour. Res.* **27**(6): 1049–1057.
6. Kueper, B.H. and Frind, E.O. (1991). Two phase flow in heterogeneous porous media 2. Model application. *Water Resour. Res.* **27**(6): 1059–1070.
7. Brooks, R.H. and Corey, A.T. (1964). *Hydraulics Properties of Porous Media*. Hydrol., Pap. 3. Civil Engineering Department, Colorado State University, Fort Collins, CO.
8. van Genuchten, M.T. (1980). A closed form solution for predicting the hydraulic conductivity in soils. *Soil Sci. Am. J.* **44**: 892–898.
9. Buckley, S.E. and Leverett, M.C. (1942). Mechanism of fluid displacement in sands. *Trans. Am. Inst. Min. Metall. Pet. Eng.* **146**: 107–116.
10. Fokas, A.S. and Yortsos, Y.C. (1982). On the exactly solvable equation \$$S_t = [(\beta S + \gamma)^{\{-2\}}S_x]_x + \alpha(\beta S + \gamma)^{\{-2\}}S_x$\$ occurring in two phase flow in porous media. *SIAM J. Appl. Math.* **42**(2): 318–332.
11. McWhorter, D.B. and Sunada, D.K. (1990). Exact integral solutions for two phase flow. *Water Resour. Res.* **26**(3): 399–414.
12. Peaceman, D.W. (1967). *Numerical Solution of the Nonlinear Equations for Two Phase Flow Through Porous Media, Nonlinear Partial Differential Equations—Methods of Solutions*. Academic Press, New York, pp. 79–91.

13. Peaceman, D.W. and Rachford, H.H. (1962). The numerical calculation of multidimensional miscible displacement. *Trans. SPE AIME* **225**: 327–329.
14. Douglas, J., Jr. and Rachford, H.H., Jr. (1959). On the numerical solution of heat conduction problems in two and three space variables. *Am. Math. Soc. Trans.* **82**(2): 421–439.
15. Coats, K.H., Nielsen, R.L., Terhune, M.H., and Weber, A.G. (1967). Simulation of three dimensional two phase flow in oil and gas reservoir. *Trans. SPE AIME* **240**: 377–388.
16. Coats, K.H. (1968). *Elements of Reservoir Simulation*. Lecture notes, University of Texas, reprinted by Intercomp Resources Development and Engineering, Inc., Houston, TX.
17. Sheffield, M. (1969). Three phase fluid flow including gravitational, viscous and capillary forces. *Soc. Pet. Eng. J.* **9**: 232–246.
18. Douglas, J., Jr. (1960). A numerical method for the solution of a parabolic system. *Numer. Math.* **2**: 91–98.
19. Sheldon, J.W., Zondek, B., and Cardwell, W.T. (1959). One dimensional incompressible non capillary two phase fluid flow in porous medium. *Trans. SPE AIME* **216**: 290–296.
20. Stone, H.L. and Gardner, A.O., Jr. (1961). Analysis of gas cap or dissolved gas reservoirs. *Trans. SPE AIME* **222**: 92–104.
21. Peaceman, D.W. and Rachford, H.H. (1955). The numerical solution of parabolic and elliptic differential equations. *SIAM J. Appl. Math.* **3**: 28–41.
22. Fagin, R.G. and Stewart, C.H. (1966). A new approach to the two dimensional multiphase reservoir simulator. *Soc. Pet. Eng. J.* **6**: 175–182.
23. Coats, K.H., Dempsey, J.R., and Henderson, J.H. (1970). A new technique for determining reservoir description from field performance data. *Trans. SPE AIME* **249**: 66–74.
24. Croise, J., Helmig, R., and Sheta, H. (1995). Numerical simulation of DNAPL infiltration processes in saturated heterogeneous porous media. In: *Proceedings of GQ Conference*, May 15–18.
25. Helmig, R. (1995). *Theory and Numerical Methods for Multiphase Flow in Porous Medium*. Technical report, Institute for Wasserbau, University of Stuttgart, Stuttgart, Germany.
26. Huber, R. and Helmig, R. (1999). Multiphase flow in heterogeneous porous media: a classical finite element method versus an implicit pressure explicit saturation based mixed finite element–finite volume approach. *Int. J. Numer. Methods Fluids* **29**: 899–920.
27. Peszynska, M., Lu, Q., and Wheeler, M.F. (1999). Coupling different numerical algorithms for two phase fluid flow. In: *The Mathematics of Finite Elements and Applications X*. J.R. Whiteman (Ed.). Brunel University, Uxbridge, Middlesex, UK, pp. 205–214.
28. Dekker, T.J. and Abriola, L.A. (2000). The influence of field scale heterogeneity on the infiltration and entrapment of dense nonaqueous phase liquids in saturated formulations. *Journal of Contaminant Hydrology* **42**(2): 187–218.
29. Ataie–Ashtiani, B. et al. (2001). Effective parameters for two phase flow in a porous medium with periodic heterogeneities. *J. Contam. Hydrol.* **49**: 87–109.
30. Yu-Shu, W., Zhang, K., and Pruess, K. (2002). Massively parallel simulation of flow and transport in variably saturated porous and fractured media. In: *Proceedings of the XIVth International Conference on Computational Methods in Water Resources* (CMWR XIV), June 23–28, Delft, The Netherlands, Vol. 1, pp. 289–297.
31. Thomas, G.W. and Thurnan, D.H. (1983). Reservoir simulation using an adaptive implicit method. *Soc. Pet. Eng. J.* **23**: 750–768.
32. Forsyth, P.A. (1989). Adaptive implicit criteria for two phase flow with gravity and capillary pressure. *SIAM J. Sci. Static Comput.* **10**: 227–252.
33. Unger, A.J.A, Forsyth, P.A., and Sudicky, E.A. (1998). Influence of alternative dissolution methods and subsurface heterogeneity on DNAPL disappearance times. *J. Contam. Hydrol.* **30**: 217–242.

THE USE OF SEMIPERMEABLE MEMBRANE DEVICES (SPMDs) FOR MONITORING, EXPOSURE, AND TOXICITY ASSESSMENT

A. Katsoyiannis
C. Samara
Aristotle University of Thessaloniki
Thessaloniki, Greece

INTRODUCTION

Semipermeable Membrane Devices (SPMDs) constitute an innovative approach for sequestering trace/ultra-trace levels of waterborne hydrophobic chemicals and for accurate exposure estimates (1,2). Bioavailable hydrophobic organic contaminants usually occur at trace (<1 ppb) or ultra-trace (<1 ppt) concentration levels in water. The traditional active sampling methods such as liquid-liquid extraction (LLE) or solid phase extraction (SPE) are only partially successful in sequestering these pollutants from water because of the very large water volumes required for processing/analysis. The concentrations of aquatic contaminants can change daily or even hourly, therefore, a great number of samples are required to get a true picture of contaminants present. The limitations in the direct determination of such low concentrations in water have often prompted the use of biomonitoring organisms for exposure estimates. However, the organism-based approach also has inherent problems, including biotransformation and depuration of contaminants, and inapplicability in many exposure situations because of the effects of stress on the biomonitoring organisms that often lead to a lack of proportionality between biomonitoring organism tissue concentrations and ambient exposure concentrations. The SPMDs are lipid-containing polymeric membranes, which function as passive *in situ* samplers of hydrophobic organic pollutants and as physical models of the bioconcentration process (Fig. 1).

Triolein (glyceryl trioleate, $C_{57}H_{104}O_6$) containing SPMDs made of low-density polyethylene (LDPE) were first developed by scientists of the U.S. Geological Survey's Columbia Environmental Research Center (1,3), although similar membranes containing *n*-hexane were previously used by Soedergren (4,5) for simulation of the uptake of pollutants by aquatic organisms. In general, SPMDs can be advantageously used in the following broad areas of application (1):

1. Determination of pollution point sources
2. Estimation of time-weighted average dissolved or vapor phase chemical concentrations

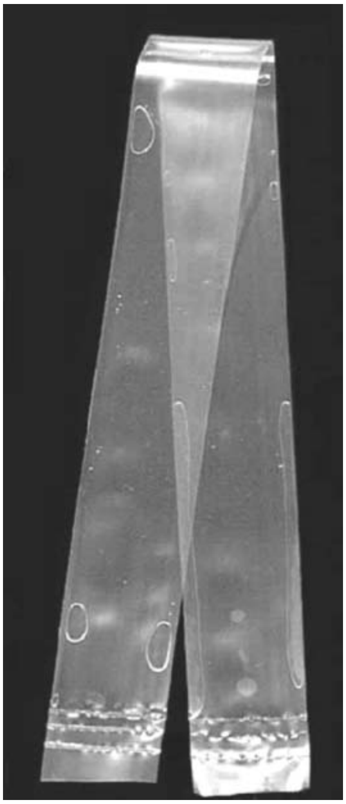

Figure 1. Photo of a semipermeable membrane device (SPMD).

3. *In situ* mimetic concentration of bioavailable chemicals for bioindicator or immunoassay tests
4. Contaminant sequestration in toxicity identification and evaluation (TIE) procedures
5. Estimation of organisms' exposure and bioconcentration potential
6. Determination of the readily bioavailable fraction (dissolved or vapor phase) of a chemical in an environmental compartment for predicting transport, fate, and residue toxicity

SPMD-BIOMEMBRANE SIMILARITIES

The organism-based bioconcentration studies are associated with several problems, such as biotransformation and depuration of contaminants, stress effects caused by noncontaminant stressors or the target contaminants, etc. The SPMD-based approach may offer significant advantages, including: (a) SPMDs do not need to be fed, (b) they do not eat each other or die from disease, lack of oxygen, or from the target contaminants, (c) an SPMD does not move around, will not escape, reproduce, and create an introduced species problem, (d) an SPMD can be used in almost any aquatic environment (cold mountain water, warm desert water, sea water, groundwater), (e) fish can actively depurate some kinds of hydrophobic contaminants, SPMDs don't, and (f) SPMDs can be used to back-calculate the concentration of a waterborne contaminant, but that usually cannot be done with fish.

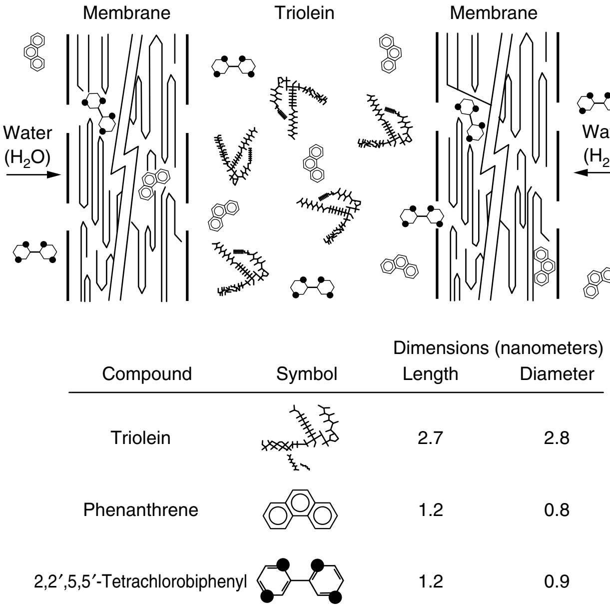

Figure 2. Schematic cross section of an SPMD in an aqueous environment (7).

SPMDs are designed to mimic the parts of animals that cause bioconcentration. The special plastic of SPMDs (LDPE) allows contaminants to pass through to the lipid while excluding water-like membranes of animal cells. The lipid inside (triolein) is similar to a highly purified fish fat. Although the composition of polymeric membranes differ substantially from that of gill membranes (i.e., bilayers of phospholipids and proteins), the diffusion of many lipophilic organic compounds through so-called "nonporous" polymeric membranes appears to simulate diffusion through biomembranes (6). Low-density polyethylene and gill membranes appear to exhibit similar steric exclusion limits with respect to the uptake of hydrophobic organic contaminants (3). Molecules greater than about 600 da (about 10 Åcross-sectional diameter in terms of molecular breadth) are excluded (Fig. 2). Contaminants larger than the exclusion limit or adsorbed on suspended particles (not dissolved) are neither bioconcentrated nor SPMD-available (2,6).

Herve et al. (8) and Prest et al. (9) compared SPMDs and mussels as monitors of bioconcentrating organochlorine compounds. Mussel-SPMD comparisons were species-specific and depended on noncontaminant stressors. However, mussels and SPMDs were found to be complementary in that mussels alone concentrated some compounds, whereas SPMDs alone concentrated others. A comparison between Green-Lipped mussels and SPMDs in Hong Kong coastal waters showed that SPMDs failed to rank the sites in the same order as mussels in terms of organochlorine contaminant concentrations, nevertheless, in localities where mussels cannot survive, SPMDs may be valuable in providing an indication of potentially bioavailable lipophilic pollutants (10). SPMDs and caged channel catfish were also compared for uptake of polychlorinated dibenzo-p-dioxins/furans (PCDD/Fs) and polychlorinated biphenyls (PCBs). Higher concentrations of analytes were found in SPMDs than in catfish (11). On the contrary, concentrations of DDT and its metabolites detected in SPMD extracts were approximately 1000 times smaller than those measured in fish (*Cyprinus caprio*) tissues in the Mississipi Delta (7).

BASICS OF SAMPLING USING SPMDS

SPMDs will, in general, concentrate from water all neutral hydrophobic chemicals having molecular masses <600 da (10 Å in terms of molecular breadth). No other sampling approach offers this broad a range of applicability with respect to chemical class or molecular mass (1,6). Examples of environmental contaminants that have been effectively sampled by SPMDs include polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), organochlorine pesticides (OCPs), polychlorinated dibenzo-*p*-dioxins and furans (PCDD/Fs), moderately polar organophosphate pesticides, alkylated phenols, pyrethroid insecticides, nitroaromatic compounds, neutral organomettalic compounds, and certain heterocyclic compounds (1,2,6,7,12–15). Other neutral chemicals that have been found in dialysates from field-deployed SPMDs include polybrominated diphenyl ethers, trifluralin, pendimethalin, etc. (7,12). The SPMDs are not designed to concentrate ionic species such as ionic metals, salts of organic acids, or very polar chemicals (1).

SPMDs have been applied successfully not only in natural aquatic environments, but also in sewerage systems and waste water treatment plants (2,14,16,17), as well as in the air phase of soils (18) and in the atmosphere (19,20).

The uptake of organic compounds by SPMDs is in accordance with polymer permeability and equilibrium partition theory. As chemical partitioning coefficients between membrane and water (K_{MW}) and between lipid and water (K_{LW}) increase with the lipophilic indicator (K_{OW}), the SPMD sampling rate is expected to rise proportionally. At saturation, the capacity of the SPMD for a hydrophobic chemical is generally related with the compound's K_{OW}, i.e., the higher a compound's K_{OW}, the greater the capacity of the SPMD for that chemical. Neutral organic chemicals that are hydrophobic (i.e., with $\log K_{\mathrm{OW}} \geq 3$) will be concentrated by the SPMD, but for compounds with $\log K_{\mathrm{OW}} \leq 3$, no significant advantage exists in using SPMDs in preference to other sampling procedures (2).

The uptake of HOCs by SPMDs is linear (or integrative) for the first days of deployment, then it becomes curvilinear (partly integrative) and, finally, when it reaches equilibrium, the amount of HOCs taken up is exactly equal to the amount of HOCs lost (Fig. 3). One of the main principals in using SPMDs in sampling is to operate only in the kinetic phase of uptake, where contaminant uptake is linear. Exposure time must therefore be optimized because, during longer exposure periods, equilibrium between the SPMD and sampling media may be approached for certain contaminants (6). The uptake of HOCs with $\log K_{\mathrm{OW}}\mathrm{s} > 4.9$ is usually linear during exposure times <30 days, whereas for compounds with $\log K_{\mathrm{OW}} < 4.9$, equilibrium concentrations of analytes may be reached or approached in ≤30 days (22). A compound with $\log K_{\mathrm{OW}} = 6$ would need 200 days at a constant water concentration and at an effective sampling rate of 10 L/d to reach 90% equilibrium. However, during the initial period, the uptake rate by the SPMD is linear and first-order (23).

In the linear uptake phase, the concentration of a contaminant in the SPMD will be proportional to the concentration in water as a time-weighted average according to the Eq. 1:

$$C_{\mathrm{SPMD}} = C_{\mathrm{W}} R_{\mathrm{S}} \mathrm{t} / M_{\mathrm{SPMD}} \tag{1}$$

where C_{SPMD} is the concentration of the analyte in the SPMD, C_{W} is the concentration of the analyte in water, t is the exposure time in days, M_{SPMD} is the mass of the SPMD, and R_{S} is the SPMD sampling rate (21). Briefly, the analyte sampling rate (R_{S}) for a given analyte is determined from laboratory or field exposures conducted under controlled conditions. As field-deployed SPMDs often become biofouled, the potential impedance to analyte uptake because of aufwuchs layers needs to be accounted for, thus the environmental sampling rate (R_{SC}) is given by Eq. 2:

$$R_{\mathrm{SC}} = R_{\mathrm{S}} F \tag{2}$$

where F is 1 minus the fractional reduction in uptake flux or sampling rate because of fouling (2). Typically, for no or little biofouling, no correction is made. For mild, medium, and heavy biofouling, uptake impedances of 10%, 30%, and 50%, respectively, are assumed.

The SPMD sampling rates have been demonstrated to be independent of water concentrations and salinity. Sampling rates increase with temperature and $\log K_{\mathrm{OW}}$ (i.e., sampling rates increase with increasing hydrophobicity) up to a $\log K_{\mathrm{OW}}$ of about 5.5 (1,21,24). A decrease in uptake rates for compounds with $\log K_{\mathrm{OW}} > 6$ has been observed in similarity to bioconcentration rates in fish (2). Using the above approach and available R_{SC} values (15,21,24), the bioavailable waterborne concentrations

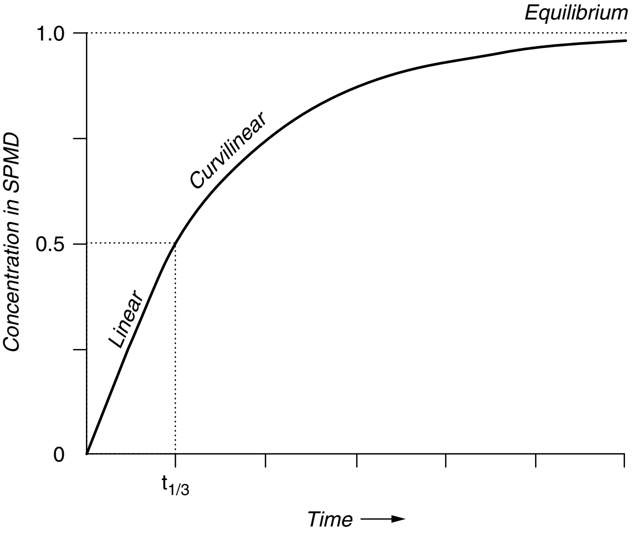

Figure 3. The three phases of SPMD uptake (21).

of typical hydrophobic contaminants have been calculated in many field locations (1,13,25–28).

Any length SPMD with a constant surface area/volume ratio (SA/V) of about 460 cm^2 per ml of triolein, an approximate 0.2 lipid-to-membrane mass ratio, and a 75–90 μm wall thickness is considered to be a "standard SPMD." The uptake rates of specific HOCs by standard SPMDs are presented in Table 1.

SPMDs are great for identifying sources of contamination on small geographical and temporal scales, i.e., when in the study area, the water flow velocities and exposure temperatures are essentially unchanged. A higher observed amount for a particular compound usually indicates a higher exposure concentration. Things become more complicated, however, when between-compound comparisons have to be made (e.g., low K_{OW} compounds may have higher aqueous concentrations, and yet have smaller absorbed amounts because the SPMD sorption capacity is limited) or when substantial between-site differences exist in water flow velocities and temperatures. A third difficulty is that the "SPMD absorbed amount" is not a quantity that is easy to use in environmental fate modeling and in monitoring programs (2). The use of Permeability/Performance Reference Compounds (PRCs) helps to identify and eliminate these problems, establishing a firm link between the absorbed amounts and the aqueous concentrations of contaminants. PRCs are analytically noninterfering compounds, such as perdeuterated PAHs, with moderate to fairly high SPMD fugacity, that are added to SPMD prior to deployment. Their losses during the field exposure compared with losses under controlled laboratory conditions are used to correct for the effects of variations of temperature, diffusion layer thickness, and degree of biofouling on SPMD uptake rates (2,29).

Table 1. R_S Values for Selected HOCs by Standard SPMDs (l/d, 18 °C)

Compound	R_S	Compound	R_S
HCB	2.6^a	PCB-153	4.2^b
α-HCH, β-HCH	0.9^a	PCB-138	4.9^b
Lindane	0.7^a	PCB-180	7.7^b
Heptachlor	3.6^a	Naphthalene	0.9^a
Heptachlor epoxide	2.9^a	Acenapthene	2.3^a
Endosulfan	1.1^a	Fluorene	1.7^a
p-p′-DDE	5.5^a	Phenanthrene	3.4^a
Dieldrin	1.8^a	Pyrene	5.2^a
Endrin	3.1^a	Benzo[a]pyrene	4.3^a
p-p′-DDD	3.1^a	2,3,7,8-TCDD	3.8^a
p-p′-DDT	3.2^a	2,3,7,8-TCDF	3.7^a
PCB-28	2.3^b	1,2,3,7,8-PeCDD	3.4^a
PCB-52	1.8^b	1,2,3,7,8-PeCDF	3.8^a
PCB-101	3.0^b	1,2,3,4,7,8-HxCDD	4.0^a
PCB-118	4.2^b	1,2,3,4,7,8-HxCDF	2.7^a

[a] (21).
[b] (24).

DEPLOYMENT, PROCESSING, AND ANALYSIS OF SPMDS

SPMDs are available commercially, but they can also be prepared in the laboratory. Once SPMDs are acquired or prepared, they should remain in vapor-tight tins until deployment, and, ideally, they should be stored to −15 °C (1). Several deployment devices have been proposed for the passive sampling with SPMDs (25,27,30), which usually are made of stainless steel or other metals. Use of plastic is not suggested, as it can be a potential interference because of the possible presence of leachable organics. Hardware must be designed so as to minimize abrasion of the membrane, especially

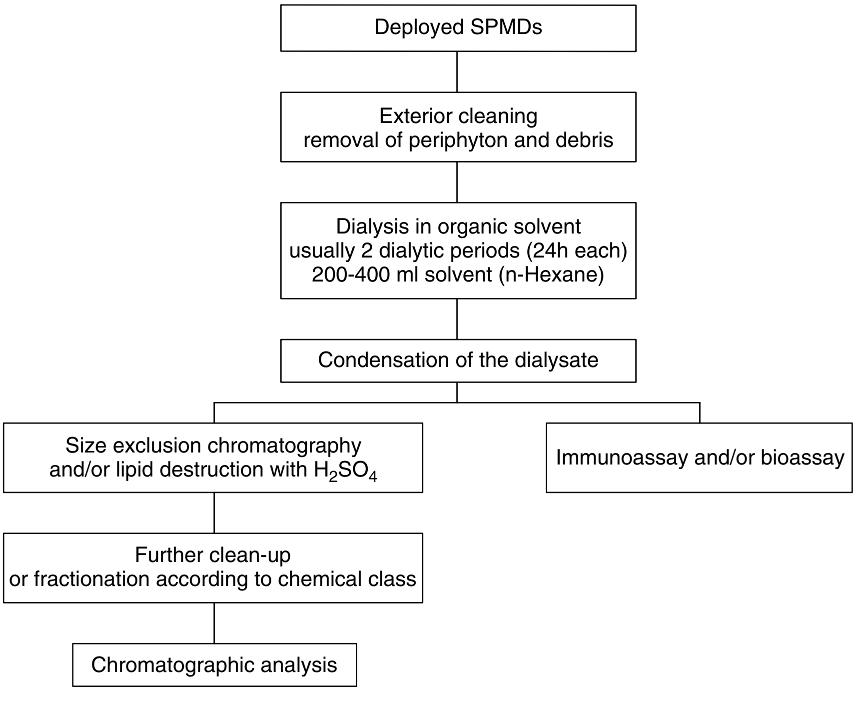

Figure 4. Analytical scheme for the processing of SPMDs.

in turbulent environments (1), and they should also be deployed in such a way that the maximum exposure of membrane surface is ensured. When the deployment period is over, the SPMDs should again be put in vapor-tight tins and kept frozen until analysis.

SPMDs that have been exposed to the aquatic environment can contain a diversity of lipophilic contaminants other than those specifically targeted (3). Also, the outsides of exposed SPMDs are often biofouled by aufwuchs (attached microscopic or macroscopic organisms) and must be cleaned before the targeted residues can be determined (6,13,14,31). Contaminants concentrated in SPMDs are easily recovered by dialysis into organic solvents such as n-hexane, cyclopentane, or cyclohexane (6,14,16,17,19,25–27,30–33). Dialysates from exposed SPMDs can conceivably contain biogenic material from aufwuchs, codialyzed triolein, methyl oleate, and impurities and oligomers from polyethylene (2,3,6). Analytical procedures are therefore required to overcome these impediments. Size exclusion chromatography (SEC) is usually employed to remove interferences, especially codialyzed polyethylene oligomers, oleic acid, and methyl oleate, although oleic acid may be present even after SEC-treatment (1). A solution to this problem may be the application of a destruction technique, like hydrolysis or acidic treatment. The problem with these treatments is that a part of target compounds may also be destroyed. Compounds that resist strongly to acidic treatment are PCBs and some OCPs (34). Fig. 4 is a generalized representation of the analytical scheme.

Quality control of sampling employing SPMDs requires the use of field and trip blanks accounting for contamination of SPMDs by airborne chemicals during the time between opening the can and deployment, and for potential interferences introduced during transport. Laboratory blanks may also be used, together with reagent blanks and duplicate samples, to ensure the quality of results.

The range of concentrations in which HOCs are present in SPMDs can vary from a few pg SPMD^{-1} (14) to μ g SPMD^{-1} (2) or higher depending on the concentration of the specific chemical in water, the size of the SPMD deployed, the uptake rate, and the detection limit of the analytical technique employed. In any case, the concentration of a chemical in the SPMD will be higher than in the ambient water, which is one of the major advantages of the use of SPMDs.

BIBLIOGRAPHY

1. Petty, J.D. et al. (2000). An approach for assessment of water quality using semipermeable membrane devices (SPMDs) and bioindicator tests. *Chemosphere* **41**: 311–321.
2. Petty, J.D. et al. (2000). Considerations involved with the use of semipermeable membrane devices for monitoring environmental contaminats. *J. Chromatogr A.* **879**: 83–95.
3. Huckins, J.N., Tubergen, M.W., and Manuweera, G.K. (1990). Semipermeable membrane devices containing model lipid: A new approach to monitoring the bioavailability of lipophilic contaminants and estimating their bioconcentration potential. *Chemosphere* **20**: 533–552.
4. Soedergren, A. (1987). Solvent-filled dialysis membranes simulate uptake of pollutants by aquatic organisms. *Environ. Sci. Technol*. **21**(9): 855–859.
5. Soedergren, A. (1990). Monitoring of persistent, lipophilic pollutants in water and sediment by solvent-filled dialysis membranes. *Ecotox. Environ. Safe* **19**(2): 143–149.
6. Lebo, J.A. et al. (1992). Use of semipermeable membrane devices for *in situ*monitoring of polycyclic aromatic hydrocarbons in aquatic environments. *Chemosphere* **25**: 697–718.
7. Zimmerman, L.R., Thurman, E.M., and Bastian, K.C. (2000). Detection of persistent organic pollutants in the Mississippi Delta using semipermeable membrane devices. *Sci. Tot. Environ.* **248**: 169–179.
8. Herve, S. et al. (1995). Lipid-filled semipermeable membrane devices and mussels as samplers of organochlorine compounds in lake water. *Environ. Sci. Pollut. Res.* **2**: 24–30.
9. Prest, H.F., Richardson, B.J., Jacobson, L.A., Vedder, J., and Martin, M. (1995). Monitoring Organchlorines with semi-permeable membrane devices (SPMDs) and Mussels (Mutilus-Edulis) in corio Bay, Victoria, Australia. *Mar. Pollut. Bull.* **30**: 543–554.
10. Richardson, B., Zheng, G., and Lam, P. (2000). Are SPMDs as good as mussels? A comparison of the Green-Liped Mussel (*Perna Viridis*) and SPMDs for monitoring organochlorine contaminats in Hong-Kong coastal waters. 6th *International SPMD Workshop and Symposium*, Columbia, MO.
11. Gale, R.W. etal. (1997). Comparison of the uptake of dioxin like compounds by caged channel catfish and semipermeable membrane devices in the Saginaw River, Michigan. *Environ. Sci. Technol.* **31**: 178–187.
12. Petty, J.D., Huckins, J.N., Martin D.B., and Adornato, T.G. (1995). Use of semipermeable membrane devices (SPMDs) to determine bioavailable organochlorine pesticide residues in streams receiving irrigation drainwater. *Chemosphere* **30**: 1891–1902.
13. Wang, Y., Wang, Z., Wang, C.X., and Wang, W.H. (1999). Uptake of weakly hydrophobic nitroaromatics from water by semipermeable membrane devices (SPMDs) and by goldfish (Carassius auratus). *Chemosphere* **38**(1): 51–66.
14. Wang, Y. et al. (2001). Monitoring priority pollutants in a sewage treatment process by dichloromethane extraction and triolein-semipermeable membrane device. *Chemosphere* **43**: 339–346.
15. Rantalainen, A.L., Cretney, W.J., and Ikonomou, M.G. (2000). Uptake rates of Semipermeable membrane devices (SPMDs) for PCDDs, PCDFs and PCBs in water and sediment. *Chemosphere* **40**: 147–158.
16. Stuer-Laudridsen, F. and Kjolholt, J. (2000). Identification of selected hydrophobic organic contaminants in wastewater with semipermeable membrane devices (SPMDs). *Wat. Res.* **34**: 3478–3482.
17. Samara, C. and Katsoyiannis, A. (2002). Evaluation of Persistent Organic Pollutants (POPs) in wastewater by means of Semipermeable Membrane Devices (SPMDs). *Proceedings of the* 2nd *European Conference on Pesticides and Related Organic Contaminants*, September 25–29, Corfu, Greece, pp. 109–116.
18. Rantalainen, A.L., Paasivirta, J., and Herve, S. (1998). Uptake of chlorohydrocrbons from soil by lipid-containing semipermeable membrane devices (SPMDs). *Chemosphere* **36**: 1415–1427.
19. Petty, J.D., Huckins, J.N., and Zajicek, J.I. (1993). Application of semipermeable membrane devices (SPMDs) as passive air samplers. *Chemosphere* **27**: 1609–1624.

20. Prest, H.F., Jacobson, L.A., and Huckins, J.N. (1995). Passive sampling of water and coastal air via semipermeable membrane devices. *Chemosphere* **35**: 1351–1361.
21. Huckins, J.N., Petty, J.D., Prest, H.F., Clark, R.C., Alvarez, D.A., Orazio, C.E., Lebo, J.A., Cranor, W.L., and Johnson, B.T. (2000). *A Guide for the Use of Semipermeable Membrane Devices (SPMDs) as Samples of Waterborne Hydrophobic Contaminats*. Edition of Columbia Environmental Research Center, USGS, USDI, and California Analytical Division, Agilent Technologies, Inc., Palo Alto, CA.
22. Huckins, J.N. et al. (1999). Determination of uptake kinetics (sampling rates) by lipid-containing semipermeable membrane devices (SPMDs) for polycyclic aromatic hydrocarbons in water. *Environ. Sci. Technol.* **33**: 3918–3923.
23. Huckins, J.N. et al. (1993). Lipid-containing semipermeable membrane devices for monitoring organic contaminants in water. *Environ. Sci. Technol.* **27**: 2489–2496.
24. Meadows, J.C. et al. (1998). Estimation of uptake rate constatnts for PCB congeners accumulated by semipermeable membrane devices and brown trout (Salmo trutta). *Environ. Sci. Technol.* **32**: 1847–1852.
25. Petty, J.D. et al. (1995). Determination of waterborne bioavailable organochlorine pesticide residues in the lower Missouri river. *Environ. Sci. Technol.* **29**: 2561–2566.
26. Petty, J.D. et al. (1998). Determination of bioavailable contaminats in the lower Missouri River following the flood of 1993. *Environ. Sci. Technol.* **32**: 837–842.
27. Lebo, J.A. et al. (1995). Use of Semipermeable membrane devices as an in-situ sampler of waterborne bioavailable PCDD and PCDF residues at sub-parts-per-quadrillion concentrations. *Environ. Sci. Technol.* **29**(11): 2886–2892.
28. Echols, K.R. et al. (2000). Comparing polychlorinated biphenyl concentrations and patterns in the Saginaw River using sediment, caged fish, and semipermeable membrane devices. *Environ. Sci. Technol.* **34**: 4095–4102.
29. Huckins, J.N. et al. (2002). Development of the permeability/performance reference compound approach for *in situ* calibration of semipermeable membrane devices. *Environ. Sci. Technol.* **36**: 85–91.
30. Ellis, G.S. et al. (1995). Evaluation of lipid-containing semipermeable membrane devices for monitoring organochlorine contaminants in the upper Mississippi river. *Environ. Sci. Technol.* **14**: 1875–1884.
31. Rantalainen, A.L., Ikonomou, M.G., and Roge, I.H. (1998). Lipid containing semipermeable mebrane devices (SPMDs) as concentrators of toxic chemicals in the lower Fraser River, Vancouver, British Coloumbia. *Chemosphere* **37**(6): 1119–1138.
32. Lebo, J.A. et al. (1996). Use of the semipermeable membrane device (SPMD) to sample polycyclic aromatic hydrocarbon pollution in a lotic system. *Polycycl. Arom. Comp.* **8**(1): 53–65.
33. Van Stee, L.L.P. et al. (2002). Use of semi-permeable membrane devices and solid phase extraction for the wide-range screening of microcontaminants in surface water by GC-AED/MS. *Water Res.* **36**: 4455–4470.
34. USEPA. (1998). *Standard operating procedure for acid cleanup of organic extracts for PCB analysis*. CRL Method 3600—Acid. *CRL SOP GC016*.

READING LIST

Chiou, C.T. (1995). Partition coefficients of organic compounds in lipid-water systems and correlations with fish bioconcentration factors. *Environ. Sci. Technol.* **19**: 57–62.

RIVER-CONNECTED AQUIFERS: GEOPHYSICS, STRATIGRAPHY, HYDROGEOLOGY, AND GEOCHEMISTRY

Tom A. Al
Kerry T. MacQuarrie
Karl E. Butler
Jean-Christophe Nadeau
Melissa R. Dawe
Larry Amskold
University of New Brunswick
Fredericton, New Brunswick,
Canada

GENERAL

When water wells are constructed in aquifers that occur adjacent to rivers, pumping from the wells commonly causes infiltration of river water toward the wells, thereby increasing the total available water yield. This type of water production has been used extensively in many European countries (1) and, to a lesser extent, in North America.

The flow of river water through permeable riverbed and aquifer sediments, commonly referred to as riverbank filtration, can provide filtration and general purification benefits. Riverbank filtration provides some protection to water supplies from persistent and shock-loaded biological and chemical contaminants that may be present in the river water (2), and commonly provides a first step in the treatment of water supplies (1,3,4); however, in some cases, water quality problems associated with the river water may be transmitted through the aquifer to break through to the wells (5). Riverbank filtration can completely or partially remove suspended solids and pathogens from the infiltrating river water (1,6–11). In addition, the infiltrating river water undergoes chemical modifications, some of which may be beneficial, such as adsorption of metals (6), denitrification (12–15), and biodegradation of organic compounds (15,16). Some chemical changes may be detrimental such as increased hardness (10) and increased concentrations of dissolved iron and manganese (10,12,13,17–19).

In the development and management of river-recharged aquifer systems, the two most important considerations are the sustainability of water quantity and water quality. The quantity of water available is dependent on the river discharge and on the hydrostratigraphy between the riverbed and adjacent pumping wells. The quality of water extracted from the wells is influenced by the quality of the river water, the nature of flow paths from the river to the wells, the travel time in the riverbed sediments and aquifer (9), and the geochemical and mineralogical characteristics of the riverbed and aquifer sediments (19). Development of an understanding of these characteristics is therefore a prerequisite for managing well fields that use induced infiltration of river water.

In 2001, a multidisciplinary investigation of the Fredericton Aquifer was initiated with the objectives of determining:

- stratigraphic controls on the flow path and infiltration flux for river water entering the aquifer,
- the travel time between the river and the pumping wells, and
- the spatial and temporal variation in geochemical reactions that occur as river water infiltrates the aquifer, leading to changes in water quality.

In this contribution, the Fredericton Aquifer, which is considered typical of many alluvial-valley water-supply aquifers, is used as a case study to illustrate possible approaches to the characterization of river-recharged aquifer systems, and describe some of the important stratigraphic, hydrogeologic, and geochemical features of these systems. The combined use of geophysical, hydrogeological, and geochemical methods of investigation has led to a more complete understanding of river water infiltration toward municipal wells and the resulting water quality implications.

FREDERICTON AQUIFER CASE STUDY

The study area is located in Fredericton, New Brunswick, Canada (Fig. 1) (45.87 °N, 66.53 °W). The Saint John River at Fredericton lies in a broad bedrock valley that is partially filled with up to 65 m of unconsolidated sediments. The City of Fredericton obtains its entire water supply (approximately 25,000 m^3/day) from a semiconfined aquifer that lies within the river valley. As shown in Fig. 2, drilling has revealed that the coarse-grained sediments that make up the aquifer have the morphology of a buried esker ridge that runs approximately SE-NW underneath the city and is thought to follow the general drainage direction of the bedrock valley floor. The ridge is up to 30-m thick and contains a large range in particle sizes from sand to boulders (20). The deposit thins as it spreads along the north and south valley walls.

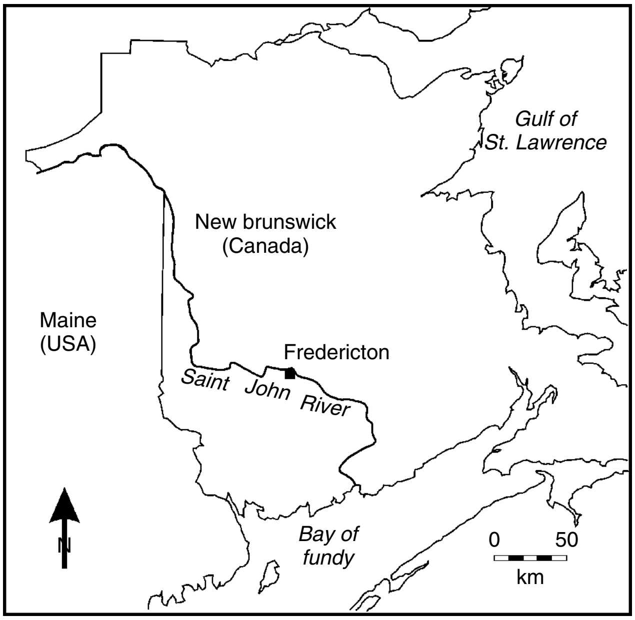

Figure 1. Location of the Fredericton Aquifer case study.

The city currently has eight wells located on the south side of the Saint John River (Fig. 2), and all wells are completed in a sand and gravel aquifer with screen depths between 25 m and 48 m below ground level. Four of the eight wells are located within 150 m of the river. A discontinuous silt/clay unit overlies the aquifer, and in locations where it is discontinuous, referred to as windows, the aquifer is recharged by the river or by precipitation. The aquifer may also be recharged by an upward flux of groundwater from the underlying bedrock.

A history of elevated Mn concentrations has existed in several of the Fredericton supply wells, and Mn concentrations in groundwater from some of the production

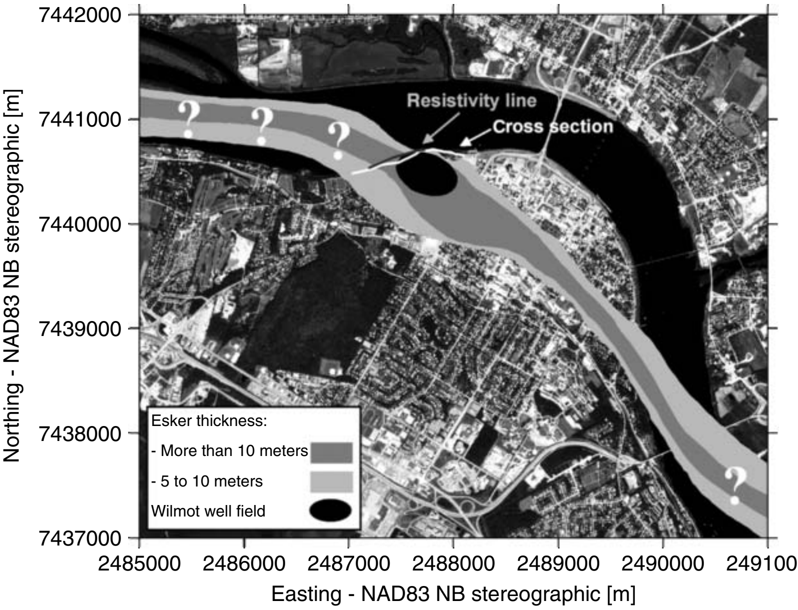

Figure 2. Satellite photo with superimposed map of the buried esker ridge that forms the main part of the aquifer underlying the City of Fredericton and adjacent Saint John River. The well field is highlighted in black [from (21)].

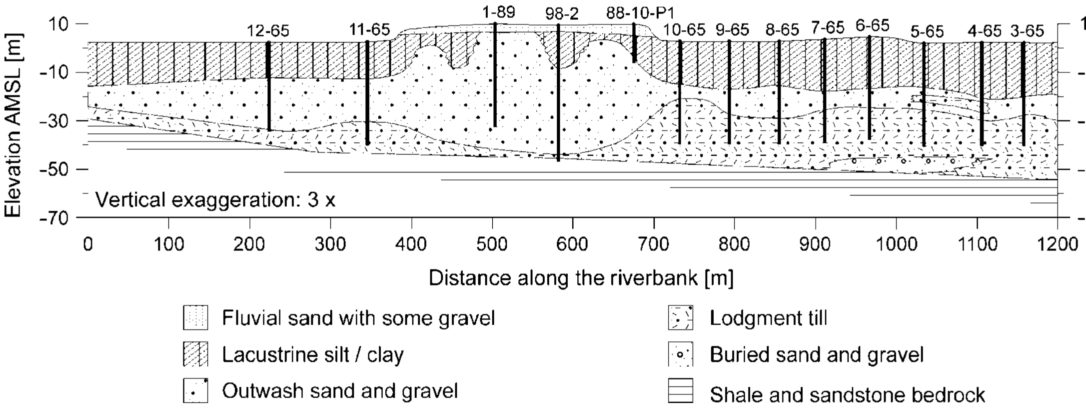

Figure 3. Cross section located along the south shore of the Saint John River and passing obliquely over the buried esker at the edge of the well field as shown in Fig. 2. The rise in surface elevation over the esker ridge reflects the fact that boreholes in that area were drilled atop the river's bank whereas most others were drilled at the river's edge. Borehole logs are from (22) and the City of Fredericton. Undulations in the esker surface are inferred from the resistivity imaging results in Fig. 4.

wells exceed the Canadian Drinking Water Guideline of 0.05 mg/L by up to 3 mg/L. Previous work (19,21) has demonstrated that dissolved organic carbon (DOC) present in the infiltrating river water promotes biogeochemical reactions in the adjacent aquifer, leading to elevated Mn concentrations.

Stratigraphic Controls on River Water Infiltration

In the approximately 50 years since the first water wells were developed in the Fredericton Aquifer, there have been over 700 drill holes completed to sample the unconsolidated glacial and postglacial deposits in the bedrock valley. The information obtained from these drilling activities is extremely useful in developing a general understanding of the glacial stratigraphy, including the aquifer and its enclosing geologic units (Fig. 3). However, despite the large amount of drill-hole data available, the detailed knowledge of stratigraphy below the river that is required to identify discrete flow paths from the river to the wells could not be derived from the drill-hole data. A very large number of additional drill holes would have been required to obtain the necessary geologic information, and such a conventional approach was not feasible because of financial constraints. Consequently, an alternative approach was taken using a variety of near-surface geophysical techniques to determine the subsurface distribution of important stratigraphic units.

Shoreline Resistivity Imaging. The clay content of the silt/clay aquitard unit creates a contrast in electrical conductivity/resistivity with the surrounding sands and gravels. This contrast allows for the effective use of electrical and electromagnetic geophysical techniques for mapping discontinuities (windows) in the aquitard that overlies the aquifer. Resistivity data were first acquired along a 500-m profile located on the shoreline of the river beside the Wilmot Well Field (Fig. 2). The data were acquired in dipole-dipole mode using 10-m dipoles and five apparent resistivity measurements for every current injection location (i.e., $a = 10$ m, $n = 5$) in order to provide a dataset suitable for two-dimensional inversion.

Figure 4 shows the minimum structure resistivity model obtained by inversion (23). The model indicates that the aquitard is laterally continuous near either end of the line and confirms the presence of a window zone in the area where we would expect it based on the cross section in Fig. 3. The interpreted thicknesses of clay/silt on either side of the window (~12–15 m) are in agreement with geologic logs from three boreholes that had previously been drilled on the shore. Below the high resistivity layer that is interpreted to represent the aquifer sand and gravel, a slight decrease occurs in resistivity depths greater than 30 m, which may represent the glacial till layer; however, the depth of penetration and/or resistivity contrast are not sufficient to conclusively define this contact. Most notably, the resistivity image reveals the presence of two silt/clay pods within the window, neither of which had been well defined by drilling. These pods reduce the area available for river water infiltration through the window into the aquifer.

Figure 4 also shows a profile of apparent electrical conductivity as measured along the shoreline with an electromagnetic terrain conductivity meter (Geonics EM31) having an effective depth of exploration of approximately 6 m (24). Comparison of this profile to the resistivity inversion result shows that apparent conductivities are relatively high in areas underlain by the clay/silt aquitard and low where it is absent. The EM method's speed as a reconnaissance tool and sensitivity to the presence of the aquitard window encouraged us to try adapting it for use on the river in the search for the river bottom recharge area.

Riverine Electromagnetic and Seismic Surveys. Although the resistivity survey revealed the width of the potential recharge zone along the shore and its two-dimensional morphology, a need still existed to define this zone underneath the river where the majority of recharge was

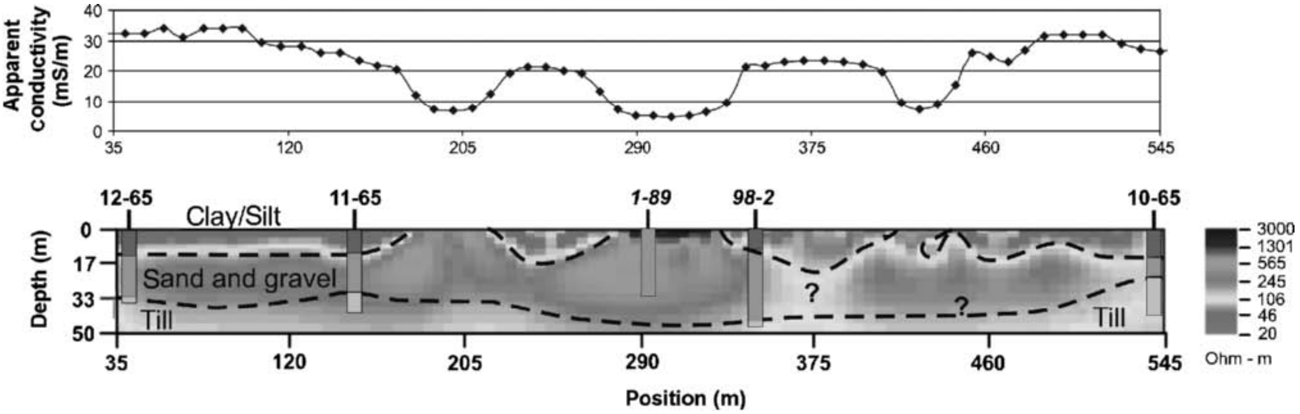

Figure 4. EM31 apparent conductivity profile (top), and resistivity depth section model (bottom) for a 500-m transect located along the shoreline of the Saint John River at the edge of the well field as shown in Fig. 2. Low resistivity zones (30 to 60 ohm-m) indicated by hot colors are interpreted as the conductive silt/clay aquitard, whereas the higher resistivities (>200 ohm-m) are interpreted as sand and gravel. In comparison, the measured resistivity of the river water was approximately 100 ohm-m.

expected to occur. Riverine EM and seismic reflection surveys were carried out for that purpose.

Electromagnetic Surveys. EM31 and deeper penetrating EM34 terrain conductivity meters were adapted for use on the water by deploying them in a nonmetallic canoe and inflatable raft. A DGPS was used to navigate along survey lines and record the locations of measurements made at 2–5 second intervals as the craft moved slowly along the water surface. In order to compensate for variations in apparent conductivity related to variations in water depth (~0.5 to 6 m), we devised an approximate bathymetric correction based on a one-dimensional earth model (23). A measurement of the river water's conductivity (10.3 mS/m) and a bathymetric map extracted from the seismic data described below were used as the basis for these corrections. Finally, the bathymetry-corrected apparent conductivity data were "leveled" to remove small miss-ties at the line intersections that were attributed to temporal drift in the instrument's response.

Figure 5 is a map of the leveled, bathymetry-corrected apparent conductivity measured using the EM34 in such a way as to achieve an effective depth of penetration of approximately 15 m (i.e., vertical dipole mode measurements with a transmitter-receiver coil separation of 10 m). The map created from EM31 data was similar but more strongly influenced by the varying thickness of fluvial sands (~0 to 6 m) overlying the aquitard. The most striking feature in Fig. 5 is the oblong, low conductivity zone trending under the river in a northwesterly to westerly direction. Based on a comparison of the EM31 and EM34 maps, this zone is interpreted as a window in the aquitard where sand and gravel extends upward through the surrounding silt/clay layer. The two elongated high conductivity anomalies within the window were previously detected by the shoreline resistivity survey and are interpreted as silt/clay pods that were deposited within depressions on the paleo-surface of the esker. The strong, erratic anomalies centered 250 m NE of the window are probably caused by metal cables that are known to have been discarded on the riverbed.

Single Channel Seismic Reflection Surveys. Single channel seismic reflection surveys were carried out with the expectation that the interface between aquitard and aquifer would be detectable because of differences in the seismic velocities or densities of the silt/clay and sand lithologies. The instrumentation consisted of an IKB Seistec sub-bottom profiler with an electrodynamic boomer source and integrated hydrophone receiver that was well suited for operation in the shallow riverine environment. Depth penetration was dependent on river bottom type, but many areas yielded excellent records showing reflections with dominant frequencies of 2–2.5 kHz from depths as great as 50 m (23).

Seismic profile C1040, shown in Fig. 6, lies subparallel to the river edge, 100–200 m offshore (Fig. 5). The eastern half of the line, where we expected aquitard material to be present, exhibits excellent reflectivity extending to two-way times as great as 30 ms, which corresponds to a depth of 24 ± 1 m using the average P-wave velocity of 1600 ± 50 m/s deduced for this depth range from borehole seismic surveys carried out on shore (25).

There are two prominent reflectors in this and most other records. The shallower one appears to come from the middle of the silt/clay aquitard where an abrupt increase in P-wave velocity was observed in borehole seismic records (25). The deeper one is thought to be the contact between the aquitard and underlying sand and gravel aquifer. Vestiges of slightly deeper reflectors exist in the record, one of which may represent the contact between the aquifer and glacial till. However, the relatively high frequencies employed in this survey were unable to penetrate far through the coarse-grained sediments underlying the aquitard, and no evidence exists of a reflection from bedrock.

In the central part of the seismic profile below position 460 m, the prominent aquitard reflectors terminate

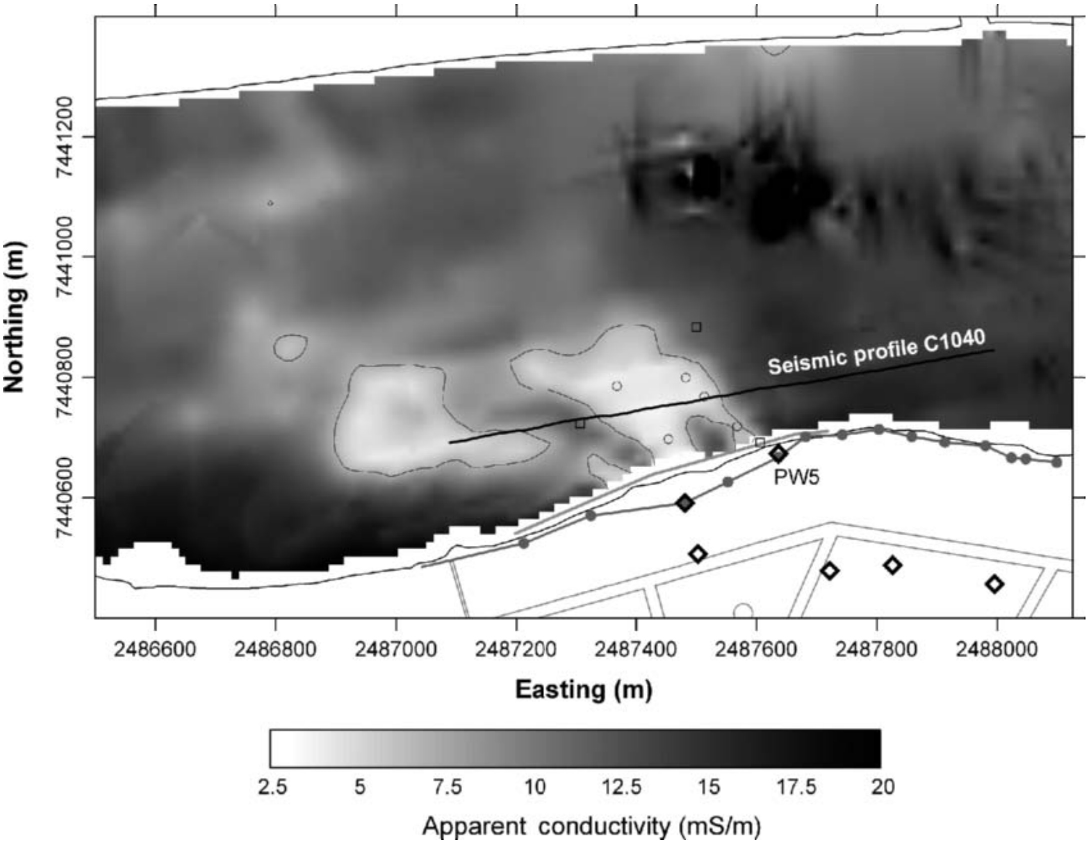

Figure 5. Bathymetry-corrected apparent conductivity map for the bottom of the Saint John River based on data collected with an EM34 terrain conductivity meter having an effective exploration depth of ~15 m. The map reveals a low conductivity zone of irregular shape (approximated by the 7 mS/m contour) interpreted as an aquitard window. The open circles and squares represent eight boreholes, three of which encountered aquitard (squares) and five of which did not. The locations of pumping wells (diamonds) and of the shoreline resistivity transect and cross section are also shown.

abruptly against a zone of low reflectivity. The coincident drops in EM31 and EM34 apparent conductivities shown in Fig. 6 lead us to conclude that an electrically conductive and laminated aquitard is here draped on the edge of the more resistive and massive-textured sand and gravel esker ridge. The region of shallow reflectivity and elevated conductivity observed in the western half of the profile, below positions 140 to 280 m, is similarly attributed to the presence of a silt/clay pod embedded in the top of the esker.

Confirmation of the Geophysical Interpretation. Confirmation of the presence of the aquitard window was provided by a drilling program on the river designed to core the aquifer and install piezometers for hydrogeological and hydrogeochemical investigations. The locations of eight holes drilled from a floating barge are shown in Fig. 5. The five holes identified as open circles were sampled continuously using a split-spoon through uninterrupted sands and gravels to depths ranging from 6 to 25 m below the river bottom. The absence of clay/silt is in agreement with the low bathymetry-corrected apparent conductivities obtained at those locations. In contrast, the three holes shown as open squares in Fig. 5 all encountered clay/silt aquitard material, consistent with their location in areas of elevated apparent conductivities.

Water Travel Times

Travel times between the riverbed and the wells can be assessed using tracer tests; however, because of the logistical difficulties of using artificial tracers in a large river such as the Saint John River, we have used two natural tracers, ^{222}Rn and water temperature. Velocities obtained from these tracers were compared with velocities in engineered slow-sand filtration systems, which are very effective for the removal of micro-organisms and pretreatment of drinking water (26). It must be noted that the velocities obtained to date are based on the assumption of vertical infiltration of river water into the underlying sediments; the actual flow paths will also have a horizontal component that has not been considered. The actual flow paths will therefore be longer, and the corresponding resultant velocities higher, than the vertical velocities we present here.

Water Temperature. The water temperature tracer technique is based on the occurrence of temporal temperature contrasts between the river water and the

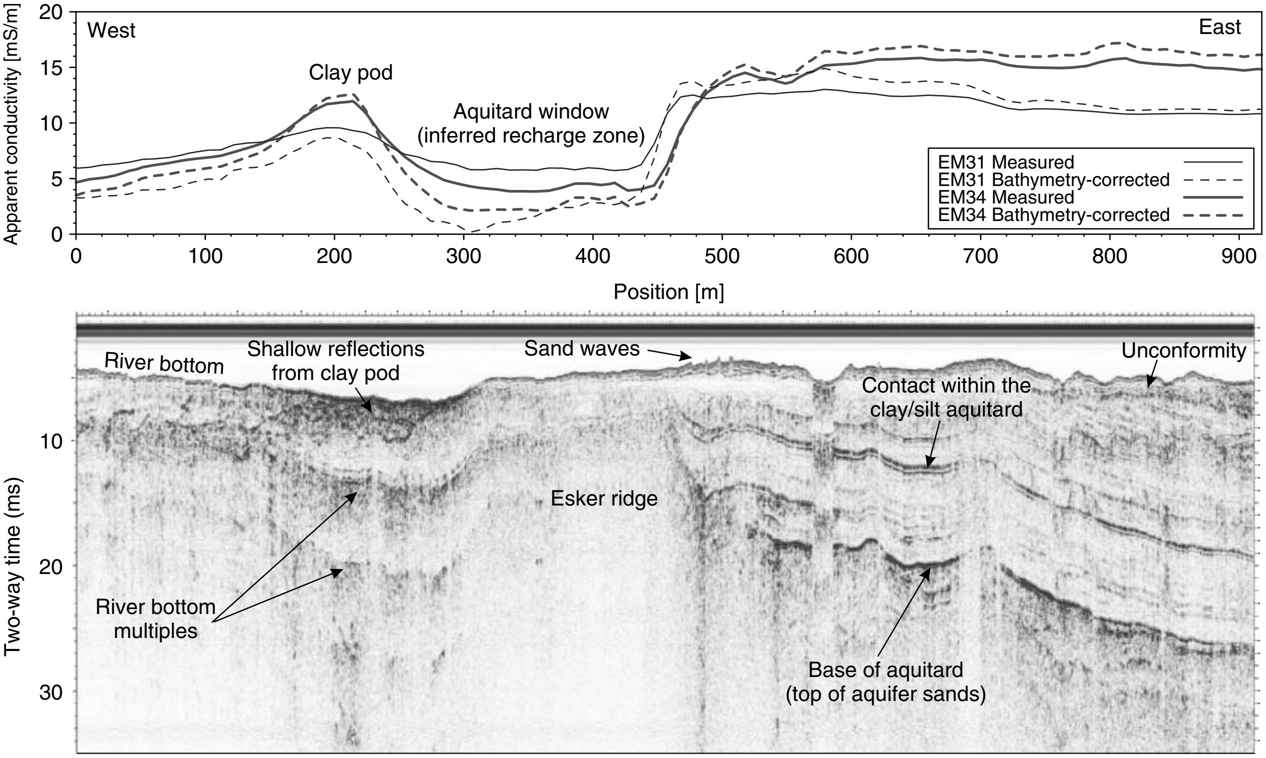

Figure 6. Seismic section and apparent conductivity profiles along the 920-m line C1040 shown in Fig. 5. The seismic record extends to approximately 28-m depth and is displayed with a vertical exaggeration of 12 times. Two strong reflectors, corresponding to interfaces within and at the base of the aquitard, are evident in the seismic data and are draped on the flank of the esker ridge below position 460 m. The apparent conductivity profiles reveal the same contact as well as the presence of a clay pod incised into the esker.

underlying aquifer. The method requires time-series measurements of the groundwater temperature along a flow path from the river toward the wells, and numerical modeling of these time series to determine groundwater flow velocities (27).

Seasonal time-series temperature data were collected at three locations (Fig. 7) within the aquitard window that was identified by EM surveys. At each location, data were collected with temperature loggers located at five or six depths distributed between 1.2 m and 18 m below the riverbed. In addition, one temperature logger was located at the riverbed to record the river water temperature changes.

Figure 8 presents representative data from location T1. The river water temperature decreases in the fall, approaches 0 °C during winter ice cover, and then gradually increases after ice break up reaching a maximum in midsummer. Groundwater temperatures recorded at T1 also exhibit significant seasonal variation with temperatures ranging from slightly greater than 0 °C to 22 °C. With increasing depth below the riverbed, groundwater temperature time series display an increasing lag relative to river water temperature. The lag reflects the time required for infiltration-induced transport of heat from the riverbed to the underlying sediments, and confirms the initial stratigraphic interpretation that, at this location, river water moves downward into the aquifer. Similar trends were also observed in data from locations T2 and T3.

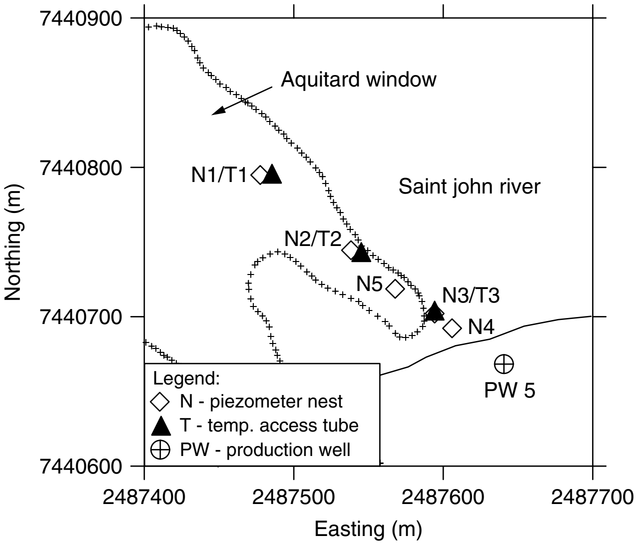

Figure 7. Location map of eastern portion of aquitard window showing the position of multilevel piezometer nests, temperature access tubes, and production well PW5.

A finite-element solution (28) for the one-dimensional partial differential equation for heat convection and

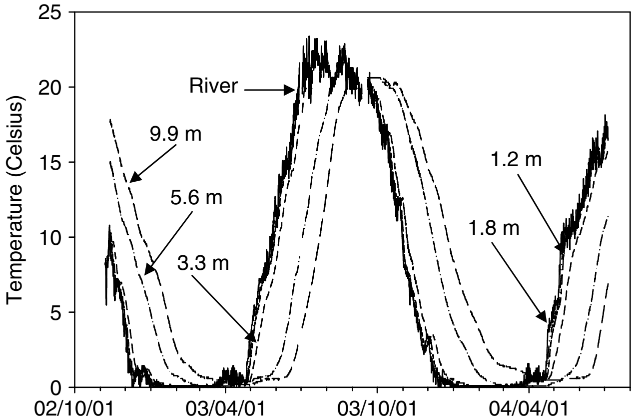

Figure 8. Saint John River and subsurface temperature time series for location T1. Locations T2 and T3 have similar temperature trends.

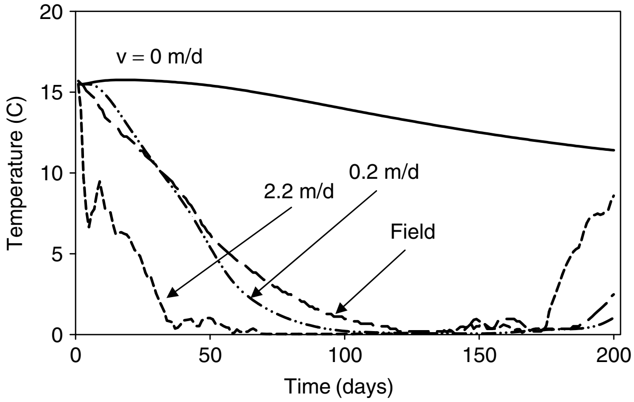

Figure 9. Comparison of simulated and measured temperatures at a depth of 4.7 m below the riverbed at location T3. To demonstrate the sensitivity to the infiltration velocity, simulation results are shown for three average linear velocity values, with v = 0 m/d representing pure heat conduction.

conduction:

$$[\varepsilon\rho^f c^f + (1-\varepsilon)\rho^s c^s]\frac{\partial T}{\partial t} + \rho^f c^f q^f \frac{\partial T}{\partial z} - \frac{\partial}{\partial z}\left(\lambda \frac{\partial T}{\partial z}\right) = 0 \quad (1)$$

where T is the system temperature, ε is the porosity, $\rho^f c^f$ is the volumetric heat capacity of the fluid, $\rho^s c^s$ is the volumetric heat capacity of the solids, q^f is the vertical Darcy flux, and λ is the thermal conductivity, was used to simulate temperatures in the recharge zone. Using the measured river water temperatures as the upper boundary condition, downward Darcy velocities were obtained by trial-and-error adjustment of the hydraulic head boundary conditions to reproduce the temperature time series observed beneath the riverbed. Thermal parameters were derived from correlations between sediment physical properties and thermal properties (27). An example of a comparison between simulated and measured temperature time series for several Darcy velocity values is presented in Fig. 9. For this particular location and depth (T3, 4.7 m depth), a velocity of 0.2 m/day provides a good agreement with the field data. Using this approach, vertical groundwater velocities have been estimated from each of the time-series records at locations T1, T2, and T3, and the results are summarized in Table 1.

Radon. The use of ^{222}Rn as a tracer for determining groundwater travel times is based on the knowledge that soluble ^{222}Rn is formed by the radioactive decay of ^{226}Ra. Given that ^{222}Rn is a gas that partitions readily from water into the air, and that ^{226}Ra is present dominantly in the mineral constituents of rocks and sediments, the activity of ^{222}Rn is generally very low in surface water and increases with travel time after infiltration of surface water into sediments. Steady-state activities are generally reached in saturated groundwater flow systems after residence times of approximately 30 days (29).

Water samples were collected for ^{222}Rn analyses from nested piezometers installed in the riverbed sediments to depths of between 4 and 25 m (Fig. 7). To estimate the steady-state ^{222}Rn activity for the aquifer, samples were also collected from several on-land monitoring wells.

Assuming that ^{222}Rn progenitors (e.g., uranium and radium) are homogeneously distributed throughout the aquifer sediments and that a steady-state ^{222}Rn activity, A_e, is attained, the residence time for infiltrated water can be estimated using (29):

$$t = \ln(1 - A_t/A_e)/\lambda' \qquad (2)$$

where t is the residence time, A_t is the measured ^{222}Rn activity in each piezometer, A_e is the steady-state ^{222}Rn activity, and λ' is the radioactive decay constant for ^{222}Rn.

A steady-state value of $A_e = 36$ Bq/L was used in Eq. 2 because this was the highest activity found in a series of on-land monitoring wells adjacent to the river. Activities of ^{222}Rn varied with depth and location from 13 to 23 Bq/L at N1, 14 to 22 Bq/L at N2, and 23 to 29 Bq/L at N3. These activities correspond to residence times of 2.4 to 5.5 days at N1, 2.8 to 5.3 days at N2, and 5.5 to 8.8 days at N3. Downward flow velocities were calculated by dividing the piezometer screen (sample) depth by the corresponding residence time, which resulted in a total of eleven estimates of vertical velocity in the infiltration area, which are summarized in Table 1. Heat transport-derived velocities and those based on ^{222}Rn activity are similar in magnitude.

Comparison with Slow-Sand Filtration. Average linear velocities for slow-sand filters are typically in the range of 5 to 20 m/day with sand bed thicknesses of approximately one meter (26). The vertical velocity estimates we have obtained are generally lower that slow-sand filtration

Table 1. Ranges of Vertical Average Linear Velocity (m/day)

Method	T1/N1	T2/N2	T3/N3
Temperature	1.5–2.6	0.2–1.3	0.1–0.2
^{222}Rn	1.1–4.0	1.6–5.0	0.7–2.7

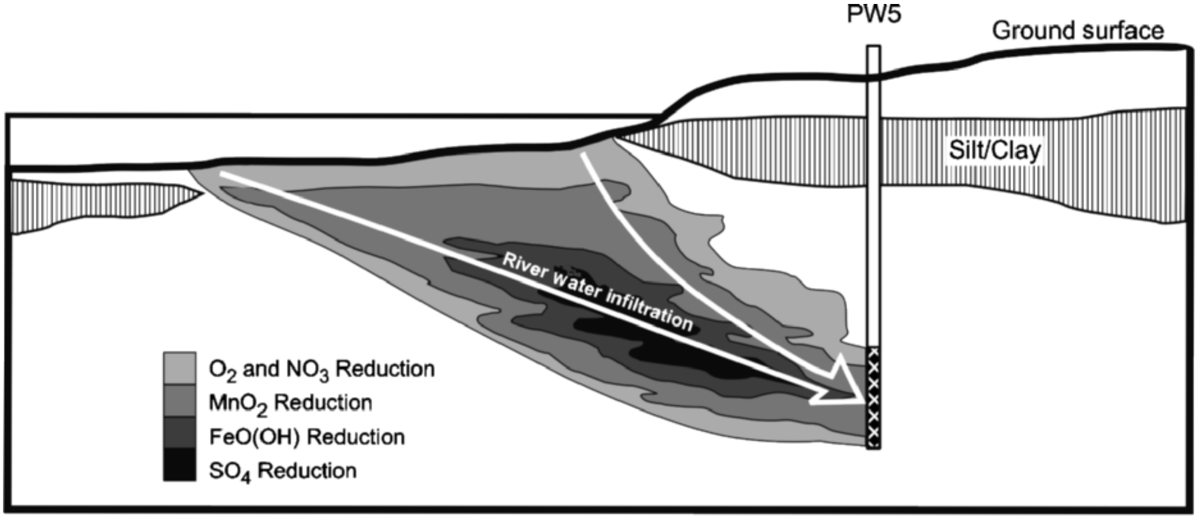

O_2 Reduction
$CH_2O + O_2 => HCO_3^- + H^+$

NO_3 Reduction
$5CH_2O + 4NO_3^- + 4H^+ => 5HCO_3^- + 2N_2 + 2H_2O + 5H^+$

MnO_2 Reduction
$^1/_2CH_2O + MnO_{2(s)} + {}^1/_2H^+ => Mn^{(II)} + {}^1/_2HCO_3^- + OH^-$

FeO(OH) Reduction
$^1/_4CH_2O + FeO(OH)_{(s)} + {}^7/_4H^+ => Fe^{(II)} + {}^1/_4HCO_3^- + {}^3/_2H_2O$

SO_4 Reduction
$2CH_2O + SO_4^{2-} + H^+ => 2HCO_3^- + HS^- + 2H^+$

Figure 10. Thermodynamically based conceptual model for the redox zonation that is expected to develop in groundwater along the infiltration flow path from the river to wells. In these reactions, CH_2O is used to represent labile DOC.

velocities, and the aquifer sediments under investigation have a much larger "bed" thickness (>20 m) than engineered systems. These characteristics suggest that, in this case, riverbank filtration is conducive to providing pretreatment of the drinking water supply. However, it has also been noted (26) that cold water temperatures (near 1 °C) inhibit biological processes such as the production of natural polymers that are important for micro-organism removal during slow-sand filtration. We have observed such cold temperatures to significant depths in this case study, and therefore it might be expected that micro-organism removal during riverbank filtration will also be seasonally variable.

Groundwater Geochemistry

In order to improve the knowledge of the long-term geochemical evolution of groundwater from wells adjacent to the river, hydrogeochemical investigations were designed to provide an understanding of the influence of the DOC flux from the river on the redox conditions within the aquifer. The approach was to monitor the geochemical conditions along a flow path from the river toward supply well PW5 (Fig. 10). The monitoring was intended to identify spatial and temporal (seasonal) variations in the concentrations of redox-sensitive solutes such as dissolved O_2 (DO), NO_3, Mn, Fe, and SO_4. Aqueous geochemical sampling was conducted during the spring, summer, and fall at piezometers installed in the aquifer below the river bed (Fig. 7).

Thermodynamic principles suggest that the infiltration of DOC to an aquifer system would result in a sequential or stepwise decline in redox conditions, with reactions that provide the greatest energy advantage to microbial catalysts proceeding first, followed by reactions that provide less energy. The general sequence that is expected (30,31) would be the oxidation of DOC coupled sequentially to the reduction of O_2, NO_3, Mn(IV) oxide, Fe(III) oxide, SO_4, and finally CO_2. The O_2, NO_3, and SO_4 would generally be supplied by the infiltrating river water, and the Mn(IV) and Fe(III) are present as oxide or oxyhydroxide minerals in the aquifer sediment, with Fe(III) minerals usually in greater abundance. In accordance with thermodynamic principles, a conceptual model for the redox zonation that may be expected to develop along the flow path during river bank filtration is presented in Fig. 10.

Seasonal Geochemical Variations. The DOC concentration in the Saint John River ranges from 7.9 mg/L in July to 10.9 mg/L in September. Along the infiltration pathway, a general decrease occurs in DOC concentration with depth and distance toward the well, which is greatest (30–50%) in August and September. The concentrations of DO, NO_3,

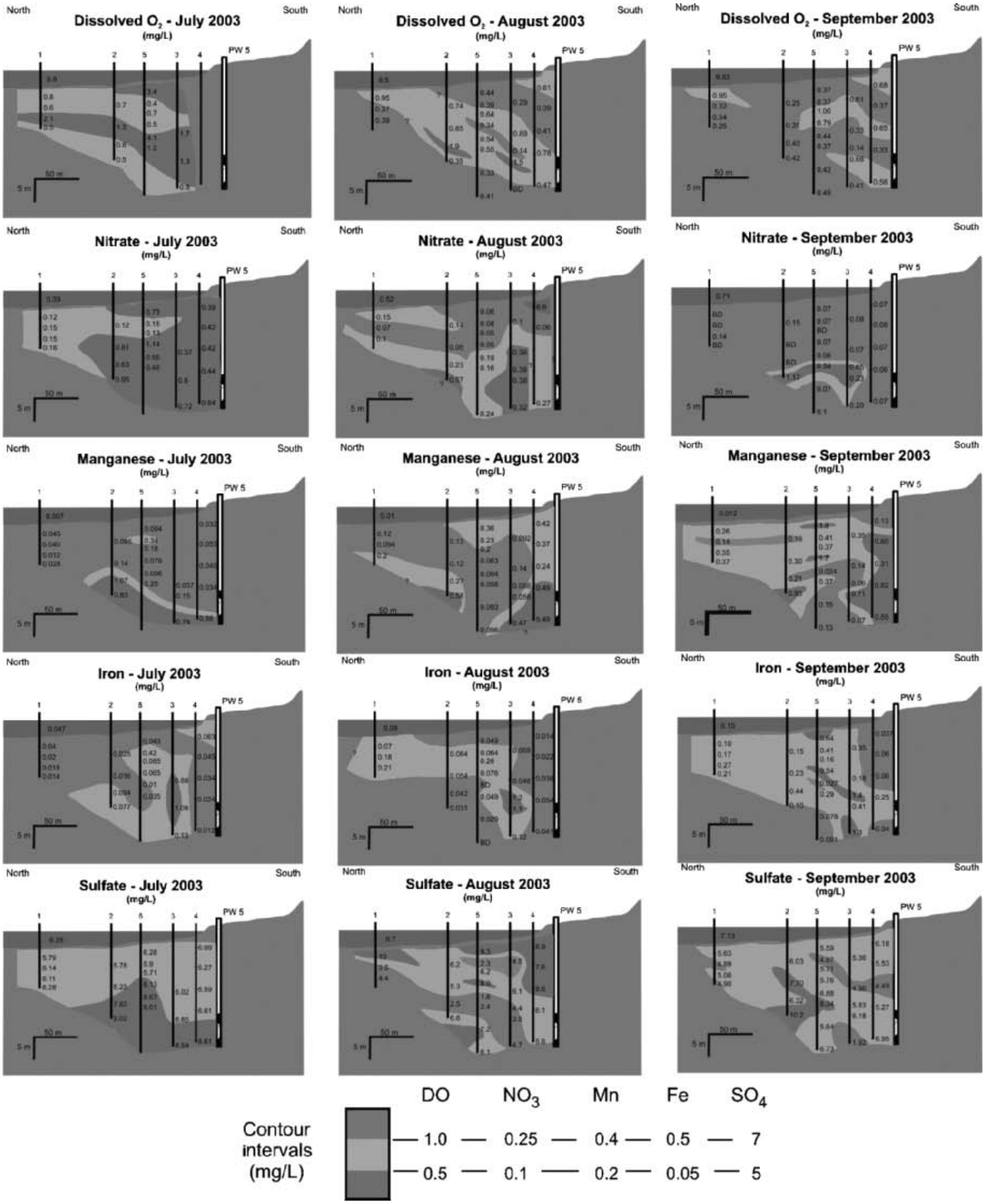

Figure 11. Seasonal (left to right) and spatial variations in redox-sensitive geochemical parameters along the flow path between the Saint John River and City of Fredericton municipal well PW5. Piezometer nests are indicated by vertical lines and individual piezometers (sample locations) are indicated by points on the lines.

Mn, Fe, and, to some extent, SO_4 display significant seasonal variations (Fig. 11). The redox conditions cycle on a seasonal time scale, likely in response to temperature changes in the infiltrating river water. Through the spring and into midsummer, the aquifer temperature is generally less than 5 °C (Fig. 8). The cold temperatures inhibit microbiologically mediated redox reactions between DOC and electron acceptors such as O_2, NO_3, Mn(IV), and Fe(III). As a result, conditions are relatively oxidizing with elevated concentrations of dissolved O_2 and NO_3, and low concentrations of Mn and Fe. Toward late summer and into the fall, the aquifer temperature increases and the system tends toward more reducing conditions, with concentrations of dissolved O_2 and NO_3 declining, and concentrations of Mn and Fe increasing. Localized zones of elevated HS concentrations suggest that SO_4 reduction occurs; however, the seasonal trend toward reducing conditions is not manifest by a widespread decline in SO_4 concentrations as it is for O_2 and NO_3 (Fig. 11).

Spatial Geochemical Variations. Although the decline in dissolved O_2 and NO_3 concentrations and increase in the concentrations of Mn and Fe toward late summer and fall suggests that the reactions noted in Fig. 10 occur, the substantial overlap of the reactions in space that is observed (Fig. 11) indicates that they do not proceed sequentially as suggested by equilibrium thermodynamics, on the contrary, the data suggest that the system is far from equilibrium, which is likely attributable to a combination of variable reaction kinetics for the respective reactions (probably related to mineralogy and microbiological abundance and activity) and a heterogeneous grain-size or permeability distribution that may lead to the development of localized redox gradients at a scale that can not be resolved by the sampling.

With sustained pumping over several decades, the flux of DOC from the river will lead to depletion of the Mn(IV) and Fe(III) oxide minerals in the sediment. The pattern of depletion will be a function of the initial mineral abundance and distribution, the respective rates of oxide-mineral reductive dissolution, and the spatial variations in groundwater flux within the aquifer. However, over time, there should be a general trend toward depletion of Mn-oxides, followed by depletion of Fe oxides. The distribution of dissolved Mn and Fe in August 2003 is consistent with this expectation in that Mn concentrations are highest in the down-gradient region closest to the well, suggesting that Mn-oxide minerals have been depleted along an interface that migrates toward the well with time. The trend in concentrations of dissolved Mn in the water extracted from well PW5 (Fig. 12) is also consistent with the expectation that Mn-oxide mineral depletion will occur in the aquifer sediments.

SUMMARY

The Fredericton Aquifer is a typical semiconfined river-valley aquifer composed of glaciofluvial sand and gravel that is recharged by infiltration of water from a large river. Geophysical surveys have revealed a hydraulic connection to the Saint John River through an irregular-shaped window zone where the aquitard overlying the aquifer has been removed by erosion. Continuous profiling, using electromagnetic terrain conductivity meters deployed in small boats, proved to be the most effective way to delineate the window zone. The river water's moderately low conductivity (10 mS/m) and the relatively shallow depth to the aquitard, contributed to the method's high spatial resolution. Single channel seismic reflection profiles and a shoreline resistivity imaging survey provided important stratigraphic information and evidence to support the electromagnetic interpretation.

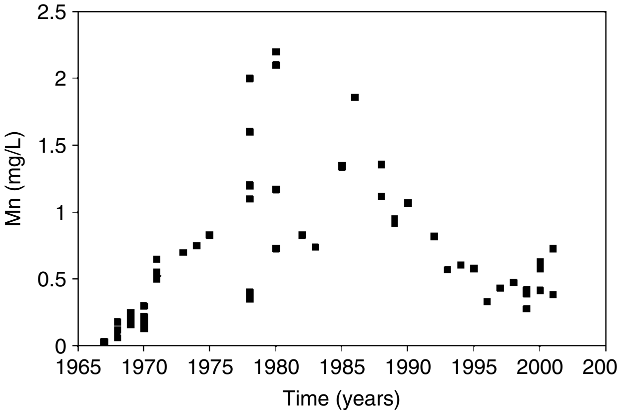

Figure 12. Manganese concentrations in the groundwater extracted from municipal well PW5.

Two natural tracers, water temperature and ^{222}Rn, were used to determine downward flow velocities within the aquitard window located beneath the riverbed. Temperature data was collected continuously to determine seasonal variations and interpreted using heat transport modeling. ^{222}Rn activities for water samples from piezometers nests gave estimates of residence time, which were then used to compute infiltration velocities. The velocities derived from heat transport modeling and ^{222}Rn activities compared well with vertical average linear velocities ranging from 0.1 m/day to 5 m/day. Although these velocities are relatively high for subsurface flow, they are lower than typical velocities in engineered slow-sand filtration systems. As such, it is expected that filtration in the sediments located between the riverbed and production wells will be effective in removing micro-organisms. However, the seasonal occurrence of cold water temperatures in the sediments may cause a lower removal of micro-organisms during certain periods of the year.

The groundwater velocities and seasonal temperature variations measured in the aquifer below the river have a significant influence on the redox reactions that control the concentrations of DO, NO_3, Mn, Fe, and SO_4. The high velocities result in a large flux of reactants (DOC, DO, NO_3, and SO_4), thereby minimizing the potential that reaction rates will be reactant limited. The seasonal temperature fluctuations in the groundwater influence the reaction kinetics for chemical and microbiologically mediated reactions. The seasonal variations in temperature are considered the controlling influence on the aquifer redox conditions, with elevated temperature in late summer and fall contributing to increased reaction rates and a trend toward more reducing conditions. The data suggest that reductive dissolution of Mn oxide minerals as a consequence of sustained pumping over the past 40 years has resulted in the near depletion of Mn oxides along the infiltration flow path to well PW5.

This case study demonstrates the use of near-surface geophysical methods to define stratigraphic complexity in three dimensions, which controls spatial variability of the recharge flux. The geophysical results were especially useful for optimizing the layout of multilevel piezometer nests and mini-piezometers used in the subsequent hydrogeological and hydrogeochemical investigations. The use of the natural tracers, temperature and ^{222}Rn, provides important information about the travel time between the river and the wells. This information allows for an assessment of risk associated with possible contamination events in the river. In addition, the temperature data are important in the interpretation of seasonal variations in groundwater chemistry. In general, riverbank filtration provides a reliable source of recharge to the wellfield in Fredericton. The filtration velocities in the aquifer sediments in the vicinity of PW5 are comparable with those in engineered slow-sand filtration systems. The principle disadvantage of riverbank filtration in this case study is elevated Mn concentrations that are related to the modification of redox conditions in the aquifer as a result of the infiltration of DOC from the river.

Acknowledgments

The authors wish to acknowledge the financial support from the Canadian Natural Sciences and Engineering Research Council (NSERC) through the Canadian Water Network (CWN), the City of Fredericton, and the New Brunswick Department of Environment and Local Government. In addition, the Geological Survey of Canada (Atlantic) provided equipment and personnel support for riverine seismic surveys.

BIBLIOGRAPHY

1. Kuehn, W. and Mueller, U. (2000). Riverbank filtration, an overview. *J. Amer. Water Works Assoc.* **92**: 60–69.
2. Ray, C., Soong, T.W., Lian, Y.Q., and Roadcap, G.S. (2002). Effect of flood-induced chemical load on filtrate quality at bank filtration sites. *J. Hydrol.* **266**: 235–258.
3. Malzer, H.-J., Gerlach, M., and Gimbel, R. (2000). Transport and fate of organic compounds during riverbank filtration. In: *Proceedings of the International Riverbank Filtration Conference, November 2000*. W. Jülich and J. Schubert (Eds.). IAWR, Amsterdam, the Netherlands, pp. 117–126.
4. Rismal, M. and Kopac, I. (2000). Riverbank filtration as pre-treatment of river water for artificial recharge of groundwater in Slovenia. In: *Proceedings of the International Riverbank Filtration Conference, November 2000*. W. Jülich and J. Schubert (Eds.). IAWR, Amsterdam, the Netherlands, pp. 183–196.
5. Duncan, D., Pederson, D.T., Shepherd, T.R., and Carr, J.D. (1991). Atrazine used as a tracer of induced recharge. *Ground Water Mon. Rev.* **11**: 144–150.
6. Bourg, A.C.M., Darmendrail, D., and Ricour, J. (1989). Geochemical filtration of riverbank and migration of heavy metals between the Deûle River and the Ansereuilles Alluvion-Chalk Aquifer (Nord, France). *Geoderma*. **44**: 229–244.
7. Waber, U.E., Lienert, C., and von Gunten, H.R. (1990). Colloid-related infiltration of trace metals from a river to shallow groundwater. *J. Contam. Hydrol.* **6**: 251–265.
8. Medema, G.J., Juhasz-Holterman, M., and Luijten, J.A. (2000). Removal of micro-organisms by bank filtration in a gravel-sand soil. In: *Proceedings of the International Riverbank Filtration Conference, November 2000*. W. Jülich and J. Schubert (Eds.). IAWR, Amsterdam, the Netherlands, pp. 161–168.
9. Wang, J.Z., Song, R., and Hubbs, S.A. (2000). Particle removal through riverbank filtration process. In: *Proceedings of the International Riverbank Filtration Conference, November 2000*. W. Jülich and J. Schubert (Eds.). IAWR, Amsterdam, the Netherlands, pp. 127–137.
10. Hiscock, K.M. and Grischek, T. (2002). Attenuation of groundwater pollution by bank filtration. *J. Hydrol.* **266**: 139–144.
11. Ray, C. (2002). Effect of biogeochemical, hydrogeological, and well construction factors on riverbank filtrate quality. In: *Riverbank Filtration: Understanding Contaminant Biogeochemistry and Pathogen Removal*. C. Ray (Ed.). Kluwer Academic Publishers, the Netherlands, pp. 1–16.
12. Jacobs, L.A., von Gunten, H.R., and Keil, R. (1988). Geochemical changes along a river-groundwater infiltration flow path: Glattfelden, Switzerland. *Geochim. Cosmochim. Acta*. **52**: 2693–2706.
13. Bourg, A.C.M. and Bertin, C. (1993). Biogeochemical processes during the infiltration of river water into an alluvial aquifer. *Environ. Sci. Technol.* **27**: 661–666.
14. Doussan, C., Ledoux, E., and Detay, M. (1998). River-groundwater exchanges, bank filtration, and groundwater quality: ammonium behaviour. *J. Environ. Qual.* **27**: 1418–1427.
15. Grischek, T., Worch, E., and Nestler, W. (2000). Is bank filtration under anoxic conditions feasible?. In: *Proceedings of the International Riverbank Filtration Conference, November 2000*. W. Jülich and J. Schubert (Eds.). IAWR, Amsterdam, the Netherlands, pp. 57–66.
16. Sacher, F., Brauch, H.-J., and Kuhn, W. (2000). Fate studies of hydrophyllic organic micro-pollutants in riverbank filtration. In: *Proceedings of the International Riverbank Filtration Conference, November 2000*. W. Jülich and J. Schubert (Eds.). IAWR, Amsterdam, the Netherlands, pp. 139–148.
17. von Gunten, H.R. et al. (1991). Seasonal biogeochemical cycles in riverborne groundwater. *Geochim. Cosmochim. Acta*. **55**: 3597–3609.
18. Bourg, A.C.M. and Bertin, C. (1994). Seasonal and spatial trends in manganese solubility in an alluvial aquifer. *Environ. Sci. Technol.* **28**: 868–876.
19. Thomas, N.E., Kan, K.T., Bray, D.I., and MacQuarrie, K.T.B. (1994). Temporal changes in manganese concentrations in water from the Fredericton Aquifer, New Brunswick. *Ground Water* **32**: 650–656.
20. Violette, G.G. (1990). *The application of a computer model to assess the potential effects of groundwater withdrawals from the Fredericton Aquifer*. Fredericton, N.B., MScE Thesis, University of New Brunswick, Department of Civil Engineering, Fredericton, NB, Canada.
21. TerrAtlantic Engineering Limited. (2000). *Development of the Groundwater Resource*. Fredericton, NB—status report: Unpublished report prepared for the City of Fredericton.
22. Petrunic, B.M., MacQuarrie, K.T.B., and Al, T.A. (2004). Reductive dissolution of Mn oxides in river-recharged aquifers: a laboratory column study. *J. Hydrol.* In press.
23. Butler, K.E., Nadeau, J.-C., Parrott, R., and Daigle, A. (2004). Delineating recharge to a river valley aquifer by riverine seismic and EM methods. *J. Env. Eng. Geophys.* **9**: 95–109.
24. McNeill, J.D. (1980). *Technical note TN-6, Electromagnetic Terrain Conductivity Measurements at Low Induction Numbers: Geonics Limited*. Mississauga, Ontario, Canada.
25. Nadeau, J.-C., Butler, K.E., Hunter, J.A., Burns, R., Douma, M., Good, R., and Veillette, N. (2004). Reflection and

Downhole Seismic Studies of a Glaciofluvial River Valley Aquifer. In: *Proceedings of 17th Symposium on Application of Geophysics to Engineering and Environmental Problems*. EEGS, pp. 1408–1419.

26. Logsdon, G.S., Kohne, R., Solomon, A., and LaBlonde, S. (2002). Slow sand filtration for small water systems. *J. Envir. Eng. Sc.* **1**: 339–348.
27. Lapham, W.W. (1989). *Use of temperature profiles beneath streams to determine rates of vertical ground-water flow and vertical hydraulic conductivity*. US Geol. Survey Water Supply Paper 2337.
28. Diersch, H.-J.G. (2004). *WASY Software FEFLOW® 5.1 Finite Element Subsurface Flow & Transport Simulation System User's Manual*. WASY, Berlin, Germany.
29. Hoehn, E. and von Gunten, H.R. (1989). Radon in groundwater: a tool to assess infiltration from surface waters to aquifers. *Water Resour. Res.* **25**: 1795–1803.
30. Champ, D.R., Gulens, J., and Jackson, R.E. (1979). Oxidation-reduction sequences in groundwater flow systems. *Can. J. Earth Sci.* **16**: 12–23.
31. Stumm, W. and Morgan, J.J. (1981). *Aquatic Chemistry*. John Wiley & Sons, New York.

Index

NOTE: Page numbers followed by f refer to figures, page numbers followed by t refer to tables. All volume numbers appearing in italics refer to volume subject material as stated in footnote.

Index terms **Links**

A

A2/O process *1:*790
Abandoned channel pools *3:*69
Abandoned mines, alkaline amendment to *5:*3
ABA plant hormone *3:*718
Abiotic redox reactions *5:*413
Abiotic solute transport *5:*527
Abiotic transformation, of chlorinated aliphatic compounds *1:*691
Abrasion number, of carbon *1:*97
Absolute humidity *4:*269
Absolute ownership doctrine *4:*630
Absorbed solar radiation *3:*375
Absorption, of organochlorine pesticides *3:*643
Absorption systems, for odor removal *1:*914
Acanthamoeba *1:*278
Acaricide Dicofol *3:*645
Accelerator mass spectrometry (AMS) technique *4:*418
Acceptable daily intake (ADI) *3:*637
Accumulation, of organochlorine pesticides *3:*643
  See also Bioaccumulation
Acequia, defined *2:*541
Acetic acid *2:*295
Acetogenesis, in landfills *1:*696
Acid, carbonate scale control with *4:*415
Acid deposition *4:*377
  control of *4:*380
  effects of *4:*378
Acid Deposition Control Program (Title IV) *4:*379 *4:*380
Acidic brines *3:*10
Acidic crater lakes *3:*9
Acidic mining lakes *3:*10
Acidic pit lakes *3:*10 *3:*11

Index terms **Links**

Acidic rivers *3:*10
Acidic tailings, from geological deposits *3:*9
Acidification. *See also* Acidization
  atmospheric *3:*8
  biological effects of *3:*9
  causes of *3:*5
  chronic *3:*1
  episodic *3:*5
  of freshwater resources *3:*7
  geogenic *3:*9
  in Indian lakes *3:*451
  industrial *3:*9
  soil *5:*377
Acidified soft waters *3:*8
Acidifying substances, life cycle emissions of *3:*548
Acidity
  negative and positive *2:*424
  relationship to pH *3:*8
  soil *3:*707
Acidization *5:*28
  See also Acidification
Acid mine drainage (AMD) *1:*609 *2:*1 *3:*1 *3:*2 *3:*10
  active and passive treatment of *3:*11
  alkaline amendment to *5:*2
  chemical treatment of *5:*3
  control of *5:*2
  costs of treating *5:*4
  geochemistry of *3:*13
  iron-oxidizing bacteria in *2:*150
  overburden analyses in *5:*1
  passive treatment of *2:*423 *5:*5
  remediation of *3:*11
  in the United States *5:*1
Acid mine wastewater *1:*113
Acid-neutralizing capacity (ANC) *3:*1 *3:*3
Acidogenesis, separation from methanogenesis *1:*909
Acidophilic iron bacteria *2:*149
Acid oxidation *2:*85
Acid phosphatase (ACP) activity *2:*415
Acid production rate *3:*14

Index terms **Links**

Acid rain *3:*1 *3:*8 *3:*292 *4:*184
  assessment of *4:*380
  countermeasures against *3:*9
  current conditions related to *4:*380
  defined *4:*377
  effects of *3:*224
  groundwater acidification due to *3:*10
  regulation of *4:*379
  social response to *4:*379
  sources and effects of *4:*378
Acids
  defined *2:*542
  of geogenic origin *3:*7
  strength of *2:*294
Acid treatment, for scale control *1:*548
Acinetobacter *2:*241
A climates *4:*256
Acoustical monitoring, underwater *4:*91
Acoustic televiewer *5:*154
Acre-foot, defined *2:*542
Action level *1:*479
Activated alumina (AA) *1:*460
  for arsenate removal *5:*20
Activated alumina sorption, for arsenic removal *1:*638
Activated carbon(s) *2:*79 *4:*507
  See also Activated carbon adsorption; Powdered activated carbon (PAC)
  acid surface functional groups on *2:*81
  adsorption capacity of *4:*381
  applications of *1:*101
  background of *2:*80
  common uses for *1:*105
  competitive adsorption on *1:*107
  competitive sorption of inorganics on *1:*110
  domestic use of *1:*105
  gas-phase applications of *1:*104
  methods for developing *1:*93
  organic micropollutant and metal ion removal by *4:*506
  physical and chemical properties of *2:*81
  precursors of *2:*80
  preparation and properties of *1:*94t *2:*82
  sorption of trace metals onto *2:*83

Index terms **Links**

Activated carbon(s) *(Continued)*
  steps in producing *1:*95f
  surface chemistry of *1:*98
  surface functional groups on *2:*82
  types of *1:*92
Activated carbon (AC) adsorption *1:*817
  domestic sewage and *1:*834
  of landfill leachates *1:*707
Activated carbon posttreatment *5:*25 *5:*327
Activated carbon samplers *5:*109
Activated carbon technology *1:*351
Activated carbon treatment, of radioactive waste *1:*804
Activated sludge *1:*566 *1:*568 *1:*861 *1:*862
  See also Activated sludge process(es)
  bacteria in *1:*668t
  floc from *1:*844
Activated Sludge Models (ASMs) *1:*733 *1:*739
  See also International Water Association (IWA) models
  variables in *1:*740t
Activated sludge plants, microbial foaming and bulking in *1:*844
Activated sludge process(es) *1:*565 *1:*668 *1:*815
  biological phosphorus removal in *1:*788
  denitrification in *1:*667
  domestic sewage and *1:*832
  microbial foaming in *1:*728
  nitrification in *1:*751
  nutrient deficiencies in *1:*729
Activated sludge treatment, POP behavior in *1:*767
Activation *2:*80
  See also Activated carbon entries; Activated sludge entries
  of carbon *1:*94
Active biomonitoring (ABM) *2:*33
  advantages and disadvantages of *2:*33
  system for *2:*28
Active direct solar water heaters *1:*65
Active disease reporting *1:*185
Active immobilization, of metals *5:*282
Active indirect solar water heaters *1:*65
Active isotopes, as groundwater tracers *5:*503
Active microwaves *4:*324

Index terms **Links**

Active real-time control, modeling *3:*340
Active rock glaciers *3:*174 *3:*175
Active sound detection *4:*570
Active storage *3:*259
Acute sediment toxicity tests *2:*351 *2:*385t *3:*51
Acute toxicity, categories of *2:*415
Adaptive management *3:*128
Adaptive modified sequential method, of DNAPL migration modeling *5:*669
Additives, odor-reducing *1:*912
Adenosine triphosphate (ATP) *1:*399 *2:*42
  detection of *1:*85
Adenovirus *1:*279
Adenylate kinase (AK) detection *1:*85
"Adequate protection by natural means" *1:*293
"Adequate protection by treatment" *1:*293
Adiabatic cooling *4:*366
Adiabatic ideal gas process *4:*368
Adsorbent dose (W_s) *3:*299
Adsorbents
  characterization of *4:*508
  for inorganic ion removal *4:*494
  micropores and mesopores in *2:*79
Adsorbent surface chemistry *4:*491
Adsorption *1:*811 *2:*374 *5:*184 *5:*186
  See also Inorganic adsorption
  for arsenic removal *1:*638
  dynamics of *3:*302
  effect of pH on *3:*299
  of hydrocarbons *1:*577
  in situ metal/organic *5:*609
  kinetic formulations for *5:*415
  in mine effluent remediation *1:*611
  models of *3:*300
  for odor abatement *1:*762
  operating variables in *3:*298
  of organic compounds *4:*384
  thermodynamics and *3:*301
  water-rock interaction and *5:*566

Index terms **Links**

- Adsorption capacity
  - correlation for *4:*382
  - graphs of *4:*382 *4:*383f
- Adsorption clarifier package plants *1:*516
- Adsorption equilibrium, of carbon *1:*99
- Adsorption isotherms *3:*299 *4:*384
  - types of *1:*99
- Adsorption models, multicomponent *1:*107
- Adsorption systems, for odor removal *1:*914
- Adsorptive colloidal processes *5:*350
- Adsorptive columns, sizing *2:*363
- Adsorptive filtration, for arsenic removal *1:*639
- Advanced impellers *1:*79f
- Advanced Microwave Scanning Radiometer (AMSR-E) *2:*589 *4:*194 *4:*216
- Advanced Microwave Sounding Unit (AMSU, AMSU-A) *4:*216 *4:*351
- Advanced nucleic acid analyzer (ANNA) *2:*339
- Advanced oxidation processes (AOP) *2:*318
  - for vinyl chloride *5:*639
- Advanced oxidation technology *1:*579
- Advanced wastewater treatment techniques *1:*816 *1:*871
  - *See also* Wastewater entries
  - advanced oxidation processes (AOP) *1:*871
  - complexation/flocculation *1:*872
  - conducting polymers *1:*872
  - ionizing irradiation *1:*873
  - membrane/sonication/wet oxidation *1:*873
  - sorption by zeolites *1:*874
  - supercritical water oxidation *1:*874
  - ultrasonic irradiation *1:*875
- Advection *5:*273 *5:*278
  - organic vapor transport via *5:*544
  - transport by *5:*516
- Advection fog *4:*232
  - variables of *4:*236
- Advection phenomena *3:*78
- Advective-dispersion model *3:*332
- Adverse effects assessments *2:*516
- Adverse social impacts, of water transfers *4:*687
- Advisors, international *1:*719
- Aeolian valleys *3:*66

Index terms **Links**

- Aerated lagoons *1:*704
- Aerated static piles *1:*648
- Aeration *1:*380 *1:*460 *2:*374
  - in domestic sewage treatment *1:*833
  - extended *1:*815
  - groundwater remediation by *5:*426
  - methods of *1:*51
  - pretreatment *1:*323
  - tapered *1:*815
  - technology for *1:*352 *4:*546
  - wastewater *1:*623
- Aeration efficiency (AE) *1:*623
- Aeration systems *1:*623
  - submerged *1:*624
  - surface *1:*623
- Aerial natural gamma radiation surveys *4:*338
- Aerobic bacteria *2:*20 *5:*581
  - in cooling systems *2:*241
- Aerobic biodegradation, optimum *5:*44
  - *See also* Aerobic degradation; Biodegradation
- Aerobic bioremediation *5:*424
  - technologies *5:*40
- Aerobic cometabolism, methyl *tertiary*-butyl ether *5:*388
- Aerobic degradation, rate of *5:*44
  - *See also* Aerobic biodegradation
- Aerobic digestion
  - of biosolids *1:*647
  - of sludge *1:*865
- Aerobic *in situ* enhanced bioremediation *5:*423
- Aerobic leachate treatment *1:*704 *1:*709
- Aerobic plate count (APC), in dishwater *2:*113 *2:*114
- Aerobic stabilization, of sludge *1:*857
- Aerobic wastewaters
  - odorants in *1:*911
  - treatment of *1:*813
- Aerobic wetlands *5:*5
- *Aeromonas* *1:*277
- Aerosols, atmospheric *4:*371
- Aesthetics, algae-impaired *3:*108

Index terms **Links**

Africa
  unsafe water in *2:*631
  monsoons in *4:*304
African foods, nitrate levels in *1:*32t
Age dating, of groundwater *4:*388
Agency management impacts, of water transfers *4:*687
Agenda 21 *3:*33 *3:*34
Aggressive metal mobilization *5:*282
Aging
  impact of water on *4:*455
  premature *4:*455
Agitation, in water treatment systems *1:*76
Agrichemical facilities, soil remediation at *3:*651
  See also Agrochemicals
Agricultural activities. *See also* Agricultural practices
  contamination from *1:*25
  as a nitrate contamination source *2:*571
Agricultural area pollution
  mathematical models for estimating *3:*658
  outflow of *3:*657
Agricultural areas, impact on groundwater quality *5:*251
Agricultural changes, influence on the hydrologic cycle *4:*281
Agricultural districts law *3:*597
Agricultural drainage *5:*94
Agricultural drainage ditches, research conducted on *3:*88 *3:*90t *3:*90t
Agricultural droughts *4:*208
Agricultural impacts, of waterlogging *3:*741
Agricultural irrigation *1:*818
Agricultural land, classification and mapping of *3:*608
Agricultural landscape
  pollutants in *3:*605t
  water quality management in *3:*604
Agricultural land use planning *3:*595
  issues in *3:*595
  national, state, and local *3:*596
Agricultural Nonpoint Source Pollution Software (AGNPS) *2:*325
  Model 2001 (AGNPS 2001) *2:*251
Agricultural practices. *See also* Agricultural activities
  algal control via *2:*3
  improving *2:*571

Index terms **Links**

Agricultural runoff, as a source of cultural eutrophication *3:*114
Agricultural soils, phosphorus status of *3:*704
Agricultural wastewater, constructed wetlands for *1:*896
Agricultural water
  Arab World share of *2:*474f
  curious uses of *4:*695
  United States use of *2:*653
Agricultural watershed, organochlorine pesticides in vegetables grown on *3:*643
Agricultural water use efficiency (WUE) *3:*558
Agriculture. *See also* Crop entries; Farming; Irrigated agriculture
  in the Arab World *2:*470
  in China *2:*488
  frost damage to *4:*242
  groundwater role in *2:*392
  of hollows *4:*695
  irrigated *2:*489
  nonpoint source pollution from *2:*186
  organic *3:*644
  raised-style *4:*697
  relation to water security *1:*438
  sediment from *3:*509
  wastewater reuse guidelines for *3:*670
  water conservation in *2:*495
  as a water consumer *2:*489 *2:*662
  water reuse in *2:*635
  water supply for *2:*617
Agrochemicals, role of *3:*643
  See also Agrichemical facilities
Agroforestry systems *3:*551 *3:*568
Agronomic/ecological models, use of field capacity in *5:*126
Agronomic practices, phytoextraction optimization using *5:*372
Agronomic soil conservation methods *3:*551
AIDS *1:*162 *1:*163 *1:*182 *1:*258
Air, cadmium in *5:*617
  See also Aeration; Aerobic entries
Air aeration systems, fine bubble diffused *1:*626
Airborne contaminant sources *5:*433
Air bubbling *1:*594
Air chambers, water hammer and *1:*262
"Air dry" soil *5:*536

Index terms **Links**

Air emissions *1:*413
  electric generating plant contaminants in *1:*554
Air entrapment effects *5:*536
Air flotation *3:*407
Air gaps *1:*155
Air-line submergence method *5:*101
Air oxidation *2:*85
Air pollution *4:*379
  chemicals in *3:*224
  effects of *4:*378
  MtBE and *5:*319
Air quality *1:*554
Air sampling techniques *2:*307
AIRS/AMSU/HSB sounding system *4:*351 *4:*352
Air-sea interaction *4:*1
Air sparging *5:*39 *5:*428 *5:*440
  See also Biosparging
Air sparring, activated carbon in *1:*105
Air stripping *1:*380 *1:*576 *1:*631 *1:*816
  of landfill leachates *1:*707
  for radioactive waste *1:*803
  of vinyl chloride *5:*639
Air temperature, determining *3:*563
Air to plant absorption *3:*643
Air-water interface *4:*452 *4:*470
  GHG fluxes at *3:*205
Alaska, in-stream flow protection in *4:*662
  See also Glacier Bay
Alaskan rain forest *2:*667
Albedo *4:*179
Alfalfa channel tests *5:*392
Algae *2:*269
  See also Blue-green algae entries; Periphyton index of biotic integrity (Periphyton IBI); Toxic algae
  acid-sensitive *3:*4
  in cooling water systems *1:*540
  excessive growths of *3:*108
  health effects of *1:*277
  removal using ozone *1:*355
Algal bioremediation *2:*44

Algal blooms *3:*114 *4:*108
  decomposition of *3:*115
  harmful *4:*55
Algal growth *2:*241
  importance of light penetration to *3:*111
Algal nutrients, sources and control of *3:*110
Algal populations
  biomanipulation of *2:*6
  chemical control of *2:*5
  environmental control of *2:*3
Algal toxins *2:*387
Alice Water Treatment Plant (AWTP) *1:*221
  alum and lime consumption and cost estimation for *1:*222
  bacterial monitoring at *1:*222 *1:*225
  clarifier and filter performance at *1:*222 *1:*224
  disinfection at *1:*222 *1:*224
  DOC and BDOC levels at *1:*225
  flow rate measurement and control at *1:*222
  performance of *1:*222 *1:*225
Alicyclic hydrocarbons, degradation of *1:*692
Alien species *3:*134
Alignment log *5:*154
Aliphatic hydrocarbons
  chlorinated *1:*690
  degradation of *1:*692
Alkaline amendment
  to abandoned mines *5:*3
  to acid mine drainage *5:*2
Alkaline leach beds, in acid mine treatment *5:*7
Alkaline phosphatase (APase) *2:*231
Alkaline recharge trenches *5:*2
Alkaline stabilization
  of biosolids *1:*649
  of sludge *1:*865
Alkalinity
  in carbonate geochemistry *4:*409
  defined *2:*542
  effect on corrosion *1:*8
  imparted by carbonaceous minerals *3:*14
  insufficient *5:*2
  in natural waters *2:*394

Index terms **Links**

Alkalinity *(Continued)*
  nitrification and *1:*754
  in ozone half-life *1:*359
  role in fish growth and production *3:*131
  of soil *3:*707
  of water *1:*902
Alkalinity addition, chemicals suitable for *1:*755t
Alkaloid cyanobacterial toxins, structures of *3:*190f
Alkene biodegradation *1:*692
Alkenones *4:*93
Allergies, dehydration and *4:*725
Allochthonous inputs, watershed-related *3:*207
Allogenic recharge *5:*236
Allowable exposures *1:*427
Allowable headwater (*HW*) *3:*77
Allowable velocity *3:*77
Alluvial valleys *3:*66
Alluvium, defined *2:*542
Alpha emitters, health risks of *1:*396
Alpine climatic zone *4:*258
Alpine hydrochemical model (AHM) *5:*298
Alternative collecting systems *1:*840
Alternative cost approach *4:*609
Alternative formulation, in a systems approach *2:*686
Alternative Maximum Contaminant Level (AMCL) *1:*479
Alternative stable states theory *3:*272
Alternative wastewater treatment, systems *1:*840
Altitude effect *4:*439
Alum *1:*138
Alum coprecipitation *5:*20
Alumina, activated *5:*20
Aluminum (Al)
  coagulation using *1:*636
  episodic acidification and *3:*7
  toxicity of *3:*131
Aluminum industry *1:*562
Aluminum recovery, from water treatment plant residuals *1:*140f
Amargosa Desert, recharge in *5:*74
Amazonia region *4:*642
Ambient Groundwater Monitoring and Assessment Program (GAMA) studies *5:*62

Index terms **Links**

Ambient groundwater monitoring network *5:*313
  monitoring frequency at *5:*317
  monitoring measurements and samples for *5:*316
  site inventory for *5:*315
  site selection for *5:*315
  size of *5:*316
  strategies and designs for *5:*314
  types of wells for *5:*316t
  using data from *5:*317
Ambient water quality, optimal level of *2:*129
AMD generation, rate of *3:*13
  See also Acid mine drainage (AMD)
American Farmland Trust (AFT) *3:*598
American Meteorological Society *4:*329
American National Standards Institute (ANSI) *1:*381
  standards of *1:*198
American Rule *4:*631
American Society of Testing and Materials (ASTM) *2:*459
  Soil Classification Flow Chart *3:*690
American Society of Agricultural Engineers (ASAE) microirrigation standards *3:*615
American Society of Civil Engineers - Environmental Water Resources Institute (ASCE-EWRI) *3:*573
American water resource development *2:*498
American Water Works Association (AWWA) *1:*146 *1:*285 *1:*437
  standards of *1:*261 *1:*403
  survey *2:*91
Americas, transboundary water treaties in *4:*637 *4:*641t
  See also Latin America; North American Free Trade Agreement (NAFTA)
Americium (Am). *See* Am-humic colloids
Am-humic colloids, dissociation kinetics of *2:*105
Aminolevulinate dehydratase (ALAD)
  activity *2:*236
  protein *2:*236 *2:*238f
Aminolevulinic acid (ALA) *2:*236
Ammonia *4:*390
  in acid mine treatment *5:*4
  analysis of *4:*393
  bromate ion formation and *1:*360
  cycling of *4:*391
  eutrophication and *4:*392

Index terms	**Links**	
Ammonia *(Continued)*		
impact on aquatic organisms	*2:*123	
in inhibition of biological iron removal	*2:*153	
measurement of	*4:*393	
in ponds	*3:*485	
sources and concentrations of	*4:*392	
toxicity of	*4:*392	
in urban stormwater runoff	*3:*434	
Ammonia-ammonium equilibrium	*4:*519	
Ammonia nitrogen, in recirculating aquaculture systems	*3:*544	
Ammonia removal	*1:*346	
steady-state model for	*1:*348	
Ammonia toxicity	*3:*696	
water chemistry and	*3:*132	
Ammonification	*1:*893	*4:*517
Ammonium. *See also* Ammonia-ammonium equilibrium		
distribution of	*4:*74	
environmental pathways of	*3:*538	
in source waters	*2:*226	
Ammonium perchlorate	*5:*631	
Amorphous hydrous ferric oxide	*5:*609	
Amorphous ice	*4:*189	
Amoebas	*2:*313	
Amoebic dysentery	*2:*313	
Amorphous organic matter domains, sorption in	*4:*386	
Amperes	*4:*442	
Amphidromous fish species, in the Hawaiian Islands	*4:*803	
Amphipods		
characteristics of	*2:*409	
commonly used	*2:*409t	
Amphipod sediment toxicity tests	*2:*408	
application of	*2:*409	
experimental design in	*2:*410	
future research related to	*2:*411	
organism selection in	*2:*409	
significance of	*2:*408	
Anacapa Island pelican colony survey	*3:*517	
Anadromous fish	*3:*127	
Anaerobes	*5:*581	
See also Anaerobic bacteria; Anaerobic microbial transformation		

Index terms	**Links**	
Anaerobic-aerobic treatment, for maize processing plants	*1:*581	
Anaerobic attached film expanded bed (AAFEB) system	*1:*519	
Anaerobic bacteria	*2:*20	
See also Anaerobes		
iron-reducing	*2:*151	
oil-field brine and	*2:*288	
Anaerobic bioremediation technologies	*5:*40	
Anaerobic biotransformation, of PCE/TCE	*5:*584	
Anaerobic dechlorination	*1:*690	
Anaerobic decomposition	*1:*701	
odor from	*1:*760	
Anaerobic degradation pathways	*5:*582f	
See also Biodegradation		
Anaerobic digestion	*1:*518	
of biosolids	*1:*647	
of sludge	*1:*856	*1:*865
Anaerobic Digestion Model	*1:*733	
Anaerobic filter (AF)	*1:*906	
Anaerobic leachate treatment	*1:*709	
Anaerobic microbial transformation, of methyl *tertiary*-butyl ether	*5:*386	
Anaerobic reactor systems	*1:*906	
Anaerobic sewage treatment	*1:*517	
See also Anaerobic wastewater treatment		
examples of	*1:*519	
future of	*1:*520	
Anaerobic systems, for treating leachates	*1:*704	
See also Anaerobic reactor systems; Anaerobic sewage treatment		
Anaerobic wastewater treatment	*1:*813	*1:*904
See also Anaerobic sewage treatment		
application of	*1:*909	
separation of acidogenesis from methanogenesis	*1:*909	
technology of	*1:*905	
Anaerobic wetlands	*2:*425	*5:*5
Analysis		
of antimony	*4:*591	
of colloids	*3:*74	
in fish consumption advisories	*3:*119	
Analysis of variance (ANOVA) comparison technique	*2:*411	
systems approach to	*2:*685	
Analytical modeling	*5:*33	

Index terms	**Links**
Analytical solutions	*5:*621
Anatoxin-a	*2:*390
Anatoxins	*2:*388
Ancient lakes	*3:*270
Ancient world, water in	*4:*726
Animal facilities, managing	*2:*186
Animal farming, groundwater quality related to	*3:*538
Animal husbandry	*3:*124
Animal manure, nitrogen in	*3:*695
Animals, rain forest	*4:*240
Anionic organoclay	*1:*779
Anionic polymeric antiscalants	*1:*417
Anionic surfactants	*1:*670
Anions, in seawater	*4:*159
Anisotropic heterogenetic flow systems, testing and analysis of	*5:*176
Anisotropy	*5:*509
Annual status reports (EPA)	*5:*39
Annual water quality report (AWQR)	*2:*346
Anodic stripping voltammetry (ASV)	*2:*207
Anomalous propagation (AP)	*4:*308
Anoxia	
factors affecting	*4:*65
quantification of	*4:*64
variations in	*4:*65
Anoxic biodenitrification	*5:*326
Anoxic conditions	*1:*667
Anoxic factor (AF)	
determining	*4:*65
prediction of	*4:*67
Anoxic limestone drain (ALD)	*2:*424
in acid mine treatment	*5:*5
Anoxic periods (denitrification), controlled	*1:*752t
Antarctica, microbial ecosystems in	*3:*504
Antecedent moisture conditions	*3:*473
Anthra/sand	*1:*314
Anthropogenic activities, toxic metals accumulation from	*2:*433
Anthropogenic arsenic	*5:*18
Anthropogenic CO_2, global partitioning of	*4:*85
Anthropogenic compounds, toxic effects of	*2:*427

Index terms	**Links**	
Anthropogenic dammed pools	*3:*69	
Anthropogenic disturbances, to wetlands	*3:*173	
Anthropogenic factors, coastal wetland loss from	*3:*73	
Anthropogenic GHG emissions	*3:*204	
See also Greenhouse gases (GHGs)		
duration of	*3:*207	
political aspects of	*3:*208	
Anthropogenic load	*2:*443	
Anthropogenic nonturbulent unit	*3:*70	
Anthropogenic scour pools	*3:*68	
Anthropogenic selenium contamination	*5:*397	
Anthropogenic sources		
of arsenic	*1:*82	
of heavy metals	*5:*276	
of lead	*2:*433	
of radioactive contamination	*1:*803	
of vinyl chloride	*5:*635	
Anthropogenic tritium	*5:*69	
Anthropogenic turbulent unit	*3:*70	
Antibiotic resistance	*1:*178	
Antibiotics, in animal farming operations	*3:*540	
Antibody technology, for identifying *Cryptosporidium*	*1:*160	
Anti-DDT propaganda	*3:*513	
Antidegradation, precautionary principle regarding	*2:*601	
Antidegradation policy, Great Lakes system	*4:*624	
Antifluoridationist movement	*1:*255	
Antifoulant design	*1:*415	
Antimony (Sb)		
abundance of	*4:*589	
analytical methods for	*4:*591	
in aquatic systems	*4:*589	
atmospheric emission values for	*4:*589	*4:*590t
concentrations of	*4:*592	*4:*593
kinetic issues related to	*4:*592	
maximum contaminant level of	*2:*234t	
solution chemistry of	*4:*591	
speciation in natural waters	*4:*592	
uses of	*4:*590	
Antiscalants, anionic polymeric	*1:*417	
Antoine’s equation	*4:*265	

AOC (assimilable organic carbon) level, microbial regrowth and *1:*344
A/O (anaerobic/oxic) process *1:*790
Apalachicola-Chattahoochee-Flint (ACF) River Basin water allocation dispute *2:*501
  Army Corps of Engineers role in *2:*506
  crisis related to *2:*503
  role of interest groups in *2:*505
  water allocation and *2:*504
  water marketing proposals related to *2:*506
APE surfactants *1:*670
Aphotic zone *3:*265 *3:*267
Aplysiatoxin *2:*389
Apoptosis, characterizing *3:*116
Apparent density *1:*96
Apparent soil electrical conductivity (EC_a). *See also* EC*a* entries
  edaphic factors influencing *5:*466
  mobile measurement equipment for *5:*467
Apparent temperature effect *4:*439
Appliances, water-saving *2:*497
Application portals *2:*670
Appropriation doctrine *2:*542
Approximate infiltration models *4:*487
Aquacultural facilities
  separation processes for *1:*683t
  waste loads in *1:*681
Aquaculture *3:*123 *3:*124 *3:*127
  defined *2:*542
  technology related to *3:*540
Aquaculture ponds *3:*540
  constituents of *3:*541
Aquaculture systems
  flow-through *3:*542
  net pen *3:*543
  recirculating *3:*544
Aqua earth observing system (EOS) *4:*214
  See also Aqua spacecraft
  data processing and distribution related to *4:*216
  goals and objectives of *4:*216
  management of *4:*216
  spacecraft and instruments in *4:*215
Aquarium fish *3:*123

Aqua spacecraft *4:*194
  water cycle and *4:*193
  weather forecasting and *4:*351
Aquatic animals, harvesting of *3:*123
Aquatic bacteria *1:*639
Aquatic biota. *See also* Aquatic organisms
  antimony concentrations in *4:*593
  trace elements in *3:*456
Aquatic cadmium *5:*614
Aquatic colloids *4:*27
Aquatic ecosystem resources, Great Lakes region *3:*178
Aquatic ecosystems *2:*617
  See also Aquatic environments/ habitats
  availability of nutrients in *3:*109
  biotic integrity of *3:*36
  damage to *1:*555
  health of *3:*116
  role of macrophytes in *1:*714
Aquatic environments/habitats. *See also* Aquatic ecosystems
  alteration of *3:*134
  characteristics of *3:*123
  classification and environmental quality assessment in *2:*94
  effect of sediment on *3:*508
  effect of urban stormwater runoff on *3:*436
  impact of road salt on *2:*320
  lipophilic contaminant monitoring in *2:*170
  trace metal pollution in *2:*64
Aquatic foodweb, plankton in *4:*157
Aquatic health, assessing *3:*40
Aquatic humic acid, origin, stability, and mobility of *5:*190
Aquatic life
  effects of acidification on *3:*3
  environmental contaminants of *3:*115
  impact of chronic acidification on *3:*4
Aquatic Life Criterion *2:*355
Aquatic life toxicity. *See also* Aquatic toxicity entries
  testing *4:*600
  of urban stormwater runoff *3:*433
Aquatic macroinvertebrates, effect of road salt on *2:*321

Index terms	**Links**
Aquatic macrophytes	*3:*107
See also Aquatic vegetation; Aquatic weeds; Macrophytes; Submerged aquatic plants	
as biomonitors	*2:*65
hyperaccumulation potential of	*2:*66
research opportunities related to	*2:*66
Aquatic nitrogen	*4:*518
Aquatic organisms. *See also* Aquatic animals; Aquatic biota; Submerged aquatic plants	
bioaccumulation of hazardous chemicals in	*3:*435
as indicators of aquatic ecosystem health	*3:*40
toxicity of lead to	*2:*434
Aquatic plants. *See also* Aquatic vegetation; Aquatic weeds; Submerged aquatic plants	
excessive growth of	*1:*788
in mine effluent remediation	*1:*612
Aquatic resources, renewable	*3:*121
Aquatic sediments	*2:*384
Aquatic systems	
antimony in	*4:*589
assessing the quality of	*2:*25
impact on	*2:*24
water reduction in	*2:*465
Aquatic toxicity. *See also* Aquatic life toxicity	
measures of	*2:*381
standardized test protocols for evaluating	*2:*278
Aquatic toxicity testing	*2:*380
microscale	*3:*116
Aquatic vegetation	*3:*108
Aquatic weeds	*3:*742
biology and ecology of	*3:*743
types of	*3:*743
Aqueducts	
defined	*2:*542
Roman	*4:*266
Aqueous electrons	*1:*579
Aqueous phase partition	*1:*761
Aqueous solutions	
analysis of	*4:*443
phenol adsorption from	*4:*404
removal of organic micropollutants and metal ions from	*4:*506
Aqueous systems, removal of metal ions from	*1:*918

Index terms **Links**

Aquicludes *5:*10

Aquifer flow *5:*237

Aquifer matrix, sorption and desorption of CFCs from *4:*422

Aquifer parameters *5:*164

  steady-state methods of determining *5:*492

Aquifers *5:*9

  See also Artesian aquifers; Coastal aquifers; Confined aquifers; Fossil aquifers; Fractured aquifers; Granular aquifers; Karst aquifers; Limestone aquifers; Metamorphic rock aquifers; ''Native'' aquifers; River-connected aquifers; Semiconfined aquifer entries; Sloping aquifer; Surficial aquifers; Unconfined aquifers

  characterization of *5:*10

  confined *4:*167

  contamination of *3:*228

  defined *2:*542 *5:*600

  drawdown in *5:*103

  enhanced flushing of *5:*440

  groundwater in *5:*71

  hydraulic head in *5:*171

  microbial degradation of CFCs in *4:*422

  natural replenishment of *5:*11

  overpumped *2:*648

  properties of *5:*492

  remediation of *4:*675

  river-connected *5:*677

  safe yield of *5:*575

  sensitivity to contamination *5:*56

  size of *5:*9

  storativity in *5:*129

  surficial *5:*80

  travel time of water through *5:*65

  types of *5:*515

  unconfined *4:*168 *5:*11 *5:*483 *5:*497

  variability of *5:*664

  water quality and *5:*9

Aquifer-storage-recovery (ASR) systems *2:*515

Aquifer-storage-recovery technique *5:*134

Aquifer test analysis models *5:*663

Aquifer test data, analysis of *5:*183

Aquifer tests *5:*182

  of hydraulic conductivity/transmissibility *5:*513

Aquifer tests *(Continued)*
  steady-state flow *5:*491
Aquifer thermal capacity *5:*173
Aquitards *5:*10
Arab-Israeli water resources *4:*699
Arable farming *3:*568
Arab World
  conservation and demand management in *2:*472
  renewable water resources by state *2:*471t
  supply and demand in *2:*471
  water law in *4:*634
  water management solutions in *2:*472
  water resources in *2:*470
^{39}Ar age dating method *4:*389
Aragonite *4:*415
Aral Sea basins
  irrigation in *3:*17t
  water sharing in *3:*18
Aral Sea disaster *3:*15
  effects of *3:*16
Arbitrary fixed radius (AFR) *1:*525
Archaeology, marine *4:*62
Arctic Ocean
  metal levels in *4:*100
  sea ice in *4:*69
Area method, of base flow separation *3:*26
Argon (Ar). *See* ^{39}Ar age dating method
Arid areas, sediment production in *3:*31
  See also Arid countries/lands; Arid regions
Arid climates, soil erosion in *3:*408
Arid climatic zone *4:*257
Arid countries/lands. *See also* Arid areas; Dry environments; Dry-land agriculture
  crop irrigation in *2:*549
  effluent water regulations in *2:*475
  wastewater applications in *1:*632
  water regulation in *2:*478
Arid regions. *See also* Arid areas defined *5:*409
  precipitation and recharge variability in *5:*409
  precipitation versus evapotranspiration in *5:*410
  recharge in *5:*72 *5:*408

Index terms **Links**

Arid regions *(Continued)*
  subclassifications of *5:*409t
Aristotle *4:*352
  observations on weather *4:*348
Arizona
  cienegas in *3:*58
  in-stream flow protection in *4:*662
  overdraft in *5:*341
Arizona spillway *4:*733
Armonica *4:*758
  future of *4:*762
  as a healing water instrument *4:*760
  resurrection of *4:*761
Army Corps of Engineers *4:*612 *4:*616
  permits issued by *4:*665
  role in Apalachicola-Chattahoochee-Flint River Basin water allocation dispute *2:*501
  water marketing and *2:*506
  working relationship with Congress *2:*523
Aromatic compounds
  chlorinated *1:*688
  degradation of *1:*693
Aromatic hydrocarbon *2:*110
Arrhenius kinetics *4:*565
Arsenates *2:*16 *2:*17
  See also As(V)
  iron coprecipitation of *5:*20
Arsenic (As). *See also* Arsenic removal entries
  anthropogenic sources of *1:*82
  in ash pond water *1:*851
  common minerals of *1:*81t
  detecting in drinking water *1:*2
  groundwater *5:*17
  health effects of *2:*15
  immobilization of *5:*18
  inorganic *2:*8
  in situ electrokinetic treatment of *5:*28
  maximum contaminant level of *2:*234t
  metabolism and disposition of *2:*16
  minerals with *5:*18
  natural distribution of *1:*82
  in natural waters *1:*81

Index terms **Links**

Arsenic (As) *(Continued)*
  particulate *2:*11
  regulations related to *1:*82
  role in metal tolerance *3:*612
  toxicity mechanisms of *2:*17
  treating in water supply wells *5:*22 *5:*23
  uncharacterized *2:*11
  uptake of *2:*217
Arsenic compounds *2:*7
  methylated *2:*11
Arsenic drinking water crisis, Bangladesh *1:*1
Arsenic exposure, acute and chronic *2:*17
Arsenic removal
  from drinking water *1:*2
  ion exchange for *5:*21
  mechanisms for *1:*83
  via reverse osmosis *5:*21
Arsenic removal studies *1:*636
Arsenic removal technologies *1:*636
  adsorption processes for *1:*638
  coagulation-precipitation processes for *1:*636
  membrane processes for *1:*637
Arsenic species *2:*7t
  organic *2:*12t
Arsenic standard (USEPA) *5:*17
Arsenic water treatment methods, effectiveness of *5:*20t
Arsenite *2:*17
  See also As(III)
  oxidation *5:*19
Arsenothiols *2:*11
Artesian aquifers *5:*104 *5:*600
Artesian pressure *5:*30
Artesian water *5:*29
  defined *2:*542
Artesian well *5:*600
Artificial aquatic systems *3:*311
‘‘Artificial’’ communities *2:*379
Artificial groundwater tracers *5:*504
Artificial neural networking (ANN) *4:*323

Index terms **Links**

Artificial recharge
  defined *2:*542
  flood spreading for *3:*164
  of unconfined aquifers *5:*11
Artificial valleys *3:*66
Artificial water retention *1:*405
Arumugam Plan *4:*776
As(III) *1:*81 *1:*82
  oxidation of *1:*639
As(V) *1:*83
Asbestos
  analysis of *2:*310
  maximum contaminant level of *2:*234t *2:*235t
Asbestos-containing material (ACM) *2:*310
Aseptic meningitis *1:*178
Ash, from coal combustion *1:*555
Ash percent, of carbon *1:*97
Ash pond water quality *1:*850
Asia, urban water resource management in *2:*552
Asian cities, flooding in *3:*160
Asian plains, developing *3:*16
ASM models *1:*739 *1:*740
  See also Activated Sludge Models (ASMs)
Assays, standardization and reproducibility of *2:*378
Assessment
  biological impact *4:*111
  of contaminant sensitivity *5:*57
  source water *1:*444
  underground pipeline *1:*884
Asset sale privatization *1:*389
Assimilable organic carbon (AOC) *2:*175 *2:*224 *2:*244
Asthma, dehydration and *4:*725
Astragalus bisulcatus *3:*35
Astrovirus *1:*279
Astrovirus enteritis *1:*179
Atkins, Anna *4:*767
Atlantic Oceanographic and Meteorological Laboratory (AOML) *4:*4
Atmosphere *4:*179
  compressing weight of *4:*360
  dynamics of *4:*296

Atmosphere *(Continued)*
  hydrological cycle of *4:*181
  long-wave radiation from *3:*192
  ocean interaction with *4:*22
  pressure and convection in *4:*360
  vapor conditions in *4:*224
  wave propagation in *4:*295
  weather and *4:*360
Atmospheric acidification *3:*8
Atmospheric Boundary Layer (ABL) *4:*181 *4:*296
Atmospheric cadmium *5:*614
Atmospheric carbon dioxide, increase of *4:*176
Atmospheric Infrared Sounder (AIRS) *4:*215 *4:*351
Atmospheric nitrogen *3:*695
Atmospheric/oceanic circulation, changes in *4:*174
Atmospheric pollutant deposition *3:*282
Atmospheric scientists *4:*328
  employment of *4:*328
  job outlook and earnings for *4:*329
  responsibilities and working conditions of *4:*328
  training, qualifications, and advancement of *4:*329
Atmospheric temperatures, global-scale measurements of *4:*173
Atmospheric tides *5:*498
Atmospheric trace constituents, anthropogenic *4:*176
Atmospheric tritium, production and decay behavior of *5:*65
Atomic absorption spectrophotometer (AAS) *4:*508
Atomic absorption (AA) spectroscopy *2:*308 *4:*498
ATP content assays *2:*415
ATP sulfurylase *5:*399 *5:*400
ATP-sulfurylase (*APS*) gene *5:*288
Atrazine *2:*312 *4:*506
Attached growth processes *1:*827
Attainable uses, for surface waters *2:*476
Attenuation, as a watershed function *3:*478
Attenuation factor *5:*597
Attenuation processes,
phytoremediation enhancement of *5:*374
  See also Monitored natural attenuation (MNA); Natural attenuation processes
Audubon Society *3:*513 *3:*514
Australian Drought Watch System *4:*214

Index terms	**Links**		
Australian River Assessment System (AusRivAs)	*3:*40		
Australian Water Resources Council (AWRC)	*4:*449		
Authority-polluter negotiations	*4:*647		
Autoanalyzer	*4:*514		
Autoconvection	*4:*366		
Autogenic recharge	*5:*236		
Automated local evaluation in real time (ALERT) systems	*3:*148		
Automatic model calibration	*2:*332		
Automation, of microirrigation	*3:*623		
Autonomous CTD profiler (ACP)	*4:*71		
Autothermal thermophilic aerobic digestion (ATAD)	*1:*648		
Average daily flow (ADF)	*3:*106		
Average water requirement	*1:*506		
Averting behavior water value estimation model	*4:*609		
Avoided cost water valuation approach	*2:*655		
Awesome Aquifers science project	*4:*714		
Axial flow pumps	*3:*380		
Axial impellers	*1:*76	*1:*78f	
flow patterns induced by	*1:*78f		
Azimuthal resistivity method	*5:*147		

B

Backfill, hydrology of	*5:*2		
Backflow	*1:*203		
devices	*1:*343		
prevention of	*1:*155		
Background pollutant concentration	*2:*18		
inorganic pollutants	*2:*19		
organic pollutants	*2:*20		
water chemistry and	*2:*19		
Backpressure	*1:*155		
Backscatter	*4:*321		
Backsiphonage	*1:*155		
Backup electrical power	*1:*303		
Backwashable depth filtration	*1:*459		
Backwater pools	*3:*69		
Backwater profiles	*1:*57		
Backyard sewerage	*2:*633		
Bacteria	*1:*903	*3:*668	*5:*580
See also Marine bacteria			

Index terms	**Links**	
Bacteria *(Continued)*		
activated sludge	*1:*668t	
algal control using	*2:*6	
biofilm-associated	*1:*603	
chemical transformation via	*5:*378	
denitrifying	*1:*667	
dissolved oxygen and	*3:*445	
in domestic sewage	*1:*831	
effect on bioventing	*5:*429	
in foaming and bulking	*1:*846	
health effects of	*1:*277	
heterotrophic	*1:*567	
identification of	*2:*21	
indicator	*2:*293	
iron-oxidizing	*2:*149	
luminescent	*2:*172	
MAG tests for	*1:*84	
on Mars	*4:*747	
metabolite production by	*1:*601	
metal-resistant	*2:*446	*2:*448
methanogenic	*5:*47	
molecular and biochemical techniques related to	*2:*23	
motility of	*5:*351	
nitrifying	*1:*753	
nutrient levels and	*1:*598	
origins of	*5:*463	
pH and	*3:*14	
planktonic	*1:*596	
removal of	*1:*485	
resistance to heavy metals	*2:*443	
role in heavy metal phytoremediation	*5:*376	
waterborne	*2:*20	
as water quality indicators	*2:*23t	
Bacteria count	*5:*184	
Bacterial agents	*1:*178	
Bacterial bioassay	*2:*389	
Bacterial biofilm	*1:*539	
formation of	*1:*539	
Bacterial bioluminescence	*2:*453	
toxicity and	*2:*47	

Index terms	**Links**	
Bacterial biosensors		
availability of	*2:*455	
constructed	*2:*453	
luminescent	*2:*453	
Bacterial concentration decreases, dilution equivalent to	*4:*104	
Bacterial disinfection	*1:*469	
Bacterial/fecal contamination, indicators of	*2:*21	
Bacterial growth		
byproducts of	*5:*23	
dependence on metal concentration	*2:*446f	
Bacterial growth factor	*1:*223	
Bacterial infections	*2:*371	
Bacterial isolates, identification of	*1:*664	
Bacterial monitoring, water treatment plant	*1:*222	
Bacterial neuston	*2:*442	
Bacterial populations, increase in	*3:*114	
Bacterial regrowth		
analysis of	*1:*223	
BDOC impact on	*1:*225	
Bacterial screening tests	*1:*566	
Bacterial source tracking (BST)	*2:*21	
Bacterial transport		
factors affecting	*5:*351	
in groundwater	*5:*350	
in groundwater bioremediation	*5:*43	
in groundwater systems	*5:*350	
Bacteria-plant-contaminant interactions	*5:*378	
Bactericides, drinking water disinfection using	*1:*382	
Bacteriological content, of sludge	*1:*863	
Bacteriological monitoring	*2:*267	
Bacteriological parameters, of River Yamuna groundwater	*2:*396	
Bacteriophages	*3:*312	
Bacteriophage transport studies	*1:*368	
Bacterioplankton	*4:*155	
progression in reservoirs	*3:*207f	
BACTOX test	*2:*414	
Bad cholesterol, myth of	*4:*701	
Bag filtration	*1:*459	*1:*488
Bahls' periphyton index of biotic integrity	*3:*40	
Baird, Spencer	*4:*141	

Index terms	**Links**	
Balanced ecological community	*2:*193	
Balanced ecosystems	*1:*639	
Balantidium coli	2:313	
Bald eagles, effects of DDT on	*3:*519	
Ballasted flocculation	*1:*454	*4:*426
Ballooning, meteorology and	*4:*164	
Balloons, meteorological	*4:*164	*4:*328
BA-metal complex separation	*1:*587	
Band-sar	*3:*708	
Bangkok, flooding in	*3:*160	
Bangkok Metropolitan Administration rain gauges	*1:*124f	
Bangladesh		
arsenic drinking water crisis in	*1:*1	
flooding in	*3:*160	
Bank filtration	*1:*486	
Bank storage, base flow and	*3:*22	
Baptism, as a water ritual	*4:*709	
Barbiturates, adsorption of	*1:*118	
Bardenpho process	*1:*790	
Bare metal mode	*4:*445	
Barium (Ba), maximum contaminant level of	*2:*234t	
Barium sulfate, co-precipitation of radium with	*1:*398	
Barley straw, in algal control	*2:*5	
Barnacles	*1:*541	
Barometric efficiency	*4:*166	
Barometric loop	*1:*157	
Barometric pressure, effect on wells	*5:*182	
Barrages	*3:*201	
Bar screens	*1:*784	
Bartlett-Lewis rectangular pulse model	*3:*428	
Basalt aquifers	*5:*145	
Base cation concentrations	*3:*3	
Base cation depletion	*3:*3	
Base cation dilution, acidification and	*3:*5	
Base flow	*3:*22	
defined	*2:*542	
factors affecting	*3:*22	
recession of	*3:*22	
Base flow curves, ideal	*3:*23	

Index terms **Links**

Base flow discharge (*Q*) *3:*23
Baseflow hydrograph *3:*476
Base flow index (BFI) *3:*27
Base flow recession
  applications of *3:*26
  as auto regressive processes *3:*24
  constants determining *3:*24
Base flow recession curve, master *3:*25
Base flow separation *3:*26
  graphical methods of *3:*27
Baseline stations *2:*178
Bases
  defined *2:*542
  strength of *2:*294
Basin area, peak discharge per unit of *3:*29
Basin characteristics, role in flooding *3:*143
Basin circularity ratio *3:*94
Basin elongation *3:*29
Basin geomorphology *4:*222
Basins
  inflow from and outflow to *5:*167
  retarding *3:*150
BASINS. *See* Better Assessment Science Integrating Point and Nonpoint Sources (BASINS)
Basin slope *3:*29
Batch adsorption experiments *4:*508
Batch membrane photoreactors, 4-nitrophenol degradation in *1:*793
Batch sorption isotherm studies *1:*112
Batch systems, activated carbons in *1:*101
Bautzen Reservoir, biomanipulation trials in *2:*57t
BAX® system *2:*140
Bays *4:*25
Bazin’s roughness coefficient, values for *3:*196t
BBC Water Portal *2:*669
B climates *4:*257
Beaches *4:*26
Beaches Environmental Assessment and Coastal Health (BEACH) Act *4:*39
Beach raking *4:*41
Bear Trap Dam *3:*44
Beaufort Laboratory, coastal fishery and habitat research at *4:*55

Index terms **Links**

Beaver dam pools *3:*69
Bed depth-service time model (BDST) *2:*363
Bedding, in pipeline repair *1:*889
Bed life *1:*104
Bed load *3:*400
  See also Bed-material load
  transport of *3:*418
Bed material, particle size distribution of *3:*420
Bed-material load *3:*397
Bedrock *5:*55
  defined *2:*542
  dissolution *5:*243
Bedrock valleys *3:*66
Bed sediment metal ion adsorption *3:*295
  experimental methodology for *3:*296
  operating variables in *3:*298
  results of *3:*298
  river system used in *3:*296
Behavioral endpoints, in sediment toxicity testing *3:*54
Behavioral screen fish passage facility *3:*531
Belt pressing *1:*855
Benazolin
  adsorption isotherms of *4:*510
  adsorption of *4:*509
  properties and structural formula of *4:*507t
Benchmark process *1:*440 *1:*441t
Bench terraces *3:*550 *3:*567
Bench tests *5:*424
Beneficial use classification *2:*191
Beneficial uses
  protection of *2:*478
  for surface waters *2:*476
Benefit transfer process *4:*610
Benthic algal mats *3:*277
Benthic chambers *2:*352
Benthic environment, in net pen aquaculture systems *3:*543
Benthic nutrient regeneration *4:*6
  temporal scaling of *4:*11
Benthic organisms/species
  exposure tests with *2:*212

Index terms	**Links**			
Benthic organisms/species *(Continued)*				
sediment as habitat for	*2:*408			
surface water metals and	*2:*217			
Benthic zone	*3:*267			
Bentonite	*1:*771	*1:*778		
Benzene	*5:*427			
biodegradation of	*1:*693			
chemical structure of	*5:*626			
electrolysis of	*5:*122			
exposure to	*5:*627			
in groundwater	*5:*626			
hazardous effects of	*5:*627			
in situ electrokinetic treatment of	*5:*116			
production and uses of	*5:*627			
remediation of	*5:*627			
Benzene adsorption, on zeolites	*4:*406			
Benzene-contaminated water	*1:*580			
Benzene contamination	*5:*627			
Benzene, toluene, ethylbenzene, and xylene (BTEX)	*5:*529	*5:*424	*5:*425	
plumes	*5:*80			
soil vapor concentration of	*5:*552			
Benzoic acid, adsorption of	*1:*117			
Berms				
landside	*3:*290			
landside seepage	*3:*289			
stability	*3:*288			
Bernoulli equation	*3:*194	*3:*195	*5:*170	*5:*180
Berylliosis	*4:*394			
Beryllium (Be)				
chemistry of	*4:*395			
maximum contaminant level of	*2:*234t			
source material for	*4:*395			
in water	*4:*394			
Beryllium concentrations, range of	*4:*395			
Best available control technology (BACT)	*1:*101			
Best available technology (BAT)	*1:*360	*1:*479	*1:*681	*1:*757
Best conventional pollutant control technology (BCT) standards	*4:*656			
Best management practices (BMPs)	*1:*54	*1:*681	*2:*192	*3:*503
	*4:*669			
in aquaculture	*3:*542			

Best management practices (BMPs) *(Continued)*
  to minimize groundwater nitrate contamination *2:*571
  for nitrogen *3:*694 *3:*696
  for nonpoint source control *2:*187
  in nutrient runoff control *3:*110
  riparian buffer *3:*391
  solidification/stabilization processes and *1:*838
  stormwater management *3:*443
  for urban stormwater runoff treatment *3:*432
  for water resources *2:*570
Best practicable control technology (BACT, BPT) *4:*596 *4:*656
Best professional judgment (BPJ) *4:*656
  triad weight-of-evidence as *4:*600
Beta-gamma radiation *4:*547
Beta emitters, health risks of *1:*396
Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) *2:*328 *2:*249
BGP carbon *2:*83
Bhuj earthquake *5:*478
Bible, water symbolism in *4:*786
Bicarbonate alkalinity *4:*410
  calculation of *4:*410
Bilevel drains, falling water table between *5:*96
Bingham plastics *5:*556
Bioaccumulated metals *2:*421f
Bioaccumulation *2:*44 *2:*69 *3:*34
  in assessing sediment quality *2:*426
  bioassays *2:*351
  characteristics of *2:*70t
  of chemical contaminants *2:*125
  of chemicals in edible aquatic organisms *3:*435
  in Indian lakes *3:*451
  recommendations concerning *2:*431
  of selenium *2:*357t
  use in regulation *2:*429
Bioaccumulation Factor (BF) *1:*717
Bioaccumulation testing *4:*600
Bioassay batteries *2:*383
Bioassays
  in marine environmental monitoring *2:*443 *2:*444
  purpose of *2:*418

Bioassays *(Continued)*
  sediment toxicity *2:*458
  subcellular *2:*278
  use of polychaetes in *4:*42
Bioassessment procedures
  current *2:*25
  development of *2:*24
  preserving biotic integrity and *2:*26
Bioaugmentation *5:*352
Bioavailability *4:*525
  in aquatic ecosystems *2:*170
  metal *2:*235
  of zinc and cadmium *5:*370
Bioavailability biosensors, for arsenic detection *1:*2
Biochemical field surveys *5:*119
Biochemical groundwater parameters *5:*184
  field testing and analysis of *5:*186
Biochemical oxygen demand (BOD) *1:*455 *1:*623 *1:*639 *1:*701 *1:*733 *2:*37 *2:*175 *2:*332
  See also BOD$_5$ test; Rapid BOD technologies
  carbonaceous *1:*754t
  control strategies for *1:*642
  historical context of *2:*37
  of industrial wastewater *1:*569
  measurement of *1:*641 *2:*37
  microbes and *1:*640
  in natural and engineered systems *1:*639
  operation of *2:*38
  uses for *2:*38
Biochemical oxygen demand test *1:*829
Biochemical processes *5:*116
Biochemistry, role of dissolved gases in *4:*451
BIOCHLOR program *5:*141
Biocides
  dosage levels of *1:*602t
  “environmentally friendly” *1:*541
  resistance to *1:*602
  types of *2:*242
Bioconcentration factor (BF) *3:*35
Biocorrosion *1:*600
Biodegradable capacity, of soil *3:*706

- Biodegradable dissolved organic carbon (BDOC) *2:*224 *2:*244 *2:*245
  - *See also* Biodegradable organics
  - concentration of *1:*226
  - impact on bacterial regrowth *1:*225
- Biodegradable dissolved organic carbon analysis *1:*223
- Biodegradable organic carbon control of *2:*224
- Biodegradable organics *1:*903
  - *See also* Biodegradable dissolved organic carbon (BDOC)
- Biodegradation *2:*41 *3:*35 *5:*528
  - *See also* Aerobic biodegradation
  - acclimation phase of *2:*43
  - algal bioremediation and *2:*44
  - bioavailability and *2:*43
  - of detergents *1:*672
  - enrichment cultures and *2:*42
  - enzymatic mechanisms in *2:*42
  - factors limiting *2:*42
  - groundwater bioremediation and *2:*44
  - marine oil spills and *2:*44
  - mechanisms of *5:*582
  - microbial diversity and *2:*42
  - modeling *5:*31
  - naturally occurring processes of *5:*579
  - nonaqueous phase liquids and *2:*44
  - of organic vapors *5:*545
  - recalcitrant molecules and *2:*43
  - in subsurface systems *5:*415
- Biodegradation microbiology, anaerobic *5:*40
- Biodiversity
  - monitoring in wastewater treatment plants *1:*642
  - precautionary principle regarding *2:*600
  - in rain forests *4:*240
  - reducing *2:*637
- Biodrainage *5:*98
- Bioenhancement methods *5:*117
- Biofilm(s) *1:*205 *1:*576
  - bacterial *1:*539
  - channel structures in *2:*230
  - composition of *5:*36
  - control of *1:*541
  - detachment of *2:*232

Index terms **Links**

Biofilm(s) *(Continued)*
development of *1:*596
in distribution systems *2:*244
formation of *2:*228
growth curve for *1:*540
heavy *1:*602
metabolic heterogene1ty of *2:*230
microbial dynamics of *2:*228
microbial structure and dynamic heterogeneity in *2:*230
microbiology of *3:*307
nitrification/denitrification *2:*231
oxygen limitation in *2:*231
removal of *1:*602
specialized *3:*308
structural heterogeneity of *2:*229
sulfate-reducing *2:*231
Biofilm bacteria *1:*539 *1:*597 *1:*601
Biofilm formation, boundary layers and *1:*600
Biofilters, for odor abatement *1:*763
Biofouling *1:*415
See also Biofouling monitoring
analysis *1:*84
caused by iron bacteria *2:*151
in a cooling tower *2:*241
effects on production efficiency and water quality *5:*35
effects on system hydraulic performance *5:*37
evaluating microbial components of *1:*83
of industrial cooling water *1:*538
in situ electrokinetic treatment of *5:*28
in a membrane water treatment system *2:*242
methods of controlling *2:*242
microbial forms in *2:*239
nutrient adsorption in *2:*239
surfaces and *1:*539
symptoms of *1:*83
treating in water supply wells *5:*22 *5:*27
in water wells *5:*35
Biofouling microflora *5:*464
Biofouling monitoring
practical issues in *1:*86
sampling methods for *1:*86

Index terms	**Links**	
Biofouling layers, relationship of microorganisms inside	*2:*240	
Biofuel alternatives	*3:*545	
as “climate neutral”	*3:*546	
impacts of	*3:*547	
substance flows from	*3:*547	
Biofuel life cycle	*3:*547	
Biogas		
controlling	*1:*860	
production of	*1:*905	
in sludge treatment	*1:*856	
Biogeochemical cycling	*3:*173	
‘‘Biogeochemical reflux’’ scheme	*2:*359f	
Biogeochemistry	*1:*898	
hydrosphere	*4:*284	
Bioindication approach, to marine environmental monitoring	*2:*443	
Bioindicators	*2:*29	
species used as	*2:*34	
versus biomonitors	*2:*64	
Bioleaching	*5:*434	
Biological Activity Reaction Test (BART™)	*1:*85	
Biological agents	*4:*295	
Biological carbonates	*4:*415	
Biological contaminants	*2:*343	
detection in surface waters	*2:*441	
Biological control, of the water hyacinth	*3:*482	
Biological control systems, for odor removal	*1:*914	
Biological CSO treatment	*1:*787	
Biological degradation, of pesticides	*3:*648	
Biological denitrification	*5:*325	
Biological dual-nutrient removal	*1:*816	*1:*834
Biological effects, of acidification	*3:*9	
Biological filtration, organic removal by	*1:*248	
Biological growth, Monod equation and	*1:*735	
Biological impacts, assessing	*4:*111	
Biological integrity	*2:*24	
Biologically produced chemicals	*4:*400	
Biologically stable water	*1:*226	*2:*224
parameter values that define	*2:*224t	
Biological monitoring	*2:*26	*2:*378
See also Biomonitoring		

Index terms	**Links**	
Biological monitoring *(Continued)*		
with bivalve molluscs	*2:*35	
submitochondrial particle assay as a tool for	*2:*376	
Biological nutrient removal (BNR)	*1:*399	
See also Biological dual-nutrient removal		
Biological nutrients, measurement of	*4:*515	
Biological phosphorus removal	*1:*816	*1:*833
in the activated sludge process	*1:*788	
Biological pollutant removal processes	*3:*366	
Biological process modeling	*1:*735	
Biological pump	*4:*86	*4:*414
Biological reactors		
POP behavior in	*1:*768	
sewers as	*3:*331	
Biological samples, of vinyl chloride	*5:*637	
Biological soil treatment	*5:*434	
‘‘Biological supermarkets,’’ coastal wetlands as	*3:*72	
Biological systems		
dissolved gases in	*4:*451	
redox potential ranges in	*1:*399	
temperature and	*5:*174	
Biological toxins, in source waters	*1:*89	
Biological treatment, of landfill leachates	*1:*703	*1:*709
Biological treatment technologies		
for odor abatement	*1:*762	
for wastewater	*1:*813	
Biological wastewater treatment systems	*3:*366	
Biological water-quality parameters	*1:*903	
Biological weed control	*3:*746	
Bioluminescence tests	*2:*47	
Bioluminescence toxicity tests, marine	*2:*48	
Bioluminescent assay systems	*2:*455	
Bioluminescent biosensors. *See also* Recombinant luminescent bacterial biosensors		
in direct toxicity assessment	*2:*47	
for toxicity testing	*2:*45	
in WWTP monitoring	*2:*48	
Bioluminescent recombinant bacteria	*2:*174	*2:*454
Bioluminescent reporter genes	*2:*453	
Biomagnification	*3:*35	
in Indian lakes	*3:*451	

Biomanipulation *2:*50
  lake restoration methods and *2:*50
  principles of *2:*50
Biomanipulation trials, review of *2:*51
Biomarkers *2:*28 *3:*116
  in assessing sediment quality *2:*429
  classes of *2:*429
  metallothioneins as *2:*407
  recommendations concerning *2:*431
  relevance of *2:*32
  use in regulation *2:*429
  use in sediment toxicity assessment *2:*430t
Biomass
  in foulants *1:*417
  of submerged vegetation *3:*276
Biomass immobilization *2:*72
Biomass production, plant nutrients in *3:*548
Biomass water productivity *3:*560
Biomonitoring. *See also* Biological monitoring; Biomonitors
  developing tools for *2:*62
  genomic technologies/tools in *2:*58
  high throughput technologies and *2:*59
  importance of *2:*67
  of polychlorinated biphenyls *1:*715
Biomonitors. *See also* Biomonitoring
  trace metal *2:*64
  versus bioindicators *2:*64
Biooxidation stoichiometry *5:*428t
BIOPLUME III program *5:*33 *5:*141
Bioreactor landfill design *5:*257
Bioremediation *4:*748
  See also In situ bioremediation
  of cadmium *5:*618
  of contaminated soil *3:*653
  enhanced *5:*319 *5:*424
  of groundwater *5:*423 *5:*442
  iron-reducing bacteria in *2:*151
  of metals *5:*282
  of selenium-contaminated waters *2:*355
  of uranium *5:*642
Bioretention systems, for storm water treatment *1:*868 *1:*869f

Index terms **Links**

BIOSCREEN program *5:*141
Bioscreen spreadsheet model *5:*33
Bioscrubber, for odor abatement *1:*763
Biosecurity. *See* Security; Water biosecurity
Biosensors *1:*567 *4:*525
  future development of *2:*455
  microbial aminolevulinate dehydratase *2:*236
  non-toxicant-specific *2:*454t
  respirometric *1:*565
  toxicant-specific *2:*455t
Biosolids *1:*646 *1:*863
  alkaline stabilization of *1:*649
  composting of *1:*648
  digestion of *1:*647
  as a source of contaminants *5:*433
  thermal drying of *1:*649
  uses for *1:*649
  use standards for *1:*650
Biosorbent materials
  broad-range *2:*69
  comparison of *2:*72t
  immobilization of *2:*72
Biosorption *2:*69 *4:*565
  future of *2:*73
  of heavy metal anions *2:*72
  by living microorganisms *2:*71
  mathematical modeling of *2:*71
  technologies *2:*73
  of toxic metals *2:*68
Biosorption treatment plants *2:*73
Biosparging *5:*429
  See also Air sparging
  case history of *5:*430
  technology *5:*41
Biosurfactants, solubilization by *5:*377
Biosurveys, microbial *2:*442 *2:*444
Biosynthesis, of phytochelatins *3:*610
Biota transfer *4:*686
Biotechnology, precautionary principle regarding *2:*600
Bioterrorism Act *2:*347 *2:*348

Index terms **Links**

Biotic assemblages, direct measures of *3:*37
Biotic communities, evolution of *3:*36
Biotic diversity *3:*36
Biotic indicators *3:*37
Biotic integrity
  factors affecting *3:*36
  multimetric indexes of *3:*38
  reasons for using *3:*37
Biotic integrity index *3:*36
Biotic interactions, of microorganisms *3:*312
Biotic pollution *3:*134
Biotic processes, effect on solute transport *5:*528
Biotic redox reactions *5:*413
Biotoxins, health effects of *1:*277
Biotransformation, of chlorinated aliphatic compounds *1:*690
Biotreatment(s)
  of groundwater *5:*120
  for methyl *tertiary*-butyl ether *5:*389
  passive *5:*120
Biotrickling filter, for odor abatement *1:*763
Bioturbation *3:*312
Bioventing *5:*428
  soil *5:*120 *5:*121
  technology *5:*40
  versus soil vapor extraction *5:*430
Biovolatilization *2:*359
Birds
  data related to *3:*514
  effects of DDT on *3:*513
  impact of oil on *4:*106
  importance of coastal wetlands to *3:*72
  toxicity of lead to *2:*434
Birth defects, nitrate-related *1:*37
Bivalve assemblages, dense *4:*73
Bivalve-dominated tidal flat, temporal scaling of benthic nutrient regeneration in *4:*11
Bivalve excretory activity, coupling with porewater nutrient concentrations *4:*75
Bivalve molluscs
  active biomonitoring of *2:*33
  biomonitoring with *2:*34
  chemical monitoring with *2:*35

Index terms	**Links**		
Bivalve molluscs *(Continued)*			
illnesses associated with	*2:*360		
nutrient excretion by	*4:*6	*4:*12	
Bivariate spectral analysis	*4:*339		
Black Mesa monitoring program	*5:*48		
Blackstone River model	*2:*256		
Blackstone River Watershed	*1:*499		
Blackwater	*1:*53	*1:*841	
cesspit for	*1:*843		
coral reef stress and	*4:*133		
MODIS measurements of	*4:*134		
Blue baby syndrome	*2:*219		
Blue Book	*2:*315		
Blue Book of Water Quality Criteria	4:598		
Blue-green algae, toxin producing	*3:*190f		
Blue-green algae blooms			
control of	*3:*189		
health issues associated with	*3:*188		
toxins in	*3:*189		
Blue Planet Project	*4:*770		
Blueprint for Sustainable Development	*2:*624		
Blue Revolution	*2:*627		
BOD_5 test. *See also* Biochemical oxygen demand (BOD)			
components of	*2:*39		
measurement protocol of	*2:*38		
problems associated with	*2:*40		
quality controls for	*2:*39		
quantity measured by	*2:*39		
standards for	*2:*39		
Body. *See also* Human body			
fluid fluctuations in	*4:*458		
water-regulatory mechanisms of	*4:*790		
Body concentrations, time-dynamic speciation and	*2:*215		
Body feed	*1:*232	*1:*247	*1:*251
Bonding, of toxic metal ions	*1:*586		
Bonding agents	*1:*586	*1:*588	
iron-based	*1:*587		
Bones, Jim	*4:*768		
Boreal reservoirs, studies of	*3:*180		
Borehole condition log	*5:*153		

Index terms	Links		
Borehole methods	*5:*151		
Boreholes, amount of gas hydrate in	*4:*61		
Borehole storage	*4:*168		
Boron (B)			
in irrigation water	*2:*159		
irrigation water quality and	*5:*208		
isotopes	*5:*220	*5:*232	
Boron-11	*5:*218		
Boston, extraterritorial land use control in	*1:*316		
Bottled water	*1:*3		
cost of	*1:*5		
historical uses of	*1:*4		
regulation of	*1:*4		
types of	*1:*4		
Boundary conditions			
for coupled flow domains	*5:*660		
flow system	*5:*14		
types of	*5:*561		
Boundary-dependent shock formation	*3:*240		
Boundary-element method	*5:*311		
Boundary layer, atmospheric	*4:*181		
Boundary Waters Treaty of 1909	*4*:618	*4*:620	*4*:686
Bowen ratio	*3:*193		
Box-Cox (BC) transformation	*4:*299		
Box-Jenkins models	*3:*424		
Boyle's law	*4:*360		
Braided channels	*3:*32		
Brain, lead-induced damage to	*2:*435		
Brakensiek's flood routing models	*3:*255		
Bransby-Williams formula	*3:*471		
Brazil, water markets in	*2:*568		
Br^-/DOC ratio	*2:*75	*2:*76	
Breakers, types of	*4:*137		
Breaking waves	*4:*18		
Break-point chlorination	*1:*471		
Breakwaters	*4:*14		
floating	*4:*20		
vertical	*4:*17		
Bridges	*2:*568		
Bridge slot screens	*5:*573		

Index terms	**Links**	
Brilliant Blue FCF	*3:*98	*3:*99
Brine(s)		
acidic	*3:*10	
deposits	*5:*51	
occurrence of	*5:*52	
oil-field	*2:*284	
production of	*5:*54	
products	*5:*54	
resources	*5:*53	
types of	*5:*52	
Brine injection technology	*2:*285	
Brinkman's equation	*5:*659	
British Geomorphological Research Group	*4:*330	
British water companies	*1:*305	
See also U.K. National Rivers Authority		
British water management model	*1:*389	
Bromate ion	*1:*358	
health effects of	*1:*270	
minimization strategies for	*1:*360	
removal of	*1:*361	
Bromate ion formation	*1:*358	
reaction time for	*1:*360	
Bromide ion	*1:*358	
concentration of	*1:*359	
health effects and regulations related to	*2:*74	
influence on trihalomethane and haloacetic acid formation	*2:*74	
Bromine-containing disinfection byproducts	*1:*214	
Bromine hydrate	*4:*473	
Bromoacetic acid, health effects of	*1:*266	
Bromoacetonitrile, health effects of	*1:*267	
Bromochloroacetic acid, health effects of	*1:*266	
Bromochloroacetonitrile, health effects of	*1:*268	
Bromochloromethane, health effects of	*1:*264	
Bromodichloroacetic acid, health effects of	*1:*267	
Bromodichloromethane, health effects of	*1:*265	
Bromoform, health effects of	*1:*265	
Bromomethane, health effects of	*1:*264	
Brooks and Corey equation	*5:*462	
Brown pelicans, effects of DDT on	*3:*516	
Brunauer, Emmett, Teller (BET) isotherm	*1:*101	

Index terms **Links**

Bubble-diffuser contactor *1:*363
Bubbler microirrigation *3:*621
Buckingham-Darcy law *5:*126
  See also Darcy entries
Budgetary models, estuarine *4:*52
Buffering capacity *4:*378
Buffering systems *3:*8
Buffer land *2:*168
Buffer mechanisms *3:*273
Buffer solutions *2:*295
Buffer strips *2:*3
Buffer zones *3:*476
Build-operate-transfer (BOT) contracts *1:*388
Build-own-operate (BOO) contracts *1:*388
Build-own-[operate]-[train]-[transfer] model *1:*50
Build-own-operate-transfer (BOOT) method *1:*174
Bulk density *4:*484
Bulking. *See* Microbial foaming and bulking
Bulk viscosity *5:*560
Bulk water detention time *1:*205
Buoyancy accelerations *4:*367
Bureau of Indian Standards (BIS) *5:*195
Bureau of Reclamation (BOR) *4:*687
Burg algorithm *4:*221
Burgers' flux law *3:*249 *3:*250
Burmister Soil Identification System (BS) *3:*690
Bushwalking *4:*788
Business sector, interests of *2:*626
Butler and Ockrent adsorption model *1:*107
Bycatch *3:*126

C

Cable-tool water well drilling *5:*105
Cadmium (Cd). *See also* Cd-resistant bacteria
  bioavailability in soils *5:*370
  dissolution and precipitation of *5:*616
  fate and transport of *5:*617
  geochemistry of *5:*616
  in groundwater *2:*148 *5:*613
  health risks of *5:*615

Index terms	**Links**	
Cadmium (Cd) *(Continued)*		
maximum contaminant level of	*2:*234t	
physical and chemical characteristics of	*5:*613	
phytoextraction of	*5:*369	
precipitation	*5:*619	
remediation of	*5:*617	
role in metal tolerance	*3:*611	
sorption and desorption of	*5:*616	
sources of	*5:*614	
species of	*5:*616	
terrestrial impacts of	*5:*615	
Cadmium hyperaccumulation, molecular basis of	*5:*371	
Cadmium ions, concentration of	*3:*298	
Cadmium levels, in marine biota	*4:*100	
Cadmium-zinc interactions	*5:*370	
Caffeine, as a dehydrating substance	*4:*724	
Caged animals, use at dredged sites	*2:*351	
^{14}C age dating method	*4:*389	
See also Carbon (C)		
Cage farms	*3:*580	
Caisson sand-abstraction system	*3:*417	
Calcite	*4:*408	*4:*415
solubility of	*4:*411	
solubility products from	*4:*408	
Calcium carbonate coating, formation of	*1:*10	
See also Lime entries		
Calcium carbonate scaling indexes	*4:*415	
Calcium hardness	*4:*453	
See also Mg/Ca ratio		
Calcium hypochlorite	*1:*457	
Calculated fixed radius (CFR)	*1:*525	
Calderon-Sher Safe Drinking Water Act (California)	*2:*345	
Calibration	*1:*134	
Calibration studies. *See* Hydraulic network models		
Calicivirus	*1:*279	
Calicivirus enteritis	*1:*179	
California		
canals in	*3:*87	
coastal wetlands in	*3:*72	
in-stream flow protection in	*4:*662	

Index terms	**Links**	
California *(Continued)*		
overdraft in	*5:*342	
source-water protection in	*2:*312	
water use in	*2:*478	
California Bay Delta Authority (CBDA) Contaminant Stressors Workshop	*3:*371	
California Department of Health Services (CDHS)	*2:*345	
California Department of Water, demand forecasting by	*2:*530	
California Fish & Game Department	*3:*518	
"California Rule"	*4:*631	
California Storm Water Quality Task Force	*3:*433	
California Toxics Rule	*2:*356	
Caliper log	*5:*154	
Calorimetry	*4:*401	
CALVUL statistical method	*5:*564	
Campylobacter	1:277	
Campylobacteriosis	*1:*179	
Canada. *See also* Environment Canada (EC)		
annual precipitation in	*2:*660	
contamination problems in	*5:*604	
continental watersheds in	*2:*657	
flood management in	*2:*681	
hydropower generation in	*3:*199	
infrastructure in	*2:*662	
water conservation in	*2:*660	
water misuse in	*2:*661	
water prices in	*2:*604	
water quality in	*2:*663	
water resources in	*2:*656	
water supply in	*2:*661	
water use in	*2:*661	
Canadian Environmental Protection Act	*2:*319	
Canadian households, water consumption in	*1:*506*f*	*1:*507f
Canadian landscape, role of water in	*3:*507	
Canadian rivers, suspended sediment load in	*3:*509	
Canal data	*5:*164	
Canal irrigation systems	*3:*581	*3:*583
Canals, Mekong Delta	*4:*748	
Canal seepage recharge (R_c)	*5:*166	
Cancer(s). *See also* Carcinogenic agents; Lung cancer		
arsenic-related	*1:*2	

Cancer(s) *(Continued)*
  gastric *1:*36
  vinyl chloride and *5:*636
Candida albicans *2:*240
Canopy reflectance models *3:*720
Canopy transpiration *4:*346
Capacity building *1:*651 *1:*655 *2:*628
  See also Integrated capacity building (ICB)
Capacity to inflow ratio (CIR) *3:*409
‘‘Cap and trade’’ program *2:*133 *2:*134
Capillary action *2:*542
Capillary flow, vadose zone *5:*73
Capillary forces, head and *5:*180
Capillary fringe zone, soil samples from *3:*688
Capillary lysimeter *2:*340
Capillary rise equation *3:*570 *5:*180
Capillary rise/pressure *5:*535
Capillary zone *5:*514
Capping systems *5:*433 *5:*617
Carbon (C). *See also* Activated carbon entries; ^{14}C age dating method; Carbon adsorption; Carbons; Coal entries; Dissolved organic carbon (DOC); Organic entries control of *2:*224
  density *1:*96
  effective size, mean particle diameter, and uniformity coefficient of *1:*97
  migrating *3:*180
  uptake capacity of *2:*83t
Carbon-13 *5:*231
Carbon-14 *5:*227 *5:*231
  See also ^{14}C age dating method; Radiocarbon
  in groundwater tracing *5:*503
Carbonaceous BOD (cBOD) *1:*667 *1:*754t
Carbonaceous material, water adsorption on *4:*402
Carbonaceous minerals *3:*14
Carbon adsorption *4:*381
  of landfill leachates *1:*707
Carbonate. *See also* Carbonate geochemistry
  alkalinity *3:*131
  biochemical aspects of *4:*414
  dissolution and deposition of *4:*414
  environmental impact of *4:*416

Index terms **Links**

Carbonate *(Continued)*
  in natural waters *4:*413
  as a pH buffer *4:*414
  precipitation *1:*705
  water treatment and utilization related to *4:*415
Carbonate alkalinity, calculation of *4:*470
Carbonate aquifers *5:*72
Carbonate geochemistry *4:*408
  calcite solubility products and *4:*408
  measured alkalinity and *4:*409
  saturation index and *4:*409
Carbonate-poor geological regions *3:*8
Carbonate rocks *5:*55
Carbonate scale control, with acid *4:*415
Carbonate sediments *4:*415
Carbonate species *4:*411
Carbon bed volume (V_b) *1:*103
Carbon Chloroform Extract (CCE)
  method *1:*297
  standard *1:*288
Carbon cycle, geochemical *4:*414
Carbon dating *5:*231
  See also C:N ratio
Carbon dechlorination *1:*169
Carbon dioxide (CO_2). *See also* CO_2 production
  atmospheric *4:*176
  as an atmospheric acid *3:*8
  global partitioning of *4:*85
  ocean storage of *4:*416
  residence time of *4:*85
Carbon dioxide uptake, ocean *4:*86
Carbon dioxide water *4:*799
Carbon fluxes *3:*204t
Carbon fractionation *4:*502
Carbonic acid, calcite solubility and *4:*412
Carbon isotopes *5:*220
  See also Carbon-13; Carbon-14; Radiocarbon
Carbonization (charring) *1:*94 *2:*80
Carbon regeneration *1:*104

Index terms **Links**

Carbons, mechanisms of water adsorption on *4:*400
  See also Carbon (C)
Carbon stocks, biogenic *3:*547
Carbon surface chemical composition, effect on phenol adsorption *4:*404
Carbon usage rate (CUR) *1:*103
Carbonyl sulphide (COS) *4:*88
Carcinogenic agents *1:*180
  See also Cancer(s)
  arsenic and arsenic compounds as *2*:7 *2*:17
  luminescent bacteria for determining *2:*174
Carcinogenic byproducts *2:*116
Carcinogenic effects, advisories based on *3:*119
Carcinogenic responses *4:*674
Cardiovascular effects, of lead *2:*436
Caribbean basin, oil in *4:*99
Carp breeding, origins of *4:*718
Carp culture *3:*137
Carson, Rachel *3:*514
"Cartographic anxiety" *4:*683
Cartridge filters *1:*247 *1:*488
  advantages and disadvantages of *2:*155
  for iron removal *2:*152
Cartridge filtration *1:*229 *1:*232 *1:*459
Cascades *3:*70
''Cash-for-grass'' programs *3:*556
Casing *1:*149
Caspian Sea level rise *2:*480
  climate change and *2*:482
  hydraulic construction and *2*:482
  nature of *2*:481
  societal impacts of *2*:482
  tectonic plate movement and *2*:482
Catadromous fish *3:*127
Catalysis, activated carbon in *1:*105
Catchment management *3:*33
Catchment water quality modeling software *2:*325
Catenary grid systems *1:*353
Cathodic protection *1:*154
  to limit corrosion *1:*11
Cation exchange, radionuclide removal via *1:*397

Index terms **Links**

Cation exchange capacity (CEC) *1:*778 *5:*261 *5:*278
Cationic/anionic solubility relationships *5:*281
Cationic organoclay *1:*779 *1:*780t
Cationic polymers, adsorption of *1:*118
Cationic surfactants *1:*670 *1:*672
Cauchy boundary condition *5:*14
Caustic soda, in acid mine treatment *5:*4
Caustic soda softening *1:*323
Cave formation *4:*412
Cave passageways *5:*245
Cavitation *1:*484 *3:*489
  acoustic *1:*579
C climates *4:*258
CDF proteins *5:*286
cDNA libraries *2:*62
  See also Complementary DNA (cDNA)
cDNA microarrays *2:*63
Cd-resistant bacteria *2:*447
  See also Cadmium (Cd)
Cell biology, effect of magnetization and water on *4:*456
Cell construction, in landfills *1:*697
Cell culture-IFA, for measuring inactivation of UV-treated *C. parvum* oocysts *1:*167
Cell culture polymerase chain reaction (CC-PCR) *1:*161
Cell lines, in sediment toxicity testing *3:*54
Cell membrane, metal transport across *2:*70
Cells
  electron stability of *4:*457
  heavy metal binding to *5:*378
  integrity of *4:*460
  passage of water in and out of *4:*458
Cell stability, relationship of magnetism to *4:*457
Cellular biomarkers *2:*31
Cellular sheet piling *4:*19
Cell wall binding, of metals *5:*283
Cement-based binders *1:*836
Cement hydration product leaching *5:*364
  effect on groundwater pH *5:*365
Cementitious coatings *1:*882
Cementitious liners *1:*883
Cement mortar pipe linings *1:*879 *1:*880

Index terms **Links**

Cement pastes, dissolution of *5:*362
Cement replacements *5:*363
Cements, rapid setting hydraulic *1:*882
Center for Watershed Protection *3:*432
Centers for Disease Control (CDC) *2:*343
  safe water system *1:*26
Central America, transboundary waters in *4:*641
Central Asian republics
  recession and social unrest in *3:*19
  water disputes among *3:*18
Central Ground Water Authority (CGWA) *2:*181
Centralized wastewater treatment *1:*678
Central nervous system, effect of lead on *2:*435
Central Pollution Control Board (CPCB), ‘‘designated best use’’ yardstick of *2:*183 *2:*184
Central Valley, California, overdraft in *5:*342
Centrifugal pumps *1:*392 *3:*380
  operation of *1:*394
Centrifugal separation, wastewater *1:*810
Centrifuge, in sludge treatment *1:*855
CE-QUAL-RIV1 software *2:*328
CE-QUAL-W2 software *2:*328
Ceramic membranes *1:*591
Certifying authority *1:*293
Cesspit waste, disposal and agricultural use of *1:*843
CFC dating method
  versus H-He dating *5:*68
  versus ^{85}Kr dating method *5:*249
CH4, global role and sources of *4:*86
Chadha’s diagram *2:*396 *2:*397f
Change, equations of *5:*558
‘‘Changing mean’’ models *3:*215
Channel cross sections, geometric properties of *3:*347t
Channel flow routing
  kinematic wave formulation for *3:*253
  in unit hydrograph models *4:*358
Channel heads, identification of *3:*93
Channels
  networks of *3:*32
  reaches of *3:*31 *3:*67
  side slopes of *3:*348t

Index terms	**Links**	
Channels *(Continued)*		
slope of	*3:*438	
systems of	*3:*31	
terrace systems of	*3:*567	
time of travel in	*3:*244	
Channel structures, in biofilm	*2:*230	
Channel tests, with alfalfa	*5:*392	
Channel units	*3:*67	
nonturbulent and turbulent	*3:*70	
Chara (shrimp grass)	*3:*276	
Charge neutralization	*4:*425	
Check dams	*2:*550	*3:*412
Check valves	*1:*482	
Chelates, synthetic	*5:*383	
Chelating agents	*5:*282	
synthetic	*5:*383	
Chelation	*3:*609	*5:*377
CHEMFLO-2000 program	*5:*141	
Chemical adsorption/desorption	*1:*725	
Chemical agents	*4:*295	
Chemical analysis, of rainwater	*4:*372	
Chemical analyzers	*4:*514	
Chemical carbon activation	*1:*95	
Chemical coagulants	*4:*424	
Chemical coagulation process	*2:*374	
Chemical compounds, use in acid mine drainage	*5:*4t	
See also Chemicals; Compounds		
Chemical concentration. *See also* Concentration techniques		
information related to	*4:*600	
measurement of	*2:*59	
Chemical contaminants	*2:*344	
See also Contaminants		
bioaccumulation of	*2:*125	
Chemical control, of biofilms	*1:*541	
Chemical CSO treatment	*1:*786	
Chemical degradation, of pesticides	*3:*648	
Chemical destabilization	*1:*231	*1:*246
Chemical drinking water standards	*1:*529	
Chemical exchange, in heavy water production	*4:*465	
Chemical factors, in episodic acidification	*3:*5	

Index terms **Links**

Chemical fingerprinting *3:*169
Chemical-free mussel control *1:*510
‘‘Chemical-free’’ water treatment *1:*330
Chemical industry, activated carbon in *1:*105
Chemical information, incorporating
  into water quality evaluation *4:*602
Chemically enhanced flushing methods *5:*440t
Chemically enhanced primary treatment (CEPT) *1:*659
Chemical monitoring *2:*267
  with bivalve molluscs *2:*35
  in situ *4:*514
Chemical monitoring reform (CMR) *4:*678
Chemical oxidation *1:*370 *2:*300
  See also Oxidation entries
  of landfill leachates *1:*706
  of vinyl chloride *5:*638
Chemical oxidation technologies
  application practicality of *5:*347
  Fenton’s Reagent treatment *5:*345
  for groundwater remediation *5:*344
  health and safety precautions for *5:*348
  ozone treatment *5:*347
  permanganate treatment *5:*346
Chemical oxygen demand (COD) *1:*519 *1:*623 *1:*701 *1:*733 *2:*40
  of dishwater *2:*113 *2:*114
Chemical pollutant removal processes *3:*366
Chemical pollutants
  pre-1940 *1:*529
  present and future *1:*531
Chemical precipitation *4:*586
  of landfill leachates *1:*705
  of radioactive waste *1:*804
  of uranium *5:*641
  uses of *4:*586
Chemical pretreatment *1:*250
Chemical reactions, as watershed functions *3:*477
Chemical reactors, sewers as *3:*331
Chemical regeneration *1:*917
Chemical rehabilitation *3:*125

Index terms	**Links**
"Chemical revolution"	*1:*286
Chemicals. *See also* Chemical compounds; Compounds	
in an agricultural landscape	*3:*607
bioavailability of	*2:*427
biodegradation potential for	*3:*390
effect on corrosion	*1:*8t
impact of	*2:*28
microirrigation and	*3:*617
odorous	*1:*760t
partitioning and bioavailability of	*4:*521
solubility in water	*4:*555
unregulated	*3:*371
uptake and elimination of	*3:*35
weed control via	*3:*745
Chemical sensors	*4:*514
impairment of	*2:*283
residence time of	*4:*159
Chemical standards, U.S.	*2:*268
Chemical techniques, for base flow separation	*3:*27
Chemical toxicants, mechanisms of action of	*3:*117
Chemical treatment	
of acid mine drainage	*5:*3
for iron bacteria	*2:*151
of soil	*5:*434
of sludge	*1:*857
Chemical wastewater treatment technologies	*1:*811
Chemigation	*3:*623
Chemiometabolic enzyme systems	*3:*117
Chemisorption	*1:*99
Chemistry	
of coastal waters	*4:*25
marine	*4:*57
in transport modeling	*5:*31
Chemoautotrophic nitrification	*3:*641
Chemograph analysis	*5:*239
Chemographs	*5:*240
Chemosynthetic bacteria	*2:*21
Chernaya Bay, underwater weapon tests in	*4:*102
Chesapeake Bay watershed, nutrient trends in	*3:*110
Chezy formula	*3:*195

Index terms	**Links**
Chicago, Illinois, overdraft in	*5:*342
Chicago River reversal	*3:*41
channel work related to	*3:*43
commission formed for	*3:*42
Joliet project and	*3:*44
South Branch improvements and	*3:*44
Chicago Station, SAMSON Database statistics for	*5:*79t
Chick's law	*1:*193
Chihuahuan Desert, recharge in	*5:*74
Children's Groundwater Festival	*4:*713
Children's water rights	*4:*771
Chile	
tradable rights in	*2:*645
water markets in	*2:*645
water reallocation in	*2:*607
China	
legal system in	*2:*486
policy, laws, and regulations in	*2:*487
water management financial	
sources in	*2:*487
water management in	*2:*484
water management institutions in	*2:*484
water scarcity in	*2:*488
water symbolism in	*4:*786
Chinampas	*4:*697
Chinese water consciousness	*4:*707
Chinook	*4:*170
Chironomids	
metal concentrations in	*2:*421
in sediment-based endocrine disruption	*3:*54
in sediment toxicity testing	*3:*50
species of	*3:*51
Chironomus	
future sediment toxicity testing using	*3:*54
heavy metal uptake by	*2:*215
Chironomus riparius	
biology of	*3:*51
in life cycle testing	*3:*52
as a test organism	*2:*420

Index terms **Links**

Chironomus tentans
  biology of *3:*51
  in life cycle testing *3:*52 *3:*53
Chloraminated systems, nitrification in *2:*226
Chloramines *1:*195 *1:*198 *1:*457 *1:*471 *2:*89 *2:*118
  advantages of *2:*244
Chlorate, health effects of *1:*270
Chlordane *3:*352
  residues *3:*649
Chloride. *See also* Chlorine (Cl)
  brines *5:*52
  concentration *5:*197
  in drinking water *2:*395
  levels *2:*321
  mass-balance methods *5:*74
Chloride ions, atmospheric *4:*373
Chlorinated aliphatic compounds *1:*688 *5:*579
  degradation of *1:*690
Chlorinated alkene cometabolism *1:*690
Chlorinated aromatic compounds *1:*689 *1:*691
  degradation of *1:*688
Chlorinated dibenzofurans (CDFs) *2:*107
  chemical properties of *2:*109t
Chlorinated dibenzo-*p*-dioxins (CDDs) *2:*107
  chemical properties of *2:*109t
Chlorinated disinfection byproducts (DBPs) *2:*115
Chlorinated ethene contaminants *5:*82
Chlorinated hydrocarbons *2:*109 *2:*266 *5:*322
Chlorinated methanes *5:*529
Chlorinated solvents *2:*317 *5:*91 *5:*116 *5:*117 *5:*581
  biodegradation of *5:*584
  biotreatment of *5:*117
  degradation of *5:*587
  electrokinetic treatment of *5:*123
  Fenton's Reaction and *4:*446
  in situ electrokinetic treatment of *5:*116
  leaking vapors from *5:*552
  metabolism of *5:*529
Chlorinated trihalomethanes *2:*77

Index terms **Links**

Chlorinated xenobiotic compounds, degradation of *5:*47
Chlorination *2:*88 *2:*374
  advantages and disadvantages of *2:*398
  application and design of *2:*90
  break-point *1:*471
  carcinogenic byproducts of *2:*398
  chemistry of *2:*88
  domestic sewage *1:*834
  Escherichia coli O157:H7 susceptibility to *2:*138
  factors affecting *2:*89
  gas *1:*197
  for hepatotoxins *2:*390
  increasing *4:*524
  introduction of *1:*286
  products of *2:*89
  sodium hypochlorite solution *1:*198
  solid calcium hypochlorite *1:*198
  systems for *1:*912
Chlorination byproducts *2:*91
  health effects of *2:*92
Chlorinator *1:*197f
Chlorine (Cl) *1:*127
  See also Chloride; Chlorine (gas); Chlorine reactions; Cl. contamination; Cl- phenols; Dechlorination; Free chlorine entries
  compounds of *2:*88
  disinfection using *2:*244
  dose response of mussels to *2:*401
  effect on corrosion *1:*8
  efficacy as an antifoulant *2:*402
  estimating *1:*128
  health effects of *2:*92
Chlorine (gas) *1:*457
  See also Free chlorine
  in leachates *1:*706
  safety of *2:*90
Chlorine-36 *5:*232
Chlorine-37 *5:*220
Chlorine byproducts, nonoxidizing *1:*129
Chlorine demand (CD) *1:*128
Chlorine dioxide *1:*194 *1:*214 *1:*458 *2:*119
Chlorine disinfection, at CSO facilities *1:*786

Index terms	Links		
Chlorine isotopes, fractionation in	*4:*502		
See also ^{36}Cl			
Chlorine-produced oxidants (CPO)	*1:*128		
Chlorine reactions			
kinetics of	*1:*131		
parameter estimation for	*1:*134		
Chlorine residual modeling	*1:*131		
See also Chlorine transport model			
dynamic-state	*1:*133		
steady-state	*1:*132		
within a water distribution system	*1:*132		
Chlorine residuals	*1:*128	*2:*398	
See also Chlorine residual modeling			
environmental discharge limits of	*1:*130		
measurement of	*1:*130		
Chlorine simulation model	*1:*136		
Chlorine toxicity, factors influencing	*2:*402		
Chlorine transport model	*1:*133		
Chlorite, health effects of	*1:*269		
Chloroacetic acid, health effects of	*1:*266		
Chloroacetonitrile, health effects of	*1:*267		
Chloroanilines, degradation of	*1:*689		
Chlorofluorocarbons (CFCs)	*4:*185	*4:*420	*5:*232
calculating the age of	*4:*421		
concentration of	*4:*420		
dating technique limitations and errors related to	*4:*422		
effect of	*4:*184		
in groundwater studies	*4:*422		
measurement laboratories for	*4:*423		
sampling groundwater for	*4:*421		
Chloroform	*2:*92		
health effects of	*1:*265		
Chloromethane, health effects of	*1:*264		
Chloro-organics, degradation of	*1:*688		
Chlorophyll absorption in reflectance index (CARI)	*3:*722		
Chlorophyll-α concentrations (CHL), in water quality models	*2:*262		
Chlorophyll fluorescence (CF)	*3:*722		
Chloropicrin, health effects of	*1:*269		
Choked flow	*1:*484		

Index terms	Links	
Cholera	*1:*24	*1:*179
outbreaks of	*1:*283	*1:*290
Cholesterol, vital roles of	*4:*702	
Cholesterol lowering, role of water in	*4:*701	
Chromated copper arsenate (CCA)	*1:*83	
Chromates, sorption of	*1:*588f	
Chromatographic columns, standard packed	*2:*304	
Chromium (Cr)	*5:*290	
See also Hexavalent chromium		
in groundwater	*2:*147	
in situ electrokinetic treatment of	*5:*28	
maximum contaminant level of	*2:*234t	
in soluble oxyanionic complexes	*5:*281	
treating in water supply wells	*5:*22	*5:*25
Chromophores	*4:*533	
organic	*4:*530	
Chronic acidification	*3:*1	
relationship to episodic acidification	*3:*5	
waters sensitive to	*3:*3	
Chronic bioassays	*2:*351	
Chronic dehydration	*4:*722	
Chronic sediment toxicity testing	*3:*52	
"Chronic water scarcity"	*2:*495	
Cienega Creek, Arizona	*3:*58	*3:*59f
Cienega de Santa Clara, Colorado River Delta	*3:*58	
Cienegas	*3:*57	
environmental conditions associated with	*3:*57	
examples of	*3:*58	
hydrologic conditions related to	*3:*57	
terms related to	*3:*59	
Ciliated protists, as toxicity assessment test organisms	*2:*413	
Ciliates	*2:*313	
Ciliophora	*2:*313	
Circular Poiseuille flow	*5:*654	
Circular settling tanks	*1:*453	
Circular water meter register	*1:*339	
Cities		
real-time hydrological information system for	*1:*121	
water demand in	*2:*490	

Index terms	**Links**		
Civilizations			
decline of	*4:*739		
irrigation and	*4:*737		
Civil War, effect on water projects	*2:*522		
^{36}Cl. *See also* Chlorine isotopes			
case studies on	*4:*418		
dating method for	*4:*417		
measurement of	*4:*417		
old groundwaters and	*4:*416		
production of	*4:*417		
Cladocera	3:273		
^{36}Cl age dating method	*4:*389		
Clapeyron equation	*4:*264	*4:*365	
Clarification, water treatment plant	*1:*224		
Clarifiers			
design of	*1:*815		
modeling	*1:*733		
Clathrate hydrates	*4:*471		
in natural and industrial environments	*4:*473		
Clark hydrograph	*3:*60	*3:*61	
Clark rainfall-runoff transformation	*3:*60		
See also Isochrone spacing methods			
parameters of	*3:*61		
principles of	*3:*60		
Class ''A'' pathogen reduction	*1:*649		
Classification, of agricultural land	*3:*608		
Classification system, for wetlands and deepwater habitats	*3:*496		
Clathrate hydrate structure complex	*4:*472		
Clathrate hydrate structures	*4:*471		
novel	*4:*472		
Clausius-Clapeyron equation	*3:*193	*4:*366	*4:*369
Clausius-Clapeyron method	*4:*272		
Clay minerals, trace elements in	*3:*455		
Clays, in membranes	*1:*418		
See also Organoclay			
Clay-supply industry, impact on surface water quality	*3:*374		
Cl. contamination	*3:*2		
Clean Air Act Amendments, acid rain under	*4:*377		
Clean Lakes Program	*2:*315		

Index terms **Links**

Clean Water Act (CWA) *1:*755 *2:*96 *2:*194 *2:*292 *2:*315 *4:*40 *4:*595 *4:*671
  amendments and water protection laws under *4:*597
  effluent limitations and water quality standards under *1:*755 *4:*596
  goals of *2:*127
  implementation procedures for *4:*670
  improvements under *2:*191
  1977 and 1987 amendments to *4:*654
  nonpoint source control regulations under *2:*185
  numeric effluent system of *1:*756
  plans, grants, and funding under *4:*597
Clean Water Act regulations, early *4:*598
Clean water plants *1:*660
Clean Water Restoration Act of 1966 *4:*653
Clear water zone (CWZ) depth *1:*452
Clepsydras *4:*704 *4:*705
Climate
  Columbia Basin *4:*293
  extremes in *4:*175
  Great Lakes *3:*177
  of Hawaii *4:*255 *4:*802
  “normal” *4:*187
  plankton influence on *4:*157
  role of ocean in *4:*21
  sea level and *4:*117
  society and *4:*183
  of the stratosphere *4:*182
  understanding *4:*169
Climate change *2:*600
  Caspian Sea level rise and *2:*482
  coping mechanisms for *4:*185
  detecting *4:*177
  effect on aquatic environments *3:*124
  global *4:*171
  influence on the hydrologic cycle *4:*282
  observations of *4:*172
  water cycle and *4:*193
Climate models, evaluating *4:*172
Climate prediction, long-term *4:*215
Climate-related transboundary issues *4:*185

Index terms **Links**

Climate system *4:*179
  sea ice and *4:*69
Climatically-active trace-gas cycles, role of oceans in *4:*85
Climatic conditions, in the Mediterranean islands *2:*642
Climatic issues, interest in *4:*92
Climatic zones, lake temperatures and *3:*266
Climatologists *4:*328
Climatology *4:*186
  drought and *4:*187
Clocks
  early *4:*702
  elements of *4:*703
Clofibrate *1:*373
Clogging, in diffused air aeration systems *1:*630
Close-fit lining *1:*878
Closure, of municipal solid waste landfills *2:*168
Cloth-Plate Method (CPM) *4:*202
Cloud effects *4:*169
Clouds *4:*293
  long-wave radiation from *3:*192
  understanding *4:*169
Clouds and the Earth's Radiant Energy System (CERES) *4:*169 *4:*194 *4:*216
  commercial applications of *4:*170
  educational outreach and *4:*170
  future missions of *4:*170
Cloud seeding *4:*187
  effectiveness of *4:*188
  future of *4:*188
Cl-phenols, microbial detection of *2:*442
Clumping *1:*667 *1:*668
Cnoidal wave theory *4:*135
C:N ratio *5:*329
CO_2 production *3:*181
Coagulant dosing, water treatment plant *1:*223
Coagulant recovery, from water treatment plant residuals *1:*139
Coagulants *4:*425
Coagulation *1:*137 *1:*370 *1:*370 *1:*811 *2:*98 *4:*424 *4:*586
  disinfection byproduct precursor removal by *2:*116
  efficiency of *4:*427

Coagulation *(Continued)*
  enhanced *2:*120
  of landfill leachates *1:*705
  mechanisms of *1:*138 *4:*424
  ozone effects on *1:*355
  particulate matter removal by *1:*137
  in precursor minimization *2:*116
Coagulation/filtration, for arsenic treatment *5:*24
Coagulation-precipitation processes *1:*706
  arsenic removal *1:*636
Coagulation/sedimentation process *1:*816
  domestic sewage in *1:*834
Coal ash byproduct contaminants, from electric generating plants *1:*556
Coal combustion *1:*555
  contaminants released by *1:*555t
Coal extraction, environmental degradation from *1:*553
  See also Coal mining
Coal mining, impact on surface water quality *3:*374
  See also Coal extraction
Coarse gravel filtration *1:*239
Coarse screens *1:*784
Coastal aquifers *5:*497
  groundwater contamination in *5:*223
Coastal fisheries
  management *3:*126
  research on *4:*55
Coastal fog. *See also* Fog
  along the Northern Gulf of Mexico *4:*230
  conditions that help *4:*231
  scales occurring with *4:*230
  sounding profiles present during *4:*237
  variables present in production of *4:*232
Coastal hurricane forecast system *4:*275
Coastal regions, salt water intrusion in *3:*678
Coastal resources *2:*489
Coastal shelf waters *4:*26
Coastal water pollutants *4:*96
Coastal waters *4:*23
  arsenic in *2:*11
  CH_4 concentrations in *4:*87

Index terms	**Links**
Coastal waters *(Continued)*	
chemical characteristics of	*4:*25
importance of	*4:*26
indicators of organic contamination in	*2:*442
physical processes in	*4:*23
types of	*4:*25
Coastal watershed, coral reefs and	*4:*113
Coastal wetlands	*3:*71
See also Wetlands	
ecological importance of	*3:*72
loss trends related to	*3:*72
research related to	*3:*73
Coastal Wetlands Planning, Protection, and Restoration Act	*3:*73
Coastal Zone Act Reauthorization Amendments	*2:*185
Coast Guard Maritime Differential GPS (DGPS) system	*4:*249
CoastWatch program (NOAA)	*4:*55
Coatings	
cementitious	*1:*882
pipe	*1:*154
protective	*1:*882
Cobalt (Co)	
chemical reactions with	*5:*611
contamination by	*5:*612
fate and transport of	*5:*611
in groundwater	*5:*610
health and safety of	*5:*612
isolation of	*5:*611
physical properties of	*5:*611
remediation techniques for	*5:*612
sources of	*5:*610
COD relationships, in modeling	*1:*735
See also Chemical oxygen demand (COD)	
Coefficient of storage	*5:*492
Coefficient of variation (COV)	*5:*238
Cold cloud seeding	*4:*187
Cold fronts, fog and	*4:*231
Cold vapor atomic absorption spectrophotometry	*3:*119
Cold-water corals	*4:*32
distribution of	*4:*34

Index terms	Links		
Coliform bacteria	*1*:661	*5*:200	*5*:407
See also Escherichia coli			
in drinking water	*1*:661		
removal performance for	*1*:663	*1*:664f	
strains of	*1*:662t		
Coliform counts	*1*:662	*1*:664f	
Coliform group	*1*:293	*1*:295	
as indicator organisms	*2*:293		
Coliforms, in groundwater	*2*:396		
See also Escherichia coli; Fecal coliforms (FC)			
Colitis			
hemorrhagic	*1*:180		
shigellosis	*1*:181		
Collapsing breakers	*4*:137		
Collecting systems, alternative	*1*:678	*1*:840	
Collection, as a watershed function	*3*:473		
Collector wells, horizontal	*5*:88		
Colleges, land-grant	*3*:597		
Collision-induced dissociation (CID) technique	*4*:443		
Colloidal fouling	*1*:415		
Colloidal processes			
adsorptive	*5*:350		
mechanical	*5*:350		
Colloidal silica, in membranes	*1*:418		
Colloidal stability	*4*:424		
Colloid compounds removal	*1*:684		
Colloids	*2*:99		
biophysicochemical interactions with filters	*2*:100		
freshwater	*3*:74		
marine	*4*:27		
properties of	*3*:74		
techniques for characterizing	*3*:74		
Colloid science, role of dissolved gases in	*4*:451		
Colloid transport	*5*:349		
models	*1*:368		
Colluvial valleys	*3*:66		
Color, water	*1*:901		
Color abatement, using ozone	*1*:355		
Colorado			
in-stream protection in	*4*:660		

Index terms **Links**

Colorado *(Continued)*
  overdraft in *5:*341
  Standardized Precipitation Index use in *4:*210
Colorado Drought Plan *4:*213
Colorado River *2:*555
  See also Cienega de Santa Clara, Colorado River Delta
  salinity profile of *3:*680
  United States-Mexico conflict over *4:*636
Color-coding, in reclaimed water systems *1:*808
Colorimetric assays *2:*415
Colorimetric titration method *4:*453
Colorimetry, in pH measurement *2:*295
Colpidium campylum, as a test organism *2:*413
Columbia River *2:*554 *2:*555
Columbia River Basin *2:*589
  climate of *4:*293
Column experiments/studies *1:*723 *1:*772
  categories of *2:*103
  column preparation in *2:*103
  in contaminant transport *2:*103
  examples of *2:*105
  procedure for *2:*104
Columns
  dimensions of and materials for *2:*104
  nutrients in *3:*277
Comanagement *3:*125
Combined chlorine/combined available chlorine *1:*128
Combined chlorine residual *2:*398
Combined sewer overflows (CSOs) *1:*496 *1:*497 *1:*748 *1:*782 *1:*799 *3:*281 *3:*332
  See also CSO entries
Combined sewer overflow treatment *1:*782
  biological *1:*787
  chemical *1:*786
  physical *1:*783
Combined sewers *3:*331
Combustible watersheds,
  conflagration prediction for *3:*461
Cometabolic oxidation *2:*300

Index terms		**Links**
Cometabolic processes, for bioremediation of chlorinated solvents and petroleum hydrocarbons	*5:*117	
Cometabolism	*2:*42	*5:*584
Comets, water ice in	*4:*190	
‘‘Command and control’’ regulation	*2:*132	
Command area studies, remote sensing and GIS application in	*2:*533	
Commerce, importance of coastal waters to	*4:*27	
Commercial activated carbon (CAC)	*4:*495	
Commercial fishing	*3:*123	*3:*533
Commercial meters	*1:*340	
Commercial storm water discharges	*1:*867	
Commercial water use	*2:*542	
in the United States	*2:*652	
Comminution	*1:*814	
domestic sewage	*1:*832	
Commission on Drainage and Water Supply	*3:*42	
Commissions, inland water, river basin, and lake catchment	*3:*293t	
Committee on the Biological Effects of Ionizing Radiation (BEIR) VI	*4:*543	
Commodification	*1:*42	*1:*216
Common law	*4:*630	
Common-Pool Resources (CPR)	*2:*625	
Communication, role in fish consumption advisories	*3:*120	
Community, water conservation in	*2:*664	
Community analysis, molecular probes in	*1:*643	
Community-based fisheries management	*2:*636	
Community-based organizations (CBOs)	*1:*655	
Community-based resource management	*2:*629	
Community-based water management	*2:*628	
Community groups, involvement in source water assessment	*1:*447	
Community involvement		
in groundwater protection	*2:*518	
strategies for	*2:*519	
Community profiling methods	*1:*643	
Community water systems (CWSs)	*4:*676	
Community Water Supply Study (CWSS)	*1:*289	
Competing ligand exchange technique	*2:*420	

Competitive adsorption
  on activated carbon *1:*107
  statistical design for *1:*109
Competitive sorption
  of inorganics on activated carbon *1:*110
  of organics *1:*115
Complementary DNA (cDNA) *2:*59
  See also cDNA entries
Complete mix reactor *1:*815 *1:*833
Complexation constants *2:*206
Complexation/flocculation *1:*872
Complexation reactions *1:*725
Complex chlorinated hydrocarbons *5:*43
Complex fluids *4:*448
Complexing agents *5:*281
Complexing ligands, destruction of *5:*283
Complex waste waters, mercury removal from *1:*722
Compliance technologies *1:*457
  for the Total Coliform Rule *1:*463t
Composite materials, sorption characteristics of *2:*364
Composite programming (CP) *2:*521
Composting
  of biosolids *1:*648
  of manure *2:*572
  of sludge *1:*857 *1:*865
Compost toilets *1:*679
Compound meters *1:*338
Compounds. *See also* Chemical compounds; Chemicals; Organic compounds
  in groundwater *2:*316
  most frequently detected *5:*61f
  partition variables for *4:*523t
Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) *2:*96 *5:*368
Comprehensive models *2:*272
Comprehensive Soil Classification System (CSCS) *3:*690
Computer codes, for geochemical modeling *5:*140
Computer models *1:*122
  See also Computer simulations
  water-management *2:*620
Computer simulations, for vulnerability assessment *5:*563

Index terms **Links**

Concentrated recharge *5:*236
Concentrates *1:*413
  disposal methods for *1:*174t
Concentrating solar collector *1:*66
Concentration measurement *2:*59
Concentration techniques *2:*102
  See also Chemical concentration
"Concept of Ecotones" *3:*292
Conceptual models *3:*328 *3:*337 *5:*589
Concession contract model *1:*50
Concessions, private sector *1:*388
Concrete caissons *4:*19
Condensation *4:*188 *4:*196
  defined *2:*542
  for odor abatement *1:*762
Condensed organic matter domains, sorption in *4:*386
Conditional entropy *4:*219
Conditioning, of sludge *1:*864
Condominial sewerage *2:*633
Conducting polymers *1:*872
Conductivity, electric *4:*429
Conduit flow springs *5:*239
Conduit flow systems *5:*237
Cone penetrometer (CPT) rigs *5:*455
Confined aquifer flows *5:*493
Confined aquifers *5:*11 *5:*145 *5:*481 *5:*497 *5:*600
  barometric efficiency of *4:*167
  drawdown in *5:*103
  miscible displacement in *5:*161
  storage in *5:*515
  storativity in *5:*129
''Confined'' disposal of dredged sediments *2:*124
Conflagration prediction, for combustible watersheds *3:*461
Conflict management, in the Middle East *4:*755 *4:*756 *4:*757
Conglomerates *5:*55
Conjunctive use management *4:*632
Connate water *5:*54
Conservation *1:*509
  in the Arab World *2:*472

Index terms	**Links**		
Conservation *(Continued)*			
industrial	*2:*496		
water	*1:*307		
water cycle and	*4:*433		
Conservation aquaculture	*3:*128		
See also Species and habitat preservation/protection			
Conservation laws, for fluid flow systems	*5:*651		
Conservative pollutants	*4:*50		
Consolidated water-bearing rocks	*5:*55		
Constant-head boundary	*5:*419		
Constant-rate pumping test	*5:*574		
Constituents of concern (COC)	*5:*344		
Constructed wetlands	*1:*787	*1:*843	*1:*892
See also Wetlands			
in acid mine treatment	*5:*5		
advantages and disadvantages of	*3:*366		
algal control via	*2:*3		
application of	*1:*896		
costs of	*1:*843		
for mine wastewater treatment	*1:*897		
phytoremediation by	*3:*364		
pollutant removal mechanisms of	*3:*366		
types of	*1:*894	*3:*365	
water quality improvement in	*1:*892		
Constructed wetland systems, for storm water treatment	*1:*868		
Construction			
stormwater discharges from	*4:*659		
well	*5:*87		
Construction activities, stream quality and	*3:*442		
Construction and operating agencies, federal	*4:*650t		
Construction companies, service concession operation by	*1:*390		
Construction Grants Program	*4:*654		
Construction materials, for drinking water storage facilities	*1:*410		
Construction sites, water impacts from	*1:*537		
Consumer confidence reports (CCRs)	*1:*145	*2:*346	
Consumer information regulations	*4:*678		
Consumption advisories	*3:*118		
Consumptive use, defined	*2:*542		
Consumptive water use	*1:*561		
in the United States	*2:*647		

Contact angle hysteresis *5:*536
Contact flocculation *1:*254
Contactor hydrodynamics, role in bromate ion formation *1:*361
Contact precipitation *4:*437
Contact stabilization *1:*815
Contact time (CT) *1:*354
Containment
  as a remediation strategy *2:*357
  of vinyl chloride *5:*639
Contaminant distributions, vertical characterization of *5:*80
Contaminant hydrogeology environmental isotopes in *5:*217
  isotopes in *5:*216
Contaminant mixtures, effects of *2:*429
Contaminant monitoring regulations, unregulated *4:*678
Contaminant movement and transformation
  information needed for predicting *5:*579t
  role of perched groundwater in *5:*354
Contaminant-plant-bacteria interactions *5:*378
Contaminant plumes *5:*47 *5:*255
  vertical delineation of *5:*85
Contaminant release, from redeposited sediment *2:*125
Contaminant removal, thermal
  treatment methods for *5:*441t
Contaminants. *See also* Chemical contaminants; Contamination
  attenuation of *5:*56
  in bioproductivity zones *4:*105
  categories of *5:*216
  characteristics of *5:*57t
  chemical properties of *1:*54t
  cometabolism of *5:*379
  degradation modes of *5:*581
  degradation products of *5:*32
  from electric generating plants *1:*553
  field sampling and monitoring of *2:*263
  globally distributed *4:*97
  groundwater *2:*183
  inorganic *1:*463t
  investigations of *3:*169
  lipophilic *2:*170
  macrophyte uptake of *1:*715

Index terms	**Links**	
Contaminants *(Continued)*		
metabolic pathways of	*5:*529	
microbial processes affecting attenuation of	*5:*578	
microorganism	*1:*54	
natural	*2:*181	
radioactive	*5:*221	
reducing through reverse osmosis	*5:*25	*5:*326
reduction processes for	*5:*442	
regulatory standards for	*2:*346t	
in sediments	*3:*50	
sensitivity to	*2:*409	
sewage	*1:*829	
spills of	*5:*439	
toxicity of	*5:*57	
use of biodegradation to remove	*2:*42	
Contaminant sensitivity		
assessing	*5:*57	
limitations of	*5:*57	
Contaminant transport		
column experiments in	*2:*103	
mathematical model of	*5:*517	
modeling	*5:*30	
Contaminant vapors		
downward transport of	*5:*551	
transport of	*5:*550	
transport rates of	*5:*550	
Contaminated groundwater. *See also* Groundwater contamination		
in situ bioremediation of	*5:*38	
organoclay/carbon sequence for treating	*1:*777t	
Contaminated sediments, characterizing	*2:*460	
Contaminated sediment testing, species appropriate for	*2:*427t	
Contaminated soil. *See also* Soil contamination		
excavation of	*3:*652	
land application of	*3:*652	
remediation options for	*3:*652	*5:*432
selection of the remedial approach for	*3:*654	
Contaminated waters, remediation of	*5:*278	
Contaminated water sources, arsenic removal from	*1:*636	
Contamination. *See also* Contaminants		
aquifer sensitivity to	*5:*56	
Cryptosporidium	1:164	

Index terms **Links**

Contamination *(Continued)*
  drinking water distribution system *1:*342
  groundwater *2:*182 *5:*56
  karst area *5:*247
  mathematical formulation for *5:*663
  mercury *1:*722
  prevention of *1:*438
  septic system *1:*62
  susceptibility to *1:*446
  technetium-99 *4:*580
  vulnerability to *2:*313 *5:*247
Contamination sources inventory *1:*446
Continental effect *4:*439 *5:*230
Continental rifts *3:*20
Contingency planning (CP) *1:*527
Contingent valuation (CV) method *2:*130
Contingent water valuation method (CVM) *4:*607
Continually stirred tank reactors (CSTRs) *1:*740
Continuity equation *3:*197 *5:*481 *5:*558 *5:*651
Continuous chlorination *2:*398
Continuous flow analysis (CFA) *4:*514
Continuous membrane photoreactors
  degradation of dyes in *1:*794
  at high pollutant concentrations *1:*794
  4-nitrophenol degradation in *1:*793
Continuous monitoring technique *3:*400
Continuous permafrost *4:*306
Continuous systems, granular activated carbons in *1:*102
Continuous watershed response, modeling *3:*343
Continuous waves *3:*248
Continuous wire wrap screen *5:*573
Contour ridging *3:*551
Contour trenches *2:*550
Contracts, private sector participation *1:*389t
Control costs *2:*131
Controlled-tension lysimeter *2:*341
Control valves *1:*483
Control water, selecting *2:*382
Control-water exposures *2:*382
Convection, atmospheric *4:*360

Index terms **Links**

Convective storms, forecasting *4:*354
Conventional roof drainage *1:*55
Conventional sanitation, limitations of *1:*677
Conventional wastewater treatment *1:*840
  systems for *1:*678
Convention on Biological Diversity 3:122
Convergence *4:*716
Convergence pools *3:*68
Conveyance loss, defined *2:*542
Conveyance system structures *2:*568
Co-occurrence-based approaches *4:*601
Co-occurrence-based sediment quality guidelines *4:*601
Cooking, without power *2:*529
Cooling, adiabatic *4:*366
Cooling tower, biofouling in *2:*241
Cooling water contaminants, from electric generating plants *1:*556
Cooling water systems, water treatment for *1:*560
Cooperative agreements *2:*401
Cooperative model *1:*50
Coordinated freshwater fisheries management *3:*134
Coordination and study agencies, federal *4:*650t
Copenhagen permeable reactive barrier site *5:*521 *5:*522
Copernicus, Nicolas, water theory of *4:*712
Copper (Cu)
  biocidal effects of *1:*382
  concentrations *2:*422t
  dietary *1:*383
  in groundwater *2:*147
  maximum contaminant level of *2:*234t
  recovery of *1:*686
  role in metal tolerance *3:*612
  sorption capacity *2:*83
  toxicity to aquatic life *2:*147
Copper industry *1:*563
Copper removal
  by sorption *2:*362
  using polymers *1:*919
Copper tolerance, metallothioneins and *3:*612

Coprecipitation
  um *5:*20
  on *5:*20
Coprecipitation technique *4:*438
Coral decline, causes of *4:*114
Coral reefs *4:*113
  impacts on *4:*114
  industrial pollution and *4:*107
  protecting *4:*115
Coral reef stress, from black water *4:*133
Corals
  cold-water *4:*34
  deep-water *4:*32
  impact of oil on *4:*106
Corin Dam catchment fire *3:*466
Coriolis effect *4:*24 *4:*361
Corporate Average Fuel Economy standards (CAFE) *4:*479
Corporations, American *2:*499
Corporatization, of water *2:*569
Corrective Action, Phase III *5:*437
Corrective maintenance *2:*595
Correlative rights doctrine *4:*631
Corrosion *1:*142
  See also Corrosion control
  assessing levels of *1:*9t
  conditions leading to *1:*544
  consequences of *1:*7
  diagnosing *1:*152
  factors impacting *1:*7
  in industrial cooling water *1:*542
  inhibitors of *1:*10 *1:*153
  measuring *1:*8
  microbiologically induced *1:*600
  pipe breaks and *1:*400
  process of *1:*5
  soft water and *1:*324
  types of *1:*6 *1:*7t
Corrosion cell *1:*5
Corrosion control
  in drinking water systems *1:*5

Index terms **Links**

Corrosion control *(Continued)*
  methods of *1:*9
  microbial regrowth and *1:*345
  in pilot systems *1:*11
  in small drinking water systems *1:*459
  system design and *1:*153
  in water distribution systems *1:*152
"Corrosion fatigue cracking" *1:*545
Corrosion indexes (indices) *1:*9 *1:*10*t* *1:*152
Cosmic water *4:*189
Cosolvation *5:*527
Cost(s). *See also* Cost estimation; Economics; Life cycle costs; Pricing; Transaction costs; Water pricing
  of acid mine drainage treatment *5:*4
  of bottled water *1:*5
  of chemical oxidation technologies *5:*347
  of desalinated water *1:*174
  of diffused air aeration systems *1:*630
  of droughts *2:*579
  of fluoridating water systems *1:*256
  of *in situ* bioremediation *5:*42
  of irrigation *3:*592
  of magnetic water conditioning *1:*536
  of membrane technology *1:*309
  MIEX®. plant *1:*329
  of microirrigation *3:*619
  of multistage filtration *1:*242
  of the 1930s drought *2:*513
  point-of-use/point-of-entry system *1:*379t
  of regulatory compliance *1:*305
  of sand abstraction *3:*417
  of small-scale wastewater treatment *1:*842
  of solidification/stabilization *1:*839
  of sprinkler irrigation *3:*713
  of urban flooding *3:*161
  of water distribution system design *1:*211
  of water quality management *2:*369
Cost analysis, for treating disease *1:*427
Cost-benefit analysis *2:*129
Cost effectiveness *4:*675
Cost-effectiveness analysis *2:*131

Index terms **Links**

Cost-effective technology *2:*569
Cost estimation
  as a water conservation measure *1:*147
  water treatment plant *1:*222
Council of Great Lakes Governors (CGLG) *4:*618
Coupled flow domains, boundary conditions for *5:*660
Coupled flows
  mathematical models for *5:*659
  nodal replacement scheme for *5:*660
Coupled free flow *5:*655
Coupled Markov-chain (CMC) model *5:*309
Coupled model *4:*22
Coupled nonconservative pollutants *4:*50 *4:*51
Coupled reaction-transport models *5:*139
Coupon weight loss corrosion measurement *1:*8
Cousteau, Jacques *4:*792
Cox, Wim *4:*769
Cr(VI) *4:*515 *4:*587
  See also Hexavalent chromium
  instrument development for *in situ* monitoring of *4:*515
Cradle Mountain-Lake St. Clair National Park *4:*788
Craig line equation *4:*439
Crater lakes, acidic *3:*9
CREAMS model *2:*250
Creeping flow *5:*654
Critical depth *3:*196
Critical period techniques *3:*383
Critical storm period analysis
  basis for *4:*339
  case study in *4:*339
  management implications related to *4:*341
  statistical approach to *4:*338
Crop coefficient *2:*537
  factors affecting *3:*576f
  values of *3:*578t
Crop development stages, lengths of *3:*577t
Crop evaporation *3:*720
Crop evapotranspiration *3:*571
  estimates of *3:*572
  measurements of *3:*572

Index terms	**Links**		
Crop evapotranspiration *(Continued)*			
water table contribution to	*3:*570		
Cropland, pesticides applied to	*3:*607		
Crop management			
as a nitrogen best management practice	*3:*696		
techniques of	*2:*496		
Crop metabolism	*3:*721		
Crop Moisture Index (CMI)	*4:*212		
Crop phenology	*3:*722		
Crop photosynthesis	*3:*721		
Cropping patterns	*5:*164		
Crop quality, microirrigation and	*3:*618		
Crop removal, nitrogen losses from	*3:*696		
Crop residues, nitrogen in	*3:*695		
Crops. *See also* Crop water requirements			
adaptation to water stress	*3:*719		
microirrigation of	*3:*625		
salt tolerance of	*3:*684		
Crop tolerance, groundwater depth and	*3:*601		
Crop water requirements	*3:*557		
Crop water stress			
chronic	*3:*722		
detection using remote sensing	*3:*719		
Crop water stress index (CWSI)	*3:*719	*3:*720	
See also Idso-Jackson crop water stress index (CWSI)			
Crop yields			
evapotranspiration and	*4:*227		
waterlogging and	*3:*600		
Cross-flow ultrafiltration	*4:*29		
Cross connection control programs	*1:*157		
Cross connections	*1:*155	*1:*203	*1:*342
"Crossing distance"	*3:*81		
Cross-media contamination, from solidification/ stabilization	*1:*838		
Cross-media transfer potential, of pollutants	*1:*838		
Cross subsidization	*1:*216		
Crown width, levee	*3:*287		
Crustaceans. *See also* Amphipod entries			
in sediment impact testing	*2:*426		
stock enhancement of	*4:*125		
Cryosphere	*4:*183		

- Cryptosporidiosis *1:*180
- *Cryptosporidium* *1:*162 *1:*278 *1:*522 *1:*523
  - detection via polymerase chain reaction *1:*160
  - diagnosis of *1:*163
  - epidemiology and prevention *1:*164
  - in karstic aquifers *1:*366
  - life cycle and morphology *1:*163
  - molecular detection methods for *1:*161t
  - transport of *1:*366
  - treatment of *1:*164
- *Cryptosporidium parvum* *2:*343
  - antibody technology for identifying *1:*160
  - measuring oocyst inactivation following UV disinfection *1:*165
  - molecular-based detection of *1:*158
- Crystal Lake, biomanipulation trials in *2:*57t
- Crystallization *1:*810
  - of heavy metals *5:*378
  - scale formation and *1:*547
- Crystal modification, for scale control *1:*548
- CSO retention basins (RBs) *1:*783
- Cubic feet per second *2:*542
- Cultivated ponds *4:*697
- Cultural diversity, in relation to water *4:*791
- Cultural eutrophication *3:*114
  - effects of *3:*115
  - in ponds *3:*486
- Cultural media, prepackaged *1:*84
- Culturing methods, for detecting biofouling *1:*84
- Culvert design *3:*75
  - steps in *3:*77
- CulvertMaster software *3:*78
- Culvert materials *3:*75
- Culverts, flow velocities from *3:*77
- Cumulative absolute relative error (CARE) *2:*335
- Cumulative probability plots *2:*20
- Cured-in-place lining *1:*878
- Cured-in-place pipe (CIPP) *1:*802 *1:*878
- Current circulation, importance of *3:*320
- Currents *4:*24
  - effect on oceanic dilution *4:*104

Next Page

Index terms **Links**

Currents *(Continued)*
  methods of determining *3*:321
  numerical modeling of *3*:320
Curved surfaces, shocks on *3*:241
Curve number (CN) method *3*:343
Cutoffs, levee *3*:289
Cyanide
  analysis *2*:309
  maximum contaminant level of *2*:234t *2*:235t
Cyanobacteria *1*:277 *2*:21 *5*:463
  See also Blue- green algae entries
  blooms of *2*:390
  toxic *2*:388
Cyanobacterial toxins, structures of *3*:190f
Cyanosis *2*:219
Cyanotoxins *2*:387
  detection of *2*:389
  health problems related to *2*:388
  removal of *2*:390
  stability of *2*:390
Cycloalkanes, biodegradation of *1*:693
Cyclones/cyclonic storms *4*:194
  role in flooding *3*:144 *3*:146
Cyclospora *1*:278
Cyclosporiasis *1*:180
Cylindrical coordinates, groundwater flow equations in *5*:13
Cylindrospermopsins *2*:388 *2*:390
CYP1A induction *2*:109
Cytochrome P450 enzymes *2*:107
Cytochrome P450 monooxygenase
  as a bioindicator *2*:110
  as an indicator of pollutants in fish *2*:110
  as a PCB/dioxin-like compound indicator *2*:106
Cytotoxicity *2*:413

D

Dahlberg model *3*:455
Daily cover practice *1*:698
Dairy cow diets, improving *2*:573
  See also Livestock entries

Previous Page

Index terms	**Links**		
Currents *(Continued)*			
methods of determining	*3:*321		
numerical modeling of	*3:*320		
Curved surfaces, shocks on	*3:*241		
Curve number (CN) method	*3:*343		
Cutoffs, levee	*3:*289		
Cyanide			
analysis	*2:*309		
maximum contaminant level of	*2:*234t	*2:*235t	
Cyanobacteria	*1:*277	*2:*21	*5:*463
See also Blue- green algae entries			
blooms of	*2:*390		
toxic	*2:*388		
Cyanobacterial toxins, structures of	*3:*190f		
Cyanosis	*2:*219		
Cyanotoxins	*2:*387		
detection of	*2:*389		
health problems related to	*2:*388		
removal of	*2:*390		
stability of	*2:*390		
Cycloalkanes, biodegradation of	*1:*693		
Cyclones/cyclonic storms	*4:*194		
role in flooding	*3:*144	*3:*146	
Cyclospora	*1:*278		
Cyclosporiasis	*1:*180		
Cylindrical coordinates, groundwater flow equations in	*5:*13		
Cylindrospermopsins	*2:*388	*2:*390	
CYP1A induction	*2:*109		
Cytochrome P450 enzymes	*2:*107		
Cytochrome P450 monooxygenase			
as a bioindicator	*2:*110		
as an indicator of pollutants in fish	*2:*110		
as a PCB/dioxin-like compound indicator	*2:*106		
Cytotoxicity	*2:*413		

D

Dahlberg model	*3:*455		
Daily cover practice	*1:*698		
Dairy cow diets, improving	*2:*573		
See also Livestock entries			

Index terms **Links**

Dairy herd size reduction *2:*572
Damavand watershed *3:*157
Dammed pools *3:*67
Dam removal *3:*387
  engineering considerations related to *3:*388
Dams *2:*567
  adverse effect of *3:*200
  ancient *3:*382
  hydropower and *3:*199
  in the United States *2:*555
Dangerous Substances Directive (EU) *2:*267
Daphnia *3:*268
  heavy metals in *2:*423f
Daphnia magna, as a test organism *2:*420
Daphnids, metal concentrations in *2:*421
Darcy-Buckingham law *3:*249 *3:*250
Darcy equation *5:*656
  See also Darcy's Law
Darcy's Law *3:*249 *3:*453 *4:*486 *5:*63 *5:*507 *5:*515 *5:*525 *5:*554
Darcy-Weisbach formula *3:*195 *3:*197
Dastgheib and Rockstraw adsorption model *1:*108
Data collection/gathering
  for hydraulic network models *3:*316
  lake-related *3:*270
  in a systems approach *2:*685
Dataloggers *4:*71
Data management, following test results *2:*460
Data management model *2:*622
da Vinci, Leonardo, water experiments by *4:*710 *4:*717
DBCP (1,2-dibromo-3-chloropropane) *5:*358*t* *5:*359
DDT *2:*600 *5:*338
  See also Dichlorodiphenyltrichloroethane (DDT)
DDT residues, persistence of *3:*515
"Dead chlorine" *4:*418
Dead storage *1:*409
"Dead zones" *4:*25
Debris, marine *4:*38
Debris dam pools *3:*68
Decaying tropical cyclones, role in flooding *3:*144

Decay rate estimation *2:*334
Decentralized Sanitation And Reuse (DESAR) *1:*520
Decentralized wastewater treatment *1:*678
Dechlorination *1:*169 *2:*399
  domestic sewage *1:*834
  reductive *2:*299
Decibel scale *4:*570
Deciles *4:*213
Decision making
  flood-management *2:*679
  participatory multicriteria *2:*681
  precautionary *2:*596
  risk-informed *2:*599
  systems approach to *2:*685
Decision matrix *2:*679
Decision Support System for Evaluating River Basin Strategies (DESERT) *2:*328
Decision support systems (DSSs) *2:*328 *2:*535 *2:*589
  hydroelectric *2:*621
  river basin *2:*619
  for pond aquaculture *3:*375
  spreadsheets in *2:*620
  value of *2:*623
  for water resources management *2:*622
Decision variables, in a mathematical model *2:*686
Deck officers *4:*763
Declining-rate filters *1:*235
Decomposers *4:*152
Decompression sickness *4:*451
Deep-sea shipping, employment in *4:*765
Deep soil-water movement *5:*471
  forces that control *5:*471
  material properties that affect *5:*471
  water vapor movement in *5:*472
Deep-water corals *4:*32
  distribution of *4:*34
  threats to *4:*36
  types of *4:*33
Deepwater habitats, classification of *3:*496
Deepwater swamps *3:*172

Index terms	**Links**
Deep-well turbine pumps	*3:*664
efficiency of	*3:*665
tests of	*3:*666
Definite Enterprise Portal Resource	*2:*669
Defluoridation	*4:*434
Nalgonda process for	*4:*437
sorption design and	*4:*437
testing or metering	*4:*438
treatment principles for	*4:*436
Defluoridation technology	*4:*435
Degassing fluxes	*3:*206
Degradation. *See also* Pollutant photodegradation	
of chemicals	*5:*262
of chloro-organics and hydrocarbons	*1:*688
of dyes	*1:*794
4-nitrophenol	*1:*793
Degradative enzymes	*2:*43
Degree-day factor	*4:*337
Degree day (DD) method	*4:*197
De Groen’s interception method	*3:*237
Dehydration, diseases of	*4:*723
Dehydration toilets	*1:*679
Deicers, chemical	*2:*319
See also Ice control	
Delaware River Basin Commission drought indicators	*2:*581
Delaware River Basin Compact	*2:*585
Deleterious organic chemicals, in urban stormwater runoff	*3:*435
Delivery system	*5:*291
Demand (average water requirement)	*1:*506
Demand curve	*4:*606
Demand management	*1:*204
in the Arab World	*2:*472
in Canada	*2:*662
Demand reduction systems	*2:*515
Demersal fish	*3:*126
Demineralization	*1:*297
in small drinking water systems	*1:*459
Denaturing gradient gel elecrophoresis (DGGE), community profiling using	*1:*643
Denil fish passage facility	*3:*530

Index terms **Links**

Denitrification *1:*816 *1:*833 *1:*893 *3:*696 *4:*54 *4:*391
  in the activated sludge process *1:*667
  biological *5:*325
  controlled *1:*752t
  in soils *3:*539
Denitrification tank *1:*669
Dense nonaqueous phase liquids (DNAPLs) *3:*689 *5:*340 *5:*439 *5:*603
  See also DNAPL entries
  modeling of *5:*668
  physical properties of *5:*91
Density. *See also* Filter density (ρ_{GAC}); High density sludge (HDS) process; Silt density index (SDI)
  of carbon *1:*96
  of sludge *1:*862
Density currents *3:*412
Density gradients, as forcing factors *3:*322
Dental fluorosis *4:*435
Denver, Colorado, overdraft in *5:*341
Deoxygenation *3:*445
Department of Heath, Education, and Welfare (HEW) *4:*652
de Passa, Jaubert, irrigation history by *4:*736
Depletion factor *3:*23
Depth filters *2:*99
Depth filtration, role of colloids and dissolved organics in *2:*99
Depth-integrating samplers *3:*398
Depth profiling *5:*411
Dermatitis *1:*180
Dermatotoxins *2:*389
Desalinated water
  cost of *1:*174
  quality of *1:*173
Desalination *1:*170 *2:*374 *2:*493
  concentrate management in *1:*174
  defined *2:*542
  by electrodialysis *1:*171
  membrane *1:*171
  reverse osmosis *1:*171
  in the United States *2:*650
Desalination market, future of *1:*309

Index terms **Links**

Desalination system, wave-powered *4:*46

Desert cienegas *3:*58

Desert ecology, dew in *4:*200

Desertification *4:*184 *4:*199

Desert regions, recharge in *5:*72

Design
- municipal solid waste landfill *2:*168
- well *5:*87

Design discharge
- for a series of planes *3:*243
- for a V-shaped basin *3:*245

Design flows, in streams/rivers *3:*81

Desorption *2:*70

Desorption kinetic tests *4:*568

Desorptive agents *2:*70

Destruction/destructive technologies
- *in situ* 5:441
- for odor abatement *1:*762

Detachment mechanisms *1:*231 *1:*233 *1:*244 *1:*246

Detailed water quality models *2:*249

Detection
- molecular-based *1:*158
- polymerase chain reaction *1:*160
- of waterborne radon *1:*51

Detection methods, categories of *1:*90

Detection systems, flood-related *3:*148

Detectors, gas chromatography *2:*307

Detention facilities *3:*430

Detention ponds *3:*430

Detention storage *3:*474

Detention systems, for storm water treatment *1:*868

Detergents *1:*669
- biodegradation of *1:*672
- occurrence in wastewater and sewage sludge *1:*671
- regulatory standards for *1:*672
- structure and use of *1:*670
- toxicity of *1:*672

Deterministic approach, to studying sediment transport *3:*401

Deterministic models *3:*337 *5:*303
  in eutrophication management *3:*111
  of sewer water quality *3:*332
Deterministic techniques *2:*332
‘‘Detrending’’ technique *3:*219 *3:*220
Deuterium *4:*438
  See also Heavy water
  analyzing water samples for *4:*439
  discovery of *4:*464
  in the global water cycle *4:*439
  as a tracer *4:*440
Developed countries, household water consumption in *1:*506f
Developing countries/nations
  applying existing technologies to *1:*718
  centralized and decentralized treatment in *1:*720
  drinking water in *1:*290 *2:*630
  effluent quality and standards in *1:*720
  meeting water needs in *2:*643
  microirrigation in *3:*627
  organic agriculture in *3:*645
  remedial criteria for *5:*424
  safe drinking water in *1:*22t
  sanitation services in *1:*23*t* *1:*661 *2:*630
  wastewater management for *1:*718
  wastewater treatment technology in *1:*721
  water resources planning and management objectives for *2:*683
  water supply in *2:*616
  wetland values in *3:*492
Development
  managing urban runoff from *2:*187
  patterns of *2:*191
  surface water chemistry and *3:*442
  water quality and *3:*501
Dew *4:*200
  duration of *4:*204
  measurements of *4:*202
Dewatering, of sludge *1:*854 *1:*863 *1:*865
Dew deserts *4:*201
Dew point *4:*207 *4:*272
Dew point lapse rate (DPLR) *4:*369

Index terms **Links**

Dew precipitation *4:*205
Diabetes, dehydration and *4:*725
Diadromous fisheries management *3:*127
Diarrheal disease *1:*24 *2:*371
  from drinking water sources *1:*188
  Milwaukee outbreak of *1:*184
Diatomaceous earth filters *1:*175f
  elements of *1:*176
Diatomaceous earth filtration *1:*175 *1:*228 *1:*232 *1:*247 *1:*458
  for drinking water *1:*174
  feed water quality and performance capabilities in *1:*175
  monitoring and operating requirements for *1:*176
  for small systems *1:*177
  types of *1:*176
Dibromoacetic acid, health effects of *1:*266
Dibromoacetonitrile, health effects of *1:*268
Dibromochloroacetic acid, health effects of *1:*267
Dibromochloromethane, health effects of *1:*265
Dibromomethane, health effects of *1:*264
Dichloroacetic acid, health effects of *1:*266
Dichloroacetonitrile, health effects of *1:*268
Dichlorodiphenyltrichloroethane (DDT) *3:*352 *3:*645
  See also DDT entries
  in Argentina *3:*649
  beneficial results of *3:*515
  biological magnification of *3:*523
  breakdown and disappearance of *3:*516
  effects on bird abundance and reproduction *3:*513
  misidentifications of *3:*522
  role in saving forests *3:*522
  versus PCBs *3:*521
Dichloroethene (DCE) *5:*581
Dichloromethane (DCM) *2:*317
Dieldrin *3:*352
Differential scanning calorimetry (DSC) *4:*401
Differential settling, flocculation by *1:*252
Diffused aeration *1:*353 *1:*460
Diffused air aeration *1:*625
Diffused air aeration systems. *See* Fine bubble diffused air aeration systems

Index terms	Links			
Diffuse double layer (DDL) model	*2:*363	*5:*567		
Diffuse flow springs	*5:*239			
Diffuse flow systems	*5:*237			
Diffuse recharge	*5:*236			
Diffuser membrane materials	*1:*628t			
Diffuser systems, mixed	*1:*628			
Diffusion	*5:*278			
leaching	*5:*261			
organic vapor transport via	*5:*544			
transport by	*5:*516			
Diffusion-based groundwater sampling				
case study of	*5:*458			
equilibration time for	*5:*457			
membrane materials and construction examples for	*5:*456			
methods of	*5:*456			
Diffusion kinetics, modeling	*4:*567			
Diffusion porometers	*3:*717			
Diffusion transport	*3:*310			
Diffusion wave theory	*3:*249			
validity of	*3:*250			
Digestion, of biosolids	*1:*647			
Digital data, extraction of drainage networks from	*3:*93			
Digital elevation models (DEMs)	*2:*531	*3:*28	*3:*29	*3:*64
	*3:*93	*3:*157	*3:*162	
Digital filters, base flow separation using	*3:*27			
Digital models	*5:*241			
Digital terrain modelling (DTM)	*2:*534			
Dike and bleeder flow systems	*3:*165			
Diked fields	*4:*696			
Dilatant fluids	*5:*557			
Dilution				
in algal control	*2:*4			
examples of	*3:*84			
importance of	*3:*78			
as a means of wastewater disposal	*3:*79			
Dilution equivalents	*4:*104			
Dilution factor	*4:*104			
Dilution function	*4:*104			
Dilution-mixing zones	*3:*78			

Index terms	Links			
Dilution water	*2:*382			
characteristics of	*2:*383			
Dimensionless cumulative time-area diagram	*3:*64			
Dimensionless unit hydrograph (DUH)	*3:*440			
Dimethyl sulphide (DMS)	*4:*88			
Dimictic lakes	*3:*266			
Dinitrophenol, adsorption of	*1:*118			
1,4-Dioxane	*2:*317			
Dioxin exposure	*3:*120			
Dioxin-like compounds, as environmental toxicants	*2:*106			
Dipole-dipole array	*5:*445			
Direct current (DC) technology	*5:*118			
Direct electron transfer	*2:*41			
Direct filtration	*1:*228	*1:*231	*1:*235	*1:*246
	*1:*458			
Direct liquid monitoring techniques	*5:*541			
Direct potable reuse	*2:*611			
Direct push water well drilling	*5:*106			
Direct roughing filtration	*1:*238			
Direct runoff hydrograph (DRH)	*3:*440	*4:*355	*4:*356	
Direct runoff plots	*3:*26			
Direct search method	*2:*332			
Direct soil moisture measurement	*3:*692			
Direct soil monitoring techniques	*5:*542			
Direct toxicity assessment (DTA)	*2:*46			
Direct vadose zone monitoring methods	*5:*541			
Direct wastewater reuse	*1:*826			
Direct water valuation approaches	*4:*607			
Direct water valuation studies	*2:*655			
Dirichlet boundary condition	*5:*14	*5:*15	*5:*561	
Disaggregation	*3:*306			
stochastic	*3:*427			
Disaster Assistance Act of 1988	*2:*578			
Disaster management systems, development and	*3:*33			
Disaster supplies	*2:*529			
Discharge(s)	*5:*601			
defined	*2:*542			
effects of seasonal changes in	*3:*6			
to lakes	*3:*281			
mixing behavior of	*3:*81			

Index terms	**Links**	
Discharge(s) *(Continued)*		
monitoring	*4:*317	
in regional flow systems	*5:*418	
as a watershed function	*3:*476	
Disconnected pools	*3:*69	
Discontinuities		
first or second order	*3:*248	
order of	*3:*239	
Discontinuous gullies	*3:*184	
Discontinuous heterogeneity	*5:*509	
Discontinuous permafrost	*4:*306	
Discrete particle settling	*1:*260	
Discrete recharge	*5:*236	
Discrete sampling	*5:*455	
Discriminant function analysis (DFA)	*2:*282	
Disease(s). *See also* Disease outbreaks; Infectious diseases		
arsenic-caused	*1:*1	
endemic	*1:*187	
flooding-related	*3:*161	
Saratoga Springs waters treatment of	*4:*798	
soft water and	*4:*554	
wastewater-associated transmission of	*3:*669	
water-related	*1:*23	*2:*630
Disease outbreaks		
cholera	*1:*283	*1:*290
Cryptosporidium	2:343	
detection of	*1:*186	
from distribution system deficiencies	*1:*341	
Escherichia coli	1:430t	2:344
Milwaukee	*1:*184	
short-duration	*1:*189	
Disease pathogens, waterborne	*2:*343	
Disease pyramid	*1:*185f	
Disease reporting	*1:*185	
Disease surveillance		
definitions	*1:*185t	
improving	*1:*190	
Disease surveillance systems, limitations of	*1:*185	
Disease transmission, low levels of	*1:*187	
Disease treatment, cost analysis for	*1:*427	

Index terms	Links	
Dishwashing detergent, antibacteriological properties of	*2:*114	
Dishwater		
microbiological content of	*2:*114	
temperature of	*2:*113	*2:*115
Dishwater quality	*2:*112	
environmental aspects of	*2:*114	
Disinfectant application, changing the order of	*2:*120	
Disinfectant residual concentrations	*2:*226t	
Disinfectants	*1:*192	
See also Disinfection		
chemistry of	*1:*194	
comparison of	*1:*197	
efficacy of	*1:*193	
against food-poisoning bacteria	*1:*604	
history of	*1:*193	
microbial regrowth and	*1:*344	
purpose of	*1:*192	
testing of	*1:*602	
toxicology of	*1:*195	
types of	*1:*193	
Disinfection	*1:*196	*1:*380
See also Chlorination entries; Disinfectants; Ultraviolet disinfection		
achievement of	*1:*197	
byproduct control for	*1:*199	
at CSO facilities	*1:*786	
dilution equivalent to	*4:*104	
in distribution systems	*2:*244	
of domestic sewage	*1:*834	
after pipeline repair	*1:*890	
POP behavior during	*1:*768	
regulations related to	*2:*92	
of rural drinking water	*1:*382	
of sludge	*1:*857	
for small drinking water systems	*1:*457	
surface water treatment rule compliance technologies for	*1:*461t	
using ozone	*1:*354	
water safety and	*1:*196	
water treatment plant	*1:*222	*1:*224
well	*1:*151	

Disinfection byproduct precursors
  removal by coagulation *2*:116
  removal from natural waters *2*:115
Disinfection byproducts (DBPs) *1*:193 *1*:287 *1*:354 *1*:357 *2*:398
  aspects of *2*:77f
  in drinking water *1*:213
  formation and occurrence of *2*:91
  health effects of *1*:264
  importance of *1*:214
  minimization of *2*:119
  regulations related to *2*:92
  routes of exposure for *2*:92
  toxicity estimates for *1*:270 *1*:271t
Disinfection Byproducts Rule (DBPR) *2*:92
Disinfection practices, alternative *2*:118
Disinfection wastewater treatment processes *1*:817
Disk infiltrometers *5*:215 *5*:539
Dispersants, for scale control *1*:548
Dispersed-air flotation *1*:685
Dispersion *5*:46
  leaching via *5*:26
  in transport modeling *5*:331
Dispersion phenomena *3*:78
Dispersivity *5*:274
Dissociation kinetics, of Am-humic colloids *2*:105
Dissolution
  in subsurface systems *5*:416
  water-rock interaction and *5*:568
Dissolution rate (*R*), for rock *5*:448
Dissolved-air flotation (DAF) *1*:487 *1*:685 *1*:786 *3*:407
Dissolved compounds, removal of *1*:684
Dissolved gases *4*:450
  in biological systems *4*:451
  in groundwater *4*:450
  in the oceans *4*:451
  role in colloid science and biochemistry *4*:451
  role in surface tension of electrolytes *4*:452
Dissolved humic substances (DHS) *1*:872
Dissolved inorganic carbon (DIC), effect on corrosion *1*:8

Index terms	**Links**			
Dissolved ionics, removal of	*1:*54			
Dissolved iron	*4:*498			
Dissolved matter, transport phenomena of	*2:*105			
See also Dissolved organic matter (DOM)				
Dissolved metals, role in fish growth and production	*3:*131			
Dissolved organic carbon (DOC)	*1:*360	*2:*224	*3:*1	*4:*399
See also DOC entries				
acidification and	*3:*5			
analysis with	*1:*223			
chemical structure of	*4:*399			
Dissolved organic compounds. *See also* Dissolved organic carbon (DOC); Dissolved organic matter (DOM); Dissolved organics				
metals and	*1:*727			
removal of	*1:*54			
Dissolved organic matter (DOM)	*2:*212	*2:*213	*4:*533	
See also Dissolved matter; Dissolved organics				
Dissolved organics	*2:*100			
biophysicochemical interactions with filters	*2:*100			
Dissolved oxygen (DO)	*2:*142	*5:*431		
depletion of	*4:*64			
effect on corrosion	*1:*8	*1:*153		
in ponds	*3:*485			
production of	*2:*275			
role in fish growth and production	*3:*130			
in urban stormwater runoff	*3:*434			
variation in	*3:*84			
as a water quality parameter	*2:*270			
Dissolved oxygen concentration				
fish health and	*3:*130			
in tidal streams	*4:*130			
Dissolved phosphorus (DP)	*3:*704			
Dissolved solids	*4:*587			
Distillation	*1:*380	*1:*813	*2:*527	
in heavy water production	*4:*464			
multiple effect	*1:*170			
multistage flash	*1:*170			
Distilled water	*4:*441			
Distributed solar water heaters	*1:*64			
Distribution, regional	*2:*615			
Distribution line breaks, repairing	*1:*400			

Index terms	**Links**	
Distribution systems	*2:*616	
control of nitrification in	*2:*226	
disinfection in	*2:*244	
microbial nutrients in	*2:*244	
microbiological quality control in	*2:*243	
research on	*2:*511	
Distribution system water quality	*1:*204	
contamination prevention and control in	*1:*205	
corrosion control in	*1:*206	
deterioration factors in	*1:*204	
maintenance alternatives for	*1:*206	
modifications to system operation	*1:*205	
monitoring and modeling in	*1:*207	
operational factors in	*1:*205	
structural factors in	*1:*204	
treatment practice in	*1:*206	
water quality factors in	*1:*205	
District metering areas (DMA)	*1:*202	
Ditches		
in industrial areas	*3:*87	
research on	*3:*88	
Diversion pumped storage development	*3:*201	
Diversion structures	*2:*568	*3:*151
Diversity measures	*3:*37	
DNA		
shotgun sequencing of	*2:*63	
water and	*4:*458	*4:*512
DNA arrays	*2:*60	
DNA chip technology	*1:*161	
See also DNA microarrays		
DNA damage, bacterial biosensors for detecting	*2:*455	
DNA microarrays, in community analysis	*1:*644	
DNAPL migration, effect of	*5:*670	
DNAPL migration modeling, adaptive modified sequential method of	*5:*669	
DNA probes, in *Salmonella*	*2:*337	
DNA stable isotope probing	*1:*645	
DOC exchange mechanism	*1:*326f	
See also Dissolved organic carbon (DOC)		
DOC removal	*1:*225	
ion exchange for	*1:*325	

Index terms	**Links**	
Dolomite	*4:*413	
‘‘Dolphin safe’’ tuna fishing	*2:*636	
Domain membrane process	*1:*139	
Dombes basin	*4:*697	
Domestic sector, water conservation in	*2:*496	
Domestic sewage	*1:*828	*1:*830
characteristics of	*1:*830	
treatment technology for	*1:*832	
as a source of cultural eutrophication	*3:*114	
as water pollution	*4:*97	
Domestic solar water heaters	*1:*63	
distributed solar water heater	*1:*64	
integrated collector/storage solar water heater	*1:*63	
Domestic wastewater		
constructed wetlands for	*1:*896	
disposal of	*2:*182	
Domestic wastewater discharges, control of algal nutrients in	*3:*110	
Domestic water		
quality of	*2:*371	
radon in	*4:*542	
risk assessment of radon in	*4:*543	
Domestic water supply		
public-private partnerships and	*1:*42	
eutrophication in	*3:*107	
Domestic water usage	*2:*371	*2:*542
average daily	*1:*506f	
in India	*2:*565	
in the United States	*2:*652	
Dominant processes concept (DPC)	*3:*229	
Dong Nai-Saigon Rivers, wastewater discharge into	*2:*552	
Donnan membrane process (DMP)	*1:*139	*1:*140f
Doppler radar	*4:*354	
precipitation estimates via	*4:*307	
Dose-response data	*2:*381	
Dose response of mussels to chlorine	*2:*401	
Double check valve	*1:*156	
Double-composite landfill liner	*2:*165	
Double layer compression	*4:*425	
“Double low water”	*4:*24	
Downflow roughing filters in series (DRFS)	*1:*241	

Index terms **Links**

Down-hole televising log *5:*153
Downstream fining *3:*420
Downward vapor transport *5:*551
Draft from groundwater (*Tp*) *5:*168
Drainage *1:*498
  inadequate *3:*603
  lysimeters *5:*488
  road runoff *3:*498
  roof runoff *3:*498
  status and scope of *5:*99
  subsurface *3:*451 *5:*94
  vertical *5:*98
Drainage and leachate collection/recirculation/treatment, for landfills *1:*698
Drainage basins *2:*543
  bioremediation in *2:*359
  climatic characteristics of *3:*31
  rock and soil characteristics of *3:*29
  sediment movement in *3:*185
  topographic characteristics of *3:*29
  vegetation and land use characteristics of *3:*30
Drainage channel, slope of *3:*31
Drainage composition, Horton’s laws of *3:*32
Drainage density *3:*94
  of a channel network *3:*32
Drainage design, role of evapotranspiration in *5:*96
Drainage ditches *3:*87
Drainage divides *2:*657
Drainage layers, levee *3:*290
Drainage measurements, using nonweighing lysimeters *5:*488
Drainage models, urban *1:*125
Drainage networks *3:*93 *3:*438
  analysis of composition of *3:*32
  extraction from digital data *3:*93
  statistical properties of *3:*93
Drainage works, controversy over *3:*88
Drains, field measurements in *3:*338
DRASTIC assessment methodology *1:*73
DRASTIC ground water vulnerability map *5:*58 *5:*563
DRASTIC index method *5:*57 *5:*358 *5:*595

Drawdown *5:*101 *5:*103
  defined *2:*543
Drawdown information, importance of *5:*104
Drawdown zone *3:*206
Dredged material
  acute sediment toxicity tests for *2:*351
  characterization of *2:*351 *2:*354t
  disposal at sea *4:*148
  uncertainties associated with testing *2:*353t
Dredged sediment
  regulating disposal of *2:*125
  water quality aspects of managing *2:*122
Dredged sites, use of caged animals at *2:*351
Dredging, hydraulic *2:*122
Dredging/barge disposal, mechanical *2:*122
Dredging/disposal operations, impact of *2:*122
Dresden, Germany, biomanipulation trials in *2:*56t
Dried food, nitrite content of *1:*33t
Drift currents *4:*24
Drifts, as groundwater tracers *5:*504
Drilled wells *5:*572
  versus qanats *5:*486
Drilling, gas hydrate sampling through *4:*61
Drilling fluids, specific gravity of *5:*474
Drilling sites, Lake Baikal *3:*21
Drilling techniques *5:*438
  water-jetting *5:*234
Drinking water. *See also* Public health protection
  aquatic macrophytes as biomonitors of *2:*66
  detecting arsenic in *1:*2
  in developing countries *1:*290 *2:*630
  diatomaceous earth filtration for *1:*174
  disinfecting *1:*196
  disinfection byproducts in *1:*213
  disinfection byproduct standards for *2:*92t
  effects of surface water pollution on *3:*445 *3:*446 *3:*447t
  guidelines *5:*324t
  microbial contaminants and biotoxins in *1:*277
  microbes in *2:*243
  microbiological quality of *1:*225

Index terms	Links			
Drinking water *(Continued)*				
nitrate limits for	*1:*33t			
nitrates in	*2:*219	*3:*694	*5:*198	*5:*323
nutrient limitations of bacterial growth in	*2:*245			
organic compounds in	*2:*372			
pollutants in	*5:*262			
radioactive contamination in	*1:*803			
radon in	*4:*542			
reducing radon in	*1:*51			
removal of arsenic from	*1:*2			
Salmonella detection and monitoring in	*2:*337			
security of	*1:*434			
toxic chemicals in	*2:*371			
types of organic carbon in	*2:*224			
Drinking Water Contaminant Candidate List	*1:*532	*4:*678		
Drinking water distribution systems				
etiological agents of outbreaks related to	*1:*342			
microbial occurrence in	*1:*342			
microbial regrowth in	*1:*344			
microbiological concerns of	*1:*341			
Drinking water equivalent level (DWEL)	*4:*397			
Drinking water filtration	*1:*227			
cartridge	*1:*229			
diatomaceous earth	*1:*228			
direct	*1:*228			
membrane	*1:*229			
multistage	*1:*237			
packaged	*1:*229			
regulatory requirements for	*1:*233			
slow sand	*1:*228			
Drinking water protection, EPA Pretreatment Program and	*1:*799			
Drinking water purification, Br^-/DOC changes in	*2:*75			
Drinking water quality	*1:*131			
Australian framework for managing	*1:*428t			
causes of failure of	*1:*221			
Drinking water quality standards (DWQS). *See also* Drinking water standards				
compliance and supervision related to	*1:*480			
future of	*1:*481			
history of	*1:*479			
setting	*1:*480			
in the United States	*1:*476			

Index terms **Links**

Drinking water regulations, history of *2:*194
Drinking water regulations, risk assessments in crafting *1:*422
Drinking water standards *1:*2 *1:*293 *2:*345
  See also Drinking water quality standards (DWQS)
  analytical methods for *1:*297
  chemical *1:*529
  national primary *2:*373t
Drinking Water State Revolving Fund (DWSRF) *1:*445
Drinking water storage facilities *1:*408
  construction materials for *1:*410
  daily volume use *1:*409
  requirements for *1:*408
  storage reservoir location *1:*410
  storage tank shape and volume *1:*410
Drinking water systems
  corrosion control in *1:*5
  infrastructure for *1:*883
Drinking Water Systems Center *1:*91
Drinking water treatment
  household *1:*67
  MIEX® resin in *1:*327
  ozone-activated carbon *1:*357
Drip irrigation *2:*543 *3:*582 *3:*593
  in India *3:*625
Dripper tube *3:*623
Driven wells *5:*572
Drop-size distribution rain simulators *4:*331
Drought(s) *4:*208
  climatology and *4:*187
  collective action against *2:*579
  costs of *2:*513 *2:*579
  effects of *4:*209
  in India *2:*567
  internal *4:*725
  intracellular *4:*723
  legacy of *2:*513
  in the Nile Basin *2:*590
  in the 1930s *2:*511
  obstacles to planning for *2:*578
  planning tools for *2:*580
  political considerations related to *2:*580

Drought(s) *(Continued)*
  public notification of *2:*583
  recovering from *2:*512
  traditional government approach to *2:*577
  water resources management for *2:*576
  water use curtailment during *2:*583
  water-user compliance monitoring during *2:*583
  West African Sahel *4:*184
Drought-contingency-plan phasing criteria *2:*582t
Drought escape *3:*719
Drought indicators, identifying *2:*580
Drought indices *4:*209
  Crop Moisture Index *4:*212
  deciles *4:*213
  Palmer Drought Severity Index *4:*211
  percent of normal calculation *4:*209
  Reclamation Drought Index *4:*213
  Standardized Precipitation Index *4:*210
  Surface Water Supply Index *4:*212
Drought management *2:*515 *2:*585
  base flow recession in *3:*26
  planning in *2:*514
Drought planning *2:*577
  specificity of *2:*578
Drought-planning authority, designating *2:*583
Drought programs, federal *2:*578
Drought-related planning and management activities, state *2:*583
Drought relief programs *2:*512
Drought-severity index *2:*581
Drought surcharge *2:*583
Drought tolerance *3:*719
Dry adiabatic lapse rate (DALR) *4:*368 *4:*369
Dry arm discontinuous valley-floor gully *3:*184f
Dry barriers *5:*2
Dry-bulb temperature *4:*224
Dry electrolytic conductivity detectors (DELCD) *2:*308
Dry environments, disturbance of *2:*516
  See also Arid countries/lands
Drying, of sludge *1:*860
Drying beds, sludge *1:*854

Index terms **Links**

Dry-land agriculture, water conservation in *2:*496
Dry monsoons *4:*304
Dry toilets *1:*679
Dry tomb landfilling *2:*163
Dual-flush device *2:*665
Dual-nutrient removal, biological *1*:816 *1*:834
Dubinin-Astakhov (DA) adsorption isotherm equation *4:*406
Dublin principles *2:*574
Dublin Statement *4:*611
Dug wells *3:*594 *5:*571
Dumped wastes, interactions with seawater *4:*148
Duplex PCR *2:*338
Dupuit-Thiem method *5:*492
Durges, Susan *4:*767
Dye introduction trenches (DITs) *5:*109
Dyes
  degradation of *1:*794
  as hydrological tracers *3:*95
  sorption to subsurface media *3:*98
  testing of *1:*885
Dye tracers
  characteristics of *3:*95
  limitations in using *3:*98
  measurements of *3:*83f
  QSAR approach versus screening for selecting *3:*98
  selecting for specific uses *3:*98
  structure of *3:*97f
  for surface water, groundwater, and vadose zone *3:*96
Dye tracing *5:*240
  groundwater *5:*107
  at waste disposal sites *5:*110
Dynamic data *2:*622
Dynamic modeling *1:*132 *2:*272
Dynamic pressure *5:*652
Dynamic roughing filters (DyRF) *1:*237 *1:*240
Dynamic roughing filtration *1:*239
Dynamics, reservoir *3:*259
Dynamics and instabilities *4:*449
Dynamic-state modeling *1:*133
Dynamic wastewater models *1:*731 *1:*733

E

Early civilizations, water in *4*:726 *4*:746
Early-closure flush device *2*:665
‘‘Early warning’’ biomarkers *2*:429
Early warning systems *3*:33
Earth
  heat balance of *4*:92
  origin of water on *4*:283
  water content of *4*:191
  water resource renewal periods of *4*:278t
Earth energy budget, balancing *4*:169
Earth Observing System (EOS) *4*:351
  See also Aqua earth observing system (EOS)
Earth Observing System Data and Information System (EOSDIS) *4*:216
Earthquake dewatering *5*:478
Earthquakes *5*:111
  hydrogeologic responses to *5*:113
  severity *5*:478t
  stream, spring, seep, and lake responses to *5*:112
  water well responses to *5*:111
Earth Summits *2*:624 *2*:628
Earth system science *4*:214
Earth tides *5*:498
East Asian Seas, dumping in *4*:145
Easterly Waves *4*:305
Eastern states, protection of in-stream values *4*:662
East Patchogue, New York, plume diving in *5*:80
EC$_a$-directed sampling designs *5*:467
  See also Apparent soil electrical conductivity (EC$_a$)
EC$_a$-directed soil sampling survey *5*:467
EC$_a$ field survey, steps in *5*:469t
EC$_a$ measurements, limitations and precautions concerning *5*:467
Echolocation *4*:571
Echovirus infections *1*:179
E climates *4*:258
Ecohydrology *2*:368 *3*:292 *4*:267
Ecological benefit. *See* Net ecological benefit concept
Ecological effects
  assessment for water-limited environments *2*:516
  of pharmaceutical products *1*:373 *1*:376

Index terms	**Links**		
Ecological health	*2:*182		
human activities as stressors on	*2:*625		
Ecological impacts, of water transfers	*4:*686		
Ecological processes, stock assessments and	*4:*57		
Ecological research, on agricultural drainage ditches	*3:*88t		
Ecological risk assessment (ERA)	*3:*54		
Ecological Risk Assessment Guidance for Superfund (ERAGS)	*2:*517		
Ecological sanitation (EcoSan)	*1:*675		
See also EcoSan systems			
Ecological surveys	*2:*30		
Ecological wastewater management	*1:*675		
treatment systems in	*1:*677		
Ecological watershed functions	*3:*477		
Economic activity, in India	*2:*564		
Economic growth, American	*2:*499		
Economic planning (EP) model	*2:*520	*2:*521	
Economic productivity, water resources development and	*3:*33		
Economics. *See also* Cost(s); Financial performance; Global economy; Investment; Pricing; Stock price performance; Water companies; Water economic analysis; Water prices; Water utilities			
concepts in	*4:*606		
during the Great Depression	*2:*511		
of industrial water demands	*1:*549		
of residential water demands	*1:*12		
of water quality	*2:*127		
of water resource allocation	*1:*215		
of water resource valuation	*2:*653		
Economic value of water	*4:*605		
Economic theory, of industrial input demands	*1:*550		
Economies of scale, public water supply and	*1:*501		
Ecoregions, environmental management of	*2:*667		
EcoSan systems	*1:*675	*1:*676	*1:*680
advantages of	*1:*680		
Ecosystem processes, using to treat mine wastewater	*1:*897		
Ecosystems. *See also* Sensitive ecosystems			
balanced	*2:*667		
coral reef	*4:*114		
effect of water hyacinth on	*3:*481		
effluent-driven	*2:*476		
Hawaiian	*4:*802		

Index terms **Links**

Ecosystems *(Continued)*
  lake *3:*267 *3:*272
  managing fishery-related *3:*128
  sustainable management of *2:*636
  variability in *3:*124
  water and soil conservation *4:*772
Ecotoxicogenomic research, developing tools for *2:*62
Ecotoxicological assessment, of pharmaceuticals *1:*376
Ecotoxicological effects, of lead *2:*433
Ecotoxicological sediment assessment *2:*383
Ecotoxicological testing *4:*42
Ecotoxicology, purpose of *2:*379
Eddy flux method *3:*736
Eddy pools *3:*67
Edgerton, Harold *4:*768
Edible aquatic organisms, bioaccumulation of hazardous chemicals in *3:*435
EDTA method *4:*453
Education *1:*148
  See also Public education; Water education strategies
  CERES-related *4:*170
  as a water conservation measure
Educational programs *2:*497
Effective contact time *1:*103
Effective porosity *5:*184
  analysis *5:*186
Effective rainfall (ER) *4:*355 *4:*356 *4:*357
Effective rainfall hyetograph (ERH) *4:*355 *4:*356
Effective sediment erosion intensity per unit depth (ESEI) *4:*359
Effective viscosity *5:*659
"Efficient water use" *2:*491
Effluent(s). *See also* Whole effluent toxicity controls
  defined *2:*543
  diversion of *2:*3
  industrial *2:*496
  mining *1:*609
  pond *3:*541
  runoff *2:*516
  technetium-laden *4:*580
  trading *2:*134
  treating *1:*682f

Index terms **Links**

Effluent discharges *2:*373
  to dry riverbeds *2:*476
  national primary drinking water standards *2:*373t
Effluent-driven ecosystems
  attainable uses in *2:*476
  flow regime modification in *2:*477
  net ecological benefit in *2:*477
  water regulation in *2:*478
Effluent limitations *1:*755 *4:*656 *4:*664
  Clean Water Act and *1:*755
  by industrial category *1:*757
  statutes concerning *4:*664
  storm water *1:*759
  by substance *1:*757
  water quality-based *1:*759
  under the Clean Water Act *4:*596
Effluent quality, in developing countries *1:*720
Effluent-receiving water systems, pharmaceutical detection in *1:*376
Effluent remediation, for mines *1:*610
Effluent treatment
  of algal populations *2:*3
  for flow-through aquaculture systems *3:*543
  gully pot *3:*499
Effluent water regulations, in arid lands *2:*475
Eggshell thinning *3:*514
  DDT and *3:*520
Egypt
  irrigation in *4:*739
  wastewater applications in *1:*634
  water symbolism in *4:*785
Egyptian water clocks *4:*705
Egypt-Sudan water dispute *2:*591
E_h 2:464
  application of *2:*467
  difficulty in using *2:*467
  interpreting *2:*467
  measurement of *2:*466
Eh/pH conditions, changes in *5:*422
Elastic deformation *5:*482
Elastic storage coefficient *5:*482

Index terms	**Links**			
Electrical charge	*5:*184			
Electrical conductivity (EC)	*3:*682	*4:*429	*5:*184	*5:*195
alternative measurement technologies for	*4:*432			
measurement of	*4:*430			
of soil solution extracts	*3:*673			
temperature effects related to	*4:*430			
total dissolved solids and	*4:*432			
Electrical current	*4:*442			
effect on corrosion	*1:*8			
Electrical geophysical methods	*5:*146			
Electrical logs	*5:*151			
Electrical resistance blocks	*5:*541			
Electrical resistivity (ER)	*3:*675			
Electrical resistivity method	*5:*146			
Electrical sounding process	*5:*446			
Electrical tape method	*5:*101			
Electrical wastewater treatment technologies	*1:*812			
Electric fields	*4:*443			
Electric generating plants				
air emission contaminants from	*1:*554			
contaminant release by	*1:*553			
effects of contaminants from	*1:*556			
history of	*1:*553			
solid waste and water contaminants from	*1:*555			
Electricity				
as a fluid	*4:*442			
hydropower-supplied	*3:*199			
world generation of	*3:*199t			
Electricity generation/production, water use in	*1:*552	*1:*561		
risk analysis of buried wastes from	*5:*448			
Electric pump motors	*3:*380			
Electric utilities, wastewater treatment facilities at	*1:*850			
Electrobiochemical oxidation	*5:*122			
Electrobiochemical processes, engineered	*5:*116			
Electrochemical corrosion measurement	*1:*9			
Electrochemical oxidation	*5:*122			
Electrochemical processes, groundwater remediation via	*5:*283			
Electrochemical reactions, for groundwater contaminants	*5:*283			
Electrochemical regeneration	*1:*917			
Electrochemical technology, problems with	*5:*284			

Index terms	Links			
Electrode configurations, for geoelectrical operations	*5:*445			
Electrodialysis (ED)	*1:*219	*1:*297	*1:*300	*1:*459
	*1:*812			
for arsenic removal	*1:*638			
bonding agent regeneration using	*1:*590f			
desalination by	*1:*171			
Electrodialysis reversal (EDR)	*1:*218	*1:*301		
Electrodialysis reversal system	*1:*218			
Electroflotation	*1:*687			
Electrokinetic enhanced soil bioventing	*5:*121			
Electrokinetic field surveys	*5:*119			
Electrokinetic gradient	*5:*184	*5:*186		
Electrokinetic parameter testing	*5:*186			
Electrokinetic processes	*5:*118			
Electrokinetic remediation technologies	*5:*435			
Electrokinetic systems, components of	*5:*120			
Electrokinetic treatment	*5:*329			
Electrolysis	*1:*812			
in heavy water production	*4:*464			
of MtBE and benzene	*5:*122			
Electrolyte surface tension, role of dissolved gases in	*4:*452			
Electromagnetic geophysical methods	*5:*149			
Electromagnetic induction (EMI)	*3:*676	*5:*302		
Electromagnetic surveys, riverine	*5:*679			
Electromigration	*5:*284			
Electron acceptors	*1:*690	*1:*691		
Electron activity. *See* pe				
Electron beam irradiation, at CSO facilities	*1:*786			
Electron capture detector (ECD)	*2:*307			
"Electronic circuit-rider"	*1:*450			
Electronic leak testing	*1:*887			
Electronic nose				
environmental applications for	*2:*283			
future of	*2:*283			
for odor monitoring	*2:*281			
user needs for	*2:*281			
Electron transfer (ETr) assay	*2:*376			
Electron transfer reaction (ETR) assay	*2:*279	*2:*280		
Electro-optical wave propagation	*4:*296			
Electroosmosis technology	*5:*284			

Index terms	**Links**			
Electroplating industry, impact on surface water quality	*3:*374			
Electrospray ionization mass spectrometry (ESI MS), analysis of aqueous solutions using	*4:*443			
Electrospray ionization MS (ESI-MS) quadrupole	*2:*61			
Elemental mode	*4:*445			
Elements, water symbolism and	*4:*786			
Elevation effect	*5:*230			
Elevation head	*5:*171			
El NiZo phenomenon	*4:*22	*4:*43	*4:*91	*4:*184
	*4:*185			
El NiZo/Southern Oscillation (ENSO)	*4:*43			
index	*4:*174			
prediction of	*4:*23			
Eluviation	*5:*260			
Embankment design, levee	*3:*287			
Embankments, seepage through	*3:*290			
Embankment slopes, levee	*3:*288			
Embankment zoning, levee	*3:*290			
Emergencies, food and water in	*2:*526			
‘‘Emergency areas’’ policy	*2:*526			
Emergency line break repairs	*1:*401			
Emergency outdoor water sources	*2:*527			
Emergency response, coordinating actions for	*1:*871			
Emergency response plans (ERPs)	*1:*528	*2:*348		
Emergency storage	*1:*409			
Emergency supplies	*2:*529			
Emergency water purification	*2:*527			
Emergent aquatic plants	*3:*743			
Emerging waterborne infectious diseases	*1:*177			
pathogens responsible for	*1:*178			
specific	*1:*178			
Emission(s)	*4:*321			
greenhouse gas	*3:*180			
Emitters				
clogging of	*3:*618	*3:*626		
microirrigation	*3:*615			
Empirical area reduction method	*3:*410			
Empirical infiltration models	*4:*486			
Empirical methods, for computing time of concentration	*3:*471			

Index terms **Links**

Empirical models *3*:328 *3*:343
  in industrial water demand economics *1*:550
Employees, security as a priority for *1*:870
Empty bed contact time (EBCT) *1*:103 *1*:351 *1*:352 *1*:579
Enabling environment, in the integrated water resources management framework *2*:575
Enclosed recirculating systems *1*:558
Endangered Species Act (ESA) *4*:664
Endemic disease *1*:187
Endemic fluorosis *4*:435
Endemic stream animals, Hawaiian Islands *4*:803
Endocrine-disrupting compounds (EDC) *1*:819
Endocrine disruption (ED) *3*:54
Endocrine system, effects of lead on *2*:436
‘‘End of pipe’’ drainage solutions *3*:498
Endosulfan *3*:646
Endpoint selection
  in environmental risk assessment *2*:459
  for sediment toxicity tests *2*:459
Energy. *See also* Renewable energies
  conservation of *4*:264
  defined *5*:170
  weather and *4*:353
Energy balance (EB) *3*:561
  global average *4*:179
  remotely sensed *3*:562
  satellite *3*:564
Energy balance equation *3*:535
Energy balance methods *3*:572
Energy budget, balancing *4*:169
Energy characteristics, of rain *4*:331
Energy costs, pumping station *3*:380
Energy dissipation *1*:558
  using enclosed recirculating systems *1*:558
  using once through systems *1*:559
  using open recirculating systems *1*:559
Energy equation *3*:197
Energy flow, metabolism and *5*:528
Energy industry materials, water use for *1*:562
Energy-mass-momentum balances *5*:558
Energy pricing, in India *2*:557

Index terms **Links**

Energy production
  backup energy and electricity for *1:*563
  indirect water use for *1:*563
  iron-reducing bacteria in *2:*151
  material requirements for *1:*562
  sediment transport and *3:*508
  water supply for *2:*618
  water use in *1:*560
Energy use, for evapotranspiration *4:*228
Enforcement agencies, federal *4:*650t
Engineered carbons *2:*82 *2:*85
Engineered systems, biochemical oxygen demand in *1:*639
Engineering risk index *2:*675
Engineering techniques, ditch-related *3:*88
Engineering works, increasing water supply via *2:*490
English Reasonable Use Rule *4:*628
Enhanced Stream Water Quality Model (QUAL2E) *2:*326 *2:*329 *2:*334
Enhanced biodegradation, of vinyl chloride *5:*638
Enhanced bioremediation *5:*319 *5:*424
Entamoeba 1:279
Enteric microorganisms *1:*821
Enteric viruses *1:*821
Enteritis
  astrovirus *1:*179
  calicivirus *1:*179
Enterococcus group, as indicator organisms *2:*293
Enterohemorrhagic *Escherichia coli* (EHEC) *1:*180 *2:*138
  See also Hemorrhagic *Escherichia coli* outbreak
Enteroviruses *1:*70 *1:*178 *1:*279 *1:*522
Entropy
  defined *4:*218
  as a modeling tool *4:*220
Entropy-spectral analysis, for flow forecasting *4:*221
Entropy theory, for hydrologic modeling *4:*217
Environment(s)
  anthropogenic impacts on *5:*62
  benefits of water recycling to *2:*610
  contentious issues regarding *4:*717
  lead movement in *2:*433
  pesticide chemistry in *3:*647

Index terms **Links**

- Environment(s) *(Continued)*
  - pharmaceuticals in *1:*372
  - role of small water reservoirs in *1:*403
  - sources of heavy metals in *5:*276
  - water-limited *2:*516
- Environmental Assessment, Phase I *5:*437
- Environmental Assessment (EA) reports *1:*377
- Environmental contaminants
  - bioaccumulation of *2:*428
  - using fish cells in toxicological evaluation of *3:*115
- Environmental Control and Life Support System (ECLSS) *4:*572
- Environmental Defense Fund (EDF) *3:*513 *3:*514 *3:*522
  - epidemiologic study *1:*289
- Environmental economics *2:*127
- Environmental exposure, to cadmium *5:*614
- Environmental flow regimes
  - data required for *3:*167
  - minimum *3:*166
  - recommendations related to *3:*167
- Environmental flows *3:*106
  - objectives of *3:*167
  - principles of *3:*106
- Environmental Fluid Dynamics Code (EFDC) model *2:*256
- Environmental groundwater tracers *5:*502
- Environmental health, weather modification and *4:*188
- Environmental health programs *1:*799
- Environmental Hydroscience Discipline Group *3:*231
- Environmental impact(s)
  - of carbonate *4:*416
  - of fish farms *3:*579
  - of flow-through aquaculture systems *3:*542
  - integrating into water resources planning *2:*520
  - of iron in groundwater *5:*608
  - of landfills *1:*696
  - of recirculating aquaculture systems *3:*544
  - tools to evaluate *2:*516
  - of weed control *3:*744
- Environmental impact assessment (EIA) model *2:*520
- Environmental impact factors (EIF) *2:*520

Index terms **Links**

- Environmental isotopes
  - applications of *5:*232
  - in contaminant hydrogeology *5:*217
  - fractionation of *4:*501
  - as groundwater tracers *5:*502
  - in hydrogeology *5:*227
  - practical applications of *5:*221
  - stable *5:*228t
- Environmental justice *2:*668
- Environmental lapse rate (ELR) *4:*370
- Environmental management, of ecoregions *2:*667
- Environmental media, arsenic concentrations in *1:*82t
- Environmental meteorologists *4:*328
- Environmental monitoring
  - biosensors as tools for *2:*455
  - considerations for *2:*263
  - using submitochondrial particle assay *2:*378
- Environmental monitoring programs, data from *4:*113
- Environmental photochemistry, in surface waters *4:*529
- Environmental pollution, from electric generating facilities *1:*553
- Environmental problems, worldwide *1:*586
- Environmental projects, groundwater sampling techniques for *5:*454
- Environmental Protection Agency (EPA) *1:*5 *4:*379 *4:*691
  - *See also* EPA entries; United States Environmental Protection Agency (USEPA); U.S. EPA Subtitle D dry tomb landfilling
  - coral reef protection by *4:*115
  - criteria for toxics *4:*598
  - lake protection standards *3:*270
  - National Drinking Water Regulations and *1:*480
  - National Primary Drinking Water Standards *1:*478t
  - nonpoint source control and *2:*185
  - permit programs *4:*666
  - pharmaceutical release and *1:*377
  - primary and secondary water quality standards of *2:*373
  - priority pollutants list *2:*19 *4:*599
  - state reporting to *2:*315
  - water reuse guidelines of *1:*806t
  - water security and *1:*436
- Environmental Protection Council, Ho Chi Minh City *2:*553

Environmental quality
aquatic-environment assessment of *2:*94
standards for *2:*46
Environmental regulations, for fish farms *3:*580
Environmental remediation technologies *3:*364
Environmental resources management *3:*492
Environmental responsibility, new age of *4:*730
Environmental risk assessment processes (ERAs) *2:*459
Environmental samples, of vinyl chloride *5:*637
Environmental science precautionary principle regarding *2:*599
Environmental studies, isotopes used in *5:*218t
Environmental toxicants *2:*106
Environmental tracer recharge measurement *5:*74
Environmental tracers *5:*65
Environmental water quality, evaluation of *2:*450
Environmental water requirements (EWR) *3:*166
methods of determining *3:*167
Environmental weed control *3:*747
Environment Canada (EC) *2:*321 *2:*459 *3:*633 *3:*635
Enzymatic biodegradation mechanisms *2:*42
Enzyme assays *2:*110
Enzyme-linked immunosorbent assay (ELISA) *1:*90 *1:*160 *2:*139 *2:*337 *2:*390
Eosin, in groundwater tracing *5:*506
EPA National Pretreatment Program *1:*798
See also Environmental Protection Agency (EPA)
environmental/public health programs and *1:*799
future of *1:*800
partnerships and *1:*798
EPA priority pollutants *2:*305t
See also Environmental Protection Agency (EPA)
EPA Regional Ground Water Forum *5:*578
See also USEPA codes
EPA Stormwater Management Model (SWMM) *3:*256 *3:*337
EPA Superfund *1:*434 *2:*517
See also Superfund sites
EPA toxic constituent manuals *2:*381
EPDM membranes *1:*627
Ephemeral rivers, sand abstraction from *3:*414
Ephemeral streams, incisement of *3:*57

Index terms **Links**

Ephemeral waters *2:*475
Epidemics, nineteenth-century *4:*729
Epidemiological studies *1:*187
  lead-related *2:*438
Epidemiologic data, use of *1:*423
Epidemiology
  Cryptosporidium 1:164
  Escherichia coli O157:H7 2:137
  giardiasis *1:*258
Epikarst *5:*108
Epikarstic dye introduction points (EDIPs) *5:*109
Epikarstic zone *5:*236 *5:*243
Epilimnetic hypoxia *4:*65
Episodic acidification *3:*5
  biological importance of *3:*7
  causes of *3:*5
  modeling of *3:*5
  regional differences in *3:*6
Epoxy resin pipe linings *1:*879 *1:*880
EQ3/6 program *5:*141
Equation of continuity *5:*558 *5:*651
Equation of motion *5:*559 *5:*651
Equalization *1:*814
  domestic sewage *1:*832
  of landfill pollutant loadings *1:*705
  storage tanks and *1:*448
Equalizing storage *1:*409
Equations, water quality model *2:*273
Equations of change *5:*558
  in laminar flow systems *5:*651
Equilibrium
  kinetic modeling and *4:*565
  reactive transport models *5:*519
Equilibrium models *5:*414
Equilibrium analysis *2:*71
Equilibrium discharge (Q_e) *3:*63
Equilibrium freezing point *4:*585
Equilibrium phase partitioning *5:*549
Equilibrium sorption models *1:*368
Equilibrium-tension lysimeters *5:*489

Index terms	Links		
Equilibrium time (*t*)	*3:*298		
Equilibrium vapor pressure	*4:*363	*4:*365	
Equipment, rain-simulator	*4:*330		
Equipotential line	*5:*419		
Equipotentials	*5:*182		
Equity, water quality and	*2:*131		
Equivalent atmospheric concentration (EAC)	*4:*421		
Eroded sediment, deposition of	*3:*567		
Erosion	*2:*543	*5:*125	
See also Erosion control			
at construction sites	*1:*537		
gully	*3:*183		
river	*2:*657		
sediment production by	*3:*31		
Erosion control	*3:*697		
agronomic methods of	*3:*568		
mechanical methods of	*3:*567		
Erosion cycles	*3:*184		
Erosion-Productivity Impact Calculator (EPIC)	*2:*251		
Escherichia coli	1:277	2:21	2:22
lux-gene-carrying	*2:*174		
outbreaks of	*1:*341	*1:*430t	
waterborne transmission of	*1:*429		
Escherichia coli O157:H7	2:136	2:344	
clinical features and treatment for	*2:*136		
diagnostic procedures for	*2:*139		
epidemiology of	*2:*137		
history of	*2:*136		
monitoring in water	*2:*140		
pathogenesis of	*2:*137		
susceptibility to chlorination	*2:*138		
transmission routes of	*2:*138		
waterborne transmission of	*2:*138		
Espy, James Pollard	*4:*353		
Estuaries	*2:*543	*4:*25	*4:*49
budgetary models for	*4:*52		
coastal	*3:*72		
as disposal sites	*4:*50		
importance of	*4:*49		
physicochemical processes in	*4:*50		
role of nutrients in	*4:*51		

Estuaries *(Continued)*
  water quality goals for *4:*54
Estuarine coastal wetlands *3:*72
Estuarine microalgal sediment toxicity tests *4:*120
Estuarine waters, metals in *4:*29
Ethanol *3:*546
  soil vapor concentration of *5:*553
Ethics, in forensic hydrogeology *3:*170
Ethnic community, protecting *2:*667
Ethnopolitics, of hydro-borders *4:*681
Ethylenediaminetriacetate (EDTA) *1:*779
EU Freshwater Fish Directive *2:*267
Eukarya *2:*313
EU Landfill Directive *1:*695
Eulerian coordinate system *5:*558
Eulerian models *1:*133
Euler's equations *4:*352
EU Nitrate Directive *3:*637
Euphotic zone *3:*265 *3:*267
Europe
  fish ponds and fish production in *4:*721t
  land drainage in *3:*490
European Ecological Network (ECONET) *1:*404
European Environmental Agency *3:*292
European Food Safety Authority, fish advisories *3:*119
European policy, precautionary principle and *2:*597
European Union (EU) *1:*407
  See also EU entries
  flooding directives *3:*162
  pharmaceutical guidelines of *1:*377
  wastewater treatment within *1:*844
  water quality standards *2:*267
European Union-Water Framework Directive (WFD) *2:*638 *2:*639 *2:*640 *2:*641
Eutrophication *2:*2 *2:*181 *2:*269 *3:*107 *3:*704
  See also Cultural eutrophication
  ammonia and *4:*392
  controlling *3:*109
  health implication of *5:*334
  in Indian lakes *3:*451

Index terms	**Links**			
Eutrophication *(Continued)*				
in Indian rivers	*3:*449			
lake	*3:*283			
relation to organic loading	*2:*142			
in surface waters	*3:*695	*3:*696	*3:*702	
Eutrophication impacts, on water quality	*3:*107			
Eutrophic water bodies, algal control in	*2:*2			
Evacuated-tube solar collector	*1:*66			
Evaluative criteria, in a systems approach	*2:*685			
Evaporation	*1:*812	*2:*543	*3:*192	*4:*192
	*4:*223			
agricultural impact of	*4:*226			
average estimated	*4:*226t			
crop	*3:*720			
determining from interception	*3:*237			
estimating	*4:*2			
in the hydrologic cycle	*4:*276			
hydrologic impacts of	*4:*225			
mathematical estimates of	*4:*225			
measurement and estimation of	*4:*224			
physical controls on	*4:*223			
in the water cycle	*4:*196			
Evaporation/condensation, latent heat of	*3:*193			
Evaporative rates, mean daily	*4:*260			
Evaporite mineralogy	*5:*52			
Evaporites	*5:*53			
Evapotranspiration	*2:*543	*3:*30	*3:*171	*3:*222
	*3:*225	*4:*226		
in arid regions	*5:*410			
concentration effect of	*4:*374			
crop	*3:*557			
estimating	*4:*227			
importance of	*4:*226			
magnitude of	*4:*229			
measuring	*4:*227			
modeling of	*4:*259			
remote sensing of	*4:*321			
requirements for	*4:*228			
role in drainage design	*5:*96			
tomato-crop	*3:*576			
in the vadose zone	*5:*534			

Index terms	**Links**	
Evapotranspiration *(Continued)*		
versus transpiration	*3:*733	
Evapotranspiration estimates	*5:*78	
Chicago Station	*5:*79t	
reference	*3:*573t	
Evapotranspiration from groundwater (*Et*)	*5:*168	
Event-based watershed response, modeling	*3:*343	
Evolution, of water	*4:*715	
Evolutionary algorithms (EAs)	*2:*333	
EXAFS spectroscopy method	*2:*207	*2:*208
Excavation	*1:*402	*1:*887
levee	*3:*288	
Excess fishing capacity	*3:*127	
Excess rainfall intensity (*E*)	*3:*63	
Excreta, approaches for handling	*1:*676f	
Exhaustion phenomenon	*3:*23	
Exohydrology	*5:*270	
Exopolymeric substances (EPS)	*4:*27	*4:*29
Expanded Granular Sludge Bed (EGSB)	*1:*906	
‘‘Explosimeters,’’	*2:*308	
Exponential recession	*3:*22	
Export of virtual water	*2:*536	
Exposure		
presence of food during	*2:*420	
use of semipermeable membrane devices for	*5:*672	
to vinyl chloride	*5:*635	
Exposure Analysis Modelling System (EXAMSII)	*2:*327	
Exposure assessments	*1:*424	
Exposure biomarkers	*2:*28	*2:*31
Exposure concentrations	*2:*422	*2:*423
time-varying	*2:*213	
Exposure methods	*2:*420	
Exposure tests, sediment or fauna incubation experiment for	*2:*418	
Expressed sequence tag (EST) databases	*2:*62	
Ex situ physical/chemical treatment, of vinyl chloride	*5:*639	
Ex situ solidification/stabilization processes	*1:*837	
Ex situ vitrification	*5:*618	
Extended aeration	*1:*815	
External mass transfer	*4:*567	
Extracellular biodegradation	*2:*43	

Index terms	Links	
Extracellular complexation, of metals	*5:*283	
Extracellular fluid (ECF)	*4:*459	*4:*460
Extracellular matrix (ECM)		
endocrine system and	*4:*461	
water and	*4:*459	
Extracellular molecules, heavy metal binding to	*5:*378	
Extracellular polymeric substances (EPSs)	*2:*229	
Extraction, of heavy metals	*5:*435	
Extraction estimates, Chicago Station	*5:*79t	
Extraterritorial land use control	*1:*315	
in Boston	*1:*376	
choices related to	*1:*317	
in New York	*1:*316	
in Syracuse	*1:*317	
Extremely low frequency (ELF) spectrum	*1:*513	
Exudates, as chemical signals	*5:*379	
Exxon Valdez oil spill	2:292	4:597

F

Facies concept	*5:*52	
Facultative aerobes/anaerobes	*5:*581	
Failure		
defined	*2:*675	
reservoir	*3:*259	
Failure region boundary	*2:*676	
Fairness, public water supply and	*1:*501	
Fair weather flow. *See* Base flow		
Fair weather waterspouts	*4:*347	
Falling water, between two level drains	*5:*96	
Falling water table, between two bilevel drains	*5:*96	
Falls	*3:*70	
‘‘Fanya juu’’ terrace	*3:*550	
Far East		
clock making in	*4:*706	
early clocks in	*4:*704	
Far Eastern Russia seas		
biosurveys in	*2:*444	
evaluation of heavy metal pollution in	*2:*447	
seawater quality control in	*2:*449	
Farm and Ranch Lands Protection Program (FRPP)	*3:*596	

Index terms **Links**

Farming. *See also* Agricultural entries; Agriculture; Crop entries; Fish farms; Livestock entries
  irrigation techniques and *3:*591
  organic *3:*644
Farmland, irrigated *2:*549
Farm runoff *2:*400
Fast atom bombardment mass spectroscopy (FABMS) *2:*389
Fate and transport models *5:*589
Fatty acid esters (FES) *1:*672
Fatty acids *2:*442
Faucets, low-flow *2:*665
Faults, rock *5:*137
Feasibility characterization, for *in situ* electrokinetic treatment *5:*118
Fecal bacteria, as water quality indicators *2:*23t
Fecal coliforms (FC) *2:*293 *2:*344
  See also Coliforms
  in groundwater *2:*396
  as indicator bacteria *2:*361
  modeling *3:*334
Fecal contamination, in Indian rivers *3:*449
Fecal-oral infection route *1:*178
Feces, as a source of infectious agents *3:*668
Federal Emergency Management Agency (FEMA) *2:*526 *2:*529 *3:*147
Federal Fish & Wildlife Service *3:*518
Federal government. *See also* Federal regulations; United States; U.S. entries
  approach to droughts *2:*578
  drought-related water quality and *2:*585
  as waterway decision maker *2:*523
Federal Highway Administration (FHWA) *2:*186
  model *2:*253
Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) *5:*357
Federal lands, reserved water rights for *4:*689
Federal law, protection of in-stream values by *4:*663
Federal legislation, groundwater *4:*631
Federal regulations. *See also* Federal government
  history of *2:*194
Federal-state regulatory interface *4:*671
  water-quality-related *2:*190

Federal Water Pollution Control Act (FWPCA) *2:*125 *2:*194 *4:*595 *4:*664 *4:*671 *4:*691
  amendments to *2:*195 *2:*314
  1961 amendments to *4:*652
  1972 amendments to *4:*598
Federal water projects, expansion of *2:*522
Federal water resource agencies *4:*649
Federal water supply planning *4:*612
  developments in *4:*612
  federal-state cooperation in *4:*612 *4:*614
  multiprogram *4:*613
  Water Resources Council and *4:*613
Feedback, in hydrologic systems *3:*230
Feeding status, chlorine toxicity and *2:*403
Feed water
  chemistry of *1:*414
  quality of *1:*175
Feed wells, flocculating center *1:*454
Female reproductive system, effects of lead or *2:*436
FEMWATER model *5:*660 *5:*663
Fen draining *3:*88
Fenton process *1:*871
  See also Fenton’s Reaction
Fenton’s Reaction *4:*445 *5:*441
  case studies involving *4:*447
  delivery options related to *4:*447
  equations related to *4:*446
  history of *4:*446
  in situ applications of *4:*447
Fenton’s Reagent, chemical oxidation with *5:*345
Fermentation, in landfills *1:*696
Fern Lake, biomanipulation trials in *2:*57t
Ferrel, William *4:*353
Ferric hydroxides *2:*363 *5:*19
  See also Iron entries
Ferric iron *4:*496
Ferrobacillus ferrooxidans *3:*14
Ferrous iron *4:*496 *5:*609
Fertigation *3:*623

Index terms	**Links**		
Fertilization, with nitrogen	*3:*638		
See also Eutrophication			
Fertilizer application			
methods of	*2:*572		
rates of	*2:*571		
Fertilizer management, as a nitrogen best management practice	*3:*698		
Fertilizers			
impact on surface water quality	*3:*374		
leaching of	*3:*651		
nitrogen in	*3:*695		
Fiberscopes	*5:*402		
Fick’s first law	*5:*550	*5:*551	*5:*553
Field capacity concept	*5:*124	*5:*536	
applications of	*5:*126		
history of	*5:*124		
measurement deficiencies in	*5:*127		
relation to water holding properties	*5:*125		
water content predictions and	*5:*126		
Field investigation, for levee construction	*3:*286		
Field irrigation recharge (R_i)	*5:*166		
Field measurements	*3:*321		
Fields			
diked	*4:*696		
floating	*4:*695		
Field sampling			
of contaminants	*2:*263		
when and how often to sample	*2:*264		
where to sample	*2:*263		
Field studies			
on mussels	*1:*512		
on veliger settling	*1:*513		
Field surveys			
biochemical	*5:*119		
electrokinetic	*5:*119		
Filamentous bacteria	*2:*150		
Filamentous organisms, foam-producing	*1:*728		
Filaments			
identification of	*1:*846		
isolation and characterization of	*1:*847		
“Filming amines”	*1:*545		

Index terms **Links**

Filling time *3:*266
Fill material permits *4:*665
Filter density (ρ_{GAC}) *1:*103
Filter effluent turbidity *1:*224
Filtering screens *3:*748
Filter operation time (t_F) *1:*103
Filter pack *1:*151
Filter performance, water treatment plant *1:*222
Filters *1:*379
  See also Filtration
  biofouling of *2:*99
  biophysicochemical interactions with colloids and dissolved organics *2:*100
  iron-removal *2:*152
  large-pore *2:*101
  microirrigation *3:*623
  rapid granular bed *1:*234
  roughing *1:*237 *1:*486
  slow sand *1:*233 *1:*235
  types of *1:*817 *2:*100
Filter underdrain systems *1:*234
Filter (linear) velocity (u_F) *1:*103
Filtration *1:*227 *1:*817 *2:*374
  See also Diatomaceous earth filtration; Filters; Filtration systems
  coarse gravel *1:*239
  for CSO treatment *1:*786
  domestic sewage *1:*834
  drinking water *1:*227
  granular bed *1:*249
  hydrous manganese oxide *1:*398
  of iron and manganese *1:*314
  microbial regrowth and *1:*344
  multistage *1:*238
  particulate removal by *1:*243
  precoat *1:*251
  slow sand *1:*239 *1:*249 *1:*431
  for small drinking water systems *1:*458
  sorptive *2:*362
  surface water treatment rule compliance technologies for *1:*462t
  wastewater *1:*810
  water treatment plant *1:*224
  water treatment via *1:*245

Filtration hydraulics, granular media filtration *1:*233
Filtration spectrum *1:*333f
Filtration systems
  comparison of *1:*228
  package plant *1:*515
  selecting *1:*230
  for storm water treatment *1:*868
Financial performance, of water companies *1:*310
Fine bubble diffused air aeration systems *1:*626
  design considerations for *1:*629
  diffuser types *1:*626
  fouling, scaling, and clogging of *1:*630
Fine screens *1:*785
Finfish, stock enhancement of *4:*125
Fingerprint biomarkers *2:*29 *2:*31
‘‘Fingerprinting’’ analysis *2:*307
Finite differences (FD) method *5:*310 *5:*621
Finite element method/modeling (FEM) *5:*311 *5:*621 *5:*655
  benefits of *5:*660
  for incompressible flow problems *5:*657
  principles of *5:*656
Finite elements (FEs), types of *5:*656
Finite volumes (FV) method *5:*311 *5:*623
Fire demand constraints, in water distribution systems *1:*209
Fire flow tests *3:*316
  See also Combustible watersheds; Corin Dam catchment fire
Fire management *2:*201
Fire storage *1:*409
Firm yield *3:*261
First law of thermodynamics *5:*170
First-order degradation *5:*32
Fiscal solutions, in the Arab World *2:*473
Fish. *See also* Fisheries entries
  ambient exposure routes for *3:*116
  biomanipulation of *2:*51
  effect of road salt on *2:*323
  herbivorous *3:*746
  impacts of PCBs and dioxin-like compounds on *2:*108
  loss to acidification *3:*9
  pesticides in *3:*435

Index terms	**Links**
Fish *(Continued)*	
Puget Sound Basin	*3:*349
trace elements in	*3:*456
vulnerability to pollutant discharge	*3:*283
worldwide consumption of	*3:*123
Fish and Wildlife Service (FWS)	*4:*691
Fish assemblages, managing	*3:*134
Fish cells	
bioavailability and mode/mechanism of action of	*3:*117
in toxicological evaluation	*3:*115
Fish cell toxicity end point, importance and relevance of	*3:*116
Fish consumption advisories	*3:*118
example of	*3:*120
history and scope of	*3:*118
methods associated with	*3:*119
theory behind	*3:*119
Fisheries	*3:*121
See also Aquatic environments; Fisheries management	
benefits of	*3:*122
defined	*3:*122
ecosystem management and	*3:*128
effect of sediment on	*3:*508
history of	*3:*122
impact of eutrophication on	*3:*108
impact of oil on	*4:*106
species and habitat preservation and	*3:*128
Fisheries management	*3:*122
See also Freshwater fisheries management	
coastal	*3:*126
diadromous	*3:*127
evolution of	*3:*124
future of	*3:*129
lake	*3:*125
open ocean or high seas	*3:*126
riverine	*3:*125
scope of	*3:*133
sustainable	*2:*635
Fisheries oceanography	*4:*55
Fisheries Oceanography Coordinated Investigations program	*4:*90
Fisheries resources, current uses of	*3:*123
Fishery biology, problem of	*4:*143

Index terms		**Links**
Fishery stock assessments	*4:*56	
Fishery sustainability, precautionary principle regarding	*2:*600	
Fish farms		
environmental impacts of	*3:*579	
environmental regulations concerning	*3:*580	
waste treatment in	*1:*681	
water pollution from	*3:*579	
Fish growth and production, factors affecting	*3:*129	
Fish habitat, protecting	*3:*129	
Fish index of biotic integrity (Fish IBI)	*3:*38	
Fishing waters	*3:*532	
Fish lift fish passage facility	*3:*530	*3:*531f
Fish lock fish passage facility	*3:*530	
Fish models	*3:*115	
Fish passage facilities	*3:*529	
Fishpond culture, history of	*3:*135	
Fish ponds, water retention potential of	*1:*407	
Fish populations, serial depletion of	*3:*126	
Fish removal, algal control and	*2:*6	
Fish species, invasive	*3:*284	
Fish species richness, anoxic factor and	*4:*66	
Fixed bed adsorber	*1:*102	
regeneration of	*1:*104	
Fixed-bed sorption	*2:*363	
Fixed bed studies	*1:*115	
Fixed-film aerobic systems	*1:*704	
Fixed-film processes	*1:*815	*1:*833
Fixed fine screens	*1:*785	
Fixed-tension lysimeters	*5:*489	*5:*488
Fjord lakes	*3:*269	
Flagellates	*2:*313	
Flame ionization detector (FID)	*2:*307	
Flame photometric detectors	*2:*308	
Flash flood warning systems	*1:*125	
Flash floods	*3:*142	
Flashing	*1:*484	
Flashy streams	*3:*104	
Flat-plate solar collector	*1:*65	
Flat rate structure	*2:*663	

Index terms **Links**

Flat roofs *1:*55
Flattop Mountain SNOTEL (SNOw TELemetry) station *4:*334
  snowpack at *4:*334
Flav-Savr tomato *3:*686
Fleet Numerical Meteorology and Oceanography Center (FNMOC) *4:*96
Flexible repair coupling *1:*889
Floatable debris-related legislation *4:*39
Floatables control, netting systems for *1:*785
Floating agriculture *4:*695
Floating breakwaters *4:*20
Floating ice *4:*69
Floating-leaved aquatic plants *3:*743
Floating plants *3:*743
Flocculant polymers *4:*588
Flocculants *4:*425
Flocculant settling *1:*243 *3:*404
Flocculating center feed wells *1:*454
Flocculation *1:*227 *1:*252 *2:*98 *2:*374 *4:*424 *4:*586
  ballasted *1:*454 *4:*426
  efficiency of *4:*427
  mechanisms of *4:*424
Flocculation reaction *5:*619
Flood basalts, perched groundwater in *5:*353
Flood bypasses *3:*151
Flood control
  reservoir operation for *3:*386
  Yellow River Basin *3:*45
Flood Control Act of 1944 *2:*523 *4:*616
Flood control engineering system, Yellow River *3:*48
Flood control storage, design of *3:*264
Flood control structures *3:*150
  diversion structures *3:*151
  failure of *3:*145
  levees *3:*152
  peak flow reduction using *3:*150
Flood conveyance capacity, decreasing *3:*50
Flooded soils *3:*206
Flooded zones, agricultural use of *4:*695
Flood farming practices *3:*164

Index terms	**Links**	
Flood frequency analysis	*3:*46t	
Flood index	*3:*156	*3:*159
analysis	*3:*158	
Flooding. *See also* Floods; Urban flooding		
causes of	*3:*511	
Ganga Basin	*3:*234	
general conditions for	*3:*45	
human response to	*3:*149	
regulation of	*3:*382	
tolerance for	*3:*601	
Yellow River Basin	*3:*45	
Flood insurance	*3:*147	
Flood insurance rate maps (FIRMs)	*3:*528	
Flood irrigation, suitability criteria for	*3:*164	
Flood management	*3:*512	
participatory multicriteria	*2:*678	
Red River Basin, Manitoba, Canada	*2:*681	
stakeholder participation in	*2:*680	
structural and nonstructural	*3:*512	
Flood managers, Chinese	*3:*48	
Flood mapping, remote-sensing	*4:*322	
Floodplains	*2:*543	*3:*527
importance of	*3:*527	
mapping	*3:*147	
regulatory definition of	*3:*528	
remote sensing and GIS application in managing	*2:*533	
Flood prevention	*2:*525	*3:*510
See also Floodproofing		
Floodproofing	*3:*147	
Flood protection	*1:*498	
levees for	*3:*286	
“Flood Pulse Concept”	*3:*292	
Flood-related warning systems	*3:*148	
Flood retardation basins, Yellow River	*3:*48	
Flood risk, Yellow River	*3:*49	
Flood risk management	*2:*525	
optimal	*2:*526	
Flood routing models, Brakensiek’s	*3:*255	

Index terms	**Links**	
Floods	*2:*543	*3:*142
See also Flooding		
defined	*3:*142	
detection and response warning systems for	*3:*148	
factors leading to	*3:*142	
forecasting	*1:*123	
human responses to	*3:*146	
ice jam	*3:*511	
impacts of	*3:*142	*3:*153
in India	*2:*566	
as natural hazards	*3:*153	
rain-on-snow	*3:*511	
responding to	*3:*512	
snowmelt runoff	*3:*511	
storm-rainfall	*3:*511	
structural and nonstructural measures against	*3:*146	
in the United States	*2:*523	
velocities of	*3:*64	
Flood source mapping		
case study in	*3:*157	
flood index analysis in	*3:*158	
procedure in	*3:*156	
unit flood response approach to	*3:*156	
in watersheds	*3:*155	
Flood stage, defined	*2:*543	
Flood storage, reservoir	*3:*151	
Flood-susceptible land, public acquisition of	*3:*147	
Flood warning systems	*1:*125	
Floodwater harvesting	*2:*550	
Floodwater spreading	*3:*163	*3:*709
advantages and problems associated with	*3:*165	
for artificial recharge	*3:*164	
design factors in	*3:*165	*3:*711
equations governing	*3:*710	
history of	*3:*164	
schemes related to	*3:*165	
Floral/faunal estimates, semiquantitative	*4:*93	
Florida		
coral banks in	*4:*36	
coral black water and	*4:*133	
Florida Bay-Everglades National		

Index terms	**Links**		
Park	*4:*55		
Flotation	*1:*371	*1:*487	*1:*811
sedimentation and	*3:*406		
as a separation process	*1:*684		
Flotation processes, selecting	*3:*407t		
Flotation separation techniques	*1:*594		
Flow(s). *See also* Fluid flows; Pressurized flow			
creeping	*5:*654		
environmental	*3:*106		
laminar	*5:*649		
open channel	*3:*195		
through orifices and weirs	*3:*195		
steady-state and unsteady-state	*5:*491		
subsurface	*5:*498		
in pipes	*3:*195		
in the vadose zone	*5:*537		
viscous	*5:*555		
Flow analysis predictions	*5:*63		
Flow competence	*3:*419		
Flow cytometry	*1:*160		
Flow design, open channel	*3:*346		
Flow direction, identification of	*3:*93		
Flow-duration curves	*3:*102		
relationship to geohydrology	*3:*103		
shape of	*3:*103		
variability indexes and	*3:*104		
Flow equations	*5:*310		
Flow forecasting, entropy-spectral analysis for	*4:*221		
Flowforms	*4:*795		
Flowing artesian well	*5:*600		
Flowing gas-static liquid respirometers	*1:*567		
Flowing well/spring	*2:*543		
Flowmeter log	*5:*154		
Flow methods, instream	*3:*526		
Flow monitoring	*1:*886		
Flow paths, effect on water	*3:*5		
Flow phenomena, categories of	*3:*247		
Flow porosity	*5:*184		
Flow purging	*5:*405		
Flow rate measurement and control, water treatment plant	*1:*222		

Index terms **Links**

Flow rates *1:*132

Flow regime modification *2:*477

Flow regimes, minimum environmental *3:*166

Flow regulation, river *3:*381

Flow restrictor showerheads *2:*666

Flow reversal, Chicago River *3:*41

Flow routing
  channel *3:*253
  HYMO model of *3:*255
  kinematic wave *3:*253
  Kousiss model of *3:*257
  model parameters for *3:*257
  Muskingum-Cunge method of *3:*256
  Muskingum method of *3:*255
  reservoir *3:*255
  SSARR model of *3:*255
  SWMM model of *3:*256

Flowstone *5:*246

Flow systems. *See also* Fluid flow systems; Fracture flow system
  boundary conditions for *5:*14
  one-way *3:*437
  regional *5:*417
  topography and *5:*420

Flow-through (FT) aquaculture systems *3:*542

Flow-through fish farms *1:*681

Flow velocities *1:*205
  from culverts *3:*77
  effect on corrosion *1:*8
  dye tracers for measuring *3:*96

Flue gas desulfurization (FGD) material *1:*556

Flue gas desulfurization sludge disposal site
  aqueous behavior of elements in *1:*848
  FGD scrubber purge water in *1:*849
  limestone/lime-based desulfurization systems and *1:*849
  process description of *1:*848
  water quality in *1:*850

Fluid density, head and *5:*181

Fluid flow equations *5:*309

Fluid flows *3:*194

Fluid flow systems, conservation laws for *5:*651
  See also Flow systems
Fluidity, of sludge *1:*862
Fluidized bed furnace (FBF) *1:*858 *1:*860
Fluid mechanics *3:*197
Fluid momentum equation *3:*198
Fluid Physics Research Program *4:*448
Fluids. *See also* Gas entries; Liquids
  consistency coefficient for *5:*557
  dilatant *5:*557
  non-Newtonian *5:*556
  pseudoplastic *5:*557
  time-dependent *5:*557
  time-independent *5:*556
Fluorescent dyes *3:*96
  in groundwater tracing *5:*108 *5:*505
Fluorescent *in situ* hybridization (FISH) *1:*161 *1:*646
  See also Whole cell fluorescent *in situ* hybridization
  in community analysis *1:*645
  monitoring foam-forming bacteria using *1:*847
Fluorescent-labeled latex beads (FLLB) *2:*415
Fluorescent labeling *2:*59
Fluoridation *1:*254
  cost of *1:*256
  history of *1:*255
  worldwide rates of *1:*257
Fluoride *1:*255 *1:*903
  in bottled water *1:*5
  control limits on *1:*296t
  in groundwater *2:*396
  maximum contaminant level of *2:*234t
  safety of *1:*255 *4:*435
Fluoride concentration *1:*296
  safe *4:*435
Fluoride contamination
  controlling factors in *5:*133
  in groundwater *5:*130
  mechanism of *5:*134
  remedial measures for *5:*134
  sources of *5:*132
Fluorosis *5:*130

- Flushing
  - in algal control *2:*4
  - distribution system *1:*206
  - enhanced *5:*440
  - as a watershed function *3:*478
- Flushing rate, lake *3:*283
- Fluvial lakes *3:*265
- Fluvial sediments, movement of *3:*418
- Fluviogenic areas *1:*404
- Flux(es) *5:*126
  - measurement of *5:*545
  - stoichiometric analysis of *4:*53
- Flux chambers *5:*549
- Flux-concentration relation *3:*248 *3:*250
- Flux laws *3:*249
- Flux-type natural resources *2:*634
- Fly ash production, fluoride contamination and *5:*133
- Foam formation *1:*844
- Foam-forming bacteria, monitoring *1:*847
- Foaming and bulking. *See* Microbial foaming and bulking
- Foam-producing filamentous organisms *1:*728
- Foci detection method (FDM) *1:*160
- Focused electrode leak location (FELL) system *1:*887
- Fog *4:*229 *4:*293
  - coastal *4:*230
  - maritime *4:*232
- Food(s)
  - emergency *2:*526
  - low supplies of *2:*528
  - nitrate in *1:*31 *1:*33
  - shelf-life of *2:*528
  - short-term supplies of *2:*527
  - storage tips for *2:*528
- Food and Agriculture Organization (FAO) *3:*573 *3:*619
- Food and Drug Administration (FDA) *1:*4
  - pharmaceutical release and *1:*377
- Food chain(s) *4:*151
  - *See also* Foodwebs
  - biological magnification of DDT in *3:*523
- Food cycle *4:*152

Index terms	**Links**	
Food processing industry, impact on surface water quality	*3:*374	
Food production, drainage and	*5:*99	
Food security, relation to water security	*1:*438	
Food wastes, constructed wetlands for	*1:*896	
Foodwebs	*4:*152	
See also Food chain(s)		
microorganisms in	*3:*312	
properties of	*4:*153	
Force mains		
point repairs for	*1:*889	*1:*890
repairing	*1:*888	
Forcing factors, current	*3:*322	
Forecast Systems Laboratory (FSL), ground-based GPS meteorology at	*4:*244	
Forensic hydrogeology	*3:*168	
ethics and standards in	*3:*170	
issues in	*3:*169	
tools and approaches for	*3:*169	
Forest(s)	*3:*170	
acidic deposition and	*4:*378	
cutting	*2:*200	
distribution of	*3:*171	
regeneration of	*2:*200	
Forest activities, influence on the hydrologic cycle	*4:*282	
Forested landscape, water quality management in	*2:*199	
Forested wetlands	*3:*172	
Forest fire, impact on water quality	*2:*201	
Forest hydrology research	*2:*199	
Forest Legacy Program (FLP)	*3:*596	
Forest management	*2:*201	
See also Forestry		
Forestry. *See also* Forest management		
Great Lakes region	*3:*178	
nonpoint source pollution from	*2:*187	
sediment problems related to	*3:*508	
Forest soil, disturbance of	*2:*200	
Forest structure		
hydrologic cycle and	*3:*171	
implications for management	*3:*172	
FORTRAN Hydrological Simulation Program (HSPF)	*2:*249	*2:*326

Index terms	**Links**	
Fossil aquifers	*5:*294	
American	*5:*295	
international examples of	*5:*295	
Fossil-fuel combustion, contaminants released by	*1:*554t	
Fossil fuel power, greenhouse gas emissions from	*3:*182f	
Fossil fuels	*1:*560	
biofuel alternatives to	*3:*545	
trace elements in	*3:*458	
use of	*3:*546	
Foulants, types of	*1:*420	
Fouling	*1:*334	
See also Biofouling		
control of	*1:*415	
in diffused air aeration systems	*1:*630	
mussel size and	*2:*402	
Foundation stability, levee	*3:*288	
Foundation underseepage, levee	*3:*289	
Four-electrode measurement	*4:*432	
Fourier Transform Infra Red (FTIR) spectroscopy analysis	*4:*508	
Fractal analysis	*5:*176	
Fractal geometry	*5:*301	
Fractionation, of colloids	*3:*74	
Fractionation factor	*4:*500	*4:*501
Fraction of absorbed photosynthetically active radiation (fAPAR)	*3:*722	
Fractured aquifers	*5:*600	
Fractured igneous rock aquifers	*5:*145	
Fractured rock systems, groundwater flow in	*5:*175	
Fracture flow system, characterization of	*5:*177	
Francois formula	*4:*19	
Franklin, Benjamin	*4:*780	
relationship with water	*4:*741	
swim fin development by	*4:*784	
water music instrument of	*4:*758	
weather theories of	*4:*784	
Franklin/Folger Gulf Stream chart	*4:*782	
Franklin-Gordon Rivers Wild Park (FGRWP)	*4:*789	
Franklin River	*4:*788	
Fredericksborg Castle Lake, biomanipulation trials in	*2:*56t	
Fredericton aquifer case study	*5:*678	
Freeboard	*3:*287	*3:*348

Index terms	**Links**		
Free chlorine	*1:*169	*1:*194	
See also Chlorine entries			
bromide oxidation by	*2:*74		
residual	*2:*89		
Free chlorine/free available chlorine (FC/FAC)	*1:*128		
Free chlorine residual	*1:*662	*1:*663	*2:*398
Free-floating aquatic plants	*3:*743		
Free flow, in culverts	*3:*76		
Free-flowing groundwater	*5:*29		
Free-formed pools	*3:*67		
Free Ion Activity Model (FIAM)	*2:*212	*3:*75	
Free phase chlorinated solvents	*5:*92		
Free water knock out (FWKO)	*2:*285		
Free-water organisms	*3:*310		
Free-water surface flow systems	*3:*365		
Free water surface (FWS) wetlands	*1:*787		
technology for	*1:*894		
Freeze-drying process	*4:*344		
Freeze-thaw sludge conditioning	*1:*864		
Freezing, of water	*4:*585		
Frenchman's Cap	*4:*788		
French water management model	*1:*389		
Frequency domain electromagnetic induction	*5:*149		
Freshwater(s)	*2:*543	*4:*449	
chronic acidification of	*3:*1		
distribution of	*4:*753		
in Lake Vostok	*3:*506		
removing colloids from	*3:*74		
standard composition of	*3:*266		
Freshwater colloids	*3:*74		
environmental importance of	*3:*75		
Freshwater density	*5:*181		
Freshwater fisheries management			
coordinated	*3:*134		
holistic	*3:*134		
water needs for	*3:*133		
Freshwater head	*5:*160	*5:*161	*5:*181
Freshwater net pens	*3:*543		
Freshwater resources			
acidification of	*3:*7		

Index terms	**Links**			
Freshwater resources *(Continued)*				
global	*1:*437			
Hawaiian	*4:*801			
proton concentration, acidity, and dissolved solids in	*3:*8			
sustainable management of	*2:*634			
Freshwater runoff, effect on nutrient concentrations	*4:*81	*4:*82		
Freshwater streambed sediment, organic compounds and trace elements in	*3:*349			
Freshwater supply				
global	*2:*548			
sources of	*2:*515			
Freshwater swamps	*3:*172			
Fresno, pesticide assessment in	*5:*359			
Freundlich adsorption isotherms	*1:*101	*1:*113	*1:*114t	*1:*578
	*1:*777	*4:*566		
See also Isotherms				
Freundlich equation	*2:*71	*2:*87		
Freundlich model	*3:*300	*4:*385	*4:*386	
Friction velocity, estimation of	*4:*1			
Fritz and Schlunder multicomponent adsorption model	*1:*108			
Frontal fog	*4:*231			
Fronts, coastal fog and	*4:*231			
Frost	*4:*240			
damage from	*4:*241			
formation of	*4:*241			
permanent	*4:*305			
Fugacity	*4:*524			
Full cost pricing	*2:*663			
Full privatization model	*1:*50			
Full reservoir level (FRL)	*3:*411			
Full utility concessions	*1:*388			
Fulvic acid	*4:*399	*5:*190	*5:*191	
Funding, of municipal solid waste landfills	*2:*168			
See also Cost(s)				
Fungal biofouling	*2:*240			
Fungi				
in cooling water systems	*1:*540			
health effects of	*1:*278			
Furrow infiltrometer	*4:*487			
Fuss, Adam	*4:*767			
Fuzzy compromise programming (FCP)	*2:*681			

Fuzzy criteria, for system performance evaluation *2:*674

Fuzzy expected value (FEV) *2:*680 *2:*682

Fuzzy inference system *2:*680

Fuzzy membership function *2:*676

Fuzzy reliability-vulnerability criteria *2:*677

Fuzzy resiliency *2:*677

Fuzzy robustness *2:*677

Fuzzy sets *2:*675

G

GAC bed *1:*103
  See also Granular activated carbon (GAC)

Gage height *2:*543

Gaging station *2:*543

Galerkin weighted residual finite element method *5:*660

Galge site, biomanipulation trials in *2:*55t

Galilei, Galileo *4:*349
  on water music *4:*758
  water observation by *4:*711

GAMIT software *4:*245

Gamma attenuation soil moisture measurement *4:*488

Gamma density log *5:*153

Ganga Basin *3:*232
  See also Ganga-Brahmaputra-Meghna Basin; Ganga River (India); Ganges water dispute
  hydrology of *3:*233
  sediment sampling sites in *3:*297f
  water quality in *3:*235
  water resources development projects in *3:*234

Ganga-Brahmaputra-Meghna Basin *2:*560

Ganga River (India) *3:*232
  adsorption experiments on *3:*296

Ganges water dispute *4:*682
  See also Ganga entries

Gardens
  floating *4:*695
  raised *4:*698

Gardens of the war *4:*697

Garrison Diversion Unit (GDU) *4:*686

Gas abstraction, from landfills *1:*699

Index terms **Links**

Gas chlorination *1:*197
  See also Chlorine entries
Gas chromatography (GC)
  air sampling techniques in *2:*307
  detectors in *2:*307
  headspace technique in *2:*306
  for odor analysis *2:*281
  purge and trap procedure in *2:*306
  sample introduction in *2:*305
  solvent extraction in *2:*307
  in water analysis *2:*304
Gas chromatography/electron capture detection (GC/ECD) *3:*119
Gases. *See also* Greenhouse gases
  dissolved *4:*450
  in domestic sewage *1:*831
Gas exchange coefficient *5:*551
Gas hydrate mounds, observing *4:*60
Gas hydrates *4:*471
  amount in boreholes *4:*61
  distribution and dynamics of *4:*57
  formation of *4:*58
  origin of gas in *4:*58
  remote sensing for *4:*59
  sampling *4:*61
  stability of *4:*58
Gas laws *4:*521
Gas-liquid contactor hydrodynamics *1:*361
Gas migration, landfill *5:*256
Gas-phase activated carbon applications *1:*104
Gas stripping systems *1:*803
Gastric cancer, nitrate-related *1:*36
Gastritis *1:*180
Gastroenteritis *2:*138
Gas-water reactions, isotope exchange in *4:*535
GC/mass spectrometry (GC/MS) *3:*119
Gelatinous zooplankton *4:*155
GEMS/WATER program, objectives of *2:*177
Gene Chips® *2:*60
Gene-probe technology *2:*352
General Agreement on the Crisis of the Aral Sea *3:*18

Index terms	**Links**	
General factorial design	*1:*110	
Generalized autocovariance structure (GAS)	*3:*426	
Generalized likelihood uncertainty estimator (GLUE)	*4:*299	
Generalized Watershed Loading Function (GWLF)	*2:*252	
Generating plants. *See* Electric generating plants		
Genes, luminescent-marked	*2:*453	
GENESIS genetic algorithm optimization software	*2:*327	*2:*334
Genetic algorithm-based techniques	*3:*319	
Genetic algorithms (GAs)	*2:*333	
Genetically modified plants	*3:*686	
Genetic damage	*1:*289	
Genetic engineering	*2:*48	
to advance selenium phytoremediation	*5:*399	
phytoextraction optimization using	*5:*372	
Genomics	*2:*59	*2:*429
Genomic technologies		
in biomonitoring	*2:*58	
high-throughput	*2:*59	
Genomic tools	*2:*339	
in biomonitoring	*2:*63	
Genotoxicity assessment, luminescent bacterial biosensors for	*2:*455	
Genotoxic response	*2:*171	
Geobacter metallireducens	*2:*151	
Geochemical carbon cycle	*4:*414	
Geochemical fixation, of lead	*5:*648	
Geochemical modeling. *See also* Geochemical models		
computer codes for	*5:*140	*5:*142
computer models for	*5:*143	
equilibrium thermodynamics and	*5:*143	
Geochemical models	*5:*138	
See also Geochemical modeling		
components of	*5:*139	
types of	*5:*139	
validation and usefulness of	*5:*139	
Geochemical processes, in groundwater	*5:*32	
Geochemical silica cycle	*4:*550	
Geochemical temperature estimation methods	*4:*93	
Geochemical variations		
seasonal	*5:*684	
spatial	*5:*686	

Index terms **Links**

Geochemistry *5:*463
  of acid mine drainage *3:*13
  of arsenic *1:*81
  cadmium *5:*616
  of carbonate *4:*408
  of carbonate in natural waters *4:*414
  of silica *4:*549
Geoelectrical methods *5:*447
Geoelectrical operations, electrode configurations for *5:*445
Geogenic acidification *3:*9
Geographic Information System (GIS). *See also* GIS technology
  application in water resources *2:*531
  in command area studies *2:*533
  in floodplain management *2:*533
  future need for *2:*535
  ground feature categorization using *3:*363
  in groundwater studies *2:*534
  in hydrologic modeling *2:*535
  in land use classification *2:*532
  management and research uses of *4:*56
  modeling non-point source pollutants using *5:*299
  in precipitation studies *2:*532
  raster-based *3:*28
  in reservoir sedimentation *2:*534
  role in nonpoint source pollution assessment *3:*658
  in snow cover studies *2:*533
  in waterlogging and soil salinity *2:*534
  in water quality studies *2:*534
  in watershed mapping and monitoring *2:*533
Geohydrologic characterization, improvements in *5:*187
Geohydrology
  differential *3:*103
  flow-duration curves and *3:*103
Geological deposits, acidic tailings from *3:*9
Geological fluoride contamination sources *5:*132
Geological formations, resistivities of *5:*444
Geologic deposits, hydraulic conductivity of *5:*508
Geologic settings, setback distances for *1:*72
Geologic soil erosion *3:*417

Index terms **Links**

- Geology
  - effect on trace elements *3:*457
  - Great Lakes *3:*175
- Geomicrobiology *5:*463
- Geomorphology, basin *4:*222
- Geophones® *1:*401
- Geophysical Fluid Dynamics Laboratory (GFDL) *4:*43
- Geophysical groundwater parameters *5:*184
  - field testing and analysis of *5:*185
- Geophysical methods, surface methods *5:*146
- Geophysical survey design *5:*155
- Geophysics
  - groundwater *5:*151
  - remote sensing and *5:*145
- Geopolitics, of South Asia hydro-borders *4:*680
- Geosorbent domains, sorption in *4:*385
- Geosorbents, sorption of hydrophobic organic compounds on *4:*384
- Geostatistics *5:*301
- Geotechnical problems *5:*604
- Geotextiles, natural fiber *3:*568
- Geothermal geophysical methods *5:*151
- Geothermal plants *1:*563
- Geothermal water *5:*156
- ‘‘Germicidal’’ radiation *1:*469
- Geysers *2:*543 *5:*157
- GFDL modular ocean model *4:*44
  - *See also* Geophysical Fluid Dynamics Laboratory (GFDL)
- GHG dynamics, in hydroelectric reservoirs *3:*204
  - *See also* Greenhouse gases (GHGs)
- GHG emission levels, inter- and intra-reservoir *3:*206t
- GHG emission processes *3:*205
- GHG emissions *3:*203 *3:*204
  - from hydroelectric reservoirs *3:*203
- GHG fluxes *3:*204*t* *3:*205
- GHG production processes, in hydroelectric reservoirs *3:*206
- Ghijben-Herzberg equilibrium (GHE) *5:*158
- *Giardia* *1:*279 *1:*522 *1:*523
  - in karstic aquifers *1:*367
- *Giardia lamblia* *1:*189 *1:*257
  - life cycle and morphology of *1:*257

Index terms	**Links**		
Giardiasis	*1:*180	*1:*257	*2:*543
case clusters of	*1:*188t		
clinical disease related to	*1:*257		
diagnosis and treatment of	*1:*258		
epidemiology and prevention of	*1:*258		
Gibbs-Duhem equation	*4:*364		
Gibbs phase equilibrium rule	*4:*264	*4:*364	
Gilgamesh poem, story of water in	*4:*708		
GIMRT/OS3D program	*5:*141		
GIS technology	*2:*252		
to indicate heavy metal concentrations	*2:*448		
Glacial-interglacial cycles	*4:*117		
Glacial lakes	*3:*269		
Glacier Bay, oceanographic environment of	*4:*61		
Glacier ice	*4:*334		
Glaciers	*2:*543	*2:*660	
See also Rock glacier			
Canadian	*2:*660		
Glasses, musical	*4:*758		
Glass pH sensor	*2:*296		
GLEAMS model	*2:*251		
GLERL instrumentation	*4:*70		
Gliding cell formation	*2:*240		
Global average energy balance	*4:*179		
Global change			
detecting	*4:*177		
observations of	*4:*172		
radiative forcing of	*4:*176		
Global climate change	*4:*171		
Global currents	*3:*320		
importance of	*3:*321		
Global cycles, of climatically-active trace gases	*4:*85		
Global desalination market	*1:*308		
Global economy. *See also* Economics			
pharmaceutical industry in	*1:*373		
precautionary principle and	*2:*596		
Global Environment Facility (GEF)	*4:*643		
Global Horizontal Sounding Technique (GHOST)	*4:*166		
Global hydrologic cycle	*4:*266		
Global impacts, of sea-level change	*2:*483		

Globalization
- of water *2:*536
- water rights and *4:*770

Global Meteoric Water Line (GMWL) *4:*439 *5:*230
Global partitioning, of anthropogenic CO_2 *4:*85
Global Positioning System (GPS), refractivity measurements using *4:*244
- *See also* GPS entries

Global Positioning System Integrated Precipitable Water (GPS-IPW) data, effect on numerical weather prediction accuracy *4:*248
Global Positioning System Integrated Precipitable Water Demonstration Network *4:*249
- expansion and improvement of *4:*251

Global Sea Level Observing System (GLOSS) *4:*119
Global search methods *2:*333
Global warming *4:*173 *4:*184 *4:*479 *4:*480
Global water
- balance *4:*225 *4:*279
- conclusions about *4:*289
- crisis *4:*769

Global water availability, prediction of *1:*437
Global water cycle *4:*242
- *See also* Hydrologic cycle; Water cycle
- deuterium in *4:*439

Global water market *1:*3
Global water resources, efficient use of *2:*536
Global water use *2:*489
Global weather *4:*361
Glucose-glutamic acid (GGA) solution, in BOD_5 tests *2:*39
Glutathione (GSH) *3:*609
- role in plant protection *3:*610

Glycerides *2:*442
GOD vulnerability assessment *5:*563
Goff-Gratch function *4:*366
Gold Book *2:*315
Golden Horn Bight, pollution in *2:*446
Golet and Larson wetland classification system *3:*497
Goodness-of-fit statistics *4:*298
Gordon River *4:*788
Governance *1:*42

Index terms **Links**

Government. *See also* British entries; European Union (EU); Extraterritorial land use control; Federal entries; French water management model; Ireland; Mexico; National entries; New Mexico; New York entries; Poland; Public entries; Regulation; Regulations; Regulatory entries; San Diego; South Africa; Sri Lanka; State entries; Syracuse; United States; U.S. entries
  drought policies of *2:*576 *2:*577 *2:*580 *2:*583
  market operation and policies of *2:*500
  public interest water management and *2:*609
  structure of American *2:*498
  water security and *1:*434
Governmental obligations, water rights and *4:*771
Governmental water resource organizations *4:*649
GPS-IPW Demonstration Network sites, data and products from *4:*250
  See also Global Positioning System Integrated Precipitable Water entries
GPS Meteorology (GPS-Met) *4:*249
  See also Global Positioning System (GPS)
GPS orbit predictions *4:*248
GPS signal delays, calculating integrated precipitable water vapor from *4:*245
''Grab'' groundwater sampling *5:*454
Gradients *3:*309
Gradient-type flux law *3:*250
Grain size/porosity *5:*535
Granular activated carbon (GAC) *1:*92 *1:*102 *1:*231 *1:*244 *1:*380 *1:*460 *1:*575 *2:*80
  See also GAC bed
  characterization of *1:*96
  effects of treatment with *2:*76
  pecan shell-based *1:*110
  raw materials for *1:*93
  technology with *1:*350
Granular activated carbon unit *1:*52
Granular aquifers, heterogeneity in *5:*175
Granular bed filtration *1:*231 *1:*249
Granular biological activated carbon (GBAC), oil- field brine and *2:*288
Granular filtration, high-rate *1:*250 *1:*487
  See also Granular bed filtration; Granular media (medium) filtration
Granular filtration technologies, innovations in *1:*235
Granular media (medium) filtration *1:*230 *1:*233 *1:*244 *1:*827
  filter classification *1:*233
  filter media types *1:*233
  filter operation and control *1:*235

Index terms **Links**

Granular media (medium) filtration *(Continued)*
  filtration hydraulics *1:*233
  particle removal mechanisms *1:*233
Granulation *1:*906
Granulometric composition, of sludge *1:*862
Graphical method, for determining base flow recession constants *3:*24
Grass groundcover, as a living mulch *3:*708
Grassland Reserve Program (GRP) *3:*596
Grass plants, root growth in *3:*706
Gravimeter surveys *5:*148
Gravimetric method *5:*533
Gravimetric soil moisture measurement *4:*487
Gravitational geophysical methods *5:*148
Gravitational potential *5:*532
Gravity-flow units
  advantages and disadvantages of *2:*155
  for iron removal *2:*154
Gravity irrigation methods *2:*495
Gravity screen filters *3:*749
Gravity separation/sedimentation *1:*259 *1:*811
Gravity sewers *1:*887
  point repairs for *1:*888
Gravity thickening, of sludge *1:*854 *1:*864
Gray water *3:*557
  See also Greywater
  pollutant concentration in *1:*17t
  quality of *1:*16t
  recycling system for *1:*18f
  removing pollutants from *1:*18f
  reuse in households *1:*16
Grazing, microorganism strategies against *3:*312
Grease, in urban stormwater runoff *3:*435
Great Artesian Basin *5:*30
Great Britain. *See* British entries; U.K. National Rivers Authority
Great Depression *2:*511
  development of *2:*512
Great Egg Harbor River, flow-duration curve for *3:*103 *3:*104f
Great Lakes *2:*658 *3:*175
  geologic origins of *3:*175
  hydrologic flows and climate of *3:*177

Great Lakes *(Continued)*
  land and water use related to *3:*178
  physical features of *3:*176t
  preventing diversion from *4:*618
  water budgets for *3:*177t
  water quality guidance for *4:*622
  water resource management related to *3:*179
  water withdrawal from *3:*179t
  wetlands *3:*71
Great Lakes Charter *4:*618
Great Lakes Charter Annex *4:*619
Great Lakes Commission (GLC) *4:*618
Great Lakes Critical Programs Act *4:*597
Great Lakes Environmental Research Laboratory (GLERL) *3:*274
  meteorological instrumentation at *4:*72
Great Lakes Governors' Agreement *4:*617
Great Lakes Toxic Reduction Effort (GLTRE) *4:*625
Great Lakes Water Quality Agreement (GLWQA) *3:*179 *4:*618 *4:*621 *4:*622
Great Lakes Water Quality Initiative *4:*621
  history of *4:*622
Greece, water symbolism in *4:*786
Greek water clocks *4:*706
Greek water consciousness *4:*708
Green accounting *2:*626
Green algae, filamentous *3:*115
Green Book *2:*315
Green/brown roofs (sloped or flat) *1:*55
Greenfield contracts *1:*389
Greenhouse gas emissions, from hydroelectric reservoirs *3:*180
Greenhouse gases (GHGs) *4:*171
  See also GHG entries
  anthropogenic sources of *3:*203
  state of research on *3:*180
  temperature increase and *4:*282
Greenhouse warming models *4:*22
Greenland ice sheet *4:*118
Green Revolution *2:*617
Greensand filtration, for radium removal *1:*398
''Green'' technology *1:*535 *1:*536
  See also Environmental entries

Grotthuss proton transfer mechanism *4:*483
Ground-based GPS meteorology, at Forecast Systems Laboratory *4:*244
Greywater, defined *2:*543
  See also Gray water
Grid spacing *3:*257
Grid-square basin slope method *3:*29
Grit removal *1:*814
  from domestic sewage *1:*832
Ground-level pumped storage tanks *3:*316
Ground penetrating radar (GPR) *5:*150 *5:*402 *5:*541
  for pipeline assessment *1:*884
Groundwater *1:*444 *2:*95 *4:*285
  See also Contaminated groundwater; Groundwater assessment; Groundwater quality; Modern groundwaters
  acidification of *3:*10
  active management and regulation of *4:*632
  in agriculture *2:*156 *5:*205
  in Arab states *2:*471
  availability of *5:*192
  bacteriological analysis of *5:*202
  benzene in *5:*626
  beryllium in *4:*396
  biochemical parameters of *5:*184
  biotreatment of *5:*120
  cadmium in *5:*613
  Canadian *2:*660 *5:*601
  chemical analyses of *5:*449
  cobalt in *5:*610
  composition of *4:*285
  compounds in *2:*316
  contaminant vapor transport in *5:*551
  defined *2:*543 *5:*600
  dependence on *5:*451
  dissolved gases in *4:*450
  downward vapor transport into *5:*551
  effect of gas on *5:*256
  engineering and *5:*604
  EU monitoring of *2:*268
  evapotranspiration from *5:*168
  flow velocity studying *3:*96
  fluoride contamination in *5:*130

Groundwater *(Continued)*
  fluoride content in *5:*199
  free-flowing *5:*29
  geological occurrence of *5:*145
  geophysics of *5:*151 *5:*184
  harvesting of *2:*550
  as a hidden treasure *5:*599
  high pH *5:*362
  humic matter in *2:*209
  hydraulic conductivity of *5:*128
  hydraulic properties of *5:*184
  hydrochemical data on *2:*158t
  in the hydrologic cycle *5:*601
  impact of river water quality on *2:*396
  in India *2:*555
  injection of *5:*178
  ionic strength of *5:*351
  iron-oxidizing bacteria in *2:*150
  isotopic composition of *4:*576
  lead in *5:*645
  mercury in *5:*642
  metal organic interactions in *5:*258
  methane in *5:*293
  mobility of humic substances in *5:*188
  ‘‘mounding’’ of *5:*353
  movement of *5:*600
  MtBE in *5:*318 *5:*319
  nitrate contamination/content in *2:*156 *2:*570 *5:*322 *5:*628
  nitrate losses to *3:*642
  nitrate pollution of *3:*637
  as a nitrate source *1:*33
  nitrogen in *5:*333
  organic compounds in *5:*337
  particulate arsenic in *2:*11
  particulate transport in *5:*349
  perched *5:*352
  perchlorate in *5:*631
  permafrost and *5:*601
  permeability of *5:*128
  pesticide contamination of *5:*333
  pesticide vulnerability of *5:*357

Index terms	**Links**	
Groundwater *(Continued)*		
Polydex disinfection of	*1:*384	
problems with	*5:*333	
pumped	*3:*583	
reactive solute transport in	*5:*524	
residence time	*4:*388	*5:*601
skimmed	*3:*691	
sodium in	*4:*551	
solute transport modeling in	*5:*305	
as a source of energy	*5:*601	
specific yield of	*5:*129	
sulfate content of	*5:*197	
technetium in	*4:*580	
trace element contamination in	*2:*143	
transmissivity of	*5:*129	
treatment for nitrates in	*5:*323	
unconfined	*5:*662	
uranium in	*5:*640	
use in irrigation	*3:*588	
velocities	*5:*554	
vinyl chloride in	*5:*634	
wetlands and	*3:*173	*5:*605
Groundwater age	*4:*388	
See also Groundwater dating		
estimating	*5:*70	
Groundwater arsenic	*5:*17	
chemical character of	*5:*17	
treatment systems for	*5:*19	
Groundwater assessment	*5:*313	
experimental methodology for	*2:*144	
recommendations following	*2:*148	
results of	*2:*144	
using soil sampling techniques	*3:*688	
Groundwater balance	*5:*162	
equation	*5:*164	
estimation of	*5:*163	
estimation of components for	*5:*164	
Groundwater balance study, data requirements for	*5:*164	
Groundwater bioremediation	*2:*44	
pump and treat	*5:*45	
Groundwater cleanup, objectives for	*3:*652	

Index terms **Links**

Groundwater contaminants. *See also* Groundwater contamination
  defined *5:*30
  movement of *5:*426
  sources of *3:*226
Groundwater contamination *1:*438 *1:*524 *1:*556 *2:*183 *5:*603
  See also Contaminated groundwater
  assessing and defining *5:*549
  cost allocation for *3:*169
  environmental isotopes in *5:*221
  leachate releases and *5:*255
  from runoff *5:*451
  sources of *5:*216 *5:*516
  from urban runoff *5:*452
Groundwater dams *2:*551
Groundwater dating. *See also* Groundwater age
  with H-He *5:*65
  with radiocarbon *5:*64
  with tritium *5:*69
Groundwater depth, crop tolerance and *3:*601
Groundwater devolatilization, upward migration of vapors from *5:*553
Groundwater discharge, from karst systems *5:*238
Ground Water Disinfection Rule (GWDR) *1:*70 *1:*354
Groundwater disputes *4:*632
Groundwater divide *5:*419
Groundwater draft *5:*164 *5:*168
Groundwater dye tracing *3:*96
  See also Groundwater tracing
  analytical strategies for *5:*109
  in karst *5:*107
  methods of *5:*109
  selecting dye introduction points for *5:*109
  at waste disposal sites *5:*110
Groundwater Festival, Children's *4:*713
Groundwater flow *3:*96 *3:*453 *5:*29 *5:*514
  See also Base flow
  aquifers in *5:*515
  effects of recharge and discharge on *5:*83
  in heterogenetic sediments and fractured rock systems *5:*175
  mathematical model of *5:*516
  properties of *5:*128

Index terms **Links**

- Groundwater flow *(Continued)*
  - system *5:*417
  - in unit hydrograph models *4:*359
- Groundwater flow equations *5:*12 *5:*13
- Groundwater flow models, commercial *5:*660
- Groundwater Foundation, The (TGF) *2:*518 *2:*519 *4:*713
- Groundwater geochemistry *2:*19
  - of river-connected aquifers *5:*684
- Groundwater Guardian *2:*518 *2:*520
- Groundwater impacts, bioreactor landfill design and *5:*257
- Groundwater inflow (I_g) *5:*167
- Groundwater investigations, soil vapor data and *5:*548
- Groundwater iron, environmental impact of *5:*608
- Groundwater law *4:*627
  - *See also* Groundwater disputes
  - legal classifications of groundwater *4:*630
  - legal rules for groundwater use *4:*630
  - origin and nature of property rights in *4:*629
  - reform of *4:*633
- Groundwater modeling *5:*619
  - *See also* Groundwater models
  - boundary and initial conditions in *5:*620
  - contaminant transport and biodegradation in *5:*30
  - differential equations in *5:*620
  - discretization in *5:*621
  - parameter estimation in *5:*624
  - results and visualization of *5:*623
  - steps in *5:*623
- Groundwater models *5:*77
  - *See also* Groundwater modeling
- Groundwater monitoring *5:*110
  - for municipal solid waste landfills *2:*165
  - Phase IV *5:*438
- Groundwater monitor wells *5:*404
- Groundwater outflow (O_g) *5:*167
- Groundwater overdraft, in the Arab World *2:*472f
- Groundwater pollutants *3:*226
  - from animal farming operations *3:*538
- Groundwater pollution *2:*400 *3:*282
  - prevention and remediation of *5:*187

Index terms **Links**

Groundwater pollution studies
  isotopes used in *5:*219t
  resistivity methods in *5:*447
Groundwater pollution vulnerability characterization, isotope technique for *5:*219t
Groundwater protection *5:*599
  public education and community involvement in *2:*518
Groundwater protection area *1:*524
Groundwater purging/surging, low-flow *5:*404
Groundwater quality *2:*160 *2:*182 *5:*406 *5:*602
  in animal farming operations *3:*538
  common problems with *5:*406
  in Hardwar, Uttaranchal, India *5:*192
  land-use impacts on *2:*183t *5:*250
  near River Yamuna, India *2:*392
Groundwater recession plots *3:*26
Groundwater recharge *1:*818 *2:*393 *3:*603
  to karst systems *5:*236
  in karst topography *5:*247
  projects *2:*611
  regulation of water quality for *1:*291
Groundwater recovery, goals of *4:*675
Groundwater remediation. *See also In situ* groundwater bioremediation; *In situ* groundwater remediation
  activated carbon in *1:*105
  chemical oxidation technologies for *5:*344
  by electrochemical processes *5:*283
  Fenton's Reaction and *4:*445
  horizontal wells in *5:*178
  by injection and problem prevention *5:*421
  innovative technologies for *5:*438
  landfill-related *5:*257
  by no action (passive) *5:*44 *5:*423
  organoclays in *1:*772
  project life cycle *5:*436
Groundwater resources
  in India *2:*562
  quantification of *5:*163
  use of *5:*187
  vulnerability mapping of *5:*561
Ground Water Rule (GWR) *2:*195 *2:*345

Index terms **Links**

Groundwater sampling
  for CFCs *4:*421
  diffusion-based *5:*456
  for environmental projects *5:*454
  with passive diffusion samplers *5:*456
  protocol for *5:*455
  purpose of *5:*454
  quality control of *5:*455
Groundwater scheme, environmental impact assessment for *2:*521t
Groundwater sensitivity
  to contamination *5:*56
  factors affecting *5:*58t
Groundwater sources, high control of *2:*400
Groundwater storage
  change in (Δ*S*) *5:*168
  in karst systems *5:*238
Groundwater storage coefficient *5:*129
Groundwater studies
  CFCs in *4:*422
  microbiological methods in *5:*464
  remote sensing and GIS application in *2:*534
Groundwater supply, safeguarding *5:*605
Groundwater systems
  features of *5:*308
  fluid dynamic analysis of *5:*498
  role of heat in *5:*172
  subsurface heterogeneity characterization in *5:*308
Groundwater tracing *5:*501
  See also Groundwater dye tracing
  experiment *5:*506
  history and goals of *5:*502
  ideal tracer characteristics *5:*502
  through karst systems *5:*240
  tracer types *5:*502
Ground-Water Treatment Rule *1:*72
Groundwater transport
  of bacteria *5:*350
  in karst systems *5:*237
  processes *5:*516
Groundwater University (GU) *4:*714

Index terms		Links
Groundwater vulnerability	*5*:561	
assessing	*5*:562	
characterization of	*5*:563	
factors in	*5*:562	
models	*5*:358	
to pesticides	*5*:594	
Groundwater vulnerability assessment (VA)	*1*:527	
Groundwater withdrawal, stages of	*5*:576	
Group of Experts on the Scientific Aspects of Marine Pollution (GESAMP)	*4*:109	
Grouting	*1*:881	
Growing areas, shellfish	*2*:361	
Growth-linked biodegradation	*2*:42	
GSTAR computer models	*3*:389	
Guided random search method	*2*:333	
Guinea worm	*1*:26	
Gulf of Mexico, gaseous hydrocarbon seepage in	*4*:99	
Gulf Stream, discovery of	*4*:780	
Gullies		
characterization of	*3*:183	
continuous and discontinuous	*3*:184	
initiation of	*3*:187	
sediment yields from	*3*:184	
Gully erosion	*3*:183	
causes of	*3*:185	
effects of	*3*:187	
processes associated with	*3*:184	
Gully formation	*3*:87	
deductive model of	*3*:185	
Gully plugs	*2*:550	
Gully pot effluent treatment	*3*:499	
Gunite	*1*:802	*1*:880
Gutters	*1*:56	
design methods for	*1*:57	
flow division within	*1*:57	
outlet depths for	*1*:56	
GWLF model	*5*:298	

Index terms	Links	
H		
H_2 economy	*4:*477	
See also Hydrogen entries		
Habitat(s)		
alterations of	*3:*126	
coral reef	*4:*114	
deepwater	*3:*496	
microorganism	*3:*310	
research on	*4:*55	
as a watershed function	*3:*477	
Habitat conservation, precautionary principle regarding	*2:*600	
See also Environmental entries; Habitat preservation		
Habitat preservation, fishery-related	*3:*128	
See also Habitat conservation		
Habitat restoration/enhancement, water reuse for	*1:*818	
Habitats methods, for determining environmental water requirements	*3:*167	
Hack's law	*3:*94	
Hagen-Poiselle formula	*3:*195	
Haloacetaldehydes, health effects of	*1:*268	
Haloacetic acids (HAAs)	*2:*91	
bromide influence on formation of	*2:*74	
brominated	*2:*75	
health effects of	*1:*264	*1:*266
Haloacetonitriles, health effects of	*1:*267	
Halocarbons, marine sources of	*4:*149	
Halogenated byproducts	*1:*199	
Halogenated methanes, measurements of	*4:*149	
Halogenated organic chemicals (HOC), phytoremediation research on	*3:*368	
Haloketones, health effects of	*1:*268	
Halomethanes, health effects of	*1:*264	
Halophytes	*3:*686	
Hamoon Wetlands hydraulic-hydrologic system	*3:*563	
Hantush and Jacob method	*5:*494	
Harbor surveys, New York City	*1:*745	
Hardness	*5:*406	
of activated carbons	*1:*96	
defined	*2:*543	
of drinking water	*2:*394	
forms of	*4:*553	
role in fish growth and production	*3:*131	

Index terms	**Links**			
Hardness salts	*1:*545			
removal of	*1:*548			
Hardwar, Uttaranchal, India				
groundwater contamination in	*5:*202t			
groundwater quality in	*5:*192			
groundwater sampling locations in	*5:*206t			
hydrochemical data from	*5:*195t			
irrigation water quality in	*5:*204			
Hardware development, Marine Instrumentation Laboratory (MIL)	*4:*70			
Hard water	*1:*144	*1:*322	*1:*534	*1:*902
	*4:*452			
definition and specifications of	*4:*453			
forms of	*4:*453			
softening of	*4:*454			
worldwide cases of	*4:*454			
Hargreaves method	*3:*575			
Hartree-Fock (HF) method	*4:*401			
Harvesting, nitrogen losses from	*3:*696			
Harvest regulation	*3:*124			
in river fisheries	*3:*125			
Harvest restrictions, in fisheries	*3:*129			
Hatchery animals, success of	*4:*126			
Hathorn Spring	*4:*799			
Hawaii. *See also* Hawaiian Islands				
climate and water balance in	*4:*255			
Koppen climatic zones in	*4:*256			
soil moisture storage capacity for	*4:*261			
Hawaiian freshwater resources, history of	*4:*801			
Hawaiian Islands. *See also* Hawaii				
amphidromous fish species in	*4:*803			
climate patterns of	*4:*802			
ecosystems and endemism in	*4:*802			
endemic stream animals in	*4:*803			
freshwater resource use and management in	*4:*804			
geological origin of	*4:*802			
stream animal communities in	*4:*803			
surface water removal in	*4:*804			
water law in	*4:*805			
Hawaiian stream habitats, degradation of	*4:*806			
Hawaiian streams, freshwater animals of	*4:*804t			
Hawk Mountain Sanctuary Association	*3:*518	*3:*519		

Index terms **Links**

Hazard assessment critical control point (HACCP) risk assessment approach *1:*425f
Hazardous air pollutants (HAPs) *1:*555
Hazardous chemical bioaccumulation, in urban stormwater runoff *3:*435
Hazardous solid wastes, solidification/ stabilization of *1:*835
Hazardous substance liability *4:*665
Hazardous waste management, bioremediation in *5:*116
Hazardous waste sites, monitoring *5:*587
Hazen-William roughness factors *3*:314 *3*:315
Hazen-Williams formula *3:*197
  values of C for *3:*198t
H (basic) carbons *1:*98
Head *5:*180
  See also Hydraulic head
  capillary and osmotic forces and *5:*180
  fluid density and *5:*181
  measurement *5:*474
  spatial and temporal variation of *5:*181
Head loss, in diffused air aeration systems *1:*629
Headspace technique *2:*306
Headwater *2:*543
^{4}He age dating method *4:*390
Healing, water-instrument *4:*760
Health. *See also* Human health
  ecological *2:*182
  impact of water on *4*:455 *4:*722 *4:*729
  public water supply and *1:*501
  soft water and *4:*554
  waterborne pathogens and *1:*521
Health Advisories *1:*479
Health Assessment on Perchlorate *2:*345
Healthcare reforms *1:*190
Health effects
  of arsenic *2:*15
  of bromide *2:*74
  of chlorine and chlorination byproducts *2:*92
  of cobalt *5:*612
  of disinfection byproducts *1:*264
  of lead *2:*432
  of microbial contaminants and biotoxins in drinking water *1:*277
  of nitrate *1:*30

Index terms	**Links**			
Health effects *(Continued)*				
of polychlorinated biphenyls	*2:*107			
of uranium	*5:*641			
of vinyl chloride	*5:*636			
of water pollution	*2:*617			
Health hazards	*1:*293			
See also Health risks; Hygienic hazards				
Health issues. *See also* Human health				
algae-related	*3:*188			
iron-related	*4:*497			
Health precautions, for chemical oxidants	*5:*348			
Health risk reduction and cost analysis (HRRCA)	*1:*425	*1:*479	*4:*541	*4:*546
for radon	*4:*546			
Health risks. *See also* Health hazards; Health risk reduction and cost analysis (HRRCA)				
of cadmium	*5:*615			
of heavy metals	*5:*279t			
of lead	*5:*646			
of mercury	*5:*643			
of MtBE	*5:*318			
of nitrate	*5:*629			
of perchlorate	*5:*632			
of radionuclides	*1:*395			
of radon	*4:*543	*4:*544		
reduction of	*1:*427			
of wastewater reclamation	*1:*826			
Heat				
effect on physical properties and reaction kinetics	*5:*173			
enhancement	*5:*117			
role in groundwater systems	*5:*172			
Heat balance				
formulation of	*3:*191			
meteorologic driving factors in	*3:*191			
of open waterbodies	*3:*190			
Heat dissipation soil moisture measurement	*4:*488			
Heated waters, discharge of	*4:*105			
Heat flux densities, sensible and latent	*3:*563			
Heat fluxes				
of long-wave radiation	*3:*192			
relationship to meteorologic fields	*3:*191			
Heat of vaporization	*4:*263			

Index terms	**Links**		
Heat transfer/equipment problems, water treatment and	*1:*599		
Heavy metal anions, biosorption of	*2:*72		
Heavy metal contamination			
in situ groundwater remediation for	*5:*290		
prevention of	*5:*278		
Heavy metal efflux, landfill disposal practices and	*1:*726		
Heavy metal ion removal, by ion exchangers	*4:*494t		
Heavy metal pollution	*2:*407		
in Far Eastern Russia seas	*2:*447		
Heavy metal removal technologies, performance characteristics of	*2:*69t		
Heavy metals	*4:*506		
See also Iron entries; Manganese entries; Mercury removal; Metal entries			
in an agricultural landscape	*3:*607		
in constructed wetlands	*1:*894		
disease and	*3:*295		
environmental sources of	*5:*276		
health risks associated with	*5:*279t		
immobilization processes for	*5:*378		
in Indian rivers	*3:*449		
management and remediation alternatives for	*5:*278		
microbial biosurvey detection of	*2:*442	*2:*445	*2:*446
mobility of	*5:*277		
phytoremediation of	*5:*376		
precipitation/crystallization/mineralization of	*5:*378		
regulation of	*4:*598		
removal of	*3:*368		
in river systems	*3:*295		
risk assessment of	*5:*278		
soil and water contamination by	*5:*275		
sorption of	*1:*111		
toxicity	*3:*609		
transfer of	*3:*608		
uptake and intercellular sequestration of	*5:*378		
uptake rates of	*2:*211		
in urban stormwater runoff	*3:*433		
as water pollution	*4:*99		
as water quality indicators	*2:*266		
Heavy metal sorbents	*1:*586		
Heavy metal toxicity detection, microbial enzyme assays for	*2:*233		
Heavy oil sorption	*2:*363		

Index terms	**Links**		
Heavy water	*4:*462		
See also Deuterium			
applications of	*4:*463		
behavior of	*4:*463		
biological and biotechnological applications of	*4:*464		
characteristics of	*4:*462		
production of	*4:*464		
worldwide production of	*4:*465		
HEC-HMS hydrologic model	*3:*157	*3:*158	
Hedonic price model	*4:*609		
Helley-Smith pressure-difference bed load sampler	*3:*400		
Helminths	*1:*904		
in domestic sewage	*1:*831		
Hematopoietic effects of lead	*2:*437		
Hemolytic uremic syndrome (HUS)	*1:*181	*2:*136	
Hemorrhagic colitis (HC)	*1:*180	*2:*136	
Hemorrhagic *Escherichia coli* outbreak	*1:*341		
Henry, William	*4:*466		
Henry’s law	*1:*576	*1:*577	*4:*466
limitations and application of	*4:*467		
Henry’s law constants	*1:*761	*4:*522	*5:*428
Hepatitis	*1:*181	*1:*341	
Hepatitis virus	*1:*279		
Hepatotoxins	*2:*388	*3:*189	
Heptachlor epoxide	*3:*352	*3:*649	
Herbicides	*3:*482		
agricultural productivity and	*4:*507		
in algal control	*2:*5		
alternative	*4:*506		
commonly observed	*5:*252		
tile drainage and	*3:*729		
weed-control	*3:*745t		
Herbivorous fish	*3:*746		
Heterogeneity	*5:*509		
Heterogeneous freezing	*4:*585		
Heterogeneous photocatalysis	*1:*791		
Heterogenetic sediments, groundwater flow in	*5:*175		
Heterotrophic bacteria	*2:*21	*2:*175	
Heterotrophic microorganisms/microbes	*2:*38	*2:*448t	*3:*311
Heterotrophic nitrification	*3:*641		

Index terms	**Links**	
Heterotrophic plate count bacteria	*1:*222	
Heterotrophic plate counts (HPCs)	*1:*86	
Hexachlorobenzene (HCB)	*3:*352	
Hexavalent chromium	*4:*515	
samples	*2:*309	
removal of	*4:*587	
Hg (II) ion sorption	*4:*494	
H-He dating	*5:*65	
See also Tritium/helium ratios		
analytical measurement methods for	*5:*68	
principles of	*5:*66	
versus CFC dating	*5:*68	
versus ^{85}Kr dating	*5:*68	*5:*249
Hidden water sources, in the home	*2:*527	
Higee aeration	*1:*353	
High-density polyethylene (HDPE) landfill liners	*2:*164	
High density sludge (HDS) process	*1:*610	
Higher-order degradation behavior	*5:*32	
Higher plants		
metal tolerance and accumulation in	*5:*284	
trace elements in	*3:*456	
Higher water retention	*1:*406	
High level waste, versus average rock	*5:*450	
Highly impacted soil zone (HISZ)	*3:*688	
High-performance liquid chromatography (HPLC)	*2:*308	*2:*389
High pH groundwater	*5:*362	
High-precision deuterium analysis	*4:*439	
High rate anaerobic treatment systems	*1:*908t	
High-rate granular filtration	*1:*250	*1:*487
High rate treatment systems	*1:*841	*1:*842t
High-resolution sea floor surveys	*4:*77	
High-salt-content water sources, microirrigation and	*3:*618	
High seas fisheries management	*3:*126	
High-throughput technologies, biomonitoring and	*2:*59	
High volcanic islands, perched groundwater in	*5:*353	
High water table, ill effects of	*3:*601	
Himalayas	*2:*559	
river systems in	*2:*560	
Hin primers	*2:*338	

Index terms **Links**

Historic Area Remediation Site (HARS), New York City *4:*77
  changes in *4:*79
  images of *4:*78
  sea floor of *4:*78
H-1i primers *2:*338
Hoarfrost *4:*241
Ho Chi Minh City
  water environment and pollution in *2:*552
  water resource management in *2:*552
  water supply and distribution in *2:*553
Hofmeister effects *4:*468
  explanation of *4:*469
Holistic management *3:*33
  of freshwater fisheries *3:*134
Holistic methods, for determining environmental water requirements *3:*167
Hollows, agriculture of *4:*695
Hollow-stem auger water well drilling *5:*106
Holometabolous insects *3:*51
Holomixis *3:*266 *3:*267
Holoplankton *4:*155
Home
  hidden water sources in *2:*527
  water conservation in *2:*664
Home aeration methods *1:*51
Homeland Security Act *2:*347
Homeland security initiatives *1:*800
Homeostasis, impact of water on *4:*455
Homogeneous nucleation *5:*569
Hoover Dam *2:*554
  history of *4:*732
Hopper-dredge disposal *2:*123
Horizontal collector wells *5:*88
Horizontal directionally drilled (HDD) wells *5:*408
Horizontal drainage layers, levee *3:*290
Horizontal drains, design criteria for *5:*97
Horizontal drilling *5:*179
Horizontal-flow constructed wetlands *1:*895
Horizontal groundwater flow *5:*511
Horizontal louver screens *5:*572
Horizontal pipe drainage systems *3:*731

Index terms	**Links**		
Horizontal profiling	*5:*446		
Horizontal(-flow) roughing filters (HRFs)	*1:*237	*1:*239	*1:*241
Horizontal subsurface drainage	*5:*95		
Horizontal wells	*5:*177		
advantages of	*5:*180		
drilling	*5:*106		
in groundwater remediation	*5:*178		
Hormones, in animal farming operations	*3:*540		
Horn, Roni	*4:*768		
Horology	*4:*704		
Hortillonnages	*4:*697		
Hortonian runoff	*3:*566		
Horton’s double exponential	*3:*23		
Horton’s laws of drainage composition	*3:*32	*3:*93	*3:*438
Hot springs	*5:*475		
Hourly digital product (HDP) rainfall estimates	*4:*310		
Household drinking water, treatment and safe storage of	*1:*67		
Households, gray water reuse in	*1:*16		
Household wastewater, characteristics of	*1:*677t		
Household water consumption, developed countries	*1:*506f		
Household water efficiency	*2:*635		
Household water meters	*1:*490		
HSPFEXP system	*5:*297		
HSPF model	*5:*297	*5:*298	*5:*332
Huayuankou			
composition of floods in	*3:*47t		
flood frequency analysis at	*3:*46t		
Hubenov Reservoir, biomanipulation trials in	*2:*53t		
Hudson’s formula	*4:*15	*4:*16	
Huelva coast, storm record on	*4:*339		
Human activities			
contamination from	*2:*489		
effect on trace elements	*3:*457		
impacts on groundwater quality	*5:*250		
influence on the hydrologic cycle	*4:*280		
most water demanding	*2:*489		
pollution caused by	*3:*223		
role in river water chemistry	*4:*374		
sedimentation and	*4:*107		
as a threat to corals	*4:*33		

Human activities *(Continued)*
  in the water cycle *4:*244
Human body, water shortage in *4:*789
  See also Body
Human disturbance, quantifying *3:*39
Human exposure, to reclaimed irrigation *3:*670
Human/fish relationship, history of *3:*122
Human health. *See also* Health entries
  basic water requirements for *1:*22
  nitrogen loading and *3:*696
  role of water scarcity and stress in *1:*20
  water and *1:*19
Human infections, *Escherichia coli* shedding patterns in *2:*137
Human interference, runoff pollution caused by *3:*225
Human rights, water-related *4:*769
Humans
  effect on the hydrosphere *4:*285
  effects of cadmium on *5:*615
  marine debris hazard to *4:*40
  toxicity of lead to *2:*435
  vulnerability to pollutant discharge *3:*284
Human safety, weed control and *3:*744
Human waste disposal *1:*27
Human water consumption, categories of *2:*370 *2:*614
Humic acid(s) *1:*794 *4:*399 *5:*188
  isolation and purification of *5:*189
  origin, stability, and mobility of *5:*190
  properties of *2:*205 *5:*189 *5:*190
Humic colloids *2:*105
Humic/fulvic acid
  complexation to *2:*207
  inorganic colloids stabilized by *2:*208
Humic substances (HS) *4:*399 *5:*337
  concentrations of *4:*450
  mobility in groundwater *5:*188
Humidity
  absolute *4:*269
  relative *4:*270
Humidity indexes, alternative *4:*273
Humidity Sounder for Brazil (HSB) *4:*216 *4:*351

Index terms **Links**

Humidity statistics, Chicago Station 5:79t
Humid tropical climatic zone 4:256
Humin 4:399
Hurricane Emily 4:274*f* 4:275f
  forecasting 4:275
Hurricane Research Division (NOAA) 4:5
Hurricanes
  estimating maximum significant wave height during 4:2
  estimating maximum sustained wind speed during 4:2
  modeling of 4:274
  theory of 4:353
Hurst coefficient 3:212
Hurst phenomenon 3:210 3:261
  importance of 3:215
  physical explanations of 3:212
  stochastic representation of 3:211
  time series for reproducing 3:217
Hyallela azteca 3:51 3:52
Hydrant flow data collection table, sample 3:317t
Hydrant flow tests 3:316
Hydrant-flushing program 1:401
Hydrants, water distribution system 1:208
Hydrated lime 5:4
Hydrates, clathrate 4:471
Hydration 4:475
  of hydrophobic molecules 4:475
  of ions 4:476
Hydraulic age 4:389
Hydraulic analysis 3:431
Hydraulic calculations, in culvert design 3:77
  See also Hydraulic methods
Hydraulic conductivity (*K*) 3:172 4:485 5:64 5:184 5:492 5:526
  conversion factors 5:509t
  of groundwater 5:128
  unsaturated 5:126
  in the vadose zone 5:537
Hydraulic conductivity/transmissibility 5:507
  Darcy's law and 5:507
  effective 5:510

Index terms	**Links**	
Hydraulic conductivity/transmissibility *(Continued)*		
estimation and measurement methods for	*5:*512	
geologic factors influencing	*5:*508	
permeability and	*5:*508	
representative values of	*5:*508	
units of	*5:*508	
variability in	*5:*508	
Hydraulic conductivity values, statistical treatment of	*5:*510	
Hydraulic construction, Caspian Sea level rise and	*2:*482	
Hydraulic depth	*3:*347	
Hydraulic design, of water distribution storage tanks	*1:*448	
Hydraulic dredging	*2:*122	
Hydraulic flocculation methods	*1:*253	
Hydraulic gradeline (HGL) values	*1:*210	
Hydraulic gradient	*5:*171	
Hydraulic habitats	*3:*106	
Hydraulic head	*5:*169	*5:*554
components of	*5:*171	
in a 3-D aquifer	*5:*171	
Hydraulic loading	*1:*347	
Hydraulic methods, for computing time of concentration	*3:*470	
Hydraulic network models	*3:*313	
calibration data sources for	*3:*315	
calibration model parameters for	*3:*314	
data collection for	*3:*316	
evaluating results of	*3:*316	
intended use of	*3:*314	
macrocalibration in	*3:*318	
microcalibration in	*3:*319	
sensitivity analysis in	*3:*319	
Hydraulic potential	*5:*171	
hydraulic head and	*5:*171	
Hydraulic properties characterization	*5:*184	
Hydraulic radius	*3:*64	
Hydraulic resistance	*5:*492	
Hydraulic retention time (HRT)	*1:*753	
Hydraulics	*3:*194	
of pressurized flow	*3:*196	
roof drainage	*1:*54	
well	*5:*182	

Hydraulic shock. *See* Water hammer
"Hydraulic society" *3:*587
Hydraulic turbine *3:*487
Hydraulic works, origin of *4:*738
HYDRO-35 method *4:*315
Hydroborders
  ethnopolitics and *4:*681
  in South Asia *4:*681
Hydrocarbon fuel vapors, leaking *5:*552
Hydrocarbon phytoremediation, selecting plants and microorganisms for *3:*628
Hydrocarbon plumes *5:*44
Hydrocarbon remediation *5:*281
  See also Hydrocarbon phytoremediation
  water-jetting drilling technologies for *5:*234
Hydrocarbons. *See also* Alicyclic hydrocarbons; Aliphatic hydrocarbons; Petroleum hydrocarbons; Polycyclic aromatic hydrocarbons (PAHs)
  adsorption of *1:*577 *5:*46
  classes of *5:*42
  degradation of *1:*692
  destruction of *1:*579
  dissolved and adsorbed *5:*46
  plants that spur degradation of *3:*629
  polycyclic aromatic *1:*571
  separation of *1:*576
  solubility in salt water *4:*559
  solubility in water *4:*561
  treatment techniques for *1:*575
Hydrochemical classification *2:*396 *2:*397f
Hydrochemical models *5:*295
  parameterization of *5:*297
  properties of *5:*296
  used in watershed ecology *5:*297t
Hydroclones *1:*683
Hydrodynamic and Water Quality Model for Streams (CE-QUAL-RIV1) software *2:*328
Hydrodynamic dispersion *5:*33 *5:*70 *5:*273 *5:*517 *5:*525 *5:*526
Hydrodynamic parameters, determining *5:*526
Hydrodynamic processes *5:*525
Hydrodynamic simulation *2:*256

Index terms	**Links**			
Hydroelectric dams				
environmental disruptions related to	*4:*267			
South American	*4:*642			
Hydroelectric decision support system	*2:*621			
Hydroelectricity water use	*1:*561			
Hydroelectric plants	*2:*568			
Hydroelectric potential, Ganga Basin	*3:*234			
Hydroelectric power	*3:*199	*4:*729		
See also Hydropower entries				
advantages and disadvantages of	*3:*200			
dams and	*3:*199			
generation issues related to	*3:*200			
in India	*2:*565			
quantitative facts concerning	*3:*199			
water use for	*2:*543			
Hydroelectric reservoirs	*3:*182			
See also Hydropower reservoirs				
as anthropogenic sources of GHGs	*3:*203			
GHG dynamics in	*3:*204			
GHG emissions from	*3:*180	*3:*203	*3:*205	
GHG production processes in	*3:*206			
Hydroelectric storage plants	*3:*201			
Hydrogen (H)	*4:*189			
See also H_2 economy				
Hydrogen bonds	*4:*406	*4:*511	*4:*583	
Hydrogen ion	*4:*480			
characteristics of	*4:*481			
importance of	*4:*481			
production in water	*4:*481			
Hydrogen ion concentration	*2:*294	*4:*482		
measurement of	*4:*481			
Hydrogen isotopes	*4:*438	*4:*501		
as groundwater tracers	*5:*503			
Hydrogen peroxide	*1:*603	*1:*913	*4:*533	*5:*344
	*5:*345	*5:*348		
bromate ion formation and	*1:*361			
groundwater remediation and	*4:*445			
health and safety precautions for	*5:*348			
Hydrogen sulfide	*3:*341			
gas	*5:*407			
generation	*1:*911			

Hydrogen sulfide *(Continued)*
  modeling *3:*333
Hydrogeologic responses, to earthquakes *5:*113
Hydrogeologic settings, effect on viral movement *1:*71
Hydrogeology. *See also* Contaminant hydrogeology
  effect on natural attenuation *5:*586
  environmental isotopes in *5:*227
  forensic *3:*168
Hydrographic analysis *3:*26
Hydrographic area, as an operational unit of analysis *2:*613
Hydrograph recession selection algorithm
  automatic *3:*25
Hydrographs
  incremental *3:*65
  unit *3:*221
Hydroids, on surfaces *1:*541
Hydrologic activity, rates of *4:*287
Hydrological information system, real-time *1:*121
Hydrological methods, for determining environmental water requirements *3:*167
Hydrological networks *3:*461
Hydrological processes *3:*222 *3:*659
  peculiarities of *3:*425
Hydrological Simulation Program-FORTRAN (HSPF) *2:*249 *2:*326
Hydrological storage *3:*173
Hydrological tracers, dyes as *3:*95
Hydrologic cycle *2:*95 *2:*656 *3:*327 *3:*477 *3:*688 *4:*275 *4:*287 *4:*584 *4:*711
  See also Global water cycle; Water cycle
  atmospheric *4:*181
  changes in *4:*173
  components of *4:*276
  defined *2:*544
  forest structure and *3:*171
  groundwater in *5:*601
  history of *4:*191
  influences on *4:*280
  interception process in *3:*236
  mathematical representation of *4:*277
  as a multi-phased journey *4:*192
  role of aquifers in *5:*9

Index terms	**Links**	
Hydrologic cycle *(Continued)*		
scales for studying	*4:*277	
water reserves in	*4:*282t	
Hydrologic data		
transforming to real-time information	*1:*123	
uncertainty in	*2:*621	
Hydrologic drought	*4:*208	
Hydrologic ensembles	*2:*621	
Hydrologic equation	*4:*279	
Hydrologic feasibility assessment and design		
contaminant distribution characterization	*5:*80	
implications of	*5:*85	
phytocontainment	*5:*76	
phytohydraulic containment and treatment	*5:*81	
in phytoremediation	*5:*76	
vadose zone transport	*5:*77	
Hydrologic flow paths, acidification and	*3:*5	
Hydrologic flows, Great Lakes	*3:*177	
Hydrologic forecasting model	*2:*622	
Hydrologic history	*4:*265	
Hydrologic impacts, of evaporation	*4:*225	
Hydrologic information, availability of	*2:*500	
Hydrologic modeling (HYMO)	*3:*255	*3:*533
See also Hydrologic models		
entropy theory for	*4:*217	
macroscale	*3:*534	
remote sensing and GIS application in	*2:*535	
Hydrologic models, use of field capacity in	*5:*126	
Hydrologic processes, in arid or semiarid lands	*5:*75	
Hydrologic networks, design of	*4:*222	
Hydrologic persistence	*3:*210	
effect of	*3:*263	
Hydrologic problems and perspectives	*4:*267	
Hydrologic sciences, advancements in	*4:*266	
Hydrologic soil groups	*3:*30	
Hydrologic systems		
engineering decisions concerning	*4:*217	
modeling of	*4:*267	
water contamination in	*5:*60	

Index terms	Links		
Hydrologic thresholds	*3:*229		
identification and quantification of	*3:*231		
interdisciplinary research on	*3:*231		
Hydrologic watershed functions	*3:*473		
Hydrology			
Ganga Basin	*3:*233		
isotopic	*4:*440		
use of Mariotte bottle in	*4:*503		
order of discontinuity in	*3:*239		
postmining	*5:*2		
soil pipe	*5:*403		
submodel	*5:*296		
watershed	*3:*472		
Hydrology applications, remote sensing of	*4:*319		
Hydrology research, forest	*2:*199		
Hydrolysis	*1:*518	*5:*528	
calcite solubility and	*4:*412		
in landfills	*1:*696		
in mine effluent remediation	*1:*611		
Hydrometeorological mechanisms, as a cause of flooding	*3:*511		
Hydrometer	*5:*474		
Hydronium ion	*4:*482		
Hydrophilic soil conditioners	*3:*552		
Hydrophobic bacteria	*5:*351		
Hydrophobicity	*1:*761	*4:*475	*4:*523
biodegradability and	*1:*761	*4:*524	
toxicity and	*4:*524		
Hydrophobic molecules, hydration of	*4:*475		
Hydrophobic organic compounds (HOCs)	*4:*384		
sorption on geosorbents	*4:*384		
Hydrophobic organic compound sorption, rate considerations in	*4:*387		
Hydroponic solutions, MtBE uptake and	*5:*389		
Hydropower	*2:*554		
See also Hydroelectric power			
greenhouse gas emissions from	*3:*182f		
major producers of	*3:*200t		
Hydropower potential			
developed	*3:*199t		
estimating	*3:*202		

Index terms	**Links**		
Hydropower projects			
components of	*3:*200		
load of	*3:*202		
types of	*3:*201		
Hydropower reservoirs	*3:*202		
See also Hydroelectric reservoirs			
Hydropsychology	*4:*733		
research in	*4:*735		
Hydrosphere	*4:*283		
biogeochemistry of	*4:*284		
effect of humans on	*4:*285		
Hydrostatic equilibrium, disturbance of	*4:*367		
Hydrostatic forces	*3:*194		
Hydrosystems, stochastic simulation of	*3:*421		
Hydrothermal phenomena	*5:*156		
Hydrous manganese oxide (HMO) filtration	*1:*398		
Hydroxide alkalinity	*4:*410		
calculation of	*4:*410		
Hydroxyl free radical	*4:*446	*4:*447	*4:*533
Hydroxyl radicals	*5:*344		
HYDRUS code	*5:*564		
HYDRUS model	*5:*663		
Hyetograph analysis	*4:*318		
Hygienic hazards, sources of	*1:*677		
See also Health hazards			
Hygrometers	*4:*273		
HYMO flow routing model	*3:*255		
See also Hydrologic modeling (HYMO)			
Hyperaccumulating plants. *See also* Hyperaccumulators			
nutrient requirements of	*5:*379		
phytoextraction using	*5:*369		
Hyperaccumulation, lead	*5:*382		
Hyperaccumulation potential, of aquatic macrophytes	*2:*66		
Hyperaccumulators	*3:*629	*5:*366	*5:*370
See also Hyperaccumulating plants			
metal	*5:*288		
selenium	*5:*398		
Hyperendemic disease	*1:*187		
Hyperfiltration	*1:*631	*1:*810	
Hypertension, water intake and	*4:*725		

Index terms	**Links**		
Hyperthermophilic bacteria	*3:*311		
Hypertonic solution	*4:*459		
Hypochlorites	*1:*457	*2:*88	
Hypolimnetic hypoxia	*4:*65		
Hypothesis testing, entropy theory and	*4:*220		
Hypoxia			
factors affecting	*4:*65		
quantification of	*4:*64		
Hypoxic factor (HF), determining	*4:*65		
Hypsometric curve	*3:*29		
Hysteresis	*5:*536		

I

Ice	*4:*285		
floating	*4:*69		
Ice Ages, remains of	*4:*305		
Ice control, alternatives to NaCl for	*2:*320t		
See also Deicers			
Ice-covered lakes, heat transfer across	*3:*191		
Ice-jam floods	*3:*45	*3:*46	*3:*511
Ice-melt recession	*3:*24		
Idaho			
source-water protection in	*2:*312		
stream flow statute	*4:*661		
Ideal adsorbed solution (IAS)	*1:*115		
Ideal adsorbed solution theory (IAST)	*1:*109	*1:*119	
‘‘Ideal’’ water tracer	*3:*95		
Idso-Jackson crop water stress index (CWSI)	*3:*720		
Igneous rocks	*5:*55		
Ignorance, types of	*4:*300		
Illegal discharges, elimination of	*1:*867		
Illinois, overdraft in	*5:*342		
Immobilization technologies	*5:*434	*5:*439	
Immunodetection methods	*2:*109		
Immunofluorescence assay (IFA)	*1*:159	*1*:160	*1*:165
Immunohistochemistry	*2:*109		
Immunologic assays/testing	*1:*90		
for *Escherichia coli* O157:H7	2:139		
Immunomagnetic separation (IMS)	*1:*159		
Impact reduction, for net pen aquaculture systems	*3:*543		

Index terms **Links**

Impellers *3:*664
  advanced *1:*79f
  characteristics of *1:*80t
  performance data for *1:*80
  types of *1:*76
Impermeable layer *2:*544
Impermeable materials *5:*600
Impervious cover *3:*363
Implicit pressure explicit saturation (IMPES) approach *5:*669
Import of virtual water *2:*536
Impoundments, blue-green algae blooms in *3:*188
Impregnated activated carbon (IAC) *1:*93
Impulse turbines *3:*488
Inactive (dead) storage *3:*259
Inactive rock glaciers *3:*174
In-channel pumped storage development *3:*201
Incidental recharge *5:*75
Incineration, of sludge *1:*858 *1:*866
Inclined drainage layers *3:*290
Inclined plates *1:*454
Incompressible flow problems, finite element schemes for *5:*657
Incompressible fluids *5:*559
Index-and-overlay methods *5:*563
Index models *5:*597
Index of Biotic Integrity (IBI) *2:*25 *3:*36
  See also Biotic integrity entries
India. *See also* Indian water markets
  climate of *2:*559
  dilution in streams/rivers of *3:*84
  floods and droughts in *2:*566
  fluorosis in *5:*130
  groundwater resources in *2:*562
  groundwater quality in *5:*192
  hydro-borders in *4:*681
  irrigation water quality in *2:*155 *5:*204
  lake water pollution in *3:*451
  major rivers in *3:*449*f* *3:*450t
  marginal cost pricing in *2:*558
  national identity of *4:*684
  promoting self-governing institutions in *2:*558

Index terms	**Links**	
India *(Continued)*		
property rights in	*2*:557	
regions of	*2*:559	
river water pollution in	*3*:447	
subsurface drainage in	*5*:94	
surface water pollution in	*3*:445	
surface water resources of	*2*:561	
trace element groundwater contamination in	*2*:143	
water quantity regulation in	*2*:557	
water resource requirements for	*2*:564	
water resources of	*2*:559	
water stress and water scarcity in	*3*:445	*3*:448f
well-spacing norms in	*2*:558	
Indian 1-MW OTEC plant	*4*:48	
Indian lands, reserved water rights for	*4*:689	
Indian river basins, assessment of nutrients from	*3*:659	
See also Ganga entries		
Indian water markets	*2*:555	
economic aspects of	*2*:556	
extent of	*2*:556	
legal aspects of	*2*:556	
nature of	*2*:555	
Indian wave energy program	*4*:45	
Indicator forms, harmful characteristics in	*2*:443	
Indicator groups, response to pollution	*2*:441	
Indicator monitoring	*2*:267	
Indicator organisms	*2*:25	*2*:292
coliform group	*2*:293	
in domestic sewage	*1*:831	
enterococcus group	*2*:293	
future of	*2*:293	
Indicators		
biotic	*3*:37	
efficient	*3*:34	
Indigenous people, water rights of	*4*:771	
Indirect potable reuse	*2*:611	
Indirect search method	*2*:332	
Indirect soil moisture measurement	*3*:692	
Indirect vadose zone monitoring methods	*5*:539	
Indirect wastewater reuse	*1*:307	*1*:826

Index terms **Links**

Indirect water valuation approaches *4:*608

Indirect water valuation studies *2:*655

Indoor residential water use *2:*652

Induced polarization method *5:*147

Induction, of cytochrome P450 monooxygenase *2:*108

Induction logs *5:*152

Inductive conductivity measurement *4:*432

Inductively coupled plasma (ICP) *2*:308 *2*:309

Indus River, salinity profile of *3:*679
  See also Indus Waters Treaty

Industrial acidification *3:*9

Industrial activities, stormwater discharges from *4:*658

Industrial biocides *1:*602

Industrial byproducts, disposal of *5:*276

Industrial cooling water
  biofouling of *1:*538
  corrosion in *1:*542
  scale formation in *1:*545

Industrial development, Great Lakes region *3:*178

Industrial discharges. *See also* Industrial effluents
  Clean Water Act and *1:*755
  daily toxicity variation of *1:*570t

Industrial effluents. *See also* Industrial discharges
  evaluation by on-line respirometry *1:*565
  limits on *1:*552

Industrial environments, clathrate hydrates in *4:*473

Industrial field studies
  on mussels *1:*512
  on veliger settling *1:*513

Industrial fluoride contamination sources *5:*132

Industrial hydrocarbons, complex *5:*43

Industrial input demands, economic theory of *1:*550

Industrial ion discharges *1:*752t

Industrialized countries, water resources planning and management objectives for *2:*683

Industrial meters *1:*340

Industrial pollutants *3:*281
  macrophytes as biomonitors of *2:*66

Industrial pollution, impact of *4:*107

Index terms **Links**

- Industrial Revolution
  - public health during *1:*283
  - sanitation during *1:*282
- Industrial storm water discharges *1:*867
- Industrial wastes
  - as a source of cultural eutrophication *3:*115
  - as water pollution *4:*97
- Industrial wastewaters *2:*375
  - constructed wetlands for *1:*896
  - microfiltration of *1:*591
  - nutrient-deficient *1:*729t
  - pretreatment of *4:*586
  - reuse of *1:*818
  - toxicity assessment of *1:*566
- Industrial water demand economics *1:*549
  - empirical models in *1:*550
  - future research related to *1:*552
- Industrial water demands, structure of *1:*551
- Industrial water requirement, in India *2:*566
- Industrial water use *2:*544
  - United States *1:*620 *2:*652
- Industry
  - water conservation in *2:*496
  - water consumption in *2:*370
  - water-efficiency improvement in *2:*635
- Indus Waters Treaty *4:*682
  - *See also* Indus River
- Infantile cyanosis (methemoglobinemia) *1:*34 *1:*35
- Infant mortality *1:*27
- Infaunal amphipods *2:*409 *2:*410
- Infection, fecal-oral cycle of *2:*112
- Infection monitoring, versus disease monitoring *1:*189
- Infectious diseases, waterborne *1:*177
  - *See also* Disease entries
- Infectivity assays, *C. parvum* *1:*166
- Infiltration *5:*212
  - controls on *4:*484
  - defined *2:*544
  - for estimating stomatal aperture *3:*716
  - estimation of *4:*486

Index terms	**Links**		
Infiltration *(Continued)*			
excess	*4:*316		
measurement of	*4:*487		
modeling	*4:*486		
quantification of	*5:*663		
soil moisture processes and	*4:*484		
soil water processes and	*5:*210		
storage capacity	*5:*125		
in the vadose zone	*5:*534		
Infiltration capacity	*3:*474	*4:*486	*5:*212
Infiltration-excess overland flow	*3:*452		
Infiltration galleries	*3:*416	*5:*89	
Infiltration rates	*5:*212		
on forest soils	*2:*199		
Infiltration systems, for storm water treatment	*1:*868		
Infiltrometers	*4:*487	*5:*214	*5:*539
Inflow-outflow method	*5:*166		
Influent and effluent seepage	*5:*167		
Informal sector provision model	*1:*51		
Information			
concerning water quality issues	*1:*509		
corrosion control	*1:*154		
water conservation	*1:*148		
Information portals	*2:*670		
Infrared propagation	*4:*296		
Infrared satellite measurements	*4:*323		
Infrasound	*4:*571		
Infrastructure			
investing in improvements in	*1:*871		
in the Nile Basin	*2:*593		
pretreatment programs and protection of	*1:*800		
Ingested radon diffusion model	*4:*544		
Inhibition, effect on nitrification	*1:*754t		
Initial and boundary problem	*5:*621		
Initial dilution	*4:*104		
Initial public offerings (IPOs), corporatized utility	*1:*389		
Injection wells	*2:*544		
drinking water contamination by	*2:*286		
In-lake blue-green algae controls	*3:*189		

Inland waters
  data, reports, and publications on *3:*11
  rain-acidified *3:*8 *3:*292
In-line drippers *3:*622
Inorganic adsorption, ion exchange and *4:*490
Inorganic chemical health issues *1:*822
Inorganic chemical monitoring *2:*267
Inorganic chemicals *5:*185
Inorganic coagulants *4:*425
Inorganic colloids, stabilized by humic/fulvic acid *2:*208
Inorganic compounds
  competitive sorption on activated carbon *1:*110
  in domestic sewage *1:*831
  technologies for *1:*463t
Inorganic-humic agglomerates, metal ion binding to *2:*208
Inorganic ion removal
  adsorbents for *4:*494
  factors affecting *4:*490
  ion exchangers for *4:*494
Inorganic monomeric aluminum, episodic acidification and *3:*7
Inorganic phosphate corrosion inhibitors *1:*154
Inorganic phosphorus *3:*702
Inorganic phytoremediation verification protocol *3:*633
Inorganic pollutants, background concentration of *2:*19
Inorganic wastes, inhibitory threshold concentrations of *1:*754t
Input (forcing) uncertainty *3:*326
Insertion valves *1:*890
In situ biodegradation *2:*44
In situ biological treatment
  of perchlorate *5:*633
  of vinyl chloride *5:*638
In situ bioremediation *5:*442
  of contaminated groundwater *5:*38
  costs of *5:*42
  technologies for *5:*39
In situ chemical monitoring *4:*514
In situ chemical oxidation (ISCO) *4:*445 *4:*446
In situ chemical oxidation technologies *5:*344
In situ destruction technologies *5:*441
In situ dredged material characterization *2:*351

Index terms	**Links**	
In situ electrokinetic remediation, of		
cadmium	*5:*618	
See also In situ groundwater bioremediation; *In situ* groundwater remediation; *In situ* metal remediation		
In situ electrokinetic treatment	*5:*116	*5:*329
See also In situ electrokinetic remediation; *In situ* physical/chemical treatment; *In situ* treatment technology		
achievable cleanup levels associated with	*5:*121	
case histories in	*5:*121	
of chromium, arsenic, and biofouling	*5:*28	
design of	*5:*120	
economics of	*5:*121	
field feasibility characterization for	*5:*118	
In situ groundwater bioremediation. *See also In situ* groundwater remediation; Natural attenuation		
bacterial transport in	*5:*43	
limitations of	*5:*42	
In situ groundwater remediation	*5:*423	
See also Groundwater remediation; *In situ* groundwater bioremediation		
by aeration and volatilization	*5:*426	
for heavy metal contamination	*5:*290	
hydrogeologic characterization for	*5:*425	
In situ iron filing curtain treatment	*5:*329	
In situ lining systems	*1:*883	
In situ metal/organic adsorption	*5:*609	
In situ metal remediation, biological methods of	*5:*292	
In situ monitoring, of odorous emissions	*2:*282	
In situ physical/chemical treatment, of vinyl chloride	*5:*638	
In situ rate coefficients	*5:*523	
In situ sediment toxicity assessments/test methods	*2:*384	*3:*54
In situ soil flushing	*5:*435	
In situ solidification/stabilization technology	*1:*837	
In situ treatment technology	*5:*291	
Instantaneous peak demand (IPD) curves	*1:*213	
Instantaneous samplers	*3:*398	
Instantaneous unit hydrograph (IUH)	*4:*356	*4:*357
Instantaneous unit sediment graph (IUSG)	*4:*359	
Institutional solutions, in the Arab World	*2:*473	
Institutional water conservation measures	*2:*497	
Institutions, irrigation and	*4:*739	

Index terms **Links**

Instream flow incremental methodology (IFIM) *3:*526

Instream flow methods *3:*526

Instream flows *3:*106

legal protection for *4:*659

In-stream values, federal law protection of *4:*663

Instream water use *2:*650 *2:*661

Instrumentation

- CERES *4:*169
- satellite-borne *4:*311

Instrument development, for *in situ* Cr(VI) monitoring *4:*515

Insurance, flood *3:*147

Intake systems, riverbed/seabed filtration *5:*89

Intake wells *5:*88

- vertical *5:*88

Integrated capacity building (ICB)

- multicapital context of *1:*655
- operational considerations for *1:*654
- participatory *1:*652
- role in watershed planning *1:*655f
- sanitation-related *1:*651

Integrated collector/storage solar water heater *1:*63

Integrated Economic and Environmental Satellite Account (IEESA) *2:*627

Integrated modeling *3:*340

- future of *3:*341

Integrated municipal watershed management (IMWM) *1:*497

Integrated Pollution Prevention and Control (IPPC) Directive *2:*45

Integrated precipitable water vapor (IPW), calculating from GPS signal delays *4:*245

Integrated Risk Information System (IRIS) *1:*426

Integrated treatment systems *1:*684

Integrated water quality modeling software tools *2:*328

Integrated water resources management (IWRM) *1:*122 *2:*263 *2:*574 *2:*627

Integrated watershed management (IWM) *1:*651 *2:*613 *2:*627

Intensity-depth-frequency (IDF) curves *4:*315

Intensity-duration-frequency (IDF) curve *3:*157

Intensive farm systems *3:*579

Intensive partial molar internal energies *4:*363

Interbasin Water Transfer Act *4:*688

Interception *3:*235

- consequences of underestimating *3:*238
- determining evaporation from *3:*237

Index terms **Links**

Interception *(Continued)*
  importance of *3:*236
Interdisciplinary research, on hydrologic thresholds *3:*231
Interdisciplinary water studies. *See* SUSTAINIS advanced study course
Interfacial phenomena *4:*449
Intergenerational water resources *4:*771
Intergovernmental organizations (IGOs), water rights and *4:*771
Intergovernmental Panel on Climate Change (IPCC) *2:*640 *3:*208 *4:*120
  Report by *2:*483
Interim Enhanced Surface Water Treatment Rule (IESWTR) *1:*233 *1:*428 *2:*195 *2:*344
Interlaboratory precision tests *2:*459
Intermittent waters *2:*475
International Bottled Water Association (IBWA) *1:*4
‘‘Internal improvements,’’ American *2:*498
International Boundary and Water Commission (IBWC) *4:*620 *4:*644
International Coastal Cleanup (ICC) *4:*39
International Conference on Freshwater *3:*33
International Conference on Water and the Environment *4:*606
International Coral Reef Initiative (ICRI) *4:*115
International Decade for Natural Disaster Reduction (IDNDR) *3:*153 *3:*154
International experts/advisors *1:*719
International fishing waters *3:*532
International freshwater management *4:*636
International GPS Service (IGS) tracking stations *4:*250
International Humic Substances Society (IHSS) *4:*399
International Joint Commission (IJC) *3:*179 *4:*618 *4:*620
International Maritime Organization—London Convention (IMO-LC) ocean dumping reports *4:*144
International Network to Promote Household Drinking Water Treatment and Safe Storage *1:*68
International standards, water reuse *3:*670
International Strategy for Disaster Reduction (ISDR) *3:*154
International treaties *2:*473 *2:*597
  See also International water treaties
International Union for the Conservation of Nature (IUCN) *2:*626
International Union of Pure and Applied Chemistry (IUPAC) *4:*491 *4:*525
International virtual water flows, calculating *2:*540
International Water Association (IWA) models *1:*732 *1:*736
International water market, trends in *2:*568

Index terms	**Links**	
International water resources management, Great Lakes region	*3:*179	
International Water Services Association (IWSA)	*2:*497	
International water treaties	*2:*618	
Internet access gap	*2:*670	
Internet portals, water-related	*2:*668	
Internet sources, for inland-water-related data, reports, and publications	*3:*11t	
Interparticle bridging	*4:*425	
Interstate commission compacts, drought management by	*2:*585	
Interstate Commission on the Potomac River Basin (ICPRB)	*2:*577	
Interstate compacts	*4:*613	
See also Interstate water compact		
Interstate water compact	*2:*503	
Interstellar space, water in	*4:*189	
Interstitial water	*2:*384	
Intertidal macrofauna, seasonal coupling with	*4:*73	
Intertropical convergence zone (ITCZ)	*4:*225	
Intestinal parasites	*1:*159	
Intracellular accumulation, of metals	*5:*283	
Intracellular biodegradation	*2:*43	
Intracellular fluid (ICF)	*4:*459	
Intraparticle diffusion	*3:*303	*4:*567
Intrastate policies	*4:*649	
Intrastate regional authorities, drought management by	*2:*585	
Intrinsic permeability	*5:*537	
Intrinsic susceptibility	*5:*595	
Inventory agencies, federal	*4:*650	
Inverse mass balance models	*5:*139	
Inverse model, for parameter estimation	*1:*135	
Inversion, fog and	*4:*239	
Invertebrate Community Index (ICI)	*3:*39	
Invertebrates, toxicity of lead to	*2:*434	
Investment, in security and infrastructure improvements	*1:*871	
In vitro bioassays	*2:*280	
In vitro cytotoxicity techniques	*2:*413	
In vitro fish studies, challenges and obstacles in	*3:*116	
In vitro response tests	*3:*116	
In vitro toxicology models	*3:*117	
In vivo bioassays	*2:*280	
Iodinated trihalomethanes, health effects of	*1:*266	

Index terms	**Links**			
Iodine number	*1:*98			
Ion chromatography	*1:*358			
Ion discharges, industrial	*1:*752t			
Ion exchange	*1:*297	*1:*380	*1:*725	*1:*812
	*1:*827	*5:*527		
adsorption mechanism in	*4:*493			
advantages and limitations of	*1:*297			
for arsenic removal	*1:*638	*5:*21		
for DOC removal	*1:*325			
inorganic adsorption and	*4:*490			
for iron and manganese removal	*1:*315			
in landfill leachates	*1:*707			
potential	*5:*185			
for radioactive waste	*1:*804			
for radionuclide removal	*1:*397			
in small drinking water systems	*1:*459			
Ion exchange resins	*1:*413			
Ion exchangers				
for inorganic ion removal	*4:*494			
removal of heavy metal ions by	*4:*494t			
Ion exchange technology, for arsenic treatment	*5:*23			
Ionic concentration, in groundwater tracing	*5:*503			
Ionization	*5:*527			
Ionizing irradiation	*1:*873			
Ion precipitation, kinetic formulations for	*5:*416t			
Ions. *See also* Hydrogen ion entries				
hydration of	*4:*476			
in seawater	*4:*159			
Ion-selective field effect transistor (ISFET)	*2:*298			
IRC International Water and Sanitation Centre websites	*2:*669			
Ireland, mine wastewater treatment in	*1:*897			
Iron (Fe)	*4:*496			
See also Ferric hydroxides; Iron removal				
arsenic removal using	*1:*636			
biofouling	*5:*36			
chemistry of	*4:*496			
complexing	*5:*609			
concentration of	*4:*498			
environmental impact of	*5:*608			
forms of	*1:*313			
in groundwater	*5:*406			

Iron (Fe) *(Continued)*
  health and regulatory issues related to *4:*497
  levels and solubility in water *4:*497
  methods for determining *4:*498
  oxidation *5:*609
  oxidation states of *4:*534
  problems caused by *1:*313
  removing from water *4:*499
Iron acrylate foulant *1:*418t
Iron bacteria *5:*22 *5:*407
  See also Iron-oxidizing bacteria; Iron-related bacteria (IRB)
  biofouling and water color caused by *2:*151
  metabolic activities of *2:*153
Iron-bearing minerals *5:*345
Iron coprecipitation, of arsenate *5:*20
Iron corrosion *1:*544
Iron filing curtain treatment *5:*329
Iron industry, impact on surface water quality *3:*374
Iron oxidation rates *2:*152
Iron oxide coated sand, adsorptive filtration using *1:*639
Iron oxides *5:*609
Iron-oxidizing bacteria *2:*149
  See also Iron- reducing bacteria
  in acid mining drainage *2:*150
  in groundwater and surface water *2:*150
Iron precipitation *5:*422
Iron pyrite (FeS_2) *3:*13
  See also Pyrite
Iron-reducing bacteria *2:*151
Iron-reducing to sulfate-reducing conditions *5:*583
Iron-related bacteria (IRB) *1:*84
Iron removal *1:*312
  alternative treatments for *1:*314
  ammonia and manganese inhibition of *2:*153
  cartridge filters for *2:*152
  common treatment processes for *1:*313
  using ozone *1:*354
Iron removal systems, biological *2:*152
Iron salts, coagulation using *1:*636
Iron sulfides, formation of *5:*36

Index terms	**Links**		
Irradiation, ultraviolet	*1:*469		
See also Ultrasonic irradiation; Ultraviolet entries			
Irreversible specific toxicity	*2:*415		
Irrigated agriculture	*2:*635	*3:*559	*3:*673
sustainability and future of	*3:*584		
Irrigated land, maintaining salt balance on	*3:*677		
Irrigated regions, tile drainage in	*3:*731		
Irrigation. *See also* Irrigation management; Irrigation performance assessment; Irrigation water; Microirrigation			
agricultural	*1:*818		
in the Central Asian plains	*3:*17		
civilization and	*4:*737		
defined	*2:*544		
diversions of	*2:*617		
efficiency of	*2:*495	*3:*555	
extent of	*3:*581		
first world history of	*4:*736		
groundwater for	*2:*156		
in India	*2:*565		
influence on nitrate content	*3:*638		
managing	*2:*186		
methods of	*2:*375	*3:*581	
objectives of	*3:*582		
sand abstraction for	*3:*413		
sprinkler	*3:*712		
trends in	*3:*586		
in the United States	*3:*586		
water management and	*4:*771		
in Western states	*3:*587		
Irrigation areas, major	*3:*581		
Irrigation management	*3:*583		
Irrigation performance assessment	*3:*583		
Irrigation practices, reclaimed	*3:*669		
Irrigation regulations, Islamic	*4:*635		
Irrigation-system maintenance	*3:*554		
Irrigation techniques, farming and	*3:*591		
Irrigation water	*2:*370		
annual use of	*2:*653		
efficiency and management of	*2:*572	*2:*617	
electrical conductivity for	*2:*159t		
presence of salts in	*3:*678		

Index terms **Links**

Irrigation water *(Continued)*
  quality of *2:*155 *2:*371
  use of *2:*544
Irrigation water quality, in Hardwar, Uttaranchal, India *5:*204
Irrigation wells *3:*594
Islamic law, water management prescriptions in *2:*474
  See also Islamic water law
Islamic water law *4:*634
Island Reports (Mediterranean) *2:*640
Isochrones *3:*61
Isochrone spacing methods *3:*63
Isoelectric point (IEP) *4:*509 *4:*511
Isohyetal map *4:*291
  software *4:*292
Isohyetal method *4:*290
Isolation technologies *5:*433
Isoparametric transformation mapping scheme *5:*657
Isotherm models *4:*385t
Isotherms, adsorption *1:*99
  See also Freundlich adsorption isotherms; Langmuir adsorption isotherms; Mathews and Weber multicomponent isotherm model
Isotope equilibration *4:*540
Isotope exchange, in gas-water reactions *4:*535
Isotope exchange rate, S-state dependence on *4:*537
Isotope fractionation *4:*500
  See also Isotope separation
  artificial *4:*502
  curve *4:*501
Isotope ratio mass spectrometry (IRMS) *1:*645
Isotopes *4:*499
  See also Carbon isotopes; Chlorine isotopes; Environmental isotopes; Nitrogen isotopes; Oxygen isotopes; Radioactive isotopes; Radionuclides; Stable isotopes; Sulphur isotopes; Uranium series isotopes
  in contaminant hydrogeology *5:*216
  as groundwater tracers *5:*502
  hydrogen *4:*438
  hydrology of *5:*217
  naturally occurring *5:*232
  new applications of *5:*218
  stable and unstable *1:*803

Index terms	**Links**	
Isotopes *(Continued)*		
strontium	*4:*574	
technetium	*4:*578	
Isotope separation	*4:*500	*4:*502
See also Isotope fractionation		
techniques for	*4:*502	
Isotopic composition, dependencies of	*5:*230	
Isotopic fractionation	*5:*229	
Isotopic hydrology	*4:*440	
Isotopic techniques, for base flow separation	*3:*27	
Israel		
public works projects of	*4:*699	
reclaimed water usage regulations in	*3:*672	
water resources of	*4:*699	
Israeli-Arab water resources	*4:*699	
Israel-Jordan Treaty of Peace	*4:*700	
Israel-Syria water conflict	*4:*754	
IWR-MAIN water demand forecast model	*2:*529	*2:*530
J		
Jacob’s well	*4:*735	
Jain and Snoeyink adsorption model	*1:*108	
Jaffna Peninsula, ancient wells in	*4:*775	
Jakarta, flooding in	*3:*160	
Jamin effect	*5:*421	
Japan, research on wave energy in	*4:*45	
Japan Sea, radioactive contaminant transport in	*3:*322	
Java activated sludge process simulator (JASS)	*1:*731	
‘‘Jet effect,’’	*3:*80	
Jetting	*5:*234	
See also Water-jetting drilling technologies		
technology	*5:*234	*5:*235
Joint entropy	*4:*219	
Joints		
rock	*5:*136	
water distribution system	*1:*208	
Joint venture water/sewerage contracts	*1:*391	
Joint Water Committee (JWC)	*4:*700	
Joliet Army Ammunition Plant survey	*5:*77	
Joliet project	*3:*44	

Jordan Basin *4:*755
Jordan River watershed *4:*755
“Joseph effect” *3:*211

K

‘‘Kair’’ farming *3:*164
Kansas, source-water protection in *2:*312
Karren *5:*244
Karst *5:*600
  See also Epikarst; Karst topography
  defined *5:*108
  groundwater dye tracing in *5:*107
Karst hydrology *5:*235
  groundwater discharge in *5:*238
  groundwater recharge in *5:*236
  groundwater storage in *5:*238
  groundwater tracing in *5:*240
  groundwater transport in *5:*237
  modeling of *5:*241
  spring hydrograph and chemograph analysis in *5:*239
Karstic aquifers *1:*365 *5:*237
  calculating flow in *1:*368
  Cryptosporidium in *1:*366
  Giardia in *1:*367
  Microsporidium in *1:*367
  parasite fate and transport in *1:*365
  volume of water in *5:*238
Karst landscapes *5:*236
  See also Karst topography
  vulnerability of *5:*241
Karst topography *5:*243
  bedrock dissolution in *5:*243
  development of *5:*243
  groundwater recharge in *5:*247
  setting types in *5:*247
  springs in *5:*245
  subsurface features in *5:*245
  surface features in *5:*244
  valleys in *5:*245
  vulnerability to contamination and pollution *5:*247
Kesterson Wildlife Refuge, heavy metals in *5:*276

Index terms **Links**

‘‘Keystone’’ predator *4:*153
‘‘Khaki’’ farming *3:*164
Kick-sampling *3:*39
“Killer fog” *4:*379
Killing frosts *4:*241 *4:*242
Kilowatt-hours *3:*202
Kinematic cascade *3:*249
  shocks in *3:*241
Kinematics *3:*247
Kinematic shock *3:*239
  determining *3:*249
  formation of *3:*248
Kinematic time to equilibrium *3:*63
Kinematic viscosity *5:*556
Kinematic wave equation *3:*248
Kinematic wave flow routing *3:*253
  numerical solution for *3:*254
Kinematic wave flux laws *3:*249
Kinematic wave method, for storm drainage design *3:*242
Kinematic wave model *3:*470
Kinematic wave speed *3:*253
Kinematic wave theory *3:*242 *3:*247
  validity of *3:*250
Kinetic analysis *2:*71
Kinetic modeling *5:*414
  equilibrium and *4:*565
Kinetic processes, in water quality models *2:*271
Kinetic reactive transport models *5:*519
Kinetics
  chlorine reaction *1:*131
  of metal ion-humic colloid interaction *2:*209
  multicomponent *1:*119
  sorption *4:*564
Kinetic studies *4:*565
Kirpich formula *3:*471
Klepsydras *4:*704 *4:*705
Knowledge portals *2:*670
Knowledge transfer *1:*719
Kona coast, rainfall in *4:*257
Koppen climatic zones *4:*256

Koran, water references in *4:*746
Kousiss flow routing model *3:*257
^{85}Kr
  anthropogenic sources and atmospheric distribution of *5:*248
  detecting modern groundwaters with *5:*248
  production mechanisms and decay behavior of *5:*248
Krasheninnikov Bay, environmental quality of *2:*451t
^{85}Kr dating method
  characteristics of *5:*248
  versus CFC dating methods *5:*249
  versus H-He dating *5:*68 *5:*249
Krypton. *See* ^{85}Kr entries
Kyoto Protocol *1:*560

L

Labeled substrate probes *1:*645
Lability, of metal species *2:*212
Laboratory safety *2:*161
Laboratory testing
  for levee construction *3:*286
  toxicity *4:*112
Lacustrine valleys *3:*66
Ladyzhenskaya-Babuska-Brezzi (LBB) stability condition *5:*658
Lagergren rate constants *3:*302 *3:*303t
Lagooning *1:*324
Lagrangian coordinate system *5:*558
Lagrangian drifters *4:*71
Lagrangian models *1:*133
Lake and pond water *2:*95
Lake Baikal
  climatic variations in *3:*21
  environmental and geological settings of *3:*20
  future drilling sites in *3:*21
Lake Baikal rift system
  cooperative studies of *3:*20
  studies of *3:*20
  understanding *3:*20
Lake basins, origin of *3:*265
  See also Basin entries
Lake Bleiswijkse Zoom, biomanipulation trials in *2:*55t

Index terms **Links**

Lake Bysj.on, biomanipulation trials in *2:*54t
Lake Cristina, biomanipulation trials in *2:*54t
Lake districts, from low altitude glaciations *3:*265
Lake Duiningermeer, biomanipulation trials in *2:*55t
Lake ecosystems *3:*267
  alternative stable states theory regarding *3:*272
Lake Feldberger, biomanipulation trials in *2:*57t
Lake fisheries management *3:*125
Lake Froylandsvatn, biomanipulation trials in *2:*56t
Lake Gjersjoen, biomanipulation trials in *2:*54t
Lake Haugatjern, biomanipulation trials in *2:*54t
Lake Lilla Stocke-lidsvatten, biomanipulation trials in *2:*53t
Lake Mead *2:*555
Lake Michigan *3:*274
  biomanipulation trials in *2:53t* *2:*56t
  NOAA lake level forecast for *3:*274
Lake Michigan Wireless Environmental Observatory *4:*72
Lake Mosvatn, biomanipulation trials in *2:*54t
Lake Number *3:*266
Lake Okeechobee *3:*279
Lake Orta, Italy, industrial acidification of *3:*9
Lake plankton, food chains, filtrators, and size spectra in *3:*268
Lake Pontchartrain and Vicinity Hurricane Protection Project *3:*284
Lake purification *2:*50
Lake restoration methods, classification of *2:*50
Lake Rusutj. arvi, biomanipulation trials in *2:*56t
Lakes *3:*265 *4:*285
  See also Lacustrine valleys
  acidic hypersaline *3:*9
  acidified *3:*9 *3:*10
  blue-green algae blooms in *3:*188
  Canadian *2:*658
  chronic acidification of *3:*3
  climatic zones and *3:*266
  composition of *4:*285
  discharges to *3:*281
  ecosystems of *3:*267
  glacial *3:*269
  management of *3:*125
  man-made *3:*270 *3:*293

Index terms **Links**

Lakes *(Continued)*
  physical characteristics of *3:*282
  pollutant fate in *3:*283
  polluted discharges into *3:*281
  programs for protecting *3:*270
  regional limnology of *3:*269
  responses to earthquakes *5:*112
  stratification and mixing in *3:*266
  submerged aquatic plants in *3:*275
  trophic state and productivity of *3:*267
  tropical and ancient *3:*270
  water budget and salin1ty of *3:*266
  water quality impairment in *3:*275
  waves and oscillations in *3:*267
Lake sediments, pollutant deposition into *3:*283
Lake Severson, biomanipulation trials in *2:*53t
Lake SøbygÅrd, biomanipulation trials in *2:*55t
Lake Tenkiller model *2:*260
Lake Trummen, biomanipulation trials in *2:*53t
Lake Væng, biomanipulation trials in *2:*54t
Lake Vostok *3:*503
  bacterial cell densities in *3:*506
Lake Washington, biomanipulation trials in *2:*53t
Lake water(s)
  degradation of *3:*283
  sodium in *4:*552
Lake water pollution, in India *3:*451
Lakewide Management Plans (LaMPs) *3:*179
Lake Wolderwijd, biomanipulation trials in *2:*55t
Lake Zwemlust, biomanipulation trials in *2:*55t
Lambert-Beer law *3:*192
Laminar flow *3:*195 *5:*649
  determination of *5:*63
Laminar flow systems, equations of change in *5:*651
Land and water management agencies, federal *4:*650t
Land Data Assimilation System (LDAS) *2:*588
Land degradation processes, remote sensing and GIS application in *2:*534
Land development activities, stream quality and *3:*442
Land disposal *5:*254
Land drainage *5:*99

Index terms **Links**

Land Evaluation and Site Assessment (LESA) system *3:*596
Landfill buffer land *2:*166
Landfill design, bioreactor *5:*257
Landfill disposal practices, influence on metal speciation and mobility *1:*723
  See also Landfilling
Landfill gas
  collecting *2:*168
  from municipal solid waste landfills *2:*165
Landfill gas migration *5:*256
  controls on *5:*256
Landfilling
  of sludge *1:*860
  using pesticide-contaminated soil *3:*653
Landfill leachates *1:*699 *5:*277
  See also Landfill leachate treatment; Leachate entries
  characteristics of *1:*701
  constructed wetlands for *1:*896
  contamination by *5:*255
  generation of *1:*700
  recirculation of *1:*708
Landfill leachate treatment
  activated carbon in *1:*105
  biological processes for *1:*703
  physical-chemical *1:*704
  selecting processes for *1:*708
Landfill pollutant loadings, equalization of *1:*705
Landfill-polluted environments, metal-organic interaction in *5:*259
Landfill-polluted water *5:*259
Landfill regulations *2:*166
Landfills *1:*695
  bioreactor technology for *1:*727
  components of *1:*700f
  contamination from *5:*452
  degradation cycle of *1:*725
  design and construction of *1:*697
  environmental impacts of *1:*696
  groundwater contamination from *5:*253
  groundwater remediation related to *5:*257
  inefficiency of *1:*696
  municipal solid waste *2:*163
  pollution prevention related to *5:*256

Landfills *(Continued)*
  refuse decomposition microbiology in *1:*695
  uncontrolled runoff impacts from *5:*256
Landfill system *1:*724f
Land-grant colleges/universities (LGUs) *3:*597
Land improvement works *1:*403
Land management practices, in the 1930s *2:*511
Land-ocean-atmospheric system *3:*230
Land-ocean interaction in the coastal zone (LOICZ) methodology *4:*53
Land reclamation, in acid mine drainage *5:*2
Land resources. *See* Poor-quality land resources
Land runoff, as a nutrient source *3:*110
Landscape
  irrigation of *2:*651
  root zone components of *3:*706
Landscape analysis *4:*315
Landscape plants *3:*706
Landscape process studies *4:*267
Landscape water-conservation techniques *3:*553
  definitions related to *3:*553
  drought or maximum heat-day practices *3:*555
  postdrought or maximum heat-day practices *3:*555
  predrought/premaximum heat-day practices *3:*554
  two-track strategy for *3:*554
  water budget program *3:*554
Landscape water management *3:*751
Landside berms *3:*290
Landside seepage berms *3:*289
Landslide dam pools *3:*69
Landslide warning systems *1:*125
Land surface, rainfall partitioning at *4:*315
Land surface modeling *3:*533
  evolution of *3:*534
  framework for *3:*534
  inputs and outputs related to *3:*536
Land Surface Models (LSMs) *2:*588
Land surface precipitation, changes in *4:*173
Land topographic features, influence on waterlogging *3:*742

Index terms	**Links**			
Land use				
effect on water quality	*2:*169			
Great Lakes area	*3:*178			
Land-use activities, as threats to groundwater quality	*2:*183t			
Land-use changes, influence on the hydrologic cycle	*4:*280			
Land use classification, remote sensing and GIS application in	*2:*532			
Land use control, extraterritorial	*1:*315			
Land use data	*5:*164			
Land use impacts, on groundwater quality	*5:*250			
Land-use ordinances, flood-related	*3:*147			
Langelier saturation index (LSI)	*1:*206	*4:*415		
Langmuir adsorption isotherms	*1:*101	*1:*112	*1:*113	*1:*114t
See also Isotherms				
Langmuir coefficient	*3:*99			
Langmuir coefficient QSAR model	*3:*99			
Langmuir isotherm	*4:*567			
Langmuir isotherm constants	*1:*114t			
Langmuir isotherm equation	*2:*87	*4:*566		
Langmuir isotherm model	*3:*300	*4:*385		
La Paz Agreement	*4:*646			
La Plata Basin	*4:*642			
Large area surface energy balance estimation	*3:*560			
data accuracy in	*3:*564			
Larvae, in marine toxicological testing	*4:*42			
Laser scans, of levees	*3:*284			
LAS surfactants	*1:*670			
Las Vegas, Nevada, overdraft in	*5:*343			
Latent heat flux, estimating	*4:*2			
Latent heat flux density	*3:*563			
Lateral connection repair	*1:*888			
Lateral radial wells	*5:*106			
Latin America. *See also* Americas				
transboundary waters in	*4:*636			
Web use in	*2:*671			
Latitude effect	*4:*439			
Law, irrigation	*4:*739			
See also Groundwater law; Water law				
Law of basin areas	*3:*94			
Law of stream lengths	*3:*93			

Index terms **Links**

Law of the Sea Convention *3:*126
Law of stream numbers *3:*93
Lawrence Livermore National Laboratory (LLNL) report *4:*674
Laws, water sustainability *2:*627
  See also Legal entries; Legislation; Regulations
Lax-Wendroff scheme *3:*239 *3:*241
Layered heterogeneity *5:*509
L (acidic) carbons *1:*98 *1:*99
Leachate collection and removal, from municipal solid waste landfills *2:*164
Leachate composition, in landfills *1:*696
  See also Landfill leachates
Leachate landfill *5:*255
Leachate quality, treatment techniques based on *1:*709
Leachates, subtitle D landfill *5:*258
Leachate toxicity evaluation *2:*280
Leachate treatment plant *1:*710f
Leach beds, alkaline *5:*7
LEACH code *5:*564
Leaching *2:*544 *5:*260
  of cement hydration products *5:*364
  of nitrogen *3:*696
Leaching Estimation and Chemistry Model (LEACHM) *5:*597
Leaching Potential Index (LPI) *5:*597
Leaching practices, control of *3:*679
LEACHM model *3:*680
Lead (Pb). *See also* Pb bioavailability
  cardiovascular effects of *2:*436
  chemistry *5:*367
  ecotoxicological effects of *2:*433
  effects on central nervous system *2:*435
  effects on endocrine system *2:*436
  effects on peripheral nervous system *2:*436
  effects on renal system *2:*436
  effects on reproduction *2:*436
  as an environmental contaminant *5:*365
  fate and transport of *5:*647
  future research on *2:*437
  in groundwater *2:*148 *5:*645
  guidelines and standards for *5:*647
  health effects/risks of *2:*432 *5:*381 *5:*646

Lead (Pb) *(Continued)*
  hematopoietic effects of *2:*437
  intoxication by *2:*435
  maximum contaminant level of *2:*234t
  movement in the environment *2:*433
  neurological effects of *2:*435
  poisoning by *2:*437
  properties and uses of *5:*646
  remediation of *5:*648
  sources of *2:*433 *5:*381
Lead arsenate *5:*646
Lead compounds, insoluble *5:*382
Lead concentration, along the Blackstone River *2:*259
Lead-contaminated soils
  chemistry of *5:*381
  phytoextraction for *5:*382
  phytostabilization for *5:*383
  rhizofiltration for *5:*383
Lead phytoremediation *5:*381
  advantages and limitations of *5:*385
  future of *5:*368
  plants and mechanisms used in *5:*384t
  role of synthetic chelates in *5:*383
trends in *5:*385
  Lead uptake, from sediment pore water *2:*217
Leaf area index *3:*734
Leaf level transpiration *3:*735
Leaf stomatal conductance *3:*735
Leakage factor *5:*492
Leak detection and repair *1:*317 *2:*663 *2:*664
  benefits of *1:*318
  coordinating with other activities *1:*319
  strategy for *1:*318
  as a water conservation measure *1:*146
Leaking vapors *5:*552
Leaks. *See also* Leak detection and repair
  causes of *1:*318
  sealing *1:*881
  testing *1:*886
"Leaky aquifer" *5:*11
Lease contract (affermage) model *1:*50

Index terms **Links**

Least squares method *2:*332
  for determining base flow recession constants *3:*25
Lee partitioning array *5:*445
Legal considerations, in water quality management *2:*183
Legal protection, for in-stream flow *4:*659
Legal solutions, in the Arab World *2:*473
Legal system, Chinese *2:*486
Legendre transformation *4:*363 *4:*364
Legionella *1:*278 *1:*560
Legionella pneumophila *1:*18
Legionnaires' disease *1:*181
Legislation. *See also* Federal legislation; Laws; Regulations
  drought-related *2:*578
  extinction-related *3:*128
  floatable debris-related *4:*39
  water *2:*96
  water-quality *2:*586
  wetlands-related *4:*691
''Legislative'' toxicity *2:*47
Le Gray, Gustave *4:*766
Legumes, nitrogen in *3:*695
Length heterogeneity PCR (LH-PCR), community profiling using *1:*643
Lentic waters *2:*544
Leptothrix ferrooxidans *2:*149
Levan and Vermeulen adsorption model *1:*109
Levee construction *3:*290
  field investigation and laboratory testing related to *3:*286
Levee embankments, seepage through *3:*290
Levees *2:*544 *3:*152
  design of *3:*286
  embankment design for *3:*287
  for flood protection *3:*286
  laser scans of *3:*284
  rehabilitating *3:*290 *3:*291t
  settlement of *3:*289
  stability increase measures for *3:*288
Level drains, falling water table between *5:*96
Levenberg-Marquardt method *4:*567
Lidar *3:*284
Lidar systems *4:*296

Index terms **Links**

Life cycle costs, of diffused air aeration systems *1:*630
Life cycle tests *3:*52 *4:*42
  endpoints measured in *3:*53
  setup and activity schedule for *3:*53t
Life history method *4:*323
Life of a reservoir, sedimentation and *3:*409
Lifting condensation level (LCL) *4:*369
Ligand chemicals *2:*236
Ligand groups *2:*204
Ligands, colloidal *4:*29
Light, microorganism growth and *1:*606 *3:*311
  See also Photon emitters; Ultraviolet entries
Light nonaqueous phase liquids (LNAPLs) *5:*216 *5:*340 *5:*439
Light Detection And Ranging (lidar) *3:*284
Lightning rod experiment (Franklin) *4:*742
Light penetration, effect on algal growth *3:*111
Light production, measuring *2:*48
Lime
  in acid mine treatment *5:*4
  precipitation *5:*20
Lime neutralization *1:*610
Lime-soda ash processes *1:*320
  chemical reactions of *1:*320
  implementation and limitations of *1:*322
Lime-soda ash softening plant *1:*315
Lime softening *1:*322
  See also Softening
  for arsenic removal *1:*636 *5:*24
  for corrosion control *1:*153
  pretreatment for *1:*323
  of radioactive waste *1:*804
  radionuclide removal via *1:*397
  requirements for *1:*323
  residuals from *1:*324
  in small drinking water systems *1:*460
Limestone aquifers *1:*365 *5:*145
  karstic *1:*365
Limestone channels, open *5:*6
Limestone drains, anoxic *5:*5
‘‘Limited use zone’’ (LUZ) *3:*81

Index terms **Links**

Limiting nutrients *4:*52
Limits to Growth, The *2:*625
“Limnic ratio” *3:*269
Limnodrilus, heavy metal uptake by *2:*215
Limnology *3:*291
  regional lake *3:*269
Limnotrissa *3:*269 *3:*270
Linear flux-concentration law *3:*250
Linearization procedures *5:*13
Linear nonthreshold (LNT) model *4*:541 *4*:543
Linear polarization resistance (LPR) *1:*9
Linear programming techniques *1:*551
Linear reservoir storage coefficient *3:*62
Linear shallow water theory *4:*135
Line breaks
  digging up pipe with *1:*402
  emergency repair of *1:*401
  in laying new pipes *1:*402
  repairing *1:*400
Liners, landfill *1:*698
Lines of evidence (LOEs) *2:*353 *2:*429
  biomarkers and bioaccumulation as *2:*426
Line stops *1:*890
Linings
  close-fit *1:*878
  cured-in-place *1:*878
  pipe *1:*154
  spray-on *1:*879
Lipids, as indicators of organic contamination *2:*442
Lipophilic contaminants, SPMD-TOX paradigm monitoring of *2:*170
Liquid chromatography, high performance *2:*308
Liquid junction *2:*466
Liquid/liquid solvent extraction *2:*307
Liquid monitoring techniques, direct *5:*541
Liquid-phase carbon adsorption *2:*364
Liquid-phase granular activated carbon adsorption (GACA) *1:*101
Liquids. *See also* Fluids
  dense nonaqueous phase *5:*91
  properties of *3:*194
Liquid traps *2:*307

Index terms	**Links**		
Liquid water/water vapor system	*4:*263		
Litigation, wetlands-related	*4:*691		
Litter, marine	*4:*38		
Littoral zone	*3:*267		
Livestock production, impact on surface water quality	*3:*374		
Livestock grazing, managing	*2:*186		
See also Dairy entries			
Livestock safety, weed control and	*3:*744		
Livestock watering, sand abstraction for	*3:*413		
Livestock water use	*2:*544		
Living filters	*2:*102		
Local agricultural land use planning programs	*3:*596		
Local consumption advisories	*3:*120		
Local currents	*3:*320		
Local fishing waters	*3:*532		
Local government, approach to droughts	*2:*577		
Local management agencies	*2:*193	*2:*195	
Local meteoric water line (LMWL)	*4:*439	*4:*440	
Local search methods	*2:*332		
Local water resource agencies	*4:*649		
"Local water use efficiency"	*2:*536		
Local weather	*4:*361		
Locus of enterocyte effacement (LEE)	*2:*137		
Logan's method	*5:*495		
Logarithmic transformation	*4:*299		
Logging			
rain forest	*2:*667		
Great Lakes region	*3:*178		
Logistic regression	*5:*596		
Log K_{oc}	*4:*524		
Log K_{ow}	*4:*523		
Log removal	*1:*249		
London Convention (LC)	*4:*40		
ocean dumping reports	*4:*144		
London-van der Waals bonding	*5:*18	*5:*21	*5:*350
Longitudinal dispersion	*5:*46	*5:*273	
Longitudinal scour pools	*3:*68		
Long pathlength absorbance spectroscopy (LPAS)	*4:*515		
Long Term Ecological Research (LTER) site	*3:*391		

Index terms **Links**

Long-term persistence *3:*210 *3:*211 *3:*212

Long-term simulation, of pollution *3:*335

Longwave radiation *4:*179

Loop system weight loss corrosion measurement *1:*8

Loose sand densification *3:*288

Lophelia pertusa

  distribution of *4:*34

  threats to *4:*36

Los Angeles County, water withdrawals by *2:*479

Loss trends, coastal-wetland-related *3:*72

Lotic aggregates, microbiology of *3:*305

Lotic waters *2:*544

Louver screens, horizontal *5:*572

Low altitude glaciations, lake districts from *3:*265

Low dosage hydrate inhibitors *4:*474

Lower explosion limit (LEL) *4:*561

Low flow. *See* Base flow

Low flow aerators *2:*666

Low-flow groundwater purging/surging *5:*404

  techniques for *5:*406

Low-flow sampling *5:*455

Low-frequency active sonar (LFAS) *4:*571

Low impact development (LID) *2:*188

Low-income countries, residential water demands in *1:*14

Low level organic waste compounds, water contamination by *5:*60

Low rate treatment systems *1:*841 *1:*842t

  anaerobic *1:*907f

Low-tech wastewater reuse system *5:*329

Low-water-solubility compounds *2:*44

‘‘Low water-use’’ plants *3:*556

Luminescent bacteria

  for determining mutagenic and carcinogenic agents *2:*174

  for determining nutrients in water *2:*175

  in online water toxicity monitoring *2:*174

  in water quality determination *2:*172

Luminescent bacterial biosensors

  for genotoxicity assessment *2:*455

  recombinant *2:*454

  for toxicant detection *2:*453

Luminescent-marked genes *2:*453

Index terms **Links**

Luminometers *2:*49f

Luminometry *2:*48

Lumped modeling approach, in flood studies *3:*156

Lung cancer, cadmium inhalation and *5:*615

*lux*CDABE reporter system *2:*453

lux genes
  for determining stress-inducing agents or toxic chemicals *2:*174
  in water quality determination *2:*172

Lyngbyatoxin A *2:*389

Lysimeters *3:*572 *5:*74 *5:*487
  drainage *5:*488
  equilibrium-tension *5:*489
  history of *5:*487
  installation and sampling procedures for *5:*490
  nonweighing drainage *5:*488

Lysimeter soil water sampling *2:*340

Lysosomes *2:*31

M

Maar lakes *3:*265

Machine-slotted (milled) screens *5:*573

Macroalgae, trace elements in *3:*456

Macroarrays *2:*60

Macrocalibration *3:*318

Macrofauna, intertidal *4:*73

Macrofaunal communities *4:*11

Macrofouling, of surfaces *1:*541

Macroinvertebrate assemblages *3:*40

Macroinvertebrate index of biotic integrity (Macroinvertebrate IBI) *3:*39

Macroinvertebrates, aquatic *2:*321

Macrolite *1:*314

Macrophyte buffer mechanisms *3:*273

Macrophytes. *See also* Aquatic macrophytes
  in the aquatic ecosystem *1:*714
  as biomonitors *1:*716 *2:*66
  contaminant exposure and uptake by *1:*715
  halogenated compounds and *4:*150
  as PCB biomonitors *1:*714
  role of *1:*894
  as trace metal biomonitors *2:*64

Index terms	**Links**	
Macropore flow, vadose zone	*5:*73	
Macropores	*1:*97	*3:*474
Macropore throughflow	*3:*453	
Macroscale hydrologic modeling	*3:*534	
Madhya Ganga Canal	*3:*234	
Magnesium hardness	*4:*453	
See also Mg/Ca ratio		
Magnetic geophysical methods	*5:*148	
Magnetic ion exchange (MIEX®) resin	*1:*325	
for DOC removal	*1:*325	
in water treatment	*1:*327	
Magnetic meters	*1:*338	
Magnetic water conditioning	*1:*141	*1:*534
applications for	*1:*535	
scientific support for	*1:*536	
Magnetism, relationship to cell stability	*4:*457	
Magnetization, effect on cell biology	*4:*456	
Magnetotelluric (MT) method	*5:*150	
Magnus equation	*4:*270	
Magnus-Teten function	*4:*366	
MAG tests, for bacteria	*1:*84	
Mahakali Treaty of 1996	*4:*683	
Maintenance, of municipal solid waste landfills	*2:*168	
Maize processing plants, anaerobic-aerobic treatment for	*1:*581	
Malachite, adsorption of	*1:*118	
Malaria, DDT use for	*2:*600	
Male reproductive system, effects of lead on	*2:*436	
Mammals		
effects of cadmium on	*5:*615	
marine	*4:*57	
Management		
forest	*3:*172	
municipal watershed	*1:*497	
as a remediation strategy	*2:*358	
of residential water demands	*1:*15	
of water infrastructure	*2:*594	
of waterlogged lands	*3:*603	
of water treatment plant residuals	*1:*411	
Management activities, impacts of	*2:*248	
Management contract model	*1:*50	

Index terms **Links**

Management issues, Great Lakes region 3:179
Management measures (MMs), well head protection 1:527
Management models 2:369 2:620
Manchester, Connecticut, drought-contingency-plan phasing criteria 2:581 2:582t
Manganese (Mn). *See also* Manganese removal; Mn(II)
  arsenic removal using 1:636
  forms of 1:313
  in groundwater 2:147
  in inhibition of biological iron removal 2:153
  microbial oxidation of 2:226
  problems caused by 1:313
Manganese biofouling 5:36
Manganese greensand 1:314
Manganese-reducing to iron-reducing conditions 5:583
Manganese removal 1:312 2:425
  alternative treatments for 1:314
  common treatment processes for 1:313
  using ozone 1:354
Manhole connection repair 1:888
Manhole renovation, using trenchless techniques 1:882
Manholes, repair and renewal of 1:889
Manifold and well points, sand abstraction system 3:416
Manmade lakes 3:270 3:293
Manmade reservoirs 2:367
Manning, John C. 4:454
Manning coefficient values 3:196*t* 3:242 3:244 3:347t
Manning equation 3:196
Manning roughness 3:63 3:64 3:76
Manual valves 1:482
Manufacturing, water use for 2:662
Manufacturing processes, effect on corrosion 1:8
Manure
  composting/exporting 2:572
  contents of 3:538
  efficient use of 3:698
Manure application, nitrate contamination and 2:572
Mapping
  of agricultural land 3:608
  parametric 5:657
''Marble-in-a-cup'' theory 2:50 2:52f

Index terms	**Links**	
Marginal cost pricing, in India	*2:*558	
Margin of safety	*2:*676	
Mariculture, impact of oil on	*4:*106	
Marine animals		
debris hazard to	*4:*40	
mercury levels in	*4:*107	
Marine bacteria, morphological features of	*2:*446	
Marine bioluminescence toxicity tests	*2:*48	
Marine biota, cadmium levels in	*4:*100	
Marine chemistry	*4:*57	
Marine colloids	*4:*27	
physicochemical parameters of	*4:*30	
residence times of	*4:*30	
types of	*4:*28	
Marine debris		
abatement fo	*4:*38	
coral reefs and	*4:*114	
hazard to humans	*4:*40	
hazard to marine animals	*4:*40	
hazard to navigation	*4:*40	
reducing	*4:*41	
sources of	*4:*38	
Marine ecosystems, cultural eutrophication in	*3:*114	
See also Marine environment		
Marine environment. *See also* Marine ecosystems; Marine habitats		
gas hydrates in	*4:*57	
heavy metals in	*4:*100	
monitoring quality of	*2:*443	
monitoring water quality in	*2:*447	
pollutant input into	*4:*96	
Marine habitats. *See also* Marine environment		
effects of pollutants on	*4:*105	
impact of oil on	*4:*106	
Marine Instrumentation Laboratory (MIL), technology and hardware development at	*4:*70	
Marine mammals	*4:*57	
Marine microalgal sediment toxicity tests	*4:*120	
Marine oilers	*4:*763	
Marine oil spills	*2:*44	*2:*291
effects of	*4:*105	
Marine organisms, juvenile	*3:*72	

Index terms	**Links**	
Marine pollutants, classes of	*4:*50	
Marine pollution, sources and effects of	*4:*108	*4:*109t
Marine Protection, Research, and Sanctuaries Act	*2:*125	
Marine seeps, as a CH_4 source	*4:*87	
Marine sediments, heavy metal concentrations in	*2:*448	
Marine Stewardship Council	*2:*636	
Marine stock enhancement techniques	*4:*124	
Marine toxicological testing	*4:*42	
Mariotte bottle		
applications for	*4:*504	
use in hydrology	*4:*503	
Maritime fog	*4:*232	
Maritime history, protection of	*4:*62	
Marketable permits system	*2:*507	
Market prices	*4:*607	
Market-related transaction costs	*2:*500	
Markovian process	*3:*214	*3:*215
MARPOL Convention	*4:*40	
Mars		
life on	*4:*747	
megawatersheds on	*5:*270	
Mars Exploration Rover mission	*4:*504	
Marshes	*1:*404	
coastal	*3:*72	
Marsh vegetation, impact of oil on	*4:*106	
Martin wetland classification system	*3:*497	
Mass conservation equations	*2:*271	
Mass curve analysis	*3:*261	
Mass-flow leaching	*5:*261	
Mass-flow porometers	*3:*717	
Massively parallel signature sequencing (MPSS) technology	*2:*60	
Mass-momentum-energy balances	*5:*558	
Mass movement events, role in flooding	*3:*146	
Mass spectrometry, electrospray ionization	*4:*443	
Mass spectroscopic methods, advanced	*2:*206	
Mass spectroscopy (MS)	*2:*308	*2:*389
Mass transfer rate constant	*4:*568	
Mass transport, in saturated media	*5:*273	
Master base flow recession curve	*3:*25	

Master meters *1:*340
Mastigophora *2:*313
Mathematical modeling. *See also* Mathematical models
  in artificial aquifer recharge *5:*12
  of contaminant transport *5:*517
  for coupled flows *5:*659
  of groundwater flow *5:*516
Mathematical models *2:*333
  See also Mathematical modeling; Modeling; Numerical modeling
  for estimating agricultural area pollutants *3:*658
  leachate metal transport and *1:*725
  role in water storage *2:*368
  sediment transport *3:*421
  in water resources systems
  analysis *2:*686
Mathematical transport modeling approaches *5:*33
Mathews and Weber multicomponent isotherm model *1:*108
  See also Isotherms
Matlab program *1:*732
Matric potential *4:*485 *5:*125 *5:*533
Matric tension *5:*180 *5:*462
Matrix-assisted laser desorption ionization time of flight (MALDI-TOF) *2:*61 *2:*206
Matrix throughflow *3:*453
Maximum acceptable toxicant concentration (MATC) *2:*434
Maximum adsorption QSAR model *3:*99
Maximum basin relief *3:*29
Maximum concentration limit (MCL), of arsenic *2:*15
Maximum contaminant level goals (MCLGs) *1:*422 *1:*425 *1:*427 *1:*477 *2:*195 *2:*268
Maximum contaminant levels/limits (MCLs) *1:*70 *1:*82 *1:*291 *1:*312 *1:*378 *1:*422 *1:*477 *2:*233 *2:*268 *2:*373 *4:*669 *4:*677 *5:*278
  for arsenic *2:*8
  defined *2:*544
  developing *1:*425
  identifying *1:*427
  judicial review of *4:*678
  for perchlorate *2:*345
  water contaminants with *2:*234t
Maximum entropy principle *4:*218

Maximum entropy spectral analysis (MESA) *4:*221
Maximum extent practicable (MEP) pollution control *3:*432
Maximum permissible concentrations *1:*287
Maximum permissible limits *1:*25
Maximum Residual Disinfectant Level (MRDL) *1:*477
Maximum Residual Disinfectant Level Goal (MRDLG) *1:*477
Maximum significant wave height, estimating *4:*2
Maximum sustained wind speed, estimating *4:*2
Maximum vapor pressure *4:*271
McCarran Act *4:*690
MDL QSAR program *3:*99
Mean basin elevation *3:*29
Mean cell residence time (MCRT) *1:*753 *1:*846
Meander wavelength *3:*31 *3:*32
Mean liquor suspended solids (MLSS) *2:*48
Mean residence time *4:*287
Mean sea level (MSL) *4:*118
Measured pollutant loads *3:*223
Measurement programs, optimizing *2:*177
Measurement systems, prototype *4:*296
Mechanical aeration *1:*353 *1:*460
  See also Mechanical surface aeration
Mechanical colloidal processes *5:*350
Mechanical dewatering, of sludge *1:*855
Mechanical dispersion of contaminants *5:*517
Mechanical dredging/barge disposal *2:*122
Mechanical flocculators *1:*253
Mechanical move microirrigation *3:*621
Mechanical seals *1:*882
Mechanical soil conservation methods *3:*550
Mechanical submerged aeration systems *1:*624
Mechanical surface aeration *1:*624
Mechanical treatment, for iron bacteria *2:*152
Mechanical weed control *3:*744
Media filters, for microirrigation *3:*752
Median effective concentration (EC_{50}) *2:*377
Mediated sensors *2:*41
Media thickness *5:*184
Medical science, illogical foundations of *4:*723

Medications, nitrate-containing *2:*220
Medicinal activated carbon *1:*105
Medicinal properties, of Saratoga Springs waters *4:*797
MEDIS project *2:*639
Mediterranean islands, sustainable water management on *2:*638
  See also SUSTAINIS advanced study course
Mediterranean Sea, mercury cycling simulation for *3:*323
Megawatersheds *5:*266
  discovery of *5:*267
  on Mars *5:*270
  Somalia case study of *5:*267
  Trinidad and Tobago case studies of *5:*269
Megawatersheds exploration program *5:*267
Mekong Delta
  ancient settlement of *4:*750
  French colonial conquest of *4:*751
  geology of *4:*749
  subregions of *4:*749
  Vietnamese settlement of *4:*750
Mekong Delta canals, history of *4:*748
Melioidosis *1:*88
Membrane backwashing *1:*334
Membrane-based reverse osmosis *1:*308
Membrane bioreactors (MBRs) *1:*822
Membrane cleaning
  choosing cleaners *1:*419
  control and improvement of *1:*421
  multistage systems for *1:*421
  on-line procedure for *1:*421
  in reverse osmosis *1:*419
  strategies for *1:*420
Membrane filtration *1:*229 *1:*232 *1:*247 *1:*295 *1:*331 *1:*458 *1:*488
  of landfill leachates *1:*707
  for radioactive waste *1:*803
  role of colloids and dissolved organics in *2:*99
  waste stream disposal and *1:*336
Membrane filtration systems, comparing *1:*331 *1:*332f
Membrane foulants, in reverse osmosis *1:*416

Membrane fouling *1:*301
  limiting *1:*594 *2:*364
  mechanisms of *1:*415
Membrane materials *1:*708
  for diffusion-based groundwater sampling *5:*456
  reverse osmosis *5:*26
Membrane ‘‘package’’ plants *1:*334
Membrane performance, in reverse osmosis *1:*173
Membrane processes *1:*792
Membranes. *See also* Membrane cleaning; Membrane filtration
  desalination via *1:*171
  integrity testing of *1:*336
  pretreatment of *1:*334
  reverse osmosis *1:*308 *1:*810
  synthetic *1:*591
  washing and reuse of *1:*920
Membrane separation processes *1:*791
Membrane/sonication/wet oxidation hybrid system *1:*873
Membrane systems *1:*817
  domestic sewage *1:*834
Membrane techniques, for perchlorate removal *5:*633
Membrane water treatment system, biofouling in *2:*242
Menhaden research *4:*56
Meningitis, aseptic *1:*178
Mercury (Hg). *See also* Hg (II) ion sorption
  atmospheric deposition of *3:*282
  fate and transport of *5:*644
  in groundwater *5:*642
  health risks of *5:*643
  maximum contaminant level of *2:*234t
  poisoning by *3:*118
  regulatory guidelines for *5:*645
  remediation of *5:*644
  in urban stormwater *3:*435
Mercury compounds, inorganic and organic *5:*643
Mercury cycling, simulation of *3:*323
Mercury levels
  in marine animals *4:*107
  in marine biota *4:*101
Mercury removal, from complex waste waters *1:*722

Index terms	**Links**	
"Merit goods"	*1*:216	*1*:217
Meroplankton	*4:*155	
Mesmer, Franz, water healing by	*4:*760	
Mesophilic processes	*1:*647	
Mesopores	*1:*97	
Mesopotamia, water symbolism in	*4:*785	
Mesoscale currents	*3:*320	
Mesoscale fog	*4:*230	
Mesozooplankton	*4:*156	
Meta-analysis	*4:*610	
Metabolic heterogeneity, in biofilms	*2:*230	
Metabolic pathways, of contaminants	*5:*529	
Metabolic potential	*2:*42	
Metabolism		
crop	*3:*721	
energy flow and	*5:*528	
Metabolomics	*2:*62	
Metal accumulation, by plant roots	*3:*35	
Metal arsenides	*5:*22	
Metal binding sites, in aquatic biomacromolecules	*4:*29	
Metal biosorption. *See also* Metal sorption		
characteristics of	*2:*70t	
mechanisms of	*2:*70	
Metal cations	*4:*444	
Metal chemistry	*5:*283	
Metal complex formation	*2:*204	
Metal concentration(s)		
in acid mine drainage	*5:*1	
dissolved	*2:*422	
impact on plants	*2:*65	
pervasiveness of	*2:*237	
Metal contaminants, toxicity/mobility reduction of	*5:*434	
Metal-contaminated groundwater, remediation of	*5:*283	
Metal-contaminated soils, remedial approaches for	*5:*433	
Metal detectors	*5:*150	
Metal discharges, regulations concerning	*2:*68	
Metal hyperaccumulation	*5:*371	
Metal hyperaccumulators	*3:*35	
Metal-induced humic agglomeration	*2:*209	
Metal ion adsorption, on bed sediments	*3:*295	

Index terms	**Links**
Metal ion binding, to inorganic-humic agglomerates	*2:*208
Metal ion concentrations, in groundwater	*2:*144
Metal ion-humic colloid interaction	*2:*205
kinetics of	*2:*209
mechanisms of	*2:*207
modeling approaches to	*2:*206
speciation methods related to	*2:*206
Metal ions	
adsorption of	*4:*492
removal by activated carbons	*4:*506
removal from aqueous systems	*1:*918
Metal ion speciation, estimating	*2:*206
See also Metal speciation	
Metallic corrosion	*1:*154
Metallic ion nitrate treatment	*5:*328
Metallothioneins (MTs)	*5:*287
Metallic salt coagulants	*1:*138
Metalloenzymes, microbial	*2:*233
Metallothioneins (MTs)	*3:*609
biological function of	*2:*406
as biomarkers	*2:*407
classes of	*2:*406
factors affecting levels of	*2:*407
as indicators of trace metal pollution	*2:*406
quantification of	*2:*407
role in metal tolerance	*3:*612
Metal mining	
impact on surface water quality	*3:*374
pollutants from	*1:*897
Metal-organic complexes, stability of	*5:*259
Metal-organic interactions	
in landfill-polluted environments	*5:*259
in subtitle D landfill leachates and ground waters	*5:*258
Metal-oxide semiconductor field effect transistor (MOSFET) technology	*2:*298
Metal phytoremediation, selecting plants and microorganisms for	*3:*628
Metal pyrites	*3:*13
Metal remediation	*5:*290
water-jetting drilling technologies for	*5:*234
Metal removal, role of phytoremediation in	*3:*368

Index terms	**Links**	
Metal-resistant bacteria		
distribution of	*2:*448	
in Peter the Great Bay	*2:*447f	
in Rudnaya Bight	*2:*449f	
selection for	*5:*379	
Metal response elements (MREs)	*2:*407	
Metals	*5:*280	
See also Heavy metal entries		
active immobilization of	*5:*282	
in an agricultural landscape	*3:*607	
in ash and FGD sludge ponds	*1:*852	
bioavailability of	*2:*211	
bioremediation of	*5:*282	
chemical binding by colloids	*3:*75	
detoxification in plants	*5:*287	
effect on corrosion	*1:*8	
hyperaccumulation in plants	*5:*285	
immobilization of	*1:*836	*5:*282
in key enzyme functions	*2:*237t	
microbe-induced reduction of	*5:*293	
mobility of	*5:*281	
mobilization of	*5:*377	
in natural waters	*1:*903	
passive immobilization of	*5:*283	
permissible concentrations of	*2:*235t	
plant-based remediation of	*5:*292	
plants that accumulate	*3:*629	
precipitation of	*2:*70	
reduction of	*2:*300	
removal by constructed wetlands	*3:*368	
sample preparation of	*2:*309	
sequestration in vacuoles	*5:*287	
in sewage treatment plant influent	*1:*831	
simulation	*2:*256	
in streambed sediment	*3:*355	*3:*356t
toxicity of	*5:*259	
in whole sculpin tissue	*3:*361t	
Metal sorption	*1:*111	*2:*363
See also Metal biosorption		
Metal speciation		
chemical reactions affecting	*2:*203	

Index terms	**Links**
Metal speciation *(Continued)*	
measurement of	*2:*204
modeling	*2:*204
sensors	*2:*419t
Metal speciation and mobility, landfill disposal practices and	*1:*723
Metal surface complexes	*4:*509
Metal tolerance	
in plants	*3:*609
role of phytochelatins in	*3:*611
Metal treatment	*5:*290
Metal uptake, mechanism of	*5:*286
Metamorphic rock aquifers	*5:*145
Metamorphic rocks	*5:*55
Meta River	*4:*642
Meteoric water line	*5:*230
Meteorological data, in land surface modeling	*3:*536
Meteorological drought index	*4:*211
Meteorological droughts	*4:*208
Meteorological instrumentation	*4:*349
GLERL	*4:*72
Meteorological measurements	*4:*164
Meteorological observation networks	*4:*349
Meteorological observation stations	*4:*72
Meteorologic conditions, heat balance and	*3:*191
Meteorologic fields, relationship to heat fluxes	*3:*191
Meteorologic variables, southwestern Spanish coast	*4:*341
Meteorologists	*4:*328
functions of	*4:*293
Meteorology	*4:*292
ballooning and	*4:*164
ground-based GPS	*4:*244
scientific progress in	*4:*352
Meteors, water in	*4:*190
Metering	*2:*496
See also Water meters	
in Canada	*2:*663
as a water conservation measure	*1:*146
Meter yokes	*1:*339
Metha, Gita	*4:*768

- Methane. *See also* CH_4; Halogenated methanes
  - in groundwater *5*:293
  - migration of *5*:256
- Methane hydrate *4*:473 *4*:474
- Methanogenesis
  - in landfills *1*:696
  - separation from acidogenesis *1*:909
- Methanogenic metabolism *1*:689
- Methanogenic systems *5*:47
- Methanotrophic bacterial systems *5*:47
- Methemoglobinemia *2*:219 *4*:519 *5*:324
  - diagnosis and treatment of *2*:221
  - hereditary causes of *2*:220
  - nitrate-related *1*:35
  - susceptibility to *2*:219
  - water-related and non-water-related causes of *2*:220
- Methylated antimony species *4*:592
- Methylated arsenic compounds *2*:11
- Methylation, of metals *5*:378
- Methyl bromide
  - cycling study *4*:80
  - study of sources of *4*:150
- Methylene blue number *1*:98
- Methyl halides, measurements of *4*:150
- Methylmercury exposure *5*:643 *5*:644
- Methyl-*tertiary*-butyl ether (MTBE) *2*:312 *2*:317 *2*:344 *5*:318
  - aerobic cometabolism of *5*:388
  - alfalfa channel tests and *5*:392
  - anaerobic microbial transformation of *5*:386
  - chemical and physical properties of *5*:388t
  - degradation pathway of *5*:387f
  - electrolysis of *5*:122
  - fate in the atmosphere *5*:393
  - field studies involving *5*:394
  - *in situ* electrokinetic treatment of *5*:116
  - phytoremediation of *5*:385
  - plumes *5*:80
  - potential biotreatments for *5*:389
  - properties of *5*:386
  - soil vapor concentration of *5*:553
  - uptake by hybrid *Populus* *5*:389

Methyl-*tertiary*-butyl ether (MTBE) *(Continued)*
  uptake efficiency and toxicity for *5:*391
Metropolitan Washington Council of Governments (MWCOG) "simple method" *2:*253
Mexican National Water Commission (CNA) *1:*652
Mexico
  coastal fog along the northern Gulf of *4:*230
  wastewater applications in *1:*633 *3:*672
Mexico-Guatemala water conflicts *4:*641
Mexico-U.S., transboundary waters in *4:*636
  See also United States-Mexico border waters
Mg/Ca ratio *4:*93
  See also Magnesium hardness
Microalgal sediment toxicity tests, marineand estuarine *4:*120
Microarray analysis *2:*60
Microarray experiments *2:*62
Microarray Gene Expression Data (MGED) Society *2:*60
Microarray (DNA chip) technology *1:*161
  See also DNA microarrays
Microbes. *See also* Bacteria; Marine bacteria; Microorganisms
  biochemical oxygen demand and *1:*640
  in drinking water *2:*243
  plants as energy sources for *5:*378
Microbe transport models *1:*368
Microbial activity
  aerobic tank *1:*789 *1:*790f
  in ''dirty water'' formation *2:*226
  disinfection to control *2:*225
  effect of temperature on *2:*226
  in groundwater *5:*32
Microbial activity management *2:*223
  biofilm control *2:*223
  control of biodegradable organic carbon *2:*224
Microbial aminolevulinate dehydratase (ALAD) *2:*236
Microbial assessments *2:*450
  historical background of *2:*441
  for monitoring marine environment water quality *2:*447
Microbial biofouling *1:*83
  See also Microbial fouling
  direct analysis of *1:*84
Microbial biomass, as a sorbent *2:*365

Index terms	Links		
Microbial biosurveys			
in Far Eastern Russia seas	*2*:444		
for heavy metal detection	*2*:442	*2*:445	*2*:446
for organic pollution assessments	*2*:449		
Microbial BOD biosensors	*2*:40		
See also Biochemical oxygen demand (BOD)			
Microbial communities	*3*:306		
Microbial contaminants, health effects of	*1*:277		
Microbial corrosion	*1*:6		
Microbial detection, of phenols	*2*:442		
Microbial/direct chemical contamination zone (Zone A2)	*1*:526		
Microbial diversity	*2*:42	*4*:746	
Microbial dynamics, of biofilms	*2*:228		
Microbial enzyme assays, for heavy metal toxicity detection	*2*:233		
Microbial foaming and bulking			
in the activated sludge process	*1*:728	*1*:844	
bacteria involved in	*1*:846		
factors affecting	*1*:847		
prevention and control of	*1*:846		
problems caused by	*1*:845		
Microbial forms, in biofouling events	*2*:239		
Microbial fouling	*5*:422		
See also Microbial biofouling			
Microbial genes, using in selenium phytoremediation	*5*:400		
Microbial growth, corrosion and	*1*:544		
Microbial loop	*3*:312	*4*:152	
Microbially available phosphorus (MAP)	*2*:245		
Microbial metabolism	*2*:299		
Microbial metalloenzymes	*2*:233		
Microbial motility	*2*:228		
Microbial nutrients, in distribution systems	*2*:244		
Microbial pathogens	*2*:371		
in aquatic systems	*2*:352		
Microbial physiology	*5*:580		
Microbial pollutant detection, water quality monitoring and	*2*:440		
Microbial processes			
in biofilms	*3*:307		
in mine effluent remediation	*1*:612		
Microbial products, algal control and	*2*:6		
Microbial quality, of reclaimed irrigation	*3*:667		

Index terms **Links**

Microbial structure, in biofilms *2:*230
Microbial transformation
  of chlorinated aliphatic compounds *1:*690
  of chlorinated aromatic compounds *1:*689
  of petroleum hydrocarbons *1:*692
Microbiological analyses, of rural drinking water *1:*383
Microbiological behavior *1:*566
Microbiologically induced corrosion (MIC) *1:*596 *1:*600
Microbiological mechanisms *1:*596
Microbiological methods, in groundwater studies *5:*464
Microbiological monitoring *2:*267
Microbiological processes, soil *3:*706
Microbiological quality control, in distribution systems *2:*243
Microbiological standards, U.S. *2:*268
Microbiology
  iron and *5:*609
  of lotic aggregates and biofilms *3:*305
  soil *5:*464
in transport modeling *5:*31
Microcalibration *3:*319
Microcolumns (MIC) *2:*212
Microcontrollers *4:*71
Microcoulometric detectors *2:*308
Microcystins *2:*388 *2:*390 *3:*189
Microencapsulation *1:*835
Microfiltration (MF) *1:*331 *1:336t* *1:*444 *1:*459 *1:*488 *1:*591 *1:*810
  of industrial wastewaters *1:*591
Microfiltration and reverse osmosis (MF-RO) *1:*218 *1:*220
Microflora, lipolytic *2:*442
Microgravity networking *5:*411
Microhabitats *3:*70
Microirrigation *3:*615
  advantages and disadvantages of *3:*617 *3:*621
  components and system design for *3:*622
  current irrigated area using *3:*619
  development of *3:*620
  efficiency of *3:*620
  expansion of *3:*627
  extent of *3:*624

Index terms **Links**

Microirrigation *(Continued)*
  future scope of *3:*627
  limitations of *3:*626
  media filters for *3:*752
  research trends in *3:*626
  screen filters for *3:*748
  strategies for promoting *3:*627
  types of *3:*620
  water saving and yield increase related to *3:*617 *3:*622
Microirrigation systems, control station of *3:*617
MicroMac-ToxScreen *2:*174
Microorganism contaminants *1:*54
Microorganisms. *See also* Bacteria; Lotic aggregates; Microbes; Microbial entries; Single-celled organisms; Protozoa; Viruses
  abiotic factors related to *3:*311
  adhesion of *2:*239
  biotic interactions of *3:*312
  deactivation of *1:*605f
  dew use by *4:*201 *4:*205
  genetically modified *4:*525
  geochemical ecology of *2:*442
  in geothermal waters *5:*157
  habitats of *3:*310
  as indicators of oil pollution *2:*441
  inside biofouling layers *2:*240
  iron distribution and *4:*497
  metabolic diversity of *2:*44
  monitoring *3:*669
  in natural environment *3:*309
  oil-destroying *2:*449
  persistence data for *1:*522
  in raw and treated wastewater *3:*670t
  role in biodegradation *2:*42
  role in water vectored diseases *2:*112
  selecting for hydrocarbon and metal phytoremediation *3:*628
  selective binding and absorption of metals by *2:*443
  types of *3:*668
Microphytobenthos *4:*120
  bioassays involving *4:*122
  removal from sediments *4:*121
Micropollutants, organic *4:*506

Micropores *1:*97 *3:*474
Micropurging *5:*455
Microscale currents *3:*320
Microscale flow manifolds *4:*515
Microscale fog *4:*231
Microscale toxicity testing, in natural systems *2:*379
Microscopic corrosion measurement *1:*9
Microscopic examination/analysis, of biofouling structures *1:*84
Microscreen filters *1:*683
Microspora *1:*279
Microsporidiosis *1:*367
Microsporidium, in karstic aquifers *1:*367
Microstrainers *1:*486
Microthrix parvicella *1:*729
MICROTOX® bioassay *2:*171 *2:*172 *2:*173 *2:*278 *2:*351 *2:*413 *2:*453
Microtox-OS Test System *2:*174
Microwave precipitation estimation *4:*323
Microwave soil moisture equations *4:*324
Midchannel pools *3:*68
Middle Ages
  sanitation during *1:*282
  water in *4:*728
Middle East
  conflict and water use in *4:*753
  international river basins in *4:*754t
Midoceanic ridge basalts (MORB) *4:*577
Midrange water quality models *2:*248
MIEX® resin plant *1:*326f
  See also Magnetic ion exchange (MIEX®) resin
  capital and operating costs of *1:*329
Military
  use of activated carbon in *1:*105
  role in prototype measurement systems *4:*296
Military applications, research for *4:*295
Milled screens *5:*573
Mine drainage
  acidification of surface waters by *3:*10
  composition of *1:*609t
"Mineral (nitrogen) balance" *3:*639

Index terms	Links	
Mineral domain, sorption in	*4:*386	
Mineralization	*2:*42	
of heavy metals	*5:*378	
Mineral matrix	*5:*172	
Mineralogy, evaporite	*5:*52	
Minerals, in soil	*3:*706	
Mineral water, historical uses of	*1:*4	
Mines, abandoned	*5:*3	
Mine waste	*1:*609	
acid mine drainage generation	*1:*609	
effluent remediation and	*1:*610	
Mine waters, discharge standards for	*2:*425	
Minicolumn method	*1:*774	*1:*779
Minimum cross entropy principle	*4:*219	
Minimum environmental flow regimes	*3:*166	
Minimum Information About a Microarray Experiment (MIAME)	*2:*60	
Minimum National Standards (MINAS)	*2:*183	
Mining		
groundwater acidification due to	*3:*10	
impact on surface water quality	*3:*374	
ocean water contamination by	*4:*100	
Mining lakes, acidic	*3:*10	
Mining water	*1:*552	
Mining water use	*2:*544	*2:*662
in the United States	*2:*652	
Ministry of Agriculture (China)	*2:*486	
Ministry of Construction (China)	*2:*486	
Ministry of Geology and Mineral Resources (China)	*2:*486	
Ministry of Public Health (China)	*2:*486	
Ministry of Transportation (China)	*2:*486	
Ministry of Water Resources (China)	*2:*485	*2:*488
MINTEQA2/PRODEFA2 program	*5:*141	
MINTEQ database	*5:*144	
Minute process	*4:*644	
Mirror carp	*4:*719	
Mississippi River Commission (MRC)	*4:*613	
Mississippi River water quality report	*1:*289	
Mixed clathrate hydrates	*4:*471	
Mixed-flow impellers	*1:*76	*3:*664

Index terms	**Links**	
Mixed liquor suspended solids (MLSS)	*1:*730	*1:*828
Mixing		
lake	*3:*266	
in water treatment systems	*1:*76	
Mixing fog	*4:*231	
Mixing ratio (*r*)	*4:*273	
Mixing zones	*3:*80	
Mixing zones ban, Great Lakes system	*4:*624	
Mixotrophic organisms	*4:*152	
Mn(II)	*5:*585	
See also Manganese entries		
MNA monitoring well network	*5:*587	
See also Monitored natural attenuation (MNA)		
Model calibration, alternatives to	*3:*342	
Modeling. *See also* Activated Sludge Models (ASMs); Biological process modeling; Chlorine residual modeling; Empirical models; Mathematical models; Models; Residential water use models; Steady-state modeling; Wastewater modeling and treatment plant design		
adsorption	*3:*300	
COD relationships in	*1:*735	
of contaminant transport and biodegradation in groundwater	*5:*30	
diffusion kinetics	*4:*567	
entropy as a tool for	*4:*220	
of episodic acidification	*3:*5	
of fecal coliform in sewers	*3:*334	
finite element	*5:*655	
groundwater	*5:*619	
of hurricanes	*4:*274	
hydrogen sulfide	*3:*333	
infiltration	*4:*486	
integrated	*3:*340	*3:*341
karst	*5:*241	
kinetic	*4:*565	
land surface	*3:*533	
of nonconservative pollution	*3:*332	
ocean	*4:*22	
of pond aquaculture	*3:*375	
predictive	*3:*40	*5:*589
of quasi-conservative pollution	*3:*332	
rainfall-runoff	*4:*297	*4:*357

Index terms **Links**

Modeling *(Continued)*
  real-time *3:*340
  of salt balance *3:*680
  sedimentation *3:*404 *3:*421
  of sewer water quality *3:*331
  of solute transport in groundwater *5:*305
  subsurface redox chemistry *5:*413
  of transpiration *3:*736
  of ungauged watersheds *3:*342
  unit hydrograph *4:*357
  of unsaturated soil properties *5:*538
  of urban drainage and stormwater *3:*337
  of urban flooding *3:*339
  water quality *3:*338 *3:*341
  watershed *3:*327
Modeling programs *4:*323
Model parameter uncertainty *3:*325
Models. *See also* Mathematical models; Modeling
  calibration methods for *2:*332 *5:*297
  comprehensive *2:*255
  geochemical *5:*138
  verification/validation of *5:*298
Model sensitivity analysis *3:*318
Model structure (equation) uncertainty *3:*325
"Model Water Use Act" *2:*580 *2:*583
Moderate Resolution Imaging Spectroradiometer (MODIS) *2:*589 *4:*194 *4:*216 *4:*351
  measurements with *4:*134
"Modern biomass" *3:*546
Modern groundwaters, detecting with ^{85}Kr *5:*248
MODFLOW *1:*368 *5:*564
  packages *5:*33
  simulation *5:*84f
MODFLOW-RT3D simulator *5:*520
Modified chlorophyll absorption in reflectance index (MCARI) *3:*722
MODSIM water quantity network flow allocation model *2:*327
Modular groundwater flow (MODFLOW) code *5:*83
  See also MODFLOW entries
Moffett Federal Airfield permeable reactive barrier site *5:*520 *5:*521
Moist adiabatic lapse rate (MALR) *4:*369

Index terms	Links		
Moisture, influence on nitrification	*3:*641		
See also Soil moisture measurement			
Moisture characteristic curves	*5*:462	*5*:472	
Moisture distribution, using microirrigation	*3:*618		
Moisture enhancement	*5:*117		
Mojave Desert, recharge in	*5:*73		
Molasses, composition of	*1:*616t		
Molasses number	*1:*98		
Molecular assays, for *Escherichia coli* O157:H7	*2:*139		
Molecular-based detection, of *Cryptosporidium parvum*	*1:*158		
Molecular-biological tools	*1:*847		
for biodiversity monitoring	*1:*642		
community profiling methods	*1:*643		
labeled substrate probes	*1:*645		
molecular probes	*1:*643		
Molecular-biology-based assays	*1:*91		
Molecular diffusion	*5:*174	*5:*544	
Molecular fingerprints	*1:*643		
Molecular hydrogen	*4:*480		
See also H_2 economy			
Molecular methods	*2:*352		
Molecular oxygen	*4:*535		
See also O_2 evolution			
Molecular probes, in community analysis	*1:*643		
Molecular water network, changes in	*4:*513		
Molecular weight cutoff (MWCO)	*1*:331	*1*:334	*1*:335
Mole drains	*5:*95		
Molluscs. *See* Bivalve molluscs; Mussel entries; Shellfish growing; Zebra mussel control			
Momentum balances	*5:*558		
Momentum-mass-energy balances	*5:*558		
Monitored natural attenuation (MNA)	*5*:319	*5:*578	
See also MNA monitoring well network			
bioremediation technology	*5:*40		
degradation rate constant	*5:*588		
modeling as a predictive tool for	*5:*589		
model of	*5:*579		
statistical validity of data for	*5:*588		

Monitoring
  of contaminants *2:*263
  of municipal solid waste landfills *2:*168
  use of semipermeable membrane devices for *5:*672
Monitoring parameters, criteria for *2:*266
Monochlorophenols *1:*689
Monoclonal antibodies *1:*160
Monod degradation kinetics *5:*32
Monod equation *1:*735
Monod rate expression *5:*416
Monohalogenated haloacetic acids *2:*75
Monomictic lakes *3:*266
Monsoons *4:*304
  African *4:*304
  Indian *2:*560
Montana, in-stream flow protection in *4:*662
Monte Carlo simulations *3:*262 *3:*263
Monticello permeable reactive barrier site *5:*520
Moody diagram *3:*197
Morbidity, from water related diseases *1:*24
Morphological changes, studies of *2:*65
Mortality, from water-related diseases *1:*24
MOSQITO model *3:*339
Mosses, trace elements in *3:*456
Most probable number polymerase chain reaction (MPN-PCR) *1:*161
Most probable number (MPN) technique *2:*337
‘‘Most sensitive species’’ assay *2:*376
Mother-daughter (shedding) cell formation *2:*239
Motion, equation of *5:*559 *5:*651
Mouse bioassay *2:*389
MOUSE TRAP model *3:*334 *3:*339
Moving bed system *1:*104
Moving water, flow (fall) of *2:*554
MTT assay *2:*415
Mud Dump Site (MDS) *4:*77 *4:*79
Mulberry tree dikes *4:*697
Multiangle Imaging SpectroRadiometer (MISR) satellite sensor *5:*479
Multicomponent kinetics *1:*119
Multiconcentration tests *2:*382

Index terms **Links**

Multidimensional pump test analysis *5:*186
Multidrug-resistant tuberculosis (MDRTB) *1:*182
Multi-jet meters *1:*338
Multimedia mitigation (MMM) programs *4:*546
Multimetric biotic integrity indexes *3:*38
Multimetric periphyton index of biotic integrity *3:*40
''Multiobjective evaluation'' concept *2:*595
Multiphase flow and heat transfer *4:*449
"Multiple-barrier concept" *1:*341
''Multiple convergent datasets'' method *5:*270
Multiple effect distillation (MED) *1:*170
Multiple-hearth furnace (MHF) *1:*858
Multiple stage flash (MSF) distillation units *1:*308
Multiple tray aeration *1:*353 *1:*460
Multipurpose reservoirs *3:*382
  conflicts in *3:*384
  flood control storage capacity in *3:*384
  operation of *3:*384
  storage in *3:*383
  systems engineering for managing *3:*386
Multirate meters *1:*493
Multireservoir system, operation of *3:*385
Multi-Sector General Permit (MSGP) *4:*658
Multisite generalized D'Arcy and Watt (MSGDW) model *4:*403
Multispecies Freshwater Biomonitor (MFB) *2:*31
Multistage drinking water filtration *1:*237
Multistage filtration (MSF) *1:*238
  cost of *1:*242
  criteria selection in *1:*239
  performance of *1:*241
  roughing filter components *1:*240
Multistage flash distillation *1:*170
Multivariate periodic autoregressive order 1 [PAR(1)] model *3:*425
Municipal and industrial needs (MAIN) model *1:*15
Municipal landfills, groundwater contamination from *5:*253
Municipal multiutilities *1:*390
Municipal Separate Storm Sewer System (MS4) *3:*392
Municipal sewage *1:*828
Municipal solid waste (MSW) *1:*695

Index terms	**Links**	
Municipal solid waste landfills	*2:*163	
See also Landfill entries		
cover for	*2:*164	
dry tomb	*2:*163	
groundwater monitoring for	*2:*165	
improving	*2:*167	
landfill gas from	*2:*165	
leachate collection and removal from	*2:*164	
postclosure monitoring and maintenance of	*2:*166	
reuse, recycling, and reduction related to	*2:*167	
siting issues related to	*2:*167	
Municipal Storm Sewer Systems (MS4)	*4:*659	
Municipal storm water management	*1:*866	
evaluating	*1:*868	
monitoring in	*1:*870	
public participation in	*1:*867	
Municipal wastewaters	*1:*678	
analysis of	*1:*829t	
chemical precipitation of	*4:*586	
organic compounds in	*1:*766	
Municipal water cycle	*1:*495	*1:*496f
Municipal watersheds	*1:*495	
impacts on	*1:*495	
management of	*1:*497	
types of	*1:*495	
Municipal water supplies		
disinfection byproducts in	*1:*264	
ozonation of	*1:*362	
in the United States	*2:*651	
Municipal water supply and sewerage		
corporate strategies in	*1:*389	
private sector participation in	*1:*387	
Municipal water supply system	*2:*371	*2:*544
Municipal water use	*2:*662	
Munsell Rock Color Chart	*3:*690	
Mupfure catchment	*3:*237	
Murray River, salinity profile of	*3:*680	
Musculista senhousia	*4:*11	*4:*75
excretion rates of	*4:*7	*4:*12
Musical glasses	*4:*758	
Muskingum-Cunge flow routing method	*3:*256	

Next Page

Index terms **Links**

Muskingum flow routing method *3*:255

Mussels
  chemical-free control of *1*:510
  dose-response to chlorine *2*:401
  industrial field studies on *1*:512
  on surfaces *1*:541

"Mussel Watch Program" *2*:35

Mutagenic agents, luminescent bacteria for determining *2*:174

Mutagenicity assay *2*:175

Mutatox assay *2*:171 *2*:172 *2*:174 *2*:455

MX [3-chloro-4-(dichloromethyl)-5-hydroxy-2(5H)- furanone], health effects of *1*:269

Mycobacterium *1*:278

Mylar balloons *4*:165

Mytilus californicus *4*:153

Mytilus galloprovincialis, as a bioindicator *2*:407

N

N_2O, global role and sources of *4*:86

NADH, monitoring *2*:376
  See also Nicotinamide-adenine dinucleotide (NAD^+)

Naegleria *1*:279

Nalgonda defluoridation process *4*:437

Nannoplankton *4*:93

Nanofiltration (NF) *1*:333*t* *1*:335 *1*:336t *1*:458 *1*:488 *1*:828
  effect of *2*:77

N aquifer (Black Mesa, Arizona) *5*:48

NASA Aqua mission, water cycle and *4*:193
  See also Aqua earth observing system (EOS); National Aeronautics and Space Administration (NASA)

Nash model *4*:358

Nashua River Watershed *1*:499

National Academy of Science (NAS)
  Committee on Mitigating Wetlands Loss *4*:693
  risk assessment paradigm *1*:423

National Acid Precipitation Assessment Program (NAPAP) *4*:379

National Aeronautics and Space Administration (NASA), water-flow satellite data from *2*:587
  See also NASA Aqua mission

National agricultural land use planning program *3*:596

Previous Page

Index terms **Links**

Muskingum flow routing method 3:255
Mussels
  chemical-free control of 1:510
  dose-response to chlorine 2:401
  industrial field studies on 1:512
  on surfaces 1:541
"Mussel Watch Program" 2:35
Mutagenic agents, luminescent bacteria for determining 2:174
Mutagenicity assay 2:175
Mutatox assay 2:171 2:172 2:174 2:455
MX [3-chloro-4-(dichloromethyl)-5-hydroxy-2(5H)- furanone], health effects of 1:269
Mycobacterium 1:278
Mylar balloons 4:165
Mytilus californicus 4:153
Mytilus galloprovincialis, as a bioindicator 2:407

N

N_2O, global role and sources of 4:86
NADH, monitoring 2:376
  See also Nicotinamide-adenine dinucleotide (NAD^+)
Naegleria 1:279
Nalgonda defluoridation process 4:437
Nannoplankton 4:93
Nanofiltration (NF) 1:333t 1:335 1:336t 1:458 1:488 1:828
  effect of 2:77
N aquifer (Black Mesa, Arizona) 5:48
NASA Aqua mission, water cycle and 4:193
  See also Aqua earth observing system (EOS); National Aeronautics and Space Administration (NASA)
Nash model 4:358
Nashua River Watershed 1:499
National Academy of Science (NAS)
  Committee on Mitigating Wetlands Loss 4:693
  risk assessment paradigm 1:423
National Acid Precipitation Assessment Program (NAPAP) 4:379
National Aeronautics and Space Administration (NASA), water-flow satellite data from 2:587
  See also NASA Aqua mission
National agricultural land use planning program 3:596

Index terms	**Links**			
National Aquaculture Act	*3:*540			
National Drinking Water Advisory Council (NDWAC)	*1:*425			
National Drinking Water Clearinghouse (NDWC)	*1:*465			
Tech Briefs	*1:*151	*1:*158	*1:*177	*1:*207
National Drinking Water Regulations (NDWRs)	*1:*480			
National Environmental Policy Act	*2:*594			
National Environmental Protection Agency (China)	*2:*486			
National Estuary Program (NEP)	*2:*185			
National fishing waters	*3:*532			
National Flood Insurance Program (NFIP)	*3:*147	*3:*528		
National groundwater nitrate vulnerability assessment	*5:*565			
National Infrastructure Protection Center (NIPC)	*1:*451			
Nationalist language	*4:*681			
National Marine Debris Monitoring Program (NMDMP)	*4:*39	*4:*116		
National Oceanic and Atmospheric Administration (NOAA)	*2:*185	*4:*329		
See also NOAA entries				
National Ocean Service (NOS) sanctuaries	*4:*62			
National Park Service (NPS)	*4:*687			
National Policy on Water Management	*2:*526			
National Pollutant Discharge Elimination System (NPDES)	*1:*756	*1:*798	*1:*866	*2:*96
	*2:*315	*3:*432	*4:*595	*4:*655
	*4:*669			
See also NPDES permits				
pollution control policy of	*2:*127	*2:*128		
water quality-based limitations and	*1:*759			
discharge limitations under	*4:*656			
statutes related to	*4:*665			
National Pretreatment Program (EPA)	*1:*798			
National Primary Drinking Water Regulations (NPDWRs)	*1:*379	*1:*422	*1:*425	*1:*428
National primary drinking water standards	*2:*373t			
National Research Council (NRC)	*4*:541	*4*:544	*4*:599	
National Secondary Drinking Water Regulations (NSDWRs)	*4:*678			
National Sedimentation Laboratory	*3:*89			
National Shellfish Sanitation Program (NSSP)	*2:*361			
National Stormwater Program	*4:*596			
National Toxics Rule	*4:*599			
National Tsunami Hazard Mitigation Program	*4:*89			
National Water Carrier	*4:*699	*4:*700		

National Water Commission (NWC) *4:*614
National water policies, India *2:*556
National Water-Quality Assessment (NAWQA) program *2:*311 *5:*359
National Water Quality Inventory Report *5:*332
National Weather Service *4:*329
  Climate Prediction Center (CPC) *4:*335
National Wetlands Coalition (NWC) *4:*692
National Wetlands Inventory (NWI) *3:*72 *4:*654
National Wildlife Refuge *4:*690
Nationwide Urban Runoff Program (NURP) *3:*433
‘‘Native’’ aquifers *5:*44
NATSGO database *5:*302
Natural aeration *1:*623
Natural attenuation processes *5:*424 *5:*443
  See also Monitored natural attenuation (MNA); Natural contaminant attenuation
  effect of pump and treat remediation *5:*45
  groundwater remediation by no action *5:*44
  methanogenic systems *5:*47
  phytoremediation enhancement of *5:*374
  transverse dispersion *5:*46
Natural bioremediation, of selenium *2:*358
  See also Bioremediation
Natural coagulants *4:*426
Natural contaminant attenuation
  factors affecting the demonstration of *5:*586
  microbial processes affecting *5:*578
Natural contaminants *2:*181
Natural dams, as a cause of flooding *3:*511
Natural disasters
  economic losses due to *3:*153 *3:*154t
  number and economic losses of *3:*153t
Natural drainage networks *3:*94
Natural environments
  clathrate hydrates in *4:*473
  microorganisms in *3:*309
Natural gamma logs *5:*152
Natural gas hydrates *4:*474
Natural groundwater tracers *5:*502
Natural hazards, floods as *3:*153
Natural heritage, protecting *2:*667

Naturally occurring radioactive materials (NORM) *2:*287
Naturally soft water, availability of *4:*554
Natural organic material/matter (NOM) *1:*248 *1:*335 *1:*357 *1:*358 *1:*370 *2:*74 *2:*75
  reaction of ozone with *1:*359
  removal of *2:*116
Natural phenomena, pollution caused by *3:*223
Natural resources. *See also* Natural water resource systems
  monitoring integrity of *2:*440
  sustainable management of *2:*633
  types of *2:*634
Natural Resources Conservation Service *2:*513
Natural Resources Inventory (NRI) *3:*72 *3:*73
Natural sources, of runoff pollution *3:*225
Natural retention *1:*404
Natural systems
  biochemical oxygen demand in *1:*639
  microscale toxicity testing in *2:*379
Natural wastewater treatment *1:*840
  systems for *1:*678
Natural water resource systems *2:*586
Natural waters
  antimony speciation in *4:*592
  arsenic in *1:*81
  carbonate in *4:*413
  cause of alkalinity in *2:*394
  classes of organic compounds in *5:*339t
  disinfection byproduct precursor removal from *2:*115
  pollution of *3:*444
  silica in *4:*548
  sodium in *4:*551
  trace metals in *2:*202
Nature Conservancy, Arizona Chapter, The (TNCA) *3:*57 *3:*58
Nature-like fish passage facilities *3:*531 *3:*532f
Navier-Stokes equations *5:*560 *5:*561 *5:*652 *5:*656 *5:*659
Navigation, role of sediment in *3:*508
Near-infrared (NIR) domain *3:*719
Negev Desert, dew amounts in *4:*203
Negombo Estuary, Sri Lanka *4:*53

Index terms **Links**

Negotiations, authority-polluter *4:*647

Neoprene balloons *4:*165

Nephelometric turbidity units (NTUs) *1:*236 *1:*599 *2:*544

Nernst equation *1:*399

Net ecological benefit concept *2:*477

Net head *3:*201

Netherlands, flood control history in *2:*524

Net pen aquaculture systems *3:*543

Net radiation *3:*563

Netting systems, for floatables control *1:*785

Neumann boundary conditions *5:*14 *5:*561

Neurological effects, of lead *2:*435

Neurotoxins *2:*388 *3:*189

Neutralization facilities *1:*610

Neutral water *4:*481

Neutrino detector, heavy water as *4:*464

Neutron density logs *5:*153

Neutron moderator, heavy water as *4:*463

Neutron moisture probe *5:*539
  moisture measurement with *3:*692

Neutron thermalization soil moisture measurement *4:*488

Nevada
  in-stream flow protection in *4:*662
  overdraft in *5:*343

New development, management of *1:*867

Newfoundland, aquatic habitat and construction in *3:*509

New Madrid, Missouri earthquakes *5:*477 *5:*479

New Mexico, wastewater applications in *1:*634

New Jersey, source water protection in *2:*312

New Source Performance Standards (NSPS) *4:*656

Newton, Isaac *4:*352

Newtonian fluids *5:*556

Newton-Raphson method *5:*142

Newton’s law of viscosity *5:*555

New York, extraterritorial land use control in *1:*316

New York City Harbor Survey *1:*745
  physical setting of *1:*746
  public policy and *1:*750
  sewage treatment infrastructure and programs *1:*748

Index terms	**Links**		
New York City Harbor Survey *(Continued)*			
sociological setting of	*1:*746		
water quality trends	*1:*748		
New York City Historic Area Remediation Site (HARS), mapping the sea floor of	*4:*77		
New York State Department of Environmental Conservation (NYSDEC) criteria	*3:*351		
NEXRAD observations	*4:*307	*4:*309	*4:*310
Nhieu Loc-Thi Nghe Project	*2:*553		
Niagara Falls	*2:*554		
dewatering of	*3:*345		
NICA-Donnan model	*2:*206		
Niche portals	*2:*669		
Nickel (Ni), in groundwater	*2:*147		
Nickpoints	*3:*184		
Nickzones	*3:*184		
Nicotinamide-adenine dinucleotide (NAD^+)	*2:*42		
See also NADH			
Nier/McKinney mass spectrometer	*5:*228		
Nile Basin			
solutions to water conflicts in	*2:*593		
transboundary water conflicts in	*2:*590		
water allocation in	*2:*591		
water scarcity in	*2:*590		
water-sharing framework for	*2:*592		
Nile Basin Initiative (NBI)	*2:*592		
Nile civilization	*4:*727		
Nile River basin	*4:*756		
Nile River, salinity profile of	*3:*680		
Nile valley, recent development of	*4:*729		
Nile Valley Plan	*2:*591		
Nile waters, management of	*4:*756		
Nile Waters Agreement	*2:*591		
Nile Waters Treaty	*4:*757		
Nitrate(s). *See also* NO_3^-			
acidification and	*3:*5		
anoxic biodegradation of	*5:*325		
diseases related to	*1:*34		
in drinking water	*2:*219	*2:*395	
fate and transport of	*5:*629		
groundwater and food sources of	*1:*33	*5:*628	
health effects/risks of	*1:*30	*5:*324	*5:*629

Nitrate(s) *(Continued)*
  history of *5:*628
  innovative *in situ* treatment for *5:*328
  ions *1:*668
  in ponds *3:*486
  population exposed to *1:*34t
  properties of *5:*628
  regulation of *5:*629
  remediation of *5:*630
  reverse osmosis treatment of *5:*326
  toxicity of *4:*519
  treatment for *5:*323
  in water and foods *1:*31
  water use criteria for *5:*324
Nitrate concentrations *5:*252
Nitrate-contaminated agricultural runoff, treatment of *5:*329
Nitrate contamination. *See also* Nitrate pollution
  of groundwater *2:*570 *5:*322
  investigating *5:*630
  maximum level of *2:*234*t* *2:*235t
Nitrate leaching *2:*200
Nitrate Leaching and Economic Analysis Package (NLEAP) *2:*252
Nitrate levels
  high *5:*407
  water quality and *1:*32
Nitrate-nitrite poisoning *1:*38
Nitrate pollution. *See also* Nitrate contamination
  prevention of *3:*637 *5:*325
  sources of *5:*323
Nitrate-reducing to manganese-reducing conditions *5:*583
Nitrifer studies *2:*231
Nitrification *1:*789 *1:*833 *3:*640 *4:*518
  See also Nitritification
  in the activated sludge process *1:*751
  in distribution systems *2:*226
  factors influencing *3:*641
  forms of *1:*752 *1:*753t
  management of *3:*642
  oxygen consumption during *1:*754t
  of potable water *1:*346
Nitrification/denitrification biofilm *2:*231

Nitrifiers, common species of *3:*640t
Nitrifying trickling filters *1:*346
  effect of support material on *1:*347
  influence of recirculation on *1:*347
  modeling and design of *1:*348
Nitrite
  health effects of *5:*324
  maximum contaminant level of *2:*235t
  in ponds *3:*486
  water use criteria for *5:*324
Nitrite ions, accumulation of *1:*752t
  See also Nitrate-nitrite poisoning
Nitrite-oxidizing bacteria (NOB) *2:*231
Nitrite pollution, preventive measures for *5:*325
Nitritification, dynamic model of *1:*349
Nitrobenzene *4:*510
Nitrogen (N) *4:*517
  See also Ammonia nitrogen; C:N ratio; Nonconservative dissolved inorganic P and N fluxes; N/P (nitrogen/phosphorus) ratio; Organic nitrogen
  in an agricultural landscape *3:*604
  algal growth and *3:*109
  ammonification of *4:*517
  in animal farming operations *3:*538
  aquatic *4:*518
  in constructed wetlands *1:*893
  in domestic sewage *1:*831
  in domestic wastewaters *3:*110
  fixation of *4:*517
  forms of *2:*269
  fractionation of *4:*502
  isotopes of *4:*519
  maximum contaminant level of 2:234t *2:*235t
  as a nutrient of concern *2:*269
  retention of *2:*200
  sources of *3:*1 *3:*695
  tile drainage and *3:*729
  water contamination by *4:*519
Nitrogen availability reporter system *2:*175
Nitrogen balance *3:*639
Nitrogen best management practices (N-BMPs) *3:*696
Nitrogen cycle *3:*308 *4:*517

Index terms **Links**

Nitrogen deficiency *3:*694
Nitrogen deposition *4:*380
Nitrogen fertilizer *2:*572 *3:*698
Nitrogen fixation *4:*517
Nitrogen fluxes, estuarine *4:*53
Nitrogen isotopes *5:*219
  as groundwater tracers *5:*503
Nitrogen loading
  effect on human health *3:*696
  identifying areas with *2:*571
Nitrogen management, in soil *3:*694
Nitrogen species, role in fish growth and production *3:*131
p-Nitrophenol (PNP), adsorption of *1:*117
4-Nitrophenol degradation *1:*793
‘‘Nixtamalization’’ process *1:*582 *1:*583f
NO_3^-. *3:*6
  See also Nitrate(s)
NOAA Atlantic Oceanographic and Meteorological Laboratory *4:*4
  See also National Oceanic and Atmospheric Administration (NOAA)
NOAA Forecast Systems Laboratory (FSL), ground- based GPS meteorology at *4:*244
NOAA Geophysical Fluid Dynamics Laboratory (GFDL) *4:*43
NOAA lake level forecast *3:*274
NOAA methyl bromide cycling study *4:*80
NOAA Pacific Marine Environmental Laboratory (PMEL) *4:*89
NOAA Profiler Network (NPN) *4:*249
No-action remediation *5:*44
  See also Passive groundwater remediation
Nocardioforms *1:*728
Nodal demands, distribution of *3:*315
Nodal replacement, for coupled flows *5:*660
Nodularins *2:*388 *2:*390
NOEC calculations *2:*411
  See also ‘‘No observed effect concentration’’ (NOEC)
‘‘No exposure’’ exclusion *4:*658
No-flow boundary *5:*419
Nonaqueous phase liquids (NAPLs) *2:*44 *5:*216 *5:*439
Nonbreaking wave *4:*17
Nonbiodegradable organics *1:*903
Noncarbonate hardness *1:*323 *4:*453

Index terms	**Links**	
Noncarcinogenic effects		
advisories based on	*3:*120	
fish consumption advisories based on	*3:*119	
Noncarcinogenic PAHs	*3:*350	
Noncarcinogenic responses	*4:*674	
Noncommunity public water systems (PWSs)	*4:*676	
Nonconservative dissolved inorganic P and N fluxes	*4:*53	
Nonconservative pollutants	*4:*50	
Nonconservative pollution, modeling of	*3:*332	
Nonculturing biofouling analytical methods	*1:*85	
Nonengineering water conservation interventions	*2:*472	
‘‘No-net-loss’’ goal	*4:*693	
Nonexclusion principle	*1:*216	
Nongovernmental organizations (NGOs)	*1:*655	
agriculture and land use planning	*3:*598	
water resource	*4:*651	
‘‘Nonhazardous’’ waste management	*2:*163	
Nonindigenous fish species	*3:*128	
Nonindustrial leaching	*5:*260	
Nonionic surfactants	*1:*670	*1:*671t
Nonlinear recession equation	*3:*23	
Nonnative species, introduction of	*3:*125	
See also Nonindigenous fish species		
Non-Newtonian fluids	*5:*556	
Nonoxidizing biocides	*2:*242	
Nonoxidizing chlorine byproducts (CBP)	*1:*129	
Nonoxygen-based sensors	*2:*41	
Nonpathogenic microbes	*2:*243	
Non-point fluoride contamination sources	*5:*132	
See also Non-point source contamination		
Non-point source (NPS) contamination	*5:*603	
See also Non-point fluoride contamination sources; Non-point source (NPS) pollution		
Nonpoint source control	*2:*184	
federal programs for	*2:*185	
regulations related to	*2:*185	
by source categories	*2:*186	
total maximum daily load and	*2:*188	
Nonpoint source estimation, equations for variables used in	*3:*662f	

Non-point source pollutant model

  components of *5:*300

  scale of *5:*303

  in the vadose zone *5:*299

Nonpoint source pollutants *2:*373

Nonpoint source (NPS) pollution *2*:181 *2*:184 *2*:544 *2*:614 *3:*37 *3:226t* *3:*228 *3:*502 *5:*331

  See also Non-point source (NPS) contamination

  control of *4:*595 *5:*333

  during and after development *3:*501

  management of *2:*311

  terrestrial *3:*282

Nonpoint source pollution assessment, role of Geographical Information System and remote sensing in *3:*658

Nonpoint sources *2:*169 *2:*190

Nonpoint source water quality models *2:*248

Nonpolar molecules, hydration of *4*:475 *4*:476

Nonpolar organic compounds, sorption of *4:*385

Nonpotable wastewater reuse *1:*821 *1:*826

Nonpriority pollutant petroleum hydrocarbon analyses *2:*310

Nonrenewable natural resources *2:*634

Nonrival consumption *1:*216

Nonskeletal fluorosis *4:*435

Nonsparkling water *1:*4

Nonsteady state, of drainage water *3:*678

Nonstructural flood adjustment measures *3:*146

  web pages related to *3:*150

Non-toxicant-specific luminescent biosensors *2*:454t

Nontoxic chemicals *2:*278

Nontransient, noncommunity water systems (NTNCWS) *4:*676

Nonturbulent channel units *3:*70

Nonweighing drainage lysimeters

  drainage measurements using *5:*488

  recent usage of *5:*488

Nonyphenol aquatic toxicity *1*:673t

No-observable-adverse-effect level (NOAEL) *1:*33 *1:*426

''No observed effect concentration'' (NOEC) *2:*381

  See also NOEC calculations

''No observed effect level'' (NOEL) *2:*381

Normalized Phaeophytinization Index (NPQI) *3:*722
Normalized Pigment Chlorophyll Index (NPCI) *3:*722
North American Free Trade Agreement (NAFTA) *4:*645 *4:*646
North Atlantic Oscillation Index *4:*340 *4:*341
Northern Gulf of Mexico, coastal fog along *4:*230
Northwest Power Planning Council (NPPC) *4:*614
Norway, coral reefs in *4:*34
''Not in my backyard'' (NIMBY) *2:*167
NPDES permits *2:*132 *2:*188 *2:*373 *4:*655
  See also National Pollutant Discharge Elimination System (NPDES)
N/P (nitrogen/phosphorus) ratio *2:*143
Nramp proteins *5:*286
NSF International listing program *1:*381
Nuclear detonations, underground *4:*102
Nuclear fuel reprocessing *1:*561 *4:*101
Nuclear fuel storage. *See* Spent nuclear fuel storage
Nuclear power water use *1:*561
Nuclear reactors, activated carbon in *1:*105
Nuclear techniques *5:*165
Nucleation temperature *4:*585
Nucleic acid probes *1:*161
Nucleic acid sequence-based amplification (NASBA) *2:*337
Nucleic acid testing *1:*85
  See also DNA entries; Polymerase chain reaction (PCR) assay
Nuisance plants, characteristics of *3:*744
Numerical modeling *3:*321 *5:*33
  See also Mathematical models; Numerical models
  case studies in *3:*322
  of currents *3:*320
  dimensions in *3:*321
  for Orlando Naval Training Center *5:*83
  verification, calibration, and validation in *3:*322
Numerical models *2:*272
  See also Numerical modeling
  flow condition *1:*58
Numerical solutions, for subsurface drainage *5:*97
Numerical weather prediction (NWP) *4:*350 *4:*354
Numeric effluent system *1:*756
Nutating disc meters *1:*338
Nutrient adsorption, in biofouling *2:*239

Nutrient availability, in aquatic systems *3:*109

Nutrient concentrations
  in natural environments *3:*311
  in sediment column porewater *4:*73
  tidally mediated changes in *4:*81

Nutrient control *3:*111

Nutrient control programs, issues in developing *3:*111

Nutrient deficiencies, in activated sludge processes *1:*729

Nutrient excretion, bivalve *4:*12

Nutrient excretion rates, bivalve *4:*6

Nutrient export coefficients *3:*110

Nutrient load criteria *3:*111

Nutrient loading *2:*3

Nutrient management *2:*571

Nutrient/phosphorus load reduction, rate of recovery after *3:*112

Nutrient pollution, tile drainage and *3:*729

Nutrient-related water quality *3:*111

Nutrient research, on agricultural drainage ditches *3:*89

Nutrient runoff control *3:*110

Nutrients, *See also* Microbial nutrients
  for algal growth *2:*269 *3:*110
  biological removal of *1:*399
  from cage farms *3:*580
  in coastal waters *4:*25
  control of *3:*483 *5:*334
  distributional pattern of *4:*81
  enhanced biodegradation and *5:*320
  enhancement and distribution of *5:*117
  in estuaries and bays *4:*26
  impact on coral reefs *4:*114
  limiting *3:*109
  in lotic aggregates *3:*306
  luminescent bacteria for determining *2:*175
  “luxury uptake” of *3:*273
  managing *2:*186
  in marine environments *4:*51
  removal by constructed wetlands *3:*368
  role in estuaries *4:*51
  sources of *3:*115
  tidal stream variability and *4:*131

Nutrients *(Continued)*
  in urban stormwater runoff *3:*434
  in water *1:*903
  in water bodies *2:*3
  as water pollution *4:*97
  as water quality indicators *2:*266
Nutrient transformation processes *2:*270
Nutrition, emergency *2:*528

O

O_2 evolution, implications of substrate water binding for *4:*540
  See also Molecular oxygen
Observation wells, characteristics of *5:*496
Obukhov length *4:*1
Occupational Safety and Health Administration (OSHA) water handling standards *2:*162
Occupations, water-transportation *4:*762
Occurrence-based reliability *3:*261
Ocean(s). *See also* Ocean water(s)
  capacity to absorb carbon dioxide *4:*86
  debris abatement in *4:*38
  dissolved gases in *4:*451
  factor in determining surface
  temperatures *4:*183
  frozen water in *4:*194
  interaction with atmosphere *4:*22
  measuring *4:*22
  modeling *4:*22
  processes in *4:*21
  properties of *4:*21
  renewable energies from *4:*44
  role in climate *4:*21
  role in climatically-active trace-gas cycles *4:*85
  salinity in *4:*21
Ocean Acoustics Division (NOAA) *4:*6
Ocean-atmosphere models
  coupled *4:*86
  three-dimensional *4:*172
Ocean Chemistry Division (NOAA) *4:*6
Ocean CO_2 uptake *4:*86
Ocean Conservancy *4:*39
Ocean Drilling Program (ODP) *4:*61

Ocean dumping, types of *4:*145

Ocean fishery stocks, overexploitation of *4:*124

Oceania, water symbolism in *4:*786

Oceanic chalk, carbonate as *4:*414

Oceanic circulation, changes in *4:*174

Oceanic processes *4:*104

Oceanic sound, fate of *4:*570

Oceanographic environment, Glacier Bay *4:*61

Oceanographic processes, stock assessments and *4:*57

Oceanography
- fisheries *4:*55
- GLERL instrumentation in *4:*70
- physical *4:*95

Ocean pollution, response of indicator groups to *2:*441

Ocean ranching *3:*128

Ocean shipments/dumping, precautionary principle regarding *2:*600

Ocean surface waters, antimony concentration in *4:*592

Ocean thermal energy conversion (OTEC) *4:*47

Ocean thermal energy conversion plants *4:*47

Ocean waste disposal
- effects of *4:*103
- increase in *4:*146

Ocean water(s) *4:*285
- heating of *4:*584
- composition of *4:*285

Octanol-water partition coefficient *1:*574

Oculina varicosa *4:*34
- distribution of *4:*35
- threats to *4:*36

Odor abatement. *See also* Odor-reducing additives
- aqueous phase partition and *1:*761
- biodegradability and *1:*761
- biological treatment technologies for *1:*762
- causes of odor *1:*760
- choices related to *1:*763
- competing technologies for *1:*762
- odor measuring systems *1:*760
- volatility and *1:*761
- in wastewater treatment plants *1:*760

Odor analysis, state-of-the-art developments in *2:*281

Odor control, using ozone *1:*355

Odor measurement methods, basis of *2:*281

Odor monitoring, electronic nose for *2:*281

Odor-reducing additives *1:*912

Odors. *See also* Odor abatement; Sewerage odors

  of domestic sewage *1:*830

  landfill *2:*166

  of water *1:*902

^{18}O exchange kinetics *4:*538 *4:*540

  See also ^{18}O-labeled water

  rapid *4:*536

Off-channel pumped storage development *3:*201

Office of Environmental Health Hazard Assessment (OEHHA) *1:*479

Office of Oceanic and Atmospheric Research (NOAA) *4:*4

Office of Solid Waste and Emergency Response (EPA) *2:*517

Off-site sewage treatment systems *1:*830

Offstream reservoirs *3:*411

Offstream storage *1:*486

Offstream water use *2:*650

Ohio River Sanitary Commission (ORSANCO) compact *4:*651

Ohm's law *4:*442

Oil

  toxic components in *4:*106

  in urban stormwater runoff *3:*435

  as water pollution *4:*98

Oil-destroying microorganisms *2:*449

Oil-field brine (OFB) *2:*284

  controlling by wastewater treatment methods *2:*288

  historical control practice for *2:*285

  impact on water quality *2:*287t

  pollution aspects of *2:*286

  use in oil production *2:*285

Oil-methylesters *3:*546

Oil pollution *2:*290

  See also Exxon Valdez oil spill; Oil spills

  microorganisms as indicators of *2:*441

Oil Pollution Act *4:*597

Oil price increases *4:*479

Oil recovery *1:*561

Oil-related accidents *4:*99

Index terms		Links
Oil reservoir monitoring	*5:*54	
Oil-resistant microorganisms	*2:*448t	
Oil slicks, surface	*4:*106	
Oil spills. *See also Exxon Valdez* disaster		
biggest	*2:*291t	
effects of	*4:*105	
case study of	*2:*172	
Oklahoma, drought index used in	*4:*213	
^{18}O-labeled water	*4:*537	
Old groundwaters, ^{36}Cl and	*4:*416	
Olfactometry	*2:*281	
Oligochaetes, accumulation patterns for	*2:*215	
Oligotrophic systems	*3:*268	
Oligotrophy	*1:*248	
Ombrogenic areas	*1:*404	
Omega equation	*4:*354	
Once through systems	*1:*559	
One-dimensional numerical models	*3:*321	
One-dimensional transport equations	*5:*525	
One-way flow systems, rivers and streams as	*3:*437	
On-line drippers	*3:*622	
Online monitors, bioluminescence-based	*2:*174	
On-line respirometry	*1:*565	
Online water toxicity, monitoring luminescent bacteria in	*2:*174	
On-site disposal systems, managing runoff for	*2:*187	
On-site sewage treatment systems	*1:*520	*1:*830
OnSite system	*5:*81	
On-site water quality determination, luminescent bacteria for	*2:*173	
Open channel flows	*3:*195	
design of	*3:*346	
Open-cut trench construction	*1:*876	
Open limestone channels	*2:*424	
in acid mine treatment	*5:*6	
Open ocean fisheries management	*3:*126	
Open ocean N_2O emissions	*4:*87	
Open recirculating systems	*1:*559	
Open waterbodies, heat balance of	*3:*190	
Operating storage	*1:*409	
Operational meteorologists	*4:*328	*4:*329

Index terms	**Links**	
“Operational storage”	*1:*202	
Operations and maintenance (O & M) contracts	*1:*388	*1:*389
Optical pH sensors	*2:*299	
Optical properties		
of plant canopies	*3:*720	
of plant leaves	*3:*719	
Optical televiewer	*5:*154	
Optimization, science of	*3:*386	
Optimization algorithms	*4:*298	
Optimized soil-adjusted vegetation index (OSAVI)	*3:*722	
Optimum sustainable yield	*3:*124	
Option contracts	*2:*607	
Option prices	*2:*607	
ORBAL plant	*1:*742	
ORCHESTRA program	*5:*141	
Oregon, in-stream protection in	*4:*660	
Organic acid anions	*3:*2	
Organic agriculture	*3:*644	
Organically cultivated vegetables, organochlorine pesticides in	*3:*645	
Organic arsenic species	*2:*12	*2:*13t
Organic binders	*1:*836	
Organic carbon. *See also* Carbon (C)		
control of	*2:*224	
in distribution systems	*2:*244	
types of	*2:*224	
Organic chemical monitoring	*2:*267	
Organic chemicals/compounds	*5:*185	*5:*258
See also Organic removal; Organic materials; Organic waste compounds		
adsorption of	*4:*384	
adsorption on carbons	*4:*407	
aqueous reactions with ozone	*1:*765	
classes of	*5:*339t	
competitive sorption of	*1:*115	
in domestic sewage	*1:*830	
in drinking water	*2:*372	
in groundwater	*5:*337	
in Puget Sound Basin sediment and fish	*3:*349	
removal by biological filtration	*1:*248	
sorption of	*4:*386	
transport of	*4:*384	

Index terms		**Links**
Organic chemicals/compounds *(Continued)*		
in urban stormwater runoff	*3:*435	
water contamination from	*5:*339	
year first available	*3:*170t	
Organic contaminants	*1:*375t	
See also Organic wastewater contaminants		
in Indian rivers	*3:*449	
metabolism of	*5:*529	
removal by constructed wetlands	*3:*368	
Organic contamination, indicators of	*2:*442	
Organic loading, relation to eutrophication	*2:*142	
Organic materials, sorption onto	*5:*5	
Organic matter	*2:*544	
See also Organic soil matter		
in an agricultural landscape	*3:*606	
in constructed wetlands	*1:*892	*1:*894
degradation of	*1:*726	*3:*180
flooded	*3:*206	
in foulant samples	*1:*417	
microbially mediated oxidation of	*2:*464	
oxygen deficiencies and	*3:*130	
removal of	*4:*426	
in water	*1:*903	
Organic micropollutant removal, by activated carbons	*4:*506	
Organic nitrogen	*3:*538	
Organic phosphorus	*3:*702	
Organic pollutants, microbial indication of	*2:*441	
Organic pollution assessments, microbial biosurveys for	*2:*449	
Organic pollution measures, alternative	*2:*40	
Organic polymers, as coagulants	*4:*426	
Organic removal	*1:*350	
by organoclays	*1:*772	
in small systems	*1:*351	*1:*460
Organic soil matter, nitrogen in	*3:*695	
Organic vapor analyzers (OVAs)	*5:*549	
Organic vapors	*5:*545	
biodegradation of	*5:*545	
sorption of	*5:*545	
sources of	*5:*543	
transport and fate of	*5:*544	

Index terms **Links**

- Organic waste compounds
  - most frequently detected *5:*61f
  - inhibitory threshold concentrations of *1:*754t
  - water contamination by *5:*60
- Organic wastewater contaminants, in U.S. streams *5:*605
- Organism assemblage information *4:*600
- Organisms. *See also* Indicator organisms; Microorganisms; Pathogens/pathogenic organisms; Sediment dwelling organisms
  - cadmium in *5:*617
  - effects of pharmaceuticals on *1:*376
  - hatchery-raised *4:*126
  - labeled substrate probes for *1:*645
  - selection for sediment toxicity tests *2:*459
  - trace element concentrations in *4:*111
- Organization for Economic Cooperation and Development (OECD)
  - BOD_5 standard *2:*39
  - eutrophication study data *3:*112
- Organizations, water resource *4:*648
- Organoantimony compounds *4:*592
- Organochlorine compounds
  - in Puget Sound Basin sediment and fish *3:*349
  - in sediment *3:*353 *3:*354t
  - in tissue *3:*360t
- Organochlorine pesticides (OCPs) *3:*649
  - absorption and accumulation of *3:*643
  - occurrence of *3:*643
- Organoclay
  - oil removal by *1:*774 *1:*776f
  - role in water cleanup *1:*771
  - in treating contaminated groundwater *1:*777t
- Organoclay indicating partition *1:*773f
- Organolead, uptake of *2:*434
  - *See also* Lead (Pb)
- Organophosphorous compounds *1:*548t
- Orifice meters *1:*338
- Orifices, flows through *3:*195
- ‘‘Orifice’’ type roof drainage *1:*56
- Orinoco basin *4:*642

Index terms	**Links**			
Orlando Naval Training Center				
feasibility investigation at	*5:*81			
model application for	*5:*83			
Orthokinetic flocculation	*1:*252			
Orthophosphate	*1:*788	*1:*789		
Oscillating water columns (OWC)	*4:*45			
Oscillations, in lakes	*3:*267			
Oslo process	*4:*700			
Osmolality	*4:*521			
Osmolarity	*4:*459			
Osmosis	*2:*544	*4:*458	*4:*459	*4:*520
Osmotic concentrations, units for	*4:*521			
Osmotic forces, head and	*5:*180			
Osmotic potential	*5:*181	*5:*532		
Osprey, effects of DDT on	*3:*518			
OSPREY oscillating water column	*4:*45			
OSPREY research vessel	*1:*748			
Ostwald's experiment	*5:*649			
OTIS program	*5:*141			
Our Common Future	*2:*625	*2:*633		
Outbreak detection, early	*1:*186			
Outdoor residential water use	*2:*652			
Outdoor water-conservation measures	*3:*553			
Outdoor watering equipment	*2:*666			
Outdoor water sources, emergency	*2:*527			
Outfall, defined	*2:*544			
Outflow rate, river	*3:*392			
Outflow structure design	*3:*500			
Outgassing	*4:*159			
Overall existence ranking index (OERI)	*2:*681			
Overburden analyses, in acid mine drainage	*5:*1			
Overcapitalization, during the Great Depression	*2:*511			
Overdraft	*5:*340	*5:*604		
Overfishing, impact of	*4:*107			
Overland flow (OLF)				
infiltration-excess	*3:*452			
saturation-excess	*3:*452			
Overlay and index approach, to pesticide assessment	*5:*595			
Overtopping	*4:*16	*4:*19		

Oxidants, chlorine-produced *1:*128

Oxidation *1:*380 *1:*811

 of arsenite [As(III)] *1:*639 *5:*19

 chemical *2:*300

 cometabolic *2:*300

 electrobiochemical *5:*122

 electrochemical *5:*122

 of iron *1:*313 *5:*609

 of landfill leachates *1:*706

 in limestone-based FGD systems *1:*849

 of manganese *1:*313

 in mine effluent remediation *1:*611

 for odor abatement *1:*762

 ozone-induced *1:*765

Oxidation and reduction (redox) *5:*185 *5:*527

 See also Oxidation-reduction entries; Redox entries

 of metals *5:*283

 potential *5:*321 *5:*584

Oxidation processes *5:*441

 advanced *1:*871

Oxidation-reduction potential (ORP) *2:*466

 See also E_h; Oxygen redox potential (ORP); Redox entries

Oxidation-reduction potential values *1:*850

Oxidation-reduction (redox) reactions *1:*399 *1:*725 *4:*518

 See also Redox potentials

Oxidation states, of wetlands *3:*367

Oxidative capacity (OXC) *2:*468

Oxidative precursor removal *2:*120

Oxidized carbons, sorptive capacity and selectivity of *2:*84

Oxidizing biocides *1:*602 *2:*242

Oxygen (O) *4:*189

 See also Dissolved oxygen (DO); Ozone

 delivery *5:*424

 isotopes *4:*501

 profiles *4:*451

 singlet *4:*533

Oxygenated hydrophilic groups *4:*401

Oxygenation, of sludge *1:*857

Oxygen-based sensors *2:*40

Oxygen complexes, in activated carbons *2:*81

Index terms	Links			
Oxygen concentration. *See also* Dissolved oxygen (DO)				
influence on nitrification	*3:*642			
in ponds	*3:*485			
Oxygen demand	*2:*544			
biochemical	*1:*639			
Oxygen-demanding wastes (ODWs), as water quality indicators	*2:*266			
Oxygen demand threshold	*5:*424			
Oxygen depletion, in waterbodies	*2:*181			
Oxygen enhancement	*5:*117			
Oxygen fugacity approach	*5:*414			
Oxygen isotope exchange measurements	*4:*536			
Oxygen isotopes, as groundwater tracers	*5:*503			
Oxygen limitation, in biofilms	*2:*231			
Oxygen-reducing to nitrate-reducing conditions	*5:*583			
Oxygen redox potential (ORP)	*2:*231			
See also Oxidation-reduction potential (ORP); Redox entries				
“Oxygen-sag curve”	*3:*449			
Oxygen transfer efficiency (OTE)	*1:*623			
in diffused air aeration systems	*1:*629			
Oxygen transfer rate (OTR)	*1:*623			
Oxygen uptake rates (OUR)	*1:*565	*1:*566	*1:*568	
Ozonation	*1:*199	*1:*458	*1:*817	*2:*119
	*2:*244	*2:*374		
See also Ozone				
byproducts of	*1:*355	*1:*363		
domestic sewage	*1:*834			
municipal water supply	*1:*362			
personnel requirements for	*1:*355			
process and equipment for	*1:*355			
Ozone	*1:*195	*1:*313	*1:*354	*1:*603
	*4:*270			
chemistry of	*1:*362			
at CSO facilities	*1:*786			
disinfection byproducts and	*1:*214			
disinfection using	*1:*354			
health and safety precautions for	*5:*349			
injection	*5:*442			
organic compound reactions with	*1:*765			
safety of	*1:*356			

Index terms	**Links**		
Ozone-bromide interactions	*1:*357		
bromate formation	*1:*358	*1:*359	
bromate ion minimization strategies	*1:*360		
bromate ion removal	*1:*361		
bromide ion	*1:*358		
Ozone bubble-diffuser contactor	*1:*363		
Ozone contactors	*1:*356	*1:*363	
for odor removal	*1:*914		
Ozone dose, in bromate ion formation	*1:*359		
Ozone generation, feed gas preparation for	*1:*356		
Ozone leachate treatment	*1:*706		
Ozone off-gas destruction	*1:*356		
Ozone treatment	*5:*347		

P

Pacific Fisheries Environmental Laboratory (PFEL), research at	*4:*96		
Pacific Guano Works	*4:*140		
Pacific Marine Environmental Laboratory (PMEL)	*4:*89		
Pacific-North American (PNA) circulation pattern index	*4:*174		
Pacific Northwest Electric Power Planning and Conservation Act	*4:*614		
Pacific Tropical Atmosphere Ocean (TAO) array	*4:*91		
Packaged filtration	*1:*229		
Package plants	*1:*514		
advantages and limitations of	*1:*515		
filtration systems for	*1:*515		
operation and maintenance of	*1:*517		
selecting	*1:*515		
Packed column aeration (PCA)	*1:*51	*1:*352	*1:*460
Packing cell formation	*2:*240		
PACT process	*2:*87		
PAH compounds	*1:*573	*1:*574	*1:*575
See also Polycyclic aromatic hydrocarbons (PAHs)			
Pain, thirst and	*4:*724		
Pakistan, national identity of	*4:*683		
Paleoceanography, seawater temperature estimates in	*4:*92		
Paleoclimatology	*4:*93		
Palmer Drought Severity Index (PDSI)	*4:*211		
limitations of	*4:*212		
Palmer Hydrological Drought Index (PHDI)	*4:*211		
Palmer Index	*2:*581		

Index terms	Links		
Pamlico-Albemarle Estuarine Complex	*4:*57		
Pan American Center for Sanitary Engineering and Environmental Sciences (CEPIS/PAHO)	*2:*671	*2:*673	*2:*674
Pan American Health Organization (PAHO/WHO)	*2:*671	*2:*672	*2:*673
Pan evaporation method	*3:*573		
Pantanal wetland	*4:*642		
Paper industry, impact on surface water quality	*3:*374		
Paraguay-Parana River hydro-via	*4:*642		
Parallel plate electrode arrangement	*4:*430		
Parallel scour pools	*3:*68		
Paralytic shellfish poisons (PSPs)	*2:*389		
Parameter calibration	*4:*298		
Parameter estimation	*4:*221		
chlorine reaction	*1:*134		
inverse model for	*1:*135		
Parameter optimization	*2:*332	*4:*299	
Parametric mapping	*5:*657		
Parasites	*3:*668		
in constructed wetlands	*1:*894		
fate and transport in karstic aquifers	*1:*365		
waterborne	*2:*344t		
Parasitic infections	*1:*25		
Partial private-sector responsibility model	*1:*50		
Participatory integrated capacity building (PICB)	*1:*652	*2:*628	
Participatory multicriteria flood management	*2:*678		
Particle density	*1:*96		
Particle settling velocity	*1:*259		
Particle size (d_p)	*2:*545	*3:*300	
Particulate arsenic	*2:*11		
Particulate matter, in wetlands	*1:*892		
Particulate metal contamination	*5:*281		
Particulate phosphorus (PP)	*3:*704		
Particulate removal	*1:*370		
by coagulation	*1:*137		
by filtration and sedimentation	*1:*243		
by granular media filtration	*1:*233		
by high-rate granular filtration	*1:*250		
Particulate transport, in groundwater	*5:*349		
Partitioning	*4:*521		
properties that determine	*4:*522		

Index terms	**Links**	
Partnering	*1:*42	
EPA Pretreatment Program	*1:*798	
Passive biomonitoring (PBM)	*2:*33	
Passive biotreatment	*5:*120	
See also Passive groundwater remediation		
Passive diffusion samplers		
case study of	*5:*458	
groundwater sampling with	*5:*456	
Passive disease reporting	*1:*185	
Passive filter systems	*1:*237	
Passive groundwater remediation	*5:*423	
See also No-action remediation; Passive biotreatment		
Passive immobilization, of metals	*5:*283	
Passive *in situ* remediation technologies (PIRT)	*5:*423	
Passive microwaves	*4:*324	
Passive radioactivity log	*5:*152	
Passive reactive treatment, of vinyl chloride	*5:*638	
Passive remote sensing, spectral regions used for	*4:*320	
Passive sound detection	*4:*570	
Passive treatment		
of acid mine drainage	*2:*423	
selecting	*2:*425f	
systems	*2:*424f	
Passive vapor survey	*5:*437	
Pasteurization, of sludge	*1:*858	
Pathogenic bacteria. *See also* Pathogens/pathogenic organisms		
detection in surface waters	*2:*441	
removal of	*1:*485	
Pathogenic microorganisms, wastewater-associated	*3:*668t	
Pathogenic pollution	*2:*182	
Pathogens/pathogenic organisms	*1:*903	*2:*545
See also Pathogenic bacteria		
in animal farming operations	*3:*539	
in constructed wetlands	*1:*894	
in domestic sewage	*1:*831	
drinking water	*1:*277	
impact on coral reefs	*4:*115	
opportunistic	*1:*342	
persistence of	*1:*521	
waterborne	*2:*344t	

Index terms	**Links**			
Pathogens/pathogenic organisms *(Continued)*				
as water pollution	*4:*97			
as water quality indicators	*2:*266			
Patuxent Wildlife Research Center	*3:*519			
Paul Lake, biomanipulation trials in	*2:*56t			
Pb bioavailability, microbial aminolevulinate dehydratase as a biosensor for	*2:*236			
See also Lead (Pb)				
PCB-contaminated sediments, Hudson River	*2:*126			
PCB/dioxin-like compounds, cytochrome P450 monooxygenase as an indicator of	*2:*106			
PCR-based monitoring, of *Salmonella*	*2:*338			
Penalty finite element scheme	*5:*658			
pe	*2:*466			
See also E_h; pH				
Peak discharge	*3:*143			
Peak flow	*2:*545			
Peaking plants	*3:*202			
Pearl River	*4:*697			
Pedersen-Burcharth formula	*4:*17			
Pelagic fish	*3:*126			
Pelagic organisms	*3:*310			
Pelagic zone	*3:*267	*3:*268		
Penman-Monteith energy balance equation	*3:*574	*3:*576	*3:*721	*5:*78
Penman-Monteith model	*4:*227			
Penneable material	*5:*600			
Pennsylvania, source-water protection in	*2:*312			
Pennsylvania Drought Contingency Plan	*2:*581			
Pens, role in trout hatching	*3:*458			
Penstock	*2:*554			
Pentachlorophenol (PCP)	*1:*689			
Pentavalent arsenic compounds	*2:*11	*2:*16		
People's World Water Forum (PWWF)	*1:*215			
Peracetic acid	*1:*604			
at CSO facilities	*1:*786			
Per capita use	*2:*545			
Percent of normal calculation	*4:*209			
Perceptual model (PM) uncertainty	*3:*325			
Perched groundwater	*5:*352			
examples of	*5:*353			
hydraulic properties of	*5:*352			
hydrologic conditions related to	*5:*353			

Index terms	Links		
Perched groundwater *(Continued)*			
role in contaminant movement	*5:*354		
Perched water table	*5:*11		
Perchlorate	*2:*317	*2:*345	
characteristics and persistence of	*5:*632		
fate and transport of	*5:*632		
in groundwater	*5:*631		
health risks of	*5:*632		
physical removal technologies for	*5:*633		
remediation of	*5:*632		
as an unrecognized pollutant	*3:*372		
uses for	*5:*631		
Perchlorate Community Right to Know Act of 2003	*2:*345		
Perchloroethene (PCE)	*5:*91		
Perchloroethylene (PCE)	*2:*317		
removal of	*2:*299		
‘‘Percolating’’ groundwater	*4:*630		
Percolation	*2:*545		
Percolation tank	*2:*551		
Perennial flow, artificial	*2:*477		
Perennially draining epikarsts	*5:*237		
Perennial streams, base flow and	*3:*22		
Perennial waters	*2:*475		
Performance, impeller	*1:*80		
Performance evaluation, for water resource systems	*2:*674		
Performance failure	*3:*259		
Performance indicators, for developing countries	*1:*721		
Perhydroxyl radical	*4:*446	*5:*345	
Perikinetic flocculation	*1:*252		
Periodicity analysis	*4:*338		
Peripheral nervous system, effects of lead on	*2:*436		
Peripheral neuropathy	*2:*436		
Periphyton index of biotic integrity (Periphyton IBI)	*3:*40		
Permanent frost (permafrost)	*4:*305		
geography of	*4:*306		
groundwater and	*5:*601		
Permanganate salts, health and safety precautions for	*5:*348		
Permanent magnetic water conditioning (PMWC)	*1:*141	*1:*143	*1:*144
Permanent Service for Mean Sea Level (PSMSL)	*4:*118		
METRIC dataset from	*4:*119		

Index terms **Links**

Permanent spot market *2:*606
Permanent water hardness *4:*453
Permanent wilting point *4:*485
Permanganate treatment *5:*346
Permeability (*k*) *2:*545 *5:*355
  conversion factors *5:*509t
  groundwater *5:*128
  soil *3:*30
  values of *5:*356t
Permeable reactive barriers (PRBs) *5:*518 *5:*423 *5:*442
  field data sources and model setup for *5:*519
  technology of *5:*41
Permeameters *5:*512
Permit statutes *4:*632
Peroxyl radicals *4:*532
Persistent organic pollutants (POPs) *4:*99
  in the activated sludge treatment process *1:*767
  in the wastewater treatment process *1:*766
Personal care products (PPCPs) *1:*819
  See also Pharmaceuticals and personal care products (PPCPs)
  as unrecognized pollutants *3:*371
Pertechnetate anion *4:*579
Perturbed continuity modeling scheme *5:*658
Pervious toe drain *3:*290
Pervious toe trench *3:*289
PES capillary membranes *1:*593
Pesticide assessment
  approaches to *5:*595
  uncertainty in *5:*358
Pesticide-caused toxicity, of urban stormwater runoff *3:*433
Pesticide chemistry *3:*647
Pesticide-contaminated soil, remediation of *3:*651
Pesticide groups *3:*648
Pesticide leaching assessment *5:*359t
Pesticide research, on agricultural drainage ditches *3:*90t
Pesticide Root Zone Model (PRZM) *2:*252
Pesticides
  agricultural *3:*607 *3:*643
  behavior of *3:*647
  biological and chemical degradation of *3:*648

Pesticides *(Continued)*
  control of *5:*334
  effect on surface water *3:*374 *3:*650
  environmental pathways of *3:*655
  groundwater vulnerability to *5:*357 *5:*594
  important *5:*358t
  managing *2:*186
  occurrence/use relationship for *3:*655
  organochlorine *3:*352 *3:*643 *3:*649
  runoff and leaching of *3:*648 *3:*651
  sorption to soil components *3:*648
  in surface waters *5:*332
  tile drainage and *3:*729
  transport and transformation processes of *3:*655
  as unrecognized pollutants *3:*372
  volatilization of *3:*648
  water quality and *2:*266 *2:*303
Pesticide-use issues, key *3:*656
Pesticide vulnerability assessment *5:*595
Pesticide vulnerability maps *5:*598
Pest Management Regulatory Agency (PCPA) *1:*382
Peter the Great Bay *2:*445
  features of marine bacteria in *2:*446
heavy metal monitoring in *2:*446
Petroleum hydrocarbons *5:*43 *5:*116 *5:*322 *5:*428
aerobic biodegradation of *5:*320
degradation of *1:*692
plants that spur degradation of *3:*629
Petroleum pollution *2:*290
  See also Oil entries
Petroleum products *2:*304
Petroleum residues *4:*98
Pfiesteria *3:*108 *4:*108
Pfiesteria piscicida *3:*702
pH *2:*294 *2:*544 *5:*184
  See also High pH groundwater; pH adjustment
  of acid mine drainage *2:*1
  acid-neutralizing capacity and *3:*1
  bacteria and *3:*14 *5:*321
  bromate ion concentration and *1:*360

Index terms	Links	
pH *(Continued)*		
dishwater	*2:*113	
effect of carbon on	*1:*97	
effect on adsorption	*3:*299	
effect on corrosion	*1:*8	
impact on DBP concentrations	*2:*119	
influence on nitrification	*3:*641	
inorganic ion removal and	*4:*491	
measurement of	*2:*295	
microbial growth and	*3:*312	
organic matter removal and	*4:*427	
of pond water	*3:*487	
relationship to acidity	*3:*8	
research on	*2:*352	
role in fish growth and production	*3:*131	
soil	*5:*367	
theory of	*2:*294	
of urban stormwater runoff	*3:*434	
pH adjustment		
chemicals used for	*1:*10t	
corrosion control via	*1:*9	*1:*153
of landfill wastewater	*1:*705	
Pharmaceutical industry, activated carbon in	*1:*105	
Pharmaceuticals		
in animal farming operations	*3:*540	
characteristics of	*1:*374t	
detection in effluent-receiving water systems	*1:*376	
effect on human health	*1:*377	
in the environment	*1:*372	*1:*376
persistence of	*1:*375	
regulations concerning	*1:*377	
in sewage treatment facilities	*1:*373	
in the United States	*1:*373	
in water systems	*1:*372	
Pharmaceuticals and personal care products (PPCPs)	*3:*435	
See also Personal care products (PPCPs)		
as unrecognized pollutants	*3:*371	
Pharmaceutical water-for-injection systems	*1:*602	
Phase I Environmental Assessment	*5:*437	
Phase I stormwater discharge regulations	*4:*657	
Phase II stormwater discharge regulations	*4:*658	

Index terms	Links	
Phase II Subsurface Investigation	*5:*437	
Phase III Corrective Action	*5:*437	
Phase IV groundwater monitoring	*5:*438	
Phasing practice, for landfills	*1:*697	
pH balance, in the body	*4:*456	
pH buffer(s)	*2:*295	
carbonate as	*4:*414	
Phenol adsorption		
effect of carbon surface chemical composition on	*4:*404	
by organoclay	*1:*777f	
porous structure and	*4:*407	
on zeolites	*4:*406	
Phenolic compounds	*1:*530	
See also Phenols		
Phenol number	*1:*98	
Phenology, crop	*3:*722	
Phenol-resistant microorganisms	*2:*448t	
Phenols		
adsorption of	*1:*117	*1:*119
microbial detection of	*2:*442	
toxicity of	*4:*524	
pH glass element	*2:*296	
Phoenix, Arizona, overdraft in	*5:*341	
PhoP primers	*2:*338	
Phosphate concentration	*5:*199	
Phosphate fluxes, estuarine	*4:*53	
Phosphates		
distribution of	*4:*74	
inorganic	*1:*154	
Phospholipid Fatty Acid (PLFA) analysis	*1:*85	
Phospholipid Fatty Acid Methyl Ester (PL-FAME) analysis	*1:*85	
Phospholipids	*2:*442	
Phosphorus (P). *See also* Nonconservative dissolved inorganic P and N fluxes; N/P (nitrogen/phosphorus) ratio; Phosphates; Phosphorus removal; Soil phosphorus		
in an agricultural landscape	*3:*605	
algal growth and	*3:*109	
in animal farming operations	*3:*539	
in constructed wetlands	*1:*893	
in distribution systems	*2:*245	
in domestic sewage	*1:*831	

Index terms	**Links**		
Phosphorus (P) *(Continued)*			
forms of	*2:*269		
in foulant samples	*1:*418		
as a nutrient of concern	*2:*269		
tile drainage and	*3:*730		
Phosphorus-accumulating organisms (PAO)	*1:*788		
Phosphorus availability indices	*3:*703		
Phosphorus loading, reduction in	*3:*113		
Phosphorus precipitation, in algal control	*2:*4		
Phosphorus removal			
in the activated sludge process	*1:*788		
biological	*1:*816	*1:*833	
Phostrip process	*1:*790		
Photocatalysis	*1:*872		
heterogeneous	*1:*791		
Photocatalyst	*1:*792		
See also Photocatalysis			
Photocatalytic membrane reactors			
in pollutant photodegradation	*1:*793		
in water purification	*1:*791		
Photocatalytic oxidation	*2:*300		
Photochemical Reflectance Index (PRI)	*3:*722		
Photochemistry, environmental	*4:*529		
Photography, water in the history of	*4:*766		
Photoionization detectors (PIDs)	*2:*307	*3:*689	
Photon emitters, health risks of	*1:*396		
Photophysical processes	*4:*530		
Photoreactivation	*1:*471		
Photospheres, water in	*4:*190		
Photosynthesis			
crop	*3:*721		
sensitivity to Pb toxicity	*2:*434		
Photosynthetic bacteria	*2:*21		
Photosynthetic O_2 evolution, rapid ^{18}O exchange measurements during	*4:*536		
Photosynthetic water productivity	*3:*560		
Photosystem II (PSII) enzyme complex	*4:*535		
Phragmites research	*4:*57		
PHREEQC geochemical code	*5:*33		
PHREEQC program	*2:*363	*5:*140	*5:*143
pH reference element	*2:*296		

Index terms **Links**

PHRQCGRF program *5:*140
PHRQPITZ program *5:*140
Physical adsorption *1:*99
Physical barriers, to vinyl chloride *5:*639
Physical carbon activation *1:*94
Physical-chemical treatment
  of landfill leachates *1:*709
  of leachates *1:*704
  of vinyl chloride *5:*638
Physical control, of biofilms *1:*541
Physical CSO treatment *1:*783
Physical Habitat Simulation (PHABSIM) models *3:*106 *3:*526
Physical infiltration models *4:*487
Physical intrusion, guarding against *1:*870
Physical meteorologists *4:*328
Physical models *3:*328 *3:*337
Physical oceanography *4:*95
Physical Oceanography Division (NOAA) *4:*6
Physical pollutant removal processes *3:*366
Physical reactors, sewers as *3:*331
Physical recharge measurement *5:*74
Physical screen fish passage facility *3:*530
Physical separation techniques *5:*435
Physical wastewater treatment technologies *1:*809
Physical water conditioning *1:*141
Physicochemical monitoring *2:*267
Physiologically based extraction test (PBET) *4:*526
Physiology, microbial *5:*580
Physisorption *4:*386
Phytoaccumulation, of cobalt *5:*612
Phytin *1:*788
Phytochelatins (PCs) *2:*406 *5:*287 *5:*371 *5:*376
  role in plant protection *3:*610
Phytochelatin (PC)-synthase activity *5:*287
Phytocontainment, assessment and characterization of *5:*76
Phytodetoxification *5:*399
Phytoextraction *5:*285 *5:*365 *5:*399 *5:*434
  economic considerations related to *5:*367
  future of *5:*372

Index terms	**Links**		
Phytoextraction *(Continued)*			
for lead-contaminated soils	*5:*382		
of metals	*5:*288		
optimization of	*5:*372		
regulatory considerations related to	*5:*368		
using hyperaccumulator plants	*5:*369		
Phytoextraction coefficient	*3:*35	*5:*383	
Phytofiltration	*5:*399		
Phytohydraulic containment	*5:*77		
Phytohydraulic containment and treatment, feasibility investigation for	*5:*81		
“Phytolignification”	*5:*383		
PHYTOPET database	*3:*628		
Phytoplankton	*2:*142	*2:*269	*4:*154
estuarine	*4:*52		
trace elements in	*3:*456		
PHYTORAD database	*3:*635		
PHYTOREM database	*3:*629		
contents and update recommendations for	*3:*633		
fields, description, and access for	*3:*632	*3:*633t	
hyperaccumulation and accumulation criteria for	*3:*631t		
Phytoremediation	*5:*366		
See also Remediation			
applications of	*3:*629		
of cadmium	*5:*618		
components of	*5:*377		
by constructed wetlands	*3:*364		
of contaminated soil	*3:*653		
genetic engineering of plants for	*5:*288		
hydrocarbon and metal	*3:*628		
hydrologic feasibility assessment and design in	*5:*76		
of lead	*5:*648		
of lead-contaminated soils	*5:*381		
of metals	*5:*284		
of methyl *tertiary*-butyl ether	*5:*385		
need for	*3:*739		
of selenium-laden soils	*5:*397		
soil	*5:*434		
types of	*5:*375t		
of uranium	*5:*641		
Phytoremediation enhancement, of natural attenuation processes	*5:*374		
Phytoremediation reactions, influences on	*3:*367		

Index terms **Links**

Phytostabilization *5:*365 *5:*399 *5:*434
  economic considerations related to *5:*367
  for lead-contaminated soils *5:*383
  regulatory considerations related to *5:*368
Phytovolatilization *5:*399
Pick-Sloan plan *4:*616
Piedmont lakes *3:*265 *3:*270
Piezometric head *5:*160
Piezometric surface *5:*600
Pilot systems, corrosion control in *1:*11
Pipe bedding *1:*890
Pipe coatings/linings *1:*154
Pipe diameter
  measuring *1:*401f
  pipe roughness versus *3:*314
Pipeline renovation, using trenchless techniques *1:*876
Pipelines. *See also* Pipes; Underground pipeline
  cleaning techniques for *1:*206
  inadequate *1:*317
  televising of *1:*884
Pipeline sediment disposal *2:*124
Pipe materials, microbial regrowth and *1:*345
  See also Pipe coatings/linings
Pipe repair, problems during *1:*801
Pipe roughness values *3:*314
Piper trilinear diagram *2:*396
Pipes. *See also* Line breaks; Piping
  examining *1:*153
  flows in *3:*195
  laying *1:*402
  steady-state flow in *5:*654
  water distribution system *1:*208
Pipe scrapers *1:*880
Pipe throughflow *3:*453
Pipe wall biofilm communities *2:*223
Piping *5:*403
  See also Pipes
  leaking *5:*603
Pippin, Steven *4:*768
Pisaster ochraceus *4:*153

Index terms **Links**

Piston meters *1:*338

Pit lakes *3:*270

Pitless adaptors *1:*151

Planar flows, shocks in *3:*241

Plane Couette flow *5:*653

Plane Poiseuille flow *5:*653

Plane series
  design discharge for *3:*243
  time of concentration for *3:*243

Planetary boundary layer (PBL) models *4:*322

Planetoids, water in *4:*191

Plankton *4:*154
  categories of *4:*154
  collecting and studying *4:*155
  food chains, filtrators, and size spectra in *3:*268
  global importance of *4:*157
  impact of oil on *4:*106
  seasonal patterns of *4:*156
  spatial patterns of *4:*155
  temporal patterns of *4:*156

Planktonic algae *3:*107
  blooms of *3:*133

Plankton Survey System (PSS) *4:*72

Planned water recycling *2:*610

Planning
  decision support systems and *2:*619
  for droughts *2:*578
  river basin *3:*33
  water-infrastructure *2:*594

Planning tools, drought-related *2:*580

Plan Puebla-Panama *4:*641

Plant available water *5:*124

Plant-bacteria-contaminant interactions *5:*378

Plant-based remediation, future of *3:*635
  See also Phytoremediation entries

Plant biomass *3:*480

Plant biomonitors *1:*716

Plant canopies
  optical properties of *3:*720
  transpiration and *4:*346

Index terms	**Links**	
Plant growth, tile drainage and	*3:*731	
Plantings, anchored	*3:*148	
Plant leaves, optical properties of	*3:*719	
Plant metabolites, secondary	*5:*379	
Plant pathogens	*5:*379	
Plant nutrients, in ponds	*3:*486	
Plant root system, augmentation of	*5:*380	
Plants. *See also* Higher plants; Phytoextraction; Phytoremediation; Vegetables; Vegetated entries; Vegetation		
in constructed wetlands	*1:*894	
detoxification of metals in	*5:*287	
effects of cadmium on	*5:*616	
as energy sources for microbes	*5:*378	
genetically modified	*3:*686	*5:*288
in heavy metal extraction	*5:*279	
hyperaccumulator	*5:*369	
metal hyperaccumulation in	*5:*285	
metal tolerance in	*3:*609	
metal uptake mechanism in	*5:*286	
nitrate level in	*3:*637	
in rocky shores and beaches	*4:*26	
salt tolerance of	*3:*682	
as sediment traps	*3:*277	
selenium in	*5:*398	
symptoms of high salinity in	*3:*683	
tolerance to soil salinity	*3:*676	
trace elements in	*3:*456	
vulnerability to pollutant discharge	*3:*283	
wave energy	*4:*45	
Plant selection		
appropriate	*3:*556	
for hydrocarbon and metal	*3:*628	
Plant species, invasive	*3:*480	
Plastic bottles, safety of	*1:*5	
Plastic filter media	*1:*347	
Platonic solids, water symbolism and	*4:*786	
Pleasant Pond, biomanipulation trials in	*2:*53*t*	*2:*56t
PLOAD model	*2:*253	
Plug flow reactor	*1:*815	*1:*833
Plumbing, water-saving	*2:*497	

Index terms **Links**

Plumbism *2:*432
Plume behavior *5:*589
Plume diving *5:*80
  Orlando Naval Training Center *5:*82
Plume diving calculation *5:*83
Plume movement *5:*579
Plunge pool *3:*67
Plunging breakers *4:*137
Plutonium (Pu), in Chernaya Bay *4:*102
p-nitrophenol (PNP), adsorption of *1:*117
Point CIPP *1:*882
Point estimation techniques *2:*381
Point-integrating samplers *3:*398
Point-of-entry drinking water treatment devices *1:*67
  See also Point-of-use/point-of-entry (POU/POE) systems
Point-of-entry radon removal *1:*52
Point-of-use drinking water treatment devices *1:*68
Point of use/point of entry (POU/POE) devices, for arsenic treatment *5:*24
Point-of-use/point-of-entry systems *1:*378
  operation and maintenance of *1:*380
  regulations affecting *1:*379
  role of *1:*378
  safety of *1:*381
  selecting *1:*381
  types of *1:*379
Point-of-use treatment methods *2:*343
Point-of-zero charge (PZC) *2:*98
Point-pollution sources *2:*265
Point process models *3:*428
Point repairs. *See also* Point source repairs
  for force mains *1:*889 *1:*890
  for gravity sewers *1:*888
Point source contamination *5:*603
Point-source pollution *2:*545 *3:*226
  control of *3:*133
  during and after development *3:*501
  terrestrial *3:*281
Point source repairs *1:*881
Point sources *2:*169 *2:*190

Index terms **Links**

Poland
- fishpond culture in *3:*135
- pond fisheries in *4:*718

water retention methods in *1:*404 *1:*406 *1:*407

Polarity *4:*523

Polar lipid-derived fatty acid-based SIP *1:*645

Policies, water industry *1:*510
- *See also* Public policy

Policy decisions, environmental principles needed to formulate *3:*106

Political solutions, in the Arab World *2:*473

Politics, of wetland management *4:*691

Poljes *5:*245

Pollutant concentration, in gray water *1:*17t

Pollutant concentration profiles *3:*78

Pollutant dispersal, in estuaries *4:*50

Pollutant loads
- in groundwater *3:*226
- measured *3:*223

Pollutant photodegradation, using photocatalytic membrane reactors *1:*793

Pollutant removal mechanisms *3:*366

Pollutants. *See also* Chemical pollutants; Pollution
- affinity to suspended particulate matter *2:*384
- agriculturally derived *3:*604
- background concentration of *2:*18
- coastal water *4:*96
- conservative *4:*50
- cross-media transfer potential of *1:*838
- degradation by wetlands *3:*369
- effects on marine habitat *4:*105
- erosion-related *1:*537
- human health and *4:*379
- inorganic *2:*19
- microbial detection of *2:*440
- nonconservative *4:*50
- organic *2:*20
- in ponds *3:*485
- unrecognized *3:*371
- versus pollution *2:*190

Polluted discharges, into lakes *3:*281

“Polluted runoff” *3:*281

Index terms	**Links**		
Polluted waters	*1:*382		
See also Water pollution entries			
Polluter-authority negotiations	*4:*647		
''Polluter pays'' principle	*1:*589	*2:*597	*2:*638
Pollution	*1:*293		
See also Pollutants			
arsenic	*1:*1		
biotic	*3:*134		
groundwater	*2:*182		
indicator group response to	*2:*441		
large-scale	*4:*105		
long-term simulation of	*3:*335		
nitrate	*1:*30	*1:*39	*3:*637
nonpoint source	*2:*181	*2:*184	*2:*186
oil	*2:*290		
from oil-field brine	*2:*286		
pathogenic	*2:*182		
sediment-attached	*3:*333		
of surface waters	*3:*373		
trace element	*4:*109		
trace metal	*2:*64		
transformation modeling of	*3:*333		
urbanization-related	*2:*181		
vulnerability of karst areas to	*5:*247		
water	*1:*901		
water recycling and	*2:*612		
Pollution control	*2:*97		
affordable	*2:*132		
maximum extent practicable	*3:*432		
in stored water	*2:*367		
technologies for	*1:*554		
''Pollution exclusion'' clauses	*3:*169		
Pollution load, Ganga Basin	*3:*235		
Pollution loading, variations in	*2:*264		
Pollution management			
in Ho Chi Minh City	*2:*553		
tile drainage and	*3:*730		
in urban landscapes	*2:*190		
Pollution measurement			
in rainfall	*3:*223		
in runoff	*3:*225		

Index terms **Links**

Pollution monitoring programs, biomarker approach in *2:*429
‘‘Pollution offsets,’’ *2:*133
Pollution outflow, from agricultural areas *3:*657
Pollution prevention *1:*143
  landfill-related *5:*256
Pollution stress, response of biota to *1:*715
Pollution taxes *2:*132
Pollution tracers *5:*232
Poltruba backwater, biomanipulation trials in *2:*56t
Polybrominated diphenyl ethers (PBDE), as unrecognized pollutants *3:*372
Polychaetes, in marine toxicological testing *4:*42
Polychlorinated biphenyl (PCB) compounds *1:*689 *2:*545 *3:*352
  chemical properties of *2:*109t
  as environmental toxicants *2:*106
  fish consumption advisories for *3:*118
  macrophyte biomonitoring of *1:*714
  in streambed sediment *3:*357 *3:*359t
  versus DDT *3:*521
Polycyclic aromatic hydrocarbons (PAHs) *1:*571 *3:*352 *4:*98
  See also PAH compounds
  biodegradation of *1:*693
  in chironomids *2:*421
  detection limits for *1:*573t
  in Puget Sound Basin sediment and fish *3:*350
  in urban stormwater runoff *3:*435
Polycyclic organic matter (POM) compounds *1:*571
Polydex
  bactericidal efficiency of *1:*382
  test water disinfection using *1:*383
Polymerase chain reaction (PCR) *2:*139
  relative quantitation of products of *2:*59
Polymerase chain reaction assay *1:*89f *1:*160
Polymer-assisted ultrafiltration (PAUF) *1:*916 *1:*917
Polymer coagulants *1:*138
Polymer coated activated carbons (POAC) *1:*93
  in metal ion removal *1:*918
Polymeric flocculants *2:*98
Polymers
  copper removal using *1:*919
  organic *4:*426

Polymers *(Continued)*
  water-soluble *1:*919
Polymictic lakes *3:*266
Polynuclear aromatic compounds *2:*310
Polynuclear aromatic hydrocarbons (PAHs) *5:*91
Poly-P bacteria *1:*788t
Polyphosphates *1:*314
Polypropylene (PP) membranes *1*:592f *1*:593f
Polyurethane (PU) grouts *1:*881
Pond aquaculture *3:*375
Pond carp culture *3:*137
Pond construction, in Silesia *3:*135
Ponded ring infiltrometer *5:*214
Pond effluent *3:*541
Pond fisheries, in Poland *4:*718
Ponding method *5:*166
Ponds
  blue-green algae blooms in *3:*188
  classification of *3:*541
  cultivated *4:*697
  detention *3:*430
  water quality in *3:*484
Pond systems, for aquaculture production *3:*540
Pond water, color of *3:*485 *3:*486t
Pool and weir fish passage facility *3:*529
Pools, types of *3:*69t
Poor people, water for *4:*771
Poor-quality water resources, salt tolerance and management of *3:*685
Population
  in developing countries *1:*504
  effect on coastal wetland losses *3:*73
  freshwater availability and *1:*437
  Great Lakes region *3:*178
Population growth *2:*574
  efficiency and *2:*493
  impacts of *4:*289
  in the Nile Basin *2:*590
  water scarcity and *2:*495
Population growth impairment, as a sublethal toxic endpoint *2:*414
Pore size distribution, of activated carbon *1:*97

Index terms **Links**

Pore spaces *5:*444
Pore water *2:*384
  concentrations *2:*421
Porewater nutrient concentrations
  coupling with bivalve excretory activity *4:*75
  seasonality of *4:*73
Porosity *2:*545 *4:*484
  soil *3:*30
  testing *5:*186
Porosity log *5:*152
Porous cup samplers *2:*341
Porous flow *5:*655
Porous media *5:*600
Portals
  general *2:*669
  water-related *2:*668
Porter-Cologne Water Quality Act *2:*358
Portland cement, hydration chemistry of *5:*362
Portland steam ship, exploration of *4:*63
Positive-displacement meters *1:*338
Positive-displacement pumps *1:*393
Possible contaminating activities (PCAs) *1:*527
Posterior probability distribution *5:*596
Postindustrial Revolution Era, water in *4:*728
Post-treatment sewage treatment systems *1:*519
Potable wastewater reuse *1:*821
Potable water *2:*545
  augmenting supplies of *1:*819
  filter design for treating *1:*348
  inhibition of biological ammonia removal from *1:*348
  nitrification of *1:*346
  regulatory and security
  requirements for *2:*343
  reuse of *1:*826
  security requirements for *2:*347
Potassium (K)
  concentration *5:*196
  in groundwater *2:*395
Potassium perchlorate *5:*631
Potassium permanganate *1:*313 *2:*300 *5:*346

Potential additional total exposed population (PATEP) *1*:36

Potential evapotranspiration (PET) *4*:259 *5*:409

Potentiometric surface *5*:184

Potomac River Low Flow Agreement *2*:585

Powdered activated carbon (PAC) *1*:92 *1*:351 *1*:817 *2*:86

  See also Activated carbon(s)

  advantages of *2*:88

  dosages of *2*:87 *2*:88

  effect on water adsorption *4*:402

  oil-field brine and *2*:288

  in water treatment *2*:86

Powdered activated carbon contactor, analysis and design of *2*:87

Power generation, technological revolution in *4*:729

Power laws *3*:212

Power plants. *See* Electric generating plants; Hydropower

Precautionary principle *3*:125

  application to water science *2*:595

  concerns related to *2*:598

  current examples of *2*:599

  drivers and antidrivers related to *2*:598

  global elements of *2*:596

  global scope of *2*:596

  interpretation of *2*:596

  recommendations for *2*:601

  relevance to water and environmental science issues *2*:599

  risk and *2*:601

  risk tolerance and *2*:599

Precipitate flotation *1*:686

Precipitating reagents *1*:705

Precipitation *1*:811 *2*:516 *2*:545 *4*:434

  See also Rain entries

  acid *3*:225

  in arid regions *5*:409 *5*:410

  assessment of *4*:290

  of cadmium *5*:616

  chemical *4*:586

  forest interception of *3*:171

  of heavy metals *5*:378

  in India *2*:560 *2*:561f

  isohyetal method for estimating *4*:290

  of landfill leachates *1*:705

Index terms **Links**

Precipitation *(Continued)*
  of lime *5:*20
  of metals *5:*282
  remote sensing of *4:*322
  role in flooding *3:*143
  Sr isotope ratios in *4:*371
  in subsurface systems *5:*416
  throughfall of *3:*171
  tritium in *5:*67f
  in the vadose zone *5:*534
  variation in *2:*517 *2:*574
  versus streamflow *4:*433
  in the water cycle *4:*196
  water-rock interaction and *5:*569
  wet-season *4:*304
  Yellow River Basin *3:*45
Precipitation analysis, regional *4:*221
Precipitation depth, computing *4:*290
Precipitation-dissolution reactions *5:*277
Precipitation-effectiveness (P-E) index *3:*32
Precipitation estimation from remotely sensed information using artificial neural networks (PERSIANN) *4:*323
Precipitation rate effect *5:*230
Precipitation routes *4:*193
Precipitation studies, remote sensing and GIS application in *2:*532
Precoat filters *1:*245
Precoat filtration *1:*251 *1:*488
Precursors, physical removal of *2:*120
Predictive models *2:*26 *3:*40 *3:*41
  use for *3:*330
Predictive well maintenance *5:*263
Prefilters *1:*332
Prehydrolyzed coagulants *4:*425 *4:*427
Preliminary wastewater treatment processes *1:*814
Presedimentation *1:*323
  with coagulation *1:*487
Pressure
  atmospheric *4:*360
  dynamic *5:*652
  hydrostatic *3:*194

Pressure *(Continued)*
  water hammer *1:*891t
Pressure compensating drippers *3:*622
Pressure constraints, in water distribution systems *1:*209
Pressure contour plotting *3:*318
Pressure filters *1:*176 *1:*247f
Pressure head *5:*171
Pressure leak testing *1:*886
Pressure relief valves *1:*483
Pressure relief wells *3:*289
Pressure supplies, in water distribution system design *1:*212
Pressure transducer probe *5:*101
Pressure transducers *5:*474
Pressure transients *1:*213 *1:*343
Pressurized filters
  advantages and disadvantages of *2:*155
  for iron removal *2:*154
Pressurized flow, hydraulics of *3:*196
Pretreatment, for lime softening *1:*323
Pretreatment effluent standards *4:*665
Pretreatment Program (EPA) *1:*798
Pretreatment roughing filter *1:*237
Preventive maintenance (PM) *2:*595
  of wells *5:*263
Price elasticity, of water demand *4:*606
Pricing, tiered *2:*615
  See also Cost(s); Economics; Water pricing
Primary raw sludge *1:*861
Primary settling tanks *1:*452 *1:*455 *1:*456
  design considerations for *1:*455
Primary wastewater treatment *1:*809 *1:*815 *1:*827 *2:*545
Primary water quality standard *2:*373
Principle of maximum entropy (POME) *4:*217 *4:*218
Principle of minimum cross entropy (POMCE) *4:*219
Prior Appropriation Doctrine *2:*545 *4:*628 *4:*631
  water transfers under *4:*686
Prioritized objectives method *2:*623
‘‘Priority elements,’’ effect on fish *3:*131
‘‘Priority Pollutant’’ list (EPA) *2:*304 *2:*305t *4:*599

Priority pollutants *3:*371
Priority system rule *4:*631
Private goods *1:*42
Private Participation in Infrastructure (PPI) database *1:*46
Private sector participation (PSP) *1:*50
  extent of *1:*389
  forms of *1:*388
  history of *1:*387
  in municipal water supply and sewerage *1:*387
Private value of water *2:*653
Private well owners, in Arab states *2:*471
Privatization
  increase in *1:*508
  of water *2:*569
  in the water industry *1:*302
“Privatization by default” *2:*632
Proactive groundwater pumping *5:*45
Probabilistic approach *2:*675
  to incipient motion *3:*418
  to studying sediment transport *3:*401
Probabilistic-topological stream order approach *3:*32
Probability distributions, derivation of *4:*221
Probable effects level (PEL) *3:*351
Probable maximum precipitation (PMP) analyses *4:*309
Probable maximum precipitation rates *4:*315
Problematic weeds *3:*479
Problem definition, in a systems approach *2:*685
Process-based computer simulations *5:*563
Process-based pesticide vulnerability models *5:*596
Process solid waste, as a source of contaminants *5:*433
Production efficiency, biofouling effects on *5:*35
Productivity, lake *3:*267
Product testing *1:*381
Professional water resource organizations *4:*651
Profitability measures *1:*312
Profundal *3:*267
Programs, conservation *2:*493
Propeller meters *1:*338
Property rights *2:*500 *2:*506
  in India *2:*556

Index terms **Links**

Property transfer issues *3:*169
PROSPECT model *3:*720
PROSPER model *3:*738 *3:*739
Protection zones *1:*316
Protective coatings *1:*882
Protective irrigation systems *3:*583
Protective linings, corrosion and *1:*11
Protein array technologies *2:*61
Protein expression analysis, high throughput *2:*61
Protein phosphatase inhibition assay *2:*389
Proteins
  as indicators of organic contamination *2:*442
  role in metal transport *5:*286
Proteomics *2:*61 *2:*429
Protists, as toxicity assessment test organisms *2:*413
Protonation *5:*377
"Proton reducers" *5:*40
Prototype measurement systems *4:*296
Protozoa *1:*904 *2:*313
  in domestic sewage *1:*831
  health effects of *1:*278
  removal of *1:*485
Protozoal agents *1:*178
Protozoan parasites *1:*166
PRZM code *5:*564
Pseudomonas *1:*278 *2:*241
Pseudoplastic fluids *5:*557
Pseudosolubilization *2:*44
Psychrometers *5:*540
Psychrometry *5:*462
PT benzene (propylene tetramer benzene sulphonate) *1:*671
Public. *See also* Public health entries
  participation in municipal storm water management by *1:*867
  providing real-time hydrological information to *1:*123
Public agency activities, effect on storm water pollution *1:*867
Public awareness/concerns, growth of *1:*304
Public concerns, precautionary principle and *2:*599
Public education *3:*503
  in Canada *2:*663
  in groundwater protection *2:*518

Public education *(Continued)*
  strategies for *2:*519
Public goods *1:*42
Public health. *See also* Public health protection
  biofouling and *5:*36
  programs *1:*799
  water reuse research to protect *1:*819
Public health goal (PHG) *2:*345
Public health protection
  chemical revolution in *1:*286
  in developing countries *1:*290
  drinking water and *1:*281
  early water treatment *1:*286
  history of *1:*282
  water as an issue in *1:*283
Public Health Security and Bioterrorism Preparedness and Response Act *1:*158 *1:*476 *1:*527 *2:*347
Public Health Service (PHS) *2:*194
Public Law 660 *4:*653
Public Law 92-500 *4:*653
Publicly Owned Treatment Works (POTWs) *1:*755 *1:*757 *1:*798 *1:*801 *2:*96 *4:*596 *4:*597 *4:*670
Public notification, of repairs *1:*891
Public Notification Rule *2:*347
Public outreach, role in fish consumption advisories *3:*120
Public policy, New York City Harbor Survey and *1:*750
  See also Policies
Public-private partnerships (PPPs) *1:*42
  challenges to *1:*46
  current *1:*46
  defined *1:*43
  durable *1:*45
  obstacles to forming *1:*45
  start-up of *1:*43
  successful *1:*44
  trust creation by *1:*47
Public programs, for agricultural land-use planning *3:*596
Public safety, pretreatment programs and *1:*800
Public supply *2:*545
Public trust, water managed in *2:*608
Public trust doctrine *4:*661

Index terms **Links**

Public water supply. *See also* Public water systems (PWSs)
  conflicts and needs related to *1:*503
  data and indicators related to *1:*505
  defined *1:*501
  disinfection of *2:*118
  future of *1:*290
  global *1:*500
  national differences in *1:*501
in the United States *2:*651
Public water systems (PWSs) *1:*201 *1:*476 *1:*524 *2:*96 *2:*371 *4:*669
  regulation of *4:*676
Public water use *2:*545
Puget Sound Basin
  data collection from *3:*350
  organic compounds and trace elements in fish in *3:*349
Pulp industry, impact on surface water quality *3:*374
Pulsed bed *1:*104
Pulse microirrigation *3:*621
Pulses, in groundwater tracing *5:*504
Pump and treat technology *5:*41
Pumped storage facilities *3:*201
Pumping rates, equalization of *3:*380
Pumping stations *3:*379
  power in *3:*380
  pumps in *3:*379
  water inlet in *3:*379
Pumping tests *5:*185 *5:*513
  practical aspects of *5:*496
Pumping water level *5:*101 *5:*104
Pumps *1:*203
  centrifugal *1:*392
  positive-displacement *1:*393
  priming of *1:*393 *1:*394
  reciprocating *1:*394
  rotary *1:*394
  safety of *1:*394
  submersible *1:*393 *3:*664
  terms related to *1:*392
  types of *1:*392
  vertical turbine *1:*393

- Pumps *(Continued)*
  - water distribution system *1:*208 *1:*391
  - water hammer and *1:*262
- Pump samplers *3:*400
- Pump-test packages *5:*663
- Purchase of agricultural conservation easements (PACE) programs *3:*596 *3:*597
- Pure water *5:*602
- Purgeable halocarbons, gas chromatogram of *2:*306f
- Purge-and-bail sampling *5:*459
- Purge and trap procedure *2:*306
- Purging, low flow groundwater *5:*404
- ‘‘Put-and-take’’ fisheries *3:*128
- PVC repair clamps *1:*889
- Pyrites *1:*609
  - *See also* Iron pyrite (FeS_2)
  - acid production and *3:*14
  - weathering of *2:*1
- Pyrometallurgical extraction technologies *5:*435

Q

- Qanats *5:*483
  - constructing *5:*485
  - costs of *5:*486
  - history and geographic extent of *5:*484
  - statistics of *5:*487
  - versus drilled wells *5:*486
  - water regulations and *5:*486
- Qochas *4:*696
- QSAR approach, in dye tracer selection *3:*98
  - *See also* Quantitative structure-activity relationships (QSARs)
- QSAR models
  - soil sorption using *3:*99
  - using triarylmethane dyes *3:*99
- Quabbin Reservoir *1:*316
- Quality. *See also* Distribution system water quality; Water quality
  - desalinated water *1:*173
  - feed water *1:*175
  - gray water *1:*16t
  - irrigation water *2:*155
  - raw water *1:*440

Index terms	Links	
Quality assessment. *See* Environmental quality assessment; Quality control		
Quality control		
in amphipod sediment toxicity tests	*2*:410	
of groundwater samples	*5*:455	
Quality criteria, for raw waters	*3*:227t	
See also Quality parameters		
Quality parameters, of water	*1*:901	
See also Quality criteria		
Quality standards, drinking water	*1*:476	
Quantitative foodwebs	*4*:152	
Quantitative groundwater law	*4*:627	
Quantitative microautoradiography (QMAR)	*1*:847	
Quantitative reverse transcription polymerase chain reaction (qRT-PCR)	*2*:59	
Quantitative structure-activity relationships (QSARs)	*2*:415	*3*:95
See also QSAR entries		
Quantity-based reliability	*3*:261	
Quartz sand	*1*:347	
Quasi-conservative pollution, modeling of	*3*:332	
Quasi-Nonhydrostatic (QNH) model	*4*:247	
Quasi-riparian doctrines	*4*:630	
"Quickflow"	*5*:236	

R

Index terms	Links	
Race-to-pump mentality	*4*:632	
Radar		
attenuation of	*4*:308	
development of	*4*:354	
ground penetrating	*1*:884	
polarimetric measurements with	*4*:311	
reflectivity factor of	*4*:322	
use in rainfall measurements	*4*:306	*4*:310
Radial impellers	*1*:76	
flow patterns induced by	*1*:77f	
Radially entrained vapor extraction (REVEX)	*2*:286	
Radial wells	*3*:595	*5*:407
Radiation. *See also* Ultraviolet (UV) radiation		
fog and	*4*:231	
radioactivity and	*1*:802	
Radiation balance	*3*:561	
Radiation fog variables	*4*:232	

Index terms	Links	
Quality assessment. *See* Environmental quality assessment; Quality control		
Quality control		
in amphipod sediment toxicity tests	*2*:410	
of groundwater samples	*5*:455	
Quality criteria, for raw waters	*3*:227t	
See also Quality parameters		
Quality parameters, of water	*1*:901	
See also Quality criteria		
Quality standards, drinking water	*1*:476	
Quantitative foodwebs	*4*:152	
Quantitative groundwater law	*4*:627	
Quantitative microautoradiography (QMAR)	*1*:847	
Quantitative reverse transcription polymerase chain reaction (qRT-PCR)	*2*:59	
Quantitative structure-activity relationships (QSARs)	*2*:415	*3*:95
See also QSAR entries		
Quantity-based reliability	*3*:261	
Quartz sand	*1*:347	
Quasi-conservative pollution, modeling of	*3*:332	
Quasi-Nonhydrostatic (QNH) model	*4*:247	
Quasi-riparian doctrines	*4*:630	
"Quickflow"	*5*:236	
R		
Race-to-pump mentality	*4*:632	
Radar		
attenuation of	*4*:308	
development of	*4*:354	
ground penetrating	*1*:884	
polarimetric measurements with	*4*:311	
reflectivity factor of	*4*:322	
use in rainfall measurements	*4*:306	*4*:310
Radial impellers	*1*:76	
flow patterns induced by	*1*:77f	
Radially entrained vapor extraction (REVEX)	*2*:286	
Radial wells	*3*:595	*5*:407
Radiation. *See also* Ultraviolet (UV) radiation		
fog and	*4*:231	
radioactivity and	*1*:802	
Radiation balance	*3*:561	
Radiation fog variables	*4*:232	

Index terms	**Links**		
Radiative forcing, global	*4:*176		
Radiative opacity	*4:*190		
Radioactive contaminants	*5:*221		
See also Radioactive contamination			
Radioactive contaminant transport, numerical modeling of	*3:*322		
Radioactive contamination. *See also* Radioactive contaminants			
drinking water limits for	*1:*803		
manmade sources of	*1:*803		
Radioactive isotopes. *See also* Radionuclides			
environmentally important	*5:*221t		
naturally occurring sources of	*1:*802		
sources of	*5:*221t		
Radioactive waste	*1:*802		
activated carbon treatment for	*1:*804		
air stripping for	*1:*803		
chemical precipitation of	*1:*804		
ion exchange for	*1:*804		
lime softening of	*1:*804		
membrane filtration processes for	*1:*803		
treatment technologies for	*1:*803		
as water pollution	*4:*101		
Radioactivity	*5:*184	*5:*186	
Radioactivity mapping	*5:*186		
Radiocarbon, groundwater dating with	*5:*64		
See also Carbon-14; Carbon isotopes			
Radiocesium levels	*4:*103		
Radiographic corrosion measurement	*1:*9		
Radiometric resolution	*4:*320		
Radionuclide monitoring	*2:*267		
Radionuclides	*1:*395	*1:*803	*5:*221
See also Radioactive isotopes			
health risks of	*1:*395		
regulations governing	*1:*396t		
technologies for	*1:*465t		
treatment technologies for	*1:*396		
Radionuclide standards, U.S.	*2:*268		
Radionuclide study	*4:*581		
Radiosondes	*4:*349		
Radium (Ra), health risks of	*1:*395		
Radium-barium sulfate co-precipitation	*1:*398		

Index terms **Links**

Radon (Rn)
  case-control studies *4:*544
  in domestic water *4:*542
  fatalities from *4:*544
  health risks of *1:*395
  reducing in drinking water *1:*51
  as a tracer *5:*683
  treatment technologies for *4:*546
  in water *4:*541
Radon risk policies *4:*542
Rain *4:*293
Rain-acidified waters
  inland *3:*8
  regional distribution of *3:*8
Rain acidity, neutralization of *3:*7
Rainfall *4:*309
  See also Rainwater
  as energy spectrum *4:*307
  estimation by isohyetal method *4:*292t
  flooding induced by *3:*160
  in India *2:*559 *2:*561f
  intensity-duration relationship for *3:*243
  observation and estimation of *4:*315
  pollution measurement in *3:*223
  radar measurement of *4:*310
  remote sensing of *4:*322
  runoff and *4:*315
  satellite monitoring of *4:*311
Rainfall data *5:*164
Rainfall estimates, Chicago Station *5:*79t
Rainfall measurements
  errors in *4:*297
  radar use in *4:*306
Rainfall partitioning, at the land surface *4:*315
Rainfall plot, coverage simulation of *4:*331
Rainfall rates *4:*486
  computed *4:*308
Rainfall recharge (*Rr*) *5:*165
Rainfall-runoff models *4:*357 *5:*74
  input data and model parameter uncertainties in *4:*298
  ‘‘lack of knowledge’’ uncertainty in *4:*300

Index terms	**Links**	
Rainfall-runoff models *(Continued)*		
observation and model errors in	*4:*300	
propagation of input errors through	*4:*298	
residuals in	*4:*299	
serial dependence in	*4:*300	
simplification of	*4:*297	
uncertainties in	*4:*297	
Rain-fed agriculture	*3:*560	
Rain forests	*4:*239	
threats to	*4:*240	
Rain gages	*4:*309	
See also Rain gauges		
Rain gauge network density	*4:*298	
Rain gauges, Bangkok Metropolitan Administration	*1:*124f	
See also Rain gages		
Rain-on-snow floods	*3:*511	
‘‘Rain shadow’’ effect	*4:*293	
Rain-simulator equipment	*4:*330	
Rain simulators	*4:*330	
rainfall intensity used in	*4:*332	
types of	*4:*330	
typical measurements by	*4:*331	
Rainsplash	*3:*566	
Rain terminal velocity, simulation of	*4:*331	
Rainwater	*2:*95	*4:*285
See also Rainfall		
chemistry of	*4:*371	
composition of	*4:*285	
technetium in	*4:*580	
Rainwater collection gutters	*1:*56	
Rainwater harvesting	*2:*548	
components of	*2:*549	
defined	*2:*549	
need for	*2:*548	
structures related to	*2:*550	
techniques for	*2:*551	
Rainwater systems, conventional	*1:*58	
Raised agriculture	*4:*697	
Raised gardens	*4:*698	
Rajarata, ancient ecosystems in	*4:*774	

Index terms	Links	
Ramganga dam	*3:*234	
Random-coordinate basin slope method	*3:*29	
Random metering	*1:*494	
Random search method	*2:*333	
Range analysis	*3:*261	
Ranney collector wells	*5:*88	
Rapid Bioassessment Protocols (RBPs)	*2:*25	
Rapid BOD technologies	*2:*40	
See also Biochemical oxygen demand (BOD)		
Rapid granular bed filters	*1:*234	
Rapid granular bed filtration	*1:*244	
Rapid gravity filter box	*1:*371f	
Rapid gravity filtration	*1:*245	
Rapidly draining epikarsts	*5:*236	
Rapid ^{18}O exchange, measurements of	*4:*536	
Rapid rate pressure filters	*1:*247	
Rapids	*3:*70	
Rapid Update Cycle (RUC-2) model	*4:*248	
Rate increases	*1:*147	
Rate of time of preference	*4:*607	
Rating curve	*2:*545	*4:*222
Rationale wastewater model	*1:*731	
Raw sewage	*3:*445	
Raw water, quality and testing for	*1:*440	*3:*227t
rBAT-induced amino acid transport	*2:*436	
Reaction kinetics		
effect of heat on	*5:*173	
oxygen demand and	*1:*640	
Reaction-path models	*5:*139	
Reaction phenomena	*3:*78	
Reaction turbines	*3:*488	
Reactive intermediates, in the irradiation of natural waters	*4:*531	
Reactive solutes	*5:*274	
Reactive solute transport, in soil and groundwater	*5:*524	
Reactive transport, in the saturated zone	*5:*518	
Reactive transport models		
equilibrium and kinetic	*5:*519	
hardware for	*5:*34	
Reactive transport simulator	*5:*519	

Index terms	**Links**	
Reactors, POP behavior in	*1:*767	
Reactor systems, anaerobic	*1:*906	
Real-time control, modeling	*3:*340	
Real-time hydrological information		
future of	*1:*126	
providing to the public	*1:*123	
urban drainage models and	*1:*125	
weather radar for	*1:*126	
Real-time hydrological information system	*1:*121	
Real-time measurements, observations of climate and global change from	*4:*172	
Real-time modeling	*3:*340	
Real-time polymerase chain reaction	*1:*160	
Real-time reservoir operation	*3:*387	
Reasonable Use Doctrine	*4:*628	*4:*631
Receiving waters, assessment of	*4:*103	
Recession constants	*3:*23	
base flow	*3:*24	
Receiving water body samples, analyses of	*1:*662	
Recession-selecting algorithms	*3:*25	
Recharge		
in arid regions	*5:*408	
defined	*2:*545	
in desert regions	*5:*72	
estimation of	*5:*411	
flood spreading for	*3:*164	
incidental	*5:*75	
measurement techniques for	*5:*73	
methods of	*5:*12	
in regional flow systems	*5:*417	
subsurface	*5:*73	
of unconfined aquifers	*5:*11	
in the vadose zone	*5:*534	
Recharge areas	*5:*601	
Recharge capacity, impact of air on	*5:*421	
Recharge coefficient, establishment of	*5:*168	
Recharge/discharge boundaries	*5:*184	
Recharge interactions	*4:*371	
Recharge zones	*4:*670	*5:*417
Recharging, excess	*5:*12	
See also Recharge		

Index terms **Links**

Reciprocal SSH libraries *2:*62
Reciprocating pumps *1:*394
Recirculating aquaculture systems *1:*683 *3:*544
Recirculating systems
  enclosed *1:*558
  open *1:*559
Reclaimed irrigation
  human exposure to *3:*670
  microbial quality of *3:*667
  types of practices of *3:*669
Reclaimed wastewater *2:*545 *2:*649
Reclaimed water *1:*805
  benefits of *1:*805
  constituents of *1:*806 *1:*807t
  federal guidelines for *1:*806
  planning for *1:*807
  technical considerations related to *1:*808
  uses of *1:*805
Reclamation
  in acid mine drainage *5:*3
  of waterlogged lands *3:*603
Reclamation Drought Index (RDI) *4:*213
Recombinant bacteria *2:*454 *2:*455
  bioluminescent *2:*174
Recombinant luminescent bacterial biosensors *2:*454
Recording, of hydrogeologic responses to earthquakes *5:*114
Recovery/removal treatments, for odor abatement *1:*762
Recovery tests *5:*575
Recreation, algae-impaired *3:*108
Recreational fishing *3:*123
Recreational wastewater reuse *1:*818
Rectangular settling tanks *1:*453
Recycled water *2:*545
  in the United States *2:*613
Recycling, on the International Space Station *4:*572
Recycling systems *1:*509
Red Book *2:*315
Red edge reflectance-ratio indexes *3:*722
Redeposited sediment, contaminant release from *2:*125
Redfield ratios *4:*52

Index terms	Links	
‘‘Redfield’’ stoichiometric ratio	*3:*109	
Redox buffering capacity	*2:*468	
Redox chemistry, sub-surface	*5:*413	
Redox couple approach	*5:*414	
Redox gradients	*2:*213	
Redox phenomena, alternative methods of examining	*2:*468	
See also Oxygen redox potential (ORP); Reduction-oxidation (redox) reaction		
Redox potentials, use in wastewater treatment	*1:*399	
See also E_h; Oxidation-reduction (redox) reactions		
Redox reactions	*5:*569	*5:*583f
See also Oxidation and reduction (redox)		
ranges of E_h for	*2:*468	
‘‘Red-rot’’ disease of water	*2:*147	
Red tides	*4:*55	*4:*134
Reduced pressure principle backflow preventer (RPBP)	*1:*155	
Reduction chemistry	*5:*291	
Reduction-oxidation (redox) reaction	*2:*464	
See also Redox entries		
Reduction reactions, in mine effluent remediation	*1:*611	
Reductive dechlorination	*2:*299	
Reductive dehalogenation	*1:*689	
Reed bed	*1:*787	
Reef breakwater	*4:*15	
Reef fish research	*4:*56	
Reference crop evapotranspiration	*2:*537	*3:*572
estimating	*3:*573	
Reference dose (RfD)	*1:*426	
Reference site criteria	*2:*26	
Reference toxicant tests	*2:*411	
Reflectance	*4:*321	
Reflectivity images, “real-time”	*4:*308	
Refrigerants	*4:*420	
Refuse decomposition, microbiology of	*1:*695	
Regional flow systems	*5:*417	
boundary conditions in	*5:*418	
characteristics of	*5:*419	
discharge in	*5:*418	
recharge in	*5:*417	
Regional precipitation analysis and forecasting	*4:*221	
Regional watershed protection	*2:*613	

Index terms **Links**

- Regional water supplies *2:*613
- Regional water treatment *2:*615
- Regional water use *2:*489
  - in the United States *2:*646
- Registry of Equipment Suppliers of Treatment Technologies for Small Systems (RESULTS) database *1:*154 *1:*315 *1:*325 *1:*461
- Regression methods *2:*253
  - in odor analysis *2:*282
- Regulated riparianism *4:*632
- Regulated Riparian Model Water Code *4:*629
- Regulated rivers *3:*106 *3:*381
- Regulation. *See also* Regulations; Water regulation
  - of bottled water *1:*4
  - groundwater *4:*632
  - of mercury *5:*645
  - of nitrate *5:*629
  - of power plant wastewater
  - discharges *1:*557
  - of public-private partnerships *1:*44
  - of public water systems *4:*676
  - of technetium *4:*580
  - of urban stormwater runoff *3:*432
  - using biomarkers and bioaccumulation for *2:*429
  - of vinyl chloride *5:*637
  - of wastewater treatment *1:*844
  - of water markets *2:*500
- Regulation and enforcement agencies federal *4:*650t
- Regulations. *See also* Effluent water regulations; Federal regulations; Laws; Regulation
  - arsenic-related *1:*82
  - broadening of *1:*508
  - bromide-related *2:*74
  - Chinese *2:*487
  - Clean Water Act *1:*756
  - consumer information *4:*678
  - detergent-related *1:*672
  - disinfection-related *1:*197 *2:*92
  - drinking water *1:*53 *1:*422 *2:*194 *2:*345
  - enforcement of *1:*508
  - filtration-related *1:*227
  - introduction of *1:*287

Index terms	**Links**	
Regulations *(Continued)*		
membrane technologies and	*1:*331	
metal-discharge-related	*2:*68	
nitrate/nitrite-related	*3:*639	
nonpoint source control	*2:*185	
ozone	*1:*354	
pharmaceutical-related	*1:*377	
point-of-use/point-of-entry system	*1:*379	
radionuclide	*1:*396t	
risk assessments as	*1:*428	
Subtitle D	*2:*166	*2:*167
sugarcane industry	*1:*619t	
water quality	*1:*282	*2:*315
water quantity	*2:*557	
water supply degradation and	*2:*401	
water sustainability and	*2:*614	*2:*627
for water treatment plant residuals	*1:*413	
Regulatory agencies, early	*1:*285	
Regulatory controls, growth of	*1:*305	
Regulatory impact analysis (RIA)	*1:*427	
Regulatory interface, federal-state	*4:*671	
Regulatory issues	*4:*674	
iron-related	*4:*497	
Regulatory purposes		
sediment toxicity tests for	*2:*383	
test development for	*2:*458	
Regulatory requirements, drinking water filtration	*1:*233	
Regulatory reviews, of NPDWRs	*1:*428	
Relative humidity	*4:*270	
vapor pressure and	*4:*271	
Relative humidity statistics, Chicago Station	*5:*79t	
Relative risk concept	*1:*287	
Release fractures	*5:*177	
Reliability		
defined	*3:*260	
in reservoir design	*3:*259	
Relief drains. *See* Tile drainage		
Relief valves	*1:*262	
Remedial Action Plans (RAPs)	*3:*179	*4:*622

Index terms	**Links**
Remedial investigation/feasibility study (RIFS)	*2:*318
Remediation. *See also* Phytoremediation	
of cadmium	*5:*617
of contaminated soils	*5:*432
of fluoride contamination	*5:*134
of lead	*5:*648
limiting geochemical factors in	*5:*319
of mercury	*5:*644
of MtBE	*5:*319
of metal-contaminated soils	*5:*433
of selenium	*5:*397
of selenium-contaminated waters	*2:*355
of technetium	*4:*581
of uranium	*5:*641
Remediation injection process (RIP®)	*5:*234
Remediation target setting	*2:*355
Remining	*5:*3
Remotely sensed data	
analyzing	*4:*319
methodologies for analyzing	*4:*319
Remotely sensed energy balance (EB)	*3:*562
Remote monitoring	*1:*381
Remote sensing	*5:*411
advantages of	*3:*561
of algal blooms	*4:*55
alternative applications of	*4:*325
application in water resources	*2:*531
in command area studies	*2:*533
crop water stress detection using	*3:*719
of evapotranspiration	*4:*321
flood mapping via	*4:*322
in floodplain management	*2:*533
future need for	*2:*535
in groundwater studies	*2:*534
in hydrologic modeling	*2:*535
of hydrology applications	*4:*319
in land use classification	*2:*532
in precipitation studies	*2:*532
of rainfall and snowfall	*4:*322
in reservoir sedimentation	*2:*534
reservoir surveys and	*3:*411

Index terms **Links**

Remote sensing *(Continued)*
  resolution and spectral regions in *4:*319
  role in nonpoint source pollution assessment *3:*658
  in snow cover studies *2:*533
  in soil moisture assessment *2:*533 *4:*323 *4:*488
  for submarine gas hydrates *4:*59
  of surface snow *4:*324
  terms and equations related to *4:*321
  of waterbodies *4:*324
  in waterlogging and soil salinity *2:*534
  in water quality studies *2:*534
  in watershed mapping and monitoring *2:*533
  of wetlands *4:*325
Remote sensor technologies *1:*147
Remote telemetry units (RTUs) *1:*450
Removal processes *1:*485 *5:*618
  for bacteria, viruses, and protozoa *1:*485
  principal *1:*487
  as a remediation strategy *2:*358
  supplemental *1:*488
Removal technologies *5:*439
Renal system, effects of lead on *2:*436
Renewable aquatic resources *3:*121
Renewable energies *1:*562 *4:*44
  ocean thermal *4:*47
  tidal energy *4:*46
  wave energy *4:*45
Renewable energy projects, drawbacks of *4:*47
Renewable freshwater resources *2:*634
Renewable natural resources *2:*634
Renewable water resources, in Arab states *2:*471t
Repair clamps *1:*889
Repair couplings *1:*889
Repairs
  point source *1:*881
  public notification of *1:*891
Repair systems, structural *1:*883
REPIDISCA Technical Dissemination Sheets *2:*672
Reporter genes, bioluminescent *2:*453
Reporting agencies *1:*293

Index terms	**Links**			
Reports, consumer confidence	*1:*145			
Representative elementary area (REA)	*5:*306			
Representative elementary volume (REV)	*5:*306			
Reproducing and feeding ecology, classifying organisms by	*3:*37			
Reproduction, effects of lead on	*2:*436			
Research				
on agricultural drainage ditches	*3:*88	*3:*89t	*3:*90t	
on aquatic macrophytes	*2:*66			
biodegradation	*2:*42			
biomarker	*2:*30			
on coastal fisheries and habitats	*4:*55			
on greenhouse gases	*3:*180			
on hydrologic thresholds	*3:*231			
hydropsychology	*4:*735			
limnological	*3:*291			
lysimeter	*5:*487			
on microirrigation	*3:*626			
for military applications	*4:*295			
on phytoremediation	*3:*368	*5:*285		
related to coastal wetlands	*3:*73			
tsunami	*4:*162			
on wastewater reclamation and reuse	*1:*819			
Research and inventory agencies, federal	*4:*650t			
Research on Watering among Ancients	*4:*736			
irrigation and governance in	*4:*737			
Reserved water rights	*4:*689			
Reservoir design				
reliability concepts in	*3:*259			
simulation procedure for	*3:*263			
traditional	*3:*261			
Reservoir drawdown	*3:*412			
Reservoir dynamics	*3:*259			
Reservoir flood storage	*3:*151			
Reservoir flow routing	*3:*255			
Reservoir management	*3:*408			
systems engineering for	*3:*386			
Reservoir operations models	*2:*623			
Reservoirs	*1:*405	*2:*545	*2:*367	*2:*568
	*3:*270			
See also Multipurpose reservoirs; Small water reservoirs				
boreal	*3:*180			

- Reservoirs *(Continued)*
  - classification of *1:*406t
  - ecosystems below *1:*406
  - fisheries management and *3:*125
  - functions of *1:*408
  - hydroelectric *3:*180 *3:*202
  - locating *1:*410
  - need for *3:*382
  - priority concept for *3:*386
  - sediment inflow into *3:*411
  - trap efficiency of *3:*409
  - tropical *3:*181
- Reservoir sedimentation *3:*408
  - *See also* Reservoir sediments
  - factors influencing *3:*409
  - loss of storage capacity and *3:*409
  - remote sensing and GIS application in *2:*534
- Reservoir sediments
  - distribution of *3:*410
  - unit weight of *3:*410
- Reservoir site, permeability investigation for *5:*356
- Reservoir size, inflow characteristics and *3:*262
- Reservoir storage zones, design of *3:*264
- Reservoir surveys *3:*411
- Reservoir system, components of *2:*367
- Residential meters *1:*340
- Residential water demands *1:*12
  - in low-income countries *1:*14
  - management of *1:*15
- Residential water use
  - categories of *2:*652
  - models of *1:*13
- Resident time *1:*103
- Residual method of price estimation *4:*608
- Residual sodium carbonate (RSC), irrigation water quality and *5:*208
- Resistance soil moisture measurement *4:*488
- Resistivity, measurement of *5:*444
- Resistivity logs *5:*152

Index terms **Links**

Resistivity methods *5:*443
  in groundwater pollution studies *5:*447
  theory and application of *5:*444
Resource Conservation and Recovery Act (RCRA) *2:*163
Resources. *See also* American water resource development; Coastal resources; Community- based resource management; Flux-type natural resources; Freshwater resources; Global water resources; Groundwater resources; Integrated water resources management (IWRM); Natural resources; Renewable entries; Water resource entries
  public-private partnership *1:*45
  wastewater *1:*676
Respiratory problems *4:*379
Respirographic biosensor, industrial effluent examination using *1:*568
Respirometers *2:*40
  types of *1:*567
Respirometric biosensors *1:*565
  characteristics of *1:*570
Respirometric techniques, toxicity assessment using *1:*566
Response warning systems, flood-related *3:*148
Response watershed functions *3:*477
Restatement rule *4:*631
Retarding basins *3:*150
Retention, higher water *1:*406
Retention basins (RBs) *1:*783
Retention storage *3:*474
Retention systems, for storm water treatment *1:*868
Retention time *1:*103
Retrofitting *2:*665
Return activated sludge (RAS) *1:*454 *1:*456
Return flow *2:*545
Reverse electron transfer (RET) assay *2:*279 *2:*280 *2:*376 *2:*377
Reverse osmosis (RO) *1:*298 *1:333t* *1:*335 *1:*380 *1:*459 *1:*488 *1:*707 *1:*810 *2:*545 *5:*326
  See also Reverse osmosis membranes; Seawater RO plant
  antifoulant design and *1:*415
  for arsenic removal *1:*637 *5:*21
  in chromium treatment *5:*25
  desalination *1:*171
  efficiency of *5:*327
  feedwater chemistry and *1:*414

Index terms **Links**

Reverse osmosis (RO) *(Continued)*
  membrane materials in *5:*327
  oil-field brine and *2:*289
  process chemistry of *1:*414
  radionuclide removal via *1:*398
Reverse osmosis membranes
  cleaning *1:*416 *1:*419
  foulants in *1:*416
  fouling of *1:*415
  materials for *5:*26
Reverse osmosis membrane vessel *1:*172f
Reverse osmosis systems *1:*297
  choosing *5:*27 *5:*328
  design of *5:*25
  efficiency of *5:*26
  maintenance of *5:*26 *5:*328
  for nitrate treatment *5:*326
Reverse sample genome probing (RSGP), in community analysis *1:*644
Reverse transcription polymerase chain reaction (RT-PCR) *1:*161
Reversible nonspecific toxicity *2:*415
Reversibly adhering (rolling) cell formation *2:*240
Revised Universal Soil Loss Equation (RUSLE) *2:*251
Reynolds Number (Re) *5:*650
Rheopectic fluids *5:*557
Rhepoxynius abronius, sediment toxicity tests using *2:*410
Rhizofiltration, for lead-contaminated soils *5:*383
Rhizofiltration agents *3:*630
Rhodamine WT *3:*96
Ribosome content, cellular *2:*230
Rice, flooding tolerance of *3:*601
Rice growing, floating *4:*696
Richards equation *4:*486 *5:*538
Richardson, Lewis Fry *4:*350
Riffles *3:*70
Rights, water-related *4:*769
Rigid porous fine bubble diffusers *1:*626
Rills *3:*183
Ring infiltrometer *4:*487
Ring tests *2:*459
Rio Grande Basin *2:*589

Rio Tinto, acidity of *3:*10
Riparian buffers *3:*391
Riparian buffer zones (RBZs) *3:*391
  See also Riparian zone
Riparian doctrine, water transfers and *4:*686
Riparian habitats. *See also* River entries
  wastewater discharge and *2:*475
  water recycling and *2:*612
Riparian marshland. *See* Cienega entries
Riparian rights systems, transfers under *4:*685
  See also Riparian water rights
Riparian systems *3:*390
  inputs, outputs, and transformations related to *3:*390
  vegetated buffers in *3:*391
Riparian water rights *2:*501 *2:*546
Riparian weeds *3:*743
Riparian zone *3:*172
  See also Riparian buffer zones (RBZs)
Risk(s). *See also* Security
  public perception of *4:*675
  of radon in domestic water *4:*543
  of waterborne transmission of *Escherichia coli* *1:*429
Risk analysis/assessments
  of buried wastes from electricity generation *5:*448
  in contaminated groundwater bioremediation *5:*38
  in crafting drinking water regulations *1:*422
  entropy and *4:*220
  health risk reduction and cost analysis *1:*427
  of heavy metals *5:*278
  as regulations *1:*428
  Safe Drinking Water Act and *1:*424
  toxicological methods for *2:*278
Risk-based corrective action (RBCA) *5:*586 *5:*587
Risk-based decision making *4:*675
Risk-based fish consumption advisories *3:*119
Risk-based screening levels (RBSLs) *4:*675
Risk-informed decision making *2:*599
Risk management. *See* Flood risk management
Risk sharing, flood-associated *3:*512
Risk tolerance *2:*599

Ritchie second-order equation *4:*566
River and reservoir operations optimization models *2:*623
River basin decision support systems *2:*619
River basins *3:*28
  characteristics of *3:*28
  Indian *2:*560 *2:*562*t* *2:*564t
  influences on *3:*221
  as management units *3:*128
  planning and coordination related to *3:*33
  transformation of *4:*729
Riverbed/seabed filtration intake systems *5:*89
River chemistry, global *2:*19
River-connected aquifers *5:*677
  groundwater geochemistry of *5:*684
  water travel times in *5:*681
“River Continuum Concept” *3:*292
River data *5:*164
River discharge measurements *2:*622
River flow, runoff process controls of *3:*454
River flow regime grouping *4:*222
River for Jaffna proposal *4:*776
River health, predictive models for *3:*40
River-induced flooding *3:*160
Riverine ecosystems, ecological integrity of *3:*106
Riverine electromagnetic/seismic surveys *5:*679
Riverine fisheries management *3:*125
River inflow, as a forcing factor *3:*322
River Invertebrate Prediction and Classification System (RIVPACS) *3:*40
River Kali (India) *3:*660f
  nonpoint source nutrients from *3:*661f
River mapping, ‘‘cartographic anxiety’’ associated with *4:*683
River model development *2:*333
River pollution, in India *3:*84 *3:*447
River reaches *3:*66
  identifying and naming *3:*67
River restoration, dam removal as *3:*387
Rivers *3:*392
  See also Regulated rivers; Streams
  acidification of *3:*10
  cadmium-contaminated *5:*615

Index terms **Links**

Rivers *(Continued)*
  Canadian *2*:657
  channel patterns of *3*:31
  currents in *3*:320
  design flow in *3*:81
  dilution in *3*:84
  facts related to *3*:394
  fluctuations of *3*:134
  Indian *2*:560 *2*:563f
  lengths of *3*:397
  limnology of *3*:292
  as one-way flow systems *3*:437
  pollution of *4*:268
  salinity of *3*:679
  sediment content in *3*:408
  sodium in *4*:552
  topography of *3*:392
  trace elements dissolved in *4*:110
  velocity of *3*:393
  water and sediment loads of *3*:393t
Rivers and Harbors Act of 1899 *4*:595 *4*:651
Riverside blankets *3*:289
Riverside levee slopes, protecting *3*:290
"River snow" *3*:306
River systems, classification of *3*:392
River (stream) water(s) *2*:95 *4*:285
  composition of *4*:160 *4*:285
River water chemistry *4*:371
  atmospheric inputs into *4*:373
  contribution of lithology in *4*:374
  methodology of studying rocks *4*:371
  natural and anthropogenic inputs into *4*:373
  quantification of marine and terrestrial contributions to *4*:372
  role of human activities in *4*:374
River water infiltration, stratigraphic controls on *5*:679
River water quality, impact on groundwater *2*:396
River water quality model calibration *2*:331
  genetic algorithms for *2*:333
  procedure for *2*:334
  single optimized parameter set selection in *2*:335

Index terms	**Links**
River water quality modeling software tools	*2:*325
evaluation and selection of	*2:*328
River watershed, natural and anthropogenic export in	*4:*374
River Yamuna, India, ground water quality near	*2:*392
See also Yamuna River	
RNA stable isotope probing	*1:*645
Road construction, managing	*2:*187
Road crossing culvert fish passage facility	*3:*531
Road runoff drainage	*3:*498
Road salt	*2:*319
alternatives to	*2:*320
effects on aquatic macroinvertebrates	*2:*321
effects on zooplankton and fish	*2:*323
field and laboratory studies of	*2:*322t
impacts on aquatic environments	*2:*320
indirect effects of	*2:*321
loading	*2:*321
properties of	*2:*320
runoff	*2:*321
Robin’s boundary condition	*5:*561
Robotic pipe repair techniques	*1:*881
Rock(s)	
concentration of trace elements in	*3:*457t
consolidated water-bearing	*5:*55
dissolution rate (*R*) for	*5:*448
igneous	*5:*55
metamorphic	*5:*55
methodology of studying	*4:*371
sedimentary	*5:*55
specific yield of	*5:*577t
squeezing water from	*5:*477
strontium isotopes in	*4:*574
Rock glacier	*3:*174
taxonomy of	*3:*175t
Rock faults	*5:*137
Rock fracture	*5:*136
Rock joints	*5:*136
Rock systems, groundwater flow in	*5:*175
Rock samples, analysis of strontium isotopes in	*4:*575
Rock-water interaction	*5:*566

Index terms	**Links**		
Rocky shores	*4:*26		
ROMAG™ screen	*1:*785		
Roman aqueducts	*1:*282		
Roof drainage, threats to	*1:*60		
Roof drainage hydraulics	*1:*54		
future of	*1:*60		
Roof drainage systems, siphonic	*1:*58		
‘‘Roofnet’’ model	*1:*58		
Roof runoff drainage	*3:*498		
Roof surface design	*1:*55		
Rooftop harvesting	*2:*551		
‘‘Room for Rivers’’ policy report	*2:*526		
Root Bioaccumulation Factors (RBFs)	*1:*717		
Root growth, in grass plants	*3:*706		
Root zone	*3:*559	*3:*560	*3:*570
components of	*3:*706		
salt balance of	*3:*678		
Root Zone Water Quality Model (RZWQM)	*2:*252		
Rotary fine screens	*1:*785		
Rotary kiln furnace (RKF)	*1:*858	*1:*860	
Rotary pumps	*1:*394		
Rotary water well drilling	*5:*105		
Rotating biological contactors (RBCs)	*1:*833		
Rotavirus	*1:*279		
Roughing filters	*1:*237	*1:*486	
components of	*1:*240		
Roughness length, parameterization of	*4:*1		
Round Lake, biomanipulation trials in	*2:*57t		
Rover mission, Mars exploration	*4:*504		
Rubber membrane fine bubble diffusers	*1:*627		
Rubble-mound structures	*4:*15		
Ruditapes philippinarum	*4:*11	*4:*75	
excretion rates of	*4:*7	*4:*12	
Ruhunurata, ancient irrigation works in	*4:*775		
Rule curves	*3:*384		
Rule of Capture	*4:*628		
Rule of Constant Proportions	*4:*159		
Ruminants, effect of nitrates on	*1:*38		

Index terms	**Links**		
Runoff	*2:*546		
See also Roof runoff drainage; Surface runoff			
contaminated	*3:*606		
control of	*3:*454	*5:*256	
effluent	*2:*516		
farm	*2:*400		
groundwater contamination from	*5:*451		
observation and estimation of	*4:*376		
pollution measurement in	*3:*225		
quantity and quality of	*4:*318		
rainfall and	*4:*315		
recharge from	*5:*410		
road salt	*2:*321		
routing of	*4:*318		
stormwater	*2:*190		
surface water	*2:*516		
urban	*2:*186	*2:*190	*2:*192
use of	*3:*164		
Runoff estimation, methods for	*4:*317		
Runoff farming	*3:*164	*3:*708	
Runoff hydrology, special topics in	*4:*319		
Runoff impacts, from landfills	*5:*256		
Runoff management, constructed wetlands for	*1:*896		
Runoff mechanisms, forest	*3:*171		
Runoff pollution			
caused by human interference	*3:*225		
natural sources of	*3:*225		
Runoff water, harvesting	*3:*708		
Run-of-river projects	*3:*201		
Run-of-the-river system	*2:*554		
Run up phenomenon	*4:*16		
Rural areas, drinking water disinfection in	*1:*382		
Rural development, in the Nile Basin	*2:*593		
Rural sanitation, in developing countries	*2:*632		
Rural small water supply, drinking water quality failure in	*1:*221		
Rural water supply, in developing countries	*2:*631		
Rural water systems, design of	*1:*213		
Russia. *See also* Far Eastern Russia seas			
Lake Baikal studies by	*3:*20	*3:*21	
marine environmental monitoring in	*2:*443		

S

SAA ($SO_4^{2-} + NO_3^- + Cl^-$) 3:1
  anion concentration in 3:2
Safe Drinking Water Act (SDWA) of 1974 1:4 1:70 1:82 1:424 1:476 2:195 2:286 2:345 2:585 4:597 4:676
  Amendments of 1986 2:8 2:299
  definitions related to 4:669
  events leading to 1:288
  history of 1:479
  implementation procedures for 4:670
  regulations related to 4:668
  risk mandates from 1:424
  small drinking water systems and 1:457
  standards and goals of 4:677
  statutes related to 4:667
  underground injection control under 4:679
  variances and exemptions related to 4:678
Safe Drinking Water Hotline 1:145
Safe Drinking Water Information System (SWDIS) 4:542
  database 1:481
Safe storage, household drinking water 1:67
Safety
  of chlorine gas 2:90
  of cobalt 5:612
  of ozone 1:356
  of plastic bottles 1:5
  of point-of-use/point-of-entry systems 1:381
  pump 1:394
Safety precautions, for chemical oxidants 5:348
Safety valve 1:484
‘‘Safe yield’’ concept 2:514
Sailors 4:763
Saint-Venant (SV) equations 3:250
Saline groundwater 3:691
Saline lakes 3:266
Saline seep 5:453
Saline water 2:546
  sprinkler irrigation with 3:713

Index terms	**Links**	
Salinity. *See also* Desalination; Salt entries; Soil salinity; U.S. Salinity Laboratory Classification		
in animal farming operations	*3:*539	
biological response to	*3:*682	
in global oceans	*4:*21	
of irrigation water	*2:*157	*5:*207
lake	*3:*266	
profiles	*3:*679	
research on	*2:*352	
river	*3:*679	
sources of	*5:*223	
of tidal streams	*4:*128	*4:*130
in water bodies	*2:*181	
Salinity measurements, wetland	*3:*497	
Salinity threshold	*3:*684	
Salinization		
in Indian rivers	*3:*451	
soils affected by	*3:*571	
Salmonella	*1:*278	*2:*337
detecting using molecular tools	*2:*337	
direct detection of	*2:*338	
DNA probes in	*2:*337	
future options in detecting	*2:*339	
PCR based monitoring of	*2:*338	
Salmon species	*3:*127	
Salt(s). *See also* Road salt; Saline water; Salinity; Salinization; Sea salt effect		
effect of excess	*3:*682	
as groundwater tracers	*5:*505	
ion effects of	*3:*682	
leaching of	*3:*679	
in ponds	*3:*485	
scaling and	*1:*545	
as water quality indicators	*2:*266	
Salt balance		
on irrigated land	*3:*677	
measuring and modeling	*3:*680	
of the root zone	*3:*678	
Salt buildup, using microirrigation	*3:*618	
Salt loading	*1:*633	
Salt mines, oil-field brine and	*2:*288	
Salt stress	*3:*683	

Index terms	**Links**	
Salt tolerance	*3:*681	
characterization and assessment of	*3:*684	
of crops	*3:*684	
factors influencing	*3:*683	
of plants	*3:*682	*3:*686
recent developments in	*3:*686	
Salt-tolerant plants, breeding	*3:*686	
Salt water, solubility of hydrocarbons in	*4:*559	
Saltwater contamination, mathematical formulation for	*5:*664	
Saltwater intrusion	*5:*602	
Sample contamination	*2:*342	
Samples. *See also* Sampling		
soil vapor	*5:*548	
vinyl chloride	*5:*637	
Sampling. *See also* Water sampling		
aquatic environmental	*2:*170	
of colloids	*3:*74	
design	*2:*265	
in fish consumption advisories	*3:*119	
frequency and time of	*2:*179	
locations and points related to	*2:*178	
low-flow	*5:*406	
raw water	*1:*440t	
role in river water chemistry	*4:*371	
of sediment	*3:*509	
timing of	*2:*180	
water supply	*1:*294	
Sampling parameters, criteria for	*2:*266	
Sampling site, documenting	*2:*162	
SAMSON Database statistics	*5:*79t	
Sand abstraction	*3:*412	
advantages and disadvantages of	*3:*414	
cost effectiveness of	*3:*417	
from ephemeral rivers	*3:*414	
history and use of	*3:*413	
methods of	*3:*415	
sustainability of	*3:*415	
water quality and	*3:*415	
Sand and gravel aquifers	*5:*145	
Sand filters	*3:*752	
Sand filtration	*1:*225	

Index terms	**Links**		
San Diego, water supply demands in	*1:*218		
Sandstone	*5:*55		
Sandstone aquifers	*5:*145		
Sand wells	*3:*413		
Sandy media, in column experiments	*2:*104		
‘‘Sanitary awakening,’’	*1:*283		
Sanitary landfilling	*1:*699	*2:*163	
Sanitary quality, of urban stormwater runoff	*3:*434		
Sanitary sewerage system, design of	*3:*434		
Sanitary Sewer Overflows (SSOs)	*1:*799		
Sanitary sewers	*3:*331		
Sanitary surveys	*1:*294		
Sanitary systems	*5:*452		
Sanitation	*1:*661		
in developing countries	*2:*630		
ecological	*1:*675		
impact of	*1:*25		
integrated capacity building for	*1:*651		
issues in	*1:*518		
private sector participation models providing	*1:*50		
sustainable	*1:*653	*1:*654t	
wastewater	*1:*498		
Sanitation Connection Internet resource	*2:*669		
Sanitation surveys	*2:*361		
Sanitation systems			
criteria for	*1:*675		
implementation of	*1:*676		
Sanmenxia reservoir	*3:*46	*3:*48	
Santa Barbara Oil Spill	*3:*517		
Sap-flow techniques	*3:*736	*3:*737	*3:*739
Saprotrophic organisms, in ponds	*3:*487		
Saratoga Springs waters, medicinal properties of	*4:*797		
Sarcodina	*2:*313		
Satellite-borne instruments	*4:*311		
Satellite data, water-flow-related	*2:*587		
Satellite energy balance (EB), sources of	*3:*564		
Satellite imagery, large area surface energy balance estimation using	*3:*560		
Satellites. *See also* Aqua earth observing system (EOS)			
rainfall monitoring using	*4:*311		
water-resource	*4:*326t		

Index terms **Links**

Satellite sensors, for waterbody detection *4:*325
Satellite sensor technology *4:*350
Saturated adiabatic lapse rate (SALR) *4:*370
Saturated media, mass transport in *5:*273
Saturated porous media, modeling of DNAPL migration in *5:*668
Saturated river sediment, water from *3:*412
Saturated soil, role in flooding *3:*142
Saturated zone *5:*514
  soil samples from *3:*689
  reactive transport in *5:*518
Saturating recycle equation *2:*277
Saturation excess *4:*316
Saturation-excess overland flow *3:*452
Saturation index *4:*409
Saturation vapor pressure *3:*574 *4:*271
“Saturnine gout” *2:*436
Saxitoxins *2:*389
SBC ($Ca^{2+} + Mg^{2+} + Na^{+} + K^{+} + NH_4^{+}$) *3:*1 *3:2*
SBR biological treatment unit *1:*711 *1:*712t
SCADA systems *1:*437
  See also System Control and Data Acquisition (SCADA)
Scale *5:*175
  problem of *1:*534
Scale formation. *See also* Scaling entries
  crystallization leading to *1:*547
  in industrial cooling water *1:*545
  on pipe surface *1:*6
“Scale modifiers” *1:*548
Scaling *1:*415
  in diffused air aeration systems *1:*630
Scaling control, in industrial cooling water *1:*547
Scanning Doppler lidar *4:*297
Scanning electron microscopy (SEM) *2:*310 *4:*491
SCARPS, case study of *5:*98
Schauberger, Viktor *4:*794
Schlumberger array *5:*445
Schmutzdecke *1:*234 *1:*235 *1:*244 *1:*246
School conservation programs *1:*148
Schwenk, Theodor *4:*768 *4:*793
Schwenk, Wolfram *4:*796

Index terms	**Links**		
Science magazine	*3:*515		
Science Olympiad (SO)	*4:*714		
Scientific fisheries management	*3:*124		
Scientific water management models	*2:*620		
Scientists, atmospheric	*4:*328		
Scour hole	*3:*67		
Scour pools	*3:*67	*3:*68	
Screened hydraulic probes	*5:*405		
Screen filters, for microirrigation	*3:*748		
Screening	*1:*814		
at CSO facilities	*1:*784		
domestic sewage	*1:*832		
in dye tracer selection	*3:*98		
wastewater	*1:*809		
Scrubbing, for odor abatement	*1:*762		
SCS lag equation	*3:*471		
Sculpin tissue. *See also* Tissue			
metals in	*3:*361t		
organochlorine compounds in	*3*:349	*3*:350	*3*:360t
trace elements detected in	*3*:362		
Seabed filtration intake systems	*5:*89		
Sea floor			
gas hydrate observations on	*4:*60		
high-resolution surveys of	*4:*77		
Sea floor mapping, of New York City Historic Area Remediation Site	*4:*77		
Sea fog	*4:*232		
Seagrass restoration	*4:*56		
Sea ice	*4:*21	*4:*194	
Sea ice cycle	*4:*70		
Sea level, climate and	*4:*117		
Sea-level changes, potential	*4:*118		
Sea-level rise			
as a global problem	*2:*483		
societal impacts of	*2:*482		
Seals, mechanical	*1:*882		
Sea salt effect, acidification and	*3:*5		
See also Salt entries			
Seasonal cropping	*4:*698		
Seasonal crop water requirements, estimating	*3:*558		
Seasonally draining epikarsts	*5:*237		

Index terms	**Links**			
Seasonal tidal stream variability	*4:*130			
Seasonal-vertical solar radiation, distribution equation for	*3:*376			
Sea Surface Temperature (SST)	*4:*335			
Seattle Water Department, water use reduction plan	*2:*581	*2:*582t		
Sea turtles	*4:*57			
Sea volume, effects of decrease in	*3:*15			
Seawater	*2:*95	*4:*285		
chemistry of	*1:*414			
composition of	*4:*160	*4:*285		
contamination criteria for	*2:*450t			
ions in	*4:*159			
microbial detection of oil pollution in	*2:*441			
microbial detection of phenols in	*2:*442			
molecular structure of	*4:*513			
quality control of	*2:*449			
similarity to blood plasma	*4:*712			
sodium in	*4:*552			
Seawater brines	*5:*52			
Seawater CH_4, sources of	*4:*87			
Seawater desalination plant	*5:*91			
Seawater N_2O, sources of	*4:*87			
Seawater RO plant	*1:*308			
Seawater temperature estimates, in paleoceanography	*4:*92			
Seawater temperature information	*4:*93			
SEBAL (surface energy balance algorithm for land)	*3:*561			
Secondary drinking water regulations	*4:*678			
Secondary ion mass spectrometry	*1:*646			
Secondary MCLs (SMCLs)	*1:*477			
Secondary physicochemical monitoring	*2:*267			
Secondary settling tanks	*1:*453	*1:*456		
design considerations for	*1:*456			
Secondary wastewater systems, powdered activated carbon in	*2:*88			
Secondary wastewater treatment	*1:*809	*1:*815	*1:*827	*2:*546
Secondary water quality standard	*2:*373			
Second law of thermodynamics	*5:*170			
Security				
prioritizing	*1:*870			
SCADA system	*1:*451			
of wastewater utilities	*1:*870			
water distribution system	*1:*434			

Index terms	Links			
Security improvements, investing in	*1:*871			
Security requirements, for potable water	*2:*347			
Security vulnerability checklist	*1:*434			
Sediment(s). *See also* Bed sediment metal ion adsorption; Sand abstraction; Sedimentation; Sediment transport; Streambed sediment				
in an agricultural landscape	*3:*606			
aquatic	*2:*384			
characteristics of	*3:*298t			
contaminated	*2:*426			
data and information related to	*3:*509			
defined	*2:*546			
from forest water	*2:*199			
heterogeneity of	*2:*411			
importance of	*3:*508			
organochlorine compounds in	*3:*353	*3:*354t		
physical properties of	*3:*417			
as a pollutant	*3:*607			
properties of	*2:*213t			
Puget Sound Basin	*3:*349			
routing	*3:*412			
sampling	*3:*509			
sealing and removal of	*2:*5			
sources of	*3:*417			
trace elements in	*3:*455			
in the Yellow River Basin	*3:*46			
Sedimentary deposits, of continental rifts	*3:*20			
Sedimentary rock(s)	*2:*546	*5:*55		
Sediment assays, sediment or fauna incubation experiment for	*2:*418			
See also Sediment quality assessment/characterization; Sediment toxicity assessment/tests				
Sedimentation	*1:*227	*1:*370	*1:*487	*1:*811
	*2:*374	*3:*401	*5:*619	
See also Reservoir sedimentation				
analysis and modeling of	*3:*404			
classes of	*3:*404			
dilution equivalent to	*4:*104			
domestic sewage	*1:*832			
flotation and	*3:*404			
general conditions for	*3:*45			
impact of	*4:*107			
managing	*2:*186			

Sedimentation *(Continued)*
  particulate removal by *1:*243
  primary *1:*815
  process design and implementation of *3:*406
  as a threat to fisheries resources *3:*134
Sedimentation processes, selecting *3:*407t
Sedimentation tanks *1:*259f *1:*260 *2:*546
  design criteria for *1:*456t
Sediment-attached pollution *3:*333
Sediment-based endocrine disruption (ED) *3:*54
Sediment bypass *3:*412
Sediment cleanup, dredging for *2:*126
Sediment column porewater nutrient concentrations *4:*73
  seasonal coupling with intertidal macrofauna *4:*73
Sediment cores *4:*92
Sediment data, trace element pollution and *4:*112
Sediment detention system *1:*538
Sediment discharge ratio *3:*419
Sediment-dwelling organisms
  body concentrations in *2:*215
  exposure tests with *2:*212 *2:*420
  heavy metal uptake rates among *2:*211
Sediment inflow, into reservoirs *3:*411
Sediment load measurements *3:*397
Sediment loads *3:*403
  river *3:*393t
Sediment management alternatives, dam decommissioning and *3:*388t
Sediment Or Fauna Incubation Experiment (SOFIE$^{®}$) *2:*212 *2:*418
  exposure method for *2:*420
  results of *2:*421
Sediment oxidation, algal control via *2:*4
Sediment pollutants, ecotoxicological effects of *2:*428f
Sediment production, by erosion *3:*31
Sediment quality
  criteria for *2:*126
  future research on *2:*430
Sediment quality assessment/characterization
  bioaccumulation in *2:*426
  biomarkers in *2:*429
  chronic, bioaccumulation, and sublethal bioassays for *2:*351

Index terms **Links**

Sediment quality assessment/characterization *(Continued)*
  future developments in *2:*352
  historical background of *2:*408
  weight of evidence approach to *2:*350
Sediment quality criteria/guidelines (SQGs) *2:*428 *2:*430 *3:*54
  co-occurrence-based *4:*601
Sediment Quality Triad (SQT) *2:*350 *2:*351
Sediment research/studies
  on agricultural drainage ditches *3:*90t
  Lake Baikal *3:*21
Sediments
  antimony concentrations in *4:*592
  groundwater flow in *5:*175
  specific yield of *5:*577t
  xenobiotic accumulation and *4:*120
Sediment samplers *4:*71
Sediment screening *4:*601
Sediment simulation *2:*256
Sediment testing, incentive for *2:*408
Sediment toxicity assessment/testing *4:*122
  acute *3:*51
  amphipod *2:*408
  application of *2:*460
  behavioral endpoints in *3:*54
  bioassays checked for use in *2:*462t
  biological guidelines for *2:*463t
  biomarkers used for *2:*430t
  chironomids in *3:*50
  chronic *3:*52
  design for *2:*459
  development and application of *2:*383
  development for regulatory purposes *2:*458
  future *3:*54
  marine and estuarine microalgal *4:*120
  systems used for *2:*384
Sediment transport *3:*49 *3:*401 *3:*417 *4:*359
  downstream fining in *3:*420
  functions of *3:*419
  incipient motion and *3:*418
  modeling of *3:*421
  modes of *3:*397

Sediment transport *(Continued)*
  movement of fluvial sediments *3:*418
  by size fraction *3:*403
  stream aggradation and degradation related to *3:*420
  stream capacity for *3:*419
Sediment transport capacity *3:*659
  changes in *3:*393
Sediment yield *3:*419
  gully *3:*184
Seed microorganisms *2:*39
Seepage *2:*546
Seepage control, levee *3:*289
Seepage flow. *See* Base flow
Seepage meter method *5:*166
Seepage meters *5:*74
Seepage velocity *5:*525
Seeps, responses to earthquakes *5:*112
Seismic geophysical methods *5:*147
Seismic reflection methods *5:*147
Seismic refraction method *5:*147
Seismic sea wave *4:*160
Seismic surveys, riverine *5:*679
Select Committee on Transportation Routes *2:*523
Selective dissemination of information (SDI) *2:*672
Selective metering *1:*494
Selenate *5:*398
Selenide *1:*852
Selenium (Se)
  Aquatic Life Criterion for *2:*355
  in ash pond water *1:*851
  bioaccumulation of *2:*357t
  biogeochemistry of *2:*355
  biogeochemical cycling of *2:*356f
  hyperaccumulators and nonaccumulators of *5:*398
  maximum contaminant level of *2:*235t
  volatilization of *2:*359
Selenium-contaminated waters, remediation and bioremediation of *2:*355
Selenium contamination, sources of *5:*397
Selenium-laden soils, phytoremediation of *5:*397
Selenium transformation mechanisms *5:*398

Index terms **Links**

Self-governing institutions, promoting *2:*558
Self-supplied water *2:*546
  users of *1:*551
Semi-arid climatic zone *4:*257
Semi-Arid Hydrology and Riparian Areas (SAHRA) program *5:*72
"Semiconfined aquifer" *5:*11
Semiconfined aquifer flows *5:*494
Semi-intensive farm systems *3:*579
Semiperched groundwater *5:*354
Semipermeable membrane devices (SPMDs) *2:*170 *5:*672
  deployment, processing, and analysis of *5:*675
  sampling using *5:*674
  similarity to biomembranes *5:*673
Semivolatile organic compounds (SVOCs) *1:*838 *5:*338
  in streambed sediment *3:*357 *3:*359t
Senec gravel pit, biomanipulation trials in *2:*55t
Sensible heat flux *3:*193
Sensible heat flux density *3:*563
Sensitive ecosystems, water recycling and *2:*612
Sensitivity analysis *3:*319 *3:*329
Sensor-based pH measurement *2:*295
Sensor drift *2:*283
Sentinel species *3:*116 *4:*113
Separate Municipal Storm Sewer Systems (MS4) *4:*659
Separation, of hydrocarbons *1:*576
Separation processes, flotation *1:*684
Septic systems *5:*604
  See also Septic treatment systems
  contaminants from *3:*228
Septic tank *2:*546
  systems *1:*34 *1:*61
Septic treatment systems *2:*663
Sequencing batch reactor *1:*815
Sequential sediment samplers *4:*71
Sequential stream flow routing (SSR) *3:*202
Sequent-peak analysis *3:*261
Sequestering agents *1:*548t
Sequestration
  of iron and manganese *1:*314
  for scale control *1:*548

Index terms	Links	
Serial analysis of gene expression (SAGE)	*2:*60	
Serological tools	*1:*189	
Serpulid worms	*1:*541	
Service contract model	*1:*50	
Service pipes, freezing and clamping of	*1:*890	
Setback distances, for geologic settings	*1:*72	
Setting velocity	*3:*405	
Settling basins	*1:*682	
Settling pond	*2:*546	
Settling tanks	*1:*452	
design considerations for	*1:*455	
enhancements to	*1:*454	
types of	*1:*452	
Severe cyclonic storm	*4:*195	
Sewage	*1:*828	
See also Domestic sewage; Sewage treatment entries; Sewerage odors		
composition and contaminants	*1:*829	*2:*182
defined	*1:*517	
dioxin-contaminated	*2:*110	
reasons for treating	*1:*517	
treatment plant influent metals in	*1:*831	
Sewage control, during repairs	*1:*888	
Sewage liquid, electronic nose analysis of	*2:*282	
Sewage sludge		
disposal at sea	*4:*147	
occurrence of detergents in	*1:*671	
precautionary principle regarding	*2:*600	
Sewage treatment	*1:*829	
anaerobic	*1:*517	
Sewage treatment facilities, fate of pharmaceuticals in	*1:*373	
Sewage treatment plant	*2:*546	
Sewage treatment systems		
off-site	*1:*830	
on-site	*1:*830	
Sewerage, simplified	*2:*633	
Sewerage odors		
causes of	*1:*911	
controlling	*1:*910	
preventing	*1:*912	
removing	*1:*913	

Index terms	Links	
Sewer overflows	*2:*263	
Sewer overflow treatment. *See* Combined sewer overflow (CSO) treatment		
Sewer overflow volumes, reduction of	*3:*338	
Sewer pipeline. *See* Underground pipeline		
Sewers		
field measurements in	*3:*338	
modeling fecal coliform in	*3:*334	
as reactors	*3:*331	
Sewer separation	*1:*783	
Sewer systems	*1:*884	
holistic analysis of	*3:*336	
urban	*2:*96	
Sewer water quality modeling	*3:*331	
challenges in	*3:*336	
deterministic	*3:*332	
procedure for	*3:*334	
Shaded relief sea floor images	*4:*78	
Shadow clock	*4:*703	
Shallow aeration unit	*1:*51	
Shallow groundwater contamination	*5:*426	
Shallow lake ecosystems, alternative stable states theory regarding	*3:*272	
Shallow lakes	*3:*265	
Shallow sewerage	*2:*633	
Shallow tray aeration (STA)	*1:*460	
Shallow water equations (SWEs)	*3:*710	*5:*659
Shallow water flow (SWF)	*3:*710	
Shallow water habitat, eutrophication in	*3:*108	
Shallow water theories	*4:*135	
Shallow water waves	*4:*135	
Shannon entropy	*4:*218	
Shear stress, microbial growth and	*3:*312	
See also Shields shear stress approach		
Sheet piling breakwater	*4:*19	
Sheindorf *et al.* adsorption model	*1:*109	
Shelf waters	*4:*26	
Shellfish, vulnerability to pollutant discharge	*3:*284	
Shellfish growing, water classification for	*2:*360	
Shelterbelts	*3:*551	*3:*568
Shields shear stress approach	*3:*418	
Shiga-like toxins	*2:*136	

Index terms	**Links**	
Shiga toxin-producing *Escherichia coli* (STEC)	*2:*136	*2:*137
Shiga toxins	*2:*137	
Shigella	*1:*278	
Shigellosis colitis	*1:*181	
Ship engineers	*4:*763	
Ship pilots	*4:*763	
Shoaling, wave refraction and	*4:*136	
Shoaling depth, estimating	*4:*3	
Shock fitting	*3:*241	
Shock formation	*3:*239	*3:*248
Shocks		
on curved surfaces	*3:*241	
first-order	*3:*240	
in a kinematic cascade	*3:*241	
in planar flows	*3:*241	
Shore Protection Act of 1989	*4:*40	
Short-term persistence	*3:*210	*3:*212
Short-wave infrared (SWIR) domain	*3:*719	*3:*720
Short-wave radiation, into open waterbodies	*3:*191	
Shotcrete	*1:*880	
Showerheads, low-flow	*2:*665	
Shuffled complex evolution algorithm (SCE-UA)	*4:*298	
Sieve analysis	*1:*96	
Signal delay measurements, slant-path	*4:*247	
Silent Spring, The (Carson)	*4:*653	
Silesia		
fishpond culture in	*3:*135	
pond fish farming in	*4:*719	*4:*721
Silica		
biochemical aspects of	*4:*550	
classification of	*4:*549	
deposition of	*4:*550	
geochemistry of	*4:*549	
in natural waters	*4:*548	
Silica compounds, dissolved	*4:*550	
Silica cycle, geochemical	*4:*550	
Silicate corrosion inhibitors	*1:*154	
Silicate rocks, dissolution of	*4:*549	

Index terms	**Links**	
Silicates		
distribution of	*4:*74	
inorganic	*1:*11	
Silicic acid	*4:*548	*4:*549
Silicon dioxide (SiO_2)	*4:*548	
Siltation rate, reservoir	*3:*410	
Silt density index (SDI)	*1:*418	
Silt traps	*3:*500	
Silvester, Hans	*4:*768	
Simple clathrate hydrates	*4:*471	
Simple random groundwater monitoring network	*5:*314	
Simple Ratio Pigment Index (SRPI)	*3:*722	
Simple water quality models	*2:*248	
Simulation, long-term	*3:*335	
See also Stochastic simulation		
Simulation models	*3:*386	
physically and chemically based	*3:*5	
Simulation of single sludge processes (SSSP) model	*1:*731	
Simulated oil spill case study	*2:*172	
Simulator for Water Resources in Rural Basins (SWRRB) model	*2:*250	
Singh and Stall base flow separation method	*3:*26	
Single-celled organisms, as test subjects	*2:*413	
Single channel seismic reflection surveys	*5:*680	
Single-component phase diagram	*4:*365	
Single dimensional pump test analysis	*5:*185	
Single pass adsorption resins technology	*5:*633	
Single point resistance log	*5:*152	
Single-stage samplers	*3:*399	
Singlet oxygen	*4:*533	
Sinkhole	*2:*546	
Sinking bowl water clock	*4:*705	
Sinuosity	*3:*31	
Sinusoidal oscillations	*5:*498	
SIPHONET numerical model	*1:*59	
Siphonic roof drainage	*1:*55	*1:*56
systems	*1:*58	
Site assessment, Orlando Naval Training Center	*5:*82	
Site selection, for wells	*5:*88	
Site-specific sampling	*2:*162	

Index terms **Links**

Siting issues, for municipal solid waste landfills *2*:167
Size fractionation procedures *2*:102
Skeletal fluorosis *4*:435
Skimmed groundwater *3*:691
Skimming wells *3*:594
Skin cancer, arsenic-related *2*:17
Sky emissivities *3*:563
Slant-path signal delay measurements *4*:247
Slime layers. *See* Biofilms
Slimes, extracellular *5*:23
Slip coefficient *5*:500
Sliplining *1*:801 *1*:877
Slope-area law *3*:94
Sloped roofs *1*:55
Slope protection, riverside *3*:290
Sloping aquifer, groundwater flow equations for *5*:13
Sloping lands, drainage of *5*:97
Slots *3*:70
Slow migratory (spreading) cell formation *2*:240
Slow moving fronts, fog and *4*:231
Slow neutron count ratio (CR) *3*:693
Slow neutron detector *3*:692
Slow sand filter pilot, testing *1*:432
Slow sand filters *1*:231 *1*:233 *1*:246 *1*:432f
  advantages and limitations of *1*:431
  design summary of *1*:433t
  treatment performance of *1*:431t
Slow sand filtration *1*:228 *1*:235 *1*:239 *1*:249 *1*:431 *1*:458 *1*:488
  monitoring and operation requirements for *1*:433
  process of *1*:432
Sludge(s) *1*:413
  See also Sludge treatment and disposal; Wastewater sludge
  chemical composition of *1*:862
  discharge *1*:324
  floc from *1*:844
  odors from *1*:850
  pharmaceutical products in *1*:375
  physical properties of *1*:862
  quantity and moisture of *1*:862

Index terms **Links**

Sludge(s) *(Continued)*
  reuse and disposal of *1:*684
  septicity of *1:*455
  as a source of contaminants *5:*433
  types of *1:*861
Sludge age *1:*730
Sludge digestion, persistent organic pollutants and *1:*769
Sludge processing *1:*864
  system for *1:*682f
Sludge separation problems *1:*844
Sludge treatment and disposal *1:*687 *1:*853
  codisposal with municipal solid wastes *1:*861
  land application *1:*858
  options for *1:*854
  sewage sludge categories *1:*854
  thermal processing *1:*858
  usable materials production *1:*861
Sludge treatment stream, fate of POPs in *1:*768
Slug test *5:*183 *5:*575
Small drinking water systems
  corrosion control in *1:*459
  disinfection methods for *1:*457
  filtration for *1:*458
  ion exchange and demineralization in *1:*459
  lime softening in *1:*460
  organic removal in *1:*351 *1:*460
  SCADA and *1:*450
  treatment technologies for *1:*457
Smallpox *1:*88
Small-scale industry, sand abstraction for *3:*414
Small-scale wastewater treatment *1:*840
  costs of *1:*842
  methods of *1:*840
  requirements for *1:*840
Small water reservoirs, role in environment *1:*403
Small water retention *1:*405 *1:*407
Smoke testing, in pipeline assessment *1:*885
SMP reagents *2:*376
SNOTEL snowpack, Flattop Mountain *4:*334

Index terms	**Links**			
Snow	*4:*293	*4:*336		
density of	*4:*333			
remote sensing of	*4:*324			
Snow control, alternatives to NaCl for	*2:*320t			
Snow course	*4:*338			
Snow cover studies, remote sensing and GIS application in	*2:*533			
Snowfall				
in Canada	*2:*660			
remote sensing of	*4:*322			
Snowmelt	*4:*337			
acidic deposition from	*3:*225			
computing	*4:*337			
episodes of	*3:*5			
role in flooding	*3:*145			
runoff floods	*3:*511			
Snowpack	*4:*336			
Snow surveys	*4:*337			
instruments used for	*4:*338			
SO_3 groups, QSAR modeling and	*3:*101			
SO_4^{2-}	*3:*1	*3:*2	*3:*3	*3:*6
See also Sulfate				
Social capital	*1:*655			
Social impacts, of water transfers	*4:*687			
Social justice, in the Arab world	*4:*634			
Social responsibility	*1:*45			
Social standards, sharing of	*4:*738			
Social value of water	*2:*653			
Society				
acid rain and	*4:*377			
benefits of weather forecasting to	*4:*351			
climate and	*4:*183			
impact of urban flooding on	*3:*161			
importance of water to	*4:*289			
"Sociological drought"	*2:*576			
Socio-spatial consciousness	*4:*680			
Soda ash, in acid mine treatment	*5:*4			
Sodium (Na)				
exchangeable	*3:*682			
in groundwater	*2:*394			
irrigation water quality and	*5:*207			

Index terms **Links**

- Sodium (Na) *(Continued)*
  - in natural waters *4:*551
  - treatment of *4:*552
- Sodium adsorption ratio (SAR) *5:*204
- for irrigation water *2:*157 *5:*207 *5:*210
  - values of *2:*160t
- Sodium carbonate, in irrigation water *2:*159 *5:*208
- Sodium chloride, properties of *2:*320t
  - *See also* Road salt; Salt entries
- Sodium concentration *5:*196
- Sodium hexametaphosphate (SHMP) *1:*419 *4:*415
- Sodium hypochlorite *1:*457
- Sodium ion *4:*551
- Sodium nitrate *5:*322
- Sodium perchlorate *5:*632
- Sodium pertechnetate *4:*579
- SOFIE. *See* Sediment Or Fauna Incubation Experiment (SOFIE®)
- ‘‘Soft’’ carbon *4:*386
- Softening. *See also* Lime softening
  - alternatives for *1:*323
  - benefits and concerns about *1:*324
  - lime-soda ash *1:*320
- “Softening membranes” *1:*335
- Software. *See also* Spreadsheet entries
  - analysis *3:*387
  - catchment water quality modeling *2:*325
  - culvert hydraulic computation *3:*78
  - drainage/sewer system modeling *3:*161
  - integrated water quality modeling *2:*328
  - reservoir-analysis *3:*203
  - river water quality modeling *2:*325
  - steady-state *2:*326
  - unsteady-state *2:*327
  - wastewater modeling *1:*732
- Soft water *4:*553
  - acidified *3:*8
  - availability of *4:*554
  - health and *4:*554

Soil(s). *See also* Permanent frost (permafrost); Soil conservation; Soil erosion
  acidification of *3:*8
  acidity and alkalinity in *3:*707
  as a biodegradable agent *3:*707
  cadmium sorption and desorption in *5:*616
  chemistry of *5:*367
  cleanup of *3:*652
  detoxification of *5:*380
  effective utilization of *5:*466
  electrical conductivity of *3:*675
  forms of nitrogen in *3:*695
  infiltration capacity of *3:*171 *5:*272 *5:*213
  lithologic description and interpretation of *3:*689
  metal burden in *5:*433
  mineral components of *3:*706
  nitrogen losses from *3:*695
  nitrogen management in *3:*694
  pesticide-contaminated *3:*651
  phosphorus in *3:*110
  premining analysis of *5:*1
  reactive solute transport in *5:*524
  redox potential of *3:*599
  relationship to water *3:*641
  as a semipermeable barrier *5:*532
  as a source of contaminants *5:*433
  spatial and temporal variability in *4:*488
  water issues related to *3:*706
  water retention properties of *5:*260t *5:*472
  zinc and cadmium bioavailability in *5:*370
Soil analysis and improvements *3:*556
Soil and Water Assessment Tool (SWAT) *2:*250
Soil bioengineering, flood-related *3:*148
Soil bioventing *5:*120
  electrokinetic enhanced *5:*121
Soil components, pesticide sorption to *3:*648
Soil conditioners *3:*552 *3:*569
Soil conservation *3:*549
  approaches to *3:*549
  implementation and adoption of measures for *3:*552
  need for *3:*549
  soil management methods of *3:*551 *3:*569

Index terms **Links**

Soil conservation ecosystems, ancient Sri Lankan *4:*772
Soil Conservation Service (SCS) *2:*513
  soil group classification *3:*30
Soil contaminants, sources of *5:*433
Soil contamination. *See also* Contaminated soil
  causes of *5:*432
  by heavy metals *5:*275
  sources of *3:*651
  treatment with Fenton's Reaction *4:*445
Soil disturbance, reducing *3:*552
Soil erosion *3:*565
  controlling *3:*567
  impacts of *3:*566
  nitrogen losses from *3:*696
  rates of *3:*566 *3:*567t
  as a source of sediments *3:*417
Soil fauna *3:*706
Soil fertility, components of *3:*550t
Soil flushing, *in situ* *5:*435
Soil gas *5:*548
Soil groups *3:*30
  CEC values for *5:*261t
Soil lysimeter *5:*463
Soil management *4:*433 *4:*434
Soil management practices, water quality models for developing *2:*248
Soil metals, chemical fate and mobility of *5:*433
Soil microbiology *5:*464
Soil moisture *5:*600
  content *3:*600 *3:*692 *5:*320
  controls on *4:*484
  measurement of *4:*487
  remote sensing of *4:*323
  soil capacity to hold *3:*707
Soil moisture assessment, remote sensing and GIS application in *2:*533
Soil-moisture-data-based methods *5:*165
Soil moisture measurement, using neutron probe *3:*692
Soil moisture processes, infiltration and *4:*484
Soil monitoring techniques, direct *5:*542
Soil nutrients, leaching of *4:*380

Index terms **Links**

Soil phosphorus. *See also* Phosphorus (P)
  availability of *3:*702
  environmental management of *3:*704
  impact on surface water quality *3:*701
  transport of *3:*703
Soil pipes *5:*401
  hydrology and erosion of *5:*403
  initiation and identification of *5:*401
  morphology of *5:*402
Soil-plant systems, variation in *3:*602
Soil salinity *3:*673
  field methods of measuring *3:*674
  laboratory methods of measuring *3:*673
  plant tolerance to *3:*676
  remote sensing and GIS application in *2:*534
Soil sampling
  groundwater assessment using *3:*688
  methods of *3:*689
Soil solution extracts, electrical conductivity of *3:*673
Soil sorption parameters *3:*99
Soil spatial variability
  apparent soil electrical conductivity and *5:*466
  characterizing *5:*465
  EC$_a$-directed soil sampling survey to characterize *5:*467
Soil surface, evaporation at *5:*211
Soil tests, to predict available phosphorus *3:*703
Soil texture, influence on water penetration *3:*707
Soil to plant absorption *3:*644
Soil treatment/groundwater treatment boundary *5:*39
Soil vapor concentrations *5:*552
Soil vapor data, groundwater investigations and *5:*548
Soil vapor extraction (SVE) *5:*40 *5:*426 *5:*427 *5:*440 *5:*628
  activated carbon in *1:*105
  versus bioventing *5:*43
  of vinyl chloride *5:*639
Soil vapor/groundwater contamination, sources of *5:*550
Soil vapor methods *5:*548
Soil vapor samples, collection of *5:*548
Soil-vegetation-atmosphere-transfer (SVAT) models *4:*322

Soil venting *5:*426 *5:*430
Soil washing *3:*653 *5:*435
  for lead *5:*648
Soil water *5:*461 *5:*531
  lysimeter sampling of *2:*340
  measurement of *5:*533
  solute concentration of *2:*340
  storage of *5:*124
Soil water balance method *5:*165
Soil-water characteristic curve *5:*535
Soil water movement *5:*210 *5:*471
  principles *4:*486
Soil water potential *5:*532
  components of *5:*532
  factors affecting *5:*532
Soil water processes, infiltration and *5:*210
Soil water redistribution *5:*211
Soil water sampling *5:*463
Solar actinic flux *4:*531
Solar and Meteorological Observation Network (SAMSON) *5:*78 *5:*79t
Solar electric power plants *4:*479
Solar energy, evaporation and *4:*223
Solar radiation
  hydrologic cycle and *4:*277
  in shallow stagnant waters *3:*375
Solar system, liquid water in *4:*190 *4:*191
Solar thermal plants *1:*563
Solar water heaters, domestic *1:*63
Solar water heating collectors *1:*65
Solidification process *5:*618
Solidification/stabilization (S/S) *1:*837f
  application and performance of *1:*839
  cost of *1:*839
  defined *1:*835
  of hazardous solid wastes *1:*835
  immobilization technologies *5:*434
  key features of *1:*838
  likelihood of cross-media contamination from *1:*838
  pollutant cross-media transfer potential after *1:*838
Solid/liquid solvent extraction *2:*307

Index terms	**Links**			
Solid phase extraction (SPE) technique	*4:*508			
Solids				
in domestic sewage	*1:*830			
role in fish growth and production	*3:*130			
Solids loading rate (SLR)	*1:*456			
Solid state pH sensors	*2:*298			
Solid trap materials	*2:*307			
Solid wastes. *See also* Hazardous solid wastes				
from electric generating plants	*1:*555			
removal of	*1:*682			
Solitary wave theory	*4:*135			
Soligenic areas	*1:*404			
Solubilization, by biosurfactants	*5:*377			
Solubility				
of chemicals in water	*4:*555			
of hydrocarbons and sulfur compounds in water	*4:*561			
of hydrocarbons in salt water	*4:*559			
Solubility graphs	*4:*556	*4:558f*	*4:*559	*4:*560
	*4:*561f	*4:*562	*4:*564f	
Soluble fatty acids, anaerobic tank	*1:*789			
Solute mobility, issues involved in	*5:*307	*5:*308t		
Solutes	*2:*547	*4:*458		
reactive	*5:*274			
Solute transport				
abiotic	*5:*527			
biotic processes affecting	*5:*528			
equations governing	*5:*309			
physical mechanisms of	*5:*525			
scales of observation for	*5:*306			
in soil and groundwater	*5:*524			
Solute transport models	*5:*303			
groundwater	*5:*305			
Solutions				
defined	*2:*547			
freezing of	*4:*586			
redox potential of	*2:*466			
Solvent extraction	*1:*812			
Solvent remediation, water-jetting drilling technologies for	*5:*234			

Index terms	**Links**	
Solvents	*2:*317*t*	*2:*547
extraction of	*2:*306	
stabilizers for	*2:*317	
SOLVER function	*2:*621	
Somalia, megawatersheds exploration in	*5:*267	
Sonar, for pipeline assessment	*1:*884	
Sonic logs	*5:*153	
Sorbents, chemically modified	*2:*72	
Sorbitol-containing media	*2:*139	
Sorption	*2:*362	*5:*527
in amorphous and condensed organic matter domains	*4:*386	
arsenic	*1:*82	
defluoridation and	*4:*437	
of dissolved chemicals	*5:*261	
fixed-bed	*2:*363	
in geosorbent domains	*4:*385	
of hydrophobic organic compounds	*4:*387	
kinetics	*2:*72	
in mineral domain	*4:*386	
of organic vapors	*5:*545	
rates of	*4:*387	
by zeolites	*1:*874	
Sorption coefficient, of cadmium onto sea sand	*2:*105	
Sorption-desorption kinetics	*4:*388	
Sorption equilibrium	*4:*384	
factors affecting	*4:*387	
Sorption kinetics	*4:*564	
Sorptive filtration	*2:*362	
combined process in	*2:*364	
Sorptive flotation	*1:*588	
SOTE, of diffused air aeration systems	*1:*629	
Sound		
attenuation of	*4:*570	
oceanic	*4:*570	
in water	*4:*569	
Sound detection, underwater	*4:*570	
Sounding profiles, present during coastal fog	*4:*237	
Sound waves, frequency and amplitude of	*4:*569	
Source Loading and Management Model (SLAMM)	*2:*253	
Source-pathway-receptor analysis	*5:*38	

Index terms **Links**

Source water assessment *1*:444 *1*:524
  area delineation *1*:445
  contamination inventory *1*:446
  determining contamination susceptibility *1*:446
  releasing assessment results *1*:447
  using *1*:447
Source water assessment programs (SWAPs) *1*:445 *1*:525
Source-water protection *2*:120 *2*:311
Source water quality, degraded *2*:400
Source water quality management *2*:399
  future of *2*:401
South Africa
  drinking water disinfection in *1*:382
  drinking water quality failure in *1*:221
  inadequate wastewater treatment in *1*:661
  water quality guidelines in *1*:385 *1*:386
South America
  cienega in *3*:57
  transboundary waters in *4*:642
South Asia, hydro-borders in *4*:680
South Atlantic Bight Recruitment Experiment (SABRE) *4*:55 *4*:57
South Branch Raritan River, flow-duration curve for *3*:103 *3*:104f
South Coast Track *4*:788
South Dakota, in-stream flow protection in *4*:662
Southeast, water demands in *2*:501
Southern Arizona, cienegas in *3*:58
Southern Hemisphere circumpolar vortex, strength of *4*:175
Soviet style water management *3*:19
Soxhlet extractor *2*:307
Space Station, water on *4*:572
Spain
  acidic rivers in *3*:10
  dredged material characterization in *2*:351
  permanent spot market in *2*:606
  sediment toxicity tests in *2*:460
Spanish coast, storm record on *4*:339
Sparging, air *5*:428
  See also Biosparging
Sparging mixers *1*:625
Sparkling water *1*:4

Index terms **Links**

Spate irrigation *3:*709
Spatial grid spacing *3:*257
Spatial heterogeneities *5:*308
Spatial resolution *4:*319
Spatial variability *5:*472
Spawning season, chlorine sensitivity and *2:*402
Speciation, effects of treatment processes on *2:*75
Species, taxonomical identification of *2:*459
Species and habitat preservation/protection *3:*124
  fishery-related *3:*128
Specific conductance *2:*547
Specific capacity *5:*101 *5:*460
  tests *5:*102 *5:*183
Specific gravity *5:*473
  importance of *5:*474
  measuring *5:*474
Specific heat *4:*367 *4:*369
Specific humidity (q) *4:*274
Specific storage *5:*129 *5:*481 *5:*482
Specific throughput volume (V_{sp}) *1:*103
Specific vulnerability *5:*595
Specific yield (Sy) *5:*481 *5:*483 *5:*576
  of groundwater *5:*129
  storage equation *5:*576
Spectral actinic flux *4:*531
Spectral analysis, for periodicity identification *4:*339
Spectral expansion method *5:*311
Spectral gamma log *5:*152
Spectral libraries *2:*308
Spectral radiance *3:*563
Spectral resolution *4:*320
Spectroscopic methods *2:*207
Spectroscopy, atomic absorption *2:*308
Specular reflectance *4:*321
Speleothems *5:*245 *5:*246
Spent nuclear fuel storage, water treatment in *1:*595
Spherical activated carbon (SAC) *1:*93
Spilling breakers *4:*137
Spillways *2:*568

Index terms **Links**

Spirostomum ambiguum, biotest using *2:*413
Spirostomum teres, as a test organism *2:*414
SPIROTOX test *2:*413
SPMD-TOX paradigm, monitoring lipophilic contaminants using *2:*170
Spontaneous potential log *5:*151
Spontaneous potential method *5:*146
Sporadic disease *1:*187
Spore transmission, *Microsporidium* *1:*367
Sporozoans *2:*314
Spot markets *2:*606
Spot repairs *1:*882
Spray aeration *1:*460
Spray aeration unit *1:*51
Spray irrigation *2:*547
Spray microirrigation *3:*621
Spray-on linings *1:*879
Spreader ditch flow system *3:*165
Spreadsheet models *2:*253
Spreadsheets, decision support system *2:*620
Spreadsheet wastewater model *1:*731
Spring discharge, in regional flow systems *5:*420
Spring hydrograph analysis *5:*239
Springs
  geothermal *5:*156
  hydrogeologic characteristics of *5:*476
  responses to earthquakes *5:*112
‘‘Spring’’ tides *4:*24
Sprinkler infiltrometer *4:*487
Sprinkler irrigation *3:*582 *3:*712
  methods *2:*495
Sprinklers *2:*666
Sri Lanka
  ancient water and soil conservation ecosystems of *4:*772
  modern development in *4:*776
  water disinfection in *1:*471
SSS autocorrelation function *3:*218
S state cycle, substrate water binding during *4:*536
S-state dependence, on isotope exchange rate *4:*537
SSURGO database *5:*302

Index terms	**Links**		
Stability analysis, for levee construction	*3:*287		
Stability length, parameterization of	*4:*1		
Stabilization	*5:*618		
of sludge	*1:*855	*1:*865	
Stabilization ponds	*1:*704	*1:*816	*1:*833
Stable isotope probing (SIP)	*1:*645		
See also Polar lipid-derived fatty acid-based SIP			
DNA	*1:*645		
RNA	*1:*645		
Stable isotopes			
as groundwater tracers	*5:*503		
in the water molecule	*5:*229		
Stagnant waters, solar radiation and temperature in	*3:*375		
Stakeholder approach	*3:*124		
Stakeholders			
integration among	*2:*575		
participation in flood management	*2:*680		
Standard aeration efficiency (SAE)	*1:*623		
Standard bioassay, as an exposure method	*2:*420		
Standardized bioassays	*2:*419	*2:*420	
Standardized Precipitation Index (SPI)	*4:*209	*4:*210	
Standard linear operating policy (SLOP) reservoir	*3:*384		
Standard mean ocean water (SMOW)	*5:*229		
Standard oxygen transfer rate (SOTR)	*1:*623		
Standards			
disinfection byproduct	*2:*92t		
need for	*1:*288		
water-handling	*2:*162		
water-quality	*2:*302		
Standard-state reduction potentials	*2:*465t		
Stand level transpiration	*3:*736		
Starch, as an indicator of organic contamination	*2:*442		
Star photospheres, water in	*4:*190		
State agricultural land use planning programs	*3:*596		
State clean water act programs, exemptions from	*4:*673		
State consumption advisories	*3:*120		
State government, approach to droughts	*2:*577	*2:*580	
State groundwater law	*4:*630		
State planning and management activities, drought-related	*2:*583		
State programs, private and market-based augmentation of	*4:*661		

Index terms **Links**

States
  drought policies of *2*:576
  industrial water use by *1*:621 *1*:622t
  protection of in-stream values by *4*:660
State treatment capacity *2*:616
State variable interactions *2*:274f
State water, allocation laws *4*:673
State water laws *2*:584
State Water Pollution Control Revolving Fund *4*:597
State water resource agencies *4*:649
State water supplies *2*:613
Static data *2*:622
"Static head" *3*:200 *3*:201f
Static level *5*:101
Static wastewater models *1*:731
Stationary fronts, fog and *4*:231
Statistical analysis, in amphipod sediment toxicity tests *2*:411
Statistical approaches/methods
  to pesticide assessment *5*:595
  to vulnerability assessment *5*:564
Statistical methods, for determining base flow recession constants *3*:25
Statistical properties, of drainage networks *3*:93
STATSGO database *5*:302
sT clathrate hydrate *4*:473
Steady fluid flows *3*:194
Steady-state calibration *3*:319
Steady-state flow aquifer tests *5*:491
Steady-state flows *5*:491
  between parallel plates *5*:652
  in a pipe *5*:654
Steady-state hydraulic theory *1*:58
Steady-state methods, for determining aquifer parameters *5*:492
Steady-state modeling *1*:132 *2*:272
Steady-state nonuniform flow *3*:251
Steady-state software tools *2*:326
Steady-state subsurface drainage theory *5*:95
Steady©. wastewater model *1*:731
Steam, natural *5*:157
Steam flooding, of oil-field brine *2*:286

Index terms **Links**

Steel industry impact, on surface water quality *3:*374
Steel tape method *5:*101
Stefan-Boltzmann law *3:*192
Stemflow *3:*171
Step-drawdown tests *5:*101 *5:*102 *5:*183 *5:*574
Step feed *1:*815
Step feed reactor *1:*833
Step pools *3:*67
Steps *3:*70
Steric stabilization *4:*424
Stewart and Kantrud wetland classification system *3:*497
Stiff and Davis stability index (S&DSI) *4:*415
Still-water channel units *3:*69t
  types of *3:*69t
St. John’s, Newfoundland, aquatic habitat and construction in *3:*509
Stochastic disaggregation techniques *3:*427
Stochastic methods *2:*333
Stochastic models *3:*337 *5:*309 *5:*664
Stochastic simulation *3:*263
  Box-Jenkins models in *3:*424
  components and solution procedure of *3:*423
  history of *3:*422
  hydrologic process representation in *3:*425
  multivariate stationary model for *3:*426
  of hydrosystems *3:*421
  point process models of *3:*428
  utility of *3:*422
Stock enhancement
  examples of *4:*125
  goals of *4:*125
  optimizing *4:*126
  problems of *4:*126
Stock enhancement techniques, marine *4:*124
Stockholm Convention on Persistent Organic Pollutants (SCOPOP) *2:*106
Stock price performance, water company *1:*312t
Stoichiometric analysis, of fluxes *4:*53
Stokes Law/Equation *3:*404 *3:*405 *5:*560 *5:*561 *5:*652 *5:*654
Stokes wave theory *4:*138
Stomatal aperture *3:*717 *3:*718

Index terms	**Links**			
Stomatal conductance model	*3:*714			
Stomatal density	*3:*714			
Stomates	*3:*714			
anatomically explicit models of	*3:*716			
behavior regulation by	*3:*718			
diffusion porometers and	*3:*717			
function of	*3:*717			
mass- or viscous-flow porometers and	*3:*717			
occurrence and dimensions of	*3:*715t			
study of	*3:*715			
Stone quarrying, impact on surface water quality	*3:*374			
Stony Brook River, flow-duration curve for	*3:*104	*3:*105f		
Storage				
off-stream	*1:*486			
as a watershed function	*3:*474			
Storage capacity, reservoir	*3:*409			
Storage coefficient (S_S)	*5:*480			
groundwater	*5:*129			
Storage facilities	*1:*408	*3:*430		
‘‘Storage lag’’ policy	*3:*386			
Storage reservoirs. *See also* Storage zoning reservoir				
classification of	*1:*406t			
functions of	*1:*408			
locating	*1:*410			
Storage system	*2:*554			
Storage tanks	*1:*203	*1:*448		
hydraulic design of	*1:*448			
leaking	*5:*603			
shape and volume of	*1:*410			
Storage volume, water quality and	*2:*367			
Storage water quality	*2:*367			
management of	*2:*368			
Storage zoning reservoir	*3:*385			
Storativity (S)	*5:*184	*5:*481	*5:*483	*5:*516
in confined and unconfined aquifers	*5:*129			
STORET system	*2:*315			
Storm-caused floods	*3:*45			
Storm drainage design, kinematic wave method for	*3:*242			
Storm hydrograph	*3:*476			
sample	*3:*439	*3:*440f		

Index terms	**Links**		
Storm peak runoff, predicting	*4:*317		
Storm period analysis, statistical approach to	*4:*338		
Storm periodicity, southwestern Spanish coast	*4:*341		
Storm-rainfall floods	*3:*511		
Storm record, southwestern Spanish coast	*4:*339		
Storm sewers	*2:*547	*3:*331	
Storm statistics, Chicago Station	*5:*79t		
Storm surge, estimating	*4:*3		
Storm water. *See also* Storm water management			
effluent limitations	*1:*759		
preliminary treatment of	*3:*500		
treatment	*1:*867		
Storm water discharges	*4:*657		
industrial/commercial	*1:*867		
Storm water management	*1:*499	*1:*799	*2:*192
constructed wetlands for	*1:*896		
municipal	*1:*866		
Storm Water Management Model (SWMM)	*2:*325	*3:*256	*3:*337
Stormwater management practices	*3:*443		
Storm water pollution, public agency activities and	*1:*867		
Stormwater pollution prevention plans (SWPPPs)	*4:*658		
Stormwater runoff	*2:*190		
Stormwater systems	*3:*498		
modeling	*3:*337		
Stormwater-wastewater collection system	*1:*748		
Strahler stream order system/method	*3:*32	*3:*93	
Straight water meter register	*1:*339		
Strategic analysis	*2:*629		
Stratification, lake	*3:*265	*3:*282	
Stratified lakes, oxygen depletion in	*4:*65		
Stratified random groundwater monitoring network	*5:*314		
Stratigraphic controls, on river water infiltration	*5:*679		
Stratosphere, climate of	*4:*182		
Straw, in algal control	*2:*5		
Stream animal communities, Hawaiian Islands	*4:*803		
Stream-aquifer interaction	*4:*359		
Streambed sediment			
metals in	*3:*355	*3:*356t	
organic compounds detected in	*3:*352		
polycyclic aromatic hydrocarbons in	*3:*352t		

Index terms **Links**

Streambed sediment *(Continued)*
  semivolatile organic and polychlorinated biphenyl compounds in *3:*357 *3:*359t
  trace elements detected in *3:*362
Stream classification *3:*65
  See also River reaches
  valley segment level of *3:*66
  watershed level of *3:*66
Streamflow *2:*547 *3:*439
  base flow contribution to *3:*22
  versus precipitation *4:*433
Streamflow recession *3:*22
  plots *3:*26
Stream-flow records, synthetic *3:*422
Stream-flow runoff, vegetative cover and *3:*30
Stream-flow simulation and reservoir regulation (SSARR) flow routing model *3:*255
Streamflow source zone *4:*670
Streamflow variability, quantitative measure of *3:*104
Streamflow variability indexes *3:*104
Stream gaging *5:*74
Stream health
  human impacts on *3:*39
  using fish index of biotic integrity to measure *3:*38
Streamlines *5:*182
Stream order *3:*32
StreamPlan (Spreadsheet Tool for River Environmental Assessment Management and Planning) *2:*328
Stream/river dilution capacity *3:*80
Streams. *See also* Rivers
  aggradation and degradation of *3:*420
  chronic acidification of *3:*3
  design flow in *3:*81
  dilution in *3:*84
  downstream fining in *3:*420
  as one-way flow systems *3:*437
  responses to earthquakes *5:*112
  sorting processes in *3:*420
  transport capacity of *3:*419
  versus gullies *3:*183
Streamside Management Areas (SMAs) *2:*187
STREAMS resource centers *2:*670

Index terms	**Links**	
Streamwater, acidification of	*3:*4	
Stress-inducing agents, *lux* genes for determining	*2:*174	
Stressor-based assessments	*2:*516	
Stress tensor	*5:*560	*5:*652
Stripping	*1:*707	
Strontium isotopes		
case studies of	*4:*576	
in water and rock	*4:*574	
Strontium isotope systematics	*4:*373	
Structural changes, influence on the hydrologic cycle	*4:*282	
Structural failure	*3:*259	
Structural flood adjustment measures	*3:*146	*3:*148
Structural repair systems	*1:*883	
Structure Intensive Pigment Index (SIPI)	*3:*722	
Study agencies, federal	*4:*650t	
Sturgeon species	*3:*127	
Stx-producing *Escherichia coli*	*2:*137	
See also Escherichia coli O157:H7		
Subchronic toxicity tests, sediment	*2:*385t	
Subcutaneous zone	*5:*243	
Subglacial lakes		
life in	*3:*506	*3:*507
sampling of	*3:*506	
Subglacial Lake Vostok	*3:*503	
Sublethal bioassays	*2:*351	
Sublimation	*4:*192	*4:*343
Submarine gas hydrates, remote sensing for	*4:*59	
Submerged aeration systems	*1:*624	
Submerged aquatic plants	*3:*743	
biomass and extent of	*3:*276	
lake water quality and	*3:*275	
processes in	*3:*276	
promoting	*3:*278	
types of	*3:*276	
Submerged flow, in culverts	*3:*76	
Submerged pens, role in trout hatching	*3:*458	
Submersible pumps	*1:*393	
Submitochondrial particle (SMP) assay	*2:*278	
as a biological monitoring tool	*2:*376	
environmental monitoring with	*2:*378	

Submitochondrial particle (SMP) assay *(Continued)*
  literature review of *2:*377
  sensitivity of *2:*377
  variants of *2:*378
Subparametric transformation mapping scheme *5:*657
Subsidence *2:*547
Substrate angle, dew amounts and *4:*204
Substrates, microbial growth in *3:*311
Substrate water binding, implications for O_2 evolution *4:*540
Subsurface, contaminant attenuation in *5:*578
Subsurface barriers *5:*434 *5:*617
Subsurface brines *5:*53
Subsurface contaminants *5:*579
Subsurface contaminant vapor, transport in *5:*550
Subsurface "dimples" *5:*92
Subsurface drainage *3:*451 *5:*94
  history of *5:*94
  horizontal *5:*95 *5:*97
  numerical solutions for *5:*97
  status and scope of *5:*99
  theories of *5:*95
Subsurface drip irrigation (SDI) *3:*621
  advantages of *3:*622
  extent of *3:*626
  systems *3:*615
Subsurface flow
  combined free and porous *5:*498
  governing model equation for *5:*499
Subsurface-flow constructed wetlands *1:*895
Subsurface flow systems *3:*365
Subsurface-flow wetlands *1:*787
Subsurface flushing methods, chemically enhanced *5:*440t
Subsurface intake system, well *5:*87
Subsurface Investigation, Phase II *5:*437
Subsurface pipe drainage. *See* Tile drainage
Subsurface recharge *5:*73
Subsurface redox chemistry *5:*413
Subsurface virus transport *1:*70
Subsurface water, vertical distribution of *5:*514
Subsurface water quality, soil nitrogen management and *3:*694

Index terms	**Links**	
Subtitle D landfill leachates, metal organic interactions in	*5:*258	
Subtitle D landfill regulations	*2:*166	*2:*167
Suburban watersheds, water quality in	*3:*441	
Subwatersheds, ranking of	*3:*158	
Successive alkalinity producing system (SAPS)	*2:*424	
Suction cups	*2:*341	
Suction lift	*3:*380	
Sudan-Egypt water dispute	*2:*591	
Sugar, water use, wastewater generation, and treatment proposals for	*1:*617f	
Sugarcane industry		
effluent characteristics of	*1:*619t	
ethanol production in	*1:*618f	
raw materials, products, and byproducts of	*1:*616f	
regulations in	*1:*619t	
wastewater treatment in	*1:*614	
Sulfate. *See also* SO_3 groups; SO_4^{2-}		
acidification and	*3:5*	
as an electron acceptor	*1:*691	
in groundwater	*2:*395	
Sulfate assimilation pathway, manipulation of	*5:*399	
Sulfate-reducing bacteria (SRB)	*1:*84	*1:*601
reactors	*5:5*	
Sulfate-reducing biofilm	*2:*231	
Sulfate-reducing to methanogenic conditions	*5:*583	
Sulfate transporter	*5:*398	
Sulfur-based reductants	*5:*291	
Sulfur compounds, solubility in water	*4:*561	
See also Sulphur compounds		
Sulfur deposition	*4:*380	
Sulfur dioxide dechlorination	*1:*169	
Sulfur fractionation	*4:*502	
Sulfur gases	*4:*378	
Sulfur mining, impact on surface water quality	*3:*374	
Sulfur-oxidizing bacteria	*2:*241	
Sulfur removal	*1:*913	
Sulphur compounds, volatile	*4:*88	
See also Sulfur compounds		
Sulphur isotopes	*5:*220	
as groundwater tracers	*5:*503	
Sumerian water consciousness	*4:*708	

Index terms **Links**

- Sun clocks *4:*703
- Sun-powered water cycle *4:*584
- Sun Valley, water conservation in *2:*635
- Supercooling, of water *4:*585
- Supercritical water oxidation (SCWO) *1:*579 *1:*874
- Superfund. *See* EPA Superfund
- Superfund sites
  - adsorptive filtration at *2:*365
  - groundwater treatment at *5:*39
- Supermicropores *1:*97
- Superoxide *4:*532
- Superparametric transformation mapping scheme *5:*657
- Supervisory Control and Data Acquisition (SCADA) system *2:*510
- Supplies, disaster *2:*529
- Supply equation *2:*615
- Supply side groundwater management *4:*632
- Supported capillary membrane sampler (SCMS) *4:*515
- Suppression subtractive hybridization (SSH) libraries *2:*62
- Surface aeration *1:*623
- Surface area, of porous solids *1:*97
- Surface complexation models *2:*363
- Surface drainage *5:*94
- Surface emissivities *3:*563
- Surface-flow constructed wetlands *1:*894
- Surface flow systems, free-water *3:*365
- Surface-flow wetlands *1:*787
- Surface functional groups, on activated carbons *2:*82
- Surface geophysical methods *5:*146
- Surface mines, abandoned *5:*3
- Surface runoff *3:*451
  - nitrogen losses from *3:*696
  - overland flow and *3:*452
  - river flow control and *3:*454
  - throughflow and *3:*453
- Surfaces
  - biofilm *3:*311
  - biofilm formation on *2:*228
  - biofouling and *1:*539
  - macrofouling of *1:*541
- Surface snow, remote sensing of *4:*324

Index terms	**Links**		
Surface stagnation			
stress due to	*3:*602		
from waterlogging	*3:*600		
Surface supplies, variability in	*2:*400		
Surface temperature	*3:*563		
Surface tension	*2:*547		
of chlorinated solvents	*5:*92		
Surface tension of electrolytes, role of dissolved gases in	*4:*452		
Surface trickle microirrigation	*3:*621		
Surface water(s). *See also* Surface water acidification; Surface water pollution			
algal-prone	*1:*444		
beryllium concentration in	*4:*395		
constructed wetlands for	*1:*897		
defined	*2:*547		
detection of pathogenic bacteria and biological contamination in	*2:*441		
effect of pesticides on	*3:*650		
environmental photochemistry in	*4:*529		
EU monitoring of	*2:*267		
form of iron in	*4:*497		
impacts on	*4:*266		
iron-oxidizing bacteria in	*2:*150		
metals in	*3:*443		
nitrate losses to	*3:*642		
Polydex disinfection of	*1:*384		
problems with	*5:*332		
quality of	*2:*585		
relationship to groundwater	*2:*396		
retention of	*1:*405		
technetium in	*4:*580		
typical analysis of	*3:*373t		
use-based classification of	*2:*184		
Surface water acidification	*3:*1	*3:*10	*4:*286
problem areas	*3:*8		
research on chronic effects of	*3:*4		
Surface water DDT, effects on bird abundance and reproduction	*3:*513		
Surface water pollution	*3:*373	*3:*444	
effects on drinking water	*3:*445		
in India	*3:*445		
sources of	*3:*374	*3:*444	*3:*657
treatment of	*3:*445		

Index terms **Links**

Surface water quality
  impact of soil phosphorus on *3:*701
  soil nitrogen management and *3:*694
  tile drains and *3:*729
Surface water resources
  Ganga Basin *3:*233
  Great Lakes *3:*177
  of India *2:*561
Surface water runoff *2:*516
  from construction sites *1:*538
  in the vadose zone *5:*534
Surface water sources, high control of *2:*399
Surface Water Supply Index (SWSI) *4:*212
Surface water tracers, dyes as *3:*96
Surface Water Treatment Rule (SWTR) *1:*70 *1:*175 *1:*196 *1:*197 *1:*244 *1:*457 *2:*140 *2:*195
Surface water treatment rule compliance technologies
  for disinfection *1:*461t
  for filtration *1:*462t
Surface water use, in arid lands *2:*476
Surfactants *1:*670 *5:*377
  analysis of *1:*671
  biodegradation of *1:*672
  water-soluble *1:*111
Surficial aquifers, characteristics of *5:*80
Surges *1:*202
Surge tanks *1:*262
Surging breakers *4:*137
Surveys
  electromagnetic *5:*680
  gravimeter *5:*148
  high-resolution sea floor *4:*77
  reservoir *3:*411
  riverine electromagnetic and
  seismic *5:*679
  single channel seismic reflection *5:*680
  snow *4:*337
  soil vapor *5:*548 *5:*549
Survey sampling *2:*161
Susceptibility assessment *1:*527

Index terms **Links**

Suspended growth biological treatment systems *1:*703
Suspended growth processes *1:*827
Suspended particles, chemical precipitation of *4:*588
Suspended sediment *2:*547
  concentration of *2:*547
  discharge *2:*547
  in urban stormwater runoff *3:*435
  vertical, lateral and temporal distribution of *3:*397
Suspended sediment load *3:*398 *3:*418
Suspended sediment samplers, types of *3:*398
Suspended solids *1:*901 *2:*547 *4:*587 *5:*421
  removal of *1:*54
Sustainability
  concepts related to *2:*625
  defined *2:*633
  imperatives and preconditions related to *2:*626t
  of irrigated agriculture *3:*584
  in the Nile Basin Initiative *2:*592
  practices in *2:*628
  thermodynamic aspects of *2:*626
  water resource *2:*624
Sustainable development *1:*651 *2:*624
  precautionary principle regarding *2:*601
  of water resources *5:*163
Sustainable flood-management decision making *2:*678
Sustainable management. *See also* Sustainable water management; Water resource sustainability
  of ecosystems *2:*636
  of fisheries *2:*635
  of freshwater resources *2:*634
  of natural resources *2:*633
''Sustainable urban drainage systems'' (SUDS) *1:*54
Sustainable wastewater management, in developing countries *1:*721
Sustainable water management. *See also* MEDIS project
  on Mediterranean islands *2:*638
  practices in *2:*639
Sustainable water supply, elements in *1:*653 *1:*654t
SUSTAINIS advanced study course *2:*641
Su Sung water clock tower *4:*704

Index terms **Links**

Swamps *4:*690
  freshwater and deepwater *3:*172
SWAP model *3:*680
Sweep coagulation *4:*425
Sweep-floc coagulation *1:*138
Swirl settlers (hydroclones) *1:*683
Swirl/vortex technologies *1:*784
Switching tensiometer *3:*725
SWMM model *5:*298
SWRRB model *5:*298
Symbolism, water *4:*785
Symmetrical moving average (SMA) scheme *3:*219 *3:*426
Symmetry boundary *5:*419
SYMTOX4 (Simplified Method Program-Variable Complexity Stream Toxics Model) software *2:*327
Synoptic meteorologists *4:*328
Synthetic chelates, role in lead phytoremediation *5:*383
Synthetic dilution water *2:*383
Synthetic organic chemicals/compounds (SOCs) *1:*248 *1:*350 *2:*268 *5:*250
  elimination using ozone *1:*355
  technologies for *1:*464t
Synthetic polymers
  as coagulants *4:*426
  in foulants *1:*417
Syntrophism/syntrophy
  in landfills *1:*696
  microorganisms and *3:*312
Syracuse, extraterritorial land use control in *1:*317
Syria-Israel water conflict *4:*754
Syrup-pan flow system *3:*165
System Control and Data Acquisition (SCADA) *1:*449
  See also SCADA systems
  buying *1:*451
  components and terminology *1:*451
  security and *1:*451
  small water systems and *1:*450
  water operations enhancement with *1:*450
System hydraulic performance, biofouling effects on *5:*37
System performance evaluation, fuzzy criteria for *2:*674
System problems, from corrosion *1:*152

Next Page

Index terms **Links**

System response data uncertainty *3*:326
Systems analysis *2*:683
  techniques *3*:384
Systems approach *2*:685
Systems engineering, for reservoir management *3*:386

T

Talbot, William Henry Fox *4*:766
Tamankaduwa, hydrography of *4*:778f
Tampa Bay water program *1*:309
Tank data *5*:164
Tank recharge (R_t) *5*:167
Tanks
  diffused air aeration system *1*:629
  settling *1*:452
  water distribution system *1*:208
Tanzania, ‘‘privatization by default’’ in *2*:632
Tapered aeration *1*:815
Tapproge system *1*:542
Tara Mines treatment wetlands *1*:898
  success of *1*:899
‘‘Target concentration,’’ *2*:357
Tariffs, water delivery *2*:473
Tasmanian Wilderness World Heritage Area (TWWHA) *4*:788
Taste, of water *1*:902
Taste control, using ozone *1*:355
Taxa richness *3*:37
Taylor-Galerkin time-stepping technique *5*:659
Taylor-Hood modeling scheme *5*:658
Taylor’s series expansion *4*:365
TCA *2*:317 *2*:318t
Tech Briefs *1*:151 *1*:158
  for drinking water treatment *1*:177 *1*:207
Technetium (Tc)
  chemistry of *4*:579
  regulatory standards for *4*:580
  remediation of *4*:581
  in water *4*:578
Technetium-99 *4*:578

Previous Page

Index terms	Links	
System response data uncertainty	*3*:326	
Systems analysis	*2*:683	
techniques	*3*:384	
Systems approach	*2*:685	
Systems engineering, for reservoir management	*3*:386	
T		
Talbot, William Henry Fox	*4*:766	
Tamankaduwa, hydrography of	*4*:778f	
Tampa Bay water program	*1*:309	
Tank data	*5*:164	
Tank recharge (R_t)	*5*:167	
Tanks		
diffused air aeration system	*1*:629	
settling	*1*:452	
water distribution system	*1*:208	
Tanzania, ‘‘privatization by default’’ in	*2*:632	
Tapered aeration	*1*:815	
Tapproge system	*1*:542	
Tara Mines treatment wetlands	*1*:898	
success of	*1*:899	
‘‘Target concentration,’’	*2*:357	
Tariffs, water delivery	*2*:473	
Tasmanian Wilderness World Heritage Area (TWWHA)	*4*:788	
Taste, of water	*1*:902	
Taste control, using ozone	*1*:355	
Taxa richness	*3*:37	
Taylor-Galerkin time-stepping technique	*5*:659	
Taylor-Hood modeling scheme	*5*:658	
Taylor’s series expansion	*4*:365	
TCA	*2*:317	*2*:318t
Tech Briefs	*1*:151	*1*:158
for drinking water treatment	*1*:177	*1*:207
Technetium (Tc)		
chemistry of	*4*:579	
regulatory standards for	*4*:580	
remediation of	*4*:581	
in water	*4*:578	
Technetium-99	*4*:578	

Technical Cooperation Committee for the Promotion of the Development and Environmental Protection of the Nile Basin (TECCONILE) *2:*592

Technical treatment system *1:*842

Technocracy, irrigation before *4:*740

Technologies
  applying to developing countries *1:*718
  desalination *1:*170
  high-throughput *2:*59
  for inorganic contaminants *1:*463
  new *1:*509
  public-private partnerships and *1:*47
  for radionuclides *1:*465t
  role in water markets *2:*568
  solidification/stabilization *1:*836
  for synthetic organic compounds *1:*464t
  telemetry *1:*147
  for volatile organic contaminants *1:*464t
  wastewater treatment *1:*808
  water reclamation *1:*806

Technology-based effluent limitations (TBELs) *4:*656
  variances from *4:*657

Technology development, Marine Instrumentation Laboratory *4:*70

Technology standards *1:*381

Technology transfer *1:*719

Tectonic lakes *3:*265

Tectonic plate movement, in the Caspian basin *2:*482

Telemetry *1:*202
  data *3:*316
  technologies *1:*147

Televising, pipeline *1:*884

Tempe pressure cells *5:*462

Temperature. *See also* Thermal entries
  bromate ion formation and *1:*360
  chlorine sensitivity and *2:*404
  of domestic sewage *1:*830
  effect on corrosion *1:*8
  effect on microorganisms *1:*522
  effect on nitrification *1:*753
  in groundwater tracing *5:*504
  hydrocarbon degradation and *5:*320
  influence on nitrification *3:*642

Index terms	**Links**			
Temperature *(Continued)*				
logs	*5:*155			
microbial growth and	*3:*311			
microbial regrowth and	*1:*344			
organism growth and	*1:*605			
of pond water	*3:*486			
profiling	*5:*411			
role in fish growth and production	*3:*130			
in shallow stagnant waters	*3:*375			
time-vertical distribution of	*3:*377			
Temperate climatic zone	*4:*258			
Temperature effect	*5:*230			
inorganic ion removal and	*4:*492			
Temperature fluctuation, of tidal streams	*4:*129			
Temperature of crystallization	*4:*585			
Temperature statistics, Chicago Station	*5:*79t			
Temperature stratification, in open waterbodies	*3:*190			
Temperature variability, interannual	*4:*175			
Temporal discretization	*5:*658			
Temporal fluctuations	*3:*310			
Temporal grid spacing	*3:*258			
Temporal resolution	*4:*319			
Temporary water hardness	*4:*453			
10-d survival and growth assay	*3:*51			
Tennessee Valley Authority (TVA)	*4:*613			
Tensiometer readings, interpreting	*3:*726t			
Tensiometers (TE)	*2:*341	*5:*180	*5:*533	*5:*540
with electronic reader	*3:*725			
installing, servicing, and troubleshooting	*3:*727			
limitations of	*3:*728			
preparing for field use	*3:*727			
vacuum gauge	*3:*724			
working principle of	*3:*726			
Tensiometric soil moisture measurement	*4:*488			
Tension disk infiltrometers	*5:*215			
Tension-free lysimeter	*2:*340			
Tension infiltrometers	*4:*487	*5:*539		
Tensionless column lysimeters	*5:*488			
Tension lysimeters	*2:*341			
installation and sampling procedures for	*5:*490			

Index terms **Links**

Tension-saturated zone *5:*354
Terminal electron acceptor (TEA) *5:*319 *5:*528 *5:*580 *5:*581
Tentatively identified compounds (TICs) *2:*308
Ten Thousand Islands estuary *3:*72
Terminal electron-accepting processes (TEAPs) *2:*469 *5:*585
Terminal electron acceptor (TEA) *1:*640
Terminal restriction fragment length polymorphism (tRFLP), community profiling using *1:*643
Ternary complexes *2:*208
Ternary effluent tests *1:*775
Terraces, soil-conservation *3:*550
Terrestrial bacteria, luminescent marking of *2:*453
Terrestrial cadmium *5:*615
Terrestrial nonpoint source pollutants *3:*282
Terrestrial organisms, toxicity of lead to *2:*434
Terrestrial point source pollutants *3:*281
Terrorism
  point-of-use/point-of-entry systems and *1:*381
  valves and *1:*484
  water security and *1:*158
Terrorist threats, against wastewater utilities *1:*870
Tertiary wastewater treatment *1:*827 *2:*547
Test development, for regulatory purposes *2:*458
Testing
  of aquifers *5:*182
  of meters *1:*340
  of pipelines *1:*888
  for point repairs *1:*889
  raw water *1:*440
  of repaired pipe *1:*890
  waterborne radon *1:*51
water meter *1:*491
Test kits, commercial *2:*384
Test organisms, ciliated protists as *2:*413
Test ruggedness *2:*459
Test systems, for sediment toxicity assessment *2:*384
Test validation *2:*459
Tetrachlorodibenzo-*p*-dioxin (TCDD) *2:*108
Tetrachloroethene (PCE) *5:*581
Tetrachloroethylene. *See* Perchloroethylene (PCE)

Index terms	**Links**		
Tetrahymena, in laboratory research	*2:*413		
Tetrahymena pyriformis, as a test organism	*2:*413	*2:*414	
Tetrahymena termophila, in toxicity assessment	*2:*413		
TETRATOX test	*2:*414		
Texas vulnerability assessment case study	*5:*565		
Thales' water theory	*4:*708		
Thallium (Tl), maximum contaminant level of	*2:*235t		
The Groundwater Foundation (TGF)	*2:*518	*2:*519	*4:*713
Theis equation, in aquifer test data analysis	*5:*183		
The Nature Conservancy, Arizona Chapter (TNCA)	*3:*57	*3:*58	
Thermal carbon activation	*1:*94		
Thermal conductivity	*5:*173		
Thermal conductivity detectors	*2:*308		
Thermal cycles, fish production and	*3:*134		
Thermal desalination	*1:*170		
Thermal diffusivity	*5:*173		
Thermal drying			
of biosolids	*1:*649		
of sludge	*1:*865		
Thermal infrared (TIR) domain	*3:*720		
Thermal *in situ* groundwater treatment	*5:*440		
Thermalization process	*3:*692		
Thermal kinetic window (TKW)	*3:*719	*3:*721	
Thermal oxidation systems, for odor removal	*1:*915		
Thermal pollution	*2:*547		
as a water quality indicator	*2:*266		
Thermal power generation, water use for	*2:*662		
Thermal processing, of sludge	*1:*858		
Thermal properties, of water	*4:*583		
Thermal regeneration	*1:*918		
Thermal theory of cyclones	*4:*353		
Thermal velocity-groundwater velocity ratio	*5:*173		
Thermal wastewater treatment technologies	*1:*812		
"Thermochemical cycle"	*4:*477		
Thermodynamics			
effect on adsorption	*3:*301		
entropy in	*4:*218		
laws of	*5:*170		

Thermoelectric power water use *2:*547
  in the United States *2:*653
Thermoelectric water use *1:*560
Thermometer, first *4:*711
Thermophysical characteristics, of sludge *1:*863
Thermosiphon solar water heaters *1:*65
Thessaloniki Industrial Area, wastewater values from *1:*569t
θ time-stepping method *5:*659
THF hydrate *4:*473
Thickening, of sludge *1:*854 *1:*864
Thiem method, Ernst's modification of *5:*495
Thiem's equilibrium equation *5:*493
Thiobacillus ferrooxidans *2:*149 *2:*150 3:14
Thiol-containing peptides, roles of *3:*609
Thirst, pain and *4:*724
Thixotropic fluids *5:*557
THM formation potential (THMFP) *2:*117
  See also Trihalomethanes (THMs)
THM precursor removal *2:*116
Thornthwaite potential *4:*260
Thornthwaites index *3:*32
Thornthwaite water balance diagrams *4:*262
Threat agents *1:*87
  monitoring water for *1:*89
  nature of *1:*87
  rapid detection of *1:*91
  survival in water *1:*88
  water-supply *1:*88
Threat contaminant (TC) *2:*348
Threats, to water resources *1:*438
Three-component base flow separation *3:*26
Three-dimensional numerical models *3:*322
3-Dimensional voltage distribution *4:*443
"Three measures on river treatment (flood management)" *3:*47
Threshold bromide ion concentration *1:*359
Threshold dose *4:*674
Threshold effects level (TEL) *3:*351
Threshold limit value (TLV) *4:*555 *4:*561
Threshold odor number (TON) *1:*355 *1:*760

Index terms	**Links**		
Threshold treatment, for scale control	*1:*548		
Thrombotic thrombocytopenic purpura (TTP)	*2:*137		
Throughfall, precipitation	*3:*171		
Throughflow	*3:*453		
Throughput bed volume (BV)	*1:*103		
Throughput volume (V_t)	*1:*103		
Thu Duc Water Station, Ho Chi Minh City	*2:*553		
Thunder Bay National Marine Sanctuary and Underwater Preserve	*4:*64		
Thunderstorms	*4:*362		
role in flooding	*3:*145		
Tidal currents	*4:*24		
Tidal efficiency	*5:*497		
Tidal energy	*4:*46		
Tidal stream variability			
physical and chemical	*4:*128		
seasonal	*4:*130		
short-term	*4:*128		
Tidal wetlands	*3:*72		
Tides	*4:*24		
as a forcing factor	*3:*322		
Tiered testing	*2:*430	*2:*460	*2:*461f
Tigris-Euphrates River Basin	*4:*754		
Tigris-Euphrates waters, conflicts over	*4:*755		
Tijuana River waters, allocating	*4:*645		
Tile drainage	*3:*729	*3:*742	
impacts of	*3:*731		
pollutant management and	*3:*730		
Tillage			
traditional	*3:*650		
wind erosion and	*3:*552		
Timber Culture Act of 1873	*2:*512		
Timber harvesting, managing	*2:*187		
Time-area curve (TAC)	*3:*62		
Time-area (TA) diagram	*3:*62		
dimensionless cumulative	*3:*64		
Time-area histogram (TAH)	*3:*60	*3:*64	*3:*65
constructing	*3:*61		
total hydrograph derived from	*3:*62t		
Time-dependent fluids	*5:*557		
Time-dependent models	*2:*272		

Time-dependent partial differential equations *5:*517
''Time-discharge'' histogram (TDH) *3:*65
Time-domain electromagnetic induction *5:*150
Time-domain reflectometry (TDR) *3:*676 *4:*488 *5:*533 *5:*540
  devices using *5:*461
Time-dynamic speciation, body concentrations and *2:*215
Time immemorial water rights *4:*690
Time-independent fluids *5:*556
Time of concentration (T_c) *3:*61
  methods and formulas for computing *3:*470
  for a series of planes *3:*243
  in watersheds *3:*469
Time of travel, in a channel *3:*244
Time-resolved laser fluorescence spectroscopy (TRLFS) *2:*207
Time series analysis *3:*422 *3:*423 *4:*339
Time-variable isochrone method *3:*64
Time-varying exposure concentrations *2:*213
Tissue. *See also* Sculpin tissue
  metals in *3:*361t
  organochlorine compounds in *3:*352 *3:*360t *3:*362f
  trace elements detected in *3:*362
Tissue analysis, for trace metal concentration *2:*65
Tissue quality guidelines (TQGs) *2:*428
Tissue quality values (TQVs) *2:*351
Title IV. *See* Acid Deposition Control Program (Title IV)
TMDL plan *2:*188
  See also Total maximum daily load (TMDL)
Tobago, megawatersheds exploration in *5:*269
Toe drain, pervious *3:*290
Toe trenches, pervious *3:*289
Toilet dam *2:*665
Toilet flushing, water quality criteria for *1:*17t
Toilet leaks *2:*664
Toilet retrofits *2:*665
Toilets, water-free *1:*678
Tolerance measure *3:*37
Tomatoes, genetically engineered *3:*686
Tool development, for ecotoxicogenomic research and biomonitoring *2:*62
  See also Software tools

Index terms **Links**

Tooth decay *1:*254
  See also Fluoridation
Topogenic areas *1:*404
Topographic impacts, of
  waterlogging *3:*741
Topography
  regional flow systems and *5:*420
  role in flooding *3:*143
Topologically distinct channel networks (TDCN) *3:*32
Tordon 75D investigation *2:*280
Tornadic waterspouts *4:*347
Tort claims, contamination-related *3:*169
Tory Lake, biomanipulation trials in *2:*57t
Total ammonia nitrogen (TAN) *3:*131
Total available chlorine (TAC) *1:*128
Total chlorine residual *2:*399
Total Coliform Rule *2:*140
  compliance technology for *1:*463t
Total dissolved inorganic carbon (TDIC) *5:*224
Total dissolved solids (TDS) *1:*538 *1:*810 *1:*812 *1:*902 *2:*157 *5:*281
  effect on corrosion *1:*8
  electric conductivity and *4:*432
  in freshwater *4:*449
  in groundwater *2:*394
  salinity and *5:*207
Total load *3:*418
Total Maximum Daily Loads (TMDLs) *1:*305 *1:*759 *1:*801 *2:*21 *2:*185 *2:*188 *2:*233 *4:*596 *4:*598 *4:*671
  cost-effectiveness analysis for *2:*131
  of impaired waters *2:*476
  implementation of *4:*599
  programs *2:*127 *2:*128 *2:*188
Total maximum daily load program *4:*663
Total oil and grease analysis *2:*311
Total organic carbon (TOC) *2:*40
  in urban stormwater runoff *3:*435
Total petroleum hydrocarbon (TPH) oil and grease *2:*311
Total phenolics analysis *2:*309

Index terms **Links**

Total residual chlorine/total residual oxidant (TRC/TRO) *1:*128
Total suspended solids (TSS) *1:*455 *1:*538 *1:*599 *1:*683 *1:*739 *1:*830
  along the Blackstone River *2:*258
Total water hardness *4:*453
"Total Water Management" *2:*587
TOXcontrol-BioMonitor *2:*174
Toxic agents/chemicals/substances *2:*29
  in drinking water *2:*371
  lux genes for determining *2:*174
  precautionary principle regarding *2:*600
Toxic algae *3:*108
Toxicant detection, luminescent bacterial biosensors for *2:*453
Toxicant exposure, forms of *3:*117
Toxicants
  defined *2:*170
  in sediment *2:*426
Toxicant-specific luminescent biosensors *2:*455t
Toxic chemicals, in sediment *3:*508
Toxic constituent manuals (EPA) *2:*381
Toxic cyanobacteria, ecology of *2:*388
Toxic elements, in biodegradation processes *5:*321
Toxicity. *See also* Aquatic toxicity
  arsenic *1:*1 *2:*15
  bacterial bioluminescence and *2:*47
  chlorine *2:*402
  of detergents *1:*672
  of industrial effluents *1:*565
  nitrate *4:*519
  numbers associated with *2:*47
  persistence of *2:*381
  predicting *2:*415
  uranium *5:*641
Toxicity assays, of SPMD dialysates *2:*171
  See also Toxicity assessment; Toxicity bioassays; Toxicity testing; Toxicological assays
Toxicity assessment *4:*602
  See also Toxicity testing
  ciliated protists in *2:*413
  direct *2:*46
  of industrial effluents *1:*566

Toxicity assessment *(Continued)*
  of industrial wastewaters *1:*566
  new approaches in *2:*414
  quantitative structure-activity relationships in *2:*415
  use of semipermeable membrane devices for *5:*672
Toxicity-based effluent limitations *4:*657
Toxicity/bioaccumulation testing *4:*600
Toxicity bioassays *2:*453
  See also Toxicity assays
Toxicity biomarkers *2:*29 *2:*31
Toxicity Characteristic Leaching Procedure (TCLP) *1:*853
Toxicity consents *2:*46
Toxicity controls, whole effluent *2:*382
Toxicity identification and evaluations (TIEs) *2:*380 *2:*383 *2:*411 *4:*602
‘‘Toxicity reduction evaluation’’ (TRE) *2:*381
Toxicity testing. *See also* Sediment toxicity testing; Toxicity assays; Toxicity assessment
  bioluminescent biosensors for *2:*45
  environmental quality standards and *2:*46
  legislative drivers for *2:*45
  of sediment *2:*351 *2:*383
  trends in *3:*116
  use of fish species in *3:*116
  variations in *3:*117
Toxic metal biosorption *2:*68
  mechanisms of *2:*70
Toxic metal ions, bonding of *1:*586
Toxic metals, persistence in the water environment *2:*233
Toxicogenomic experiments *2:*62
Toxicological assays, problems with *2:*415
  See also Toxicity assays; Toxicological evaluation
Toxicological database *2:*409
Toxicological evaluation, using fish cells in *3:*115
  See also Toxicological assays
Toxicological methods *2:*278
Toxicological testing. *See* Marine toxicological testing; Toxicity testing; Toxicological assays
Toxicology, of disinfectants *1:*195

Index terms	**Links**	
Toxic pollutants	*4:*670	
in Indian rivers	*3:*449	
pursuant to Clean Water Act	*4:*596t	
Toxic Release Inventory (TRI)	*1:*799	*5:*381
Toxic sediments, transport of	*3:*15	
Toxic Substances Control Act	*1:*531	
Toxic Substances Hydrology Program study	*5:*605	
Toxin-producing blue-green algae	*3:*189	
Toxins		
algal	*2:*387	
in blue-green algae blooms	*3:*189	
Toxoplasma	*1:*279	
Toxoplasmosis	*1:*182	
ToxScreen test	*2:*173	
Trace element contamination, in groundwater	*2:*143	
Trace element pollution	*4:*109	
Trace elements	*3:*454	
abundance, distribution, and behavior of	*4:*110	
in aquatic biota	*3:*456	
effects of geology on	*3:*457	
effects of human activity on	*3:*457	
in Puget Sound Basin sediment and fish	*3:*349	
in River Yamuna groundwater	*2:*396	
in sediments	*3:*455	
sources of	*4:*111	
in streambed sediment and whole sculpin tissue	*3:*362	
toxic effects of	*4:*111	
in water	*3:*455	
Trace gases		
atmospherically active	*4:*85t	
fluxes of	*4:*149	
Trace metal biomonitors, macrophytes as	*2:*64	
Trace metal pollution		
in the aquatic environment	*2:*64	
metallothioneins as indicators of	*2:*406	
Trace metals, sorption onto activated carbon	*2:*83	
Trace metal speciation	*2:*202	
chemical reactions affecting	*2:*203	
measurement of	*2:*204	
modeling	*2:*204	

Index terms **Links**

Trace substances, health implications of *1:*289
Tradable permits system *2:*133
Tradable rights
  issues related to *2:*643
  market allocation of water and *2:*644
  meeting water needs with *2:*643
  security of tenure and *2:*644
“Tragedy of the Commons, The” *2:*556 *2:*580 *2:*625
Transaction costs, water market *2:*499
Transboundary diagnostic analysis (TDA) *4:*643
Transboundary water conflicts, Nile Basin *2:*590
Transboundary waters
  in Central America *4:*641
  in Latin America *4:*636
  managing *4:*643
  Mexico-U.S. *4:*636
  in South America *4:*642
Transboundary water treaties, in the Americas *4:*637 *4:*641t
Transcription factor overexpression, in selenium phytoremediation *5:*400
Transect groundwater monitoring network *5:*314
Transfer functions *4:*93
Transfer of development rights (TDR) programs *3:*596 *3:*597
Transfer rules *2:*501
Transient analysis, of rapid flow changes *1:*213
Transinformation *4:*219
Transitional subunit groundwater monitoring network *5:*315
Translocation factor (TF) *3:*35
Transmembrane pressure (TMP) pulsing *1:*594
Transmissibility, ground water *2:*547
Transmission electron microscopy (TEM) *2:*310
Transmissive media systems, biofilms in *5:*36
Transmissivity (T) *5:*507
  of an aquifer *5:*492
  groundwater *5:*129
Transpiration *2:*547 *3:*718 *3:*732 *4:*192 *4:*345 *4:*434
  controls on *3:*733
  effects of *4:*346
  effects on contaminant movement *5:*378
  estimates of *3:*738

Transpiration *(Continued)*
  factors affecting *4:*345
  fast versus delayed *3:*237
  leaf level *3:*735
  measuring *3:*737 *4:*346
  modeling *3:*736
  quantification methods for *3:*733 *3:*735
  soil characteristics and *3:*735
  stand level *3:*736
  tree level *3:*735
  vegetation characteristics and *3:*734
  in the water cycle *4:*196
Transpiration rate (*Tr*) *4:*345
Transport
  of cadmium *5:*617
  of contaminant vapors *5:*550
  from groundwater to human stomachs *5:*449
  of lead *5:*647
  of mercury *5:*644
  of MtBE *5:*318
  of nitrate *5:*629
  particulate *5:*349
  of perchlorate *5:*632
  of uranium *5:*640
  through the vadose zone *5:*77
  of vinyl chloride *5:*635
Transport equations, one-dimensional *5:*525
Transport mechanisms *1:*246 *3:*181
  fish as *3:*283
Transport modeling, basic concepts in *5:*31
Transport phenomena, of dissolved matter *2:*105
Transport simulation programs *5:*34t
Transversal flow membrane (TFM) modules *1:*593
Transverse dispersion *5:*46 *5:*273
Transverse scour pools *3:*68
Trash, in urban stormwater runoff *3:*435
Trash-Trap™ system *1:*785
Travel cost method *4:*609
Travel cost models *2:*130

Index terms **Links**

Travel time(s)
  variation in *3:*64
  in watersheds *3:*469
Travel time (T_w) formula *3:*63
Travel-time map, cumulative *3:*65
Tray aerator *1:*313
Treated effluent, characteristics of *1:*819
Treated water
  contaminants in *5:*421
  reinjection of *5:*421
Treaties, transboundary water *4:*637 *4:*641t
Treatment
  effect on assimilable organic carbon and biodegradable dissolved organic carbon *2:*225
  dilution equivalent to *4:*104
  effect on speciation *2:*75
  loss of *2:*48
  as a remediation strategy *2:*358
Treatment practices, for wastewater reuse *3:*670
  See also Treatment technique entries
Treatment processes
  domestic sewage *1:*832
  fate of POPs throughout *1:*769
Treatment systems. *See also* Integrated treatment systems
  alternative *1:*678 *1:*679
  stressed *2:*663
Treatment technique (TT) *1:*477
Treatment techniques/technologies. *See also* Treatment practices; Treatment technique (TT)
  developing *1:*425
  for hydrocarbons *1:*575
  for small drinking water systems *1:*457
Treatment Work Construction Grants *4:*664
Treaty of Guadalupe Hidalgo *4:*643
Tree crops, microirrigated *3:*625
Tree level transpiration *3:*735
Trees, transpiration of *5:*85
Trenches, alkaline recharge *5:*2
Trenchless repair and rehabilitation techniques *1:*876
  manhole renovation using *1:*882
  pipeline renovation using *1:*876

Trench pools *3:*67
Trench shoring *1:*402
Trending heterogeneity *5:*509
Trend stations *2:*178
Triarylmethane dyes, QSAR case study using *3:*99
Tribal reserved rights *4:*689
Tribromoacetic acid, health effects of *1:*267
Tributary *2:*547
Trichloroacetaldehyde monohydrate, health effects of *1:*269
Trichloroacetic acid, health effects of *1:*267
Trichloroacetonitrile, health effects of *1:*268
Trichloroethene (TCE) *5:*581
  cometabolism of *5:*584
Trichloroethylene (TCE) *2:*317
  aerobic degradation of *5:*43
  biotransformation of *1:*691
  Fenton's Reaction and *4:*446
  removal of *3:*368
2,4,6-Trichlorophenol (TCP) adsorption *1:*116
Trickling filters *1:*833
  advantages and disadvantages of *2:*155
  for iron removal *2:*154
Trihalomethane formation potential (THMFP) *1:*139
  See also THM formation potential (THMFP)
Trihalomethanes (THMs) *1:*129 *1:*313 *2:*91 *2:*115 *2:*116 *2:*268
  See also THM entries
  bromide influence on formation of *2:*74
  brominated *2:*74
  health effects of *1:*264
  precursors of *1:*432
Trimethylbenzene (TMB) *5:*587
Trinidad, megawatersheds exploration in *5:*269
Trinitrotoluene (TNT) plumes *5:*77
Triolein-containing SPMDs *5:*672
Tritium *5:*231
  anthropogenic *5:*69
  concentration of *5:*70
  groundwater dating with *5:*65 *5:*69
  in groundwater tracing *5:*503

Tritium *(Continued)*
  measurement methods for *5:*68
  in precipitation *5:*67f
  production and decay behavior of *5:*65
Tritium/helium ratios, dating young groundwater using *5:*71
  See also H-He dating
Tritium units (TU) *5:*66
Trondheim Biomonitoring System (TBS) *2:*28 *2:*30
"Trophic-dynamic system" *3:*269
Trophic levels, disaggregation of *4:*151
Trophic state, of lakes *3:*267
Trophic state classification *4:*66
Trophic upsurge *3:*206
Tropical cyclone activity *4:*176
Tropical cyclones, role in flooding *3:*144
Tropical lakes *3:*270
Tropical Ocean-Global Atmosphere (TOGA) *4:*22
Tropical Rainfall Measuring Mission (TRMM) *4:*305 *4:*309 *4:*312
Tropical reservoirs, studies of *3:*181
Tropospheric lapse rate *4:*370
Trout, innovative hatching program for *3:*458
Trout farms, in Poland *4:*722
"Tsunameters" *4:*90
Tsunami Hazard Mitigation Program *4:*89
Tsunamis *4:*160
  research on *4:*162
Tsunami simulations *4:*162
Tsunami Warning System *4:*160 *4:*162
Tuberculation *1:*400
Tuberculosis (TB) *1:*182
Tube-type clarifier package plants *1:*515
Tube well *3:*594
Tubewell drainage *3:*732 *5:*98
Tucson, Arizona, overdraft in *5:*341
Tucson Basin, perched groundwater in *5:*353
Tumen River Economic Development Area Project (TREDA) *2:*444
Tumorigenicity, of arsenic *2:*8

Index terms	Links		
Turbidity	*1:*485	*1:*599	*1:*901
defined	*2:*547		
inorganic	*3:*108		
lake	*3:*278		
measurements of	*2:*344		
removal of	*1:*250		
of urban stormwater runoff	*3:*435		
Turbidity tube	*1:*599		
Turbine design	*3:*199		
Turbine meters	*1:*338	*1:*340	
Turbine pumps, deep-well	*3:*664		
Turbine spillways, degassing fluxes at	*3:*206		
Turbulence			
in estuaries and bays	*4:*25		
interaction with propagation	*4:*296		
thermal and mechanical	*4:*1		
Turbulent channel units	*3:*70		
Turbulent Eddy Profiler	*4:*296		
Turbulent vertical convection	*4:*367		
Turf areas, practical	*3:*750		
Turfgrass	*3:*706	*3:*707	
benefits of	*3:*750		
reducing irrigation of	*3:*751		
selecting	*3:*751		
sod	*3:*557		
xeriscape principles	*3:*751		
Turf watering	*3:*556		
Turkish South East Anatolia Project (GAP)	*4:*754		
Turnover number (TON)	*1:*791		
Two-dimensional numerical models	*3:*322		
2DLEAF model	*3:*716		
Two-electrode measurement technique	*4:*432		
Two-moment information theory	*4:*476		
Type II diabetes, dehydration and	*4:*725		
Typhoid fever	*2:*337		
Typhoid rates	*1:*286t		
Typhoons	*4:*194		
Typhoon wave	*4:*139		

Index terms	Links
U	
U.K. National Rivers Authority, water supply categories used by	2:614 2:615t
See also British entries	
Ultrafast vibrational energy	4:513
Ultrafiltration (UF)	1:333t 1:334 1:336t 1:459 1:488 1:707 1:810 1:817 1:828 1:916
See also Polymer-assisted ultrafiltration (PAUF)	
history of	1:916
pore sizes for	2:101
Ultrafiltration methods	4:29
Ultrafiltration tests	1:919
Ultrahigh purity (UHP) nitrogen	4:421
Ultra-low-volume (ULV) toilets	2:665
Ultramicropores	1:97
Ultrasonic irradiation	1:875
Ultrasonic meters	1:338
Ultrasound	4:571
Ultraviolet (UV) radiation	1:199 1:458 2:119 2:244 2:374
See also UV entries	
at CSO facilities	1:786
technology of	1:471
water disinfection using	1:471
Ultraviolet disinfection	1:165 1:380 1:466 1:523 1:604 1:817 1:834
See also Ultraviolet irradiation	
advantages and limitations of	1:466 1:472
domestic sewage	1:834
experiments in	1:167 1:168
monitoring and operation requirements for	1:468
process of	1:467
Ultraviolet irradiation	1:469
See also Ultraviolet (UV) radiation	
Unaccounted-for water (UFW), calculating	1:318
Unaccounted for water indicator	2:492 2:496
Uncertainty	2:599
in hydrologic data	2:621
participatory multicriteria decision-making under	2:681
in pesticide assessments	5:358
in pesticide vulnerability assessment	5:598

Index terms	**Links**			
Uncertainty *(Continued)*				
in rainfall-runoff modeling	*4:*297			
sources of	*2:*674	*3:*325		
types of	*2:*680			
Uncertainty analysis	*3:*117			
approaches to	*3:*326			
in watershed modeling	*3:*325			
Unconfined aquifer flows	*5:*492			
Unconfined aquifers	*5:*11	*5:*88	*5:*145	*5:*483
	*5:*497			
artificial recharge of	*5:*11			
barometric efficiency of	*4:*168			
drawdown in	*5:*103			
storage in	*5:*515			
storativity in	*5:*129			
water flow in	*5:*662			
Unconfined groundwater	*5:*662			
case study of	*5:*665			
Underground injection control (UIC)	*4:*679			
Underground pipeline	*1:*883			
assessment of	*1:*884			
repair and renewal of	*1:*887			
Underground water storage	*3:*164			
Undersea “windmills”	*4:*24			
Underwater acoustical monitoring	*4:*91			
See also Underwater sound detection				
Underwater light, availability of	*3:*276			
Underwater sound detection	*4:*570			
See also Underwater acoustical monitoring				
Underwater volcanoes	*4:*90			
Underwater weapon tests, as water pollution	*4:*102			
Undeveloped areas, impact on groundwater quality	*5:*250			
Ungauged watersheds, modeling	*3:*342			
Uniform fluid flows	*3:*194			
“Uniform numeric limitations”	*1:*755			
Unilateralism, in water management	*3:*19			
Un-ionized ammonia concentrations	*3:*131			
United Nations, public trust guidelines of	*2:*609			
United Nations Framework Convention on Climate Change (UNFCCC)	*3:*208			

Index terms	**Links**			
United Nations Industrial Development Organization (UNIDO), planning analysis objectives	*2:*683			
United Nations World Water Development Report	*2:*548	*2:*549		
United Soil Classification System (USCS)	*3:*689			
calculation	*3:*61t			
United States. *See also* Federal entries; U.S. entries; Western United States				
acid mine drainage in	*3:*10	*5:*1		
ambient standards for waters of	*1:*755			
aquatic environment clean up by	*2:*184			
classification of wetlands and deepwater habitats in	*3:*496			
disinfectant use in	*1:*195			
distribution of pharmaceuticals in	*1:*373			
drinking water quality in	*2:*373			
evolution of water use in	*2:*646t			
flood damages in	*3:*146			
groundwater contamination from municipal landfills in	*5:*253			
history of Water Pollution Control Act in	*4:*651			
irrigation in	*3:*586			
Lake Baikal studies by	*3:*21			
new water sources in	*2:*649			
1990 industrial water use in	*1:*620			
nitrate concentrations in	*1:*31			
nitrate data for	*1:*32t			
percentage of precipitation in	*4:*226			
pesticide use in	*3:*655			
selenium containment strategy in	*2:*358			
source-water protection in	*2:*312			
total water withdrawals in	*2:*646t			
water consumption patterns in	*2:*651			
water markets in	*2:*500			
water quality management in	*2:*193			
water quality policy in	*2:*127			
water quality standards in	*1:*476	*2:*268		
water saving measures in	*2:*649			
water use in	*1:*805	*2:*645	*2:*650	
wetlands in	*3:*493	*4:*690		
United States Environmental Protection Agency (USEPA)	*1:*33	*1:*145	*1:*287	*1:*531
See also Environmental Protection Agency (EPA); USEPA entries				
drinking water standards	*1:*183			
fish advisories	*3:*119			
maximum contaminant levels	*2:*233			

Index terms	Links		
United States Environmental Protection Agency (USEPA) *(Continued)*			
water quality standards	*2:*373		
water reuse guidelines	*2:*611		
United States-Mexico border waters	*4:*636		
binational agreements related to	*4:*645		
current developments regarding	*4:*646		
treaties and conventions regarding	*4:*643		
United States Weather Bureau	*4:*347		
Unit flood response approach	*3:*156		
Unit hydrograph (UH)	*3:*221		
dimensions of	*4:*356		
Unit hydrograph models	*4:*357		
Unit hydrograph theory	*4:*355		
limitations of	*4:*356		
mathematical representation in	*4:*356		
Univariate spectral analysis	*4:*339		
Universal metering	*1:*494		
Universal soil loss equation (USLE)	*2:*535		
Universities, land-grant	*3:*597		
Unloading fractures	*5:*177		
Unplanned water recycling	*2:*610		
Unrecognized organic chemicals, in urban stormwater runoff	*3:*435		
Unrecognized pollutants	*3:*371		
Unregulated Contaminant Monitoring Rule/ Regulations (UCMR)	*2:*345	*4:*678	
Unsaturated hydraulic conductivity	*5*:126	*5*:471	
Unsaturated soil properties, modeling	*5*:538		
Unsaturated zone	*2:*548	*5:*600	
vapor transport in	*5:*543		
water in	*5:*531		
Unsteady-state flows	*5:*491		
Unsteady-state software tools	*2:*327		
Unsteady-state subsurface drainage theory	*5:*96		
Unsteady-state uniform flow	*3:*251		
Unsustainable water supply systems	*1:*505		
Upflow anaerobic sludge bed (UASB)	*1:*906		
Upflow anaerobic sludge bed reactor	*1*:518	*1*:519	*1*:618f
Upflow filtration	*2:*366		
Upflow roughing filters	*1:*237	*1:*239	*1:*240
Upland-‘‘confined’’ dredged sediment disposal	*2:*124		
Upper Mississippi River Source Water Protection Initiative	*2:*312		

Index terms	Links		
Upward contaminant vapor flux	*5:*553		
Upward vapor migration, from groundwater devolatilization	*5:*553		
Upwelling	*4:*24		
Uranine	*3:*96		
in groundwater tracing	*5:*505		
Uranium (U)	*4:*542		
fate and transport of	*5:*640		
in groundwater	*5:*640		
health risks of	*1:*395	*5:*641	
processing of	*5:*640		
remediation of	*5:*641		
uses of	*5:*641		
Uranium series isotopes	*5:*232		
Urban creeks	*3:*432		
water quality in	*3:*433		
Urban Drainage and Sewer Master Plan, Ho Chi Minh City	*2:*553		
Urban drainage models	*3:*337		
calibration and verification of	*3:*338		
future of	*3:*341		
real-time hydrologic information and	*1:*125		
Urban drainage systems, modeling real-time control and active real-time control of	*3:*340		
Urban flooding	*3:*159		
causes of	*3:*160		
cities that experience	*3:*339		
extent of	*3:*160		
future perspectives and challenges related to	*3:*162		
impact on society	*3:*161		
modeling	*3:*161	*3:*339	
understanding and reducing	*3:*161		
Urbanization			
effects of	*3:*430		
influence on the hydrologic cycle	*4:*281		
pollution due to	*2:*181		
Urban landscape, water quality management in	*2:*189		
Urban land use, impact on groundwater quality	*5:*250		
Urban river sites case study	*2:*172		
Urban runoff	*2:*190	*2:*192	*3:*498
See also Urban stormwater runoff			
causes of	*5:*451		
groundwater contamination	*5:*452		
gully pot effluent treatment and	*3:*499		

Index terms	**Links**		
Urban runoff *(Continued)*			
indicator parameters of	*2:*192t		
managing	*2:*187		
nonpoint source pollution from	*2:*186		
as a source of cultural eutrophication	*3:*115		
into urban watercourses	*3:*499		
Urban services, efficiency of	*2:*491		
Urban stormwater runoff			
evaluating impacts of	*3:*433		
regulating	*3:*432		
water quality issues in	*3:*432		
Urban wastewater reuse	*1:*817		
Urban Wastewater Treatment Directive	*2:*45		
Urban water demand	*2:*490		
Urban water distribution systems, modeling chlorine residuals in	*1:*131		
Urban water resource management, in Asia	*2:*552		
Urban watershed protection approach	*2:*191		
Urban water studies	*3:*501		
Urban water supply, in developing countries	*2:*632		
Urinals, water-free	*1:*678		
Urine separation	*1:*843		
U.S. Agency for International Development	*2:*489		
U.S. Army Corps of Engineers. *See* Army Corps of Engineers			
U.S./Canadian Boundary Waters Treaty	*4:*620		
U.S. coastal waters, oil spills in	*2:*291t		
U.S. Constitution, water planning and	*4:*604		
U.S. Department of Agriculture (USDA)	*2:*186		
See also United States agricultural land use planning program	*3:*596		
U.S. Department of Agriculture Natural Resource Conservation Service (USDA NRCS), agricultural drainage ditch research	*3:*88		
U.S. Department of the Interior (DOI)	*2:*186		
U.S. Department of Transportation (DOT)	*2:*186		
USEPA codes	*5:*141		
See also Arsenic standard (USEPA); EPA Regional Ground Water Forum			
USEPA ‘‘priority elements,’’ effect on fish	*3:*131t		
U.S. EPA Subtitle D dry tomb landfilling	*2:*163		
USEPA Test and Evaluation (T&E) Facility	*2:*509		
U.S. Geological Survey (USGS)	*2:*311	*2:*312	*3:*349
codes	*5:*140		
contaminant study	*5:*60		

U.S. Geological Survey National Water-Quality Assessment (NAWQA) Program *3:*608

U.S. National Environmental Policy Act (NEPA) *2:*624

U.S. Public Health Service (USPHS) *1:*287
  1962 standards of *1:*288 *1:*292

U.S. Salinity Laboratory Classification *2:*159 *2:*161t *5:*208

USS Monitor, discovery of *4:*62

U.S. streams, organic wastewater contaminants in *5:*605

U.S. waterways, contamination of *2:*122

UV dosage *1:*469
  See also Ultraviolet (UV) radiation

UV light components *1:*467

U-V-P finite element scheme *5:*658

UV sources *1:*470

UV treatment system *1:*470

V

Vacuoles, sequestration of metals in *5:*287

Vadose models *5:*77

Vacuum breakers *1:*156

Vacuum compression *1:*171

Vacuum filters *1:*176

Vacuum flotation *3:*407

Vacuum gauge tensiometers *3:*724
  handling of *3:*728
  response time of *3:*725

Vacuum toilet systems *1:*679

Vadose zone *5:*105
  averaged transport through *5:*77
  capillary flow in *5:*73
  contaminant vapor transport through *5:*551
  flow processes in *5:*537
  hydrology *5:*533
  macropore flow in *5:*73
  materials *5:*535
  modeling non-point source pollutants in *5:*299
  processes in *5:*534
  soil samples from *3:*688

Vadose zone gas monitoring, direct *5:*542

Index terms **Links**

Vadose zone monitoring techniques *5:*538
  direct *5:*541
  indirect *5:*539
Vadose zone tracers, dyes as *3:*96
Vaiont Dam collapse *3:*147
Valley-floor gullies *3:*183
Valley-floor slope *3:*187
VALIMAR monitoring project *2:*31 *2:*32
Valtin, Heinz *4:*789
Valuation approaches, direct *4:*607
‘‘Value-added’’ ocean observations *4:*23
Value estimation methodologies *4:*607
‘‘Valvecard’’ software *1:*484
Valve-exercising program *1:*401
Valves *1:*482
  common problems with *1:*483
  control of *1:*483
  double check *1:*156
  high velocities through *1:*484
  insertion *1:*890
  operation and maintenance of *1:*484
  safety of *1:*484
  types of *1:*482
  water distribution system *1:*208
  water hammer and *1:*262 *1:*263*f* *1:*484
Van der Meer expressions *4:*16
van der Waals forces *1:*99 *4:*387 *4:*469
Vaporization, heat of *4:*263
Vapor migration, from groundwater devolatilization *5:*553
Vapor pressure *4:*271 *4:*366 *4:*522 *4:*528
  determining *4:*272
Vapors, leaking *5:*552
Vapor sampling *5:*545
Vapor transport, in the unsaturated zone *5:*543
Variability indexes, streamflow *3:*104
Variable uncertainty *3:*325
Variance technologies *1:*457
Vegetable crops, microirrigation of *3:*625

Index terms	**Links**	
Vegetables		
nitrate and nitrite levels in	*3:*639	
organochlorine pesticides in	*3:*643	
Vegetated riparian buffers	*3:*391	
Vegetated submerged bed (VSB) systems	*1:*787	
Vegetated systems, for storm water treatment	*1:*868	
Vegetation. *See also* Plant entries		
acid deposition and	*4:*378	
concentration of MtBE released through	*5:*394	
submerged	*3:*275	
use in water erosion control	*3:*551	
water stored in	*3:*475	
Vegetation index (VI)	*3:*722	
Vegetation Index/Temperature (VIT) trapezoid	*3:*721	
Vegetation survey	*5:*78	
Vegetative filter strips (VFS)	*3:*391	
Veliger settling, industrial field studies on	*1:*513	
Velocities, vector plots of	*5:*624	
Velocity approach, to incipient motion	*3:*418	
Velocity constraints, in water distribution systems	*1:*209	
Velocity meters	*1:*338	*1:*340
Vendor-supplied systems, for storm water treatment	*1:*868	
Venezuela, flooding in	*3:*142	
Ventilation, sewer system	*1:*912	
Vents, underwater	*4:*90	
Venturi meters	*1:*338	
Vertical auger mixing	*1:*837	
Vertical breakwaters	*4:*17	
Vertical drainage	*3:*732	*5:*98
Vertical-flow constructed wetlands	*1:*896	
Vertical flow pond	*2:*424	
Vertical flow wetlands, in acid mine treatment	*5:*6	
Vertical hydraulic conductivity analysis	*5:*186	
Vertical intake wells	*5:*88	
Vertical mixing	*3:*81	
Vertical roughing filters	*1:*237	
Vertical slot fish passage facility	*3:*530	
Vertical soil vapor	*5:*550	
Vertical turbine pumps	*1:*393	

Very low frequency induction *5:*149
Veterinary pharmaceuticals *1:*377
Vibrio *1:*278
Vibrio fischeri luminescent bacterium *2:*453
Vietnam, water treatment in *2:*553
Vietnam War era, ecoterrorism during *4:*653
Vinasses, composition of *1:*619t
Vinyl chloride (VC) *2:*266
  characteristics and production of *5:*634
  detection of *5:*637
  disposal of *5:*635
  exposure to *5:*635
  fate and transport of *5:*635
  in groundwater *5:*634
  health effects of *5:*636
  regulation of *5:*637
  remediation of *5:*638
Viral agents *1:*178
Viral disease *2:*371
Viral threat agents *1:*89
Virtual Library on Environmental Health (VLEH) *2:*671
  difficulties with *2:*673
Virtual temperature *4:*369
Virtual water *2:*536 *4:*787
  import and export of *2:*538 *2:*540t
  importation in the Nile Basin *2:*593
Virtual water balance *2:*537f
  national *2:*540
Virtual water content, calculating *2:*537
Virtual water flows
  calculating *2:*537
  global *2:*540
Viruses *1:*343 *1:*904 *3:*669
  in domestic sewage *1:*831
  health effects of *1:*279
  movement of *1:*71
  removal of *1:*249 *1:*485
  sources of *1:*71

Index terms	**Links**			
Virus transport				
modeling	*1:*72			
subsurface	*1:*70			
Viscoelastic fluids	*5:*557			
Viscosity	*5:*649			
Newton's law of	*5:*555			
Viscous floc	*1:*730			
Viscous flow	*5:*555			
Viscous-flow porometers	*3:*717			
Visibility degradation	*4:*378			
Vitamin D metabolism, effects of lead on	*2:*436			
Vitotox® assay	*2:*175	*2:*455		
Vitrification	*1:*836	*5:*434		
VLEACH code	*5:*564			
Void volume (α)	*1:*103			
Volatile aromatic compounds	*2:*310			
Volatile fatty acids (VFA)	*1:*701	*1:*702		
Volatile organic				
chemicals/compounds (VOCs)	*1:*105	*1:*353	*1:*760	*1:*761
	*2:*268	*2:*343	*4:*378	*5:*543
See also Odor abatement; Perchloroethylene (PCE)				
adsorption of	*1:*117			
detection of	*2:*283			
occurrence of	*5:*250			
options for treating	*1:*762f			
removal of	*3:*368			
soil vapor extraction of	*5:*121			
as water quality indicators	*2:*266			
Volatile organic contaminants, technologies for	*1:*464t			
Volatile sulphur compounds	*4:*88			
Volatilization	*5:*528			
groundwater remediation by	*5:*426			
of heavy metals	*5:*378			
of nitrogen	*3:*696			
of persistent organic pollutants	*1:*768			
Volcanic activity, as a source of acidity	*3:*7			
Volcanic emissions	*4:*159			
Volcanic islands, perched groundwater in	*5:*353			
Volcanic water, acidification of rivers and lakes by	*3:*10			
Volcanoes, underwater	*4:*90			

Vollenweider-OECD eutrophication model *3:*110 *3:*111 *3:*112
Voltage *4:*442
Voltage difference *4:*442
Volume Imaging Lidar *4:*296
Volumetric flow rate (*Q*) *1:*103
Volumetric reliability *3:*261
Voluntary metering *1:*494
Vortex technologies *1:*784
Vortex-type CSO solids separation unit *1:*784
VS2D model *5:*663
V-shaped basin, design discharge for *3:*245
Vulnerability assessment (VA) *1:*424 *2:*347
  groundwater *1:*527
  limitations of *5:*565
  verification and postaudits of *5:*566
Vulnerability Index (VI) *5:*597
Vulnerability mapping, of groundwater resources *5:*561
  See also Groundwater vulnerability

W

Walawe basin development *4:*779f
"Walking clocks" *4:*704
Walk-through, role in pipe assessment *1:*886
Warm cloud seeding *4:*187
Warm fronts, fog and *4:*231
Warning systems for flooding, flash floods, and landslides *1:*125
Washington State
  in-stream flow protection in *4:*661 *4:*662
  source-water protection in *2:*311 *2:*312
Wash load *3:*397
Waste, aquaculture *3:*541
Waste activated sludge (WAS) *1:*456 *1:*862
Waste deposits model *1:*726
Waste disposal, from arsenic treatment *5:*25
Waste disposal sites, dye tracing at *5:*110
Waste loads
  allocation of *3:*81
  transport mechanism for *3:*79
Wasteload allocations (WLA) *2:*188

Index terms	**Links**	
Waste management practices	*2:*374	
revising	*1:*900	
Wastes, from electricity generation	*5:*448	
Waste stabilization ponds (WSP)	*1:*518	
Waste stream disposal	*1:*336	
Waste treatment, in fish farms	*1:*681	
Wastewater(s)	*1:*656	
See also Complex waste waters; Wastewater management; Wastewater modeling and treatment plant design; Wastewater reclamation and reuse entries; Wastewater treatment entries		
aeration of	*1:*623	
analyses of	*1:*657t	*1:*658t
biological characteristics of	*1:*658	
characteristics of	*1:*677	*1:*900
chemical characteristics of	*1:*657	
chemically enhanced primary treatment of	*1:*659	
concentration of microorganisms in	*3:*670t	
contaminants of	*1:*829t	
defined	*2:*548	
direct reuse of	*1:*306	
disposal practices for	*3:*79	
EU monitoring of	*2:*267	
land applications of	*1:*632	
occurrence of detergents in	*1:*671	
from oil production	*2:*285	
physical characteristics of	*1:*657	
pollution by	*3:*292	
pollution in	*4:*507	
reclamation of	*2:*497	
as a resource	*1:*676	
sample analyses of	*1:*662	
toxic evaluation of	*1:*565	
Wastewater-associated pathogenic microorganisms	*3:*668t	
Wastewater bacteria, poly-P-containing	*1:*788t	
Wastewater biofilms	*2:*230	
Wastewater contaminants, in U.S. streams	*5:*605	
Wastewater discharge numeric limitations, industry sectors subject to	*1:*758t	
Wastewater discharge permits	*1:*755	
Wastewater discharges. *See also* Wastewater effluents		
from power plants	*1:*556	
regulatory controls on	*1:*557	

Wastewater effluents *3:*80
Wastewater management. *See also* Ecological wastewater management
  centralized and decentralized *1:*720
  for developing countries *1:*718
Wastewater modeling and treatment plant design *1:*730
  applications of *1:*738
  biological process modeling *1:*735
  information gathering step *1:*738
  input and output file preparation step *1:*744
  International Water Association models *1:*732 *1:*736
  laboratory and data organization step *1:*742
  model compilation and setup step *1:*744
  model parameter initialization step *1:*745
  model selection step *1:*739
  parameter adjustment step *1:*745
  reactor hydraulics step *1:*740
  static and dynamic models *1:*731 *1:*733
Wastewater quality, for developing countries *1:*719
Wastewater reclamation and reuse *1:*825
  treatment technology related to *1:*826
Wastewater reclamation and reuse research *1:*819
  future needs for *1:*823
  to protect public health *1:*819
  public perception and *1:*823
  in wastewater treatment technologies *1:*822
Wastewater reuse *3:*667
  See also Wastewater reclamation and reuse entries
  guidelines and treatment practices for *3:*670 *3:*671*t* *3:*672t
Wastewater reuse system, low-tech *5:*329
Wastewater sanitation, enhancement of *1:*498
Wastewater sludge *1:*861
  See also Sludge
Wastewater streams, pollutants in *4:*506
Wastewater toxicity, biotests for *1:*566
Wastewater treatment *1:*814
  See also Advanced wastewater treatment techniques; Small-scale wastewater treatment; Wastewater treatment plants (WWTPs); Wastewater treatment technologies
  anaerobic *1:*904
  biofilm in *2:*230
  centralized versus decentralized *1:*678

Wastewater treatment *(Continued)*
  disinfection *1:*817
  fate of persistent organic pollutants in *1:*766
  inadequate *1:*661
  powdered activated carbon in *2:*87
  preliminary *1:*814
  primary *1:*815
  regulation of *1:*844
  removal of microorganisms by *3:*669
  secondary *1:*815
  sugarcane industry *1:*614
  use of redox potentials in *1:*399
Wastewater treatment facilities, Ho Chi Minh City *2:*553
Wastewater treatment industry *4:*654
Wastewater treatment methods, oil-field brine control by *2:*288
Wastewater treatment plants (WWTPs) *1:*565 *1:*623 *1:*766 *1:*767
  monitoring biodiversity in *1:*642
  odor abatement in *1:*760
  performances of *1:*663 *1:*664f
Wastewater-treatment return flow *2:*548
Wastewater treatment technologies *1:*808
  biological *1:*813
  chemical *1:*811
  electrical *1:*812
  physical *1:*809
  thermal *1:*812
Wastewater use studies, site characteristics of *1:*633t
Wastewater utilities, security of *1:*870
WATEQ4F program *5:*140
WATEQ database *5:*144
Water. *See also* Aquatic entries; Drinking water; Heavy water; Natural waters; Raw water; Reclaimed water; Wastewater entries; Water conservation measures; Water distribution system entries; Water industry; Waterlogging; Water quality; Water resource entries; Water supply entries; Water treatment entries; Water use
  absolute (dynamic) viscosity of *4:*527
  algal toxins in *2:*387
  bacteria, viruses, and pathogens in *2:*303
  as a better medication *4:*724
  balancing the need for *2:*574
  beryllium in *4:*394
  between Arabs and Israelis *4:*699

Index terms	**Links**	
Water *(Continued)*		
biological structure and function of	*4:*512	
bottled	*1:*3	
cadmium in	*5:*617	
characteristics of	*1:*900	
chemistry of	*4:*711	
as a cholesterol-lowering medication	*4:*701	
clarity/turbidity of	*1:*599	*3:*108
classification of	*2:*95	*4:*449t
competing uses for	*2:*613	
complex dynamics of	*4:*512	
cosmic	*4:*189	
crystals in	*4:*712	
density of	*4:*527	
discovery of	*4:*711	
distilled	*4:*441	
distribution and quantity of	*4:*283	
economic value of	*4:*605	
effect on cell biology	*4:*456	*4:*458
electric generating plant contaminants of	*1:*555	
emergence as a public health issue	*1:*283	
emergency	*2:*526	
evolution of	*4:*715	
extracellular matrix and	*4:*459	
facts related to	*3:*394	
freezing and supercooling of	*4:*585	
future of	*4:*713	
globalization of	*2:*536	
gross value of	*4:*608	
heat capacity of	*4:*529	
in history	*4:*726	
in the history of photography	*4:*766	
impact on health and aging	*4:*455	*4:*729
impurities accumulated by	*1:*900	
in interstellar space	*4:*189	
as key to health and healing	*4:*722	
kinematic viscosity of	*4:*527	
‘‘magical’’ properties of	*4:*583	
man’s history and	*4:*745	
market allocation of	*2:*644	
marketing proposals for	*2:*506	

Index terms	**Links**	
Water *(Continued)*		
measurement of	*2:*658	
measurement of ^{36}Cl in	*4:*417	
microbial growth in	*3:*311	
modulus of elasticity of	*4:*529	
molecular network dynamics of	*4:*511	
monitoring for threat agents	*1:*89	
nitrate in	*1:*31	
nitrogen pollution of	*3:*605	
oil removal from	*1:*774	*1:*776f
organic matter contamination of	*3:*606	
origin of	*4:*283	
osmotic movement of	*4:*520	
oxidation of	*4:*535	
passage in and out of cells	*4:*458	
prescriptive use of	*4:*798	
pricing	*3:*585	
production of hydrogen ion in	*4:*481	
protozoa in	*2:*313	
radon in	*4:*541	
regulatory standards for	*1:*576	
relationship to soil	*3:*641	
as a renewable resource	*2:*661	
role in enzyme functioning	*4:*512	
role in the Canadian landscape	*3:*507	
solubility of chemicals in	*4:*555	
solubility of hydrocarbons and sulfur compounds in	*4:*561	
as a solvent	*4:*725	
sound in	*4:*569	
as a source of wisdom	*4:*793	
on the Space Station	*4:*572	
specific gravity of	*4:*527	
specific weight of	*4:*527	
speed of sound of	*4:*529	
squeezing from rock	*5:*477	
storing	*2:*527	
strontium isotopes in	*4:*574	
surface tension of	*4:*528	
survival of threat agents in	*1:*88	
technetium in	*4:*578	
trace elements in	*3:*455	

Index terms	**Links**	
Water *(Continued)*		
in the unsaturated zone	*5:*531	
varieties of	*4:*462	
vinyl chloride in	*5:*637	
Water accounting, as a water conservation measure	*1:*146	
Water Act of 1974	*2:*183	
Water adsorption		
on carbons	*4:*400	
studies of	*4:*400	
Water allocation	*2:*504	
See also Water distribution		
authority over	*4:*660	
efficient	*2:*643	
Water allocation law	*2:*558	
Water allocation system	*2:*644	
Water allotments, purchasing	*2:*508	
Water analysis	*1:*86	
corrosion measurement and	*1:*9	
tests for	*2:*178	*2:*180t
trace element pollution and	*4:*112	
Water analysis methods	*2:*304	
gas chromatography	*2:*304	
Water assessment and criteria	*2:*24	
biotic integrity and	*2:*26	
current	*2:*25	
development of	*2:*24	
institutional arrangements for	*3:*34	
reference sites and biocriteria	*2:*26	
Water audits	*1:*401	*2:*492
Water availability, impact of biofuels on	*3:*547	
Water availability index	*1:*404f	
Water balance	*1:*202	
global	*4:*279	
in Hawaii	*4:*259	
Water balance equation	*3:*259	
Water balance method/techniques	*3:*572	*5:*163
Water balance/modeling, recharge measurement	*5:*74	
Water Bank Act	*4:*691	
Water-bearing rocks, consolidated	*5:*55	

Index terms **Links**

Water biosecurity *1:*87
  future directions for *1:*91
Waterbodies. *See also* Open waterbodies
  advancements in understanding *4:*730
  anoxia and hypoxia in *4:*64
  effluent-based *2:*475
  excessive fertilization of *3*:107 *3*:108
  "impaired" *2:*476
  impact of cadmium on *5:*615
  remote sensing of *4*:324
Waterborne bacteria *2:*20
  identification of *2*:21
Waterborne disease outbreaks *2:*293
Waterborne disease surveillance *1:*183
  background of *1:*184
  early outbreak detection *1:*186
  endemic disease and *1:*187
  improving *1:*190
  of infection versus disease *1:*189
  limitations of *1*:185
Waterborne infectious diseases. *See* Emerging waterborne infectious diseases
Waterborne pathogens
  health and *1*:521
  persistence of *1*:521
Waterborne radon *1*:51
Waterborne transmission
  of *Escherichia coli* *1*:429
  of *Escherichia coli* O157:H7 *2*:138
Water budget *2:*507 *3:*222
  lake *3:*266
  program *3:*554
Water "carriage" *1:*510
Water chemistry
  factors determining *2*:19
  in regional flow systems *5:*419
Water cleanup, organoclay in *1:*771
Water clocks *4:*703
Water color, caused by iron bacteria *2*:151
Water column
  bacterial activity in *3*:181
  heat flux into *3*:192

Water companies, financial performance of *1:*310
Water conditioners, magnetic *1:*144
Water conditioning
  magnetic *1:*534
  physical *1:*141
Water conflicts. *See also* Water disputes
  in the Nile Basin *2:*590
  over scarce water *2:*502
Water consciousness, evolution of *4:*707
Water conservation, during a drought *4:*209
Water conservation ecosystems, ancient Sri Lankan *4:*772
Water conservation interventions, nonengineering *2:*472
Water conservation measures *1:*53 *1:*146 *1:*510 *2:*489 *2:*495
  See also Landscape water-conservation techniques; Water saving technologies; Water use
  in Canada *2:*660
  costing and pricing *1:*147
  efficiency and *2:*492
  in the home, community, and workplace *2:*664
  information and education *1:*148
  in integrated water resources management *2:*575
  metering *1:*146
  water accounting and leak repair *1:*146
Water Conservation Plan Guidelines (EPA) *1:*146
Water-conserving toilets *1:*679
Water consumption *2:*370
  metering coverage and *1:*494f
  reducing *1:*678 *2:*664
  rise in *2:*375
Water contamination
  by heavy metals *5:*275
  by low level organic waste compounds *5:*60
  sources of *3:*651
Water content
  measuring *5:*461
  predicting *5:*126
Watercraft discharges *3:*281
Water crisis, global *4:*769

Index terms **Links**

Water cycle *2:*548 *4:*191 *4:*196
  See also Global water cycle; Hydrologic cycle
  climate change and *4:*193
  conservation and *4:*433
  deuterium in *4:*439
  NASA Aqua mission and *4:*193
  sun-powered *4:*584
Water debts *2:*617
Water Deficit Index (WDI) *3:*719 *3:*721
Water demand
  derived *4:*608
  estimation of *1:*496
  management of *1:*146
  price elasticity of *4:*606
  residential *1:*12
  urban *2:*490
Water demand forecasting *2:*529
  cases of *2:*530
  model *2:*622
Water dependency, global *2:*538 *2:*540t
Water discharge permits *5:*3
Water disinfection, using UV radiation *1:*471
Water disputes. *See also* Water conflicts
  averting *2:*501
  constitutional solutions to *2:*503
  water allocation and *2:*504
Water distribution, fair *2:*491
  See also Water allocation entries
Water distribution storage tanks, hydraulic design of *1:*448
Water distribution system design *1:*207
  constraints on *1:*209
  iterative method for *1:*210
Water distribution systems (WDSs) chlorine residual modeling within *1:*132
  components of *1:*208
  corrosion control in *1:*152
  desirable operating conditions for *1:*201
  elements of *1:*201
  maintenance and repair functions of *1:*203
  operation, maintenance, and design of *1:*200
  prevention and safety functions of *1:*203
  principal operations of *1:*201

Index terms	**Links**	
Water distribution systems (WDSs) chlorine residual modeling within *(Continued)*		
public relations and management functions of	*1:*203	
pumps in	*1:*391	
risk reduction for	*1:*436	
securing	*1:*434	
urban	*1:*131	
water delivery and monitoring by	*1:*202	
Water distribution system simulators (DSS)	*2:*509	
Water dowsing (witching)	*5:*92	
Water dwellers, exposure method using	*2:*420	
Water economic analysis, components of	*2:*128	
Water education		
impacts on	*4:*714	
strategies for	*4:*713	
Water efficiency, technological options for	*2:*493	
Water efficient residential technology	*2:*663	
"Water emergency price schedule"	*2:*583	
Water engineering technology, development of	*4:*728	
Water environment, persistence of toxic metals in	*2:*233	
Water Environment Federation (WEF)	*1:*863	
Water erosion	*1:*537	*3:*566
control of	*3:*551	*3:*568
Water Erosion Prediction Project (WEPP)	*2:*251	
Water facilities, capital investment in	*2:*594	
Water filtration	*1:*230	
Water flow, in unconfined aquifers	*5:*662	
Water holding properties, field capacity and	*5:*125	
Water flow patterns, using dyes to delineate	*3:*96	
Water flow rates, biofilm growth and	*1:*598	
Water-flow satellite data, from NASA	*2:*587	
Water-flush toilets	*1:*678	
Waterfowl, lead toxicity to	*2:*434	
Water-free toilets	*1:*678	
Water Framework Directive (EU)	*2:*45	*3:*292
Water/graphite model potentials	*4:*401	
Water habitat, eutrophication in	*3:*108	
Water hammer	*1:*261	*1:*401
causes and effects of	*1:*261	*1:*891
computation of	*1:*891	

Index terms	**Links**
Water hammer *(Continued)*	
control of	*1:*892
solutions to	*1:*261
valves and	*1:*484
Water hardness	*4:*413
Water harvesting	
choosing techniques for	*2:*551
components of	*2:*549
defined	*2:*549
forms of	*2:*550
need for	*2:*548
and reuse	*3:*557
structures related to	*2:*550
Water heaters, solar	*1:*63
Water hyacinth	
control of	*3:*481
as a problematic weed	*3:*479
Water impacts, from construction sites	*1:*537
Water industry	*2:*587
companies with interests in	*1:*302t
consolidation of	*1:*509
future of	*1:*307
key trends in	*1:*508
market drivers of	*1:*304
market size and growth characteristics of	*1:*304
policies and approaches of	*1:*510
privatization of	*1:*306
state of	*1:*301
trends in	*1:*305
Water infrastructure	*2:*567
planning and managing	*2:*594
Water Infrastructure Network	*1:*303
Watering equipment, outdoor	*2:*666
Water-jetting drilling technologies	*5:*234
Water Law of China	*2:*486
Water laws/legislation	*2:*96
Hawaiian	*4:*805
Islamic	*4:*634
state	*2:*584
water transfers and	*4:*686

Index terms	**Links**	
Water-level drawdown	*5:*102	
in aquifers	*5:*103	
information related to	*5:*104	
Water-level oscillation	*5:*112	
Water levels, in wells	*5:*572	
Water-limited environments, ecological effects assessment for	*2:*516	
Water loads, river	*3:*393t	
Waterlogged areas, types of	*1:*404	
Waterlogged lands, management and reclamation of	*3:*603	
Waterlogging	*3:*599	*3:*741
causes of	*3:*599	
effects of	*3:*600	
field measurement assessments for	*3:*599	
impacts of	*3:*741	
losses due to	*3:*601t	
problem identification observations related to	*3:*599	
remote sensing and GIS application in	*2:*534	
surface stagnation and crop yields from	*3:*600	
Water loss		
controlling	*1:*317	
tracing	*1:*401	
Water mains	*1:*203	
Water management	*4:*433	
in the Arab World	*2:*472	
British and French models of	*1:*389	
in China	*2:*484	
computer models for	*2:*620	
institutional structure of	*2:*614	
irrigation and	*4:*771	
lessons related to	*4:*268	
in the Mediterranean islands	*2:*642	
in the public trust	*2:*608	
Soviet style	*3:*19	
Water management agencies, federal	*4:*650t	
Water-management plans, drought strategy in	*2:*581	
Water management problems, severity of	*4:*289	
Water markets	*2:*499	
See also International water market; Spot markets		
Chilean	*2:*645	
in India	*2:*555	
institutional options related to	*2:*500	

Water markets *(Continued)*
  reform of *2:*508
  regulation of *2:*500
  transaction costs of *2:*499
Water meters *1:*337 *1:*489
  classes of *1:*492
  compound *1:*338
  densities, costs, and impacts of *1:*494
  history of *1:*489
  household *1:*490
  installation of *1:*338
  operating principles of *1:*493f
  positive displacement *1:*338
  precision tests for *1:*491
  reading *1:*339
  selecting and sizing *1:*340 *1:*492
  technical aspects of *1:*493
  testing and maintenance of *1:*340
  as tools for management and charging *1:*490
  types of *1:*338
  velocity *1:*338
Water molecules *4:*527
  stable isotopes in *5:*229
Water music instrument *4:*758
Water needs
  for freshwater fisheries management *3:*133
  tradable rights and *2:*643
Water networks, behavior of *4:*511
Water oxidation, supercritical *1:*874
Water penetration, soil texture and *3:*707
Water phase diagram *4:*344f
Water philosophy, evolution of *4:*708
Water pipeline. *See* Pipeline entries; Underground pipeline
Water planning, U.S. Constitution and *4:*604
Water policies, supply-oriented *2:*574
Water pollution *1:*901 *2:*95
  See also Polluted waters; Surface water pollution
  causes of *4:*519
  from fish farms *3:*579
  health risks of *2:*617
  reducing *4:*647

Index terms **Links**

Water pollution *(Continued)*
  tests used for measuring *2:*180t
  types and forms of *4:*97
Water Pollution Control Act of 1948 *4:*652
  history of *4:*651
  before the mid-1960s *4:*652
Water pollution control plants (WPCPs) *1:*747 *1:*748 *1:*750
  New York City *1:*746 *1:*748t
Water Portal of the Americas *2:*669
Water portals *2:*668
  types of *2:*669
Water Poverty Index *1:*439
Water pressure, Space Station *4:*573
Water pressure reduction *1:*157
Water prices, rise in *1:*508
Water pricing *1:*215 *1:*216 *2:*497 *2:*603
  See also Cost(s); Option prices
  assessment of *2:*605
  in Canada *2:*663
  components of *2:*603
  functions of *2:*604
  global *2:*603
  increase in *2:*604
Water problems
  awareness of *4:*606
  current *2:*574
Water productivity (WP) *3:*558
  improving *3:*585
Water project operations *2:*621
Water projects, federal *2:*522
Water protection laws *4:*597
Water provision, private sector participation models for *1:*50
Water purification *4:*441
  adsorption capacity of activated carbon for *4:*381
  emergency *2:*527
  photocatalytic membrane reactors in *1:*791
Water purification machines, Space Station *4:*573
Water quality *2:*96 *2:*301 *2:*314 *5:*602
  See also Sewer water quality; Storage water quality; Water quality management; Water quality modeling; Water supply quality

Index terms **Links**

Water quality *(Continued)*
  aquifer *5:*9
  ash pond and FGD sludge pond *1:*850
  biofouling effects on *5:*35
  in Canada *2:*663
  chemical parameters *1:*902
  chronology of events related to *2:*314
  complaints concerning *1:*152t
  of constructed wetlands *1:*892
  costs and benefits of *2:*129
  criteria for *2:*315
  data on *2:*192
  defined *2:*302 *2:*548
  in distribution systems *1:*204
  in dredged sediment management *2:*122
  early pollutant detection and *2:*440
  economics of *2:*127
  effect of human activities on *2:*303
  effect of natural processes on *2:*302
  effect of scale on *2:*190
  electronic nose applications related to *2:*282 *2:*283
  ethic for *2:*315
  evaluating *2:*368
  federal regulations regarding *2:*190
  in forested land *2:*199
  Ganga Basin *3:*235
  guidelines for *2:*337
  impacts of eutrophication on *3:*107
  impacts on *1:*496
  impaired *2:*663 *3:*111
  incentive-based mechanisms for improving *2:*132
  in India *2:*566
  information about *2:*303
  issues related to *2:*181
  land use effects on *2:*169
  microbial criteria for controlling *2:*449
  in municipal solid waste landfills *2:*163
  in net pen aquaculture systems *3:*543
  nitrate levels and *1:*32
  nitrogen loading and *3:*696
  nutrient-related *3:*111

Water quality *(Continued)*
  oil-field brine and *2:*287
  parameters of *1:*901 *2:*180t
  in ponds *3:*484
  problems with *1:*304
  protection measures for *2:*120
  public awareness of *2:*194
  regulation of *4:*671
  sand abstraction and *3:*415
  soil pipe erosion and *5:*403
  in suburban watersheds *3:*441
  in urban stormwater runoff *3:*432
  uncertainties in analyzing *2:*369
  in urban areas *5:*250
  using luminescent bacteria and *lux* genes for determining *2:*172
  variations in *2:*179
  in watersheds *3:*460
  wetlands and *3:*491
Water Quality Act of 1965 *2:*194 *2:*314
Water Quality Act of 1987 *4:*654
Water Quality Analysis Simulation Program (WASP5) *2:*327 *2:*329
Water quality assessment, classification and mapping of agricultural land for *3:*608
Water Quality Assessment program (USGS) *5:*386
Water Quality Association (WQA) *1:*381
Water quality-based effluent limitations (WQBELs) *1:*557 *4:*657
Water quality constraints, in water distribution systems *1:*209
Water quality criteria
  Great Lakes system *4:*623
  for toilet flushing *1:*17
Water Quality Criteria and Standards Plan (EPA) *3:*436
Water quality evaluation, incorporating chemical information into *4:*602
Water quality goals, for estuaries *4:*54
Water quality impairment, urban-landscape-related *2:*191t
Water Quality Index (EPA) *2:*368
Water quality indicators *2:*266
Water quality indices *2:*368
Water quality issues, major *4:*286t
Water quality “ladder” *2:*655

Index terms **Links**

Water quality management *2*:176 *2*:188 *5*:334
  in an agricultural landscape *3*:604
  approach to *2*:183
  determinants related to *2*:178
  in a forested landscape *2*:199
  Great Lakes region *3*:179
  legal considerations in *2*:183
  nonpoint source control and *2*:184
  in the United States *2*:193
  in an urban landscape *2*:189
Water quality modeling *3*:338
  See also Water quality models
  case studies in *2*:255
  of Lake Tenkiller, Oklahoma *2*:260
  software tools for *2*:325
  future of *3*:341
Water quality models *2*:248
  See also Water quality modeling; Water quality simulations
  Agricultural Nonpoint Source Pollution Model 2001 *2*:251
  chemical principles related to *2*:269
  CREAMS and GLEAMS *2*:250
  detailed *2*:249
  equations for *2*:273
  Erosion-Productivity Impact Calculator *2*:251
  Generalized Watershed Loading Function *2*:252
  hydrologic simulation program *2*:249
  kinetic parameters used in *2*:276t
  mathematical framework of *2*:273
  midrange *2*:248
  Nitrate Leaching and Economic Analysis Package *2*:252
  Pesticide Root Zone Model *2*:252
  regression methods and spreadsheet models *2*:253
  review of *2*:249
  Revised Universal Soil Loss Equation *2*:251
  Root Zone Water Quality Model *2*:252
  selecting *2*:255
  simple *2*:248
  Soil and Water Assessment Tool *2*:250
  Source Loading and Management Model *2*:253
  variables in *2*:270
  Water Erosion Prediction Project *2*:251

Index terms **Links**

Water quality models *(Continued)*
  Watershed Modeling System *2:*250
Water quality monitoring *2:*177
Water quality parameters *4:*325t
Water quality policy *2:*127
  economic instruments for informing *2:*129
  instruments of *2:*132
Water quality programs *2:*315
Water quality protection, integration of *4:*672
Water Quality Protection Center *1:*91
Water quality regulations *2:*314 *4:*665
Water quality simulations *2:*261
Water quality standards (WQS) *2:*267 *2:*302 *4:*657 *5:*324
  Clean Water Act *4:*596
  violations of *3:*107
Water quality statutes *4:*664
Water quality studies, remote sensing and GIS application in *2:*534
Water quality trends, in New York City *1:*748
Water quality variables, notations and source terms for *2:*275t
Water quantity
  allocation of *4:*671
  impacts on *1:*495
Water quantity constraints, in water distribution systems *1:*209
Water quantity regulation, in India *2:*557
Water recovery/reclamation *1:*53
  See also Water reuse
  achievable treatment levels for *1:*807t
  technologies for *1:*806
Water recycling
  defined *2:*610
  environmental benefits of *2:*610
  future of *2:*613
  increased focus on *1:*306
  pollution reduction and *2:*612
  for sensitive ecosystems *2:*612
  technologies for *1:*808
  for wetlands and riparian habitats *2:*612
Water reduction methods, in Arab states *2:*472
Water regulation, history of *2:*193

Index terms **Links**

Water-related diseases *1:*23 *2:*111
  global morbidity and mortality from *1:*24
Water requirements
  average *1:*506
  for human health *1:*22
Water resource development, American *2:*498
Water resource development projects, Ganga Basin *3:*234
Water resource health, measuring *3:*37
Water resource integrity, evaluation using biotic integrity index *3:*36
Water resource management *2:*586 *4:*267
  in America *2:*608
  Chinese laws related to *2:*488
  decentralization of *2:*473
  decision support system for *2:*622
  for drought conditions *2:*576
  Great Lakes area *3:*179
  structures for *2:*567t
Water resource organizations *4:*648
  governmental *4:*649
  nongovernmental *4:*651
  professional *4:*651
Water resource remote sensing *4:*321
Water resource requirements, for India *2:*564
Water resources *2:*370 *4:*288
  See also Poor- quality water resources
  anthropogenic threats to *3:*37
  in the Arab World *2:*470
  best management practices for *2:*570
  Canadian *2:*656
  depletion of *4:*289
  economics of allocating *1:*215
  Geographic Information System and remote sensing in *2:*531
  health indicators for *3:*37
  in India *2:*559
  intergenerational *4:*771
  management of *1:*303
  protection of *2:*374
  satellites useful for *4:*326t
  stress on *2:*495
  sustainable development of *5:*163
  threats to *1:*438

Index terms **Links**

Water Resources Development Act of 1986 *4:*619

Water resources planning,
  integrating environmental
  impacts into *2:*520

Water resources systems *2:*586
  components of *2:*619
  performance evaluation of *2:*674

Water resources systems analysis *2:*683
  defined *2:*684
  mathematical modeling in *2:*686

Water resource sustainability *2:*624
  concepts related to *2:*625
  practices in *2:*628
  stages of *2:*629f

Water resource valuation *2:*653
  economic perspective on *2:*653
  empirical estimates of *2:*654

Water retention. *See also* Retention
  of fish ponds *1:*407
  maintaining and increasing *1:*406
  in river valleys *1:*406

Water retention characteristic *4:*486

Water reuse *1:*53 *1:*817
  See also Water recovery/reclamation
  in agriculture *2:*496
  increase in *1:*509
  in industry *2:*496
  international standards for *3:*670
  technologies applicable to *1:*54

Water rights *2:*500 *4:*769
  declarations of *4:*770
  distribution of *2:*643
  Hawaiian *4:*805
  issues related to *4:*770
  reserved *4:*689
  security of *2:*644

Water rituals *4:*709
  ancient *4:*709

Water-rock interaction *5:*566
  adsorption and *5:*566
  dissolution and *5:*568

Water-rock interaction *(Continued)*
  precipitation and *5:*569
Water safety, precautionary principle regarding *2:*600
Water sampling *2:*161
  See also Water samples
  security and integrity of *2:*162
Water samples. *See also* Water sampling
  analysis of strontium isotopes in *4:*575
  analyzing for deuterium *4:*439
Water savings, using microirrigation *3:*617
Water-saving techniques, domestic *2:*497
Water-saving technologies, in the Nile Basin *2:*594
Water scarcity *1:*304 *2:*181 *2:*490
  conflicts over *2:*502
  global *2:*538 *2:*540t
  human behavior and *4:*733
  in India *3:*445 *3:*448f
  in the Nile Basin *2:*590
  role in human health *1:*20
Water science. *See also* Water studies issues and risks related to *2:*602t
  precautionary principle regarding *2:*595
  terms related to *2:*541
Water security *1:*437
  See also Water resources
  global water availability *1:*437
  government actions regarding *1:*434
  relation to food security and agriculture *1:*438
  terrorism and *1:*158
Water self-sufficiency, global *2:*538 *2:*540t
Water service(s)
  data and indicators related to *1:*505
  valuation of *4:*607
Watershed(s) *2:*548 *3:*460
  See also Combustible watersheds; Megawatersheds; Suburban watersheds
  allochthonous inputs from *3:*207
  assessments of *2:*264
  changes in *4:*280
  flood potential of *3:*460
  flood source mapping in *3:*155
  municipal *1:*495
  processes related to *3:*327

Index terms	**Links**		
Watershed(s) *(Continued)*			
in stream classification	*3:*66		
time of concentration and travel time in	*3:*469		
water control in	*2:*567		
Watershed approach	*2:*189	*4:*115	
Watershed ecology, hydrochemical models used in	*5:*297t		
Watershed functions	*3:*473		
Watershed hydrogeomorphologic index	*3:*63		
Watershed hydrology	*3:*472		
modeling of	*3:*342		
role of wetlands in	*3:*172		
terminology related to	*3:*472		
units used in	*3:*473		
Watershed lands, conversion to agriculture	*3:*36		
Watershed management	*2:*192	*2:*248	*3:*502
plan for	*2:*629	*5:*335	
Watershed mapping/monitoring, remote sensing and GIS application in	*2:*533		
Watershed modeling	*3:*327		
See also Watershed models			
procedures in	*3:*328		
uncertainty analysis in	*3:*325		
Watershed Modeling System (WMS)	*2:*250		
Watershed models. *See also* Watershed modeling			
calibration of	*3:*329f		
classification of	*3:*328		
Watershed nutrient export coefficients	*3:*110t		
Watershed outlet, travel time to	*3:*63		
Watershed response			
continuous	*3:*343		
event-based	*3:*343		
Watershed restoration plans	*4:*596		
Watershed spatial scale	*2:*613		
Watershed water balance, forest quality and	*3:*172		
Water shortage, in the human body	*4:*789		
Water shortage planning tools, Palmer Index	*2:*581		
Water-shortage plans, state	*2:*583		
Water softening	*4:*454		
See also Magnetic water conditioning			

Index terms	**Links**
Water softening *(Continued)*	
for scale control	*1:*548
Water, Soil and Hydro-Environmental Decision Support System (WATERSHEDSS)	*2:*328
Water solubility, correlations for	*4:*382t
Water solubility graphs	*4:*556 *4:*558f *4:*559 *4:*560 *4:*561f *4:*562 *4:*564f
Water source emergencies, planning for	*2:*401
Water sources	
groundwater supplies variability in	*2:*400
little control of	*2:*400
Waterspout	*4:*347
Water spreading	*3:*163 *3:*708
methods of	*3:*164
Water stocks, improving	*2:*635
Water storage, technological revolution in	*4:*729
Water stress	
crop adaptations to	*3:*719
in India	*3:*445 *3:*448f
Water-stressed countries, health in	*1:*20
Water studies, interdisciplinary	*2:*641
Water subsidies	*1:*217
Water supplies. *See also* Water supply improvement	
for agricultural production	*2:*617
bacteriological quality of	*1:*294
chemical characteristics of	*1:*295
in developing countries	*2:*616
effect of sediment on	*3:*508
for energy production	*2:*618
enhancement of	*1:*497
failures in	*2:*514
indicators of	*1:*505
meeting demands for	*1:*218
physical characteristics of	*1:*295
political and legal impacts on	*2:*618
privatization of	*4:*770
radioactivity of	*1:*296
source and protection of	*1:*294
state and regional	*2:*613
treatment of	*2:*374
types of	*2:*371
uncertainty of	*2:*400

Water supplies *(Continued)*
  value of *1:*500
Water supply allocation *2:*569
Water supply improvement
  impact of *1:*25
  metering and *1:*490
Water supply planning/management *2:*515 *2:*614
  federal *4:*612
Water supply protection, extraterritorial land use control for *1:*315
Water supply quality *2:*370
  factors affecting *2:*373
Water supply subsystems *2:*515
Water-supply threat agents *1:*88
Water supply-wastewater sanitation cycle *1:*651f
Water supply wells *5:*22
  treatment of arsenic in *5:*22 *5:*23
  treatment of biofouling in *5:*22 *5:*27
  treatment of chromium in *5:*22 *5:*25
Water surface, long-wave radiation from *3:*192
Water Survey of Canada *2:*658
Water symbolism *4:*785
  in art and nature *4:*787
  postmodern metaphors and *4:*787
Water system operation, decisions about *2:*595
Water systems
  components and evolution of *1:*501
  iron and manganese removal from *1:*312
  pharmaceuticals in *1:*372
  rural *1:*213
Water table *2:*548 *5:*600
  See also Falling water table
  constant *3:*602
  contribution to crop evapotranspiration *3:*570
  data *5:*164
  “perched” *5:*11
  studies *3:*601
Water table levels, tile drainage and *3:*731
WaterTechOnline portal *2:*669
Water temperature tracer technique *5:*681
Water testing, frequency of *5:*325

Water towers *1:*410
Water tracing *5:*240
Water transfers *4:*685
  disputes over *4:*687
  impacts of *4:*686
  motives and issues in *4:*685
  return flow and *4:*685
  riparian doctrine and *4:*686
Water transportation, inefficiency in *2:*492
Water transportation occupations *4:*762
  employment in *4:*764
  job outlook and earnings in *4:*765
  nature of *4:*762
  training, qualifications, and advancement in *4:*764
  working conditions in *4:*763
Water travel times, in river-connected aquifers *5:*681
Water treatment. *See also* Drinking water treatment; Package plants
  beginning of *1:*286
  for cooling water systems *1:*560
  filtration *1:*245
  microbiological mechanisms and *1:*596
  ozone in *1:*362
  powdered activated carbon for *2:*86
  rationale for *1:*599
  in recirculating aquaculture systems *3:*544
  regional *2:*615
  in spent nuclear fuel storage *1:*595
Water treatment methods *1:*601
  See also Water treatment processes
  advantages and efficacy of *1:*194t
Water treatment plant residuals
  aluminum recovery from *1:*140f
  categories of *1:*411
  coagulant recovery from *1:*139
  management of *1:*411
  regulations governing *1:*413
Water treatment processes *1:*412t
  See also Water treatment methods
  selection guide for *1:*439

Index terms **Links**

Water treatment systems
  mixing and agitation in *1:*76
  stages in *2:*616
Water treatment technology, selecting *1:*242
Water turbine *3:*487
  specific speed of *3:*489
Water turnover *4:*276
Water use in California *2:*478
  categories of *2:*651
  defined *2:*548
  economic efficiency in *2:*575
  efficient *1:*307 *1:*509
  for energy industry materials *1:*562
  in energy production *1:*560
  excessive *2:*492
  in flow-through aquaculture systems *3:*542
  forecasting *1:*15
  global and regional *2:*489
  Great Lakes area *3:*178
  hydroelectricity *1:*561
  impact of biofuels on *3:*547
  indirect *1:*563
  in the Middle East *4:*753
  national trends in *2:*650
  nuclear power *1:*561
  in recirculating aquaculture systems *3:*544
  regulatory mechanisms regarding *2:*490
  thermoelectric *1:*560
  in the United States *2:*645
Water use categories, regional *2:*615t
Water use curtailment, during droughts *2:*583
Water use efficiency (WUE), agricultural *3:*558
Water use management, in the United States *2:*647
Water-use-priority system *2:*583
Water-user compliance, monitoring during droughts *2:*583
Water users, categories of *2:*587
Water-use sectors
  main *2:*495
  water demand forecasting for *2:*530

Index terms **Links**

Water utilities
  charging practices for *1:*490
  initial public offerings (IPOs) of *1:*389
Water valuation
  application of *4:*610
  estimation methodologies for *4:*607
  as an intermediate good *4:*608
  motivation for *4:*605
  studies *2:*655
Water valuation approaches
  direct *4:*607
  indirect *4:*608
Water vapor *4:*243 *4:*244 *4:*362
  changes in *4:*174
  condensing *4:*240
  distribution of *4:*270
  effects of *4:*170
  importance of *4:*717
Water vectored disease *2:*112
"Water wars" *2:*649
Water washed diseases *2:*112
Water waves *4:*138
  See also Waves
  characteristics of *4:*138
Waterways
  gross pollution of *2:*314
  South Asian *4:*681
Water well drilling techniques *5:*105
  cable-tool drilling *5:*105
  direct push drilling *5:*106
  hollow-stem auger drilling *5:*106
  horizontal drilling *5:*106
  rotary drilling *5:*105
  wire line coring *5:*105
Water wells. *See also* Well entries
  biofouling in *5:*35
  deterioration of *5:*23
  earthquake responses in *5:*111
  for groundwater sampling *5:*455
  installation of *5:*106

Index terms	**Links**	
Water withdrawal(s)	*2:*478	*3:*582t
See also Withdrawal		
in Canada	*2:*661	
by country	*2:*648t	
global	*2:*538	*2:*540t
Great Lakes	*3:*179t	
for public supplies	*2:*651	
in the United States	*2:*645	*3:*588
worldwide	*3:*381	
Water year	*4:*318	
Wave breaking	*4:*137	
Wave celerity	*3:*258	
Wave energy	*4:*45	
Wave height, estimating	*4:*2	
Wave overtopping	*4:*16	*4:*19
Wave-powered desalination system	*4:*46	
Wave reflection	*4:*17	
shoaling and	*4:*136	
Wave theory	*4:*45	
Waves		
in coastal waters	*4:*23	
continuous	*3:*248	
in lakes	*3:*267	
shallow water	*4:*135	
Wave theory	*3:*246	
Wave transmission	*4:*17	*4:*19
Weak add respiratory uncoupler (WARU)	*4:*524	
Weapon tests, as water pollution	*4:*102	
Weather		
atmosphere and	*4:*360	
energy and	*4:*353	
extremes in	*4:*175	
local and global	*4:*361	
Weather balloons	*4:*164	*4:*328
Weather Bureau, United States	*4:*347	
Weather forecasting		
advances in	*4:*354	
Aqua spacecraft and	*4:*351	
benefits to society	*4:*351	
foundations of	*4:*352	

Index terms **Links**

Weather forecasting *(Continued)*
  history of *4:*348
  modern tools of *4:*350
  numerical prediction in *4:*350
Weather modification
  health concerns related to *4:*188
  uses of *4:*188
Weather patterns, Great Lakes region *3:*177
Weather prediction accuracy, effect of GPS-IPW data on *4:*248
Weather radar, for hydrologic real-time information *1:*126
  See also Weather Surveillance Radar 1988 Doppler (WSR-88D)
Weather Surveillance Radar 1988 Doppler (WSR-88D)
  observations *4:*310 *4:*311
  radar stations *4:*307
Weather systems *4:*352
  understanding *4:*353
Wedderburn Number *3:*266
Weed control strategies *3:*742
  alternative *3:*747
  biological *3:*746
  chemical *3:*745
  environmental *3:*747
  factors affecting *3:*744
  mechanical *3:*744
Weeds, problematic *3:*479
Weighting factor *3:*258
Weight of evidence (WOE) approach *2*:381 *2*:429 *4:*600 *4:*602
  to sediment quality characterization *2*:350
Weir and orifice fish passage facility *3:*529
Weirs, flows through *3:*195
‘‘Weir’’ type roof drainage *1:*56
Welland Canal *3:*177
Well-being, water and *4:*791
Well casing perforated in place *5:*574
Well contamination, in developing countries *2:*631
Well contamination prevention *1:*149
  casing *1:*149
  disinfection procedures *1:*151
  filter pack *1:*151
  pitless adaptors *1:*151

Well contamination prevention *(Continued)*
  site selection *1:*149
  well design for *1:*149
  well screen *1:*151
  well seals *1:*151
Well drawdown *5:*101
Wellhead protection (WHP) *1:*524
  See also WHP zones (WHPZs)
  delineation criteria and methods for *1:*525
  management measures for *1:*527
Wellhead protection area (WHPA) *1:*524
Wellhead protection plan (WHPP) *1:*524
  developing *1:*525
Well hydraulics *5:*101 *5:*182
Well installation, water-jetting drilling technologies for *5:*234
Well jetting *5:*234
Well logs, studying *5:*81
Well maintenance *5:*263
  data collection intervals related to *5:*265t
  methods *5:*264t
  predictive and preventive *5:*263
  records and *5:*264
Wells. *See also* Intake wells; Irrigation wells; Water supply wells; Water well entries
  defined *2:*548
  design and construction of *5:*87
  horizontal *5:*177
  installation of *5:*106
  monitoring *5:*586
  performance characteristics of *5:*87
  radial *5:*407
  site selection for *5:*88
  types of *5:*571
  water levels in *5:*572
Well screens *3:*415 *5:*88 *5:*572
  lateral *5:*407
Well site control zone (WSCZ) *1:*526
Well-spacing norms, in India *2:*558
Well test *5:*574
Well water
  analysis of *5:*407
  blue baby syndrome and *2:*219

Index terms	**Links**	
Well yield	*5:*101	
Welwitschia mirabilis	*4:*200	
Wenner array	*5:*445	
West Antarctic ice sheet	*4:*118	
Western blotting	*2:*109	
Western United States		
irrigation in	*3:*587	
protection of in-stream values	*4:*660	
water markets in	*2:*500	
water rights adjudications and permit proceedings in	*4:*672	
water use in	*2:*646	
Wet ash discharges	*1:*556	
Wet covers	*5:*3	
Wet desulfurization processes	*1:*849	
Wetland ecosystems	*3:*57	*3:*59
productivity of	*3:*490	
Wetland hydrology	*4:*577	
Wetland management, politics of	*4:*691	
Wetlands	*3:*489	
See also Coastal wetlands; Constructed wetlands		
benefits of	*3:*364	
characteristics of	*4:*695	
classification of	*3:*496	
constructed	*5:*5	
EPA protection of	*3:*494	
federal agencies involved in protecting	*4:*692t	
forested	*3:*172	
functions of	*3:*490	
Great Lakes region	*3:*177	
groundwater and	*5:*605	
N and C transformations in	*5:*329	
organic acidity in	*3:*4	
overview of	*3:*496	
passive treatment of	*2:*423	
physicochemical properties of	*3:*367	
protecting and restoring	*3:*494	
remote sensing of	*4:*325	
role in watershed hydrology	*3:*172	
sanitary impact of	*1:*404	
support of flora and fauna	*3:*492	
terms related to	*3:*59	

Index terms **Links**

Wetlands *(Continued)*
  tidal *3:*72
  values attributed to *3:*492
  vertical flow *5:*6
  water recycling for *2:*612
Wetlands loss *3:*73 *3:*173
  precautionary principle regarding *2:*600
Wetlands policy
  conservation- and restoration-related *4:*693
  legislation and litigation related to *4:*691
  in the United States *4:*690
Wetland treatment *5:*328
Wet periods, role in flooding *3:*144
Wet signal delay (ZWP) *4:*245
Wet sluicing *1:*851
Wet soils *5:*461
Wetted steel tape method *5:*101
Wet thermometer *4:*273
Whole cell fluorescent *in situ* hybridization, in community analysis *1:*645
Whole effluent toxicity controls *2:*382
Whole effluent toxicity testing protocols (EPA) *2:*382
Whole pipe repair and renewal *1:*887
Whole water exchange rates *4:*540
WHO water reuse guidelines *3:*670
WHP zones (WHPZs) *1:*524 *1:*526
Wick lysimeter *2:*340
Width function *3:*94
Wiener-Khintchine theorem *4:*339
Wildlife, vulnerability to pollutant discharge *3:*284
Wilkes*, 4:*John *4:*794
Willis sand zone *3:*67
Wind
  creation of *4:*293
  as a forcing factor *3:*322
Windage *1:*559
Windermere humic aqueous model (WHAM) *2:*206
  version 6 program *5:*141
Wind erosion *3:*566 *5:*125
  control of *3:*551 *3:*568

Index terms **Links**

Windmills *1:*73
  history of *1:*74
Wind pumps *1:*74
  advantages and disadvantages of *1:*75
Windrow piles *1:*648 *1:*649
Wind speed, estimating *4:*2
Wind turbines *1:*75
Wind wave forecasting *4:*139
Wire line coring technique *5:*105
Wire wrap screen *5:*573
Wisconsin Department of Natural Resources wastewater model *1:*731
Wisdom, free flowing water as a source of *4:*793
''Wise use'' policy, for wetlands *3:*492
Witching *5:*92
Withdrawal. *See also* Water
withdrawal(s)
  defined *2:*548
  uses for *2:*661
Women's water rights *4:*771
Woods Hole
  as base of the fisheries service *4:*143
  financial disaster in *4:*141
  fishing activity at *4:*142
  history of *4:*139
  as a scientific center *4:*141
  zoological work at *4:*144
Worker safety, pretreatment programs and *1:*800
Workplace, water conservation in *2:*664
Works Progress Administration (WPA) *2:*511 *2:*513
World Bank Water Research Team *1:*14
World Health Organization (WHO) *1:*20 *1:*23 *1:*31
  See also WHO water reuse guidelines
  arsenic-related guidelines *2:*8
  basic water requirement of *1:*497
  radioactive contamination
  guidelines of *1:*803
  water quality standards *2:*267
World Meteoric Water Line (WMWL) *4:*439
World Ocean Circulation Experiment (WOCE) *4:*22
World public water supply *1:*500

Index terms	**Links**	
World Water Council	*1:*387	
World water industry, key trends in	*1:*508	
Wortley, Stuart	*4:*766	
WWTP monitoring	*2:*48	

X

XANES spectroscopy method	*2:*207	
Xanthates	*1:*686	
Xenobiotic industrial compounds	*5:*43	
Xenobiotics	*2:*42	
Xeriscape design	*3:*750	
benefits of turfgrass in	*3:*750	
challenge of	*3:*752	
practical turf areas in	*3:*750	
principles of	*3:*751	
Xeriscape landscaping, principles of	*3:*555	
Xeriscaping	*2:*548	
Xiaolangdi Reservoir	*3:*48	*3:*49
X-ray photoelectron spectroscopy (XPS)	*4:*491	

Y

Yamuna River	*3:*234	
See also River Yamuna, India		
Yangtze River, flooding of	*3:*143	
Yarra River, Australia, water quality software for	*2:*328	
Yarra River Water Quality Model (YRWQM)	*2:*334	*2:*335
linking with genetic algorithm software	*2:*334	
Yellow River		
current flood control in	*3:*48	
flood analysis in the lower reach of	*3:*49	
historical flood management in	*3:*47	
reducing flood risk for	*3:*49	
Yellow River Basin		
flood control in	*3:*45	
sediment in	*3:*46	
Yersinia enterocolitica	*1:*278	
Yield	*2:*548	
water productivity	*3:*560	
Young groundwater, dating using tritium/helium ratios	*5:*71	

Index terms **Links**

Z

Zebra mussel control, chemical-free *1:*510
Zenith-scaled hydrostatic delay (ZHD) *4:*245
Zenith tropospheric delay (ZTD) *4:*245
Zeolites, sorption by *1:*874
Zero matric potential *5:*533
Zero-order degradation *5:*32
Zero-pressure balloons *4:*165
Zero-tension lysimeters *2:*340 *5:*488 *5:*489
Zero valent metals (ZVMs) *2:*300
Zero Waste America *4:*148
Zinc (Zn)
  bioavailability in soils *5:*370
  in groundwater *2:*148
  phytoextraction of *5:*369
  uptake of *2:*217
Zinc hyperaccumulation, molecular basis of *5:*371
Zinc ions, sorption of *1:*588f
ZoBell solution *2:*466
"Zone of noncompliance" *3:*81
"Zone of passage" *3:*81
"Zone" settling *3:*404
Zoogloeal growth *1:*730
Zoonoses *1:*178
Zooplanktivorous fish, algal control and *2:*6
Zooplankton *4:*154
  effect of road salt on *2:*323
  trace elements in *3:*456
Zooxanthellae *4:*33
Zwitterionic surfactants *1:*670